IP QoS evaluation using interoperability of differentiated and integrated services

BRUNONAS DEKERIS, LINA NARBUTAITE
Faculty of Telecommunications and Electronics
Kaunas University of Technology
Studentu str. 50, Kaunas
LITHUANIA

Abstract: - Both QoS architectures, Integrated and Differentiated services, have their own advantages and disadvantages. By combining the advantages from both models it might be possible to develop a flexibile system which would provide predictable services. We present the network architecture that provides end-to-end quantitative QoS evaluation consists of IntServ regions at the access of the network and Diffserv regions in the core of the network. The paper describe the mathematical model and simulation results, which explore what kind of quality of service the combination of Integrated and Differentiated Services can offer.

Key-Words: QoS, delay, integrated services, differentiated services , priority, scheduling

1. Intoduction

Technical revolution in telecommunications and computer science have led to the development of new types of distributed applications such as multimedia and co-operative system with interactive simulation. These applications present problems to network designers, such as higher bandwidths, the need for bounded delays, etc. Quality of Service (QoS) describes the assurance of data transfer that fits to the application requirements. Traditionally, the best-effort service has been the main type of service available over networks. While this type of service has contributed much towards the rapid growth of the Internet, it can not support applications that have real-time requirements. As the Internet resources become more and more constrained, the Best Effort (BE) model is increasingly proving less capable of providing the required Quality of Service. This area is one of the main topics of research and development in data networks. In the Internet community, IETF1 Integration services (IntServ) [1] and differentiation services (DiffServ) [2] working groups have been carried out in order to develop a QoS framework for the TCP/IP protocols. Several projects have also been initiated to target the QoS problem; let's cite the TF-TANT activity and the GEANT, TEQUILA, CADENUS, AQUILA and GCAP projects.

In order to exploit the individual advantages of IntServ (per flow QoS guarantee) and DiffServ (good scalability in the backbone), interconnection of IntServ and DiffServ, requires a mapping from IntServ trafficc. Flows to DiffServ classes have to be transformed at the ingress to the DiffServ network. Some works has been carried out in the area of interconnecting IntServ and DiffServ. In [3] a concept for the integration of IntServ and DiffServ is presented, and describes a prototype implementation using commercial routers. Some ideas on a packet forwarding algorithm to forward packets from the Guaranteed Class traffic of IntServ to Expedited Forwarding class of DiffServ are suggested [4-6]. Other authors show that packet loss in the DiffServ network can result in bursty loss to the IntServ applications. In [7] results are presented which determine performance differences between IntServ and DiffServ, as well as some characteristics about their combined use. A new DiffServ class for carrying RSVP signaling originating from the edge IntServ domain is proposed in [8]. However, the above studies do not present any numerical result to evaluate the QoS guarantee that can be achieved by end applications. The authors are not aware of study which quantitatively shows the QoS that can be achieved by IntServ end applications when IntServ and DiffServ are interconnected. The objective of this paper is to *quantitatively evaluate* the QoS guarantees that can be obtained by end applications when IntServ is run over DiffServ. We present the mathematical model and simulation results, which explore what kind of quality of service the combination of Integrated and Differentiated Services can offer as well.

2. Interoperability of Differentiated and Integrated Services

Integrated (Intserv) and Differentiated (Diffserv) services are popular Quality of Service (QoS) models for the internet that receives most of the attention. The Integrated Services model is characterized by resource reservation. For real-time applications, before data are transmitted, the applications must first set up paths and reserve resources. RSVP is a signaling protocol for setting up paths and reserving resources. Differentiated services (DiffServ) are intended to enable the deployment of scalable service discrimination in the Internet without the need for per-flow state and signaling at every hop. The assumption of DiffServ networks is that routers in the core network handle packets from different traffic streams by forwarding them using different per-hop behaviors (PHBs).The DiffServ architecture is shown in Figure 1.

