Jon Crowcroft James Roberts
Mikhail I. Smirnov (Eds.)

Quality of Future Internet Services

First COST 263 International Workshop, QofIS 2000
Berlin, Germany, September 25-26, 2000
Proceedings

Springer

Series Editors

Gerhard Goos, Karlsruhe University, Germany
Juris Hartmanis, Cornell University, NY, USA
Jan van Leeuwen, Utrecht University, The Netherlands

Volume Editors

Jon Crowcroft
University College London, Department of Computer Science
Gower Street, London WC1E 6BT, United Kingdom
E-mail: jon.crowcroft@cs.ucl.ac.uk

James Roberts
France Telecom R&D
38 rue de Général Leclerc, 92794 Issy-Moulineaux, Cedex 9, France
E-mail: james.roberts@cnet.francetelecom.fr

Mikhail I. Smirnov
GMD FOKUS
Kaiserin-Augusta Allee 31, 10589 Berlin, Germany
E-mail: smirnov@fokus.gmd.de

Cataloging-in-Publication data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Quality of future Internet services : first COST 263 international workshop ; proceedings / QofIS 2000, Berlin, Germany, September 25 - 26, 2000. Jon Crowcroft ... (ed.). - Berlin ; Heidelberg ; New York ; Barcelona ; Hong Kong ; London ; Milan ; Paris ; Singapore ; Tokyo : Springer, 2000
(Lecture notes in computer science ; 1922)
ISBN 3-540-41076-7

CR Subject Classification (1998): C.2, H.4, H.3, J.1

ISSN 0302-9743
ISBN 3-540-41076-7 Springer-Verlag Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable for prosecution under the German Copyright Law.

Springer-Verlag Berlin Heidelberg New York
a member of BertelsmannSpringer Science+Business Media GmbH
© Springer-Verlag Berlin Heidelberg 2000
Printed in Germany

Typesetting: Camera-ready by author, data conversion by DA-TeX Gerd Blumenstein
Printed on acid-free paper SPIN: 10722832 06/3142 5 4 3 2 1 0

Priority Queueing Applied to Expedited Forwarding: A Measurement-Based Analysis

Tiziana Ferrari[1], Giovanni Pau[2], and Carla Raffaelli[2]

[1] INFN - CNAF,
viale Berti Pichat 6/2, I-40127 Bologna, Italy
Tiziana.Ferrari@cnaf.infn.it
[2] DEIS, University of Bologna
viale Risorgimento 2, I-40136 Bologna, Italy
{gpau,craffaelli}@deis.unibo.it

Abstract. The priority queueing mechanism is analysed to verify its effectiveness when applied for the support of Expedited Forwarding-based services in the Differentiated Services environment. An experimental measurement-based methodology is adopted to outline its properties and end-to-end performance when supported in real transmission devices. A test layout has been set up over a metropolitan area for the estimation of one-way delay and instantaneous packet delay variation.
The effect of relevant factors like the buffering architecture, the background traffic packet size distribution and the EF traffic profile are considered. In particular, the complementary one-way delay probability function is computed for a given packet size distribution and the Aggregation Degree parameter is defined to quantify the effect of traffic aggregation on end-to-end QoS.[1]

Keywords: Priority Queueing, Differentiated Services, Expedited Forwarding, Performance measurement, One-way delay, Instantaneous packet delay variation

1 Introduction

The Differentiated Services framework has been recently considered by the scientific community as a scalable solution for the support of Quality of Service to time-sensitive applications. In the Differentiated Services architecture (diffserv) [1,2] traffic is classified, metered and marked at the edge of the network so that streams with similar requirements are placed in the same class, i.e. are

[1] This work has been partially funded by M.U.R.S.T, the Italian Ministry of University and scientific and technological research, in the framework of the research project *Quality of service in multi-service telecommunication networks* (MQOS). This work is a joint activity carried out in collaboration with the TF-TANT task force in the framework of the EC-funded project QUANTUM[17,18,19].

