Deployment of a DiffServ-based Priority Service in a MAN/WAN Environment

Christos Bouras[1,2], Afrodite Sevasti[1,2]
[1] *Research Academic Computer Technology Institute-RACTI, Kolokotroni 3, 26221 Patras, Greece*
[2] *Dept. of Computer Engineering and Informatics, University of Patras, 26500, Patras, Greece*
{bouras, sevastia}@cti.gr

Abstract

The increased need for Quality of Service (QoS) in today's networks has concentrated a lot of research and implementation efforts. Particular interest exists for the exploitation of the DiffServ framework towards the provision of qualitative services in MANs and WANs worldwide. Although individual mechanisms have been widely investigated, not many references to studies combining all levels of priority and QoS provisioning exist. This work aims at presenting a thorough and unified approach to the dimensioning and provisioning of a high priority service for high-quality demanding traffic over a MAN/WAN environment. Our approach anticipates for bounds in specific quality metrics and is demonstrated by the experimental case-study of a high-priority service over a MAN topology.

1. Introduction

The DiffServ framework has been designed for the provision of QoS services in large-scale environments, where the extensive aggregation of flows does not allow solutions for QoS provisioning on a per-flow scale. At the same time, DiffServ is a framework and as such, it provides guidelines rather than strictly defined service models. Network designers have, thus, at their disposal multiple different individual mechanisms and alternatives for QoS services' implementation. Our approach attempts to define a set of guidelines for the provision of a high-priority service to QoS demanding traffic over a DiffServ-enabled domain, while at the same time ensuring deterministic upper bounds to critical metrics, such as one-way delay.

Dimensioning and provisioning QoS on the DiffServ framework basis has been the topic of many research initiatives in the last years. Some of them focus in DiffServ-based QoS service models and definitions ([5],[7],[10]) while others perform experimental studies on individual mechanisms and modules of DiffServ-based QoS services ([6],[8],[9]). Our approach aims at combining both a service definition and a set of mechanisms, towards an integrated solution for QoS provisioning to IP multimedia traffic, with bounded quality guarantees.

With respect to the related work done so far, the contribution of this work is two-fold: we initially provide a model for dimensioning and provisioning a high-priority service over a backbone topology. We use a simulation framework in order to verify our proposed provisioning model through a case study. In section 2 of this paper, the model used for both the analytical and simulation work is provided. In section 3, the analytical approach of DiffServ-based QoS service is provided. In section 4 a set of scenarios verifying the analysis of section 2 and demonstrating the performance of the service are provided. This paper concludes with our intended future work on this topic and conclusions drawn from the work carried out so far.

2. The model

The case that will be further investigated in this work is that of dimensioning and providing a high-priority, low latency QoS service for the aggregated IP multimedia traffic entering a MAN/WAN. This service, referred to as 'Gold' service from now on, is built according to the Expedited Forwarding Per-Hop-Behaviour ([1]) of the DiffServ framework. The Gold service is provided to the transport networks attached to a MAN/WAN, which will be referred to as 'customers' of the service. The Gold service aims at offering the equivalent of an end-to-end virtual leased line (VLL) service at the IP layer across multiple domains. In general, traffic requesting for such a service is sensitive to delay, jitter and packet losses, so the proposed service and provisioning method will be evaluated against these metrics. The customers are modeled as sources of Gold traffic. Between each of the customers (C_i) and the MAN/WAN there exists a Service Level Agreement (SLA) that specifies the characteristics (traffic envelope) of the marked as Gold traffic injected by C_i into the MAN/WAN. Also, apart from guaranteeing a specific rate of service for each aggregate, the Gold service guarantees a specific bounded end-to-end delay (D), bounded jitter and minimal packet losses to all the legitimate aggregates served by it. Legitimate aggregates are those the traffic envelope of which is examined (and enforced if necessary) by the use of a policer for each one of them.

One of the most common policers used and supported by commercial network equipment is that of a token bucket (r, b), that imposes conformance to an average rate and maximum burst size to the incoming traffic. After Gold aggregates from the different customers are policed, they are further aggregated and served as a unified aggregate, inserted to the ingress node's queue used for scheduling of Gold traffic and served through the domain with no more discrimination and policing.

