

Развитие Центрального вычислительно-информационного комплекса ОИЯИ в 2008-2009 гг. и текущее состояние программно-аппаратной среды

Н.С. Астахов, А.Г. Долбилов, В.В. Иванов, В.В. Кореньков, В.В. Мицын,
С.В. Мицын, В.В. Трофимов

Лаборатория информационных технологий ОИЯИ, Дубна

Центральный информационно-вычислительный комплекс (ЦИВК) ОИЯИ организован как единый информационно-вычислительный ресурс, предназначенный для обеспечения всех направлений деятельности сотрудников ОИЯИ. Счетные ресурсы и ресурсы для хранения данных управляются базовым программным обеспечением (ПО), позволяющим использовать средства ЦИВК как удаленным пользователям в рамках международных проектов (WLCG, EGEE), так и локальным пользователям ЦИВК ОИЯИ.

Настройки ПО ЦИВК обеспечивают оптимальное использование счетных ресурсов и поддержку наиболее универсальных и защищенных методов доступа к хранилищам данных. Распределение и учет счетных ресурсов реализован на основе системы пакетной обработки (batch) torque и планировщика ресурсов maui. Доступ к данным выполняется через dCache и частично через NFS. Доступ к общему ПО и домашним директориям пользователей реализован через AFS и NFS. Для регистрации и идентификации локальных пользователей ОИЯИ применяется система kerberos5.

Для массовой установки базового ПО на вычислительные узлы (счетные машины) в ЛИТ разработана система удаленной установки (СУУ) ПО. СУУ построена на стандартных средствах OS Linux с использованием элементов из ПО Warewulf. Она позволяет автоматизировать массовую установку ПО на новые машины и замену старых версий операционной системы.

В 2008 и 2009 годах проводилась большая работа по приобретению нового оборудования для инфраструктуры ЦИВК: счетных машин, дисковых серверов, серверов общего назначения, сетевых маршрутизаторов, систем питания и других средств. Перед закупкой оборудования конкретного типа проводилось его тестирование в ЛИТ, в среде грид-сегмента ОИЯИ. Обычно тестировалось оборудование сразу нескольких компаний. Для выбора наилучшего варианта проводился сравнительный анализ оборудования как по его техническим характеристикам, так и по соотношению качество/цена. Результаты проведенных испытаний оборудования представлены на сайтах [1,2].

Текущее состояние ресурсов ЦИВК

В настоящее время в состав ЦИВК входят следующие счетные машины:

- 60 64-х битных машин:
2 процессора Xeon 5150 (2 ядра на процессор);
тактовая частота 2.66 ГГц;
4096 Кбайт кэш на ядро;
8 Гбайт оперативная память;
160 Гбайт диск на 20 и 250 Гбайт на 40 машинах;
1 Гбит Ethernet.
- 30 64-х битных машин:
2 процессора Xeon E5430 (4 ядра на процессор);
тактовая частота 2.66 ГГц;
6144 Кбайт кэш на ядро;
16 Гбайт оперативная память;
250 Гбайт диск;
1 Гбит Ethernet.
- 10 64-х битных машин:
2 процессора Xeon X5450 (4 ядра на процессор);
тактовая частота 3.00 ГГц;
6144 Кбайт кэш на ядро;
16 Гбайт оперативная память;
250 Гбайт диск;
1 Гбит Ethernet;
2x infiniband.
- 10 64-х битных машин:
2 процессора Xeon E5410 (4 ядра на процессор);
тактовая частота 2.33 ГГц;
6144 Кбайт кэш на ядро;
16 Гбайт оперативная память;
2 × 160 Гбайт диск;
1 Гбит Ethernet;
2x infiniband.
- 40 64-х битных машин, по 2 машины в одном корпусе высотой 2 дюйма (U2):
2 процессора Xeon E5420 (4 ядра на процессор);
тактовая частота 2.50 ГГц;
6144 Кбайт кэш на ядро;
16 Гбайт оперативная память;
250 Гбайт диск;
1 Гбит Ethernet.

Так как 2-х и 4-х ядерные процессоры содержат, соответственно, два и четыре независимых процессора на одном кристалле, то на данный момент парк счетных машин ЦИВК-а включает 960 64-х битных ЦПУ. Все эти ЦПУ доступны пользователям ЦИВК ОИЯИ и пользователям грид-среды через единую систему пакетной обработки заданий - batch.

Для разработки собственного ПО и обеспечения доступа к ресурсам ЦИВК для пользователей ОИЯИ установлены 5 машин с интерактивным доступом:

- 2 x 64-х битных машин (Core 2 Duo, 4 Гбайт RAM, 120 Гбайт + 250 Гбайт HDD, 1 Гбит Ethernet);
- 2 x 32-х битных машин (Dual Athlon 2400+, 2 Гбайт RAM, 40 Гбайт + 120 Гбайт HDD, 1 Гбит Ethernet);

- 64-х битная машина (Xeon E5420, 8 x Core, 16 Гбайт RAM, 250 Гбайт HDD, 1 Гбит Ethernet).

