

Storage Element

Vanessa Hamar

Noviembre 21-25 del 2005

Mérida Venezuela

- **Todas las láminas aquí presentadas han sido tomadas del presentaciones y documentos del EGEE disponibles en Internet.**

- **Archivos y replicas**

- Los usuarios y los programas producen y requieren datos.
- La data puede ser almacenada en *Datasets* (Conjuntos de archivos)
 - *Se localizan en los Storage Element*
 - *Se puede tener una copia de un archivo en muchos sitios*
 - *Deben ser accesibles por los usuarios y las aplicaciones del Grid desde cualquier sitio.*
 - *Deben poder ser localizadas por el WMS (requerimientos de datos en los JDL)*
- *Otra manera*
 - *El Resource Broker puede enviar (pequeñas cantidades) datos para los trabajos Input Sandbox y Output Sandbox*
 - *La data puede ser copiada desde sistemas de archivos locales (WNs, Uis) hacia el grid*

- **Básicamente un disco con:**
 - Servidor GridFTP
 - Demonio RFIO
- **Capacidades que no posee**
 - Manejo de tiempo de gracia de archivos Lack of lifetime management (archivos permanentes y volátiles)
 - No permite realizar reservaciones de espacio o identificación de archivos (Pinning file).
- **Es muy limitado para el Grid**
 - Por lo que se introdujo el SRM

- **Cern Advanced STORage Manager**
- **Manejador de almacenamiento jerárquico**
 - Cintas + Discos Cache
 - Espacios permanentes y volátiles
- **La data es accesible por:**
 - GridFTP
 - **Plugin** RFIO – GSI
- **Implementa la interfaz SRM**

- **Maneja pool de discos**
 - La data puede ser distribuida entre una gran cantidad de servidores de discos.
 - Soporta copias internas y externas de un único archivo
 - Balanceo de carga automático utilizando métricas de costos y transferencias entre pools.
 - Maneja Archivos permanentes y volátiles.
- **Data accesible por**
 - GridFTP
 - Gsidcap (GSI-seguro)
- **Sigue siendo una solución pesada para:**
 - Sitios pequeños
 - Solo para soluciones pre-probadas

- **Nodos del dCache**
 - Un nodo admin + varios nodos pool
 - El nodo admin = centro del sistema + puntos de acceso
 - SRM, gridftp y servidores gridcap
 - Nodos pool: servidores de discos con espacios dedicados
 - Las pool pueden ser agrupados para facilitar su administración.
 - El nodo admin también puede ser un nodo pool.
- **Un único árbol de espacios jerárquicos**
 - Pnfs virtual del sistema de archivos (Utiliza NFS)
 - Soportado por una base de datos PostgreSQL en el nodo admin.
- **Terminología**
 - La funcionalidades son divididas en dominios (servicios)
 - Los dominios están subdivididos en celdas
 - Doors o puertas son puntos de entrada para acceder a la data (GridFTP door, gsidcap door....)

- **El Storage Resource Manager (SRM) es una especificación que define la manera en la cual un server que corre sobre un recurso de almacenamiento debe ser capaz de interactuar con aplicaciones alcanzables a través del Grid.**
- **Estas aplicaciones deben ser capaces de invocar un conjunto de métodos y esperar respuestas estándares y el rol de la interfaz SRM es asegurar que cualquier implementación de un sistema de manejo de almacenamiento puede ser capaz de utilizar esos métodos y respuestas.**

- **SRM ha sido el resultado de la colaboración del European DataGrid, CERN, Fermilab y LBNL.**
- **La implementación de los protocolos SRM se han hecho para numerosos sistemas de almacenamiento: HPSS, Enstore, JasMINE, CASTOR, EDG SE, ATLAS y arreglos RAID.**

- **Los Archivos SRM y los tipos de espacios de almacenamiento**
 - Se basan en el tiempo de vida de un sistema de almacenamiento:
 - **Permanente:** en un sistema de almacenamiento son de la misma naturaleza que los archivos permanentes en un sistema de archivos típico. Se garantiza que estos son almacenados y permanecen sin modificaciones a menos que los propietarios decidan borrarlos.
 - **Volátil:** son aquellos archivos que tienen un tiempo de vida específico, va a estar disponible para el usuario solo por ese tiempo específico. Si este expira el archivo puede ser removido tan pronto el sistema de almacenamiento requiera el espacio.
 - **Durable:** son archivos que usualmente son de naturaleza temporal y que contiene data importante. También tiene tiempo de vida pero cuando expira el sistema no lo remueve. La idea de esta información es que el sistema pueda accederlo rápidamente.

