

PROPUESTA ARQUITECTÓNICA DE SERVICIOS FUNCIONALES A TRAVÉS DE UNA PLATAFORMA DE COMPUTACIÓN GRID PARA LA UNIVERSIDAD DISTRITAL

Carlos Enrique Montenegro Marín

Universidad Distrital Francisco José de Caldas, Bogotá, Colombia, cemontenegro@udistrital.edu.co

Paulo Alonso Gaona García

Universidad Distrital Francisco José de Caldas, Bogotá, Colombia, pagaonag@udistrital.edu.co

RESUMEN

Este documento describe inicialmente las tecnologías que se utilizan en la propuesta de la plataforma GRID para la Universidad Distrital, posteriormente se ahonda en la arquitectura WLCG/EGEE describiendo sus componentes de seguridad, interfaz de usuario, elementos de cómputo, elementos de almacenamiento, servicio de información, para finalmente describir como es el flujo de trabajo en una arquitectura GRID de este tiempo.

Palabras claves: GRID Computing, WLC/EGEE, Glite, Grid Architecture, Globus.

ABSTRACT

The next project describe the technologies that were originally used in the proposal from the GRID platform for the Distrital University, then depends into the architecture WLCG / EGEE describing the components security, user interface, computer components, elements of storage service information, to finally describe how is the workflow in a GRID architecture of this time.

Keywords: GRID Computing, WLC/EGEE, Glite, Arquitectura Grid, Globus.

1. INTRODUCCIÓN

El siguiente proyecto tiene como finalidad dotar de una arquitectura general que soporte diversas plataformas de computación Grid tales como Condor , Globus , ó Glite entre otras, pero a la vez busca conferir a la arquitectura de una plataforma robusta para su funcionamiento, planteando de esta manera una propuesta basada en un modelo computacional que implanta la plataforma de computación Grid Glite y que es entendible y operable sobre otras plataformas.

2. CARACTERIZACIÓN DE PLATAFORMAS GRID.

El proyecto EGEE (Switzerland, 2008) tiene como propósito principal proporcionar a los investigadores acceso a una infraestructura de computación Grid distribuida geográficamente, disponible las veinticuatro (24) horas del día. Este se enfoca en el mantenimiento del Middleware gLite y en el funcionamiento de una gran infraestructura de cómputo para el beneficio de una comunidad de investigadores vasta y diversa.

El Proyecto World wide LHC Computing Grid (WLCG) (Switzerland, 2008) se creó para preparar la infraestructura de cómputo sobre entornos de simulación, procesamiento y análisis de datos de los experimentos

del Large Hadron Collider (LHC). El LHC, el cual está siendo construido en el Laboratorio Europeo de Física de Partículas (CERN), sería el más grande y poderoso acelerador de partículas del mundo¹.

Los proyectos WLCG y EGEE comparten gran parte de su infraestructura y la operan conjuntamente. Por esta razón, se referirá a la misma como infraestructura WLCG/EGEE.

El Middleware gLite 3.0 surge de un sin número de proyectos Grid, como DataGrid (European Union Project, 2009), DataTag (Switzerland, 2008), Globus (University of Chicago, 2008), GriPhyN (Paul Avery, Ian Foster, 2008), iVDGL, EGEE y LCG. Actualmente, este Middleware está instalado en sitios que participan sobre la infraestructura WLCG/EGEE.

Sobre el WLCG existen otras infraestructuras Grid, específicamente la Open Science Grid (OSG) (Ruth Pordes, 2008), que utiliza el Middleware distribuido por VDT (National Science Foundation and the Department of Energy, 2008), y la NorduGrid (Farid Ould-Saada, 2008), que utiliza el ARC middleware, las cuáles para esta propuesta arquitectónica no se tendrán en cuenta para el manejo de servicios.

Los usuarios de una infraestructura Grid se dividen en Organizaciones Virtuales (VOs) según (Ian Foster, Carl Kesselman, Steven Tuecke, 2008), entidades abstractas que agrupan usuarios, instituciones y recursos en el mismo dominio administrativo (M. Dimou, 2008).

3. LA ARQUITECTURA DEL WLCG/EGEE:

En esta sección, se proporciona una perspectiva general breve de la arquitectura y los servicios del WLCG/EGEE.

