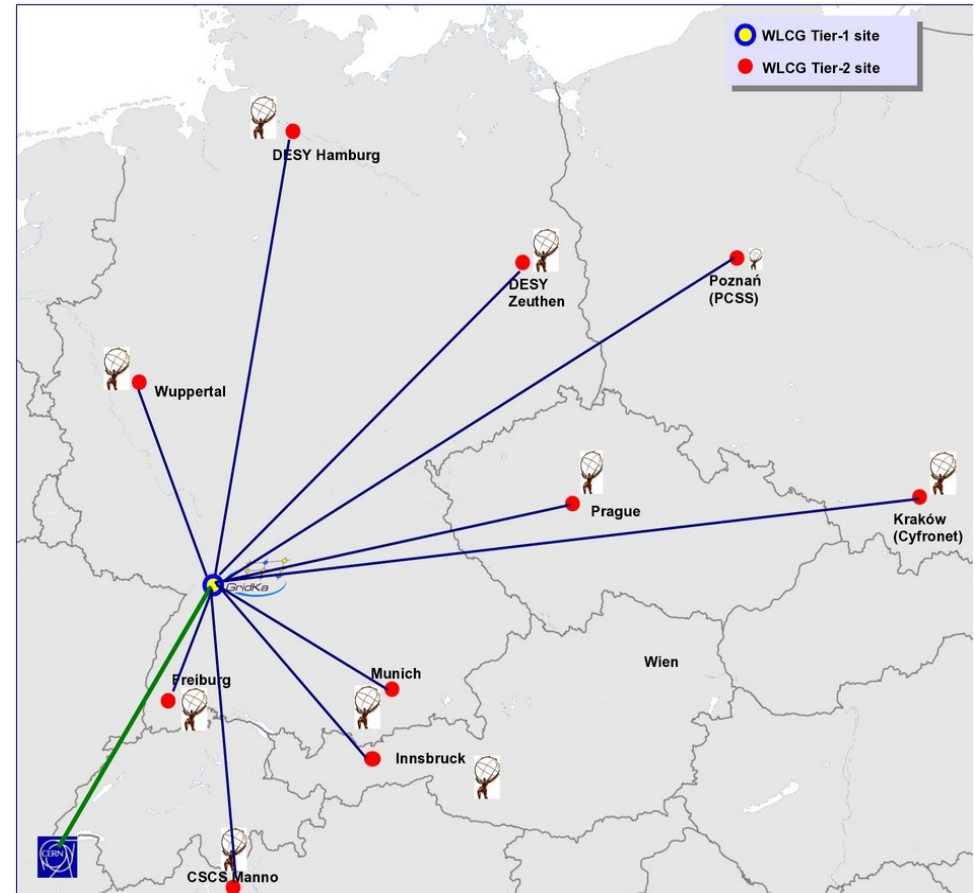


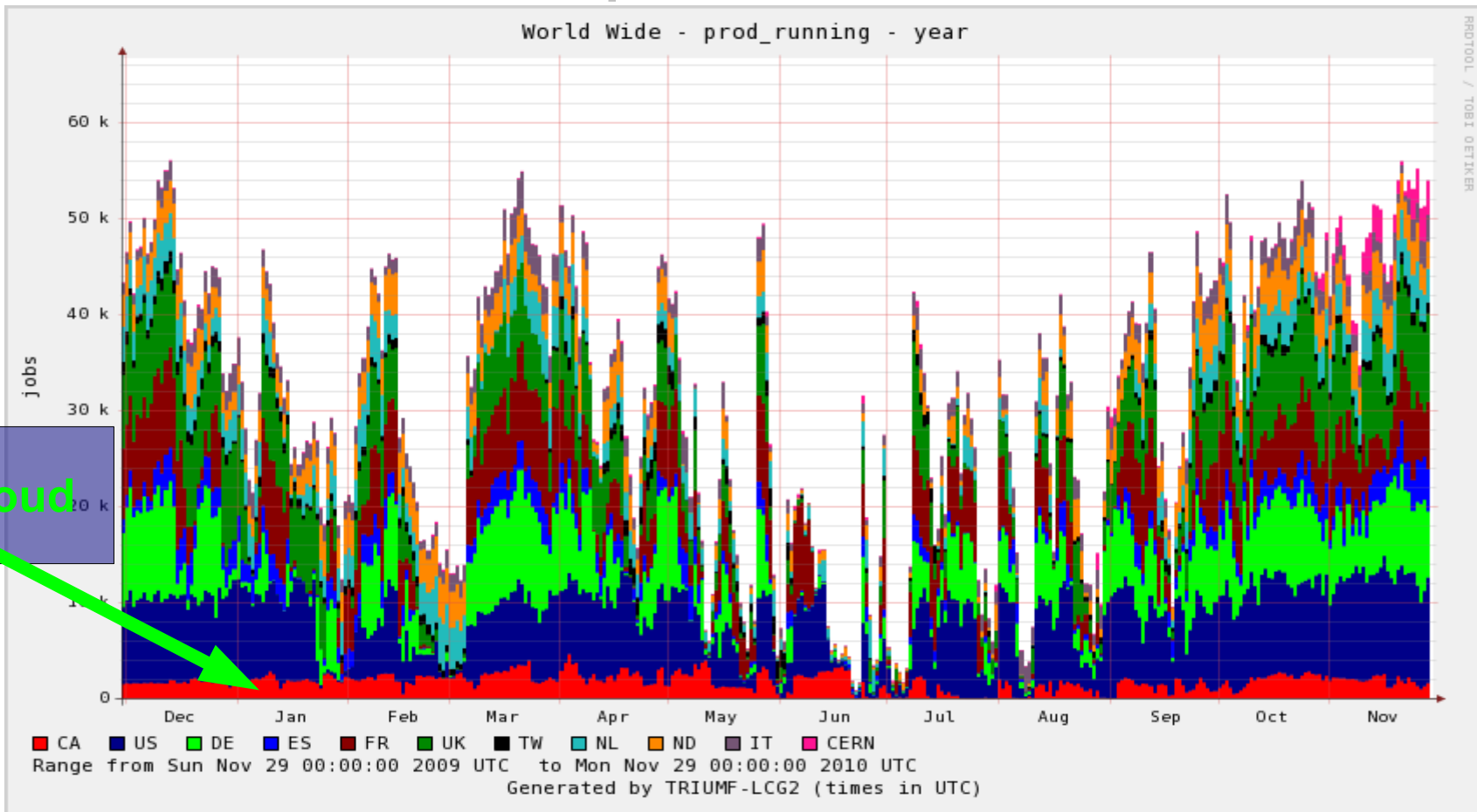
ATLAS-D Tier-2 Report

Dec 01, 2010
HGF-Grid-PB
Dresden

- ATLAS Production and Analysis
- Data distribution
- Resources
- Organization
- Problems and Issues

