Available online at www.sciencedirect.com
ScienceDirect
Computer Networks 51 (2007) 3549–3563
Computer Networks
www.elsevier.com/locate/comnet

An analytical QoS service model for delay-based differentiation

Christos Bouras [a,b,*], Afrodite Sevasti [b,c]

[a] RA Computer Technology Institute – RACTI, Kolokotroni 3, 26221 Patras, Greece
[b] Department of Computer Engineering and Informatics, University of Patras, 26500 Rion, Patras, Greece
[c] Greek Research and Technology Network – GRNET, 56 Mesogion Av., 11574 Athens, Greece

Received 6 February 2006; received in revised form 27 September 2006; accepted 19 February 2007
Available online 15 March 2007

Responsible Editor: G. Ventre

Abstract

The increased need for Quality of Service (QoS) in today's IP networks has concentrated a lot of research and implementation efforts. Carefully designed and managed priority services are essential for quality-demanding traffic, especially in large-scale IP-based environments where aggregation of flows is extensive and a variety of traffic types co-exist. Although individual mechanisms have been widely investigated, not much related work exists on integrated approaches to QoS provisioning that are also feasible to implement. This work presents a thorough approach to the design, dimensioning and provisioning of a high-priority service for high-quality-demanding traffic over an IP network. Our approach employs efficient scheduling and a dynamic admission control scheme while demonstrating an outstanding performance in terms of the quality offered to QoS-demanding traffic. A novel feature is also introduced: while other existing schemes only focus on provisioning of service rate guarantees, ours achieves in addition differentiation of the end-to-end delay perceived by IP flows.

Keywords: IP QoS; High-priority service; Delay-based service differentiation

1. Introduction

Despite the success of the best-effort service model of the Internet, the provision of Quality of Service guarantees in IP networks has become a trend in packet-based networks' research. Approaches vary from detailed per flow resources' provisioning to mechanisms that operate on IP flow aggregates in an effort to reduce complexity and achieve scalability. Theoretical models that have been proposed employ either deterministic or statistical traffic representations in order to provide deterministic or stochastic qualitative guarantees to IP traffic. Other proposals simply consider over-provisioning [1]. Still most of the theoretical models are rarely implemented in practice due to practical limitations.

The goal of this work is to build upon and extend existing research in order to (a) devise a service

* Corresponding author. Address: RA Computer Technology Institute, 61 Riga Feraiou Str., 26221 Patras, Greece. Tel.: +30 2610 960375; fax: +30 2610 960358.

E-mail addresses: bouras@cti.gr (C. Bouras), sevasti@grnet.gr (A. Sevasti).

1389-1286/$ - see front matter
doi:10.1016/j.comnet.2007.02.010

model that delivers advanced QoS to IP flow aggregates and is applicable to production networks, (b) define a set of explicit provisioning principles for the realization of this model that can be implemented by network administrators without administrative overhead and scalability implications, (c) include an acceptance control mechanism for incoming QoS service requests without service disruption and last but not least (d) enhance the performance of existing, comparable schemes by introducing a novel feature: delay-based service differentiation.

As stated in [2], the main difficulty for provisioning statistical QoS for a multi-node network lies in addressing the complex correlation of traffic at downstream multiplexing points. A lot of work in the area has focused on analytically characterizing aggregated traffic within a network. However, most of the models are restricted to the single node case or only to a specific type of traffic (such as VoIP traffic) in the multi-node case, due to the fact that modelling of aggregated traffic, especially in the cases where the statistical characteristics of individual flows are heterogeneous and unpredictable, is infeasible. So relying purely on analytical models of traffic in the core is not possible for domain-wide provisioning of QoS services.

Other efforts have focused on reconstructing traffic characteristics inside the network so that modelling at core nodes can be done assuming the same traffic pattern as in the ingress node, where traffic is fully compliant to its envelope. Reconstructing traffic characteristics at each node introduces complexity that we wish to avoid, let alone introducing artificial delays and worsening the statistical QoS achieved. Another approach that also introduces complexity and scalability problems in large networks with multiplexed flows is to explicitly reserve resources for the subset of flows traversing the exact same path throughout a topology [2]. This removes a large part of the complex correlation of traffic in the core, as it only exploits the multiplexing gain of flows on the same path and is easier to model. However, it introduces an amount of complexity, including reconfigurations of resource allocation among intersecting paths in a node whenever a new flow has to be admitted in one of them.

In this work, we assume aggregated IP traffic complying with deterministic traffic envelopes imposed at the network edge. Our approach does not employ traffic profile reconstruction mechanisms in the core neither distinguishes between flows following the same path across the network. We propose a set of domain-wide mechanisms operating on all flow aggregates traversing an IP backbone network, in an effort to provide QoS guarantees and efficient service differentiation to IP packets in a scalable manner. We focus on deploying a single service, in addition to Best Effort, that offers guaranteed service rate, bounded end-to-end delay, minimal jitter and no packet loss to the most exigent applications in IP networks. We call this a Gold service but other names, such as Premium IP service, have been used to refer to the same notion.

For the deployment of a service for traffic with stringent QoS requirements, related work in the literature takes for granted some amount of service rate over-provisioning or at least exact service rate provisioning for the total of Gold traffic served through each node. Thus, guaranteed throughput is not an issue as there is everywhere enough capacity to serve Gold traffic. What is more important and comprises the main topic of related research is how to achieve bounded, minimal end-to-end delay, jitter and losses for Gold traffic packets. In [3], it is emphasized that the limiting factor for a Gold-like service is not bandwidth but end-to-end delay. This issue has to be dealt with in a domain-wide manner, as the aggregation and dis-aggregation of real-time traffic flows, which are usually bursty enough at the source, can magnify burstiness and inter-packet delay variation in an uncontrollable manner at the core.

In this work, we control the allocation of resources along the path that a Gold flow follows in a way that guarantees a bounded worst-case perceived end-to-end delay and jitter, zero packet loss and at the same time allows for QoS differentiation without compromising quality. Our approach ensures that the end-to-end queuing delay perceived by each Gold flow is relative to the queuing delay tolerance of the flow, which is in fact an input parameter to the Gold service provisioning process. In this way, the packets of a Gold flow with a queuing delay tolerance of 20 ms perceive an actual worst-case queuing delay which is larger than the one perceived by the packets of a Gold flow with a queuing delay tolerance of 10 ms whether, assuming the flows use the same path or at least share the same ingress and egress points in the domain.

We provide an efficient Call Admission Control (CAC) mechanism that achieves this performance profile and comprises a tool in the hands of the network operator for dimensioning and providing the

Gold service. Apart from the novelty of delay-based service differentiation, the proposed CAC has the nice property of ensuring zero packet losses for Gold traffic as the admission control criterion set by the scheme is that no packet ever misses its service time deadline at each node along its path and thus that no packet of an accepted Gold flow is ever dropped.

Our approach employs some existing results of QoS related research. We use Dynamic Packet State (DPS) as originally proposed in [4] to store at the ingress information for each packet's treatment in the core. We use a version of Earliest Deadline First (EDF) scheduling, that assigns local deadlines of packets at each node derived from time-shifting of each packet's arrival time at the ingress node, similar to DC-EDF of Zhu et al. [5]. Our proposed dynamic CAC scheme is based on DPS and an arrival-times based EDF. However, unlike DC-EDF we propose a domain-wide aware policy for dividing the end-to-end delay budget of flows along a path.

