A framework for flow-based accounting on the Internet

Hans-Werner Braun
hwb@sdsc.edu
San Diego Supercomputer Center
San Diego, CA 92186-9784

Kimberly C. Claffy and George C. Polyzos
kc@cs.ucsd.edu, polyzos@cs.ucsd.edu
Computer Systems Laboratory
University of California, San Diego
La Jolla, CA 92093-0114

Abstract

We describe steps toward an accounting mechanism to attribute Internet resource consumption based on service quality. Our objective is not to describe a complete accounting and billing system. Rather we advocate taking advantage of existing Internet instrumentation to implement incremental improvements in the short to medium term. Experience from these improvements can enable us to make more educated progress towards a comprehensive Internet accounting system.

1 Introduction

Most architecture and instrumentation for accounting in the Internet reflects its historical status as bulk-funded good for the academic community, rather than the free market pay-per-service environment toward to which it is currently evolving. As a result the current Internet architecture is not conducive to pricing resource consumption among multiple entities and at multiple qualities of service. Quadratic traffic volume increases that still characterize many network clients reflect this current perception of the network as a research environment with usage-insensitive cost which are often transparent to the end-users.

In this paper we discuss two aspects of accounting and pricing in a "free market Internet", and propose how to incorporate these aspects into the current Internet infrastructure. We first discuss a macro-aspect of Internet accounting: decentralized attribution of network resource consumption. The second aspect we discuss is the implementation of multiple service qualities utilizing an existing mechanism in the IP architecture, specifically the IP *precedence* field. We judge these two components as essential to a realistic Internet accounting architecture that is implementable within short to medium time horizons.

In our description we separate the task of usage accounting from that of billing and concentrate on the former task: the attribution of resource usage to network clients. We view billing as a higher level application built on top of an accounting architecture, incorporating additional objectives and constraints. We do not contend that the scheme we propose will resolve, or even address, all the difficulties with Internet service billing. However we advocate taking advantage of existing instrumentation, albeit sparse, to make incremental steps toward more equitable attribution of Internet resource consumption and eventually cost recovery mechanisms. The proposal we outline here reflects this objective.

In Section 2 we present motivating factors for implementing reasonable Internet accounting as soon as possible. In Section 3 we discuss the the requirement for distributed accounting. Section 4 discusses the history of and motivation for offering multiple levels of service precedence on the Internet. Section 5 discusses how current Internet instrumentation can support these requirements, and present relevant statistics from an example data set. Section 7 summarizes our proposals and discusses directions for further research.

2 Motivation

Two ever-advancing trends motivate our objective of deploying more effective Internet accounting mechanisms in the short-term. First, as interconnectivity and resource sharing among network providers continues to increase [3], providers will need mechanisms to determine the extent and fairness of their traffic exchange. Both commercial and non-commercial, domestic and international, providers will track how they "use each other's bandwidth".

The second and perhaps more critical trend is the demand which newer real-time multimedia applications place on the Internet. The intense and continuous flows of these applications will increase the gap between high and low end Internet users. Low end users, who require only bursty, low-priority traffic and yet find themselves suddenly unable to get a packet through the network will

This research is supported by a grant from the National Science Foundation (NCR-9119473). To appear in Proceedings of SICON '93, ©IEEE Singapore Section.

find suboptimal a scheme where they pay the same flat fee for their slightly used connection as their multimedia neighbor pays for his continuously packed one. Instrumenting the network so that providers can base service costs on individual client needs will allow for a more equitable overall network service.

3 Decentralized accounting

The principal objective of accounting is attribution of resource consumption between a service provider and a client, likely for a further objective such as equitable cost allocation and billing. Currently, many service providers charge clients a flat fee for a fixed size pipe into their network cloud, independent of the amount of bandwidth the client actually consumes. For example, the National Science Foundation is currently funding all 45Mbps access lines from the attached NSFNET clients into the backbone, independent of actual line utilizations. Such a policy is feasible when a single agency funds the network service. But as a wider variety of commercial and non-commercial providers offer Internet services, and inequities among client usage of those services increase, the traditional flat-fee approach to service will become untenable.[1] Usage-sensitive accounting will become imperative.

