Inżynieria Internetu Przyszłości

część 1

Praca pod redakcją
Wojciecha Burakowskiego
i Piotra Krawca

Warszawa 2012

Projekt nr POIG.01.01.02-00-045/09 współfinansowany ze środków Europejskiego Funduszu Rozwoju Regionalnego w ramach Programu Operacyjnego Innowacyjna Gospodarka 2007–2013

Recenzent
Tadeusz Łuba

ISBN 978-83-7814-042-9

Oficyna Wydawnicza PW, ul. Polna 50, 00-644 Warszawa. Wydanie I. Zamówienie nr 397/2012
Druk i oprawa: Drukarnia Oficyny Wydawniczej Politechniki Warszawskiej, tel. 22 234-55-93

Przedmowa

Niniejsza książka zawiera zbiór artykułów opisujących główne osiągnięcia krajowego projektu "Inżynieria Internetu Przyszłości - IIP" po ok. dwóch latach działania, tj. za okres 2010-2012. Aktywność projektu IIP skupia się na czterech głównych tematach.

Pierwszym zagadnieniem jest zaprojektowanie, zaimplementowanie i przetestowanie nowej architektury sieciowej dla Internetu następnej generacji, nazywanego Internetem Przyszłości. Architektura ta zakłada wirtualizację zasobów sieciowych oraz ustanowienie na tej podstawie wielu tzw. Równoległych Internetów. W kolejnych artykułach omawiających to zagadnienie przedstawione są szczegóły architektoniczne, wirtualizacji zasobów w sieciach szkieletowych i dostępowych, implementacji węzłów, zarządzania nową infrastrukturą, elementy zapewnienia bezpieczeństwa w sieci i koncepcje realizacji trzech Równoległych Internetów, tj. Internetu IPv6 QoS będącego siecią opartą na protokole IP ale zapewniającą jakość przekazu pakietów, Internetu realizującego sieć świadomą przekazywanej treści (ang. *CAN - Content Aware Networks*) i Internetu bazującego na emulacji komutacji kanałów (ang. *DSS - Data Switching Streams*).

Następnym zagadnieniem realizowanym w projekcie jest budowa krajowej sieci badawczej, nazywanej PL-LAB, która ma umożliwić przeprowadzanie eksperymentów praktycznych dotyczących proponowanych rozwiązań dla Internetu Przyszłości.

Kolejne artykuły omawiają aplikacje, które są projektowane w projekcie i które będą wykorzystane dla zbadania efektywności działania Systemu IIP. Ogólnie, w projekcie tworzonych jest 16 aplikacji, które dotyczą sieci otoczenia człowieka, ochrony zdrowia, edukacji i dostarczania treści. Pierwszy z artykułów dotyczy krótkiego omówienia wszystkich aplikacji, zaś kolejne artykuły opisują wybrane aplikacje głównie z obszaru ochrony zdrowia.

Ostatnim zagadnieniem rozważanym w projekcie jest stworzenie wsparcie technicznego dla transformacji obecnej sieci Internet z protokołu IPv4 do protokołu IPv6. W szczególności przedstawiono opracowywany przewodnik migracji z IPv4 do IPv6, sposoby testowania usług w sieci IPv6, zapewnienie mobilności w sieci IPv6 oraz środowisko IPTV/IMS w sieci IPv6.

Na zakończenie należy nadmienić, iż projekt jest przewidziany na 4 lata i bierze w nim udział ponad 120 pracowników naukowych i inżynierów z 9 organizacji. W skład konsorcjum wchodzą: Politechnika Warszawska (koordynator), Instytut Łączności - Państwowy Instytut Badawczy, Politechnika Wrocławska, Politechnika Poznańska, Instytut Chemii Bioorganicznej PAN - Poznańskie Centrum Superkomputerowo Sieciowe, Instytut Informatyki Teoretycznej i Stosowanej PAN, Politechnika Śląska, Politechnika Gdańska oraz Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie.

Wojciech Burakowski
Koordynator projektu "Inżynieria Internetu Przyszłości"

Badania, których wyniki są przedstawione w tej książce, są prowadzone w ramach projektu „Inżynieria Internetu Przyszłości" częściowo finansowanego przez Ministerstwo Nauki i Szkolnictwa Wyższego z Europejskiego Funduszu Rozwoju Regionalnego w ramach Programu Operacyjnego Innowacyjna Gospodarka na lata 2007-2013, umowa nr POIG.01.01.01-00-045/09-00.

Spis treści

Sieć badawcza PL-LAB

Aplikacje testowe dla Systemu IIP

Transformacja z IPv4 do IPv6

Architektura Systemu IIP

Wojciech Burakowski, Halina Tarasiuk, Andrzej Bęben

Instytut Telekomunikacji, Politechnika Warszawska

Streszczenie W artykule przedstawiono architekturę dla obecnie implementowanego Systemu IIP. System IIP spełnia wymagania stawiane Internetowi Przyszłości (*Future Internet*) i jest istotnie różny od obecnie działającego Internetu opartego na protokole IP (*Internet Protocol*). W szczególności, w omawianym systemie zakłada się, iż infrastruktura sieciowa jest infrastrukturą wirtualną budowaną na bazie urządzeń umożliwiających wirtualizację. Ponadto, przyjęto iż na tej infrastrukturze są ustanowione trzy tzw. Równoległe Internety (RI) działające w oparciu o węzły i łącza wirtualne i różniące się pomiędzy sobą innymi rozwiązaniami dotyczącymi formatów przesyłanych bloków i rozwiązaniami dotyczącymi sterowania. Równoległymi Internetami są: RI-IPv6 QoS (*Internet Protocol Quality of Service*), RI-CAN (*Content Aware Network*) i RI-DSS (*Data Streams Switching*). Rozwiązania RI-CAN i RI-DSS wpisują się w rozwiązania post-IP.

1 Wprowadzenie

Internet przyszłości jest obecnie tematyką wiodącą, realizowaną w ramach 7. Programu Ramowego Unii Europejskiej. Rokrocznie odbywają się dwa spotkania, dotyczące Internetu przyszłości (*Future Internet Assembly*) zrzeszające wykonawców projektów europejskich i narodowych, reprezentantów gremiów standaryzacyjnych (**ITU-T**, **IETF**, **ETSI**) oraz przemysłu, operatorów sieci i dostawców usług. Jednym z podstawowych zagadnień jest opracowanie architektury dla Internetu przyszłości. Rolę koordynacyjną w pracach nad tym zagadnieniem pełni Grupa **FIArch** (*Future Internet Architecture Group*).

Dlaczego ta tematyka jest obecnie tak ważna? Internet w obecnej wersji odniósł ogromny sukces na świecie. Uważa się, iż jego rozwój jest wręcz wskaźnikiem rozwoju cywilizacji. Oczekuje się, iż w dającej się przewidzieć przyszłości (tj. do ok. roku 2020) liczba nowych urządzeń przyłączonych do Internetu będzie 10 razy większa niż obecnie. Wiąże się to z nowymi oczekiwaniami w wyniku wprowadzenia nowych zastosowań, w takich dziedzinach, jak e-zdrowie, e-edukacja, sieci otoczenia człowieka (domowe, samochodowe) itd. Jednocześnie powszechnie uznano, iż dalszy rozwój Internetu oparty na protokole **IP** będzie stanowił istotną barierę dla wprowadzenia nowych zastosowań. Podsumowanie ograniczeń obecnie działającego Internetu można znaleźć np. w [1]. Główne przyczyny tego stanu rzeczy wynikają z własności sieci **TCP/IP**. Ogólnie ujmując, sieć ta jest zbyt prosta, co powoduje iż dodanie do niej nowych funkcji powoduje powstanie skomplikowanych systemów, trudnych do zaimplementowania i utrzymania

przez operatorów sieci. Przykładem rozwiązań sieciowych, które zostały przetestowane w prototypach, a jeszcze nie zaimplementowane, jest sieć **NGN** (*Next Generation Network*), która wymaga wprowadzenia sygnalizacji i rezerwacji zasobów do realizacji połączeń z tzw. gwarancją jakości przekazu pakietów, co jest z kolei wymagane dla wielu aplikacji.

Jak wspomniano, zagadnienie opracowania nowej architektury dla Internetu przyszłości jest obecnie zagadnieniem podstawowym. Nowa architektura powinna być kamieniem milowym w porównaniu z architekturą TCP/IP, co oznacza iż powinna być otwarta m.in. na wprowadzanie nowych rozwiązań sieciowych, modeli biznesowych i nowych aplikacji.

W artykule przedstawiono architekturę Systemu IIP, który jest obecnie implementowany w ramach projektu POIG: *Inżynieria Internetu Przyszłości* [2], [3]. W architekturze tej zakłada się użycie urządzeń, umożliwiających wirtualizację. Dzięki nim jest możliwe zaimplementowanie wielu jednocześnie działających łączy i węzłów wirtualnych. W wyniku tego na jednej infrastrukturze fizycznej można zbudować wiele równolegle działających internetów. Rozwiązanie przyjęte dla systemu IIP zakłada ustanowienie trzech Równoległych Internetów (**RI**), których działanie jest oparte na węzłach oraz łączach wirtualnych i różniących się pomiędzy sobą innymi rozwiązaniami dotyczącymi formatów przesyłanych bloków i mechanizmami sterowania. Równoległymi Internetami są: **RI-IPv6 QoS** (*Internet Protocol Quality of Service*), **RI-CAN** (*Content Aware Network*) i **RI-DSS** (*Data Streams Switching*).

Podsumowanie dotychczasowej aktywności w poszukiwaniu nowej architektury dla Internetu można znaleźć np. w [4]. Oprócz aktywności w zakresie projektów europejskich, są również prowadzone prace w USA pod nazwą Internet 3.0 oraz w Japonii w ramach projektu Akari. Dyskutowanymi obszarami są aplikacje, modele biznesowe, architektury, protokoły i zarządzanie. Przedstawiona w tym artykule architektura Systemu IIP uwzględnia doświadczenia z następujacych architektur: Internet 3.0: *Generalized Networking Architecture* (**GINA**) [5], Akari: *New Generation Networks* (**NWGN**) [6], 4WARD FP7 project: *Virtual Networks* (**VNET**) [7], i *Future Internet Assembly* (**FIA**): *Management and Service-Aware Networking Architectures for Future Internet* (**MANA**) [8]. Powszechnie uważa się, iż rozwiązanie VNET jest dotychczas najlepszym przykładem prototypu systemu i testowanej architektury. VNET opiera się na nowych zasadach dotyczących modelu biznesowego, jak również jest nowym podejściem do zarządzania systemem. Wprowadzono nowych „graczy", w tym przypadku dostawcy fizycznej infrastruktury (*Physical Infrastructure Provider*), dostawcy sieci wirtualnych (*Virtual Network Provider*), operatora sieci wirtualnych (*Virtual Network Operator*) i dostawcy usług (*Service Provider*). Za pewna wadę tego podejścia uważa się, iż zostało ono przetestowane w środowisku sieci IPv4 z wykorzystaniem usługi *best effort*. Z drugiej strony architektury GINA I NWGN są obecnie w fazie prototypowania. Z kolei grupa MANA zdefiniowała do chwili obecnej jedynie ogólne ramy architektury, ale z jasno określonymi problemami oczekującymi na rozwiązanie. Podsumowując, prace nad zdefiniowaniem nowej architektury dla Internetu Przyszłości trwają. Autorzy artykułu

wierzą, że przedstawiona architektura Systemu IIP wspiera aktywności związane z projektowaniem Internetu Przyszłości.

W następnym rozdziale przedstawiono ogólną charakterystykę systemu IIP. Kolejno opisano architekturę systemu, poszczególne Równoległe Internety oraz podano informacje o podejściu do implementacji systemu. Na koniec przedstawiono wnioski.

2 System IIP

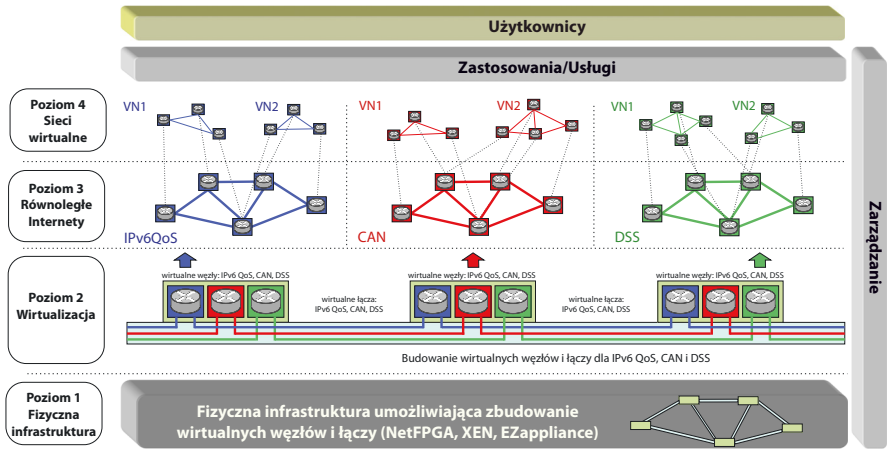
System IIP istotnie zwiększa możliwości infrastruktury sieciowej dla zapewnienia bardziej efektywnego przekazu danych w porównaniu w tym co oferuje obecny Internet działający w oparciu o stos protokołów TCP/IP. Nowe możliwości systemu to: (1) ustanowienie wielu jednocześnie działających RI, (2) poszczególne RI różnią się realizacją płaszczyzny danych i sterowania. W szczególności, każdy z RI może być specjalizowany w obsłudze pewnej grupy aplikacji wymagającej np. szczególnych gwarancji QoS lub bezpieczeństwa przekazu danych.

W Systemie IIP budujemy RI przy wykorzystaniu urządzeń umożliwiających wirtualizację. Zatem poszczególne RI będą działały w oparciu o węzły i łącza wirtualne, co jest nowym podejściem. Nowym rozwiązaniem jest zwłaszcza tworzenie węzłów wirtualnych, co można wykorzystać w przyszłości np. do dynamicznego planowania sieci, gdyż węzły wirtualne można zaimplementować w krótkim czasie i zmieniać w ten sposób topologie sieci. Dla zbudowania łączy i węzłów wirtualnych wykorzystuje się takie narzędzia jak NetFPGA, EZappliance i XEN [9].

Na bazie infrastruktury wirtualnej ustanawia się trzy Równoległe Internety, czyli IPv6 QoS [10], CAN [11] i DSS [12]. Techniki te różnią się pomiędzy sobą zarówno formatami przekazywanych bloków **PDU** (*Protocol Data Unit*) jak i zasadami sterowania ruchem w sieci. Należy przypomnieć, iż w Systemie IIP ww. Internety działają w oparciu o wirtualne węzły i łącza. Zatem podstawowym wymaganiem dla urządzeń umożliwiających wirtualizację jest zapewnienie izolacji pomiędzy tymi węzłami i łączami. Izolacja w tym przypadku oznacza, że obsługa ruchu w jednym z RI nie powoduje pogorszenia jakości obsługi ruchu w innych RI. Dodatkowo, dla odpowiedniego wymiarowania poszczególnych RI, konieczne jest odpowiednie przydzielenie im zasobów infrastruktury fizycznej. Należy zwrócić uwagę, iż przy podziale łącza transmisyjnego musimy uwzględnić fakt, iż w ramach każdego RI przesyłamy PDU o różnych długościach.

3 Architektura Systemu IIP

W architekturze Systemu IIP zakłada 6 poziomów, tak jak to przedstawia rysunek 1. Niższe cztery poziomy, od 1 do 4, dotyczą infrastruktury telekomunikacyjnej i obejmują obszary od zasobów fizycznych po sieci wirtualne. Wyższe poziomy są związane z aplikacjami (poziom 5) i użytkownikami (poziom 6). Również częścią systemu jest system zarządzania, który dotyczy każdego ww. poziomu i komunikacji pomiędzy poziomami.



Rysunek 1. Architektura Systemu IIP.

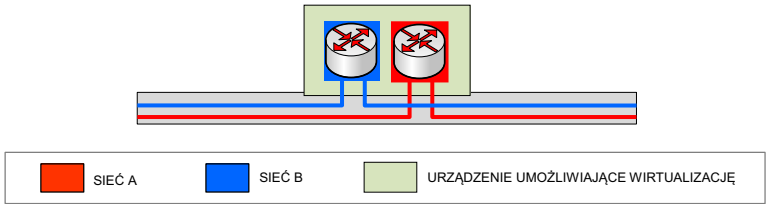
Poniżej pokrótce scharakteryzowano każdy z tych poziomów.

Poziom 1. Infrastruktura fizyczna: Infrastruktura fizyczna Systemu IIP zawiera węzły umożliwiające wirtualizację oraz łącza transmisyjne. Topologia tej sieci jest stała i może stanowić jedną lub wiele domen. Sieć ta ma oddzielną adresację i zarządzanie. Ponadto, sieć zawiera część dostępową (przewodową lub bezprzewodową) oraz szkieletową. Przy budowie prototypu Systemu IIP zakładamy, iż przekaz danych pomiędzy urządzeniami umożliwiającymi wirtualizację odbywa się za pomocą sieci Ethernet.

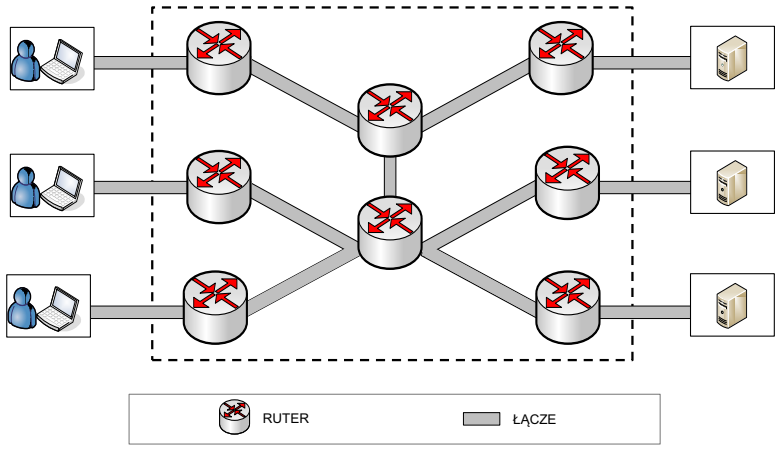
Poziom 2. Wirtualizacja: poziom ten odnosi się do urządzeń umożliwiających wirtualizację, które powinny pozwolić na stworzenie wirtualnych łączy i węzłów należących do poszczególnych RI. Również na tym poziomie przydzielane są fizyczne zasoby dla poszczególnych RI. W celu dedykowania określonych szybkości bitowych dla poszczególnych RI korzystamy ze specjalnie zaprojektowanego mechanizmu szeregowania pakietów opartego na cyklu [13]. W przypadku przydzielenia zasobów dla realizacji węzłów wirtualnych, takich jak CPU i pamięci, zagadnienie nie zostało jeszcze rozwiązane.

Rysunek 2 pokazuje przykład z dwoma wirtualnymi węzłami (powiedzmy A i B) i łączami usytuowanymi na pojedynczym węźle umożliwiającym wirtualizację. Każdy taki węzeł odbiera/wysyła jednostki danych związanych z wirtualnymi węzłami ustanowionymi na łączu fizycznym.

Poziom 3. Równoległe Internety: termin *Równoległy Internet* określa sieć działającą na wirtualnych zasobach, w której wykorzystuje się specyficzne rozwiązania dla realizacji płaszczyzny danych i sterowania. Uważamy, że dzięki utrzymaniu pewnej liczby RI w sieci będzie możliwy bardziej efektywny przekaz danych w porównaniu z obecnie używanym stosem protokołów TCP/IP. Jak



Rysunek 2. Przykład z dwoma węzłami i łączami wirtualnymi ustanowionymi w jednym urządzeniu umożliwiającym wirtualizację.

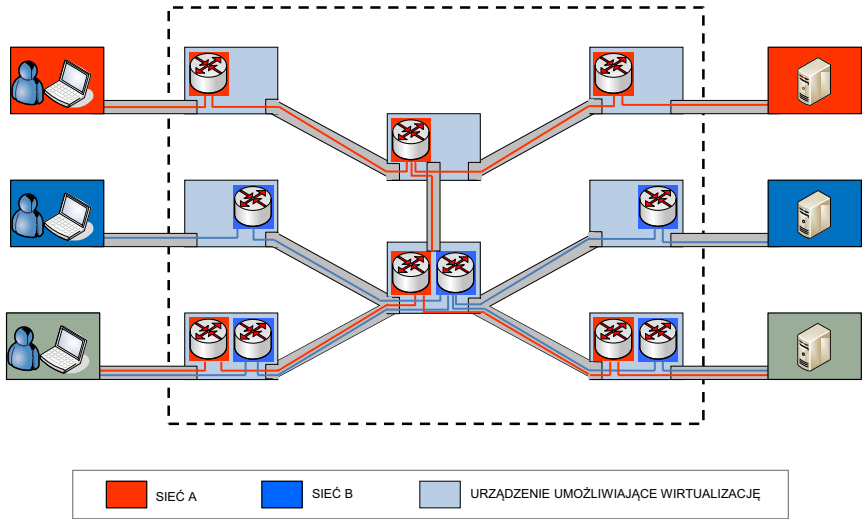


Rysunek 3. Obecny Internet z ruterami.

to stwierdzono powyżej, w Systemie IIP rozważamy trzy RI, a więc IPv6 QoS, CAN i DSS. Sieci te różnią się formatami przekazywanych danych, adresacją, rutingiem, sterowaniem ruchem i zarządzaniem.

Na rysunkach 3 i 4 przedstawiono zasadnicze różnice pomiędzy obecnym Internetem działającym w oparciu o TCP/IP a proponowanym rozwiązaniem przykładowo z dwoma RI, powiedzmy A i B.

Należy podkreślić, iż raczej na jednej infrastrukturze fizycznej/wirtualnej będzie dostępna ograniczona liczba RI pracujących równolegle. To będzie wynikało z trudności przy definiowaniu zbyt wielu istotnie różniących się stosów protokołów. Na podstawie tego, co proponuje się w literaturze, można oczekiwać (przynajmniej w dającej się przewidzieć przyszłości) co najwyżej 3-5 ostatecznych propozycji. Z drugiej strony, można również oczekiwać ograniczeń odnośnie sprawności działania urządzeń umożliwiających wirtualizację, jeżeli będzie



Rysunek 4. Przykład dwóch Równoległych Internetów (A i B) z urządzeniami umożliwiającymi wirtualizację.

na nich ustanowionych zbyt dużo węzłów wirtualnych. Przynajmniej na dzień dzisiejszy, jest raczej niemożliwym do przyjęcia, że liczba wirtualnych węzłów będzie liczona w dziesiątkach i więcej. Biorąc powyższe pod uwagę, RI pozwolą nam na agregacje ruchu obsługiwanego przez dany stos protokołów. Takie rozwiązanie oczywiście wydaje się rozwiązaniem skalowalnym. Jeśli rozróżniamy jedynie trzy RI, zatem na jednym urządzeniu umożliwiającym wirtualizację będą co najwyżej trzy węzły wirtualne.

Poziom 4. Sieci Wirtualne: poziom ten odpowiada za utworzenie pewnej liczby dedykowanych sieci wirtualnych działających równolegle na danym RI. Należy zauważyć, iż w porównaniu do obecnego Internetu jedyna różnica w Systemie IIP jest taka, iż sieci te korzystają nie z zasobów fizycznych (ruterów, łączy) ale z zasobów wirtualnych poziomu 3. Mimo wszystko wydaje się, że do zaprojektowania tego poziomu możemy posłużyć się doświadczeniami z obecnego Internetu. Ten poziom ma styk z aplikacjami/użytkownikami.

4 Typy Równoległych Internetów

Poniżej przedstawimy główne założenia dotyczące rozważanych RI w Systemie IIP.

RI-IPv6 QoS: Rozwiązanie sieci IP QoS bazuje na architekturach DiffServ i NGN. Jest to sieć IP, w której zakłada się utrzymanie pewnej liczby tzw. klas usług. Dana klasa usług jest przewidziana dla obsługi wybranych aplikacji

i zapewnia ścisłe gwarancje przekazu pakietów. Zakłada się, iż w porównaniu z siecią *best effort* w tym przypadku należy zaimplementować mechanizmy QoS działające w różnych skalach czasowych. Dla każdej klasy usługi dedykuje się odpowiednie zasoby sieci (przepustowości łączy, wielkości buforów w ruterach). Istotnym elementem jest zapewnienie izolacji pomiędzy klasami, co oznacza, iż ruch obsługiwany w jednej klasie nie może zakłócić jakości obsługi ruchu obsługiwanego w innej klasie. Za zapewnienie jakości przekazu pakietów w ruterach odpowiadają mechanizmy **PHB** (*Per Hop Behaviour*), które realizują funkcje klasyfikatora, monitorowania, mutipleksacji, szeregowania pakietów. Następne mechanizmy działają na poziomie wywołań, gdzie realizowana jest funkcja przyjmowania nowych wywołań. Nowe wywołanie może być przyjęte jedynie w przypadku, kiedy w sieci jest dostateczna wielkość zasobów. Następnie wymagane jest odpowiednie wymiarowanie zasobów i dobór rutingu. Ogólnie dąży się, aby przyjmowanie nowego wywołania odbywało się jedynie na brzegu sieci, co powoduje konieczność odpowiedniego wymiarowania szkieletu sieci. Sieć IPv6 QoS w przypadku Systemu IIP spełnia rolę szczególną. Nie jest to oczywiście rozwiązanie *post-IP* ale pokazuje, iż przejście z obecnego Internetu do Internetu Przyszłości może odbywać się w sposób ewolucyjny.

CAN: Sieć typu CAN jest obecnie szeroko dyskutowana w literaturze światowej. Ogólnie ujmując, sieć CAN ma za zadanie polepszenie działania sieci Internet w sytuacjach, gdy użytkownicy chcą przesłać różne zbiory, które są umieszczone na różnych serwerach. Chodzi o to, aby wyszukiwanie tych zbiorów odbywało się w krótkim czasie jak również, żeby przesyłać zbiory z tych serwerów, które są umiejscowione jak najbliżej użytkownika. Przyjęte rozwiązanie dla Systemu IIP zakłada zastosowanie nowego rutingu w sieci. Przewiduje się również zaimplementowanie wielu klas usług (różnych niż w IPv6 QoS) w zależności od przesyłanych zbiorów. Ponadto, w sieci CAN możemy stosować zupełnie inny format przekazywanych bloków niż pakiety IP. Wynika to z faktu, iż do każdego przesyłanego bloku dodaje się specjalny nagłówek CAN, który to nagłówek jest analizowany przy przekazie bloku przez węzły sieci.

DSS: Motywacją dla wprowadzenia sieci DSS jest zapewnienie gwarancji strumieniom ruchu generowanym przez źródła o stałej szybkości nadawania, raczej o dużych szybkościach. Przykładem takich aplikacji jest przekaz 2D lub 3D. Taką klasę usług mamy w sieci IPv6 QoS ale ze względu na efekty multipleksacji, jakość jaką możemy oferować jest niezadowalająca dla wspomnianych aplikacji. Dlatego też jest konieczne utworzenie specjalizowanej sieci obsługującej jedynie jeden rodzaj ruchu. Najlepszym rozwiązanie jest użycie klasycznej komutacji kanałów, ale z kolei ta technika wymaga ścisłego synchronizmu czasowego pomiędzy węzłami sieci. Dlatego też zdecydowano się na realizację tej sieci w oparciu o wiedzę z obsługi strumieni ruchu w sieciach pakietowych. DSS opiera się na multipleksacji strumieni, w których stosujemy bardzo krótkie bufory, a poziom strat pakietów regulujemy poprzez regulowanie dopuszczalnego obciążenia łączy.

5 Implementacja

System IIP jest implementowany z użyciem trzech platform umożliwiających wirtualizację, czyli NetFPGA, EZappliance i XEN. Na każdej platformie implementuje się węzły IPv6 QoS, CAN i DSS. Urządzenia te są połączone łączami transmisyjnymi Ethernet. W ramach łącza transmisyjnego wydziela się szczeliny czasowe dla przekazu danych (pole danych ramek Ethernet) należących do danego RI.

Nawiązując do architektury Systemu IIP, należy zwrócić uwagę na istotnie różną funkcjonalność realizowaną na poziomie 1/2 oraz na poziomie 3 i 4. Poziom 1/2 jest wspólny dla wszystkich RI i jego zadaniem jest przygotowanie infrastruktury wirtualnej dla RI, czyli łączy i węzłów wirtualnych wraz z przydzielonymi zasobami. Warunkiem koniecznym jest takie przydzielenie zasobów, aby osiągnąć izolację pomiędzy działaniami poszczególnych RI. Istotnym elementem dla Systemu IIP poziomu 1/2 jest zarządzanie, które odpowiada m.in. za przydzielenie i wymiarowanie zasobów z podziałem na RI. Ostatecznie, implementacja poziomu 3 i 4 jest w zasadzie niezależna od implementacji poziomu 1/2.

6 Podsumowanie

W artykule przedstawiono architekturę Systemu IIP, która opiera się na wirtualizacji zasobów i ustanowieniu na jednej infrastrukturze fizycznej wielu równolegle działających Internetów, każdy możliwie działający o inny stos protokołów. W szczególności, przedstawiono koncepcję utworzenia trzech RI, tj. IPv6 QoS, CAN i DSS. Sieć IPv6 QoS bazuje na protokole IP, podczas gdy rozwiązania CAN i DSS są rozwiązaniami *post-IP*, gdyż dopuszczają inne formaty przekazywanych bloków i inne mechanizmy sterowania.

Proponowane rozwiązanie zapewnia również nowe możliwości tworzenia modeli biznesowych z racji stworzenia nowych poziomów sieci. Przykładowo, inny może być właściciel infrastruktury fizycznej a inny infrastruktury wirtualnej.

Pierwsze prototypy implementacji Systemu IIP należy oczekiwać pod koniec 2011.

Literatura

1. Zahariadis, T. i inni: Towards a Future Internet Architecture. The Future Internet (ed. John Domingue et al.) LNCS 6656, 2011
2. Burakowski, W. i inni: The IIP System: Architecture, Parallel Internets, Virtualization and Applications. Future Network and Mobile Summit Conference, Warszawa, 15–17 June 2011
3. Burakowski, W., Tarasiuk, H., Bęben, A., Zwierko, P.: Future Internet architecture based on virtualization and co-existence of different data and control planes. 5^{th} Workshop on Future Internet Cluster, Warszawa, 15 June 2011
4. Paul, S., Pan, J., Jain, R.: Architectures for future networks and the next generation Internet: A survey. Computer Communications 34(2011), pp. 2–42

5. Jain, R.: Internet 3.0: Ten Problems with Current Internet Architecture and Solutions for the Next Generation. Military Communications Conference, Washington, DC, October 23–25 2006
6. Aoyama, T.: A New Generation Network: Beyond the Internet and NGN. ITU-T Kaleidoscope, IEEE Communications Magazine, May 2009
7. Schaffrath, G. i inni: Network Virtualization Architecture: Proposal and Initial Prototype. VISA'09, August 2009, Barcelona, Spain
8. Galis, A. i inni: Management and Service-aware Networking Architectures (MANA) for Future Internet. System Functions, Capabilities and Requirements, Position Paper, Version V6.0, 3^{rd} May 2009
9. Zwierko, P. i inni: Platformy wirtualizacji dla Systemu IIP. Krajowe Sympozjum Telekomunikacji i Teleinformatyki, Łódź, 14–16 wrzesień 2011
10. Tarasiuk, H. i inni: Architektura i mechanizmy Równoległego Internetu IPv6 QoS. Krajowe Sympozjum Telekomunikacji i Teleinformatyki, Łódź, 14–16 wrzesień 2011
11. Bęben, A. i inni: Architektura sieci wiadomych treści w systemie IIP. Krajowe Sympozjum Telekomunikacji i Teleinformatyki, Łódź, 14–16 wrzesień 2011
12. Danilewicz, G. i inni: Koncepcja Internetu opartego na komutacji strumieni danych dla Systemu IIP. Krajowe Sympozjum Telekomunikacji i Teleinformatyki, Łódź, 14–16 wrzesień 2011
13. Burakowski, W. i inni: Idealne urządzenie umożliwiające wirtualizację infrastruktury sieciowej w Systemie IIP. Krajowe Sympozjum Telekomunikacji i Teleinformatyki, Łódź, 14–16 wrzesień 2011

Idealne urządzenie umożliwiające wirtualizację infrastruktury sieciowej w Systemie IIP

Wojciech Burakowski, Halina Tarasiuk, Andrzej Bęben, Witold Góralski, Piotr Wiśniewski

Instytut Telekomunikacji, Politechnika Warszawska

Streszczenie Wprowadzenie wirtualnej infrastruktury sieciowej na niższych poziomach sieci niż to obecnie ma miejsce, czyli na poziomie sieci wirtualnych, stanowi istotny element rozszerzający możliwości tworzenia nowych architektur przewidzianych do zastosowania w Internecie Przyszłości. Artykuł omawia wymaganą funkcjonalność oraz algorytmy i mechanizmy konieczne dla wykorzystania urządzeń umożliwiających wirtualizację infrastruktury sieciowej na poziomie L1/L2 zgodnie z przyjętą architekturą Systemu IIP [1].

1 Wprowadzenie

Wirtualizacja infrastruktury sieciowej polega na utworzeniu wielu wirtualnych łączy i węzłów na bazie wspólnych fizycznych łączy i węzłów, przy czym każde łącze i węzeł wirtualny powinno być widzialne jako niezależnie działający obiekt. Pojęcie wirtualizacji pojawiło się po raz pierwszy w telekomunikacji w sieciach pakietowych, w których to sieciach jednym z podstawowych pojęć jest "połączenie wirtualne". Poprzez połączenie wirtualne rozumiemy, w odróżnienia od tzw. połączeń fizycznych realizowanych w sieciach z komutacją kanałów, takie połączenie, dla którego nie koniecznie przydziela się zasoby na łączach fizycznych. Dane o połączeniu wirtualnym są zapisane w węzłach sieci wchodzące w skład ścieżki ustalonej przez ruting. Ścieżki takie wyznacza się w fazie zestawiania połączenia. Techniką wykorzystującą połączenia wirtualne na poziomie sieci są przykładowo techniki X.25 i ATM (ang. *Asynchronous Transfer Mode*). w sieciach opartych na protokole IP (ang. *Internet Protocol*), połączenia wirtualne są realizowane na poziomie połączeń TCP (ang. *Transmission Control Protocol*) i zatem nie są zapisane w węzłach sieci. Innym obszarem, gdzie wykorzystujemy pojęcie wirtualizacji, jest tworzenie pewnej liczby sieci wirtualnych ustanawianych na danej infrastrukturze fizycznej sieci, czyli węzłach i łączach. Przykładowo, w sieci IP definiujemy sieci wirtualne, w których dla przekazu pakietów korzysta się ze wspólnych zasobów infrastruktury sieciowej. Zazwyczaj w takich sieciach stosuje się inne pule adresowe i również jest możliwym dedykowanie tym sieciom wydzielonych (statystycznie, dzięki zastosowaniu odpowiednich mechanizmów multipleksacji) zasobów łączy. W artykule tym opiszemy wymaganą funkcjonalność tzw. idealnego urządzenia umożliwiającego wirtualizację zarówno węzłów jak i łączy. Funkcjonalność taką wymagamy w Systemie

IIP, w szczególności na poziomie warstwy L2 zgodnie z przyjętą architekturą [1], [2], [3]. System IIP jest propozycją dla Internetu Przyszłości i zakłada ustanowienie trzech Równoległych Internetów (RI), tj. IPv6 QoS (ang. *Quality of Service*), CAN (ang. *Content Aware Network*) i DSS (ang. *Data Streams Switching*) działających w oparciu o łącza i węzły wirtualne. Organizacja artykułu jest następująca. Rozdział 2 omawia pokrótce wymagania dotyczą wirtualnych łączy i węzłów dla ww. Równoległych Internetów. w rozdziale 3. przedstawimy koncepcję idealnego urządzenia umożliwiającego wirtualizację. Mechanizmy i algorytmy wymagane dla tego urządzenia opisujemy w rozdziale 4. Ostatecznie, rozdział 5 podsumowuje artykuł.

2 Wymagania na wirtualizację urządzeń sieciowych

W przypadku Systemu IIP, infrastrukturę fizyczną sieci stanowią połączone łączami transmisyjnymi urządzenia umożliwiające wirtualizację, definiując w ten sposób topologię sieci fizycznej. Na bazie tej infrastruktury ustanawiamy trzy RI, czyli IPv6 QoS [4], CAN [5] i DSS [6]. Techniki te różnią się pomiędzy sobą zarówno formatami przekazywanych bloków PDU (ang. *Protocol Data Unit*) jak i zasadami sterowania ruchem w sieci. Należy podkreślić, iż w Systemie IIP ww. Internety działają w oparciu o wirtualne węzły i łącza. Zatem, podstawowym wymaganiem dla urządzeń umożliwiających wirtualizację jest zapewnienie izolacji pomiędzy ww. węzłami i łączami wirtualnymi. Izolacja w tym przypadku oznacza, że obsługa ruchu w jednym z RI nie powoduje pogorszenia jakości obsługi ruchu w innych RI. Dodatkowo, dla odpowiedniego wymiarowania poszczególnych RI, koniecznym jest odpowiednie przydzielenie im zasobów infrastruktury fizycznej. Ponadto, przy podziale łącza transmisyjnego musimy uwzględnić fakt, iż w ramach każdego RI przesyłamy PDU o różnych długościach.

3 Koncepcja idealnego urządzenia umożliwiającego wirtualizację

Przypomnijmy, że węzły wirtualne utworzone na bazie urządzeń umożliwiających wirtualizację powinny zapewnić funkcjonalność i możliwości obsługi ruchu na podobnym poziomie jak węzły fizyczne. Oznacza to, że korzystamy z węzłów i łączy wirtualnych nie powinno wpływać na jakość działania RI. Ponadto, dla każdego węzła i łącza wirtualnego musimy zapewnić jego izolację od innych. Rys. 1 przedstawia schemat blokowy idealnego urządzenia umożliwiającego wirtualizację w przypadku, kiedy mamy dwa porty wejściowe (A i B), dwa porty wyjściowe (C i D), trzy różne węzły wirtualne należące do Internetów IPv6 QoS, CAN i DSS oraz agenta dla zarządzania. Bloki PDU, przykładowo przychodzące z łącza a i przesyłane w ramach RI-IPv6 QoS/RI-CAN/RI-DSS, są przekierowywane do obsługi w "swoim" węźle wirtualnym, tj. w węźle IPv6 QoS/CAN/DSS. w przypadku, kiedy odbierane są wiadomości zarządzające Systemem IIP, są one przesyłane do agenta zarządzania (MGT agent). Należy przypomnieć, iż omawiane wiadomości zarządzające dotyczą wyłącznie poziomu L1/L2, natomiast

nie dotyczą zarządzania w ramach poszczególnych RI. Wiadomości zarządzające danym Internetem są przesyłane w PDU należących do danego Internetu. Poprawne działanie urządzenia umożliwiającego wirtualizację warunkowane jest znalezieniem rozwiązań dotyczących następujących problemów:

- Problem 1: zdefiniowanie nagłówka pozwalającego na rozróżnienie przynależności odbieranych PDU do danego RI lub zarządzania L1/L2. Zakładamy, iż w łączach transmisyjnym przesyłamy ramki Ethernet. Założenie takie zostało podyktowane popularnością Ethernetu i przyjętymi założeniami dla prototypu Systemu IIP.
- Problem 2: zapewnienie izolacji pomiędzy wirtualnymi łączami przy przekazie PDU przez łącza transmisyjne.
- Problem 3: zapewnienie izolacji w obsłudze ruchu pomiędzy wirtualnymi węzłami i agentem zarządzania.

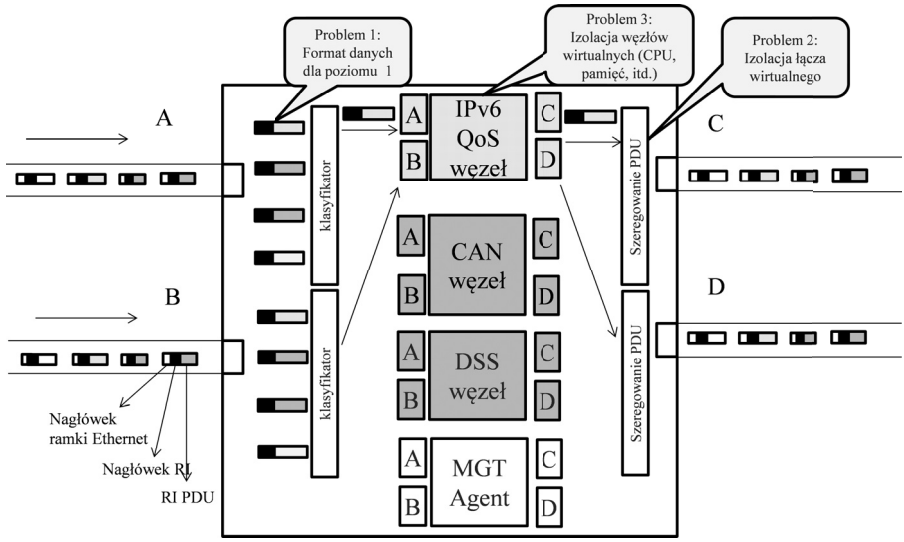
W rozdziale 4 opiszemy proponowane rozwiązania dla problemów 1 i 2, które to problemy są związane z przekazem PDU w ramkach Ethernet. Rozwiązanie problemu 3 to dedykowanie oddzielnych zasobów takich jak CPU, pamięć operacyjna, pamięć masowa dla każdego wirtualnego węzła i agenta zarządzania. w tym przypadku, koniecznym jest również zdefiniowanie modelu zasobów. Niestety, w obecnie dostępnych urządzeniach umożliwiających wirtualizację nie wydaje się, aby to było możliwe. Zatem, pierwsze rozwiązania problemu zapewnienia izolacji będą dopiero po zaimplementowaniu Systemu IIP i jego przetestowaniu. Należy oczekiwać, iż zapewnienie izolacji pomiędzy węzłami wirtualnymi jest ściśle związane z zastosowanym typem urządzenia/narzędzia umożliwiającego wirtualizację. Omówienie poszczególnych typów urządzeń/narzędzi zawarte jest w [7].

4 Mechanizmy i algorytmy dla przekazu PDU przez system transmisyjny

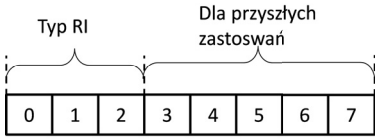
W rozdziale tym omówimy proponowane rozwiązania dla uzyskania izolacji pomiędzy poszczególnymi RI oraz zarządzaniem L1/L2 w procesie transmisji PDU przez łącze transmisyjne. Powyższe uzyskujemy przez zastosowanie odpowiedniego klasyfikatora na łączu wejściowym i odpowiedniej metody szeregowania w dostępie do łącza wyjściowego.

4.1 Klasyfikacja PDU na łączu wejściowym

Łącze transmisyjne jest wspólnie wykorzystywane przez RI-IPv6 QoS, RI-CAN, RI-DSS oraz zarządzanie L1/L2. Dla rozróżnienia przynależności danego przesyłanego PDU do określonego RI lub zarządzania L1/L2 musimy zdefiniować nowy nagłówek PIH (ang. *Parallel Internet Header*), którego długość przyjęto 8 bitów jak to przedstawia Rys. 2. Pierwsze trzy bit stanowią pole nagłówka - "typ RI" (typ Równoległego Internetu) dla rozróżnienia przynależności PDU do RI lub



Rysunek 1. Idealne urządzenie umożliwiające wirtualizację.



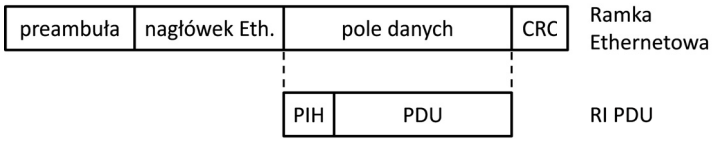
Rysunek 2. 8 bitowy nagłówek PIH: 3 bity dla klasyfikacji PDU do poszczególnych RI lub zarządzania L1/L2, 5 bitów nie wykorzystanych.

zarządzania L1/L2. Pozostałe 5 bitów zostało zarezerwowane dla przyszłych zastosowań. 8 bitowe pole dla nagłówka PIH powinno być wystarczająco duże, gdyż raczej należy spodziewać się jedynie kilku typów RI. Przypomnijmy, że każdy z RI powinien być istotnie różny od pozostałych w rozwiązaniach dotyczących płaszczyzny danych i płaszczyzny sterowania.

Jak uprzednio wspomniano, w projektowanym Systemie IIP zakładamy, że przez systemy transmisyjne przesyłamy ramki Ethernet. Na Rys. 3 przedstawiono format przesyłanych ramek Ethernet obejmujący preambułę i nagłówek ramki Ethernet, RI PDU (zawierające nagłówek i PDU) oraz CRC.

4.2 Mechanizm szeregowania PDU przed wysłaniem do łącza wyjściowego

Mechanizm szeregowania PDU decyduje o kolejności wysyłania ich w zależności od przynależności do RI. Głównym celem algorytmu szeregowania RI-PDU



Rysunek 3. Enkapsulacja PDU RI w ramce Ethernet.

jest podział pojemności łącza wyjściowego C pomiędzy poszczególne RI i zarządzanie L1/L2. w przypadku trzech RI przydzielamy każdemu odpowiednią część pojemności łącza C. Ponadto powinniśmy przydzielić odpowiednią pojemność łącza dla ruchu związanego z zarządzaniem Systemem IIP (ruch typu MGT). W szczególności powinien być spełniony następujący warunek:

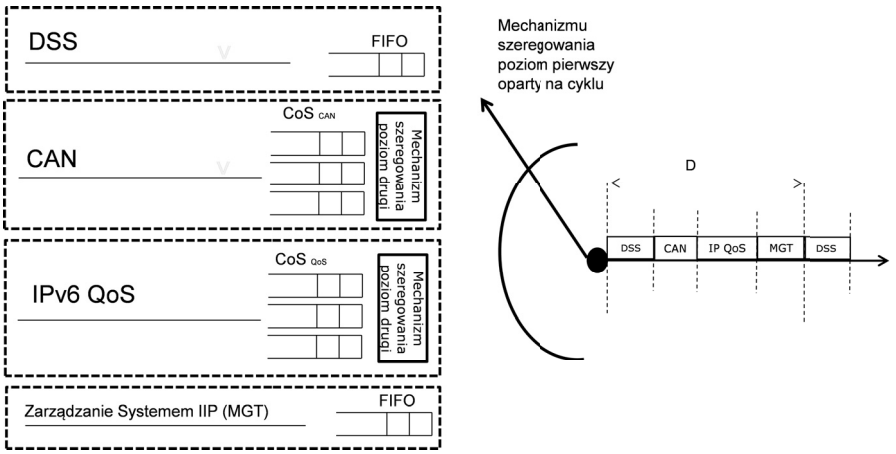
$$C_{IPv6QoS} + C_{CAN} + C_{DSS} + C_{MGT} \leqslant C \tag{1}$$

gdzie C oznacza pojemność łącza wyjściowego zaś $C_{IPv6QoS}$, C_{CAN}, C_{DSS} i C_{MGT} oznaczają odpowiednio przydzielone pojemności dla RI-IPv6 QoS, RI-CAN, RI-DSS i zarządzania.

Zapewnienie izolacji pomiędzy strumieniami danych należącymi do różnych RI i zarządzania L1/L2 jest warunkiem koniecznym zapewnienia izolacji pomiędzy nimi. Na Rys. 4 przedstawiono dwupoziomowy algorytm szeregowania RI-PDU wykorzystywany w Systemie IIP. Pierwszy poziom mechanizmu odpowiada za podział dostępu do łącza pomiędzy RI-IPv6 QoS, RI-CAN, RI-DSS oraz zarządzanie (MGT). Działanie tego mechanizmu szeregowania oparte jest na cyklu. w naszym przypadku cykl składa się ze stałej liczby faz, z których każda jest dedykowana odpowiedniemu RI lub zarządzaniu L1/L2. Czas trwania każdej z faz jest ustalany w fazie wymiarowania Systemu IIP. Należy podkreślić, iż w tym przypadku, niewykorzystana przepływność przez jeden z RI nie może być wykorzystana przez innych. Rozwiązanie to wpisuje się w klasę algorytmów "nie zachowujących ciągłości obsługi" (ang. *non-work conserving*). Mechanizm szeregowania PDU na poziomie drugim jest właściwy dla danego RI. Przykładowo, sieci IPv6 QoS czy też CAN definiują różne klasy usług i zatem rozważają struktury wielokolejkowe.

5 Podsumowanie

W artykule pokrótce przedstawiono niezbędną funkcjonalność idealnego urządzenia umożliwiającego zbudowanie wirtualnych węzłów i łączy. Przedstawiona funkcjonalność jest niezbędna dla zbudowania Systemu IIP. Podstawowym wymaganiem dla takiego urządzenia jest zapewnienie izolacji pomiędzy łączami przy przesyłaniu bloków PDU przez łącza transmisyjne i zapewnienie izolacji pomiędzy wirtualnymi węzłami przy obsłudze ruchu. Dla rozwiązania problemu związanego z przesyłaniem PDU przez łącza transmisyjne, zaproponowano



Rysunek 4. Proponowany dwupoziomowy mechanizm szeregowania.

użycie klasyfikatora na łączu wejściowym działającego w oparciu o nowy nagłówek, zaś dla wydzielenia zasobów na łączu wyjściowym zaproponowano metodę szeregowania opartą na cyklu. Rozwiązania te są obecnie implementowane i testowane. Drugie z wymienionych zagadnień nie ma jeszcze rozwiązania co wynika bezpośrednio z faktu, iż problem wirtualizacji węzłów nie jest jeszcze zbadany. Potrzebujemy w tym przypadku modelu zasobów w urządzeniach umożliwiających wirtualizację. Pierwszych propozycji rozwiązań należy oczekiwać w niedalekiej przyszłości.

Literatura

1. Burakowski W. i inni, Architektura Systemu IIP, Krajowe Sympozjum Telekomunikacji i Teleinformatyki, Łódź, 14-16 września 2011
2. Burakowski W. i inni, The IIP System: Architecture, Parallel Internets, Virtualization and Applications, Future Network and Mobile Summit Conference, Warszawa, 15-17 czerwca 2011
3. Burakowski W., Tarasiuk H., Beben A., Zwierko P., Future Internet architecture based on virtualization and co-existence of different data and control planes, 5th Workshop on Future Internet Cluster, Warszawa, 15 czerwca 2011
4. Tarasiuk H. i inni, Architektura i mechanizmy Równoległego Internetu IPv6 QoS, Krajowe Sympozjum Telekomunikacji i Teleinformatyki, Łódź, 14-16 września 2011
5. Beben A. i inni, Architektura sieci świadomych treści w systemie IIP, Krajowe Sympozjum Telekomunikacji i Teleinformatyki, Łódź, 14-16 września 2011
6. Danilewicz G. i inni, Koncepcja Internetu opartego na komutacji strumieni danych dla Systemu IIP, Krajowe Sympozjum Telekomunikacji i Teleinformatyki, Łódź, 14-16 września 2011
7. Zwierko P. i inni, Platformy wirtualizacji dla Systemu IIP, Krajowe Sympozjum Telekomunikacji i Teleinformatyki, Łódź, 14-16 września 2011

Platformy wirtualizacji: implementacja węzła Systemu IIP

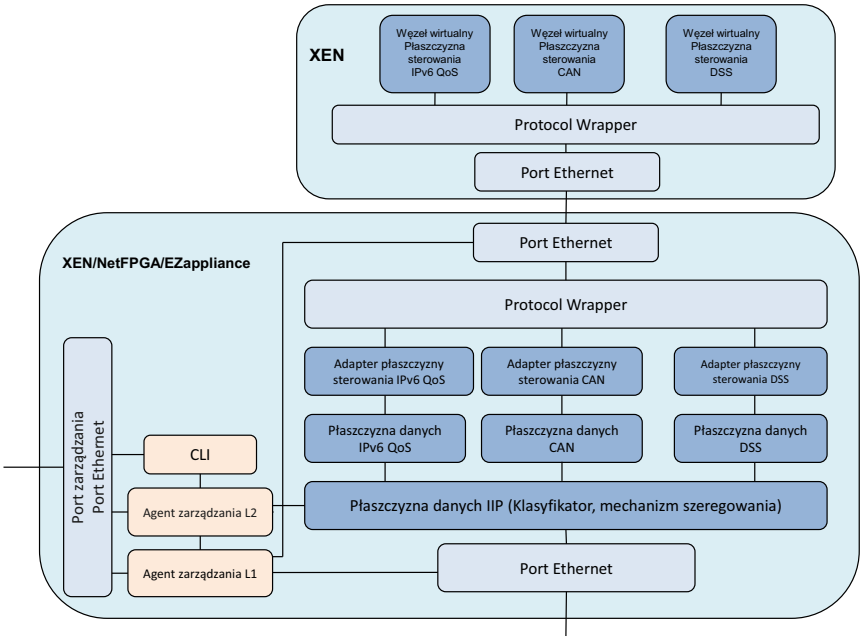
Piotr Zwierko[1], Halina Tarasiuk[1], Mariusz Rawski[1], Piotr Wiśniewski[1], Damian Parniewicz[2], Artur Juszczyk[2], Błażej Adamczyk[3], Adam Kaliszan[4]

[1] Instytut Telekomunikacji, Politechnika Warszawska
[2] Poznańskie Centrum Superkomputerowo-Sieciowe
[3] Politechnika Śląska
[4] Politechnika Poznańska

Streszczenie Artykuł przedstawia realizację węzłów Systemu IIP i koncepcję wirtualizacji węzła z wykorzystaniem trzech różnych platform implementacyjnych: czysto-programowej wykorzystującej oprogramowanie Xen, układu programowalnego NetFPGA oraz urządzenia EZappliace z programowalnym procesorem sieciowym EZchip NP-3.

1 Wprowadzenie

System IIP jest budowany w ramach projektu Inżynieria Internetu Przyszłości (IIP) [1] [2]. System jest tworzony w oparciu architekturę zaproponowaną w [3]. Dwie pierwsze warstwy architektury (fizyczna i wirtualizacji) mają dostarczyć infrastrukturę dla tworzenia sieci w ramach trzech Równoległych Internetów (*PI – Parallel Internet*): IPv6 QoS (*Quality of Service*) – sieć z gwarancją jakości przekazu pakietów, CAN (*Content Aware Network*) – sieć świadoma przekazywanej treści oraz DSS (*Data Stream Switching*) – komutacja strumieni danych. Rozważane Równoległe Internety różnią się zarówno warstwą protokołów dla przekazu danych, jak i dla sterowania. Aby zapewnić możliwość zbudowania węzłów Równoległych Internetów na jednej wspólnej infrastrukturze fizycznej stworzono koncepcję idealnego urządzenia umożliwiającego wirtualizację [4]. Urządzenie to realizuje dwie podstawowe funkcjonalności: przekaz danych oraz sterowanie (ruting). Do implementacji wybrane zostały trzy platformy: czysto-programowa wykorzystująca oprogramowanie Xen, układ programowalny NetFPGA oraz urządzenie EZappliance z programowalnym procesor sieciowy EZchip NP-3. Xen i NetFPGA są również stosowane dla rozwiązań prototypowych w obszarze technik wirtualizacji w projektach międzynarodowych. Celem implementacji jest zbadanie izolacji między wirtualnymi węzłami i łączami na różnych platformach przy realizacji trzech Równoległych Internetów. Główne bloki funkcjonalne płaszczyzny przekazu danych Systemu IIP, które są implementowane w urządzeniach umożliwiających wirtualizację to klasyfikator w porcie wejściowym urządzenia oraz mechanizm szeregowania pakietów w porcie wyjściowym. W celu realizacji płaszczyzny danych i sterowania dla każdego z Równoległych Internetów zaprojektowano odpowiednie bloki funkcjonalne. W artykule przedstawiamy w rozdziale 2 oryginalną propozycję ogólnego modelu implementacyjnego



Rysunek 1. Ogólny model implementacyjny węzła Systemu IIP.

węzła Systemu IIP opartego na koncepcji idealnego urządzenia umożliwiającego wirtualizację. Zaś w kolejnych rozdziałach realizację ogólnego modelu z wykorzystaniem trzech platform: Xen (rozdział 3), NetFPGA (rozdział 4) i EZappliance (rozdział 5). Rozdział 6 stanowi podsumowanie.

2 Ogólny model implementacyjny węzła Systemu IIP

Rysunek 1 przedstawia proponowany model implementacyjny węzła Systemu IIP. W modelu ogólnym przyjęto założenie o podziale węzła Systemu IIP na dwa moduły funkcjonalne: moduł pierwszy zależny od zastosowanej platformy (Xen, NetFPGA lub EZappliance), spełniający głównie funkcje płaszczyzny przekazu danych (DP – Data Plane), oraz moduł drugi, wspólny dla wszystkich rodzajów węzłów Systemu IIP, spełniający rolę płaszczyzny sterowania (CP - Control Plane). Płaszczyzna przekazu danych składa się z części wspólnej dla Systemu IIP (klasyfikator - port wejściowy, mechanizm szeregowania pakietów IIP - port wyjściowy) oraz części właściwych dla poszczególnych Równoległych Internetów. Platforma Xen jest realizacją całkowicie programową, zaś EZappliance i NetFPGA zawierają elementy sprzętowe umożliwiające zwiększenie wydajności funkcji przekazu danych w porównaniu z platformą programową Xen. Implementacja płaszczyzny sterowania na platformach NetFPGA i EZappliance jest dla zasto-

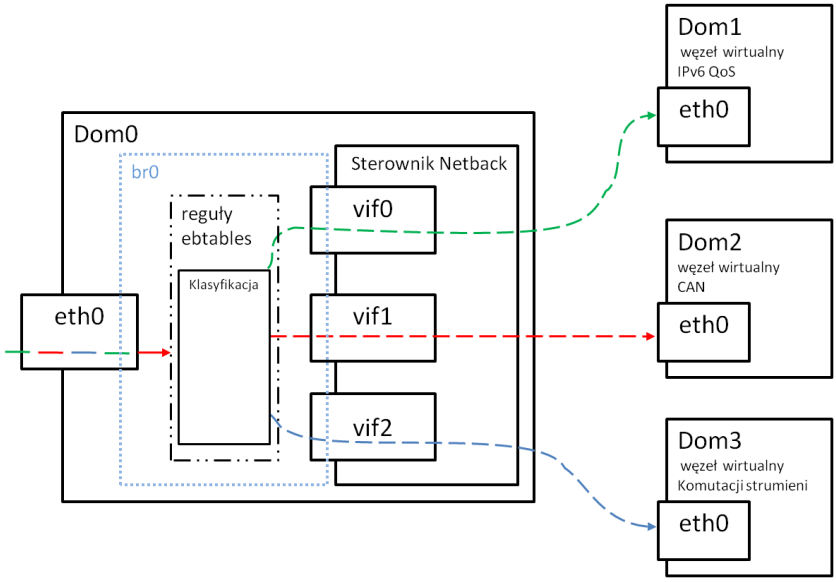
sowanych urządzeń ograniczona ze względu na znaczną złożoność algorytmów sterowania, co pociąga za sobą nieproporcjonalnie duże nakłady na implementację oraz zużycie zasobów urządzeń bez istotnego wzrostu wydajności węzła. Dlatego też w projekcie przyjęliśmy, że funkcje płaszczyzny sterowania zostaną zaimplementowane na oddzielnej platformie Xen, w której dla każdego Równoległego Internetu zostanie przeznaczona oddzielna maszyna wirtualna. Platforma ta będzie połączona za pomocą łącza Ethernet z platformą realizującą funkcje przekazu danych. Ponieważ komunikacja CP z DP w ogólnym przypadku jest zależna od rodzaju platformy użytej dla realizacji płaszczyzny DP, koniecznym okazało się użycie Adapterów Płaszczyzn Sterowania, uniezależniających komunikację CP – DP od rodzaju platformy sprzętowej. Na początkowym etapie implementacji przyjęto również upraszczające założenie o oddzieleniu komunikacji dla potrzeb zarządzania od Równoległych Internetów i przeznaczeniu na ten cel dedykowanej sieci. W pierwszej fazie implementacji, dla uproszczenia realizacji funkcji zarządzania, w każdym węźle Systemu IIP zdefiniowano interfejs (CLI - Command Line Interface) umożliwiający konfigurację podstawowych modułów przekazu danych dla poziomu 1 i 2 Systemu IIP. Konfiguracja ta odbywa się odpowiednio za pomocą agentów zarządzania L1 oraz L2.

3 Realizacja węzła Systemu IIP na platformie XEN

Jedna z trzech platform wirtualizacji węzłów i łączy w Systemie IIP wykorzystuje technologię Xen [5] [6]. Xen jest to opracowany na Uniwersytecie Cambridge monitor maszyn wirtualnych, rozpowszechniany na licencji *open source*. Xen umożliwia w pełni programową wirtualizację zasobów, dlatego uzyskiwana wydajność może być mniejsza niż w przypadku platform sprzętowych. Takie rozwiązanie ma jednak również istotne zalety. Jest tanie, może zostać uruchomione na dowolnym komputerze umożliwiającym wirtualizację i nie wymaga innego sprzętu. W dalszej części przedstawiony zostanie szczegółowy opis możliwej realizacji klasyfikatora, węzła wirtualnego oraz mechanizmu szeregowania w węźle Systemu IIP realizowanym na platformie Xen.

3.1 Klasyfikator

Klasyfikator Systemu IIP zostanie uruchomiony na głównej maszynie wirtualnej systemu Xen o nazwie *Domena 0* (rysunek 2). Ta domena, jako jedyna, posiada bezpośredni dostęp do interfejsu sieciowego *eth0* i jest odpowiedzialna za dzielenie tego interfejsu pomiędzy pozostałe maszyny wirtualne. Podział ten jest realizowany za pośrednictwem sterownika *Netback*, który dla każdej maszyny wirtualnej (o nazwie *Domena U*) tworzy wirtualny interfejs sieciowy (*vif*) oraz mostek (*br0*) łączący wszystkie interfejsy *vif* oraz *eth0*. Każdy interfejs *vif* posiada własny adres MAC co sprawia, że wszystkie domeny U są osiągalne z zewnętrznej sieci. Mostek sieciowy w architekturze Xen jest standardowym mechanizmem *bridging* systemu Linux. Dzięki temu administrator może łatwo zarządzać ruchem za pomocą dobrze znanych narzędzi, takich jak *iptables* czy

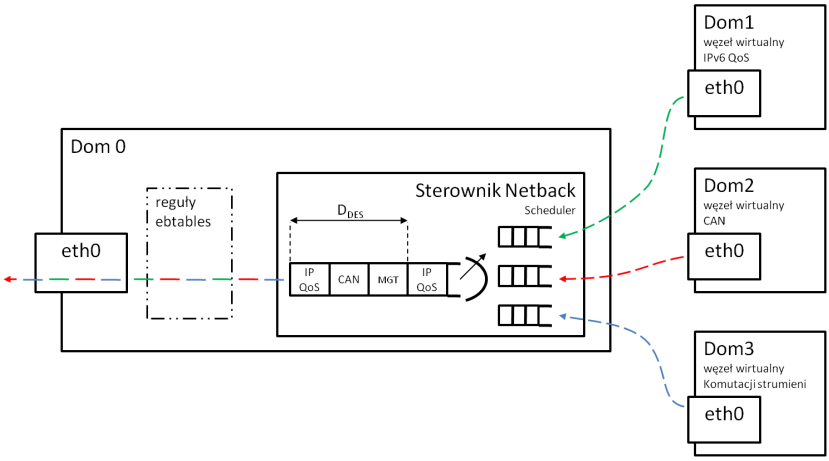


Rysunek 2. Klasyfikator Systemu IIP na platformie Xen.

ebtables. Taka architektura pozwala na implementację klasyfikatora Systemu IIP za pomocą zmodyfikowanych reguł *ebtables*. Aby to osiągnąć, należy stworzyć specjalne rozszerzenie mechanizmu *ebtables* zwane *matching extension*. Rozszerzenie to pozwoli na identyfikację ramek należących do danego Równoległego Internetu przy pomocy nagłówka PIH (Parallel Internet Header). Właściwa klasyfikacja może zostać następnie określona za pomocą reguł wykorzystujących to rozszerzenie, dodanych do łańcucha *BROUTING* tabeli *BROUTE* [7]. Na najniższym poziomie Systemu IIP wszystkie ramki Ethernet są adresowane na adres MAC węzła fizycznego (np. MAC interfejsu *eth0*). Wspomniane reguły muszą więc zmienić adres docelowy na adres MAC węzła wirtualnego, zgodnie z zawartością nagłówka PIH, zanim ramki dotrą do mostka *br0*. Ten z kolei standardowo prześle ramki do odpowiedniego interfejsu wirtualnego *vif*. W tym celu można wykorzystać standardowe rozszerzenie *DNAT* mechanizmu *ebtables*, które pozwala na zmianę adresu docelowego MAC. Takie rozwiązanie ukrywa adresy MAC interfejsów wirtualnych, dzięki czemu w każdym węźle fizycznym będzie można użyć tych samych adresów, co znacznie ułatwi zarządzanie.

3.2 Węzeł wirtualny

Jedną z zalet rozwiązania opartego o wyłącznie programową wirtualizację jest to, że każdy węzeł wirtualny będzie uruchomiony na osobnej maszynie wirtualnej z własnym systemem operacyjnym, zasobami i specyficznym oprogramowaniem.



Rysunek 3. Mechanizm szeregowania Systemu IIP na platformie Xen.

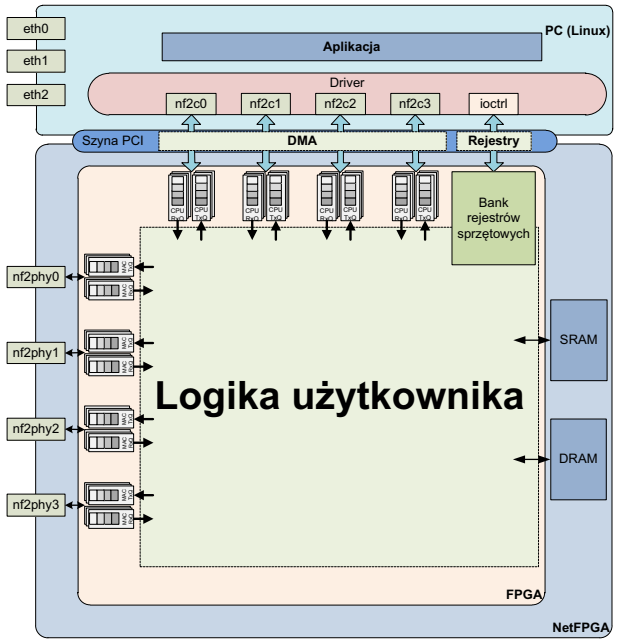
Oznacza to, że wszystkie węzły będą odizolowane i nie będą zależeć od siebie, co znacznie ułatwia implementację oraz zarządzanie.

3.3 Mechanizm szeregowania

W standardowej konfiguracji Xen szereguje ruch wychodzący wewnątrz procesu wspomnianego wcześniej sterownika *Netback*. Algorytm szeregowania wykorzystuje standardową kolejkę FIFO rozszerzoną o tzw. *Credit Scheduler*, który pozwala na określenie maksymalnej przepustowości dla dowolnego interfejsu wirtualnego za pomocą parametru konfiguracyjnego *rate*. Aby zaimplementować algorytm szeregowania dla węzła Systemu IIP, standardowy mechanizm szeregowania Xen musi zostać zmieniony. Następnie, podobnie jak w przypadku klasyfikatora, adres źródłowy ramek wychodzących z węzła musi zostać ustawiony na adres MAC węzła fizycznego (domyślnie będą one mieć adres MAC węzła wirtualnego). Można to osiągnąć stosując kolejną regułę *ebtables*, dodaną do łańcucha *POSTROUTING* tabeli *NAT* używając standardowego rozszerzenia *SNAT*. Schemat blokowy przedstawia rysunek 3.

4 Realizacja węzła Systemu IIP na platformie NetFPGA

Platforma NetFPGA składa się z karty PCI wyposażonej w układ programowalny FPGA Virtex-II Pro 50 firmy Xilinx, cztery porty Ethernet 1 Gbit/s, pamięci SRAM i DRAM (rysunek 4) oraz z oprogramowania umożliwiającego zarządzanie nią [8]. Karta ta jest instalowana w komputerze klasy PC działającym pod kontrolą systemu Linux. Konfiguracja i zarządzanie kartą odbywa się poprzez szynę PCI. W tym celu opracowane zostały moduły jądra systemu Linux (w

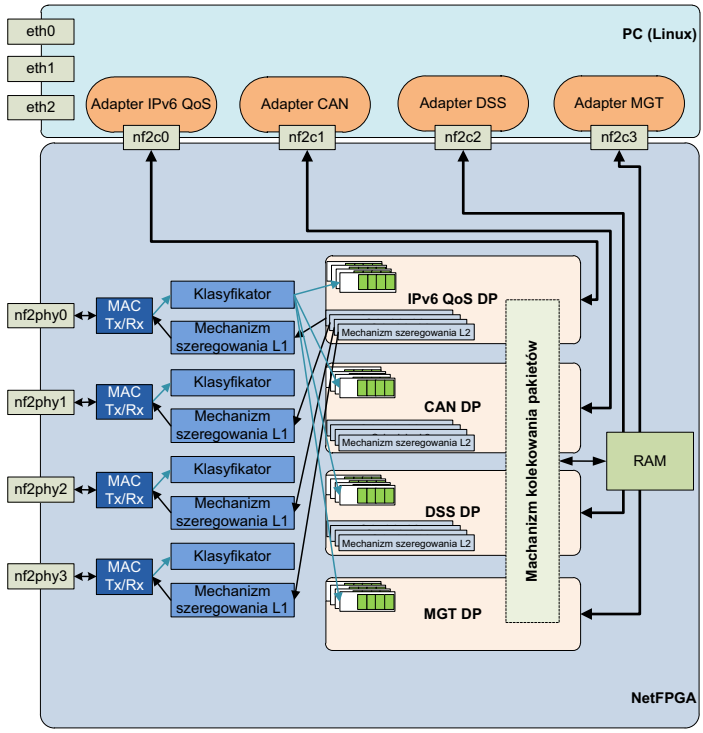


Rysunek 4. Schemat blokowy platformy NetFPGA.

języku C), które zapewniają komunikację między częścią sprzętową realizowaną w układzie FPGA karty NetFPGA a częścią programową realizowaną w systemie Linux. Moduły jądra udostępniają w systemie Linux 4 wirtualne porty sieciowe i zapewniają mechanizmy, które przekazują dane zapisywane/odczytywane z tych portów do/z odpowiednich kolejek FIFO zrealizowanych w układzie FPGA. Dodatkowo pozwalają one dozorować rejestry sprzętowe w pamięci systemu operacyjnego. Dzięki platformie NetFPGA istnieje możliwość realizacji w układzie FPGA elementów składowych węzła IIP dla płaszczyzny przekazu danych natomiast część elementów płaszczyzny sterowania może być zrealizowana w postaci oprogramowania działającego w systemie Linux. Takie rozwiązanie pozwala na uzyskanie bardzo dużej szybkości przetwarzania dzięki sprzętowej akceleracji kluczowych elementów systemu, a jednocześnie nie traci się możliwości dopasowania systemu do własnych potrzeb co gwarantuje programowa realizacja części funkcjonalności systemu i możliwość rekonfiguracji układu programowalnego [9].

4.1 Klasyfikator

Klasyfikator Systemu IIP zostanie zrealizowany jako sprzętowy moduł przetwarzający strumień danych otrzymywanych z danego portu fizycznego. Podstawowym zadaniem tego modułu jest analiza nagłówka odebranej ramki oraz wyod-



Rysunek 5. Schemat blokowy realizacji wirtualnego węzła z wykorzystaniem platformy NetFPGA.

rębnienie informacji, dzięki której ramka będzie przekierowana do odpowiedniego węzła wirtualnego (rysunek 5). Z tego względu, że klasyfikator nie dokonuje żadnych modyfikacji ramki tego typu realizacja będzie bardzo wydajna. Dodatkowo moduły klasyfikatora dla każdego portu będą działały całkowicie niezależnie w sposób w pełni zrównoleglony.

4.2 Węzeł wirtualny

Podstawową zaletą realizacji węzła IIP z wykorzystaniem platformy NetFPGA jest możliwość implementacji elementów składowych węzła jako niezależnych modułów sprzętowych co oznacza całkowitą izolację wszystkich węzłów wirtualnych, a co za tym idzie brak konieczności współdzielenia zasobów sprzętowych układu FPGA (rysunek 5. Dodatkowo, brak zależności węzłów wirtualnych od siebie znacznie ułatwia implementację oraz zarządzanie. Z tego względu, że istnieje konieczność realizacji dla każdego z wirtualnych węzłów szeregu kolejek FIFO przechowujących przetworzone ramki niezbędna będzie realizacja mecha-

nizmu współdzielenia zewnętrznej pamięci RAM oraz algorytmu arbitrażu żądań zapisu/odczytu.

4.3 Mechanizm szeregowania

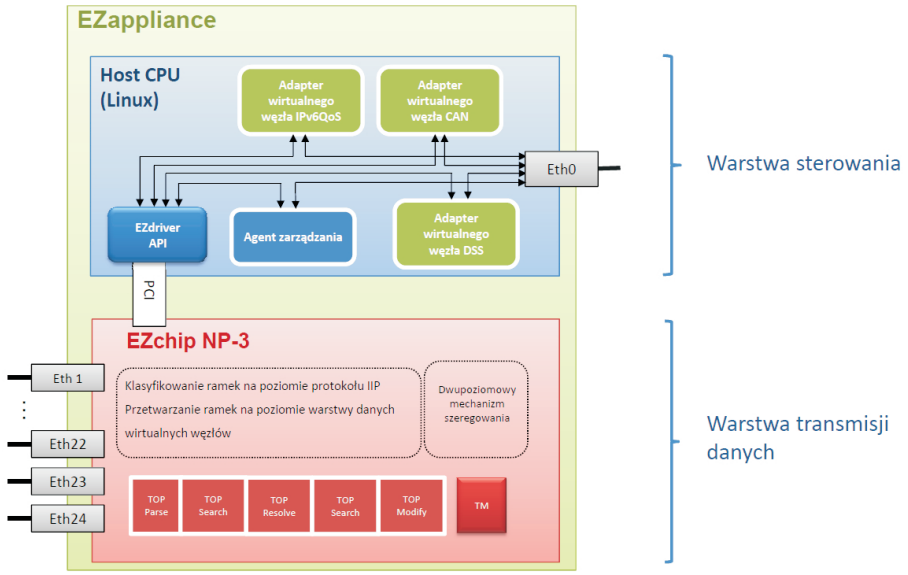
Algorytm mechanizmu szeregowania dla węzła Systemu IIP będzie także zrealizowany w taki sposób, iż część złożona obliczeniowo zostanie zaimplementowana całkowicie sprzętowo jako maszyna stanów FSM w układzie FPGA. Natomiast kontrola pracy tego modułu będzie się odbywać poprzez rejestry sprzętowe, których wartość będzie ustalana przez część sterującą algorytmu zrealizowaną w postaci programu i działającą w systemie Linux. Z tego względu, że mechanizm szeregowania węzła Systemu IIP musi współpracować z mechanizmami szeregowania poszczególnych węzłów wirtualnych taka realizacja pozwoli na zapewnienie efektywnej komunikacji pomiędzy tymi modułami.

5 Realizacja węzła Systemu IIP na platformie EZappliance

EZappliance jest kompaktową platformę sprzętową do szybkiego wdrażania aplikacji sieciowych [10]. Rozwiązanie składa się z w pełni programowalnego procesora sieciowego EZchip NP-3 pozwalającego na klasyfikowanie, modyfikowanie, przełączanie i zarządzanie przetwarzanymi pakietami [11]. EZchip NP-3 podłączony jest za pomocą interfejsu PCI do komputera wbudowanego (Host CPU) z zainstalowanym systemem Linux. EZappliance jest wyposażony w 24 porty Gigabit Ethernet. EZchip NP-3 zbudowany jest z kilku wyspecjalizowanych, programowalnych procesorów zoptymalizowanych zadaniowo (TOP) oraz dwóch mechanizmów zarządzających ruchem (TM). W systemie IIP do przetwarzania pakietów danych wykorzystano wszystkie rodzaje procesorów: TOPparse, TOPsearch, TOPresolve i TOPmodify. Mechanizmy modułu TM pozwalają na sprzętowe sterowanie ruchem. Obejmują one: pomiar, znakowanie, formowanie i szeregowanie pakietów. Konfiguracja modułu TM, wczytywanie programów dla poszczególnych procesorów oraz komunikacja z EZchip NP-3 odbywa się za pomocą oprogramowania EZdriver API dostarczonego przez producenta.

5.1 Klasyfikator

Platforma EZappliance zostanie użyta w Systemie IIP do realizacji węzła brzegowego. W przypadku węzła brzegowego można rozróżnić dwa podstawowe typy portów fizycznych: porty w kierunku sieci szkieletowej (porty szkieletowe), oraz porty w kierunku sieci dostępowych (porty dostępowe). Po odebraniu ramki na porcie fizycznym urządzenia, zostanie wykonana klasyfikacja ramki do odpowiedniego Równoległego Internetu. Klasyfikacja będzie przeprowadzona na podstawie nagłówka PIH w przypadku portów szkieletowych oraz na podstawie numeru portu fizycznego lub pola VLan Tag (nagłówek 802.1Q) w przypadku portów brzegowych. Po wykonaniu klasyfikacji, pole danych odebranej ramki



Rysunek 6. Realizacja wirtualnych węzłów systemu IIP na platformie EZappliance.

(bez nagłówka PI) zostanie przekazane do odpowiedniego portu wirtualnego, który pełni funkcję styku pomiędzy Systemem IIP i węzłem wirtualnym. Zasady klasyfikacji (nagłówek PI, numer portu fizycznego lub VLan Tag) dla każdego portu fizycznego mogą być dynamicznie modyfikowane (tzn. podczas działania EZappliance) przez zarządzanie Systemu IIP.

5.2 Węzeł wirtualny

W systemie IIP wirtualne węzły składają się z funkcjonalności warstwy danych i warstwy sterowania. EZchip NP-3 pozwala przede wszystkim na implementację warstwy danych, podczas, gdy warstwa sterująca każdego węzła wirtualnego musi być umieszczona na zewnętrznym serwerze podłączonym do EZappliance (Host CPU urządzenia EZappliance nie posiada niezbędnych zasobów do uruchomienia środowisk wirtualnych). Realizacja wirtualnych węzłów systemu IIP na platformie EZappliance jest pokazana na rysunku 6. EZchip NP-3 nie pozwala niestety na pełną wirtualizację warstw danych. Funkcjonalność przesyłania danych każdego z wirtualnych węzłów i przetwarzanie na poziomie samego protokołu IIP musi być implementowana w postaci pojedynczego zintegrowanego programu dla każdego procesora specjalizowanego TOP. Budowa procesorów TOP pozwala jednocześnie przetwarzać przychodzące ramki na poziomie protokołu IIP jak i na poziomie wirtualnego węzła. Przetwarzanie ramki danych w przynależnej do danego wirtualnego węzła odbywać się będzie z wykorzystaniem struktur przeszukiwania, w zależności od potrzeb tablice przełączania lub

tablice routingu. Powyższe tablice będą kontrolowane przez adaptery warstwy sterowania uruchomione w Host CPU i wykorzystujące interfejs programowy EZdriver API obudowujący funkcjonalnie interfejs fizyczny PCI.

5.3 Mechanizm szeregowania

Algorytm szeregowania bloków danych Systemu IIP będzie odwzorowany w EZappliance w postaci kombinacji dwóch mechanizmów: mechanizmu szeregowania pakietów typu WFQ (*Weighted Fair Queuing*) oraz mechanizmu kształtowania profilu ruchowego strumienia pakietów (typu *Single Leaky Bucket*). Celem działania tych mechanizmów jest zapewnienie zadanej wartości pasma dla poszczególnych Równoległych Internetów oraz niedopuszczenie do jej przekroczenia. Parametry działania powyższych mechanizmów będą mogły być dynamicznie modyfikowane przez agenta zarządzania Systemu IIP. Funkcjonalność szeregowania danych (na poziomie Systemu II, a także węzłów wirtualnych) będzie realizowana w EZappliance przez mechanizm TM.

6 Podsumowanie

W artykule przedstawiono realizację węzła Systemu IIP z wykorzystaniem trzech platform realizujących funkcje płaszczyzny przekazu danych: programistycznej Xen oraz sprzętowych NetFPGA i EZappliance. W każdym przypadku funkcje sterowania są realizowane przez dedykowaną platformę wykorzystującą oddzielną platformę Xen. Węzeł realizowany na platformie Xen jest najprostszy w implementacji ze względu na możliwość wykorzystania szeregu znanych rozwiązań. Jednak platformy NetFPGA i EZappliance ze względu na sprzętową realizację funkcji przekazu danych mogą pozwalać na uzyskanie zdecydowanie większej wydajności (ze względu na równoległy sposób przetwarzania danych). Obecnie projekt IIP jest na etapie implementacji i integracji węzła Systemu IIP na wymienionych platformach. Badania wydajnościowe węzłów Systemu IIP będą możliwe po zakończeniu tej fazy projektu.

Literatura

1. W. Burakowski, et al., Architektura Systemu IIP, KSTiT 2011, Łódź 2011.
2. W. Burakowski (ed.), et al., Specyfikacja systemu IIP – poziom 1 i 2 – wersja 1, Raport projektu „Inżynieria Internetu Przyszłości", POIG.01.01.02-00-045/09-00, 2011.
3. W. Burakowski, H. Tarasiuk, A. Bęben, System IIP for supporting „Parallel Internets (Networks)", Future Internet Assembly meeting, Ghent, 2010, http://fi-ghent.fi-week.eu/files/2010/12/1535-4-System-IIP-FIA-Ghent-ver1.pdf.
4. W. Burakowski, et al., Idealne urządzenie umożliwiające wirtualizację infrastruktury sieciowej w Systemie IIP, KSTiT 2011, Łódź 2011.
5. Citrix Systems Inc, Hawlett-Packard, IBM Corp.: Xen Users' Manual, 2008.
6. P. Barham, et al., Xen and the art of virtualization. in Proc. of the 19th ACM SOSP, New York, 2003.

7. B. De Schuymer, et al., Ebtables, http://ebtables.sourceforge.net/.
8. John W. Lockwood, et al., NetFPGA - An Open Platform for Gigabit-rate Network Switching and Routing; Proceedings of the 2007 IEEE International Conference on Microelectronic Systems Education, San Diego, June 2007.
9. M. Bilal Anwer, N. Feamster, Building a Fast, Virtualized Data Plane with Programmable Hardware, SIGCOMM VISA, Barcelona, Spain, August 2009.
10. EZchip Technologies, EZappliance Product Brief, http://www.ezchip.com/p_ezappliance.htm.
11. EZchip Technologies, NP-3 Product Brief, http://www.ezchip.com/p_np3.htm.

Koncepcja systemów dostępowych do Internetu Przyszłości

Rafał Krenz[1], Krzysztof Wesołowski[1], Paweł Szczepański[2], Paweł Gdula[2], Katrin Welikow[2], Ryszard Piramidowicz[2], Piotr Zwierko[3], Marek Natkaniec[4]

[1] Katedra Radiokomunikacji, Politechnika Poznańska
[2] Instytut Mikroelektroniki i Optoelektroniki, Politechnika Warszawska
[3] Instytut Telekomunikacji, Politechnika Warszawska
[4] Katedra Telekomunikacji, Akademia Górniczo-Hutnicza

Streszczenie W pracy przedstawiono koncepcję systemów dostępowych do Internetu Przyszłości. Przedstawiono podstawową organizację transmisji przewodowej bazującą na systemie Ethernet IEEE 802.3 z uwzględnieniem rozróżnienia transmisji w trzech rodzajach równoległych Internetów. Następnie przedstawiono specyfikę podstawowego bezprzewodowego systemu transmisyjnego w sieci dostępowej opartego na standardzie IEEE 802.11 przedstawiając stos modułów i interfejsy dla punktu dostępu do tej sieci. Z kolei podano podstawową koncepcję zaawansowanego bezprzewodowego systemu transmisyjnego w konfiguracji sieci kratowej bazującego na standardzie IEEE 802.11s. Przedstawiono możliwość dopasowania działania tej sieci do specyfiki systemu Internetu Przyszłości. Jako alternatywne rozwiązanie systemu dostępowego przedstawiono światłowodową sieć dostępową, w szczególności z zastosowaniem włókien polimerowych, przedstawiając zalety tej technologii i wyniki wykonanych w ramach projektu badań.