J. Crowcroft, J. Roberts, and M. Smirnov (Eds.): QofIS 2000, LNCS 1922, pp. 167–181, 2000.
© Springer-Verlag Berlin Heidelberg 2000

marked with the same label –the Differentiated Services Code Point (DSCP) [3]–, and treated in an aggregated fashion so that no per-flow differentiation is required.
QoS guarantees are applied to the class on the whole instead of the single streams. Class differentiation is achieved through queueing, which is placed at critical points of the network so that datagrams with identical DSCP are stored in the same queue. Queues are drained according to the service order defined by the scheduling algorithm adopted by the queueing system [4]. Several packet treatments - the so-called Per-Hop Behaviours (PHBs) - have been standardized so far: the Expedited Forwarding PHB (EF) (RFC 2598) for the support of delay - and jitter-sensitive traffic and the Assured Forwarding PHB Group (AF) (RFC 2597) for the differentiation into relative levels of priority. The traditional Best-Effort (BE) packet treatment is an additional valid PHB.

The experimental approach to the problem of delay and jitter offers the challenge of verifying the influence of scheduling [5,6], one of the main QoS building blocks, on end-to-end traffic performance when applied in a production-like environment.

In this paper we study the performance of a specific queueing algorithm: Priority Queueing (PQ), when applied to delay- and jitter-sensitive traffic. Several test scenarios are considered: end-to-end performance is analysed as a function of the background traffic pattern and of the priority traffic characteristics like the frame size, the number of concurrent flows and their profile.
The goal is to identify the requirements for an effective deployment of PQ when applied to the Expedited Forwarding PHB: We derive a general law describing the queueing delay introduced by PQ under different traffic profiles and we evaluate the PQ nodal jitter as a function of the traffic profile when several EF streams run concurrently.

Section 2 introduces the network testbed deployed for end-to-end performance measurement, while in Sections 3 and 4 we provide a high-level description of the queueing system under analysis and of the measurement methodology and metrics adopted in this paper. In Section 5 focuses on the nodal delay introduced by PQ in presence of different best-effort traffic scenarios, while in Section 6 we develop the details of the effect of EF aggregation on both one-way delay and instantaneous packet delay variation. The EF packet size, an additional important factor, is evaluated in Section 7 and the article is concluded by Section 8, in which we summarize the main achievements here discussed.

2 Network Layout

A metropolitan network based on 2 Mbps ATM connections was deployed as illustrated in Figure 1. Packet classification, marking and policing are enabled on router C7200 [2] (experimental IOS version 12.0(6.5)T7) and PQ is the scheduling

[2] No shaping of IP traffic was applied in the test scenarios analysed in this paper. However, ATM shaping was supported in router C7200. The study of the impact of shaping on delay and jitter is subject of future research.

algorithm configured on its output interface. The first router C7200 on the data path is the only one that enables QoS features, while the two remaining routers C7500 are non-congested FIFO devices introducing a constant transmission delay which is only a function of the EF packet size. The two routers C7500 have no influence on jitter, since the minimum departure rate at the output interface is always greater or equal to the maximum arrival rate, as a consequence the input stream profile is not distorted when handled by the C7500s.
The round trip time (RTT) of a 64-byte packet is approximately 2 msec, but RTT linearly increases with the packet size.
The SmartBits 200 by Netcom Systems is deployed as measurement point and traffic generator, while EF and BE background traffic is introduced by test workstations both to congest the egress ATM connection of router C7200 and to create aggregation when needed. Both EF and BE traffic are received by the C7200 from the same interface and they share the same data path to the destination. The SmartBits 200 is a specialized platform which performs packet time-stamping in hardware with a precision of 100 nsec and it is capable of gathering measures for a large range of performance metrics.
Sources and destinations are all located in Site 1 and connected through a switched Ethernet. Both EF and BE traffic are looped back to Site 1 so that the measurement point can be deployed as source and receiver at the same time. In this way accuracy in one-way delay measurement is not affected by clock synchronization errors.

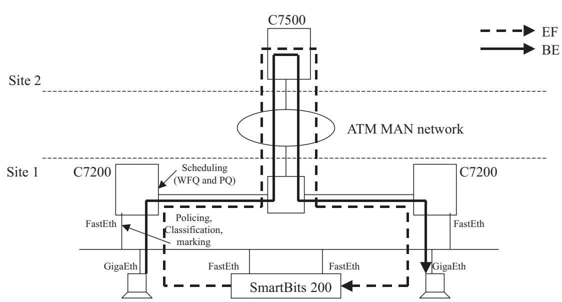

Fig. 1. Diffserv test network

3 Diffserv Node Architecture

In this section we present the queueing architecture of the ingress diffserv node whose performance is the subject of our analysis. The system can be represented as the combination of a queueing system coupled with a FIFO transmission queue (Figure 2).

