Quality of Services for Internet Multicast Application Traffic

Jefferson Manhães de Azevedo, Luci Pirmez, Luiz Fernando Rust da Costa Carmo, Oswaldo Vernet de Souza Pires.
Núcleo de Computação Eletrônica – UFRJ Caixa Postal 2324, Rio de Janeiro, RJ, Brasil
E-mail: jazevedo@ucam-campos.br
{luci, rust,oswaldo}@nce.ufrj.br

ABSTRACT

This paper describes an architecture to use multicast with QoS in differentiated services model. That architecture aims at creating an Internet infrastructure for distance learning projects where multimedia interactive applications, like multimedia teleconferences and retrieving multimedia documents, are used. To perform the proposed multicast traffic, the architecture allows an advanced network resource verification and reserve in distribution tree branches of multicast session. This architecture is composed by a reservation system of network resources and differentiated service routers. These routers implement the "multicast premium" service through Expedited Forwarding PHB. Three queuing disciplines were developed in the router schedule module: Priority Queuing, Deficit Round-Robin and to Start-Time Fair Queuing (a variant of WFQ). Two multicast routing protocols types are used such as MASC/BGMP interdomain protocol and a DVMRP intradomain protocol.

Keywords: Integrated Services, Multimedia, Differentiated Services, QoS.

1 – Introduction

The framework reported in this paper is part of a research project about Internet quality of service that has been performed at Universidade Federal do Rio de Janeiro (UFRJ). The project aims at creating an Internet infrastructure for distance learning project at Centro Federal de Educação Tecnológica de Campos (CEFET-Campos). This project is based on the Internet intensive use and of its multimedia applications that stimulate interaction, mainly through video/audio teleconference, as well as multimedia documents retrieval. Such applications need QoS (Quality of Service) levels different from the best-effort packets delivery service model provided by the IP protocol. On account of it, technologies are being investigated so that they can provide QoS in networks based on the IP protocol ("Internet Protocol").

The framework is based on the Differentiated Services Model (Diffserv). The Diffserv model brings a new perspective for the QoS offer in the Internet, making a distinction among the services offered by the network without the need of creating a state and a specific signaling for each flow in each router between the flow source and target, like in Integrated Services Model [1]. By contrast, in the diffserv model the QoS attribution is giving to the flows aggregation (flows that have similar features and must receive the same treatment in each router of the diffserv network).

An important concept in Diffserv model is the one called Domain. The DS (Differentiated Service) Domain is formed by a group of nodes with DS capacity, implementing the same behavior and services supply policies (PHB - Per-Hop-Behavior) in its interior nodes. The DS domain has well defined limits composed by edge nodes (border routers) in charge of classification and conditioning policies of input traffics. These nodes are called ingress nodes, and inside them, the packets belonging to several entrance flows will receive a codepoint [2], identifying the treatment received by those packets in the interior domain nodes (interior routers). Traffic conditioning policies in ingress nodes must assure that flows entering in the domain are according to the characteristics of TCA (Traffic Conditioning Agreement) related to the interlinked domain or network.

One or more networks can compose a DS domain. Usually, a DS domain is formed by a DS service provider network and its customers network. Service Level Agreements (SLA) are estabilished among a provider and its customers. Those SLAs are called intradomain SLAs. The Differentiated services will be able to extend through a group of DS domains, forming a DS Region, since they establish SLA between each upstream and downstream domain of the region. Those SLAs are called interdomain SLAs. The SLA includes the TCA.

The multicast traffic is still an IETF Work Group study issue. This traffic type has two features that create difficulty to its implementation in Diffserv: the packets replication and the multicast groups dynamism. Those features may compromise the resources reserved to unicast traffic, as when a multicast packet arrives to a DS domain ingress node, the packet will be able to leave the domain through several egress nodes. This multicast traffic behavior is able to overload some links and violate some peering SLAs. Violations happen either when a SLA is non- previewed multicast traffic or when the agreement traffic limit is exceeded. In RFC 2475 [3] two considerations are presented for this traffic type in order to preserve the resource reserved commitment to unicast traffic. The former shows multicast traffic should have a set of codepoints different from those addressed to the unicast traffic. The latter indicates multicast traffic should have a different SLA from unicast traffic SLA.

In order to avoid overload of links and SLAs violations, before accepting a new receiver to the multicast session (admission process), it is needed to check available resources along the new path (branch of the multicast session's distribution tree to reach the new receiver) where the flow will pass through. That process consists of checking whether any SLA is violated and whether there are available resources in routers along the path. The SLA verification should happen either for intradomain SLAs (among Diffserv provider and its customers) or for interdomains SLAs (among Diffserv providers).

