Towards a global monitoring system for CMS computing

Lothar A. T. Bauerdick
Andrea P. Sciabà

Computing in High Energy and Nuclear Physics (CHEP) 2012
21-25 May, 2012
New York, NY, USA

Outline

- Introduction
- CMS monitoring areas
  - Service monitoring
  - Facility monitoring
  - Global overview
  - Historical accounting
- The CMS Monitoring Task Force (MTF)
  - Purpose and goals
  - Work plan and activities
- Evolution
- Final thoughts

Introduction

- Operation of the CMS computing system requires complex monitoring to cover all its aspects
- A multitude of monitoring systems are used, developed inside and outside of CMS
- Existing monitoring was successful in allowing CMS to operate and understand its computing system
- Still, there is a need to
  - Adapt to changes in the computing tools
  - Rationalize effort and tools
  - Have a common vision
- The purpose of this talk is
  - To give an overview of CMS monitoring
  - To show its current and future evolution

Monitoring areas

Area	Description	Users
Service monitoring	Provided by each service Describes service status and health	Service operators
Facility monitoring	Status of distributed computing infrastructure: health, used and available resources	Computing operators Sites
Global overview	Status of the computing activities across the infrastructure	Production coordinators Physicists
Historical accounting	Historical trends in activity levels, resource usage and health	Computing management

- The areas are quite distinct conceptually but the same monitoring tool may address more than one
  - E.g. the PhEDEx monitoring addresses the four of them

Service monitoring

- Each service should come with its own monitoring
  - To allow operators to understand if the service is working and what it is doing
  - Services providing their monitoring are for example
    - PhEDEx
    - CRAB server
    - WMAgent
    - Frontier
- Other services provide very little in terms of native monitoring
- All services rely quite heavily on Lemon and SLS (developed by CERN IT)
  - Lemon for host-centric monitoring
  - SLS for service-centric monitoring

PhEDEx and CRAB server

PhEDEX reports a detailed picture of its status
It sets an example for other services!

Infrastructure Agents			
Node	FileRouter	FileIssue	FilePump
PhEDEx Central	UP	UP	UP

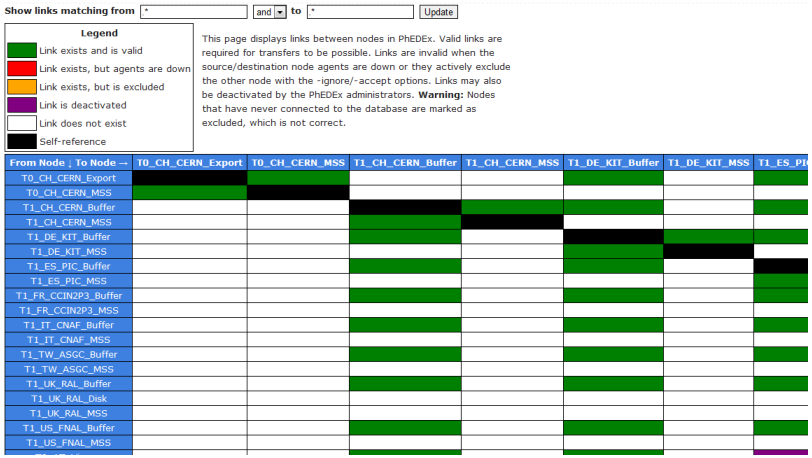
Workflow Agents						
Node	RequestAllocator	BlockAllocator	BlockMonitor	BlockDelete	BlockActivate	BlockDeactivate
PhEDEx Central	UP	UP	UP	UP	UP	UP

Support Agents					
Node	BlockDownloadVerifyInjector	InfoFileSize	InfoStatesClean	InvariantMonitor	PerfMonitor
PhEDEx Central	UP (0h59 ago)	DOWN (0h36 ago)	UP	UP	UP

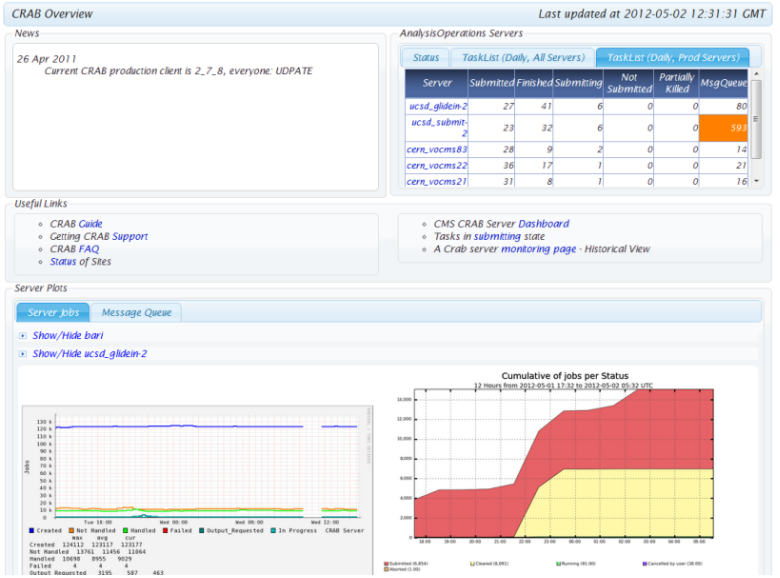
Site Agents						
Node	FileDownload	FileExport	FileStager	FileRemove	BlockDownloadVerify	Watchdog
T0_CH_CERN_Export		UP (2/2 agents)	UP	UP	UP	UP (2/2 agents)
T0_CH_CERN_MSS	UP	UP		UP		
T1_CH_CERN_Buffer	UP (0h15 ago) (2/2 agents)	UP (2/2 agents)	UP	UP	UP	UP (0h29 ago)
T1_CH_CERN_MSS	UP	UP		UP		
T1_DE_KIT_Buffer	UP (0h19 ago)	UP	UP (0h14 ago)	UP (0h11 ago)	UP	UP
T1_DE_KIT_MSS	UP	UP		UP (0h11 ago)	UP	
T1_ES_PIC_Buffer	UP (3/3 agents)	UP	UP	UP	UP	
T1_ES_PIC_MSS	UP	UP		UP		
T1_FR_CCIN2P3_Buffer	UP	UP (0h44 ago)	UP	UP (0h10 ago)	UP (2/2 agents)	

PhEDEx agent status

Native CRAB server status is less detailed
Users better served by an aggregation page collecting info from all servers



PhEDEx link status



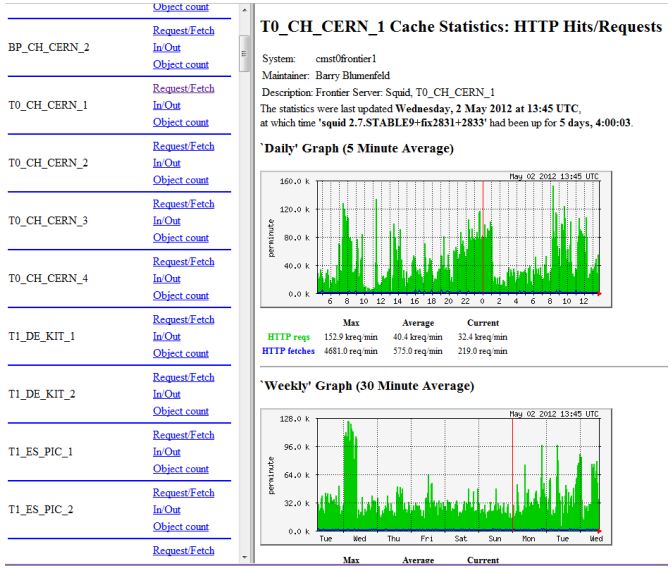
WMAgent and Frontier

- WMAgent is the new production system since 2011
- All information is defined in various databases and can be retrieved via REST API
- Monitoring was very basic at the beginning, development fully driven by operations

