CERN-Data Handling Division
DD/80/5
J.M. Gerard
February 1980

[redacted]

THE EFFECT OF THE CERNET NETWORK ON COMPUTING AT CERN

(presented at the SEAS Winter Project Meeting, Antwerp,
14-18 January 1980)

DD-jt

THE EFFECT OF THE CERNET NETWORK ON COMPUTING AT CERN

J. M. Gerard

European Organisation for Nuclear Research,
1211 Geneva 23
Switzerland

January 14, 1980

ABSTRACT

A high speed packet network has been in operation at CERN for two years, connecting experimental mini-computers and large mainframes over an extensive site.

The use of the network has increased extremely rapidly, both in terms of amount of traffic and variety of applications. It is now true that the usage for which the network was designed forms only a small percentage of the total load.

This paper examineS the effect which the network has had on the functional distribution of computer centre resources. Some personal views are also given on why the project has been a success and what further developments might affect the user community.

INTRODUCTION

CERN (the European Organisation for Nuclear Research), situated on the French-Swiss border covers an extensive site and provides advanced high energy physics facilities to teams of scientist from European institutes.

Computing services are part of these facilities and more than 200 computers, either mini-computers or large mainframes, are spread over the site.

In order to provide mini-computers with access to facilities available on mainframes, and more generally to offer a communication means between any pair of connected computers, CERN decided in 1975 to design and develop a general purpose packet switching network: CERNET, which has been in operation since 1978.

CERNET is a high speed data network consisting of

- a communication sub-network,
- a set of communication services available to the end users.

<u>The communication sub-network</u> is mesh structured with ample redundancy and provides high speed inter-node communications (2.5 Megabits/second) through CERN-designed asynchronous data link interfaces over distances of, at present, up to 5 kilometres.

Data transfer is under the control of a line protocol which handles packets of up to 2046 bytes and offers flow control, detection and recovery of transmission errors.

<u>The communication services</u> are supplied by an end-to-end protocol controlling communications between end user processes. They provide communication management facilities for multi-message exchanges with message assembly/disassembly, flow control and error recovery. The end-to-end protocol is available within user computers in the form of a portable Transport Manager.

While the use and size of CERNET is still increasing, the sub-network is currently composed of 12 nodes connecting more than 40 user machines.

HISTORY

The developments which led to the decision, taken in 1975, to create a packet-switching network are not particularly relevant to this paper. Reference [1] gives some information on these developments, as well as an overview of CERNET.

The essential point of departure for this paper is the project definition for a network with the specific goal of allowing minicomputers to access the computing facilities of the computer centre, and with the general goal of providing a general-purpose communications network. For the specific goal the stated intentions were limited to the sending of samples of experimental data for immediate analysis and the return to the minicomputer of the results. Centralised data recording was excluded in favour of the traditional method of local recording of data by the minicomputers on magnetic tapes.

The first use of CERNET was in March 1978, when one experimental group with a multiple PDP11 configuration was given access to the computer centre IBM 370/168. Since then CERNET has grown to include connections to all of the computer centre mainframes (CDC 6400,6500,7600 and IBM 3032,370/168) and over 40 smaller computers, mostly minicomputers in use by experimental groups. The range of services which has been made possible via CERNET has also grown tremendously as both users and computer centre personnel have come to realise the extent to which a general communications network can be used.

TECHNICAL DESCRIPTION

In order to understand the possible uses of CERNET it is necessary to give a brief technical description of its design and implementation.

CERNET is designed to allow communication between a pair of programs, each running in a computer connected to CERNET. It is not intended for access from terminals to computers directly, although in some specific cases such access can be obtained.

In essence, the basic communications network of CERNET is simply a datagram service attempting to deliver packets of data according to the destination specified by the sender. There is a recommended transmission protocol, currently used by virtually everybody and obligatory for communications with the computer centre mainframes, which converts the datagram service into a virtual call service. In addition, on the computer centre mainframes there is a permanently running program called the File Manager, to which any CERNET-connected user can connect for the purpose of accessing files in that mainframe.

