

The ATLAS GridKa Cloud

John Kennedy
LMU Muenchen
on behalf of many other people



Introduction

- Brief Intro to a Cloud
- Sites
- Organization
- Service Challenge
- Production

A Cloud

- A Cloud is more than just a collection of sites
- Provide a service to the community
 - Data storage and distribution
 - MC production (and aggregation)
 - Distributed Analysis
 - Re-processing of RAW Data
- Central services provided by T1 are essential to the Clouds health
- T2's and their associated manpower are providing much of the driving force

A Cloud in the comp Model

- Data from CERN to T1
- AOD flow T1->T2
 - Need tools to do assignment and decide how we distribute the data
- MC prod at T2 with data aggregation at T1
- Distributed analysis at T2
 - With fraction at T1 ESD
- Re-processing of Raw Data at T1
 - Need development of tools to do this

The Tier1

- In many ways the cornerstone of the whole Cloud
- Recieve Data from CERN + push AOD's to T2's
- DDM Cloud services run on T1 (DQ2/FTS/LFC)
- Re-processing of RAW data
- Don't have an ATLAS person at GridKa, and their staff can't read our minds (we've tried)

The Tier2's

- Currently a large community of T2's
 - DESY Zeuthen
 - DESY Hamburg
 - Wuppertal
 - Freiburg
 - Munich LMU/MPI
 - CSCS
 - PL CYF,poznana
 - CZ Prague
- Varied setups: different OS, batch system, storage
- You'll hear much more tomorrow

Organization

- How do we organize a Cloud
- Group by tasks and/or by sites?
- Site contacts present on mailing list
- Define major tasks and responsibilities
 - Distributed Data Management(DDM)
 - Monte-Carlo Production
 - SW installation
- Monthly phone/VRVS meetings, share experience and horror stories (1hr)

Tasks within a Cloud

- Defining the tasks we need to do helps
 - Factorize and Organize
 - Decide where we need the manpower
 - Identify and gaps in our knowledge
- Some of the main areas are
 - SW installation
 - Data Management
 - Production
- By joining these groups and in some cases aiding with development we become stronger

ATLAS SW installation

- SW installation is done via LCG jobs
- Problems **DO** occur
 - Sqlite at DESY Zeuthen/Hamburg
 - Currently ZN/HH/FZK missing latest release due to installation problems!
- Important to have input here
 - Problem sites are often focused on last and sometimes miss sw releases
 - By being part of the install team and having good contact with sites we can ensure that we get the sw installed on our Cloud
- Also need to clean up old releases

DDM

- Aim to have a team of experts (~2FTE) who support the Cloud's data management and seek out problems. Being passive here is not good enough
- Support/Monitor the Site Services provided at the T1
 - DQ2 instance (vobox)
 - FTS/LFC
- Develop knowledge base and perform independent tests

