Eileen Berman

Condor in the Fermilab Grid Facility

Fermilab

- Fermi National Accelerator Laboratory is a high energy physics laboratory outside of Chicago
- Many different experiments, comprised of 1000's of users working for many years, rely on Fermilab to provide the core services and software necessary to enable the research that leads to scientific discoveries
- The Fermilab Grid Facilities are strategic in the execution of this mission

Fermilab Grid Facility

The Fermilab Grid Facility supports many different activities in support of these scientists.

The scale of these activities places demands that often require close collaborations.

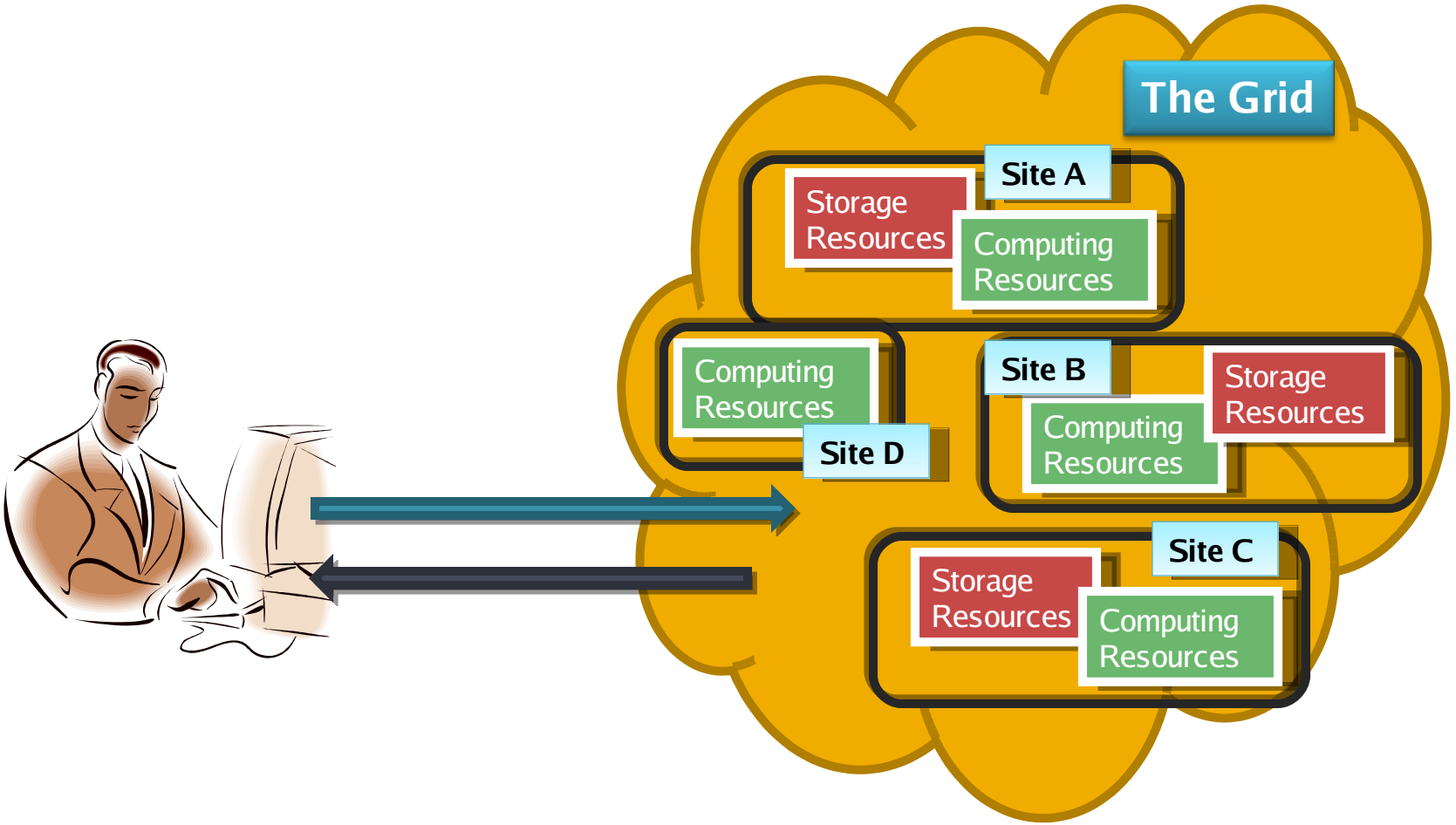
- Petabytes of data to store and manage
- Billions of events to analyze
- Tens of millions of jobs to run per year

Collaboration with the Condor team over the years, has enabled Fermilab to deliver the quality software and services that this large scale science demands.

