

# **Интеграция эксперимента ALICE и суперкомпьютера Titan, с применением системы управления потоками заданий PanDA**

**А. О. Кондратьев**

Объединенный Институт Ядерных Исследований,  
Россия, 141980, г. Дубна, Московская область, ул. Жолио-Кюри, д.6

E-mail: <sup>a</sup> kondratyev@jinr.ru

Вычислительная среда эксперимента ALICE на Большом Адронном Коллайдере позволяет обрабатывать различные задачи, используя ГРИД сайты, расположенные по всему миру. Тем не менее, следующий запуск БАК подразумевает использование больших ресурсов, чем может предоставить ГРИД. Вследствие чего ALICE ищет способы наращивания ресурсов, в частности суперкомпьютеров.

Ключевые слова: ALICE, Titan, PanDA, AliEn

© 2016 Кондратьев Андрей Олегович

## Введение

Эксперимент ALICE (A Large Ion Collider) [ALICE Collaboration, 2004] – был основан для изучения взаимодействия тяжелых ионов в ядерных столкновениях в Большом Адронном Коллайдере (БАК). С момента старта БАК, эксперимент накопил и обработал десятки петабайт данных по всему миру. Около 1000 ученых ALICE из 28 стран мира ежедневно обращаются к этим данным для анализа.

ALICE использует более 100 000 процессоров, развернутых в ГРИД [Adam et al, 2015; Foster and Kesselman, 1999], охватывающих более чем 80 сайтов по всему миру. Однако, следующий запуск БАК подразумевает использование намного больших ресурсов, чем может предоставить ГРИД. Инфраструктуры будет достаточно для анализа и обработки данных, но она будет отставать от требований, предъявляемых к исследованиям. Поэтому очень важно найти дополнительные ресурсы. Для решения данной проблемы, ALICE занимается проектом по расширению существующей вычислительной модели, с целью включить в неё дополнительные ресурсы, такие как суперкомпьютеры.

## Суперкомпьютер Titan

Выбор пал на суперкомпьютер Titan [Titan at OLCF], являющийся третьим суперкомпьютером в мире по данным списка Top 500 [Top500 List], и расположен в Oak Ridge Leadership Computing Facility в Национальной Лаборатории Оук Ридж в США. Он имеет пиковую теоретическую производительность в 29 петафлопс. Titan стал первой крупномасштабной системой с гибридной архитектурой, одновременно использующей рабочие узлы с 16-ядерными процессорами AMD Opteron 6274 и графическими ускорителями NVIDIA Tesla K20. Он имеет 18,688 рабочих узлов с 299 008 процессорами. Каждый узел имеет 32 Гб оперативной памяти и использует Cray Gemini для передачи MPI сообщений другим узлам, но не имеет подключения к внешней сети. Titan обслуживается распределенной файловой системой Lustre, которая имеет 32 Пб дискового пространства и 29 Пб HPSS ленточного хранилища. Рабочие узлы Titan используют Compute Node Linux, который запускает программное окружение, основанное на ядре Linux, полученное с сервера SUSE Linux Enterprise.

## PanDA – Production and Distributed Analysis System

PanDA является акронимом от Production and Distributed Analysis System. Эта система разработана для эксперимента ATLAS, являющимся одним из экспериментов БАК, с целью управления и распределения больших объемов данных. PanDA имеет масштабируемую и гибкую структуру, что позволяет быстро адаптироваться к новым технологиям в области обработки, хранения, а также промежуточного программного обеспечения.

PanDA является системой, основанной на Пилотах [Nilsson, 2010]. В жизненном цикле задачи PanDA, пилотная задача запускается на вычислительных сайтах. Когда эти пилотные задачи запускаются на вычислительных узлах, они связываются с центральным сервером для получения payload(пользовательской задачи) и выполняют её. Использование систем, основанных на Пилотах, позволяет оптимизировать использование ресурсов, а также сводит к минимуму многие проблемы, связанные с неоднородностями ГРИД.

