Information and Computing Infrastructure of JINR (05-6-1118-2014/2019)

(Written report)

Leader: V.V. Korenkov – korenkov@jinr.ru

Deputy: T.A. Strizh – strizh@jinr.ru

April, 2019

CONTENTS

Introduction

The LIT activity within the theme 05-6-1118-2014/2019 "Information and Computing Infrastructure of JINR" was carried out in frames of the direction "Networking, Computing, Computational Physics". The main emphasis was on further development of the JINR network, information and computing infrastructure for the research and production activities of the Institute and its Member States on the basis of state-of-the-art information technologies in accordance with the JINR Seven-Year Plan of development. The major direction within the theme is the development of the JINR LIT Multifunctional Information and Computing Complex (MICC) presented as **the Project**. The JINR MICC consisting of four key components, namely, the grid infrastructure, the central information and computing complex, the computing cloud and the high-performance platform HybriLIT with the "GOVORUN" supercomputer, ensures the implementation of a whole range of competitive research conducted at the world level at JINR in the experiments: MPD, BM@N, Alice, ATLAS, CMS, NOvA, BESIII, STAR, COMPASS and others. The MICC includes the Tier-1 level grid center, which is the only one in the JINR Member States and one of the seven world data storage and processing centers of the CMS experiment (CERN). The JINR Tier-1 and Tier-2 grid sites are elements of the global grid infrastructure used in the WLCG project for processing data from the LHC experiments and other grid applications. The MICC project is aimed at:

1. Development and improvement of the JINR telecommunication and network infrastructure: extension of the bandwidth of the main JINR-Moscow communication link; creation of backup data channels; organization of dedicated lines of data transfer connecting the data centers of organizations in frames of international projects conducted by JINR scientists both at the JINR basic facilities and in cooperating countries; modernization and maintenance of the JINR local area network to provide data storage and processing resources for the NICA project.

2. Modernization of the MICC engineering infrastructure: new power supply systems including uninterruptible power supplies (UPS) and a diesel generator 3 unit (DGU); enhancement of the cooling and conditioning system of both the servers and the server room; the fire suppression system; the system of monitoring of all the infrastructure elements.

3. Modernization, development and creation of the MICC new components for data storage, processing and analysis. Five key components are the following:

    a. NICA project IT-infrastructure foreseeing the creation of a robust storage system for experimental data (BM@N, MPD, SPD), a reliable and effective system for experimental data processing and analysis;

    b. extension of capacity and data storage systems of the MICC basic grid component - the CMS Tier-1 center in frames of the RDMS CMS collaboration and the CMS experiment as a whole;

    c. upgrade of outdated computing resources and data storage systems being part of the Tier-2/CICC integral component which provides support for the experiments at LHC (ATLAS, Alice, CMS), FAIR (CBM, PANDA) and other large-scale experiments using the grid environment and cooperating with physical groups in JINR as well as for non-grid JINR users and its Member States (MPD/NICA, BESIII, LRB, FLNR, DLNP, BLTP, LNP);

    d. extension of the cloud component in order to enlarge a range of services provided to users as well as to create an integrated cloud environment for experiments of JINR (NICA, ALICE, BESIII, NOvA, Daya Bay, JUNO, etc.) and its Member States;

    e. enlargement of the heterogeneous computing cluster HybriLIT as a major resource for high-performance computing.

It is noteworthy that within the report period the "GOVORUN" supercomputer being a natural development of the heterogeneous cluster HybriLIT was put into operation. The creation of the supercomputer was not initially implied in the MICC project and the JINR Seven-Year Plan of development, however, it was the result of requests from scientific groups of JINR Laboratories, primarily BLTP, with the aim of cardinally accelerating complex theoretical and experimental studies in nuclear physics and condensed matter physics underway at the Institute as well as for the development of computing for the NICA megaproject. It became possible due to the significant support of the JINR Directorate. The introduction of the "GOVORUN" supercomputer into MICC was based on the accumulated successful experience in operating the HybriLIT cluster, which supports novel computing architectures, as well as on the enormous work of the heterogeneous computing team on the implementation of parallel calculations to solve difficult tasks facing the Institute and train specialists in high-performance computing.

The other key objectives of the theme are:

— **Information and software environment for the research and production activities of JINR.**

One of the major goals of the theme is implementation and development of the JINR corporate information system (CIS). CIS consists of the following services: the system 1C: Manufacturing Enterprise Management (1C: MEM) is aimed at providing a solution to the problems of accounting and management accounting, payroll, personnel records; the corporate system of project management (PMIS (APT EVM for NICA)) allows one to carry out high-level monitoring and management of JINR large-scale projects such as the NICA project in VBLHEP; the information search system (ISS) provides various reports on personnel and financial information; the system of electronic signing, storage and search of documents of the JINR main office administration as well as documents of MES&CC (Management of Economic Services and Capital Construction) and PLS (Procurement and Logistics Service) (Document Base); the institutional open access repository of articles, preprints and other documents (JINR Document Server (JDS)) reflecting and contributing to research activities underway at JINR; the information system of scientific event management Indico provides the full cycle management of conferences, seminars, workshops, meetings, lectures, etc.; the information system for storing and managing data on the results of research activities of JINR employees (Personal Information System (PIN)); the system of management accounting (ADB2); the JINR electronic document system (JINR EDS). The development of this direction is aimed at simplifying and optimizing scientific and administrative activities of the Institute.

— **Development of the system for training and retraining IT-specialists on the basis of the JINR MICC and its educational components.**

This section of the theme is aimed at creating a system for training and retraining IT-specialists based on the MICC cloud, heterogeneous and grid infrastructures. In particular, the HybriLIT platform is used not only for parallel calculations on novel computing architectures, but also as a polygon for training students, post-graduate students and young scientists.

In the framework of this theme section, the following training courses are held:

- regular courses on modern IT-technologies both for the Institute staff and students and young scientists from the JINR Member States in frames of practices organized by the UC (JINR University Centre);
- special courses from leading software developers;
- special courses and seminars in frames of conferences and schools organized by JINR;
- special courses organized in the JINR Member States in frames of international cooperation programs.

The degree of interest towards the LIT activities in the JINR Member States is high. We have a number of collaborators from the JINR Member States: protocols of cooperation with INRNE (Bulgaria), FZK Karlsruhe GmbH (Germany), IHEPI TSU (Georgia), NC PHEP BSU (Belarus), IFIN-HH (Romania), etc.

Project MICC

The JINR Multifunctional Information and Computing Complex meets the requirements for a modern highly performant scientific computing complex: multi-functionality, high performance, task adapted data storage system, high reliability and availability, information security, scalability, customized software environment for different existing user groups, high-performance telecommunications and modern local network. During the project realization in the 2016-2019 time period, the focus was on the development and modernization of the JINR telecommunication and network infrastructure including the modernization of the JINR local area network to provide data storage and processing resources for the NICA project as well as on the modernization of the MICC engineering infrastructure including and uninterruptible power supplies systems, the conditioning and ventilating system. Special attention was paid to the creation of the NICA project IT-infrastructure including both a long-term storage system for experimental data (BM@N, MPD, SPD) and a reliable and effective system for off-line data processing. During the report period, the MICC monitoring systems were also actively developed allowing to get information from different levels of the computing complex, namely, engineering infrastructure, network, computing nodes, task launching systems, data storage elements, grid services, etc.; all the enumerated above will ensure a high level of the MICC reliability.

Furthermore, within the report period the performance and data storage systems of the MICC basic grid component, i.e. the JINR CMS Tier-1 center, were extended as planned, which ensured the second place in the world for the JINR Tier-1 among other Tier-1 sites by the number of processed events.

The JINR Tier-2 level center was actively developed; it provides the processing of data from four LHC experiments (Alice, ATLAS, CMS, LHCb) and, moreover, supports a whole range of non-LHC virtual organizations (VOs) (BES, BIOMED, COMPASS, MPD, NOvA, STAR, ILC). The MICC also provides calculations including parallel ones outside the grid environment, which is essential for the experiments such as NOVA, PANDA, BES, NICA/MPD and others as well as for local users from JINR Laboratories.

During the report period, work on the integration of cloud structures of the JINR Member States was carried out as well as the maintenance of common IT-services and computing infrastructures was provided, training courses were held, the deployment and support of additional IT-services were ensured upon request in frames of the JINR neutrino program. In particular, the NOvA experiment, which was the first to actively use the LIT cloud resources, now has the largest number of allocated resources, the Baikal-GVD and JUNO experiments follow it. Due to the DLNP neutrino program, the number of computing cores in the cloud infrastructure increased by two times.

The "GOVORUN" supercomputer commissioned in the impressively short period in 2018 enabled to significantly increase the number of high-performance calculations and investigations requiring resource-intensive computations on the heterogeneous computing paradigm at JINR, especially in the field of quantum chromodynamics on grids. At present, the supercomputer is actively used by various JINR user groups; since July 2018, within the operation period, over 100 thousand tasks have been performed on the supercomputer. During the report period, appreciation for using computing resources of the HybriLIT heterogeneous platform including the "GOVORUN" supercomputer was expressed in over 60 publications.

In the frame of the MICC project, work on creating computing for the NICA megaproject at all its stages was carried out – from software modeling to creating the prototype of the data storage and processing complex including the use of the "GOVORUN" supercomputer.

One can find below more detailed information on the results obtained during the implementation of the MICC project in the 2016-2019 time period.

Networking

JINR telecommunication data links

Work dedicated to the reliable operation and development of the JINR network and information infrastructure was performed. The key components of the infrastructure are telecommunication data links, the JINR local area network, the network of the MICC/Tier2 and Tier1 sites.

The reliable operation of the Dubna-Moscow high-speed communication channel was ensured. The external JINR channel was built on the DWDM technology and used one 100 Gbit/s lambda and two lambda (two frequencies) 10 Gbit/s each. Optical equipment of Nortel and Transmode was used for the construction of the channel. The connection with scientific networks and the Internet used the following telecommunication links: LHCOPN/LHCOPN/CERN (100 Gbps), Russian scientific networks and international scientific networks (4 x 10 Gbps), RUNnet (10 Gbps), RETN (6 Gbps). The bandwidth of the reserve data link was raised up to 20 Gbps. The reliability of external channels was improved due to the addition of two supplementary routers Cisco ASR-1006-X.
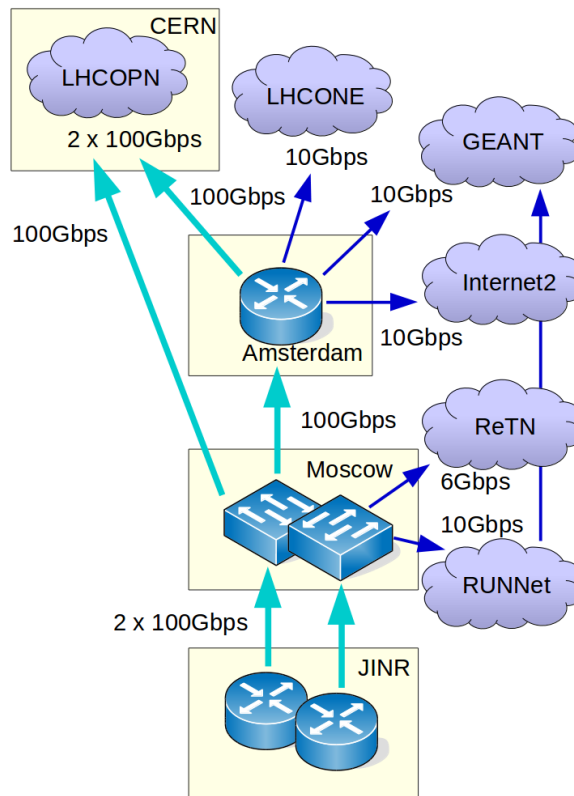
Fig.1. Logical structure of the JINR external network.

The configured DWDM equipment for the organization of an additional fiber-optic channel Dubna-Moscow with a capacity of 100 Gbit/s was installed at the MMTS-9 site in Moscow, which will give a new direct channel to CERN. Thus, the external distributed network of JINR will be represented by the JINR-CERN direct channel and a backup channel passing through MMTS-9 in Moscow and Amsterdam ensuring the operation of LHCOPN (JINR-CERN) for connection between Tier0 (CERN) and Tier1 (JINR) and the external superimposed network LHCONE intended for the JINR Tier2 center, direct channels for connection using the RU-VRF technology with the collaboration of RUHEP research centers and networks Runnet, RASnet (Fig.1).

Table 1 shows the distribution of the ingoing and outgoing traffic (more than 55PB of the incoming traffic) in 2016-2018 among JINR subdivisions.

Table 1.

Subdivision	2016		2017		2018	
	Ingoing (TB)	Outgoing (TB)	Ingoing (TB)	Outgoing (TB)	Ingoing (TB)	Outgoing (TB)
Laboratory of High Energy Physics	102.43	90.68	115.36	89.71	161.99	77.77
Laboratory of Nuclear Problems	125.98	154.99	158.23	82.88	197.80	67.84
Laboratory of Neutron Physics	49.20	63.15	54.38	37.56	76.84	39.77
Laboratory of Information Technologies	62.13	34.79	58.15	18.89	87.55	22.91
General access servers	90.99	17.72	299.05	131.81	392.76	299.80
Laboratory of Radiation Biology	8.76	5.26	7.71	4.39	10.65	3.47
Laboratory of Theoretical Physics	19.71	16.73	22.75	8.68	29.08	7.91
JINR Directorate	26.82	58.20	39.65	74.95	41.31	28.24
Laboratory of Nuclear Reactions	37.60	18.18	41.18	21.10	67.09	11.49
Remote Access Node	24.05	5.74	35.96	4.95	50.10	5.21
JINR's Hotel&Restaurant Complex	25.76	5.90	30.54	4.05	67.72	8.11
Joint Stock Company "Dedal"	10.87	5.97	18.96	3.84	16.86	1.25
Medical Unit-9	10.65	0.82	13.88	1.57	19.92	1.14
Dubna University	48.98	38.43	79.08	57.35	17.69	20.02

In 2018 the overall incoming JINR traffic including general access servers, Tier-1, Tier-2 and the MICC amounted 18.3 PB (23.3 PB in 2017). The weights of various incoming traffic categories are shown in Table 2.

Table 2.

Scientific and educational networks	File exchange (p2p)	Web-resource	Social networks	Software	Multimedia
91.96%	4.75 %	2.79 %	0.49%	0.03%	0.01%

The weights of the traffic by top 10 scientific organizations are shown in Table 3 (year 2018).

Table 3.

Institute	To JINR (IN)	From JINR (OUT)	% IN	% OUT
Istituto Nazionale di Fisica Nucleare (INFN)	501.5 Tb	544.55 Tb	13.24 %	12.07 %
Fermi National Accelerator Laboratory (FNAL)	403.95 Tb	101.34 Tb	10.67 %	2.25 %
Institute for High Energy Physics (IHEP, Protvino)	346.46 Tb	83.98 Tb	9.15 %	1.86 %
European Organization for Nuclear Research (CERN))	292.5 Tb	152.18 Tb	7.72 %	3.37 %
Forschungszentrum Karlsruhe (KIT)	192.89 Tb	235.19 Tb	5.09 %	5.21 %
Institut National de Physique Nucleaire (IN2P3)	174.33 Tb	228.08 Tb	4.6 %	5.06 %
Rutherford Appleton Laboratory (RAL)	155.42 Tb	281.92 Tb	4.1 %	6.25 %
Deutsches Elektronen-Synchrotron (DESY)	114.2 Tb	175.65 Tb	3.02 %	3.89 %

Institute	To JINR (IN)	From JINR (OUT)	% IN	% OUT
California Institute of Technology, Pasadena (US)	111.98 Tb	73.34 Tb	2.96 %	1.63 %
Imperial College London (GB)	102.27 Tb	98.42 Tb	2.7 %	2.18 %

JINR local area network (LAN)

Work on the development and improvement of network components of the JINR IT-structure designed to increase its efficiency was in progress. Work on the introduction of the 100GB network covered approximately 10% of the network equipment of JINR subdivisions.

Currently, the main optical data transport medium (backbone) operates at 100 Gbps. It consists of three optical rings, which are based on two network switches of the third level (L3) according to the reference network model ISO OSI (International Standard Organization Open System Interface). As switches in the core of the network structure used two (2) Cisco NCS-5011 devices and ten (10) Cisco Catalyst 9500-32C in eight (8) research laboratories, management of JINR and MICC. Three fiber optic rings cover spaced over the JINR campus - sites DLNP, VBLHEP and "City". The JINR local area network (LAN) was transferred to DHCP (Dynamic Host Configuration Protocol). At present, the JINR LAN comprises 8008 network elements and 14154 IP-addresses, 7715 users registered within the network, 2677 mail.jinr.ru service users as well as 1597 digital library users and 396 remote VPN and EDUROAM users. A project on a new link using 4x100 Gbps between the territories of DLNP and VBLHEP with a double redundancy to improve the reliability of the optical transport highway was elaborated. Figure 2 shows the logical structure of the transport channel DLNP & VBLHEP.

The JINR network like any other network functions properly due to the established system of network services. These services include DNS (Domain Name System), DHCP (Dynamic Host Configuration Protocol), SMTP (Simple Mail Transfer Protocol), SNMP (Simple Network Management Protocol), SSO (Single Sign-On), registration (user), authorization (devices), authentication (users), switching, routing, security, video conference communication, VoIP (Voice over Internet Protocol), IPDB (Internet Protocol Data Base), WebMail, etc. Many of these system-wide services such as DNS, DHCP, SMTP, SNMP, switching (STP – Spanning Tree Protocol, VLAN –Virtual Local Area Network), routing (EIGRP Enhanced Interior Gateway Routing Protocol BGP – Border Gateway Protocol), are working and have to work in a 7x24 mode without interruption. At the same time, other services perform upon request video conference communication, VoIP, access to resources of digital libraries, some special applications, i.e. a user initiates a session using the service.

Network security is a layered structure, which is implemented on three (3) levels:

1) on the external network - by means of edge routers;

2) at the local network core – by the equipment for all servers;

3) on the user (terminal) level - by means of operating systems, firewalls, work stations as well as specialized antiviral agents, in particular, NOD32.

On the border (edge) routers (in case of JINR the company Cisco Systems is a vendor) by IOS (Internetwork Operating System) mechanisms traffic can be filtered by many criteria: by IP addresses, ports, traffic type (content), etc.

During the report period, a reconfiguration of the JINR network was carried out systematically; its goal was the improvement of quality and availability of the entire JINR IT-infrastructure. New data transfer protocols, in particular, the latest version of the IPv6 IP-protocol, was implemented. The introduction of new protocols such as MPLS (Multiprotol Label Switching) increased the security of the JINR backbone network against the impact of possible failures of the hardware component of the network structure and meets strict requirements imposed on the work of the Tier1 center by the CMS collaboration, regarding the almost constant availability of the network infrastructure. The main feature of the MPLS technology is the separation of the process of switching packet data and the process of analysis of the IP-address in the packet header, which

reduces the data transmission time. This technology (MPLS) makes it possible to perform services of virtual private networks (VPN) of a new generation. It ensures the quality of transfer over IP-networks, the transfer of delay-sensitive traffic and can become a foundation for effective support of such services as the transmission of real-time voice and video stream in the JINR network. MPLS VPN allows using services such as video telephony, videoconferencing and remote video surveillance. The key VPN service is to provide the security of the virtual network of a user, i.e. a JINR employee, and prevent unauthorized access to the resources of employees' networks. The multiprotocol label switching technology for VPN ensures the same degree of protection as second-level connections, without the complexity of encryption.

The introduction of the IPv6 protocol, which has some advantages over its predecessor, i.e. the IPv4 protocol, made it possible to set the priority of a number of packets of data flows, thereby to accelerate the routing of required flows and increase the speed of data exchange between devices. An important feature of IPv6 is that it incorporates the possibility of data encryption and support for the service quality, the property, which is essential for multimedia applications.

The development of the DHCP service at JINR, being systemwide and centralized, allowed distributing necessary IP-addresses for users only through the mechanisms of this protocol. Thus, the allocation process of IP-addresses became automated avoiding a human factor, which led to decreasing configuration errors and increasing the availability of the work of the entire system.

Within the report period, the e-mail service was improved taking into account the suggestions of the network service team: 1) the service was transferred to two new powerful servers with a large disk memory; 2) a backup copy of e-mail messages received by the Institute staff was made.

To create a secure access mechanism in the network, a system to authorize the WiFi equipment for all categories of users was developed, at the same time the process of self-registration or registration via operator requires the establishment of user attributes (username and password), and at the first positive completion of the authentication process, the system performs the authorization of a user device (detects a hardware address MAC - Media Address Control). In the future, every time the network, taking in legal authentication attributes, will bind them to the hardware (MAC) address of the device. Thus, it is a two-level security control.

At present, the service is JINR-wide and performs a unified centralized policy of WiFi proliferation over the JINR LAN campus. It will provide the Institute guests, the participants of scientific conferences and seminars as well as the JINR staff with a reliable WiFi service, especially for access to the Internet via their own mobile devices.

Engineering Infrastructure

The MICC engineering infrastructure is designed to ensure the reliable, uninterrupted and fault-tolerant operation of the computing subsystem and the network infrastructure, maximally protecting the complex from the effect of negative factors during the equipment operation.

The JINR LIT engineering service uses an integrated approach to building the MICC engineering infrastructure, which allows one to elaborate algorithms of the equipment operation and interaction of separate systems both in a normal operation mode and in emergencies, and ensures the uninterrupted information flow regardless of external factors.

The integrated approach to the design of engineering systems allows the interaction of existing engineering systems at the design stage, which is crucial for the modernization of the outdated equipment in the mode of the uninterrupted work of the MICC computing equipment.

An analysis of the growth prospects of the server infrastructure for 3-4 years is performed by LIT specialists and engineers. The LIT engineering component is designed with the possibility of scaling, based on the analysis. It will allow easily adapting the engineering systems of the LIT building to new requirements and growing power of the IT-equipment in accordance with the Laboratory current needs.

The developed centralized power supply scheme allows solving several tasks:
1. To consolidate all UPS equipment in one room;

2. To free up space for hardware racks in the machine hall;
3. To minimize the total amount of the power supply equipment;
4. To organize a more reliable power supply scheme for MICC consumers with a smaller amount of equipment;
5. To simplify the electrical distribution schemes for all types of load;
6. To increase the manageability of power supply systems and the efficiency of troubleshooting.

In frames of the MICC development plan, the electrical capacity was increased from 1mVA to 2,5 mVA in 2017. In 2016 technical conditions from JINR MED (Main Energy Department) were obtained, and at the end of 2016 the modernization of the JINR LIT power supply system, together with the replacement of the transformers with a nominal value of 1000 kVA with new transformers of 2500kVA (Fig.2-3), started. Therefore, the grounding system was also upgraded.
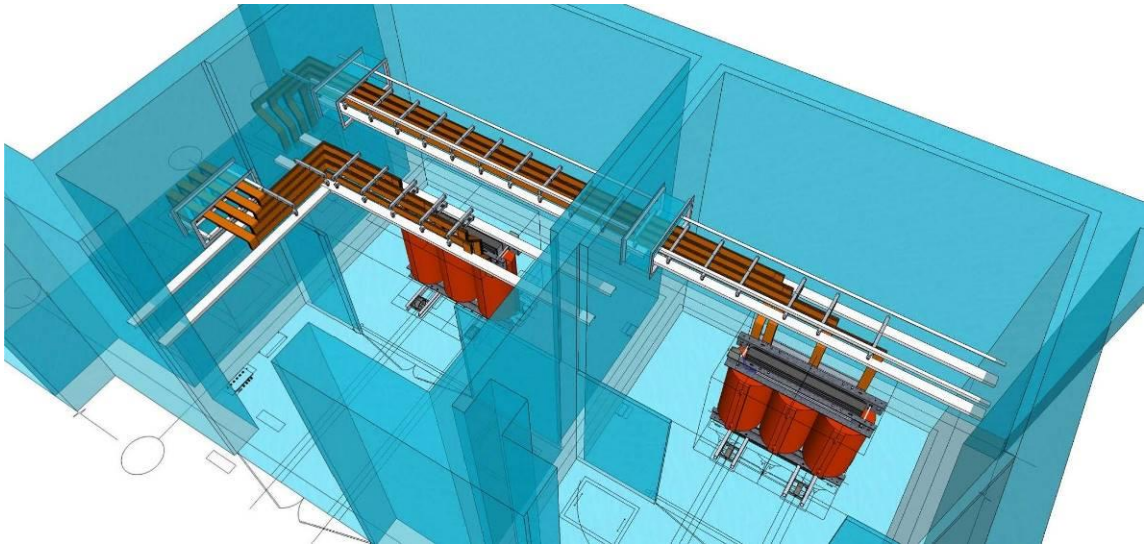
Fig. 2. Architectural view of two transformers with a busline.
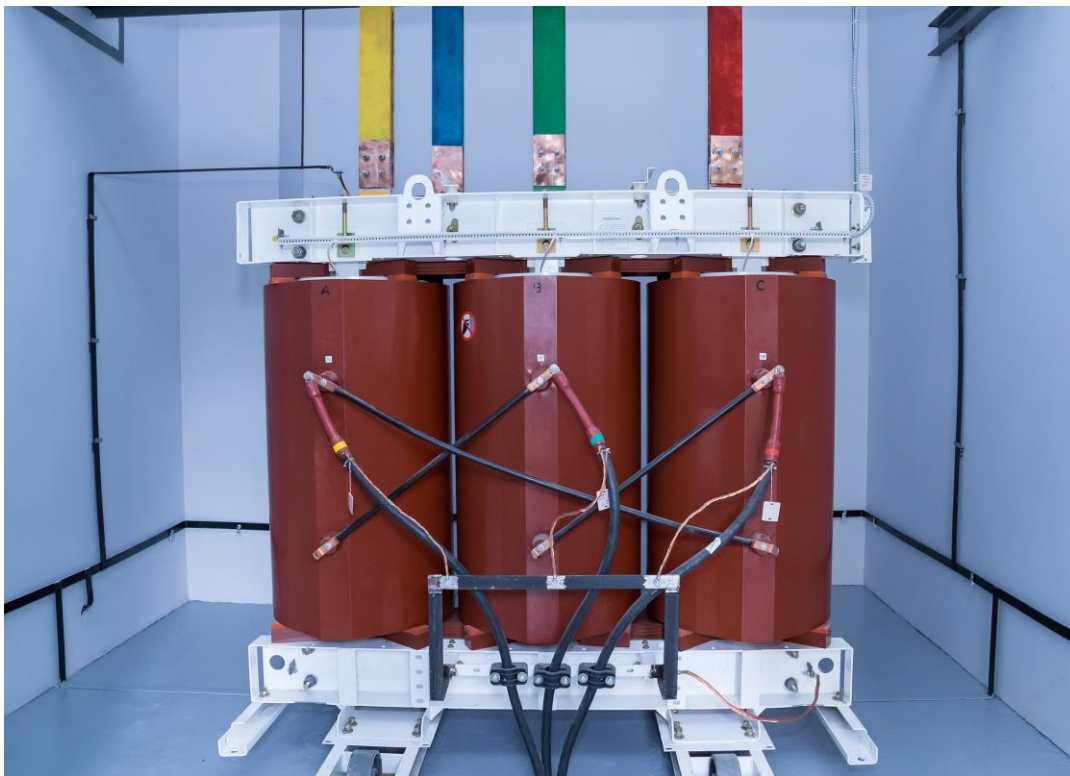
Fig.3. Type of the transformer TSGL 2500 kVA located in the 134 LIT transformer building.

