



# Tema 1: Arquitectura General de un Entorno Grid

Conceptos de la Computación en Grid y Cloud



- Definir los conceptos claves de un Grid.
- Describir los componentes y la arquitectura de un entorno Grid.
- Describir la Arquitectura General de dos de los Middlewares más extendidos (globus y gLite).

- **Concepto de Grid.**
- Requisitos y Funcionalidad de los Grids.
- Descripción de los Elementos Básicos.
- Infraestructuras Grid.
- Middlewares Grid.



- El Término Grid se Acuñó a Principios de los 90 para Denotar una Metodología de Acceso a Recursos Distribuidos Orientada a la “Computación Bajo Demanda”.
- El Concepto de Grid se Refiere a:
  - “Coordinar Recursos que no son sujetos a un control centralizado utilizando protocolos e interfaces estándar, abiertos y de propósito general, para proporcionar una calidad de servicio No Trivial”

**Ian Foster, What is the Grid? A Three Point Checklist” GRID Today, 20 de Julio de 2002.**

- Las Tecnologías Grid Pretenden Por Tanto Integrar un Entorno Virtual de Computador Único Capaz de:
  - **Realizar Tareas Complejas y de Alto Coste Computacional** que normalmente requerirían Supercomputadores.
  - Proporcionar una Estructura para el Registro y Almacenamiento de los Datos Producidos en Grandes Instalaciones Científicas.
  - Utilizar de Forma Efectiva Recursos que Permanecerían inactivos la Mayor Parte del Tiempo Ociosos.

- Concepto de Grid.
- **Requisitos y Funcionalidad de los Grids.**
- Descripción de los Elementos Básicos.
- Infraestructuras Grid.
- Middlewares Grid.



# Requisitos y Funcionalidad de los Grids

## Virtualización de recursos y Organizaciones Virtuales

- Grid Proporciona un Paso Más en la Integración de Recursos Heterogéneos con Respecto a los Sistemas Distribuidos.
  - **A Nivel de Usuarios.** Éstos Integran Organizaciones Virtuales que Comparten Permisos de Acceso, Facilitando la Gestión de los Recursos.
  - **A Nivel de Datos.** El Almacenamiento Distribuido se Organiza a Través de Catálogos que Involucran Múltiples Recursos e Instituciones.
  - **A Nivel de Proceso.** Los servicios de Planificación y Metaplanificación Proporcionan la Visión de Computador Único.

# Requisitos y Funcionalidad de los Grids

## Fiabilidad y Alta Disponibilidad

- Muchos Procesos Requieren de un **Funcionamiento Continuo, 24x7**.
- La Fiabilidad (Baja Tasa de Fallos) y Alta Disponibilidad (Capacidad de Mantenerse Operativo) Requieren de Elementos Redundantes que se Encuentren Activos en Todo Momento y Puedan Retomar el Control del Sistema.
- Los Grids Proporcionan **Recursos y Servicios Redundantes** que Aumentan Ambos Aspectos.
  - Relanzamiento de Trabajos Ante Fallos.
  - Monitorización del Estado de los Recursos.
  - Manejo de Datos Replicados.



# Requisitos y Funcionalidad de los Grids

## Gestión y Planificación de Recursos

- La Gestión de los Recursos en un Grid es uno de los Aspectos Más Complejos.
- Por un Lado, es Necesario **Conocer el Estado y Configuración de los Recursos** de Forma Dinámica.
- Por Otro Lado, los Sistemas de Planificación Deben Analizar las Características de los Trabajos y **Encontrar el Recurso o Recursos Más Apropriados**.
- Los Objetivos que Se Persiguen son:
  - Equilibrio de la Carga.
  - Reducción del Tiempo de Respuesta.
  - Eficiencia.
  - Fiabilidad.

# Requisitos y Funcionalidad de los Grids

## Explotar Recursos Infrautilizados

- Compartir Recursos en un Grid Permite **Aprovechar los Recursos No Utilizados Disponibles** en la Infraestructura.
- Muchos Recursos se Encuentran por Debajo del 5% de Utilización.
- Los Recursos son Tanto Tiempo de Proceso como Espacio de Almacenamiento.
- Para Que esto sea Posible y Efectivo:
  - Los Programas Deben Poder Ejecutarse en Cualquiera de los Recursos Disponibles y Seleccionables. Esto Implica Utilizar Arquitecturas Homogéneas, Instalaciones Compatibles, Programas Multiplataforma o Incluso Recompilar Fuentes.
  - Los Planificadores Deben Disponer de la Información Sobre el Estado y las Características de los Recursos.

# Requisitos y Funcionalidad de los Grids

## Reducir Tiempos de proceso

- Compartir Centenares o Miles de Recursos de Proceso Proporciona la Capacidad de Abordar Problemas de Mayor Dimensión. Sin Embargo, no Todos los Procesos se Pueden Beneficiar del Procesamiento Paralelo, e Incluso Muchos Programas Paralelizables no Son Eficientes en una Infraestructura Grid.
- Son Generalmente Adecuados para HTC
  - Los Problemas Multiparamétricos.
  - Los Programas Paralelos de Grano Grueso y con Buena Escalabilidad.
- No Son Generalmente Adecuados para HPC
  - Los Programas Paralelos de Grano Fino.
  - Los Programas Paralelos con Alto Coste de Comunicaciones.
  - Los Programas con Requerimientos de Tiempo Real.

