

PSS

Physics Services Support

CERN IT
Department

Grids, middleware, tools and experiment activities at CHEP07

Andrea Sciabà



- Introduction
- Production Grids
- Workload Management Systems
- Data Management
- Monitoring
- Experiment Tools
- Experiment Challenges and Activities
- Conclusions

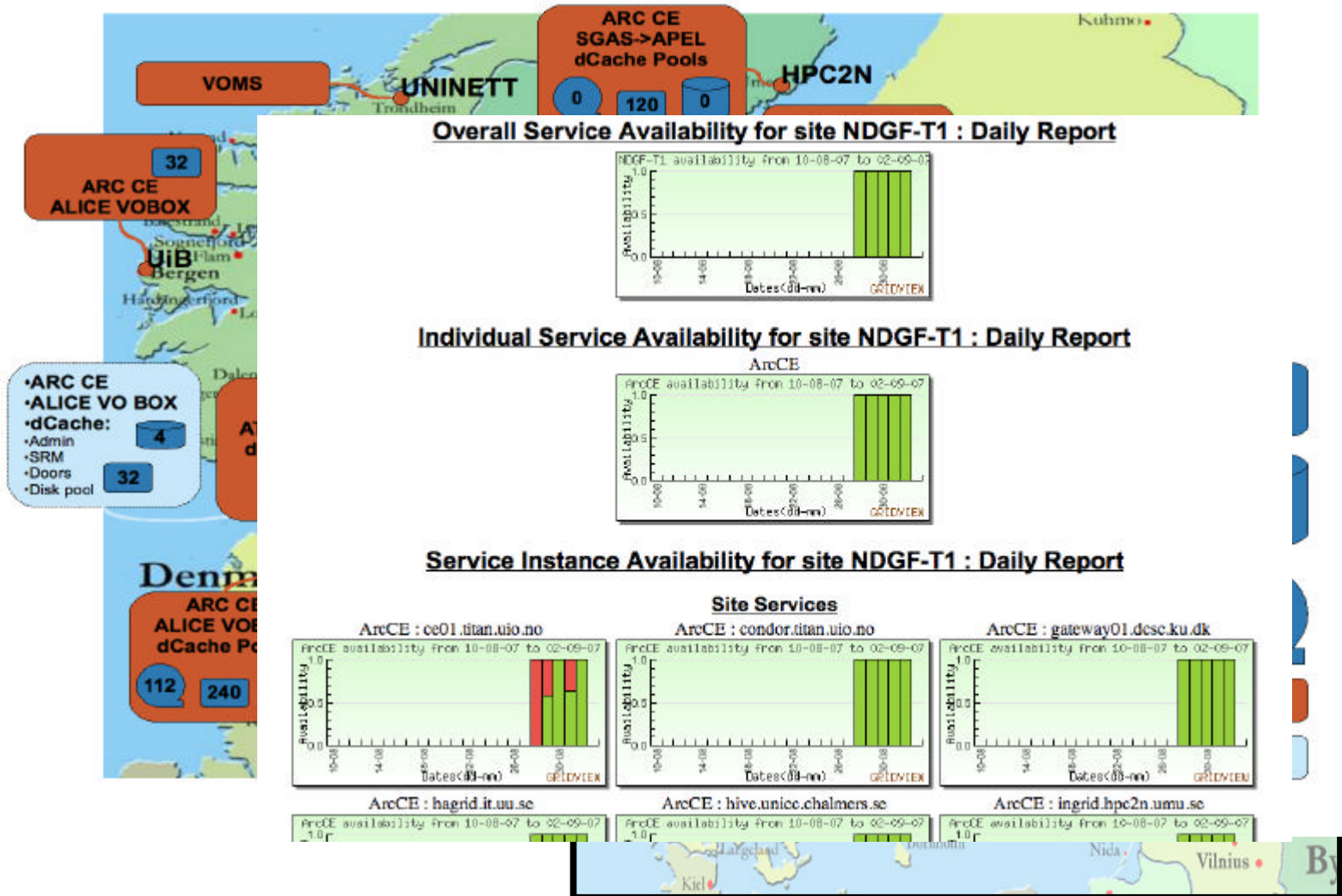
- Several sessions at CHEP07 were centered on Grid computing and its applications
 - Computer Facilities, Production Grids and Networking
 - Grid Middleware and Tools
 - Distributed Data Analysis
- This presentation tries to summarize a personal selection of the reports on some of the areas covered
 - The main focus will be though on the usage of the Grid by the LHC experiments
 - Many topics have been left out...

- OSG (R. Pordes)
- GridPP (J. Coles)
- Nordic Data Grid Facility (NDGF) (L. Fischer)

- Supports ATLAS and CMS
- But also CDF, D0, neutrino physics, STAR, LIGO, etc.
- Operates the infrastructure and supports site administrators and users
- 75 processing resources in US, 4 in Brazil, 1 in Mexico, UK, Taiwan and South Korea
 - ~30,000 CPU cores
 - 75,000 jobs/day
 - 240,000 CPU wallclock hours/day
 - CPU time/wallclock time ~0.75
 - 20%-30% usage from non-owners!
- 15 storage resources
 - ~4 PB of disk
 - 100 TB of data moved per day
 - 1 PB/day next year
- ~2x increase in accessible resources by end of 2009

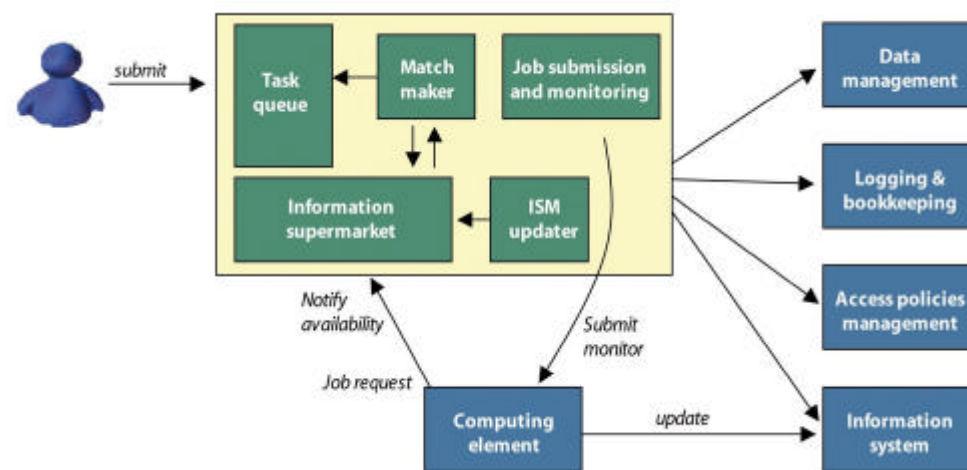
- The middleware stack is the Virtual Data Toolkit (VDT)
- Interfaces
 - Job execution
 - Condor-G
 - Submitting to WS, pre-WS GRAM, Condor glide-in
 - Storage
 - SRM
 - Full support to SRM 2.2 from next February
 - Data transfer
 - GridFTP
 - VO management
 - VOMRS, VOMS shared with EGEE
- Achieving interoperability with WLCG
 - Information system publishes to the WLCG BDII
 - OSG CEs can accept jobs from gLite WMS

