Proportional Service Differentiation with MPLS

João Neves, Paulo Rogério Pereira, Augusto Casaca
INESC ID, Rua Alves Redol, 9. 1000-029 Lisboa, Portugal.
Phone: +351-213100345. Fax: +351-213145843.
Email: {Joao.Neves, Paulo.Pereira, Augusto.Casaca}@inesc-id.pt

Abstract. This paper describes two traffic engineering techniques for implementing proportional differentiated services based on Multiprotocol Label Switching constraint based routing. Both use a dynamic bandwidth allocation scheme to modify the bandwidth reserved by each traffic class according to the current network load. The first scheme uses an adaptive algorithm that qualitatively determines the required average throughput per source for each class and moves bandwidth between classes for each path as necessary. The second scheme mathematically divides the existing bandwidth through the traffic classes for each path. The quality of service that users get with both techniques is assessed by simulation and compared with a fixed bandwidth allocation scheme.

1 Introduction

Quality of service (QoS) has been a major research topic in the networking community for the last 20 years [1]. Different user applications have different QoS requirements. Depending on the user application, parameters such as end-to-end delay, jitter, packet loss and throughput may be critical. The simplest answer to provide the QoS that the user applications need is to increase the capacity of the links. This is known as overprovisioning and results in a waste of resources, especially for end-to-end QoS.

The Differentiated Services architecture [2] offers a scalable solution for providing end-to-end QoS. Packets from different user applications with different QoS requirements are classified into different traffic classes and marked accordingly. Only ingress edge routers perform classification, marking and policing functions. The network core only deals with the aggregate of flows in each traffic class, giving them a common queue or scheduling behavior. This allows a scalable service differentiation according to the traffic class.

However, if a fixed bandwidth allocation is used for each traffic class, a priority inversion may occur when a high priority class has higher load than a low priority class. A possible solution is to use flow admission control to avoid this situation. It has the disadvantage of refusing the requested traffic class to flows that would degrade the QoS of existing flows. Although the admission control is necessary to limit the maximum total network load and avoid QoS degradation, if the network resources in the core devoted to each traffic class are made to depend on the actual load of the traffic classes, fewer flows will see their traffic class changed by the admission control.

Towards this objective, the paper proposes two kinds of dynamic bandwidth allocation schemes based on traffic engineering (TE) techniques for implementing differentiated services supported by Multiprotocol Label Switching (MPLS) constraint based routing (CR) [3].

MPLS uses a label switching technique that consists in the addition of a label to each packet. This label indexes a forwarding table in each router, determining the outgoing link and the new label to use. The processing involved can be made much faster than looking at the destination address in the packet as in a routing process. Constraint-based routing (CR) is a mechanism used to meet Traffic Engineering requirements. It allows for a path to be setup based on explicit route constraints, QoS constraints and other constraints. The use of MPLS-CR was selected because it allows reserving and dynamically modifying the bandwidth in a path between a network core ingress and egress node. It can be used to configure the bandwidth, as necessary, for every traffic class in each path of the core network.

The first scheme proposed in this paper uses an adaptive algorithm that determines the required average throughput per source for each traffic class, according to the current network load, and moves bandwidth between classes for each path in the network core as necessary.

The second scheme proposed in this paper mathematically divides the existing bandwidth through the traffic classes for each possible path in the network core according to the current network load.

We propose to use an Olympic service model with three traffic classes: Gold, Silver and Bronze. We propose also to use a proportional differentiation mechanism between the Gold and Silver classes to give the Gold class the best QoS. The Bronze class is a best effort class that gets a very small fraction of the existing bandwidth along with the unused bandwidth of the other classes. This nomenclature is slightly different from the Olympic service nomenclature proposed in [4]. A proportional differentiation mechanism [5] is a refinement of a differentiated services mechanism, where an adjustable and consistent differentiation between classes is provided. The proportional differentiation has the advantages of being controllable, consistent and scalable. The proportional differentiation mechanism should try to give to each Gold flow the double of the bandwidth as to each Silver flow. This factor of two was chosen arbitrarily, as an example. In practice, it should depend on the relation of the prices charged to clients in the different classes.

