The Open Science Grid and the Management of Rights

Miron Livny
University of Wisconsin - Madison

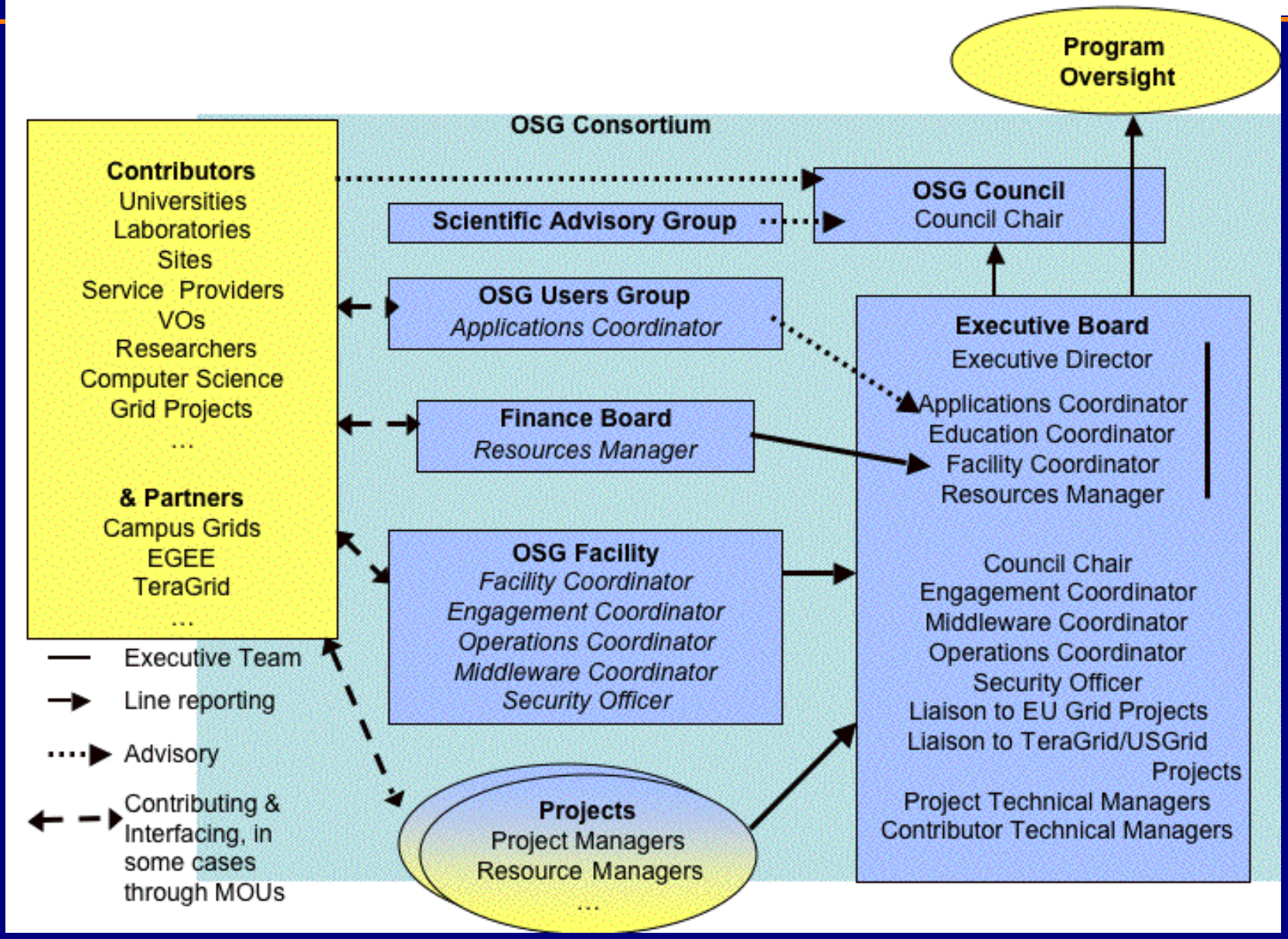
What is OSG?

The Open Science Grid is a US national distributed computing **facility** that supports scientific computing via an **open collaboration** of science researchers, software developers and computing, storage and network providers. The OSG Consortium is building and operating the OSG, bringing resources and researchers from universities and national laboratories together and cooperating with other national and international infrastructures to give scientists from a broad range of disciplines access to shared resources worldwide.

