Paying for High Speed Network Services

WB Smuts
Dept of Computer Science and Information Systems
University of South Africa, PO Box 392, Pretoria
E-mail: smutswb@alpha.unisa.ac.za
Facsimile: 012 429-3434

Abstract: The real-time services offered by modern high speed networks are complex enough to prompt a new look at how these services should be paid for. In this article, it is motivated that paying for the use of guaranteed services can ensure fairness in the availability of limited data communication resources.

The limitations of conventional accounting frameworks for communication services are discussed. A new framework for accounting in high speed networks is proposed and the advantages thereof pointed out. A quotation and the payment are included as part of the basic protocol to establish a channel. This saves the cost of recording usage information for every user, sending out invoices and collecting payments. It furthermore makes it possible for users to use and pay for services provided by distant service providers with whom they do not have an account. It is proposed that users pay a flat rate connection fee for unguaranteed services, but pay more if the services are guaranteed.

Finally the importance of a service reservation protocol, and of the guarantees offered by these protocols are pointed out.

Keywords: Computer Networks, Invoicing, Billing, Digital Cash.

Computing Review Category: C.2.0, C.2.2, C.2.3.

1 Introduction

Paying for the use of a computer communication network may be a controversial idea. Some researchers attribute INTERNET's success to the fact that it has been virtually free of charge for most of its users.

Although this investigation is not focused on the INTERNET, it is still important to take note of the recent awareness of the cost of providing INTERNET services. In an article on charging for INTERNET services [Dav94], Dai Davies[1] mentions that:

> "Freeness leads to a lack of respect for facilities and discourages investment. The problem that needs to be addressed is a creative approach to pricing."

High speed networks (Gigabit+) are just emerging from the laboratories. Real-time applications such as video conferencing are a reality. According to studies by Parris, these high speed networks offer qualities of service (QOS) which are complex enough to warrant a new look at pricing in integrated networks [PF92].

[1] Dai Davies is joint manager of DANTE, the company set up by the European National Research Network to provide international services.

Pricing is the process by which the service provider determines the price that the user must pay. It usually follows a process of *costing* by which the cost of providing the service is determined. Pricing and its effect on network utilization have been studied quite extensively [CPF92, Zha91].

Although pricing and costing may be important, these are a financial issue best left to the economists. In this investigation, *costing* and *pricing* have not been studied directly. The study concentrated more on the technical *mechanisms* (protocols, etc.) needed to implement a network accounting system. For the purposes of this article, accounting is defined to include the processes of *recording usage information*, *sending out invoices* and *collecting payments*. In particular, the issues involved in implementing an accounting system as part of the *basic* communication protocol were studied.

Instead of relying on the big service providers to decide how the users will pay for services (which favours the formation of monopolies), one of the underlying objectives of this study was to investigate a payment system which will make it possible for smaller service providers to become part of a larger networking service — and which will give the users the freedom to choose which provider they want to use.

The cost of accounting can be very high and can form a substantial part of the total cost of providing a networking service. Because the user carries this cost, he benefits directly if it can be kept to a minimum. The cost of accounting must be balanced by the benefits it provides for both the service providers and the users.

In this article, we motivate why paying for guaranteed communication services is essential when the resources are limited. We continue to discuss the conventional framework used for accounting in communication networks, after which a new framework is proposed. Finally a few concluding remarks are made.

2 Motivation

A free data communication service does not exist. The infrastructure cannot be put into place and maintained without money. Somewhere someone pays. The question is not whether the services should be paid for or not, but *who* should pay for it. More specifically: Should the end-user be involved in footing the bill or should he be buffered from the fact that his actions on the network cost money?

One of the conclusions from our investigation is that the effect of a communication network accounting system goes beyond the collection of money to pay for installation and maintenance.

A basic assumption which underlies our research is that communication resources are limited and in many cases even scarce. If communication resources were completely unlimited, there would be little motivation for doing this research.

Three objectives for invoicing the users have been identified:

- **Earn revenue.**

  If earning revenue were the only objective, a simple flat-rate per network connection (regardless of the amount of traffic generated) can be used. This results in a very simple and cheap invoicing system.

