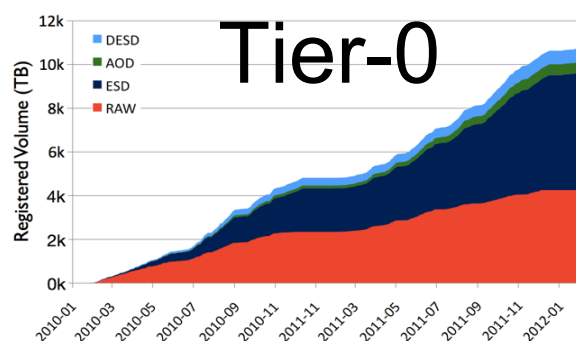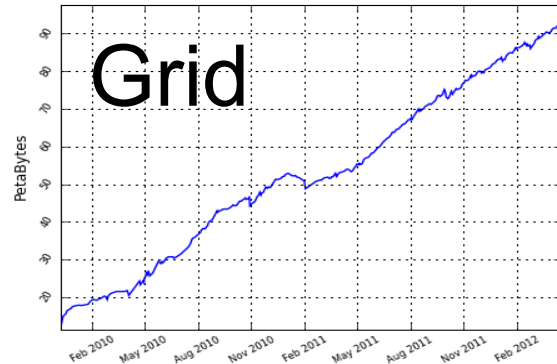# ATLAS Distributed Computing Operations:
Experience and improvements
after 2 full years of data-taking

Data management

Tier-0

Grid
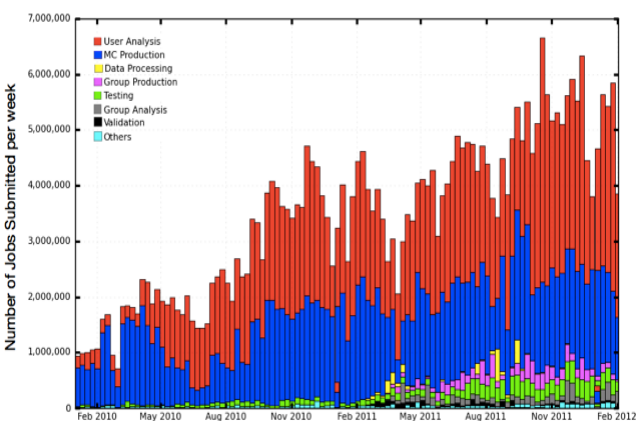
* 3.5 billions events collected
and processed by 5k cores at Tier-0
* RAW compression introduced in 2011
(gain factor ~2)

* Constant creation of data
* Number of files 0.5 → 2.5 million

Job activity

* Submitted job rate increased by factor 6
   * Main contribution from User Analysis
* Number of job slots multiplied by factor 2

* New activities
   1) Group production (Physics and performance groups' data analysis)
   2) Permanent site validation (Ensure that jobs are not stuck at problematic clusters)

Data distribution

Tier-1

pp 2010 reprocessing

Tier-2

pp 2010 reprocessing

More data at T2s
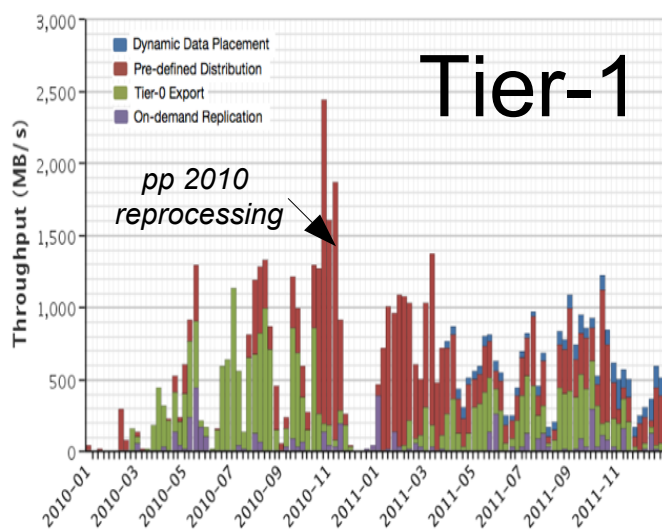
* Changes in Data Replication policy which impact transfer rates
   * Dynamic Data Placement : Replication according to data popularity
   * No ESD in 2011 : 'Life without ESD'
* Since summer 2011, Tier-2 storage hosting larger fraction of derived data
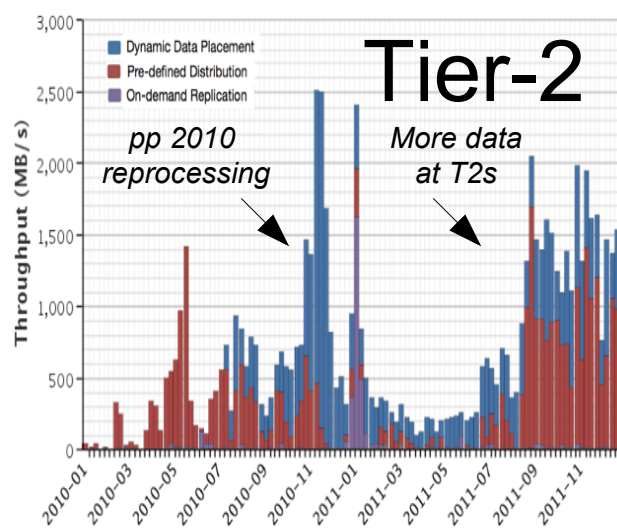
Breaking cloud model

