

New Tools for Monitoring CMS Tier 3 LHC Data Analysis Centers

**In Cooperation With:
The Texas A&M Tier 3 CMS Grid Site on the Brazos Cluster**

Texas A&M University:

David Toback

Guy Almes

Steve Johnson

Vaikunth Thukral

Daniel Cruz

Sam Houston State University:

*** Joel Walker**

Jacob Hill

Michael Kowalczyk



Motivation

- **LHC data analysis is distributed globally across a grid of computing centers**
- **Requires cooperation of complex machines under diverse management**
- **Cutting edge technology always has many possible points of failure**
- **Enormous data transfer and storage needs**
- **Vast quantities of jobs must be run**
- **Site availability and optimization depend on accurate and current monitoring feedback**



Monitoring Goals

- **Monitor data transfers, data holdings, job status, and site availability**
- **Optimize for a single CMS Tier 3 site**
- **Provide a convenient and broad view**
- **Unify grid and local cluster diagnostics**
- **Give current status and historical trends**
- **Realize near real-time reporting**
- **Email administrators about problems**
- **Improve the likelihood of rapid resolution**



Implementation Goals

- **Host monitor online with public accessibility**
- **Provide rich detail without clutter**
- **Favor graphic performance indicators**
- **Merge raw data into compact tables**
- **Avoid wait-time for content generation**
- **Avoid multiple clicks and form selections**
- **Harvest plots and data with scripts on timers**
- **Automate email and logging of errors**



Email Alert System Goals

- **Operate automatically in background**
- **Diagnose and assign a “threat level” to errors**
- **Recognize new problems and trends over time**
- **Alert administrators of threats above threshold**
- **Remember mailing history and avoid “spam”**
- **Log all system errors centrally**
- **Provide daily summary reports**



**View the working development
version of the monitor online at:**

collider.physics.tamu.edu/tier3/mon/

**The next four slides provide a tour of the
website with actual graph and table samples**

(Plots and grid data are harvested from dashboard.cern.ch)



Monitoring Category I: Data Transfers to the Local Cluster

- Do we have solid links to other sites?
- Is requested data transferring successfully?
- Is it getting here fast?
- Are we passing load tests?

Production	PhEDEx Data Transfers	Load Test	Select	Hour	Day	Week	Month
Link Status	Linked Node	Link Status	Rate	Bytes	Files	Expired	Errors
Valid	T1_CH_CERN_Buffer	Valid	-	-	-	-	-
Valid	T1_DE_KIT_Buffer	Valid	-	-	-	-	-
Valid	T1_ES_PIC_Buffer	Valid	-	-	-	-	-
Valid	T1_FR_CCIN2P3_Buffer	Valid	-	-	-	-	-
Valid	T1_IT_CNAF_Buffer	Valid	15.9 MiB/s	1.3 TiB	577	638	1834
Valid	T1_TW_ASGC_Buffer	Valid	-	-	-	-	-
Valid	T1_UK_RAL_Buffer	Valid	-	-	-	-	-
Valid	T1_US_FNAL_Buffer	Valid	8.2 MiB/s	688.1 GiB	459	0	311
Valid	T2_IT_Pisa	Valid	-	-	-	-	-
Valid	T2_US_Caltech	Valid	-	-	-	-	-
Valid	T2_US_Florida	Valid	-	-	-	-	-
Valid	T2_US_MIT	Valid	-	-	-	-	-
Valid	T2_US_Nebraska	Valid	-	-	-	-	-
Valid	T2_US_Purdue	Valid	1.9 MiB/s	159.5 GiB	16	0	175
Valid	T2_US_UCSD	Valid	-	-	-	-	-
Valid	T2_US_Vanderbilt	Valid	-	-	-	-	-
Valid	T2_US_Wisconsin	Valid	-	-	-	-	-
Valid	T3_US_Colorado	Valid	-	-	-	-	-
Valid	T3_US_Rice	Valid	-	-	-	-	-
Valid	T3_US_TTU	Excluded	-	-	-	-	-
Totals			26.0 MiB/s	2.1 TiB	1052	638	2320