Queueing system It represents the diffserv nodal component traffic differentiation relies on. In this study we assume that it is based on two scheduling algorithms: *Priority Queueing* and *Weighted Fair Queueing* [13,14,15]. We restrict the analysis to a single-level Priority Queueing model, in which only one priority queue is present, while the Weighted Fair Queueing system can be composed of one or more different queues. Packets are distributed among queues according to the label carried in the IP header, which identifies the Behaviour Aggregate (BA) the packet belongs to.
For simplicity we restrict our analysis to a scenario based on only two per-hop behaviours: the *Expedited Forwarding* PHB (EF) and the *Best-Effort* PHB (BE). PQ is used to handle delay- and jitter-sensitive traffic, i.e. EF traffic, while background BE traffic is stored in a WFQ queue. Nevertheless, results are generally applicable to a WFQ system with multiple queues, since PQ is independent of the number of WFQ queues by definition of the algorithm.
Despite of the above-mentioned restrictions, the WFQ system can be composed of an arbitrarily large number of queues. The combined presence of a PQ queueing module and of a WFQ queueing module in an output interface gives the possibility to support at the same time services for delay and jitter-sensitive traffic as well as services for loss, delay and jitter sensitive traffic.
Queues handle aggregated traffic streams, not per-flow queues, in other terms, each queue in the queueing system illustrated in Figure 2 collects packets from two or more micro-flows. Aggregation is a key feature in diffserv networks introduced to increase the scalability of the architecture.
The length of the priority queue is limited in comparison with the BE queue: the priority queue length is limited to 10 packets, while the best-effort queue length is set to the maximum allowed by the system, i.e. 64 packets. The availability of queue space is not relevant in a high-priority queue since the instantaneous accumulation of data is avoided through policing and shaping in order to minimize the corresponding queueing delay.

Transmission queue It is an additional buffering stage introduced to capture the common architecture of many transmission devices in which the line adapter is separated from the queueing block by an internal transmission element, e.g. a bus. The transmission queue gathers packets issued by the queueing system according to the order defined by the scheduler and it is emptied at line rate. In this study we assume that the transmission queue is serviced in a FCFS fashion as this is the service policy commonly adopted for production devices.
Given the relatively small rate of the wide area ATM connection (2 Mbps), the time needed to dequeue a packet and to place it in the transmission queue is assumed to infinitely small in comparison with its transmission time. In this paper the set of WFQ and PQ queues was configured on an ATM interface and the transmission queue is emptied at line rate, i.e. at the PVC rate, which is 2 Mbps.
Memory in the transmission queue is allocated so that as soon as one unit is freed, the unit can be immediately allocated to store a new packet, the one selected by the first queueing stage for transmission. If the memory unit size is

not enough to store the whole packet, then, additional units can be temporarily added to the transmission queue so that the packet can be entirely stored. This means that if the transmission queue is 5 unit long and the MTU size is 1500 by, then the maximum instantaneous queue length is 7 units when the queue stores two best-effort packets and 1 EF packet.

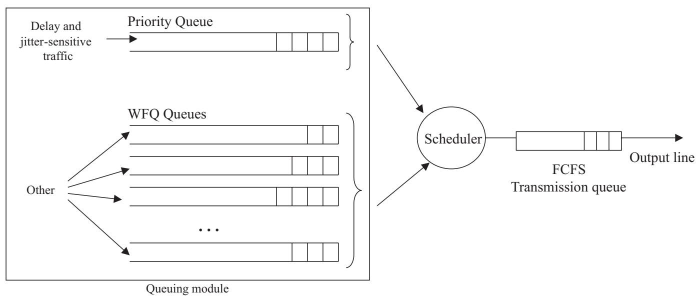

Fig. 2. Architecture of the diffserv node queueing system

The original contribution of this paper comes from the fact that two queueing algorithms (PQ and WFQ) are coupled with a finite-length and discrete transmission queue like in real-life network systems and from the fact that different traffic models are used at the same time to feed different queues.
Analytic studies of both PQ and WFQ can be found in literature [7,8].

4 Measurement Methodology

The priority queueing algorithm is evaluated by focusing on two metrics: *one-way delay* and *instantaneous packet delay variation* (IPDV). The two above-mentioned parameters were selected to verify the suitability of PQ as scheduling algorithm when applied to the EF PHB.

One-way Delay is defined in RFC 2679. This metric is measured from the wire time of the packet arriving on the link observed by the sender to the wire time of the last bit of the packet observed by the receiver. The difference of these two values is the one-way delay. In our experimental scenario one-way delay is derived from the *cut-through latency* measured by the SmartBits 200 according to RFC 1242.

