Trillium & Grid2003
Status and Plans

Ruth Pordes,
Fermilab

Sept 26th 2003

Acronyms abound.. Please ask me to explain if I use ones you don't know..

Note: Applications using Standard Grid tools are less fully developed as end to end systems compared to SAM..

Trillium Status - PPDG

- Experiment and Computer Science Team based. Significant DOE Lab involvement.
  - Focus on end to end applications on the grid. Made impact over the past two years in U.S. experiments acceptance of grids e.g. STAR, JLAB, US CMS.
  - Data management, job management are focus areas.
    - But we have six (seven?) implementations of file replication and management services.
    - Migrating to the use of VDT.
    - Web portals for data replication and job submission prototypes.
    - Keep trying for a common Job Description Language (a la EDG) but so far no real engagement.
    - Storage Resource Management (SRM) interface in collaboration with SRM DOE project (LBNL, Fermilab) quite successful.
  - Prototyping and requirements for grid analysis environments.
    - PPDG CS-11 providing a productive forum for cross U.S. experiment discussion of terminology and requirements.
    - Working also with the CERN-LCG Grid Analysis Group (GAG).
  - Facilitate cross-experiment sharing and collaboration. E.g. PHENIX showing some initial “interest” in SAM.
  - Successful DOE SciDAC review last April.

Trillium Status - GriPhyN

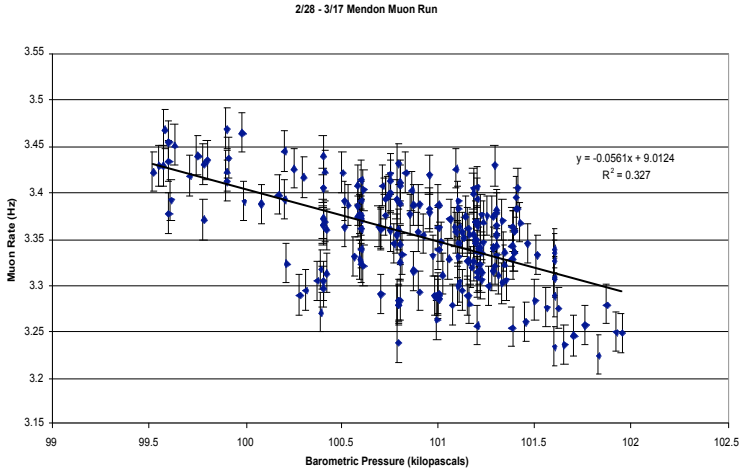
- GriPhyN - institution based:
  - Delivers software to the Virtual Data Toolkit catch all packaging, and testing for shared grid middleware. (VDT http://www.lsc-group.phys.uwm.edu/vdt/, based on NSF Middleware Initiative (NMI) releases http://www.nsf-middleware.org/).
  - Virtual Data System (Chimera http://www.griphyn.org/chimera) .
    - Scripts to generate scripts for running multiple and dependent jobs.
    - Provenance, Transformations and Derivations driven through declarative descriptions.
    - Initial releases in use for ATLAS, SDSS and LIGO demonstrator and production applications. Changes being made as deployments proceed.
    - Now included in VDT.
    - Overlaps in functionality with Runjob. Interfaced to CMS version of Runjob for in development testbeds.
  - Other computer science research in Planning (Pegasus), Partial Results (from LBNL), Fault Tolerance (UCSD) ongoing.
  - Successful NSF review last February.

