Implementation of proportional loss rate differentiation in EDS using Proportional Loss Rate and RED

Benjamin Gaidioz, Pascale Primet

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Implementation of proportional loss rate differentiation in EDS *using Proportional Loss Rate and* RED

Benjamin Gaidioz, Pascale Primet

N° 5029

December 2003

THÈME 1

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--5029--FR+ENG

Implementation of proportional loss rate differentiation in EDS using Proportional Loss Rate and RED

Benjamin Gaidioz*, Pascale Primet†

Thème 1 — Réseaux et systèmes
Projet RESO

Rapport de recherche n° 5029 — December 2003 — 10 pages

Abstract: The TCP/IP stack has been mainly designed for elastic traffic (file transfers). It is nowadays recognized that it is not able to efficiently support traffic patterns with completely differing requirements (e.g. applications with delay requirements). Service differentiation at the flow agregate level (DiffServ) is a promising way to implement some form of IP QoS because it is robust and scalable. The EDS PHB is a Diffserv PHB based on both loss rate and delay proportional differentiation. In this report, we present and compare two implementations of the loss rate differenciation in EDS. One is based on the PLR tail-drop like system, the second uses a RED like dropping mechanism.

Key-words: Diffserv, EDS, loss rate differenciation

* Université Lyon 1, LIP RESO (UMR CNRS-INRIA-ENS-UCBL N°5668)
† LIP RESO (UMR CNRS-INRIA-ENS-UCBL N°5668), INRIA

Unité de recherche INRIA Rhône-Alpes
655, avenue de l'Europe, 38330 Montbonnot-St-Martin (France)
Téléphone : +33 4 76 61 52 00 — Télécopie +33 4 76 61 52 52

Mise en œuvre de la différenciation proportionnelle en taux de perte dans EDS avec Proportional Loss Rate et RED

Résumé : La pile de protocoles TCP/IP a été conçue principalement pour transporter du trafic « élastique » (transferts de fichier). Aujourd'hui, il est admis qu'elle ne convient pas pour de nouvelles catégories de trafic aux besoins très différents (par exemple des flux temps-réel avec des contraintes de délais). La différenciation de services au niveau d'agrégats de flux (DiffServ) est une architecture prometteuse pour mettre en oeuvre une forme de qualité de service IP car elle est robuste et s'étend bien à un grand nombre de noeuds. EDS est un PHB Diffserv qui s'appuie sur la différenciation proportionnelle en délai et taux de perte. Dans ce rapport, nous présentons et comparons la mise en œuvre de la différenciation en taux de perte dans EDS avec le système de gestion de file PLR type *tail-drop* et PLR type RED.

Mots-clés : Diffserv, EDS, différenciation en taux de perte

1 Introduction

The network layer of Internet (IP) provides a simple, robust and effective packet forwarding service. The TCP/IP stack has been mainly designed for elastic traffic (e.g. FTP). It is nowadays recognized that it is not able to efficiently support new traffic patterns with different requirements in terms of speed, latency and reliability (e.g. real-time applications, interactive applications, bulk transfers, etc.). It is commonly accepted that IP needs to be extended with a form of quality of service (QoS). In the case of IP, adding QoS has to be done carefully in order to maintain the efficiency and robustness of the network.

The DiffServ architecture [1] is convincing as a QoS approach on a large scaled network. Its principles keep the network layer simple and robust. In the core network, IP packets are expected to be marked with a specific *class identifier*. This identifier selects a forwarding treatment met by the packet each time it crosses a router. From the core network point of view, there is a very small number of classes, there is no resource reservation between hops. Marking is done at the edge routers, where the number of flows is sufficiently low to apply marking rules without losing much performance.

A characteristic that makes TCP/IP so appealing is its ease of use: a host can plug into the network and use it immediately. According to its design philosophy [2], IP attempted to provide a basic *building block* out of which a variety of types of service could be built. The decision was an extremely successful one, which allowed the Internet to meet its most important goals.

In this report, we rely on the implementation of the EDS service differentiation [8, 7] at the IP level (see in the technical report [7] for details). This architecture provides a best-effort service differentiation which is not sufficient to satisfy most of the applications needs. Thus, it needs to be extended with end-to-end protocols which use the services it provides in order to ensure some higher level guarantees to applications.

