

Fermilab's Grid & Cloud Computing Dept. & KISTI: Opportunities for Collaborations

Overview

- FermiGrid
- FermiCloud
- Federation of Resources in Korea
- Fast Networks
- Distributed Storage
- Computing Accounting

Meeting with KISTI

Nov 2, 2011

Gabriele Garzoglio

Grid & Cloud Computing Department, Associate Head

Computing Sector, Fermilab  **Fermilab**



KISTI and GCC: working together...

- Maintaining frequent in-person visits at KISTI and FNAL
- Working side-by-side in the Grid and Cloud Computing department at FNAL
 - Seo-Young: 3 mo in the Summer 2011
 - Hyunwoo Kim: 6 months since Spring 2011
- Sharing information about Grid & Cloud computing and FNAL ITIL service management
- Consulting help on operations for CDF data processing



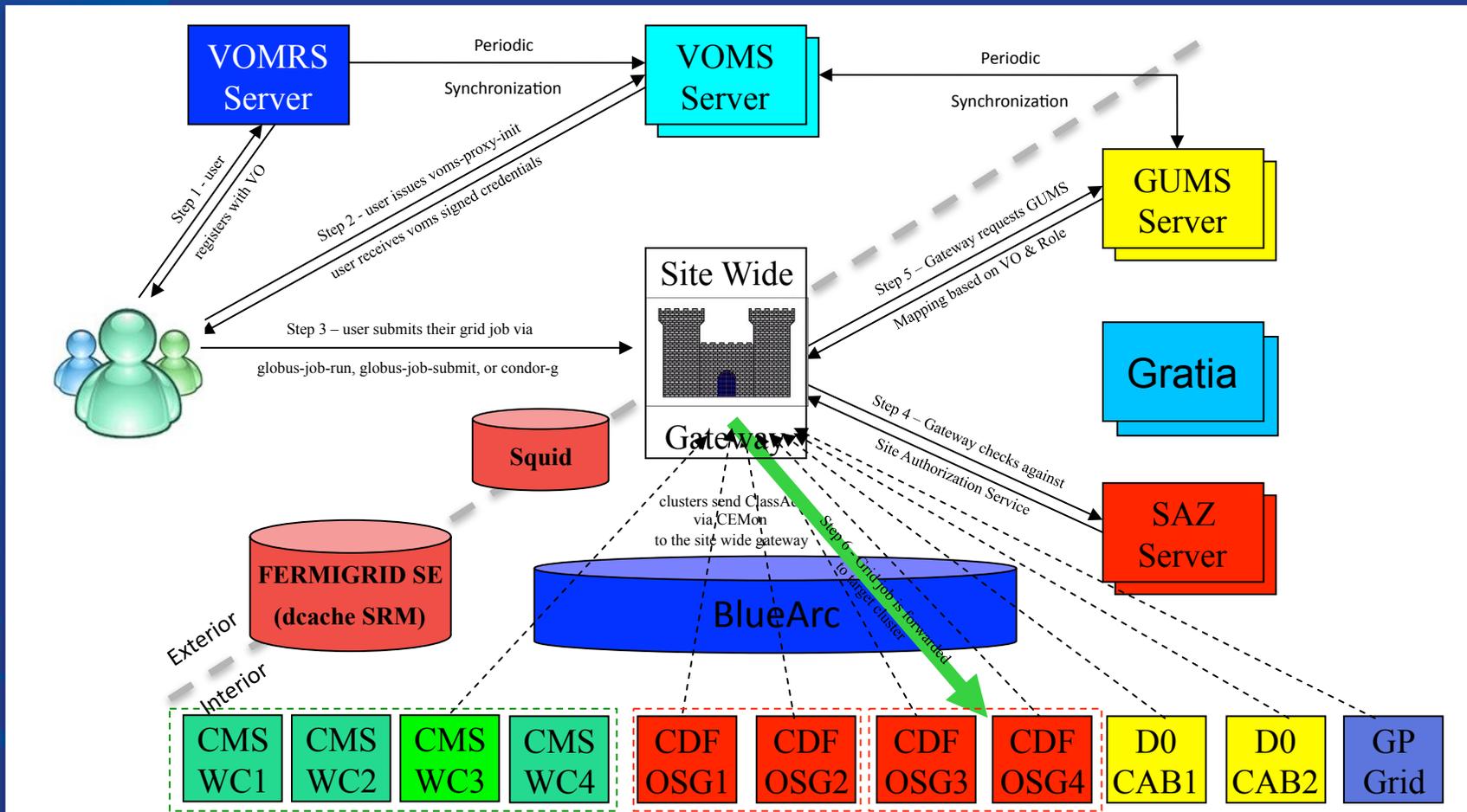
Opportunities for Collaboration

- **FermiGrid**
 - The Fermilab Campus Grid
- **FermiCloud**
 - The Fermilab Infrastructure as a Service
- **Federation of Resources in Korea**
 - Infrastructures for resource sharing
- **Fast Networks**
 - Preparing computing services to interface 100 Gbps links
- **Distributed Storage**
 - Storage Management across Grid sites
- **Computing Accounting**
 - Gratia: Inter-Grid and Cloud accounting

