

**В.В. Кореньков¹, Н.А. Кутовский¹, В.В. Мицын¹, А.А. Мошкин², И.С. Пелеванюк¹,
Д.В. Подгайный¹, О.В. Рогачевский², В.В. Трофимов¹, А.Ю. Царегородцев³**

¹ Лаборатория информационных технологий им. М.Г. Мещерякова ОИЯИ

² Лаборатория физики высоких энергий им. В.И. Векслера и А.М. Балдина ОИЯИ

³ Марсельский центр физики элементарных частиц (СРРМ), Университет Экс-Марсель (Aix-Marseille University), Национальный центр научных исследований, Французский национальный институт ядерной физики и физики элементарных частиц, Марсель, Франция

На конкурс работ ОИЯИ по разделу научно-методических и научно-технических исследований выдвигается цикл работ **“Разработка и внедрение единого доступа к гетерогенным распределенным ресурсам ОИЯИ и стран-участниц на платформе DIRAC”**.

Интеграция вычислительных ресурсов на базе платформы DIRAC – это важный шаг в развитии распределённой обработки данных в ОИЯИ. Такая интеграция позволяет включить практически любой вид вычислительных ресурсов в единую систему и предоставить их пользователям через единый веб-интерфейс, интерфейс командной строки или программный интерфейс. Актуальность представленных исследований и реализованных подходов связана в первую очередь с реализацией экспериментов мегасайнс проекта NICA: VM@N, MPD, SPD. По оценкам, приведённым в документе «TDR MPD: Data Acquisition System» от 2018 года, поток данных с детектора MPD составит, как минимум 6,5 ГиБ/с. Для эксперимента SPD предварительные оценки по количеству получаемых данных приближаются к 20 ГиБ/с. Для обработки, передачи, хранения и анализа таких больших объёмов данных потребуется значительное количество вычислительных ресурсов и ресурсов хранения.

ОИЯИ обладает большим количеством разных вычислительных ресурсов: кластеры Tier1/Tier2, суперкомпьютер «Говорун», облако, NICA кластер. Ресурсы каждого из них можно использовать для достижения целей компьютинга экспериментов на коллайдере NICA. Основная трудность в данном случае состоит в том, что эти ресурсы являются разными с точки зрения архитектуры, процедур доступа и авторизации, способов использования. Для того, чтобы обеспечить их эффективное использование, необходимо, с одной стороны, интегрировать ресурсы в единую систему, а с другой стороны, не помешать их текущей эксплуатации и выполнению других задач.

Исследования выполнялись в соответствии с научно-тематическим планом научно-исследовательских работ ОИЯИ совместно Лабораторией информационных технологий, Лабораторией физики высоких энергий и Центром физики частиц университета Марселя (г. Марсель, Франция).

Для интеграции перечисленных выше разнородных ресурсов было решено использовать платформу DIRAC Interware [1]. DIRAC (Distributed Infrastructure with Remote Agent Control) выполняет функцию промежуточного слоя между пользователями и различными вычислительными ресурсами, обеспечивая эффективное,

прозрачное и надёжное использование, предоставляя общий интерфейс к поставщикам разнородных ресурсов. Изначально платформа DIRAC разрабатывалась силами коллаборации LHCb для организации компьютеринга. В 2008 г. она стала развиваться как продукт с открытым исходным кодом, предназначенный для организации распределенных вычислений на основе гетерогенных вычислительных ресурсов.

Платформа DIRAC была развернута в ОИЯИ в 2016 г. в экспериментальном режиме [2]. Для оценки эффективности ее работы использовались типовые задачи моделирования для экспериментов VM@N и MPD, а также не привязанные к экспериментам тестовые задачи.