The project of the modernization of the main switchboard (Fig.4) providing power supply of computing systems and the MICC engineering infrastructure, as well as the power distribution system, was implemented simultaneously with the project of external power supply.
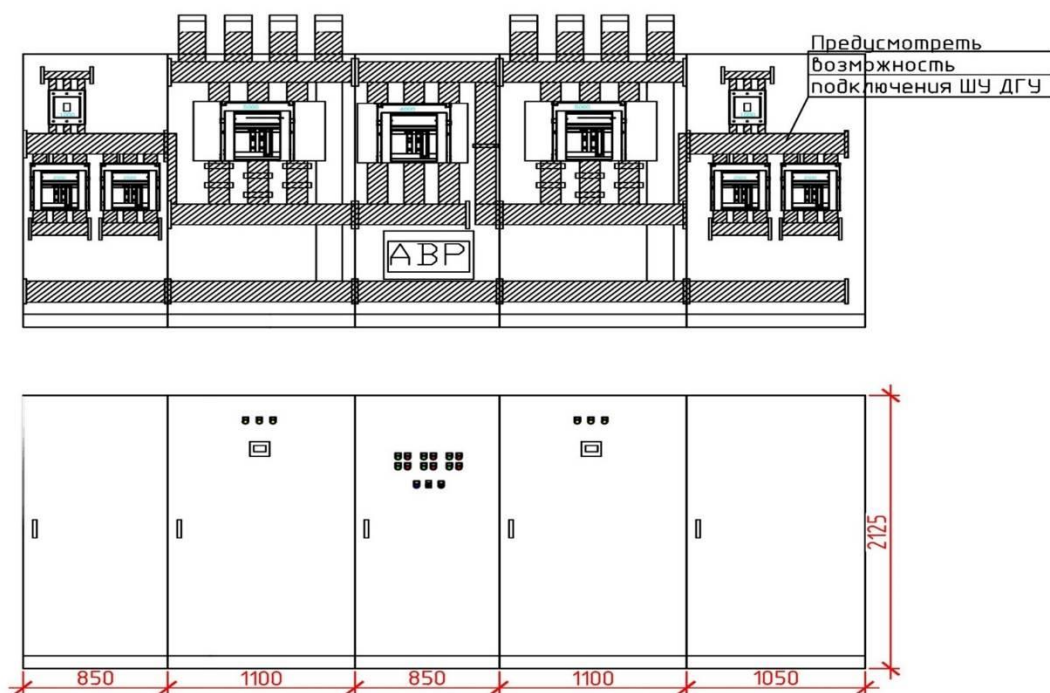
Fig.4. Main switchboard.

In frames of this project, the current power supply scheme was transformed to ensure a secure and uninterrupted power supply of all types of loads from the general electricity buses separating consumers in several categories on the reliability of power supply (Fig.5).

The following groups of users are fed from the main distribution panel:

- Household consumers (own needs of the building, which do not influence the operation of MICC and its engineering infrastructure: the appliance receptacle net, the power for goods elevators, lighting, etc.);
- Technological consumers (own needs of the building which indirectly or partially influence the reliability of MICC and the Institute: the cooling and thermal installation providing a climate control in computer and server halls, the power supply of the passage to DLNP, etc.);
- Consumers of the uninterrupted power supply (UPS) (the consumers of the "A" category are server equipment and pumps, monitoring systems, etc., which are fed from UPS, and the consumers of the "B" category are chillers, chiller machines, a cooling tower installed outside).
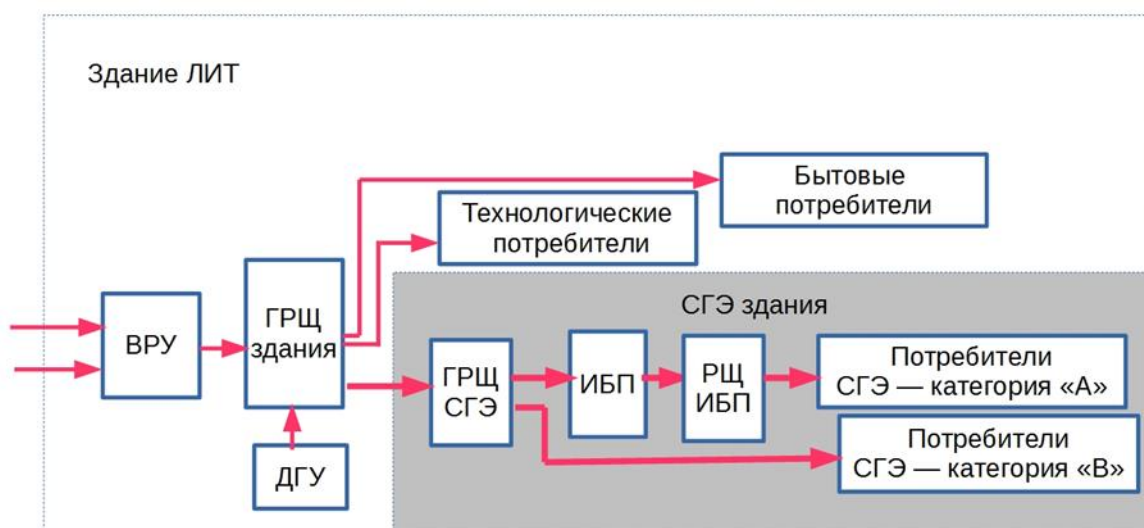
Fig.5. Basic diagram of separating consumers in the LIT building: ES – External switchgear, MS – Main switchboard, DGS – Diesel-generator set, UPS – Uninterruptable power source, S– Switchboard, GPSS – Guaranteed power supply system.

The modernization of the external power supply system allowed proceeding to the next stage, i.e. the creation of the MICC guaranteed power supply system. Two diesel-generator sets (DGS) based on the engine Mitsubishi S12R-PTAA2 and the generator Mecc Alte ECO46-1S/4 in two containers with a total capacity of 3300 kVA (Fig.6) were installed and put into operation. The computing systems with telecommunication equipment as well as the complex engineering infrastructure are supplied by DGSs. The reservation is performed according to the 2N scheme providing the guaranteed generation in a required volume even in case of the failure of one group of generators. DGSs are located close to the building, which allows significantly optimizing the cable infrastructure for their connection to the general power supply of the LIT building.

Fig.6. Location of DGS containers.

DGSs work on the principle of a leading and driven diesel-generator. The given principle means that both DGSs are launched, thereby ensuring the reliability of the MICC engineering infrastructure to the level of the 4th highest level of reliability.

In parallel with the modernization of the power supply system, UPS Riello 160 kVA was purchased and installed; it provides the guaranteed power supply to the Tier2 server nodes (Fig.7).

Fig.7. Riello 160 kVA.

Within the modernization of power supply systems, the means of monitoring the state of the main automatic switches of the power distribution system, as well as the basic parameters of the power supply equipment, were developed and implemented.

MICC air-conditioning system

In order to ensure the reliable 24-hour operation of MICC, it is necessary to provide special conditions for the microclimate, special requirements for its power supply and organize a backup protected fiber optic network in the premises where the server equipment and the data storage systems are located. The given engineering infrastructure should be built on the reliable and trusted equipment and serviced by the highly qualified personnel with experience in its operation.

In order to provide the microclimate, it is necessary to drive out a large amount of hot air released during the operation of computing modules and disk storage units. A multi-stage cooling system is used for this purpose. It includes:

- a system of inter-row cooling of racks;
- special cooling units that provide excess pressure of cold air under the raised floor;
- specialized engineering devices operating from cooling towers;
- exhaust ventilation that ensures the outflow of the remaining hot air out of the hall;
- a specialized engineering infrastructure to provide the highly efficient technology of liquid cooling of the GOVORUN supercomputer with hot water.

The creation of such a multi-stage cooling system is related to the fact that the computer hall was designed in the '70s of the last century and its engineering infrastructure needs an upgrade to meet modern requirements.

Modernization of the air-conditioning and ventilation system

The realization of all tasks of air-conditioning requires a complex approach. One needs to adjust the interaction of separate systems in a normal mode of operation as well as to work out algorithms of the equipment operation in case of emergencies, thus ensuring the guaranteed conditions of the equipment exploitation regardless of external factors. The complex approach to designing the MICC air-conditioning and ventilation systems allows one to maximize the profitability of the project due to energy conservation and cost cutting. The plan presupposes a stage-by-stage implementation of the project on the modernization of the air-conditioning and ventilation system.

The first stage of the modernization of the air-conditioning system is to replace the outdated air conditioners EDPAC by air conditioners STULZ (Fig.10).

13

Fig.10. Precision air-conditioning STULZ.

The outdated equipment in the amount of four air conditioners EDPAC with a capacity of 40 kW each and the CARRIER chiller with a capacity of 200 kW was replaced with four air conditioners StulzASD 622G with a capacity of 60 kW each and two chillers StulzCSI 981G with a capacity of 100 kW each. The operation of condensers of this equipment is possible from water up to 40°C. The work on replacing air condensers of four air conditioners StulzASD 531A with water cooling condensers was carried out as well.

It allowed:

- providing air conditioning and ventilation of the hall at an external temperature of 39°C;

- increasing the cooling capacity in the hall by 80 kW using the same installation location for new Stulz air conditioners as it was before for EDPAC;

- using the reserve of existing equipment for the designed part of the hall if necessary.

Additional opportunities and new requirements to the system are associated with the commissioning of the GOVORUN supercomputer in 2018, which required the creation of a precision liquid cooling system balanced for constant work with a high-temperature coolant (up to 63°C at the entrance to the computer cabinet). In accordance with the equipment placement conditions for JINR, an optimum operation mode of the computing cabinet at a constant coolant temperature of +45°C at the entrance to the computing nodes was chosen.

The work in a "hot water" mode for this solution allowed us to apply a year-round free cooling (24x7x365) using only dry cooling towers, which cool liquid with the help of ambient air on any day of the year, and also to completely get rid of the freon circuit and chillers. As a result, the average annual PUE indicator of the system reflecting the level of energy efficiency is less than 1,03. It means that cooling consumes 3% of all consumed electricity, which is an outstanding result for the HPC industry.

The direct liquid cooling technology of the RSC company provides precision heat removal from the server using a cooling plate, which completely covers the entire element-containing surface of the computing node and in its turn is cooled by liquid (Fig.11).

Fig.11. Supercomputer computing nodes with liquid cooling.

This approach ensures the most complete heat removal from the entire area of the server components, excluding local overheating and air pockets, which prolongs the service life of electronic components, as well as increases the fault tolerance of the solution.

A dry cooling tower was installed for the implementation of such a cooling system (Fig.12). Its task is to provide cooling of the coolant involved in the process. It is achieved due to the following: the fluid supplied to the heat exchanger lowers its temperature under the influence of the airflow taken by the ventilation unit from the environment.

Fig. 12. Installation of a dry cooling towel.

Due to the climate peculiarities, it was decided to use a solution of non-freezing liquid, i.e. ethylene glycol, as a coolant. Ethylene glycol from the cooling tower enters the collector and then the heat exchanger, which absorbs thermal energy from the water circulating through the supercomputer computing nodes.

The water cooled to a temperature of 45 degrees enters the supercomputer. After passing through the entire circuit in the supercomputer, the water heated to 50 degrees returns to the heat exchanger, where it is cooled transferring thermal energy to the hydraulic circuit of the dry cooling tower. The cooling system has a smooth performance tuning, which allows one to increase or decrease the power of the cooling system in accordance with actual load. It ensures a significant reduction of power consumption during partial load. The elements of the supercomputer liquid cooling system are presented in Figure 13.

Fig.13. Elements of the GOVORUN supercomputer liquid cooling system: the collector, the rack with a pump and a heat exchanger.

After the modernization of the cooling supply including the replacement of all freon conditioners of the machine hall with water ones and the introduction of a new system of central cooling and ventilation of the building and machine halls, the fault tolerance of the refrigeration equipment is significantly increased. During the report period, a project of the development of air conditioning, ventilation and cooling supply was elaborated. The possibility of obtaining additional power from the dry cooling tower GUNTNER of the GOVORUN supercomputer system for the operation of conditioners in the hall in case of reserve, as well as for a possible increase of power for the supercomputer in future or in case of a failure of the regular dry cooling tower from designed dry cooling towers for the operation of the second part of the hall, was tested.

JINR grid environment

The JINR grid infrastructure is represented by the Tier-1 center for the CMS experiment at LHC and the Tier-2 center, which supports a whole number of virtual organizations (VOs), particularly Alice, ATLAS, CMS, LHCb, BES, BIOMED, COMPASS, MPD, NOvA, STAR and others.

The data processing system at the JINR CMS Tier-1 consists of 415 64-bit machines: 2 x CPU, 6–16 cores/CPU that forms altogether 9200 cores for batch processing and provides a performance of 132.8 kHS06. The storage system consists of disk arrays and long-term data storage on tapes and is supported by the dCache-3.2 and Enstore 4.2.2 software. The total usable capacity of disk servers is 8.3 PB, the IBM TS3500 tape robot is 11 PB. The evolution of the JINR CMS Tier-1 main resource parameters for the last three years is presented in Figure 14.

T1_RU_JINR Logical CPU

T1_RU_JINR HEPSPEC06

T1_RU_JINR Disk (TB)

T1_RU_JINR Tape (TB)
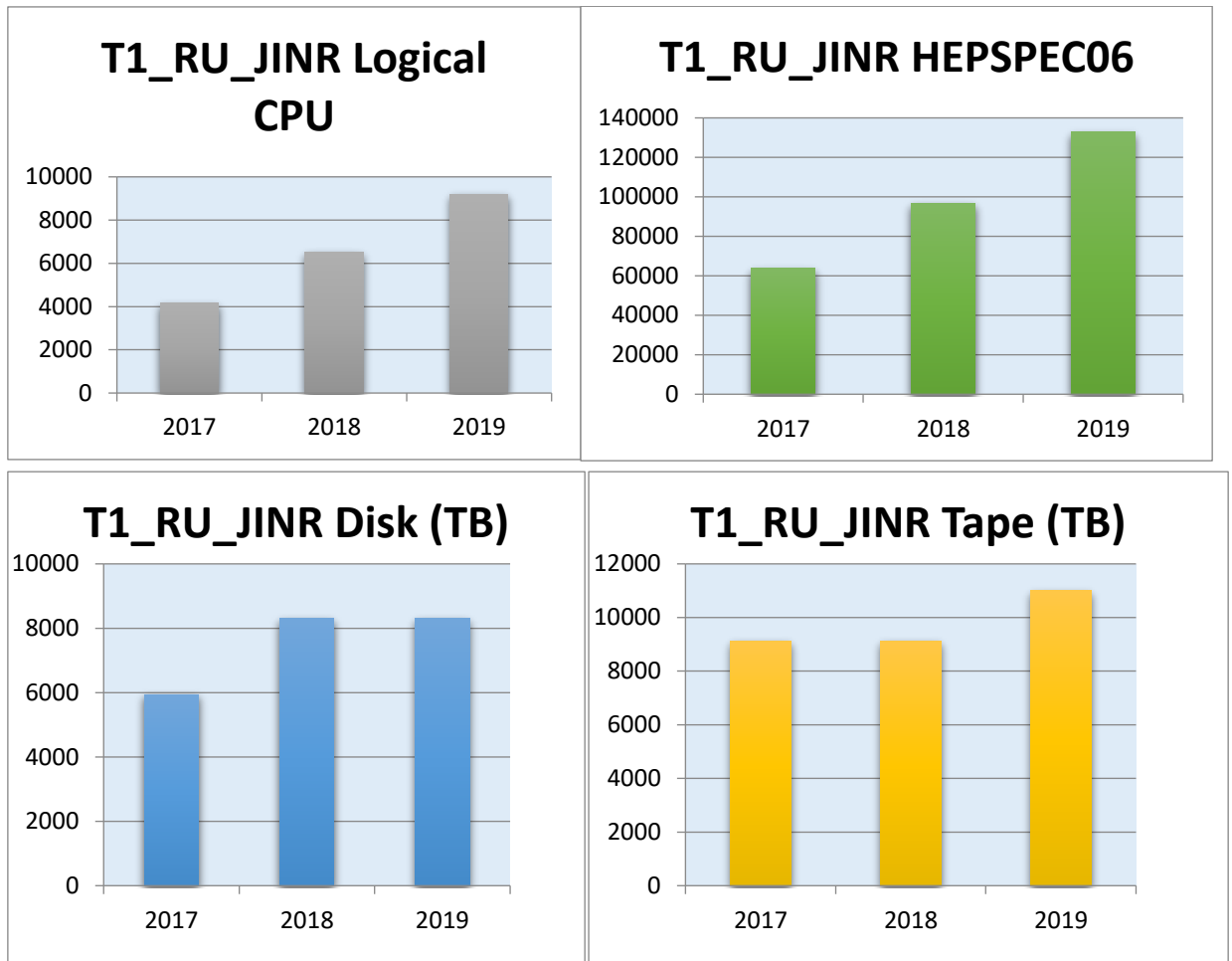
Fig.14. Evolution of the JINR CMS Tier-1 resources.

The Torque 4.2.10/Maui 3.3.2 software (custom build) is used as a resource manager and a task scheduler. The PhEDEx software is used as a tool for management of CMS data placement. The standard WLCG program stack such as 2 x CREAM, 4 x ARGUS, BDII top, BDII site, APEL parsers, APEL publisher, EMI-UI, 220 x EMI-WN + gLExec-wn, 4 x FTS3, LFC, WMS, L & B, glite-proxyrenewal is used for data processing. The service support system provides the operation of a computing service, a data storage service, a grid service, a data transfer service (FTS - File Transfer System), a distributed computing management system (PBS - Portable Batch System), an information service (monitoring of services, servers, storage, data transmission, information sites).

The CMS Tier-1 center at JINR has demonstrated stable work through the entire period since its launch into full operation [5, 7]. Since the beginning of 2017 16 506 640 tasks have been completed, 371 617 million events have been processed which accounts for 19,11% of the total number of CMS experiment events processed by all Tier-1 sites. Figure 15 demonstrates the contribution of the world first level centers to the processing of CMS experimental data starting with 2017. The JINR site is ranked second in its performance in the world.
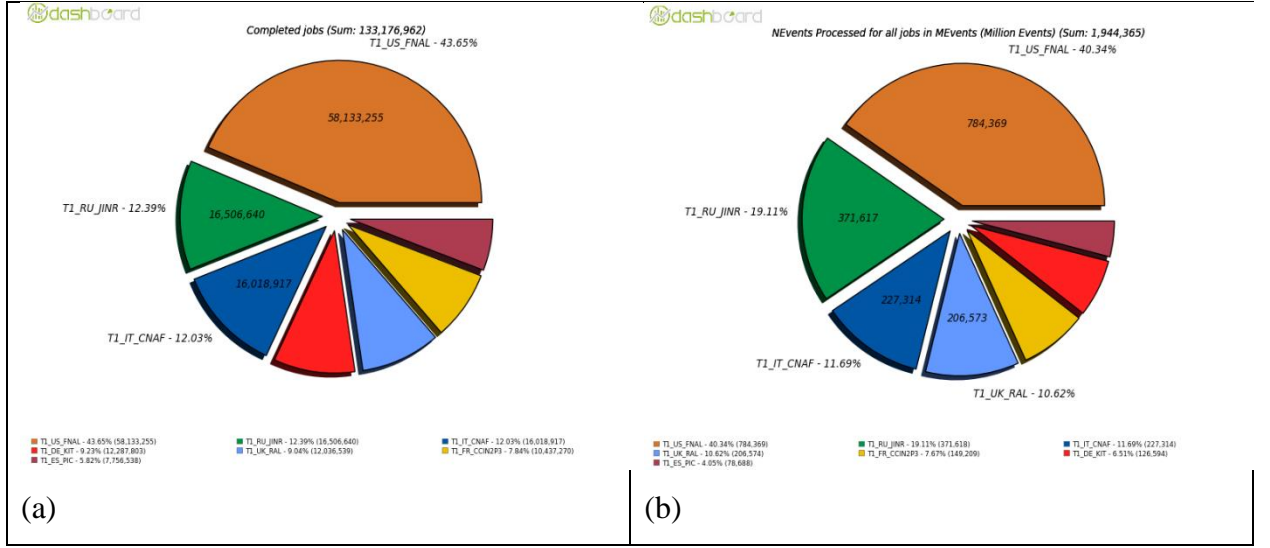
Fig.15. The number of jobs (a) and the number of events (b) processed for the CMS Tier-1 centers (in millions of events) in the 2017-2019 time period.

Figure 16 shows the number of jobs submitted to the JINR CMS Tier-1 center from the beginning of 2017 by type of tasks (production, reprocessing, analysis, etc.).
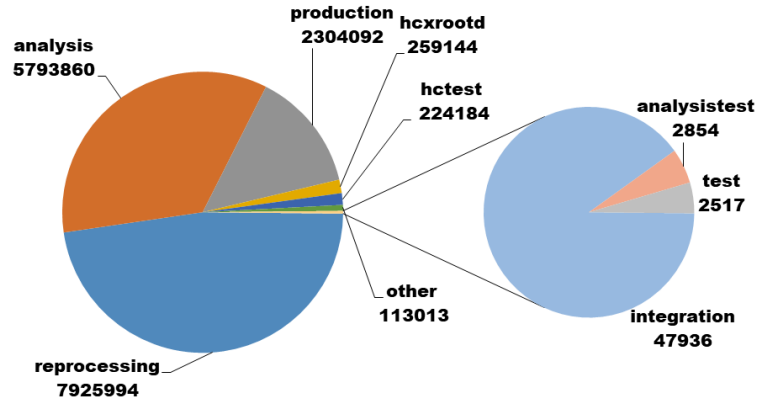

Fig.16. Distribution of jobs submitted to the JINR CMS Tier-1 center since January 2017 by type of processing (reprocessing, analysis, production, etc.)

One of the main functions of the Tier-1 centers is to provide data exchange with all global sites operating on the CMS experiment and storage of raw experimental and simulated data. Figure 17 illustrates the data transfer rates and the amount of data transferred for recording and processing from the Tier-1 and Tier-2 centers to the JINR CMS Tier-1 (Fig.17 a and b) as well as the data transfer rates and the amount of data transferred for recording and processing from the JINR CMS Tier-1 to the Tier-1 and Tier-2 centers (Fig.17 c and d) within 2018. The average data transfer rate for the JINR CMS Tier-1 is 400-520 MB/s.

18
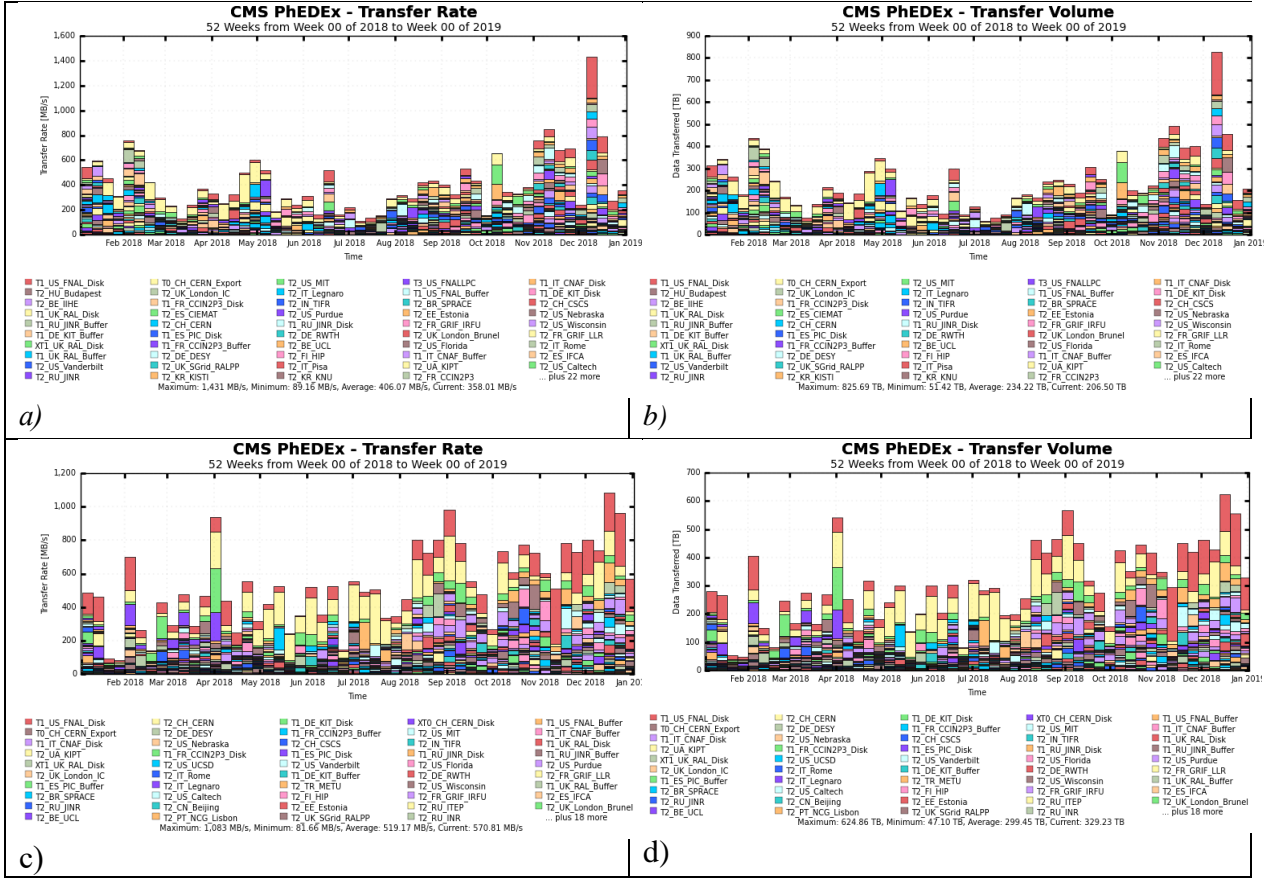
Fig.17. The data transfer rate and the amount of data transferred for recording and processing (Fig.17 a) and (Fig. 17 b) from the Tier-1 and Tier-2 centers to the CMS Tier-1 at JINR and the transfer rate and the amount of data transferred for recording and processing (Fig.17 c) and (Fig. 17 d) from the JINR CMS Tier-1 to the Tier-1 and Tier-2 centers in 2018.

