Squad Activities

L. Poggioli, LAL Orsay
On behalf of FR-Squad

- Framework
- Squad operations
- Interaction with T1

The Actors

- ADC Central Operations Team
  - Group of experts of various ADC components
- ADC Expert on call (AMOD)
  - Interface between shifters & experts
- ADCoS shifters
  - ATLAS wide
  - Notify sites (ggus) & squads (email, savannah)
- Cloud squads
  - Treats cloud issues & ATLAS-specific issues at sites
  - Interface between sites & Central operation
- Sites
  - Site issues, may consult squad for ATLAS-specific issues

ADC shift teams

- ADC@Point 1
  - Data export from T0 to T1s
  - Central services (DDM)
- ADCoS: Shifts on Distributed Computing
  - ATLAS wide, across clouds
  - Data distribution T1-T1, T2
  - Official production (MC, Reprocessing)
  - Central services (DDM, ProdSys)
- DAST: Shifts on Distributed Analysis
  - ATLAS wide
  - User analysis on Grid
  - User Data access & Replication

Cloud Squads

- Treats issues within the cloud
  - Cloud-wide issues
  - ATLAS-specific site issues in cloud (ATLAS files, ATLAS jobs)
- Interface between sites & Central Operation
  - Translates ATLAS language -> site/WLCG language
  - Supplements lack of knowledge in central operation about sites & cloud
  - Follows-up implementation of central decisions into sites/clouds
    - eg CVMFS, T2D
- In place since 2010
  - 7/7 activity

FR-Squad

- Operation 7/7 on a week-time basis
- 5 persons
  - Sabine Crépé-Renaudin (LPSC Grenoble)
  - Emmanuel Le Guirriec (CPPM Marseille)
  - Tristan Beau (LPNHE)
  - Camelia Visan (Bucharest)
  - LP (LAL Orsay)
  - Previous members: Irena Nikolic (LPNHE Paris) & Wenjing Wu (IHEP Beijing)
- In addition
  - E. Lançon constantly following all cloud issues

Squad activities

- T3 share part implementation
- T2D implementation
- Network performance/issues
- Factory/SchedConfigs/voboxes
- CREAM-CE implementation
- Follow-up new batch system at Lyon
- Interplay, VL queues & T1/T2 at Lyon
- Monitoring improvement
- Overall communication
- CVMFS implementation
- Hammercloud tests follow-up

Inputs (1)

- Directly from sites
  - Temporary problem (clim, server, disk)
  - Please inform Squad when changes at site (hw, sw)!!
- By users
  - Analysis jobs crash/Data transfer issues

  -> Might indicate a more general problem
- By DAST
  - Handled by Squad if Site/Cloud specific (ie; Site in unscheduled downtime)
- Functional tests
  - Centralized for analysis & production

  -> Might indicate a more general problem

Inputs (2): ADCoS

- Centralized for all ATLAS
  - 24/24 7/7
- Summary 3/24 all ATLAS
  - Liste of problems in clouds, alarms, pilots, ongoing tickets
- Summary 1/24 for FR
  - Sites status (offline, brokeroff)
  - Ongoing tickets (ggus, savanah)
- Tickets+ Elogs
  - ggus, savanah

Success implies Proactivity

Meetings & Communication (1)

- Central ATLAS
  - ADC weekly follow-up
  - S&C week
- FR-Cloud
  - Monthly CAF (Computing ATLAS France)
- Communication to Squad
  - Email: atlas-support-cloud-fr
  - GGUS tickets
    - Setting offline/Brokeroff site

Communication (2)

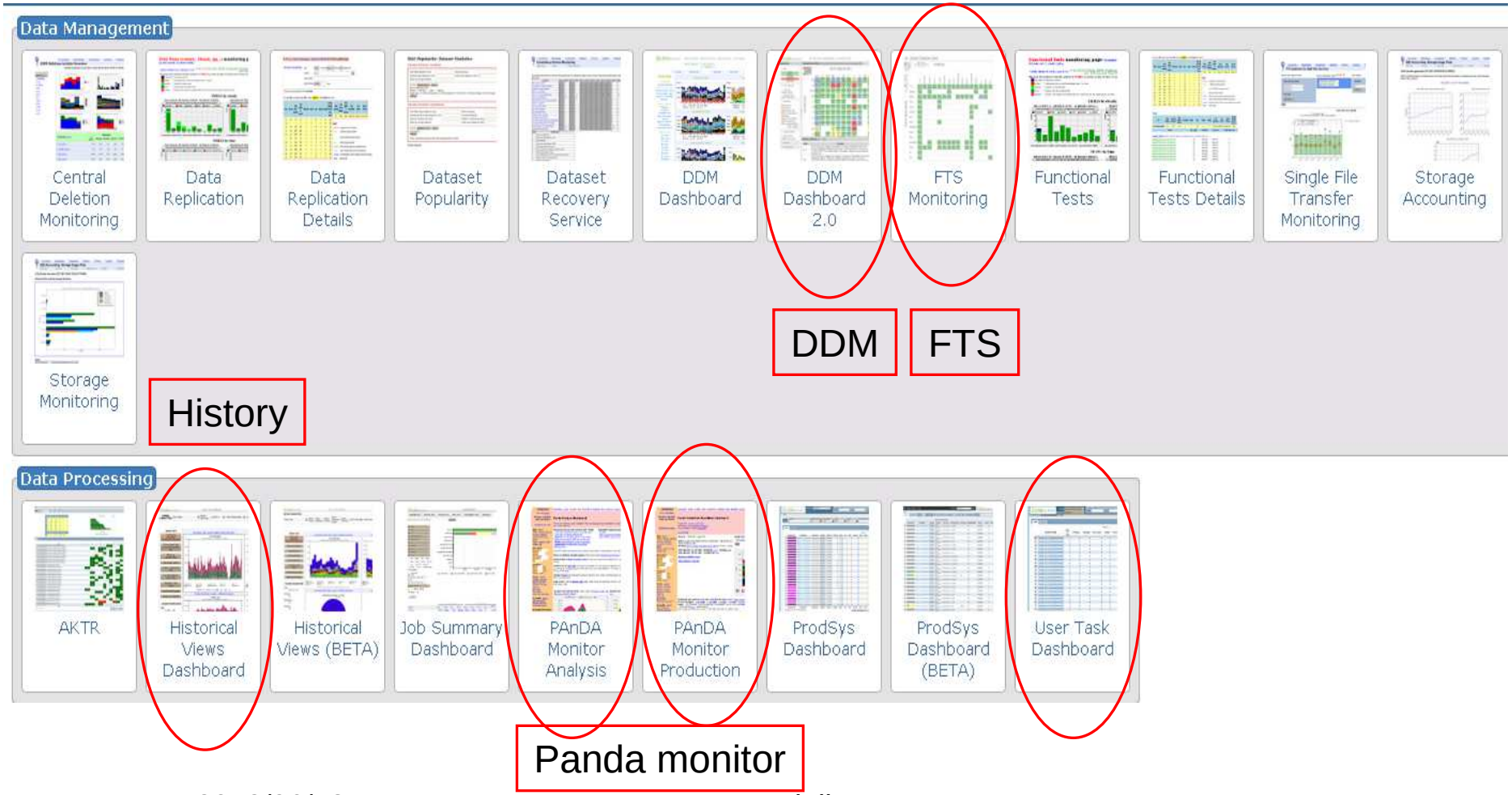
- Savannah tickets
  - Athena failure at sites, DDM exclusion
- Dedicated tickets to Lyon
  - For problems concerning, eg for Reprocessing
- 'French' Elog

http://isnpx0162.in2p3.fr/elog/ATLAS-FR-Cloud/

  - Squad weekly reports
  - Specific actions (Factory, SchedConfigs)
  - Meetings minutes, eg ADC weekly
  - All sites from FR-cloud are included

