

EQ-BGP: an efficient inter-domain QoS routing protocol

Andrzej Beben
Institute of Telecommunications
Warsaw University of Technology
Nowowiejska 15/19, 00-665 Warsaw, Poland
abeben@tele.pw.edu.pl

Abstract

In this paper we propose the Enhanced QoS Border Gateway Protocol¹ (EQ-BGP) aimed at performing the inter-domain QoS routing in a multi-domain IP network. The objective of EQ-BGP is to establish end-to-end paths that offer the most suitable QoS guarantees taking into account both the QoS capabilities of particular domains as well as inter-domain links. Thanks to this, the EQ-BGP supports providing of end-to-end QoS, what cannot be achieved with the aid of currently used BGP-4 protocol. Therefore, we enhance the existing BGP-4 protocol with four QoS components that are: (1) new path attribute conveying information about QoS capabilities of the path, (2) the QoS aggregation function used to calculate the path's QoS capabilities based on contributions of particular domains and inter-domain links, (3) the QoS aware decision algorithm that allows us to select the best path fulfilling the QoS requirements, and (4) multiple routing tables that allow border routers to keep separate paths that are optimised for different QoS objectives.

In the paper we provide description of the proposed EQ-BGP protocol and then we focus on performance evaluation corresponding to network convergence and protocol scalability. The obtained results say that EQ-BGP gives stable routing and its performance is at similar level as the BGP-4 protocol.

1. Introduction

Providing Quality of Service (QoS) in multi-domain IP network is currently widely investigated. The commonly recognised approach follows the concept of end-to-end Classes of Service (CoS) [9], [3], [6] that represent a set of traffic requiring from the network similar guarantees cor-

responding to packet transfer characteristics as packet delay, losses and jitter. This approach assumes that globally known e2e CoSs are supported inside particular domains as well as on inter-domain links. In Table 1 we present a representative set of e2e CoSs jointly with their QoS objectives that are proposed inside the EuQoS project [4] based on the proposal from IETF and ITU-T.

Despite that e2e CoSs are globally well known and have strictly defined the e2e QoS requirements the QoS level assured in particular domains and inter-domains links is usually different. The main reason is that domains are designed, provisioned and administrated by independent providers that usually have their own policy for providing QoS. Taking into account that the QoS experienced by users is influenced by all domains being on the path, therefore we point out that development of an efficient inter-domain QoS routing protocol becomes important for providing the e2e QoS. The objectives of such QoS routing is to select and establish paths that provide the most attractive QoS level from the point of view of e2e CoS requirements.

Let us recall, that while there is a number of proposals for performing intra-domain QoS routing, where main role plays QOSPF (QoS Routing Mechanisms and OSPF Extensions) proposed in [2], the solutions for performing inter-domain QoS routing are still under investigations. Most of the proposed approaches assume enhancement of the currently used BGP-4 protocol. The first extension of BGP-4

e2e Class of Service	e2e CoS objectives		
	mean IPTD [ms]	IPDV [ms]	IPLR
telephony	100	50	10^{-3}
real time interactive	100	50	10^{-3}
mm streaming	1000	-	10^{-3}
high throughput data	1000	-	10^{-3}
standard	-	-	-

Table 1. e2e Classes of Services.

¹This work was partially sponsored by EU Commission under the 6FR IST EuQoS project.

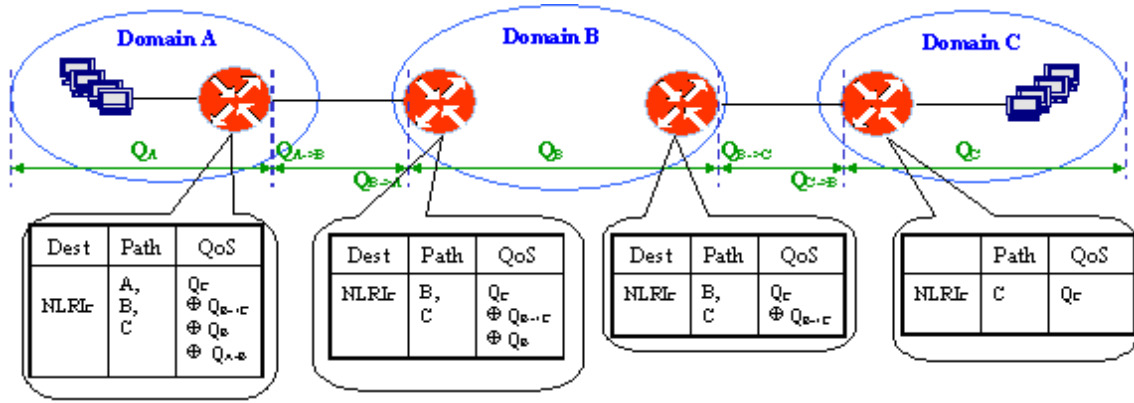


Figure 1. Example of EQ-BGP operation.

