Paris Metro Pricing

Henri Koskinen
Networking laboratory, HUT
henri.koskinen@hut.fi

Abstract

The need for differentiated levels of Quality of Service in the Internet has motivated the development of several QoS mechanisms. This paper describes Paris Metro Pricing (PMP), the simplest differentiated services solution, proposed by Odlyzko [1]. In PMP, the service differentiation would result implicitly from dividing the network into several logically separate best-effort channels with different usage prices. Channels with higher prices would attract less traffic, and hence provide better service.

In addition to giving an overview of the PMP proposal itself, this paper addresses its feasibility in real markets with reference to some game theoretical studies.

1 Introduction

It has been widely recognized that the future Internet must be able to provide differentiated services to cater to the needs of various applications with different requirements. As a result, several QoS mechanisms have been developed. Since the task of any such mechanism is to allocate limited network resources, it is safe to say that some kind of usage sensitive pricing scheme has to be involved. The purpose of this is to prevent the so-called "tragedy of the commons" where all traffic is sent with the highest priority.

Many of the proposed QoS mechanisms involve complicated and costly implementations. Paris Metro Pricing maximizes simplicity by integrating pricing with traffic management. Service differentiation is achieved essentially as a side effect of price differentiation. PMP fits consumer preferences for simplicity, both in terms of pricing and user interaction with the network. This is achieved at the expense of some of the utilization efficiency of the network.

The PMP concept is presented in detail in section 2. Section 3 discusses some design and implementation issues that have to be addressed when adopting PMP. The proposal is viewed from a business perspective in section 4. The results of two game theoretical studies are presented in the latter part of the section. Finally, section 5 sums up some conclusions.

2 Paris Metro Pricing

2.1 Background

The inspiration for the PMP proposal was – as the name implies – the Paris Metro system. Until some 20 years ago, users were offered the choice of travelling in either first or second class cars. The cars were completely identical in quality and number of seats. The only difference was that tickets to first class cars cost twice as much as to the second class cars. The first class cars thus attracted fewer people and were, on average, less crowded. Those users with a strong preference for, e.g., obtaining a seat and not being disturbed by noisy teenagers were willing to pay the higher price. The system was self-regulating, in that whenever the first class cars became too popular and nearly as crowded as the second class cars, fewer people considered them worth the extra cost, and the differential in the quality of service was maintained. This same system was in fact used on the Paris rail suburban lines until September 1, 1999.

2.2 PMP proposal

The PMP proposal is simply to apply the Paris Metro scheme in the Internet. The idea is to partition a network into several logical channels, each charging a different usage-based price, and treat packets on every channel on a best-effort basis just as in the Internet today. Users would choose the channel to use for their packets and pay accordingly. The channels with higher prices could be expected to be less congested and therefore provide better service. As in the Paris Metro, there would thus be no guaranteed quality of service, but the expected level of service on the channels would be different.

PMP inverts the traditional order followed in network design. The conventional approach is to initially try to determine the QoS requirements of various applications and then design the network so that these requirements can be met. Finally, prices are set for the different QoS levels. In PMP, pricing is the main method of traffic management that ultimately determines, through affecting user demand, the resulting QoS.

3 Technical issues

As price is the key control parameter in PMP, it is involved in the discussion in this section, even though it is not traditionally regarded as a technical issue.

3.1 Design - setting the parameters

What makes PMP so simple compared to other QoS schemes is that the main idea is merely to divide the present Internet into several smaller copies of the original, each operating exactly as before in parallel. There is however a price to pay: subdividing a network into several pieces loses some of the advantages of statistical multiplexing that large networks offer. Bearing this in mind, the number of channels in PMP should be small. Having few channels is also preferable because the simplicity appeals to consumers. Furthermore, it has been observed in other contexts that most of the economic gains from service differentiation can be achieved with just a few classes. In [1], the author suggests the number of channels should be three or four.

