CERN - CN 93-11

CERN - Computing and Networks Division
CN/93/11
June 1993

ne 9327
c2

CERN LIBRARIES, GENEVA

CM-P00065739

IP Traffic Measurements and Analysis at CERN

Jean-Michel Jouanigot
Olivier H. Martin
Jessica Yu

(Presented at the INET'93 conference, San Francisco,
August 17-20, 1993)

IP Traffic Measurements and Analysis at CERN.

Jean-Michel Jouanigot[1]
Olivier H. Martin[2]
Jessica Yu[3]

Abstract

Because of the heavy concentration of lines into CERN a significant fraction of the total European IP traffic is crossing routers located on our site. CERN is thus ideally placed to collect information about the traffic flows, both at the macro and micro levels. In particular, an analysis of the trans-european and transatlantic (over the CERN/ANS line) flows has been made, and the results have been compared with those available from MERIT, the only other source of global Internet traffic statistics known to us.

I. INTRODUCTION

CERN, the European Laboratory for Particle Physics, is at the crossroads of several physical networks (e.g. EBONE, HEPnet) and is also actively involved in a number of logical networks (e.g. EARN, EUnet, LISTSERV Distribute Backbone, MBONE). Because of the nature of its activity CERN is also a major source of data, hence its early involvement in networking activities.

Since late 1989, when the external IP service was officially opened, a very rapid growth of the IP traffic has been observed. A significant milestone was reached shortly afterwards when a T1 line to the NSFnet backbone at Cornell University was installed in March 1990, as a result of IBM's European Academic Supercomputer Initiative (EASI). Since December 1992, the T1 line has been relocated to New-York (ANS) and it is planned to connect it to the "proto-GIX" at MAE-Eeast later this year.

CERN has approximately 10 Mb/s of external bandwidth dedicated to IP, which is distributed over 21 lines in 12 countries and 3 continents, and is, by far, the largest centre of academic and research networking in Europe. According to the latest measurememts, it is estimated that 33 % of the total EBONE traffic is traversing routers located at CERN.

[1] CERN, 1211 Geneva 23, Switzerland
[2] CERN, 1211 Geneva 23, Switzerland
[3] MERIT, Network Inc., 1071 Beal Ave. Ann-Arbor, MI 48109

CERN is therefore ideally placed to study the traffic flowing through the routers and the interconnection LAN, at present an Ethernet, to be evolved shortly into a "neutral" FDDI ring. Throughout this paper, we use the acronym CERN/IXP (CERN/Internet eXchange Point) to refer to this facility.

The data that formed the basis of the analysis for this paper are of two kinds:

- IP accounting data collected on cisco. It measures the total volume of the traffic in terms of packets/bytes that traverse the CERN/IXP, and it is used to analyze the geographic aspects of the traffic.
- TCPDUMP data collected using TCPDUMP[1] developed by Van Jacobson, et al. collected on a DEC Ultrix system.

The paper is therefore divided into two main parts:

- a macro analysis
- a micro analysis

The macro analysis uses the cisco accounting data collected at CERN, as well as the T3 backbone statistics provided by MERIT (ftp server: nis.nsf.net, files: /statistics/nsfnet/1993/nsf-9303.*) for cross-checking purposes, but also to add a world wide dimension not available in the CERN/IXP environment alone.

While the results of the macro analysis has been the subject of several presentations in various forums, the micro analysis, which uses the TCPDUMP data, was never made at CERN before, because of the lack of manpower and monitoring equipment. The micro analysis was conducted by Jessica Yu (MERIT), while she was a scientific associate at CERN, and its results form the bulk of this paper. Some results from the macro analysis are also provided for illustrative and cross referencing purposes.

II. Topology of the CERN/IXP

The traffic which could be captured and measured at the CERN/IXP falls into the following categories:

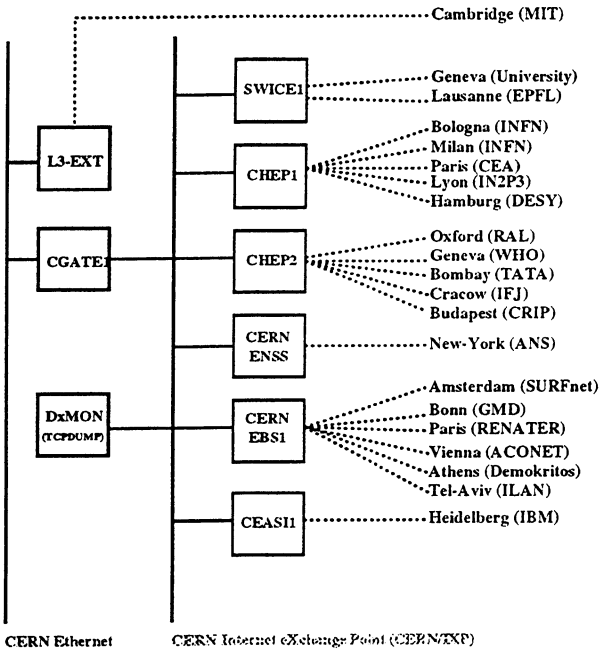

Figure 1: IP Routing environment at CERN

- Intra-European traffic
- Inter-Continental traffic (Europe - US & Europe - Other countries via the T1 line to the US)

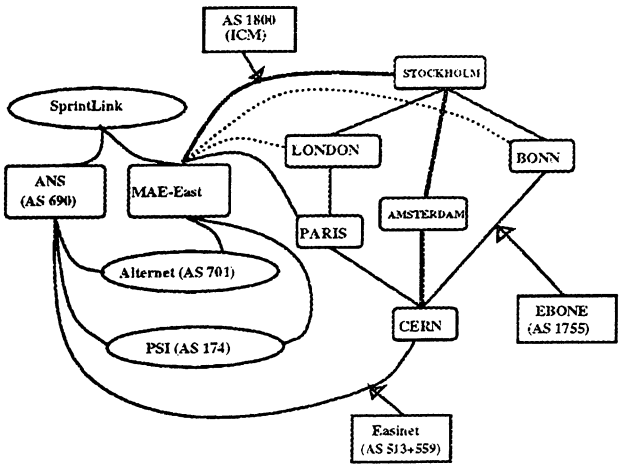

Figure 2: Ebone Backbone Map and interconnection with US backbones.

