Distributed Systems for Data Handling

Vladislav P. Shirikov

HAL Id: hal-01568405
https://inria.hal.science/hal-01568405

Submitted on 25 Jul 2017

Distributed Systems for Data Handling

Vladislav P. Shirikov

Joint Institute for Nuclear Research
Dubna, Moscow, Russia
shirikov@jinr.ru

Abstract. This article provides a brief historical review in the field of creation and use of distributed systems for computerized accumulation and information processing (handling) in Soviet Union, and especially in its scientific centers such as Joint Institute for Nuclear Research and not only. The author tries to demonstrate- how the chosen methods and ideas permitted to realize the systems, which could be named as "PreGrid-complexes" in Russia.

Keywords: Data handling, distributed computing, GRID

1. Introduction

The creation and use of distributed systems for computerized accumulation, transmissions, and information processing in interests of scientific research started less than fifty years ago. Apparently, the first such system in Soviet Union occurred in 1961 for radio-channel transmission of Atlantic Ocean objects information to first Soviet semiconductor universal computer DNEPR in the Institute "Cybernetics" (Ukrainian Academy of Science) in Kiev. Basic realizations of multicomputer complexes for distributed jobs and data handling started in 1960s and the 1970s practically in all prominent research organizations in Soviet Union and abroad.

At the "top of success" of their development, we would have Grid systems in the nearest future; that is, complexes that correspond to three criteria according to the definition of one of Grid technology's ideologists, Ian Foster. He states [1]:

" The Grid is a system that:
- coordinates resources that are not subject to centralized control …
- using standard, open, general-purpose protocols and interfaces …
- to deliver nontrivial qualities of service, relating for example, to response time, throughput, availability, security and/or co-allocation of multiple resource types to meet complex user demands, so that the utility of the combined system is significantly greater than that of the sum of its parts."

As an ideal, such a Grid system permits the ordinary user to input his/her job for batch or interactive processing without the necessity to define the job's "passport" (JDL-file) via attribute requirements. This includes the address of certain computer computing elements (CEs) as part of computing Grid resources. This is true if the system`s middleware has high-level services such as resource brokers (RB) and many of CEs, as it is now in large-scaled realized projects such as Enabling Grids for E-sciencE (EGEE). For example, see http://lcg.web.cern.ch/LCG/ and http://egee-

jra1.web.cern.ch; however, there is no such service in the American Open Science Grid (OSG) project. We talk about distributed systems of twentieth century, but it is necessary to say that many ideas realized in contemporary Grid structured were not quite unknown to specialists, which created systems more than thirty years ago, as it will be shown below. In other words, it is possible to say, that the Grid is the realization of "old" ideas based on new technologies.

2. Early Activities at JINR

At such a physical scientific centre in the Russian territory, as the Joint Institute for Nuclear Research (JINR), which was organized in 1956 and closely collaborated with Western physical centers, especially with European Centre for Nuclear Research (CERN, Geneva), problems arose. These situations demanded a common approach for applications realization and matched software environment in computing servers; this would make easier for the data and the programs to exchange at least with use of magnetic tapes in time of direct computer links absence. As to main type of processed information, then at JINR, CERN and other centers (e.g. Brookhaven, Berkeley, and Daresbury), this was a large amount of experimental data from accelerators detectors. At JINR, which also carried out theoretical and experimental research in fields of low and intermediate energy physics, it was necessary, in particular, to register and process data taken from spectrometers in experiments with the use of a fast neutron pulsed reactor.

According to these application problems, one of main tasks of the 1960s and later was the technical and software support of a whole sequence of measuring, preliminary and final data processing for different fields of experimental physics and also for theoretical physics. The chosen method to resolve this task was essential. It was necessary to start building distributed multileveled computing complexes with the use of perspective and standard measuring equipment and standardized programming tools for applications (such as compilers for high-level languages as FORTRAN. The tools had to be compatible with Western versions of program libraries for universal and special destinations including programming tools for algorithms parallelization and tools for user access to computing resources of different types. Main tasks at JINR, mentioned below, were initiated and guided (through July of 1989) by a corresponding member of the Academy of Science (USSR) N.N. Govorun, the chief of mathematical division in the JINR Computing Centre (1963-1966), vice-director (1966-1988) and director (1988-1989) of Laboratory of Computing Technique and Automation (LCTA).

