Techniques for DiffServ-based QoS in Hierarchically Federated MAN Networks – the GRNET Case

Angelos P. Varvitsiotis, Vasilios A. Siris, Dimitris N. Primpas,
Georgios I. Fotiadis, Athanassios Ch. Liakopoulos, and Christos J. Bouras

Abstract—**DiffServ is the basis of contemporary QoS-enabled networks. Setting up DiffServ QoS requires extensive engineering effort in dimensioning and provisioning, especially for adjacent networks under different administrations linked in a "federated" hierarchy. In this paper we present a case study for QoS techniques employed in the GRNET MAN networks of Athens and Crete. After introducing the supported QoS mechanisms and service types, we discuss our dimensioning methodology and present two algorithms for worst-case dimensioning. We explain the provisioning mechanisms of GRNET and we present in brief our new automated provisioning ANS tool. Finally, we deal with the extension of our mechanisms and tools in hierarchically federated networks and give some future directions of our work.**

Index Terms—**QoS dimensioning, provisioning, federated QoS.**

I. INTRODUCTION

QUALITY of service (QoS) is a crucial ingredient of today's multi-service packet networks. QoS-enabled networks can accommodate simultaneously various differing traffic types, such as data, voice, and video, by handling time-critical traffic appropriately at congestion points. DiffServ [1] is becoming the prevalent QoS architecture in today's IP-based packet networks. DiffServ introduces the concept of traffic classes, where each traffic class is mapped to a Per-Hop Behavior (PHB). PHBs are implemented at routers by means of queuing and scheduling at congestion points, where queues are formed; thus, by mapping different traffic types into different PHBs, routers are able to ensure service guarantees.

A. P. Varvitsiotis (e-mail: avarvit@grnet.gr) and A. C. Liakopoulos (e-mail: aliako@grnet.gr) are with the Greek Research and Technology Network, Mesogeion 56, Athens, GR 115 27 (phone: +30-210-7474254; fax: +30-210-7474490).

V. A. Siris (e-mail: vsiris@ics.forth.gr) and G. I. Fotiadis (e-mail: fotiadis@ics.forth.gr) are with the Institute of Computer Science, FORTH, and the Dept. of Computer Science, University of Crete, P.O. Box 1385, Heraklion, Crete, GR 711 10 (phone: +30-2810-391726, 393524, Fax: +30 2810 391601).

Ch. J. Bouras (e-mail: bouras@cti.gr) and D. N. Primpas (e-mail: primpas@cti.gr) are with the Research Academic Computer Technology Institute and Computer Engineering and Informatics Department of University of Patras, 61 Riga Feraiou Str., Patras, GR 262 21 (phone: +30-2610-960375; fax: +30-2610-969016).

In broadband networks, congestion does not necessarily occur at the edge of the network (the link interconnecting the subscriber to the network core): congestion is equally likely to occur at the edge and in the core of the network. A common congestion cause in broadband networks is capacity mismatch in different parts of the network core. For example, a MAN may have capacities ranging from 10 Gbps to a few Mbps on low-speed DSL links. Moreover, while network routers are ideal for implementing DiffServ PHBs at capacity mismatch points, MAN networks are increasingly based on Layer-2 switches. Inexpensive as they may be, Layer-2 switches complicate matters, because capacity mismatches now occur beyond the Layer-3 DiffServ domain. Even when a hybrid QoS scheme, e.g., a translation from DiffServ to 802.1p [9] is used, Layer-2 switches are usually not as versatile as routers in the implementation of various PHBs.

Another difficulty in broadband MAN networks stems from the fact that the aggregate traffic from a handful of high-speed subscribers can very easily exhaust the available bandwidth in the core of the network. This calls for protection measures at the network perimeter. Engineering QoS at the network perimeter involves (a) service *dimensioning*, where a fair scheme has to be adopted that satisfies a set of demand scenarios and (b) service *provisioning*, where mechanisms must be provided to set up network and service protection measures according to specific subscriber needs.

When independently managed networks are interconnected, additional difficulties exist in ensuring interoperability of DiffServ-based QoS across network boundaries. This involves (a) adopting interoperable conventions about DiffServ traffic classes and the respective PHBs, (b) adopting interoperable dimensioning and provisioning mechanisms and (c) linking together functions of the two domains, such as provisioning, policing and admission control functions.