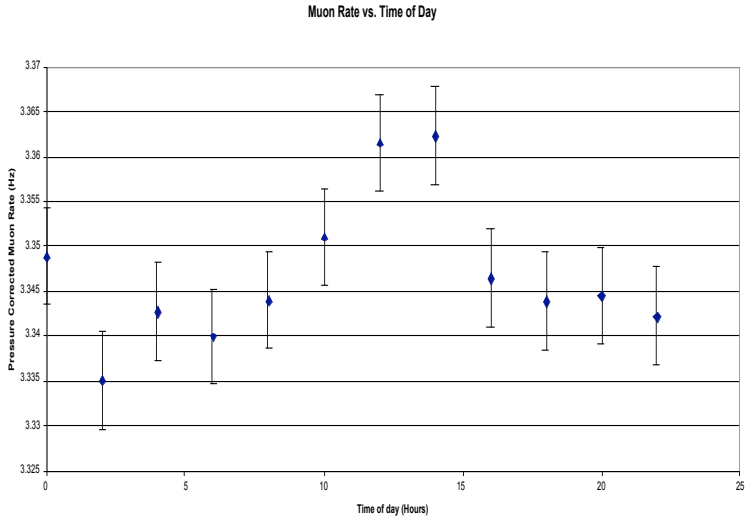
Trillium Status - iVDGL

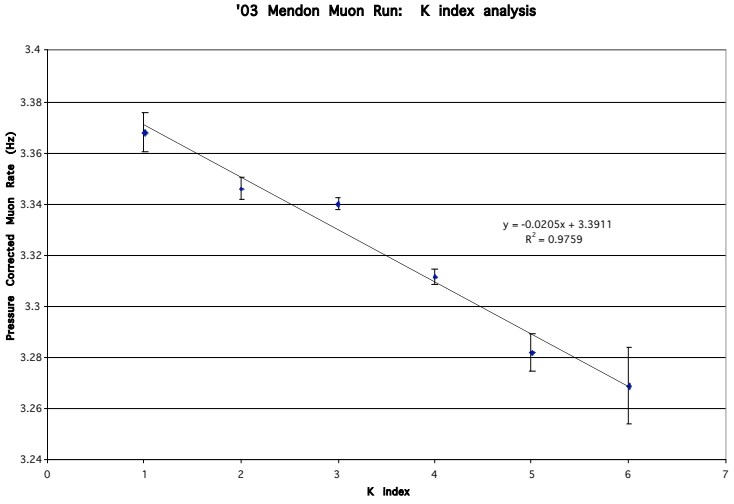
- Work organized as work teams across institutions:
  - Tier 2 facilities for ATLAS, CMS, LIGO, SDSS. Multi organization laboratory for testing and using grids. Open to new groups and sites on request and agreement by Steering Committee.
  - Deployment and support of VDT. Operations center at Indiana.
  - Currently concentrating on Grid2003 or Grid3 Functional Demonstrator Grid project. http://www.ivdgl.org/Grid2003/
  - Quarknet - Grid project merging existing Cosmic Ray detectors with Virtual Data techniques for data transformation and derivation. With analysis available through a web porgal
    - Expectations of working systems early next year.
    - From Marge Bardeen: “Grid Resources Add Value to Education”
      - New style of collaborative learning. Online “posters” & discussions with students at other schools
      - Single portal to distributed resources
      - Access to computing power & file space
      - Participation for students without cosmic ray detectors

Mendon Research 2003

- Single three-week data run
- Three class projects:
  - Muon rate vs. pressure
  - Muon rate vs. time of day
  - Muon rate vs. geomag. activity
    - (Using NOAA Internet data)







As Trillium we

- Participate in Global Grid Forum.
  - Particle and Nuclear Physics Research group just “approved” by executive committee.
- Work with LHC Computing Grid Project and other European projects such as the European Data Grid, DataTAG etc.
- Grid2003..

Grid2003 is a Joint Project between US ATLAS, US CMS, PPDG, GriPhyN, iVDGL

- Demonstrate that the 2004 Data Challenge needs of US CMS and US ATLAS can be effectively partially met by a distributed grid infrastructure, and production and analysis applications running **at scale** in a **common** grid environment.
- Demonstrate computer science technologies at scale in a common grid environment e.g. sustained throughput using third party GridFTP transfers at all sites, first U.S. production use of Replica Location Service (RLS).
- Common grid environment with policies applied
  - for LIGO and SDSS applications, and possibly other experiment or scientific applications.
  - US CMS jobs to run on US ATLAS sites and vice versa subject to control of the % of the throughput each gets on each others sites.
- To operate a grid of >500 machines ("worker" or farm nodes) and >12 sites. The complexity is in the diversity of sites and applications, not in the scale.
- Data replication and data movement services are in early stages: based on SRM, GridFTP and RLS, and perhaps DCache.

A six month project from July-Dec 2003 with demonstrations at SC2003 conference - Nov 16th

Principles - a Federated Grid

- Experiments must be able to effectively interoperate and run their applications on non-dedicated resources.

- Applications must be able to install themselves dynamically, thereby imposing minimum requirements on grid facility managers.

