

Tecnologías Grid

Plataformas

Master en Sistemas y Servicios Informáticos para Internet
Área de Arquitectura y Tecnología de Computadores
Universidad de Oviedo



Plataformas

Introducción

□ Middleware grid

□ Definición:

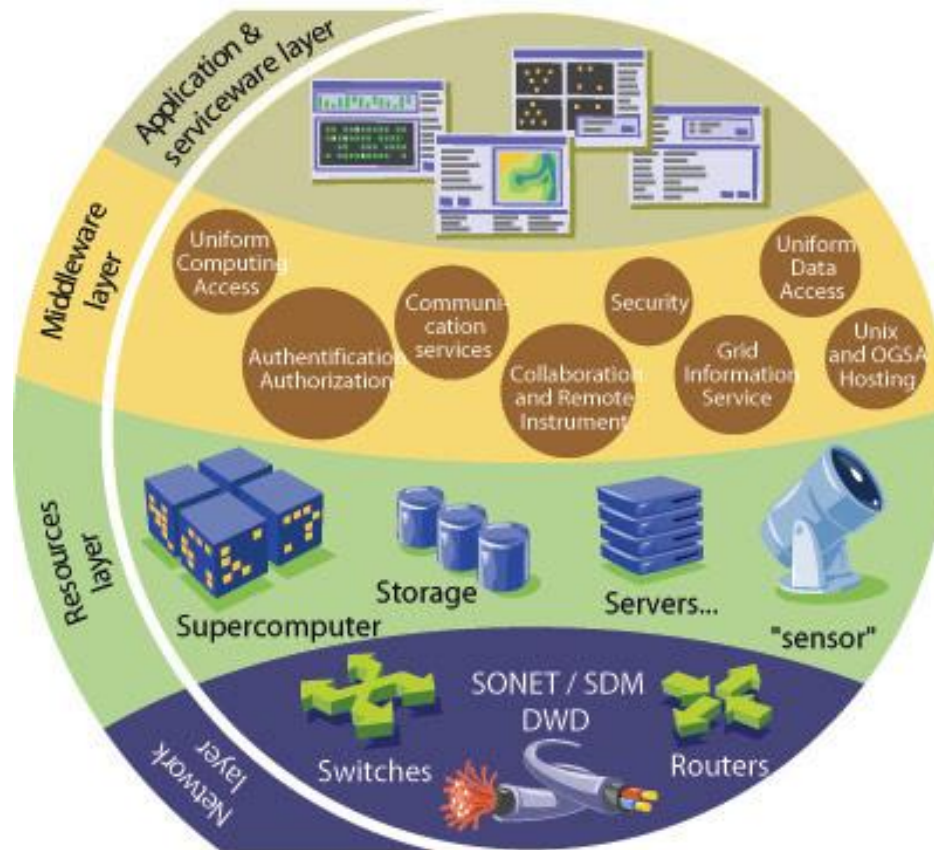
- Una capa de mediación que proporciona un acceso consistente y homogéneo a recursos locales heterogéneos

□ Múltiples implementaciones/arquitecturas:

- Se incluyen servicios comunes:
 - Gestión de trabajos, datos, descubrimiento
- Originalmente basados en protocolos y mecanismos propios
 - Tendiendo hacia la utilización de estándares grid

