



# 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 RODRIGUEZ (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

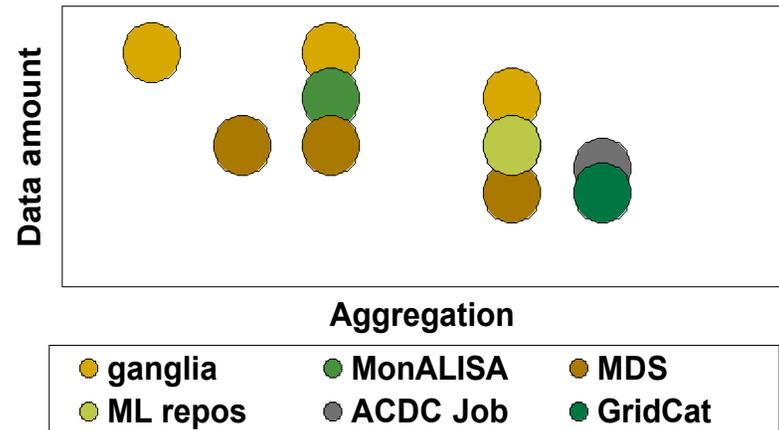
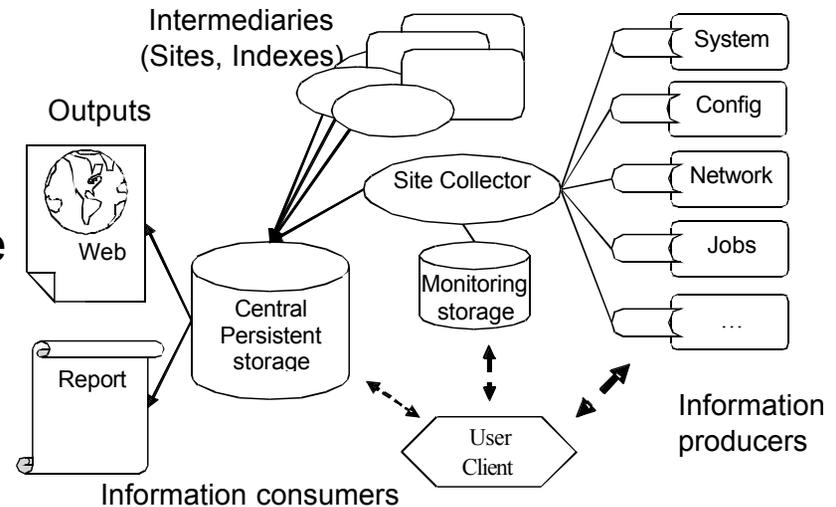
# Example of metrics specification

