ATLAS Distributed Computing: Experience and Evolution

Armin NAIRZ (CERN)

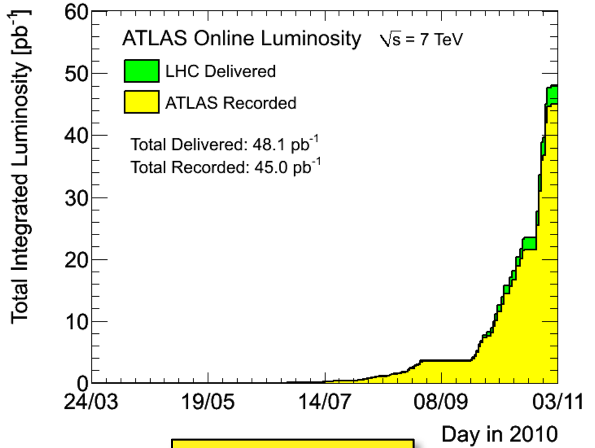
for the ATLAS Collaboration

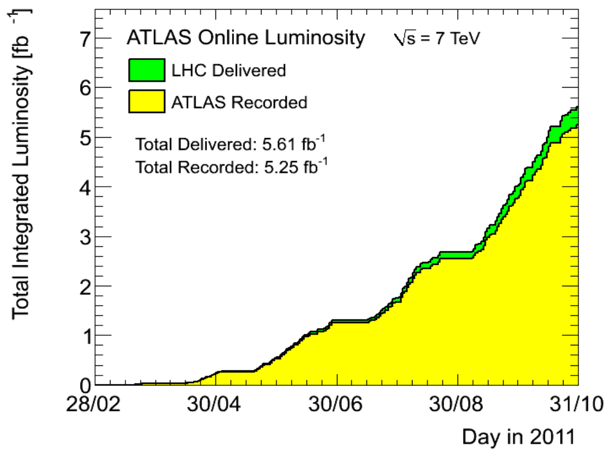
15th International Workshop on advanced computing and analysis techniques in physics

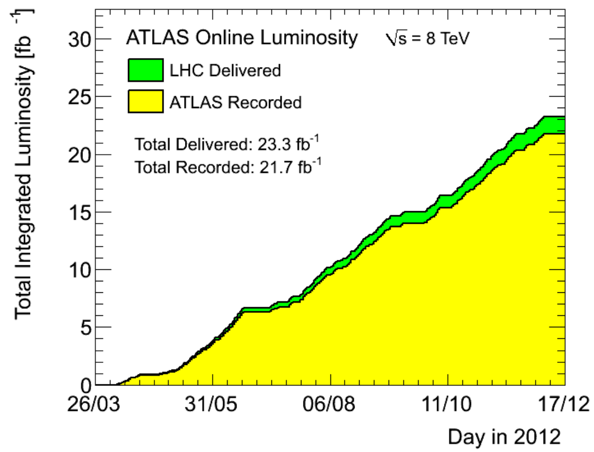
May 16-21, 2013, Beijing, China

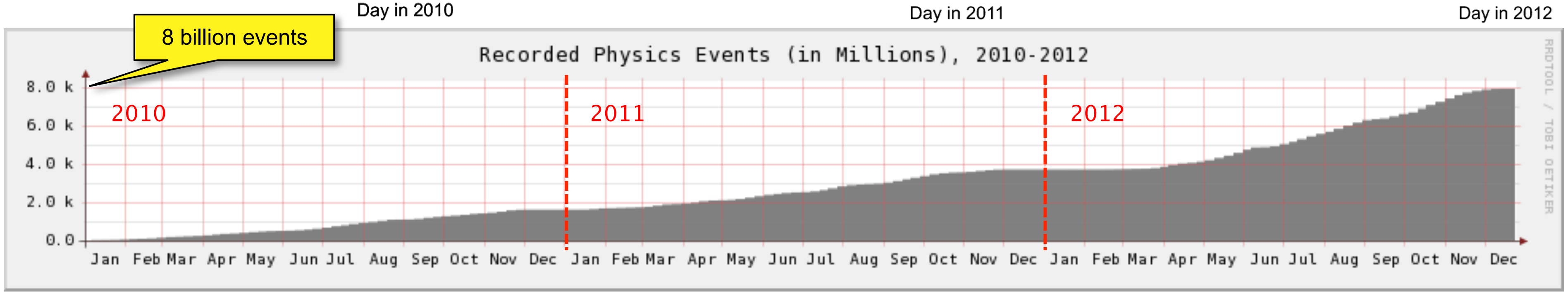
http://acat2013.ihep.ac.cn

Run-1: Successful ATLAS Data Taking









In LHC Run-1, ATLAS recorded

- 5.3 fb^{-1} of p-p data at 7 TeV (2010-11)
- 21.7 fb^{-1} of p-p data at 8 TeV (2012)
- 167 μb^{-1} of Pb-Pb data (2010-11)
- 29.8 nb^{-1} of p-Pb data (2013)