The first stage of distributed systems creation started at JINR in 1962, when the simple structure of two connected by cable computers M-20 and KIEV in JINR Computing Centre could collect (via cable also) and process information from spectrum analyzer, located in measuring centre of Neutron Physics Laboratory shown in Figure 1. They also processed the punched paper tapes with information, prepared with use of special measuring devices, where the events, which happened in accelerator detectors and automatically photographed (in High Energy Physics Laboratory), were preliminary processed by operators. In 1965, this structure was modernized; they mounted an additional M-20 and they replaced the KIEV, which

was not reliable, by the MINSK-22. They also used this structure for batch processing jobs in other fields of research. Its schema looked like the graphical presentation shown in Figure 2.

1962-1965. The first initial steps, when at least two connected computers became responsible for experimental data handling and other calculations
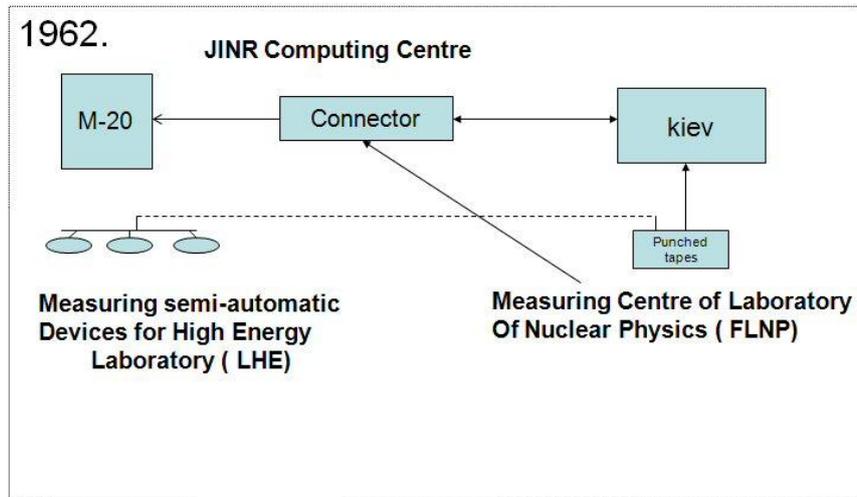

**Figure 1.** Initial version of the JINR Grid
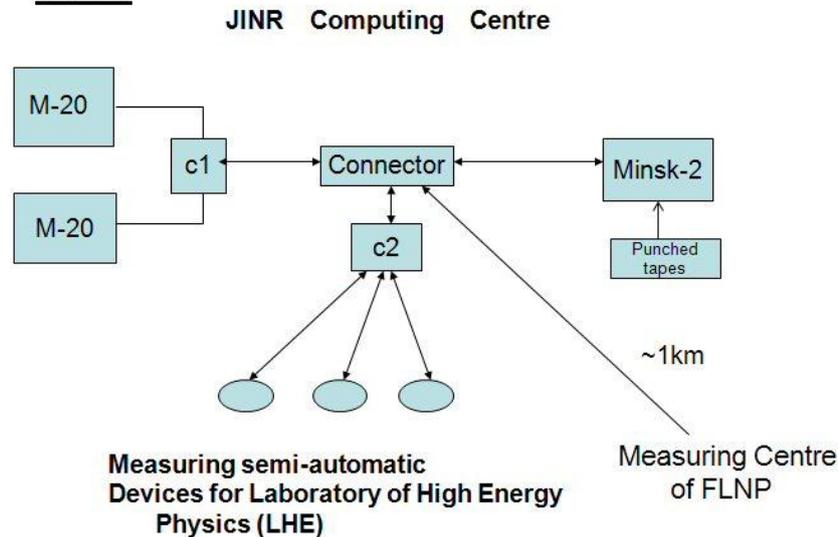
Figure 2. Improved version of the JINR Grid