1 Wprowadzenie

Systemy transmisyjne są podstawowym środkiem wymiany danych w sieciach telekomunikacyjnych i komputerowych. Ich przepustowość jest jednym z głównych czynników, które określają pojemność sieci, a w konsekwencji różnorodność oferowanych za jej pomocą usług. Internet Przyszłości będzie wszechobecną siecią komputerową, która będzie korzystać z różnych systemów transmisyjnych. Ponieważ ilość danych do przesłania w sieci rośnie w sposób dramatyczny, zwiększają się również wymagania odnośnie systemów transmisyjnych. Sieć szkieletowa bazuje zazwyczaj na technice światłowodowej, która przy obecnym stanie techniki zapewnia bardzo wysoką przepływność. Rzeczywistym „wąskim gardłem" dla Internetu Przyszłości na poziomie systemu telekomunikacyjnego są więc systemy dostępowe, zarówno przewodowe jak i bezprzewodowe, szczególnie wobec stawianych wymagań na zachowanie odpowiedniej jakości obsługi (QoS). Należy przewidywać, że dostęp do Internetu Przyszłości będzie się odbywać za pomocą różnych sieci dostępowych takich jak: lokalna sieć kablowa

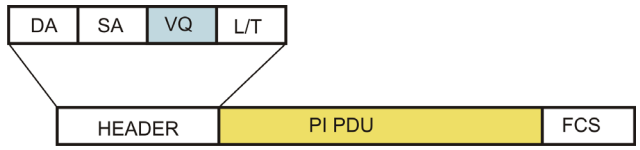
Ethernet, sieć w technologii ADSL/VDSL i ich odmianach, sieci WLAN według standardów IEEE 802.11, szerokopasmowe sieci bezprzewodowe bazujące na sieci komórkowej, np. HSPA, LTE, LTE-Advanced, itp., sieci WiMAX w różnych wersjach, czy wreszcie dedykowane sieci bazujące na technologii optoelektronicznej. Z przyczyn pragmatycznych badania nad sieciami dostępowymi realizowane w ramach projektu PO IG pn. *Inżynieria Internetu Przyszłości* skoncentrowały się na wybranych, lecz bardzo istotnych sieciach dostępowych: sieciach bezprzewodowych działających według standardów IEEE 802.11 a/g/n oraz s a także sieciach przewodowych wykorzystujących włókna optyczne.

W pracy tej przedstawiamy koncepcję systemów dostępowych do Internetu Przyszłości opisując specyfikę tych systemów wynikającą z idei wirtualizacji i koncepcji trzech Równoległych Internetów (RI): pakietowego IPv6 z zachowaniem ustanowionych poziomów jakości obsługi QoS (PS – *Packet Switching*), z komutacją strumieni (PI-DSS – *Parallel Internet – Data Stream Switching*) oraz sieci świadomej zawartości (CAN – *Content Aware Networks*). W kolejnych paragrafach prezentujemy w skrócie koncepcję systemów podstawowych bazujących na standardzie Ethernet oraz IEEE 802.11, bezprzewodowego systemu zaawansowanego opierającego się na standardzie IEEE 802.11s, dającego możliwości działania w trybie komutacji pakietów i emulującego system z komutacją strumieni a także opisujemy media transmisyjne dla światłowodowej sieci dostępowej pierwszej generacji. Artykuł zakończony jest wnioskami.

2 Podstawowy system transmisyjny

2.1 Przewodowy system transmisyjny

W Systemie IIP przyjęto założenie, że podstawowym sposobem dla łączności przewodowej będzie system Ethernet IEEE802.3 [1] w wersji 1Gb/s 1000Base-T. Podstawową przyczyną takiego wyboru jest powszechna dostępność tej technologii dla różnych typów fizycznych mediów, wysoka przepływność (1Gb/s) oraz łatwość definiowania nowych formatów danych, opartych na standardowych ramkach. Ramka ethernetowa w systemie IIP zawiera standardowy nagłówek Ethernetu zawierający adres przeznaczenia (DA), adres własny (SA), wyróżnik protokołu (L/T) oraz opcjonalne pole VLAN-tag (VQ).



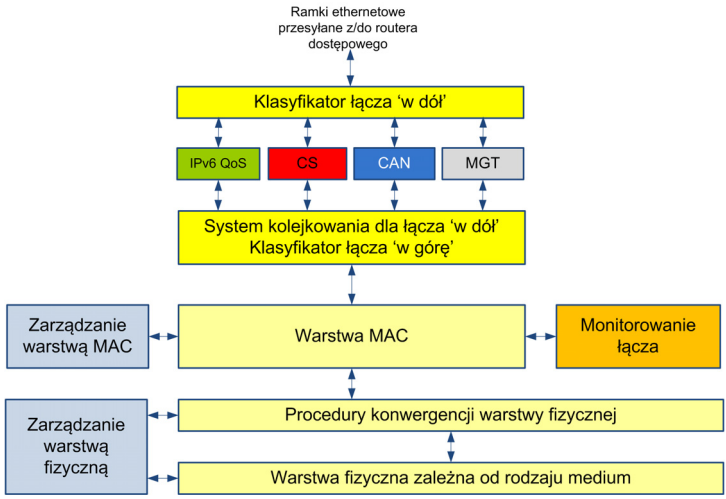
Rysunek 1. Format ramki Ethernet w Systemie IIP

Pole danych (*payload*, ozn. PI PDU, PI – *Parallel Internet*) jest zdefiniowane przez specyfikację systemu IIP i składa się z jednobajtowego wyróżnika równo-

ległego Internetu (PI) oraz pola danych PDU charakterystycznego dla równoległego Internetu. W niektórych typach połączeń, ze względu na ograniczenia implementacyjne zakłada się oznaczanie przynależności danych do równoległego Internetu poprzez wykorzystanie opcjonalnego pola VLAN.

2.2 Bezprzewodowy system transmisyjny

Sieci standardu IEEE 802.11 [2] stanowią bardzo ważny element systemu IIP, zapewniając bezprzewodowy dostęp użytkownikom w ramach lokalnych sieci komputerowych. W typowym scenariuszu pracy terminal mobilny użytkownika jest przyłączony do sieci bezprzewodowej z użyciem tzw. punktów dostępu (APs – *Access Points*). Punkty te są przyłączone do sieci rdzeniowej Internetu Przyszłości łączem przewodowym zgodnym ze standardem IEEE 802.3. W obecnej wersji standardu IEEE 802.11 uwzględniono możliwość zastosowania wielu różnych warstw fizycznych (rozszerzenia IEEE 802.11a/b/g/n), pracujących w pasmach 2,4 GHz oraz 5 GHz. W infrastrukturze sieciowej systemu IIP może zostać zastosowana każda z tych warstw. Rozszerzenie IEEE 802.11e definiuje dla podwarstwy MAC (*Medium Access Control*) dwie funkcje dostępu do medium, które posiadają mechanizmy pozwalające świadczyć usługi z odpowiednią jakością QoS (*Quality of Service*): EDCA (*Enhanced Distributed Channel Access*) i HCCA (*Hybrid Coordination Function Controlled Channel Access*). Pierwsza z nich może zostać użyta zarówno w trybie pracy *ad-hoc* jak i w sieci z infrastrukturą (w której działa punkt dostępu). Druga z funkcji dostępu jest możliwa tylko w sieci z infrastrukturą. W ramach praktycznej realizacji Internetu Przyszłości przewiduje się użycie wyłącznie funkcji EDCA. Funkcja EDCA pozwala na różnicowanie czterech klas ruchu (głosu, wideo, ruchu o najwyższej możliwej jakości dostarczania oraz ruchu tła) dzięki zastosowaniu czterech sprzętowych kolejek umieszczonych w karcie lokalnej sieci bezprzewodowej. Prawidłowy sposób różnicowania ruchu w funkcji EDCA jest dodatkowo zapewniany poprzez odpowiednią konfigurację kilku parametrów definiowanych na poziomie warstwy MAC. Każdy z punktów dostępu sieci standardu IEEE 802.11 jest przyłączany z zastosowaniem sieci Ethernet do dedykowanego portu routera dostępowego (AR – *Access Router*) w sieci szkieletowej. Podczas realizacji projektu założono, że wszystkie urządzenia sieci bezprzewodowej, w tym również punkty dostępu będą umożliwiały działanie trzech rodzajów równoległych Internetów rozważanych w projekcie, tzn. IPv6 QoS, CAN i PI-DSS. Informacja związana z typem równoległego Internetu będzie przesyłana zarówno w sieci przewodowej IEEE 802.3 (między routerem dostępowym (brzegowym) do sieci bezprzewodowej a punktem dostępu) jak również w sieci bezprzewodowej IEEE 802.11 (między punktem dostępu a terminalem mobilnym) z użyciem nagłówka VLAN (*Virtual Local Area Network*) standardu IEEE 802.1q. Na rys. 2 przedstawiono niezbędne moduły oraz interfejsy umieszczone w punkcie dostępu do sieci standardu IEEE 802.11. Wszystkie funkcje warstwy fizycznej oraz MAC są zgodne ze standardem IEEE 802.11. Dodatkowymi modułami są: klasyfikatory dla łącza „w górę" oraz „w dół", system kolejkowania ramek dla łącza „w dół" oraz moduł monitorowania łącza radiowego. Klasyfikatory oraz system kolejkowania



Rysunek 2. Stos modułów oraz interfejsy dla punktu dostępu do sieci standardu IEEE 802.11

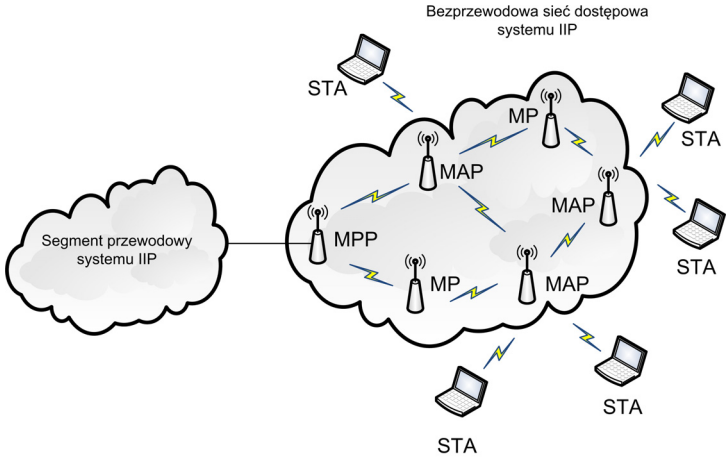
ramek mają za zadanie separację ruchu w ramach poszczególnych równoległych Internetów, a także separację ruchu pomiędzy łączem „w górę" oraz „w dół". System monitorowania łącza radiowego dzięki analizie wszystkich ramek przesyłanych w kanale radiowym, realizuje szereg pomiarów pozwalających określić m.in. osiąganą przepływność dla każdego z równoległych Internetów oraz pozostałe do wykorzystania zasoby. Wykonywane w czasie rzeczywistym pomiary są następnie przesyłane z użyciem interfejsu zarządzającego do routera dostępowego, który na ich podstawie decyduje o przyjęciu bądź też odrzuceniu nowych zgłoszeń (*Admission Control*).

2.3 Zaawansowany bezprzewodowy system transmisyjny

Zaawansowana bezprzewodowa sieć dostępowa systemu IIP zostanie zbudowana z wykorzystaniem architektury kratowej (rys. 3), do której zalet należą m.in. niski koszt implementacji (brak infrastruktury), łatwa rozbudowa w razie konieczności zwiększenia pokrycia oraz wysoka niezawodność. W jej skład wchodzi kilka typów węzłów:

- węzły MP (Mesh Point), pośredniczące w przesyłaniu pakietów pomiędzy pozostałymi węzłami sieci kratowej,
- węzły MAP (Mesh Access Point), pełniące funkcję MP oraz umożliwiające stacjom końcowym (STA) dostęp do sieci kratowej,
- węzły MAP (Mesh Access Point), pełniące funkcje MP oraz zapewniające dostęp do segmentu przewodowego systemu IIP.

Bezprzewodowa sieć kratowa będzie zgodna z wdrażanym obecnie standardem IEEE 802.11s (w końcowej fazie standaryzacji [3]), który zostanie uzupeł-



Rysunek 3. Architektura zaawansowanej bezprzewodowej sieci dostępowej systemu IIP.

niony o specyficzne funkcje, niezbędne np. do wirtualizacji łączy. Jest on rozszerzeniem powszechnie wykorzystywanego w bezprzewodowych sieciach komputerowych standardu 802.11 i korzysta ze sprawdzonych rozwiązań warstwy fizycznej (802.11a/b/g/n). Ze względu na ograniczone zasoby (moc obliczeniowa, pamięć) węzłów sieci kratowej, nie będzie w niej realizowana wirtualizacja węzłów. W bezprzewodowej sieci o architekturze kratowej bardzo ważnym zagadnieniem mającym decydujący wpływ na jakość jej funkcjonowania jest wybór odpowiedniego protokołu wielodostępu (MAC). Powszechnie stosowany w bezprzewodowych sieciach komputerowych protokół dostępu przypadkowego EDCA, oprócz niewątpliwej zalety, jaką jest jego prostota, posiada również wady ujawniające się szczególnie przy transmisji wieloetapowej (multi-hop) – na skutek licznych kolizji przepustowość sieci drastycznie maleje nawet w sieci o niewielkiej liczbie węzłów [4]. Dlatego standard 802.11s przewiduje opcjonalne wykorzystanie szczelinowego protokołu MCCA, który umożliwia dynamiczną rezerwację szczelin przez poszczególne węzły, zgodnie z ich zapotrzebowaniem na pasmo transmisyjne [5]. W opisywanej sieci dostępowej systemu IIP wykorzystany zostanie właśnie ten protokół MAC, gdyż oprócz lepszej wydajności umożliwia on również efektywną emulację połączeń z komutacją strumieni (PI CS) oraz dostarcza mechanizmów QoS dla transmisji pakietowej (PI IP QoS). Drugim elementem, decydującym o jakości funkcjonowania bezprzewodowej sieci kratowej, są zastosowane mechanizmy przydziału zasobów radiowych (RRC – Radio Resource Control). Jest to związane z dużymi wahaniami chwilowej przepustowości łączy (na skutek zmian parametrów kanału radiowego), która w opisywanym rozwiązaniu może zmieniać się od kilku do kilkuset Mb/s. Dlatego konieczne jest opracowanie efektywnych metod zarządzania pasmem, z jednej strony spełniających wymagania równoległych Internetów, a z drugiej strony dopasowanych do specyfiki wieloetapowej

transmisji radiowej. Ze względu na wspomniane powyżej specyficzne problemy występujące w zaawansowanej bezprzewodowej sieci dostępowej systemu IIP, nie będzie w niej spełnionych wiele założeń przyjętych dla segmentu przewodowego, w szczególności ograniczony zostanie zakres usług oferowanych w dostępie bezprzewodowym. Dodatkowo, może się on zmieniać wraz lokalizacją stacji końcowej w systemie.

3 Media transmisyjne dla światłowodowej sieci dostępowej nowej generacji

W sieciach dostępowych, stanowiących najniższą warstwę systemów telekomunikacyjnych, włókna optyczne systematycznie wypierają rozwiązania miedziane, pozwalając na radykalne zwiększenie przepustowości i niezawodności. Należy jednak podkreślić, że o ile nie ma żadnych wątpliwości co do wyboru medium transmisyjnego dla sieci dalekosiężnych i metropolitalnych, to włókna dla sieci o charakterze dostępowym, a szczególnie dla sieci wewnątrz budynków są ciągle przedmiotem intensywnych badań i analiz, prowadzonych zarówno w jednostkach R& D największych operatorów i producentów włókien, jak też wiodących ośrodkach naukowych [6], [7]. Wybór optymalnego medium transmisyjnego dla sieci dostępowych jest ciągle sprawą otwartą, a prowadzone badania koncentrują się zarówno na aspektach optymalizacji i modyfikacji włókien klasycznych (przede wszystkim kwarcowych jedno- i wielomodowych, ale również wielomodowych włókien polimerowych), jak też na poszukiwaniach zupełnie nowych mediów transmisyjnych (jak np. włókna mikrostrukturalne).

Optyczna sieć dostępowa rozwijana w Instytucie Łączności w ramach projektu Inżynieria Internetu Przyszłości zapewnia środowisko pomiarowe do testowania włókien optycznych, zarówno kwarcowych jak i polimerowych, starannie wybranych pod kątem zastosowania w systemach dostępowych funkcjonujących w sieci IIP. Równolegle trwają intensywne prace w zakresie projektowania nowych typów włókien optycznych, które byłyby wolne od ograniczeń cechujących rozwiązania konwencjonalne i mogły zaoferować nową jakość w systemach dostępowych. Wśród badanych rozwiązań szczególne miejsce zajmują mikrostrukturalne włókna polimerowe.

Głównymi zaletami włókien polimerowych (POF – *Polymer Optical Fiber*) są ich znakomite właściwości mechaniczne, łatwość konstruowania światłowodów o dużych rozmiarach rdzenia, niskie koszty produkcji i niezwykle prosty sposób obróbki. Podstawową wadą jest wysoka tłumienność oraz wysokie wartości dyspersji modowej, będące konsekwencją rozmiarów rdzenia [8].

Na korzyść technologii POF silnie przemawia to, że spełnia ona zarówno wymagania niskich kosztów podzespołów, instalacji okablowania i późniejszego utrzymania sieci, jak też wymagania dotyczące możliwości prostego i zapewniającego bezpieczeństwo użytkownikowi samodzielnego montażu. Dotychczasowe doniesienia dotyczące pracy włókien POF w systemach dostępowych pozwalają przypuszczać, że jeśli ww. czynniki ograniczające zostaną przezwyciężone, włókna plastikowe zdominują światowy rynek mediów transmisyjnych dedyko-

wanych systemom dostępowym. Prace prowadzone w ramach projektu IIP mają na celu próbę opracowania włókien plastikowych, przeznaczonych do zastosowań w systemach wewnątrz-budynkowych, o istotnie polepszonych parametrach transmisyjnych - ograniczonej dyspersji modowej i poprawionych parametrach strat zgięciowych.

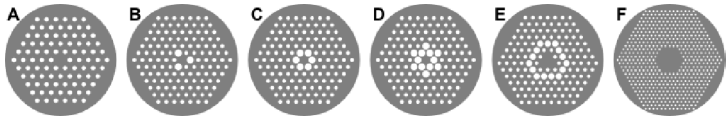
Poniżej przedstawione wyniki analiz numerycznych koncentrują się na parametrach transmisyjnych mikrostrukturalnych włókien polimerowych o siatce heksagonalnej, projektowanych pod kątem ograniczenia ilości modów przy zachowaniu stosunkowo dużego pola modowego, oraz zredukowania strat zgięciowych. We wszystkich rozważaniach zakładano, że materiałem bazowym jest polimetakrylan metylenu (PMMA), powszechnie stosowany jako materiałrdzeniowy we włóknach POF, charakteryzujący się dobrze opanowaną technologią syntezy i pozwalający na wytwarzanie klasycznych włókien POF o odpowiednio dobrych parametrach optycznych. Wszystkie symulacje zostały przeprowadzone dla długości fali 650 nm, typowo stosowanej w systemach polimerowych ze względu na dobre dopasowanie do minimum tłumienności PMMA i dostępność tanich źródeł promieniowania. Założono, że wprowadzana w obrębie płaszcza mikrostrukturyzacja powinna spełniać następujące zadania:

- znacząco ograniczać liczbę prowadzonych modów przy zachowaniu maksymalnie dużej powierzchni rdzenia,
- znacząco ograniczać straty makro-zgięciowe, które powinny być porównywalne z wartościami określonymi normą ITU G.652 dla typowych włókien,
- zapewniać tłumienie na poziomie nie wyższym niż poziom strat materiałowych PMMA.

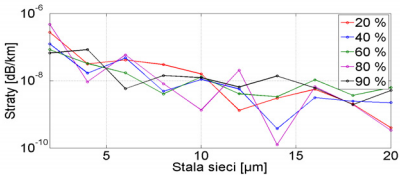
W pierwszym etapie optymalizacji geometrii mikrostrukturalnego włókna POF przyjęto metodę polegającą na analizie możliwości zaadaptowania typowych struktur stosowanych dotychczas w kwarcowych włóknach fotonicznych. Z tego powodu poddano testom powszechnie stosowaną w przypadku włókien kwarcowych strukturę heksagonalną z rdzeniem dielektrycznym, przedstawioną na rys. 4.A oraz jej liczne modyfikacje, z których przykładowe pokazano na rys. 4.B-4.F. We wszystkich przebadanych przypadkach centralnie położony rdzeń otoczony jest przez kilka pierścieni złożonych z otworów powietrznych ułożonych w sieć heksagonalną. Wszystkie struktury (zdefiniowane w materiale polimerowym) poddano symulacjom numerycznym, zmieniając w szerokim zakresie ich podstawowe parametry geometryczne, takie jak stała sieci Λ i współczynnik wypełnienia otworów d/Λ, gdzie d oznacza średnicę otworu powietrznego, co dało to w rezultacie setki tysięcy przeanalizowanych kombinacji.

Całkowity zestaw wyników uzyskanych dla poszczególnych przypadków jest w oczywisty sposób niemożliwy do przedstawienia "przypadek po przypadku", dlatego w niniejszej pracy przedstawione zostaną jedynie przykładowe zestawy wyników (rys. 5), uzyskane dla geometrii oznaczonej jako C i charakteryzującej się jak dotąd największym potencjałem z punktu widzenia zdefiniowanych wyżej wymagań.

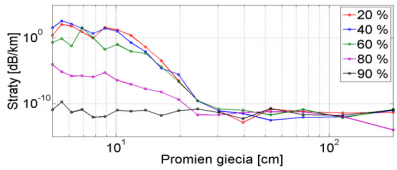
Geometria C wykazuje wszystkie pożądane właściwości, takie jak redukcja strat zgięciowych, niskie straty propagacyjne oraz duże pole modu. Dokładniej-



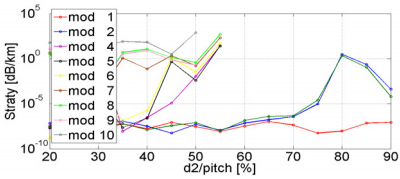
Rysunek 4. Schematy przykładowych, analizowanych geometrii przekroju poprzecznego włókna mikrostrukturalnego z rdzeniem dielektrycznym, otwory tworzą siatkę heksagonalną.



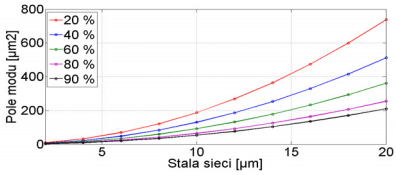
(a) Straty zgięciowe w funkcji stałej sieci dla różnych współczynników wypełnienia d_1/Λ



(b) Straty makrozgięciowe w funkcji promienia gięcia dla różnych współczynników wypełnienia d_1/Λ



(c) Straty propagacyjne w funkcji współczynnika wypełnienia wewnętrznych otworów, kolejne krzywe odpowiadają modom coraz wyższych rzędów

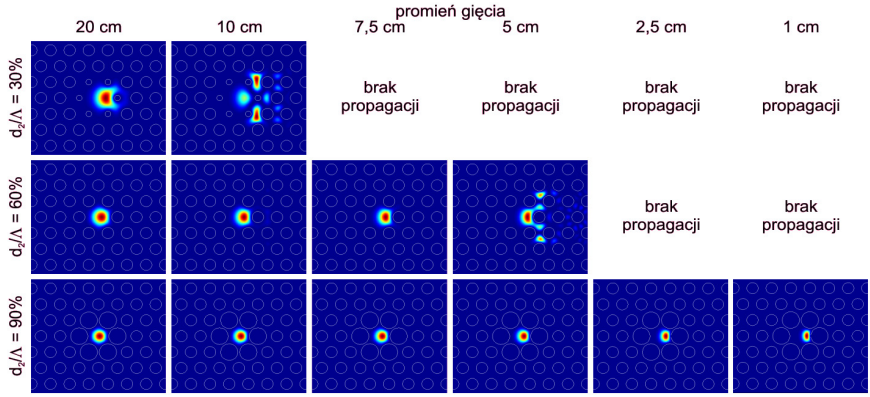


(d) Rozmiar pola modowego w funkcji stałej sieci, kolejne krzywe odpowiadają zmieniającym się współczynnikom wypełnienia wewnętrznych otworów d_1/Λ

Rysunek 5. Zestaw przykładowych wyników dla geometrii C, przy czym $\Lambda = 6, 8, ..., 16$ mm; d_2/Λ =60 %; $d_1/\Lambda = 20, 40, 60, 80, 90$ %

sze zobrazowanie wpływu współczynnika wypełnienia na straty zgięciowe przedstawiono na rys. 3 dla trzech wybranych struktur o współczynniku wypełnienia zewnętrznych otworów równym 45% i i współczynniku wypełnienia otworów wewnętrznych wynoszącym odpowiednio 30, 60 i 90%. Dla kolejnych promieni gięcia zapisano rozkład poprzeczny modu podstawowego.

Można zaobserwować, że ze zmniejszaniem się promienia gięcia, maksimum rozkładu pola przesuwa się w kierunku obszaru o wyższym współczynniku załamania, którym jest fragment ściskany (wewnętrzna część krzywizny zgięcia). Jeśli wewnętrzne otwory są małe, pole przenika między nimi i koncentruje się w mostkach polimerowych. Przy dalszym zmniejszaniu promienia pole ulega wypromieniowaniu i nie propaguje się żaden mod. Powiększanie otworów redukuje obszary polimerowe pomiędzy nimi i stwarza dookoła rdzenia bufor zapobiegający ucieczce promieniowania nawet przy zginaniu na bardzo małych krzywiznach.



Rysunek 6. Analiza zmian rozkładu poprzecznego pola dla modu podstawowego w zależności od promienia gięcia włókna mPOF, współczynnik wypełnienia pierścieni od 2 do 6, d_2/Λ=45%.

Dla dużych otworów wewnętrznych kierunek zginania nie ma wpływu na wartość strat zgięciowych.

Wyniki obliczeń potwierdzają możliwość efektywnego sterowania stratami propagacyjnymi i zgięciowymi, manipulacji liczbą modów i polem modowym włókien mPOF – w ich wyniku zaproponowano również pierwsze struktury o polepszonych parametrach propagacyjnych – ograniczonym tłumieniu, obniżonej modowości i zredukowanych stratach zgięciowych przy zachowaniu stosunkowo dużych rozmiarów rdzenia. Staranna optymalizacja geometrii zaproponowanego włókna powinna pozwolić na dalszą poprawę jego parametrów i w konsekwencji umożliwić zaprojektowanie testowych struktur do pierwszych eksperymentów technologicznych. Przeprowadzona wstępna analiza uwarunkowań technologicznych wskazuje na możliwość wykonania próbnej serii włókien mPOF przy wykorzystaniu potencjału Instytutu Technologii Materiałów Elektronicznych. Uzyskane wyniki będą przedmiotem kolejnych publikacji.

Literatura

1. IEEE 802.3-2008 IEEE Standard for Information technology-Specific requirements - Part 3: *Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications*. Revision of IEEE Std. 802.3-2008 26 December 2008.
2. IEEE Standard for Information technology-Telecommunications and information exchange between systems-Local and metropolitan area networks-Specific requirements – Part 11: *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE Std 802.11-2007 (Revision of IEEE Std 802.11-1999), pp. C1-1184, 2007.

3. IEEE, Wireless LAN medium access control (MAC) and physical layer (PHY) specifications: Amendment s: Mesh networking, IEEE P802.11s/D4.01, February 2010.
4. G. R. Hiertz, S. Max, Y. Zang, T. Junge, D. Denteneer, IEEE 802.11s MAC Fundamentals, IEEE International Conference on Mobile Adhoc and Sensor Systems, MASS 2007, str. 1-8 5.
5. R. Krenz, IEEE 802.16 Wireless Mesh Networks Capacity Assessment Using Collision Domains, IARIA International Journal On Advances in Internet Technology, Vol. 3, No. 1-2, pp. 77-87, 2010.
6. M. Shimizu, *Optical Access Network Technology (Optical Media Technology) Toward Expansion of FTTH*, NTT Technical Review, 2008;
7. M. Li, *Bend-Insensitive Optical Fibers Simplify Fiber-to-the-Home Installations*, Optoelectronics & Optical Communications, 2008
8. Y.Koike, T. Ishigure, and E. Nihei, *High-Bandwidth Graded-Index Polymer Optical Fiber*, J. of Lightwave Technology, vol. 13, 1995, pp.1475-1489; S. Moradi et al. *A Fast and Simple Method for Fabrication of Polymer Optical Fiber*, TEMU2006, Heraklion 2006

Zarządzanie i wymiarowanie zasobów w L1/L2 Systemu IIP

Janusz Granat[1], Jakub Bielecki[1], Wojciech Szymak[1],
Krzysztof Wajda[2], Janusz Gozdecki[2], Halina Tarasiuk[3]

[1] Instytut Łączności - Państwowy Instytut Badawczy
[2] Akademia Górniczo-Hutnicza
[3] Instytut Telekomunikacji, Politechnika Warszawska

Streszczenie Celem rozdziału jest przedstawienie systemu zarządzania dzia lającego w warstwie L1 i L2 architektury Systemu IIP. Na wstępie została krótko scharakteryzowana architektura Systemu IIP [1] oraz umiejscowiono moduły zarządzania warstwy L1/L2 w architekturze zarządzania Systemu IIP. Następnie omówiono zadania modułów zarządzania dla warstwy L1/L2 takich jak zarządca zasobów, zarządca konfiguracji, zarządca awarii, zarządca polityki oraz modułmonitoringu.

1 Wprowadzenie

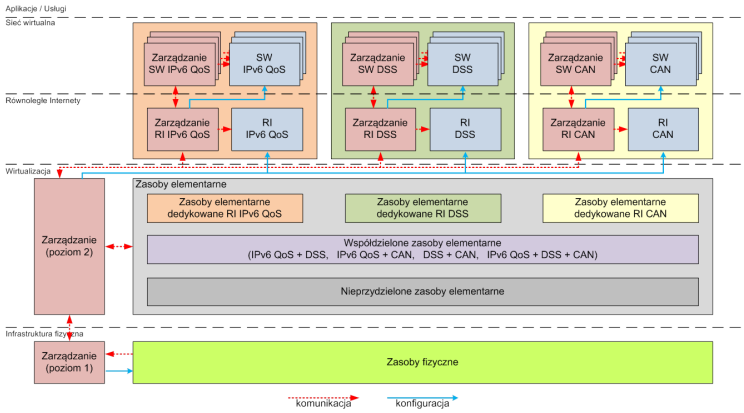
System IIP jest tworzony w ramach projektu Inżynieria Internetu Przyszłości (IIP). Architektura systemu [1,2,3,4] składa się z 4 podstawowych warstw: fizycznej, wirtualizacji, Równoległych Internetów (RI) oraz sieci wirtualnych. Dwie pierwsze warstwy (L1/L2) stanowią infrastrukturę dla stworzenia tzw. Równoległych Internetów w warstwie 3. Koncepcja Równoległych Internetów [1] polega na możliwości działania niezależnych sieci, które mogą mieć różne protokoły dla płaszczyzny danych i sterowania, a jednocześnie współdzielą tą samą infrastrukturę fizyczną sieci dzięki technice wirtualizacji węzłów i łączy. W ramach projektu trwają prace nad definicją trzech Równoległych Internetów: IPv6 QoS (*Quality of Service*), CAN (*Content Aware Networks*) i DSS (*Data Streams Switching*). Ponadto, w systemie wyróżniamy płaszczyznę zarządzania, która jest definiowana dla każdej z warstw architektury wraz z interfejsami do komunikacji między warstwowej. Niniejszy artykuł opisuje ogólną specyfikację oraz aspekty definiowania i realizacji systemu zarządzania dla dwóch najniższych warstw architektury Systemu IIP, nazwanych L1 i L2 (odpowiednio warstwy fizycznej i wirtualizacji).

Obecnie są prowadzone prace w ramach inicjatyw i projektów o zasięgu krajowym i międzynarodowym, które są poświęcone definicji architektur, w tym systemów zarządzania dla Internetu Przyszłości (ang. Future Internet). Wśród inicjatyw zasługują na wyróżnienie prace mające na celu stworzenie architektury dla systemu zarządzania prowadzone w ramach inicjatywy MANA (ang. Management and Service-aware Networking Architectures) [7]. Kolejną inicjatywą jest grupa robocza ITU-T (Future Network), która opracowała założenia

i cele projektowe dla Internetu Przyszłości [7]. Jednymi z głównych celów projektowych jest wirtualizacja zasobów oraz zarządzanie siecią. Oba z tych aspektów są również przedmiotem prac w ramach definiowania architektury dla Systemu IIP.

2 Koncepcja systemu zarządzania

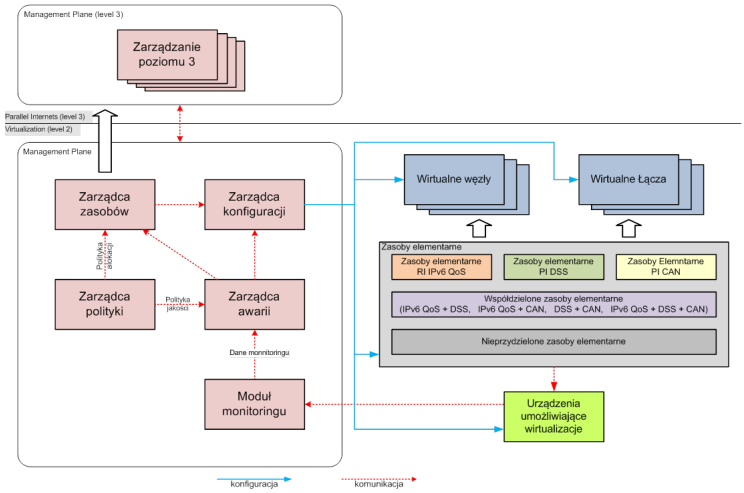
System zarządzania dla Systemu IIP realizuje następujące funkcje: (1) tworzenie wirtualnych węzłów i łączy oraz topologii dla Równoległych Internetów, (2) przydział zasobów na podstawie przyjętego mechanizmu wymiarowania (3) monitorowania. Zarządzane zasoby są reprezentowane w postaci obiektów opisanych za pomocą właściwych parametrów. Przyjęty model systemu zarządzania opisuje trzy główne aspekty: organizacyjne, informacyjne oraz funkcjonalne.



Rysunek 1. Architektura systemu zarządzania

Architektura zarządzania systemem IIP (Rys.1) obejmuje cztery warstwy architektury: infrastruktury fizycznej, wirtualizacji, Równoległych Internetów i sieci wirtualnych. W pierwszej warstwie zarządzaniu podlegają zasoby fizyczne (infrastruktura fizyczna), czyli węzły i łącza, które podlegają procesowi wirtualizacji. Węzły i łącza są podstawowymi zasobami dla tworzenia Równoległych Internetów. System zarządzania na poziomie L2 architektury (wirtualizacji) odpowiada za obsługę środowisk wirtualizacyjnych. Do jego zadań należy określanie topologii RI oraz wymiarowanie i przydział zasobów dla poszczególnych sieci. Zarządzanie na poziomie L3 architektury (RI) jest zbiorem podsystemów zarządzania dedykowanym do obsługi konkretnego RI i dostarczanym przez operatora danego RI. Funkcjonalność tej warstwy zarządzania jest ściśle powiązana z architekturą danego RI. Poziom czwarty zarządzania, czyli zarządzanie sieciami wirtualnymi jest zestawem kolejnych podsystemów dedykowanych do zarządzania siecią wirtualną w danym RI. Warstwa ta może być zarządzana przez innego

operatora niż warstwa L3 architektury, Artykuł przedstawia szczegóły systemu zarządzania dla poziomu L1 i L2 architektury Systemu IIP. Opis podsystemów zarządzania warstw L3 i L4 wykracza poza ramy tego artykułu. System zarządzania poziomu L2 rozróżnia trzy poziomy zasobów: fizyczne, elementarne i wirtualne (Rys. 2). Zasoby fizyczne są to fizyczne urządzenia podłączone do serwerów (dyski twarde, CPU, pamięć, itp.). Zasoby elementarne są wysokopoziomowym opisem zasobów fizycznych, abstrahującym od platformy sprzętowej; są to zasoby możliwe do przydzielenia w środowisku wirtualizacyjnym. Zasoby wirtualne, są to łącza i węzły utworzone w środowiskach wirtualizacyjnych z zasobów elementarnych.



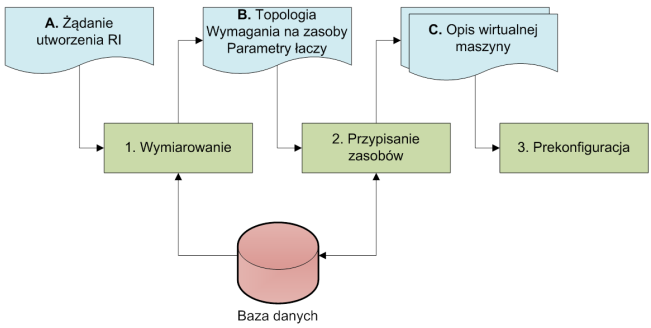
Rysunek 2. Architektura systemu zarządzania L2 systemu IIP

System zarządzania poziomu L2 ma budowę modularną. Składa się z 5 połączonych modułów: zarządcy zasobów, zarządcy konfiguracji, zarządcy awarii, zarządcy polityki oraz modułu monitoringu. Zarządca zasobów przechowuje informacje o zasobach. Odpowiada za przydział zasobów do wszystkich RI oraz obsługę awarii węzłów wirtualnych. Moduł dostarcza także informacji o zasobach dla zarządzania poziomu L3. Zarządca zasobów zawiera w sobie logikę wymiarowania (patrz punkt 4 niniejszego artykułu). Zarządca konfiguracji odpowiada za wstępną konfiguracje wszystkich RI. Konfiguracja ta jest ograniczona tylko do tych elementów, które są konieczne do uruchomienia i podłączenia się do zarządzania poziomu L3. Moduł ten odpowiada także za zarządzanie maszynami wirtualnymi w ramach zarządcy wirtualizacji. Zarządca awarii odpowiada za wykrywanie niewłaściwego funkcjonowania sieci i zasobów. Wykrywanie problemów oparte jest o dane z modułu monitoringu. W razie wykrycia awarii, moduł realizuje reguły określone przez zarządcę polityki i powiadamia odpowiedni moduł konfiguracyjny (zarządca konfiguracji lub zarządca zasobów). Zarządca polityki

odpowiada za bezpieczeństwo i nieprzerwane działanie systemu. Moduł definiuje kilka typów reguł: politykę jakości dla detekcji błędów i zapobiegania awariom, politykę reakcji na błędy, reguły dla przywracania działania po awarii, politykę alokacji zasobów.

3 Procesy zarządzania

Moduły zarządzania poziomu drugiego realizują kilka procesów. Wśród nich należy wyróżnić: wymiarowanie, przydział zasobów, wstępną konfigurację węzłów oraz monitoring (Rys. 3). Proces wymiarowania definiuje topologie każdego RI. Na podstawie informacji o dostępnych zasobach, lokalizacji łączonych węzłów oraz wymagań co do wydajności sieci, proces określa które węzły i łącza wirtualne będą stanowiły topologię wymiarowanego RI, oraz jaką część zasobów węzła należy przypisać dla danego RI.



Rysunek 3. Procesy w zarządzaniu

Proces przypisania zasobów decyduje o przydziale konkretnych zasobów do RI. Stanowi on łącznik między wysokopoziomowym opisem sieci uzyskanym w procesie wymiarowania, a niskopoziomowym opisem dostępnych urządzeń. W praktyce ten proces generuje pliki z opisem wirtualnych węzłów, a także aktualizuje bazę danych informacji o zasobach. Proces wstępnej konfiguracji węzłów ("prekonfiguracja") odpowiada za utworzenie węzła wirtualnego w środowisku wirtualizacyjnym. Na podstawie pliku z opisem węzła, konfiguruje zarządcę wirtualizacji i uruchamia maszyny wirtualne.

4 Wymiarowanie zasobów

Niniejsza sekcja zawiera opis zasad przydziału zasobów w sieci (czyli wymiarowania - *provisioning*) niezbędnych w procesie tworzenia wirtualnych węzłów i łączy dla poszczególnych Równoległych Internetów. Zasoby sieci podlegające wymiarowaniu obejmują: przepustowości łączy, pamięć RAM, pamięć masową,

moc obliczeniową w węzłach systemu oraz bufory (kolejki). Podstawowe założenia uwzględniają statyczny przydział zasobów do poszczególnych Internetów (tzn. nie podlegający ciągłym zmianom, a co najwyżej okresowo weryfikowany) oraz brak współdzielenia (w tym również brak czasowego odstępowania) zasobów pomiędzy poszczególnymi Równoległymi Internetami. Przydział zasobów jest dokonywany w długiej skali czasowej. W prototypie systemu zarządzania zaprojektowany jest w tym celu moduł wymiarowania, znajdujący się na poziomie L2 systemu zarządzania, dysponujący pełną wiedzą o parametrach sieci, a więc będący w stanie także dokonać optymalizacji przydziału zasobów.

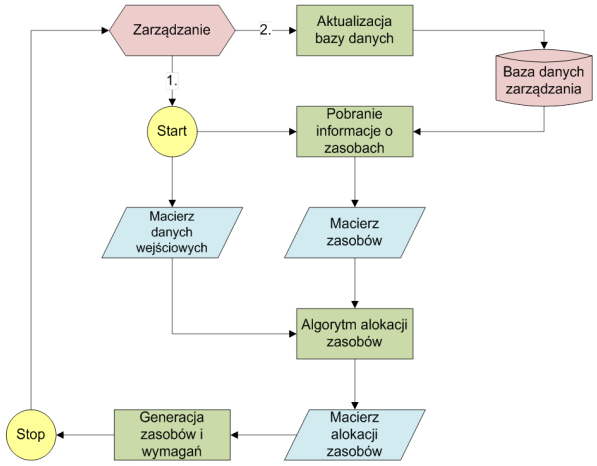
Podstawową zasadę opisującą przydział zasobów do Równoległych Internetów opisuje nierówność 1.

$$R_{resource} \geqslant R_{IPv6_QOS} + R_{CAN} + R_{DSS} + R_{mgt} \tag{1}$$

gdzie R z odpowiednim indeksem oznacza zagregowane zasoby przydzielone do konkretnego Równoleg lego Internetu.

Na obecnym etapie rozwoju systemu trwają prace nad szczegółowymi algorytmami wymiarowania. Główne problemy do rozwiązania w tym obszarze to definicja pojemności sieci, definicja pojemności węzła sieci, rozpoznanie zależności między pojemnością węzła fizycznego a pojemnością węzłów wirtualnych, metodologia dla wymiarowania sieci, która uwzględniałaby zależność między wymaganą pojemnością łącza wirtualnego a pojemnością węzła wirtualnego.

Podstawowe sformułowanie problemu wymiarowania obejmuje identyfikację potrzeb poszczególnych RI, rozpoznanie zasobów dostępnych w węzłach oraz w łączach a następnie przydzielenie zasobów, ze sprawdzeniem możliwości realizacji przydziału zasobów. Realizacja wymiarowania może odbywać się w pętli sprzężenia, wskazanej na Rys. 4, przy czym kryterium inicjowania procedury wymiarowania jest generowane przez system zarządzania.



Rysunek 4. Relacja pomiędzy systemem zarządzania a procesem wymiarowania

Moduł wymiarowania jest wywoływany przez moduł zarządcy zasobów systemu zarządzania. Po jego wywołaniu następuje pobranie informacji z bazy danych dotyczącej danej domeny o dostępnych zasobach (przepustowości łączy, mocy obliczeniowej, wielkości pamięci itd.) i ich dotychczasowej alokacji. Na podstawie danych wejściowych dotyczących przydziału zasobów (na Rys. 4 oznaczonych jako macierz danych wyjściowych) oraz informacji o zasobach (macierz zasobów) algorytmu alokacji zasobów oblicza nową alokację zasobów. Dane te są następnie wykorzystywane do generacji wymaganych wirtualnych zasobów. W razie potrzeby (np. w przypadku instalacji nowych zasobów lub awarii węzła) procedura wymiarowania może być zrealizowana ponownie na żądanie systemu zarządzania.

5 Podsumowanie

Architektura i funkcjonalność systemu zarządzania dla Systemu IIP mają znaczenie dla określenia elastyczności i efektywności działania systemu w warstwach L1/L2 oraz dla reakcji na zmianę zapotrzebowania czy też sytuacji awaryjnych. Zaproponowane rozwiązanie jest modularne. Elementem systemu zarządzania jest moduł wymiarowania, który rozdziela zasoby pomiędzy poszczególne RI. Obecnie, po etapie specyfikacji rozwiązań, są prowadzone prace implementacyjne mające na celu przygotowanie kompletnego środowiska zarządzania i jego sprawdzenie w eksperymentalnym systemie.

Literatura

1. Burakowski W. i inni, *Architektura Systemu IIP*, Krajowe Sympozjum Telekomunikacji i Teleinformatyki, Łódź, 14-16 wrzesień 2011
2. W. Burakowski et al., *The IIP System: Architecture, Parallel Internets, Virtualization and Applications*, Future Network & Mobile Summit 2011, 15 - 17 June 2011, Warsaw, Poland (presentation).
3. W. Burakowski, H. Tarasiuk, A. Beben, *Future Internet Architecture Based on Virtualization and Co-existence of Different Data and Control Planes*, Fifth Future Internet Cluster Workshop focused on Future Network Architectures, Warsaw, Poland, June 2011 (presentation).
4. W. Burakowski, H. Tarasiuk, A. Beben, *System IIP for supporting "Parallel Internets (Networks)"* , Future Internet Assembly Meeting, Ghent, 2010.
5. M. Bourguiba, K. Haddadou, G. Pujolle, *Evaluating and Enhancing Xen-based Virtual Routers to Support Real-time Applications*, Proceedings of IEEE CCNC 2010, Las Vegas, 9-12 January 2010.
6. A. Kumar, R. Rastogi, A. Silberschatz, B. Yener, *Algorithm for Provisioning Virtual Private Networks in the Hose Model*, IEEE/ACM TRANSACTIONS ON NETWORKING, Vol. 10, No. 4, August 2002.
7. J. Ding, *Advances in Network Management*, CRC Press, Boca Raton, 2010.
8. A. Galis, et al., *Management and Service-aware Networking Architectures (MANA) for Future Internet. System Functions, Capabilities and Requirements*, Position Paper, Version V6.0, 3rd May 2009.
9. ITU-T Rec. Y.3001, *Future Networks: Objectives and Design Goals*, (05/2001).

Architektura bezpieczeństwa Systemu IIP

Krzysztof Cabaj[1], Grzegorz Kołaczek[2], Jerzy Konorski[3], Zbigniew Kotulski[4], Łukasz Kucharzewski[4], Piotr Pacyna[5], Paweł Szałachowski[4]

[1] Instytut Informatyki, Politechnika Warszawska
[2] Instytut Informatyki, Politechnika Wrocławska
[3] Wydział Elektroniki, Telekomunikacji i Informatyki, Politechnika Gdańska
[4] Instytut Telekomunikacji, Politechnika Warszawska
[5] Katedra Telekomunikacji, Akademia Górniczo-Hutnicza

Streszczenie Artykuł przedstawia koncepcję architektury bezpieczeństwa na poziomie wirtualizacji zasobów Systemu IIP. Omawiane są trzy linie mechanizmów obronnych, w tym ochrona integralności informacji, wykrywanie anomalii i zasady pracy systemu budowania metryk zaufania węzłów wirtualnych.

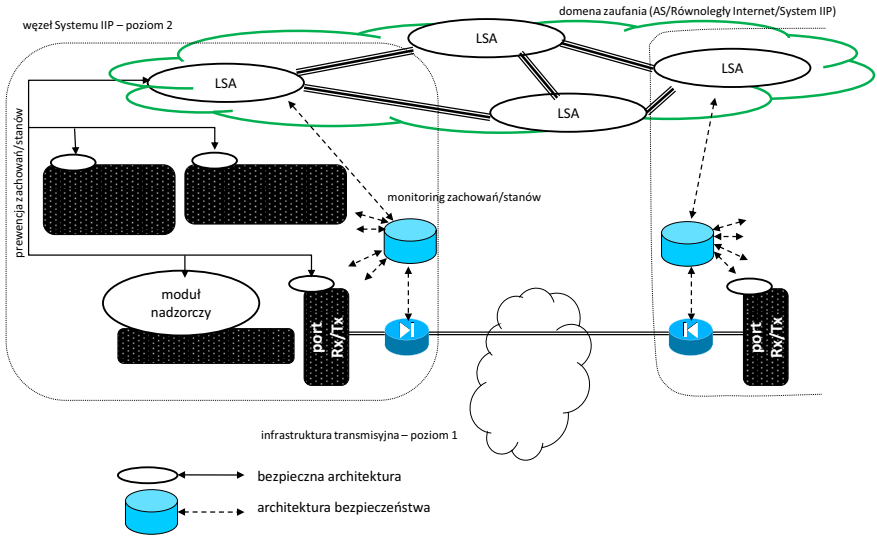
1 Wprowadzenie

Niniejszy rozdział opisuje architekturę bezpieczeństwa dla Systemu IIP - prototypowej instalacji powstającej obecnie w Polsce w ramach projektu Inżynieria Internetu Przyszłości (IIP) [1, 2], integrującej tzw. Równoległe Internety o różnych technikach transmisji. Podstawową metodą integracji jest wirtualizacja łączy, węzłów i serwerów w ramach 4-poziomowej architektury logicznej. Przedstawione tu propozycje architektury bezpieczeństwa ograniczone są do poziomu 2, odpowiedzialnego za tworzenie i funkcjonowanie zasobów wirtualnych.

Bezpieczeństwo sieciowe znajduje się w centrum uwagi wielu projektów europejskich dotyczących Internetu Przyszłości. Zauważyć w nich można zwłaszcza:

- podnoszoną konieczność uwzględnienia problematyki bezpieczeństwa sieciowego juz we wczesnych stadiach projektu, by przełamać paradygmat retrospektywnego przeciwdziałania zagrożeniom oraz
- powiązanie wybranych schematów bezpieczeństwa z mechanizmami budowy zaufania pomiędzy elementami sieci. Koncepcje architektury systemu zaufania rozwijane są m. in. w projektach X-ETP i 4WARD [3–8].

Wirtualizacja zasobów sieci, stanowiąca motyw przewodni Systemu IIP stwarza nowe wyzwania, do których należy zaliczyć większą podatność na nieuprawniony dostęp podmiotów spoza logicznej struktury sieci wirtualnej (związany ze współdzieleniem fizycznej infrastruktury transmisyjnej przez wielu użytkowników), trudności w zakresie ochrony programowych maszyn wirtualnych oraz trudności zabezpieczania komunikacji pomiędzy maszynami wirtualnymi w węźle fizycznym i pomiędzy węzłami.



Rysunek 1. Bezpieczna architektura i architektura bezpieczeństwa na poziomie 2

Obecnie brakuje spójnej wizji rozwiązań w zakresie bezpieczeństwa w środowisku wirtualizowanych zasobów sieciowych, chociaż próby sformułowania ogólnych zasad bezpieczeństwa dla takich środowisk znajdujemy w wielu źródłach. Na przykład w pracy [9] opisano możliwy sposób migracji współczesnej sieci w kierunku Internetu Przyszłości poprzez rozmieszczenie we wszystkich elementach sieci modułów softwarowych, zwanych *Cognitive Managers*. Każdy z nich zarządza abstrakcjami wirtualnych zasobów sieci oraz posiada wbudowany moduł pod nazwą *Supervisor and Security Module*, zapewniający ustalone atrybuty bezpieczeństwa (np. uwierzytelnianie, integralność i poufność przekazu danych, wykrywanie włamań). W ten sposób tworzona jest *bezpieczna architektura* Internetu Przyszłości; cechuje ją immanentny i proaktywny charakter mechanizmów bezpieczeństwa, co w zasadzie zapobiega wykroczeniom poza zakres normalnej pracy sieci. W odróżnieniu od tego przez *architekturę bezpieczeństwa* należy rozumieć zespół mechanizmów wzbogacających funkcjonalność już istniejącego bądź uprzednio zaprojektowanego systemu transmisyjno-komutacyjnego. Mechanizmy te są zwykle specyficzne względem przewidywanych zagrożeń i rozmieszczane jedynie w wybranych elementach systemu; w ogólności nie zapobiegają one wykroczeniom poza zakres normalnej pracy, a jedynie je sygnalizują i uruchamiają odpowiednie procedury reaktywne.

Rysunek 1 pokazuje koncepcję bezpiecznej architektury Systemu IIP na poziomie 2, gdzie każdy moduł węzła Systemu IIP (dla uproszczenia zaznaczono jedynie kilka modułów, w tym port nadawczo-odbiorczy) wyposażony jest w moduł nadzorczy bezpieczeństwa komunikujący się z lokalnym agentem bezpieczeństwa (LSA - *local security agent*). Moduły nadzorcze zapobiegają wykroczeniom odpowiednich modułów węzła poza dopuszczalne zakresy stanów lub zachowań specyfikowane przez LSA zgodnie z aktualną polityką bezpieczeństwa. Jednocześnie na podstawie danych zebranych z modułów nadzorczych LSA uzyskuje obraz stanu bezpieczeństwa własnego węzła i węzłów sąsiednich, który następnie współdzieli z LSA w innych węzłach w obrębie ustalonej domeny zaufania (np. systemu autonomicznego (AS), Równoległego Internetu, bądź Systemu IIP jako całości). Współpraca pomiędzy różnymi LSA w ramach domeny zaufania umożliwia w szczególności eliminację węzłów, w których - wskutek uszkodzenia bądź ataku intruza - LSA sam wykracza poza zakres normalnej pracy.

W proponowanym rozwiązaniu przyjęto podejście architektury bezpieczeństwa, gdyż koncepcja Systemu IIP koncentruje się na integracji technik transmisji, nie uwzględniając jawnie problematyki bezpieczeństwa. Na rysunku 1 zilustrowano fakt, że wprowadzenie kryptograficznego zabezpieczenia komunikacji na łączach wirtualnych pomiędzy sąsiednimi węzłami Systemu IIP oraz monitoring zachowań lub stanów poszczególnych modułów węzła pozwala nadal zasilać LSA obrazem stanu bezpieczeństwa węzła, jakkolwiek nie wyklucza wykraczania poza zakres normalnych zachowań lub stanów. Z tego powodu domena zaufania realizowana jest jako system reputacyjny i dostarcza jedynie rekomendacji dotyczących izolacji wybranych fragmentów Systemu IIP, co do których istnieje podejrzenie wykroczenia poza zakres normalnej pracy. Idee te zostaną bardziej szczegółowo omówione w kolejnych punktach rozdziału.

2 Architektura bezpieczeństwa Systemu IIP na poziomie 2

Proponowana architektura bezpieczeństwa bierze pod uwagę zagrożenia, które można klasyfikować jako przypadkowe bądź celowe, a także jako wewnętrzne (pochodzące od współużytkowników fizycznej infrastruktury transmisyjnej) bądź zewnętrzne (stwarzane np. przez przejętą przez intruza maszynę wirtualną implementującą wirtualny węzeł Systemu IIP). Incydenty bezpieczeństwa będą wspólnie określane jako *atak* dla podkreślenia ich jakościowo nieodróżnialnych konsekwencji na poziomie 2. Rozważane na tym poziomie ataki mają przynajmniej jedną z następujących cech:

- pochodzą ze źródła zewnętrznego w stosunku do Systemu IIP i są skierowane przeciwko poprawnemu przepływowi jednostek danych (*PDU - Protocol Data Unit*) w łączach wirtualnych, lub
- pochodzą z przejętej przez intruza maszyny implementującej wirtualny węzeł Systemu IIP.

Przykłady obejmują: fałszowanie PDU przez zewnętrznych intruzów (np. przechwytywanie i modyfikacje ruchu Systemu IIP, bądź wprowadzanie obcego ruchu); wprowadzanie niebezpiecznego ruchu przez przejęte węzły wirtualne Systemu IIP dążące do dezorganizacji pracy lub degradacji wydajności węzłów odbiorczych; zakłócenia relacji czasowych lub kolejnościowych w strumieniach PDU spowodowane przyczynami obiektywnymi (niewystarczające pasmo łącza wirtualnego, niedoskonała izolacja zasobów wirtualnych), bądź działaniami celowymi (np. atakami typu *jellyfish, replay* itp.), wreszcie zakłócenia izolacji względnie profili użytkowania zasobów wirtualnych spowodowane np. atakami typu *VM escape.*

2.1 Polityka bezpieczeństwa

Klasyczne podejścia oparte na znanych modelach zagrożeń i repozytoriach sygnatur ataku nie wydają się przydatne na poziomie 2. Systemu IIP. Mechanizmy znanych ataków są specyficzne względem protokołów poziomów 3. i 4. Systemu IIP - tzw. Równoległych Internetów oraz sieci wirtualnych - których znajomości ani dostępności odpowiednich informacji sterujących nie zakłada się na poziomie 2. Wskutek współdzielenia infrastruktury transmisyjnej z użyciem zróżnicowanych technik transmisyjnych ataki na poziomie 2. są trudniejsze do przewidzenia, zaś ich objawy w mniejszym stopniu charakterystyczne. Uzasadnia to stosowanie podejścia opartego na wykrywaniu anomalii. Zarazem ataki na poziomie 2. mają potencjalnie większą skalę oddziaływania: np. atak na węzeł wirtualny jednego z Równoległych Internetów może oddziaływać na pozostałe Równoległe Internety. W ramach proponowanego podejścia definiuje się typy obserwowalnych zdarzeń wskazujących na wystąpienie ataku (SRE - *security related events*). Szczególnie istotne są SRE związane z obserwacją strumieni PDU, a także z profilem wykorzystania zasobów węzłów wirtualnych. Celem polityki bezpieczeństwa jest:

- *prewencja* fałszowania PDU przez intruzów zewnętrznych w stosunku do Systemu IIP,
- *wykrywanie* wybranych ataków pochodzących z wewnątrz lub z zewnątrz Systemu IIP.

2.2 Trzy linie mechanizmów obronnych

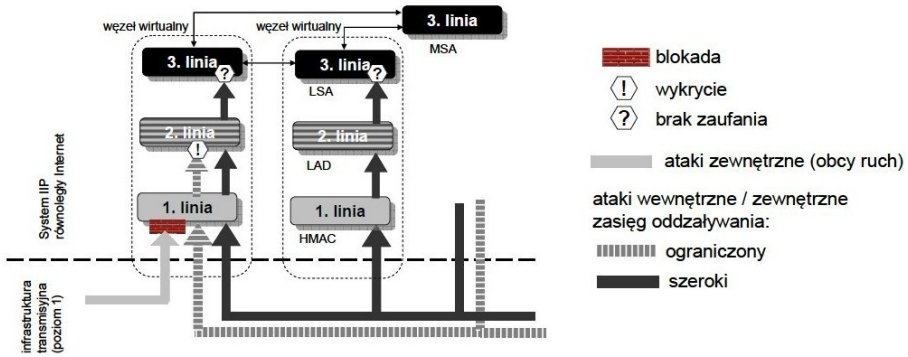
Rysunek 2 przedstawia linie obrony w węźle wirtualnym Równoległego Internetu (strzałki blokowe wskazują możliwe źródła i zasięg oddziaływania ataków).

Pierwsza linia mechanizmów obronnych zatrzymuje obcy ruch próbujący przeniknąć do Systemu IIP (neutralizując ataki przez fałszowanie PDU oraz typu *replay*). Realizuje ona ochronę integralności i uwierzytelnianie jednostek PDU wymienianych przez łącze wirtualne przy pomocy mechanizmu HMAC pracującego z nominalną przepływnością łącza, wyspecyfikowanego jako podzbiór standardu IEEE 802.1ae (MACSec). Moduły HMAC, przewidziane do implementacji

w obrębie narzędzi wirtualizacyjnych realizowanych na platformie netFPGA, eliminują PDU niespełniające testu HMAC oraz raportują zdarzenia sygnalizujące podejrzenie ataku.

Drugą linię stanowią mechanizmy *soft security* lokalne dla węzła wirtualnego. Ich zadaniem jest filtracja SRE wskazujących na ataki o zasięgu lokalnym, dla których pierwsza linia obrony nie jest wystarczającą przeszkodą (np. generujące ruch zewnętrzny o charakterze jedynie "nękającym" typu DoS, bądź ruch wewnętrzny z przejętej maszyny wirtualnej Systemu IIP). Rejestrowane są SRE obserwowane przez moduły pierwszej linii (nieudane testy HMAC, źle sformatowane nagłówki PDU, niepoprawne wiadomości w płaszczyźnie zarządzania itp.), a także odchylenia od profili użytkowania zasobów wirtualnych węzła. Lokalny moduł wykrywania anomalii (LAD - *local anomaly detection*), implementowany jako fragment kodu węzła wirtualnego, interpretuje rejestr SRE przekształcając go w rejestr lokalnych anomalii, ponadto za pośrednictwem mechanizmów trzeciej linii obrony kontaktuje się z innymi węzłami celem wykrycia anomalii o szerszym zasięgu.

Trzecia linia obrony również jest typu *soft security* i przeciwdziała atakom o szerokim zasięgu oddziaływania (np. typu *slow scanning*, nieuprawniony dostęp itp.), niewykrywalnym przez LAD; niweluje też skutki niepoprawnego działania LAD w przejętym węźle wirtualnym. Lokalny moduł kooperacji między węzłami (LSA - *local security agent*) zbiera i przetwarza informacje o zaobserwowanych SRE, przekształcając je w lokalne metryki reputacyjne i przekazując wraz z rejestrem SRE do centralnego węzła (MSA - *master security agent*). MSA wylicza globalne metryki reputacyjne i wykrywa anomalie o szerszym zasięgu, po czym rozgłasza je w obrębie Równoległego Internetu jako wiadomości Secure SNMP w formie odpowiednich alarmów oraz miar zaufania. W ten sam sposób przekazywane są również uaktualnienia konfiguracji lokalnych filtrów SRE.



Rysunek 2. Zarys architektury bezpieczeństwa na poziomie 2 Systemu IIP

3 Ochrona integralności na poziomie systemu transmisyjnego IIP

Oczekiwania w zakresie bezpieczeństwa, kierowane pod adresem systemu transmisyjnego rozwijanego przez projekt IIP, są związane z istotną grupą zastosowań sieci wirtualnych, w których do głównych wymagań zalicza się dużą odporność na zakłócenia pracy sieci. Mowa tutaj w pierwszej kolejności o zakłóceniach będących wynikiem świadomej ingerencji niepowołanych osób trzecich w pracę sieci, głównie poprzez aktywne ingerowanie w strumienie danych użytkowych przenoszonych przez system transmisyjny, lub w dane wymieniane przez systemy sterowania i zarządzania pracą sieci, jeżeli sieć zarządzająca korzysta z infrastruktury stanowiącej sieć zarządzaną.

Moduł HMAC-SHA-1 odpowiada za identyfikację i usuwanie ruchu obcego próbującego przeniknąć do Systemu IIP, neutralizując przez to ataki polegające na ingerowaniu w zawartość jednostek PDU oraz na wprowadzaniu do systemu transmisyjnego fałszywych jednostek PDU. Rozwiązanie HMAC-SHA-1 stanowi zatem pierwszą linię obrony Systemu IIP.

Rozwiązanie HMAC-SHA-1 zostało zrealizowane w postaci sprzętowego modułu szyfrującego, wykonanego w oparciu o platformę sprzętową netFPGA 1G, wyposażoną w układ programowalny Xilinx Virtex-II Pro. Zadaniem tego modułu jest przetwarzanie strumienia napływającego w postaci ramek systemu transmisyjnego System IIP, przenoszących dane użytkowe jak również dane sygnalizacyjne, sterujące pracą sieci. Celem przetwarzania jest zapewnienie ochrony integralności przesyłanych danych. Stosowany jest tu algorytm HMAC-SHA1 [10] pracujący z nominalną przepływnością łącza. Można uznać, że Moduł HMAC-SHA1, stanowi istotne narzędzie wirtualizacyjne należące do grupy narzędzi System IIP zrealizowanych na platformie netFPGA.

Moduł HMAC-SHA-1 realizuje ochronę ramek transmisyjnych Systemu IIP na poziomie drugim, według modelu warstwowego przyjętego w projekcie IIP, w ten sposób, że przepuszcza ramki transmisyjne spełniające test integralności oraz eliminuje ramki transmisyjne niespełniające tego testu. W tym drugim przypadku raportuje zdarzenia sygnalizujące niepomyślną weryfikację, co jest odbierane jako podejrzenie wystąpienia ataku i przekazywane, wraz z dodatkowymi informacjami do dalszej analizy. Ochrona ta może być włączona osobno na każdym łączu Systemu IIP.

Dzięki usytuowaniu tego elementu w warstwie drugiej system IIP zyskuje uniwersalny mechanizm, który może być stosowany do ochrony każdej sieci wirtualnej, niezależnie od typu Równoległego Internetu w obrębie którego taka sieć została wydzielona, a więc zarówno tam, gdzie wykorzystywane są protokoły internetowe, jak również i tam, gdzie stosowane są protokoły inne niż IP. Jest to rozwiązanie niezależne od warstw wyższych systemu transmisyjnego IIP oraz transparentne dla nich.

Wdrożenie rozwiązania w prototypowym systemie polega na wprowadzeniu do węzłów systemu transmisyjnego IIP modułu przetwarzającego HMAC-SHA-1, który wykonuje sprzętowe przetwarzanie ramek IIP w układzie FPGA pracującym z nominalną prędkością 1 Gbit/s (docelowo 10 Gbit/s).

4 Wykrywanie anomalii

Zastosowanie metod wykrywania anomalii daje możliwość identyfikacji wcześniej nierozpoznanych metod ataku lub problemów bezpieczeństwa, dla których jeszcze nie została zdefiniowana odpowiednia sygnatura. Po zdefiniowaniu bądź wyznaczeniu charakterystyki normalnego zachowania systemu przeprowadza się identyfikację zdarzeń znacząco od niej odbiegających. Jedną z metod postępowania jest generowanie alarmów po wykryciu statystycznie rzadkich stanów systemu.

Moduł LAD wykrywa anomalie na poziomie węzła Systemu IIP. Możliwe jest też identyfikowanie anomalii osobno dla każdego z Równoległych Internetów tworzących System IIP, rozumianych jako zbiory wirtualnych węzłów. Odpowiada za to moduł PI-AD w MSA analizujący zmiany wartości wybranych atrybutów stanu danego Równoległego Internetu.

Proponowane podejście do projektowania modułów LAD wykorzystuje kompleksowe informacje o stanie Systemu IIP i opiera się na metodach eksploracji danych oraz na statystycznej analizie szeregów czasowych.

4.1 Wykrywanie anomalii z wykorzystaniem metod eksploracji danych

Analizie podlega tutaj rejestr SRE – na przykład zapisy nieudanego logowania, odrzucenia PDU przez moduł HMAC, prób odczytania poprzez SNMP klucza z bazy MIB bez stosownych uprawnień, czy też odrzucenie niezgodnego z polityką bezpieczeństwa połączenia w płaszczyźnie zarządzania. Z każdym SRE, oprócz jego typu i czasu wystąpienia, związane są pewne atrybuty, np. nazwa konta, źródłowy i docelowy adres IP w odrzuconym PDU, czy OID klucza w bazie MIB. SRE generowane przez programy działające pod kontrolą systemu Linux w węźle Systemu IIP są logowane za pomocą standardowego podsystemu syslog, skąd moduł LAD pobiera dane do analizy. Takie rozwiązanie pozwala w przyszłości łatwo rozszerzyć zbiór zdefiniowanych SRE.

Proponowana metoda opiera się na wykrywaniu tzw. zbiorów częstych (*frequent sets*), por. [11], tj. powtarzających się często wzorców zachowań. Jeśli dane do analizy traktujemy jako zbiory odpowiednich atrybutów, to zbiorem częstym nazywamy podzbiór występujący co najmniej *minSup*-krotnie w analizowanych danych. Parametr *minSup*, nazywany minimalnym wsparciem, jest parametrem wejściowym algorytmu wyszukiwania zbiorów częstych. Zbiory częste można wykrywać za pomocą różnych algorytmów. W ramach projektu planuje się implementację algorytmu a priori. W tabeli 1 zaprezentowano kolejne kroki tego algorytmu.

Analizując tą metodą SRE dotyczące np. odrzuconych prób dostępu do kluczy MIB można wykryć wzorzec związany z próbą poznania przez danego użytkownika wartości pewnych elementów MIB. Jeśli wykryty zbiór częsty zawiera informację o adresie IP, oznacza to atak przeprowadzany z jednej maszyny. Brak w wykrytym zbiorze częstym identyfikatora OID świadczy o próbie dotyczy poznania różnych elementów MIB. Gdy wykryty zbiór częsty zawiera atrybut

Tabela 1. Algorytm a priori, służący do wykrywania zbiorów częstych

1. START
2. Przygotowanie C1 – listy jednoelementowych kandydatów
   Przejrzenie wszystkich zbiorów podlegających analizie w danym okresie czasu i dodanie do C1 jako jednoelementowych zbiorów wszystkich pojawiających się elementów.
3. Przygotowanie F1 – listy jednoelementowych biorów częstych
   Wyliczenie wsparcia wszystkich kandydatów z listy C1, dodanie do listy F1 tych których wsparcie jest większe lub równe minSup.
4. Przygotowanie C2 – listy dwuelementowych kandydatów
   Połączenie wszystkich zbiórów jednoelementowych z F1 w dwuelementowe zbiory w ten sposób zawsze identyfikator pierwszego elementy był mniejszy niż drugiego.
5. Przygotowanie F2 – listy dwuelementowych zbiorów częstych
   Wyliczenie wsparca wszystkich kandydatów z C2 jako liczby tych zbiorów w analizowanym okresie czasu, które zawierają oba elementy. Dodanie do F2 tylko tych zbiorów, których wsparcie jest większe niż założony parametr minSup.
6. LOOP
   (a) Przygotowanie CN – listy kandydatów o N elementach,
   (b) Łączenie tylko tych których N-1 pierwszych elementów jest identycznych,
   (c) Czyszczenie CN,
   (d) Usunięcie tych zbiorów z CN, które mają choćby jeden N-1 elementowy podzbiór nie występujący w F(N-1),
   (e) Wyliczenie wsparcia każdego zbioru z CN,
   (f) Przygotowanie FN – listy zbiorów częstych o N elementach,
   (g) Dodanie do FN tylko tych zbiorów z CN, których wsparcie jest większe lub równe parametr minSup.
7. REPEAT UNTIL lista *FN list* zawiera jakieś nowe zbiory częste.
8. KONIEC

związany z OID przy braku atrybutu związanego z adresem IP, atak był przeprowadzany z wielu maszyn i dotyczył określonego elementu bazy MIB.

W wyniku analizy uzyskanych zbiorów częstych, do modułu obliczania reputacji (*reputation calculator*) zostanie przekazana ocena ryzyka (*severity*) związanego z daną anomalią oraz prawdopodobieństwo trafności oceny dokonanej przez moduł LAD. Informacje te (na rysunku 5 oznaczone odpowiednio jako c^l i p_n^l, gdzie l jest symbolem anomalii a n numerem okresu obserwacji) wykorzystywane są przez system zarządzania reputacją omawiany w punkcie 5 tego rozdziału.

4.2 Wykrywanie anomalii z wykorzystaniem analizy szeregów czasowych

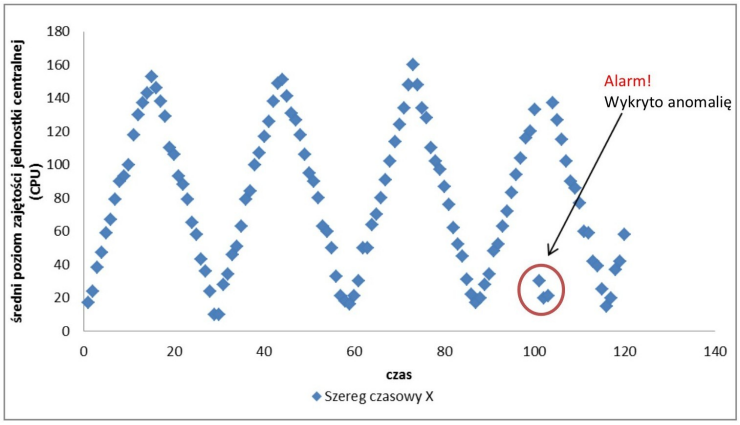
Celem modułu LAD implementującego wykrywanie anomalii z wykorzystaniem analizy szeregów czasowych jest wykrycie nietypowych zjawisk w zachowaniu pojedynczego węzła Systemu IIP. W tym celu badane są cechy charakteryzujące węzeł Systemu IIP oraz cechy charakteryzujące wirtualne węzły poszczególnych

Równoległych Internetów (PI). Agent dokonuje analizy następujących cech charakteryzujących stan węzła Systemu IIP:

1. intensywność nadchodzenia ramek dla danego typu Równoległego Internetu,
2. średni rozmiar sekcji danych w ramce dla danego typu Równoległego Internetu (PI payload),
3. średni stopień zajętości jednostki centralnej (CPU),
4. średni stopień wykorzystania pamięci operacyjnej,

oraz następujących cech charakteryzujących poszczególne wirtualne węzły Równoległych Internetów (*PI node*):

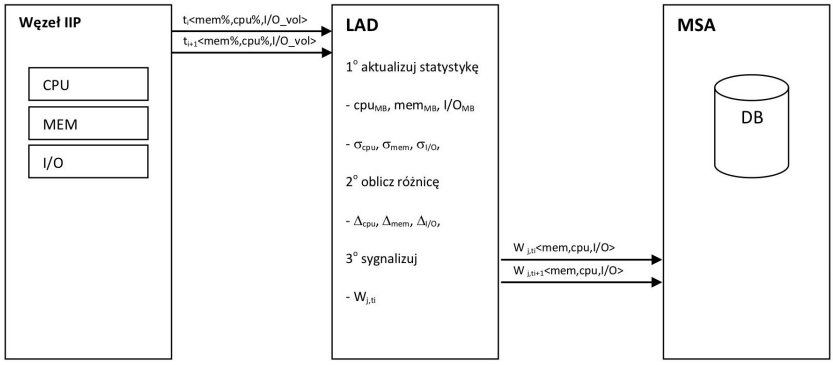
1. średnia liczba bajtów wysyłanych przez węzeł w jednostce czasu,
2. średnia liczba bajtów odbieranych przez węzeł w jednostce czasu,
3. średni stopień zajętości jednostki centralnej (CPU) w jednostce czasu,
4. średni stopień wykorzystania pamięci operacyjnej w jednostce czasu.



Rysunek 3. Anomalie w szeregach czasowych

W pierwszej kolejności planuje się implementację LAD dla węzła Systemu IIP typu Xen. W planowanej implementacji LAD dla środowiska wirtualizatora Xen bieżące wartości wybranych cech wirtualnych węzłów Równoległych Internetów są odczytywane z lokalnej bazy MIB oraz z narzędzi systemowych *vir-top* oraz *xentop*. Rozbieżności wartości aktualnych ze stanami historycznymi analizowane natomiast są przy pomocy wybranych i przetestowanych algorytmów analizy szeregów czasowych (metoda Burgessa [12]).

Komunikat $t_i < mem\%, cpu\%, I/O\_vol >$ na Rysunku 4 zawiera aktualne wartości obserwowanych zmiennych charakteryzujących stan węzła Systemu



Rysunek 4. Przetwarzanie informacji o stanie węzła IIP

IIP, które są zbierane i przetwarzane przez LAD w chwilach $t_i, t_{i+1}, ...$, o stałym i ustalonym odstępie. LAD aktualizuje na podstawie odczytanych wartości stanu węzła Systemu IIP swoje wewnętrzne struktury danych, a następnie oblicza zgodnie z metodą Burgessa wartości statystyk – wartości średnie $(cpu_{MB}, mem_{MB}, I/O_{MB})$ oraz odchylenia standardowe $(\sigma_{cpu}, \sigma_{mem}, \sigma_{I/O})$. Na tej podstawie obliczana jest wartość $\Delta_{cpu}, \Delta_{mem}, \Delta_{I/O}$ określającą poziom zgodności aktualnej obserwacji $t_i < mem\%, cpu\%, I/O\_vol >$ z historycznym zachowaniem się węzła Systemu IIP. W rezultacie LAD przesyła wartość $W_{j,t_i} < mem, cpu, I/O >$ (por. rysunek 4) z przedziału $< 0, 1 >$ będącą informacją poziomie zaobserwowanej anomalii w zachowaniu węzła j Sytemu IIP w chwili t_i do agenta MSA, gdzie 0 oznacza zachowanie zgodne z historycznymi obserwacjami węzła (brak anomalii), a 1 maksymalną różnicę (maksymalny poziom anomalii). LAD przesyła również analogiczny komunikat z wykorzystaniem protokołu SNMP do systemu zarządzania reputacją, co umożliwia aktualizację wartości reputacji węzła Systemu IIP również z uwzględnieniem informacji statystycznych.

Obserwacja stanów węzła Systemu IIP i wykrywanie anomalii z wykorzystaniem analizy szeregów czasowych ma na celu wykrywanie ataków na mechanizmy umożliwiające wirtualizację zasobów w Systemie IIP. Potencjalnymi atakami tej kategorii, które mogą być wykryte przez LAD są ataki typu przejęcie systemu gościa, atak na dostępność usług realizowane poprzez systemy goszczące lub bezpośrednio przeciw systemowi umożliwiającemu wirtualizację.

5 System zarządzania reputacją

Mechanizmy zarządzania, routingu i bezpieczeństwa w Systemie IIP wymagają stałej oceny poprawności pracy węzłów sieci i połączeń między węzłami. Różnorodność czynników wpływających na ich działanie oraz wiele możliwości precyzyjnego opisu wpływu każdego z tych czynników sprawia, że zdecydowano się na przyjęcie systemu reputacyjnego jako jednolitego narzędzia oceny poprawno-

ści pracy węzłów i połączeń w Systemie IIP. Rozwiązanie takie jest charakterystyczne dla systemów bez scentralizowanej administracji - sieci bezprzewodowych *mesh*, systemów obliczeń w chmurach (*cloud computing*), czy sieci społecznościowych [13, 14].

Model reputacji wykorzystany w Systemie IIP powinien zapewniać:

- elastyczność ze względu na zakres dostępnych informacji o węzłach i ruchu sieciowym, z preferencją dla danych obiektywnych z bieżącej obserwacji działania Systemu IIP,
- wrażliwość na pojawiające się zagrożenia przy zachowaniu możliwości poprawy reputacji w przypadku ustąpienia zagrożeń,
- możliwość modyfikacji obliczania reputacji (w tym zmiany modelu matematycznego) przy zachowaniu systemu przesyłania, przechowywania i udostępniania niezbędnych informacji,
- łatwość pozyskiwania danych do obliczania reputacji od wszystkich modułów analitycznych i zarządzających Systemu IIP,
- dostępność informacji o obliczonych wartościach reputacji wszystkich węzłów i łączy z innych podsystemów Systemu IIP (zarządzanie, routing, inne systemy bezpieczeństwa).

Jako model matematyczny obliczania reputacji przyjęto model heurystyczny inspirowany metodami bayesowskimi [15], por. rysunek 5. Jego podstawowym założeniem jest dwustopniowa ocena węzłów w ustalonych chwilach $n = 1, 2, ...$ z użyciem pojęć zaufania (*trust*) i reputacji (*reputation*). Zaufanie $T_n^{i,j}$ jest indywidualną oceną danego węzła j przez jego sąsiada i w topologii poziomu 2 Systemu IIP na podstawie danych o incydentach I_n (n oznacza numer okresu obserwacji) dostarczonych przez inne moduły (HMAC, LAD).

Ocena ryzyka I_n^l incydentu o identyfikatorze l obliczana jest jako $I_n^l = c_l \times p_n^l$, gdzie $c_l \in [0, 1]$ oznacza koszt incydentu (severity) a p_n^l prawdopodobieństwo (probability), że ten incydent rzeczywiście wystąpił. Przyjęto, że węzeł gromadzi informację o każdym z sąsiadów z okresie n i oblicza zaufanie do każdego z nich według nastepujących wzorów:

$$\begin{aligned} \check{I}_n &= \max_l \big(c_l \times p_n^l\big), \\ k &= \#\big(I_n^l : {I^l}_n > \frac{1}{2}\check{I}_n\big), \\ I_n &= \check{I}_n + \big(1 - \check{I}_n\big) \times \big(1 - a^{-(k-1)}\big), \\ T_n &= 1 - I_n, \end{aligned} \tag{1}$$

gdzie a jest stałą wyznaczaną emiprycznie. Przyjęto taki model zaufania, aby rownoczesnie uwzględnic decydujący wpływ na zaufanie incydentu o największym ryzyku oraz, jeżeli występują, licznych incydentów o ryzyku znaczącym. W kolejnym kroku węzeł o numerze i przesyła informację o zaufaniu do węzla j odpowiednio ją oznaczając, $T_n^{i,j} = T_n$. Na podstawie informacji zebranych od wszystkich sąsiadów agent centralny oblicza skumulowane zaufanie do węzła j w ocenianym okresie obserwacji n według wzoru:

$$T_n^j = \frac{\sum_{i \in A_j} T_n^{i,j} R_n^i}{\sum_{i \in A_j} R_n^i}, \tag{2}$$

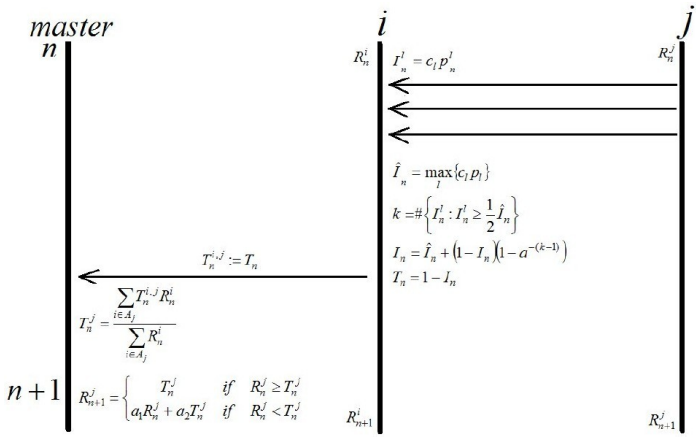
gdzie

- A_j jest zbiorem węzłów sąsiadujących z węzłem j i oceniających go,
- R_n^j jest reputacją węzła j w analizowanym okresie[6] n.

Reputacja R_{n+1}^j węzła j w kolejnej chwili $n+1$ jest obliczana na podstawie skumulowanej informacji o zaufaniu T_n^j do tego węzła, obliczonej centralnie na podstawie danych o zaufaniu $T_n^{i,j}$ dostarczonych przez jego sąsiadów, z uwzględnieniem reputacji tych sąsiadów R_n^i i wartości reputacji ocenianego węzła R_n^j. Przyjęto, że reputacja w kolejnej chwili ma wartość:

$$R_{n+1}^j = \begin{cases} T_n^j, & \text{jeżeli } R_n^j \geqslant T_n^j, \\ a_1 R_n^j + a_2 T_n^j, & \text{jeżeli } R_n^j < T_n^j, \end{cases} \tag{3}$$

gdzie stałe $0 < a_1, a_2 < 1$, $a_1 + a_2 = 1$ są dobrane eksperymantalnie. Taka postac wyrazenia dla reputacji sprawia, że reputacja zaatakowanego węzła zmniejsza się szybko, a w czasie jego poprawnej pracy odbudowywana jest powoli.



