Rapporto dalle VO LHC

M. Paganoni

Workshop CCR

Rimini, 7/5/2007

The goal of LHC VO

- ✓ The Data Management and Workload Mamagement of the LHC experiments are designed for a Grid-enabled world
- ✓ Commissioning towards fully functional distributed Computing System
- ✓ From development
  - service/data challenges (both WLCG wide and experiment specific) of increasing scale and complexity

  to operations
  - data distribution
  - MC production
  - analysis
- ✓ Primary needs:
  - Smoothly running Tier1's and Tier2's
  - Streamlined and automatic operations
  - Full monitoring and automatic recovery
  - Support continously significant load

Continuous loop

- ✓ Develop and commission tools for Operations
- ✓ Integrate and test middleware
- ✓ Commission system at scale
- ✓ Good monitoring and detailed analysis of all use cases is the key to success
  - Stress test <-> Measure <-> Debug
- ✓ For each service on each site establish "is working" tool and "what's wrong tool"
  - Prevent user's problem
  - Detct user's problem
  - Follow up on user problems, answer tickets and questions

CMS and EGEE middleware

- ✓ Data management
  - Central (single instance) CMS-own data catalogs (DBS/DLS)
  - PhEDEx layered above FTS and SRM
  - Trivial File Catalog at sites for LFC→PFN
- ✓ Workflow management
  - Direct submission via to Condor-G and/or gLite WMS, no pilots
  - Central production request repository with few Production submission agents
  - Users submit directly to Grid (CRAB), analysis server for users under development
- ✓ Databases: FroNtier solution within LCG-3D for T1/T2/T3
- ✓ Monitor: CMS-own Dashboard harvest info from all possible sources
- ✓ Accounting: rely on WLCG solution (must be same for all LHC)
- ✓ Policies: by VOMS groups/roles. Something done already with Unix groups/users. Rely on Grid for long term solution

CMS data management

- ✓ PhEDEx layered above FTS and SRM
  - need SRMv2 (still not there)
  - interoperability still a big problem since years !
  - need much better FTS then we have now (channel mgt, monitor)
- ✓ Dataset Bookeeping Service (What data exist ?)
  - Define and discover CMS data sets
- ✓ Dataset Location Service (Where are data located?)
  - Locate replicas of data sets in the distributed system
  - DLS currently layered on LFC, will probably move to a couple of tables in larger DBS
- ✓ Local Catalog at sites (LFN to PFN mapping)
  - Presently using a Trivial File Catalog (parsing rules)

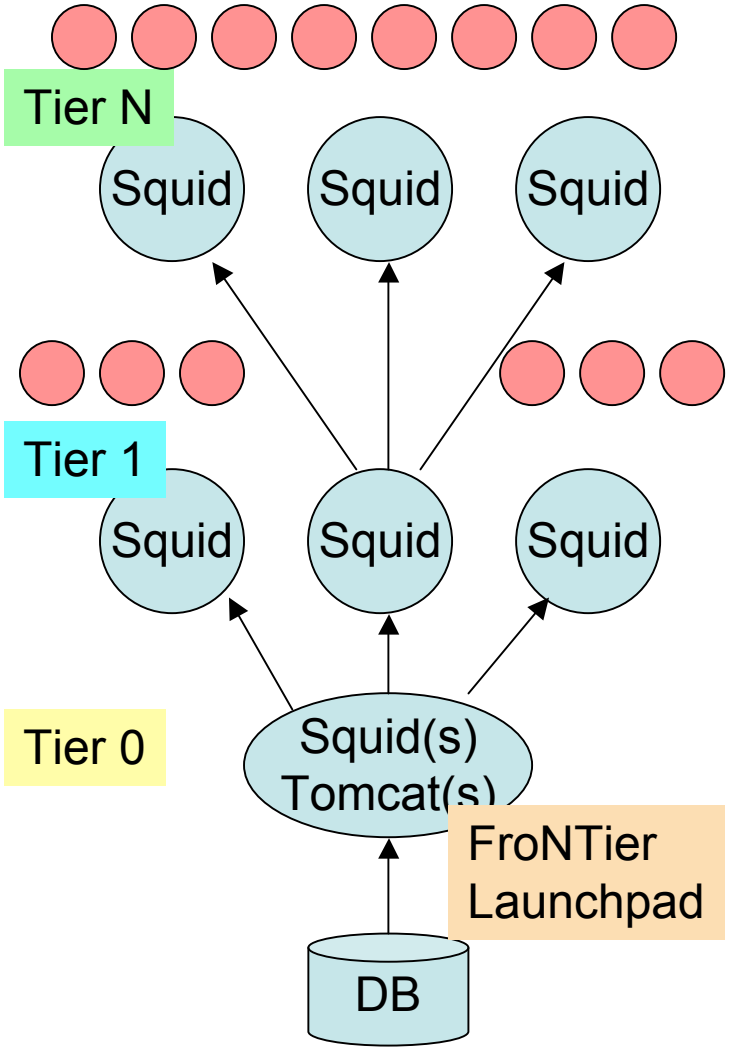
CMS - distributed database

- ✓ Frontier transforms SQL queries into http URL access to allow use of Squid. Tested successfully in CSA06

- ✓ Tier-0 (CERN):
  - Central offline calibration and alignment DB in place.
  - Frontier stateless server sends payloads over HTTP to clients.

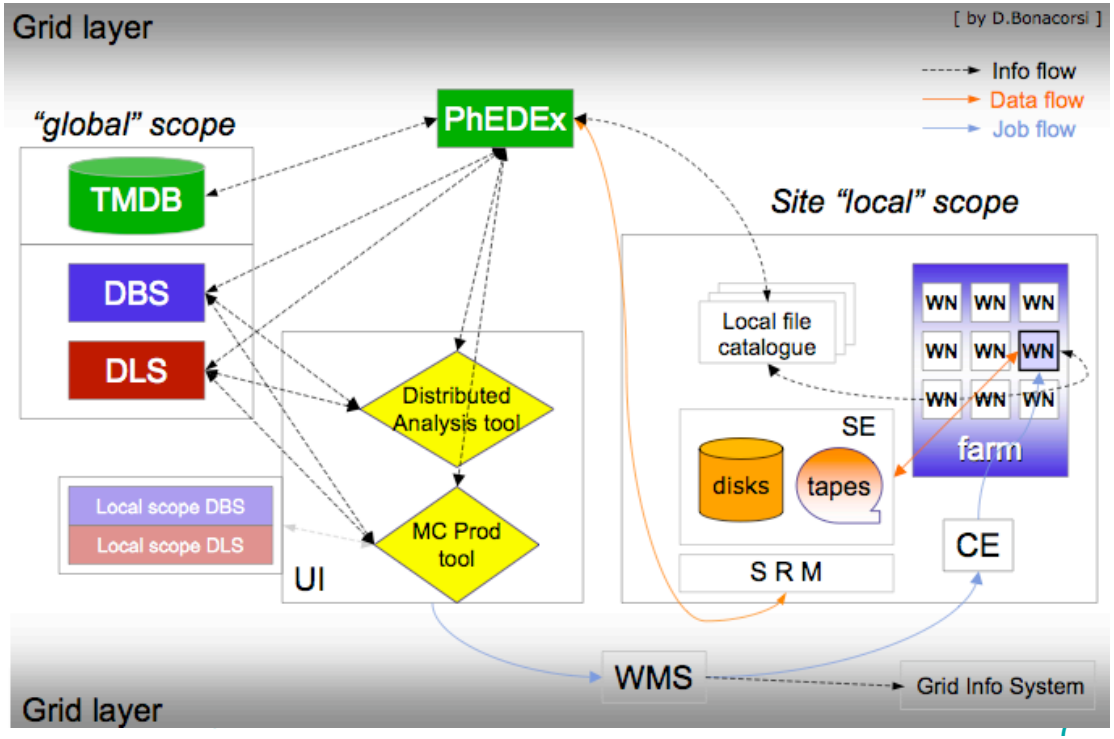
- ✓ Tier-1:
  - Caching servers (Squid) cache and deliver data to clients, and Tier-N sites.
  - Squid servers are stateless. Machines can failover without data corruption or loss.

- ✓ Tier-N (N > 1):
  - Caching servers (Squid) cache and deliver data to clients.
  - Single machine sufficient.



