



Fabric Monitoring at the INFN Tier1

Felice Rosso on behalf of
INFN Tier1
Felice.Rosso@cnafr.infn.it

Joint OSG & EGEE Operations WS, Culham (UK) - 28-9-2005



Outline

- CNAF-INFN Tier1
- FARM Monitoring
- Local Queues Monitoring
 - Local statistics
- Storage Monitoring
- Summary



Introduction

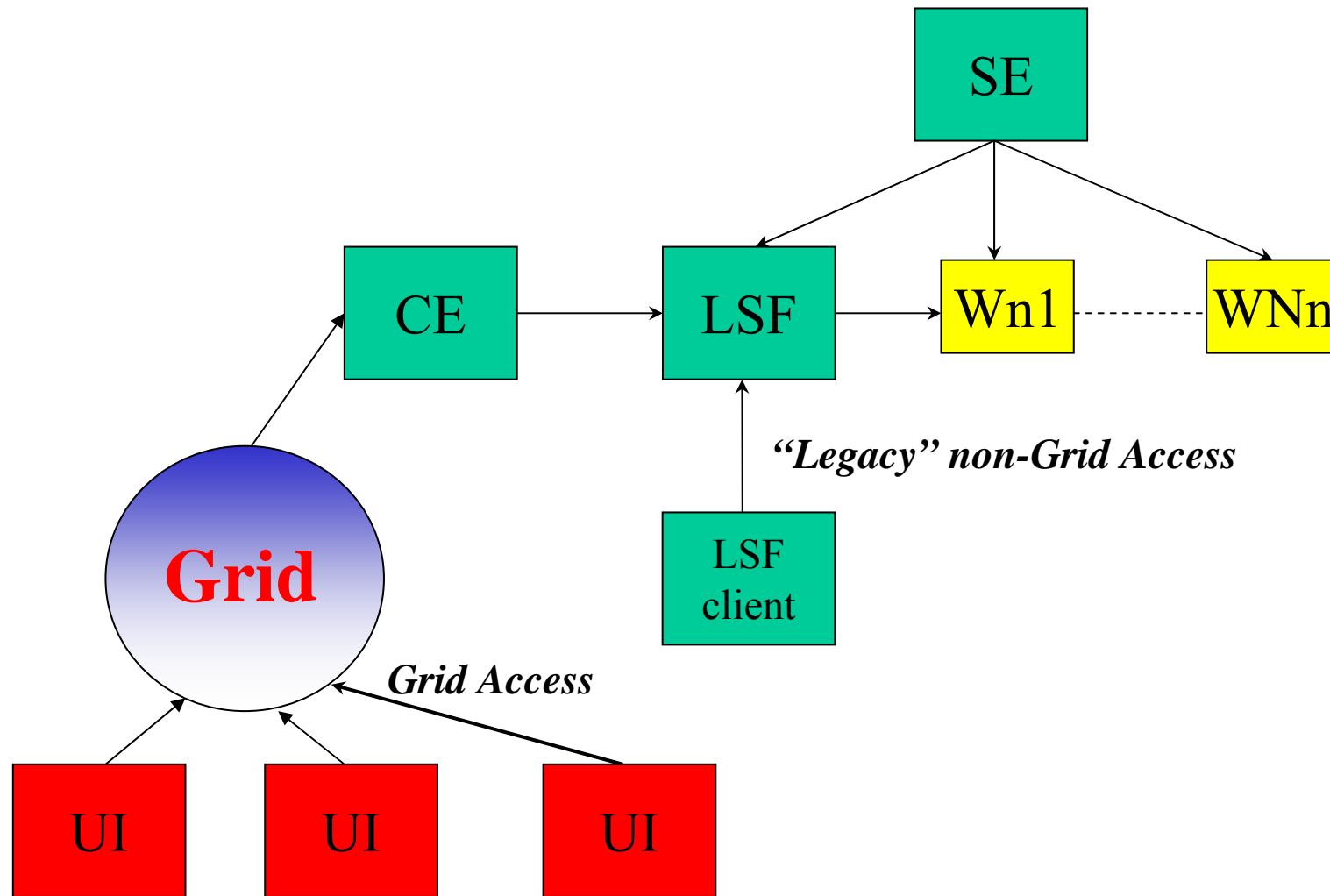
- Location: INFN-CNAF, Bologna (Italy)
 - one of the main nodes of GARR network
- Computing facility for INFN HNEP community
 - Participating to LCG, EGEE, INFN GRID projects
- Multi-Experiments TIER1
 - LHC experiments (Alice, Atlas, CMS, LHCb)
 - CDF, BABAR
 - VIRGO, AMS, MAGIC, ARGO, Pamela ...
- Resources assigned to experiments on a yearly Plan.



The Farm in a Nutshell

- SLC 3.0.5, LCG 2.6.0, LSF 6.1
- ~ 500 WNs LSF pool (484 GRID usable)
 - 1958 CPU slots usable (ht, 1878 GRID usable ~96%)
- Common LSF pool: 1 job per logical cpu (slot)
 - MAX 1 process running at the same time per job
- GRID and local submission allowed
 - On the same WN can run GRID and not GRID jobs
 - On the same queue can be submitted GRID and not GRID jobs
- For each VO/EXP one or more queues
- Since 24th of April almost 900.000 jobs were executed on our LSF pool (~455.000 GRID)

Access to Batch system





Farm Monitoring Goals

- Scalability to Tier1 full size
- Many parameters for each WN/server
- DataBase and Plots on Web Pages
- Data Analysis
- Report Problems on Web Pages
- Share data with GRID tools
- RedEye: T1 local tool monitoring