Одним из успешных решений применения PanDA является объединение вычислительных ресурсов эксперимента ATLAS и суперкомпьютера Titan.

Интеграция реализована с помощью модифицированного Пилота PanDA для запуска задач в пакетной очереди Titan и локального управления данными. Это дает PanDA новую возможность сбора информации о неиспользуемых рабочих узлах Titan в режиме реального времени,

что позволяет точно определить размер и продолжительность задач, запущенных на Titan, в соответствии с имеющимися свободными ресурсами. Эта возможность существенно снижает время ожидания задачи PanDA и одновременно повышает эффективность использования Titan[CERN CVMFS].

## **AliEN – Alice Environment**

AliEn (ALice ENvironment) [Buncic, Peters, Saiz, 2003] – распределенная вычислительная среда, разработанная для проекта ALICE Offline. Она позволяет получить доступ к распределенным вычислительным ресурсам и ресурсам хранения всем участникам эксперимента ALICE на Большом Адронном Коллайдере (БАК). Основная идея создания AliEn состоит в обеспечении функциональной вычислительной средой, удовлетворяющей потребности эксперимента, на этапе подготовки и, в то же время, предоставляющей устойчивый интерфейс для конечных пользователей.

AliEn был построен из множества компонентов с открытым исходным кодом. Это было сделано для того, чтобы повторно использовать их функциональные возможности, не изменяя их. Из всех 3 миллионов строк кода, только 1% является исходным кодом AliEn, остальные же 99% кода были импортированы в виде компонентов Open Source.

Для связи AliEn с ГРИД – инфраструктурой используется сервис VOBOX [Lorenzo, 2008]. Он позволяет эксперименту запускать собственные сервисы на сайтах, обеспечивает прямое взаимодействие с очередью пакетной обработки информации для запуска задач, а также связь между задачами и центральными сервисами ALICE. VOBOX получает и хранит всю информацию о задачах, сервисах и системах хранения, запущенных на сайтах.

## **Технология взаимодействия AliEN и Titan**

Взаимодействие AliEn и Titan происходит с помощью PanDA WMS, сервер которой расположен в облаке Amazon EC2 и позволяет отправлять задачи на Titan. Выбор облака Amazon EC2 был обусловлен тем, что в нем имелась установка сервера PanDA с возможностью отправки задач на Titan. Использование сервера осуществлялось с использованием технологий SSH, SSL и X509. Для прохождения авторизации на сервере необходимо создание прокси-сертификата пользователя.

Так как рабочие узлы Titan не имеют подключения к внешней сети, было выполнено копирование сегмента ALICE файловой системы CVMFS [CERN CVMFS] для того, чтобы обеспечить работу программного обеспечения ALICE.

Была выполнена установка сервиса VOBOX и среды AliEn на виртуальную машину в ЦЕРН, а также выполнена установка клиента файловой системы CVMFS для корректной работы VOBOX.

AliEn на 99% состоит из компонентов с открытым кодом. Поэтому был разработан компонент, позволяющий отправлять задачи на сервер PanDA, обеспечивающий, в дальнейшем, их выполнение на Titan.

VOBOX автоматически циклично осуществляет проверку задач, готовых к отправке на вычисление. Как только появляется задача, готовая к отправке, запускается компонент AliEn отправляющий эту задачу на сервер PanDA, осуществляющий дальнейшее выполнение задачи. На Рис. 1 представлена схема взаимодействия среды AliEn и суперкомпьютера Titan.

