Policy Based QoS support using BGP Routing

Priyadarsi Nanda and Andrew James Simmonds
Department of Computer Systems
Faculty of Information Technology
University of Technology, Sydney
Broadway, NSW
Australia

Abstract - Routing protocols are important to exchange routing information between neighboring routers. Such information is used to update routing tables and to share information about status of the network so that traffics to appropriate destinations will be fast and efficient. Different types of routing protocols are in widespread use across the Internet. Apart from determining optimal routing paths and carrying traffics through the networks, these routing protocols should have additional functionalities to support network policies, traffic engineering and security. In this paper we discuss the use of one such routing protocol, the Border Gateway Protocol (BGP) which is the industry standard. We also present an algorithm in which each Autonomous System (AS) decides how to forward its traffic satisfying end-to-end- QoS for its users and services. Our proposed algorithm is dynamic in that network status and route advertisements, which change with time and based on traffic loads in the network, are monitored and taken as input to the final decision on traffic forwarding between ASs.

Key words: BGP, QoS, Autonomous System (AS)

1.0 Introduction

Current Internet architecture is based on the Best Effort (BE) model, where packets can be dropped indiscriminately in the event of congestion. Such architecture attempts to deliver all traffic as soon as possible within the limits of its abilities, but without any guarantee about throughput, delay, packet loss, etc. Though such a model works well for certain traditional applications such as FTP, E-mail and less QoS constrained applications, it can be intolerable for newly emerged real-time, multimedia applications such as Internet Telephony (VoIP), Video-Conferencing and Video on-Demand, as well as future services. Hence, with massive deployment of Internet based applications in recent years and the need to manage them efficiently, current Internet structure needs a major shift from the BE model to a service oriented model with support for desired QoS. Current research in this direction is focused towards providing better than BE service over the Internet through a new architecture. Also the new architecture should be both scalable and guarantee end-to-end QoS for different services/applications while supporting different levels of performance.

Current Internet architecture lacks standardization while deployed across various domains, hence affecting end-to-end QoS significantly. In this paper our effort is to find a scalable and uniform solution mainly addressing routing and its effect on end to end QoS. In this regard, we consider current inter-domain routing based on BGP as the central component and develop an algorithm allowing QoS domains to be easily identified and enable policy based routing to support QoS for various applications.

One of the main objectives in setting up an end-to-end path for any service over the Internet is providing support for its service requirements to achieve necessary QoS, and such tasks are difficult to achieve through current Internet architecture. In this regard, our algorithm is designed to address such heterogeneous service parameter requirements for different services between ASs, and tries to find a viable solution by integrating network policies with routing and traffic engineering objectives. We mainly focus on Inter-domain traffic engineering issues in resolving the policy requirements of different services. In doing so, we have identified and addressed two core problems in the Internet today in relation to QoS:

- How do neighboring ASs learn about whether, how, and when their services can be satisfied for QoS by other ASs?
- How can we make sure that all the participating ASs in the end-to-end path of the flow can act upon same policy binding rule as close as possible?

This paper is organized into the following sections. Section 2 gives a general view on BGP based Internet routing and its support for network level policies. In Section 3 we have explored the possibilities for supporting a centralized routing decision within a network domain while still using a distributed model for Internet wide routing, and how such models can be used to satisfy QoS over the Internet. Section 4 describes our algorithm based on which BGP routing decisions are influenced between neighboring domains in order to achieve scalability in supporting end to end service guarantees for various applications. Section 5 concludes our paper with a proposal for future work.

2.0 BGP and policy based routing

BGP uses autonomous system path information between neighboring routers in different domains to inform network reachability and hence is a path vector protocol. Such network reachability information includes list of ASs through which end nodes can communicate. One important feature of BGP is its routing decisions are influenced by traffic flow policies within ASs. I.e. individual AS can implement network policies to determine whether to carry traffics from different users (mostly users from other AS) with diverse QoS requirements. Such network policies are not part of BGP, but provide various criteria for best route selection when multiple alternative routes exist and help to control redistribution of routing information resulting in a rich support of BGP for traffic engineering in the Internet. One of the key features of BGP is the decision process [3] through which each BGP router determines the path to destination prefixes. The process in brief is as below, where only if there is a tie will the next stage to be considered:

1. Find the path with highest Local-Preference
2. Compare AS-path length and choose the one with least length
3. Look for the path with Lowest Multi-Exit-Discriminator (MED) attribute
4. Prefer e-BGP (exterior) learned routes over i-BGP (interior) routes
5. Choose the path with lowest IGP metric to next hop

Because of this, current Internet Traffic Engineering depends heavily on both Intra and Inter Domain routing protocols (Interior BGP within the domain and Exterior BGP between the domains), using network policy in order to configure the routers across various domains. The support for policy based routing using BGP can provide source based transit provider selection, whereby ISPs and other ASs can route traffic originating from different sets of users through different connections across the policy routers. Also QoS support for Diffserv networks can be supported using policy routing through the use of the TOS field in IP packets. Hence a combination of traffic engineering for load balancing across network links offered by destination based routing and policy based routing can enable implementation of policies that distribute traffic among multiple paths based on traffic characteristics.