3. Late 1960s and Early 1970s Activities

From 1966 to 1974 saw the realization of the first stage of the three-leveled complex according to general principles and conditions (in addition to typical modern, up-to-date distributed systems for physics). The principles provide:

1. Meeting the requirements of online experiments;
2. Remote batch and interactive processing for jobs;
3. Interactive mode for users and their terminals;
4. Algorithm parallelism with use of different computing elements in complex;
5. Maintenance of archives for programs and data.

The realization of this set of services was defined at JINR by the appearance of a sizeable set of new computers. These included:

- o "small" computers (first level) for on-line experiments (equipment control and data registration) or remote input/output stations (RIOS), used for batch processing support;
- o "middle" computers (second level) for preliminary processing of data, taken from first level, and also for RIOS regime support; and
- o "large" computers (third level) as basic computing servers.

This new set included some TPA (Hungarian production) computers, SM-3/4, M-6000 (analog HP-2116B), original HP-2116Bs, BESM-3M and BESM-4, MINSK-

2/22, CDC-1604A, and the BESM-6. The BESM-6 started to be used as a basic computing server not only at JINR, but also in all noted scientific Soviet organizations. In fact, it was necessary to develop a preliminary software system. This is the reason the" DUBNA" appeared as a multi-language system with FORTRAN and other compilers and a programming library that was compatible with CERN library [2,3,4].

Another problem concerned the necessity to develop technical and corresponding software tools in BESM-6 operating system (OS) for connectivity with various local external devices and other computers of first and second levels. As a result, at JINR they modernized the BESM-6's peripheral control unit and new version of the operating system started to serve eight fast cabled lines. The lines had speeds up to 500 Kbyte/sec per line and up to 4 abonents connected to each line. In particular, they deployed remote data processing computers and RIOSes (also called "FORTRAN Stations" or FS) equipped with their own peripheral devices such as magnetic tape units, card readers, printers and terminals.

As far as the "middle" computers in laboratories are concerned, measuring centers had to participate interactively for information processing together with the BESM-6 with special "algorithm parallelism" tools. This is the reason to include a special "extracode" facility to serve external links embedded in the OS of the BESM-6. Additionally, they added subroutines to the DUBNA monitoring system that used this extracode useful in calling FORTRAN-written programs. Use of external interruption apparatus in the BESM-6 and peripheral computers (such as the BESM-4 that also used an analog of the FORTRAN-oriented monitoring system) helped to realize an analog of a modern Message Passing Interface (MPI) service. The schema of the complex in 1974 appears in Figure 3, where the CDC-6200 was added in 1972 as a BESM-6 partner on level 3 of the whole complex.
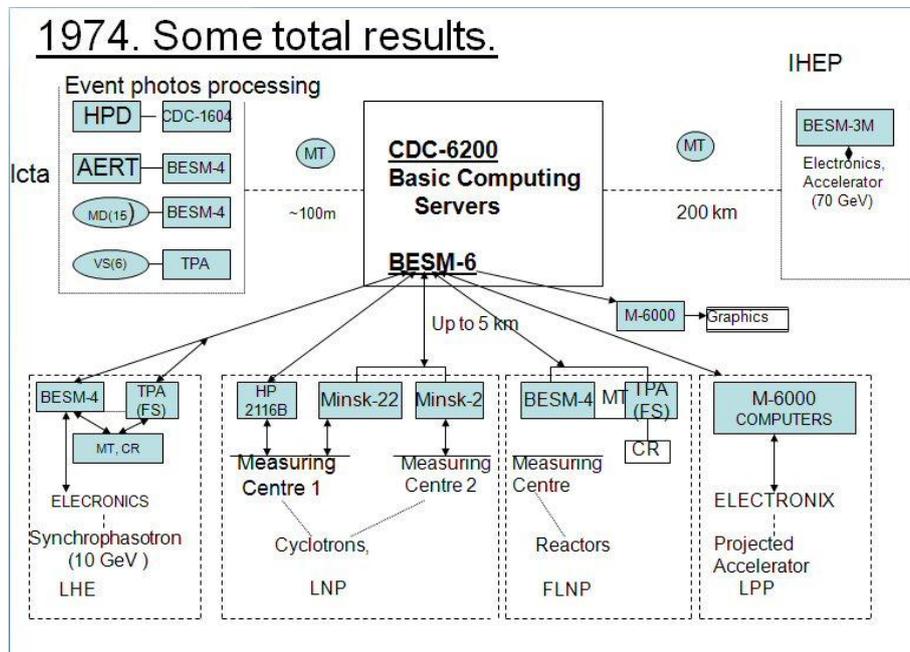
Figure 3. Use of "middle" computers in the 1970s.

