

RO-LCG workshop

Magurele, 29th of November 2011

Sabine Crépé-Renaudin for the ATLAS FR Squad team

ATLAS news

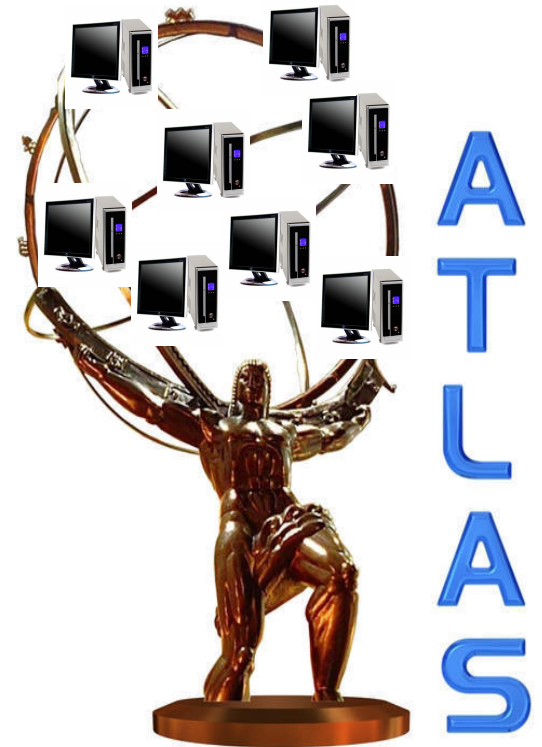
- Computing model : new needs, new possibilities :
 - Adding a third dimension
 - Software and database access

Grid operation : tools and teams

- Sites and activities
- Teams
- Romanian participation to Squad

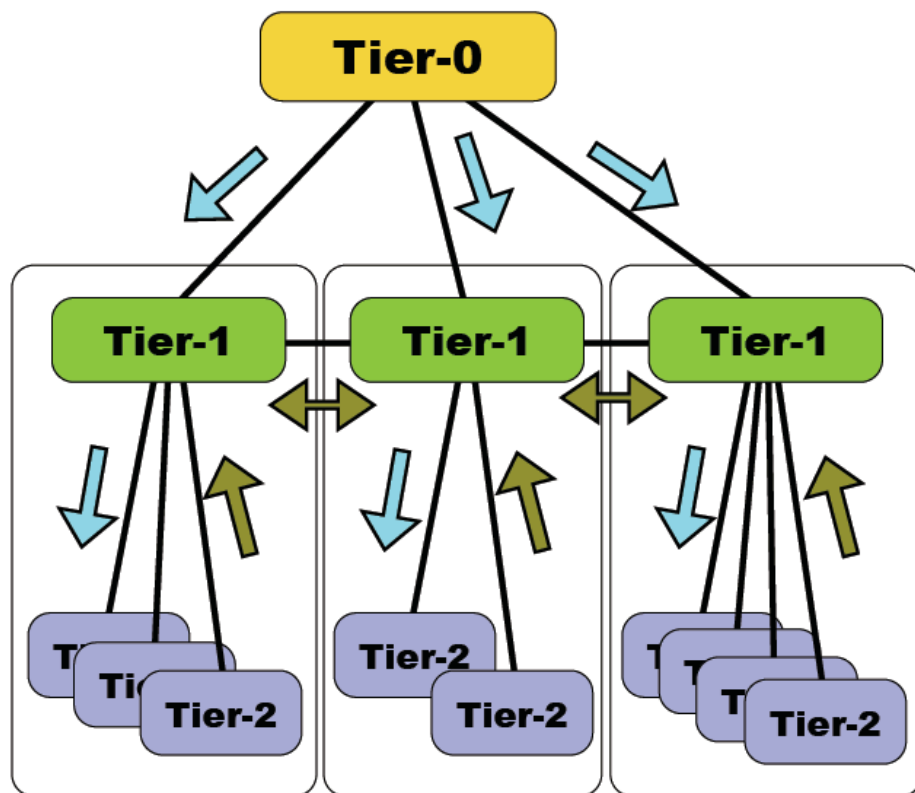
Grid activities over the year

- Statistics vs activities
- Statistics vs clouds/sites



- Hierarchical model : clouds are \sim independent

- T0 has predefined channels with T1s
- T1s have predefined channels with others T1
- T1s communicate with the T2s of their clouds
- T2s communicate only via the T1 of their cloud



- Pre-placement of data

Pre-placement of data did not correspond to user needs

AOD and dESD were pre-placed in the sites but

- ESDs for data performance study were more used than AODs
- ESDs were more used than derived ESD which were not enough tuned

Usage of data types were different than expectation, most of datasets were not used

A lot's of data are moving around

Needs of T2-T2 transfers : Output of user analysis on a T2 should be stored at an other T2 often belonging to an other cloud

- ⇒ Change of distribution policies
- ⇒ Add new paths between sites
- ⇒ Hierarchical model for data distribution abandoned
- ⇒ Break down clouds

Adding a third dimension ...

Dynamic data placement : **PD2P**

- Subscribe data on demand from jobs
- Run jobs where data is now
- Re-broker jobs when new replicas arrive

⇒ PD2P in place since one year and is working well

PD2P is used for RO-07

New paths between T2s

- New independent T2 network in preparation : **LHCONE**
 - For T2-T1, T2-T2 traffic
- ATLAS decided to set new communication channels for T2 without waiting for LHCONE : **T2Ds**

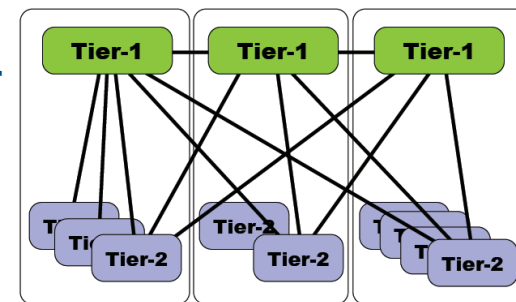
– T2Ds are T2s directly (DDM) connected to T1s of different clouds

⇒ T2Ds can contribute to production in other clouds

⇒ ATLAS will use T2Ds to connect some

T2D : all French sites are T2Ds but GRIF-IRFU, CPPM and Romanian sites

Good network needed



To enable distributed

- Frontier servers have squid installed at sites they are connected to
- Squid are installed at sites
- Data are cached in the Frontier launchpad server and the squid before going to the WN
- => allow to shield Oracle server for overload and decrease significantly access time to condition data for site far away of T1 or with large latency time

Squid server installed at RO-02
=> To serve RO-02, RO-07 and RO-14
and at RO-16

IN2P3-CC):

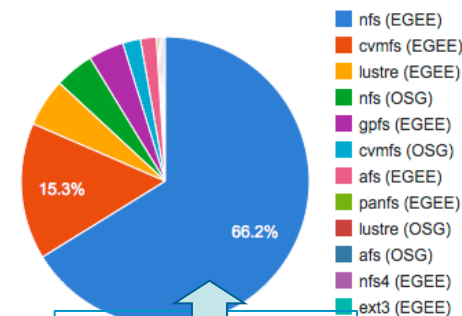
Access to software through CVMFS

- It's a web-based file system which allows caching of release at sites
- => no more need to install software on all sites
- in deployment, see next slide