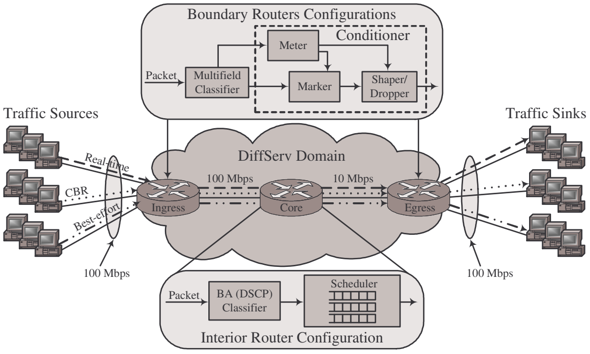

Fig. 1 DiffServ architecture

Both QoS architectures, Integrated and Differentiated services, have their own advantages and disadvantages. Integrated Services together with RSVP signaling provide dynamic admission control, end-to-end predictable services and an efficient use of a network resources. On the other hand, the processing overload in the high speed networks will be unacceptable. Differentiated Services offers a scalable way to provide QoS also in large networks. The lack of a dynamic admission control mechanism and difficulties in resource allocation can make the services quite unpredictable. By combining the advantages from both models it might be possible to build a scalable system which would provide predictable services. Since DiffServ focuses on the needs of a large network, it is clear that DiffServ approach should be used in high speed transit networks.

Another use for IntServ and RSVP might be found in the bottleneck links of the Internet, for example in access networks. Since IntServ isolates flows from each other, it provides efficient protection against misbehaving flows. Consequently, quantitative services can also be offered over the low-speed bottleneck links. Since a narrow bandwidth link can not offer several simultaneously reserved connections, scalability will not be the problem here. By using IntServ and RSVP in links like this, the link utilization increases and several customers can be offered quantitative services on demand. The main necessity of cooperation InterServ and DiffServ is that we have possibility to manage the QoS . The co-operation InterServ and DiffServ model is shown in the Figure 2 .

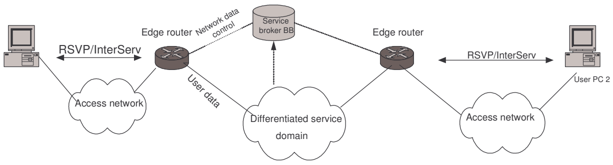

Fig. 2 The cooperation InterServ and DiffServ model

The network architecture which provides end-to-end quantitative QoS consists of IntServ regions at the access of the network and Diffserv regions in the core of the network. Nevertheless a network administrator is free to choose which regions of the network act as IntServ regions and which ones act as DiffServ regions. Basically IntServ regions are customers for DiffServ regions, which offer transport services. When IntServ and RSVP are used to perform admission control to DiffServ network, the most important node in the network is the edge router in the boundary between the network regions. The router can be seen as consisting of two halves: the standard RSVP half, which interfaces the stub networks and the DiffServ half, which interfaces the transit network. The RSVP half is able to process PATH and Resv messages, but it is not expected to store full RSVP state. In the simplest case admission control has the knowledge of how much bandwidth has been used and how much bandwidth is still left. By using this information and the token bucket parameters in RSVP Resv message, the router is able to decide if a new connection can be granted.

2.1 Co-operation Intergrated and Differentiated services mathematical model

At first we analyze the integrated services model using RSVP and token bucket scheduling.

The model under consideration is shown in Figure3 [10]

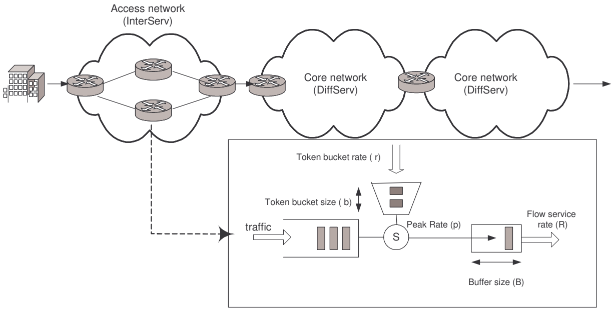

Fig.3 Integrated services model

For the simple RSVP case , the end – end delay bound is given [10] by

$$d_{\max} = \frac{b}{R} + \frac{C_{tot}}{R_t} + D_{tot}. \quad (1)$$

where : (r; b) -these are the parameters of the token bucket; *r* denotes the rate at which bytes will be produced and *b* is the depth of the token pool (in bytes); R - the rate in bytes per second (note this is the server rate of the fluid model); C_{tot} - the total rate-dependent delay computed end-to-end on the path from the sender to the receiver; D_{tot} - additional delay which is not-rate – dependent, is a result of time spent waiting for transmission through a node (unit is in microseconds).