The CMS PhEDEx (Physics Experiment Data Export) is responsible for large-scale data transfers across the grid ensuring transfer reliability, enforcing data placement policy and accurately reporting results and performance statistics. According to the PhEDEx dashboard, since 2017 the total data exchange with the tape robot has been 12.7 PB of which 5.2 PB of new data was transferred to tapes. The disk storage was used more actively: 27.7 PB of data was transferred to the JINR CMS Tier-1 site and 18.1 PB was transferred from it giving us a total of 45.8 PB of data exchange. A more detailed view on the 2018 transfers is presented in Figure 18.

Since commissioning, the JINR Tier-1 has steadily increased its productivity and maintained the level of availability and reliability required by the CMS experiment. Figure 19 shows the JINR Tier-1 (blue) monthly reliability results compared to the average Tier-1 reliabilities (orange) as well as to the WLCG target for site reliability (yellow area), which has been set to 97% since 2009, according to the WLCG MoU.

19
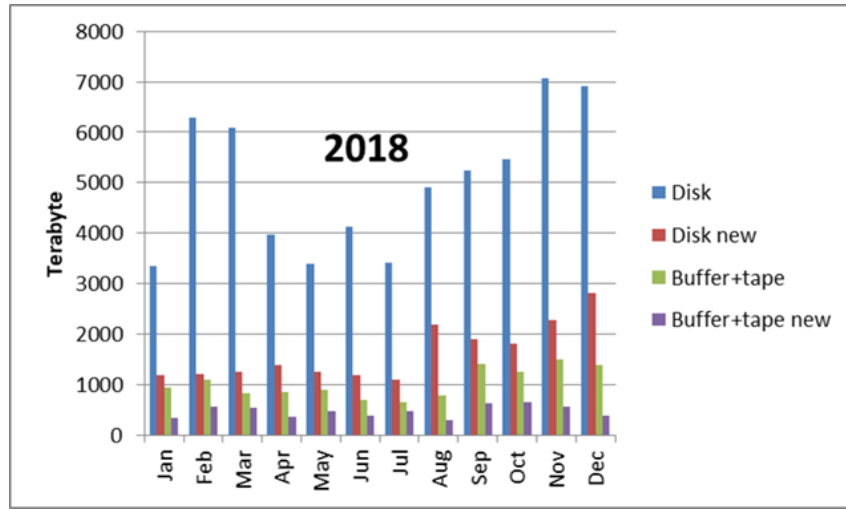
Fig.18. Statistics of the JINR CMS Tier-1 data exchange with other grid centers within 2018: Disk - all exchanges with the disk array (reading and recording), Disk new - new data from the total volume (recording), Buffer+tape - all exchanges with the tape robot (reading and recording), Buffer+tape new - new data from the total volume transferred to the tape robot (recording).
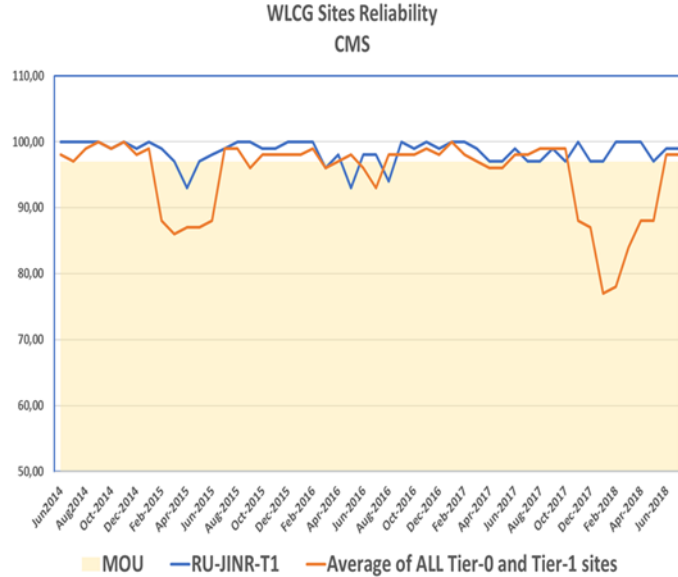
Fig. 19. WLCG Tier-1 site reliability results from 2014 to 2018.

The Tier-2 center at JINR provides data processing for all four experiments at LHC (Alice, ATLAS, CMS, LHCb) and apart from that, supports a number of virtual organizations (VOs) that are not members of LHC (BES, BIOMED, COMPASS, MPD, NOvA, STAR, ILC). The computing resources of the Tier-2 center consist of 4,128 cores. The data storage system is installed in two versions of the software: two dCache installations and two XROOTD installations. Figure 20 shows data on the use of the JINR Tier-2 site (JINR-LCG2) by virtual organizations in the framework of grid projects for the 2017-2019 time period.
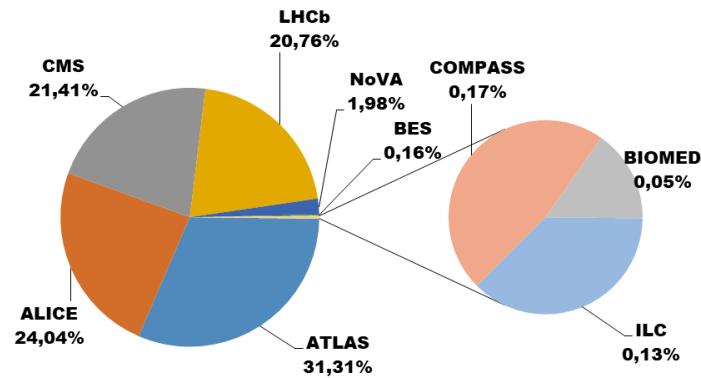
*Fig.20. Use of the JINR-LCG2 grid-site by virtual organizations within RDIG/WLCG/EGI.*

Special servers supporting the WLCG grid environment for VOs were installed. Several WLCG services were installed on physical machines, some of them – on virtual ones. WLCG services use UMD software for compatibility with the grid software environment in WLCG. Currently, 23 WLCG services have been installed. There are five settings of user interface (UI) for a job run into the distributed grid environment.

The computing element OSG HT-CONDOR was integrated in the Tier2 center infrastructure and the use of a long-lived grid proxy to authenticate was introduced. It provides a way for VO STAR to process data using Tier-2 with an over 90% first pass efficiency.

The support for VO JUNO was added. The computing tests were performed and showed good results. Thus, the JINR Tier-2 is ready to be used for processing data from the JUNO experiment.

High-performance computer system

The MICC allows users to perform calculations including parallel ones outside the framework of the grid environment. It is necessary for experiments such as NOvA, BES, NICA/MPD, COMPASS, etc. as well as for local users of JINR Laboratories. JINR and grid users have access to all computing facilities via a unified batch processing system. Figure 21 shows the time distribution of jobs performed on the computing cluster by JINR subdivisions and user groups.
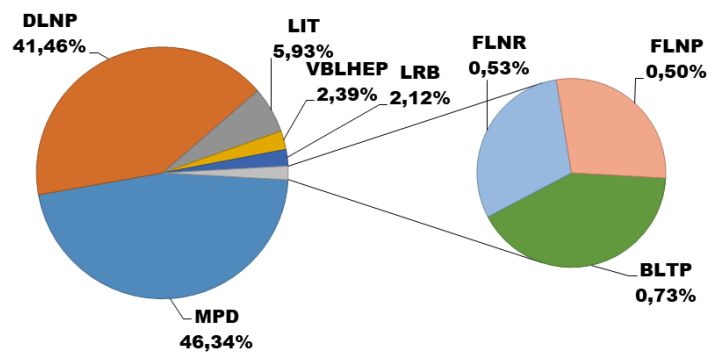
Fig.21. Statistics of using the computing cluster by subdivisions and experiments underway at JINR within 2017-2019 not taking into account grid users.

The systems of data storage and access dCache, EOS and XROOTD ensure work with data for JINR local users as well as WLCG users and collaborations. Two dCache installations are supported: dCache-1 for the CMS and ATLAS experiments; dCache-2 for local users, groups of users and international projects such as NICA/MPD, HONE, FUSION, BIOMED, COMPASS.

All storage systems are constructed under the hardware data protection mechanism RAID6 and the software mechanism RAIDZ2 that is not inferior to the reliability of the RAID6 hardware.

The CVMFS subsystem was put into operation. It is used mainly for access to big software packages used by collaborations with the WLCG computing model. The NICA MPD and BM@N software packages were also placed into CVMFS. The repository keeps several versions of the same software, and it can be changed on the CVMFS server only by an authorized person. The read access from the repository was set up on all the MICC working nodes, and it can be used worldwide by means of the CVMFS client. The MPD and BM@N software repository occupies 9.5 GB on the MICC CVMFS servers.

JINR joined the group of research centers (Fig.22) that develop the WLCG data lake prototype for HL-LHC [27]. The data lake prototype was built as a distributed EOS storage system. EOS is used for storing and accessing big arrays of information. It can be applied for collective data simulation, storage of raw data gathered from experimental setups, data processing and analysis. Four main LHC experiments use EOS for data storage and processing within CERN. The global access to EOS is carried out by means of the WLCG software.
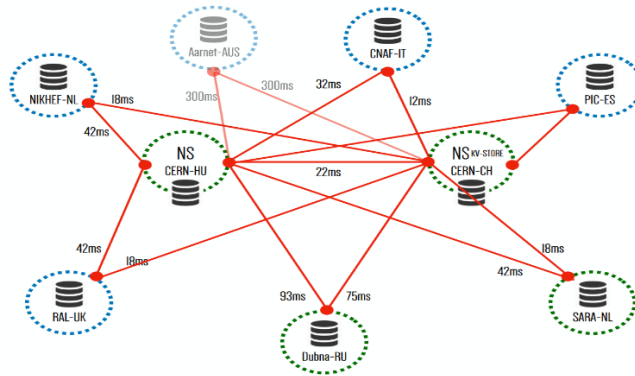

Fig.22. Data lake prototype of participating research centers.

The data lake prototype turned out to be small in terms of resources, but geographically distributed. The performance of different data placement scenarios with different access patterns was measured. The results were positive, and the CMS integration is a logical next step. The deployment and testing experience on the data lake is meant to help overcome the HL-LHC computing challenge.
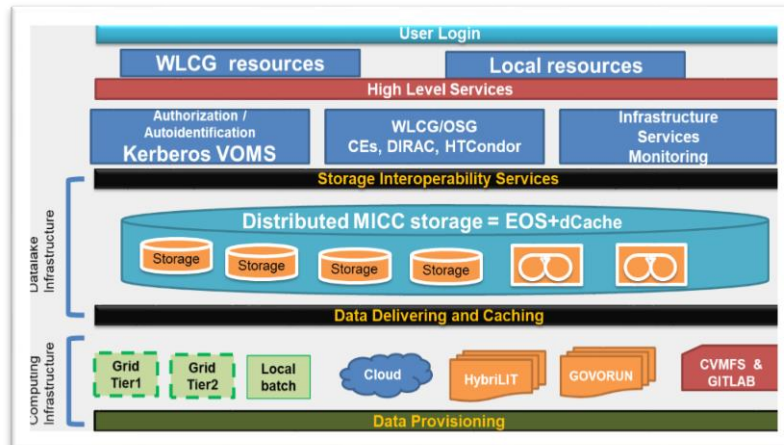

Fig.23. EOS integration into MICC.

The EOS was integrated into the MICC structure successfully (Fig.23). There is currently 3740TB of disk space available for EOS. The NICA experiments have already used EOS for data storage. At present, there is ~ 81TB of raw BM@N data and ~84GB of simulated MPD data stored

in the EOS instance. EOS is visible as a local file system on the MICC working nodes and allows authorized users (by kerberos5 protocol) to read and write data.

Heterogeneous platform HybriLIT

The commissioned "GOVORUN" supercomputer named after Nikolai Nikolaevich Govorun, with whom the development of information technologies at JINR has been connected since 1966, became a natural elaboration of the heterogeneous cluster HybriLIT being the MICC component. The "GOVORUN" supercomputer is a joint project of the Bogoliubov Laboratory of Theoretical Physics and the Laboratory of Information Technologies supported by the JINR Directorate. The project is aimed at the significant speed-up of complex theoretical and experimental studies in the field of nuclear physics and condensed matter physics underway at JINR including the development of the NICA megaproject computing.

The supercomputer commission led to the significant performance increase of CPU- as well as GPU-components of the HybriLIT heterogeneous cluster and, together with it, the supercomputer formed a heterogeneous platform.
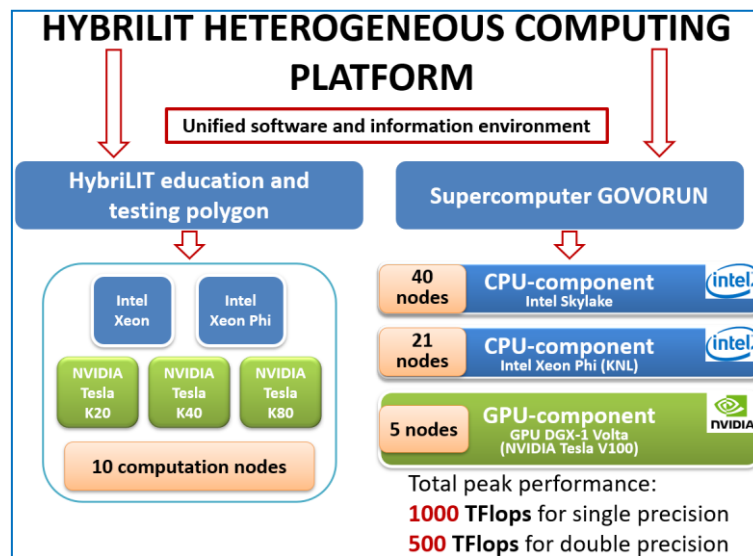

Fig.24. HybriLIT platform structure.

The platform consists of two elements (Fig.24), i.e. the education and testing polygon and the "GOVORUN" supercomputer, combined by the unified software and information environment (Fig.25).
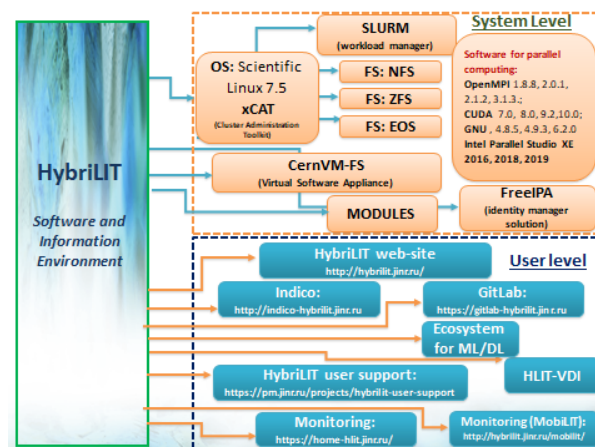

Fig.25. Structure of the software and information environment.

The "GOVORUN" supercomputer is designed to carry out resource-intensive and massive parallel calculations for the solution of a wide range of challenges facing JINR, which becomes

possible due to the heterogeneity (presence of different types of computing accelerators) of the supercomputer computing architecture. The education and testing polygon is aimed at exploring the possibilities of novel computing architectures, IT-solutions as well as conducting training courses on parallel programming technologies, modern tools of the development, debugging and profiling of parallel applications, application packages. Platform users are able to develop and debug their applications on the education and testing polygon, furthermore, carry out calculations on the supercomputer, which allows them to effectively use the supercomputer resources. The given possibility is provided by the unified software and information environment including the unified system level (the operation system, the job scheduler, file systems and software) as well as a set of services allowing users to quickly get the answers to their questions, jointly develop parallel applications, receive information about conferences, seminars and meetings dedicated to parallel programming technologies (Fig.25).

"GOVORUN" supercomputer

The "GOVORUN" supercomputer is a heterogeneous computing system containing the GPU-component based on the NVIDIA graphics accelerator and the CPU-component based on two Intel computing architectures. The GPU-component consists of 5 NVIDIA DGX-1 servers. Each server has 8 GPU NVIDIA Tesla V100 based on the latest architecture NVIDIA Volta. Moreover, one server NVIDIA DGX-1 has 40960 cores CUDA, which are equivalent to 800 high-performance central processors. A whole number of novel technologies are used in DGX-1, including the NVLink 2.0 wire with the bandwidth up to 300 Gbit/s.

The supercomputer CPU-component is implemented on the high-density architecture "RSK Tornado" with direct liquid cooling developed by specialists of the Russian group of RSK companies. Novel universal computing cabinets "RSK Tornado'" with a record energy density and a system of precision liquid cooling balanced for constant work with a high-temperature coolant (up to +63°C at the entrance to the computer cabinet) are installed in LIT. The work in the "hot water" mode for the given solution allowed to apply the year-round freecooling mode (24x7x365) using only dry cooling towers operating at an ambient temperature of up to +50°C and completely get rid of the freon circuit and chillers. As a result, the average annual PUE of the system reflecting the level of energy efficiency is less than 1,03. Thus, cooling consumes less than 3% of all the consumed electricity which is an outstanding result for the HPC industry.

CPU computing nodes are based on the Intel server products, namely, the most powerful 72-core server processors Intel® Xeon Phi™ 7290, processors Intel® Xeon® Scalable (models Intel® Xeon® Gold 6154) and novel high-speed solid-state disks Intel® SSD DC P4511 with the NVMe interface and a capacity of 1 TB. For high-speed data transfer between computing nodes the supercomputer uses an advanced switching technology Intel® Omni-Path providing the speed of non-blocking switching up to 100 Gbit/s based on 48-port switches Intel® Omni-PathEdgeSwitch 100 Series with 100% liquid cooling. The use of Intel® Omni-PathArchitecture allows not only meeting the current needs of resource-intensive user applications but also providing the necessary network bandwidth for the future.

The peak supercomputer performance is 1 PFlops for single-precision operations and 500 TFlops for double-precision operations.

At present, the supercomputer is used to solve problems that require massive parallel calculations in various fields of nuclear physics and high energy physics, particularly in lattice quantum chromodynamics to study the properties of hadronic matter at high energy density and baryon charge and in the presence of supramaximal electromagnetic fields, for mathematical modeling of interactions of antiprotons with protons and nuclei using DPM, FTF and UrQMD+SMM generators developed at JINR and being of interest for the NICA-MPD experiment, for modeling the dynamics of collisions of relativistic heavy ions as well as for solving applied problems such as calculating Josephson junctions, modeling the dynamics of many-particle Boson systems in magnetic optical traps, calculating corrections for the matrix element in the first Born approximation in case of a direct ionization of a helium atom by a fast proton taking into account

different models of the final state, etc. In addition, work related to the development of the NICA megaproject computing is carried out on the supercomputer basis.

The average load on computing components is the following: the component based on Skylake is 80,58% (maximum 100%), the component based on KNL is 38,41% (maximum 74%), the component with GPU computing accelerators is 73,58% (maximum 100%).

Currently 40 users from all the Institute Laboratories are performing calculations on the supercomputer. In total, within the commission period over 135000 tasks on all computing components were completed by all groups performing calculations on the supercomputer (Fig.26).
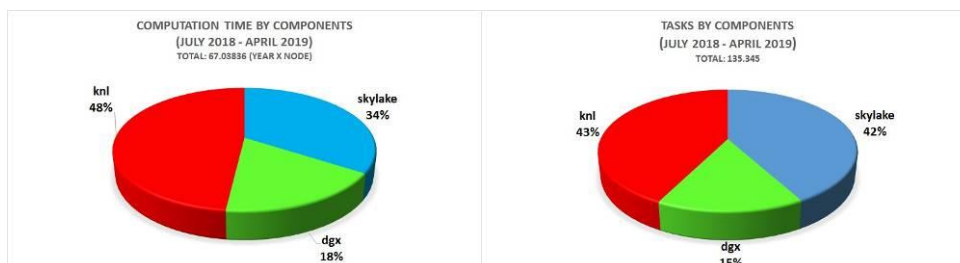
Fig.26. Usage of the computational partitions of the GOVORUN supercomputer from the moment of commissioning in July 2018 up to April 2019.

Education and testing polygon

At present, the education and testing polygon contains 10 computing nodes with graphics accelerators NVIDIA Tesla K20X, K40, K80, coprocessors Intel Xeon Phi 5110P, 7120 and processors Intel Xeon E5-2695 V2 and V3. The total number of CUDA cores is 77184, processor cores – 252, coprocessor cores – 182, the total memory volume is 2,5 TB, the total performance for single-precision calculations is 140 Tflops and 50 Tflops for double-precision calculations.

Unified software and information and hardware environment of the HybriLIT platform

Information technology is one of the most rapidly developed fields in terms of the hardware development (development of computing architectures, data storage systems, network solutions) as well as the development of methods, algorithms, software for calculations on novel computing architectures and the software development using novel frameworks and libraries. All the enumerated above creates serious requirements to the software and information environment, namely, *flexibility*, the ability to quickly form an IT-environment for solving specific applied tasks; *scalability*, for a fast expansion/reduction of the computing field; *user-friendly*, a prompt response to user requests from the platform (community) solving problems in various fields developed at the Institute. The indicated requirements are satisfied through different mechanisms. *Flexibility* is achieved due to the formed dynamic pool of virtual machines designed for debugging and launching jobs (VM user interface supporting only the SSH-protocol) and VMs for solving graphical tasks (HLIT-VDI service based on the Citrix technology). *Scalability* is attained due to the possibilities of the SLURM job scheduler allowing, if necessary, to reallocate computing resources in queues. The *user-friendly* requirement of the software and information environment is realized by the rapid deployment of necessary IT-environments and work with heterogeneous computing group users through services, which are actively developed and supported by the team.
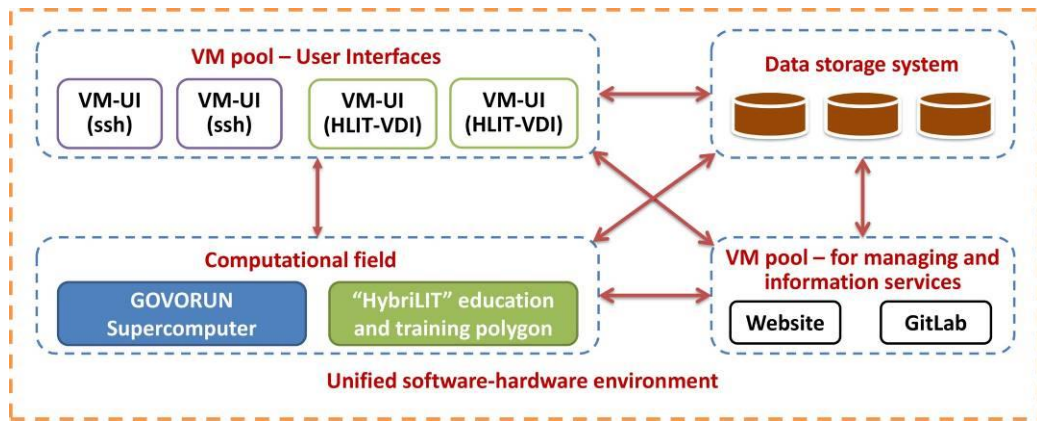
Fig.27. Structure of the software and hardware environment.

To increase the application development efficiency and carry out calculations using the HybriLIT resources, the heterogeneous computing team develops and supports the software and information environment (Fig.27), which can be divided into two parts.

The first level is considered to be a system one and includes basic software:

• Scientific Linux 7.5 OS (the choice was based on its compatibility with the other MICC components)

• NFS4 and EOS file systems;

• SLURM 17.02.9 workload manager and job scheduler, which manages jobs between the nodes of the heterogeneous cluster. Its wide use in different systems from the TOP500 list (almost one-half of supercomputers from TOP500 uses SLURM) has become a weighty argument for its choice;

• The module 3.2.10 package enables dynamic modification (interrogation, allocation and deallocation) of the user environment, i.e. it loads paths to header and library files, documentation files and adds paths to catalogues with the executive file;

• CVMFS (CERN Virtual Machine File System).

This methodology of developing the cluster provides the possibility for the efficient system administration such as the dynamic expansion of the platform with further implementation of new computing nodes, the possibility for synchronous updates and software development on all current and future computing nodes, the fast node setting including faults and reloading.