Because BGP can be used to enforce network level policies involving routing decisions by different ASs, a minor error in configuring those parameters may result in disastrous consequences affecting end-to-end flow properties for various applications. While the objectives of Internet connectivity are to provide some sort of universal interconnections amongst ASs (ISPs etc), these ASs sometimes behave as fierce competitors [2]. Individual ASs always aim to maximize their profit by minimizing service guarantees to their customers, whereas customers aim for maximized service guarantee at a minimize price. This gives rise to service differentiations and different pricing hierarchies between different providers. There is a wide community of researchers investigating traffic engineering parameter tuning in the last few years [3,4,5].

While such schemes in the Internet always create conflicts amongst the service providers, there is no concrete solution at present to our knowledge. This is not only due to lack of proper routing strategies and service policies amongst the ASs at the business level, but also due to the fact that the current Internet architecture has no central control. The issue of establishing proper AS level relationships, their connectivity based on both network and device level policies, and the need for a better Internet architecture are some of the major areas of research in recent times. The objective of our research on the above

mentioned issues is to provide robustness, scalability and better management of the current Internet in the face of heterogeneous policies, application growth and both management and technical challenges.

Policy based routing may be applied centrally within an AS while policy issues between ASs should be distributed and organized hierarchically in order to achieve scalability over the Internet. Such schemes are discussed in detail in the following section of the paper.

3.0 Network policy control

Current Internet needs a centralized (within each AS) and coordinated (between the ASs) approach for network policy management in order to facilitate better support for policy based routing and maximize benefits for individual ASs. We believe such approaches will overcome various problems related to uniform policy management for route selection and traffic engineering over the Internet. In this paper we first describe a centralized approach using BGP within the AS and then make use of the BGP decision process along with our algorithm to implement policy based routing over the Internet.

3.1 Centralized Policy control

Information relating to link status, neighboring routers, their scheduling policies and service support, plays an important role determining the traffic path for flows both within and between the ASs. Link status describes the status of the links connecting the routers and such information can be obtained by running link monitoring agents at each router within the AS. Neighboring router information can be obtained through the IBGP path attributes for specific paths. Each router on activation sends their routing policies to a centralized server to help determine a decision on a route. The centralized server is similar to a Bandwidth Broker (BB) [8], where routing policies for both incoming and outgoing traffic are stored in order to facilitate routing decision within an AS. Once the database of the server is populated with all such information, it can create a set of QoS paths between the ingress-egress pairs within the AS. Such QoS paths may each have a status field indicating current status and can be used to select optimized routing.

In a policy based network, one of the important objectives is to have a network wide view for the QoS flows i.e. which QoS path to select between end nodes and then determine an end to end policy frame work, based on which traffic may be forwarded between different domains. Also ASs can influence each other's routing decision through the use of BGP attributes such as MED and AS path pre-pending [1,3,6], mostly through a trial and error basis. Such methods can only work in the short term at the expense of long term problems such as path instabilities. But if the ASs can negotiate on some of their desired objectives to route their traffic before taking any further decision on a changed routing strategy, a much better result can be achieved.

Most ASs use default routing strategy e.g. Early-Exit policy [10] where the exit point selection is based on the point closest to the source. Through proper co-ordination and understanding between the ASs (without revealing their internal topology), a better solution can be achieved for the exit point that may minimize the total overhead (delay, cost) for each flow across the ASs and achieve end to end QoS. Such an approach using negotiation based Internet routing is given in [9] where several experiments between ISP pairs are conducted to achieve a case which is the social optimum. An anti-social and rigid ISP, who never shares its policy (both administrative and routing) with others, will be the one to fall behind others in maximizing its own revenue. In the next section we first discuss the issue of policy based networking on two different levels: Network level and AS level, and then define the relationship between them in terms of policy requirements to combine both routing and traffic engineering in an integrated approach.

