QoS-based Routing and IP Multicasting: A Framework

Gábor Rétvári, Róbert Szabó
{retvari, szabor}@ttt-atm.ttt.bme.hu
Tel: +36-1-4633110
Fax: +36-1-4631763
High Speed Network Laboratory
Department of Telecommunications and Telematics
Technical University of Budapest
H-1111 Budapest, Stoczek u. 2 - HUNGARY

Abstract

The Internet Engineering Task Force (IETF) has initiated the Integrated Services (IS) Internet framework to elaborate a new service, which supports multipoint-to-multipoint communication and QoS-based routing in order to serve the demands of multimedia applications running at Internet hosts. To achieve the goals of the IS Internet two new protocols have to be introduced over the existing Internet Protocol (IP) platform: first, due to the connectionless nature of IP a resource reservation and management mechanism has to be invoked in order to satisfy the QoS requirements of the IP flows consequently, second, a multicast capable routing protocol is desirable, which takes into account the amount of available QoS resources at network links when trying to find the best path to accommodate an IP flow. In this paper we investigate the question of QoS-based routing and IP multicasting in an OSPF and RSVP protocol environment IS model.

Keywords: Multicast, QoS Routing, RSVP

1. INTRODUCTION

The main goal of this paper is to give a brief tutorial on the current developments aiming to combine multicast capability and QoS sensitivity in intra-domain IP routing, and to attempt to outline the basic features of a QoS/multicast routing protocol by rising and answering the most generic questions.

QoS demands of multicast application include tolerance to jitter, delay and lost packets. In order for a network to satisfy the emerging QoS requirements applications must be able to specify their traffic's characteristics. This implies that finding the shortest path to a destination when constructing a data delivery tree is no longer sufficient; the selected route must satisfy application QoS demands as well. This adds considerable complexity to the routing task. The general metric used by best-effort routing algorithms is the count of the number of hops (or a special length metric assigned to the hops) along the path from a source to a destination. In QoS environment during path calculation, parameters such as delay, jitter, bandwidth, loss probability and cost must also be considered. Since path optimization with more than two constraints is an NP-complete problem (apart from a few very special cases), a possible approach is to take solely the most important QoS parameter into account: the amount of available bandwidth over the network links.

While resource reservations can be achieved by using the existing RSVP (Resource reSerVation Protocol, [1]), there exists no QoS sensitive, multicast capable IP routing protocol. This paper briefly outlines the restrictions, tradeoffs, constraints and contradictions that arise when attempting to combine multicast capability and QoS sensitivity. These problems are discussed in terms of real examples: how two extensions of OSPF [2], QoS sensitive QOSPF [3, 4] and multicast capable MOSPF [5, 6] work. We will reveal that MOSPF and QOSPF can not be easily 'merged' to support the demands of QoS-multicast flows. Thus a new routing protocol must be defined, which handles these demands while also supporting unicast QoS-based routing.

There are quite a lot of emerging difficulties when trying to define a multicast capable and QoS sensitive intra-domain IP routing protocol.

Firstly, questions directly related to the multicast QoS routing algorithm have to be answered. This algorithm is responsible for the establishment of data delivery path by selecting those links in the network that advertised enough free capacity to accommodate the IP flow in question. The question related to whether construct a pre-computed routing table or perform on-demand routing must be decided. Can IP's traditional next-hop routing conception satisfy the demands of multimedia applications, or it is necessary to adopt the PNNI approach of ATM i.e. to use source routing. Nevertheless proper modifications have to be carried out on known shortest-path-first algorithms (Dijkstra, Bellman-Ford) to support multicast and QoS.

Even more problems arise when attempting to co-operate with RSVP. One of the main problems is how to provide each "reservation style" defined in RSVP.

QoS routing requires route-pinning. This implies, that the existing QoS route will not be replaced unless it becomes unavailable (due to topology change directly affecting the used path) in order to avoid routing transients and route oscillations. Upon the change of a multicast group membership the routing table has to be re-calculated in a way ensuring that the QoS routes to the residual destinations remain the same (i.e., they remain "pinned").