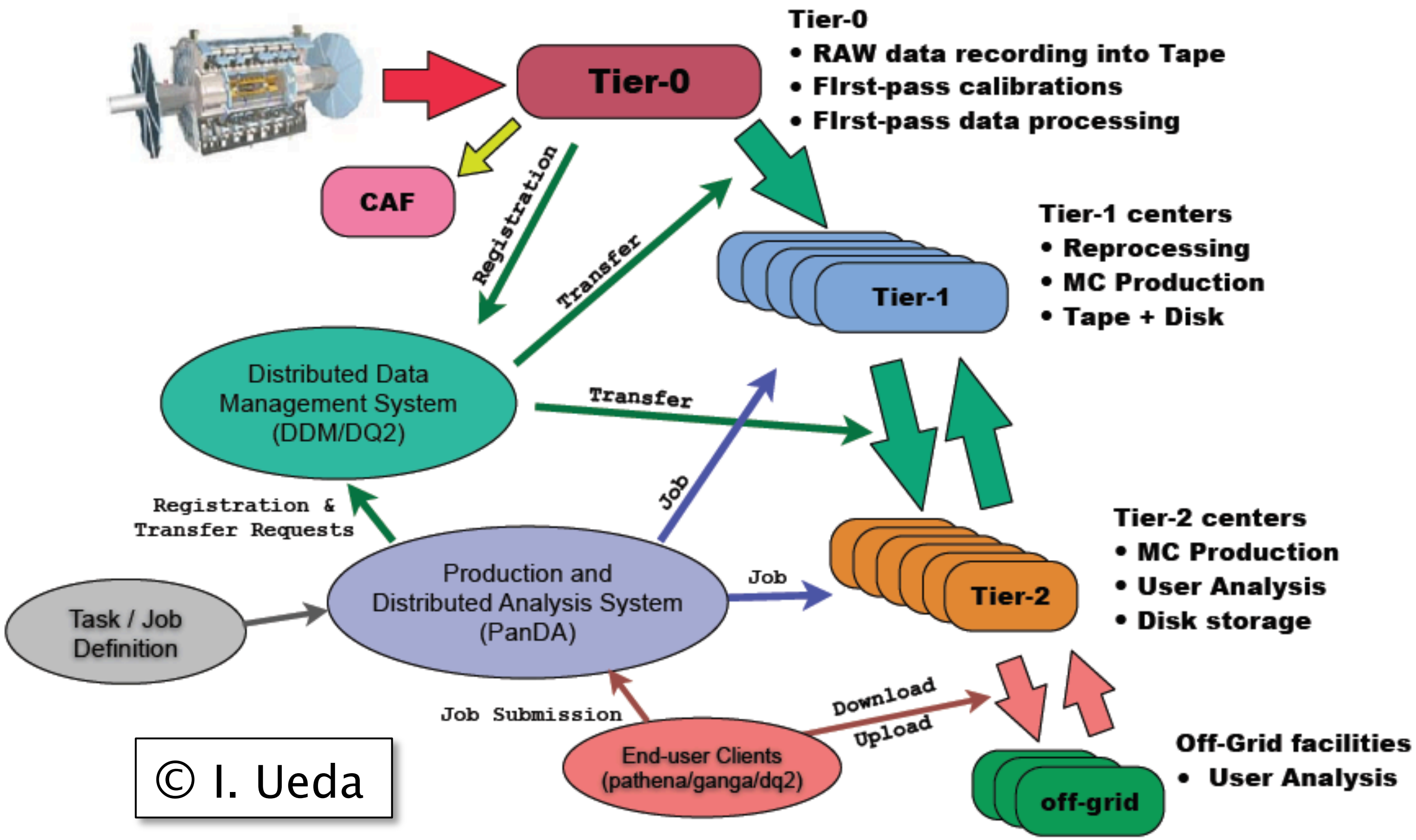
This corresponds to

- 7.6 billion p-p events, 7.4 PB data volume
- 720 million Pb-Pb and p-Pb events, 740 TB data volume

ATLAS Computing System Overview

The ATLAS Distributed Computing System manages the world-wide

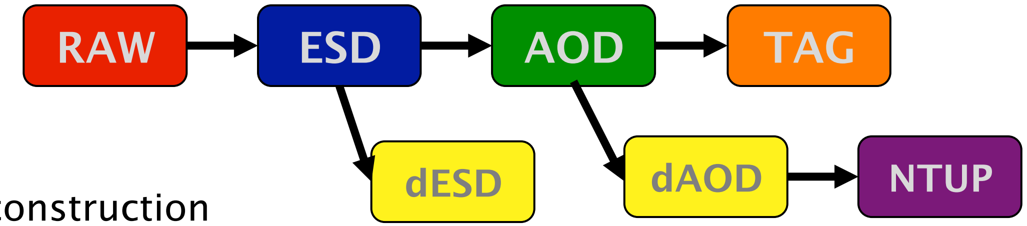
- data processing, MC production, and user analysis jobs, running on up to 150k computing cores
- data transfers to and accesses from ~70 PB disk space (and ~50 PB tape), on over 100 sites set up with LCG (LHC Computing Grid) middleware



© I. Ueda

The ATLAS Computing Model

- LHC Computing Grid: world-wide distributed computing facilities organised in "tiers"
  - Tier-0 (CERN): record RAW data on tape, first-pass processing
    - CERN Analysis Facility: mainly for calibration
  - Tier-1: store replicas of RAW on tape, reprocessing
  - Tier-2: group and end-user analysis
  - MC simulation: wherever possible (and capable)
  - Data distribution / replication over the Grid
    - for redundancy (to secure data with replicas)
    - for accessibility (more replicas for data used frequently)



- Data Types
  - RAW: raw data from the detector
  - ESD (Event Summary Data): output of reconstruction
  - AOD (Analysis Object Data): event representation with reduced information for physics analysis (ATLAS-wide format)
  - DPD (Derived Physics Data): representation for end-user analysis. Produced for working groups or individual end-users (group-specific format)
    - dESD (performance groups), dAOD (physics groups), NTUP (physics groups and end-users)
  - TAG: event-level metadata (event tags), short event summaries primarily for event selection