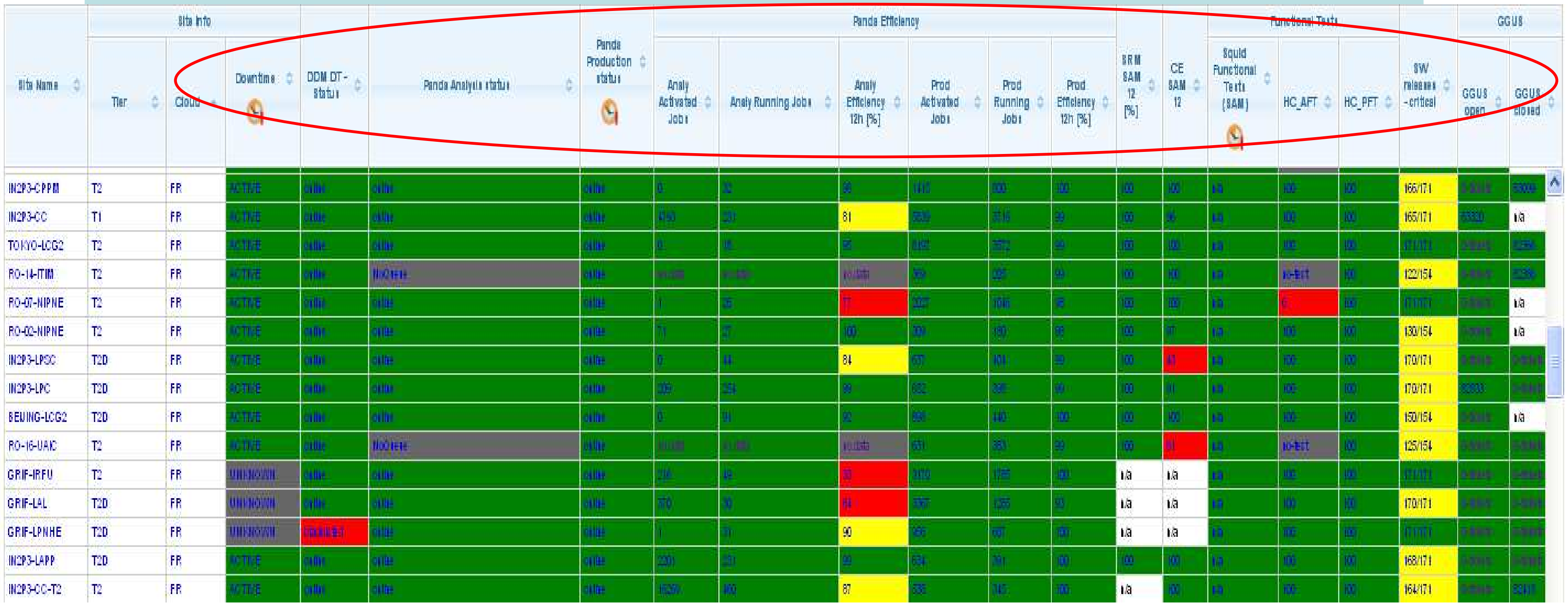
ATLAS Monitoring(1)

- http://adc-monitoring.cern.ch/
- Data Management, Data Processing, Databases, Sites & Services



Monitoring ATLAS (2)

SSB: Cloud view (new)



- Downtime
- DDM status (black/whitelisted)
- Panda status & efficiency (Analysis & Production)
- SAM tests (SRM & CE)
- Functional tests status (Squid, HC, PFT)
- SW releases installation status
- ggus tickets status

Follow-up CVMFS deployment

SITE	ALREADY/FRAC.	BEFO. XMAS	AFT. XMAS
RO-16	NO	YES	
RO-02	NO	YES	
RO-07	NO	NO	YES (T2/T3)
RO-14	NO	NO	YES (T2/T3)
LPC	YES (33%)	YES	
LPSC	NO	YES	
LAPP	YES (100%)		
CPPM	YES		
BEIJING	NO	YES	
TOKYO	NO	NO	YES (T2/T3)
LYON	YES (100%)		
LAL	YES (100%)		
LPNHE	YES (100%)		
IRFU	YES (100%)		

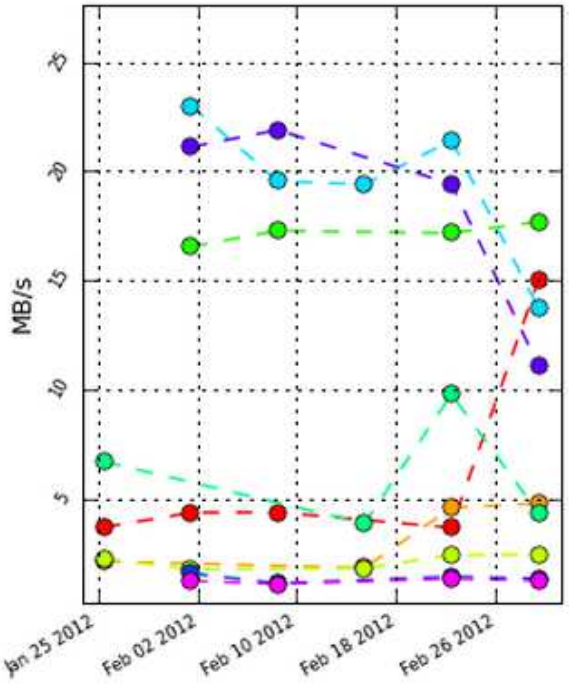
Today:
All sites CVMFS except
RO-02, RO-14, RO-16

Follow-up of deployment:
http://dashb-atlas-ssb.cern.ch/dashboard/request.py/siteview?view=cvmfs#currentView=cvmfs&highlight=false

Networking: T2D implementation

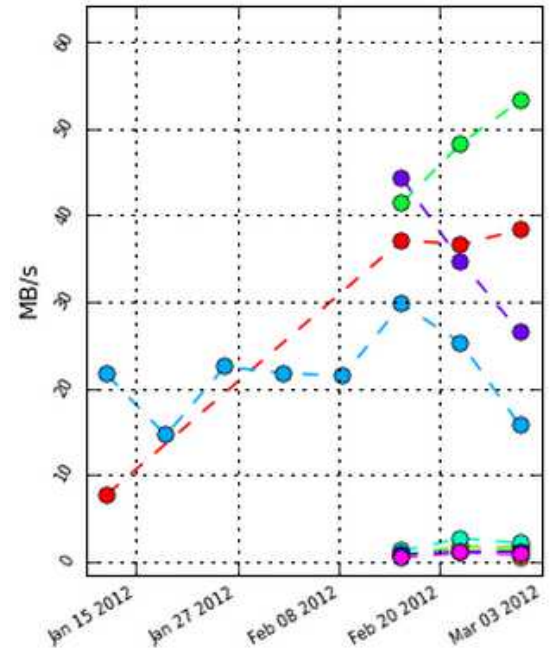
- Still IRFU & CPPM not T2D in French sites
- Criteria
  - All transfers T2D -> 10/12 T1s for big files to be above 5 MB/s during last week & 3/4 last weeks
  - Idem in other direction
- Improvements
  - Reevaluation every month - Include in SSB metrics
  - Review criteria (eg pbs at T1 affect candidate)
- For Squad
  - Follow perfSonar tests & interact with sites
  - http://dashb-atlas-ssb.cern.ch/dashboard/request.py/siteview#currentView=Sonar&highlight=false
  - Update FTS channels at IN2P3-CC

