Grid2003 Monitoring, Metrics, and Grid Cataloging System

Leigh GRUNDHOEFER, Robert QUICK, John HICKS (Indiana University)
Robert GARDNER, **Marco MAMBELLI**, Andrew ZAHN (University of Chicago)
Paul AVERY, **Bockjoo KIM**, Craig PRESCOTT, Jorge RODGRIGUEZ (University of Florida)
Mark GREEN (University at Buffalo)
Ian FISK, John WEIGAND (FERMI NATIONAL LABORATORY)
Iosif LEGRAND (CERN)

CHEP
Interlaken, CH - Sept 30 2004

Integrated Data Monitoring and Analysis

- Application domain requirements
  - Accounting VO activity
  - Compare with 'traditional batch production'
- Grid environment constraints
  - Heterogeneity
  - Scalability
  - Limited resources (bandwidth, connectivity)
  - Small footprint (light, unprivileged, compatible)
  - Reliability
- Deployment
  - Multi-level architecture
  - Using existing tools
  - Extend and integrate
  - Develop new tools to fill gaps

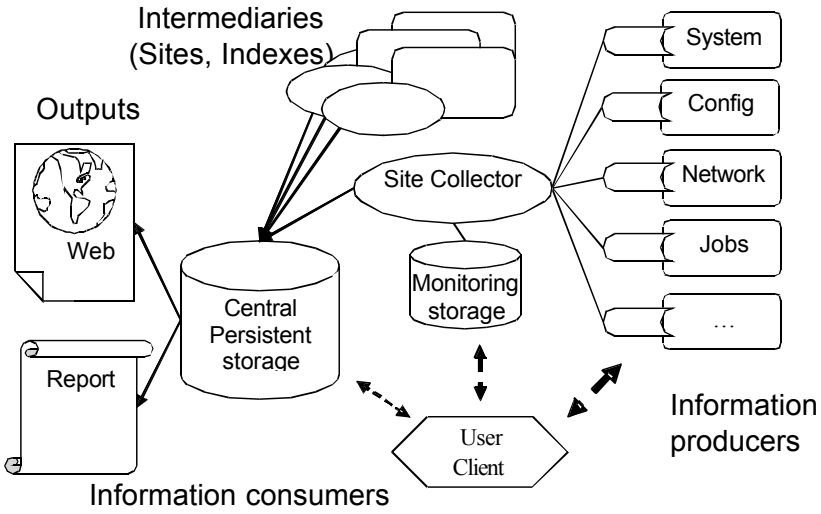
Example of metrics specification

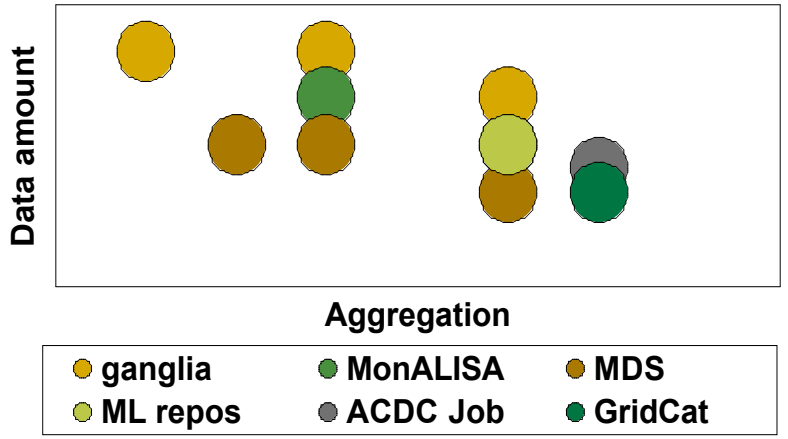
Measure	Category	T
Number of sites (per VO)	R	E
Total number of CPU (per site, VO)	R, U	P
CPU usage (per site)	R	P
Total I/O (per site, per VO)	R, U, D	P
Available disk -space	U, D	P
Memory and swap usage	D	P
Load, as number of running jobs (per VO, per site)	R, U, D	E
Number of running jobs (per VO, per Site)	U, D	E
Number of queued jobs (per VO, per Site)	U	E
CPU usage per VO	U	O
Downtime (per site – a system will be considered up when able to pass a defined test)	R, U	P
Submission efficiency as successful over submitted jobs ratio	U	O

- desired measures and short description
  - user category requiring it (**R**eporting, **U**ser, software **D**evelopment)
  - time sampling requirement (**O**ffline, **P**eriodic, **E**vent-based)
- Per-VO metrics:
  - Execution on more resources
  - Partial use of a resource