- The Tier 1 is distributed between the 7 element

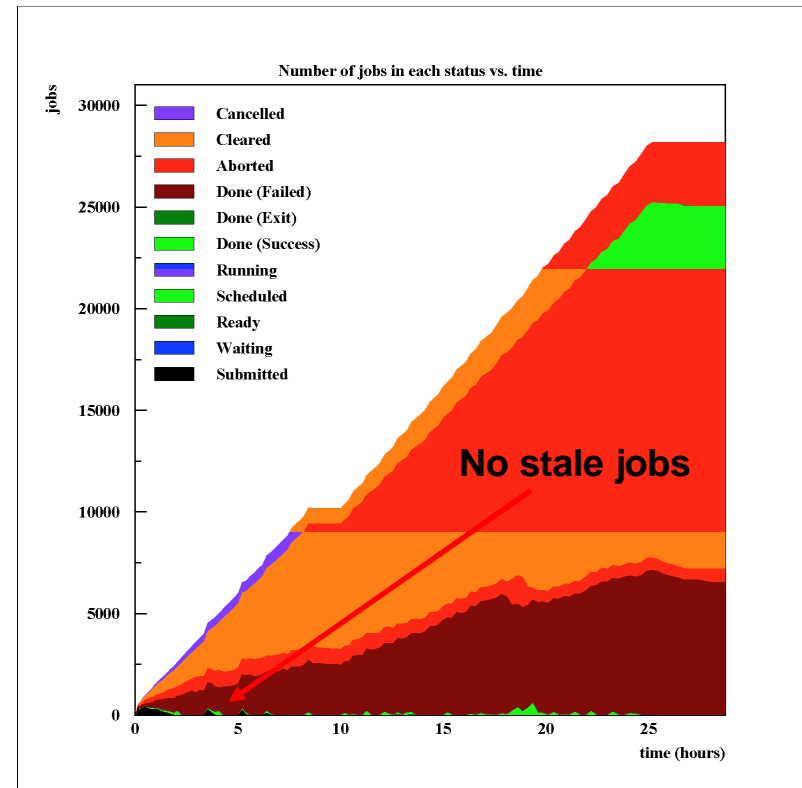


- A plethora of WM systems exist and are used
 - "Push" models
 - **gLite WMS** (M. Cecchi, S. Campana)
 - Condor-G
 - "Pull", or "Pilot" models
 - A job sits on a job slot and pulls "real" jobs from a central task queue
 - "standard" pilots
 - DIRAC (S. Paterson)
 - **PaNDA** (Y. Smirnov, P. Nilsson, T. Maeno)
 - AliEN (P. Saiz)
 - Condor Glide-in
 - **glideinWMS** (I. Sfiligoi)
 - **CRONUS** (S. Padhi)
 - Other Grids
 - Xgrid from Apple (A. Kocoloski)
 - Sun Grid (M. Potekhin)
- Several talks spread over different sessions

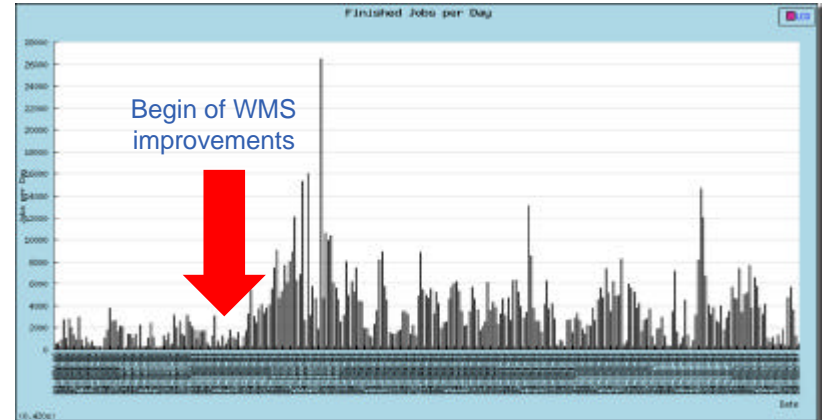
- Main functions
 - Match user jobs to available resources
 - Submit the job to the resource through the appropriate interface
 - Condor-G for the LCG CE
 - ICE for CREAM
 - "Manage" the job on behalf of the user
 - Log events to the Logging & Bookkeeping service
 - Automatically resubmit failed jobs
- Several functionalities make it better than the old LCG Resource Broker
 - **Bulk submission and bulk matchmaking**
 - Shared input sandbox among similar jobs
 - Local cache of the Information System ("Information Supermarket")
 - Task queue to hold submission requests for which resources are not available
 - Load limiter
 - A new Web Service interface ("WMPProxy")



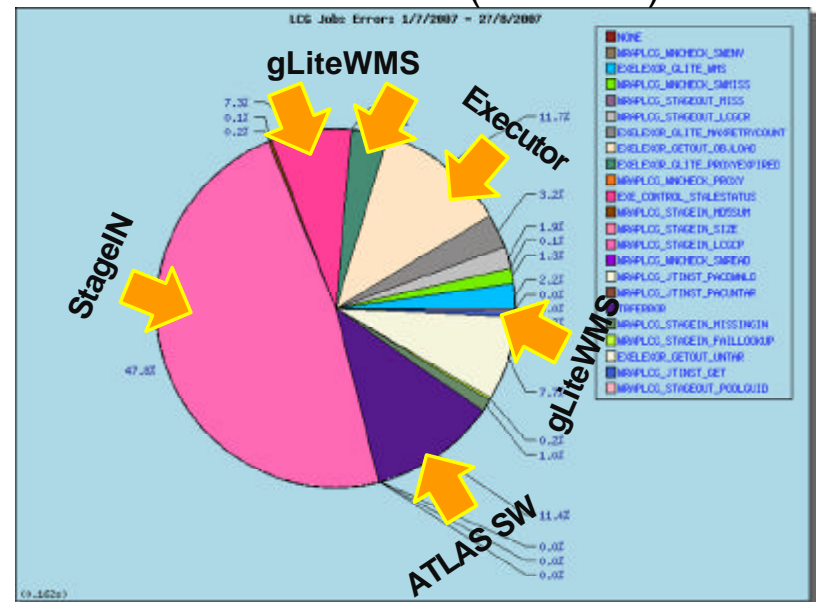
- After a very intense testing and bug fixing, the WMS demonstrated performance levels compatible with the experiment needs
 - ~16K jobs/day over one week
 - Stable behaviour and performance
 - Negligible fraction (0.3%) of jobs with problems due to the WMS
- Used in some way or another by all the LHC experiments
 - ATLAS, CMS to submit real jobs
 - ALICE, LHCb to submit pilot jobs



- The ATLAS production system submits part of the work to EGEE using the gLite WMS
 - Extremely painful at the beginning
- Good performances after the recent WMS improvements
 - Up to 20K jobs/day
 - Job efficiency ~60%
 - 13% due to the “WMS” (including problems with the CEs)
 - **Negligible** wallclock time inefficiency
 - Most jobs fail at the beginning and are resubmitted

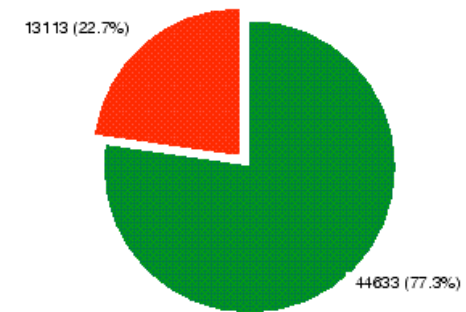
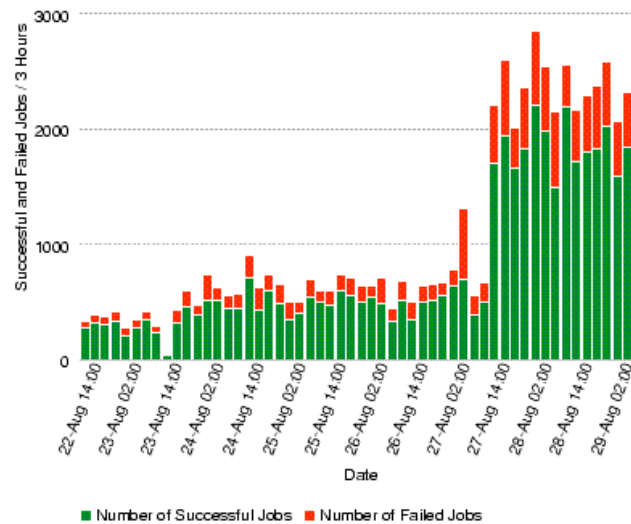
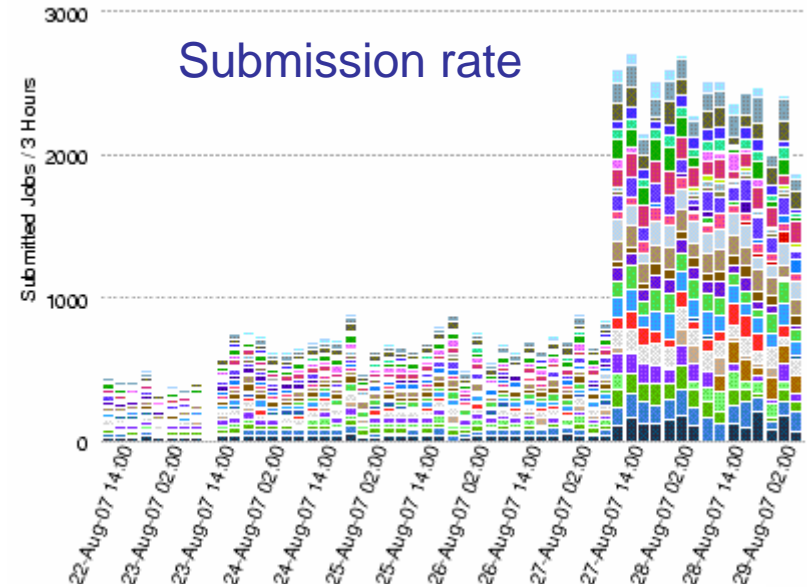