In this paper we present a case study of the Athens and Crete MANs of GRNET [2] and the techniques involved in QoS services in GRNET. In section II we explain the service types and implementation mechanisms in use today. In section III we discuss the QoS dimensioning issues that we have faced and we present two algorithms for worst-case dimensioning; we then discuss service provisioning and we present briefly a new tool that we have developed in GRNET

for automated QoS provisioning. In section IV we discuss how the above fit in a landscape where independently managed neighbor networks interoperate using compatible – but not necessarily identical – DiffServ-based QoS services. To this end, we explore the hierarchical structure that is formed by subscriber networks and GRNET – for which we toss the term *hierarchically-federated networks* – to achieve QoS service compatibility across the constituent networks. In section V we provide some numerical data and usage results; finally, in sections VI and VII we position our work with respect to pre-existing related efforts and we provide some pointers to future work.

II. THE NETWORK AND THE QOS SERVICE: DEFINITIONS AND MECHANISMS

GRNET is the Greek National Research and Education Network (NREN). GRNET is a mixed IP- and Ethernet-based network, operating at Gigabit speeds. Together with the high-speed LANs of its subscribers (universities and research institutes) and the European academic and research backbone, Géant [3][1], GRNET forms a set of hierarchically-federated networks.
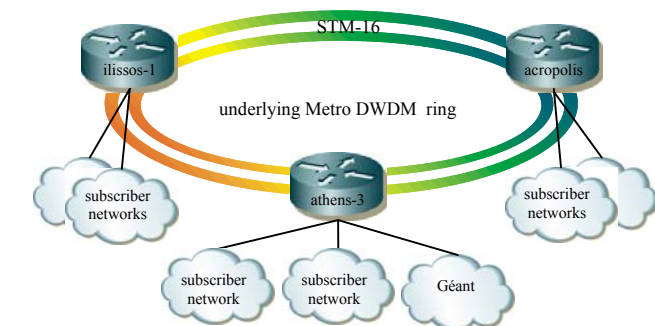
(a) The Athens L-3 (router-based) MAN
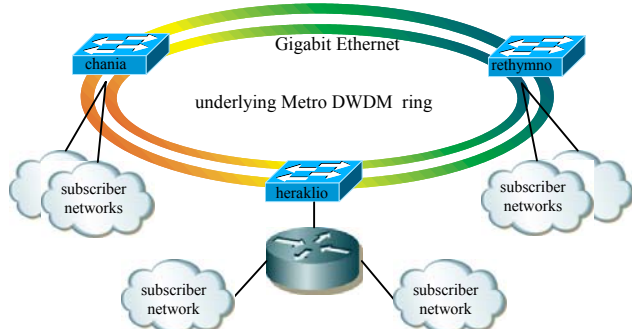
(b) The Crete L-2 (switch-based) MAN

Fig. 1. The two Metro-area networks of GRNET.

The topology of the two major MANs of GRNET is shown in Fig. 1(a) and (b). As seen in the figures, the Athens MAN is router-based, whereas the newer Crete MAN is switch based, with a router in the main aggregation site (Heraklio). Both networks are built on unprotected DWDM rings; the Athens MAN uses STM-16 *lambdas*, whereas the Crete MAN operates on 1-Gigabit Ethernet *lambdas*. Géant is shown in

Fig. 1 (a) as a subscriber of the Athens MAN.

The QoS services of GRNET are based on the traffic classes and the queuing mechanisms defined in this section.

A. Traffic Classes and Per-Hop Behaviors

In order to be compliant to the Géant QoS implementation, GRNET has adopted three basic traffic classes as explained below:
1) The *Premium* class, based on the EF PHB [4], is given absolute priority over any other class; very low delay/jitter plus negligible packet loss guarantees are provided for this class.
2) The *Best-Effort* class is forwarded as its name implies, in a best-effort manner; no guarantees exist for this class.
3) The *Less-Than-Best-Effort* class is treated less favorably than the *Best-Effort* class and is intended for specific scavenger applications.[2]

GRNET supports two more variants of the *Premium* class, the *Premium Transparent* class and the *VoIP* class. The *Premium Transparent* class is handled as *Premium* traffic by GRNET but as *Best-Effort* traffic by Géant. The *VoIP* class is handled like the *Premium* class by GRNET and is used by convention for voice-over-IP traffic.