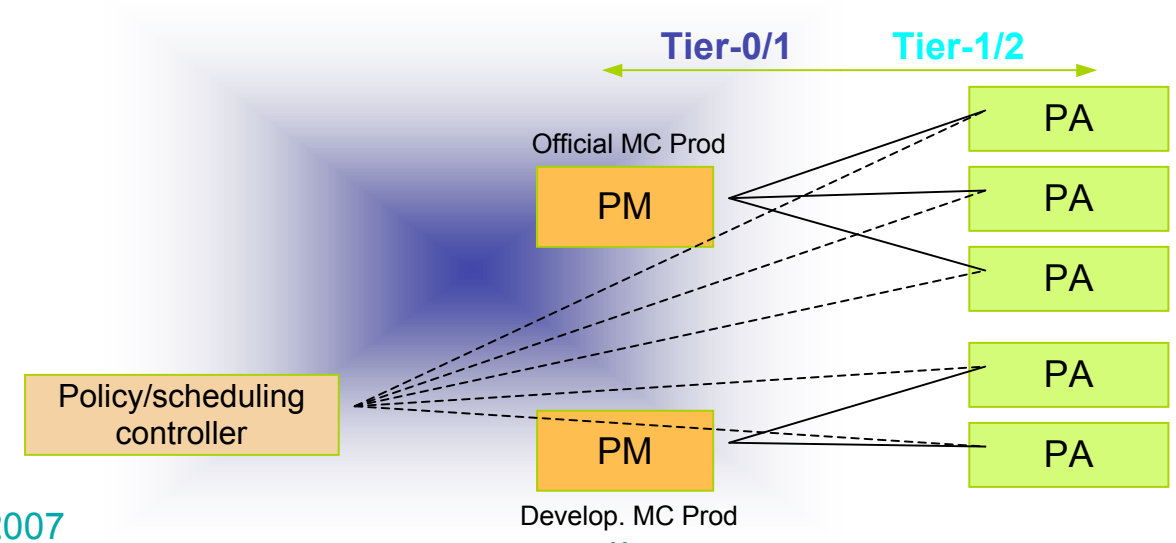
CMS - data placement (PhEDEx)

- ✓ Data placement system for CMS (in production since 3 years)
  - large scale reliable dataset/fileblock replication
  - multi-hop routing following a transfer topology (T0 → T1's ↔ T2's), data pre-stage from tape, monitoring, bookkeeping, priorities and policy, etc
- ✓ Managing transfers of several TB/day
- ✓ PhEDEx integration with gLite services File Transfer Service (FTS)
- ✓ During CSA06 for T0 --> T1s
  - 70 TB in 1 month
  - up to 6 TB/day



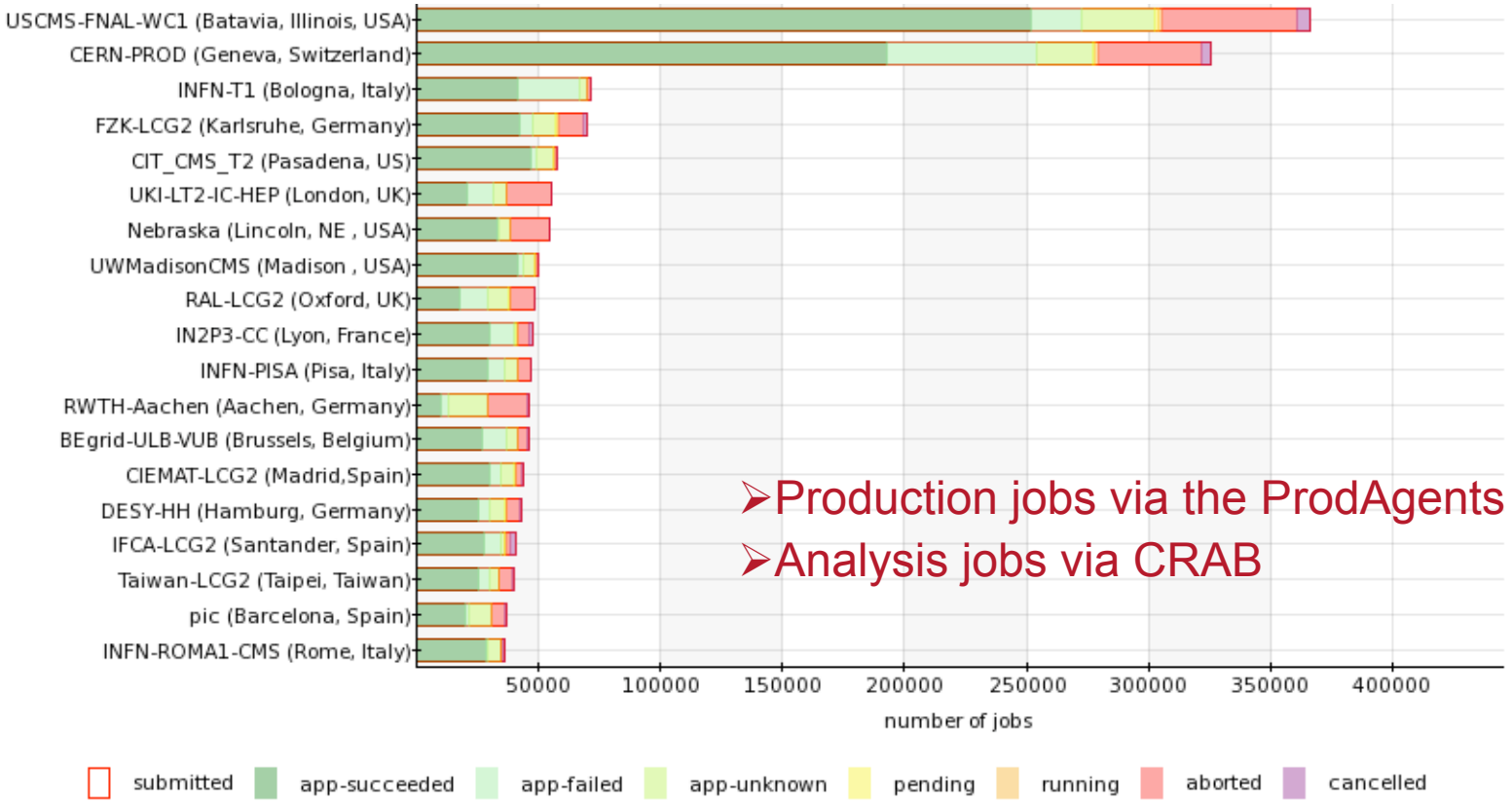
CMS - MC production

- ✓ New MC production system developed in 2006, in production already
  - less man-power consuming, better handling of Grid-sites unreliability, better use of resources, automatic retrials, better error report/handling
  - More flexible and automated architecture
    - **ProdManager** (*PM*) (+ the policy piece)
      - manage the assignment of requests to 1+ *ProdAgents* and tracks the global completion of the task
    - **ProdAgent** (*PA*)
      - Job creation, submission and tracking, management of merges, failures, resubmissions,
  - Integrate with new Event Data Model and new DMS



CMS workload management

- ✓ Direct jobs submission to Condor-G and/or gLite WMS
  - Presently using only EDG-RB in EGEE sites and Condor-G in OSG sites, but moving to inter-operability
  - Analysis server to automatize error recovery, small files merging and interaction with DM for largish user's tasks.



CMS - user analysis

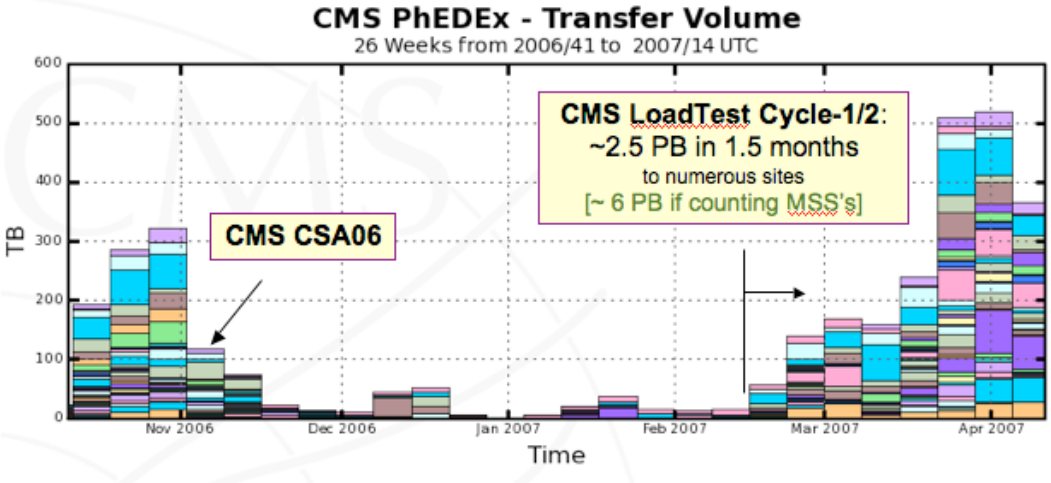
- ✓ CRAB provides an user friendly interface for end user interaction with the grid for CMS, including interaction with data management, middleware, remote computing element, basic monitoring functionalities, etc. . . hiding as much as possible the grid complexities to final user
- ✓ Used successfully since 2 years and tested in CSA06
- ✓ New version (CRAB server) being deployed
  - Automatize as much as possible the interactions with the grid, including submission, resubmission, error handling, output retrieval, etc . . .
  - Improve scalability of the whole system
  - More than 6600 jobs (256 tasks) submitted to server in 2h
  - Good initial success ratio: 99,6%
  - Foreseen job submission rate to Grid: about 10 kjob/day

