THE CENTRALIZED COMPUTER RESOURCES IN EXPERIMENTS ON HIGH ENERGY PHYSICS

A.E. Chevel [1], A.N. Lodkin [2], A.A. Oreshkin [3]

Introduction

Here we would like to cover briefly the key points of centralized computer infrastructure in high energy physics experiments (HEP) of PNPI RAS. The features of providing HEP experiments from the computer point of view were repeatedly considered (e.g. R.M. Brown and others, IEETF, 1966, pp. 114-120, H.E. Devies,1976 CERN school of computing, pp. 161-172). In publications devoted to computer aspects of scientific reseach automation the specific technical tasks and the methods of their solution are frequently discussed. It would like now to point out important stages of achieving common goal which was pursued by all papers discussed in this review, namely: increasing man-machine systems efficiency in scientific information proccessing in the case of HEP experiments in PNPI.

To bear in mind increasing efficiency includes a lot of components. In other words it is meant increasing volume of really available possibilities for scientific information proccessing at the same or less price which can be expressed variously:

- in increasing technical and functional capabilities of man-machine system as a whole;
- in increasing reliability of hardware and software;
- in reducing operating staff necessary for supporting computer complex in operational state.

It is supposed that while improving one parameter of the system other parameters don't change or change in direction of increasing efficiency.

The source of scientific data in nuclear physical experiment is measuring experimental setups. For the data to be proccessed on centralized computer capacities it can be acquired through communication lines or read from secondary memory where it was stored earlier.

The use of centralized computer resources is caused by relatively high cost of computer setups and a large number of users which varied from 10 to 200 people at different times. There is no doubt that centralized computer setups with unique technical parameters within institute (or division) framework always existed, exist and will exist. The subject of discussion or optimization can be only the way of integration of such setup into institute (devision) computer infrastructure.

The centralized computer setup is designed not only for computing but for effective use of more expensive resources than setup itself: PNPI synchrocyclotron and working time of PNPI reseachers. The latter determined important feature – centralized computer was used in two

[1]e-mail:*shevel@pnpi.spb.ru*
[2]e-mail:*lodkinan@pnpi.spb.ru*
[3]e-mail:*oreshkin@pnpi.spb.ru*

modes: for serving experiments on-line and typical running batch jobs. On-line mode was always of higher priority than running batch jobs.

The use of a computer for work on-line supposed that a program complex was run which was used for specific physical experiment. Typically a program tasks were:

- acquiring information from experimental setup and storing it on mass storage (as a rule, on magnetic tapes);
- partial or complete proccessing information acquired;
- displaying partial results.

Here it is important to note the volume and the intensity of experimental data acquiring. Of course these values varied depending on experiments character, however in many cases the volumes were estimated at hunreds megabytes per day. Typical measurements duration was from 1-2 days (debbuging and calibration) through 20 days and more.

Below the main basic technological problems and methods of overcoming them are discussed which were caused by the computer technics generation available.

The first computer of the third generation

In autumn of 1973 the computer ES-1020 with main memory 64 Kb was put into operation. It had two disk drives with 7.25 Mb each, four tape bobbin drives (24 Mb on bobbin), card reader, card punch, line printer. The average capacity on mixture of instructions gave about 20 thousands instructions per second. Disk Operating System for computers of Edinoi Serii (DOS ES) was run which had a number of compilers for such languages as Fortran, PL/I, COBOL and other typical program components: varies utilities, linkage editor (task builder) and so on.

Note the problems facing at that time.

- DOS ES version used in period 1973-1974 was poorly suitable for acquiring measuring data from experimental setups. In particular there were limited capabilities of proccessing non-standard devices interrupts.
- The limited space on disks and in main memory also made it much difficult to transfer applied programs which were available for example at CERN.
- The serious problem was the reliability of hardware (tape drives, disk drives). Naturally it caused the fear to lose software and data.

To a considerable extent the same problems remained also with putting into operation next computer setup – the complex of two computers ES-1030 in the middle 1977. This setup had 16 disk drives with 7.25 Mb each, 8 tape drives and other periphirials. It took almost 200 square metres of premises for deployment of this hardware and the labour of 20 people (hardware ingineers, programmers, operators) was required to support this setup in operational state. ES-1030 computer based computing complex was employed as centralized computing resource in experiments up to 1989 including.