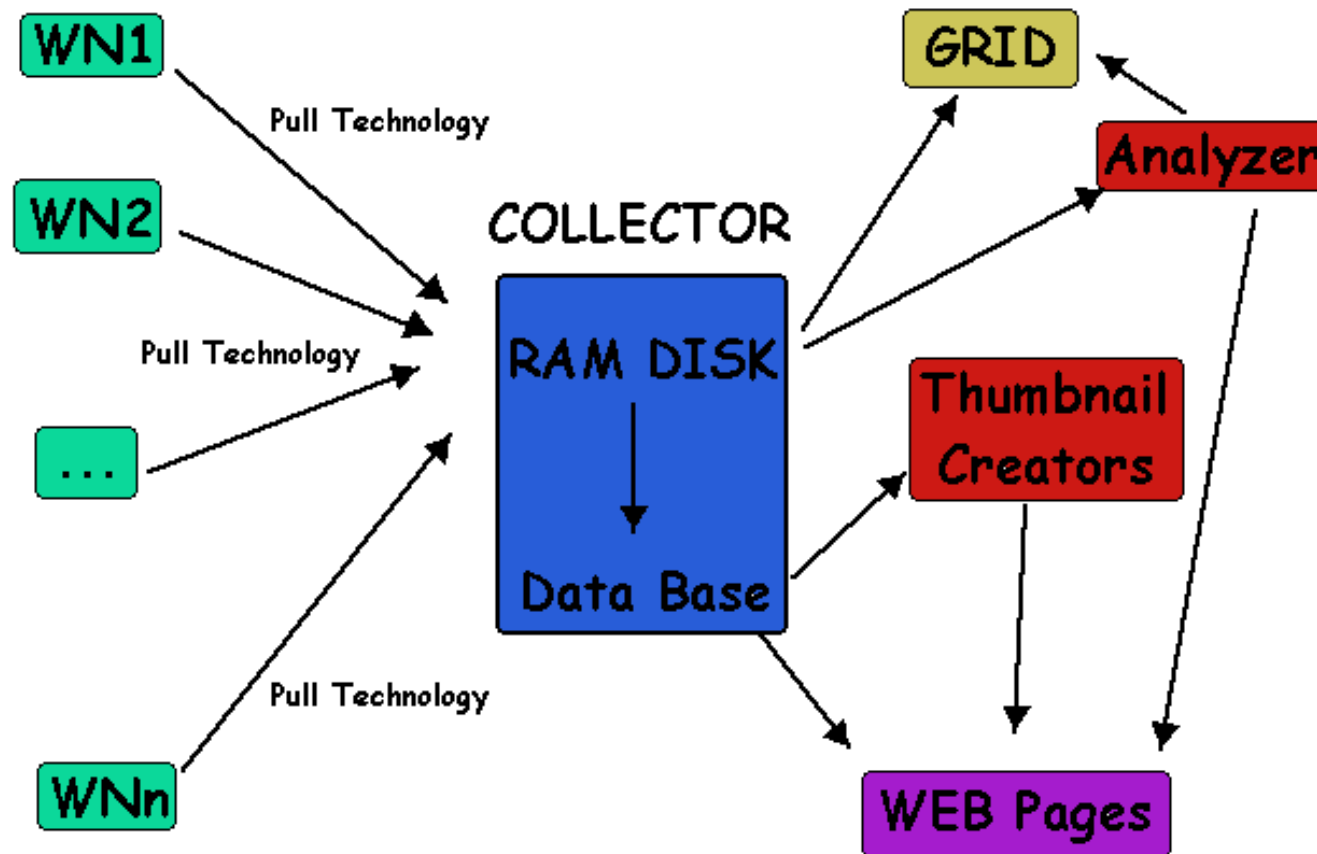


Farm Monitoring

What do we get?

- CPU load, status and jiffies
- IDE and SCSI I/O
- Ethernet I/O, packets, errors, overruns
- Temperatures, RPM fans
- Total and type of active TCP connections
- Processes created, running, zombie etc
- RAM and SWAP memory
- Users logged in

Local WN Monitoring



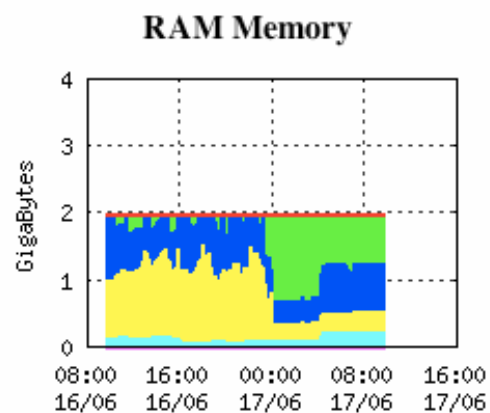
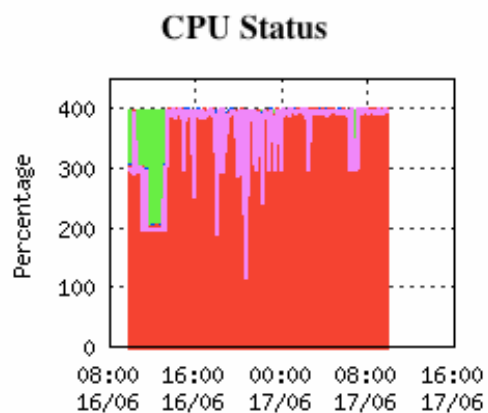
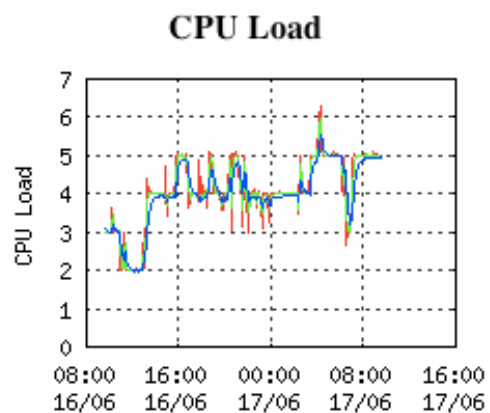


Local WN Monitoring

- On each WN every 5 min (local crontab) infos are saved locally (<3KBytes --> 2-3 TCP packets)
- 1 minute later a collector "gets" via socket the infos
 - "gets": tidy parallel fork with timeout control
- To get and save locally datas from 500 WN ~ 4 sec (best case). 15 sec worst case (timeout knife)
- Upgrade DataBase (daily, weekly, monthly, yearly)
- Each WN --> 1 file (possibility of cumulative rack plots)
- Check values (temperatures too high etc.)
- Local thumbnail cache creation (web clickable)