Instantaneous Packet Delay Variation is formally defined by the IPPM working group Draft [9]. It is based on one-way delay measurements and it is defined for (consecutive) pairs of packets. A singleton IPDV measurement requires two

packets. If we let D_i be the one-way delay of the i^{th} packet, then the IPDV of the packet pair is defined as $D_i - D_{i-1}$.
According to common usage, jitter is computed according to the following formula: $jitter = |D_i - D_{i-1}|$.
In our tests we assume that the drift of the sender clock and receiver clock is negligible given the time scales of the tests discussed in this article. In the following we will refer to jitter simply with ipdv.
It is important to note that while one-way-delay requires clocks to be synchronized or at least the offset and drift to be known so that the times can be corrected, the computation of IPDV cancels the offset since it is the difference of two time intervals. If the clocks do not drift significantly in the time between the two time interval measurements, no correction is needed.

One-way delay and IPDV are analysed by computing their frequency distribution over a population of 10000 samples.
When several EF concurrent streams are run in parallel, performance measurement is only applied to one stream, which in this paper is referred to with the term reference stream. Such a stream is generated by the SmartBits through the application SmartWindows 6.53, while additional EF flows are generated through the application *mgen 3.1* [10].
The load of the EF class is kept limited to a small fraction of the line rate (32%). Both the priority queue size and the transmission queue size are constant: the former is equal to 10 packets, the latter to 5 memory units. Both EF and BE streams are constant bit rate, unidirectional UDP flows.
While in principle EF traffic can be both UDP and TCP, we restrict our analysis to UDP streams because we want to study a queueing system which does not include traffic conditioning modules (like shapers and/or policers), which are needed in case of bursty traffic. In this paper we assume that input traffic is correctly shaped and does not exceed the maximum EF reservation.

In an ideal system end-to-end performance of a flow belonging to a given class should be independent of the traffic profile of both its behaviour aggregate and of other behaviour aggregates present in the queueing system.
However, test results show that one-way delay experienced by packets subject to priority queueing is influenced by three main factors:

1. The *packet size frequency distribution* of background traffic,
2. The *instantaneous length* of the priority queue,
3. The *EF packet size*.

In the former case the packet size has an influence on the queueing delay introduced by both the transmission queue and the priority queue itself. In fact, an EF packet has to wait for the completion of the current packet transmission before it is selected next by the scheduler for transmission. However, also the profile of the reference behaviour aggregate can impact the end-to-end performance: In presence of burstiness (for example stemming from the packet clustering introduced by aggregation) the priority queue length can be instantaneously non-zero. As a consequence the nodal queueing delay introduced by the priority queue can be different depending on the position of an EF packet within a burst. In the following sections the impact of the above-mentioned factors on one-way delay is analysed in detail.

5 Packet Size Distribution of Best-Effort Traffic

In this section EF one-way delay is evaluated in presence of different background BA packet size distributions: deterministic and *real* (when background traffic is generated according to a real packet size distribution determined through the monitoring of production traffic).
The test characteristics are summarized in Table 1. In order to reduce the complexity of our analysis only two BAs run in parallel: a best-effort BA composed multiple best-effort micro-flows, each issuing data at constant rate, and a single constant bit rate EF flow, serviced by the priority queue.

Table 1. One-way delay test parameters with different background traffic profiles

EF traffic			**BE traffic**		
Load (Kbps)	*Frame Size* (bytes)	*Prot*	*Load* (Kbps)	*Frame Size Distribution*	*Prot*
300	128	UDP	> 2000	Deterministic, Real	UDP

5.1 Constant Packet Size

To begin with, an ideal test scenario has been considered in which the BE traffic is characterized by multiple streams each issuing packets of identical constant fixed payload length in the range: [100, 1450] by. This case study is of interest to show the importance of the background packet size in presence of transmission queue-based systems (Figure 2), in which the queue space is allocated in units of fixed length. This length is equal to 512 by in this paper. The EF packet always spans a single memory unit, while the BE packet can occupy one or more depending on its size.
One-way delay is made of several components:

1. *PQ waiting time*: the time which elapses between arrival time and the time the packet moves from the first queueing stage to the FIFO transmission queue. With PQ the maximum waiting time corresponds to the transmission time of a MTU best-effort packet.
2. *FIFO queueing time*: it represents the time needed wait until a packet is dequeued from the FIFO transmission queue. It depends on the maximum length of the transmission queue, which in our test scenario is equal to 7 memory units of 512 by.