As shown in table 1 and figure 2, Easinet and ICM are the two main providers of transatlantic connectivity. The two dotted lines between Washington DC and Bonn & London are "unofficial" transatlantic Ebone lines, also managed by ICM (Sprint), but of lower capacity than the two T1 lines between Washington DC and Paris & Stockholm. The Easinet/T1 line terminates in New-York. The multi-homed AS connections on the US side (e.g. Alternet, PSI) and the lack

# nets with:	Easinet (AS 513 + AS 559)	ICM (AS 1800)
Primary access	710	959
Second. access	19	1106
Ter. access	586	91
# Eur. countries	19	12
# non-Eur.	1	11
Incoming (GBytes)	119	175
Outgoing (GBytes)	211	241
Total (GBytes)	330	416

Table 1: Some figures about Easinet vs ICM (March 1993)

of "route server" explain the asymmetric traffic spotted at the CERN/IXP.

Given the above topologies (i.e. CERN/IXP & Ebone), the Europe to US traffic referred to, throughout this paper, is due to the subset of the European networks that used the CERN/ANS transatlantic T1 link as their primary (or backup) access into the NSFnet backbone [1]. As this is a significant portion of the traffic, it can be assumed to reflect the main characteristics of the traffic between Europe and the US.

III. Traffic Collecting Mechanism and Methodology

III.A. cisco IP accounting

One major source of data regarding the IP traffic crossing the CERN/IXP, is cisco's IP accounting, which allows to collect packets and bytes counts on a host by host basis. There are drawbacks, however:

1. only switched traffic is accounted.
2. storage space is not adequate to cope with heavy traffic involving many source destination pairs.
3. high cisco CPU consumption.
4. duplication of records when multiple routers are crossed on the same site.
5. the amount of data to be stored and processed is huge at hub sites with daily traffic in excess of 20 Gigabytes/day.
6. as accounting is done prior to forwarding IP packets the accuracy is doubtful on lines

[1] the situation is unfortunately more complex as some US routes (e.g. Alternet & PSI) are announced on Ebone, whereas other destinations are reached by defaulting to the ANS/T3 backbone.

with high drop rates.

Over the years, substantial effort has been invested at CERN, in order to collect, process and analyze the cisco IP accounting records. In practice, we collect the IP accounting data from every cisco router under our control, namely: CGATE1, CHEP1, CHEP2 and CERN-EBS1, every 30 minutes. The only significant source/sink of traffic which we do not gather directly is the traffic between SWITCH (National Research Network in Switzerland) and the US, as well as the traffic between CERN and MIT (Cambridge) & Caltech (via the L3-EXT router) [2]. However, due corrections have been made to the figures provided in this paper in order to take account of ALL traffic crossing CERN.

Duplicate records are then removed and the data is compressed online, such that there is only one record left for every source/destination pair during the time interval. We also take this opportunity to record the internal paths, that is the routers crossed on site, in order to spot asymmetric traffic, which, for various reasons intrinsic to Europe, is still rather common.

More data reduction is then performed after midnight when the data is further reduced to one record per source/destination pair per day, and a daily report is produced showing the main sources and sinks of data as well as the asymmetric traffic.

For practical reasons (lack of disk space and cpu power) we are also grouping "small" records (i.e. less than 200 packets or 10 kbytes) and "overflow records" (i.e. the part of the traffic which cannot be collected because of cisco buffer overflow and for which only the number of bytes and packets is provided) into "dummy" source/destination records, namely: 0.0.0.0/255.255.255.255. This kludge allows us to process large data samples, e.g. one month, at reasonable cost and without losing accuracy as far as byte and packet counts are concerned, but at the expense of losing the real origin/destination of this type of the "noisy" traffic.

Whereas the online data collection and data reduction is done in PERL, the regular reports are produced with SAS (Stastical Analysis System) on the IBM/9000 VM system at CERN. Extensive use of, so called, SAS formats is made in order to map IP networks into:

- country
- network name
- autonomous system number (ASN)
- Ebone router, etc.

As the Internet is growing so fast, maintaining the above mappings would have been beyond our capabilities had MERIT not supplied 90% of the required information in the file /nsfnet/announced.networks on the nis.nsf.net FTP server.

III.B. TCPDUMP

Three (3) data samples were collected by TCPDUMP:

1. Feb. 2 17:23 - 18:23 (1 hour sample)
2. Apr. 2 16:00 - 22:00 (6 hours sample)
3. May. 10 14:00 - 17:30 (3.5 hours sample)

Due to the huge volume of data and the disk space limitations (1.5 GByte), the data samples cover a relatively short period of time.

We discovered that the number of lost packets in the first and second data samples was quite high, as the system collecting the data was saturated by the very high traffic rates (i.e. about 1,454 packet/s). We then used another more powerful system which allowed to reduce the loss rate to 2.6%. Therefore, the third data sample is the most accurate one, in terms of its completness, and our analysis is, hence, largely based on it.

The TCPDUMP data gives information for protocol analysis of the traffic.

With the combination of the two sets of independently collected statistics we were able to do the following analysis:

1. analysis based on volume of traffic in terms of Geographical locations
2. analysis based on protocol and applications

IV. Privacy Considerations

We are well aware of the concerns of privacy issues related to networking traffic capturing and analysis. The statistical information we collected does not include the data portion of any packet. In the cisco accounting data, we compressed the information into net to net instead of host to host matrix. In the TCPDUMP data, we do have host to host information but, in general, we did not look into such information except to deal with well-known FTP servers and traffic related to network management such as ICMP packets, etc.

[2] refer to figure 1 for more details about the IP routing environment at the CERN/IXP.

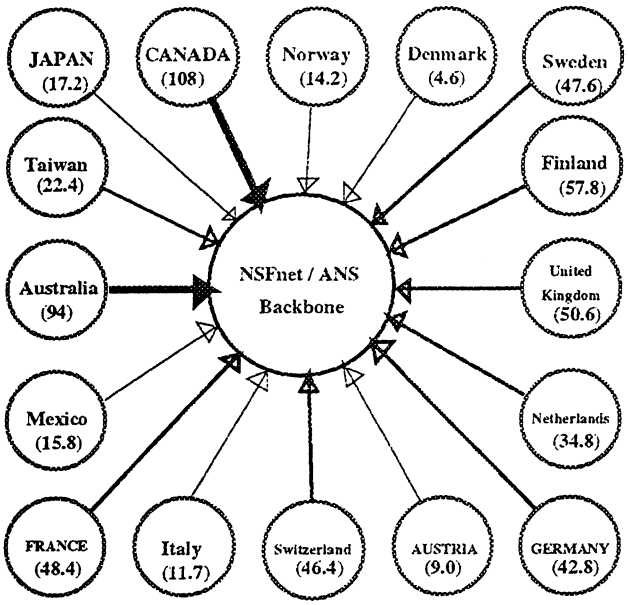

Figure 3: NSFnet backbone (incoming international traffic in Gigabytes (March 93).

