



# **Tema 5. Sistemas de Información y Descubrimiento de Recursos en Entornos Grid**

Conceptos de la Computación en  
Grid y Cloud



- Describir la Forma en que Se Publica el Estado y las Características de los Recursos en un Entorno Grid.
- Describir el proceso de Seleccionar un Recurso Compatible y Adecuado en una Infraestructura Grid.
- Definir la Arquitectura del Sistema de Monitorización y Descubrimiento de Recursos (MDS) de Globus toolkit.

- **Sistemas de Información y Descubrimiento de Recursos en Entornos Grid**
  - Necesidades de Sistemas de Información en Grid.
  - Arquitecturas.
- **Soluciones Propuestas por Globus Toolkit**
  - Monitoring and Discovery Service (MDS).
  - Arquitectura MDS.
  - Integración MDS en la Arquitectura de Globus toolkit.
  - Grid Laboratory Uniform Environment (GLUE) Schema
  - Consulta de Información en el MDS

# Sistemas de Información y Descubrimiento de Recursos en Entornos Grid Necesidades

## Para el Usuario

Obtiene Información de los Recursos Grid y su Estado.

Obtiene Información del Estado de los Trabajos.

## Para el Desarrollador de Middleware

### Gestor de Carga:

Adecuar los Requerimientos de los Trabajos y de los Recursos del Grid.

### Monitorización:

Recuperación de Información del Estado y Disponibilidad de los Recursos.

## Para el Administrador de un Site o un Servicio

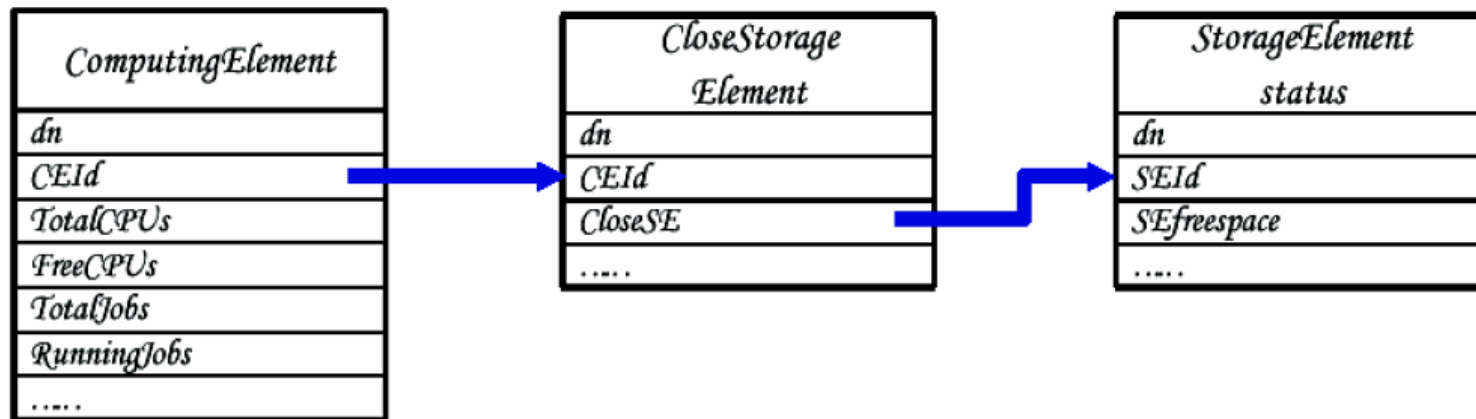
Publicar la Información Relativa a las Capacidades de un Site o Servicio

- Representación Jerárquica
  - Estructura en Árbol (Cada Hijo Tiene Sólo un Padre).
  - Escalable a una Identificación Única Global (p.e. Basada en el DNS)
  - Modelado Natural de Datos Físicos (p.e. Países, Instituciones, Recursos,...)
  - Acceso a Datos Rápido y Eficiente.

- **Representación Relacional**
  - Modelo de Datos Más Rico Con Relaciones Más Complejas.
  - Soporte de Búsquedas Complejas y Eficientes (SQL), Información Agregada, etc.
  - Diseñado Para Actualizaciones Frecuentes y Soporte de Transacciones.
  - No Maneja Tipos de Datos Complejos.