Our approach does not require any kind of measurements for its operation. No monitoring statistics need to be retrieved from the network in order to make a decision about accepting a new flow or for configuring network elements. We consider this to be an advantage of our proposed model, as premium QoS is achieved with minimum implementation complexity and overhead.

An initial set of results of the work presented here, has already been demonstrated in [6]. As far as we know, there is no similar work on the issue of delay-based differentiation for high-priority traffic in IP networks.

2. Related work

There is an extensive amount of related research work on scheduling algorithms, admission control and resource provisioning schemes for QoS services. The approaches that are most relevant to or have inspired our work are briefly presented here. Our proposed schema exploits some of these previous results but provides a novel delay-based differentiation QoS service. It offers rate and worst-case delay guarantees as well as an efficient flow admission control, relying upon domain-wide knowledge. As the authors of [7] point out, multi-hop worst-case delay can only be achieved with controlling and constraining the global network topology.

A number of papers [8,9] assume a Generalized Processor Sharing (GPS) scheduler that isolates flows in order to provide service differentiation and QoS. GPS or rate-based scheduling and its variants (e.g. Weighted Fair Queuing – WFQ) ensure that each flow or aggregate transiting a node is served with a dedicated, configured service rate which is unaffected of the load imposed to the node by other flows and traffic in general. GPS needs a different queue for each aggregate to ensure a certain service rate. In a real world scenario, the latter imposes practical limitations as there is an upper limit in the number of queues that can be configured in a router and thus in the number of aggregates that can be simultaneously served. Variations of dynamically adjusting per flow GPS weights in a node according to the current situation in order to achieve higher utilization or optimise packet delay are proposed in [1,8,9]. However, although these highly dynamic schemes make a better usage of the available resources and thus achieve tighter delay bounds than static systems using GPS, the trade-off between scalability and efficiency becomes evident.

GPS and its variants are often met in the literature as scheduling mechanisms that provide service differentiation at each node of a topology. Our focus in this work is more on QoS service models with a domain-wide validity. Thus, the domain-wide, parameter-based flow admission control scheme of Fidler and Sander [3], which is based on GPS and uses network calculus, the theory of deterministic queuing systems, is directly comparable to our approach. The authors of [3] assume statically configured core routers, destination-aware service provisioning and parameterisation of the service at the edge routers. Arrival and service curves are used to derive generic deterministic bounds on backlog and delay, while specific bounds are provided under the assumption of rate-latency service curves and leaky bucket regulated arrival curves. This work is compared in terms of performance with our proposed schema in Section 4 of this paper.

Another work that exploits GPS for scheduling but also attempts path-wide optimisation of resources' allocation is presented in [10]. The authors partition the end-to-end QoS requirement of a network flow along the links of a given path such that the deviation in the loads of the network links is as small as possible.

Going back to schedulers for QoS provisioning, the work of Zhu et al. [5], where a proposal for efficient EDF scheduling is presented, is particularly relevant to our work. EDF has been shown to provide optimal delay bounds at a single node and outperform WFQ in the end-to-end case, if per node traffic shaping is exercised. Variations of EDF exist, such as Rate-Controlled EDF (RC-EDF), which employs shapers at each node, enforcing delays to packets in order to make each flow fully conformant to its traffic envelope. In an effort to address such shortcomings, the authors of [5] proposed deadline-curve-based EDF (DC-EDF) where the local deadline of each packet in a node is a strict time-shifting of the arrival of the packet at the ingress node of a network. When DC-EDF is used, a packet's large local delays at some nodes are balanced with small local delays at some other nodes in order for the bound on end-to-end delay to be met.

In [11], the authors use EDF in the context of coordinating packet deadlines among core nodes, in another approach that uses EDF for domain-wide service provisioning. The key principle here is that the priority index of each packet at a downstream node depends on its priority index at upstream nodes so that all nodes in the network cooperate to provide an end-to-end service. Our approach is somewhat different. We explicitly distribute resources available to flows in proportion to the load of the nodes traversed, during the admission control phase. Thus, the benefits of DC-EDF are reinforced when combined with a domain-wide-load-aware provisioning policy. We propose an admission control mechanism that exploits DC-EDF principles in order to efficiently implement the Gold service. A strong point of our proposal is the use of global load knowledge in order to assign per node packet deadlines.

Another closely related work to ours is presented in [12]. The proposal is to use Coordinated EDF (CEDF) to denote local deadlines of packets at each node traversed in proportion to the server speed at each hop (over the sum of speeds in the end-to-end path). However, this model assumes exact knowledge of the traffic pattern at its source which is rather unrealistic.

A number of measurement-based approaches for implementing QoS-enabled services and corresponding admission control schemes exist in the literature. A measurement-based CAC scheme (MBCAC) for serving bursty multimedia traffic is compared against two model-based CAC schemes that assume a priori statistical knowledge of traffic in [13]. In [14] a measurement-based admission control scheme is proposed, for processing service requests of new flows in a network serving flow aggregates. The goal is to provide a decision on whether to accept a service request or not based on an estimation of the current load on each node and the load that the acceptance of the new flow will impose. Another measurement-based admission control scheme that exploits statistical multiplexing is proposed in [15]. This scheme allows aggregates that can tolerate more delay violations to reserve fewer resources than those that tolerate less, even though they require the same delay bound.

The idea of using the IP packet header for carrying state initialised by the ingress router and exploited by core routers for processing packets was first proposed in [4] with the name of Dynamic Packet State (DPS). Based on the Virtual Clock algorithm, according to which packets are held in the rate controller until they become eligible and are served in the order of their locally assigned deadlines at each node, Core-Jitter-VC (CJVC) is defined. CJVC allows for encoding eligible and deadline times for each packet traversing a path at the ingress node. The approach of Stoica and Zhang [4] is directly comparable to our work, since we also use DPS and the notion of pre-computing per node packet deadlines at the ingress node in combination with our proposed scheduling and admission control per-domain behaviour so as to ensure tight statistical delay bounds and global optimisation of resource usage. For this reason, CJVC is the second scheme that is compared in terms of performance with our proposal in Section 4 of this paper.

In another work that deals with designing networks with end-to-end statistical QoS guarantees [2], the authors assume the use of delay jitter controllers at each hop ensuring that traffic experiences its maximum allocated per node delay. This model, apart from delay jitter controllers which introduce additional delays to packets, assumes also per-class buffering and guaranteed service rate for each class (as defined within the scope of Patek et al. [2]) in the core. These assumptions make the number of flows, and thus of different queues and states preserved at each core node, excessive and thus the model lacks scalability and is not possible to compare with our approach.

Finally, in [16] a resource provisioning algorithm is proposed for achieving domain-wide load balance while simultaneously meeting the QoS requirements for bandwidth and end-to-end delay of individual flows. In Link Criticality Based Routing (LCBR) a link's cost is defined so that a link with a small residual capacity or large expected load is considered more expensive. LCBR does not focus on mechanisms for QoS on each individual node, but in domain-wide efficient path selection for serving QoS-demanding flows.

3. Proposed framework

We consider an IP network with N nodes. In every node i with $i \in 1,\ldots,N$, a service queue $Q_{i,l}^{G}$ is configured on each outgoing link l that serves aggregated Gold traffic over l separately from best-effort traffic. $Q_{i,l}^{G}$ has strict non-preemptive priority (PQ) over other queues on link l, rate-limited to a rate of $C_{i,l}^{G}$ so that best-effort traffic is not starved. Alternatively, $Q_{i,l}^{G}$ can be assigned a guaranteed service rate $C_{i,l}^{G}$, using any GPS-compliant service property like WFQ. We will assume the Priority Queuing with non-preemptive configuration for the rest of our analysis.