was proposed by B.Abarbanel in IETF draft [1]. He assumes that BGP update message includes additional parameter, called TE (Traffic Engineering) weight that influence the routing decision process. This approach was further enhanced in [5] by formal definition of an optional QoS_NLRI (Quality of Service Network Layer Reachability Information) path attribute designed for conveying information about QoS experienced along the path. In [5] the authors describe an optional attribute that conveys a single QoS performance characteristic per qBGP announcement, while this limitation has been relaxed in [1] since it might be necessary to carry a list of QoS performance characteristics in a single q-BGP update message.

In this paper, we propose the Enhanced QoS Border Gateway Protocol (EQ-BGP) aimed at performing the inter-domain QoS routing in a multi-domain IP network. The objective of EQ-BGP is to establish the the most suitable e2e paths with respect to the offered QoS level. The investigated EQ-BGP protocol extends currently used BGP-4 protocol with four components, that are: (1) QoS_NLRI optional path attribute, (2) the QoS aggregation function, (3) the QoS aware decision algorithm, and (4) multiple routing tables that allow routers to keep separate paths for each e2e CoSs.

To prove the efficiency of the EQ-BGP, we perform a number of simulation experiments focusing on protocol performance and scalability. In our experiments we analyse the impact of basic stressing events in the networks of different topology.

The organisation of the paper is the following. In section 2 we provide detailed description of the EQ-BGP protocol pointing out on the main features having impact on the protocol performance. The description of the carried out experiments, obtained results and conclusions are included in section 3. Finally, section 4 summarises the paper and gives an outline on the further work.

2. EQ-BGP protocol

The EQ-BGP is the extension of BGP-4 inter-domain routing protocol that fixes routing paths taking into account the QoS capabilities of particular domains. For that purpose, the border routers running EQ-BGP protocol, called EQ-BGP routers, advertise to their neighbours the information about reachable destination addresses jointly with information about supported types of e2e CoSs and guaranteed QoS level. For each e2e CoS, this information is used in decision process to select such routing path that offers the most attractive QoS guaranties, named QoS path. Finally, the EQ-BGP sets the QoS paths between any pair of source and destination network called end-to-end (e2e) QoS paths that for a given e2e CoS offer the best QoS guaranties. Summarising, the EQ-BGP protocol allows us to create a QoS road map in a multi-domain network with respect to available e2e CoSs.

In Figure 1 we show an example of setting the e2e QoS path using the EQ-BGP protocol. For clarity of presentation, we assume a simple network consisting of tree domains A, B and C that support the same e2e CoS. We assume that all EQ-BGP routers know the values of QoS parameters that are assured both inside particular domains (Q_A , Q_B , Q_C) as well as on the appropriate inter-domain links ($Q_{A \rightarrow B}$, $Q_{B \rightarrow A}$, $Q_{B \rightarrow C}$, $Q_{C \rightarrow B}$). These QoS values are fixed during network provisioning process as a values corresponding to the maximum admissible load allowed by admission control function. Now, let us consider situation when domain C advertises new address, say $NLRI_C$. Then, this information will be propagated towards the domain A through the domain B and appropriate e2e QoS paths will be established as we show in the routing tables of particular EQ-BGP routers. During this process EQ-BGP routers will derives the values of QoS parameters associated with e2e QoS path taking into account the QoS contributions of particular parts of the path corresponding to

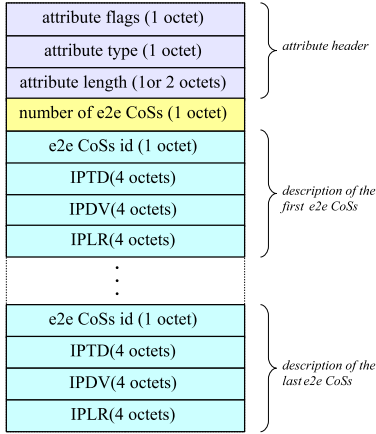


Figure 2. Format of QoS_NLRI path attribute.

particular domains as well as inter-domain links. For example, the domain A sees the e2e QoS path towards destination $NLRI_C$ with QoS level corresponding to aggregated value of QoS related with domain C, interdomain link B-C, domain B and inter-domain link A-B that is equal to $Q_C \oplus Q_{B \rightarrow C} \oplus Q_B \oplus Q_{A \rightarrow B}$, where the operator \oplus denotes an appropriate QoS assembling function.