- *Facilities/Sites* - Site Administrators control setup and management of systems, install and configure the middleware, and support execution of applications.

- *Services* - collaborative approach to bringing up cross-site services, e.g. VO management, monitoring, configuration management supported by the whole Grid3 team.

- *Applications* - submitted and operated by Application Administrators. Applications group responsible for end-to-end operations and bringing up the production services.

- The middleware is based on VDT 1.1.11. Set of used components of common middleware driven by each individual application.

- MDS used for providing and displaying information. Glue schema + Grid3 schema extensions. Ganglia, and Monalisa for the monitoring.

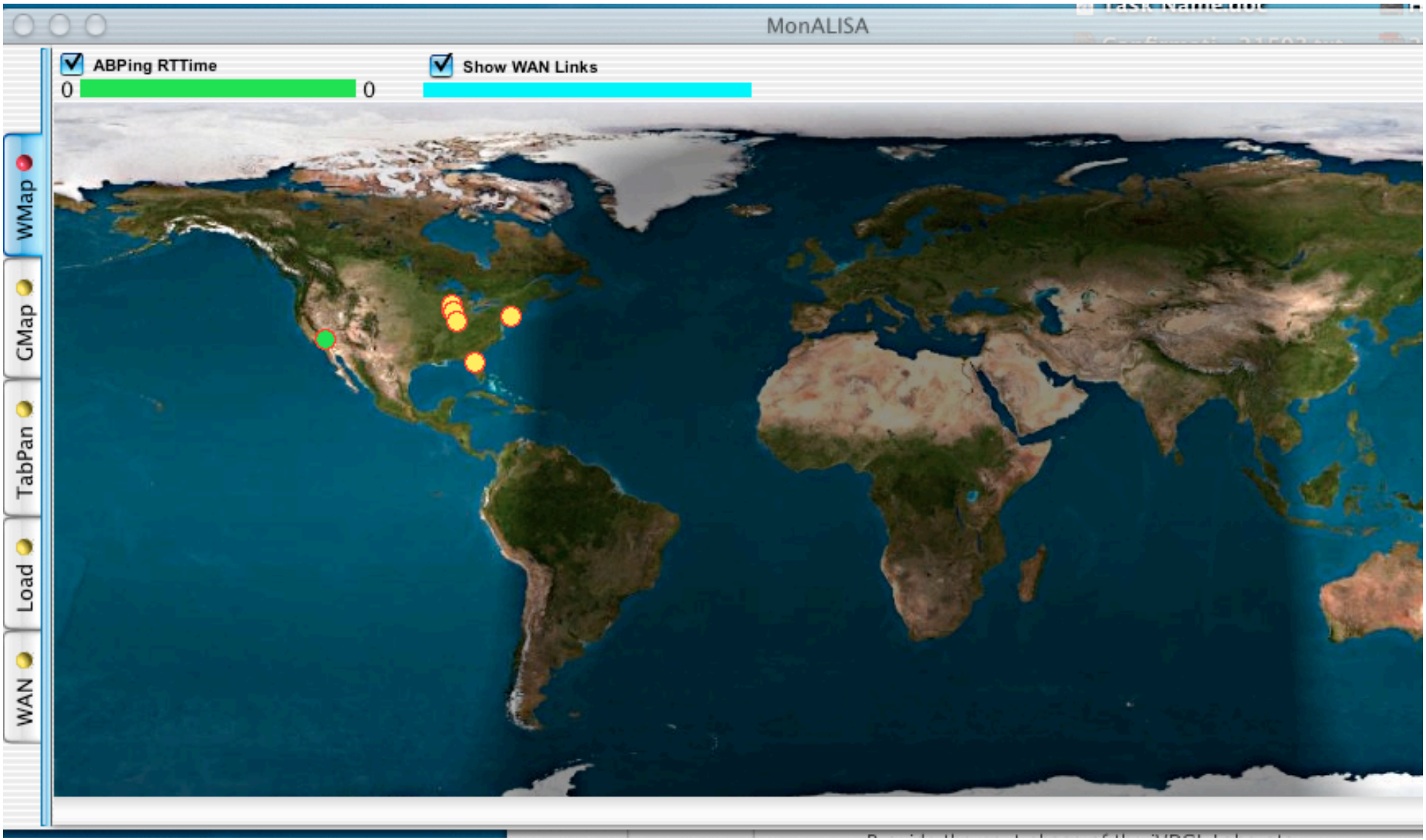
Grid 3 Demonstrator Metrics

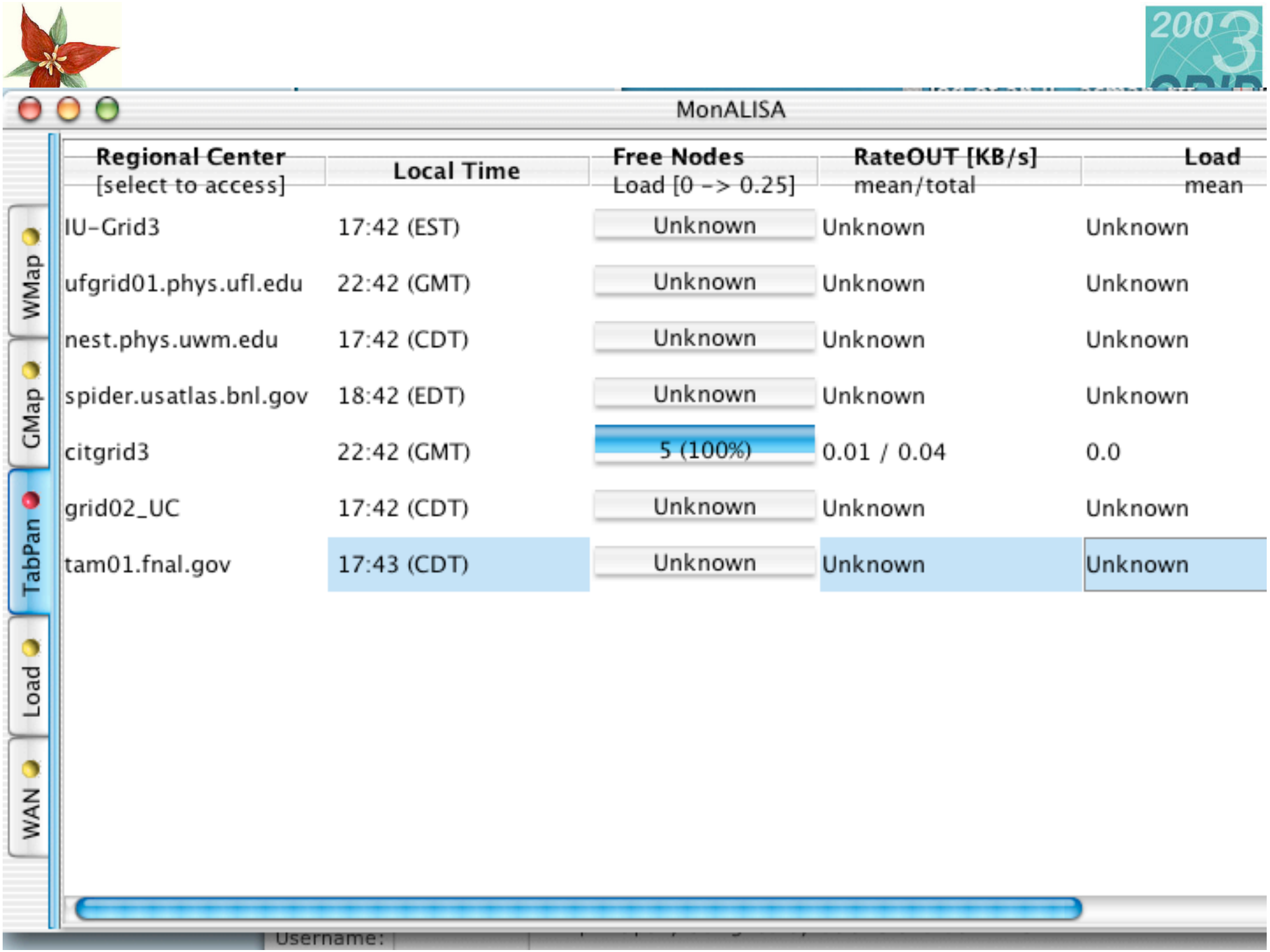
Number of processors	up to 500	Possibly collected through MDS, archived and time-stamped
Data transferred per day	>2-3 TB	Data "flux" will need to be defined carefully.
Peak number of concurrent jobs	Up to 1000	Collect the total number of jobs running on Grid3, sorted by VO, archived so that time-dependent plots can be made.
Percentage of resources used	up to 90%	
Efficiency of job completion	up to 75%	Success to be defined. Other efficiency metrics could be identified.
Number of users	>10	Collect and sort by VO origin.
Number of different applications	>4	By an application "registry", such as proposed by the WorldGrid "project" mechanism, this could be simplified.
Number of sites running multiple applications	>10	Collect with time stamps, intervals to be determined.
Rate of Faults/Crashes	<1/hour	Measure, and perhaps categorize by degree of severity.
Operational Support Load of full demonstrator	<2 FTEs	How many people (and in which roles) were operating Grid3?

