GRID AT JINR

A.S. Baginyan, A.I. Balandin, A.G. Dolbilov, A.O. Golunov, N.I. Gromova, I.S. Kadochnikov, I.A. Kashunin, V.V. Korenkov, V.V. Mitsyn, D.A. Oleynik, I.S. Pelevanyuk, A.S. Petrosyan, S.V. Shmatov, T.A. Strizh, V.V. Trofimov, A.S. Vorontsov, N.N. Voytishin, V.E. Zhiltsov

Joint Institute for Nuclear Research, 6 Joliot-Curie, 141980, Dubna

E-mail: strizh@jinr.ru

The JINR grid infrastructure is represented by the Tier1 centre for the CMS experiment at the LHC and the Tier2 centre. The grid centre resources of JINR are part of the global grid infrastructure WLCG (Worldwide LHC Computing Grid), developed for the LHC experiments. JINR LIT actively participates in the WLCG global project. The work on the use of the grid infrastructure within the WLCG project is carried out in cooperation with collaborations such as CMS, ATLAS, Alice and major international centres, which operate as Tier1 centres of the CMS experiment (CH-CERN, DE-KIT, ES-PIC, FR-CCIN2P3, IT-INFN-CNAF, US-FNAL-CMS) and as Tier2 grid centres located in more than 170 computing centres of 42 countries worldwide. Since the beginning of 2015, a full-scale WLCG Tier1 centre for the CMS experiment at the LHC has been operating in JINR LIT. The CMS Tier1 centre at JINR has demonstrated stable work through the entire period since its launch into full operation and is ranked second in its performance in the world Tier1 centres for CMS. The Tier2 centre supports a whole number of virtual organizations, particularly Alice, ATLAS, CMS, LHCb, BES, BIOMED, COMPASS, MPD, NOvA, STAR and others.

Keywords: WLCG, CMS, Tier1, Tier2

Andrey Baginyan, Anton Balandin, Andrey Dolbilov, Alexey Golunov, Natalia Gromova, Ivan Kadochnikov, Ivan Kashunin, Vladimir Korenkov, Valery Mitsyn, Danila Oleynik, Igor Pelevanyuk, Artem Petrosyan, Sergei Shmatov, Tatiana Strizh, Alexey Vorontsov, Vladimir Trofomov, Nikolay Voytishin, Victor Zhiltsov

1. Introduction

The current state of computing in high-energy and nuclear physics over the past 15 years is mainly determined by the distributed computing infrastructure created for processing and storing data of the LHC experiments. Each of four experiments, i.e. ATLAS, CMS, Alice and LHCb, represents not only a unique and sophisticated scientific tool, but also a collaboration, each comprising several thousands of scientists from several hundreds of institutes located all around the world.

Figure 1. Worldwide LHC computing Grid (WLCG) global infrastructure. The JINR Tier1 centre is marked with a white circle

Within the framework of the WLCG infrastructure (Fig.1), about 1,000,000 processors geographically distributed in more than 170 data centres from 42 countries are combined into a single computing environment [1]. The resources of the JINR grid centres are part of this global WLCG infrastructure. JINR LIT plays an active role in the WLCG project. Developments related to the use of the grid infrastructure are carried out in cooperation with the CMS, ATLAS and Alice collaborations and major international centres, which operate as Tier1 centres for the CMS experiment (CH-CERN, DE-KIT, ES-PIC, FR-CCIN2P3, IT-INFN-CNAF, US-FNAL-CMS), as well as Tier2 level grid centres. Since 2000, JINR has been actively involved in the creation of the grid segment in Russia and its integration into the global grid infrastructure of the LCG (LHC Computing Grid) and EGEE projects. During these years, the largest Tier2 centre in Russia was created at JINR for processing data from all four experiments at the LHC. Since the beginning of 2015, a full-scale Tier1 WLCG centre has been operating at JINR LIT for the CMS experiment at the LHC [2].

