



ATLAS

2008 Report and 2009/2010 Plans

reprocessing
simulation production
planning

WLCG Workshop in Prague

March 21-22

Slides from LHCC review
and ADC internal review
and some more ...

Kors Bos
CERN & NIKHEF

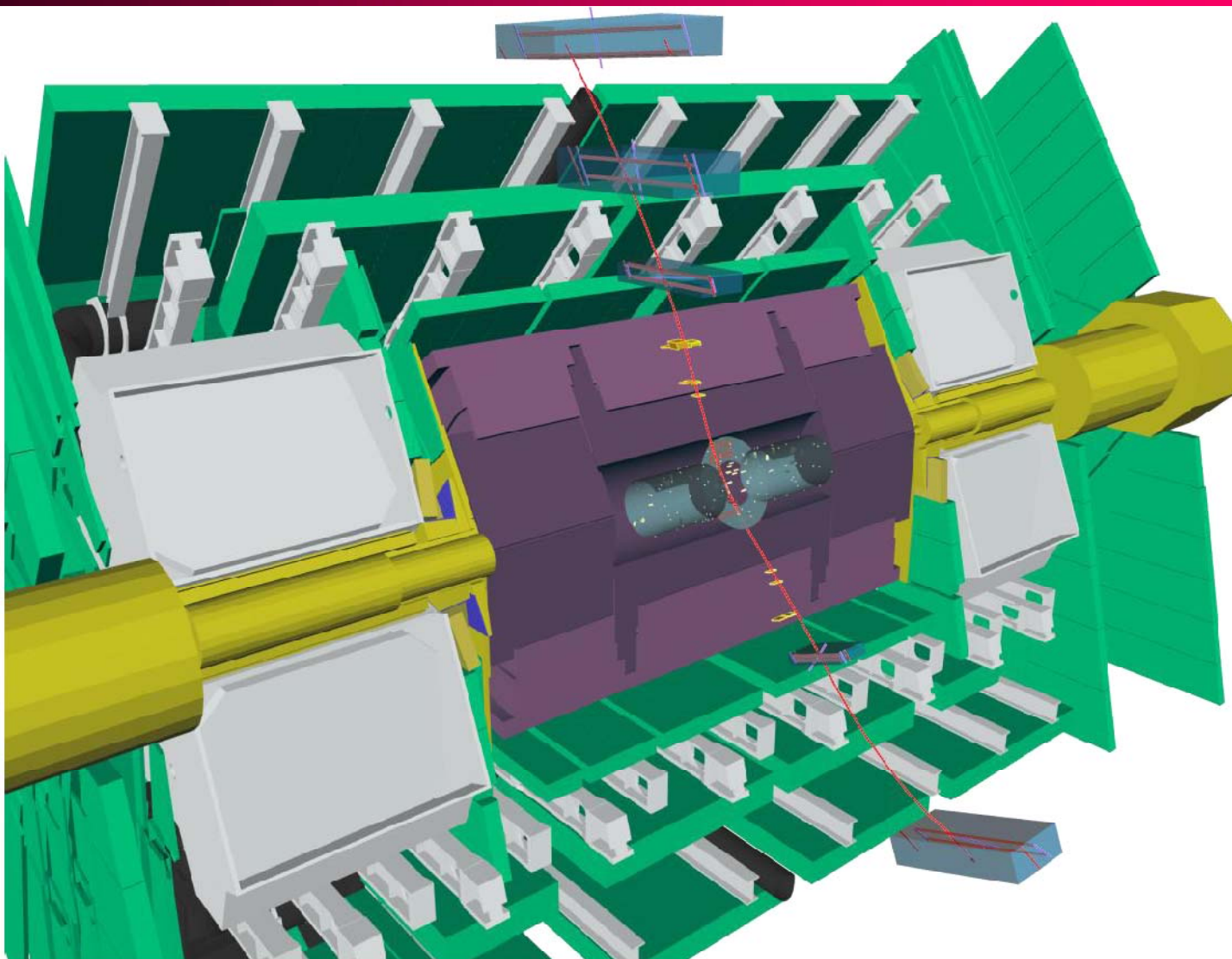
Slide 1

KB1

Kors Bos, 3/20/2009



2008 cosmics data re-processing



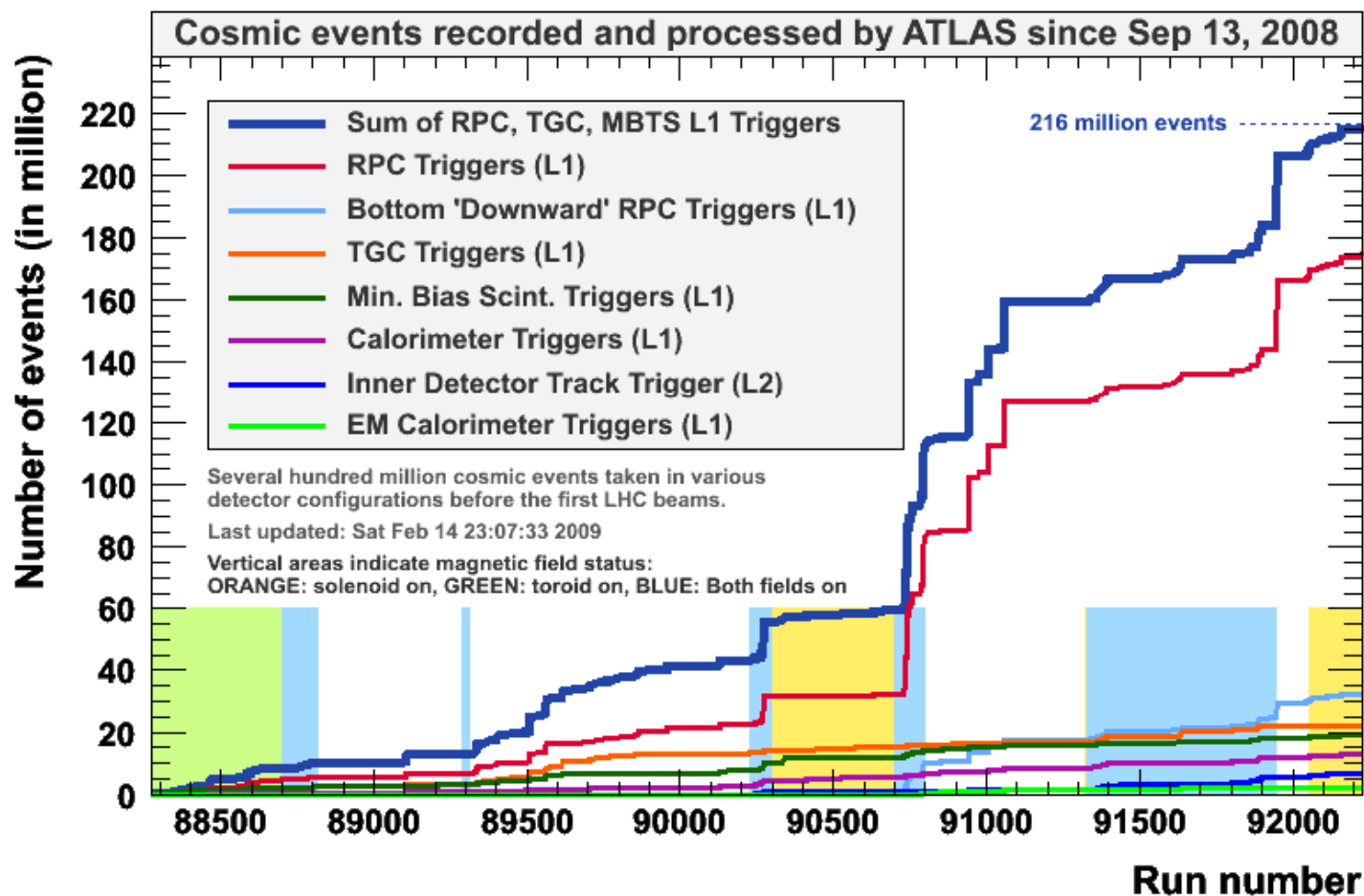


Cosmics data

284 million events

513 TB input data: 127 runs, 11 streams, 771 datasets, 330 559 files

117 TB output data: 8321 containers, 2 164 416 files



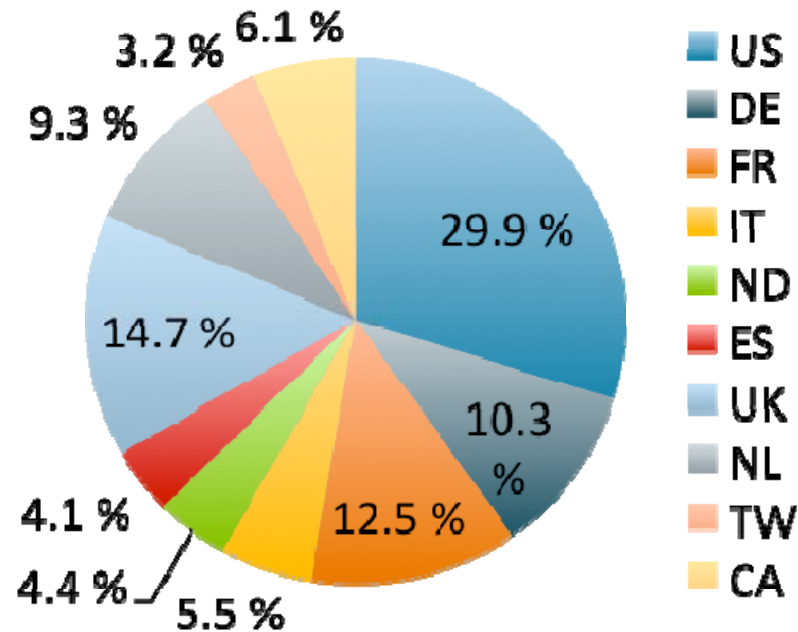


(Re-) Processing

- Data:
 - Cosmics data taken in August-November
 - Few runs of single-beam events the week of Sept 10th
- All RAW data were processed at Tier-0 → ESD, AOD, DPD, TAG
- Distributed according to the Computing Model
 - But extra RAW copy was kept on disk also in T1's
- Good data selected (about 35%)
- Software validated on Dec. 22
- Sites validated
 - Taipei and Germany failed
- Good data re-processed
 - Mainly (90%) using the disk copy
 - Using conditions data files i.s.o. database
 - Many Tier-1's were working on a best effort basis



Input data distribution



- Data from Taipei and Karlsruhe were re-distributed
- Later also a fraction of the SARA data and PIC data



Job efficiency

Tier-1:	CERN	CA	ES	FR	IT	ND	NL	UK	US	Sum
Total jobs	26348	20707	364	48288	13619	12561	23472	54360	128764	329609