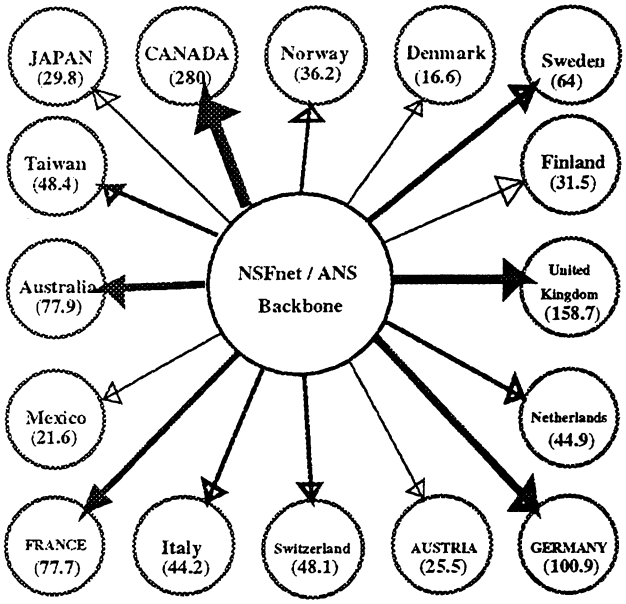

Figure 4: NSFnet backbone (outgoing international traffic in Gigabytes (March 93).

V. Analysis of Traffic based on Volume

Figure 3 through 6 are directly derived from the file /statistics/nsfnet/1993/nsf-9303.country on nis.nsf.net and are intended to help visualize the most active source & sink of international traffic in & out of the NSFnet backbone.

Incoming traffic from non-US sources was 10.27 % of the total traffic through the NSFnet backbone in march 93, namely 658 GBytes, whereas the outgoing traffic represented 19.66 % of it, namely 1.265 Terabytes. This simple result illustrates very well the imbalance between the incoming and the outgoing traffic, which is further analyzed in section VII.

Figure 5 shows the balance of traffic, namely the ratio between the incoming and the outgoing traffic. Only 2 countries, Australia and Finland, had a positive traffic balance with NSFnet in march 93. Some other countries (e.g. Mexico, Netherlands, Sweden and Switzerland (CERN included)) had a reasonably well balanced traffic, which presumably shows that they are either "net" exporters of data or that they are well organized, in terms of mirroring popular anonymous ftp servers located in the US or elsewhere.

Figure 6 is another way of attempting to measure the degree of organization of the various countries involved in the Global Internet by dividing the total traffic in & out of the NSFnet backbone by the number of networks announced (i.e. active) during march 93. Not surprisingly, this somewhat arbitrary indicator confirms the findings of figures 3 through 5. Hopefully, this paper will either be read by service providers from the countries "at fault", or better automated tools will be deployed in time to reduce the unnecessary and costly dependence on ftp servers exclusively located in the US.

Figure 7 & 8 show the incoming & outgoing traffic at the CERN/IXP.

Figure 9 show the incoming & outgoing traffic at the CERN/IXP grouped by "geographical regions" rather than by countries. A surprising result there, is the low level of traffic between Europe & the Pacific region as well as between Europe & Canada or South America. It is comforting to note that the intra-European traffic flows dominate, which has not always been the case in the past.

While carefully checking the TCPDUMP data against the cisco accounting data we discovered abnormally high traffic between a host in the US and a host in Austria. The traffic was heavy to the extent that essentially NO other traffic was able to make its way to Austria or countries beyond Austria anymore, and this lasted for more than 20 hours between April 1st & April 2nd! This "anomaly" illustrates the fragility of the IP layer when used by non-TCP based applications as well as the fragility of the IP service when paths with low capacity lines are used. It also proves that much stronger monitoring of the end to end service must be done by the various network providers in order to spot and stop such misbehaviors as early as possible.

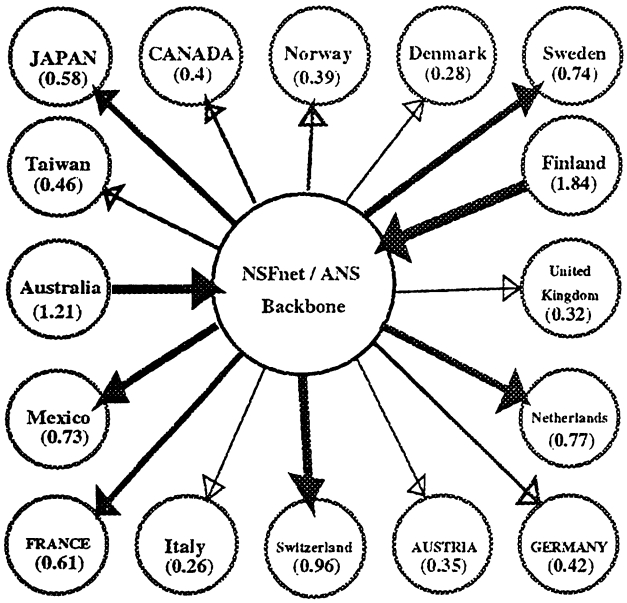

Figure 5: NSFnet backbone (international traffic balance (in/out)).