Plataformas

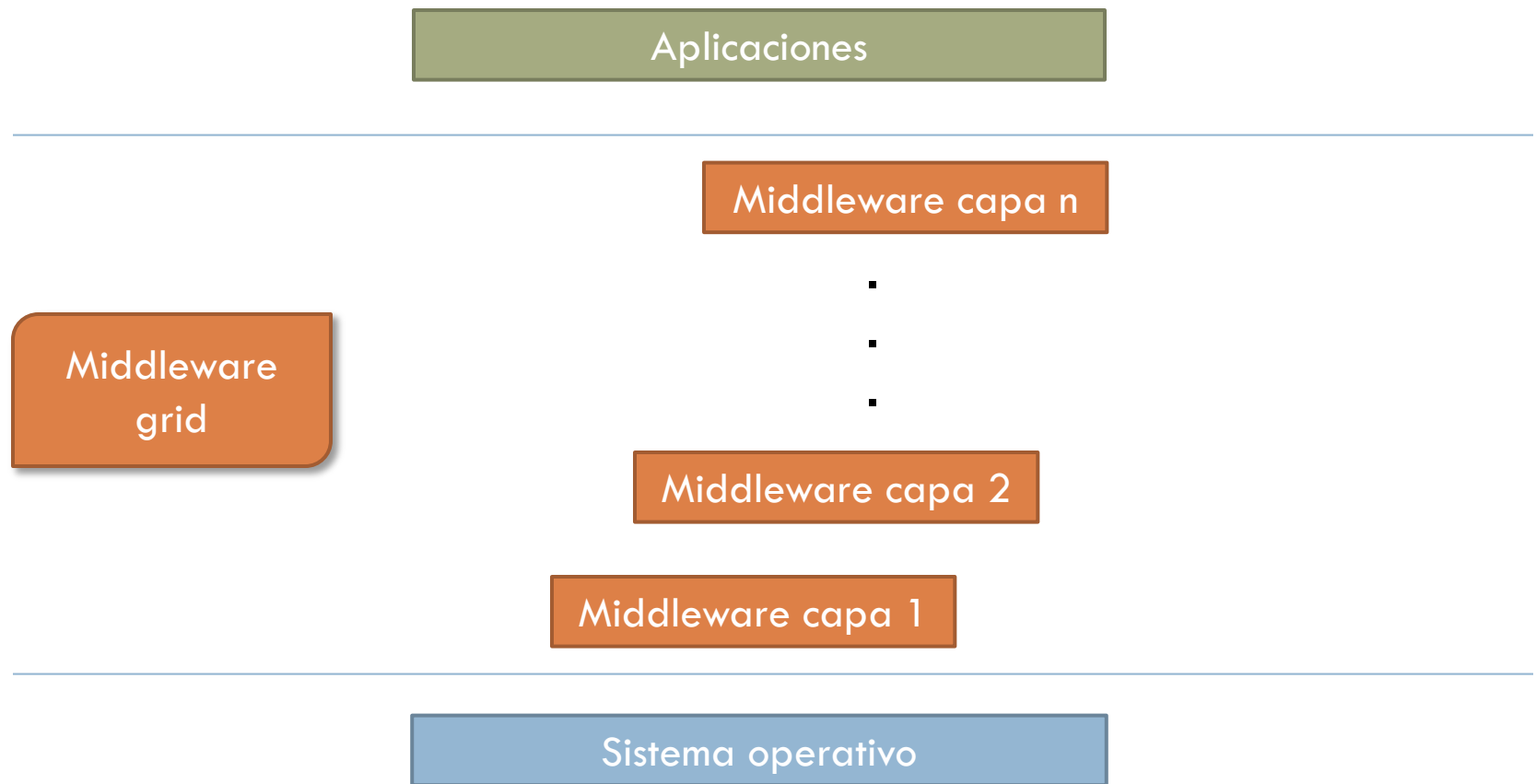
□ Middleware grid



<http://www.gridcafe.org>

Plataformas

□ Middleware grid



Plataformas

- Algunas características deseables en el Middleware grid:
 - ▣ Robustez
 - ▣ Transparencia
 - ▣ Facilidad de uso
 - ▣ Seguridad
 - ▣ Escalabilidad
 - ▣ Utilización de protocolos abiertos

Plataformas

- Algunos de los middleware grid más importantes:





Plataformas

Globus

□ Globus:



□ De forma simplificada:

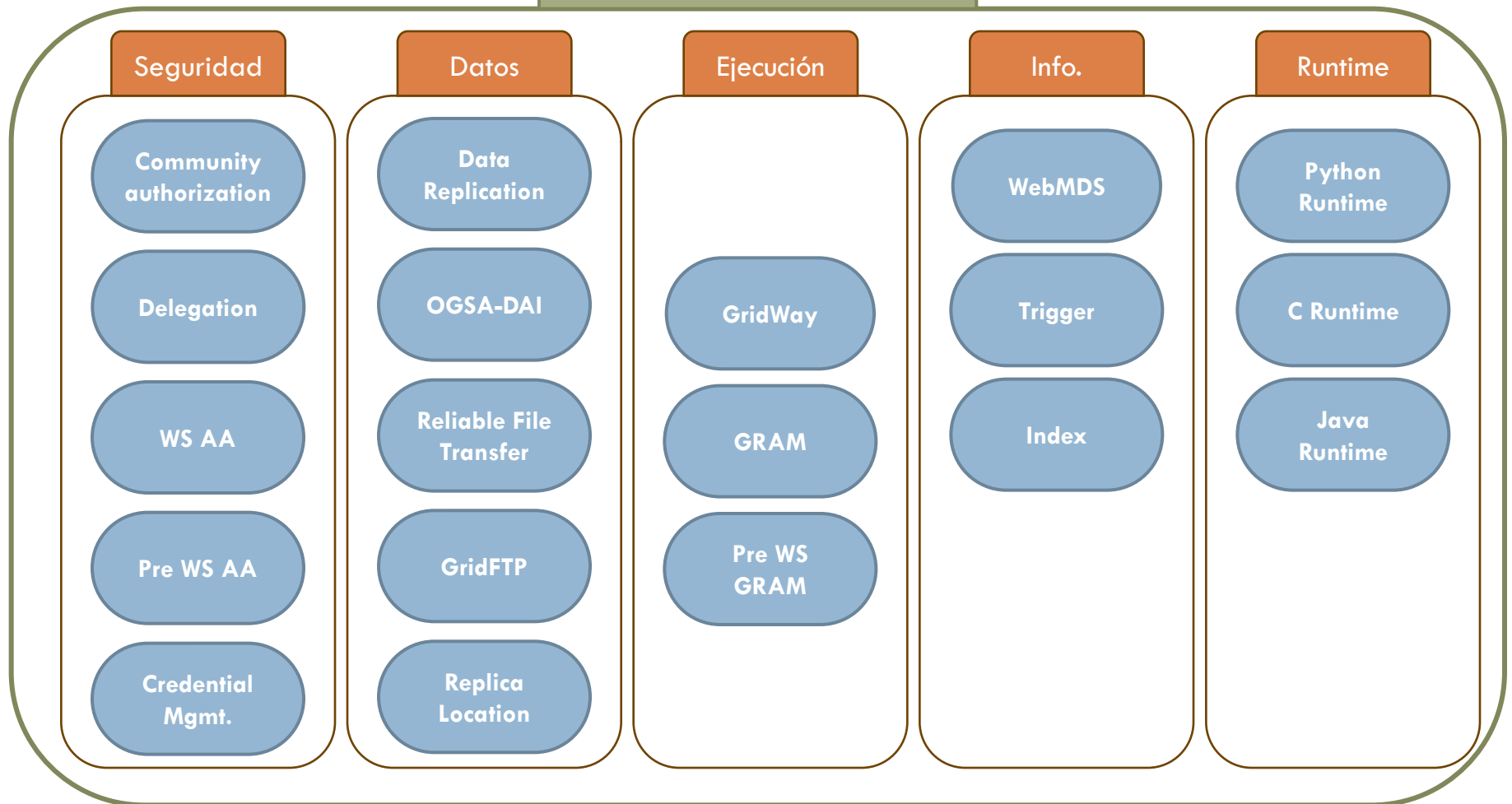
- Conjunto de componentes fundamentales para construir grids

□ Funcionalidad que ofrece:

- Gestión de la seguridad
- Gestión de trabajos
- Gestión de datos
- Gestión de la información

Globus

Componentes Globus



- Origen y desarrollo
 - ▣ El proyecto comenzó en el año 1996
 - ▣ Desarrollado por la Globus Alliance
 - Liderado por la Universidad de Chicago
 - ▣ Muy utilizado:
 - En grids
 - Integrado en otros middleware
 - ▣ Implementa un conjunto de estándares OGSA, OGSF
 - WSRF, WSDL, SOAP, JSDL

- Acceso y licencias
 - ▣ <http://www.globus.org/toolkit/>
 - ▣ Software libre bajo la Apache License version 2 (tipo BSD)
 - ▣ Disponible para:
 - Linux, Solaris, Mac OS X, HPUX, AIX...
 - También hay una versión para Windows aunque sólo soporta la parte implementada en java

□ Publicaciones:

■ Globus: A metacomputing infrastructure toolkit

- Foster I., Kesselman C., *International Journal of High Performance Computing Applications*, 1997, 11 (2), pp. 115-128.