3. Analytical study

For the Gold service provisioning, it is proposed to use an appropriate policing profile for each aggregate of Gold traffic injected by each customer attached to the MAN/WAN. It is also proposed to use appropriate configuration of all the routers comprising the MAN/WAN to ensure the existence of proper buffering space and capacity in order for the Gold traffic to be served with the required quality.

The basic principles on which the service dimensioning is based are:

- The configuration of the MAN/WAN routers so that Gold traffic will receive at least a strictly bounded minimum service rate at each router, regardless of the load imposed on the router
- The policing of the total of Gold traffic entering each Point-of-Presence (PoP) of the MAN/WAN is such that the arrival rate of the aggregated Gold traffic entering the PoP router, created by the merging of all Gold aggregates of customers attached to a specific PoP does not exceed the minimum rate of service that Gold traffic receives at each egress interface of the router (according to [1]).

Based on these principles, the goal of Gold service provisioning would be:

- To assure that each flow participating in a Gold aggregate originating from a customer attached to a PoP demonstrates a throughput equal or larger than that of entering the MAN/WAN as it exits.
- Delay and jitter for the packets of this flow are both bounded between the points where the packets enter and exit the MAN/WAN, while the flow perceives negligible or zero packet losses

Within the notion of Gold service, providing strict guarantees of delay and jitter in an end-to-end fashion is essential and therefore the service provided will be provisioned on an aggregate basis but evaluated on a per flow basis.

3.1. Policing profile

In every PoP of the MAN/WAN it is proposed to configure appropriate token bucket policers (supported by major vendors' routers e.g. Cisco's Committed Access Rate) for the Gold traffic aggregate of each customer attached to the specific PoP. The service definition proposes that the MAN/WAN could serve one or two different Gold aggregates from each customer.

The alternatives are:

- To provide each customer with the ability to inject a unique aggregate of Gold traffic to a PoP to which he is attached. This aggregate should at any moment obey to a traffic profile of token bucket type denoted by (r_i, b_i).
- To provide each customer with the ability to inject two aggregates of Gold traffic to the PoP to which he is attached. One of them will comprise solely by jitter sensitive (e.g. VoIP) traffic while the other will concentrate all the rest of IP multimedia traffic that the customer wishes to receive Gold treatment. Each of these aggregates should at any moment obey to a traffic profile of token bucket type denoted by (r_i^{voice}, b_i^{voice}) and (r_i, b_i) correspondingly.

In order for the determination of the values for the parameters r_i, b_i, an analytical approach will initially be attempted.

3.1.1. Capacity dimensioning Gold service provisioning is based on resources' over-provisioning and therefore the provision of capacity guarantees can be obtained by ensuring that for the total capacity C_q appointed to each queue serving Gold traffic (e.g. through the quantum configured for it) it holds that:

$$C_q \geq \sum_i r_i \Rightarrow \sum_i r_i \leq kC_q, \quad \text{with } k \in (0,1) \qquad (1)$$

where r_i is the rate of the token bucket profile for aggregate i served through this specific queue, with i enumerating all these possible aggregates. Factor k determines the over-provisioning factor for the implementation of the Gold service on a particular interface of a router.

If the MAN/WAN chooses to implement a 'destination-unaware' scheme for the Gold service, this means that each interface in the MAN/WAN must be able to support with enough over-provisioning the maximal sum of Gold traffic aggregates that are possible to be served through this particular interface. Otherwise, it is possible to dimension the service, provisioning at all interfaces with the resources required for the specific sum of Gold traffic that according to the destination-aware data will cross the interface. In the analysis to follow, the destination-unaware option is adopted as this choice ensures minimal administrative overhead.

In order to keep our analysis simple, we assume that all customers attached to a MAN/WAN that has to be configured for Gold service provisioning have an equal connection speed. That is

$$r_i = l \times r_{access}, \forall i \text{ with } l \in (0,1) \quad (2)$$

where r_{access} is the capacity of the interface through which each customer is attached to the MAN/WAN for each customer C_i and r_i the token bucket profile rate with which the Gold aggregate injected by customer i to the MAN/WAN is policed. Equation (2) sets the prerequisite for the capacity of the aggregated Gold traffic that each customer i injects to the MAN/WAN not to exceed the $l\%$ of the capacity of the transmission link with which the customer is attached to the MAN/WAN.