The JINR was not unique in Russia, where they already had created multicomputer complexes in the 1960s. It is necessary to mention other systems in use. These include the AIST in the Computing Centre (Siberia), the complexes at the Institute of Applied Mathematics in the Russian Academy of Sciences (IAM/RAS), the Computing Centre at the Institute of Theoretical and Experimental Physics (ITEP/RAS), the Institute of Precise Mechanics and Computing Techniques (IPMaCT) in Moscow; and the Institute of High Energy Physics (IHEP) in Protvino.

In principle, the complex systems at JINR and the aforementioned systems had some similar features. For example, at CERN in system FOCUS, there existed concentrator (CDC-3200) for lines to servers CDC-6500/6400 (four cabled lines, up to 1Mbyte/sec and four peripheral computing or measuring abonents per line), RIOS stations, a terminal concentrator, and other peripherals. However, life was easier for Western scientific centers because there was no need for them to create for themselves the software environment as mentioned above. Regarding the type of computing servers, physicists in West always preferred to use 60-bit processor techniques for high-precise complicated calculations. This is why they used IBM computers for data registration and CDC serial 6000/7000 for processing and other calculations. Such computers were under embargo in the 1960s, but even the 48-bit BESM-6 was good enough.

It is necessary to mention that Grid started with the use of 32-bit technique in clusters, but the appearance of the 64-bit AMD and Itanium 2 processors immediately initiated the use of new PC-clusters in scientific centers. Hence, already by 2003, for the CERN "Openlab" project [5], several major physics software packages have been

successfully ported on the 64-bit environment. For example, the Parallel ROOT Facility (PROOF) is a version of the popular CERN developed software ROOT for interactive analysis of large data files. In Russia, approximately at the same time, the Institute of Computing Mathematics in Moscow started to use a similar cluster with Myrinet 2000 bus for simulations, which earlier were based on use of access to the RAS Supercomputer Centre.