# Sistemas de Información y Descubrimiento de Recursos en Entornos Grid

## Arquitecturas – Representación Relacional



*“ Select a ComputingElement with at least 1 free CPU that also has a CloseStorageElement with at least 1000 MB of free space”*

```
SELECT DISTINCT ComputingElement.CEId FROM  
ComputingElement, CloseStorageElement, StorageElementStatus WHERE  
ComputingElement.FreeCPUs > 0 AND  
(ComputingElement.CEId = CloseStorageElement.CEId AND  
CloseStorageElement.CloseSE = StorageElementStatus.SEId AND  
StorageElementStatus.SEfreespace > 1000)
```



- Sistemas de Información y Descubrimiento de Recursos en Entornos Grid
  - Necesidades de Sistemas de Información en Grid.
  - Arquitecturas.
- **Soluciones Propuestas por Globus Toolkit**
  - Monitoring and Discovery Service (MDS).
  - Arquitectura MDS.
  - Integración MDS en la Arquitectura de Globus toolkit.
  - Grid Laboratory Uniform Environment (GLUE) Schema
  - Consulta de Información en el MDS



# Soluciones Propuestas por Globus Toolkit

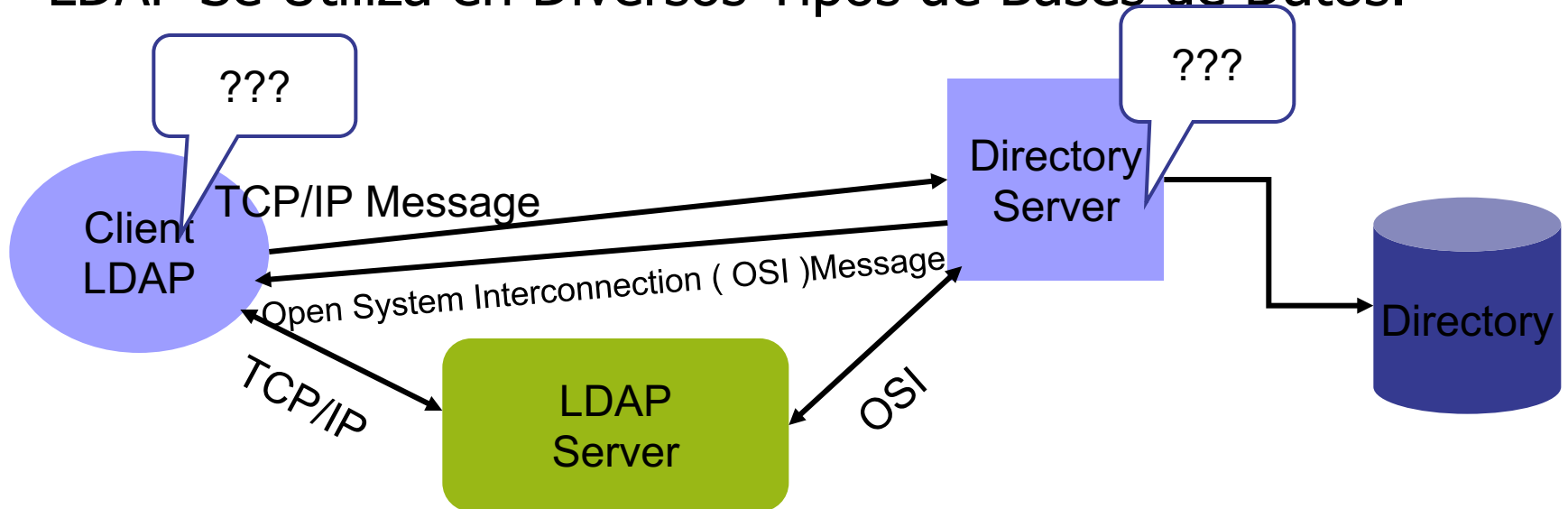
## Monitoring and Discovery Service (MDS)

- MDS es el Sistema de Información Para los Recursos en un Grid Desarrollado para Globus.
- Corresponde con una arquitectura Jerárquica.
- MDS a Grandes Rasgos Realiza Cuatro Tareas:
  - Obtener la Información Actualizada de cada Recurso.
  - Integrar esa Información en una Base de Datos General.
  - Permitir el Acceso Remoto a Esta Información.
  - Estructurar la Información Disponible.
- MDS es También el Sistema de Información de Otros Middlewares Utilizados en infraestructuras Avanzadas Grid como gLite.
- MDS se implementa a través de servicios LDAP.