The second level includes tools for the development, debugging and profiling of parallel applications and carrying out resource-intensive calculations. One of the main components of the second level is information support provided for users. This support is necessary for organizing users' work in a more efficient way and providing access to useful information while working on the cluster. Information support tools include the following services:

• HybriLIT web page (http://hybrilit.jinr.ru): it contains the detailed information about resources provided for platform users. Particular emphasis is put on hardware, which contains hardware specifications, and software installed on the platform. It should be noted that the registration procedure is crucial. Therefore, the detailed description on the registration is provided at the web page. Moreover, users can have a look at the instruction for running tasks. The instruction includes the basic information about the platform.

• Indico system (http://indico-hybrilit.jinr.ru): it is used by the HybriLIT team in order to organize conferences, seminars and meetings dedicated to parallel programming technologies. The system allows creating different events that will take place in the Institute. The created event provides the basic information about the event time and place. There is also a possibility to upload all the materials of lectures so that every user can download them.

Among services that include the detailed information about the platform and the possibilities it provides, there are also services aimed at the close interaction between users and the HybriLIT team. Such services include:

• HybriLIT User Support: it is a project developed in the Project Management Service system (http://pm.jinr.ru), which allows to answer user questions, upload useful materials, publish news, etc. The project is being developed for a more efficient interaction between users and the HybriLIT team. The system provides possibilities for the information distribution about the upcoming events and the current state of the cluster. Users can also create tasks on upgrade and debugging. Thus, the interaction between users of the cluster and developers is carried out quickly and efficiently.

• GitLab (http://gitlab-hybrilit.jinr.ru): it is a service aimed at the joint parallel development of applications. It is a version control system and allows one to follow the changes in the code of projects. The system functional is wide and includes such possibilities as the separation of access among users, the task management system, wiki, the code review, etc. The HybriLIT team developed several projects within the system; among the main ones are the project on writing documentation and the project on the development of a monitoring system.

With regard to the fact that users of the platform are scientists and specialists from different countries, bilingual support (Russian and English) for all resources is crucial.

The software and information environment of the platform allows users to develop applications and carry out calculations using the latest computing architectures.

As mentioned above, SLURM is used on the platform as a job scheduler and a workload manager; its settings allow distributing computing nodes in groups (queues) formed on the basis of computing architectures used in calculations. A mobile application MobiLIT was developed for receiving information on available resources in different SLURM queues, node characteristics, the status of running jobs, i.e. for a more effective organization of calculations on the education and testing polygon. It allows the following:

- to monitor users' jobs;
- to look through files (*.out, *.dat, *.in or *.sh);
- to be able to terminate the job;
- to monitor the estimated time of the end of calculations;
- to monitor available resources on different polygon nodes and in different queues organized for using various computing architectures;
- to have a full information on the occupied resources for all jobs running on the polygon, etc.

The MobiLIT application is developed on the NativeScript framework for users of mobile phones and tablets with Android and iOS operating systems. At present, the application is available for downloading through the Google Play Store for Android users and through the App Store for iOS (in a test mode). To log in, the authorization of the HybriLIT platform is used. The mobile application MobiLIT developed for polygon users enriches the cluster software and information environment with a convenient and up-to-date IT-service simplifying parallel calculations and provides an additional optimization for the use of computing resources.

In frames of the project on the development of the HybriLIT heterogeneous platform, the configuration consisting of virtual machines for user access to HybriLIT hardware and software resources was presented in 2017. The first implementation was based on the QEMU/KVM hypervisor and the SPICE remote access protocol. The major difference of the second implementation was the change of the remote access protocol on X2Go. The main disadvantage of the used implementations is the non-centralized management of virtual machines and the lack of hardware support for the graphics accelerator, which leads to the additional load on the central processor. In turn, the use of virtual machines made it possible to increase the placement density in comparison with the personal analog (local desktop), namely, 12 VMs in one server located in LIT JINR instead of 12 personal computers with access to high-performance software. Thus, one of the main requirements to future virtual machines is the use of a graphics accelerator for image processing and remote access with the centralized management.

The transition to Citrix software and the XenServer hypervisor provided a compact, universal solution with seamless remote client access to a virtual desktop, which is the major project motivation.

A schematic presentation of VM implementations for remote access is shown in Figure 28.
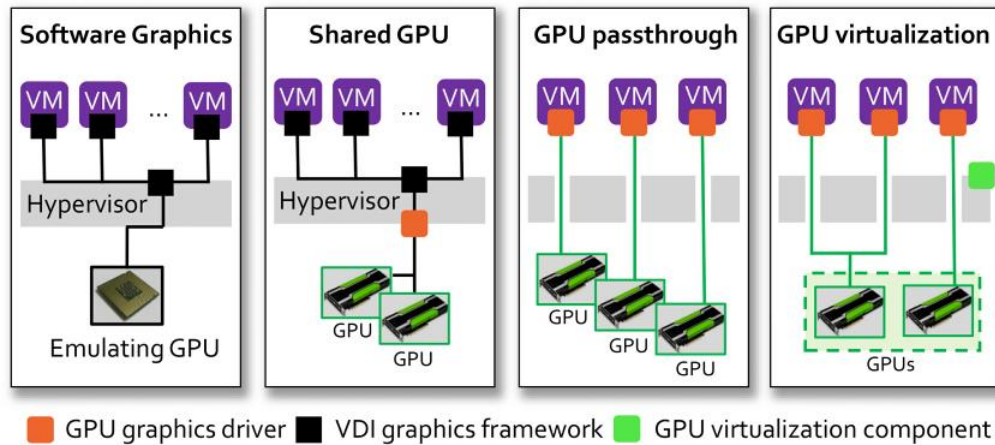
Fig.28. Schematic presentation of VM implementations with a remote access function, support for the central processor and the graphics accelerator.

Taking into account that the number of managing virtual machines in the basic configuration comprises from 2 to 5, it is recommended to separate them from future virtual machines for remote access by creating on another server. To prepare the server for managing VMs it is sufficient to install a hypervisor compatible with software, XenServer in this case, and assign it a network address. Citrix XenCenter software is used for managing hypervisors and virtual machines. In the current configuration the minimal number of managing VMs was involved, that is two with 4 cores, 10 Gb of RAM and 100 Gb of disk space on each. The first virtual machine with Windows Server 2016 contains the domain controller with Active Directory for user and computer authorization, DHCP and the NS server. The second one is the Delivery controller, was entered into the domain and contains Citrix XenStudio for creating and managing all VMs with remote access, StoreFront for user access to VMs, Citrix and NVIDIA license managers. The following actions were performed for the VM placement: the XenServer was installed, the work mode of graphics accelerators NVIDIA Tesla M60 was modified using the gpumodeswitch utility, the vGPU driver was installed with support for Citrix profiles, XenServer was added into XenStudio. The configured hypervisor and the model of the graphics accelerator NVIDIA Tesla M60 (8 pieces) can locate up to 64 virtual machines with minimal characteristics. Such a decision would have the maximum density and the minimum performance, that is why the search for an optimal configuration is a subject of separate studies. The major part of the settings is dedicated to creating a master image with the HybriLIT environment and setting up XenStudio, i.e. creating a virtual machine catalog with a choice of profile and delivery group with access to certain VMs. The creation of a master image takes several steps:

1. Creating VM with 4 cores, 20 Gb of RAM, 40 Gb of disk space;
2. Installing prerequisites from the repository;
3. Installing XenTools extensions;
4. Disabling the module *nouveau*;
5. Installing the NVIDIA driver;
6. Licensing the graphics accelerator;
7. Installing the HybriLIT environment (cvmfs, modules, mounts, firewall);
8. Installing Citrix Virtual Delivery Agent (VDA).

When installing VDA, the virtual machine will be entered into the domain, after that it will remain to create a virtual machine catalog, a group for access in XenStudio on the VM basis and reboot VM. After rebooting one should check the status of virtual machines in XenStudio, the field "Registered" in the control tab shows that VM is ready for connections. In the browser at

http://delivery-controller/Citrix/StoreWeb the Citrix Site will be available; it will propose the user to download Citrix Receiver, i.e. a client for remote access to VM. After installing the client and authorizing, the page with available VMs will appear. Clicking on any of them will lead to the .ica file download, the launch of which will open access to the remote desktop. Furthermore, several modules for the work with high-performance software such as Maple, Comsol, Wolfram Mathematica, Matlab will be available to the user in the familiar HybriLIT environment. The list of available modules can be viewed using the command *module avail*.

Ecosystem for tasks of machine learning, deep learning and data analysis

The active implementation of the neural network approach, methods and algorithms of machine learning and deep learning (ML/DL) for solving a wide range of problems is defined by many factors. The development of computing architectures, especially while using DL methods for training convolutional neural networks, the development of libraries, in which various algorithms are implemented, and frameworks, which allow building different models of neural networks (see the Venn diagram shown in Fig.29), can be referred to the main factors. To provide all the possibilities both for developing mathematical models and algorithms and carrying out resource-intensive calculations including graphics accelerators, which significantly reduce the calculation time, an ecosystem for tasks of ML/DL and data analysis has been created and is actively developing for HybriLIT platform users.
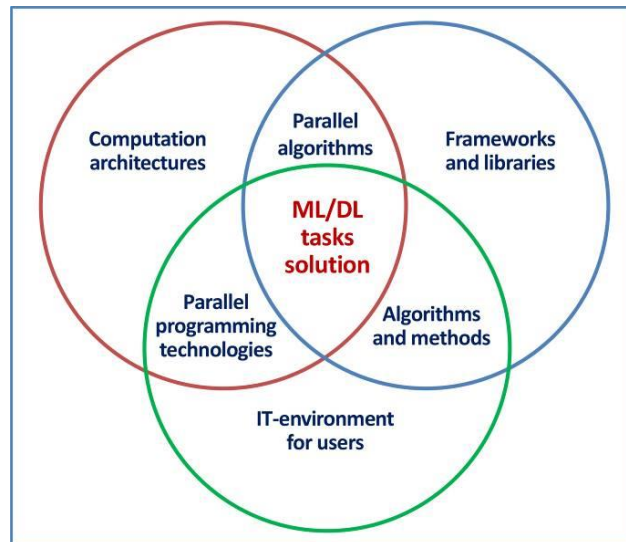
Fig.29. Venn diagram on the implementation of the neural network approach, methods and algorithms of ML/DL to solve applied tasks.

The created ecosystem has two components (Fig.30):
- the computing component is aimed at carrying out resource-intensive, massive parallel tasks of neural network training using NVIDIA graphics accelerators;
- the component for the development of models and algorithms on the JupyterHub basis, i.e. a multi-user platform for working with Jupyter Notebook (known as IPython with the possibility to work in a web browser), including several libraries and frameworks presented in Figure 31.
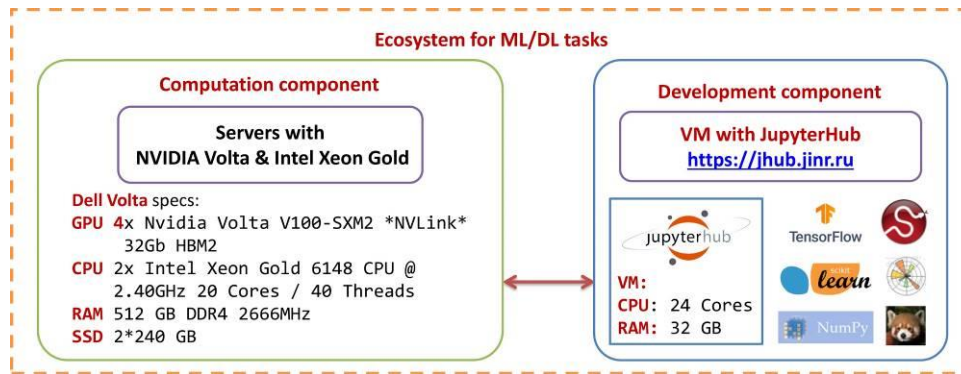
Fig. 30. Two-component ecosystem for tasks of ML/DL and data analysis.
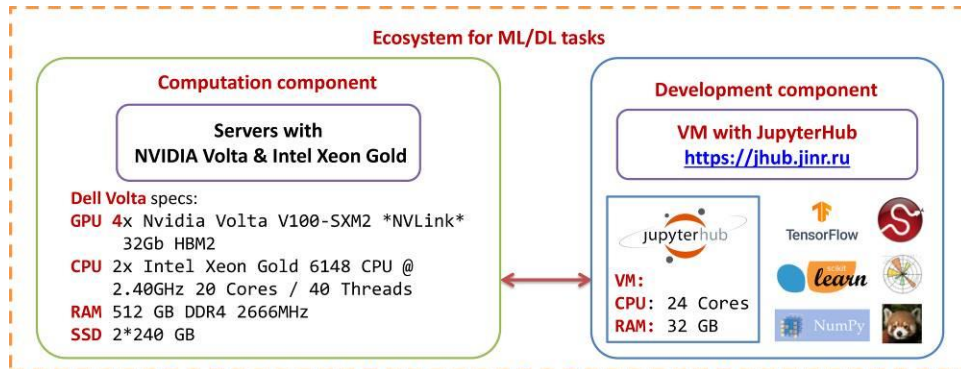

Fig. 31. Ecosystem for ML/DL tasks built on the JupyterHub multi-user service (a multi-user version of the Notebook).
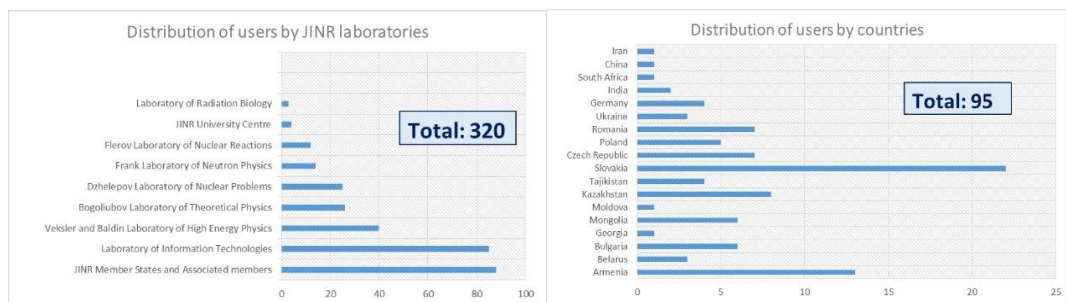

Fig.32. Distribution diagram of users of the HybriLIT platform among JINR Laboratories and its Member States.

All the components of the HybriLIT platform, i.e. the "GOVORUN" supercomputer and the education and testing polygon, are actively used by JINR scientific groups to solve a wide range of challenges facing the Institute; the total number of platform users is 320, the user distribution among laboratories is shown in Figure 32. Gratitude for the use of the computing resources of the platform, including the resources of the supercomputer, as well as the services deployed on the platform, was expressed in more than 60 publications.

Cloud environment

New high availability setup based on Raft consensus algorithm

Several software updates to the latest stable available versions of the OpenNebula platform were performed on the JINR cloud infrastructure within the period from 2017 until February 2019.

Starting from the OpenNebula 5.4 release there is a new built-in mechanism for the high availability (HA) setup based on the so-called Raft consensus algorithm. According to the OpenNebula documentation, a consensus algorithm relies on two concepts:

- System State, which in case of OpenNebula-based clouds means the data stored in database tables (users, ACLs or VMs in the system);
- Log, which is a sequence of SQL statements consistently applied to the OpenNebula DB in all servers to evolve the system state.

To preserve a consistent view of the system across servers, modifications to the system state are performed through a special node called a leader. The OpenNebula cloud front-end nodes (CFNs) select a single node to be the leader. The leader periodically sends heartbeats to other CFNs called followers to keep its leadership. If the leader fails to send the heartbeat, the followers are promoted to candidates and a new selection begins.

Whenever the system is modified (e.g. a new VM is added to the cluster), the leader updates the log and replicates the entry in a majority of followers before actually writing it to the database. It increases the latency of DB operations, but enables a safe replication of the system state, and the cluster can continue its operation in case of the leader's failure.

During a software upgrade on the JINR cloud to the OpenNebula 5.4 release the HA setup based on the Raft consensus algorithm was implemented. Following the OpenNebula documentation recommendations the JINR cloud has the odd number of front-end nodes (it equals three in our case). They are presented in Figure 33 and marked by the black digit "2" in the same color square.
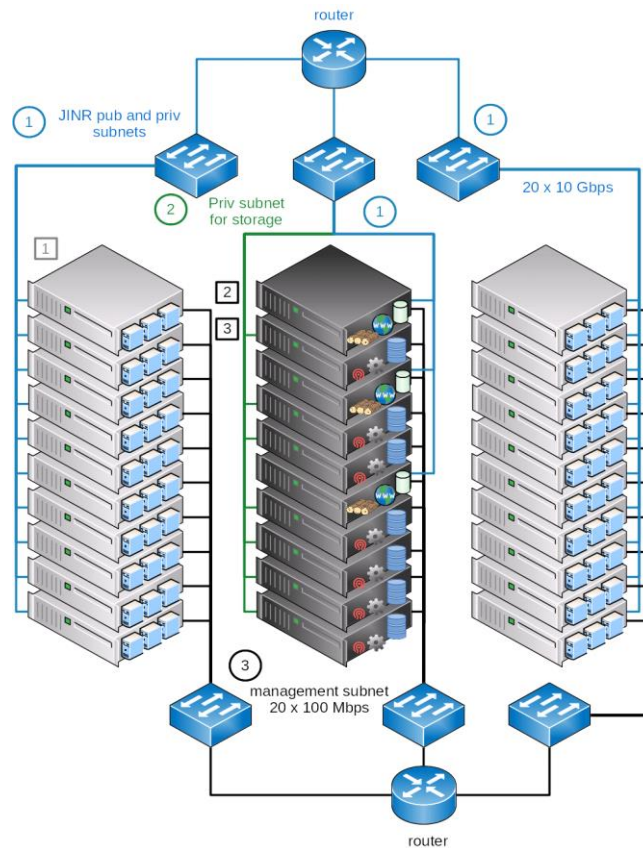
Fig.33. Schema of the high availability setup of the JINR cloud based on the Raft consensus algorithm.

There are the following types of servers shown in Figure 33:
- Cloud worker nodes (CWNs), which host virtual machines (VMs) and containers (CTs), are marked in the figure by the digit "1" in the grey square;
- Cloud front-end nodes (CFNs) where all OpenNebula core services including the database, the scheduler and some others are deployed (such hosts are marked in the figure by the black digit "2" inside the same color square);

- Cloud storage nodes (CSNs) based on Ceph SDS for keeping VMs' and CTs' images as well as user data (marked by the black digit "3" inside the same color square).
  All these servers are connected to the same set of networks:
- JINR public and private subnets (they are marked in the figure by blue lines near which the digit "1" is placed in the same color circle and signed as "JINR pub and priv subnets");
- Isolated private network dedicated for the SDS traffic (dark green lines with the digit "2" in the circle of the same color and signed as "Priv subnet for storage");
- Management network (black lines with the digit "3" in the circle of the same color singed as "management subnet").
  All network switches excluding those which are for the management network have 48 ports with 10 GbE each as well as six 40 Gbps SPF-ports for uplinks.

Hardware resources

The JINR cloud resources were increased up to 1564 CPU cores and 8.1 TB of RAM in total (see Fig.33 with changes in CPU cores and the total amount of RAM of the JINR cloud for both KVM- and OpenVZ-based clusters).
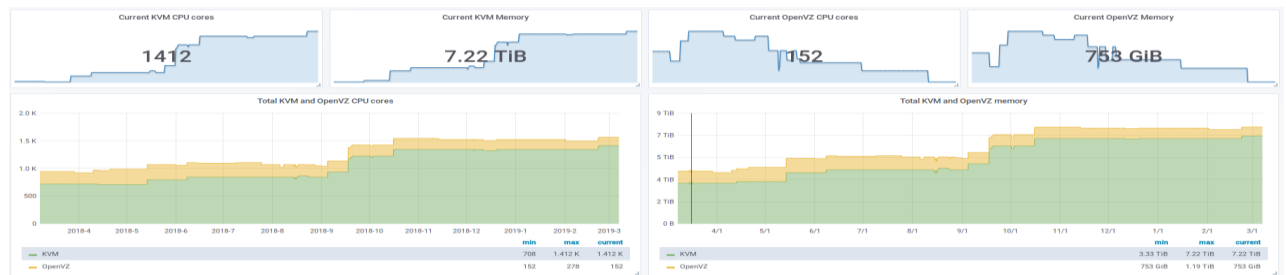

Fig.34. Changes in CPU cores and the total amount of RAM of the JINR cloud.
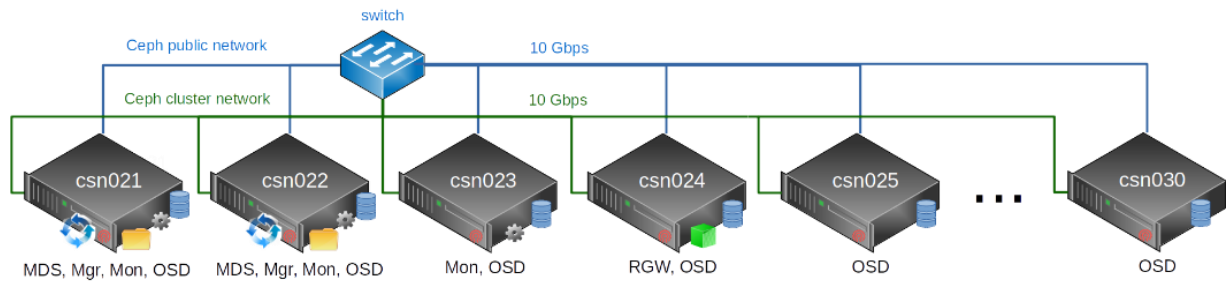
Ceph-based software defined storage

One of the key components of the JINR cloud infrastructure is a software-defined storage (SDS) based on Ceph. It has been run in a production mode since the end of 2017. It delivers object, block and file storage in one unified system. According to the Ceph documentation, each Ceph server can play one or several roles which are the following:
- Monitor (mon). It maintains maps of the cluster state including the monitor map, the manager map, the OSD map and the CRUSH map. These maps are critical cluster states required for Ceph daemons to coordinate with each other. Monitors are also responsible for managing authentication between daemons and clients. At least three monitors are normally required for redundancy and high availability;
- Manager (mgr). It is responsible for keeping track of runtime metrics and the current state of the Ceph cluster including storage utilization, current performance metrics and system load. The Ceph Manager daemons also host python-based plugins to manage and expose Ceph cluster information including a web-based dashboard and REST API. At least two managers are normally required for high availability;
- Object storage daemon (OSD). It stores data, handles data replication, recovery, rebalancing and provides some monitoring information to Ceph Monitors and Managers by checking other Ceph OSD daemons for a heartbeat. At least three Ceph OSDs are normally required for redundancy and high availability;
- Metadata Server (MDS). It stores metadata on behalf of the Ceph Filesystem. Ceph MDSs allow POSIX file system users to execute basic commands (like ls, find, etc.) without placing an enormous burden on the Ceph Storage Cluster.

Apart from these daemons, there is the RADOS gateway (RGW) which provides interfaces for interacting with the storage cluster.

The JINR Ceph-based SDS deployment schema as well as each server roles are shown in Figure 35

Fig. 35. Schema of the Ceph-based software-defined storage deployed at LIT JINR.

At the time of writing the report, the total amount of raw disk space in this SDS has been about 1 PB. Due to triple replication, an effective disk space available for users is about 330 TB.

Distributed information and computing environment based on the JINR Member State organizations' resources

A necessity to process and analyze huge data volumes as well as a luck of computing and storage resources in a unified place for such kind of tasks led to the development of worldwide distributed information and computing infrastructures which integrate resources all over the world.

JINR participates in the large number of research projects in many of which computing infrastructures play important role in obtaining meaningful scientific results. In this connection, an integration of JINR Member State organizations' resources in a uniform distributed information and computing environment (DICE) is an important and topical task, the solution of which would significantly accelerate scientific research.

Each organization participating in joining its local computational resources has created or is creating a private cloud infrastructure for the effective utilization of resources.

Initially a driver for cloud integration following the so-called "cloud bursting" model [22] was developed by the fellows of LIT JINR. During several years the driver was successfully used as a tool for joining clouds from such JINR Member State organizations as Plekhanov Russian University of Economics (Moscow, Russia), Bogolyubov Institute for Theoretical Physics (Kiev, Ukraine) and the Institute of Physics of the National Academy of Sciences of Azerbaijan (Baku, Azerbaijan).

Enormous interest of the JINR Member States in a uniform distributed computational environment led to the growing number of its participants. It revealed low scalability of the used approach due to the sufficiently increased complexity of such infrastructure maintenance (every new cloud integration requires changes in configuration files of each integrated cloud as well as appropriate service restarts). That is why research work on evaluating possible alternatives started.

Among existing software platforms for distributed computing and data management the DIRAC (Distributed Infrastructure with Remote Agent Control) [18] platform was chosen because of the following reasons:

- it provides the whole needed functionality including both job and data management;
- cloud as computational back-end support (although an appropriate plugin required some development);
- easier service deployment and maintenance in comparison with other platforms with similar functionality (e.g. EMI).

The schemas of cloud integration based on both approaches mentioned above are shown in Figure 36.
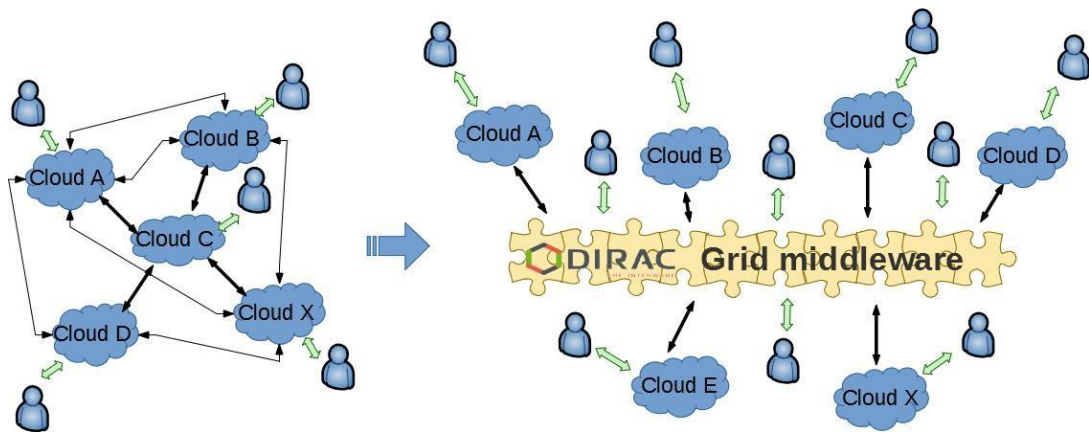
Fig 36. Two schemas of cloud integration: 1) using the cloud bursting driver (left) and 2) using the DIRAC interware (right).