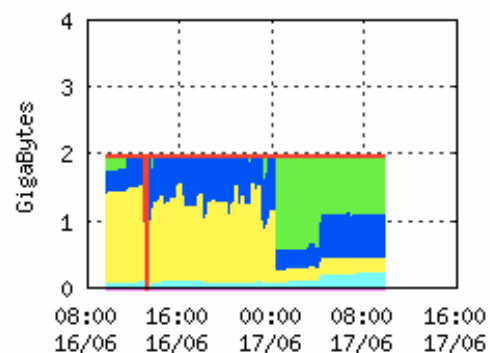
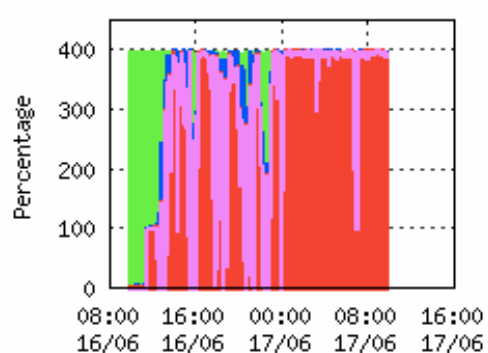
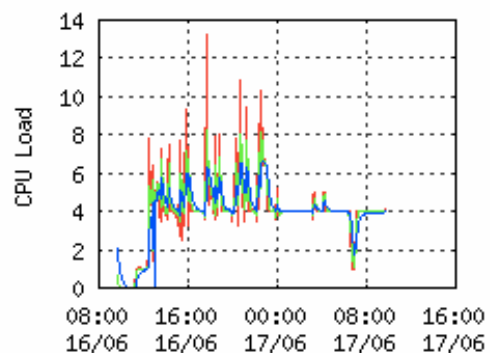