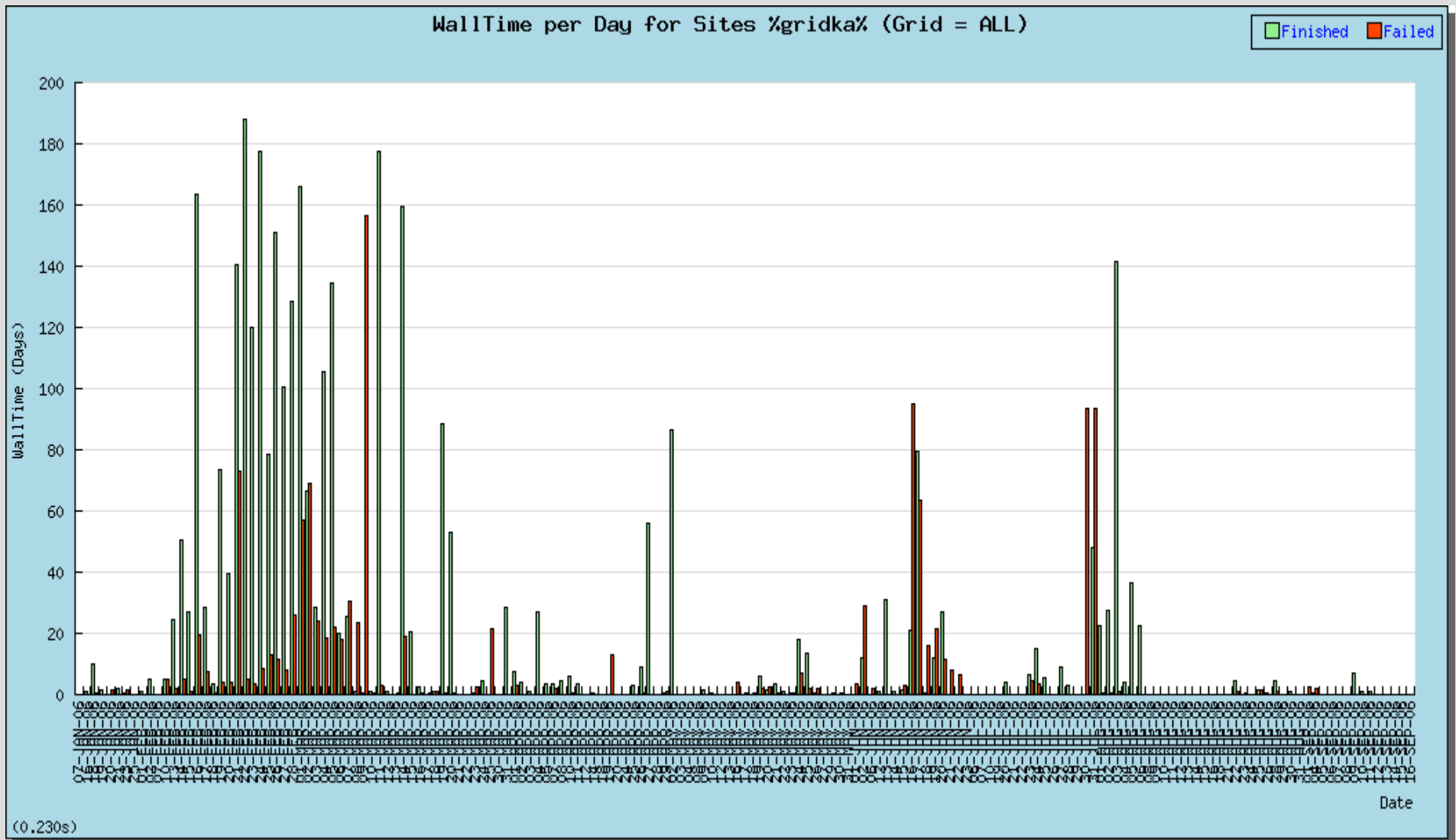
MC Production

- Production Jobs sent via central team of executors
- Until recently production managers responsible for all aspects of production
 - Assignment of tasks to a Cloud
 - Submission of jobs
 - Tracking and reporting/solving of problems
- Now aim to have shift teams to cover the various tasks.