The OSG Project

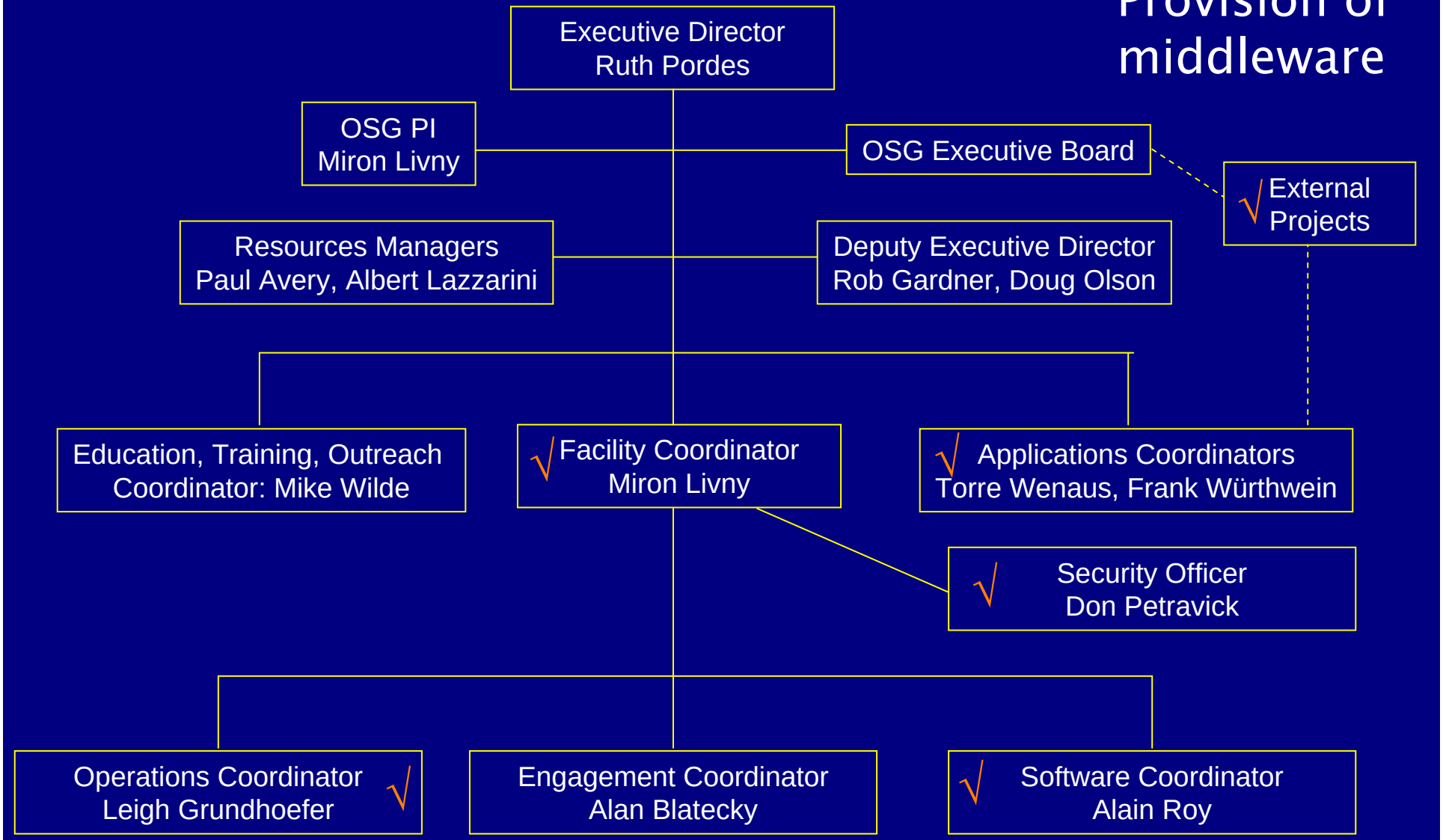
- Co-funded by DOE and NSF at an annual rate of ~$6M for 5 years starting FY-07
- Currently main stakeholders are from physics - US LHC experiments, LIGO, STAR experiment, the Tevatron Run II and Astrophysics experiments
- A mix of DOE-Lab and campus resources
- Active “engagement” effort to add new domains and resource providers to the OSG consortium