The Gold and Silver classes could be further divided according to the specific QoS needs of the applications. For simplicity, we assumed that all applications have similar QoS requirements, except for the transport protocol used. TCP applications tend to increase their transmission rate until there is a packet loss, which should not affect the transmission rate of UDP applications. Accordingly, we propose to have different classes for TCP and UDP traffic, resulting in four MPLS-CR reservations: Gold TCP, Silver TCP, Gold UDP and Silver UDP.

Both proposed dynamic bandwidth allocation schemes were analyzed by simulation and compared with a fixed bandwidth allocation scheme for the QoS that users get. Section 2 describes the dynamic bandwidth allocation schemes. Section 3 describes the simulation scenario and the simulation results. Finally, section 4 draws conclusions and raises further work topics.

2 Dynamic Bandwidth Allocation Schemes

2.1 *Adaptive Dynamic Bandwidth*

In the first dynamic bandwidth allocation technique proposed, an adaptive algorithm divides the bandwidth between the traffic classes according to a few rules that determine when the reserved bandwidth should be moved from a traffic class to another.

First, a value for the required throughput per source (TpS) is determined. This is done independently for the Gold and Silver classes and for the UDP and TCP protocols. The TpS values are tabled, as shown in table 1, using steps that become smaller as the number of sources in a node increases, so as to accommodate the throughput of all the sources in a given path. The required TpS for the Gold class is about twice the TpS for the Silver class, favoring the Gold class for higher loads, so that the QoS is higher for the Gold class. The values chosen depend on the characteristics of the sources that were used.

Table 1. Throughput per Source (TpS) values according to the number of sources

Number of Sources	Gold Class	Silver Class
up to 4	100 Kbps	50 Kbps
5 to 8	50 Kbps	25 Kbps
9 to 10	40 Kbps	20 Kbps
11 to 12	30 Kbps	15 Kbps
13 or more	25 Kbps	10 Kbps

If the total bandwidth required for a class, determined as (TpS × number of sources), is higher than what is currently reserved with MPLS-CR, the dynamic algorithm searches in all the links of the path if it is possible to increase the reserve, considering the sources of the different paths sharing the same links. If possible, the reserve for the path is increased. If not possible, the algorithm will try to decrease the reserves of other paths. The Silver class is always the first one to be decreased, even if the predefined TpS steps for the Silver class are not respected. The Gold class is only decreased if the TpS allows it. If no decrease is possible, the bandwidth remains unchanged, which usually happens with high congestion. For all classes, there is always a minimum bandwidth guaranteed, considering the number of existing sources.

2.2 *Mathematical Dynamic Bandwidth*

In the second dynamic technique, a mathematical approach to divide the bandwidth among the classes is used. The idea is that the bandwidth available per source, for the Gold class should be twice the bandwidth available per source for the Silver class. The bandwidth available per source is related to the QoS users get.

The corresponding expressions are:

$$\begin{cases} bwGoldTCP + bwSilverTCP + bwGoldUDP + bwSilverUDP = TOTALbw \\ \dfrac{bwGoldTCP}{NrSourcesGoldTCP} = 2\times\dfrac{bwSilverTCP}{NrSourcesSilverTCP} \\ \dfrac{bwGoldUDP}{NrSourcesGoldUDP} = 2\times\dfrac{bwSilverUDP}{NrSourcesSilverUDP} \\ \dfrac{bwGoldTCP}{NrSourcesGoldTCP} = \dfrac{bwGoldUDP}{NrSourcesGoldUDP} \end{cases}$$

The first expression says that the total available bandwidth in a link has to be divided through the 4 existing aggregated flows, one for each combination of the existing classes with the different application protocols. The 4 aggregated flows follow the same source-destination path, whose bottleneck bandwidth is given by the value TOTALbw. This restriction does not exist in the adaptive algorithm. There are also fixed fractions of the link bandwidth that are reserved for signaling and best effort traffic, both excluded from the above expressions.

The second expression says that the bandwidth per Gold TCP source should be twice the bandwidth per Silver TCP source. The factor of two was chosen as an example of QoS proportionality between the Gold and Silver classes. The bandwidth per source is determined by dividing the total bandwidth of the aggregated flow by the number of flows existing in the aggregated flow. If the TCP applications continuously generate traffic, this expression results in the throughput QoS parameter for each Gold TCP flow to be twice as in the Silver TCP class.

The third expression is similar for the UDP protocol: the bandwidth per Gold UDP source should be twice the bandwidth per Silver UDP source. This usually results in a better QoS (lower delays, jitter and loss) for the Gold class.

