

ОБНОВЛЕНИЕ РОССИЙСКИХ ГРИД-САЙТОВ ЭКСПЕРИМЕНТА ALICE В 2022–2023 ГГ. ДЛЯ РАБОТЫ В РЕЖИМЕ ВЫСОКОЙ СВЕТИМОСТИ LHC

*А. К. Зароченцев**

Санкт-Петербургский государственный университет, Санкт-Петербург, Россия

Представлены результаты работ по обновлению программного обеспечения российских грид-сайтов для продолжения обработки данных ALICE и в будущем, включая этап высокой светимости работы Большого адронного коллайдера (High-Luminosity LHC, HL-LHC). Также представлены опыт работы с одной из грид-платформ и некоторые практические кейсы масштабирования обновленного промежуточного программного обеспечения на другие российские площадки в 2022–2023 гг.

We present our results on updating the middleware of Russian GRID sites to be able to continue processing ALICE data in the future, including the HL stage of the Large Hadron Collider operation. We will share our experience with one of the GRID sites and discuss some practical cases of scaling the updated middleware to other Russian sites in 2022–2023.

PACS: 89.20.Ff; 07.05.Tp

ВВЕДЕНИЕ

В 2021 г. начался третий этап работы (Run 3) Большого адронного коллайдера (LHC) после длительного перерыва. Этот этап знаменателен модернизацией детекторных систем установок для увеличения их чувствительности и возможности работы в повышенных радиационных условиях и большого потока событий, а также, как следствие, роста объемов данных. Это привело как к повышенным требованиям к системам хранения, так и к новым требованиям к системе обработки данных. Во время подготовки к Run 3 основное внимание было уделено обновлению до новых, более надежных версий программного обеспечения (ПО), включая операционные системы, менеджеры управления задачами и системы хранения данных [1]. После старта третьего этапа появились задачи по обработке большего объема данных как моделирования методом Монте-Карло, так и реальных экспериментов. Поезда hyperloop [2]

* E-mail: andrey.zar@gmail.com

при запуске задач на грид-сайты стали использовать более одного ядра на задачу. Потребовалась необходимость запуска на грид-сайтах уровней Tier1–Tier2 многоядерных задач. К концу 2023 г. новый стандарт вычислительной системы эксперимента ALICE подразумевал запуск только 8-ядерных задач.

К 2022 г. также появилась необходимость обновления системы авторизации и аутентификации пользователей. В связи с использованием в проекте Worldwide LHC Computing Grid (WLCG) [3] нового ПО и новых протоколов, таких как OAuth2.0 [4] и OIDC [5], возрос интерес к применению более современных и гибких моделей аутентификации и авторизации. Современные реалии характеризуются стандартными потоками авторизации через интернет, инфраструктурой аутентификации и авторизации (AAI) [6] в соответствии с архитектурой проекта AARC [7]. Все это привело к необходимости сменить устаревшую систему авторизации и аутентификации, основанную на сертификатах x509, к чему-то новому. Возник план отказа от устаревшей авторизации GSI [8] и перехода на WLCG Token [9] к 2024 г.

МНОГОЯДЕРНЫЕ ЗАДАЧИ

Согласно рекомендациям IT-департамента коллаборации ALICE к концу 2023 г. для возможности поддержки вычислительной системы ALICE необходимо было привести сайты Tier1–Tier2 к новому стандарту запуска задач. Для российских сайтов эта задача выполнялась осенью 2022 – весной 2023 г.

В принципе, большие российские сайты, которые поддерживают несколько экспериментов, включая ATLAS и CMS, уже имели опыт запуска многоядерных задач для этих экспериментов. Поэтому для сайтов Tier1 в Курчатовском институте, Tier2 в ИФВЭ и ОИЯИ переход на 8-ядерные задачи прошел для эксперимента ALICE без особых затруднений. Для Tier2 в ПИЯФ и СарФТИ НИЯУ МИФИ пришлось делать некоторую реконфигурацию: ПИЯФ использовал 4-ядерные виртуальные машины, которые при управлении через Torque [10] блокировали очередь. СарФТИ имел 24-ядерные машины, но все равно было принято решение обновить систему управления задачами на более современный Slurm [11]. В итоге на этих двух сайтах переход на 8-ядерные задачи сопровождался переходом с Torque на Slurm.

К сожалению, количество ЦПУ на сайтах Tier2 в Троицке и ИТЭФ, на которых имело бы смысл запускать 8-ядерные задачи, невелико. В эксперименте ALICE отказались от использования вычислительных элементов этих сайтов.

Сайт Tier2 СПбГУ в течение 10 мес. по техническим причинам был выведен из работы, но летом 2023 г. был восстановлен и тоже настроен на 8-ядерные задачи без смены системы управления задачами.

WLCG TOKEN

В 2022–2023 гг. в WLCG был запланирован переход от авторизации GSI с помощью сертификатов x509 на авторизацию WLCG Tokens. При этом устаревшая система VOMS [12] была заменена на группы в IAM [13].