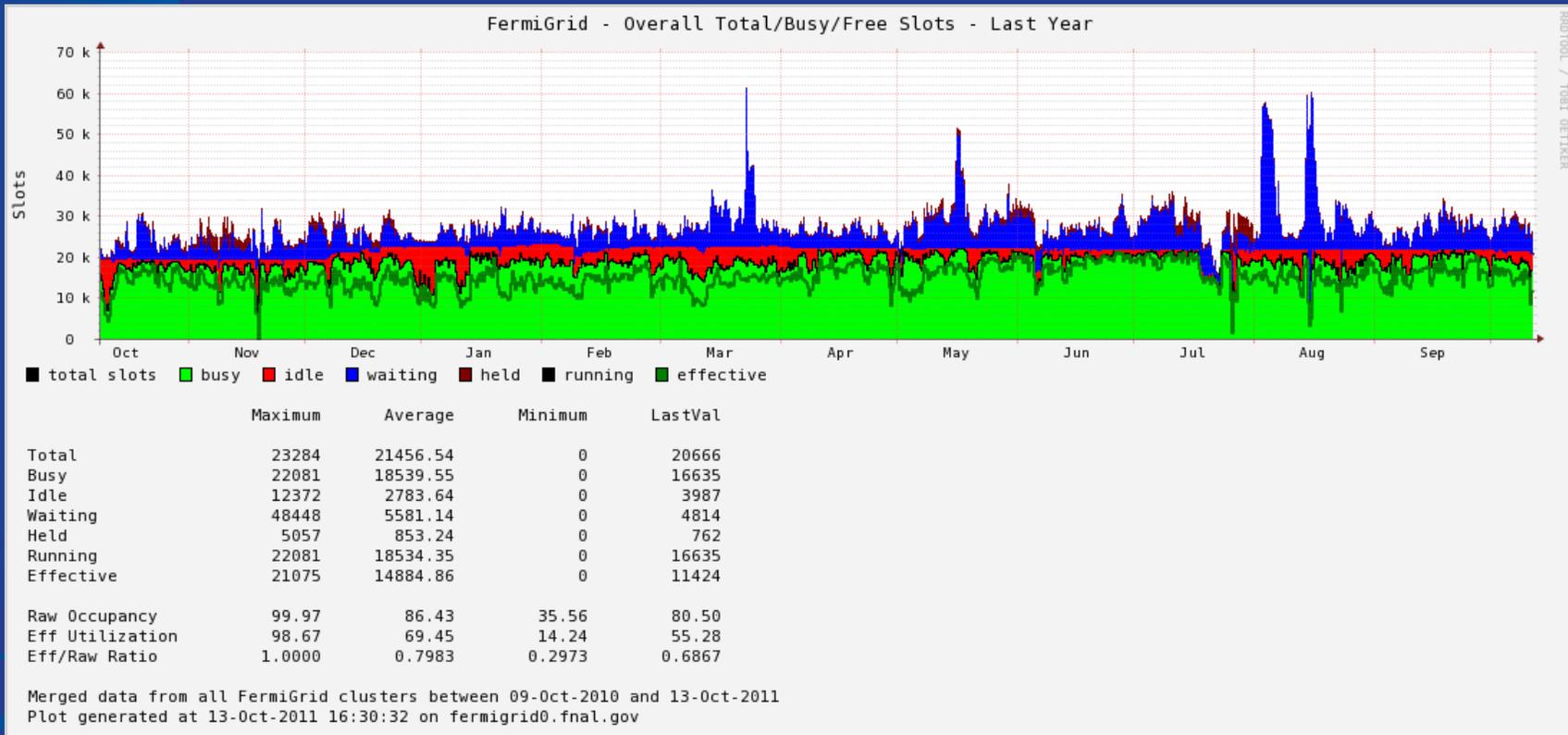
FermiGrid Characteristics

- A Meta-facility that provides Grid infrastructure for scientific computing at Fermilab
- Provides highly-available centralized authorization and authentication services
- Provides Campus Grid gateway for job submission
- Coordinates interoperability among stakeholders
- Provides Grid-enabled mass storage services
- >28,500 CPU-years recorded since Gratia accounting started, most in OSG (as of Oct 2011)
- Currently have about 22,000 batch slots, 10 compute elements