- Concepto de Grid.
- Requisitos y Funcionalidad de los Grids.
- **Descripción de los Elementos Básicos.**
- Infraestructuras Grid.
- Servicios Básicos en Entornos Grid.
- Middlewares Grid.



- Conceptos Básicos en los Grids
  - Organización Virtual
  - Recurso
  - Servicios
  - Trabajos y Aplicaciones
  - Infraestructura Grid

# Descripción de los Elementos Básicos

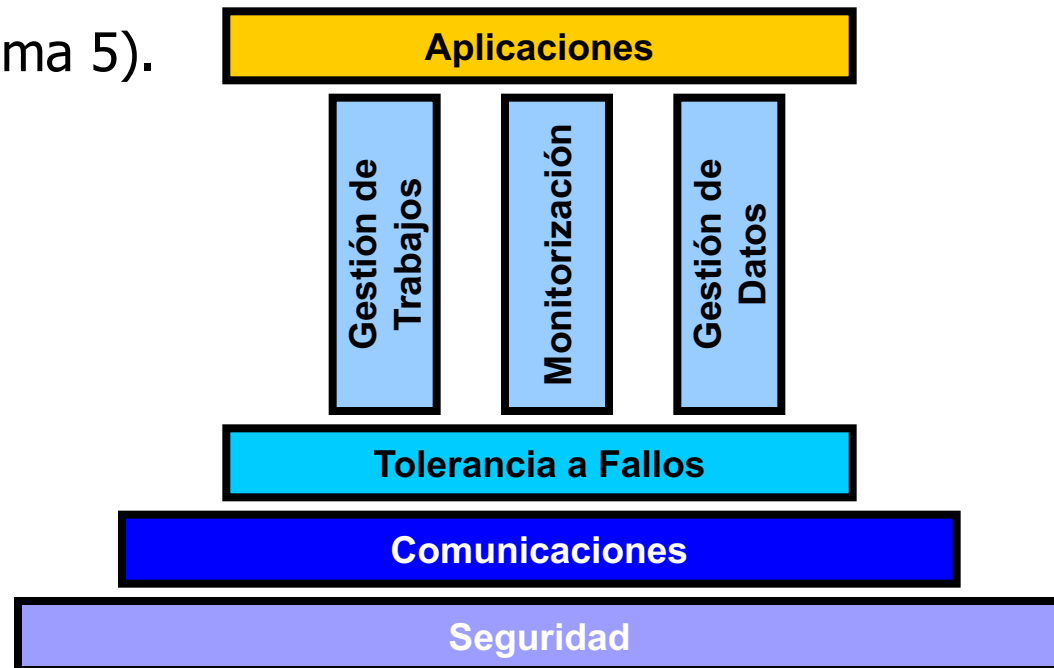
## Organización Virtual

- Una **VO** es una Asociación de Usuarios de Diferentes Organizaciones (Dominios Administrativos) Para la Resolución de Problemas Comunes.
- Generalmente Consiste en un Directorio de Identificaciones de Diferentes Usuarios a los que se les Autorizará el Uso de Recursos de una Forma Global (mas o menos).

## ● Tipos de Recursos

- Proceso: Ciclos de Computación Disponibles en los computadores del Grid.
- Almacenamiento: Primaria (RAM) o Secundaria.
- Comunicaciones: Anchos de Banda de Diferentes Conexiones y Aprovechamiento de la Topología.
- Software y Licencias: Aplicaciones Instaladas en Computadores.
- Equipos Especiales: Radiotelescopios, Equipos de Adquisición de Imágenes, Aceleradores de Partículas.

- Los Servicios Permiten Gestionar de Forma Global los recursos Grid.
  - Seguridad (Tema 2).
  - Gestión de Trabajos (Tema 3)
  - Gestión de Datos (Tema 4)
  - Sistema de Información (Tema 5).





# Descripción de los Elementos Básicos

## Trabajos y Aplicaciones

- Tipos de Trabajos
  - Multiparamétricos: Muchas Repeticiones de una Ejecución Variando los Datos de Entrada.
  - Paralelos: Una Única Ejecución Distribuida Entre Múltiples Computadores.
- Un Trabajo Contiene Ejecutables, Datos de Entrada, Especificación del Destino de los Resultados, Parámetros de Ejecución, Requerimientos y Priorizaciones.

# Descripción de los Elementos Básicos Infraestructura Grid

- European Grid Infrastructure (EGI)
  - [www.egi.eu](http://www.egi.eu)
- Worldwide LHC Computing Grid (WLCG)
  - <https://home.cern/science/computing/grid>

- Concepto de Grid.
- Requisitos y Funcionalidad de los Grids.
- Descripción de los Elementos Básicos.
- Infraestructuras Grid.
- **Middlewares Grid**



# Middleware Grid Globus Toolkit

- Por Tanto, Para el Estudio de los Entornos Grid se ha Seleccionado (Globus Toolkit <http://www.globus.org>).