LFC consolidation at CERN

- All LFC's will be aggregated in one LFC at CERN + one readonly replica

FileSystem Types (active sites)



1 month ago :
now 53%

CVMFS deployment on Fr Cloud

Site	Installation (fraction)	Deployment before Xmass	Deployment after Xmass
RO-16	NO	YES	
RO-02	NO	YES	
RO-07	NO	NO	YES
RO-14	NO	NO	YES
LPC	YES (33%)	YES	
LPSC	NO	YES	
LAPP	YES (100%)		
CPPM	YES (100%)		
BEIJING	NO	YES	
TOKYO	NO	NO	YES
LYON	YES (100%)		
GRIF-LAL	NO	YES	
GRIF-LPNHE	YES (100%)		
GRIF-IRFU	YES (100%)		

Supervision is needed at all stage : there are tools and web pages to monitor the whole system and tests are sent regularly to check its performances

- Sites
 - Site status board
 - SAM/NAGIOS tests
 - Dashboard with history
- Jobs
 - Panda pages
 - Dashboard user jobs monitoring
 - Hammercloud tests => automatic blacklisted of analysis queues
- Data
 - Transfer monitoring (Dashboard)
 - FTS channels
 - SONAR tests

⇒ Particular attention has been paid to improve analysis jobs efficiency (success rate)

Grid monitoring and tests

See <http://adc-monitoring.cern.ch/>

Data Management



Central Deletion Monitoring



Data Replication



Dataset Popularity



Dataset Recovery Service



DDM Dashboard



DDM Dashboard 2.0 (BETA)



Functional Tests



Single File Transfer Monitoring



Storage Accounting



Storage Monitoring

Data Processing



AKTR



Historical Views Dashboard



Historical Views (BETA)



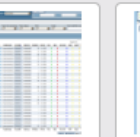
Job Summary Dashboard



PAnDA Monitor Analysis



PAnDA Monitor Production



ProdSys Dashboard (BETA)



User Job Monitoring Dashboard

Sites and Services



AGIS



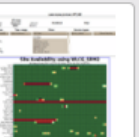
Central Services Availability



Hammercloud



Pilot factory



SAM Visualization



Site Status Board

The system does not work alone, different teams look after it and its users :

- T0 and export : Point 1 T0 (shift)
- User support : DAST (shift)
 - hn-atlas-dist-analysis-help@cern.ch
- Grid production and analysis, Data transfer : ADCOS shifters and ADC experts
- Cloud support : Squad team (shift/expert)
 - Interface between ATLAS and sites support
 - Ensures that all cloud sites functioning is optimal
 - Participate to the cloud operation
- Site support

There is one Squad per cloud

- Squad team is the main interface between ATLAS and the sites of one cloud
 - Don't hesitate to contact them !
- Squad team ensures that all cloud sites functioning is optimal
 - has a closer look than AdCoS to sites of their cloud
 - Look at ATLAS monitoring (panda, DDM,sam tests, ...)
 - works to minimize offline or blacklist periods
 - Prepares for sites schedule downtime
 - Please inform us if you plan a downtime !
 - works to prevent problems
 - Disk space, software installation
 - Manage lost or corrupted files
- Squad team participates to the good operation of the cloud
 - Manage the pilot factory
 - Look at FTS channels

How to contact them ?

atlas-support-cloud-fr@cern.ch

The team

- Emmanuel Le Guirriec (CPPM)
- Irena Nikolic (LPNHE)
- Luc Poggioli (LAL)
- Wenjing Wu (IHEP)
- Sabine Crépe-Renaudin (LPSC)

+ bonus : Éric Lançon, Stéphane Jézéquel, Ikuo Ueda

Shifts

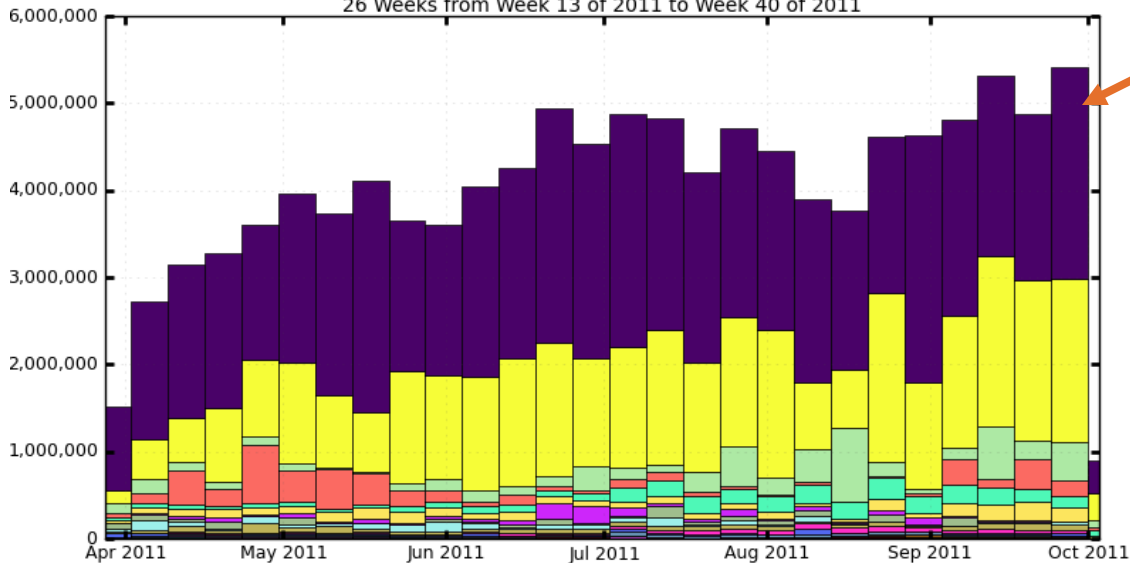
- In general for 1 week Monday-Sunday

Don't hesitate to contact them for any question, comment, suggestion or recrimination about ATLAS distributed computing operation ...

- We need to reinforce the FR Squad team
- Romania has to participate to this effort
 - It will help us to maintain high quality of support to users and sites
 - It will help Romanian sites to have up to date ATLAS information and local expertise
- Participation to Squad can start now
 - Camelia has already been trained to ATLAS monitoring
 - Can start to take shift as shadow with one of us to get expertise
- Squad duties
 - Shift as expert one week Monday-Sunday every 5->6 week
 - Stay up to date with ATLAS : attend CAF meetings, ADC meetings, ATLAS software weeks, stay tuned to our mailing list
 - Take care of some particular tasks that we share in the long term (ie CVMFS installation, factory updates ...)

