



Open Science Grid and its Interoperation with the EGEE

Ruth Pordes
Fermilab
May 1st 2006



Open Science Grid



Outline



- **Open Science Grid**
- **Common Services with EGEE**
 - Authentication
 - VO Management
 - Data Transfer
 - Information System
 - Job Execution
- **Interoperation with EGEE**
 - Authorization
 - Data Management
 - Job Scheduling/ Workload Management
 - Validation and Monitoring
- **Campus Regional and Federating Grids:**

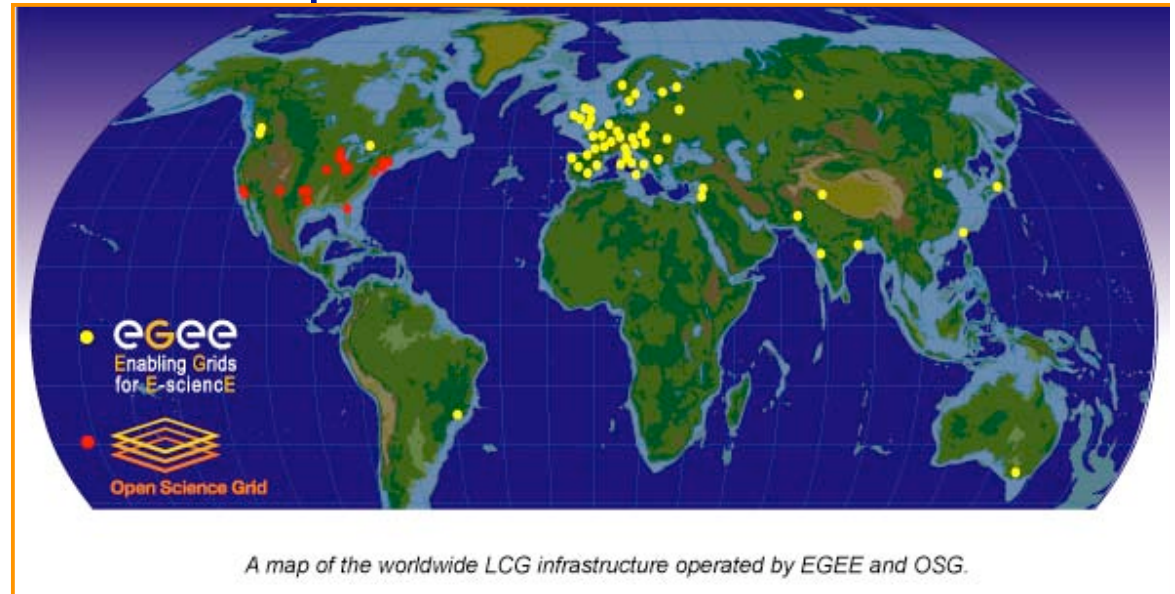
Open Science Grid



A Grid Infrastructure contributing to the World Wide LHC Computing Grid and the LHC Experiments.

Peers with EGEE
TeraGrid.