  The potential uplifting effect of a completely free data communication service may even call for a system in which tax money is used to provide a *'free-for-all'* system to the users.

  One of the limitations of paying for access only, is that it becomes difficult, if not impossible, for the user to use and directly pay for services on a remote network. The user must rely on the local service provider to choose which remote service provider to use, and to negotiate contracts with those service providers. Whereas most users would probably prefer not to be bothered by these choices, we do believe that many would like to have a choice if different prices and qualities of service form part of the deal.

- **Equalize the demand.**

  Studies by Parris [PF92] have shown that a differentiated pricing system which takes the amount of traffic offered to the network, as well as the time of the day into account, can spread the load offered to the network more evenly and reduce peak loads. It should be clear that spreading the load is only feasible if there exist times when the network is not fully utilized.

- **Create fairness in availability.**

  If communication resources were completely unlimited, these services would have been always available to everyone. However, if resources (link bandwidth, node buffer space, etc.) are limited, either performance, or availability decreases when more users make use of the network.

  Unguaranteed services such as those offered by TCP/IP, are shared by all users on an equal basis. If the load increases, the performance decreases for all users. The effect of this may range from mild frustration, to a complete collapse in the case of real-time applications (eg. video conferences).

  Performance guarantees (provided by either deterministic or statistical reservation of resources) can ensure that real-time applications operate satisfactorily in the face of an increase in demand, but this limits the availability of services to the rest of the users.

  If resources are limited, free to the users, but can be reserved, users will have no incentive to let go of any channels they have managed to reserve for themselves. In fact, there will be a strong incentive to never let go of any reserved resources. The *'free-for-all'* system will still be free – but not for all. The lucky first-comers will be the only ones served by such a system.

  There are more than one way to force users to let go of reserved services which they may not obtain easily again. A time-out on connections, for example, may free the resources for other users after a reasonable period of time has elapsed since establishment of the connection. The problem is how to determine what a 'reasonable' time-out period is. If the time-out is too short, it may cut off a 'live' connection. If the time-out is too long, resources are wasted.

  A time-out which starts only after the last packet (or cell) has been sent, merely shifts the problem to determining how long a user can reasonably be expected to remain 'silent' on an open connection, before the connection is cut off. Again, if the time-out is too short, it may cut off a 'live' connection and if the time-out is too long, resources are wasted. This solution will furthermore prompt users to periodically send an *'I'm-alive'* packet to keep an idle connection open for future use.

  We believe that services can be made available on a fair basis with an invoicing system which differentiates at least in terms of the type of service, the amount of data sent, and the duration of guaranteed connections. The purpose of the invoicing system is not only to generate revenue[2], but to create fairness in the availability of the resources. It creates an incentive to terminate connections as soon as they are not needed, and it gives users a choice to spend their money (or any suitable currency such as 'network tokens') on the type and quality of services which are important to them.

  When resources are limited, a communication network without an option to reserve guaranteed services and without the infra-structure to invoice users for these guaranteed services, leaves the users at the mercy of their fellow network users. In such a network, the performance is determined by what other users are doing at that moment. Even if a user is willing to part with some of his money to obtain a better service, he has no other choice, but to try again late at night when the demand decreases.

[2] If earning revenue were the only objective, a flat rate per connection provides the cheapest solution.

The objective of this study is not to limit the freedom of people using a data communication network by asking them to pay for every byte they send, but to add the freedom to choose a higher quality service if it is important enough. A flat rate per connection for all unguaranteed services with a differentiated invoicing system for guaranteed services is proposed.

Although the main objective of the proposed invoicing system is to create fairness in the availability of resources, the other two objectives (earning revenue, and equalizing demand) must be served equally well.

2.1 Conventional Accounting Framework

Conventional accounting frameworks for communication services are based on the concept of a *user account*. A user must register (by filling in a form and sometimes paying a deposit) before he can gain access to the network. Depending on the pricing policy used by the service provider, the usage information for every user is recorded and stored. At regular intervals (usually monthly), an invoice is sent to the user. The user can then check the validity of the information on the invoice before paying the correct amount. If the user does not pay, the service provider can terminate the services until payment has been received.