Components of an Integrated Monitoring Framework

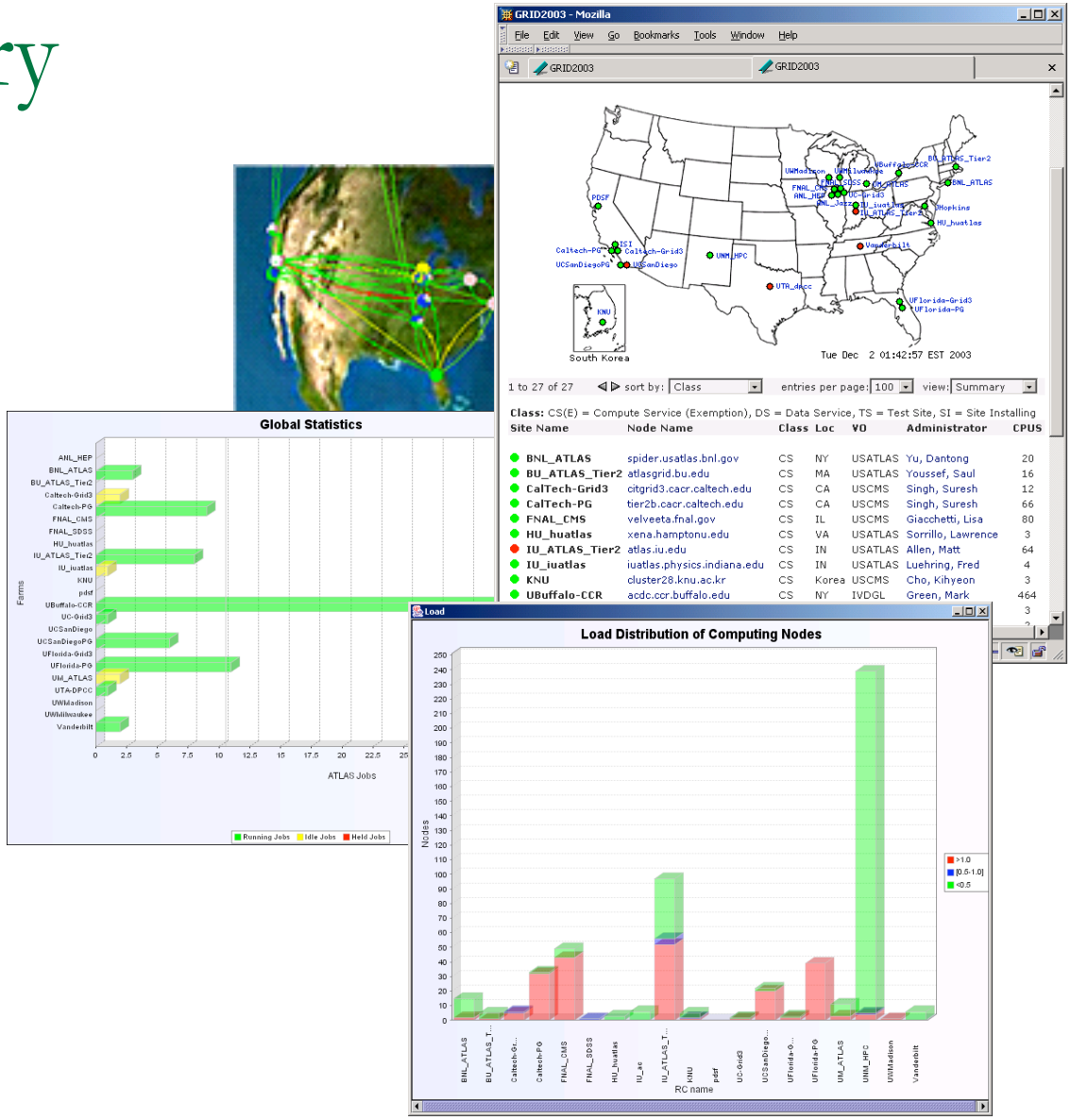
- *Globus* ***Meta Directory System (LDAP directory)***
- ***MonALISA***, Monitoring Agents in Large Integrated Service Architecture (Pub/Sub)
- ***MonALISA repository (WS/WAP)***
- **Ganglia** performance monitoring (Multicast/Hierarchical)
- **Job Monitoring System at the Advanced Center for Distributed Computing** (non invasive archive)
- The **Grid Site Status Cataloging System** at iGOC (human/automatic managed DB)