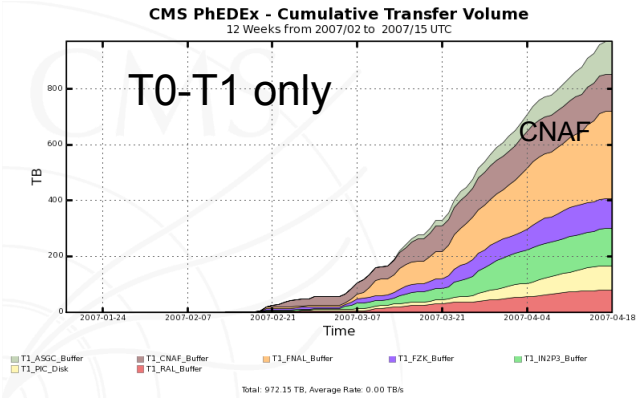
CMS - monitoring

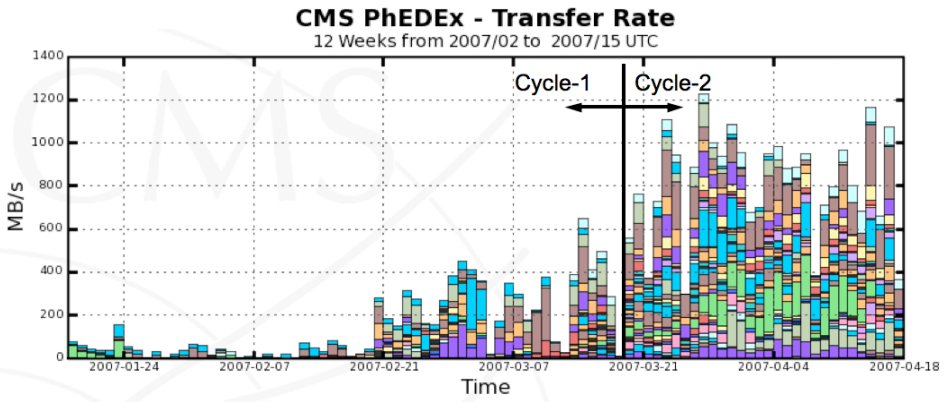
- ✓ CMS relies on monitoring (Dashboard) to run operations
- ✓ Redundant logging of CMS information to central Oracle DB http://arda-dashboard.cern.ch/cms/ (used also by Atlas, LHCb)
  - All CMS jobs report to via MonaLisa from WN
  - All CMS workload tools (CRAB, ProdAgent) report via MonaLisa at submission/check/retrieval
  - Allow to correlate using CMS variables
  - Grid/application exit code, submission tool used, executable (version), activity (test, production, analysis…), dataset, I/O rate, waiting/pending/running time, group, role, user, execution site, submission site…
  - Central DB tested to scale up to 200K jobs/day (can do more)
  - Extensive "browsing" by interactive query and web server
  - Pre-defined views using RRDTool
- ✓ PhEDEx has extensive monitoring and graphics, production grade since years, currently being improved (UK)

CMS LoadTest 2007

- ✓ An infrastructure by CMS to help Tiers to exercise transfers
  - Based on a new traffic load generator
  - Coordination within the CMS Facilities/Infrastructure project
- ✓ Exercises
  - T0→T1(tape), T1↔T1, T1↔T2 ('regional'), T1→T2 ('non-regional')

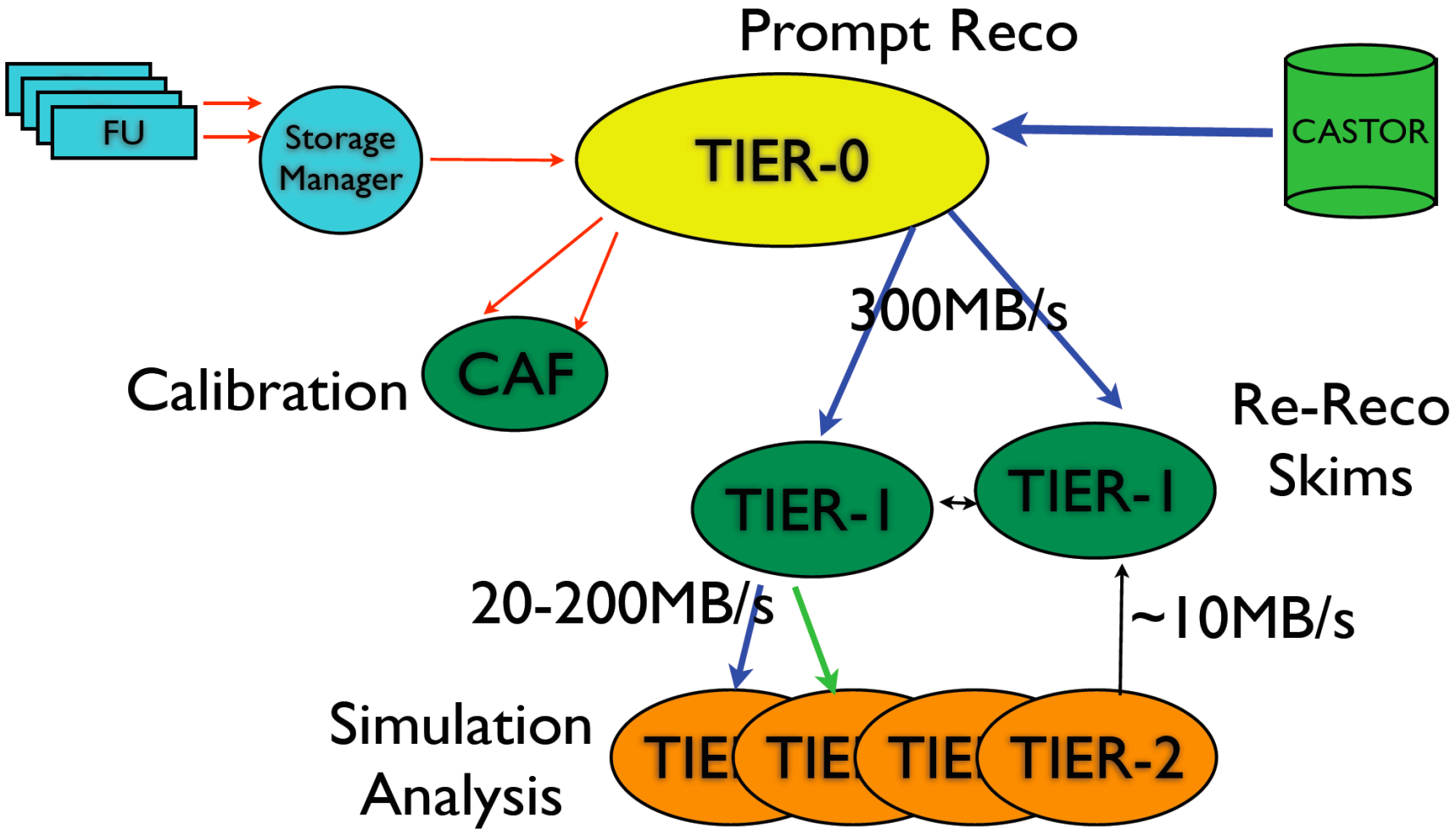






- ✓ Important achievements
  - **routinely** transferring
  - all Tiers report it's **extremely useful**
  - **higher participation** of Tiers (see plot)
  - less efforts, **improved stability**

CSA07 workflow



CSA07 success metrics

Service	2008 Goal	2007 Goal	%
Tier-0 Reco Rate	150Hz - 300Hz	100Hz	50%
Network Transfers between T0-T1	600MB/s	300MB/s	50%
Network Transfers between T1-T2	50-500 MB/s	20-200 MB/s	40%
Network Transfers T1-T1	100MB/s	50MB/s	50%
Job Submission to Tier-1s	50k jobs/d	25k jobs/d	50%
Job Submissions to Tier-2s	150k jobs/d	75k jobs/d	50%
MC Simulation	1.5 10^9 events/year	50M per month	50%

ATLAS - WMS

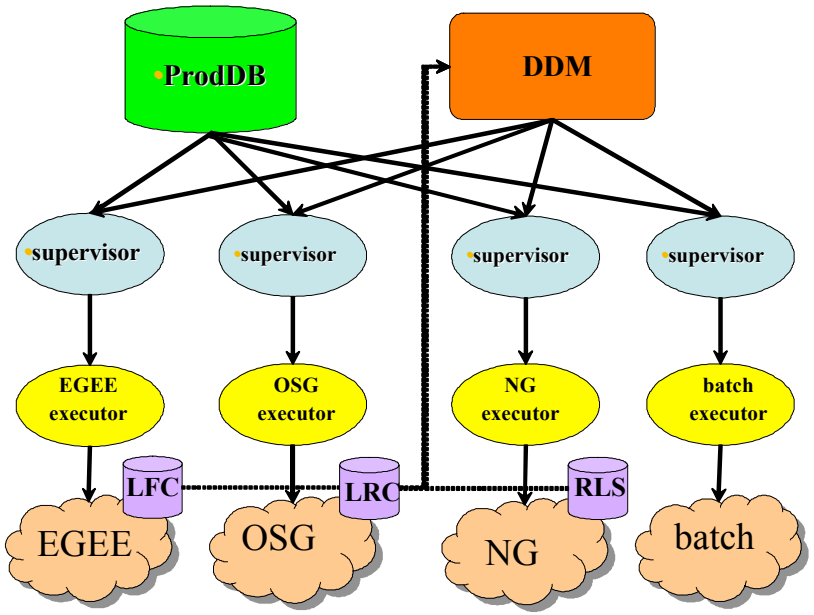
- ✓ ATLAS utilizza i tool di Grid per:
  - produzione di dati simulati
  - Analisi distribuita
- ✓ Attualmente è in corso il CSC (Computing System Commissioning) per la validazione dell'intera catena dalla simulazione all'analisi sulle 3 Grid LCG/EGEE (Europa), OSG/Grid3 (US), America, NorduGrid (Nord Europa)
- ✓ Negli ultimi 10 mesi :
  - 100 siti coinvolti
  - 1.7 Mjobs eseguiti
  - 175 Mevts
  - 820 CPU/giorno con picchi di 3500 CPU/giorno
- ✓ Per la produzione si utilizzano:
  - Condor-G
  - Lexor (basato su gLite WMS in Europa)
- ✓ 60% dei job eseguiti in EGEE
- ✓ > 50 % dei job in EGEE sottomessi via Lexor