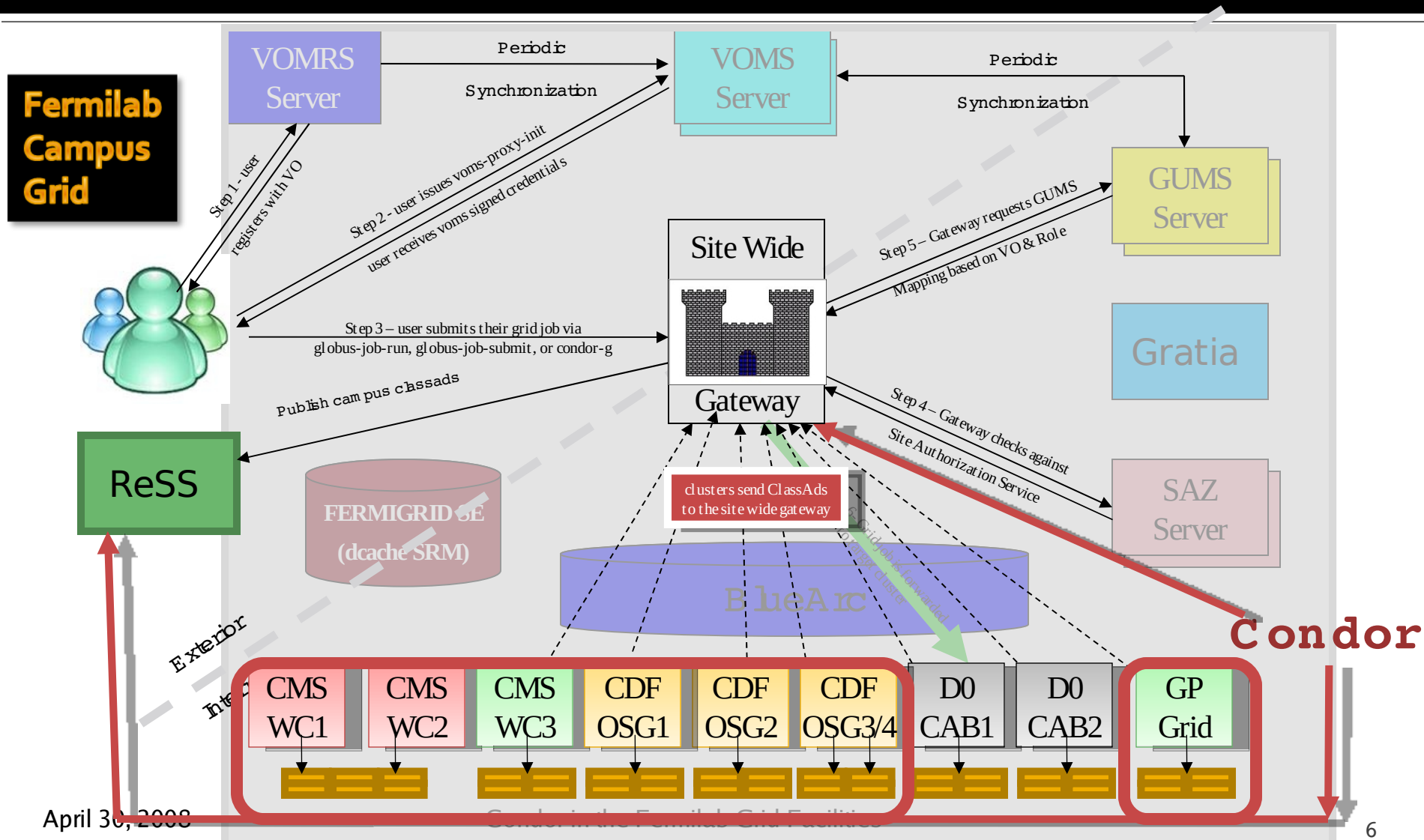
What are we talking about?