К началу работ авторизацию с помощью WLCG Token официально поддерживали два менеджера управления задачами (CE) — HTCondor [14] и ARC-CE [15]. Но реально задачи, подписанные только Token, мог принимать только HTCondor. ARC-CE дополнительно требовал подпись сертификатом x509. Так как все российские сайты использовали именно ARC, возникла необходимость разобраться с возможностью дальнейшего использования ARC при полном переходе на WLCG Tokens.

Обновленная версия ARC-7, по крайней мере официально, согласно документации поддерживает такой запуск задач. Однако при первых тестах был выявлен ряд ошибок — оставшаяся привязка к прокси x509 в клиенте ARC и на серверной части. После выявления и решения первых очевидных ошибок на виртуальном стенде было принято решение установить обновленный ARC-7 на работающий российский сайт CapFTI, участвующий в плановой обработке данных. После запуска ARC-7 на «живом» сайте обнаружился ряд новых ошибок, в основном они связаны с тем, что ARC-7 полностью перешел на LDAP-схему glue2 [16] системы мониторинга ресурсов и полностью отказался от протокола gridftp для передачи данных задач. Если клиента грид-системы эксперимента ALICE JAlien [17] удалось исправить для новых схем, то с системой мониторинга GOCDB [18] возникли проблемы. Данная система сконфигурирована еще под параметры ARC-5 и очень инерционна. В итоге пришлось вывести CapFTI из списка зарегистрированных российских сайтов, перевести его в статус «приостановлен» («suspended») в GOCDB. К сожалению, это не первый российский сайт ALICE, который вынужден был прекратить регистрацию в GOCDB.

ЗАКЛЮЧЕНИЕ

Благодаря проделанной работе в настоящее время основные российские сайты продолжают участвовать в обработке данных эксперимента ALICE и имеют техническую возможность осуществлять эту работу и в будущем. Есть сложности с регистрацией сайтов в GOCDB, но они связаны, скорее, с медленным развитием самой структуры GOCDB. Для решения этой проблемы требуется привлечение как разработчиков ARC, так и заинтересованных специалистов со стороны эксперимента ALICE. Данная работа была запланирована на конец 2023 г. – начало 2024 г.

Работа выполнена при поддержке Санкт-Петербургского государственного университета (грант ID: 94031112).

СПИСОК ЛИТЕРАТУРЫ

1. *Kiryarov A., Zarochentsev A.* Participation of Russian Institutes in the Processing and Storage of ALICE Data // CEUR Workshop Proc. 2021. V. 3041. P. 49–52.
2. *Quishpe R., Grosse-Oetringhaus J.F., Cruceru R., Grigoras C.* Hyperloop — the ALICE Analysis Train System for Run 3 // PoS (LHCP2021). 2021. V. 397. P. 250; arXiv: 2109.09594.
3. Welcome to the Worldwide LHC Computing Grid. <https://wlcg.web.cern.ch/> (accessed 22.09.2023).
4. OAuth 2.0 Is the Industry-Standard Protocol for Authorization. <https://oauth.net/2/> (accessed 22.09.2023).
5. What Is OpenID Connect. <https://openid.net/connect/> (accessed 22.09.2023).
6. Authentication and Authorization Infrastructure. <https://elixir-europe.org/platforms/compute/aai> (accessed 22.09.2023).
7. Authentication and Authorization for Research and Collaboration Blueprint Architecture. <https://aarc-project.eu/architecture/> (accessed 22.09.2023).
8. *Cai Z.* A Password-Based Grid Security Infrastructure // Proc of the Second Intern. Conf. on the Digital Society. 2008. P.157–164; doi: 10.1109/ICDS.2008.39.
9. *Bockelman B., Ceccanti A., Dack T., Dykstra D., Litmaath M., Sallé M., Short H.* WLCG Token Usage and Discovery // Eur. Phys. J. Web Conf. 2021. V. 251. P.02028.
10. Torque Resource Manager. <https://adaptivecomputing.com/cherry-services/torque-resource-manager/> (accessed 22.09.2023).
11. Slurm Is Highly Scalable Cluster Management and Job Scheduling System. <https://slurm.schedmd.com/overview.html> (accessed 22.09.2023).
12. The Virtual Organization Membership Service. <https://wlcg-ops.web.cern.ch/voms> (accessed 22.09.2023).
13. What Is Identity and Access Management (IAM)? <https://www.onelogin.com/learn/iam> (accessed 22.09.2023).
14. HTCondor Software Suite. <https://htcondor.org/> (accessed 22.09.2023).
15. Advanced Resource Connector. <https://www.nordugrid.org/arc/> (accessed 22.09.2023).
16. *Andreozzi S. and GLUE Working Group.* GLUE v. 2.0 — Reference Realization to LDAP Schema. <https://redmine.ogf.org/attachments/161/ogf-glue-2-to-LDAP-v7.3-draft5.pdf> (accessed 22.09.2023).
17. JAliEN — ALICE Environment GRID Framework. <https://alien.web.cern.ch/> (accessed 22.09.2023).
18. GOCDB Is the Official Repository for Storing and Presenting EGI Topology and Resources Information. <https://goc.egi.eu/> (accessed 22.09.2023).