Such an approach allows one to share resources of each cloud between external grid users and local non-grid ones.

The services of the DIRAC platform itself are deployed in the JINR cloud, which, together with the clouds of the organizations of the JINR Member States, provides computing resources for the grid environment.

DIRAC is a software framework for distributed computing providing a complete solution to one (or more) user community requiring access to distributed resources.

DIRAC builds an intermediate layer between users and resources offering a common interface to a number of providers of heterogeneous resources.

There is a built-in mechanism for interaction with clouds in DIRAC, which is based on the use of one of the implementations of the Open Cloud Computing Interface (OCCI), for example, ruby-written rOCCI. As cloud integration with the help of the cloud bursting driver also relied on utilizing rOCCI-servers locally deployed at each partner cloud, adding clouds into the DIRAC-based infrastructure required much less effort.

Due to a limited set of operations which a client was capable to perform with cloud via the rOCCI protocol, it was decided to extend the functionality of the corresponding DIRAC module (named "VMDIRAC") to be able to interact with OpenNebula-based clouds using their native XML-RPC API. Such a module was developed and successfully validated for the use in the DIRAC production release.

At present, the integration of the JINR Member State organizations' clouds into the DIRAC-based distributed platform is at different stages. Their technical integration has been fulfilled, while work on testing and debugging is in progress for the Astana branch of the Institute of Nuclear Physics, i.e. the private establishment "Nazarbayev University Library and IT Services" (Astana, Kazakhstan), the Scientific Research Institute for Nuclear Problems of the Belarusian State University (Minsk, Belarus), Yerevan Physical Institute (Yerevan, Armenia), the Institute of Physics of the National Academy of Sciences of Azerbaijan (Baku, Azerbaijan) and Plekhanov Russian University of Economics (Moscow, Russia). The clouds of the Institute for Nuclear Research and Nuclear Energy of BAS (Sofia, Bulgaria), Sofia University "St. Kliment Ohridski" (Sofia, Bulgaria), Georgian Technical University (Tbilisi, Georgia) and the Institute of Nuclear Physics (Tashkent, Uzbekistan) are in the process of integration. Negotiations with the University "St. Kliment Ohridski" (Bitola, Macedonia) are in progress. Figure 37 provides information on the geographical location of the organizations listed above.
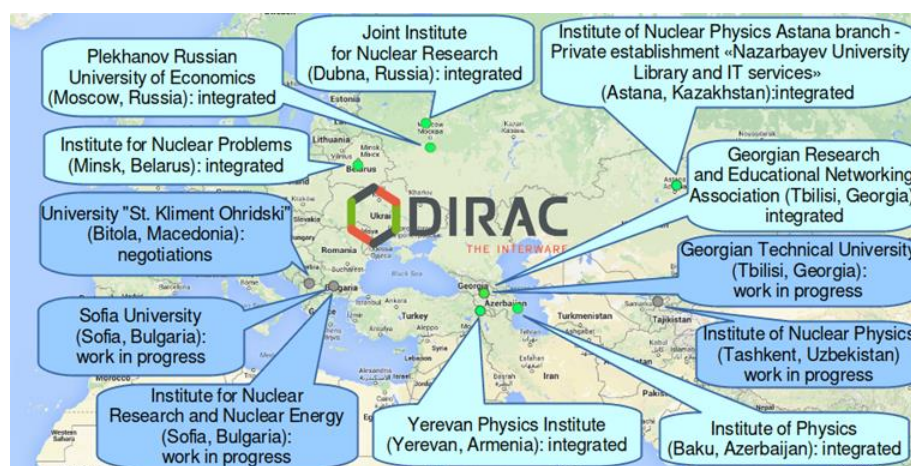
Fig.37. Geographical location of cloud infrastructures of the organizations from the JINR Member States involved in the integration of cloud resources.

The increasing number of the JINR DICE participants demonstrates significance and relevance of such an activity. It is related to the willingness of the parties to join their efforts and resources in solving fundamental and applied tasks, which require modern and novel approaches in distributed computations and data management.

Cloud service for scientific and engineering computations

Quite often small research scientific groups from JINR and its Member State organizations do not have access to sufficiently powerful resources required for their work to be productive. Global computational infrastructures can be challenging for small research teams because of bureaucracy overheads as well as usage complexity of underlying tools.

It is a widespread practice to do computations which may last days or even weeks on own laptops or PCs. Others buy powerful desktop servers and offload computational work there.

Possible drawbacks of such an approach are the following:

- Necessity to take care of the proper hosting environment for such hardware.
- Equipment maintenance (operating system installation and updates, firewall settings, shared network storage, etc.) which requires a certain level of expertise.
- Underutilization of resources (researchers need to spend a certain amount of time to prepare computations and analyze results, moreover, they do not always need all the resources of modern multicore CPU servers).

So taking into account the drawbacks listed above the JINR cloud team has developed a service which provides small research groups from JINR and its Member State organizations with access to computational resources via a problem-oriented (i.e. application-specific) web interface.

A motivation behind the JINR cloud infrastructure and the HybriLIT heterogeneous cluster deployment was to provide users with access to more powerful computational resources as well as to increase hardware utilization.

Normally a user interacts with the JINR cloud via a web interface which allows to manage virtual machines (VMs) within user quotas and use these VMs as intended. Resources of the HybriLIT heterogeneous cluster are available for researchers via a command line interface from submit machines of this complex. User interactions with both computational components mentioned above can be simplified by creating application-specific web forms where a user just needs to set a proper value for the parameters of a certain application, select a computational backend, define the amount of required resources and specify a location for job results. An implementation of such an approach transforms the JINR cloud and the HybriLIT cluster from the Infrastructure-as-a-Service platform (IaaS) into the Software-as-a-Service (SaaS) platform – the JINR service for scientific and engineering computations (or JINR SaaS for short).

The main components of the service as well as a basic workflow schema are shown in Figure 38 and described below.
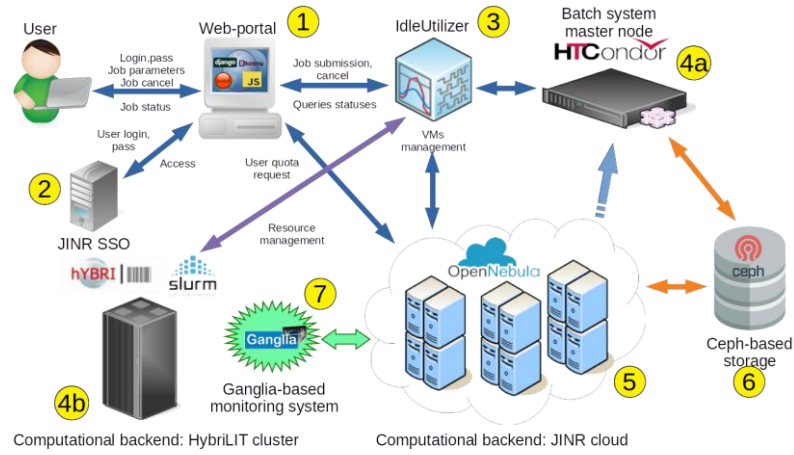
Fig.38. Basic workflow schema and main components of the JINR SaaS service.

A user interacts with the service via a web portal accessible only at URL http://saas.jinr.ru (this component is marked by the digit "1" in the yellow circle in Fig.37). It is a client-server web application. Its server part is based on the Django framework, which implements the so-called "Model-Template-View" architecture. The web application communicates with the IdleUtilizer (IU) component and the JINR cloud via remote procedure calls (RPC) encoded in the XML format (i.e. via "XML-RPC").

A client part of the service with which a user interacts via a browser was developed using the software stack based on Bootstrap 3, AJAX, JavaScript, HTML5 and CSS. The given part allows one to select a particular application and a computational backend, define input parameters for the application and the amount of requested resources (it also validates them) as well as to submit a job. Apart from that, it displays information about available and busy resources within user quotas. The server part of the web application receives such information and forms a request to the IU on its basis.

The web portal provides an opportunity to cancel a job after it was submitted.

A user authenticates on the web portal via the JINR Single Sign-on (SSO) service (marked by the digit "2" in the yellow circle in Fig. 37). Such an approach mitigates the risk for access to third-party sites (user passwords are not stored or managed externally), reduces the password fatigue from a different username and password combinations, decreases time spent re-entering passwords for the same identity as well as reduces IT expenses due to the smaller number of IT help desk calls and tickets concerning passwords. One of the main advantages of using SSO-based authentication is a possibility of getting access to the JINR services using a uniform authentication service.

The IdleUtilizer (marked by the digit "3" in the yellow circle in Fig.37) is a Python script which either deploys on the JINR cloud virtual machines becoming worker nodes (WNs) of the HTCondor-based virtual cluster for incoming jobs' execution or submits these jobs to the HybriLIT Slurm-based cluster depending on the computational backend. Initially the IU was developed to increase an overall efficiency of the JINR cloud facility utilization at the cost of loading its idle resources by jobs from the HTCondor batch system used by the NOvA experiment. However, when a necessity to develop the JINR cloud service for scientific and engineering computations appeared, it was decided to extend the IU functionality.

In this context, the IdleUtilizer serves the following main tasks:
- it gets information defined by a user for computational jobs' execution;
- it manages VMs in the cloud (instantiates or terminates VMs, checks their status, etc.);

- it handles user jobs in a batch system – HTCondor or Slurm (prepares and submits a user job, checks its status, cancels the submitted job upon a user request, etc.).

All the IU elements were designed to be flexible, easy to be tested and modified in case of changes in the interfaces of external services. The IU supports OpenNebula-based clouds [34] as well as HTCondor- and Slurm-based batch systems.

HTCondor is a specialized workload management system for compute-intensive jobs. HTCondor provides a job queuing mechanism, scheduling policy, priority scheme, resource monitoring and resource management.

Each separate user job initiates a creation of a dedicated set of computational resources – virtual machines in the JINR cloud infrastructure. The given VMs become worker nodes for the HTCondor-based batch system which master node (marked by the digit "4a" in the yellow circle in Fig. 38) is deployed in the JINR cloud as well.

The HybriLIT heterogeneous cluster (marked by the digit "4b" in the yellow circle in Fig.38) is a computing component of the JINR Multifunctional Information and Computing Complex which contains computational nodes with NVIDIA graphical accelerators and Intel Xeon Phi coprocessors apart from multicore CPUs.

Resources of the HybriLIT cluster operated by the Slurm workload manager can be used as another computational backend for user jobs coming from the SaaS web portal.

Ceph-based software-defined storage also plays an important role by providing shared storage space for all HTCondor nodes, which is used by parallel jobs for writing the job output simultaneously as well as keeping job results to be retrieved by users after job completion.

In order to get information about various metrics (CPU and memory usage as well as disk and network i/o) of HTCondor WNs, which are deployed and run on cloud VMs, a ganglia-based monitoring system is used (marked by the digit "7" in the yellow circle in Fig.37). One of its advantages over other similar systems is its topology-agnostic. It means that nodes monitored by ganglia-agents (called gmond) can dynamically appear and disappear without a necessity to restart core services on the ganglia head node, i.e. a topology of the monitored system can be dynamically changed. It is exactly the case when HTCondor WNs are created on cloud VMs upon user requests by submitting a job via the JINR SaaS web portal.

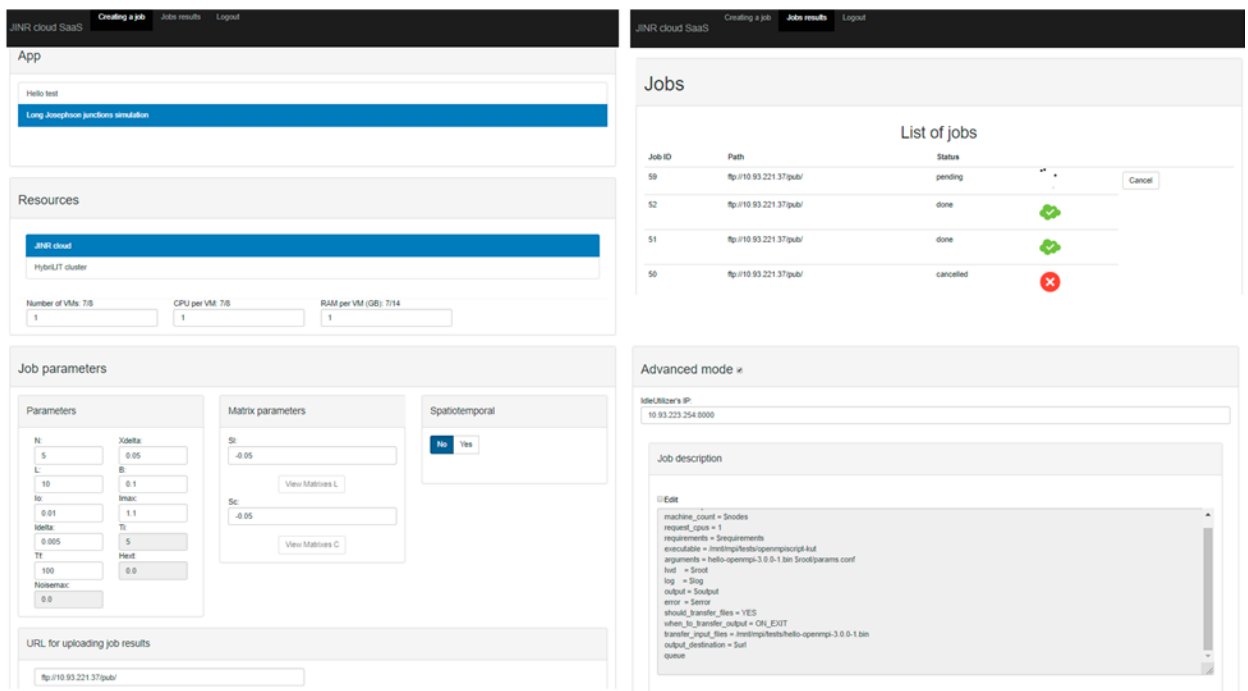Several screenshots of the JINR SaaS web portal are shown in Figure 39.

Fig. 39. Screenshots of the JINR SaaS web portal.

Smart cloud scheduler

The complexity of modern software libraries and applications makes it hard to predict possible workloads generated by software, which may lead to significant underutilization of hardware in the Infrastructure-as-a-Service (IaaS) clouds. In order to increase the efficiency of JINR cloud resource utilization the JINR cloud team developed an approach including the software framework development and an example algorithm aimed to deal with resource underutilization in cloud environments.

One of the most commonly used techniques to reduce underutilization of allocated resources is to "overcommit" VMs, i.e. to allocate more virtual resources than physically available ones. This approach has two major drawbacks. Firstly, workloads generated by VMs are dynamic, i.e. changing with time; therefore, applying this technique may lead to hardware overload. It results in a number of negative effects, ranging from reduced performance of VMs to their total failure (depending on the hypervisor in use and its configuration). Secondly, it is possible that there will be periods when multiple VMs generate low load, still resulting in high underutilization of resources. To reduce these negative effects and further improve hardware utilization, an overcommitment technique can be combined with the dynamic relocation of resources.

A live-migration technology makes it possible to transfer running virtual machines between different servers with minimal downtime and preserve the state of all running processes, so that the migration process is barely noticeable and has a minimal impact on VM functionality. This technique allows one to offload some virtual machines when hardware overload occurs (or prior to it when load reaches defined thresholds). In addition, it can be used to dynamically track cloud workloads and relocate VMs to keep resource utilization higher even under dynamically changing workload conditions.

There are several different algorithms that manage dynamic workloads within cloud environments using overcommitment and live-migration technologies. A simple heuristic algorithm based on the JINR cloud workload profile was implemented.

The analysis of JINR cloud workloads showed that VMs could be split into two classes: low active and highly active, which we will refer to as "cold" and "hot" VMs respectively. Based on this, the algorithm takes four main stages:

1) VM classification. In case of the JINR cloud, the classification is based on the VM CPU load and memory usage only, assuming that the other major characteristics, such as disk and network I/O, can be safely ignored. Figure 40 shows an example of such classification for one of the JINR cloud cluster.
2) Server classification. In this stage we define the amount of physical resources required to accommodate VMs of each class and assign each server to the corresponding class.
3) Building of a migration map. Choosing which VMs to migrate to which servers can be treated as a bin-packing problem. A modification of the Best-Fit Decreasing (BFD) algorithm may be used to handle the task.
4) VM migration. The migration process itself consumes computing resources: during every migration, snapshots of memory (and in some cases VM data disks, depending on the migration scheme) should be transferred over the network between two servers. If only the small number of VMs has to be migrated, all migrations can be performed at once. However, when the large number of migrations is required, spreading them in time and doing the small number of simultaneous migrations becomes reasonable.
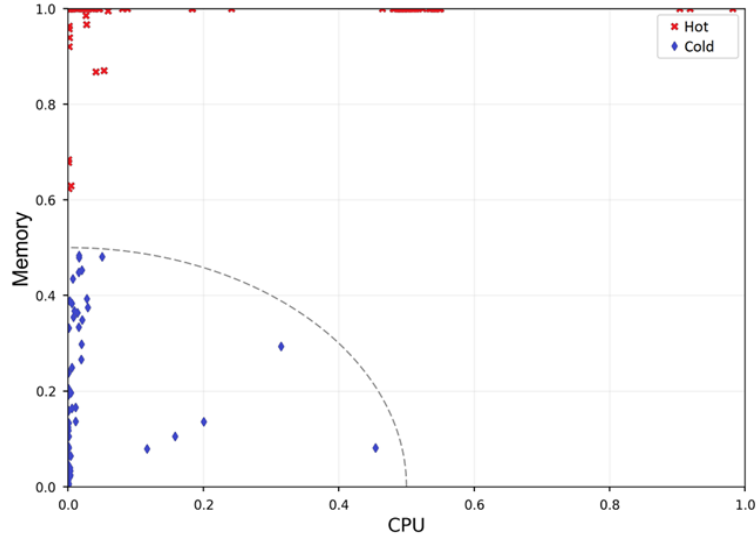
Fig. 40. Example of VM classification in the JINR cloud. The X-axis shows the real CPU usage of VM as a fraction of the total allocated CPU for VM, the Y-axis shows the real memory usage by VM as a fraction of the total allocated memory for VM.

By giving "cold" servers higher overcommitment thresholds and lower (or zero) values to "hot" server thresholds, we make the overcommitment more reasonable and safe. Cycling through four main stages forms a simple heuristic algorithm, the detailed information on which can be found in one of the publications [32]. A schematic representation of this process is illustrated in Figure 41.
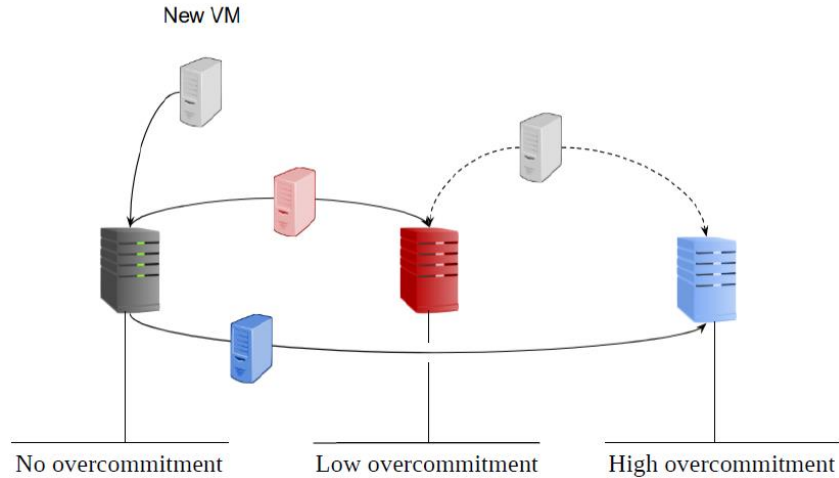
Fig. 41. Schema of the dynamic relocation algorithm.

To implement the described algorithm, a software framework was developed by the JINR cloud team. The framework acts as a meta-scheduler complementing the built-in cloud scheduler, leaving the latter the responsibility of allocating the initial VM. Figure 42 shows an overview of the framework and includes the following main components:

- A monitoring system gathers required metrics and stores them in the time-series database. It uses Icinga 2 for monitoring, InfluxDB for storing and Grafana for visualizing data.
- A custom Linux daemon that runs the scheduling algorithm.
- A library, which implements basic cloud operations and interaction with the monitoring system. Currently, it implements only the OpenNebula XML-RPC interface.
- An additional relational database for storing extra meta-information needed by algorithms, such as host and VM classes.

The framework is built completely on open-source systems, so it can be modified freely to suit any needs.
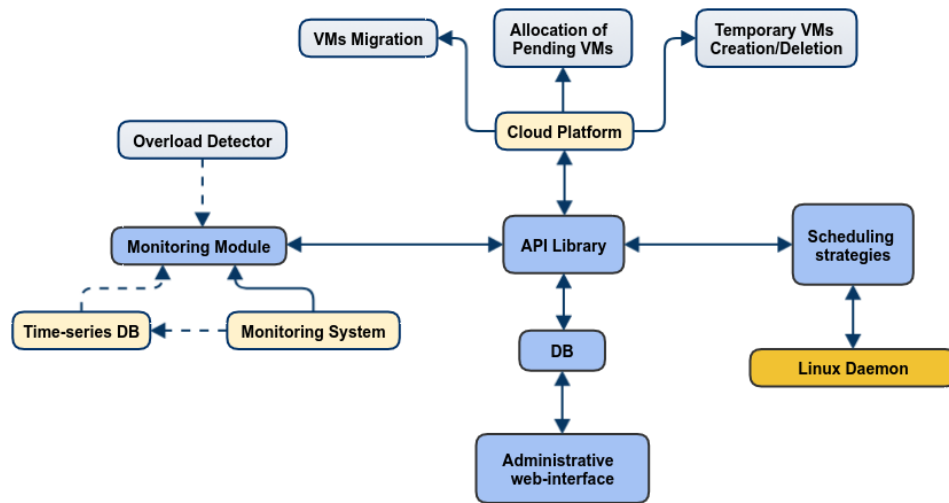

Figure 42. Main components of the software framework.
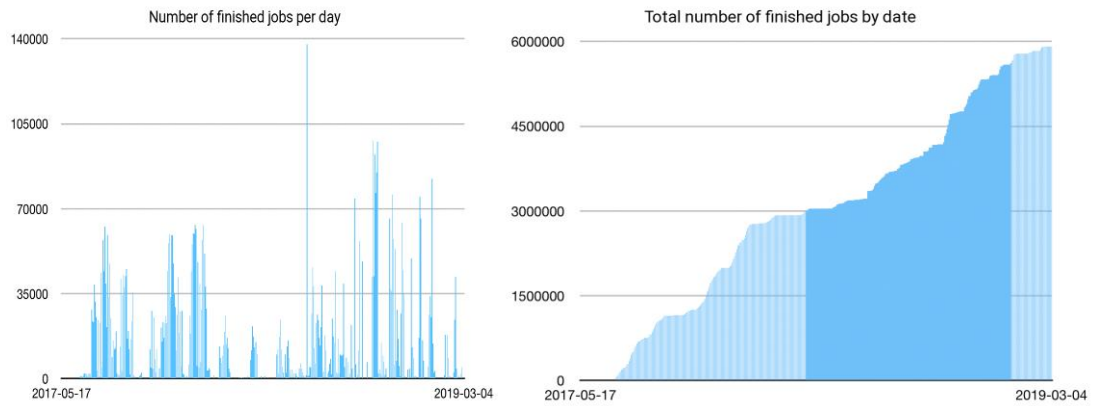
COMPASS Grid Production System

Another set of services deployed in the JINR cloud during the report period forms the «COMPASS Grid Production System». It was developed to provide an automated task processing from defining jobs to archiving them. The whole production management was done via Web UI which allows to define a task, send, follow and manage it during its lifecycle.

Jobs are sent via the PanDA job execution layer to any available type of computing resources such as HTCondor, LSF, PBS, etc. For the time of writing the report the following computing sites have been able to accept COMPASS jobs from the given system: CERN HTCondor, JINR Cream CE, BlueWaters HPC as well as storage resources of EOS and CASTOR at CERN where experimental data are kept.

The COMPASS Grid Production System has been in production since August 2017. Since then the following statistics has been reached: ~3 000 000 chunks of raw data processed, ~80 billions of events, ~500 TB of data produced, ~6 000 000 jobs processed (reconstruction, ddd filtering; merging of mDST, histograms and event dumps).

The number of finished via COMPASS Grid Production System jobs per day (on the left side) and the total number of finished jobs by date (on the right side) are shown in Figure 43.

Fig.43. COMPASS Grid Production System statistics on finished jobs: the number of finished jobs per day is on the left side, the total number of finished jobs by date is on the right side.

Multi-level monitoring system for the Multifunctional Information and Computing Complex at JINR

The multi-level monitoring system was created for the JINR MICC. It is based on different technologies such as Nagios, Icinga2, Grafana, and the systems which were developed from scratch in JINR. The Nagios/Icinga2 system proved to be suitable and reliable tools for hardware monitoring of all the MICC components. Grafana as a visualization tool is used only by the JINR Cloud team, however, it can be useful for other components. The Tier-1 service monitoring system and the HybriLIT monitoring system were developed in JINR. In order to provide the secure accessibility to all monitoring and control systems mentioned above, the MICC Operational Center was thoroughly designed and created. It is a reliable place where operators can effectively control the MICC in a wide range of critical situations like a global power cut or a network failure.