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

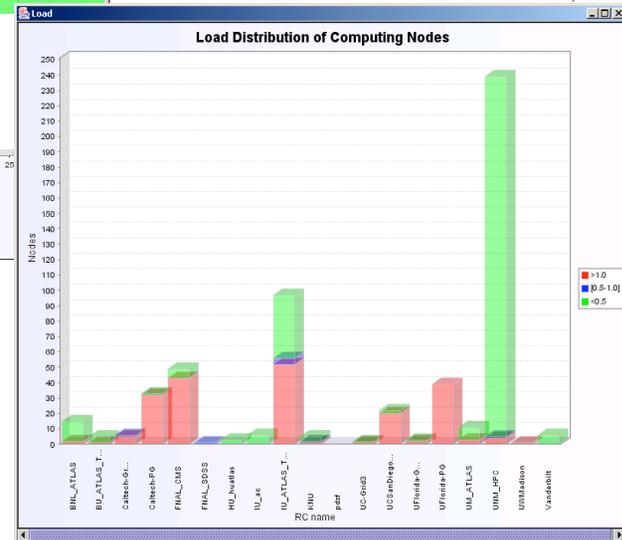
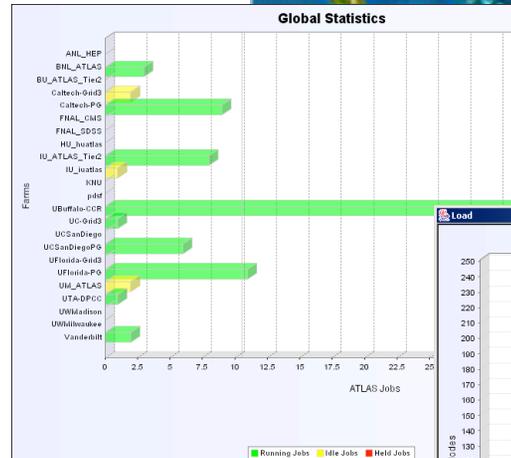
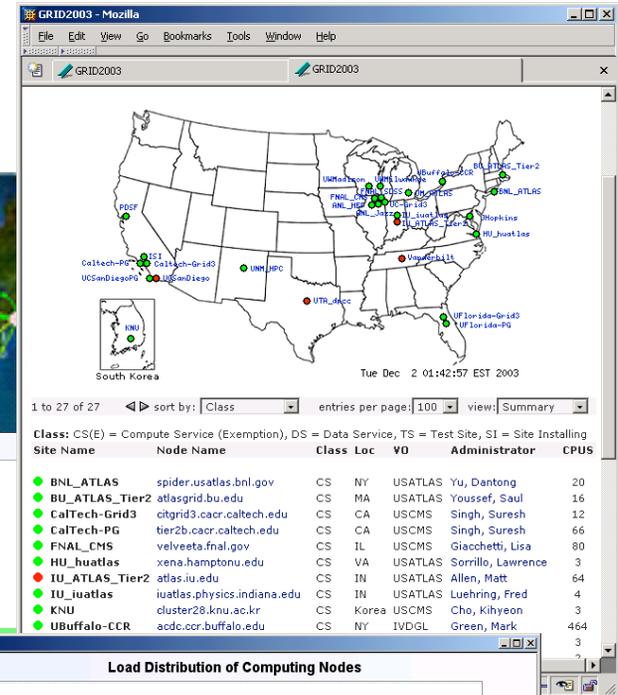
# Components of an Integrated Monitoring Framework

- **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

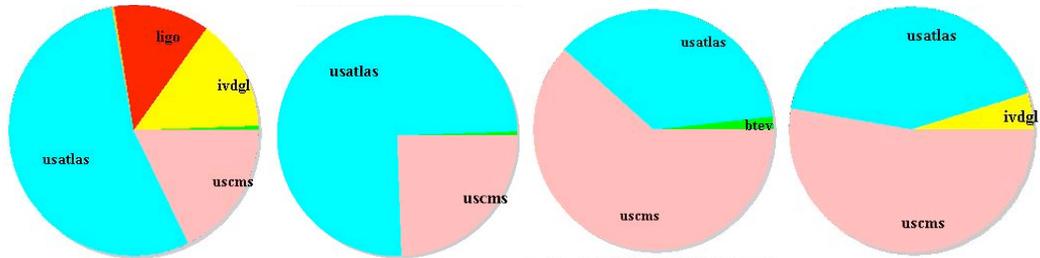
- Information
  - Site list
  - Test result
  - Load
  - Jobs running
  - Jobs queued
- Heterogeneous
- Redundant