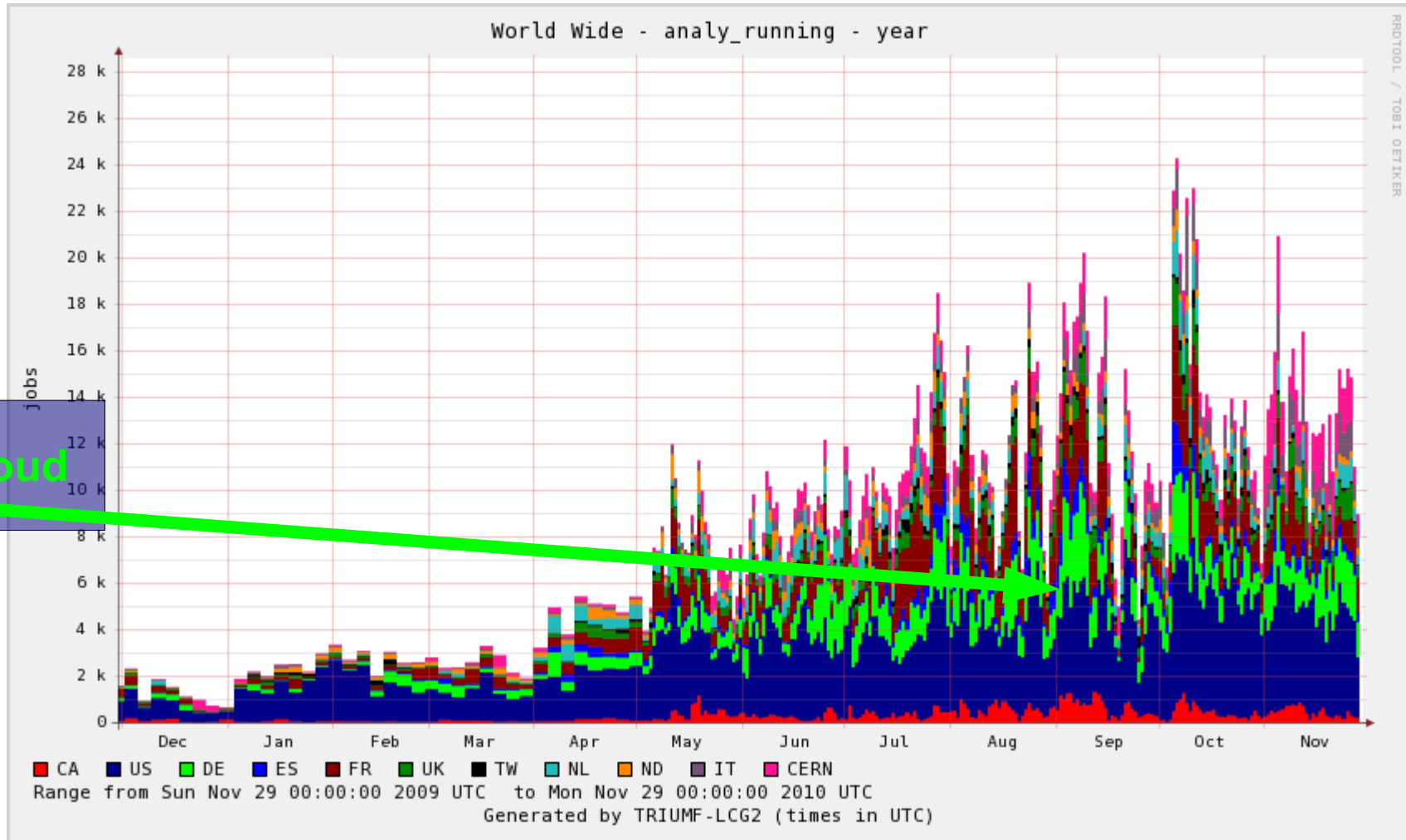


ATLAS worldwide production last 12 months



- up to 50k jobs running simultaneously
- DE fraction ~13%

ATLAS worldwide analysis last 12 months



- strong increase of analysis jobs since LHC start, up to 25k jobs

What is “analysis” in this context?

- For central computing operations “analysis” is any job submitted by ATLAS users (without production role) into the ATLAS Grid
 - job-submission with /atlas/de voms-group allows usage of privileged, non-pledged resources. In use at GridKa and Desy
- Jobs can be
 - ESD/AOD → DPD skimming/slimming
 - ESD/AOD → histo full analysis
 - DPD/ntuple analysis
 - but also MC/G4 private production ...
 - or any executable script
- Much more analysis activity on T3 or local sites
 - no central monitoring, not included in job statistics

DE cloud production statistics

- dashboard statistics DE cloud Jan 1 – Oct 7 2010 (~280 d)
 - Nominal 50% T2 site has 3440 HS06
 - With 8 HS06 per core this corresponds to ~ 1 e10 WallT sec
 - Take ¼ of this (=2.5e9 WallT secs) as ballpark target to account for periods w/o production jobs and analysis activity
 - → HGF T2s easily pass nominal production share