For the complex TSpec:

$$\begin{cases} d_{\max} = \dfrac{(b-M)(p-R)}{R((p-r)} + \dfrac{M+C_{tot}}{R} + D_{tot}, \text{if } p \ge R \ge r \\ d_{\max} = \dfrac{M+C_{tot}}{R} + D_{tot} \text{ if } R \ge p \ge r \end{cases}. \quad (2)$$

where: M is the maximum datagram size that will be sent by the flow; *p* - this is the peak rate at which packets will be sent.

The differentiated services model is shown in Figure 4 [12]

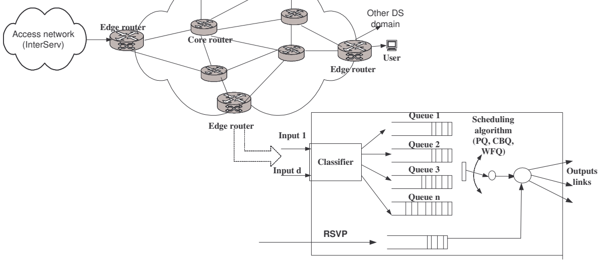

Fig.4 Differentiated services model

Let's denote 0<α<1 the max priority traffic utilization factor allowed on any link. Assume that a single output has capacity C_{out} and a single flow reserve rate is *r*. This flow arrive from input link of the same capacity as the output link. The further *n* other flows sharing the output link with the non-priority flow. All of these flows are equally distributed among the remaining *d* input links and that each of the priority (real-time) and non-priority (best effort) flows conforms to a token bucket *(b,r)*. Then the maximum number of priority flows sharing the output link can be [11]

$$n = \frac{\alpha C_{out}}{r}; \quad (3)$$

where α is utilization of the output link.

Let's consider time *t* in which a burst of size $\frac{nb}{d}$ can arrive in input link C_{in}. During this time the total of *nb* priority flows can arrive at the router and

$$tC_{out} = \frac{nbC_{out}}{dC_{in}}. \quad (4)$$

Then the router delay is given by

$$D = \frac{\alpha * b}{r}\left(1 - \frac{C_{out}}{dC_{in}}\right). \quad (5)$$

However the network consists of *H* nodes and we transmit the *M* bytes size packets . Therefore, if we want to calculate the total delay, we need to define the new parameter – *degree of hop*. A router is called an *h-degree hop*, if there is one flow at least for which this hop is exactly the *h*-th hop in its route and for all others this hop is no more then *h*-th hop in their route.

Then the delay at hop of degree *j* can be written as

$$D_j = \left(\frac{\alpha b}{r}\frac{(d-2)}{d} + \frac{M}{C}\right)\left(1 + \alpha\frac{(d-2)}{d}\right)^{j-1}. \quad (6)$$

Now , the total queuing delay is

$$D_N = \sum_{j=1}^{H} D_j = \left(\frac{\alpha b}{r}\frac{(d-2)}{d} + \frac{M}{C}\right) \times$$

$$\times \left(\frac{\left(1 + \alpha\frac{(d-2)}{d}\right)^{H} - 1}{\alpha\frac{(d-2)}{d}}\right). \quad (7)$$

where *H* is number of nodes (hop), d –number of inputs link, *M*-packet size.

3. Simulation

In this section simulation results are presented, which explore what kind of quality of service the combination of Integrated and Differentiated Services can offer. The scenario used in the simulation assumes that RSVP and IntServ are employed in access networks, and DiffServ - in core networks. The network structure is presented in Figure 5.