Of course, under normal operating conditions, differences in the demand for Gold capacity can occur among different customers, so that

$$r_i = l_i \times r_{access} \quad (3)$$

denoting that each customer will devote a different percentage of his access capacity to the MAN/WAN to Gold traffic. According to (1), the sum of l_i's cannot exceed a value calculated as follows:

$$\sum_i r_i \le kC_q \Rightarrow \sum_i l_i \times r_{access} \le kC_q$$
$$\Rightarrow \sum_i l_i \le \frac{kC_q}{r_{access}} \quad (4)$$

Assume that S_q is the set of customers that are able to inject a Gold traffic aggregate through router queue q. If q has a minimum guaranteed capacity equal to C_q, there is also a limitation to the value of C_q, which should not exceed a percentage of the capacity of the MAN/WAN backbone links. This means that

$$C_q \le m \times C, \text{ with } m \in (0,1) \quad (5)$$

so that under heavy load, best effort traffic will not face denial of service. From (1),(2) and (6) it follows that

$$\sum_{S_q} l \times r_{access} \le k \times C_q \le k \times m \times C \Rightarrow$$
$$l \times S_q \times r_{access} \le k \times m \times C \Rightarrow$$
$$l \le \frac{m \times C}{S_q \times r_{access}} \times k \quad (6)$$

Thus, the percentage of Gold capacity on their access links to the MAN/WAN is limited for each customer according to k, the over-provisioning factor of Gold traffic on the MAN/WAN backbone. The experimental analysis that will follow demonstrates how k affects perceived quality by the IP multimedia traffic served with the Gold service. Recommendations from related work ([11],[12],[13]) lead to the selection of small values for k, in the range of (0.05, 0.2).

3.1.2. Burstiness dimensioning. The quality demands and the over-provisioning nature of the Gold service impose restrictions on the allowed maximum burst size b_i for each customer's Gold traffic aggregate. At the same time, bounded end-to-end delay can be ensured for the whole of Gold traffic crossing the MAN/WAN by selecting the b_i values for each aggregate i served by a MAN/WAN router so that the worst-case delay that Gold traffic packets will face during their stay in each router will not exceed a maximum value $D_{\max}$. Accordingly, the guaranteed bounded delay that a packet can face from the moment it enters a MAN/WAN PoP until it exits the MAN/WAN is denoted by:

$$D_{tot} \le \sum_i D_{\max} + \sum_k D_{prop\,k} + \sum_k D_{trans_k} \quad (7)$$

where D_{trans_k} is the transmission delay and D_{prop_k} is the propagation delay of the packet on transmission line k. The following analysis deals with the value of $D_{\max}$, since the values of D_{trans_k} and D_{prop_k} cannot be affected by the implementation of a service over existent equipment/transmission lines. There is a common belief that ensuring low utilization of the capacity reserved for Gold traffic over a backbone link achieves low values of the delay perceived. However, this has been proved to not always hold ([11]).

The maximum value $D_{\max}$ for a packet can be considered to be equal to:

$$D_{\max} = D_{head}^{\max} + D_{queue}^{\max} \quad (8)$$

where $D_{head}^{\max}$ is the maximum delay of a packet on the top of the queue that serves the Gold traffic, ready to be placed on the transmission medium and $D_{queue}^{\max}$ the maximum delay of a packet located in lower positions of the Gold traffic queue before it reaches the top of the queue.

Depending on the scheduling mechanism used on the equipment of the MAN/WAN, the value of $D_{queue}^{\max}$ can be calculated by taking into consideration the maximum number of packets that a packet entering the Gold traffic queue at a router's egress interface finds already there. This number is upper bounded by the number of that equals the sum of bursts of all Gold aggregates using this queue. In any case, with the prerequisite that the corresponding r_i values are determined according to the analysis already made, $D_{queue}^{\max}$ can be expressed as:

$$D_{queue}^{\max} = \frac{f(\sum_i b_i)}{C} \qquad (9)$$

where i enumerates all the customers entitled to inject Gold traffic to the MAN/WAN, each with a token bucket profile of depth equal to b_i and the function f depending on the method of scheduling used at the MAN/WAN routers. We claim that

$$f(\sum_i b_i) < \sum_i b_i \qquad (10)$$

and therefore it is obvious how the selection of b_i values for the policing of each Gold aggregate injected by a customer to a MAN/WAN PoP can affect the delay guarantees provided by the MAN/WAN to its Gold service customers. Inversely, provided that specific need for bounded delay exists, e.g. 40ms for VoIP traffic crossing the MAN/WAN, one can determine the sum of the b_i values that can be offered to SLAs with customers.