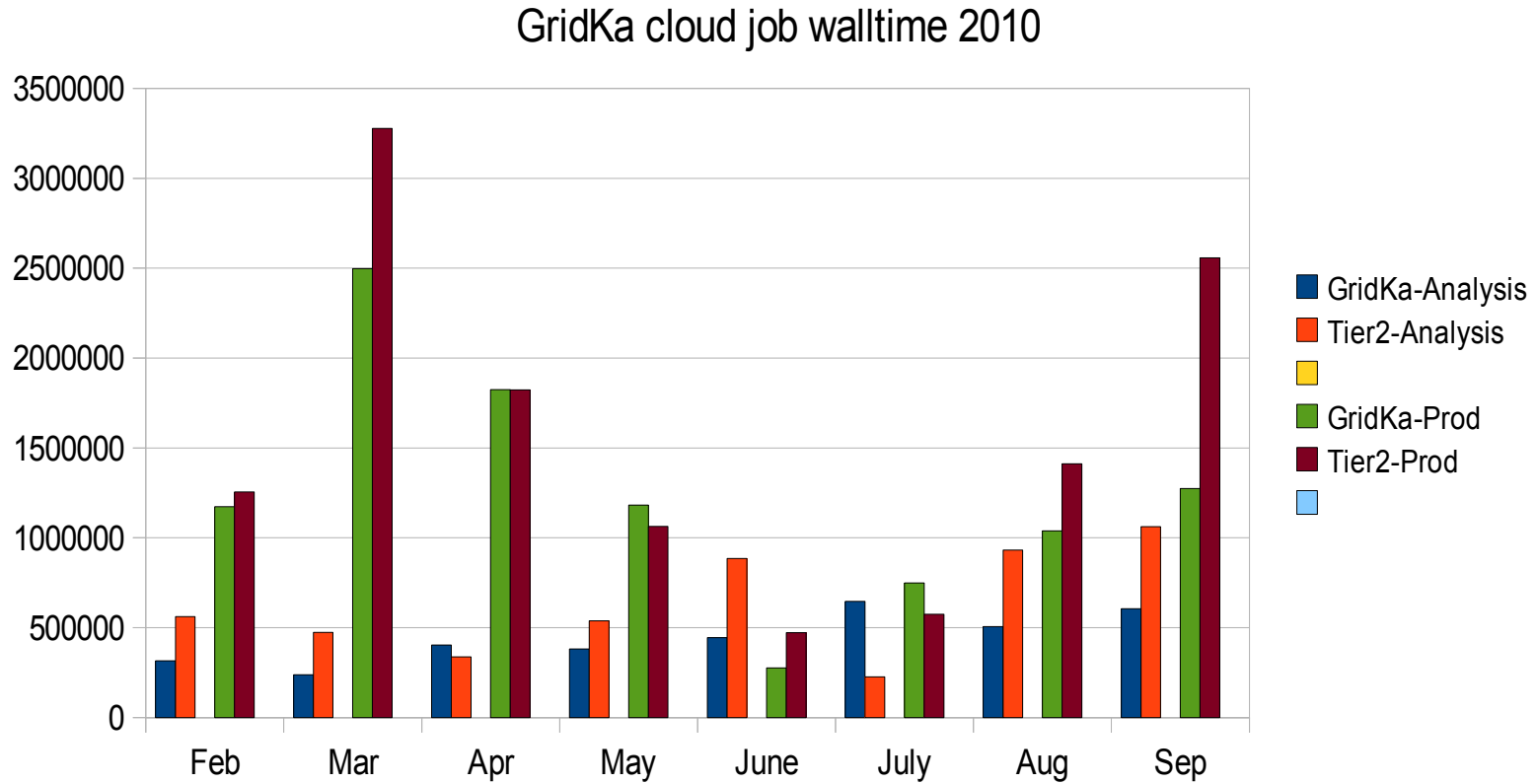
site	success	failure	success (walltime)	failure (walltime)	efficiency	efficiency (walltime)
FZK-LCG2	1971447	170303	31034854712	2034237635	92%	93.8%
DESY-HH	658856	24544	11671985254	342537775	96.4%	97.1%
LRZ-LMU	520771	24626	9101852646	192099577	95.5%	97.9%
wuppertalprod	279565	20735	5373121293	258986855	93.1%	95.4%
DESY-ZN	250404	6461	4586912611	110449758	97.5%	97.6%
GoeGrid	209532	22580	3227165327	857150849	90.3%	79%
MPPMU	216829	6724	3730286710	134424376	97%	96.5%
praguelcg2	132376	9416	2366719052	105772497	93.4%	95.7%
UNI-DORTMUND	88586	19205	1480806399	125441039	82.2%	92.2%
HEPHY-UIBK	81068	2892	1339837101	28560943	96.6%	97.9%
CSCS-LCG2	66460	3398	1378614071	42628869	95.1%	97%
PSNC	38181	6658	666196224	41946439	85.2%	94.1%
UNI-FREIBURG	27171	1522	477861410	70781987	94.7%	87.1%
CYFRONE/LCG2	14671	4219	346284286	48828036	77.7%	87.6%
total	4555917	323283	7.6782497096e+10	4.393846635e+09	93.4%	94.6%

12/02/10

Guenter Duckeck, LMU

Known problem with dashboard for FR, numbers for FR are wrong

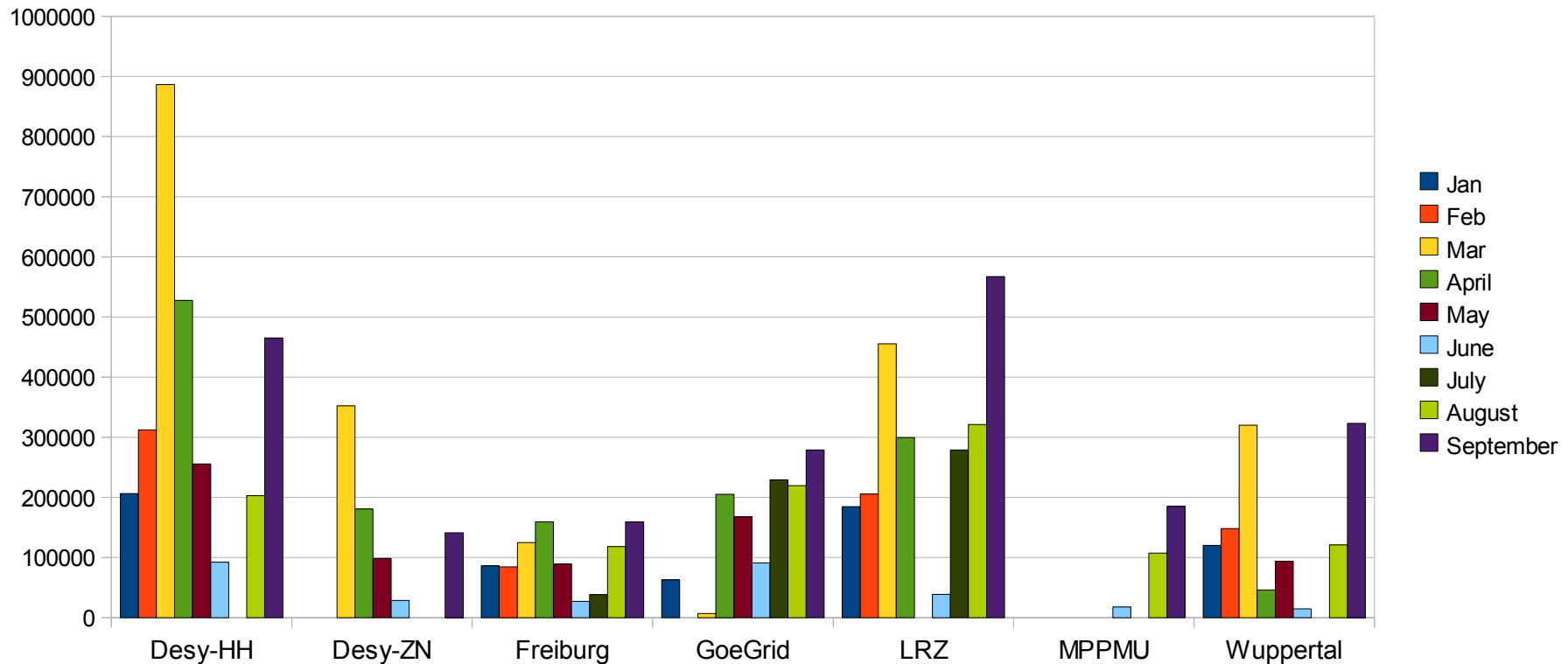
GridKa cloud job statistics - overall Jan-Sep 2010



