Service for parallel applications based on JINR cloud and HybriLIT resources

Nikita Balashov [1], *Maxim* Bashashin [1], *Pavel* Goncharov[1], *Ruslan* Kuchumov [2], *Nikolay* Kutovskiy [1], *Kiril* Kulikov [1], *Mikhail* Matveev [1], *Andrey* Nechaevskiy[1], *Gennady* Ososkov[1], *Ilkhom* Rahmonov[1], *Yuri* Shukrinov[1], *Ivan* Sokolov[1,*], *Oksana* Streltsova[1] and *Elena* Zemlyanaya[1]

[1] Joint Institute for Nuclear Research, Dubna, Russia
[2] Saint Petersburg State University, Saint Petersburg, Russia

Abstract. Cloud computing has become a routine tool for scientists in many fields. The JINR cloud infrastructure provides JINR users with computational resources to perform various scientific calculations. In order to speed up achievements of scientific results the JINR cloud service for parallel applications has been developed. It consists of several components and implements a flexible and modular architecture which allows to utilize both more applications and various types of resources as computational backends. An example of using the Cloud&HybriLIT resources in scientific computing is the study of superconducting processes in the stacked long Josephson junctions (LJJ). The LJJ systems have undergone intensive research because of the perspective of practical applications in nano-electronics and quantum computing. In this contribution we generalize the experience in application of the Cloud&HybriLIT resources for high performance computing of physical characteristics in the LJJ system.

1 Introduction

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 approach are the following:

• Necessity to take care of 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 analyse results, moreover, they do not always need all the resources of modern multicore CPUs servers).

---

* Corresponding author: isokolov@jinr.ru

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.

2 Implementation

A motivation behind the JINR cloud infrastructure [1] and the HybriLIT heterogeneous cluster [2] deployment was to provide users with access to more powerful computational resources as well as 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. Users' 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 an amount of required resources and specify a location for job results. An implementation of such 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 1 and described below.
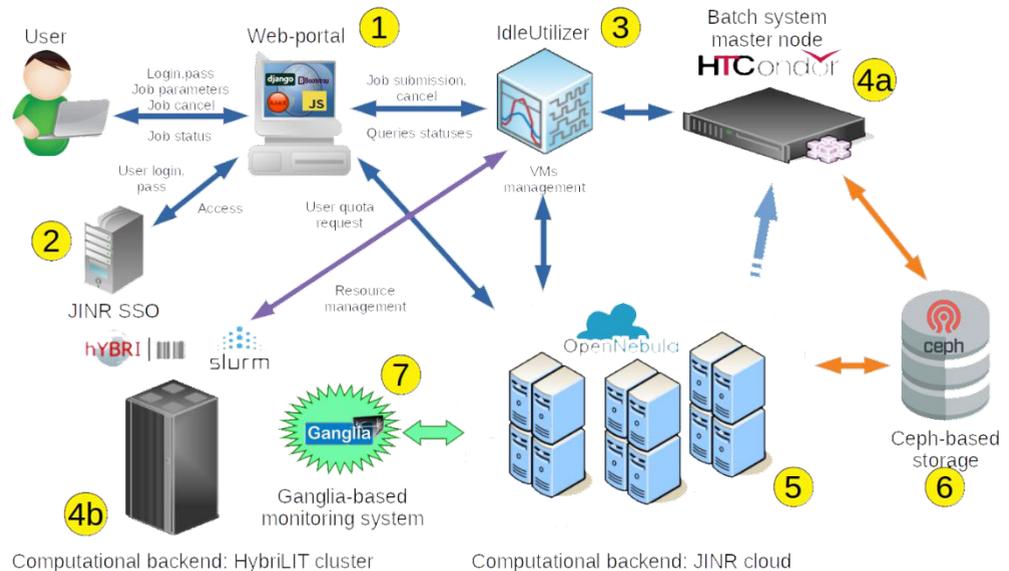

Fig. 1. A basic workflow schema and main components of the JINR SaaS service.

2.1 Web portal

A user interacts with the service only via a web portal (marked by the digit "1" in a yellow circle in Fig. 1) which is a client-server web application. Its server part is based on the Django framework [3] which implements the so-called "Model-Template-View" architecture. The web application communicates with both the IdleUtilizer (IU) component and computational

backends: either the JINR cloud via remote procedure calls (RPC) encoded in a XML format (i.e. via the so-called "XML-RPC") or the HybriLIT cluster via a command line interface of its batch system.

A client part of the service with which a user interacts via a browser has been developed using the software stack based on Bootstrap 3, AJAX, JavaScript, HTML5 and CSS. The given part allows to select a particular application and a computational backend, define input parameters for the application and an amount of requested resources (it also validates them) as well as submit a job. Apart from that, it displays information about free 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 a mechanism to cancel a job after it was submitted.

2.2 JINR SSO

A user authenticates on the web portal via the JINR Single Sign-on (SSO) service [4] (marked by the digit "2" in a yellow circle in Fig. 1). Such approach mitigates risk for access to third-party sites (user passwords are not stored or managed externally), reduces 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 a 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. A screenshot of the JINR SSO login screen is shown in Figure 2.