Rysunek 5. Przykład sposobu obliczania reputacji węzłów

Przedstawione powyżej podejście do oceny sieci komunikujących się wezłów za pomocą systemu reputacyjnego zakłada ocenę danego węzła przez jego sąsiadów na podstawie obserwacji ruchu nadchodzącego do nich od ocenianego węzła. Ocena taka nie jest wystarczająca, jeżeli chcemy uwzględniać w ocenie również

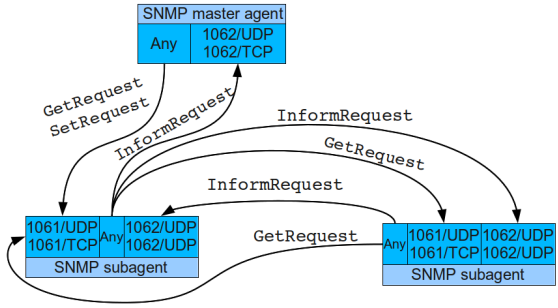
[6] W chwili rozpoczynania obliczeń przyjmujemy reputację węzłów jako 1.

wewnętrzne parametry węzła, takie jak obciążenie procesora, zajętość pamięci operacyjnej czy wykorzystanie pamięci masowej. W takiej sytuacji lokalny agent bezpieczeństwa powinien oceniać działanie swojego węzła obliczając w każdym cyklu obserwacji własną ocenę zaufania do węzła T_n^{*j} w sposób analogiczny do oceny węzła sąsiedniego, por. (1). Skumulowane zaufanie do węzła j ma w takim rozbudowanym modelu zaufania następującą postać:

$$T_n^j = \frac{\sum_{i \in A_j} T_n^{i,j} \times R_n^i + T_n^{*j} \times R_n^j}{\sum_{i \in A_j} R_n^i + R_n^j}. \tag{4}$$

Przesyłanie komunikatów o zaufaniu i sposób obliczenia reputacji są takie jak w poprzednim przypadku, podanym we wzorze (3).

Jako protokół komunikacyjny dla zarządzania reputacją wybrano SNMPv3, który pracuje jako scentralizowany system agentowy w paradygmacie klient-serwer, por. rysunek 6, ponadto ma możliwość bezpiecznej komunikacji między agentami. Protokół zapewnie poufność, integralność oraz świeżość danych [16] jak również zaawansowaną kontrolę dostępu dla użytkowników [17]. W podstawowej wersji systemu reputacyjnego wykorzystywana będzie jedynie komunikacja między agentami lokalnymi w węzłach Systemu IIP (agent SNMP, LSA) a centralnym agentem obliczającym reputacje węzłów (master agent SNMP, MSA). Agent zarządzający reputacją współdzieli lokalną bazę danych z agen-



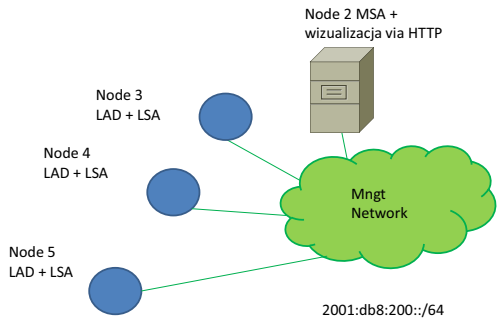
Rysunek 6. Schemat komunikacji agentów protokołu SNMP dla zarządzania reputacja

tem SNMP, stanowiąc odrębny moduł programowy. Umożliwia to niezależną aktualizację obu systemów agentowych. Protokół SNMP wykorzystuje bazę danych MIB [18], w której będą zapisywane informacje o incydentach wykrytych przez lokalne moduły analityczne węzła. LSA pobiera z bazy dane o incydentach i na ich podstawie oblicza poziom zaufania do swoich sąsiadów. Następnie zapisuje wynik do bazy, a agent SNMP przesyła rezultat do MSA poprzez master agenta SNMP. W bazie tej znajdzie się również, uzyskana od MSA informacja o aktualnej reputacji węzła. Zarówno ta informacja, jak i zapisane w bazie MSA informacje o reputacji wszystkich węzłów sieci będą udostępniane mechanizmom

zarządzania Systemu IIP poprzez dostęp do odpowiednich MIB master agenta SNMP i lokalnych agentów SNMP.

6 Eksperymenty

W ramach pierwszego etapu implementacji powstały prototypy modułu LAD, LSA i MSA. Główny nacisk został położony na implementację interfejsów pomiędzy modułami, z uproszczonymi algorytmami wykrywania anomalii (moduł LAD) czy możliwościami dynamicznej konfiguracji liczby i adresów węzłów (moduły LSA, MSA). Uproszczone na tym etapie implementacji elementy będą rozwijane w ramach dalszych prac nad projektem. Powstałe oprogramowanie zostało zintegrowane w sieci testowej grupy bezpieczeństwa na Politechnice Warszawskiej. Na potrzeby eksperymentów zostały uruchomione w środowisku Xen cztery maszyny wirtualne, reprezentujące cztery węzły systemy IIP. Wszystkie elementy komunikowały się jedynie za pomocą protokołu IPv6, wykorzystując adresy z prefixem 2001:db8:200::/64. Na rysunku 7 znajduje się logiczna topologia sieci testowej. Na węźle o nazwie Node 2 uruchomiony został moduł MSA oraz moduł wizualizacji poziomu zaufania. Na pozostałych trzech węzłach o nazwach Node 3, Node 4 i Node 5 uruchomione zostały moduły LAD i LSA.



Rysunek 7. Schemat topologii sieci testowej w której były przeprowadzano eksperymenty

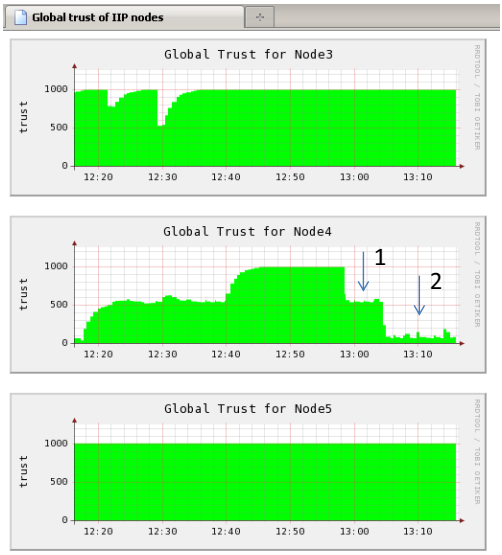
Na potrzeby testów komunikacji między modułami i wyliczania poziomu zaufania analizowane były jedynie zdarzenia SRE związane z naruszeniem polityki bezpieczeństwa w sieci zarządzania. Zdarzenia te były generowane i wysyłane

do podsystemy syslog bezpośrednio przez zaporę ogniową systemu Linux skonfigurowaną za pomocą programu *ip6tables*. Eksperymenty bazowały na danych uzyskiwanych w czasie rzeczywistym. Do symulacji ataków były wykorzystane standardowe narzędzia administracyjne ping6 oraz nmap, które mogą posłużyć do przeprowadzenia wstępnej fazy rozpoznania rzeczywistych ataków. Wynikiem eksperymentów był obliczony przez MSA globalny poziom zaufania do poszczególnych węzłów – liczba z zakresu $0 \ldots 1000$ (jest to przeskalowana reputacja, oryginalnie przyjmująca wartości z zakresu $0 \ldots 1$). Podczas eksperymentów wartość kluczowych parametrów wzorów (1) i (3) przyjmowały wartości $a = 4$, $a_1 = 0,7$ oraz $a_2 = 0,3$. Przeprowadzane w czasie rzeczywistym ataki powodowały zgłaszanie do systemu reputacji anomalii o parametrach *severity* $c_l = 0,15$, *probability* $p_n^l = 0,7$ dla *ping6* oraz *severity* $c_l = 1$ i *probability* $p_n^l = 1$ dla *nmap*.

Wyliczony przez MSA poziom zaufania do poszczególnych węzłów był wizualizowany w czasie rzeczywistym. Poniżej znajdują się przykładowe wykresy poziomu globalnego zaufania dla węzłów uzyskane podczas eksperymentów wraz z komentarzem opisującym przebieg eksperymentu. Na rysunku 6 przedstawiony jest wpływ liczby węzłów raportujących anomalie dotyczącą wybranego węzła, na uzyskany globalny poziom reputacji. Podczas tego eksperymentu węzeł Node 4 rozpoczął skanowanie węzła Node 3. Wywołało to spadek jego globalnego zaufania do poziomu bliskiego 500. Sytuacja ta jest widoczna na drugim wykresie i oznaczona strzałką z numerem 1. W momencie kiedy z węzła Node 4 rozpoczęto skanowanie kolejne węzła (Node 5) jego reputacja spadła do poziomu bliskiego 0, co zaznaczone jest strzałką z numerem 2. Eksperyment ten potwierdza wpływ liczby raportujących węzłów na globalny poziom zaufania. W realnym systemie odzwierciedla to sytuację kiedy jeden z węzłów rozpoczyna ataki na wiele innych węzłów. Wykrycie tego faktu, poprzez raporty od wielu maszyn, znacznie obniża reputację atakującego. Drugi eksperyment miał na celu potwierdzenie wpływu globalnego zaufania węzła raportującego anomalię do zmiany poziomu zaufania podejrzanej maszyny. Podczas tego eksperymentu z węzła Node 3 wykonywano dwa skanowania z wykorzystaniem programu nmap: systemu który ma zaufanie 1000 oraz systemu o zaufaniu bliskim 0. Na rysunku 6 strzałka z numerem 1 pokazuje sytuację kiedy anomalie dotyczące węzła atakującego są zgłaszane przez węzeł o minimalnym zaufaniu (Node 4). W takim przypadku globalne zaufanie do atakującego zostaje zmniejszone minimalnie. Zupełnie inne efekt obserwujemy, gdy raport na temat anomalii został wygenerowany przez zaufaną stację (Node 5). W takim przypadku spadek zaufania potencjalnego atakującego jest bardzo istotny. Taką sytuację wskazuje strzałka z numerem 2.

7 Podsumowanie

Architektura systemu bezpieczeństwa jest częścią specyfikacji poziomu 2 Systemu IIP, który udostępnia mechanizmy i techniki wirtualizacji dla tworzenia Równoległych Internetów. Dzięki takiemu usytuowaniu mechanizmy bezpieczeństwa będą wprowadzone do Systemu IIP w sposób jednolity i będą jednakowo dostępne dla Równoległych Internetów. Oparcie systemu bezpieczeństwa na wy-

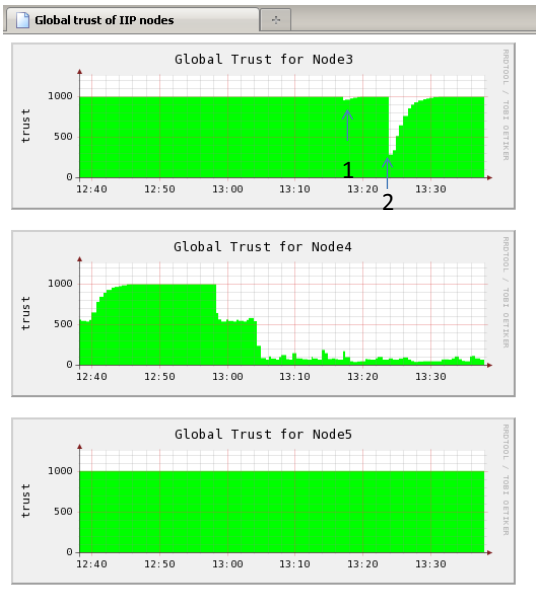


Rysunek 8. Wpływ liczby węzłów raportujących anomalie na poziom globalnego zaufania atakującego

krywaniu anomalii eliminuje wykorzystanie znanych modeli zagrożeń i repozytoriów sygnatur ataku. Te są bowiem nieprzydatne na poziomie 2, gdyż mechanizmy ataków są specyficzne względem wyższych poziomów Systemu IIP, zaś ich objawy mniej charakterystyczne wskutek współdzielenia infrastruktury transmisyjnej przez Równoległe Internety. Prowadzone prace projektowe zmierzają do praktycznej weryfikacji proponowanego rozwiązania w ogólnokrajowym środowisku eksperymentalnym, aktualnie budowanym w oparciu o specyfikację Systemu IIP.

Literatura

1. W. Burakowski, H. Tarasiuk, A. Beben, System IIP for supporting "Parallel Internets (Networks)", Future Internet Assembly meeting, Ghent 2010, fi-ghent.fi-week.eu/files/2010/12/1535-4-System-IIP-FIA-Ghent-ver1.pdf.
2. `https://www.iip.net.pl/`
3. A. Gavras et al., "Future Internet Research and Experimentation: The FIRE Initiative", ACM SIGCOMM Computer Communication Review, 37, 3, July 2007.



Rysunek 9. Wpływ reputacji węzła raportującego anomalie na poziom globalnego zaufania atakującego

4. "Future Internet - Strategic Research Agenda", ver. 1.1, FI X-ETP Group, January 2010.
5. M. Soellner, "The 4WARD Approach to Future Internet", ITG FG 5.2.3 Meeting Eschborn, 2010.
6. `http://www.syssec-project.eu/`
7. `http://www.nessos-project.eu/`
8. `http://www.effectsplus.eu/`
9. M. Castrucci, F. Delli Priscoli, A. Pietrabissa, and V. Suraci, "A Cognitive Future Internet Architecture", J. Domingue et al. (Eds.): Future Internet Assembly, LNCS 6656, pp. 91-102, 2011.
10. RFC 2104, "HMAC: Keyed-Hashing for Message Authentication", IETF 02.1997.
11. R. Agrawal, R. Srikant, "Fast algorithm for mining association rules", In: J.B. Bocca, M. Jarke, and C.Zaniolo, editors. Proceedings of 20th International Conference on Very Large Databases, pp. 487-499, (1994).
12. M. Burgess, "Two dimensional time-series for anomaly detection and regulation in adaptive systems", in: IFIP/IEEE 13th Int. Workshop on Distributed Systems: Operations and Management, DSOM 2002, pp. 169-185

13. A. Srinivasan, J. Teitelbaum, Jie Wu, M. Cardei, H. Liang, "Reputation-and-Trust-Based Systems for Ad Hoc Networks", pp. 375-403, in A. Boukerche [ed.], Algorithms and Protocols for Wireless and Mobile Ad Hoc Networks, J. Wiley & Sons, Inc. 2009.
14. K. Hoffman, D. Zage, and C. Nita-Rotaru, "A survey of attack and defense techniques for reputation systems", ACM Computing Surveys, vol. 42, no. 1, pp. 1-31, 2010.
15. T. Ciszkowski, W. Mazurczyk, Z. Kotulski, T. Hossfeld, M. Fiedler, D. Collange, "Towards Quality of Experience-based reputation models for future web service provisioning", Telecommunication Systems, DOI: 10.1007/s11235-011-9435-2.
16. RFC 5591, "Transport Security Model for the Simple Network Management Protocol (SNMP)", IETF 06.2009.
17. RFC 3415, "View-based Access Control Model (VACM) for the Simple Network Management Protocol (SNMP)", IETF 12.2002.
18. RFC 3418, "Management Information Base (MIB) for the Simple Network Management Protocol (SNMP)", IETF 12.2002.

Architektura i mechanizmy Równoległego Internetu IPv6 QoS

Halina Tarasiuk[1], Witold Góralski[1],
Janusz Granat[2], Jordi Mongay Batalla[2], Wojciech Szymak[2],
Paweł Świątek[3],
Sławomir Hanczewski[4],
Robert Szuman[5], Michał Giertych[5],
Krzysztof Gierłowski[6],
Marek Natkaniec[7], Janusz Gozdecki[7]

[1] Instytut Telekomunikacji, Politechnika Warszawska
[2] Instytut Łączności - Państwowy Instytut Badawczy
[3] Instytut Informatyki, Politechnika Wrocławska
[4] Politechnika Poznańska
[5] Poznańskie Centrum Superkomputerowo-Sieciowe
[6] Politechnika Gdańska
[7] AGH Akademia Górniczo-Hutnicza

Streszczenie Artykuł przedstawia propozycję architektury i mechanizmy Równoległego Internetu IPv6 QoS, który jest rozważany jako jeden z trzech Równoległych Internetów w Systemie IIP tworzonym w ramach projektu Inżynieria Internetu Przyszłości (IIP). Artykuł zawiera funkcje i mechanizmy, które umożliwią działanie sieci wirtualnych dla wybranych typów aplikacji, takich jak e-zdrowie, monitorowanie i bezpieczeństwo publiczne, zdalne nauczanie, w tym wideo-konferencję HD.

1 Wprowadzenie

Równoległy Internet IPv6 jest jednym z trzech Równoległych Internetów (RI), który jest rozważany dla zastosowania w Systemie IIP opracowywanym w ramach projektu Inżynieria Internetu Przyszłości (IIP) [2][3]. System IIP tworzy infrastrukturę wirtualnych węzłów i łączy, na której będzie możliwe uruchomienie Równoległych Internetów. Główne założenia dla Systemu IIP są następujące. System powinien dostarczyć infrastrukturę wirtualną, która umożliwi poszczególnym RI działanie w oparciu o różne stosy protokołów zarówno dla przekazu danych jak i dla sterowania. W szczególności rozważane są: RI oparty na technice IPv6 oraz RI oparte na zupełnie nowych podejściach określanych mianem *post-IP*, tj. Sieci Świadome Przekazywanej Treści oraz Sieci z Komutacją Strumieni Danych. Dla tak zdefiniowanych założeń dla Systemu IIP w projekcie zaproponowano architekturę IIP, która obejmuje 4 główne warstwy: fizyczną, wirtualizacji, RI, sieci wirtualnych.

Równoległy Internet IPv6 QoS (RI IPv6 QoS) jest definiowany w warstwie 3 architektury IIP. RI IPv6 QoS jest Internetem opartym na protokole IPv6. Gwarancje jakości obsługi QoS (*Quality of Service*) w sieci dla wybranych strumieni

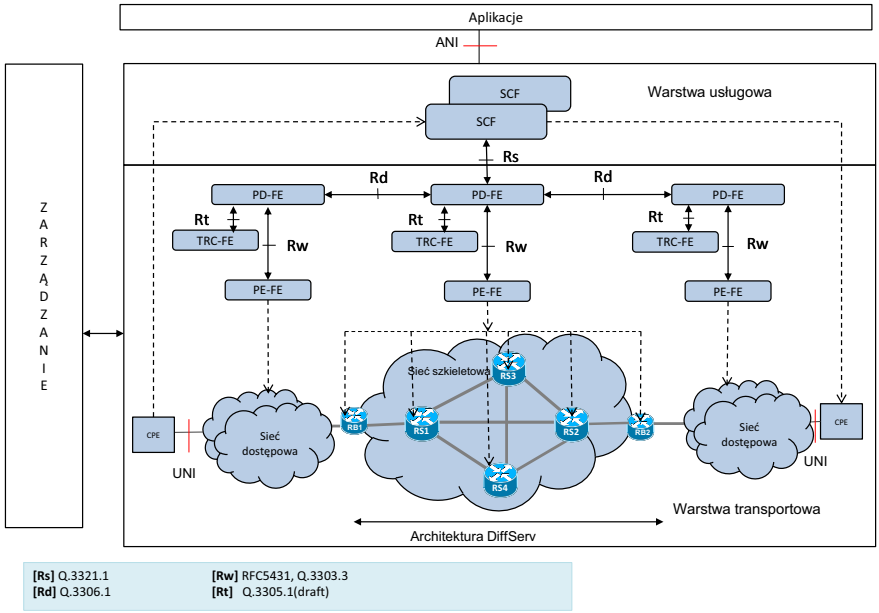
pakietów są zapewniane w ramach zdefiniowanych klas usług. W szczególności klasy usług, dzięki zastosowaniu mechanizmów PHB (*Per Hop Behaviour*) architektury DiffServ [8] oraz funkcji warstw usługowej i transportowej architektury NGN (*Next Generation Networks*) [5][6] wspierają ścisłe gwarancje QoS dla wybranych strumieni pakietów w ramach grupy aplikacji IPv6 opracowywanych w projekcie IIP. Do wspieranych aplikacji należą głównie aplikacje z grupy tzw. aplikacji Internetu Rzeczy (*Internet of Things - IoT*), które używają czujników np. dla monitorowania stanu zdrowia pacjentów lub treningu sportowców (aplikacje e-zdrowie). Druga grupa aplikacji IoT wspieranych przez RI IPv6 QoS, to aplikacje domowe np. aplikacje dla monitorowania poziomu zużycia energii w środowisku domowym. Kolejne grupy aplikacji to: monitorowanie i bezpieczeństwo publiczne (z przekazem obrazu z monitoringu mobilnego) oraz aplikacje zdalnego nauczania z możliwością zestawiania wideo-konferencji z jakością HD (*High Definition*).

Koncepcja Równoległego Internetu IPv6 QoS stanowi rozszerzenie wcześniejszych prac związanych z Systemem IP QoS dla protokołu IPv4 [1]. W odróżnieniu od rozwiązań zaproponowanych w [1], obecne prace koncentrują się na tworzeniu systemu dla RI IPv6 QoS opartego na protokole IPv6. W ramach RI IPv6 QoS realizowana jest również koncepcja sieci wirtualnych. W szczególności dla każdej aplikacji można wydzielić sieć wirtualną z dedykowanymi zasobami oraz można zastosować mechanizmy sterujące przyjmowaniem nowych wywołań w ramach klas usług dla każdej sieci wirtualnej. Z kolei dla danej sieci wirtualnej można realizować połączenia w trybie na żądanie lub połączenia planowane. Należy podkreślić, że w dotychczas prowadzonych pracach nad architekturą NGN [6] i klasami usług [10][11] skupiano się głównie na zastosowaniach architektury dla tzw. aplikacji domowych np. IPTV lub VoIP, które stosują sygnalizację SIP. W przedstawionym rozwiązaniu w warstwie usługowej zostaną zaproponowane np. rozwiązania dla aplikacji IoT. Problem opracowania wymagań na jakość przekazu dla tego typu aplikacji (np. dla aplikacji e-zdrowie) stanowi obecnie temat wielu dyskusji panelowych na konferencjach międzynarodowych, m.in. współorganizowanych przez ITU-T (*ITU-T Kaleidoskope, IEEE ICC Business Forum on eHealth Support in the Future Internet Viable Business Models for Quality of Service Assurance*). W ramach prac nad RI IPv6 QoS podejmowana jest próba opracowania odpowiednich klas usług sieciowych wraz z mechanizmami dla zapewnienia QoS dla przyjętych scenariuszy aplikacyjnych.

Artykuł jest zorganizowany następująco. Rozdział 2 przedstawia architekturę i klasy usług w Systemie IPv6 QoS, który jest tworzony dla RI IPv6 QoS. W rozdziale 3 przedstawiono zaproponowany system zarządzania. Rozdział 4 dotyczy aspektów rutingu w sieci oraz wymiarowania zasobów dla klas usług. Rozdział 5 opisuje proces tworzenia sieci wirtualnych, natomiast w rozdziale 6 przeanalizowano zagadnienia związane ze współpracą RI IPv6 QoS z siecią IPv6 typu *best effort*. Rozdział 7 stanowi podsumowanie.

2 Architektura i usługi w Systemie IPv6 QoS

Architektura Systemu IPv6 QoS jest oparta na koncepcji architektury DiffServ w warstwie sieci oraz NGN w warstwie systemu sygnalizacji i zarządzania. Należy zauważyć, że nie wszystkie elementy architektury NGN są specyfikowane w ramach Systemu IPv6 QoS. W szczególności zdefiniowano te funkcje, które są niezbędne dla zapewnienia gwarancji QoS dla wybranych aplikacji. Funkcje związane z mobilnością, identyfikacją klienta oraz naliczaniem opłat nie są objęte specyfikacją. Architektura proponowana w projekcie (Rys. 1) jest definiowana dla topologii sieci, w której wyróżniono terminale/serwery (*Customer Premises Equipment - CPE*), sieci dostępowe oraz jedno-domenową sieć szkieletową. Sieci dostępowe są podłączone do sieci szkieletowej za pomocą ruterów brzegowych. W szczególności, dla sieci dostępowej rozważane są rozwiązania oparte na standardzie IEEE 802.11. Zarówno rutery brzegowe (RB), jak i rutery szkieletowe (RS) są projektowane specjalnie dla tej sieci odpowiednio w oparciu o urządzenia EZappliance [15] (rutery brzegowe), NetFPGA [16] i Xen [17] (rutery szkieletowe). Ponadto, na zewnętrznym urządzeniu będzie realizowana warstwa sterowania dla rutingu. Takie podejście umożliwi integrację rozwiązań RI IPv6 QoS z Systemem IIP, który jest implementowany w wymienionych urządzeniach.



Rysunek 1. Architektura Systemu IPv6 QoS

Rys. 1 przedstawia proponowaną architekturę Systemu IPv6 QoS wraz z interfejsami (Rs, Rd, Rw, Rt) zgodnymi z rekomendacjami ITU-T z serii Q oraz IETF [19]. Architektura ta jest podzielona funkcjonalnie na warstwę usługową (*Service Stratum*), która umożliwia aplikacjom sygnalizację żądań QoS do sieci za pomocą funkcji SCF (*Service Control Function*). Następnie, żądania są obsługiwane przez warstwę transportową (*Transport Stratum*). Ponadto, w architekturze wyróżniamy płaszczyznę zarządzania, która z jednej strony odpowiada za konfigurację i monitorowanie rutingu i zasobów w sieci (warstwa 3 architektury IIP), z drugiej strony odpowiada za komunikację z płaszczyzną zarządzania Systemu IIP (warstwa 1 i 2 architektury IIP) oraz z płaszczyzną zarządzania sieci wirtualnych (warstwa 4 architektury IIP).

W Systemie IPv6 QoS wyróżniamy cztery typy procesów, które mogą się odbywać w różnych skalach czasowych:

- Proces zarządzania: związany z tworzeniem/usuwaniem i wymiarowaniem sieci dla RI IPv6 QoS (długa skala czasowa: miesiące, kwartały);
- Proces zarządzania: związany z tworzeniem/usuwaniem i wymiarowaniem sieci wirtualnych (długa skala czasowa: miesiące);
- Proces sygnalizacji żądań: zestawienie/usunięcie połączenia w ramach sieci wirtualnej i wybranej klasy usług (krótka skala czasowa: sekundy/minuty);
- Proces transmisji pakietów: przekaz pakietów w sieci (skala czasowa: zależna od szybkości bitowej łączy).

W referacie przedstawiono funkcje Systemu IPv6 QoS, które pozwolą na realizację w/w procesów w Systemie.

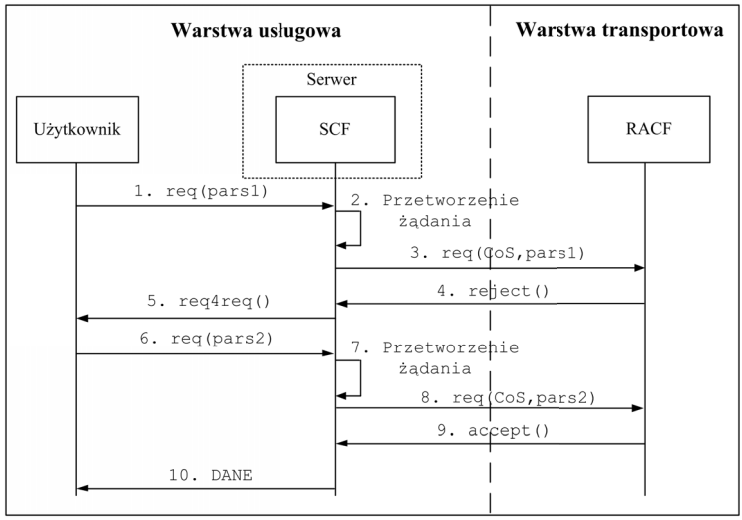
2.1 Warstwa Transportowa

Warstwa transportowa składa się z modułów funkcjonalnych systemu sygnalizacji oraz urządzeń sieciowych. Główną funkcją projektowanego systemu sygnalizacji jest funkcja RACF (*Resource and Admission Control Function*) odpowiedzialna za sterowanie dostępem do zasobów sieci. Funkcja RACF składa się z dwóch modułów funkcjonalnych PD-FE (*Policy Decision Functional Entity*) i TRC-FE (*Transport Resource Control Functional Entity*). Moduł PD-FE jest odpowiedzialny za komunikację z warstwą usługową (SCF) i przekaz wiadomości o przyjęciu lub odrzuceniu nowego żądania. Moduł TRC-FE realizuje funkcję przyjmowania nowych wywołań do poszczególnych klas usług na podstawie informacji o zasobach dostępnych w węzłach brzegowych sieci. Moduł PE-FE (*Policy Enforcement Functional Entity*) odpowiada za komunikację z danym urządzeniem sieciowym w celu ustawienia parametrów monitorowania ruchu w węzłach brzegowych sieci dla przyjmowanego połączenia (proces sygnalizacji) lub w celu przydziału zasobów dla poszczególnych klas usług w węzłach sieci (proces wymiarowania). Tym samym, moduły PE-FE dostarczają odpowiednie interfejsy dla systemu sygnalizacji do komunikacji z urządzeniami działającymi w danej sieci.

2.2 Warstwa usługowa

Warstwa usługowa jest odpowiedzialna za realizację funkcji związanych z obsługą sygnalizacji, w szczególności za sterowanie wymianą wiadomości sygnalizacyjnych aplikacji, negocjację wartości parametrów usług specyficznych dla aplikacji (np. rodzaju kodeka) oraz przekazywanie żądań rezerwacji zasobów dla poszczególnych połączeń do warstwy transportowej w celu zapewnienia wymaganego poziomu jakości usług. Dla realizacji obsługi aplikacji warstwa usługowa Systemu RI IPv6 QoS może korzystać z dwóch modułów zdefiniowanych dla tej warstwy w ramach architektury NGN [14]:

- modułu sterowania usługami (*Service Control Functions - SCF*) - odpowiedzialnego za zarządzanie zasobami, rejestrację, uwierzytelnianie i autoryzację klientów oraz usług, a także za sygnalizację specyficzną dla aplikacji;
- oraz modułu dostarczania treści (*Content Delivery Functions - CDF*) - działającego pod kontrolą modułu SCF i odpowiedzialnego za przechowywanie, przetwarzanie oraz dostarczanie użytkownikom końcowym żądanej treści.



Rysunek 2. Przykładowy scenariusz realizacji zestawienia połączenia w RI IPv6 QoS

Rys. 2 przedstawia przykładowy scenariusz wymiany wiadomości sygnalizacyjnych w Systemie IPv6 QoS dla przykładowej aplikacji. W scenariuszu tym użytkownik/terminal wysyła do serwera warstwy usługowej żądanie (1) zestawienia połączenia. Żądanie to jest odbierane i przetwarzane (2) przez moduł SCF. Na podstawie informacji zawartej w żądaniu (1) SCF generuje żądanie (3) o zestawienie połączenia w ramach klasy usług koniec-koniec (Class of Service -

CoS), czyli między użytkownikami końcowymi i wysyła je do warstwy transportowej przez interfejs SCF-RACF. W przypadku, gdy sieć nie jest w stanie obsłużyć tego żądania (4), uruchamiana jest procedura renegocjacji wartości jego parametrów (5-6). Komunikaty (8-9) reprezentują zakończoną sukcesem kolejną próbę zestawienia połączenia w ramach klasy usług koniec-koniec. Po akceptacji żądania (10) użytkownik może rozpocząć transmisję żądanych danych (11). W projekcie scenariusze wymiany wiadomości sygnalizacyjnych w warstwie usługowej są bardziej złożone np. dla aplikacji e-zdrowie, gdyż bierze w nich udział wiele elementów aplikacyjnych (więcej szczegółów np. w [13]), jednakże na ogół wszystkie takie scenariusze można zdekomponować, na mniej złożone, polegające na zestawieniu połączenia w ramach klasy usług koniec-koniec pomiędzy parą węzłów. Istotną zaletą płynącą z wydzielenia warstwy usługowej w architekturze RI IPv6 QoS jest uniezależnienie sposobu żądania zestawienia połączenia od poszczególnych komponentów aplikacji, gdyż moduł SCF będący na styku aplikacji i sieci odwzorowuje żądania aplikacji wynegocjowane dowolnym (być może bardzo złożonym) protokołem na jednolity dla wszystkich aplikacji interfejs SCF-RACF.

2.3 Klasy usług

Tab. 1 przedstawia odwzorowanie typów aplikacji na klasy usług koniec-koniec w sieci IPv6 QoS.

Tabela 1. Odwzorowanie aplikacji na klasy usług dla sieci IPv6 QoS

Typ aplikacji	Klasy usług koniec-koniec	Pole nagłówka IPv6(Traffic Class)	Wymagania QoS		
			IPLR	Średnie IPTD	IPDV
Głosowe (VoIP)	Telephony	101110	10^{-3}	100/350ms (lokalne/ międzynarodowe)	50 ms
Wideo konferencja	RT Interactive	100000	10^{-3}	100/350ms (lokalne/ międzynarodowe)	50 ms
Sygnalizacja Zarządzanie	Signalling	101000	10^{-3}	100 ms	-
Wideo na żądanie	MM Streaming	011010 011100 011110	10^{-3}	1s Nie jest krytyczne	-
FTP	High Throughput Data	001010 001100 001110	10^{-3}	1s nie jest krytyczne	-
Aplikacje sensorowe (IoT)	Low Latency Data	110000	10^{-3}	400 ms	-
Best effort	Standard	000000	-	-	-

W tabeli zawarte zostały również wymagania QoS na poziomie sieci dla klas usług koniec-koniec. Wymagania te zostały opracowane w oparciu o zalecenia

ITU-T dla obecnie istniejących typów aplikacji [4]. Wymagania QoS odnoszą się do poziomu strat pakietów (*IPLR - IP Packet Loss Ratio*), średniego opóźnienia przekazu pakietów (*IPTD - IP Packet Transfer Delay*) oraz zmienności opóźnienia (*IPDV - IP Packet Delay Variation*). W praktyce scenariusze aplikacyjne przewidziane do zastosowania mogą się odnosić do kilku typów aplikacji, np. dla monitorowania zdrowia pacjenta lub monitorowania parametrów życiowych sportowca.
Należy podkreślić, że w dotychczasowych rozważaniach w ramach ITU-T aplikacje sensorowe typu IoT nie były przedmiotem rekomendacji dla architektury NGN. Ponadto, w projekcie rozważane są scenariusze aplikacyjne złożone z wielu typów aplikacji. W związku z tym, dla każdej z grup aplikacji są opracowywane specjalne uniwersalne platformy dla zintegrowania tych aplikacji z architekturą NGN. Wymaga to opracowania podsystemów sygnalizacji w ramach realizacji warstwy usługowej architektury. Należy również podkreślić, iż opracowywane aplikacje będą tzw. aplikacjami świadomymi jakości oferowanej przez sieć (*QoS-aware*). Czyli będą miały wbudowane elementy funkcjonalne odpowiedzialne za sygnalizację żądań do warstwy usługowej architektury NGN.

W sieci dostępowej, bezprzewodowej (tj. pomiędzy punktem dostępu a CPE), QoS w warstwie drugiej będzie realizowany z użyciem funkcji EDCA (*Enhanced Distributed Channel Access*) standardu IEEE 802.11. Funkcja EDCA wprowadza cztery klasy ruchu związane z istnieniem czterech kolejek sprzętowych umieszczonych w karcie lokalnej sieci bezprzewodowej. Różnicowanie ruchu jest realizowane dzięki pracy mechanizmu szeregowania, który wybiera kolejną ramkę z odpowiedniej kolejki sprzętowej i przygotowuje ją do nadania w kanale radiowym. Prawidłowe działanie funkcji EDCA jest również możliwe dzięki zdefiniowaniu czterech parametrów ($AIFSN, CW_{min}, CW_{max}, TXOP_{limit}$), pozwalających na prawidłowe różnicowanie ruchu w przypadku rywalizacji o kanał radiowy większej liczby stacji. Przykładowy sposób odwzorowania klas usług w sieci dostępowej do klas usług funkcji EDCA standardu IEEE 802.11, a także typowe wartości parametrów funkcji EDCA zostały przedstawione w Tab. 2.

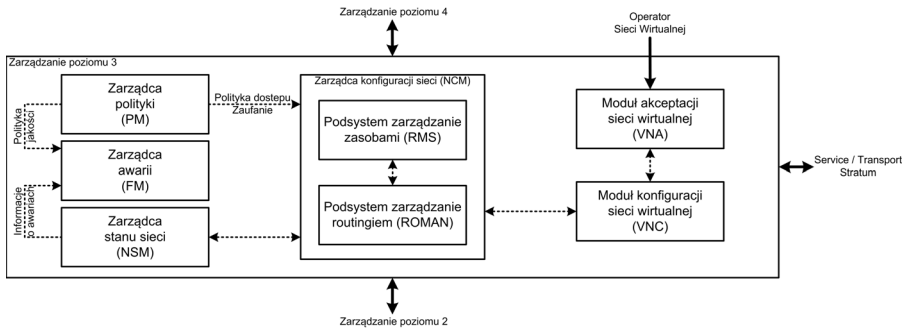
Tabela 2. Odwzorowanie klas usług w sieci dostępowej przy użyciu standardu IEEE 802.11e

Klasa EDCA	**Klasa usług**	**AIFSN**	**CWMIN**	**CWMAX**	$TXOP_{Limit}$ [μs]
Vo (Voice)	Real time	2	7	15	3264
Vi (Video)	Streaming	2	15	31	6016
BE (Best Effort)	Low Latency Data	3	31	1023	0
BK (Background)	Standard	7	31	1023	0

3 System zarządzania

Równoległy Internet IPv6 QoS ma dedykowany system zarządzania. W systemie tym wyróżniamy dwa rodzaje funkcji: nadzorujące warstwę trzecią architektury,

oraz uruchamiające komunikację z warstwą drugą i czwartą architektury Systemu IIP. Ponadto, system ten ma budowę modularną. Główna funkcjonalność systemu jest zawarta w module zarządcy konfiguracji sieci (*Network Configuration Manager - NCM*). Moduł ten zawiera podsystemy zarządzania rutingiem i zasobami. Zarządca stanu sieci (*Network Status Manager - NSM*) gromadzi informacje o działaniu sieci oraz monitoruje stan sieci. Zarządca awarii (*Fault Manager - FM*) odpowiada za reakcje na zdarzenia nieprawidłowego działania systemu. Zarządca polityki (*Policy Manager - PM*) definiuje politykę dostępu do sieci oraz reguły reakcji na sytuacje awaryjne. Pozostałe moduły umożliwiają komunikację z warstwą drugą architektury Systemu IIP w celu utworzenia żądanej topologii oraz tworzenie i usuwanie sieci wirtualnych (komunikacja z warstwą czwartą architektury Systemu IIP). Moduł akceptacji sieci wirtualnej (*Virtual Network Acceptance - VNA*) obsługuje żądania utworzenia sieci wirtualnej w warstwie czwartej, sprawdza dostępność zasobów i zapewnia szeregowanie i transakcyjność żądań. Moduł konfiguracji sieci wirtualnej (*Virtual Network Configurator - VNC*) odpowiada za uruchomienie sieci wirtualnych.



Rysunek 3. Architektura systemu zarządzania RI IPv6 QoS

4 Ruting i wymiarowanie

Topologia sieci RI IPv6 QoS jest wynikiem wirtualizacji zasobów Systemu IIP w warstwie 1 i 2 architektury. Sieć ta jest podstawą działania sieci wirtualnych warstwy 4, do których przyłączani są użytkownicy/terminale scenariuszy aplikacyjnych. Zgodnie z założeniami to sieć RI IPv6 QoS zapewnia funkcje rutingu oraz mechanizmy gwarantujące jakość obsługi na założonym poziomie. W celu zapewnienia gwarantowanych parametrów jakości obsługi oraz zagwarantowania funkcji rutingu w sieci, z każdego rutera brzegowego zostaną zestawione ścieżki poprzez sieć szkieletową do pozostałych ruterów brzegowych. Oznacza to, że z jednego rutera brzegowego zostanie zestawionych N-1 ścieżek, gdzie N jest liczbą ruterów brzegowych (całkowita liczba ścieżek jest równa N(N-1)). Zasoby

przydzielone poszczególnym klasom ruchu będą współdzielone pomiędzy wszystkie sieci wirtualne. Należy pamiętać, że określone zasoby dla poszczególnych klas ruchu muszą być również zagwarantowane w sieciach dostępowych. Wykorzystanie połączeń typu punkt-do-dowolnego punktu, z punktu widzenia efektywności wykorzystania zasobów sieci, może wydawać się nienajlepsze. Jednak z punktu widzenia uruchomienia prototypu sieci jest to rozwiązanie najprostsze.

System IIP tworzony jest na bazie wirtualizacji realizowanych programowo (Xen) i sprzętowo (karty NetFPGA oraz EZappliance). Dlatego też do realizacji funkcji rutingu (*Control Plane*) wybrano ruter programowy Quagga (stosowany z powodzeniem również w innych projektach europejskich). Rutery programowe będą odpowiedzialne za budowanie informacji o aktualnej topologii sieci (OSPFv3). Następnie tablice rutingu będą przekazywane do urządzeń realizujących funkcje węzłów sieci (*Data Plane*). Dzięki takiemu rozwiązaniu ruting realizowany jest identycznie dla każdego węzła. Każdy węzeł będzie posiadał własny ruter programowy odpowiedzialny za realizację funkcji rutingu w ramach RI IPv6 QoS oraz w przypadku ruterów brzegowych dodatkowo wszystkich sieci wirtualnych. Funkcje te będą realizowane w ramach jednego procesu protokołu OSPFv3 na zasadzie tworzenia tzw. instancji (grupy zadań realizowanych w jednym procesie). Dzięki temu realizacja funkcji rutingu znacznie się upraszcza i jest skalowalna. Jednak ze względu na różne technologie wykorzystane do tworzenia węzłów IIP należy zapewnić aby informacje o tablicach rutingu trafiały do urządzeń w sposób dla nich zrozumiały. Dlatego też konieczne jest wykorzystanie elementu pośredniczącego w postaci tzw. adaptera (*Control Plane Adapter*). Ponieważ do ruterów brzegowych mogą być dołączane całe sieci dostępowe, to aby zapewnić ruting pomiędzy urządzeniami przyłączonymi do jednej sieci wirtualnej konieczne jest tunelowanie poprzez RI IPv6 QoS wiadomości OSPFv3 pomiędzy odpowiednimi instancjami tego protokołu uruchomionymi na różnych ruterach brzegowych.

Zatem działanie sieci RI IPv6 QoS, z punktu widzenia realizacji funkcji rutingu można opisać w następujący sposób. Po utworzeniu sieci RI IPv6 QoS (wydzielenie z zasobów Systemu IIP węzłów i połączeń pomiędzy nimi), należy uruchomić serwery umożliwiające działanie Równoległego Internetu (m.in. NCM). Następnie uruchamiane są rutery programowe wraz z protokołem OSPFv3 (obsługuje IPv6). NCM gromadzi informacje o ścieżkach pomiędzy ruterami wejściowymi i wyjściowymi poprzez odpytanie wszystkich ruterów w domenie. Ścieżki te obejmują także sieć dostępową, jeśli taka istnieje. Następnie, dla każdej klasy ruchu NCM wyznacza parametry buforów i łączy a także wymusza odpowiednią konfigurację ruterów.

Wymaganą pojemność łączy dla każdej klasy usług można wyznaczyć na podstawie macierzy zapotrzebowań. Początkowo, wartości w macierzy zapotrzebowań są względne. W szczególności, wszystkie zapotrzebowania pomiędzy ruterami wejściowymi i wyjściowymi i dla wszystkich klas usług mogą być równe; w tym wypadku macierz zapotrzebowania $M[k, p]$ (k jest numerem klasy usług, p jest numerem ścieżki pomiędzy ruterami wejściowymi i wyjściowymi) jest równa jeden ($M[k, p] = [1]$). Metoda rezerwacji polega na przeliczeniu wartości

w macierzy $M[k,p]$ z pierwotnych, względnych wartości na wartości bezwzględne wyrażone w kb/s. Na podstawie wartości $M[k,p]$, możemy znaleźć wartość przepustowości dla klasy usług k w łączu l stosując wzór (1).

$$C_{i,k} = \sum_{p} M[k,p] \times \delta_{ip}$$

Oprócz wymiarowania pojemności łączy należy również przypisać odpowiednie zasoby buforów dla poszczególnych klas usług. W Tab. 3 zostało przedstawione przykładowe wymiarowanie buforów dla wybranych klas ruchu, w którym kryterium wymiarowania buforów jest maksymalne opóźnienie pakietów lub zmienność opóźnienia pakietów. Dla klasy Signalling rozmiar bufora zapewnia brak strat dla maksymalnej liczby jednoczesnych procedur konfiguracji w systemie [18]. W przypadku innych klas usług, rozmiar bufora zależy od maksymalnych wartości IPTD lub IPDV, w zależności od konkretnej klasy usług.

Tabela 3. Przykładowe wymiarowanie buforów klas usług

Klasa usług	Formuła	Wartości
Signalling	$B_{SIG} = \frac{C_{SIG}[kbps]}{50[kbps]} \times N[packets]$	C_{SIG}=przepustowość N=liczba wiadomości wysłanych bezpośrednio po sobie w jednej procedurze zestawienia połączenia
RT Interactive	$B_{RT} = \frac{IPDV_{RT}[s] \times C_{RT}[kbps]}{d_{RT}[Bytes/packets] \times 8 \times hop}$	$IPDV_{RT}$=wartość maksymalnej zmiany opóźnienia C_{RT}=przepustowość d_{RT}=długość największego pakietu hop=liczba węzłów na najdłuższej ścieżce
MM Streaming	$B_{MMS} = \frac{IPDT_{MMS}[s] \times C_{MMS}[kbps]}{d_{MMS}[Bytes/packets] \times 8 \times hop} - 1$	$IPTD_{MMS}$=wartość maksymalnego opóźnienia C_{MMS}=przepustowość D_{MMS}=długość największego pakietu hop= liczba węzłów na najdłuższej ścieżce
HTD	$B_{HDT} = \frac{IPTD_{HTD}[s] \times C_{HTD}[kbps]}{d_{HDT}[Bytes/packets] \times 8 \times hop} - 1$	$IPTD_{HTD}$=wartość maksymalnego opóźnienia C_{HTD}=przepustowość D_{HTD}=długość największego pakietu hop= liczba węzłów na najdłuższej ścieżce

5 Sieci wirtualne

Proponowana koncepcja sieci wirtualnych w ramach RI IPv6 QoS jest oparta o dostępne techniki tworzenia sieci typu VPN (*Virtual Private Network*) [9] przewidziane dla łączenia sieci klientów poprzez sieci publiczne. Głównym celem wprowadzenia koncepcji sieci wirtualnych jest przede wszystkim logiczne odseparowanie ruchu w ramach poszczególnych sieci oraz użycie algorytmów przyjmowania nowych wywołań dla zapewnienia określonej jakości połączeń w ramach

pojedynczej sieci wirtualnej. System dopuszcza jedynie możliwość funkcjonowania sieci wirtualnej w długiej skali czasowej (proces wymiarowania), natomiast zestawianie połączeń w ramach poszczególnych sieci może odbywać się na żądanie użytkownika/terminala aplikacji (proces sygnalizacji). Ponadto, żądanie może mieć charakter planowanej rezerwacji zasobów lub natychmiastowej. Ze względu na złożoność obsługi obu typów żądań w ramach danej klasy usług, przyjęto, że dana sieć wirtualna w ramach danej klasy usług może realizować tylko jeden typ żądań. Takie podejście pozwala na obsługę w ramach danej klasy usług różnych typów żądań dzięki wydzieleniu zasobów dla każdej sieci wirtualnej. W konsekwencji zasoby w postaci pasma dla każdej klasy usług są wydzielone i podlegają funkcji przyjmowania nowych wywołań w ramach każdej sieci wirtualnej.

Pojedyncza sieć wirtualna składa się z wirtualnych węzłów reprezentowanych przy pomocy dedykowanych tablic VRF (*Virtual Routing and Forwarding*) używanych na węzłach brzegowych, do których dołączone są sieci klientów. Pomiędzy węzłami wirtualnymi zestawiane są tunele w obrębie sieci szkieletowej IPv6 QoS.

Założono realizację sieci wirtualnych w warstwie trzeciej modelu ISO/OSI przy użyciu tunelowania *IPv6 in IPv6* bazującego na specyfikacji *Generic Packet Tunneling in IPv6* [7]. Dzięki temu rozwiązaniu nie jest wymagane użycie innych dodatkowych protokołów (np. GRE czy IPSec) lub korzystanie z niższych warstw przykładowo przy pomocy VLL, VPLS na warstwie drugiej modelu ISO/OSI.

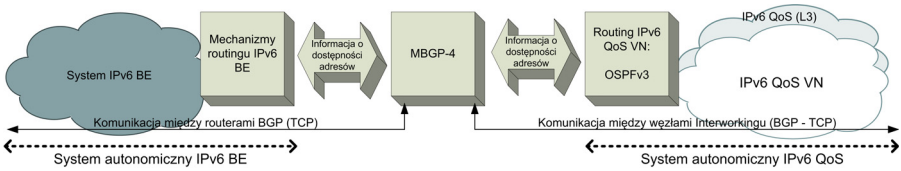
Każdy pakiet IPv6 przybywający do węzła brzegowego sieci jest identyfikowany, na podstawie portu z którego został otrzymany lub (ewentualnie w późniejszej wersji systemu) pola nagłówka *Flow Label* zawierającego identyfikator sieci wirtualnej, do której sieci wirtualnej należy. Pakiety z sieci klientów przychodzące do węzłów brzegowych są tunelowane przy pomocy opakowania ich nowymi nagłówkami IPv6, których:

- pola adresów zawierają adresy dwóch węzłów brzegowych (początkowy i końcowy tunelu);
- w polu *Flow Label* nagłówka IPv6 przesyłany jest identyfikator sieci wirtualnej, który posłuży później docelowemu węzłowi brzegowemu na wybranie odpowiedniej tablicy VRF i przekazanie pakietu do właściwej sieci klienta w zależności od adresu docelowego z opakowanego pakietu;
- do pola *Traffic Class* nagłówka IPv6 jest kopiowana wartość tego samego pola z opakowywanego nagłówka, dzięki temu węzły w sieci szkieletowej będą w stanie obsługiwać pakiety zgodnie ze zdefiniowanymi mechanizmami QoS dla danej klasy usług.

W każdej sieci wirtualnej działa instancja protokołu IGP - OSPFv3 [12], który służy nie tyle do rutowania tras pomiędzy przyłączonymi sieciami klientów, co do wymiany informacji o nich pomiędzy wirtualnymi węzłami. Proces rutowania pakietów w tunelach pomiędzy węzłami brzegowymi jest realizowany przez sieć szkieletową.

6 Współpraca z siecią *best effort*

Jednym z interfejsów tworzonego Systemu IPv6 QoS jest interfejs odpowiedzialny za współpracę Systemu IPv6 QoS z zewnętrznymi systemami sieciowymi, w naszym przypadku z siecią typu *best effort*. W tym celu w sieci dla Systemu IPv6 QoS została wydzielona klasa usług Standard dla obsługi ruchu typu *best effort*. Wprowadzenie mechanizmów, które umożliwią taką współpracę pozwoli użytkownikom sieci wirtualnych Systemu IPv6 QoS, którzy np. znajdują się poza zasięgiem tego systemu na skorzystanie np. z dostępu do wydzielonych zasobów. Integracja taka jest pewnym krokiem w procesie migracji obecnie funkcjonujących systemów IP *best effort* (BE) w kierunku Systemu IPv6 QoS. Ogólna architektura Systemu IPv6 QoS wskazuje, iż elementem, który podlega łączeniu z zewnętrznymi systemami typu *best effort* są sieci wirtualne, tworzące podstawowe środowisko sieciowe udostępniane użytkownikom i określające udostępniane użytkownikom zasoby Systemu. W celu zapewnienia możliwie dużej elastyczności rozwiązania, pozwalającej na połączenie Systemu IPv6 QoS z potencjalnie jak najszerszą grupą systemów zewnętrznych, zdecydowano się potraktować system IPv6 QoS jako niezależny system autonomiczny IPv6. W celu realizacji swej funkcji mechanizmy współpracy muszą spełnić dwa podstawowe wymagania: (1) zapewnić wymianę informacji rutingowej pomiędzy systemem autonomicznym IPv6 QoS a zewnętrznymi systemami autonomicznymi *best effort* oraz (2) umożliwić przekazywanie pakietów danych pomiędzy nimi. Zapewnienie wymiany informacji rutingowej wymaga, aby zastosowane rozwiązania współpracy były w stanie uczestniczyć w wewnętrznych mechanizmach rutingu łączonych systemów autonomicznych oraz realizowały wymianę informacji pomiędzy nimi.



Rysunek 4. Schemat dla współpracy sieci *best effort* z siecią IPv6 QoS

Jak przedstawiono na powyższym rysunku (Rys.4), do przeprowadzenia wymiany informacji rutingowej pomiędzy siecią wirtualną IPv6 QoS a zewnętrznym systemem typu best effort, zaplanowano wykorzystanie protokołu MBGP-4 [20]. Stanowi on rozszerzenie bazowego protokołu BGP-4 [20] umożliwiające, między innymi, zastosowanie go w przypadku sieci pracującej z wykorzystaniem protokołu IPv6. Poza uniwersalnością, efektywnością działania oraz zgodnością ze standardami, tego rodzaju rozwiązanie pozwala na zachowanie dużej kontroli nad informacją rutingową wymienianą między łączonymi systemami, dzięki np. możliwości łatwego prowadzenia polityk importu i eksportu tras. Dla realizacji

drugiego z wymienionych zadań, przekazu pakietów między zewnętrzną siecią IPv6 *best effort* a siecią IPv6 QoS, można zastosować klasyczne mechanizmy ich przekazywania, dzięki zachowaniu przez System IPv6 QoS kompatybilności z podstawowymi standardami IPv6 określonymi w dokumentach IETF. Należy jednak pamiętać, że na skutek znaczących różnic w realizacji mechanizmów sieciowych w obu łączonych systemach może zaistnieć konieczność modyfikacji formatu i/lub zawartości niektórych pól przekazywanego pakietu. Z tego powodu konieczne jest rozszerzenie podstawowych mechanizmów przekazywania pakietów o tzw. moduł translacji, odpowiedzialny za realizację powyższych funkcji.

7 Podsumowanie

W referacie przedstawiono propozycję architektury i mechanizmy RI IPv6 QoS, który jest rozważany jako jeden z trzech Równoległych Internetów w Systemie IIP. Referat zawiera funkcje i mechanizmy, które umożliwią działanie sieci wirtualnych oraz klas usług dla wybranych typów aplikacji, takich jak e-zdrowie, monitorowanie i bezpieczeństwo publiczne, zdalne nauczanie, w tym wideokonferencję HD. System dla RI IPv6 QoS jest budowany w oparciu o architekturę DiffServ i NGN. Obecnie System IPv6 QoS jest w pierwszej fazie implementacji, która związana jest z implementacją mechanizmów DiffServ na trzech platformach sprzętowych EZappliance, NetFPGA i Xen. Druga faza implementacji obejmuje system sygnalizacji i zarządzania. Kolejnym etapem będzie integracja Systemu IPv6 QoS z warstwami 1 i 2 Systemu IIP.

Literatura

1. W. Burakowski, et al., "Specyfikacja Systemu IP QoS opartego na architekturze DiffServ", XXV Krajowe Sympozjum Telekomunikacji i Teleinformatyki KSTiT'2009, September 2009, Warszawa.
2. W. Burakowski (ed.), "Specyfikacja Systemu IIP - poziom 1 i 2 - wersja 1", Raport projektu "Inżynieria Internetu Przyszłości", POIG.01.01.02-00-045/09-00, 2011.
3. W. Burakowski, H. Tarasiuk, A. Beben, "System IIP for supporting Parallel Internets (Networks)", Future Internet Assembly meeting, Ghent, 2010, http://fi-ghent.fi-week.eu/files/2010/12/1535-4-System-IIP-FIA-Ghent-ver1.pdf.
4. W. Burakowski, et al., "Provision of End-to-End QoS in Heterogeneous Multi-Domain Networks", Annals of Telecommunications, Springer Eds., Vol. 63, Issue 11, 2008.
5. ITU-T Rec. Y.2001, "General Owerview of NGN", 2004.
6. ITU-T Rec. Y.2111, "Resource and Admission Control Functions in Next Generation Networks", (11/2008).
7. A. Conta, S. Deering, "Generic Packet Tunneling in IPv6 Specification", Internet RFC 2473, December 1998.
8. Y. Bernet, et al., "An Informal Management Model for Diffserv Routers", Internet RFC 3290, May 2002.
9. L. Andersson, T. Madsen, "Provider Provisioned Virtual Private Network (VPN) Terminology", March 2005.

10. J. Babiarz, et al., "Configuration Guidelines for DiffServ Service Classes", Internet RFC 4594, August 2006.
11. K. Chan, J. Babiarz, F. Baker, "Aggregation of Diffserv Service Classes", Internet RFC 5127, February 2008.
12. R. Coltun, et al., "OSPF for IPv6", Internet RFC 5340, July 2008.
13. P. Świątek, P. Rygielski, A. Grzech, Resources management and services personalization in Future Internet applications, Proceedings of the First European Teletraffic Seminar, ETS 2011, February 14-16, 2011, Poznań, Poland, s. 108-112.
14. ITU-T Recommendation Y.2012 (2010), Functional requirements and architecture of next generation networks.
15. EZappliance: www.ezchip.com/p_ezappliance.htm.
16. G. Gib, et al., "NetFPGA - An Open Platform for Teaching How to Build Gigabit-Rate Networks Switches and Routers", IEEE Transactions on Education, vol. 51, no. 3, August 2008.
17. XEN: www.xen.org.
18. J.Mongay Batalla, J. Śliwiński, H.Tarasiuk and W.Burakowski, Impact of Signaling System Performance on QoE In Next Generation Networks, Journal of Telecommunications and Information Technology, 4/2009, str. 124-137.
19. D. Sun, "Diameter ITU-T Rw Policy Enforcement Interface Application", Internet RFC 5431, March 2009.
20. T. Bates et al., "Multiprotocol Extensions for BGP-4", Internet RFC 4760, January 2007.

Architektura sieci świadomych treści w Systemie IIP

Andrzej Bęben[1], Jarosław Śliwiński[1], Piotr Krawiec[1],
Piotr Pecka[2], Mateusz Nowak[2],
Bartosz Belter[3], Jakub Gutkowski[3], Łukasz Łopatowski[3],
Paweł Białoń[4], Jordi Mongay Batalla[4],
Maciej Drwal[5], Piotr Rygielski[5]

[1] Instytut Telekomunikacji, Politechnika Warszawska
[2] Instytut Informatyki Teoretycznej i Stosowanej PAN
[3] Poznańskie Centrum Superkomputerowo-Sieciowe
[4] Instytut Łączności - Państwowy Instytut Badawczy
[5] Instytut Informatyki, Politechnika Wrocławska

Streszczenie Artykuł dotyczy sieci świadomych treści CAN (ang. *Content Aware Network*), które są jednym z kluczowych rozwiązań oferowanych przez Internet Przyszłości. W artykule przedstawiono sieć PI CAN opracowaną w ramach projektu Inżynieria Internetu Przyszłości (IIP). W szczególności, omówiono środowisko sieci PI CAN wraz z podstawowymi operacjami związanymi z publikowaniem i pobieraniem treści, przedstawiono funkcje, bloki funkcjonalne i styki pomiędzy nimi, obejmujące płaszczyznę sterowania, zarządzania i przekazu danych. Ponadto, przedstawiono wybrane mechanizmy i algorytmy związane z adresowaniem i wyszukiwaniem treści, monitorowaniem stanu serwerów i sieci, a także procesem decyzyjnym. Zamieszczone wyniki symulacyjne ilustrują zalety proponowanego rozwiązania w porównaniu do obecnie przyjętych modeli dystrybucji treści.

1 Wprowadzenie

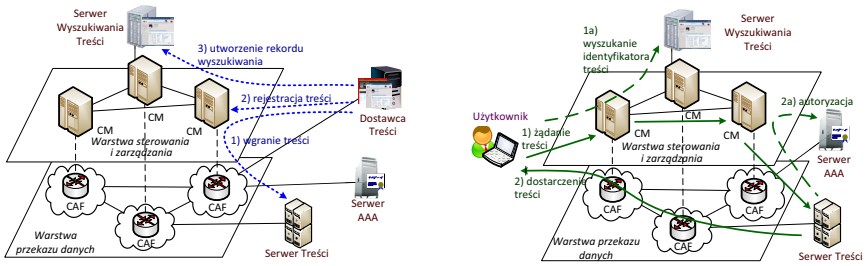
W ostatnich latach w sieci Internet można zaobserwować znaczący wzrost ruchu związanego z przekazem treści multimedialnych, takich jak filmy wideo, programy telewizyjne i radiowe, zdjęcia. Zgodnie z wynikami badań przedstawionymi w [1], przekaz treści multimedialnych stanowi obecnie około 50% ruchu, a jego udział jest stale rosnący. Stosowany powszechnie model dystrybucji treści zakłada wykorzystanie systemów pośredniczących, takich jak YouTube, Flickr, Content Delivery Networks (CDN) [2], oraz systemów wymiany danych typu peer-to-peer (p2p) [3]. Systemy te umożliwiają przechowywanie publikowanej treści oraz udostępnianie treści na żądanie użytkowników. Główne ograniczenia obecnego modelu dystrybucji treści wynikają z: (1) ograniczonej współpracy pomiędzy operatorami systemów pośredniczących a operatorami sieci, co przyczynia się do nieefektywnego wykorzystania zasobów sieci i serwerów, (2) braku

jednolitego i powszechnego systemu identyfikatorów treści, który to brak ogranicza dostępność publikowanej treści jedynie do grona użytkowników danego systemu pośredniczącego, (3) braku odpowiednich mechanizmów wspierających przekaz treści, które umożliwiałyby zapewnienie wymaganej jakości przekazu pakietów, wykorzystanie transmisji rozsiewczej, przechowywanie replik popularnej treści w pamięci podręcznej węzłów, czy też dynamiczny wybór ścieżek pomiędzy serwerem treści a użytkownikiem.

Powyższe ograniczenia, a także prognozowany dalszy wzrost ruchu związanego z przekazem treści multimedialnych, przyczyniły się do rozpoczęcia prac nad nową architekturą sieci, tzw. sieci świadomej treści - CAN (ang. *Content Aware Networks*), która jest specjalizowana dla przekazu treści. Głównym celem sieci CAN jest dostarczenie żądanej przez użytkownika treści od najlepszego źródła biorąc pod uwagę wymagania dotyczące przekazu treści, warunki ruchowe panujące w sieci oraz obciążenie serwerów. Spośród proponowanych rozwiązań należy wyróżnić rozwiązania radykalne, które wywodzące się z propozycji DONA (ang. *Data Oriented Network Architecture*) [4] oraz CCN (ang. *Content Centric Networks*) [5]. Rozwiązania te zakładają, iż węzły sieci przesyłają treść na podstawie identyfikatorów treści, a jednocześnie zachowują jej replikę w pamięci podręcznej węzłów. Ponadto, w tych rozwiązaniach założono, że sieć powinna mieć strukturę drzewa, aby usprawnić proces wyszukiwania treści. Drugim nurtem proponowanych rozwiązań dla sieci świadomych treści są tzw. rozwiązania ewolucyjne, proponowane w [6][7][8]. Rozwiązania te rozszerzają funkcjonalność sieci Internet, poprzez wprowadzenie węzłów sterujących, a także ruterów świadomych przesyłanej treści, które są odpowiedzialne za wybór serwera oraz odpowiedniej drogi w sieci w oparciu o informacje dotyczące wymaganej jakości przekazu treści, aktualnego stanu serwerów oraz warunków ruchowych w sieci.

W artykule przedstawiono architekturę sieci PI CAN (ang. *Parallel Internet CAN*) opracowaną w ramach projektu Inżynieria Internetu Przyszłości [9]. Przedstawione rozwiązanie obejmuje: (1) ujednolicony dostęp do treści, (2) nowy schemat nadawania identyfikatorów treści oraz metodę przechowywania i wyszukiwania informacji o treści, (3) mechanizmy sterowania siecią i pozyskiwania informacji o obciążeniu serwerów i warunkach ruchowych w sieci, oraz (4) proces decyzyjny umożliwiający wybór optymalnego serwera i ścieżki w sieci, które zostaną użyte do obsługi żądania wysłanego przez użytkownika. Istotną cechą opracowanego rozwiązania jest realizacja sieci PI CAN na urządzeniach umożliwiających wirtualizację.

Artykuł jest zorganizowany w następujący sposób. W rozdziale 2 przedstawiono środowisko sieci PI CAN, omówiono podstawowe operacje związane z publikowaniem treści i dostarczaniem treści, a także przedstawiono architekturę sieci PI CAN. Opis mechanizmów i algorytmów realizujących podstawowe funkcje sieci PI CAN przedstawiono w rozdziale 3. W rozdziale 4 przedstawiono wyniki badań symulacyjnych ilustrujących efektywność proponowanego rozwiązania w porównaniu do obecnie stosowanych modeli dystrybucji treści. Podsumowanie artykułu zamieszczono w rozdziale 5.



(a) publikowanie treści (b) pobieranie treści

Rysunek 1. Scenariusze podstawowych operacji sieci PI CAN związanych z publikowaniem treści oraz pobieranie treści.

2 Architektura sieci PI CAN

Sieć PI CAN realizuje dwie podstawowe operacje, tj. publikowanie treści oraz pobieranie treści. Publikowanie jest realizowane przez dostawców treści i ma na celu udostępnienie treści użytkownikom sieci PI CAN. Pobieranie treści inicjuje użytkownik przez wysłanie do sieci PI CAN żądania pobrania treści. W odpowiedzi sieć PI CAN wyszukuje najlepsze źródło żądanej treści, przygotowuje sieć do przekazu treści, a następnie inicjalizuje jej przekaz. Realizacja powyższych operacji, oprócz funkcji realizowanych przez węzły sieci PI CAN, wymaga dostępu do następujących usług zewnętrznych: usługi przechowywania treści, wyszukiwania identyfikatora treści w oparciu o dane opisujące treść, jak również usługi uwierzytelnienia, autoryzacji i naliczania. Usługa przechowywania treści publikowanej przez dostawców treści jest realizowana przez Serwery Treści (ang. *Content Server*). Usługa wyszukiwania identyfikatora żądanej treści jest oferowana przez Serwery Wyszukiwania (ang. *Content Search Server*) przechowujące opis danej treści w postaci rekordów wyszukiwania zawierających tzw. metadane powiązane z daną treścią (m.in. informacje o tytule, wydawcy czy roku produkcji). Usługę uwierzytelnienia i autoryzacji udostępnia Serwer AAA (ang. *Authentication, Autorization, Accounting*), umożliwiając w ten sposób kontrolę dostępu do poszczególnych materiałów, np. oferowanych w ramach płatnych serwisów typu Wideo na Żądanie.

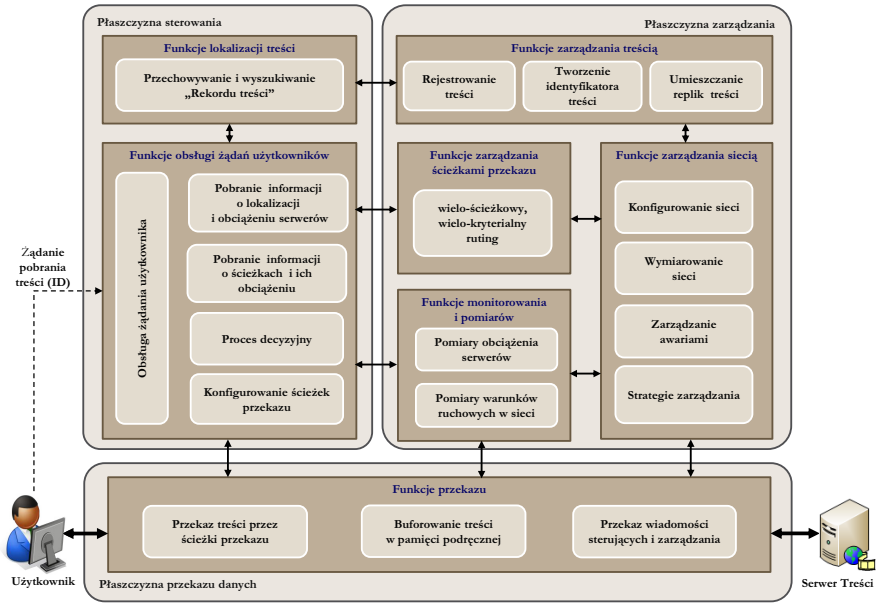
Rys. 1 przedstawia środowisko sieci PI CAN oraz scenariusze podstawowych operacji. Proces publikowania treści, przedstawiony na rys. 1a, jest inicjowany przez dostawcę treści i obejmuje trzy kroki. W pierwszym kroku, dostawca treści umieszcza repliki publikowanej treści na wybranych Serwerach Treści (jednym lub wielu). Następnie dostawca treści rejestruje publikowaną treść w węzłach sterujących CM (ang. *Content Mediator*) warstwy sterującej sieci PI CAN. W procesie rejestrowania węzły CM tworzą rekord treści CR (ang. *Content Record*), który zawiera unikalny identyfikator treści, informacje o lokalizacji replik treści na serwerach treści, a także informacje o profilu ruchowym, wymaganiach doty-

czących jakości przekazu treści oraz meta-dane opisujące treść. Ostatnim etapem procesu publikowania treści jest przekazanie opisu treści wraz z jej identyfikatorem do Serwera Wyszukiwania. Opublikowana w ten sposób treść jest dostępna dla użytkowników sieci PI CAN.

Pobranie treści wymaga, aby użytkownik znał identyfikator żądanej treści. Jeżeli identyfikator jest nieznany, proces pobierania treści jest poprzedzony procedurą wyszukiwania identyfikatora. W tym celu użytkownik wysyła do Serwera Wyszukiwania zapytanie zawierające kluczowe słowa opisujące interesującą go treść. Serwer Wyszukiwania zwraca użytkownikowi listę identyfikatorów treści odpowiadających opisowi przesłanemu w żądaniu. Jeżeli użytkownik zdecyduje się pobrać daną treść, wówczas wysyła do węzła sterującego CM żądanie pobrania treści zawierające jej identyfikator oraz kod autoryzacyjny użytkownika. W ten sposób użytkownik inicjalizuje operację pobierania treści przedstawioną na rys. 1b. Następnie węzły sterujące CM wyszukują rekord treści CR odpowiadający identyfikatorowi przesłanemu w żądaniu. Na podstawie informacji o lokalizacji serwerów zawartych w rekordzie CR oraz informacji o obciążeniu serwerów i warunkach ruchowych panujących w sieci, węzeł CM podejmuje decyzję, z którego Serwera Treści i jaką ścieżką w sieci będzie przesyłana treść do użytkownika. W kolejnym kroku, węzły sterujące konfigurują brzegowe węzły CAF (ang. *Content Aware Forwarders*) tak, aby zapewnić przekaz treści w warstwie przekazu danych zgodnie z podjętą decyzją. Następnie, węzeł CM przekazuje żądanie pobrania treści do wybranego Serwera Treści. W przypadku publicznie dostępnych treści, serwer rozpoczyna transmisję bezpośrednio po odebraniu żądania. Natomiast w przypadku treści o ograniczonym dostępie, Serwer Treści rozpoczyna transmisję po uprzedniej weryfikacji praw użytkownika do pobrania żądanej treści. Weryfikacja uprawnień użytkownika jest realizowana za pomocą Serwera AAA.

Rys. 2 przedstawia architekturę sieci PI CAN opracowaną w projekcie IIP. W architekturze tej wyróżniamy trzy płaszczyzny, tj. płaszczyznę sterowania, płaszczyznę zarządzania oraz płaszczyznę przekazu danych. Płaszczyzna sterowania obejmuje funkcje związane z lokalizacją treści (w tym przechowywaniem i wyszukiwaniem „rekordów treści") oraz obsługą żądań użytkowników. W ramach funkcji związanych z obsługą żądań użytkowników wyróżniono: (1) obsługę żądania użytkownika, (2) funkcje pobierania informacji o lokalizacji treści i obciążeniu serwerów, (3) funkcje pobierania informacji o ścieżkach i warunkach ruchowych panujących w sieci, (4) proces decyzyjny oraz, (5) funkcje związane z konfiguracją wybranej ścieżki przekazu treści. Należy zwrócić uwagę, iż konfigurowanie ścieżki dotyczy jedynie urządzeń brzegowych zlokalizowanych po stronie serwera oraz użytkownika.

W płaszczyźnie zarządzania wyróżniamy funkcje związane z zarządzaniem treścią, zarządzaniem ścieżkami przekazu treści, zarządzaniem siecią, a także funkcje monitorowania i pomiarów. Zarzadzanie treścią dotyczy rozmieszczenia replik danej treści na serwerach treści, przydzielenia unikalnego identyfikatora treści oraz zarejestrowania danej treści w sieci PI CAN. W ramach zarządzania ścieżkami przekazu są tworzone ścieżki wewnątrz poszczególnych domen oraz



Rysunek 2. Architektura sieci PI CAN opracowana w projekcie IIP.

pomiędzy domenami. Funkcje te będą realizowane przez algorytmy oraz protokoły wielościeżkowego i wielokryterialnego rutingu. Zarządzanie siecią obejmuje funkcje umożliwiające konfigurowanie urządzeń, wymiarowanie sieci, zarządzanie awariami uwzględniając różne strategie zarządzania. Ponadto, w płaszczyźnie zarządzania wyróżniono funkcje związane z monitorowaniem i pomiarami obciążenia serwerów oraz warunków ruchowych panujących w sieci.

W płaszczyźnie przekazu danych wyróżniono funkcje przekazu danych oraz buforowania treści w pamięci podręcznej węzłów. Ze względu na specyficzny format zastosowany do przekazu treści, zdefiniowano dwie niezależne funkcje przekazu, tj. funkcję przekazu treści, zgodnie z ustalaną ścieżką oraz funkcję przekazu wiadomości sterujących i zarządzających. Funkcja buforowania przesyłanej treści w pamięciach podręcznych węzłów przekazujących treść umożliwia zwiększenie efektywności sieci PI CAN poprzez skrócenie drogi, którą musi pokonać strumień pakietów, aby dotrzeć do użytkownika. Ponadto, ich zastosowanie eliminuje niekorzystny efekt obserwowany w dzisiejszym Internecie związany z wielokrotnym przesyłaniem tych samych danych po tych samych ścieżkach w tym samych czasie [10]. Przechowywanie replik przesyłanej treści w pamięci podręcznej węzłów CAF uzupełnia funkcjonalność replikacji treści na serwerach, znacznie redukując obciążenie łączy transmisyjnych, jak i procesorów w węzłach systemu. Zastosowanie pamięci podręcznej jest jednym z powszechnie przyjętych rozwiązań w sie-

ciach CAN [11][12]. W sieci PI CAN mechanizm buforowania danych zakłada zliczanie częstości odwołań do poszczególnych identyfikatorów danych przesyłanych w każdym węźle CAF i zarządzanie zawartością pamięci podręcznej przy pomocy algorytmu LRU (ang. *Least Recently Used*).

3 Mechanizmy i algorytmy warstwy sterowania i zarządzania

W rozdziale tym przedstawiono mechanizmy i algorytmy zastosowane w warstwie sterowania i zarządzania związane z adresowaniem treści, przechowywaniem i wyszukiwaniem lokalizacji treści („rekordów treści"), zbieraniem informacji o obciążeniu serwerów i warunkach ruchowych w sieci, a także procesem decyzyjnym.

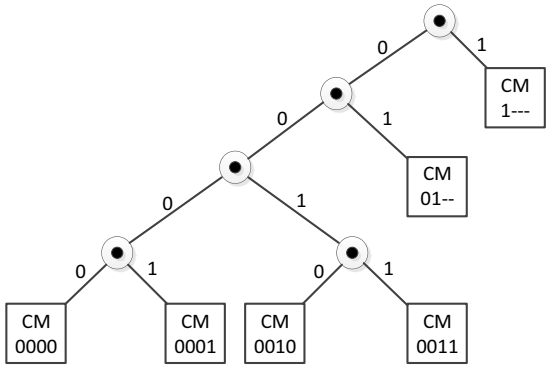
3.1 Adresowanie treści i proces wyszukiwania treści

Pierwszym etapem operacji pobierania treści w sieci PI CAN jest lokalizacja serwerów przechowujących żądaną przez użytkownika treść, wskazaną unikalnym identyfikatorem. Jeżeli użytkownik nie zna identyfikatora treści, ale zna jej atrybuty (cechy), etap ten może zostać poprzedzony procedurą wyszukiwania identyfikatorów treści spełniającej zadane przez użytkownika kryteria.

Z wyborem metody lokalizacji treści ściśle wiąże się problem formatu identyfikatora treści. Identyfikator musi być unikalny w ramach danej sieci PI CAN, gdyż pełni on rolę adresu, jednoznacznie wskazującego na daną treść. Ponadto, identyfikator musi posiadać pojemność wystarczającą do odróżnienia wszystkich potencjalnych obiektów, które zostaną udostępnione w sieci PI CAN. Z drugiej strony, identyfikator będzie umieszczany w pakietach przesyłanych przez sieć, więc powinien być jak najkrótszy dla zmniejszenia długości nagłówka. Biorąc pod uwagę powyższe cechy, korzystny jest brak struktury wewnętrznej identyfikatora, gdyż pozwala w pełni wykorzystać pojemność wynikającą z założonej maksymalnej jego długości. Dlatego w opracowanej sieci PI CAN przyjęto identyfikator binarny o długości 128 bitów, bez wewnętrznej struktury. Identyfikator ten będzie generowanych przy pomocy popularnych algorytmów, np. MD5 na podstawie binarnej zawartości obiektu lub GUID, jeżeli binarna zawartość treści nie jest znana (np. w przypadku przekazów na żywo). Algorytmy te zapewniają duże prawdopodobieństwo wygenerowania unikalnych identyfikatorów. Dodatkowo, unikalność identyfikatora jest weryfikowana podczas rejestrowania treści.

Adresy serwerów przechowujących treść są zawarte w rekordach treści, które są tworzone w podczas rejestrowania publikowanej treści. Do przechowywania i wyszukiwania rekordów treści opracowany został oryginalny algorytm COLO-CAN (ang. *COntent LOcalisation for CAN*). Algorytm ten jest oparty na idei rozproszonych tablic mieszających (ang. *Distributed Hash Table*) stosowanych w sieciach p2p, jednakże został przystosowany do specyfiki sieci PI CAN.

Należy zwrócić uwagę, iż sieci p2p charakteryzują się dużą dynamiką zmian struktury sieci ze względu na krótki czas życia węzłów, którymi zazwyczaj są stacje robocze użytkowników. W sieci PI CAN węzły sterujące CM, odpowiedzialne



Rysunek 3. Zasada tworzenia drzewa węzłów CM w algorytmie COLOCAN.

za przechowywanie i wyszukiwanie rekordów treści, są realizowane na dedykowanych serwerach, które tworzą rozproszoną bazę danych o stabilnej strukturze. Algorytm COLOCAN [13] wykorzystuje informację o strukturze połączeń pomiędzy węzłami CM. W pamięci węzła tworzone jest drzewo węzłów CM, gdzie każdy CM opatrzony jest binarnym identyfikatorem o długości zależnej od położenia w drzewie. Przykład takiego drzewa przedstawiono na rys. 3.

Identyfikator treści jest dopasowywany bit po bicie do kolejnych gałęzi drzewa, tak aby odnaleźć właściwy dla danego obiektu rekord treści. W rekordzie znajdującym się na liściu drzewa znajduje się fizyczny adres odpowiedniego węzła CM, co pozwala na bezpośrednie odwołanie do niego w celu zapytania o lokalizację danej treści. Algorytm przewiduje odpowiednie procedury dodawania i usuwania węzłów CM tak, aby działanie sieci pozostało niezakłócone także podczas zmian w topologii sieci. Każda zmiana (dodanie lub usunięcie węzła CM) jest propagowana w sieci, jednak zanim odpowiednie informacje dotrą do wszystkich CM mogących wysyłać zapytania, działanie sieci również jest zapewnione, chociaż wysłane zapytanie zostanie przekierowane do kolejnych CM i uzyskanie informacji może potrwać dłużej niż w jednym kroku.

3.2 Zbieranie informacji o stanie sieci i serwerów treści

W architekturze sieci PI CAN zdefiniowano procesy odpowiedzialne za monitorowanie i pomiary warunków ruchowych w sieci oraz stanu obciążenia serwerów. Informacje te wykorzystywane są przez proces decyzyjny przy wyborze najlepszej pary <serwer, ścieżka> z punktu widzenia efektywności przekazu treści pomiędzy serwerem treści a użytkownikiem. Monitorowanie i pomiary warunków ruchowych w sieci wymagają odpowiednich interfejsów, przy pomocy których umożliwiona zostanie komunikacja z modułami pomiarowymi umieszczonymi

w węzłach CAF przesyłających treść, jak i z komponentem odpowiedzialnym za realizację procesu decyzyjnego. Proces zbierający informacje o stanie sieci powinien być w stanie na bieżąco pobierać z modułów pomiarowych wartości parametrów związanych ze statusem portów, obciążeniem łączy, obciążeniem ścieżek, czy też poziomem strat pakietów na łączu. Informacje te są udostępniane na żądanie procesu decyzyjnego, co pozwala na wybór najlepszej pary <ścieżka, serwer treści> dla danego żądania treści.

Monitorowanie stanu serwerów jest realizowane przez funkcję bloku monitorowania i pomiarów. Należy zwrócić uwagę, iż Serwery Treści mogą być uruchamiane bezpośrednio na sprzęcie lub w postaci maszyn wirtualnych udostępnianych np. za pomocą narzędzia wirtualizacyjnego XEN. W związku z tym należy monitorować stan zasobów fizycznych, a także wirtualnych zasobów. Pomiary stanu serwerów będą dotyczyły wydajności serwerów wyrażonej przez obciążenie procesora, obciążenie łączy dostępowych, obciążenie systemu przechowującego treść, czy też liczbę jednocześnie realizowanych połączeń. Monitorowanie wydajności serwerów treści pozwala na optymalne wykorzystanie systemu tak, aby zapobiegać przeciążeniom i spadkom wydajności. Właściwe algorytmy podejmowania decyzji powinny wybierać mniej obciążone serwery treści w celu równoważenia obciążenia serwerów. Ponadto, repliki popularnej treści powinny być umieszczane w pamięci podręcznej węzłów, aby najczęściej żądane materiały były dostępne w wielu miejscach. Proces monitorowania stanu serwerów dostarcza aktualnych wartości pomiarów w odpowiedzi na żądania procesu decyzyjnego.

3.3 Proces decyzyjny

Zadaniem procesu decyzyjnego jest wybór serwera z treścią i ścieżki umożliwiających efektywne dostarczenie treści od serwera do użytkownika. Proces ten jest inicjowany jest w ramach obsługi żądania treści, po zebraniu informacji o deskryptorze ruchowym oraz wymaganiach dotyczących przekazu treści, dostępnych serwerach oraz potencjalnych ścieżkach pomiędzy serwerami a użytkownikiem. Serwery i ścieżki są charakteryzowane zarówno przez statyczne parametry, które reprezentują cechy stałe, np. maksymalna przepływności ścieżki, jak i wartości dynamicznie zmieniających się parametrów opisujących ich aktualny stan. Każda z n ($n \geqslant 1$) par <ścieżka, serwer treści> jest charakteryzowana poprzez wartości pewnych kryteriów liczbowych, związanych z profilem treści, z serwerem (np. obciążenie serwera, przepływność łącza dostępowego do serwera), z klientem (przepływność łącza abonenckiego) oraz ze ścieżką (parametry jakości przekazu pakietów oraz długość ścieżki). Wobec trudności w dokładnym prognozowaniu skutków wyboru konkretnej ścieżki na przyszły stan sieci, a także wymagania podjęcia decyzji w krótkim czasie, proponowany algorytm decyzyjny wykorzystuje do oceny ścieżek ranking oparty na analizie wielokryterialnej używającej poziomów odniesienia [14]. W proponowanym algorytmie wybierana jest para <serwer, ścieżka> znajdująca się najwyżej w rankingu. Poziomy rezerwacji i aspiracji dla każdego j-tego kryterium ustalane są w fazie wymiarowania danej domeny sieci: poziom rezerwacji można interpretować jako najgorszą, ale ciągle

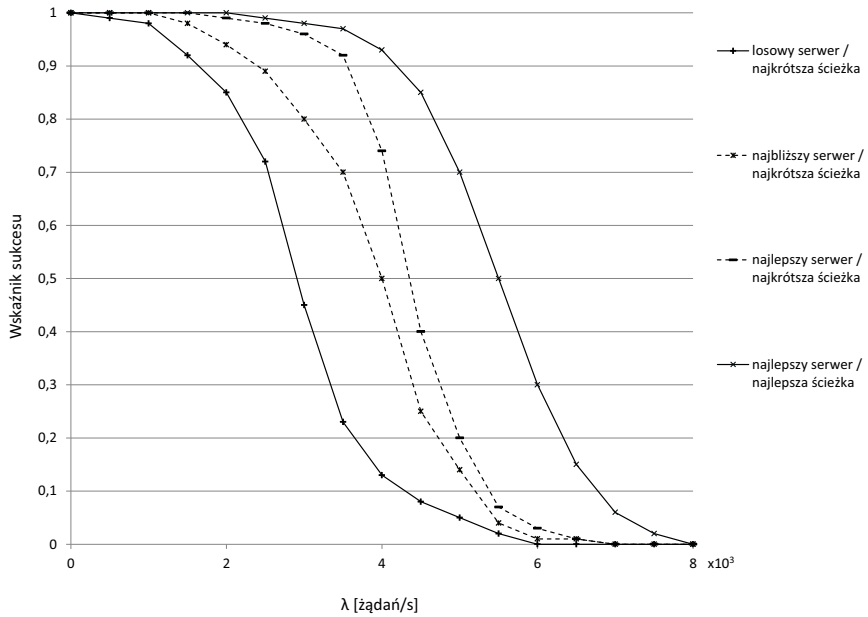
akceptowalną wartość danego kryterium, natomiast poziom aspiracji dotyczy poziomu „nasycenia satysfakcji", którego nie warto już przewyższać.