- fluctuations throughout year
- stable sharing between T1 and T2 sites (40/60)

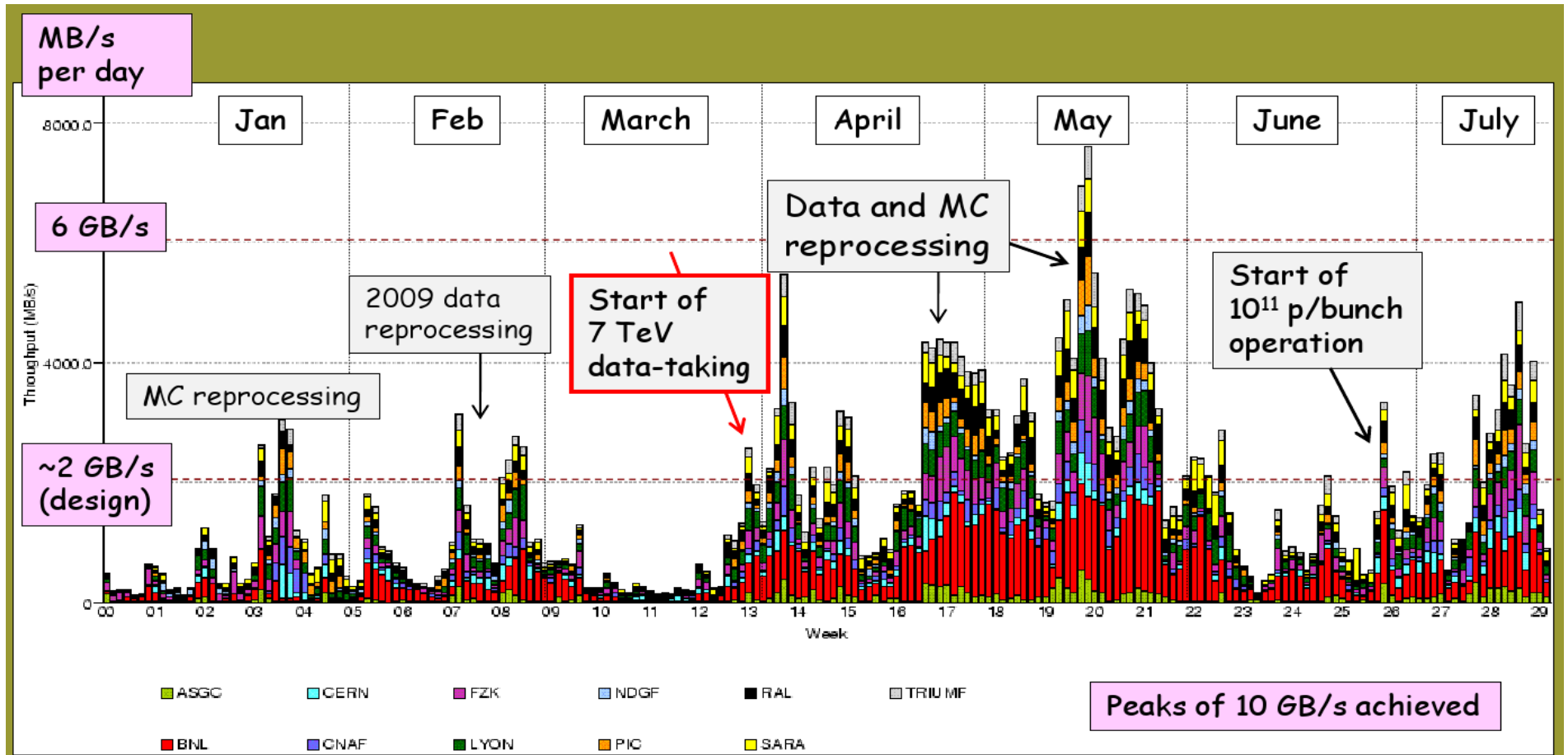
ATLAS-DE T2 production Jan-Sep 2010

ATLAS-DE T2 Produktion Jan-Sep10



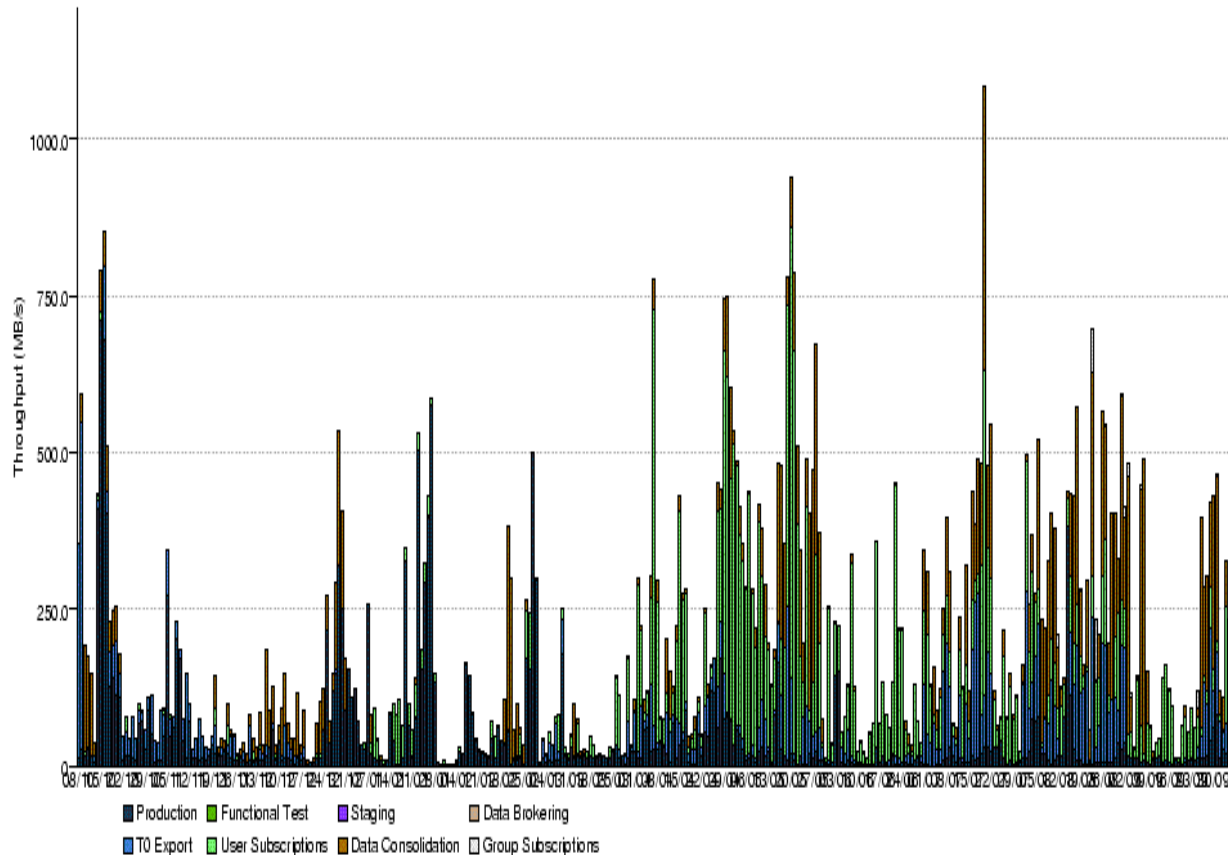
- monthly manual collection from site batch statistics
 - to cross-check Grid accounting and include local usage
 - not all months filled from all sites