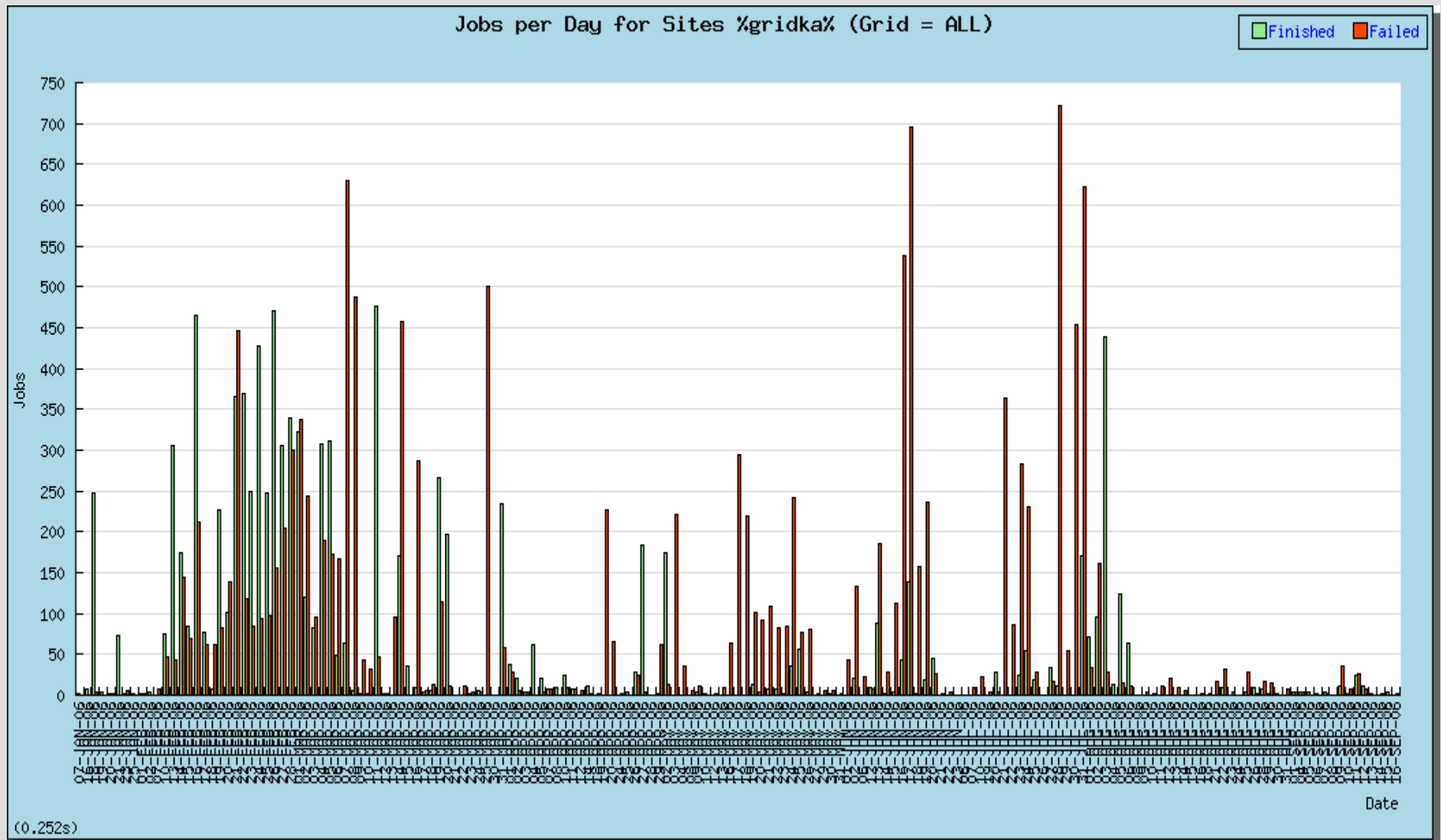
Production Shifts

- Several teams working in weekly shifts covering LCG
- Teams
 - Task assignment
 - Job Submission (should just be running the executor)
 - Job Managers, track jobs and errors both grid and athena related
 - Data Managers, track the data movement(DDM). Ensure datasets are aggregated correctly and input data is available for pending jobs
- Weekly meetings via phone

GridKa Production



GridKa Production



Production in total (ATLAS)

- 50M events produced in Total during 2006
- Currently at approx 1M evts/week
- Aim is to double the rate within next month(s)
- 794220 Finished jobs @ 163 sites with 45% job efficiency
- Factor of 2 increase every 4 months planned (hoped for)
- Some of the bottlenecks will be manpower related!

Our Production Totals 2006

Site	Finished jobs	Failed Jobs	Eff(%)
GridKa	9665	14206	40
CZ	5091	10317	33
PL	7202	28648	20
CSCS	3754	5448	41
Frei	5551	12122	31
DesyH	3228	6368	34
DesyZ	3465	3171	52

- Efficiencies in terms of walltime tend to be O 70% (so why did I show the jobs?)