Done jobs	26015	20150	364	46937	13018	12281	23167	51344	124667	317943
Fraction [%]	94.7	97.3	100	97.2	95.6	97.8	98.7	94.5	96.8	96.5
Aborted jobs	1459	557	0	1351	601	280	305	3016	4097	11666
Fraction [%]	5.3	2.7	0	2.8	4.4	2.2	1.3	5.5	3.2	3.5

Re-tries	CERN	CA	ES	FR	IT	ND	NL	UK	US	
	1.16	1.02	1.18	1.11	1.83	1.14	2.85	2.31	1.39	

Re-assigned jobs:

- ES 15932
- NL 6033
- IT, CA, UK 6409



Second try: tomorrow

- Patches to the software
 - Increased the memory footprint with 100 MB ☹
 - Not essentially different (better calibration)
- Run from tape now
 - Bulk pre-staging tested except in Lyon
- Karlsruhe in now, but Taipei's share done @CERN
- Input data
 - 370MB DB file - same for every job
 - 16MB DB file - 138 different files, 1 per run
 - 2GB RAW data - all different
- Pcache used in some places: RAL, NIKHEF
- Some Tier-2's participating
- Now we will all be watching



Memory

- Many jobs failed because of memory usage (re-tries)
- With newer release situation got worse
 - Even more jobs failed in site validation
 - Needed to switch off some functionality (monitoring)
 - This is not acceptable for the real data
- Memory usage not quite understood
 - Only for short period in initialization and first event
 - Less later on
- Will run at RAL with 3 GB virtual memory limit
 - Carefully watch CPU/Wall time for swapping
 - Expect little effect
- If so we may ask other sites to do the same