ATLAS - produzione distribuita

• Per produrre i dati necessari alle produzioni è stato sviluppato un sistema di produzione automatizzato ⇒

ProdSys

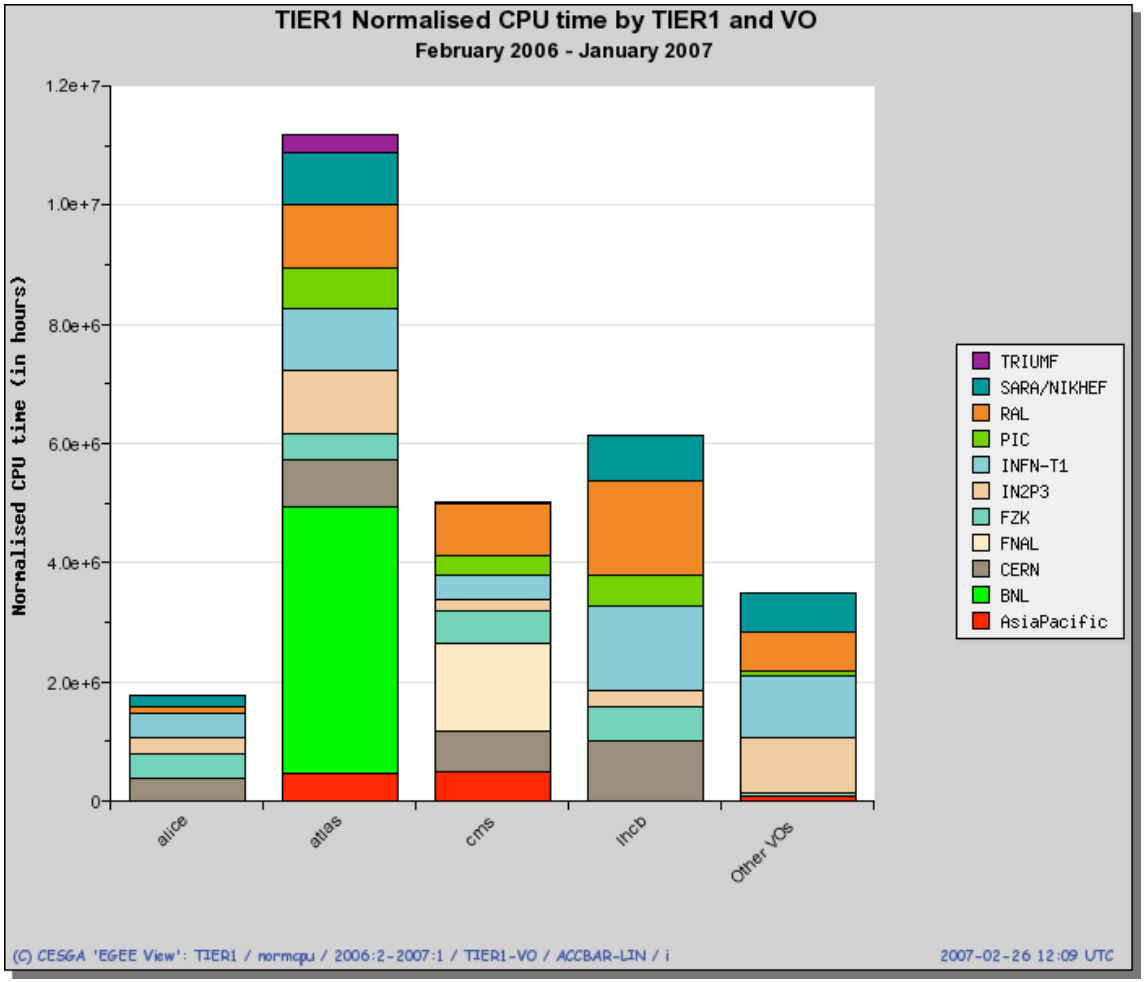
• l'architettura consiste di 4 componenti ⇒

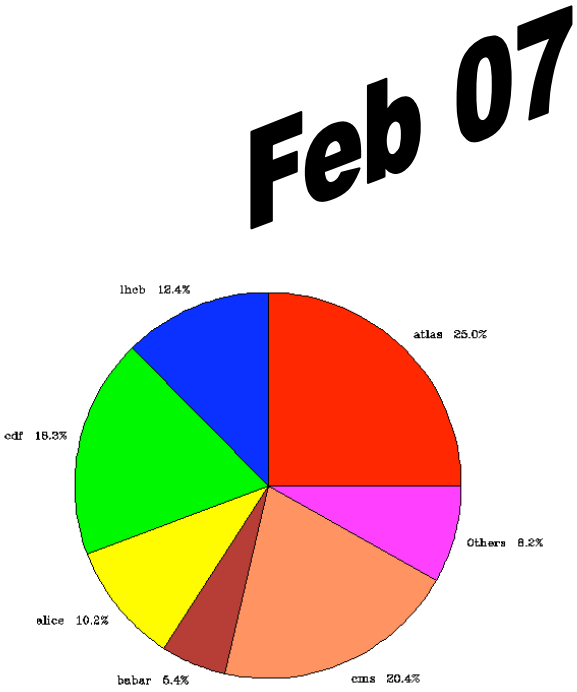


10 Mevt/week is the goal

- Il Production Database (DB ORACLE al CERN) in cui sono definiti i job di ATLAS raggruppati in:
  - Datasets : definiscono il contenuto fisico di una collezione di job
  - Tasks : identificano tutti i job dello stesso dataset (es: gen,simu, reco)
- Un Supervisor (Eowyn) che seleziona un job dal ProdDB e lo manda ai differenti Grid executors e ne verifica lo stato. Risottomette i job in caso di failure
- Gli Executors (moduli Python), uno per ogni Grid flavour, ricevono i job-definitions in formato XML, li convertono nel linguaggio della particolare Grid e li sottomettono individuando le risorse ottimali sulle infrastrutture di calcolo
- il Data Management System, DonQuijote (DQ2), si occupa della gestione e distribuzione dei dataset.