Job errors (7/8 2007)



gLite WMS: ~13% Data Management: 47% ATLAS SW: 11%

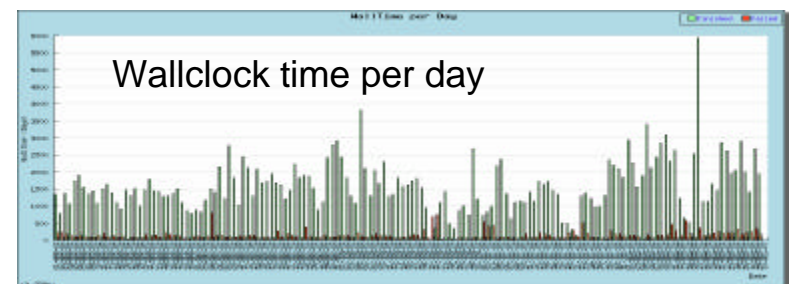
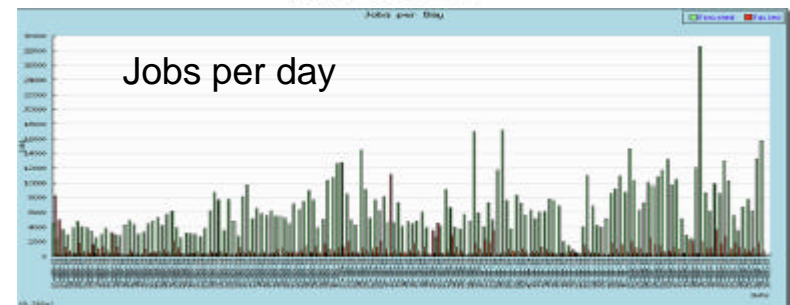
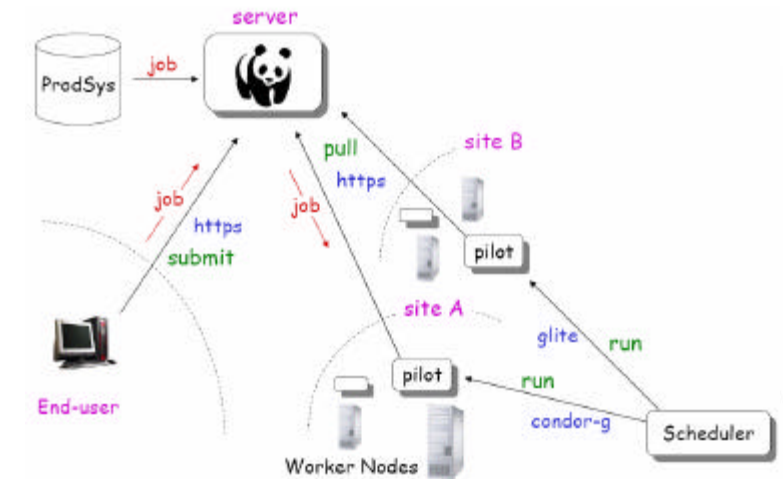
- CMS supports submission of analysis jobs via WMS
 - For CSA07 the goal is to submit at least 50K jobs/day via WMS
 - The Job Robot (a load generator simulating analysis jobs) is successfully submitting up to 20K jobs/day



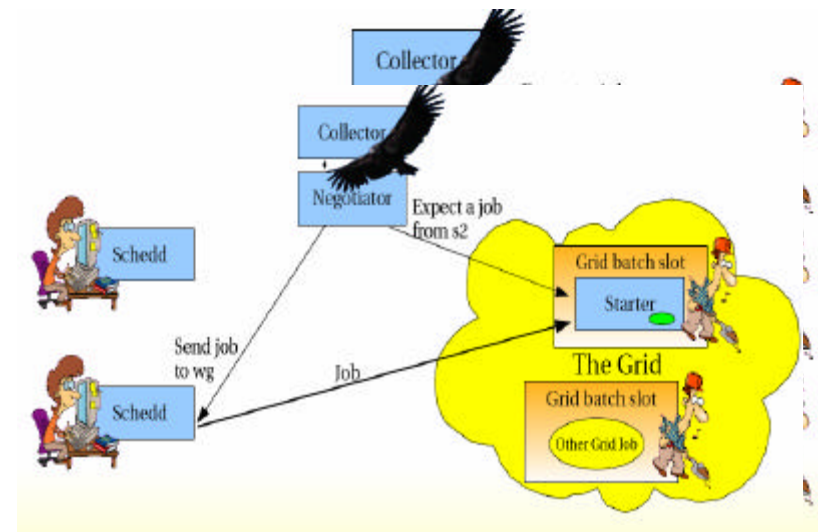
Success rate



- Used by ATLAS in OSG for production and analysis
 - Submits to Condor-G, PBS and Condor
- Central server Apache-based
 - Priority control
 - Resource allocation
 - Job scheduling
- For production jobs it takes care to send data transfer requests to DQ2 when needed
 - Data transfers managed asynchronously
- Performances
 - ~30M simulated events, >30% of total
 - Currently ~15K jobs/day, scales up to at least 40K jobs/day
 - Panda errors < 3% of total
 - 85% job efficiency
 - 91% WCT efficiency
- Plans
 - Extend support to EGEE
 - Add glideins (see next slide)

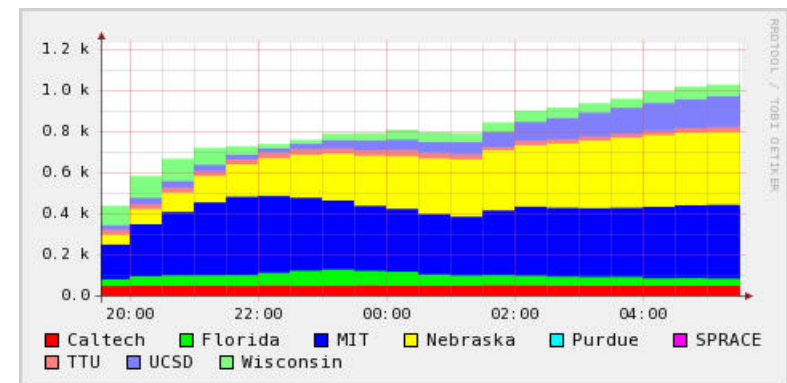
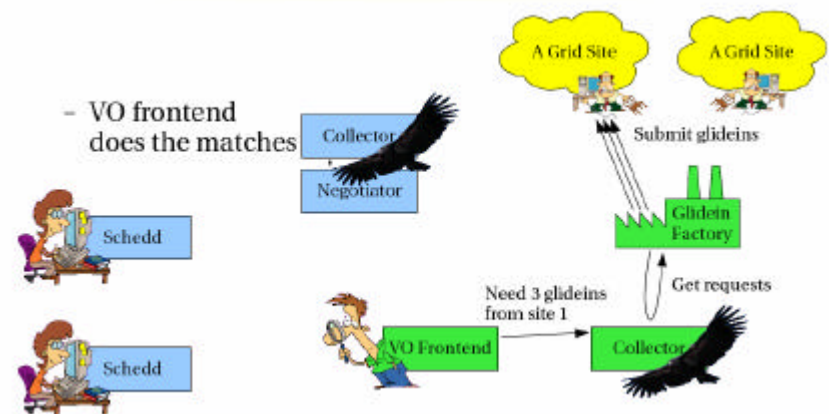


- Based on the Condor architecture
- A glidein is a starter submitted as a Grid job that runs on a worker node
 - A virtual private Condor pool running on WNs!
- Glideins can also take care of
 - Sanity checks
 - Discovery and publish batch slot characteristics
 - Preparation of the job environment



- Advantages
 - Easy to use for Condor users
 - Hides the Grid intricacies
 - Protects from obvious errors
- Disadvantages
 - Starters need an open connection to a central Condor proxy server (the GCB)
 - The user and the starter run under the same local account
 - Resource hungry (1.5 MB/running process on submit node)
 - Some components still immature (the GCB)
 - Debugging can be challenging!