For the user wishing to write a program for his own computer, which is going to communicate with a partner program in another computer, there is a subroutine package provided. By means of suitable calls the user's program can create a logical link to the partner program and exchange messages. In the case of communication with a File Manager there is a higher-level subroutine package capable of dealing with the higher level of dialogue needed for File Manager access.

GENERAL TYPES OF USAGE

There is basically only a limited number of types of use of CERNET. The simplest type is the use of a standard program, executing in either a minicomputer or the IBM, to transfer a complete file to or from a host computer. Here the term 'host computer' means a computer connected to CERNET and having a running File Manager program - currently only the IBM and CDC qualify.

The next level of complexity concerns the user who wishes to insert into his own program the code necessary to communicate with the File Manager in a host computer. Such a requirement is common in minicomputers doing data acquisition and wishing to send a percentage of the data for subsequent analysis. However, this is by no means the only use; a minicomputer may also access Job/Print files via the File Managers.

Most complex of all is program to program communication between two partner programs. This allows a minicomputer immediate access to the power of a large computer, or allows two (or more) minicomputers to collaborate in a common problem.

The main uses of CERNET belong essentially to the first two types, namely those involving a File Manager. In the course of time the capabilities of the File Manager have steadily increased according to the demand. Thus, now, the File Managers can deal with both permanent files and Job/Print files, as well as being capable of giving information on the existence, or otherwise, of files according to certain criteria.

SPECIFIC CURRENT CERNET USAGE

In the following sections and subsections are described specific uses of CERNET. Any statistics given, such as the amount of data transfer or number of logical links, is based on a survey of traffic to or from the main IBM during the last full week of 1979. The total traffic through CERNET would of course include connections between minicomputers and the CDC mainframes. However this extra traffic would not change the statistics very much, since the main IBM is involved in most of the CERNET traffic.

The analysis of the total usage involved a breakdown of the total traffic into various categories. These categories are described in the following sections in decreasing order of traffic quantity. In addition, the first three categories are sufficiently large as to be worth splitting into sub-categories, which are in turn described in subsections of the main sections.

In the various tables that accompany the text, the percentage figures given always refer to percentage of the total week's traffic. Table 1 gives the overall usage according to category. In it are shown the number of logical links established, plus the number of messages into and out of the IBM. It should be remembered that a single logical link may in practice be used for anything from the transfer of a single message to the transfer of several complete files. A message can also vary in size from a few bytes to several thousand bytes. Thus the real amount of data transferred cannot be calculated without knowing the characteristics of the corresponding type of usage.

Table 1: 'General types of use of the IBM'

Type of use	Logical links		Messages into IBM		Messages out of IBM	
	No.	%	No.	%	No.	%
1. File Manager	5663	59.6	88798	66.2	75482	47.8
2. To/from CDC	2371	25.0	30578	22.8	62828	39.8
3. Echo tests	1304	13.7	13317	9.9	13312	8.4
4. Wylbur Bridge	116	1.2	1184	0.9	6026	3.8
5. On-line analysis	10	0.1	10	0.0	0	0.0
6. Miscellaneous	36	0.4	155	0.1	162	0.1

IBM File Manager usage

In terms of either logical links or messages, over half of the total traffic into or out of the IBM occurs on logical links created between a minicomputer and the IBM File Manager. In terms of quantity of data this traffic is only about half, since almost all of the messages are short and fit into a single packet. As a comparison, at the end of 1978 this traffic was only about one third of the current amount; at that time it represented about 75% of the IBM usage.

