

IHEP TIER-2 COMPUTING CENTER: STATUS AND OPERATION

V. Kotliar^a, V. Gusev, V. Kukhtenkov

Institute for High Energy Physics named by A.A. Logunov of National Research Center "Kurchatov Institute", Nauki Square 1, Protvino, Moscow region, Russia, 142281

E-mail: ^a viktor.kotliar@ihep.ru

RU-Protvino-IHEP site is one of three biggest WLCG Tier-2 centers in Russia. The computing infrastructure serves for "big four" high energy physics experiments such as Atlas, Alice, CMS, LHCb and local experiments at IHEP such as OKA, BEC, radio biology stands and others. In this work the current status of the computing capacities, networking and engineering infrastructure will be shown as well as the contribution of the grid site to the collaboration experiments.

Keywords: WLCG, GRID, Tier-2

Viktor Kotliar, Victor Gusev, Vladimir Kukhtenkov

Copyright © 2021 for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

1. Introduction

RU-Protvino-IHEP site participates in the Worldwide LHC Computing Grid [1] for almost two decades. Started from a few Pentium III computers with several services for grid infrastructure IHEP participated in the EGEE I – EGEE III projects and became one of the core site in the Russian Data Intensive Grid (RDIG). After increasing network bandwidth to 2x10Gb/s and connecting to LHCONE network grid site in Protvino became one of the biggest Tier-2 site in Russia after JINR with more than 150 computing nodes (~3k CPU or ~26k HEP-SPEC06) and 2.5PB disks space. Now site serves for big four LHC experiments (Atlas, Alice, CMS, LHCb) and many physical experiments inside the Institute. Used schema for CPU usage with sharing them among Grid and not Grid experiments allowed to achieve effective resources usage over the years in 24x7 mode.

2. Grid site overview

IHEP cluster is a cluster with shared CPU resources among Grid and Internal experiments [2]. In the core of the cluster is the Kerberos5 “Single Sign on” system. It is used by local users to get an access to the resources of the cluster directly without any grid services. It means that a batch system on the cluster supports Kerberos tickets and these tickets are forwarded to all Working Nodes too. As soon as we also use Andrew File System for user’s home directories the cluster must provide support for AFS tokens which are based on Kerberos tickets. The next major service is a Lustre parallel cluster file system which is shared across all working nodes. It allows to local users to perform data analysis as faster as possible. Both for the long data store and for the store of the RAW data we use CASTOR2 [3] (the CERN Advanced STORage manager). It is a HSM system that has been using in IHEP for fifteen years. We store data on LTO4 and LTO5 tapes with tape libraries and stand alone tape drives. In the core of the cluster a multicomponent CMS [4] (cluster management system) is used [fig.1].

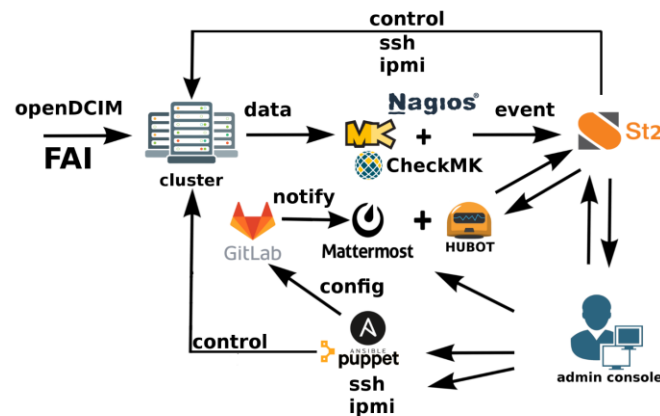


Figure 1. IHEP cluster management system

All grid services including ARC CE, BDII, APEL, VO BOX and all services of CMS works under Xen hypervisors on the high availability clustered system as virtual machines. This technique allows using resources as much as needed for the current setup and dynamically adjusting them if it will be necessary in the future ensuring availability. On the site several physical servers provide services for the site infrastructure:

- Network gateway servers which provide NAT (network address translation or masquerade) for the site internal network;
- DNS (Domain Name Servers) for the site internal network and for caching DNS queries on the cluster;

- Squid http proxy servers for caching CVMFS [5] (CERN Virtual Machine File System) requests and special frontier proxies for LHC experiments to cache Oracle DB (data base) queries to the central repositories.