Component heartbeat

name	pid	worker_name	ago	alarm	warning
ErrorHandler	18472	ErrorHandlerPoller	162	5838	
DashboardReporter	18367	DashboardReporterPoller	120	5880	
PhEDExInjector	11992	PhEDExInjectorPoller	73	5927	
TaskArchiver	18510	TaskArchiverPoller	66	5934	
JobTracker	18457	JobTrackerPoller	56	5944	
JobStatusLite	18464	StatusPoller	55	5945	
JobCreator	18423	JobCreatorPoller	54	5946	
JobArchiver	18496	JobArchiverPoller	52	5948	
JobAccountant	18413	JobAccountantPoller	26	5974	
WorkQueueManager	18376	WorkQueueManagerLocationPoller	15	5985	
RetryManager	18486	RetryManagerPoller	11	5989	
DBSUpload	18386	DBSUploadPoller	1	5999	
JobSubmitter	18436	JobSubmitterPoller	0	6000	

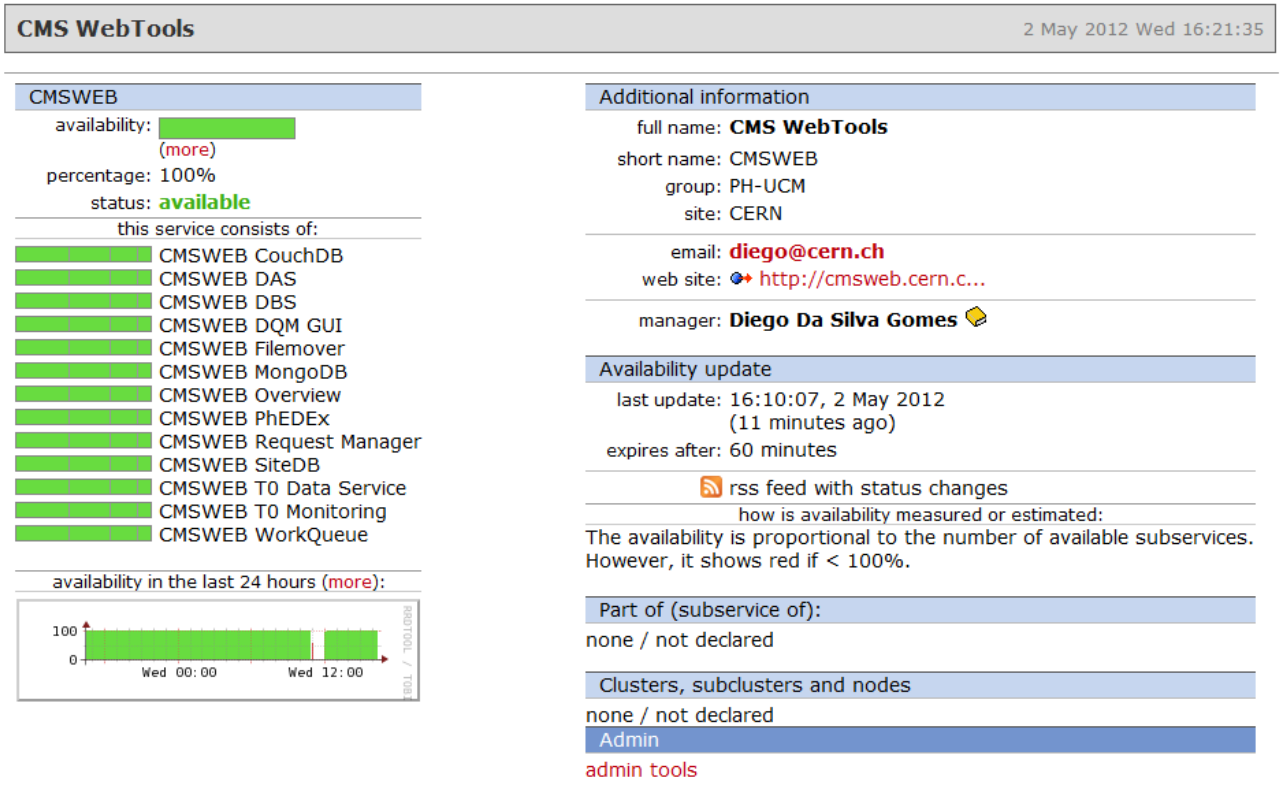
WMAgent heartbeat monitor



- Frontier monitoring relies on probing Squid servers with SNMP requests and on log file parsing