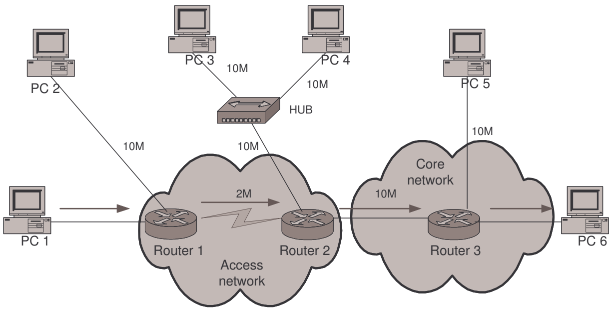

Fig. 5 The simulation network structure

All the traffic was generated by software package *Arena*. At first, the test network was lightly loaded and about 15% of the link capacity was real-time traffic. In the second and the third cases the network was overloaded consisting of 30% and 50% of real-time traffic, respectively. Table 1 presents the detailed traffic characteristics of the network in each case. In each test case the network was configured to represent four different kinds of QoS architecture. The network types were as follow: 1) best-effort, 2) IntServ, 3) access link IntServ and core link DiffServ, 4) access link IntServ and core link best-effort. In addition, to the case when the core link was DiffServ, two different scheduler types were used: priority queuing (PQ) and class based queuing(CBQ).

Table 1 Traffic characteristics of the network

Traffic	Packet size [byte]	Packet rate [pp/s]	Data rate [kbit/s]
VoIP	520	10	41.6
Video	1250	20	200
BE	1500	125	1500

Total amount of real-time data in the access link changes from 369 kbit/s to 1164 kbit/s and in the core link - 1164÷ 3692 kbit/s. Best-effort network used FIFO queuing with the queue size of 40 packets. Priority queuing separated real-time and best-effort traffic into two queues. The higher priority real-time queue was served first and if that queue was empty, the normal priority best-effort queue was then served. The queue size was configured to be 80 packets for the real-time traffic and 140 packets for best-effort traffic. For CBQ scheduling - queue sizes were 100 packets and 200 packets for real-time and best-effort traffic, respectively.

The traffic under consideration was transferred from PC1 to PC6. The PC2÷PC4 create additional load to access network, and PC5 – to core network. We calculate QoS parameters (delay, loss) for three cases and five different QoS combinations. The simulation result are shown in figure 6 - 8

At first, the test network was lightly loaded and about 15% of the link capacity was real-time traffic.

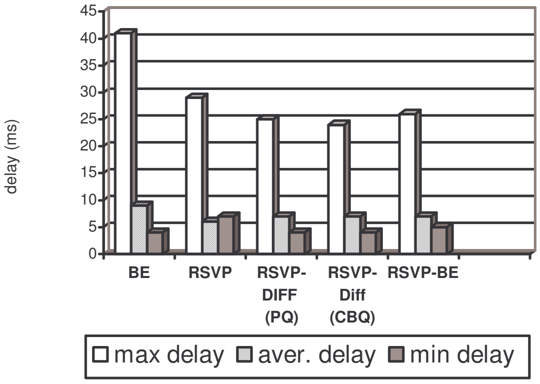

Fig 6 Delay values using five different QoS combinations (case1-real-time traffic 15%)

The Fig.6 shows that when the network is lightly loaded all mechanisms provided almost equal services and we could support QoS parameters target values. The second case - 30% real-time traffic of the link capacity.

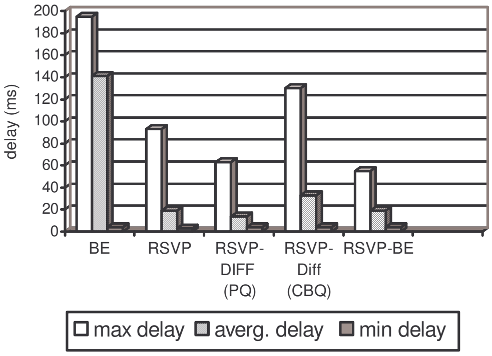

Fig 7 Delay values using five different QoS combinations (case1-real-time traffic 30%)

The pure best-effort network could not provide sufficient service for real-time traffic anymore (Fig.7). The pure RSVP and combination of RSVP and DiffServ implemented by priority queuing were still able to offer reasonable QoS. The CBQ implementation increased delays, but bandwidth was still large enough for the VoIP. The most interesting observation is that the combination of RSVP and the best-effort configurations seemed to offer quite satisfactory QoS. This is because only the access link was considerably overloaded and there RSVP offered the reservation. The core link was only slightly overloaded and queue depth was kept short. Therefore 1.06 % packets were dropped and delays remain small. The last simulation - the network consisting 50% real-time traffic of the link capacity.