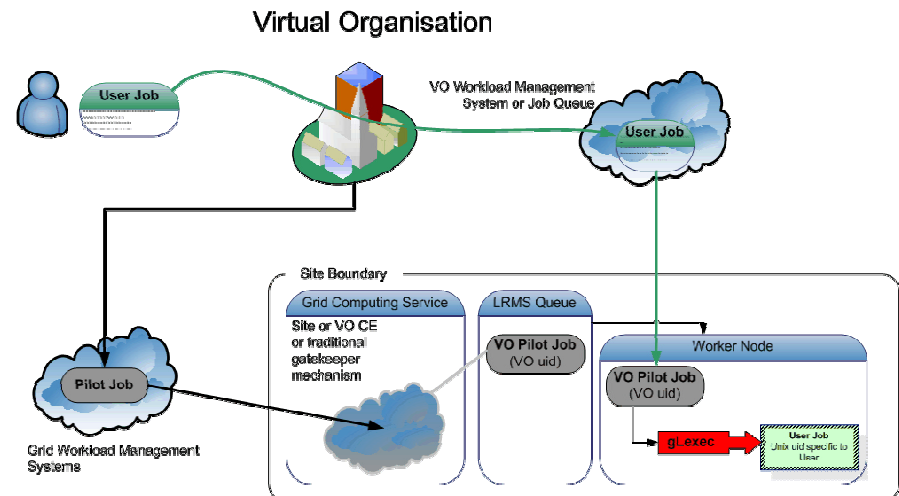
- A WMS built on top of glideins by CMS
 - A VO frontend matches jobs in the queue to sites
 - A glidein factory submits the needed glideins on the matched site
 - The rest is the same and the Condor negotiator has the final word
- Advantages
 - Keeps a steady number of idle glideins in site as long as there are waiting jobs
- Do glideins scale?
 - Not very well
 - Up to 1K jobs in parallel for CMS
- And glideinWMS?
 - No problem there



- Pilots submitted to OSG (via Condor-G) and to EGEE (via gLite WMS)
 - NDGF recently added
- The resulting virtual cluster is uniform across Grids
- Performances
 - ~645K jobs in 2007
 - WCT efficiency ~87%
 - Most errors related to data I/O and ATLAS software
 - Scaled up to ~5200 concurrent jobs
- Major problems
 - Loss of connection between starter and submit host

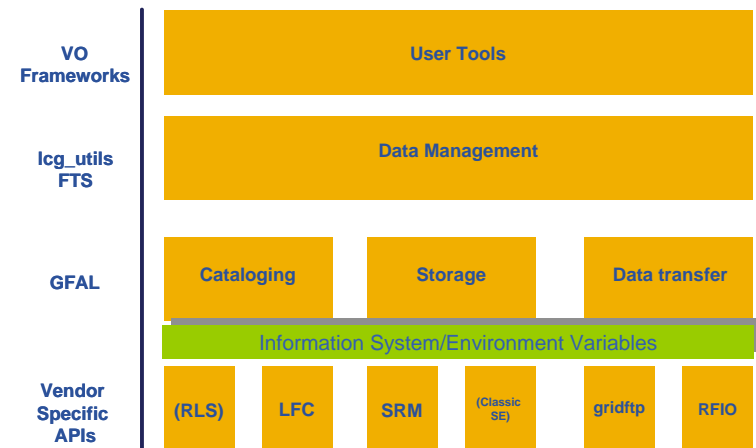


- Problem: pilot jobs run their payload (= user jobs) with the pilot's Grid credentials
 - Impossible to trace things back to the real job owner
 - Impossible to ban single users from sites
 - Users can steal or corrupt each other's data
- Solution: gLExec allows to switch UID for the running job given the real user's credentials
 - It will also keep a log of all the authorization requests
- Advantages
 - Solves the problem!
- Disadvantages
 - The VO must do the accounting
 - Policies are on local files
- Deployed and tested at FNAL with good results so far

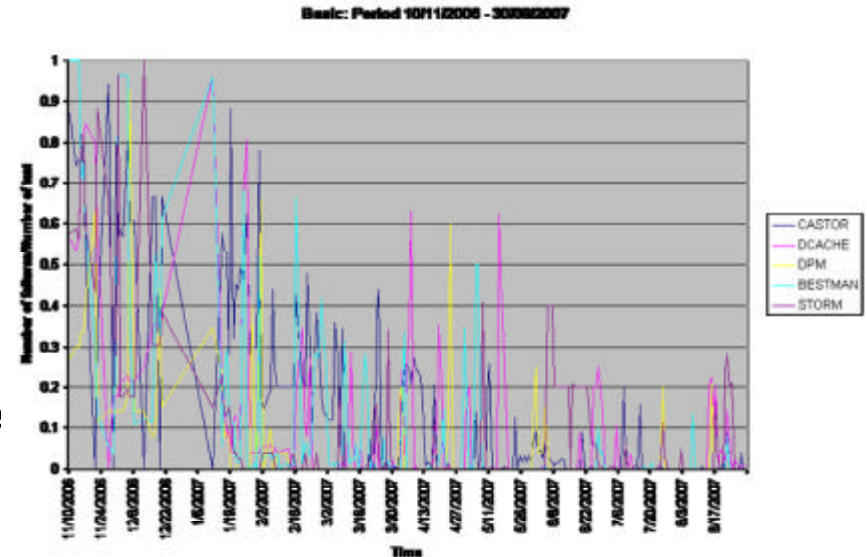


- DIRAC
 - Used by LHCb in EGEE for production and analysis
 - Pilots are submitted via LCG RB or gLite WMS
 - Up to 10K concurrent jobs
 - Uses standard batch system components (e.g. Maui) for prioritization of jobs in the Task Queue
- AliEn
 - full middleware stack developed in ALICE

- Interfaces
 - SRM 2.2 (F. Donno)
- Storage solutions
 - dCache (T. Perelmutov, P. Fuhrmann, G. Berhmann)
 - DPM (M. Schulz)
 - CASTOR2 (G. Lo Presti, J. Van Eldik)
- File catalogues
 - LFC (M. Schulz)
- Data transfer tools
 - FTS (M. Schulz)
- User clients
 - GFAL, lcg_util (M. Schulz)



- SRM is a standard interface for space allocation and file management on the Grid
 - CASTOR2, DPM, dCache, StoRM, BeStMan
- Tests done with the S2 suite developed at RAL, to ensure
 - Validation of use cases
 - Consistency
 - Compliance with the specification
 - Performance and reliability
 - 52 tests run 4 times/day on 21 endpoints with different configurations



- Other groups testing also
 - Higher level tools
 - FTS, lcg_util, etc.
 - Transparent access (v1 vs. v2)
 - handling of VOMS groups
 - VO data access patterns
 - ...

- dCache
 - Used at 7 Tier-1's and several Tier-2's
 - Supports GridFTP, (gsi)dCap, xRoot, SRM, nfs2/3
 - Working on Chimera, a new namespace system to replace PNFS
 - >2 GB files, ACLs, user defined queries, fast
 - Working also on NFS 4.1
 - Faster, GSS authentication, ACLs
 - Clients come with all major OS vendors
- DPM
 - Disk-only SE for Tier-2's
 - 96 in production
 - Supports GridFTP, secure RFIO, xRoot (limited), http(s)
 - Full support of VOMS for authorization (file permissions, ACLs)
 - Working on file encryption for medical applications

- LCG File Catalog
 - Stores user file names, location, owner and permissions
 - Hierarchical namespace
 - Supports GSI security and VOMS based ACL
 - LFC is used as internal catalogue by DPM
 - The effort is mainly on optimization
 - Bulk methods can reduce round trips \Rightarrow up to 20x faster
- lcg_util
 - C API + CLI to cover most common use cases
 - Replication, catalogue interaction, etc.
 - Work ongoing to make them work independently from LFC
 - The goal is to have a unique and official "SRM client" for users
- GFAL
 - For POSIX-like file access
 - Supports now SRM 2.2

- The tool to transfer data between endpoints
 - Handles retries in case of failures
 - Failures of the SRM
 - Channels are "quoted"
 - Inter-VO and intra-VO
 - Different roles
 - User, VO production
- Deployed at Tier-0 sites
 - Over 9 petabytes of data
- FTS 2.2 improvements
 - Better security model
 - Improved monitoring
 - Beta SRM 2.2 support
 - Database optimization

FTS Report

Disclaimer: This page contains a report generated from information stored in the FTS Database and is intended for reporting purposes only. Since the format will probably change in the future, it's therefore recommended not to use parsing robots on it.