В 2018 г. проводились работы по интеграции облачных инфраструктур ОИЯИ и стран-участниц Института в распределенную платформу на основе DIRAC [3]. Для этого потребовалось разработать специальный модуль, который позволял бы DIRAC инициировать создание виртуальных машин в системе OpenNebula, на базе которой построено вычислительное облако ОИЯИ и облака стран-участниц. Модуль был разработан специалистами Лаборатории информационных технологий и добавлен в исходный код DIRAC [4]. На данный момент разработанный модуль активно используется не только в ОИЯИ, но также включён в инфраструктуры экспериментов BES-III и JUNO. Интеграция облаков партнерских организаций стран-участниц ОИЯИ в распределённую платформу на основе DIRAC (рис. 1) открывает странам-участницам новые возможности для участия в вычислениях для экспериментов мегасайнс проекта NICA [3].

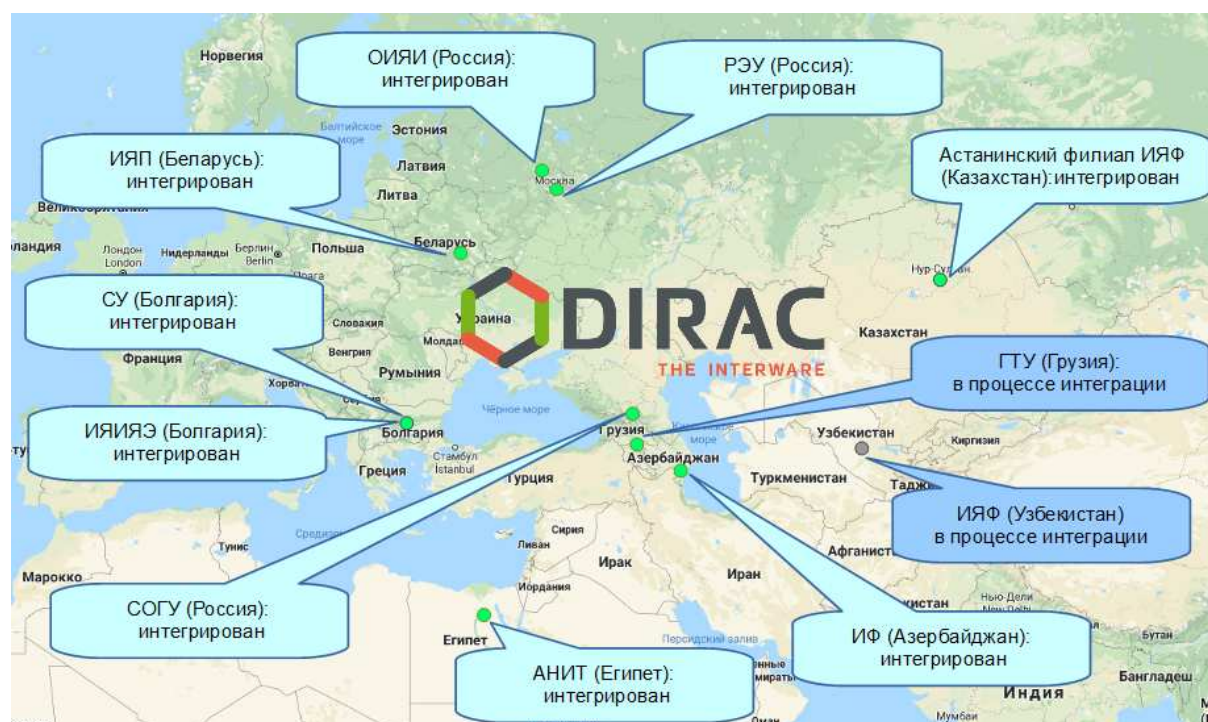


Рис. 1. Облака организаций стран-участниц ОИЯИ, интегрированные в распределённую информационно-вычислительную среду на основе платформы DIRAC

Одновременно с интеграцией вычислительных облаков прорабатывалось решение об интеграции разнородных вычислительных ресурсов на основе DIRAC для

действующего в настоящее время эксперимента VM@N, а также для будущего эксперимента MPD на базе строящегося коллайдера NICA. Были изучены пропускные способности систем хранения на дисковых и ленточных накопителях, проведено стресс-тестирование всех основных ресурсов, разработаны подходы к решению стандартных задач моделирования, обработки и передачи данных [5,6].

