

Ресурсы и сервисы центрального распределенного информационно-вычислительного комплекса ОИЯИ: текущий статус и перспективы ¹

Н.С.Астахов, В.В.Иванов, В.В.Кореньков, В.В.Мицын, А.П.Сапожников,
Е.А.Тихоненко

Лаборатория информационных технологий, ОИЯИ

Созданный за последние годы распределенный комплекс на базе операционных систем семейства Linux [1] продолжает развиваться и совершенствоваться в соответствии с современными тенденциями и особенностями построения компьютерных инфраструктур крупных научных центров.

За 2004-2005 годы был выполнен ряд работ по развитию аппаратной инфраструктуры ЦИВК. Особое внимание было уделено развитию системы дисковых массивов. На первом этапе были приобретены и запущены в эксплуатацию два файловых сервера с аппаратными RAID-массивами общей емкостью 4 Терабайта (ТВ), что позволило довести общее дисковое пространство для пользователей ЦИВК до 17 ТВ. На втором этапе исследовалась оригинальная архитектура системы дисковых массивов CertonRAID 100, которая была разработана в университете Гейдельберга (Германия). Эта система обладает следующими достоинствами: высокая производительность, гибкость, масштабируемость, дополнительные средства защиты информации, умеренная стоимость. Был подготовлен контракт на приобретение системы дисковых массивов CertonRAID 100, состоящей из 13 комплектов по 3 ТВ. Таким образом, общая емкость системы составляет 39 ТВ, из них примерно 33 ТВ могут быть доступны для пользователей. В настоящее время происходит наладка системы и ее включение в состав ЦИВК. После ввода системы в эксплуатацию, она будет тщательно протестирована в условиях высокой загрузки. После ввода этой системы в эксплуатацию общее дисковое пространство для пользователей ЦИВК будет составлять около 50 ТВ.

Вычислительная мощность кластера за последний год тоже возросла, но не так существенно. Были введены в эксплуатацию 16 современных двухпроцессорных вычислительных узлов (из них 8 использовались для замены морально устаревших узлов фермы общего назначения). Были приобретены несколько серверов, в том числе новый сервер баз данных, работающий под управлением СУБД ORACLE 10G, сервер для Грид-мониторинга, несколько серверов и рабочих станций для тестирования нового программного обеспечения. Выполнена модернизация нескольких вычислительных узлов для улучшения характеристик работы кластера. Завершается перевод инфраструктуры ЦИВК на технологию 1 GB Ethernet. Установлена система поддержки бесперебойного питания APC Symmetra, которая позволит увеличить время непрерывной работы всех машин кластера ЛИТ.

В настоящий момент комплекс имеет в своем составе более 110 процессоров суммарной производительностью около 100 kSI2K (1 kSI2K соответствует процессору Intel Xeon 2.8 ГГц), используемых как счетные узлы, и ряда серверных машин, предоставляющих общее дисковое пространство, базы для регистрации и авторизации, почтовые и другие сервисные услуги. Общее дисковое пространство для пользователей

¹Работа частично поддержана РФФИ, гранты № 04-07-90162-в, № 04-07-90174-в, № 05-07-08018-офи-а.

составляет около 17 Терабайт. Основная часть счетных ресурсов доступна пользователям через систему пакетной обработки, а некоторая часть ресурсов подключена к глобальной инфраструктуре LCG-2 и доступна с помощью средств Грид. Подготовка и редактирование файлов, программ, заданий, документов, работа с сетевыми сервисами, приложениями, анализ и визуализация результатов и т.д. осуществляется пользователем во время сеансов работы на специализированных интерактивных машинах.

Произведен переход от операционной системы LINUX (CERN Red Hat 7.3) к операционной системе Scientific Linux CERN Release 3.0.5 (SLC3) с соответствующим дополнительным программным обеспечением (ПО):

- компиляторы Gnu C, C++, F77;
- компиляторы Intel C, C++, F90;
- отладчики gdb, ddd;
- графические оболочки gnome, kde;
- различные утилиты пользователей, включая ROOT;
- различные библиотеки, включая CERN library.

ОС SLC3 - это вариант RedHat Enterprise Linux 3 (RHEL3), собранный из исходных текстов с некоторыми модификациями от FNAL и CERN и являющийся в настоящее время основной базовой операционной системой в CERN и многих других научных организациях.

На ферме из 20 вычислительных узлов, включенных в инфраструктуру LCG-2, на протяжении 2004-2005 гг. дважды обновлялись версии ПО LCG: осуществлялся переход на LCG 2.4 и LCG 2.6; установлены следующее прикладное ПО для экспериментов на LHC: VO-alice-geant42ndProd_p01, VO-cms-CMKIN_4_4_0_dar, VO-cms-OSCAR_3_6_5_SLC3_dar, VO-cms-ORCA_8_7_1_SLC3_dar, VO-cms-ORCA_8_4_0, VO-lhcb-RTTC-v1, VO-lhcb-Gaudi-v15r5.