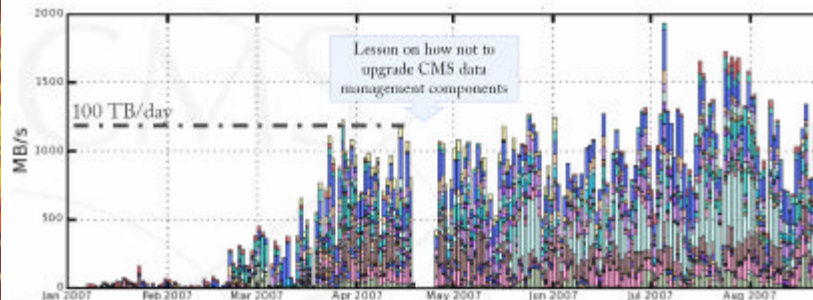
Statistics concerning all the transfers performed yesterday
Between 2006-10-12 08:00:00 +02:00 and 2006-10-13 08:00:00 +02:00

CERN-* Filter Show VO details

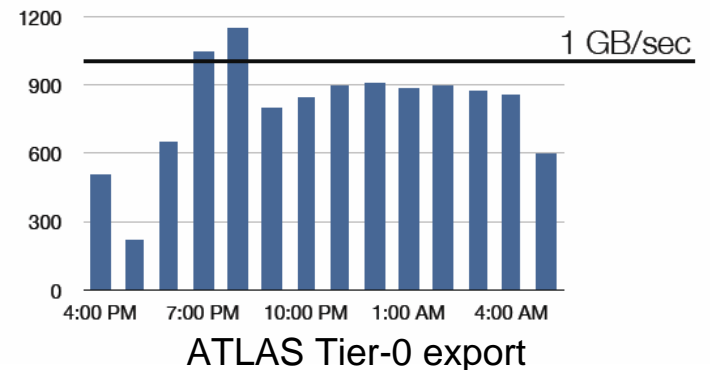
Channel Name	VO Name	Total	% Failures	# Succ.	# Fail.	1st Failure Reason	% 1st Failure Reason	2nd Failure Reason	% 2nd Failure Reason	Avg. Size (GB)	Avg. Duration (sec)	Avg. Tx Rate (MB/sec)	Eff. Tx Bytes (GB)	Tx Bytes (GB)
CERN-PIG	[All]	11282	73.97	3192	9070	Dest SRM	58.22	Other	37.53	0.93	263.03	1.82	1709.41	1700.41
	atlas	8932	99.92	7	8925	Dest SRM	57.13	Other	38.08	0	220	0	0	0
	cms	208	0	208	0					2.7	767.55	3.64	561.28	561.28
	dteam	974	0.51	969	5	Other	80	Source SRM	20	0.95	356.31	2.88	923.83	923.83
	lhcb	2145	6.53	2005	140	Source SRM	99.29	Other	0.71	0.11	165.85	0.81	215.32	215.32
	ops	3	0	3	0					0	202.67	0	0	0
CERN-RAL	[All]	8699	59.26	3644	5155	Other	84.91	Source SRM	14.88	0.85	478.22	2.58	3026.81	3027.57
	alice	1155	82.6	201	954	Other	99.58	Dest SRM	0.31	1.86	1895.05	1.11	372.95	372.95
	atlas	4512	88.52	518	3994	Other	84.85	Source SRM	15.15	1.79	1428.94	1.49	926.26	926.57
	cms	227	3.09	220	7	Dest SRM	85.71	Source SRM	14.29	2.53	348.65	10.09	555.61	555.61
	dteam	1077	3.09	1034	43	Other	86.05	Source SRM	9.3	0.95	276.64	4.01	980.47	980.91
	lhcb	1725	9.1	1568	157	Source SRM	99.36	Other	0.64	0.12	146.03	1.16	191.52	191.52
	ops	3	0	3	0					0	27	0.01	0	0
CERN-BARA	[All]	8792	42.55	5051	3741	Dest SRM	93.77	Source SRM	12.22	1.94	109.03	15.4	6777.95	6784.92
	alice	3134	15.12	2680	474	Source SRM	57.17	Dest SRM	41.14	1.86	189.53	18.43	4426.44	4430.29
	atlas	2018	53.32	942	1076	Dest SRM	72.4	Source SRM	16.54	1.15	144.44	9.42	1085.07	1087.6
	dteam	3488	61.32	1349	2139	Dest SRM	98.74	Other	0.98	0.93	91.91	14.66	1260.71	1261.32
	lhcb	148	35.14	96	52	Dest SRM	92.31	Other	3.85	0.06	76.1	0.93	5.7	5.7
	ops	4	0	4	0					0	97.25	0.02	0	0
CERN-INFN	[All]	11492	42.31	6630	4862	Dest SRM	43.95	Other	37.7	1.13	395.77	3.21	7514.29	7614.84
CERN-CERN	[All]	1538	39.71	926	610	Source SRM	58.38	Dest SRM	15.9	0.07	287.71	0.38	67.89	68.08
CERN-ASCC	[All]	6051	23.54	5236	1813	Source SRM	50.84	Other	28.89	1.14	1098.6	1.08	5955.81	6080.58
CERN-GRICKA	[All]	12755	21.38	10028	2727	Source SRM	64.38	Other	32.53	0.87	371.87	3.19	8782.02	8787.53
CERN-TRIUMF	[All]	2244	20.83	1781	463	Other	61.77	Source SRM	31.1	1.04	395.15	3.63	1847.25	1917.13
CERN-BNL	[All]	13975	19.42	11261	2714	Source SRM	69.97	Other	24.24	0.44	190.38	3.41	4951.58	4960.34
CERN-IN2P3	[All]	11697	13.76	10087	1610	Source SRM	48.57	Other	47.46	1.22	295.21	5.33	12329.63	12329.63
CERN-FNAL	[All]	917	4.58	875	42	Transfer	97.62	Other	2.38	0	378.88	0	0	0

Click on the Channel Name to show the VO details

- PhEDEx (L. Tuura)
 - The data movement system of CMS, very robust and performant by design
 - Manages not single files but blocks (5-10 TB)
 - Protects the system from failures
 - All communications among agents through a central DB, very strongly optimized
 - Thorough validation procedure
 - Routinely transfers ~100 TB/day, 50K files/day
 - Focus is now on just improving the interface