B. Queuing and Policing Implementation

Following the guidelines of [5], GRNET has chosen to implement a very simple queuing model. Queuing is applied in the outgoing direction at all router interfaces, both in the network core and at the edge; no queuing is applied in the incoming direction. The underlying implementation mechanism is Modified Deficit Round-Robin (MDRR [6], [7]). Using the priority queue of MDRR, the *Premium* class is given absolute priority. A second queue, with a very low bandwidth reservation (1% of the link's capacity), is used to implement the *Less-Than-Best-Effort* class. Finally, default FCFS or RED queuing is used for the *Best-Effort* class.

Policing is used to protect the network from excess *Premium* traffic. Thanks to our dimensioning methodology presented in section III.A, policing is not needed in the network core. Moreover, policing is not needed in the outgoing direction of edge interfaces. Thus, all policing is applied to the incoming traffic at the network edge. As part of policing, out-of-profile *Premium* traffic is either re-colored to *Best Effort* or dropped at the subscriber's choice.

The actual queuing and policing configuration is generated automatically, with the aid of our web-based ANS tool, which we present in section III.B.

III. SERVICE DIMENSIONING, PROVISIONING AND ENFORCEMENT

A. Service Dimensioning

Given our "binary" queuing scheme explained in section

---

[1] In this paper, Géant is considered as just another subscriber of the Athens MAN.

[2] Because the *Less-Than-Best-Effort* class is a "yielding" (non-competing) class, in the following discussion about dimensioning we omit that class and consider only the *Best-Effort* and the *Premium* class (plus the latter's variants).

II.B (priority versus ordinary traffic), we can characterize the dimensioning of the QoS service by means of the maximum priority-traffic load share a_l on each link l in the set L of core links, $0 < a_l < 1$. By keeping a_l below a certain limit, we can provide maximum delay and jitter guarantees for the *Premium* family of services. However, our only tool for restricting a_l in the core is the police function at the network perimeter. Thus, service dimensioning essentially amounts to (a) devising a consistent *Premium* bandwidth allocation policy for all subscriber networks and (b) expressing this policy in terms of quantitative parameters for the perimeter police functions.

Given the above targets, we formulate the service dimensioning problem as follows. For the sake of simplicity, let us fix $a_l = a \ \forall \ l \in L$, $0 < a < 1$, and let us refer to a with the term *core maximum*. Let $C = \{c\}$ be the set of all subscriber networks c and let a *dimensioning policy* $A = \{a_c\}$ be the set of the maximum *Premium* traffic shares allowed on each subscriber link, so that if b_c denotes the bandwidth of link c, the maximum rate of *Premium* traffic that subscriber c is allowed to send to the network is $a_c \cdot b_c$.[3] Let h_l represent the maximum *Premium* bandwidth that can be aggregated over a core link $l \in L$ under the policy A. We then define A to be *worst-case-feasible* with respect to the core maximum a if $h_l \leq a \cdot b_l \ \forall \ l \in L$. We further define A to be *fair* if there is a function $f(\cdot)$ such that $a_c = f(b_c) \ \forall \ c \in C$ (in other words, if all subscriber links with equal bandwidth are entitled to equal amounts of *Premium* bandwidth). The core network topology G, the subscriber set C, $B_C = \{b_c\}$, $B_L = \{b_l\}$, a and f (so that $A = \{f(b_c)\}$) form the problem's input. We define the service dimensioning problem as that of answering whether A is worst-case-feasible with respect to a. In the Appendix, we provide two equivalent algorithms to solve this problem.

Some intuition may help understanding the above problem formulation: a is chosen by the service planner according to the desired service guarantees. On the other hand, f is chosen with pragmatic criteria.[4] For example, in GRNET, we chose f to be as shown in the plot of Fig. 3 and set an upper bound of 0.2 for a; we then verified for the worst-case h_l that $h_l \leq a \cdot b_l$.

