

# Introducción a la Tecnología Grid.

*Jose I. Abreu Salas, Pedro A. Sánchez Mirabal, Patrick García Pérez. y Rainer León Morera.*

*Facultad de Informática. Universidad de Matanzas "Camilo Cienfuegos".*

2006

## Resumen

El término Grid se refiere a una infraestructura que posibilita el empleo de forma integrada y cooperativa de ordenadores, redes, bases de datos e instrumentos científicos que pertenecen y son administrados por múltiples organizaciones; de manera que el sistema permita a los usuarios un acceso uniforme a recursos distribuidos en un entorno heterogéneo. De esta forma se hace posible manejar problemas que presentan un alto costo computacional o de memoria - como la modelación de la estructura espacial de proteínas o el procesamiento de datos recolectados por radio telescopios – empleando un grupo de ordenadores conectados en red de forma que se integren en una arquitectura virtual capaz de distribuir tareas a través de una plataforma paralela. Los sistemas de Computación Grid permiten realizar tareas de cálculo que involucren grandes cantidades de datos fragmentado estos en conjuntos más pequeños; además brindan la capacidad para ejecutar muchas más instrucciones por unidad de tiempo que los ordenadores corrientes – incluso que muchos supercomputadores – empleando un modelo de procesos que pueden ser llevados en paralelo.

## 1. Antecedentes

Desde el surgimiento de la primera computadora se hizo evidente que el poder de cálculo de esta no era suficiente para dar respuesta en un tiempo razonable a múltiples problemas que por su naturaleza presentan un elevado costo computacional o que deben manejar enormes volúmenes de datos. El enfoque tradicional para lidiar con estos inconvenientes ha seguido un modelo centralizado basado en costosas supercomputadoras. Sin embargo durante la última década se ha podido apreciar un notable incremento en el rendimiento de las redes de ordenadores como resultado del desarrollo de hardware cada vez más rápido y software más eficiente, lo cual ha permitido que se empezaran a difundir otras alternativas distribuidas capaces de obtener - para determinados perfiles de aplicación - rendimientos comparables a los ofrecidos por los modelos centralizados más avanzados pero a un costo mucho menor, dando lugar a lo que se conoce como **computación distribuida (distributed computing)**

## 1.1 Computación distribuida.

Consiste en un modelo de computación en paralelo donde típicamente intervienen dos o más computadoras que se comunican a través de un red para cumplir una tarea u objetivo común. El tipo de hardware, lenguajes de programación y sistemas operativos pueden ser muy variados. Existen diferentes modelos que se agrupan bajo el término de computación distribuida, entre los que se destacan **Intranet** e **Internet computing** y los **Clusters de Computadoras (computer clusters)** aunque normalmente los clusters no son tan flexibles en cuanto a la heterogeneidad de hardware y sistemas operativos sobre los cuales se implementan.

## 1.2 Intranet e Internet computing.

Tanto Intranet como Internet computing son modelos que se construyen de manera que permitan gestionar un sistema heterogéneo de equipos que no tienen por que ser dedicados – es decir no tienen que estar disponibles en todo momento - para tratar de aprovechar todo el tiempo de procesador que normalmente es desperdiciado. Como su nombre indica, Intranet computing persigue estos fines dentro del marco de un único dominio administrativo, siendo esta la diferencia fundamental con los sistemas de computación sobre Internet, los cuales incorporan recursos físicos y virtuales que pueden estar geográficamente dispersos y que probablemente pertenezcan a dominios de administración diferentes; lo cual complica entre otros problemas los relacionados con las políticas de seguridad. Un ejemplo muy