IRFU



GRIF-IRFU_DATADISK - CERN-PROD_DATADISK (25 file
GRIF-IRFU_DATADISK - SARA-MATRIX_DATADISK (20 fil
GRIF-IRFU_DATADISK - NIKHEF-ELPROD_DATADISK (25
GRIF-IRFU_DATADISK - FZK-LCG2_DATADISK (20 files)
GRIF-IRFU_DATADISK - RAL-LCG2_DATADISK (20 files)
GRIF-IRFU_DATADISK - IN2P3-CC_DATADISK (471 files)
GRIF-IRFU_DATADISK - PIC_DATADISK (20 files)
GRIF-IRFU_DATADISK - INFN-T1_DATADISK (20 files)
GRIF-IRFU_DATADISK - NDGF-T1_DATADISK (20 files)
GRIF-IRFU_DATADISK - BNL-OSG2_DATADISK (0 files)
GRIF-IRFU_DATADISK - TRIUMF-LCG2_DATADISK (0 file
GRIF-IRFU_DATADISK - TAIWAN-LCG2_DATADISK (0 file

- IRFU -> Lyon, FZK, CNAF OK (connected to LHCONE?)
- Around March 1st: CERN OK, but drop for CC&CNAF
- Bad for other T1s



CERN-PROD_DATADISK - GRIF-IRFU_DATADISK (54 files)
BNL-OSG2_DATADISK - GRIF-IRFU_DATADISK (5 files)
SARA-MATRIX_DATADISK - GRIF-IRFU_DATADISK (24 files
NIKHEF-ELPROD_DATADISK - GRIF-IRFU_DATADISK (15 fi
FZK-LCG2_DATADISK - GRIF-IRFU_DATADISK (100 files)
RAL-LCG2_DATADISK - GRIF-IRFU_DATADISK (12 files)
IN2P3-CC_DATADISK - GRIF-IRFU_DATADISK (37586 files
PIC_DATADISK - GRIF-IRFU_DATADISK (16 files)
INFN-T1_DATADISK - GRIF-IRFU_DATADISK (20 files)
NDGF-T1_DATADISK - GRIF-IRFU_DATADISK (15 files)
TRIUMF-LCG2_DATADISK - GRIF-IRFU_DATADISK (0 files)
TAIWAN-LCG2_DATADISK - GRIF-IRFU_DATADISK (0 files)

Lyon, FZK, CNAF, CERN -> IRFU OK
- Stats missing for other T1s

Networking: LHCONE deployment

- LHCONE evaluation managed -> gain expertise in liaison with network providers
- List of T1s &T2s to be followed up
  - T2s: 8 T2s incl. LAL, Tokyo for FR
  - T1s: ASGC, BNL, CERN, PIC, SARA, TRIUMF
- For these sites
  - Deploy Perfsonar-ps
  - Central coordination with network providers
- For other sites
  - If already connected to LHCONE, OK but responsible for network problems (ADC will follow)
  - If not connected to LHCONE, sites should hold off
- Cloud Squad in contact with T2s

Space Tokens

- Tokens management
  - By ATLAS centrally: DATADISK, HOTDISK, PRODDISK SCRATCHDISK
  - By ATLAS groups: GROUPDISK
  - By site: LOCALGROUPDISK

http://bourricot.cern.ch/dq2/accounting/cloud_view/FRANCESITES/30/

- When spacetoken getting full, alert is sent to squad
  - For ATLAS managed space
    - Squad checks pledge have been deployed
    - Squad contacts ATLAS to solve the problem
  - For Site managed space
    - Squad contacts the site

CREAM-CE deployment

- Implementation done by Squad
  - In Factory
  - In SchedConfigs (production & analysis queues)
  - LCG-CE queues decomissioning
- Problems with CREAM-CE
  - Configuration & Middleware stability
  - More visible when LCG-CE are decommissioned (eg GRIF-IRFU)
  - Followed by Squads (errors at pilot level)

Panda Pilots & Factory

- Maintenance of pilots & factories done by Squad
  - Check uniformity across 2 voboxes
  - Factory updates using rpm
- Pilots submission parameters to be optimized
  - pilotLimit, depth, Boost
- Problems tracking
  - Follow-up Condor status (eg pilots in unsubmitted state)
  - Tracking in pilot logs submission errors
  - Very useful for CREAM-CE running

Panda queues: What is new

- Analysis & Production queues status, ie online/test/ are handled CENTRALLY
- Fiducial tests are run many times a day to black/whitelist queues
- Squad can decide to put site offline if pbs
- Scheduled Downtimes handled now CENTRALLY
  - More generally all downtimes published in AGIS

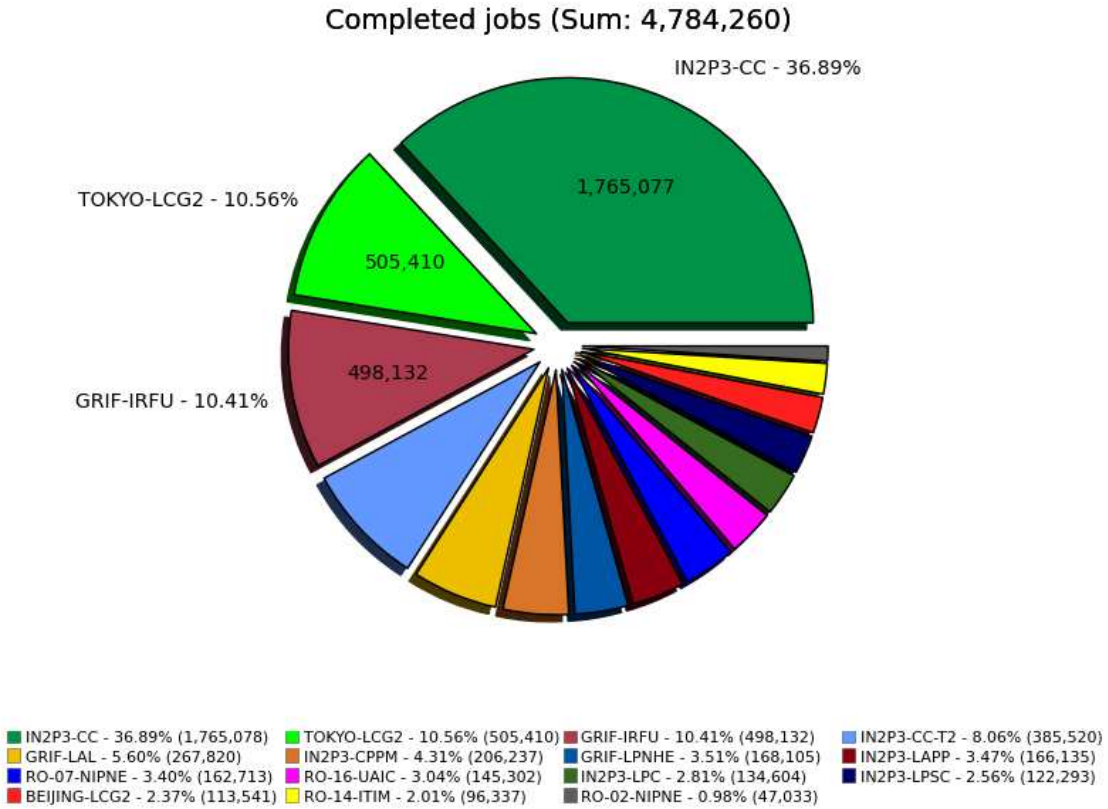
Production Functional Tests (PFT)

- Identical to AFT for Analysis queues
- In place since 1 week for all clouds (but ND)
  - Only AtlasG4 tests used (3 tests)
  - Auto-exclusion/whitelisting computed every 30'
  - Low failure rate, pbs for a few sites (~10)
    - >90% sites efficiency >90%, >80% sites effic. >95%
  - Notification OK, as for AFT
- Next
  - Reco & Evgen test soon used

SSB switcher (new)