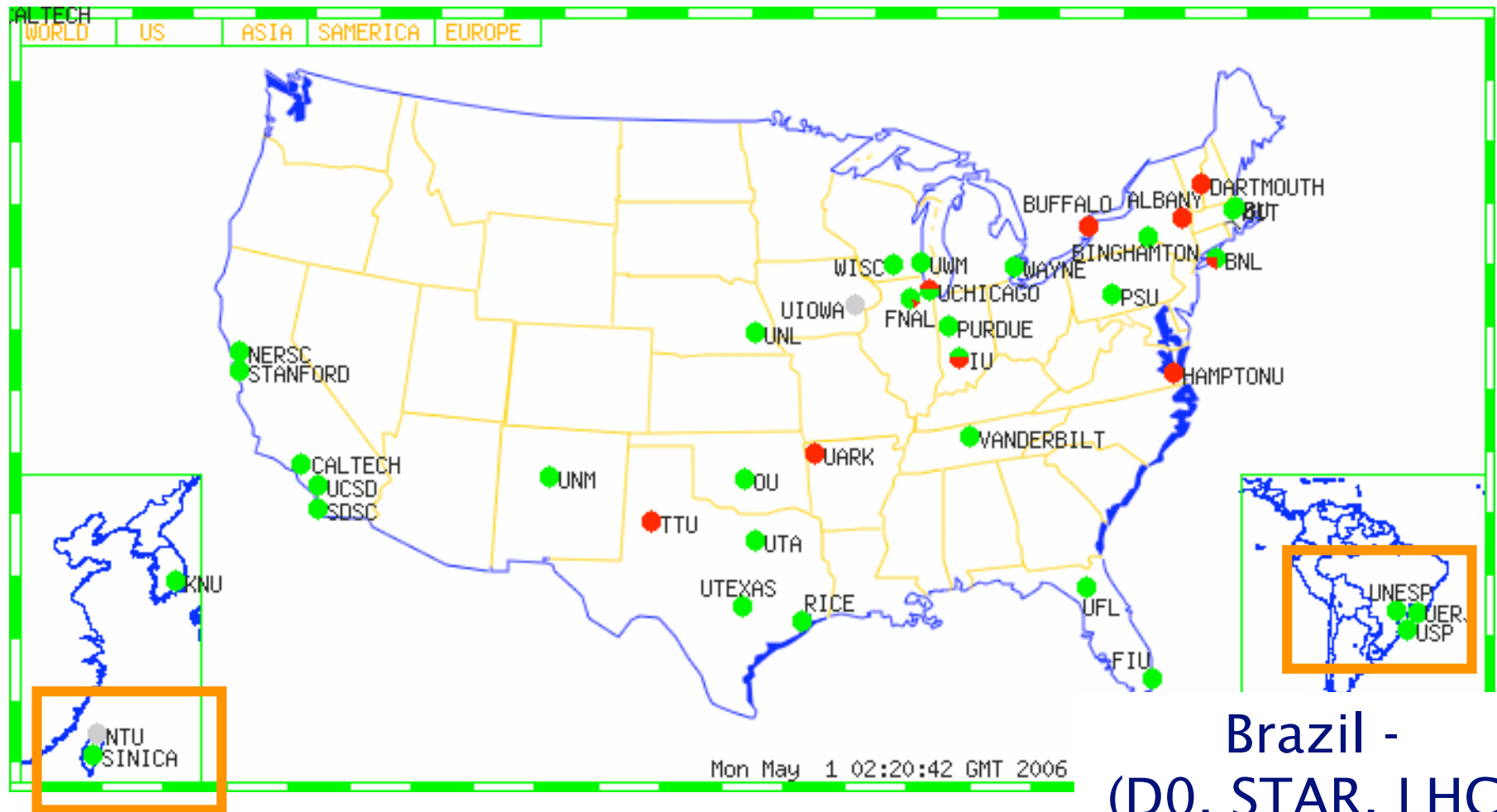
Federates with
Campus Grids



A High Throughput OPEN Distributed Facility

- Shared opportunistic access to existing clusters, storage and networks.
- Owner controlled resources and usage policies.
- Low threshold of entry for new resources & users.

More than a US Grid



Taiwan - (CDF, LHC)

Brazil -
(D0, STAR, LHC)

OSG at EGEE Tutorial

OSG people have history in Grids:



I am the *Executive Director*: Associate head of Fermilab Computing Division, US CMS Grid Services coordinator, PI of Particle Physics Data GridGrid3.

Miron Livny is Manager of the OSG Distributed Facility: Head of the Condor Project and Virtual Data Toolkit, Coordinator of US federation in EGEE, member of EGEE/gLITE design team.