Moreover, a special registry entry (the so called "flow state") has to be maintained at routers along the multicast tree, in order to accurately deliver incoming IP datagrams. The flow state contains the set of next-hops for the datagrams belonging to a particular stream. The method of building up and maintaining these states must be elaborated as well.

The rest of this document is structured as follows. In Section 2 a short overview is given of the QoS-based routing framework. Section 3 introduces IP multicasting. Section 4 and 5 describes RSVP and OSPF, respectively. In Section 6 we identify some problems of crucial importance concerning QoS/multicast routing in more detail. After analyzing these issues in Section 7 we draw conclusions.

2. QOS-BASED ROUTING IN THE INTERNET

The IETF has elaborated the "Framework for QoS-based Routing in the Internet" project [7] to call the fundamentals of the Integrated Services Internet into existence on the basis of the arguments discussed previously: a routing scheme, where a resource management mechanism collaborates with some enhanced capability routing processes.

There is a fundamental interaction between resource reservation set up and routing, since reservation requires the installation of flow state along the route of data packets. If a route changes, there must be a mechanism to set up a reservation along the new route. The document [7] proposes a framework for QoS-based routing, with the objective of fostering the development of an Internet-wide solution while encouraging innovations in solving the many problems that arise.

This approach follows the traditional separation between intra- and inter-domain routing. It allows solutions like QOSPF [3, 4] to be deployed for intra-domain routing without any restriction, other than their ability to interact with a common, and perhaps simple, inter-domain routing protocol. This routing model is the most practical one from the evolution point of view. It is superfluous to say that the eventual success of a QoS-based Internet routing architecture would depend on the ease of evolution.

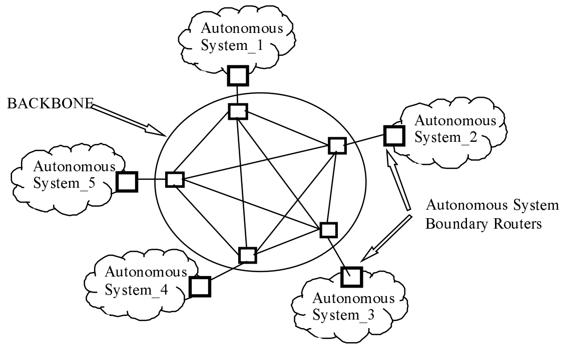


Figure 1. Illustration of the two-pronged approach initiated by IETF.

Figure 1. illustrates the separation of exterior and interior gateway protocols, as proposed by IETF. Separate Autonomous Systems can utilize independent intra-domain (IGP) routing protocols (like QOSPF). The Internet Backbone employs a universal inter-domain (EGP) routing protocol (like QoSMIC [8]). Autonomous System Boundary Routers (ASBRs) interacts between the separate routing domains and the backbone.

To conclude it seems to be the most obvious way of deploying IS services over the entire Internet to develop and test multiple, independent multicast capable, QoS sensitive intra-domain IP routing protocols with the ability of communicating with each other through a well-defined interface of an Exterior Gateway Protocol. The design of such an intra-domain routing protocol is aimed by our efforts.

3. IP MULTICAST ROUTING REVISITED

There are three essential types of IPv4 addresses: unicast, broadcast, and multicast [9, 10]. A unicast address is designed to transmit a packet to a single destination. A broadcast address is used to send a datagram to an entire subnetwork. A multicast address is designed to enable the delivery of datagrams to a set of hosts that have been configured as members of a multicast group in various scattered subnetworks.

IP multicasting is not connection oriented. Considering a standard IP infrastructure, a multicast datagram is delivered to destination group members with the same best-effort reliability as a standard unicast IP datagram. This means that a multicast datagram is not guaranteed to reach all members of the group, or arrive in the same order relative to the transmission of other packets.