4. Late 1970s and Early 1980s Activities

In 1975 to the beginning of the 1980s at JINR, saw the installation of new ES-seria servers that were compatible with IBM-360/370; this led to the development of tools for terminal access to level 3 in computer complexes as shown in Figure 4.
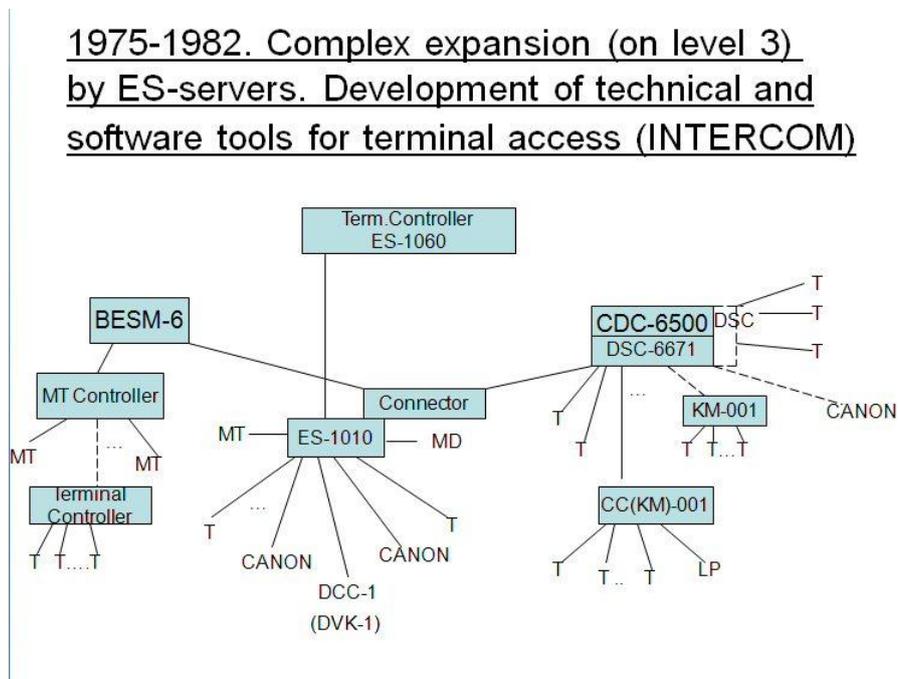
Figure 4. Complex expansion using ES-servers

This development included creation of special controllers (multiplexers) for "large" ES-computers with the use of microprocessors. At JINR, they also developed the terminal concentrator based on the ES-1010 and a programming subsystem called INTERCOM. The system had the following features:
   o Unified language (INTERCOM), which was originally used only for terminal access to CDC seria 6000) and tools for access to different computing servers via terminal concentrator ES-1010, equipped by its own

external memory on disks and tapes and by a special operating system developed in JINR Laboratory of Computing Technique and Automation, former JINR Computing Centre. Alternatively, it used the ES-1060/1061/1066 microprogramed multiplexers. These ES-servers started to understand INTERCOM language by means of a process developed in LCTA software subsystem TERM.

- Independent (of BESM-6, CDC-6000) service for terminals in the process of information receiving, editing, and archiving, when users prepared their job texts.
- Execution of user directives (commands) to send a prepared job to a certain server (after automatically using the JDL-file edition) to show the status of jobs queues, to stop temporarily or kill a job in the server, or to print an output file on printer or terminal screen.

In principle, the concentrator could, in essence, execute functions, similar to functions of program services UI, NS and Resource Broker (RB) in modern Grid systems, especially if the user used a FORTRAN subset understood by all computing servers such as the BESM-6, the CDC, or the ES. The complex had the possibility to choose a server (this or that) simplified by the existence of compatible versions of library CERNLIB in the server's software.

5. Activities in the 1980s

The period 1983 to 1985 marked the starting point for the introduction of a standard means for the creation of local networks in scientific organizations and their external channels. Regarding local network structure and technology, the possible choices were the Token Ring/Token Passing or the then "young" Ethernet technologies. At JINR, they devised a local network in 1985. This was the JINET (Joint Institute NETwork) with a marker (token) method of access to a communication line of 12 km of 75-Ohm coaxial cable [4]. The standard adapter nodes were connected to it in various JINR buildings; each node served eight to twelve serial (standard RS-232C) ports for terminals or computer multiplexer ports. About thirty adapters were bought from the Swiss firm Furrer-Gloor; they had built-in Z-80 microprocessors and built-in memories.

Afterwards, the engineers at JINR (our LCTA engineers) produced a set of their modernized analogs and they used them in the frames of JINET. The LCTA programmers at JINR completely produced the supporting software for all adapters, because the firm's software was too expensive. Moreover, we foresaw the possible modernizations and wished to have not only built-in binary software, but also source texts. This software supported virtual communication channels (communication sessions) according to a user's application through a common cable with a throughput of about 1 Mbit/sec.