- Fermilab Campus Grid (FermiGrid)
- DZero Data Handling and Job Submission System (SAMGrid)
- Resource Selection Service (ReSS)
- Glidein Workload Management System (WMS)
- Grid/Site Security Boundary (OSE)
- End To End Security

The Grid


The Grid
Site A
Storage Resources
Computing Resources
Site B
Storage Resources
Computing Resources
Site C
Storage Resources
Computing Resources
Site D
Computing Resources

FermiGrid Architecture



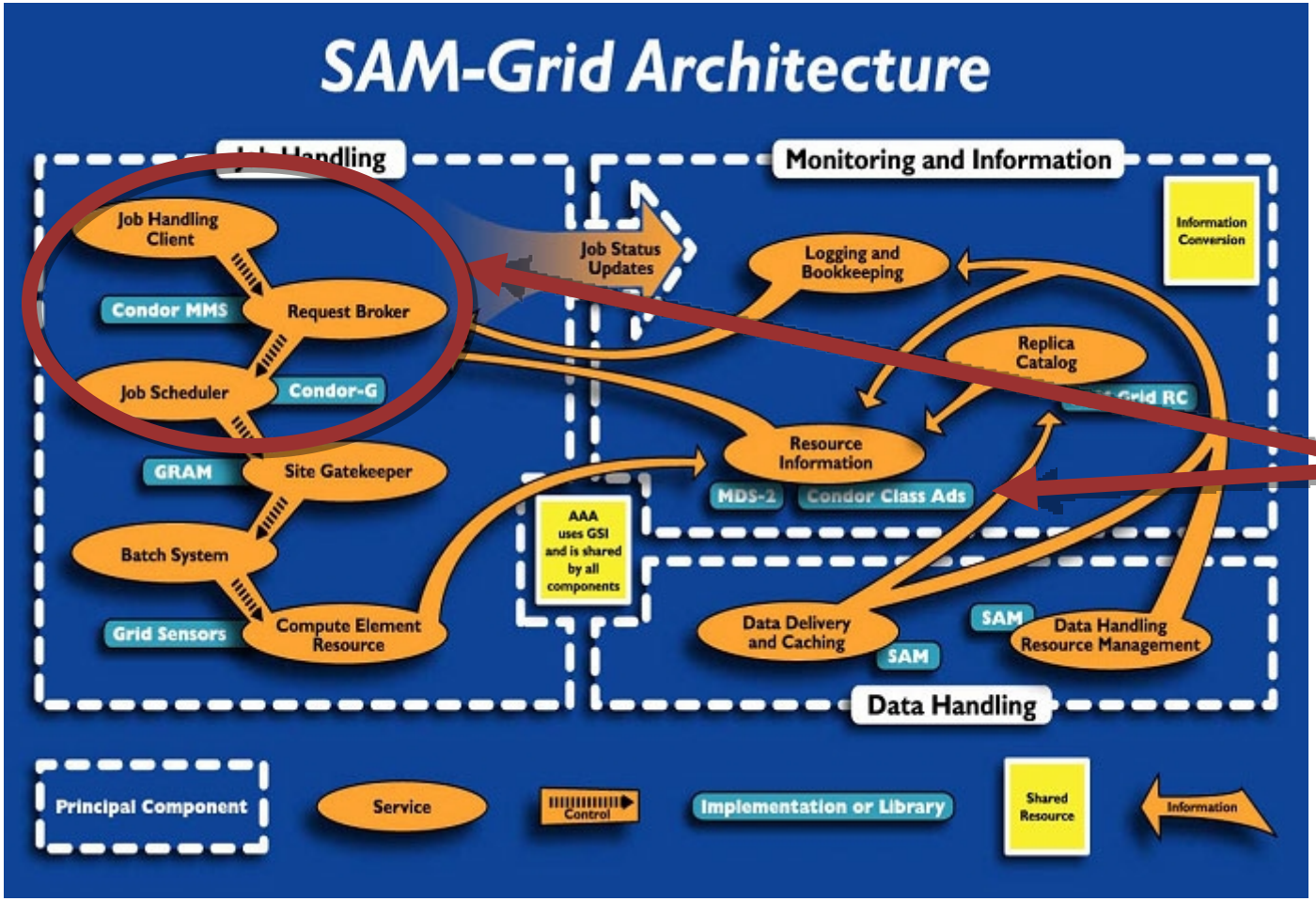
FermiGrid

- As data volumes increase, demands for increased data processing power has grown beyond the ability to support in a diverse disconnected environment.
- FermiGrid is a strategic cross campus grid supporting the experiments at Fermilab
  - 11,786 cores amongst the clusters
    - 7586 managed by Condor
  - Access to several Petabytes (10^{15}) of storage
  - O(1000) users
  - Provides a common mechanism for supporting the users while minimizing the expense of service delivery.