3. *transmission time*: the time needed to place the packet on the wire, it is a function of the line speed (2 Mbps) and of the packet size. The end-to-end transmission time is additive and depends on the number of hops.

As Figure 3 shows, the one-way delay sample mean curve is non-monotone: the discontinuity points correspond to specific BE payload sizes, namely 485 and 1000 by, i.e. for integer multiples of the transmission queue memory unit size, as configured in this test (512 by).

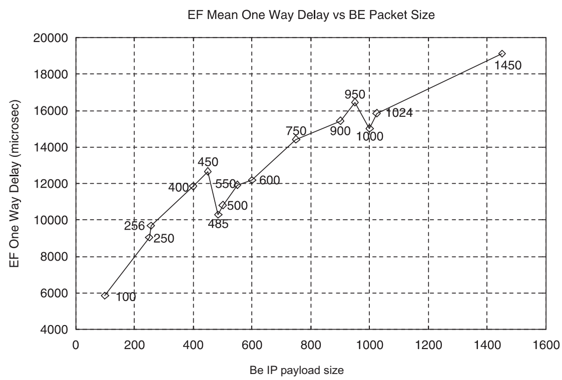

Fig. 3. One-way delay sample mean for different BE payload packet sizes

The overall increase in one-way delay is due to the small EF rate (15% of the line capacity) in comparison with the BE rate. When an EF packet arrives into the priority queue and a BE packet is under transmission, the EF waiting time varies in relation with the BE packet length, being the transmission queue occupancy characterized by a miz of BE and EF packets. In any case if a single memory unit is allocated to best-effort traffic and the BE packet size increases, the average queueing delay introduced by PQ increases, with a maximum when a BE datagram completely fills the memory unit.

However, if the packet size is such that a single memory unit is not sufficient to store it, an additional partially filled unit has to be allocated to store the remaining part of the BE packet. As a consequence, queue memory gets fragmented as completely filled units are followed in the queue by non-complete ones. The time needed to empty the transmission queue is less and the average EF queueing delay introduced by the transmission queue decreases. After a minimum in the curve, i.e. when the BE packet length further increases, the second memory unit increasingly gets more completed and the fraction of queue space allocated to BE packets becomes greater with a consequent increase in delay.

A worst case evaluation of the average EF one-way-delay has been performed using the general model for priority queueing presented in [7]. By assuming all units in the transmission queue completely full the local maxima of Figure 3 are validated by the delay formula presented in [7].

5.2 Real BE Packet Size Distribution

While in the previous section the analysis focuses on statistical distributions of the BE packet size, in this section performance is estimated when the BE packet size is modelled according to the *real* cumulative distribution plotted in Figure 4. In what follows we call it the *real distribution*.

We computed this frequency over a population of more that 100 billion of packets belonging to traffic exchanged on an intercontinental connection in both ways. For the following test scenarios it was chosen as reference distribution to model packet size according to the pattern of traffic exchanged at typical potential congestion points.
In this study, packet size is the only traffic parameter which is considered for real traffic emulation, since we focus on the worst-case scenario in which the queueing system under analysis described in section 3 is assumed to be under permanent congestion. As such, rate variability and autocorrelation of best-effort traffic are not considered, even if they can influence the performance of the system when deployed in production.

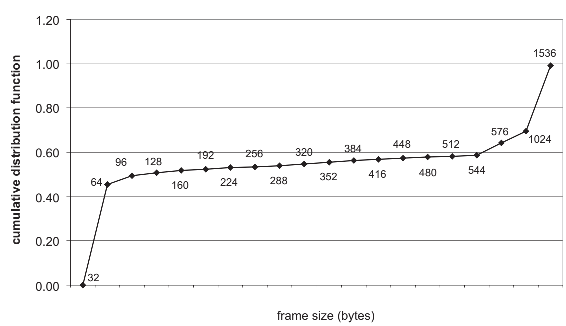

Fig. 4. *Real* cumulative BE packet size distribution

Figure 5 plots the complementary probability distribution we derived from the frequency distribution computed experimentally during the test session, i.e. the probability that the delay d experienced by a packet is greater than a given value D ($p(d) \geq D$). We express the variable delay in transmission units, i.e. in integer multiples of the transmission time of an EF packet of known size at a given line rate (for an EF payload size of 128 by it corresponds to 0.636 msec). Figure 5 shows that in the system under analysis we can assume that the probability that one-way delay is greater than 36 transmission units is negligible. This threshold can be adopted as an upper bound of the playout buffer size used by interactive multimedia applications and is useful for the optimization of memory utilization.