The fourth expression says that each flow should get the same bandwidth, whether it is using TCP or UDP. Only the class should influence the QoS, not the protocol chosen.

Solving these equations, we get the bandwidth allocation for each class, corresponding to the reservations to be made in the network, as:

$$\begin{cases} bwSilverUDP = \dfrac{TOTALbw \times NrSourcesSilverUDP}{2.NrSrGoldTCP + NrSrSilverTCP + 2.NrSrGoldUDP + NrSrSilverUDP} \\ bwSilverTCP = bwSilverUDP \times \dfrac{NrSourcesSilverTCP}{NrSourcesSilverUDP} \\ bwGoldUDP = 2 \times bwSilverUDP \times \dfrac{NrSourcesGoldUDP}{NrSourcesSilverUDP} \\ bwGoldTCP = 2 \times bwSilverTCP \times \dfrac{NrourcesGoldTCP}{NrSourcesSilverUDP} \end{cases}$$

The mathematical algorithm periodically modifies the bandwidth reserved for each source-destination path, according to these expressions.

3 Simulations and Results

3.1 *Network Configuration*

The network topology used for the simulations is shown in figure 1. The network has a core containing 12 nodes with at least 5 link-disjoint paths between each pair of nodes. This allows for some redundancy, so that there are several alternative paths between each source-destination pair that can be used by traffic engineering. Additionally, there are 3 ingress routers and 3 egress routers where traffic sources and traffic sinks are placed. The core links were configured with 1 Mbit/s capacity and 2 ms delay. Since we only wish to study the behavior of the core, the access links were configured with a sufficient large capacity. The network topology is described and discussed in [6]. It was simulated in the NS2 network simulator [7].

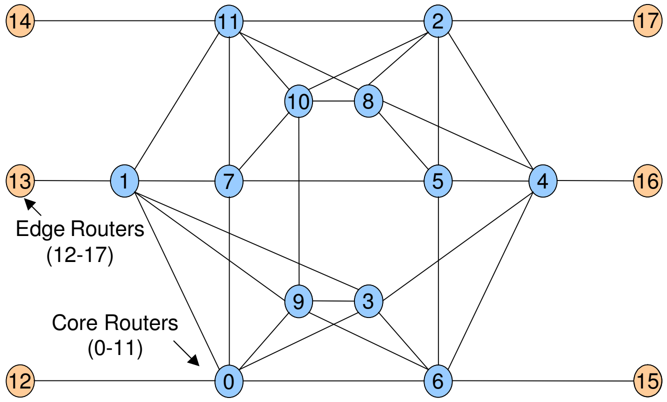

Fig. 1. Network topology

Three types of traffic sources were used simultaneously: long lived FTP/TCP sources, constant bit rate (CBR) sources over UDP and Pareto On/Off sources over UDP. Sources are randomly generated with equal proportions of the three traffic types. Therefore, there might be paths more congested than others in some simulations.

Paths are reserved in the core through MPLS-CR for each pair of source-destination nodes. The Bronze class has no paths reserved through MPLS. For Gold and Silver classes, there are 2 classes × 2 protocols × 3 source nodes × 3 destination nodes = 36 aggregated flows mapped onto MPLS reservations. All these aggregated flows are created initially. The dynamic bandwidth allocation schemes only modify the MPLS-CR bandwidth reservations for each path as necessary.

3.2 *Fixed Bandwidth*

The fixed bandwidth scenario is used only as a term of comparison in this paper. It divides the bandwidth along the classes as follows: 30% for Gold UDP, 30% for Gold TCP, 15% for Silver UDP, 15% for Silver TCP, 1% for signaling and 9% for Bronze. These values were chosen so that at least 10% of the bandwidth is available for the best effort traffic, the Gold class has the double of the bandwidth of the Silver class and the UDP and TCP classes have the same bandwidth available. Naturally, these values should depend on the services sold by the Internet Service Provider and the expected traffic pattern. But, in this scenario, they are fixed, whatever the actual traffic pattern is, while in the following two scenarios, they change according to the two algorithms proposed.

From the different simulations performed, the most interesting case is the one where the number of Gold class sources increases with a fixed number of 6 Silver sources and 60 Bronze sources. The results for the three different types of traffic sources (FTP/TCP, Pareto/UDP, CBR/UDP) are shown in figures 2, 3 and 4, respectively. Several simulations runs were made with different random number seeds and the same number of traffic sources. The results presented correspond to the average of the different simulations.