At first stages of using third generation computers in physics experiments it was required to create software-hardware base for using the computers of such class in making measurement.

Thus the special telecommunicational I/O controller for transmitting data from experimental setup to central computer was developed which enabled to connect up to 16 experimental setups [1]. By this step hardware base was created for reliable link between ES computer and mini computer (SM-3, SM-4) which were operational element of all experimental setups. Architecturally this solution allowed to have relatively simple mini computers be included in experimental setups. On the other hand relatively expensive and complicated in servicing devices such as high speed bobbin tape drives, high capacity disk storage (of that time) could be concentrated on centralized computer.

In order to make computer infrastructure more reliable and friendly to man the developments were undertaken in the area of system software and hardware architecture.

One of the first works on software for computers of such a class was the supervisor modernization (resident part of operating system) [2].

For increasing reliability of work with tape drives special algorithm was developed which prevented a loss of experimental data in most cases of device hardware failures and operator errors: invalid tape positioning, invalid tape bobbin [3].

For improving convenience of working with secondary storage on disks special access method was implemented which made it possible to consider secondary storage as main memory with random access [4]. The data exchange was performed by calling appropriate subroutines with parameters.

The neccessity to improve the quality of interface between man and computer led to the development of the dialog system [5], which was basic during a number of years. In this system there were used: a dialog initiated by computer (mode "menu"), a set of patterns for preparing user programs of serving specific experiment. The system enabled to implement quite complex scenarios of work in an experiment which caused maximal use of all centralized computer resources available at that time.

The great significance was attached to applied program system developments which reperesented sets of programs optimized on resources used and degree of stability to hardware failures. These systems implemented complete simple scenarios which were easily formulated in substantial terms for example as follows.

1. acquiring a portion of experimental data from communication line.
2. writing a data portion acquired on magnetic tape.
3. writing a data portion acquired on disk.
4. providing synchronization with user data proccesing program.
5. goto to point 1.

Here it is supposed that user program may be (or not) started which has a possibility to read experimental data from disk and proccess it. The synchronization is needed to prevent proccessing the same data more than once.

For implementing such scenarios several applied systems were developed which provided various in volume service for experimentator, at the same time they consumed different shares of centralized resources according to provided opportunities [6]. Thus there was a quite large number of programs which could be used in specific experiment. One of the examples of using centralized computer infrastrucure is given in [7].

In period 1980-1981 it became to be obvious that the bottle-neck of centralized computer resources was preparing and debugging programs. Most part of program was punched and input into computer using card reader. On the other hand in experiments the electronic means of displaying – displays were widely used which were connected mainly to mini computers of

class SM. All these circumstances stimulated the development of local terminal network [8]. As a controller of local terminal station computer SM-4 was used which several tens of terminals were connected to. The terminals could be connected at the distance of up to several hundreds metres on two twisted pairs that made it possible to place terminals where neccesary. Important feature of terminal system was that any terminal could be logically connected to any computer: ES-1030/1, ES-1030/2 and SM-4. This enabled to prepare and debug programs for several computers literally without rising from the table.

The experience of program development led naturally to employing special technology of developing large programs and program systems [9], which was used in a number of works [5, 6, 8]. The technology named metaprogramming allows to describe an architecture of large program systems regardless of programming languages used and computer type. In metaprograms only minimal set of operators is used: check, condition brunch, control transfer, comment, subroutine call. Thus quite simple means enable to describe a logic of complex program systems.

Among others was important problem of long term keeping program texts, descriptions and other documents on computer media [10]. It is a matter of archive scenario in traditional sense: when neccessary to write same programs or data for long term keeping (many days, months or years).

Main capabilities of this program package are as follows:

- writing information on magnetic archive tape in OS ES standard with possibility to append information to partialy filled archive tape;
- reading any module or module group by name, date, version number and so on from archive tape and writing on another tape or disk.;
- at last this system had special methods of error proccessing which enabled to restore information even in case of very severe magnetic tape failures.