SEGURIDAD

Como se explicó anteriormente, la comunidad de usuarios y desarrollo trabajada por el WLCG/EGEE está agrupada dentro de Organizaciones Virtuales (VO). Antes de que los recursos del WLCG/EGEE se puedan utilizar, un usuario debe leer y estar de acuerdo con las reglas de uso del WLCG/EGEE y otras reglas adicionales que la VO desee incluir, y registrar algunos datos personales en un Servicio de Registro.

Una vez que el registro de usuario se ha completado, el usuario puede acceder al WLCG/EGEE. La Grid Security Infrastructure (GSI) en el WLCG/EGEE permite autenticación y comunicación segura en una red abierta [University of Chicago, 2008]. GSI se basa en una clave pública encriptada, certificados X.509, y el protocolo de comunicación Secure Sockets Layer (SSL), con extensiones para single sign-on y delegación. A continuación se realiza a representación gráfica de la firma digital conocida como genérica dentro de cualquier tipo de especificación, mecanismo y /o algoritmo criptográfico utilizado para tal fin para tener una idea clara del proceso de firmado digital que se realiza dentro de esta arquitectura:

¹ Proyecto CERN <http://cdsweb.cern.ch/> [Última Consulta Enero 2009]

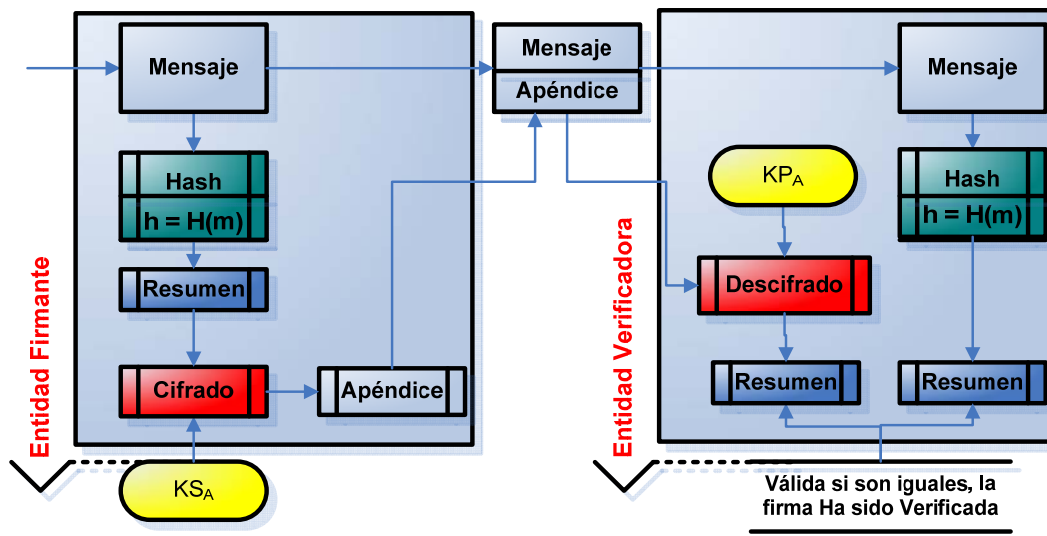


Figura No 1 Esquema de Firma Digital Genérico

La función Hash, es parte fundamental en la estructura de los algoritmos de firma digital al utilizar funciones unidireccionales en la autenticación de los mensajes, lo que garantiza que una vez ha sido cifrado el mensaje no se puede descifrar. Esta función garantiza la “huella dactilar” del documento, por lo tanto este tipo de funciones genera un gran valor agregado en el mundo de la informática y las telecomunicaciones. Pasos para firmar digitalmente un documento:

1. El **valor Hash** es conocido también como resumen de m (longitud del mensaje original), logrando obtener la siguiente función para su cálculo:

$$h = H(m)$$

Este proceso genera una firma que se trabaja mediante un **resumen** del Mensaje.

