SIMA – Simple Integrated Media Access applied to Internet charging

Carl Eklund
Nokia Research Center
Communication Systems Laboratory
P.O. Box 422
FIN-00045 Nokia Group, Finland
Email: carl.eklund@nokia.com

Abstract

Simple Integrated Media Access is a network service compatible with the proposed Differentiated Services 'Dynamic RT/NRT Per-Hop-Behavior(PHB) Group'. Traffic is divided into two categories: a real-time(RT) class and a non-real-time(NRT) class. Within each class packets are assigned one of six drop precedence levels. The traffic control is designed to provide shorter transmission delays for RT traffic while trying to keep the drop probability for NRT packets low. Link bandwidth is shared among the classes based on a single parameter, the Nominal Bit Rate (NBR). The information about the class and drop precedence is carried in the DS field of the IP header. Decisions on dropping packets are taken on a per packet basis. SIMA network nodes keep only local state and there is no signaling between nodes. It is easy to implement both in the core network and on the edges. SIMA provides a understandable way for the user to relate the quality of service to the money spent by relating the charge only to the NBR. This papers describes SIMA and how it can be used in Internet charging.

1 Introduction

Differentiated services(DS) have recently been proposed for providing Quality of Service. The Integrated Services approach has been found to scale badly and be overly complex. Simple Integrated Media Access (SIMA) is a traffic control scheme originally proposed for traffic conditioning in packet networks, especially for ATM. The development has, however, since then been directed towards the Internet and recently a proposal has been submitted to the IETF to standardize a PHB group tailored for SIMA.

2 SIMA

2.1 SIMA service model

Three main features of the SIMA service are visible to the end-users. These are the Nominal Bit Rate (NBR), the selection between time critical and non-time critical service and quality expectations.
The fundamental principle in a SIMA network is that the customer is willing to buy for a certain amount of money, a connection with a certain Nominal Bit Rate (NBR). The customer is entitled to send any amount of data into the network but the quality of the connection depends on the actual bit rate to NBR ratio, the smaller the ratio the better the quality.
Important for the SIMA concept is also the ability for the user to determine whether a flow is real-time or non-real-time. In practice, this decision can be automatically made at the application level. The network offers for real time connections as short a delay and as small a delay variation as possible by using small dedicted buffers in the network nodes.
The expectations on the QoS of the network connection is mainly determined by the dollars spent by the customer and the impression the customer has of the QoS of an equally priced competing service. The actual quality in a SIMA network depend on two issues: the NBR to actual bit ratio and the total load in the network. This of course means that it is impossible for the customer to know the exact QoS the connection will receive and rapid traffic variations may bring about unexpected changes of QoS in existing connections. Therefore the SIMA network must be implemented in a way that users can rely on the fairness of the service. Of course the provider must dimension the capacity of the network appropriately if they want to retain their customers.
The fairness of SIMA based on the fact that all (aggregated) flows with the same actual bit rate NBR ratio perceives a similar QoS, i.e. a home user with a 10kbit/s NBR receives the same QoS as a large company with a 10 Mbit/s NBR provided that both are

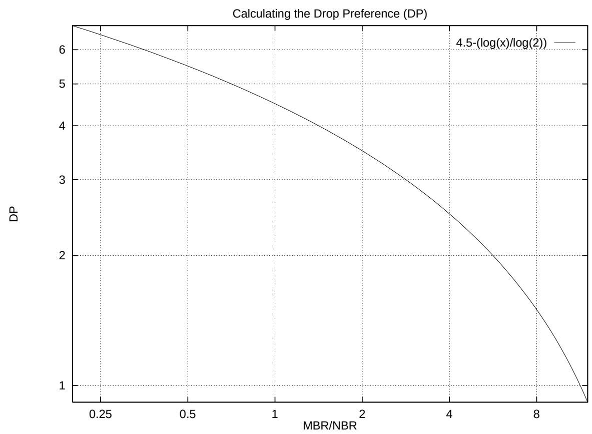

Figure 1: DP as a function of the MBR/NBR ratio

transmitting at their respective NBRs[1, 2].