Utilizzo di cpu ai T1





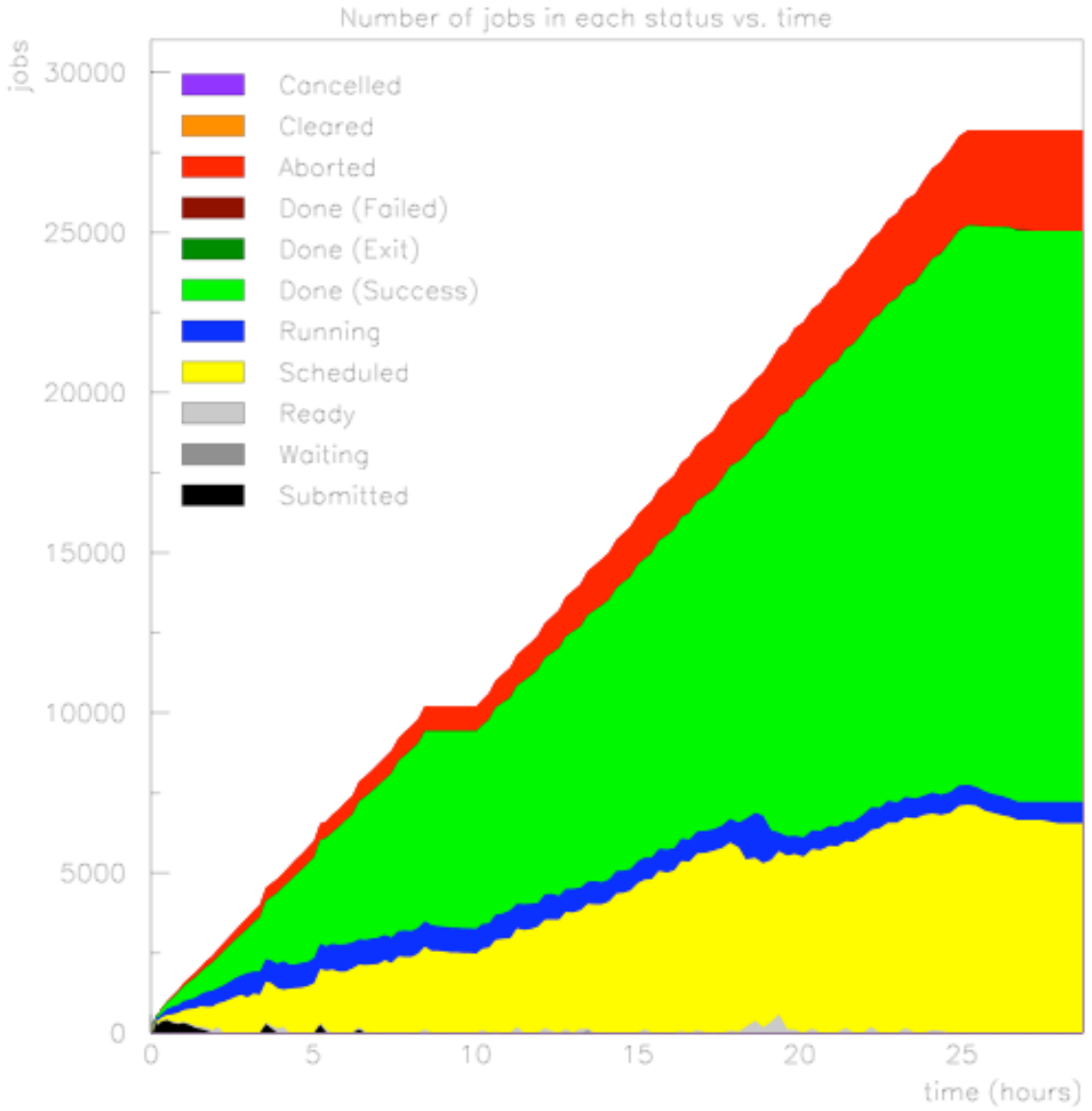
ATLAS - test WMS

- ✓ Problemi del WMS gLite risolti (vedi F. Giacomini):
  - Sottomissioni collezioni via DAGMan
  - Problemi nel L&B
  - Memori leak di Lexor
- ✓ ***Testbeds used to test new fixes and features:***
  - ***WMS devel09.cnaf.infn.it -> LB devel11.cnaf.infn.it***
  - ***WMS devel10.cnaf.infn.it -> LB devel12.cnaf.infn.it***
- ✓ The new tag will allow the WMS to handle bulk submission without using DAGMan
- ✓ A New LB for handling collections without DAGMan has to be used, in particular to retrieve the status and logging-info of the whole collection via the parent id
- ✓ Tests on the WMS dagless version using the two experimental WMSes at CNAF are on-going since March 2007
  They are both pointing to the CNAF repository goldrake.cnaf.infn.it
- ✓ Since then two patches to collect tags (and changes) were created
- ✓ A build of these patches is being produced at CERN

gLite 3.1

- ✓ Acceptance test on single WMS 3.1
- ✓ Reached more than 15 kjobs/day

(A. Sciaba', S. Campana, V. Miccio)



ATLAS - DM

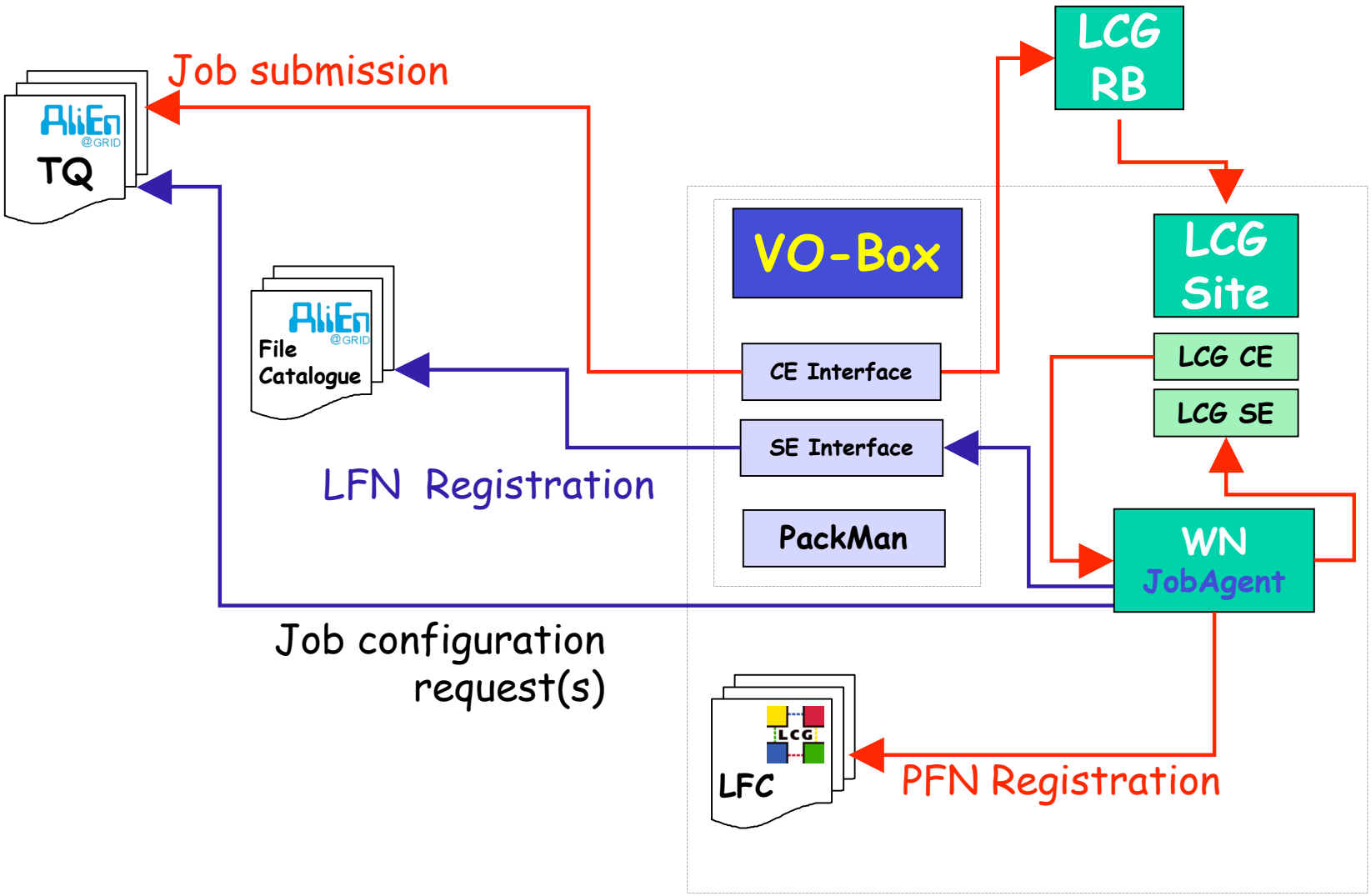
- ✓ Il Distributed Data Management (DDM) di ATLAS, Don Quixote (DQ2) provvede a:
  - distribuire tra i Tier i dati, reali e simulati, nei vari formati
  - catalogare i dataset
- ✓ Si basa sulle seguenti componenti:
  - FTS (File Transfer Service): per gestire i trasferimenti tra siti
  - SRM : interfaccia comune per gestire i vari SE nei siti
  - LFC (LCG File Catalog): Local Replica Catalog
- ✓ Comprende un sistema di sottoscrizione ai dataset automatizzato
- ✓ Obbiettivo di 1 GB/s aggregato T0 --> T1s

ATLAS - altro

- ✓ GANGA per l'analisi distribuita:
  - sottomette ad EGEE Resource Broker
  - ha interfaccia con DQ2
  - esegue automaticamente lo splitting del job, le sottomissioni, il merging degli output
- ✓ Problemi con CASTOR al CERN in via di miglioramento con nuova versione stager.
  - Configurazione di CASTOR al CNAF potrebbe essere critica
- ✓ Sistema di installazione automatica del software in tutte le sedi EGEE molto efficace (A. De Salvo)

  https://atlas-install.roma1.infn.it/atlas_install
- ✓ CSC towards fully commissioned computing under way

ALICE VO-Box concept



ALICE - integration with LCG

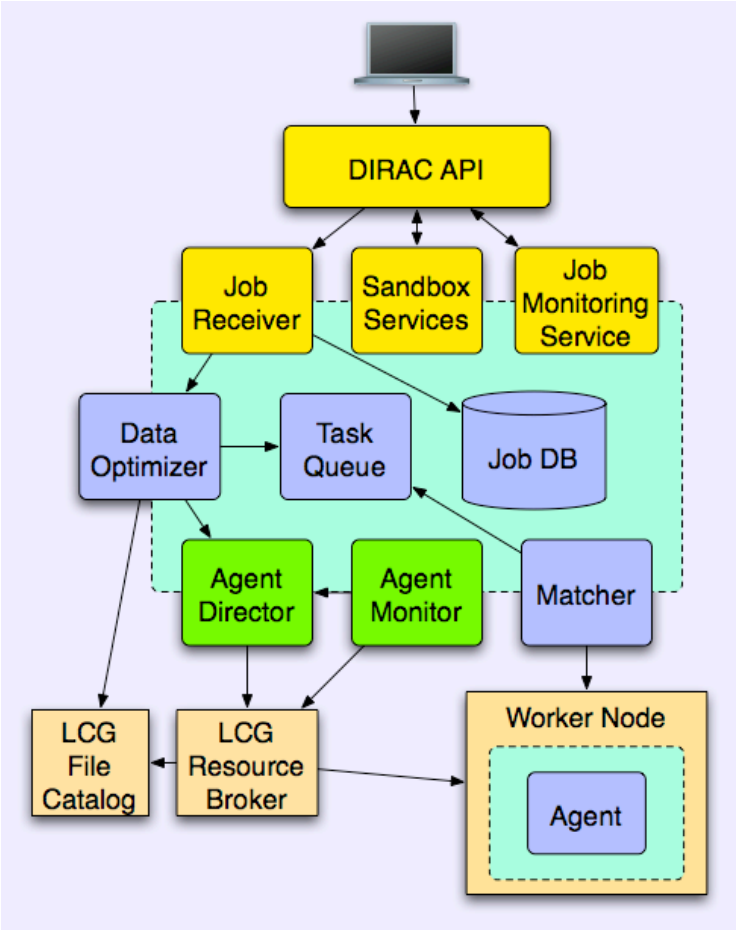
- ✓ Implement as much as possible thin interface services
  - To (stable) LCG standard services
  - Be "good citizens" of the Grid – no backdoors
  - Fairly satisfied with stability, performance and features
- ✓ Service interfaces on the VO-Box:
  - Job Submission (WMS clients)
  - Storage (SRM clients, xrootd redirector)
  - Data transfer (FTS clients)
  - Local Catalogue (LFC clients)
- ✓ Jobs are submitted to LCG Resource Brokers
  - Automatically generated JDL
  - Failover mechanism
  - Bulk submission in next release
- ✓ Job is taken from a central Task Queue
  - Priorities and shares are set in the Task Queue.
  - Takes also care of VO internal accounting
- ✓ The server of choice for data access is xrootd
  - Should work with CASTOR-2, DPM and dCache
- ✓ Also SRM needed
  - Used through FTS