В августе 2019 г. первый пакет задач моделирования данных для эксперимента MPD была отправлена через DIRAC на ресурсы грид-кластеров Tier1 и Tier2. Далее в платформу распределенных вычислений (ПРВ) был интегрирован суперкомпьютер «Говорун». Летом 2020 г. были добавлены кластер NICA и кластер Национального автономного университета Мексики (UNAM). В качестве систем хранения были интегрированы системы dCache, под управлением которой находятся дисковые и ленточное хранилища, и EOS. Стоит отметить, что кластер UNAM стал первым вычислительным ресурсом, находящимся за границей Европы или Азии и включённым в инфраструктуру DIRAC в ОИЯИ. Схема интеграции географически распределённых гетерогенных ресурсов на основе DIRAC Interware представлена на рис.2.

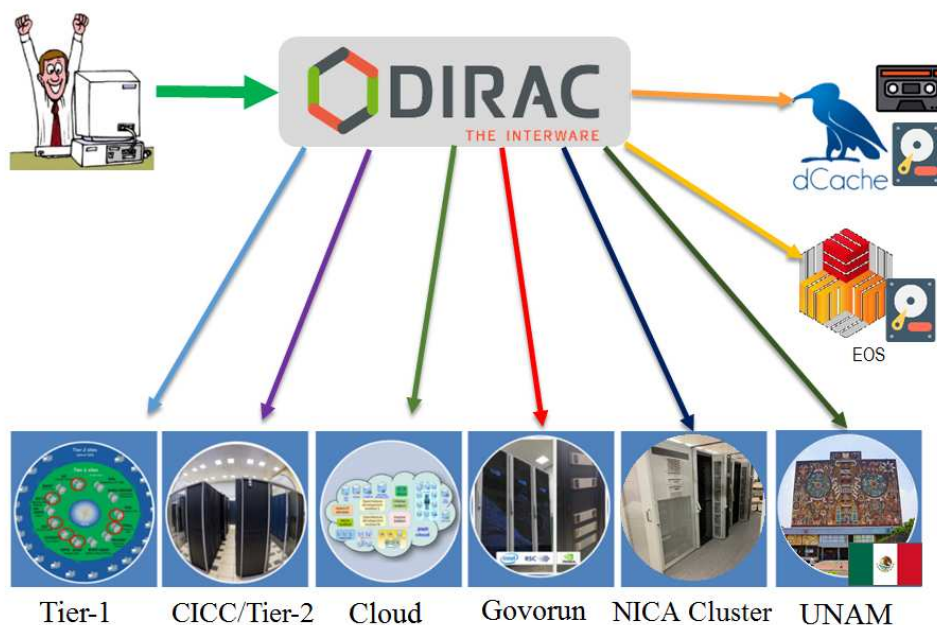


Рис. 2. Схема интеграции географически распределённых гетерогенных ресурсов на основе DIRAC

Созданная платформа распределенных вычислений позволяет распределить нагрузку по всем задействованным ресурсам, тем самым ускорив проведение расчётов и увеличив эффективность использования интегрированных ресурсов [6]. В качестве примера можно привести суперкомпьютер «Говорун». В зависимости от его загрузки различного типа задачами стало возможным часть временно свободных ресурсов направлять на решение задач, связанных с экспериментами мегасайнс проекта NICA [7].

MPD успешно использует распределённую платформу на базе DIRAC [8] не только для управления задачами, но и для управления данными. Все данные регистрируются в файловом каталоге, и некоторым из них присваивается

метаинформация, благодаря которой возможно выбирать файлы по таким характеристикам как энергия столкновений, состав пучка, имя генератора и другим.

Пример MPD и опыт, полученный в процессе внедрения новых подходов, позволил предложить платформу распределенных вычислений на основе DIRAC в ОИЯИ как стандартное решение для массового запуска задач. ПРВ использовалась экспериментом Baikal-GVD - одним из ключевых экспериментов нейтринной программы ОИЯИ. Полученные данные были использованы для исследований, связанных с наблюдением мюонных нейтрино, и доложены на профильной конференции [9]. Эксперимент VM@N протестировал DIRAC для запуска задач, связанных с первичной обработкой данных, полученных на детекторе, и успешно выполнил 13 тысяч задач, связанных с моделированием.