However, the range of overlapping service providers that differ drastically in financial structure and policy will make the task of usage-accounting incredibly complex. We illustrate the scenario using the NSFNET backbone as an example. By providing the NSFNET as a high speed interconnection backbone, ss well as some funding to regional networks, the U.S. National Science Foundation currently supplies a significant fraction of wide-area interconnectivity for the U.S. research and education community. The backbone and regional networks function as transit networks to reach client campuses, hosts, users, and eventually applications. Each of these service territories: backbones, regional networks and campus clients, represents an autonomous networking entity or *Administrative Domains* (ADs). The Internet consists conceptually of a global collection of interconnecting ADs, which roughly map to network service providers. Furthermore, although originating in the research and education community, the collection of AD's now extends to commercial service providers and to an international scale.

The complexity of today's widely distributed global Internet make any centralized accounting scheme impractical at best; a reasonable mechanism must be *decentralized.* Optimally, a given service provider would only need to attribute consumption of his network resources to immediately adjacent clients, using a traffic matrix among the interfaces to those immediately attached clients. The provider could then bill these clients without regard for other components of the Internet system. This *adjacency* requirement, depicted in Figure 1, is not a new notion. Mills *et. al* [4] describe a decentralized scheme for the Internet environment which relies on interactions between adjacent Administrative Domains (ADs).

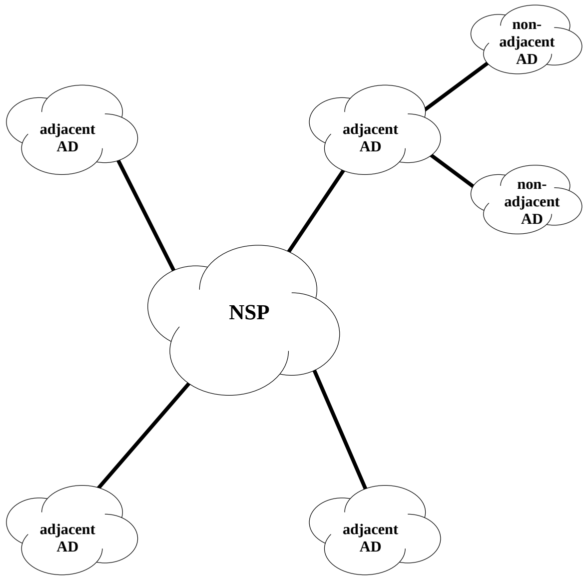

Figure 1: Adjacency

This scheme is *hierarchical* in a certain sense: each service provider constitutes the root of a tree, with adjacent ADs as branches at the next level that receive accounting information from their clients.[2] These specifications allow distributed accounting in the meshed Internet environment among many Administrative Domains, but do not specify accounting details for non-adjacent ADs.

Unfortunately, an accounting scheme derived from only client matrices of service providers may prove insufficient if detailed accounting information is required at the Internet systems level. The complex nature of interconnection among Internet ADs requires that traffic typically cross several network service providers to get from source to ultimate destination. When Internet traffic crosses many ADs (Figure 2), an accounting matrix that con-

[1] Some commercial service providers have already introduced transport service products (ANS Litespeed, Alternet Low Volume) with lower access charges for infrequent or bursty use. ANS gateway customers can elect to pay according to a tiered structure if their average link utilization (SNMP 15 minute averages measured) falls within the 10%, 20%, 30% bands. Mills *et. al.* discusses other experiments that link statistics to billing [4].

[2] The hierarchy does not reflect the hierarchical architecture of the U.S. component of the Internet, which reflects a backbone–regional–campus–host–application 'hierarchy" which is itself only a rough abstraction of actual Internet connectivity.

