Tracking Long-term Growth of the NSFNET

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

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

Abstract

We present the architecture for data collection for the NSFNET backbone and difficulties with using the collected statistics for long-term network forecasting of certain traffic aspects. We describe relevant aspects of the NSFNET backbone architecture and the instrumentation for statistics collection. We then present long-term NSFNET data to elucidate long-term trends in both the reachability of Internet components via the NSFNET as well as the growing cross-section of traffic. We focus on the difficulties of forecasting and planning in an infrastructure whose protocol architecture and instrumentation for data collection was not designed to support such objectives.

I. Introduction

While initially conceived as a demonstration project of a then new networking technology for the United States federal government, today's Internet aggregates traffic from a far wider set of constituencies. As the number of client networks of the Internet heads into the tens of thousands, the image of a ubiquitous network, relying on globally shared resources, has already become a reality. This paper reports work done to examine traffic aspects of an existing Internet backbone: the National Science Foundation Network (NSFNET) backbone.[1] We present an overview of the data collection architecture for NSFNET and then evaluate some collection mechanisms now being used to make long-term forecasts of certain of its traffic aspects. Data was collected using two methods: IP network numbers were used to track the geographic and administrative scope and growth of the Internet, and port numbers to assess the growing cross-section of traffic. We focus on the limitations of these two methodologies, both of which were designed initially to support short-term engineering and planning needs, such as routing and tracking the rough cross-section of traffic. Aspects of their architecture and implementation prevent effective usage for some long-term forecasting and planning objectives. For example, the Internet architecture makes it inherently difficult to track many applications by TCP/IP port number. This situation poses a serious obstacle to long-term planning because of the growth of real-time continuous media applications, which can consume significant fractions of the available bandwidth over long periods of time. This paper discusses these issues in the following sections.

II. NSFNET backbone architecture

Figure 2 depicts the T3 NSFNET backbone architecture, which replaced the earlier T1 architecture in 1992. (See sidebar for a description of the NSFNET.) The backbone nodes are core packet switches in the NSFNET infrastructure. On the T3 network, backbone nodes are designated as either Exterior Nodal Switching Subsystems (ENSSs) or Core Nodal Switching Subsystems (CNSSs). ENSSs are located on the client network premises, and CNSSs are co-located at carrier switching centers known as "points-of-presence" (POPs) or "junction points." Co-location of the core cloud packet switches within POPs provides several advantages. First, because these locations are major carrier circuit switching centers, they are staffed around

[1] A model of the Internet interconnectivity architecture is shown in Figure 1.