Virtual Machines System

In 1988 the first of two ES-1046 computers was installed. It was running more powerful operating system - Virtual Machines System (VM).

There was developed the network of terminals (about 70). They were mainly alphanumeric terminal of type ES-7970 and ES-7920. Although computing resources were insufficient for simulating complex experiments the terminal network made it possible to work to many scientific workers and engineers, to prepare papers, descriptions as well as to formulate problems solutions in language of computer programs.

In the second half of 1980-s there was no doubt of future of wide using personal computers (PC). In order to simplify transition from traditional computers to PC there were used PCs running emulation programs which enabled to emulate terminals ES-7970, to copy files from ES-1046 to PC and vice versa, as well as to use ES-1046 disk storage as a additional PC disk drive. Thus it was provided the possibility of smooth transition to desktop computers.

One of main massive problems which we were faced with in conditions of considerable growth of users number was low reliability of data storing on disk media. That was caused by unreliable hardware and software components of computer setup. As any information stored on secondary storage (programs or data) represents by itself materialized working time therefore a loss of information is a obvious large waste of institute.

With the goal of decreasing probability of information loss it was undertaken the development of program complex [11], which would enable to make backup copying (including incremental backup copying) and restoring under full control of this backup program complex.

Making backup copying or restoring didn't require skilled worker. Any worker who was able to mount tape in tape drive could perform any operations: the backup program itself said which tapes to be installed and controlled the correctness of tape mounting. During execution of one or another operation the protocol was logged with which not only all actions of operator but the individual directories status could be restored. A such mechanism was used from 1989 through 1992 on ES-1046 and later on communication and computing servers.

In such a way it was developed quite comprehensive backup system which in such routine operations as copying and restoring left for choice of staff only absolutely neccessary minimum for example which tape drive to choose.

Operating system VM gave a lot of possibilities to operating staff: it was possible in a number of cases to run different operating systems at the same time for different users, to test equipment while system being operating and so on. However for end users no qualitative changes happened in character of using centralized computers. Computer was mainly used as a big calculator.

UNIX

The transition to new operating platforms (systems like UNIX) for high performance and average workstations was predetermined by advent of microproccessors manufactured on RISC technology. As at once many types of microproccessors (MIPS, DEC, HP and others) appeared so in order to minimize expenses on basic software operating platform had to be relatively easily installed on machines with different architecture at the same time preserving all wealth of capabilities. To that time system UNIX was operating platform which met such requirements. Up to now operating platform UNIX is de facto standard – universal program platform for all workstations and many supercomputer installations.

The gradual transition to new platforms (hardware and software) took place around the world in all areas of using computers including scientific area. So in CERN traditional architecture computers mainframes (such as ES series computers) were closed down in 1995, in DESY they are planning to be closed in 1996. In PNPI this transition took place at rapid rate due to lack of financial resources for gradual transition to new platforms and finished in the late 1992.

For the first time UNIX (Xenix) in PNPI became to be avaiable for users in 1989 on PC/386. CERN library was installed on this computer. Then the programs installed on PC/386 were transferred to PC/486/33 [12] and now they are used on centralized communication server.

The important feature in transition from centralized systems of type mainframe to computers based on microproccessors was that this transition was made by the same specialists who had already the experience of serving a large number of users on centralized computer setups that in turn saved time and labour of experimentators during period of transition to new computer platforms.

Local network

On the very early stages of using computers in nuclear physics it was assumed that centralized

computer setup had to be connected to experimental setups. With the purpose of connecting computers into united coherent environment there were carried out the experiments and in area of local networks. The first Ethernet network was developed in 1990 (network "MIFIR"). The network consisted of 4 nodes: ES-1046, SM-1300, SM-4, DVK-1. About at the same time the new comprehension of institute computer infrastructure was shaped [13].

In 1993 it was realized first Ethernet TCP/IP network segment based on thin coaxial cable 50 om. There were three workstations and some X-terminals in the segment. At present institute scientific local network is based on Ethernet 10 Mbit/s technology with UTP category 5 cabling system. The interfaces between main institute buildings and central network node are implemented using fibre optic cables. There are more than 200 nodes in institute network. The main network protocol is TCP/IP.