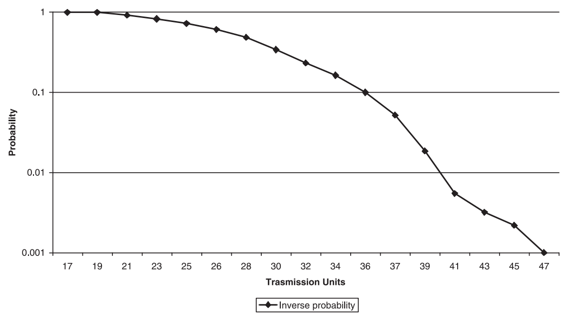

Fig. 5. EF complementary delay probability for a real BE packet sizes distribution

6 Performance with EF Aggregation

Aggregation is one of the fundamental properties which characterize the differentiated services architecture and we want to estimate its impact to verify in which cases and to which extent the differentiated services can provide effective end-to-end services as opposed to the integrated services, which can provide per-flow performance guarantees through signalling.

In the previous scenarios PQ performance was analysed under the assumption that at any time its occupancy is not greater than one packet (represented by the datagram waiting for the end of the current BE transmission). This hypothesis holds only when input traffic is evenly shaped. However, in presence of bursty traffic the nodal delay introduced by the priority queue becomes non-negligible. Burstiness can stem from traffic aggregation, i.e. from packet clustering, which occurs when packets of the same BA arrive at around the same time from different input interfaces and are destined to the same output interface. Even a single source injecting multiple streams can be a potential source of burstiness since independent streams are not synchronized and they can produce instantaneous bursts of packets, even if each single stream is perfectly shaped.

Test results confirm that in this case the priority queue size can instantaneously hold two or more packets and performance depends on the percentage of traffic injected by a given source and by its size. Results show that in absence of shaping and policing aggregation can propagate through a wide area network by increasing the burst size step by step in an avalanche fashion, as also confirmed by the simulation results presented in [11].

In this test scenario aggregation is produced by injecting several EF streams from different interfaces.
The reference stream load decreases from 300 Kbps (when the EF reference stream is present) to 50 Kbps (when several EF streams are run in parallel). The test conditions are summarized in Table 2. We introduce a new metric: the *Aggregation Degree* (which we reference with letter A) to quantify the amount of bandwidth shared among different streams within a BA: $A = 1 - \frac{l_{max}}{L_{BA}}$. where l_{max} is the maximum load of a micro-flow and L_{BA} the overall BA load, i.e. the total traffic volume generated by streams belonging to the class). A is equal to 0 if just a single stream is active, while it is close to 1 when only the BA load is divided among many tiny flows.

In this test scenario a decreasing load injected by the reference stream indicates that A increases, in fact competing EF streams consequently issue a greater amount of traffic and this implies a higher packet clustering probability. BE traffic is modelled according to the real distribution described in Par. 5.2.

Table 2. Test parameter for IPDV under different aggregation patterns

EF traffic (UDP)				BE traffic (UDP)		
BA Load (Kbps)	*Number of streams*	*Ref. stream load*	*Ref. stream frame size*	*BA Load* (Kbps)	*Number of streams*	*Frame size distribution*
300	variable	[50, 300] Kbps	128 by	> 2000	20	real, [0,1500]

6.1 One-Way Delay

Delay distributions measured in presence of multiple EF flows but with constant BA load show that when the aggregation degree A increases, the delay distribution gets more spread giving rise to greater delay variability and larger average one-way delay values. In particular for A equal to $\frac{5}{6}$, $\frac{2}{3}$, $\frac{1}{3}$ and 0, the average delay is equal to 18.29, 17.66, 17.23 and 16.75 msec respectively.
Thus, we can conclude that also the aggregation degree is a design parameter that needs to be upper bounded in order to achieve acceptable performance. This is important during the admission control phase.

In Figure 6 the complementary probability distributions derived from the above-mentioned frequency distributions are plotted. The graphs show that for a given delay D the probability significantly varies with the aggregation degree A. If stringent design requirements need to be met, then the number of flows must be bounded through admission control.