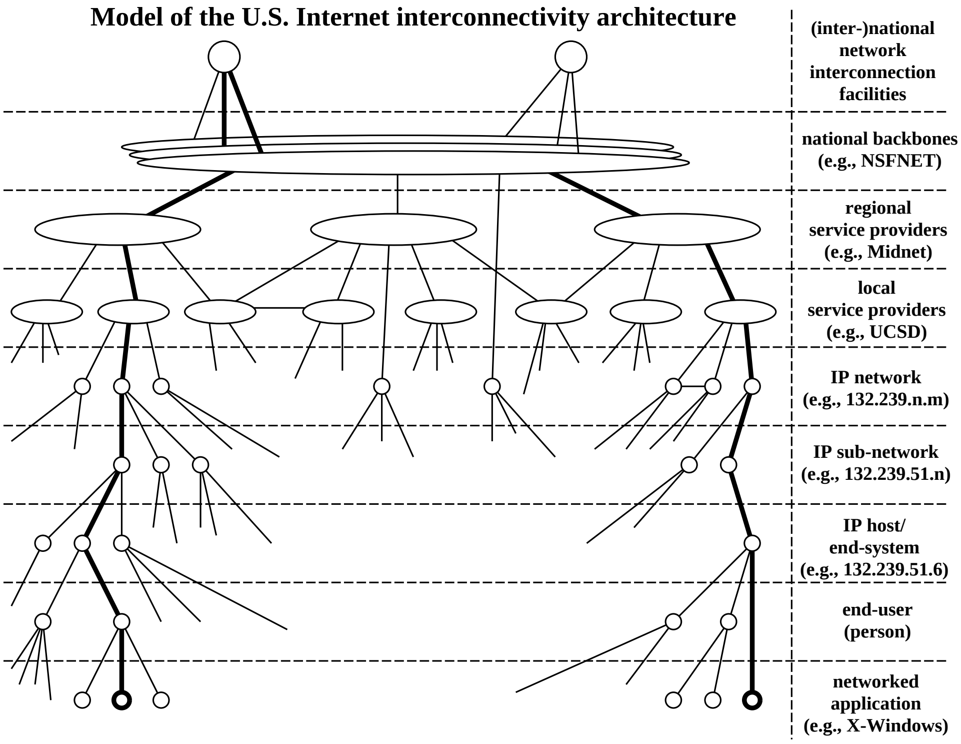

Figure 1: Model of U.S. Internet interconnectivity architecture

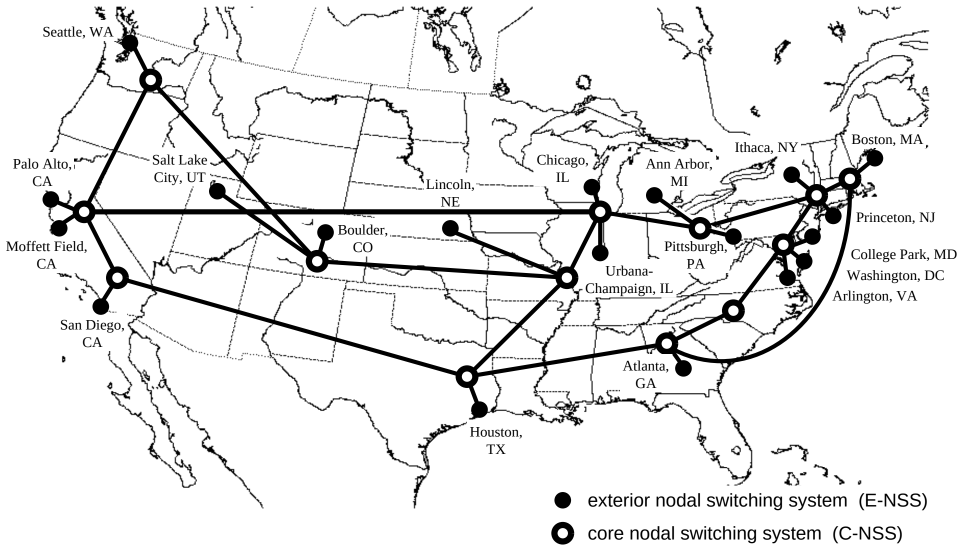

Figure 2: 1993 NSFNET T3 Backbone Service Logical Topology

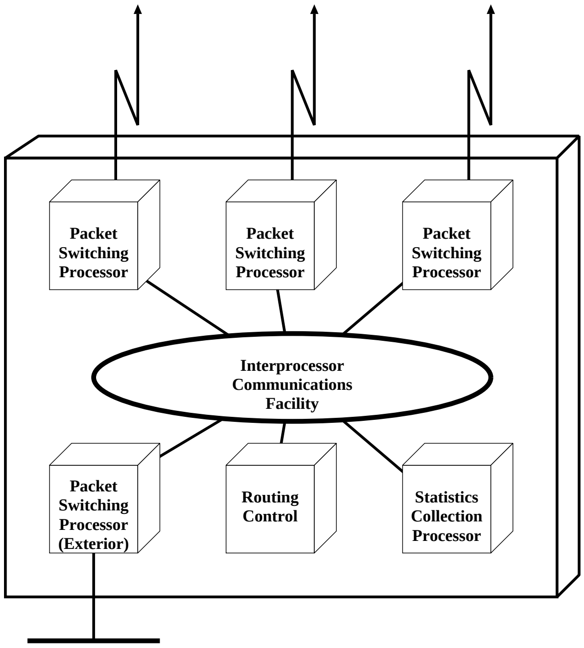

Figure 3: T1 NSS architecture

the clock and have full backup power, which is essential to network stability. Second, this co-location allows the addition of new clients (e.g., ENSS nodes) to the network by connecting them to a CNSS with minimal or no service disruption to other CNSS/ENSS clients. Co-location also allows network designers to align the carrier-provided, circuit-switched network topology more closely with the packet-switched backbone topology, facilitating path redundancy.

Figures 3 and 4 illustrate the Nodal Switching Subsystem (NSS) architectures for the core backbone nodes on the T1 and T3 backbones, respectively. The T1 NSS architecture consisted of multiple, typically nine, IBM PC/RT processors connected by a common token ring. In contrast, the T3 backbone packet forwarding routers are based on the IBM RISC System/6000 architecture, with special modifications including high-performance adapter cards and software. Initially, the interfaces to this uniprocessor architecture switched packets through to the outgoing interfaces via the main CPU. In the current implementation, the packet forwarding process is off-loaded onto intelligent subsystems. Each external interface, including T3 serial lines as well as connected Ethernet and FDDI LANs, lies on such a dedicated subsystem card. These cards have a built-in Intel 960 microcontroller on board and have local access to all information needed to switch a packet, including routing tables and relevant code. The cards can thus exchange packets among each other directly via the IBM Microchannel bus without the intervention of the main processor.

III. NSFNET statistics collection mechanisms

There are three classes of network statistics gathered from NSFNET information sources: interface statistics; packet categorization; and internodal delays. Interface statistics are produced from programs using the Simple Network Management Protocol (SNMP) [5]. Specialized software packages perform packet categorization: the T1 backbone used NNStat [3] for collection; the T3 backbone uses ARTS (ANSnet Router Traffic Statistics) [1], a package with similar functionality. Internodal delays were measured with the ping utility for the T1 backbone; on the T3, ANS collects roundtrip delay statistics between both external and internal access points with the yet-another-ping (yap) utility.

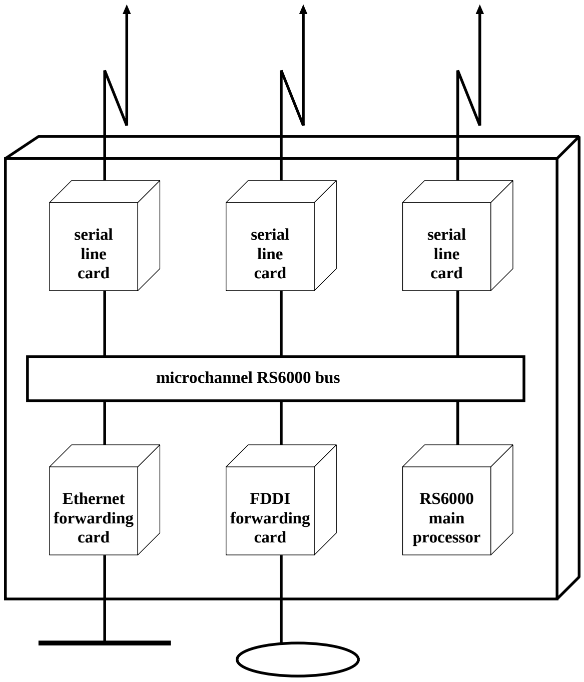

Figure 4: T3 C-NSS architecture

III.A. Interface performance

The mechanism for collecting interface performance statistics did not change from the T1 to the T3 backbone. To maintain data regarding packets and bytes transmitted and received, errors, delay times, and down times, all backbone nodes record statistics about the packets which traverse each of their interfaces. Each backbone node runs SNMP servers which respond to queries regarding SNMP Management Information Base (MIB) variables. Centralized collection of the data from each backbone interface on each NSS occurs once every 15 minutes. The counters are cleared in only two cases: when the machine is restarted; and when the 32 bit counters overrun. Cumulative counters, retrieved using SNMP, include those for packets, bytes, and errors transmitted in and out of each interface.[2] Table 1 compares the SNMP objects collected on the T1 and T3 backbones. Among other changes, the T3 backbone now supports counters of non-unicast packets.[3]

III.B. IP packet categorization

Unlike the SNMP statistics, data collection for packet categorization was modified for the T3 backbone. As described above, each T1 backbone node (NSS) was a set of interconnected IBM RT/PC processors, one of which was dedicated to statistics collection, in particular, to categorize IP packets entering the T1 backbone. The statistics gathering processor examined the header of every packet traversing the intra-NSS processor intercommunication facility and used a modified version of the NNStat package [3] to build statistical objects from the collected information. Because all packets traveled across the interconnection facility on their way through the node, the collection processor passively collected data without affecting switching throughput. Nonetheless, the nodal transmission rate did eventually surpass the capability of the collection processor to keep up with the statistics collection in parallel, and this processor had to eventually revert to sampling [7].

The T3 backbone design required significant data collection modification. In the first statistics collection design, all forwarded packets had to traverse the main RS/6000 processor itself, imposing a burden on the

[2] Error conditions on the interface include HDLC checksum errors, invalid packet length, and queue overflows resulting in discards. The single counter does not distinguish among these error conditions.

[3] Object definitions found in McCloghrie and Rose [15] [16].