ATLAS Computing Model: Evolution

During data-taking, the originally laid out Computing Model (as sketched above) had to be adjusted to new use cases, changing run conditions, technology advances, …

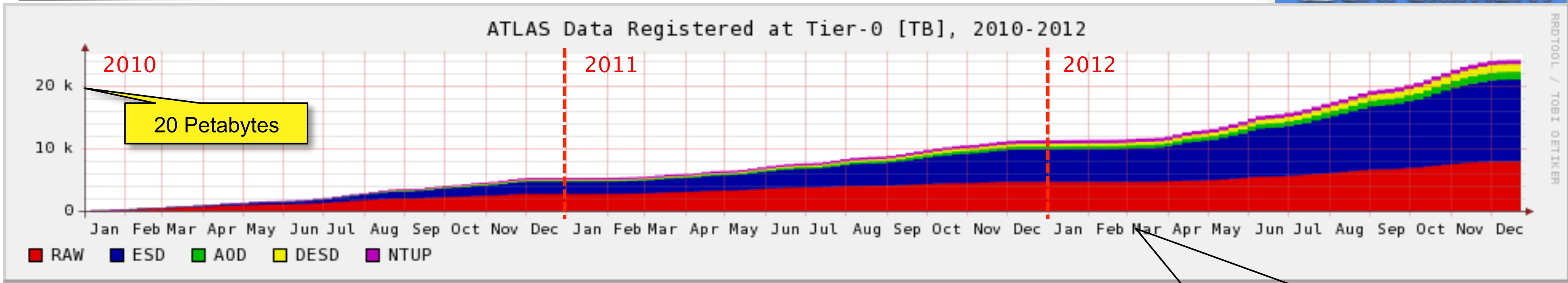
Examples:

- Unexpected usage pattern of data types
  - Revision of data distribution plans
  - Dynamic data placement following the usage
- Higher online data-taking rates (200 → 400 Hz), higher pile-up, larger event sizes
  - RAW data compression at Tier-0 introduced
  - Bulk ESDs with limited lifetime, deleted afterwards
- Data transfer revisited to enable direct transfers, breaking the original hierarchy model
  - From “tree-like” topology with Tier-1/Tier-2 association to “mesh” of any Tier-* combination, based on the measured network performance
  - Faster transfer path gives better efficiency, less load on the system, more end-user convenience

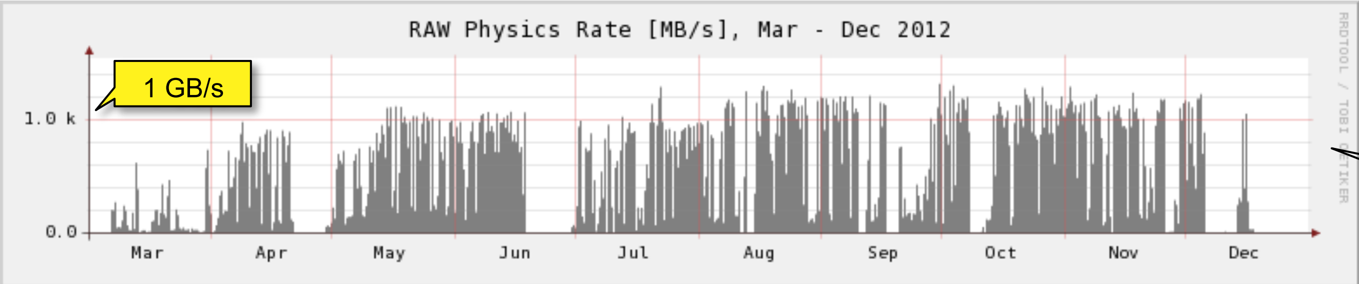
Tier-0 Processing

- The ATLAS Tier-0 system has been running reliably, stably and successfully
  - First-pass data processing has kept up with LHC performance
    - Extension of dedicated resources each year, both CPU and storage
    - Flexible usage of shared CPU batch resources (up to extra 4.5k cores)
  - High-quality data reconstruction already from first-pass processing
    - "Express stream" processing and "calibration loop" before bulk processing
      - Calibration loop: calibration and alignment processing by detector and data-quality monitoring groups
      - Bulk (physics) processing: usually 48h delayed, uses updated calibration
    - Most 2012 data were used directly in physics analysis and publications, without need of prior reprocessing
  - Comprehensive monitoring suite for operations and shift teams

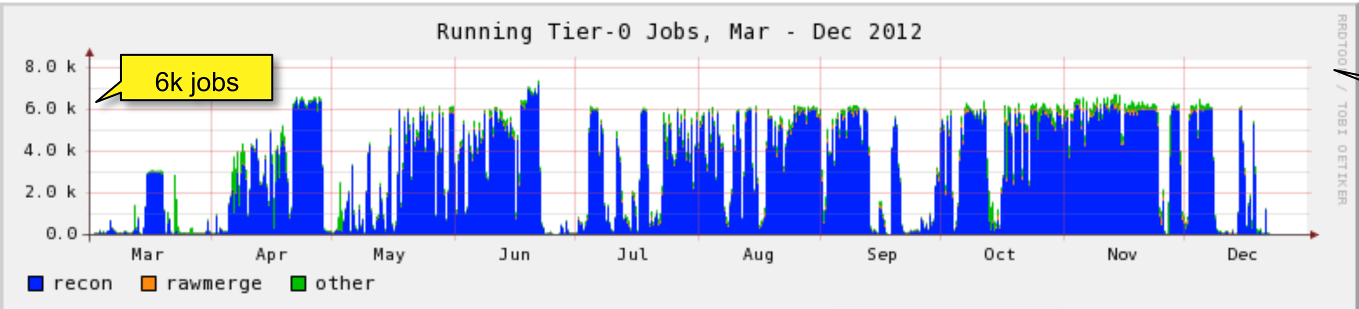
Tier-0 Processing



RAW and derived data products registered for export (2010–12)