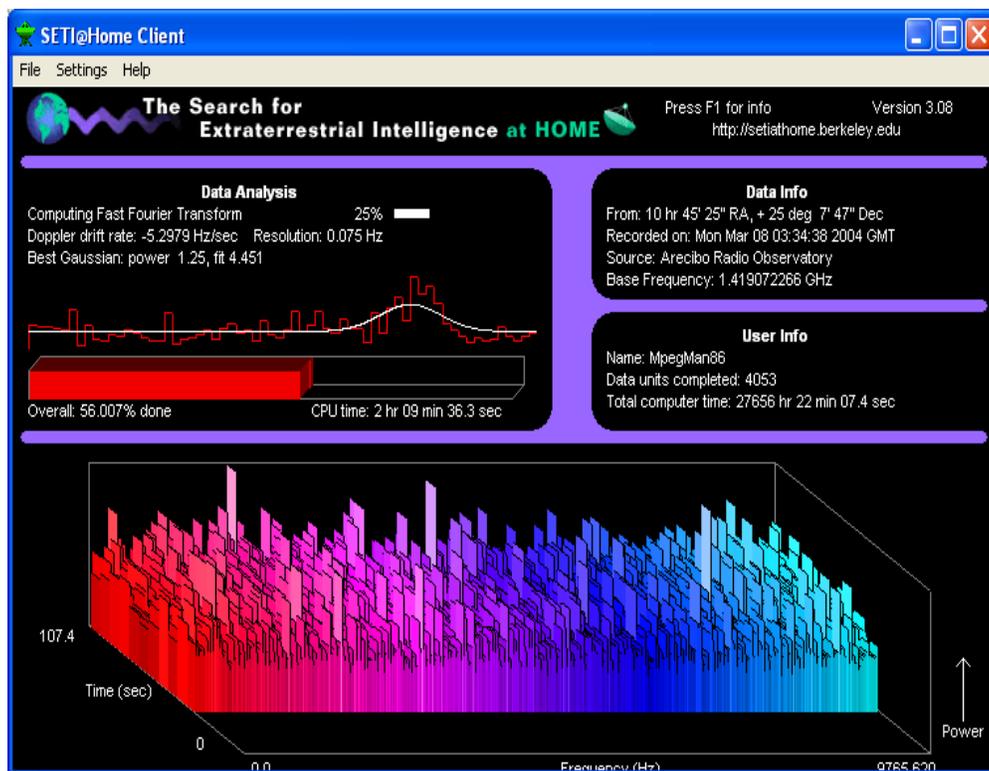


Fig. 1 SETI@home: Gráfico de la transformada de Fourier.  
Fuente: Sitio web [www.wikipedia.org](http://www.wikipedia.org)

conocido de esta infraestructura es el proyecto [SETI@home](#), el cual hace uso del tiempo de ocioso de los ordenadores conectados para analizar señales de radio en busca de vida extraterrestre.

### 1.3 Clusters de computadoras.

Un cluster es un grupo de ordenadores comúnmente dedicados y conectados a una red local de banda ancha que trabajan de manera conjunta, cuyos recursos son controlados de forma centralizada presentándose como un único sistema.



*Fig. 2 Cluster operado sobre Linux, Universidad de Purdue. Fuente: [www.wikipedia.org](http://www.wikipedia.org)*

## 2. Tecnología Grid. Conceptos.

En los últimos tiempos el término Grid ha dejado los laboratorios para hacerse muy popular, siendo frecuente encontrar trabajos sobre Data Grids, Knowledge Grids, Campus Grids etc. Naturalmente todo esto conduce a preguntarse, ¿que se entiende por tecnologías Grid?. Para dar respuesta a esta interrogante puede hacerse un análisis de una de las técnicas Grid más de moda, la **computación Grid (Grid Computing)**<sup>1</sup>, de manera que puedan identificarse las características fundamentales que distinguen a un sistema como Grid.

### 2.1 Computación Grid.

Al igual que para todos los miembros de la familia Grid, pueden encontrarse diferentes definiciones para la computación Grid, término surgido en los 90 para designar de forma metafórica la idea de hacer que el acceso a los recursos de cómputo fuese de forma tan sencilla como emplear la red de suministro eléctrico donde los usuarios no necesitan conocer donde ni como se produce la electricidad que consumen, solo tener la capacidad para hacer uso de ella tanto como sea necesaria. He aquí algunas de ellas.

- Un servicio para compartir capacidad de cómputo u almacenamiento a través de Internet. [CERN, 2005]
- La computación Grid, sin más complicaciones, es el modelo de computación distribuida llevado a un nivel superior de desarrollo. El objetivo central es crear la ilusión de una simple pero poderosa computadora virtual – cuyo mecanismo de manejo debe ser interno - a través de una colección heterogénea de equipos que comparten diferentes recursos. [Berstis, 2002].

---

<sup>1</sup> Algunos autores emplean indistintamente los términos **Grid Computing** y **Tecnología Grid** como sinónimos. Aquí se hace distinción para indicar que las potencialidades ofrecidas por el modelo Grid pueden ser extendidas a otros entornos no virtuales. De cualquier manera las definiciones empleadas entienden los términos como equivalentes.

- Un tipo de sistema paralelo y distribuido que permite compartir, seleccionar e integrar recursos autónomos geográficamente dispersos, todo en tiempo real dependiendo de su disponibilidad, capacidad, desempeño, costo y requerimientos sobre la calidad de los servicios que establecen los usuarios. [Buyya et al., 2005]

Estas definiciones encierran varios puntos que identifican aspectos fundamentales que permiten conceptuar con mayor rigor qué es la computación Grid.

i. Existen diferentes recursos compartidos:

los usuarios son capaces de compartir y tener acceso a recursos disponibles como tiempo de procesador y capacidad de almacenamiento.

ii. Los equipos integrados en el Grid son de naturaleza heterogénea:

no tiene por que existir uniformidad de hardware o software de los elementos que se integran en el sistema, siendo posible encontrar ordenadores de diferentes arquitecturas corriendo diferentes sistemas operativos como Linux, Solaris o Windows.

iii. Los recursos compartidos se encuentran geográficamente dispersos:

los elementos que conforman el sistema no tienen por que estar localizados en un área geográficamente próxima, ni tienen por que pertenecer a una misma entidad o dominio administrativo.

iv. Los equipos pertenecientes al Grid son autónomos:

cada nodo del sistema posee sus propias normas de seguridad, acceso, y control de la disponibilidad de los recursos compartidos. Todo esto debe ser transparente a los usuarios, por lo que el sistema debe funcionar como una computadora virtual que internamente sea capaz de integrar las particularidades de cada nodo y presentarse al usuario como un único sistema.