Centralized servers

PNPI computer infrastructure is moving towards specialization of centralized servers connected to institute local network. Enumerate briefly available servers.

Communication server of institute is implemented on base of Intel/Express Pentium/60, memory 64 Mb, EISA, modems battery, disks 5.3 Gb, Exabyte 8200, floppy 3.5", CD ROM. About 340 users are registered on the server. About 200 people use the server every month.

Computing server is implemented on base of SGI (R3000), memory 224 Mb, three SCSI adapters, disks 16.4 Gb, Exabyte 8500, 6 proccessors. There are 80 users registered on the server and more than 60 users use the server every month.

Data base server is implemented on DEC Alpha 4/1000 266 MHz, memory 128 Mb, disks 14.9 Gb, CD ROM. The server was installed in the early June of 1996. The server is supposed to be used by all institute laboratories for storing institute scientific data base.

Internet

In PNPI for a long time experimental works in the area of computer communication with large foreign physiscs centres were carried out. So in 1987 more than ten communication sessions with CERN were carried out. The communication was performed from alphanumeric terminal via modem 2400 bits/s to Petersburg (SPIIRAN), then through network AKADEMSET to Moscow (node VNIIPAS), further via dedicated telephone line to communication node in Vienne (Austria). The arrangement of every session required most qualified attention and as a whole represented itself nontrivial procedure.

At last in March of 1994 the first IP channel on base of switched telephone line "ISKRA-2" [PNPI - ITEP, later Radio-MSU Moscow] was realized which gave possibility of using world computer network Internet. The real achieved capacity of line "ISKRA-2" [Gatchina-Moscow] with using modems Telebit WorldBlazer was 12 Kbits/s.

Since 23 September of 1995 PNPI has satellite IP channel with bandwidth 64 Kbits/s (full duplex) which was implemented with help of international scientific foundation INTAS according to the project INTAS 93-56. That enabled to widely use new possibilities having opened for institute in Internet.

The main institute tasks in information area are a conversion of information of the matter structure from natural form into human one, i.e. articles, books, manuals ans so on. It is natural that effective ways of proccessing all kind data using computer are neccessary component of scientific reseach. However making decision on proccessing scientific information is done by human being but not machine. One of the powerful stimulus for accelerating making decision is an information exchange between reseachers. Such exchange is fulfilled according to various scenarios which are realized by appropriate software: mail, ftp, news, www, irc, ytalk, white board, vat (audio), nv (video) and so on. [14].

Main directions

As it is seen from brief review of works made on developing centralized computer resources a computer infrastructure never can be completely accomplished. In connection with this it is useful to note perspective areas in PNPI computer infrastructure.

- It is natural that for convenience of proccessing all information has to be represented in electronic computerized form. As there is a lot of electronic forms so it should be provided the tools and methods for free conversion from one electronic form into another one. There is no doubt that one should have wide range of equipment for conversion from traditional forms of representation into electronic and vice versa - scanners, printers, video and audio recording and playing equipment. Among them there is of course an equipment for seminars and presentations.

- It is important to store data archives on well protected and preserving media with a convenient network access to it.

- Internet channel of sufficient bandwidth is a neccessary condition of participating of the institute in international scientific labour sharing.

  According to the estimation the institute should have up to 2000 year main Internet channel with bandwidth no less than 2 Mbit/s (full duplex).

- Information infrastructure (infostructure) of the institute - Intranet.

  The availability of approriate software for supporting a variety of data exchange scenarios makes it possible to use these tools for internal institute needs, for example, such as institute announcement board:
  - seminars;
  - scientific and coordinating boards: announcements and decisions;
  - administrative instructions, orders;
  - technical information: time-table of work of reactor, accelerator, warehouses, publishing and foreign departments, other service;
  - telephone book;
  - private announcements, other information.

- Virtual reality.