# Soluciones Propuestas por Globus Toolkit

## Monitoring and Discovery Service (MDS) - LDAP

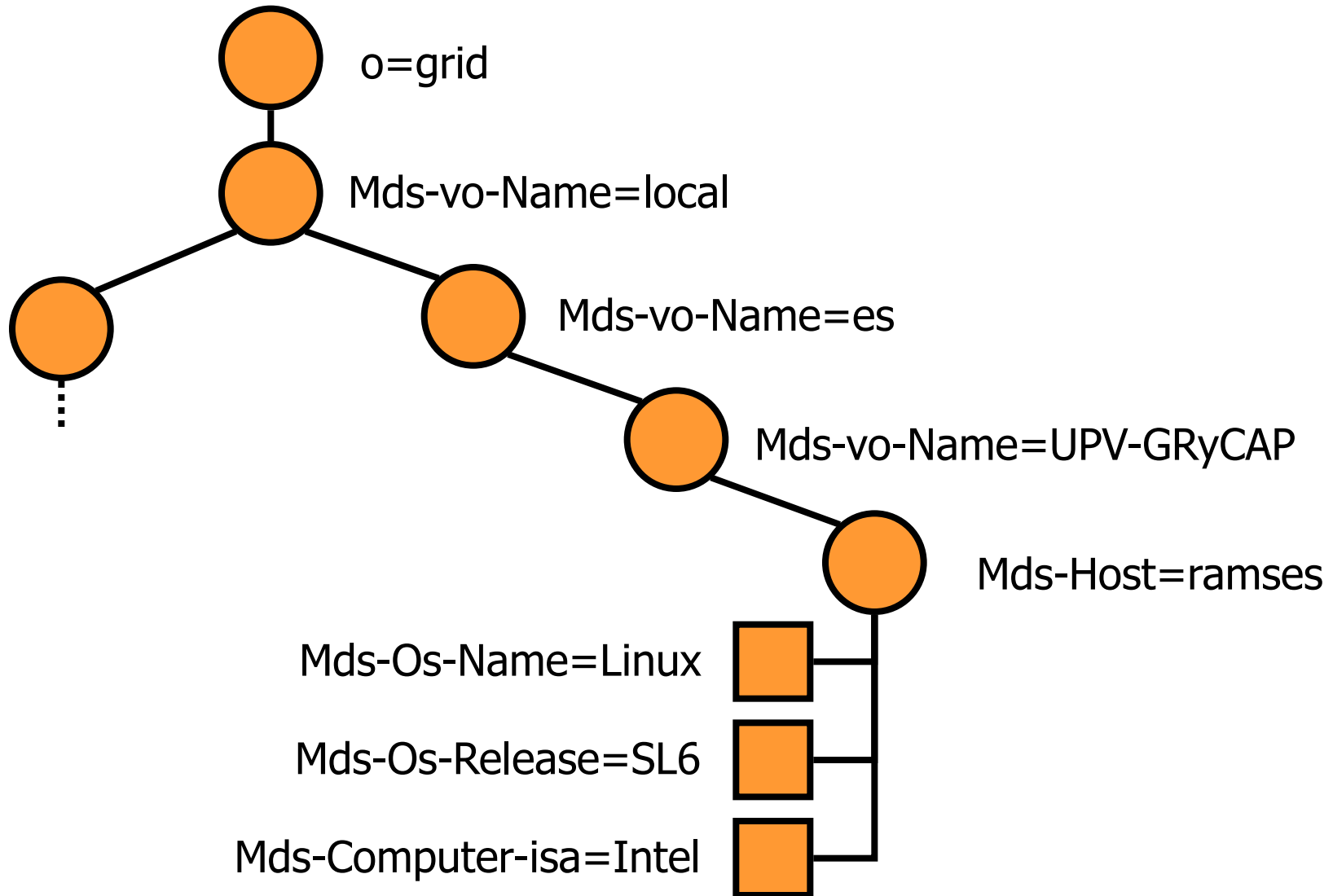
- LDAP Proporciona un Modelo Estándar Para el Almacenamiento y Recuperación de los Datos Con Estructura de Directorio en Árbol (Directory Information Tree - DIT).
- LDAP Establece el Formato de los Mensajes que un Cliente Utilizará para Acceder a un Directorio (LDIF).
- LDAP Se Utiliza en Diversos Tipos de Bases de Datos.