- **Existen muchos tipos de sistemas de almacenamiento, como cache de discos, sistemas de almacenamiento de cintas, arreglos de discos. Un “storage facility” puede contener mas de un tipo de sistemas de almacenamiento. En adición estos sistemas pueden manejar diferentes operaciones aplicar diferentes políticas.**
- **Las operaciones básicas que cualquier sistema manejado por una implementación SRM debe soportar se dividen en 5 categorías:**
 - Manejo de espacio de funciones
 - Manejo de espacio de directorios
 - Funciones de transferencia
 - Funciones de permisología
 - Funciones de status

- Estas funciones permiten reservar y actualizar el espacio en sistemas de almacenamiento, así como también encontrar información acerca del espacio y los archivos que contiene (espacio libre, tiempo de vida, etc.) y modificar estos parámetros.

- **Son las funciones que pueden ser necesarias para el manejo de directorios de sistemas de archivos de Unix. Por estas razones tienen los mismos nombres que los comandos de shell de Unix: srmMkdir, srmRmdir, srmMv, etc)**
- **Los directorios son un constructor virtual que provee esencialmente al usuario una manera lógica de agrupar archivos dentro del grid.**

- **Las funciones de transferencia son utilizados para transferir archivos de y para los sistemas de almacenamiento manejados por el SRM. Los nombres de los archivos utilizados por estas funciones se refieren a las replicas en el Grid son las SURLs y los TURLs.**

- **Esta son la funciones encargadas de proteger los archivos de accesos no autorizados. La permisología de los archivos son completamente análogos a las permisologías en los sistemas de archivos Unix.**

- **Son aquellas que pueden ser utilizadas para seguir el progreso de una operación SRM. La duración de subir o bajar desde o hacia un sistema de almacenamiento, comenzando en el momento de la petición del cliente, la información del estatus de la operación, etc.**

- **En la práctica debe garantizar la extensión del tiempo de vida de un archivo de naturaleza temporal, con el fin de incrementar el chance de que esta este disponible después que vence su tiempo de vida.**
- **Coloca una marca en el archivo temporal y va borrando el archivo con la marca mas antigua.**

- Existen diferencias entre el SE classic y la implementación del SE con SRM, la más importante es que la última soporta transferencia de datos asíncrona y la identificación de los archivos (pinning file)
- En el caso de la implementación de un SE Classic el cliente solo puede hacer la petición hasta que la previa haya sido terminada.



Instalación SE

- **Procedimiento para cada nodo:**
 - Instale el Sistema Operativo (SL3 or RH7.3)
 - Asegúrese que el NTP y el apt-get estén instalados y configurados.
 - Obtenga el certificado para cada host
 - Copie *hostcert.pem* y el *hostkey.pem* dentro del directorio */etc/grid-security*
 - Instale YAIM (como root)
 - `wget`** ;
 - `rpm -ivh lcg-yaim-...-noarch.rpm`**
 - Copie el archivo site config (previamente configurado)
 - Corra los scripts de instalación:
 - `> install_node mySite-info.def lcg-SECLASSIC`
 - `> configure_classic_SE mySite-info.def`
 - Posiblemente tendrá unos huecos de seguridad en el corta fuegos (GridFTP, SRM, RFIO...)
 - Las configuraciones locales es mejor hacerlas a mano.
 - Verifique que las CAs deben estar instaladas