Cada uno de estos aspectos muestra claramente qué se desea lograr con la tecnología Grid - además de hacer evidentes ciertas diferencias respecto a otros modelos de computación distribuida como Intranet computing o los clusters de ordenadores – pero aún no proveen de una guía concisa y rigurosa que permita clasificar un sistema como Grid. En su artículo *What is The Grid*, Ian Foster abunda sobre la naturaleza de los diferentes procesos que se desarrollan dentro del Grid.

Cuando hablamos de compartir no se trata de un simple intercambio de ficheros, más bien se trata de ofrecer acceso a computadoras, software, datos y otros recursos de forma directa, como requieren los modelos para la solución cooperativa de problemas y de negociación de recursos que aparecen en la industria, la ciencia y la ingeniería. Esta forma de hacer disponibles los recursos debe ser por necesidad rigurosamente controlada, de forma que los proveedores y los consumidores queden claramente identificados, además de establecer cuidadosamente qué recursos se comparten, quién está autorizado a compartir y las condiciones bajo las cuales se realiza el proceso... [Foster, 2002]

Las definiciones mostradas ofrecen una visión bastante clara sobre las características esenciales de un sistema Grid; particularmente, los planteamientos de Foster permiten vislumbrar determinadas líneas que persiguen aclarar no solo qué funcionalidades deben brindarse, sino también presentar ciertos papeles y reglas que deben quedar perfectamente definidos dentro de un sistema Grid. Todas las ideas anteriores, pueden sintetizarse en tres puntos o guías [Foster, 2002] de acuerdo con las cuales un sistema es Grid si:

- i. coordina recursos que no están sujetos a un control centralizado...

el sistema integra y coordina recursos y usuarios localizados en dominios distintos; diferentes dependencias administrativas pertenecientes a una o varias compañías; y dirige las políticas de seguridad, membresía, pagos etc. De otra forma se está en presencia de un sistema de administración local.

- ii. ... usando protocolos e interfaces estándares, abiertas y de propósito general

el sistema Grid está basado en protocolos e interfaces multipropósitos que definen cuestiones fundamentales como la autenticación, autorización, localización de recursos y acceso a los recursos. Es importante que dichos protocolos e interfaces sean estándares y abiertos, de lo contrario lo que se tiene es un sistema de aplicación específica.

- iii. ... para proveer servicios de una alta calidad y prestación

el sistema Grid permite que los recursos en él integrados sean usados de forma coordinada para proveer varias calidades de servicios, relacionados por ejemplo con el tiempo de respuesta, rendimiento, disponibilidad y seguridad; o la asignación de múltiples recursos para satisfacer demandas complejas de los usuarios de manera que la utilidad del sistema es significativamente mayor que la suma de sus partes.

Solo resta - a modo de ejemplo - ver como pueden aplicarse dichas reglas a un sistema, preferiblemente real, para discernir si es este Grid o no. Puede tomarse para esto la

aplicación del conocido proyecto [Seti@home2](#). SETI aúna los esfuerzos de gran número de personas en el mundo que donan el tiempo ocioso de CPU para emplearlo en el análisis de datos colectados a través del radio telescopio de Arecibo Puerto Rico, en busca de vida extraterrestre. Para esto los voluntarios deben instalar un protector de pantallas (**screensaver**) que contiene el código necesario para procesar las señales de radio y el código para manejar las solicitudes y el envío de datos o resultados. Cuando el procesador está disponible automáticamente el protector de pantalla inicia el trabajo.

A simple vista - sobre todo teniendo en cuenta las definiciones menos rigurosas - SETI clasifica perfectamente como sistema Grid:

- Dispersión geográfica y heterogeneidad: la red de voluntarios puede distribuirse en todo el planeta
- Autonomía: Los recursos compartidos pertenecen a dominios de administración diferentes teniendo sus propias políticas de funcionamiento.

Es necesario analizar entonces la respuesta a las siguientes preguntas:

- Puede un voluntario emplear el “protocolo” implementado en el screensaver de SETI para acceder a algunas de las funcionalidades del sistema?
- Puede el voluntario sin necesidad de cambiar el screensaver donar su tiempo de CPU para otro proyecto diferente?

De la forma en que se ha descrito SETI, la primera pregunta tiene a todas luces respuesta negativa, pues el sistema conoce solo como comunicarse con su contraparte cumpliendo los requisitos definidos por la regla ii. La respuesta a la segunda interrogante puede encontrarse atendiendo al hecho de que el código de SETI es residente en la maquina, para cambiarlo sería necesario cambiar la aplicación - el screensaver -, todo lo contrario a lo que se espera de un sistema Grid, donde solo debería quedar en el ordenador las funciones necesarias para enviar o recibir resultados o trabajo, es decir la respuesta también es negativa. Por lo que puede concluirse que a pesar de ser conocido como tal, SETI no hace uso de muchos de los conceptos que definirían a un sistema como Grid.