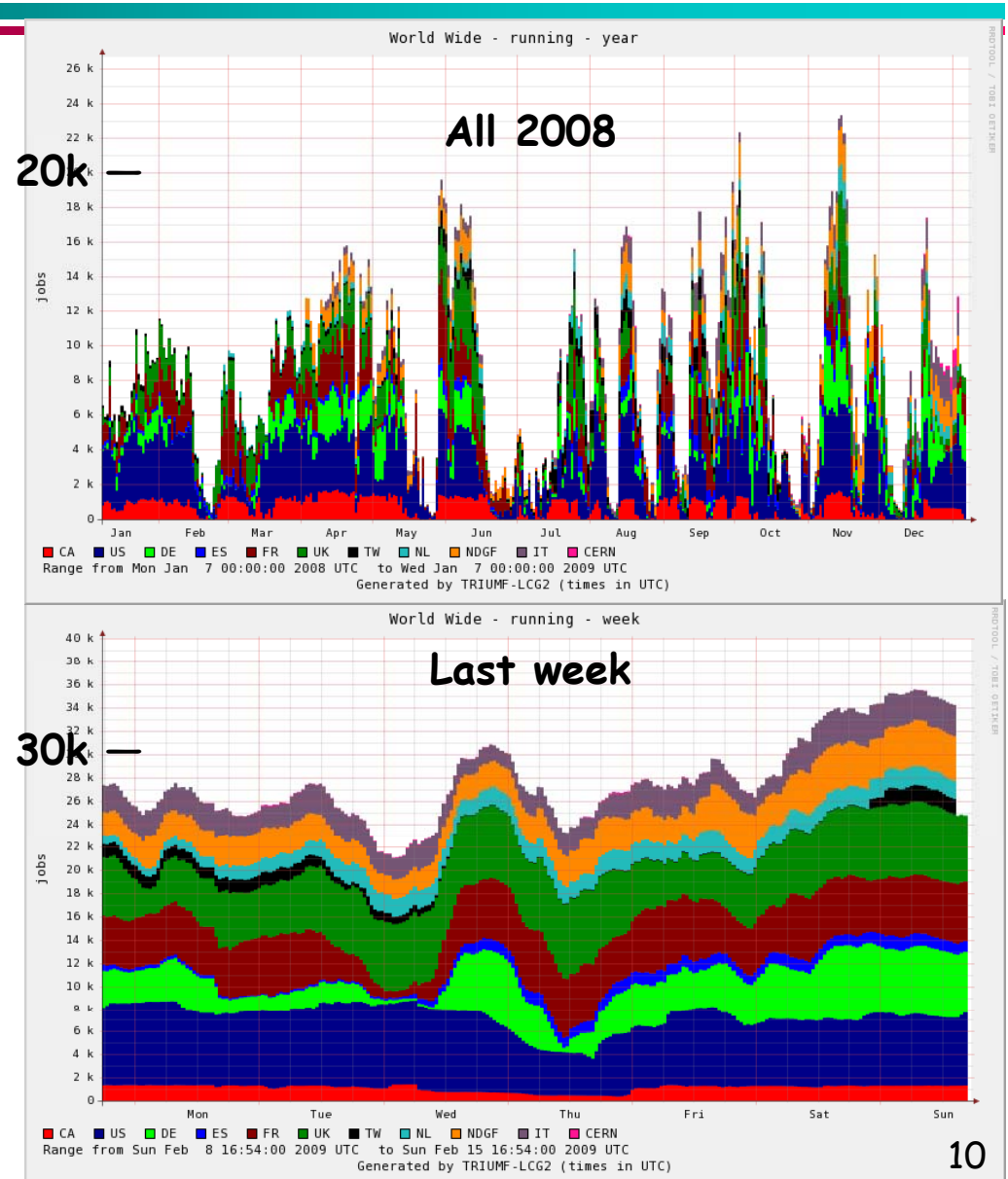
Future plans

- We may try again during the summer
- Needs much more automation
 - Not move data around by hand
- Better monitoring
 - Shifters should be involved and report errors
- Taipei must participate & Lyon must pre-stage automatically
- Physically remove input data copy from disk
- Output ESD data must be archived on tape
- Reduce #retries, reduce job failures, process 100% of all events
- Test together with writing (FT/HITS/ESD) data to tape
- Hopefully other experiments will be using tapes as well



Simulation Production

- Simulation production in 2008 used most of the available CPU capacity
 - But not all the time
- Production will continue, using now Geant4.9
 - The Geant4 step now takes ~1000 kSI2k-s/event
- Fast simulation samples are produced in addition, for large statistics studies
 - AtIfast-II takes on average ~100 kSI2k-s/event





Experience in 2008 (1)

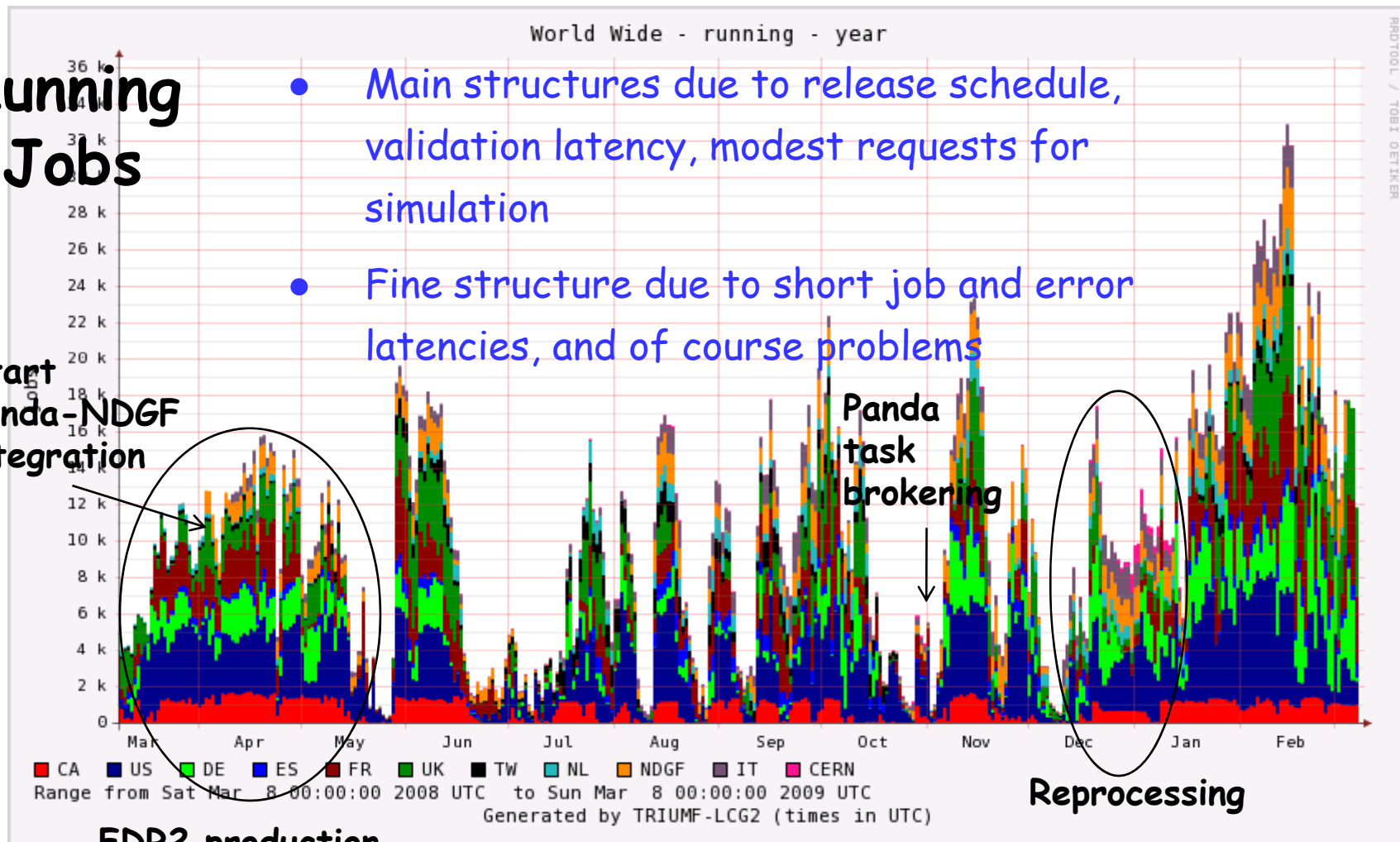
Running Jobs

- Main structures due to release schedule, validation latency, modest requests for simulation
- Fine structure due to short job and error latencies, and of course problems