Grid middleware is UMD4 [6] and base OS (operating system) on the cluster is CentOS 7 64bit. For GPU's we use CUDA 7.5. There are Ansys 19.0 and Mathematica 8.0 installed on several nodes on the cluster and Intel® Fortran Composer XE 64bit is installed on UI (User Interface). All this software allows to use more flexibly computing farm at IHEP as a multi-purpose computing cluster.

3. Available resources

Grid center at IHEP has storage and computing resources that it declares for usage to the WLCG infrastructure. At the moment there are:

- 3044 CPU or 26875 HEP-SPEC06;
- 2183TB disks space by experiments: ATLAS 1185TB, CMS 395TB, ALICE 314TB, LHCb 289TB;
- 2x10Gb/s Internet channels (LHCONE shared with Russian Data Intensive Grid);
- Manpower is three system administrators.

All four LHC experiments have their own disk resources and CPU resources shared among all of them on a fair share basis. The main storage technologies which are used on site are dCache and xrootd. They allow to build distributed storage systems for grid computing. Access to outside world is done by two 10Gb/s links where IHEP is connected to the LHCONE computing network for scientific research. Manpower for the site support is three system administrators with an administration rota on weekly basis. Such schema makes administrators available for research tasks, programming, learning new techniques. The site counts as third big Grid site in Russia [fig. 2].

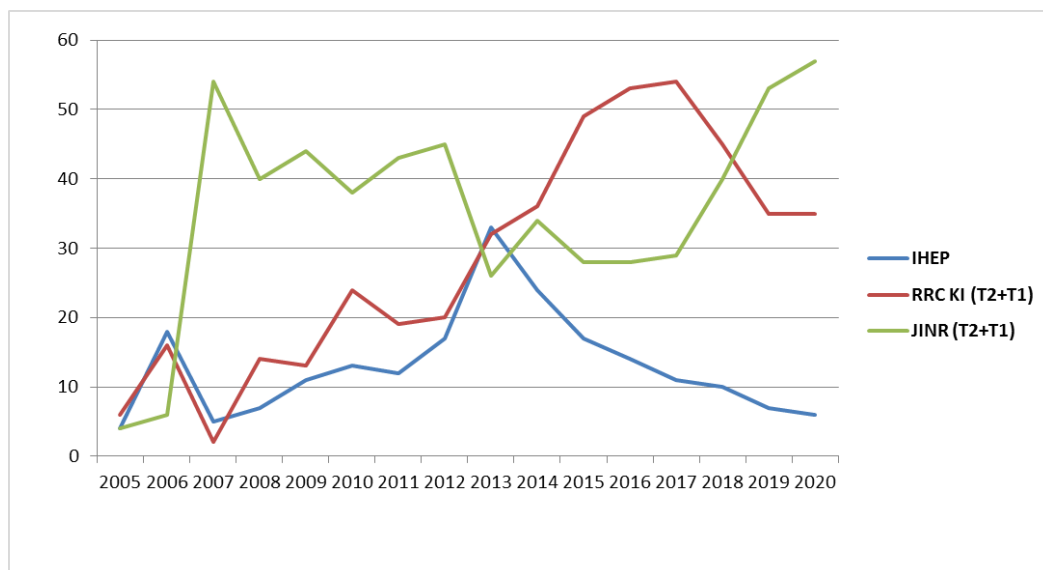


Figure 2. A history of the contribution in CPU time for big Russian Grid sites

4. Site usage by virtual organizations

As soon as the computing cluster at IHEP is shared among four Grid experiments, following fair share schema implemented for CPU resources: ATLAS 52%, CMS 30%, ALICE 12%, LHCb 6%. It needs to be taken into account that the CPUs also shared with the local IHEP experiments like BEC, OKA, etc. So the real usage for the last year represents as ATLAS 40%, CMS 17%, ALICE 13%, LHCb 30%. The main consumer for the site is the ATLAS experiment where RU-Protvino-IHEP site has a status of Tier-2D alpha which means a Tier-2 site with availability and accessibility more than