### 3. Tecnología Grid. Arquitecturas.

Desarrollar una infraestructura capaz de satisfacer cada uno de los requerimientos inherentes a un sistema Grid ha sido el centro de los trabajos desarrollados por una creciente comunidad de investigadores provenientes tanto de la industria como del sector académico. Los inconvenientes fundamentales que trae aparejado el auge de la tecnología Grid están

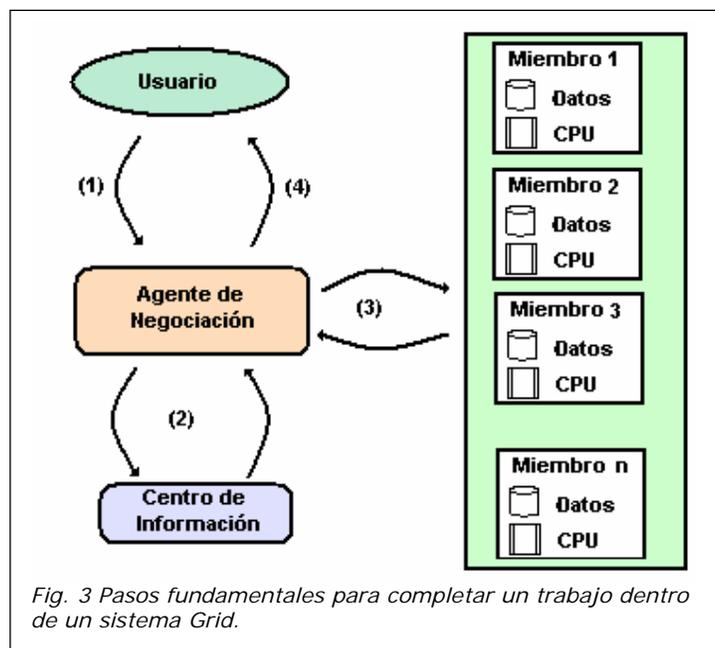
---

2 Aquí se hace referencia a la versión que funcionó desde 1999 hasta diciembre del 2005. Esta aplicación fue reemplazada por BOINC (**Berkeley Open Infrastructure for Network Computing**) que además permite a los usuarios contribuir con otros proyectos como Einstein@Home.

relacionados principalmente con el aspecto social, siendo uno de los puntos comunes en todos los proyectos de investigación sobre el tema, entre estos puede contarse: [Buyya et al., 2005]

- Perfeccionar la gestión distribuida al mismo tiempo que se mantiene el control total sobre los recursos administrados localmente.
- Acrecentar la disponibilidad de datos e identificar problemas y soluciones a cerca del acceso a los datos.
- Proveer a los desarrolladores e investigadores con un entorno uniforme y amigable que permita el acceso a un amplio espectro de recursos distribuidos.

Antes de explicar detalladamente los componentes que conforman un sistema Grid así como la forma en que estos se integran, es útil explicar aunque sea de manera bastante simplificada, la dinámica dentro de un entorno Grid de alcance global. En el sistema los recursos disponibles son registrados en uno o más centros de información (**Grid Information Services**). Los usuarios envían sus requerimientos hacia algún agente intermediario (**Grid Resource Broker**) encargado de “negociar” los recursos (1); el cual debe localizar los medios disponibles consultando los servicios de información (2), agendar (**schedule**) el trabajo para ser ejecutado y monitorearlo hasta que sea completado (3) para finalmente enviar los resultados al usuario (4). [Fig. 3][Smith, 2004]



Este entorno involucra la implantación de servicios de seguridad, información, asignación e integración de recursos, distribución de aplicaciones, administración de ejecución y programación de tareas. El conjunto de software utilitario y servicios que proveen estas capacidades se conoce como **Grid middleware**, el cual con vistas a ofrecer al usuario un entorno homogéneo debe resolver varias dificultades inherentes a las características de un sistema Grid. Primeramente es necesario solucionar los contratiempos originados por la heterogeneidad que resulta del amplio espectro de tecnologías de software y hardware integradas en el sistema. Además es necesario manejar un grupo de recursos dispersos geográficamente, sujetos a políticas administrativas reguladas por diferentes entidades, de

donde se desprende que la disponibilidad y el desempeño de los recursos es impredecible desde el momento en que las peticiones locales tengan mayor prioridad que las procedentes del exterior, lo cual dota al sistema de una naturaleza dinámica que constituye un nuevo problema a tratar. Estas dificultades han guiado a los investigadores a proponer una arquitectura para el sistema Grid basada en la creación de Organizaciones Virtuales (**Virtual Organizations o VOs**) formadas por diferentes organizaciones reales cooperando para alcanzar una meta común.[Buyya et al., 2005]