Start Panda-NDGF integration

Panda task brokering

Reprocessing



FDR2 production



Experience in 2008 (2)

Jan-Dec 2008						
<i>grid</i>	<i>success</i>	<i>failure</i>	<i>success (walltime)</i>	<i>failure (walltime)</i>	<i>efficiency</i>	<i>efficiency (walltime)</i>
LCG	7173536	2519024	133475049397	32325414698	74%	80.5%
OSG	3493728	841093	65063597186	13219670831	80.6%	83.1%
Nordugrid	1147429	93355	19796763840	631522980	92.5%	96.9%
None	316444	333398	6170728191	1404697337	48.7%	81.5%
<i>total</i>	<i>12131137</i>	<i>3786870</i>	<i>2.24506138614e+11</i>	<i>4.7581305846e+10</i>	<i>76.2%</i>	<i>82.5%</i>

- CPU's are under-utilized on average
 - Capacity ~60k jobs/day, 14-20k CPU's (25-30k peak)
 - Utilized: 33k jobs/day, 6.9k CPU's (successful jobs)
 - Nearly 20% of walltime used for failed jobs
- 25% of jobs and 20% of walltime are used for failed jobs (mixture of grid and athena)

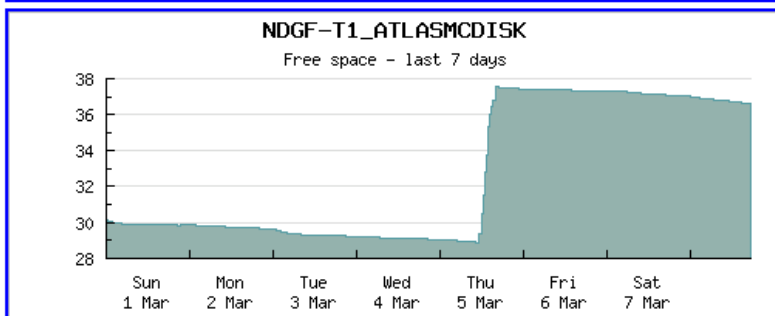
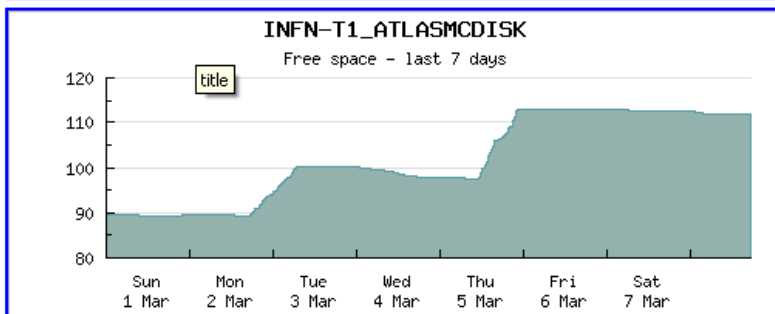
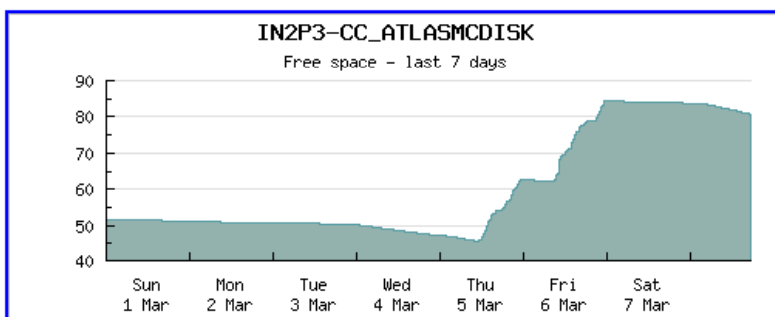


Experience in 2008 (3)

- Successes
 - Saturation of CPU resources (for long simulation jobs only)
 - Production system can produce 100 TB/week of MC
 - Efficiencies over 90% (for short periods of time)
 - Task brokering by software introduced Q3/Q4 2008
 - Complaints from physics groups about tasks with lack of progress dramatically reduced
- Issues
 - Insufficient space for the MC results Production System is capable of producing
 - More or less always at least one Tier-1 unavailable
 - Only a few weeks a year with all 10 Tier-1's + clouds working well
 - Huge rate of elog and GGUS tickets, hard for anybody to follow the many issues (related to previous point)