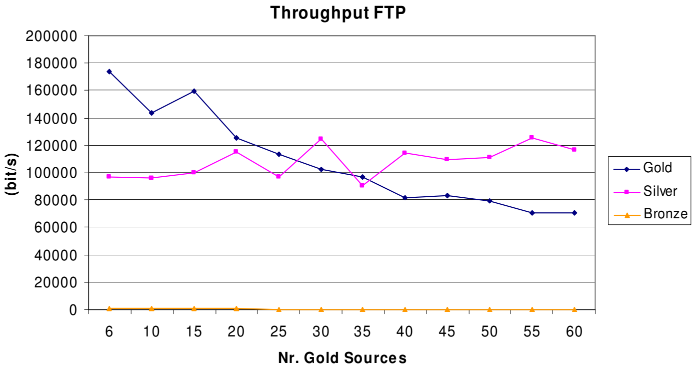

Fig. 2. FTP throughput per source for fixed bandwidth scenario

The results show a priority inversion between the throughput for the Gold and Silver classes for high loads: the fixed bandwidth available for the Gold class is divided by a large number of sources, resulting in less throughput per source and worse QoS than for the Silver class.

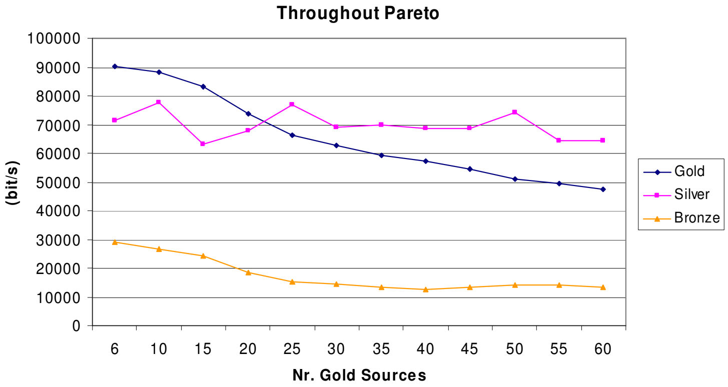

Fig. 3. Pareto throughput per source for fixed bandwidth scenario

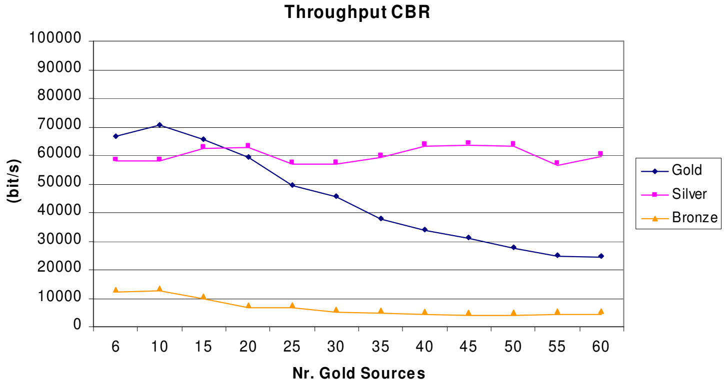

Fig. 4. CBR throughput per source for fixed bandwidth scenario

As the order in which the MPLS-CR paths between sources and destinations are created influences the paths chosen, two orders are used.

The first possibility is to create first the UDP paths for Gold and Silver classes and then the TCP paths for Gold and Silver classes. This will enable that several paths with different source-destination pairs share links in the core. This possibility was used in the adaptive algorithm.

The second possibility is for all source-destination pairs, no matter what their class or protocol, to use the same path. This will enable a better division of the bandwidth between the different classes. This possibility was used in the mathematical algorithm.

The first possibility results into a slightly worse delay and jitter for TCP flows, as they use the longest paths. Otherwise, there are no significant differences in the QoS.

3.3 *Adaptive Dynamic Bandwidth*

In the first dynamic technique, an adaptive algorithm moves bandwidth between the traffic classes, as necessary.

Figures 5, 6 and 7 show the throughput with this technique for the same situation as in figures 2, 3 and 4.