2. Network infrastructure

A high-performance network infrastructure is crucial for the development of WLCG centres. Last year the projects related to increasing the bandwidth of the Moscow-JINR telecommunication channel to 3x100 Gb/s, installing and configuring the equipment of the Institute backbone computing network to 2x100 Gb/s and the distributed computing cluster network between JINR facilities to 400 Gb/s were carried out at JINR. Thus, a new direct channel to CERN was put in operation. Now the JINR external distributed network is represented by a direct JINR-CERN communication channel and
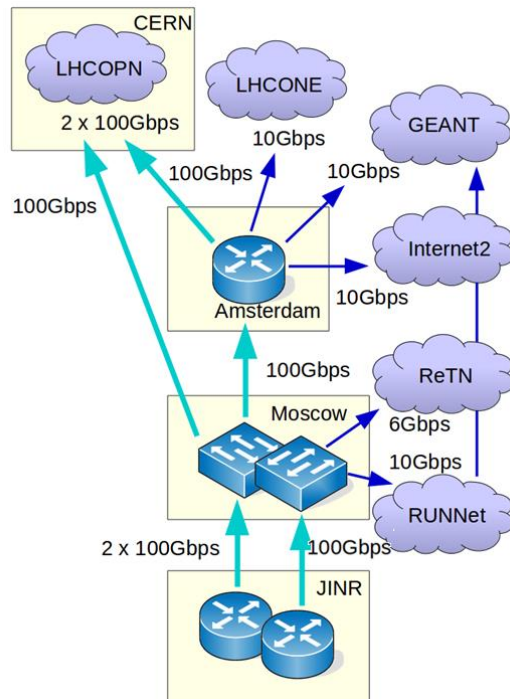
Figure 2. Logical structure of the JINR external network

a backup channel passing through MMTS-9 in Moscow and Amsterdam, ensuring the operation of LHCOPN (JINR–CERN) for the communication of the Tier0 (CERN) and Tier1 (JINR) centres and an external LHCONE overlay network designed for the JINR Tier2 centre, direct communication channels using the RU-VRF technology with a collaboration of RUHEP research centres, RUNNet and ReTN networks (Fig.2).

3. JINR grid resources

The JINR grid infrastructure is represented by the Tier1 centre for the CMS experiment at the LHC and the Tier2 centre, 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 JINR CMS Tier1 consists of 400 64-bit machines: 2 x CPU, 6–16 cores/CPU that forms altogether 10,688 cores for batch processing and provides a performance of 151.97 kHS06. The storage system contains disk arrays and long-term data storage on tapes and is supported by the dCache-5.2 and Enstore 4.2.2 software. The total usable capacity of disk servers is 10.4 PB; the IBM TS3500 tape robot is 11 PB. The evolution of the JINR CMS Tier1 main resource parameters for the last three years is presented in Figure 3.

	2017	2018	2019
DISK(TB)	5590	7216	9316
MSS(TB)	9839	11153	11153
MSD(TB)	1103	1103	802
core	4160	4720	10688
kHS06	63,7	72,31	152

Figure 3. Evolution of the JINR CMS Tier1 main resource parameters

The Torque 4.2.10/Maui 3.3.2 software (custom-built) is used as a resource manager and a task scheduler. The PhEDEx software is used as a tool for management of the 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 (Torque) and an information service (monitoring of services, servers, storage, data transmission, information sites).

The data processing system at JINR Tier2 consists of 239 64-bit machines: 2 x CPU, 6–16 cores/CPU that forms altogether 4,128 cores for batch processing and provides a performance of 55.489 kHS06. The storage system contains disk arrays and is supported by dCache-5.2 and EOS. One of the dCache instances is used by CMS and ATLAS. The second dCache is used by JINR users and user groups. Besides, this instance is used to store data of several third-party experiments (BIOMED, BESIII, FUSION). One XROOTD installation is used by Alice. The total usable capacity of disk servers is 2,789 TB for ATLAS, CMS and Alice, and 140 TB for other VOs. WLCG services are the same as in Tier1, and the software stack is mostly the same too.

4. JINR Tier1 for CMS

The CMS Tier1 centre at JINR has demonstrated stable work through the entire period since its launch into full operation [3,4]. Since the beginning of 2019, 284,762 pilot jobs have been completed, which accounts for 18% of the Sum CPU Work (HS06 hours) of CMS experiment data processed by all Tier1 centres. Figure 4 shows the contribution of the world first level centres to processing of CMS experiment data in 2019. The JINR centre is ranked second in its performance in the world.