2. El **resumen** sufre un proceso de **Cifrado** junto con su clave privada KS_A el cual se envía junto con el mensaje original al receptor. Para cifrar se utiliza un algoritmo criptográfico (clave pública ó privada), el cual garantiza la autenticación del mensaje mediante la generación de una firma digital.
3. El proceso de verificación de la firma es **Descifrando** el valor del Apéndice enviado por emisor, utilizando para ello la clave pública y este proceso debe generar la **firma resumen** el cuál debe ser comparada con la firma original generada por el emisor para identificar validez de la misma.

AUTENTICACIÓN

Con el fin de autenticarse para utilizar los recursos Grid, un usuario necesita tener un certificado digital X.509 expedido por una Autoridad de Certificación (CA) reconocida por el WLCG/EGEE. Los recursos Grid, por lo general, se emiten también con certificados que los autentican ante usuarios y otros servicios.

El certificado de usuario, cuya clave privada esta protegida por un password, se utiliza para generar y firmar un certificado temporal, denominado certificado proxy (o simplemente proxy), que se usa para la autenticación real de los servicios Grid y no necesita un password. Puesto que el poseer un certificado proxy es una prueba de identidad, el archivo que lo contenga deberá ser legible sólo por el usuario. Un proxy tiene, de manera predeterminada, un periodo de duración corto (por lo general 12 horas) para reducir los riesgos de seguridad ante posibles plagios.

La autorización de un usuario en un recurso Grid específico se puede hacer de dos formas distintas. La primera es la más simple, y depende del mecanismo grid-mapfile. El recurso Grid tiene un grid-mapfile local que mapea

certificados de usuario a cuentas locales. Cuando la solicitud de un servicio por parte de un usuario llega a un anfitrión, el subject del certificado de usuario (contenido en el proxy) se compara con lo que aparece en el grid-mapfile local, para determinar a cuál cuenta local (si existe alguna) se mapea el certificado de usuario, y esta cuenta es luego utilizada para llevar a cabo la operación solicitada [University of Chicago, 2008]. La segunda forma se basa en el Servicio de Membrecía de la organización Virtual (VOMS) y el mecanismo LCAS/LCMAPS, que cuenta con una definición más detallada de los privilegios del usuario.

Un usuario necesita un proxy válido para enviar trabajos. Esos trabajos conservan sus propias copias del proxy para estar en capacidad de autenticarse con los servicios Grid mientras se ejecutan. Para trabajos de larga ejecución, el proxy del trabajo puede expirar antes de que el trabajo haya finalizado, causando la interrupción del mismo. Para evitar que esto suceda, existe un mecanismo de renovación vía proxy que mantiene el proxy del trabajo válido tanto tiempo como sea necesario. El servidor MyProxy es el componente que proporciona esta funcionalidad.

INTERFAZ DE USUARIO

El punto de acceso al WLCG/EGEE Grid es la Interfaz de Usuario (UI). Esta puede ser alguna máquina en la cual los usuarios tengan una cuenta personal y en la cual sus certificados de usuario estén instalados. Desde una UI, un usuario puede recibir autenticación y autorización para usar los recursos del WLCG/EGEE, y puede tener acceso a las funcionalidades ofrecidas por los sistemas de Información, de gestión de Carga de trabajo y Datos. Esta proporciona herramientas de la CLI para llevar a cabo algunas operaciones básicas de Grid:

Hacer una lista de todos los recursos apropiados para ejecutar un trabajo determinado;
Enviar trabajos para su ejecución;
Cancelar trabajos;
Recuperar la salida de trabajos finalizados;
Mostrar el estado de trabajos enviados;
Recuperar la información de registro y bookkeeping de trabajos;
Copiar, replicar y eliminar archivos desde el Grid;
Recuperar el estado de diferentes recursos desde el Sistema de Información.

Además, las APIs del WLCG/EGEE también están disponibles en la UI para permitir el desarrollo de aplicaciones Grid-enabled.

ELEMENTO DE CÓMPUTO

Un Elemento de Cómputo (CE), en terminología Grid, es una serie de recursos de cómputo localizada en un sitio (es decir, un cluster, una granja de cómputo). Un CE incluye un Grid Gate (GG) (Para los CEs basados en Globus, se denomina Gatekeeper.), que actúa como una interfaz genérica para el cluster; un Local Resource Management System (LRMS) (llamado algunas veces sistema por lotes), y el cluster en sí mismo, una colección de Nodos de Trabajo(WNs), los nodos donde se ejecutan los trabajos.