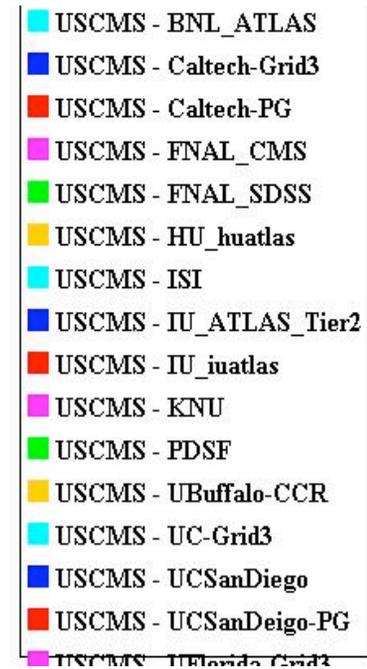
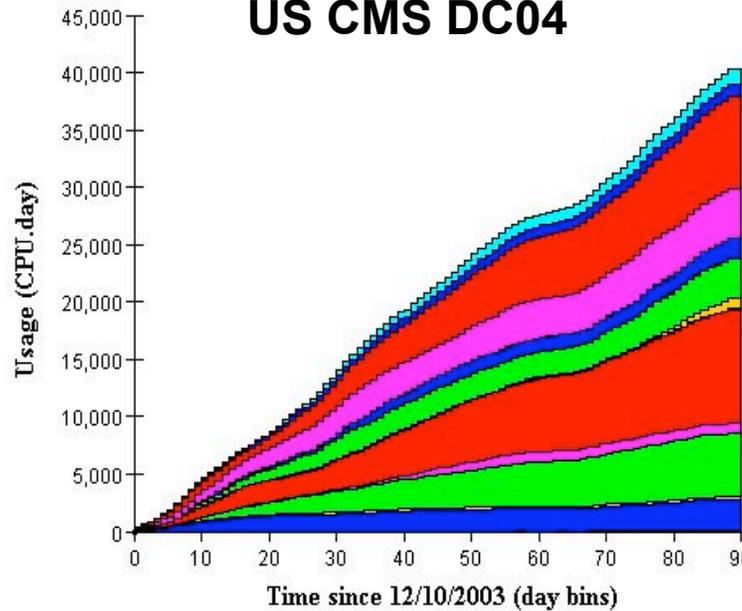
# Historic information analysis

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

CPU by VO



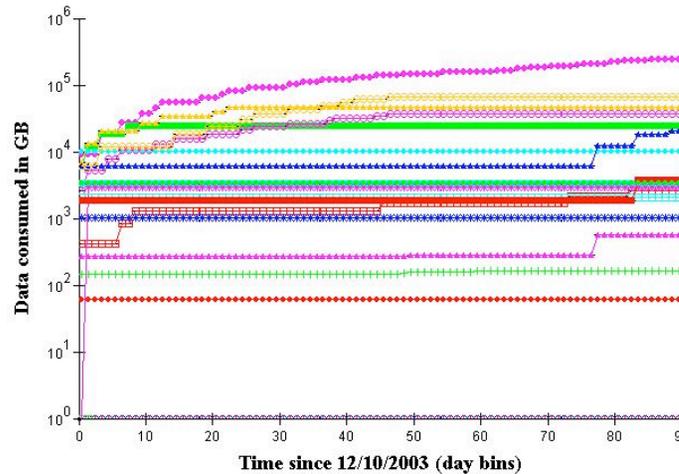
US CMS DC04



# Historic information analysis(2)

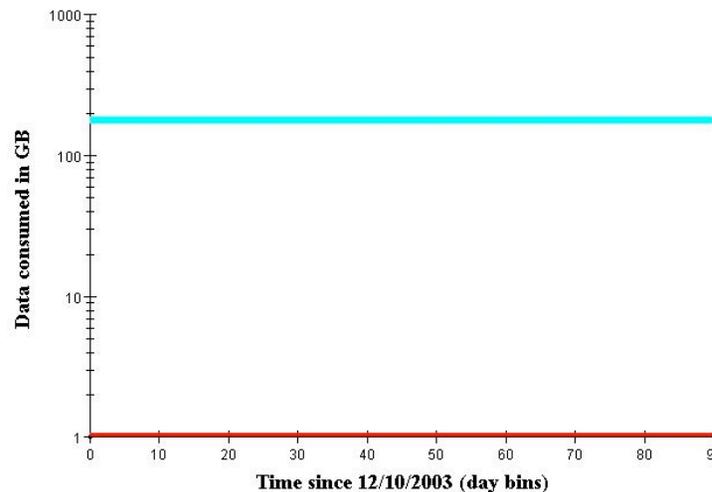
- Usage patterns:
- Effective grid storage usage
- Dedicated resource usage