- Auto-queue exclusion based on downtimes
  - Idea: less manual ops, faster queue recovery
  - SRM/CE DT queue automatically set to test
  - Only if DT > 4hrs
- Practically
  - Production queues
    - T-8h: set offline
    - T: Downtime starts: keep offline
    - T+D: Downtime finishes: set test
  - Analysis queues
    - T-6h: set brokeroff
    - T-2h: set offline -Then as for prod

Production in FR



Since January 1st

- 4.8M jobs
- Lyon-T1, Tokyo, IRFU

Production follow-up (1)

- Monitoring
  - Basis : Panda monitor

http://panda.cern.ch:25980/server/pandamon/query?dash=prod

    - Errors monitoring : Per site, per task, per type
    - Jobs status
      - If there are jobs (« defined »), no « assigned « , not too many « activated » (-> # pilots)
  - Also panglia monitoring (jobs/state) http://gridinfo.triumf.ca/panglia/
  - Also Summary dashboard

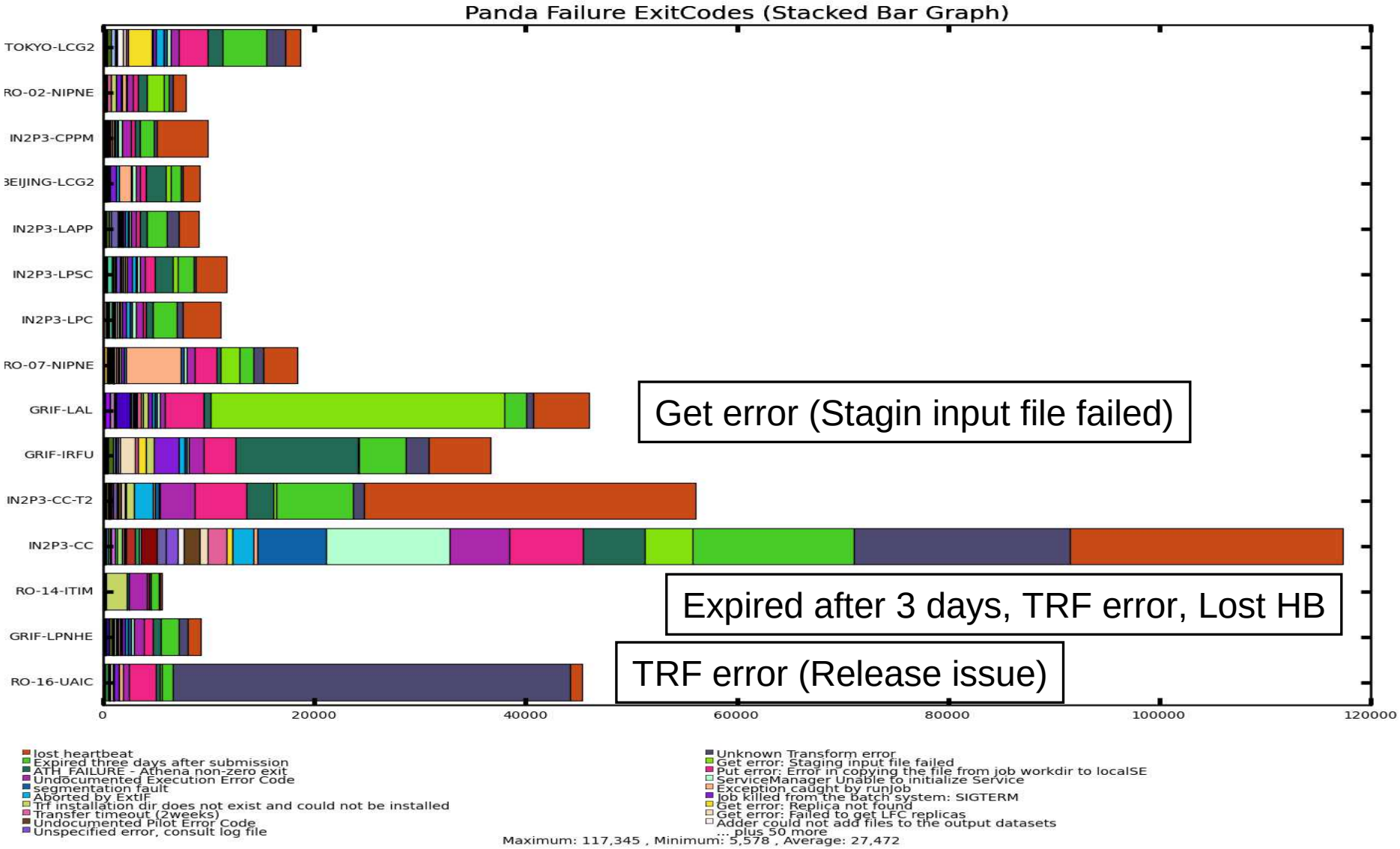
http://dashb-atlas-jobdev.cern.ch/dashboard/request.py/jobsummary

  - Also Lyon monitor (eg dcache limit reached) http://cctools2.in2p3.fr/mrtguser/mrtguser/ccin2p3/lcg_helene_atlas.html

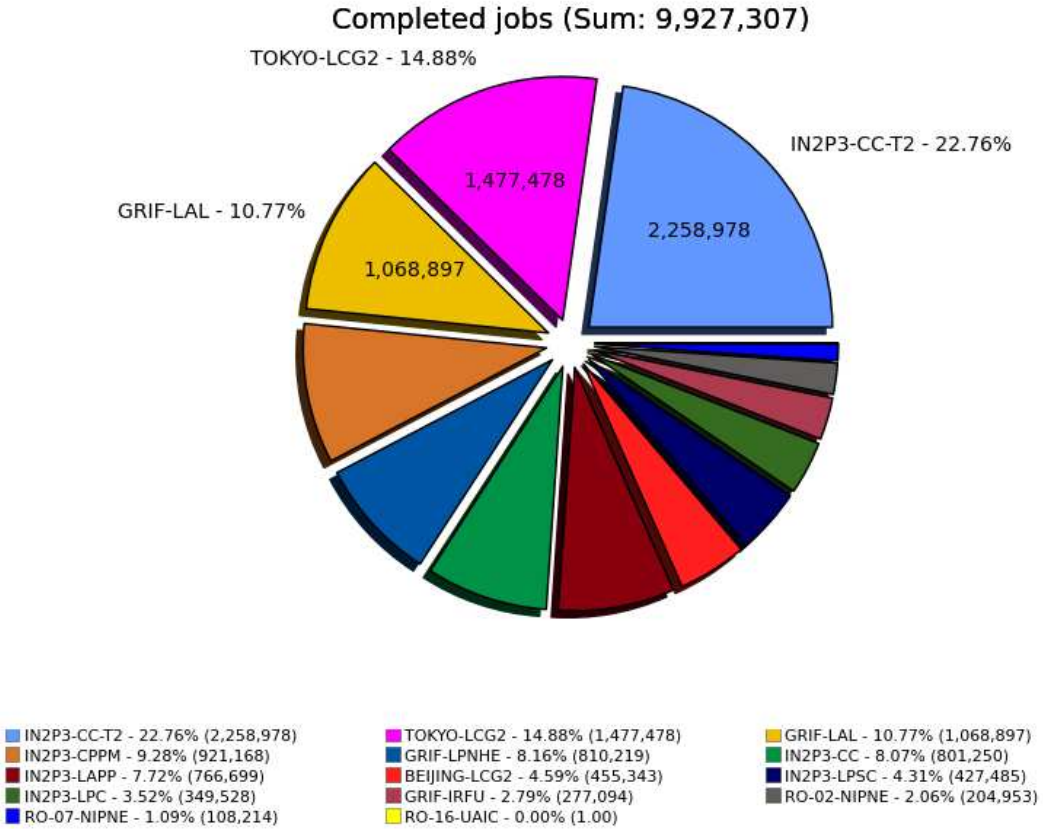
Production follow-up (2)