Space for MC production

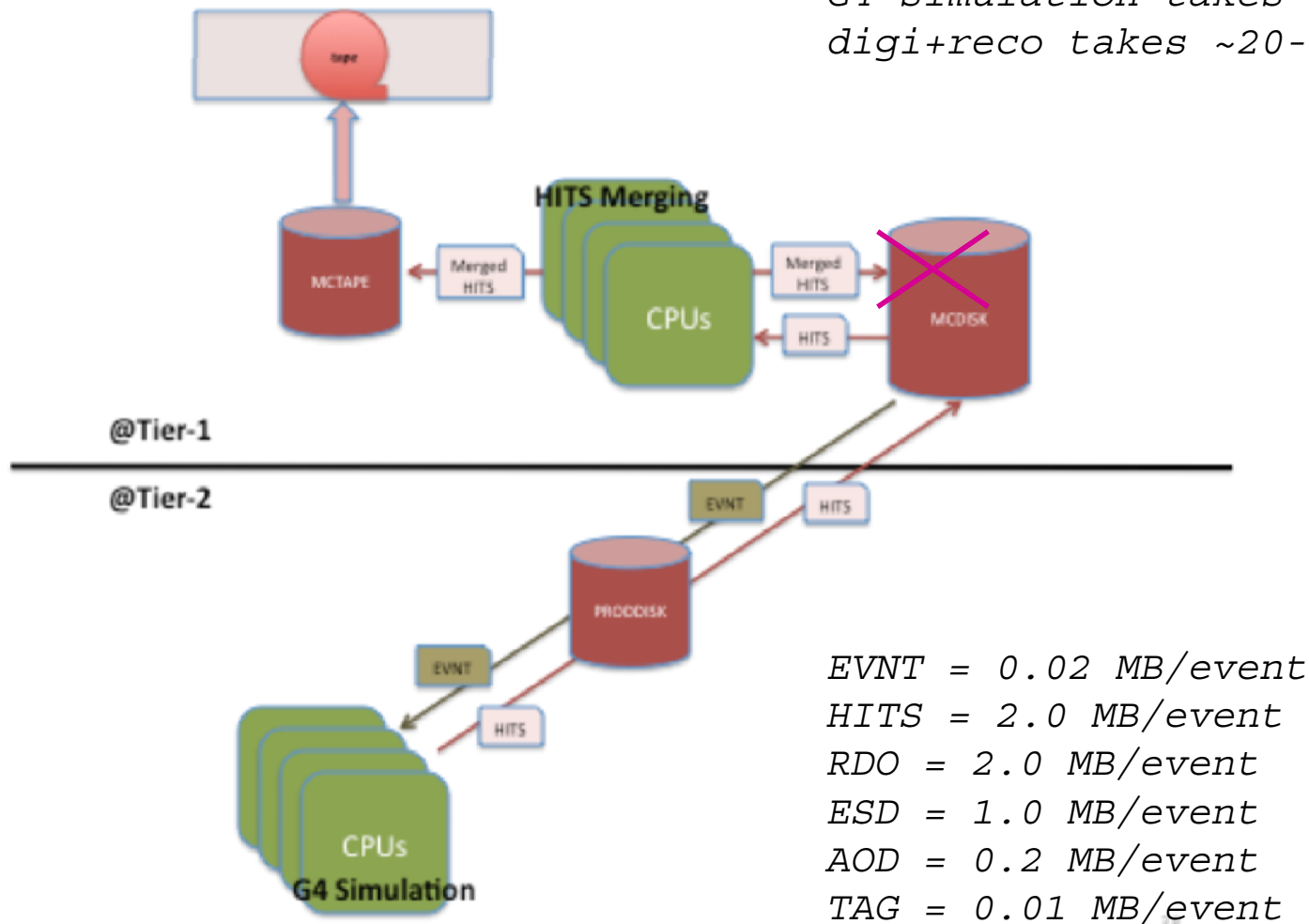


- We are now deleting old MC results to make room for new ones



HITS production

G4 simulation takes ~1000 sec/evt
digi+reco takes ~20-40 sec/evt





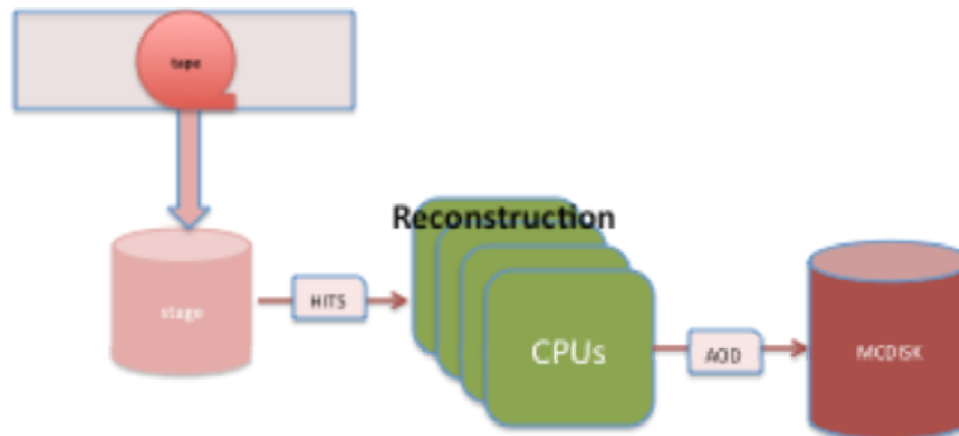
HITS merging

- Geant4 job in T2: 50 events/job → takes 14 hrs
- Output HITS file: $50 * 2\text{MB} = 100\text{ MB}$ uploaded from T2 to T1
- Merge in T1: 20 HITS files → 1 JumboHITS file of 1000 events (2 GB)
- JumboHITS file written to MCTAPE
- Small HITS files deleted from disk
- No HITS left on MCDISK
- That is a big difference as that is about 50% of all data on that disk

- This now works
- Digi + Reco can be ran from tape and produces the same output
- needs same pre-staging as for re-processing
- Also logfiles are merged and written to tape



digireco from tape



@Tier-1

@Tier-2

JumboHITS file with 1000 events staged from tape

Reconstruction takes 20 - 40 seconds/event

digireco job takes 6 - 12 hours/file

Output AOD file: $1000 * 0.2 \text{ MB} = 200 \text{ MB}$ still too small!



AOD merging

- At production peak we replicated 1.5M AOD files/day
- This is too much, no room left soon for other data transfers
- Must merge 10 AOD files into JumboAOD files
- JumboAOD file has 10,000 events and is 2 GB
- This would bring the number of file replications down
- We could even go to 20:1
- But we need to see with analysis groups

- (Almost) all components exist
- But has not been done yet
- Need to be tried a.s.a.p.