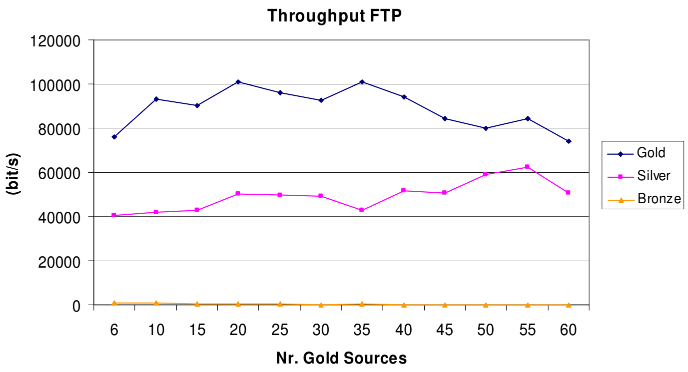

Fig. 5. FTP throughput per source for the adaptive dynamic technique

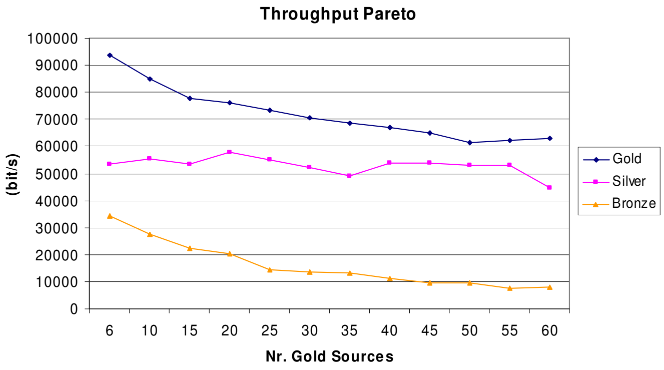

Fig. 6. Pareto throughput per source for the adaptive dynamic technique

Now there is no priority inversion for the throughput of FTP and Pareto applications, as the load increases. However, as a certain minimum bandwidth is assured to each class, the throughput variation is small as the load increases. This minimum bandwidth assured for the Silver class still causes a QoS priority inversion for high loads on the Gold class for CBR traffic. Although this priority inversion is a disadvantage from the proportional differentiation point of view, it might make sense, depending on the service being sold to users, since a minimum assured bandwidth is also of some value to users. The mathematical algorithm removes this minimum, as it does not use any steps.

For the FTP sources, the relation between the Gold throughput and the Silver throughput is 1.8 on the average, not far from the configured value of 2. But for the other types of sources, the relation is far from 2 on the average.

The other QoS parameters (end-to-end delay, jitter and packet loss) will be analyzed in section 3.5, so that all proposed techniques are evaluated together for these parameters and compared to the fixed bandwidth scenario.

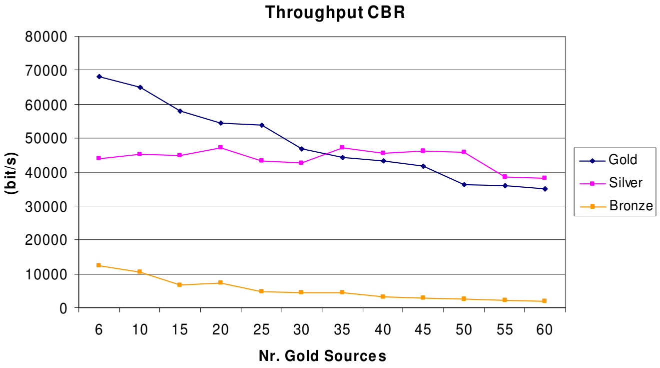

Fig. 7. CBR throughput per source for the adaptive dynamic technique

3.4 *Mathematical Dynamic Bandwidth*

In this second dynamic technique, the existing bandwidth is divided between the traffic classes according to the proposed mathematical expressions.

Figures 8, 9 and 10 show the throughput with this technique for the same situation as in figures 2, 3 and 4.

Now there is a better proportionality between the QoS in the Gold and Silver classes, with an increased throughput for low loads as compared with the adaptive algorithm.

For the FTP sources, the relation between the Gold throughput and the Silver throughput is 2.4 on the average, not far from the configured value of 2. For low loads, the Silver class is not much loaded and gets more than half the throughput of the Gold class. For high loads, there are many packet retransmissions, lowering the Silver class throughput.