Fig. 2. A screenshot of the JINR SSO login screen.

2.3 IdleUtilizer

The IdleUtilizer (marked by the digit "3" in a yellow circle in Fig. 1) is a Python script which either deploys on the JINR cloud virtual machines becoming worker nodes (WNs) of the HTCondor-based virtual cluster [5] 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 [6]. 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 have been 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 [7] as well as HTCondor- and Slurm-based batch systems.

2.4 HTCondor batch system

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 a yellow circle in Fig. 1) is deployed in the JINR cloud as well.

2.5 OpenNebula-based cloud infrastructure

The JINR SaaS service implements a modular architecture in order to enable job submission to various computational backends. It is possible to use OpenNebula-based clouds where VMs are created for running a real workload. The JINR cloud (marked by the digit "5" in a yellow circle in Fig. 1) is used as one of the computational backends for the service for scientific and engineering computations. Cloud infrastructure uses a distributed consensus protocol to provide fault-tolerance and state consistency across its services.

Apart from the local resources, the JINR cloud has some amount of external ones from the partner organizations of the JINR Member States (see [8] for more details).

2.6 HybriLIT Resources

The HybriLIT heterogeneous cluster (marked by the digit "4b" in a yellow circle in Fig. 1) 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 [2] apart from multicore CPUs.

Resources of the HybriLIT cluster operated by the Slurm workload manager [9] can be used as another computational backend for user jobs coming from the SaaS web portal.

2.7 Ceph-based software-defined storage

Another key component of the JINR cloud infrastructure and the JINR SaaS service is a software-defined storage (SDS) based on Ceph [10] (marked by the digit "6" in a yellow circle in Fig. 1). It delivers object, block and file storage in a unified system.

2.6 Ganglia-based monitoring system

A ganglia-based monitoring system [11] (marked by the digit "7" in a yellow circle in Fig. 1) is used 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. 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 a 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.

3 Workflow

The web portal of the JINR SaaS is available at the URL https://saas.jinr.ru. A user logs in on the web portal via a web browser (see Figure 3). To start working a user should be authenticated by the JINR SSO service mentioned above.

Fig. 3. A screenshot of the JINR SaaS login page.

At present, there are interfaces in the web portal for the following applications:
• "Hello test" which is used for testing and debugging purposes;
• "Long Josephson junctions simulation" which allows to simulate Long Josephson junctions using the in-house application [12, 13] developed by the collaboration of colleagues from the Bogoliubov Laboratory of Theoretical Physics and the Laboratory of Information Technologies (LIT).

Initially after successful authentication the first tab "Creating a job" (Figure 4) is active where a user needs to do several actions:

Fig. 4. A screenshot of the "Creating a job" tab on the web portal.

1. choose a certain application to run;
2. choose a computational backend and define an amount of requested resources;
3. define a set of values for the parameters of the selected applications if there are any;
4. specify a URL where output data including job stderr and stdout files should be uploaded by the service;
5. press the "Submit" button to run a job.

After the last step a user is redirected to the tab "Jobs results" (Figure 5) where the progress of the job execution can be tracked or job itself can be cancelled by pressing the "Cancel" button.
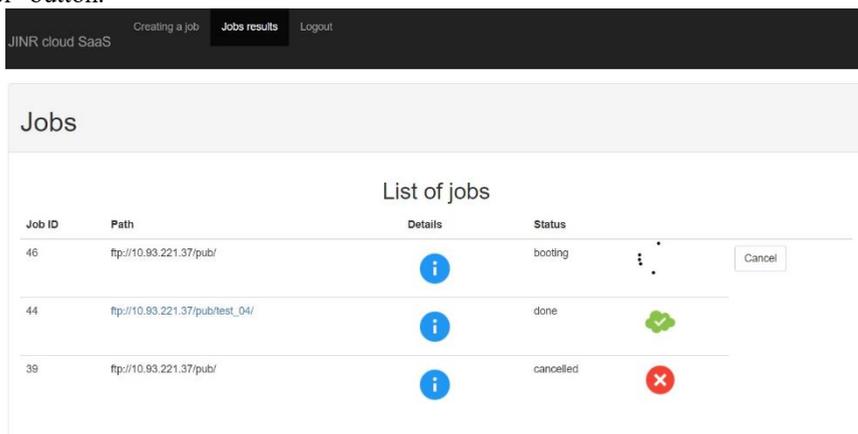

Fig. 5. A screenshot of the "Jobs results" tab with job statuses.

When the job reaches the status "Done" the output files become accessible at the specified URL.