We consider a number of Gold service customers in the domain. We consider aggregated traffic from each such customer to be a Gold flow as long as it adheres to the following definition:

Definition 1. A Gold flow f consists of a single or multiple micro-flows of Gold traffic grouped to a single aggregate with the following properties:

- All packets of the Gold flow enter the domain at the same ingress node (I_f) and exit from the same egress node (E_f).
- The aggregated traffic comprising the Gold flow f is policed at the ingress node I_f with a dedicated token bucket policer (r_f, b_f), imposing a long term average rate r_f for the Gold flow as well as a maximum burst size b_f.

Based on this definition, a Gold flow can be a single micro-flow or an aggregate of micro-flows, several Gold flows can be defined between the same ingress–egress node pair and one or more Gold flows can be simultaneously scheduled over the same link.

It is important to stress out at this point that $Q_{i,l}^{G}$ is used to temporarily store (if needed) and serve packets of all Gold flows multiplexed over the outgoing link l of node i. Thus, the model ensures that no specific queue configuration or service rate provisioning per Gold flow takes place in individual nodes.

For the rest of this paper, for simplicity in notation, we will consider a single outgoing link for each node in our domain, thus $Q_{i,l}^{G}$ becomes Q_{i}^{G}, $C_{i,l}^{G}$ becomes C_{i}^{G} and so on.

As demonstrated in our previous work [17] but also in related research work, C_{i}^{G} must always be larger than the sum of the long term average rates of the Gold flows being served by Q_{i}^{G}. Thus

$$\sum_{j\in F} r_j t \leqslant C_i^G t \Rightarrow \sum_{j\in F} r_j t = \rho_i C_i^G t, \quad \text{with } \rho_i \leqslant 1, \text{ as } t \to \infty, \tag{1}$$

where F is the set of Gold flows served through Q_{i}^{G}. This is a stability condition assuring that load does not build up in Q_{i}^{G}.

3.1. Load-aware EDF scheduling (LA-EDF)

Apart from the externally observable service rate C_{i}^{G} of the Gold traffic queue, packets within Q_{i}^{G} are ordered (and thus served) in an EDF manner. According to our proposed policy defined in this section, each packet k of Gold flow j served through node i is assigned a local deadline $d_i^{k,j}$. In compliance to EDF principles, the packet with the smallest $d_i^{k,j}$ value is the one to be served first from Q_{i}^{G}.

As already mentioned, DC-EDF scheduling (initially proposed in [5] and also addressed in [12]) assigns deadlines on a delay bound provisioning perspective and results in larger schedulability than rate-based EDF. DC-EDF has the nice property of allowing to analytically model the deadlines of packets along a path based on the traffic profile at the network ingress. This property has been very useful in our approach and has made DC-EDF-based scheduling appropriate for our Gold service model, as it allows service dimensioning and admission control based on information available at the network edges.

We propose a novel policy for obtaining $d_i^{k,j}$ values for each hop along the packets' end-to-end path. These $d_i^{k,j}$ values still occur from a time-shifting of each packet's arrival time ($a_I^{k,j}$) in the corresponding ingress node I_j, as DC-EDF requires. However, the selection of deadlines attempts to optimise global schedulability, maximize the number of Gold flows that can be safely accommodated and achieved delay-based service differentiation.

We denote by DB^j the required end-to-end delay bound from I_j to E_j for the packets of flow j under the Gold service model. We propose that in order to determine how much $a_I^{k,j}$ is shifted in time to give $d_i^{k,j}$ we consider the load of all Gold traffic queues along the path p that the packets of flow j follow. We call this the Load-Aware EDF (LA-EDF). Assuming Gold traffic queues with equal C^G values, $a_I^{k,j}$ is shifted more to provide $d_i^{k,j}$ at node i when Q_i^G admits a larger ρ_i value (see (1)), compared to how much $a_I^{k,j}$ is time-shifted to provide $d_n^{k,j}$ in node n where $\rho_i > \rho_n$.

The delay due to policing that each packet of flow j endures at the ingress policer cannot be considered part of DB^j, because the packets facing this additional delay are those that are non-conforming to the corresponding policer. According to the Gold service provisioning model, the DB^j end-to-end delay bound for the packets of flow j is guaranteed only for in-profile packets of the flow. Thus, we define $\overline{a_{I_j}^{k,j}}$ to denote the in-profile arrival time of packet k at node I_j, according to the policer for flow j. We also define the term DP_p as the propagation delay of each packet along p, where p represents the end-to-end path for the packets of flow j. Based on these definitions we can proceed with our first proposition.

Proposition 1 (Load-Aware EDF). *We propose a quantization of the end-to-end delay budget DB^j for the packets of flow j that is dependent upon the ratio of load of each node that the packets of flow j traverse when compared to the sum of loads of the nodes along the end-to-end path from I_j to E_j. The local deadline for packet k of Gold flow j served at node i is*

$$d_i^{k,j} = \overline{a_{I_j}^{k,j}} + \sum_{q=1}^{i} \sigma_q^j. \tag{2}$$

In this proposition, σ_i^j is defined as the 'local slack term' for each packet of flow j at node i, thus the maximum delay a packet of j can endure while at Q_i^G. It is defined as

$$\sigma_i^j = (DB^j - DP_p)\frac{\rho_i}{\sum_{l \in p} \rho_l}. \tag{3}$$

It has to be stressed out at this point, that the individual ρ_i values in (3) are based on the total Gold traffic subscription on each topology link, thus the sum of the accepted flows' profile rates. Thus LA-EDF does not require real-time measurements for its operation and is therefore robust against instant traffic fluctuations.

According to Proposition 1, the delay budget for the packets of a flow is distributed along the path in a way proportional to the load encountered at each node when compared to the accumulated load along the end-to-end path. Obviously, the same local deadline d_i^j is assigned to all packets of a flow j at node i.

We also provide the following definition:

Definition 2. The 'latest arrival time' of packet k in node i is defined as

$$\overline{a_i^{k,j}} = d_{i-1}^{k,j} \tag{4}$$

thus as the deadline of the packet in the upstream node.

The proposed model anticipates for dropping of a packet at a node when it exceeds its local deadline. Thus, a packet can never arrive at a node i later than its 'latest arrival time' $\overline{a_i^{k,j}}$ because otherwise it will have violated its local deadline $d_{i-1}^{k,j}$ and have been dropped at node $i-1$. As the definition of Gold traffic requires elimination of the packet losses and a hard worst-case delay bound, our proposal for load-aware assignment of deadlines, aims to maximize the probability that a packet meets all its local deadlines along the end-to-end path.

3.2. Delay-Based Admission Control (DBAC)

With LA-EDF implemented in each queue Q_i^G, a domain-wide Gold service implementation must also anticipate for a requests' admission control scheme. The admission control process should provide a positive or negative result as to whether a new Gold flow f can be added to the set of flows served through a path p. Admission control should ensure that the service rate and end-to-end delay guarantees of flow f along p and the existing flows using all or part of p are respected. It should also provide a zero packet loss probability when the new flow is served over p.

Assuming that a request for serving a new flow with arrival process A_f and a deterministic envelope $\overline{A_f} = (r_f, b_f)$ emerges. We want to examine whether the flow can be accepted along a path p where I_p is the ingress node and E_p is the egress node for all packets of the flow. As a first step, the local slack term at each node of p for the packets of f is obtained, based on (2) and (3). Our proposed Delay-Based Admission Control (DBAC) then focuses on the delay tolerance of the packets of f at each node i of p separately, taking into account the existing set of flows served through Q_i^G.