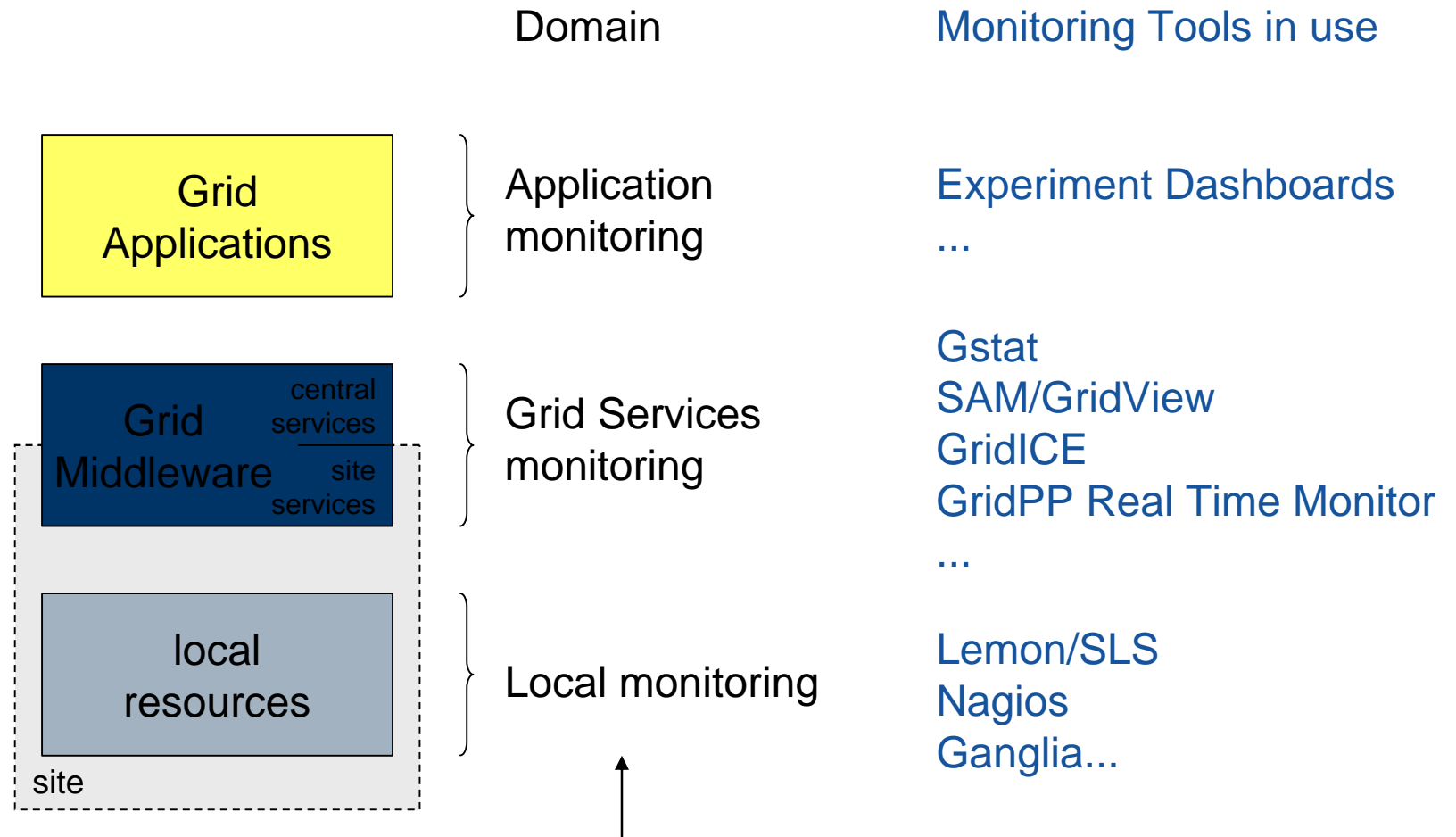


PhEDEx aggregate data transfers



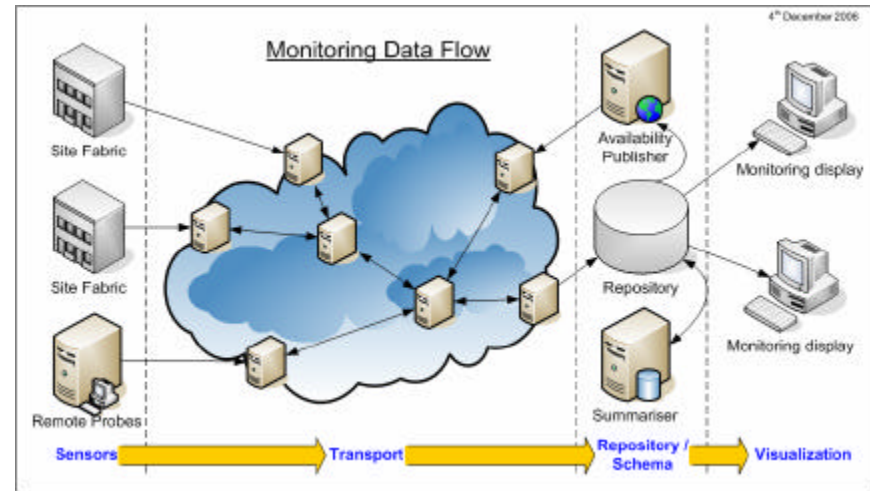
ATLAS Tier-0 export

- ATLAS DDM (M. Lassnig)
 - Takes care of data movement and data bookkeeping
 - Based on a central catalogue and some site services
 - Some recent improvements
 - Improved transfer reliability (failover strategies, etc.)
 - Improvement of fairshare
 - New database schema
 - More efficient SQL queries
 - **Almost ready for data taking**

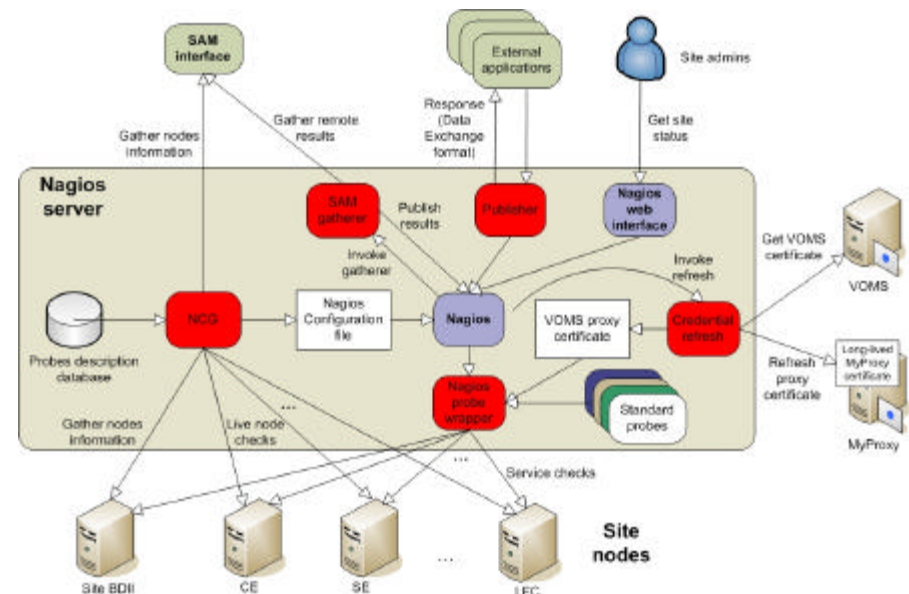


3 WLCG Monitoring Working Groups

- Provide monitoring views suited for
 - Site admins
 - Grid operators
 - VO management
 - Grid project management
- Allow the site to see information collected also by remote probes



- Standardize the monitoring of Grid services
 - Define "standard" probes
 - Define a standard Data Exchange format for publishing and consuming
- Prototype based on Nagios
 - Can be already be tried out!



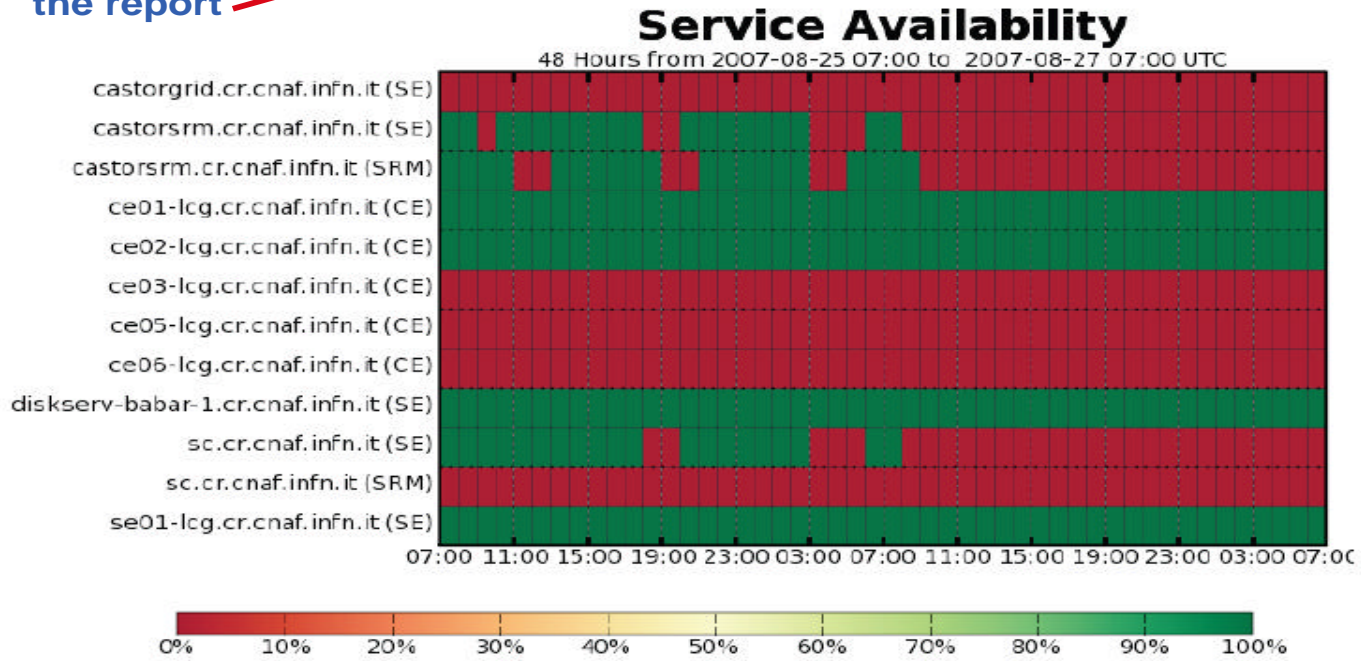
SAM VISUALIZATION

Summary View of latest Results

Historical Summary View

View	Time Range	Sites	Service Types
Service Availability	Last 48 Hours	All Sites	VO critical
		INFN-LNL-2 infn-napoli-cms INFN-PADOVA INFN-PISA INFN-ROMA1-CMS infn-roma2 INFN-T1	Select All CE SE SRM