FermiGrid Architecture



Overall Occupancy & Utilization

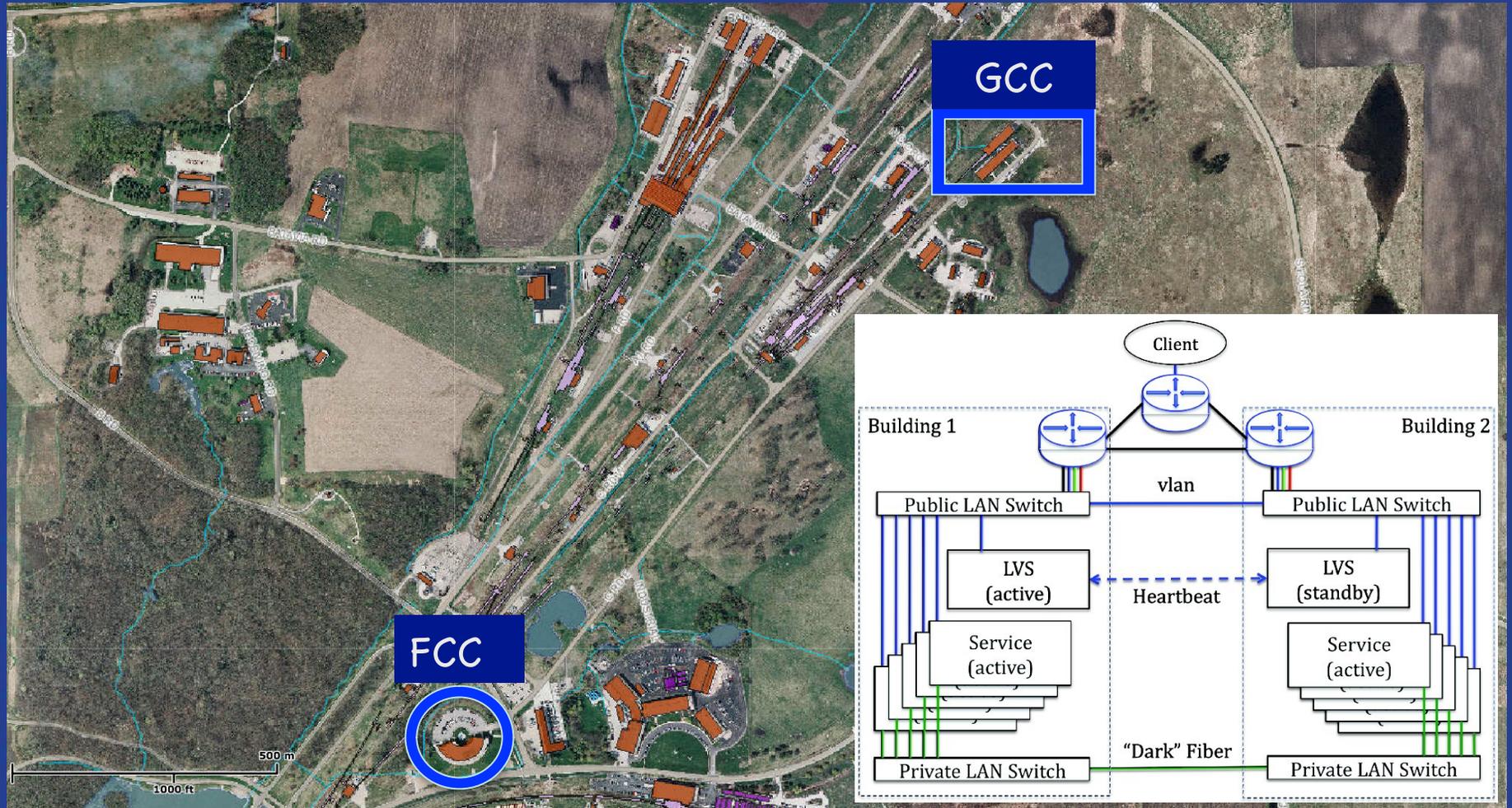


FY12 Focus Areas

- Maintenance and Operations
 - Operate the FermiGrid services at a very high level of reliability.
 - Maintain and extend the Core and OSG Grid Services infrastructure.
 - Maintain and enhance the support for the existing FermiGrid user and stakeholder communities.
 - Provide support for the Laboratory's computer security efforts.
- Complete deployment of FermiGrid-HA2 network features and make services less dependent on NFS service.

We are happy to continue sharing our operational experience as part of this collaboration

FermiGrid-HA2 Network and Geographical Redundancy



FermiGrid-HA2 Service Availability

Service	Raw Availability	HA Configuration	Measured HA Availability	Minutes of Downtime
VOMS – VO Management Service	99.657%	Active-Active	100.000%	0
GUMS – Grid User Mapping Service	99.652%	Active-Active	100.000%	0
SAZ – Site AuthoriZation Service	99.657%	Active-Active	100.000%	0
Squid – Web Cache	99.640%	Active-Active	100.000%	0
MyProxy – Grid Proxy Server	99.954%	Active-Standby	99.954%	240
ReSS – Resource Selection Service	99.635%	Active-Active	100.000%	0
Gratia – Fermilab and OSH Accounting	99.365%	Active-Standby	99.997%	120
Databases	99.765%	Active-Active	99.988%	60

Opportunities for Collaboration

- ✓ **FermiGrid**
 - The Fermilab Campus Grid
- **FermiCloud**
 - **The Fermilab Infrastructure as a Service**
- **Federation of Resources in Korea**
 - Infrastructures for resource sharing
- **Fast Networks**
 - Preparing computing services to interface 100 Gbps links
- **Distributed Storage**
 - Storage Management across Grid sites
- **Computing Accounting**
 - Gratia: Inter-Grid and Cloud accounting