3.2 Network level /AS view for policy control

Based on the present structure of the Internet, with multiple ASs and sources of traffic, our first classification for traffic type can be grouped under either local or transit type. Local traffic either originates or terminates in the domain under consideration while transit traffics are outside any specific domain. Based on such high level classification, local traffic always involves a cost function which it pays either directly or indirectly to its service provider, and transit traffic does not have a cost function involved with it unless there is a peering relationship. Also, present Internet is a connection of stub, multi-homed and transit ASs [7]. But while using Inter-domain routing protocols to receive path related information from other ASs,

both stub and multi-homed AS may not necessarily advertise their own network structure outside their own domain. From here we go further and propose that, all the customer domains which are stub AS are excluded from announcing their own traffic policies outside their domain. BGP is policy based and one of the primary goals for using BGP should be aimed reducing transit traffic between domains in the Internet through policy based routing.

We have addressed both dynamic load balancing capabilities based on destination prefixes and policy based routing to achieve optimization within the network where traffic distribution policies relate to such traffic characteristics among multiple paths. In order to clearly manage traffic flows based on QoS objectives associated with individual traffic flow across the Internet, it is important to understand the standard frame-work on which our proposed scheme is built. Consider a path level abstraction for traffic flows across networks where a traffic flow may be considered to be forwarded through the routers as shown below:

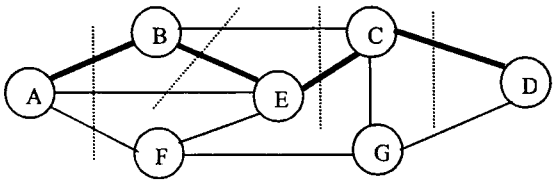

Fig.1. Network view of Path selection and Routing

In the figure above, the bold line represents the path (A-B-E-C-D) selected for a flow between A and D. Such a path set up between the routers situated in a single AS may be viewed as a directed graph for the flow under consideration and can be established through routing algorithms at different routers. A routing protocol such as IGP or IS-IS may consider the traffic engineering policies for individual links before choosing the next-hop path across the network. Hence the network level policy may include parameters such as bandwidth constraints over each link, cost for each link, load balancing feature, and routing policy attributes.

We have considered a simplistic approach for gathering as much information as we can through the use of a centralized server within the network. Assuming a single traffic flow entering the ingress at A, router A looks for the destination prefix carried within the packet. We assume that the flow is authorized and authenticated by a central server such as policy server (or BB), and authorization for network resources has been issued by it to router A. After determining the next hop in the path towards the egress router (in our case it is router D) router A now contacts the route module of the policy server within the domain. If the path selected by router A already exists within the server's database, then the server compares the performance metrics (such as delay) with other similar paths if present and determine the best path for the flow. The server also extracts all the link characteristics for that path and sends back the result to the ingress router A. Such co-ordination among the routers and central policy server within an AS is easy to manage when working through individual routing policy. However such strategy may not work well when applied between ASs in the Internet. Hence In order to explore the issue of policy based routing in the Internet the problem can be stated as following: Whether the ASs situated in the path of traffic flow could apply the same policy across their domains? Looking at the above mentioned problems of Internet routing, the routing component should have mechanisms for computing paths through the Internet that honor performance and resource utilization policies and forwarding mechanisms for sending traffic over the computed paths. Section 4 presents our algorithm based on which we may achieve scalability in policy based routing when deployed across the Internet.

4.0 Inter-domain policy management and policy based routing

Our proposed scheme is designed to address diverse policy requirements between ASs and works towards a solution by integrating policies with routing, i.e. using policy based routing, but combined with Inter-domain traffic engineering. In doing so, we have identified and addressed two core procedures in relation to QoS:

1. Each AS should investigate its own policy requirement for the services it is authorized to carry based on various statistics on the flows both entering and existing its domain. Then based on these policies, devices at network level are configured within the AS to guarantee network wide QoS.
2. Eeach AS should interact with its neighbor to discover what resources their neighbors require in order to closely satisfy their QoS requirements. Such a scheme may be implemented through resource discovery and signaling schemes similar to the one proposed in [11].

Policy co-ordination using policy based routing between ASs may play an important role while deciding the service qualities for both existing and future applications. In a scenario consisting of multiple ASs between end systems, it is important to discover the policies that each domain may support and then communicate (policy propagation) them across a chain of AS domains. Policy propagation can be thought similar to BGP based path advertisements between neighboring AS domains i.e. each AS may be configured to announce to its neighbor AS about costs to the destination AS that it knows about. In this scenario, instead of announcing the cost, AS announces policies of other ASs it knows between itself and the destination ASs (which can be done by using one of the BGP community attributes). In case multiple path announcements are received by an AS for the same destination prefix, we must have a mechanism in place which can determine the best path supporting flow policies. Hence end to end QoS depends on adopting such optimal policy decisions which we describe through an example below.