ATLAS data transfers Jan-July 2010



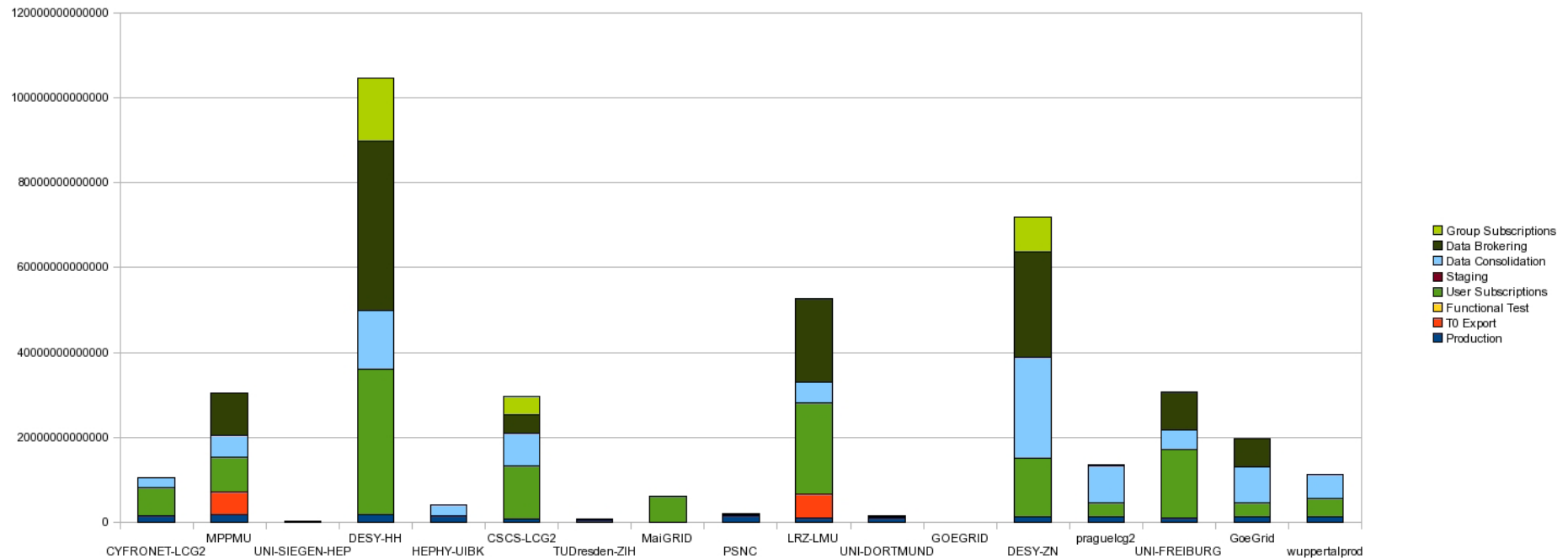
- data distribution generally working well
- substantial headroom

GridKa cloud data distribution overall Sep09-Sep10



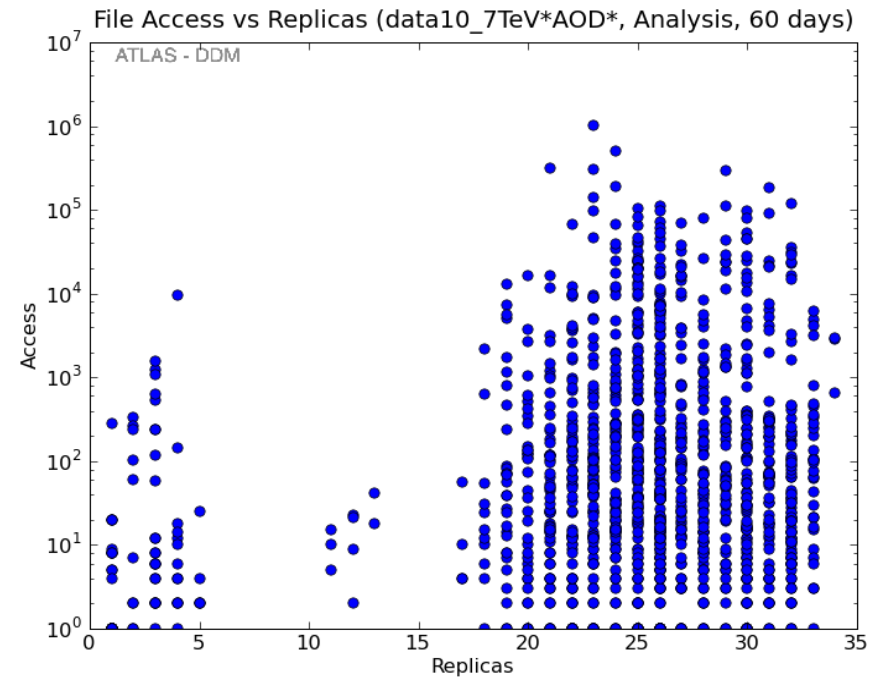
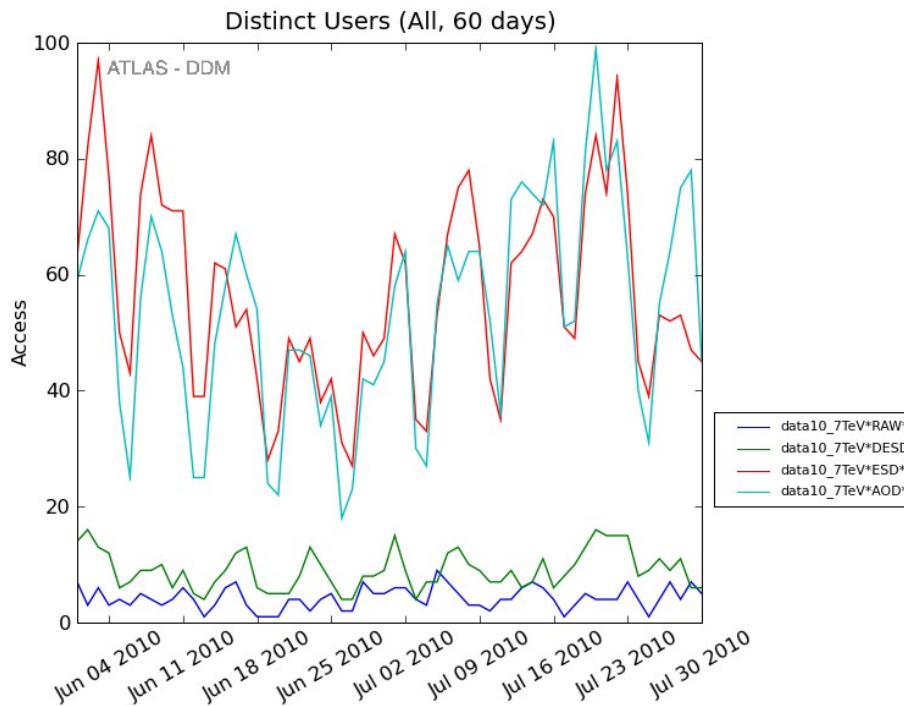
- Since the start of LHC both data volume and traffic increased substantially
- Overall >4 PB data are stored in the cloud
- Peak rates exceeded 1 GB/s
- replication of user data contributes significantly
- data loss at sites due to disk server failures happen quite often
- automatic procedures in place for quick recovery (if data available elsewhere on the Grid)