- Queues validated via automated PFT
  - As for analysis
- Inform Sites if problems
  - Put site in test-> now done automatically
  - GGUS Team ticket + Elog
- Downtimes (T2,T3) (T1 via ADC experts)
  - Now handled centrally
- Make sure tasks are assigned to FR

Production FR: Errors



Analysis in FR



Since January 1st

- 9.9jobs
- Lyon-T2, Tokyo, GRIF-LAL
- Analysis also at Lyon-T1 (reduced share)

Analysis Follow-up

- Same basis as Production

  http://panda.cern.ch/server/pandamon/query?dash=analysis

  - Less systematic than for Production
  - Check errors/site rate
  - Interface with DAST
- HammerCloud: centralized test setup
  - Test performed regulary on all analysis sites
    - Use benchmark jobs
  - Test used to:
    - Define availability & reliability of sites
    - Define white/blacklisting of sites

T2s sites availability & ranking (1)

- For analysis queues
  - To optimize analysis activity from users
- Ranking based on HC results wrt Site availability
  - Reevaluated every month
  - Impacts amount of data dirtibuted to T2s

		N_sites		Volume (PB)		Share per site	Share per group
alpha	T2D, >= 90%	21	32.31%	7.8	36.41%	2.92%	61.31%
bravo	T2, >= 90%	18	27.69%	3.13	14.58%	1.46%	26.28%
charlie	>= 80%	17	26.15%	5.87	27.37%	0.73%	12.41%
delta	< 80%	9	13.85%	4.64	21.65%	0.00%	0.00%
		65	100.00%	21.4	100.00%		100.00%

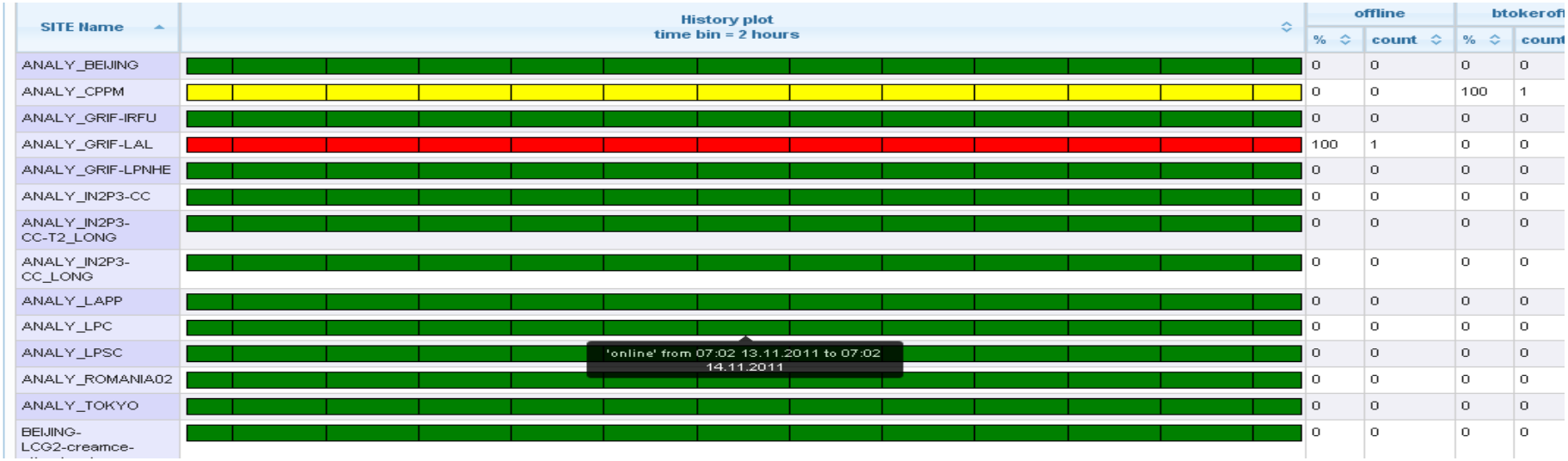
- For squads
  - Missing DS: HC switch to be more careful, clouds to be more proactive, slow transfers-> potential problem using sites (Cf. RO-02, RO-07)
    - Script implemented for informing clouds on missing DS

T2s sites availability & ranking (2)

- To follow site availability

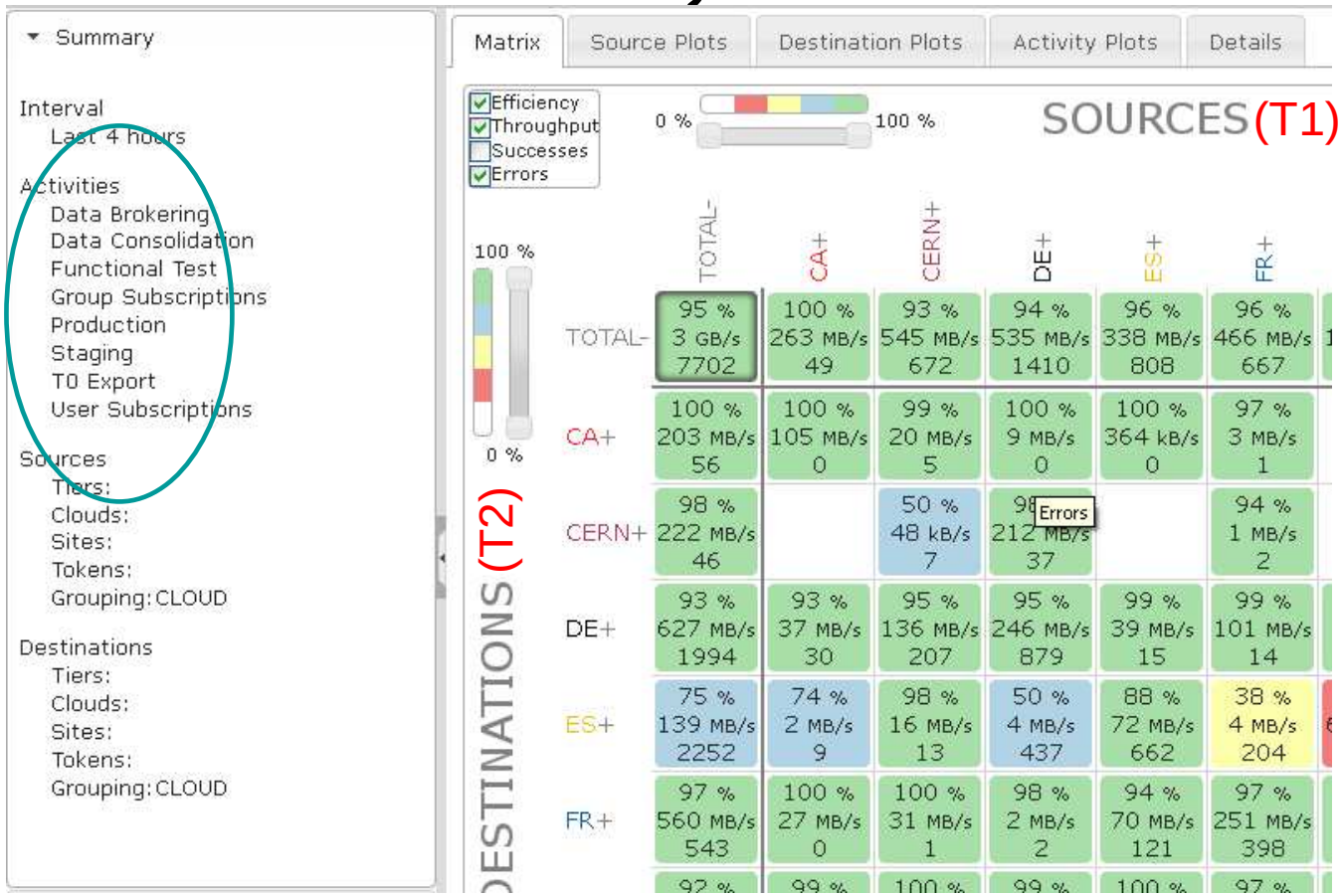
http://dashb-atlas-ssb.cern.ch/dashboard/request.py/siteviewhistorywithstatistics?columnid=562