Table 1: SNMP objects collected per node on T1 and T3 backbones

object	description	**T1**	**T3**
ifOperStatus	operational status	Y	N
sysUpTime	system uptime	Y	Y
ifDescr	interface descriptors	Y	Y
ipAdEntIfIndex	IP address corresponding to interfaces	Y	Y
is-isIndex	remote address to interface index mapping	N	Y
ifInErrors	incoming errors occurring interface	Y	Y
ifOutErrors	outgoing errors occurring on interface	Y	Y
ifInOctets	bytes entering interface	Y	Y
ifOutOctets	bytes exiting interface	Y	Y
ifInUcastPkts	unicast packets entering interface	Y	Y
ifOutUcastPkts	unicast packets exiting interface	Y	Y
ifInNUcastPkts	non-unicast packets entering interface	N	Y
ifOutNUcastPkts	non-unicast packets exiting interface	N	Y

Table 2: Packet categorization objects on T1 and T3 backbone nodes

Object	**T1**	**T3**
relative to exterior nodal interface		
source-destination matrix by network number (packets, bytes)	Y	Y
TCP/UDP port distribution, well-known subset (packets, bytes)	Y	Y
distribution of protocol over IP (e.g., TCP, UDP, ICMP) (packets, bytes)	Y	Y
packet-length histogram at a 50-byte granularity	Y	N
packet volume going out of backbone node	Y	N
NSS-centric (entire node)		
per second histogram of packet arrival rates	Y	N
NSS (intra-NSFNET) transit traffic volume	Y	N

single packet forwarding engine which intensified with comprehensive statistics collection. Figure 4 illustrates the current design of the backbone nodes, which offloads the forwarding capability to the cards as described above. Because the packets do not necessarily traverse the main processor, T3 statistics collection required moving the software that selects IP packets for traffic characterization into the firmware of the subsystems. Each subsystem copies its selected packets, currently every fiftieth, to the main CPU, where the collection software performs the traffic characterization based on these sampled packets. Note that multiple subsystems, including those connected to T3, Ethernet, and FDDI external interfaces, forward to the RS/6000 processor in parallel.

Because the main CPU performs the categorization, the statistics aggregation mechanism does not affect NSS switching throughput. The sampling and the forwarding of the samples, however, can burden the subsystem-to-CPU bandwidth, and it can potentially interfere with other critical responsibilities of that bus, such as transferring routing information between the subsystem and the CPU. Although the packet categorization mechanism at each node differs on the T1 and T3 backbones, the centralized collection of the data is the same. Every fifteen minutes, a central agent queries each of the backbone nodes, which report and reset their object counters. The collection host is an IBM RS/6000 at the ANS Network Operations Center (NOC). As an example of the memory requirements, this machine collected approximately 25 MB of ARTS traffic characterization data during a typical workday in mid-February 1993; daily figures for January 1994 exceeded 40 MB. Table 2 illustrates the traffic characterization objects collected on the T1 and T3 backbones. The first item in Table 2, the matrix of network-number-to-network-number traffic counts, forms the basis for characterizing traffic across the NSFNET backbone, by both individual network numbers and countries.

III.C. Internodal latency

On the T1 backbone, Merit Network, Inc. (formal manager and operator of the NSFNET backbone), used the ping utility for internodal latency assessments. Ping probes from one endpoint of the network to another using the ICMP Echo functionality [18] to record the round-trip times (RTT) between the two endpoints.

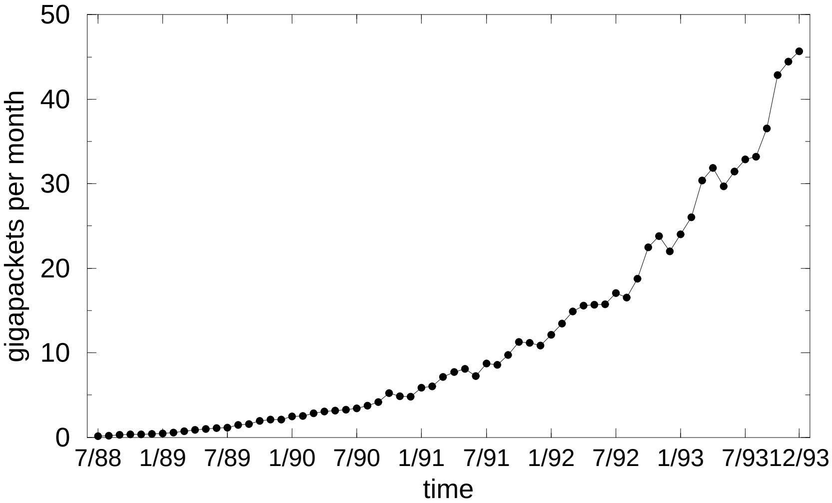

Figure 5: Long-term growth of packet volume into the NSFNET *(Data source: Merit/NSFNET operations)*

Since 1 February 1993 ANS has collected similar delay data between nodes using the yap utility, which runs on each backbone node and can measure delay to the microsecond level using the AIX system clock. Probe measurements occur five times at the beginning of every fifteen minute interval between all pairs of backbone access points (i.e., ENSSs and CNSSs). Each month ANS produces several tables of medians, interquartile differences (IQDs), and filtered delay values between all pairs of ENSSs; their analysis highlights features or changes in backbone delay performance that may require attention.

IV. Using IP addresses to show network growth

Figure 5 shows quadratic growth in packet volume during the last five years of the NSFNET backbone operation. Besides sheer traffic volume, another growth metric is the number of connected networks as seen through traffic reaching the NSFNET backbone.

Figure 6 shows the long-term growth of network numbers configured for communication via the NSFNET backbone [17]. These NSFNET numbers are the only destinations to which the NSFNET backbone will route packets; every network client must have an assigned IP network number to receive traffic. The figure shows dramatic growth over the last few years, including substantial increases in the international area.

One must be careful in mapping the metric of assigned IP network numbers to the growth of the Internet; not all IP addresses are created equal. Each IP address is four bytes long, part of which identifies the IP network and the remainder of which identifies the host within that network. The IP protocol specifies three commonly used address classes which differ in the size of the host address component within the four byte IP address. Class A, B, and C networks have three, two, and one-byte host field, allowing for a maximum of 2^{24}, 2^{16}, and 2^{8} individual addresses or hosts, respectively. The number of allocatable Class A, B and C network numbers is 2^{7}, 2^{14}, and 2^{21}, respectively [14].

Over the years the Network Information Center (NIC) assigned IP network numbers to clients according to the number of hosts to be supported. Class B addresses were most attractive due to their size. Eventually the InterNIC imposed restrictions on Class B address assignment, using instead groups of Class C addresses unless a Class B space were really justified. The policy resulted in a deceleration of Class B assignments, but accelerated the assignment of Class C addresses significantly.

