The OSG – a VO Centric VO

Miron Livny
Computer Sciences Department
University of Wisconsin-Madison
miron@cs.wisc.edu

Hierarchical resource management in the polder metacomputing initiative

Volume 24 , Issue 12-13 (November 1998)

Special issue on applications
Pages: 1807 - 1825 Year of Publication: 1998
ISSN:0167-8191

Authors A. W. van Halderen B. J. Overeinder P. M. A. Sloot R. van Dantzig D. H. J. Epema M. Livny

I am not sure what it really means but …

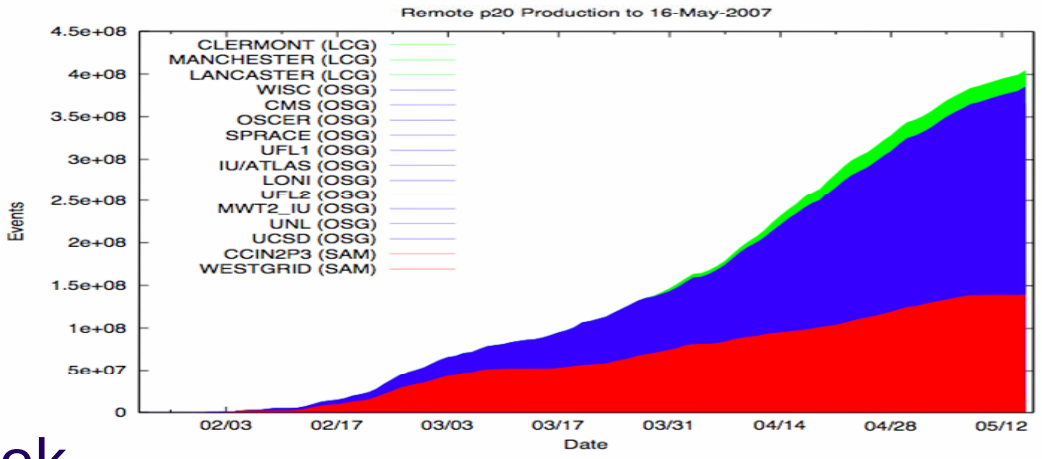
The blue print for the Open Science Grid (OSG) that was developed four years ago states that "*The OSG architecture is Virtual Organization based*". A VO is considered as party to contracts between Resource Providers & VOs which govern resource usage & policies and may consist of sub-VOs which operate under the contracts of the parent.