Имеется несколько серверов поддержки работы пользователей и служб ОИЯИ: batch, WWW, БД mysql и Oracle; e-mail; DNS и др. Эти серверы имеют различные аппаратные платформы – от 32-х битных Pentium-3, до 64-х битных Xeon и Opteron.

Основной системой для хранения больших объемов информации в ЦИВК является аппаратно-программный комплекс dCache. В настоящий момент в него включены:

- 12 серверов основных интерфейсов системы dCache;
- 32 системы хранения данных (Pool).

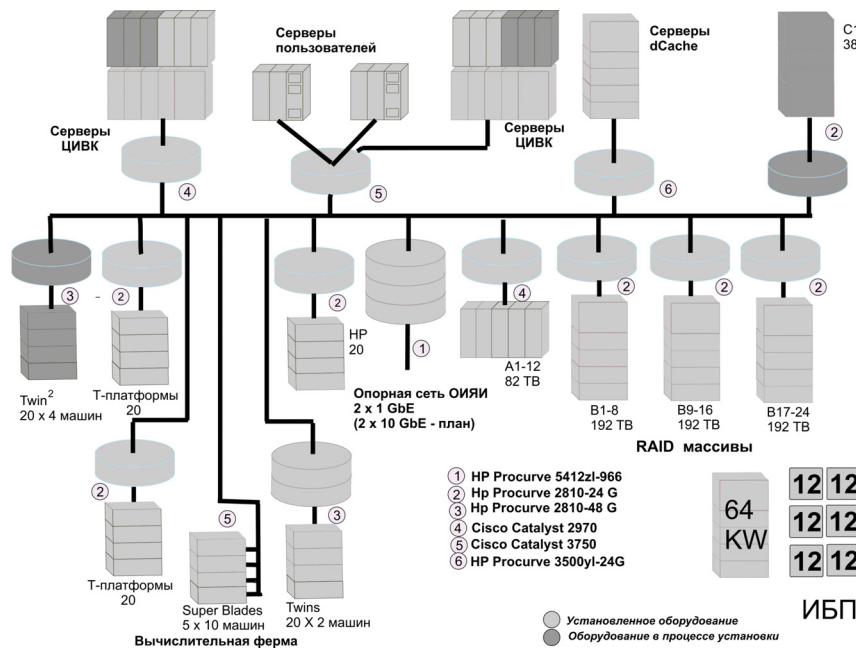


Рис. 1: Структурная схема ЦИВК ОИЯИ

Несколько пользовательских групп ЦИВК-а используют систему доступа к удаленной информации XROOTD. Для обеспечения работы этой системы создан аппаратно-программный комплекс, содержащий 1 сервер обработки запросов к системе, 6 систем хранения данных. В этом комплексе используются разные аппаратные платформы: Athlon64 X2, Athlon 2000+, Pentium 4, Pentium D, Xeon. Все системы хранения построены с использованием аппаратного механизма RAID5. Общая доступная емкость системы составляет ~500 Тбайт.

В ЦИВК ОИЯИ установлено 6 серверов AFS – высокозащищенной, распределенной файловой

системы, которая используется для сопровождения домашних директорий пользователей и для доступа к общему ПО для всего ОИЯИ. Общее пространство AFS в ОИЯИ составляет ~3.56 Тбайт.

Для обслуживания сайта WLCG в ОИЯИ (сайт - это отдельный кластер в распределенной системе WLCG и EGEE) установлено 24 сервера под системой gLite (ПО EGEE/WLCG). Кроме функций поддержки работы самого сайта JINR-LCG2, часть серверов реализуют важные сервисы и функции поддержки Российского сегмента проектов LCG и EGEE.