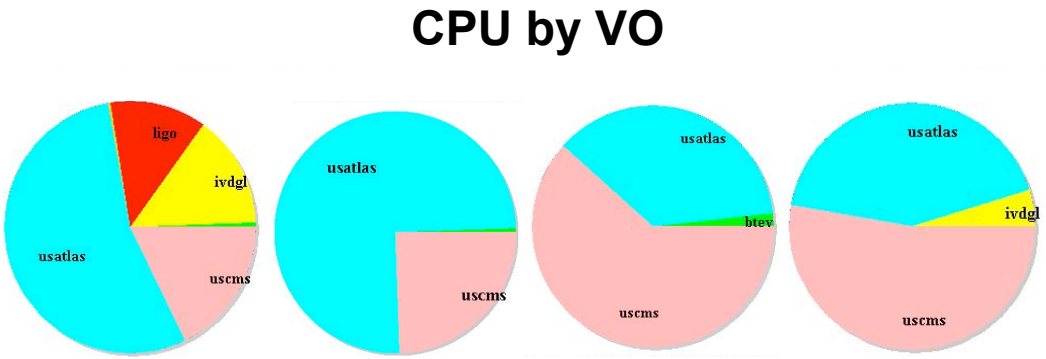
Grid Telemetry

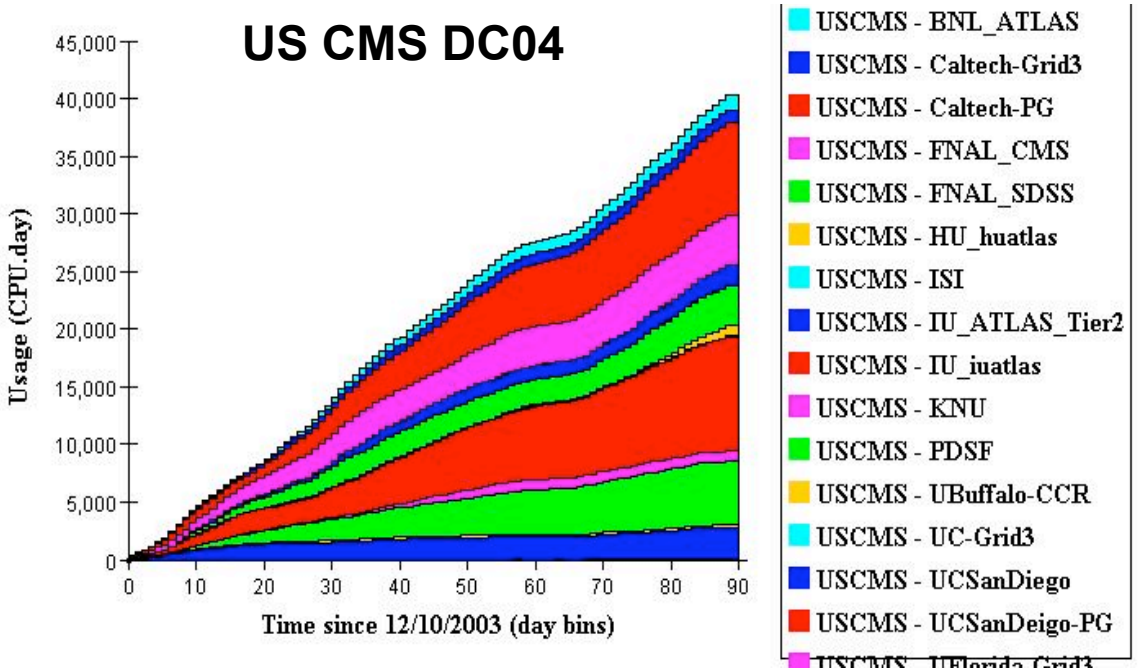
- Information
  - Site list
  - Test result
  - Load
  - Jobs running
  - Jobs queued
- Heterogeneous
- Redundant



Historic information analysis

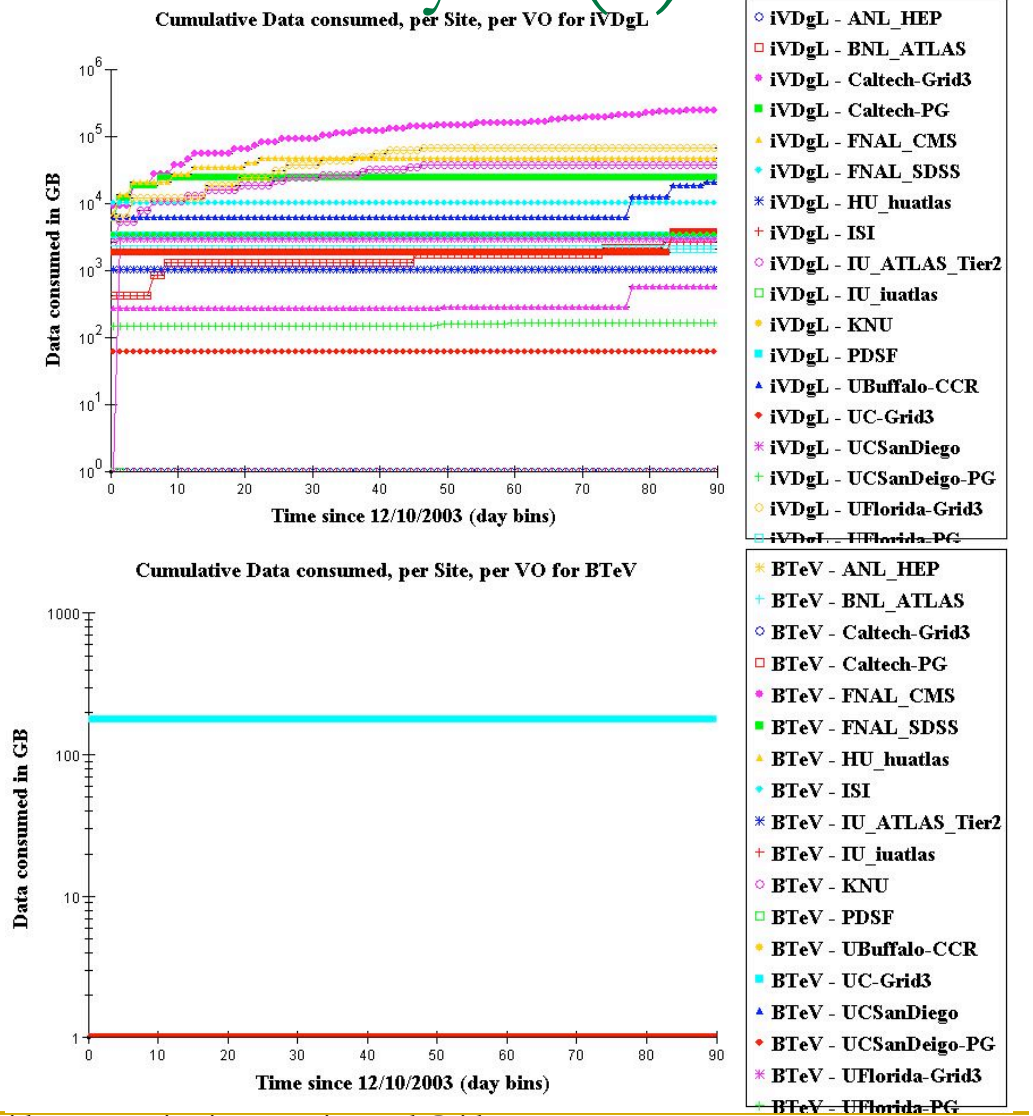
- Processed information
  - Aggregation
  - Filters
  - Analysis
- Provided CPUs
- CPU.hours consumed