Since addresses from different classes absorb different amounts of the available 32-bit address space, Figure 6, which depicts the growth in *network number* addresses regardless of address capacity, offers only part of the story. Figure 7 presents more of the picture, showing the growth in *total committed address space*. Currently, about 60% of the total available address space is assigned. In terms of total committed address

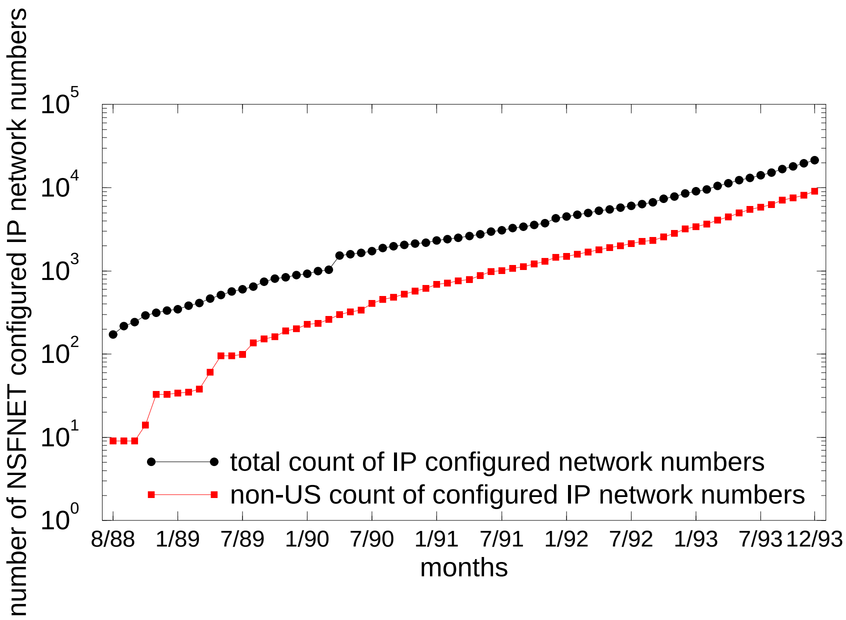

Figure 6: Long-term growth of network numbers served by NSFNET *(Data source: Merit/NSFNET operations)*

space, not network numbers, the Class A and Class B growth has in the past dwarfed the Class C growth. However the figure depicts an exponential growth in Class C committed space, which is accelerating since the new InterNIC policy. If current growth patterns persist according to this figure, the exponential growth of the Class C *committed address space* will outperform the effects of the Class A and Class B changes. See the side-bar "What's in an IP address?" for more on the significance of the IP addresses structure to the infrastructure and service reachability of the NSFNET.

More recently, in an attempt to diminish the size and associated processing requirements of the routing tables that Internet routers must exchange, the Internet community is adopting Classless InterDomain Routing (CIDR) [10]. CIDR breaks the classic bit-boundaries among Class A, B and C networks and aggregates blocks of network numbers to assign to service providers in order to allow the routing table to store store address identifiers consisting of a 32 bit number field and an associated 32 bit mask to identify the CIDR block portion of the 32 bit field.

Assigned IP network numbers are necessary but not sufficient for communication across the NSFNET backbone, which is an important component of the global Internet. The NSFNET backbone uses a policy routing database as a truth filter to verify the validity of dynamic routing information that its backbone clients have explicitly specified. This database represents the set of NSFNET-configured network numbers that the NSFNET serves, a proper subset of the assigned network numbers. However, even though a network may be in this NSFNET database, the backbone still will not be able to service that network until it receives a dynamic announcement from that network via a router of a directly attached NSFNET client by means of an inter Administrative Domain protocol such as BGP or EGP. The announcement from an NSFNET client (either a mid-level or some other network connected to the backbone) reaches the NSS, which evaluates each incoming announcement, accepting those for configured nets that come from appropriate peers in an appropriate Administrative Domain (identified by its autonomous system number). This action turns an NSFNET-configured network into an NSFNET-announced network. The configuration database thus serves to sanitize dynamically announced routing information before the backbone actively utilizes it. This filtering is essential to the sanity of the larger infrastructure, and other networks often use similar mechanisms to accomplish the same task. Upon acceptance of the announcement, the NSS tags a path priority value, or metric, to the network number, to enable comparison to other announcements of the same network number.

After a network is assigned, configured, and announced, it can both send and receive traffic over the NSFNET backbone as an active network. A network remains active as long as connectivity exists to the destination and the appropriate service provider(s) announces the network directly to the backbone.

As mentioned above, Internet reality is not entirely faithful to this model. Theoretically, any network

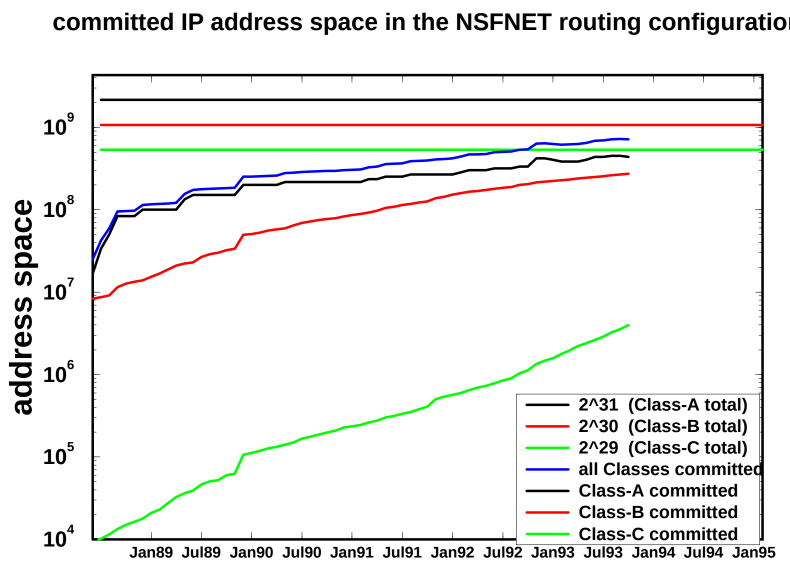

Figure 7: Long-term growth of assigned Class A, B, and C IP network numbers *(Data source: Internic* `rs.internic.net`*)*

that sends traffic is active, even if it is not assigned, configured, and announced. To make these categories unambiguous, we call an illegitimately active network, i.e., a network missing any one or more of the three essential properties (assigned, configured, and announced), a leaky network. Leaky networks, particularly from unassigned networks, pose difficulties for network operators because they can inject incorrect information into inter-administrative domain routing protocol exchanges. Operators, such as those of the NSFNET backbone, must sanitize their network topology information. For this reason, the Internet Assigned Numbers Authority (IANA) discourages the TCP/IP community from using self-selected network numbers, considering it "uncivilized" behavior in the increasingly, and often transparently, interconnected world.

Also problematic is the case of silent networks. Silent networks are those configured or announced but not active; i.e., they have not sent traffic across the backbone during the month they were configured and announced. The NSFNET project has analyzed the NSFNET Policy Routing Database (PRDB) and has developed methods to eliminate silent nets to prevent operational difficulties because of routing table size. But the number of silent networks has increased since announcement of the addressing guidelines outlined above. When service providers receive large blocks of Class C addresses in anticipation of and aligned with CIDR [10] requirements, they immediately configure the addresses in the policy routing database before assigning them to customers. This increases the number of silent networks in the NSFNET backbone configuration database. Eventual CIDR deployment will rely on net masks to reduce such blocks to a single entry in the routing table, but until that time, they still pose an obstacle to efficient configuration.