Safekeeping of a proportional share of raw and reconstructed data, large-scale reprocessing and safekeeping of the corresponding output, the distribution of data to Tier2s and safekeeping of a share of simulated data produced at these Tier2s are among the most important tasks of the Tier1

Figure 4. World Tier1 for CMS — Sum CPU Work (HS06 hours) by Tier1 Node and Year 2019

centre. In 2019, 157 WLCG centres downloaded 3.5 PB of data from our tape robot and 140 centres uploaded 0.969 PB to it. The largest data exchange occurred with FNAL (USA), CERN, KIT (Germany), INFO-CHEF (Italy), CCIN2P3 (France) and RAL (UK). The Tier1 online disk storage was also heavily used. Within the year, 16.8 PB of data was uploaded, and 36.4 PB was downloaded.
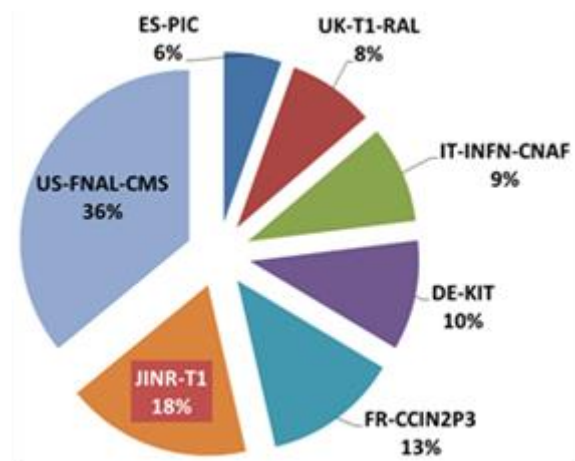
5. JINR Tier2

Tier2 at JINR currently provides the majority of JINR users and user groups, as well as users of virtual organizations (VOs) of the grid environment, with computing power, storage systems and access to data. To support a batch processing system, a special server with a cluster resource allocation system and a job scheduler, which allows processing data for all four experiments at the LHC (ALICE, ATLAS, CMS, LHCb) and supports some virtual organizations working within the grid (BESIII, BIOMED, COMPASS, FUSION, MPD, NOvA, STAR and others), was installed. The JINR Tier2

centre is the best one in the Russian Data Intensive Grid (RDIG) federation and provides 52% of the total CPU work of RDIG. More than 3 million jobs were processed in 2019.

6. Integration of distributed resources

In recent years, a number of works have been carried out at JINR LIT to apply grid technologies for processing of non-LHC experiments data.

A data processing management system for the COMPASS experiment on SPS has been developed and is in operation. It provides automatic data processing and management on different types of computing resources [5]. Currently, the computing resources available to the collaboration are the CERN Batch Service, the JINR computing centre and the Frontera supercomputer of the Texas supercomputer centre.

The DIRAC Interware was used to enable the integration of distributed heterogeneous computing and storage resources for NICA MPD experiment simulation. More than 50,000 of Monte-Carlo simulation jobs of the NICA MPD experiment were performed at JINR Tier1 and Tier2 using DIRAC.

7. Monitoring

To ensure a correct and reliable operation of the computer complex equipment and provide relevant timely information about its work, a special monitoring system has been designed [6]. This system allows one to keep in check climate control and power supply systems, the local network equipment, telecommunication links and computing nodes, running jobs, disk and tape storage systems. An operation centre has been launched to allow real-time monitoring of the whole JINR grid infrastructure. The number of nodes included in monitoring is more than 1,200. To ensure such extensive monitoring, a cluster monitoring system based on the Icinga2 software is used. Visualization is done with Grafana and NagVis.

8. Future plans

In terms of hardware, a linear increase in the characteristics of Tier1 and Tier2, i.e. an increase in computing resources of CMS Tier1 up to 350 kHS06 and of Tier2 up to 170 kHS06, an expansion of the storage system of CMS Tier1 on disks up to 16 PB and of the storage system on tapes up to 42 PB, is planned for the next years.

References

[1]   WLCG (The Worldwide LHC Computing Grid): http://wlcg.web.cern.ch/LCG

[2]   Baginyan A.S., Balandin A.I., Belov S.D. et al. The CMS TIER1 at JINR: five years of operations, CEUR-WS.org/Vol-2267/1-10-paper-1.pdf, 2018.

[3]   Astakhov N.S., Baginyan A.S., Belov S.D. et al. JINR Tier-1 centre for the CMS Experiment at LHC. Particles and Nuclei, Letters, v.13,no 5, pp.1103-1107, 2016.

[4]   Astakhov N.S., Baginyan A.S., Balandin A.I. et al., CEUR-WS.org/Vol-2023/68-74-paper-10.pdf, 2017.

[5]   Petrosyan   A.Sh.,   COMPASS   Production   System   Overview, https://doi.org/10.1051/epjconf/201921403039, EPJ Web of Conf., Vol. 214, 2019.

[6]   Baginyan A.S., Balashov N.A., Baranov A.V. et al. Multi-level monitoring system for multifunctional information and computing complex at JINR, CEUR-WS.org/Vol-2023/226-233-paper-36.pdf, 2017.