FermiGrid

- Condor is the underlying batch system on most of the clusters
- The Site-Wide Gateway is a natural pairing of Condor-G and a hierarchical grid deployment
- Gateway uses jobmanager-cemon
  - Based on jobmanager-condor
  - Matches jobs against cluster information via a local condor matchmaking service (ReSS)
  - Forwards to matched sub-cluster
  - Enhances scalability
  - Fault tolerance under development
- Additional Gateways will provide job access to Teragrid resources
- Significant scalability work done by Condor team to support usage patterns

SamGrid Architecture



DZero data handling and grid job submission system

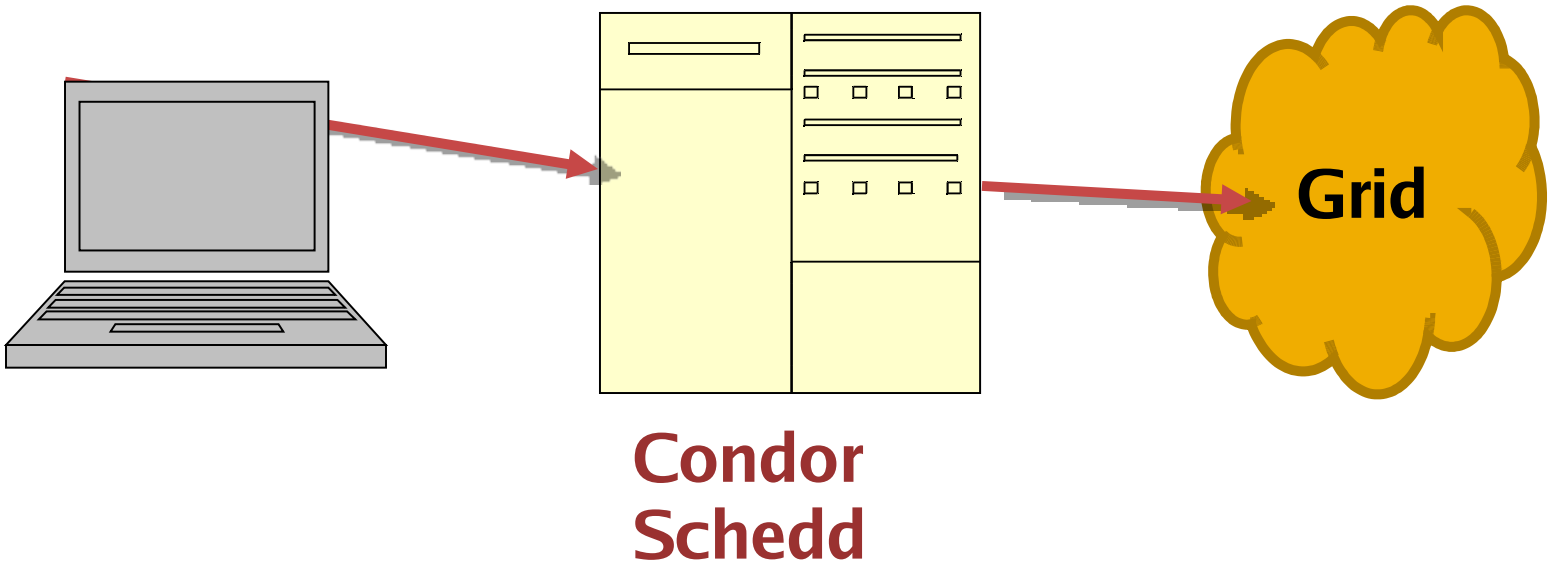
Condor

SamGrid

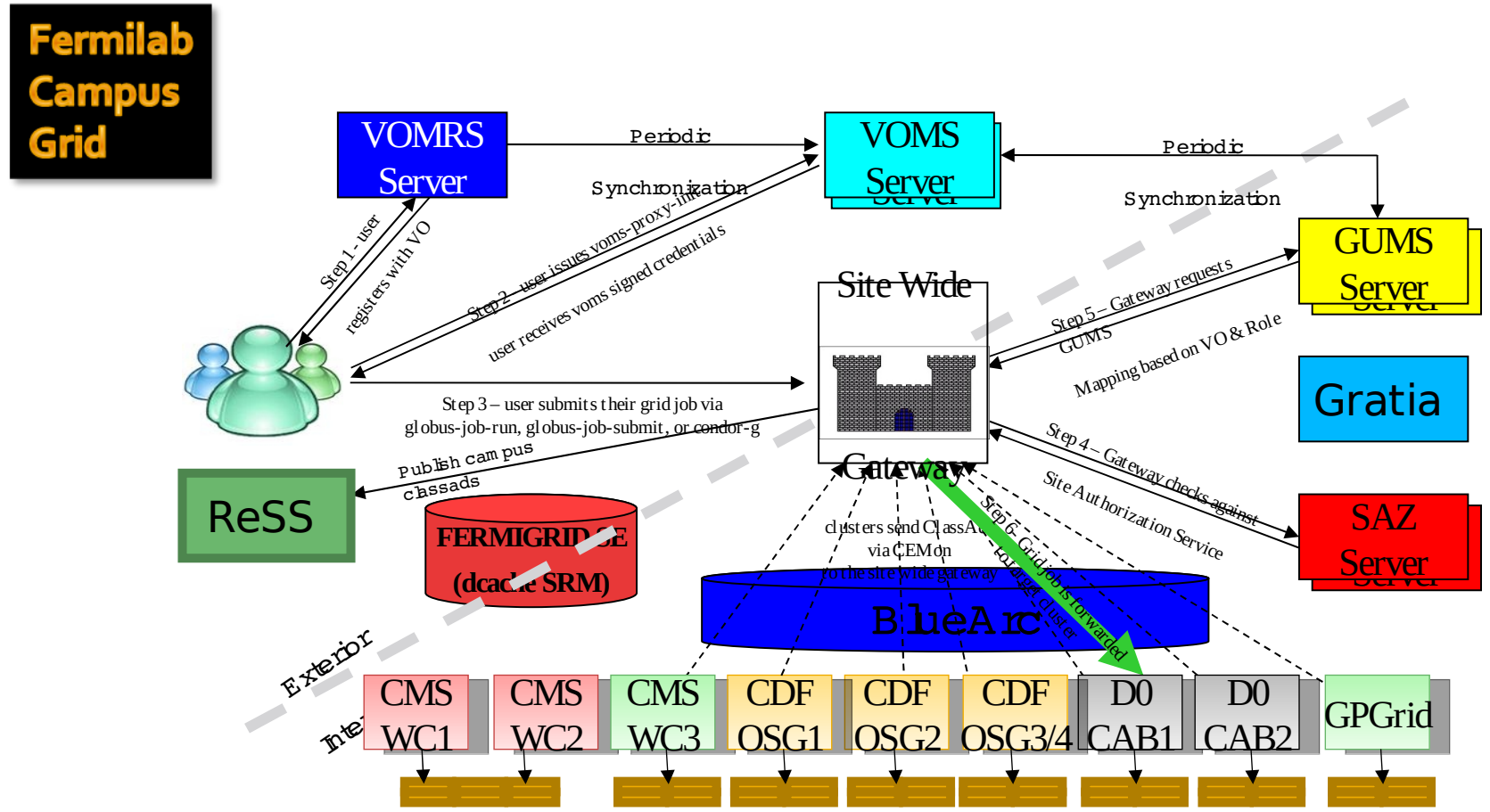
- SamGrid is the data handling and job submission system supporting the Dzero collaboration (500 physicists) data processing and simulated event generation needs
  - 100's of Terabytes of data needing processing
  - Billions of events with data and metadata to manage and deliver
  - Needed a widely distributed system to support the data delivery and data processing needs of the experiment, located across the globe
- Early adopter of Condor matchmaking service with Condor-G to select clusters instead of machines.
  - Had to learn how to describe clusters
- Uses Condor matchmaking capabilities to throttle the flow of jobs from the schedulers to the sites (stable, configurable)
  - Uses cluster classad requirements
- Policies can be implemented based on job status values (e.g. running, held…)

SamGrid

- Condor team implemented multi-tiered job submission mechanism necessary in order to take advantage of the diverse clusters on the Grid.