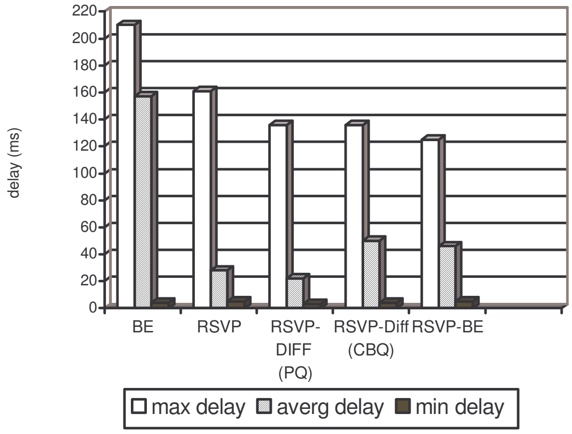

Fig 8 Delay values using five different QoS combinations (case1-real-time traffic 50%)

The most interesting feature of RSVP network its behavior . When the number of flows increased, the processing overhead became probably too high and some packets were dropped. However, the priority queue configuration did not drop any packets and the average delay stayed within a reasonable value. The Figures 9 - 11 illustrate these delays for each case.

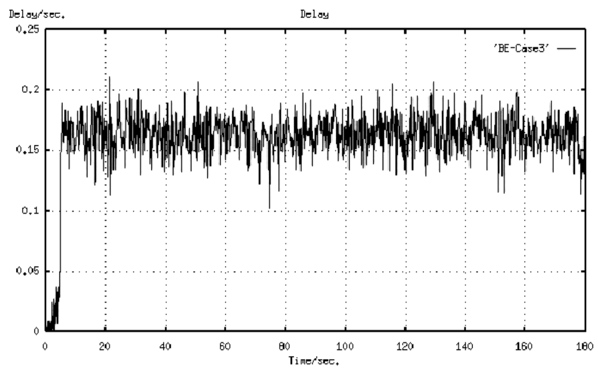

Fig 9 Delay values in best-effort configuration, case 3

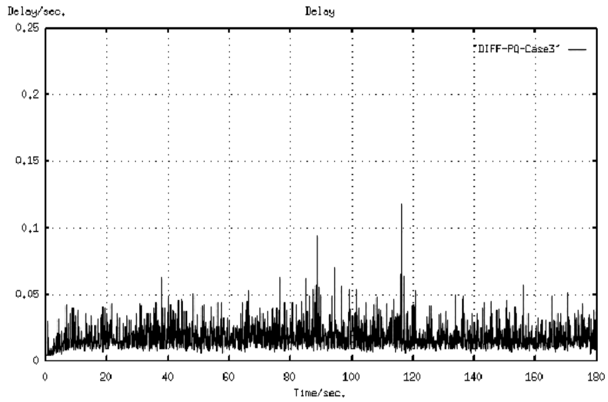

Fig 10 Delay values of combined network, for DiffServ PQ, case 3

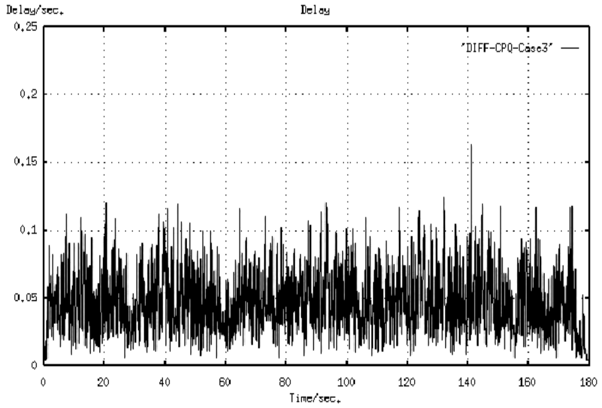

Fig 11 Delay values of combined network, for DiffServ CBQ, case 3

The priority queuing implementation gave the best QoS values for the test flow. The packets were not dropped and the delays stayed quite close to the case, which could be seen by comparing Figures 10 and 11. The few delay peaks might be caused by bursts of large video stream packets inside the real-time aggregate flow. Figure 11 shows that when the size of the aggregate flow was equal to the configured rate of the CBQ scheduler, the delays and jitter increased remarkably. The jitter might be caused by the bursty video stream inside the aggregate. Actually there should be a traffic shaper at the edge node, which smooths the traffic stream before it enters into DiffServ network. By comparing Figures 10 and 11 one can see that PQ gives a much better QoS values for real-time traffic than CBQ. The similar delay results give using RSVP-DiffServ (CBQ) and RSVP-BE, but packet loss are very different: RSVP-DiffServ (CBQ) - 2.7%, RSVP-BE – 10.3%. Therefore we could use combination RSVP-BE , because it did not support QoS target value (loss - 5%).

4. Conclusions

In this paper, we have evaluated the QoS that can be obtained by end applications when Integrated Services (IntServ) access network are connected together using Differentiated Services (DiffServ) network. Traffic from IntServ are inserted into various classes with different priorities DiffServ services and QoS can guaranteed to individual applications. The simulations results show that :

- when network is lightly loaded all networks type provided almost equal services and we can support QoS parameters target values;
- when network load is 0.8 and 30% of the link capacity are real-time traffic, the best combination is RSVP-BE, because only the access link was considerably overloaded and there RSVP allow to eliminate this overload using the resource reservation. The core links are only slightly loaded and there we can use BE ;
