

Поддержка и развитие МИВК ОИЯИ

Валерий Мицын

Составные части МИВК (не все)

- ЦИВК - сервисы пользователей.
- Тир1, Тир2 — грид.
- Batch (SLURM) — пакетная обработка задач.
- AFS — домашние директории пользователей и общее ПО.
- EOS — основное хранилище данных.
- CVMFS — распределённый доступ к специализированному ПО.
- СТА — архивное хранение данных.
- dCache, Enstore — данные для грид.

Сервисы ЦИВК

- **Обеспечение поддержки работы пользователей.**
Интерактивный вход (*ssh*), *gateway* на другие компоненты МИВК.
Аутентификация и авторизация: *kerberos*, *X509*, *OID* токен.
AFS — *HOME* директории, небольшие проекты ПО с распределённым доступом.
EOS — работа со своими и авторизованными данными.
Разработка и отладка ПО.
Запуск задач в *batch*.
Работа в *grid*.
Документация для пользователей.

Сервисы ЦИВК (продолжение)

- **Системные сервисы.**

DNS кэш, NTPD.

DHCP, PXE — автоматизация установки системы.

Puppet — автоматизация настройки систем и сервисов.

Squid серверы — кэширование http запросов.

NFS для системных сервисов и их конфигураций.

Серверы мониторинга и хранения истории работы сервисов.

Серверы хостинга виртуальных машин (KVM).

Сервисы Тир2, Тир1

- Интерфейсы доступа к batch и данным из грид.
 - CE — доступ к локальным batch системам, 2 сервиса.
 - SE — dCache (с Enstore) — доступ к данным, обеспечиваются все требуемые протоколы и методы авторизации в WLCG.
 - VOMS — сервис организации групп пользователей для координации и авторизации удалённого доступа в грид.
 - Сервис организации передачи данных в грид — FTS.
 - Сервисы инфраструктуры грид сайта: site/top bdii, apel (аккаунтинг) и другие.
 - Frontier squid — кэширование доступа к CVMFS.
 - Специальное ПО грид на интерактивных и счётных машинах.

SLURM — пакетная обработка задач

- Сервис доступа к счётным ресурсам.

Единый интерфейс запуска и выполнения счётных задач.

Две счётных фермы — два сервера SLURM.

Принимает описание задачи пользователя: число ядер, память, дисковое пространство.

Обновляет авторизацию этого пользователя, ищет затребованные ресурсы, отправляет запрос на счётные машины.

На счётной машине: резервирует ресурсы, запускает и отслеживает процесс выполнения и завершения задачи.

Сохраняет коды завершения и статистику потреблённых ресурсов с базе данных.

Обрабатывает запросы пользователей по состоянию и статистике выполнения задач.

- **Распределённая файловая система.**

Основная функция — доступ к домашним директориям пользователей (HOME).

4 сервера хранения, 1 сервер резервного копирования.

Авторизация доступа — kerberos.

Используется для небольших проектов с ПО, ≤ 2 GB.

Доступны на всех машинах ЦИВК/Тип2/Тип1 и по всему миру.

EOS

- **Распределённая система хранения больших объёмов данных.**
Основная функция — хранения и доступ к большим объёмам данных.
Основной внутренний протокол — xroot.
Есть возможность монтирования на машинах пользователей и счётных.
Для внешнего доступа — xroot, http[s]/webdav.
Авторизация доступа — kerberos, X509, OID token.
Доступны на всех машинах ЦИВК/Тип2/Тип1 и остальных частях МИВК.
Объём одиночного хранилища может достигать 100 PB, у нас ~22 PB в 2-х установках.

CVMFS

- **Распределённая файловая система, только чтение.**

Основная функция — распределённый доступ к проектам ПО общего пользования.

Протокол доступа — http.

Иерархия хранения и доступа: st-0 (основное хранилище) → st-1 (репликация st-0) → squid (кэширование) → машина с доступом к репозиторию (fuse).

Каждый st-0 обычно содержит несколько репозиторийов различного ПО, репозиторий содержит несколько версий ПО.

Доступны на всех машинах ЦИВК/Тир2/Тир1 и остальных частях МИВК.

Объём хранилища практически не ограничен, у нас ~30 ТВ.

СТА

- Система долговременного хранения на ленточных библиотеках.
Состоит из 2-х частей: буфер — EOS, серверы ленточных драйвов.
Доступ пользователей — только через запись/чтение на/из EOS, через обычные протоколы EOS.
Минимальное время на буфере: новые файлы сразу ставятся в очередь записи на ленты, считанные с лент будут удаляться с буфера после чтения их пользователем.
Внутренний протокол обмена - xroot.
Доступ пользователей к СТА ограничивается конфигурацией СТА.
Объём хранилища практически не ограничен, у нас ~100 PB разделяемого с Enstore.

dCache, Enstore

- Системы хранения данных грид на дисках и на ленточных библиотеках.

Основная функция — хранение и доступ к большим объёмам данных.

Протоколы доступа: dcap, xroot, webdav, gridftp.

Авторизация доступа — kerberos, X509, OID token.

Объём одиночного хранилища 1-100 PB, у нас ~20 PB в 3-х установках.

dCache работает в качестве буфера для Enstore.

DevLab — тестовые установки

Для тестирования новых версий и конфигураций рабочих сервисов созданы тестовые установки:

- SLURM и CE
- dCache и Enstore
- EOS и СТА.
- Новые версии ОС как правило тестируются на виртуальных машинах.