B. Service Types and Provisioning

Having calculated A, we can provision subscriber links after subscribers' requests. We support requests for two *service types*: the first type is a circuit-like subscriber-to-subscriber service, where both subscriber end-networks and the necessary bandwidth allocation are specified in the service request; the second type is a subscriber-to-network service, where only one end-network and its bandwidth are specified. We call the first service type *Source- and Destination-Aware* (SDA) and the second one *Source-Aware-Destination-*

Unaware (SADU).[5] The SDA service type accommodates the exchange of high-priority traffic between two subscriber networks, whereas SADU service type is best suited for IP telephony or arbitrary point-to-point videoconference traffic (by convention, GRNET uses the *Premium* class for the first type of service and the *VoIP* class for the second one). It is important to understand that the SADU model requires that the subscriber networks contribute an admission control function, as explained further in section III.D.

Fig. 2. GRNET ANS tool: *Premium* service request form (SDA service).

Provisioning is accomplished by means of a web-based tool developed by GRNET, which we call the Advanced Network Services (ANS) tool [8]. The ANS tool is accessible by the administrators of all subscriber networks, and performs various provisioning functions besides QoS, as for example Layer-3 and Layer-2 MPLS VPN provisioning. The tool is based on a topology database which models the Layer-2 and Layer-3 network topology and stores subscriber link bandwidth. The database also stores the values of the maximum allowed reservation rates a_c, for each subscriber link c, as pre-computed by the algorithms of the Appendix.

Through the submission of simple forms like the one shown in Fig. 2, the administrator of a subscriber network can request any of the available service types, its endpoints (two endpoints for the SDA service, one for the SADU service), the start and end date of the request validity period and access control lists. Because a_c's are pre-computed, the run-time processing that the tool must perform is minimal: upon submission of an SDA request for a traffic amount b, the tool checks if there is sufficient room for the request in both the source and the destination subscriber links s and d, that is, if min $\{a_s \cdot b_s - g_s - b, a_d \cdot b_d - g_d - b\} \geq 0$, where g_s and g_d are the already

---

[3] We assume a symmetric traffic model: the maximum rate at which a subscriber network is allowed to send *Premium* traffic to the network is the same as the maximum rate at which it is allowed to receive *Premium* traffic from the network.

[4] In practice, we have found it useful to define f as a monotonically decreasing function, since allocating larger percentages of the access link bandwidth for *Premium* traffic makes more sense for smaller link bandwidths.

[5] Note that the terms *source* and *destination* are not related with the source or the destination of the traffic; they rather represent the source and the destination subscriber networks contained in the service request.

allocated bandwidths for previous requests involving subscriber links *s* and *d*, respectively. If so, the tool automatically grants the request. Similarly, upon receipt of a SADU request for traffic amount *b*, the tool checks if there is sufficient room for the request in the source subscriber link *s*, that is, if $a_s \cdot b_s - g_s - b \geq 0$ and, if so, it grants the request.

Besides the above, the tool performs many more administrative tasks, such as parsing the actual QoS configuration on the routers, comparing the provisional and the actual configurations, reporting on inconsistencies and providing for automated decommissioning of expired requests.

C. Handling of Layer-2 Core

As shown in Fig. 1 (b), part of the network core (and consequently, part of the network border equipment) consists of Layer-2 switching devices. In order to ensure a uniform implementation of PHBs, our design choice was to implement all PHB- and police-related functions at the L-3 network boundary. To this end, we have optimized the L-2 core using the spanning-tree protocol [9], so as to ensure that no link capacity mismatches occur in the L-2 core, and that topology redundancy is explored to provide adequate bandwidth to all subscribers. We use one or more VLANs per subscriber, so that we can manage subscribers both at the L-2 domain using the SPT protocol and at the L-3 domain.

Having ensured that no congestion points exist in the L-2 core, we use scripting to query the speed and bandwidth settings at each L-2 border interface. We then reflect the speed setting of the border interface into a traffic shaping queue for the respective VLAN at the L-3 border. Using this technique, we make sure that the congestion points occur only at the L-3 border. Thus, subscribers can use the provisioning tool just as if they were connected to the L-3 core.