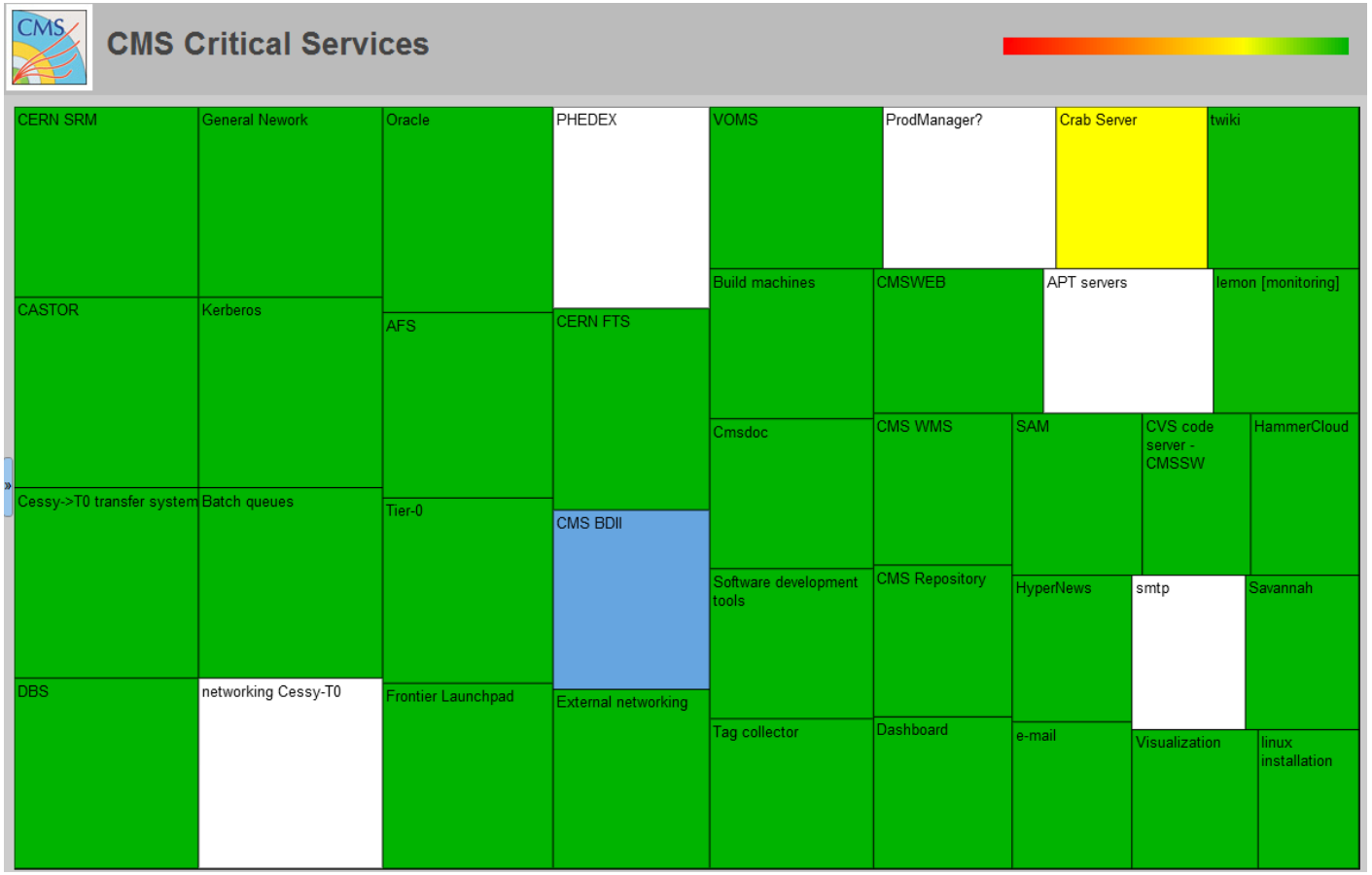
Lemon and SLS

- Substantial effort was put in ensuring that the CMS services take advantage of Lemon monitoring and alarms
- SLS is widely used at CERN IT and by LHC experiments for service monitoring
- Very convenient as interface as it provides a uniform and user-friendly interface CERN-wide



Critical services

- The “ultimate” service monitoring page, developed by CERN IT
- Big picture of the status of all services provided or used by CMS
- Based on SLS information, could be used by any experiment

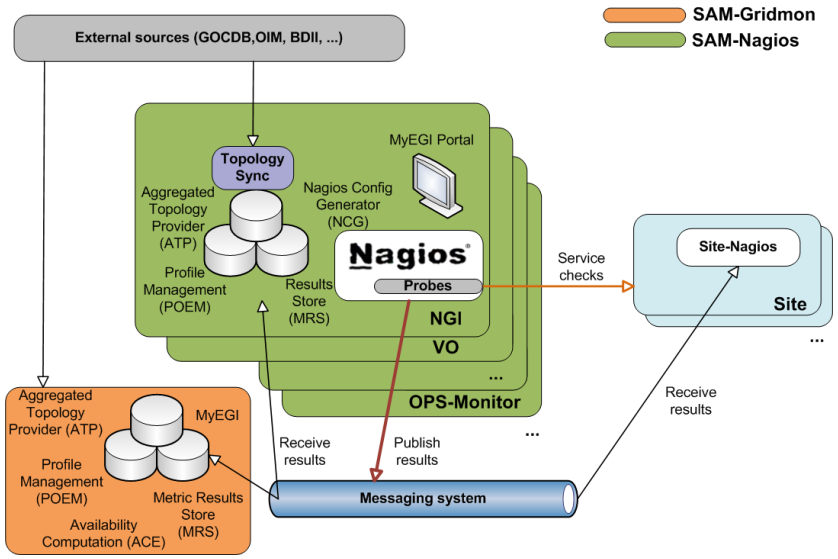


Facility monitoring

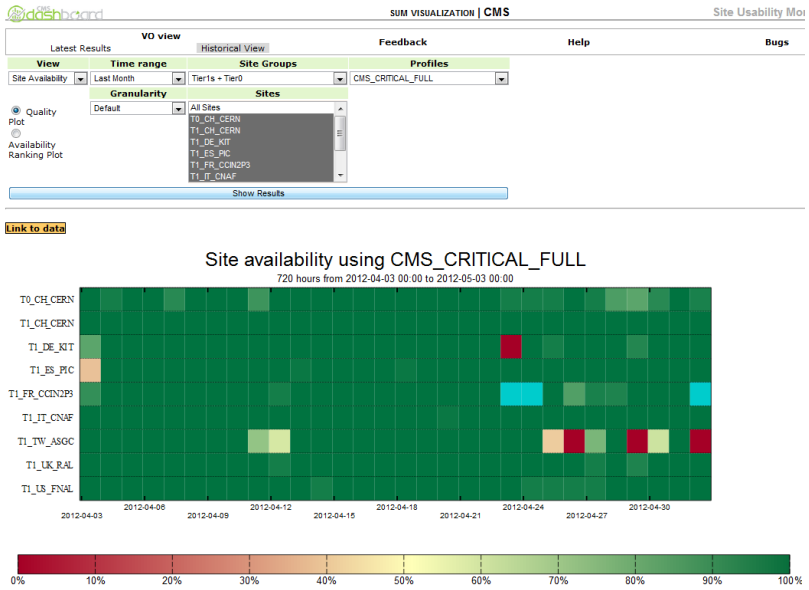
- This area covers the distributed computing infrastructure
- Health of the sites and their services
  - Service Availability Monitor (SAM)
  - HammerCloud
  - PhEDEx monitoring (transfer quality)
  - CMS Site Readiness
- Resource utilization
  - Tier-1 farm monitoring (HappyFace)

SAM

- The framework used by WLCG and EGI and developed by IT to run functional tests on site services
- Used by CMS since 2007 to run CMS-specific tests
- Completely rewritten in 2009, submission based on Nagios
- Moved to production for CMS in 2011-2012

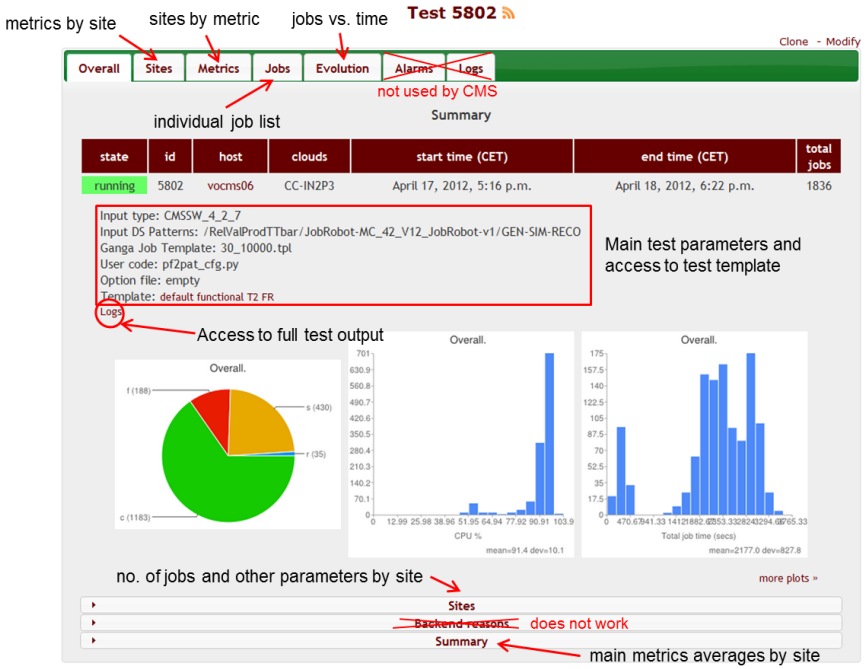


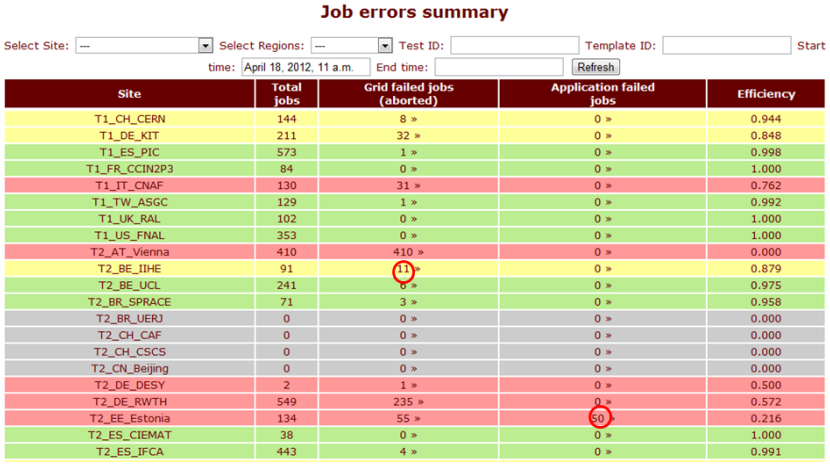
- Web visualization (SUM) developed by Dashboard team
- Used by the four LHC experiments