The impact of the EF stream on burstiness depends on the number of hops as explained in [12]. In fact, the presence of a single aggregation point limits the burst accumulation phenomenon, which is visible in multi-hop data-path with a chain of aggregation and congestion points. In addition, the presence of a short

EF queue (10 packets in this test scenario) limits the maximum instantaneous burst to 1280 by which corresponds to a maximum queueing delay of 5.12 msec.

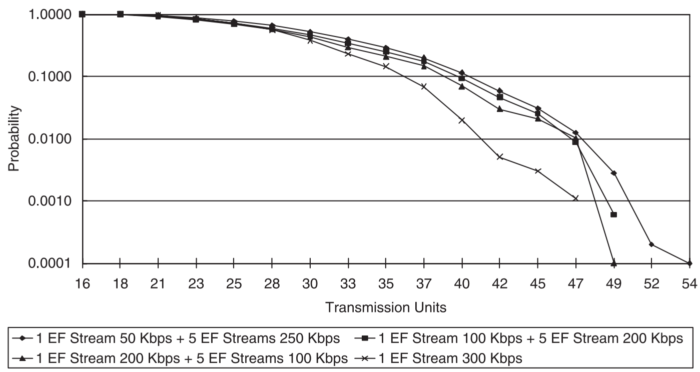

Fig. 6. Complementary one-way probability functions for different Aggregation Degrees

6.2 IPDV

IPDV frequency distribution curves were calculated for different aggregation patterns with a reference stream rate decreasing from 300 Kbps to 50 Kbps and a constant aggregate EF load equal to 300 Kbps.

Distributions[3] present two peaks which are presumably due to the real distribution of the BE packet size, which is mainly concentrated around two packet size values.

In this test the transmission time of a reference packet is 0.636 msec and the maximum IPDV was 30 times the transmission time of the packet itself, independently of the aggregation pattern. However, for a higher value of A IPDV decreases more rapidly and is more densely distributed around a small value.

High aggregation degrees produce an increase in IPDV, in fact for an aggregation degree A equal to $\frac{2}{3}$, $\frac{1}{3}$ and 0 the average IPDV value is equal to 6.05, 5.16 and 3.63 msec respectively.

7 EF Packet Size

IPDV increases with the EF packet size. In this test only one EF stream is generated (A is equal to 0) and used as reference stream. The frame size (i.e. the EF IP packet size plus the layer 2 overhead is constant for a given test, and varies in the range: [128, 512, 1024] by as indicated in Table 3. IPDV frequency

[3] For more information about these distributions and graphs we refer to the long version of this paper.

Table 3. Test parameters for IPDV with different EF frame sizes

EF traffic (UDP)				BE traffic (UDP)		
BA Load (Kbps)	*Number of streams*	*Ref. stream load*	*Ref. stream frame size*	*BA Load* (Kbps)	*Number of streams*	*Frame size distribution*
300	1	300 Kbps	128, 512, 1024 by	> 2000	20	real, [0,1500]

distribution is better for smaller EF frame sizes as illustrated by the IPDV frequency distribution curves in Figure 7: For example with frame size equal to 128 by, 60% of the IPDV values are concentrated in a single interval.
The standard deviation increases with the packet size: for packets of 128, 512 and 1024 by it takes the values: 2016.9, 2500.3 and 3057.0 respectively.

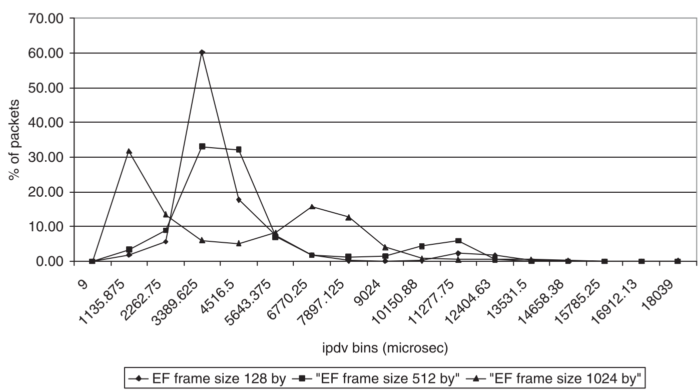

Fig. 7. IPDV frequency distribution *vs* EF frame sizes (BE size distribution: real)

8 Conclusions and Future Work

This paper provides an in depth measurement-based analysis of the Priority Queueing algorithm when adopted for the support of end-to-end QoS to delay- and jitter-sensitive applications, which require EF-based services. Experimental results show that the end-to-end performance of the queueing system is strongly related to the system components - for example the additional FCFS buffering stage (the transmission queue) - as well as to traffic-related factors like the traffic pattern of active BAs and the traffic profile of the EF stream itself. While Priority Queueing proved to be one of the most effective algorithms to minimize queueing delay, a careful system design should be adopted in order to provide strict one-way delay and IPDV guarantees.