Each abonent (terminal user or computer) could give various commands to his connection node such as calling the text editor in his node, checking status of different nodes and their ports, setting the suitable transmission speed at his port or defining the terminating symbols for messages (end of text line, of package if the defaults were

inadequate. If no specification was necessary, the abonent simply commanded to establish a virtual channel for communication with the given address (e.g. node and port number or the other computer's symbolic name), started to use this channel and disconnected at the end of session. Abonents were like computers and they could exchange files using the standard KERMIT program. A special node-administrator or his "hot reserve" partner automatically controlled the marker's (token) movements, collected statistics of nodes connections and faults, recreated the lost markers, disconnected every other node in case of noticed frequent malfunctions.

In 1988, JINET became the abonent of international X.25 standard IASNET, being connected to its communication center in Moscow [4]. This permitted to exchange files directly with physical centers, to use the telnet regime for access to certain servers; many JINR physicists had passwords for such access since they were members of different international collaborations. By the end of 1989, the standard JINR Ethernet was devised and connected to JINET as shown in Figure 5.
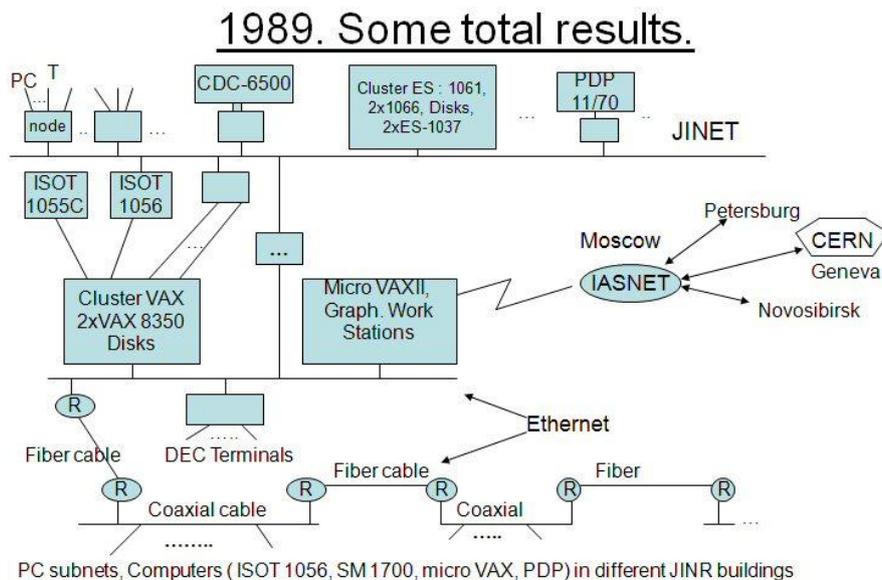
**Figure 5.** Use of Ethernet connections at the end of the 1980s.

It was time to dismount the BESM-6 and connect new servers to the Ethernet. A cluster of two VAX-8350 that contained shared disk memory (about 12 Gbyte), cartridge units (200 Mbyte), and high density (up to 6250 bit/inch) magnetic tape units. Additionally, the system contained the powerful MEGATEK/WIZZARD 7555 station with three-dimensional graphics.