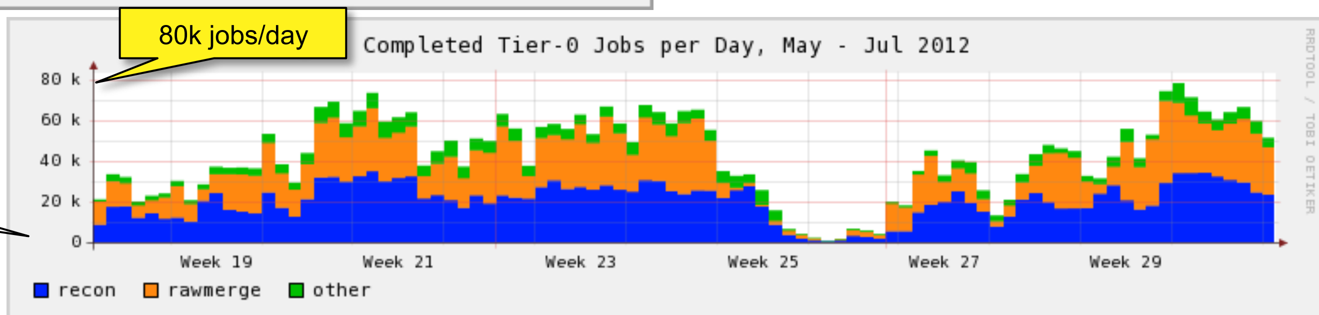


Online to Tier-0 transfer of physics data (2012)



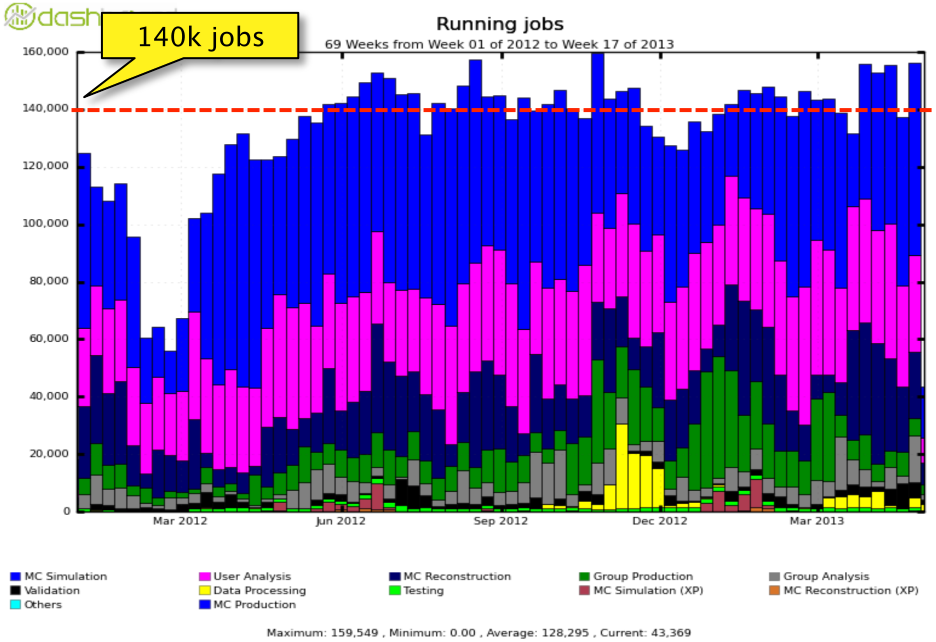
Running parallel Tier-0 jobs (2012)

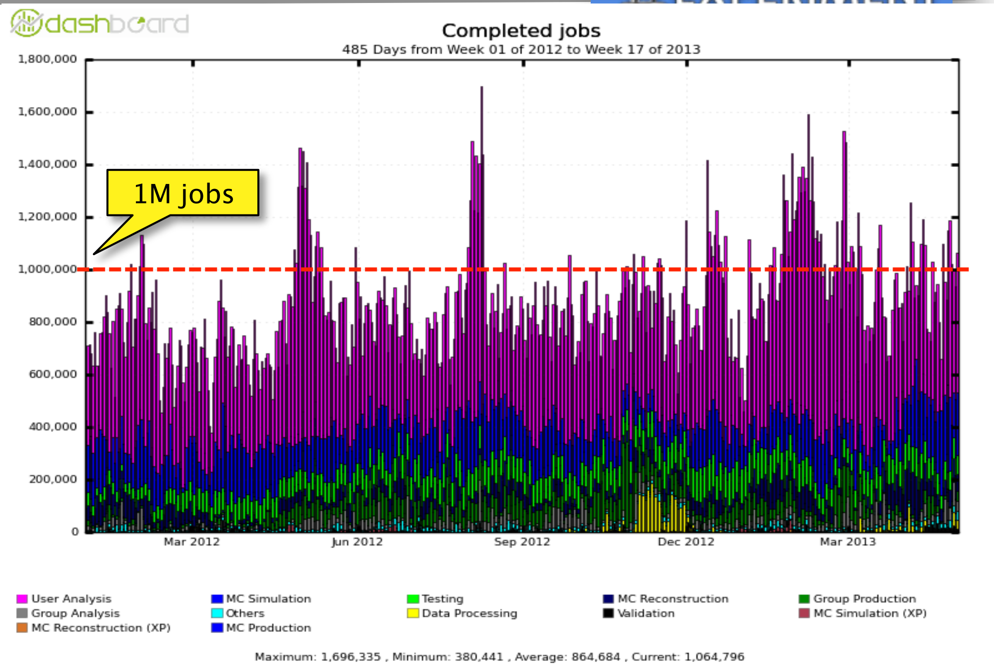
Completed Tier-0 jobs (May to July 2012): note the majority of short merging jobs.