Web Snapshot CPU-RAM



Node

wn-03-07-35-a

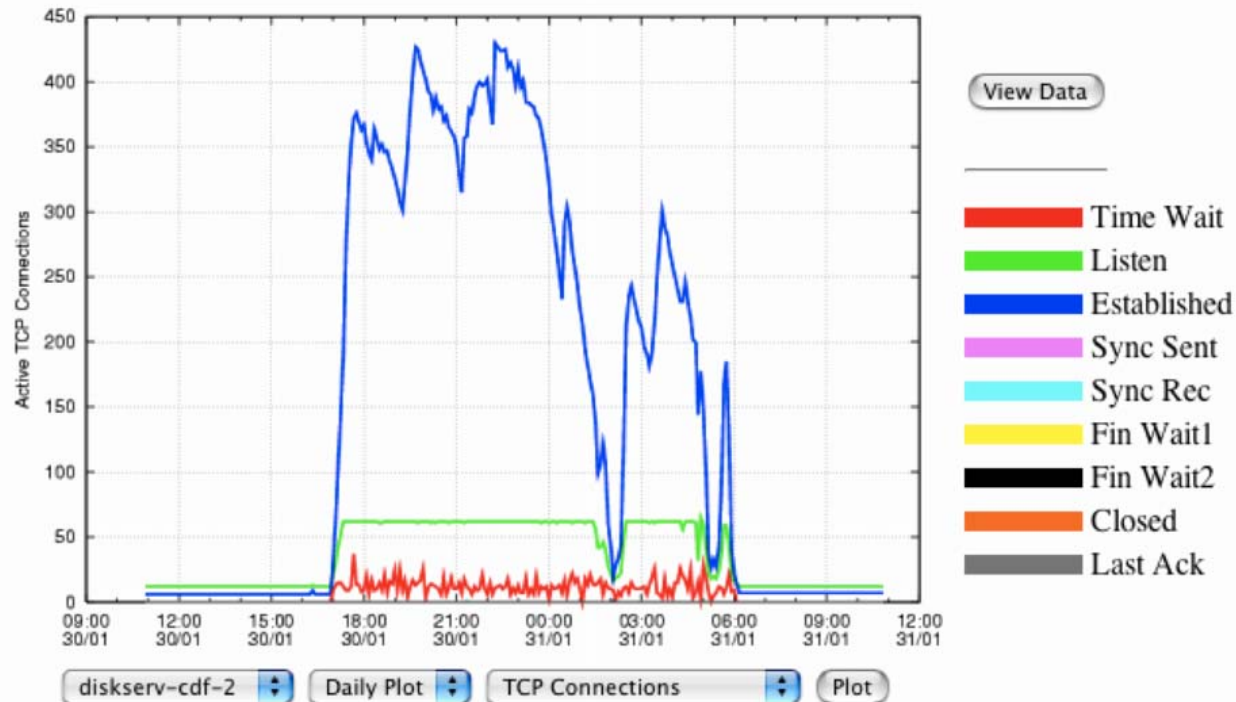


wn-03-07-34-a



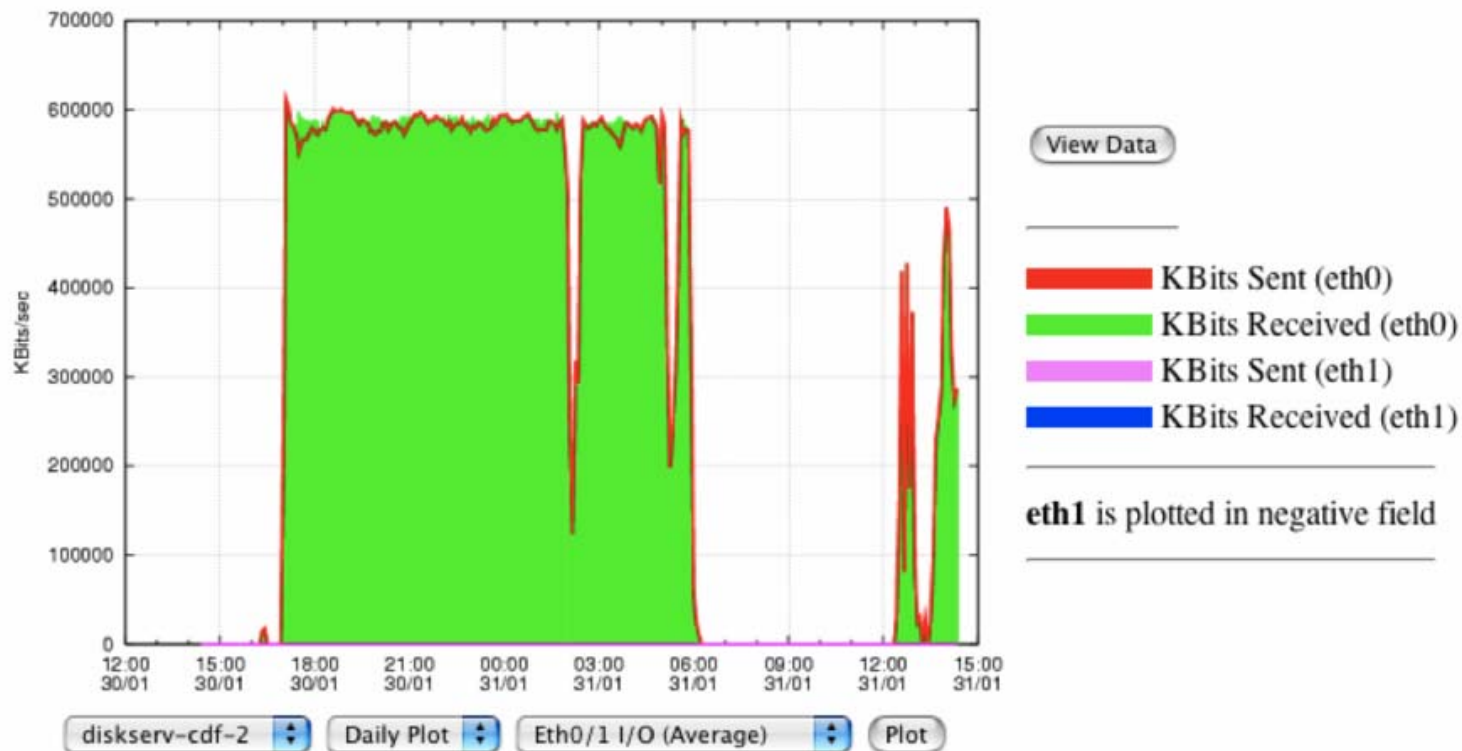
Web Snapshot TCP connections

TCP connections on diskserv-cdf-2
(Updated every 5 minutes)



Web Snapshot NIC

Eth0/1 I/O on diskserv-cdf-2
(Updated every 5 minutes)





Fabric → GRID Monitoring

- Effort on exporting relevant fabric metrics to the Grid level e.g.:
 - # of active WNs,
 - average CPU load
 - etc...
- GridICE integration
 - Configuration based on Quattor
- Avoid duplication of sensors on farm



Local Queues Monitoring

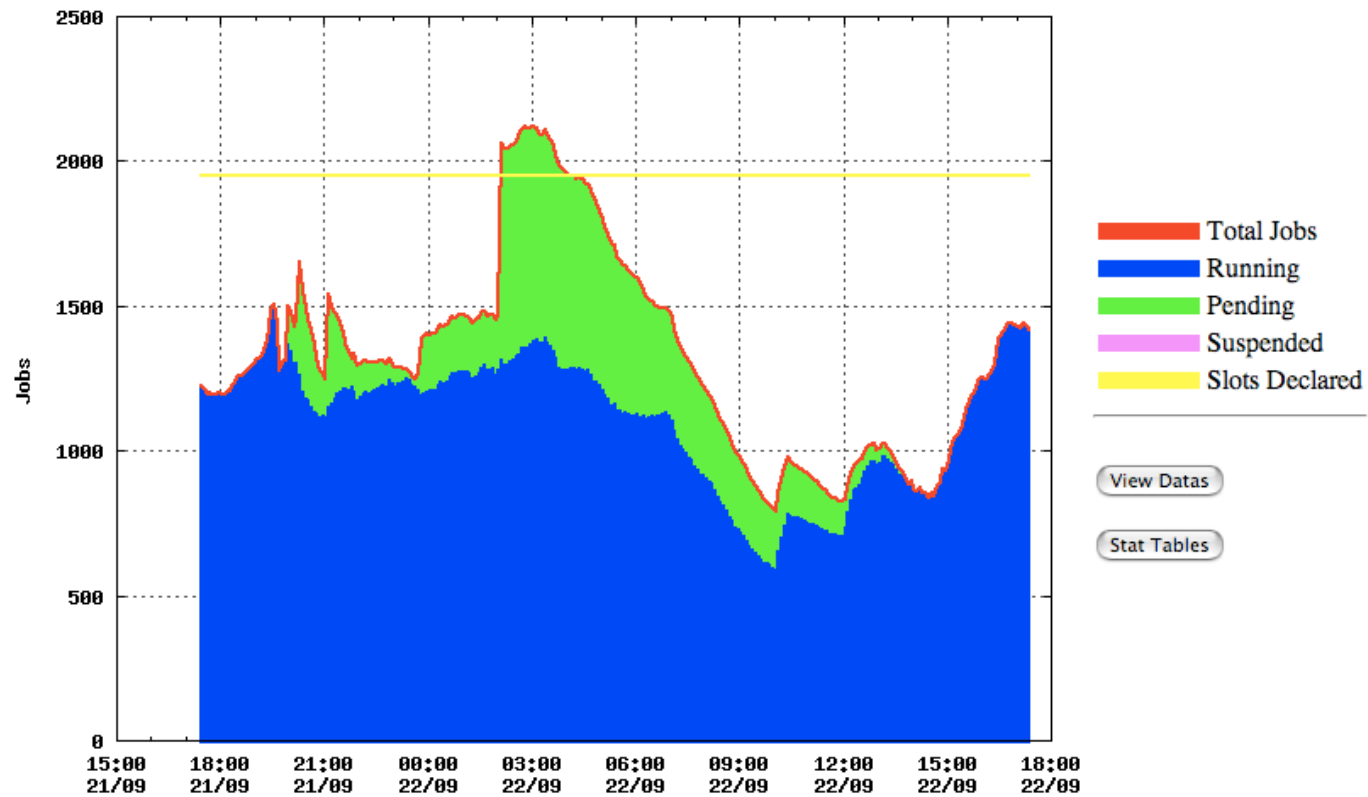
- Every 5 minutes on batch manager is saved queues status (snapshot)
- A collector gets the infos and upgrades the local database (same logic of farm monitoring)
 - Daily / Weekly / Monthly / Yearly DB
 - DB: Total and single queues
- 3 classes of users for each queue
- Plots generator: Gnuplot 4.0
- <http://tier1.cnaf.infn.it/monitor/LSF/>