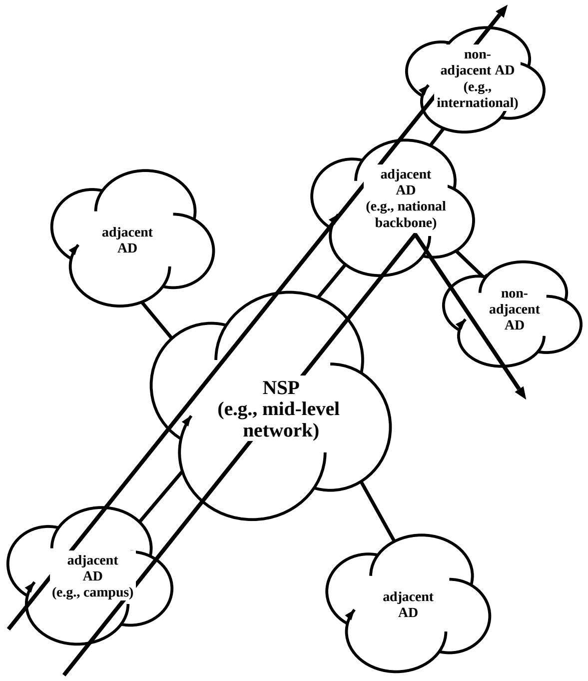

Figure 2: Accounting and end-to-end flows

siders only entry/exit flows relative to the local service provider will be insufficient to reattribute resource consumption to non-adjacent clients. Attribution across multiple ADs would require information of finer granularity, such as a matrix among end-systems in the network, or even based on IP address/mask pairs.

For example, a packet from a host at the University of California, San Diego to the University College in London would start at the UCSD host and via the local department, enter the campus network service provider, then cross the regional network (CERFnet), the NSFNET backbone, and then via some international connectivity to the University College London campus, to a particular department and finally a host on that campus. The NSFNET would be able to, using the client matrix, attribute the flows from the regional network to the international connection. However without more detailed accounting the NSFNET would not be able to provide the regional network with information as to who within the regional network was responsible for the international traffic.

Thus, for attached clients of backbone transit networks to reattribute resource consumption within their sphere, those clients will need finer-grained accounting information, e.g., hosts or address/mask pairs. In Section 5 we describe an example where a transit network is able to collapse finer granularity traffic matrices into the *entry/exit point matrix* discussed above, using internally-known routing configuration information. The transit network could then supply the finer-grained information to each access point to allow reattribution.

4 Precedence

Today's Internet is inherently based on a datagram architecture with no admission control in packet forwarders. Most entrance points into transit networks can not sufficiently provide back pressure to peer points that inject more traffic than the transit network can handle. This characteristic pervades the Internet environment, leaving end systems able to unfairly monopolize available bandwidth and cause significant congestion in the transit networks they use.

During the mid-80s on the 56kbps NSFNET backbone, this state of congestion developed to a dangerous degree, and in response the NSFNET engineers deployed an emergency measure to provide certain interactive network applications, specifically Telnet, preferential treatment over other traffic. The priority transit allowed interactive users requiring better network responsiveness to continue working under highly congested circumstances. At the same time the NSFNET backbone established support for separate queues in the routers according to the *precedence* value in the IP header field.

When the NSFNET was upgraded to T1 capacity, offering a 24-fold bandwidth increase and a richer topology, the designers did not re-introduce the priority queuing for end-user traffic. The new infrastructure used multiple queues only to differentiate between user traffic and network management traffic. An overabundance of bandwidth rendered superfluous the use of multiple queues. In the case of the NSFNET backbone, the project partners bore all the costs of maintaining this bandwidth ahead of demand. The subsequent upgrade to the T3 network exemplified further this method of coping with anticipated network congestion by increasing the bandwidth and switching capacity.

However, today software developers continue to build advanced network applications which can consume as much bandwidth as network engineers provide. In particular, real-time applications using packet voice and video do not exhibit the same burstiness characteristics of more conventional applications such as file transfer and electronic mail, but rather require prolonged delivery of large amounts of traffic in real-time, and thus continuously consume significant amounts of bandwidth. Clearly usage of such applications will not scale in the current Internet architecture, which may potentially need to support many such continuous point-to-point connections simultaneously. Most video and audio applications are currently still in a prototype phase; their wider-spread deployment bodes ominously for an infrastructure not able to preferentially deal with certain traffic types. Indeed, it is difficult to overestimate the dramatic impact which digital continuous media will have on the Internet fabric. No other phenomenon could more strongly drive the research community to instrument the network for admission control, as well as accounting and billing.