Table 2: 'Links to IBM File Manager'

	Type of use	Logical links No.	%	Messages into IBM No.	%	Messages out of IBM No.	%
1.	Experiment minis (file transfer XFL)	426	4.5	21536	16.1	21534	13.6
2.	Experiment minis (miscellaneous)	455	4.8	32819	24.5	6766	4.3
3.	HP21mx development	42	0.4	561	0.4	561	0.4
4.	NORD10 development	46	0.5	823	0.6	825	0.5
5.	PDP11 development	265	2.8	6369	4.8	6374	4.0
6.	PDP11 support	292	3.1	11588	8.6	11592	7.3
7.	PDP11 station	85	0.9	1647	1.2	244	0.2
8.	Graphics polling	1922	20.2	4226	3.2	4226	2.7
9.	CALCOMP plot	160	1.7	430	0.3	7175	4.5
10.	BENSON plot	21	0.2	60	0.0	3360	2.1
11.	CERNET log	1667	17.5	7331	5.5	7329	4.6
12.	CERNET task load	187	2.0	658	0.5	1414	0.9
13.	Core load/dump	95	1.0	750	0.6	4082	2.6

In general it is true that the IBM File Manager has more traffic from incoming files than from outgoing ones, despite the existence of certain types of usage which only deal in outgoing files. This is evident from a study of Table 1 . It is less evident than the detailed study of Table 2 might indicate, since some of the types of use involve a symmetric scheme in which a message in one direction calls for an acknowledgement message in the reverse direction.

The following subsections give a reasonably complete breakdown of traffic to and from the IBM File Manager. Table 2 shows this breakdown in terms of links and messages. It can be seen there that the use from minicomputers in experimental setups exceeds that from non-experimental minicomputers in terms of quantity of messages, though not in terms of the number of logical links due to two particular applications which open links at very frequent intervals for the purpose of very short file transfer or enquiry.

<u>File Manager traffic from experimental minicomputers</u>: The minicomputers at experiments may use the IBM File Manager in many different ways. They may, in parallel with data-taking, send a percentage of the data to the IBM for storage in special-purpose short-duration data sample files or in normal permanent files. Alternatively they may perform off-line transfer of complete files between the IBM and their own magnetic tapes or disc storage. They may even send across complete disc dumps to the IBM (normally destined for the IBM mass store) as an insurance against local system crashes. Finally they may act as a remote print station by having a program which regularly polls the IBM to see if any jobs have produced output files destined for the station.

In view of all of these possible uses it is difficult to know exactly what the users are doing at any time. The only subdivision of this category is into the use of the standard file transfer program (called XFL) and the miscellaneous user programs. These are the first two entries in Table 2 . Note that XFL is a symmetric small message transfer utility, so that it is not possible to see which way the file is going. The miscellaneous programs, however, are non-symmetrical and it can be seen that the general direction is to the IBM, mainly due to file or disc dump types of transfer to the mass store.

<u>Support and development minicomputers</u>: For each standard type of minicomputer there tends to be one of that type used for development and one for support of other, less well equipped, minis. In the case of Hewlett Packard 21mx and Norsk Data NORD-10 minis, at the time of the analysis there was only a single combined development/support machine. For PDP-11 the two uses are separated on different minis, and the support PDP-11 also has software to allow it to act as an output station to the CDC or IBM via CERNET.

The amount of use made by these minis for the purposes of file transfer is shown in Table 2 Types 3,4,5,6 and 7. All except the station use still retain the symmetric message protocol, so that the direction of transfer cannot be deduced. It is obvious that the PDP-11's make much more use than the others, simply because they have had the software available for much longer. However there is little doubt that the NORD-10 and HP21mx will increase their use in future.

<u>Graphic output service</u>: The computer centre has two plotters for graphic output, one made by Calcomp and the other by Benson. They are both controlled from a Nord-10 which in turn has a CERNET connection. The NORD-10 regularly polls both the IBM and the CDC to see if there are any files waiting to be plotted. If so, the file is read via the appropriate File Manager and plotted. This improves the turnaround time for plotting, as well as making transfer via magnetic tape unnecessary. In Table 2, types 8,9 and 10 show the various connections of the NORD-10 to the IBM only. Normally, when there are files to be plotted, a single file requires a single logical link, so it can be seen that the total file transfer in the week was around 200 files.