The following disadvantages and shortcomings of the conventional accounting framework have been identified:

- The accounting system is in many cases very expensive. Structures must be put into place to record and store the usage information for every user. An invoice must be compiled and sent to the user, before payment can be collected. An expensive (eg. the postal system) parallel communication system is sometimes used to send out invoices and collect payments. This cost is directly passed on to the end-user.
- There is usually a long delay between the time at which the service is provided and the arrival of the invoice. If one of the objectives is to affect the usage patterns of the users in a way which will ensure that the demand is equalized, there should be a more direct link between using a specific service and paying for it. In order to adjust usage patterns in response to higher prices, a user needs to be able to distinguish which actions were expensive and which were cheap. Ideally, the user should know how much a service will cost even *before* it is used (generally known as a *quotation*).
- Most existing accounting systems do not stimulate competition among the service providers. Few applications are limited to the domain covered by one communication service provider. However, most existing accounting systems make it impossible for a user to *directly* use the services of a distant service provider. In some cases, the local service provider has contracts and agreements with the distant ones and the user is not even aware of the fact that his traffic is routed via 'foreign' nodes and links. The cost of using the distant service provider is reflected in the local invoice and payable directly to the local service provider. In these cases the user must simply accept the price negotiated between the service providers.

  The fact that the accounting systems are structured around the concept of an user account, with a long-term contract, makes it difficult (if not impossible) for users to use the services offered by service providers with whom they have no contract. At the same time, service providers cannot easily provide services directly to distant users, mainly because they have no direct way to collect payment.

3 Payment Methods

If the payment for communication services is to be included as part of the basic communication protocol, it has to be both fast and secure. The proposed framework for accounting in a data communication network is made possible by recent advances in techniques to implement the digital equivalent of cash. Modern cryptographic techniques make it possible to securely exchange payments in an open communication network. Bruce Schneier [Sch94] and David Chaum [Cha85] provide detailed discussions of the theory and implementation of digital cash payment systems.

When properly implemented, digital cash and its equivalents such as digital traveler's cheques provide several advantages for payment over networks. Paying by cash is one of the simplest ways to do business. Cash may be accepted on face value without taking the payer's credentials into account. Using digital cash to pay for communication services forms a logical extension of existing cash payment systems.[3]

4 A Framework for Accounting in a Data Communication Network.

In order to alleviate some of the limitations of conventional accounting systems, a new framework for accounting in data communication networks is defined. (It is important to bear in mind that it is assumed that the communication resources are limited in comparison to the demand for services.)

The following features describe the proposed accounting framework:

- In order to create fairness in the availability of resources, it is essential that users pay for the use of *guaranteed* services. *Unguaranteed* services may be paid for directly, or indirectly as part of a flat rate, or it may be free of charge.
- Services are paid for by the entity setting up the connection.
- The complete accounting process is implemented as part of the basic protocol used to establish and maintain a connection.
- A maximum-charge field is added to the QOS specification (also called a *flow-spec* in cell networks) which forms part of the header of the first packet (cell) used to establish the connection.
- The first acknowledgment packet from the destination contains the agreed-upon QOS parameter values, as well as a value for the total charge of the connection.
- In a similar way to which the delay specification is 'used up' by the nodes on the route, all nodes may add to the charge of the connection. The total charge for the connection is the sum of all the charges 'collected' on the route.
- Not every node needs to add to the charge. A big service provider may charge connections only at the nodes where its domain is entered. A single charge can be calculated depending on where in the domain the connection terminates, or where the connection exits the domain. How to calculate the charge is a pricing issue which is left for the service providers to decide for themselves.
- After receiving a quotation for the connection, the user can either reject the quotation, or accept it by including the payment in the next packet or cell.
- Digital cash (or any other suitable form of electronic tokens) is used for payment. These are cryptographically encoded to be payable only to the intended recipient.

[3] Digital cash, in more than one format, is already available to pay for merchandise on INTERNET. Payment schemes offered by the DigiCash Company, the European Community's ESPRIT Program's *Project CAFE*, CyberCash with their INTERNET Payment Scheme, the NetCash system and the First Digital Bank founded by David Chaum in the Netherlands, are examples of existing systems. (Information from WWW-buyinfo home page on INTERNET.)