Further, individual hosts are free to join or leave a multicast group at any time. There are no restrictions on the physical location or the number of members in a multicast group. A host may be a member of more than one multicast group at any given time and does not have to belong to a group to send messages to members of a group. Multicast routers execute a multicast routing protocol to define delivery paths that enable the forwarding of multicast datagrams across an internetwork.

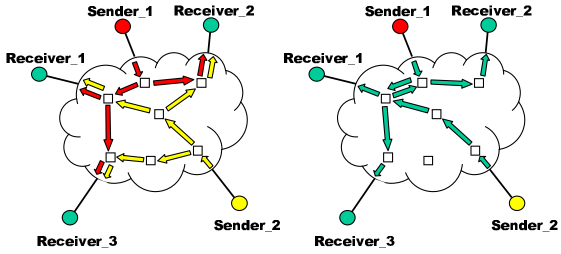
Today, the majority of Internet applications rely on point-to-point transmission. The utilization of point-to-multipoint transmission has traditionally been limited to local area network applications. Over the past few years the Internet has seen a rise in the number of new multimedia applications that rely on multicast transmission. Multicast IP conserves bandwidth by forcing the network to do packet replication only when necessary, and offers an attractive alternative to unicast transmission for the delivery of network multiparty videoconferencing, and shared whiteboard applications (among others). It is important to note that applications for IP multicast are not solely limited to the Internet. Multicast IP can also play an important role in large distributed commercial networks.

3.1 Multicast tree establishment and management

Next, the two fundamental approaches of multicast tree establishment and management are going to be reviewed. The two basic conceptions are: “per-source trees”, and “shared trees”.

In case of arranging a multicast communication on a per-source-tree basis (in much of the literature, it is also called source-based tree) the multipoint-to-multipoint connection can be divided into separate point-to-multipoint connections. Thus, there is a single point-to-multipoint data forwarding tree from each source, as it is shown in Figure 2a. Each tree is managed independently from the others by its source. This approach results in high level of load balancing and optimal shortest path routing.

The so-called shared tree is a special spanning tree of all senders and receivers of a multicast session (see Figure 2b.). Data communication takes place within the spanning tree, data are forwarded from senders to receivers over the links which are branches of the tree. The shared tree is built in an incremental way: an endpoint willing to join to the multicast communication originates a “Join” message in the direction of a known point (the so called “core”) of the tree. The new branch added to the shared tree will fork at the point where the Join message hits the tree. Since there is only one spanning tree per a multicast group the shared tree approach implies less resource consumption, on the other hand it causes grave traffic concentration and sub-optimal routing (packets do not travel the shortest path from senders to receivers).



a.) Per-source multicast forwarding tree

b.) Shared tree

Figure 2: Illustration of per-source and shared multicast data forwarding trees

4. RESOURCE RESERVATION PROTOCOL (RSVP)

RSVP is a resource reservation setup and management protocol, which is designed for the Integrated Services Internet model [10]. The RSVP protocol is used by a host to request a specific quality of service from the network for particular application data streams or flows. RSVP is also used by routers to deliver QoS requests to all nodes along the paths of the flows and to establish and maintain state to provide the requested service. RSVP requests will generally result in resources being reserved in each node along the data path.

The RSVP protocol is designed to co-operate with any current and future unicast and multicast routing protocols without any restrictions i.e. it is a routing independent resource management signaling protocol. RSVP requests resources for simplex flows; i.e. it requests resources in only one direction. Therefore, RSVP treats a sender as logically distinct from a receiver, although the same application process may act as both a sender and a receiver at the same time.

In order to efficiently accommodate large and dynamic group membership, and heterogeneous receiver requirements, RSVP makes receivers responsible for requesting a specific QoS. A QoS request from a receiver host application is passed to the local RSVP process. The RSVP protocol then carries the request to all the nodes (routers and hosts) along the reverse data path(s) to the data source(s), but only as far as the router where the receiver's data path joins the multicast distribution tree. As a result, RSVP's reservation overhead is in general logarithmic rather than linear in the number of receivers.