Our experience shows that modern open source monitoring systems like Nagios are reliable and useful. These systems can monitor critical metrics, reliably provide notifications and dashboards for administrators. That is why they can be used for the center wide monitoring of hardware. However, sometimes they are not flexible enough to be suitable for all groups of monitoring consumers, especially when a complex visualization is required. It leads to new software development, which allows providing a required complex system.

One of the crucial requirements to any monitoring or control system is its reliability. It can be satisfied by the system itself. There is also another important requirement, that is availability. To satisfy it, special arrangements are required to ensure stable access to monitoring and control systems during power cuts, network failures or other emergencies. The MICC Operational Center was formed to provide 24x7 availability of MICC systems [24]. The MICC Operational Center is located at a room dedicated for MICC operators. At present, 1316 elements of the computing complex are monitored and more than 10000 checks are performed in real time. The interface of monitoring systems for the JINR CMS Tier-1 and JINR Tier-2 sites is presented in Figure 44.
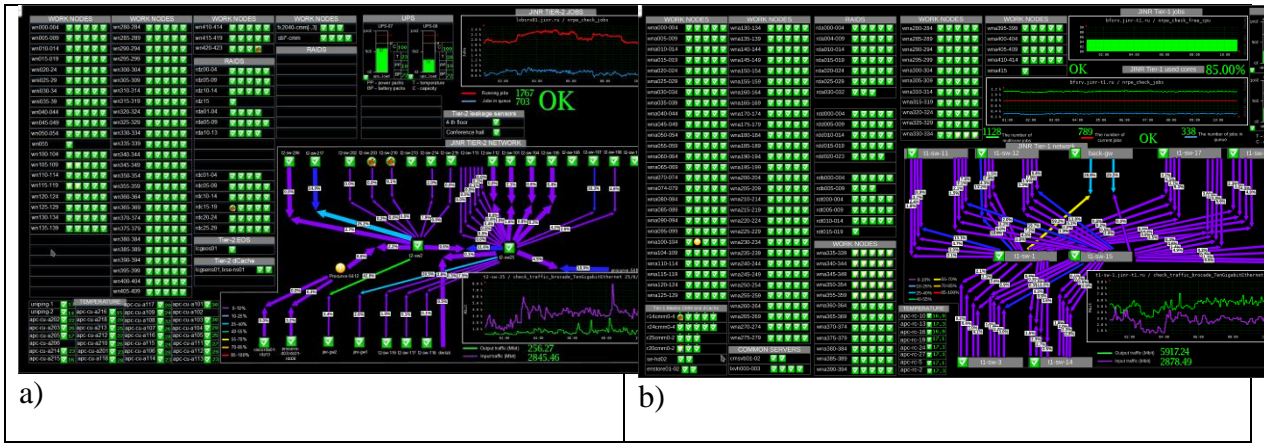
Fig.44. Interface of the monitoring system of the MICC Tier-1 (Fig.8 a) and Tier-2 (Fig.8 b) elements.

Apart from infrastructure monitoring, there is a need for consolidated service monitoring. Top-level services, which accept jobs and data from the grid, depend on lower-level storage and processing facilities relying on the underlying infrastructure. There is a variety of sources of information about the status and the activity of Tier-1 services. The decision to develop a new monitoring system was made (Fig.45). Its goals were the following: to retrieve monitoring information about services from various sources, process data into events and statuses and react according to a set of rules, e.g. to notify service administrators or restart a service.
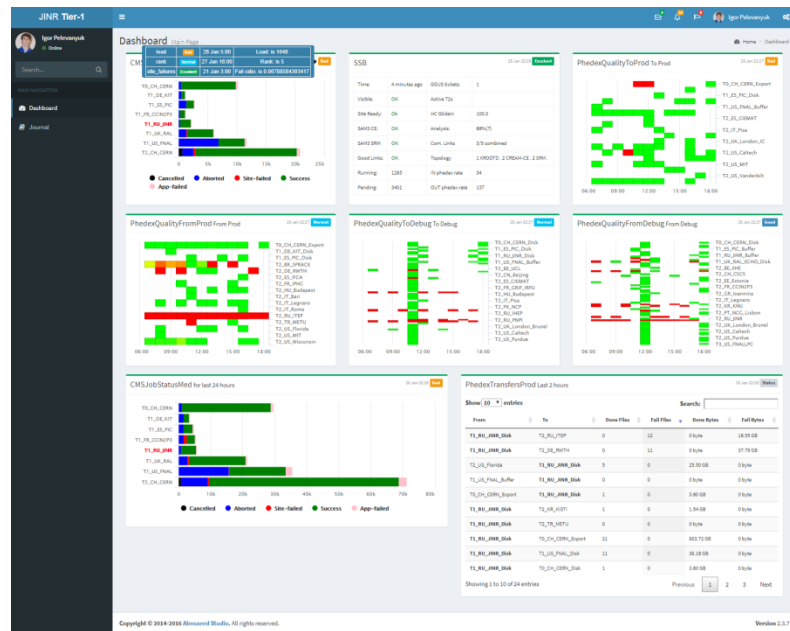
Fig.45. Tier-1 service monitoring web interface.

"GOVORUN" supercomputer monitoring

For effective work organization there is a need in a service that collects information about the use of computing resources, the status of virtual machines and related services, network traffic and user activity.

To increase the "GOVORUN" supercomputer performance, the following components are required:

— Supercomputer's accounting system – helps top managers and system administrators to evaluate the supercomputer performance;

— Nodes utilization and services monitoring system – helps users to optimize the use of supercomputer nodes and improve their own code;
— Servers monitoring – notifies system administrators of crash statistics.

The supercomputer's accounting system and the servers monitoring are parts of the LITMON project implemented with the help of the ICINGA system, for which scripts (checks) for the necessary information acquisition with its further presentation in a convenient graphical form were written. The supercomputer has its own closed subnet, where computing nodes and servers with storages are located. That is why it is impossible to acquire data from each node of another subnet. For this, the HLITMON virtual machine with access to the supercomputer subnet is used.

The list of checks:

Node status check (UP, DOWN). If the status changes, a notification is spent to system administrators.

Slurm status check – the node status check in the SLURM job scheduler.

SMART check – the status check of HDD/SSD on the nodes.

File systems check – the status check of network file systems according to certain parameters: availability of the network file system, the total volume of disk space, how much disk space is used, how much is available for recording disk space.

Accounting check – the queue status check: the number of running jobs for all queues, the number of completed jobs for all queues, detailed information about the load of each queue (skylake, knl, dgx).

Creation of the off-line computer complex for modeling, processing, analysis and storage of data within the NICA project

The distributed information and computing cluster of the NICA complex in its basic configuration should provide the processing and storage of up to 10 petabytes of data per year. The complex consists of territorially distributed on-line and three off-line clusters connected by the high-speed computer network with a bandwidth 4 x 100 Gbit/s. The prototypes of all clusters are implemented.

The equipment of the central telecommunication node, i.e. the core of the switching and routing system of on-line and off-line clusters, is implemented on four multifunctional switches Cisco Nexus 9504 with the full-mesh topology for maximum reliability and performance. Figure 46 illustrates the implemented computing network schema of the distributed multi-site cluster of the NICA complex.
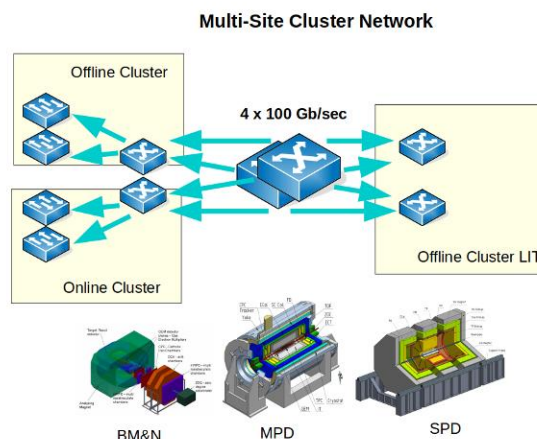
*Fig.46. Scheme of the multi-site cluster network.*

The off-line computer cluster of the NICA complex, as all large-scale centers of experimental data processing and storage, is territorially distributed and combines all the components located in VBLHEP and LIT in a unified local multi-site computer network with 4 x 100 Gb/s.

Figure 47 illustrates the storage systems of the LIT off-line cluster (data storage of the LIT cluster on the left and ultra-fast disk memory on the right). The LIT cluster prototype includes up to 500 CPU and 2,5 PB of disk and tape memory.
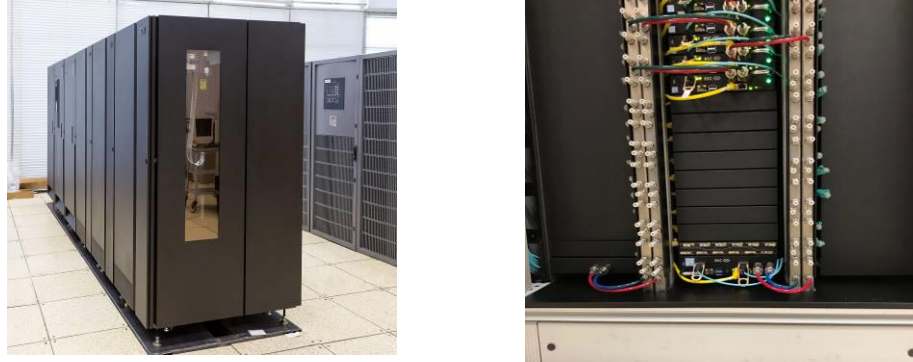
Fig.47. LIT data storage (left) and ultra-fast disk drives of the LIT off-line cluster (right).

Another component of the off-line cluster for NICA is an ultra-fast data storage system (UDSS) implemented in the GOVORUN supercomputer under the Lustre file system (Fig.47). At present, UDSS has five storage servers with 12 SSD drives with the NVMe connection technology, which reduces the time of access to data. The total capacity of UDSS is currently 120 TB and the reception/yield rate of data is 30 TB per second. It is noteworthy that UDSS has the ability to linearly increase productivity (speed of working with data) and the storage volume without changing the principles of the architectural design of the system. At the same time, the hyperconvergence and software-defined architecture of UDSS allows providing the maximum flexibility of data storage system configurations, including creating data storage domains, which meet the requirements of specific tasks, as well as connecting external disk fields to expand and develop the UDSS capacities.
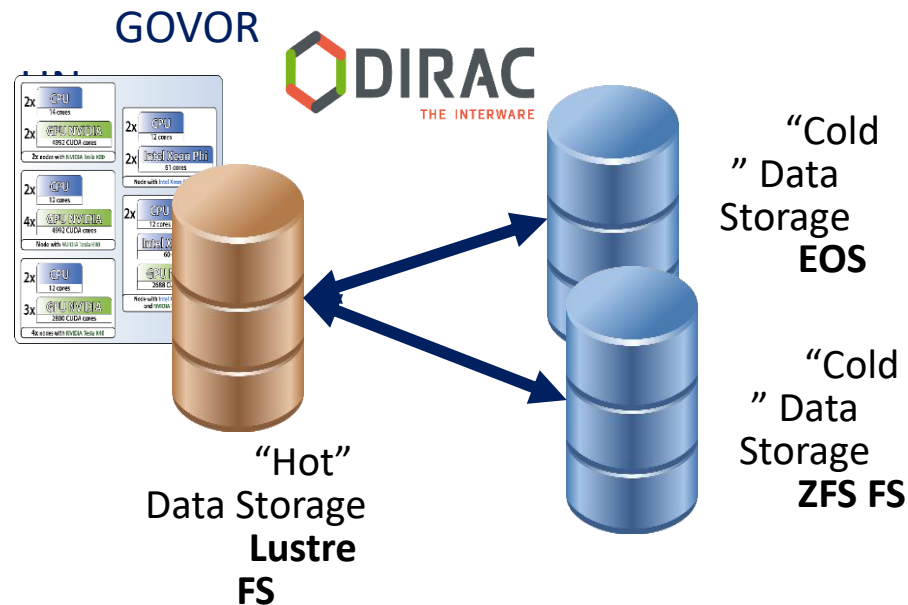

Fig.48. Scheme of using the GOVORUN supercomputer for modeling the computing for the NICA megaproject and event modeling for the MPD experiment.

At the first stage of modeling the computing for the NICA megaproject, it is supposed to use computing resources of the GOVORUN supercomputer, the Skylake and KNL partitions to generate data, save them on UDSS and transfer to cold storages managed by the EOS and ZFS file systems (Fig.48). It will allow checking the basic stack of data storage and transmission technologies as well as modeling data flows, choosing optimum distributed file systems and increase the efficiency of event modeling and processing.

It is also planned to use the supercomputer resources including UDSS for modeling of events of the MPD experiment, moreover, it is supposed to use the DIRAC software for managing jobs and the process of data read-out/record from various types of storages and various types of file systems (Fig.48).

Optimization of the Distributed Data Processing System for the NICA BM@N experiment by Using Simulation

A typical scheme of a hierarchical distributed computing grid infrastructure arranged in tiers (Korenkov, et al., 2015) can be used in the NICA project. The planned data storage and processing scheme for the BM@N experiment is shown in Figure 49.
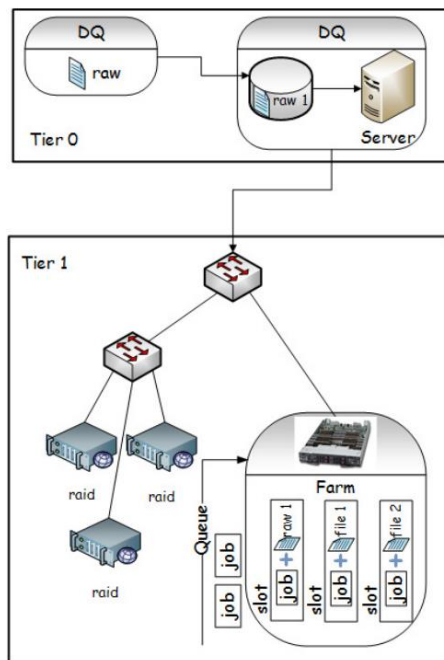
Fig.49. Data storage and processing scheme for the BM@N experiment.

The Tier0 module situated at JINR VBLHEP performs data acquisition. Raw data are transmitted to the Tier1 module via a communication link. The Tier1 center is located at the JINR Laboratory of Information Technologies (LIT). According to experimentalists' assumptions, the total BM@N data flow is expected to be about 1 PB per run. The data will be written to RAID (Redundant Array of Independent Disks) and then processed on a computer farm. Data accumulation and processing are done simultaneously.

The number of events in a file is restricted due to the convenience of further manipulations. Each second two files are sent from BM@N. 5% of events in each file are events for one express analysis job. Each event requires 500 MB of supplementary files.

Simulation Results

A simulation of 270 minutes of the BM@N run shows that the maximum load in respect of the communication link is 600 MB/s. Figure 49 (left) illustrates the load through the communication link between the BM@N experiment and LIT. By now a 10 Gb/s channel between BM@N and LIT already exists. Therefore, the bandwidth of the channel is sufficient to transfer the full amount of BM@N data to LIT disk storage.

The LIT computing farm contains 1000 cores for job processing. Figure 50 (right) shows the number of free cores during processing express analysis jobs, which results in about 880 cores being free. It means that if the rate of data accumulation is constant, 120 cores are enough for the express processing and we can occupy 880 cores for the full processing, which is sufficient as well.
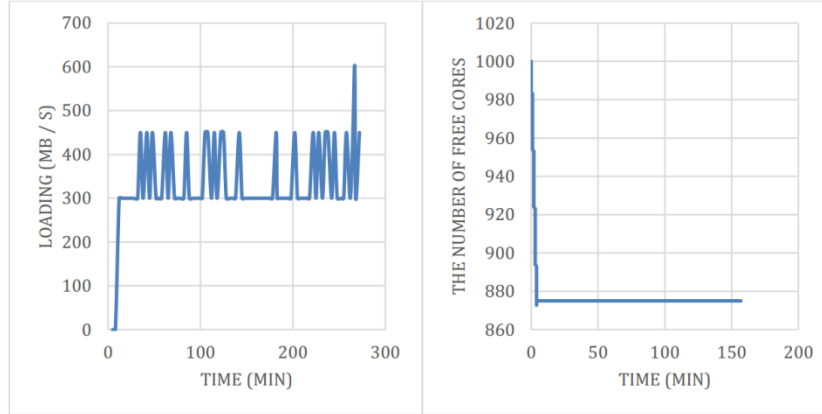

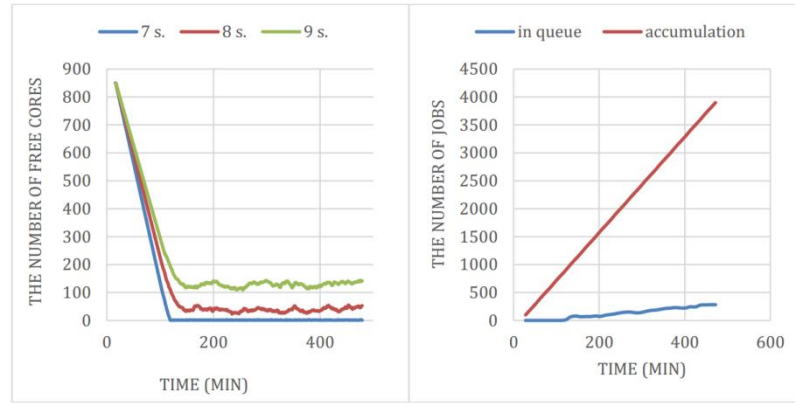*Fig.50. Left — Load of BM@N – LIT communication link. Right — Express analysis processing on the LIT farm.*

Fig.51. Left — Reconstruction jobs processing at different run rates on 880 cores. Right — Jobs arrival. Run rate – 7s.

Figure 51 (left) shows the number of free cores on the LIT farm during the reconstruction jobs processing on 880 cores at different rates of job processing runs. It is a simulation situation of all events being fully processed with the simultaneous receipt of events. The rate of job processing run rates is the number of seconds between processing each piece of data. If one job runs every 9 seconds (the green curve in Fig.51 (left)), more than 100 cores will be free and the farm will not be used fully. If one job runs every 7 seconds (blue curve), than the farm will be loaded fully, but the simulation shows that there will be no sufficient time to process all jobs due to overload. Figure 51 (right) illustrates a jobs queue emerging in this situation. The red curve shows the number of accumulated jobs and the blue one — jobs in the queue waiting to be executed. Long runs of the experiment may lead to a shortage of the resources. However, slowing down the rate of the jobs arrival (e.g. the red curve in Fig.51 (left)) or speeding up job processing will help to avoid this problem.

BM@N experimentalists propose the experiment duration as 1500 hours. The first 120 hours are assigned for the equipment adjustment, while experimental data will not be recorded. Then, for the next 1000 hours, data will be accumulated and transferred to the LIT disk storage. The data flow intensity will be 70% of its full possible value. The next 380 hours data will be accumulated with 100% intensity. The express processing of jobs will occupy 120 cores and the full processing — 880 cores simultaneously with data accumulating. After the data accumulation will be completed, i.e. at the end of the experiment run, the full processing will be done on 1000

46

cores. Physicists are going to process either 100% or 50% events from the file, i.e. either each event or each second event respectively.

Two possible scenarios of event processing were simulated. The first one involving 100% of events to be processed and the second one — 50% of them. The average processing time of an event is one second. We suppose about 2500000 files will be received from one run of the experiment. In both scenarios online event processing and data accumulation are done simultaneously. Therefore, the farm does not have enough time to process all events. If physicists process each event from files (purple curve), an events queue will emerge. Therefore, the processing of 100% events on 1000 cores would require 2000 hours to finish processing unprocessed events after the experiment run is completed. In case of the second scenario, when each second event will be processed on-line, the final processing requires 500 hours only.

In frames of the creation of a system for storing and processing data from the BM@N and MPD experiments included in the NICA complex, a new macro-modeling schema was developed and implemented as a program using a probabilistic approach to evaluate various equipment configurations which determine the probabilities of loss of information arriving from the detectors for each of these configurations, and the one for which this probability does not exceed the specified limit and the price is minimal, is selected taking into account economic factors. The program was tested on calculations of the efficiency of the data acquisition and storage system of the BM@N experiment from price investments in the system of the disk memory of intermediate data storage [53].

Applied research on MICC

LIT developed a set of interconnected cloud services and tools for managing and processing biomonitoring data allowing to simplify and automate the stages of monitoring starting from selecting sites for collecting samples and ending with generating maps of pollution distribution and predicting environmental changes. This is the first solution in Russia and Europe that allows all specialists involved in the biological monitoring network to get access to data on environmental pollution. In 2018 the development of the architecture and technologies of the web portal moss.jinr.ru was in progress. The ability to compare the concentrations of heavy metals in samples of different years in the form of diagrams as well as statistical calculations was added. The module for the comparison of concentrations in countries and regions was optimized. The module for calculating geo-indexes was improved – now it is possible to set the level of background pollution. The possibility of creating pollution maps by users of the platform was expanded and the ability to import data into Google Earth was added. Maps and statistical calculations for the atlas 2015-2016 were formed. An alpha-version of a mobile application for registering data on sampling sites and uploading them to the platform was developed. The application will reduce the number of errors when filling out forms and provide the potential to verify the correctness of the definition of moss type by participants through the use of machine learning methods. More than 6000 sampling sites from 40 regions of different countries are presented in DMS now.

The web interface of DMS is available at URL http://moss.jinr.ru, and several screenshots are shown in Figure 52.
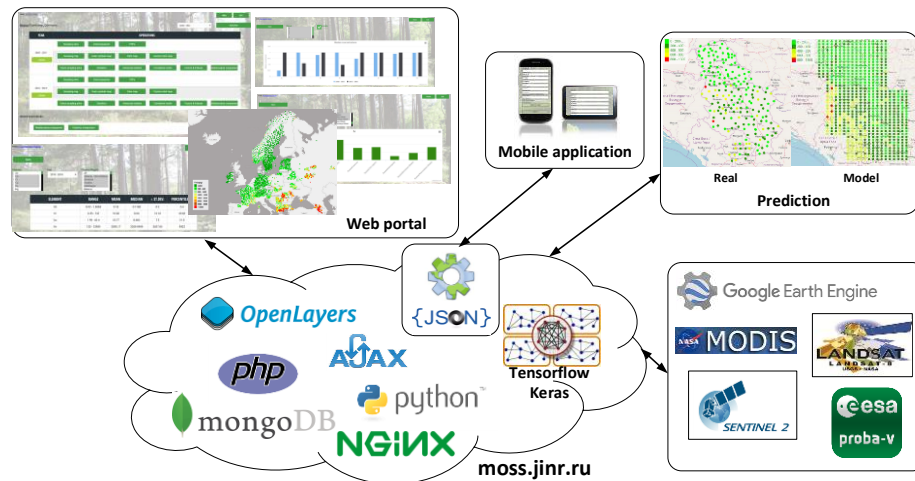
Fig.52. Architecture and technologies of moss.jinr.ru.

The increasing number of smartphones and advances in the deep learning field open new opportunities in the crop disease detection. The aim of the plant disease detection service (PDDS) is to facilitate the detection and prevent diseases of agricultural plants by both deep learning and programming services. PDDS is a multifunctional platform that uses modern organization and deep learning technologies to provide a new level of service to the agricultural community.

The PDDS architecture (shown in Fig.53) was developed. Some of its components, including a web portal available at URL http://pdd.jinr.ru, are deployed at the JINR cloud. The service has a basic detection functionality and provides an opportunity to download the self-collected database of grape leaves consisting of five image sets. With the help of the deep Siamese convolutional network it became possible to achieve 99% accuracy in the disease detection.
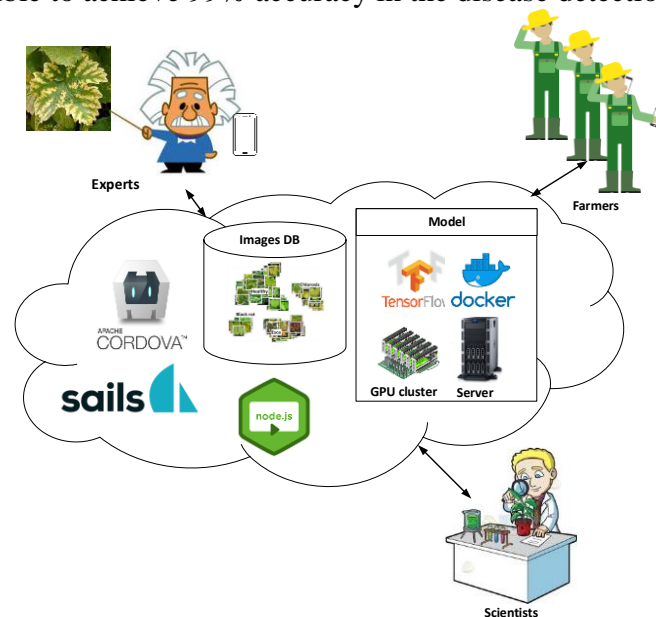

Fig.53. Components of the plant disease detection service.

In frames of the RFBR (Russian Foundation for Basic Research) project the methods and the tools for using high-performance computing infrastructures and software applications for processing textual and graphical information about plant diseases to minimize losses in agriculture were developed [33]. A platform architecture (Fig.53) for identifying plant diseases using modern software and organizational technologies was proposed. A web site of the project (http://pdd.jinr.ru/) was implemented. A special model of the deep convolutional network with a

Siamese architecture allowing one to evade the problem of a small sample was developed. The recognition accuracy on the test sample of images was more than 95%.

During the report period work on refining the software package for modeling of distributed data storage and processing systems SyMsim (Synthesis of Monitoring and SIMulation) designed for grid and cloud data storage and processing systems was in progress. In 2017 the program was first used for the interprocessor interaction simulation during performing HPC applications in the cloud based on the MPI implementation of the program for calculating the current-voltage characteristics of long Josephson junctions. The comparison of empirically obtained results with simulation results showed that the simulation model correctly simulated parallel calculations performed using the MPI technology and proved that for quick computing of the given problems one should simultaneously increase the network bandwidth and the number of processors. The results illustrated that the SyMSim program could be successfully used to estimate the performance time of MPI algorithms in the cloud environment taking into account interprocessor connections. In frames of this problem a service with a user interface for modeling of long Josephson junctions was developed; calculations are performed on virtual machines dynamically created in the cloud and connected as working nodes of the HTCondor batch system.