We advocate the application of the experience gained with NSFNET in 1986 on a wider scale, where Internet hosts and routers integrate support for the IP precedence field into their design specifications, and end-systems begin active use of this field as quickly as possible. Currently rarely used (set to 0 by default), the IP precedence field allows the originator of traffic to select among multiple service levels. An end-user can thus actively request a traffic flow with a chosen priority level, along with its associated ramifications (price, AUP compliance needs, etc.). One possible scenario is for the service provider to provide a volume insensitive baseline service, at minimal or zero cost, with no performance guarantees. A value different from the baseline value in this field would then indicate the purchase of a higher level service, and clients using those values would pay higher service fees. Although current routers do not necessarily support this component of the IP specification, the IP standard [5] documents it, which should allow vendors to incorporate it into their IP implementations.

We note that in a usage-insensitive environment, such as the traditionally bulk-funded NSFNET, users would likely misuse the power to specify a precedence value, routinely demanding higher service precedence. As various components of the commercializing Internet begin to introduce usage- and service-sensitive accounting systems, through which users eventually pay for the services they request, precedence makes considerably more sense.

5 Current instrumentation

In this section we discuss the applicability and difficulty of the two aspects we propose given current Internet instrumentation and architecture, and present relevant accounting statistics from active Internet components.

Unfortunately, a substantial fraction of current accounting instrumentation on the Internet only provides for traffic volume counters (packets/bytes) at individual interfaces. While these statistics allow for general assessment of traffic flow from/to a particular client network, one cannot use them to assess flows through the network cloud. The latter task requires a matrix of entrance and exit points discussed in Section 3. A network service provider could derive such a matrix using the point at which the traffic came into the provider's cloud, the IP destination network number, and internal routing information mapping the destination IP address to the (principal) transit network exit point. Support for IP Precedence based queuing as described in Section 4 would allow the provider to derive such intra-NSP matrices for multiple service levels.

Although much of the Internet could not currently support an entry/exit point matrix or precedence based accounting, some service providers do currently have a mechanism to derive the former.[3]

6 Entry/exit point matrix

We offer an example using operationally collected data from the NSFNET backbone. The NSFNET currently collects source/destination IP network number matrices of packet and byte counts at backbone inflow nodes.[4] We collapsed the month of matrices (originally collected in fifteen-minute intervals) down to a single matrix of flows among all external access points of the NSFNET cloud, which we offer as a prototype NSP/client matrix. Because the collection occurs at inflow node the resulting matrices automatically indicate the source inflow point of the traffic. The NSFNET also maintains a database of primary and backup exit ADs for individual network numbers, which can then map to backbone traffic exit nodes. We thus derive an entry/exit traffic matrix relative to the backbone based on the assumptions that packets typically follow their primary path through the backbone itself.

Given such a matrix, the NSFNET could conceivably send a usage bill to each access point listing the charges attributed by traffic sent to all other access points in the backbone. Figure 3 illustrates an example distribution of traffic volume from one access point to other NSFNET access points for the given time period. We do not go into a derivation of a line item cost here, leaving that to a specific pricing and billing implementation. The transit service provider, in this case NSFNET, could supply with the bill the set of mappings between IP numbers/masks and exit points which it used to derive its statistics. If an access point were to map to a regional service provider, the provider could then use this information to reattribute costs to its own clients appropriate to their usage.

[3] For example, Cisco routers implement an accounting scheme which maintains packet and bytes counts between host pairs at each outbound interface. The NSFNET backbone also collects packet and byte volume matrices, at the traffic inflow node, but bases them on network numbers rather than hosts.

[4] For performance reasons the T3 backbone routers which the NFSNET uses currently support this statistics collection by sampling every 50th packet. Any accounting scheme is subject to the accuracy of the collected data. In particular, sampling will affect the property of verifiability, i.e., the capability of a network client to independently verify his assessed charges at any time. Claffy *et. al* [2] [1] discusses sampling issues further.