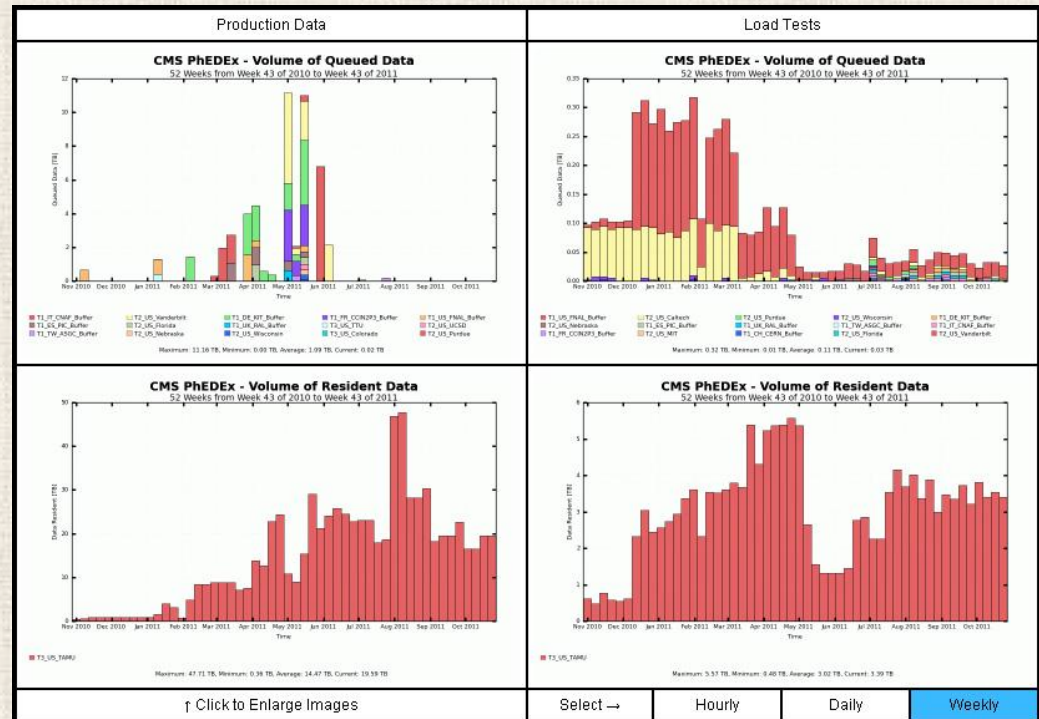
Monitoring Category II:

Data Holdings on the Local Cluster

- How much data have we asked for? Actually received?
- Are remote storage reports consistent with local reports?
- How much data have users written out?
- Are we approaching disk quota limits?

HEPX Disk Store Usage	
Directory	Bytes
→ PhEDEx Monte Carlo	19.1 TiB
→ PhEDEx CMS Data	539.8 GiB
→ PhEDEx Load Tests	49.8 GiB
↓ User Output	3.9 TiB
Indara Suarez	2.6 TiB
Diamond	529.7 GiB
Alexx Perloff	475.1 GiB
Roy J. Montalvo	146.0 GiB
Tai Sakuma	58.6 GiB
Alfredo Gurrola	45.4 GiB
Ji Eun Kim	10.2 GiB
Vaikunth Thukral	43.7 MiB
Daniel J. Cruz	6.9 MiB
→ Miscellaneous	556.0 KiB
Total	23.5 TiB

↑ Click to Expand or Collapse Table



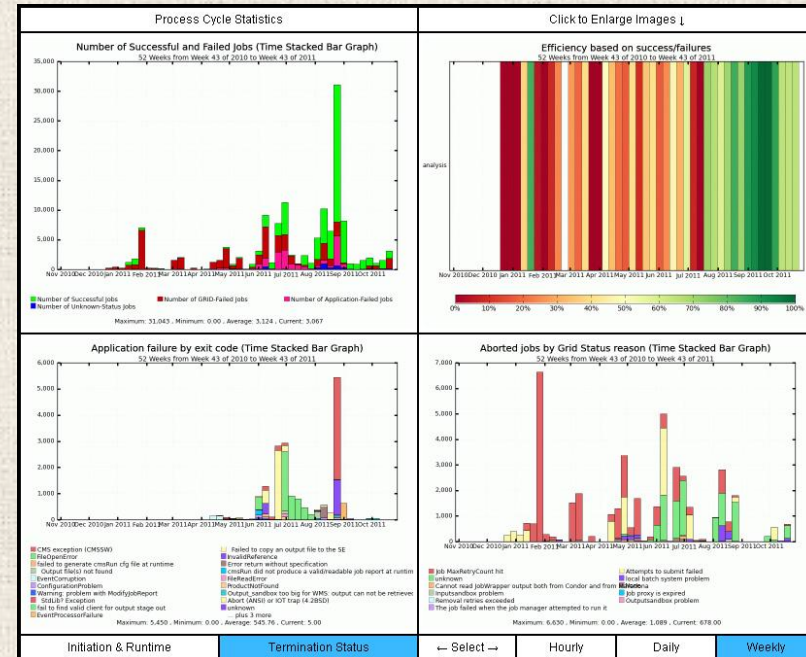
Monitoring Category III: Job Status of the Local Cluster

- How many jobs are running? Queued? Complete?
- What percentage of jobs are failing? For what reason?
- Are we making efficient use of available resources?
- Which users are consuming resources? Successfully?
- How long are users waiting to run?