HammerCloud

- End-to-end testing tool for Grid sites developed by CERN IT
  - Used by ATLAS, CMS and LHCb
- Used for functional tests and stress tests
- Highly configurable, powerful UI
- Convenient for validation and performance studies
  - Tens of metrics collected, stored and plotted

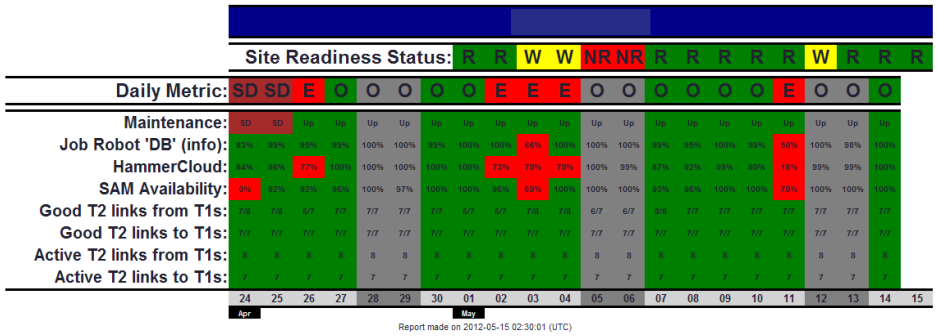


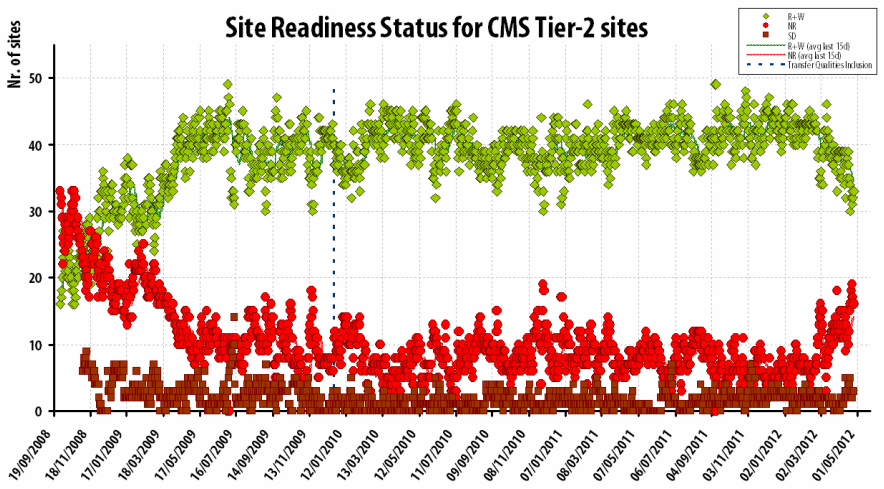


Site	Total jobs	Grid failed jobs (aborted)	Application failed jobs	Efficiency
T1_CH_CERN	144	8 »	0 »	0.944
T1_DE_KIT	211	32 »	0 »	0.848
T1_ES_PIC	573	1 »	0 »	0.998
T1_FR_CCIN2P3	84	0 »	0 »	1.000
T1_IT_CNAF	130	31 »	0 »	0.762
T1_TW_ASGC	129	1 »	0 »	0.992
T1_UK_RAL	102	0 »	0 »	1.000
T1_US_FNAL	353	0 »	0 »	1.000
T2_AT_Vienna	410	410 »	0 »	0.000
T2_BE_IIHE	91	11 »	0 »	0.879
T2_BE_UCL	241	6 »	0 »	0.975
T2_BR_SPRACE	71	3 »	0 »	0.958
T2_BR_UERJ	0	0 »	0 »	0.000
T2_CH_CAF	0	0 »	0 »	0.000
T2_CH_CSCS	0	0 »	0 »	0.000
T2_CN_Beijing	0	0 »	0 »	0.000
T2_DE_DESY	2	1 »	0 »	0.500
T2_DE_RWTH	549	235 »	0 »	0.572
T2_EE_Estonia	134	55 »	50 »	0.216
T2_ES_CIEMAT	38	0 »	0 »	1.000
T2_ES_IFCA	443	4 »	0 »	0.991

Site Readiness

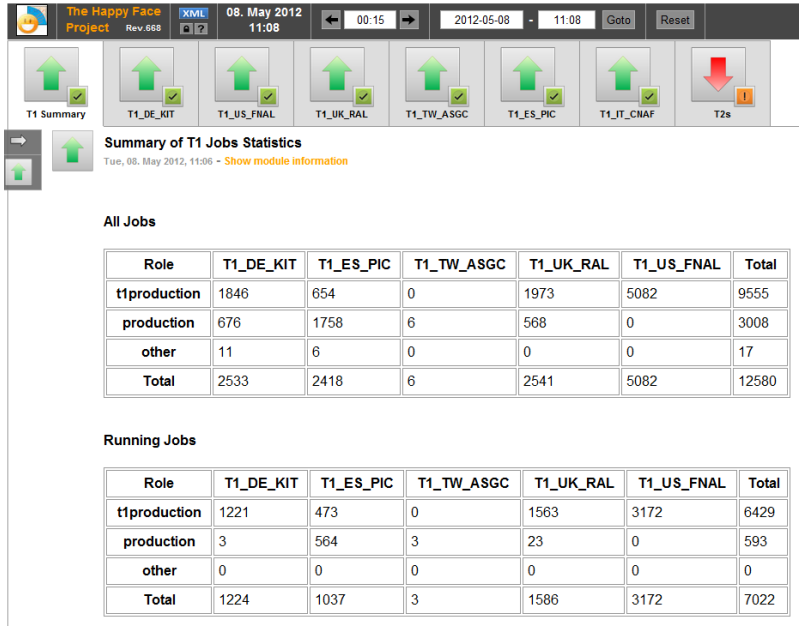
- An algorithm and a tool developed by CMS to combine different site quality metrics into a single estimator
  - Site availability, HammerCloud success rate, data transfer quality, site downtime information
- Instrumental in improving the overall quality of the infrastructure
- Some ideas reused by ATLAS
- Recently integrated in the CMS central service infrastructure at CERN





Tier-1 batch monitoring

- Provides a unique entry point to the batch system information for all Tier-1's
  - Sites provide standardized XML files with information on current jobs
  - Visualization via HappyFace, a monitoring portal used by German sites
- Eliminated the need to use 7 different monitoring pages