El GG es responsable de aceptar trabajos y enviarlos para su ejecución en los WNs a través del LRMS.

Los WNs generalmente tienen los mismos comandos y librerías que la UI, además de los comandos de gestión de trabajo. El software de aplicación específico de una VO se puede preinstalar en los sitios en un área dedicada, generalmente en un sistema de archivo compartido que sea accesible desde todos los WNs.

ELEMENTO DE ALMACENAMIENTO

Un Elemento de Almacenamiento (SE) proporciona acceso uniforme a los recursos de almacenamiento. El Elemento de Almacenamiento puede controlar servidores de disco simples, grandes arreglos de disco o Sistemas de Almacenamiento Masivo basado en cinta (MSS). La mayoría de los sitios del WLCG/EGEE proporcionan al menos un SE.

Los Elementos de Almacenamiento pueden soportar diferentes protocolos de acceso de datos e interfaces, En palabras sencillas, GSIFTP (un GSI-secure FTP) es el protocolo utilizado para la transferencia completa de archivos, mientras que el acceso local y remoto a archivos se lleva a cabo usando el RFIO o elgsidcap.

Algunos recursos de almacenamiento son administrados por un Administrador de Recursos de Almacenamiento (SRM) [University of Chicago, 2008], un módulo de middleware que proporciona capacidades tales como migración transparente de archivos de disco a cinta, bloqueo de archivo, reserva de espacio, etc. Sin embargo, diferentes SEs pueden soportar distintas versiones del protocolo SRM y las capacidades pueden variar.

SERVICIO DE INFORMACIÓN

El Servicio de Información (IS) proporciona información sobre los recursos Grid del WLCG/EGEE y su estado. Esta información es esencial para el funcionamiento de todo el Grid, puesto que es a través del IS que se descubren los recursos. La información publicada se utiliza también, con propósitos de monitoreo y contabilidad.

La mayoría de la información publicada para el IS se ajusta al GLUE Schema, el cual define un modelo de datos conceptuales común que se usa en el monitoreo y descubrimiento de recursos Grid.

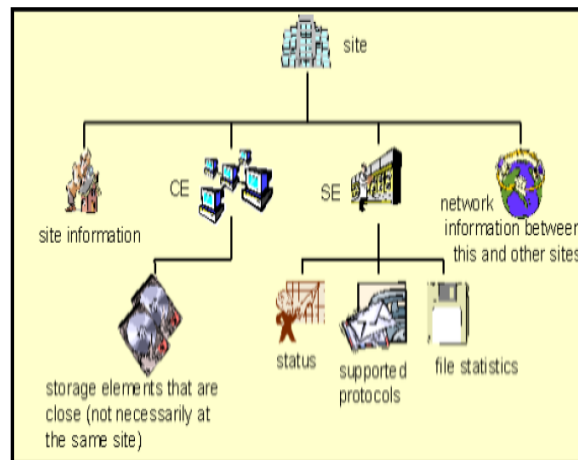


Figura No 2 Árbol de Información de Directorio. Fuente (Burke Stephen)

GESTIÓN DE CARGA DE TRABAJO

El propósito del Sistema de Gestión de Carga de Trabajo (WMS) es aceptar trabajos de usuario, asignarlos al Elemento de Cómputo más apropiado, registrar su estado y recuperar su salida. El Administrador de Recursos (RB) es la máquina donde se ejecutan los servicios del WMS [Campana, 2008] [Pacini, 2008].

4. CARGA Y PROCESAMIENTO TRANSACCIONAL SOBRE UNA GRID

Aquí se describirá brevemente lo que ocurre cuando un usuario envía un trabajo al WLCG/EGEE Grid para procesar algunos datos, y se explica cómo interactúan los diferentes componentes.

