LOOSE COUPLING OF HETEROGENEOUS MAINFRAMES
IN THE CERN COMPUTING CENTRE

(To be presented at the SEAS Spring Technical Meeting, Bari, 7-10.4.1981)
Loose Coupling of Heterogeneous Mainframes in the CERN Computing Centre
D. Wiegandt

CERN DD Division, CH-1211 Geneva 23, Switzerland

ABSTRACT

The computing centre of CERN runs an IBM dual mainframe configuration (370/168 + 3032) side by side with a complex of two CDC CYBER machines (720 + 730) front-ending a CDC 7600.

Both the IBM and the CDC parts of the centre have machines connected to CERNET, a local packet-switching network, primarily to provide support facilities for the many minicomputers on site.

Only quite recently the potential of the communication between the two parts of the computing centre started to be fully exploited. Cooperating, permanently running processes in the two computer systems permit job submission, file transfer, and output spooling in either direction between IBM and CDC machines.

These facilities help to optimize the utilization of special peripherals (laser printer, microfiche, mass storage system, remote stations etc.) or special features (high CPU power, good terminal response, user-friendly editor, etc.) available on one system by users of the other system. There is, however, no automatic resource sharing or load balancing.

The implementation was done in an asymmetrical fashion, with the centre of activity being on the IBM 3032.
COMPUTING AT CERN

CERN, the European Organization for Nuclear Research, was established in 1954 as an international research laboratory for fundamental elementary particle physics research. CERN provides sophisticated research tools for the community of physicists trying to understand nature and to analyze the constituents of matter. Four large accelerators are available on the site:

- a synchro-cyclotron of 680 MeV (SC)
- a proton synchrotron of 28 GeV (PS)
- a proton synchrotron of 400 GeV (SPS)
- the intersecting storage rings (ISR)

Ever since CERN came into being, computing has played a major role among the services provided for physicists by the organization. Over the past 20 years the experimental techniques in high energy physics have become more and more sophisticated and have required more and more sophisticated computing methods for analysis. In parallel with this, we have seen the amount of data produced by the experiments rising enormously due to the invention of new detection methods. Some 20 years ago, physicists would read a few mechanical scalers and write down the results on a sheet of paper. Nowadays they record high energy physics events on a magnetic tape as fast as the tape drive can manage.

So it is quite understandable that computers in CERN are counted by hundreds and that the CERN computing centre is one of the most powerful ones in Europe. The bulk of computing power is at present provided by computers from two manufacturers: CDC and IBM. A CDC 7600 is front-ended by a CYBER 170-720 and a CYBER 170-730. The IBM configuration consists of a 370/168 and a 3032, running MVS and sharing user data sets and spool files.

The rapid growth in the number of minicomputers on site is mainly due to the fact that physics experiments are getting more and more complex. Quite often, physicists are looking for rare events almost completely obscured by a wealth of background events which are of very little or no interest. Only rapid and flexible decision making can avoid recording of uninteresting data. Modern detectors are already made to be used with computers and present results in digitized form. So most of to-day's physics experiments have one or even a cluster of on-line computers for experiment control, data acquisition and data recording on magnetic tape.

The computing power of these minicomputers is, however, rarely sufficient for a full analysis of the events. To monitor their experiments, physicists would therefore take a tape with raw data to the computing centre (a method nicknamed "bicycle on-line"), have part of these data analyzed by a batch job and the results
printed either centrally or on a printer on one of the Remote I/O Stations (RIDS) near their experiment. Although the transport of a full magnetic tape even on a bicycle represents quite a substantial data rate, it was felt that this data communication process was not quite up to to-day's technical standards and that improvement was necessary. So the development of CERNET started about five years ago.

**CERNET**

CERNET is a packet-switching network, consisting of a number of switching nodes (MOCOMP computers), interconnected by high speed (2.5 .. 5 Mbits/s) data links. User computers (subscribers) are connected through CAMAC links developed at CERN, or, in the case of the computer centre machines, through channel adaptors.

Information is exchanged over logical links established between communicating programs according to the rules of the CERNET End-to-End Protocol. The basic unit of information exchange is a message, the basic unit of information transport is a packet. A CERNET program in every connected computer, called the Transport Manager, takes care, among other things, of assembling packets into messages or disassembling messages into packets, when the message size exceeds the maximum data contents of a packet. Packets are limited to 2046 bytes. There is no theoretical limit for the size of messages. For practical purposes all connected computers have established such a limit (32K bytes on IBM computers). Packets between subscribers are always sent on a "best route", which is updated automatically whenever a topological change affecting this route occurs in the network.