Supporting the D0 VO

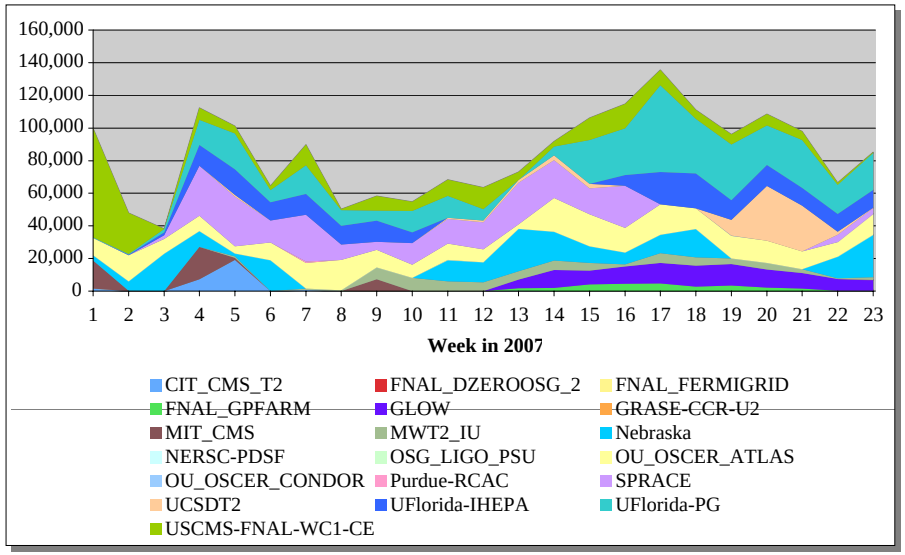
Open Science Grid

12 sites contributed up to 1000 jobs/day

Total Events



OSG CPUHours/Week



2M CPU hours
286M events
286K Jobs on OSG
48TB Input data
22TB Output data

The Open Science Grid vision

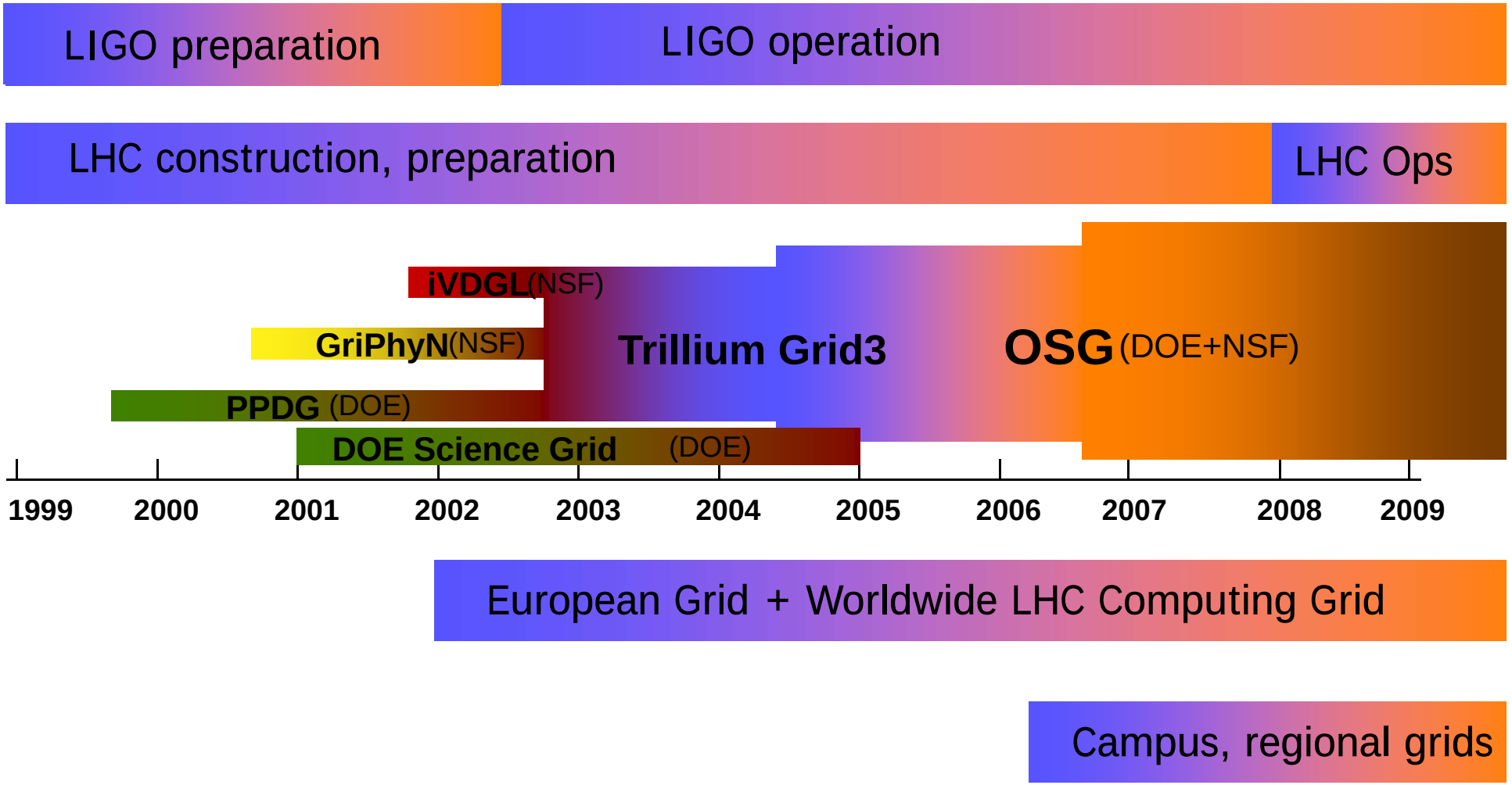
Transform processing and **data** intensive science through a cross-domain self-managed **national distributed** cyber-infrastructure that **brings together** campus and community infrastructure and facilitating the needs of **Virtual Organizations** (VO) at all scales