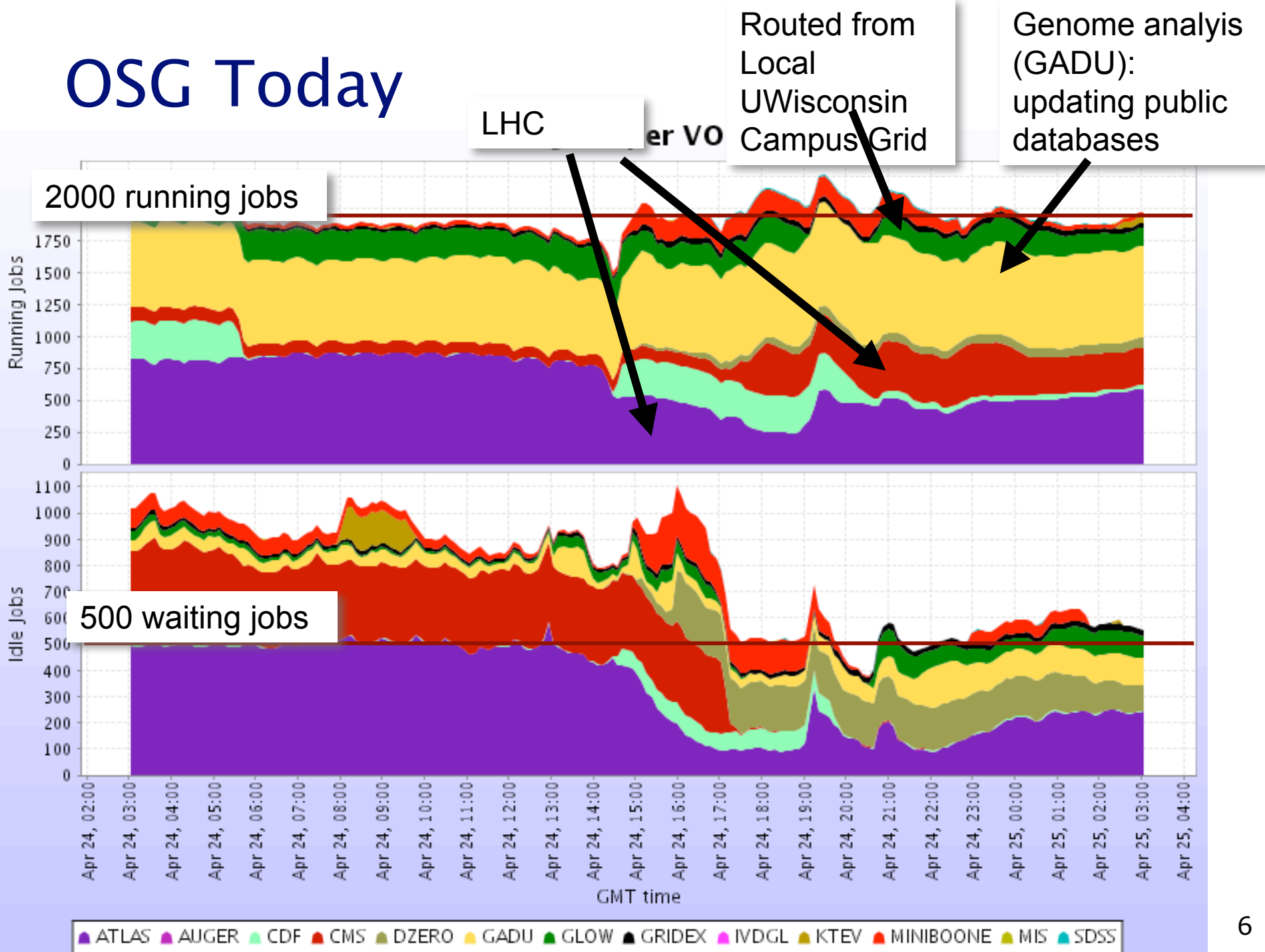
Bill Kramer is the Chair of the Science Council; Head of Lawrence Berkeley National Laboratory NERSC supercomputing facility.

Ian Foster is co-PI of the OSG Proposal: responsible for Globus and Computer Science research contributions and partnerships.

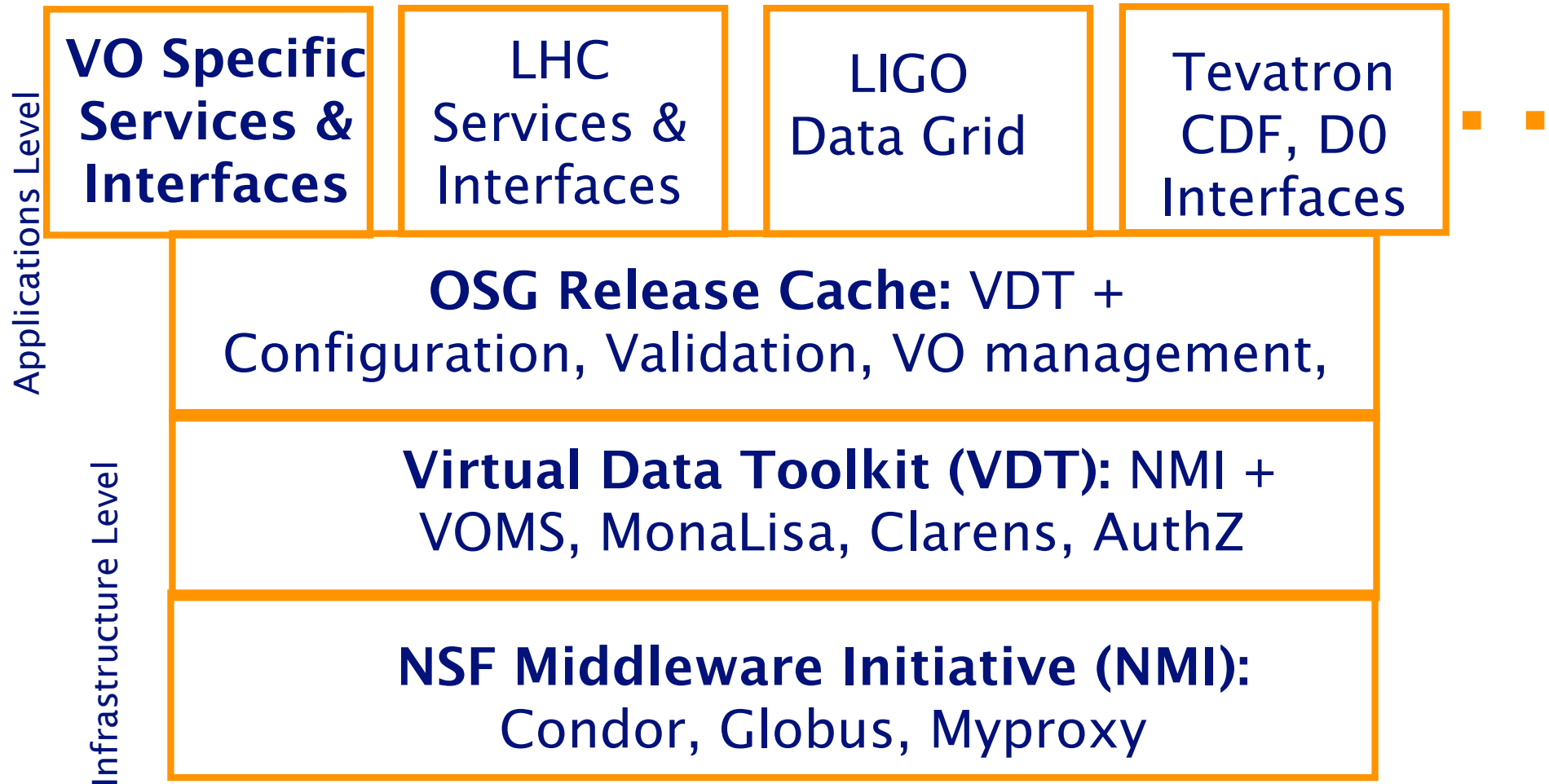
Harvey Newman represents Advanced Network project contributions and collaborations.

