

Н. Балашов, А. Баранов, Р. Гайнанов, Н. Кутовский, Р. Семенов

Развитие облачной инфраструктуры ОИЯИ

В течение 2016 г. облачная инфраструктура ОИЯИ [1] развивалась по следующим основным направлениям:

- 1) увеличение количества ресурсов, доступных ее пользователям;
- 2) увеличение количества решаемых на ней задач;
- 3) реализация сбора и визуализации статистики использования облачных ресурсов на базе программного обеспечения (ПО) Influx DB и Grafana.

Общее количество ресурсов, доступных пользователям облака ОИЯИ, было увеличено как за счет добавления дополнительных серверов в качестве рабочих узлов, так и за счет интеграции части облачных вычислительных ресурсов организаций-партнеров.

Для объединения облачных ресурсов организаций-партнеров из стран-участниц ОИЯИ с целью решения совместных задач, а также для распределения пиковых нагрузок по интегрированным ресурсам командой сотрудников ЛИТ был разработан специальный драйвер [2]. Он позволяет объединять ресурсы

партнерских организаций, развернутые как в облаке на базе OpenNebula (и в этом случае возможно мониторить в реальном времени загрузку внешнего облака), так и на других облачных платформах, которые поддерживают «open cloud computing interface» (ОСЦИ-интерфейс), но в этом случае доступные для использования внешними пользователями ресурсы задаются статическими значениями. Драйвер построен на ruby-реализации ОСЦИ-интерфейса — rОСЦИ и использовании «extensible mark up language remote procedure call» (XML-RPC) функций самой платформы OpenNebula. Причиной одновременного применения двух указанных интерфейсов является то, что ОСЦИ позволяет увеличить повторное использование кода для различных популярных облачных платформ, которые поддерживают данный интерфейс (например, Open Stack), тогда как функции XML-RPC платформы OpenNebula являются специфическими конкретно для данной платформы и не могут быть использованы для интеграции с другими. В то же время ОСЦИ не предо-

N. Balashov, A. Baranov, R. Gainanov, N. Kutovskiy, R. Semenov

JINR Cloud Infrastructure Evolution

During this year the JINR cloud infrastructure [1] has been developing in several directions:

- 1) increasing the amount of resources its users have access to;
- 2) increasing the number of tasks it is used for;
- 3) re-implementing an aggregation and visualization statistics on cloud resources utilization using InfluxDB and Grafana software.

The amount of resources available for JINR cloud users grew up in 2016 due to 1) maintaining additional servers as cloud nodes and 2) integration of part of the computing resources of partner organizations' clouds.

In order to join the cloud resources of partner organizations from JINR Member States to solve common tasks as well as to distribute a peak load across them, a cloud bursting driver [2] has been designed by the JINR cloud team. It allows one to integrate the JINR cloud with

the partner clouds, either OpenNebula-based one (and in this case it is possible to enable real time external cloud resources monitoring) or any other cloud platform which supports Open Cloud Computing Interface (OCCI), but in this case, statically defined values of external cloud resources are only possible. The driver is based on the ruby implementation of OCCI (rOCCI) and OpenNebula eXtensible Markup Language Remote Procedure Call (XML-RPC) interface. The reason why both interfaces are used at the same time is that OCCI allows one to reuse the code better, since it is provided by other popular cloud platforms like OpenStack, while OpenNebula XML-RPC is OpenNebula-specific and cannot be used for integration with other platforms. At the same time, OCCI does not provide monitoring operations, while OpenNebula XML-RPC does. It was decided to use rOCCI for management operations and XML-RPC for monitoring.

ставляет функции мониторинга, в отличие от XML-RPC платформы OpenNebula.

С облаком ОИЯИ интегрированы облака следующих партнерских организаций из стран-участниц Института: Института физики Национальной академии наук Азербайджана (Баку); Института теоретической физики им. Н.Н.Боголюбова Национальной академии наук Украины (Киев); Российского экономического университета им. Г.В.Плеханова (Москва).

Кроме того, облако ОИЯИ интегрировано в федеративное облако европейской грид-инфраструктуры (ЕГИ), тем самым делая возможным взаимное использование облачных ресурсов нашего Института и ЕГИ.

