



ATLAS Operations

Ueda I.

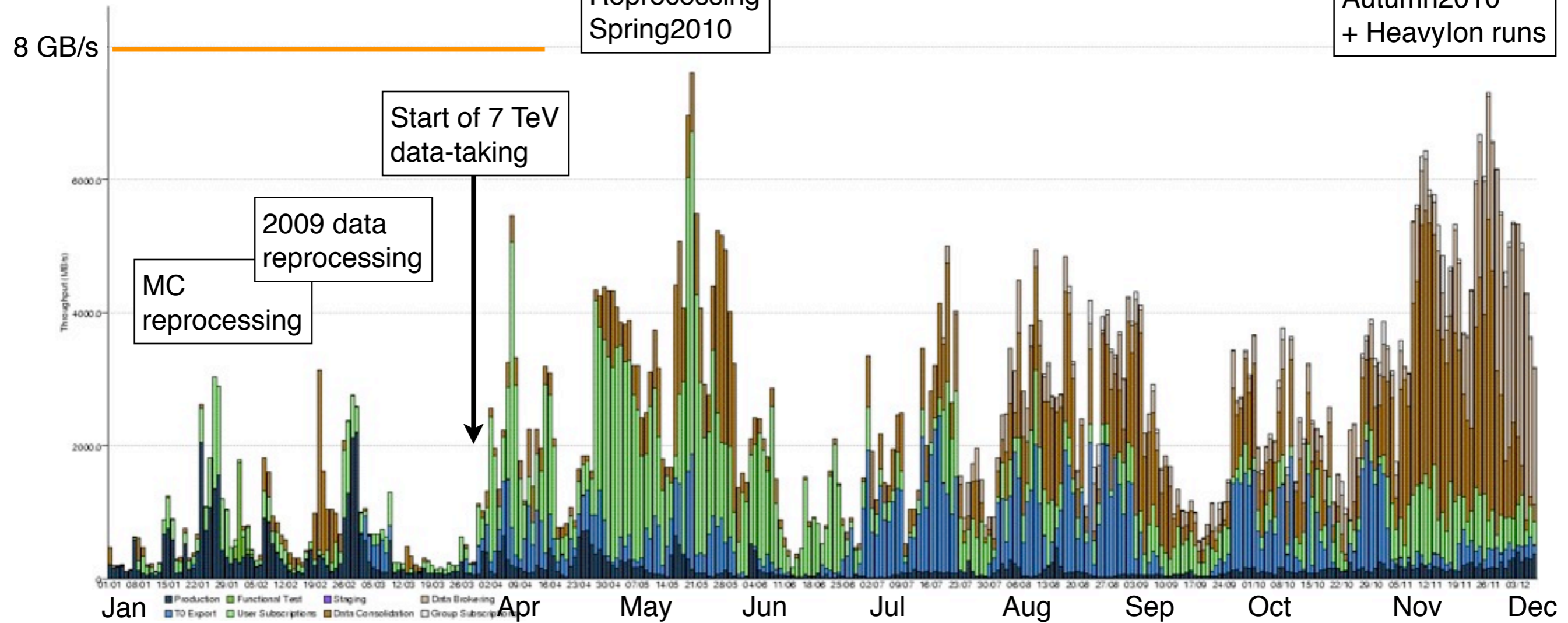
ATLAS activities 2010



Overall Data Transfer Rate

Reprocessing
Spring2010

Reprocessing
Autumn2010
+ HeavyIon runs



Heavy Ion Runs

Different data distribution patterns

- Larger event sizes
- Use cases different from pp data

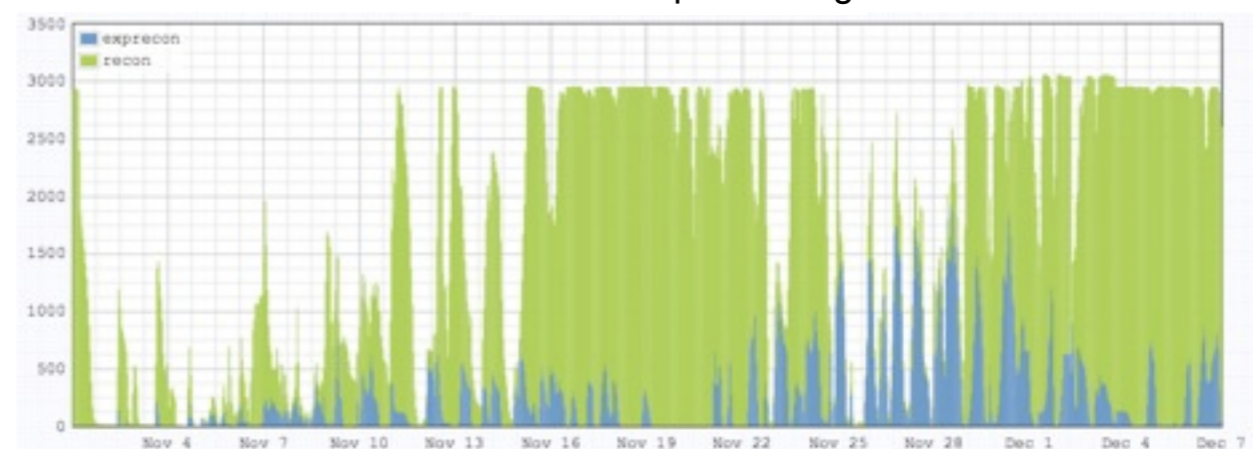
First pass processing at Tier-1s

- Started on Nov 24
- Reducing the load on Tier-0
- Taking more RAW data than expected