4.1 Optimal policy enforcement from multiple policy announcements

We describe the mechanism based on which optimal policy can be enforced between neighbors with the same destination pre-fix. Such a situation occurs with a large number of ASs having multiple providers connected to them. Consider the figure below where multiple announcements are received for the same destination by other ASs and set of policies (P1 – P4) assigned for different traffic:

Policy	Price	Resource (BW limit)	Feature
P1	X1	Y1	Gold
P2	X2	Y2	Silver
P3	X3	Y3	Bronze
P4	X4	Y4	Best Effort

After defining the above mentioned policies for different services, each AS announces a set of policies it can support with a tag of pricing, available resources and service features to its neighboring ASs. The table above gives an example which shows information when a neighboring AS receives the advertisements:

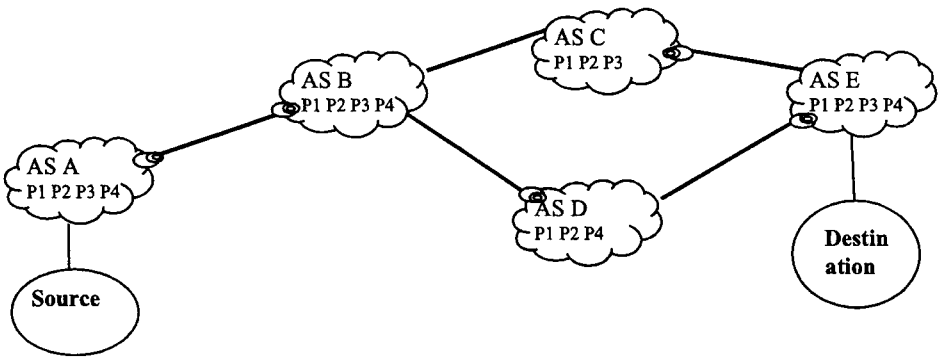

Fig.2. Optimal Policy enforcement by AS

Based on the information, priority and service quality a simple policy rule can be stated for price as:

X1>X2>X3>X4

In this case, both AS C and AS D receive policy advertisements from AS E which announces: AS E can support all the policies (P1 P2 P3 P4) for traffic with the destination prefix.

Now, after receiving the policy advertisements from AS E, both AS C and AS D prepare their policy advertisement lists and send policy advertisement to AS B. AS B receives two different policy announcements to the same destination creates the following policy vectors represented by:

AS B (P1 P2 P3 P4) AS C (P1 P2 P3) AS E (P1 P2 P3 P4) and

AS B (P1 P2 P3 P4) AS D (P1 P2 P4) AS E (P1 P2 P3 P4)

Then, AS B may perform either of the following:

- Evaluate on the basis of Capacity, Pricing, Performance and, most important, its own benefit and accept the best policy announcement rejecting other announcements.
- Accept all the announcements from down stream neighbors and announce all of them to its up-stream neighbors. Then the upstream neighbor can decide whether to accept them or not based upon its own preference and use.

The first approach is more similar to BGP decision making principle between ASs for selecting the best possible path to any prefix. In our proposed scheme the tie-breaking rules as described in section 2 are used for path advertisements between ASs. But in addition, we have added the two rules a. and b. below to support policy announcements and policy based decision for AS route selection process where ASs announce routes along with the types of services they can offer on these paths. Hence each AS after receiving the path announcements from their neighbors checks for which policies (related to the service offerings, e.g. Gold, Silver...) can be supported on those paths. Then they use an algorithm to determine the best path based upon maximum number of services they can support. Hence, before accepting any policies from neighbors, an AS must check the following:

a. A similar set of policies are supported by the AS itself within its domain
b. The policy and path advertisements contribute towards optimal resource allocation and routing information

In any case, it is essential to hold and monitor policies by individual ASs before announcing them to their neighbors. For this each AS should have a centralized policy repository directly interacting with the routing process within it.

4.2 Policy Co-ordination between multiple domains

In our proposed scheme, each AS maintains its own policy repository representing the set of policies and their characteristics impacting the nature of the services it is going to carry over its links. Because ASs act on independent policies aimed at maximizing their own profits, there may be certain differences among these policy sets based on what factors they have been developed. Hence these ASs may categorize such policies based on the service importance. Such may consider the following example.