Experiment software leadership: US LHC leaders, LIGO, CDF, D0, STAR etc.

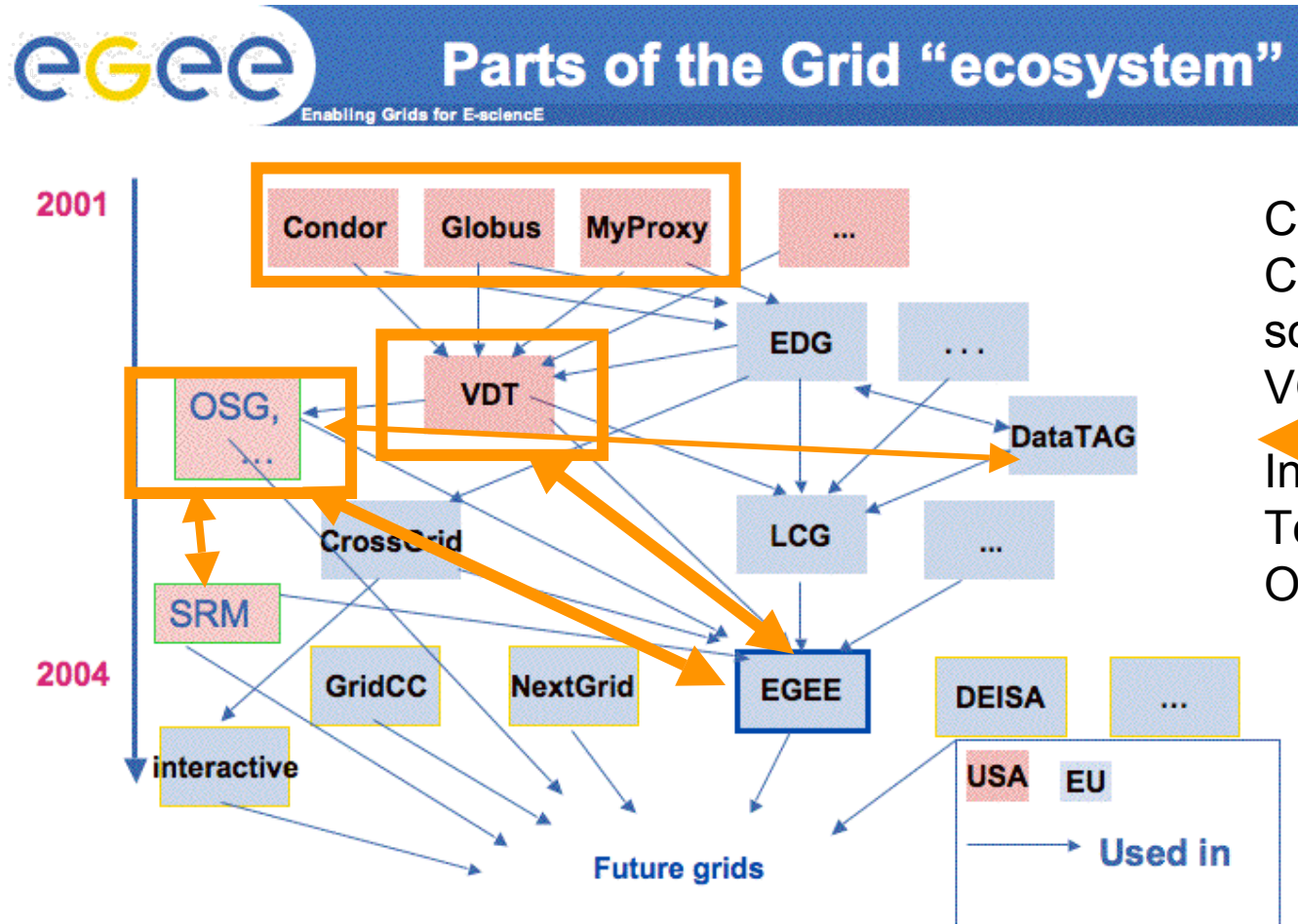
OSG Today



OSG Middleware



Common middleware with EGEE



Commonality and
Co-dependence -
software, users,
VOs

Interoperation -
Technical +
Operations,

Virtual Data Toolkit - a collection of components that integrate into a Distributed System for Researchers. Easy to download, install and use.



Apache HTTPD 2.2.0	jClarens Web Service Registry 0.6.1
Apache Tomcat 5.0.28	JobMon 0.2
Clarens 0.7.2	KX509 20031111
ClassAds 0.9.7	MonALISA 1.4.12
Condor/Condor-G 6.7.18	MyProxy 3.4
DOE and LCG CA Certificates v4 (includes LCG 0.25 CAs)	MySQL 4.1.11
DRM 1.2.10	Nest 0.9.7-pre1
EDG CRL Update 1.2.5	Netlogger 3.2.4
EDG Make Gridmap 2.1.0	PPDG Cert Scripts 1.7
Fault Tolerant Shell (ftsh) 2.0.12	PRIMA Authorization Module 0.3
Generic Information Provider 1.0.15 (Iowa 15-Feb-2006)	PRIMA Authorization Module For GT4 Web Services 0.1.0
gLite CE Monitor (INFN prerelease from 2005-11-15) 1.6.0	pyGlobus gt4.0.1-1.13
Globus Toolkit, pre web-services 4.0.1	pyGridWare gt4.0.1a
Globus Toolkit, web-services 4.0.1	RLS 3.0.041021
GLUE Schema 1.2 draft 7	SRM Tester 1.1
GSI-Enabled OpenSSH 3.6	UberFTP 1.18
GUMS 1.1.0	Virtual Data System 1.4.4
Java SDK 1.4.2_10	VOMS 1.6.10.2
jClarens 0.6.1	VOMS Admin (client 1.2.10, interface 1.0.2, server 1.2.10) 1.2.10-r0
jClarens Discovery Services registration scripts 20060206	

Common with EGEE/WLCG

OSG at EGEE Tutorial

Getting Software into Use on OSG



Domain science requirements.

Globus, Condor,
EGEE etc

OSG stakeholders and middleware
developer (joint) projects.

Test on “VO specific grid”

e.g. LIGO data grid

OSG Integration Grid
Scalable Testbed of
~15 sites.
(peers with EGEE
Preproduction System).

Integrate into VDT Release.
Deploy on OSG integration grid

Include in OSG release &
deploy to OSG production.

OSG Common Services with EGEE

- Users can route data and jobs across Grid boundaries.



Authentication - X509 Certificates, Globus GSI Security infrastructure.

- ❑ Member US TAGPMA registration authority;

VO Management - VOMRS, VOMS.

- ❑ Registration must be trusted and secure since gives users access to resources.
- ❑ Extended attribute certificates for Role based access control.

Data Transfer - GridFTP, SRM storage resource interface;

Information System - Glue Schema, BDII, Globus GT2 MDS

- ❑ Collect information from multiples grids and present uniformaly to VOs and Users.

Job Execution - Condor-G, Globus GRAM, Generic Batch Systems.

OSG Interoperation with EGEE

- Infrastructures or VOs Provide Bridges across Grid boundaries.



Authorization - Use Extended Attribute Certificates to allow VOs to manage access based on Roles.

- Enable Experiments to prioritize Jobs based on type and manage access to Storage based on Role.

Data Management - OSG Common service is Globus Reliable File Transfer and Replica Location Services;

- In practice VOs provide their own services across Grids: LIGO LDR, Run II SAMGRID, CMS PHEDEX, ATLAS DQ, .

Accounting - new service in OSG being deployed now; looking to share probes with EGEE.

- Many questions remain: How to define Value across Organizations? Account for using Storage?