In fact, PMP does not necessarily have to be based on the channel division. One alternative is to use packet priorities. The advantage of this approach is that the full gain from aggregating all traffic on a single channel would be maintained. However, this would require careful consideration for the fair distribution of service among the different prioritites.

One crucial problem is how to allocate the capacities to and set prices for the channels. Near-optimal solutions could probably be achieved iteratively in a piloting phase by utilizing feedback from a test population. At any rate, these parameters should stay constant for extended periods to allow usage patterns to stabilize. This would promote predictability in both price and service quality, a key factor in terms of consumer satisfaction. It would however make sense to have different price and capacity assignments for different times of day, to allow for better network utilization. How frequently to vary these assignments is yet another question.

According to experience from the Paris Metro, prices should jump by a substantial factor, i.e. around two, from one channel to the next, in order to achieve sufficient differentiation in the service quality. This would also alleviate potential network instability resulting from degradation on one channel propagating to other channels as users would shift their channel choices according to the perceived service quality (see Figure 1). A sufficient price difference between successive channels would raise the barrier to such shifting.

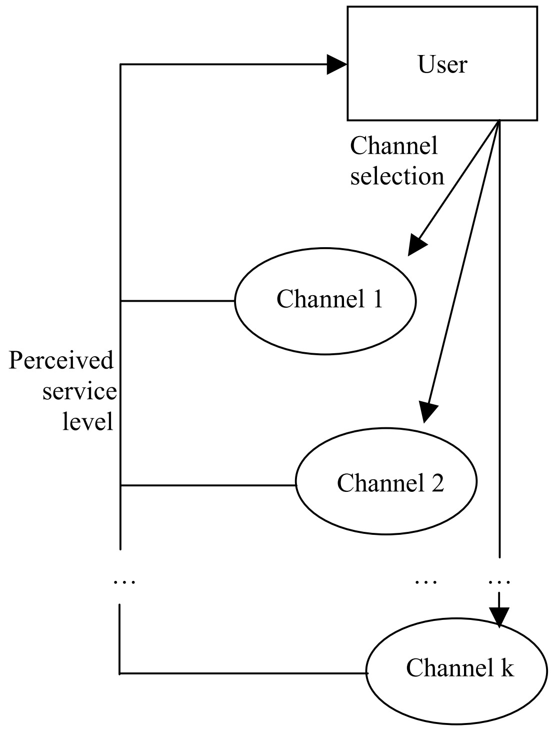

Figure 1: The iterative channel selection process in PMP

3.2 Implementation

The major change required in a network by PMP is the charging infrastructure. In order to implement usage-based pricing, there would have to be some entity keeping track of traffic for each user. Most of this functionality could be located at the edge of the network, although accounting between different ISPs would have to be performed at inter-ISP gateways.

Apart from the charging mechanism, only minor changes to the network infrastructure would be required. Inside the network, router software would have to be modified to maintain logically separate queues or to give appropriate priorities or weights to packets from different channels. This would require nothing new in light of the readily existing mechanisms developed for DiffServ.

QoS-sensitive applications would have to be rewritten to enable the users to select the channel from within each application. No new protocols would be required since packets in the current IP version 4 already have three unused priority bits, which is more than enough to indicate the desired channel out of the projected four.

In the transition phase, before the appearance of rewritten applications, some kind of “wrapper” software would be needed to set the priority bits of all packets to the desired level. Also, the price on the lowest channel would be zero to make it look like the current Internet and handle traffic from networks not implementing PMP, to smoothen the transition. (In fact, having zero price on the lowest channel could be feasible in the long run as well.)

There are some open questions that need to be addressed regarding PMP and networks of different service providers. How does a network that implements PMP interoperate with one that does not, or with one with different capacity and price assignments? How should revenues be divided among different providers?