FermiCloud Characteristics

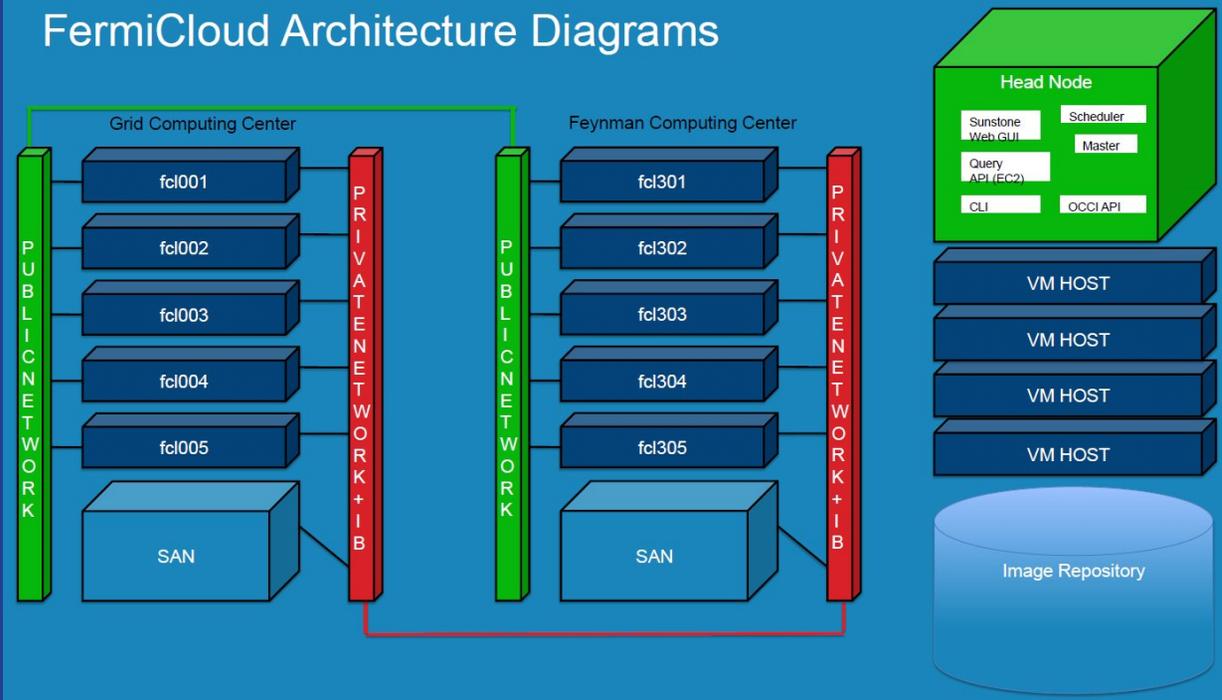
- Infrastructure-as-a-service private cloud for the Fermilab Scientific Program.
- Integrated into the Fermilab site security structure.
- Virtual machines have full access to existing Fermilab network and mass storage devices.
- Scientific stakeholders get on-demand access to virtual machines without system administrator intervention.
- Virtual machines created by users and destroyed or suspended when no longer needed.
- Testbed for developers and integrators to evaluate new grid and storage applications on behalf of scientific stakeholders.
- Ongoing project to build and expand the facility:
 - Phase 1 - Technology evaluation, requirements, deployment,
 - Phase 2 - Scalability, monitoring, accounting, performance improvement,
 - Phase 3 - High availability and reliability.

FermiCloud Hardware and Architecture

- 2x Quad Core Intel Xeon E5640 CPU
- 2 SAS 15K RPM system disk 300GB
- 6x 2TB SATA disk
- LSI 1078 RAID controller
- Infiniband card
- 24GB RAM
- 23 machines total, arrived June 2010
- +25TB Bluearc NAS disk
- Just delivered – 84 TB of Nexsan SAN disk



FermiCloud Architecture Diagrams



FermiCloud Software Technologies

- OS: Scientific Linux 5 & 6
- Hypervisor: Paravirtualized KVM
 - Fully virtualized KVM available as an option.
 - KVM allows sharing of read-only memory sections across multiple VMs with copy on write.
- Cloud Management: OpenNebula
- Modifications to OpenNebula CLI, Query API, GUI to use X.509 authentication to launch virtual machines.
- Secure credential store,
 - All security secrets loaded at boot time only.
- Site-wide patching and vulnerability scanning facilities.

Current Technology Investigations

- Testing storage services with real neutrino experiment codes.
- Evaluate ceph as a FS for the image repository.
- Testing dCache NFS4.1 support with multiple clients in the cloud.
- * Using Infiniband interface to create sandbox for MPI applications.
- * Batch queue look-ahead to create worker node VM's on demand.
- * Submission of multiple worker node VM's / grid cluster in the cloud.
- Idle VM detection and suspension, backfill with worker node VM's.
- * Bulk launching of VMs and interaction with private nets
- Leverage site “network jail” for new virtual machines.
- * IPv6 support.
- Interest in OpenID/SAML assertion-based authentication.
- Design a high-availability solution across buildings
- Interoperability: CERNVM, HEPiX, Glidein WMS, ExTENCI, Desktop

* In Collaboration with KISTI

– Seo-Young and Hyunwoo: Summer 2011

– Seo-Young / KISTI: Summer 2011; Proposed Fall / Winter 2011

– Hyunwoo: Proposed Fall / Winter 2011

Ongoing Software Development

- Accounting and billing
 - Cloud accounting add-ons to Gratia accounting project.
- Monitoring
 - How many machines are idle, are running, who is running them, is everything up that should be up?
- Authorization
 - Apply well-tested and interoperable grid authorization tools to cloud authorization as well.

All of above in collaboration with other projects and standards bodies.