- If Blacklisting by HC
  - Track back incident with time -> HC test # http://hammercloud.cern.ch/atlas/incidents/
  - With HC test #, check the failure reason



DDM & FTS Follow-up (1)

- DDM dashboard http://dashb-atlas-data.cern.ch/dashboard/ddm2/
  - Gives overall traffic inter-clouds (ie includes all DDM activities)

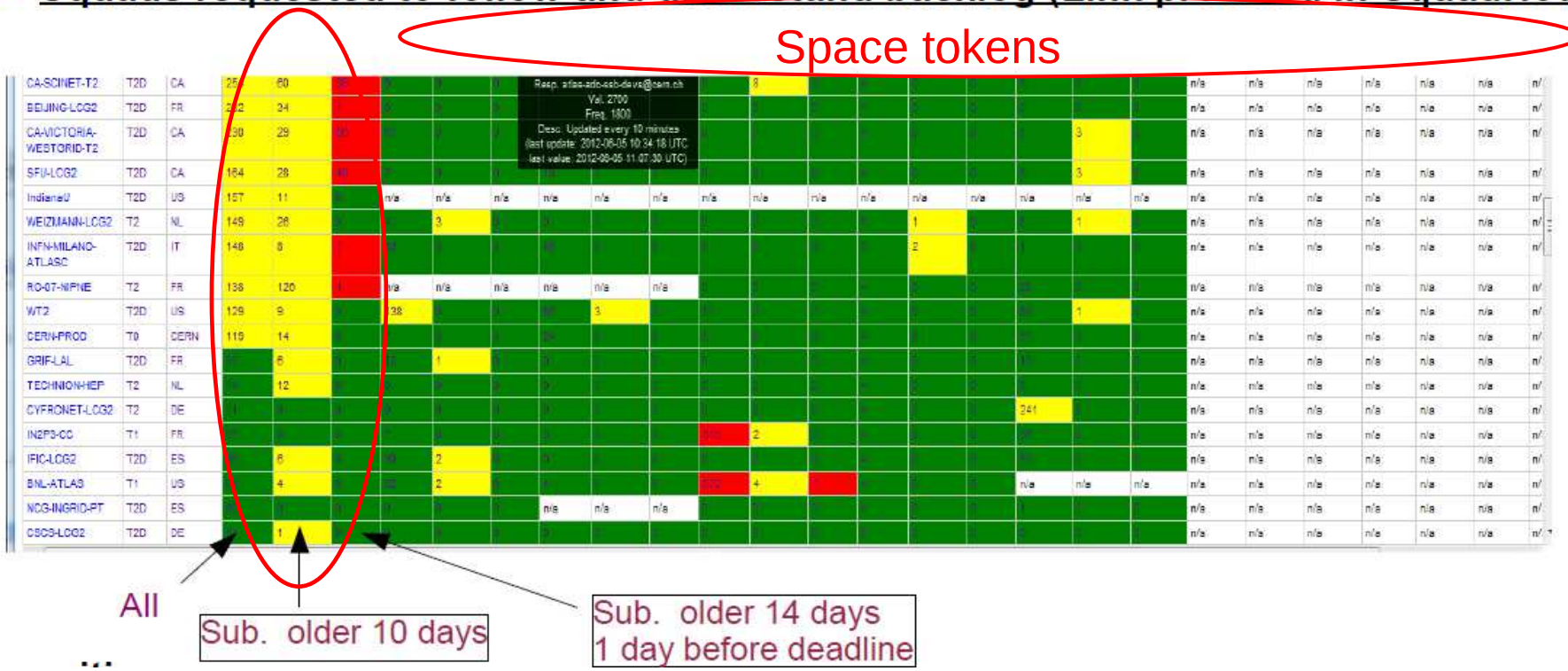


Access to:
- Throughput
- Efficiencies
- Errors
- Channel saturation

DDM Transfer backlog

S. Jézéquel

- Monitoring tool presented by Jarka few months ago
  - Embedded in SSB
  - Squads requested to follow and understand backlog (Link provided in SquadHowTo)



DDM & FTS Follow-up (2)

- ATLAS FTS monitor http://bourricot.cern.ch/dq2/ftsmon/test_view/
  - Allows follow-up of **1 file** transfer
  - Useful for debugging (Cf. Lyon-Tokyo transfer)
- IN2P3 FTS monitor
  https://cctools2.in2p3.fr/stockage/fts/monitoring/ftsmonitor.php?vo=atlas&timescale=12

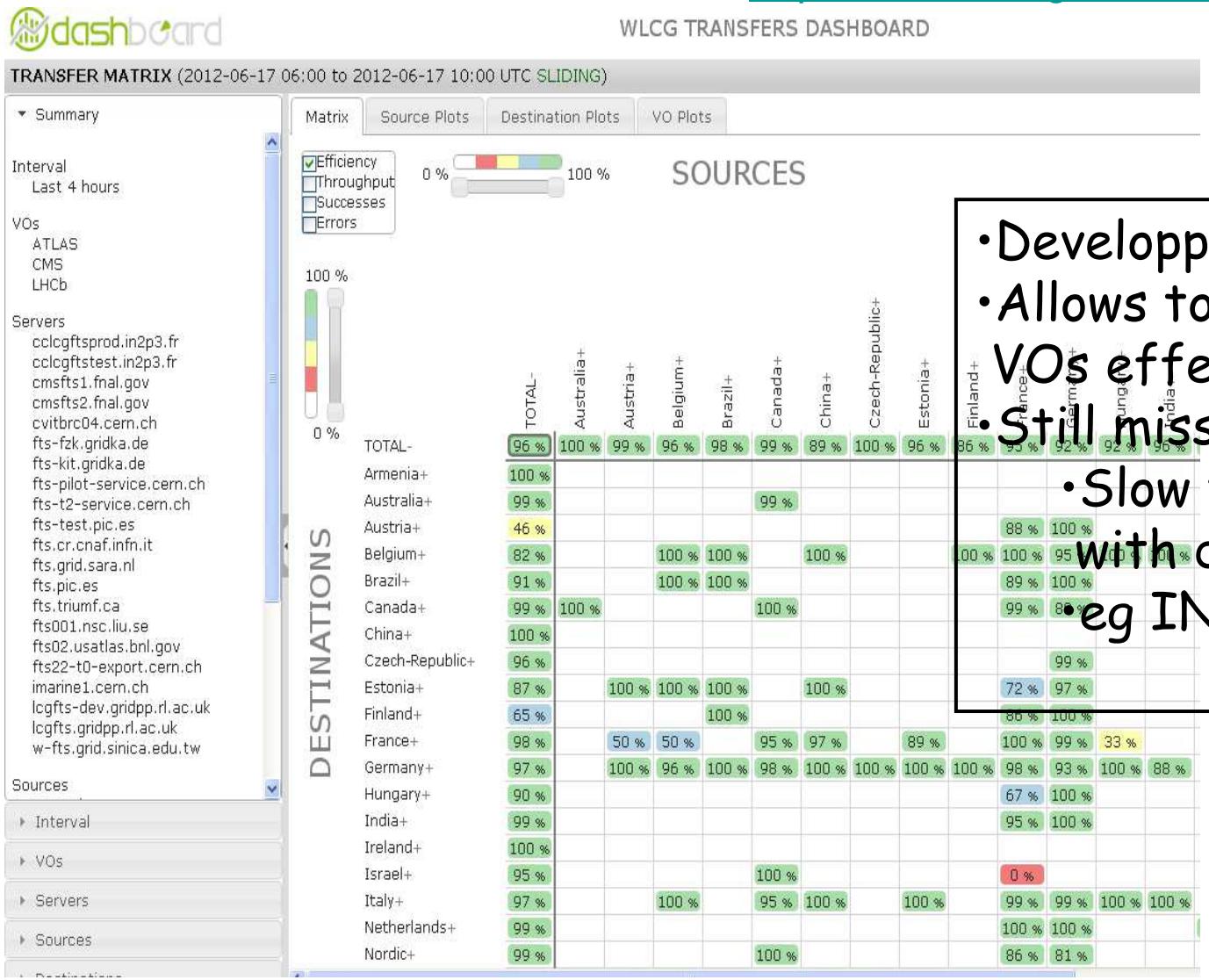
Jobs statistics (submitted last 12 h)