4 Business issues

4.1 Consumer response

4.1.1 Best-effort service

One major question is whether consumers would find the lack of guaranteed quality of service acceptable, to the extent that they would still be willing to pay for that service. It is not difficult to find examples in other contexts where purchases are made based on expected, not guaranteed, quality of service. There is a lot of uncertainty involved when bying a book, a record, or when going to see a movie, a concert, or a sports event. Moreover, consumers are willing to accept occasional large deviations from the expected quality of service. An airplane passenger in first class may have an uncomfortable trip, if there is a crying child in the seat behind, while an economy class passenger may have a whole row of seats to himself and have a much more pleasant flight than those in first class. On average, however, a first class passenger does enjoy superior service, and that is enough to maintain a huge price differential. Figure 2 illustrates the variations in quality that can take place in different service classes.

It seems therefore that consumers could accept the lack of guaranteed QoS in PMP, provided that the average quality of different channels is predictable enough. However, the unpredictable nature of data traffic implies that sporadic congestion is bound to occur on all channels. All that can be expected is that the higher-priced channels would experience service degradation less frequently. Sufficiently infrequent congestion would probably be deemed acceptable.

Moreover, guaranteed QoS can be regarded as an unrealistic illusion. Apart from very simple networks, the only absolute guarantees that can be made are for constant bandwidth. On the other hand, the utilization efficiency of capacity in data networks is based on statistical multiplexing of sources with varying and unpredictable data rates. Guaranteeing all bandwidth requirements and utilizing network resources efficiently are two contradicting objectives. Furthermore, Ethernet-like networks are still going to have a strong foothold in the overall infrastructure for an as yet undefined period, and they are hard-wired to operate on a best-effort basis. The justification for PMP is that, after all, the Internet does work, and when lightly loaded, even real-time applications can be run. This has been demonstrated on many closed networks. This suggests that PMP, as a best-effort system with several channels of different congestion levels, might satisfy most needs.

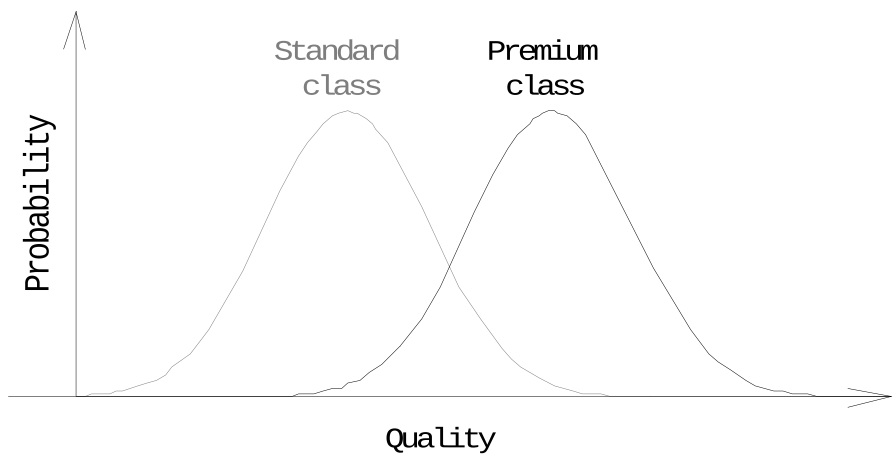

Figure 2: A schematic representation of varying service quality in different classes

4.1.2 Usage-based pricing

As stated before, providing differentiated services requires usage-sensitive pricing. This is in contrast with consumers' strong preference for flat-rate pricing. It has been observed that consumers are actually willing to pay more for a flat-rate plan than they would end up paying under a usage-based pricing scheme.

It seems that consumers might accept usage-based pricing if the benefits are made clear. Since the lowest-priced channel in PMP could – at least initially – be free of charge and would thus behave just like today's Internet, customers would themselves have the freedom of choice in experimenting the use of the priced channels.

However, the best way to accommodate the preference for flat-rate pricing would probably be a compromise, namely, block pricing. Blocks of transmission capacity could be sold in advance, giving the user the right to send or receive a certain amount of data over a certain period of time over the lowest-priced channel, or a bit smaller amount on a higher channel, etc. Predefined blocks could be issued and charged on a regular, e.g. monthly, basis according to an agreed contract. Any usage exceeding the block would then be charged separately. Block pricing has worked well in long-distance telephony in the United States, with consumers typically paying for more service than they actually use.