4 Badanie efektywności

W rozdziale przedstawiono wyniki badań dotyczące efektywności opracowanej sieci PI CAN. Głównym celem eksperymentów symulacyjnych jest zbadanie efektywności przekazu treści przez sieć PI CAN w porównaniu do modeli dystrybucji treści stosowanych w obecnej sieci Internet. W badaniach symulacyjnych analizowano proces żądania oraz dostarczania materiałów typu „wideo na życzenie" w wielodomenowym modelu sieci Internet. Dla każdego wywołania sprawdzano czy treść została poprawnie dostarczona do użytkownika. Przyjęto, że poprawne dostarczenie treści jest możliwe jedynie w przypadku, gdy w czasie trwania przekazu nie wystąpi przeciążenie serwera lub przeciążenie łączy na ścieżce pomiędzy serwerem a użytkownikiem. Kluczową kwestią przeprowadzonych badań jest ocena strategii decyzyjnych stosowanych w sieci PI CAN oraz w obecnych sieciach dystrybucji treści. Strategie te różnią się zakresem wiedzy o stanie serwerów, topologii i warunkach ruchowych w sieci a także wymaganiach dotyczących przekazu treści. W badaniach porównane zostały cztery strategie wyboru serwera oraz ścieżki: (1) losowy wybór serwera oraz najkrótsza ścieżka przekazu (metoda ta odpowiada powszechnie stosowanemu modelowi dystrybucji treści w obecnej sieci Internet); (2) najbliższy serwer oraz najkrótsza ścieżka przekazu (metoda ta jest stosowana w większości sieci CDN [15]); (3) najmniej obciążony serwer (najlepszy serwer) oraz najkrótsza ścieżka przekazu, oraz (4) najmniej obciążony serwer (najlepszy serwer) oraz ścieżka o najmniejszym obciążeniu ruchem (najlepsza ścieżka). Ostatnia z wymienionych strategii ta jest zastosowana w opracowanej sieci PI CAN.

W wielodomenowym modelu sieci Internet przyjęto topologię sieci oraz liczbę użytkowników w poszczególnych domenach na podstawie prac [16][17], natomiast rozkład serwerów oraz charakterystykę serwerów na podstawie założeń przyjętych w [18]. Rozmieszczenie materiałów multimedialnych oraz ich charakterystyki (czas trwania, wymagana przepływność) przyjęto na podstawie wyników prac prezentowanych w [19][20]. Żądania pobrania treści były generowane przez użytkowników zgodnie z rozkładem Poissona o zadanej intensywności (intensywność napływu zgłoszeń w poszczególnych domenach jest proporcjonalna do liczby użytkowników). W każdym żądaniu identyfikator żądanej treści był wyznaczany w sposób losowy zgodnie z rozkładem zipfa. Rozkład ten modeluje popularność treści.

Rys. 4 przedstawia zależność współczynnika poprawnego dostarczenia treści, nazwanego wskaźnikiem sukcesu, od intensywności żądań generowanych przez użytkowników uzyskaną dla przedstawionych powyżej czterech strategii decyzyjnych. Uzyskane wyniki pozwalają stwierdzić, że wykorzystanie w sieci PI CAN wiedzy dotyczącej stanu serwerów, warunków ruchowych panujących w sieci, a także wymagań dotyczących przekazu danej treści pozwala zwiększyć efektyw-



Rysunek 4. Zależność współczynnika poprawnego dostarczenia treści od intensywności napływu żądań użytkowników.

ność przekazu treści w porównaniu do obecnie stosowanych modeli dystrybucji treści.

W efekcie sieć PI CAN pozwala obsłużyć większą liczbę żądań przy tych samych zasobach sieci i serwerów. Efekt ten jest szczególnie widoczny w przypadku dużego obciążenia sieci. Należy zwrócić uwagę, iż w sieciach dystrybucji treści dość trudno prawidłowo prognozować zainteresowanie użytkowników publikowaną treścią. Z tego powodu należy oczekiwać dość częstego występowania zjawiska natłoku (ang. *flush crowd*). Z drugiej strony, w przypadku małego obciążenia, dostarczanie treści w sieci PI CAN jest tak samo efektywne jak w przypadku obecnie stosowanych modeli dystrybucji treści.

5 Podsumowanie

W artykule przedstawiono sieć PI CAN opracowaną w ramach projektu Inżynieria Internetu Przyszłości. Sieć ta definiuje nowe zasady dystrybucji treści multimedialnych, które w porównaniu do obecnego modelu zapewniają ujednolicony dostęp do treści, umożliwiają efektywne wyszukiwanie serwerów przechowujących treść żądaną przez użytkownika, jak również wykorzystują wiedzę o stanie serwerów i warunkach ruchowych w sieci przy wyborze źródła z którego będzie

dostarczana treść. Wymienione cechy sieci PI CAN umożliwiają zwiększenie efektywności przekazu treści.

W artykule omówiono środowisko opracowanej sieci PI CAN wraz z podstawowymi operacjami związanymi z publikowaniem i pobieraniem treści. Ponadto, przedstawiono architekturę sieci PI CAN wraz z realizowanymi funkcjami, zdefiniowanymi blokami funkcjonalnymi oraz stykami pomiędzy blokami funkcjonalnymi. Dodatkowo, przedstawiono wybrane mechanizmy i algorytmy związane z adresowaniem i wyszukiwaniem treści, monitorowaniem stanu serwerów i sieci, a także procesem decyzyjnym. Przedstawione wyniki symulacyjne wskazują, iż zastosowanie sieci PI CAN pozwala zwiększyć efektywność dostarczenia treści w porównaniu do modeli dystrybucji treści stosowanych w obecnej sieci Internet.

Literatura

1. Labovitz, C., Iekel-Johnson, S., McPherson, D., Oberheide, J., Jahanian F., Karir, M.: ATLAS Internet Observatory 2009 Anual Report. Dostępny na stronie: `http://www.nanog.org/meetings/nanog47/presentations/Monday/Labovitz_ObserveReport_N47_Mon.pdf`
2. Pathan, M., Buyya, R.: A Taxonomy and Survey of Content Delivery Networks. Technical Report, GRIDS-TR-2007-4, Grid Computing and Distributed Systems Laboratory, The University of Melbourne, Australia, February, 2007
3. Keong Lua, E., Crowcroft, J., Pias, M., Sharma, R., Lim, S.: A survey and comparison of peer-to-peer overlay network schemes. IEEE Communications Surveys & Tutorials, vol.7, no.2, pp. 72–93 (2005)
4. Koponen, T., et al.: A Data-oriented (and Beyond) Network Architecture. Proc. ACM SIGCOMM '07, Kyoto, Japan, pp. 181–92, August 2007
5. Jacobson, V., et al.: Networking Named Content. CoNEXT 2009, Italy, December 2009, pp. 1—12.
6. Ambrosio, M.D., et al.: Second NetInf architecture description. Deliverable 6.2 of project 4WARD, FP7-ICT-2007-1-216041-4WARD/D-6.2, January 2010, dostępny na stronie `http://www.4ward-project.eu/index.php?s=file_download&id=70`
7. García, G., Beben, A., Ramón, F. J., Maeso, A., Psaras, I., Pavlou, G., Wang, N., Śliwiński, J., Spirou, S., Soursos, S., Hadjioannou, E.: COMET: Content mediator architecture for content-aware networks. Future Network & Mobile Summit 2011, 15 - 17 June 2011, Warsaw, Poland
8. Mathieu, B., Ellouze, S., Schwan, N., Griffin, D., Mykoniati, E., Ahmed, T., Ribera Prats, O.: Improving end-to-end QoE via close cooperation between applications and ISPs. IEEE Communications Magazine, March 2011.
9. Burakowski, W., et al.: Architektura systemu IIP. Krajowe Sympozjum Telekomunikacji i Teleinformatyki, Łódź 2011
10. Leighton, T.: Improving performance on the Internet. Communications of the ACM, vol. 52(2), 2009
11. Jacobson, V., Smetters, D.K., Thornton, J.D., Plass, M.F., Briggs, N.H., Braynard, R.L.: Networking named content. Proc. of the 5th international conference on Emerging networking experiments and technologies, 2009
12. Zhang, L., et al.: Named Data Networking (NDN) Project. PARC Technical Report NDN-0001, 2010

13. Pecka, P., Nowak, M., Nowak, S.: Content localization for non-overlay content-aware networks. The 11th Int. Conference on Next generation Wired/Wireless Networks NEW2AN 2011, LNCS, Springer-Verlag 2011 (to appear).
14. Wierzbicki, A.: On the completeness and constructiveness of parametric characterizations to vector optimization problems. OR Spektrum, 8:73-87, 1986
15. Hofmann, M., Leland, R., Beaumont, R.: Content Networking: Architecture, Protocols, and Practice. Morgan Kaufmann Publisher, ISBN 1-55860-834-6, 2005
16. University of Oregon Route Views Archive Project David Meyer. Dostępny na stronie http://archive.routeviews.org/
17. RIPE Network Coordination Centre: Routing Information Service. Dostępny na stronie http://www.ripe.net/data-tools/stats/ris/ris-peering-policy
18. Akamai Technologies, Inc.: Facts & Figures. Dostępny na stronie http://www.akamai.com/html/about/facts_figures.html
19. Film Web Inc., homepage: http://www.filmweb.pl/
20. Florance, K.: Netflix Performance on Top ISP Networks. Netflix Content Delivery, January 2011 http://techblog.netflix.com/2011/01/netflix-performance-on-top-isp-networks.html

Warstwa sieci w sieciach świadomych treści

Bartosz Belter[1], Jakub Gutkowski[1], Łukasz Łopatowski[1],
Andrzej Bęben[2], Jarosław Śliwiński[2], Piotr Krawiec[2],
Jordi Mongay Batalla[3], Paweł Olender[3],
Maciej Drwal[4], Piotr Rygielski[4]

[1] Poznańskie Centrum Superkomputerowo-Sieciowe
[2] Instytut Telekomunikacji, Politechnika Warszawska
[3] Instytut Łączności Państwowy Instytut Badawczy
[4] Instytut Informatyki, Politechnika Wrocławska

Streszczenie Artykuł przedstawia warstwę sieci opracowaną dla sieci świadomej treści (ang. *Content Aware Networks*), która jest oferowana jako jedna z równoległych sieci opracowanych w projekcie Inżynieria Internetu Przyszłości. W artykule przedstawiono zastosowane mechanizmy rutingu oraz zarządzania ruchem, jak również pokazano zasadę przekazu treści bazującą na rutingu źródłowym, która umożliwia przekaz treści przez dynamicznie wybierane ścieżki. Dodatkowo omówiono mechanizmy zarządzania replikami treści oraz ich przechowywania w pamięciach podręcznych węzłów.

1 Wprowadzenie

Gwałtowny rozwój technologii multimedialnych oraz powiązanych z nimi zaawansowanych usług cyfrowych oferowanych użytkownikom końcowym zasadniczo zmienił sposób postrzegania technologii multimedialnych przez odbiorców i dostawców usług, operatorów sieci telekomunikacyjnych czy właścicieli infrastruktury. W roku 2008 ponad 4000 kin cyfrowych zostało wyposażonych w ekrany 3D wysokiej rozdzielczości. W roku 2009 zostało wyprodukowanych kilkanaście filmów pełnometrażowych z wykorzystaniem technologii 3D, równocześnie wiele wydarzeń było transmitowanych w czasie rzeczywistym na ekrany o wysokiej rozdzielczości.

Rozwój technologii medialnych w zakresie renderingu, przechwytywania obrazów oraz prezentacji wideo, wraz z postępem w dziedzinie składowania danych, przetwarzania i obliczeń, zrewolucjonizował sposób, w jaki tworzymy, udostępniamy oraz korzystamy z usług multimedialnych. Rozwój tych usług odgrywa obecnie kluczową rolę w obszarach ekonomicznych, społecznych oraz kulturalnych społeczeństwa informacyjnego, przekładając się na przełomowy rozwój powiązanych z nimi obszarów, takich jak przemysł, komunikacja, edukacja, kultura, rozrywka czy biznes.

Powyższe zjawisko przyczyniło się do rozpoczęcia badań nad metodami efektywnego dostarczania treści, które są prowadzone między innymi w ramach projektów współfinansowanych przez Komisję Europejską. Wiele z nich bezpośrednio skupia się na warstwie prezentacji, pomijając aspekty związane z przekazem

treści multimedialnych w sieciach. Gwałtowny rozwój technologii sieciowych oraz skojarzonych z nimi usług pozwolił na analizę potencjalnych możliwości integracji świata mediów z sieciami nowej generacji oferowanymi przez operatorów komercyjnych, jak również przez operatorów sieci naukowo-badawczych.

Artykuł prezentuje wyniki badań prowadzonych w projekcie Inżynieria Internetu Przyszłości (IIP) [1], dotyczące warstwy sieci opracowanej dla sieciach świadomych treści PI CAN (ang. *Parallel Internet Content Aware Networks*). Sieć PI CAN została opracowana jako jedna z równoległych sieci Internet oferowanych w systemie IIP [2]. W rozdziale 2 przedstawiono architekturę sieci PI CAN oraz podstawowe operacje związane z publikowaniem i pobieraniem treści. Opis mechanizmów rutingu wewnątrz-domenowego, między-domenowego oraz metod zarządzania ruchem zamieszczono w rozdziale 3. Rozdział 4 przedstawia zasadę przekazu treści zastosowaną w sieci PI CAN. Zagadnienia związane z zarządzaniem replikami treści przedstawiono w rozdziale 5. Ostatecznie, rozdział 6 podsumowuje artykuł.

2 Architektura sieci PI CAN i podstawowe operacje

W projektowanej sieci PI CAN wyróżniamy dwie główne operacje, tj. publikowanie treści oraz pobieranie treści przez użytkownika. Proces publikowania treści ma na celu udostępnienie treści użytkownikom sieci PI CAN. Proces ten jest inicjowany przez dostawcę treści i rozpoczyna się od zarezerwowania identyfikatora treści, który jest unikalny w obrębie sieci PI CAN. Następnie, repliki publikowanej treści są umieszczane na jednym lub większej liczbie serwerów treści CS (ang. *Content Server*). Ostatnim krokiem tego procesu jest zarejestrowanie treści w sieci PI CAN. Podczas rejestracji treści, sieć PI CAN tworzy tzw. rekord treści CR (ang. *Content Record*), który zawiera powiązanie pomiędzy przydzielonym identyfikatorem treści a lokalizacją poszczególnych kopii na serwerach CS. Ponadto, rekord treści zawiera informacje o profilu ruchowym, wymaganiach dotyczących jakości przekazu, protokole transportowym i jego parametrach, a także inne metadane opisujące daną treść.

Operacja pobierania treści obejmuje cztery działania: (1) gromadzenie wiedzy o sieci i stanie serwerów, (2) wyszukanie serwerów CS udostępniających żądaną treść oraz wybór serwera obsługującego dane żądanie, (3) przygotowanie węzłów sieci do przekazu treści pomiędzy wybranym serwerem a użytkownikiem, oraz (4) przesłanie żądania pobrania treści do serwera i rozpoczęcie przekazu treści. Gromadzenie wiedzy o stanie sieci i serwerów jest realizowane niezależnie od żądań generowanych przez użytkowników. Wiedza ta pozwala sieci PI CAN "świadomie" wybrać, z którego serwera i którą ścieżką przesłać żądaną treść do użytkownika, tak aby spełnić wymagania dotyczące przekazu treści i jednocześnie efektywnie wykorzystać dostępne zasoby. Świadomość ta jest charakterystyczną cechą sieci PI CAN i odnosi się do: a) świadomości stanu serwerów przechowujących treść, np. obciążenia, b) świadomości dostępnych dróg pomiędzy serwerami a użytkownikami, oraz c) świadomości stanu tych dróg, bądź też poszczególnych

łączy je tworzących (tj. obciążenia oraz wartości parametrów jakości przekazu pakietów: opóźnienia, jego zmienności i poziomu strat).

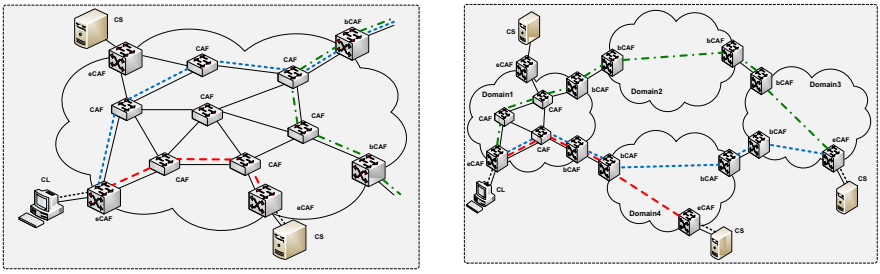
Kolejne działanie jest związane z obsługą żądania pobrania treści przesłanego przez użytkownika. W pierwszym kroku, sieć PI CAN wyszukuje rekord treści skojarzony z identyfikatorem treści przesłanym w żądaniu. Na podstawie informacji zawartych w rekordzie treści, zostają ustalone adresy serwerów przechowujących repliki żądanej treści. W następnym kroku, sieć PI CAN ustala dostępne drogi pomiędzy serwerami a użytkownikiem. Ostatecznie, sieć PI CAN podejmuje decyzję o wyborze serwera oraz odpowiedniej drogi dla danego żądania na podstawie wymagań dotyczących przekazu treści zawartych w rekordzie treści, zgromadzonej wiedzy o stanie serwerów i dróg dostępnych pomiędzy serwerami a użytkownikiem.

Trzecie działanie związane z operacją pobierania treści dotyczy przygotowania węzłów sieci do przekazu treści pomiędzy wybranym serwerem a użytkownikiem. W szczególności niezbędnym jest ustawienie mechanizmów, które umożliwią w węzłach brzegowych przyporządkowanie pakietów generowanych przez serwer do ścieżki wyznaczonej w procesie decyzyjnym, oznaczanie pakietów zgodnie z przynależnością do danej klasy usług, czy też monitorowanie ruchu generowanego z serwera. Ostatecznie, gdy serwer i sieć są już przygotowane do przesyłania treści, następuje przesłanie żądania pobrania treści do serwera, który w odpowiedzi rozpoczyna jej transmisję.

Szczegółowy opis architektury sieci PI CAN, wraz z algorytmami generowania identyfikatorów treści, wyszukiwania rekordów treści oraz algorytmami decyzyjnymi został przedstawiony w [2]. W tym artykule zaprezentowano protokoły, mechanizmy i algorytmy sieci PI CAN związane z warstwą sieci i przekazem treści.

3 Mechanizmy rutingu oraz zarządzanie ruchem

Zarówno biznesowe, jak również techniczne aspekty dzisiejszego Internetu powodują, że mechanizmy generowania, przetwarzania i przechowywania treści są niezależne od procesu dostarczania treści do konsumenta. Niezależność ta, w wielu przypadkach, prowadzi do nieefektywnego wykorzystania zasobów. Z tego względu w sieciach CAN zaproponowano odmienne rozwiązanie, zakładające ścisłą współpracę pomiędzy operatorem sieci a usługodawcami oferującymi treść. Rozwiązanie to umożliwia wybór odpowiedniego serwera oraz ścieżki, którą będzie dostarczana treść w zależności od lokalizacji konsumenta, lokalizacji serwerów przechowujących żądaną treść, a także aktualnego obciążenia serwerów oraz warunków ruchowych w sieci. Wymaga to wprowadzenia modyfikacji do mechanizmów rutingu oraz zarządzania wydajnością sieci i serwerów przechowujących treść. W dalszej części opisano wyżej wymienione mechanizmy zaproponowane w PI CAN.



(a) ruting wewnątrz-domenowy (b) ruting między-domenowy

Rysunek 1. Zastosowanie mechanizmów rutingu do wyznaczania ścieżek w sici PI CAN.

3.1 Mechanizmy rutingu

Podstawową rolą rutingu jest wyznaczenie trasy w sieci, wzdłuż której nastąpi przesłanie danych między dwoma węzłami. W przypadku sieci PI CAN trasy te są wyznaczane jako ścieżki między serwerami przechowującymi treść a konsumentami treści. W sieci PI CAN przyjęto wielo-domenową strukturę sieci, dlatego też wyodrębnić można dwa rodzaje rutingu, tj. ruting wewnątrz-domenowy oraz ruting między-domenowy.

Na rys. 1 przedstawiono koncepcję rutingu w sieci PI CAN, wewnątrz-domenowego (rys. 1(a)) oraz rutingu między-domenowego (rys. 1(b)).

Za ruting wewnątrz-domenowy, czyli wyznaczanie ścieżek wewnątrz pojedynczego systemu autonomicznego, jest odpowiedzialny algorytm rutingu wielościeżkowego (ang. *multipath routing*). Zadaniem tego algorytmu jest wyznaczenie zestawu najlepszych ścieżek między dwoma punktami w sieci na podstawie informacji o topologii, dostępnych zasobach, a także wymaganiach dotyczących przekazu treści. Informacje te będą zbierane, aktualizowane i dostarczane przez protokół rutingu typu OSPF-TE [3]. Rys. 1(a) przedstawia trzy rodzaje ścieżek, które będą wyznaczane w trakcie działania procesu rutingu wewnątrz-domenowego:

- ścieżki między serwerami treści (CS) a konsumentami treści (CL),
- ścieżki między serwerami/konsumentami treści a węzłami brzegowymi (bCAF),
- ścieżki między węzłami brzegowymi (bCAF).

W przypadku rutingu między-domenowego, wyznaczanie ścieżek między-domenowych (przechodzących przez kaskadę systemów autonomicznych) odbywać się będzie przy użyciu odpowiednio zmodyfikowanego protokołu rutingu BGP-4 [4]. Na rys. 1(b) przedstawiono przykład ścieżek między-domenowych wyznaczane między serwerami treści (CS) znajdującymi się w różnych domenach a użytkownikiem (CL) żądającym daną treść.

W trakcie działania wyżej wymienionych procesów, wybierany jest potencjalny zestaw ścieżek (między serwerami przechowującymi treść a konsumentami

treści), które mogą być wykorzystane do dostarczenia żądanej treści. Zestaw ten będzie użyty w trakcie ostatecznego wyboru ścieżki dostarczenia żądanej treści. Odpowiedzialny jest za to proces decyzyjny, którego zadanie polega na wyborze najlepszej pary serwer treści, ścieżka z punktu widzenia efektywności przekazu treści, biorąc pod uwagę nie tylko statyczne informacje o ścieżkach, ale również dane opisujące aktualny stan sieci (np. bieżące obciążenie łączy) oraz serwerów treści.

3.2 Zarządzanie wydajnością/stanem sieci

Jednym z elementów pozwalających na wybór ścieżki dostarczania treści w ramach sieci PI CAN jest proces zarządzania stanem sieci. Mechanizm ten polega na dostarczaniu na żądanie danych opisujących aktualny stan sieci do procesu decyzyjnego (pisząc tu o stanie sieci, trzeba być świadomym, że żądanie dotyczy konkretnego zestawu ścieżek w sieci). Na ten mechanizm składają się następujące procesy:

- okresowe odpytywanie modułu monitorującego węzły sieci CAN o takie informacje, jak status portu, obciążenie łącza czy współczynnik strat pakietów,
- składowanie tych danych w lokalnej bazie danych zawierającej informacje o stanie sieci,
- przetwarzanie, dopasowywanie (do poziomu rozpoznawalnego i wymaganego przez proces decyzyjny) i przekazywanie tych informacji na żądanie procesu decyzyjnego.

Wprowadzenie wyżej wymienionego mechanizmu pośredniczącego między procesem decyzyjnym a modułem monitorującym, ma na celu wyeliminowanie opóźnień w procesie wyznaczania ścieżki dostarczania treści, które mogłoby wystąpić w przypadku bezpośredniego i częstego odpytywania (o stan sieci) modułu monitorującego przez proces decyzyjny.

3.3 Zarządzanie wydajnością/stanem serwerów PI CAN

Z punktu widzenia efektywności dostarczenia treści, oprócz wydajności mechanizmów sieciowych jest istotna także wydajność serwerów, na których przechowywana jest treść. Z tego względu niezbędne jest monitorowanie i zarządzanie zasobami serwerów. Do podstawowych parametrów, których wartości będą podlegały ciągłej analizie zaliczać się będą:

- obciążenie procesora
- intensywność strumienia pakietów
- zajętość zasobów dyskowych

W sieci PI CAN podstawowym mechanizmem przeciwdziałającym przeciążeniom serwerów jest odpowiednie zarządzanie replikami treści na podstawie zaobserwowanego obciążenia serwerów oraz zapotrzebowania na treść wyrażanego przez użytkowników.

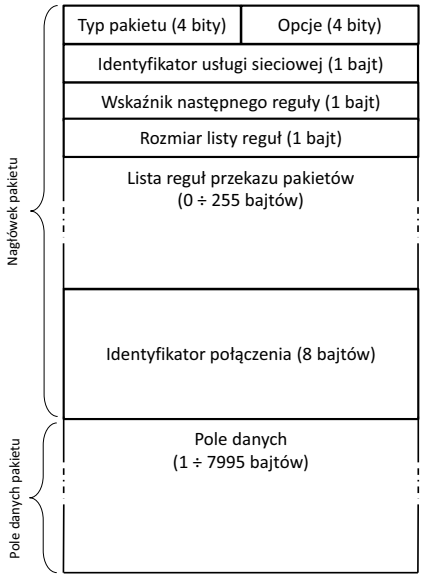
4 Warstwa przekazu danych

W sieci PI CAN wyróżniamy dwa rodzaje ruchu. Pierwszy z nich jest związany ze sterowaniem, zarządzaniem i utrzymaniem sieci, np. wyszukiwaniem i odnajdywaniem treści, ustalaniem dróg w sieci oraz zarządzaniem siecią. Drugi rodzaj to ruch związany z procesem dostarczania treści. W sieci PI CAN ruch związany ze sterowaniem jest obsługiwany przez protokół IPv6. Jednak w odróżnieniu od tradycyjnych sieci IP, schemat adresacji przyjęty w sieci PI CAN jest niezależny od adresacji stosowanej w sieciach dostępowych. W wyniku tego sieć PI CAN zapewnia separację danych wysyłanych przez użytkowników oraz serwery. Rozwiązanie to umożliwia również ukrycie miejsca, w których znajdują się serwery sterujące oraz serwery przechowujące treść (poprzez ukrycie adresów). Z drugiej strony podejście to ukrywa adresy użytkowników umożliwiając anonimowy dostęp do treści. Jednocześnie sieć PI CAN oferuję usługę uwierzytelnienia użytkownika w przypadku gdy dostawca treści wymaga tej funkcjonalności. Ze względu na to, że przekaz ruchu sterującego pomiędzy węzłami CAN odbywa się za pomocą protokołu IPv6, węzły te muszą wspierać protokoły rutingu, np. protokół OSPF jako protokół wewnątrz-domenowy, oraz BGP jako protokół rutingu stosowany na styku pomiędzy domenami.

Drugi rodzaj pakietów przesyłanych w sieci PI CAN jest związany z przekazem treści wysyłanej z serwera przechowującego treść. Ruch ten jest przekazywany w sieci za pomocą zmodyfikowanego mechanizm rutingu źródłowego (ang. *source routing*). Zakłada on, że każda domena w sieci rozpowszechnia informację o dostępnych ścieżkach a decyzja o wyborze ścieżki jest podejmowana przez proces decyzyjny niezależnie dla każdego żądania. W ten sposób ścieżka, którą będzie przesyłana treść dla danego użytkownika bierze pod uwagę stan serwerów z replikami, warunki ruchowe panujące w sieci oraz wymagania dotyczące przekazu treści. Następnie, abstrakcyjna informacja opisująca ścieżkę, czyli lista numerów domen, zostaje zamieniona na listę reguł przekazu pakietów. Lista reguł, które mają być zastosowane do przekazu treści zgodnie z wybraną ścieżką, jest umieszczana w nagłówku pakietów treści. Nagłówek ten jest dodawany przez węzeł brzegowy sieci PI CAN znajdujący się po stronie serwera treści. Klasyfikuje on strumień pakietów, w których przesyłana jest treść, korzystając z odpowiednich filtrów (adresy IP docelowy i źródłowy oraz numery portów protokołu warstwy transportowej) dodając nowy nagłówek, a jednocześnie zachowuje niezmienioną zawartość pola danych. Na wyjściu zbiorczego strumienia wykonuje się odwrotne działanie: nagłówek sieci CAN jest usuwany, a pole danych jest przesyłane do terminala użytkownika.

Na rys. 2 przedstawiono format pakietu zastosowanego do przekazu treści w sieci PI CAN. Pakiet ten zawiera następujące pola:

- typ pakietu (4 bity), zawartość tego pola umożliwia rozróżnienie pakietów przenoszących treść od pakietów zawierającymi informacje sterujące i zarządzające,
- opcje (4 bity), pole to umożliwia zidentyfikowanie opcjonalnych pól w pakiecie, np. identyfikatora połączenia,

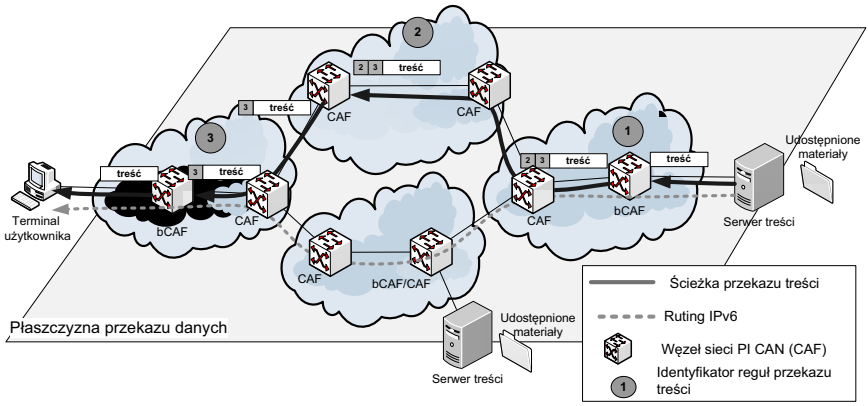


Rysunek 2. Format pakietu zastosowanego w sieci PI CAN.

- identyfikator usługi sieciowej (1 bajt), zawiera identyfikator usługi sieciowej w ramach której będzie obsługiwany pakiet,
- wskaźnik następnej reguły (1 bajt), zawiera wskaźnik następnej reguły przekazu pakietów,
- rozmiar listy reguł (1 bajt), określa liczbę reguł przekazu umieszczoną na liście,
- lista reguł przekazu pakietów (0 - 255 bajtów), zawiera identyfikatory reguł przekazu pakietów w węzłach na ścieżce przesyłania treści,
- pole danych w którym przenoszona jest treść.

Należy zwrócić uwagę, iż nagłówek pakietu sieci PI CAN ma zmienną długość, która jest uzależniona od liczby reguł przekazu pakietów wymaganych dla przesłania pakietu po ścieżce pomiędzy serwerem i terminalem użytkownika. Typowa wartość tego parametru odpowiada liczbie domen na ścieżce. Całkowita długość pakietu jest uzależniona od wartości MTU (ang. *Maximum Transfer Unit*) dopuszczalnej przez systemy transmisyjne zastosowane pomiędzy węzłami Systemu IIP. W przypadku zastosowania systemów transmisyjnych zgodnych ze standardem Ethernet, rozmiar MTU jest ograniczony do 1499 bajtów (lub 7999 bajtów w przypadku stosowania Jumbo Frames), ze względu na dodatkowy nagłówek systemu IIP o rozmiarze 1 bajtu.

Rys. 3 przedstawia zasady przekazu treści pomiędzy serwerem treści, a terminalem użytkownika zlokalizowanymi w różnych domenach sieci PI CAN. Przekaz treści umożliwiają specjalizowane węzły CAF (ang. *Content Aware Forwardes*),



Rysunek 3. Przykład przekazu treści w sieci PI CAN.

które przesyłają pakiety zawierające treść na podstawie listy reguł przekazu zawartych w nagłówku pakietów treści. Lista reguł jest wpisywana przez brzegowy węzeł CAF, oznaczony jako bCAF, który klasyfikuje strumień pakietów generowanych przez serwer treści do ścieżki przekazu treści wybranej przez proces decyzyjny. Węzły pośredniczące CAF przesyłają pakiet na podstawie reguły przekazu treści wyznaczonej przez wskaźnik następnej reguły. Ostatni węzeł CAF na ścieżce przesyła pakiet bezpośrednio do użytkownika usuwając nagłówek stosowany w sieci PI CAN.

Węzły CAF umożliwiają również realizację połączeń rozgłoszeniowych (ang. *multicast*), oraz przekaz pakietów sterujących i zarządzających. W przypadku pakietów sterujących i zarządzających przekaz jest realizowany zgodnie z tablicami rutingu IPv6.

5 Zarządzanie replikami treści i ich buforowanie

W rozdziale tym przedstawiono mechanizmy i algorytmy związane z zarządzaniem replikami treści i buforowaniem replik treści w pamięci podręcznej węzłów.

5.1 Rozmieszczenie replik treści

Podstawowym celem sieci PI CAN jest zwiększenie efektywności dostarczania treści. Jednym z mechanizmów, powszechnie stosowanym również w sieciach dystrybucji treści [5] (ang. *Content Delivery Networks*) jest zarządzanie replikami danej treści. Algorytmy zarządzania decydują o liczbie replik i ich rozmieszczeniu na serwerach treści. Odpowiednie zarządzanie treścią pozwala skrócić czas oczekiwania użytkownika na żądaną treść oraz ograniczyć ruch w sieci. W uproszczeniu można powiedzieć, że algorytmy te dążą do umieszczenia treści w pobliżu

sieci dostępowych zainteresowanych użytkowników, jednak w rzeczywistości nie jest to problem trywialny. Należy pamiętać, że większa liczba kopii wymaga większej ilości zasobów pamięciowych, co przekłada się bezpośrednio na koszty infrastruktury. Dlatego należy znaleźć odpowiedni kompromis. Przy formalnym formułowaniu problemu można wyróżnić 3 podstawowe grupy elementów: węzły z pamięcią masową, treści umieszczane w węzłach oraz użytkowników, którzy zgłaszają zapotrzebowanie na wybrane treści. Problem można przedstawić jako zadanie optymalizacji ze zmiennymi decyzyjnymi określającymi rozmieszczenie treści w węzłach przechowujących, zainteresowaniem użytkowników daną treścią, przyporządkowaniem terminali użytkowników do węzłów brzegowych sieci PI CAN oraz minimalizowaną funkcją celu wyrażającą koszt danego rozwiązania. Takie podejście pozwala na analizę problemu na różnym poziomie szczegółowości w zależności od stopnia agregacji. Tak zdefiniowane zadania są jednak NP-trudne (patrz problem rozmieszczenia treści przy założonym poziomie wydajności i minimalizacji kosztów infrastruktury w [6]). W związku z tym do rozwiązywania problemów o większym wymiarze stosowane będą heurystyki [7]. Przykładem podejścia uproszczonego może być pominięcie informacji o rozmieszczeniu treści w innych węzłach i rozpatrywanie każdego z nich osobno. Inne możliwości to relaksacja Lagrange'a oraz podejście rankingowe, w którym rozpatruje się możliwości umieszczenia dodatkowej treści w porządku wyznaczonym przez wybrane kryterium, określające konsekwencje dodania tej treści.

5.2 Buforowanie danych

Buforowanie danych w pamięciach podręcznych jest realizowane w brzegowych węzłach bCAF w celu zwiększenia efektywności dostarczania treści. Mechanizm ten redukuje liczbę transmisji tych samych danych po tych samych ścieżkach, co często występuje w obecnym Internecie [8]. Każdy pakiet w sieci CAN zawiera identyfikator treść dzięki czemu węzły CAF mogą bez wykonywania głębszej inspekcji pakietów określić jaka treść jest przesyłana i zliczać częstość odwołań. Pamięci podręczne zazwyczaj wykorzystują algorytm LRU (ang. *Least Recently Used*), dodatkowo uzupełniony o liczniki czasowe dla poszczególnych wpisów, aby zapobiec ich dezaktualizacji. Dzięki temu, pewna ograniczona liczba plików danych zostaje skopiowana i zapamiętana w węźle CAF, a następnie może być z niego wysyłana w odpowiedzi na żądania użytkowników, bez obciążania oryginalnego serwera treści. Mechanizm ten zostaje uaktywniony w odpowiedzi na żądanie użytkownika o pobranie treści. Jeżeli jednak w węźle brzegowym zbuforowane są żądane dane, wówczas ścieżka ta zostaje skrócona, co dodatkowo powoduje zredukowanie ruchu w sieci. Należy zwrócić uwagę na fakt, że taki mechanizm buforowania uzupełnia funkcjonalność algorytmów zarządzania replikami treści na serwerach, która jest jednym z podstawowych paradygmatów sieci świadomych przekazywanej treści [9].

6 Podsumowanie

W artykule przedstawiono warstwę przekazu danych opracowaną dla sieci PI CAN, która to sieć jest oferowana jako jedna z równoległych sieci Internet w Systemie IIP. Cechą charakterystyczną opracowanej warstwy przekazu danych jest zastosowanie różnych formatów danych oraz mechanizmów umożliwiających przekaz pakietów. W szczególności, wiadomości sterujące i zarządzające są przesyłane zgodnie z protokołem IPv6, natomiast treść jest przesyłana za pomocą specjalnych pakietów umożliwiających niezależny wybór ścieżki dla każdego żądania. Proponowane rozwiązanie wykorzystuje ruting źródłowy na poziomie domen, w którym węzły przesyłają pakiety na podstawie listy reguł przekazu umieszczonej w nagłówku każdego pakietu. Dodatkową funkcją opracowanej warstwy przekazu danych jest buforowanie danych w pamięci podręcznej węzłów, która pozwala na istotne zredukowanie ruchu w sieci.

Literatura

1. Burakowski, W., et al.: Architektura systemu IIP. Krajowe Sympozjum Telekomunikacji i Teleinformatyki, Łódź 2011
2. Bęben, A., et al.: Architektura sieci świadomych treści w systemie IIP. Krajowe Sympozjum Telekomunikacji i Teleinformatyki, Łódź 2011
3. Katz, D., Kompella, K., Yeung, D.: Traffic Engineering (TE) Extensions to OSPF Version 2. IETF RFC 3630, September 2003
4. Rekhter, Y., et al., : A Border Gateway Protocol 4 (BGP-4). IETF RFC 4271, January 2006
5. Pathan, M., Buyya, R.: A Taxonomy and Survey of Content Delivery Networks. Technical Report, GRIDS-TR-2007-4, Grid Computing and Distributed Systems Laboratory, The University of Melbourne, Australia, February, 2007.
6. Karlsson, M., Karamanolis, Ch.: Choosing replica placement heuristics for wide-area systems. In International Conference on Distributed Computing Systems (ICDCS), pp. 350-359, Hachioji, Japan 2004.
7. Karlsson, M., Mahalingam, M.: Do we need replica placement algorithms in Content Delivery Networks?. In Proceedings of the International Workshop on Web Content Caching and Distribution (WCW), pp. 117-128, 2002.
8. Leighton, T.: Improving performance on the Internet. Communications of the ACM, vol. 52(2), 2009
9. Zhang, et al., L.: Named Data Networking (NDN) Project. PARC Technical Report NDN-0001, 2010

Koncepcja Internetu opartego na komutacji strumieni danych dla Systemu IIP

Grzegorz Danilewicz[1], Mariusz Żal[1], Wojciech Burakowski[2], Andrzej Bęben[2], Piotr Wiśniewski[2], Mariusz Mycek[2], Henryk Gut[3], Jordi Mongay Batalla[3], Maciej Stróżyk[4], Maciej Głowiak[4], Bartłomiej Idzikowski[4], Piotr Ostapowicz[4], Krzysztof Wajda[5]

[1] Katedra Sieci Telekomunikacyjnych i Komputerowych, Politechnika Poznańska
[2] Instytut Telekomunikacji, Politechnika Warszawska
[3] Instytut Łączności – Państwowy Instytut Badawczy
[4] Poznańskie Centrum Superkomputerowo-Sieciowe
[5] Katedra Telekomunikacji, Akademia Górniczo-Hutnicza

Streszczenie W artykule przedstawiono koncepcję Równoległego Internetu, jednej z sieci przeznaczonej do pracy w Systemie IIP. System IIP jest obecnie projektowany przez uczestników polskiego projektu dotyczącego architektury Internetu Przyszłości. Podstawowym założeniem projektu jest wykorzystanie płaszczyzn sterowania oraz płaszczyzn danych w Internecie Przyszłości dla realizacji różnych usług przez różnych operatorów wirtualnych. Różne płaszczyzny danych będą pracowały równolegle i będą korzystały z innych protokołów komunikacyjnych. Stąd, w Systemie IIP nazywane są one Równoległymi Internetami (RI). W artykule przedstawiono koncepcję jednego z takich Równoległych Internetów, który ma służyć do realizacji usług z jakością obsługi zbliżoną do tej znanej z klasycznej komutacji kanałów. Nazywany jest on Równoległym Internetem z Komutacją Strumieni Danych (RI-KSD). Jest to pierwsza, wstępna prezentacja koncepcji RI-KSD, która wskazuje umieszczenie RI-KSD w Systemie IIP, usługi realizowane w tej sieci, budowę węzłów oraz zasady sterowania i zarządzania konfiguracją.

1 Wprowadzenie

Internet zmienił się znacząco od swej najwcześniejszej postaci ARPANET [1] do postaci znanej dzisiaj. Nawet zmiany z ostatnich lat odmieniły oblicze Internetu, który stał się platformą dla różnych aktywności takich jak komunikacja, biznes, edukacja czy rozrywka. Do Internetu można łączyć się z wielu typów urządzeń stacjonarnych i przenośnych takich jak komputery osobiste, laptopy i smartfony. Poza wieloma zaletami Internet ma też jednak wady. Sieć ta powstała początkowo jako narzędzie dla naukowców, w której przesyłanie informacji odbywała się na zasadzie *best effort* bez zapewnienia mechanizmów jakości obsługi i bezpieczeństwa przesyłanych informacji. Mechanizmy takie stały się potrzebne, gdy w sieci Internet zaczęto realizować różne usługi multimedialne

(VoD, IPTV, VoIP, wideotelefonia itp.) oraz inne usługi takie jak bankowość internetowa i zakupy w sieci. Liczba urządzeń przyłączonych do sieci Internet zwiększyła się gwałtownie. Z tych powodów do sieci Internet wprowadzono lub wprowadza się nowe rozwiązania takie jak: DiffServ [2], [3], IntServ [4], obsługa bezpieczeństwa [5], [6] i IPv6 [7]. Niezależnie od zmian wprowadzanych w celu usprawnienia sieci Internet, w ostatnich latach pojawiło się wiele inicjatyw, które mają za zadanie zaprojektować zupełnie nową sieć Internet. Sieć rozważana w tych inicjatywach jest często nazywana Internetem Przyszłości FI (ang. *Future Internet*).

Usprawnienie obecnej sieci Internet może być zrealizowane na dwa sposoby. W wyniku zmian ewolucyjnych, metodą małych kroków podąża się w kierunku nowej koncepcji sieci. W wyniku, Internet powinien zmienić się w najbliższej przyszłości. Koncepcja ta nosi miano Sieci Następnej Generacji NGN (Next Generation Network). Inne podejście zakłada zbudowanie całkowicie nowej sieci od podstaw. Obecnie, wiele projektów na całym świecie skupia się na koncepcji sieci FI. Jednym z celów siódmego europejskiego programu ramowego [8] jest "Sieć Przyszłości". W ramach tego celu uruchomionych zostało wiele projektów, z których najważniejsze to: TRILOGY [9], 4WARD [10], GEYSERS [11]. Przykładami innych projektów są japońskie AKARI [12] i amerykańskie GENI [13]. W tych projektach zakłada się nowe modele biznesowe, z nowymi graczami na rynku, którzy odpowiadają za dostarczanie treści i usług. Jednym z takich modeli jest wirtualizacja sieci, wprowadzenie sieci wirtualnych i operatorów sieci wirtualnych.

W polskim projekcie poświęconym przyszłej sieci Internet także zakłada się wirtualizację sieci [14]. Uczestnicy projektu projektują i zmierzają do budowy prototypu Systemu IIP (Inżynieria Internetu Przyszłości). Założeniem projektu jest wykorzystanie różnych technik transportowych w sieci gdzie wirtualizacja jest realizowana na dwóch poziomach: na poziomie transportowym, przez wykorzystanie różnych technik transportowych oraz wewnątrz technik transportowych przez utworzenie różnych sieci wirtualnych korzystających z tej samej techniki transportowej. Takie podejście pozwoli na wykorzystanie protokołów mających swoje korzenie w obecnym Internecie oraz zupełnie nowych protokołów, które pojawią się w przyszłości.

Zaproponowany w projekcie model składa się z czterech poziomów [14], [15]. Wirtualizacja bazująca na technikach transportowych jest realizowana na poziomie 2. Sieci wirtualne korzystające z różnych technik transportowych są nazywane Równoległymi Internetami. W Systemie IIP zdefiniowano trzy RI – RI-IPv6 QoS (bazujący na IPv6), RI-CAN (bazujący na koncepcji sieci świadomej przekazywanej treści – Content Aware Network) i RI-KSD (bazujący na koncepcji zbliżonej do komutacji kanałów, przeznaczony głównie dla usług czasu rzeczywistego). W każdym RI możliwe jest tworzenie wielu sieci wirtualnych. Więcej informacji o projekcie można znaleźć w [14], [15].

W artykule przedstawiono koncepcję Równoległego Internetu z Komutacją Strumieni Danych (RI-KSD) (ang. *Parallel Internet - Data Streams Switching, PI-DSS*). Sieć ta powinna zapewnić jakość obsługi zbliżoną do sieci z komuta-

cją kanałów. Problemem jaki należy rozwiązać w sieci realizującej usługi czasu rzeczywistego jest konieczność zapewnienia małego opóźnienia i małej zmienności opóźnień. Dane użytkownika umieszczane są w ramkach RI-KSD (inaczej nazywanych PDU) i wysyłane są do sieci jako strumień następujących po sobie, wysyłanych w określonych odstępach czasu ramek.

W dalszej części artykułu w rozdziale drugim przedstawiono podstawowe założenia sieci RI-KSD oraz usługi jakie będą w niej realizowane. W rozdziale trzecim omówiono zagadnienia płaszczyzny danych w tym strukturę ramki oraz węzła sieci RI-KSD. W rozdziale czwartym zawarto informacje o płaszczyźnie sterującej. W piątym rozdziale omówiono zarządzanie konfiguracją w sieci RI--KSD. Na końcu zamieszczono wnioski i wskazania do dalszej pracy.

2 Koncepcja sieci z KSD

2.1 Ogólna koncepcja sieci KSD

Jak uprzednio wspomniano, celem rozważania sieci Komutacji Strumieni Danych (KSD) jako jednego z Równoległych Internetów jest zbudowanie sieci pozwalającej na przekaz strumieni o stałej szybkości z zapewnieniem jakości przekazu zbliżonym do poziomu oferowanego przez klasyczne sieci z komutacją kanałów. W technice IP, do obsługi takiego ruchu jest zdefiniowana np. klasa RT Interactive [16]. Jednakże, użycie tej klasy powoduje konieczność multipleksacji z innym ruchem przekazywanym w sieci IP i ograniczeniem się do formatu pakietów IP. Powyższe powoduje, iż w konsekwencji w technice IP będzie ograniczona możliwość zapewnienia jakości przekazu pakietów, mierzona przez takie parametry jak opóźnienie pakietów IPTD (ang. *IP Packet Transfer Delay*), zmienność opóźnienia pakietów IPDV (ang. *IP Packet Delay Variations*) i poziom strat pakietów IPLR (ang. *IP Packet Loss Ratio*). To co może być oferowane w sieci IP, to wartości IPTD < 100ms, IPDV < 50 ms oraz IPLR $< 10^{-4}$. Od sieci KSD oczekuje się mniejszych wartości tych parametrów, szczególnie dla IPLR i IPDV, co wynika np. z wymagań na przekaz stereoskopowych strumieni wideo.

RI-KSD przyjmuje, że dla przekazu strumieni danych wykorzystuje się multipleksację typu REM (ang. *Rate Envelop Multiplexing* [16]), która zakłada jedynie buforowanie danych (wymagane są krótkie bufory) napływających w tej samej chwili czasowej, a nie jest ona projektowana do absorpcji ruchu chwilowo przewyższającego szybkość łącza. Długość buforu określa wartość parametru IPDV (który przyjmuje się za równy IPTD) zaś dopuszczalne obciążenie na łączu ρ jest funkcją długości bufora (B) i założonej wartości IPLR. Powyższe wielkości łączy następująca zależność (1):

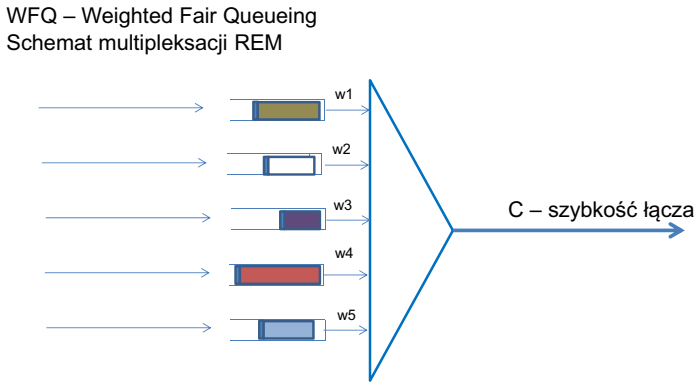
$$\frac{2B}{2B - \ln(\text{IPLR})} = \rho. \tag{1}$$

Jak można zauważyć z zależności (1) przez wartość parametru ρ można regulować wartość IPLR. Nowe wywołanie wymagające szybkości bitowej PR_{new} w łączu o szybkości bitowej C i założonym obciążeniu ρ może być przyjęte do obsługi jedynie, kiedy spełniony jest warunek:

$$PR_{new} + \sum_{i=1}^{N} PR_i \leqslant \rho \cdot C \tag{2}$$

gdzie N jest liczbą już realizowanych połączeń, z których każde wymaga szybkości bitowej PR_i $(i = 1, \ldots, N)$.

Schemat działania węzła przedstawiono na rys. 1. Dla danego połączenia obsługiwanego w danym węźle (częściej dla ścieżki wirtualnej o dedykowanej szybkości bitowej obsługującej grupę połączeń wirtualnych) dedykowany jest oddzielny bufor z dostępem do łącza o określonej szybkości. Dostęp do danej szybkości w łączu realizowany jest przez mechanizmy szeregowania takie jak WFQ (ang. *Weighted Fair Queueing*). W szczególności, danemu i-temu połączeniu nadawana jest waga w_i, a zatem to połączenie będzie mogło korzystać z szybkości $w_i C$.

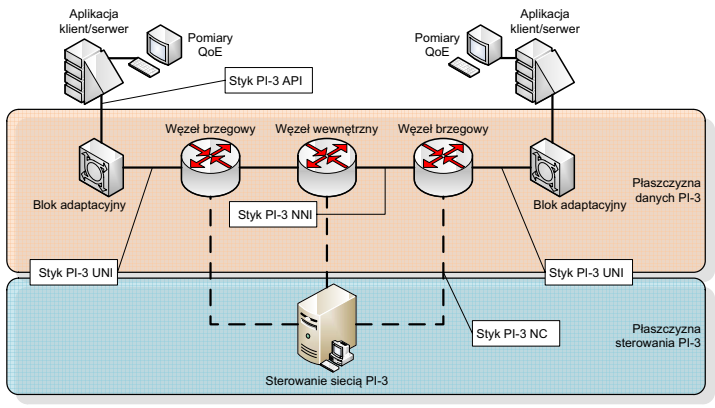


Rysunek 1. Przykład dostępu do łącza wyjściowego o szybkości C obsługującego 5 połączeń o wagach od w_1 do w_5

2.2 Architektura sieci RI-KSD

Na rysunku 2 przedstawiono architekturę sieci RI-KSD. Węzły sieci łączone są z wykorzystaniem mechanizmów transportowych Systemu IIP. W sieci RI-KSD wyróżnia się dwa rodzaje węzłów – brzegowe i wewnątrzsieciowe, różniące się funkcjonalnością w zakresie obsługi w płaszczyźnie danych jak i w płaszczyźnie sterującej. Na przykład funkcja przyjęcia zgłoszenia do sieci będzie realizowana w węźle brzegowym. W pierwszym etapie realizacji sieci RI-KSD zakłada się scentralizowane sterowanie siecią. Zadaniem węzłów zarówno brzegowych jak i wewnątrzsieciowych będzie przekazywanie informacji sygnalizacyjnych w sieci do sterowania siecią w celu realizacji podstawowej obsługi połączeń. Aplikacje dla

poszczególnych usług przyłączane są do sieci z wykorzystaniem *Bloków Adaptacyjnych*, których zadaniem będzie dostosowywanie formatu danych do transmisji w sieci RI-KSD.



Rysunek 2. Architektura sieci RI-KSD

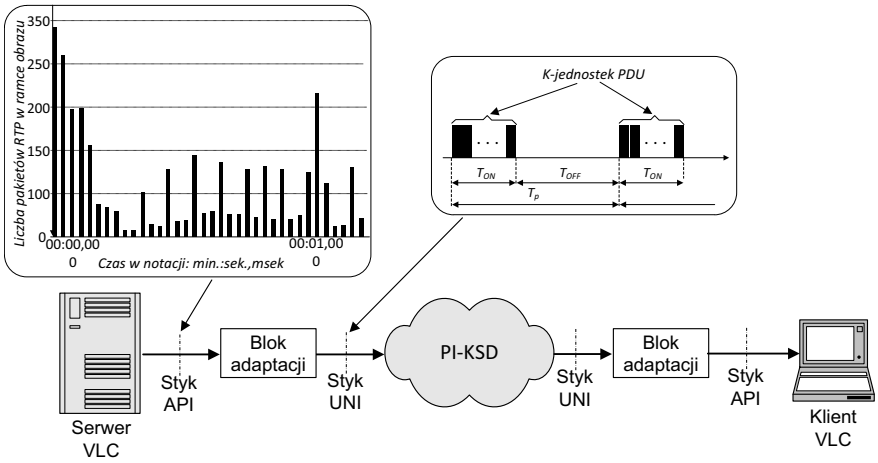
2.3 Charakterystyka aplikacji dla usług w sieci RI-KSD

Rozważa się dwa rodzaje aplikacji, a mianowicie: aplikacje istniejące na rynku (*Legacy Applications*), które nie są przystosowane do bezpośredniej współpracy z nową siecią RI-KSD oraz aplikacje, które są tworzone od podstaw z uwzględnieniem specyfiki i wymagań tej sieci (*Native Applications*), i które w sposób świadomy korzystają z możliwości oferowanych przez tę sieć.

Jedną z usług multimedialnych udostępnianych z wykorzystaniem sieci RI-KSD będzie usługa wideo na żądanie, z obrazem dwu- lub trój-wymiarowym, zwana dalej usługą VoD 2D/3D. Jeden ze scenariuszy zakłada, że usługa ta będzie tworzona z wykorzystaniem aplikacji *VLC Media Player* dla strony klienta i serwera strumieniującego. Ponieważ aplikacje te są typu *Legacy Applications*, ich dołączenie do sieci RI-KSD wymaga zastosowania dodatkowych elementów, takich jak: Serwer Proxy i Blok Adaptacyjny.

Serwer Proxy działa na poziomie warstwy aplikacji i wykonuje funkcje związane z obsługą żądania rezerwacji zasobów w sieci RI-KSD, niezbędnych dla strumieniowania wybranego filmu z aplikacji serwera VoD 2D/3D do aplikacji klienta inicjującej to żądanie. Serwer ten dokonuje konwersji odpowiednich sekwencji wiadomości sterujących wymienianych między aplikacjami na odpowiednie polecenia protokołu sygnalizacyjnego ze styku UNI, prowadzące do: rezerwacji zasobów i zestawienia ścieżki w sieci RI-KSD dla strumieniowania wybranego materiału filmowego, czy też – do rozłączenia ścieżki zestawionej i zwolnienia przydzielonych dla niej zasobów sieciowych.

Blok adaptacyjny działa na poziomie przekazu danych, jest przeźroczysty dla wymiany wiadomości między współpracującymi aplikacjami VLC, ma wbudowaną funkcjonalność rozróżniania wiadomości przekazywanych między aplikacjami oraz może obsługiwać wiele niezależnych sesji uruchomionych na jednym urządzeniu. Blok adaptacyjny realizuje dwie grupy funkcji. Jedną stanowią operacje związane z dopasowywaniem pakietów nadawanych przez aplikacje VLC do formatu jednostek PDU i na odwrót. Drugą są mechanizmy kształtowania ruchu, przekształcające strumień pakietów (z interfejsu nadawczego API serwera VLC) o zmiennej szybkości przekazu VBR[6] (ang. *Variable Bit Rate*), na strumień jednostek PDU (na interfejsie UNI) przenoszących te pakiety, które to jednostki są przesyłane w grupach ze stałym odstępem między grupami, a także – mechanizmy odtwarzania źródłowego rytmu pakietów odbieranych na interfejsie odbiorczym API po stronie klienta VLC. Zasada kształtowania ruchu przez blok adaptacji jest pokazana na rys. 3.



Rysunek 3. Zasada przekształcania strumienia pakietów VBR na strumień CBR przez blok adaptacji (oznaczenia wyjaśniono w tekście referatu)

2.4 Przykładowa aplikacja

Budowa sieci uwzględniającej model komutacji strumieni z gwarantowaną przepływnością jest szansą stworzenia optymalnego środowiska transmisyjnego dla

[6] Należy rozróżnić między CBR dla materiału filmowego, oznaczające stałą ilość danych w jednostce czasu (najczęściej 1 s), ale niekoniecznie równo rozłożonych w tej jednostce czasu, a CBR dla sieci, oznaczające równomierną, chwilową ilość transmitowanych danych w przeliczeniu na jednostkę czasu. CBR dla materiału filmowego oznacza najczęściej VBR dla sieci. Dla uniknięcia nieporozumień zastosowane określenia CBR/VBR dotyczą sieci.

aplikacji przekazujących dane w szerokim paśmie transmisyjnym, przy utrzymaniu stałej przepływności bitowej transmitowanego strumienia.

Zaproponowana aplikacja strumieniująca dane multimedialne o wysokich rozdzielczościach z wykorzystaniem natywnych mechanizmów sieci RI-KSD jest przykładem oprogramowania doskonale wpasowanego w ramy powyższych założeń, jednak przystosowanie jej do pracy w trybie natywnym wymaga odpowiedniego nakładu prac rozwojowo-badawczych.

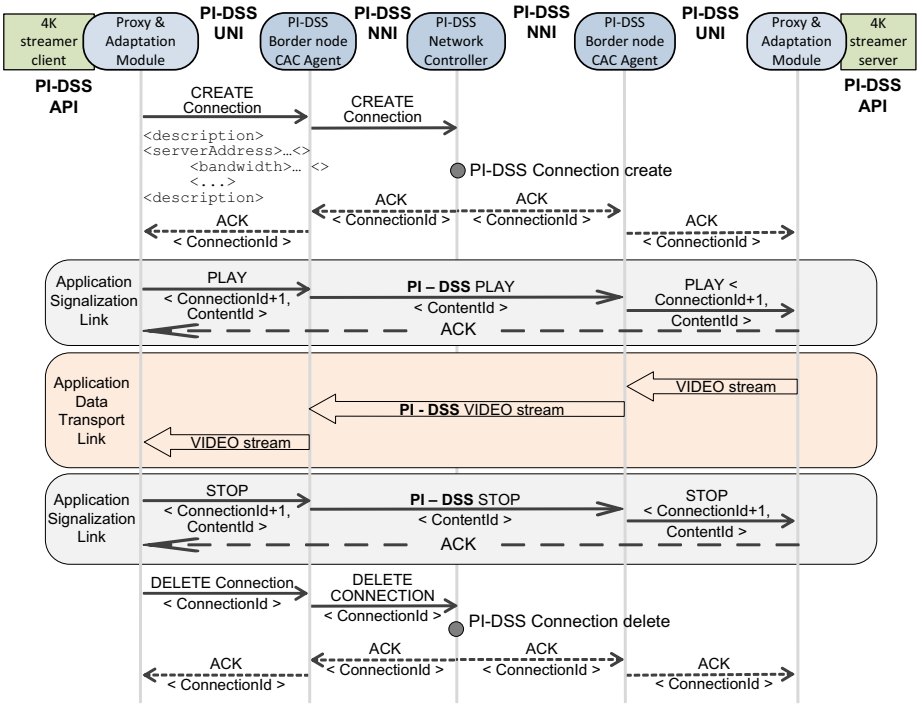
Aplikacja ma możliwość strumieniowania materiału *audio-video* zarówno z repozytoriów (materiał pre-kodowany) jak i *live* (z wykorzystaniem dedykowanego modułu sprzętowego do kamery RedOne). Materiały *live* pobierane z urządzeń zewnętrznych są kodowane w czasie rzeczywistym przez sprzętowe kodery JPEG2000, natomiast materiały składowane w repozytoriach treści mogą być zakodowane zarówno algorytmem JPEG jaki i JPEG2000. Wykorzystanie algorytmu JPEG2000 zapewnia wysoką jakość obrazu, a ponadto pozwala na narzucenie ograniczeń dotyczących pasma transmisyjnego dla kodowanych treści w granicach od 40Mbit/s do 250Mbit/s, w zależności od rodzaju materiału wejściowego (HD, HD3D, 4K, 4K3D).

Aplikacja korzystająca z sieci RI-KSD wyposażona zostanie w moduł bloku Proxy i Adaptacji, który na drodze formowania struktur utrzymanych w ścisłej zgodności ze specyfikacją modelu RI-KSD, będzie w stanie zestawić połączenie oraz transmitować dane wykorzystując podstawowe zalety sieci RI-KSD. Za odbiór, obsługę i odpowiednie kierowanie danych i komunikatów sygnalizacyjnych aplikacji odpowiedzialny będzie pierwszy brzegowy węzeł sieci RI-KSD pełniący również funkcję Agenta CAC. Podstawowy schemat komunikacji aplikacji strumieniującej z siecią RI-KSD przedstawiony jest na rys. 4.

Przesłanie strumienia danych w sieci RI-KSD wymagało będzie od aplikacji zestawienia dedykowanego połączenia o zadanej przepływności pomiędzy serwerem i klientem na czas transmisji. Moduł proxy związany z aplikacją klienta, korzystając ze zdefiniowanego kanału sygnalizacyjnego, prześle do Agenta CAC odpowiednie żądanie, które następnie skierowane zostanie do centralnego węzła zarządzającego siecią.

Po pomyślnym zakończeniu procedury zestawienia kanału komunikacyjnego, do modułu proxy zwrócony będzie identyfikator strumienia, który jednoznacznie identyfikuje połączenie. *Blok Adaptacyjny* będzie go używał w ramkach do transmisji danych. Dodatkowo ze strumieniem danych skojarzony zostanie strumień sterowania niezależny od strumienia danych, w którym przesyłane będą aplikacyjne komunikaty sygnalizacyjne.

Po zakończeniu transmisji, na żądanie aplikacji, kanał transmisyjny zostanie usunięty i wszystkie zasoby sieci RI-KSD, wykorzystane do zestawienia danego połączenia, zostaną zwolnione i będą mogły być ponownie wykorzystane do utworzenia nowych kanałów.



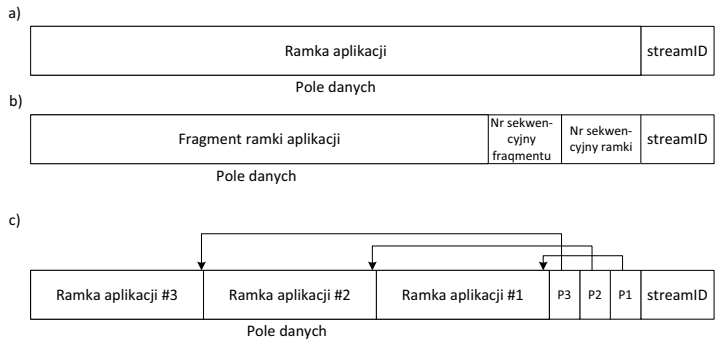
Rysunek 4. Przykładowy schemat komunikacji aplikacji z siecią RI-KSD

3 Płaszczyzna danych w sieci RI-KSD

3.1 Format danych w sieci RI-KSD

Dane pochodzące od aplikacji korzystających z usług Równoległego Internetu RI-KSD zorganizowane są w strumień ramek o określonej długości. Należy rozważyć trzy scenariusze współpracy aplikacji i sieci RI-KSD. Pierwszy, gdy długość ramki pochodzącej od aplikacji jest nieznacznie mniejsza od maksymalnego pola danych ramki RI-KSD. Jest to przypadek najłatwiejszy do implementacji, gdyż nie wymaga on żadnych dodatkowych mechanizmów transmisji. Takie ramki są bezpośrednio umieszczane w polu danych ramki RI-KSD i przesyłane są do użytkownika docelowego. Jeśli ramka RI-KSD może pomieścić kilka ramek danych aplikacji, wówczas należy rozważyć możliwość enkapsulacji kilku ramek pochodzących od aplikacji w jednej ramce RI-KSD. Enkapsulacja może wymagać mechanizmów pozwalających na wyodrębnienie poszczególnych ramek aplikacji, to znaczy wskazania ich początku i końca w polu danych RI-KSD. Wadą tego rozwiązania jest wprowadzanie dodatkowego zmiennego opóźnienia w transmisji ramek. W trzecim scenariuszu rozważane są ramki, które nie mieszczą się w jednej ramce RI-KSD. Przypadek ten wymaga implementacji mechanizmów dzielenia i ponownego składania ramek aplikacji oraz wykrywania utraty fragmentu ramki.

Formaty ramek RI-KSD zostały przedstawione na rys. 5. Każda ramka rozpoczyna sie identyfikatorem strumienia *streamID* o strukturze hierarchicznej. Pozwala ona na wyróżnienie ramek zawierających dane aplikacji oraz wiadomości sygnalizacyjne przeznaczone dla węzłów RI-KSD oraz na wskazanie węzła i interfejsu, do którego dołączony jest użytkownik docelowy, oraz aplikacji działającej w obrębie danego użytkownika. Rodzaj scenariusza współpracy określany jest w chwili negocjacji kontraktu między aplikacją a siecią. W zależności od określonego scenariusza za identyfikatorem *streamID* znajdują się dane aplikacji, wskaźniki do początków poszczególnych ramek aplikacji lub numer sekwencyjny ramki i numer sekwencyjny fragmentu ramki.



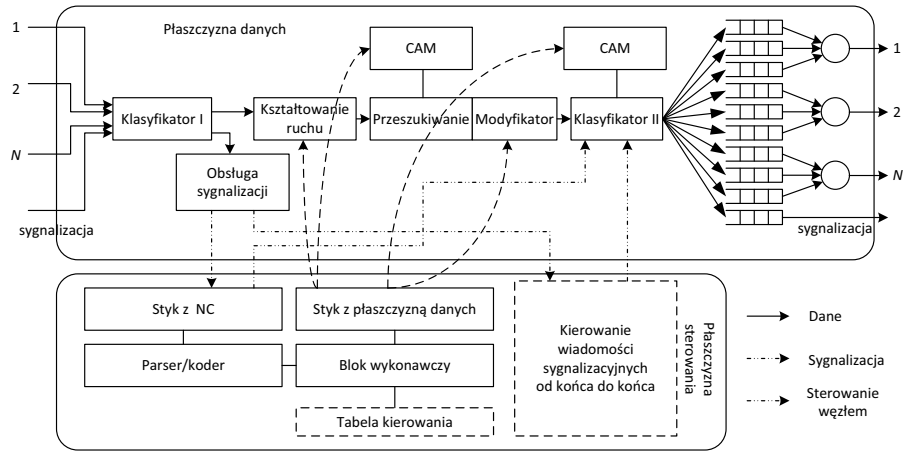
Rysunek 5. Budowa ramek RI-KSD: (a) ramka aplikacji jest równa lub nieznacznie mniejsza od ramki RI-KSD, (b) ramka aplikacji jest większa od ramki RI-KSD oraz (c) ramka aplikacji jest wielokrotnie mniejsza od ramki RI-KSD

3.2 Struktura węzła

Struktura węzła równoległego Internetu RI-KSD przedstawiona została na rys. 6. Strumienie danych oraz wiadomości sygnalizacyjne sterujące pracą węzła, jak również przeznaczone do innych węzłów, po odebraniu z poziomu 1-2 Systemu IIP podawane są na blok klasyfikatora w płaszczyźnie danych. W bloku tym następuje sprawdzenie poprawności ramki oraz rozróżnienie ramek przynależnych do strumienia danych od ramek zawierających wiadomości sygnalizacyjne. Ramki wchodzące w skład strumienia danych podawane są do bloku kształtowania ruchu, którego głównym zadaniem jest kontrola realizacji kontraktu zawartego między użytkownikiem a siecią. W przypadku przekroczenia przez użytkownika parametrów kontraktu (np. ilości przesyłanych danych), blok kształtowania ruchu może odrzucić nadmiarowe ramki. Zaakceptowane ramki przekazywane są do bloku przeszukiwania i modyfikacji. W bloku tym, na podstawie danych zawartych w pamięci adresowanej zawartością CAM (ang. *Content-Addressable Memory*), następuje wyznaczenie łącza wirtualnego, którym powinna zostać wysłana ramka. Blok ten pozwala również na modyfikację nagłówka RI-KSD oraz

dodanie danych związanych z bezpieczeństwem sieci. Następnie ramka przekazywana jest do drugiego klasyfikatora, gdzie następuje wyznaczenie bufora, do którego powinna ona zostać zapisana. Przed opuszczeniem węzła ramka przechodzi przez układ planowania, w którym zaimplementowano algorytm WFQ.

Ramki, które w klasyfikatorze zostały wydzielone jako przenoszące wiadomości sygnalizacyjne, przesyłane są do bloku obsługi sygnalizacji. Po określeniu, czy ramka przeznaczona jest dla danego węzła, czy też powinna zostać przesłana dalej, przesyłana jest ona do płaszczyzny sterowania do bloku realizującego styk z *NC* lub do bloku kierowania wiadomości sygnalizacyjnych od końca do końca. W bloku tym, na podstawie tabeli kierowania, następuje określenie łącza wirtualnego, którym wiadomość powinna być wysłana, po czym przekazywana jest ona ponownie do płaszczyzny danych do klasyfikatora II. Z kolei wiadomości przeznaczone dla danego węzła przez styk z *NC* przesyłane są do układu parsera/kodera, gdzie dokonywana jest interpretacja otrzymanej wiadomości. Na podstawie zdekodowanych informacji w bloku wykonawczym realizowane jest żądanie, w wyniku czego może nastąpić aktualizacja wpisów w pamięci CAM, rejestrów w układzie kształtowania ruchu oraz modyfikatorze. Innym efektem realizacji żądania może być wygenerowanie wiadomości przeznaczonej do *NC*. W takim przypadku wiadomość formowana jest w układzie parsera/kodera, następnie przesyłana jest przez styk z *NC* do układu klasyfikatora II.

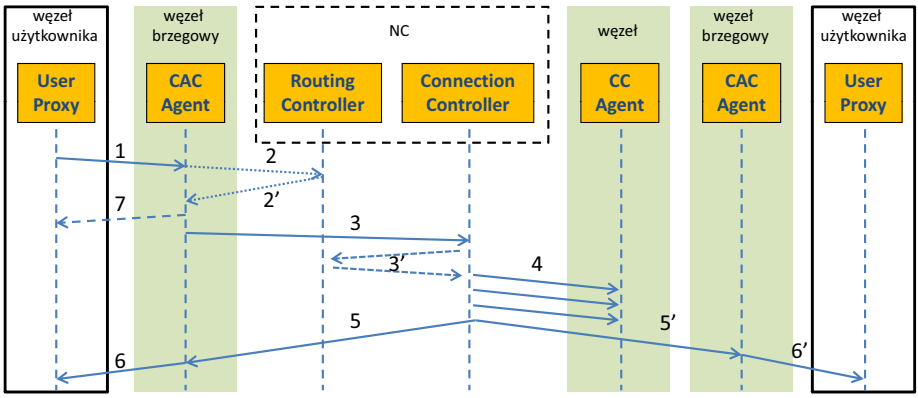


Rysunek 6. Struktura węzła RI-KSD

4 Płaszczyzna sterowania w sieci RI-KSD

Podstawowym zadaniem płaszczyzny sterowania sieci jest realizacja obsługi połączeń – tj. procedury przyjmowania połączeń CACF (ang. *Connection Admission Control Function*) oraz procedury zestawiania połączenia CCF (ang. *Con-*

nection Control Function). W realizacji tych zadań w sieci RI-KSD uczestniczą bloki *User Proxy* (implementowany w węźle użytkownika), *Agent CAC* (w węzłach brzegowych sieci) oraz *Routing Controller* i *Connection Controller* (implementowane w dedykowanym węźle sterowania/zarządzania *NC – Network Controller*).



Rysunek 7. Scenariusz realizacji funkcji CAC i CC

Zakładamy, że procedura przyjmowania połączeń (porównaj rys. 7) w sieci RI-KSD będzie realizowana w sposób zdecentralizowany, gdzie decyzję o przyjęciu/odrzuceniu połączenia podejmuje *Agent CAC* działający w węźle brzegowym sieci. Po odebraniu żądania zestawienia połączenia (wiadomość 1), agent samodzielnie, bądź we współpracy z *NC* (wiadomości 2 i 2′) dokonuje oceny możliwości jego realizacji. W przypadku odrzucenia połączenia, *Agent CAC* odsyła potwierdzenie negatywne (wiadomość 7). W przypadku przyjęcia połączenia do realizacji, agent inicjuje procedurę zestawiania połączenia (sekwencja wiadomości 3, 3′ ,4 i 5), po zakończeniu której wysyła do modułu *User Agent* potwierdzenie pozytywne (wiadomość 6).

Procedura zestawiania połączenia jest nadzorowana przez *Connection Controller* (w węźle sterowania *NC*) po otrzymaniu żądania (wiadomość 3) od agenta CAC. Jeżeli wiadomość nie definiuje ścieżki realizacji połączenia, *Connection Controller* zleca wyliczenie tej ścieżki modułowi *Routing Controller* (sekwencja wiadomości 3′). Samo połączenie jest zestawiane w wyniku wymiany wiadomości sygnalizacyjnych pomiędzy modułem *Connection Controller* a agentami sterowania połączeniem (*CC Agent*) działającymi w węzłach sieci. Pomyślne zakończenie procedury jest raportowane (wiadomości 5 i 5′) modułom CAC związanym z węzłami końcowymi połączenia.

5 Zarządzanie konfiguracją w sieci RI-KSD

System zarządzania sieci RI-KSD powinien docelowo realizować wszystkie funkcje klasycznie zdefiniowane w formie zestawu funkcji określanego jako FCAPS (failure, configuration, accounting, performance, security). W pierwszym etapie realizacji najbardziej istotne będą funkcje związane z konfiguracją ("configuration"), wspartego przez monitorowanie stanu sieci, a zarazem realizację funkcji "failure".

W zakresie zarządzania konfiguracją sieci RI-KSD realizowane będą funkcja dostarczania sieci NPF (ang. *Network Provisioning Function*), funkcja inwentaryzacji zasobów sieciowych NIF (*ang. Network Inventory Function*) oraz funkcja tworzenia sieci wirtualnych NVF (ang. *Network Virtualization Function*), przy czym dwie ostatnie z wymienionych funkcji implementowane będą w węźle sterowania/zarządzania *NC*.

Funkcja NIF realizowana jest przez węzeł sterownia *NC* we współpracy z agentami działającymi we wszystkich węzłach sieci RI-KSD. Po uruchomieniu, agent węzła sieci RI-KSD zgłasza swoją aktywność i przekazuje informacje dotyczące łączy incydentnych z jego węzłem do *NC*. Po otrzymaniu powiadomień od wszystkich węzłów sieci, *NC* jest tworzy model całej sieci. Zakładamy, że każda zmiana konfiguracji sieci RI-KSD (np. zarezerwowanie pojemności na łączu, zestawienie połączenia w polu komutacyjnym węzła) wykryta przez agenta węzła jest raportowana (w formie asynchronicznych powiadomień) do *NC*, i na podstawie tylko tych powiadomień *NC* modyfikuje zawartość modelu sieci.

Funkcja NVF ma na celu tworzenie zasobów sieci wirtualnych w sieci RI--KSD. Sieć wirtualna definiowana jest przy tym jako zestaw wirtualnych łączy (pomiędzy wybranymi węzłami brzegowymi sieci RI-KSD) oddanych do dyspozycji (sterowanie, zarządzanie) jednego operatora. Tworzenie sieci wirtualnych wymaga dokonania wyboru sekwencji łączy sieci RI-KSD przewidzianych do realizacji poszczególnych łączy wirtualnych. Podkreślamy, że istnienie systemu inwentaryzacji sieci w węźle *NC* a więc dostęp do kompletnej informacji o zasobach sieci, pozwala na zastosowanie znacznie bardziej zaawansowanych algorytmów routingu niż jest to możliwe w przypadku routingu rozproszonego.

6 Podsumowanie i przyszłe prace

W artykule przedstawiono koncepcję Równoległego Internetu bazującego na komutacji strumieni danych. Jest to jeden z Równoległych Internetów rozważany w architekturze Internetu Przyszłości w projekcie IIP. W artykule zawarto pierwsze założenia dotyczące tego Internetu wypracowane w projekcie. Podstawowym celem rozważań tego typu sieci jest zapewnienie wysokiej jakości obsługi rozumianej jako minimalizowanie opóźnień oraz minimalizowanie zmienności opóźnień dla usług czasu rzeczywistego (transmisje wideo 3D , 4K).

W projekcie przewiduje się także implementację proponowanego rozwiązania i sprawdzenie go w działaniu. Ponieważ węzły sieci RI-KSD będą wymagały specjalnych technik obsługi pakietów, przewiduje się ich implementację w układach FPGA z wykorzystaniem kart NetFPGA z układami Xilinx Virtex [17].

Literatura

1. History of the Internet, Dostępne na: http://wikipedia.org/wiki/History_of_the_Internet [Ostatni dostęp: 2011-05-28].
2. Zalecenie IETF – RFC 2474. *Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers*. 1998.
3. Zalecenie IETF – RFC 2475. *An Architecture for Differentiated Services*. 1998.
4. Zalecenie IETF – RFC 1633. *Integrated Services in the Internet Architecture: an Overview*. 1994.
5. Zalecenie IETF – RFC 1825. *Security Architecture for the Internet Protocol*. 1995.
6. Zalecenie IETF – RFC 4301. *Security Architecture for the Internet Protocol*. 2005.
7. Zalecenie IETF – RFC 2460. *Internet Protocol, Version 6 (IPv6) Specification*. 1998.
8. Decision No. 1982/2006/Ec of the European Parliament and of the Council of 18 December 2006 concerning the *Seventh Framework Programme of the European Community for research, technological development and demonstration activities (2007-2013)*. Grudzień 2006.
9. TRILOGY, Dostępne nat: http://www.trilogy-project.org [Ostatni dostęp: 2011-05-28].
10. 4WARD, Dostępne na: http://www.4ward-project.eu [Ostatni dostęp: 2011-05-28].
11. GEYSERS, Dostępne na: http://www.geysers.eu [Ostatni dostęp: 2011-05-28].
12. AKARI, Dostępne na: http://akari-project.nict.go.jp [Ostatni dostęp: 2011-05-28].
13. GENI, Dostępne na: http://www.geni.net [Ostatni dostęp: 2011-05-28].
14. IIP System, Dostępne na: http://www.iip.net.pl [Ostatni dostęp: 2011-05-28].
15. W. Burakowski, H. Tarasiuk, and A. Bęben, "Future Internet architecture based on virtualization and co-existance of different data and control planes," *IEEE GLOBECOM 2011*, (Houston, TX), zgłoszone do publikacji.
16. W. Burakowski, A. Beben, H.Tarasiuk, J. Sliwinski, R. Janowski, J. Mongay Batalla, and P. Krawiec, "Provision of end-to-end qos in heterogeneous multi-domain networks", *Annals of Telecommunications - Annales des Télécommunications, Springer Eds.*, vol. 63, nr 11, str. 559, 2008.
17. NetFPGA, Dostępne na: http://www.netfpga.org [Ostatni dostęp: 2011-05-28].

PL-LAB: sieć eksperymentalna projektu Inżynieria Internetu Przyszłości

Jarosław Śliwiński[1,2],
Radosław Krzywania[3], Łukasz Dolata[3]

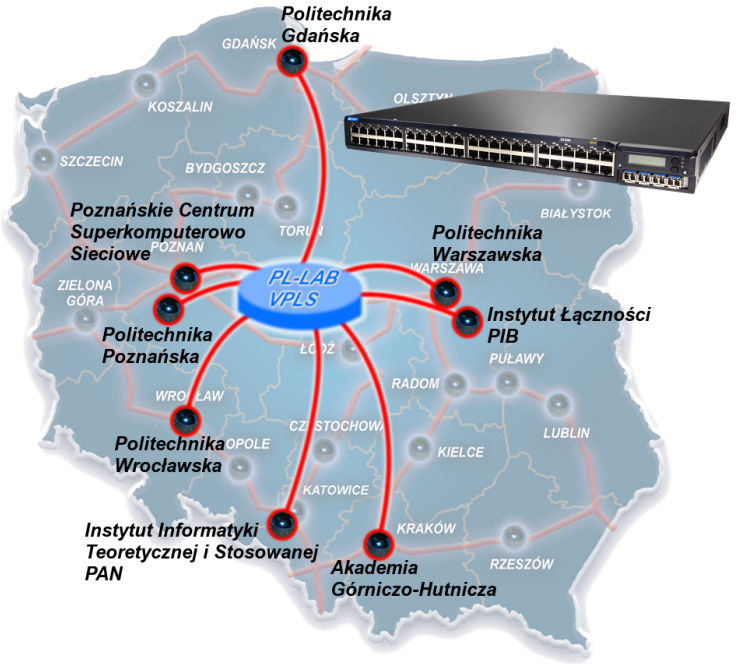
[1] Instytut Telekomunikacji, Politechnika Warszawska
[2] Instytut Łączności – Państwowy Instytut Badawczy
[3] Poznańskie Centrum Superkomputerowo-Sieciowe

Streszczenie PL-LAB to sieć eksperymentalna projektu Inżynieria Internetu Przyszłości. Sieć ta została utworzona z połączenia 8 laboratoriów z wykorzystaniem ogólnopolskiej sieci PIONIER. Artykuł ten przedstawia strukturę sieci PL-LAB oraz podstawowe narzędzia wykorzystane do zbudowania tej sieci: część operacyjną, część badawczą oraz system dostępu.