Summary of T1 Jobs Statistics

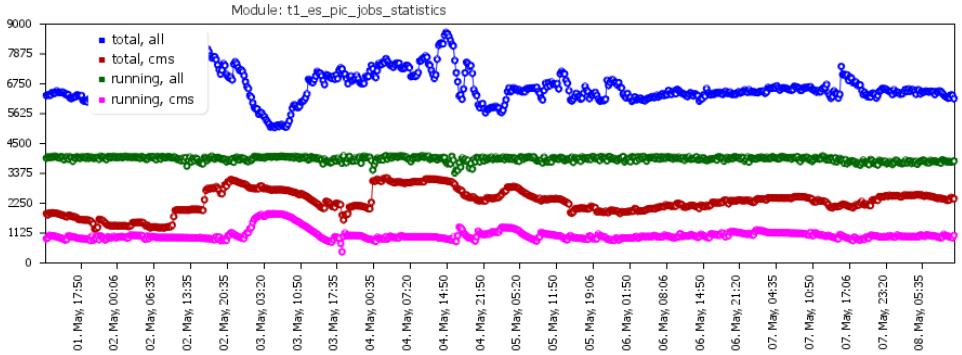
Tue, 08. May 2012, 11:06 - Show module information

All Jobs

Role	T1_DE_KIT	T1_ES_PIC	T1_TW_ASGC	T1_UK_RAL	T1_US_FNAL	Total
t1production	1846	654	0	1973	5082	9555
production	676	1758	6	568	0	3008
other	11	6	0	0	0	17
Total	2533	2418	6	2541	5082	12580

Running Jobs

Role	T1_DE_KIT	T1_ES_PIC	T1_TW_ASGC	T1_UK_RAL	T1_US_FNAL	Total
t1production	1221	473	0	1563	3172	6429
production	3	564	3	23	0	593
other	0	0	0	0	0	0
Total	1224	1037	3	1586	3172	7022



Global Overview

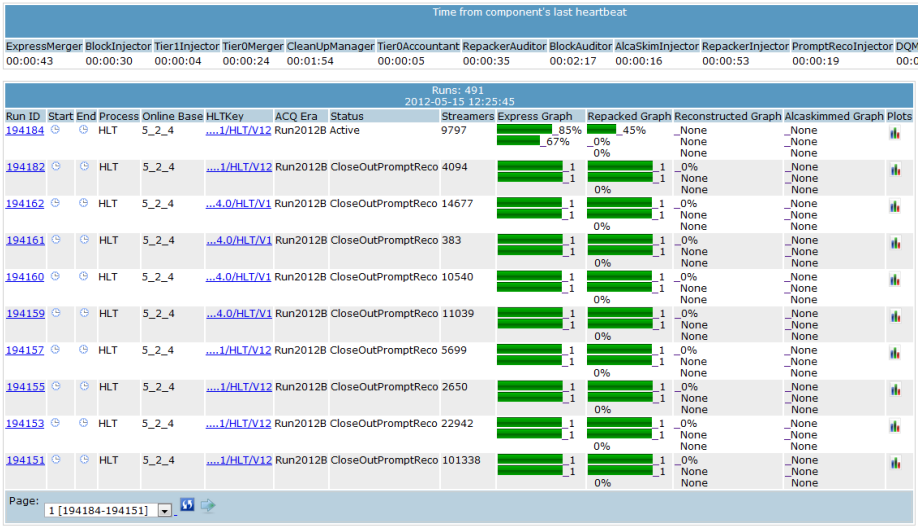
- Includes all monitoring that shows “what is happening now”
- Main examples are

Source	Description
TOMon	Tier-0 activities
Global Monitor	WMAgent workflows
PhEDEx monitoring	Data transfers
FTS monitoring	Data transfers
Storage accounting	Disk usage
CMS Dashboard applications	Cumulative job information

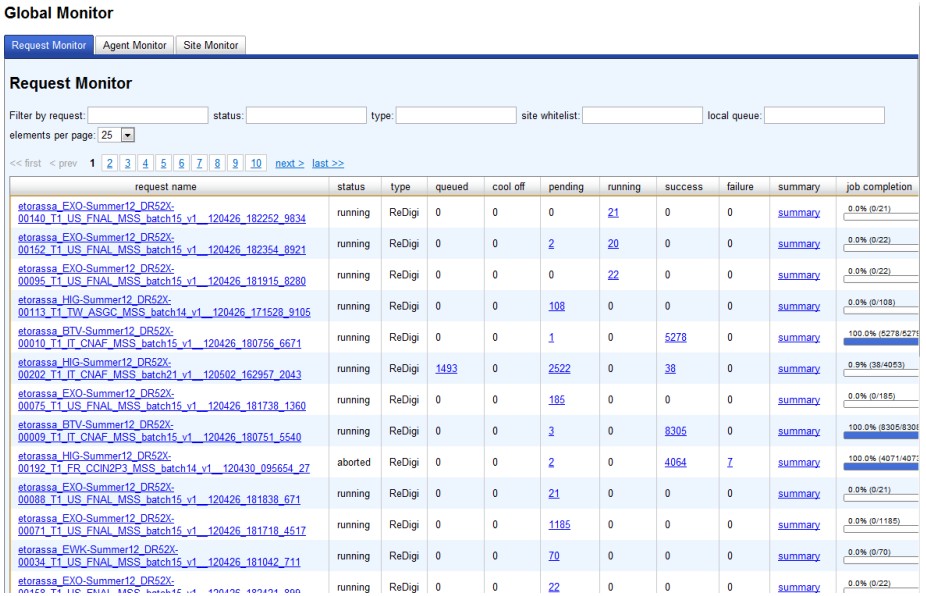
Not yet operational

T0Mon and Global Monitor

- T0Mon is a fully contained monitoring system for the Tier-0
- Being obsoleted with the transition to WMAgent

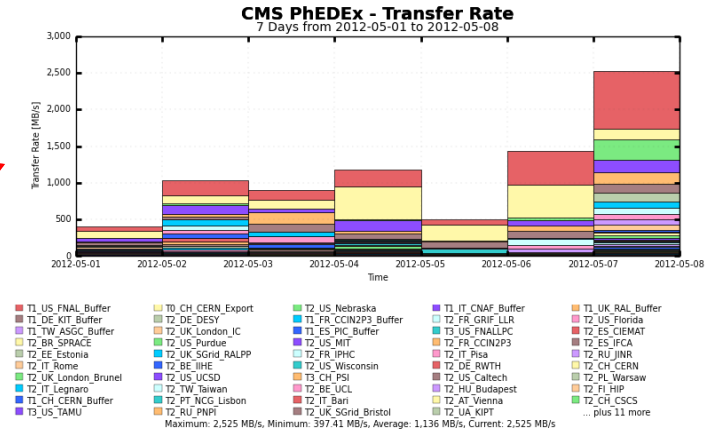


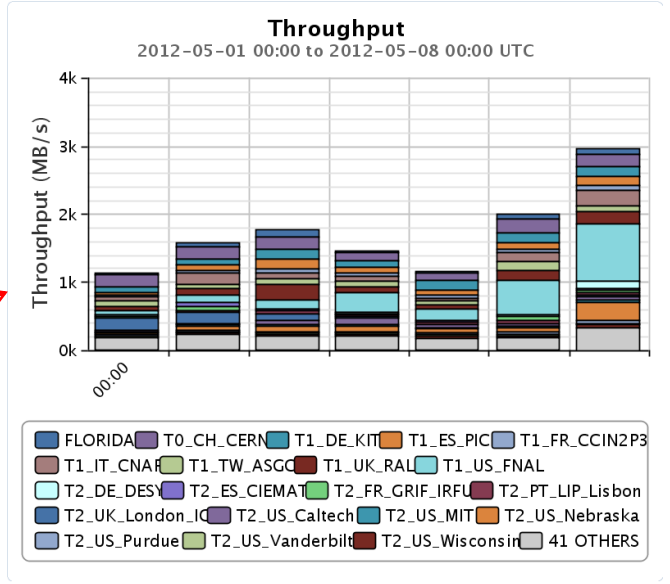
- The Global monitor tracks the status of all production workflows
- Full access to workflow metadata and job outputs
- Aggregates information from several WMAgent instances