OSG Interoperation with EGEE

- Infrastructures or VOs provide Bridges across Grid boundaries.

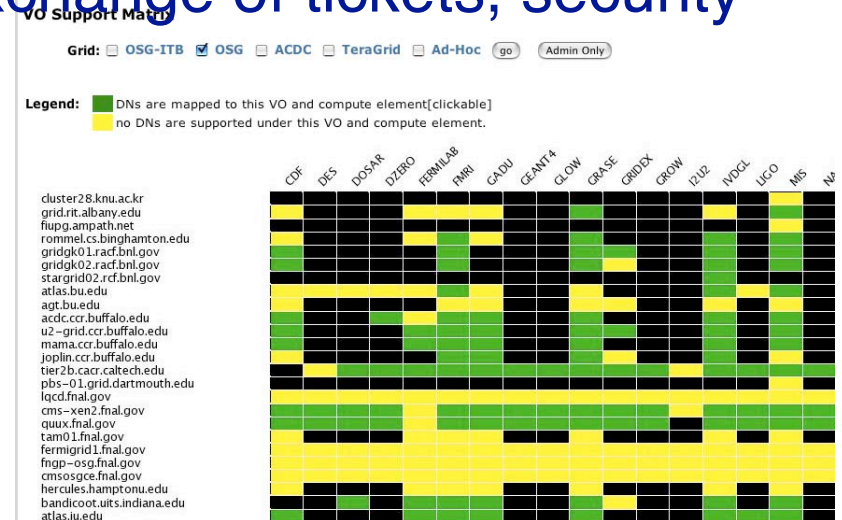


Job Scheduling / Workload Management - “Pull” model and “Glideins” for higher throughput requiring new security components.

Operations - Joint workshops, exchange of tickets, security response etc.

Validation and Monitoring

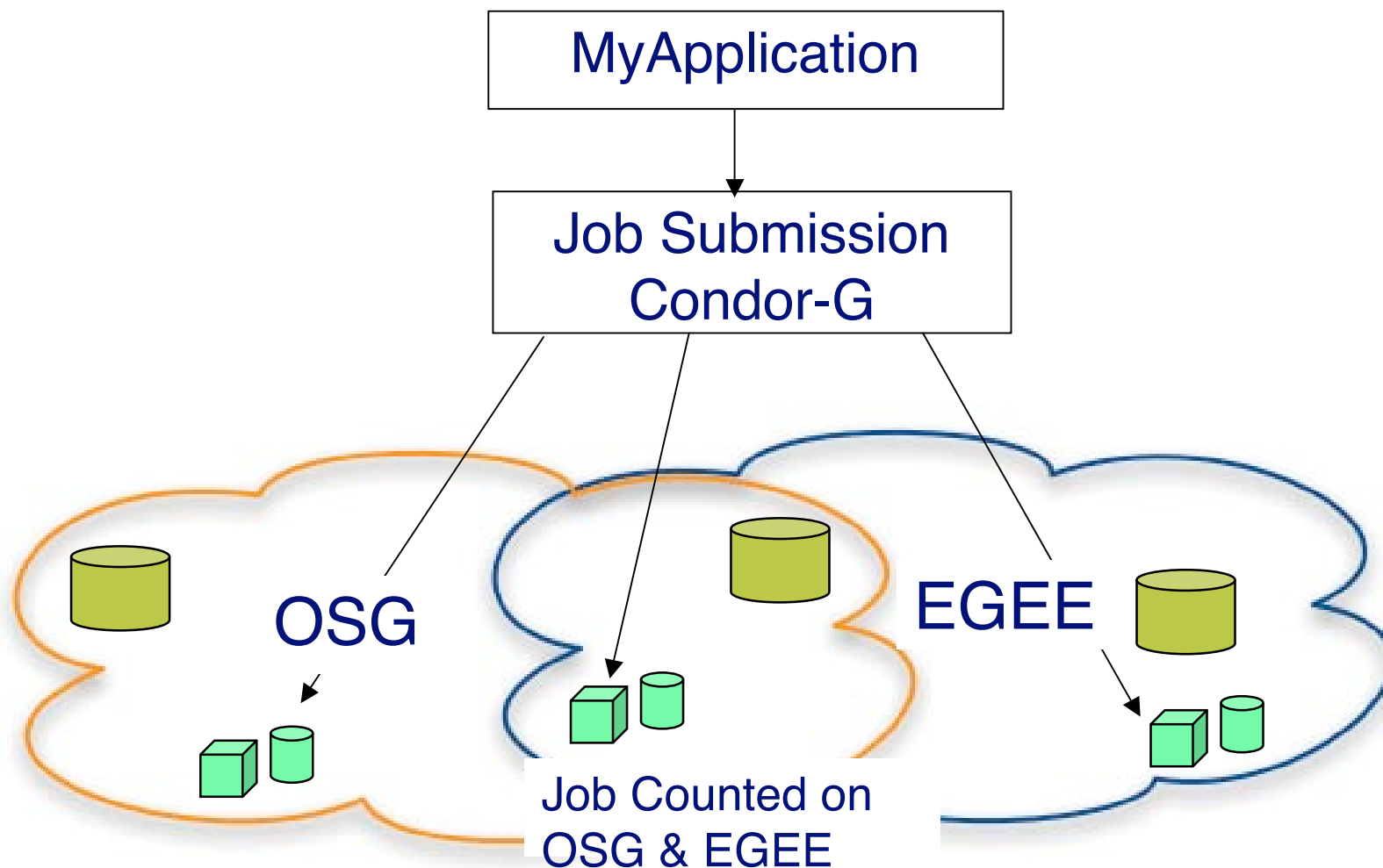
- Grid exerciser, daily validation scripts.