4. Experimental study

The experiments carried out aim at investigating the Gold service provisioning on a backbone link of a MAN/WAN, located between two PoPs to which a number of customers with different access capacities are attached. The specific topology used for the experiments is depicted in Figure 1. It is obvious from Figure 1 that the experiments were performed for three different customers (C_1, C_2, C_3) with access links of differing capacities. Gold traffic on the backbone link co-exists with background traffic served as best effort and inserted to the topology as cross-traffic ([2],[3]). Background traffic, as depicted in Figure 1, has average throughput in the range of 85%-100% of the backbone link capacity. Moreover, each customer injects through his access link to PoP0 not only Gold traffic but also background traffic that fills up the unused by Gold traffic capacity of the access link.

More details about the simulated Gold and background traffic used in the following experiments is provided in [14]. For most of the experiments that will be presented the mixture of Gold traffic presented in Table 1 was injected from customers C_1, C_2, C_3 in PoP0. Column 4 of Table 1 accounts for l_i in (3). During the experiments jitter measurements were made according to [4] and were evaluated separately for each different type of traffic, since the effect of jitter differs according to the traffic type to which it is imposed.

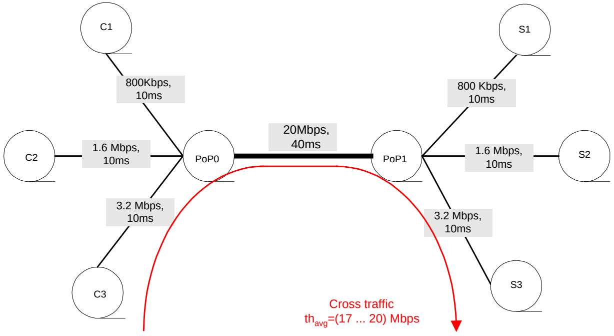

Figure 1. Experimental topology

In column 5 of Table 1, the token bucket profiles that were selected for the Gold traffic mix used in the experiments are presented. PoP0 and PoP1 were initially configured in such a way so as to be able to serve with high priority the total of Gold traffic that can possibly be injected to the simulated topology by C_1, C_2, C_3. The scheduling mechanism simulated at the MAN/WAN routers was that of Modified Deficit Round Robin of Cisco GSRs.

Table 1. Synthesis of Gold traffic injected by the customers of the simulated topology

	Access link capacity	**Gold traffic**	l_i	(r_i, b_i)
C_1	0.8Mbps	2 VoIP flows	20%	160 Kbps, 378 bytes
C_2	1.6Mbps	1 MPEG video flow	20.8%	1.3 Mbps, 10700 bytes
C_3	3.2Mbps	1 H.323 flow 2 voice flows	19%	850 Kbps, 1870 bytes

4.1. Baseline measurements

This first series of experiments concerns the comparison of the performance of the strict priority MDRR scheduling (MDRR-STR) when compared to that of best-effort service. For these experiments best-effort service was simulated by FIFO scheduling for both Gold and background traffic on PoP0 and PoP1. In Table 2 the losses imposed by best-effort FIFO scheduling used for both Gold and background traffic are shown. With the introduction of MDRR-STR scheduling, losses were equal to zero.

Table 2. Gold traffic packet losses during FIFO scheduling

Traffic aggregate	**Packets lost/ packets transmitted**	**Packet loss percentage**
Voice C1->S1	1633/5377	30%
Voice C3->S3	1067/5978	17.8%
MPEG (C2->S2)	34657/38210	90.4%
H.323 (C3->S3)	3952/5461	72%

During the experiments it was observed how the two types of scheduling affected the traffic pattern of aggregated VoIP traffic. FIFO scheduling decreased the throughput and changed the temporal characteristics of the VoIP aggregate. MDRR-STR scheduling preserved the VoIP aggregate pattern demonstrating a high level of quality perceived by the end user.