A. Nairz

Distributed Computing

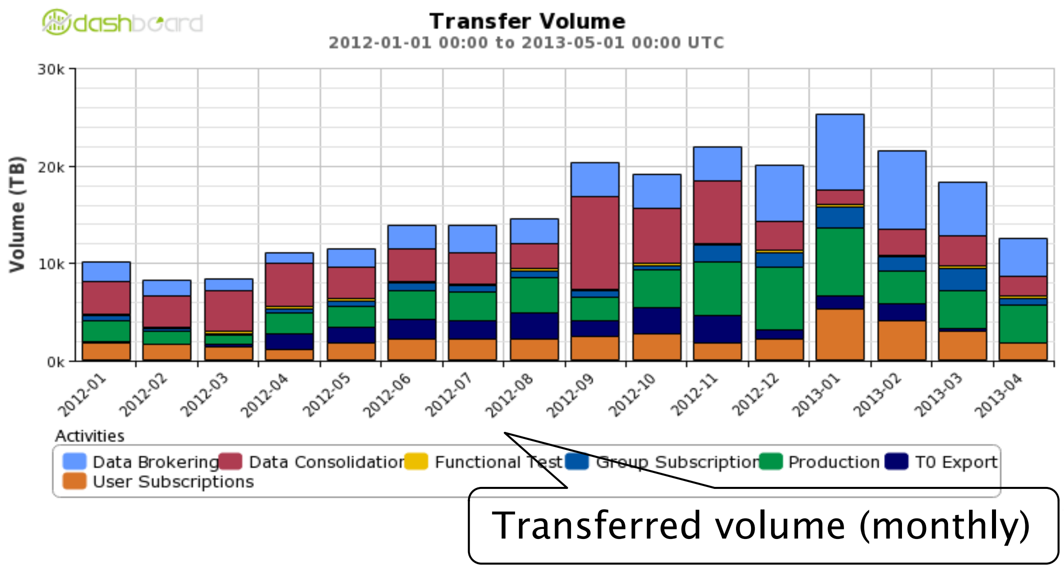


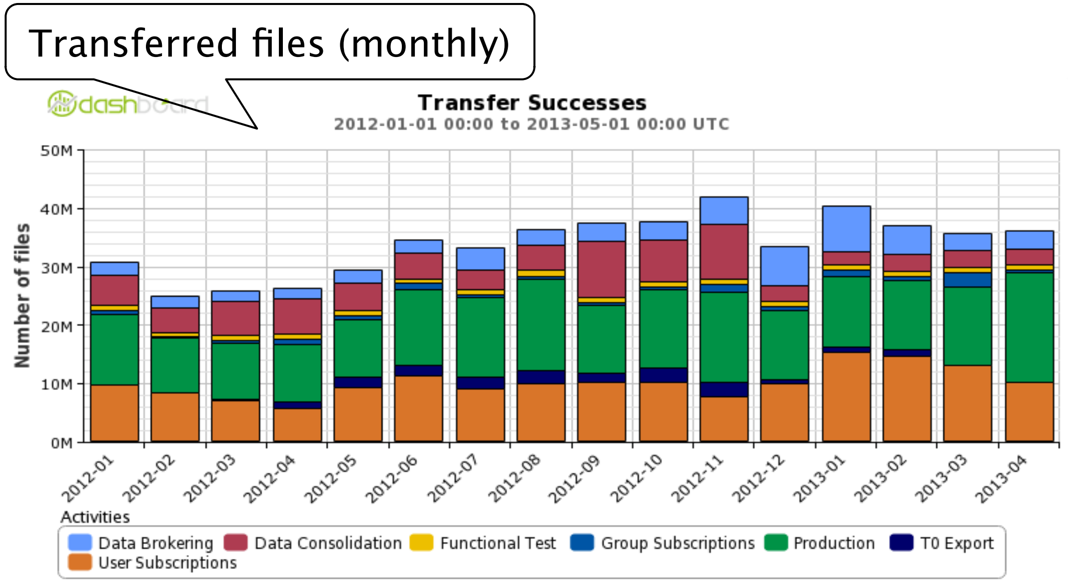


- Data processing on Grid sites
  - Processing scale:
    - Up to 150k computing cores
    - About 1 million data processing and user analysis jobs per day
- End-user analysis
  - Simple and powerful CL interface
  - Output can be either transferred to user's "home" site on the Grid (manually or automatically) or downloaded to off-Grid computers (not monitored)