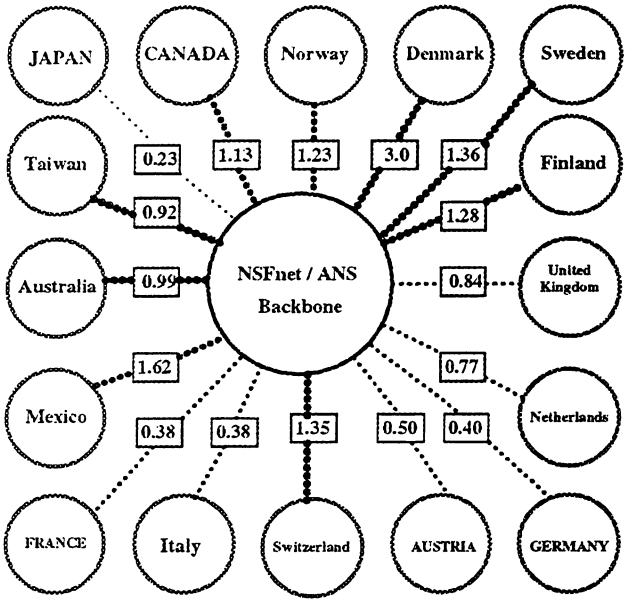

Figure 6: NSFnet backbone Total International traffic (in+out/number of announced networks).

In practice, the Internet traffic, as we can watch it at the CERN/IXP, still seems to be largely dominated by NSF AUP compliant networks. For example, for the whole of march 1993, where the record level of 550 GBytes through the CERN/IXP was observed (which is roughly 1/12th of the NSFnet backbone traffic during the same period) we had to build a data base of 12191 networks out of which 11149 were already configured in the MERIT data base, in order to map ALL the networks having generated "significant" amount of traffic through CERN, according to the "aggregation" technique explained above, into: Country, ASN, Ebone router, etc. It is worthwhile to note that out of these 1042 "unconnected to NSFnet" networks, 89 were illegal (i.e. not registered) network addresses, 35 belonged to IANA (i.e. reserved), 220 were in the US, often "connected" to Alternet, and the rest were in Europe connected to IP providers such as Pipex, Swipnet, various national instances of Eunet, etc, or elsewhere in the world.

VI. Protocol Analysis

The previous section discussed the volume of traffic traversing the CERN/IXP and the main traffic flows. This section will do the protocol analysis of the traffic based on the TCPDUMP data samples mentioned above.

Table 2 shows traffic break down into protocols in packets and percentage of total using the 3.5 hour data sample.

Comparing the traffic protocol distribution in

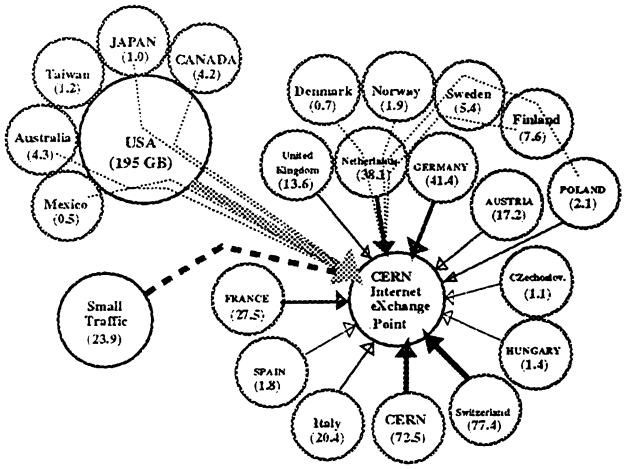

Figure 7: Incoming traffic at the CERN/IXP.

the March statistics of NSFNET [2], our 6 hour, 3.5 hour and 1 hour data samples, the top 10 service protocols are almost the same as shown in Table 2, 3 & 4. Since the data samples have durations of 1 month, 6 hours, 3.5 hours and 1 hour respectively, they provide sufficient generality to conclude that the listed protocols are the most popular amongst network users today[3].

VII. FTP Traffic Analysis

The File Transfer Protocol (FTP) traffic tops the traffic among all applications. About 40.23 % of all the traffic in terms of packets is FTP-DATA (tcp port 20) and about 4.13 % of the traffic is

[3] – indicates that the particular protocol is not within the top 10 of that specific data sample.

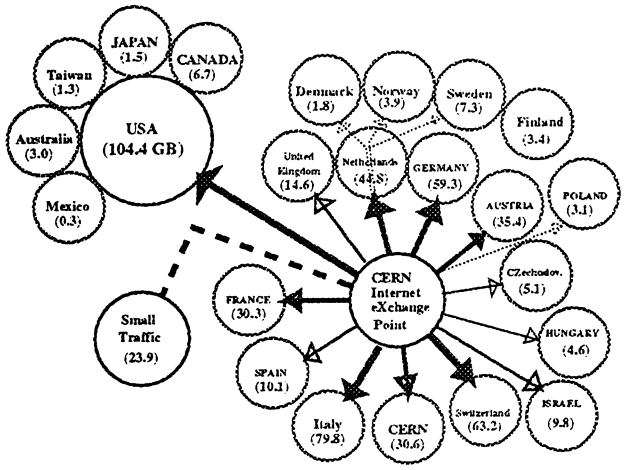

Figure 8: Outgoing traffic at the CERN/IXP.

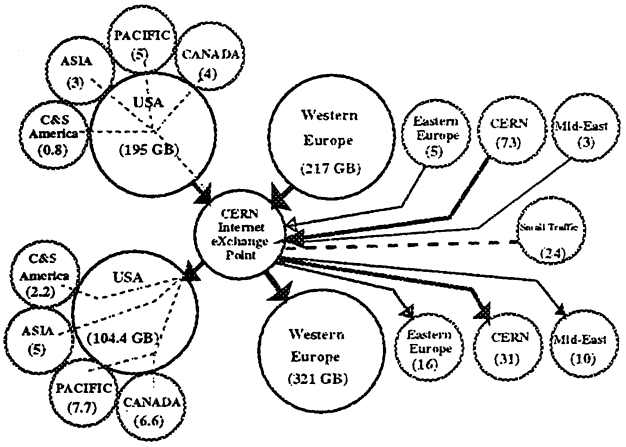

Figure 9: Traffic at the CERN/IXP classified by region.

Protocols	No. of Pkts	Percentage
ftp-data	7369398	40.23 %
telnet	3546596	19.36 %
ftp	757092	4.13 %
dns	737439	4.03 %
smtp	685284	3.74 %
irc	464072	2.53 %
x0	323573	1.77 %
nntp	318288	1.74 %
gopher	286213	1.56 %
icmp	243498	1.33 %
vmnet	197318	1.08 %
who	171113	0.93 %
p1023	101181	0.55 %
finger	74595	0.41 %
ntp	64794	0.35 %
www	56287	0.31 %
snmp	49099	0.27 %
p1022	33166	0.18 %
syslog	18066	0.10 %
bgp	17247	0.09 %
ntalk	8519	0.05 %
egp	6500	0.04 %
arp	3896	0.02 %
whois	3510	0.02 %
other	2783367	15.18 %
Total	18320111	100.00 %

Table 2: Protocol Distribution (3.5 hour sample).