- El Modelo de Información LDAP está Basado en Entradas.
- Las Entradas son Colecciones de Atributos Definidas por un Nombre Distintivo (DN) Único y Global.
- La Información se Organiza en Forma de Árbol.
  - Existe un Atributo Especial (objectclass) que Permite Definir el Árbol Correspondiente a Cada Entrada.
  - La Información Puede Filtrarse Atendiendo a la Posición en el Árbol.
- La Información se Importa y se Exporta desde y hacia el Servidor LDAP Mediante Ficheros LDIF.

# Grid Laboratory Uniform Environment (GLUE) Schema

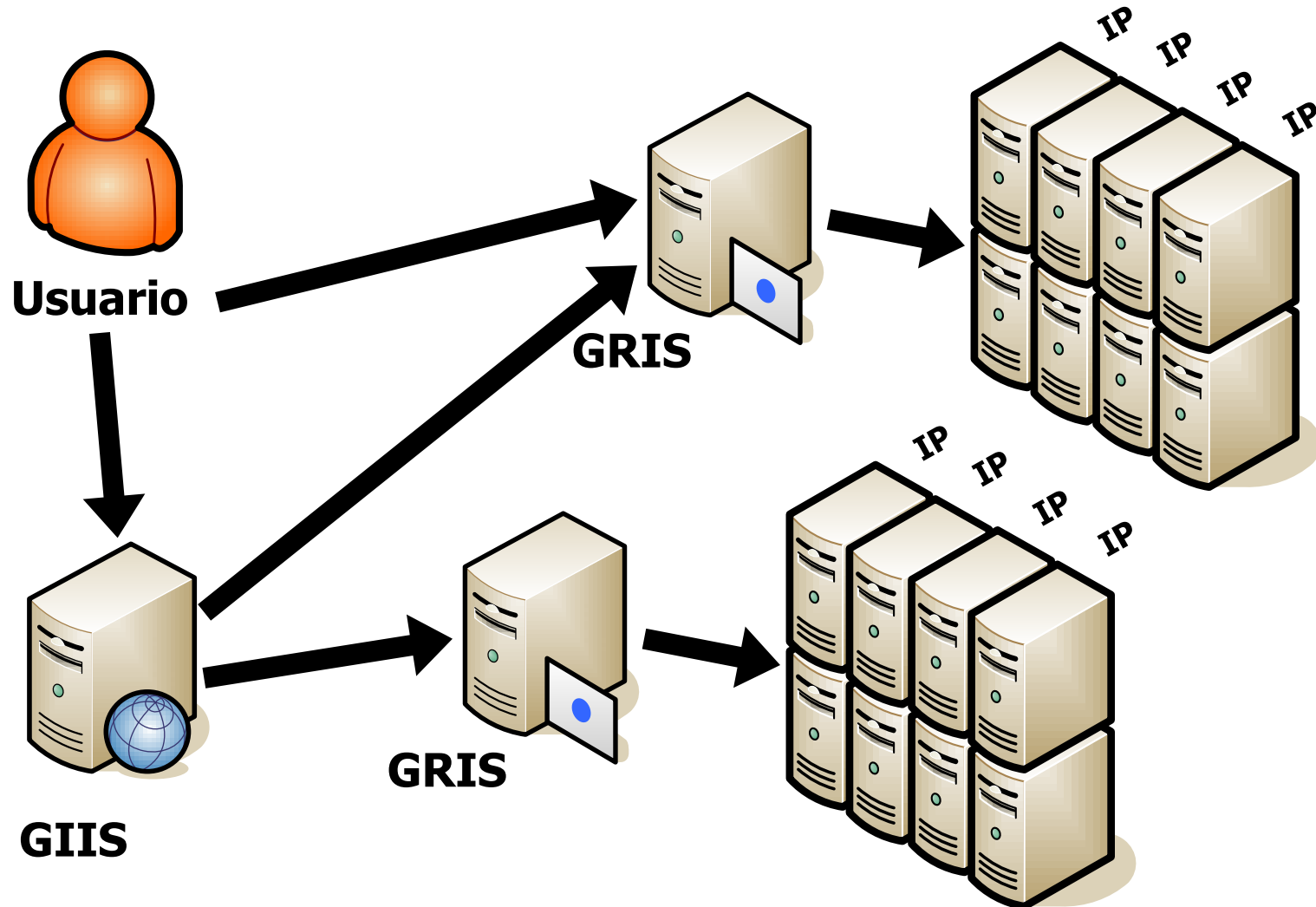
## Conceptos Generales - LDAP



- MDS se Compone de Tres Elementos Principales:
  - Proveedores de Información (Information Providers ó IP)
    - Se Ejecutan en Cada Recurso y Obtienen la Información del Estado del Mismo.
  - Grid Resource Information Service (GRIS)
    - Registra la Información Obtenida por los Diferentes IP para su Consulta.
  - Grid Index Information Service (GIIS)
    - Agrega la Información Publicada por Diferentes Servicios GRIS.

# Soluciones Propuestas por Globus Toolkit

## Arquitectura MDS



- Existen Dos Tipos de Information Providers
  - Core Ips:
    - Proporcionan Información Básica del Sistema
      - Tipo de Plataforma y Arquitectura
      - Nombre y Versión del Sistema Operativo
      - Número, Versión, Velocidad, Caché, etc. de la CPU.
      - Tamaño, Espacio Libre, etc. de la Memoria (Física y Virtual).
      - Nombre de la Máquina e IP.
      - Tamaño, Espacio Libre, etc. del Almacenamiento Secundario.
    - Existen Versiones de los IP Para Diferentes Plataformas
  - Custom Ips
    - Proporcionan la Información que el Usuario Considere Oportuna.
  - Los IPs Pueden Proporcionar la Información Atendiendo al MDS **core schema**.

- Los IPs se Registran en el GRIS que Periódicamente les Solicita Información Actualizada.
- La Información la Publica en unos Ficheros en Formato LDIF (LDAP).
- Los GRIS Contienen la Información de un Sólo Recurso y Pueden ser Consultados Directamente por el Usuario.