Figure 8 presents a schematic of the categories of network numbers we have discussed. Engineers on operational networks must contend with these issues in the design of their traffic collection mechanisms. For example, each NSFNET T3 backbone router samples every fiftieth packet to build traffic characterization objects, in particular to create a source-destination matrix by network number. The router samples these packets before routing them, and thus, it is conceivable that routers can capture packets from IP network numbers not in the routing database. We call these unassigned networks, although anarchically selected network numbers may be a more fitting term.

For the purpose of statistics analysis, there are a variety of ways to treat inactive networks, including treating them all as unconfigured. This, however, risks attributing the traffic they impart to the backbone entry point. As an example, consider December 1992. During this month, the data collection mechanism on the T3 backbone nodes recorded traffic from more than 14,000 networks. This set included networks from the entire set of available network numbers. Of these, about 9,700 were networks in the set of NIC-assigned network numbers. The number of active networks that had also been configured in the NSFNET/ANSnet

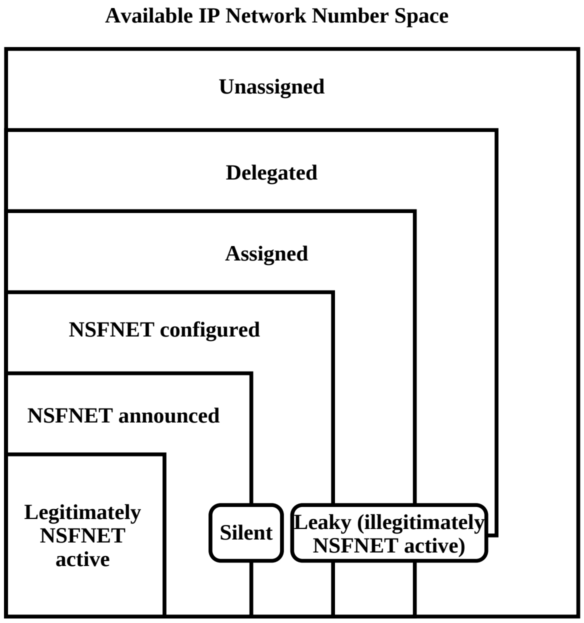

Figure 8: Descriptive categories of IP network numbers

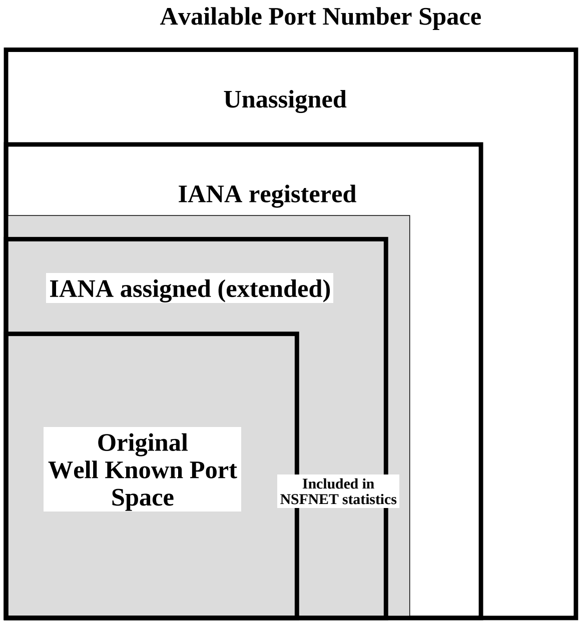

Figure 9: Descriptive categories of port numbers

topology database that month was about 6,400.

To explain the large number of non-configured networks represented in this traffic matrix, we use terminology outlined above to describe several violations of our model. A non-configured, leaky network may send traffic into the NSFNET backbone. Those packets may actually get delivered to the remote location, if the remote location is assigned, configured, and announced. But traffic for the return path to the original source will not be delivered via the NSFNET.

Alternatively, a network could send traffic to another network unrecognized by the NSFNET backbone because the network is configured but not announced, because the network is assigned but not configured; or because the network is not assigned at all. The NSFNET may have such traffic, for example, when network service providers use a default route pointing to the NSFNET. However, because routing information for these destinations will not exist in the NSFNET forwarding tables, as soon as such packets reach the NSFNET, the backbone node filters them out during the routing decision.

V. Growth in application and service diversity

Another difficulty in characterizing long-term traffic trends on the NSFNET is the increasing number of applications. Assessment of the kind and scope of these applications is critical to network service planning. For example, how do continuous media such as audio and video impact the performance of conventional bursty traffic. Tracking the traffic cross-section is becoming increasingly difficult.

Most applications on the NSFNET are built on top of the Transmission Control Protocol (TCP) or the User Datagram Protocol (UDP). Both TCP and UDP packets use port numbers to identify the Internet application that each packet supports. Each TCP or UDP header has two 16-bit fields to identify the source and destination *ports* of the packet. Originally, the Internet Assigned Numbers Authority (IANA), at ISI (Information Sciences Institute, University of Southern California), on behalf of DARPA, administered a space of 1 to 255 as the group of port numbers assigned to specific applications. For example, telnet received port assignment 23 [13]. To open a telnet connection to a remote machine, the packet carries the destination IP address of that machine in its destination IP address field, and the value of 23 in the destination port field.[4]

During the early years of the TCP/IP-based Internet, particularly in the early eighties, Unix developers injected a bit of anarchy into the IANA system when they unilaterally began using numbers between 512 and 1024 to identify specific applications. For example, they used port 513 for rlogin. Eventually network users started to use numbers above 1024 to specify more services, extending the lack of community coordination. In July 1992 [13], the IANA extended the range of port assignments they manage to 0-1023. At this time, they also began to track a selected set of registered ports within the 1024-65535 range. IANA does not attempt to control the assignments of these ports; they only register port usage as a convenience to the community [13]. Figure 9 presents a schematic of the port number categories discussed.

These port numbers are the only mechanism through which the NSFNET can monitor statistics on the aggregated distribution of applications on the backbone. Specifically, Merit (and now ANS) collects port-based information in the ranges 0-1023, 2049 (for NFS) and 6000-6003 (for X-window traffic). Merit/ANS categorizes packets into these ports if either the source or destination port in a given packet matches one of these numbers. However, even within this range, not all ports have a known assignment. So packets using undefined ports go into the *unknown* port category [17].

Figure 10 uses this data to classify the proportion of packet traffic on the network by category since August 1989, based on categories in use by the NSFNET backbone service provider. These figures indicate an increasing diversity in the cross-section of NSFNET traffic, and a decreasing reliance on conventional protocols such as telnet, FTP, and SMTP, relative to the overall network traffic. [5]

The categories in these figures correspond to:

- File exchange: FTP data and control (TCP ports 20, 21)

[4] In the case of telnet, the packet uses some arbitrarily assigned source port that has significance only to the originating host. Often these "return address ports" have values greater than 1000.