<u>CERNET use of File Manager</u>: The minicomputers which make up the CERNET communications subnetwork themselves make use of the IBM File Manager in different ways. In so doing, they are essentially their own customers. The major user is a logging program which writes a message onto a log file on the IBM every time that there is some interesting event in the network. In addition, the IBM is used to hold non-resident tasks (overlays) and to transfer complete core loads/dumps in either direction. These uses are shown as types 11,12 and 13 in Table 2

<u>IBM-CDC file transfers</u>

The volume of traffic represented by file transfers between the IBM and the CDC is currently the fastest growing load on CERNET. This stems from the fact that each mainframe has some facilities not available on the other one, but of interest to all users. In particular, the IBM has a laser printer and a mass-store, whereas the CDC has a collection of remote Input/Output stations distributed all over the CERN site. In addition, with the existence of large collaboration experiment groups, it is not infrequent to have programs which may execute on either the CDC or the IBM, and so it becomes necessary to move both programs and data between the two systems.
Communication between the two computers is always under the control of a program which executes in the IBM and communicates with the CDC File Manager. Detailed figures of this type of use are given in Table 3 . In practice the real amount of data

<u>Table 3</u>: 'File transfer between CDC and IBM'

Type of use	Logical links No.	%	Messages into IBM No.	%	Messages out of IBM No.	%
1. Print on CDC RIOS	855	9.0	5308	4.0	49044	31.1
2. 95 char ASCII to IBM	305	3.2	5633	4.2	1704	1.1
3. 63 char CDC to IBM	305	3.2	676	0.5	559	0.4
4. IBM to CDC standard	236	2.5	872	0.7	2424	1.5
5. CDC to IBM standard	230	2.4	2642	2.0	617	0.4
6. Miscellaneous	440	4.6	15447	11.5	8480	5.4

transfer is at least 50% of the total CERNET traffic, since the messages tend to be long ones using an optimised protocol which does not involve symmetric message transfer.

The main use is the transfer of IBM output files to the CDC so that they can be printed at a suitable CDC Remote I/O station. This is shown as type 1 in Table 3 . During the week in question there were over 2000 such files, totalling approximately 170 megabytes. This probably corresponds to a sustained rate of about one kilobyte/second during most of the standard working hours.

The basic method for transferring these print files to the CDC, and also for print files in the reverse direction, is an IBM program which is scheduled to run every few minutes. The program interrogates the File Manager on the machine which is the source of the print files in question. If necessary, it then creates a suitable logical link for the transfer of one or more such files.

For the transfer in the reverse direction the object is to print CDC output on the IBM laser printer. There are two such types of output, 95 character ASCII or normal 63 character CDC codes. For each of these types there is a corresponding program run regularly in the IBM. In Table 3 these are types 2 and 3. Again they are normally made up of large messages, so that for that week the total transfer was about 700 files totalling some 40 megabytes.

For the average user wanting to transfer a single file of a standard type, there are ready-prepared jobs which are invoked by catalogued procedures after the user has specified the relevant parameters (file name on CDC and IBM, type, direction of transfer etc.). In Table 3 these are shown as types 4 and 5.

Finally, it may happen that a user either has a slightly non-standard file or wishes the file transfer to be only a part of a complete job execution on the IBM. In this case there are utility programs which he can execute via suitable JCL statements in the job. These are shown as type 6 in Table 3 . It can be seen that this use is quite popular (over 400 files per week) and is mostly used for obtaining files from the CDC.