Другой сферой использования платформы DIRAC в ОИЯИ стало участие в добровольных вычислениях, связанных с исследованиями вируса SARS-CoV-2 в рамках проекта Folding@Home [9]. Для исследований по COVID-2019 успешно задействованы свободные от основной деятельности ОИЯИ облачные ресурсы как Института, так и организаций его стран-участниц. Вклад всех облачных инфраструктур учтён как в информационной системе Folding@Home в рамках группы “Joint Institute for Nuclear Research”, так и в рамках системы учёта потреблённых ресурсов самого DIRAC.

На сегодняшний день благодаря интеграции ресурсов с помощью DIRAC на мощностях распределенной платформы было выполнено 1,07 миллиона задач. Количество проведённых вычислений оценивается в 5,66 миллионов HEPSPEC2006 дней, что является эквивалентом 874 лет расчетов на одном ядре центрального процессора. 90% процентов всех расчётов проведено на суперкомпьютере «Говорун», Tier1 и Tier2, 6 % - на облачной инфраструктуре ОИЯИ, 3% на NICA кластере.

Основным пользователем распределенной платформы является эксперимент MPD, доля которого составляет 90%. С использованием DIRAC выполняется программа сеансов массового моделирования данных эксперимента MPD. Более 750 миллионов событий были успешно смоделированы с помощью UrQMD, GSM, 3 Fluid Dynamics, vHLLU-UrQMD и других генераторов, 220 миллионов событий были впоследствии реконструированы. Общий объём полученных данных превышает 200 ТБ. 5% вычислений были выполнены экспериментом Baikal-GVD, 3% - Folding@Home, 2% - VM@N.

Помимо возможности проводить массовый запуск задач, интеграция большого количества гетерогенных ресурсов открыла возможность их централизованного анализа. Был разработан подход к оценке производительности разных вычислительных ресурсов, по результатам реализации которого предложен способ оценки на основе пользовательских задач, а не искусственных тестов [10]. Появилась возможность качественно оценить производительность разных ресурсов. В некоторых случаях это позволило понять, когда ресурс используется неэффективно. Также было обнаружено, что стандартный тест DIRAC (т.н. DIRAC Benchmark, DB) занижает производительность для некоторых моделей процессоров, о чём нами было впервые доложено разработчикам этого теста.

С ростом интенсивности использования платформы распределенных вычислений потребовалось создать дополнительные сервисы для пользователей. Один из сервисов

связан с оценкой производительности на уровне конкретной пользовательской задачи, что даёт возможность пользователю лучше понять режимы использования вычислительных ресурсов, сети, оперативной памяти. Также это позволяет на раннем этапе выявлять проблемы в программных пакетах, используемых для обработки данных. Второй сервис был специально разработан для эксперимента MPD. Он предоставляет пользователям интерфейс для запуска задач, оформленный не в терминах количества задач, исполняемых процессов, передаваемых аргументов, а в физических: количество событий, генератор, энергия, состав пучка [11]. Таким образом, физикам предоставляется интерфейс, который полагается на более близкие им понятия. Это первый пример создания проблемно-ориентированного интерфейса в рамках инфраструктуры DIRAC в ОИЯИ. В дальнейшем возможна разработка подобных интерфейсов для других научных групп данной платформы распределенных ресурсов.