Cumulative Data consumed, per Site, per VO for iVDgL



- ◇ iVDgL - ANL\_HEP
- iVDgL - BNL\_ATLAS
- ◇ iVDgL - Caltech-Grid3
- iVDgL - Caltech-PG
- ▲ iVDgL - FNAL\_CMS
- ◆ iVDgL - FNAL\_SDSS
- \* iVDgL - HU\_huAtlas
- + iVDgL - ISI
- ◇ iVDgL - IU\_ATLAS\_Tier2
- iVDgL - IU\_iuAtlas
- iVDgL - KNU
- iVDgL - PDSF
- ▲ iVDgL - UBuffalo-CCR
- ◆ iVDgL - UC-Grid3
- \* iVDgL - UCSanDiego
- + iVDgL - UCSanDiego-PG
- ◇ iVDgL - UFlorida-Grid3
- iVDgL - UFlorida-PG

Cumulative Data consumed, per Site, per VO for BTeV



- \* BTeV - ANL\_HEP
- + BTeV - BNL\_ATLAS
- ◇ BTeV - Caltech-Grid3
- BTeV - Caltech-PG
- ◇ BTeV - FNAL\_CMS
- BTeV - FNAL\_SDSS
- ▲ BTeV - HU\_huAtlas
- ◆ BTeV - ISI
- \* BTeV - IU\_ATLAS\_Tier2
- + BTeV - IU\_iuAtlas
- ◇ BTeV - KNU
- BTeV - PDSF
- BTeV - UBuffalo-CCR
- BTeV - UC-Grid3
- ▲ BTeV - UCSanDiego
- ◆ BTeV - UCSanDiego-PG
- \* BTeV - UFlorida-Grid3
- BTeV - UFlorida-PG

---

# 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

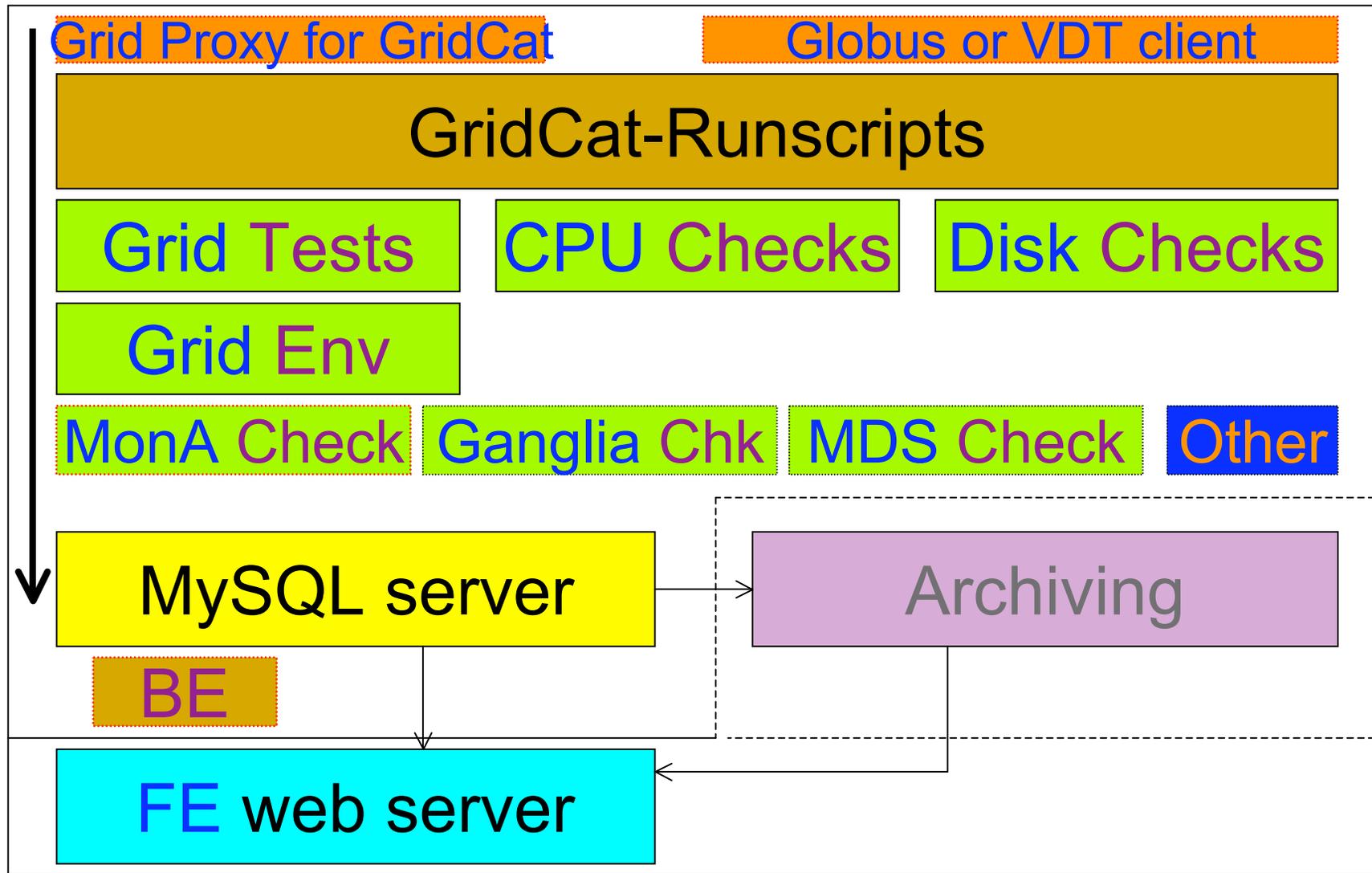
## ■ Front-end :