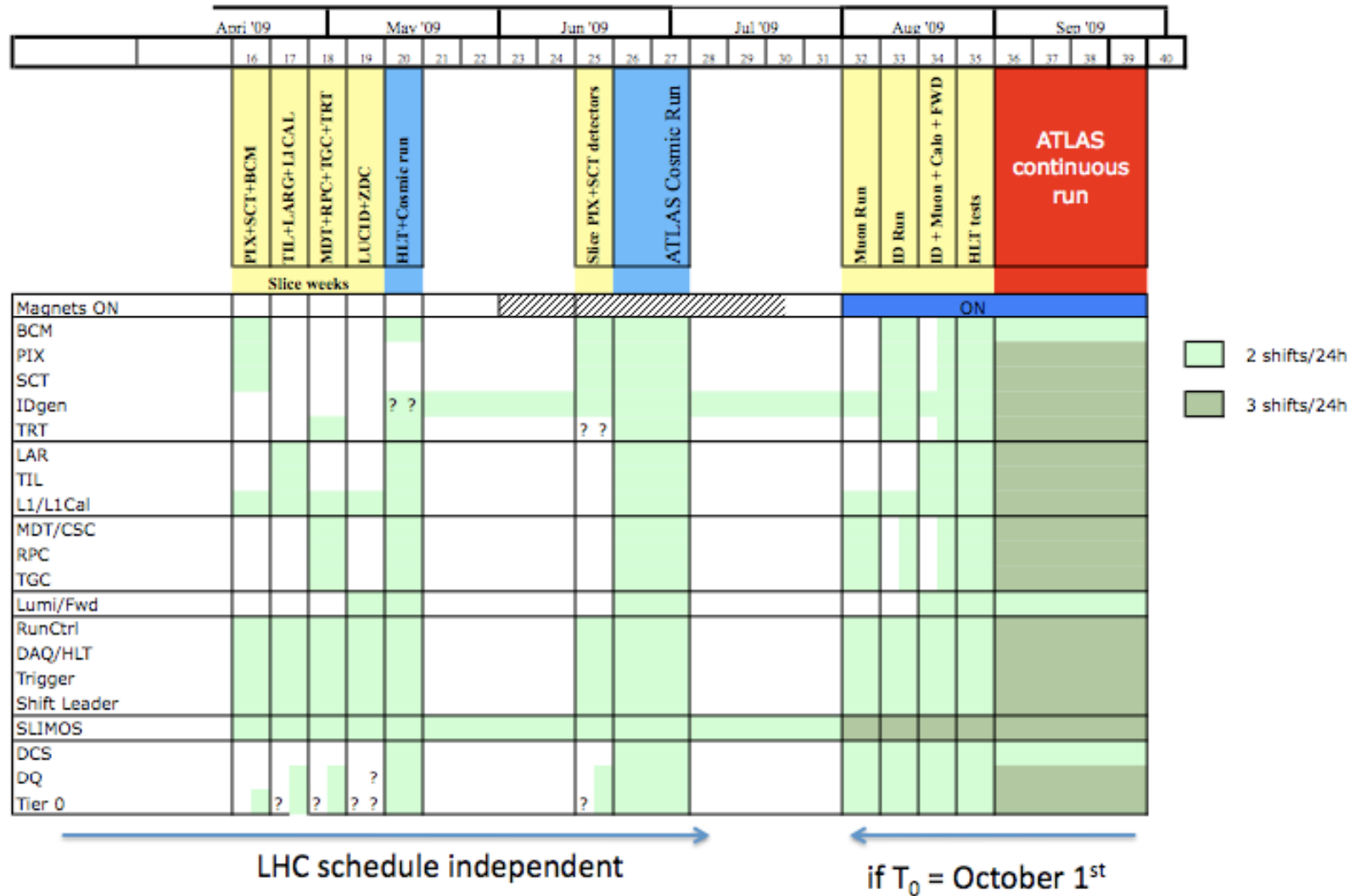


Future plans

- MC09 will soon start (wait for final release, final geometry)
- Using new Geant4 release
- Produce samples for 10 TeV running
- Produce samples for cosmic ray running
- Improvements on
 - More atomization (blacklisting, re-scheduling, brokering)
 - Reduce latencies due to transfer time-outs
 - Better matchmaking (not everywhere 2 days queues with 2+ GB)
 - ..
- Serve all clouds from Panda dbase on Oracle@CERN
- Panda for NDGF using ControlTower



Detector Commissioning Schedule





Backup slides

- I only show 20 slides now
- But have more slides if appropriate for during the workshop



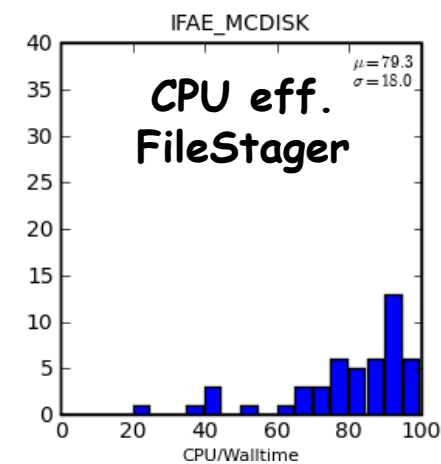
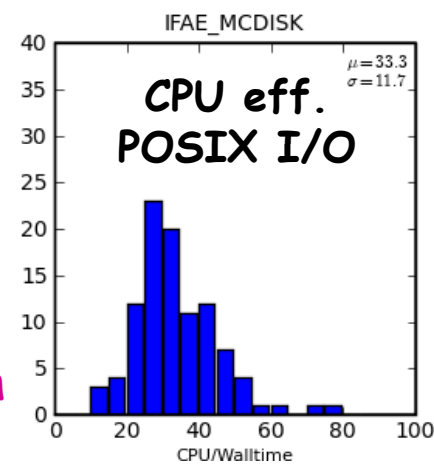
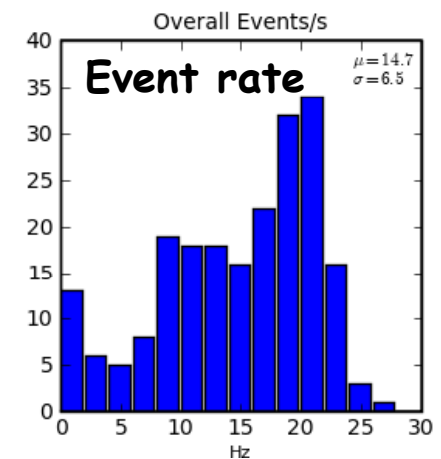
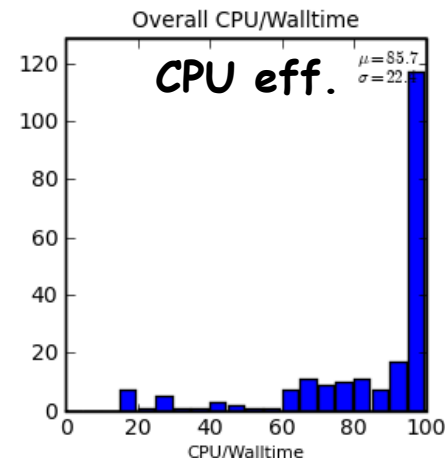
Stress-tests of distributed analysis (1)

- HammerCloud (HC) is a Distributed Analysis stress testing system
 - It's a database with command-line and web front-ends used for scheduling, running, and reporting on Ganga-based DA stress tests.
 - One specifies where, what, when, and how many jobs to run; HC runs the test and reports the results.
 - At present it can run tests on EGEE/WMS and NDGF/ARC sites.
 - Panda support needs work (in progress).
- Presently we submit large numbers of input data intensive jobs.
 - Focusing on finding site limitations, e.g. heavy loads on storage and network resources.
 - But looking out for grid limitations, e.g. broker, WMS, client.
- Specifically, HC runs an AOD muon analysis everywhere right now
 - Each HC run processes an entire dataset (200 files)
- Soon analyses of real cosmics datasets will be added (with database access)



Stress-tests of distributed analysis (2)

- The initial target for these tests was for sites to achieve >80% success efficiency and >15 Hz event rate
 - Example of a good site:
- In several cases we found out that default access methods could be changed or the local hardware setup had limitations
 - Example of improvements:
- Every site has a different hardware setup
 - Every site needs a separate investigation and optimisation
 - Adding CPUs and disks is not enough if the internal network is not correspondingly upgraded at the same time





Evolution of the computing model (1)

- The computing model is fundamentally unchanged since the Computing TDR (2005)
- Some model parameters have been adjusted all along, based on our experience with real data (cosmics) and intensive simulation campaigns during the last few years, and to reflect reality:
 - ESD and AOD sizes doubled to 1 MB and 0.2 MB/event to include more information needed to study detector performance
 - We try to avoid to go back to RAW (from ESD) or ESD (from AOD) when not necessary
 - Sizes will most probably decrease again as we gain experience with real data
 - Full simulation processing times increased from 100 to 1000 kSI2k-s/ev to allow for a very detailed detector geometry and much better physics models in Geant4
 - This level of detail is needed for some physics processes
 - We have developed a complete suite of full, medium and fast simulations that we will use and mix according to the physics needs and the availability of CPU capacity
 - Tuning the Geant4 parameters on real data will help us to optimise CPU usage
 - For reconstruction, CPU time is in the correct ballpark (15 kSI2k-s/event) at low luminosity, but more work is needed for higher luminosities ($>2 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$)
 - Targeted effort for memory reduction is now starting to pay off (improvements in the forthcoming s/w release)