### 3.1 Organizaciones Virtuales (VOs).

La naturaleza externa de los recursos compartidos en el entorno Grid – CPU, capacidad de almacenamiento, datos, software etc. – es decir su disponibilidad más allá del dominio administrativo local conducen a la creación de un nuevo dominio de administración llamado Organización Virtual que implementa un nuevo conjunto de políticas, usualmente formado por políticas locales en adición a políticas externas de administración sobre los recursos del sistema. De esta forma, a través de los estándares ofrecidos por el arquitectura Grid los usuarios pueden agruparse dinámicamente en estas VOs definiendo cada una sus propios requerimientos y políticas de acceso y empleo de los recursos.[Berstis, 2002]. Es decir dentro del sistema Grid pueden formarse subsistemas relativamente independientes formados por grupos de usuarios - no necesariamente disjuntos- que definen sus propias reglas de cooperación en consonancia con sus necesidades.

Hasta este punto, se han podido identificar algunas de las componentes primarias que han de conformar la arquitectura de un sistema Grid. En primer lugar, la propia dinámica dentro del entorno, hace necesario la implantación de los mecanismos que permitan a los usuarios cumplimentar la secuencia de operaciones ilustrada en la Fig. 3. Así mismo, las problemáticas generadas por la concepción del propio sistema conducen a la creación de las Organizaciones Virtuales lo que a su vez da paso a la necesidad de implementar las funciones necesarias capaces de dar soporte a este mecanismo.

### 3.2 Calidad de los Servicios (QoS).

Sin embargo, la luz bajo la cual se han observado los sistemas Grid hasta el momento, no permite identificar otra de las piezas que podrían ser fundamentales dentro del entorno. Al hablar sobre el proyecto [SETI@home](#) - ciertamente en todo momento - se ha dado por sentado que al sistema se incorporan usuarios que voluntariamente donan sus recursos, pero como puede sospecharse esto no es así del todo. Sin necesidad de enumerar la extensa lista de beneficios que brinda la computación Grid, salta a la vista una perspectiva comercial del modelo. Es decir, los sistemas Grid pueden verse como la oportunidad de crear un enorme centro de procesamiento cuyos servicios puedan venderse a clientes interesados. Entre otros elementos atractivos, esta facilidad libra a los compradores de tener que adquirir, configurar y administrar grandes redes de ordenadores, incentivo que aumenta cuando el cliente no se ve necesitado de realizar proyectos de gran envergadura computacional con

mucha frecuencia [Fordhal,2005]. Claramente, esta nueva perspectiva viene acompañada con la necesidad de implementar un modelo de Calidad de los Servicios (**Quality of Services o QoS**) capaz de manejar información sobre precios, disponibilidad y características de los recursos, así como definir los términos sobre los que se basa el contrato cliente-proveedor que incluye ciertos compromisos entre las partes – como el tiempo máximo para completar una tarea – así como las medidas que se desprendan de su incumplimiento.

### 3.3 Arquitectura de Capas.

Además de las herramientas para dar soporte a los modelos VOs y QoS, la infraestructura de un entorno de computación Grid debe comprender las siguientes facilidades: [Buyya et al., 2005]

- Almacenamiento o replicación remota de almacenes de datos.
- Publicación de almacenes de datos usando nombres y atributos globales en los correspondientes catálogos.
- Seguridad, autorización de acceso y mecanismos de autenticación y acceso uniforme a los recursos remotos.
- Publicación o localización de los servicios y recursos aptos para su empleo así como su costo de uso.
- Creación de aplicaciones distribuidas empleando diferentes componentes de software.
- Planificar y agendar tareas así como monitorear las que se encuentran en ejecución.
- Traspaso de código y datos entre los usuarios y los recursos distribuidos.
- Medir monto del empleo de los recursos.

Los componentes del modelo Grid que permiten las características mencionadas han sido organizados en un sistemas de capas [Buyya et al., 2005], donde los servicios de cada una de estas se sustentan en los de la capa inferior y en la interacción entre los servicios dentro de la misma capa. Aunque existen diferentes propuestas - modelo referido por [Smith,2005] - en cuanto a las funciones correspondientes a cada capa así como al número de estas la arquitectura más difundida consta de cuatro capas [Fig. 4]

- **Fabric Layer:** Consiste en el conjunto de recursos distribuidos, tales como clusters, supercomputadoras, servidores, PCs, dispositivos de almacenamiento, instrumentos científicos, etc. .