1 Wstęp

Obszar badań wokół tematyki Internetu Przyszłości (ang. *Future Internet*) obejmuje dwa główne nurty: rewolucyjny i ewolucyjny. Nurt rewolucyjny zakłada znaczące zmiany w działaniu sieci Internet, które w praktyce wymagają istotnych, a zatem kosztownych, zmian w infrastrukturze sieci. W zamian oczekuje się istotnych korzyści związanych ze zwiększoną wydajnością, uproszczonym zarządzaniem siecią, lepszym wykorzystaniem zasobów, itp. Z kolei nurt ewolucyjny skupia się na wyszukaniu elementów obecnej architektury sieci, których funkcjonalność można rozszerzyć, zmodyfikować lub zastąpić. Potencjalny zasięg korzyści jest ograniczony w porównaniu z nurtem rewolucyjnym, jednak z drugiej strony pozwala na zachowanie obecnie działających usług i aplikacji. Zagadnienia poruszane w projekcie Inżynieria Internetu Przyszłości (akronim IIP) reprezentują oba te nurty: 1) poprzez tworzenie sieci z nowymi zasadami przesyłania danych i zasadami zarządzania, oraz 2) poprzez rozszerzanie możliwości istniejących sieci IP, w tym sieci IPv6.

Jednym z kluczowych elementów w badaniach nad Internetem Przyszłości są sieci eksperymentalne, które umożliwiają praktyczne testowanie proponowanych rozwiązań, np. [1], [2] i [3]. Sieć eksperymentalna projektu IIP, nazwana **PL-LAB**, łączy 8 laboratoriów, w których znajduje się sprzęt umożlwiający badania z szerokiego zakresu: od aplikacji aż do programowalnych układów obsługujących przekaz danych. Artykuł ten opisuje sieć PL-LAB i został podzielony na następujące części. Cześć druga opisuje infrastrukturę sieci z uwzględnieniem połączeń poprzez sieć PIONIER. Część trzecia jest poświęcona systemowi dostępu, poprzez który użytkownicy tworzą eksperymenty. Artykuł podsumowuje część poświęcona wnioskom.



Rysunek 1. Rozmieszczenie laboratoriów sieci PL-LAB.

2 Infrastruktura sieci

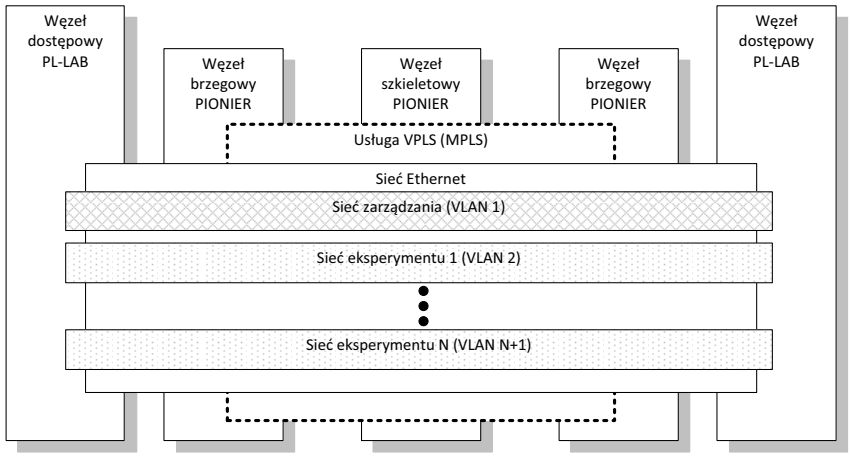
Sieć PL-LAB obejmuje 8 laboratoriów w Polsce, które znajdują się w ośrodkach akademickich i instytutach badawczych zajmujących się badaniami w obszarze Internetu Przyszłości. Rysunek 1 przedstawia rozmieszczenie geograficzne laboratoriów na mapie Polski.

Infrastruktura sieci PL-LAB została podzielona na 2 części:

- Operacyjną, która zapewnia podstawowe połączenia pomiędzy laboratoriami. Połączenia te muszą być skonfigurowane dla działania w długiej skali czasu i często wykorzystują urządzenia, do których administratorzy sieci PL LAB nie mają bezpośredniego dostępu.
- Badawczą, która obejmuje pozostałe urządzenia w laboratoriach. Administratorzy sieci PL LAB mają kontrolę nad konfiguracją tych urządzeń.

2.1 Część operacyjna

Część operacyjna sieci PL LAB jest związana z ogólnopolską siecią PIONIER, która umożliwia przesyłanie danych pomiędzy laboratoriami. Każde laboratorium jest podłączone do węzła sieci PIONIER za pomocą łącza Gigabit Ethernet (lub większej liczby łączy w zależności od lokalnych uwarunkowań).



Rysunek 2. Techniki sieciowe w połączeniu laboratoriów poprzez sieć PIONIER.

Łącze dostępowe od strony laboratorium jest zakończone na przełączniku dostępowym (Juniper EX3200), natomiast od strony sieci PIONIER na przełączniku brzegowym (przełączniki Brocade serii MLX). Przełącznik brzegowy odbiera ramki wysłane z laboratorium i opakowuje je w etykiety MPLS usługi VPLS (ang. *Virtual Private LAN Service*), która przesyła je do celu (czyli innego laboratorium). Usługa VPLS jest kontrolowana przez sieć PIONIER i sieć PL-LAB nie ma możliwości sterowania parametrami tej usługi (ścieżki, dodatkowe etykiety, itp.). Aby przyporządkować poszczególne ramki Ethernet do aktualnie działających eksperymentów wykorzystano mechanizm sieci VLAN, który jest zarządzany w przełączniku dostępowym. Zastosowanie oznaczania etykietami VLAN jest związane z zastosowaniem poszczególnych sieci:

- Sieć zarządzająca, która służy zarządzaniu siecią PL-LAB. Sieć ta korzysta z pojedynczej etykiety VLAN i umożliwia dostęp do wszystkich urządzeń laboratoryjnych pozwalających na zdalne zarządzanie. Konfiguracja tej sieci jest statyczna.
- Sieci poszczególnych eksperymentów, które są specjalnie stworzone na potrzeby konkretnego scenariusza badawczego. Konfiguracja takich sieci może zmieniać się w czasie w zależności od zapotrzebowania.
- Sieć do robienia kopii zapasowych, która umożliwia kopiowanie obrazów dysków dla systemów wirtualnych oraz kopiowanie wyników eksperymentów na urządzenia składowania danych.

Rysunek 2 obrazuje zastosowanie poszczególnych technik sieciowych w połączeniu laboratoriów.

2.2 Część badawcza

Część badawcza sieci PL-LAB obejmuje kilka rodzajów urządzeń jednak ich wspólną cechą jest dołączenie do węzła dostępowego PL-LAB za pomocą techniki Ethernet.

Pierwszą grupą urządzeń są programowalne urządzenia komutacji pakietów, które opierają się na dwóch technologiach: układach FPGA (np. karta NetFPGA) i programowalnych procesorach sieciowych (np. urządzenie EZapplicance). Pozwalają one w szerokim stopniu zmieniać działanie węzłów, w tym umożliwiają zastosowanie nowych reguł obsługi strumieni danych. W porównaniu z programowalnymi procesorami sieciowymi, układy FPGA oferują dużo większe możliwości modyfikacji w tym stworzenia rozwiązania całkowicie "od zera". Oczywiście kosztem tego są dużo większe nakłady pracy, gdyż w procesorach sieciowych część rozwiązań jest gotowa, np. mechanizm szeregowania ramek dla podziału pasma na łączach.

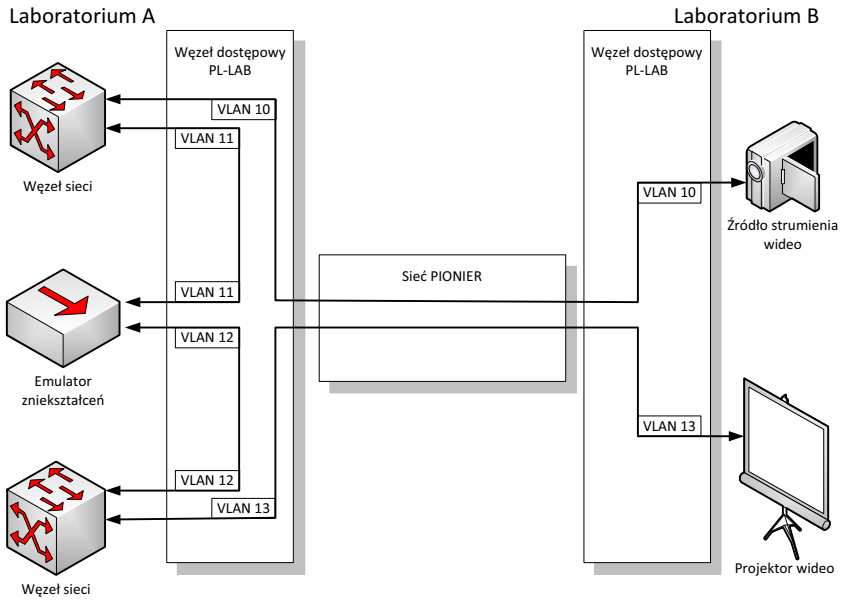
Kolejną grupą urządzeń są serwery wspierające wirtualizację, które umożliwiają badania nad technikami wirtualizacji systemów operacyjnych. Rozwiązania te składają się z części sprzętowej, wyposażonej w procesory (CPU) dostosowane do wirtualizacji, oraz z systemów zarządzania maszynami wirtualnymi (ang. *hypervisor*). Ponadto serwery te posiadają kilka kart sieciowych, tak aby mogły spełniać rolę węzłów sieci.

W części badawczej znajdują się również źródła ruchu, wśród których wyróżniono źródła związane z aplikacjami oraz źródła związane z narzędziami pomiarowymi. W przypadku aplikacji zakres dostępnych źródeł jest szeroki: od sensorów, poprzez wideokonferencję i wideo HD/3D, aż do wideo w bardzo wysokiej rozdzielczości 4K, tj. strumienie wideo o rozdzielczości sięgającej 4096x3112. Z kolei narzędzia pomiarowe umożliwiają wysyłanie i analizę ruchu zgodnie z założonymi profilami (dla zamkniętych narzędzi sprzętowych) lub ruchu dowolnego typu (dla otwartych narzędzi programowalnych).

Rysunek 3 przedstawia przykładowy eksperyment stworzony z połączenia zasobów części operacyjnej i części badawczej PL-LAB. W przykładzie tym laboratorium A posiada specjalizowane węzły sieciowe oraz emulator zniekształceń przechodzącego ruchu, natomiast laboratorium B posiada źródło wideo oraz system prezentacji. Część operacyjna PL-LAB pozwala stworzyć wirtualną topologię, która rozszerza możliwości badań poza pojedyncze laboratorium. Laboratorium A może wykorzystać źródło wideo o rzeczywistym profilu ruchu znajdujące się innym miejscu, natomiast laboratorium B ma możliwość zbadania wpływu zniekształceń strumienia po przejściu przez sieć, korzystając z urządzeń laboratorium A.

3 System dostępu

Sieć PL-LAB umożliwia tworzenie eksperymentów łącząc zasoby badawcze za pomocą węzłów dostępowych. Jednak aby w pełni skorzystać z tej funkcjonalności potrzebne jest narzędzie do określenia wymagań użytkownika dotyczących

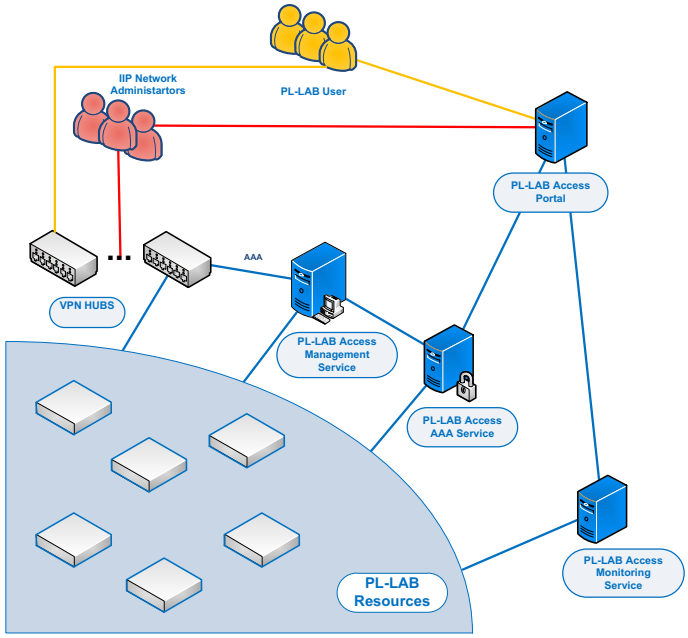


Rysunek 3. Przykład eksperymentu w sieci eksperymentalnej PL-LAB.

samego eksperymentu. Tą rolę pełni system dostępu do sieci PL-LAB, którego główne elementy przedstawiono na rysunku 4.

Elementem systemu, który jest widziany przez użytkowników, jest **portal**, czyli strona internetowa pozwalająca zarządzać siecią PL-LAB. Użytkowników podzielono na dwie podstawowe grupy: 1) tworzących eksperymenty (nazwanych **użytkownikami**) i 2) zarządzających siecią PL-LAB (**administratorów**). Pierwsza grupa obejmuje większość użytkowników systemu dostępu, natomiast w drugiej wyróżniono specjalizacje związane z zakresem obowiązków i uprawnień (np. opiekun laboratorium).

Użytkownicy PL-LAB dołączają się do sieci PL-LAB za pomocą dedykowanych **węzłów VPN** (ang. *Virtual Private Network*). Każdy węzeł VPN jest bezpośrednio podłączony do tzw. sieci zarządzania, czyli sieci umożliwiającej dostęp do interfejsów zarządzania w każdym urządzeniu. Aby zabezpieczyć eksperymenty tworzone przez użytkowników **usługa AAA** kontroluje dostęp do urządzeń wykorzystując serwer protokołu RADIUS. Z kolei **usługa zarządzania** przeprowadza konfigurację urządzeń i topologii sieci dla potrzeb każdego eksperymentu (automatycznie lub pół-automatycznie). W przypadku gdy użytkownik korzysta z opcji tworzenia kopii zapasowych, system korzysta z **repozytorium kopii zapasowych**. Ostatnim elementem jest usługa monitorowania, która obserwuje i archiwizuje stan urządzeń i połączeń pomiędzy laboratoriami.



Rysunek 4. Elementy systemu dostępu do sieci PL-LAB.

W czasie przygotowywania eksperymentu użytkownik musi określić następujące parametry:

- które urządzenia chce użyć (rodzaj, liczba, itp.),
- jaką powinna wyglądać topologia sieci pomiędzy urządzeniami,
- czy wymagane jest tworzenie kopii zapasowych konfiguracji urządzeń po zakończeniu okresu eksperymentu,
- datę rozpoczęcia oraz czasu trwania eksperymentu.

Po potwierdzeniu, że zasoby są dostępne w określonym czasie, opis eksperymentu jest przesyłany do administratorów sieci PL-LAB. Eksperyment jest ostatecznie zatwierdzany a administratorzy przystępują do konfiguracji części operacyjnej i badawczej.

Użytkownik po zakończeniu badań ma możliwość utworzenia kopii konfiguracji urządzeń, które brały udział w eksperymencie. Pozwala to również na archiwizację wyników eksperymentów, jednak należy uwzględnić, że polityka przechowywania kopii może być ograniczona czasowo.

4 Podsumowanie

Sieć eksperymentalna PL-LAB jest istotnym fragmentem projektu Inżynieria Internetu Przyszłości. Jest ściśle powiązana z urządzeniami wykorzystywanymi

w pracach badawczych, zarówno od strony sieci jak i zastosowań w aplikacjach. Wprowadziło to szereg wymagań, np. dostęp do fizycznych urządzeń, które zostały uwzględnione przy projektowaniu systemu dostępu do sieci PL-LAB. Z drugiej strony możliwość poprzez wykorzystanie ogólnopolskiej sieci PIONIER, sieć PL-LAB łączy laboratoria specjalizujące się w różnych obszarach badań związanych Internetem Przyszłości. W wyniku tego w kraju powstała infrastruktura znacząco rozszerzająca możliwości eksperymentalne w tej dziedzinie. Dodatkowo sieć PL LAB przewiduje rozszerzenie dostępu do swoich zasobów dla innych organizacji z Polski, oraz połączenie z podobnymi inicjatywami na Świecie.

Literatura

1. Projekt FEDERICA, `http://www.fp7-federica.eu/`
2. Projekt OFELIA, `http://www.fp7-ofelia.eu/`
3. Projekt SMART SANTANDER, `http://www.smartsantander.eu/`

Aplikacje w projekcie Inżynieria Internetu Przyszłości

Adam Grzech[1], Krzysztof Juszczyszyn[1], Paweł Świątek[1], Cezary Mazurek[2], Arkadiusz Sochan[3]

[1] Instytut Informatyki, Politechnika Wrocławska
[2] Poznańskie Centrum Superkomputerowo-Sieciowe
[3] Instytut Informatyki Teoretycznej i Stosowanej PAN w Gliwicach

Streszczenie W artykule przedstawiono cztery grupy aplikacji opracowywanych w ramach projektu Inżynieria Internetu Przyszłości (IIP) charakterystycznych dla aktualnych i perspektywicznych zastosowań Internetu: sieci otoczenia człowieka, ochrony zdrowia, edukacji i dostarczania treści. Opracowywane aplikacje będą wykorzystywać własności trzech Równoległych Internetów wykorzystujących różne, pracujące równolegle i korzystające z różnych zasobów komunikacyjnych, płaszczyzny sterowania i danych do realizacji różnych usług udostępnianych w sieciach wirtualnych przez różnych operatorów.

1 Wprowadzenie

Technologie informacyjne i komunikacyjne umożliwiają coraz tańsze, elastyczniejsze i szybsze gromadzenie, przechowywanie, przekazywanie, przetwarzanie i prezentowanie danych umożliwiając i powodując coraz większą gęstość połączeń w systemach społecznych i gospodarczych. Konsekwencją wspomnianych zmian w technologiach są masowe zastosowania Internetu jako narzędzia dostępu do usług, co niesie za sobą zarówno wiele różnych, nowych możliwości, ale także jest źródłem wielu nowych wyzwań z zakresu skalowalności, pojemności, przepływności, mobilności, bezpieczeństwa, itd. rozwiązań sieciowych.

Warunkiem realizacji koncepcji „społeczeństwa informacyjnego" i „gospodarki cyfrowej" jest jednoczesne użycie wielu różnych technologii; od wszechobecnych sieciowych systemów sieciowych i obliczeniowych poprzez różnego rodzaju sieci czujników do nowych metod współdziałania użytkowników sieci. Internet przyszłości jest powszechnie rozumiany jako platforma umożliwiająca integrację aplikacji wspierających wszystkie możliwe aktywności społeczne i gospodarcze. Dotyczy to w szczególności platformy integracji otwartych, elastycznych i skalowalnych systemów sieciowych dla potrzeb wspomagania działalności gospodarczej (globalizacja rynku towarów i usług), zarządzania wykorzystaniem zasobów (surowce naturalne i ich wykorzystanie, ograniczenie emisji zanieczyszczeń), wspomaganiem systemów podejmowania decyzji w systemach ochrony zdrowia (ograniczenie liczby i kosztów konsekwencji wypadków komunikacyjnych), zapewnieniem bezpieczeństwa w sferze prywatnej i publicznej (komfort życia) oraz

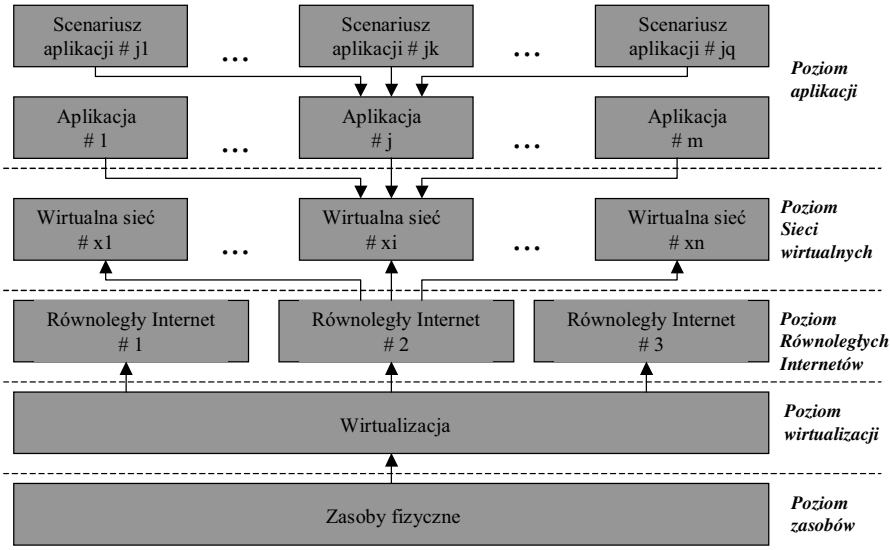
nieskrępowanym dostępem do treści (edukacja, zachowanie dziedzictwa narodowego) [4].

Internet Przyszłości jest definiowany jako dynamiczna, globalna infrastruktura wykorzystująca standardowe protokoły współpracy z wbudowanymi mechanizmami samo-konfiguracji, w której rzeczywiste oraz wirtualne usługi i rzeczy posiadają tożsamość, fizyczne atrybuty, wirtualne osobowości, wykorzystują inteligentne interfejsy i mogą być bezproblemowo integrowane w sieci informacyjnej. Wizja Internetu Przyszłości bazującej na standardowych protokołach komunikacji zakłada połączenie sieci komputerowych, Internetu Mediów (*Internet of Media*), Internetu Usług (*Internet of Services*) i Internetu Rzeczy (*Internet of Things*) w jedną globalną platformę sieci i rzeczy *(networked „things"*). Oczekiwane jest, że infrastruktura Internetu Przyszości integrującego istniejące i nowe sieci będzie dynamicznie rozszerzaną i rozbudowywaną o usługi terminali (rzeczy) wzajemnie połączonych infrastrukturą publiczno-prywatną. Przewidywane jest, że Internet Rzeczy - gdzie „rzecz" jest rozumiana jako identyfikowalny (poprzez nazwę i/lub adres) zmienny w czasie i przestrzeni podmiot fizyczny lub wirtualny - pozwoli na komunikację zawsze i wszędzie (*anytime, anyplace*) wszystkiego ze wszystkim (*anything, anyone*) korzystając z dowolnych sieci (*any path/network*) i ich usług (*any service*). Koncepcja Internetu Rzeczy jest postrzegana jako koncepcja rozszerzenia funkcjonalności znanych sposobów interakcji pomiędzy człowiekiem i aplikacją oraz między aplikacja a infrastrukturą sieciowo-komunikacyjną [3].

2 Aplikacje w projekcie Inżynieria Internetu Przyszłości

Aplikacje, planowane do wdrożenia w projekcie Inżynieria Internetu Przyszłości, zostały zaprojektowane i będą najpierw testowane, a następnie używane w trzech Równoległych Internetach (RI) w architekturze Systemu IIP: RI IPv6 QoS (bazujący na IPv6), RI-CAN (bazujący na koncepcji sieci świadomej przekazywanej treści - Content Aware Network) i RI-KSD (Komutacja Strumieni Danych bazująca na koncepcji zbliżonej do komutacji kanałów z przeznaczeniem dla usług czasu rzeczywistego). Zaproponowany w projekcie model systemu zawiera cztery poziomy [1, 5]. Poziom pierwszy to zasoby fizyczne, poziom drugi to wirtualizacja z wykorzystaniem której na poziomie trzecim organizowane są sieci wirtualne (Równoległe Internety), w każdej z których możliwe jest tworzenie sieci wirtualnych (poziom czwarty) udostępnianych operatorom usług poszczególnych aplikacji, scenariuszy poszczególnych aplikacji lub grupom aplikacji i ich scenariuszy (Rysunek 1).

Aplikacje, przygotowywane w projekcie Inżynieria Internetu Przyszłości, są aplikacjami dziedzinowymi i dotyczą usług charakterystycznych dla sieci domowych i samochodowych, sieci e-zdrowie, systemów dostarczania treści oraz systemów kształcenia na odległość. Aplikacje te są aplikacjami odpowiadającymi wybranym procesom biznesowym, dla obsługi (wsparcia) których integrowane są różne, ale wspólne dla tych aplikacji, techniki prezentacji, przetwarzania i przesyłania danych, informacji i wiedzy.



Rysunek 1. Ogólna architektura systemu Inżynieria Internetu Przyszłości (IIP).

W projekcie Inżynieria Internetu Przyszłości (IIP) wyróżniono cztery grupy aplikacji: sieci domowe i samochodowe (sieci otoczenia człowieka), sieci e-zdrowie, Internet 3D, kino cyfrowe i UHD oraz sieci edukacyjne i społecznościowe [2–4].

Wymienione grupy aplikacji, związane są z różnymi sposobami gromadzenia, przekazywania, przetwarzania, przechowywania i prezentowania różnych klas (podklas) ruchu strumieniowego i elastycznego (sekwencyjnie i/lub równolegle), charakteryzują się różnymi wymaganiami dotyczącymi ilości i jakości usług komunikacyjnych (a dalej transmisyjnych) oraz jakości ich postrzegania. Obsługa ruchu generowanego przez wymienione dziedzinowe aplikacje może być realizowana przez różne rozwiązania sieciowe.

Dla potrzeb uwzględnienia różnych sposobów używania funkcjonalności aplikacji oraz testowania funkcjonalności Równoległych Internetów opracowano różne scenariusze użycia aplikacji, które różnią się w szczególności możliwością wykorzystania własności trzech Równoległych Internetów oraz zapotrzebowaniem na ilość i jakość usług sieci.

2.1 Sieci domowe i samochodowe

Liczba urządzeń, możliwych do identyfikacji oraz mogących się komunikować w otoczeniu człowieka rośnie wraz ze wzrostem roli jaką nowoczesne technologie spełniają w życiu codziennym. Techniki łączenia takich urządzeń migrują od prostych połączeń punkt-punkt do bardziej złożonych topologii sieciowych pozwa-

lając na implementację złożonych rozwiązań teleinformatycznych w zadaniach monitorowania aktywności człowieka (np. lokalizacja, parametry funkcji życiowych, itp.) i jego otoczenia (stan techniczny używanych urządzeń technicznych, sposoby ich użytkowania, itp.) gromadzone i przechowywane w takich urządzeniach jak telefon komórkowy, komputer osobisty, kamera, sensory itp. Efektywne wykorzystanie możliwości wzrastającej liczby heterogenicznych urządzeń dla potrzeb monitorowania, zarządzania i sterowania wymaga implementacji uniwersalnych infrastruktur komunikacyjnych uwzględniających interoperacyjność tych urządzeń i sieci w otoczeniu człowieka, pozwalających na dostarczenie transparentnych usług komunikacyjnych, przetwarzania danych i sterowania. Uniwersalność infrastruktur wymaga implementacji procedur i algorytmów umożliwiających jednocześnie zarządzanie pasmem, dostępem do sieci zawsze i wszędzie, dostępem do wszystkich urządzeń, jakością usług, zapewnieniem bezpieczeństwa, niezawodnością i efektywnym wykorzystaniem energii.

Usługi sieci otoczenia człowieka (w tym sieci domowych i samochodowych) wymagają infrastruktury teleinformatycznej pozwalającej na gromadzenie danych z wielu różnych, heterogenicznych źródeł danych, przekazywanie i przetwarzanie wolno i szybkozmiennych strumieni ruchu teleinformatycznego, różnicowanie sposobu komunikacji w zależności od szybkozmiennych wymagań użytkowników aplikacji, integrację funkcji zarządzania i sterowania, komunikacji świadomej dostarczanych treści i kontekst jej użycia za pomocą prostych i intuicyjnych interfejsów, personalizację użytkowników i bezpieczeństwo usług, możliwość wyboru spośród wielu dostawców treści i usług, otwartość na nowe urządzenia z wykorzystaniem wirtualnej rzeczywistości i sztucznej inteligencji.

W grupie sieci domowe i samochodowe przygotowywane są trzy aplikacje:

HomeNetMedia - spersonalizowany dostęp do treści multimedialnych pochodzących z zasobów domowego repozytorium i zasobów dostawców treści przez Internet oraz odtworzenie treści na wybranym urządzeniu końcowych podłączonym do sieci domowej,

HomeNetEnergy - monitorowanie zużycia i inteligentne zarządzanie zużyciem energii elektrycznej w środowisku domowym z wykorzystaniem informacji kontekstowej i profili użytkowników,

MobiWatch - integracja rozproszonych systemów bezpieczeństwa i monitorowania z danymi z systemu dozoru wizyjnego oraz sieci sensorowych w zorientowany usługowo sposób.

2.2 Sieci e-zdrowie

Wykorzystywanie systemów informatycznych i telekomunikacyjnych w opiece zdrowotnej związane jest interdyscyplinarną dziedziną jaką jest telemedycyna, której celem jest projektowanie efektywnych systemów (wsadowych i czasu rzeczywistego) zdalnego monitorowania i diagnozowania stanu pacjentów, terapii oraz wspomaganie podejmowania decyzji wykorzystujących różne struktury informacji z wykorzystaniem metod sztucznej inteligencji.

Systemy teleinformatyczne e-zdrowia posiadają wszystkie cechy mobilnych, heterogenicznych sieci sensorowych zintegrowanych z usługami niezawodnych,

bezpiecznych i trwałych infrastruktur komunikacyjnych, za pomocą których dostępne są usługi: baz danych pacjentów, chorób i historii chorób oraz różnych systemów informatycznych wspomagania podejmowania decyzji obejmujących szeroki zakres różnych rozwiązań informatycznych: od systemów przetwarzania wielowymiarowych struktur danych do multimedialnych systemów konsultacji w czasie rzeczywistym.

W zadaniach projektowania systemów monitorowania, diagnostyki i terapii w systemach telemedycznych obiecującym jest podejście, zgodnie z którym implementowane w systemie teleinformatycznym funkcjonalności imitują sprawdzone, wymagane i często standardowe praktyki i procedury medyczne. Warunkiem efektywności systemów telemedycznych jest zakres i możliwość personalizacji automatyzowanych procedur, profilowania usług, akwizycji informacji i wiedzy oraz automatycznego wnioskowania.

Usługi sieci e-zdrowie wymagają dostępu do infrastruktury teleinformatycznej pozwalającej na integrację usług systemów zdalnego monitorowania, diagnostyki, terapii i sterowania, różnicowanie jakości usług w zależności od stanu usługobiorcy, dostępu do usług wrażliwych na dostarczaną treść, kontekst jej wykorzystania i bezpieczeństwo danych, równoległego korzystania z wielu różnych usług z możliwością wywłaszczania jednych przez drugie, adaptacji przepustowości sieci w dużym zakresie zmienności w zależności od obsługiwanego scenariusza monitorowania, diagnostyki lub terapii, odporności na uszkodzenia w krytycznych scenariuszach dostarczania usług, itp.

W grupie sieci e-zdrowie przygotowywanych jest pięć aplikacji:

Universal Communication Platform - specyfikacja i opis protokołów sygnalizacyjnych dla potrzeb komponowania złożonych usług e-zdrowie, negocjacji wymaganych wartości parametrów QoS i QoE oraz alokacji zasobów komunikacyjnych i obliczeniowych,

e-Asthma - integracja systemów monitorowania pacjentów chorych na astmę oraz środowiska naturalnego chorych na astmę dla potrzeb wspomagania diagnozowania,

e-Diab - wspomaganie podejmowania decyzji w zarządzaniu i przetwarzaniu danych wykorzystywanych w procesie monitorowania, diagnozowania i terapii chorych na cukrzycę,

SmartFit - akwizycja danych i wspomaganie decyzji w treningu sportowym zgodnie ze standardami obowiązującymi w sporcie wyczynowym,

Private Family e-Health Network - implementacja medycznej biblioteki cyfrowej w środowisku rozproszonym.

2.3 Internet 3D, kino cyfrowe, UHD

Kolejne, pojawiające się technologie rejestracji i przetwarzania obrazów wideo 2D (HD, UHD) oraz rosnące zapotrzebowanie na dostęp do treści z wykorzystaniem tych technologii wymagają zwiększenia przepustowości sieci oraz skalowalności rozwiązań sieciowych.

Wzrost dostępności urządzeń do akwizycji i prezentacji 3D powoduje wzrost zapotrzebowania na rozwiązania pozwalające efektywnie składować i przesyłać

informacje o przestrzeniach i obiektach 3D, które - w odróżnieniu od formatów projektowanych przede wszystkim do zapisu pojedynczych obiektów o regularnych kształtach lub ich niewielkich grup - będą przystosowane do efektywnego zapisu informacji o obiektach pochodzących np. ze skanowania 3D, przystosowane do progresywnej prezentacji informacji i zapewnią przenaszalność informacji pomiędzy poszczególnymi formatami.

Dla potrzeb dostarczania treści, rejestrowanych i przetwarzanych w wykorzystaniem nowych technologii, potrzebne są nowe lub zmodernizowane formaty danych 3D oraz protokoły sieciowe do ich przekazu. Konieczność uwzględniania progresywnego przesyłu i prezentacji informacji, w tym także poddanej kompresji stratnej, wymaga nowych procedur kodowania, transmisji i prezentacji pozwalających na uzyskiwanie wymaganej przez użytkownika jakości.

Usługi systemów dostarczania treści (w zadaniach kształcenia, rozrywki, ochrony zdrowia itp.) są możliwe pod warunkiem istnienia infrastruktury teleinformatycznej pozwalającej na efektywne wyszukiwanie treści i jej efektywny przekaz, integrację różnych strumieni ruchu w czasie rzeczywistym, efektywne i adaptacyjne zwielokrotnianie i rozmieszczanie treści dla potrzeb jej efektywnego wyszukiwania i dostarczania, personalizowanie dostarczanych treści, dopasowywania dostarczanej treści do urządzeń końcowych użytkownika, itd.

W omawianej grupie przygotowywane są trzy aplikacje:

QoE Assessment for 3D Content - sekwencje wzorcowe materiałów 3D do oceny jakości przekazu treści multimedialnych,

4K and 3D Video Streaming - akwizycji materiałów wideo 3D oraz prezentacji materiałów wideo 3D i 4K,

Interactive Delivery of 3D Content (Virtual Museum) - akwizycja obiektów, przygotowanie sceny i ekspozycja oraz progresywny przesył obiektów 3D.

2.4 Sieci edukacyjne i społecznościowe

Rosnące możliwości komunikacji oraz dostępu do treści w sieci pozwalają na efektywne wykorzystywanie zaawansowanych mechanizmów komunikacyjnych dostępnych w architekturach internetowych. Mechanizmy te mogą być wykorzystane zarówno dla potrzeb efektywnej realizacji znanych, tradycyjnych scenariuszy nauczania (interakcja nauczający - nauczany), jak i budowanie nowych scenariuszy nauczania, w których wspomniana interakcja jest uzupełniana możliwościami dostępu w czasie rzeczywistym do źródeł treści (wirtualne laboratorium) oraz różnych sposobów kontekstowego przetwarzania i prezentowania treści.

Poszerzone scenariusze nauczania, zakładające dostęp do treści wymagających znacznych zasobów obliczeniowych i komunikacyjnych (obliczenia rozproszone, video HD, jednoczesna transmisja do użytkowników w wielu lokalizacjach) wymagają zapewnienia odpowiedniej jakości usług oraz skutecznych mechanizmów rezerwacji zasobów.

Efektywność sieciowych systemów nauczania zależy zatem także od implementowanych w takich systemach metod modelowania i predykcji zachowań użytkowników, wykorzystujące teorie sieci społecznych oraz inżynierię wiedzy

(w przypadku semantycznego opisu zasobów, usług i użytkowników). Zagadnienia te są bezpośrednio związane z problemem personalizacji dostępu do usług oraz zapewnieniem ich jakości z uwzględnieniem indywidualnych charakterystyk użytkownika, usług i sposobów ich użytkowania w otwartych systemach sieciowych. W systemach takich aktywność użytkowników, wyrażona wzorcami określającymi sposób korzystania z usług systemu, determinuje w wielu przypadkach jakość oraz dostępność świadczonych usług. Wzorce takie określane są często jako tzw. zjawiska emergentne - charakterystyczne dla działającego systemu i nie dające się wywieść z założeń dotyczących projektu i architektury. Przyczyną ich pojawiania się jest istnienie złożonych struktur (społecznych, ekonomicznych, funkcjonalnych) wpływających na funkcjonowanie systemu sieciowego i determinujących zachowanie jego użytkowników. Do najbardziej istotnych należą sieci społeczne. Standardowa definicja sieci społecznej, definiująca ją jako graf składający się ze zbioru węzłów (aktorów) oraz łączących ich relacji społecznych. Teoria sieci społecznych oferuje szereg metod pozwalających na określenie istotnych parametrów sieci oraz charakterystykę aktywności użytkowników i ich społeczności.

Obecnie do najważniejszych potrzeb związanych z nowoczesnymi systemami zdalnego nauczania należą: zapewnienie jakości procesu nauczania z wykorzystaniem komunikacji sieciowej oraz zapewnienie jakości usług w przypadku wykorzystania w nauczaniu mediów strumieniowych oraz aplikacji pozwalających na dostęp do zasobów obliczeniowych współczesnych systemów sieciowych.

W grupie systemów wspierających proces nauczania przygotowywane są trzy aplikacje:

Computer Aided Assessment - system weryfikacji wiedzy oraz oceny jakości procesów zdalnego nauczania,

EduVideoHD - usługa wideokonferencji dla potrzeb systemów zdalnego kształcenia,

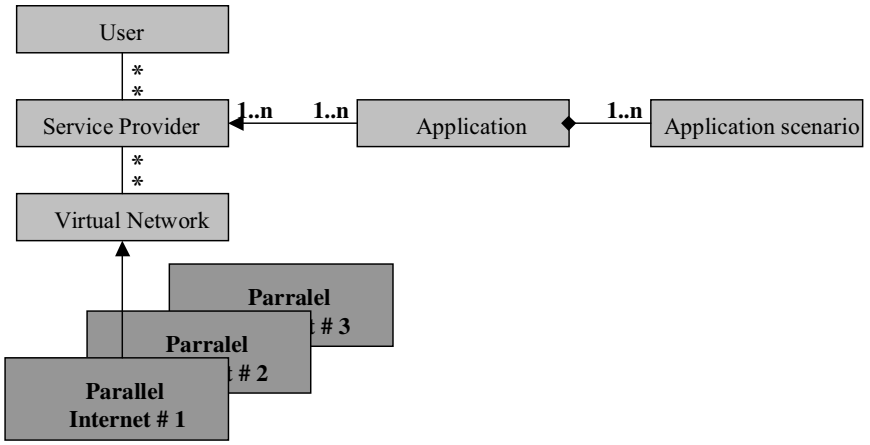
OnLineLab - wirtualne laboratorium obliczeniowe.

3 Modele biznesowe aplikacji w systemie IIP

Ogólna architektura systemu IIP (Rysunek 1) daje wiele potencjalnych możliwości testowania aplikacji. Zakres możliwych scenariuszy testowania aplikacji wyznaczony jest z jednej strony przez najprostszy model biznesowy (jeden użytkownik - lub grupa użytkowników - korzysta z usług jednego scenariusza jednej aplikacji, dostarczanej przez jednego dostawcę usług w jednej sieci wirtualnej w jednym z Równoległych Internetów), a z drugiej strony przez złożony model biznesowy (jeden użytkownik korzysta z usług różnych scenariuszy różnych aplikacji dostarczanych przez różnych dostawców usług, w różnych sieciach wirtualnych utworzonych w różnych Równoległych Internetach) [6].

Zbiór możliwych modeli biznesowych udostępniania użytkownikom usług aplikacji w ich różnych scenariuszach przez dostawców usług korzystających z sieci wirtualnych Równoległych Internetów IIP ilustruje Rysunek 2.

Dla potrzeb testowania aplikacji wykorzystujących właściwości RI IPv6 QoS przeprowadzono analizę możliwych modeli biznesowych. Do testowania proto-



Rysunek 2. Możliwe warianty użycia scenariuszy aplikacji w Równoległych Internetach.

typu RI IPv6 QoS wybrany został prosty model biznesowy, zgodnie z którym grupa użytkowników korzysta z kilku scenariuszy jednej aplikacji dostarczanych przez jednego operatora usług aplikacji wykorzystującego jedną sieć wirtualną w jednym Równoległym Internecie.

Przyjęcie takiego modelu biznesowego oznacza, że do wymiarowania sieci wirtualnej dla operatora usług aplikacji (poziom czwarty modelu) wymagana jest znajomość a priori sumarycznego zapotrzebowania na wyróżnione klasy usług RI IPv6 QoS. Losowość żądań dostępu do usług aplikacji w jej różnych scenariuszach powoduje, że możliwe do zastosowania są różne strategie wymiarowania sieci wirtualnych zależnie od wiedzy o przewidywanych i prognozowanych zachowaniach użytkowników aplikacji. Możliwe strategie obejmują wiele różnych: od najprostszej, gdzie zapotrzebowanie jest sumą zapotrzebowania najbardziej zasobochłonnych scenariuszy udostępnianych aplikacji do złożonych, w których ilość zasobów będących do dyspozycji w ramach sieci wirtualnej nie pozwala na dostarczenie wszystkich usług aplikacji w przypadku jednoczesnego wystąpienia wszystkich użytkowników z żądaniami dostępu do tych usług.

Przyjęcie pierwszej z wymienionych wyżej strategii wymiarowania oznacza, że brak ryzyka niedostępności usługi jest kompensowany niską efektywnością wykorzystania zasobów. Druga z wymienionych strategii wymiarowania wymusza możliwość lepszego wykorzystania zasobów związanego z ryzykiem braku dostępu do żądanych usług spowodowanego chwilowym brakiem zasobów. W takim przypadku niezbędne jest zarządzanie zasobami i usługami, które zminimalizuje ryzyko niedostępności usługi i ograniczy konsekwencje chwilowej niedostępności żądanej usługi. Ma to szczególne znaczenie wtedy, gdy rozważany scenariusz

aplikacji jest scenariuszem krytycznym (np. w sieciach otoczenia człowieka lub sieciach e-zdrowie).

Opracowywane aplikacje będą testowane w ramach scenariusza zakładającego realizację następujących etapów:

- zdefiniowanie repozytorium treści dla usług aplikacji,
- zdefiniowanie testowego zbioru użytkowników,
- udostepnienie treści i usług dla testowego zbioru użytkowników w ramach proponowanego modelu biznesowego,
- akwizycja i analiza danych o aktywności użytkowników oraz wykorzystaniu zasobów Równoległego Internetu.

Przeprowadzone dla każdej z aplikacji testy pozwolą na opracowanie efektywnych metod rezerwacji zasobów Równoległych Internetów oraz przygotowanie scenariuszy wdrożeń poszczególnych aplikacji.

4 Testy aplikacji

Opisane wyżej aplikacje współdzielą zasoby sieciowe i obliczeniowe sieci wirtualnych Równoległych Internetów, które są platformę współdzielonych zasobów wykorzystywanych przez użytkowników aplikacji współzawodniczących o te zasoby. Współdzielone zasoby to zasoby komunikacyjne (przepływności łączy), obliczeniowe (czas procesorów) i pamięciowe (pojemność pamięci) oraz udostępniana treść (bazy danych, repozytoria danych, itp.).

Opracowane aplikacje będą testowane z wykorzystaniem norm i standardów; w tym normy ISO/IES 29119 dotyczącej testowania m.in. aplikacji oraz norm: ISO/IEC 9126-1:2001 dotyczącej modelu jakości oraz metryki (*Software Engineering - Software Product Quality*), normy ISO/IEC 12207:1995 dotyczącej cyklu życia w wytwarzaniu oprogramowania (*Information Technology - Software Life Cycle Processes*) oraz normy ISO/IEC 14598 dotyczącej ewaluacji aplikacji (*Information Technology - Software Product Evaluation*). W procesie testowania aplikacji użyteczne będą także standardy IEEE: IEEE 730:2002. dotyczący zapewnienia jakości (*Standard for Software Quality Assurance Plans*), IEEE 829:1998. dotyczący dokumentacji testowania (*Standard for Software Test Documentation*), IEEE 1012:1986. dotyczący walidacji i weryfikacji aplikacji (*Standard for Verification and Validation Plans*), IEEE 1028:1997. dotyczący przeglądów i audytów (*Standard for Software Reviews and Audits*), IEEE 1044:1993. dotyczący klasyfikacji anomalii (*Standard Classification for Software Anomalies*) oraz IEEE 1061:1998. dotyczący metryk jakości (*Standard for a Software Quality Metrics Methodology*).

5 Podsumowanie

W wyniku przeprowadzonych testów określone zostaną charakterystyczne dla każdej ze zdefiniowanych w projekcie grup aplikacji wymagania dotyczące jakości usług komunikacyjnych oraz klas ruchu udostępnianych przez system IIP. Ze

względu na różnorodność aplikacji, także w ramach każdej z wyróżnionych grup, uzyskane wyniki posłużą do opracowania ogólnych strategii wdrożeń dla klas aplikacji wykorzystujących Równoległe Internety jako komunikacyjne platformy sieciowe. Ponadto będą one podstawą do opracowania zestawu tzw. "dobrych praktyk" w zakresie wzbogacenia logiki aplikacji o wykorzystujące interfejs systemu IIP funkcje rezerwacji zasobów w sposób właściwy dla cech danej aplikacji. Pozwoli to na zachowanie jakości usług świadczonych przez poszczególne aplikacje oraz właściwą obsługę generowanego przez nie ruchu, a w następstwie - zademonstrowanie unikalnych własności opracowywanej w projekcie IIP infrastruktury sieciowej.

Literatura

1. Burakowski, W. i in.: Specification of the IIP System - levels 1 and 2, Raport Projektu Inżynieria Internetu Przyszłości, Warszawa 2011.
2. Xiao, Y., Chen, H.,: Mobile Telemedicine; A Computing and Networking Perspective , Taylor and Francis Group, New York 2008,
3. EUROPEAN FUTURE INTERNET INITIATIVE (EFII):White paper on the Future Internet PPP Definition, January 2010.
4. Haibo, B.L., Sohraby, L.K., Wang, C.: Future internet services and applications, IEEE Network 24(5), 4-5, 2010.
5. System IIP: http://www.iip.net.pl
6. Świątek, P., Drwal, M., Grzech, A.: Providing Strict QoS quarantines for flows with time-varying capacity requirements. In proc. of 21st International Conference on Systems Engineering, s. 279–284, Las Vegas, 2011

Automatyczna kompozycja usług monitorowania dla systemu wspomagania treningu wytrzymałościowego sportowców

Krzysztof Brzostowski, Piotr Rygielski

Instytut Informatyki, Politechnika Wrocławska

Streszczenie W niniejszej pracy scharakteryzowano rozproszony system wspomagania treningu sportowca. Wyróżniono właściwości platformy przetwarzającej sygnały życiowe pozyskiwane z czujników umieszczonych na jego ciele. Przy użyciu technologii transmisji danych opracowywanych w ramach projektu Inżynieria Internetu Przyszłości rozproszone elementy opisywanej aplikacji wymieniają dane stanowiąc system wspomagania treningu jako całość. W pracy przedstawiono dekompozycję architektury aplikacji na elementy atomowe a następnie zamodelowano je jako usługi systemu opartego na paradygmacie SOA. Przetwarzanie w systemie usługowym jest elastyczne, co należy rozumieć tak, że poszczególne komponenty systemu mogą być zwielokrotnione a konkretne wersje zduplikowanych elementów powinny być wybierane do uruchomienia podczas procesu kompozycji usługi złożonej. W pracy formułuje się problem kompozycji złożonej usługi monitorowania oraz podaje się sposoby efektywnego rozwiązania tego problemu.

1 Wprowadzenie

Postęp w mikroelektronice i telekomunikacji pociąga za sobą m.in. na rozwój w dziedzinie nowych urządzeń pomiarowych charakteryzujących się niską ceną, niewielkimi rozmiarami oraz niskim zużyciem energii. Wielu producentów takich jak Shimmer, Zephyr Technology lub Polar oferuje rozwiązania, które znalazły zastosowanie w różnorodnych systemach związanych z telemedycyną. W literaturze opisano wiele aplikacji wspomagających terapię diabetyków [4, 7, 14], osób cierpiących na chorobę Parkinsona [12] oraz analizy i diagnostyki chodu [13]. Oprócz rozwiązań ściśle związanych z medycyną w literaturze opisanych zostało wiele systemów wspomagających trening osób uprawiających sport zawodowo i amatorsko. W tej grupie możemy wskazać na dwa główne nurty, z których jeden związany jest z systemami przeznaczonymi do wspomagania treningu wytrzymałościowego ogólnego przeznaczenia. Systemy takie charakteryzują się tym, że mogą być wykorzystywane zarówno przez sportowców amatorów jak i zawodowców. Celem stosowania takich rozwiązań jest monitorowanie przebiegu treningu wytrzymałościowego oraz wspomaganie podejmowania decyzji z nim związanych. Charakterystyczne cechy takich systemów to pomiar, w czasie aktywności ruchowej, takich wielkości fizjologicznych jak puls, częstość oddechów, aktywność

mięśni. Pomiary tych wielkości pozwalają na ocenę intensywności treningu i/lub oszacowanie czy zostały spełnione cele treningu wytrzymałościowego [1, 3, 9, 10]. Do drugiej grupy rozwiązań zaliczyć można systemy wspomagające trening dla konkretnej dyscypliny sportowej. Specjalizacja związana jest głównie z treningiem technicznym w mniejszym stopniu z treningiem wytrzymałościowym. Zadaniem treningu technicznego jest poprawa techniki wykonywanych ruchów lub ich sekwencji. Oczywiste jest, że projektowany system wspomagania treningu technicznego musi być dedykowany do konkretnej dyscypliny sportowej. W literaturze opisano systemy wspomagania m.in. treningu tenisistów stołowych [2] i tenisistów ziemnych [5, 18].

W niniejszej pracy przedstawiono system wspomagania treningu wytrzymałościowego bazującego na pomiarach z bezprzewodowych czujników i wynikach ich późniejszego przetwarzania. Ostatnim etapem jest prezentacja wyników pozyskiwania i przetwarzania danych na urządzeniach dostępowych zarówno trenera jak i sportowca. Całość procesu przetwarzania pomiarów została zamodelowana jako system usługowy. W pracy przedstawiono i podano rozwiązania problemu kompozycji złożonej usługi monitorowania treningu.

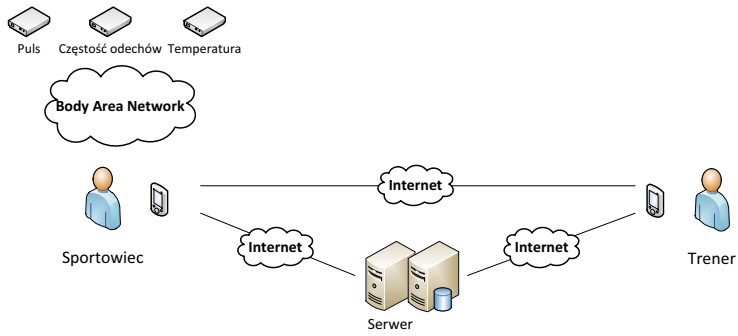
2 System wspomagania treningu wytrzymałościowego

W podpunkcie przedstawiono architekturę systemu oraz aplikacji do wspomagania treningu wytrzymałościowego sportowców. Głównym wymaganiem dla systemu jak również aplikacji jest to by wspomagała ona zarówno mobilnego sportowca jak i trenera. W tym przypadku przez wspomaganie będziemy rozumieli wspomaganie podczas pobierania danych (zadania po stronie urządzenia dostępowego sportowca) oraz prezentacji wyników przetwarzania danych (zadania zarówno po stronie urządzenia dostępowego sportowca jak i trenera).

2.1 Architektura systemu

Proponowana architektura systemu do wspomagania treningu sportowego została przedstawiona na rysunku 1. Główne elementy zaproponowanego rozwiązania to urządzenia dostępowe dla sportowca oraz trenera, którymi mogą być telefony komórkowe, PDA (ang. Personal Digital Assistant) lub (w przypadku trenera) komputer przenośny typu laptop. Kolejnym elementem proponowanego systemu jest serwer, którego głównym zadaniem jest przetwarzanie zgromadzonych danych oraz wysyłanie wyników przetwarzania danych do urządzeń dostępowych użytkowników. Od strony sportowca urządzenie dostępowe pozwala na prezentację wyników przetwarzania danych na serwerze, oraz wysyłanie pomiarów z czujników bezprzewodowych ulokowanych na ciele sportowca. W celu gromadzenia danych z tych czujników zaprojektowano dodatkową sieć typu BAN (ang. Body Area Network). Zadaniem urządzenia dostępowego trenera jest wizualizacja przetworzonych pomiarów z czujników pracujących w sieci BAN sportowca, prezentacja wyników przetwarzania danych np. wydatkowanej podczas treningu

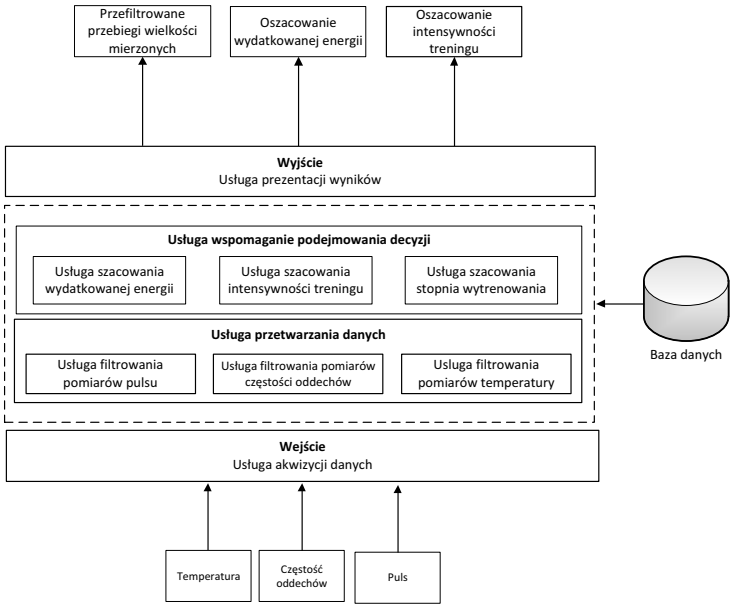
przez sportowca energii itp. Proponowany system umożliwia również zestawianie połączeń audio i wideo do przeprowadzania tele i wideo konsultacji pomiędzy sportowcem i trenerem.



Rysunek 1. Architektura systemu do wspomagania treningu sportowego.

2.2 Architektura aplikacji

W podrozdziale omówiono architekturę aplikacji, która została zaprojektowana na potrzeby systemu przedstawionego w poprzednim podrozdziale. Na rysunku 2 zaprezentowano schemat aplikacji z wyszczególnieniem trzech głównych elementów. Pierwszy z nich związany jest z akwizycją danych z czujników bezprzewodowych. Tak jak zaznaczono na rysunku w skład usługi monitorowania wchodzą trzy komponenty będące źródłem pomiarów tj. usługa pomiaru temperatury, częstości oddechów i pulsu. W drugiej sekcji zaznaczono dwie kolejne usługi tj. jedna związana z przetwarzaniem danych i druga dotycząca wspomagania podejmowania decyzji. Z pierwszą z nich związane są trzy usługi atomowe dotyczące przetwarzania danych zmierzonych z wykorzystaniem czujników bezprzewodowych. Natomiast w skład usługi wspomagania podejmowania decyzji wchodzą trzy usługi atomowe do szacowania wydatkowanej podczas treningu energii, intensywności treningu oraz stopnia wytrenowania. Ostatni z elementów proponowanej aplikacji związany jest z usługą prezentacji wyników, która to usługa zaprojektowana jest tak, by możliwa była obsługa różnych typów urządzeń dostępowych wykorzystywanych zarówno przez trenera jak i sportowca. Zaprezentowana aplikacja została zaprojektowana jako zgodna z paradygmatem SOA w celu rozproszenia i uelastycznienia rozwiązania. Architektura aplikacji została przedstawiona na rysunku 2. W sekcji trzeciej opisano sposób kompozycji usługi monitorowania z usług składowych w oparciu o przedstawiony przykład.



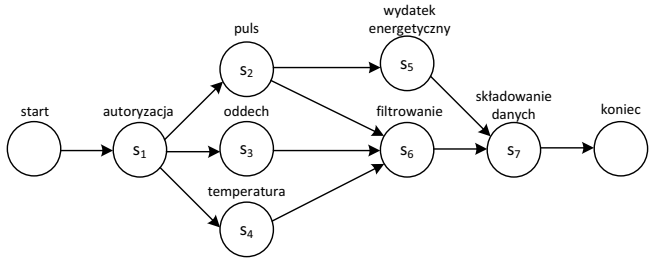
Rysunek 2. Architektura aplikacji do wspomagania treningu sportowego.

3 Usługa wspomagania treningu wytrzymałościowego

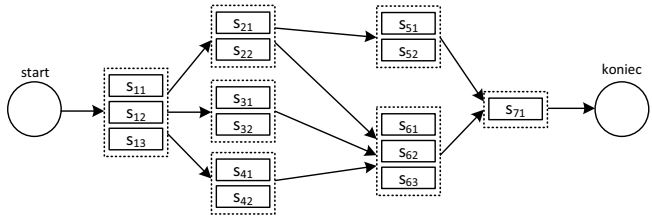
Architektura aplikacji przedstawiona na rysunku 2 została przekształcona do postaci scenariusza wykonania usługi złożonej, który zaprezentowano na rysunku 3. Wyróżnione wcześniej komponenty systemu monitorowania zostały zamodelowane jako usługi atomowe. Każdy z węzłów grafu (rysunek 3) odpowiada za przetwarzanie właściwej porcji danych i w ogólności może być umiejscowiony w dowolnej lokalizacji sieciowej. Dodatkowo wyróżnia się węzeł początkowy i końcowy, który modeluje odpowiednio pobranie surowych danych z czujników oraz przesłanie już przetworzonych na urządzenia dostępowe sportowca i trenera.

Każda z usług przetwarzających może być umieszczona na różnych węzłach obliczeniowych w systemie wykonawczym. Ponadto niektóre usługi mogą być zduplikowane bądź mogą różnić się wartościami parametrów jakościowych, podczas gdy funkcjonalność pozostaje bez zmian. Na rysunku 4 wyróżniono przykładowe wersje poszczególnych usług biorących udział w procesie monitorowania omawianym wyżej.

Aby usługa monitorowania mogła zostać uruchomiona należy wybrać odpowiednią wersję każdej usługi atomowej w celu wskazania konkretnych kandydatów, którzy będą brali udział w procesie przetwarzania strumienia monitorowanych danych. Graf powstały w wyniku wyboru jednej z kandydujących wersji dla każdej usługi atomowej nazywa się planem wykonania usługi złożonej. W kolejnym podrozdziale omówiony został problem wyznaczenia optymalnego



Rysunek 3. Scenariusz usługi złożonej monitorowania.



Rysunek 4. Scenariusz wykonania złożonej usługi monitorowania z wyróżnionymi wersjami usług atomowych.

planu wykonania usługi złożonej, który dla uproszczenia nazwano problemem kompozycji usługi złożonej.

4 Problem kompozycji usługi złożonej

Zadanie wyznaczenia optymalnego planu wykonania usługi złożonej jest zadaniem rozwiązywanym na ostatnim etapie kompozycji usługi. Na ogół zadanie wyznaczenia optymalnego planu wykonania usługi złożonej można opisać następująco. Zadaniem systemu jest wybranie jednej z wersji każdej atomowej usługi określonej jednoznacznie przez scenariusz w taki sposób, że wymagania jakościowe są spełnione. Na uwagę zasługuje fakt, że wybranie dwóch wersji usługi atomowej precyzyjnie określa również ścieżki komunikacyjne, które będą użyte do przesyłania danych między tymi usługami - zakładając, że zawsze jest wybierana najlepsza możliwa ścieżka, jeśli istnieje więcej niż jedna możliwość wyboru. Podejmowany problem jest uogólnieniem problemu kompozycji polegającej na wyznaczeniu najkrótszej ścieżki w ważonym grafie rozpatrywanym np. w pracy [8]. W tym przypadku, różnica jest taka, że rozwiązanie jest reprezentowane jako graf a nie jako pojedyncza ścieżka przez co poprzednich algorytmów nie można stosować. Formalnie zadanie wyznaczenia optymalnego planu wykonania usługi złożonej można sformułować w następujący sposób:

Dla danych:

- Scenariusz wykonania usługi złożonej dany grafem $GC = (VC, EC)$ (patrz rysunek3)
- Żądanie wykonania usługi reprezentowane przez SLA_l
- Wymagania niefunkcjonalne Ψ_l zawarte w SLA_l
- Kryterium jakości planu wykonania $Q(G, \Psi_l)$

Znaleźć: Taki zestaw wersji usług atomowych dla scenariusza GC, który minimalizuje zadane kryterium jakości Q:

$$as_{1j_1^*}, \ldots, as_{kj_k^*} = \underset{as_{1j_1}, \ldots, as_{kj_k}}{\arg\min} Q\big(G(\{as_{1j_1}, \ldots, as_{kj_k}\}, E), \Psi_l\big),$$

gdzie as_{1j_k} oznacza j_k-tą wersję k-tej usługi atomowej.

Sformułowany wyżej problem można łatwo przetransformować do problemu wielowymiarowego wielokryterialnego problemu plecakowego (ang. Multpile choice multidimensional knapsack problem, MMKP), który jest problemem z klasy NP-trundych. Zadaniem jest wybrać dokładnie jeden element (wersja usługi atomowej) z każdej grupy (usługa atomowa) w taki sposób, że jakość reprezentowana przez wybrane przedmioty (usługa złożona) jest optymalna w sensie wybranego kryterium jakości Q. Rozwiązanie tego problemu jest znane a niektóre algorytmy rozwiązujące problem zostały przedstawione np. w pracy [11]. Tak sformułowanie zadanie jest problemem NP-trudnym, ponieważ transformuje się do problemu MMKP w czasie wielomianowym. Dokładne rozwiązanie postawionego problemu przedstawiono w pracy [17]. Z racji, na niewielki rozmiar problemu rozważanego w tej pracy rozwiązanie dokładne uzyskiwane jest w krótkim czasie. Ponadto w pracy [17] przedstawiono dokładną iteracyjną procedurę wyliczania wartości kryterium jakości danego wzorem powyżej. Inne rozwiązania problemu MMKP można znaleźć np. w pracach [6, 15, 16].

5 Uwagi końcowe

W pracy przedstawiono system monitorowania i wspomagania treningu wytrzymałościowego sportowca. Zaprezentowano architekturę aplikacji i scharakteryzowano system akwizycji i rozproszonego przetwarzania aktualnych wartości pomiarów parametrów życiowych. Aplikację zamodelowano przy użyciu paradygmatu systemów zorientowanych na usługi i wyróżniono potrzebę duplikacji niektórych usług przetwarzających. Wyróżnienie wielu wersji usług mogących dostarczać jednakową funkcjonalność spowodowało potrzebę efektywnego składania (kompozycji) złożonych usług monitorowania. W pracy sformułowano problem kompozycji usługi złożonej na wybranym przykładzie. Problem został scharakteryzowany oraz podano przykładowe algorytmy znajdujące jego rozwiązanie. Dalsze prace w zakresie omawianej pracy będą polegały głównie na implementacji i testach jakościowych omawianego rozwiązania. Przeprowadzone zostaną również rzeczywiste eksperymenty podczas treningu tenisisty ziemnego.

Literatura

1. J.V.Alonso, et al., Ambient intelligence systems for personalized sport training, Sensors, Vol. 10, pp. 2359 – 2385
2. A. Baca, P. Kornfeid, Rapid Feedback systems for elite sports training, Sports Technologies, pp. 70 - 76
3. F. Buttussi, L. Chittaro, MOPET: A Context-Aware and User-Adaptive Wearable System for Fitness Training, Vol. 42, pp. 153 - 163
4. K. Brzostowski, J.M. Tomczak, J. Świątek, Wspomaganie przeprowadzenia wywiadu lekarskiego dla systemu zarządzającego terapią cukrzycy, XVII Krajowa Konferencja Automatyki – KKA'2011, Kielce-Cedzyna (w druku)
5. D. Connaghan et al., A sensing platform for physiological and contextual feedback to tennis athletes, Body Sensor Networks, 2009.
6. J. Egeblad and D. Pisinger. Heuristic approaches for the two- and three-dimensional knapsack packing problems. Computers and Operations Research, 36:1026–1049, 2009.
7. L. Grandinetti, O. Pisacane, Web based prediction for diabetes treatment, Future Generation Computer Systems, Future Generation Computer Systems, Vol. 27, 2011, pp. 139-147
8. A. Grzech. P. Rygielski, P. Świątek, QoS-aware infrastructure resources allocation in systems based on service-oriented architecture paradigm, Performance modelling and evaluation of heterogeneous networks, 6th Working International Conference HET - NETs 2010, Zakopane, January 14th-16th, 2010 s. 35-47.
9. H. J. Hermens, M.M.R. Vollenbroek-Hutten, Towards remote monitoring and remotely supervised training, Journal of Electromyography and Kinesiology, Vol. 18, pp. 908 – 919
10. Y.J. Hong, et al., Mobile health monitoring system based on activity recognition using accelerometer (in press)
11. Shahadat Khan and Kin F. Li and Eric G. Manning, The Utility Model For Adaptive Multimedia Systems, In International Conference on Multimedia Modeling, 1997 pp.111-126.
12. K. Lorincz, et al., Mercury: A Wearable Sensor Network Platform for High-Fidelity Motion Analysis
13. M. J. McGrath, SHIMMERTM Validation and Applications
14. S.G. Mougiakakou, et al., SMARTDIAB: A Communication and Information Technology Approach for the Intelligent Monitoring, Management and Follow-up of Type 1 Diabetes Patients, IEEE Transactions on Information Technology in Biomedicine, Vol. 14, 2010, pp. 622 – 633
15. D. Pisinger. Where are the hard knapsack problems? Computers & Operations Research, 32:2271–2282, 2005
16. D. Pisinger. A minimal algorithm for the Multiple-choice Knapsack Problem. European Journal of Operational Research, 83:394–410, 1995.
17. P. Rygielski, P. Świątek, Graph-fold: an efficient method for complex service execution plan optimization, Systems Science vol. 36, no.3, pp. 25-32, 2010.
18. K. Urawaki et al., Development of the Learning Environment for Sport form Education with the Visualisation of Biophysical Information, ICAT, 2000

Taksonomia systemów e-learningowych w aspekcie zapotrzebowania na zasoby techniczne oraz usługi sieciowe

Lesław Sieniawski, Krzysztof Juszczyszyn

Dział Zdalnej Edukacji, Instytut Informatyki Politechnika Wrocławska

Streszczenie W artykule zaproponowano taksonomię systemów e-learningowych, odwołującą się do zasobów wykorzystywanych przez świadczone przez nią usługi edukacyjne. Przedstawiono przykładowe wyniki jej zastosowania do klasyfikacji systemów Działu Zdalnej edukacji Politechniki Wrocławskiej. Wskazano na stosowalność zaproponowanej taksonomii do klasyfikacji usług e-learningowych.

1 Wprowadzenie

Usługi realizowane z wykorzystaniem sieci komputerowych stanowią obecnie dominującą część wszystkich usług świadczonych przez systemy komputerowe. Swoisty wyścig pomiędzy zapotrzebowaniem na ilość i jakość tych usług a możliwościami realizacyjnymi wymaga dokonywania pomiarów pozwalających na zobiektywizowaną ocenę potrzeb oraz stopnia możliwości ich zaspokojenia przez systemy komputerowe i infrastrukturę komunikacyjną. W tym celu definiuje się odpowiednio pojęcia jakości usług (ang. *Quality of Service*; *QoS*) oraz jakości odbioru (postrzeganej jakości usług) (ang. *Quality of Experience*; *QoE*). Definicje miar QoS wprowadzone np. przez zalecenia [2] i [3] nie odnoszą się do rodzajów aplikacji, lecz do charakterystycznych dla nich typów przekazywanych treści (plików i komunikatów) i określają wartości progowe, bez określania gradacji ocen. Rzeczywiste aplikacje operują jednak mieszankami treści różnych typów co powoduje, że bezpośrednie zastosowanie tych miar do oceny jakości usług i ich odbioru jest utrudnione. W niniejszym opracowaniu zostanie przedstawiona propozycja kilku taksonomii systemów e-learningowych związanych z oceną zapotrzebowania na zasoby techniczne i usługi sieciowe. Celem przypisania danego systemu e-learningowego do określonej klasy - w sensie przyjętej taksonomii - jest wskazanie wymaganego dla niego poziomu jakości usług. Wyniki badań wykorzystane zostaną do klasyfikacji oraz szacowania zapotrzebowania na zasoby usług wirtualnego laboratorium obliczeniowego Online Lab realizowanego na Politechnice Wrocławskiej w ramach projektu POIG "Inżynieria Internetu Przyszłości" [4].

2 Badanie przykładowego systemu e-learningowego

Przez system e-learningowy (SeL) będziemy rozumieli aplikację typu LMS (ang. *Learning Management System*) wraz z uczestniczącym w jej wykonywaniu sys-

temem komputerowym zawierającym m.in. serwer WWW, serwer bazy danych oraz infrastrukturą teleinformatyczną i stanowiskami pracy użytkowników.

Działanie aplikacji LMS obejmuje m.in.:

- przesyłanie do użytkowników żądanych (statycznych) stron WWW,
- interpretację danych zawartych w formularzach wysyłanych do LMS przez użytkowników,
- generowanie (dynamicznych) stron WWW,
- dwukierunkowe przesyłanie plików (strumieni) danych.

Czynności te stanowią usługi edukacyjne systemu SeL. Funkcje systemu LMS wykonywane są z udziałem serwera WWW oraz współpracującej z nim bazy danych. Poszczególne strony WWW, zarówno statyczne, jak i dynamiczne, składają się z jednego lub wielu plików, których zawartość jest interpretowana przez przeglądarkę internetową, tworząc obraz postrzegany przez użytkownika za pomocą zmysłów. Faktyczny sposób obsługi plików stron WWW, plików (strumieni danych) przesyłanych do systemu LMS wpływa na ocenę jakości odbioru.

Przez system e-learningowy (SeL) będziemy rozumieli aplikację typu LMS, przy czym jego klasyfikacja wymaga od taksonomii spełnienia następujących warunków:

- Taksonomia SeL winna być stosowalna - na podstawie danych o funkcjonowaniu konkretnego systemu zebranych w tym systemie lub jego otoczeniu powinno dać się jednoznacznie wskazać klasę, do której dany SeL należy.
- Klasa x, do której zostanie zaliczony dany SeL nie jest ustalona w sposób trwały: zależy ona od dostarczanych treści oraz preferencji użytkownika. W konsekwencji, przypisanie SeL do danej klasy wybranej taksonomii może zmieniać się w jego cyklu życia.
- Ocena jakości systemu e-learningowego będzie dokonywana całościowo - bez wyodrębniania oferowanych przezeń treści i funkcji (kursów, modułów, lekcji, ćwiczeń, itd.).
- Poszczególne typy elementarnych usług edukacyjnych związane są z rodzajami zawartości plików przetwarzanych przez SeL w ramach odpowiednich usług. Pliki można grupować według rodzajów zawartości, np. dla odzwierciedlenia ich przeznaczenia lub sposobu prezentacji.
- Przyjmuje się, że na jakość odbioru usług edukacyjnych ma wpływ objętość przetwarzanych przez SeL plików i że jest to zależność monotoniczna słabo malejąca (większy plik - jakość odbioru nie większa). Zakłada się przy tym, że wymagania jakościowe dotyczące realizacji przez system SeL usługi dotyczącej pliku o zerowej długości są spełnione zawsze, bez względu na rodzaj zawartości.

Przyjmiemy następujące oznaczenia:
$T = \{t_1, t_2, \ldots, t_N\}$ - ciąg typów plików określony przez wybrany model QoS przetworzonych przez LMS; t_i oznacza i-ty typ pliku;
N - liczba typów plików,
$C = \{c_1, c_2, \ldots, c_N\}$ - rozkład empirycznej liczby plików poszczególnych typów;

c_i jest liczbą przetworzonych plików typu i,
$V = \{v_1, v_2, \ldots, v_N\}$ - rozkład łącznej objętości plików poszczególnych typów; v_i określa łączną objętość przetworzonych plików typu i w bajtach.

Rozkłady C i V obejmują wszystkie przetworzenia plików, łącznie z powtórzeniami. W dalszym ciągu rozważania ograniczymy do zbioru $T = \{$*dane, tekst, grafika, audio, wideo*$\}$ (dla którego N=5), ponieważ taka klasyfikacja typów dostarczanych treści jest właściwa dla wspomnianych wcześniej taksonomii usług sieciowych. Poszczególne elementy zbioru T reprezentują (grupują) następujące formaty zawartości plików uwidocznione zazwyczaj w postaci rozszerzeń nazw plików (Audio: WAV, MP3, OGG, ... itd.). Źródłem danych do oceny SeL będzie kronika serwera WWW. W badanych systemach SeL kronika ta miała postać pliku tekstowego. Algorytm przetwarzania kroniki, można przedstawić następująco (szerzej problem analizy danych z logów platform edukacyjnych opisano w [1]):

1. Grupowanie rekordów wg typu przetwarzanego pliku.
2. Wyznaczenie statystyk: rozkładu liczby plików w grupach i rozkładu łącznej objętości plików.
3. Dla każdej grupy $i \in \langle 1, N \rangle$, wyznaczenie rozkładu objętości plików danej grupy oraz średniej $E(v_i)$ i maksymalnej $max(v_i)$ objętości pliku grupy.

Czynności wstępne poprzedzające właściwą analizę danych obejmują:

1. Importowanie kroniki serwera WWW systemu e-learningowego do bazy danych zawierającej tabelę opisaną powyżej,
2. Filtrowanie kroniki polegające na usuwanie rekordów błędnych, niekompletnych oraz zbędnych z punktu widzenia celu analizy, np. zarejestrowane usługi serwera WWW nie dotyczące badanego systemu e-learningowego.

Przez analizę zawartości tabeli *Accesslog* utworzonej na podstawie kroniki serwera WWW portalu edukacyjnego A (należącego do zasobów Działu Zdalnej Edukacji Politechniki Wrocławskiej) uzyskano następujące dane:

Mnożąc odpowiednio wartości z kolumn fileCount (liczba plików) i avgFileSize (zaokrąglona średnia objętość pliku) można uzyskać oszacowanie łącznej objętości plików o danych rozszerzeniach nazwy (nieuwidocznione w tabeli). Wartość "/" w kolunie fileExtension oznacza żądanie dostarczenia przez serwer WWW ze wskazanego katalogu pliku o nazwie domyślnej, ustalonej w konfiguracji tego serwera (np. index.php, index.html, itp.). Wartość "?" koduje niemożność ustalenia rodzaju zawartości żądanego pliku, z powodu błędu w składni żądania lub niejednoznaczności interpretacyjnych. Pogrubiono największe wartości w kolumnach fileCount, avgFileSize oraz maxFileSize (największa objętość pliku). Na podstawie zawartości analizowanej kroniki uzyskano również dane na temat średniego pasma transmisji danych w kierunku od serwera WWW do klientów.

Należy zwrócić uwagę, że zastosowany tu bezpośredni sposób szacowania średniego pasma prowadzi do uzyskania wartości zaniżonych w stosunku do rzeczywistych potrzeb. Chwilowe zapotrzebowania na pasmo mogą być znacznie większe.

fileExtension	fileCount	avgFileSize	maxFileSize
/	19 385	7 842	149 724
?	4 687	253	11 508
aln	2	2 195	2 195
bmp	267	1 192 082	1 243 278
c	41	5 182	23 160
class	4 296	627	7 289
mdb	1	2 891 776	2 891 776
...	...	...	...

fileExtension	fileCount	avgFileSize	maxFileSizee
mpg	1	3 845 715	3 845 715
o	1	7 064	7 064
odt	5	17 758	26 561
pdb	1	1 332	1 332
pdf	3 025	293 193	22 077 788
php	177 363	4 566	4 541 488
zip	18	**5 129 901**	**47 436 405**
...	...	...	...

Tabela 1. Statystyki liczby i objętości plików- portal A (fragment)

Okres obserwacji	od 2010-06-16 04:05:04
Okres obserwacji	do 2010-06-21 13:46:48
Długość okresu obserwacji	466.904 [s]| 5,404 [dni]
Objętość dostarczonych plików	7.341.745.405 [B]
Średnie pasmo wykorzystane do dostarczenia plików	15.742 [B/s]

Tabela 2. Globalne statystyki serwera WWW portalu A

W kolejnych punktach przedstawione zostaną propozycje taksonomii, wykorzystujące zapisy kroniki serwera WWW obsługującego dany system e-learningowy. Do klasyfikacji plików o określonych rodzajach zawartości wykorzystano tabelę fileTypes w bazie danych.

3 Taksonomie zależne od profilu dynamicznego WWW

1. Taksonomia częstościowa. Systemy e-Learningowe klasyfikuje się według rozkładu liczby plików poszczególnych typów; dany SeL jest zaliczany do określonej klasy według typu pliku o największym prawdopodobieństwie wystąpienia $c_i / \sum c_i$ (lub, co jest równoważne, największej wartości samego c_i).
2. Taksonomia tolerancyjna. Systemy e-Learningowe klasyfikuje się według rozkładu średniej objętości poszczególnych typów plików; dany SeL jest zaliczany do określonej klasy według typu pliku o największej średniej objętości $E(i)$.

Taksonomia	Częstościowa	Tolerancyjna	Redundancyjna	-
t_i	c_i	E(i)	max(i)	v_i
?	4 685	253	11 508	1 185 305
audio	7	**2 205 447**	3 673 887	15 438 129
dane	5 272	898 464	**47 436 405**	4 736 702 208
grafika	247 655	4 377	1 466 674	1 083 985 935
tekst	**306 615**	4 344	4 541 488	1 331 935 560
wideo	4 235	40 248	9 176 144	170 450 280

Tabela 3. Dane przykładowego SeL

3. Taksonomia redundancyjna. Systemy e-Learningowe klasyfikuje się według rozkładu maksymalnej objętości poszczególnych typów plików; dany SeL jest zaliczany do określonej klasy według typu pliku o największej maksymalnej objętości $max(i)$.

Przykład 1.

Dla portalu A wyznaczono rozkłady c_i, v_i oraz $E(v_i)$ i $max(v_i)$, $i \in \langle 0, 5\rangle$, $i = 0$ oznacza nieustalony typ pliku (nazwą typu pliku jest "?").

Wyróżnione komórki zawierają wartości maksymalne w swoich kolumnach. Oznacza to, że - w sensie taksonomii częstościowej - w portalu A dominują pliki typu tekstowego, w sensie tolerancyjnej - pliki typu audio, a w sensie redundancyjnej - dominują dane.

4 Taksonomie zależne od profilu dynamicznego WWW i preferencji użytkownika

Przyjmujemy, że preferencje użytkownika przedstawione są za pomocą ciągu wag przypisanych do poszczególnych typów plików, $W = \{w_1, w_2, \ldots, w_N\}$, $0 \leqslant w_i \leqslant 1$. Waga mniejsza od 1 oznacza, że użytkownik jest skłonny pogodzić się z tym, że jakość usługi związanej z przetwarzaniem i dostarczaniem plików danego typu może być obniżona. W szczególności, waga równa zeru wyraża fakt, że jakość usługi związanej z danym typem plików jest nieistotna.

1. Taksonomia ważona tolerancyjna. Systemy e-Learningowe klasyfikuje się według rozkładu iloczynu wagi i średniej objętości poszczególnych typów plików; dany SeL jest zaliczany do określonej klasy według typu pliku o największej wartości iloczynu $w_i \cdot E(i)$.
2. Taksonomia ważona redundancyjna. Systemy e-Learningowe klasyfikuje się według rozkładu iloczynu wagi i maksymalnej objętości poszczególnych typów plików; dany SeL jest zaliczany do określonej klasy według typu pliku o największej wartości iloczynu $w_i \cdot max(i)$.

Przykład 2.

Wykorzystamy zestaw danych z poprzedniego przykładu, uzupełniając zamieszczoną tam tabelę o trzy nowe kolumny.

Taksonomia	Częstościowa	Telerancyjna	Redundancyjna
t_i	c_i	E(i)	max(i)
?	4 685	253	11 508
audio	7	**2 205 447**	3 673 887
dane	5 272	898 464	**47 436 405**
grafika	247 655	4 377	1 466 674
tekst	**306 615**	4 3444	4 541 488
wideo	4 235	40 248	9 176 144

-	-	Ważona tolerancyjna	Ważona redundancyjna
V_i	w_i	$w_i \cdot E(i)$	$w_i \cdot max(i)$
1 185 305	1,00	253	11 508
15 438 129	0,90	**1 984 902**	3 306 498
4 736 702 208	1,00	898 464	**47 436 405**
1 083 985 935	0,75	3 283	1 100 006
1 331 935 560	1,00	4 344	4 541 488
170 450 280	0,65	26 161	5 964 494

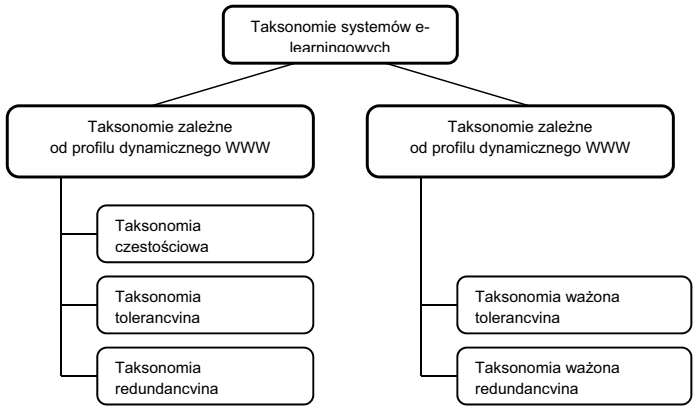
Tabela 4. Rozszerzone dane przykładowego SeL

Kolumna w_i zawiera współczynniki wagowe reprezentujące znaczenie, jakie użytkownik nadaje jakości odbioru poszczególnych typów treści systemu e-learningowego; w niniejszym przykładzie wartości wag przypisano arbitralnie. Ten sam zbiór danych jak poprzednio oraz przyjęty układ wag wskazują teraz, że w badanym SeL dominują pliki audio (w sensie taksonomii ważonej tolerancyjnej) lub że dominują dane (dla taksonomii ważonej redundancyjnej). W obydwu przykładach pliki o nieokreślonym typie nie wpływają na klasyfikację danego SeL.

5 Wnioski

W pracy przedstawiono próbę zdefiniowania taksonomii systemów e-learningowych odnoszącej się do wykorzystania przez ten system zasobów technicznych oraz usług sieciowych. Jednym z kluczowych wymagań dla taksonomii była jej stosowalność, a więc możliwość zaliczenia danego SeL do odpowiedniej klasy wyłącznie na podstawie danych dostępnych w kronice serwera WWW, nie obejmujących jednak informacji o wykorzystaniu procesora, pamięci operacyjnej i innych zasobów. W konsekwencji, przedstawione taksonomie wiążą się z danymi opisującymi wykorzystanie zasobów i usług pośrednio - poprzez taksonomie usług sieciowych opisujące minimalne wymagania QoS dla poszczególnych usług związanych z udostępnianiem treści różnych typów. Na rys.1. zostało przedstawione drzewo proponowanej taksonomii systemów e-learningowych.

Klasa, do której - zgodnie z przyjęta taksonomią - zostanie przypisany dany SeL wskazuje wymagania QoS, jakie winny być mu zapewnione. Tak więc, jeśli zostanie ustalone, że dominującym typem zawartości plików są dane audio, to systemowi e-learningowemu należy zapewnić jakość obsługi wymaganą



Rysunek 1. Drzewo taksonomii systemów e-learningowych.

dla usługi tego właśnie rodzaju. Zakwalifikowanie SeL na podstawie taksonomii tolerancyjnej oznacza przy tym akceptację skutków obniżenia jakości obsługi plików większych niż średnie. Podobnie, użycie taksonomii redundancyjnej wiąże się ze świadomym "przewymiarowaniem" potrzebnej jakości obsługi danego SeL. Taksonomia częstościowa jest najprostszą z przedstawionych wyżej koncepcji. Z powodu ilościowej dominacji plików tekstowych w strumieniu żądań przetwarzanych przez serwer WWW, klasyfikacja systemu e-learningowego według tej taksonomii może jednak prowadzić do zaniżenia wymagań dotyczących jakości obsługi.