The admission process of a new receiver becomes complex, as the multicast groups are dynamic. It means that at any moment a new receiver may take part in a multicast session and it is able to receive multicast flows. Thus, it is very difficult to know in advance which are the paths where the multicast flows pass through and the amount of resources for a multicast session.

Another difficulty happens when there is a split of the multicast flow inside the domain. As a multicast flow comes into a domain ingress node and leaves it through several others ones, there is a flow split in some interior router inside the domain. In this case, the multicast flow is replicated and forwarded to at least two paths. It is possible there is not enough resources for the multicast flow in one of the paths, what may cause overload of links and SLAs violations.

The mentioned difficulties are found by multicast traffic implementation, concerning Diffserv. Therefore, a previous forecasting and checking of network resources for the distribution tree of multicast sessions would be needed. A framework to implement the multicast traffic in DS networks is reported below.

2 - Framework for Multicast Traffic in Diffserv

The framework proposed in this work aims at allowing multicast traffic in diffserv networks by making a previous verification and reservation of available resources in paths where the multicast traffic passes through. This framework is composed by a reservation system of network resources and DS (Differentiated Service) routers.

The reservation system is based on Resources Reservation Servers (RRS). These servers are known as “Bandwith Brokers” in the Differentiated Services Model model (Diffserv). The RRS exchanges information with others RRSs, in order to check and reserve network resources in domains, as well as configuring the DS domain routers. Each network should have at least one RRS. The RRSs are aware about its network topology, and they perform the same multicast routing algorithm (intradomain multicast routing protocol) used in their networks, allowing the links discovery where multicast traffic passes through inside the domain.

Routers implement the "multicast premium" service through Expedited Forwarding PHB (Per-Hop-Behavior) [4]. The premium service [5] is an end-to-end service with low delay and jitter, whose resources allocation is made in terms of the peak rate. Despite the recommendation in RFC 2475 [3], the proposed architecture is using the same codepoint and SLA (Service Level Agreement) to unicast and multicast flows.

A multicast session is formed by sources and receivers of multicast flows. However, in the proposed framework only one multicast traffic source with QoS in each session is possible. The data flows of this multicast source are marked with the multicast traffic codepoint. The data flows of other multicast session sources are not marked. Those flows receive best-effort treatment (the packets codepoint value is zero). So that a member of the multicast session receive a multicast traffic with QoS, it is necessary to request a reservation to RRS of its network. This RRS will make an authentication process, verifying whether this receiver has permission to make a reservation. If the receiver has permission, the RRS performs an admission control, verifying whether the receiver has available amount of requested resources. In positive case, the RRS sends a resource request message to domain provider RRS.

When the domain provider RRS receives the reservation request message, it must also make an authentication and admission control to the requester host at the provider network. The admission control performs the intradomain multicast routing algorithm to discover the branch of the multicast session distribution tree binding the receiver (receiver's branch). This RRS checks all links and SLAs with its customer networks in this branch in order to verify whether there is no aggrement traffic violation or overload link.

It is also needed to identify the network adapters of the interior router where the multicast flow is split, because these routers may receive a special configuration. In case the receiver's branch spans to others domains, the RRS of provider exchange information with the RRSs of those domains in order to perform the same procedures of the authentication and admission control. If the multicast tree spans through several domains, it is necessary exchange information among RRSs of its domains so that the same procedures can be performed. In this case, the RRS needs to perform the interdomain multicast routing algorithm to discover the others RRSs.

The proposed framework uses two types of multicast routing protocols. The former is an intradomain multicast routing protocol that is in charge of the multicast packets transport inside a DS domain. The latter is an interdomain multicast routing protocol that is responsible for the multicast packets transport among domains. That approach seems to be more suitable to Internet, since it does not force the domains in a DS region to use the same type of intradomain multicast routing protocol.

Even though the proposed architecture is able to use others interdomain protocols, MASC/BGMP (Multicast Address-Set Claim / Border Gateway Multicast Protocol) [6] interdomain protocol was chosen. That protocol has been currently developed by IETF and has been implemented in some commercial routers. The BGMP protocol allows routes advanced discovery through where multicast flows transit, creating a bidirecional interdomain multicast tree of domains. The domains tree is created from the leaf domains (domains where there are multicast session's members) to Root Domain (domain being the root of the interdomain multicast tree).

The chosen intradomain protocol was the Distance Vector Multicast Routing Protocol (DVMRP) [7], as it has been extensively used in the Internet and it allows to specify the source of the multicast session. The last feature allows to make a reservation to a specific source in the routers.