The outline of the paper is as follow. In sect. 2 we present EDS. In sect. 3 we present the implementation of EDS with proportional service differentiation. Sect. 4 is related to the implementation of loss rate differentiation with a RED-like system. Then, in sect. 5 we emphasize some similar work to EDS and conclude in sect. 6.

2 Quick overview of EDS

Fig. 1 gives the intuition of the concept of Equivalent Differentiated Services. It shows both the service provided by a best-effort router and an EDS router. Roughly, the EDS router provides a set of N classes, each one experiencing a different level of performance in both queuing delay and loss rate. Considering a single performance criteria (delay or loss rate), the range of performance is centered around an average performance, which is the performance all packets would have obtained through a plain best-effort router. Thus, the performance is directly linked to the router load. Moreover, there is an asymmetry between delay and loss rate performance: The class which obtains the i^{th} best performance in delay obtains the i^{th} worst performance in loss rate.

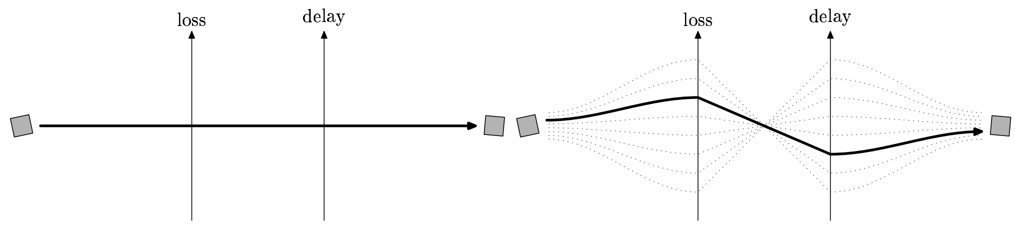

Figure 1: A packet (gray square) is crossing a best-effort router (left) or an EDS router (right). Through the best-effort router, the packet experiences a given queuing delay and a given loss probability. Through an EDS router, the packet is member of a service class among N classes, where it experiences a specific queuing delay and a specific loss probability, both being relatively better or worst than the performance through a best-effort router.

The EDS service differentiation proposal has been designed by starting from the Internet protocol design principles which gave to IP its robustness, ease of deployment and ease of use. Thus, EDS provides *best-effort* service differentiation.

Stronger guarantees on the plain best-effort IP are implemented in end-to-end protocols. For example, TCP guarantees reliability. These protocols use the building block provided by IP to provide a specific service matching the need of specific applications. Since we have defined a best effort service differentiation system from the same design rules, the natural way to implement stronger QoS guarantees has to be done in the protocol layer.

This report does not present a transport protocol on top of EDS but proposes two different implementations of EDS.

3 Implementation of EDS

In this section, we present the implementation of EDS with proportional differentiation schedulers presented in [3]. Delay differentiation is practiced with the WTP scheduler (Waiting-Time Priority). Loss rate differentiation is practiced with the PLR (Proportional Loss Rate) queue management system.

3.1 Implementation of Waiting-Time Priority scheduling

The scheduler is illustrated on fig. 2 with four classes. Each class is associated to a specific queue. When a packet is admitted in the router, the date is stored in it. This date permits to compute the waiting time δ of a packet when a packet is going to be emitted. A coefficient d_i is associated to queue i. When the router has to dequeue a packet, it computes the *normalized* waiting time $\overline{\delta}_i$ of each first packet of each queue by multiplying its real waiting-time δ_i by d_i. The router chooses to dequeue the packet with the highest *normalized* waiting time. Thus, the higher the coefficient d_i is, the quickest the class.