For Pareto and CBR sources, the proportionality is not so easily obtained, but the Gold class always gets better QoS. For low loads, all Gold and Silver source applications can transmit all their packets, so the throughput is similar for both classes. When the load increases, the Silver class is much more affected than the

Gold class. The relation between the Gold throughput and the Silver throughput is 2.2 on the average for the Pareto sources and 1.9 for the CBR sources.

The other QoS parameters (end-to-end delay, jitter and packet loss) are analyzed in the next section.

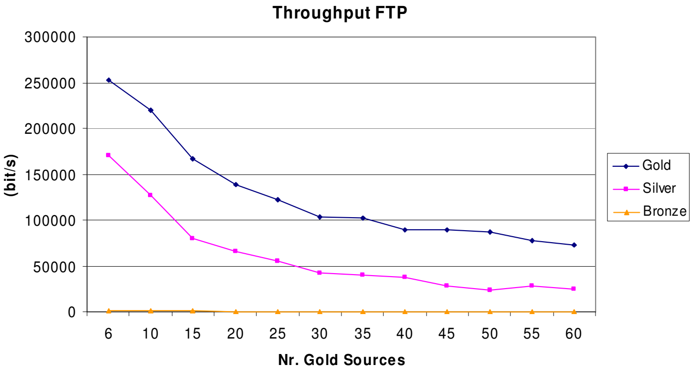

Fig. 8. FTP throughput per source for the mathematical dynamic technique

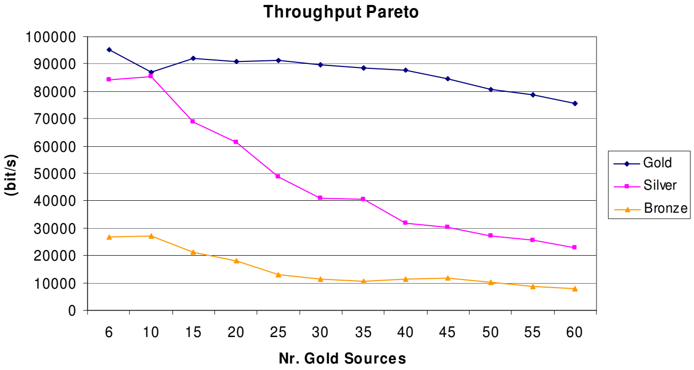

Fig. 9. Pareto throughput per source for the mathematical dynamic technique

3.5 *QoS Comparison*

Figure 11 shows the end-to-end delay averaged over the different types of user applications, for the different algorithms, for the Gold class on the left and for the Silver class on the right. The evolution of the remaining QoS parameters (jitter and packet loss) is very similar to the evolution of the delay, so they are not shown. The conclusions are the same as for the delay.