MonALISA


ALICE Grid Monitoring with MonALISA
http://alimonitor.cern.ch:8889/show?page=index.html
CERN Users' pages
Grid
CICAP
Software
Tools
Physics
News
SNIP THIS!
MonALISA Repository
ALICE
MonALISA
MONitoring Agents using a Large Integrated Services Architecture
MonALISA Client
Click on the button below to start the Monalisa Client.
Client
ALICE Repository
Interactive Map
Job Information
SE Information
Services
Network Traffic
FTD Transfers
CAF Monitoring
Original
close all
ALICE Reports
Site Administration
Farm Colour Configuration
Job Status
Utilisation: Busy / Free Nodes
1039
Jobs
Jyvaskyla
UiB
SPbSU
PNPI
Aalborg
NBI
NIKHEF
Birmingham
RAL
SARA
Muenster
ITEP
Troitsk
RRC-KI
JINR
IHEP
SINP
KNU
BITP
GRIF_DAPNIA
IPNO
GSI
FZK
Prague
Cyfronet
Kosice
KFKI
Subatech
CERN_gLite
ALICE-CAF
CCIN2P3
CERNMAC
CERN-L
CERN
Clermont
Torino
Legnaro
CNAF
Bologna
NIHAM
ISS
Madrid
Trujillo
Bari
Cagliari
TriGrid_Catania
Catania
Athens
Done

LHCb and the Grid

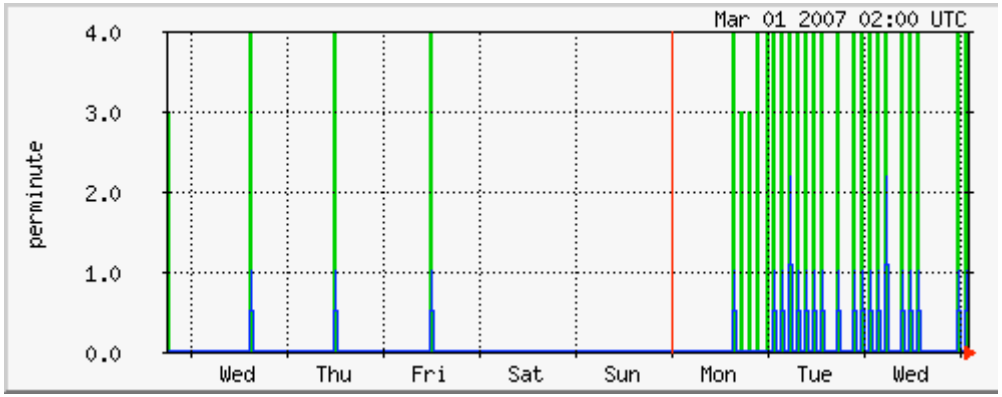
- Jobs submitted to the Grid thorugh the RB by the DIRAC Agent Director
- Production, Reconstruction and Pre-selection are scheduled activities, centrally managed
- For the analysis:
  - Use Pilot Agents and a centralized task queue and prioritization mechanism as in the previous cases..
  - Access the Grid directly.
  - Using the Ganga UI in both the cases with different backends.
- ✓ Separate disk & tape SRM endpoints already in place.
- ✓ Together with Castor2 need stoRM available in production at CNAF.
  - Tests of stoRM under way



Site Availability Monitor

- ✓ Collaboration with LCG and CERN/IT
- ✓ Run tests at each EGEE sites every 2 hours now
- ✓ Specific diagnostic for sites, target individual components
- ✓ Early detection of problems
- ✓ Site admins have access to status, history and logs via web
- ✓ Improved web interface being developed
- ✓ OSG sites being added

Squid traffic induced by SAM job
green = to WN
Blue = from CERN



Accounting

- ✓ Accounting scope is really accounting:
  - Verify for RRB times that sites delivered the pledged resources
- ✓ Hence it is global to WLCG
- ✓ Must be the same for all experiments
- ✓ It is discussed (routinely now) in WLCG-MB and GDB
- ✓ Users are classified in groups and roles
- ✓ GOC DataBase is the selected tool
- ✓ It can be filled in various ways
  - Apel (UK)
  - DGAS (INFN)

Policies

- ✓ DONE: Central job queue for Production
  - used by Atlas, LHCb, Alice
  - Coming in CMS as central repository of MC requests (pre-job level)
- ✓ NEED: Something for access granularity to disk.
  - Example: give to each user/group some "storage on the grid"
  - partly done at CERN on Castor (access from outside? FTS ?)
  - what about Tier2's ?
- ✓ NEED: Local mapping of Grid users to Unix users and use LRMS
  - on EGEE made complex by pool account with account recycling
- ✓ NEED: Every site implement fair share among users in a group
  - If my job is sent to the site where my group jobs runs better
  - Then fair share within this group is achieved globally
  - Maybe not perfectly, but good enough
  - Adopted as initial strategy by EGEE's Job Priority Working Group
- ✓ NEED: so called VOViews (better called GroupViews): gLite 3.1+

  This will allow the breaking down for groups and roles inside a VO

  Is that the final solution ? Is it flexible enough ?

  Does it provide enough granularity ?

Summary of middleware issues

- ✓ Data Transfer: operational issue
  - Throughput ~OK. Need much more work at fabric level: interoperability, reliability, disk→tape→disk
- ✓ Workload Management: middleware issue
  - Reliability ~OK.
  - Need throughput : scalability + performance
- ✓ Databases: deployment issue
  - Production level already
- ✓ Monitor: Never enough, need development and better error reporting from WMS
- ✓ Accounting: deployment issue
  - WLCG's GOCDB claims to have all desired functionalities
- ✓ Policies: deployment/operation (middleware issue long term)

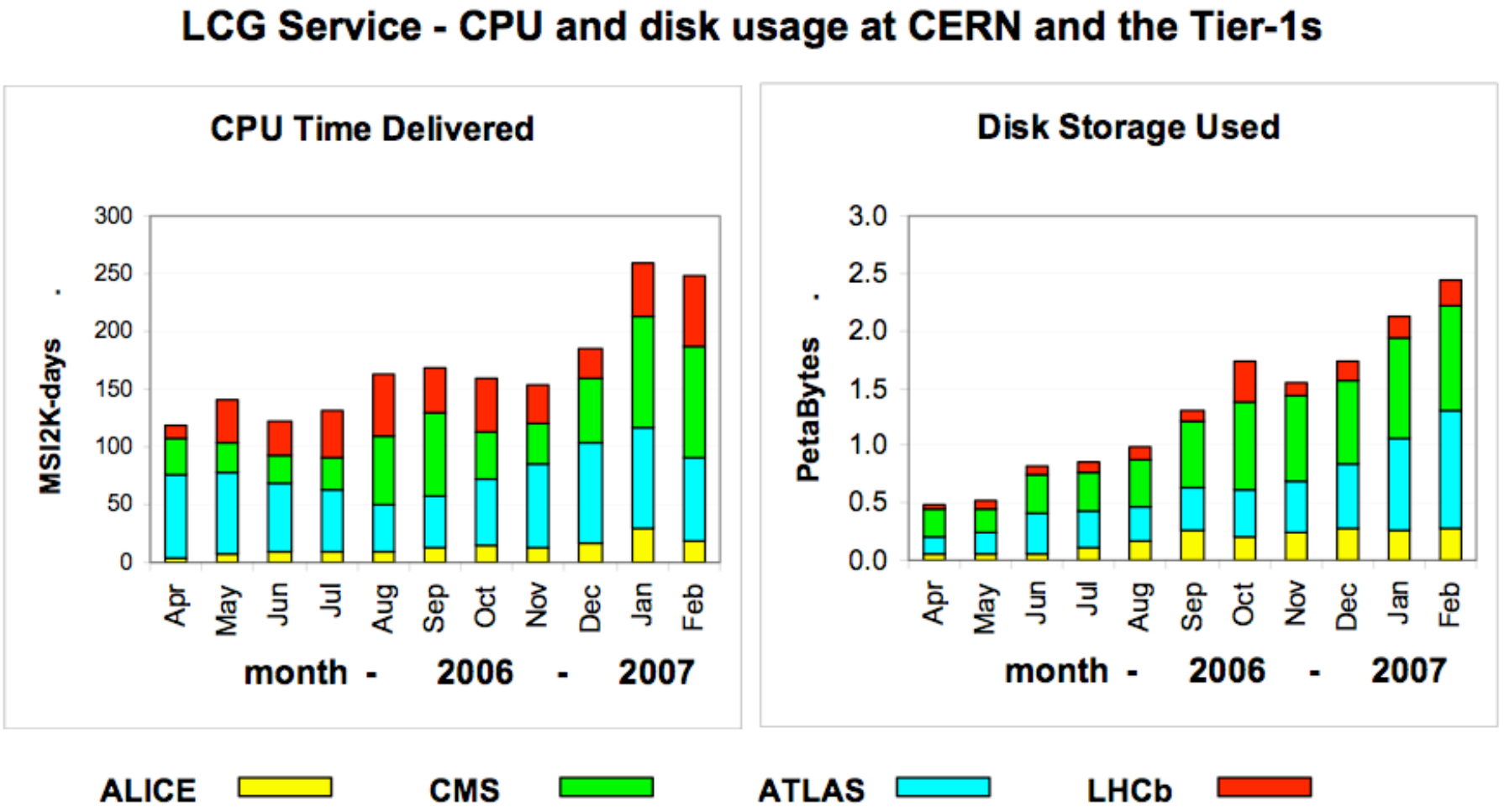
Essential services from the Grid

- ✓ WMS with:
  - Bulk submission and bulk matchmaking
  - High avaibility: one RB goes down without 100K jobs being lost
  - Scalability: transparent pool of hosts ? How reach 1M jobs/day ?
  - Global fair share
  - Local priorities
- ✓ Disk-only storage at T2 (DPM/dCache will do)
- ✓ Disk+tape at T1 (different solutions at different sites de-facto, Castor2 and STORM at CNAF)
  - Castor2 has to work and meet more demands then at CERN
- ✓ FTS that works like clockwork when sites are up.
- ✓ RGMA or similar is needed to collect information from monitors, L&B etc.
- ✓ Most urgent needs
  - WMS
  - Priorities via VOVIEWs
  - Tape related operations at Tier1's
  - Scale Tier2's at 2008 level