Table 3. Quality metrics' values for MDRR-STR scheduling of Gold traffic

Traffic aggregate	**max delay (ms)**	**jitter (ms)**
Voice C1->S1	73.98	10.14 (1.9)
Voice C3->S3	70.85	2.5 (0.49)
MPEG (C2->S2)	67.04	4.9 (0.5)
H.323 (C3->S3)	70.85	2.5 (0.49)

Table 3 displays the values measured for the quality metrics of end-to-end delay and jitter under MDRR-STR scheduling. It is important to stress out how, differences in the capacities of the interfaces in PoP1 (e.g. PoP1->S1 and PoP1->S3) differentiate the maximum end-to-end delay and jitter measured for equivalent Gold traffic aggregates (Voice C1->S1, Voice C3->S3 correspondingly) served through these interfaces and thus, confirming formula (9).

4.2. Interaction of gold and background traffic

The experiments presented in this section deal with the effect of increasing the load imposed by background traffic to the performance of MDRR-STR scheduling as well as the quality perceived by Gold traffic. Two experiments were carried out with the throughput of background traffic equal to 85% and 100% of the backbone link's capacity.

Table 4 demonstrates the throughput of Gold traffic for the two different cases. It is obvious how MDRR-STR scheduling protects Gold traffic and also how Gold traffic seems to gain a portion of the available on the background link capacity in the 100% background load case.

Table 4. The effects of background traffic's increased load in Gold traffic

Interface	**Gold traffic mean throughput (85% background load)**	**Gold traffic mean throughput (100% background load)**
PoP0	1006	1090
PoP1->S1	52	56
PoP1->S2	436	502
PoP1->S3	518	531

Figure 2 demonstrates how the quality metrics in the two cases are affected. The quality provided is slightly worse as the background traffic load increases, by the shifting of delay and jitter distribution curves to the right.

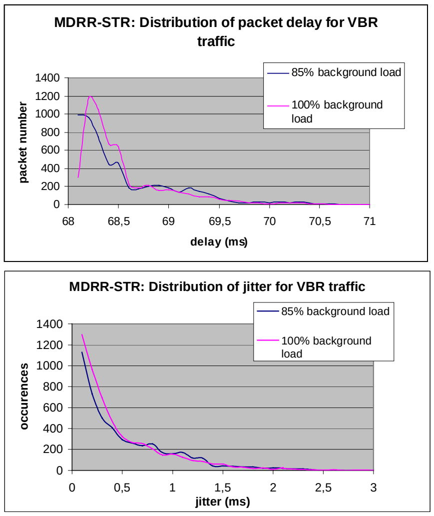

Figure 2. Delay & jitter perceived by VBR traffic

4.2.1. Different scenarios' testing

In these experiments, comparison of the perceived by Gold traffic quality was made in the two different cases described in section 3.1. PoP0 and PoP1 were configured with MDRR-STR scheduling so that:

- Case 1: The total of Gold traffic was served by the strict priority queue
- Case 2: the VoIP flows were served by the strict priority queue and the rest of Gold traffic was served by one of the rest of the queues available with a high queue weight value (97), ensuring access to 97% of the available bandwidth in the absence of strict priority traffic

The experiments were performed measuring both the throughput of the total Gold traffic at each PoP but also separately for the Gold VoIP traffic and the rest of Gold traffic at each PoP. The total throughput of Gold traffic was increased in case 2, while the LLQ queue 'protected' VoIP traffic. More specifically, VoIP traffic reached the maximum rate of 320Kbps at PoP0, which is equivalent to the accumulation of peaks (80Kbps) of the four VoIP flows transmitted through the PoP). This observation holds for interfaces PoP1-S1 (2 VoIP flows -165 Kbps maximum throughput) and PoP1-S3.

Case 1

Traffic aggregate	**max delay (ms)**	**max jitter (ms)**
Voice C1->S1	73.98	10.14 (1.9)

Voice C3->S3	70.85	2.5 (0.49)
MPEG	67.04	4.9 (0.5)
H.323	70.85	2.5 (0.49)

Case 2

Voice C1->S1	74.05	10.19 (1.94)
Voice C3->S3	65.29	4.2 (1.07)
MPEG	71.964	9.07 (0.713)
H.323	74.83	6.15 (0.9)

Table 5. Quality metrics' values for cases 1 & 2

Table 5 shows the quality metrics' values for the two cases. End-to-end delay perceived by VoIP traffic was improved in case 2. As far as the other types of Gold traffic are concerned, a slight decrease in the quality perceived was noticed. This confirms the theoretical expectations, since this traffic is transferred from the strict priority queue to a regular queue of service.