95%. The site contributes with 20% of all ATLAS jobs in RU-cloud and 10% of all RU-cloud data stored at IHEP. For CMS the percent of jobs computed on the site is around 6% inside Russian segment. For ALICE it is 7% and for LHCb 10%. For LHCb, IHEP has also a status of the site Tier-2D (with data) and it is only Tier-2 site with such status in Russia. IHEP site shows very good stability and has CPU usage efficiency near 94% [fig. 3] which is above average CPU efficiency for Grid sites in WLCG.

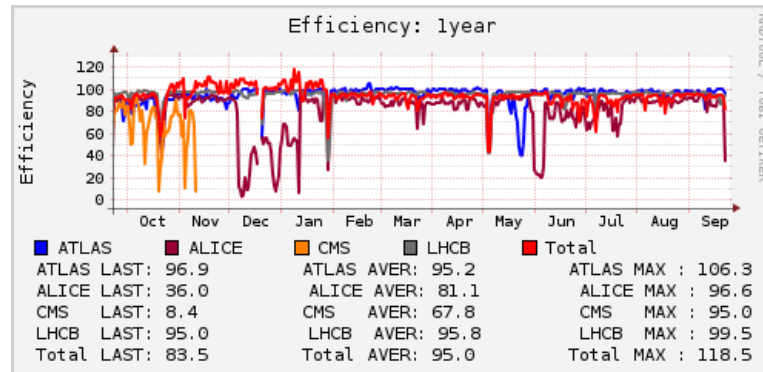


Figure 3. Running jobs and efficiency per year

To achieve such parameters and to be able to run without downtimes some works performed:

- Power connections to IHEP data center were modified enabling auto switching on input power sources in case of power cuts;
- Constant system and software upgrades seamlessly made with a help of the cluster management system;
- Main works focused on supporting current infrastructure availability, reliability and efficiency.

5. Conclusion

The WLCG Grid site at IHEP shows a high stability with a big pledge of resources to the LHC experiments. Performed works and gained experience in the grid computing makes the site ready for LHC run 3, which planned for next year.

IHEP uses leading technologies in the computing science for computing clusters with batch systems to run jobs. There are many research works and tests inside the site infrastructure ongoing. Self-healing systems, containers and high availability systems and technologies are going to increase stability, availability and interoperability of the computing infrastructure at IHEP making it ready to further increasing resources and system complexity to achieve new goals in the HEP computing.

References

- [1] The Worldwide LHC Computing Grid project [WLCG]. Available at: <https://wlcg.web.cern.ch> (accessed 22.09.2021)
- [2] Kotliar V. IHEP cluster for Grid and distributed computing // CEUR Workshop Proceedings. February 2017: Vol. 1787.- pp. 312-316
- [3] The CERN Advanced STORage manager [CASTOR]. Available at: <https://castor.web.cern.ch/castor/> (accessed 22.09.2021)
- [4] Ezhova, V., Kotliar, A., Kotliar, V., Popova, E. Multicomponent cluster management system for the computing center at IHEP // CEUR Workshop Proceedings. 2018: Vol. 2267. - pp. 37-43.
- [5] The CernVM File System [CVMFS]. Available at: <https://cernvm.cern.ch/fs/> (accessed 22.09.2021)
- [6] The Unified Middleware Distribution [UMD]. Available at: <https://wiki.egi.eu/wiki/Middleware> (accessed 22.09.2021)