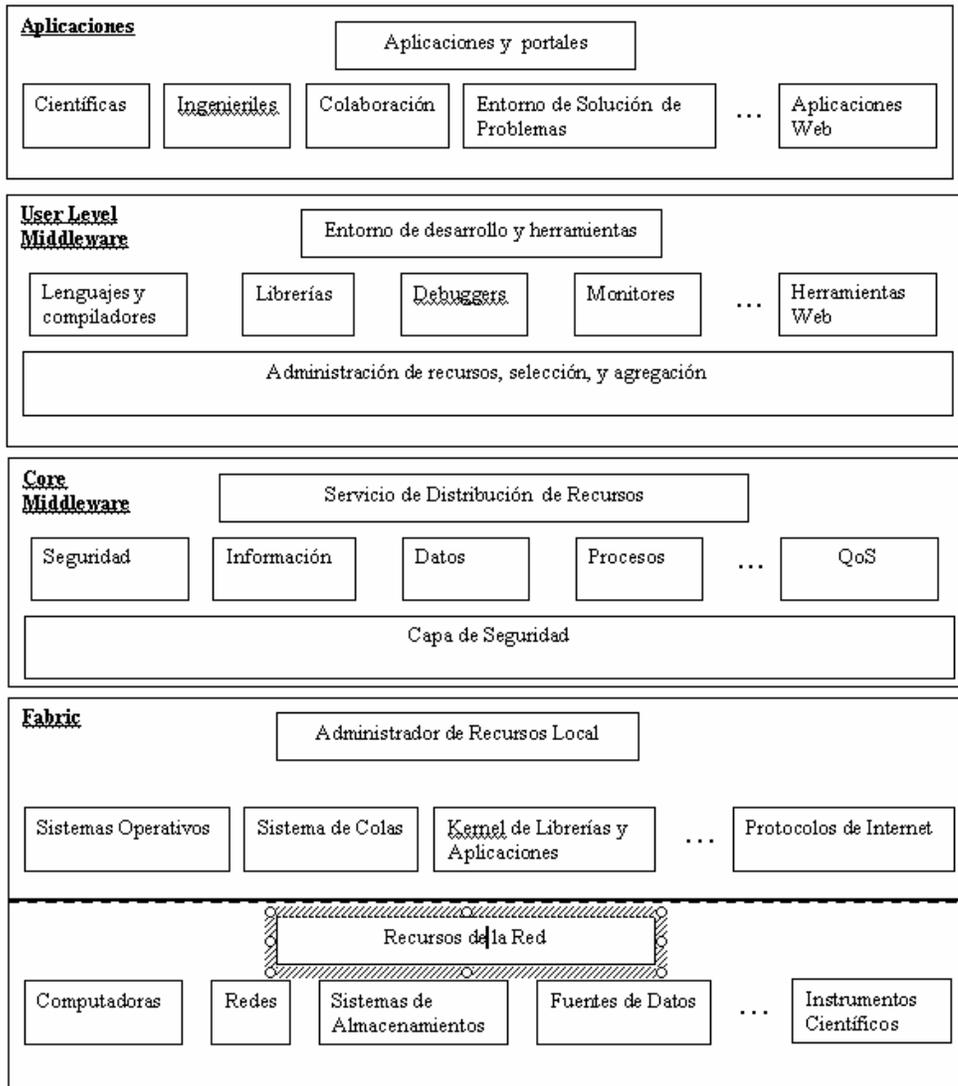


Fig. 4 Arquitectura de capas de un sistema Grid.

- **Core Middleware Layer:** En esta se agrupan los servicios que posibilitan la gerencia remota de procesos, asignación de recursos, acceso de almacenamiento, registro y recuperación de información, seguridad, aspectos de QoS como reservación y negociación de recursos. Estos servicios brindan una abstracción respecto a la heterogeneidad intrínseca de la capa precedente - Fabric Layer - brindando métodos consistentes y uniformes de acceso a los recursos distribuidos.

- **User-level Middleware Layer:** Esta capa provee servicios con un mayor nivel de abstracción tales como entornos de desarrollo, herramientas de programación y agentes (**brokers**)<sup>3</sup>, mecanismos que se encargan de negociar los recursos y agendar tareas para su ejecución en el sistema.
- **Applications Layer:** Aplicaciones o portales desarrollados empleando entornos de programación Grid así como las interfaces y servicios brindados por la capa inferior. En el caso de los servicios Web, los usuarios pueden enviar y recibir los resultados de su trabajo a través de la Web.

Una vez realizada la disección del sistema, es posible examinar nuevamente la secuencia de pasos que se habían mostrado como necesarios para cumplimentar una tarea dentro de un entorno Grid con vistas a ilustrar más detalladamente la forma en que interactúan los diferentes componentes del sistema. En primer lugar es necesario incorporar los recursos dentro del entorno y hacerlos accesibles en diferentes dominios de administración, para esto pueden instalarse diferentes módulos - Core Middleware - como Globus para Unix o Linux y Alchemi para Windows. Los recursos compuestos por varios nodos, como los clusters, deben ser presentados como un solo elemento esto puede lograrse empleando herramientas como Condor, PBS o Sun Grid Engine. En los recursos responsables de proveer los servicios de "brokering" es necesario implantar las herramientas de usuario - User Middleware - pertinentes. Resueltos los detalles de instalación y configuración de los servicios, puede verse con mayor claridad la manera en que interactúan los componentes durante el cumplimiento de una tarea. [Buyya et al., 2005]