2.2 SIMA implementation

There are two kinds of building blocks in a SIMA network: access nodes and core network nodes. The access node measures the traffic per aggregated flow basis and sets the DS code points. All traffic must enter a SIMA network through an access node.
The access node works as follows: Assume an aggregated flow,i of IP packets. A nominal bit rate, $NBR(i)$, is associated to the flow and no assumptions are made about the traffic process. At the user/network interface the momentary bit rate is measured by a device. An exponential average or a token bucket can be used for this purpose. The measured rate is denoted $MBR(i,j)$ at the arrival of the j:th packet. A drop preference $DP(i,j)$ is given to the packet based on the formulas:

$$x = 4.5 - \ln\left(\frac{MBR(i,j)}{NBR(i)}\right)/\ln 2$$

$$DP(i,j) = \begin{cases} 6 & \text{if } x \geq 6 \\ \text{Int}(x) & \text{if } 0 < x < 6 \\ 0 & \text{if } x \leq 0 \end{cases}$$

where Int(x)is the integer part of x [1].
The drop precedence as a function of the MBR/NBR ratio is shown in figure 1. From the figure it can be seen that a for a amount of traffic equal to the NBR the DP will be 4 while at only 17% of NBR all packets will be of highest priority. The access node maps the RT/NRT information and the DP to the DS field bit [3] in the IP headers. The current proposal for the mapping can be found in [4]. The packets are then passed to a router implementing the SIMA core router functionalities.
In a SIMA network core network functionalities shown in figure 2 have to be implemented in every output buffer of each router interface. The architecture of a SIMA network is shown in figure 3 Before a packet is allowed into either output queue its DP is compared against the lowest accepted drop precedence value, DP_a, and only packets with a value geater or equal to that are allowed into the queues. The lowest accepted DP can be calculated as a function of the queue occupancies. One choice could be:

$$DP_a = \text{Int}\left(7 \times \max\left(M_{RT}, M_{NRT}\right)\right)$$

where $M_{RT,NRT}$ are the normalized occupancies of the queues respectively, but more refined ways are not precluded.

[1]The terminology of used in this paper and the formulas are those of the original SIMA authors. A drop precedence of zero means lowest priority traffic. Keeping in mind the meaning of the word precedence (in some papers they use preference), their definitions defy all conventional logic and cause serious confusion

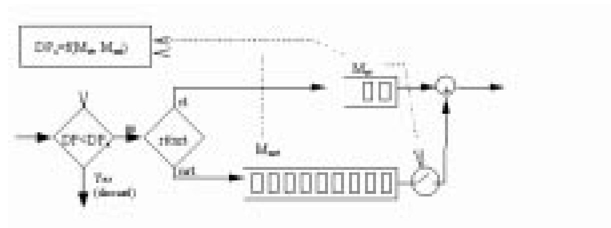

Figure 2: The SIMA core functionality([2])

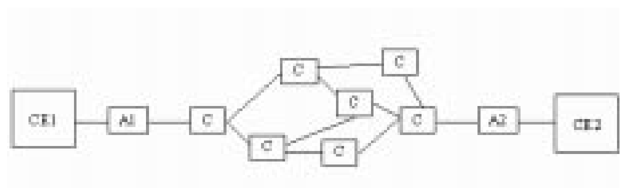

Figure 3: A SIMA network

3 SIMA and DS

SIMA fits well into the DS philosophy[5] . A DS PHB-Group called Dynamic RT/NRT PHB Group has been proposed to the IETF. The Dynamic RT/NRT PHB group is tailored for SIMA[4]. The proposal is to have two classes of traffic, real-time and non-real-time. Within each class there are at least four different levels of drop precedence. The delay suffered by the RT class must be lower than the NRT class. DS code points within the standards track code point space have also been proposed.

4 Internet charging

A good Internet charging scheme will at least have the following properties:

- simple implementation
- no complicated tariff tables
- suitability for traffic with different requirements
- suitability for connection oriented and connectionless applications
- possibility for flat-rate as well as time-dependent charging
- independence of underlying network technology
- difference between busy and idle hours are taken into account
- it is incentive
- it is equitable

Also it is often perceived that Internet rates should be of the form

$$X = F(p_1, p_2, \ldots, p_n)$$

where X is the amount paid for the service and the $p_1, p_2, \ldots, p_n$ are the essential technical parameters of the connection. This, however, usually leads to a number of problems. The tariff tables tend to be complicated, the traffic is unpredictable and it rules out flat rate charging. Setting the parameters is also a daunting task.