Storage Management

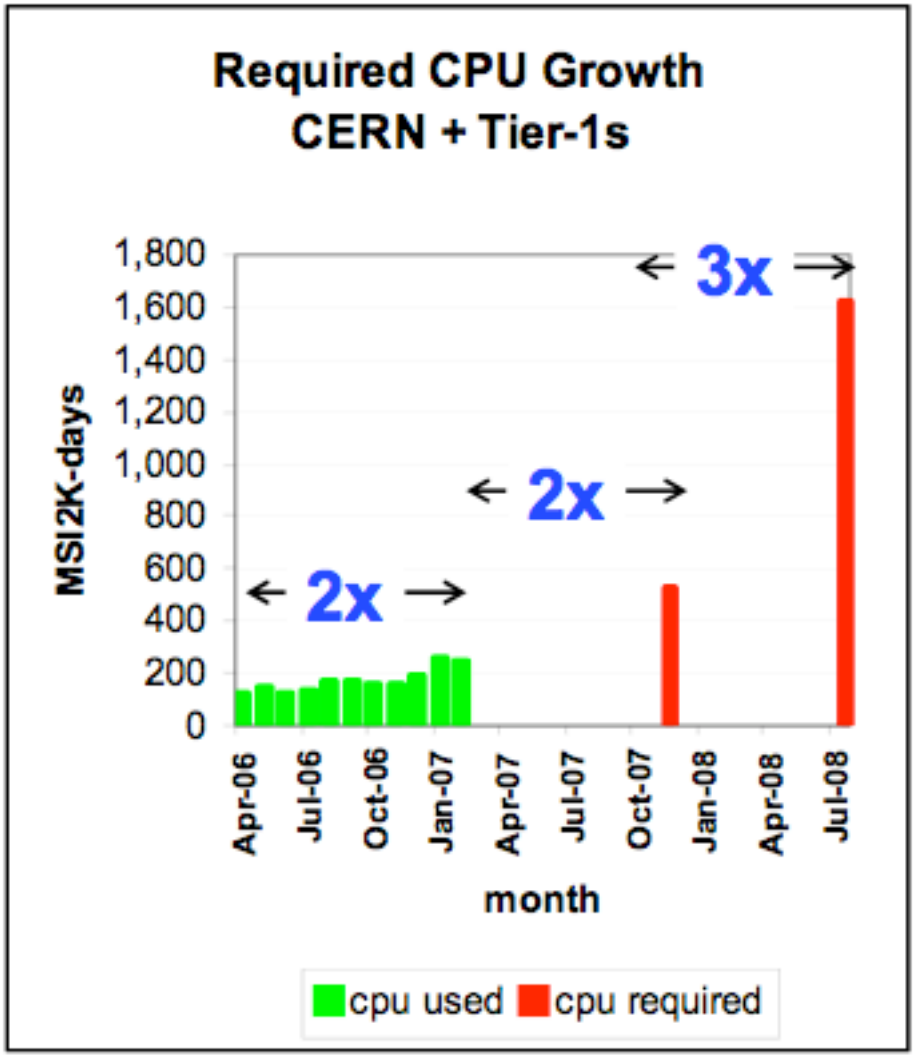
- ✓ New disk/tape standard interface coming: SRM v2
  - site interoperability
  - better control at Tier1 of disk/tape, pin/unpin
  - Plan for a smooth (!) transition
- ✓ Work with Tier1's to help them setup a disk/tape service that fits needs of LHC experiments
  - Tests on STORM
  - Commissioning of CASTOR
- ✓ Data serving at Tier2 sites
  - Validate/monitor capacity of sites to serve data for analysis
  - Provide requirement to dashboard and WM/DM tools
  - Dedicated program for dCache stress test
  - Integrate with other monitoring/accounting systems (local site monitoring, grid accounting …)

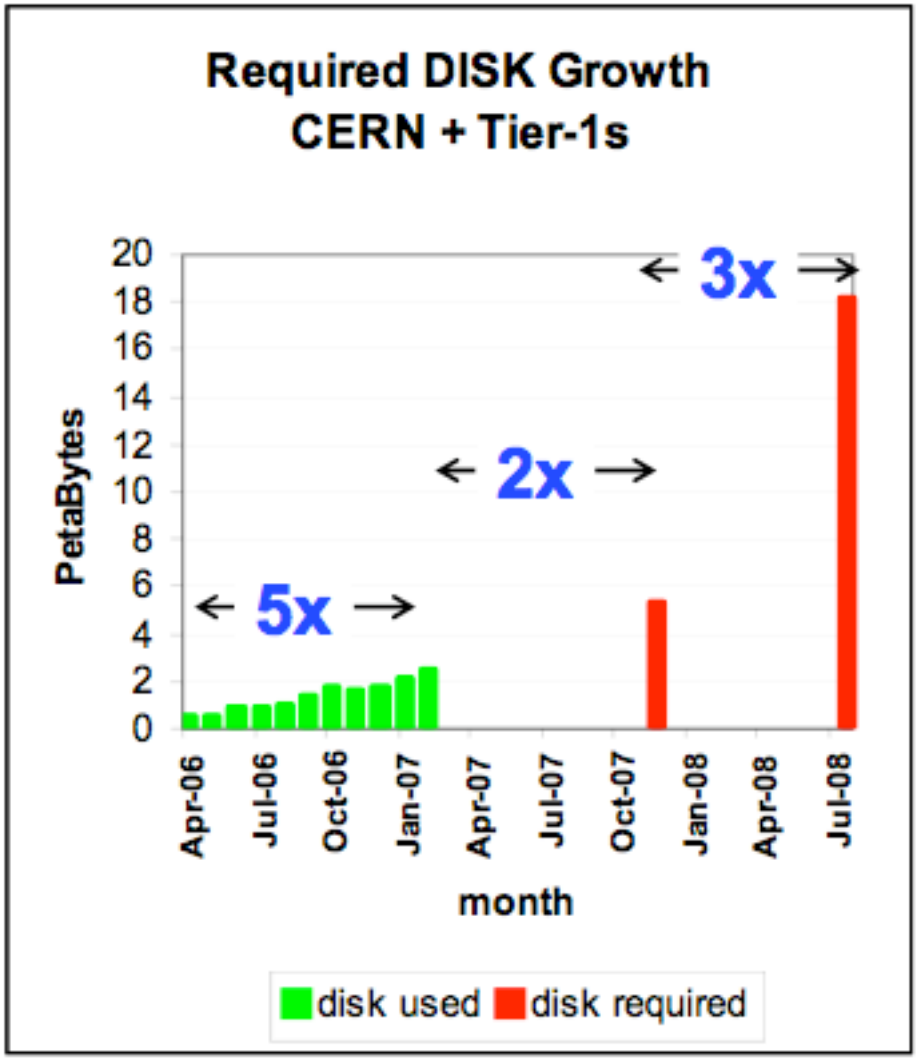
LCG - CRRB Apr07



LCG - CRRB Apr07

✓ # jobs needs 5 x from 1.8 Mjobs/month





CASTOR (CERN's mass storage system)

- Concern over stability and performance
  - Last year Castor achieved all of the Tier-0 and Data Distribution milestones – emphasising throughput
  - BUT – a significant software weakness was uncovered when handling large numbers of requests, mixed workloads (CASTOR is able to process 500K requests per day)
  - A prototype fix showed promising results in September – - but production ready code took longer to develop than expected
  - ATLAS testing in 2007 immediately generated overload on the ATLAS stager – making this issue critical
- Task Force set up to focus on this and other stability and performance issues in Castor
  - Working closely with ATLAS, bringing in additional expertise

  BUT --
  - May delay other Castor developments, including SRM 2.2 deployment

23-Apr-07

les.robertson@cern.ch

Status of other new, upgraded services

- **3D – Synchronised Distributed Oracle Database Service**
  - All but 3 sites now certified or awaiting certification by experiments.
- **FTS – File Transfer Service**
  - Version 2 in test at CERN, deployment to Tier-1s planned for May
  - Tier-1s are responsible for operation of the Tier-1/Tier-2 traffic – experiments organizing testing of all connections during first half of year
- **VOMS-based Job Scheduling Priorities**
  - Being deployed at Tier-1s
- **gLite Workload Management System**
  - In final stages of test – agreed performance and stability requirements achieved – entering certification
- **gLite Compute Element**
  - Problems encountered during tests – discussions on changing strategy
- **Port of EGEE Middleware to Scientific Linux 4**
  - Worker node components ready for distribution
  - Problems with other some of the other components