FTP control packets (tcp port 21) in our 3.5 hour data sample.

The total data transmitted via FTP during this 3.5 hour period is 7,369,398 packets and 2,216,124,112 bytes. Discounting the 20 bytes of IP and 20 bytes of TCP header information, there are 1,921,340,000 bytes of actual data transmitted. Assuming that a standard page contains 80*64 bytes, it transmitted about 375,262 pages at the rate of about 30 pages per second on average during the sampling period.

The average number of bytes per ftp-data packet is 301 including the 40 bytes of IP and TCP header information in each packet.

The total traffic for FTP application during the 3.5 hour period is 8,126,490 packets or 2,233,205,523 bytes. Given the fact that each byte contains 8 bits, the average FTP traffic transfer rate during this period is 1.42 Mb/s which is almost equivalent to T1 (1.5 Mb/s) speed in one direction.

VII.A. Most active FTP Host Pairs

The top 20 pairs of hosts that exchanged the most information during that period via FTP are shown in Table 5. The Source host is the FTP server while the destination host is its client. Among them, only 5 pairs are located in Europe. One pair is between Asia and Europe and the rest is between US and Europe.

The first pair on the list has a total FTP traffic of 79,119,984 bytes which is equivalent to 50 kbit/s on average during this 3.5 hour time period. It shows how much of the network bandwidth can be taken by individuals.

The average transfer rate of the 14 pairs between US and Europe and the one pair between Asia and Europe which also used the transatlantic link is about 201 kb/s which is 13% of the total T1 bandwidth of the transatlantic link between US and CERN.

VII.B. Well-known FTP Servers

Counting the FTP servers by identifying the hosts which use the well-known ftp ports (i.e. 20 & 21) and which exported data, we obtained the

protocols	6 hr data %	rank	3.5 hr %	rank
ftp-data	31.35	1	40.23	1
telnet	14.07	2	19.36	2
ftp	3.14	6	4.13	3
dns	3.33	5	4.03	4
smtp	3.36	4	3.74	5
irc	2.84	8	2.53	6
x0	2.28	9	1.77	7
nntp	2.94	7	1.74	8
gopher	–		1.56	9
icmp	4.51	3	1.33	10
vmnet	1.31	10	–	
who	–		–	

Table 3: Top 10 Applications (Samples 2 & 3).

protocols	1 hr data %	rank	NSFnet %	rank
ftp-data	38.51	1	23.74	1
telnet	19.09	2	15.10	2
ftp	3.51	4	2.10	6
dns	3.67	3	4.97	5
smtp	3.42	5	7.01	4
irc	3.23	6	2.02	7
x0	1.57	8	–	
nntp	1.91	7	8.39	3
gopher	–		0.94	10
icmp	0.01	10	1.82	8
vmnet	–		1.31	9
who	1.26	9	–	

Table 4: Top 10 Applications (Data samples 1 & NSFnet March 93).

list of the top 20 FTP servers shown in table 6 & 7. The column '# Pkts' includes the FTP packets, both ftp-data and control, exchanged between the server listed and other hosts in our data sample. The column '# Bytes' is the same as the '# Pkts' column with byte counts instead of packet counts. In table 7, the column '#ftp'd' counts the total number of different hosts with at least one FTP session with the listed server.

It is interesting to note how popular some of the FTP servers are. For example, during the 3.5 hour period, 327 individual hosts actually have FTP sessions with wuarchive.wustl.edu and fetched file(s) from it. That means, on average, there are about 93 different hosts per hour that actually FTP files from this server from only the part of Europe whose traffic towards the US traverses the transatlantic line at CERN.

This figure also indicates that if one likes to provide such a file service, a very powerful system is needed in order to meet the demands of

Country of the Source Host	Country of the Dest. Host	#Pkts	#Bytes
US	CH	154594	79119984
US	NL	49144	25024315
US	FR	48884	24726644
US	CH	47816	23805027
US	IE	43752	22346174
CH	US	38108	20410272
CH	FR	35021	18748978
US	FR	36220	18453572
US	FR	34264	17349886
US	IT	29426	14871119
CH	CH	28738	14706860
US	CH	26174	13334432
UK	DE	23423	12527214
US	ES	23651	12083244
CH	DE	23209	11616346
TW	IT	22900	11603842
US	IT	21775	11147688
US	HU	21512	11012418
US	IT	21375	11012418
FR	IT	21072	10742609

Table 5: Most active FTP Hosts (Top 20 Pairs).

users. It also indicates the need for mirroring some of the popular file servers at every corner of the network, in order to reduce the traffic.

Out of the 20 most ftp'd servers by users whose traffic could be counted at the CERN Ethernet (European users but not all), 13 are in the US, 1 is in Australia and 6 are in Europe. This shows that the US is still the main data source for network users around the world, as many similar studies have already pointed out.

The total FTP traffic between the listed top 20 FTP servers and other hosts who fetched data from them are 1,002,857,115 bytes or 44.91% of the total FTP traffic in our logged TCPDUMP data sample.

Table 8 shows that among the 1,002,857,115 bytes exchanged between the top 20 FTP servers and other hosts, 64.18% are between the FTP servers located in the US and the client located in Europe with the US servers exporting data. 9.48% are between Europe and US with the European FTP servers as the data source. 2.49% of that are between a server in Australia and its FTP clients in Europe. Since the traffic between Australia and Europe traverses the US network infrastructure, it also passes through the transatlantic line at CERN. Therefore, 76.15% of the total FTP traffic exchanged between the top 20 FTP servers and their clients is between US and

Server Name	#Pkts	#Bytes
wuarchive.wustl.edu	683973	188349810
nic.switch.ch	473538	127065545
doppler.ncsc.org	265122	79378924
donau.et.tudelft.nl	260533	78617924
cs.uwp.edu	237590	55320762
nic.funet.fi	180184	48995397
ftp.cica.indiana.edu	176920	50142899
sumex-aim.Stanford.edu	173582	51286959
amiga.physik.unizh.ch	130685	34713012
SunSite.unc.edu	122061	32750835
aeneas.mit.edu	119028	35430372
tucana.tuc.noao.edu	114794	34747543
rigel.acs.oakland.edu	102105	26720633
fiji.wri.com	100345	25701160
plaza.aarnet.edu.au	95819	24931019
export.lcs.mit.edu	93603	19079353
hobbes.nmsu.edu	89005	24427334
garbo.uwasa.fi	85557	23051950
mcafee.com	80831	20315581
asis01.cern.ch	78118	21830103

Table 6: Top 20 FTP servers (ordered by traffic).