We begin our definition of DBAC by making the following observation: Assume a non-schedulable set of packets in Q_i^G, based on the actual arrival times at the ith node of p and the assigned packet deadlines at the same node. All packets can be delayed until their corresponding latest arrival times in Q_i^G before being considered eligible to serve. Thus packet k is assumed to arrive at the ith node of the path at time $\overline{\alpha_i^{k,j}}$. This certainly does not violate the local delay bound for any packet, since due to Definition 2, $\overline{a_i^{k,j}} = d_{i-1}^{k,j} = d_i^{k,j} - \sigma_i^{k,j} < d_i^{k,j}$. If we can ensure that the aggregated arrival process denoted by the latest arrival times of all packets is schedulable, so is the actual aggregated arrival process.

DBAC examines schedulability conditions at each node separately, based on the following corollary:

Corollary 1. *Under the assumption of any EDF scheduling based on time-shifting of arrival times (such as LA-EDF), if the packet of a flow f is schedulable in node m when all existing flows' packets arrive at m at their latest arrival times and the packet itself arrives at m at its latest arrival time* $\overline{a_m^{k,f}}$, *then the packet is schedulable at node m under any circumstances.*

Corollary 1 allows us to conduct admission control for each node of p based on the latest arrival times of packets from existing flows, the latest arrival time of packets of the new flow f, the traffic envelopes of flows as well as the nice properties of LA-EDF. To make things even more convenient, all the information required for DBAC's operation can be kept at the ingress node (or some Gold service broker operating for the whole domain) and encoded in the packet headers. No state information or configuration changes are required in core nodes.

The intuitive schedulability condition for the case of EDF schedulers, which is refined for the case of deadlines that are time-shifting of each packet's arrival time (like LA-EDF) in [5] is as follows:

$$\sum_{j\in F} \max\{0, D_j(v+\tau) - A_j(v)\} \leqslant C_i^G\tau - l_i^{\max} \quad \text{for each } \tau, \tag{5}$$

where $D_j(t)$ is the continuous function providing the cumulative amount of traffic from flow j with a deadline up to time t. $D_j(t)$ is also known as the deadline curve of j. Similarly, $A_j(t)$ is the continuous function providing the amount of flow j traffic from that has arrived to Q_I^G up to time t (arrival curve of j). Finally, F is the set of Gold flows traversing node i and $l_i^{\max}$ is the maximum sized packet served in node i. The latter is part of the right-hand side in (5) due to the non-preemptive assumption for Q_I^G. The intuition behind (5) is that for all flows using node i and for each interval of length τ, the amount of Gold traffic that has arrived after the beginning of the interval and has no deadlines after the end of the interval has to be less than the amount of Gold traffic that can be served by the node in the same interval.

Here, we extend this schedulability criterion for the purposes of performing DBAC for a single additional flow f on node i of p. We order already accepted flows in node i so that $k < l$ whenever $\sigma_i^k < \sigma_i^l$.

For simplicity we consider DBAC taking place at ingress node I_p. The basic assumption is that we have a set F of Gold flows in node I_p, for which the inequality of (5) holds for each τ. We wish to define DBAC to ensure that whenever a new Gold flow f is being accepted, the set $F \cup f$ remains schedulable. For the purposes of our DBAC, we exploit the property of LA-EDF provided in (2). It allows to relate the deadline and arrival curves of flow j as follows

$$D_j(v) = A_j(v - \sigma_i^j). \tag{6}$$

Eq. (6) implies that all flow j packets with a deadline up to moment v must have arrived at I_p (the latest) by moment $v - \sigma_{I_p}^j$. Based on that,

$$\begin{aligned} D_j(v+\tau) - A_j(v) &= A_j(v+\tau-\sigma_i^j) - A_j(v) \\ &= A_j(\tau - \sigma_i^j). \end{aligned} \tag{7}$$

Also, the deterministic envelope $\overline{A_j} = (r_j, b_j)$ for each flow ensures that

$$A_j(\tau) \leqslant r_j\tau + b_j \quad \forall \tau. \tag{8}$$

Using Eq. (7) and (8), an alternative form of (5) can be derived such that:

$$\sum_{k=1}^{j} [r_k(\tau - \sigma_i^k) + b_k] \leqslant C_i^G\tau - \max_{l>j} L_l^{\max} \quad \text{with } \sigma_i^j \leqslant \tau < \sigma_i^{j+1},\ 1 \leqslant j < |F|, \tag{9a}$$

$$\sum_{k=1}^{|F|} [r_k(\tau - \sigma_i^k) + b_k] \leqslant C_i^G\tau \quad \text{with } \tau > \sigma_i^{|F|}. \tag{9b}$$

As long as Eq. (1) is satisfied, it is straightforward to show that the schedulability condition for a flow j is

$$\sigma_{i(\min)}^f \geqslant \frac{b_f + \sum_{k=1}^{j-1}(b_k - r_k\sigma_k) + \max_{l>f} L_l^{\max}}{r_f}. \tag{10}$$

DBAC also has to make sure that the addition of flow f in the set of Gold flows served in node $i = I_p$ does not cause the schedulability criterion of (10) for each flow with a local slack term such that $\sigma_i^j > \sigma_{i(\min)}^f$ to be violated. Finally, based on the following lemma, it needs to be ensured that $F \cup f$ is schedulable $\forall \tau : \tau > \sigma_i^{|F|}$.

Lemma 1. *If a set of flows F is schedulable for a random t such that* $t > \sigma_i^{|F|}$, *then F is schedulable* $\forall \tau : \tau > \sigma_i^{|F|}$.

The proof of the Lemma is straightforward and therefore omitted.

So there is a set of these three steps for the proposed DBAC to ensure that $F \cup f$ is schedulable for any interval τ and here follows our proposition.

Proposition 2 (Delay-Based Admission Control). *A flow f with a local slack term of* $\sigma_{i(\min)}^f$ *is schedulable in node i, assuming a set of already accepted flows F in i, under the following conditions*:

- *The schedulability condition of* (10) *holds with* $\sigma_i^{j-1} < \sigma_{i(\min)}^f < \sigma_i^j$.
- *The schedulability condition of* (10) *is verified for all flows with* $\sigma_i^j > \sigma_{i(\min)}^f$.
- *The schedulability condition of* (9b) *is verified for* $F \cup f$ *and a random t such that* $t > \sigma_i^{|F|}$.

For the specific case of ingress node I_p, we can set $i = I_p$ in (9) and (10). In all other nodes than I_p, according to the DBAC, if σ_i^f as produced from (3) is equal to or greater than $\sigma_{i(\min)}^f$ as produced from (10), then f is schedulable in node i.

DBAC provides a result on whether a Gold flow can be accepted on a single path p. However, when a new flow f can be served by more than one paths, then path selection should be made so that load for Gold traffic, thus ρ_i, is balanced for each link i of the network. Thus links with a higher ρ_i value will still be able to serve some future requests. A proposal for efficient path selection, in addition to LA-EDF and DBAC, is presented in the following section.

3.3. Domain-wide optimisation of Gold service provisioning

Our aim is to maximize the number of Gold flows that can be simultaneously served over a network topology G, by properly selecting the path over which each new Gold flow is served, without compromising the QoS provided. Intuitively, the shortest path is the most proper choice for each flow, since it is expected to consume the least amount of resources and thus result in the maximum possible domain-wide amount of residual resources for next flows. However a network topology can be such that a large number of shortest paths between ingress–egress node pairs use one or more common network links. In such cases these links become soon saturated and future flow requests are impossible to serve using simply the shortest path between the corresponding ingress–egress node.