Opportunities for Collaboration

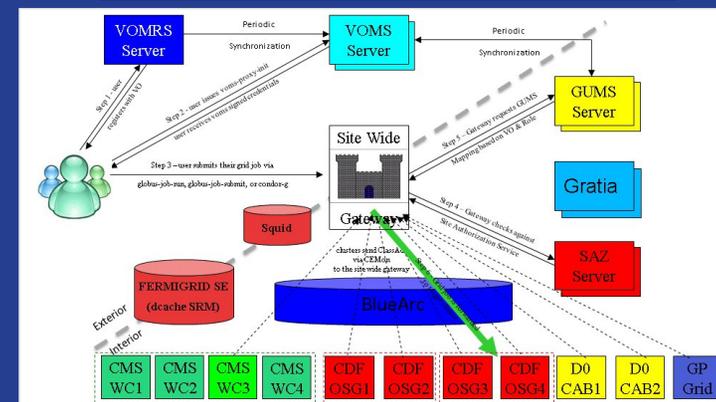
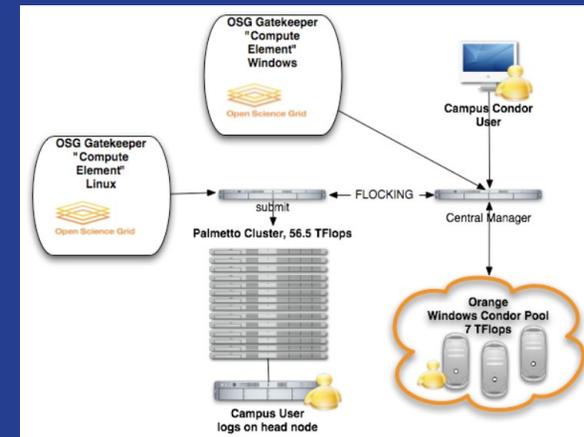
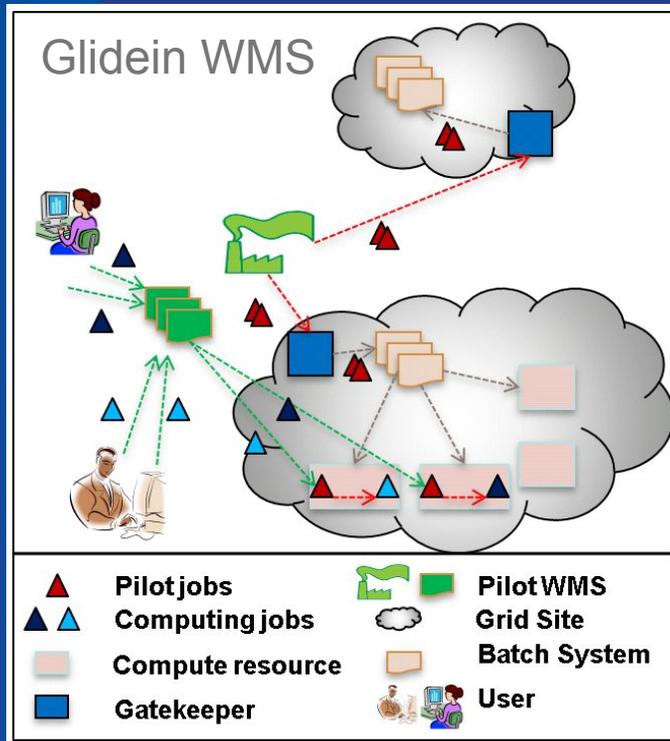
- ✓ **FermiGrid**
 - The Fermilab Campus Grid
- ✓ **FermiCloud**
 - The Fermilab Infrastructure as a Service
- **Federation of Resources in Korea**
 - Infrastructures for resource sharing
- **Fast Networks**
 - Preparing computing services to interface 100 Gbps links
- **Distributed Storage**
 - Storage Management across Grid sites
- **Computing Accounting**
 - Gratia: Inter-Grid and Cloud accounting

Federation of Resources in Korea

- Research institutions often benefit from sharing resources
- Clusters are built with over-provisioned resources to address the needs of the stakeholders
- Resource sharing minimizes idle resources as stakeholders with peak needs can overflow to the federated resources.

Fermilab and OSG are leaders in resource federation

- We offer to consult on configuring and deploying resource federation solutions.