ENVÍO DE TRABAJO

La Figura 3 ilustra el proceso que se lleva a cabo cuando se envía un trabajo al Grid. Esta hace referencia al WMS de LCG-2, sin embargo, el WMS de gLite es similar. Los pasos son los siguientes:

1. Después de obtener un certificado digital de una Autoridad de Certificación confiable, registrarse en una VO y obtener una cuenta en una Interfaz de Usuario, el usuario está listo para usar el WLCG/EGEE Grid. El usuario ingresa a la UI y crea un certificado proxy para autenticarse en subsiguientes interacciones seguras.
2. El usuario envía un trabajo desde la UI a un Administrador de Recursos. En el archivo de descripción de trabajo, se puede especificar uno o más archivos que se vayan a copiar desde la UI al WN. Esta serie de archivos se denomina Input Sandbox. Se registra un evento en el LB (Servicio de Registro y de Bookkeeping) y el estado del trabajo es SUBMITTED.
3. El WMS busca el mejor CE disponible para ejecutar el trabajo. Con este fin, este interroga el BDII para consultar el estado de los recursos cómputo y de almacenamiento, y el Catálogo de Archivos para encontrar la ubicación de todos aquellos archivos de entrada requeridos. Se registra otro evento en el LB y el estado del trabajo es WAITING.

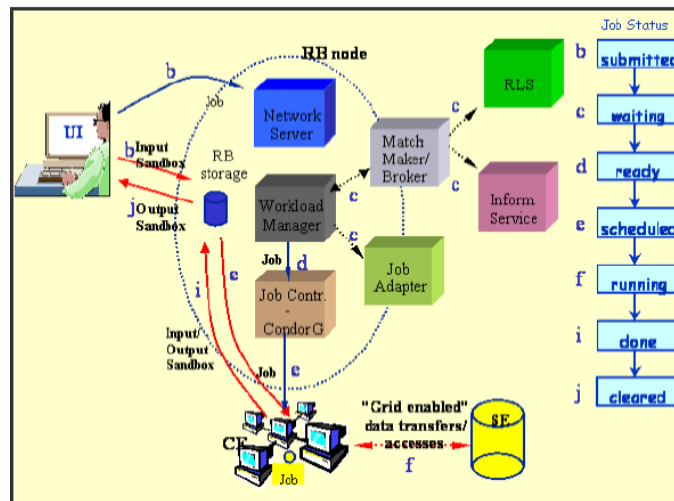


Figura No 3 Flujo de Trabajo en el WLCG/EGEE Grid (Burke, 2008).

4. El RB prepara el trabajo para su envío, creando un wrapper script que pasará, junto con otros parámetros, al CE escogido. Se registra un evento en el LB y el estado del trabajo es READY.
5. El CE recibe la solicitud y envía el trabajo al LRMS local para su ejecución. Se registra un evento en el LB y el estado del trabajo es SCHEDULED.
6. El LRMS se encarga de la ejecución del trabajo en los nodos de trabajo de la granja local disponible. Se copian los archivos de usuarios del RB al WN donde se ejecuta el trabajo. Se registra un evento en el LB y el estado del trabajo es RUNNING.

7. Mientras el trabajo se ejecuta, se puede acceder directamente a los archivos Grid desde un SE, usando los protocolos secure RFIO o gsidcap, o después de copiarlos al sistema de archivo local en los WN con las herramientas del Manejo de Datos.
8. El trabajo puede producir nuevos archivos de salida que se pueden transferir al Grid y poner a disposición de otros usuarios Grid.
9. Si el trabajo finaliza sin errores, la salida (no de archivos de datos extensos, sino sólo de archivos de salida pequeños que el usuario haya especificado en la denominada Output Sandbox) se transfiere nuevamente al nodo del RB. Se registra un evento en el LB y el estado del trabajo es DONE.
10. En este momento, el usuario puede recuperar la salida de su trabajo de la UI. Se registra un evento en el LB y el estado del trabajo es CLEARED.
11. Las consultas sobre el estado del trabajo se pueden dirigir al LB desde la UI. Asimismo, es posible consultar el BDII desde al UI, para conocer el estado de los recursos.
12. Si el sitio al cual se envía el trabajo no puede aceptarlo o ejecutarlo, el trabajo puede reenviarse automáticamente a otro CE que satisfaga los requisitos del usuario. Después de alcanzar el máximo número permitido de reenvíos, el trabajo será señalado como cancelado. Los usuarios pueden obtener información sobre la historia de un trabajo consultando el servicio de LB.