Echo response

Table 4: 'connections to IBM Echo program'

Type of use	Logical links No.	%	Messages into IBM No.	%	Messages out of IBM No.	%
1. Normal monitor	1057	11.1	5279	3.9	5279	3.3
2. Extra monitor	154	1.6	1530	1.1	1530	1.0
3. Miscellaneous	93	1.0	6508	4.9	6503	4.1

Both the IBM and the CDC have a permanently running program willing to accept a logical link request from any CERNET user and then to echo back every incoming message. This is useful for program debugging, but can also be used to test the availability if the associated mainframe. Table 4 shows the use of the Echo program on the IBM for the week in question. The main user is the PDP support computer, which can then distribute the information on availability over a complete network of PDP computers. In addition one particular experiment with a complex PDP configuration does its own verification at less frequent intervals. These are types 1 and 2 in Table 4 . Type 3 is then the remaining miscellaneous debugging, hardware verification etc.

Remaining use

The remaining use is shown as types 4,5 and 6 in Table 1 . This essentially comprises interactive uses of the IBM from a computer connected to CERNET. Type 5 is particularly important because it involves using a large number-crunching program in the IBM to check individual events acquired by experimental data acquisition computers at experiments, and to give a response in a matter of seconds. This is known as Real Time Backup Facility (RTBF). For this particular week the usage was small, but this use varies according to the conditions of the CERN accelerators and the particular experimental group involved.

As has been stated earlier, CERNET is not intended for the connection of terminals to the IBM. However, there is sometimes a case where a terminal is hard-wired to a minicomputer, but the user would occasionally like to use it to connect to the Wylbur terminal system on the IBM. This has been permitted by including in the IBM a task which can interact with the Milten subsystem, and hence with Wylbur. The complete entity is called the Wylbur bridge, and its use is shown as type 4 in Table 1 . Each link represents a Wylbur session, and the majority of traffic is out of the IBM.

Finally, the real miscellaneous shown as type 6 in Table 1 is exactly what it says. It is perhaps interesting to note that if a CERNET-connected mini is suitably programmed then a user on its locally connected terminal can use the IBM interactively. This is also possible with TSO, but the use via CERNET may perhaps cause less overhead in the IBM. It is currently not CERN policy to run a widespread TSO service.

GENERAL REMARKS ON CURRENT USAGE

It is clear from the previous chapters that the current usage of CERNET is extremely widespread and intensive. The variety of types of usage has even surprised the most ardent supporters and the network designers. In essence what exists now is a form of practical distributed computing adapted to the particular needs of the CERN community.

Despite the success, it is clear to us that there is a long way to go yet before the use of such a network becomes automatic and natural. It is currently unnatural because the programming for the use of CERNET requires the programmer to know that he is using it. Thus, for instance, he must set up logical links and call specific routines to send or receive messages, and of course deal specifically with any errors which might arise. Such is not the case for a simple Fortran program which only needs to do read or write statements.

In one sense, what CERNET is providing is the means to deal with a distributed data base by allowing access to the various data bases. Of course the user must specifically keep track of the whereabouts of any relevant files and cater for their modification. This is not the same thing as a distributed data base, for which, ideally, a user would not have to know the whereabouts of files. In fact CERNET could be the tool by which a genuine distributed data base system could be developed. However the current state of the art is perhaps not yet advanced enough to understand and solve the problems which such a system would pose when applied to a heterogeneous system of computers.

Another area onto which CERNET has thrown some light is in the correspondence across different computers of standard types of file. The multiplicity of file types, character codes, word lengths, internal organisation etc. is at first sight daunting. The approach taken has been to concentrate on standard types of file whose concept is easily understood, and to make the use of such files as easy as possible. A simple example is a source file in a suitable format for the particular computer's terminal editor. On the IBM this is a special compressed format only understood by Wylbur, whereas the CDC has its own particular internal display code. In fact it is amazing how much can be done with a small set of file types.