D. Service enforcement

After granting a request, the ANS tool provisions the police functions at the network perimeter. The tool performs this task automatically, by generating appropriate router configuration commands. For the SDA service type, each request results in a separate police function. Admission control for SDA is done statically, by means of access-control lists that filter packets at the Layer-3 boundary of the network based on the source and/or the destination IP address, protocols and ports. The police functions resulting from each request are cascaded into a larger, composite policer.

For the SADU service type, admission control is done both statically and dynamically. The static police function is similar to the SDA service type, but may contain only checks on the source IP address, protocols and ports. For the dynamic part, it is important to note that the subscriber networks have to contribute a run-time admission control function. For the IP telephony and other H.323 applications for which the SADU service type is intended, this function is accomplished by H.323 gatekeepers located at each subscriber network. Gatekeepers keep track of the actual bandwidth traversing the

subscriber link at each instant and permit or deny the establishment of new calls accordingly. Gatekeeper-based bandwidth control is built into the H.323 protocol, so this step does not require extensive engineering besides specifying an upper *bandwidth* limit on each subscriber's gatekeeper.[6] Because both the source and the destination gatekeepers are queried before establishing an H.323 call and injecting colored traffic into the network, neither the provisioned *VoIP* capacity of the source subscriber, nor that of the destination subscriber can be overwhelmed with excess priority traffic, provided of course that all subscriber networks obey the above rules.

IV. HIERARCHICALLY FEDERATED MAN NETWORKS

In this section we examine the issues discussed so far in the context of a set of hierarchically federated MAN networks. We assume that compatible DiffServ-based QoS service definitions and mechanisms exist in all member networks; thus, we only discuss the issue of service provisioning across hierarchical boundaries, and the issue of advanced admission control and advanced run-time signaling techniques.

A. Provisioning

We support automated provisioning with our ANS tool; however, provisioning in a hierarchical MAN federation requires the cooperation of our tool with other similar tools. This involves application-to-application cooperation for which XML [10] and the Web Services framework [11], [12] form an ideal toolset. As of this writing, development based on the WS toolset is underway by both GRNET and the SA3 activity of the GN2 project [18]. Development efforts aim at transforming existing web-based provisioning systems into a set of interoperating Web Services (WS), where each WS will manage one level in the hierarchical federation. It is expected that subscriber networks will adopt this paradigm and adapt their internal provisioning systems accordingly.

While the exact information that needs to be exchanged among provisioning tools at the various hierarchy levels has not yet been fixed, it is expected that for each service request, the tools involved will require the exchange of topology information, service type, endpoints for the request, amount of *Premium* bandwidth requested and the police function details.

B. Admission Control and Signaling

Admission control is implemented via police functions at the perimeter of each hierarchical level of the federation and is thus relatively easy to implement, at least for the SDA service type. For the SADU service type, an *aggregate admission control function* is necessary at each hierarchy boundary. For example, considering the case of H.323 and gatekeepers, a "federal" (e.g., national) bandwidth-controlling gatekeeper function is required to keep track of calls that traverse the hierarchical boundary between GRNET and Géant. This function is the subject of future development.

---

[6] Review procedures in GRNET exist to make sure that subscriber networks comply with this rule before granting them a *VoIP* QoS request.

A more advanced technique allows networks lower in the hierarchy to manage admission control functions in higher-level networks via signaling. To this end, we have considered the QoS Policy Propagation on BGP (QPPB) technique. Using QPPB, a subscriber network may modify its own admission control function at run-time, by allowing or denying traffic to specific destination addresses within its own network. QPPB is ideal for hierarchical implementation because its basis, BGP, can aggregate prefixes to receive *Premium*-type traffic at all hierarchy levels, thus making the whole approach scalable and viable.

V. NUMERICAL AND USAGE DATA

As of this writing, the QoS service of GRNET is being used by a number of academic institutions. Table I summarizes the currently active QoS requests (only MAN nodes are listed).

