

# Capítulo 1

## Introducción

### 1.1. Motivación

Los grids computacionales permiten el uso de equipos de cómputo que en su conjunto proporcionan un gran poder de cálculo y de almacenamiento. Estos equipos están distribuidos en diferentes lugares y comúnmente están conectados por redes con un gran ancho de banda que permite mantener ocupados a los equipos del grid computacional.

Debido a las características de los grids computacionales y a su potencial, el cómputo en grid permite estudiar problemas complejos y con mayor precisión, utilizando una infraestructura que se expande a través de límites administrativos y geográficos. Estos problemas serían imposibles de estudiar utilizando un sólo equipo de cómputo por muy robusto que fuera. Este hecho queda de manifiesto en los experimentos llevados a cabo en los detectores del LHC (ver sección 2.2.1) ya que no existe actualmente un equipo con el poder de cálculo suficiente para almacenar y analizar todos los datos generados por dichos experimentos.

Una de las motivaciones de este trabajo es proporcionar una interfaz amigable, accesible y disponible a los usuarios de un grid computacional. Disponible en casi cualquier lugar por ser una interfaz web, amigable en el sentido de que no requiere mucho tiempo de aprendizaje y que es fácil de utilizar.

El uso de interfaces Grids amigables permite que los usuarios, no expertos en cómputo, dediquen sólo atención a la ejecución de sus aplicaciones y a la administración de su información. Es decir, las interfaces Grid hacen posible que los usuarios se concentren sólo en problemas científicos y no en cómo funciona un grid computacional.

Las interfaces grid benefician y fortalecen a los *science gateway*, grupos de expertos o de investigación en una determinada área del conocimiento que trabajan en conjunto aunque se encuentren geográficamente dispersos. Entre las áreas del conocimiento de los *science gateway* se encuentran: astronomía, química, bioquímica, ciencias de la tierra, sismología, genética, biomedicina, geografía, geofísica, investigación atmosférica, investigación de materiales, neuronciencia, física de altas energías, física teórica, física de astro partículas, visualización y procesamiento de imágenes.

Diversos grupos en México están interesados en grids computacionales. En particular la Delta Metropolitana es un proyecto que conjuntará a las tres supercomputadoras más rápidas del país, ubicadas en la UNAM, UAM y el CINVESTAV. La delta, que puede ser administrada como un grid computacional, puede tener como una interfaz de acceso a los equipos de supercómputo un portal grid.

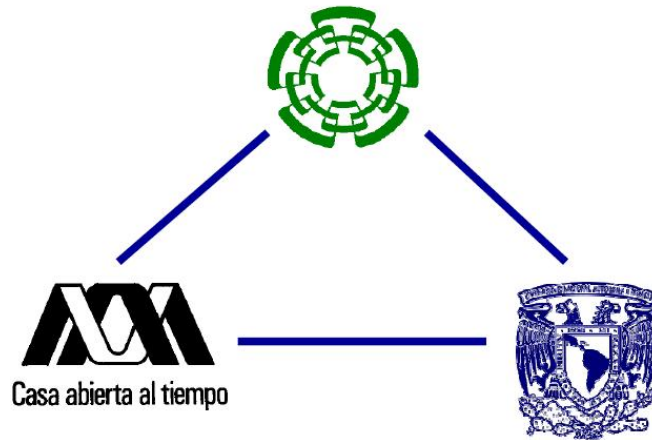


Figura 1.1: Fuente: Redes Ópticas Metropolitanas Académicas[1]

## 1.2. Descripción del problema

Los servicios proporcionados en un grid computacional son comúnmente mostrados a los usuarios en una interfaz de línea de comandos (CLI) o sesiones remotas de “shell”, donde cada comando tiene diferentes opciones a utilizar. Las CLI proporcionan una interfaz con mayor control y más herramientas, en un modo de uso más cercano con los recursos de cómputo, pero debido a su complejidad, al tiempo necesario para aprender a utilizarlas y a que los usuarios (físicos, químicos, ingenierías, etc) de los grids

computacionales no son expertos en cómputo, sino en las aplicaciones que utilizan, se buscaron mecanismos que proveerán un ambiente más amigable a los usuarios.

De esa necesidad, surgieron los portales Grid: interfaces web para uso de las grids computacionales y portales de acceso a recursos de supercómputo. Este tipo de interfaz se caracteriza por proporcionar un acceso a recursos que resulta más amigable que las CLI, ya que desde un mismo punto, a través de interfaces gráficas y de forma segura, permite consultar información relacionada con equipos de cómputo, envío y monitoreo de trabajos, así como manejo de archivos. Esta interfaz permite que los usuarios de los grids computacionales se dediquen y concentren más en el uso de sus aplicaciones y problemas de su área de conocimiento y no inviertan mucho tiempo conociendo la infraestructura (grid computacional), sobre la que yace su aplicación.

Existe la posibilidad de desarrollar una interfaz o portal grid propia para el uso de grids computacionales. Sin embargo, esto requeriría mucho tiempo de desarrollo e implicaría conocer con detalle todas las diferentes tecnologías involucradas para construir este tipo de interfaz. Además, se volvería a hacer un trabajo que grupos de investigación (por ejemplo, OGCE y GENIUS [2]) ya han desarrollado en diferentes plataformas; adicionalmente, si por alguna razón existiera un fallo en la seguridad del portal, se comprometería a los equipos de cómputo inmersos en el grid computacional.

Otra alternativa es implementar una interfaz o portal grid ya desarrollado, de preferencia de código abierto. La ventaja de esto es que su tiempo de implementación es menor comparado con desarrollar una interfaz propia, porque no se tienen que conocer todas las tecnologías involucradas a profundidad, además de que varias instituciones y centros de cómputo (Universidad de Indiana, el Centro de cómputo de Renaissance, el Instituto de Tecnología de Rochester, la Universidad de San Diego, el Centro de Supercómputo de San Diego y el Centro de Cómputo Avanzado de Texas) han invertido en su desarrollo, han probado la seguridad y confiabilidad de este tipo de interfaz.

### **1.3. Resumen del trabajo**

En este trabajo se presentarán los conceptos básicos de un grid computacional. Asimismo se mostrarán los componentes principales del software más utilizado para implementar grids computacionales (Globus Toolkit). En este mismo contexto, se describirán algunos de los principales proyectos de grids computacionales en la actualidad.

En el marco de los portales grids, se profundizará en el conocimiento de este tipo de interfaces web, a través del estudio de su funcionamiento y las tecnologías involucra-

das en su implementación. Asimismo, se estudiarán los componentes principales que constituyen a un portal grid para su uso e integración en un grid computacional.

De forma práctica, se implementará un portal grid para equipos de supercómputo, se describirá la infraestructura necesaria, en hardware y software, para la implementación del portal grid; se explicará el proceso de instalación y configuración de los componentes de un grid computacional y de un portal grid; asimismo, se harán las pruebas para mostrar el correcto funcionamiento de un portal grid.

La organización del trabajo es como sigue:

- El capítulo 2 trata sobre los grids computacionales y las tecnologías inmersas en ellas.
- El capítulo 3 versa sobre los portales grids.
- El capítulo 4 presenta la implementación y pruebas de funcionamiento de un Portal Grid.
- El capítulo 5 muestra los resultados de este trabajo.

Finalmente, se concluirá sobre los resultados esperados de este trabajo. Los resultados que se esperan de este trabajo son:

- Descripción de la arquitectura de los portales Grid (portales de acceso a equipos de supercómputo)
- Implementación de un Portal Grid para equipos de supercómputo de la Universidad Nacional Autónoma de México.