Literatura

1. A. Syguła, O czym mówią logi platformy edukacyjnej. Możliwości personalizacji kursów e-learningowych w oparciu o logi systemowe, X Konferencja Wirtualny Uniwersytet. Model, narzędzia, praktyka, Warszawa, 16-18 czerwca 2010
2. K. Babiarz, K. Chan, F. Baker, Configuration Guidelines for DiffServ Service Classes, The Internet Society. The Network Working Group, 2006
3. P. Coverdale, Multimedia QoS requirements from a user perspective, International Telecommunication Union, ITU-T Study Group, Workshop on QoS and user-perceived transmission quality in evolving networks, 18 Oct 2001, http://www.itu.int/itudoc/itu-t/workshop/qos/ s2p1.pdf (pobrany: 2010-05-30)
4. K. Juszczyszyn, M. Paprocki, A. Prusiewicz, L. Sieniawski: Personalization and Content Awareness in Online Lab - Virtual Computational Laboratory. ACIIDS (1) 2011: 367-376

Nowe media i Internet 3D

Arkadiusz Sochan, Ryszard Winiarczyk

Instytut Informatyki Teoretycznej i Stosowanej Polskiej Akademii Nauk

Streszczenie W rozdziale przedstawiono w syntetyczny sposób charakterystykę rozwiązań stosowanych do akwizycji i wizualizacji treści w wysokich rozdzielczościach i treści 3D. Opisano przykładowe usługi bazujące na wykorzystaniu tego typu treści w odniesieniu do rozwiązań komunikacyjnych opracowywanych w ramach projektu Inżynieria Internetu Przyszłości.

1 Wprowadzenie

Dzisiejsze scenariusze użycia Internetu i występujące w nim charakterystyki ruchu coraz bardziej różnią się od pierwotnych założeń jego zastosowań. Stały się one przyczynkiem do rozpoczęcia prac badawczych i rozwojowych nad Sieciami Przyszłości (ang. *Future Networks*), których celem jest zapewnienie funkcjonalności i usług, które są obecnie trudne do zrealizowania [1]. Przewiduje się, że w niedługiej perspektywie nastąpi znaczący wzrost zapotrzebowania na przekaz treści multimedialnych wymagających nowych technik i form prezentacji uwzględniających między innymi: zwiększenie interaktywności, personalizację, bezpieczny i inteligentny dostęp do rozproszonych źródeł, zachowanie poziomu jakości z uwzględnieniem wzrostu mobilności użytkowników. Ten zbiór nowych technik i form przekazu nazywany jest Nowymi Mediami lub Mediami Przyszłego Internetu (ang. *Future Internet Media*) [2].

Do technologii reprezentatywnych dla Nowych Mediów można zaliczyć między innymi: strumieniowanie treści w wysokich rozdzielczościach oraz dystrybucję treści 3D (zarówno w formie strumieni 3D, jak obiektów 3D), która stanowi bazę dla Internetu 3D. Wzrost świadomości użytkowników, dotyczącej potencjalnych możliwości dostępnego sprzętu i oprogramowania do przygotowywania, dystrybucji i wizualizacji tego typu treści, przyczynił się do uruchomienia w Internecie pierwszych serwisów udostępniających treści z wykorzystaniem Nowych Mediów, jednak w ograniczonym zakresie ze względu na brak odpowiednich usług.

Poniżej zawarto syntetyczny przegląd technik i technologii związanych z akwizycją, reprezentacją i wizualizacją treści xK (w wysokich rozdzielczościach) oraz 3D. Opisano reprezentatywne usługi i aplikacje dla przekazu tego typu treści oraz możliwość wykorzystania ich w połączeniu z technologiami sieciowymi opracowywanymi w ramach projektu Inżynieria Internetu Przyszłości.

2 Akwizycja i wizualizacja treści xK i 3D

Rozwój systemów akwizycji strumieni wideo w wysokiej rozdzielczości jest stymulowany z jednej strony specjalistycznymi zastosowaniami o charakterze diagnostyczno-dokumentacyjnym, z drugiej strony upowszechnianiem się technologii do wielkoformatowych prezentacji cyfrowych takich jak: kina cyfrowe, ściany wideo czy telebimy. Dla zachowania, a nawet poprawy jakości prezentacji przy zwiększających się wymiarach ekranów wymagany jest więc wzrost rozdzielczości podczas akwizycji. W dalszej części przez treści xK są rozumiane materiały wideo rejestrowane z rozdzielczością 2K (2048 pikseli w linii obrazu) lub większą.

Obecnie na rynku bardzo mało urządzeń umożliwia akwizycję w rozdzielczości większej od 2K. Wśród profesjonalnych kamer telewizyjnych i filmowych dominują rozwiązania oferujące rozdzielczości HD i 2K. Kamery 4K mają przede wszystkim zastosowania kinematograficzne. Często są to jeszcze urządzenia prototypowe i dodatkowo rzadko udostępniają sygnał w pełnej rozdzielczości podczas rejestracji. Nie zmienia to faktu, że rozwiązania do rejestracji materiałów w wysokich rozdzielczościach są permanentnie rozwijane. Przykładowo japońska telewizja NHK dysponuje już prototypem kamery 8K, która we współpracy z Hitachi rejestruje i prezentuje testowo treści w tej rozdzielczości.

Wizualizacja xK może być realizowana z wykorzystaniem różnych technologii. Najprostszym i zarazem najtańszym rozwiązaniem pozwalającym na prezentację obrazu o rozdzielczościach xK lub większej jest połączenie wielu niedrogich monitorów komputerowych lub telewizorów HD. Jednakże zastosowanie nawet cienko-ramkowych lub bezszwowych monitorów nie eliminuje efektu przerw między wyświetlanymi fragmentami obrazu. Nieco droższym rozwiązaniem jest skorzystanie w rzutników HD wraz z zastosowaniem tzw. systemu edge-blending, który powoduje, że łączenia między obrazami są niemal niewidoczne, ale kalibracja wielu projektorów w takich koniguracjach jest dosyć czasochłonna. Jednak na rynku zaczynają być już oferowane dedykowane urządzenia do prezentacji w rozdzielczościach 2K i 4K.

Technologie 3D są już wykorzystywane od wielu lat w różnych dziedzinach - na przykład do wspomagania projektowania, wizualizacji naukowych, symulatorach czy oczywiście w grach komputerowych, które w znaczącym stopniu przyczyniają się do popularyzacji tego typu technologii. Często przez aplikacje 3D rozumiane są rozwiązania, w których przetwarza się „informacje 3D" i przygotowuje się ich trójwymiarową graficzną reprezentację, jednak sama wizualizacja jest realizowana w klasyczny, dwuwymiarowy sposób. Dlatego są one określane jako 2,5D w odróżnieniu od coraz popularniejszych rozwiązań, w których stosuje się faktycznie wizualizacje 3D. Dynamiczny rozwój tych ostatnich jest związany nie tylko z coraz większym potencjałem sprzętu do tworzenia treści 3D, ale także z możliwościami sprzętu do akwizycji 3D.

Ze względu na źródło pochodzenia treści 3D można je podzielić na dwa rodzaje:

1. naturalne - rejestrowane na przykład z wykorzystaniem systemów akwizycji stereoskopowej czy skanerów 3D,

2. syntetyczne - generowane przy użyciu oprogramowania przykładowo do modelowania 3D lub do wspomagania projektowania.

Ponadto można wyróżnić dwie podstawowe formy stosowane do zapisu i przekazywania tego typu treści:

1. reprezentacja rastrowa - wykorzystywane są grupy obrazów otrzymywanych z wielu punktów obserwacji (w szczególności obrazy stereoskopowe),
2. reprezentacja wektorowa - opis scen i obiektów 3D realizowany jest z użyciem np. siatek i tekstur.

Obecnie na rynku jest dostępnych coraz więcej urządzeń oraz systemów do akwizycji obiektów 3D. W przypadku obiektów statycznych są stosowane skanery optyczne i laserowe oraz digitalizatory, natomiast do rejestracji obiektów dynamicznych używane są systemy akwizycji ruchu (ang. *motion capture, mocap*), o różnym stopniu szczegółowości odwzorowania ruchu (ubrania *mocap*, kontrolery Wii, Kinect). Rozwiązania tego typu mogą być z powodzeniem wykorzystywane do budowania systemów wirtualnej rzeczywistości 3D. Również oferta rozwiązań do rejestracji wideo 3D wzrosła znacząco w ostatnich kliku latach. Dostępne są zarówno urządzenia dla użytkowników domowych (np. kamery internetowe), jak i profesjonalne systemy do rejestracji materiałów w wysokich rozdzielczościach (Full HD, 2K).

Akwizycja wideo 3D dokonywana jest najczęściej za pomocą dwóch zsynchronizowanych ze sobą kamer, zamontowanych w specjalnym jarzmie 3D (ang. *rig 3D*). Stosowane są dwa rodzaje takich jarzm:

1. równoległe (ang. *side-by-side*) - kamery umieszczane są obok siebie z zachowaniem równoległości ich osi optycznych do osi optycznej akwizycji,
2. prostopadłe (ang. *beam splitter*), czyli z rozdziałem toru optycznego poprzez zastosowanie półprzepuszczalnych luster - oś optyczna przynajmniej jednej z kamer jest prostopadła do osi optycznej akwizycji.

Dostępne są także rozwiązania zintegrowane (w ramach jednej obudowy), w których konstrukcji stosuje się zasadniczo trzy podejścia:

1. niezależna optyka i dwie matryce,
2. wspólna optyka i dwie matryce,
3. wspólna optyka i jedna matryca.

Kamery, w których część optyczna jest zintegrowana w jednym korpusie obiektywu, są nazywane cyklopami. W przypadku rozwiązań zintegrowanych zakres zmiany odległości (tzw. baza) oraz konwergencji pomiędzy osiami optycznymi jest bardzo mały lub w ogóle niedostępny. Dlatego kamery zintegrowane mogą być stosowane do akwizycji w bardzo ograniczonym zakresie odległości, a oferowane w nich rozdzielczości nie przekraczają HD.

Dla obu rodzajów reprezentacji treści 3D stosowane są metody kodowania, które z jednej strony służą do usunięcia z nich nadmiarowej informacji, z drugiej

strony usprawniają przesył i prezentację treści. Dla materiałów filmowych jest to: stosowanie strumieniowego charakteru zapisu oraz wprowadzenie metod skalowania wielkości przesyłanego strumienia. Dla scen i obiektów 3D stosowane są metody „progresywnego uszczegółowiania", polegające na przesyłaniu informacji niezbędnych do odtworzenia widoku widzianego tylko z określonego punktu obserwacji.

Ze względu na użyte formy zapisu treści 3D można stosować różne technologie do jej wizualizacji. W przypadku reprezentacji rastrowej dostępne są systemy stereoskopowe i autostereoskopowe. Pierwsze wymagają wyposażenia każdego z obserwatorów w urządzenie, które zapewnia separację obrazów z prezentowanej stereopary odpowiednio dla lewego i prawego oka przy współpracy z właściwym urządzeniem, które wyświetla obraz. Do rozwiązań tego typu można zaliczyć gogle 3D (wyświetlanie i separacja następuje w tym samym urządzeniu), telewizory i monitory 3D oraz systemy do projekcji 3D. Dodatkowo rozwiązania te można podzielić na aktywne, które wymagają synchronizacji indywidualnego urządzenia obserwatora z urządzeniem wyświetlającym i pasywne. W systemach pasywnych, zamiast okularów „migawkowych" zsynchronizowanych z urządzeniem wyświetlającym, są stosowane okulary z filtrami (np.: polaryzacyjnymi lub barwnymi). W przypadku rozwiązań autostereoskopowych obserwator nie musi być wyposażany w żaden dodatkowy osprzęt. Wykorzystywane są w tym przypadku zjawiska załamania światła przy przejściu przez płaszczyzny lentikularne lub bariery paralaksy. Obecnie stosowane są w niektórych telewizorach i monitorach 3D, lecz ze względu na brak zachowania ciągłości punktów obserwacji rozwiązania te są uznawane za gorsze od stereoskopowych pod względem doświadczanej jakości prezentacji 3D (QoE 3D).

W przypadku reprezentacji wektorowej zwiększa się grupa technik wizualizacji 3D. Można do nich zaliczyć technologie wolumetryczne lub pseudoholograficzne, które służą do prezentacji obiektów 3D w naturalnej przestrzeni tak, aby mogły być obserwowane równocześnie z wielu punktów widzenia przez wielu obserwatorów. Obecnie są to głównie rozwiązania o charakterze laboratoryjnym lub prototypowym. Z tego powodu do wizualizacji 3D informacji o takiej reprezentacji stosuje się przede wszystkim te same techniki, jak dla reprezentacji rastrowej. Jednak łatwość modyfikacji ich przestrzennego opisu umożliwia wizualizację w szerokim zakresie w systemach projekcji z ekranami cylindrycznymi, sferycznymi czy też w tzw. jaskiniach 3D. Ponadto oprócz wizualizacji niematerialnej reprezentacja wektorowa pozwala na „zmaterializowanie" obiektów z wykorzystaniem różnych technologii druku 3D.

Dostępność urządzeń do akwizycji i wizualizacji 3D zarówno do zastosowań profesjonalnych, jak dla użytkowników domowych może wpływać na pojawianie się nowych lub szybki rozwój już istniejących usług sieciowych, w których wykorzystywane jest przetwarzanie i udostępnianie tego typu treści.

3 Usługi dostarczania treści xK i 3D

W zależności od reprezentacji treści 3D usługi związane z ich przesyłaniem dzieli się na dwa rodzaje:

1. Strumieniowanie,
2. Dostarczanie interaktywne.

Pierwszy rodzaj jest związany z reprezentacją rastrową treści. W tym przypadku usługi te są rozwinięciem klasycznych usług strumieniowania, od których wymaga się zapewnienia wyższych przepływności (w przypadku treści xK) oraz ewentualnie zastosowania nowych metod kodowania dla treści 3D (przykładowo Profil High Stereo z norm H.264/MPEG-4 AVC, opisany w rozszerzeniu dotyczącym wielowidokowego kodowania wideo - MVC). Maksymalna rozdzielczość jest określana na etapie przygotowywania treści. Usługi tego typu charakteryzują się dosyć niskim poziomem interaktywności. Chociaż, prowadzone ostatnio prace rozwojowe nad różnymi technologiami związanymi z prezentacją strumieni wideo (takie jak: video click, hypervideo) mogą zmienić ten obraz.

Drugi rodzaj usług dotyczy wektorowej reprezentacji treści 3D. Opracowane do tej pory formaty bazujące na takiej reprezentacji, najczęściej nie są dostosowane do wykorzystania w interaktywnych aplikacjach sieciowych. W tego typu aplikacjach do pełnej rekonstrukcji sceny i obiektów znajdujących się na niej wymagane byłoby przesyłanie znaczących ilości danych. W praktyce część z tych danych podczas wizualizacji nie musi być potrzebna, np. z powodu przesłaniania się obiektów z punktu wiedzenia obserwatora. Dlatego w celu ograniczenia ruchu w sieci i wykorzystania lokalnych zasobów, w interaktywnych aplikacjach 3D stosuje się metody kodowania pozwalające na realizację progresywnego przesyłu danych zależnego od punktu widzenia obserwatora na scenie oraz podejmowanych prze niego akcji. Maksymalna rozdzielczość jest określana na etapie wizualizacji treści.

3.1 Usługi bazujące na strumieniowaniu

Ze względu na kierunkowość przesyłu strumienia multimedialnego i możliwość sterowania źródłem jego nadawania można wyróżnić następujące klasy usług bazujące na strumieniowaniu:

- konferencyjne - przesył dwukierunkowy pomiędzy wieloma uczestnikami z brakiem możliwości sterowania,
- dystrybucyjne - przesył jednokierunkowy
  - na żądanie - możliwość sterowania (np.: wstrzymywanie, przewijanie),
  - emisyjne - brak możliwości sterowania.

Usługi konferencyjne są wykorzystywane do transmisji na żywo (ang. *live*). W tym przypadku treść trafia do systemu strumieniowania bezpośrednio ze źródła rejestrującego, następnie jest kodowana w czasie rzeczywistym i przesyłana

przez sieć do odbiorców. Dla transmisji na żywo wymagane jest źródło (np. kamera) umożliwiające akwizycję i przesyłanie materiału w czasie rzeczywistym. W przypadku usług na żądanie (ang. *on demand*) materiały są wstępnie przygotowane do transmisji i składowane w repozytoriach. Użytkownik ma możliwość wyboru konkretnego materiału z katalogu dostępnych treści i zażądania rozpoczęcia strumieniowania. Usługi emisyjne mogą być realizowane zarówno z wykorzystaniem wcześniej przygotowanych treści, jak i na żywo. Użytkownik nie ma możliwości sterowania transmisją, może ewentualnie wybrać kanał informacyjny. Dla usług emisyjnych ważna z punktu widzenia gospodarowania zasobami sieci jest zdolność tworzenia połączeń w topologii punkt - wiele punktów.

Wymienione usługi są klasycznymi w zastosowaniach systemów multimedialnych. W strumieniach multimedialnych dominującym pod względem wymagań wydajnościowo-technicznych, w tym wartości parametrów QoS dla kanału transmisyjnego, jest strumień wideo. Dla transmisji wideo wymagane są wysokie przepływności, małe opóźnienia i ich ograniczone fluktuacje oraz małe wartości współczynnika utraty pakietów. Ponadto w przypadku transmisji strumieni z treścią 3D w trybie symultanicznym przez dwa kanały jest wymagane zapewnienie synchronizacji pomiędzy strumieniami, co jeszcze w większym stopniu wymusza ograniczenie zmienności opóźnienia.

W ramach projektu Inżynieria Internetu Przyszłości budowana jest aplikacja, która ma realizować strumieniowanie wideo o wysokich rozdzielczościach w dwóch scenariuszach:

1. wideo na żądanie,
2. emisja na żywo.

Biorąc za punkt odniesienia standard RFC 4594 „Configuration Guidelines for DiffServ Service Classes" [3] oraz rekomendację ITU.T G.1010 [4] możemy określić podstawowe parametry QoS dla tych dwóch typów transmisji. Usługę wideo na żądanie można zaliczyć do kategorii „Multimedia streaming" wymagającej utrzymania wartości współczynnika utraty pakietów poniżej 1%, ale jednocześnie tolerującej zmienność opóźnienia, dzięki możliwości ich kompensacji z użyciem buforów. W tym przypadku akceptowalne jest również opóźnienie rozpoczęcia transmisji na poziomie 5-10s. Emisję na żywo można zaliczyć do grupy „Broadcast video" nakładającej większe wymagania głównie na wartości parametrów opóźnienia w sieci. Często aplikacje tego typu są wrażliwe na zmienność opóźnienia za względu na wykorzystanie bardzo małych buforów lub nawet całkowity brak buforowania. Przy transmisji na żywo użytkownik jest w stanie zaakceptować opóźnienie rozpoczęcia transmisji ok. 1s i artefakty obrazu wynikające z wartości współczynnika utraty pakietów na poziomie do 2%.

Oba typy transmisji strumieniowej treści o wysokiej rozdzielczości wymagają, zgodnie ze standardem DCI, przepustowości sieci na poziome do 250Mbit/s dla treści 4K (500Mbit/s dla 4K3D) [5]. Elastyczność algorytmów kompresji pozwala jednak na zredukowanie strumieni do wartości ok. 80-100Mbit/s i wysłanie materiałów 4K3D w kanale 250Mbit/s.

Odpowiednim rozwiązaniem dla takich scenariuszy ma być opracowywany w ramach projektu IIP jeden z Równoległych Internetów (ang. *Parallel Internet*,

PI), którego koncepcja bazuje na komutacji łączy i ma być realizowana jako sieć z komutacją strumieni danych (ang. *data streaming switching*, PI-DSS). Główną cechą różnicą komutację strumieni danych od komutacji łączy jest brak wymogu zapewnienia synchronizacji w sieci. Celem PI-DSS jest zapewnienie kanałów transmisyjnych o wysokich przepływnościach, przy niskim poziomie utraty pakietów dla strumieni danych o stałej szybkości bitowej.

Aplikacja do strumieniowania działająca w PI-DSS przed rozpoczęciem transmisji będzie wysyłać do sieci komunikat z żądaniem zestawienia kanału dla transmisji. Podstawowymi parametrami, przesłanymi w takim żądaniu, będą identyfikatory węzłów końcowych strumienia oraz parametry transmisji. Żądanie aplikacyjne zostanie przesłane do Adaptera - modułu, który będzie odpowiedzialny za translację aplikacyjnego protokołu strumieniowania (np. RTP) na jednostki transmisyjne w sieci z PI-DSS. Adapter, korzystając z predefiniowanego kanału zarządzania infrastrukturą, wyśle w pierwszym kroku żądanie zestawienia kanałów transmisyjnych dla aplikacji. Mechanizmy zarządzania siecią zestawią kanał o zadanych parametrach i prześlą informację zwrotną do Adaptera, który poinformuje aplikację o możliwości rozpoczęcia strumieniowania. Dla każdej transmisji strumieniowej zestawione zostaną dwa kanały. Jeden z nich będzie posiadał przepustowość niezbędną dla transmisji danego rodzaju danych multimedialnych (np. 250Mbit/s dla strumienia 4K), a drugi kanał o minimalnej przepustowości (np. 64Kbit/s) będzie służyć do wymiany informacji z wykorzystaniem aplikacyjnego protokołu zarządzania.

3.2 Usługi bazujące na dostarczaniu interaktywnym

Rosnąca dostępność urządzeń do akwizycji obiektów 3D oraz rozwój oprogramowania do ich syntezy i modelowania, przyczyniają się do popularyzacji zastosowań wykorzystujących treści 3D. Jedną z charakterystycznych funkcjonalności tych zastosowań jest interaktywna prezentacja treści 3D. Aktualnie są już dostępne w Internecie usługi pozwalające na „interaktywne oglądanie" scen i obiektów 3D (np.: obracanie obiektów, płynna zmiana punktu obserwacji). Są także dostępne rozwiązania, w których możliwe jest przemieszczanie się w wirtualnym świecie, lecz wciąż są to rozwiązania ze scentralizowanym źródłem danych bez wizualizacji stereoskopowej.

Naturalnym etapem rozwoju tego typu aplikacji jest umożliwienie użytkownikowi tworzyć lub modyfikować wirtualne sceny poprzez dołączanie do nich obiektów 3D wyszukiwanych na bieżąco w sieci na podstawie zadanych parametrów (np. identyfikatory, meta dane). Odnalezione obiekty można następnie, w zależności od przeznaczenia aplikacji, porównywać, łączyć ze sobą albo też wykorzystywać w kompozycji sceny. Poziom szczegółowości odwzorowania tego samego obiektu może być różny w zależności od obszarów zastosowań, których przykładami mogą być:

- wspomaganie projektowania (architektura, ergonomia stanowisk pracy, mechanika),
- analizy i wizualizacje naukowe (medycyna, archeologia, kryminalistyka),

– zdalne nauczanie,
– wystawy (muzea wirtualne, targi wirtualne).

Również w zależności od przeznaczenia aplikacji w różnym zakresie może być realizowana w niej interaktywność. W rozwiązaniach prostszych wystarczające mogą się okazać: klawiatura i mysz, natomiast w bardziej zaawansowanych, potrzebne mogą być manipulatory 3D (o sześciu stopniach swobody) lub systemy akwizycji ruchu (ręki, głowy, itd.), w tym wykorzystujące sprzężenie zwrotne dla imitacji odczucia dotyku (ang. *haptic devices*).

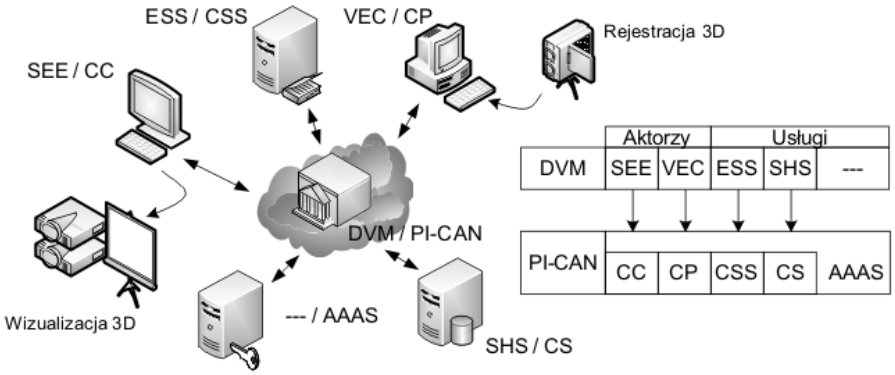
Przykładem aplikacji, której funkcjonalność bazuje na interaktywnym dostarczaniu treści 3D, jest Rozproszone Wirtualne Muzeum (ang. *Distributed Virtual Museum*, *DVM*) tworzone w IITiS PAN w ramach projektu IIP. Zakłada się, że będą z niej korzystały dwa rodzaje użytkowników: kurator wystawy i zwiedzający. Rolą kuratora jest z przygotowanie niezależnych scen z eksponatami i połączenie ich w ekspozycję. Do budowy scen mogą być wykorzystane istniejące obiekty 3D, których lokalizacja w sieci może być rozproszona. Podobnie do przygotowania ekspozycji mogą zostać wykorzystane istniejące sceny, których opisy też mogą mieć rozproszoną lokalizację. Zwiedzający ma możliwość wyszukania interesującej go tematycznie ekspozycji, a następnie interaktywnego zwiedzania scen, które ją tworzą. Sceny i mieszczące się na niej obiekty mają być wyszukiwane i pobierane w zależności od akcji podejmowanych przez użytkownika, w szczególności w zależności od położenia jego punktu widzenia w stosunku do sceny.

Do realizacji takiego scenariusza zostały zdefiniowane cztery komponenty programowe:

– Serwer Ekspozycji i Scen (ang. *Exhibition and Scen Server* , *ESS*) - zawiera opisy ekspozycji i scen wchodzących w ich skład,
– Serwer Magazynowy (ang. *StoreHouse Server*, *SHS*) - zawiera eksponaty reprezentowane w postaci wektorowej wraz z informacjami uzupełniającymi,
– Kurator Wirtualnej Wystawy (ang. *Virtual Exhibition Curator*, *VEC*) - służy do tworzenia opisów wystaw i scen (w tym wyszukiwania scen i obiektów) oraz do umieszczania ich, wraz z wymaganymi eksponatami na odpowiednich, wyżej wymienionych serwerach,
– Inteligentny Eksplorator Wystawy (ang. *Smart Exhibition Explorer*, *SEE*) - jest oprogramowaniem klienckim pozwalającym użytkownikowi wyszukać interesującą go wystawę, a następnie zwiedzać składające się na nią sceny 3D wizualizowane z wykorzystaniem technologii stereoskopowych. Eksponaty są pobierane z wykorzystaniem technik kodowania progresywnego [6].

Odpowiednim rozwiązaniem dla takiego scenariusza jest wykorzystanie idei sieci świadomej przesyłanej treści (ang. *Content Aware Network* , *CAN*), która w szczególności zapewnia funkcjonalność:

– publikowania,
– replikacji,
– wyszukiwania



Rysunek 1. Odwzorowanie komponentów DVM na aktorów i usługi PI-CAN

– i dostarczania różnych typów treści.

Przykładem takiej sieci jest Równoległy Internet CAN (ang. *Parallel Internet CAN*, *PI-CAN*) opracowywany w ramach projektu IIP.

W PI-CAN wyróżniono następujących aktorów:

– Konsument Treści (ang. *Content Consumer*, *CC*) - użytkownik; ma możliwość wyszukiwania treści oraz ich pobierania w zależności od przyznanych uprawnień.
– Wydawca Treści (ang. *Content Publisher*, *CP*) - twórca i zarządca treścią.

Z punktu widzenia aplikacji dostępne są następujące usługi w PI-CAN:

– Serwer Wyszukiwania Treści (ang. *Content Search Server* , *CSS*) - wyszukuje (lokalizuje) treść w odpowiedzi na zapytanie Użytkownika.
– Serwer Treści (ang. *Content Server*, *CS*) - przechowuje treści otrzymane od Wydawcy Treści i udostępnia je użytkownikowi po zlokalizowaniu.
– Serwer Uwierzytelniania, Autoryzacji, Rozliczania (ang. *Authentication, Authorization, Accounting Server*, *AAAS*) - uwierzytelnia Użytkowników i Wydawców Treści oraz autoryzuje podejmowane przez nich działania.

Odwzorowanie komponentów Rozproszonego Wirtualnego Muzeum na aktorów i usługi PI-CAN zostało zilustrowane na Rysunku 1.

Obecnie architektura PI-CAN jest zorientowana na publikację, wyszukiwanie i dostarczanie treści, dla których wystarczające jest tzw. działanie bezstanowe (ang. *stateless*). Ważne jest wyszukanie treści spełniającej określone kryteria, natomiast ich lokalizacja z punktu widzenia aplikacji może być dowolna. Kolejne wyszukiwania mogą skutkować pobieraniem tej samej treści z różnych Serwerów Treści. Nie ma więc gwarancji, że wyszukiwany obiekt będzie miał tę sama lokalizację, a co za tym idzie nie ma możliwości powiązania z nim pewnego stanu po stronie serwera. Nie udostępnia ona więc w tej wersji bezpośrednio wsparcia

dla aplikacji zorientowanych na usługi inne niż dostarczanie treści w topologii punkt-punkt. Wprowadzenie możliwości wyszukiwania nie tylko treści, ale także usług sieciowych, które będą wspierać działanie aplikacji z uwzględnieniem zmiany stanów (ang. *statefull*) oraz komunikację w topologii punkt - wiele punktów pozwolą na implementację aplikacji o charakterze konwersacyjnym czy multiemisyjnym. Architekturę tego typu można określić jako świadomą lokalizacji usług i przesyłanej treści (ang. *Service and Content Aware Network*, *SCAN*).

4 Nowe media w Równoległych Internetach

Nowe Media, rozumiane jako zbiór technik i form przekazu, oprócz przekazu treści o wysokich rozdzielczościach lub o wektorowej reprezentacji przestrzennej, pozwalają na ich łączne wykorzystanie. Może to być zastosowane do zwiększania realności tworzonych systemów wirtualnej rzeczywistości lub wprowadzania do nich nowych funkcjonalności. Łączenie różnych form przekazu pozwala na realizację zaawansowanych, interaktywnych aplikacji multimedialnych. Przykładem takiej aplikacji może być Rozproszone Wirtualne Muzeum rozbudowane o możliwość:

1. odbioru strumieni multimedialnych,
2. wykorzystania usług konwersacyjnych.

Korzystając z takiej możliwości użytkownik oprócz wirtualnej, tematycznej przestrzeni, zrealizowanej z wykorzystaniem interaktywnego dostarczania treści (na bazie PI-CAN), będzie mógł na przykład odtwarzać strumienie audio z komentarzem do oglądanych scen lub odtwarzać strumienie wideo (także 3D) wzbogacające prezentowane ekspozycje (na bazie PI-DSS).

Natomiast wprowadzenie usług konwersacyjnych pozwoli na współdziałanie grupy podczas korzystania z aplikacji, w tym korzystania z komentarzy i wyjaśnień od wirtualnego przewodnika po wystawie. Organizacja grup i współdziałanie w jej ramach może być realizowane z wykorzystaniem funkcjonalności PI-IPv6 QoS.

5 Podsumowanie

W rozdziale przedstawiono wymagania odnośnie sieci dla strumieniowania treści w wysokiej rozdzielczości i przykład realizacji aplikacji, która będzie wykorzystywać PI-DSS do tego celu. Opisano również funkcjonalność Rozproszonego Wirtualnego Muzeum jako przykład aplikacji do interaktywnego dostarczania treści 3D w środowisku rozproszonym na bazie usług PI-CAN implementowanych w ramach projektu IIP. Przedstawiono propozycję rozbudowy tej aplikacji o dodatkowe funkcjonalności umożliwiające równoczesne wykorzystanie trzech Równoległych Internetów.

Urządzenia dostępne na rynku są przeznaczone zarówno dla użytkowników indywidualnych, jak i dużych grup użytkowników. Rozwój technologii xK i 3D

z jednej strony wpływa na podniesienie jakości takich systemów, z drugiej - służy ich popularyzacji. Istnieje dużo potencjalnych obszarów zastosowań przekazu Nowych Mediów, do których można zaliczyć:

- analizy i wizualizacje naukowe,
- wspomaganie projektowania (architektura, ergonomia stanowisk pracy, mechanika),
- telemedycynę,
- zdalne nauczanie,
- wystawy (muzea wirtualne, targi wirtualne).
- transmisje z wydarzeń kulturalnych i sportowych,
- ochronę dziedzictwa kulturowego,
- rozrywkę, itp.

Zainteresowanie wykorzystaniem takich technologii potwierdza powstawanie kolejnych serwisów Internetowych udostępniających między innymi treści 3D. Ponadto możliwość scalania treści 3D o reprezentacji rastrowej i wektorowej przed ich wizualizacją może przyczynić do wzrostu zainteresowania nowymi usługami i aplikacjami takimi, jak: rzeczywistość rozszerzona (ang. *augmented reality*) czy teleobecność 3D (ang. *3D telepresence*).

Literatura

1. ITU FG-FN OD-72, Draft Deliverable on "Future Networks: Design Goals and Promising Technologies", Rev.1, Październik 2010
2. Zahariadis T., Daras P., Laso-Ballesteros I.: Towards Future 3D Media Internet, NEM Summit 2008, St. Malo, 13-15 Październik 2008
3. Babiarz J., Chan K., Baker F.: RFC 4594 - Configuration Guidelines for DiffServ Service Classes, August 2006
4. ITU-T G.1010 - Series G: TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS. Quality of service and performance. End-user multimedia QoS categories, November 2001
5. Digital Cinema System Specification Version 1.2, Digital Cinema Initiatives, March 2008
6. Skabek K., Winiarczyk R., Sochan A.:Implementation of the Progressive Meshes for Distributed Virtual Museum, Human-Computer System Interaction. Backgrounds and Applications 2, Advances in Soft Computing, Springer 2011

Integracja aplikacji e-zdrowie w sieciach z gwarancją jakości usług

Paweł Świątek[1], Adam Grzech[1], Krzysztof Brzostowski[1], Jarosław Drapała[1], Marek Natkaniec[2]

[1] Instytut Informatyki, Politechnika Wrocławska
[2] Katedra Telekomunikacji, Akademia Górniczo-Hutnicza

Streszczenie W niniejszej pracy przedstawiono koncepcję integracji aplikacji e-zdrowie w równoległym internecie IPv6 QoS. W ogólności zaproponowana koncepcja integracji polega na odpowiedniej sygnalizacji pomiędzy komponentami aplikacji. W pracy przedstawiono ogólną koncepcję realizacji usług e-zdrowie w sieciach z gwarancją jakości usług opartych na architekturach NGN i DiffServ. Przykładem takiej sieci jest równoległy internet IPv6 QoS Systemu IIP. Koncepcja integracji aplikacji została zilustrowana na przykładzie trzech różnych aplikacji e-zdrowie dotyczących zdalnego monitorowania: sportowców, diabetyków i astmatyków.

1 Wprowadzenie

Nowoczesne technologie wymiany informacji umożliwiają wykorzystanie niedostępnych do tej pory sposobów wymiany informacji, zwiększenie jakości przekazu danych a co z tym idzie, zastosowania nowych aplikacji, które nie działały by właściwie przy obecnych metodach komunikacji. W niniejszej pracy omawiane są metody integracji aplikacji e-zdrowie w sieci z gwarancją jakości usług, na przykład w równoległym internecie IPv6 QoS, który jest przykładem zastosowania nowoczesnych metod wymiany danych w Internecie Przyszłości [9]. Rozpatrywane są przykładowe aplikacje e-zdrowie jako ilustracja użycia mechanizmów gwarantowania jakości usług w sieciach opartych na architekturze NGN i DiffServ. Mechanizmy obecne w RI IPv6 QoS umożliwiają aplikacji działanie zgodnie z założeniami, tj. dostarczanie usług o wymaganej jakości w dowolnym momencie.

Integracja poszczególnych komponentów aplikacji e-zdrowie zakłada negocjację parametrów wymiany danych w momencie inicjowania użycia aplikacji na żądanie użytkownika. Ponadto ustala się usługi biorące udział w przetwarzaniu, ustala kolejność ich wywołania oraz konfiguruje ich parametry. Tak dostrojone komponenty aplikacji są gotowe do przetworzenia danych pomiarowych z czujników użytkownika.

Niniejsza praca skomponowana jest następująco. W rozdziale drugim przedstawiono architekturę uniwersalnej platformy komunikacyjnej (UPK) odpowiedzialnej za integrację aplikacji w warstwie usługowej. W kolejnym rozdziale

przedstawiono przykładowe scenariusze użycia aplikacji e-zdrowie ze szczególnym uwzględnieniem przepływu komunikatów pozwalających na pełne wykorzystanie możliwości oferowanych przez sieci z gwarancją jakości usług. W rozdziale czwartym podsumowano dotychczasowe prace nad integracją aplikacji i wskazano kierunki dalszych badań.

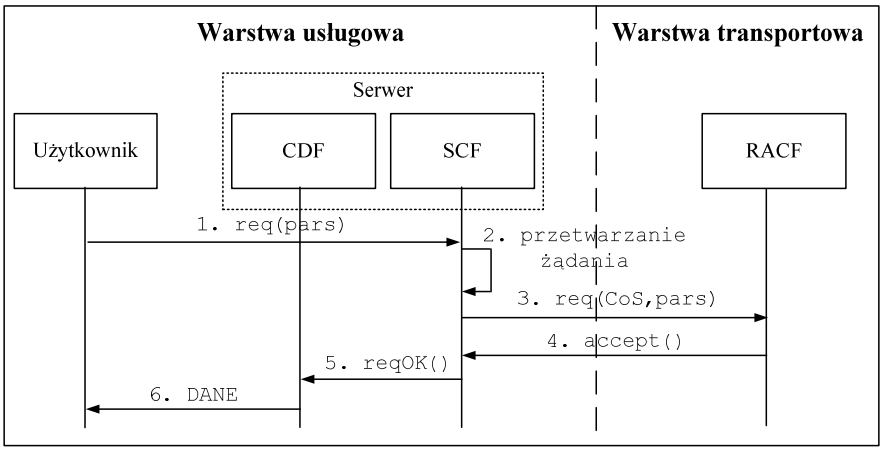
2 Uniwersalna Platforma Komunikacyjna

Warstwa usługowa (ang. *service stratum*) jest elementem architektury Równoległego Internetu IPv6 QoS, w której realizowane są funkcje związane ze sterowaniem przepływem komunikatów sygnalizacyjnych aplikacji, negocjacji wartości parametrów usług specyficznych dla aplikacji oraz przekazywaniu żądań rezerwacji zasobów komunikacyjnych do warstwy transportowej w celu zapewnienia wymaganego poziomu jakości usług. W celu poprawnego zarządzania wykonaniem aplikacji warstwa usługowa RI IPv6 QoS oparta na architekturze NGN [5] korzysta z dwóch modułów: sterowania usługami (ang. *Service Control Functions - SCF*) oraz dostarczania treści (ang. *Content Delivery Functions - CDF*).

Na rysunku 1 przedstawiony został przykładowy scenariusz realizacji usługi komunikacyjnej w RI IPv6 QoS dla potrzeb pewnej aplikacji. W senariuszu tym użytkownik wysyła do serwera aplikacyjnego żądanie (1) o pewne zasoby (np. treść). Żądanie to jest odbierane i przetwarzane (2) przez moduł SCF. Na podstawie informacji zawartej w żądaniu (1) SCF generuje żądanie (3) o zestawienie usługi koniec-koniec i wysyła je do warstwy transportowej przez interfejs SCF-RACF. Po otrzymaniu potwierdzenia rezerwacji zasobów komunikacyjnych (4) moduł SCF przekazuje do CDF informacje niezbędne do rozpoczęcia komunikacji z użytkownikiem (5), po czym CDF rozpoczyna dostarczanie użytkownikowi żądanej treści (6).

Można wyobrazić sobie dużo bardziej skomplikowane scenariusze dla różnego typu aplikacji (np. e-zdrowia [8]), jednakże na ogół wszystkie takie scenariusze można zdekomponować, na podscenariusze polegające na zestawieniu usług koniec-koniec pomiędzy parą węzłów. Istotną zaletą płynącą z wydzielenia warstwy usługowej w architekturze RI IPv6 QoS jest uniezależnienie sposobu żądania usług koniec-koniec od aplikacji, gdyż moduł SCF stojący na styku aplikacji i sieci odwzorowuje żądania aplikacji wynegocjowane dowolnym (być może bardzo złożonym) protokołem na jednolity dla wszystkich aplikacji intefejs SCF-RACF.

Przykładową implementacją warstwy usługowej RI IPv6 QoS jest Uniwersalna Platforma Komunikacyjna (UPK) dla aplikacji e-zdrowie [7]. W UPK zakłada się, że każdy element aplikacji (np.: użytkownik, serwer) biorący udział w wymianie danych jest traktowany jako usługa oraz że każda usługa może być klientem innej usługi. Dodatkowo zakłada się, że żądanie realizacji usługi złożonej (np. zdalnego monitorowania sportowca) zawiera funkcjonalny i niefunkcjonalny opis wymaganej usługi [2]. Opis funkcjonalny dotyczy w szczególności kolejności wykonywania odpowiednich czynności (modułów aplikacyjnych) oraz schematu przepływu danych pomiędzy modułami i jest zdefiniowany jako pewien spój-

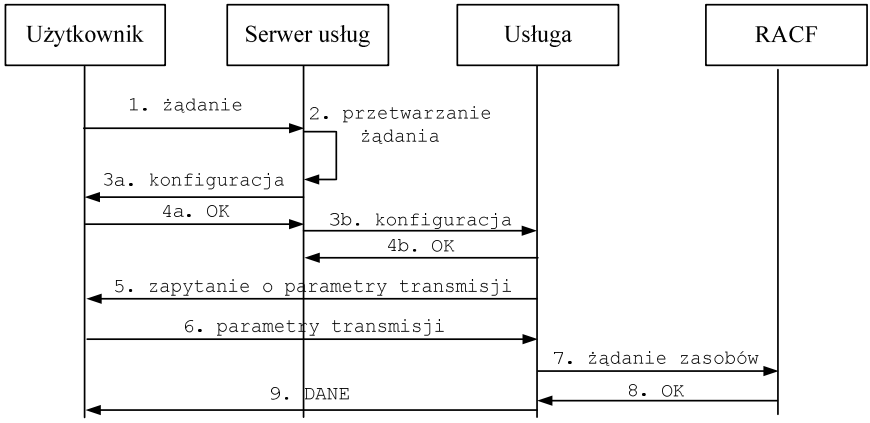


Rysunek 1. Przykładowy scenariusz realizacji usługi komunikacyjnej w RI IPv6 QoS.

ny graf skierowany, którego węzłami są usługi, a krawędzie definiują kolejność wykonania usług [6]. Opis niefunkcjonalny dotyczy na ogół wymaganych wartości parametrów jakościowych (np.: czas odpowiedzi, dostępność, QoS). Żądanie dostarczenia usługi sformułowane w powyższy sposób wysyłane jest do wyróżnionego w systemie serwera usługowego, którego zadaniem jest wybór odpowiednich usług spełniających funkcjonalne i niefunkcjonalne wymagania klienta [4], rezerwacja zasobów komunikacyjnych i obliczeniowych [3] oraz sterowanie wykonaniem usługi poprzez odpowiednią sygnalizację [7].

W celu dostarczania klientom żądanych usług zaproponowano dwuetapową sygnalizację w warstwie usługowej. Pierwszy etap sygnalizacji dotyczy kompozycji złożonej usługi spełniającej funkcjonalne i niefunkcjonalne wymagania klienta z dostępnych w systemie usług oraz poinformowania odpowiednich modułów rozproszonej aplikacji jak i z kim powinny nawiązać komunikację, aby możliwe było zrealizowanie żądanej usługi. Drugi etap sygnalizacji ma na celu ustalenie szczegółów komunikacji pomiędzy każdą parą modułów biorących udział w dostarczaniu usługi klientowi. Parametry te są specyficzne dla funkcjonalności poszczególnych modułów i mogą dotyczyć między innymi: formatu przesyłanych danych, kodeka audio i wideo, wymaganego protokołu transportowego, sposobu szyfrowania, etc.

Na rysunku 2 przedstawiono przepływ komunikatów sygnalizacyjnych dla zestawiania przykładowej usługi polegającej na przesłaniu żądanych danych od usługi do użytkownika. W pierwszym etapie realizacji żądania użytkownika (1) serwer usług przetwarza je (2) w celu skopomonowania usługi spełniającej wymagania użytkownika. Rezultatem kompozycji jest zbiór konktretnych usług dostępnych w systemie rozproszonym, które będą brały udział w realizacji żądania (1). W kolejnym etapie serwer usług konfiguruje wszystkie usługi (w tym aplika-



Rysunek 2. Przykładowy scenariusz realizacji usługi komunikacyjnej w RI IPv6 QoS.

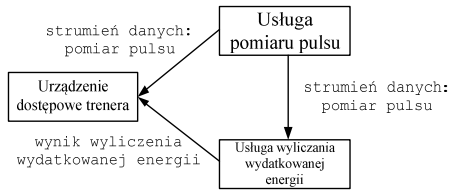
cję kliencką użytkownika) poprzez przesłanie im informacji dotyczących: źródła danych niezbędnych do realizacji usługi, sposobu przetwarza tych danych oraz adresata przetworzonych danych (3a i 3b). Po potwierdzeniu konfiguracji (4a i 4b) każda komunikująca się ze sobą para usług negocjuje wartości parametrów transmisji (5 i 6) w celu rezerwacji niezbędnych zasobów komunikacyjnych (7 i 8). Ostatecznie następuje transmisja żądanych danych (9).

Pierwszy etap sygnalizacji rozpoczynający się wysłaniem przez użytkownika żądania (1), a kończący się konfiguracją wszystkich usług (4a i 4b) realizowany jest z wykorzystaniem protokołu XML RPC. Drugi etap sygnalizacji polegający na uzgodnieniu wartości niezbędnych parametrów komunikacyjnych oraz innych parametrów specyficznych dla aplikacji (5 i 6) jest realizowany z wykorzystaniem protokołu XMPP. Realizacja żądań usług koniec-koniec na styku aplikacji i sieci (7 i 8) odbywa się z wykorzystaniem interfejsu SCF-RACF.

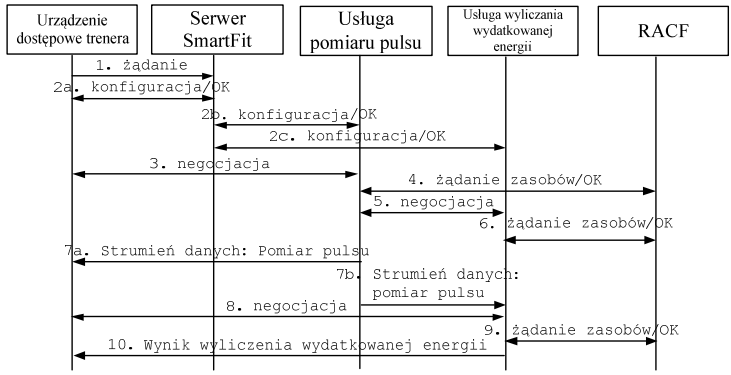
3 Przykładowe zastosowania dla aplikacji e-zdrowie

3.1 Zdalne monitorowanie sportowca - SmartFit

Aplikacja SmartFit przeznaczona jest do wspomagania treningu wytrzymałościowego i technicznego sportowców. Jedną z głównych jej cech jest to, że wspomaga ona zarówno mobilnego sportowca jak i trenera. Wspomaganie użytkownika aplikacji związane jest zarówno z akwizycją danych (zadania po stronie urządzenia dostępowego sportowca) jak również prezentacją wyników ich przetwarzania (zadania zarówno po stronie urządzenia dostępowego sportowca jak i trenera). Inną, równie ważną cechą aplikacji SmartFit jest udostępnienie funkcjonalności umożliwiającej taki dobór usług, by możliwe było skomponowanie usługi złożonej dopasowanej do konkretnego użytkownika i jego wymagań.



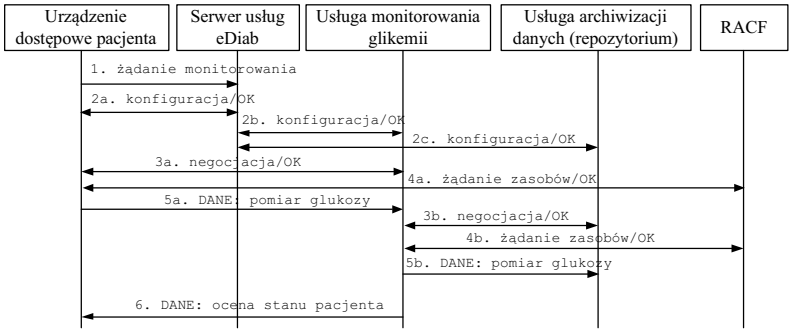
Rysunek 3. Schemat transmisji danych dla usługi monitorowania treningu wytrzymałościowego.



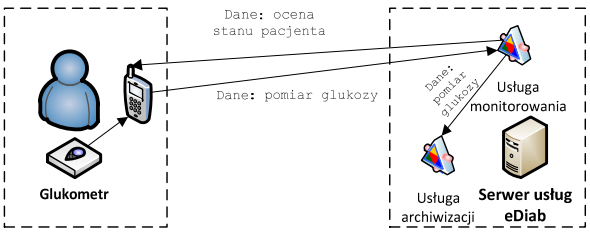
Rysunek 4. Scenariusz usługi monitorowania treningu wytrzymałościowego dla aplikacji SmartFit.

Jednym z elementów treningu sportowego zawodowych sportowców jest monitorowanie treningu wytrzymałościowego. Ma ono na celu oszacowanie intensywności treningu poprzez ciągły pomiar pulsu oraz wyliczenie wydatkowanej energii podczas wysiłku fizycznego. Na rysunku 3 przedstawiono główne elementy wchodzące w skład procesu monitorowania tj. usługa pomiaru pulsu, usługa wyliczania wydatkowanej energii oraz urządzenie dostępowe użytkownika (trenera). Usługa pomiaru pulsu przesyła dane pomiarowe do urządzenia dostępowego użytkownika w celu ich prezentacji na bieżąco oraz do usługi wyliczającej, która gromadzi je a następnie wylicza na ich podstawie wydatkowaną podczas treningu energię. Wyniki działania drugiej usługi są dostępne po zakończeniu treningu.

Na rysunku 4 opis usługi monitorowania został uzupełniony o komunikaty sygnalizacyjne, które są niezbędne do zestawienia usługi monitorowania treningu wytrzymałościowego.



Rysunek 5. Scenariusz realizacji podstawowych usług systemu eDiab.



Rysunek 6. Ilustracja poglądowa działania podstawowych usług systemu eDiab.

3.2 Zdalne monitorowanie diabetyka - eDiab

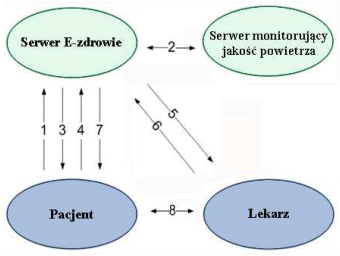
Diagnostyka, leczenie i monitorowanie pacjentów chorych na cukrzycę wiąże się z koniecznością wykonywania regularnych pomiarów glikemii, tj. stężenia glukozy we krwi oraz sporadycznie (dwa razy na rok) pomiarów innych wielkości fizjologicznych. Celem systemu eDiab jest wspomaganie terapii cukrzycy, a w szczególności: przesyłania, przetwarzania i archiwizacji informacji związanych ze stanem pacjenta.

Scenariusz przedstawiony na rysunku 5 ilustruje przepływ informacji związany z realizacją podstawowych zadań systemu eDiab. W etapach (1)-(2) przetwarzane jest żądanie monitorowania oraz konfigurowane są usługi biorące udział w realizacji zadania monitorowania. W etapach (3a)-(3b) usługi monitorowania glikemii i archiwizacji negocjują wartości parametrów transmisji z urządzeniem dostępowym pacjenta oraz ze sobą. Pomiary wykonywane przez glukometr wyposażony w interfejs bluetooth przesyłane są automatycznie do urządzenia dostępowego pacjenta. Trafiają one stąd do usługi monitorowania glikemii (4a-5a) oraz do usługi archiwizacji danych (4b-5b). Usługa monitorowania dokonuje prostej oceny stanu pacjenta, wyznaczając odchylenie zmierzonego stężenia glukozy od wartości należnej i wysyła odpowiednie powiadomienie do urządzenia dostę-

powego (6). Na rysunku 6 przedstawiono ilustrację omówionego scenariusza. Do systemu eDiab wpływają dane pomiarowe z glukometru, a miejscem docelowym tych danych i wyników ich przetwarzania jest urządzenia dostępowe pacjenta oraz repozytorium danych, dostępne za pośrednictwem usługi archiwizacji.

3.3 Zdalne monitorowanie astmatyka - Astma

Astma należy do grupy chorób, które są bardzo trudne lub wręcz niemożliwe do wyleczenia. Terapia wiąże się ze stałym monitorowaniem stanu pacjenta, informowaniem o zagrożeniach zaostrzenia się stanu choroby oraz odpowiednim zażywaniem leków. Właściwy monitoring stanu pacjenta może prowadzić do znacznej poprawy komfortu życia pacjenta. Dzięki odpowiednim serwerom e-zdrowie, lekarze specjaliści mają dostęp do aktualnych danych pacjenta, a także do historii jego choroby. Postawiona na podstawie zebranych danych diagnoza może być bardziej dokładna, a samo leczenie znacznie wydajniejsze. Możliwość całodobowej konsultacji z lekarzami przy użyciu telefonu komórkowego, laptopa lub dowolnego komputera powoduje zwiększenie komfortu i poczucia bezpieczeństwa pacjenta. Taka forma komunikacji umożliwia również nadzorowanie przez lekarza samodzielnych testów wykonywanych przez pacjenta.

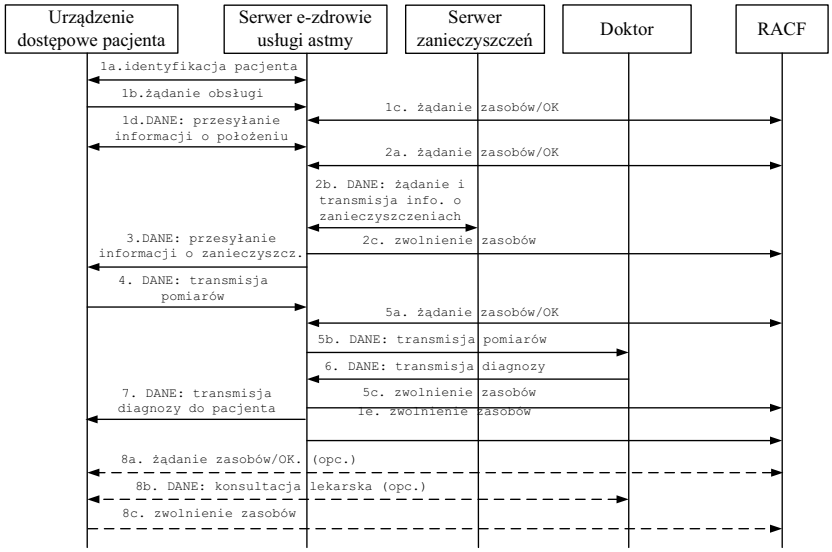


Rysunek 7. Ilustracja poglądowa działania podstawowych usług systemu astmy.

Scenariusz przedstawiony na rysunku 7 pokazuje typowy przepływ informacji związany z realizacją podstawowych zadań systemu nadzoru chorych na astmę. W kroku (1) wykonywana jest identyfikacja pacjenta wraz z określeniem jego położenia przy użyciu systemu GPS. W kroku (2) z serwera zanieczyszczenia powietrza, który znajduje się najbliżej miejsca położenia pacjenta pobierane są informacje o wielkościach zanieczyszczeń najważniejszych parametrów (NOx, COx, SO2, PM10, PM2.5, itp.). Są one również przesyłane w kroku (3) do urządzenia dostępowego pacjenta. Następnie, pacjent przy użyciu dostępnych urządzeń pomiarowych (np. spirometru lub pikflometru) dokonuje właściwych pomiarów swojego stanu zdrowia. Wyniki zostają odpowiednio w krokach (4) oraz (5) przesłane do serwera e-zdrowie oraz do lekarza specjalisty. Dwa kolejne kroki (6) i (7)

pozwalają na przesłanie diagnozy od lekarza do serwera e-zdrowie oraz samego pacjenta. Opcjonalnym krokiem (8) jest przeprowadzenie konsultacji między lekarzem a pacjentem.

Warto również zaznaczyć, że w przypadku monitorowania astmy wyróżnia się zwykle tzw. monitorowanie 'ciągłe', wykonywane przez dłuższy okres czasu oraz 'na życzenie', wykonywane zwykle rano oraz wieczorem, a także w przypadku każdego pogorszenia się stanu pacjenta. Rysunek 8 przedstawia przykładową wymianę komunikatów sygnalizacyjnych w teście 'na życzenie', niezbędnych do zestawienia usługi monitorowania astmy.



Rysunek 8. Scenariusz realizacji podstawowych usług systemu monitorowania astmy.

4 Podsumowanie

W pracy zaprezentowano ogólną koncepcję integracji aplikacji e-zdrowie w równoległym internecie IPv6 QoS tworzonym w ramach Systemu IIP (Inżynieria Internetu Przyszłości). Koncepcja integracji oparta jest na wprowadzeniu wspólnych mechanizmów sygnalizacji niezbędnych do zestawiania, konfigurowania, zarządzania oraz udostępniania usług e-zdrowie, a także do żądania rezerwacji zasobów komunikacyjnych dla potrzeb aplikacji. W przedstawionym podejściu dla potrzeb sygnalizacji wykorzystano mechanizmy kompozycji usług złożonych znanych z architektury zorientowanej na usługi (ang. *Service Oriented Architecture - SOA*) [4]. Zaproponowano dwuetapową sygnalizację, w której zadaniem

pierwszego etapu, realizowanego z wykorzystaniem protokołu XML-RPC, jest konfiguracja komponentów aplikacji dostarczających poszczególne funkcjonalności usługi e-zdrowie. Zadaniem drugiego etapu sygnalizacji, realizowanego z wykorzystaniem protokołu XMPP, jest przede wszystkim negocjacja odpowiednich wartości parametrów dostarczanej usługi (np.: protokół komunikacyjny, format danych, kodeki audio i wideo, etc.).

Przedstawiona koncepcja uniwersalnej platformy komunikacyjnej nie jest dedykowana jedynie dla aplikacji e-zdrowie. Przedstawione mechanizmy mogą być z powodzeniem wykorzystane w zadaniach dostarczania dowolnych usług aplikacyjnych. Z uwagi na wysokie wymagania dotyczące jakości usług w krytycznych scenariuszach zastosowań usług e-zdrowie właśnie te aplikacje zostały wykorzystane w celu prezentacji możliwości i zalet równoległego internetu IPv6 QoS.

Literatura

1. Bernet, Y. et al., An Informal Management Model for Diffserv Routers, Internet RFC 3290, May 2002.
2. Grzech, A., Rygielski, P., Świątek, P.: Translations of service level agreement in systems based on service-oriented architectures. Cybernetics and Systems, vol. 41, no. 8, 610-627, 2010.
3. Grzech, A., Rygielski, P., Świątek, P.: QoS-aware infrastructure resources allocation in systems based on service-oriented architecture paradigm. Proceedings of 6-th Working Conference on Performance Modeling and Evaluation of Heterogeneous Networks, Zakopane, Poland, 2010. pp. 35-48
4. Grzech, A., Świątek, P.: Modeling and optimization of complex services in service-based systems, Cybernetics and Systems, vol. 40, no 8, pp. 706 - 723, 2009
5. ITU-T Rec. Y.2012, Functional requirements and architecture of next generation net-works, (04/2010).
6. Rygielski, P., Świątek, P.: Graph-fold: an efficient method for complex service execution plan optimization, Systems Science vol. 36, no.3, pp. 25 - 32, 2010.
7. Świątek, P. et al., Specification of Universal Communication Platform for e Health services, IIP project Consortium, February 2011.
8. Świątek, P., Rygielski, P., Grzech, A.: Resources management and services personalization in Future Internet applications, Proceedings of the First European Teletraffic Seminar, [ETS 2011], February 14-16, 2011, Poznań, Poland, s. 108-112.
9. Tarasiuk, H. et al., Future Internet Architecture - state of the art and requirements - first version, IIP project Consortium, August 2010.

System zdalnego nadzoru pacjentów chorych na astmę

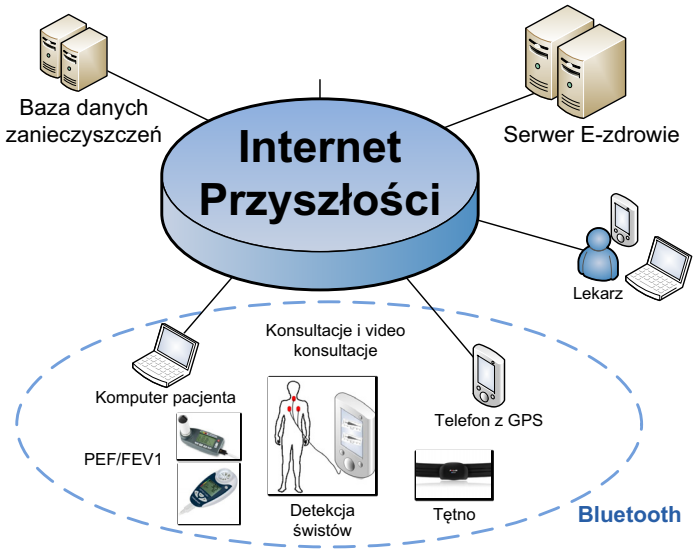
Marcin Wiśniewski, Tomasz Zieliński, Marek Natkaniec, Janusz Gozdecki, Krzysztof Łoziak, Katarzyna Kosek-Szott, Szymon Szott

Katedra Telekomunikacji, Akademia Górniczo - Hutnicza w Krakowie

Streszczenie Systemy zdalnego nadzoru chorych na choroby przewlekłe zaczynają odgrywać znaczącą rolę w procesach leczenia pacjenta. Właściwy monitoring pacjenta chorego na astmę powoduje praktycznie całkowity zanik choroby, a co za tym idzie znaczne poprawienie komfortu życia pacjenta. W rozdziale tym przedstawiony został system do zdalnego monitoringu chorych na astmę. Opisana została metodologia pozyskania sygnałów potrzebnych do zdalnej diagnozy, serwer e-zdrowie oraz sposób przyłączenia systemu do sieci Internetu Przyszłości.Przedstawiona została również przykładowa sekwencja przepływu danych dla IPv6-QoS podczas testu pacjenta.

1 Wprowadzenie

Systemy zdalnego nadzoru chorych na choroby przewlekłe zaczynają odgrywać znaczącą rolę w procesach leczenia pacjenta. Głównym zadaniem takich systemów, jest zebranie jak największej liczby informacji o stanie zdrowia chorego i przesłanie ich do zewnętrznych serwerów, które przetwarzają i przechowują potrzebne informacje. Lekarz specjalista może zdalnie przeprowadzić diagnozę danego pacjenta analizując aktualne dane lub historię choroby zapisaną na serwerze. W szczególnych przypadkach zalecenia mogą być wystawiane automatycznie, bez udziału lekarza. Współczesne systemy e-zdrowie potrafią kontrolować choroby serca [1], cukrzycę [2] oraz schorzenia płuc. Niestety systemy skupiające się na chorobach serca oraz cukrzycy są znacznie popularniejsze niż systemy dotyczące chorób płuc. Astma jest specyficzną chorobą, bardzo trudną lub wręcz niemożliwą do wyleczenia. Terapia powiązana jest z monitoringiem, informowaniem pacjenta o zagrożeniach zaostrzenia się stanu oraz stosownym zażywaniu leków. Właściwy monitoring pacjenta powoduje praktycznie całkowity zanik choroby, a co za tym idzie znaczne poprawienie komfortu życia pacjenta. W artykule przedstawiony został system do zdalnego monitoringu chorych na astmę. Opisana została metodologia pozyskania sygnałów potrzebnych do zdalnej diagnozy, serwer e-zdrowie oraz sposób przyłączenia systemu do sieci Internetu Przyszłości. Przedstawiona została również przykładowa sekwencja przepływu danych dla IPv6-QoS podczas testu pacjenta.



Rysunek 1. Architektura systemu do monitoringu chorych na astmę.

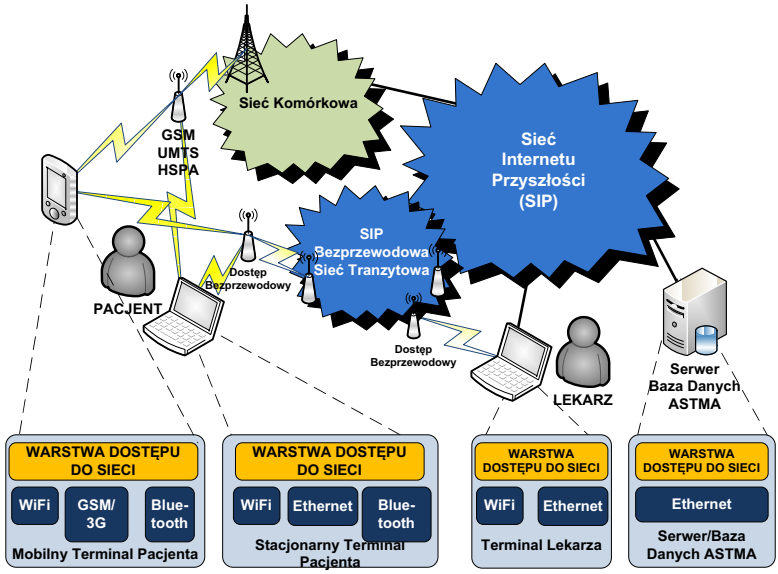
2 Architektura systemu

Sercem systemu jest smartfon **HTC HD2 mini**. Zapewnia on komunikację na trzech różnych płaszczyznach tworząc terminal mobilny pacjenta (PMT - *Patient Mobile Terminal*). Pierwsza z nich to komunikacja pomiędzy pacjentem a lekarzem. Urządzenie umożliwia przeprowadzenie konsultacji, wideo konsultacji lub przeprowadzenie wywiadu lekarskiego o stanie zdrowia pacjenta. Kolejną warstwą komunikacyjną urządzenia to zbieranie i wstępne przetworzenie informacji o pacjencie z różnych urządzeń zewnętrznych (np. Spirometr, pulsometr). Telefon tworzy z zewnętrznymi urządzeniami sieć WPAN (*Wireless Personal Area Network*) opartą na technologii Bluetooth. Przesył informacji odbywa się automatycznie lub na żądanie pacjenta (w zależności od rodzaju urządzenia). Ostatnia warstwa to transmisja zebranych informacji o pacjencie do serwerów e-zdrowie. Urządzenie za pomocą sieci GSM lub Wi-Fi przekazuje niezbędne do diagnozy dane o pacjencie. Dodatkowo wykorzystano przenośny komputer klasy PC jako urządzenie wspomagające. Zadaniem komputera jest zwiększenie wygody użytkowania aplikacji kosztem mobilności i jest to tak zwany terminal stacjonarny pacjenta (PST - *Patient Stationary Terminal*). Strukturę systemu przedstawia rys. 1. Aby zautomatyzować proces zbierania informacji o pacjencie, stworzona została specjalna aplikacja - AstmaApp. Pierwszym etapem pozyskiwania informacji o pacjencie jest wywiad lekarski. Informacje o stanie zdrowia zbierane są na podstawie ustandaryzowanych pytań z punktowanymi odpowiedziami. Po wywiadzie lekarskim, na podstawie zebranych punktów, pacjent otrzymuje informacje o stanie zdrowia za pomocą ustandaryzowanych bramek

w kolorach: zielony - dobra kontrola, żółty - łagodne zaostrzenie, czerwony - poważne zaostrzenie. Kolejnym krokiem jest pomiar wydolnościowy płuc. Jednym z ważniejszych pomiarów podczas leczenia i kontroli astmy są badania spirometryczne lub pikflometryczne. Obserwowanie takich wielkości jak PEF (*Peak Expiratory Flow*)/FEV1 (*Forced Expiratory Volume in 1 second*) pozwala na dokładną ocenę stanu zdrowia pacjenta. W systemie użyty został przenośny spirometr **Spirobank II** z interfejsem Bluetooth i USB oraz Pikflometr **Asma-1 USB Vitalograph** z możliwością podłączenia do komputera poprzez USB. Aplikacja AsthmaApp pozwala na ręczne lub automatyczne wprowadzenie potrzebnych do monitorowania wielkości. Zebranie informacji o kondycji płuc, uaktualnia informacje o stanie zdrowia pacjenta (bramki). Dodatkowym urządzeniem monitorującym stan zdrowia jest analizator osłuchu płuc. Po badaniach spirometrem lub pikflometrem pacjent nagrywa osłuch klatki piersiowej za pomocą 4 kanałowej nagrywarki z interfejsem Bluetooth. Pacjent przesyła dane do telefonu za pomocą aplikacji. Zebrane w ten sposób dane o pacjencie przesyłane są do zewnętrznego serwera e-zdrowie gdzie zostają poddane analizie. Kolejnym bardzo ważnym czynnikiem jest jakość powietrza. Pacjenci chorzy na astmę niekorzystnie reagują na wzrost zanieczyszczeń powietrza [3]. System zdalnego monitoringu astmy pobiera informacje o zanieczyszczeniach w okolicy pacjenta z zewnętrznego monitoringu zanieczyszczeń [4]. Na podstawie pozycji GPS telefonu, informacje o jakości powietrza są przekazywane pacjentowi oraz zapisywane w bazie danych e-zdrowie. W przypadku ciężkiej astmy oraz zaostrzenia stanu zdrowia, pacjent ma możliwość wykorzystania dodatkowego monitoringu tętna za pomocą pulsometru **Bluetooth Polar WearLink+**. W zaostrzonym stanie, pacjent może dostać ataku astmy w sytuacjach stresowych lub podczas wysiłku [5] . Monitoring tętna informuje pacjenta o możliwym zagrożeniu. Cała procedura monitorowania pacjenta oparta jest na codziennych testach zdrowia. Testy wykonywane są rano i wieczorem, ale oprócz tego, pacjent może wykonać test "na żądanie". Każdy krok testu powoduję aktualizację wyniku stanu zdrowia pacjenta (bramek), przy czym testy wydolnościowe płuc mają największy priorytet. W każdym momencie pacjent również może sprawdzić stan zanieczyszczenia powietrza w swoim otoczeniu, lub przeprowadzić konsultację głosową lub wideo-kolnsultacje.

3 Połączenia sieciowe

System zdalnego monitoringu wykorzystuje różne techniki przesyłu danych. Jak już to zostało wspominane wcześniej, do połączeń urządzeń zewnętrznych w sieci personalnej wykorzystywana jest technologia Bluetooth, dodatkowo do urządzenia stacjonarnego (PST), niektóre urządzenia zewnętrzne można połączyć poprzez USB. Transmisja informacji od pacjenta do serwera e-zdrowie w przypadku mobilnego terminalu (PMT) odbywa się za pomocą sieci Wi-Fi lub GSM/3G. Wybór konkretnej technologii zależeć będzie od dostępności sieci w otoczeniu pacjenta i odbywać się będzie automatycznie, bez ingerencji użytkownika. Bazową technologią przesyłu danych jest Wi-Fi, która umożliwi bezpośredni dostęp do

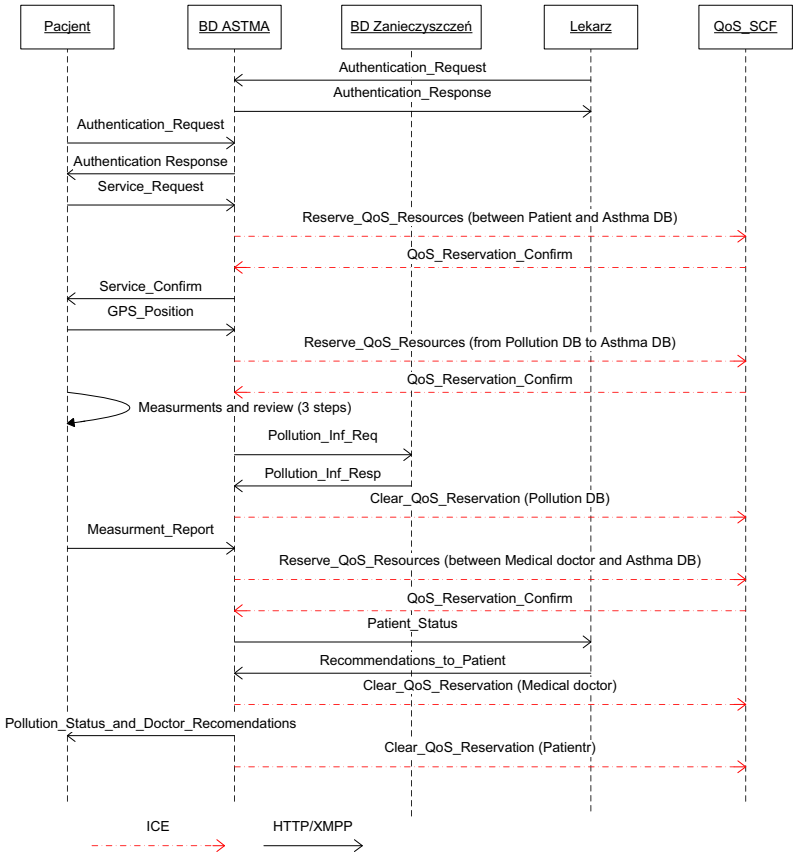


Rysunek 2. Konfiguracja interfejsów sieciowych aplikacji.

Internetu Przyszłości i jego funkcji. W przypadku gdy pacjent nie będzie w zasięgu punktu dostępowego Wi-Fi Internetu Przyszłości, może być użyta każda inna sieć, włącznie z GSM/3G. Użycie innej sieci niż sieć Internetu Przyszłości może spowodować, że nie wszystkie usługi monitoringu astmy będą dostępne. W przypadku urządzenia stacjonarnego (PST), połączenie również dokonuje się automatycznie spośród sieci Wi-Fi, Ethernet oraz GSM/3G. Dostęp poprzez Internet Przyszłości zapewnia odpowiedni QoS oraz wymagane bezpieczeństwo dla aplikacji monitoringu astmy. Serwer e-zdrowie jest podłączony do sieci poprzez sieć Internetu Przyszłości wykorzystując interfejs Ethernet. Rozwiązanie to pozwoli na w pełni wykorzystanie wszystkich usług monitoringu astmy jak i Internetu Przyszłości. Na rys. 2 przedstawiony jest ogólny schemat konfiguracji interfejsów sieciowych aplikacji.

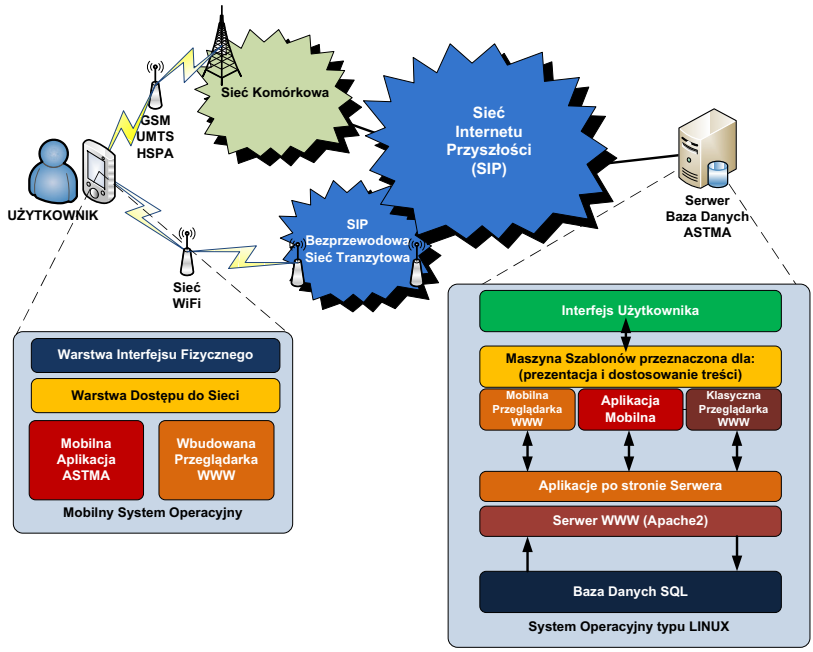
4 Zapewnienie jakości usług sieciowych w Internecie Przyszłości

Sieci wykorzystywane do realizacji transmisji danych pomiędzy elementami systemu zdalnego nadzoru nad stanem pacjenta chorego na astmę powinny gwarantować jakość przesyłania danych. Jest to ważny aspekt systemu, ponieważ poprawność przekazu danych może mieć znaczenie dla zdrowia, a nawet życia pacjenta. Implementacja systemu zdalnego nadzoru pacjentów chorych na astmę w sieci Internetu Przyszłości umożliwia wykorzystanie wbudowanych me-



Rysunek 3. Sygnalizacja dla testów porannych w systemie zdalnego nadzoru stanu pacjenta.

chanizmów umożliwiających zapewnienie jakości usług przesyłu danych. Rys. 3 przedstawia przykładową sekwencję wymiany informacji z siecią zbudowaną na równoległym Internecie typu IPv6-QoS, implementującym różne klasy gwarancji jakości transmisji danych. W rozwiązaniu przyjęto scentralizowane rozwiązanie polegające na rezerwacji zasobów sieci centralnie, przez moduł w serwerze astmy (moduł DB ASTMA na rys. 3). Rezerwacje wykonywane są za pomocą interfejsu do modułu QoS-SCF (QoS *Service Control Function*), który to moduł jest odpowiedzialny za komunikacje z aplikacjami w warstwie sterowania sieci IPv6-QoS. W pokazanym scenariuszu dla zrealizowania scenariusza testów porannych lub wieczornych wymagane są trzy rezerwacje pomiędzy głównymi elementami systemu, przy czym zasoby pomiędzy serwerem ASTMA a lekarzem jest zestawiana tylko w przypadkach nagłych. Spośród 6 klas usług zdefiniowanych w Internecie typu IPv6-QoS system nadzoru nad pacjentami chorymi na astmę korzysta



Rysunek 4. Architektura systemu bazy serwera e-zdrowie.

z trzech klas usług: "Low Latency Data" dla transmisji danych i "Telephony" oraz "RT Interactive" dla audio i video konsultacji.

5 Bazy danych

Na drugim końcu aplikacji znajduje się serwer bazy danych e-zdrowie. W bazie danych przechowywane są podstawowe informacje o pacjencie (tj. Imię, Nazwisko, płeć, wiek, wzrost, waga lub BMI) Na bazie takich informacji określane są odpowiednie progi dla pomiarów wydolnościowych płuc. Głównym zadaniem serwera jest zbieranie, zapisywanie i przetwarzanie danych. Serwer zapewnia przechowywanie wyników dziennych badań wraz z nagraniami audio osłuchu płuc. Aby ułatwić dostęp do odczytu danych przez lekarza, serwer wykorzystuje protokoły HTTP/HTTPS z autentykacją. Takie rozwiązanie pozwala na przesył informacji nie tylko poprzez sieci Internetowe, ale również za pośrednictwem prostych wiadomości SMS generowanych przez system. Głównymi komponentami serwera e-zdrowie jest warstwa do rozpoznawania rodzaju połączenia (aplikacja telefonu, przeglądarka mobilna, standardowa przeglądarka), mechanizm doboru wyświetlania zawartości w zależności od rodzaju połączenia, aplikacja do przetwarzania danych, baza danych SQL do przechowywania danych pacjenta oraz serwer WWW kompatybilny z Apache2.Oprócz funkcji przechowywania, doboru

zawartości i dostarczania treści, najważniejszym modułem serwera e-zdrowie jest aplikacja umożliwiająca weryfikację i autentykację użytkownika (zawartość wyświetlana jest w zależności od użytkownika tj. pacjent czy lekarz), preprocesing danych przed ich zapisaniem do bazy danych oraz statystyczna analiza danych i detekcja stanów krytycznych pacjenta. Architektura systemu bazy serwera e--zdrowie przedstawiona jest na rys. 4

6 Wnioski

W rozdziale tym przedstawiono innowacyjny system do monitoringu pacjentów chorych na astmę. Rozwiązanie to idealnie wpisuje się w wymagania jakie stawia standardowa kuracja chorych na tę przewlekłą chorobę. Dane o pacjencie, które zbierane są przez system mogą w pełni opisać aktualny stan zdrowia pacjenta. Poprzez zdalny monitoring pacjent ogranicza wizyty u lekarza do niezbędnego minimum, co korzystnie wpływa na komfort jego życia. Pełne spektrum badań, od wywiadu lekarskiego, poprzez badania wydolnościowe płuc do analizy osłuchu płuc daje kompletny obraz stanu zdrowa pacjenta, a dodatkowy monitoring jakości powietrza oraz tętna pozwalają pacjentowi na wcześniejszą reakcję i zapobieganie niebezpiecznych ataków. Dodatkową cechą jaką wyróżniają tę aplikacje spośród innych aplikacji e-zdrowie jest fakt, że system ten od początku projektowany jest pod wymagania Internetu Przyszłości. System ten w pełni wykorzystywać będzie protokół IPv6, ale również działać będzie na starszym typie sieci. System tworzony jest przy konsultacji z pulmonologami z krakowskiego szpitala im. Jana Pawła II i tam też będzie testowana jego pierwsza jaki kolejne wersje.

Literatura

1. F. Sufi, I. Khalil, "An Automated Patient Authentication System for Remote Telecardiology", International Conference Series on Intelligent Sensors, Sensor Networks and Information Processing (2008) 279 284
2. Seto, E.; Istepanian, R.S.H.; Cafazzo, J.A.; Logan, A.; Sungoor, A.; "UK and Canadian perspectives of the effectiveness of mobile diabetes management systems" Engineering in Medicine and Biology Society. Annual International Conference of the IEEE 2009 , Page(s): 6584 - 6587
3. http://www.who.int/topics/asthma/en/
4. http://monitoring.krakow.pios.gov.pl/iseo/
5. Gregerson, M.B. "The historical catalyst to cure asthma". Adv. Psychosom. Med. 24, 16-41 2003

Biblioteka Cyfrowa Pacjenta: zarządzanie danymi medycznymi w Internecie Przyszłości

Michał Kosiedowski, Cezary Mazurek, Aleksander Stroiński

Poznańskie Centrum Superkomputerowo-Sieciowe

Streszczenie Postępująca informatyzacja społeczeństwa oraz powszechny dostęp do informacji mają ogromny wpływ także na możliwości realizacji ochrony zdrowia. Nowoczesne technologie pozwalają w jeszcze większym stopniu włączyć pacjenta w proces profilaktyki oraz leczenia. W tym celu rozwijane są złożone systemy teleinformatyczne, których główną funkcją jest integracja danych medycznych zapisanych w różnego rodzaju repozytoriach. Jednym z najważniejszych wyzwań związanych z realizacją tego typu usług jest zapewnienie efektywnego zarządzania rozproszonymi danymi medycznymi. Rodzi to liczne wymagania dla warstwy sieciowej, które wydają się niezbędne do budowy nowoczesnych i innowacyjnych aplikacji medycznych. W referacie przedstawiono wyniki prac mających na celu opracowanie i wdrożenie prototypowej aplikacji Biblioteki Cyfrowej Pacjenta zanurzonej w Internecie Przyszłości integrującej dane medyczne z systemów obsługujących rekordy pacjenta.