5. Future work – Conclusions

Our future work on the Gold service will concentrate on providing stricter bounds for the quality metrics of our analytical approach, based on related work ([11]). We also intend to deploy and evaluate an algorithm for the dimensioning of the Gold service that, according to the analysis made here, will be useful for network administrators in order to introduce the service to their networks. At the same time, we intend to deploy and test a provisioning and management infrastructure that will regulate the provisioning of Gold capacity and traffic profiles automatically.

The work presented in this paper aims at providing guidelines for the provisioning of high-quality service to IP multimedia traffic over MAN/WANs of today. We provide guidelines for dimensioning the service and estimating worst-case bounds on quality metrics. We have also developed a thorough simulation environment that can be reliably used to simulate high quality service provision to IP multimedia traffic and observed the quality perceived as the dimensioning or service parameters change.

For the specific case-study, we have used the analytical approach to dimension the Gold service in a simulated topology. The elimination of packet losses and the significant improvement of the quality metrics account for a successful provisioning. Although our work attempts to provide a number of guidelines, more work is required in the direction of studying and 'automating' the dimensioning and configuration of DiffServ-based services such as the Gold service in contemporary MANs/WANs. These efforts will be essential in establishing such services and bringing forward their beneficial aspects.

6. References

[1] V. Jacobson et al., 'An Expedited Forwarding PHB', IETF RFC 2598, June 1999

[2] C. Dovrolis, D. Stiliadis and P. Ramanathan, 'Proportional Differentiated Services: Delay Differentiation and packet Scheduling', IEEE/ACM Transactions in Networking, February 2002

[3] M. Yuksel, B. Sikdar, K.S. Vastola, and B. Szymanski, 'Workload generation for ns Simulations of Wide Area Networks and the Internet', In Proc. of CNDS 2000, San Diego, CA, 2000

[4] IETF, 'Instantaneous Packet Delay Variation Metric for IPPM', found at: http://www.ietf.org/internet-drafts/

[5] M. Goyal, A. Durresi, C. Liu and R. Jain, 'Performance Analysis of Assured Forwarding', Internet Draft (draft-goyal-diffserv-afstdy-00), February 2000

[6] I. Yeom, N. Reddy, 'Impact of marking strategy on aggregated flows in a differentiated services network', proceedings of IWQoS workshop, June 1999

[7] S. Shalunov, B. Teitelbaum, 'Qbone Scavenger Service (QBSS) Definition', Internet2 Technical report, Proposed Service Definition, March 16, 2001

[8] M. May, J-C. Bolot, A. Jean-Marie and C. Diot, 'Simple performance models of tagging schemes for service differentiation in the Internet', In Proc. of Infocom '99, New York, NY, 1999

[9] TF-TANT Task Force, Tiziana Ferrari, Editor, 'Differentiated Services: Experiment report', Phase 2, May 15, 2000

[10] M. Campanella, 'Implementation Architecture specification for the Premium IP service', Deliverable D2.1-Addendum 1, SEQUIN Project (IST-1999-20841)

[11] Y. Le Boudec and P. Thiran, 'Network Calculus: A Theory of Deterministic Queuing Systems for the Internet, LNCS 2050, Springer Verlag, 2001

[12] A. Charny and J.-Y. Le Boudec, 'Delay bounds in a network with aggregate scheduling', in Proc. of QofIS 2000, Germany, 2000

[13] C. Bouras, M. Campanella, M. Przybylski, A. Sevasti, 'QoS and SLA aspects across multiple management domains: The SEQUIN approach', to appear in Journal of Future Generation Computer Systems 939, Elsevier Publishing, 2002, pp. 1-15

[14] C. Bouras, D. Primpas, A. Sevasti, A. Varnavas, 'Enhancing the DiffServ architecture of a simulation environment', In Proc. of DS-RT 2002, Texas, USA, October 11-13, 2002