Grid activities : last 6 months

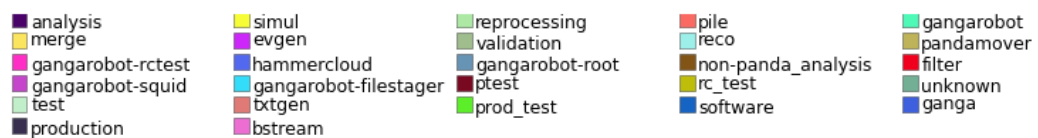
Terminated jobs
26 Weeks from Week 13 of 2011 to Week 40 of 2011



5 M jobs/week !

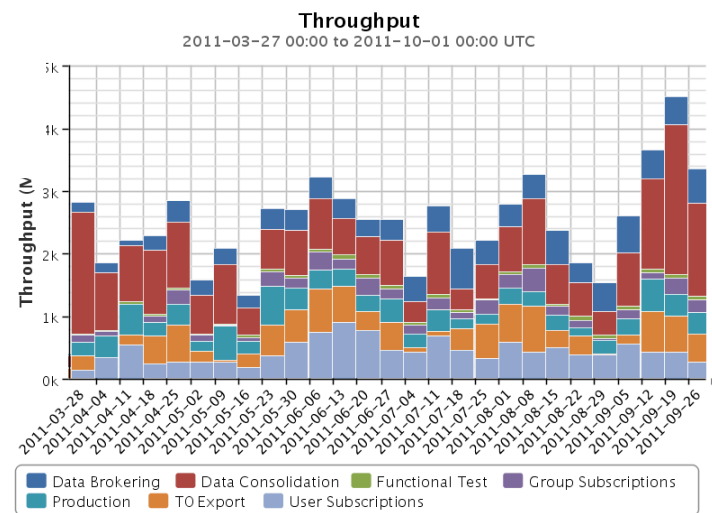
Job number is steadily increasing

- Analysis
- MC Production
- Reprocessing
- ...