Work on refining the software package for modeling distributed data storage and processing systems (SyMSim) with the aim of using it to improve the topology of the computer network and the performance of the data center of the Institute of High Energy Physics of the Chinese Academy of Sciences was carried out. The parameters from the database of the computing infrastructure of the Institute as well as some data from the BESIII experiment were used as input data for simulation. The simulation was carried out on three aspects of the Institute's computing configuration (Fig.54). The first results showed that the proposed approach would allow making an optimal choice of the network topology increasing its performance and saving resources.
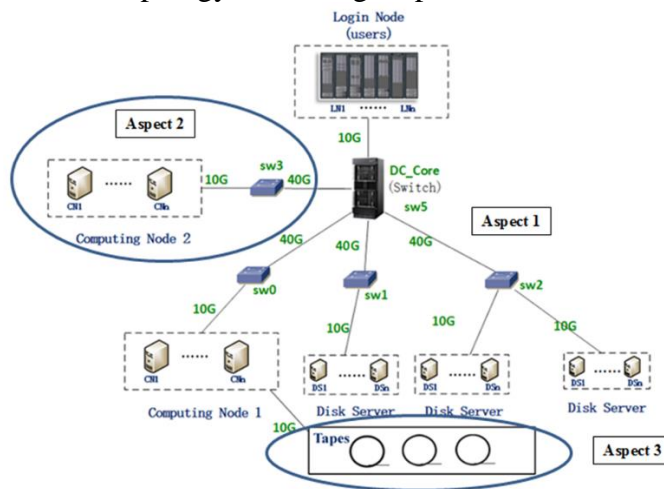
Fig.54. Organization of the computing complex of the Institute of High Energy Physics of the Chinese Academy of Sciences.

In 2018 the state registration certificate on the computer program "Software complex of intelligent scheduling and adaptive self-organization of virtual computing resources" was received. N.A.Balashov, A.V.Baranov, I.S.Kadonchikov, V.V.Korenkov, N.A.Kutovskiy, I.S.Pelevanyuk are the authors. The software complex is designed to optimize the consumption of computing resources in cloud infrastructures and can be used by cloud providers in infrastructures built on the IaaS model on the OpenNebula platform. The complex includes the VM planning strategy based on VM and server rankings, API for interacting with the OpenNebula platform, an administrative web interface, net-snmp extensions and corresponding polling modules for Icinga 2, which allow collecting information on the current use of VM resources based on KVM as well as OpenVZ containers.

A new production system to run COMPASS data processing in a distributed computing environment was developed. Now ProdSys is a production service, being used by the collaboration

to run jobs in Grid. During the last two years, more than 6 million jobs were processed. The infrastructure of the ProdSys consists of ProdSys UI, data management services, PanDA&AutoPilotFactory server, MySQL server at JINR; and CRIC for COMPASS information system, deployed at CERN. The most contributing computing elements are CERN HTCondor and JINR T2. A method of translating from local to distributed processing, formulated during that work, may be also used to enable a distributed data processing of other experiments, which are still processing their data within only one computing center. All the necessary infrastructure was deployed on JINR Cloud service.

Information and software environment for the research and production activities of JINR

JINR CIS

The data exchange between the elements of the JINR integrated corporate information system (CIS) through the universal exchange gateway was established. At present, the following subsystems are involved in the exchange through the gateway: 1C, ADB2, EDS "Dubna", PIN, ISS, Document Base, HR. All the given systems, except for 1C, were transferred to authentication through the SSO (Single sign-on) system. The PIN weak point was discovered and fixed.

EDS "Dubna". In frames of the development of the electronic document system "Dubna" the following documents were developed and put into operation: "Procurement commission protocol", "Delivery agreement", "Capital construction agreement", "Repair construction agreement", "Institute service agreement", additional agreements to these agreements, "Application form for the purchase of personal computers, office equipment, components and consumables", "Procurement plan for PC, office equipment and components", "Protocol of disagreements to the supply agreement". Moreover, the following documents were prepared for commissioning: "Agreement on the paid provision of services for the Institute", "Agreement on execution of research", "Act of completed works", "Advance invoice", "Supplier invoice", "Order on the admission of persons invited to visit JINR", "Application form for admission to RHC (radiation hazardous conditions)". These documents were not put into operation due to some unresolved administrative issues.

A reduced version of EDS adapted for mobile devices was developed and put into operation:

A number of specialized reports for monitoring the timing of the passage of documents and tracking their status were elaborated:
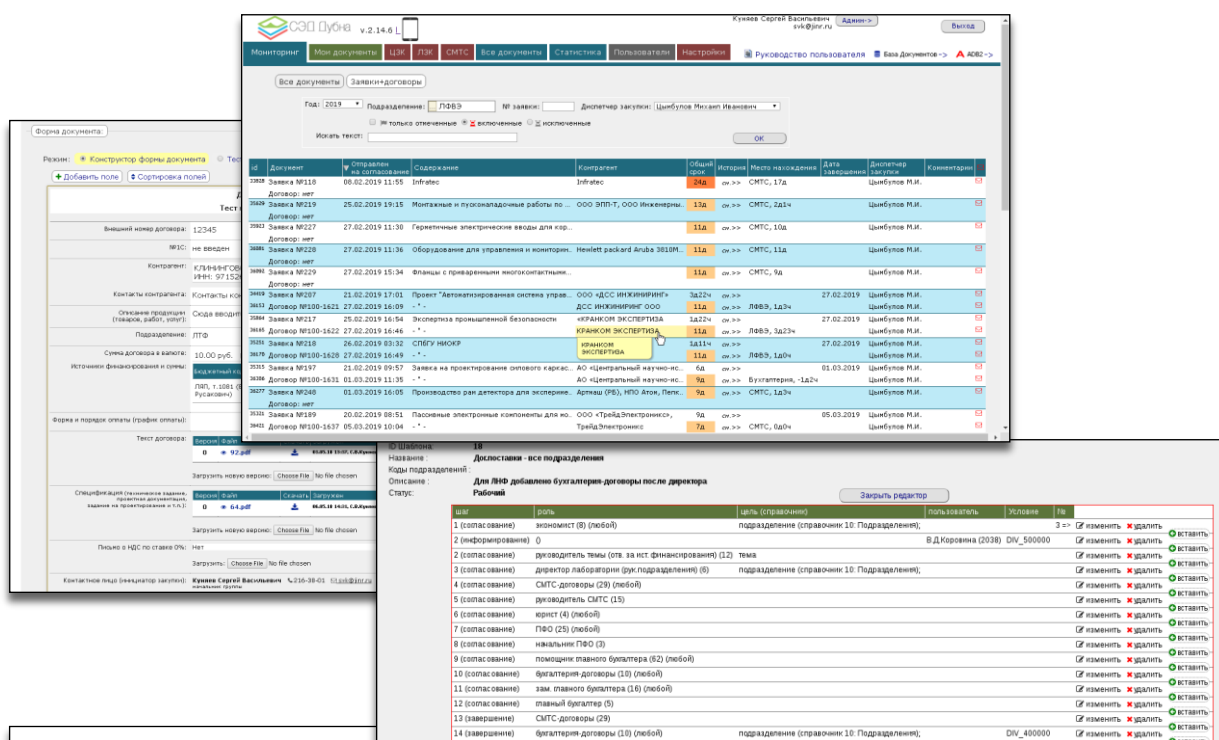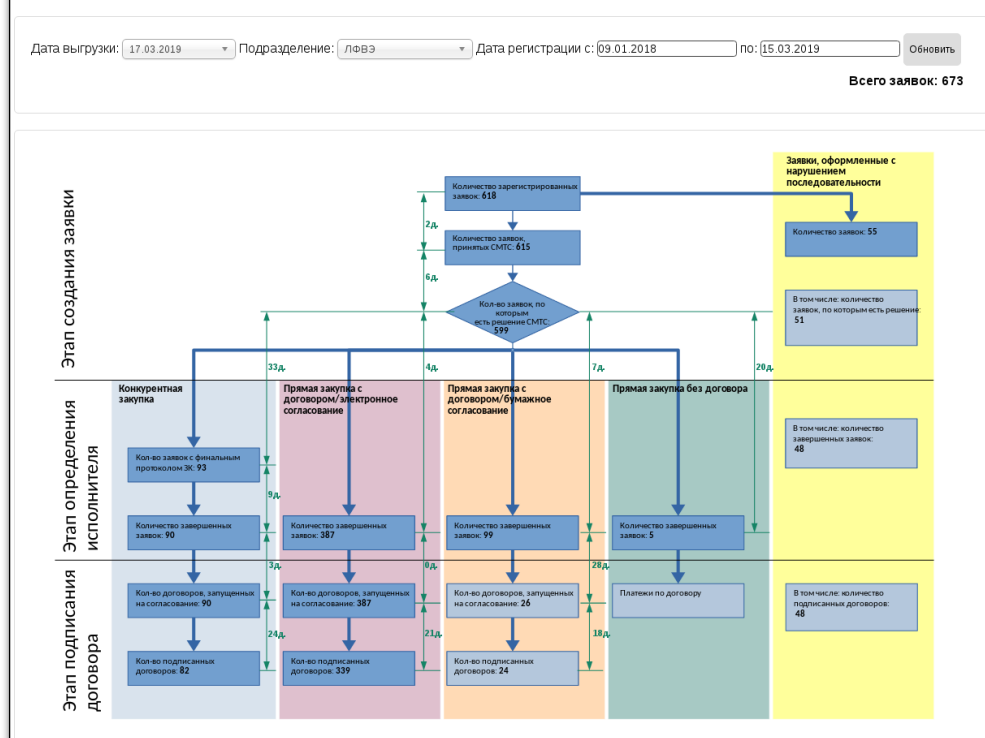
The modules for administering the EDS "Dubna", "Designer of forms of documents", "Coordination route template designer", which significantly reduce the complexity of elaborating new types of EDS documents, were developed:

A subscription module for events during the processing of documents and sending out notifications about them was elaborated.

The system was supported; a number of other works were performed upon user requests.

In 2017 1921 documents were processed in the EDS "Dubna", in 2018 – 3883 documents, within 2,5 months of 2019 – 1004 documents.

For the tasks assigned by VBLHEP, the following works on the development of the project management system **APT EVM for NICA** were carried out:

- A data entry block for forming the Cost Book of the project and reports on it was elaborated:

- A block of the data entry and formation of specialized reports under the agreement with the Russian Federation on financing the NICA megascience project was developed:

ОСНОВНЫЕ ФИНАНСОВЫЕ РАСХОДЫ

на реализацию базовой конфигурации проекта комплекса сверхпроводящих колец на встречных пучках тяжелых ионов NICA

цены 2013г. (млн.руб.)

|	Российская Федерация	Объединенный институт ядерных исследований	Страны-участники	Итого
1. Ускорительный комплекс				
нуклотрон и каналы	0.0	159.5	0.0	159.5
инжекционный комплекс	0.0	337.1	0.0	337.1
бустерный синхротрон	244.4	1,068.9	107.6	1,420.9
коллайдер	2,036.0	1,029.1	18.0	3,083.0
Итого:	**2,280.4**	**2,594.6**	**125.5**	**5,000.5**
в том числе:				
опытно-конструкторские работы	227.9	107.4	0.0	335.2
капитальные затраты на создание	2,052.5	2,487.2	125.5	4,665.3
2. Экспериментальные установки				
многоцелевая установка MPD для проведения исследований плотной барионной материи на встречных пучках коллайдера NICA	789.2	1,764.2	0.0	2,553.4
установка BM@N для физических исследований по изучению плотной барионной материи на выведенных пучках Нуклотрона	106.6	438.0	0.0	544.6
установка SPD для изучения спиновой структуры нуклона на встречных пучках коллайдера NICA	544.6	138.5	0.0	683.1
Итого:	**1,440.4**	**2,340.7**	**0.0**	**3,781.1**
в том числе:				
опытно-конструкторские работы	32.4	134.3	0.0	166.7
капитальные затраты на создание	1,408.0	2,206.4	0.0	3,614.4
3. Научно-исследовательская и инженерная инфраструктура				
здания коллайдерного комплекса с инженерной инфраструктурой	889.2	1,894.2	0.0	2,783.3
здания установки MPD с инженерной инфраструктурой	402.6	7.0	0.0	409.6
здания установки SPD с инженерной инфраструктурой	333.6	0.0	0.0	333.6
центр инновационных разработок проекта комплекса NICA - "Центр NICA", с инженерной инфраструктурой	757.9	1.4	0.0	759.3
экспериментальный павильон и зоны для прикладных исследований	378.5	113.6	0.0	492.1
инфраструктуры криогенного комплекса	761.2	811.0	70.3	1,642.5
инфраструктуры энергосберегающих инженерных систем	377.4	340.7	0.0	718.2
Итого:	**3,900.5**	**3,167.9**	**70.3**	**7,138.6**
в том числе:				
опытно-конструкторские работы	1.3	2.0	0.0	3.3
капитальные затраты на создание	3,899.2	3,165.8	70.3	7,135.3
4. Компьютерно-информационный комплекс - центр хранения и обработки данных, сетевой инфраструктуры, информационных сервисов	590.7	79.8	0.0	670.4

- A module of tracking the status of contracts was elaborated.

The system support and other ongoing work were carried out.

At the request of PLS (Procurement and Logistics Service) the MES&CC (Management of Economic Services and Capital Construction), CCD (Capital Construction Department) system **"Document Base"** was adapted for storing the MTS contracts:

Unified information system 1C

Within the report period the development of a unified system 1C 8.2 MEM was in progress; regular support for system users was provided. Currently, 1C includes 5 conferences and 7 databases, about 200 users of the system can read information, and about 300 users can fill information, every month, about 40 thousand new documents are created in the system. The current configuration of 1C and the modules included in the system are shown in the Figure 55.

In particular, in frames of the further development of "1C Manufacturing Enterprise Management" a module for the purchase of computer equipment was created. The module for forming regulated reports of self-supporting units was improved. Also, conference accounting module and secondment module was implemented. In the first half of 2018 the programs "1C:ERP Enterprise management 2", "1C: SPM Salary and personnel management" were purchased, now they are being set up and employees of the 1C group are being trained to use the given products. On the basis of the program "1C: Retail" cash registers were launched at the "Nauka" stadium and the Athlete's House. Creation of a specialized program for the sale of tickets and management of places and events in the house of culture «MIR»
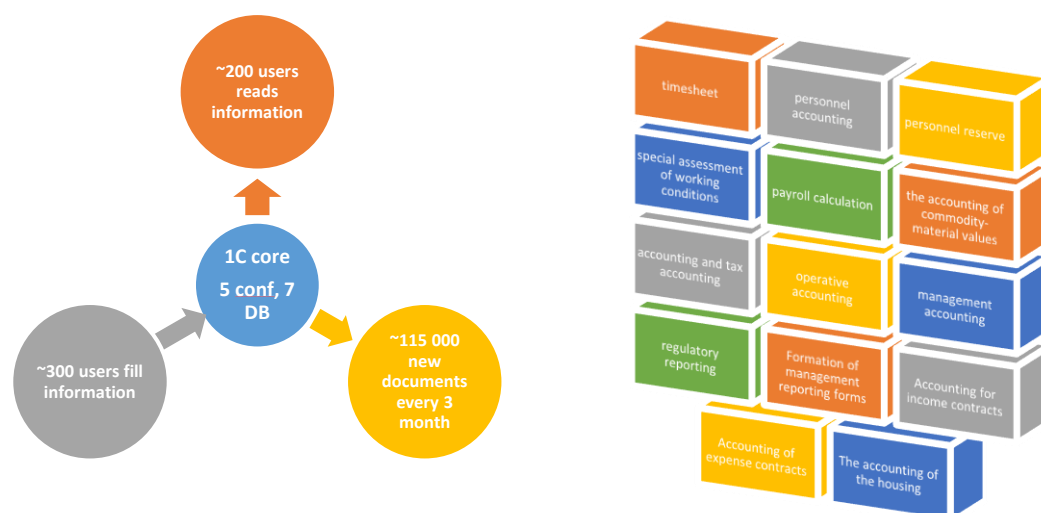
Fig.55. Structure and main modules of the unified information system 1C.

JINR Document Server

In frames of the development of the JINR corporate information system (IS) works on the development of the CRIS&OAR component based on the information systems such as the JINR Document Server (JDS) and Personal Information System (PIN) providing information support for scientific research were in progress. The goal of the given IS is to increase the effectiveness of scientific studies using modern information and computer technologies as well as to create, develop and integrate electronic resources in a unified information environment.

The JDS IS is an Open Access (OA) repository of articles, preprints and other documents reflecting and contributing to research activities underway at JINR. The development of the JDS information services is carried out in frames of the JOIN2 project (Just anOther INvenio INstance) [26] aimed at combining the efforts of partner organizations for the development of the Invenio software platform [25] taking into account their specificity. The project partners are the libraries of scientific centers and universities in Germany, namely, DESY (Hamburg/Zeuthen), the German

Cancer Research Center DKFZ (Heidelberg), the Helmholtz Center for Research of Heavy Ions GSI (Darmstadt), the Institute of Experimental Nuclear Physics KIT (Karlsruhe) and others [25]. The JOIN2 project is one of the largest Invenio users worldwide and open to new partners. Collaborative research in this field is regulated by the DESY-JINR Cooperation Agreement.

In order to ensure effective reuse of data and minimize manual entry, the procedure for controlling authoritative records of the Organization, Division, Grants and Personalities was automated (Fig.56-57) (https://lt-jds-join2.jinr.ru).

Fig.56. Information card of a normative record "Grant".

Authoritative/normative records are search elements of bibliographic records (names of authors and personalities, names of organizations, generic names, thematic terms, geographical names, trade names, unified titles, etc.) presented in accordance with certain rules.

Fig.57. Information card of a normative record "Author".

To solve this problem, a set of API was developed and implemented; it provides the acquisition and syntactic analysis of data on 1) grants with the participation of JINR [45]; 2) authors with JINR affiliation – as well as the formation of files in the marcxml format for batch loading into the system.

Web forms with extended functionality were introduced into the data entry module; it allows automatically importing metadata into publications from external bibliographic databases

for a number of identifiers such as DOI, arXiv, Inspire, WoS (Fig.58-59) and entering metadata on video resources hosted in external repositories. As shown in Figure 58, when entering the correct DOI for an article in the field Import data, the form fields are automatically filled with data from the external bibliographic database. Using the toolbar, a content manager can confirm or edit the information about the authors (Fig.59) and fill in the missing form fields.

The results were presented at the VI Annual Conference of Young Scientists and Specialists Alushta'2017 and at the JOIN2 Workshop'2018.

Fig.58. Entry form for an article.

Fig.59. Window for editing the authors' information.

Automation of STL

With the participation of the LIT staff, work on improving the information services of the JINR STL was in progress. The design of the column "History of JINR and Dubna on the pages of books, journals and national newspapers" was updated on the website of the Science and Technology Library, which contains publications about JINR, JINR laboratories and JINR scientists. The project is replenished and developed, as materials become available (http://who-is-who.jinr.ru/catalog3/main.html) (Fig.60). Furthermore, an information service for forming a literature order for the interlibrary loan and sending it to the STL e-mail was developed and implemented (http://lib.jinr.ru/mba/mba.html) (Fig.61).

Fig.60. Updated web interface of IS "Publications about JINR Scientists".

Fig.61. Web form of a literature order for the interlibrary loan.

Traditional work on maintenance and development of the JINRLIB program library was in progress. The library was replenished with the DFM-POTM/DFM-POTM_MPI software package which implements the construction of a nucleus-nucleus optical potential of elastic scattering based on the double folding model and including the DFM-POTM sequential version and the DFM-POTM_MPI parallel version written in C++ and developed using the MPI technology (authors K.V. Lukyanov, E.V. Zemlyanaya, M.V. Bashashin, http://wwwfofo.jinr.ru/programs/jinrlib/dfm-potm/index.html). Furthermore, the library was enlarged by educational materials on parallel programming technologies and two programs, namely, PSD2SAS, a program for converting data of a position-sensitive detector of a small-angle neutron spectrometer in case of an isotropic scattering sample (authors A.G. Soloviev, S. A. Kutuzov, O.I. Ivankov, A.I. Kuklin), and IntroOMP, a selection of educational sample programs on the OpenMP parallel programming technology (authors M.V. Bashashin, T.F. Sapozhnikova and E.V. Zemlyanaya). At present, the JINRLIB library has more than 60 software packages, most of which solve the problems of automating experimental data processing and computational mathematics. The section for programs with the use of the MPI technology was created. In order to get statistics on the use of programs a page visit counter and a program text download counter were installed. The list of the program library authors was prepared.

The LIT group provides support for the components of the TDAQ system developed in LIT and included in Release 5. The resource manager (RM), one of the components included in the system core, was significantly ungraded [3] based on the RM experience in Run1. The code was simplified by using new possibilities of C++ and other libraries and reduced by 40% without the loss of functionality and quality of work. Each job starting in TDAQ requests its resources

from RM and launches only if all resources are available. While running ATLAS, many applications must be launched and stopped for a short period of time. RM is designed in such a way as to process about 30 thousand requests from O(1k) clients within several seconds so as not to affect the TDAQ start time. Figure 62 illustrates the results of a series of tests (start on 48 hosts from 1 to 30 applications on each host with (blue plot) and without (red plot) RM) proving minimal changes in the start time during the RM work.
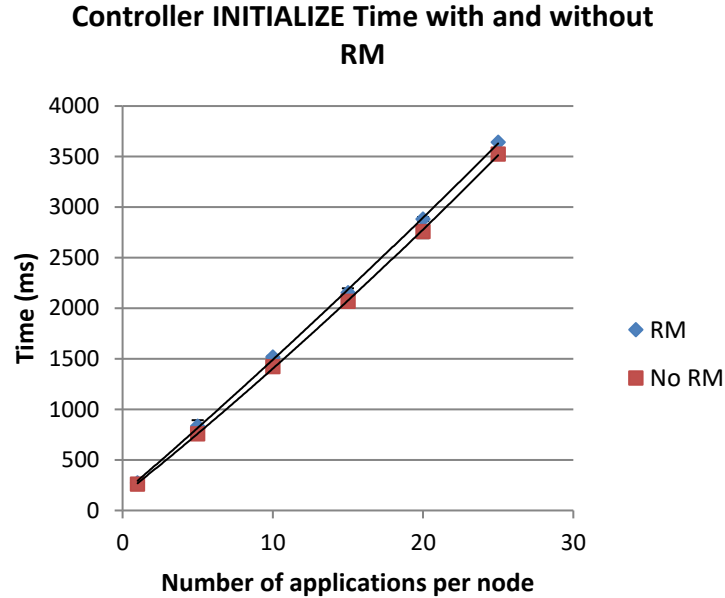
Controller INITIALIZE Time with and without RM

Fig.62. Plot of the Controller initialization time.

The group also participated in the development of a new version of *Log Manager online* of the TDAQ ATLAS system [28].

Since autumn 2017, the LIT group has taken part in the ATLAS EventIndex project. A system for monitoring the ATLAS EventIndex project based on the Grafana system was developed and put into operation, which speeded up receiving panels in ten times [4]. We participated in the development of a trigger information processing system for indexing in EventIndex for Monte Carlo data [35]. In frames of the project Conditions DB ATLAS, last autumn the development of software for the conversion of COOL data to CREST data in order to upgrade ConditionDB for RUN3 and the creation of the corresponding user libraries Aphena started. The first version of software was submitted to CERN in March 2019.

A system for monitoring the ATLAS networks was created. The work on its improvement and commissioning is in progress.

In 2017-2018 work on maintenance and development of the functionality of services rendered to users of the "Visit-center" web portal providing on-line support for the implementation of the JINR visitors' reception process (https://visitcentre.jinr.ru/) and the modernization of the JINR journals' web portal "Physics of elementary particles and atomic nucleus" (EPAN) and "Letters to EPAN" (http://pepan.jinr.ru/) was carried out. Traditionally the development, creation and maintenance of websites of conferences, symposia at the request of laboratories and other JINR subdivisions were carried out (European Cyclotron Progress Meeting ECPM-2018, http://ecpm2018.jinr.ru; International Symposium on Exotic Nuclei EXON-2018, http://exon2018.jinr.ru). Hosting with the transfer of all the information of past conferences (from 2006 to 2017) to create the site of the International Conference "Mathematical Modeling and Computational Physics" MMCP-2019 (http://mmcp.jinr.ru) was organized.

Development of the system for training and retraining IT-specialists on the basis of the JINR MICC and its educational components

Educational program on Cloud infrastructure

Tutorials on how to build local cloud or grid infrastructures, integrate them into a distributed environment and use them for own purposes are held on a regular basis or upon request from the JINR Member State organizations.

A set of training events in creating cloud platforms based on OpenNebula and using resources of distributed information and computing environments was organized and conducted for specialists from various countries (the information on conducted tutorials is shown in Table 4).

The purpose of training courses is to provide information to university students, employees of various organizations in the form of easy-to-understand and use in everyday needs, both personal and organizational.

Tutorials include the following aspects:

* providing information about main cloud services;
* providing information on the use of virtualization, advantages and disadvantages;
* hands-on tutorial with access to the JINR cloud infrastructure and the opportunity to get practical experience in VM management: creation, deployment, suspending, resuming, making VM disk snapshots, terminating, etc.).

Table 4. The information on conducted trainings.