In order to preserve the nice property of consuming the least amount of resources that shortest paths demonstrate, we will make the selection of the path that will serve each new flow f among the k-shortest paths between the pair $(I_f - E_f)$. This selection will be based on domain-wide optimisation criteria in order to achieve a maximum Gold flow acceptance ratio. The initial incentive for such an approach was inspired by the work presented in [16]. However, in [16] constant monitoring of the load imposed between each $I - E$ pair is required as a means to predict future load and path selection is based on this load prediction. In our case, we avoid monitoring and base path selection decisions on the load of the links as estimated using the number of already accepted Gold flows.

For the purposes of domain-wide optimisation of our proposed Gold service provisioning, we consider two important properties of each link in the topology (similarly to the proposal of Gopalan et al. [16]):

- The criticality c_l of each link l expressed as the ratio of the sum of the k-shortest paths between all possible ingress–egress nodes of the topology that use this link over the total number of k-shortest paths in the topology.
- The residual Gold capacity R_l^G on each link l, provided by $R_l^G = (1 - \rho_l)C_l^G$.

We define a link cost metric $cost_p$ that is composed of these two parameters. Initially we normalize the residual capacity R_l^G and the total Gold capacity C_l^G of each link l using the minimum residual capacity in G as follows:

$$\overline{R_l^G} = \frac{R_l^G}{\min_i R_i^G} \quad \text{and} \quad \overline{C_l^G} = \frac{C_l^G}{\min_i R_i^G}. \tag{11}$$

The link cost is then defined as

$$\text{cost}_l = \frac{c_l}{R_l^G}. \tag{12}$$

Finally, we define the domain-wide cost as the squared magnitude of the distance vector between the actual link costs and the minimum link costs in G:

$$\text{cost}_G = \sum_l \left(\text{cost}_l - \frac{c_l}{C_l^G}\right)^2. \tag{13}$$

The path selection process for the new flow f, begins by obtaining the set K of the k-shortest paths from I_f to E_f. DBAC is then used on each of these paths to find out the set of paths S over which $F \cup f$ is schedulable, with $S \subseteq K$. For each of the paths $p_j \in S$ with $j \in [1, |S|]$, the value of $\cos t'_G$ is obtained, assuming that f is already accepted along p_j and using $R_l^{G'} = R_l^G - r_f$ in (11) for each link of p_j.

Finally, flow f is served over the path among those in S for which $cost'_G$ obtains its minimum value.

4. Experimental evaluation

For the purpose of evaluating the effectiveness of our proposed scheme in QoS provisioning and end-to-end delay differentiation, a number of experiments were conducted in a simulation environment, the components of which are analytically presented in [18]. The experiments were carried using ns-2 simulator [19].

The topology used is depicted in Fig. 1. It comprises of a number of edge nodes (E_i) and three core nodes (C_i) interconnected in a topology where the backbone links' capacities are not at all times adequate to carry the accumulated traffic of all adjacent peripheral links. The topology has been designed to provide at least two alternative paths for serving a traffic flow. DBAC is used to regulate whether arriving flows are accepted or not. At each node a queue is configured to serve Gold traffic with absolute non-preemprive priority (PQ) over all other traffic. Within each Gold traffic queue, LA-EDF is implemented. Each Gold flow f comprises of a random number v, $1 \leqslant v \leqslant 10$ of VoIP micro-flows with $r_j = 80$ Kbps ($j \in [1, v]$) each and a random number m, $1 \leqslant m \leqslant 5$ of streamed MPEG video micro-flows with $r_i = 700$ Kbps each ($i \in [1, m]$). For the simulation of VoIP flows an exponential ON/OFF distribution was used, with an average duration of ON periods equal to 1.004 s, average duration of OFF periods (idle-time) equal to 1.587 s, packet size of 188 bytes (8 byte UDP header + 20 byte IP header + 160 byte voice data) and transmission rate during the 'on' period equal to 80 Kbps. For the simulation of MPEG flows, a number of MPEG video traces were used, with a packet size of 200 bytes. In this way, each Gold flow comprises of a realistic mix of micro-flows in order to simulate burstiness and statistical properties of delay-sensitive traffic in IP networks.

Gold traffic flows' inter-arrival intervals are uniformly distributed with a mean of 10 ms, the duration of each flow has a uniform distribution with a mean of 10 s. For each flow that is accepted for service, a policer is configured at the ingress node with a rate equal to the sum of the individual VoIP and MPEG flows' average rates ($\sum_1^v j + \sum_1^m i$) and a burst size of

$$a \times (v \times L_{\text{VoIP}} + m \times L_{\text{MPEG}}), \tag{14}$$

where L_{VoIP} and L_{MPEG} are the sizes of the VoIP and MPEG traffic packets. The value of a (an integer) is varied in some of the experiments in order to observe how the different mechanisms react when increasing the burstiness of each admitted flow.

The simulated duration of each experiment was 120 min, while each experiment was repeated 10 times.

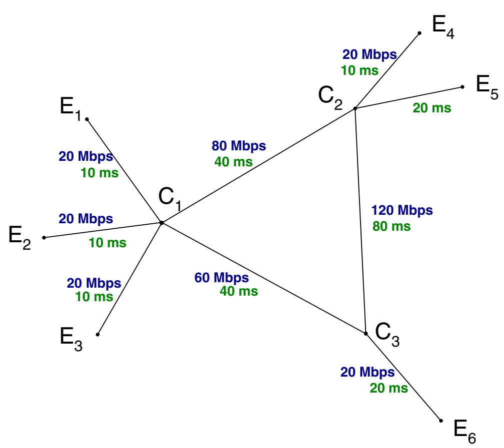

Fig. 1. Topology of simulation experiments.

4.1. LA-EDF and DBAC performance in achievable end-to-end queuing delay

In order to evaluate the effectiveness of LA-EDF as a scheduling mechanism and DBAC as an admission control mechanism in provisioning rate and end-to-end delay guarantees in a backbone network, we are comparing how our mechanisms perform when compared with equivalent QoS provisioning schemes. More specifically, we compare LA-EDF and DBAC with the Core-Jitter-Virtual Clock (CJVC) of Stoica and Zhang [4] and the approach of Fidler and Sander [3], based on Network Calculus. These two approaches are, to the best of our knowledge, the ones that can be directly compared with our work, as they provide QoS mechanisms and admission control for serving quality-demanding traffic in a network serving aggregated IP traffic.

Provisioning rate, end-to-end delay and minimum packet loss guarantees are the requirements that Gold-like service models, including ours, must adhere to. The purpose of the initial experimental evaluation in this section is to show that our proposed schema performs equally well, if not better, to existing equivalent service models in terms of these basic requirements. This is achieved by comparing the performance of the three different schemes as perceived by the same Gold flow in terms of the perceived end-to-end queuing delay. Packet loss statistics are not provided here for brevity's sake, however, LA-EDF and DBAC have verified their expected performance, demonstrating zero packet losses in all experiments presented throughout Section 4. In the subsequent sections, we will demonstrate how our approach, in addition to what existing service models provide, also achieves end-to-end delay differentiation.