■ Globus Toolkit Version 4: Software for Service-Oriented Systems

- I. Foster, *IFIP International Conference on Network and Parallel Computing*, 2006, LNCS 3779, pp 2-13

■ Globus Toolkit 4: Programming Java Services

- Borja Sotomayor, Lisa Childers, *Morgan Kaufmann*, 2005



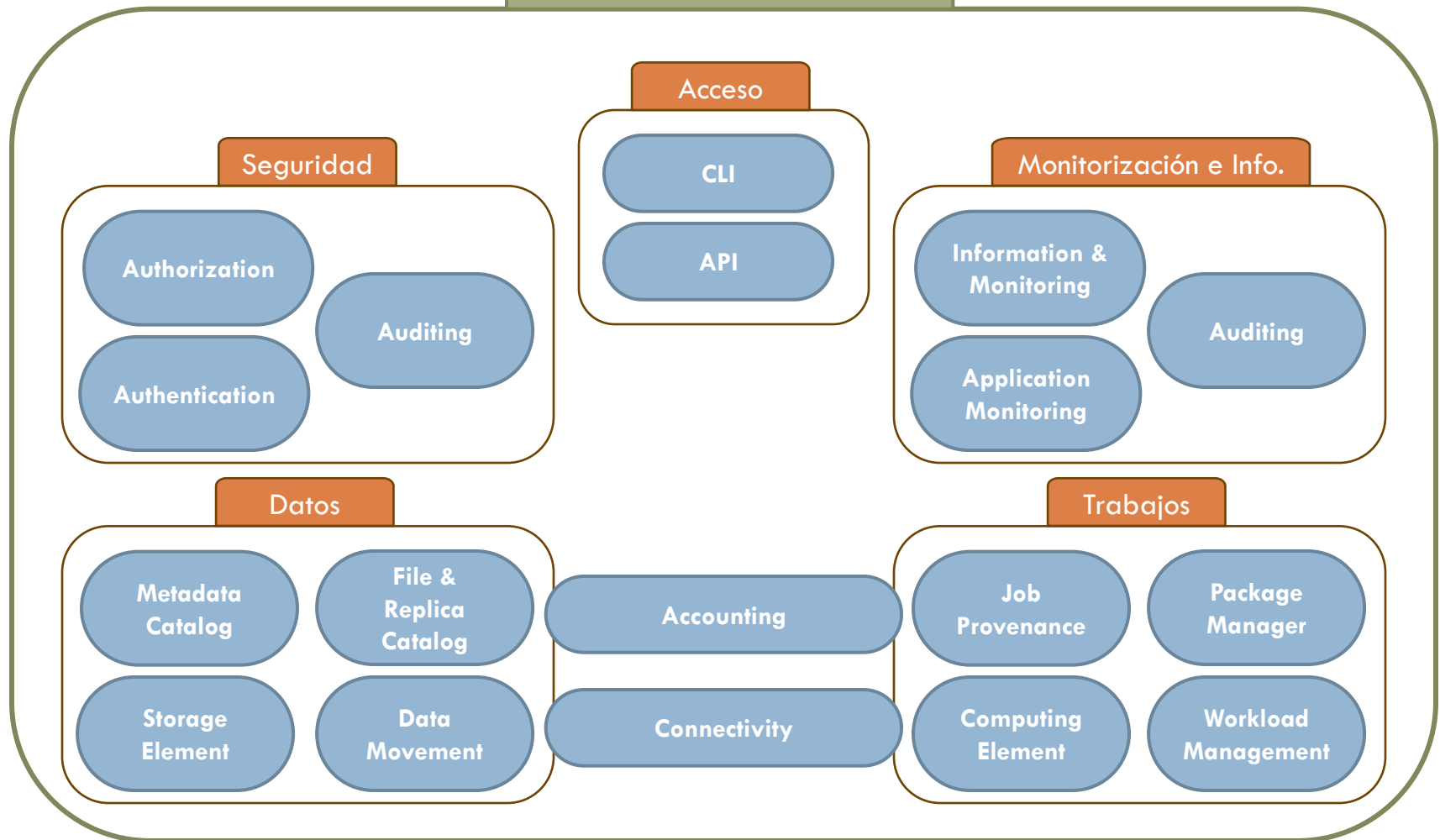
Plataformas

gLite



- gLite:
 - ▣ De forma simplificada:
 - Es el middleware desarrollado para dar soporte al proyecto EGEE (Enabling Grids for E-sciencE)
 - ▣ Funcionalidad que ofrece:
 - Gestión de datos
 - Gestión de trabajos
 - Gestión de acceso
 - Gestión de la seguridad
 - Gestión de la información

Servicios gLite



- Origen y desarrollo
 - ▣ El proyecto comenzó en el año 2004
 - Basado parcialmente en el LCG (LHC (Large Hadron Collider) Computing Grid) y en otros proyectos existentes
 - Primera versión en 2006
 - ▣ Desarrollado por múltiples organizaciones alrededor del mundo
 - ▣ Instalado en 120 organizaciones europeas, 47 países, 100000 CPUs,
 - ▣ Da servicio a 8000 usuarios con 150000 trabajos al día

- Acceso y licencias
 - ▣ <http://glite.web.cern.ch/glite/>
 - ▣ Software libre bajo la licencia Apache License 2.0 (compatible con GPL 3)
 - ▣ Disponible para:
 - Scientific Linux: Distribución de Linux basada en Red Hat desarrollada en el Fermi National Accelerator Laboratory y el Organisation Européenne pour la Recherche Nucléaire (CERN)

□ Publicaciones:

□ gLite 3.1 User Guide

- Stephen Burke, Simone Campana, Elisa Lanciotti, Patricia Mendez Lorenzo, Vincenzo Miccio, Christopher Nater, Roberto Santinelli, Andrea Sciaba, 2009.

□ The glite workload management system

- Marco, C. Fabio, C. Alvise, D. Antonia, G. Francesco, G. Alessandro, M. Moreno, M. Salvator M., Fabrizio, P. , Luca, P., Lecture Notes in Computer Science, 2009, Volume 5529, 2009, Pages 256-268



Plataformas

UNICORE

- UNICORE (Uniform Interface to Computing Resources):