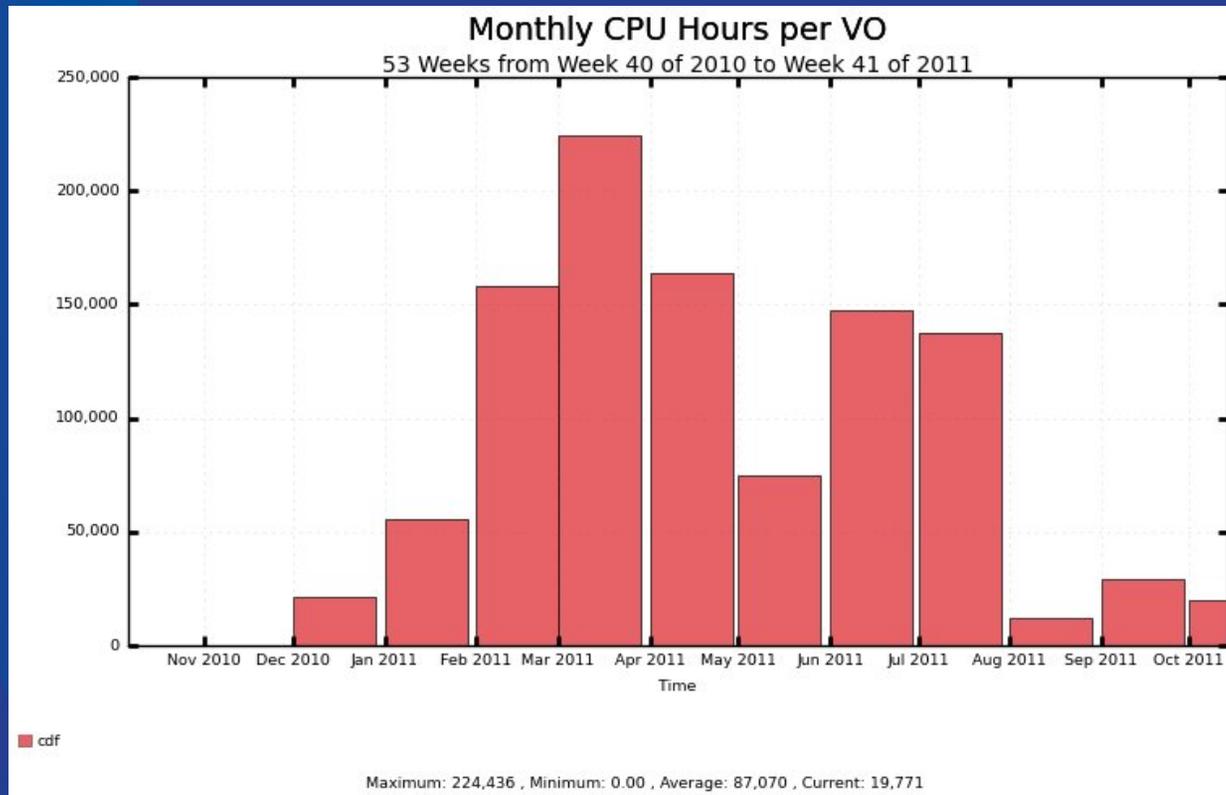


Grid-wide Federation Solutions

Campus Grid Federation Solutions

Resource Sharing across Resource Federations

- A vibrant federation of resources fosters resource sharing across Grids, including opportunistic resource usage.



Usage of KISTI by OSG VOs: CDF

KISTI is part of the OSG federation

Opportunities for Collaboration

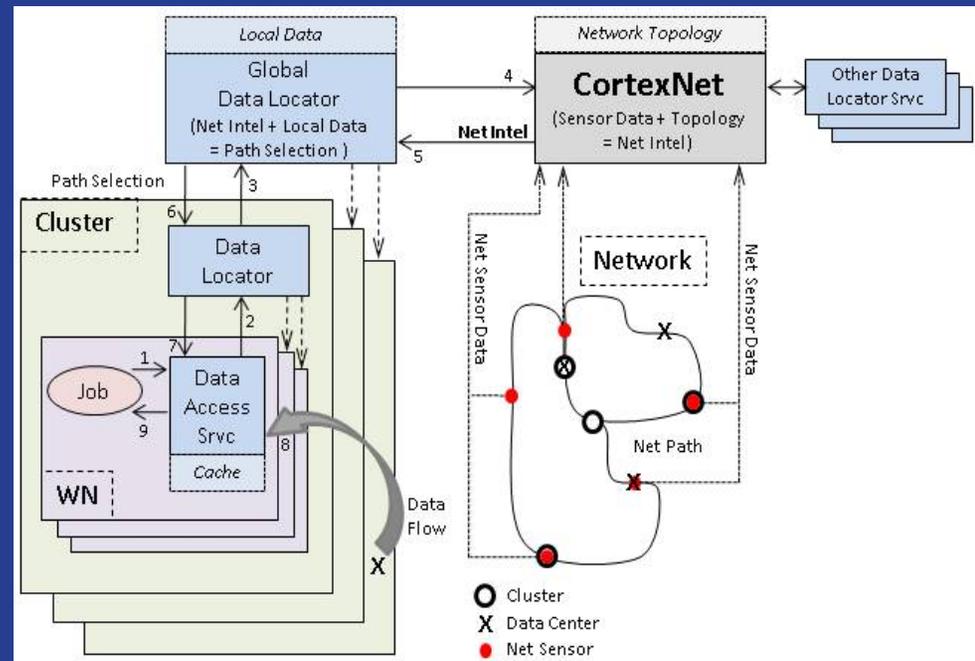
- ✓ **FermiGrid**
 - The Fermilab Campus Grid
- ✓ **FermiCloud**
 - The Fermilab Infrastructure as a Service
- ✓ **Federation of Resources in Korea**
 - Infrastructures for resource sharing
- **Fast Networks**
 - Preparing computing services to interface 100 Gbps links
- **Distributed Storage**
 - Storage Management across Grid sites
- **Computing Accounting**
 - Gratia: Inter-Grid and Cloud accounting

Focus on Fast Networks: Motivation

- The GCC dept. has a vibrant program of work on the cutting edge of technology
- The focus of the DOE Office of Science for the 2010-2020 funding is transitioning to...
 1. Exascale Computing
 2. Fast Networks
 3. Data
- The department is aligning its focus to these priorities, starting with fast networks
- Fermilab will be connected to 100 Gbps network in Dec 2011

GCC program of work on Networks

- Identifying gaps in middleware interfacing 100 Gbps networks
- Seeking funds to integrate network intelligence in site selection for data movement
- Cloud Computing:
 - Use of InfiniBand on FermiCloud (Hyunwoo Kim)
 - IPv6
 - Novel network topologies