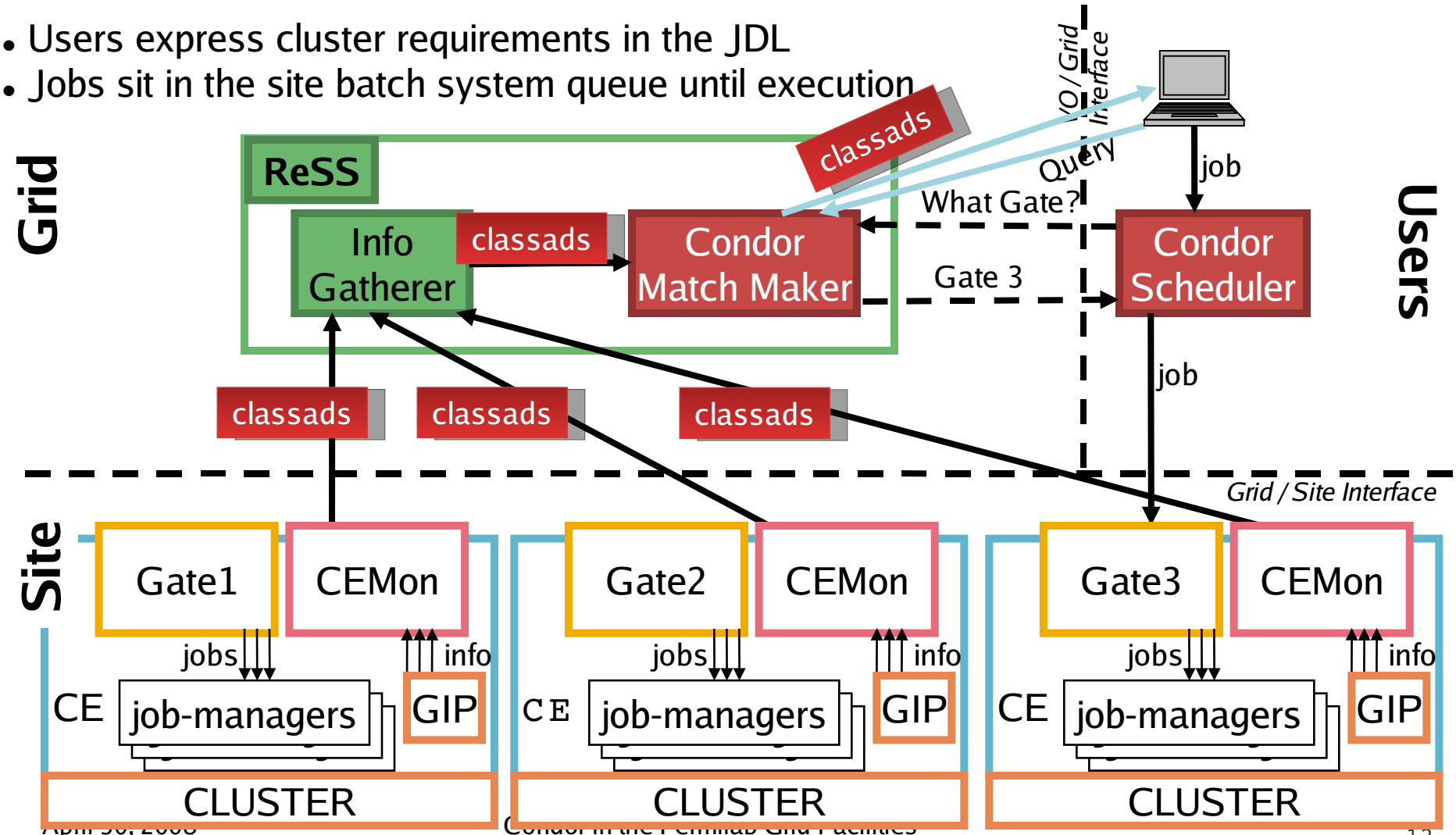


ReSS


Fermilab Campus Grid
VOMRS Server
Periodic Synchronization
VOMS Server
Periodic Synchronization
GUMS Server
Step 1 - user registers with VO
Step 2 - user issues voms-proxy-init
user receives voms signed credentials
Site Wide Gateway
Step 5 – Gateway requests GUMS
Mapping based on VO & Role
Gratia
Step 3 – user submits their grid job via globus-job-run, globus-job-submit, or condor-g
Publish campus classads
Step 4 – Gateway checks against Site Authorization Service
SAZ Server
ReSS
FERMIGRID SE (dcache SRM)
clusters send ClassAds via CEMon to the site wide gateway
Step 6 - Grid jobs forwarded
BlueArc
Exterior
CMS WC1
CMS WC2
CMS WC3
CDF OSG1
CDF OSG2
CDF OSG3/4
D0 CAB1
D0 CAB2
GPGrid