4.1.3 Fairness

Consumer behavior is in large part driven by the notion of fairness. In the design of PMP, assigning fixed capacity to different channels is likely to appeal to consumers more than using for example priorities. It avoids the appearance of the widely disliked auction mechanism where users willing to pay higher prices would get all the bandwidth. It also shifts the responsibility for congestion to other users and not on the network provider, which is deemed more acceptable.

4.2 Feasibility of PMP

The PMP proposal has received a lot of attention in the networking research community and inspired further studies on the matter. In [2], Gibbens, Mason and Steinberg assess the viability of service class differentiation proposals in general and PMP in particular, by analyzing competition between two ISPs, either or both of which may choose to offer multiple service classes. The authors consider three cases: the social optimum, where total benefit from network usage is maximized, monopoly, where a single network maximizes profit, and a duopoly, where two networks compete to maximize their individual profits. In the game theoretical model developed, both a social planner, interested in maximizing the welfare of users, and a profit-maximizing monopolist will wish to use several service classes, but duopolists will not. The reason to the latter is that when networks offer multiple service classes in order to increase user benefits and hence charge higher prices, they effectively increase the number of points in the service quality range at which they compete. In the model, the costs of increased competition as more subnetworks are introduced always outweigh the benefits from greater segmentation of the market, and consequently, in any equilibrium competitive outcome, both ISP's offer a single service class. This suggests that PMP may not be viable under competition.

In [3], Cao and Shen arrive at some contradicting results. They adopt a leader-follower game framework: an ISP sets prices as the leader and the consumers respond with demands. The objective of the ISP is to maximize profits by optimizing the prices to induce desirable demands from the consumers. The setting therefore corresponds to the monopoly case considered in [2]. To study PMP, the authors assume the ISP divides its capacity into two classes. The solution of a numerical example indicates that the leader-follower game leads to an optimal solution with the same price for both classes in PMP, i.e. the ISP is better off by not differentiating the service classes. This is inconsistent with the conclusion arrived at in [2]. Apparently, the general result seems to depend on the chosen model, which makes one question the viability of game theoretical studies of PMP. Indeed, in the concluding sections of both [2] and [3], the authors point out that allowing the free entrance of competitors into the market may produce different equilibrium results.

5 Conclusion

This paper presented Paris Metro Pricing, the simplest differentiated services solution. For comparison, DiffServ, the most popular of the QoS mechanisms developed, would imply a lot more complex network control, as it does not by itself say anything about assignment of priorities and pricing. It only specifies how the network should deal with packets with different priorities.

Complexity is the main problem with most QoS schemes, given the fact that networking is already seen to be too complicated. In fact, it has been argued that instead of working on complicated QoS schemes, all resources should be put into improving capacity (the so-called "fat dumb pipe" model). Recent studies have shown that most of the Internet is very lightly utilized, and that the main demand is for low delay, not high throughput. It appears that in the backbones of the Internet, providing uniformly high quality of service might be not only feasible but also optimal, given the implementation costs that most QoS measures, even PMP, would cause.

On the edges of the Internet, on the other hand, constraining traffic with pricing will probably be needed. With wireless access technologies, for example, the resources are likely to remain so scarce that not all demand can otherwise be handled.

References

[1] Odlyzko, A.: Paris metro pricing for the internet, in *Proceedings of the first ACM conference on Electronic commerce*, pp. 140-147, November 1999, ISBN 1-58113-176-3

[2] Gibbens, R.; Mason, R.; and Steinberg, R.: Internet service classes under competition, *IEEE Journal on Selected Areas in Communicarions*, vol. 18, no. 12, pp. 2490 –2498, December 2000, ISSN 0733-8716

[3] Cao, X. R. and Shen, H. X.: Internet pricing: comparison and examples, in *Proceedings of the 39th IEEE Conference on Decision and Control*, vol. 3, pp. 2284-2289, December 2000