На момент написания статьи общее количество ядер центрального процессорного устройства в облаке ОИЯИ, доступное его пользователям, равно 330

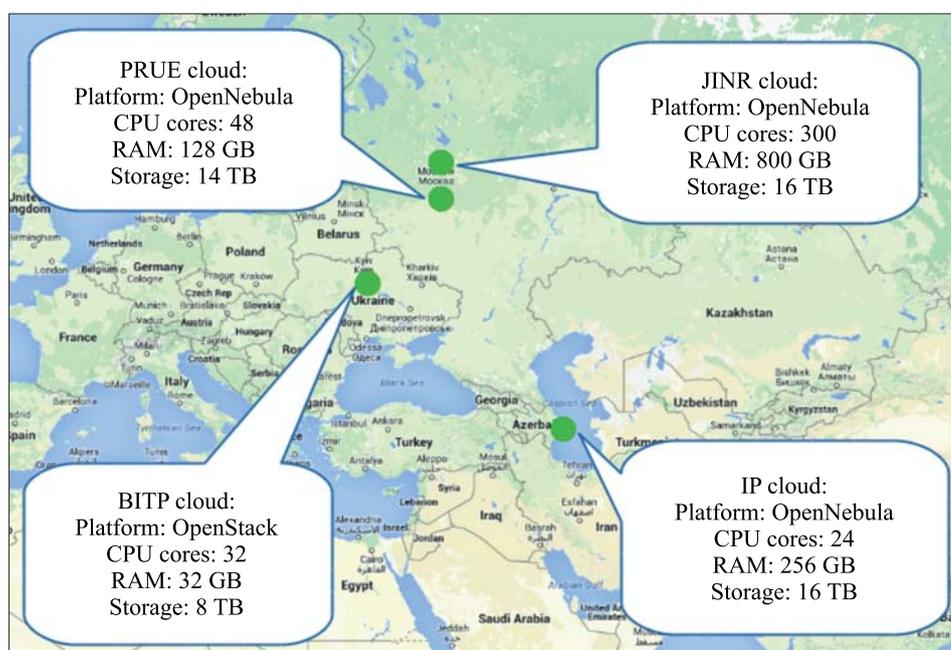
(в 2015 г. оно составляло 200) и общий объем ОЗУ равен 840 ГБ (в 2015 г. было 400).

Спектр решаемых задач на облачных ресурсах ОИЯИ также увеличился в 2016 г. Для этого на данной инфраструктуре были установлены следующие компоненты:

- полигон PanDA (для развития ПО PanDA и его использования для решения задач экспериментов ATLAS и COMPASS);
- полигон на базе промежуточного программного обеспечения (ППО) DIRAC (используется для разработки средств мониторинга распределенной вычислительной инфраструктуры эксперимента BES-III, а также как один из ее вычислительных ресурсов);
- набор контейнеров для пользователей — участников эксперимента NOvA;

Географическое расположение облачных инфраструктур организаций из стран-участниц ОИЯИ, предоставляющих часть своих ресурсов для интеграции с облаком ОИЯИ: IP — Институт физики НАН Азербайджана; BITP — Институт теоретической физики им. Н.Н.Боголюбова НАН Украины; PRUE — Российский экономический университет им. Г.В.Плеханова

Geographical locations of the partner cloud infrastructures from JINR Member States which provide part of their computational resources being integrated into the JINR cloud



The clouds of the following partner organizations from JINR Member States are integrated with the JINR cloud: the Institute of Physics (IP) of Azerbaijan National Academy of Sciences (Baku); the Bogolyubov Institute for Theoretical Physics (BITP) of the National Academy of Sciences of Ukraine (Kiev); the Plekhanov Russian University of Economics (PRUE) (Moscow).

Apart from that, the JINR cloud is integrated into EGI Federative Cloud thus enabling a possibility to use a part of JINR computing resources by EGI FedCloud Virtual Organizations.

At the moment of writing the article, the total number of the CPU cores available for users in the JINR cloud is 330 (200 in 2015) and the total amount of RAM is 840 GB (400 GB in 2015).