ReSS Architecture

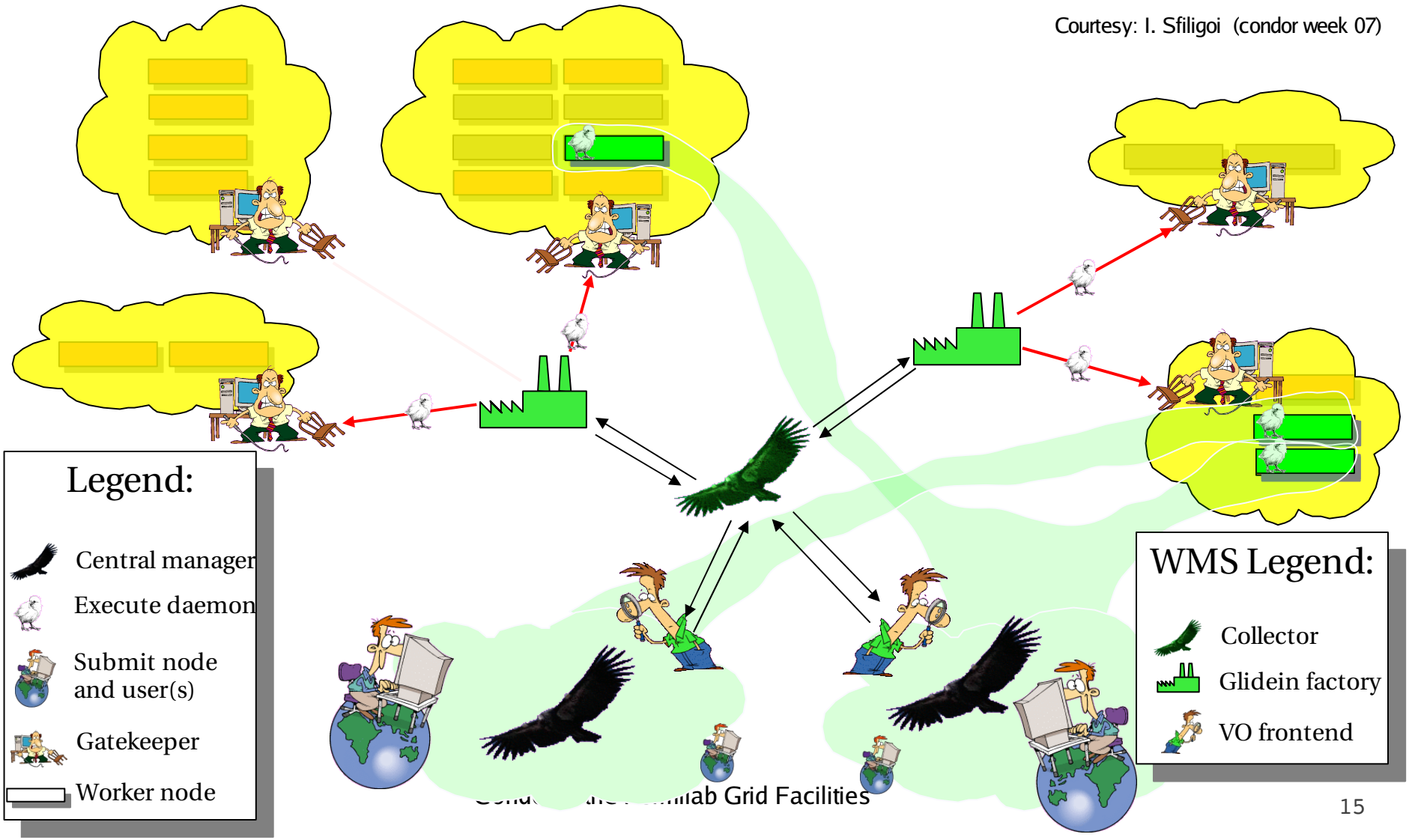
- Users express cluster requirements in the JDL
- Jobs sit in the site batch system queue until execution



ReSS

- Needed the ability to match diverse job needs with available resources across an entire grid (OSG) without requiring users to know the details
- Resource Selection Service using same concept as SamGrid, generalized for the Open Science Grid (OSG)
  - Condor-G Matchmaking
- ReSS is the production OSG resource selection service – **Grid Wide**
- FermiGrid uses ReSS internally for matching jobs to the campus clusters.
- For more information see – Mats Rynge's talk "OSGMM and ReSS – Matchmaking on OSG"

Glidein Workload Management System (WMS)



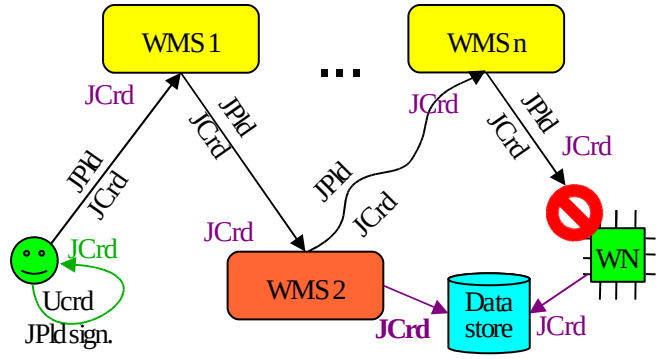
Glidein WMS

- Provides glideins in a more transparent manner insulating the user from the diverse nature of the grid
- Supports vanilla and standard universe jobs
- Works across firewalls
  - Required GCB be made production quality to support this
- Support for calling glExec built in to Condor directly
  - Allows separation of execution of the Glidein pilot and the user job
- Scaling upgrades were done to support the O(10,000) jobs running simultaneously
- Creates a virtual condor pool so distributed grid resources look like dedicated local resources