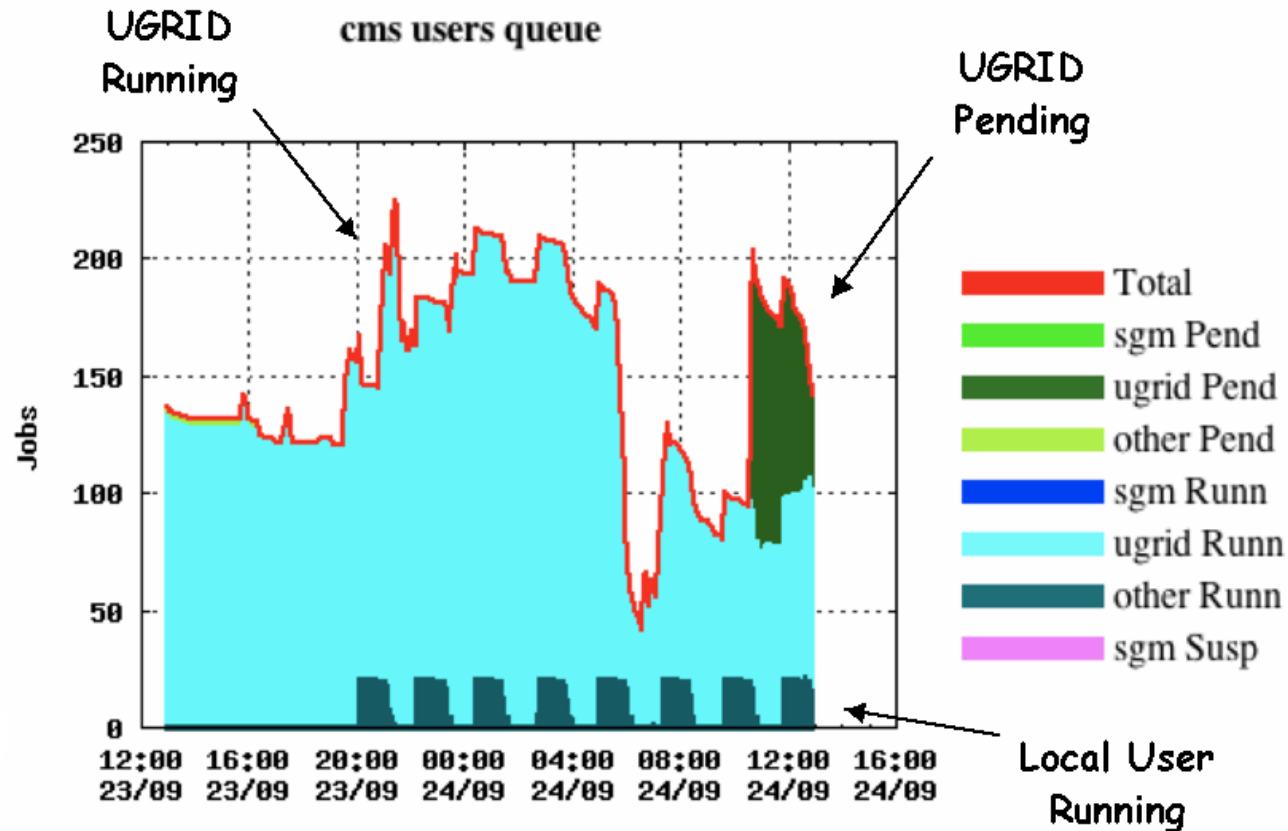


Web Snapshot LSF Status

[Daily Plot](#) [Weekly Plot](#) [Monthly Plot](#) [Yearly Plot](#)

LSF Queues @ CNAF
Auto-refresh every 5 minutes





UGRID: general GRID user (cms001, cms030...)
 SGM: Software GRID Manager (cmssgm)
 OTHER: local user



RedEye - LSF Monitoring

- Almost real time slot usage ☺
- Fast, few CPU power needed, stable, works on WAN ☺
- RedEye simple user, not root ☺

BUT...

1. all slots have the same weight
2. Jobs shorter than 5 minutes can be lost
3. No info about RAM or CPU time

AND

We need something good for ALL jobs.

We need to know who and how uses our FARM.

Solution:

Offline parsing LSF log file (lsb.acct) one time per day



Job-related metrics

From LSF log file we got the following non-GRID info:

- LSF JobID, local UID owner of the JOB
- "any kind of time" (submission, WCT etc)
- Max RSS and Virtual Memory usage
- From which computer (hostname) the job was submitted (GRID CE/locally)
- Where the job was executed (WN hostname)
- We complete this set with KspecINT2K of the slot

Queues usage report

Statistics for all jobs in week 37/2005 (from 12/09/2005 to 18/09/2005)

Queue	Jobs Completed	Total CPU Time Used (hours)	Total Time Used (hours)	Efficiency (percentage)	Average CPU Time per Job (seconds)	Average Virtual Memory per Job (MBytes)	Average RAM Memory per Job (MBytes)	KSpecINT2k (WCT) Test!
babar_objy	6018	13805.92	14937.81	92.42	8258.77	308.81	235.33	49.60
argo	765	4268.04	4626.82	92.25	20084.90	188.88	117.25	15.66
babar	8068	2372.54	2520.21	94.14	1058.65	479.76	455.39	7.67
pamela	10	0.28	0.51	54.38	99.84	182.53	54.83	0.00
babar_test	5	3.94	4.21	93.55	2838.05	234.12	164.45	0.01
lhcb	29419	43791.68	47255.64	92.67	5358.78	2062.18	1175.49	155.67
dteam	726	1.27	15.55	8.15	6.28	39.58	14.93	0.05
babar_build	349	0.92	2.28	40.36	9.49	38.41	30.67	0.01
quarto	82	2802.85	2808.72	99.79	123051.99	110.06	90.16	17.17
babar_xxl	5	4.64	4.68	99.14	3337.35	71.75	34.55	0.01
test	2885	3242.92	3323.42	97.58	4046.63	2053.30	1850.43	11.27
magic	10	70.70	70.81	99.86	25453.64	35.71	13.60	0.24
atlas	328	1581.23	1692.80	93.41	17354.96	218.02	120.24	5.61
cms_align	5	0.33	0.38	85.42	234.90	0.00	0.00	0.00
alice	211	1822.49	2018.43	90.29	31094.56	1146.78	838.03	6.83
cms	1092	10173.37	11078.15	91.83	33538.59	596.41	450.52	36.37
infngrid	15	0.02	0.09	18.13	3.83	28.35	10.53	0.00