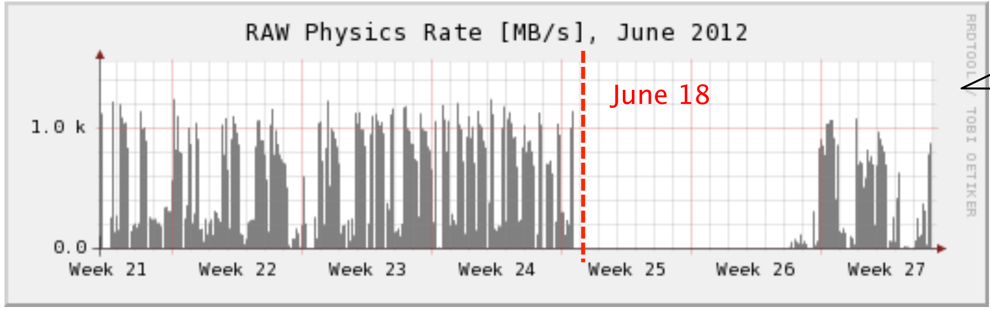
Distributed Computing

- Data distribution to and between Grid sites
  - Transfer scale:
    - About 1 million files per day
    - 400–800 TB per day
  - Includes transfer of input and output of production jobs and user analysis jobs
- Powerful, comprehensive monitoring suite is in place
  - Essential for day-to-day operations, shift service, overviews, etc.