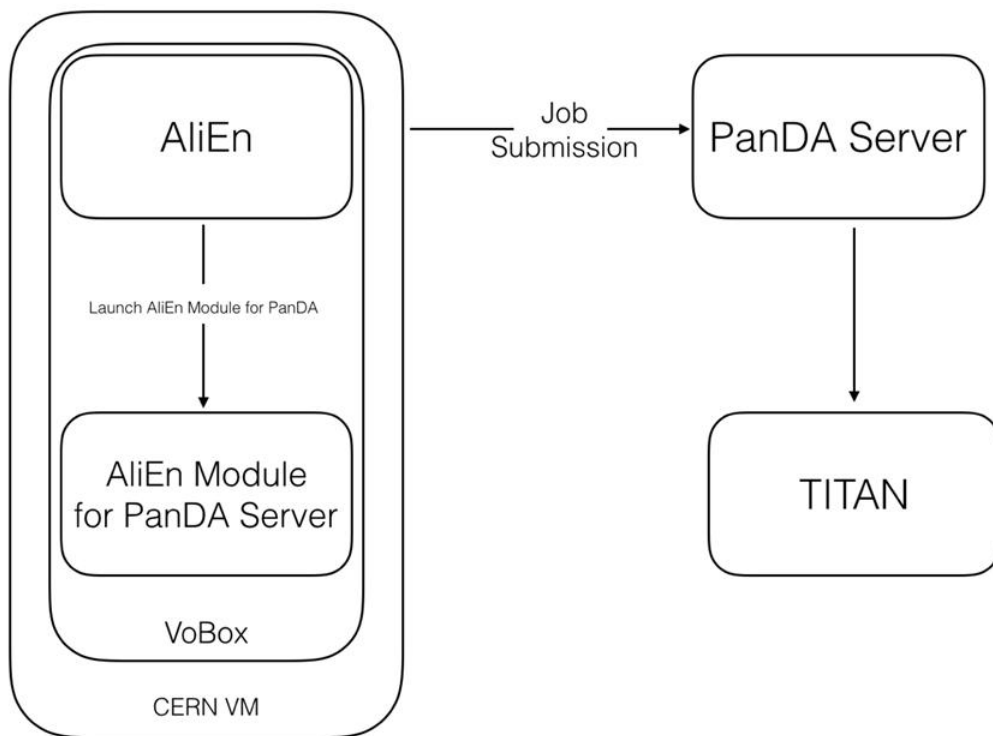


Рис. 1. Схема взаимодействия среды AliEn и суперкомпьютера Titan

Данная схема работает следующим образом, задача ALICE представляет собой файл с расширением `.jdl`. В данном файле содержится информация необходимая для запуска Пилота. Как только этот файл отправляется в сервис VOBOX, задача считается сформированной и готовой к отправке на вычисление. Получив задачу, VOBOX обращается к среде AliEn для её дальнейшей обработки. После получения всех необходимых входящих параметров от сервиса, AliEn запускает компонент формирования задачи для отправки на сервер PanDA.

Все компоненты AliEn написаны на языке программирования высокого уровня Perl. Основной причиной использования этого языка программирования является наличие большого количества компонентов с открытым кодом, обеспечивающим поддержку криптографии, платформы клиент-сервер, а также простую интеграцию с Веб-сервисами для мониторинга. Разработанный компонент AliEn формирует задачу PanDA, используя входящие параметры, переданные сервисом VOBOX, и отправляет их на сервер PanDA.