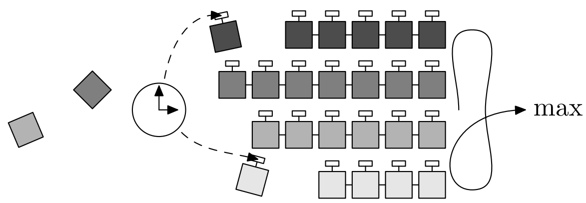

Figure 2: The WTP scheduler for proportional delay differentiation

3.2 Implementation of Proportional Loss Rate

The queue management system is illustrated on fig. 3. The queuing discipline bases its differentiation on the average loss rate of all classes upon the *Loss History Table* (LHT), which gets the loss history of the last K received packets. On fig. 3, it is represented by

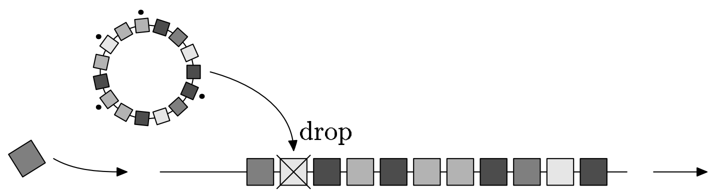

Figure 3: The PLR queuing discipline for proportional loss differentiation

the circular list. When a packet is received, an entry is added in the LHT (by removing the oldest one) which contains the class of the packet (the color of the squares on the figure). Since the packet is in the queue, it is also marked in the LHT as *admitted*. Packet that are later dropped are marked as *dropped* (on the figure, packets which have a little black dot are marked as *dropped*). The average loss rates ρ_i are updated according to new packets admitted in the router and old packets removed from the LHT. When a congestion occurs, the router computes the *normalized* loss rates $\bar{\rho}_i$ by multiplying ρ_i by the l_i coefficient. The class that obtains the lowest $\bar{\rho}_i$ loses a packet while the new one is enqueued. The higher l_i is, the lower probability a class has of losing a packet.

3.3 Combination of both

Since the implementation of EDS requires both schedulers to run at the same time, we have combined them into a single system, shown on fig. 4.

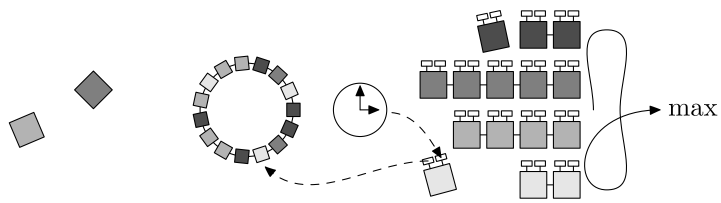

Figure 4: The EDS queuing discipline for delay and proportional loss differentiation

4 Loss rate differentiation implementation

In the next subsections, we run a test with the previous loss rate differentiation scheduler and the new one. EDS has been implemented in NS and runs on a router crossed by random traffic generated according to a Pareto distribution. There are eight EDS classes and eight Pareto sources (one per class). The average load generated by Pareto sources is a bit higher than what the output link can handle, so that packets experiences various queuing delays and random losses probability.

The router queue is 300 packets long. Delay and loss rate are differentiated so that classes obtain a relative delay ranging from 1 to 8 and a loss rate ranging from 8 to 1. The RED mechanisms starts to drop packets when queue length is five packets long with a maximum threshold of 300 (queue capacity) where the maximum average drop probability is 25%.

4.1 With the original implementation of loss rate differentiation

Graphs of fig. 5 shows performance of classes in delays and loss rate. Since differenciation

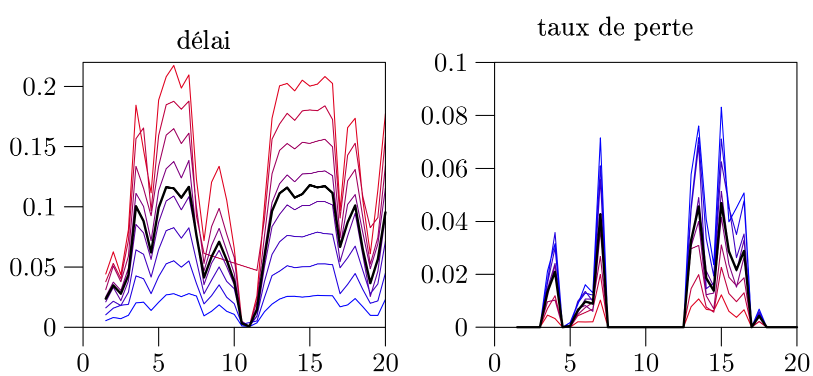

Figure 5: Queuing delays and loss rate for all classes (gray lines) and the average performance (thick lines). We do not show which class is obtaining which performance because this is not the main purpose.

is set up, delay and loss rate are differentiated. Delay differentiation takes place whenever there are more than one packet in the router and more than one class. Even for short queue length, delay differentiation occurs. Loss rate differentiation does not occurs whatever the queue length, it occurs only when the capacity is reached, in the case of what, dropped packets are chosen to ensure the proportionality. This is visible because there are several places where there is no loss rate while queuing delay is not null. This corresponds to places where the queue is not empty and does not reach the capacity.