5. Activities during the 1990s

At the very beginning of the 1990s, the throughput of the Russian terrestrial external channels was inadequate for territorially distributed information exchange and processing in high-energy physics, which had practical limits of 32K bits/sec. This is why two direct satellite channels were organized at JINR and started to be exploited in 1994 for access to the DFN (Germany) networks and the High Energy Physics NETwork (HEPNET) with an entry point in Italy and a terrestrial channel to CERN. The throughput of each channel was up to 64K bits/sec; we used our own antenna and the Soviet "RADUGA" satellite for first channel and the terrestrial link to a nearby Cosmic Communication Station "DUBNA" and international satellite INTELSAT for second channel. We had a plan to give some other Soviet scientific institutes such as the IHEP in Protvino the possibility to share this second channel with JINR. It description appears in [2] and shown in Figure 6. Of course, terrestrial channels, with the same throughput as satellite channels, have an advantage in many cases and when new facilities to increase their throughput appeared (up to 1 Gbit/sec nowadays), we had taken the opportunity and ended the use of satellite channels.
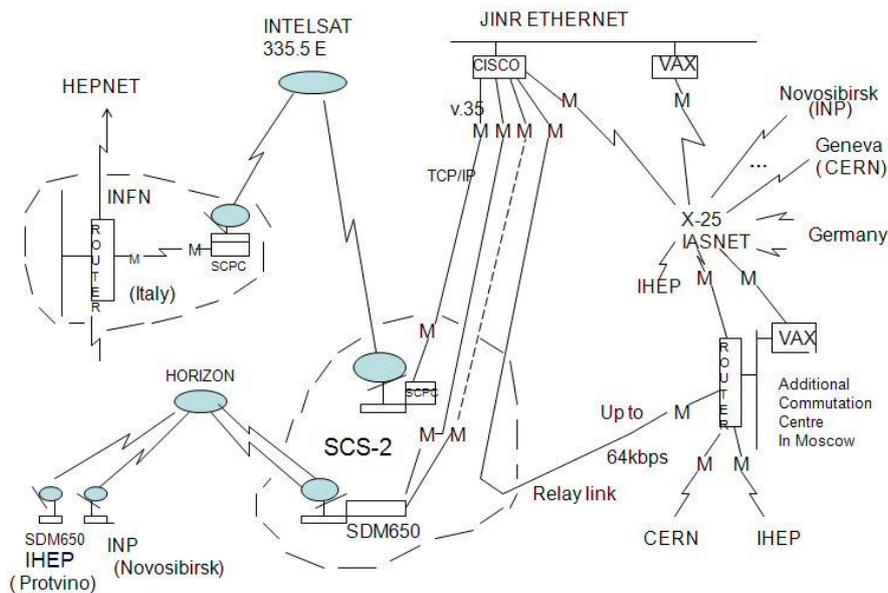
Figure 6. Shared resources during the 1990s.

6. Conclusion

In conclusion, we have seen some close perspectives in the use of distributed systems by a community of scientific organizations. They are, in particular, the final realization of an international project called the Worldwide LHC Computing Grid (WLCG) [6]. Eight institutions in the Russian Federation (in Moscow, Moscow district, St. Petersburg) became the participants in this project. The JINR Scientific Research Institute for Nuclear Physics of Moscow State University (SRINPofMSU), ITEP, IHEP, IAM, Research Scientific Centre of Kurchatov`s Institute (RSC KI), Petersburg`s Institute of Nuclear Physics (PINP) and the Institute of Mathematical Problems in Biology (IMPB) as institutes of Russian Academy of Science …

Acknowledgments

References

1. Ian Foster. What is the Grid? A Three Point Checklist. In:
   http://www.gridtoday.com/02/0722/100136.html
2. V.P.Shirikov. Matematicheskoye obespechenie vychislitel`nyh kompleksov i setei.[The Mathematical Support of Computation Complexes and Networks]. Moscow: Nauka. In: Programmirovaniye, Number 3, 1991.
3. Igor A.Apokin. The Development of Electronic Computers in the USSR. Germany: Vieweg. In: Computing in Russia. The History of Computer Devices and Information Technology revealed, 2001.
4. V.P.Shirikov. Scientific Computer Networks in the Soviet Union. Germany: Vieweg. In: Computing in Russia. The History of Computer Devices and Information Technology revealed, 2001.
5. The CERN openlab: a novel testbed for the Grid. In: CERN COURIER, October 2003.
6. Memorandum of Understanding for Collaboration in the Development and Exploitation of the Worldwide LHC Computing Grid. In: http://lcg.web.cern.ch/LCG/C-RRB/MoU/WLCGMoU.pdf