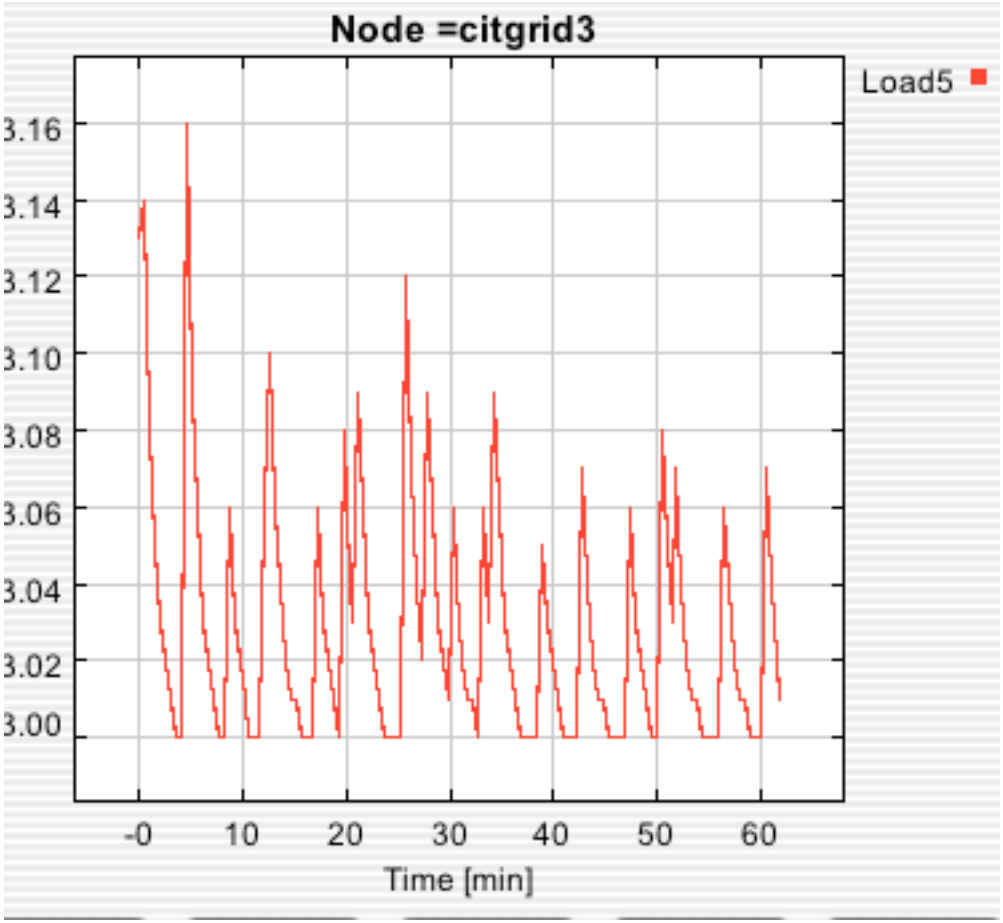
Operate the Grid stably for 1 week.

Much additional work in Policies (registration, priority), Procedures (information required and provided through MDS), Service Level Agreements (SLAs) with operations and support groups, Documentation and Web presence.