Maximum: 5,409,938 , Minimum: 0.00 , Average: 3,875,317 , Current: 895,478

Data transfer and replication



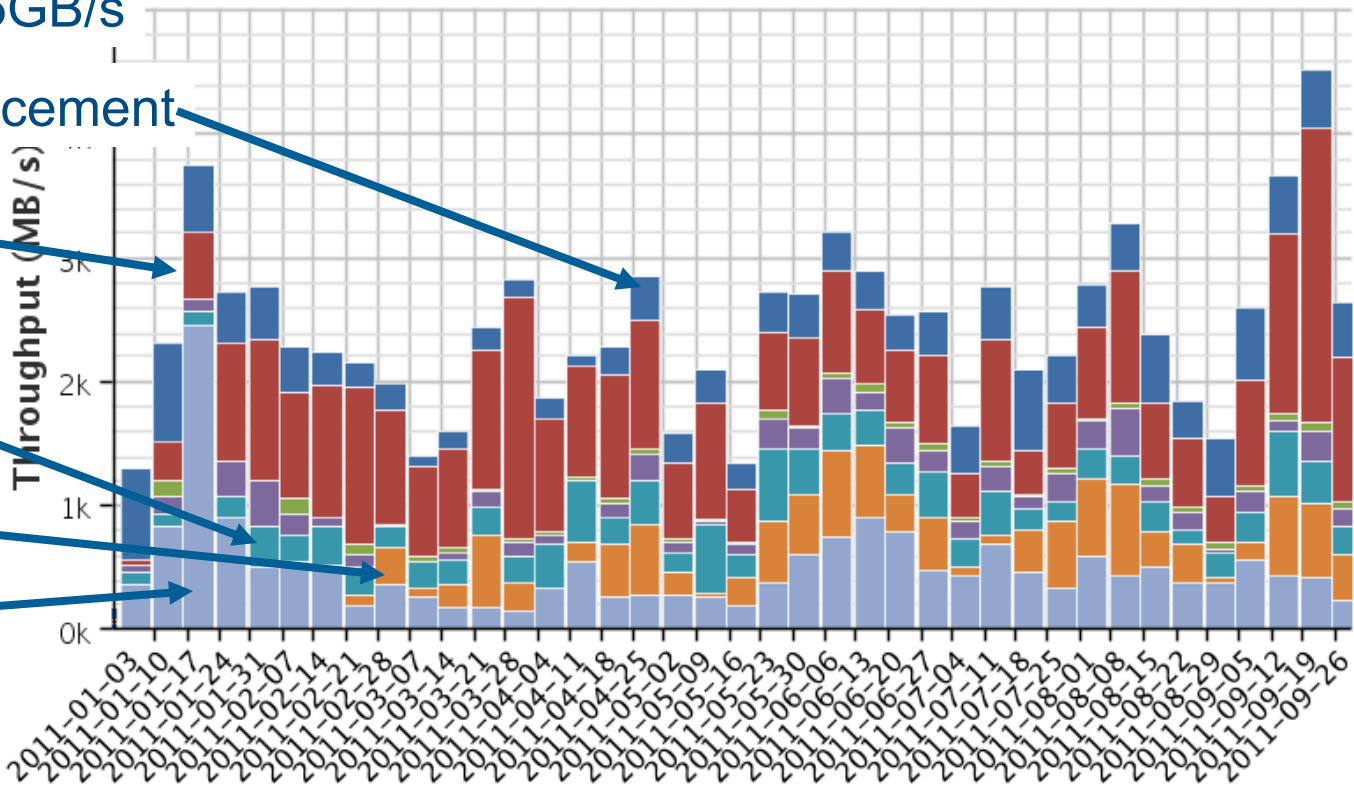
Data distribution on the grid

Efficiency : 93% success rate, first retry always succeeds

Throughput

2011-01-01 00:00 to 2011-09-30 00:00 UTC

5GB/s



Dynamic data placement

Pre-placement

Production

T0 export

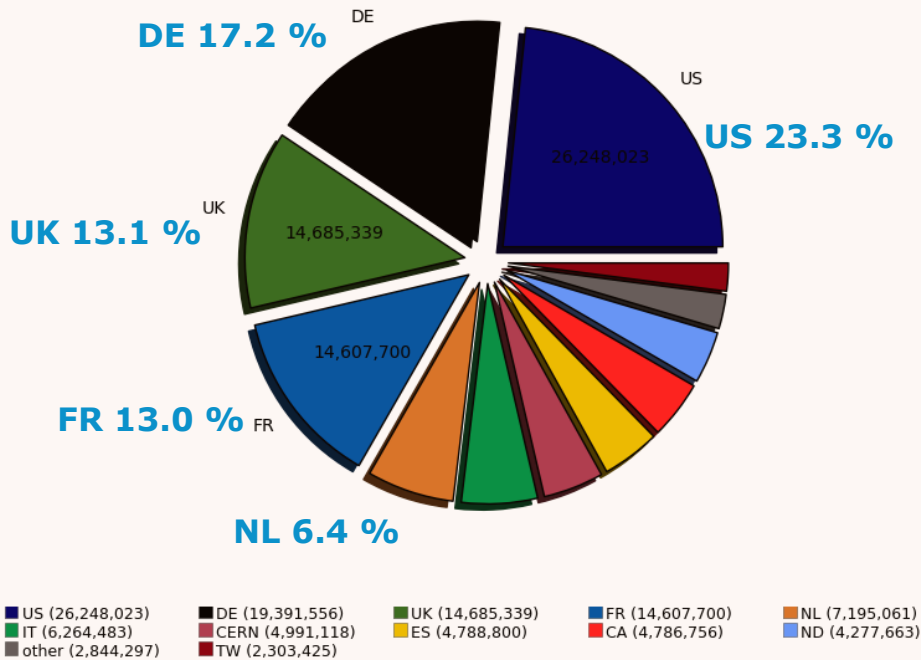
User requests



Activity sharing vs clouds and sites last 6 months

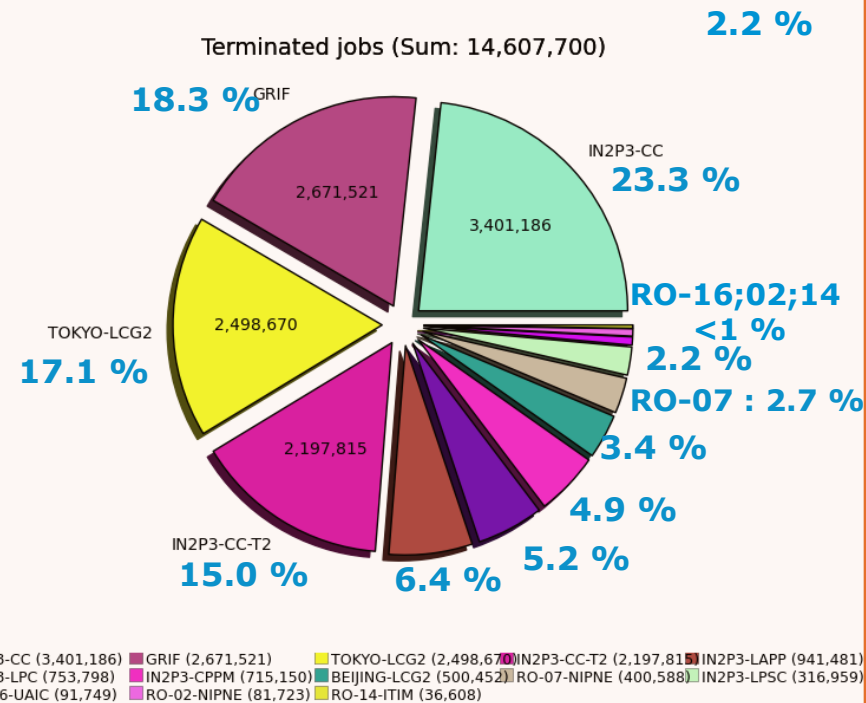
Over the grid
112 millions of jobs
600 000/day

Terminated jobs (Sum: 112,384,221)



On the French Cloud
French cloud = 13%
14.6 millions jobs
77 300 jobs/day

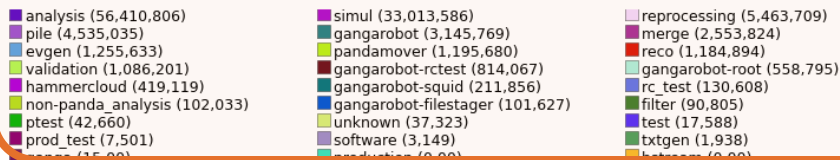
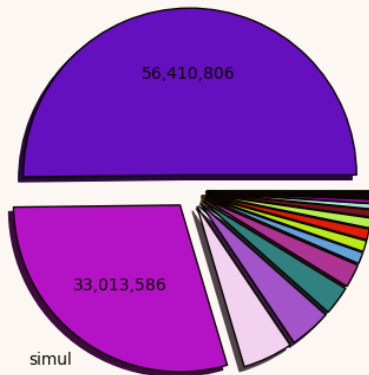
Terminated jobs (Sum: 14,607,700)



Jobs

Analysis = 50.2 %
 Simulation = 29.4 %
 Reprocessing = 4.9 %
 Pile-up = 4.0 %
 Validation = 1 %
 Tests = 2.8 %

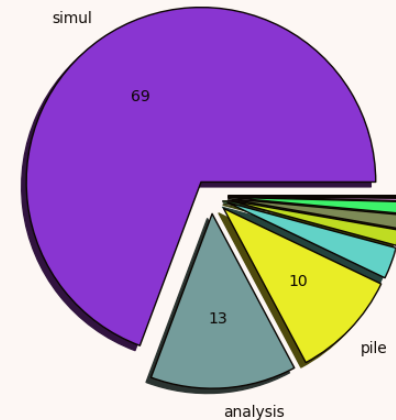
Terminated jobs (Sum: 112,384,221)
analysis



CPU

Analysis = 13.5 %
 Simulation = 69.3 %
 Reprocessing = 1.1 %
 Pile-up = 10.1 %
 Validation = 1.4 %
 Tests = 0.05 %

CPU consumptions Good Jobs (Pie Chart in percentage) (Sum: 99.00)