Server Name	# ftp'd
wuarchive.wustl.edu	327
nic.switch.ch	219
amiga.physik.unizh.ch	164
cs.uwp.edu	129
nic.funet.fi	127
ftp.cica.indiana.edu	101
sumex-aim.Stanford.edu	100
donau.et.tudelft.nl	70
mcafee.com	57
plaza.aarnet.edu.au	55
SunSite.unc.edu	50
rigel.acs.oakland.edu	49
hobbes.nmsu.edu	41
aeneas.mit.edu	38
export.lcs.mit.edu	38
tucana.tuc.noao.edu	38
garbo.uwasa.fi	34
asis01.cern.ch	30
doppler.ncsc.org	5
fiji.wri.com	5

Table 7: Top 20 FTP servers (ordered by # of requests).

Europe, whereas 23.85% of the traffic is intra-European.

This phenomenon suggests that if the contents of all the popular US FTP archive servers are duplicated in one or more, well managed and well publicized, servers located in Europe, the traffic on scarce resources such as transatlantic links would be greatly reduced.

VII.C. Mirroring Well-Known FTP Servers

According to our data sample, mirroring the contents of the following popular FTP servers needs to be specially considered in order to reduce the traffic traversing the nearly congested transatlantic link at CERN, and most likely other similar links, as well.

The well-known FTP server at wuarchive.wustl.edu exported 188,349,810 bytes during this 3.5 hour period to hosts in Europe through the US connection via CERN. This is about 120 kbit/s. The volume and the popularity (measured by the number of different hosts ftp'ing from it) strongly suggests that it is worthwhile to copy the contents of this server to one or more servers in Europe.

The server plaza.aarnet.EDU.AU in Australia, according to its login banner, is primarily for users within Australia. It mirrors file servers around the world in order to reduce the load on the Australia to US link. Presumably, the information on the system could be fetched from

Pair in Country	# MBytes	%	Source
US & Europe	643652	64.18	US
Australia & Europe	24931	2.49	AU
Europe & US	95086	9.48	Eur.
Europe & Europe	23918	23.85	Eur.

Table 8: Top 20 FTP data exchange between Europe and other Countries.

servers outside of Australia. However, there are still 55 individual hosts in Europe which ftp'ed 24,895,948 bytes worth of information from this server during the 3.5 hour period. The traffic consumed was 15 kbit/s of bandwidth and had to traverse the transatlantic link as well as the congested US to Australia link, unfortunately.

The server amiga.physik.unizh.ch in Switzerland exported 18,094,488 bytes to US which is 48.4% of the total bytes it exported during the period. It suggests that the contents of this server should be duplicated at servers in the US. This appears to be happening, in our data sample, it shows that wuarchive.wustl.edu imported 3,093,583 bytes worth of data from this particular server, in order to, presumably, mirror some of the files from the amiga server.

The server donau.et.tudelft.nl is another one that needs to be considered. It exchanged

78,617,924 with its FTP clients as the data source. 99.90% of the data it exported goes to non-European countries with 87.67% to US, 3.75% to Canada and 8.48% to countries in the Pacific region. This amount of traffic all traverses the transatlantic link at CERN.

Of course, duplicating files also consumes bandwidth, but this could be done during off-peak hours when the usage of the link is low such as, night hours or weekends. Merit's statistics of the transatlantic link utilization [3] shows that on weekends the utilization of the link is significantly lower (about 50% lower) than on week days.

Mirroring popular servers at many network locations to economize the bandwidth has been suggested in other papers describing related studies [4]. We believe that there are some efforts to implement this. In our short period data sample, we already saw instances where a popular Europe FTP server was being mirrored by another well known FTP server in the US, as mentioned above.

However, our data sample still shows that a great deal of the FTP traffic travels over very long distances and, in so doing, consumes scarce network resources in order to fetch data from the various file servers around the world in a "semi-random" manner. This behaviour could be due to the following reasons:

- Not many "good quality" mirrored FTP servers available, at this stage.
- "mirroring" is neither organized nor coordinated.
- Users are not aware that "mirrored FTP servers" exists in servers closer to them.

We suggest the formation of a body, for example, an IETF Working Group, to coordinate such effort and to deploy other suggested techniques such as file caching [5].

We also suggest the development of mechanisms to encourage users to use the FTP server which is closest to their network connection or through a path which is not bandwidth critical. One way to implement this is through advertisements, such as placing banners on those popular FTP servers and to redirect users to servers nearer to them. In addition, mechanisms such as checking the source address of each FTP request and refusing access if a better located and properly managed server exists, or automatically redirecting the FTP requests to such servers could be used.

The study of the top 20 most ftp'd servers in our short period data sample suggests that 76.15% of transatlantic FTP traffic would be eliminated if mirror images of well-known file servers were located in proper places AND utilized by users. Given the fact that FTP traffic between the top 20 servers and other hosts occupies about 44.91% of the total FTP traffic and that the FTP traffic is always the top contributor to the total volume of the traffic (see table 2), the potential savings in bandwidth are probably very significant.

VIII. Telnet Traffic

Over the 3.5 hour data collecting period, 4224 host pairs established or attempted to establish telnet sessions. This resulted into a total of 3,546,596 packets. On average, this is about 839 packets exchanged between each telnet user/host pair.