RAW Data Writing



Tier-0 data processing

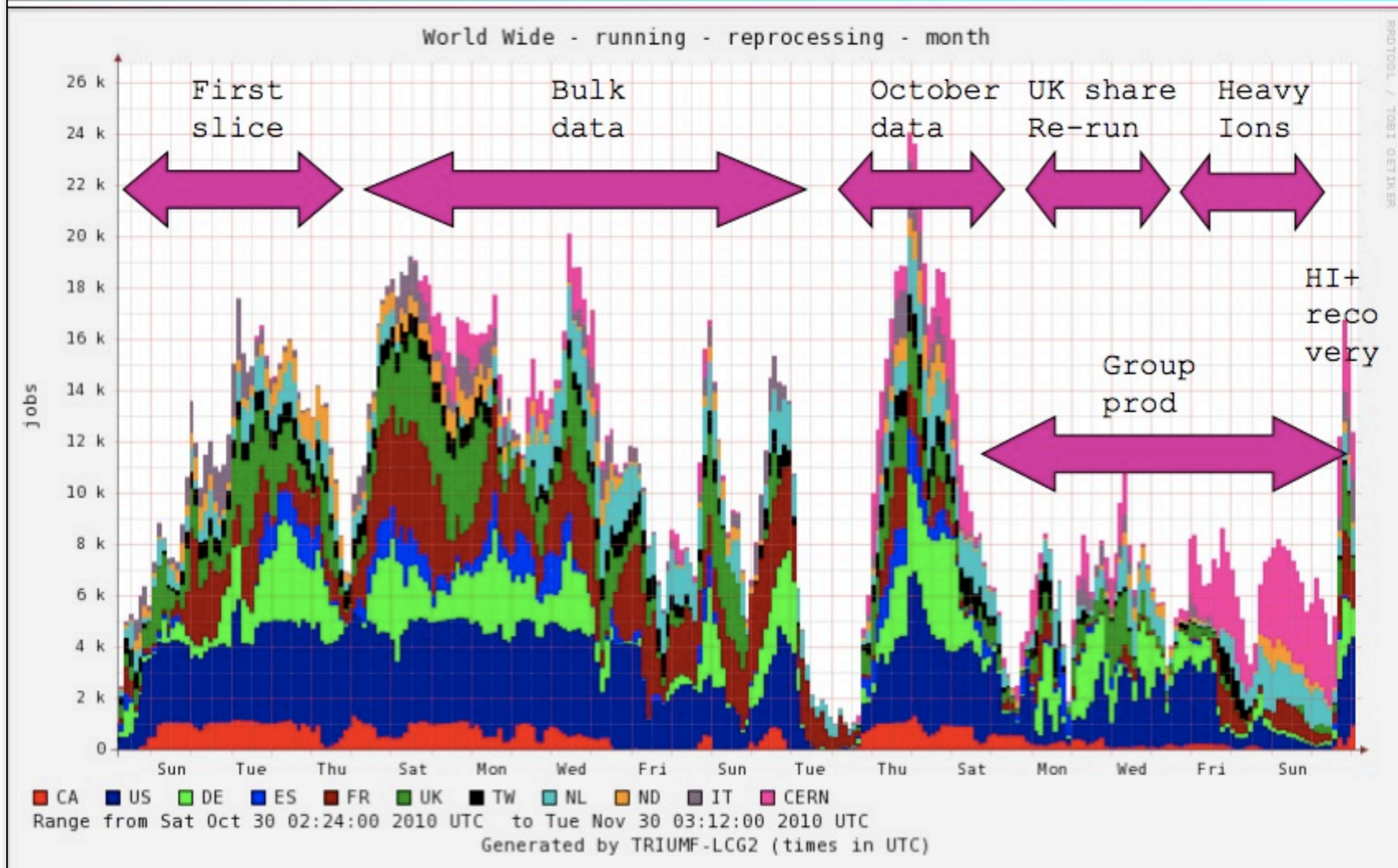


Reprocessing Oct-Nov

S&C Workshop, December 02, 2010



Main Reprocessing Cycles



ES

dCache instabilities: status

CERN IT
Department

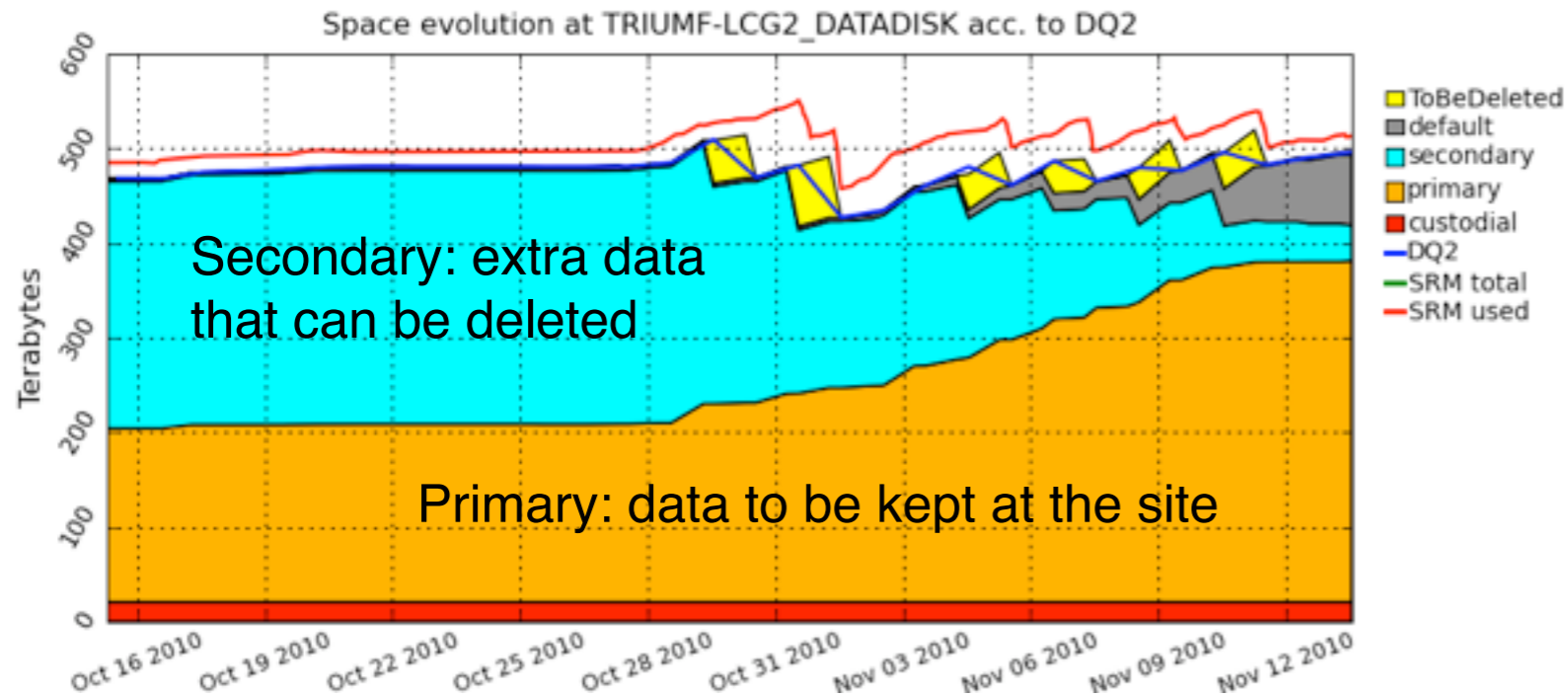
- ATLAS receives now one status report every day
 - Extremely appreciated, very helpful to understand the progress
 - Could Lyon present/append the same report to 15:00 WLCG daily?
- Various activities have been restarted incrementally
 - Basically full speed now
 - Something still throttled (number of jobs, active transfers)
- No signs of instabilities for the moment
 - But no more reprocessing (finished)

CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it



T1_DATADISK

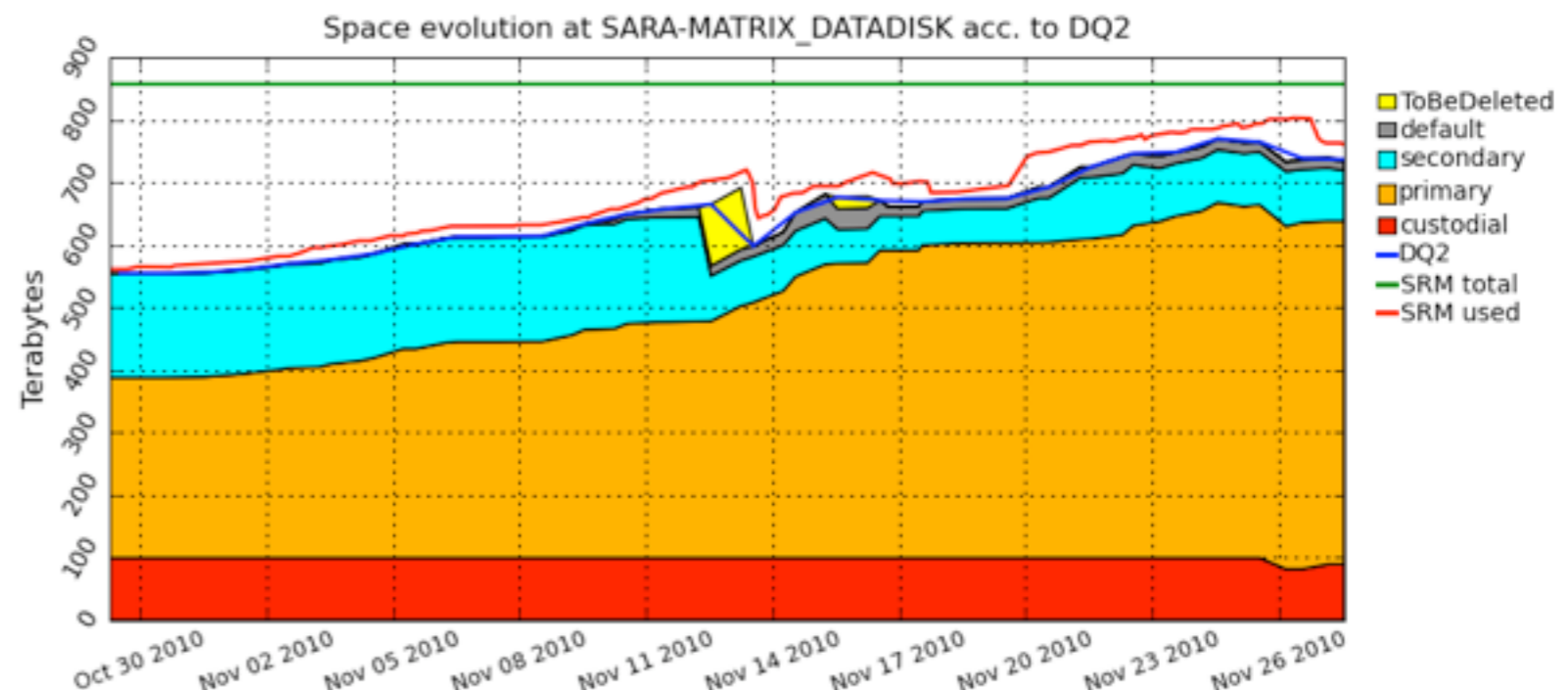
Some Tier-1s had ATLASDATADISK nearly full



Twice more data than the nominal volume of the reprocessing output were stored on Tier-1 DATADISK (due to 'transient' data)

Some old data should have been declared as obsolete before starting the reprocessing campaign

We will try to coordinate better, but anyway disks are getting full...