Cloud : Tier-1+collection of associated Tier-2s (0-15)

task
job input
T1
T2 T2

2010
* Data collected at Tier-2 only from/through its Tier-1
* Tier-2 running only jobs for tasks assigned to its Tier-1
   → Hosted simulated data scales with its cloud CPUs
* Tier-2 LFC hosted at T1
   Tier-2 activity closely linked to Tier-1 availability

task
job input
T1 T1
T2 T2 T2

2011
* Identified Tier-2s well connected to many Tier-1s (T2D) :
   → Can transfer data directly from almost all Tier-1s
   Selection rule : 9 T1/T0 with file transfer rate > 5 MB/s (FTS report)

+ Faster transfers - less load on Tier-1 SE – Remove bottleneck
+ Tier-2s process data for most urgent tasks within many Tier-1s
   → Hosted simulated data scales with Tier-1 size

* LFC aggregation at CERN (still under way)

Site distribution

Europe >80
America >20
Africa : 2
Asia : 3
Oceania : 1
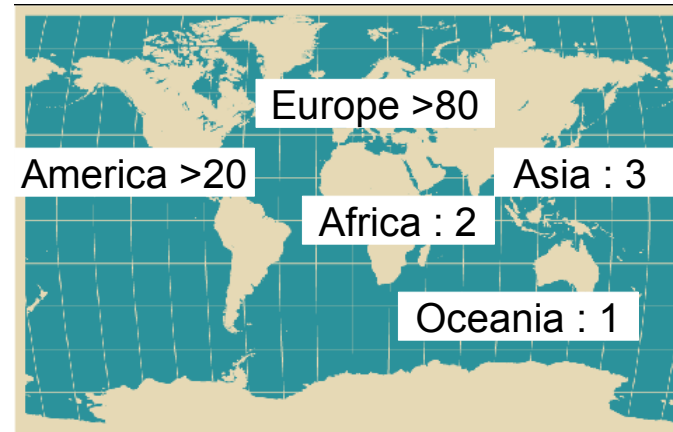
New in 2010-2011:
   * ATLAS Tier-3 Grid site:
      1) Definition of T3 operational model
      2) Creation/support of > 50 sites

CPU consumption of top 20 sites: 61%
→ Careful follow-up of big sites
(especially CERN+Tier-1s) from central operations
+ other sites followed by cloud squads

Grid operation tools and shifts

* The need for improved monitoring tools highlighted in ATLAS Grid survey
* Automatic blacklisting reduced amount of manual operations (still pending automation of announcement to site)
   * Reduced load on expert
   * Reduced human errors
   * Easier to spot fundamental issues

T2 site subset
In production
Under testing
Site in downtime
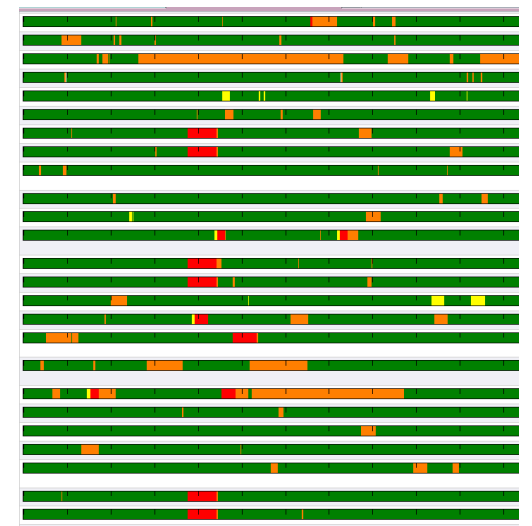
* First level support:
   * Shift in ATLAS control room: Processing of Tier-0 data + export to Tier-1s
   * Remote ATLAS Grid shifter + expert : Grid processing and data distribution
   * Remote Distributed Analysis shifter: User support on the Grid
* Expert support
   * Cloud support : ATLAS contact and support for sites within cloud
      * Follow pending issues and technology migration (CE → cream, LFC, ...)
   * AMOD : Contact for WLCG Operation and highest level of decision

Frontier/cvmfs +squid

* **Database :**
   * At the beginning of 2010 : Only direct access → limited at 10 Tier-1s/Tier-0
   * Frontier : Data access optimisation and caching at Tier-1s and Tier-0
   * Squid : Database access from any squid to Frontier
         → Oracle necessary at few Tier-1s (5 sites) and Tier-0
      → Reconstruction can be run at any site with enough memory

* **Software deployment**
   * 2010 : Grid job installing each software release on a central disk server at each site
         → New releases deployed in few days
         Too much load on site's central software area and too much space required
   * cvmfs: Highly optimised read-only filesystem over http with strong caching
   + 1 stratum0 (CERN) + 3 stratum1 + squid cache at all sites
   + Automatic deployment of releases in a few hours and caching on squid+WN
   + No more scalability issues
   - Still require development for diskless WNs
   → end 2011 : Deployed in > 50% of Grid sites

Future challenges

* Consolidation of ATLAS Grid infrastructure :
   * Improve redundancy and reduce sensitivity to site instabilities
   * Better usage of storage/CPUs
* Include new technologies : cloud, new transfer protocols,...
* More flexibility to network availability
* Keep monitoring up to date with all evolution

G. Steward and S. Jézéquel, on behalf of ATLAS collaboration

ATLAS EXPERIMENT