Production efficiency

Efficiency = ratio successful / all jobs

Average efficiency over all clouds : 92 %, Fr : 90%

The success rate varies within sites

⇒ Sites will be tested regularly and errors analysed to allow improvements

Fr Cloud sites



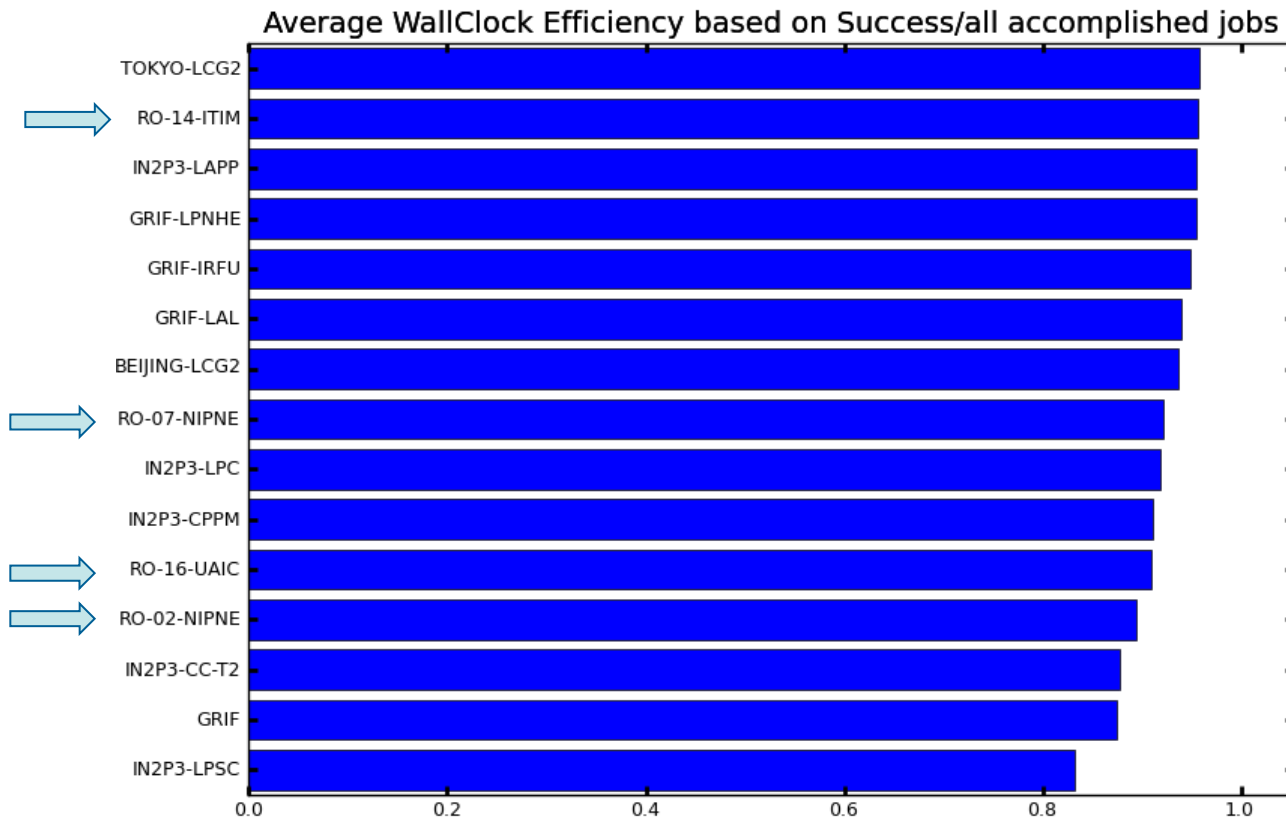
From: 2011-06-01
To: 2011-11-25

Sabine

Wallclock efficiency

T2 all over the grid : 95%

FR :91 %



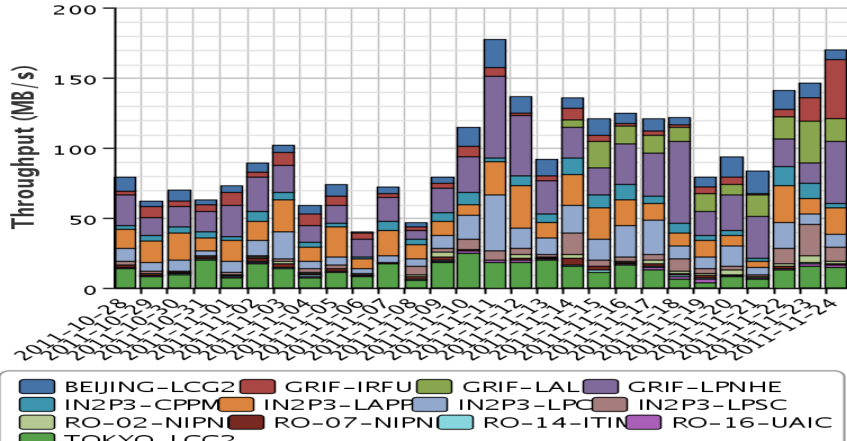
Transfers rate and failures



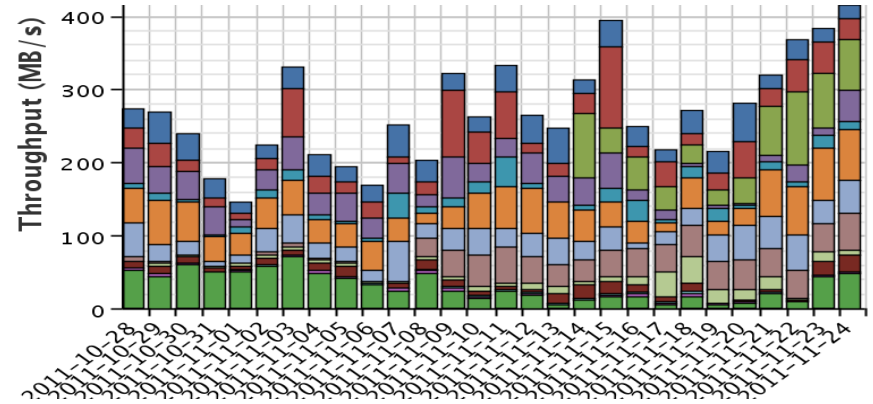
As source

Throughput

2011-10-28 00:00 to 2011-11-25 00:00 UTC

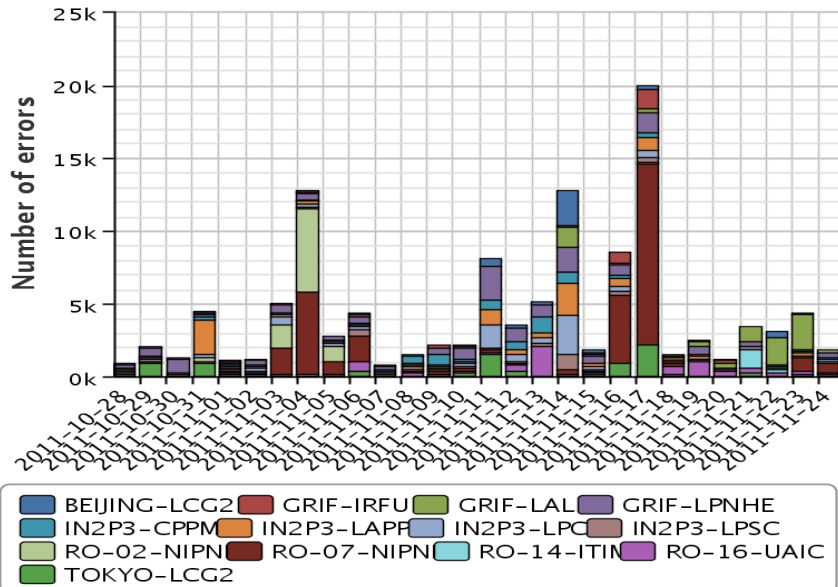


As destination



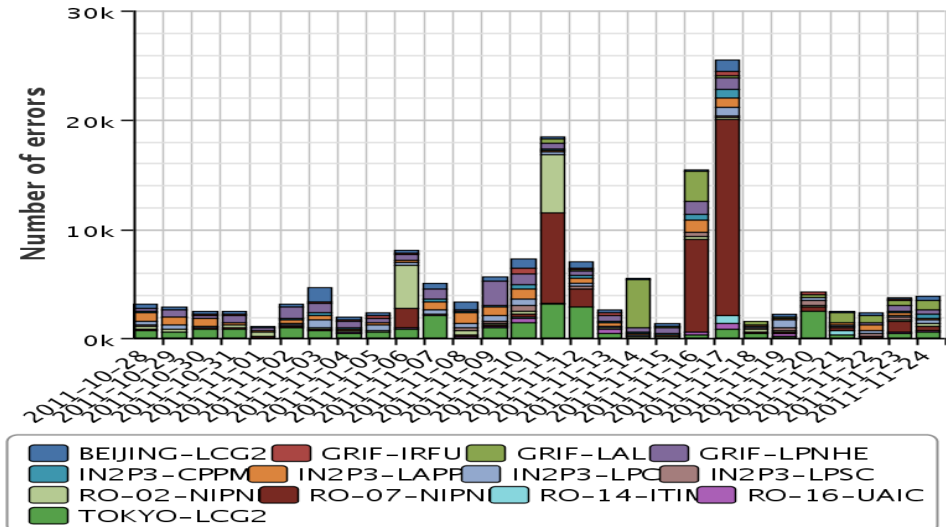
Transfer Failures

2011-10-28 00:00 to 2011-11-25 00:00 UTC



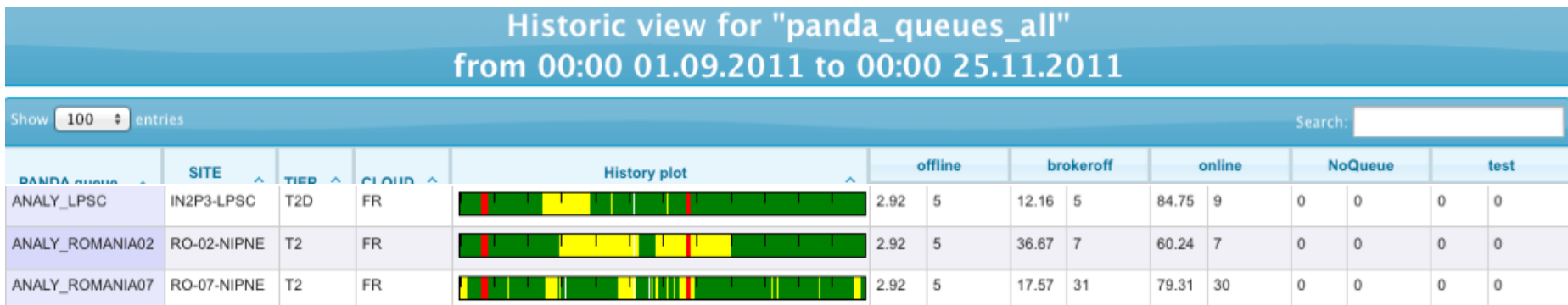
Transfer Failures

2011-10-28 00:00 to 2011-11-25 00:00 UTC



Hammercloud tests

Aim : send test jobs to analysis queues to prevent users sending jobs to non functioning sites



Need to be improved

- ⇒ Still some period due to central ATLAS problem (DB distribution ex LPSC)
- ⇒ Problems due to T1 (Cf common period, removed)
- ⇒ Measure reliability of sites
- ⇒ In discussion : could be used to distribute data to sites

- HC tests could be used to define the distribution policy among T2s: Preference to
 - Reliable sites for analysis (fraction of time with analysis queue online)
 - Well connected sites (T2Ds) to transfer datasets