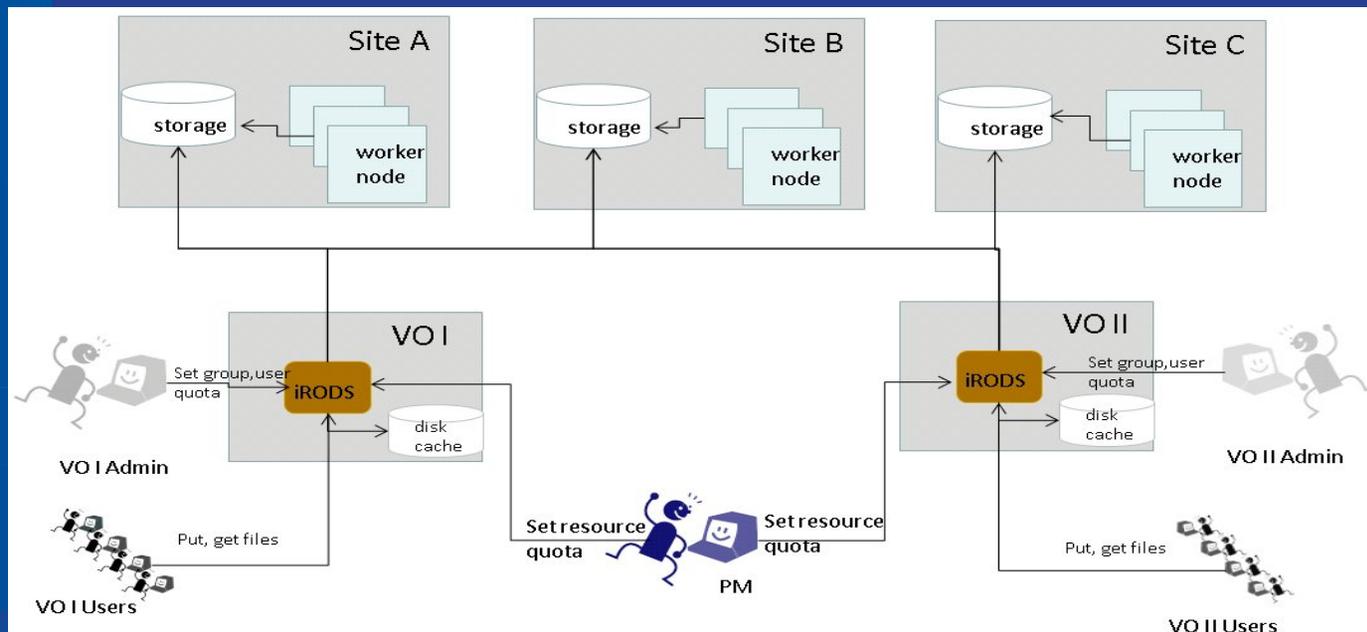


Opportunities for Collaboration

- ✓ **FermiGrid**
 - The Fermilab Campus Grid
- ✓ **FermiCloud**
 - The Fermilab Infrastructure as a Service
- ✓ **Federation of Resources in Korea**
 - Infrastructures for resource sharing
- ✓ **Fast Networks**
 - Preparing computing services to interface 100 Gbps links
- **Distributed Storage**
 - Storage Management across Grid sites
- **Computing Accounting**
 - Gratia: Inter-Grid and Cloud accounting

Public Storage Management

- OSG is investigating the use of iRODS to manage storage across Grid sites
- This will evolve in a system for VOs to handle data (with some features similar to SAM for CDF).
- Academia Sinica, Taiwan, involved in SRM to iRODS interface.