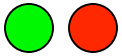


Historic information analysis(2)

- Usage patterns:
- Effective grid storage usage
- Dedicated resource usage



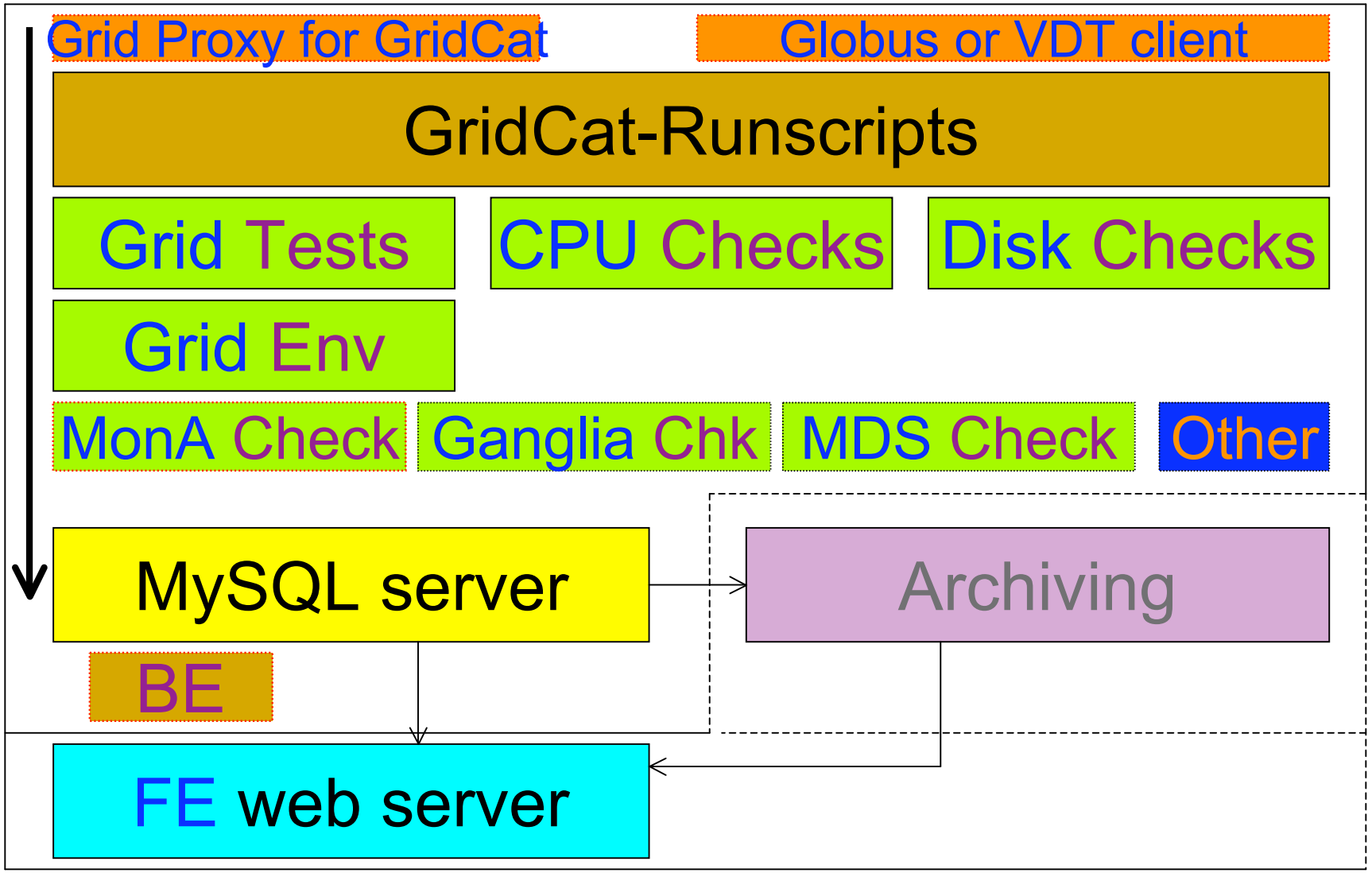
What is GridCat ?