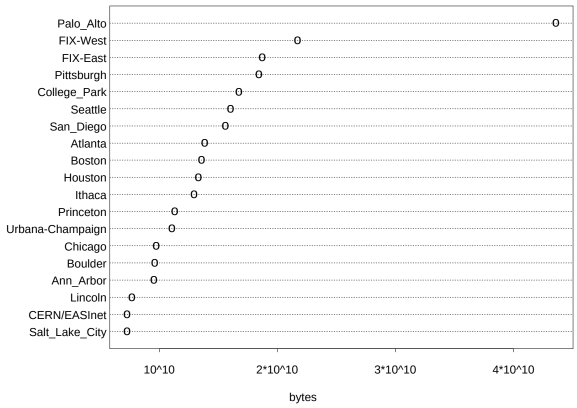

Figure 3: Distribution of traffic volume from NSFNET access point in San Diego to other NSFNET access points

Implementing the proposed scheme on the NSFNET would involve two augmentations to the current mechanism: the ability to collapse the IP source/destination information to network entry/exit point matrix information; and adding a separate matrix for each IP Precedence value.

7 Conclusion and future directions

We have proposed a decentralized multiple service-level accounting scheme for the Internet. Our two-component scheme relies on existing features of the Internet architecture and fabric. The first component of our scheme involves decentralized accounting at various required levels of granularity. Usage-based decentralized charging between service providers and clients in the near term would allow more equitable cost distribution among clients than traditional flat fee schemes. Secondly, we propose active use of the IP precedence field, in order to allow a wide spread implementation of multiple qualities of service within the realm of the current IP specification and architecture.

This scheme may prove inadequate in providing the required level of detail in the face of accounting requirements stretched across multiple ADs, and we consider it critical to Internet evolution to continue investigations into accounting requirements and limitations. Although there is elegance and simplicity in restricting accounting relationships to adjacent providers and clients, we surmise that reasonable accounting cannot withstand this restriction: exchanged accounting information must include enough information so that clients of service providers can reattribute costs to their own clients. In this paper we offer an example where a network service provider could include in a usage bill additional information regarding network numbers to allow a certain degree of reattribution. However, as ominous as it sounds, we may find that at least some network components will have to support even finer-grained accounting, perhaps based on an end-host pair granularity, in order to reasonably attribute costs in the large, non-homogeneous Internet. We are currently collecting and analyzing specific data from actual Internet components, in order to explore issues such as how to compound the resource consumptions for switched objects (packets) and their payload (bytes per packet), and qualification of the collection reliability, including computational, memory, and bandwidth burdens.

Nonetheless, we consider it critical that the community begin deployment of usage- and precedence-based accounting at least at NSP/client boundaries to gain more insight into resource consumption profiles for a wide variety of network environments. The exercise will not only facilitate future efforts in network billing, but also provide superior data for network engineering and planning.

Our objective has not been to describe a complete accounting and billing system. We focus only on the development of a sturdy accounting framework on which a billing structure can be architected. In pursuit of incremental progress toward a satisfying accounting framework we propose combining two mechanisms, decentralized accounting and multiple service queues, to yield a baseline for Internet transit billing and more equitable attribution of Internet resource consumption.

Acknowledgements

We would like to express our appreciation to Jordan Becker of ANS for making data available to us and we would like to acknowledge the helpful discussions and exchange of ideas on this topic with, Jordan Becker of ANS, and Dr. Roger Bohn and John May of UC, San Diego.

References

[1] K. Claffy and H.-W. Braun. Network analysis issues for a public Internet. To appear in Proceedings of *Public Access to the Internet*, May 1993.

[2] K. Claffy, H.-W. Braun, and G. Polyzos. Application of sampling methodologies to wide-area network traffic characterization. In *Proceedings of SIGCOMM '93*, September 1993.

[3] P. Jones. Resource allocation, control, and accounting for the use of network resources. RFC 1346, June 1992.

[4] C. Mills, D. Hirsh, and G. Ruth. Internet accounting: Background. RFC 1272, November 1991.

[5] J. B. Postel. Internet Protocol. Internet Request for Comments Series RFC 791, 1981.