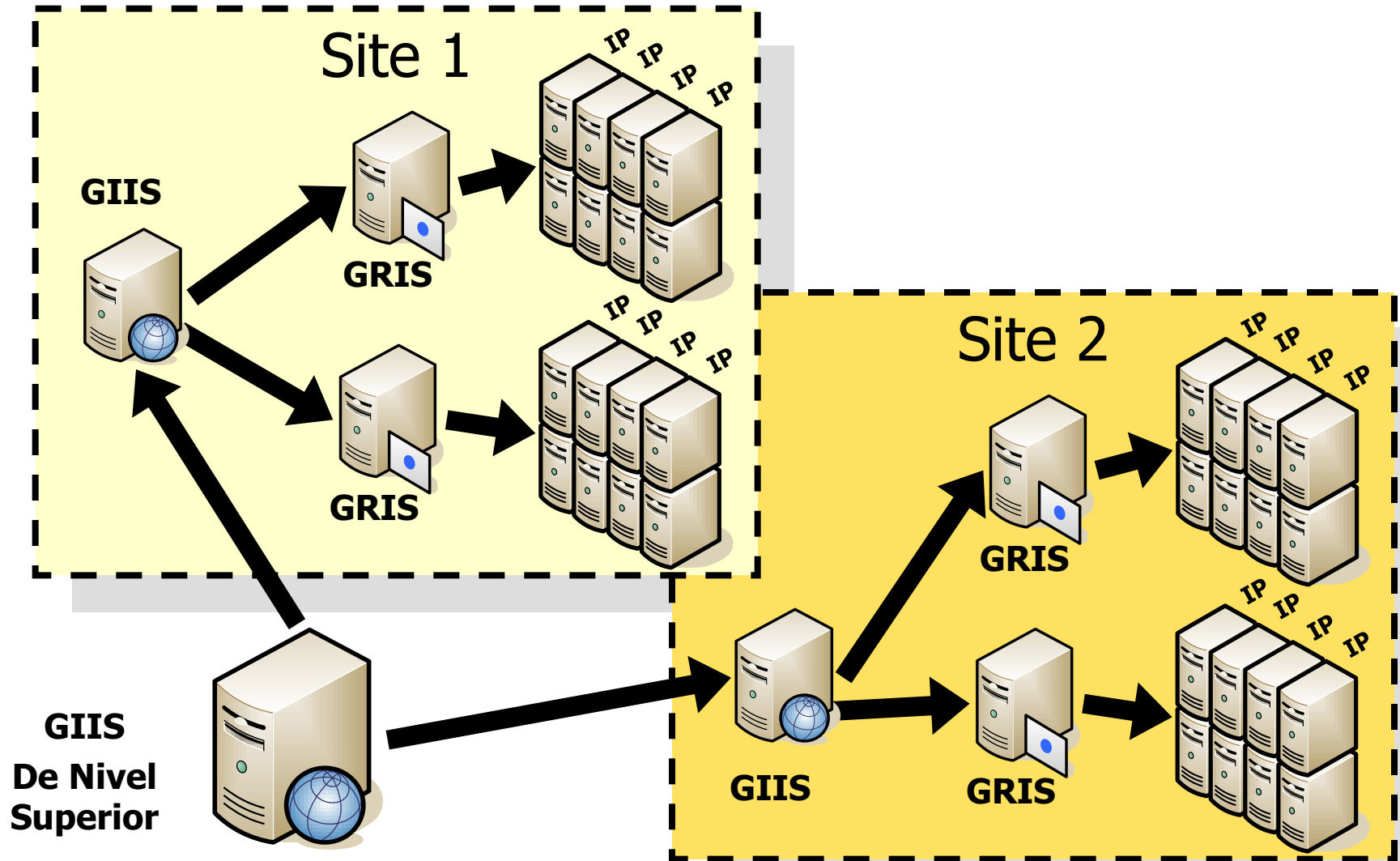
# Soluciones Propuestas por Globus Toolkit

## Arquitectura MDS – GIIS

- De Forma Análoga, los GRIS se Registran en los GIIS, Quienes Periódicamente les Requieren Información Actualizada.
- Los GIIS Actúan como “Caché” de la Información de los GRIS Consolidando la Información de Varios Recursos.
- De Esta Forma, es Posible Consultar Toda la Información de un Grid con un solo Comando.
- Los GIIS a su Vez Pueden Organizarse Jerárquicamente Aislando Diferentes Centros, Federaciones o Secciones de un Grid.
- También Pueden Configurarse de Forma Redundante, Permitiendo que un GRIS se Registre en Varios GIIS.