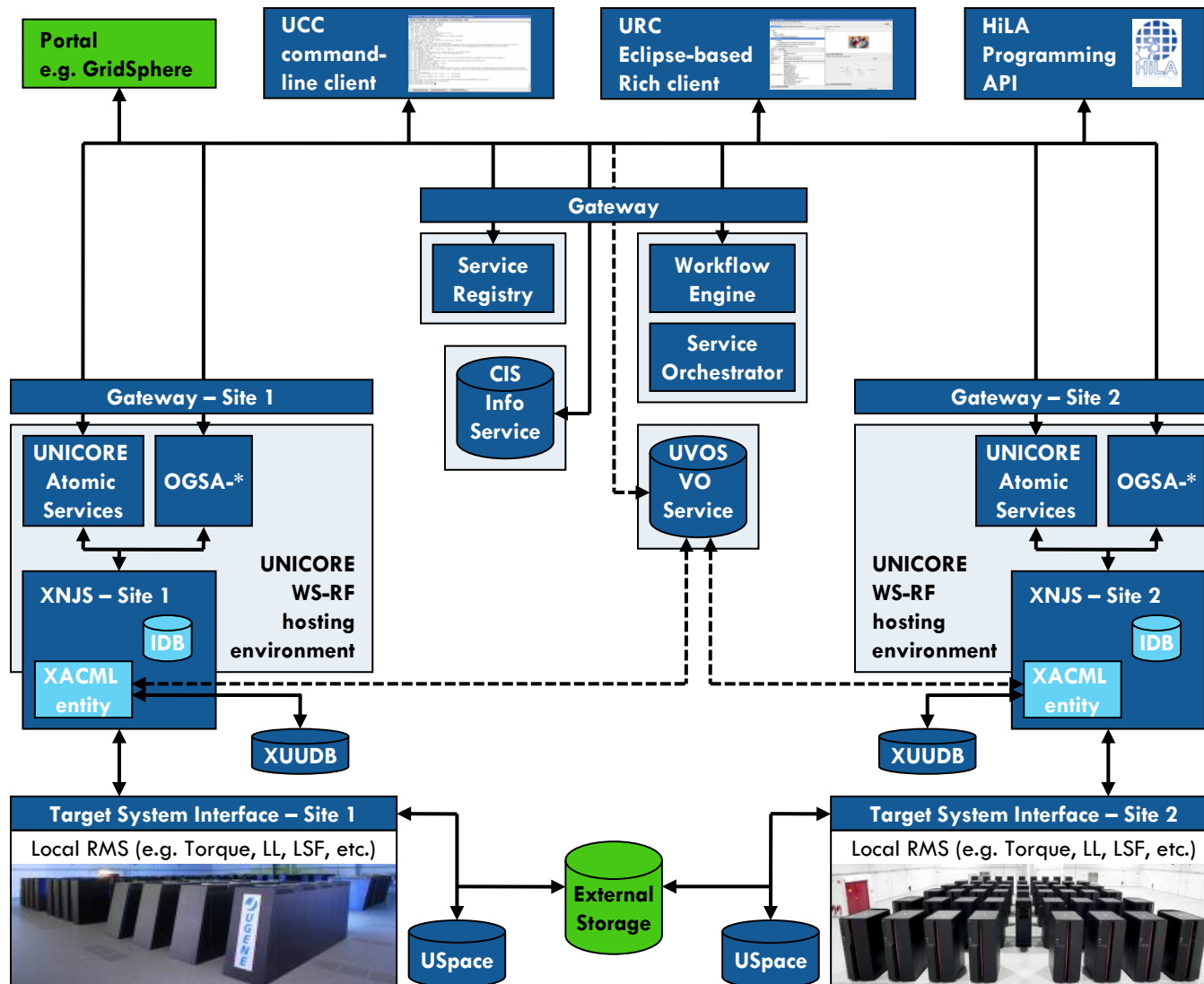
- De forma simplificada:

- Middleware de fácil uso para conectar LRMS (Local Resource Management System)

- Funcionalidad que ofrece:

- Gestión de seguridad
 - Gestión de trabajos
 - Gestión de workflows
 - Interfaces de usuario para facilitar su uso

UNICORE



<http://www.unicore.eu>

- Origen y desarrollo
 - ▣ Desarrollado desde 1997 en dos proyectos del Ministerio Alemán de Educación e Investigación
 - Para interconectar a los centros de Supercomputación Alemanes, como alternativa a Globus Toolkit
 - ▣ Su evolución ha continuado en proyectos europeos desde 2002
 - Software de código abierto desde 2004
 - ▣ Implementa un conjunto de estándares OGF, OASIS e ITF (desde la versión 6)
 - Open Grid Services Architecture (OGSA) y Web Services Resource Framework (WS-RF 1.2)

- Acceso y licencias
 - <http://www.unicore.eu/unicore/>
 - Software libre bajo la licencia BSD (alojado en SourceForge)
 - Disponible para:
 - Multiplataforma (está hecho en java)
 - LiveCD para probarlo

□ Publicaciones:

▣ UNICORE: uniform access to supercomputing as an element of electronic commerce

- Almond, J. and Snelling, D., *Future Generation Computer Systems*, 1999, Volume 15, Pages 539-548.

▣ The UNICORE grid infrastructure

- Romberg, M., *Scientific Programming*, 2002, Volume 10 (2), Pages 149-157.