TABLE I
CURRENTLY ACTIVE QoS REQUESTS BY GRNET MAN SUBSCRIBERS

Institution	Access BW	Request type	Request BW
Univ. of Athens	1 Gbps	*VoIP*	1750 Kbps
Nat. Tech. Univ. Athens	1 Gbps	*VoIP*	2000 Kbps
Isabella EGEE node	1 Gbps	*Premium*	50 Mbps
Tech. Univ. Crete	1 Gbps	*VoIP*	1374 Kbps
GRNET Headquarters	2×2 Mbps	*VoIP*	270 Kbps

The actual function f used in the GRNET ANS tool is shown in Fig. 3. The resulting worst-case h_l was calculated to ~110 Mbps for a link in the Athens MAN, corresponding to about 4.3% of the link's capacity (STM-16), well below a.
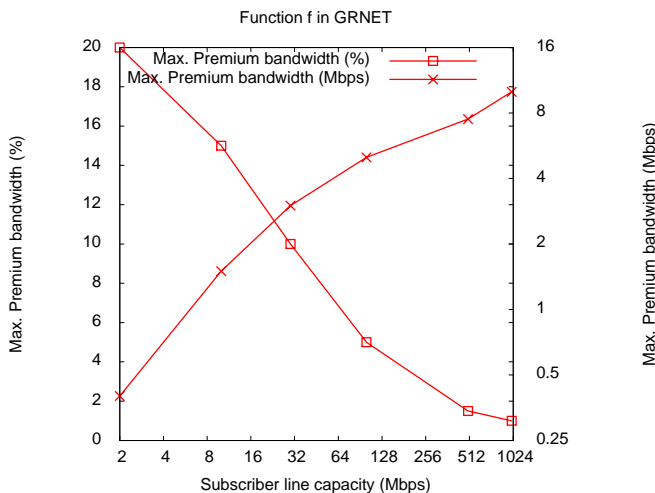

Fig. 3. The function f used in our ANS tool.

Unfortunately, detailed delay and jitter measurement data from the QoS-enabled network are not yet available. We are in the process of deploying measurement equipment and software as CPE equipment on subscriber networks. Using this equipment, we expect measurement data to be available by the end of year 2005. We are however measuring the usage of the existing QoS bandwidth reservations using MRTG [13]. Two such example plots are shown in Fig. 4 (note that due to its 5-minute sampling interval, MRTG does not capture accurately "spikes" typical in short-term voice calls; thus, the instantaneous usage may be considerably higher than that shown in the plots).
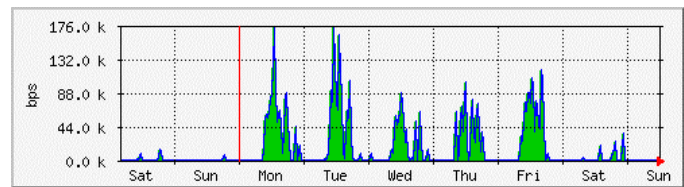
(a) University of Athens
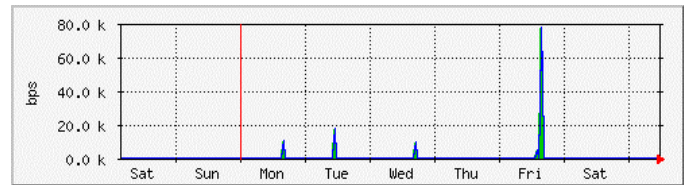
(b) GRNET Headquarter offices

Fig. 4. Two sample QoS bandwidth usage plots (last week of July 2005).

VI. RELATED WORK

Many QoS issues were initially investigated by the SEQUIN project [5]. SEQUIN has adopted the DiffServ framework in order to define an end-to-end approach to QoS over independently managed domains. The resulting "Premium IP" service (equivalent to the *Premium* service discussed in this paper), was thus conceived, tested and finally introduced in Géant [14]. Meanwhile, several other projects, e.g. [15], [16], [17] have investigated QoS issues.

The GN2 project [18] that has started recently has taken over the above work in order to provide a cross-NREN service, in line with our own work. GN2 currently envisages only an SDA service type; however, development in GN2 has also taken into account the SADU model of GRNET so that future cooperation between the two domains is likely to also cover the SADU model. GRNET participates and contributes actively in the specific "SA3" activity of the GN2 project which deals with QoS provisioning tools and services.