- ❑ 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



## Grid Service Monitoring

[Home](#) > [Grid Site Catalog](#) > all

1 to 8 of 8    sort by:     entries per page:     view:

Class: CS(E) = Compute Service (Exemption), DS = Data Service, TS = Test Site, SI = Site Installing

Site Name	Node Name	Class	Loc	VO	Administrator	CPUS
● STORAGE_ELEMENT_1	<a href="http://aftpexp01.bnl.gov">aftpexp01.bnl.gov</a>	BNL GridFtp Server	New York	USATLAS	<a href="#">Yu, Dantong</a>	2
● STORAGE_ELEMENT_1	<a href="http://aftpexp02.bnl.gov">aftpexp02.bnl.gov</a>	BNL GridFtp Server	New York	USATLAS	<a href="#">Yu, Dantong</a>	2
● BNL_RLS_DQ_JABBER	<a href="http://atlasgrid02.usatlas.bnl.gov">atlasgrid02.usatlas.bnl.gov</a>	Globus RLS & DQ Server	New York	USATLAS	<a href="#">Smith, Jason</a>	2
● BNL_USATLAS_GRID3	<a href="http://spider.usatlas.bnl.gov">spider.usatlas.bnl.gov</a>	Grid3 Gate Keeper	New York	USATLAS	<a href="#">Zhao, Xin</a>	30
● BNL_USATLAS_SECOND	<a href="http://atlasgrid01.usatlas.bnl.gov">atlasgrid01.usatlas.bnl.gov</a>	Production Gate Keeper	New York	Gabriele Carcassi	<a href="#">Zhao, Xin</a>	2
● BNL_VDC	<a href="http://db1.usatlas.bnl.gov">db1.usatlas.bnl.gov</a>	REDHAT LINUX	New York	USATLAS	<a href="#">Deng, Wensheng</a>	2
● BNL_GIIS	<a href="http://giis01.usatlas.bnl.gov">giis01.usatlas.bnl.gov</a>	USATLAS GIIS Server	New York	USATLAS	<a href="#">Yu, Dantong</a>	2
● UTA	<a href="http://heppc16.uta.edu">heppc16.uta.edu</a>	UTA_JABBER_SERVER	Texas Arlington	USATLAS	<a href="#">De, Kaushik</a>	2

total cpu's: 44

**Grid3**  
Catalog of Grid3 Production Sites

**Project**  
 ▶ Project Information (Savannah Portal)  
 ▶ Grid3 Operations  
 ▶ Applications  
 ▶ Documents  
 ▶ User Info

