

*Н. Балашов, А. Баранов, Н. Кутовский, Е. Мажитова,
И. Пелеванюк, Р. Семенов*

Интеграция вычислительных ресурсов организаций стран-участниц ОИЯИ с использованием облачных и грид-технологий

Объединенный институт ядерных исследований (ОИЯИ) участвует в большом количестве научно-исследовательских проектов, в которых важным средством для получения значимых научных результатов выступают вычислительные ресурсы. В связи с этим объединение счетных мощностей организаций стран-участниц Института в единую информационно-вычислительную среду (ИВС) представляется важной и актуальной задачей, решение которой позволило бы существенно ускорить проведение научных исследований.

Переход на современные принципы построения вычислительных центров с использованием облачных

технологий, повышающих эффективность использования аппаратного обеспечения за счет виртуального разделения ресурсов и упрощающей обслуживание инфраструктуры, что снижает стоимость владения и использования аппаратных ресурсов, является одним из стратегических направлений в работе Лаборатории информационных технологий.

Для объединения облачных вычислительных ресурсов ОИЯИ и организаций из его стран-участниц был установлен реализованный на языке ruby программный интерфейс OCCI (Open Computing Cloud Interface) — rOCCI, а также разработан драйвер для интеграции по пиринговой модели «cloud bursting» [1]. Эти раз-

*N. Balashov, A. Baranov, N. Kutovskiy, E. Mazhitova,
I. Pelevanyuk, R. Semenov*

Integration of Computing Resources of the JINR Member State Organizations by Means of Cloud and Grid Technologies

The Joint Institute for Nuclear Research (JINR) participates in a large number of research projects in many of which computer infrastructures play an important role in obtaining meaningful scientific results. That is why the integration of computational resources of the JINR Member State organizations into a unified informational and computational environment (ICE) is an important and topical task the solution of which would significantly accelerate scientific research.

The transition to modern principles of building computing centers using cloud-based technologies that increase the efficiency of hardware utilization due to the virtual separation of resources and simplify the infrastructure, thus reducing the cost of ownership and usage of the

hardware resources, is one of the strategic directions of the Laboratory of Information Technologies.

In order to integrate the cloud computing resources of JINR and its Member State organizations, implementation of the programming interface OCCI (Open Cloud Computing Interface) — rOCCI on the ruby language has been installed, and a driver has been developed for integration in peer model “cloud bursting” [1]. These developments have been successfully used for the last few years as a tool to integrate the cloud resources of organizations of the JINR Member States. In this way resources were integrated of 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,

работки успешно использовались в течение нескольких лет как инструмент для объединения облачных ресурсов организаций стран-участниц ОИЯИ. Таким способом были интегрированы ресурсы Российского экономического университета им. Г.В.Плеханова (Москва, Россия), Института теоретической физики им. Н.Н.Боголюбова (Киев, Украина) и Института физики Национальной академии наук Азербайджана (Баку, Азербайджан). Схема интеграции ресурсов по пиринговой модели представлена на рис. 1, слева.

Возросший интерес со стороны стран-участниц Института привел к увеличению количества участников такой распределенной облачной инфраструктуры, что выявило низкую масштабируемость данного подхода. В связи с этим была начата исследовательская работа по оценке возможных альтернатив для интеграции облачных инфраструктур. В качестве нового средства для интеграции облаков было выбрано решение на основе грид-платформы DIRAC [2] (рис. 1, справа). Службы самой DIRAC-платформы развернуты в облаке ОИЯИ, которое так же, как и облака организаций

стран-участниц ОИЯИ, предоставляет вычислительные ресурсы для грид-среды.

DIRAC (сокращение от английского «Distributed Infrastructure with Remote Agent Control» — распределенная инфраструктура с возможностью управления удаленным агентом) — это программная среда для распределенных вычислений, предоставляющая законченное решение для одного и более сообществ пользователей, нуждающихся в доступе к распределенным ресурсам. DIRAC выполняет функцию промежуточного слоя между пользователями и аппаратными мощностями, предоставляя общий интерфейс к поставщикам разнородных ресурсов.

Система управления рабочей нагрузкой с помощью агентов в виде пилотных задач, впервые представленная в DIRAC, в настоящее время широко используется в различных грид-инфраструктурах. Эта концепция позволяет прозрачным для пользователей образом объединить в общей системе такие вычислительные ресурсы различной природы и происхождения, как грид-среды, облака или кластеры.