Queues usage report

- GRID jobs come from CEs

Statistics for grid jobs in week 37/2005 (from 12/09/2005 to 18/09/2005)

Queue	Jobs Completed	Total CPU Time Used (hours)	Total Time Used (hours)	Efficiency (percentage)	Average CPU Time per Job (seconds)	Average Virtual Memory per Job (MBytes)	Average RAM Memory per Job (MBytes)	KSpecINT2k (WCT) Test!
dteam	726	1.27	15.55	8.15	6.28	39.58	14.93	0.05
lhcb	521	1.84	1148.83	0.16	12.73	54.50	20.19	3.54
atlas	328	1581.23	1692.80	93.41	17354.96	218.02	120.24	5.61
alice	211	1822.49	2018.43	90.29	31094.56	1146.78	838.03	6.83
cms	1092	10173.37	11078.15	91.83	33538.59	596.41	450.52	36.37
infngrid	15	0.02	0.09	18.13	3.83	28.35	10.53	0.00
cdf	1493	10433.86	10776.77	96.82	25158.67	229.23	103.34	36.55
Totale	4386	24014.07	26730	89.84	19710.59	311.12	201.56	88.96

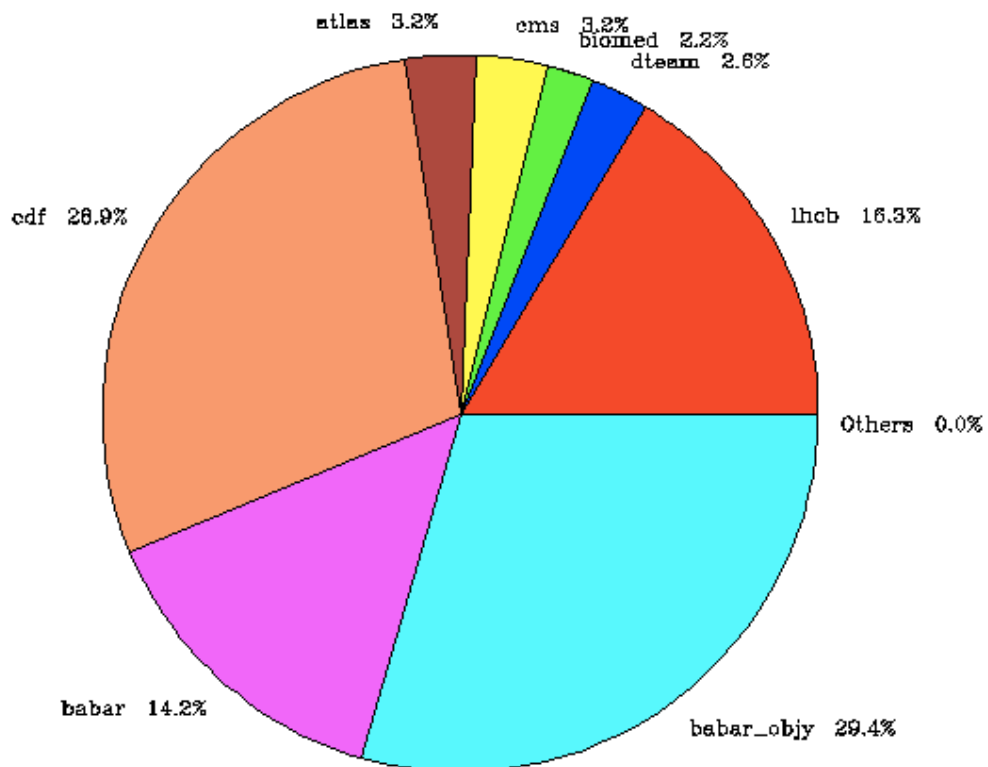
Queues usage report

Statistics for no-grid jobs in week 37/2005 (from 12/09/2005 to 18/09/2005)

Queue	Jobs Completed	Total CPU Time Used (hours)	Total Time Used (hours)	Efficiency (percentage)	Average CPU Time per Job (seconds)	Average Virtual Memory per Job (MBytes)	Average RAM Memory per Job (MBytes)	KSpecINT2k (WCT) Test!
babar_objy	6018	13805.92	14937.81	92.42	8258.77	308.81	235.33	49.60
argo	765	4268.04	4626.82	92.25	20084.90	188.88	117.25	15.66
babar	8068	2372.54	2520.21	94.14	1058.65	479.76	455.39	7.67
pamela	10	0.28	0.51	54.38	99.84	182.53	54.83	0.00
babar_test	5	3.94	4.21	93.55	2838.05	234.12	164.45	0.01
lhcb	28898	43789.84	46106.82	94.97	5455.17	2098.38	1196.32	152.13
babar_build	349	0.92	2.28	40.36	9.49	38.41	30.67	0.01
quarto	82	2802.85	2808.72	99.79	123051.99	110.06	90.16	17.17
babar_xxl	5	4.64	4.68	99.14	3337.35	71.75	34.55	0.01
test	2885	3242.92	3323.42	97.58	4046.63	2053.30	1850.43	11.27
magic	10	70.70	70.81	99.86	25453.64	35.71	13.60	0.24
cms_align	5	0.33	0.38	85.42	234.90	0.00	0.00	0.00
Totale	47100	70362.91	74406	94.57	5378.06	1538.48	957.74	253.78