  The virtual reality tools enable to shape on workstation screen 3-dimensional picture of being built or already built object from any point of view: experimental device, building, nuclear reactor body and etc. Changing a point of view it is possible to "travel" inside

an object and to view the interior which is formed dynamically on workstation screen according to working drawing with the speed of changing view point. That permits the experts to travel through tubes, crossing, walls and so on.

Conclusion

The qualitative differences between existing centralized computer resources and mainframes stage which came to an end in 1993 in PNPI make no doubts for anybody. It is useful to note what they show themselves in.

As early as in 1980-s it became clear that for building computer infrastructure one should use ready products (software and hardware) which are on computer market or free available such as GNU packet or CERNLIB. In other words it is not perspective to try making own developments of small components and adjusting them for some special features of computing environment; the ready products should be used as they are. New developments and projects in area of centralized computer resources should be made primarily in choosing available on market solutions of lower level (workstations, servers, routers, network standards, programming systems, robotized peripherials and so on.), as well as in the sphere of infostructure.

Computer is flexible tool for communicating between people and representing data in convenient for perception form: text, graphic, audio, video and so on. Fax, autoanswer, TV set, telephone, radio set, video tape recorder, device for audio playing – all this is easily combined in a personal computer. Using computer without network is considered to be not reasonable like transporting car using horse power. However while estimating the projects on changing computer infrastructure it should be taken into account an integral expenses on realization including the cost of further maintenance. One of the consequences of such statement is the conclusion that for most part of users it is more technological to use X-terminal instead of computer.

As the volume of information in WAN is more than in any library of the world so one of important new infrastructure element is special tools for importing neccessary information from WAN and exporting information to networks of different types and scales.

At last the most important difference – is a growth of a number of scenarios of interactions between people which computer is used in as a operational element.

References

[1] *B. Bochin, Yu.V. Elkin, P.V. Neustroev.* Preprint LNPI-284, Leningrad, 1976, 18p.

[2] *A.E. Shevel.* Preprint LNPI-128, Leningrad, 1974. 8p.

[3] *M.N. Andronenko, I.I. Gracheva, A.E. Shevel.* Preprint LNPI-658, Leningrad, 1981. 24p.

[4] *I.I. Gracheva, A.A. Oreshkin, A.E. Shevel, V.A. Schegelsky.* Preprint LNPI-496, Leningrad, 1979. 16p.

[5] *T.S. Serebrova, L.F. Suvorova, A.E. Shevel, S.V. Yazikova.* Preprint LNPI-302, Lenigrad, 1977. 34p.

[6] *A.N. Lodkin, B.Yu. Sokolovsky, A.E. Shevel.* Preprint LNPI-1063, Leningrad, 1985. 30p.

[7] *S.S. Volkov, Yu.A. Dotsenko, L.G. Kudin, N.P. Kuropatkin, A.A. Oreshkin, O.E. Prokofiev, A.E. Shevel.* Preprint LNPI-674, Leningrad, 1981. 42p.

[8] *A.G. Atamanchuk, V.I. Benevelsky, I.I. Gracheva, N.M. Gulina, S.G. Dolgobrodov, A.N. Lodkin, P.V. Neustroev, A.A. Oreshkin, E.M. Orischin, T.S. Serebrova, N.A. Seregin, B.Yu. Sokolovsky, E.V. Fotieva, A.E. Shevel.* Autometria, 1986. No. 4, p.45-49.

[9] *A.E. Shevel.* The Programming, 1986. No.1, p.31-37.

[10] *I.I. Gracheva, A.N. Lodkin, A.E. Shevel.* Preprint LNPI-914, Leningrad, 1983. 21p.

[11] *S.A. Kulikov, A.E. Shevel.* Controlling Systems and Machines, 1991. No.6, p.133-135.

[12] *A.N. Lodkin, A.A. Oreshkin, T.S. Serebrova, A.E. Shevel.* The Proceedings of the conference. Dubna, 1993. p.114-116.

[13] *A.E. Shevel.* Preprint LNPI-1647, Leningrad, 1990. 12p.

[14] *A.E. Shevel.* The materials of PNPI jubilee XXX winter school, St. Petersburg, 1996. p.247-263.