В отличие от пиринговой модели «cloud bursting» платформа DIRAC позволяет использовать разные вычислительные ресурсы с помощью пилотных задач.

Для того, чтобы проверить возможность подобного использования облачных ресурсов, была проведена

Рис. 1. Возможные схемы интеграции облаков

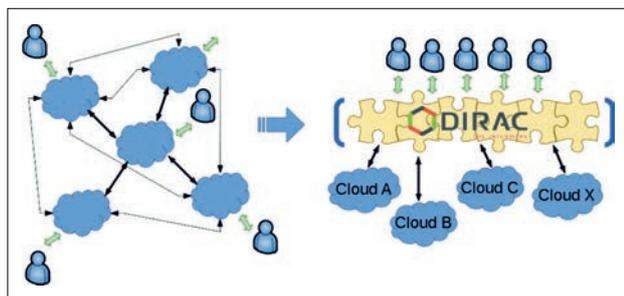


Fig. 1. Possible schemes of the clouds integration

Azerbaijan). A scheme of the clouds integration following the “cloud bursting” model is shown in Fig. 1, left.

Enormous interest of the JINR Member States in the uniform distributed computational environment has led to an increased number of participants in such a multicloud infrastructure and revealed a low scalability of this approach. That is why research work was started to evaluate possible alternatives, so the DIRAC software framework [2] has been chosen as a new tool for the distributed resources integration (Fig. 1, right). 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 (Distributed Infrastructure with Remote Agent Control) 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 the users and

the resources offering a common interface to a number of providers of heterogeneous resources.

The Workload Management System with Pilot Jobs introduced by the DIRAC project is now widely used in various grid infrastructures. This concept allows one to aggregate in a single system the computing resources of different source and nature, such as computational grids, clouds or clusters, transparently for the end users.

Unlike the “cloud bursting” model, the platform DIRAC allows one to use different computing resources with the help of pilot tasks.

In order to test the possibility of such use of cloud resources, a series of tests were conducted, which resulted in new virtual machines that were successfully created, their contextualization was performed, user tasks were started, and the results were sent to the storage element. No special actions were required to run tasks on different cloud resources. At the moment, the program code of DIRAC is being finalized, because not all the necessary functionality is present in the basic version of the system.

At present, the integration of the JINR Member State organizations clouds into DIRAC-based distributed platform is at different stages. Their technical inte-

серия тестов, в результате которых успешно создавались новые виртуальные машины, производилась их контекстуализация, запускались пользовательские задачи и результаты отправлялись на элемент хранения. При этом не требовалось никаких специальных действий для запуска задач на разных облачных ресурсах. Сейчас ведется доработка программного кода DIRAC, так как не весь необходимый функционал присутствует в базовой версии системы.

В настоящее время интеграция облаков партнерских организаций стран-участниц ОИЯИ в распределенную платформу на основе DIRAC находится на разных этапах. Выполнена техническая интеграция

и идет работа по тестированию и отладке облаков Астанинского филиала Института ядерной физики — частного учреждения «Nazarbayev University Library and IT Services» (Астана, Казахстан) [3], Научно-исследовательского института ядерных проблем Белорусского государственного университета (Минск, Белоруссия), Ереванского физического института (Ереван, Армения), Института физики Национальной академии наук Азербайджана (Баку, Азербайджан), Российского экономического университета им. Г.В.Плеханова (Москва, Россия). Находятся в процессе технической интеграции облака Института ядерных исследований и ядерной энергетики БАН (София, Болгария),

Рис. 2. Географическое положение облачных инфраструктур организаций из стран-участниц ОИЯИ, участвующих в интеграции облачных ресурсов



Fig. 2. Geographical location of the partner cloud infrastructures from the JINR Member States involved in the integration of their cloud resources

gration has been fulfilled, work on testing and debugging is in progress for the Astana branch of the Institute of Nuclear Physics — the private establishment “Nazarbayev University Library and IT Services” (Astana, Kazakhstan) [3], the Scientific Research Institute for Nuclear Problems of 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). In the process of integration are 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). Negotiations are in progress with the University “St. Kliment Ohridski” (Bitola, Macedonia). Figure 2 maps the location of the organizations listed above.