Current TORQUE Batch Queue Job Status (2011-10-20 13:20 UTC)							
User Name	Queue	Run Status	Processors	Memory	CPU Hours	Run Hours	Queue Hours
Totals	-	- / -	-	-	-	-	-
No Jobs Are Currently Queued			Select Sorting By →		User	Queue	Status

Current Condor Batch Queue Job Status (2011-10-20 13:20 UTC)							
User Name	Universe	Run Status	Processors	Memory	CPU Hours	Run Hours	Queue Hours
Totals	-	- / -	-	-	-	-	-
No Jobs Are Currently Queued			Select Sorting By →		User	Universe	Status

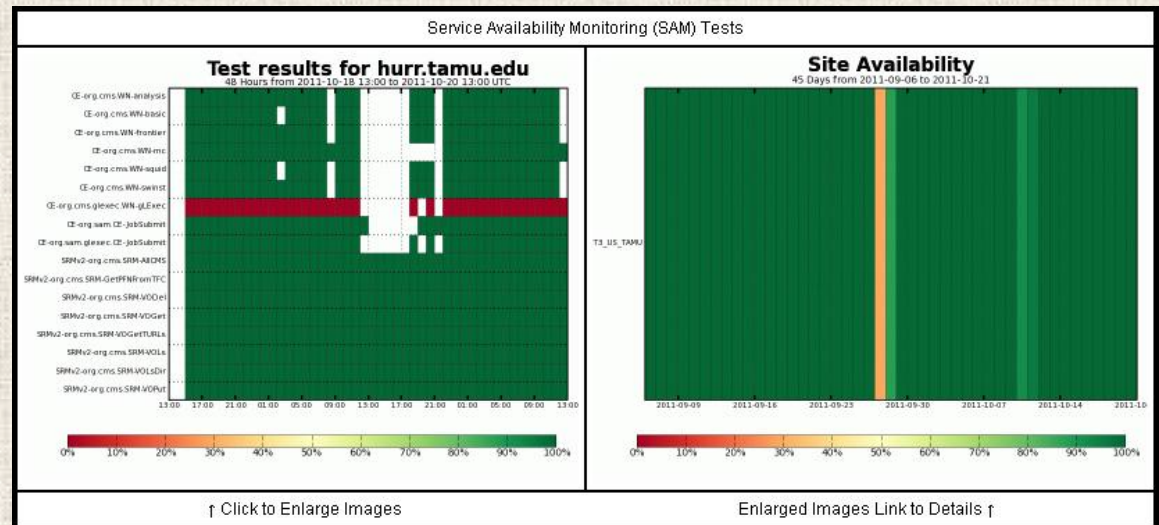
Recent Brazos Cluster Job Activity by User							
User Name	Pending	Running	Terminated	App Failed	Grid Failed	CPU Usage	Run Hours
Benedet	0	0	356	1.4 %	0.0 %	94.0 %	921.2
Kadir Ocalan	21	0	0	0.0 %	0.0 %	100.0 %	0.0
Totals	21	0	356	1.4 %	0.0 %	94.0 %	921.2
↑ Click User Row for Job Details			Select →		Hour	Day	Week



Monitoring Category IV: Site Availability

- Are we passing tests for connectivity and functionality?
- What has our uptime been for the day? Week? Month?
- Are test jobs which follow “best practices” successful?

Service Availability Percentage		
Day	Week	Month
100 %	100 %	97 %



↓ Output Host	CRAB Analysis Test Suite (CATS) Completed Job Status			
Grid Client →	gLite		CondorG	
Output Size →	Small	Large	Small	Large
Brazos Cluster	Pass:30 Fail:0 Other:0 2011-10-16 20:33 UTC	Pass:30 Fail:0 Other:0 2011-10-16 20:33 UTC	Pass:0 Fail:0 Other:20 2011-10-18 20:59 UTC	Pass:30 Fail:0 Other:0 2011-10-16 20:22 UTC
Fermi National Laboratory	Pass:30 Fail:0 Other:0 2011-10-16 20:34 UTC	Pass:30 Fail:0 Other:0 2011-10-16 20:34 UTC	Pass:0 Fail:0 Other:20 2011-10-18 20:59 UTC	Pass:30 Fail:0 Other:0 2011-10-16 20:23 UTC

↑ Test Results Link to Job Details ↑

Ongoing Work

- **Enhancement of content and real-time usability**
- **Beta testing on A&M's young Tier 3 site**
- **Vetting for robust operation and completeness**
- **Expanding implementation of the alert layer**
- **Development of suitable documentation**
- **Distribution to other University Tier 3 sites**
- **Seeking out additional funding sources**



Conclusions

- **New monitoring tools are uniquely convenient and site specific, with automated email alerts**
- **Early deployment at Texas A&M has already improved rapid error diagnosis and resolution**
- **We are working on securing funding to bring the monitor to other University Tier 3 sites**



**We acknowledge the
Norman Hackerman
Advanced Research Program
and the
Department of Energy
ARRA Program
for prior support in funding**