- Based on site reliability for analysis estimated from 'brokeroff' periods in the previous month
 - 4 groups :
 - Alpha (60% share) : T2Ds with $> 90\%$ reliability
 - Bravo (30 % share) : non-T2Ds with $> 90\%$ reliability
 - Charlie (10 % share) : Any T2 with $80\% < \text{reliability} < 90\%$
 - Delta (0% share) : Any T2 with reliability $< 80\%$

T2 grouping with Oct stats

alpha (21): 62%

AGLT2
CA-SCINET-T2
CA-VICTORIA-WESTGRID-T2
CSCS-LCG2
DESY-HH
DESY-ZN
GRIF-LAL
IFAE
IFIC-LCG2
IN2P3-LAPP
IN2P3-LPSC
INFN-NAPOLI-ATLAS
INFN-ROMA1
MWT2
SLACXRD
SWT2_CPB
UAM-LCG2
UKI-LT2-QMUL
UKI-SCOTGRID-ECDF
UKI-SCOTGRID-GLASGOW
UNI-FREIBURG

bravo (18): 26%

AUSTRALIA-ATLAS
CA-ALBERTA-WESTGRID-T2
CYFRONET-LCG2
GRIF-IRFU
IL-TAU-HEP
IN2P3-CPPM
JINR-LCG2
LIP-LISBON
OU_OCHEP_SWT2
PRAGUELCG2
RU-PROTVINO-IHEP
SFU-LCG2
TECHNION-HEP
UKI-LT2-RHUL
UKI-LT2-UCL-HEP
UKI-NORTHGRID-LIV-HEP
UKI-NORTHGRID-SHEF-HEP
UKI-SOUTHGRID-CAM-HEP

charlie (17): 12%

BEIJING-LCG2
GOEGRID
GRIF-LPNHE
IN2P3-LPC
INFN-FRASCATI
INFN-MILANO-ATLASC
LRZ-LMU
NCG-INGRID-PT
RO-02-NIPNE
RO-07-NIPNE
TOKYO-LCG2
UKI-NORTHGRID-LANCS-HEP
UKI-NORTHGRID-MAN-HEP
UKI-SOUTHGRID-OX-HEP
UKI-SOUTHGRID-RALPP
WEIZMANN-LCG2
WUPPERTALPROD

delta (9): 0%

CA-MCGILL-CLUMEQ-T2
HEPHY-UIBK
LIP-COIMBRA
MPPMU
NET2
RRC-KI
RU-PNPI
TR-10-ULAKBIM
UKI-SOUTHGRID-BHAM-HEP

Grid computing is working well

- Data are (re)processed and distributed
- Data are analysed
- Support to our large community is provided