SC4 – The Goals

- Run a full scale T0 exercise, from EF, reconstruction farm, with T1 and T2 export
- Realistic data sizes and complete flow
- 700MB/s export from T0->T1's
- Include all T1's
- T2's added later (15)
- Large scale test of DDM and also first time many users have encountered it, lots to be learned

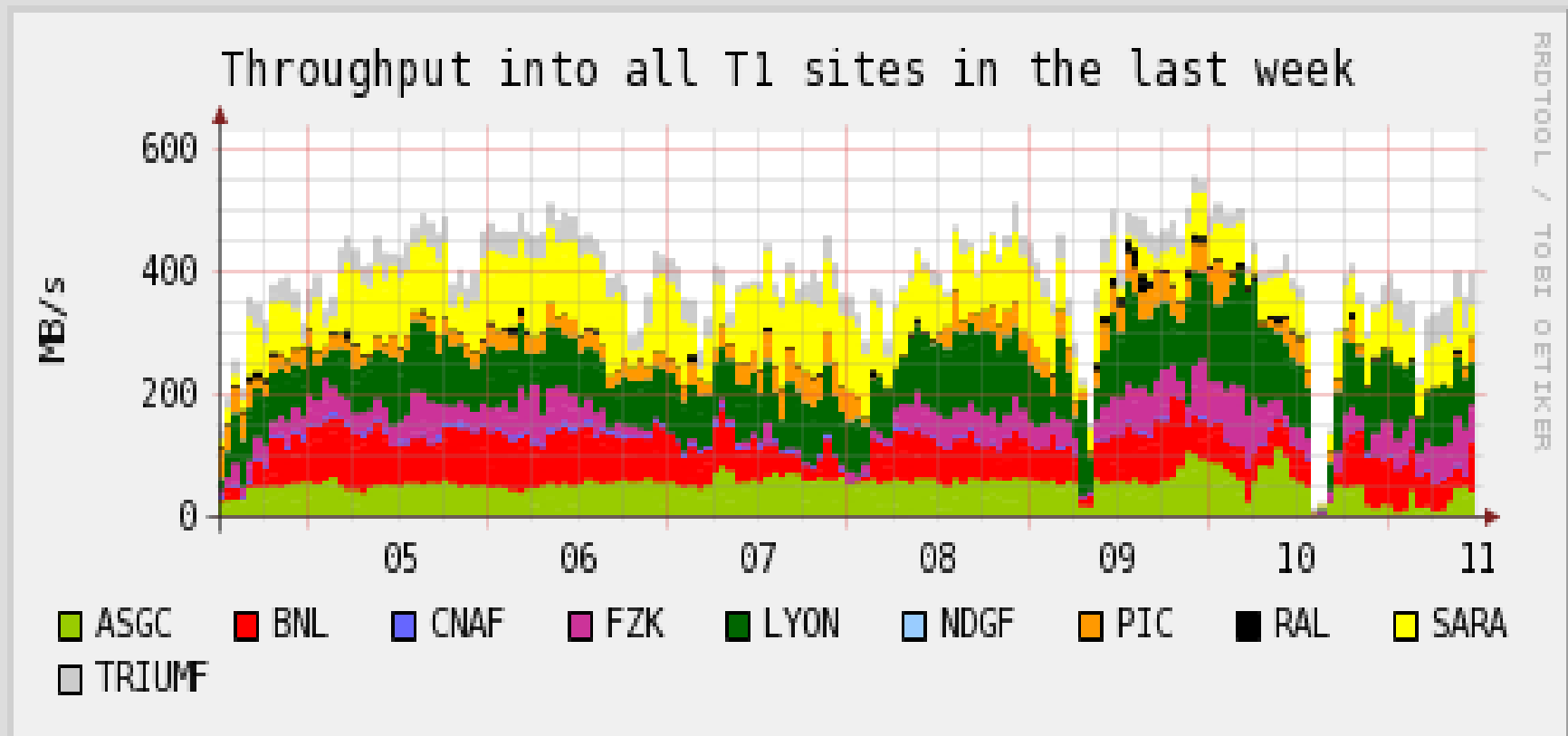
SC4 buildup

- We ran a pre-sc4 mini transfer exercise to help prepare our sites
- Pre-SC4 Transfers
 - FTS transfers tested T1<->T2: Functionality and bandwidth
 - Initially one person testing and documenting on a wiki page
 - All other sites then follow from this experience
 - This allows sites to learn about transfers and we share the workload
- Aim to continue with this pre-test idea when possible

Tiers of ATLAS

- DDM uses ToA file to define
 - Cloud structure and constituent sites
 - FTS topology, defining which FTS server is responsible for which transfers
- All sites need to be in ToA before they can be used correctly
- Coordination needed to make sure this happens and ensure it remains correct.