Наиболее важными результатами являются:

1. Разработанный и внедрённый модуль для интеграции облачных ресурсов, работающих на базе программного обеспечения OpenNebula. С использованием разработанного модуля произведено объединение облака ОИЯИ и стран участниц для проведения совместных вычислений. Объединённая инфраструктура ОИЯИ и стран-участниц была использована для участия в проекте Folding@Home для изучения вируса SARS-CoV-2, а также для запуска задач моделирования эксперимента Baikal-GVD.
2. Интеграция в DIRAC вычислительных ресурсов: суперкомпьютера «Говорун», Tier1, Tier2, NICA кластера, облака ОИЯИ, облаков стран-участниц ОИЯИ, кластера UNAM. Проведена интеграция ресурсов хранения: дисковое хранилище EOS, ленточное хранилище dCache. Произведена адаптация рабочих процессов массового моделирования данных эксперимента MPD для запуска задач и сохранения данных с использованием платформы DIRAC. С 2019 года эта платформа используется для реализации программы сеансов массового моделирования данных эксперимента MPD.
3. Новый подход к анализу производительности распределённых гетерогенных вычислительных ресурсов. Применение данного подхода позволило определить производительность вычислительных ресурсов, интегрированных в DIRAC.

Список публикаций:

1. V. Korenkov, I. Pelevanyuk, P. Zrellov, and A. Tsaregorodtsev: “Accessing Distributed Computing Resources by Scientific Communities using DIRAC Services”// CEUR workshop proceedings, 2016, Vol. 1752, pp. 110-115
2. V. Gergel, V. Korenkov, I. Pelevanyuk, M. Sapunov, A. Tsaregorodtsev, and P. Zrellov: “Hybrid Distributed Computing Service Based on the DIRAC Interware”// Communications in Computer and Information Science, 2017, Vol. 706, pp. 105-118, doi:10.1007/978-3-319-57135-5_8
3. N.A. Balashov, N. A. Kutovskiy, A. N. Makhalkin, Y. Mazhitova, I. S. Pelevanyuk, and R. N. Semenov: “Distributed Information and Computing Infrastructure of JINR Member States' Organizations”// AIP Conference Proceedings, 2021, Vol. 2377, 040001, https://doi.org/10.1063/5.0063809

4. N.A. Balashov, R. I. Kuchumov, N. A. Kutovskiy, I. S. Pelevanyuk, V. N. Petrunin, and A. Yu Tsaregorodtsev: 2019. "Cloud Integration within the Dirac Interware"// CEUR workshop proceedings, 2019, Vol. 2507, pp. 256-260
5. V. Korenkov, I. Pelevanyuk, and A. Tsaregorodtsev: "Dirac System as a Mediator between Hybrid Resources and Data Intensive Domains"// CEUR workshop proceedings, 2019, Vol. 2523, pp. 73-84
6. V. Korenkov, I. Pelevanyuk, and A. Tsaregorodtsev: "Integration of the JINR Hybrid Computing Resources with the DIRAC Interware for Data Intensive Applications"// Communications in Computer and Information Science, 2020, Vol. 1223, pp. 31-46, doi:10.1007/978-3-030-51913-1_3
7. D.V. Belyakov, A. G. Dolbilov, A. N. Moshkin, I. S. Pelevanyuk, D. V. Podgainy, O. V. Rogachevsky, O. I. Streltsova, and M. I. Zuev: "Using the "Govorun" Supercomputer for the NICA Megaproject"// CEUR workshop proceedings, 2019, Vol. 2507, pp. 316-320
8. N. Kutovskiy, V. Mitsyn, A. Moshkin, I. Pelevanyuk, D. Podgayny, O. Rogachevsky, B. Shchinov, V. Trofimov, and A. Tsaregorodtsev. 2021. "Integration of Distributed Heterogeneous Computing Resources for the MPD Experiment with DIRAC Interware"// Physics of Particles and Nuclei, 2021, Vol. 52 (4), pp. 835-841, doi:10.1134/S1063779621040419
9. N.A. Kutovskiy, I. S. Pelevanyuk, and D.N. Zaborov: "Using distributed clouds for scientific computing"// Принята к опубликованию в CEUR workshop proceedings
10. I. Pelevanyuk: "Performance Evaluation of Computing Resources with DIRAC Interware"// AIP Conference Proceedings, 2021, Vol. 2377, 040006, <https://doi.org/10.1063/5.0064778>
11. A.A. Moshkin, I. S. Pelevanyuk, and O. V. Rogachevskiy: "Design and development of application software for the MPD distributed computing infrastructure"// Принята к опубликованию в CEUR workshop proceedings