#### 1. Los usuarios diseñan su aplicación siguiendo un modelo distribuido:

Esta puede ser fragmentada en determinado número de aplicaciones individuales, las cuales a su vez podrían dividirse en diferentes sub-aplicaciones en ocasiones llamadas transacción (**transaction**) o unidad de trabajo (**work unit**). Una aplicación organizada según este esquema usualmente está diseñada para que las sub-aplicaciones sean ejecutadas en paralelo en diferentes nodos del sistema, no obstante determinadas sub-aplicaciones podrían depender de ciertas condiciones que impidiesen su ejecución en forma paralela, al menos respecto a otras unidades de trabajo, por ejemplo cuando se requiere los resultados producidos por otra sub-aplicación. Cuanto menos de estos fragmentos presenten estos inconvenientes se dice que más **escalable** es la aplicación [Berstis, 2002], pudiendo esperarse que una aplicación perfectamente escalable, termine diez veces más rápido si emplea un número de procesadores diez veces mayor.

---

3 Resource brokers: En este caso la negociación de los recursos se entiende desde una óptica social, a diferencia de los mecanismos de negociación que se brindan en la capa anterior - Core Middleware - que pueden ser un conjunto de herramientas que implementen cierto protocolo de negociación, los cuales serían empleados por el broker.

2. Los usuarios especifican sus requerimientos en cuanto a calidad de los servicios los que son enviados junto con el trabajo a realizar.
3. El agente de negociación (**resource broker**):
  - i. Localiza y hace un análisis de las características de los recursos disponibles empleando los servicios de información.
  - ii. Identifica los precios de los recursos consultando los directorios de información sobre el mercado (**Grid Market Directory**).
  - iii. Identifica una lista de almacenes de datos seleccionado los más convenientes.
  - iv. Identifica una lista de los medios de cómputo que proveen los servicios adecuados.
  - v. Se asegura que el usuario tenga el crédito o el nivel de acceso requeridos.
  - vi. Planifica el trabajo en los medios que cumplen los requerimientos de calidad de los servicios especificados por el usuario:

El mecanismo para agendar explicado sigue un esquema bien simple - distribuir cada tarea en el primer dispositivo que cumpla con los restricciones especificadas - pero en ocasiones es ventajoso contar con una modelo más avanzado, así algunos sistemas implementan un sistema de prioridades, formando diferentes colas de trabajos según la prioridad de estos. Cuando existe un medio disponible se toma una tarea de la cola de mayor prioridad. Además permiten establecer ciertas reglas de ejecución para los trabajos, por ejemplo especificar días y horas concretos en que se correrá.

Con vistas a garantizar el nivel esperado en los índices de calidad de los servicios - y cuando el sistema lo permita - el usuario puede optar por reservar un conjunto de medios con antelación para su uso exclusivo o con un mayor nivel de prioridad.

4. Los agentes de negociación en los recursos ejecutan el trabajo y envían los resultados:

Los ficheros de la aplicación pueden ser instalados con antelación en las máquinas o hacerse accesibles a través de un sistema de archivos distribuido

montable. Cuando en el sistema está integrado por un conjunto heterogéneo de medios pueden encontrarse diferentes archivos de aplicación, cada uno compilado para las diferentes plataformas. Algunos sistemas Grid realizan cierto procesamiento de los datos o el programa a ejecutar con el objetivo de adicionar rutinas de protección y control de forma que cualquier problema durante la ejecución no cause daños a la maquina donde se está ejecutando el trabajo [Berstis, 2002].

La ejecución de las sub-aplicaciones puede ser monitorizada manual o automáticamente para seguir el progreso de la aplicación o detectar fallas de hardware, software, comunicación etc. tratando de garantizar la confiabilidad del sistema al establecer esquemas de recuperación ante estas fallas.

5. El agente de negociación compagina los resultados y los pasa al usuario.
6. El sistema de metrado factura al usuario enviando la información sobre el uso de los recursos al sistema de contaduría el cual lo reporta al usuario.

#### 4. Consideraciones Finales.

El amplio catálogo de posibilidades que brinda la computación Grid ha motivado un creciente interés en esta tecnología. Estos son algunos de los principales proyectos:

- Estados Unidos:
  - TeraGrid: Integra ordenadores, bases de datos y herramientas a lo largo de todo el país. Cuenta con más de 102 teraflops de capacidad computacional y más 15 petabytes de datos organizados en alrededor de 100 bases de datos sobre diferentes disciplinas.
  - GriPhyN: Proyecto que aúna investigadores de las Tecnologías de la Información y la Física Experimental con el objetivo de crear Data Grids (**Petascale Virtual Data Grids o PVDG**) que satisfagan los requerimientos de manejo de datos a gran escala a científicos en todo el mundo.
  - Condor Project: Desarrolla, implementa y evalúa mecanismos para dar soporte a la computación de altas prestaciones a través de largas colecciones de recursos computacionales distribuidas. Provee herramientas para el manejo de colas de tareas, políticas de planificación, esquemas de prioridades, monitoreo de recursos etc.
- Europa:
  - EGEE (Enabling Grids for E-sciencE): Desarrollado dentro de la EU Research Network une expertos de más de 27 países con el fin de aprovechar los avances de la tecnología Grid para proveer a los investigadores acceso a recursos de computo independientemente de su ubicación geográfica.