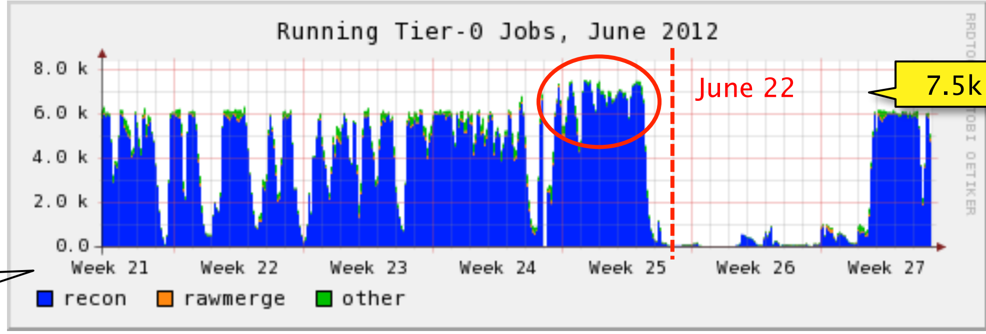




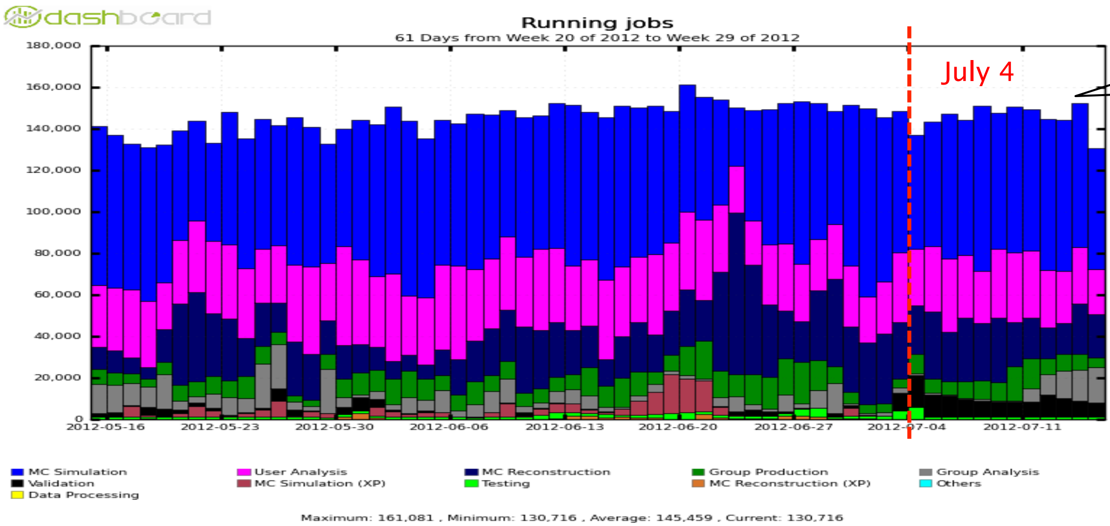
Harvesting the Fruits…



June 18, 2012: LHC Technical Stop, end of data taking

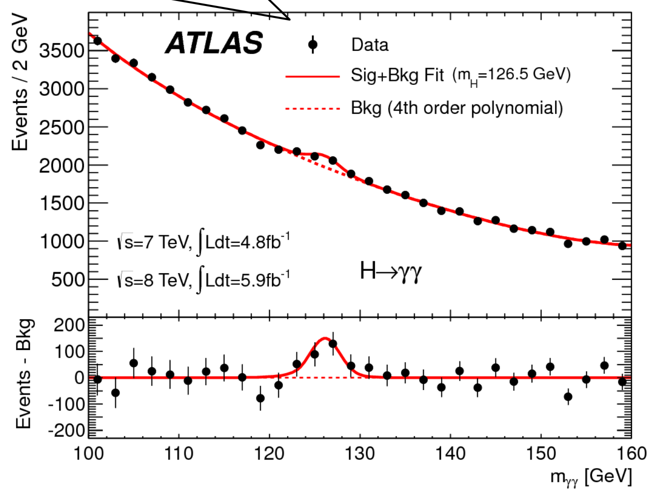


June 22, 2012: Tier-0 processing (and data export to Tier-1 sites) completed



Group production and user analysis

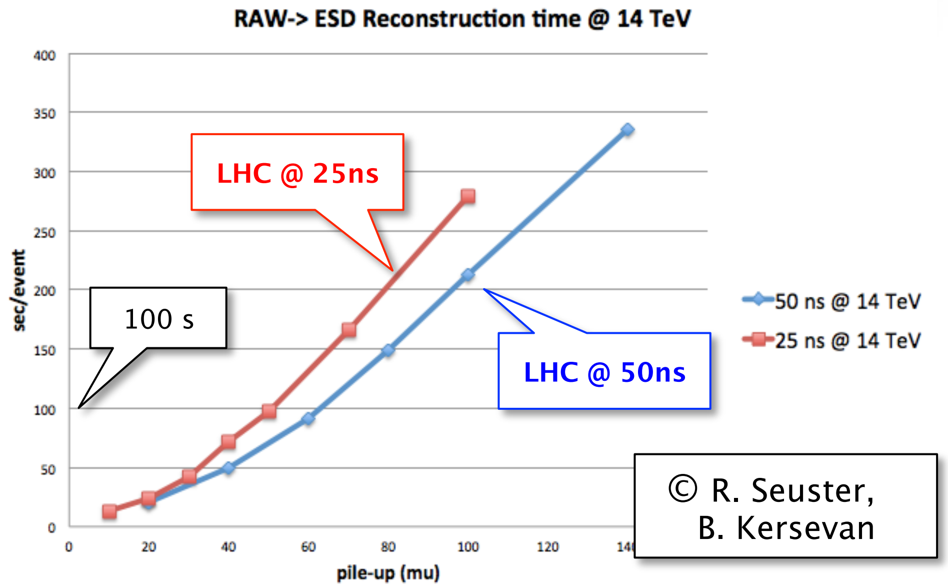
July 4, 2012: results presented at ICHEP



This was possible thanks to a well-working Computing System, but also the dedicated work and great commitment of many people!

Future Challenges: Run-2 and Beyond

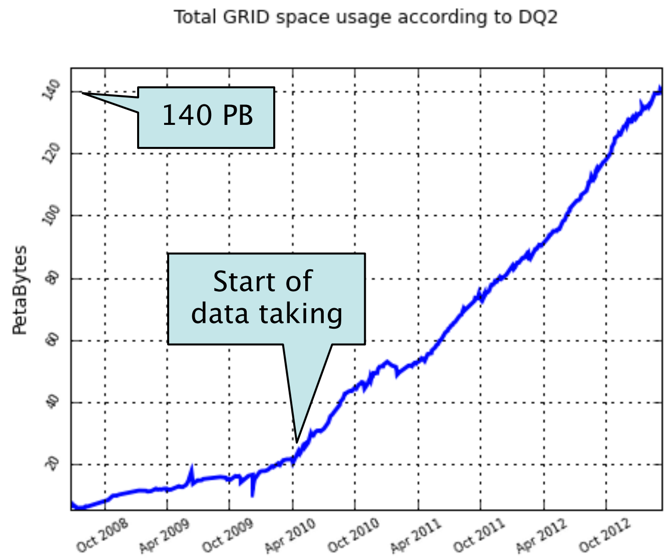
- LHC run conditions in 2015 will bring higher luminosity and pile-up
  - <μ>~40 if levelled, else higher
- The ATLAS T/DAQ system will be upgraded to sustain higher rates
  - Target figure: 1 kHz (cf. 400 Hz in Run-1)



- This will require more resources: CPU, storage (due to more data, bigger event sizes), network, …
- Realistically, there will be no increase in the Computing budget
  - Increase in Grid resources only by technological progress
- The ATLAS Software and Computing community has embarked on an ambitious upgrade and development programme, to address those issues and increase the efficiency of the system.

New DDM System: Rucio

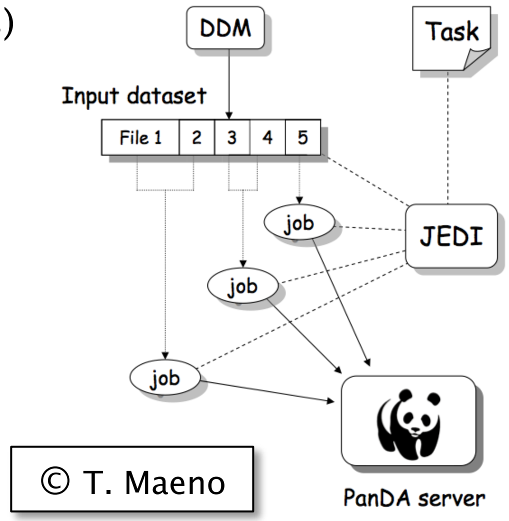
- The current DDM system (DQ2) has worked very well, lived up to expectations (and beyond)
  - 140 PB of ATLAS data on 130 Grid sites
  - >40 PB in 100 million files per year
  - About 1 million files transferred per day
- It will not continue to scale
  - Principal extensibility (design) issues
  - Computing Model and middleware changes



- Fundamental redesign and rewrite are in progress
  - Addressing scalability issues
  - New concepts for optimisation of storage occupancy, data distribution, network utilisation
  - Exploiting new technologies
    - Envisaged usage of noSQL databases (e.g., Hadoop)
    - Support for new protocols (e.g., xroot, http)
  - Better accounting, quotas