Evolution of the computing model (2)

- The DPDs (Derived Physics Data), mentioned in the C-TDR, have been in the meantime defined and implemented in the same data formats as ESD and AOD:
 - Commissioning DPDs are subsets (skims) of ESDs for particular detector studies
 - Physics DPDs are subsets (skim/slim/thin) of AODs for physics studies
 - DPDs are in principle distributed like AODs (a full set per Tier-1 and per Tier-2 cloud)
- We'll have to take operational decisions if the available disk is << than the pledge
 - We may have to delete AOD/DPD datasets from (some) Tier-1s and leave them only in Tier-2s for user analysis
 - Or recognise that not all Tier-1s (and their Tier-2 clouds) are equal and adjust AOD/DPD replicas accordingly
- In the meantime, we have developed tools to collect statistics on dataset usage and we will use it to evaluate how many replicas are needed for different dataset families



What is DDM

- Distributed Data Management System
 - *Baseline Service for Data Distribution and Consolidation, MonteCarlo production and Reprocessing, Data Analysis*
 - *Provides functionalities for Data Organization (bookkeeping) and Data Transfers*
 - Data Organization based on Datasets
 - Collection of Files
 - Unit of data location and data replication, data deletion
 - Data Transfers based on Subscriptions
 - Placement policy, which DDM tries to enforce
- DDM is a Distributed Service
 - *Central Catalogs, Accounting Service*
 - Hold dataset definitions, subscription requests, Dataset Locations
 - *Site Services, Deletion Service*
 - Triggering actions to enforce subscriptions
 - *Other WLCG services at sites*
 - FileTransferService for data movement (at T1s), StorageElements (at T1/T2s), File Catalogs (at T1s) hold information about single files placement
 - *Command Line Interfaces*



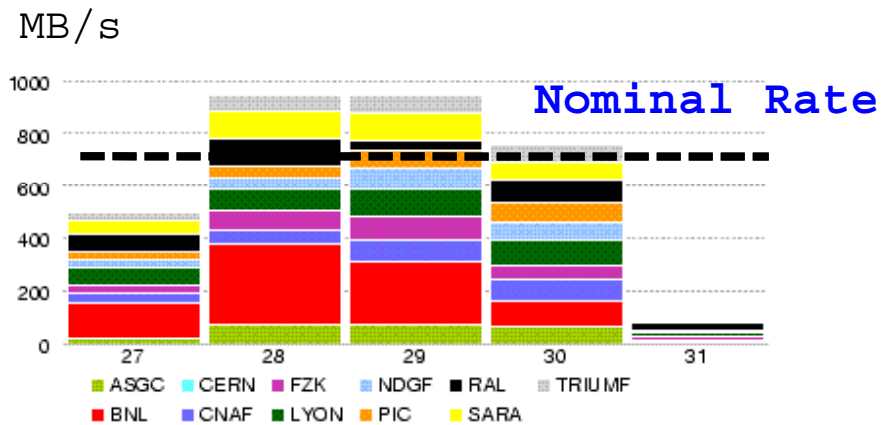
Readiness status of DDM

- Since Jan 2008, DDM is constantly under testing
 - Weekly Functional Test (low transfer rate)
 - Test both ATLAS specific and WLCG services at sites
- Challenges
 - Test particular aspects of DDM, mostly scalability and reliability under stress
 - CCRCs (Feb and May 2008)
 - Full Scale Test of the system at 150% of nominal data taking rate
 - Includes T0-T1, T1-T1 and T1-T2 transfers
 - Mostly a throughput test
 - 10M files test (Jan 2009)
 - Test of the system moving large number of files
 - 1M files imported by each T1 in 10 days
- Real Activity
 - Detector Data Distribution
 - Little beam events, lots of cosmics + reprocessing data
 - MonteCarlo data Distribution and Consolidation