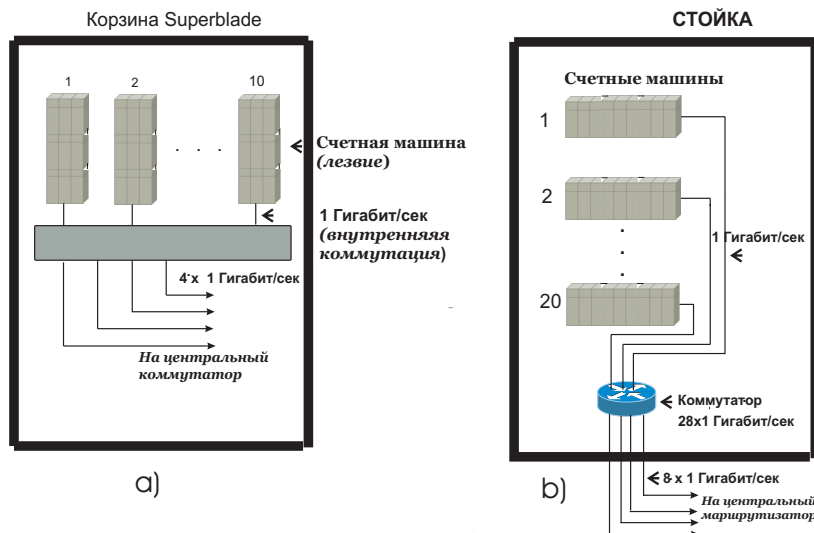


Рис. 2: Схема подключения счетных машин к локальной сети: а) -корзинный вариант Superblade, б) - стойечный вариант

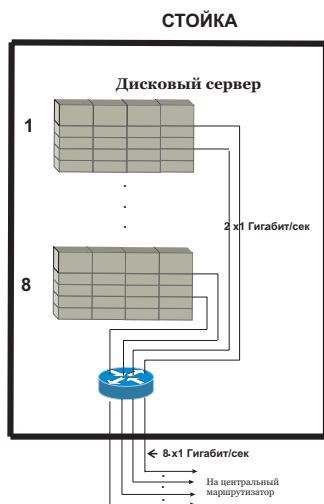


Рис. 3: Схема подключения дисковых серверов к локальной сети ЦИВК ОИЯИ

В состав ЦИВК ОИЯИ входит тестовая ферма из 8 машин, предназначенная для проверки и отладки новых версий ПО gLite.

Для обеспечения высокой пропускной способности локальной сети ЦИВК и минимального времени доступа к данным и файлам применяется агрегирование (TRUNK) нескольких соединений 1 Гбит Ethernet в единый виртуальный канал с увеличенной пропускной способностью.

На рис. 2 приведена логическая схема сетевого подключения корзины Superblade из 10-ти вычислительных лезвий и 10-ти счетных машин, установленных в стойке, к локальной сети ЦИВК. На рис. 3 показана схема подключения стойки с 8-ю дисковыми серверами к основному маршрутизатору локальной сети ЦИВК.

Методика агрегирования сетевых линков позволила удовлетворить требования по скорости доступа к данным из счетных задач как для локальных пользователей ОИЯИ, так и для пользователей среды WLCG/EGEE. Этот подход также позволил не переводить основную часть локальной сети ЦИВК-а на более мощные каналы (10

GbE, или Infiniband), что потребовало бы весьма значительных денежных затрат.

Базовое программное обеспечение ЦИВК

Базовой операционной системой для ЦИВК является OS Linux, а именно, дистрибутив Scientific Linux - SL4.8 и SL5.4. Этот дистрибутив сопровождается такими крупными научными центрами, как CERN и FNAL, распространяется бесплатно и является аналогом широкоизвестного коммерческого продукта фирмы Redhat Enterprise Linux. SL4.8 и SL5.4 также выбраны в качестве базовой ОС в проектах WLCG и EGEE.

С 2004 года ЦИВК ОИЯИ является составным элементом мировой вычислительной грид инфраструктуры - вначале в среде LCG, а с 2007 года в среде WLCG/EGEE. Для интеграции ресурсов ЦИВК-а в грид-среду на счетных машинах, ряде серверов, а также на интерактивных машинах установлено ПО промежуточного уровня - gLite. В настоящее время вклад вычислительных ресурсов ЦИВК по отношению ко всему Российскому сообществу WLCG составляет от 40 до 50 про-

ентов и чуть меньше по ресурсам для хранения данных. Сайт ОИЯИ является одним из наиболее эффективно работающих сайтов всей инфраструктуры WLCG по критериям надежности и доступности.

Для расширения возможностей работы локальных пользователей на ЦИВК установлено дополнительное ПО: оптимизирующие компиляторы Intel, система распараллеливания задач, графические утилиты и др.

В состав ЦИВК также входят несколько специализированных машин для поддержки проектов ОИЯИ и международных коллабораций, в частности, таких как NICA/MPD, PANDA-GRID и CBM. На этих машинах устанавливается специа-

лизированное ПО для конкретных экспериментов и организуется запуск счетных задач через систему пакетной обработки. Эти же машины поддерживают NFS сервис для специализированного ПО на счетных машинах.

Список литературы

- [1] 2008 год: <http://mammoth.jinr.ru/~vvm/lit-new/raids.txt>,
<http://mammoth.jinr.ru/~vvm/lit-new/wns.txt>
- [2] 2009 год: <http://mammoth.jinr.ru/~vvm/HEP-SPEC06/result.txt>,
<http://mammoth.jinr.ru/~vvm/tests/dsk/prot-dsk1.txt>,
<http://mammoth.jinr.ru/~vvm/tests/dsk/prot-dsk2.txt>