GridKa cloud GridKa ↔ T2 data flow Oct 10



- About 2 M Files transferred GridKa ↔ T2s (Oct 10)
 - average transfer-rate 180 MB/s
 - > 600 MB/s on some days
- Dynamic Data distribution (PD2P) in operation since Oct:
 - Datasets get replicated from GridKa to Tier-2 Sites in case corresponding analysis jobs got to GridKa (mostly ESD data format)

Computing model vs reality



- Dataset popularity:
 - ESD as AOD, DESD much less
 - model: user analysis mostly AOD/DESD/DPD
- huge variation of access frequency between different data-sets
 - substantial fraction of distributed data very rarely used

Computing model vs reality -2

- Distribution of analysis jobs over sites/clouds
 - preference for huge T1 sites (ESD access, processing of large dataset containers)
 - often long queues at T1 while T2s not filled
 - Model: most analysis jobs at T2s, T1 only exceptionally
- Current static predefined data distribution on T2s (push model) considered non-optimal
 - more dynamic on-demand distribution (pull) under discussion/test
 - hierarchical cloud model vs dynamic peer-to-peer
 - T2 storage as a dynamic cache
 - might require substantial changes in computing model and resources
 - in particular network bandwidth and data transfer paths

GridKa cloud Resources

Pledges	July 2010	April 2011	July 2010	April 2011
Site	CPU(HS06)		DISK(TB)	
GridKa-T1	21600	28250	2190	3125
DESY-HH/ZN	4800	6200	740	1050
Göttingen	2280	3800	250	400
Wup	3440	4633	370	633
Freiburg	3440	4609	370	518
LMU/LRZ	3440	4609	370	518
MPI/RZG	3440	4609	370	518
UIBK/A	1850	1857	94	120
PL Federation	3180	4000	285	385
CSCS/CH	3063	5420	364	469
FZU/CZ	1783	3500	150	400
Sum T2	30716	43237	3363	5011

- Status of pledges accdg WLCG:
 - http://lcg.web.cern.ch/LCG/Resources/WLCGResources-2010-2012_04OCT2010.pdf
- substantial increase (30-40%) of resources for 2011

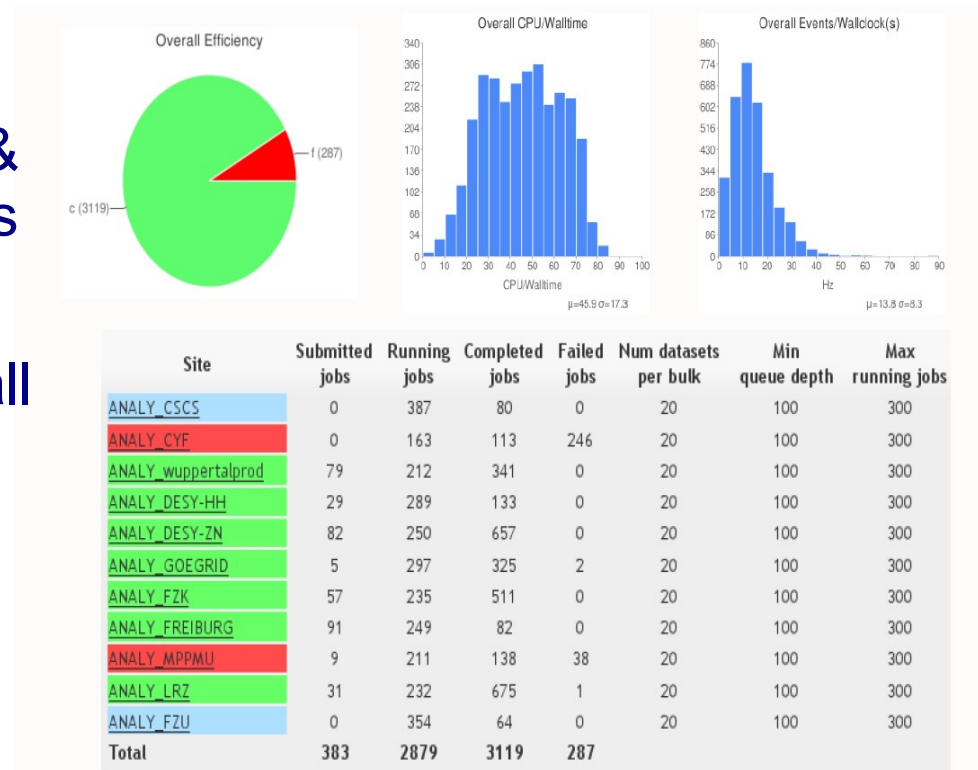
ATLAS-D Computing Organization

- ATLAS DE cloud operation
 - production, data management, site checks, analysis support
 - team of ~10 people (BN/KIT, FR, Goe, M, MZ, Wup, Zeuthen)
 - weekly phone meeting
 - contact for ATLAS central operations & shifters in case of problems, tickets, ...
 - monthly cloud meeting involving T2 sites (plus F2F 1/year)
- Mailing lists
 - gridka-atlas@lists.kit.edu
 - High traffic, used by operations team, day-to-day issues are discussed here.
 - atlas-germany-computing@desy.de
 - Low traffic, broader audience list, includes most people active in ATLAS-D computing, mainly used for announcements