Since the membership of a large multicast group and the resulting multicast tree topology are likely to change with time, the RSVP design assumes that RSVP and traffic control states are created and destroyed incrementally in routers and hosts. For this purpose, RSVP establishes "soft state"; that is, RSVP sends periodic refresh messages to maintain the state along the reserved path. In the absence of refresh messages, the state automatically times out and is deleted.

4.1 Reservation Models

An elementary RSVP reservation request specifies the desired QoS parameters and the set of data packets − the "flow" − to receive the required QoS.

RSVP messages carrying reservation requests originate at receivers and are passed upstream towards the senders. The basic RSVP reservation model is "one pass": a receiver sends a reservation request upstream, and each node in the path either accepts or rejects the request. This scheme provides no easy way for a receiver to find out the resulting end-to-end service. Therefore, RSVP supports an enhancement to one-pass service by gathering information that may be used to predict the end-to-end QoS (OPWA). The results are delivered by RSVP to the receiver hosts and perhaps to the receiver applications. The advertisements may then be used by the receiver to construct, or to dynamically adjust, an appropriate reservation request.

4.2 Reservation Styles

In RSVP a reservation request includes a set of options that are collectively called the reservation style. One reservation option concerns the treatment of reservations for different senders within the same session: whether to establish a "distinct" reservation for each upstream sender (e.g., in case of a video-conference application), or to make a single reservation that is "shared" among all packets of selected senders (for an Internet phone teleconference − where senders are unlikely to transmit simultaneously). Another reservation option controls the selection of senders; it may be an "explicit" list of all selected senders, or a "wildcard" that implicitly selects all the senders to the session. As a combination of these reservation options RSVP provides three essential reservation styles: Fixed-Filter (FF), Shared-Explicit (SE) and Wildcard-Filter (WF) [1].

4.3 RSVP Protocol Mechanisms

There are two fundamental RSVP message types: Resv and Path. Each receiver host sends RSVP reservation request (Resv) messages upstream towards the senders. These messages must follow exactly the reverse of the path the data packets will use, upstream to all sender hosts included in the sender selection. They create and maintain "reservation state" in each node along the path. Resv messages must finally be delivered to the sender hosts themselves, so that the hosts can set up appropriate traffic control parameters for the first hop.

Each RSVP sender host transmits RSVP "Path" messages downstream along the unicast or multicast routes provided by the routing protocol following the paths of the data. These Path messages store a temporary path state in each node along the way. This includes at least the unicast IP address of the previous hop node, which is used to route the Resv messages hop-by-hop in the reverse direction. Path messages are sent with the same source and destination addresses as the data, so that they will be routed correctly through non-RSVP clouds. On the other hand, Resv messages are sent hop-by-hop; each RSVP-speaking node creates a Resv message to the unicast address of a previous RSVP hop.

5. THE OSPF ROUTING ARCHITECTURE

In this paper we attempt to answer the problems of QoS/multicast routing in terms of a real intra-domain IP routing architecture: OSPF (Open Shortest Path First). OSPF is a link state, unicast, best-effort IP routing policy, which computes minimal cost paths in an Autonomous System. As a consequence of flooding topology information in Link State Advertisements (LSA), all routers are aware of the network structure within the administrative scope of a routing domain.

MOSPF (Multicast extensions to OSPF) adds the ability of distributing group membership information to the OSPF platform and performs on-demand multicast routing. QOSPF (QoS extensions to OSPF) facilitates the distribution of Link State Advertisements describing network links with the indication of free bandwidth and queuing delays in order to achieve QoS routing. Thus to build a QoS/multicast routing protocol on the basis of OSPF seems promising since the basic functionality of dynamic metric advertisement and group membership distribution is already successfully handled in this environment.