- A Grid Site Status Cataloging System - A Web App.
- High level simple status map : 
- Computing Resource Information Collector/Presenter
  - Static and dynamic information about all sites
  - Simple grid status presentation on the web
- Identifies site readiness
- A web application easy to develop and deploy
- Displays disk space and CPU slots
- Parallel information collecting, storing, and archiving among sites (BE)
- Web pages: In templated html+php+js(FE)

Components

- Back-end
  - Grid Status Tests: GITS 1.4 (default/non-default ports)
  - Site $grid3, $tmp, $data, and $wntmp directory information
  - Site disk information (space used and space left)
  - Available CPU slots
  - Site status map generation (High Level Status Map)
  - Scripts can be used for complex grid failure mode analysis
- Front-end :
  - Displays default views
  - Choice of two maps
  - Monitoring status per site (MonAlisa, Ganglia, and MDS)
- Configuration-end (autoconf + Sitepop. Script)
  - Initial site population
    - Collect site information
    - Site physical/network/pixel address determination
  - Site parameter editing

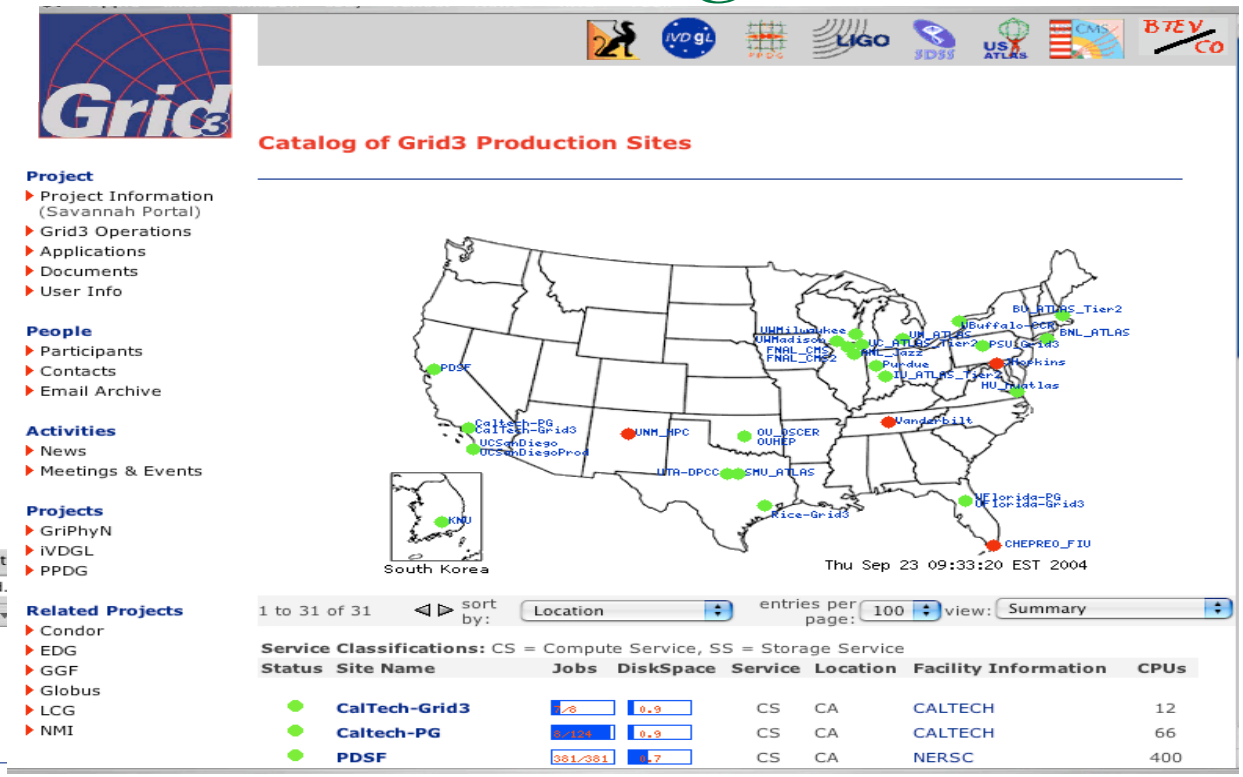
Flow of Operations

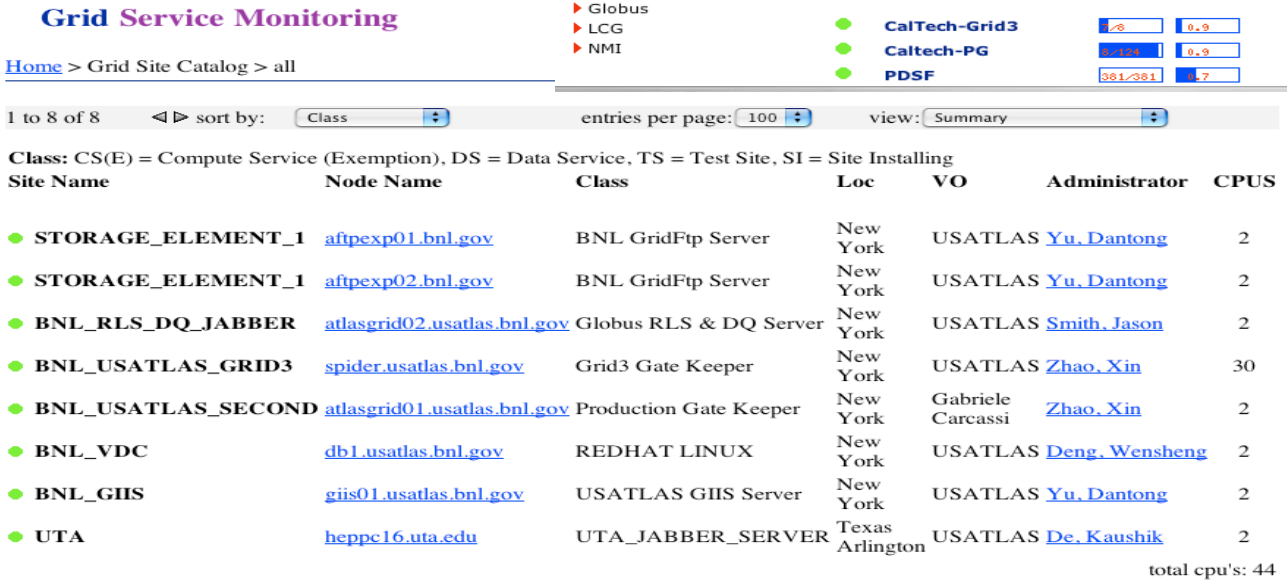


iGOC and Tier1 monitoring

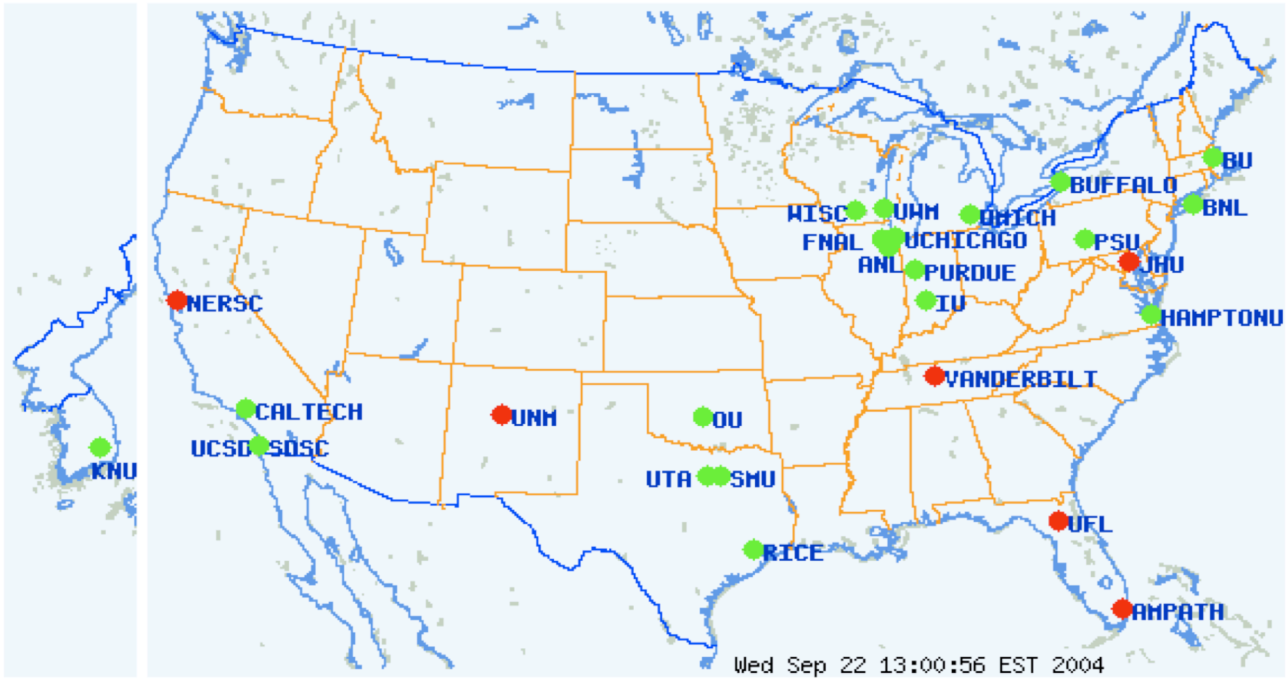
Grid3
iGOC

BNL Tier1
US ATLAS
DC2 Services





Gridcat In Action



- Status map on top
- Facility--> Sites
- Grid test results clickables
- Dynamic CPU/Disk info
- Optional views: different information
- Map : US-Korea map or Worlmap <- New release