For enabling the EQ-BGP operations, the four main components need to be added to the BGP-4, that are:

1. **QoS_NLRI attribute** that carries information about QoS capabilities of particular path. This attribute conveys information about the type of e2e CoSs as well as detailed values of QoS parameters. In our proposal we adapt the QoS_NLRI attribute proposed in [5] taking into account requirements of considered e2e CoSs to the structure depicted on Figure 2. This attribute contains header indicating the attribute type, length and flags. The main body specifies the supported e2e CoSs with e2e CoS the codepoint and values of QoS parameters. In case when a given parameter is not used its value is set to -1.
2. **QoS aggregation function** that allows an EQ-BGP router to derive the aggregated values of QoS parameters. Let us recall that considered five e2e CoSs have QoS objectives expressed by mean IP Packet Transfer Delay (mean IPTD), IP Packets Delay Variation (IPDV) defined as a quantile of a packet transfer delay distribution and IP Packets Loss Rate (IPLR). The aggregated mean IPTD value we can calculate as a simple sum of all factors because of the additive nature of expected value of random variables. For IPDV and IPLR parameters we propose to also use a simple sum function that allows us to approximate the aggregated QoS value. Note that whenever, the quantile of IPDV and values of IPLR are lower then 10^{-2} a simple sum

function gives an upper approximation of aggregated values.

3. **QoS decision algorithm** that allows router to select the best QoS path from available alternatives paths. For that purpose, we replace the AS path length criterion used in BGP decision process with a Degree of Preference (DoP) function that is calculated based on the values of particular QoS parameters. We propose to apply DoP function as presented in (1). Note that other function may also be applied.

$$DoP(.) = \sum_{i \in \text{IPTD, IPDV, IPLR}} \frac{f_i}{\max[0, T_i - q_i]} \quad (1)$$

where: i denotes a given QoS parameter from the set corresponding to e2e CoS, T_i is targeted QoS value related with e2e CoS objectives (e.g. as presented in Table 1, q_i is the QoS value offered on the path, and f_i is a QoS preference factor allowing network provider to control routing. In the proposed QoS decision algorithm, the QoS path with the lowest value of DoP is preferred. Moreover, in case when at least one of the QoS parameters of the e2e QoS path, say q_i , will go beyond the target values T_i then such QoS path will be treated as non-conforming with QoS objectives and therefore it will be excluded from this part of decision process.

4. **Multiple Routing Tables**, where each one is used for handling traffic belonging to particular e2e CoS. The separate routing tables are required because e2e CoSs differ in QoS objectives (see Table 1), and therefore it is not possible to satisfy their requirements on a single path.

3. Performance evaluation

In this section we focus on performance evaluation of EQ-BGP protocol. Our objective is to analyse the impact of new EQ-BGP components on the network convergence and the protocol scalability. The evaluation of the protocol will be performed by comparison of EQ-BGP with standard BGP-4 protocol assuming different network scenarios and stressing events such as announcement or withdrawal of a route (network). As the measure of protocols effectiveness we use two metrics that are:

1. **The network convergence time (NCT)** defined as the total amount of time that elapses between the time instant when a stressing event had occurred and the time instant when the processing of last update message caused by this event was finished.