6. QUESTIONS RELATED TO MULTICAST, QOS-BASED ROUTING

First of all, one has to decide what service classes to cover with the proposed routing protocol. Nowadays different solutions for best-effort, QoS and multicast realizations exist. We propose extensions that cover unicast and multicast QoS flows, as it is shown in Table 1. By utilizing this approach a QoS multicast routing protocol can be run as an extension of the already implemented and used OSPF and MOSPF, thus there is no need for changing the way non-QoS packets are routed. The reception of RSVP protocol messages triggers the QoS routing process independently from best-effort routing. Responsibilities of QOSPF can be taken over and merged with a new protocol, because QOSPF's implementations have not been commonly spread yet. Nevertheless the same link state database can be used by each routing protocol, OSPF's unicast routing table and MOSPF's forwarding cache are also utilized and can be used by the QoS routing process, henceforward. We assume that the maintaining of QoS flow states and managing QoS sessions are the responsibilities of RSVP; it is RSVP that triggers the QoS multicast routing process (e.g., on receiving Path messages).

SERVICE	UNICAST	MULTICAST
NON-QoS	OSPF	MOSPF
QoS	QoS extensions to OSPF supporting multicast	

Table 1: Routing protocols and data services

6.1 Next-hop vs. source routing

It has to be decided, whether to perform routing at each hop along the path as the datagram proceeds towards the destination or at the ingress router that first handles the specific QoS flow.

Next-hop routing method follows the conception of traditional IP: Path messages are routed at every router along the path, and each router determines only the next-hop towards the destination. After creating a flow state the Path message is sent along the interface identifying the next hop.

Another aspect of routing is source or explicit routing. This is the PNNI approach of ATM [11] to manage multicast/QoS communication. The router directly adjacent to the network initiating the QoS flow computes the entire multicast tree, and the completely specified data forwarding tree is distributed along the calculated tree. Intermediate routers on the path do not perform routing as they use the distributed routing information given by the ingress router.

6.2 Pre-computed or on-demand routing

On-demand routing means that the routing protocol triggers a path computation for every QoS packet previously not seen. This involves no additional issues in terms of when computations should be triggered, but running the path selection algorithm for each new request can be computationally expensive.

On the other hand route pre-computation means that paths are computed to all destinations with all possible residual bandwidths. Hence the resulting pre-computed routing table can be used to look up best routes for each new flow. On the other hand pre-computed routing table must be re-computed according to the changes occurring in resources related to the routing. Route pre-computation amortizes the computational cost over multiple requests, but each computation instance is usually more expensive than in the on-demand case. Furthermore, depending on how often paths are re-computed, the accuracy of the selected paths may be lower.

6.3 Collaboration with RSVP

Since RSVP supports various reservation styles it is not the least indifferent, how a QoS multicast routing protocol can establish such a data forwarding tree that utilizes best the capability of RSVP, especially different reservation styles. As it was detailed earlier, two fundamental styles of multicast tree establishment is relevant in our case: the shared tree conception and the per-source tree approach. Now let us see, how these reservation styles can be associated with multicast tree establishment styles.

First let us assume, that we can solely build per-source trees. The per-source tree approach is advantageous, because it results in an optimal shortest path routing with higher level of load-sharing and fault tolerance compared to the shared tree case. The drawback of this method is its inefficient nature of supporting shared reservation style. It is because shared reservations for different senders can not be merged at the routers since per-source multicast trees from different sources may be separated. Establishing a shared reservation over a per-source tree is against the concept of RSVP – although it is feasible. The result is a reservation scheme, where there are allocated resources for traffic potentially originated from each sender at the same time, while – as a consequence of the reservation style – multiple senders are unlikely to transmit along the same path. The most important problem is that this way the shared reservation waists resources since it reserves resource along the individual per source trees as if aggregated (shared) traffic were going through them. An example reservation is a multiparty teleconference where all independent per-source trees will allocate the one unit bandwidth for communication whereas only one source is likely to transmit at a time.