- when network is overload and 50% of the link capacity are real-time traffic, the combinations RSVP-BE and RSVP-DiffSer (PQ) give the similar delay, but packet drops using RSVP-BE are 10.3% (it is over QoS norm), therefore the best combination is RSVP-DiffSer (PQ) for this case.

The co-operation differentiated and integrated services gives the better QoS parameters support (1.9 time) then using these models separately . This model is preferable when network is overloaded.

References:

[1] R. Braden, D. Clark, S. Shenker. Integrated services in the internet architecture: an overview. *RFC 1633*, June 1994.
[2] S. Blake, D Black et. al. An Architecture for Differentiated Services. *RFC 2475*, December 1998.
[3] R. Balmer, F. Baumgarter, T. Braun. A concept for RSVP over Diffserv. *Proc. Nineth ICCCN'2000,* 2000
[4] B. Dekeris, L. Narbutaitė, T. Adomkus. Quality evaluation of proportional differentiated services. *Electronics and Electrical Engineering,* No 6(48)– Kaunas: Technologija, 2003, pp. 67-70
[5] B. Dekeris, L. Narbutaitė, T. Adomkus. Throughput and delay evaluation for IP networks. *Electronic and Electrical Engineering* , No 6 (41) – Kaunas: Technologija, 2002, pp. 67-71
[6] B. Budiardjo, B. Nazief, D. Hartanto, Integrated services to Differentiated services packet forwarding:Guaranteed service to expedited forwarding PHB. *25th Annual IEEE Conference on Local Compute rNetworks*, 2000, pp. 324–325.
[7] J. Harju, P. Kivimaki, Co-operation and comparison of diffserv and intserv: Performance measurements. *Proc. 25th Conference on Local Computer Networks*, 2000, pp. 177–186.
[8] G. Mamais, M. Markaki. Efficient buffer management and scheduling in a combined IntServ and DiffServ architecture; a performance study. *Second International Conference on ATM*, 1999, pp. 236–242.
[9] G. Armitage, *Quality of Service in IP Networks: Foundataion for a Multi-Service Internet*, Macmillan Technical Publishing, April 2000.
[10] B. Dekeris, Quality of services (QoS) support for IP networks. *Electronic and Electrical Engineering* , No 6 (41) – Kaunas: Technologija, 2002, pp. 67-71
[11] A. Charny, Delay Bounds In a Network With Aggregate Scheduling. *Cisco Systems*, 1998, pp. 1-31
[12] B. Dekeris, L. Narbutaitė, T. Adomkus, The analysis of service disciplines of Differentiated services flows. *Electronics and Electrical Engineering*. No 6(55)- Kaunas: Technologija, 2004, pp. 34-39.