CHANNEL	Ready	Active	Finished	FinishedDirty	Failed	Canceled
TOTAL			2855			
ATLAST1-BEIJING		1	4			
ATLAST1-LAL	12		57	22		
ATLAST1-LAPP			8		2	
ATLAST1-LPC	1	1	8			
ATLAST1-LPNHE			3	1		
ATLAST1-LPSC	7			2		
ATLAST1-TOKYO			32			
ATLAST2D-IN2P3			274			

Access to:
- Transfers TO & FROM IN2P3
- T2D included
- Errors
- Failures

DDM & FTS Follow-up (3)

- WLCG FTS monitor http://dashb-wlcg-transfers.cern.ch/ui/#



- Developped centrally
- Allows to see concurrent VOs effect
- Still miss info
  - Slow transfer <-> pbs with disk servers
    - eg IN2P3-TRIUMF

User Analysis Follow-up (1)

- Interaction with users
  - CAF activity, not only Squad
  - Regular meetings with FR-physicists (twice/year)
- Evaluate needs & problems
  - Questionnaire being sent to users
  - Access to data
  - Resource in CPU
  - Storage requirements
  - PROOF facility usage at Lyon, in collaboration with Lyon

User Analysis Follow-up (2)

- PAF: Physics ATLAS France
- Outcome
  - Analysis uniformization among groups/users
  - Difficulty to retrieve outputs (eg LPC)
  - Jobs bookkeeping & failed jobs handling
- Most delicate
  - 1st analysis phase: User production
  - -> dedicated meeting in February

‘User Prod’ meeting 6/02

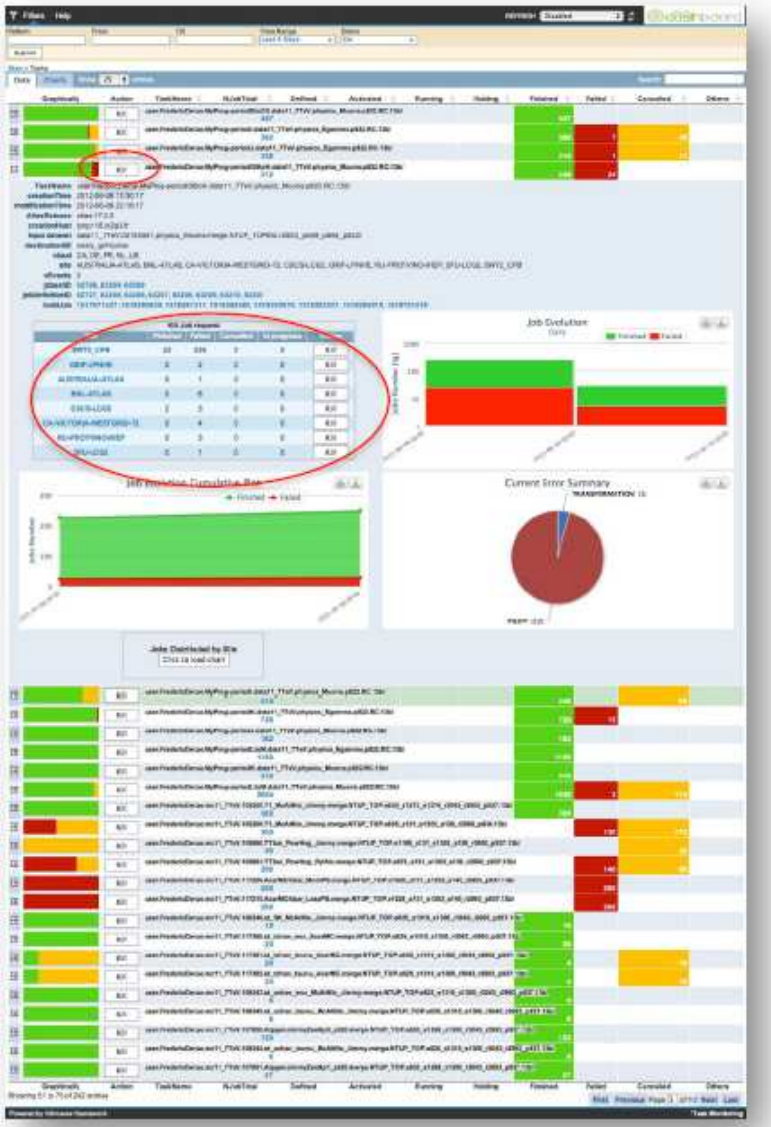
- Heavy /sps needs -> Lyon informed
- Using CHIRP?
- FR-user resources (localpool) implemented
- Simplification?
  - Move upstream: integrate **user production** into group production (monitoring done by ADCoS)
  - Share load among groups : cooperation between french groups (no such example)

J. Jaroslava
last S&C week

Analysis Task monitor

URL: https://dashb-atlas-task.cern.ch/templates/task-analysis/

Demo version is available on *http://dashb-atlas-task.cern.ch/templates/task-analysis/#demo=on*



Functionalities:

- Monitoring on the task (collection of jobs, based on output container) level and individual job level.
- User interface includes sorting and filtering possibilities:
- Tasks can be filtered by name or some pattern inside the name or by time ranges.
- Graphical representation:
  - Job Evolution, jobs distribution by site, distribution of failed jobs by failure error category (Figure 2).
- Summary charts for task and job tables.

Recent development and testing

- Enabling of cancellation on the task and individual job level from the user web interface

Development phase

- Display of tasks according to different roles of the user
  - Default (/atlas), Shifters (/atlas/Production), Group roles

Please take part on survey questions at
https://espace.cern.ch/adc-monitoring-surveys/Lists/Beta User Task Monitoring/overview.aspx

Laura Sargsyan
Manoh Jha
Lukasz Kokoszkiewicz
Edward Karavakis

Interaction with T1 (1)