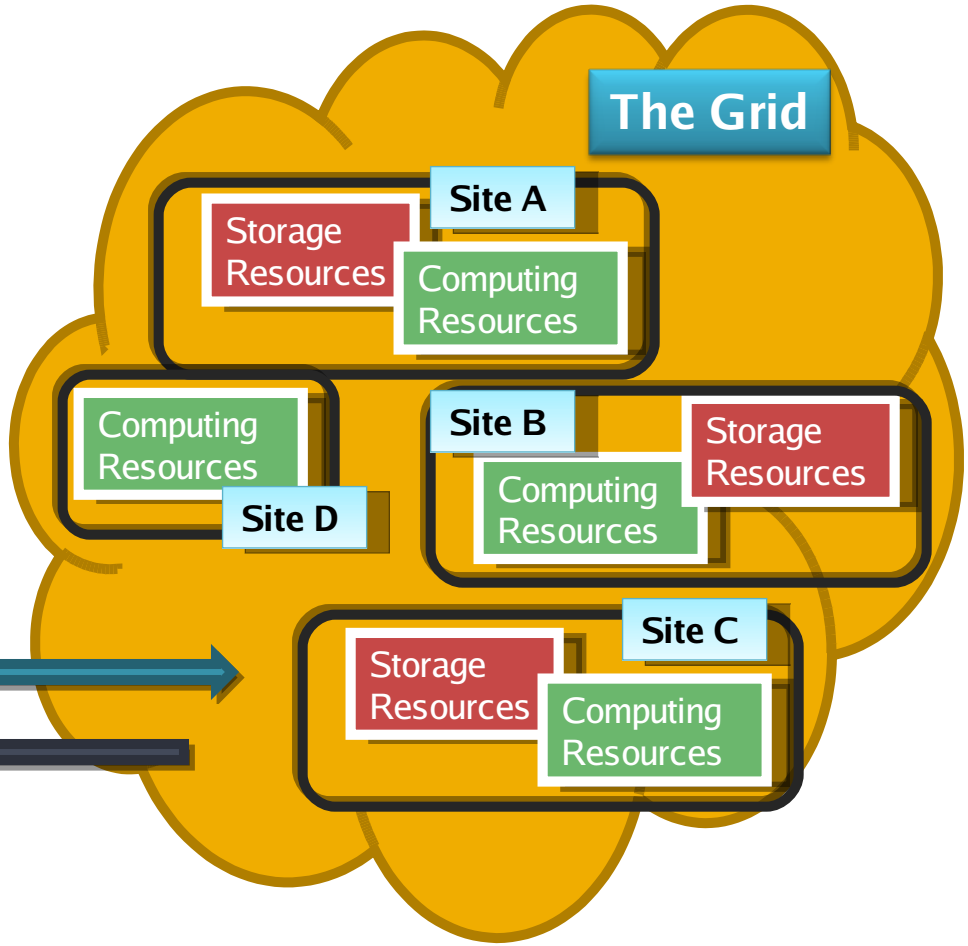
Grid/Site Security Boundary

- Work is being done at Fermilab at the boundary of the Condor-G world and Condor, determining the security considerations on this boundary.
- Defines how site installations can work securely when interfacing with a grid environment.

End To End Security



Acronyms:
UCrd - User credentials
JCrd - Job specific user credentials
JPld - Job payload



End To End Security

- With the implementation of Glidein WMS's, a job may make many stops on its way to a worker node
  - How do you insure the job integrity?
- Can determine (e.g. on a worker node), if the job has been tampered with
- Implements signed classads to help improve the integrity of the user job as it travels in the Grid
- Collaboratory investigation between Fermilab and the Condor team.
- See Ian's talk (Thursday)
  - 'End-to-end security in Condor'

Conclusion

- Fermilab provides a secure, production grid environment offering grid software and services that enables scientists to produce scientific results on the scale required by today's petascale experiments.
- The productive relationship existing between Fermilab and the Condor team has contributed to this success.
- Condor is part of the Grid Facility at Fermilab, serving the production needs of many experiments and that of the wider OSG.