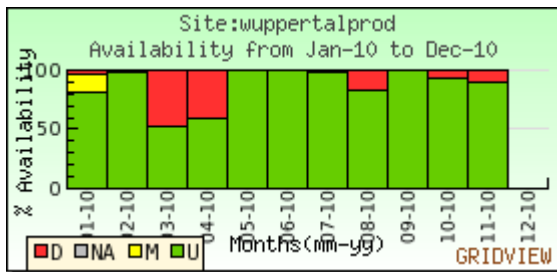
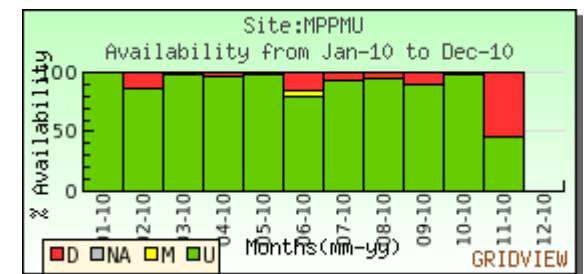
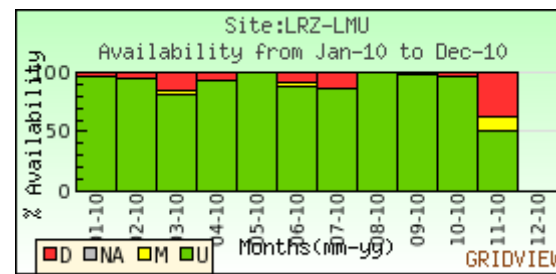
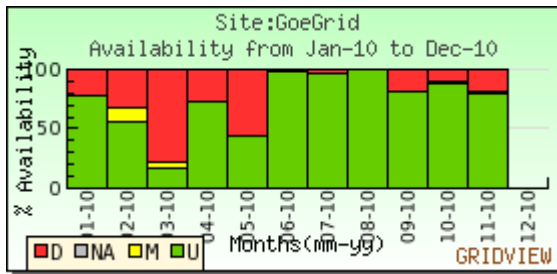
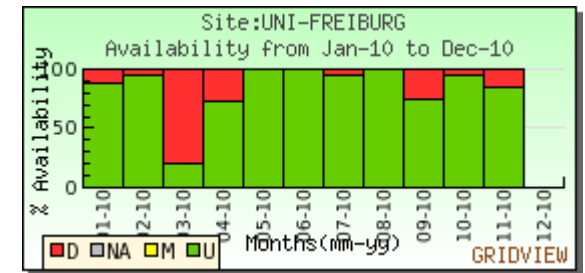
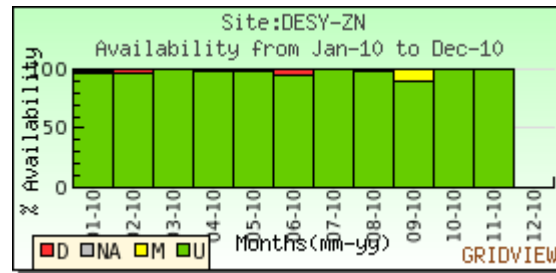
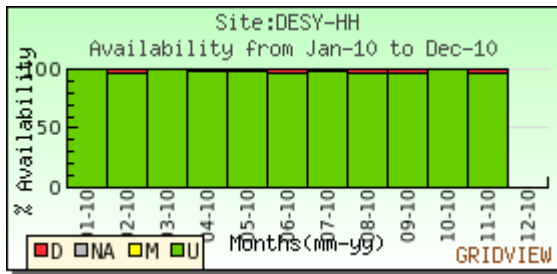
Problems/Issues at HGF T2 sites

- Setup and operation of stable and performant site non-trivial
- Close cooperation (dCache support) & extensive testing (eg Hammercloud) is crucial
- Mostly stable operation now, though all sites have to cope with more or less serious and critical problems at times, eg:

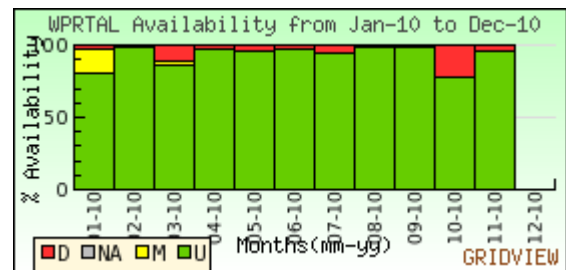
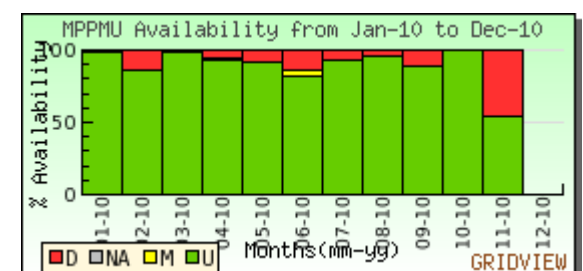
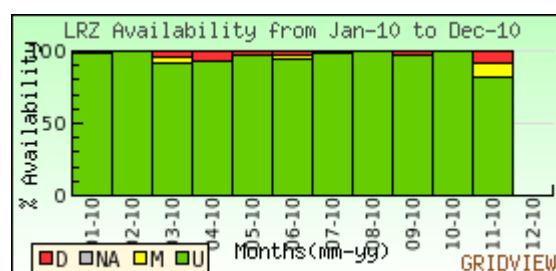
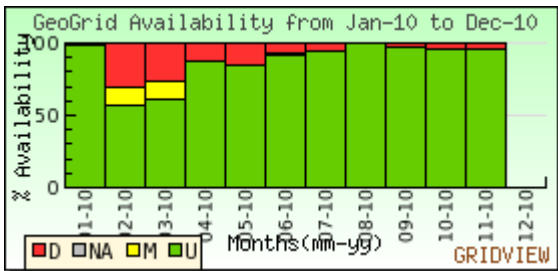
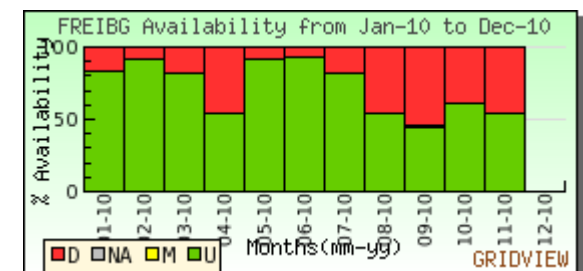
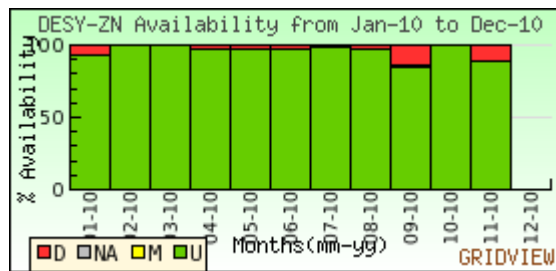
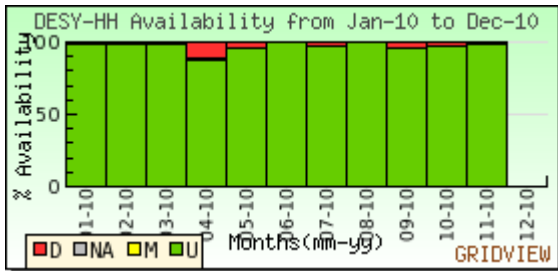
- Desy-HH: DDN servers, file loss
- Wup: deploying HP storage servers
- FR: cooling outage in summer
- LMU: continous struggle to operate ATLAS/gLite SW stack w/ Suse SLES10