**People**  
 ▶ Participants  
 ▶ Contacts  
 ▶ Email Archive

**Activities**  
 ▶ News  
 ▶ Meetings & Events

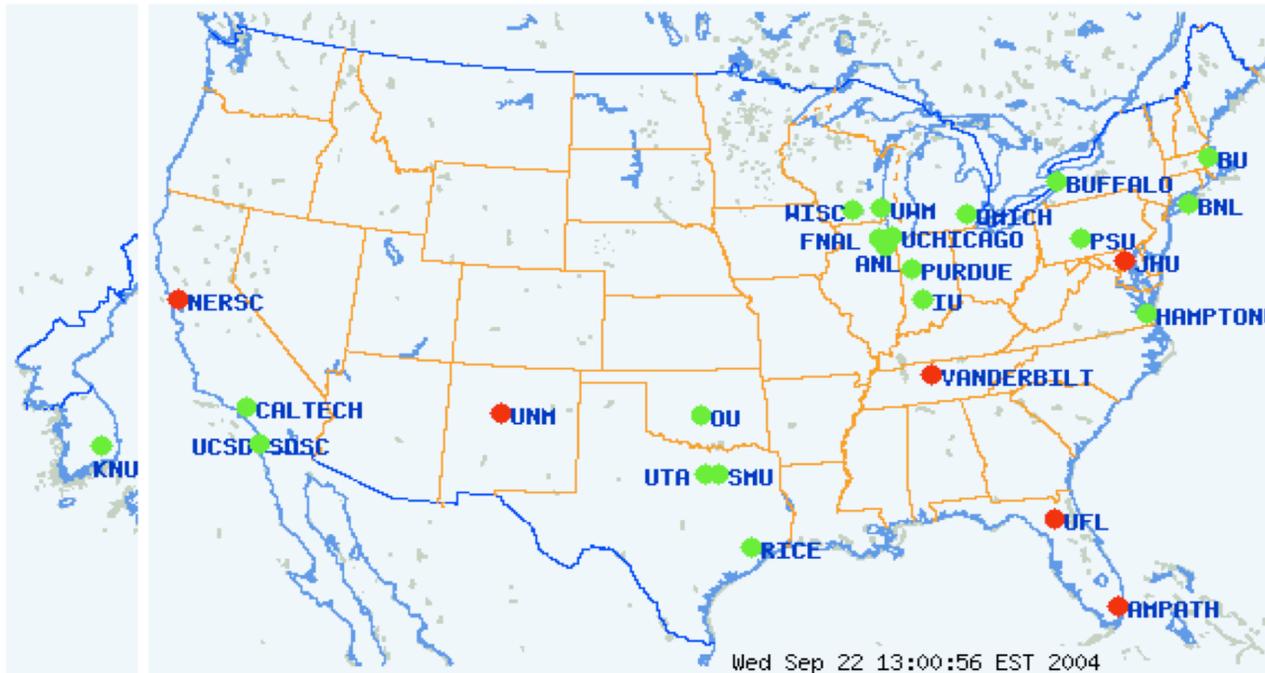
**Projects**  
 ▶ GriPhyN  
 ▶ IVDGL  
 ▶ PPDG

**Related Projects**  
 ▶ Condor  
 ▶ EDG  
 ▶ GGF  
 ▶ Globus  
 ▶ LCG  
 ▶ NMI

**Service Classifications:** CS = Compute Service, SS = Storage Service

Status	Site Name	Jobs	DiskSpace	Service	Location	Facility Information	CPUS
●	CalTech-Grid3	28	0.9	CS	CA	CALTECH	12
●	Caltech-PG	3174	0.9	CS	CA	CALTECH	66
●	PDSF	381/381	4.7	CS	CA	NERSC	400