Otros países cuentan con programas para el desarrollo de la tecnología Grid como los japoneses Ninf Project (**Tokyo Institute of Technology**) y NAREGI (**National Research Grid Initiative**), APAC Grid en Australia, en China el China National Grid y en la India el proyecto conocido como Kerala Education Grid.

El incremento notable de la comunidad Grid convierte en una necesidad la estandarización de protocolos y modelos para asegurar la interoperabilidad entre diferentes productos e implementaciones, lo cual redundaría en un mejor aprovechamiento de las prestaciones de la tecnología. Como institución, el Foro Global Grid (**Global Grid Forum o GGF**) ha guiado estos esfuerzos, estableciendo estándares para casi todos los aspectos de la tecnología Grid, siendo notables OGSi (**Open Grid Service Infrastructure**) y WSRF (**Web Services Resource Framework**) que direccionan la integración de los Servicios Web dentro de la arquitectura Grid permitiendo a los desarrolladores aprovechar protocolos y formatos estándares como HTTP y XML para la comunicación entre un conjunto no uniforme de componentes y arquitecturas.

Desgraciadamente no todas las aplicaciones son paralelizables, en otros casos el esfuerzo necesario para obtener una versión capaz de aprovechar las potencialidades de la tecnología Grid puede ser muy costoso.

## 5. Referencias Bibliográficas.

- Berstis, Viktors. 2002. "Fundamentals of Grid Computing". IBM Corp.
- BOINC. <<http://boinc.berkeley.edu>>
- Brown, Martin. 2005. "Build grid applications based on SOA". Disponible en <<http://www-128.ibm.com/developerworks/grid/library/gr-soa/index.html>>. Fecha de consulta: 24 de mayo, 2006.
- Buyya, Rajkumar y Srikumar Venugopal. 2005. "A Gentle Introduction to Grid Computing and Technologies". Computer Society of India Communications, 9 July 2005.
- CERN. 2005. "The Grid Cafe-What is the Grid?" Disponible en <<http://gridcafe.web.cern.ch/gridcafe/whatisgrid/whatis.html>>. Fecha de consulta: 1 de junio, 2006.
- Condor Project. <<http://www.cs.wisc.edu/condor>> Fecha de consulta: 2 de junio, 2006.
- EGEE. <<http://public.ey-egee.org>>. Fecha de consulta: 2 de junio, 2006.
- Einstein@home. <<http://einstein.phys.uwm.edu/>>
- Fordahl, M. 2005. "Sun Microsystems unveils grid computing initiative". Disponible en <[http://www.technologyreview.com/articles/05/02/ap/ap\\_2020205.asp](http://www.technologyreview.com/articles/05/02/ap/ap_2020205.asp)>. Fecha de consulta: 24 de mayo, 2006
- Foster, Ian. 2002. "What is The Grid". Argonne National Laboratory & University of Chicago.
- Global Grid Forum. <<http://www.ggf.org/>> Fecha de consulta: 2 de junio, 2006.
- Grid Computing Info Centre. <<http://www.gridcomputing.com>>. Fecha de consulta: 1 de junio, 2006.
- GridLab <<http://www.gridlab.org>> Fecha de consulta: 2 de junio, 2006.
- GryPhyN Project. <<http://www.griphyn.org/>> Fecha de consulta: 2 de junio, 2006.
- IBM Research. <<http://www.research.ibm.com/journal/sj43-4.html>>. Fecha de consulta: 1 de junio, 2006.
- NeuroGrid <<http://www.neurogrid.net>>. Fecha de consulta: 1 de junio, 2006.
- Seti@home. <<http://setiathome.berkeley.edu/>>. Fecha de consulta: 1 de junio, 2006.
- Smith, Roger. 2004. "Grid Computing: A Brief Technology Analysis". CTONet.org.
- TeraGrid. <<http://www.teragrid.org/>> Fecha de consulta: 1 de junio, 2006.
- Wikipedia.org. <[http://en.wikipedia.org/wiki/Computer\\_cluster](http://en.wikipedia.org/wiki/Computer_cluster)>. Fecha de consulta: 24 de mayo, 2006
- Wikipedia.org. <[http://en.wikipedia.org/wiki/Distributed\\_computing](http://en.wikipedia.org/wiki/Distributed_computing)>. Fecha de consulta: 24 de mayo, 2006
- Wikipedia.org. <[http://en.wikipedia.org/wiki/Grid\\_computing](http://en.wikipedia.org/wiki/Grid_computing)>. Fecha de consulta: 24 de mayo, 2006.
- Vazquez-Poletti, J. L., Eduardo H. Cuesta, Rubén S. Montero, Ignacio M. Llorente. 2006. "Una visión global de la Tecnología Grid." Departamento de Arquitectura de Computadoras y Automática, Universidad Complutense de Madrid.