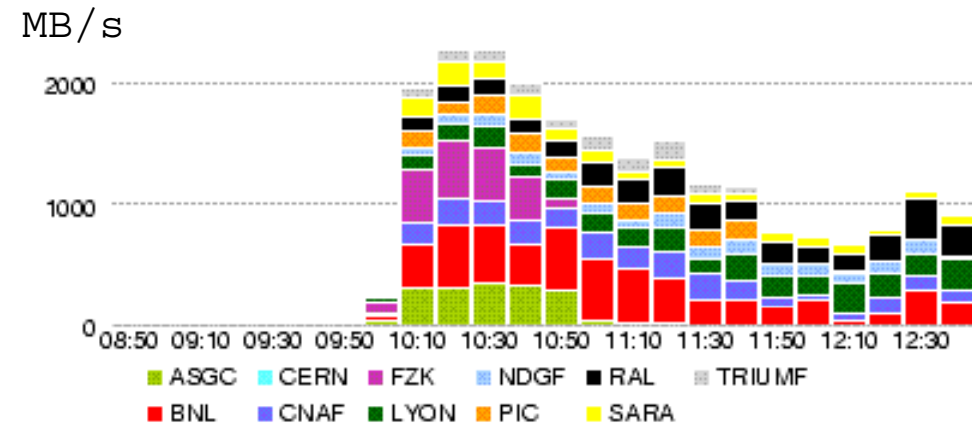


Experience in 2008: CCRC08

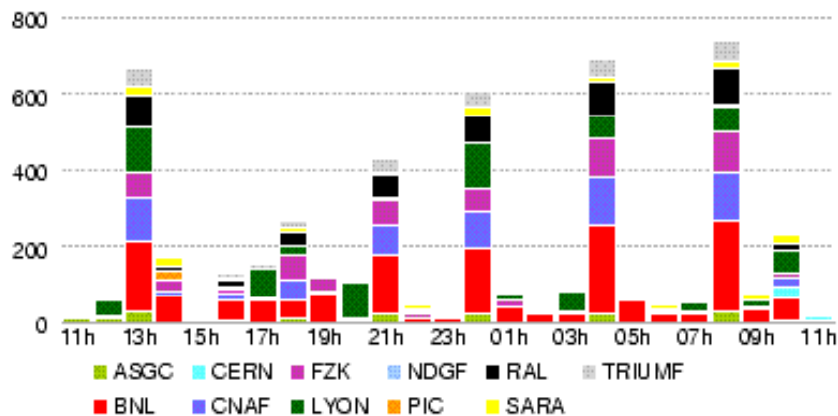
T0->T1s throughput



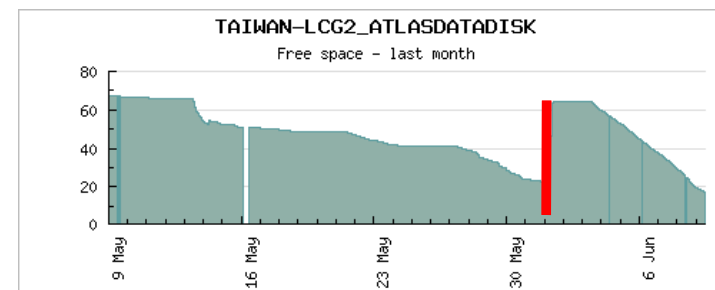
12h backlog recovered in 90 minutes



T1->T2 transfers



1 PB data deleted in 1 day

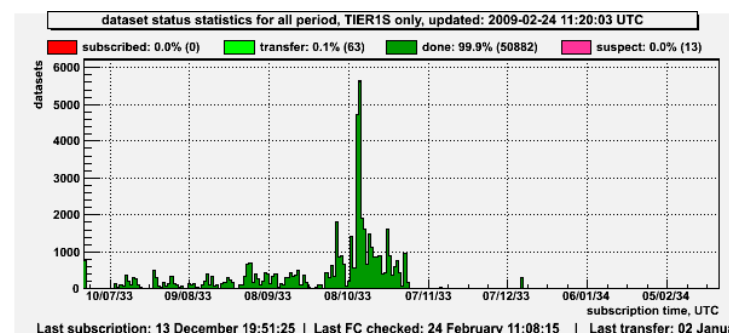




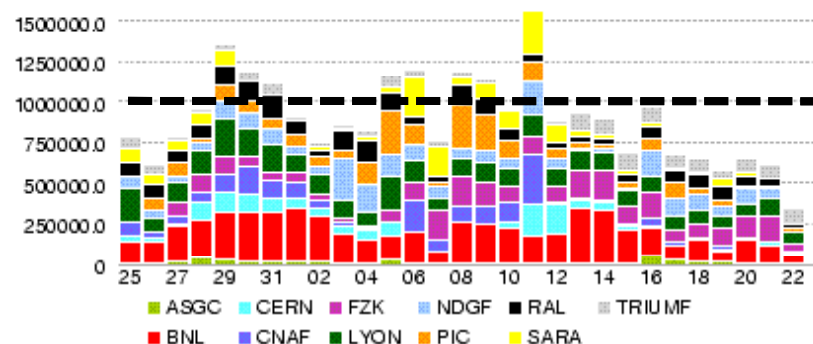
Experience in 2008: production data

- 15 days of problematic period in Detector Data Exports
 - October 2009
 - Consequence of (Agreed) change in TO workflow
 - To deal with data distribution long (1 day) runs
 - DDM limitation dealing with > 5K concurrent subscriptions
- Improvements in DDM
 - Tested in 10M files test (Jan 2009)
 - Further changes in TO workflow
 - Will be tested in a few weeks
- 10 M file test showed expected improvements in DDM scalability
 - Can now handle > 10K simultaneous subscriptions
 - 100K files transfer per day per T1
- Results are already obsolete
 - MC production and Data Consolidation moved > 30M files in 30 days.
- **This is NOT A POSITIVE ASPECT: ATLAS data are fragmented in too many files of small size.**

Detector Data Distribution



MC production and Data Consolidation





Experience in 2008: new functionalities

- DDM offers an accounting system
 - Currently per site, will be extended to groups per site
- DDM offers a central deletion service
 - Big step forward, 1 year ago dataset deletion needed intervention from the sites
 - Partial Synchronization of various components
 - DDM, ProdDB, AMI
- DDM offers a Consistency service
 - On demand verification of consistency information between catalogs
- SS now also handle data staging from Tape
 - Partially tested during reprocessing preparation
- We have a subscription interface to validate and handle user requests.
 - Placement of ATLAS data into production areas
- We now have an Integration Testbed
 - Includes all services (also monitoring)
 - Very important for validation of new SW releases, packaging, etc ..
 - Offers more clear workflow for service deployment



Experience in 2008: Users and Groups

- Clients for user dataset upload/download exist
 - Include support for logging data access patterns
 - Dataset popularity per site/user etc ...
 - very important to improve data placement and for the system modeling
 - Should be used as common library by every system
 - Production system and Analysis frameworks: work in progress, good cooperation
- But still ... DDM offers very little support for group and user analysis
 - DQ2 clients need to be more resilient against failures
 - Minimal infrastructure for User Dataset replication
 - Deletion of user datasets is very problematic
 - Needs central operation intervention
- This is due (IMHO) to
 - Difficulty to understand user requirements
 - We need to give more thinking to it, was not the primary use case so far
 - Need to collect and focus more on User Community requirements
 - Reference person from User Community?
 - Logging of access patterns is a good starting point.
 - Absence of User Support
 - Many times DDM operations and DDM development must take care of it



Experience in 2008: other issues

- Staging from Tape is part of DDM
 - The baseline service is there but is the most untested part of the system
 - Both at the level of DDM but also Storage Services at T1s
- Situation of the CERN site
 - Internal data replication works at 100MB/s but we have too much fragmentation of CASTOR disk pools
 - Not transparent to users
 - This is being followed up with CASTOR team at CERN
 - Promising setup for the beginning of April, better support for analysis activities
 - ATLAS Should appoint a CONTACT person at CERN
 - We have for all T1s and several T2s
 - This is a almost full time job
 - CERN is a very peculiar Grid/NonGrid site
- Data Management operations still suffer from service unavailability at sites
 - Need to handle this in a more automatic way
 - ATLAS Grid Information System



Work plan for 2009/10 and beyond

- Short Term Plans (next 4 months)
 - **Development**
 - Integration with ATLAS Information System (including possible extensions)
 - Full Support for Groups (accounting, quotas, placement requests)
 - Full support for Datasets deletion (might be more mid term plan)
 - Hardening of Staging Services
 - Several minor (but important) improvements at various DDM services.
 - TRAINING!!!
 - **Operations**
 - Improved Documentation for Users and Groups.
 - Aim for 100% dataset completion at final destinations
 - Includes more monitoring about problematic subscriptions and files
 - Final setup of Storage infrastructure
 - Mostly has to do with Users and Groups
 - Heroic effort in dark and inconsistent data cleanup
 - Large scale test (including new TO dataflow)
 - And more focus to staging tests.



Work plan for 2009/10 and beyond

- Medium Term Plan (2009/10): preserve the current functionality “streamlining” the system.
 - Automate operation model
 - Based on first months experience
 - Re-implementation of subscription catalogue
 - Merge of Location/Subscription/Deletion catalogues
 - Will provide a state machine for subscription requests
 - Implementation of automatic data placement policies
 - Possibly based on dataset importance and popularity
 - Revisit deployment model (could go to Long Term)
 - Fully distributed file catalogues? Centralized file catalogues?