# Gridcat In Action



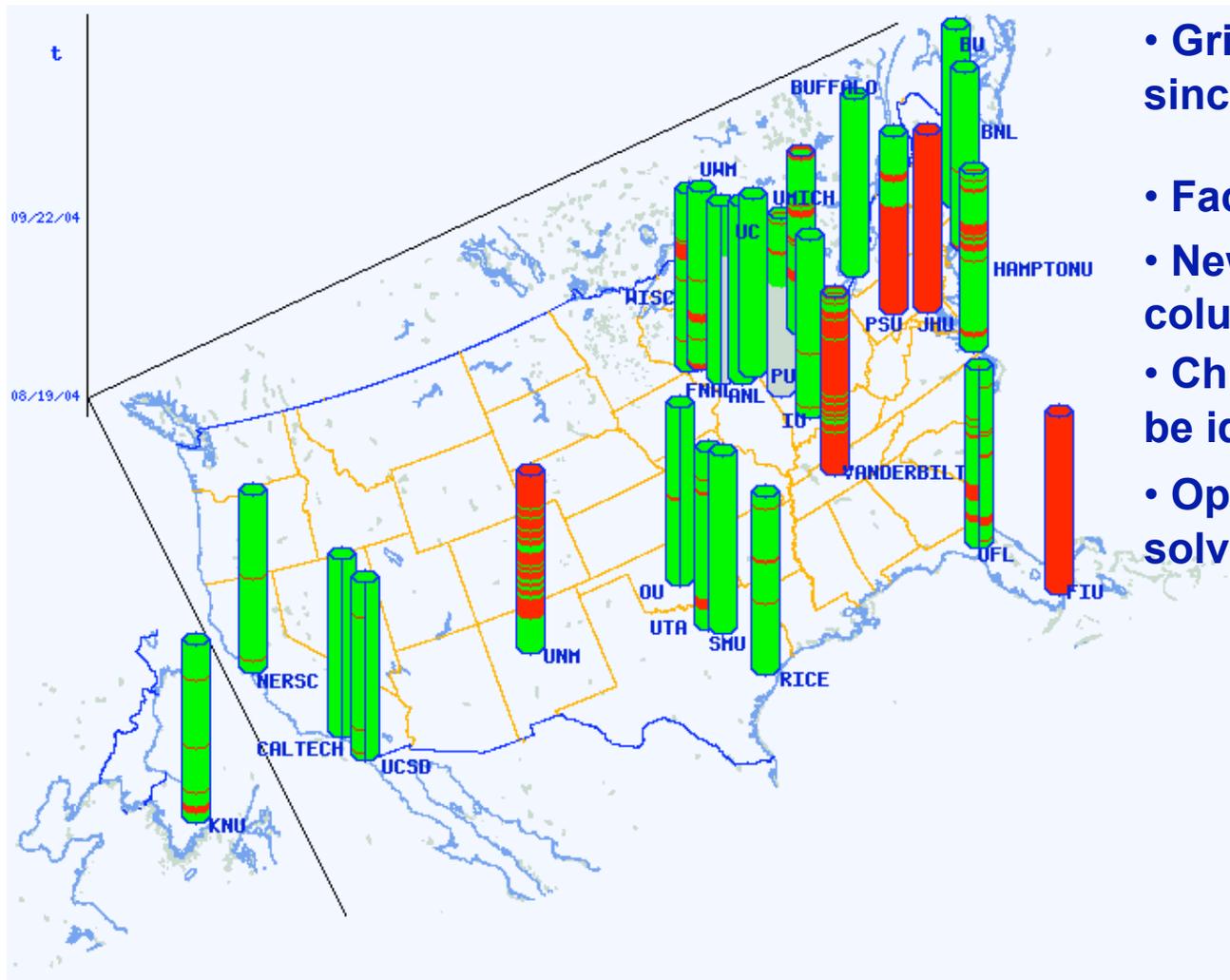
- Status map on top
- Facility--> Sites
- Grid test results clickables
- Dynamic CPU/Disk info
- Optional views: different information
- Map : US-Korea map or Worlmap <- New release

1 to 31 of 31   ◀▶ sort by: CPUs   entries per page: 100   view: Summary   map: US-Korea