When there is no loss rate in the queue, all loss rates are equal to zero. This is a particular case where proportional loss rate differentiation is ensured ($0 = l_i \times 0$). However, this is not satisfying because delay differentiation is also ensured but classes get different performance.

4.2 With the RED based implementation of loss rate differentiation

We have modified the original proportional loss rate queue management in order to be more conform to the expected properties of EDS classes. Indeed, the PLR queue management mechanism practices loss rate differentiation using a DropTail with a selective drop between classes that ensures proportionality. Since it drops when DropTail would drop, no loss occurs until the queue capacity is reached. However, delay differentiation is practiced whatever the queue length. Thus, we the queue builds, only delay differentiation is practiced. This is not a satisfactory behavior because it leads to privileged classes (quickest classes). Random early drops have been introduced by including the RED [5] active queue management in the queuing discipline. Proportionality is still ensured by the same selective drop mechanism. Thus, since random early dropping starts when the queue length reaches a low threshold, loss rate differentiation is practiced simultaneously with delay differentiation.

Fig. 6 shows the delay and loss rate evolution for all classes. There is a lower queuing

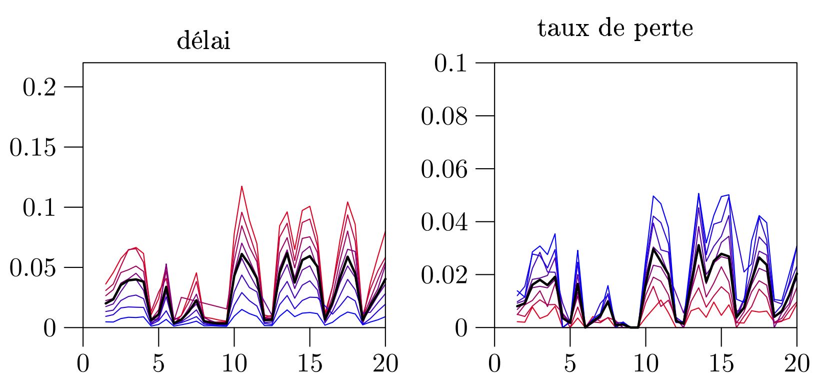

Figure 6: Queuing delays and loss rate for all classes (gray lines) and the average performance (thick lines). We do not show which class is obtaining which performance because this is not the main purpose.

than in the previous experiment delay because of early drops. The main point is that using this implementation, there is an effective loss rate even when the queue capacity is not reached. Moreover, the experienced loss rate is similar to the queuing delay experienced by the packets. We think it is closer to the service we expect to provide with EDS.

5 Related Work

There are two systems based on asymmetry in service differentiation in delay and loss rate.

- The first system is ABE (Alternative Best-Effort) [9, 10, two implementations exist] which provides a low delay class freely available and a TCP class, provided as a default class, where the performance is expected to be at least as good as on a plain best-effort network. The low delay class experiences a higher loss rate. A little drawback of ABE is that it requires all flows to conform to a TCP-friendly profile to ensure its properties.
- A similar system has been defined as Best-Effort Differentiated Services [4]. The system also provides two classes where the low delay class gets a higher loss rate. The main interest of the system compared to ABE is that the low delay class always have resources even if flows are aggressive. Tuning of a router is based on the TCP performance through it. It is possible to implement BEDS on commercial routers because they usually include the schedulers that are needed.

ABE and BEDS do not require the protocol level to be adapted because they do not provide a *range* of best-effort services. At least, a low delay protocol switches to the default class when loss rate is too high. Thus, adaptive protocols with specific requirements in delay and loss rate cannot easily be implemented on top of them. Their scheduling mechanism is a function of the TCP end-to-end performance in order to provide a default class for TCP flows which are expected to obtain the same level of performance they would have obtained on a flat best effort network. The EDS service differentiation cannot provide such a class because this would not match the third criteria of its design philosophy. It cannot takes into account the end-to-end performance of protocols because its QoS processing is limited to the router.