23-Apr-07

les.robertson@cern.ch

LCG - CRRB Apr07

Site Reliability - WLCG Tier-1s + CERN

The table shows *availability* for May through August, and *reliability* from September on BNL included in average from November
Reliability = Availability/Scheduled_Availability *(Scheduled_Availability=(1-Scheduled_Down_Time); tests are not run while scheduled down)*

	CERN-PROD	FZK-LCG2	IN2P3-CC	INFN-T1	RAL-LCG2	SARA-MATRIX	TRIUMF-LCG2	Taiwan-LCG2	USCMS-FNAL-WC1	PIC	BNL-LCG2	average reliabilities	8 best sites average			# sites ≥ target	# sites ≥90% target
													availability	reliability	(% target)		
Oct-2006	96%	54%	85%	85%	77%	74%	80%	67%	55%	84%	27%	76%	80%	81%	92%	*1*	*5*
Nov-2006	90%	85%	62%	94%	87%	77%	87%	95%	77%	79%	56%	81%	86%	87%	99%	*3*	*7*
Dec-2006	93%	63%	22%	77%	86%	82%	91%	95%	79%	90%	70%	77%	86%	87%	99%	*4*	*6*
Jan-2007	99%	85%	96%	75%	80%	93%	79%	96%	84%	86%	90%	87%	91%	91%	103%	*5*	*9*
Feb-2007	91%	90%	74%	93%	82%	83%	88%	97%	67%	86%	57%	57%	88%	89%	101%	*5*	*8*
Mar-2007	97%	75%	58%	76%	80%	47%	70%	95%	90%	96%	6%	85%	85%	85%	97%	*4*	*5*
average last three months	**96%**	83%	**77%**	82%	80%	**75%**	**79%**	**96%**	80%	**89%**	**54%**	81%	--	**88%**	100%	*3*	*8*

Piano di sviluppo Tier1

CNAF Plan March 2007																
		2006			2007			2008			2009			2010		
Experiment	%	CPU KSI2K	DISK TB-N	TAPE TB	CPU KSI2K	DISK TB-N	TAPE TB	CPU KSI2K	DISK TB-N	TAPE TB	CPU KSI2K	DISK TB-N	TAPE TB	CPU KSI2K	DISK TB-N	TAPE TB
ALICE	22%	154	*16*	77	286	110	143	1210	550	836	1870	880	1320	3520	1760	1870
ATLAS	32%	224	*40*	112	416	160	208	1760	800	1216	2720	1280	1920	5120	2560	2720
CMS	35%	245	*86*	123	455	175	228	1925	875	1330	2975	1400	2100	5600	2800	2975
LHCB	11%	77	*26*	39	143	55	72	605	275	418	935	440	660	1760	880	935
Total LHC TIER1		**700**	***168***	**350**	**1300**	**500**	**650**	**5500**	**2500**	**3800**	**8500**	**4000**	**6000**	**16000**	**8000**	**8500**
BaBar		585	*149*	0	680	200	0	1215	350	0	1215	350	0	1215	350	0
CDF		900	*66*	0	820	100	15	1161	170	15	1290	220	15	1420	270	15
LHCB TIER2		0	*0*	0	150	0	0	600	0	0	1200	350	0	1700	350	0
TOTALE GRUPPO I		**1485**	**214**	**0**	**1650**	**300**	**15**	**2976**	**520**	**15**	**3705**	**920**	**15**	**4335**	**970**	**15**
AMS2		32	*2*	16	25	5	16	32	5	24	180	16	128	180	16	128
ARGO		22	*12*	28	150	70	186	188	112	366	188	129	546	188	129	546
GLAST					5	10	0	5	10	10	5	10	20	5	10	20
MAGIC			*1*		20	5	4	20	4	8	20	4	12	20	4	12
PAMELA			*4*		20	10	16	25	10	32	25	10	48	25	10	48
Virgo		10	*25*	75	180	90	130	250	150	200	500	220	250	500	220	250
TOTALE GRUPPO II		**64**	***43***	**119**	**400**	**190**	**352**	**520**	**291**	**640**	**918**	**389**	**1004**	**918**	**389**	**1004**
All experiments		**2249**	***426***	**469**	**3350**	**990**	**1017**	**8996**	**3311**	**4455**	**13123**	**5309**	**7019**	**21253**	**9359**	**9519**
All w/ overlap factor		**1874**	***387***	**469**	**2792**	**900**	**1017**	**7497**	**3010**	**4455**	**10936**	**4827**	**7019**	**17711**	**8509**	**9519**
CNAF TOTAL (PLAN)		**1874**	***387***	**469**	**3000**	**1000**	**1000**	**7497**	**3010**	**4455**	**10936**	**4827**	**7019**	**17711**	**8509**	**9519**
CNAF ACTUAL		***1570***	***400***	***510***												
Relative Contingency						0%			30%			40%			50%	
Absolute contingency					0	0	0	2249	903	1337	4374	1931	2808	8855	4254	4760
Zoccolo duro (TOTAL-CONTINGENCY)					**3000**	**1000**	**1000**	**5248**	**2107**	**3119**	**6562**	**2896**	**4211**	**8855**	**4254**	**4760**
INFN T1 P2P 2005		**1800**	**850**	**850**	**2400**	**1200**	**1000**	**5500**	**2500**	**2100**	**8000**	**4000**	**4100**	**11500**	**5800**	**6000**
INFN T1 P2P 2007		**-**	**-**	**-**	**1300**	**500**	**650**	**4500**	**2000**	**2100**	**6500**	**3200**	**3300**	**10000**	**5000**	**5000**

Bilancio risorse

Summary of Regional Centre Capacities
2/1/2007

Tier-1 Planning for 2008		ALICE	ATLAS	CMS	LHCb	SUM 2008
CPU - MSI2K	Offered	6.7	24.0	12.0	5.0	47.7
	TDR Requirements	10.2	18.1	12.4	1.8	42.5
	Balance	-34%	32%	-3%	182%	12%
Disk - PBytes	Offered	2.8	13.1	5.7	2.5	24.1
	TDR Requirements	5.2	9.9	5.6	1.0	21.7
	Balance	-47%	32%	3%	140%	10%
Tape - PBytes	Offered	3.1	9.0	9.6	1.9	23.6
	TDR Requirements	7.0	7.7	13.1	0.9	28.7
	Balance	-56%	17%	-27%	123%	-18%

Includes current planning for all Tier-1 centres

Tier-2 Planning for 2008		ALICE	ATLAS	CMS	LHCb	SUM 2008
CPU - MSI2K	Offered	6.0	19.5	20.0	3.7	49.2
	TDR Requirements	9.6	17.5	15.2	4.6	46.9
	Balance	-37%	12%	31%	-18%	5%
Disk - PBytes	Offered	1.4	5.9	5.1	0.7	13.1
	TDR Requirements	2.5	7.7	4.2	n/a	14.4
	Balance	-41%	-23%	21%	n/a	-9%
# Tier-2 federations - included(expected)		16 (18)	25 (30)	28 (31)	11 (12)	43 (50)

Tier-0 Planning for 2008		ALICE	ATLAS	CMS	LHCb	SUM 2008
CPU - MSI2K	Offered	3.3	3.7	3.9	0.4	11.3
	TDR Requirements	3.3	3.7	3.9	0.4	11.3
	Balance	0%	0%	0%	0%	0%
Disk - PBytes	Offered	0.1	0.2	0.3	0.3	0.8
	TDR Requirements	0.1	0.2	0.3	0.3	0.8
	Balance	0%	0%	0%	0%	0%
Tape - PBytes	Offered	1.2	2.4	3.6	0.6	7.8
	TDR Requirements	1.2	2.4	3.6	0.6	7.8
	Balance	0%	0%	0%	0%	0%

CAF Planning for 2008		ALICE	ATLAS	CMS	LHCb	SUM 2008
CPU - MSI2K	Offered	3.9	2.1	3.8	0.0	9.8
	TDR Requirements	3.9	2.1	3.8	0.0	9.8
	Balance	0%	0%	0%	0%	0%
Disk - PBytes	Offered	1.0	1.0	1.3	0.1	3.3
	TDR Requirements	1.0	1.0	1.3	0.1	3.3
	Balance	0%	0%	0%	0%	0%
Tape - PBytes	Offered	1.2	0.4	1.5	0.0	3.0
	TDR Requirements	1.2	0.4	1.5	0.0	3.0
	Balance	0%	0%	0%	0%	0%

Conclusions

- ✓ Integrate data and workload management systems with EGEE middleware
  - Prepare for analysis, production, service/data challenges including regional centres
- ✓ VO users are not Grid experts --> need simple and stable middleware
- ✓ Scaling at production level, while keeping high efficiency is the critical point
  - Continuous effort
  - To be monitored in detail
  - To be shared between experiments and developers

Acknowledgements

- ✓ S.Belforte, D.Bonacorsi, S.Campana, F. Farina, S.Lacaprara, E.Miccio, U.Marconi, M.Masera, L.Perini, S.Resconi, A.Sciaba', D.Spiga