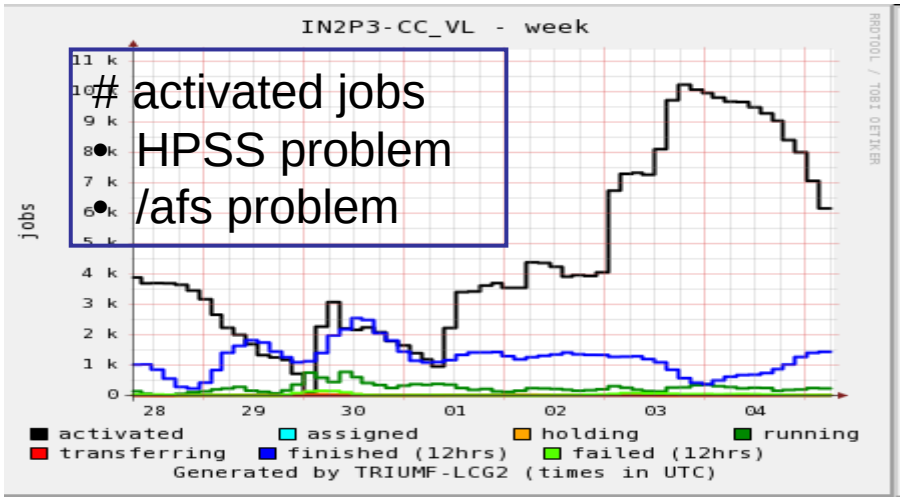
- Close collaboration between CAF (eg Squad) and Lyon on critical issues
  - Monthly meetings
  - Dedicated meeting on Reprocessing, Monitoring
- Monitoring
  - Define what are the needs from squad
    - dCache, SRM, SGE, FTS
  - Lyon -> provide us dedicated web pages with needed info
  - Aim is to give squad betters tools for debugging & gain in proactivity

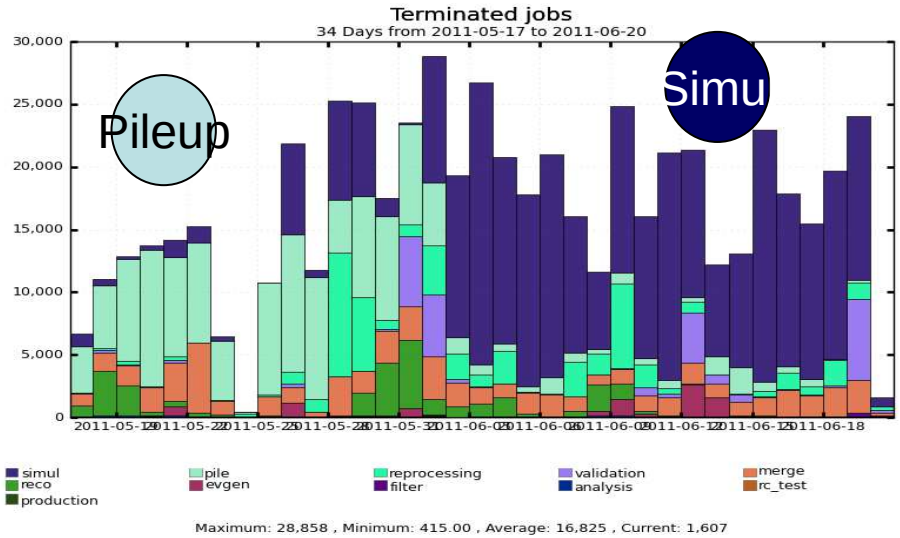
Interaction with T1 (2)

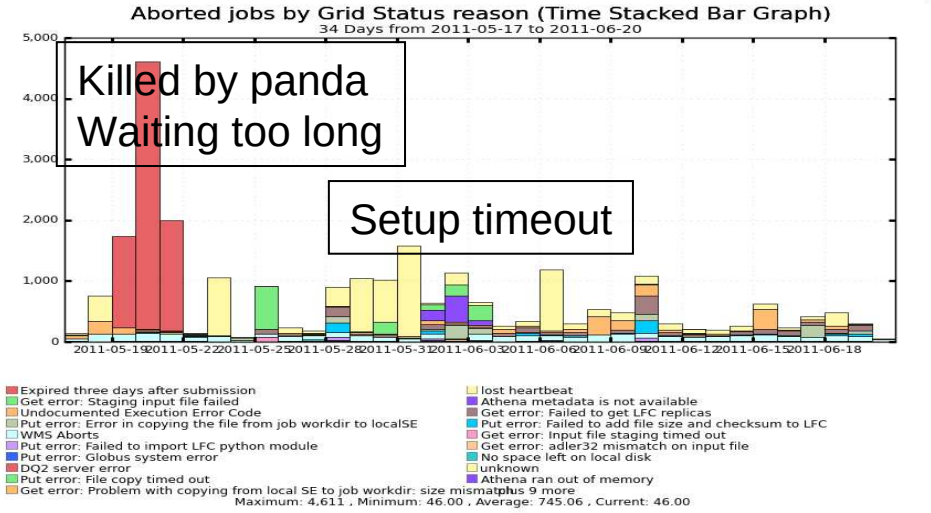
- Batch queues & CREAM deployment checks
  - CREAM-CE efficiencies, error messages
  - Implementation of new batch queues
- dCache
  - Inputs provided by Squad to Lyon experts
    - After problems in November reprocessing
  - Saturation in case of strong ATLAS activity
    - 20M ATLAS files, 50% are <10 MB
  - Actions
    - HW upgrade, SW upgrade, VO separation
  - No dCache problem in latest repro campaigns

Interaction with T1 (3)

- Production follow-up
  - Direct info on potential problems
- Tracking down low efficiency on production jobs

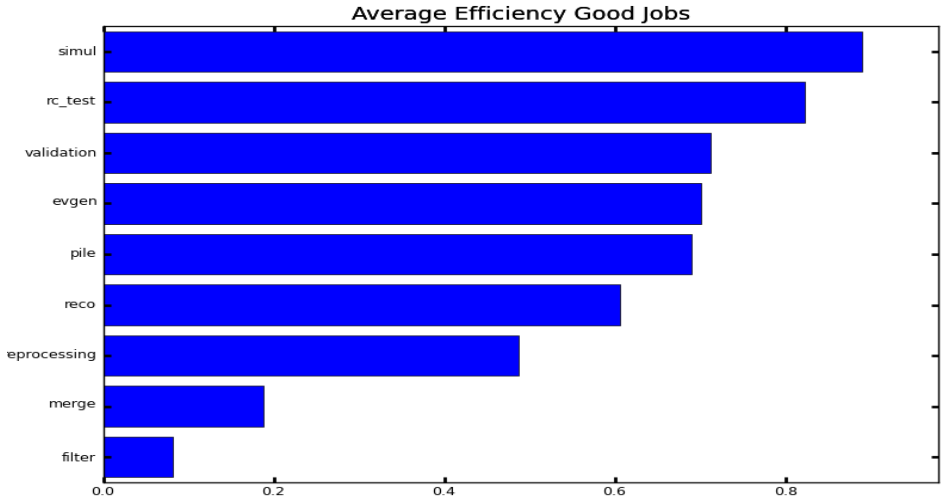






Interaction with T1 (4)

- Low efficiency issue



- Huge effort to understand
  - Eric + Pierre Girard
  - Due to # concurrent jobs @ WN level
- Solved by moving from /afs to CVMFS

- Reprocessing
  - Dedicated set-up for last Sep. reprocessing
    - Specific Elog / Limit to queues with CVMFS
    - Jabber discussion space with Lyon experts & Squad
    - Direct squad access to helpdesk during week-ends
  - Big success for latest reprocessing campaigns

Interaction with T1 (5)

- LFC migration at CERN OK
  - Even if Lyon is down, T2s still available for production & analysis
- Basic elements OK
  - dCache (storage), SGE (batch), CVMFS (releases)
- Multicore test setup
  - implemented
- Monitoring & Accounting
  - In progress but not satisfactory

Conclusions

- Squad work is made easier
  - Sites are more stable (panda queues)
  - Automated tests (HC, PFT) for prod & analysis queues
  - Downtimes handled automatically (SSB switcher)
  - CVMFS deployment
  - Please inform Squad when changes made at site (hw, sw)
- Most critical issue: Network
  - Better monitoring tools today
  - Also very complex issue (SRM, FTS, LHCONE)
  - Better expertise needed at Squad level
- Interaction with T1
  - Much higher level of reliability (CVMFS, SGE, dCache)
  - Maintain close interaction
  - Monitoring to be improved