From the proportional differentiation point of view. Schedulers presented in [3] met a large success because of there accuracy. Previous work [12] proposed a delay differentiation but in a less accurate manner. Delay differentiation has been widely studied after [3], particularly to improve its accuracy in periods of low load [11]. It is possible we move the delay differentiation implementation from the one we have today to one that has been shows to be more accurate.

6 Conclusion

We have presented the EDS best-effort service differentiation system. The EDS system has been designed by mapping the IP design philosophy to service differentiation. From the network viewpoint, it results that EDS does not affect the robustness, ease of use and deployment of IP. It can inherently be easily and incrementally deployed. We do not expect the system to replace differentiated services systems which provide stronger guarantees through provisioning and control.

From a more practical point of view, since we implemented proportional differentiation schedulers in Linux [6], we are also planning to implement the RED based implementation next.

Stronger guarantees on the plain best-effort IP are implemented in end-to-end protocols. For example, TCP guarantees reliability. These protocols use the building block provided by IP to provide a specific service matching the need of specific applications. Since we have defined a best effort service differentiation system from the same design rules, the natural way to implement stronger QoS guarantees has to be done in the protocol layer. This is the subject of parallel work.

References

[1] Steven Blake, David Black, Mark Carlson, Elwyn Davies, Zheng Wang, and Walter Weiss. An architecture for differentiated services. Internet Request For Comments RFC 2475, Internet Engineering Task Force, December 1998.

[2] David D. Clark. The design philosophy of the DARPA internet protocols. In *Proceedings of Special Interest Group on data Communication (*SIGCOMM*)*, number 4 in Computer Communication Review, pages 106–114, Stanford, CA USA, August 1988. ACM, ACM Press.

[3] Constantinos Dovrolis and Parameswaran Ramanathan. A case for relative differentiated services and the proportional differentiation model. IEEE *Network*, 13(5):26–34, September 1999.

[4] Victor Firoiu and Xiaohui Zhang. Best effort differentiated services: Trade-off service differentiation for elastic applications. In *Proceedings of International Conference on Telecommunications (*ICT*)*, Bucharest, Romania, June 2001. IEEE.

[5] Sally Floyd and Van Jacobson. Random early detection gateways for congestion avoidance. IEEE/ACM *Transactions on Networking*, 1(4):397–413, August 1993.

[6] Benjamin Gaidioz, Mathieu Goutelle, and Pascale Primet. Implementation of IP proportional differentiation with waiting-time priority and proportional loss rate dropper in linux. Technical Report RR-4511, INRIA, August 2002.

[7] Benjamin Gaidioz and Pascale Primet. The equivalent differentiated services. Technical Report RR-4387, INRIA, February 2002.

[8] Benjamin Gaidioz and Pascale Primet. EDS: A new scalable service differentiation architecture for internet. In *Proceedings of International Symposium on Computer Communication (*ISCC*)*, pages 777–782, Taormina, Italy, July 2002. IEEE.

[9] Paul Hurley and Jean-Yves Le Boudec. A proposal for an asymmetric best-effort service. In *Proceedings of International Workshop on Quality of Service (IWQoS)*, pages 132–134, London, England, June 1999. IEEE/IFIP.

[10] Paul Hurley, Jean-Yves Le Boudec, Patrick Thiran, and Mourad Kara. ABE: Providing a low-delay service within best effort. IEEE *Network*, 15(5), May 2001.

[11] Jung-Shian Li and Hsing-Chien Lai. Providing proportional differentiated services using PLQ. In *Proceedings of Globecom*, San Antonio, Texas, USA, November 2001. IEEE.

[12] Y. Moret and Serge Fdida. A proportional queue control mechanism to provide differentiated services. In *Proceedings of International Symposium on Computer and Information Sciences (ISCIS)*, Belek, Turkey, October 1998.

Unité de recherche INRIA Rhône-Alpes
655, avenue de l'Europe - 38330 Montbonnot-St-Martin (France)

Unité de recherche INRIA Futurs : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)
Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399