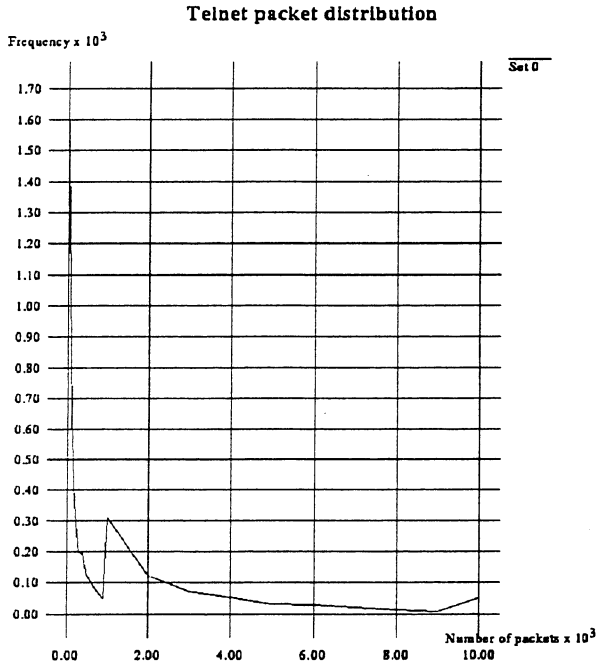

Table 9: Telnet Traffic (packet distribution).

Table 9 shows that about 1705 telnet pairs exchanged less than 100 packets. Discounting the ACK packets required by TCP connections which telnet connection is based on, those pairs only exchanged less than 50 packets during their telnet connections. That is actually 40.36% of the total 4224 telnet pairs logged in our TCPDUMP data sample. Among them, about 500 telnet pairs exchanged less than 10 packets which could be considered as non-successful telnet connection. It is hard to explain why so many telnet pairs had so little activity during the course of the telnet connection. A possible explanation may be that

users keep their telnet sessions open in order to avoid re-logging into the remote system.

IX. Routing Protocols Overhead

As shown in figure 1, there are altogether seven (7) IP routers on this ethernet; among them, Inter-domain protocols BGP and EGP are used for exchanging routing information among them. No IGP protocol is implemented.

In total, there are 9 BGP sessions including 6 EBGP sessions and 3 IBGP sessions. There are also 11 EGP sessions.

The routing information exchanged among the 11 EGP sessions are 9349 in total and among BGP sessions are 37016. Note that each individual route may be counted multiple times in this figure because they may be advertised once from a router to its M peers resulting into M occurrence of advertisements of this particular route. Since this study is concerned with the routing traffic introduced by routers exchanging routing information, the multiple occurrence of a single route should be counted.

The routing packets (bgp & egp) in our 3.5 hour data sample are shown in table 2 & 10. One could observe that the routing traffic in this routing environment comprises 0.13% of the total traffic in terms of packet count. It is also 1.89 packet/s and 204.90 bytes/s on average, as indicated in table 10. Given the amount of routing information exchanged among the routers mentioned above, the overhead of routing traffic is insignificant.

Protocol	Pkts	Pkts/s	Bytes	Byte/s
bgp pkts	17247	1.37	442582	35.13
egp pkts	6500	0.52	2139133	169.77
Total	23747	1.89	2581715	204.90

Table 10: Routing Traffic and its Rate.

We made an attempt to compare BGP and EGP in terms of amount of routing traffic each generates, when exchanging the same amount of routing information. It is known that BGP does routing loop suppression which frees the network from the topology restrictions imposed by EGP. BGP also facilitates policy routing better than EGP with the AS path information it provides. Our study only tries to compare routing traffic generated by each protocol.

We used a test router (a cisco) to first BGP with ENSS 206 (the NSFnet Backbone Router at CERN) and then EGP, in order to receive and re-advertise the full NSFnet routes (i.e. 8000 routes approximately). We logged the BGP and EGP traffic using TCPDUMP in 9 hours period each. Table 11 shows the data of the EGP and BGP traffic logged by the TCPDUMP.

	EGP	BGP	ratio
time period (hour)	9	9	
total packets	4224	24742	1:6
Packet rate (p/s)	0.13	0.76	
total bytes	3873573	1286026	3:1
Byte rate (B/s)	119.55	39.69	
Av. Packet size	917	52	

Table 11: Comparison of BGP and EGP at the CERN/IXP.

One could observe from table 11 that in terms of packet counts, BGP uses 6 times that of EGP. While, by comparing the byte counts, BGP uses a third of that of EGP.

In terms of bandwidth, the data shows that BGP uses a third of what EGP would consume in a steady state situation. By steady state, we mean that the EGP peer sessions have already been established (i.e. no initialization).

BGP's property of incremental update (i.e. it only sends changed portion of routing information), explains why there are so few bytes exchanged. EGP requires periodical updating routing information with everything it knows, which would result in the consumption of more bandwidth.

We can also explain why BGP has higher packet counts than EGP does. BGP is based on TCP and so acknowledges every packet it receives, so the packet count doubles. While EGP is UDP based, no ACK is required by its peer. Also, a BGP speaker sends updates to its peer as soon as the network topology changes while EGP has a "damping" period of 3 minutes since it sends updates every 3 minutes. Unstable network would result in BGP sending more updates while EGP "holds down" the unstable routes for 3 minutes. Because of that, BGP has a faster convergency time than EGP. It is a trade off between having faster convergence time and less packet counts. There has been discussions about whether packets generated by BGP could be reduced by "damping" the BGP response to network topology changes. Our study shows that it may not be necessary given that the absolute rate of packets generated by BGP is 0.76 packet/s in one data sample and 1.37 packet/s in another. The packet rate is not high enough to cause concern, especially if its reduction is done at the expense of slowing down the convergency time.

However, our data sample does not include

the BGP initialization state which would cause more traffic than the steady state, since, during its initialization stage, a BGP speaker has to advertise all the routes it knows to all its neighbors. Also, during the time period where the data was collected, the network could be stable, which would also reduce the BGP related traffic. Therefore, it may be better that routers provide the flexibility to set the BGP "timer" and let the network provider adjust the BGP "damping" time based on the stability and/or the load of the particular networking environment.

X. ICMP Traffic

In the 3.5 hour data sample, ICMP packets accounted for 243498 or 1.33% of the total traffic logged. The breakdown of the ICMP packet type is shown in Table 12. In this section, we will discuss some of the findings by analyzing ICMP "unreachable" and ICMP "redirect" packets.

ICMP Type	Pkt Count	Percentage
unreachable	101541	41.70 %
echo request	52352	21.50 %
echo reply	40207	16.51 %
source quench	22318	9.17 %
address mask	17142	7.04 %
redirect	127	0.05 %
other	9811	4.03 %

Table 12: Break Down of ICMP Packets.