- Expensive connections can be paid for incrementally, by sending the cash at regular intervals during the duration of the connection. In case of any network failures, disputes are limited to small amounts only.

The proposed framework has a number of advantages over existing approaches:

- Users may choose to pay for a higher quality of service when it is needed.
- Users can directly use the services of distant service providers. There is no need to open an account with those service providers.
- When source routing is used, users can select which service providers they want to use.
- In situations where privacy is important, users may remain anonymous, while still being able to guarantee their payment for the communication services.
- Service providers receive guaranteed payment for services rendered.
- Expensive logging of usage information and sending of invoices are not necessary.

5 Guarantees

Guarantees always play a role when purchasing products or services. The service provider must be able to guarantee certain properties of the service, while the customer must be able to guarantee payment. Few users will be willing to pay for unguaranteed services.

A number of different properties can be guaranteed. High speed applications typically need guarantees for three properties: [Par94]

- Worst-case loss rate,
- Worst-case bandwidth, and
- Worst-case delay.

Special mechanisms are needed to ensure that these properties are guaranteed. The guarantees may be either *deterministic*, or *statistical* in nature.

Deterministic guarantees promise that the values of the properties such as maximum delay, or minimum bandwidth will be within certain limits with a 100% probability. Resources such as link bandwidth and buffer space are reserved for every channel. One guaranteed channel is then not allowed to use the resources of another guaranteed channel, even if the second channel is not using it at that moment. Unused packets or cells in a guaranteed channel may only be used to forward unguaranteed traffic.

Statistical guarantees, on the other hand, only promise that the required parameter values will be within the limits with a probability of less than (but usually close to) 100%. Resources are not partitioned and reserved for specific channels, but remain in a pool. Total network utilization is higher than for deterministic guarantees, and thus the availability of services is increased.

A payment system cannot work without a well-designed service reservation system which provides service guarantees. The *Real-Time Protocol Suites* proposed by the Tenet Group at the International Computer Science Institute in Berkeley, California, forms an example of such a service reservation system which can provide service guarantees [MN94, GF94].

6 Conclusion

In this article, it is motivated that paying for the use of guaranteed services can ensure fairness in the availability of limited data communication services.

The limitations of conventional accounting frameworks for communication services were discussed. A new framework for accounting in high speed networks is proposed and the advantages thereof pointed out. The importance of a service reservation protocol, and the guarantees offered are also mentioned.

It must be pointed out that the proposed solution forms merely a *framework* within which a detailed accounting system can be designed.

References

[Cha85] David Chaum. Security without identification: Transaction systems to make big brother obselete. *Communications of the ACM.*, 28(10):1030–1044, October 1985.

[CPF92] Srinivasan Keshav Colin Parris and Domenico Ferrari. A framework for the study of pricing in integrated networks. Technical report, University of California, Berkely, Internet FTP: icsi-ftp.Berkeley.EDU, 1992.

[Dav94] Dai Davies. There's no such thing as a free internet. In *Proc. INET94/JENC5*. INTERNET, 1994.

[GF94] Amit Gupta and Dominico Ferrari. Resource partitioning for multi-party real-time communication. Technical Report TR-94-061, University of California at Berkeley and the International Computer Science Institute, Berkeley, California, 1994.

[MN94] Amit Gupta Wingwai Howe Mark Moran and Quyen Nguyen. Scalable resource reservation for multi-party real-time communication. Technical Report TR-94-050, University of California at Berkeley and the International Computer Science Institute, Berkeley, California, 1994.

[Par94] Craig Partridge. *Gigabit Networking*. Addison Wesley, 1994.

[PF92] Colin Parris and Domenico Ferrari. A resource based pricing policy for real-time channels in a packet-switching network. Technical report, University of California, Berkely, Internet FTP: icsi-ftp.Berkeley.EDU, 1992.

[Sch94] Bruce Schneier. *Applied Cryptography*. John Wiley and Sons, 1994.

[Zha91] R Cocchi D Estrin S Shenker L Zhang. A study of priority pricing in multiple service class networks. In *Proceedings of the ACM SIGCOM 1991 Conference.*, pages 123–130. ACM, 1991.