Network dimensioning for priority IP traffic is typically an optimization problem involving link costs. In case that delay and loss constraints are also considered, the resulting problem is NP-complete, thus its solution requires several assumptions and involves various heuristics [20]. Other works, e.g., [19], have examined optimization objectives when priority IP traffic co-exists with best effort traffic. Unlike [19] however, our approach does not assume knowledge of a traffic matrix between source-destination pairs, and finds the worst-case capacity independently of the underlying routing algorithm.

VII. CONCLUSION AND FUTURE WORK

In this paper we have presented a set of techniques in use today at the Athens and the Crete MANs of GRNET. We have described the service model and the different supported service types. We have discussed the issue of worst-case service dimensioning and we have provided two equivalent algorithms for verifying *Premium* traffic percentage

assignments in the border of the network. Both algorithms have been shown to perform satisfactorily in the special case of MAN rings. Thanks to efficient dimensioning, we have been able to automate the provisioning process. We have presented our web-based tool that handles this process and explained its functionality. Finally, we have discussed the issues of integrating the presented QoS techniques in a hierarchically-federated set of cooperating networks.

Our future plans include augmenting our provisioning tool with Web Services functionality, so that it can interoperate with peer services at different hierarchy levels. We also plan to deploy additional, more sophisticated run-time admission control schemes, depending on the availability of such mechanisms in networking equipment. Finally, as switching equipment becomes more and more powerful, we consider moving the PHB- and police-related functions to the L-2 network boundary, so as to provide a unified QoS service layer across both the L-2 and the L-3 domains of GRNET.

REFERENCES

[1] RFC 2475, "An Architecture for Differentiated Services", S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, W. Weiss, December 1998.
[2] GRNET home page, (http://www.grnet.gr/?language=en).
[3] Géant home page, (http://www.geant.net/).
[4] V. Jacobson, K. Nichols, K.Poduri, "An Expedited Forwarding PHB", *RFC 2598*, June 1999.
[5] SEQUIN: "Service Quality across Independently Managed Networks", IST Project IST-1999-20841, (http://www.dante.net/sequin/)
[6] M. Shreedhar and G. Varghese, "Efficient Fair Queuing using Deficit Round Robin", in *Proc. SIGCOMM '95*, Cambridge, USA, 1995.
[7] Cisco Systems, "Understanding and Configuring MDRR/WRED on the Cisco 12000 Series Internet Router", (http://www.cisco.com/warp/public/63/mdrr_wred_overview.html).
[8] V. Haniotakis, D.Primpas and A.Varvitsiotis, "GRNET Advanced Services Tool", *18th TERENA TF-NGN meeting*, Jul 2005, Paris (tool URL: http://anstool.grnet.gr/, demo version at http://edet.ucnet.uoc.gr/demo/html/index.php?lang=en).
[9] IEEE, "802.1D: IEEE Standard for Local and Metropolitan Area Networks – Media Access Control (MAC) Bridges", *IEEE Std 802.1D-2004*, (http://standards.ieee.org/getieee802/download/802.1D-2004.pdf).
[10] F. Yergeau, T. Bray, J. Paoli, C. M. Sperberg-McQueen and E. Maler, "Extensible Markup Language (XML) 1.0 (3rd Edition)", *W3C*, Feb 2004, (http://www.w3.org/TR/2004/REC-xml-20040204/).
[11] SOAP/1.1 TR, *W3C*, Jun 2003, (http://www.w3.org/TR/soap/).
[12] E. Christensen, F. Curbera, G. Meredith and S. Weerawarana, "Web Services Description Language (WSDL) 1.1", *W3C TR*, Mar 2001, (http://www.w3.org/TR/wsdl/).
[13] T. Oetiker, The Multi-Router Traffic Grapher tool (online at http://people.ee.ethz.ch/~oetiker/webtools/mrtg/).
[14] C. Bouras, M. Campanella, M. Przybylski and A. Sevasti, "QoS and SLA aspects across multiple management domains: The SEQUIN approach", *FGCS*, TERENA 2002 Networking Conference, Elsevier Science, Volume 19, Issue 2, February 2003, pp. 313-326
[15] EUQoS Project, (http://www.euqos.org/index.php).
[16] Howarth M. et al., "Provisioning for Interdomain Quality of Service: The MESCAL Approach", *IEEE Comm.*, Vol. 43, No 6, June 2005.
[17] Roth R. et al., "IP QoS Across Multiple Management Domains: Practical Experiences for the Pan-European Experiments", *IEEE Comm.*, Vol. 41, No 1, January 2003
[18] GN2 project home page, (http://www.geant2.net).
[19] K. Wu and D. S. Reeves, "Capacity Planning of DiffServ Networks with Best-Effort and Expedited Forwarding Traffic", *Telecommunication Systems*, Vol. 25, No. 3-4, pp. 193-207, March-April 2004.
[20] P. Trimintzios, T. Bauge, G. Pavlou, L. Georgiadis, P. Flegkas, R.Egan, "Quality of Service Provisioning for Supporting Premium Services in IP Networks", *Proc. IEEE Globecom 2002*, Taipei, Taiwan, Vol. 3, pp.2473-2477, IEEE, November 2002.
[21] V. A. Siris and G. I. Fotiadis, "Dimensioning Algorithms for Provisioning in Networks supporting Premium IP Services", *work in preparation*, 2005.