SC4 results



- Throughput during week 11/08/06
- Although at peak times the transfer rates achieved the goals this was not sustained

SC4 results - FZK

- Suffered due to VOBBOX problems
 - User login node was our vobox, loads of 90 seen
 - Quick response from GridKa to provide new box
- Sea-sawing transfer rates
 - Nominal rate achieved at times but with periods of hours with no transfers – not fully understood
 - We do have just 4TB Disk and this gets full easily. Better communication between ATLAS/GridKa needed!

T0 test post-mortem

- From Miguel Branco – DDM
 - Goals not achieved but limitations understood
 - FZK had problems which were quickly addressed-most notably the VO BOX (and here you can argue we didn't really know the requirements).
 - As for the overall stability of the site, yes, clearly we will need more stable sites but FZK is far from being any different from most other sites.
 - We do appreciate (a lot) the quick response given to ATLAS.
 - ... also one of the first 'pioneers' in testing the full T0->T1->T2s->T1 model and has found issues, some of which in our own m/w.

T0 test post-mortem

- From Miguel Branco – DDM
 - All Sites: must monitor their service proactively
 - Objectives for next iteration of tests already understood.
 - Include more T2's (especially large ones)
 - Better contact to site/cloud reps needed
 - More stable running
 - Understand implications of production and distributed analysis data access and data transfers within clouds
 - In good shape for a re-run
 - started T2 tests already, not going too well
 - T0 export planned from 25th Sept

Distributed Analysis

- Physics community needs the tools and Computing community needs to understand the implications of their (chaotic) usage.
- Through the D-GRID project we have a Ganga developer in our community
 - Advice, tutorials, feedback etc
- Aim to aggregate CSC AOD data (O 1.5TB) to T1 (T2's) and open this up for analysis

Progress this year

- **Forming a community**
 - Developing teams and expertise in key areas
 - Getting in contact/organizing
 - We have benefited a lot through the contributions of non-German sites
- **Contributed to Production throughout the year**
 - Not quite as strongly as we would have liked
- **Data Management**
 - Organized pre-tests (Jiri Chudoba)
 - Contributed well to SC4 as a Cloud
 - Increasing expertise here, aiming towards a team
 - Developing tools (see next slide)

Monitoring DDM Transfers

File Edit View Go Bookmarks Tools Help

http://pho-srv1.ifj.edu.pl/Atlas/Cyfronet/grid/wwwdq2/Atlas_DQ2_T2T1_Tests_Status.html

DQ2 T2->T1 Tests Status

last updated Sun Sep 17 18:02:36 CEST 2006

Site	Dataset 100 MB	Dataset 1 GB
FZU	●	●
CSCS	●	●
CYF	●	●
DESY-HH	●	●
DESY-ZN	●	●
UNI-FREIBURG	●	●
WUP	●	●

● - complete ● - incomplete ● - failed ● - down

Done

- Tools developed by Andrzej Olszewski

Current status and future

- A lot has been learned from the T0 exercise
- We are currently starting to perform T1 \leftrightarrow T2 transfers and the T0 tests will be re-run from 25th Sept (about 2 weeks)
- Continue development of teams to cover the core tasks
- Plan for Re-processing at GridKa (Jan 2007?)
- Better communication with GridKa: We have no ATLAS person at GridKa and so simply need to talk more!

Current status and future

- Production Global target set at 20M events by end of the year. Aiming for stronger participation in this.
- Improve organization and communication:
 - We need to push a few issues (DDM)
 - Have someone doing Global organization without being responsible for core tasks?
 - We hope to use this meeting to re-start this process