X.A. ICMP Unreachable Packets

It is worth noting that ICMP destination unreachable packets account for 101,541 in our data sample. At first sight, it may seem that, on average, in every one minute there were 483.5 cases of denial of service to the network users, which is quite high. However, when examined in detail, we found that among them there were 30,920 incidents of "port unreachable" which should not be attributed to network reachability problem. Also 31,252 out of 70,621 host unreachable incidents were generated because the target host was on an "unassigned" network number or on a network not connected to the general Internet. In other words, half of the host unreachable incidents were false events. Despite that, there were still 39,369 "host unreachable" incidents during the 3.5 hour time period which is a rate of 187, denial of service to users, per minute.

Our 3.5 hour of logged icmp unreachable packet information reveals that there were 57 unassigned network numbers that users tried to reach, 23 of them are unassigned or IANA reserved Class A network numbers. Some of the non-valid nets were "retired" nets, yet there were still people trying to reach hosts on it (e.g., network 10).

There were 165 networks which users tried to reach during the period of time, with valid network numbers but not connected to the general Internet; for example, they are not known to the NSFnet and other default free network routing tables, due to AUP constraint, intended to be only on private network, or for other reasons.

Since the ICMP "host unreachable" message is a response to a packet with destination address to the unreachable hosts, there must be users who did try to reach hosts on the 57 invalid network addresses mentioned above. This leads to the conclusion that those networks are being used and used without authorization. Our short period data sample reveals that such use is not rare. It suggests that action needs to be taken by an appropriate Internet body to stop such misusage of unassigned network addresses.

X.B. Redirect Packets

A router sends an ICMP "redirect" message to a host only when the host appears on the same network as the next hop router towards the destination the host sends a datagram to. The purpose of the redirect message is to enable the host to take a shorter path towards a destination by avoiding extra router hop.

The set up of routers and hosts on the CERN Ethernet should not result into any ICMP redirect packets. However, in our 6 hour data sample, 16,822 ICMP redirect packets were seen. That prompted us to look into further details.

We found that all the ICMP redirect packets we logged were not issued as described above. The hosts to which the redirect message were sent to, were not on the same network, with the suggested short-cut router, but were actually 'far away' across the transatlantic link. As a result, we saw many redirect packets across the transatlantic link and Ebone. Since these redirect packets are meaningless and that the source host cannot make any use of it, this is definitely a waste of bandwidth. With the heavy load on the transatlantic links as well as on the intra-European links, it is desirable to minimize such senseless packets.

In our 3.5 hour data sample, we again observed such meaningless ICMP redirect packets. Although the amount was small, they were sent by a completely different set of malfunctioning systems. That shows, if need be, that there are still many "misbehaving" systems on the Internet.

We informed the appropriate people about the improper behaviour of their systems with the

hope that it may be corrected.

XI. Conclusion

In this paper we have presented some results of the IP traffic studies carried out at CERN. The analysis was done at two levels, a macroscopic study of the total traffic during March 1993 and a microscopic study of the traffic crossing the CERN Ethernet Exchange during selected short periods in February, April and May 1993. The macro analysis, combined with the NSFnet backbone statistics from MERIT, allowed us to record the traffic flows of an important subset of the research community in Europe at a significant point in the evolution of the "Global Internet", namely some months after Ebone went into full operation. It also allowed us to spot some "anomalies" like the traffic imbalance of some countries which then led us into an in depth analysis of the FTP traffic.

A real case showing the fragility of the IP network layer, with respect to non-TCP applications, was identified, and its practical consequences, on the traffic to/from an entire Ebone country, described. We believe that the European Internet is extremely vulnerable to incidents of this kind because, due to the high PTT costs, low or medium speed lines are still predominant; this observation is particulariy true for Eastern Europe. Therefore, the need for a stronger "Global Internet Monitoring Team" has been stressed. This team could also tackle the problem of "illegal" use of network numbers and various other misbehaviours reported in this paper (e.g. ICMP redirects).

The micro analysis of the FTP traffic advocated the formation of an operational body to coordinate the "mirroring" of the most popular FTP archives worldwide and/or to deploy more advanced techniques such as file caching [5].

The micro analysis of the BGP traffic advocated the provision of a BGP "damping timer update" in the routers, in order to be better protected against "route flapping" in some parts of the Internet.

Although no major discovery was made, but we were not expecting any, we believe that the studies undertaken have been worth the effort invested into it. Indeed, we plan to continue our active involvement in local, as well as coordinated "Global Internet" traffic measurements, in order to grasp a better understanding of the evolution of the traffic and of its type (e.g. impact of multimedia applications), but also to develop and deploy tools to uncover faulty systems automatically.

XII. Acknowledgements

We are extremely grateful to Denise Heagerty and Brian Carpenter, from CERN, for their careful review of this paper and their numerous suggestions, and to Laurent Joncheray for his work on the collection of the cisco accounting data and the implementation of the "hooks" for studying the asymmetric traffic, while he was making his military service at CERN in 1992 as a "French cooperant".

XIII. References

[1] V. Jacobson, et al, "TCPDUMP (1)", Unix Manual Page

[2] NSFNET monthly statistics by Merit Network Inc. FTPable via nis.nsf.net

[3] NSFNET Link utilization statistics collected by Merit

[4] T. Asaba, K. Claffy, O.Nakamura and J. Murai "An analysis of international academic research network traffic between Japan and other nations", Proceeding of INET'92, Jun, 1992

[5] R. Danzig, R. Hall and M. Schwartz "A Case for Caching File Objects Inside Internetworks", Mar. 1993

Author Information

Jean-Michel Jouanigot graduated from Ecole Superieure d'Electricite (SUPELEC) in 1988 and joined CERN, the European Laboratory for Particle Physics, the same year. He is a member of the communications group and has been responsible for all aspects of the external IP service since 1990.

Olivier H. Martin graduated from Ecole Superieure d'Electricite (SUPELEC) in 1964 and joined CERN, the European Laboratory for Particle Physics, in 1971, where he became leader of the CDC 7600 and IBM software support teams. He is a member of the communications group since 1983 and has been in charge of the external networking team since 1989.

Jessica Yu received a Computer Engineering degree from Michigan-Ann Arbor University in 1987. She joined MERIT in 1987 and has been a member of the Internet Engineering team, where she has been closely involved with the routing architecture of the NSFnet. She came to CERN for a 6 month sabbatical as a scientific associate in December 1992.