OSG Consortium



OSG Project Execution

√ Role includes Provision of middleware



OSG Principles

- Characteristics -
  - Provide guaranteed and opportunistic access to shared resources.
  - Operate a heterogeneous environment both in services available at any site and for any VO, and multiple implementations behind common interfaces.
  - **Interface to Campus and Regional Grids.**
  - **Federate with other national/international Grids.**
  - Support multiple software releases at any one time.
- Drivers -
  - Delivery to the schedule, capacity and capability of LHC and LIGO:
    - Contributions to/from and collaboration with the US ATLAS, US CMS, LIGO software and computing programs.
  - Support for/collaboration with other physics/non-physics communities.
  - Partnerships with other Grids - especially EGEE and TeraGrid.
  - Evolution by deployment of externally developed new services and technologies:.

OSG Middleware Layering

Applications

LIGO Data Grid

CMS Services & Framework

ATLAS Services & Framework

CDF, D0 SamGrid & Framework

Infrastructure

OSG Release Cache: VDT + Configuration, Validation, VO management

Virtual Data Toolkit (VDT) Common Services
NMI + VOMS, CEMon (common EGEE components), MonaLisa, Clarens, AuthZ

NSF Middleware Initiative (NMI):
Condor, Globus, Myproxy

OSG Middleware Deployment

Domain science requirements.

Condor, Globus, Privilege, EGEE etc

OSG stakeholders and middleware developer (joint) projects.

Test on “VO specific grid”

Integrate into VDT Release.
Deploy on OSG integration grid

Provision in OSG release & deploy to OSG production.

Security Infrastructure

- Identity: X509 Certificates
  - OSG is a founding member of the US TAGPMA.
  - DOEGrids provides script utilities for bulk requests of Host certs, CRL checking etc.
  - VDT downloads CA information from IGTF.
- Authentication and Authorization using VOMS extended attribute certficates.
  - DN-> Account mapping done at Site (multiple CEs,SEs) by GUMS.
  - Standard authorization callouts to Prima(CE) and gPlazma(SE).

Security Infrastructure

- Security Process modelled on NIST procedural controls starting from an inventory of the OSG assets:
  - *Management* - Risk assessment, planning, Service auditing and checking
  - *Operational* - Incident response, Awareness and Training, Configuration management,
  - *Technical* - Authentication and Revocation, Auditing and analysis. End to end *trust* in quality of code executed on remote CPU - signatures?

User and VO Management

- VO Registers with with Operations Center
  - Provides URL for VOMS service to be propagated to the sites.
  - Several VOMS are shared with EGEE as part of WLCG.
- User registers through VOMRS or VO administrator
  - User added to VOMS of one or more VOs.
  - VO responsible for users to sign AUP.
  - VO responsible for VOMS service support.
- Site Registers with the Operations Center
  - Signs the Service Agreement.
  - Decides which VOs to support (striving for default admit)
  - Populates GUMS from VOMSes of all VOs. Chooses account UID policy for each VO & role.
- VOs and Sites provide Support Center Contact and joint Operations.
  - For WLCG: US ATLAS and US CMS Tier-1s directly registered to WLCG. Other support centers propagated through OSG GOC to WLCG.

Inter-operability with Campus grids

FermiGrid is an interesting example for the challenges we face when making the resources of a campus (in this case a DOE Laboratory) grid accessible to the OSG community

What is FermiGrid?

- Integrates resources across most (soon all) owners at Fermilab.
- Supports jobs from Fermilab organizations to run on any/all accessible campus FermiGrid and national Open Science Grid resources.
- Supports jobs from OSG to be scheduled onto any/all Fermilab sites,.
- Unified and reliable common interface and services for FermiGrid gateway - including security, job scheduling, user management, and storage.
- More information is available at http://fermigrid.fnal.gov