Monitoring Displays



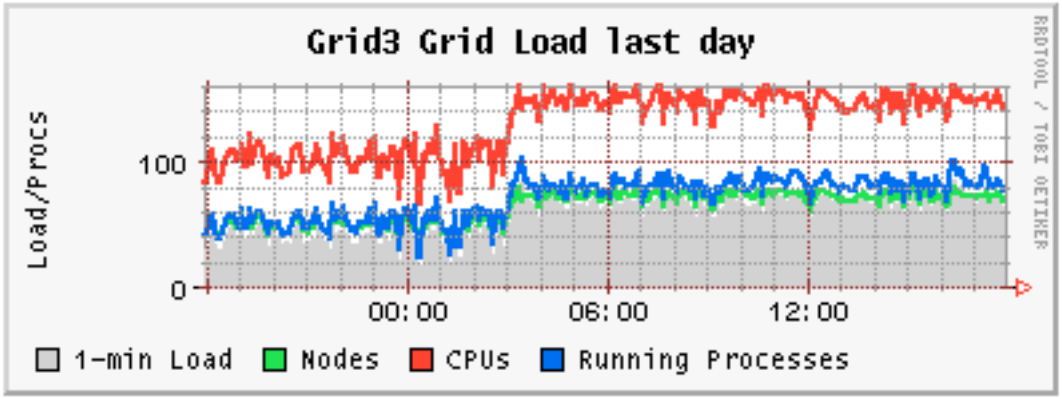

MonALISA
WMap
GMap
TabPan
Load
WAN
Regional Center
[select to access]
Local Time
Free Nodes
Load [0 -> 0.25]
RateOUT [KB/s]
mean/total
Load
mean
IU-Grid3
17:42 (EST)
Unknown
Unknown
Unknown
ufgrid01.phys.ufl.edu
22:42 (GMT)
Unknown
Unknown
Unknown
nest.phys.uwm.edu
17:42 (CDT)
Unknown
Unknown
Unknown
spider.usatlas.bnl.gov
18:42 (EDT)
Unknown
Unknown
Unknown
citgrid3
22:42 (GMT)
5 (100%)
0.01 / 0.04
0.0
grid02_UC
17:42 (CDT)
Unknown
Unknown
Unknown
tam01.fnal.gov
17:43 (CDT)
Unknown
Unknown
Unknown
Username:

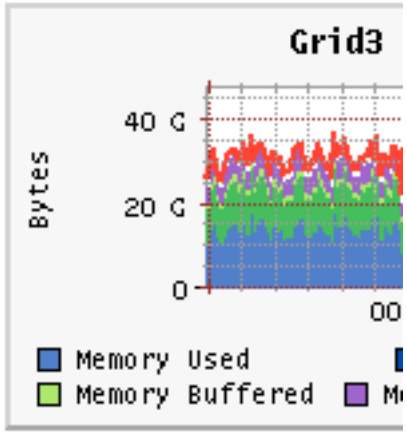

Node =citgrid3
Load5
3.16
3.14
3.12
3.10
3.08
3.06
3.04
3.02
3.00
-0
10
20
30
40
50
60
Time [min]

Grid3 Grid (8 sources) (tree view)