На смену прежней системе пакетной обработки PBS установлена его новая усовершенствованная версия, называемая Torque и интеллектуальный диспетчер Maui. В связи с появлением новых возможностей была переконфигурирована система очередей, и теперь для установки заданий в очередь имя очереди указывать не надо, поскольку задание направляется для выполнения в соответствии с указанными параметрами (требуемое время выполнения; необходимая оперативная и виртуальная память; количество процессоров для задач параллельных вычислений и т.п.).

На кластере для параллельных вычислений, где узлы связаны высокоскоростной сетью Myninet, установлено ПО поддержки распаралеливания задач пользователя - MPI.

Для улучшения характеристик работы сетевых сегментов ЦИВК протокол NFS был заменен на протокол AFS для доступа к ПО, а для доступа к данным внедряются интеллектуальные средства DPM (Disk Pool Manager) и dCache.

Пользователи, работающие в физических экспериментах ALICE, ATLAS, CMS, COMPASS, D0, DIRAC, H1, HARP, HERA-B, HERMES, IREN, KLOD, LHCb, NA48, NEMO, OPERA, STAR, (всего более 130 пользователей) объединены в отдельные

подгруппы. Для каждой подгруппы выделено дополнительное дисковое пространство для проведения моделирования, хранения и обработки экспериментальных и модельных данных. Имеется необходимое специализированное ПО для каждого из перечисленных выше экспериментов, в том числе AliRoot v4-01-Rev07, ORCA_8_7_1, Gaudi 0.16.12. Принято решение о переносе специализированного ПО коллабораций в AFS.

Использование ресурсов ЦИВК ОИЯИ

По состоянию на октябрь 2005 года количество зарегистрированных пользователей на центральном PC-Linux кластере ОИЯИ 501 человек:

Подразделение	Количество Пользователей
ЛИТ	171
ЛЯП	104
ЛФЧ	53
ЛВЭ	44
внешние пользователи (из стран-участниц ОИЯИ и московских институтов)	34
ЛЯР	28
ЛТФ	14
ЛНФ	12
Управление	9

Использование локальных ресурсов Linux кластера (по процессорному времени) за период с января по сентябрь 2005 года:

Подразделение	Процент использования процессорного времени
ЛТФ	25%
ЛИТ	23%
ЛЯП	17%
ЛВЭ	15%
ЛФЧ	7%
ЛНФ	5%
LHC-production	5%
ЛЯР	3%

Массовое моделирование событий для экспериментов ALICE [3], ATLAS, CMS и LHCb в основном производилось на ресурсах ЦИВК ОИЯИ, входящих в инфраструктуру LCG-2.

В ближайшие планы входит использование средств систем DPM и dCache для удаленного доступа к данным, издание руководства для пользователей по работе в инфраструктуре LCG. Для повышения уровня защиты работы пользователя на ЦИВК внедряются новые протоколы и ПО – Kerberos5 и LDAP. После решения вопроса

с внедрением Kerberos5 и LDAP, все имеющиеся вычислительные ресурсы, включая машины сайта LCG, будут включены в общий пул вычислительных узлов ЦИВК для доступа к ним всех пользователей. Система распределения вычислительных ресурсов Maui будет учитывать приоритеты групп пользователей и уже израсходованное счетное время.

Для успешной работы сотрудников ОИЯИ на Linux кластере было подготовлено и издано подробное руководство для пользователей [2]. Информация о всех изменениях и нововведениях на кластере публикуется на странице <http://www.jinr.ru/unixinfo/pc/news-pc.html>

Список литературы

- [1] *В.В.Кореньков, В.В.Мицын, Е.А.Тихоненко*, Развитие и использование ресурсов и сервисов центрального информационно-вычислительного комплекса ОИЯИ, в сборнике “Годовой отчет Лаборатории информационных технологий за 2003 год”, 2004-53, Дубна, 2004, с.15-21.
- [2] *В.Галактионов, Т.Голоскокова, Н.Громова, А.Гусев, В.Мицын, Ж.Мусульманбеков, И.Некрасова, В.Позе, А.Сергеев, Е.Тихоненко*, Руководство для пользователей LINUX кластера ЛИТ ОИЯИ, Р11-2004-184, Дубна, 2004, 123 стр.
- [3] *Y.Bugaenko, G.Shabratova, V.Korenkov, V.Mitcin et.al.* Participation of Russian sites in the data challenge of ALICE experiment in 2004. In Proceedings of "Computing in High Energy and Nuclear Physics 2004", CERN, 2004, vol.2, pp.802-804.