To work in any cloud infrastructure, the user must be identified in it. Each organization independently determines the identification procedure in its cloud service. JINR provides cloud access only for users who have a General Institute Account, i.e., if they are registered in the JINR User Base. Having become an AFS user at JINR, this user will automatically be able to log into the cloud service as a single Kerberos service is used for authentication.

The cloud infrastructures of JINR and its Member State organizations are grounded on a solution with open source code OpenNebula [4].

There are two interfaces available for users of the JINR cloud to access the OpenNebula service, namely, a full-featured command-line interface (CLI) and a Sunstone graphic web-interface. Access to the virtual machines is provided either by {rsa, dsa} key or by password. User authentication in Sunstone is implemented using Kerberos. To improve the security of information transfer between

Софийского университета им. Св. Климента Охридского (София, Болгария), Грузинского технического университета (Тбилиси, Грузия), Института ядерной физики АН РУз (Ташкент, Узбекистан). Ведутся переговоры по интеграции облаков с Университетом им. Св. Климента Охридского (Битола, Македония) (рис. 2).

Для работы в облачной инфраструктуре пользователь должен быть идентифицирован в ней. Каждая организация самостоятельно определяет процедуру идентификации в своем облачном сервисе. В ОИЯИ вход в облако возможен только для пользователей, которые имеют общеинститутскую учетную запись, т. е. зарегистрированы в базе пользователей ОИЯИ. В этом случае пользователь автоматически будет иметь возможность входа в облачный сервис, так как для аутентификации используется единый сервис Kerberos.

Облачные инфраструктуры Института и организаций из стран-участниц основаны на решении с открытым исходным кодом OpenNebula [4].

Для пользователей облака ОИЯИ существует два интерфейса доступа к сервису OpenNebula, а именно полнофункциональный интерфейс командной строки (CLI) и графический веб-интерфейс «Sunstone». Доступ к виртуальным машинам осуществляется либо по {rsa,dsa}-ключу, либо по паролю. Аутентификация

пользователя в Sunstone реализована с использованием Kerberos. Для повышения безопасности передачи информации между веб-интерфейсом сервиса и браузерами пользователей применено SSL-шифрование.

В ближайшем будущем планируется предоставление ученым из организаций стран-участниц Института доступа к вычислительным ресурсам суперкомпьютера ОИЯИ «Говорун». Интеграция аппаратных ресурсов суперкомпьютера и облачных инфраструктур стран-участниц позволит существенно увеличить общую вычислительную емкость всей ИВС.

Увеличивающееся количество участников распределенной вычислительной среды демонстрирует важность и актуальность данной деятельности, что, в свою очередь, связано со стремлением сторон объединить усилия и ресурсы в решении фундаментальных и прикладных задач, успешная реализация которых невозможна без применения новых подходов в проведении распределенных и параллельных расчетов и использования систем распределенного хранения больших объемов данных.

Информация о доступе к ресурсам облака ОИЯИ и работе в нем доступна на веб-портале с адресом <http://misc.jinr.ru> по ссылке «Облачный сервис».

Список литературы / References

1. Baranov A. V. *et al.* Approaches to Cloud Infrastructures Integration // Computer Research and Modeling. 2016. V.8, No.3. P.583–590.
2. DIRAC web-portal. <http://diracgrid.org>.
3. Mazhitova Ye., Balashov N., Baranov A., Kutovskiy N., Semenov R. Integrated Cloud Infrastructure of the LIT JINR, PE “NULITS” and INP’s Astana Branch // Eur. Phys. J. Web Conf. 2018. V.177. 1–5.
4. OpenNebula web-portal. <https://opennebula.org>.

the web-based interface of the service and user browsers, SSL-encryption is used.

One of the most important tasks for the nearest future is to provide scientists from the JINR Member States with an access to the computational capacities of the JINR supercomputer “Govorun”. The integration of its hardware and the cloud resources of the participating countries would significantly increase a total computing capacity of the entire ICE.

The growing number of participants of the distributed computing environment demonstrates importance and relevance of this activity, which in turn is related to the desire of the parties to join forces and resources in solving fundamental and applied problems, the successful implementation of which is impossible without the use of new approaches in conducting distributed and parallel calculations and using distributed storage for large amounts of data.

Information on how to access the JINR cloud and use its resources is available on the web-portal <http://misc.jinr.ru>, option “Cloud service”.