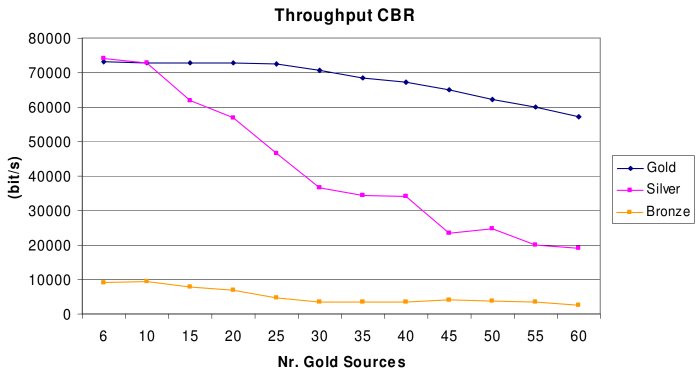

Fig. 10. CBR throughput per source for the mathematical dynamic technique

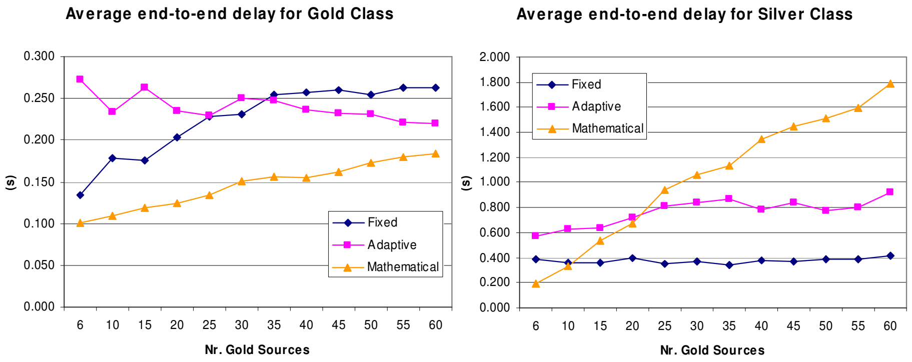

Fig. 11. Average end-to-end delay for the Gold and Silver classes

From figure 11, it can be observed that the end-to-end delay for the Gold class with the mathematical algorithm has a similar evolution as in the fixed bandwidth scenario, but with smaller values, resulting in better QoS. On the other hand, the end-to-end delay for the Silver class with the mathematical algorithm increases much faster as the load increases, when compared with the fixed bandwidth scenario. This means that the mathematical algorithm moves bandwidth from the Silver to the Gold class, increasing the Gold QoS and degrading the Silver QoS, when the load on the Gold class increases.

As regards the adaptive algorithm, the end-to-end delay for the Gold class is about constant when the load increases, while the end-to-end delay for the Silver class degrades much more slowly than for the mathematical algorithm.

The mathematical algorithm offers better QoS to the Gold class, but for high loads, the QoS of the Silver class is better for the adaptive algorithm, as it assures a minimum bandwidth to the Silver sources, independently of the load in the Gold class.

4 Conclusions and Further Work

This paper proposes and evaluates two techniques for implementing proportional differentiated services based on MPLS-CR.

The results show that with both techniques proposed it is possible to improve the QoS users get, as compared with the situation of fixed bandwidth division.

The results show that the mathematical technique makes a more uniform division of the bandwidth, according to the number of existing sources, resulting in a better overall QoS and better proportional differentiation of the throughput.

The main restriction to the mathematical approach is that the paths for the same source-destination pair need to be always the same for all the protocols and classes, otherwise the value for the TOTALbw in the equations will not make sense, since it will not correspond to one path but to several ones.

On the other hand, the adaptive algorithm does not have that kind of restriction and paths can be randomly created. However, in these simulations we had a concern which was to put a Gold and a Silver class aggregate in the same path in order to make it easier for the Gold class to get the needed bandwidth. The results also show that the difference between Gold and Silver QoS is not always proportional, but the aim is to guarantee a TpS according to the number of sources in the path. If the path is equally shared by both classes, Gold flows should usually get twice the QoS Silver flows get.

As regards the signaling traffic required for modifying the bandwidth reserved for the paths, the mathematical technique provides new values for the reservations for every flow that starts or stops, while the adaptive technique works in steps and requires fewer modifications. The development of a technique to keep the signaling traffic within a certain limit, say 1% of the network capacity, was left for further study. The removal of some simplifications was also left for further study. These include: the use of other network topologies; the simultaneous use of sources with different QoS requirements; the use of different proportionality factors between the Gold and Silver bandwidths; and the use of multiple paths for each source-destination pair.

References

[1] S. Giordano, S. Salsano, S. Van den Berghe, G. Ventre, D. Giannakopoulos, "Advanced QoS Provisioning in IP Networks: The European Premium IP Projects", *IEEE Communications Magazine*, **41**(1):30-36, January 2003.

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

[3] B. Jamoussi, L. Andersson, R. Callon, R. Dantu, L. Wu, P. Doolan, T. Worster, N. Feldman, A. Fredette, M. Girish, E. Gray, J. Heinanen, T. Kilty, A. Malis, "Constraint-Based LSP Setup using LDP", IETF RFC 3212, January 2002.

[4] J. Heinamen, F. Baker, W. Weiss, J. Wroclawski, "Assured Forwarding PHB Group", IETF RFC 2597, June 1999.

[5] Y. Chen, C. Qiao, M. Hamdi, D. Tsang, "Proportional Differentiation: A Scalable QoS Approach", *IEEE Communications Magazine*, **41**(6):52-58, June 2003.

[6] P. Pereira, P. Lepera, A. Casaca, "DiffServ Traffic Management with MPLS", 5th Conference on Telecommunications, ConfTele'2005, Tomar, Portugal, April 2005. ISBN: 972-98115-9-8.

[7] UCB/LBNL/VINT Network Simulator (version 2). http://www.isi.edu/nsnam/ns/