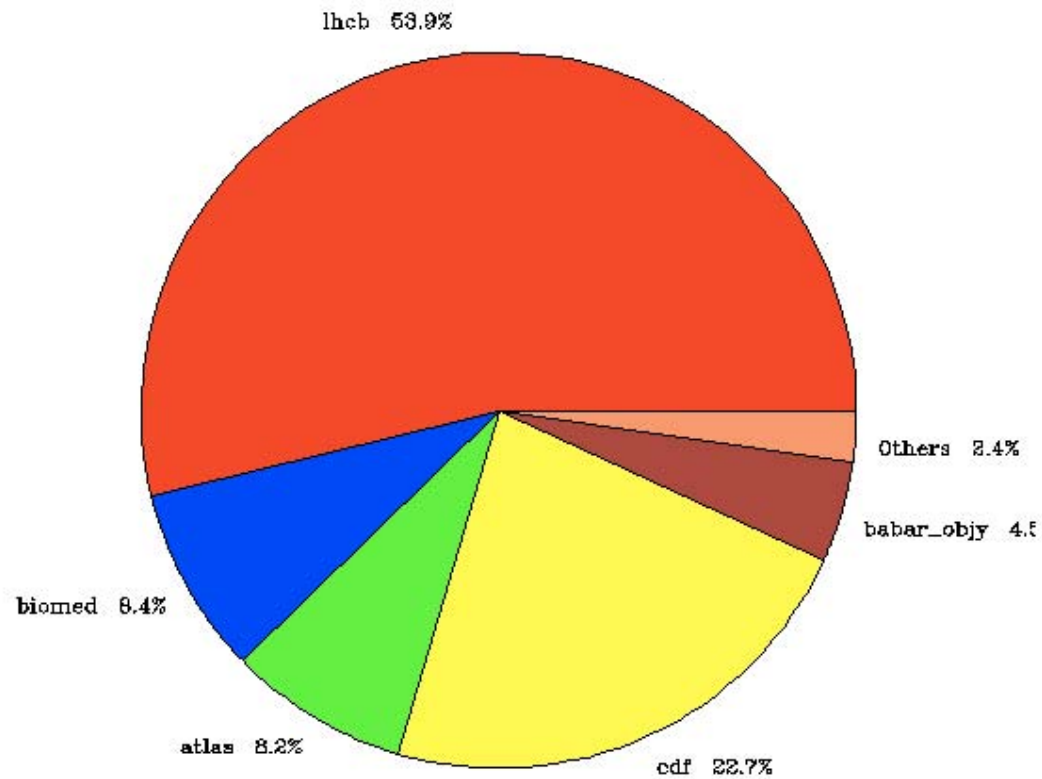
Queues usage report

Queue	Jobs Completed	%
lhcb	2532	16.326
dteam	402	2.592
biomed	337	2.173
cms	497	3.205
atlas	494	3.185
cdf	4475	28.854
babar	2206	14.224
babar_objy	4560	29.402
babar_xxl	4	0.026
babar_build	1	0.006
babar_test	1	0.006



Queues usage report

Queue	KSpec (WCT)	%
lhcb	151.36	53.889
dteam	0.03	0.012
biomed	23.69	8.435
cms	1.48	0.526
atlas	22.90	8.154
cdf	63.72	22.687
babar	4.58	1.632
babar_objy	12.59	4.481
babar_xxl	0.51	0.182
babar_build	0.00	0.000
babar_test	0.00	0.001



How we use KspecINT2K?

- 1 CPU → 1 job (XEON → 4 jobs; P3/Opteron → 2 jobs)

CPU Model	MHz	Total KspecINT2K	KspecINT2K per slot
XEON 3.06 GHz	3066.845	2.30	0.57
XEON 2.8 GHz	2799.313	2.10	0.53
XEON 2.6 GHz	2591.715	1.94	0.48
XEON 2.4 GHz	2395.991	1.80	0.45
XEON 2.2 GHz	2196.273	1.65	0.41
P3 1.4 GHz	1396.124	0.91	0.46
P3 1.2 GHz	1266.111	0.78	0.39
P3 1 GHz	993.401	0.65	0.33
P3 800 MHz	797.516	0.52	0.26

$$WeightJob = \frac{WCT(sec)}{86400(1day)} * \frac{KspecINT2K}{\# CPU(kernel)}$$



Fabric and GRID monitoring

- Effort on exporting relevant queue and job metrics to the Grid level.
 - Integration with GridICE (monitoring), DGAS (Grid accounting)
 - Grid (VO) level view of resource usage
- Integration of local job information with Grid related metrics. E.g.:
 - DN of the user proxy
 - VOMS extensions to user proxy
 - Grid Job ID



GridICE screenshots

GridICE the eyes of the Grid

Site view VO view Help About

is monitoring grid-logo.gif 70x70 change with the logo of the Grid being monitored

GridICE >> Site::ALL >> Site::INFN-T1

GridICE the eyes of the Grid

Site view VO view Help About

is monitoring grid-logo.gif 70x70 change with the logo of the Grid being monitored

GridICE >> Site::ALL >> Site::INFN-T1

CE	SE	Gris	Host	Job	Charts					
Hostname		Site	Domain	Middleware	Type	LastCheck	Conn	Since	Entries	Scheduling
sg01-lcg.cr.cnaf.infn.it		INFN-T1	cr.cnaf.infn.it	LCG-2_6_0	EX	0h2m14s	✓	2005-09-09 14:57	13	0
se01-lcg.cr.cnaf.infn.it		INFN-T1	cr.cnaf.infn.it	LCG-2_6_0	EX	0h2m14s	✓	2005-09-09 13:57	47	0
castorgrid.cr.cnaf.infn.it		INFN-T1	cr.cnaf.infn.it	LCG-2_6_0	SE	0h27m27s	✓	2005-09-16 09:42	24	0
diskserv-lhc-2.cr.cnaf.infn.it		INFN-T1	cr.cnaf.infn.it	LCG-2_6_0	SE	0h27m55s	✓	2005-09-15 02:43	6	0
ce01-lcg.cr.cnaf.infn.it		INFN-T1	cr.cnaf.infn.it	LCG-2_6_0	CE	0h4m23s	✓	2005-09-13 06:15	76	0
se01-lcg.cr.cnaf.infn.it		INFN-T1	cr.cnaf.infn.it	LCG-2_6_0	SE	0h24m38s	✓	2005-09-09 11:16	7	0
castorsrm.cr.cnaf.infn.it		INFN-T1	cr.cnaf.infn.it	LCG-2_6_0	SE	1573h37m7s	✓	2005-07-22 16:31	25	2
ce02-lcg.cr.cnaf.infn.it		INFN-T1	cr.cnaf.infn.it	LCG-2_6_0	CE	0h4m17s	✓	2005-09-13 12:11	22	0
ce01-lcg.cr.cnaf.infn.it		INFN-T1	cr.cnaf.infn.it	LCG-2_6_0	SB	0h1m15s	✗	2005-09-07 16:48	0	0
ce02-lcg.cr.cnaf.infn.it		INFN-T1	cr.cnaf.infn.it	LCG-2_6_0	SB	0h1m6s	✓	2005-09-18 16:20	131	0
sc.cr.cnaf.infn.it		INFN-T1	cr.cnaf.infn.it	LCG-2_6_0	SE	0h24m49s	✓	2005-09-15 02:46	24	0