[5] Note that Merit began to use sampling for this collection on the backbone in September 1991. In November 1991 traffic migration to the T3 backbone began; the majority of the links had migrated by May 1992 and in November 1992 the T1 backbone was dismantled. For June to October 1992 no data was available for either the T1 or T3 backbones.

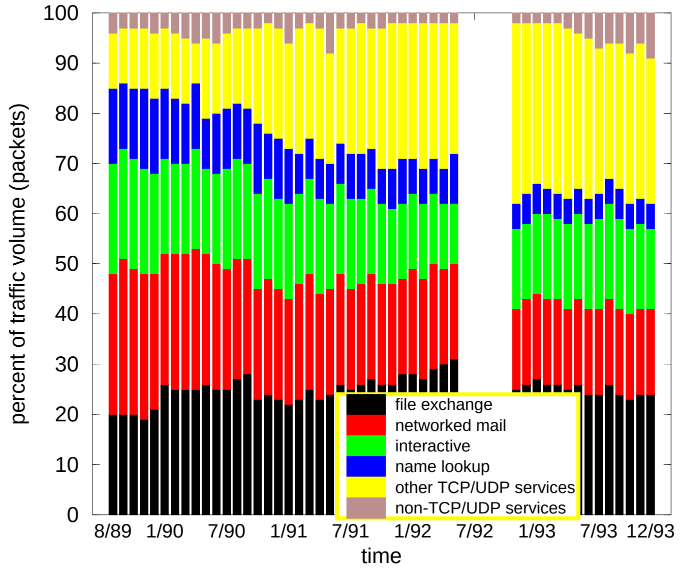

Figure 10: Distribution of the number of packets offered into NSFNET backbone by application *(Data source: Merit/NSFNET operations)*

- Mail: SMTP, NNTP, VMTP, UUCP (TCP ports 25, 119, 175, 540)
- Interactive: telnet, finger, rwho, rlogin (TCP ports 23, 79, 513, UDP port 513)
- Name lookup/DNS: (UDP port 53, TCP port 53)
- Other TCP/UDP services: all TCP/UDP ports not included above (e.g. irc, talk, X-windows)
- Non-TCP/UDP services: Internet protocols other than TCP or UDP (e.g. ICMP, IGMP, EGP, HMP, etc.)

More detailed distribution of traffic by port on the NSFNET backbone is available via ftp from `nis.nsf.net`, and shows some details regarding the growing range of applications. Several Internet resource discovery services (e.g., WAIS, WWW, gopher, prospero, mosaic [9]) have experienced tremendous growth in volume since their deployment. Other applications have also gained a greater proportion of network bandwidth: e.g., MUD (Multi-User Domain), a distributed electronic role playing environment; X-windows, for remote graphical displays across the network; and more recently real-time applications like packet video and audio. Many of these applications use multiple TCP/UDP port numbers, which often are not centrally coordinated and therefore unknown to anyone but the end sites using them. Such traffic is a subset of the "other Protocols" category in Figure 10, and add complexity to the task of Internet workload characterization. The number and traffic volume of non-categorized applications has grown much larger over the years, reflecting an increasingly multi-application environment, and a diminishing ability to assess the impact of individual new applications due to their use of non-standardized port numbers.

VI. Forecasting traffic type

The issue of unknown applications is not by itself as disturbing as the dramatically changing nature of the newly introduced traffic. The recent deployment of more widespread packet video and audio applications bodes ominously for an infrastructure that is not able to preferentially deal with certain types of traffic. The dangers of our increasing inability to monitor traffic type in a "high-end" Internet loom.

Today's Internet is based on a datagram architecture with typically no admission control in packet forwarders. Most entrance points into transit networks can not provide back pressure to other points of the network that inject more traffic than the network can handle. End systems can thus unfairly monopolize available bandwidth and cause significant congestion in the larger network.

During the mid-80s on the T1 NSFNET backbone, congestion developed to a dangerous degree. In response, the NSFNET engineers deployed an emergency scheme to provide certain interactive network applications, specifically telnet, preferential treatment over other traffic. Such priority transit allowed interactive users, who required better network responsiveness, to continue working under highly congested circumstances.

Since that time, the principal means of addressing network congestion has been to increase network capacity. However, today software developers continue to build advanced network applications that can consume as much bandwidth as network engineers can provide. In particular, applications using packet audio and video require continuous delivery of large amounts of traffic in "real-time," and, thus, continuously consume significant bandwidth. Clearly, usage of such applications will not scale in the current Internet architecture. This architecture, however, will need to support such continuous point-to-point and point-to-multipoint connections simultaneously in the near future.

It is difficult to overestimate the dramatic impact 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 and multiple service levels, as well as accounting and billing. Although increasingly critical, these topics are beyond the scope of this paper.[6] We do note that prerequisite to accounting and billing instrumentation is a more accurate model for the attribution of resource consumption, including distinguishing between sender-initiated (e.g., e-mail) and receiver-initiated (e.g., gopher, mosaic, www, ftp) traffic for charging purposes. Such a model will also have to reliably attribute applications, or traffic profiles, to the clients if multiple levels of services exist.

VII. Assessment of international flows

National or international policy may necessitate attribution of resource consumption to individual countries. We can use currently collected data to explore shifts in traffic volume over time between the U.S. and specific countries via the NSFNET backbone. Figure 11 shows the bidirectional flow of traffic for the first week of February 1993 between the U.S. and three countries in different time zones (Japan, Mexico, and Great Britain). The impact of the time zones is visible in relationship to traffic volume; flow peaks coincide with the business hours of the country, and are more prominent if the business hours of the two countries overlap.

Figure 12 is an NSFNET backbone-centric illustration of countries using the U.S. for their own domestic communications, both in terms of absolute volume as well as in relation to the overall traffic those countries exchanged with the NSFNET. This effect typically derives from multiple connections between some country and the U.S. and is, at times, addressed case-by-case by the constituents of the connections.

Attribution of international traffic flows is becoming an important issue for splitting the costs between the two end-point countries. Several recent international connection scenarios have required the reevaluation of the current interconnection model. All international networking resources contribute to the quality of the global Internet, including the emergence of major international database servers. Therefore, better instrumentation is necessary to assess the service quality and network impact of such resources.

VIII. Conclusions

We have presented the architecture for data collection for the NSFNET and the limitations of the approach being used for long-term network forecasting and planning. We have also discussed the IP address structure and its application to NSFNET growth, port numbers and their implementation limitations that prevent real tracking of service diversity, and traffic volume trends by application and country.