1 Wprowadzanie

Istniejące systemy informatyczne odpowiedzialne za zbieranie, przechowywanie oraz przetwarzanie informacji medycznych, to nie tylko systemy zlokalizowane w zorganizowanej służbie zdrowia lub innych instytucjach opieki zdrowotnej, ale także specjalne aplikacje medyczne działające na użytek pacjenta, w jego domu lub najbliższym otoczeniu. Dane medyczne wykorzystywane dla różnych celów przechowuje się również w bazach wiedzy, bazach genów i innych systemach informacji biomedycznych. Każdy z tych systemów gromadzi dane, które mogą przyczynić się do poprawy oceny stanu zdrowia pacjenta i stać się częścią rozproszonej biblioteki cyfrowej budowanej wokół jego potrzeb. Integracja takich rozproszonych danych medycznych wraz z mechanizmami inteligentnego zarządzania nimi, pozwoliłaby na pokonanie wielu barier w rozwoju nowoczesnej medycyny i jednocześnie pozwoliłaby na podniesienie jakości i zwiększenie skuteczności procesu leczenia pacjenta, a także profilaktyki. Do tej pory dominował nurt wykorzystania publicznej, globalnej sieci Internet, jako bazy dla rozwiązania powyższego problemu. Z tego powodu projektanci medycznych bibliotek cyfrowych natrafiają na liczne przeszkody związane z przetwarzaniem danych w środowisku rozproszonych systemów medycznych zarządzanych przez różne jednostki administracyjne lub różne osoby gdyż Internet nie zapewnia wystarczających mechanizmów w tym zakresie. Budowa cyfrowych bibliotek pacjenta

połączona z równoległym rozwojem sieci, która będzie wspierać proces tworzenia nowoczesnych systemów medycznych oraz proces inteligentnego zarządzania informacją medyczną jest obecnie wyzwaniem w zakresie budowy aplikacji z obszaru e-Zdrowie i jest przedmiotem wielu prac badawczo-rozwojowych.

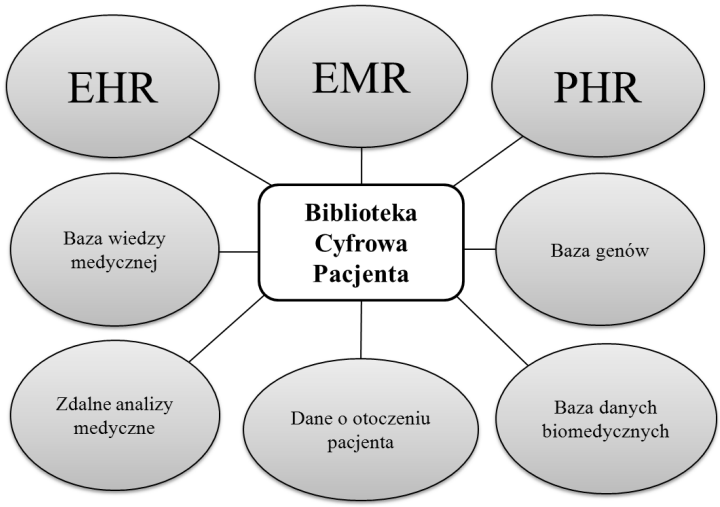
W referacie przedstawiona została koncepcja realizowanej w ramach projektu Inżynieria Internetu Przyszłości (IIP) Biblioteki Cyfrowej Pacjenta, wspierającej proces zarządzania danymi medycznymi w środowisku nowoczesnych technologii sieciowych. IIP ma zapewnić nową architekturę sieci, która spotyka się z wymaganiami Biblioteki Cyfrowej Pacjenta. pracy sformułowano te wymagania oraz wynikające z tego koncepcje dotyczące procesu zarządzania, architektury aplikacji oraz scenariuszy użycia. W rozdziale drugim opisane zostały rodzaje systemów przechowywania danych medycznych oraz zdefiniowano zadania, jakie powinna spełniać Biblioteka Cyfrowa Pacjenta. W rozdziale trzecim przedstawiono proces zarządzania danymi medycznymi w sieci e-Zdrowie oraz przedstawiono, w jaki sposób Biblioteka Cyfrowa Pacjenta może wspierać ten proces. Czwarty rozdział zawiera opis architektury oraz scenariuszy użycia budowanej w ramach projektu IIP aplikacji Biblioteki Cyfrowej Pacjenta. Rozdział piąty zawiera wnioski oraz podsumowuje dyskusję w zakresie zarządzania danymi medycznymi w Internecie Przyszłości.

2 Systemy przechowywania danych medycznych

Dane medyczne przeważnie przechowywane są w postaci rekordów pacjenta, czyli struktur zawierających różnego rodzaju informacje o pacjencie. Możemy wyróżnić kilka rodzajów systemów przechowywania takich rekordów. Dane medyczne wytworzone na terenie danej placówki służby zdrowia (np. szpital, przychodnia) archiwizowane są w systemach EMR (ang. Electronic Medical Records). Obecnie podejmuje się wiele wysiłków w celu stworzenia systemów EHR (ang. Electronic Health Records) obejmujących dane zapisane w wielu systemach EMR. Zgodnie z definicją HIMSS (Healthcare Information and Management Systems Society) „*rekordy EHR są trwałym elektronicznym zapisem informacji o zdrowiu pacjenta, wygenerowanym podczas jednej lub kilku wizyt pacjenta w placówce służby zdrowia. (...) EHR ma zdolność do generowania kompletnego zapisu wizyty klinicznej pacjenta*" [1]. Innymi słowy, pojęcie EHR zakłada integrację danych pochodzących z różnych specjalistycznych repozytoriów informacji medycznych w sieci [2]. Zdefiniowano kilka standardów określających strukturę informacji medycznej zapisanej w EHR, do których należą profile IHE, HL7 CDA i EN 13606. Obecnie w wielu krajach i regionach trwają prace nad wdrożeniem systemów EHR, np. w Danii [3], Finlandii [4], Wielkiej Brytanii [5] i Nowej Zelandii [5]. Innym rodzajem systemów przechowywania danych pacjenta są systemy PHR (ang. Personal Health Record), które „*pozwalają ludziom na zarządzanie informacjami na temat własnego zdrowia w sposób elektroniczny*" [6]. Główną różnicą między systemami EHR i EMR, a systemami PHR jest zastosowane podejście do zarządzania danymi medycznymi. W przypadku pierwszych systemów to personel medyczny jest odpowiedzialny za wprowadzenie danych do systemu i ich dalsze użycie. W przy-

padku systemów PHR to pacjent decyduje, jakie dane zostaną zarchiwizowane w systemie oraz komu zostaną one udostępnione w procesie monitorowania bądź leczenia.

Systemy EHR i PHR przechowują rekordy pacjenta. Jednak informacje w nich zgromadzone mogą okazać się niewystarczające do otrzymania poprawnej i pełnej oceny stanu zdrowia pacjenta oraz możliwości wypracowania odpowiedniej strategii postępowania w przypadku problemów ze zdrowiem. Istotne dla zdrowia pacjenta mogą się okazać dane dotyczące środowiska, warunków życia pacjenta oraz jego codziennej aktywności. Wiele tych informacji jest skorelowanych z różnymi schorzeniami, np. klimatyzacja może mieć wpływ na astmatyków, a natężenie światła słonecznego może mieć związek z chorobami skóry [7]. Dodatkowo w celu otrzymania pełnej wiedzy medycznej należy wszystkie zebrane dane zestawić z danymi referencyjnymi przechowywanymi w różnego rodzaju bazach danych biomedycznych, bazach genów (np. DrugBank [8]) oraz repozytoriach wiedzy medycznej EBM (ang. Evidence Based Medicine – np. PubMed [9], Cochrane Library [10], TRIP [11]). Dane te mogą być analizowane przez człowieka lub w sposób automatyczny przy użyciu specjalizowanych usług analitycznych.



Rysunek 1. Biblioteka Cyfrowa Pacjenta integruje dane z różnych systemów medycznych.

Podsumowując dane medyczne oraz wszystkie inne informacje związane ze zdrowiem pacjenta rozproszone są w różnych systemach. Rodzi to potrzebę integracji tych danych w celu otrzymania lepszego obrazu stanu zdrowia pacjenta oraz potrzebę efektywnego zarządzania tymi danymi w rozproszonym środowisku sieciowym. W tym celu opracowuje się spersonalizowane biblioteki cyfrowe

pacjenta (Rys. 1). Realizacja tego typu rozwiązań stanowi wyzwane nie tylko dla warstwy aplikacji, ale również dla warstwy sieciowej. W ramach projektu IIP zespół Poznańskiego Centrum Superkomputerowo-Sieciowego opracowuje prototyp Biblioteki Cyfrowej Pacjenta zanurzonej w Internecie Przyszłości. Biblioteka ta ma za zadanie wspierać zarządzanie danymi medycznymi zgromadzonymi w rozproszonej sieci wyżej opisanych systemów.

3 Zarządzanie danymi medycznymi

Ze względu na ogromne spektrum wymagań, jakie stawiają nowoczesne aplikacje medyczne, istnieje konieczność dokładnego zdefiniowania funkcjonalności oraz architektury Biblioteki Cyfrowej Pacjenta w aspekcie zarządzania danymi konkretnej osoby. Zarządzanie jest procesem monitorowania i kontrolowania dostępu do danych medycznych. W podejściu spersonalizowanym przez zarządzanie rozumie się proces przetwarzania danych konkretnego pacjenta od momentu ich powstania do momentu użycia ich w procesie leczenia lub procesie profilaktyki. Dlatego niezbędnym elementem Biblioteki Cyfrowej Pacjenta musi być mechanizm zarządzania, który będzie wspomagał szereg czynności związanych z tworzeniem oraz dalszym użyciem danych medycznych (Rys. 2). Do takich czynności należy bezpieczna akwizycja danych pochodzących z osobistych urządzeń medycznych. Zbieranie danych może odbywać się w domu pacjenta (lokalnie) oraz poza nim (zdalnie) np. w drodze do pracy. Miejsce zbierania danych może mieć wpływ na ich efektywne i bezpieczne przesłanie oraz zapisanie w Bibliotece Cyfrowej Pacjenta, dlatego pacjent powinien określać miejsca archiwizowanych danych. Ponadto pacjent powinien mieć możliwość decydowania, do których danych będzie miał dostęp personel medyczny bądź członkowie jego rodziny. Taki proces udostępniania danych i zarządzania uprawnieniami do nich powinien uwzględniać możliwość dostępu do danych przez automatyczne systemy wspomagające monitorowanie i analizę danych. Dodatkowo, w skład czynności zarządzania wchodzą różnego rodzaju sytuacje alarmowe oraz stany zagrożenia zdrowia pacjenta, w których dane powinny zostać natychmiast udostępnione konkretnym osobom niosącym pomoc medyczną lub członkom rodziny pacjenta.

Rysunek 2. Zarządzanie danymi medycznych to szereg czynności związanych z przetwarzaniem danych.

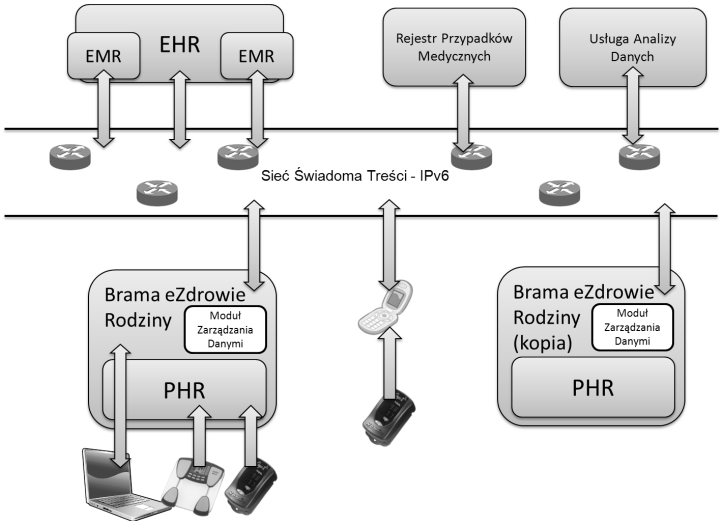
Tak zdefiniowany proces zarządzania danymi medycznymi, uwzględniający integrację treści z rozproszonych systemów medycznych za pośrednictwem Biblioteki Cyfrowej Pacjenta nie jest obecnie w łatwy sposób realizowalny. Zasadnicze powody to brak kompatybilności między systemami z uwagi na różnorodność formatów danych oraz interfejsów dostępowych w systemach medycznych, ograniczenia warstwy sieciowej związane z efektywnością udostępniania danych medycznych, brakiem warstwy semantycznej związanej z treścią medyczną w sieci i bezpieczeństwem danych medycznych w publicznej sieci Internet oraz aktualna infrastruktura sieci Internet, która nie w pełni uwzględnia wyżej wymienione aspekty zarządzania danymi. Dlatego istnieje potrzeba wykorzystania przez nowoczesne aplikacje medyczne funkcjonalności, jaką dostarczyć może warstwa sieciowa Internetu Przyszłości.

4 Biblioteka Cyfrowa Pacjenta w Internecie Przyszłości

W celu usprawnienia integracji i udostępniania danych medycznych w sieci, proponuje się, aby Biblioteka Cyfrowa Pacjenta oparta była o sieć świadomą treści (ang. CAN – Content Aware Network). Dane medyczne w takiej sieci mogą być traktowane w sposób podobny do treści multimedialnych. Dzięki temu sieć może decydować, czy i w jaki sposób należy przekazać dane treści lub z którego źródła pobrać dane. Dodatkowo do takiej sieci można wprowadzić dodatkowe informacje semantyczne wykorzystujące standardy medyczne, medyczną terminologię i normy medyczne np. ICD-9 i ICD-10. Informacje semantyczne ułatwią proces integracji danych medycznych w rozproszonej Bibliotece Cyfrowej Pacjenta. Architekturę Biblioteki Cyfrowej Pacjenta opartej na sieci świadomej treści przedstawiono na rysunku 3.

Prototyp aplikacji realizowany w ramach projektu IIP łączy systemy EHR, PHR oraz inne repozytoria wiedzy medycznej i uwzględnia scenariusze związane z usługami zdalnej analizy danych. EHR może zostać zasymulowany przy użyciu otwartego oprogramowania, a za repozytorium wiedzy medycznej posłuży rejestr przypadków medycznych Wielkopolskiego Centrum Telemedycyny. Kluczowym elementem realizowanej aplikacji jest Brama eZdrowie Rodziny. Brama ta projektowana jest jako oparte o specjalizowane oprogramowanie urządzenie zlokalizowane w domu pacjenta i stanowiące jego punkt dostępu do sieci e-Zdrowie. Na Bramie uruchomiona zostanie, obok interfejsów komunikacji z siecią świadomą treści, lokalna instancja repozytorium PHR i moduł komunikacji z osobistymi urządzeniami medycznymi. Pacjent za pośrednictwem Bramy eZdrowie Rodziny będzie mógł zarządzać danymi rozproszonymi w sieci. Ponadto, aplikacja będzie wykorzystywać urządzenia mobilne typu smartfon w celu akwizycji danych poza domem pacjenta i przesłania ich do Bramy eZdrowie Rodziny.

Aplikacja realizować będzie dwa podstawowe scenariusze obejmujące zarządzanie danymi medycznymi w Bibliotece Cyfrowej Pacjenta (Rys. 4). Pierwszy scenariusz dotyczy zbierania danych z osobistych urządzeń medycznych oraz ich archiwizacji w PHR na Bramie eZdrowia Rodziny. Należy zaznaczyć, ze zbieranie danych medycznych może odbywać w dowolnym miejscu i w dowolnym



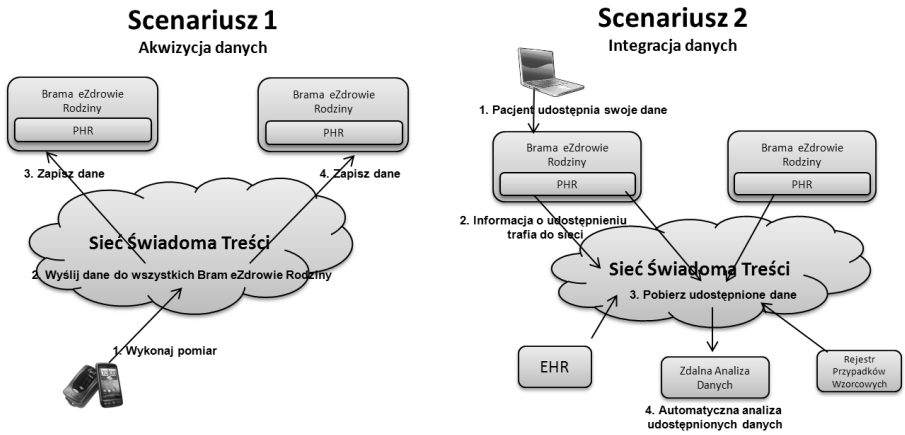
Rysunek 3. Architektura Biblioteki Cyfrowej Pacjenta.

czasie, w domu pacjenta, bądź poza nim. Drugi scenariusz dotyczy dostępu do zgromadzonych danych i zarządzania uprawieniami do treści medycznych.

W celu realizacji tych scenariuszy przyjmuje się następujące wymagania dotyczące sieci:

- aplikacja oraz wszystkie jej elementy będą korzystać z dedykowanej dla dziedziny e-Zdrowia sieci wykorzystującej protokół IPv6, który pozwoli w elastyczny sposób adresować urządzenia medyczne, usługi, itp.
- sieć powinna traktować każdy element aplikacji jako źródło udostępniania danych medycznych (w tym urządzenia mobilne, które służą do przechwytywania danych z osobistych urządzeń medycznych) (scenariusz 2)
- sieć powinna zapewnić możliwość kopiowania wybranej przez pacjenta treści medycznej pomiędzy wskazanymi lokalizacjami (scenariusz 1)
- sieć powinna przechowywać informacje o udostępnionych treściach medycznych i uniemożliwić nieautoryzowany dostęp do treści (same treści nie są przechowywane w sieci)
- sieć powinna w automatyczny sposób dostosować format przekazywanej treści do możliwości odbiorcy (automatyczna translacja jednego formatu medycznego w drugi, np. HL7 2.x i HL7 3.x)
- sieć powinna być świadoma rodzaju przesyłanych danych i zapewnić najlepszą technikę transmisji tych danych [12]

Powyższe wymagania spełniają sieci świadome treści zbudowane na bazie protokołu IPv6. Sieć taka jest tworzona w ramach projektu „Inżynieria Internetu Przyszłości" jako element nowoczesnej infrastruktury sieciowej umożliwiającej realizację efektywnych aplikacji sieciowych.



Rysunek 4. Scenariusze dla Biblioteki Cyfrowej Pacjenta.

5 Wnioski

W referacie przedstawiono założenia funkcjonalne i architekturę Biblioteki Cyfrowej Pacjenta opartej o sieć świadomą treści. Aplikacja ta integruje dane medyczne z rozproszonych i niezależnych systemów medycznych: EHR, EMR, PHR, banków wiedzy medycznej, rejestrów przypadków i innych źródeł treści medycznych. Dodatkowo, zapewnia ona mechanizmy pozwalające pacjentowi na efektywne zarządzanie własnymi danymi medycznymi zgromadzonymi w Bibliotece.

W ramach prac nad Biblioteką Cyfrową Pacjenta w projekcie IIP zaimplementowano pierwszą wersję Bramy e-Zdrowie Rodziny. W oparciu o interfejsy Bramy, zaimplementowano repozytorium PHR, moduł akwizycji danych z osobistych urządzeń medycznych oraz moduł prezentacji wyników pomiarów. Kolejnym etapem prac będzie integracja aplikacji z mechanizmami sieci świadomej treści. W dalszej części projektu przewiduje się podłączenie pozostałych elementów Biblioteki Cyfrowej Pacjenta, tj. systemu EHR i rejestru przypadków oraz uruchomienie przykładowej usługi analizującej dane. Tak skonstruowane środowisko umożliwi przetestowanie scenariuszy zarządzania danymi medycznymi na bazie systemu IIP. Pozwoli to porównać nową funkcjonalność aplikacji e-Zdrowie, uzyskaną dzięki integracji rozproszonych danych medycznych przy wykorzystaniu mechanizmów systemu IIP, z dotychczasowymi możliwościami Internetu.

Literatura

1. HIMSS, Electronic Health Record (EHR). Healthcare Information and Management Systems Society, http://www.himss.org/ASP/topics_ehr.asp, ostatni dostęp: 29. maja 2011

2. NIH, Electronic Health Records Overview, National Institutes of Health, National Center for Research Resources, http://www.ncrr.nih.gov/publications/informatics/EHR.pdf, ostatni dostęp: 29. maja 2011
3. K. Bernstein et al, Modelling and implementing Electronic Health Records in Denmark, International Journal of Medical Informatics, Vol. 74, No. 2, pp. 213-220, 2005
4. P. Ruotsalainen, P., Regional EHR systems and eAcrhives in Finland. Proceedings of European Federation for Medical Informatics, Special Topic Conference 2007, Birjuni, Chorwacja, pp. 174-179, 2007
5. M.K. Mason, What Can We Learn from the Rest of the World? A Look at International Electronic Health Record Best Practices. http://www.moyak.com/papers/best-practices-ehr.html, ostatni dostęp: 29.maja 2011, 2004
6. K. Smolij, K. Dun, Patient Health Information Management: Searching for the Right Model, In Perspect Health Inf Manag, Vol. 3, No. 10, 2006
7. M. Ichihashi et al, UV-induced skin damage, Toxicology, Vol. 189, No. 1-2, pp. 21-39, 2003
8. DrugBank, Open Data Drug & Drug Targer Database, http://www.drugbank.ca/
9. PubMed, http://www.ncbi.nlm.nih.gov/pubmed/
10. Cochrane Library, http://www.cochrane.org/
11. Trip Database, http://www.tripdatabase.com/
12. T. Koponen et al, A data-oriented (and beyond) network architecture, Proceedings of SIGCOMM'07 conference on Applications, technologies, architectures, and protocols for computer communications, Kyoto, Japonia, pp. 181-192, 2007

Przewodnik migracji IPv6 - narzędzie dla administratora sieci SOHO

Krzysztof Nowicki[1], Mariusz Gajewski[2], Wojciech Gumiński[1], Aniela Mrugalska[1], Mariusz Stankiewicz[1], Józef Woźniak[1]

[1] Politechnika Gdańska, Wydział Elektroniki, Telekomunikacji i Informatyki
[2] Instytut Łączności, Państwowy Instytut Badawczy

Streszczenie W pracy przedstawiono założenia i sposób implementacji Przewodnika migracji IPv6, narzędzia wspomagającego administratora sieci w realizacji procesu migracji sieci do protokołu IPv6. Omówiono motywy powstania proponowanego narzędzia. Zaprezentowano architekturę aplikacji. Opisano sposób gromadzenia i zarządzania wiedzą na temat aktualnego stanu sieci przygotowywanej do migracji. Przedstawiono sposób wykorzystania przewodnika.

1 Wstęp

Powszechnie znanym problemem w Internecie jest kończąca się wolna pula adresów IPv4. IANA dysponująca globalną pulą adresów IPv4 oddała 3.02.2011 ostatnie wolne adresy IPv4 [7]. Z kolei pierwszym RIRem, którego zasoby adresowe się skończyły był APNIC [2]. Od 15.04.2011 na terenie Azji nie jest możliwe pozyskanie nowych pul adresów IPv4. Prognozy przewidują, że zasoby pozostałych Regionalnych Rejestrów Internetowych skończą się wciągu najbliższych kilkunastu miesięcy. Dokonywane są różne działania mające na celu opóźnienie tego zdarzenia, w tym tak spektakularne jak próby odzyskiwania bloków adresowych od korporacji i gigantów rynku, które na początku ery Internetu otrzymywały niemalże nielimitowane pule. Działania takie są jednym z powodów wolniejszej, niż było to zakładane, migracji sieci IPv4 do IPv6 [8]. Inną przyczyną wolniejszego niż się spodziewano, procesu migracji do IPv6 jest brak wiedzy wśród administratorów, szczególnie mniejszych sieci [8]. Brak wiedzy wynika z faktu, że z jednej strony na rynku nie jest oferowanych zbyt wiele szkoleń w zakresie migracji do IPv6, a z drugiej strony oferta zwartych materiałów, które by te procesy dla konkretnych rozwiązań sieci IPv4 opisywały jest znikoma. Brakuje również "tanich specjalistów", którzy świadczyli by tego rodzaje usług dla małych sieci, w tym sieci domowych.

Vint Cerf, powszechnie nazywany ojcem Internetu, próbuje uświadamiać, że rozwój biznesu jest możliwy tylko wtedy, gdy istnieje przestrzeń adresowa, w ramach której może być on dokonywany [3]. Obecne lata stają się krytyczne w kontekście migracji do protokołu IPv6. Fakt konieczności dokonania tego procesu jest od wielu lat powszechnie znany i do tej pory był bagatelizowany. W ostatnim czasie jednak nastąpiło wiele zmian w podejściu do tego tematu.

8 czerwca 2011 został ogłoszony Światowym Dniem IPv6 [9]. Był on obchodzony dość spektakularnie. Obok największych dostawców treści internetowych pośród, których wymienić można Google [1], Facebook, Yahoo czy też gigantów w swoich branżach jak Cisco, Juniper i BBC, aktywny udział wzięły także liczne ośrodki naukowe. Na opublikowanie rekordów AAAA zdecydowały się m.in. Harvard University, Universita di Pisa a także polskie uczelnie - między innymi Politechnika Gdańska [17] i Politechnika Wrocławska. Działania takie jak Światowy Dzień IPv6 doskonale dowodzą, że oba protokoły IP pomimo słabej kompatybilności względem siebie są w stanie w sposób prawidłowy działać jednocześnie. Już nie tylko środowiska naukowe są zgodne, że protokół IPv6 [4] jest jedynym długoterminowym rozwiązaniem, które jest w stanie zapewnić dalszy rozwój Internetu. Proces migracji z IPv4 do IPv6 nie jest jednak procesem łatwym.

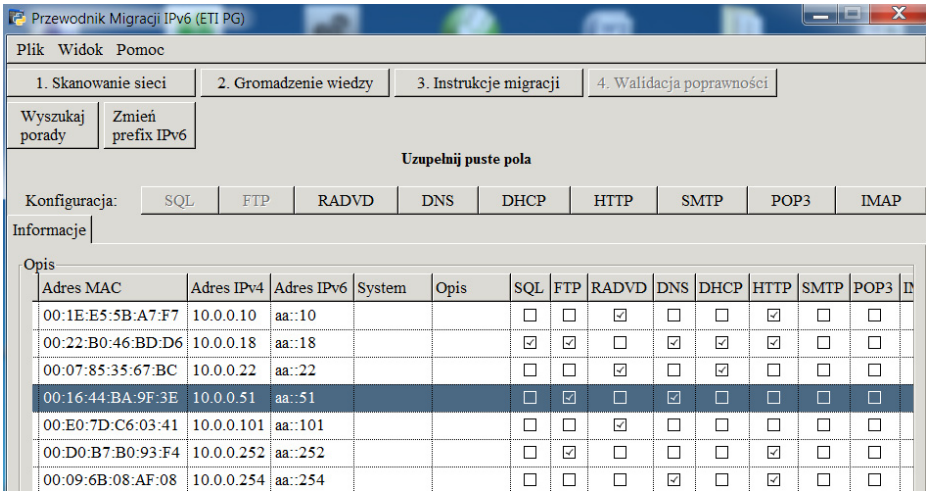
Google, będące liderem w swojej branży, potrzebowało blisko trzech lat, aby zestawić i uruchomić sieć działającą po IPv6 [3]. Jak zatem z migracją do nowego protokołu mają sobie poradzić małe firmy, które nie posiadają w swojej kadrze specjalistów dorównujących wiedzą informatykom z Doliny Krzemowej [15]? Jedną z propozycji ułatwiającej rozwiązanie problemu jest przygotowywana w ramach projektu Inżynieria Internetu Przyszłości [10] aplikacja "*Przewodnik migracji IPv6*" [18], która ma na celu ułatwienie tego procesu. *Przewodnik migracji IPv6* został opracowany ze szczególnym skupieniem na sieciach małych, typu SOHO (ang. Small Office Home Office), w przypadku których osoba zarządzająca siecią nie zawsze posiada ugruntowaną wiedzę teleinformatyczną.

2 Podstawowe założenia Przewodnika migracji IPv6

Zadaniem *Przewodnika migracji IPv6* jest dostarczenie administratorowi sieci wiedzy o czynnościach koniecznych do zmigrowania zarządzanej przez niego sieci IPv4 do IPv6. Migracja w każdej sieci, w zależności od jej budowy i wykorzystywanych zasobów będzie przebiegała w odmienny sposób i wymaga innych działań. Aplikacja *Przewodnik migracji IPv6* winna precyzyjnie dostarczać jedynie potrzebne porady wraz z łatwym dostępem do podpowiedzi zawierających usystematyzowaną wiedzę dotyczącą danego tematu. W tym celu proces migracji zdekomponowano na pojedyncze zadania [18]:

1. analiza i gromadzenie wiedzy o sieci,
2. instrukcje migracji i
3. walidacja poprawności.

Pierwsze zadanie (analiza i gromadzenie wiedzy o sieci) realizowane jest w dwóch krokach: 1. automatyczna analiza sieci oraz 2. "odpytanie" administratora. W celu rozpoznania (analizy) sieci zaimplementowany został zewnętrzny program skanujący zasoby sieci, zaś uzyskane informacje zebrane są w odpowiednim pliku XML. Przy wykorzystaniu automatycznej analizy sieci otrzymujemy informacje dotyczące zainstalowanych systemów na hostach oraz wykorzystywanych usług. Rozpoznawana jest zatem, w sposób automatyczny, budowa i topologia sieci oraz



Rysunek 1. podsumowująca informacje na temat migrowanej sieci

jej elementy składowe. Krok pierwszy nie jest obligatoryjny, choć jego uruchomienie może zaoszczędzić czas administratora ponieważ wiedza uzyskana w tym kroku wykorzystywana jest w następnym kroku - odpytywania administratora, ograniczając liczbę zadawanych mu pytań. Drugi krok obejmuje weryfikację informacji pozyskanych w sposób automatyczny w pierwszym kroku. Informacje na temat używanego w sieci oprogramowania i sprzętu są pozyskiwane bezpośrednio od administratora. Poszczególne pytania kierowane do administratora służą otrzymaniu dodatkowej wiedzy, której nie był w stanie dostarczyć zautomatyzowany proces skanowania sieci. Poszczególne pytania są powiązane ze sobą w taki sposób, aby nie zadawać pytań, gdy wcześniejsze odpowiedzi wskazują, że nie są one w danej sytuacji potrzebne. W wyniku udzielania odpowiedzi możliwe jest także wyświetlenie dodatkowych pytań.

Ostatnim działaniem w tym kroku jest podsumowanie wszystkich informacji uzyskanych w poprzednich krokach w formie zaprezentowania ich w jednym oknie wewnątrz czytelnej tabeli (patrz rysunek 1). W ten sposób możliwa staje się łatwa weryfikacja poprawności wszystkich informacji.

W tabeli przedstawione są informację o adresie IP w wersji czwartej, proponowanym adresie w wersji szóstej zgodnym z ustalonym planem adresacji [6], systemie operacyjnym, adresie MAC danego urządzenia oraz o uruchomionych na nim usługach sieciowych.

Przewodnik migracji IPv6, w drugim zadaniu - migracja, dostarcza administratorowi spisu działań, który należy podjąć aby sieć zaczęła działać w oparciu o protokół IPv6. Jest to wiedza niezbędna do rekonfiguracji urządzeń i usług wykorzystywanych w sieci.

Informacje na temat zmiany ustawień sprzętu i migracji oprogramowania zawierają dane dotyczące konfiguracji w systemach [5], które wskazane zostały

w poprzednich krokach. Przy każdej poradzie oprócz wiedzy wyjaśniającej w jaki sposób przeprowadzić konfigurację, użytkownik otrzymuje także możliwość uzyskania teoretycznego opisu na temat przeprowadzanych w danym momencie czynności. Zapoznanie się z wiedzą na temat sieci IPv6 i podstaw teoretycznych dotyczących konfigurowanych usług następuje niezależnie od dostarczanych użytkownikowi porad odnoszących się do poszczególnych działań i konfiguracji konkretnych usług sieciowych. Przykładowo podczas dokonywania adresacji urządzeń można skorzystać z bazy wiedzy i uzyskać informację o tym, czym są adresy multicastowe IPv6.

Część pól w tabeli (patrz tabela 1 zakłada możliwość edycji. Przykładowo użytkownikowi umożliwiono podanie loginu oraz hasła administratora danego systemu operacyjnego bądź też zmiany proponowanego planu adresacji IPv6. Użytkownik będzie miał też możliwość zmiany sprzętu wykorzystywanego w sieci np. jako router bądź serwer DHCP. Umożliwia to, po uprzednim zatwierdzeniu planu adresacji bądź podaniu własnych adresów, uzyskanie pliku konfiguracyjnego dla niektórych usług sieciowych, które zostały uznane za najbardziej powszechne bądź istotne. W tym wypadku wystarczy tylko skopiować dany plik konfiguracyjny, kolejno podmienić ten plik w danym serwerze i uruchomić na nowo dany proces. Po wykonaniu wszystkich zaleceń administrator ma możliwość przetestowania sieci - zarówno systemów, usług jak również urządzeń pod kątem zgodności z zaleceniami dotyczącymi pracy w sieciach opartych o IPv6. Wyniki weryfikacji poprawności działania sieci mogą zawierać informacje niezbędne do ewentualnego poprawienia konfiguracji [16].

3 Testy dla *Przewodnika Migracji IPv6*

Efektem poprawnego wdrożenia protokołu IPv6 w sieci jest fakt wzajemnej komunikacji na poziomie warstwy sieciowej pomiędzy poszczególnymi urządzeniami. Fakt ten nie oznacza jednak, że migracja danej infrastruktury sieciowej zakończyła się powodzeniem, ponieważ konieczne jest również zapewnienie poprawnej współpracy aplikacji, niezbędnych do funkcjonowania tej sieci. To z kolei jest związane z przeprowadzeniem migracji w kierunku IPv6 również usług sieciowych oraz aplikacji uruchamianych w węzłach sieci. Zawarty z *Przewodniku Migracji IPv6* system porad oraz praktycznych wskazówek wspiera uruchomienie protokołu IPv6 w poszczególnych węzłach, jak również migrację do IPv6 typowych usług sieciowych oraz aplikacji. Z kolei wersja Przewodnika instalowana w dystrybucji LiveCD zawiera również zestaw testów, pozwalający zweryfikować, czy migracja najpopularniejszych usług sieciowych i aplikacji została przeprowadzona w taki sposób, aby w rezultacie uzyskać poprawnie komunikujące się węzły. Zasadniczym założeniem przyjętym przy konstruowaniu scenariuszy testowych, jest możliwość wykonania testów:

- w funkcjonującej sieci bez dokonywania zmian konfiguracyjnych węzłów na potrzeby testów zarówno usług sieciowych jak i aplikacji;
- w taki sposób, jakby do funkcjonującej sieci dołączony został dodatkowy węzeł emulujący określoną funkcjonalność (tester);

Testowana funkcjonalność	**Sposób realizacji testów**
osiągalność IPv6	wykorzystanie narzędzi systemowych - *ping6*
autokonfiguracja	wykorzystanie testów realizowanych na platformie *v6eval*
serwer DHCPv6	sprawdzenie przypisania adresu IPv6 dla testowego węzła w sieci oraz sprawdzenie, czy serwer udostępnia dodatkowe informacje (wykorzystanie testów realizowanych na platformie *v6eval*)
serwer DNS	sprawdzenie reakcji na żądania kierowane do serwera DNS (polecenie *dig*)
serwer WWW	pobranie strony internetowej (polecenie *wget*)
serwery poczty	nawiązanie sesji IMAP i POP3 i oczekiwanie na przekroczenie czasu kontrolnego (moduły *Python*), nawiązanie sesji SMTP i próba przesłania wiadomości z wykorzystaniem błędnego adresu nadawcy (moduł *Python*)
serwer FTP	logowanie do serwera z wykorzystaniem błędnej nazwy użytkownika/hasła (moduł *Python*)
serwer SSH	logowanie do serwera z wykorzystaniem błędnej nazwy użytkownika/hasła (moduł *Python*)

Tabela 1. Zestawienie testowanych funkcjonalności

– bez konieczności interakcji pomiędzy procesem realizującym scenariusz testowy (testerem) oraz użytkownikiem tego narzędzia.

Przyjmując te założenia, sprawdzenie efektów migracji do protokołu IPv6 poszczególnych węzłów sieci IPv6 (zarówno stacji roboczych jak i serwerów) zakłada komunikację pomiędzy testerem oraz danym węzłem sieci IPv6 z wykorzystaniem określonego protokołu (np. http, ftp lub smtp) lub określonego mechanizmu - np. autokonfiguracji. Jeśli testowany węzeł komunikuje się z testerem wykorzystując te protokoły lub mechanizmy oparte na protokole IPv6, wówczas przyjmuje się, że testowane usługi sieciowe oraz aplikacje zaimplementowane w tym węźle poprawnie wykorzystują stos IPv6, choć nie jest to równoznaczne z poprawną ich konfiguracją w zakresie realizowanych przez nie zadań. Takie podejście pozwala stwierdzić, czy usługi sieciowe i aplikacje w danym węźle "współpracują" z wykorzystaniem protokołu IPv6, a jednocześnie realizacja scenariuszy testowych nie wymaga dokonywania zmian konfiguracji węzłów na potrzeby testów.

Zbiór testów jest realizowany w odniesieniu do każdego z węzłów umieszczonego na liście i wyróżnianego za pomocą adresu MAC oraz adresów sieciowych IPv6. Lista węzłów wraz ze specyfikacją adresów fizycznych i sieciowych oraz przypisanych usług stanowi zbiór danych wejściowych dla testera (plik .xml). W zależności od zadeklarowanej dla danego węzła listy usług sieciowych oraz aplikacji skrypt uruchamia odpowiednie scenariusze testowe. Dodatkowo sprawdzana jest również osiągalność węzłów sieci IPv6, które uzyskały adres IPv6 w oparciu o konfigurację statyczną oraz w procesie autokonfiguracji. Zakres sprawdzanych za pomocą tych testów funkcjonalności zestawiono w tabeli 1.

Wynikiem przeprowadzonych testów dla danego węzła jest informacja o jego osiągalności w sieci oraz informacja o wsparciu dla procesu autokonfiguracji. Wynikiem realizacji testów dla usług sieciowych i aplikacji jest z kolei informacja o wspieraniu przez dany węzeł tych usług. W przypadku, gdy test usługi daje wynik negatywny i możliwe jest jednoznaczne wskazanie przyczyny, taka informacja jest również dołączana do listy wyników. Uzyskane wyniki testów dla poszczególnych węzłów są zapisywane w pliku .xml i mogą być wykorzystane przez użytkowników *Przewodnika Migracji IPv6* do dalszej analizy i prezentacji efektów migracji.

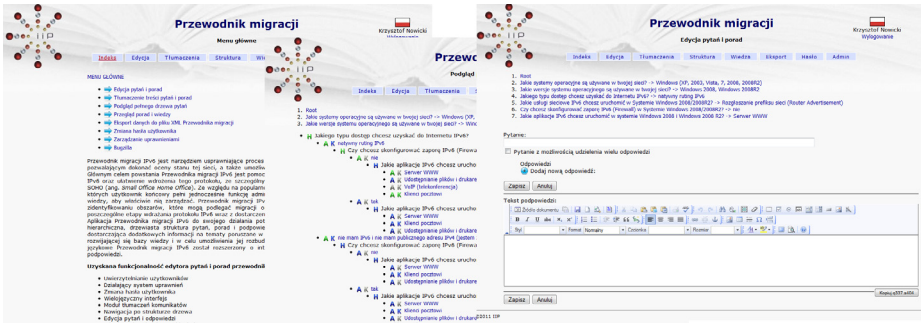
Przewodnik Migracji IPv6 wraz z testami w wersji dla dystrybucji LiveCD bazuje na systemie operacyjnym FreeBSD 8.0 R1 i wykorzystuje interpreter języka Python w wersji 2.6. Ponadto, w niektórych scenariuszach testowych wykorzystywane jest środowisko do realizacji testów IPv6 - v6eval (ver. 3.3.2) opracowane w ramach projektu Tahi [19] jak również narzędzia dostępne w systemie FreeBSD (m.in. polecenie wget, dig) i w postaci modułów Python (m.in.: klient ftp, SMTP, POP3 i IMAP).

4 Koncepcja zarządzania wiedzą

W celu sprecyzowania spisu działań koniecznych do przejścia na IPv6, które będą dopasowane do migrowanej sieci, w *Przewodniku migracji IPv6* zostały zaimplementowane mechanizmy, które umożliwiają automatyczny dobór pytań zależny od udzielanych przez użytkownika odpowiedzi oraz wyświetlenie porad przygotowanych pod kątem konkretnej sieci. W ten sposób możliwe staje się zdefiniowanie poszczególnych kroków migracji, z których można zbudować określony schemat działań. Możliwe jest to dzięki precyzyjnemu dostarczaniu jedynie potrzebnych porad wraz z łatwym dostępem do podpowiedzi zawierających usystematyzowaną wiedzę dotyczącą danego tematu który zapewnia dekompozycję na pojedyncze działania [13].

Zarządzanie wiedzą wykorzystywaną w *Przewodniku migracji IPv6* oparte jest o parsowanie plików XML (ang. Extensible Markup Language), którego zastosowanie pozwala na zachowanie właściwej struktury zależności pytań i porad [11]. Dodatkowo korzystanie z formalnych definicji struktury plików XML zapewnia automatyzację procesu walidacji ich formalnej poprawności i spójności danych. Różnorodność scenariuszy migracji może wymagać ich podziału na wiele plików. Aplikacja zatem jest w stanie otwierać wszystkie znalezione pliki pytań i porad i wykorzystywać zapisane w nich informacje w procesie migracji. Ważnym elementem *Przewodnika migracji IPv6* jest doprowadzenie do propozycji planu adresacji sieci IPv6 bazującego na rozpoznaniu dotychczasowego rozwiązania opartego o dotychczas stosowane IPv4 poprzez analizę budowy sieci.

Konfiguracja podstawowych usług sieci IPv6 obejmuje porady dotyczące konfiguracji routingu IPv6 oraz dostępu do sieci szkieletowej IPv6 zarówno w wersji natywnej jak i za pomocą tuneli takich jak 6in4, 6to4, 6rd albo Teredo. Kluczowe usługi, które są niezbędne do prawidłowego działania sieci takie jak Router Advertisement, DHCPv6, DNS czy NTP zostały omówione bardzo szczegółowo



Rysunek 2. Edytor pytań i porad

z uwzględnieniem systemów używanych w danej sieci. Przedstawiona jest także zmiana konfiguracjci większości popularnych systemów operacyjnych z rodzin Windows, Linux i BSD do pracy w sieciach IPv6 [12].

5 Edytor pytań i porad *Przewodnika migracji IPv6*

Niewątpliwą zaletą omawianego rozwiązania jest otwarta i rozwijająca się w sposób ciągły baza wiedzy *Przewodnika migracji IPv6*, która będzie obejmowała z czasem coraz większą liczbę usług i konfiguracji stosowanych w sieciach. Aplikacja *Przewodnik migracji IPv6* do swojego działania potrzebuje bazy wiedzy, na którą składa się hierarchiczna, drzewiasta struktura pytań, porad i podpowiedzi wyjaśniających używane pojęcia oraz dostarczająca dodatkowych informacji na tematy poruszane w poradach. Aby przewodnik mógł się rozwijać oraz aby jego baza wiedzy była aktualna i rozszerzalna, do aplikacji wprowadzono mechanizmy internetowej aktualizacji bazy wiedzy. Dla utrzymania spójności szybko rozwijającej się bazy wiedzy i w celu umożliwienia jej rozbudowy o kolejne porady i dodatkowe wersje językowe przewodnik został rozszerzony o internetową aplikację edycji pytań, porad i podpowiedzi (rys.2).

Założeniami do projektowania edytora pytań i porad była przenośność implementacji, obsługa wielu języków narodowych, edycja porad i podpowiedzi połączona z wygodną nawigacją po drzewie pytań, możliwość równoczesnej pracy wielu zespołów redakcyjnych oraz bezpośrednia współpraca z *Przewodnikiem migracji IPv6* [14].

Uzyskana funkcjonalność obejmuje uwierzytelnianie użytkowników z precyzyjnym rozdzieleniem uprawnień do tworzenia edycji, tłumaczeń i zatwierdzania treści bazy wiedzy *Przewodnika migracji IPv6*. Wprowadzenie systemu uprawnień do poszczególnych czynności edycyjnych pozwala na równoczesną pracę wielu osób nad edycją bazy wiedzy w wielu językach narodowych z możliwością rozgraniczenia autora, tłumacza i osoby zatwierdzającej daną treść. Wygodna nawigacja po złożonej strukturze drzewa pytań połączona z edycją porad

w trybie WYSIWYG (ang. What You See Is What You Get) daje możliwość umieszczania w treści porad i podpowiedzi grafik i internetowych odnośników.

Moduł tłumaczeń pozwala na wykonywanie lokalizacji pytań i porad bez konieczności tworzenia całej struktury bazy wiedzy w danym języku narodowym od podstaw. Zaimplementowana została również funkcjonalność eksportu bazy wiedzy w postaci wymaganego przez *Przewodnik migracji IPv6* pliku XML. Możliwość pobrania bazy porad i podpowiedzi wraz ze wszystkimi rekwizytami, takimi jak m.in. osadzone obrazy, jako pojedynczego pliku archiwum typu .tar.gz pozwala na internetową aktualizację aplikacji *Przewodnika migracji IPv6* o dodatkową wersję językową albo o zaktualizowaną bazę wiedzy.

Portal internetowy zawierający aplikację edytora pytań i porad stanie się miejscem dostarczania informacji o *Przewodniku migracji IPv6* oraz docelowym źródłem aktualizacji jego bazy wiedzy.

6 Plany rozwoju

Przewodnik migracji IPv6 jako aplikacja projektowana ze szczególnym uwzględnieniem łatwości obsługi oferuje automatyczne ograniczanie liczby pytań i porad zależne od udzielnych przez użytkownika odpowiedzi. Pozwala to na ograniczenie liczby zadawanych pytań i precyzyjne dostarczanie użytkownikom jedynie informacji związanych ze stosowanym przez nich sprzętem i oprogramowaniem.

Aplikacja tworzona w wersji przenośnej i niezwiązanej z żadną platformą systemową będzie uruchamiana i testowana pod trzema rodzinami systemów operacyjnych, którymi będzie Windows, Linux oraz FreeBSD. W planach rozwojowych *Przewodnika migracji IPv6* znajduje się wydanie go w formie płyty LiveCD pracującej pod systemem FreeBSD z przejrzystym interfejsem graficznym w połączeniu z systemem testów walidacji poprawności konfiguracji sieci IPv6. Otwarty internetowy edytor pytań i porad ma na celu ciągły rozwój bazy wiedzy *Przewodnika migracji IPv6*. Moduł tłumaczeń pozwoli na wydanie edycji *Przewodnika migracji IPv6* w innych językach, co przyczynić ma się do dotarcia do możliwie najszerszej grupy zainteresowanych użytkowników.

Obszerna baza wiedzy *Przewodnika migracji IPv6* z precyzyjnym wyszukiwaniem słów kluczowych i odpowiedziami na najczęściej zadawane pytania może stanowić dobre źródło informacji o konfiguracji usług sieciowych IPv6 wraz ze szczegółowym opisem wybranych problemów migracji.

7 Podsumowanie

W pracy zaprezentowana została metodologia wspomagania procesu migracji sieci do IPv6 poprzez analizę budowy i konfiguracji sieci, a następnie sposób dostarczenia porad i podpowiedzi ułatwiających rekonfigurację systemów i usług sieciowych. Zaproponowano narzędzie - *Przewodnik migracji IPv6*, który może być pomocny nie tylko dla administratorów małych sieci, ale również stanowić cenne źródło informacji na temat sieci IPv6 i konfiguracji usług sieciowych.

Otwarty rozwój bazy wiedzy, tłumaczenie porad i podpowiedzi ma szansę przyczynić się do zwiększenia powszechnej wiedzy na temat sieci IPv6 i dostarczenia jej do możliwie szerokiego grona odbiorców a także znacząco wpłynąć na rozpowszechnienie się IPv6 w sieciach typu SOHO.

Literatura

1. Access Google services over IPv6. In: The Official Website of Google over IPv6, http://www.google.com/intl/en/ipv6/l (retrieved on January 2011)
2. Asia Pacific Network Information Centre, The official website of APNIC, http://www.apnic.net/
3. British Broadcasting Corporation, Vint Cerf interview , "Internet pioneer Vint Cerf warns over address changes" http://www.bbc.co.uk/go/rss/int/news/-/news/technology-11736394, Nov 2010
4. S. Deering, R. Hinden, Internet Protocol, Version 6 (IPv6) Specification, RFC 2460, IETF, 1998
5. W. Gumiński, J. Światowiak, Obsługa protokołu IPv6 w systemach operacyjnych Windows i Linux, Przegląd Telekomunikacyjny 8-9/2010, pp. 797-798
6. R. Hinden, S. Deering, IP Version 6 Addressing Architecture, RFC 4291, IETF, 2006
7. IANA, The Official Website of IANA's IPv4 Address Report, http://www.ipv4.potaroo.net, 2011
8. InformationWeek, IPv6 Makes Slow Progress, http://www.informationweek.com/news/infrastructure/ipv6/212501014, Maj 2011
9. Internet Society, World IPv6 Day, The official website http://www.worldipv6day.org/
10. Inżynieria Internetu Przyszłości, The Official Website of Future Internet Engineering, http://iip.net.pl, Maj 2011
11. Inżynieria Internetu Przyszłości, Raport badawczy, zadanie Z1.2. Migration Guide Data Format, The Official Website of Future Internet Engineering, http://iip.net.pl, Maj 2011
12. N. R. Murphy, D. Malone, IPv6 Network Administration, O'Reilly Media 2005
13. K. Nowicki, M. Stankiewicz, A. Mrugalska, T. Mrugalski, J. Woźniak, Knowledge Managemenent in the IPv6 Migration Process, Computer Networks 2011, Czerwiec 2011
14. K. Nowicki, M. Stankiewicz, A. Mrugalska, J. Woźniak, T. Mrugalski, Extension Management of a knowledge base migration process to IPv6, SAINT2011 The 11th IEEE/IPSJ International Symposium on Applications and the Internet, 2011
15. K. Nowicki, J. Światowiak, B. Gajda, Problemy wdrażania protokołu IPv6, Przegląd Telekomunikacyjny, 8/9/2010
16. K. Sienkiewicz; M. Gajewski, J. Mongay Batalla, Testy IPv6 w sieci małego operatora, Przegląd Telekomunikacyjny, 8/9/2010
17. Światowy Dzień IPv6 na Politechnice Gdańskiej, the official website http://kti.net.pl/
18. W. Gumiński, A. Mrugalska, K. Nowicki, M. Stankiewicz, J. Woźniak: Przewodnik migracji do IPv6, Przegląd Telekomunikacyjny, 4/2011
19. http://www.tahi.org/

Testowanie usług IPv6 w sieci testowej opartej na urządzeniach z możliwością wirtualizacji sprzętowej

Henryk Gut[1], Jordi Mongay Batalla[1], Waldemar Latoszek[1],
Bartosz Gajda[2], Wiktor Procyk[2],
Krzysztof Chudzik[3], Jan Kwiatkowski[3], Kamil Nowak[3]

[1] Instytut Łączności – Państwowy Instytut Badawczy
[2] Poznańskie Centrum Superkomputerowo-Sieciowe
[3] Instytut Informatyki, Politechnika Wrocławska

Streszczenie W artykule omówiono zasadę tworzenia wirtualnych sieci IPv6 z użyciem wirtualnych: maszyn, pomostów, kart i interfejsów sieciowych, wirtualnych ruterów (tworzonych w środowisku ruterów fizycznych) oraz połączeń logicznych opartych na technice VLAN. Przedstawiono scenariusze testowe do badania zarówno protokołu trasowania, jak i prawidłowego funkcjonowania usług sieciowych DHCPv6 i DNS w środowisku tak tworzonych wirtualnych sieci IPv6. Omówiono również wyniki badań przeprowadzonych wg tych scenariuszy. Artykuł zawiera także opis wirtualnej sieci testowej IPv6, utworzonej wg tak zweryfikowanych zasad wirtualizacji.

1 Wprowadzenie

Jednym z głównych priorytetów Unii Europejskiej w obszarze technologii informatycznych i telekomunikacyjnych jest wdrażanie protokołu IPv6. Pomimo tego, iż protokół ten jest w dużej mierze zestandaryzowany, to nadal identyfikowane są pewne obszary niszowe wymagające dalszych badań. Jednym z takich obszarów jest szeroko rozumiana wirtualizacja zasobów środowiska IPv6, obejmująca wirtualizację urządzeń sieciowych, systemów obliczeniowych i łączących je sieci. Tak rozumiana wirtualizacja umożliwia tworzenie elastycznego środowiska sieci IPv6 i jako taka jest rozwijana w ramach jednego z zadań badawczych Projektu Krajowego „Inżyniera Internetu Przyszłości", a wyniki prowadzonych tam badań są przedstawione w raporcie [7]. Raport ten specyfikuje prototyp testowej sieci wirtualnej IPv6 i szczegółowo omawia zagadnienie testowania protokołu IPv6 w środowisku tej sieci, wraz ze specyfikacją konfiguracji sieci testowej dla różnych środowisk wirtualizacyjnych, opisem scenariuszy testowych oraz wynikami przeprowadzonych badań. Artykuł ten dotyczy jedynie wybranych aspektów zaprezentowanej w raporcie problematyki, i jako taki przedstawia zasadę budowania wirtualnych sieci IPv6, opartych na wirtualizacji serwerów, łączy i ruterów fizycznych, oraz – metodykę testowania podstawowych funkcjonalności tak budowanych sieci. W szczególności, w punkcie drugim artykułu omawia się: zasadę

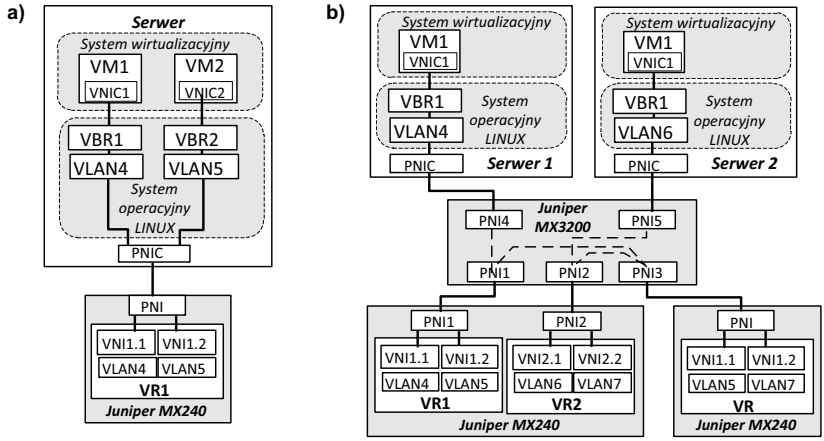
tworzenia wirtualnych sieci IPv6 z użyciem wirtualnych: maszyn, pomostów, kart i interfejsów sieciowych, a także – wirtualnych ruterów (tworzonych w środowisku ruterów fizycznych) oraz połączeń logicznych opartych na technice VLAN. W punkcie tym przedstawia się także scenariusze testowe do badania protokołu trasowania w środowisku sieci wirtualnych IPv6 oraz wyniki z przeprowadzonych badań wg tych scenariuszy. W punkcie trzecim przedstawia się testy weryfikujące możliwości implementacji różnych technologii wspierających wirtualizację sieci w ramach środowisk wirtualizacyjnych systemów operacyjnych i na pograniczu tych środowisk ze środowiskami sieci rzeczywistych, realizownych sprzętowo. W punkcie czwartym omawia się metodykę testowania usług sieciowych DHCPv6 i DNS, implementowanych w środowisku wirtualnych sieci IPv6, wraz z wynikami testów wykonanych wg tej metodyki. Punkt piąty zawiera krótki opis testowej sieci wirtualnej IPv6, utworzonej według zweryfikowanych zasad wirtualizacji. Punkt szósty to podsumowanie i wnioski z przeprowadzonych rozważań.

2 Testy weryfikacji rutingu

Opisane dalej scenariusze testowe zaprojektowano w celu sprawdzenia, czy przekazywanie pakietów w środowisku sieci IPv6 opartej na wirtualizacji zasobów sieciowych i informatycznych jest realizowane zgodnie z wcześniej zdefiniowanymi ścieżkami, czy też pakiety te są przekazywane od źródła do adresu przeznaczenia przy pomocy niekontrolowanych (przez użytkownika) procedur obejściowych, które są uaktywniane przez wirtualizacyjne systemy operacyjne, rezydujące w fizycznych serwerach i/lub ruterach użytych do wirtualizacji. Aby wykluczyć hipotezę takiego „pomostowego" przekazywania pakietów w środowisku wirtualnej sieci IPv6 proponuje się następujące dwa scenariusze testowe.

Scenariusz testowy 1 odnosi się do badania procesu przekazywania pakietów pomiędzy maszynami wirtualnymi, utworzonymi w środowisku jednego serwera fizycznego i należącymi do różnych podsieci jednego wirtualnego rutera. Badania te mogą być przeprowadzane z użyciem zestawu pomiarowego pokazanego na rys. 1a). Zestaw ten jest utworzony z jednego serwera firmy HP i jednego routera MX240 firmy Juniper, które to urządzenia są połączone bezpośrednio łączem fizycznym G-Ethernet. Serwer z tego zestawu może być wirtualizowany z użyciem jednego z powszechnie dostępnych systemów wirtualizacyjnych, takich jak np.: VMware [9], Hyper-V [10], XEN [11] lub KVM [12]. W laboratorium testowym serwer pracuje pod kontrolą systemu operacyjnego LINUX, pełniącego funkcję tzw. systemu operacyjnego gospodarza (*Host Operating System*) z systemem wirtualizacyjnym XEN wersji Citrix XenServer.

W zestawie pomiarowym z rys. 1a), z użyciem zastosowanego systemu wirtualizacyjnego, są utworzone dwie maszyny wirtualne (VM1 i VM2) z system operacyjnym tzw. gościa (*Guest Operating System*). Każda z tych maszyn ma wbudowaną własną wirtualną kartę sieciową (VNIC), z unikalnym adresem MAC, i ma przydzielony unikalny adres IPv6. Przy pomocy tzw. instancji wirtualnych pomostów (VBR1 i VBR2) oraz wirtualnych sieci LAN (VLAN4 i VLAN5),



Rysunek 1. Zestawy pomiarowe do przeprowadzania testów weryfikacji rutingu. Oznaczenia wyjaśniono w tekście referatu.

utworzonych w środowisku systemu operacyjnego gospodarza, maszyny te są logicznie dołączone do różnych wirtualnych interfejsów sieciowych (VNI1.1 i VNI1.2) wirtualnego rutera VR1 i tym samym należą do różnych podsieci tego rutera. Ruter wirtualny VR1 jest utworzony w środowisku rutera fizycznego MX240 firmy Juniper w oparciu o technologię tzw. systemów logicznych, udostępnianą w ruterach fizycznych tej klasy. W ruterze VR1 jest stosowany jedynie protokół trasowania OSPF (*Open Shortest Path First*) [1], zaś wirtualne interfejsy sieciowe tego rutera mają przydzielone (w sposób ręczny) różne adresy IPv6 [2, 4].

Po właściwym skonfigurowaniu zestawu pomiarowego i poprawnej aktywacji protokołu trasowania OSPF w ruterze VR1, w zestawie tym występuje tylko jedna ścieżka (dla każdego kierunku transmisji) pomiędzy maszynami VM1 i VM2. Dla kierunku przekazu od VM1 do VM2, ścieżka ta prowadzi poprzez: wirtualną kartę sieciową VNIC1 maszyny VM1, interfejsy sieciowe VNI1.1 i VNI1.2 rutera VR1 oraz – kartę sieciową VNIC2 maszyny VM2. W przeciwnym kierunku transmisji, pakiety IPv6 są przekazywane poprzez: kartę sieciowa VNIC2, interfejsy sieciowe VNIC2 i VNI1.1 rutera VR1 oraz kartę sieciowa VNC1 maszyny MV1. Aby sprawdzić, czy pakiety IPv6 są przekazywane tą właściwą trasą, a nie z użyciem pomostu programowego (tzn. bezpośrednio pomiędzy wirtualnymi kartami sieciowymi VNIC1 i VNIC2 tych maszyn), aktywowanego przez system wirtualizacyjny, można zastosować aplikację testową TraceRoute6 zarówno do wysyłania wiadomości ICMPv6 (*Internet Control Message Protocol for the Internet Protocol Version 6*) z maszyny VM1 do maszyny VM2 (i w kierunku przeciwnym),

jak i do rejestracji odpowiedzi urządzeń znajdujących się na wybranej ścieżce. Oczekiwanym pozytywnym wynikiem testu jest:

- raport z informacją o przechodzeniu pakietów przez ruter wirtualny VR1 wraz z parametrami tzw. „*Roundtrip Time*" dla każdego przeskoku;
- brak utraty pakietów i nie występowanie wiadomości „*Destination Unreachable*".

Postępując zgodnie z powyższym scenariuszem testowym, przeprowadzono badania z użyciem zestawu pomiarowego z rys. 1a). W zestawie tym zainstalowano system operacyjny LINUX z systemem wirtualizacyjnym XEN wersji Citrix XenServer 2.6.32.12-0.7.1.xs5.6.100.307.170586xen, protokół OSPF oraz przydzielono adresy IPv6 dla maszyn wirtualnych (VM1, VM2) oraz interfejsów sieciowych rutera VR1 zamieszczone na rys. 2a). W trakcie badań stwierdzono, że pakiety przechodzą przez ruter VR1 (rys. 2b), nie zarejestrowano utraty pakietów oraz nie odnotowano odbioru wiadomości „*Destination Unreachable*". Wynik ten jest zgodny z oczekiwanym pozytywnym rezultatem badań.

Rysunek 2. Przykładowy wynik testu wykonanego zgodnie z 1-scenariuszem testowym: a) – adresy maszyn wirtualnych i interfejsów sieciowych rutera VR1, b) – raport z aplikacji *TraceRoute6*.

Scenariusz testowy 2 odnosi się do badania procesu przekazywania pakietów pomiędzy maszynami wirtualnymi, utworzonymi w środowisku dwóch różnych serwerów fizycznych i należącymi do różnych podsieci dwóch różnych ruterów wirtualnych. Rutery te są utworzone w jednym ruterze fizycznym, ale połączenia pomiędzy nimi są wykonane poprzez trzeci ruter wirtualny, który rezyduje w środowisku drugiego rutera fizycznego. Badania te mogą być przeprowadzane z użyciem zestawu pomiarowego pokazanego na rys. 1b). Zestaw ten jest utworzony z dwóch serwerów firmy HP (Serwer 1 i Serwer 2), dwóch ruterów MX240 oraz jednego przełącznika MX3200 firmy Juniper. Urządzenia te są połączone przy pomocy pięciu łączy fizycznych G-Ethernet w sposób pokazany na rys. 1b). Obydwa serwery mogą być wirtualizowane przez zastosowanie jednego z dostępnych systemów wirtualizacyjnych (VMware, Hyper-V, XEN lub KVM).

W zestawie pomiarowym z rys. 1b), w środowisku każdego z serwerów, z użyciem wybranego systemu wirtualizacyjnego, jest utworzona jedna maszyna wirtualna VM1. Każda z tych maszyn ma wbudowaną jedną wirtualną kartę sieciową VNCI1 z unikalnym adresem MAC oraz ma przydzielony unikalny adres IPv6. Wykorzystując instancje wirtualnego pomostu (VBR1) i wirtualnych sieci LAN (VLAN4 i VLAN6), utworzone w środowisku systemu operacyjnego gospodarza, maszyny te są logicznie dołączone do różnych wirtualnych interfejsów sieciowych z wirtualnych ruterów VR1 i VR2. Przy czym maszyna VM1 z Serwera 1 jest przyłączona do interfejsu VNI1.1 z rutera VR1, zaś maszyna VM1 z Serwera 2 – do interfejsu VNI2.1 z rutera VR2. Dzięki temu maszyny te funkcjonują w różnych podsieciach, które to podsieci są przypisane do różnych ruterów wirtualnych VR1 i VR2, rezydujących w środowisku tego samego rutera fizycznego.

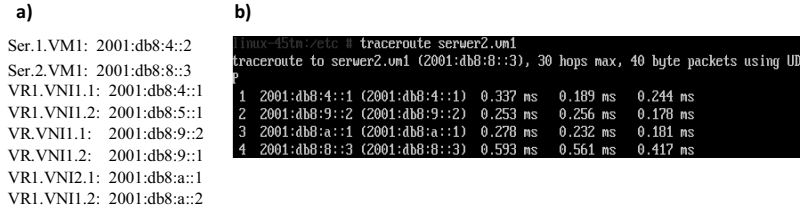
Zgodnie z rys. 1 b), drugi wirtualny interfejs sieciowy VNI1.2 z rutera VR1 jest logicznie połączony z pierwszym interfejsem VNI1.1 z rutera wirtualnego VR, rezydującego na drugim ruterze fizycznym, zaś drugi interfejs VNI1.2 z tego rutera jest połączony z drugim interfejsem wirtualnym VNI2.2 z rutera wirtualnego VR2, osadzonego w pierwszym ruterze fizycznym. Te logiczne połączenia są zrealizowane odpowiednio przez VLAN5 i VLAN7.

W omawianym zestawie pomiarowym, rutery wirtualne VR1, VR2 oraz VR są utworzone z użyciem wcześniej wzmiankowanych systemów logicznych i stosują protokół trasowania OSPF. Każdy wirtualny interfejs sieciowy z tych ruterów ma ręcznie przydzielony adres IPv6.

Po wykonaniu właściwej konfiguracji zestawu pomiarowego, tzn. w sposób zgodny z rys. 1 b), i poprawnej aktywacji protokołu trasowania OSPF w ruterach: VR1, VR2 oraz VR, w zestawie tym występuje tylko jedna ścieżka (dla każdego kierunku transmisji) pomiędzy wirtualnymi maszynami VM1 rezydującymi na Serwerze 1 i Serwerze 2. Dla kierunku przekazu od VM1 z Serwera 1 do VM1 z Serwera 2, ścieżka ta prowadzi poprzez: odpowiednie interfejsy sieciowe VNI z ruterów VR1, VR oraz VR2, zaś dla przeciwnego kierunku transmisji pakiety będą przekazywane poprzez właściwe interfejsy ruterów: VR2, VR oraz VR1. Aby sprawdzić, czy pakiety IPv6 są przekazywane tą właściwą trasą a nie z użyciem pomostu programowego (tzn. bezpośrednio pomiędzy wirtualnymi interfejsami VNI1.1 i VNI2.1 z ruterów wirtualnych odpowiednio VR1 i VR2), aktywowanego przez system wirtualizacyjny pierwszego rutera fizycznego Juniper MX240, można zastosować aplikację testową *TraceRoute6* zarówno do wysyłania wiadomości ICMPv6 pomiędzy maszynami wirtualnymi VM1 rezydującymi na Serwerach 1 i 2, jak i rejestracji odpowiedzi od ruterów znajdujących się na wybranej ścieżce. Pozytywnym wynikiem testu jest:

– raport z informacją o przechodzeniu pakietów przez rutery wirtualne: VR1, VR oraz VR2, w przypadku kiedy wiadomość ICMPv6 jest nadawana z maszyny VM1 rezydującej na Serwerze 1 do maszyny VM1 utworzonej w środowisku Servera 2, wraz z parametrami tzw. „*Roundtrip Time*" dla każdego przeskoku;

– raport z informacją o przechodzeniu pakietów przez rutery wirtualne: VR2, VR oraz VR1, w przypadku kiedy wiadomość ICMPv6 jest nadawana z maszyny VM1 utworzonej na Serwerze 2 do maszyny VM1 z Servera 1, wraz z parametrami "Roundtrip Time" dla każdego przeskoku;
– brak utraty pakietów i nie występowanie wiadomości „*Destination Unreachable*" dla obydwu kierunków transmisji wiadomości ICMPv6.



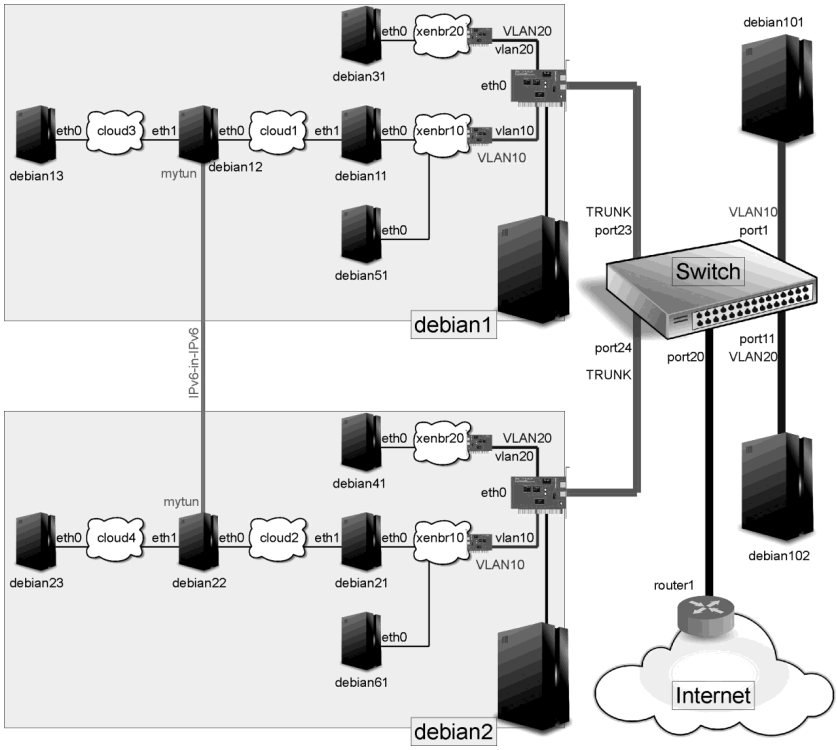
Rysunek 3. Przykładowy wynik testu wykonanego zgodnie z 2-scenariuszem testowym, dla przekazu pakietów z maszyny VM1 z Serwera 1 do maszyny VM2 z Serwera 2: a) – adresy maszyn oraz interfejsów sieciowych ruterów: VR1, VR i VR2, b) – raport z aplikacji *TraceRoute6*.

Postępując zgodnie z powyższym scenariuszem testowym, przeprowadzono badania z użyciem zestawu pomiarowego z rys. 2a). W zestawie tym zainstalowano system operacyjny LINUX, system wirtualizacyjnym XEN oraz protokół OSPF takiej samej wersji, jak w Scenariuszu testowym 1. Dla maszyn wirtualnych VM1 z serwerów 1 i 2 oraz interfejsów sieciowych ruterów VR1, VR2 i VR przydzielono adresy zamieszczone na rys. 3a). W trakcie badań stwierdzono, że pakiety przechodzą przez ruter VR1, VR i VR2 (rys. 3b), nie zarejestrowano utraty pakietów oraz nie odnotowano odbioru wiadomości „*Destination Unreachable*". Wynik ten jest zgodny z oczekiwanym pozytywnym rezultatem badań. Taki sam wynik uzyskano dla drugiego kierunku przekazu pakietów.

3 Testy połączeń sieciowych w środowisku wirtualizacyjnym XEN

Celem testów było sprawdzenie możliwości implementacji różnych technologii wspierających wirtualizację sieci w ramach środowisk wirtualizacyjnych systemów operacyjnych i na pograniczu tych środowisk ze środowiskami sieci rzeczywistych, realizowanych sprzętowo. Badano też możliwości wykorzystania tych technologii do tworzenia równoległych internetów i wirtualnych sieci aplikacyjnych powstających w ramach projektu „Inżyniera Internetu Przyszłości".

Jako środowisko wirtualizacyjne wybrano system operacyjny Debian z zainstalowanym środowiskiem XEN. Schemat przyjętego środowiska badawczego przedstawiono na rys. 4.



Rysunek 4. Zestaw testowy do badania połączeń sieciowych w środowisku wirtualizacyjnym XEN.

Przedmiotem badań były technologie wirtualizacyjne w drugiej i trzeciej warstwie modelu ISO/OSI, a mianowicie:

– sieci VLAN, jako wsparcie tworzenia równoległych internetów oraz
– sieci wirtualne (tunele) oparte na adresacji logicznej IPv6, jako wsparcie tworzenia sieci wirtualnych dla aplikacji w ramach równoległych internetów.

W skład zestawu testowego z rys. 4 wchodzi następujący sprzęt fizyczny: Switch Cisco Catalyst; dwa komputery z zainstalowanym systemem operacyjnym Debian 6 (squeeze) i środowiskiem wirtualizacyjnym XEN 4.0 – komputery debian1 i debian2; dwa komputery wykorzystywane do testów sieci VLAN –

debian101 i debian102 oraz przyłącze do Internetu za pośrednictwem routera programowego router1.

Centralnym urządzeniem sieciowym jest przełącznik (switch) Cisco Catalyst. Konfiguracja definiuje dwie sieci VLAN o numerach 10 i 20 oraz różne typy połączeń zewnętrznych. Komputery debian1 i debian2 zostały przyłączone do przełącznika za pomocą łączy typu trunk (porty 23 i 24), które mogą przesyłać jednym kablem sieciowym dane z wielu sieci VLAN. Komputery debian101 i debian102 zostały przyłączone odpowiednio do dedykowanych portów skonfigurowanych do transmisji danych w sieciach VLAN: port 1 – VLAN10 oraz port 11 – VLAN20. Switch został przyłączony portem 20 do routera router1, który jest podłączony do Internetu z wykorzystaniem translacji adresów. Port 20 był skonfigurowany do pracy w domyślnej sieci VLAN 1, skonfigurowanej jako sieć typu native. Ramki pochodzące z tej sieci nie były powiększane o znaczniki sieci VLAN używane w standardzie 802.1q. Sieć VLAN 1 przesyłała dane z głównych interfejsów komputerów debian1 i debian2, które wysyłały ramki nieznakowane.

Na komputerach ze środowiskiem wirtualizacyjnym zainstalowano i skonfigurowano: sieci wirtualne w obrębie środowiska wirtualizacyjnego, maszyny wirtualne z systemem operacyjnym Debian 6, połączenia maszyn wirtualnych z sieciami wirtualnymi oraz przyłącza typu trunk pomiędzy środowiskami wirtualizacyjnymi a przełącznikiem.

Testy VLAN obejmowały następujące aspekty: zdolność interpretacji protokołu 802.1q i współpraca z użyciem łączy typu trunk, wykorzystujących ten protokół, z innymi środowiskami wirtualizacyjnymi oraz sprzętowym przełącznikiem; poprawność izolacji pomiędzy VLAN; dystrybucja VLAN wewnątrz środowiska wirtualizacyjnego.

W ramach testu połączono komputery fizyczne z przełącznikiem za pomocą łącza typu trunk. Stwierdzono, że istnieje możliwość konfiguracji interfejsów sieciowych systemu Debian 6 tak, aby można było odbierać i poprawnie interpretować pakiety opatrzone znacznikami VLAN. Wykonuje się to definiując odpowiednio interfejsy wirtualne o nazwie vlanX, gdzie X to numer VLAN, i do nich transmitowane są odpowiednio pakiety sieci VLAN zgodne z ich nazwą. Pakiety z interfejsów VLAN można dystrybuować do maszyn wirtualnych za pomocą mostów połączonych z tymi interfejsami (xenbr10 i xenbr20), przyłączając do nich maszyny wirtualne. Sprawdzono poprawność funkcjonowania połączeń i izolację pomiędzy maszynami wirtualnymi i z użyciem dodatkowych maszyn fizycznych włączonych w odpowiednie sieci VLAN.

Podsumowując tą część testów, można stwierdzić, że w systemie Debian 6 z XEN 4.0 da się prawidłowo skonfigurować połączenia typu trunk dla interfejsów sieciowych zewnętrznych i dystrybuować transmisję wewnątrz środowisk wirtualizacyjnych z zachowaniem izolacji pomiędzy sieciami VLAN.

Test wirtualnych sieci logicznych obejmował następujące aspekty: możliwość wytworzenia tuneli dla sieci wirtualnych funkcjonujących z wykorzystaniem IPv6 oraz możliwość konfigurowania i poprawność funkcjonowania routingu wykorzy-

stującego interfejsy tuneli IPv6. Do testu wykorzystany został mechanizm tunelowania IPv6-in-IPv6.

Test przebiegał następująco: Wybrano i skonfigurowano sieci końcowe do tunelowania (cloud 3 i 4). Wybrano i skonfigurowano routery tunelu (debian 12 i 22) na skraju sieci końcowych, pomiędzy którymi miał powstać tunel. Skonfigurowano i sprawdzono poprawność funkcjonowania routingu na routerach pośredniczących pomiędzy routerami tunelu. Sieć skonfigurowano tak, aby routery pośredniczące nie posiadały informacji o sieciach końcowych. Skonfigurowano tunel (interfejsy mytun na debian 12 i 22) i routing z wykorzystaniem interfejsów tego tunelu. Stwierdzono, że mechanizm tunelowania funkcjonuje poprawnie.

Zatem, w systemie Debian 6 z XEN 4.0 da się prawidłowo konfigurować wirtualne tunele IPv6-in-IPv6 i wykorzystać interfejsy tych tuneli do kierowania pakietów.

Podsumowując wyniki badań w środowisku testowym z rys. 4 można stwierdzić, że zaprezentowane techniki w zakresie konfiguracji i uruchamiania maszyn wirtualnych oraz konfiguracji i uruchamiania sieci wirtualnych dają podstawy do konfigurowania i uruchamiania złożonych rozwiązań w ramach środowisk wirtualizacyjnych, obsługi serwerów aplikacji użytkowych, maszyn wirtualnych usług sieciowych (routery i serwery), sieci wirtualnych logicznych dla potrzeb aplikacji oraz sieci wirtualnych tworzących równoległe internety.

4 Testy prawidłowego działania usług DHCPv6 i DNS w środowisku wirtualnej sieci IPv6

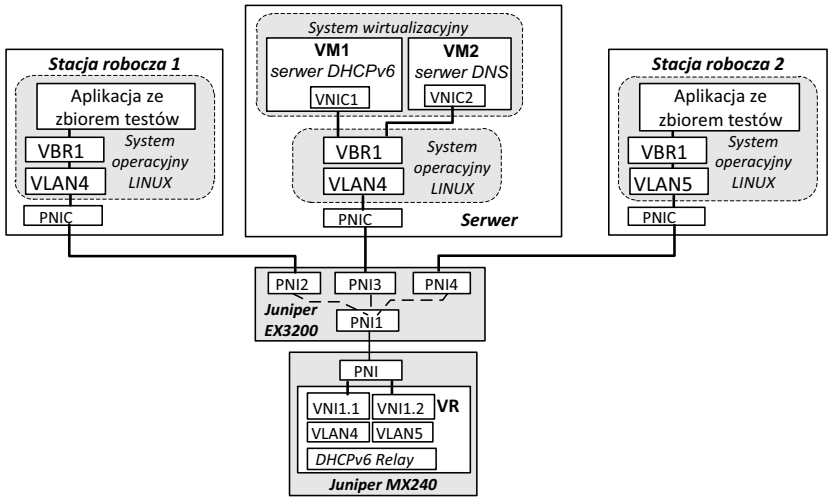
Usługa dynamicznej konfiguracji węzłów sieci IPv6, zwana dalej usługą/protokołem DHCPv6 (*Dynamic Host Configuration Protocol for IPv6 network*) [3, 12], oraz system DNS (*Domain Name System*) [5, 6] mogą być postrzegane jako usługi sieciowe ułatwiające użytkownikom korzystanie z sieci IPv6. Usługa DHCPv6 jest pewnego rodzaju protokołem komunikacyjnym, który dotyczy zarówno formatu pakietów IPv6, jak i systemu adresacji, który jest wyspecyfikowany przez RFC 3315 [3]. Umożliwia ona: przydzielanie adresu IPv6 dla stacji roboczej (host'a), uzyskiwanie adresu serwera DNS, a także – maski podsieci do której należy stacja robocza. Usługa DHCPv6 ma trzy różne implementacje: klient DHCPv6 instalowany w środowisku stacji roboczej, DHCPv6 Relay oraz Serwer DHCPv6. Funkcje serwera DHCPv6 jako głównego węzła usługowego są zwykle uruchamiane w środowisku dedykowanych serwerów. Węzeł DHCPv6 Relay ma funkcjonalność węzła pośredniczącego. Z jednej strony węzeł ten przekazuje wiadomości z żądaniami DHCPv6 odebranymi od aplikacji klienta DHCPv6 i kierowanymi do określonego serwera DHCPv6, zaś ze strony drugiej – wiadomości z odpowiedziami serwera DHCPv6 na żądania przekazuje do odpowiedniego klienta DHCPv6, generującego te żądania. Węzeł DHCPv6 Relay jest zwykle implementowany jako dodatkowa funkcjonalność ruterów fizycznych i logicznych, co ma miejsce między innymi w przypadku ruterów fizycznych serii MX240 firmy Juniper oraz wszystkich 15-ruterów logicznych, które mogą być utworzone w ruterach tej serii.

Dalej opisane scenariusze testowe zaprojektowano w celu sprawdzenia, czy implementacje serwera DHCPv6 i węzła pośredniczącego DHCPv6 Relay, a także implementacja serwera DNS, zainstalowane w środowisku wirtualizowanych sieci IPv6, spełniają wymagania wyspecyfikowane w odpowiednich zaleceniach RFC [3, 5, 6]. W celu weryfikacji tej hipotezy proponuje się przeprowadzenie, następujących dwóch scenariuszy testowych dla usług DHCPv6 (Scenariusz testowy 1 oraz Scenariusz testowy 2) oraz scenariusz dla usługi DNS (Scenariusz testowy 3). Badania według tych scenariuszy mogą być wykonywane z użyciem zestawu pomiarowego pokazanego na rys 5. Zestaw ten jest utworzony z jednego serwera firmy HP i jednego routera MX240 firmy Juniper, dwóch stacji roboczych (Stacja robocza 1 i Stacja robocza 2) oraz jednego przełącznika Juniper MX3200. Przełącznik ten ma uaktywnione jedynie cztery fizyczne interfejsy sieciowe PNI, do których to interfejsów (za pomocą łączy fizycznych G-Ethernet) są przyłączone bezpośrednio interfejsy fizyczne: rutera, serwera i stacji roboczych.

Serwer z tego zestawu może być wirtualizowany z użyciem jednego z wcześniej wzmiankowanych systemów wirtualizacyjnych: VMware, Hyper-V, XEN lub KVM. Z użyciem zastosowanego systemu wirtualizacyjnego, są utworzone dwie maszyny wirtualne (VM1 i VM2) z system operacyjnym tzw. gościa. Każda z tych maszyn zawiera własną wirtualną kartę sieciową (VNIC), z unikalnym adresem MAC, i ma przydzielony unikalny adres IPv6. Przy pomocy wirtualnego pomostu (VBR1) oraz wirtualnej sieci LAN (VLAN4), utworzonych w środowisku systemu operacyjnego gospodarza, maszyny te są logicznie dołączone do jednego wirtualnego interfejsu sieciowego (VNI1.1) z wirtualnego rutera VR i tym samym należą tej samej podsieci tego rutera. Na użytek dalej omówionych scenariuszy testowych, na maszynach VM1 i VM2 są zainstalowane i właściwie skonfigurowane wybrane implementacje serwerów DHCPv6 i DNS. **Zakłada się, że prawidłowe funkcjonowanie wybranych implementacji serwerów DHCPv6 i DNS zostało sprawdzone w środowisku fizycznej (nie wirtualnej) sieci IPv6, na przykład przy pomocy zestawu testów dostępnych na TAHI Live CD** [8].

Ruter wirtualny VR jest utworzony w środowisku rutera fizycznego MX240 firmy Juniper w oparciu o technologię tzw. systemów logicznych, udostępnianą w ruterach fizycznych tej klasy. W ruterze tym jest stosowany protokół trasowania OSPF, a wirtualne interfejsy sieciowe tego rutera (VNI1 i VNI2) mają ręcznie przydzielone różne adresy IPv6. **Zakłada się, że ruter VR ma wbudowaną funkcjonalność węzła pośredniczącego DHCPv6 Relay**, która jest automatycznie (lub w sposób ręczny) uaktywniana w czasie procedury restartu tego rutera.