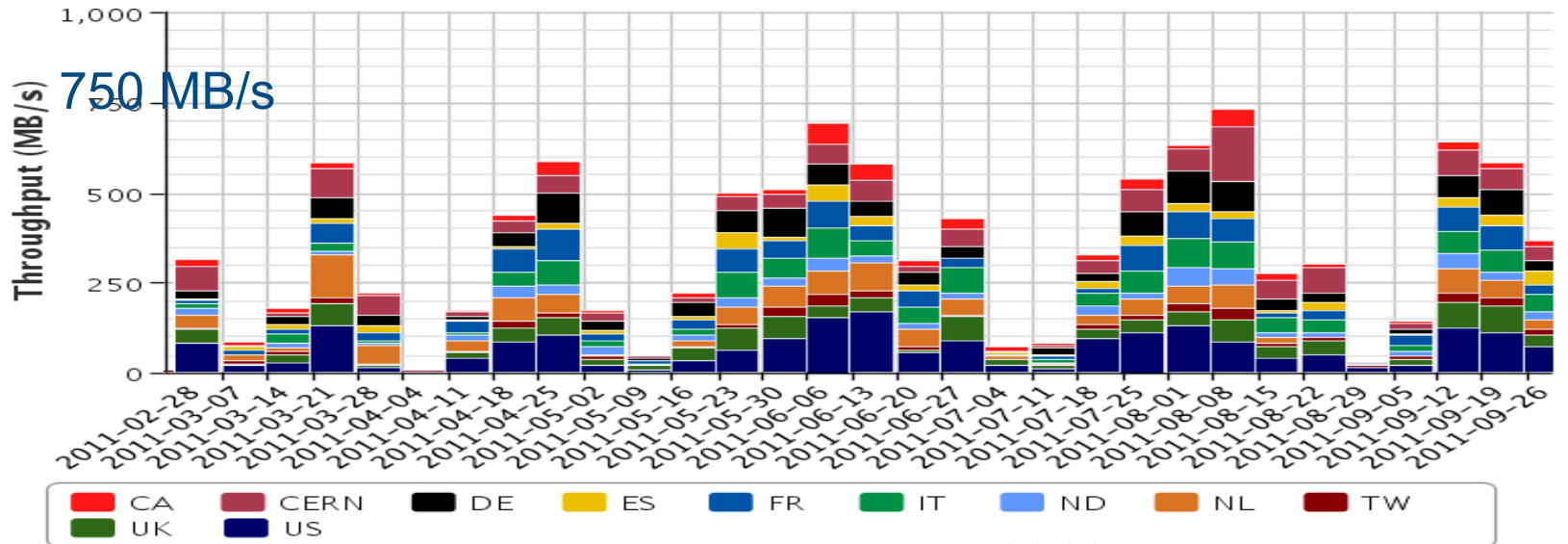
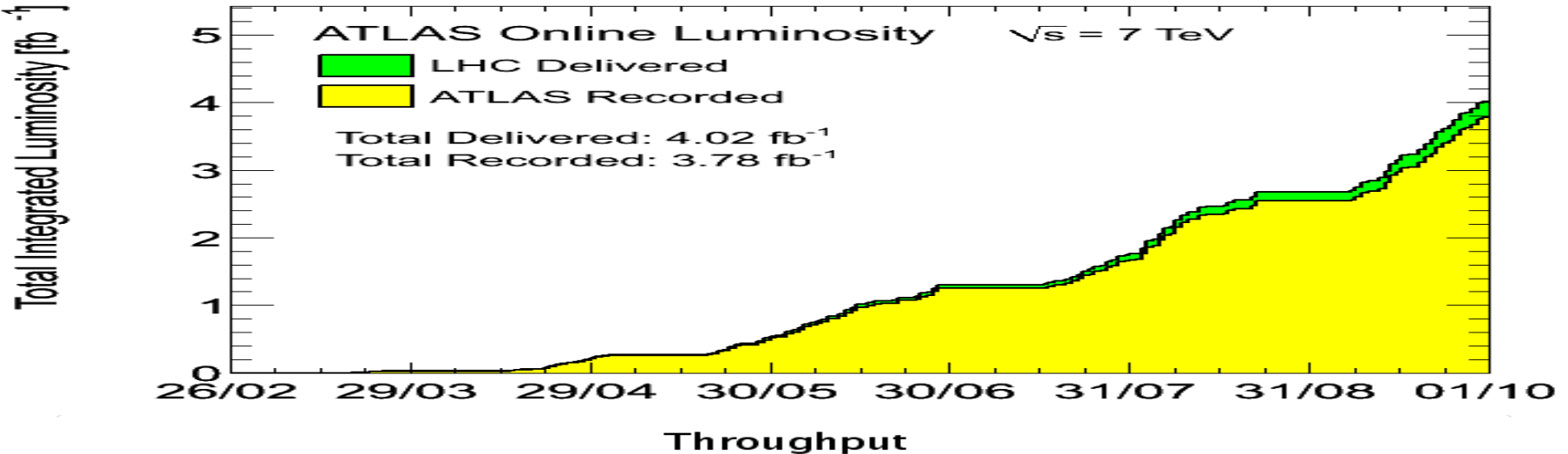
ADC world is changing permanently to improve efficiency

- Cloud boundaries have been released to benefit from good network performance
- Data distribution is closer to user needs
- Steadily work to improve sites and grid reliability, analysis efficiency and user life

=> Needs work at all level : sites, ATLAS supports for clouds and central activities

Your help is more than welcome ...

2011 data taking: T0 export



Raw data compression since July 2011

FR Dashboard



<http://happyface-goegrid.gwdg.de/cloudmon/fr/>

Cloud Monitoring 2.0

Monday, 03 October, 01:45:02 (CEST)

Home Storage Usage Cloud software Help

Cloud: FR	Panda Analysis		Panda Production		ATLAS SAM		OPS SAM Daily Month Year			Availability			Site Exclusion		GGUS		Downtimes		Ganga Robot		SW
Site	Run	Eff%	Run	Eff%	CE	SRMv2	CE	sBDII	SRMv2	Day	Oct	Sep	Panda Status	DDM Status	Open	Closed	Now	Soon	PANDA	LCG	
BEIJING-LCG2	100	95	3	100	ok	ok	ok	ok	ok	100	100	100	online	online	0	0	UP	No	ok	n/a	F
GRIF-LPNHE	116	89	20	100	ok	ok	ok	ok	ok	100	online	online	0	0	UP	No	ok	n/a	E		
GRIF-LAL	1	23	26	0	ok	ok	ok	ok	ok	online	online	0	0	2011-09-29 18:00:00 CEST	No	ok	n/a	F			
GRIF-IRFU	74	100	0	0	ok	ok	ok	ok	ok	test	online	0	0	UP	No	ok	n/a	F			
IN2P3-CC	124	9	108	91	ok	ok	ok	ok	ok	online	blacklisted	3	0	UP	2011-10-03 20:00:00 CEST	ok	n/a	F			
IN2P3-CC-T2	49	48	27	98	F	ok	ok	ok	ok	online	online	0	0	UP	No	NA	NA	E			
IN2P3-CPPM	399	99	3	100	ok	ok	ok	ok	ok	online	online	0	0	UP	No	ok	n/a	F			
IN2P3-LAPP	600	96	15	56	ok	ok	ok	ok	ok	online	online	0	0	UP	No	ok	n/a	F			
IN2P3-LPC	1	0	0	0	ok	F	F	ok	F	offline	blacklisted	74684	0	2011-09-29 19:00:00 CEST	No	n/a	n/a	F			
IN2P3-LPSC	119	100	36	0	ok	ok	ok	ok	ok	online	online	0	0	UP	No	NA	NA	F			
RO-02-NIPNE	5	98	0	0	ok	ok	ok	ok	ok	online	online	0	0	UP	No	at risk	n/a	F			
RO-07-NIPNE	175	88	42	98	ok	ok	ok	ok	ok	online	online	2	0	UP	No	ok	n/a	F			
RO-14-ITIM	NA	NA	2	100	ok	F	ok	ok	ok	online	online	0	0	UP	No	no-test	no-test	F			
RO-16-UAIC	NA	NA	0	0	ok	ok	ok	ok	ok	online	online	0	0	UP	No	no-test	no-test	F			
TOKYO-LCG2	1053	80	76	100	ok	ok	F	F	F	online	online	73623	0	UP	No	n/a	n/a	F			

Storage Usage & Data Transfers

Site	dCache/DPM	SPACE TOKENS (TB)					Data Transfer				
		Version	DATADISK	GROUPDISK	PRODDISK	SCRATCHDISK	LOCALGROUPDISK	DATADISK	GROUPDISK	PRODDISK	SCRATCHDISK
BEIJING-LCG2	1.8.0	138/224	0/1	1/7	3/5	2/53	100.0	0.0	100.0	0.0	0.0
GRIF-LPNHE	1.8.0	155/166	37/54	4/16	6/9	19/21	99.8	0.0	91.7	0.0	0.0
GRIF-LAL	1.8.0	155/166	37/54	4/16	6/9	19/21	0.0	0.0	100.0	0.0	0.0



Sonar view



Site Status for the ATLAS sites

v0.10.0_rc62 Login Found a bug? HELP

Index **Expanded Table** Gridmap Alternative views Admin ATLAS SSB Instructions Ask the ATLAS SSB people!