These statistics reflect operational collection of traffic and network registration data, both initially designed to support short term engineering and planning needs. Traditionally, statistics used in forecasting

[6] The authors discuss accounting and billing issues further in [2] [4].

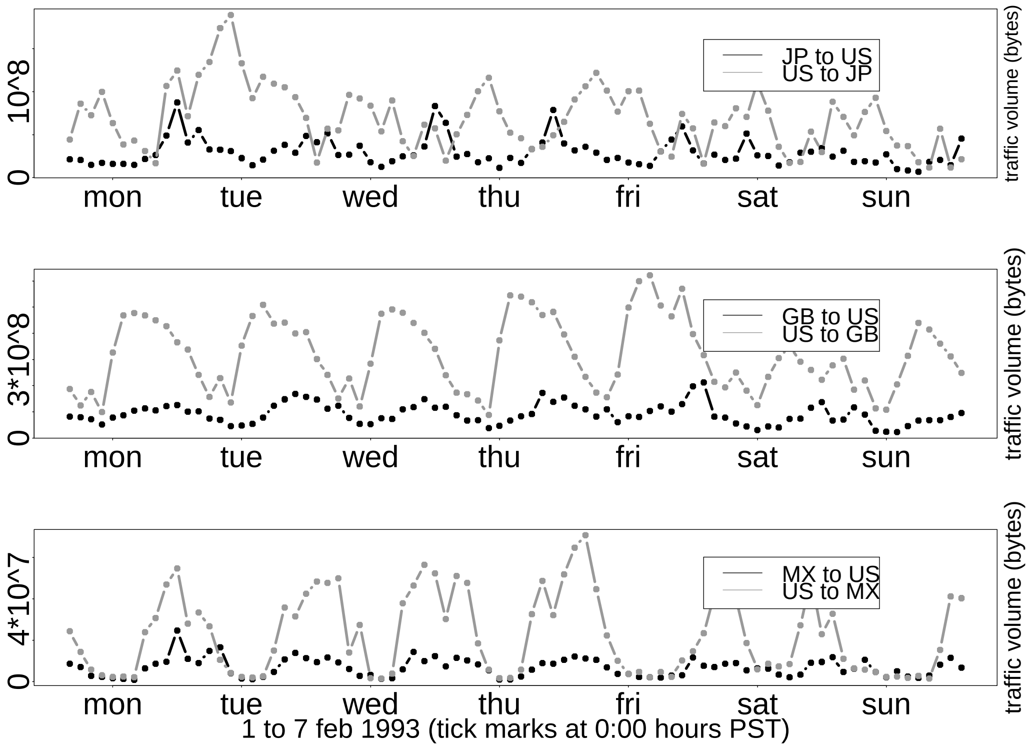

Figure 11: Traffic exchanged from Japan, Mexico, and Great Britain to United States

compounded traffic volume at network access points or individual network interfaces, which network planners extrapolate for indicators of future performance requirements. Although such statistics allow some tracking of Internet growth, they limit our ability to forecast capacity requirements in a network with ever richer functionality.

These statistics indicate significant growth of IP network number utilization, and therefore Internet clients, over the last several years. The trend is clearly continuing at a global scale; international clientele now account for over 40% of the IP network numbers known to the U.S. infrastructure. As the need to attribute network usage intensifies, e.g., for accounting and billing purposes, the currently available data sets will seem even more inadequate. Deployment of network number aggregation techniques (e.g., CIDR), which hide the interior structure of a network cluster, will further aggravate the situation.

We also investigated the growth in application or service diversity on the Internet as measured by TCP/UDP port numbers. The ever-increasing diversity in Internet application profiles, whose complexity will increase further with the newer continuous-flow multimedia applications, will require reassessment of network mechanisms such as queueing management in routers. Even within the traditional flow paradigm, subcategories of traffic such as interactive, transaction, or bulk traffic, may exhibit performance requirements which are different enough to justify priority queueing.

ANS has recently deployed software for the NSFNET service that will allow more flexibility with the port distribution assessments, though the inherent difficulty with the Internet model of application attribution remains. Furthermore, the recently established InterNIC activity may allow greater flexibility in maintaining accurate databases of network number and traffic type statistics. Concerted attention to such activities will help foster an Internet environment where network planning and traffic forecasting can rely on more than the traditional traffic counters used in the past.

IX. Acknowledgements

This research is supported by a grant from the National Science Foundation (NCR-9119473), and a joint study agreement with the International Business Machines (IBM), Inc.

We thank Merit Network, Inc., and Advanced Network Services (ANS) for their assistance and coopera-

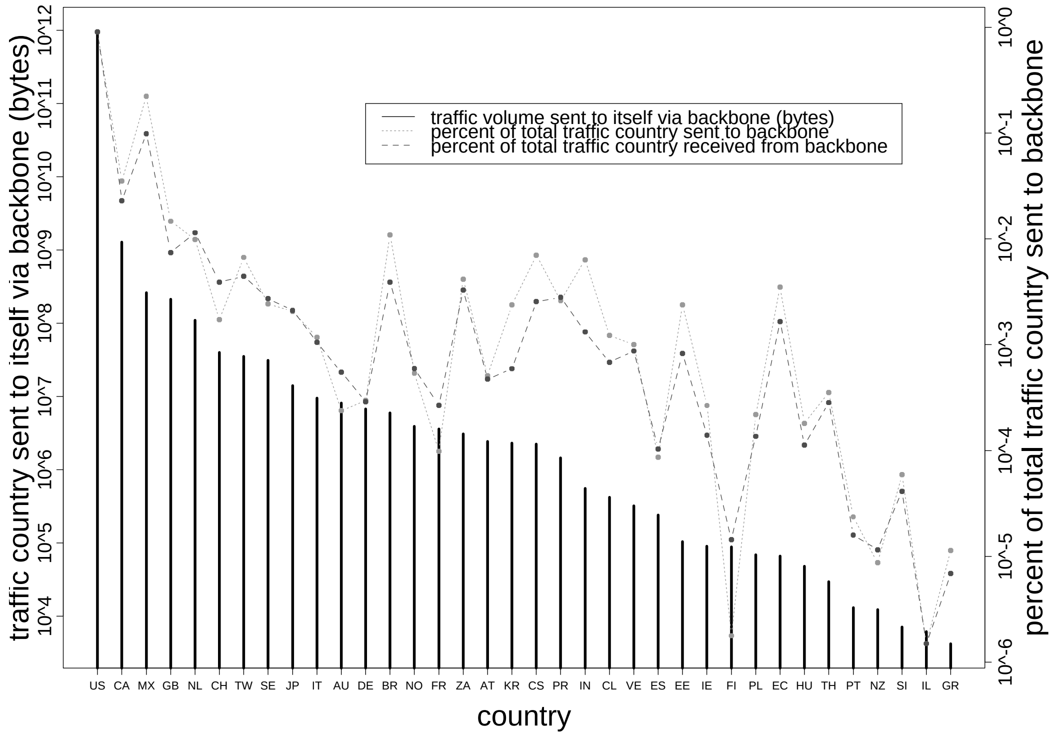

Figure 12: Traffic from countries to themselves through the NSFNET backbone

tion. We also thank David Bolen and Jordan Becker (ANS) and Mark Knopper and Susan Horvath (Merit) for many informative discussions, and Marsha Jovanovic for her editorial assistance.