Generated: Mon, 26 Sep 2005 14:40:30 +0200

GridICE Homepage



Job Check Control

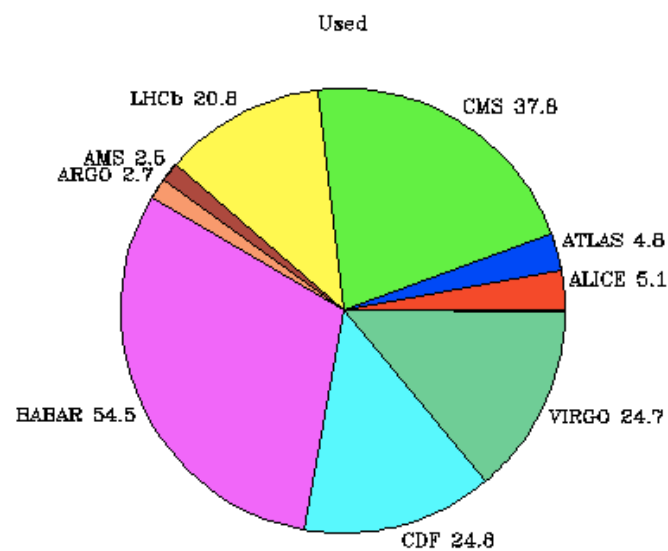
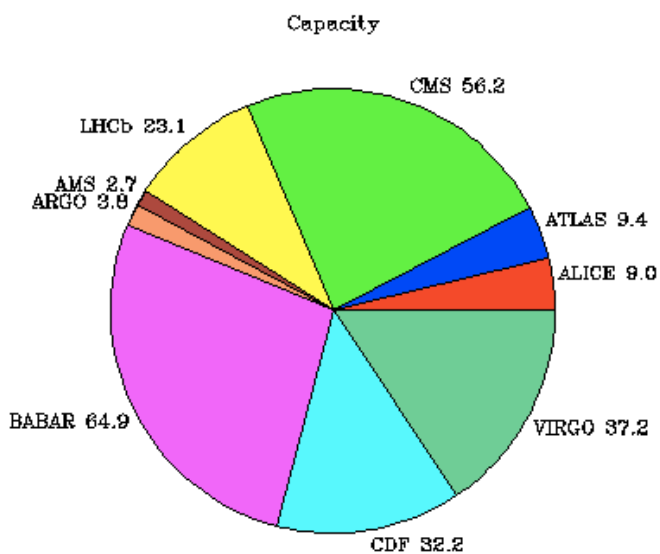
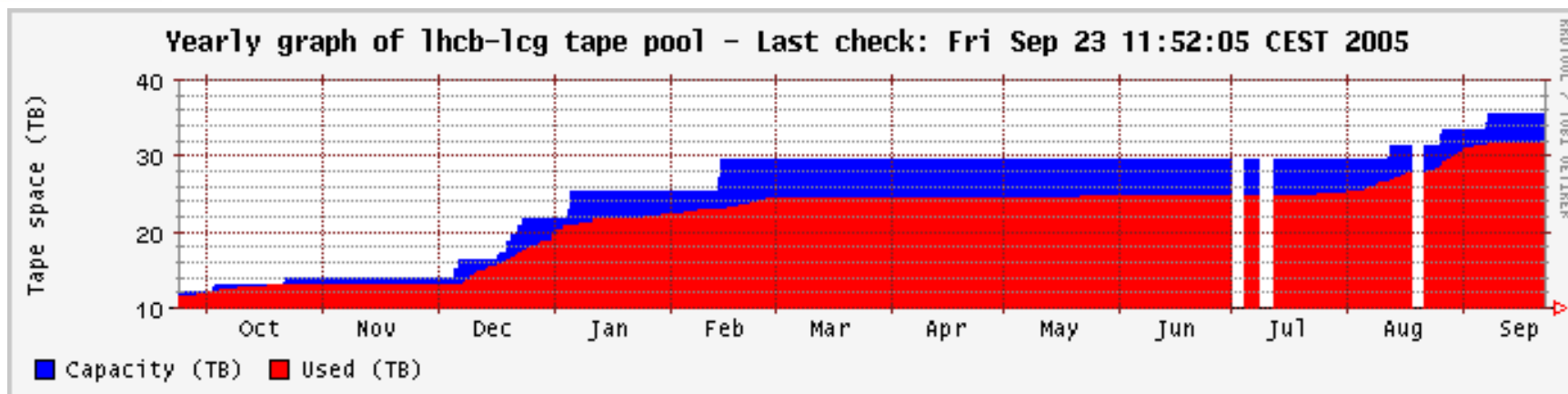
- Lsb.acct had a big bug!
 - Randomly: CPU-user-time = 0.00 sec
 - From bjobs -l <JOBID> correct CPUtime
 - Fixed by Platform at 25th of July 2005
- CPUtime > WCT? --> Possible Spawn
- RAM memory: is a job on the right WN?
- We have a daily report



The Storage in a Nutshell

- Different hardware (NAS, SAN, Tapes)
 - 275 TB HD, 130 TB Tape
- Different access methods (NFS/RFIO/Xrootd/gridftp)
- HD Volumes FileSystem: EXT3 and GPFS
- Tape access: CASTOR (50 TB of HD as stage)
- Volumes management via Postgresql DB
- 80 servers to manage and export FS
- Nagios to check hw & services status
 - gridftp, srm, rfio, castor, ssh
- RRD to plot (volume space total and used)

Storage usage report



Summary

- Fabric level monitoring is needed to ease management
- At Grid level more specific tools like GridICE and DGAS to publish the VOs, jobs and users related metrics are mandatory
- An effort to integrate these 2 levels is on-going
 - unify data collection
 - avoid duplications