Data transfer monitoring

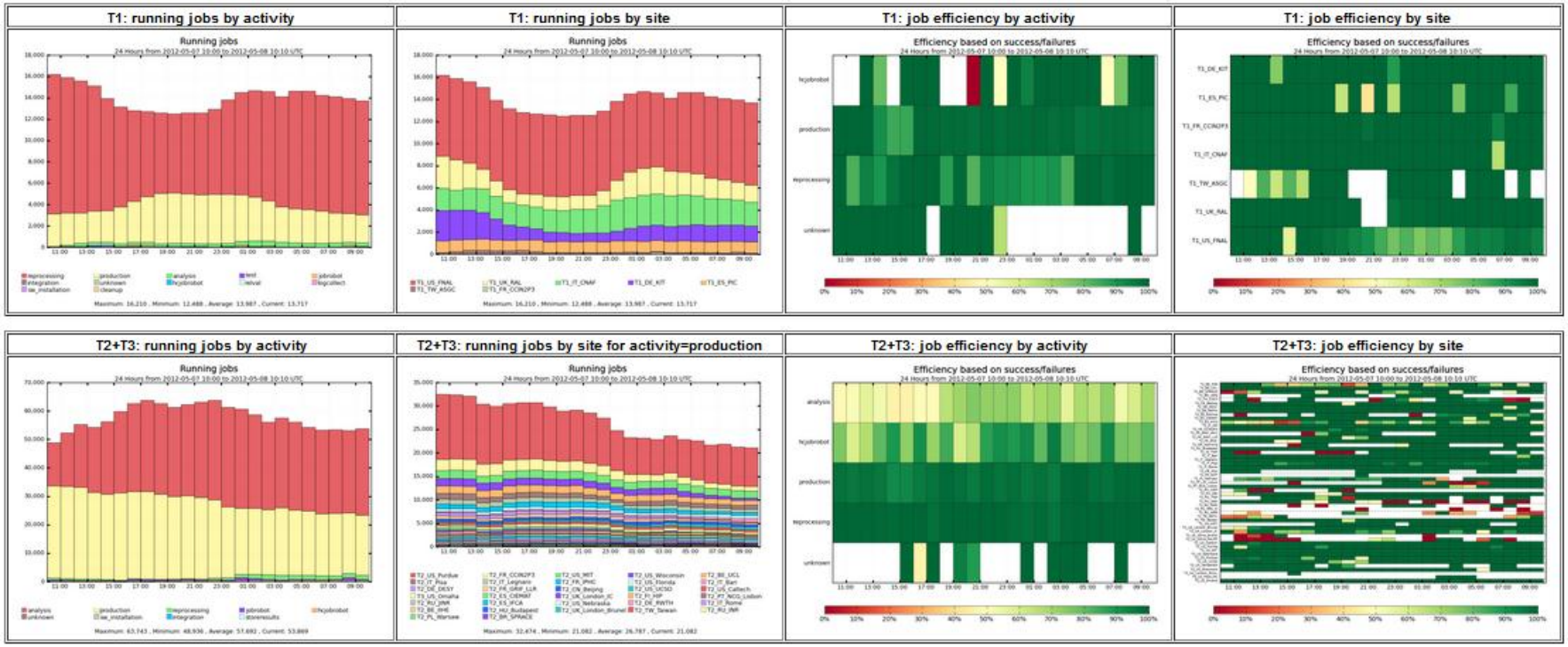
- Three data transfer monitors
  - PhEDEx monitoring: follows transfers at the dataset level
  - FTS Monitor (by CCIN2P3): aggregates information from all FTM, measures "everything"
    - Rates, duration, streams/file, SRM overheads
    - Useful for FTS configuration optimization, debugging, LHCONE and other studies
  - WLCG Global Transfer Monitor (developed by IT-ES): cross-VO view, uses MSG for data transport, powerful web UI





Dashboard

- Dashboard Historical View keeps track of all jobs
  - More global in scope than Global Overview (as it includes analysis or test jobs)
  - Many quantities monitored
  - Plots can be easily embedded in other pages or data exported



Historical Accounting

- Being able to reconstruct long term trends in resource usage, activity levels etc. is essential for accounting and planning

Source	Description
PhEDEx monitoring	Data distribution and movement
Storage accounting	Usage of disk at Tier-2's
CMS Dashboard	Activity/job accounting
Accounting portal	Central portal for accounting information

Not yet operational

Tier-2 Disk Accounting

- Goal: know how CMS uses the storage in a site
  - Any data, not just datasets registered in central catalogues
- The system will take care of collection, aggregation, DB store, visualization of information
- Source is standard WLCG-compliant storage dumps
  - Already used for storage-file catalog consistency checks
  - Compatible with all WLCG storage technologies
- Will use treemapping for interactive visualization
- Could be reused by other VOs

Accounting portal

- A project to provide a central page where to find current and historical values of several quantities
  - Many already available but scattered in several pages
  - Will be very useful at many levels (computing operators, coordinators, management)
  - Development just started

Metrics
LHC duty cycle, rate/primary dataset, pile-up
Processing times/event
Event sizes per data type
No. events/primary dataset per data type
Job processing latencies
Memory usage for jobs
CPU efficiency
Job success rate
Used vs. pledged resources (tape, disk, CPU)
Transfer efficiencies and rates

The Monitoring Task Force

- A monitoring review in November 2010 identified issues and proposed recommendations
  - Improve coordination of monitoring efforts and interaction between operations and developers
  - Aim at developing a coherent monitoring overview of CMS systems and operations
  - Appoint monitoring coordinators and define a work plan
  - Ensure that all relevant information is sent to the Dashboard and that this performs as required
- A Monitoring Task Force was started in March 2011 with an expected duration of 9 months

Monitoring work plan

Item	
Draw the overall picture of CMS monitoring	General
Identify information needed for application performance and validation studies	Improvements
Consolidate and clean up Dashboard information	
Improve Dashboard performance	
Choose an aggregation technology to create customized views of monitoring information	Development
Define requirements for a solution to aggregate and generate alarms	
Build an accounting portal	
Implement a disk space accounting system	
Put Data Popularity service in production	

Development projects

- CMS Overview chosen as presentation layer for monitoring
  - Derived from the highly successful Data Quality Monitor and completely owned by CMS
  - Two ongoing developments: computing shift page and Accounting Portal
  - *Critically* important to create the "coherent monitoring overview" that CMS needs!
- Alert Framework
  - Propagates and collects alert-worthy conditions from WMagent and visualizes them
  - Scope will be extended to other computing software (PhEDEx, RequestManager)

Monitoring evolution

- Driving forces in the monitoring evolution
  - Use systems used and supported also outside CMS when possible
    - ☞ SAM/Nagios, SUM/myWLCG, LeMON, SLS, HammerCloud, Data Popularity, Dashboard
  - Single (or few) entry point(s) to all monitoring information and less “home-made” scattered pages
- Relationship with WLCG Operations and tools technical evolution group (TEGs)
  - Mostly about infrastructure and network testing

Final thoughts

- The CMS Monitoring Task Force succeeded in improving the awareness on several monitoring issues
- Facilitated discussions, brainstorming and taking on responsibility
- Set a direction for current and future developments
- But much still needs to be done to achieve all the goals in time for the end of the long shutdown
- Looking for even more synergies with other experiments should be a priority