The primary aim of CERNET was to provide physicists with a communication medium enabling them to send a small percentage of physics events from their experimental minicomputer to the computer centre for analysis by a program running on a computer centre machine (sample analysis). Physicists have two principal methods for sample analysis:

1. On-line sample analysis. A permanently running program (quite often the standard analysis program where the tape I/O routines have been replaced by CERNET I/O routines) communicates with a partner program in the minicomputer at the experiment. It receives samples of the experimental data via CERNET, does a full analysis of the event, and sends the results back to the program in the minicomputer, often in a form suitable for display on a graphics terminal on that minicomputer.

2. Deferred sample analysis. Samples of events are written to a disk data set in the computer centre. When this data set is
full a suitable analysis program is started and the results are interrogated later.

To exploit this second method conveniently, a suitably general mechanism for remote access of disk data sets, submission of batch jobs and retrieving job output is needed. This mechanism has been implemented by a program called File Manager.

**FILE MANAGERS AND THEIR CAPABILITIES**

One of the main goals of the development of CERNET was to provide remote file access. This is achieved by the collaboration of two communicating programs, a file access or file transfer utility and a File Manager. These programs communicate according to the rules of the CERNET File Access Protocol. The commands OPEN, CLOSE, GET, PUT, etc., issued by the utility program are carried out by the File Manager (provided a certain number of checks have been passed successfully). CERNET File Managers are at present running on the CDC CYBER 170-720, the IBM 3032, several VAX11/780s, and PDP11s.

File systems on computers of different manufacturers are usually totally incompatible. Faced with this difficulty you can either restrict yourself to what might be regarded as a common subset - the transfer of card images - provided you have found a solution to the character translation problem, or you have to assume that users know what file system they are dealing with and that they want to exploit the particular properties of that system. CERNET File Managers have indeed been coded to allow most of the features of the target file system to be exploited by remote users. The rather unfortunate consequence of this implementation is, however, that it is not at all trivial to write a universal file transfer utility, permitting file transfer from any CERNET subscriber computer to any other such computer.

In addition to describing the rules for accessing data sets on disk, the CERNET File Access Protocol covers job submission and output retrieval. Here the differences between operating systems of different manufacturers are at least as big as for file systems. The possibility of submitting jobs to and retrieving output from the computing centre is, as was stated above, a very desirable feature for deferred sample analysis. Therefore job submission and output retrieval were implemented in the CERNET File Managers on CDC and IBM.

Under MVS/JES2, the "internal reader" mechanism allows a very straightforward implementation of job submission. Output retrieval, however, is by no means trivial. Fortunately, the implementation of output retrieval could be based on an interface to JES2 developed for the FETCH command of the WYLBUR editor run-
ning at CERN. This interface also accepts standard JES2 commands and returns appropriate answers to the calling program, just as if it were a console. This feature was exploited to implement job enquiry and job control via the File Manager, e.g. the commands LOCATE, CANCEL, PURGE, etc.

Once these facilities had been made available to users and had been in use for some time, a physics group requested some means for printing the output of all their jobs running on one of the IBM computers remotely on a printer attached to their mini-computer. The result of this request was twofold:

1. A File Manager command PRINTQ(Rnnn) was implemented, simulating the JES2 command $QON,Q=PPU,R=Rnnn, providing the requestor with a list of the contents of the print queue for a particular (fake) RJE station. Individual job output can then be FETCHed, printed remotely, and PURGED.

2. The only means for transfer of output files from IBM to CDC which existed at that time, was a facility called TIELINE, a HASP workstation simulation on the CDC, using a 4800 baud link and based on obsolete hardware. It was realized that one could replace it by a file transfer program talking to the File Managers on both CDC and IBM. This was the origin of CHIMP.

CHIMP.

CHIMP stands for CERN High speed Inter Mainframe Program. This program is the coupling link between the two parts of the CERN computing centre. A close look at the configuration of the CERN computing centre will show you the tasks that CHIMP has to deal with:

1. Almost all remote job entry stations on the CERN site are connected to the CDC CYBER 170-720. To print IBM output on a CDC RIOS these output files have to be sent to the CDC File Manager and their final destination, i.e. to which RIOS they should go, has to be indicated.