Another way in which CERNET has provided a good example is the extent to which code can be structured so as to be portable across a variety of computers and operating systems. In particular, quite a lot of software which runs on connected minicomputers can, and has been, run on the CERNET nodes themselves. The standardisation of software and, even more important, of the interface definition for the standard user wishing to write high level software, is by no means perfect in the CERNET context, but we have made and are continuing to make a serious effort.

Part of this effort on standard software will be particularly applicable to the current trend towards providing specialised support for a particular type of minicomputer on a dedicated support machine of that type. The provision of a standard File Manager on these support minicomputers considerably enhances their potential, as well as being another step towards the distributed file base mentioned earlier. At present there are plans to put File Manager programs on Hewlett Packard HP21mx, Norsk Data NORD-10 and Digital PDP-11 and VAX computers.

Another part of the effort on standard software stems from the fact that some minicomputers are quite small and do not have much memory even in these days of rapidly diminishing costs. On such a mini one can provide the same interface for the end user in terms of Fortran calls, but the underlying software is smaller and limited in its ability to handle multiple users, multiple links or error cases. Provided that the end user is only doing simple things he is quite willing to accept such restrictions.

Even a rapid study of preceding chapters will have revealed that the main use of CERNET is simply to move files around between minis and the computer centre. The reason is clearly that these centres have sophisticated facilities which cannot be matched by even quite complex local setups. In particular the computer centre has good file storage and backup facilities and normally a large amount of file storage capability, plus a terminal system very appealing to many (but by no means all) users. CERN's mass store, laser printer and Remote I/O station complex are also important, but other centres might have different facilities equally important in their own context. The essential point is that with a data communications network these become accessible.

No paper on CERNET can be complete without some remarks on data transfer rate. However, frequently people are mesmerised by raw data rates, whereas what really counts is the services offered and the accessibility and reliability of these services. Of course some services are useless if the data rate is not above a certain threshhold, but it is my opinion that this threshhold is often below what many people would expect. Therefore all that I will say on CERNET data rates is that they are easily adequate for all of the applications described. In fact they are so high that if any new application should require higher rates I would advise a direct channel to channel connection between the computers involved.

A further very important point can be deduced from the statistics on use, namely that we are our own users. By this I mean that the people involved in providing CERNET and its associated services are at the same time users of those services. Thus, when there are bugs, we are frequently the ones to suffer. This is a great incentive to eliminating bugs as soon as possible.

CONCLUSIONS

Most papers end with a set of conclusions, and this one is no exception. These conclusions are however the personal ones of the author alone.

A project such as CERNET cannot be a success without a definite goal and a first user willing to be a guinea pig. Networks are sometimes described as a solution in search of a problem; this one is not. However, a good designer will not only assume that the first user does not necessarily know what he wants even if he thinks that he does, but will also try to anticipate other users' requirements. Even then he will, if he has any sense, realise that he may have little more idea than the potential first user of how the use will evolve. Therefore ensure you have a definite first user but keep all options open as far as possible.

Another conclusion which I draw is that the modern computer centre is no longer a set of computers but a set of services. To treat it as such is what interests the end user most. This remains true when different types of computer are involved. Unfortunately, much of the computer industry still aims to tie the users to their own particular brand of services, and in particular to offer a complex system making all of these services available in one place. If,on the other hand, one accepts that services can be distributed in an appropriate way, then the provision of the services can be much simplified. This is essentially the way that I see distributed processing.

Finally, I would expect that we have a lot to learn yet about the possible uses of CERNET. I can foresee some further uses, such as services offered by the network itself (value-added services) or as a result of connecting to other computers and networks inside and outside CERN. I also hope to make the use much more transparent to the community at large. However I am sure that in the next few years I shall find that many other possibilities exist. I hope I shall have learnt something by then, just as I hope that you have learnt something from this paper.

REFERENCES

[1] J. M. Gerard: The Development and Use of a High Speed Packet Switching Network in a High-Energy Physics Laboratory EuroIFIP 79, London, 25-28 September 1979