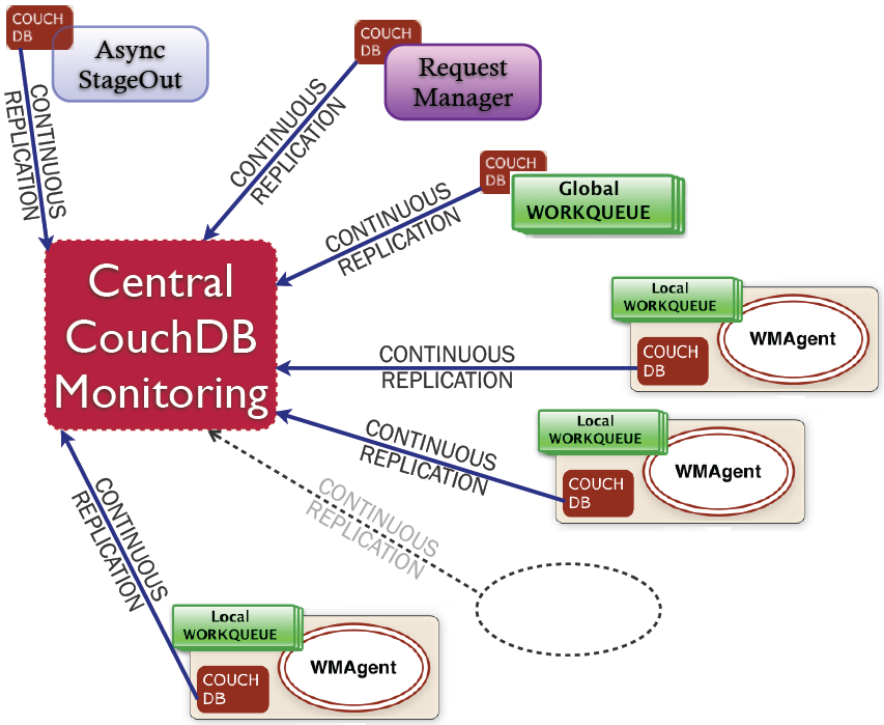
References

- Wakefield S *The CMS workload management system*, CHEP2012
- Saiz P et al, *Experiment Dashboard - a generic, scalable solution for monitoring of the LHC computing activities, distributed sites and services*, CHEP2012
- Molina-Perez J A et al *Monitoring techniques and alarm procedures for CMS services and sites in WLCG*, CHEP2012
- Van Der Ster D C et al *Experience in Grid site testing for ATLAS, CMS and LHCb with HammerCloud*, CHEP2012
- Saiz P et al *Collaborative development. Case study of the development of flexible monitoring applications*, CHEP2012
- Sciabà A et al *New solutions for large scale functional tests in the WLCG infrastructure with SAM/Nagios: the experiments experience*, CHEP2012
- Lapka W et al *Distributed monitoring infrastructure for Worldwide LHC Computing Grid*, CHEP2012
- Andreeva J et al *Providing WLCG Global Transfer monitoring*, CHEP2012
- Ratnikova N et al *Data storage accounting and verification in LHC experiments*, CHEP2012
- Maxa Z Alert *Messaging in the CMS Distributed Workload System*, CHEP2012
- Giordano D et al *Implementing data placement strategies for the CMS experiment based on a popularity model*, CHEP2012
- Tadel M et al *Xrootd Monitoring for the CMS experiment*, CHEP2012

BACKUP SLIDES

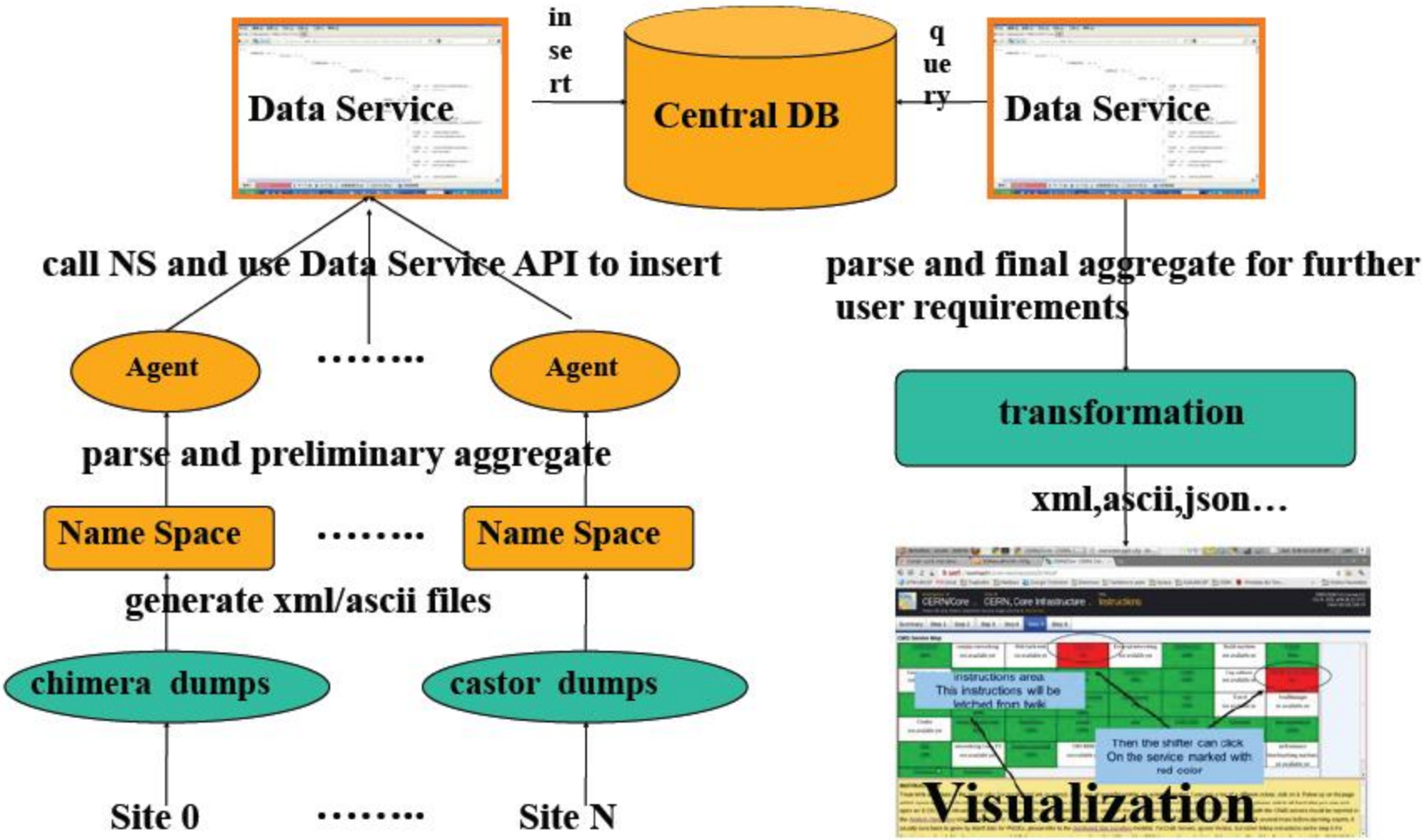
WMagent monitoring evolution

- Currently, using a drill-down model for monitoring and tracking
  - All distributed services need to be contacted to satisfy a request
  - Serious scalability, security and reliability issues
- Moving to a push model where information is aggregated at source level and pushed to a central service
  - Decouples WMAgent from monitoring load
  - Scales much better as less information is pushed centrally