It is important to mention that in the scenarios presented in this section, all schemes operate under congestion, as the rejection rate of arriving flows throughout the whole range of experiments is non-zero, or, in other words, Gold flows are generated with a higher rate than that any of the three schemes is able to accommodate. For each experiment, LA-EDF and DBAC are evaluated first and the experiment provides a subset of accepted flows, from the total of arriving flows. This subset is then fed to the experiments using CJVC and the Calculus scheme. In this way, it is ensured that CJVC and the Calculus schemes are evaluated against LA-EDF and DBAC under the same set of accepted flows and thus that the observed robustness of LA-EDF and DBAC in providing minimal queuing delay is not due to a higher flow rejection ratio but rather due to the nice properties of these mechanisms.

In the first set of experiments, the efficiency of LA-EDF and DBAC in terms of achieving minimal queuing delay for QoS-demanding traffic when compared to CJVC and the Calculus scheme is demonstrated by observing how the three different schemes perform when aggregated Gold traffic is highly bursty. We have observed and measured the queuing delay perceived by a Gold flow F (a mix of VoIP and MPEG traffic) generated at E_5 and destined to E_2, as the capacity of the $E_5 - C_2$ link varies from 20 to 50 Mbps. The selection of a specific flow F among the set of flows generated during the simulation is such that the end-to-end delay budget DB^j is one of the lowest observed, however our proposed scheme was tested and proven to demonstrate comparable performance to that of CJVC and Calculus for several Gold flows in the experimental set-up. The selection of the path along which F is served is made according to the domain-wide optimisation principles of Section 3.3.

A higher capacity value for the $E_5 - C_2$ link results in bursts of aggregated Gold traffic from E_5 reaching C_2 with a much higher rate than the capacity of the queue serving Gold traffic on the $C_2 - C_1$ link. As a result, all three schemes demonstrate an increase in the end-to-end queuing delay perceived by the packets of F (see Fig. 2), as the capacity of the $E_5 - C_2$ link increases. However, LA-EDF and DBAC demonstrate the best performance (lowest worst-case queuing delay) among the three, even as burstiness increases. In fact, when observing the queuing delay observed for a large number of flows with low delay budget, over different paths, our scheme has an equal or better performance to that of the Calculus scheme while CJVC demonstrates the poorest performance among the three approaches. However, this outcome is somewhat justified as CJVC introduces extra delay as it adds (along the hops that a packet traverses) the

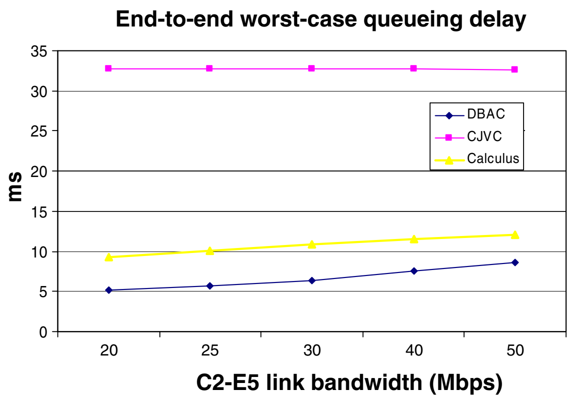

Fig. 2. Performance in terms of end-to-end delay guarantees.

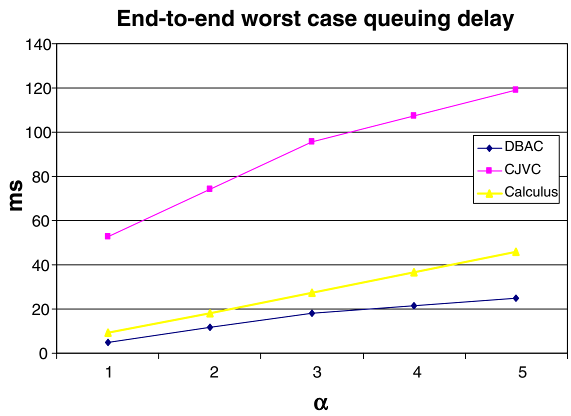

Fig. 3. End-to-end queuing delay vs. increasing burstiness.

difference between the required end-to-end delay for a packet and the slack expected under the guaranteed service rate.

A second set of experiments was conducted along the same lines. In this case the capacity of the $E_5 - C_2$ link was held constant while the a factor in (14) was varied in order to introduce increased burstiness in the output of the policer applied to F. As demonstrated in Fig. 3, the combination of LA-EDF and DBAC continuously outperforms CJVC and the Calculus scheme in terms of the worst-case queuing delay perceived by the packets of F. DBAC ensures that end-to-end queuing delay for Gold packets of a flow with a low delay budget remains low even when traffic is highly bursty, while the Calculus scheme in some cases results in a queuing delay that increases at a higher rate. CJVC demonstrates again the worst performance and seems to be much more vulnerable to burstiness of traffic.

4.2. Using DBAC for queuing delay differentiation

In this set of experiments, the efficiency of DBAC in differentiation of the end-to-end queuing delay perceived by Gold traffic flows is evaluated. Differentiation is based on the end-to-end delay budget DB^j for the packets of each flow j. Thus, DBAC is evaluated against its efficiency in ensuring for each flow a differentiated end-to-end queuing delay based on the flow's DB^j value when compared to the DB values of all other flows with which flow j shares the same path. In essence, DBAC ensures that the packets of a flow with a low delay budget will never encounter higher queuing delay than the packets of a flow with a higher delay budget.

Table 1
Queuing delay tolerance of flows

	Exp. #1 (ms)	Exp. #2 (ms)	Exp. #3 (ms)	Exp. #4 (ms)	Exp. #5 (ms)
Flow 1	25	30	55	80	130
Flow 2	130	80	55	45	40

For this experiment, we have monitored the perceived worst-case queuing delay by two different flows sharing the same path ($E_6 - C_3 - C_1 - E_1$) in the topology of Fig. 1. In Table 1, the queuing delay tolerance ($DB^j - DB_p$) of each of the flows in a series of 5 experiments is provided.

As can be seen from Fig. 4, LA-EDF and DBAC are successful in achieving relative differentiation in terms of the queuing delay perceived by Gold traffic flows sharing the same path. When the two flows have the same queuing delay tolerance value (in the case of Experiment 3), the worst-case queuing delay perceived by the packets of each flow is comparable for the two flows. As the difference between $DB^j - DB_p$ for each flow increases, so is the difference in the queuing delay perceived by the packets of each flow.

A side-observation for this experiment and the ones to follow is that the actual queuing delay observed is much lower than the theoretical queuing delay tolerance of each flow. This is the effect of over-provisioning that is inherent in the Gold service model as well as of the statistical multiplexing of traffic. This phenomenon is observed in equivalent service models in the literature, as such models employ over-provisioning in order to provide hard worst-case delay bounds, valid even in highest congestion periods of short duration, due to the critical

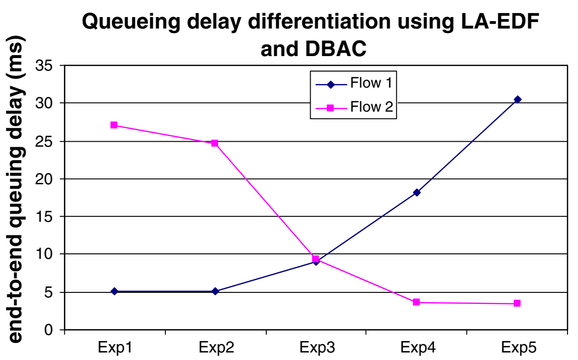

Fig. 4. LA-EDF and DBAC effectiveness in perceived queuing delay differentiation.