A spectrum of the tasks solved with the JINR cloud is also increasing and the following additional components are running now on its resources:

- PanDA testbed was deployed for PanDA software validation and extensions development for ATLAS and COMPASS experiments;
- DIRAC-based testbed (it is used for monitoring tools development for BESIII experiment distributed computing infrastructure as well as its computing facility);
- a set of VMs of NOvA experiment users for analysis and software development;
- NICA testbed for grid middleware evaluation for NICA computing model development;
- EOS testbed for research on heterogeneous cyberinfrastructures, computing federation prototype creation

- полигон для изучения и оценки ППО для построения вычислительной инфраструктуры экспериментов на ускорителе NICA;

- полигон на базе системы хранения данных EOS для выполнения обязательств в рамках участия в проекте по исследованию гетерогенных киберинфраструктур, разработки и создания прототипа компьютерной федерации на основе высокоскоростных вычислений, облачных вычислений и суперкомпьютеров для хранения, обработки и анализа больших данных;

- сервер Spark для машинного обучения и анализа больших данных.

Как было упомянуто выше, в качестве третьего направления развития облачной инфраструктуры ОИЯИ в 2016 г. выступает реализация сбора и визуализации статистики использования облачных ресурсов. Ранее данный функционал был реализован сотрудниками ЛИТ как дополнительный пункт меню графического веб-интерфейса платформы OpenNebula, который называется «Sunstone». Недостатком такой реализации была необходимость проверки совместимости данного модуля сбора статистики и ее визуализации с новыми версиями платформы OpenNebula и ее веб-интерфейса, так как последний мог измениться, и в этом случае возникала необходимость адаптации кода

модуля под новый веб-интерфейс. Чтобы избежать подобных дополнительных усилий перед обновлением ПО облачной платформы, а также с целью сохранения собираемых метрик о потребляемых ресурсах в базе данных для их последующей визуализации и анализа на предмет их изменений в течение задаваемого промежутка времени было решено реализовать сбор и визуализацию статистических данных об использовании ресурсов облака ОИЯИ с помощью таких инструментов, как Influx DB и Grafana.

Информация о доступе к ресурсам облака ОИЯИ и работе в нем доступна на веб-портале с адресом <http://miccom.jinr.ru>, где необходимо выбрать «Облачный сервис».

Список литературы

1. Баранов А.В., Балашов Н.А., Кутовский Н.А., Семенов Р.Н. // Письма в ЭЧАЯ. 2016. Т. 13, № 5. С. 672–675.
2. Баранов А.В. и др. // Компьютерные исследования и моделирование. 2016. Т. 8, № 3. С. 583–590.

and development based on high-performance computing, cloud computing and supercomputing for Big Data storage, processing and analysis;

- a standalone Spark instance for Machine Learning and Big Data analysis.

As mentioned above, the third direction of the JINR cloud development was a re-implementation of aggregation and visualization statistics on the resources utilization. Initially it was performed by the JINR cloud team as an additional item in a menu of OpenNebula graphical web interface, which is called Sunstone. However, a drawback of such implementation was a necessity to check a compatibility of that aggregation and visualization add-on against each new release of OpenNebula because its web interface might be changed. There was a need to adopt the code of visualization module for the new Sunstone in case of its incompatibility. So, in order to avoid such an extra step before software update on the JINR cloud as well as to store a collecting metrics in a database for further analysis against their changes over time and to obtain the dynamics for the selected period, it was decided to implement aggregation and visualization statistics on the JINR

cloud resources utilization using such tools as InfluxDB and Grafana software.

Information on how to access the JINR cloud and to use its resources is available on the web portal at <http://miccom.jinr.ru> where the user must select “Cloud service”.

References

1. Baranov A. V. et al. JINR Cloud Infrastructure Evolution // Phys. Part. Nucl. Lett. 2016. V. 13, No. 5. P. 672–675.
2. Baranov A. V. et al. Approaches to Cloud Infrastructures Integration // Computer Research and Modeling. 2016. V. 8, No. 3. P. 583–590 (in Russian).