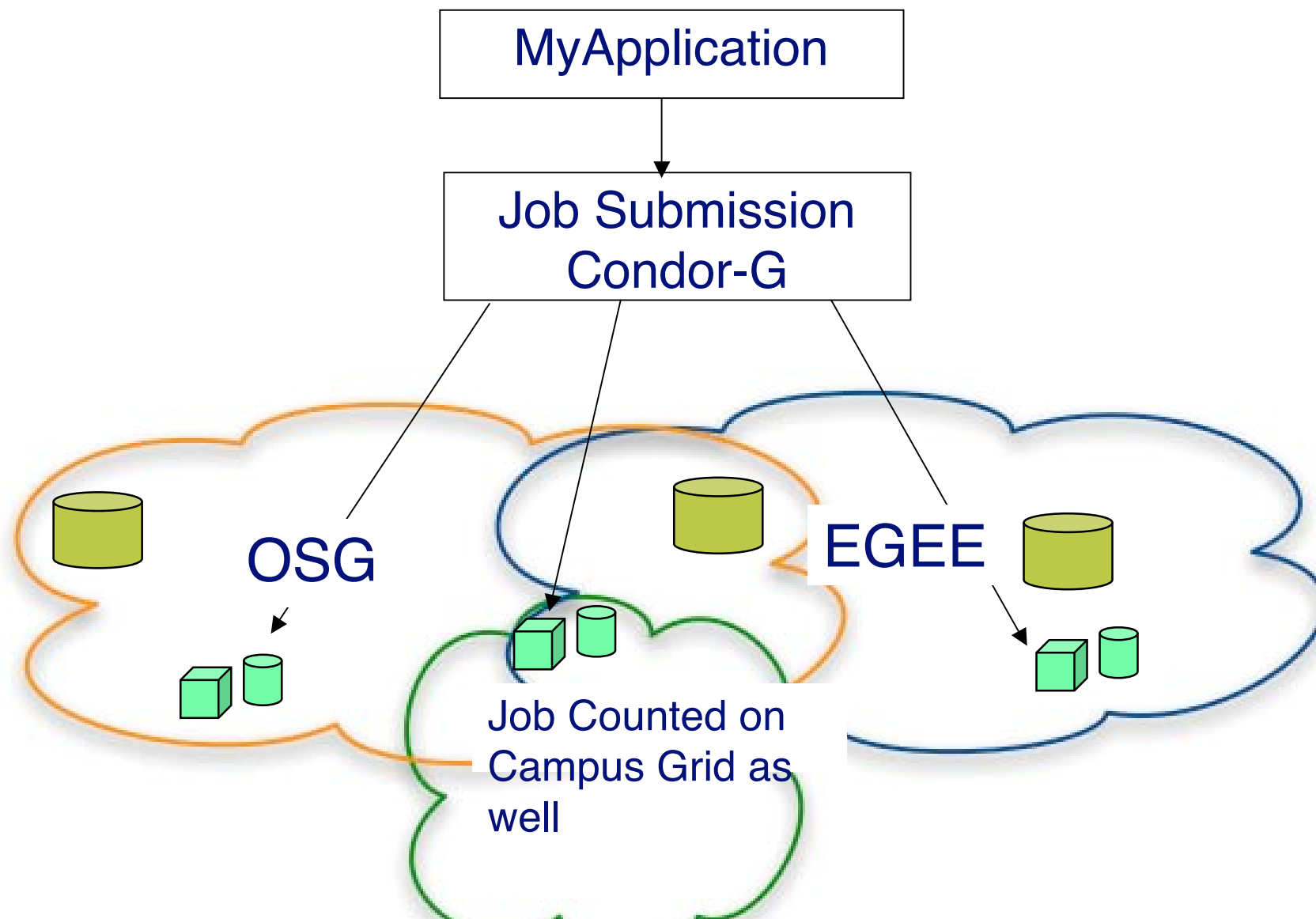
Local VO Environments & Services

- working on XEN Virtual Machine encapsulation

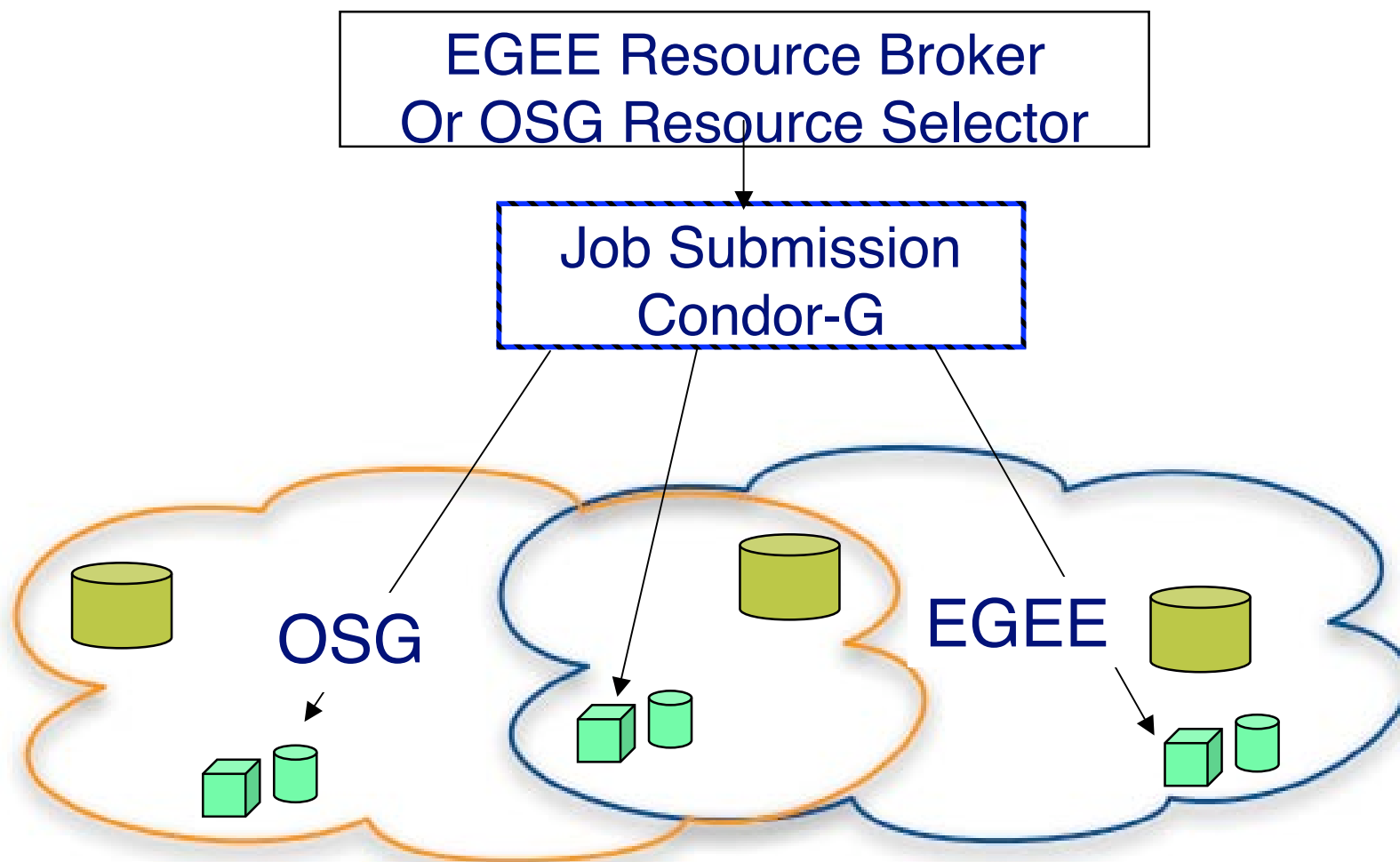
Resources can be on Multiple Grids: What is an OSG Job?



Resources can be on Multiple Grids: What is an OSG Job?



End to end Activity can cross multiple Grids: What is an OSG Job?



OSG Where to find information:



-
- **OSG Web site:** www.opensciencegrid.org
 - **Work in progress:**
<http://osg.ivdgl.org/twiki/bin/view/Integration/OverviewviewGuide>
 - **Virtual Data Toolkit:**
<http://vdt.cs.wisc.edu//index.html>
 - **News about Grids in Science in “Science Grid This Week”:** www.interactions.org/sgtw