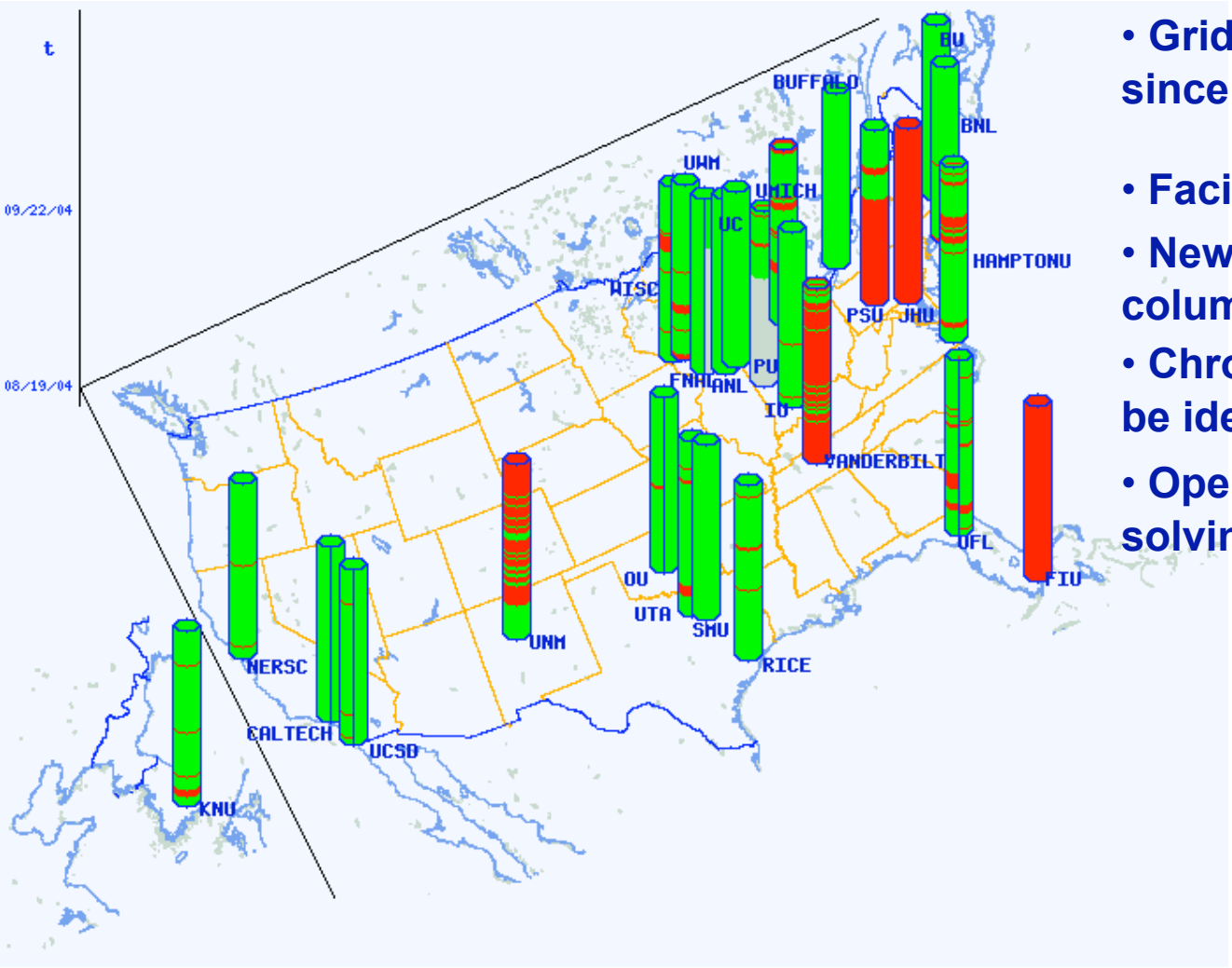
map: US-Korea

1 to 31 of 31 ◁ ▷ sort by: CPUs entries per page: 100 view: Summary

Service: CS(E) = Compute Service (Exemption), SS = Storage Service, LS = Login Service