Select site or set of sites and service types to be included in the report



- A framework developed to monitor remote Grid services for the LCG operations, and now used also by
 - Site Availability calculations
 - GridView
 - Middleware certification
 - Experiment specific tests

- CMS

- "plugs" custom tests for the CE and SRM to the worker node to the local storage, etc.
- Calculates a "CMS availability" based on the

- ALICE

- Uses SAM to publish the results of custom tests
- Results integrated in ALICE monitoring, alarm

- LHCb

- Critical tests for LHCb applications are submitted to SAM

- Test results are used to build a "production monitor"

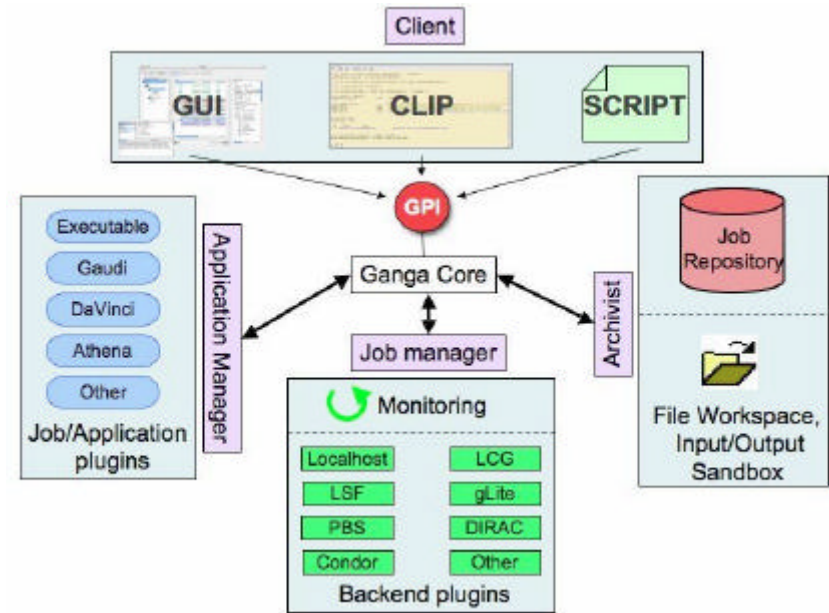
- ATLAS

- Now only "standard" tests run with an ATLAS proxy
- Developing custom tests for the SRM (approach similar to CMS)



LCG.Legnaro.it	LCG.Liverpool.uk	LCG.NIF.it	LCG.NPI-R2G.de	LCG.Manchester.uk
LCG.Milano.it	LCG.Montreal.ca	LCG.NCP.pk	LCG.NCU.tw	LCG.NIIF.hu
LCG.NIKHEF.nl	LCG.NIPNE.ro	LCG.NOVSU.ru	LCG.NSC.se	LCG.Napoli-Atlas.it
LCG.Napoli.it	LCG.OU.il	LCG.Oxford.uk	LCG.PARGRID.pk	LCG.PANUKKALE.tr
LCG.PDC.se	LCG.PIC-slc4.es	LCG.PIC.es	LCG.PNPI.ru	LCG.Padova.it
LCG.Pisa.it	LCG.QMUL.uk	LCG.RAL-HEP.uk	LCG.RAL-slc4.uk	LCG.RAL.uk
LCG.RHUL.uk	LCG.SARA.nl	LCG.SINP.ru	LCG.SPACI.it	LCG.SRCE-slc4.hr
LCG.Sheffield.uk	LCG.Sofia.bg	LCG.TAU-slc4.il	LCG.TAU.il	LCG.TCD.ie
LCG.TIFR.in	LCG.Torino.it	LCG.Toronto.ca	LCG.Trieste-slc4.it	LCG.Trieste.it
LCG.UAM-FT.es	LCG.UCL-CCC.uk	LCG.UCL-HEP.uk	LCG.UERJ.br	LCG.UIB-BCCS.no
LCG.ULAKBIM.tr	LCG.UNI-KA.de	LCG.USC.es	LCG.WARSAW.pl	LCG.WCSS-slc4.pl
LCG.WCSS.pl	LCG.WEIZMANN.il	PPS.BHAM.uk	PPS.CERN.ch	PPS.CESGA.es
PPS.CESNET.cz	PPS.CMFP.it	PPS.CTF-KR.pl	PPS.DLUGE.it	PPS.ESA.int
PPS.FRAUNHOFER.de	PPS.GRIDKA.de	PPS.IC.uk	PPS.IEETA.pt	PPS.IJS.si
PPS.IN2P3.fr	PPS.ISIT.it	PPS.IP.us	PPS.IR.it	PPS.PIC

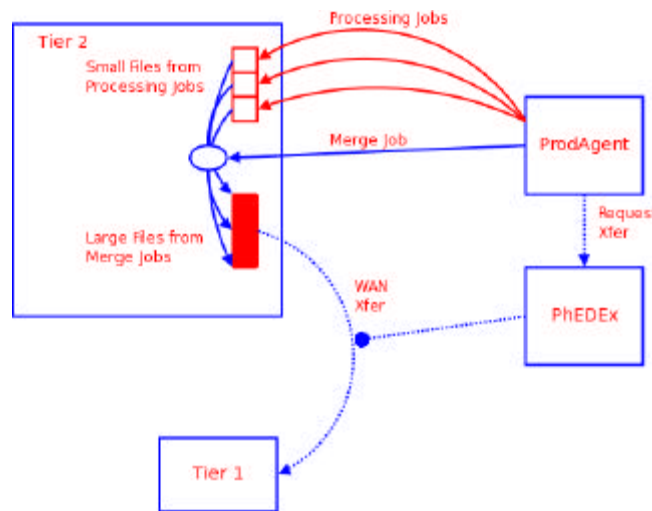
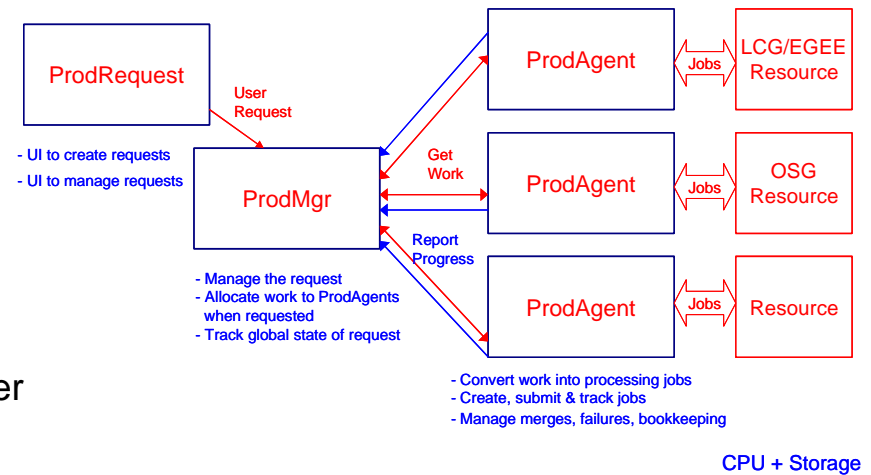
- An application to configure – prepare – submit and monitor jobs
 - Submitted to Grid or to a batch system
- Plugins for several applications
 - DaVinci, Athena, etc.
- Use case: LHCb
 - Test analysis code on local machine
 - Analyze partial sample on batch system
 - Analyze full sample on the Grid via DIRAC backend
- ~ 840 users (~10% of LCG users)
 - Mainly from ATLAS and LHCb but 20% non-HEP



- New developments
 - Ganga robot: to perform complex tasks like
 - Submitting jobs, extracting data about them, reporting stats
 - Windows port
 - Windows UI
 - Windows backend

- CMS Monte Carlo production
 - J. Hernández, D. Evans
- ATLAS Monte Carlo production
 - X. Espinal, S. Campana, Y. Smirnov

- 2007 goal: 50M events/month
- 2008 goal: $1.5 \cdot 10^9$ simulated events/year
- ProdAgent, a new production system
 - Rewritten from scratch
 - Scalable, avoids single points of failure, supports many grids, better automation
 - Used also for data processing (reconstruction, skimming, etc.)