Network type	No. of ASs	BGP-4		EQ-BGP					
				random		increasing		decreasing	
		mean	max	mean	max	mean	max	mean	max
full mesh	4	3.55 ± 0.0	3.55	3.55 ± 0.0	3.55	3.55 ± 0.0	3.55	3.55 ± 0.0	3.55
	11	3.55 ± 0.0	3.55	3.55 ± 0.0	3.55	3.55 ± 0.0	3.55	3.55 ± 0.0	3.55
	29	3.55 ± 0.0	3.55	3.55 ± 0.0	3.55	3.55 ± 0.0	3.55	3.55 ± 0.0	3.55
ring	4	6.5 ± 0.028	8.22	6.78 ± 0.03	8.22	7.64 ± 0.02	8.22	6.79 ± 0.03	8.22
	11	10.84 ± 0.0	10.85	12.94 ± 0.03	14.6	13.32 ± 0.03	14.6	13.49 ± 0.03	14.6
	29	28.2 ± 0.0	28.2	51.45 ± 0.05	53.94	47.26 ± 0.05	50.4	47.79 ± 0.06	50.4
internet	29	11.92 ± 0.03	15.31	17.9 ± 0.07	27.38	13.59 ± 0.06	20.38	18.97 ± 0.07	27.36

Table 2. Network convergence time after advertisement of new route (in msec).

Network type	No. of ASs	BGP-4		EQ-BGP					
				random		increasing		decreasing	
		mean	max	mean	max	mean	max	mean	max
full mesh	4	16.75 ± 0.03	18.47	16.83 ± 0.03	18.47	17.04 ± 0.03	18.47	12.49 ± 0.0	12.49
	11	1854 ± 6.13	2077	717.2 ± 4.21	1139	1275 ± 3.5	1544	576.5 ± 1.7	9.56
ring	4	8.22 ± 0.03	9.56	7.89 ± 0.03	9.56	6.89 ± 0.02	9.56	7.89 ± 0.03	9.56
	11	20.03 ± 0.0	20.03	17.84 ± 0.03	20.03	17.52 ± 0.02	20.03	17.52 ± 0.02	20.03
	29	53.94 ± 0.0	53.94	51.45 ± 0.05	53.94	47.26 ± 0.05	50.4	47.79 ± 0.06	50.4
internet	29	1872 ± 13	2808	1212 ± 65	1706	2654 ± 23	4042	2328 ± 26	5241

Table 3. Network convergence time after withdrawal of existing route (in msec).

2. **The number of advertisement (withdrawal) N_A (N_W)** messages that are exchanged during the network convergence time.

In our experiments we consider three types of networks with different number of ASs, that are: full mesh, ring as well as a representative topology for the Internet. The full mesh topology was selected because it allows for establishing the maximum number of alternative paths and therefore it can be treated as the “worst case”. On the other hand, the ring network (or b-clique) is commonly used for analysing the routing decision algorithm, as there exists exactly two disjointed paths between each pair of domains. Moreover, these two topologies constitutes the basic “building blocks” for other networks. To complete the performance evaluation, we consider also one topology derived from routers operating in the Internet backbone as presented in [8].

For the sake of simplicity, we assume that each AS is represented by a single EQ-BGP router that is connected with its neighbours using links of 1Mbps capacity that introduces constant delay of 1msec. Even the parameters of links were chosen arbitrary, the obtained results corresponding to NCT can be easily scaled taking into account actual link characteristics. In addition, we assume that all ASs support a single e2e CoS but differ in guaranteed value of meanIPTD. The values of IPDV, IPLR are the same for all domains. Moreover, we consider three different strategies for assignment of values of meanPTD to particular domains that are: (1)

random strategy assuming that the value of delay offered by a given AS is randomly chosen from 1msec to 10msec, (2) increasing strategy assuming that the AS with greater AS number introduces larger delay and (3) decreasing strategy assuming that AS with lower AS number introduces larger delay. Our experiments were performed using NS2 simulator [7], which was enhanced by the model of EQ-BGP protocol. All experiments were performed assuming that advertisement or withdrawal of a route, would occurred when the network was in a stable state. Each simulation run was stopped when the service of the last update message originated by considered stressing event was finished. The presented results were collected from 10 simulation runs, where a randomly chosen AS in a randomly chosen time instant advertised or withdraw the route. The reported values of convergence time includes also the 95% confidence interval.

Below, in section 3.1 and 3.2, we present results corresponding to network convergence time and assessment of protocol scalability.