Three Building Blocks

The OSG organization, management and operation is structured around three components:

- the **Consortium**
- the **Project**
- the **Facility**

The Evolution of the OSG



OSG challenges

- Develop the **organizational** and management structure of a consortium that drives such a Cyber Infrastructure
- Develop the **organizational** and management structure for the project that builds, operates and evolves such Cyber Infrastructure
- Maintain and evolve a **software stack** capable of offering powerful and dependable capabilities that meet the science objectives of the NSF and DOE scientific communities
- Operate and evolve a dependable and well managed **distributed facility**

The OSG Consortium

- **> 20 Scientific Virtual Organizations:** LHC, STAR, LIGO, NanoHub etc.
- **> 25 Resource Providers:** DOE National Labs, University Facilities etc.
  - 10 Storage focused resources
- **>10 Software Providers (including External Projects):** Condor, Globus, Storage Resource Manager, Internet2, ESNET, CEDPS, Fermilab Accounting etc.
- **>4 Partners - Ex-Officio:** EGEE, TeraGrid, NWICG, TIGRE, APAC etc.

Open Science Grid

The Project

- Annual Project Plan (including WBS) gives details of deliverables and timeline for the year.
  - Deliverables driven by the science stakeholders.
  - Buy-in through “Science Milestones” deliverables - owned by the stakeholders and included in the plan.
  - Area Coordinators responsible for milestones and deliverables under their branch.
  - Well defined Software releases part of the plan.
- WBS updated by Area Coordinators quarterly. Missed milestones subject to discussion.
- Adjust plans based on experience, requests, feedback and problems.
- Change request process for project deliverables