Organization/Event	Organization location	Training dates	Number of trainees	Training type	Training location
Internship of specialists from Uzbekistan	Uzbekistan	08.06-06.07.2018	2	usage and administration	JINR, Dubna, Russia
Internship of young scientists from CIS countries	CIS countries	24.05.2018	20	usage	JINR, Dubna, Russia
The course of cloud services and virtual environments	Dubna University, Dubna, Russia	01.09.-29.12.2017	22	usage	Dubna University, Dubna, Russia
Internship of young scientists from CIS countries	CIS countries	17-18.07.2017	6	usage	JINR, Dubna, Russia
L.N. Gumilyov Eurasian National University	Astana, Kazakhstan	22.07.2017	5	usage	JINR, Dubna, Russia
Kazakhstan, Georgia, Bulgaria scientific organizations	Kazakhstan, Georgia, Bulgaria	05-09.06.2017	7	usage and administration	JINR, Dubna, Russia

Educational program on the HybriLIT platform

The HybriLIT platform is used not only for parallel computations, but also as a polygon for training students, post-graduate students and young scientists in the field of HPC. This is another important aspect of using the HybriLIT platform. Training is conducted both by young staff members of the Institute and its Member States and by students of Dubna University. The main tool in the educational process is an education and testing polygon which is the basic platform for basic courses on "Computer System Architecture", "High Performance Computing Technologies" and "Mathematical Models in Physics"; the number of students is about 200 people per academic year. The educational program can be divided into three directions.

The main purpose of **the first direction** is to acquaint students with bases of parallel programming technologies as well as IT-solutions and tools necessary for the effective use of HPC platforms. During the 2014-2018 time period, on the basis of the platform, 41 tutorials and lectures were held; more than 341 people from JINR and 218 people from the JINR Member States took part in them (Fig.63). The tutorials were carried out in frames of the conferences and schools held by LIT JINR (MPANCS'2014, MMCP'2015 and 2017, NEC'2015 and 2017, GRID'2016 and 2018), within the programs on international cooperation at Sofia University in Bulgaria and Mongolian State University. In frames of regular training courses on parallel programming technologies organized at LIT for JINR staff, students and post-graduate students of Dubna University, the HybriLIT heterogeneous computing team held tutorials on C/C++ program languages, ROOT/PROOF program packages, parallel programming technologies such as CUDA, OpenMP, OpenCL, MPI as well as on the user-friendly GitLab web interface for joint parallel development of applications.
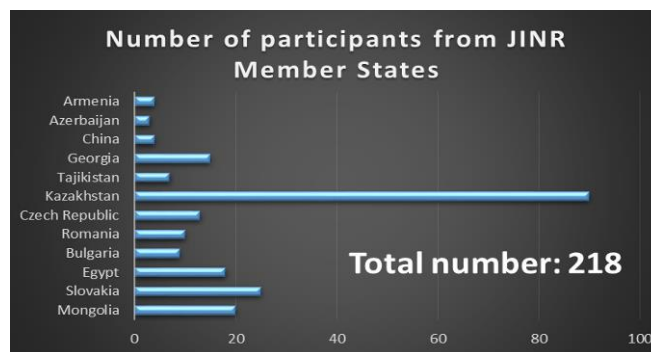

Fig.63. Number of participants from the JINR Member States.

The second direction is related to advanced learning of parallel programming technologies as part of regular training courses conducted by the HybriLIT team for JINR staff, students, post-graduate students of Dubna University. Training courses are held on C/C ++, MPI, OpenMP, CUDA, OpenCL as well as on hybrid technologies MPI + OpenMP, MPI + CUDA, etc., specialized software such as ROOT, PROOF, etc., used to solve problems in high energy physics. These courses are aimed at improving skills and retraining specialists.

For advanced learning of the latest computing technologies of parallel programming and the specifics of using the latest computing architectures, the HybriLIT group provides training courses from software developers of parallel programming. In particular, on the basis of the HybriLIT heterogeneous platform training courses on MPI parallel programming technologies and tools for debugging and profiling parallel applications (Intel company), CUDA (NVIDIA company), application programs COMSOL Multiphysics, MATLAB, etc. were held.

The third direction is a regular training of IT-specialists at the department of System Analysis and Management of Dubna University. In this direction, the courses on "Computing System Architectures", "Architectures and Computing System Technologies", "Mathematical Models in Physics" are held for second year students (bachelor's degree) and fifth year students (master's degree). During the courses, students learn about tools for working on HPC platforms and related IT-services. In addition, along with their regular studies, students of Dubna University

participate in solving specific research tasks based on the HybriLIT platform and have a possibility to choose themes for undergraduate and master theses on solving problems related to JINR research topics using HPC technologies.

This adaptive educational program based on the HybriLIT cluster allows training students and specialists taking into account the rapidly developing HPC sector in the field of scientific computing.

Conferences

On 25-29 September, the town of Budva, Montenegro, hosted the **XXVI international symposium on nuclear electronics and computing** ("Nuclear Electronics and Computing" - NEC'2017). The Symposium has been traditionally held by JINR since 1963, and for the ninth time JINR and CERN became its organizers. Co-chairmen of the Symposium were LIT Director V. V. Korenkov (from JINR) and Dr. Ian Bird (from CERN). The Symposium was attended by more than 120 leading specialists in the field of modern computer and network technologies, distributed computing and nuclear electronics from 14 countries: Belarus, Moldova, Bulgaria, Great Britain, Germany, Russia, USA, France, Czech Republic, Slovakia, Italy, China, Netherlands, Switzerland. The scientific program of the Symposium covered a wide range of issues and included the following sections: detector and nuclear electronics; triggering, data acquisition and control systems; machine learning and big data analytics; grid-technologies and cloud computing; computing for large-scale accelerator facilities (LHC, FAIR, NICA, SKA, PIC, XFEL, ELI, etc.); non-relational databases and heterogeneous repositories; research data infrastructure, computations with hybrid systems (CPU, GPU, coprocessors) as well as a traditional topic of the Symposium related to innovative IT education. Within the Symposium, a workshop "BigPanDA Technical Interchange Meeting" was held. The Symposium was organized under the sponsorship of the companies Niagara Computers, JET Infosystems, Dell-EMC and IBS Platformix.

In total, 89 reports (36 plenary and 53 sectional ones) were heard. At the educational section, 10 reports were delivered.

For the fourth time the international school on modern information technology for students and postgraduates was organized within the Symposium. This School was devoted to heterogeneous distributed computing infrastructures.

The School was attended by leading scientists from Russia, UK, USA, Italy as well as employees from JINR, NRC KI, CERN and DESY.

The school-conference was attended by 32 senior students as well as masters and postgraduates majoring in information technology from the leading Russian universities (National Research Nuclear University "MEPhI", Saint Petersburg State University, University "Dubna", Ryazan State Radio-engineering University, Magnitogorsk State Technical University, Russian University of Peoples Friendship, Tomsk Polytechnic University.

The **8th International Conference 'Distributed Computing and Grid Technologies in Science and Education'** (GRID'2018) was held on September 10-14 in the Laboratory of Information Technologies. The GRID biennial conference being traditional for the Laboratory represents a unique platform for discussing a wide range of issues related to the use and development of distributed grid technologies, heterogeneous and cloud computing in various fields of science, education, industry and business. The conference traditionally attracted a large community of Russian and foreign experts who were ready to discuss emerging issues and prospects for the development of modern information technologies. In total more than 250 scientists from scientific centers of Belarus, Bulgaria, Germany, Georgia, China, Moldova, Romania, Slovakia, the Czech Republic, France, Sweden and other countries took part in the conference. Russia was represented by participants from more than 30 universities and research centers. As part of the conference the work of eleven sections was organized where issues related to the development of distributed computing technologies, cloud technologies, heterogeneous computing, voluntary computing and big data analytics were discussed.

In total 33 plenary reports, over 120 sectional reports and 26 posters were presented at the conference. During the conference fruitful discussions were held by the participants, new IT projects aimed at developing distributed and high-performance computing and new areas of LIT cooperation with organizations, universities of Russia and JINR Member States were proposed

Presentations of the reports, an electronic version of the book of abstracts and photographic materials are posted on the conference website http://grid2018.jinr.ru. The conference proceedings are published in CEUR Workshop Proceedings (CEUR-WS.org) http://ceur-ws.org/Vol-2267/.

Within the framework of the conference an international school 'Scientific Computing, Big Data Analytics and Machine Learning Technologies for Megascience Projects' was held.

The school was aimed at attracting young scientists, PhD students and students to solve IT problems related to various aspects of megaprojects in high-energy physics as well as to get acquainted with modern methods of big data analytics, machine learning technologies on high-performance computing systems used to solve these problems. The lecture and training program of the School allowed participants to see all aspects of research related to Big Data using the example of the NICA megaproject and experiments at the Large Hadron Collider as the main sources of big data in high-energy physics. Within the school training courses on distributed systems for collecting, processing, managing and storing information, the use of high-performance and high-speed computing systems (supercomputers, computer clusters) for processing and modeling data of physical experiments including methods of machine learning, were organized. School participants had the opportunity to visit plenary reports of leading world scientists and specialists in the field of grid technologies and distributed computing presented at the conference.

The school was attended by 46 students specializing in information technologies from the leading universities of the country such as NRNU MEPhI, SPSU, Dubna University, TPU (Tomsk), TSU (Tver), VSU (Voronezh), Plekhanov REU, SUSU as well as the universities of the JINR Member States such as TSU and GTU (Georgia), TU in Kosice (Slovakia), NBU (Bulgaria), Al-Farabi KNU (Kazakhstan) and JINR associate Member Egypt, Suez University.

The conference was sponsored by the JINR Directorate and partners such as Huawei, IBSPlatformix, Niagara Computers, Supermicro, Jetinfosystems, SchneiderElectric, NVIDIA, Dell EMC, RSC Group, Intel, Cisco, ExtremeNetworks, Softline. Informational support was provided by PARALLEL.RU, the Publishing House 'Open Systems' and the International Scientific Journal 'Modern Information Technologies and IT Education'.

Publications

For the period of 2017-2019, the LIT staff members have been co-authors of 132 publications of the ATLAS collaboration, 198 publications of the CMS collaboration which are not included in the list of publications and publish 56 articles in referred journals.

[1] Adam G., Korenkov V., Podgainy D., Streltsova O., Strizh T., Zrelov P. HybriLIT – the main component of the MICC for heterogeneous computations at JINR, CEUR Workshop Proceedings (CEUR-WS.org), Vol. 2023, pp. 351-356, 2017.

[2] Agakishiev G., Balashov N., Betts W., Didenko L., Hajdu L., Korenkov V., Kuznetsov E., Lauret J., Mitsyn V., Panebratsev Y. STAR's approach to highly efficient end-to-end grid production. CEUR Workshop Proceedings (CEUR-WS.org), Vol. 2023, pp. 6-14, 2017.

[3] Aleksandrov I., Avolio G., Lehmann G., Miotto, Soloviev I. The Resource Manager the ATLAS Trigger and Data Acquisition System. Journal of Physics Conference Series, IOP Publishing LTD, England, 898(3):032016, Online Computing, 2017.

[4] Alexandrov E., Kazymov A., Prokoshin F. BigData tools for the monitoring of the ATLAS EventIndex. CEUR Workshop Proceedings (CEUR-WS.org), Vol. 2267, pp. 91-94, 2018.

[5] Astakhov N., Baginyan A., Balandin A. et al. CEUR Workshop Proceedings (CEUR-WS.org), Vol. 2023, pp. 68-74, 2017.

[6] Astakhov N., Baginyan A., Balandin A., Belov S., Dolbilov A., Golunov A., Gromova N., Kadochnikov I., Kashunin I., Korenkov V., Mitsyn V. et al. JINR Grid Tier-1@Tier-2. CEUR Workshop Proceedings (CEUR-WS.org), Vol. 2023, pp. 68-74, 2017.

[7] Badalov L.A., Belov S.D., Kadochnikov I.S. Checking foreign counterparty companies using Big Data. CEUR Workshop Proceedings (CEUR-WS.org), Vol. 2267, pp. 523-527, 2018.

[8] Baginyan A., Balandin A., Belov S., Dolbilov A., Golunov A., Gromova N., Kadochnikov I., Kashunin I., Korenkov V., Mitsyn V., Pelevanyuk I., Shmatov S., Strizh T., Trofimov V., Voytishin N., Zhiltsov V. The CMS Tier1 at JINR: five years of operations. CEUR Workshop Proceedings (CEUR-WS.org), Vol. 2267, pp. 1-10, 2018.

[9] Baginyan A., Balashov N., Baranov A., Belov S., Belyakov D., Butenko Yu., Dolbilov A., Golunov A., Kadochnikov I., Kashunin I., Korenkov V. et. al. Multi-level monitoring system for multifunctional information and computing complex at JINR. CEUR Workshop Proceedings (CEUR-WS.org), Vol. 2023, pp. 226-233, 2017.

[10] Balashov N. et al. Advanced Scheduling in IaaS Clouds. 23rd International conference on computing in high energy and nuclear physics (CHEP'2018), 9-13 July 2018, Sofia, Bulgaria [accepted for publication in EPJ Web of Conferences].

[11] Balashov N. et al. The service for parallel applications based on the JINR cloud and HybriLIT resources. 23rd International conference on computing in high energy and nuclear physics (CHEP'2018), 9-13 July 2018, Sofia, Bulgaria [accepted for publication in EPJ Web of Conferences].

[12] Balashov N., Baranov A., Belov S., Kadochnikov I., Korenkov V., Kutovskiy N., Nechaevskiy A., Pelevanyuk I. Optimization of the JINR Cloud's efficiency. The 26th Symposium on Nuclear Electronics and Computing (NEC'2017). CEUR Workshop Proceedings (CEUR-WS.org), Vol. 2023, pp. 88-91, 2017.

[13] Balashov N., Baranov A., Mazhitova Ye., Kutovskiy N., Semenov R. JINR member states Cloud Infrastructure. The 26th Symposium on Nuclear Electronics and Computing (NEC'2017). CEUR Workshop Proceedings (CEUR-WS.org), Vol. 2023, pp. 202-206, 2017.

[14] Balashov N.A., Baranov A.V., Kadochnikov I.S., Korenkov V.V., Kutovskiy N.A., Pelevanyuk I.S. Simple heuristic algorithm for dynamic VM reallocation in IaaS clouds. Modern Information Technologies and IT-Education, Vol. 14, No. 1, pp. 101-110, 2018. doi: https://doi.org/10.25559/SITITO.14.201801.101-110

[15] Balashov N.A., Bashashin M.V., Kuchumov R.I., Kutovskiy N.A., Sokolov I.A. JINR cloud service for scientific and engineering computations. Modern Information Technologies and IT-

Education, Vol. 14, No. 1, pp. 61-72, 2018, doi: https://doi.org/10.25559/SITITO.14.201801.061-072.

[16] Barberis D. et. al. The ATLAS EventIndex and its evolution based on Apache Kudu storage, CEUR Workshop Proceedings (CEUR-WS.org), Vol. 2267, pp. 18-25, 2018

[17] Barberis D. et al., The ATLAS EventIndex and its evolution towards Run 3. ATL-SOFT-SLIDE-2019-075 19th International Workshop on Advanced Computing and Analysis Techniques in Physics Research, 2019, https://cds.cern.ch/record/2666698

[18] Baranov A.V. et al. New features of the JINR cloud. The 8th International Conference "Distributed Computing and Grid-technologies in Science and Education" (GRID'2018), CEUR Workshop Proceedings (CEUR-WS.org), Vol. 2267, pp. 257-261, 2018.

[19] Belov S.I., Mazhitova Ye.M., Mazhitov M.I., Potapchuk B.G. Creation of cloud infrastructure of INP's Astana branch – private entity "NULITS" and its integration with the distributed JINR cloud infrastructure. CEUR Workshop Proceedings (CEUR-WS.org), Vol. 2267, pp. 237-242, 2018.

[20] Belov S., Filozova I., Kadochnikov I., Korenkov V., Semenov R., Smelov P., Zrelov P. Labour market monitoring system. CEUR Workshop Proceedings (CEUR-WS.org), Vol. 2267, pp. 528-532, 2018.

[21] Boytsov A., Kadochnikov I., Zuev M., Bulychev A., Zolotuhin Ya., Getmanov I. Comparison of python 3 single-GPU parallelization technologies on the example of a charged particles dynamics simulation problem. CEUR Workshop Proceedings (CEUR-WS.org), 2018.

[22] Korenkov V., Balashov N., Kutovskiy N., Dimitrov V., Kouzmov K., Hristova R., Hristov S., Clouds of JINR, University of Sofia and INRNE – current state of the project. CEUR Workshop Proceedings (CEUR-WS.org), Vol. 2267, pp. 248-251, 2018

[23] Javadzade, J. Belov, S. Data gathering and analysis for the monitoring of the Russian labour market. CEUR Workshop Proceedings (CEUR-WS.org), Vol. 2267, pp. 549-552, 2018.

[24] Golunov A.O et al. Operation center of the JINR Multifunctional information and computing complex. CEUR Workshop Proceedings (CEUR-WS.org), Vol. 1787, pp. 235-240 (in Russian).

[25] Invenio – Open Source framework for large-scale digital repositories [online resource]. URL: https://invenio-software.org/ (accessed 01.04.2019).

[26] JOIN² – Just anOther INvenio Instance [online resource]. URL: https://join2-wiki.gsi.de/cgi-bin/view, (accessed 01.04.2019).

[27] Kadochnikov I., Bird I., McCance G., Schovancova J., Girone M., Campana S., Espinal Curull X. WLCG data lake prototype for HL-LHC. CEUR Workshop Proceedings (CEUR-WS.org), Vol. 2267, pp. 393-399, 2018. URL: http://ceur-ws.org/Vol-2267/509-512-paper-97.pdf.

[28] Kazarov A., Mineev M. et al., Experience with SPLUNK for archiving and visualisation of operational data in ATLAS TDAQ system. J.Phys.Conf.Ser. 1085 (2018) 032052, 18th International Workshop on Advanced Computing and Analysis Techniques in Physics Research, pp. 032052. URL: http://cds.cern.ch/record/2290023

[29] Korenkov V., Kutovskiy N., Bondyakov A. Grid-site for CERN. Otkrytye systemy. SUBD No.3, 2018. (in Russian). URL: https://www.osp.ru/os/2018/03/13054506/

[30] Korenkov V., Kutovskiy N., Balashov N., Dimitrov V., Hristova R., Kouzmov K., Hristov S. Clouds of JINR, University of Sofia and INRNE Join Together. The 26th Symposium on Nuclear Electronics and Computing (NEC'2017), CEUR Workshop Proceedings (CEUR-WS.org), Vol. 2023, pp. 122-128, 2017.

[31] Korenkov V.V., Nikonov E.G., Popovicova M. Molecular dynamic simulation of water vapor interaction with various types of pores using hybrid computing structures, CEUR Workshop Proceedings (CEUR-WS.org), Vol. 2267, pp. 170-174, 2018.

[32] Kuchumov R., Petrunin V., Korkhov V., Balashov N., Kutovskiy N., Sokolov I. Design and implementation of a service for cloud HPC computations, Lecture Notes in Computer Science, Springer, 10963, pp. 103-112, 2018.

[33] Kutovskiy N.A., Nechaevskiy A.V., Ososkov G.A., Uzhinskiy A.V., Frontasyeva M.V. Cloud Platform for Data Management of the UNECE ICP Vegetation Environmental Monitoring Network. Geoinformatika, No.2, pp. 11-16, 2017 (in Russian).

[34] Mazhitova Ye., Balashov N., Baranov A., Kutovskiy N., Semenov R. Cloud infrastructure of INP'S Astana branch – PE "NULITS" and its integration with distributed JINR cloud infrastructure. "The XXII International Scientific Conference of Young Scientists and Specialists (AYSS-2018)", European Physical Journal (EPJ) Web of Conferences, Vol. 201, pp. 1-5, 2019. doi: https://doi.org/10.1051/epjconf/201920105003.

[35] Mineev M., Prokoshin F., Yakovlev A. Trigger information data flow for the ATLAS EventIndex. Proceedings of the VIII International Conference "Distributed Computing and Grid-technologies in Science and Education" (GRID 2018), CEUR Workshop Proceedings (CEUR-WS.org), 2018. URL: http://ceur-ws.org/Vol-2267/104-107-paper-18.pdf

[36] Nechaevskiy A., Ososkov G., Pryahina D., Trofimov V., Weidong Li. Simulation approach for improving the computing network topology and performance of the China IHEP Data Center. European Physical Journal (EPJ) Web of Conferences: http://www.epj-conferences.org/, free (accepted in print).

[37] PANDA Collaboration. Feasibility study for the measurement of pi N transition distribution amplitudes at PANDA in -pp $\rightarrow$ J/Psi pi^0. Physical Review D covering particles, fields, gravitation, and cosmology, The American Physical Society, 95, 3, 032003, 2017.

[38] Pelevanyuk I.S. et al. Service monitoring system for JINR Tier-1. CHEP2018 Proceedings (accepted for publishing).

[39] Petrosyan A. COMPASS Grid Production System. Proceeding of the 26th International Symposium on Nuclear Electronics & Computing (NEC'2017), CEUR Workshop Proceedings (CEUR-WS.org), Vol. 1111, pp. 111-112, 2018.

[40] Petrosyan A. COMPASS Production System: Processing on HPC. Proceeding of the 8th International Conference "Distributed Computing and Grid-technologies in Science and Education" (GRID 2018), CEUR Workshop Proceedings (CEUR-WS.org), Vol. 2267, pp. 139-144, 2018.

[41] Petrosyan A. COMPASS Production System Overview. Proceeding of the 23th International Conference "Computing in High Energy Physics" (CHEP 2018) (to be published).

[42] Petrosyan A., Zemlyanichkina E. PanDA for COMPASS: processing data via Grid. Proceeding of the 7th International Conference "Distributed Computing and Grid-technologies in Science and Education" (GRID 2016), CEUR Workshop Proceedings (CEUR-WS.org), Vol. 1787, pp. 385-388, 2017.

[43] Trofimov V.V., Nechaevskiy A.V., Ososkov G.A., Pryhina D.I. A probabilistic-cost approach to optimize distributed data storage systems for physical experiments. CEUR Workshop Proceedings (CEUR-WS.org), Vol. 2267, pp. 393-399, 2018, free (in Russian). URL: http://ceur-ws.org/Vol-2267/393-399-paper-75.pdf/

[44] Uzhinskiy A., Goncharov P., Ososkov G., Nechaevskiy A. CEUR Workshop Proceedings (CEUR-WS.org), Vol. 2267, pp. 200-206, 2018; Goncharov P., Ososkov G., Nechaevskiy A., Uzhinskiy A., Nestsiarenia I. Advances in Neural Computation, Machine Learning, and Cognitive Research II, Springer, pp. 151-159, 2018.

[45] Zaikina T., Egorova O., Filozova I., Kunyaev S., Musulmanbekov G., Semenov R., Shestakova G., Ustenko P. The Automation of the Content Filling for JINR DOCUMENT SERVER. Distributed Computing and Grid-technologies in Science and Education GRID'2016. CEUR Workshop Proceedings (CEUR-WS.org), 2017 (in Russian). URL: http://ceur-ws.org/Vol-1787

[46] Baginyan A.S., Dolbilov A.G., Kashunin I.A., Korenkov V.V. Traffic balancing in high-load systems using the TRILL protocol, T-Comm – Telecommunications and Transport, Media Publisher, 4, pp. 14-18, 2017. (in Russian)

[47] Baginyan A.S., Balashov N.A., Baranov A.V., Belov S.D., Belyakov D.V., Butenko Yu.A., Dolbilov A.G., Golunov A.O., Kadochnikov I.S., Kashunin I.A., Korenkov V.V., Kutovskiy N.A.,

Mayorov A.V. et al. A multi-level system for monitoring the JINR Multifunctional Information and Computing Complex, JINR News, 2, pp. 32-34, 2018. (in Russian)

[48] Balashov N.A., Baranov A.V., Kadochnikov I.S., Korenkov V.V., Kutovskiy N.A., Pelevan I.S. et al. A simple heuristic algorithm of the VM dynamic redistribution in IAAS clouds, Modern information technologies and IT-education, Foundation for assisting the development of Internet Media, IT-education, Human Potential "Internet Media League", Vol. 14, No. 1, pp. 98-108, 2018. (in Russian)

[49] Balashov N.A., Baranov A.V., Kuchumov R.I., Kutovskiy N.A., Sokolov I.A. JINR cloud service for scientific and engineering computing, Modern information technologies and IT-education, Foundation for assisting the development of Internet Media, IT-education, Human Potential "Internet Media League", Vol. 14, No. 1, pp. 57-68, 2018. (in Russian)

[50] Bondyakov A.S. Main modes of operation of an intrusion prevention system (IDS/IPS SURICATA) for the computing cluster, Modern information technologies and IT-education, Foundation for assisting the development of Internet Media, IT-education, Human Potential "Internet Media League", Vol. 13, No. 3, pp. 31-37, 2017. (in Russian)

[51] Dolbilov A.G., Zrelov P.V., Korenkov V.V., Kutovskiy N.A., Mitsyn V.V., Podgainy D.V., Streltsova O.I., Strizh T.A., Trofimov V.V. Multifunctional information and computing complex at the Laboratory of Information Technologies, JINR News, 1, pp. 22-27, 2018. (in Russian)

[52] Korenkov V.V., Kutovskiy N.A., Bondyakov A.S. Grid site for CERN. Open systems, 3, 2018. (in Russian)

[53] Kutovskiy N.A., Nechaevskiy A.V., Ososkov G.A., Trofimov V.V. Simulation of the interprocessor interaction when performing MPI applications in the cloud. Computer research and modeling, Institute of Computer Studies in cooperation with the Department of Biophysics, Biological Faculty, MSU, Vol. 9, No. 6, pp. 953-961, 2017. (in Russian)

[54] Kutovskiy N.A., Nechaevskiy A.V., Ososkov G.A., Pryahina D.I., Trofimov V.V. Modeling of MPI applications on heterogeneous computing resources, Information and telecommunication technologies and mathematical modeling of high-tech systems: Proceedings of the All-Russian conference with international participation, 2017. (in Russian)

[55] Kutovskiy N.A., Nechaevskiy A.V., Ososkov G.A., Pryahina D.I., Trofimov V.V., Simulation of the interprocessor interaction when performing MPI applications in the cloud, Computer research and modeling, Institute of Computer Research, Vol. 9, No. 6, pp. 955-963, 2017. (in Russian)

[56] Korenkov V.V., Podgainy D.V., Streltsova O.I., Educational program on HPC technologies based on the HYBRILIT heterogeneous cluster (LIT JINR), Modern information technologies and IT-education, Vol. 13, No. 4, pp. 141-146, 2017. (in Russian)