- **Important attributes for the SE:**
 - Generales:
 - CE_HOST, SE_HOST, SITE_NAME
 - Cuentas y monitoreo
 - MON_HOST, REG_HOST, MYSQL_PASSWORD, GRIDICE_SERVER_HOST
 - Configuración del apt_get
 - LCG_REPOSITORY, CA_REPOSITORY
 - VOs y usuarios
 - USERS_CONF, VOS, VO_<VO>_SGM, VO_<VO>_USERS, VO_<VO>_STORAGE_DIR
 - Especificación del Close SE
 - CE_CLOSE_SE, CE_CLOSE_<LABEL>_HOST, CE_CLOSE_<LABEL>_ACCESS_POINT

- **Gridftp server**

- Información del Logging:
 - Archivos: `/var/log/globus-gridftp.log` y `/var/log/messages`
- Quotas de discos:
 - Normalmente todas las VOs están en la misma partición
 - ➔ *No se colocan quotas por VO*
 - ➔ *No se tiene información del espacio consumido por cada VO A VO should not "steal" the space assigned to other VOs*
- Mapeo de usuarios y autenticación
- Local Centre Authorization Service (LCAS):
 - *Chequea si un usuario esta autorizado o no utilizando la información existente en el `/etc/grid-security/grid-mapfile`*
 - Local Credential Mapping Service (LCMAPS):
 - *Mapea las credenciales del grid con las locales (ejemplo: UNIX uid/gid, AFS tokens, etc.)*
 - *Utiliza el `grid-mapfile`*
 - Configuración: `/opt/edg/var/etc/gridftpd_lcas_lcmaps.conf`
 - Logging: `/var/log/gridftp-lcas_lcmaps.log`

- **Demonio RFIO**
- **RFIO no utiliza el GSi-enabled, la seguridad está basada en uid/gid's**
 - Requiere que los uid/gid de los usuarios del Grid sean los mismos en los WNs and SEs
 - Es levantado usualmente al iniciarse la máquina
(*/etc/rc.d/rc3.d/S65rfiod*)
 - */usr/local/bin/rfiod -sl* (standalone con logging info) corre como root
 - Configuración
 - */etc/sysconfig/rfiod*
 - */etc/services* contiene los puertos RFIO (“rfio 5001/tcp”, “rfio 5001/udp”)
 - Logging information: un archivo nuevo diariamente
 - */var/log/rfiod*, */var/log/rfiod.1*, */var/log/rfiod.2.gz* ...
- **Persistencia de Archivos**
 - ➔ Los archivos no deben ser borrados aun así sean viejos...

- **Problemas de Autenticación**

- Los usuarios necesitan un proxy valido, registrado en una VO y haber utilizado el grid-proxy-init
- Deben tener una entrada correcta en el grid-mapfile (*/etc/grid-security/grid-mapfile*)
 - Actualiza cada 6 horas en un minuto aleatorio mediante el cronjob.
 - Esto requiere que el archivo *edg-mkgridmap.conf* sea correcto.
- Se debe tener una lista actualizada de cada CA de los certificados revocados
 - Estas son actualizadas utilizando el crontab y bajandose los archivos de un sitio web.
 - La CA debe actualizar los CRL antes que expiren los certificados.
- Los RPMs de las autoridades de certificaci{on deben ser actualizados
 - Todos los hosts necesitan todos los certificados de todas las CAs que aceptan
 - Coloque un trabajo en el crontab para actualizar las CAs

/opt/lcg/yaim/scripts/install_node <site-conf-file> lcg-CA

- **Depurando errores de autenticación**

- Cheque el archivo de logs `/var/log/messages` para obtener los errores de autenticación
- Chequee los siguiente
 - Chequee si el DN del usuario y la VO esten en el grid-mapfile.
 - Cheque la fecha del archivo CRL (`/etc/grid-security/certificates/xxxxxx.r0`)
 - Chequee que el RPM de las CAs este instalado
 - Puede obtener mas información tratando como usuario de:

```
globus-url-copy -dbg gsiftp://<hostname>/ file:/tmp/myFile
```



Preguntas