5. CONCLUSIONES

Mediante la representación definida por (Burke, 2008) se busca determinar un modelo adecuado para cumplir adecuadamente con un proceso de comunicaciones mediante la generación de certificados digitales extremo a extremo que garantice la confiabilidad de las transacciones realizadas y el éxito de los procesos generados dentro de la comunicación.

Para enviar trabajos que necesiten procesamiento en paralelo a la Grid es necesario realizar la programación de estos antes de ser enviados, la Grid por si sola no paraleliza los trabajos, esta es una tarea del usuario que necesita de la Grid. Una Grid puede estar enfocada a diferentes servicios por ejemplo Almacenamiento o procesamiento.

G-Lite es un Middleware para una plataforma Grid y no es el único existen otros como CONDOR o Globus, pero G-Lite ofrece a través de GILDA una plataforma académica excelente para el aprendizaje de este nuevo paradigma.

Un nodo en una Grid puede ofrecer varios servicios pero es claro que hay combinaciones de servicios que no se pueden realizar en el mismo nodo como por ejemplo tener en el mismo nodo un CE con un BDII. Esto no se puede realizar.

Casi todos los nodos de una Grid y los servicios que estos prestan están basados en la misma infraestructura (SL3, Java SDK, Ntp daemon, CA_Gilda rpm, edg_VO_Gilda rpm, glite-yaim-3.0.11, gilda_ig-yaim-3.0.0-37), sobre esta si instalan y configuran los servicios necesarios.

REFERENCIAS

Burke Stephen, Campana Simone, Delgado Peris Antonio, (2008) Glite 3.0 User Guide Manuals Series, <https://edms.cern.ch/file/722398/1/gLite-3-UserGuide.pdf> [Enero 2008]

Campana, M. Litmaath, A. Sciab`a, LCG-2 Middleware overview <https://edms.cern.ch/document/498079/> [Noviembre 2008]

European Union Project, (2008). The DataGrid Project <http://www.edg.org/> [Agosto 2008]

Paul Avery, Ian Foster, (2008). Project GriPhyN – Grid Physics Network <http://www.griphyn.org/> [Julio 2008]

Farid Ould-Saada, (2008). NorduGrid <http://www.nordugrid.org/> [Enero 2009]

- Ian Foster, Carl Kesselman, Steven Tuecke, (2008). The Anatomy of the Grid: Enabling Scalable Virtual Organizations <http://www.globus.org/alliance/publications/papers/anatomy.pdf> [Octubre 2008]
- iVDgL, (2008)– International Virtual Data Grid Laboratory <http://www.ivdgl.org/> [Julio 2008]
- M. Dimou, LCG User Registration and VO Management, (2008) <https://edms.cern.ch/document/428034/> [Octubre 2008]
- National Science Foundation and the Department of Energy, (2008). The Virtual Data Toolkit <http://vdt.cs.wisc.edu/> [Octubre 2008]
- Pacini, WMS User's Guide <https://edms.cern.ch/document/572489/> [Noviembre 2008]
- Ruth Pordes, (2008). Open Science Grid <http://www.opensciencegrid.org/> [Octubre 2008]
- Switzerland, (2008). DataTAG – Research & technological development for a Data TransAtlantic Grid <http://cern.ch/datatag/> [Octubre 2008]
- Switzerland, (2008). European Commission EGEE – Enabling Grids for E-science. <http://eu-egee.org/> [October 2008]
- Switzerland, (2008), Worldwide LHC Computing Grid <http://cern.ch/LCG/CERN-LCG-GDEIS-722398> Manuals Series, Página 8. [Octubre 2008]
- University of Chicago, (2008) The Globus Alliance <http://www.globus.org/> [Noviembre 2008]
- University of Chicago (2008). MDS 2.2 Features in the Globus Toolkit 2.2 Release http://www.globus.org/toolkit/mds/#mds_gt2 [Noviembre 2008]

Autorización y Renuncia

Los autores autorizan a LACCEI para publicar el escrito en los procedimientos de la conferencia. LACCEI o los editores no son responsables ni por el contenido ni por las implicaciones de lo que esta expresado en el escrito

Authorization and Disclaimer

Authors authorize LACCEI to publish the paper in the conference proceedings. Neither LACCEI nor the editors are responsible either for the content or for the implications of what is expressed in the paper.