An ISP might want to provide different service classes each associated with a monthly fee. The requirements for these classes are:

- inside a class every customer should be served equally
- a higher class must provide significantly better service than any lower class
- the charges of the different classes must reflect the difference in service quality.

The requirement of equality can be understood to mean that all customers of a certain class can obtain the same amount of bandwidth on a given link. This amount of course varies depending on the momentary load situation and from link to link.

The assumption concerning the fairness can be interpreted in a way that on a certain link bit rates attainable for customers belonging to service classes i and j, W_i and W_j, respectively are proportional to the charges of the services X_i and X_j i.e.

$$\frac{W_i}{W_j} = \frac{X_i}{X_j}.$$

In addition any charging scheme needs an adequate traffic control mechanism to implement it. A mapping between the charge and the traffic control parameters is also required.

4.1 Integrating traffic control and charging in SIMA

The NBR concept is fundamental in SIMA. It also provides a convenient parameter for the charging. A customer subscribes to a certain NBR which is the same NBR parameter used in the ingress node of the DS network with with a DRT/NRT PHB. This of course raises the question of how the customer can be sure that (s)he actually gets the data rate (s)he paid for. In adherence to the DS philosophy no quantitative guarantees are given, only the assurance that the capacity is distributed according to the NBRs.

The competition in the ISP business is however fierce and the client will quickly change service provider [2] if the service is perceived inadequate. This means that there is an incentive for the ISP to invest in a sufficient infrastructure. The total amount of subscribed NBR serves as a more accurate indicator of the capacity needed than the mere number of subscribers.

The charging can be flat rate, time dependent or a combination of the both. A potential scheme could be that the customer buys a basic NBR for a flat rate and within certain limits can buy additional NBR during the session. This could either be achieved via RSVP signaling (only a minimal implementation is needed) and the access node or through WWW server of the ISP.

One feature in the SIMA charging scheme is that it automatically takes into account the difference between busy and idle hours. With most data applications the busy hour means a smaller portion of bandwidth is available per user. As bandwidth is divided between users according to their NBRs, subscription of a higher NBR is required to get the same level of service as during idle hours. This of course means increased revenues for the ISP. Again it has to be pointed out that this discussion assumes that the network is reasonably planned.

The charging and control system in SIMA is such that it is advantageous for the user to select real-time service only for delay sensitive applications and non-real-time services for other applications. Selection of a real-time service does not disturb the other applications.

Another incentive property is that it is useful for the customer to adapt the bit rate of a real-time application during congestion. If the user persists on a high bit rate during these times he either has to buy more bandwidth or suffer severely degraded service.[6] Whether this is a positive feature can of course be discussed.

It should also be noted that no changes in the cus-

[2]This will be even more true in the future when autoconfiguration tools will make changing providers a point and click operation

tomer hosts are needed if the access node of the ISP performs the packet classification.

5 Conclusions

SIMA is a traffic conditioning scheme compatible with DS. SIMA provides an easy way to relate the QoS with the billed amount, both for the ISP and the customer. SIMA billing exthibits the properties that are desirable in a good billing scheme. It can in addition provide savings in equipment cost for the ISP as connections at differrent speeds can be supported with common equipment.

References

[1] K. Kilkki. Simple integrated media access - a novel service concept for the internet. In *ACTS ATM Traffic Symposium '97*, 1997.

[2] K. Kilkki and J. Ruutu. Simple integrated media access (SIMA) with TCP. In *4th INFORMS Telecommuncations Conference*, 1998.

[3] K. Nichols et al. Definition of the Differentiated Services field DS field in the IPv4 and IPv6 headers. RFC 2474, IETF, December 1998.

[4] M. Loukola, J. Ruutu, and K. Kilkki. Dynamic RT/NRT PHB group. Intenet draft, IETF, November 1998.

[5] Y. Bernet et al. A framework for Differentiated services. Internet draft, IETF, February 1999.

[6] K. Kilkki, M. Kalervo, and J. Ruutu. Internet charging reconsidered. In *Networld+Interop 98*, 1998.