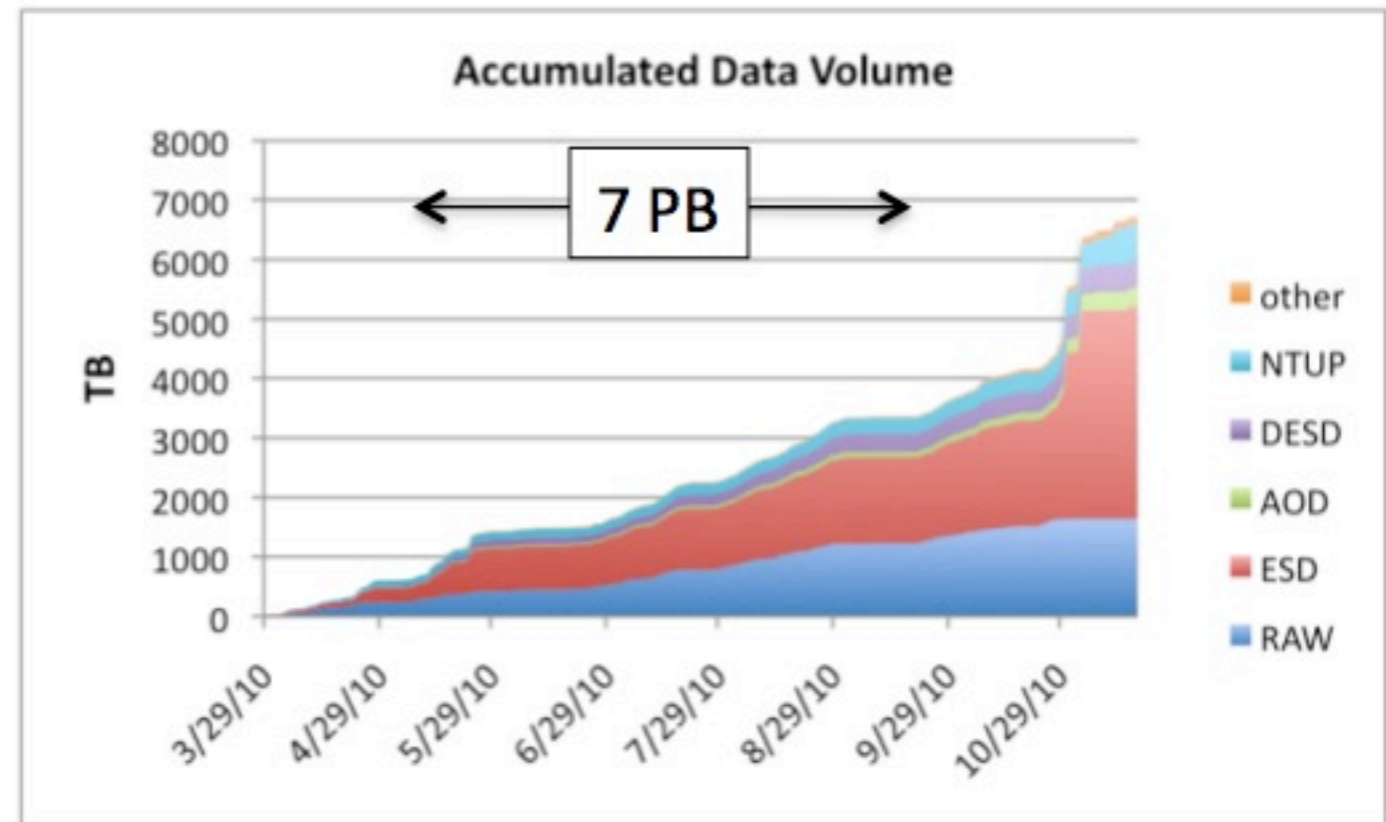
2010 Data Volumes

Data10_7TeV

- RAW 1.6 PB
- ESD 3.5 PB (1 new + 1 reproc)
- AOD + DESD + NTUP 1.4 PB (bis)

Event size in October

- RAW 1.40 MB/event
- ESD 1.48 MB/event

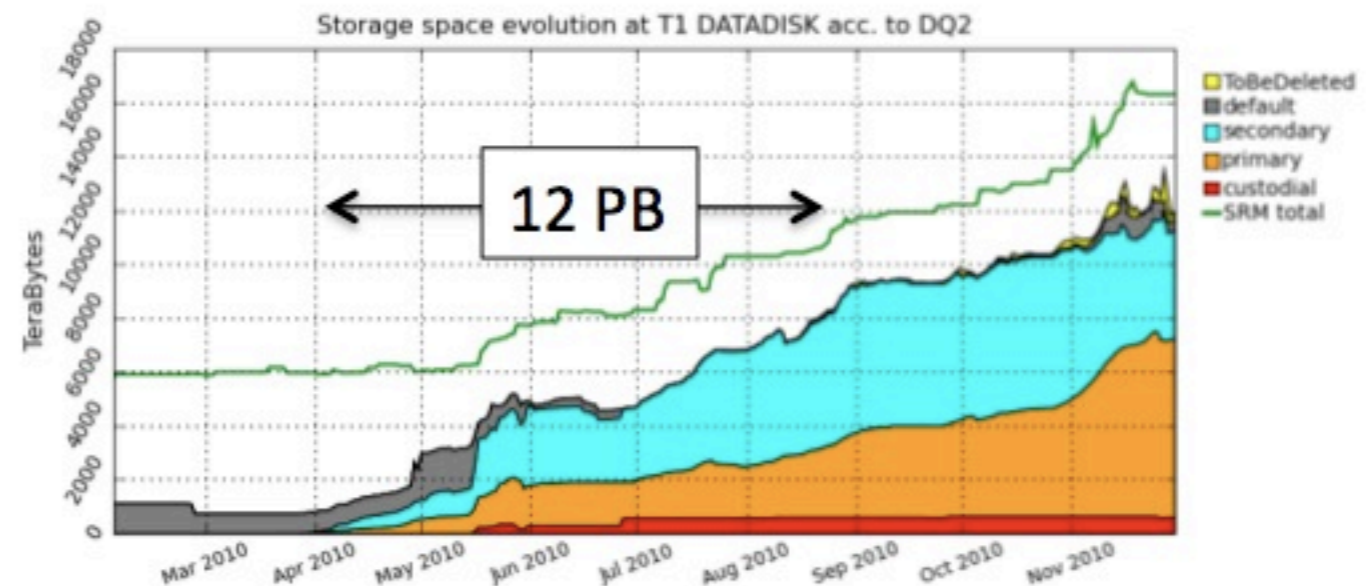


Data10_HI (on Dec.2)

- RAW 300 TB
- ESD 400 TB

Event size on Dec.2

- RAW 1.48 MB/event
- ESD 2.01 MB/event



TAPE usage for 2011/2012



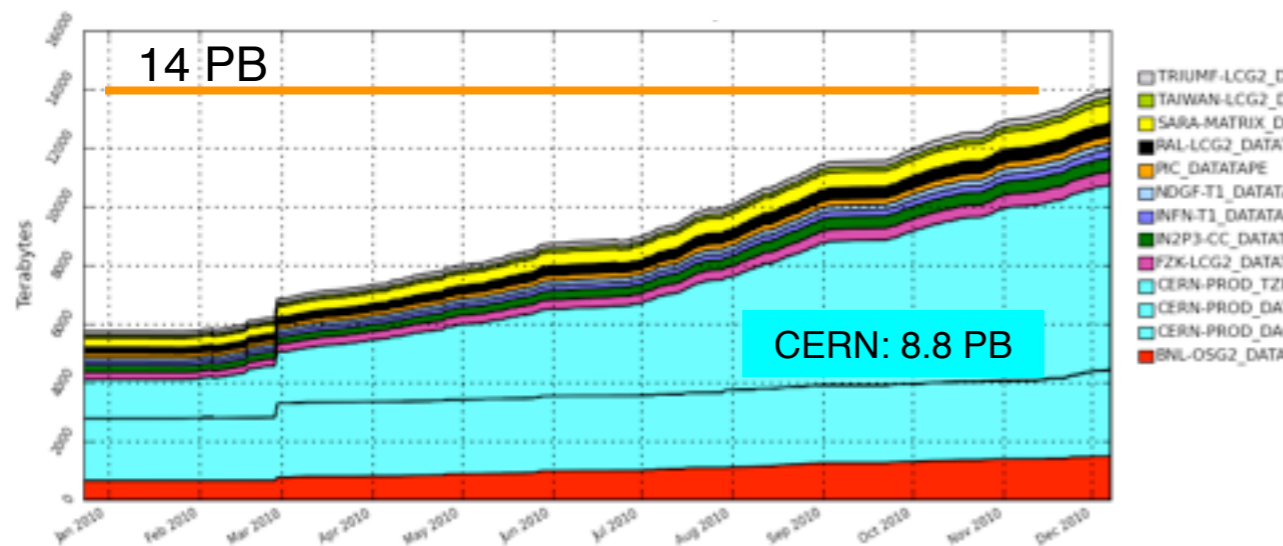
ATLAS has not been fully using the pledged TAPE resources at Tier-1s.