OSG Middleware Evolution

Domain science requirements

Condor, Globus, Privilege, EGEE, …

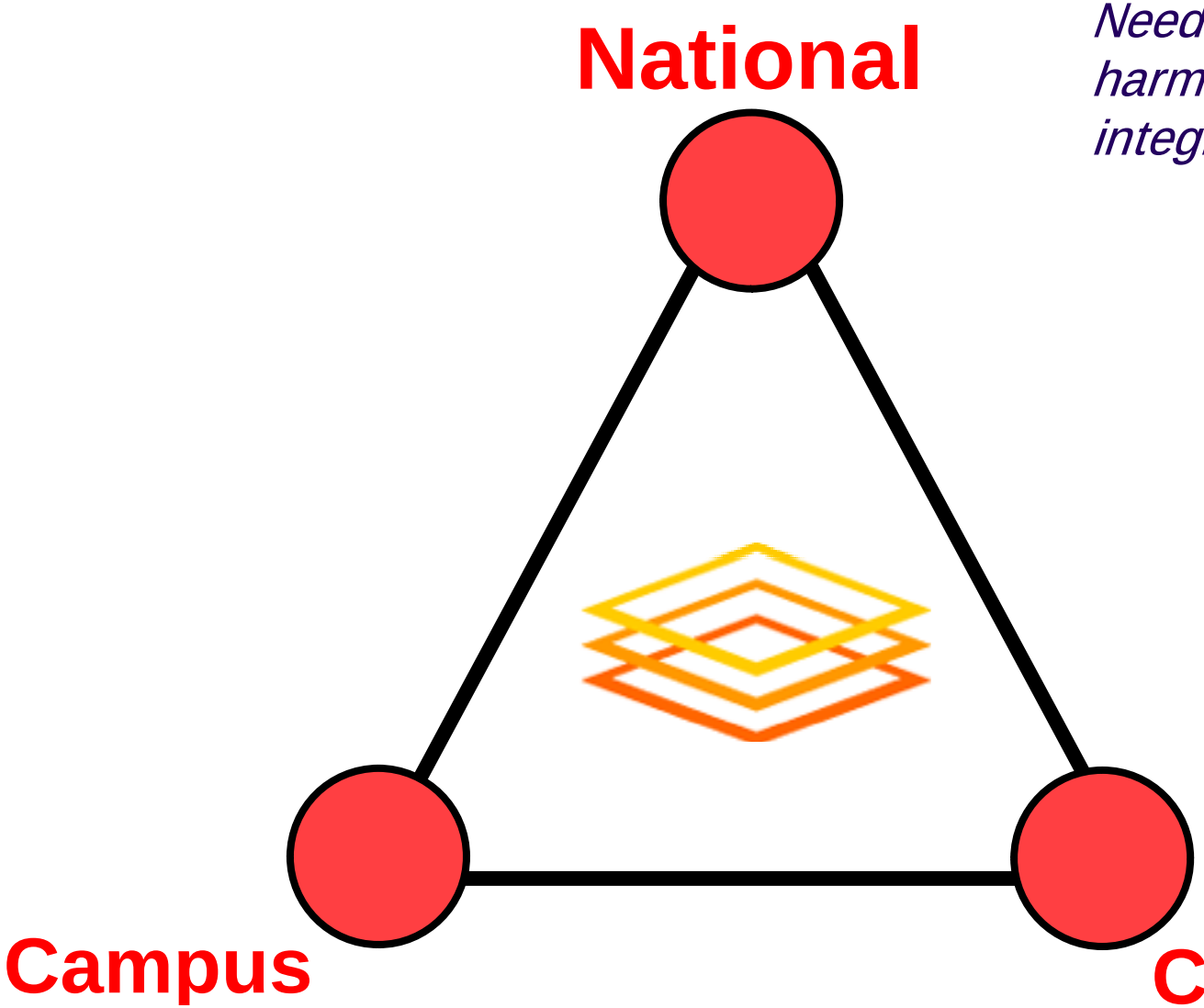
OSG stakeholders and middleware developer (joint) projects.

Test on "VO specific grid"

Integrate into VDT Release.
Deploy on OSG integration grid
Interoperability testing

Provision in OSG release & deploy on OSG sites.

The Three Cornerstones



Need to be harmonized into a well integrated whole.

OSG is VO Centric

- OSG brings together many VOs for
  - opportunistic sharing of resources in a grid environment
  - allowing for more effective use of their collective resources
  - Allowing easier use of dedicated/allocated resources that are distributed
- A Virtual Organization (VO) is a collection of people (VO members),
  - A VO's member structure may include groups, subgroups and/or roles into which it divides its members according to their responsibilities and tasks, such that they are accorded appropriate levels of authorization.
  - In order to receive the appropriate authorization at another VO's site, a user's grid job must be able to present an authentication token along with a token indicating the desired computing privileges.
- A Site is a collection computing/storage resources (sites) and services (e.g., databases). and the terms "Site," "Computing Element" ("CE"), and/or "Storage Element" ("SE") to refer to the resources owned and operated by a VO or other organization.
  - Use of the resources at a site is determined by a combination of the site's local policies and the user VO's policies. VOs are responsible for contracting individually with each other for guaranteed access to resources.
  - Groups the provide software are also known as resource providers

A Terminology Zoo

- Organizations
- Virtual Organizations
- Users/people/VO members
- Groups of users
- Roles
- Resources
- Sites
- Institutions
- Software providers
- Certificate Authorities
- Campus Grids
- National Grids
- Facility
- Consortium
- Project

What will it take to turn most (all?) of these entities in to Virtual Organizations?

Some very basic questions …

- What is an organization and what makes one virtual?
- What is the scope, authority and leagal standing of a VO?
- What kind of inter-VO relationships can we define?
- How are we going to develop the tools we need to support the (intra and inter) management and operations of VOs?