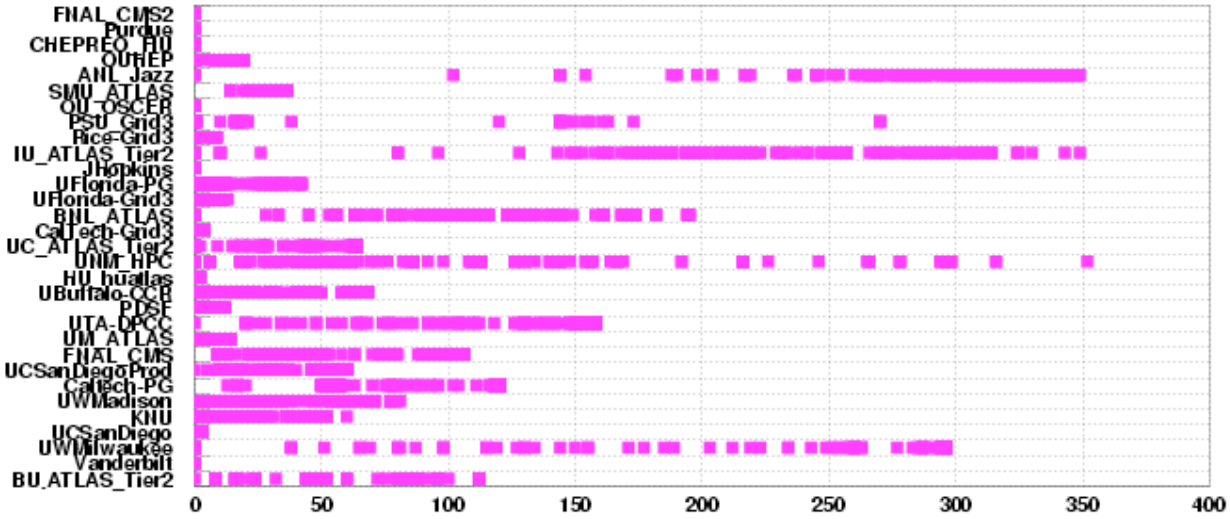
Status	Site Name	Jobs	Disks	Service	Loc	Facility	CPUs
●	IU_ATLAS_Tier2	175/376	1723/2154	CS	IN	IU	376
●	PSU_GRID3	85/312	903/1806	CS	PA	PSU	312
●	UWMilwaukee	296/306	443/492	CS	WI	UWM	306

GridCat Site History Analysis



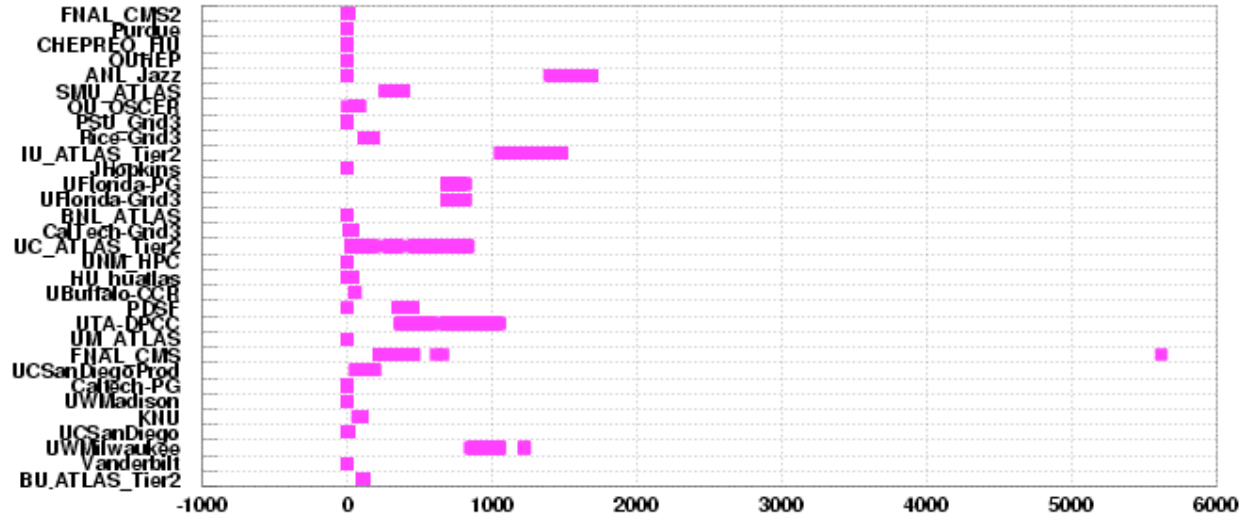
- **GridCat collected DB since 08/19/04**
- **Facility: >=1 sites**
- **New sites with grey column**
- **Chronic problems can be identified easily**
- **Operational/problem solving tool**

Gridcat DB Analysis



08/19/04 - 09/22/04
GridCat Grid3DB

Batch Jobs Running



data disk used

Conclusions and Future Development

Results

- Integrated diverse tools into a reliable monitoring infrastructure
- Collect, archive, and analyze user-defined metrics in an end-to-end grid computing system
  - resource usage by VO
- Testing in a production environment of considerable size
  - Grid3: 30 sites, ~3000 CPUs
  - feasibility of multi-VO Grids.
- Contributing to several applications
- Design and develop GridCat to display high level Grid status
  - Dynamic computing resource info monitoring at a glance
  - Map view to quickly identify and locate problems on the Grid
  - DB archiving for failure analyses
  - extensible for various site status cataloging purposes and operations

Work in progress

- Increase scalability, reliability
- Integrate monitoring, reservation and scheduling