- being afraid of writing data to be deleted (waste of tape)
- having enough DISK resources

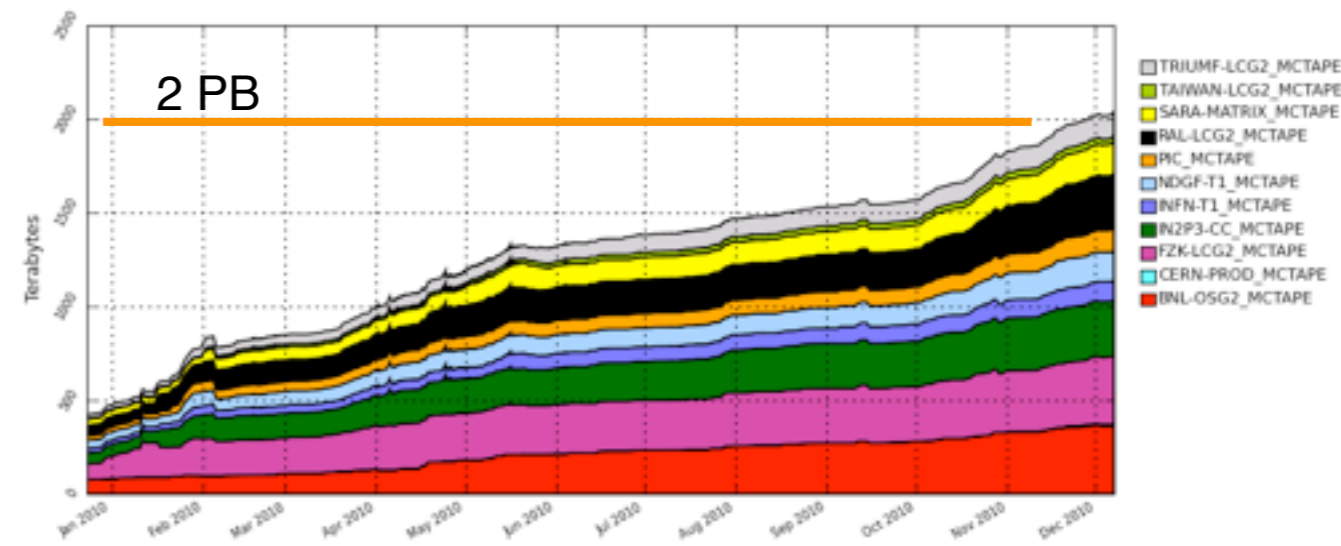
But from now on we will put more data onto tape.

- Tier-1s should foresee more tape access.