P1: Traffic flows requiring $\leq$5 ms delay and $\leq$1 ms jitter

P2: Traffic flows requiring $\leq$ 5 ms delay and $\leq$5 ms jitter

Now from the above two policies, assume P1 has a higher priority than P2 when considered in AS A while the reverse applies in AS B. Hence such decisions must be worked out within individual ASs. It is also not possible to have uniform policy sets right across these ASs which would be needed to have single tier pricing architecture across the Internet. Hence, based on the policy profiles defined by individual AS, we need to have policy co-ordination (policy conflict resolution) amongst ASs and for this reason we have developed a simple policy co-ordination mechanism, based on which inter AS routing may take place.

Such policy information is important to set up SLAs between the customer and service providers. We have mainly considered four QoS parameters (Priority, Loss, Delay and Jitter) based on which policies have been defined and the necessary entries made in the policy repository. In our central policy repository module, we have maintained two different groups of policies: one holding the actual policy set of the AS while the other one holds the negotiated policy set between the ASs. This latter group is the optimized policy which is obtained after the algorithm is run by the AS. The optimized policy list is usually constructed from the policy announcements by neighboring ASs by running a decision algorithm to determine the best possible AS in the neighbor which supports those policies.

Policy decision Algorithm

1. *Get the list of policies from neighbor announced with BGP Path advertisements*
2. *Compare the list of policy support with own policy list*
3. *If matched, check individual policy:*
   Check all parameters to match
   i. *If matched, list them in the E-E list*
   ii. *Else, tag the policy as non-confirmed and store them in temp list*
   iii. *After all policies checked, adjust tagged policies by assigning a different policy in consultation with neighbor and enter them in E-E list*
4. *If all policies in the list do not match, then find those policies which are supported by AS domain and its neighbor. These lists of policies are the ones going to be in the E-E list and then further advertised to their neighbors.*
5. *Terminate the algorithm if none of the policies are supported*

The algorithm works through a relative standard rather than an absolute one where in case of policy mismatch, no further actions are taken. Also based on amount of traffic loading between different ASs such mechanism may improve resource utilization and achieve minimum supported QoS for end to end traffic flows corresponding to various applications over the Internet.

5.0 Conclusion and future work

One of the main objectives in setting up of an end-to-end path for any service over the Internet must be to satisfy its service requirements to achieve the necessary QoS. In this paper we describe how routing policies in the Internet based on BGP may affect the QoS for various applications. Also we discussed how policy co-ordination between ASs can play an important role in order to support the service qualities for both existing and future applications. Based on this we proposed an algorithm to resolve policy conflicts between ASs in the Internet. We believe such an approach can contribute significantly towards deploying policy based Internet architecture. Our future work in this area of research will be to explore validity of our proposed scheme through a prototype using OPNET simulator and investigating further issues covering fault management, policy based overlay routing and security related to information sharing between ASs.

5.0 References

1. N.Feamster, J.Winick, J.Rexford; A model of BGP routing for Network Engineering, SIGMETRICS/Performance'04, June 12 – 16, 2004, New York, USA
2. D.Clark, J.Wroclawski, K R.Sollins, R.Braden; Tussle in cyberspace: Defining tomorrow's Internet, SIGCOMM'02, August 19-23, 2002, Pittsburgh, Pennsylvania, USA
3. B. Quoitin, C. Pelsser, O.Bonaventure, S.Uhlig ; A Performance evaluation of BGP based traffic Engineering
4. R.Yavatkar, D.Pendarakis, R.Guerin; A frame work for policy based admission control, RFC 2753, January 2000
5. S.Salsano; COPS usage for Diffserv resource allocation (COPS-DRA), Internet Draft, October 2001
6. S.Uhlig, O.Bonaventure, B.Quoitin; Interdomain traffic engineering with minimal BGP configurations, 18th International Teletraffic Congress, 31 August - 5th September 2003, Berlin, Germany
7. G.Di Battista, M.Patrignani and M.Pizzonia; Computing types of relationships between autonomous systems, IEEE INFOCOM 2003
8. K.Nichols, V.Jacobson and L.Zhang; A two-bit differentiated services architecture for the Internet, RFC 2638, July 1999
9. R.Mahajan, D.Wetherall and T.Anderson; Interdomain routing with negotiation, Technical Report, CSE 04-06-02, University of Washington, May 2004
10. N.Spring, R.Mahajan and T.Anderson; Quantifying the causes of path inflation, ACM SIGCOMM, August 2003
11. M.Li, D.B.Hoang and A.J.Simmonds: 'Fair intelligent admission control over resource-feedback DiffServ network', *Computer Communications*, Vol. 28, No. 15, pp. 1770-1777 September 2005.