▣ UNICORE-a grid computing environment

- Erwin, D. W., *Concurrency Computation Practice and Experience*, Volume 14, Issue 13-15, 2002, Pages 1395-1410



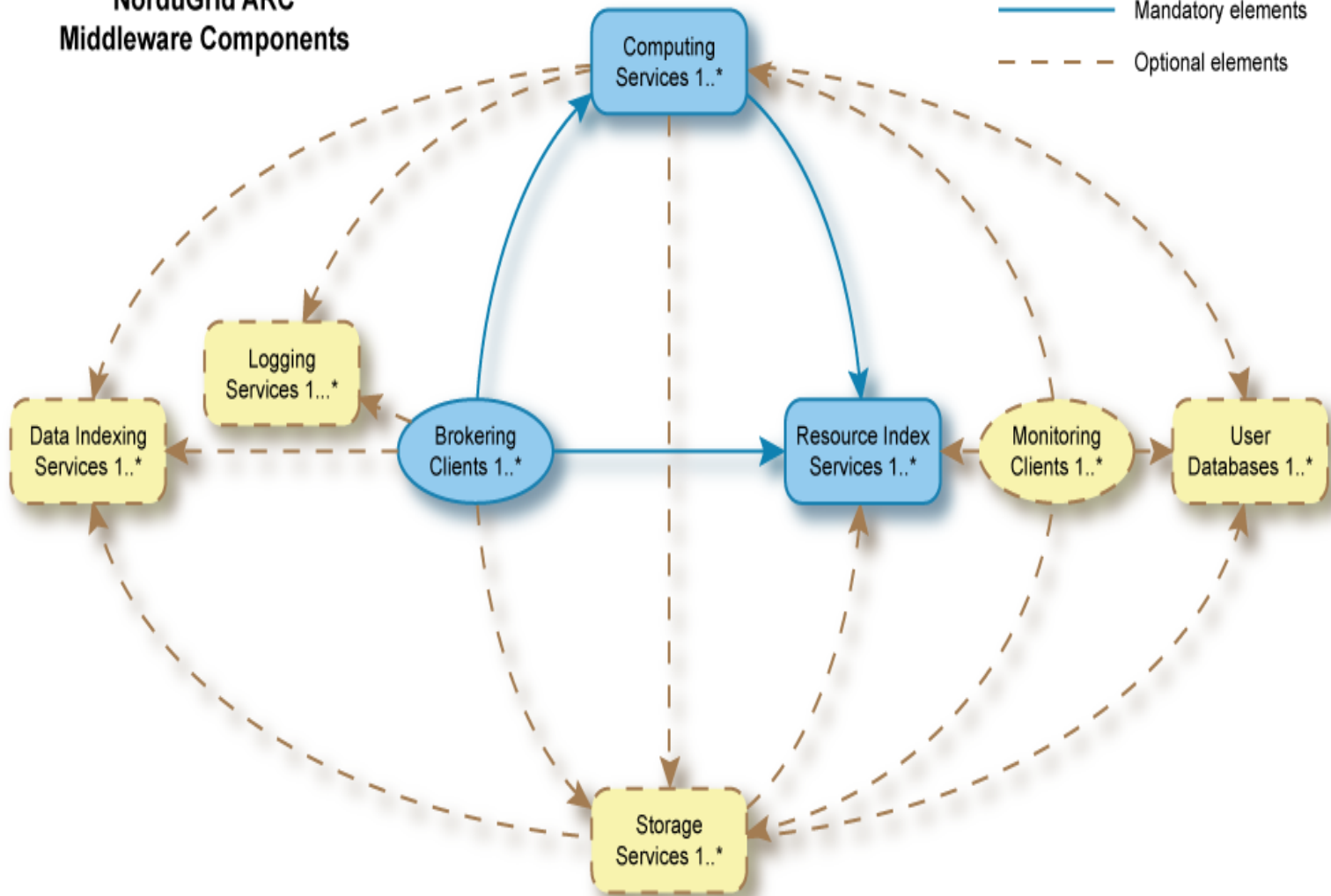
Plataformas

ARC



- ARC (Advanced Resource Connector):
 - ▣ De forma simplificada:
 - Middleware ligero desarrollado por NorduGrid
 - ▣ Funcionalidad que ofrece:
 - Gestión de trabajos
 - Gestión de información
 - Broker de recursos en el cliente