Table 2
Queuing delay tolerance of flows

	Exp. #1 (ms)	Exp. #2 (ms)	Exp. #3 (ms)	Exp. #4 (ms)	Exp. #5 (ms)
Flow 1	30	40	50	60	70
Flow 2	80	70	60	50	40

QoS requirements of traffic flows. This performance comes in expense of the resources allocated for such services, however this is a fair price to pay in return of the best possible QoS perceived. Our model does not escape this rule, however adds the delay-based differentiation to existing approaches.

We conducted similar experiments to assess the delay differentiation effectiveness of our proposed scheme in the cases where the flows share the same source and destination edge node but are routed through different paths of the topology. In this experiment, the queuing delay tolerance values for each of the two flows are set to the values of Table 2, however due to our proposed domain-wide optimisation of path selection in Section 3.3, Flow 1 is served through path $E_3 - C_1 - C_2 - E_5$ while Flow 2 is served through path $E_3 - C_1 - C_3 - C_2 - E_5$.

As demonstrated in Fig. 5, our proposed mechanisms still achieve relative differentiation on the queuing delay perceived, based on the queuing delay tolerance of flows. However, as the load of the paths serving the individual flows may significantly differ, like in the case of the current experiment, it is not trivial to achieve comparable queuing delay values for flows with the same delay tolerance but served through different paths. We aim to work on this issue as part of our future work.

Although not depicted in Figs. 4 and 5, it is obvious that the CJVC and Calculus schemes do not provide any kind of differentiation of the queuing delay perceived among different flows in either of the cases examined here.

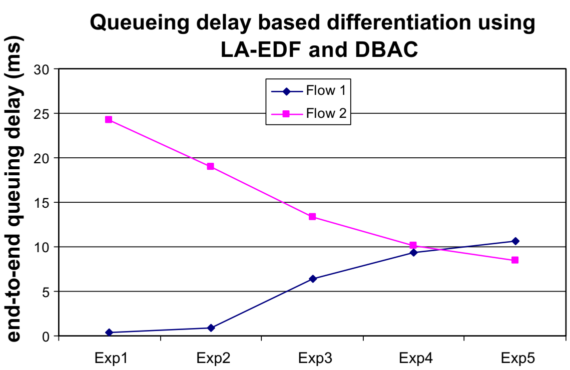

Fig. 5. LA-EDF and DBAC effectiveness in perceived queuing delay differentiation over different paths.

4.3. Path selection evaluation

For the evaluation of our proposed domain-wide optimisation of path selection in Section 3.3, targeting at Load Balancing (LB) as opposed to Shortest Path (SP) routing, we have conducted a set of experiments on the topology of Fig. 6. The topology depicts the core of the pan-European Research Network GEANT2 [20], interconnecting 16 PoPs in equivalent countries. GEANT2 provides IP connectivity among its PoPs and implements QoS mechanisms for serving QoS-demanding traffic.

Except for using a more complex topology, the experiments use the traffic models and parameters described in the beginning of Section 4. In order to achieve a feasible simulation scenario, in terms of the number of packets handled by the simulator, the value of C_i^G for each link in the experimental setup was fixed to 100 Mbps rather than the 1 Gbps value that GEANT2 uses for QoS-sensitive traffic in reality. The experiments concern varying the mean inter-arrival times of flows requesting service and the mean duration of flows that are accepted and served in the domain as seen in Tables 3 and 4 respectively. For the flows' admission control and queuing, DBAC and LA-EDF were used.

It is obvious from the results in Tables 3 and 4 that the Load Balancing approach achieves higher utilization of the available resources than traditional Shortest Path routing. This comes as no surprise in our case. However, it has to be pointed out

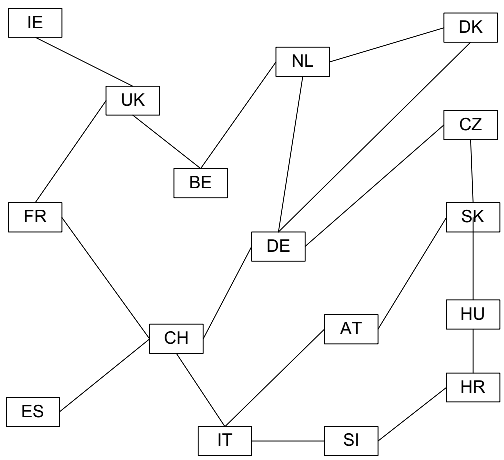

Fig. 6. The topology used to evaluate the domain-wide optimisation for path selection.

Table 3
Flow acceptance ratio with varying inter-arrival times

	Path selection algorithm							
	SP	LB	SP	LB	SP	LB	SP	LB
Inter-arrival mean	5 ms		10 ms		20 ms		30 ms	
Schedulability ratio	75%	89%	78%	90%	84%	92%	91%	95%

Table 4
Flow acceptance ratio with varying flow durations

	Path selection algorithm							
	SP	LB	SP	LB	SP	LB	SP	LB
Duration mean	10 s		20 s		30 s		50 s	
Schedulability ratio	89%	95%	84%	92%	78%	90%	61%	83%

that the proposed LB technique comes at no expense in terms of the QoS perceived by the individual flows, as LA-EDF and DBAC ensure that the bandwidth and end-to-end delay requirements of each accepted flow are never violated and the advantages of our proposed mechanisms in terms of performance and delay-based differentiation, as presented in the previous sections, still apply.

5. Implementation issues

The introduction of the proposed mechanisms in a production environment raises a number of implementation issues for the realization of the Gold service. However, these issues fulfil the requirements for keeping complexity away from the core of the network, thus achieving a scalable and practical service model.

LA-EDF only assumes the support of EDF scheduling at core routers. However, packet deadlines are computed and inserted in packet headers using DPS at the traffic ingress nodes. The computation of the slack terms occurs once for all the packets of a Gold flow, at the moment when the request for serving the flow arrives. The resulting values are then used to determine each packet's deadline for the entire lifetime of a Gold flow. The implications and implementation issues concerned with introducing DPS for storing service deadlines in packet headers are presented in detail in [4] and will not be repeated here.

Our proposal also assumes the implementation of a centralized Gold service broker, with a domain-wide view of accepted Gold flows and the network topology. The broker provides the input parameters (existing Gold flows' profile rates) for the packet deadlines' determination at each node according to LA-EDF and executes the DBAC operations upon arrival of a new request. It also applies the Load Balancing algorithm for selection of the path over which an arriving flow will be served.

Due to the fact that path selection is not based on shortest paths, explicit routing of Gold flows is required. This requirement can be fulfilled with the use of Multi-Protocol Label Switching (MPLS) as a traffic engineering technique. In MPLS, labels are assigned to packets and used in the forwarding process in each router to explicitly define the next hop in the path from the source (ingress router) to the destination (egress router). All Gold traffic flows using the same path can be explicitly routed over the same MPLS Labelled Switched Path (LSP) and thus comprise a Forwarding Equivalence Class (FEC) according to the MPLS terminology.

6. Future work

As part of our future work, we initially aim to investigate methodologies for optimising the performance of LA-EDF in terms of how the delay budget of a flow along the path is partitioned. In this work we proposed partitioning in proportion to the service load at each node, based on the intuition that the queuing delay is increased by higher subscription of flows on a certain queue. We aim to define alternative ways to partitioning, by examining different correlations of the queuing delay and the load imposed on each Gold queue.