# Soluciones Propuestas por Globus Toolkit

## Arquitectura MDS



# Soluciones Propuestas por Globus Toolkit

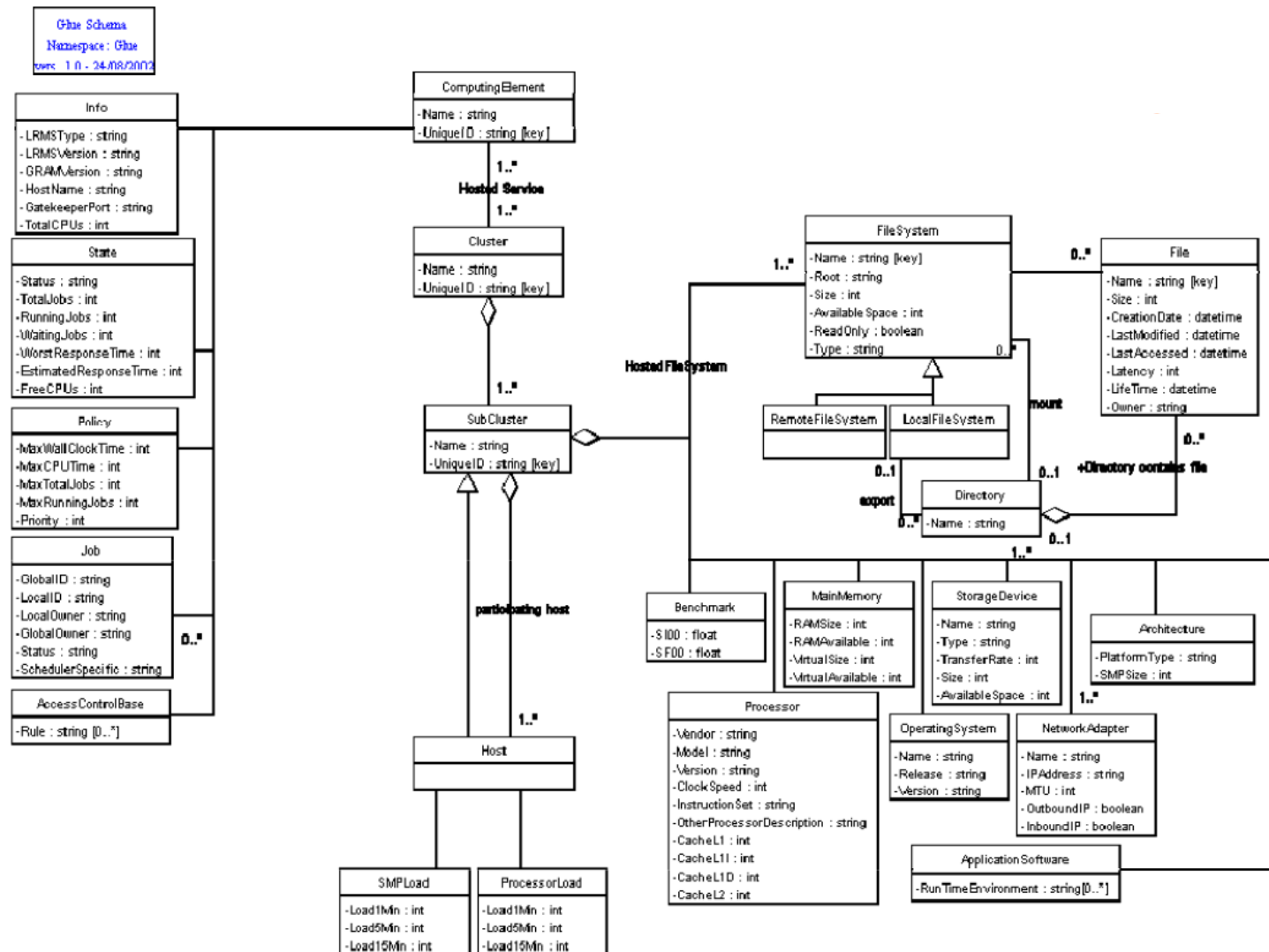
## Integración MDS en la Arquitectura de Globus toolkit

- El MDS Tiene en Cuenta la Arquitectura de Seguridad del GSI, con lo que Necesita de un Proxy Válido para la Ejecución de Comandos.
- El GRAM No Consulta el MDS Para la Selección de Recursos, Debería Hacerlo un Recurso de Planificación de Nivel Superior (o el Propio Usuario).
- Sin Embargo, Componentes del GRAM Interactúan con los IP para Proporcionar Información de Estado.