The transmission queue introduces an additional contribution to the nodal delay produced by the priority queue, a delay that is proportional to the average packet size in the background BAs. As such, the transmission queue size should be limited.
The one-way delay distribution of EF traffic is strongly dependent on the background traffic distribution and in this study we have compared two cases : the uniform and the real distribution. Generally speaking for larger background traffic packet sizes the one-way delay standard deviation increases as values are more spread over a large range. For a given distribution and a given EF packet size the complementary delay probability can be computed for delay estimation in complex network scenarios and system design purposes.

Background traffic profile is not the only relevant factor: Both the flow and the BA profile can impact performance.
Firstly, the average packet size of a flow is such that for larger datagrams IPDV standard deviation increases. In the second place, stream aggregation within a class has an influence on both the end-to-end one-way delay and IPDV experienced by each flow. In this paper we define the Aggregation Degree parameter to describe the load partitioning between flows within a class and we express performance as a function of it. A high aggregation degree produces an increase in average one-way delay and IPDV. As a consequence, in case of stringent delay and jitter requirements, aggregation has to be limited. The dimensioning of aggregation is subject of future research: The effect of the number of streams, of the nominal rate and of packet size will be investigated.

References

1. *An Architecture for Differentiated Services*, RFC 2475 167
2. Nichols, K., Jacobson, V., Zhang, L.: *A Two-bit Differentiated Services Architecture for the Internet* 167
3. *Definition of the Differentiated Services Field (DS Field) in the Ipv4 and Ipv6 Headers*; RFC 2474 168
4. Bernet, Y., Smith, A., Blake, S.: *A Conceptual Model for Diffserv Routers*, diffserv draft, work in progress. 168
5. Zhang, H.: *Service Disciplines For Guaranteed Performance Service in Packet-Switching Networks* 168
6. Ferrari, T., Chimento, H., P.: *A Measurement-Based Analysis of Expedited Forwarding PHB Mechanisms*, Proceedings of the Eighth Int. Workshop on Quality of Service, IWQoS 2000, Page(s) 127-137 168
7. Bertsekas, D., Gallager, R.: *Data Networks*, Prentice Hall, Page(s) 203-205 171, 174, 175
8. Kleinrock, L.: *Queueing Systems*, John Wiley & Sons, Page(s) 106-147 171
9. Demichelis, C., Chimento, P.: *Instantaneous Packet Delay Variation Metric for IPPM*, ippm draft, work in progress. 171
10. *The Multi-Generator (MGEN) Toolset*, Naval Research Laboratory (NRL), http://manimac.itd.nrl.navy.mil/MGEN/ 172
11. Charny A.: *Delay Bounds in a Network with Aggregate Scheduling* 176

12. Ferrari, T.: *End-to-End perfomance Analysis with Traffic Aggregation*, TNC'2000 177
13. Law, K. L. E.: *The bandwidth guaranteed prioritized queuing and its implementations law*, Global Telecommunications Conference, 1997. GLOBECOM '97, IEEE Volume: 3, 1997, Page(s) 1445-1449 170
14. Rönngren, R., Ayani, R.: *A comparative study of parallel and sequential priority queue algorithms*, ACM Trans. Mod. Comp. . Sim. 7,2 (Apr 1997), page(s) 157-209 170
15. Parekh, A., Gallager, R.: *A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Single-Node Case*; IEEE/ACM Transactions on Networking, Vol 1, No 3, June 1993 170
16. Floyd, S, Jacobson, V.: *Link-sharing and Resource Management Models for Packet Networks*; ACM Transactions on Networking, Vol 3 No. 4, Aug 1995
17. Ferrari, T.: *Differentiated Service Experiment Report*, TF-TANT interim report Jun 99 - Sep 99, http://www.cnaf.infn.it/ ferrari/tfng/ds/del-rep1.doc 167
18. *The Joint DANTE/TERENA Task Force TF-TANT*; http://www.dante.net/tf-tant/ 167
19. *Quantum Project: the Quality network Technology for User-Oriented Multi-Media*; http://www.dante.net/quantum/ 167