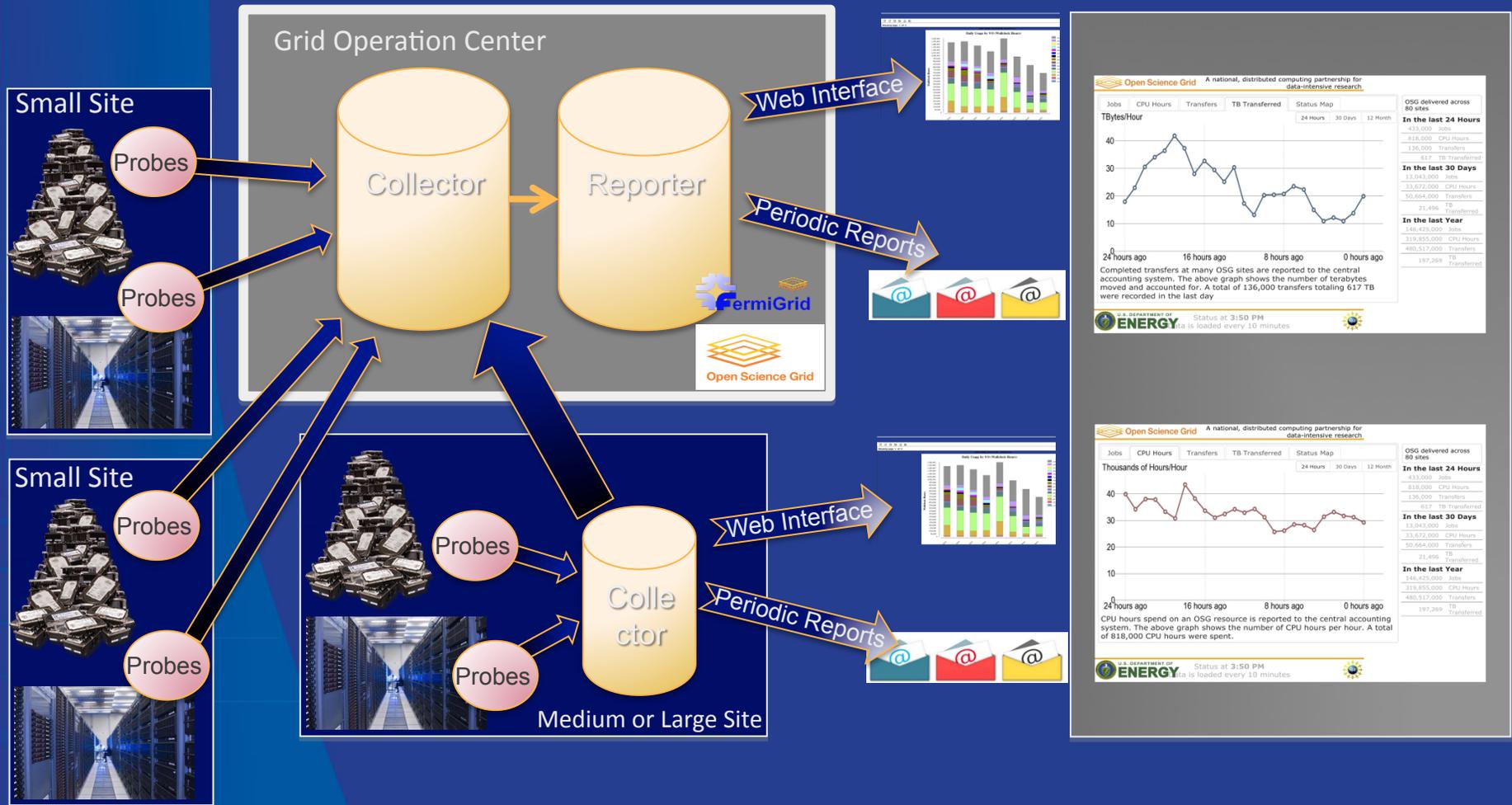
Opportunities for Collaboration

- ✓ **FermiGrid**
 - The Fermilab Campus Grid
- ✓ **FermiCloud**
 - The Fermilab Infrastructure as a Service
- ✓ **Federation of Resources in Korea**
 - Infrastructures for resource sharing
- ✓ **Fast Networks**
 - Preparing computing services to interface 100 Gbps links
- ✓ **Distributed Storage**
 - Storage Management across Grid sites
- **Computing Accounting**
 - **Gratia: Inter-Grid and Cloud accounting**

Gratia Accounting Service

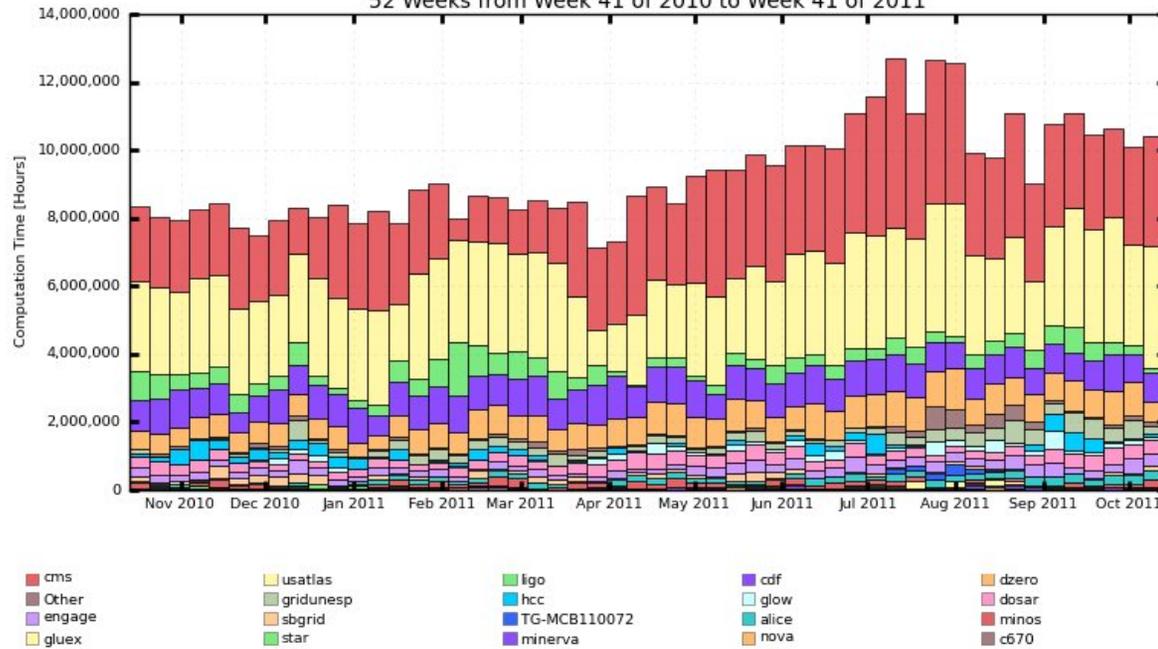
- Robust, scalable, trustable, dependable grid accounting service.
- The Gratia system consists of many probes operating on and uploading data from remote locations to a network of one or more collector-reporting systems.
- Data could be about batch jobs, grid transfers, storage allocation, campus grids, cloud utilization, site availability tests, or process accounting.
- Gratia provides summarization of accounting records
- The primary focus of the Gratia system is to provide an accounting of jobs executed on the Open Science Grid.

Gratia Architecture



Computation Hours Per Week

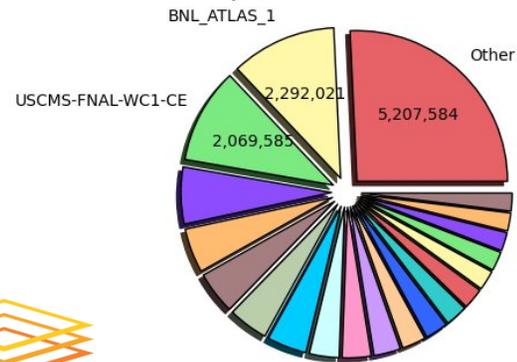
52 Weeks from Week 41 of 2010 to Week 41 of 2011



Maximum: 12,689,275 Hours, Minimum: 851,128 Hours, Average: 9,124,399 Hours, Current: 851,128 Hours

Wall Hours by Facility (Sum: 20,241,679 Hours)

14 Days from 2011-09-29 to 2011-10-13



Open Science Grid



Conclusions

- The collaboration between GCC and KISTI has resulted in a win-win opportunity for both institutions.
- The Grid Department is thrilled of continuing a strong collaboration
- We are open to continue a program of work...
 - Hosting KISTI engineers at Fermilab
 - Working on common projects remotely
 - Fostering the partnership of KISTI with OSG