If we consider the shared tree approach then the result is a routing protocol, which can provide all RSVP reservation styles, thus higher level of resource utilization. On the other hand it introduces traffic concentration at network links and increased delay because of the sub-optimal routing, which can be disastrous to delay-sensitive real-time applications at the endpoints. Another problem is that there exist no known methods to establish, manage and tear down shared trees in an optimal way in this very special intra-domain QoS environment.

We think that in order for a QoS/multicast routing protocol to support all reservation styles it has to be able to establish, manage and tear down both types of multicast trees. A per-source tree may be built for a multicast group indicating distinct reservation style, and a shared tree may be established for a shared reservation style multicast group respectively.

7. SUMMARY

Intra-domain multicast routing with QoS support was investigated in this paper. It can be stated that several problems arise when trying to combine the QoS- and multicast-based routing in a way preserving compatibility, maintaining scalability while realizing efficiency.

As a conclusion we outline, how a QoS multicast routing protocol may select paths for IP flows. All QoS flows (unicast/multicast) may be routed in a source (explicit) fashion i.e., on the reception of an RSVP Path message, RSVP triggers the QoS routing algorithm to calculate the complete path to the destination(s). This path is computed according to the reservation style determined by the multicast group's members. As the next-hop routing approach does not fit into the QoS/multicast routing scenario, the conception of source routing is strongly recommended. Further, the set of next hops is registered in a forwarding cache and Path message is forwarded in the direction of the next-hops delivering the complete data forwarding tree topology. Other routers along the path do not perform route computation. As Resv messages return from the receivers to the senders the data forwarding path for the multicast session is established, flow states are created, and data delivery can be done on the basis of the cached routing information.

Since the need for a multicast routing algorithm considering QoS parameters are vital further research in this area is encouraged.

BIBLIOGRAPHY

[1] Braden, R. (Ed.), Zhang, L., Berson, S., Herzog, S., and Jamin, S.: Resource reSerVation Protocol (RSVP) version 1, Functional Specification – RFC 2205, – ISI, UCLA, IBM, Univ. of Michigan, September 1997.

[2] Moy, J.: OSPF Version 2 –, RFC 2178, Ascend Communications Inc., April 1998.

[3] Zhang, Sanchez, Salkewicz, Crawley: Quality of Service Extensions to OSPF (QOSPF) – INTERNET-DRAFT, Bay Networks, Avici Systems, Redback Networks, GigaPacket Networks, September 1997.

[4] Guerin, R., Kamat, S., Orda, A., Przygienda T., Williams, D.: "QoS Routing Mechanisms and OSPF Extensions" – INTERNET-DRAFT, IBM, Technion, Lucent, December 1998.

[5] Moy, J.: Multicast Extensions to OSPF – Network Working Group INTERNET DRAFT, Ascend Communications, Inc., February 1999.

[6] Moy, J.: MOSPF: Analysis and Experience – RFC 1585., Ascend Communications Inc., March 1994.

[7] Crawley, E., Nair, R., Rajagopalan, B., Sandick., H.: A Framework for QoS-based Routing in the Internet, RFC 2386, Argon Networks, Arrowpoint, NEC USA, Bay Networks, August 1998.

[8] Banerjea, A., Faloutsos, M.: Designing QoSMIC: A Quality of Service sensitive Multicast Internet protoCol – Inter-Domain Multicast Routing INTERNET-DRAFT, University of Toronto, QUALCOMM, October, 1998.

[9] Stardust Technologies, Inc. – An IP multicast Initiative White Paper: Introduction to IP multicast routing, 1995-97.

[10] Braden R., Clark, D., Shenker, S.: Integrated Services in the Internet Architecture: an Overview, RFC 1633, ISI, MIT, Xerox PARC, June 1994.

[11] ATM Forum PNNI subworking group: Private Network-Network Interface Specification V1.0 (PNNI 1.0), 1996.