DATATAPE

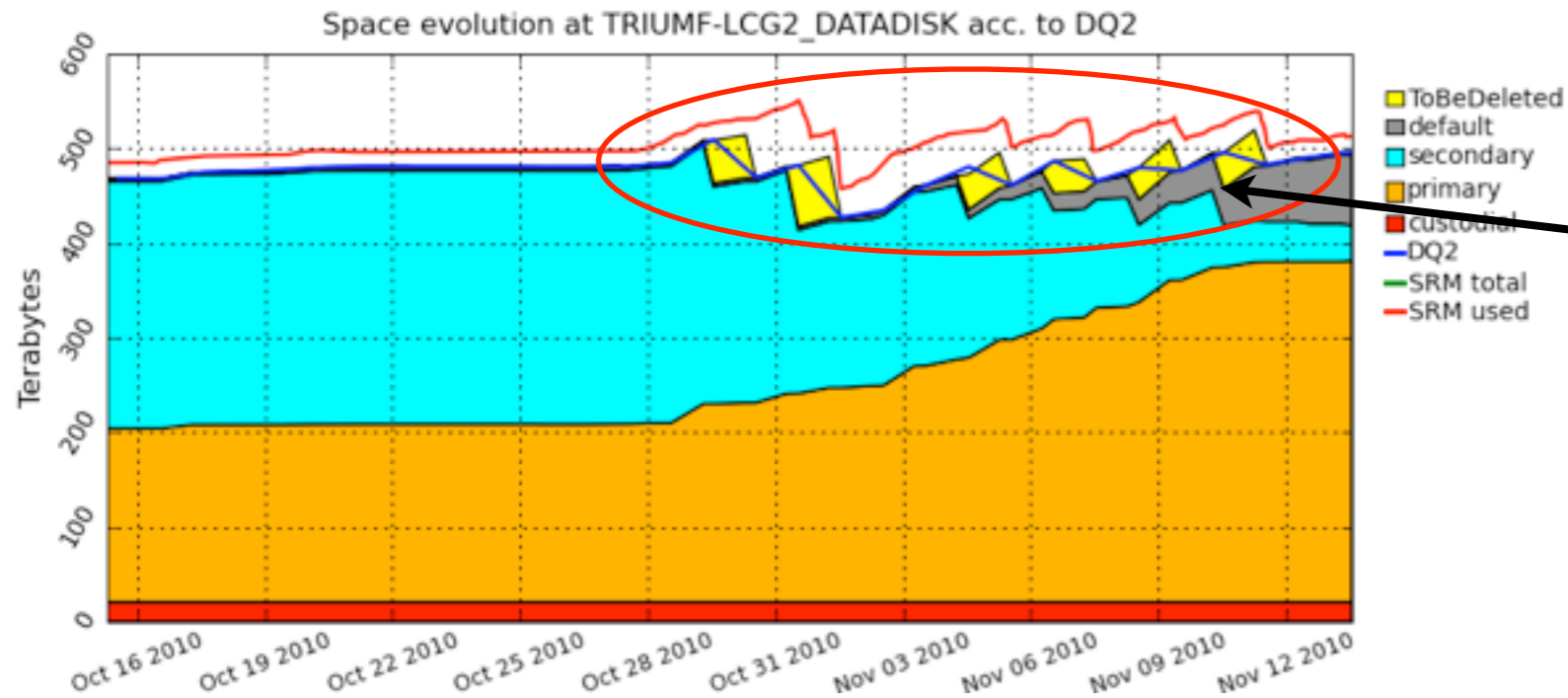


MCTAPE



Auto-Cleaning

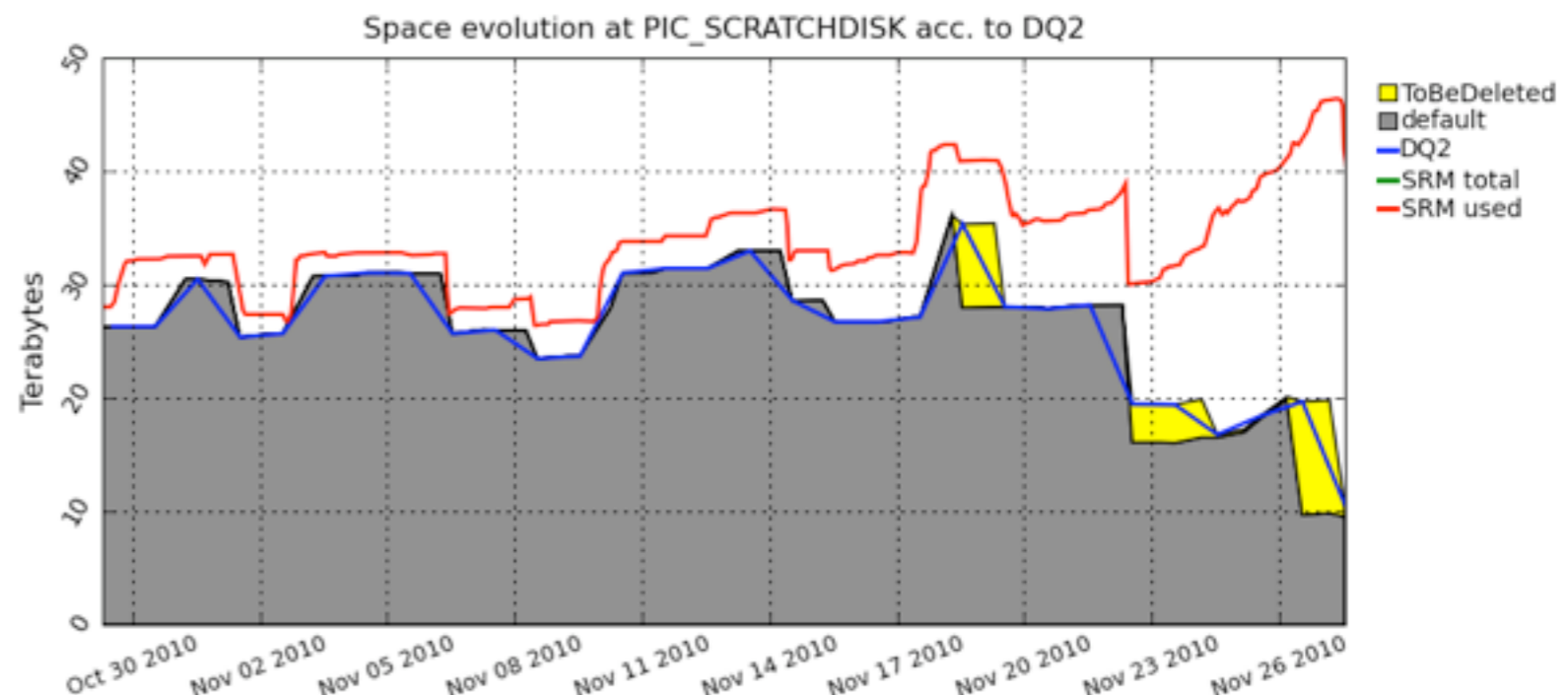
Now we have automatic cleaning agent



Auto-cleaning usually works fine as long as there are secondary data

Although in some cases where physical file deletion is stuck or slow, it cleans too much.

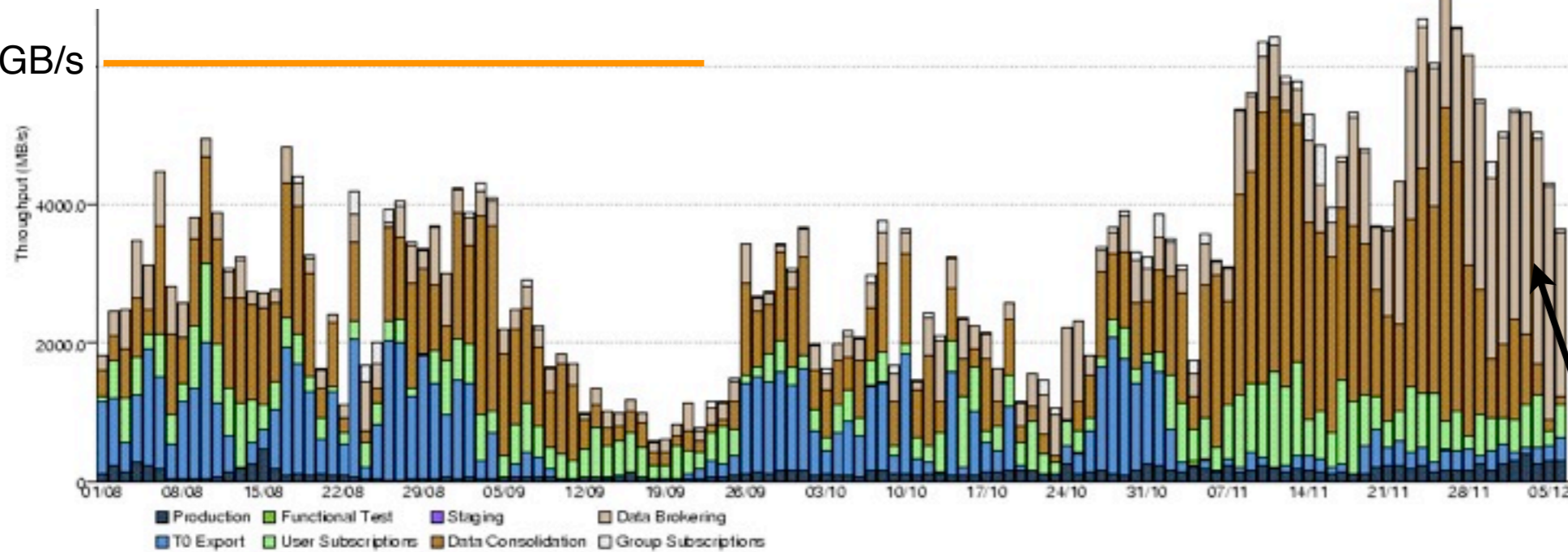
to be fixed soon.