Here search on IN2P3-LPC

Show 200 entries Copy Print Save view: Sonar Search...

Site Name	SrcSite	SrcCloud	SrcTier	DstSite	DstCloud	DstTier	AvgBRS(MB/s)	EvS	AvgBRM(MB/s)	EvM	AvgBRL(MB/s)	EvL	Prio
AGLT2 to IN2P3-LPC	AGLT2	US	T2D	IN2P3-LPC	FR	T2D	0.57+/-0.03	10	3.32+/-1.07	10	13.09+/-6.29	10	5
AUSTRALIA-ATLAS to IN2P3-LPC	AUSTRALIA-ATLAS	CA	T2	IN2P3-LPC	FR	T2D	0.39+/-0.01	10	2.51+/-0.81	10	0.00+/-0.00	0	2
BEIJING-LCG2 to IN2P3-LPC	BEIJING-LCG2	FR	T2D	IN2P3-LPC	FR	T2D	0.41+/-0.11	11	1.26+/-0.49	53	1.33+/-0.11	5	6
BNL-OSG2 to IN2P3-LPC	BNL-OSG2	US	T1	IN2P3-LPC	FR	T2D	0.43+/-0.20	10	4.98+/-0.51	10	16.91+/-3.84	10	7
CA-ALBERTA-WESTGRID-T2 to IN2P3-LPC	CA-ALBERTA-WESTGRID-T2	CA	T2	IN2P3-LPC	FR	T2D	0.48+/-0.03	10	3.28+/-0.39	10	0.00+/-0.00	0	2
CA-SCINET-T2 to IN2P3-LPC	CA-SCINET-T2	CA	T2D	IN2P3-LPC	FR	T2D	0.55+/-0.10	10	3.34+/-1.74	10	2.96+/-2.24	10	5
CA-VICTORIA-WESTGRID-T2 to IN2P3-LPC	CA-VICTORIA-WESTGRID-T2	CA	T2D	IN2P3-LPC	FR	T2D	0.55+/-0.09	10	4.41+/-1.23	10	15.42+/-3.36	10	5
CERN-PROD to IN2P3-LPC	CERN-PROD	CERN	T0	IN2P3-LPC	FR	T2D	0.77+/-0.17	10	10.17+/-2.82	119	37.93+/-10.84	289	7
CSCS-LCG2 to IN2P3-LPC	CSCS-LCG2	DE	T2	IN2P3-LPC	FR	T2D	0.89+/-0.06	10	5.87+/-1.46	5	0.00+/-0.00	0	2
CSTCDIE to IN2P3-LPC	CSTCDIE	NL	T3	IN2P3-LPC	FR	T2D	0.83+/-0.08	10	6.85+/-0.10	5	0.00+/-0.00	0	0
CYFRONET-LCG2 to IN2P3-LPC	CYFRONET-LCG2	DE	T2	IN2P3-LPC	FR	T2D	0.66+/-0.09	10	2.19+/-0.17	5	0.00+/-0.00	0	2
DESY-HH to IN2P3-LPC	DESY-HH	DE	T2D	IN2P3-LPC	FR	T2D	0.88+/-0.04	10	7.65+/-1.26	5	23.78+/-7.48	10	5
DESY-ZN to IN2P3-LPC	DESY-ZN	DE	T2D	IN2P3-LPC	FR	T2D	0.45+/-0.13	10	4.17+/-1.13	5	33.97+/-7.49	10	5
FZK-LCG2 to IN2P3-LPC	FZK-LCG2	DE	T1	IN2P3-LPC	FR	T2D	0.67+/-0.20	10	2.82+/-1.56	5	7.63+/-5.64	10	7
GOEGRID to IN2P3-LPC	GOEGRID	DE	T2	IN2P3-LPC	FR	T2D	0.68+/-0.04	10	4.39+/-1.65	5	0.00+/-0.00	0	2
GRIF-IRFU to IN2P3-LPC	GRIF-IRFU	FR	T2	IN2P3-LPC	FR	T2D	0.69+/-0.06	15	4.30+/-0.41	5	0.00+/-0.00	0	6
GRIF-LAL to IN2P3-LPC	GRIF-LAL	FR	T2D	IN2P3-LPC	FR	T2D	1.44+/-1.24	17	7.30+/-2.92	1705	45.37+/-24.95	18	6
GRIF-LPNHE to IN2P3-LPC	GRIF-LPNHE	FR	T2D	IN2P3-LPC	FR	T2D	0.95+/-0.10	12	5.87+/-3.06	240	50.35+/-31.68	74	6



At 1rst of September 2011

- IT INFN-MILANO-ATLASC
- IT INFN-ROMA1
- IT INFN-NAPOLI

- UK UKI-LT2-QMUL_DATADISK
- UK UKI-NORTHGRID-LANCS-HEP
- UK UKI-NORTHGRID-MAN-HEP
- UK UKI-SCOTGRID-GLASGOW
- **UK UKI-SOUTHGRID-OX-HEP**
- **UK UKI-SCOTGRID-ECDF**

- FR GRIF-LAL
- FR GRIF-LPNHE
- FR TOKYO-LCG2
- **FR IN2P3-LAPP**
- **FR BEIJING-LCG2**
- **FR IN2P3-LPSC**
- **FR IN2P3-LPC**

- DE DESY-HH
- DE DESY-ZN
- DE LRZ-LMU
- DE MPPMU
- **DE WUPPERTALPROD**
- **DE CSCS-LCG2**
- **DE GOEGRIDDE UNI-FREIBURG**

- ES IFIC-LCG2
- ES IFAE
- ES UAM-LCG2

- US NET2
- US SWT2_CPB
- US MWT2
- US SLACXRD
- US AGLT2

- **CA CA-VICTORIA-WESTGRID-T2**
- **CA CA-SCINET-T2.**

Frequency

- Triggered on Monday
- Stopped and cleaned on Thursday
- Check the Sonar results on Friday

Tests consists of 15 FTS transfers

- 5 small files: 20 MB each (def = 0-100MB)
- 5 medium files: 200 MB each (def = 100MB-1GB)
- 5 large files: 2 GB each (def = 1GB-∞)

Criteria

Average ByteRate

SMALL	<0.05MB/s	<0.1MB/s	≥0.1MB/s
MEDIUM	<1MB/s	<2MB/s	≥2MB/s
LARGE	<10MB/s	<15MB/s	≥15MB/s

Number of files transferred

SMALL	≤3	4	≥5
MEDIUM	≤2	3	≥4
LARGE	≤1	2	≥3

Tests results

<http://dashb-atlas-ssb.cern.ch/dashboard/request.py/siteview?view=Sonar>