- Grid Laboratory Uniform Environment (GLUE) Describe la Información de los Recursos del Grid.
- Sigue la Estructura Jerárquica del Directory Information tree (DIT) para las objectclasses y los atributos:
- Top
  - GlueTop
    - 1. GlueGeneralTop
      - 1. ObjectClasses
        - » 1. GlueSchemaVersion
        - » 2. GlueCESEBindGroup
        - » 3. GlueCESEBind
        - » 4. GlueKey
        - » 5. GlueInformationService
      - 2. Attributes
        - » 1. Attributes for GlueSchemaVersion
        - » .....
        - » 5. Attributes for GlueInformationService
    - 2. GlueCETop

# Grid Laboratory Uniform Environment (GLUE) Schema

## Conceptos Generales – Modelo de Datos



# Grid Laboratory Uniform Environment (GLUE) Schema

## Conceptos Generales – Modelo de Datos – Computing Element

- Unique ID
- Name
- Info (Requerido)
  - Name of local resource management system, GRAM version, host name, gatekeeper port, total CPUs
- State (Opcional)
  - Running jobs, total jobs, status (queuing, production, closed, draining), waiting jobs, worst response time, estimated response time, free CPUs
- Policy (Opcional)
  - MaxWallClockTime, MaxCPUTime, MaxTotalJobs, MaxRunningJobs, Priority
- Job (Opcional)
  - owner local username, owner GSI subject name, job local id, job global id, job status (submitted, waiting, ready, scheduled, running, aborted, done, cleared, checkpointed), scheduler specific info
- AccessControlBase (Opcional)
  - Authorisation rule based on VO, X.509 certificate, etc.

- Los IPs se Integran a Partir del Fichero de Configuración Idif.
- En Este Fichero se Indican los Ejecutables que Obtienen la Información y Producen una Salida en el Modelo de Datos del LDAP.
- Los Ips son Programas que se Instalan en el Directorio `$GLOBUS_LOCATION/libexec`, Dependientes de la Plataforma y la Arquitectura, que Cuando se Ejecutan Imprimen la Información Correspondiente.
- Es Posible Añadir Nuevos Providers.

- `$GLOBUS_LOCATION/libexec/grid-info-mem-linux`

dn:  
objectclass: MdsMemoryRamTotal  
objectclass: MdsMemoryVmTotal  
Mds-Memory-Ram-Total-sizeMB: 4054  
Mds-Memory-Ram-Total-freeMB: 3474  
Mds-Memory-Vm-Total-sizeMB: 1992  
Mds-Memory-Vm-Total-freeMB: 1992  
Mds-Memory-Ram-sizeMB: 4054  
Mds-Memory-Ram-freeMB: 3474  
Mds-Memory-Vm-sizeMB: 1992  
Mds-Memory-Vm-freeMB: 1992  
dn: Mds-Device-Group-name=memory,  
objectclass: MdsMemoryRamTotal

objectclass: MdsMemoryVmTotal  
objectclass: MdsDeviceGroup  
Mds-Device-Group-name: memory  
Mds-validfrom: 20060504175446Z  
Mds-validto: 20060504175446Z  
Mds-keepsto: 20060504175446Z  
Mds-Memory-Ram-Total-sizeMB: 4054  
Mds-Memory-Ram-Total-freeMB: 3474  
Mds-Memory-Vm-Total-sizeMB: 1992  
Mds-Memory-Vm-Total-freeMB: 1992  
Mds-Memory-Ram-sizeMB: 4054  
Mds-Memory-Ram-freeMB: 3474  
Mds-Memory-Vm-sizeMB: 1992  
Mds-Memory-Vm-freeMB: 1992



- En un Centro, el GIIS Recopila la Información Proporcionada por los GRISs
  - `ldapsearch -x -h <hostname> -p 2135 -b "mds-vo-name=<name>,o=grid"`
- Los GRISs Locales Extraen la Información de los IPs y Reportan Dinámicamente Esta Información
  - `ldapsearch -x -h <hostname> -p 2135 -b "mds-vo-name=local,o=grid"`

- **Comando ldapsearch**
  - -x Autenticación Simple.
  - -H ldap://grid017.ct.infn.it:2170 Uniform Resource Identifier (URI).
  - -b DN Base para las búsquedas
    - '(objectclass=XXX)' Criterio de Filtrado
  - nombre\_atributo, Atributos a Retornar
- **ldapsearch -x**
  - H ldap://ngiescream.i3m.upv.es:2170
  - b o=grid
  - '(objectClass=GlueCE)' GlueCEInfoContactString