NorduGrid ARC Middleware Components



<http://www.knowarc.eu>

□ Origen y desarrollo

- Desarrollado desde 2001 en el consorcio NorduGrid
 - Para interconectar los centros de Supercomputación nórdicos
 - Primera versión pública en 2002
- Se basa en varios proyectos existentes
 - OpenLDAP, OpenSSL, SASL and Globus Toolkit (GT2)
- KnowARC
 - Nueva versión de ARC en desarrollo a través de un proyecto europeo desde 2006 para extender y adaptar ARC a los estándares grid

- Acceso y licencias
 - ▣ <http://www.nordugrid.org/middleware/>
 - ▣ Software libre bajo la licencia Apache v2.0
 - ▣ Disponible para:
 - Múltiples distribuciones de Linux
- Publicaciones:
 - ▣ Advanced Resource Connector middleware for lightweight computational Grids
 - Ellert, M. Grønager, M. Konstantinov, A. Kónya ,B. Lindemann, J. Livenson, I. Nielsen, J. L. Niinimäki, M. Smirnova, O. Wäänänen, A., Future Generation Computer Systems, 2007, Volume 23 (2), Pages 219-240.



Plataformas

Condor



□ Condor:

▣ De forma simplificada:

- Un sistema de gestión de carga computacional a varios niveles

▣ Funcionalidad que ofrece:

- Gestión de recursos
- Gestión de trabajos
- Políticas de planificación
- Sistema de prioridades
- Monitorización de recursos y trabajos

- Origen y desarrollo
 - ▣ El proyecto comenzó en el año 1988 y se sigue desarrollando de forma activa
 - ▣ Desarrollado por la Universidad de Wisconsin-Madison
 - ▣ Actualmente trabajan en el proyecto 35 personas a tiempo completo
 - ▣ Instalado en la actualidad en $\sim 1/4$ de millón de máquinas

- Acceso y licencias
 - ▣ <http://www.cs.wisc.edu/condor>
 - ▣ Software libre bajo la licencia Apache License 2.0 (compatible con GPL 3)
 - ▣ Disponible para:
 - Linux, Solaris, Mac OS X, HPUX, AIX...
 - También hay una versión para Windows aunque recortada (no permite la migración de trabajos)

- **Publicaciones:**
 - **Condor - a hunter of idle workstations**
 - Litzkow, Michael J. and Livny, Miron and Mutka, Matt W., International Conference on Distributed Computing Systems, Volume 8, 1988, Pages 104-111.
 - **Condor-G: A computation management agent for multi-institutional Grids**
 - Frey, J. and Tannenbaum, T. and Livny, M. and Foster, I. and Tuecke, S., IEEE International Symposium on High Performance Distributed Computing, 2001, Pages 55-66.
 - **Distributed computing in practice: The Condor experience**
 - Thain, D. and Tannenbaum, T. and Livny, M., Concurrency Computation Practice and Experience, Volume 17, Issue 2-4, February 2005, Pages 323-356



Plataformas

BOINC

- BOINC (Berkeley Open Infrastructure for Network Computing):
 - De forma simplificada:
 - Middleware para crear grids basados en Volunteer computing
 - También se puede utilizar para construir los denominados grids de escritorio (ejemplo: jarifa)
 - Funcionalidad que ofrece:
 - Arquitectura cliente-servidor
 - El servidor reparte trabajo a los clientes



- Origen y desarrollo
 - ▣ El proyecto comenzó en el año 2002
 - El software de SETI@HOME era específico para ese proyecto y crearon BOINC como un software genérico que arreglase los problemas de seguridad
 - ▣ Desarrollado por la Universidad Berkeley
 - ▣ Formado por una media de 300.00 voluntarios, 500.000 ordenadores proporcionando 2 petaFlops (FLOP = FLoating point Operations Per Second)
 - Superior al IBM Roadrunner

- Acceso y licencias
 - ▣ <http://boinc.berkeley.edu/>
 - ▣ Software libre bajo la licencia LGPL
 - ▣ Disponible para:
 - Cliente: Windows, Mac, o Unix
 - Servidor: Unix
- Publicaciones:
 - ▣ BOINC: A system for public-resource computing and storage
 - Anderson, D.P., Proceedings - IEEE/ACM International Workshop on Grid Computing, 2004, Pages 4-10