X. Disclaimer

Any opinions, conclusions, or recommendations in this report are those of the authors and do not necessarily reflect the views of the National Science Foundation, other supporting organizations, General Atomics, the San Diego Supercomputer Center (SDSC), the University of California, San Diego, or the SDSC Consortium members.

References

[1] ANS. ARTS: ANSnet Router Traffic Statistics software, 1992.

[2] R. Bohn, H.-W. Braun, K. Claffy, and S. Wolff. Mitigating the coming Internet crunch: multiple service levels via precedence. submitted for publication, SDSC TR GA-A21530, November 1993.

[3] R.T. Braden and A. DeSchon. NNStat: Internet statistics collection package. Introduction and User Guide. Technical Report RR-88-206, ISI, USC, 1988. Available for a-ftp from `isi.edu`.

[4] H.-W. Braun, K. Claffy, and G. C. Polyzos. A framework for flow-based accounting on the Internet. In *Proc. Singapore International Conference on Networks (SICON'93)*, pages 847–851, September 1993.

[5] J.D. Case, M. Fedor, M.L. Schoffstall, and C. Davin. Simple Network Management Protocol (SNMP). Internet Request for Comments Series RFC 1157, May 1990.

[6] B. Chinoy and H.-W. Braun. The National Science Foundation Network. Technical Report GA-A21029, SDSC, 1992.

[7] K. Claffy, G. C. Polyzos, and H.-W. Braun. Application of sampling methodologies to network traffic characterization. In *Proc. ACM SIGCOMM '93*, pages 194–203, September 13-14 1993.

[8] K. Claffy, G. C. Polyzos, and H.-W. Braun. Traffic characteristics of the T1 NSFNET backbone. In *Proc. IEEE INFOCOM '93, San Francisco, CA*, pages 885–892, March 28 - April 1 1993.

[9] Peter Danzig, Katia Obraczka, and Shih-Hao Li. Internet resource discovery services. *IEEE Computer*, pages 8–22, September 1993.

[10] P. Ford, Y. Rekhter, and H.-W. Braun. Improving the Routing and Addressing of the Internet protocol. *IEEE Network*, May 1993.

[11] E. Gerich. Guidelines for management of IP address space. Obsoleted by RFC 1466, October 1992.

[12] E. Gerich. Guidelines for management of IP address space. Obsoletes RFC 1366, May 1993.

[13] J. Postel J. Reynolds. Assigned numbers. RFC 1340, July 1992.

[14] S. Kirkpatrick, M. Stahl, and M. Recher. Internet numbers. RFC1166, July 1990.

[15] K. McCloghrie and ed. M. T. Rose. Management Information Base for network management of TCP/IP-based internets, MIB-II. Internet Request for Comments Series RFC 1213, March 1991.

[16] K. McCloghrie and M. T. Rose. Management Information Base for network management of TCP/IP-based internets. Internet Request for Comments Series RFC 1156, May 1990.

[17] Network information services. Data available on `nis.nsf.net: /nsfnet/statistics`.

[18] J. B. Postel. Internet Control Message Protocol. RFC 792, September 1981.

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

XI. SIDEBAR – NSFNET: some history

NSFNET is a general purpose, packet-switching network that supports access to scientific computing resources, data, and interpersonal electronic communications (see [6] for a detailed description of the NSFNET). Evolved from a 56kbps network among NSF-funded supercomputer centers starting in mid-1986, today's T3 (45Mbps) network serves a much broader clientele encompassing not only the transcontinental backbone connecting the NSF-funded supercomputer centers and mid-level networks, but also mid-level networks themselves, and campus networks. The hierarchical Internet structure includes most of the research and educational community, and even extends into a global arena via international connections. Figure 2 shows the logical topology of the backbone.

Since July 1988, Merit Network, Inc., has administered and managed the T1 NSFNET backbone. For a description of the T1 NSFNET backbone and its instrumentation for data collection see [8]. In late 1990, in conjunction with its partners IBM and MCI, Merit began to deploy, in parallel to the T1 backbone, a replacement T3 network. The T3 network provides a 28-fold increase in raw capacity over the T1 network (from 1.544 Mb/sec to 44.736 Mb/sec). In November 1992 the T3 network had completely replaced the T1 network.

The status of the NSFNET has shifted through organizational restructuring among original participants in the backbone project. In 1991, Advanced Network Services (ANS) began official operation and management of their national T3 backbone infrastructure, on which they provide the NSFNET backbone service. Merit Network, Inc., administrator of the original NSFNET T1 backbone, still holds a cooperative agreement with NSF to provide backbone services.

XII. SIDEBAR – What's in an IP Address?

The IP address space architecture originated with RFC 791 [19], the initial Internet Protocol specification that defined a pool of available network numbers. Ignoring some special cases, such as multicast addresses, every network number on the Internet came from this pool of available network numbers. A large subset, although not every number in this pool, has been assigned to a requestor, typically on behalf of a company, university or other institutions, for active duty. The InterNIC Registrar, on behalf of the Internet community,[7] now formally registers these assigned network numbers in a database that also includes mappings to address information of the institution responsible for the network.

With the advent of RFC1366 [11] [12] in October 1992, the InterNIC began to assign addresses according to the geographic location of the requestor. The InterNIC also, in certain cases, delegates blocks of Class C IP network numbers to other authorities for further assignment. For example, the InterNIC assigned a large portion of the Class C space to Europe for further redistribution within their network community. From the InterNIC's point of view, these delegated numbers are no longer available but not yet formally assigned until the Europeans notify them that they have really assigned those numbers to their final IP networks.

Networks that are InterNIC-assigned do not by definition actively exchange traffic on the Internet. In fact, the set of communicating, or Internet active network numbers (see main body of paper for definitions), is not even necessarily a proper subset of the set of assigned network numbers (although in a frictionless world, it would be). Some organizations consider their local network environments wholly disconnected from the Internet, and with no plans for future connection, they sometimes even choose their own IP network numbers, independent of the InterNIC's registry, to satisfy their isolated TCP/IP protocol needs. Unfortunately, experience has shown that such disconnected environments often turn out to be quite leaky. When traffic from these networks manages to find its way into the Internet, often much to the surprise (or ignorance) of the local network administrators, these network numbers join the set of leaky unassigned numbers. Leaky unassigned numbers are members of the active set of numbers that are not in the assigned set.

[7] Prior to April 1993 the Defense Data Network's Network Information Center (DDN NIC) performed this registry function.