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

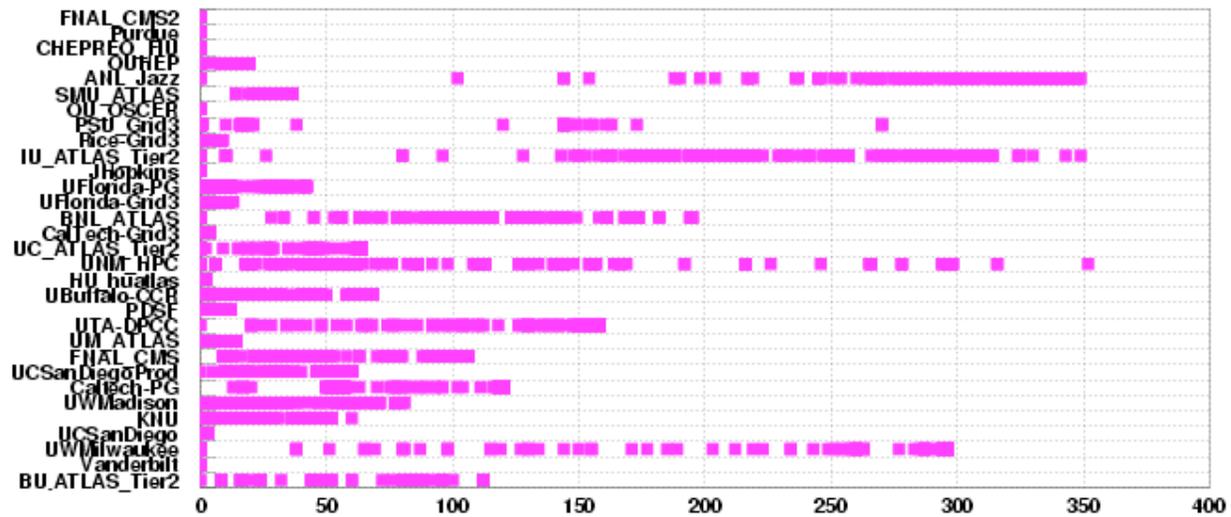
# GridCat Site History Analysis



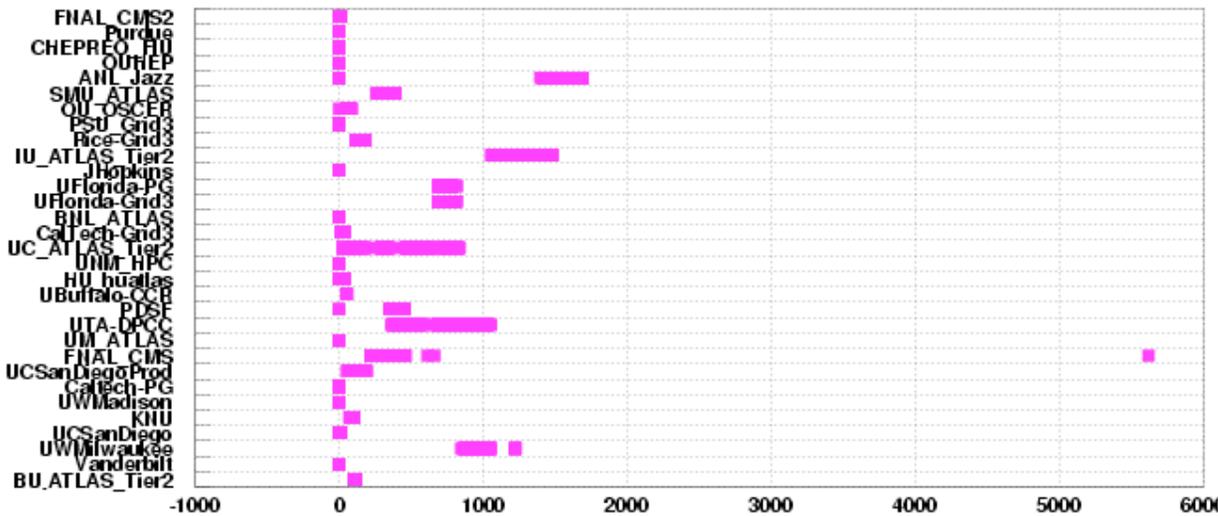
- GridCat collected DB since 08/19/04
- Facility:  $\geq 1$  sites
- New sites with grey column
- Chronic problems can be identified easily
- Operational/problem solving tool

# 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