2. The text-editing facilities on the IBM computers (i.e. WYL-BUR) are quite adequate for a FORTRAN-oriented user community, and the terminal response is far better than that of CDC's INTERCOM in our present configuration; however, the biggest machine in terms of CPU power is the CDC 7600, and this CPU power is sometimes badly needed for complex analysis programs. So what users like to do is to edit (and often also develop) their program on the IBM part of the centre, then transfer it to the CDC machines, still keeping
the habit of submitting jobs from a WYLBUR terminal, now to
a CDC machine, and to scan output on that same WYLBUR termi-
nal. CHIMP therefore takes care of transferring jobs to the
CDC File Manager for submission on any of the CDC computers.
It will also interrogate the CDC for job output to be trans-
ferred back to the user's FETCH queue.

3. There are special peripheral devices connected to the IBM
computers in the computing centre: a 3800 laser printer, a
3850 mass storage device, and a KODAK microfiche printer.
With the help of CHIMP, users of the CDC computers can make
use of these special peripherals on the IBM. Data sets can,
for example, be transferred from the CDC for storage on the
IBM 3850 MSS, or output of a CDC job can be directed to the
laser printer on the IBM, simply by requesting this on a
control card in the job.

CHIMP is a PASCAL program (with some subroutines in Assembler)
that runs as a started task on the IBM 3032. It talks to the File
Managers on CDC and IBM. This solution had to be chosen because
the CERNET implementation on the CDC does not yet enable a user
program to communicate via the network.

Since CHIMP first came into action, it has had to be modified
quite considerably to adapt it to an ever increasing set of tasks
and duties. To permit interleaving of transfers going on in
opposite directions, it has been split into two independently
operating started tasks. One of these interrogates the transfer
queues in the CDC CYBER 170-720 and transfers anything it finds
there (i.e. jobs to be submitted to the IBM computers or output
tapes to be reprinted on the laser printer or on microfiche) to the
IBM via internal readers. This task uses a local modification of
JES2 which introduced a "list offline" control card, to permit
the direct transfer of information to be printed from the JES
input queue to the appropriate SYSOUT queue, i.e. without running
a job. The other task scans transfer queues on the IBM comput-
ers, and transfers jobs and output to the CDC machines. The
queues on the IBM computers are implemented as NJE transfer
queues for a simulated NJE node. Whenever any of these CHIMP
tasks finds its work queues empty, it goes to sleep setting its
wait time according to the preceding level of activity.

What CHIMP does not provide is job enquiry and job control
from a WYLBUR terminal or an IBM console on the CDC computers.
These facilities will be implemented later this year, probably
using a third started task that establishes a direct commu-
nication between JES2 on the IBM and the File Manager on the CDC.

There has never been an attempt to use this kind of coupling
to achieve any load balancing or resource sharing between the two
parts of the CERN computing centre. It is entirely up to the user
where he wants to run his job and where he wants his output to be printed. Information which is distributed regularly (CERN COMPUTER NEWSLETTER) is meant to keep users informed about the computer centre facilities at their disposal and to help them to use these facilities in an efficient and convenient way.

**SOME NUMBERS ON USAGE**

The use of CERNET has increased considerably over the last few years. The analysis of the network statistics (user-written SMF records) for a typical week towards the end of each year shows for instance in the total number of links to the IBM computers an increase from 2058 per week in 1978 to 9500 per week in 1979 and to 14564 per week in 1980.

An even steeper increase can be found in the number of links per week to the CDC File Manager, which are included in the above totals: 119 (5.8 %) in 1978, 2371 (25.0 %) in 1979, and 5133 (35.2 %) in 1980.

The rise in the amount of data being transferred is quite spectacular. The figures for the total number of messages per week are

- 1978: 106,267
- 1979: 291,852
- 1980: 524,825

and correspondingly, the number of messages to and from the CDC File Manager

- 1978: 17,631 (16.6 %)
- 1979: 93,406 (32.0 %)
- 1980: 336,027 (64.0 %)

Currently, the number of transfers from IBM to CDC is approximately 1000/day, of which about 400 are jobs being submitted to the CDC computers, the rest being files for printing at CDC remote stations. Similarly, of the 560 transfers from CDC to IBM, about 160 are jobs for submission to the IBM, with the rest going into FETCH queues of individual users and to the laser printer.