Dynamic data placement

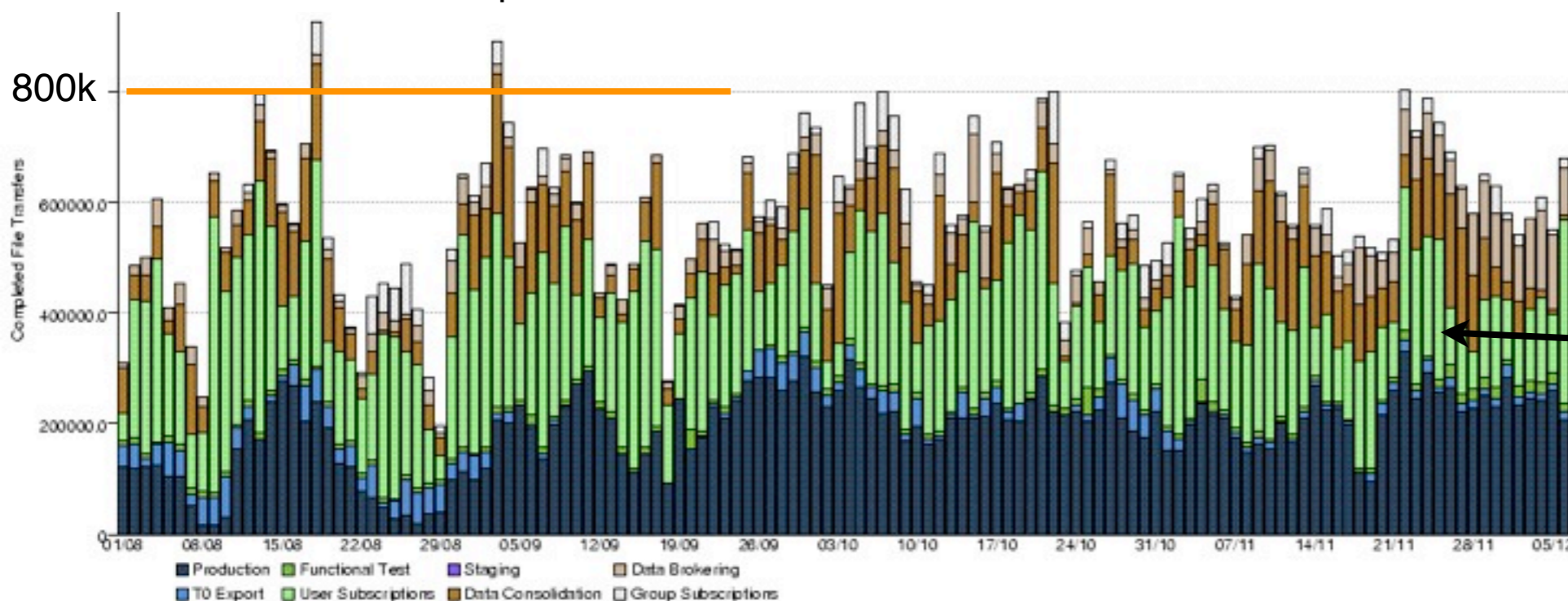


Throughput (MB/s)



- Group Subscriptions
- Data Brokering (Dynamic Data Placement)
- Data Consolidation (T1-T1, T1-T2)
- User Subscriptions
- T0 Export
- Production

Number of File Transfers Completed



More dynamic data placement activated (~all the clouds)

User data movement has been a major component in number of files

Dynamic data placement



Pre-defined distribution pattern cannot always follow the demand

Now we have more user/group data moving around

DaTRI: Data Transfer Request Interface

- On-demand Replication
- user/group dataset movement to the specified destinations

PD2P: Panda Dynamic Data Placement

- Replication triggered based on the distributed analysis statistics

Neither of them is perfect yet

- Rather chaotic
- Need to be tuned and better controlled

BDII — reconfirmation of ATLAS position



We are relying on BDII for ATLAS distributed computing operations

- We are relying on the services that use BDII (eg. FTS, WMS, SAM)
- We are using BDII to retrieve information needed by our applications
 - ▶ The information is cached in our “information system” (to be unified into AGIS = ATLAS Grid Information System)
 - ▶ This service discovery function is important.
- OSG takes an effort to place information concerning OSG sites in BDII and to keep it updated.

If in future this information can be retrieved from other sources, we will be happy to do so.

For the moment we need BDII supported.

Information system – another use case



ATLAS would like to have a place to get the information about how much data is stored on disk.

- If BDII is the place, we need this implemented and supported.

A slide titled "ES Conclusions" from the CERN IT Department. The slide contains a list of bullet points discussing data storage requirements and the feasibility of using SRM or BDII. The slide also includes contact information for the CERN IT Department and a CERN logo.

ES Conclusions CERN IT Department

- Back to requirements:
 - ATLAS would like to know how much data is stored in TxD1 space tokens
 - At least it would be good to know the size of the offline disks
- At the moment this seems not possible via SRM or BDII
- The BDII seems to offer more flexibility
 - But it needs to be agreed which info should be published and work on the information providers
- What is the right forum?

CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

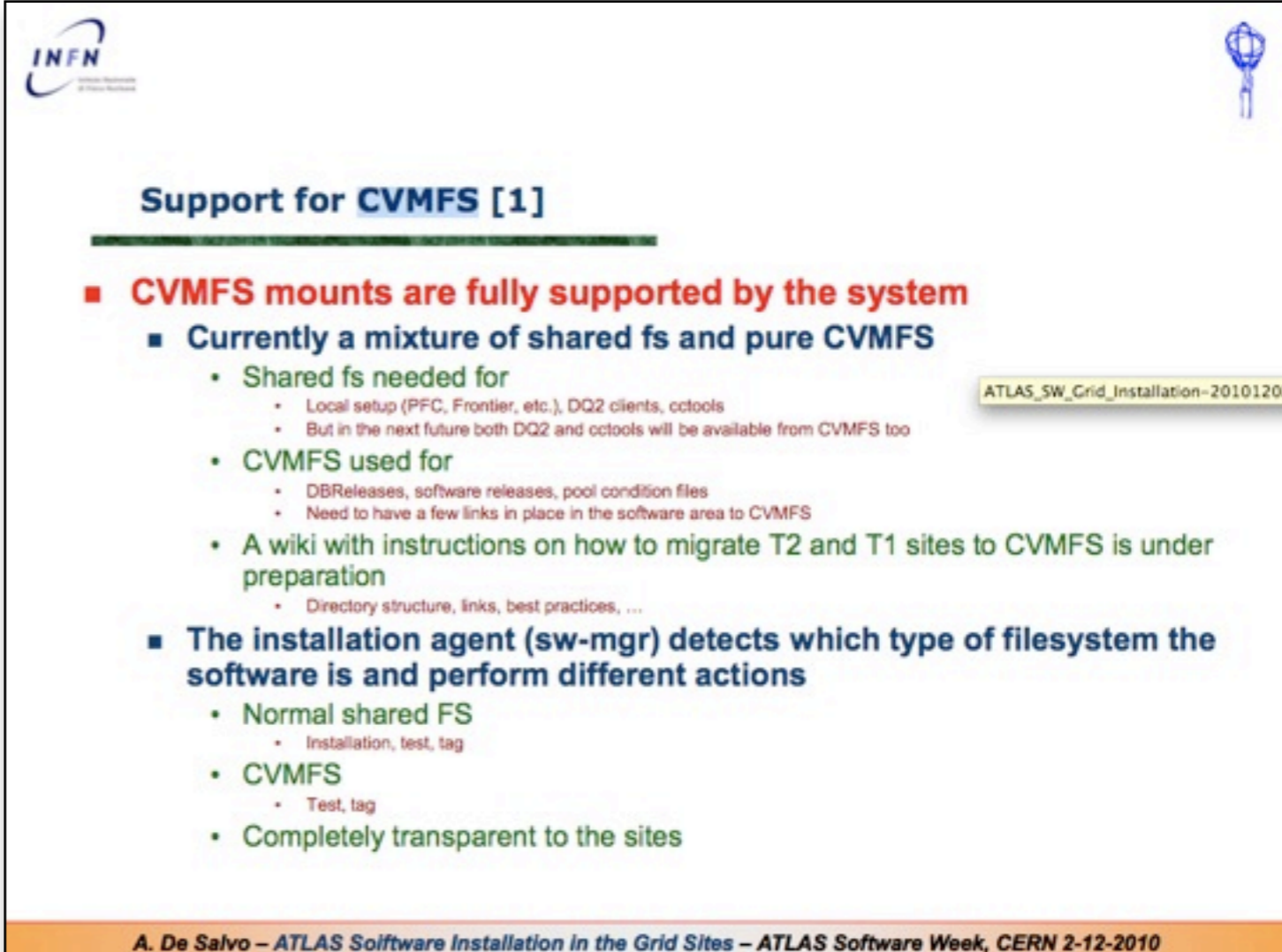
S. Campana, et al. GDB 08 September 2010