Stacje robocze 1 i 2 są fizycznymi (nie wirtualnymi) komputerami klasy PC, które pracują pod kontrolą systemu operacyjnego LINUX. Każda z tych stacji ma fizyczną kartę sieciowa PNIC z unikalnym adresem MAC. Przy pomocy instancji wirtualnego pomostu i wirtualnej sieci LAN, utworzonych w środowisku systemu LINUX, stacja robocza 1 jest logicznie dołączona (poprzez VLAN4) do interfejsu VNI1.1 z rutera wirtualnego VR, zaś stacja robocza 2 – do interfejsu VNI1.2 (poprze VLAN5) tego rutera. Dzięki tym połączeniom stacja robocza 1 razem



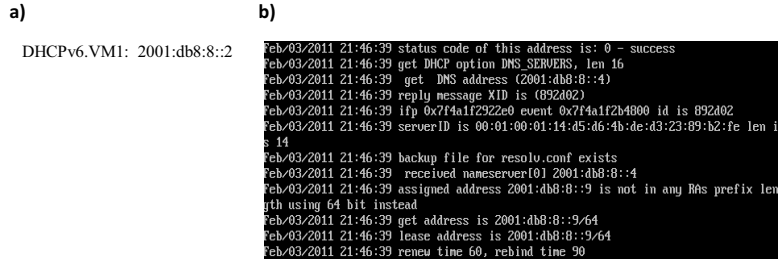
Rysunek 5. Zestaw pomiarowy do testowania prawidłowego funkcjonowania usług *DHCPv6* i *DNS* w środowisku wirtualnej sieci IPv6. Oznaczenia wyjaśniono w tekście referatu.

z serwerem DHCPv6 i serwerem DNS tworzą podsieć 1 z rutera wirtualnego VR, natomiast stacja robocza 2 należy do podsieci 2 tego rutera.

Scenariusz testowy 1 dotyczy badania prawidłowego funkcjonowania serwera DHCPv6 w środowisku wirtualnej sieci IPv6, w sytuacji kiedy serwer DHCPv6 oraz stacja robocza należą do tej samej podsieci. Test ten jest realizowany z wykorzystaniem wyżej opisanego zestawu pomiarowego, w którym aktywnymi węzłami są: Stacja robocza 1, maszyna wirtualna VM1 z serwerem DHCPv6 oraz maszyna wirtualna VM2 z serwerem DNS. W scenariuszu zakłada się, że: wszystkie połączenia logiczne są wykonane w sposób jak opisano powyżej. Weryfikacja prawidłowego działania serwera DHCPv6 jest dokonywana przez wysłanie ze Stacji roboczej 1 żądania automatycznego przydziału adresu IPv6. Pozytywnym wynikiem testu jest odpowiedź od serwera DHCPv6, która powinna zawierać: właściwy adres IPv6 przydzielony stacji roboczej 1, adres IPv6 serwera DNS oraz maskę podsieci 1.

Postępując zgodnie z powyższym scenariuszem testowym, przeprowadzono badania z użyciem zestawu pomiarowego z rys. 5. W zestawie tym zainstalowano system operacyjny Citrix XenServer w wersji 2.6.32.12-0.7.1.xs5.6.100.307.170586xen, aplikacje serwera i klienta DHCPv6 wersji 1.0.22 oraz aplikacje serwera DNS wersji Bind 9.8.0. Dla serwerów DHCPv6 i DNS przydzielono adresy zamieszczone na rys. 6a). W trakcie badań stwier-

dzono funkcjonowanie serwera DHCPv6 zgodne z oczekiwanym wynikiem testu. Potwierdzenie tego są logi klienta DHCPv6 zamieszczone na rys. 6b).



Rysunek 6. Przykładowy wynik testu serwera DHCPv6 wykonanego zgodnie z 1-scenariuszem testowym: a) – adresy serwerów DHCPv6 i DNS, b) – logi klienta DHCPv6 na stacji roboczej 1.

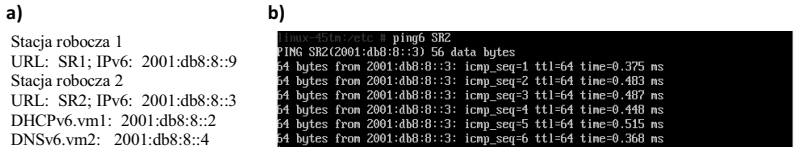
Scenariusz testowy 2 odnosi się do badania prawidłowego funkcjonowania węzła pośredniczącego DHCPv6 Relay w środowisku wirtualnej sieci IPv6, w sytuacji kiedy serwer DHCPv6 oraz stacja robocza należą do różnych podsieci. Badanie to jest wykonywane z użyciem zestawu pomiarowego z rys. 5, w którym aktywnymi węzłami są: Stacja robocza 2, ruter VR z aktywnym węzłem DHCPv6 Relay, maszyna wirtualna VM1 z serwerem DHCPv6 oraz maszyna wirtualna VM2 z serwerem DNS. W scenariuszu zakłada się, że: wszystkie połączenia logiczne są wykonane w sposób jak opisano powyżej, poszczególnym podsieciom rutera VR zostały przypisane właściwe maski, protokół OSPF oraz węzeł DHCPv6 Relay zostały poprawnie uaktywnione na ruterze wirtualnym VR. Weryfikacja prawidłowego działania węzła DHCPv6 Relay jest dokonywana przez wysłanie ze Stacji roboczej 2 żądania automatycznego przydziału adresu IPv6. Pozytywnym wynikiem testu jest odpowiedź od serwera DHCPv6, która powinna zawierać: właściwy adres IPv6 przydzielony stacji roboczej 2, adres IPv6 serwera DNS oraz maskę podsieci 2.

Postępując zgodnie z powyższym scenariuszem testowym, przeprowadzono badania z użyciem zestawu pomiarowego z rys. 5. W zestawie tym zainstalowano system operacyjny LINUX, system wirtualizacyjnym XEN, aplikacje serwera i klienta DHCPv6 oraz aplikacje serwera DNS, takie same jak w scenariuszu testowym 1. W trakcie badań stwierdzono funkcjonowanie serwera DHCPv6 zgodne z oczekiwanym wynikiem testu. Z uwagi na ograniczoną objętość artykułu nie zamieszczono tu logów klienta DHCPv6 uzyskanych w trakcie badań.

Scenariusz testowy 3 dotyczy weryfikacji poprawnego działania serwera DNS w środowisku wirtualnej sieci IPv6, w przypadku kiedy serwer DNS

i stacja robocza inicjująca sesję należą do różnych podsieci. W scenariuszu tym aktywnymi węzłami zestawu pomiarowego z rys 5 są: Stacje robocze 1 i 2, ruter VR, maszyna VM1 z serwerem DHCPv6 oraz maszyna VM2 z serwerem DNS. W scenariuszu zakłada się, że: wszystkie połączenia logiczne są wykonane w sposób jak opisano powyżej, poszczególnym podsieciom rutera VR zostały przypisane właściwe maski, a protokół OSPF został poprawnie uaktywnione na tym ruterze. Aby zweryfikować prawidłowe działania serwera DNS ze Stacji roboczej 2 jest wysyłana wiadomość ***pingv6*** z właściwym adresem URL stacji roboczej 1. Pozytywnym wynikiem testu jest zbiór odpowiedzi od wywoływanej stacji roboczej 1 z poprawnym adresem IPv6.

Postępując zgodnie z powyższym scenariuszem testowym, przeprowadzono badania z użyciem zestawu pomiarowego z rys 5. Zestaw ten skonfigurowano identycznie jak w scenariuszu testowym 2. W trakcie badań stwierdzono funkcjonowanie serwera DNS zgodne z oczekiwanym wynikiem testu. Potwierdzenie tego są odpowiedzi od stacji wywoływanej (w tym przypadku jest nią stacja robocza 1), zamieszczone na rys. 7b).

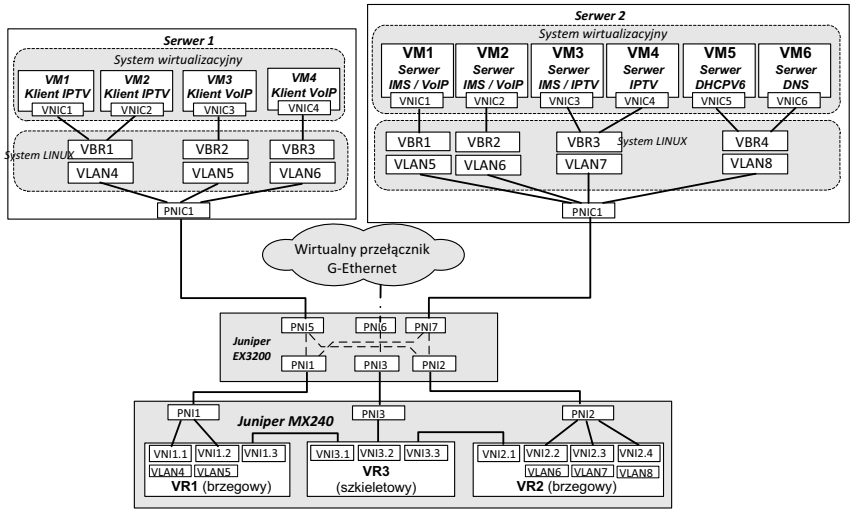


Rysunek 7. Przykładowy wynik testu serwera DNS wykonanego zgodnie z 3-scenariuszem testowym: a) – adresy stacji roboczych (URL i IPv6) oraz serwerów DHCPv6 i DNS, b) – odpowiedzi od wywoływanej stacji roboczej 1.

5 Wirtualna sieć laboratoryjna IPv6

Jedno-domenową wirtualną sieć laboratoryjna IPv6 utworzono z użyciem (rys. 8): dwóch serwerów firmy HP, jednego routera typu MX240 oraz jednego przełącznika ethernetowego typu MX3200 firmy Juniper. W sieci tej serwery 1 i 2 reprezentują zasoby wirtualne warstwy usługowej sieci, utworzone z wykorzystaniem systemu wirtualizacyjnego XEN. System ten funkcjonuje w środowisku systemu operacyjnego LINUX, pełniąc funkcje tzw. systemu operacyjnego gospodarza, i jako taki jest wykorzystywany do tworzenia maszyn wirtualnych. Każda z tych maszyn ma wirtualną kartę interfejsu sieciowego VNIC z unikalnym adresem MAC. Karty te, w obrębie pojedynczego serwera, są logicznie dołączone do odpowiednich pomostów wirtualnych VBR, a ramki ethernetowe wychodzące z ze-

wnętrznego portu każdego wirtualnego pomostu VBR są oznaczane odpowiednią etykietą „VLANk", gdzie k – jest numerem identyfikacyjnym sieci VLAN. Ramki te są przekazywane do fizycznej karty sieciowej serwera PNIC, poprzez którą są wprowadzane do łącza fizycznego wirtualnej sieci laboratoryjnej. W przeciwnym kierunku transmisji, funkcja dekodowania etykiety VLANk analizuje etykiety ramek odbieranych od karty PNIC. Jeśli etykiety tych ramek są równe etykiecie VLANk, wtedy etykiety te są usuwane z ramek odbieranych, a ramki bez tych etykiet są przekazywane do portu zewnętrznego odpowiedniego pomostu wirtualnego VBR. Zarówno mechanizmy wirtualnego pomostu VBR, jak i funkcje dopisywania (usuwania) etykiet sieci wirtualnych VLANk do (z) ramek nadawanych (odbieranych) przez kartę PNIC są realizowane przez system operacyjny gospodarza (LINUX) serwerów 1 i 2.



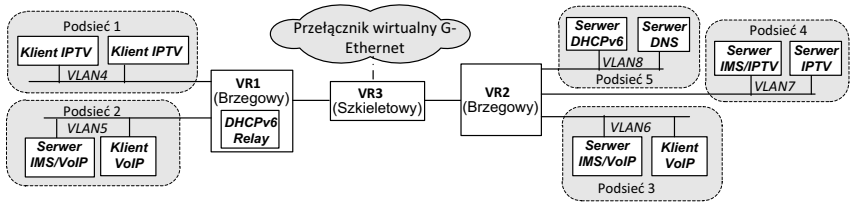
Rysunek 8. Topologia fizyczna wirtualnej sieci laboratoryjnej IPv6. Oznaczenia wyjaśniono w tekście artykułu.

W sieci laboratoryjnej IPv6, serwery usług sieciowych (IMS/VoIP, IMS/IPTV, IPTV, DHCPv6 oraz DNS), a także wybrane aplikacje klienckie (IPTV, VoIP) rezydują w środowisku poszczególnych maszyn wirtualnych VM. Dla przejrzystości rozwiązania przyjęto, że wszystkie aplikacje klienckie są zainstalowane na maszynach wirtualnych serwera 1, zaś aplikacje usług sieciowych – na maszynach wirtualnych z serwera 2.

W sieci laboratoryjnej IPv6, warstwa sieciowa jest utworzona w oparciu o technologię tzw. systemów logicznych, udostępnianą w ruterach serii MX240 firmy Juniper. Na bazie tej technologii, w pojedynczym ruterze MX240

może być utworzonych do 15-niezależnych ruterów wirtualnych, całkowicie od siebie odseparowanych. W omawianej sieci wykorzystano jedynie trzy tego typu rutery, oznaczone (na rys. 8) jako: VR1, VR2 oraz VR3. Rutery brzegowe VR1 i VR2 mają zaimplementowany protokół OSPF, natomiast ruter szkieletowy VR3 – protokół BGP (*Border Gateway Protocol*). Protokół BGP jest tu wykorzystywany do tworzenia połączeń sieci lokalnej z innymi sieciami lokalnymi. Każdy z wirtualnych ruterów jest wyposażony w zbiór wirtualnych interfejsów sieciowych VNI. Przy czym interfejsy, oznaczone (na rys. 8) jako: VNI1.3, VNI3.1, VNI3.3 oraz VNI2.1, są wykorzystywane dla tworzenia wewnętrznych połączeń logicznych ruterów brzegowych z ruterem szkieletowym, zaś interfejsy: VNI1.1, VNI1.2, VNI2.2, VNI2.3 oraz VNI2.4 realizują funkcje zakończeń odpowiednich podsieci VLAN. Interfejs wirtualny VNI3.2 rutera szkieletowego jest połączony logicznie z portem fizycznym PNI3 rutera MX240 i jako taki jest wykorzystywany do tworzenia połączeń zewnętrznych z innymi sieciami lokalnymi. Wszystkie połączenia logiczne pomiędzy interfejsami wirtualnymi VNI i fizycznymi PNI rutera MX240 są tworzone w oparciu o technologię tunelowania dostępną w tym ruterze.

W sieci laboratoryjnej IPv6, przełącznik EX3200 jest wykorzystywany do tworzenia połączeń pomiędzy interfejsami 1-GEthernet urządzeń fizycznych zastosowanych w tej sieci, a także – do odpowiedniego przełączania ruchu od poszczególnych podsieci VLAN pomiędzy tymi interfejsami. Te funkcjonalności przełącznika, w połączeniu z wirtualnymi pomostami (VBR) i funkcjami oznaczania podsieci VLAN, tworzą topologię logiczną sieci laboratoryjnej IPv6 przedstawioną na rys 9.



Rysunek 9. Topologia logiczna wirtualnej sieci laboratoryjnej IPv6. Oznaczenia wyjaśniono w tekście artykułu.

6 Podsumowanie

Problemy wdrażania protokołu IPv6 znajdują się w centrum uwagi wielu środowisk naukowych, operatorów komercyjnych oraz są jednym z ważnych priorytetów Unii Europejskiej. Z drugiej strony, wirtualizacja jest technologią praktycz-

nie wdrażaną i wykorzystywaną na coraz szerszą skalę. Niniejszy artykuł porusza mało zbadaną do tej pory problematykę wynikającą z użycia tych dwóch rozwiązań: protokołu IPv6 i wirtualizacji.

W artykule przedstawiono zagadnienia dotyczące wirtualizacji środowisk IPv6 zbudowanych w oparciu o infrastrukturę sieci wirtualnych oraz zwirtualizowane systemy operacyjne. W szczególności przedstawiono scenariusze budowy sieci wirtualnych IPv6 wykorzystujących wirtualne maszyny, pomosty, karty i interfejsy sieciowe oraz routery wirtualne bazujące na routerach sprzętowych Juniper MX240. Dla separacji ruchu pomiędzy sieciami wykorzystano technikę VLAN. Następnie zdefiniowano niezbędne scenariusze testowe dla weryfikacji poprawności protokołu trasowania i tworzenia połączeń sieciowych w środowisku wirtualizacyjnym XEN, a także - scenariusze testowe dla prawidłowego funkcjonowania usług sieciowych DHCPv6 i DNS.

Artykuł prezentuje również wyniki prac praktycznych. W tym zakresie zbudowano sieć testową wykorzystującą system wirtualizacyjny XEN w wersji Citrix XenServer oraz zrealizowano szereg testów w oparciu o zdefiniowane wcześniej scenariusze. Testy te pozwoliły potwierdzić prawidłowe działanie środowiska wirtualnego IPv6 w zakresie protokołów trasowania. Ponadto, potwierdzono poprawne działanie niezbędnych usług sieciowych takich jak: DHCPv6 oraz DNS. W ostatniej części artykułu przedstawiono topologię rozbudowanej wirtualnej sieci laboratoryjnej przeznaczonej do uruchomienia szeregu usług sieciowych i aplikacji: (IMS/VoIP, IMS/IPTV, IPTV, DHCPv6 oraz DNS), a także wybrane aplikacje klienckie (IPTV, VoIP).

Literatura

1. Coltun, R., Ferguson, D., Moy, J., Lindem, A.: RFC 5340 – OSPF for IPv6, July 2008
2. Crawford, M.: RFC 2464 – Transmission of IPv6 Packets over Ethernet Networks, December 1998
3. Droms, R., Bound, J., Volz, B., Lemon, T., Perkins, C., Carney, M.: RFC 3315 – Dynamic Host Configuration Protocol for IPv6 (DHCPv6), July 2003
4. Hinden, R., Deering, S.: RFC 2373 – IP Version 6 Addressing Architecture, July 1998
5. Mockapetris, P.: RFC 1034 – Domain Names-Concept and Facilities, November 1987
6. Mockapetris, P.: RFC 1035 – Domain Names-Implementation and Specification, November 1987
7. Mongay Batalla, J., Latoszek, W., Gajewski, M., Gut, H., Sienkiewicz, K., Binczewski, A., Gajda, B. i in.: Analysis of conditions and preparation of requirements for implementation of network prototype (M1.B.1„ M1.B.2), Raport projektu IIP, Warszawa 2011
8. Sienkiewicz, K., Gajewski, M., Batalla, J.M.: Testy IPv6 w sieci małego operatora, Przegląd Telekomunikacyjny i Wiadomości Telekomunikacyjne, SIGMA NOT, Warszawa, 2010, nr 8-9 , s. 804–805
9. VMware, `http://www.vmware.com/products/vcenter-server/`

10. Wprowadzenie do Hyper-V w systemie Windows Server 2008, http://technet.microsoft.com/pl-pl/library/wprowadzenie-do-hyper-v-w-systemie-windows-server-2008.aspx
11. Home page of XEN, http://www.cl.cam.ac.uk/research/srg/netos/xen/
12. Home page of KVM, http://www.linux-kvm.org/page/Documents
13. Thomson, S., Narten, T.: RFC 2462 – IPv6 Stateless Address Autoconfiguration, December 1998

Realizacja przełączeń terminali ruchomych przez elementy infrastruktury systemu mobilności

Michał Hoeft[1], Tomasz Gierszewski[1], Krzysztof Gierłowski[1], Józef Woźniak[1], Rafał Chrabąszcz[2], Piotr Pacyna[2]

[1] Katedra Teleinformatyki, Politechnika Gdańska
[2] Katedra Telekomunikacji, Akademia Górniczo-Hutnicza

Streszczenie W artykule przedstawione zostało rozwiązanie pozwalające zrealizować mobilność terminali, które nie posiadają zaimplementowanych zaawansowanych mechanizmów dedykowanych dla tej usługi. Opisano protokół Proxy Mobile IPv6 jako rozwiązanie wpisujące się w model network-based localized mobility management, w której to elementy infrastruktury systemu odpowiadają za zachowanie ciągłości połączenia w trakcie przełączania. Zaprezentowane zostały implementacje systemów oraz scenariusze przeprowadzonych badań, niezależnych dla każdego z ośrodków. Otrzymane rezultaty zostały opsiane w kontekście przerwy w transmisji danych odczuwalnej dla użytkownika systemu.

1 Wprowadzenie

Zwiększająca się liczba terminali ruchomych [1] oraz szacowany dalszy ich wzrost w kolejnych latach stawiają przed projektantami i administratorami systemów teleinformatycznych potrzebę implementacji mechanizmów zapewniających użytkownikom mobilnym nieograniczony, ciągły dostęp do sieci. Korzystanie z usług Triple Play, przeniesienie aplikacji użytkowych na terminale mobilne sprawiają, że istotny jest nie tylko sam dostęp do sieci, ale również jego ciągłość oraz nieprzerwana dostępność między użytkownikami w trakcie ich przemieszczania. Istotnym wymaganiem, w takim scenariuszu, jest zapewnienie użytkownikom możliwości zarówno inicjowania połączeń, jak i ich przełączania, warunkujących integralność odbioru przesyłanych wiadomości. Wymóg ten powinien być spełniony niezależnie od punktu, czy sposobu podłączenia użytkownika - również w przypadku przebywania w innych, niż ich macierzysta, sieciach. Wspomniane oczekiwania użytkowników rzutują na wymagania techniczne stawiane sieciom, z których istotne są mechanizmy zapewniające identyfikację, lokalizację, odpowiednią adresację terminali mobilnych oraz dodatkowe mechanizmy gwarantujące płynne przełączenia w trakcie zmiany punktu podłączenia do sieci. Podstawowy zestaw protokołów IPv6 nie jest wystarczający, aby zrealizować wszystkie wymagania. Na przeszkodzie stają wprowadzone na początku rozwoju Internetu mechanizmy identyfikacji oraz adresacji urządzeń końcowych, które nadal są stosowane. Powstaje zatem potrzeba nowego rozwiązania pracującego w sposób przezroczysty dla aplikacji i pozwalającego na ciągłość sesji w warstwie transportowej modelu ISO/OSI.

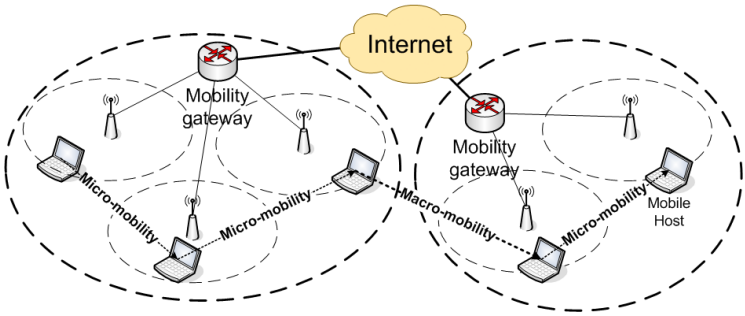
W artykule zaprezentowano dwie implementacje protokołu Proxy Mobile IPv6, który jest jednym z nowych rozwiązań mogących sprostać opisanym oczekiwaniom. Drugi rozdział pracy zawiera klasyfikację protokołów mobilności. Idea oraz architektura PMIPv6 zostały opisane w rozdziale trzecim. Szczegółowy opis procesu przełączania znajduje się w rozdziale czwartym. Opis implementacji oraz scenariusze badań zostały zaprezentowane w rozdziale piątym. W rozdziale szóstym zostały przedstawione i omówione wyniki przeprowadzonych prac. Artykuł zamyka podsumowanie.

2 Protokoły mobilności w sieciach IP

Klasyfikacji rozwiązań mobilności realizowanych w warstwie sieciowej można dokonać na kilka sposobów. Jednym z nich jest podział ze względu na rodzaj przemieszczającego się elementu systemu [2]. W takim przypadku można wyszczególnić mobilność terminali (terminal ma możliwość podłączenia do różnych punktów sieci), mobilność użytkownika (użytkownik może korzystać z różnych terminali, punktów podłączenia) oraz mobilność usługi (usługa jest dostępna dla użytkownika niezależnie do punktu podłączenia).W świetle powyższego podziału, szczególnie interesujący nas protokół Proxy Mobile IPv6 może zostać wykorzystany do realizacji mobilności terminala oraz usługi. Co istotne, umożliwia on zachowanie ciągłości świadczenia usługi realizowanej z użyciem terminala mobilnego. Jest to możliwe dzięki zachowaniu stałej adresacji przemieszczającego się terminala.

Inną, często spotykaną, klasyfikacją rozwiązań wsparcia mobilności w sieciach IP jest podział ze względu na jej zakres - mobilność międzydomenową (*inter-domain*) i wewnątrzdomenową (*intra-domain*), gdzie domena rozumiana jest jako fragment sieci podlegający jednolitemu zarządzaniu. Mobilność międzydomenowa nazywana także makromobilnością odnosi się do przełączeń między niezależnymi, pod względem zarządzania, systemami sieciowymi. W związku z tym stosowane rozwiązania okazują się daleko bardziej złożone - powinny zawierać takie elementy jak uwierzytelnienie, przydzielenie i weryfikację adresu IP oraz aktualizacje ścieżki przesyłania danych. Dla porównania, w przypadku mobilności wewnątrzdomenowej (mikromobilnoci) część z tych procedur może zostać uproszczona albo pominięta - np. mechanizmy autoryzacji czy sprawdzania adresacji.

Rozwiązanie Mobile IP [6], jako popularne rozwiązanie makromobilności, zrywa z tradycyjnym powiązaniem adresów węzła i sieci. Zastosowanie podwójnej adresacji rozwiązuje problem dostarczania pakietów w przypadku przemieszczania się użytkownika miedzy domenami. W klasycznym MIPv4 występował jednak inny znaczący problem - problem routingu trójkątnego, którego rozwiązaniem jest szeroka gama propozycji optymalizacji routingu. W przypadku MIPv6 możliwe jest np. poinformowanie węzła korespondującego o aktualnym adresie terminala mobilnego, co umożliwia obu węzłom bezpośrednią komunikację. Podejście proponowane w rozwiązaniach Mobile IP nie jest rozwiązaniem szczególnie efektywnym. Pozwala natomiast na sprawną realizację makromobilności w przypad-



Rysunek 1. Porównanie mikromobilności i markomobilności

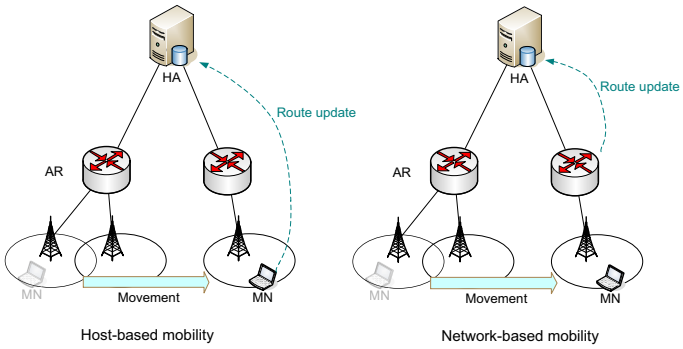
ku niezbyt częstego przełączania między domenami. W celu usprawnienia przełączania w systemach MIP proponowane są liczne mechanizmy zmniejszające opóźnienie występujące podczas przełączenia [3, 4].

Do rozwiązań pozwalających realizować mikromobilność, przy zachowaniu niewielkich opóźnień, można zaliczyć Cellular IP, HAWAII, czy TIMIP. Ich popularność jest jednak znacznie mniejsza niż protokołu MIP. Opisany w dalszej części rozdziału protokół Proxy Mobile IPv6 również jest rozwiązaniem pozwalającym realizować mikromobilność. Należy jednak zauważyć, że w tym przypadku do domeny zarządzania muszą być włączone poszczególne elementy infrastruktury systemu (LMA, MAG), lecz specjalnych odnoszących się do elementów infrastruktury sieciowej łączącej te elementy nie ma - wystarczająca jest standardowa łączność IPv6 w warstwie sieciwej modelu ISO/OSI.

Osobną kategorią rozwiązań wsparcia mobilności są rozwiązania wykorzystujące translację adresów. Mogą one zostać wykorzystane do realizacji zarówno mikro-, jak i makromobilności. Należy jednak zaznaczyć, że łatwość wdrożenia takich systemów może być powiązana z - niekiedy znaczącymi - ograniczeniami ich funkcjonalności. Przykładam może być rozwiązanie Reverse Address Translation (RAT), które realizuje makromobilność, lecz pozwala na obsługę jedynie ruchu UDP.

3 Protokół Proxy Mobile IPv6

Postępująca konwergencja sieci przewodowych i bezprzewodowych oraz sieci pakietowych z systemami komórkowymi powoduje, że nowe rozwiązania realizujące mobilność powinny być zgodne ze schematami stosowanymi w systemach komórkowych. Taką zgodność zapewnia model przełączeń obsługiwanych przez sieć (*network-based mobility management*), stanowiący alternatywę dla modelu przełączeń nadzorowanych przez terminal ruchomy (*host-based mobility management*) oraz koncepcja network-based localized mobility management (NETLMM) z protokołem Proxy Mobile IPv6 [5], [7], [8]. NETLMM stanowi alternatywę dla koncepcji realizowanej przez protokół Mobile IPv6 [6], w której



Rysunek 2. Porównanie protokołów MIPv6 oraz PMIPv6

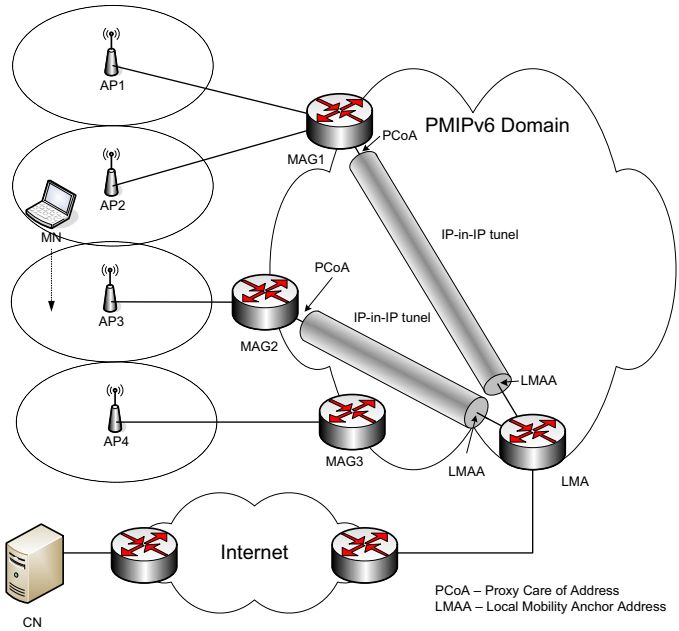
to terminal odgrywał istotną rolę w całym procesie przełączania. Ulokowanie rozwiązania mobilności w warstwie sieciowej modelu ISO/OSI pozwala na uniezależnienie od stosowanej techniki bezprzewodowej. Jej szczególna atrakcyjność związana jest z przeniesieniem obowiązku sygnalizowania przełączenia i rejestrowania użytkownika w domowej sieci IPv6 z terminala mobilnego na elementy sieci obsługującej to przełączenie. Dzięki takiemu podejściu możliwe jest świadczenie mobilności również terminalom nie posiadającym wsparcia dla tej usługi (*legacy mobile node*). Jest to istotny aspekt w kontekście zmniejszenia złożoności stosów sieciowych obsługiwanych przez terminale ruchome.

W modelu NETLMM to sieć, a nie terminal realizuje niezbędną funkcjonalność związaną z przełączeniem, pozostawiając terminalowi zadania detekcji łącza i autokonfigurację adresów. Realizacja przełączenia w warstwie sieciowej pozwala na zachowanie sesji warstwy transportowej i ciągłość transmisji danych. W przypadku PMIPv6 terminal może korzystać z topologicznie niepoprawnego adresu IPv6 - właściwego dla swojej sieci domowej, ale niepoprawnego dla aktualnego punktu podłączenia. Dostępność terminala możliwa jest dzięki zastosowaniu odpowiednich mechanizmów tunelowania i routingu realizowanych przez lokalny węzeł wsparcia mobilności (*local mobility anchor point*). Chociaż istnieją pewne podobieństwa między tradycyjnym MIPv6 oraz PMIPv6, zwłaszcza w kontekście realizowanych zadań, są to rozwiązania bazujące na odmiennych i przeciwstawnych modelach (rys. 2). Dodatkowo protokół PMIPv6 posiada szereg zalet w stosunku do MIPv6 [8], [9], do których możemy zaliczyć:

- zniesienie konieczności implementacji dodatkowych mechanizmów w terminalu mobilnym;
- przeniesienie obowiązku sygnalizacji oraz tunelowania na elementy infrastruktury systemu, skutkujące poprawą efektywności mechanizmu przełączania oraz lepszym wykorzystaniem pasma na bezprzewodowym łączu dostępowym;
- zwiększenie poufności informacji o aktualnym punkcie podłączenia użytkownika;

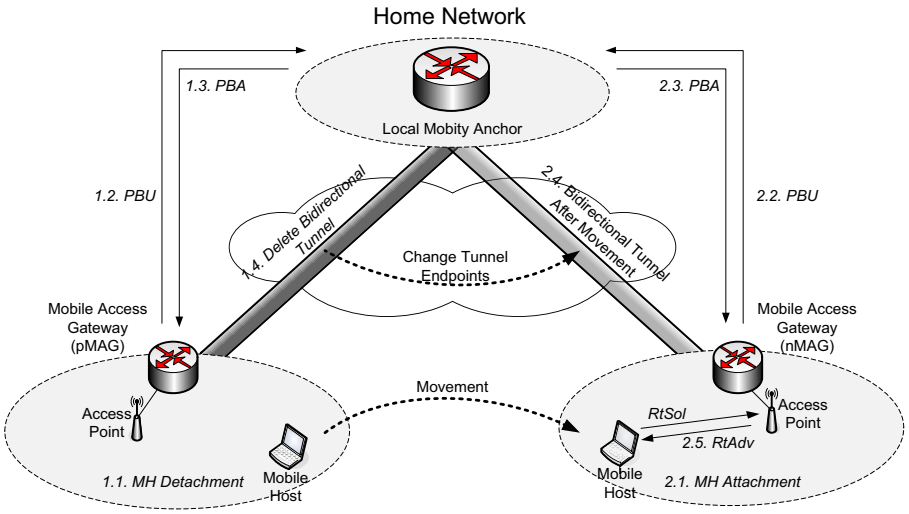
– możliwość skrócenia czasu przełączania poprzez zastosowanie oddzielnych prefiksów sieci dla poszczególnych terminali mobilnych i, w efekcie, możliwość pominięcia stosowania mechanizmów wykrywania duplikacji adresów.

Z powyższych powodów PMIPv6 wydaje się być rozwiązaniem bardzo atrakcyjnym dla operatorów sieci komórkowych. Proxy Mobile IPv6 jest włączony do specyfikacji 3GPP SAE/LTE jako jeden z protokołów mobilności.



Rysunek 3. Architektura Systemu PMIPv6

Standard Proxy Mobile IPv6 [5] wprowadza dwa dodatkowe elementy występujące w sieci - Mobile Access Gateway (MAG) oraz Local Mobility Anchor (LMA) (rys. 3). Elementy MAG współpracują z punktami dostępowymi w celu wykrycia próby przełączenia. Po wykryciu przełączenia następuje realizacja sygnalizacji obsługującej przełączenie w sposób sprzyjający zachowaniu ciągłości połączenia. Dla terminala mobilnego wszystkie elementy MAG widziane są jako ta sama brama domyślna do innych sieci (router L3), nawet po przełączeniu do innego punktu dostępowego (przełączenie L2). Pozwala to na zastosowanie niepoprawnej topologicznie adresacji, która jest jednak zgodna z adresacją sieci domowej użytkownika, co jest wygodne dla terminala, który mimo zmiany punktu przyłączenia nie musi konfigurować nowego adresu IP. Elementy LMA realizują funkcjonalność podobną do domowego agenta wsparcia mobilności (Home Agent MIPv6). Do ich zadań należy utrzymanie i monitorowanie informacji



Rysunek 4. Przełączenie w domenie PMIPv6

o trasach routingu do poszczególnych terminali mobilnych. Dodatkowo LMA określa prefiks sieci dla każdego terminala oraz nadzoruje dostępność mobilnych urządzeń w obrębie domeny.

Rozwiązanie Proxy Mobile IPv6 nie determinuje sposobu przydzielania adresów IPv6. Można zastosować stanową i bezstanową autokonfigurację adresów. W przypadku trybu bezstanowego, terminal mobilny wyznacza adres na podstawie odebranego w wiadomości Router Advertisment prefiksu sieci domowej (Home Network Prefix) oraz identyfikatora interfejsu, zgodnie ze standardowymi mechanizmami. W przypadku trybu stanowego elementy MAG muszą pełnić rolę pośrednika DHCPv6.

4 Realizacja przełączenia

Przełączenie w warstwie sieciowej nie może zostać zrealizowane bez współpracy z niższymi warstwami. Do prawidłowej realizacji mobilności, z wymaganiami opisanymi w poprzednich punktach, niezbędne są mechanizmy współpracy międzywarstwowej protokołu PMIPv6. Powinny one dostarczać informacje o wyrejestrowaniu (deasocjacji) z aktualnego punktu dostępowego oraz rejestracji (asocjacji) z nowym punktem. Współpraca ta pozwala na uzyskanie informacji dotyczących aktualnego punktu podłączenia oraz, jeżeli proces przełączania pozostaje w jego gestii, stanu procesu przełączania.

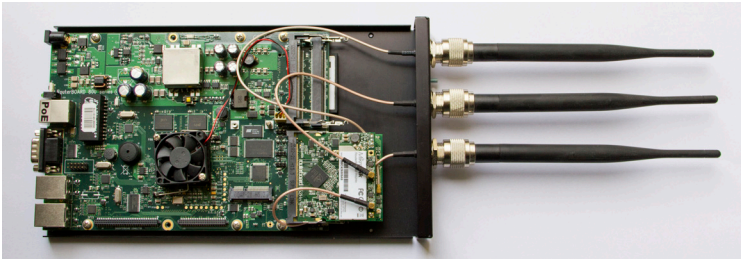
Na rysunku 4 zaprezentowano schemat przepływu wiadomości PMIPv6 w procesie przełączania między dwoma elementami MAG. Dla poprawy czytelności pominięto proces pierwszego podłączenia do sieci, opisany poniżej. Jego pierw-

szy etap to realizowane w warstwie L2 wykrycie przyłączenia użytkownika do sieci dostępowej nadzorowanej przez pMAG. Proces ten nie powinien polegać na interpretacji wiadomości przesyłanych przez terminal mobilny, lecz na komunikacji AP-MAG. Następnie pMAG po uwierzytelnieniu użytkownika pobiera jego profil. Zawarte są w nim informacje o identyfikatorze terminala, adresie LMA oraz sposobie konfiguracji adresów. W kolejnym kroku pMAG wysyła w imieniu terminala mobilnego do LMA wiadomość Proxy Binding Update. Jest ona rozszerzeniem standardowej wiadomości MIPv6 (Mobile IPv6) Binding Update. Jeżeli odebrana przez LMA wiadomość PBU jest poprawna, a użytkownik zostanie uwierzytelniony, następuje przygotowanie odpowiedzi. Najważniejsze, na tym etapie jest określenie przynajmniej jednego prefiksu sieci domowej (HNP), który zostanie przesłany do terminala mobilnego. Odpowiedź LMA przesyłana jest w postaci wiadomości Proxy Binding Acknowledge. Wymiana komunikatów PBU i PBA pozwala na uzyskanie informacji niezbędnych do zestawienia tunelu oraz konfiguracji routingu. Przydzielony prefiks sieci, tryb ustalania adresu oraz informacje routingu zostają przesłane przez pMAG do terminala w wiadomości Router Advertisement. Inaczej niż w standardowej transmisji, wiadomości RA jest przesyłana na adres link-local terminala mobilnego (transmisja unicast zamiast multicast).

Przełączenie użytkownika realizowane jest w bardzo podobny sposób. Po sygnale odłączenia od obsługiwanego punktu dostępowego ponownie obserwujemy wymianę wiadomości PBU i PBA (rys. 4 wiadomości 1.2, 1.3), których celem jest zakończenie aktualnej sesji mobilności (rys. 4 1.4). W tym samym czasie może odbywać się proces podłączenia użytkownika do kolejnego punktu dostępowego nadzorowanego przez nMAG (rys. 4 wiadomości 2.1-2.5). Dla poprawy czytelności rysunku wiadomości przesyłane między MAG i AP zostały zaznaczone tylko jedną linią.

5 Opis implementacji oraz realizowane scenariusze badań

W Katedrze Telekomunikacji AGH protokół PMIPv6 testowany był na urządzeniach fizycznych. Rolę Mobile Access Gateway (MAG) pełnił wydajny, konfigurowalny router MikroTik RouterBOARD 800 (rys. 5). System operacyjny dostarczony przez producenta wymieniono na Debian 6.0 "Squeeze" z jądrem dd-wrt w wersji 2.6.34.9, które zostało przygotowane dla tej platformy sprzętowej (PowerPC). Każdy z sześciu routerów wyposażony został w 3 aktywne interfejsy radiowe. Dzięki temu każdy router, przy użyciu oprogramowania hostapd, mógł realizować obsługę trzech niezależnych sieci bezprzewodowych. Taka konfiguracja sprzętowo-programowa umożliwiła wygodną instalację na nich dodatkowego oprogramowania, a przez to rozszerzenie dostępnych funkcjonalności. Wykorzystywana implementacja protokołu PMIPv6 to OpenAir PMIPv6 [10], która została zmodyfikowana i przystosowana do środowiska zbudowanego w Katedrze Telekomunikacji AGH. Rolę Local Mobility Anchor (LMA) pełnił serwer



Rysunek 5. Platforma testowa PMIPv6 w KT AGH

PC, pracujący pod kontrolą systemu Debian. W roli Mobile Node (MN) został użyty netbook pracujący również pod kontrolą systemu operacyjnego Debian.

Scenariusze realizowane w celu badania zachowania terminala mobilnego obejmowały przełączenia pomiędzy dwoma MAG, o różnych SSID według schematu MAG1 - MAG2. W czasie eksperymentów terminal pozyskiwał adres IPv6 w wyniku procesu samodzielnej konfiguracji adresu (*stateless address autoconfiguration*). W celu sprawdzenia wpływu przełączenia na aplikacje użytkowe posłużono się oprogramowaniem vlc do strumieniowania wideo.

Prace i badania wykonane przez zespół z Katedry Teleinformatyki PG przeprowadzone zostały z wykorzystaniem autorskiej implementacji protokołu PMIPv6. Jest ona rozszerzeniem implementacji standardowego protokołu Mobile IPv6 dostępnej w module Python Scapy [11]. Oprogramowanie zostało uruchomione w wirtualnym środowisku pod nadzorem systemu Microsoft Hyper-V. Aplikacje realizujące funkcjonalności LMA oraz MAG uruchomione zostały w systemie operacyjnym Debian Squeeze z jądrem 2.6.32. Ponieważ implementacja nie wymaga niestandardowego wparcia w postaci dodatkowych modułów jądra, nie powinno być problemów z jej uruchomieniem w innych systemach operacyjnych rodziny Linux. Elementy architektury PMIPv6 zostały połączone z punktami dostępowymi Cisco serii Aironet 1130 AG. Połączenie między terminalem mobilnym, a punktami dostępowymi zrealizowane zostało poprzez sieć 802.11a. Informacja o asocjacji i deasocjacji terminali mobilnych przesyłana była do odpowiedniego elementu MAG jako wiadomość syslog. Jak pokazały pomiary (punkt 6) taki sposób przekazywania informacji może mieć niekorzystny wpływ na czas przełączenia. Konfiguracja tuneli oraz routingu na elementach MAG i LMA realizowana jest przy pomocy narzędzia systemowego *ip*.

W badanych scenariuszach realizowanych na Politechnice Gdańskiej dokonywano przełączeń według dwóch schematów MAG1-MAG2-MAG1 oraz MAG2-MAG1-MAG2. Do każdego elementu MAG podłączone zostały, po jednym, punkty dostępowe AP rozgłaszające sieć o jednakowym SSID. Terminal mobilny konfigurował adres IPv6 automatycznie, zgodnie z automatycznym trybem bezstanowym (*stateles saddress autoconfiguration*). Zbadano terminale mobilne pracujące pod kontrolą systemów operacyjnych Linux Ubuntu oraz Windows 7.

W ramach przeprowadzonych badań dokonano testu działania systemu pod względem prawidłowego funkcjonowania w warstwie sieciowej z wykorzystaniem standardowych narzędzi - *ping6* oraz *mtr*. Przy pomocy tych aplikacji dokonywano też szacunkowego pomiaru czasu przełączania. Dodatkowo przeprowadzono badania wpływu wykorzystania rozwiązań PMIPv6 na działanie oprogramowania użytkowego. Jako przykład wybrano popularne protokoły FTP oraz SSH. Ponadto przy pomocy narzędzia *iperf* dokonano pomiarów przepływności porównanych z siłą sygnału odbieranego. Wyniki zostały omówione w rozdziale 6.

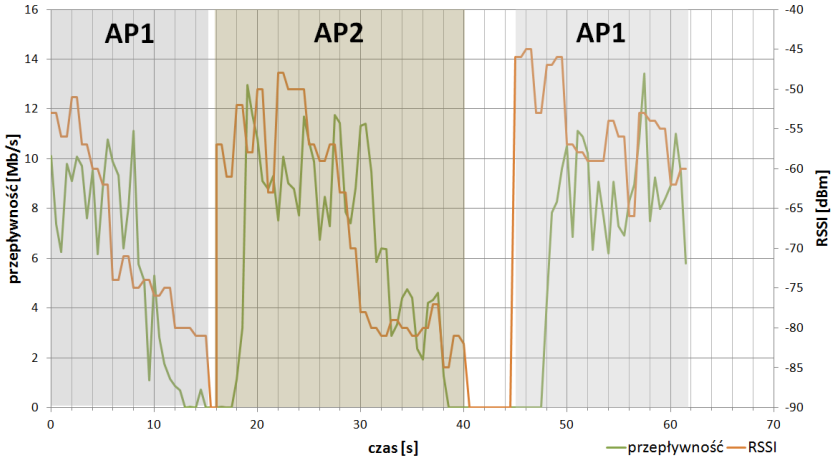
6 Omówienie wyników

W trakcie eksperymentów przeprowadzonych w KT AGH stwierdzono działanie systemy zgodne z oczekiwaniami. Uzyskano następujące pomiary czasu przełączeń obsługiwanych przez PMIPv6: czas przyłączenia do sieci – ok. 69 ms, wysłanie komunikatu Neighbor Adverisement zawierającego adres skonfigurowany przez PMIPv6 – ok. 3 sek., przebieg procedury PMIP w węźle MAG – ok. 28ms, w tym przebieg procedury PMIP w węźle LMA: ok. 13ms. Łączne opóźnienie przełączenia wynosiło około 3,7 s. W praktyce jednak opóźnienie to bywa często dłuższe i może wynosić kilka sekund. Znacząca jego część jest wnoszona przez obsługę standardowych komunikatów IPv6 wymienianych pomiędzy terminalem ruchomym i MAG, w tym Router Solicitation, które w przeprowadzonych eksperymentach wynosiło ok. 1.4s. Jest to charakterystyczne dla protokołu IPv6. Sama obsługa przeniesienia połączenia według PMIPv6 była bardzo szybka. Przełączenie było wymuszane z linii komend, co pozwoliło na uniknięcie zwykle dużego opóźnienia skanowania sieci po utracie sygnału.

Jak można przypuszczać, test polegający na strumieniowaniu wideo pokazał, że opóźnienie przełączenia jest dostrzegalne na obrazie przy transmisji z domyślnymi ustawieniami buforowania strumienia w sieci i w oprogramowaniu klienckim. Dalsze przyspieszenie procesu PMIPv6 wymagać będzie modyfikacji stałych czasowych protokołu w dopuszczalnym zakresie wartości oraz powtórzenia eksperymentów.

W trakcie badań przeprowadzonych na Katedrze Teleinformatyki PG stwierdzono poprawne działanie zaimplementowanego systemu. Przerwy w transmisji wahały w zakresie do kilku sekund. Na rysunku 6 zaprezentowano wykres przykładowego pomiaru poziomu sygnału oraz przepływności strumieni TCP w trakcie przełączania między AP1-AP2 oraz AP2-AP1. Obserwowaną przerwę w transmisji dla tych pomiarów można oszacować na odpowiednio 2s i 10s. Należy zauważyć, że czas przełączania w warstwie L2 może być różny - na zaprezentowanych wykresach wynosił on odpowiednio ok. 1s oraz 7s. Są to wartości odczuwalne dla użytkownika. Wartości te nie zależą jednak od opisywanej implementacji protokołu PMIPv6.

Spadek siły sygnału odbieranego w trakcie wychodzenia z zasięgu punktu dostępowego, ale przed odłączeniem do niego powoduje, że transmisja jest wstrzy-



Rysunek 6. Przykładowy pomiar siły sygnału oraz przepływności strumienia TCP w trakcie przełączenia

mywana jeszcze przed rozpoczęciem procesu przełączania. W związku z potencjalną możliwością znaczącej poprawy czasu przełączania w warstwie L2 mechanizmy te są przedmiotem dalszych prac zespołu.

Jak zostało to opisane w punkcie 5 sygnalizację o przełączeniu w warstwie L2 zrealizowano przy wykorzystaniu wiadomości syslog. Wykorzystywane modele punktów dostępowych przekazywały wiadomości syslog po około 1s od czasu zakończenia procesu uwierzytelnienia terminala ruchomego. Jest to dodatkowe opóźnienie, którego minimalizacja wymaga zastosowania innych mechanizmów sygnalizacji międzywarstwowej L2-L3.

Zaprezentowane na rys. 6 wyniki pomiarów pokazują, że w zależności od czasu przełączania w warstwach L2 i L3 zmienia się czas przerwy transmisji dla połączenia w warstwie transportowej. Widoczna jest różnica między czasem wznowienia transmisji TCP po przełączeniu dla pomiaru AP1-AP2 (krótkie przełączenie w L2) oraz AP2-AP2 (długie przełączenie w L2). Jest to związane z mechanizmem retransmisji protokołu TCP, w którym okresy pomiędzy kolejnymi retransmisjami są zwiększane.

Czas potrzebny na realizację przełączenia w warstwie sieciowej zgodnej z implantacją protokołu PMIPv6 trwał kilkadziesiąt milisekund, a więc był bardzo obiecujący. Należy podkreślić, że opisywana implementacja nie zawiera mechanizmów współpracy z serwerami AAA, które mogą wydłużyć ten czas.

7 Podsumowanie

Protokół PMIPv6 zgodny z koncepcją NETLMM realizowany jest w warstwie sieciowej, co skutkuje dużą uniwersalnością tego rozwiązania. W opisywanym

przypadku mobilność może być realizowana dla szeregu aplikacji korzystających z różnych protokołów warstwy transportowej oraz różnych protokołów warstwy aplikacji. Nie jest również wprowadzone żadne ograniczenie odnośnie zastosowanych technik bezprzewodowych.

Protokół PMIPv6 zaprojektowany został dla łączy point-to-point jednak, jak pokazują przeprowadzone eksperymenty, po wprowadzaniu modyfikacji mechanizmu rozgłaszania prefiksów sieci (transmisja unicast zamiast multicast pomiędzy MAG i MN), może być stosowany również w sieciach IEEE 802.11 (WLAN). Przeniesienie obowiązku realizacji mobilności z terminali mobilnych na elementy infrastruktury systemu znosi konieczność implementacji dedykowanych mechanizmów w terminalu mobilnym. Dodatkowo wprowadza możliwość realizacji mobilności również dla terminali z podstawowym zestawem protokołów IPv6 (*legacy mobile node*).

Realizowane prace pokazują skuteczne implementacje systemów PMIPv6 oferujące podstawową funkcjonalność. Jak pokazują zaprezentowane wyniki istnieje szereg zagadnień (np. optymalizacja przełączania w warstwie L2), które mogą usprawnić proces przełączania i skrócić jego czas trwania. Minimalizacja opóźnienie jest istotna zwłaszcza w kontekście współpracy z usługami multimedialnymi. Dalsze prace prowadzone będą nad implementacją platformy mobilności pozwalającej na korzystanie z takich usług.

Literatura

1. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update 2010–2015, February 1, 2011.
2. Jyh-Cheng C., IP-Based Next-Generation Wireless Networks: Systems, Architectures, and Protocols, John Wiley Sons, 2004
3. Malki El K. (ed.), "Low Latency Handoffs in Mobile IPv4", RFC 4881, June 2007.
4. R. Koodli, "Mobile IPv6 Fast Handovers", Internet Engineering Task Force, RFC 5268, Jun. 2008.
5. Gundavelli S., Leung K., Devarapalli V., Chowdhury K., Patil B., Proxy Mobile IPv6, Proposed Standard, IETF, RFC 5213, August 2008.
6. Johnson D., Perkins C., Arkko J., Mobility Support in IPv6, IETF, RFC 3775, June 2004.
7. Kong K. S., Lee W., Han Y. H., Shin M. K., You H., Mobility Management For All-Ip Mobile Networks: Mobile IPv6 vs. Proxy Mobile IPv6, IEEE Wireless Communications, April 2008.
8. Udugama A., UmerIqbal M., Toseef U., Goerg C., Fan C., Schlaeger M., Evaluation of a Network based Mobility Management Protocol : PMIPv6, IEEE VTC 2009, Barcelona, Spain, April 2009.
9. Wozniak J., Machan P., Gierlowski K., Hoeft M., Lewczuk M., Comparative analysis of IP-based mobility protocols and fast handover algorithms in IEEE 802.11 based WLANs Communications In Computer And Information Science, Springer 2011
10. Open Air Interface Proxy Mobile IPv6, Mobile Communications Department at EURECOM in Sophia-Antipolis, France, http://www.openairinterface.org/components/page1095.en.htm
11. Scapy,http://www.secdev.org/projects/scapy

Realizacja środowiska IPTV/IMS z wykorzystaniem oprogramowania typu opensource

Konrad Sienkiewicz, Jordi Mongay Batalla, Stanisław Janikowski

Instytut Łączności - Państwowy Instytut Badawczy

Streszczenie Z punktu widzenia operatorów telekomunikacyjnych architektura IMS pozwala na łatwiejsze uruchamianie nowych usług/aplikacji. Niewątpliwie jedną z takich usług jest usługa IPTV, Uruchomienie usług IPTV w środowisku IMS wymaga między innymi zapewnienia komunikacji pomiędzy platformą IMS a serwerem IPTV. W niniejszym artykule przedstawione zostały propozycje rozwiązań w obszarze w/w komunikacji, a także opisane zostały problemy i sposób ich rozwiania, które wystąpiły przy implementacji systemu IPTV/IMS w sieci laboratoryjnej IPv6.

1 Wprowadzenie

Architektura platformy IMS (ang. IP Multimedia Subsystem) została zdefiniowana przez 3GPP w celu dostarczenia usług multimedialnych opartych na protokole IP dla sieci mobilnych, po czym została ona zaadoptowana przez ETSI TISPAN dla wprowadzania nowych, multimedialnych usług z wykorzystaniem dostępów szerokopasmowych takich jak xDSL, czy FTTx w sieciach stacjonarnych. Z punktu widzenia operatorów, architektura IMS jest atrakcyjna, gdyż daje możliwość tworzenia nowych usług znacznie szybciej niż dotychczas, ponadto realizacja usług jest niezależna od technologii w sieci dostępowej (zarówno dla sieci przewodowych jak i bezprzewodowych).

Usługa IPTV (ang. Internet Protocol Television) – oznacza jedną z kluczowych usług telekomunikacyjnych jaką jest przesyłanie sygnału telewizyjnego w sieciach szerokopasmowych z protokołem IP. Realizacja usługi IPTV z wykorzystaniem możliwości jakie daje platforma IMS niesie szereg korzyści, szczególnie istotnych z punktu widzenia operatora sieci. Do korzyści tych można zaliczyć: jeden system rejestracji i autoryzacji użytkowników, sterowanie sesjami oraz wsparcie dla QoS, wspólny system billingowy, możliwość integracji z innymi usługami dostępnymi w IMS (np. presence, messaging) oraz wsparcie dla mobilności.

Wobec wyżej wymienionych korzyści, a także w obliczu coraz częściej stosowanych ofert typu multiplay, wydaje się, że dla operatora sieci (w szczególności przewodowej), który posiada uruchomioną platformę IMS, naturalnym krokiem będa działania podejmowane na rzecz uruchomienia usług IPTV z wykorzystaniem zasobów tej platformy.

W niniejszym artykule przedstawiamy propozycje rozwiązań dotyczących uruchomienia usług IPTV w sieci IPv6 z architekturą IMS, ze szczególnym uwzględnieniem zapewnienia komunikacji pomiędzy serwerem strumieniującym dla IPTV a platformą IMS, przy założeniu, że w ramach rozwiązania wykorzystywane jest jedynie oprogramowanie typu opensource. Opisujemy również problemy implementacyjne, które wystąpiły przy realizacji przez nas testowego środowiska IPTV/IMS oraz sposób ich rozwiązania.

2 Stosowane protokoły sterujące

W przypadku realizacji usługi IPTV jednym z najczęściej stosowanych protokołów sterujących jest protokół RTSP (ang. Real-Time Streaming Protocol). Protokół ten zapewnia sterowanie związanego z dostarczaniem danych czasu rzeczywistego pomiędzy terminalem IPTV a serwerem strumieniującym. Z wykorzystaniem protokołu RTSP negocjowane są miedzy innymi parametry transmisji danych użytkowych (obraz+dźwięk), która odbywa się za pomocą protokołu RTP (ang. *Real-Time Transport Protocol*), a także steruje się dostępem do usługi (np. funkcje PLAY, PAUSE, STOP). Protokół ten jednak nie zapewnia możliwości autoryzacji użytkowników.

W przypadku sieci IP opartych o platformę IMS, podstawowym protokołem sterującym jest protokół SIP (ang. *Sesion Initiation Protocol*) [7]. W oparciu o protokół SIP, realizowana jest sygnalizacja pomiędzy terminalami klientów oraz poszczególnymi elementami platformy IMS (np. CSCF, różnymi serwerami usługowymi). To za pomocą protokołu SIP zestawiane są sesje, w których realizowana jest komunikacja w ramach poszczególnych aplikacji. Co równie istotne w IMS przewiduje się, że zestawianie to powiązane jest z procedurą weryfikacji dostępności zasobów warstwy transportowej. Ponadto w sieci IMS realizowana jest autoryzacja użytkowników, a także możliwe jest tworzenie profili użytkowników z różnymi uprawnieniami w zakresie dostępu do usług.

Dla realizacji usług IPTV w sieci wykorzystującej platformę IMS należy zapewnić zarówno realizację funkcjonalności realizowanych przez IMS przy wykorzystaniu protokołu SIP (takich jak: autoryzacja, zaliczanie, sterowanie warstwą transportową) oraz funkcjonalności związanych ze sterowaniem w ramach usług IPTV, które realizowane jest z wykorzystaniem protokołu RSTP. Zakładając, że nieodzowne jest stosowanie protokołu SIP (tego wymaga architektura IMS), to możliwe są 2 rozwiązania:

- rozszerzenie specyfikacji SIP o funkcjonalności związane ze sterowaniem strumieniowanymi mediami tak, by protokół SIP mógł w całości zastąpić RTSP lub
- wykorzystanie obu protokołów wraz z określeniem szczegółowych zasad dotyczących powiązań pomiędzy tymi protokołami.

Za pierwszym rozwiązaniem (tylko SIP) przemawia stosowanie tylko jednego protokołu sterującego i stosunkowo łatwa integracja z innymi usługami, natomiast za rozwiązaniem, w którym oprócz SIP stosowany jest również protokół RTSP przemawia fakt, że większość obecnych rozwiązań korzysta z RTSP.

Warto również zaznaczyć, że na chwilę obecną w obszarze usług IPTV realizowanych w ramach platformy IMS trwają intensywne prace standaryzacyjne prowadzone między innymi prowadzone przez ETSI i Open IPTV Forum (OIPF). W przypadku ETSI ogólne zasady realizacji usługi IPTV w sieciach opartych o IMS określa standard TS 182 027[3]. Definiuje on szereg dedykowanych bloków funkcjonalnych oraz powiązań między nimi dla realizacji usługi IPTV, w skład której wchodzą: transmisja kanałów telewizyjnych (broadcast), VoD/CoD, PVR. Z kolei standard TS 183 063[2] określa szczegółowe zasady realizacji w/w usług, w tym procedury sygnalizacyjne.

Równolegle do standardów ETSI publikowane są standardy Open IPTV Forum, które skupia między innymi producentów, usługodawców. OIPF w [4] wyróżnia dwa scenariusze dla realizacji usług IPTV. Jeden dotyczy sieci operatorskich IP (zarządzanych), drugi publicznej sieci Internet (niezarządzanych). W przypadku sieci zarządzanych IP, dla realizacji usług IPTV podobnie jak w ETSI wykorzystywana jest architektura IMS.

3 Zaproponowany sposób implementacji w środowisku laboratoryjnym IPv6

Jako główny cel postawiliśmy sobie implementację rozwiązania pozwalającego na realizację usług IPTV w środowisku IPv6 z platformą IMS. Przy realizacji sieci laboratoryjnej przyjęto założenie, że całość tej sieci oparta będzie o ogólnodostępne oprogramowanie, przy czym preferowane będzie oprogramowanie typu opensource.

W ramach implementacji w naszej sieci laboratoryjnej IPv6 przewidziano do wdrożenia: (1) platformę IMS, (2) klienta IPTV, (3) serwer strumieniujący IPTV oraz (4) serwer aplikacyjny IPTV, który realizuje autoryzację i kontrolę dostępu użytkowników do treści IPTV.

W tym celu dokonano przeglądu wybranych dostępnych modułów i aplikacji dla zbudowania wymienionego systemu w sieci laboratoryjnej IPv6.

W wyniku przeprowadzonej analizy zdecydowano, że **platforma IMS** oparta zostanie o rozwiązanie OpenIMSCore. Jest to jedno z pierwszych rozwiązań platformy IMS typu opensource opracowane i udostępnione przez Fraunhofer Institute [8]. Za jego wyborem przemawia fakt, że jest to rozwiązanie stale rozwijane, a także jest ono wykorzystywane przez wiele zespołów badawczych w pracach związanych z IMS. W skład platformy OpenIMSCore wchodzą HSS [ang. *Home Subscriber Server*] oraz CSCF [ang. *Call Session Control Function*], w ramach którego zaimplementowane są: P-CSCF [ang. *Proxy - Call Session Control Function*], I-CSCF [ang. *Interrogation - Call Session Control Function*], (ang. S-CSCF [ang. *Serving- Call Session Control Function*].

Platforma OpenIMSCore zapewnia szereg standardowych interfejsów (np. do dołączenia serwerów aplikacyjnych AS) i co istotne wspiera protokół IPv6.

Niewątpliwie liczba darmowych aplikacji **klientów IMS** jest znacznie mniejsza niż na przykład aplikacji terminala SIP. W większości aplikacje te są wynikiem projektów badawczych w obszarze IMS, niemniej można też znaleźć roz-

wiązania, które w większym stopniu przeznaczone są do komercyjnego wykorzystania (np. Boghe [12]). W tabeli poniżej dokonano porównania najbardziej popularnych klientów IMS z uwzględnieniem zarówno obsługi IPTV jak i protokołu IPv6.

Tabela 1. Porównanie możliwości klientów IMS

Klient IMS	**IPv6**	**Usługi multimedialne**		**Software**		
		VoD/IPTV	VoIP	Windows	Linux	Opensource
Mercuro [9]	+[1]	-	+	+	-	-
Monster [10]	+	-	+	+	+	+
UCT Client [11]	-	+	+	-	+	+
Boghe [12]	+	-	+	+	-	+

W wyniku przeprowadzonej wstępnej weryfikacji w naszym rozwiązaniu zdecydowaliśmy się zastosować UCT (University of Cape Town) Client jako klient IPTV/IMS.

W przypadku **serwerów strumieniujących video** dostępnych jest szereg darmowych rozwiązań. Poszczególne rozwiązania różnią się oferowanymi funkcjonalnościami, rodzajami licencji, systemem operacyjnym, a także językiem, w którym stworzono kod źródłowy. Z naszego punktu widzenia istotne było by serwer strumieniujący umożliwiał strumieniowanie video z wykorzystaniem protokołów RTP/RTSP, miał kody źródłowe w języku C++ oraz możliwe było jego uruchamianie na platformach sprzętowych z systemem operacyjnym Windows i Linux. Pewnym wskazaniem na konkretne rozwiązania serwerów była też ich popularność, ze szczególnym uwzględnieniem wykorzystania w różnych projektach badawczych. Uwzględniając powyższe wymagania, można wskazać co najmniej 3 rozwiązania serwerów strumieniujących: Darwin Streaming Server [13] - wersja opensource Apple QuickTime Streaming Server dostępna dla platform Linux/Windows/Solaris, Live555 Media Server [14] – rozwiązanie opensource serwera strumieniującego obsługa protokołu RTSP Server dostępna dla platform Linux/Windows/Mac, VLC (VideoLan Client) [15] – aplikacja serwera strumieniującego dla platform Linux/Windows, opracowana w ramach VideoLAN Project oferująca poza protokołem RTSP szereg innych protokołów (np. MMSH). Do zastosowania w naszej sieci laboratoryjnej zdecydowaliśmy się wykorzystać serwer strumieniujący VLC, który w dużej części oparty jest o kod źródłowy serwera Live555. Trzeba jednak nadmienić, że żaden z w/w serwerów strumieniujących nie obsługuje protokołu SIP (połączenie z IMS).

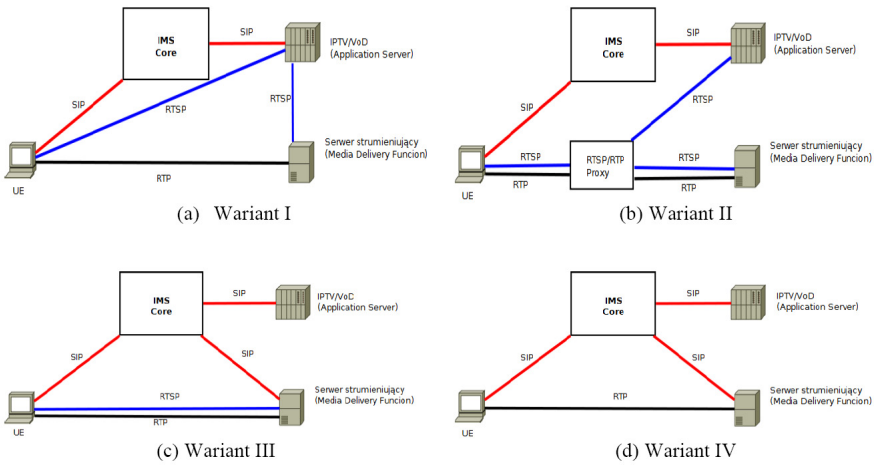
Jednym z rozwiązań dla **serwera aplikacyjnego** jest serwer UCT IPTV AS, który został opracowany przez zespół naukowców z University of Cape Town. Ograniczeniem tego rozwiązania jest fakt, że jest ono dostosowane jedynie do wykorzystania w sieciach IPv4 oraz nie jest zaimplementowana komunikacja między UCT IPTV AS i serwerem strumieniującym.

Zakładając, że do realizacji usługi IPTV w sieci z protokołem IPv6 wykorzystane zostaną wskazane powyżej oprogramowanie, to warto wspomnieć, że

zarówno serwer strumieniujący VLC jak i aplikacja UCT Client dla obsługi protokołu RTSP korzystają z bibliotek Live555, które nie obsługują protokołu IPv6. Stąd też w odniesieniu w/w aplikacji wymagane jest wprowadzenie zmian do kodu źródłowego oraz rekompilacja mająca na celu zapewnienie obsługi protokołu IPv6, co zostało szerzej omówione w rozdziale 5. Ponadto niezbędne jest rozszerzenie możliwości serwera aplikacyjnego o obsługę protokołu IPv6, a także implementację protokołu SIP w serwerze strumieniującym (w zależności od wariantu).

Integracja środowiska laboratoryjnego z wybranych elementów wykazała, że kluczowym zadaniem do realizacji jest wprowadzenie rozszerzeń do sygnalizacji, mających na celu eliminację możliwości niekontrolowanego korzystania z zasobów serwera strumieniującego przez użytkowników. Wynika to z faktu, że znajomość przez użytkownika adresu docelowego RTSP, identyfikującego żądany kontent na serwerze strumieniującym, pozwala mu (a również innym użytkownikom, którzy znają ten adres) na korzystanie w sposób nieograniczony (bez udziału IMS) z danego zasobu serwera strumieniującego.

Na etapie przygotowywania założeń dla realizacji laboratoryjnej sieci IPv6, w której możliwa będzie realizacja usług IPTV z wykorzystaniem zasobów platformy IMS zidentyfikowano 4 warianty możliwych rozwiązań, które między innymi mają na celu eliminację możliwości niekontrolowanego korzystania z zasobów serwera strumieniującego. Odzwierciedlają one zarówno kierunki prowadzonych prac badawczych (raczej w obszarze IPv4) jak i prac standaryzacyjnych.Warianty te zostały zaprezentowane na rysunku 1 i scharakteryzowane poniżej.



Rysunek 1. Realizacja usługi IPTV/IMS w poszczególnych wariantach

W ramach pierwszego z rozpatrywanych wariantów –Rysunek 1(a)– przyjęto założenie, że nie są wprowadzane żadne zmiany w obrębie protokołów sygnaliza-

cyjnych SIP i RTSP. Eliminacja możliwości niekontrolowanego korzystania przez użytkowników z zasobów serwera strumieniującego osiągana jest w ten sposób, że sygnalizacja RTSP pomiędzy UE i serwerem strumieniującym odbywa się za pośrednictwem serwera aplikacyjnego IPTV (realizuje on funkcję Proxy dla RTSP), natomiast strumienie RTP od serwera strumieniującego przesyłane są bezpośrednio do UE. Niewątpliwą zaletą tego wariantu jest jego prostota. Konieczne jest jedynie zaimplementowanie w serwerze aplikacyjnym funkcjonalności Proxy dla protokołu RTSP oraz wymagane jest by serwer aplikacyjny podawał adres UE w polu *destination* w wiadomości *Setup* protokołu RTSP. Przy realizacji funkcjonalności Proxy dla RTSP nie występuje zagrożenie natury wydajnościowej (jest to Proxy jedynie dla sygnalizacji). Problemem może być jednak znalezienie serwera strumieniującego, poprawnie obsługującego wartość pola *destination* w wiadomości *Setup* protokołu RTSP. W wyniku naszych testów ustaliliśmy, że popularne darmowe serwery strumieniujące jak VLC, Darwin nie obsługują tego parametru. Serwer VLC i ignoruje wartość pola destination (strumieniowanie mediów odbywa się na adres IP hosta, od którego przysłana została wiadomość *Setup*), natomiast serwer Darwin zwraca błąd "*Invalid Code*".

Wariant II –Rysunek 1(b)– nawiązuje do propozycji zaprezentowanej w [1] (dla sieci IPv4). Wariant ten różni się od wariantu I tym, że funkcjonalność Proxy oprócz protokołu RTSP obejmuje również pakiety wysyłane w protokole RTP. Takie rozwiązanie nie wymaga stosowania pola *destination* w wiadomości *Setup* protokołu RTSP. Co z kolei sprawia, że mogą być wykorzystywane dowolne serwery strumieniujące obsługujące IPv6 (również te nie obsługujące w/w pola).

Do zalet tego wariantu należy zaliczyć brak konieczności wprowadzania zmian po stronie zarówno UE jak i serwera strumieniującego. Niewątpliwie wadą jest fakt, że cały kontent przesyłany jest poprzez Proxy, co może powodować problemy natury wydajnościowej. Na niekorzyść tego wariantu przemawia również to, że jest to rozwiązanie niezgodne ze standardem ETSI.

Wariant III –Rysunek 1(c)–, co do ogólnej zasady, jest zgodny z rozwiązaniem rekomendowanym przez ETSI wyspecyfikowanym w [2] i [3]. W odróżnieniu od dwóch poprzednich wariantów, wariant III przewiduje, że na potrzeby zestawienia sesji, serwer strumieniujący będzie komunikował się z serwerem aplikacyjnym i UE z wykorzystaniem sygnalizacji SIP, natomiast sterowanie mediami w ramach usługi IPTV odbywa się bezpośrednio pomiędzy UE a serwerem strumieniującym z wykorzystaniem protokołu RTSP.

To rozwiązanie wymaga zaimplementowanie po stronie serwera strumieniującego przynajmniej częściowej funkcjonalności MCF (Media Control Function, komunikującej się z serwerem aplikacyjnym protokołem SIP, w ramach której zapewniona zostanie kontrola dostępu do kontentu. Do zalet tego wariantu należy zaliczyć fakt, że rozwiązanie to jest oparte o standardy ETSI, zatem można zakładać, że w przyszłości możliwe będzie wykorzystanie innych terminali klienckich. Pozwala on na pełne wykorzystanie możliwości jakie oferuje platforma IMS (np. zapewnienie zasobów warstwy transportowej). Ponadto nie są wymagane żadne zmiany w obrębie protokołu RTSP. Z drugiej strony, rozszerzenie funkcjonalności serwera strumieniującego o obsługę protokołu SIP, która wymaga

również zmian po stronie serwera aplikacyjnego IPTV, jest zadaniem znacznie bardziej skomplikowanym od modyfikacji proponowanych we wcześniejszych wariantach.

Wariant IV –Rysunek 1(d)– zakłada, że pełna sygnalizacja związana z realizacją usług IPTV przesyłana jest za pomocą protokołu SIP, natomiast nie wykorzystywany jest protokół RTSP. Rozwiązania bazujące na podobnych założeniach zostały zaproponowane w [5], [6]. Zaletą tego rozwiązania jest między innymi pełna unifikacja protokołów sterujących wykorzystywanych w ramach platformy IMS oraz pełna kontrola platformy IMS nad przebiegiem realizacji usługi. Na niekorzyść tego wariantu świadczą konieczność wprowadzania zmian w funkcjonującym od lat protokole SIP, brak jakichkolwiek aplikacji klienckich oraz serwerów strumieniujących, w których pełne sterowanie IPTV odbywa się za pomocą protokołu SIP.

W wyniku pogłębionych analiz postanowiliśmy w dalszych pracach skoncentrować się jedynie na wariantach I i III. Spowodowane to jest faktem, że wariant I jest najprostszy w realizacji przy założeniu, że dostępny jest serwer strumieniujący pozwalający na przesyłanie kontentu na adres inny niż adres IP klienta RTSP (serwer Live555). Realizacja tego wariantu pozwoli na względnie szybką implementację usługi IPTV w naszej sieci, gdyż wymaga jedynie drobnych zmian w oprogramowaniu dostępnych komponentów. Wariant III powinien być traktowany jako wariant docelowy, gdyż jest to rozwiązanie zgodne z zaleceniami ETSI. W związku z czym realizacja tego wariantu umożliwi nam w przyszłości stosowanie innych aplikacji, które będą zgodne ze standardami ETSI. W przypadku wariantów II i IV, które również realizują postawiony przez nas cel, zaprzestano dalszych prac, z uwagi na ich pracochłonność, a także fakt, że nic nie wskazuje, by rozwiązania te w przyszłości cieszyły się większym zainteresowaniem.

4 Podstawowe zagadnienia implementacji

W tym rozdziale omówione zostały główne zagadnienia implementacyjne związane z realizacją transmisji sygnalizacji (RTSP i SIP) na potrzeby IPTV w sieci IPv6.

W przypadku serwera aplikacyjnego IPTV, z uwagi na jego prostotę oraz fakt, że bazuje on na bibliotece eXosip2, która posiada już obsługę IPv6, zmiany w kodzie źródłowym ograniczyły się do wprowadzenia do kodu źródłowego linii eXosip_enable_ipv6(1); powodującej włączenie obsługi IPv6 w bibliotece eXosip2, oraz zmiany typu gniazda sieciowego z AF_INET na AF_INET6.

Zapewnienie obsługi IPv6 w kliencie UCT wymagało więcej pracy, ze względu na znacznie większą złożoność tego programu. Dodatkowo zaszła konieczność modyfikacji bibliotek Live555 w celu wprowadzenia tam obsługi IPv6, ponieważ klient UCT korzysta z bibliotek VLC do obsługi IPTV, te zaś biblioteki korzystają z bibliotek Live555 do obsługi połączeń RTSP.

W kodzie źródłowym samego klienta UCT oprócz analogicznych zmian jak w opisanym wyżej przypadku serwera aplikacyjnego IPTV konieczne było wprowadzenie dalszych modyfikacji. Z uwagi na błąd w bibliotece eXosip2 powodują-

cy, że dla IPv6 funkcja eXosip_guess_localip zwracała zawsze adres ::1 zamiast, jak specyfikuje dokumentacja biblioteki, adresu związanego ze ścieżką domyślną musieliśmy wprowadzić własny sposób określania własnego adresu IP na potrzeby protokołu SDP. W naszej implementacji wykorzystaliśmy do tego plik specjalny /proc/net/if_inet6 zawierający informacje o wszystkich skonfigurowanych w systemie adresach IPv6. Ponadto, ze względu na fakt, że pojedynczy interfejs sieciowy ma zazwyczaj w przypadku szóstej wersji protokołu IP przyporządkowany więcej niż jeden adres, w zmienionej przez nas wersji klienta UCT identyfikujemy interfejsy sieciowe po ich nazwie, a nie przypisanym im adresie IP. Podczas implementacji IPv6 zmieniliśmy również sposób rozwiązywania nazw DNS na getaddrinfo, zapewniający wsparcie dla wszystkich rodzin adresów oraz instrukcje do konwersji adresów IP między postacią binarną a postacią łańcucha znakowego, gdzie instrukcje inet_ntoa i inet_addr wspierające tylko IPv4 zastąpiliśmy bardziej uniwersalnymi instrukcjami inet_ntop i inet_pton. Zmieniliśmy również struktury używane do przechowywania adresów na sockaddr_in6, które zapewniają obsługę IPv6.

W związku z faktem, że rozwój klienta UCT IMS przez Uniwersytet w Cape Town zakończył się w lutym 2009r, zaś biblioteki VLC, na których opiera się obsługa strumieni wideo w tej aplikacji przeszła od tego czasu wiele zmian, w tym zmianę API używanego do sterowania mediami, musieliśmy również dokonać zmian nie związanych bezpośrednio z wprowadzeniem obsługi IPv6, a związanych z zapewnieniem poprawnego działania aplikacji we współczesnych systemach operacyjnych. Klient UCT IMS w wersji 0.13, ostatniej udostępnionej przez Uniwersytet w Cape Town, korzystał z biblioteki VLC w wersji 0.9.8 używając do sterowania odtwarzaniem MediaControlAPI[2] . Współcześnie dostępna jest wersja 1.1 bibliotek VLC, w której zrezygnowano ze starego interfejsu programistycznego i zalecono przejście na zaktualizowane LibVLC API[3]. Zmiany, które wprowadziliśmy polegały głównie na zamianie dołączanych plików nagłówkowych oraz zmianie w pliku rtsp.c używanych funkcji na analogiczne pochodzące z nowego API. Przy aktualizacji używanego API wykryliśmy oraz poprawiliśmy znaleziony błąd powodujący poważny wyciek pamięci, związany z brakiem deinicjalizacji biblioteki VLC po użyciu.

Kod źródłowy programu oraz bibliotek VLC nie wymagał wprowadzenia zmian, ponieważ cała obsługa interesującego nas strumieniowania RTSP/RTP jest tam realizowana przy użyciu bibliotek Live555. Niestety jednak, biblioteki te nie obsługiwały IPv6 i wymagały znacznych modyfikacji w celu zapewnienia wsparcia dla tego protokołu. Najwięcej zmian dotyczyło modułu Groupsock, odpowiedzialnego za tworzenie i zarządzanie gniazdami sieciowymi. W wielu miejscach należało wprowadzić zmiany podobne do tych w kliencie UCT, czyli użycie innego zbioru instrukcji do rozwiązywania nazw DNS oraz do konwersji zapisu adresów. Jak podają autorzy projektu Live555, założenie, że adres sieciowy ma 32 bity jest zakodowane na sztywno w wielu miejscach aplikacji. Wymagało to analizy całego kodu źródłowego bibliotek, ponieważ np. strukturą in6_addr nale-

[2] http://wiki.videolan.org/MediaControlAPI
[3] http://wiki.videolan.org/LibVLC

żało zastąpić nie tylko zmienne typu in_addr, ale również wiele zmiennych typu unsigned tam, gdzie służyły one do przechowywania adresu. Fakt, że 128 bitowy adres IPv6, w przeciwieństwie do 32 bitowego adresu IPv4 nie jest w stanie zmieścić się w pojedynczej zmiennej nie będącą strukturą, spowodował również, że inny jest sposób inicjalizacji oraz porównywania tych adresów. Podczas gdy adres IPv4 można z powodzeniem zainicjować lub nadpisać zerem, w przypadku adresu IPv6 konieczne było użycie specjalnej wartości IN6ADDR_ANY_INIT przy inicjalizacji i in6addr_any przy zamazaniu zmiennej adresem nieokreślonym. Podobnie, do sprawdzenia, czy dany adres jest adresem nieokreślonym musieliśmy użyć makra IN6_IS_ADDR_UNSPECIFIED, zamiast, jak w przypadku IPv4 po prostu porównać adres z wartością 0.

Niestety, z uwagi na zmiany w API do multicastów w przypadku IPv6, zmiany spowodowały utratę ich obsługi i w zmienionej przez nas wersji biblioteka działa tylko w przypadku transmisji unicast. Jest to jednak wystarczające na potrzeby usługi IPTV, zaś podejmując dodatkowy wysiłek programistyczny możliwe jest dodanie obsługi IPv6 multicast w przyszłości.

5 Podsumowanie

W niniejszym artykule przedstawione zostały propozycje rozwiązań dotyczących uruchomienia usług IPTV w sieci IPv6 z architekturą IMS. Scharakteryzowane zostały główne problemy jakie wystąpiły w trakcie uruchomiania systemu IPTV/IMS realizowanego oparciu o oprogramowanie typu opensource, a także przedstawione zostały propozycje rozwiązań tych problemów.

W ramach zagadnień implementacyjnych zostały szczegółowo omówione modyfikacje, które były niezbędne do uruchomienia transmisji IPTV w sieci IPv6 z wykorzystaniem wybranych modułów/aplikacji. Ponadto określone zostały możliwe warianty rozwiązań dla zapewnienia komunikacji pomiędzy serwerem strumieniującym dla IPTV a platformą IMS, które planowane są do zrealizowania w dalszej fazie implementacji.

Literatura

1. Zelalem Shibeshi, Alfredo Terzoli, Karen Bradshaw „Using an RTSP Proxy to implement the IPTV Media Function via a streaming server", International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT) 2009
2. ETSI TS 183 063 V3.5.2 "Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN); IMS-based IPTV stage 3 specification", Marzec 2011
3. ETSI TS 182 027 V3.5.1 "Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN); IPTV Architecture; IPTV functions supported by the IMS subsystem", Marzec 2011
4. Open IPTV Forum, "OIPF Functional Architecture [V2.1]", Valbonne – Francja, Marzec 2011

5. Steven Whitehead, Marie-Jose Montpetit, Xavier Marjou, "An Evaluation of Session Initiation Protocol (SIP) for use in Streaming Media Applications", draft-whitehead-sip-for-streaming-media-00.txt, 2006
6. S. Sivasothy, G. M. Myoung, and N. Crespi. "A unified session control protocol for IPTV service", In Proceedings of the 11th international conference on Advanced Communication Technology, pages 961–965, Gangwon-Do, South Korea, February 15-18, 2009.
7. J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, E. Schooler "Request for Comments: 3261 SIP: Session Initiation Protocol", 2002
8. Projekt OpenIMSCore, homepage: http://www.openimscore.org/
9. Mercuro IMS Client, homepage:http://mercuro-ims-client.co.tv/
10. Monster Client, homepage:http://www.monster-the-client.org/
11. UCT IMS Client, homepage:http://uctimsclient.berlios.de/
12. Boghe Client, homepage:http://www.doubango.org/
13. Darvin Streaming Server , homepage:http://dss.macosforge.org/
14. Serwer strumieniujący Live555, homepage:http://live555.com/mediaServer/
15. Projekt Video LAN Client, homepage:http://www.videolan.org/

Wykaz autorów