The service has an advanced mode on the tab "Creating a job" (Figure 6). It provides an ability to change the IdleUtilizer IP address as well as edit the HTCondor job description. This mode can be activated by enabling the corresponding checkbox. It is useful for debugging purposes and normally should not be used by regular users.
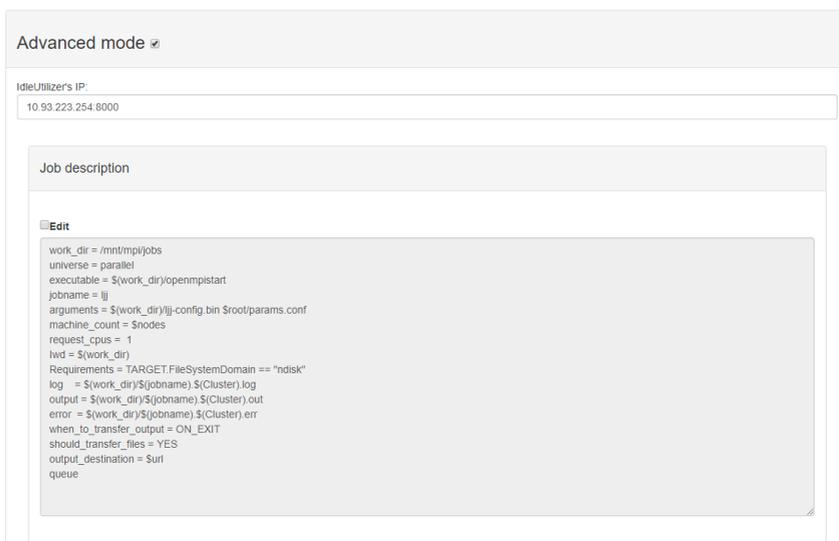

Fig. 6. A screenshot of the advanced mode on the tab "Creating a job".

Upon job completion or its cancellation a user gets an e-mail with the corresponding information.

4 Future plans

In the nearest future it is planned to develop the JINR SaaS web portal into the following directions:

• add visualization of the results directly in the service web interface;
• improve the web interface by adding a possibility to choose a certain type of LJJ, set a job description, re-submit a job, automatically calculate some of the parameters based on the values already entered by a user and known interrelations between them.

Apart from the LJJ application, more applications developed by research groups from JINR and its Member State organizations are planned to be added into the given service.

5 Conclusion

The JINR cloud team has developed the service which provides access for scientists from small research groups from JINR and its Member State organizations to computational resources via an application-specific web interface. It allows a scientist to focus on his research field by interacting with the service in a convenient way via a browser and abstracting away from underlying infrastructure as well as its maintenance. A user just sets required values for his job via the web interface and specifies a location for uploading the result. The computational workload is done on the resources of either the JINR cloud infrastructure or the HybriLIT heterogeneous cluster. Some amount of the resources is provided for researchers for free. However, if more resources are required there is an option to buy capacities at one's own expense and integrate them into the JINR LIT data centre where proper engineering infrastructure and environment for hardware already exist. In this way an overall hardware utilization efficiency is increased due to the possibility to use them by other users when such capacities are idle. The service has considerably facilitated the solution of the very important problem of superconducting processes in the stacked long Josephson junctions.

References

1. A.V. Baranov, N.A. Balashov, N.A. Kutovskiy, R.N. Semenov *JINR cloud infrastructure evolution*, PPNL **13**, No. 5, 672 (2016)

2. E.I. Alexandrov et al., *Research of Acceleration Calculations in Solving Scientific Problems on the Heterogeneous Cluster HybriLIT*, RUDN JMISP **4**, 30 (2015)

3. D. Greenfeld, A. Greenfeld *Two Scoops of Django: Best Practices for Django 1.8 (3rd ed.)*, Publisher: Two Scoops Press, 531

4. D. Drinkwater *What is single sign-on? How SSO improves security and the user experience*, CSO from IDG (2018)

5. D. Thain, T. Tannenbaum, M. Livny *Distributed Computing in Practice: The Condor Experience*, Concurrency and Computation: Practice and Experience **17**, No. 2-4, 323 (2005)

6. P. Adamson et al. (NOvA Collab.) *First measurement of electron neutrino appearance in NOvA*, Phys. Rev. Lett. **116**, 151806 (2016)

7. R. Moreno-Vozmediano, R. S. Montero, I. M. Llorente *IaaS Cloud Architecture: From Virtualized Datacenters to Federated Cloud Infrastructures*, IEEE Computer, **45**, 65, (2012)

8. N.A. Balashov et al. *JINR Member States cloud infrastructure* // CEUR Workshop Proceedings **2023**, 202 (2017)

9. J. Morris, M. Grodona *SLURM: Simple Linux Unity for Resource Management*, CWC and Expo (2003)

10. Weil, S.A., Brandt, S.A., Miller, E.L., Long, D.D.E., Maltzahn, C. *Ceph: A Scalable, High-Performance Distributed File System*, USENIX Association, 307 (2006)

11. M. Massie, B. Chun, D. Culler *The Ganglia Distributed Monitoring System: Design, Implementation, and Experience*, Paral. Comp. **30**, 817 (2004)

12. I. R. Rahmonov, Yu. M. Shukrinov, P. Kh. Atanasova, E. V. Zemlyanaya, and M. V. Bashashin, JETP 124, 131 (2017)

13. M.V. Bashashin, E.V. Zemlyanaya, I.R. Rahmonov, Yu.M. Shukrinov, P.Kh. Atanasova, A.V. Volokhova *Numerical approach and parallel implementation for computer simulation of stacked long Josephson Junctions*, CRM T.8 **4**, 593 (2016)