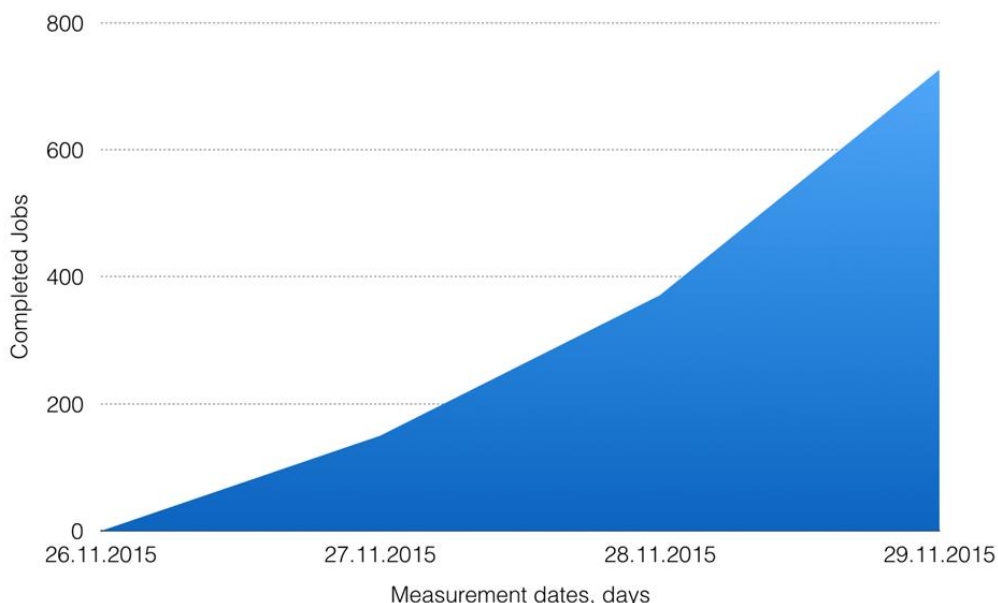


Рис. 2. Количество выполненных задач ALICE на Titan

## Результаты

Результаты работы вышеописанной технологии наглядно показаны на Рис. 2. Данная диаграмма демонстрирует количество выполненных задач ALICE на Titan в период с 26 по 29 ноября 2015 года. 1 декабря 2015 года общее число выполненных задач ALICE превысило 700.

Автор выражает благодарность коллаборациям ALICE, ATLAS и BigPanDA, а также Климентову Алексею Анатольевичу и Олейнику Даниле Анатольевичу за помощь в создании и развитии проекта по интеграции эксперимента ALICE и суперкомпьютера Titan.

## Список литературы

- ALICE Collaboration, F. Carminati, P. Foka, P. Giubellino, A. Morsch, G. Paic, J-P Revol et al.* ALICE: Physics Performance Report – Vol. 1 // J. Phys. G: Nucl. Part. Phys. - №30 — 2004 - PII: S0954-3899(04)83684-3
- Adam J. et al.* ALICE Technical Design Report: Upgrade of the Online – Offline computing system // Tech. rep. CERN-LHCC-2015-006/ALICE-TDR-019 2015
- Foster I. and Kesselman C. (eds).* The Grid: Blueprint for a New Computing Infrastructure, (Morgan Kaufman) 1999
- Titan at OLCF Web page // URL <https://www.olcf.ornl.gov/titan/>
- Top500 List // URL <http://www.top500.org>
- Nilsson P. ,* The ATLAS PanDA Pilot in Operation // in Proc. Of the 18<sup>th</sup> Int. Conf. on Computing in High Energy and Nuclear Physics - (CHEP2010) — 2010
- Buncic P. , Peters A.J. , Saiz.P.* The AliEn System, status and perspectives // Computing in High Energy and Nuclear Physics - ECONF C0303241(2003) - MOAT004 2003
- Lorenzo P. M.* The management of the VOBOX // II Brazilian LHC Computing Workshop 2008
- CERN CVMFS Web page // URL <https://cernvm.cern.ch/portal/filesystem>

# Integration of ALICE experiment and Titan supercomputer with using PanDA Workload Management System

**A. O. Kondratyev**

Laboratory of information technology, Joint institute for nuclear research  
Russia, 141980, Dubna, Moscow region, Joliot-Curie, 6

E-mail: [kondratyev@jinr.ru](mailto:kondratyev@jinr.ru)

The computing environment of the ALICE experiment at the Large Hadron Collider allows to handle different tasks using GRID sites located around the world. However, following run of the LHC involves using more resources that can provide GRID. The result is that ALICE is looking for ways to increase resources, in particular supercomputers.

Keywords: ALICE, Titan, PanDA, AliEn