CPUs Total: **156**
Hosts up: **76**
Hosts down: **1**

Avg Load (15, 5, 1m):
49%, 48%, 49%
Localtime:
2003-09-14 17:54

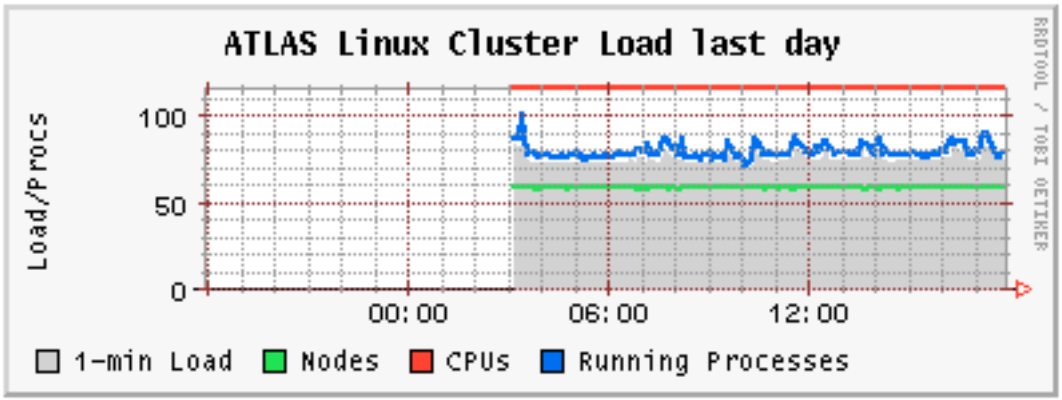


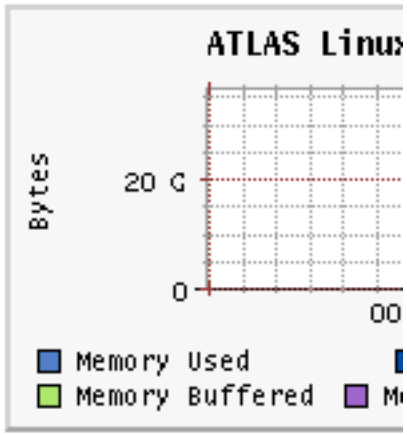


ATLAS Linux Cluster (physical view)

CPUs Total: **116**
Hosts up: **58**
Hosts down: **0**

Avg Load (15, 5, 1m):
64%, 63%, 64%
Localtime:
2003-09-14 17:51

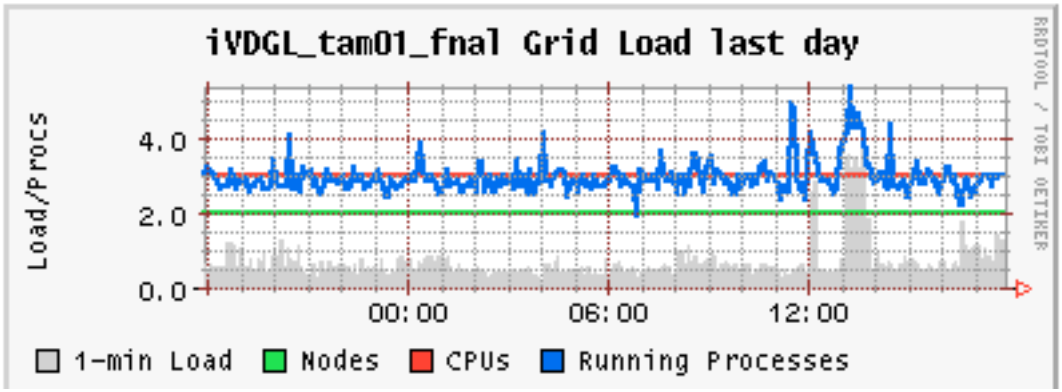


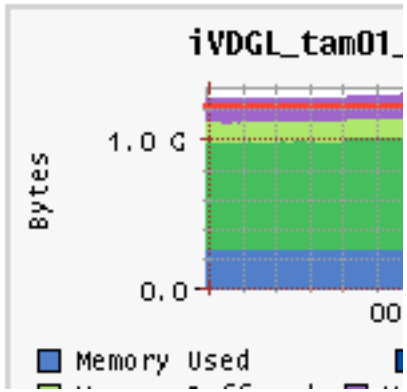


iVDGL_tam01_fnal Grid (tree view)

CPUs Total: **3**
Hosts up: **2**
Hosts down: **1**

Avg Load (15, 5, 1m):
36%, 33%, 29%
Localtime:
2003-09-14 17:52





Grid2003 is a step to..

- Proposed permanent Grid Infrastructure for the nation.
- .Peer with European EGEE and LCG

Propose the Open Science Grid:

as a strategy to build a national infrastructure for science in the U.S.

- Open Science Grid program of work
  - Federate disjoint grid resources at labs and universities into single, scalable, engineered, managed grid
  - Start with LHC resources to build peta-scale global grid
  - Include Run2, BaBar, RHIC when possible and "desired" (as have effort/funding)
  - Can serve as backbone to merge and federate with others

- Open Science Grid goals
  - Benefit maximally from investments already made or planned
  - Ensure continuing U.S. leadership in defining and operating global grids
  - Provide initial set of grid services to be enriched as others join
  - Combine resources at several DOE labs and NSF funded universities into a coherent national resource

Vicky White: "The idea is to federate, over time, all of these computers and storage systems and services together into a Grid that serves the needs of all of these physics and related disciplines. But more importantly to invite other sciences to join this Open Science Grid and to invite educators, students and others to work with us as we evolve."