Furthermore, we aim to investigate the use of statistical models (such as effective envelopes) for modelling traffic arrivals, rather than the deterministic envelopes used in this work. These models will be then used by DBAC to provide admission control in an effort to achieve higher schedulability.

It is also important to further investigate how we can further optimise delay differentiation for flows which share the same source and destination but follow different paths or, in general, for flows with the same delay budget, regardless of entrance-exit points and the path used to serve them, over a certain topology.

Finally, we plan to conduct an experimental evaluation of the proposed scheme in terms of the overhead imposed by its implementation in the operations of an IP network.

7. Conclusions

In this work, we propose an efficient EDF scheduling algorithm as well as an admission control scheme for implementing a QoS service for delay-sensitive traffic in an IP network. We also provide a simple methodology for domain-wide path selection optimisation having as ultimate goals those of higher schedulability and load balancing.

The proposed mechanisms are efficient but also scalable as all the information required for their operation can be kept at the traffic ingress nodes and encoded in the packet headers. No state information or configuration changes are required in core nodes of the topology. The performance of our approach has been experimentally validated against previous related work.

The QoS service implemented offers guaranteed service rate, bounded end-to-end delay, minimal jitter and no packet loss to aggregated IP traffic, while at the same time demonstrating efficiency in differentiation of the worst-case end-to-end delay perceived by flows with different requirements. This is a novel feature in relation to existing approaches and is achieved while demonstrating comparable or in certain cases better performance than existing QoS schemes.

References

[1] C. Fraleigh, F. Tobagi, C. Diot, Provisioning IP backbone networks to support latency sensitive traffic, in: Proc. of IEEE INFOCOM'03, vol. 22 (1), 2003, pp. 375–385.

[2] S.D. Patek, J. Liebeherr, E. Yilmaz, Tradeoffs in designing networks with end-to-end statistical QoS guarantees, Telecommunications Systems 23 (1) (2003) 9–34.

[3] M. Fidler, V. Sander, A parameter based admission control for differentiated services networks, Elsevier Computer Networks Journal, Special issue on QoS in Multiservice IP Networks 44 (4) (2004) 463–479.

[4] I. Stoica, H. Zhang, Providing guaranteed services without per flow management, ACM SIGCOMM Computer Communication Review 29 (4) (1999) 81–94.

[5] K. Zhu, Y. Zhuang, Y. Viniotis, Achieving end-to-end delay bounds by EDF scheduling without traffic shaping, in: Proc. of IEEE INFOCOM'01, (1), 2001, pp. 1493–1501.

[6] C. Bouras, A. Sevasti, A delay-based analytical provisioning model for a QoS-enabled service, in: Proc. of the 2006 IEEE International Conference on Communications (ICC), 2006.

[7] B. Davie, A. Charny, J.C.R. Bennet, K. Benson, J.Y. Le Boudec, W. Courtney, S. Davari, V. Firoiu, D. Stiliadis, RFC 3246: An Expedited Forwarding PHB (Per-Hop Behavior), March 2002 (Obsoletes RFC2598).

[8] V. Sivaraman, F.M. Chiussi, M. Gerla, End-to-end statistical delay service under GPS and EDF scheduling: a comparison study, in: Proc. of IEEE INFOCOM'01, (1), 2001, pp. 1113–1122.

[9] D. Nandita, J. Kuri, H. Jamadagni, Optimal call admission control in generalized processor sharing (GPS) schedulers, in: Proc. of IEEE INFOCOM'01, (1), 2001, pp. 468—477.

[10] K. Gopalan, T. Chiueh, Y. Lin, Delay budget partitioning to maximize network resource usage efficiency, in: Proc. of IEEE INFOCOM'04, vol. 23 (1), 2004, pp. 2061–2072.

[11] C. Li, E. Knightly, Coordinated multihop scheduling: a framework for end-to-end services, IEEE/ACM Transactions on Networking (TON) 10 (6) (2002) 776–789.

[12] M. Andrews, Probabilistic end-to-end delay bounds for earliest deadline first scheduling, in: Proc. of IEEE INFOCOM'00, (1), 2000, pp. 603–612.

[13] T.K. Lee, M. Zukerman, R.G. Addie, Admission control schemes for bursty multimedia traffic, in: Proc. of IEEE INFOCOM'01, (1), 2001, pp. 478–487.

[14] J. Qiu, E. Knightly, Measurement-based admission control with aggregate traffic envelopes, IEEE/ACM Transactions on Networking (9 2) (2001) 199–210.

[15] K. Gopalan, T. Chiueh, Y. Lin, Probabilistic delay guarantees using delay distribution measurement, in: Proc. of the 12th Annual ACM International Conference on Multimedia, 2004, pp. 900–907.

[16] K. Gopalan, T. Chiueh, Y. Lin, load balancing routing with bandwidth-delay guarantees, IEEE Communications Magazine 42 (6) (2004) 108–113.

[17] C. Bouras, A. Sevasti, Analytical approach and verification of a DiffServ-based priority service, in: Proc. of 6th IEEE International Conference on High Speed Networks and Multimedia Communications, 2003, pp. 11–20.

[18] C. Bouras, D. Primpas, A. Sevasti, A. Varnavas, Enhancing the DiffServ architecture of a simulation environment, in: Proc. of 6th IEEE International Workshop on Distributed Simulation and Real Time Applications, 2002, pp. 108–118.

[19] S. McCanne, S. Floyd, 'ns Network Simulator'. Available from: <http://www.isi.edu/nsnam/ns/>.

[20] GÉANT2: The pan-European research and education network. <http://www.geant2.net/>.

Christos Bouras obtained his Diploma and Ph.D. from the Computer Science and Engineering Department of Patras University (Greece). He is currently an Associate Professor in the above department. Also he is a scientific advisor of Research Unit 6 in Research Academic Computer Technology Institute (CTI), Patras, Greece. His research interests include Analysis of Performance of Networking and Computer Systems, Computer Networks and Protocols, Telematics and New Services, QoS and Pricing for Networks and Services, e-learning, Networked Virtual Environments and WWW Issues. He has extended professional experience in Design and Analysis of Networks, Protocols, Telematics and New Services. He has published 150 papers in various well-known refereed conferences and journals. He is a co-author of five books in Greek. He has been a PC member and referee in various international journals and conferences. He has participated in R&D projects such as RACE, ESPRIT, TELEMATICS, EDUCATIONAL MULTI-

MEDIA, ISPO, EMPLOYMENT, ADAPT, STRIDE, EUROFORM, IST, GROWTH and others. Also he is member of, experts in the Greek Research and Technology Network (GRNET), Advisory Committee Member to the World Wide Web Consortium (W3C), Task Force for Broadband Access in Greece, ACM, IEEE, EDEN, AACE and New York Academy of Sciences.

Afrodite Sevasti obtained her Diploma and her Master's degree (M.Sc.) from the Computer Engineering and Informatics Department of Patras University in Greece. She also holds a Master's Degree in Information Networking from Carnegie Mellon University. She is a Ph.D. candidate at the same department of Patras' University. She has worked as an R&D Computer Engineer at the RA Computer Technology Institute (Greece) and she is currently with the Greek Research and Technology Network (GRNET) S.A. Her main interests and expertise lie in the fields of Computer Networks, Telematics, Distributed Systems and especially in technologies and architectures of high performance networks, in traffic and network resources' management, in Managed Bandwidth Services, provisioning of Quality of Service (QoS), SLAs and pricing/billing of next generation networks. She has published 17 papers in well-known refereed conferences and journals. She has participated in several R&D projects.