Production Period	Number of MC events	Production Rate	Production Teams / PAs
Winter07	90 M	35M/month	6 / 10
Spring07	67 M	46M/month	6 / 9
Summer07	145 M	64M/month	5 / 7

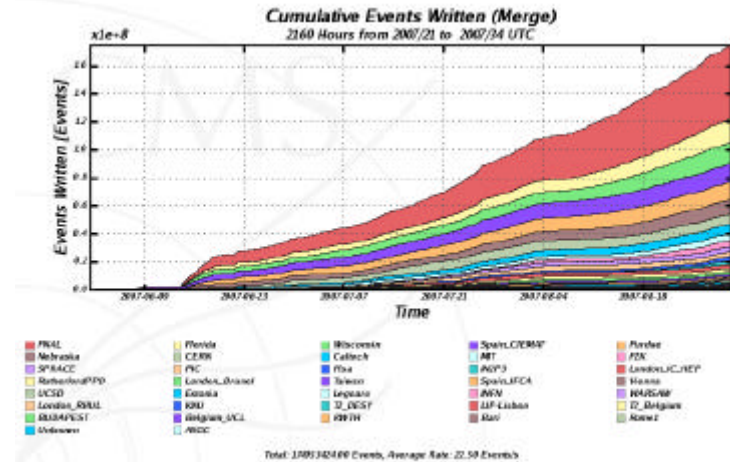
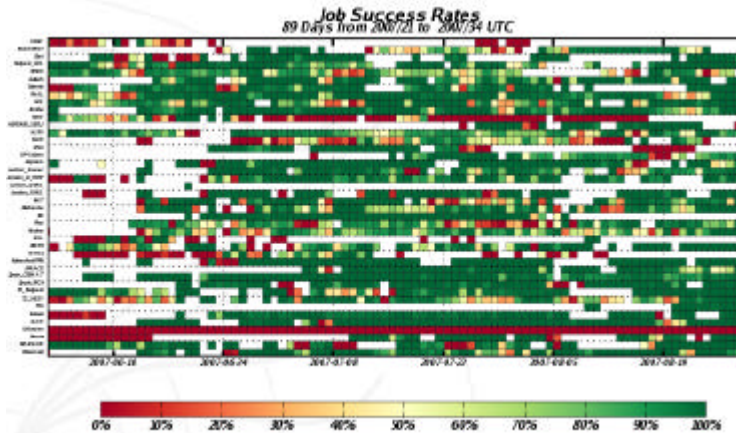
Overall MC production @ CMS in 2007

~5K slots available at T0/T1

~6K slots available at T2

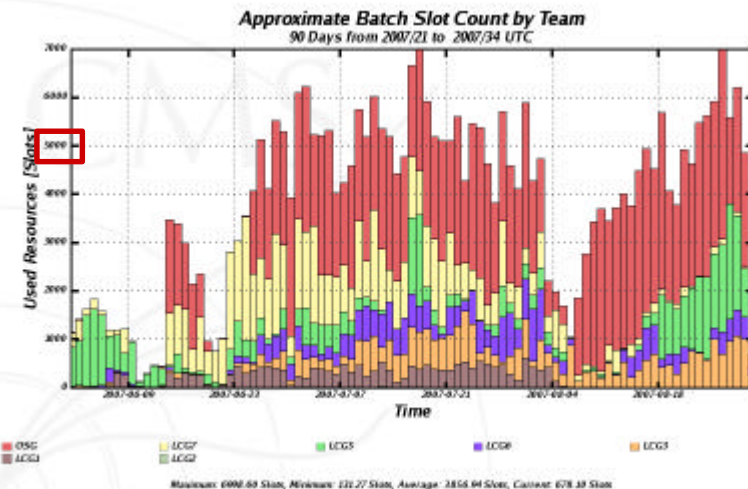


- ~175 M events in last 3 months
- ~ 2/3 done at Tier-2 sites
- ~ 50% done at OSG sites

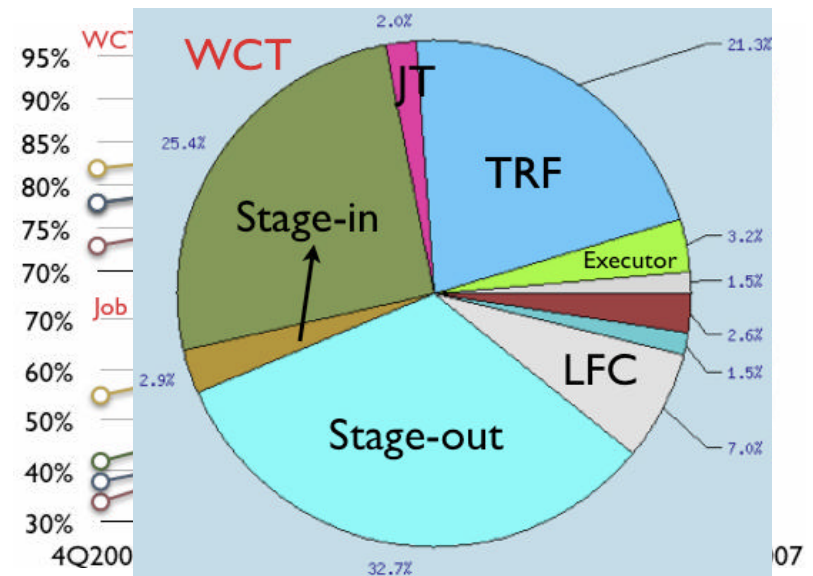
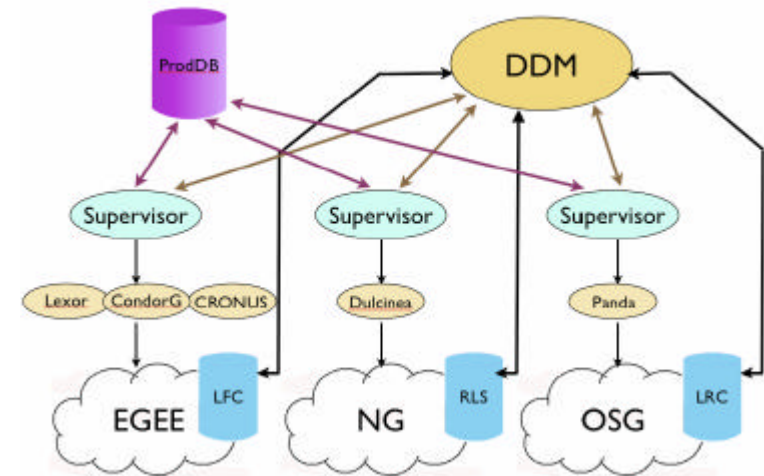


- 40+ sites used
- > 20K/day production jobs
- Average job efficiency ~ 75%, including application and Grid

- Ramp up from June (~ 5000 slots used in average)
- In average ~50% resource occupation. Production inefficiencies, no automatic resource management (manual job release by operators), workflow management (many sites, many small production requests)



- Prodsys Architecture
 - ProdDB: common database
 - Eowyn: common supervisor
 - Executors: Lxor, CondorG, CRONUS, Dulcinea, Panda
 - Integrated with the Distributed Data Management
- Operations
 - Production coordinators: assign tasks to clouds, monitor activities
 - Shifters
 - WM: monitor jobs
 - DM: control data flow
 - Need for more automation!
 - Error spotting and reporting
 - Monitoring
 - Automatic task assignment
- Performances
 - 60M events produced in 1st quarter of 2007
 - Up to 55K finished jobs/day
 - Job efficiency ~60%
 - WCT efficiency improving with time and reaching 90%
 - 65% of the inefficiency due to data I/O



- The parallel sessions provided a very comprehensive coverage of the HEP Grid computing
 - Some topics were particularly represented, even on more sessions
 - WMS systems, tools for production/analysis, storage systems, etc.
- Workload management
 - There are many systems developed by the experiments, but very few which are part of a middleware stack
 - Systems based on pilot jobs are becoming more popular, mainly because they succeed in "hiding" the Grid from the users
 - Still, "push-mode" systems are as reliable and scalable as needed

- Data Management
 - Here the top priority is on releasing production-quality SRM 2.2-compliant versions of the SEs
 - The experiments seem to be quite satisfied with their DM systems
- Monitoring
 - Still a lot of work to do to have standardized monitoring systems, but also a lot of new developments
 - SAM and the Dashboard rapidly expanding in functionality and effectiveness
- Experiment activities
 - Taking as example the MC production systems, they are able to reach the targets, but at the expense of a lot of manpower
 - Most of the development goes into improving automation and error recovery