APPENDIX

In this Appendix we provide two equivalent algorithms that solve the dimensioning problem of section III.A. Let $G = (R, L)$ be an undirected graph representing the network topology, where $R = \{r\}$ is the set of routers and $L = \{l\}$ is the set of core links. The total *Premium* traffic t_r that each router r can inject into the network[7] can be obtained by summing the maximum allowed incoming *Premium* bandwidth on all subscriber network links that connect to that router, i.e.

$$t_r = \sum_{c \in C(r)} b_c \cdot f(b_c),$$

where $C(r)$ are the subscriber networks connected to router r. A worst-case upper bound for the bandwidth needed at each core link can be calculated as follows:

1) Algorithm 1

Let $T = (R, L')$ be a spanning tree of G (L' is a subset of L); for each l in L', let $P_l = \{(s, d)\}$ be the set of all paths (s, d) in T from a source router $s \in R$ to a destination router $d \in R$ such that (s, d) contains l. Then, let

$$G_I(P_l) = \sum_{r: \exists d \in R:\, (r,d) \in P_l} t_r \quad \text{and} \quad G_O(P_l) = \sum_{r: \exists s \in R:\, (s,r) \in P_l} t_r .$$

The worst-case premium bandwidth h_l for link l is given by

$$h_l = \min \{G_I(P_l), G_O(P_l)\}.$$

The overall worst-case bandwidth $\underline{h}_l$ is the maximum h_l for all possible spanning trees T of G. In ring topologies which are common in MANs, the set of all possible spanning trees can be easily found by iteratively removing each time one link of the ring. A detailed presentation, proof and evaluation of this algorithm are contained in [21].

2) Algorithm 2

Let K be the set of all cuts k of G. Each cut k is a subset $L_k = \{l\}$ of L which, when removed from L, divide G into exactly two disconnected sub-graphs, G_{Vk} and G_{Wk}. If R_{Vk} and R_{Wk} denote the routers in G_{Vk} and G_{Wk}, respectively, let

$$h_k = \min \{ \sum_{r \in R_{Vk}} t_r, \sum_{r \in R_{Wk}} t_r \}.$$

Then, the worst-case bandwidth $\underline{h}_l$ of a link l is the maximum h_k over the set $K_l \subset K$ of all cuts k that contain link l.

Correctness and performance (second algorithm): if all links in a cut fail simultaneously except one, say l, then all *Premium* traffic from sources in G_{Vk} to destinations in G_{Wk} (or *vice versa*) will flow through the still-working link l; however, even if routers in G_{Vk} can inject more *Premium* traffic than routers in G_{Wk} can accept, there is neither a possible provisioning scenario in the SDA model, nor a possible run-time acceptance scenario in the SADU model that would allow this traffic; hence the min{·} operation. Finally, we note that in ring topologies, the set of all cuts is the set of all pairs

---

[7] Because we assume a symmetric model, this is the same amount of traffic that can sink into the router r.

of distinct links $K = \{(l_i, l_j) \mid j \neq i\}$, so the algorithm can be easily shown to perform in time $O(\|R\|^3)$. $\qquad\square$