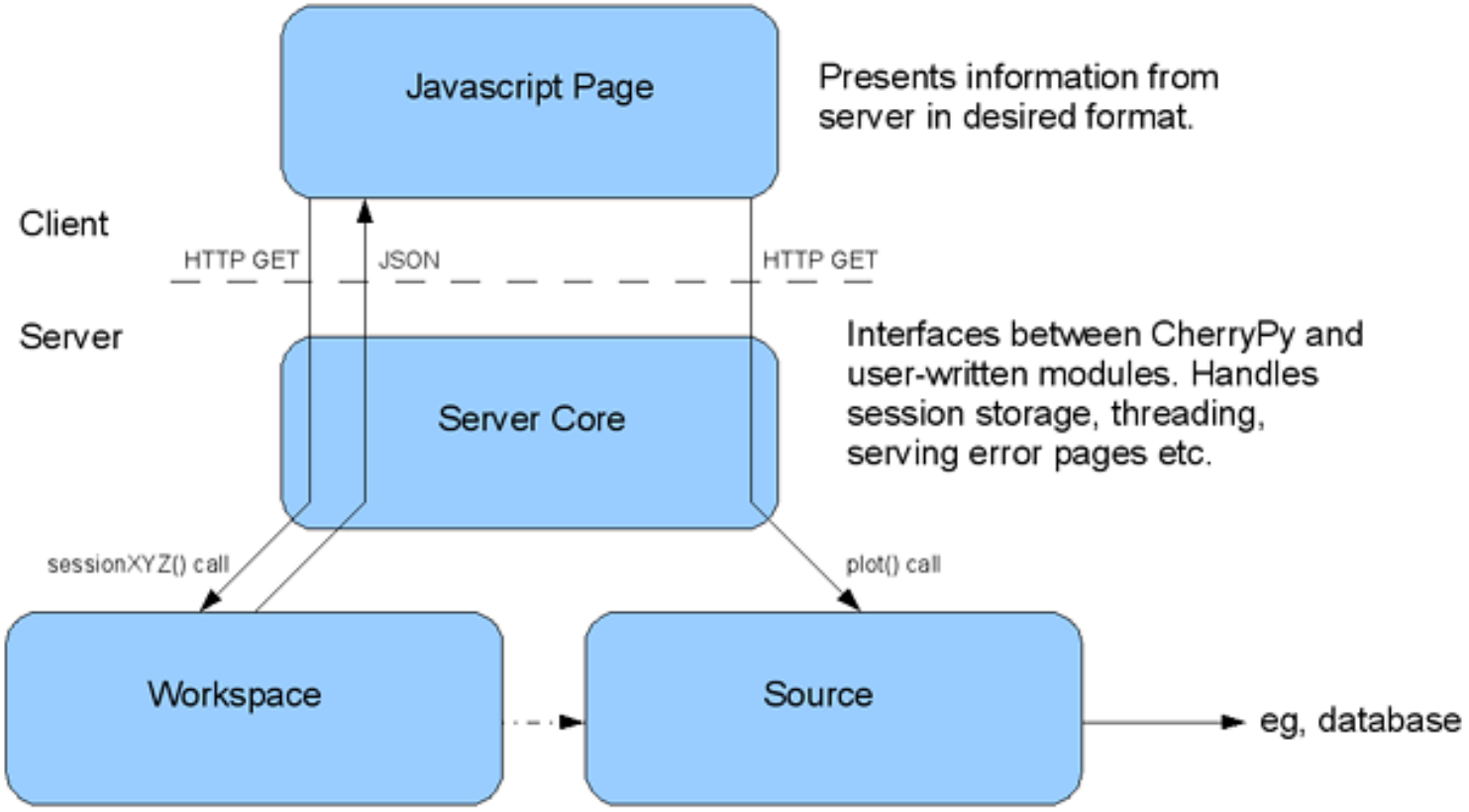
Wakefield S. *The CMS workload management system*, CHEP2012

Storage accounting architecture



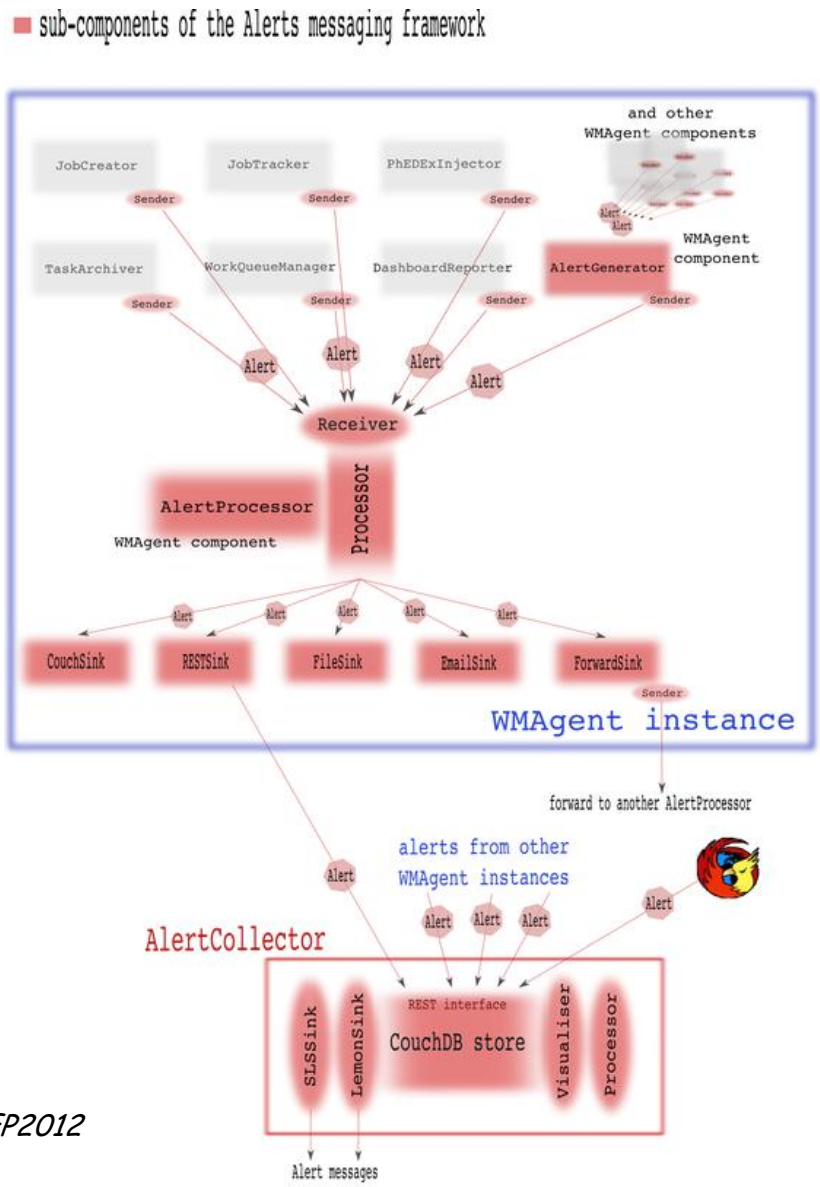
Ratnikova N. *et al, Data storage accounting and verification in LHC experiments*, CHEP2012

Overview design



Alarm framework

- AlertGenerator
  - Inside WMAgent, generates the alerts based on active monitors and defined metrics
- AlertProcessor
  - Inside WMAgent, gathers the alerts, buffers them, sends them to various sinks
- AlertCollector
  - Central store receiving alerts from many WMAgent instances (and in future other DMWM applications such PhEDEx)



Maxa Z., *Alert Messaging in the CMS Distributed Workload System, CHEP2012*