Site availabilities 2010 - OPS



Site availabilities 2010 - ATLAS



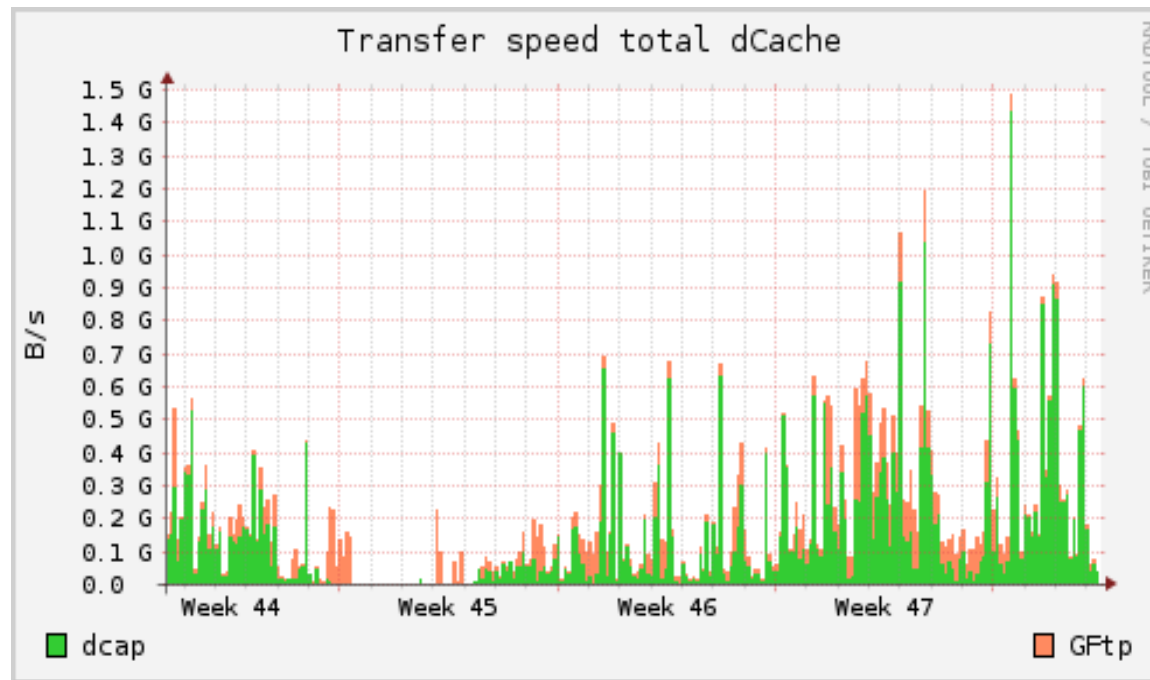
- availability for some sites low
 - in practice not so relevant for operations/analysis.
 - site usage/blacklisting based on direct criteria from data management, production and analysis.

Summary

- HGF funded ATLAS T2s are working well
 - stable contribution to production
 - well integrated in ATLAS data management
 - increasing analysis usage
- Close integration in DE support structures and ATLAS operation teams crucial
 - storage systems (dCache support)
 - Hammercloud tests
 - data consistency and recovery services
- Implications of computing model evolution still unclear
 - more information and tests needed

Add-on: networking LRZ

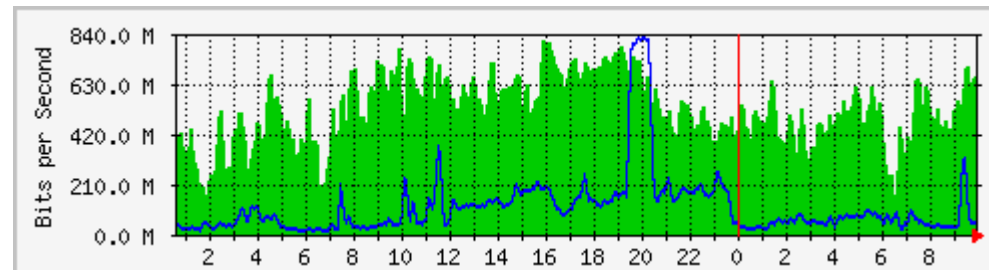
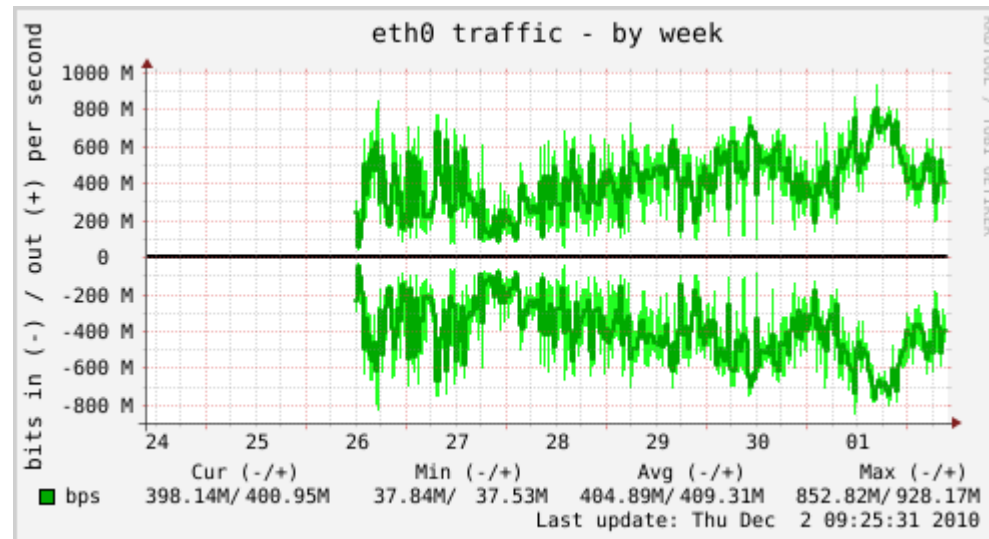
Transfer rates in/out LRZ dCache storage in Nov 10
(local/dcap=green, WAN/gridftp=orange)



- occasional spikes of 200-300 MB/s for WAN
 - on average well below 100 MB/s
 - no problem with current connection to X-Win/GridKa so far (LRZ)

Add-on: networking WUP

Wuppertal: gridftp door traffic last week, router traffic last month



-
-
- average below 500 Mb/s (<60 MB/s)