Job Forwarding and Resource Sharing

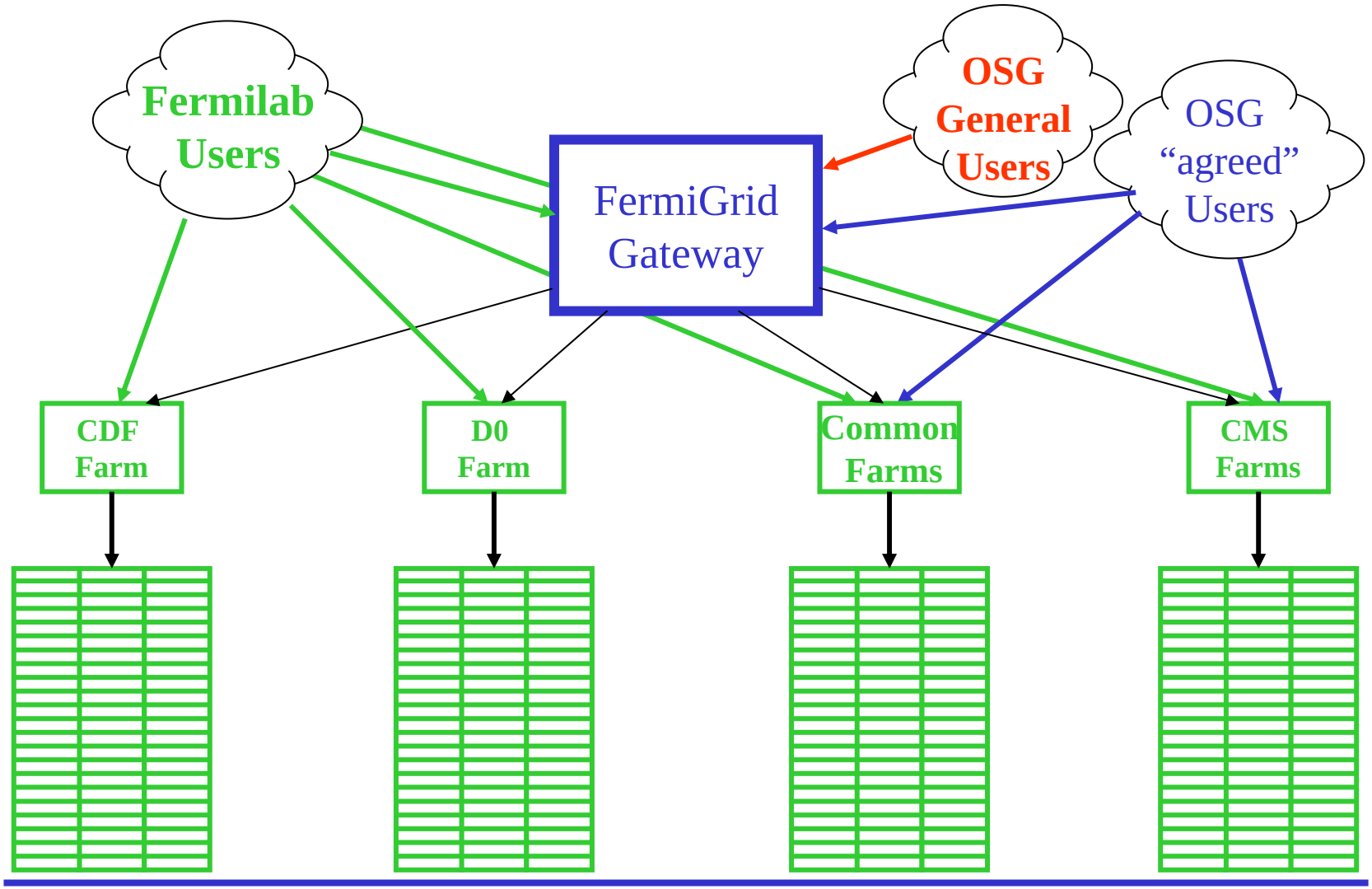
Gateway currently interfaces 5 Clusters with diverse batch and file systems and >1000 Job Slots.

Job scheduling policies and in place agreements for sharing allow fast response to changes in resource needs by Fermilab and OSG users.

Gateway provides single bridge between OSG wide area distributed infrastructure and FermiGrid local sites.

Storage and Job execution policies applied through Site-wide managed security and authorization services.

Access to FermiGrid



Issues

- Heterogeneous identity management systems – EGEE vs. OSG, grid vs. local, compute vs. storage, head-node vs. , old-version vs. new-version …
- Cross domain right management
- Right/identity management of software modules and resources
- Error/rejection propagation
- Solutions/approaches that work end-to-end