CVMFS

Validated for production and analysis in ATLAS

- Used at RAL, Wuppertal, QMUL (plus T3s)
- Looking for more volunteers

Supported in the ATLAS SW installation system



Support for CVMFS [1]

- **CVMFS mounts are fully supported by the system**
 - **Currently a mixture of shared fs and pure CVMFS**
 - Shared fs needed for
 - Local setup (PFC, Frontier, etc.), DQ2 clients, cctools
 - But in the next future both DQ2 and cctools will be available from CVMFS too
 - CVMFS used for
 - DBReleases, software releases, pool condition files
 - Need to have a few links in place in the software area to CVMFS
 - A wiki with instructions on how to migrate T2 and T1 sites to CVMFS is under preparation
 - Directory structure, links, best practices, ...
 - **The installation agent (sw-mgr) detects which type of filesystem the software is and perform different actions**
 - Normal shared FS
 - Installation, test, tag
 - CVMFS
 - Test, tag
 - Completely transparent to the sites

ATLAS_SW_Grid_Installation-2010120

A. De Salvo – ATLAS Software Installation in the Grid Sites – ATLAS Software Week, CERN 2-12-2010

Data Placement in 2011

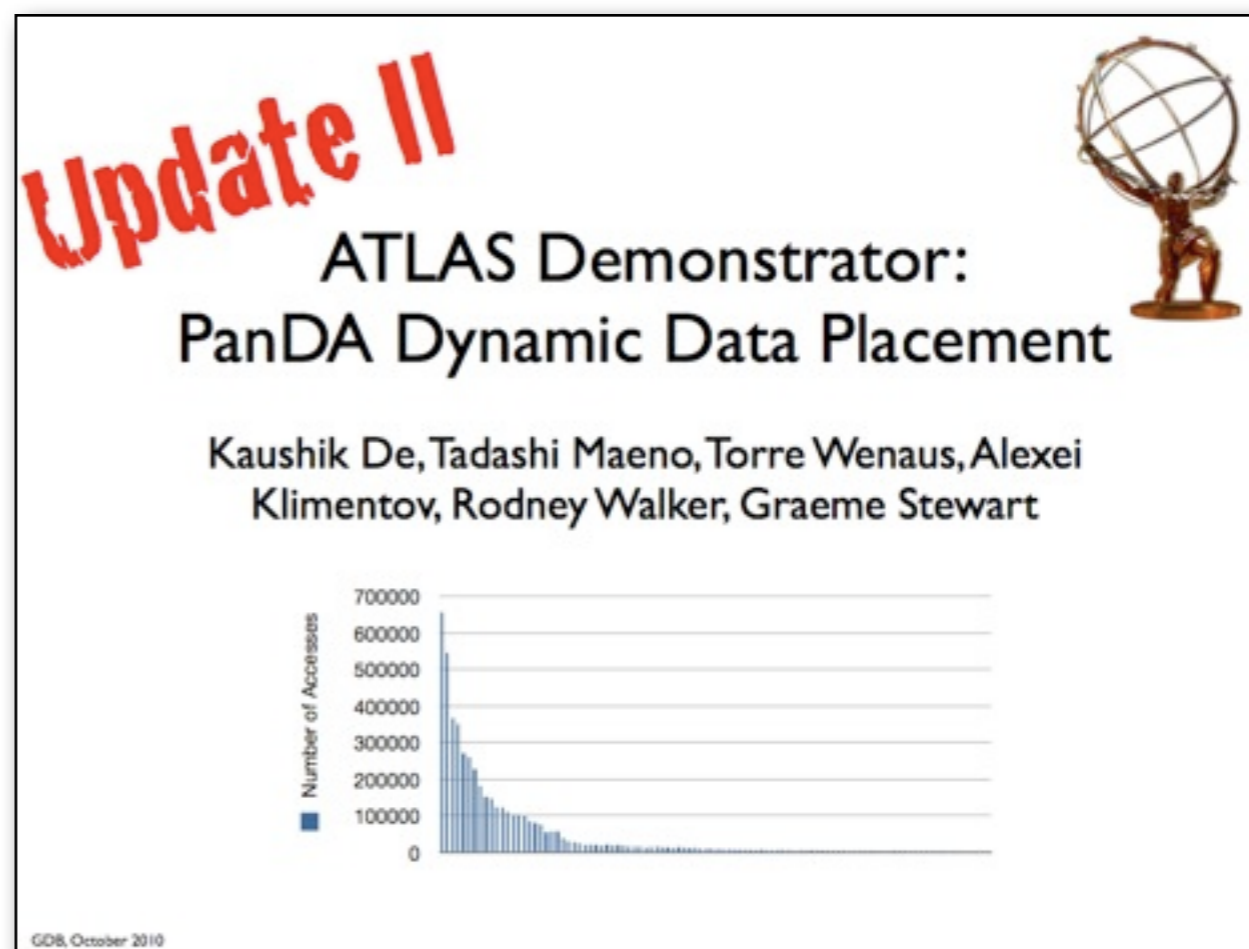


ATLAS is foreseeing much more data in 2011 than in 2010;

- We will put much more data onto TAPE than this year.
- More dynamic data placement than the current static pre-defined placement.

Expecting less traffic, but

- could be more chaotic
- yet to be tuned and to be controlled



G. Stewart, et al. GDB 13 October 2010