In case of failed reservation, multicast flow packets passing through multicast tree branch, binding receivers without reservations, should be re-marked to best-effort treatment. On the other hand, in an Intserv environment, when a receiver cannot perform a certain reservation level, it will be able to receive multicast session data flows in a lower reservation level that may be different from "best-effort" treatment. But it would be a problem for the proposed framework, because in domains where the multicast traffic is split there would be need for interior routers re-mark data flows packets to an appropriate service treatment. It might be necessary to perform a new policing, having as a consequence the need of a reservation protocol state information to be stored in interior nodes of domains, as it can be seen in Berson [8]. It increases the interior nodes complexities, what should be avoided in the Diffserv model.

By contrast, in the proposed architecture, all multicast flow packets passing through a router are re-marked to best-effort treatment. If a branch has a receiver with reservation to a multicast session, all routers in this branch does not re-mark the packets, which belong to this multicast session (these packets receive a QoS treatment). In this specific case, a flag (a special flag) should be set in the multicast routing table (this configuration is made by the RRS), indicating to the marker module which packets are not be re-marked. It is needed to set one flag in the multicast routing table to each router network adapter that has a multicast session's members with reservation.

3- Implementation

A prototype of the framework presented in this work was developed to allow the intradomain multicast traffic.

The RRS (Resources Reservation Servers) was developed by using the Delphi 4.0 developing enviroment with paradox tables and Windows 98 operating system. The RRS has the following functional modules: authentication, admission control, communication, configuration and paths discovery modules. The authentication module verifies whether a host has permission to make a reservation. The admission control module must verify the network resources availability to admit a new receiver of multicast flow with QoS. The communication module performs the exchange of reservation and configuration information among RRSs and routers. The configuration module is in charge of routers configuration. Lastly, the route discovery module is responsible for the discovery of links and routers that compose the branch linking the receiver to the multicast session distribution tree. These routes have information about the routers and routers network adapters.

The RRS is composed by three applications that implement its funcions. This approach allows a logic division of the RRS and parallel development of the applications. Moreover, this logic division allows a future multi-thread development of the RRS. The first aplication is in charge of the configuration data input. The second application is responsible for implementing the authentication module, the admission control module and the dicovery route module. That application is always awaiting actively a profile or reservation request. The third application implements the configuration module. The communication module is implemented by both in the second and third applications. Another application was developed in order to allow receivers to make multicast session profile or reservation request, called reservation agent.

The routers use TROPIX operating system (UNIX standard) [9]. There are two types of routers: border and interior; each one with three modules: classification, schedule and configuration. Besides, the border router has a condition module. The classification module is in charge of classifying flows and the aggregated flows through the packet codepoint. The schedule module is responsible for the differential treatement and it has implemented four queueing disciplines: FIFO, Priority Queueing, Deficit Round-Robin and to Start-Time Fair Queueing (a variant of WFQ). The condition module must assure that the flows entering in the domain are according to the features of TCA (Traffic Conditioning Agreement) and it uses a "Token-Bucket" algorithm. These three modules are part of the system operation kernel, because every time a packet arrives it is necessary to activate these modules.

The configuration module exchanges information with the RRS in order to make a configuration of the router. It creates reservation states or it makes the changes needed in the multicast routing table. This module is an aplication user because it is activated with low frequency.

4 - REFERENCES

[1] R. Braden, D. Clark, S. Shenker, "Integrated Services in the Internet Architecture: an Overview", RFC 1633, June, 1994

[2] K. Nichols, S. Blake, F. Baker and D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", RFC 2474, December 1998.

[3] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, "An Architecture for Differentiated Services", RFC 2475, December 1998.

[4] V. Jacobson, K. Nichols, K. Poduri, "An Expedited Fowarding PHB", RFC 2598, June 1999.

[5] K. Nichols, V. Jacobson, and L. Zhang, "A Two-bit Differentiated Services Architecture for the Internet", Internet Draft <draft-nichols-diff-svc-arch-00.txt>, November 1997.

[6] S. Kumar, P. Radoslavov, D. Thaler, C. Alaettinoglu, "The MASC/BGMP Architecture for Inter-domain Multicast Routing", SIGCOMM´98.

[7] D. Waitzman, S. Deering, C. Partridge, "Distance Vector Multicast Routing Protocol", RFC 1075, Novembro 1988.

[8] S. Berson, S. Vicent, "Aggregation of Internet Integrated Services State", Internet Draft, Agosto 1998.

[9] http://www.tropix.nce.ufrj.br/