New Grid Production System: ProdSys2

- The existing Production System has worked very well
- Similar scalability concerns as with DQ2, and difficulties to accommodate new requirements with existing design
- Fundamental redesign and rewrite are in progress
  - Addressing scalability issues
  - Better support for new/different workflows, more flexible and dynamic job definitions and scheduling
    - Support for multicore scheduling, dynamic job splitting, …
  - New idea: "Event Server"
    - Allows for finer job granularity (processing of single events)
  - Exploitation of new Rucio concepts
    - e.g., file-based data management, transfer rules, …
- Two new main components
  - DEfT (Database Engine for Tasks)
  - JEDI (Job Execution and Definition Interface)



R&D: Concurrency

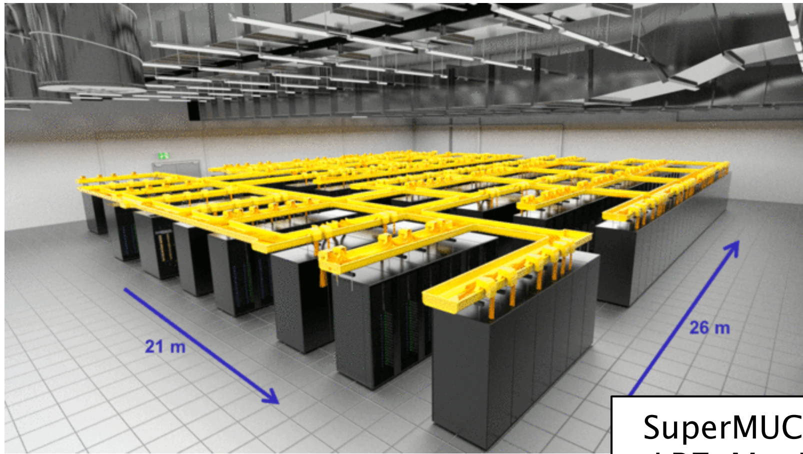
- Goal is a more efficient usage of modern multi-core processors
- Several tracks within ATLAS Software and Computing
  - Parallelisation of algorithms
  - Work on a new Concurrent Framework
    - Full threading, event-level and algorithm-level parallelism, …
    - Benefitting from collaboration with other CERN teams
  - AthenaMP (Athena = ATLAS core software, MP = MultiProcess)
    - Processing of event loop by parallel worker processes
    - Memory sharing ("Copy-on-Write" feature) between workers reduces memory footprint
- AthenaMP is advanced and will be commissioned and moved to production during LS1
  - Challenge of efficiently matching jobs to resources
    - Whole-node scheduling, multi-core queues, mixed serial/multicore sites
  - Flexible input and output file management required
    - Use case for Event Server

R&D: Cloud Computing

- Using free (e.g. academic) or “cheap” commercial clouds as “opportunistic”, additional resources for ATLAS
- Utilisation of statically allocated VMs in a cloud has already been successfully demonstrated in production
  - Adaptation of ATLAS high-level trigger farm
  - Several cloud-based Panda queues have been running production jobs
- CERN OpenStack being commissioned and is offering resources to ATLAS
  - 800 cores since October 2012
- ATLAS is already actively exploring public-private partnerships in cloud computing
  - Still R&D, but on a scale of several thousand cores, sustained over weeks/months
  - Helix-Nebula was a pioneering project
  - New large-scale projects started on commercial clouds: Amazon, Google
- Still many open issues (but Cloud Computing is a quickly evolving field)
  - e.g., efficient usage of cloud storage and data transfers to/from clouds (for user analysis jobs)

R&D: Usage of HPCs

- Using HPCs as "opportunistic" resources is being investigated
  - Very active field, intensive R&D ongoing
- First proof-of-principle tests were already run
  - USA: Intrepid (Argonne)
  - Germany: SuperMUC (Munich), MOGON (Mainz)
  - No commitment yet beyond test accounts
- Issues (examples):
  - Whole-node scheduling
  - No WN disk
  - No outbound IO
  - But workarounds are available
- Event Server (short jobs!) as promising concept



SuperMUC at LRZ, Munich

R&D: Federated Storage

- The traditional Computing Model is based on the "Data Grid" concept
  - Jobs go to data (access via LAN)
  - Replication of data (whole dataset) for higher accessibility
- "Storage Federations" provide new access modes and redundancy
  - Jobs access data on shared storage resources via WAN
  - Analysis jobs may not need all the input files of a whole dataset
    - Transfer of a part of the dataset
    - File and event level caching
  - Create a common ATLAS namespace across all storage sites, accessible from anywhere
  - Make easy to use, homogeneous access to data
- System of XRootD "redirectors" is a working solution
  - "Full Dress Rehearsal" exercise for a federated XRootD storage (FAX) is in progress
    - Currently for sites in US and Europe

Conclusions

- ATLAS Distributed Computing has been working extremely successfully in all aspects of large-scale data processing, data management and user analysis

- To get ATLAS Distributed Computing prepared for the challenging start-up of the LHC in 2015, many interesting, promising ideas and projects are being studied and followed up
- There is still a lot of work ahead of us during LS1
- For the core components (Rucio, ProdSys2) we need to finalise the development in 2013, and commission at scale during 2014

Acknowledgements

Special thanks in particular to the following colleagues who contributed material to this presentation:

I. Ueda, T. Wenaus, S. Campana, B. Kersevan, V. Garonne, H. Sakamoto, K. De, A. Klimentov, T. Maeno, M. Potekhin, T. Lecompte, R. Walker, S. Panitkin, R. Gardner, I. Vukotic, R. Seuster