3.1. Network convergence

In Table 2 we present results corresponding to the convergence time of “full mesh”, “ring” as well as “internet” networks after advertisement of a new route. The obtained results say that the full mesh network converges almost during the same time for both EQ-BGP and BGP-4 protocols,

independently of the number of ASs. This can be explained by the fact all ASs fixed routing paths on the direct links and then establish the alternative paths. As the QoS level assured on direct links is better than offered on alternative paths, therefore none of the routers would change to alternative. So, the routing process will be finished in the same time for all analysed cases. On the other hand, in case of "ring" and "internet" networks, the NCT time increases with the number of ASs as well as the number of interdomain links (see the case of ring29 and internet29 networks). Moreover, we can observe that EQ-BGP protocol needs more time to converge that in some cases is even two times greater. This effect is caused by introduction of an additional "degree of freedom" comparing to BGP-4 that comes from possibilities for assigning arbitrary values of QoS in particular ASs instead of a unit value in case of AS path length. Therefore, the chance that an alternative path is better than currently used is greater. As a consequence, the network needs more time to converge. The complementary effect, we can observe in case of route withdrawal as presented in Table 3. For all types of network, EQ-BGP converges a bit faster comparing to BGP-4. This effect comes from the fact that alternative paths that have assigned more information about its preferences and less suitable paths are removed faster. However, we stress that this reduction of convergence time is not significant. Summarising the presented results points out that EQ-BGP protocol gives a stable routing and the network convergence time is at the similar level as in case of BGP-4.

3.2. Scalability

The EQ-BGP protocol is designed to operate in a large network. Therefore, assessment of its scalability is an important part of the performance evaluation. For this evaluation, we compare the number of update messages proceeded by both EQ-BGP and BGP-4 protocols. From results presented on Figure 3 we can observe that EQ-BGP and BGP-4 need similar numbers of messages to converge. However, EQ-BGP in the case of the ring network needs a bit more messages than BGP-4, while in case of a full mesh network a little less. Therefore, we can conclude that EQ-BGP scales at the similar level as BGP-4.

4. Summary and further work

In this paper we provide a description of the EQ-BGP protocol designed for performing an inter-domain QoS routing. The EQ-BGP extends the standard BGP-4 protocol with a modified routing decision algorithm, and a QoS-aware attribute. The obtained simulation results focused on convergence as well as protocol scalability confirmed that EQ-BGP produces stable routing and its performance is at the similar

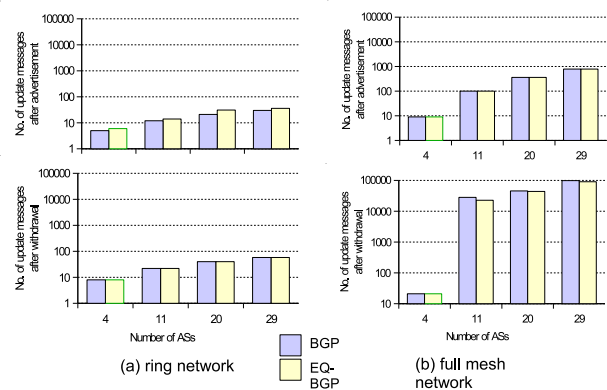


Figure 3. Scalability of EQ-BGP vs. BGP-4

level as the BGP-4. The further work we will focus on analysing alternative decision algorithms as well as evaluation of other stressing events like link or node failure, route flapping, etc.

References

- [1] B. Abarbanel. BGP-4 support for Traffic Engineering. Technical Report draft-abarbanel-idr-bgp4-te, IETF, 2000.
- [2] G. Apostolopoulos, D. Williams, S. Kamat, R. Guerin, A. Orda, and T. Przygienda. QoS Routing Mechanisms and OSPF Extensions. Technical Report RFC 2676, IETF, August 1999.
- [3] J. Babiarz, K. Chan, and F. Baker. Configuration guidelines for DiffServ service classes. Technical Report draft-ietf-tsvwg-diffserv-service-classes-00, IETF, February 2005.
- [4] W. Burakowski, M. Dabrowski, and M. Potts. Euqos classes of service. In "First European Workshop on End-to-end QoS in the Internet (EEQoS'05)", 2005.
- [5] G. Cristallo and C. Jacquenet. The BGP QOS_NLRI Attribute. Technical Report draft-jacquenet-bgp-qos-nlri-05, IETF, 2004.
- [6] IST-EUQOS. End-to-end Quality of Service support over heterogeneous networks. Technical Report 004503, IST, 2004.
- [7] The Network Simulator - ns-2. <http://www.isi.edu/nsnam/ns/>, 2005.
- [8] B. Premore. Multi-AS topologies from BGP routing tables. <http://www.ssfnet.org/Exchange/gallery/asgraph/index.html>, 2005.
- [9] I.-T. Recommendation. Network performance objectives for IP-based services. Technical Report Y.1541, ITU-T, May 2002.