



Shared IT Services for Higher Education & Research

Conference 2017



The Data Challenges of the LHC

Reda Tafirout, TRIUMF

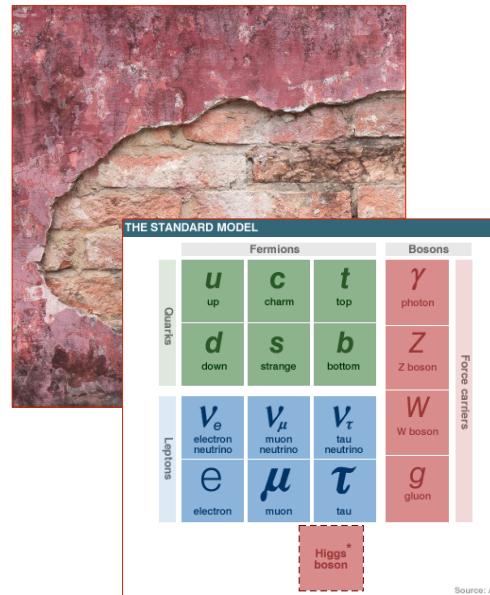
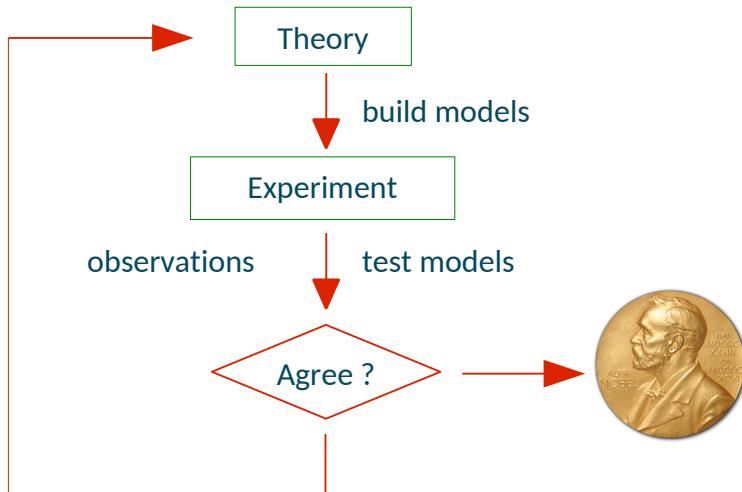


Outline

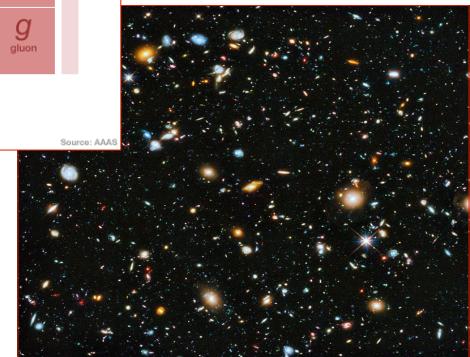
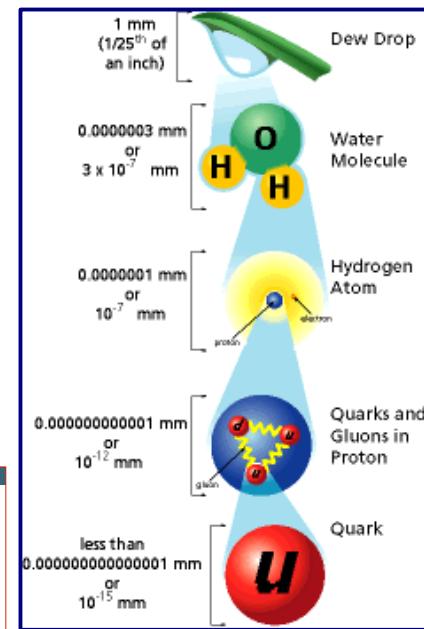
- LHC Science goals, tools and data
- Worldwide LHC Computing Grid
 - Collaboration & Scale
 - Key challenges
 - Networking
- ATLAS experiment distributed computing
 - Workload management system
 - Data handling & distribution
- Future outlook

Science goals

- Explore the fundamental nature of matter and the basic forces that shape our universe
- One successful theory: the Standard Model
 - basic building blocks and forces
- Perform precision measurements



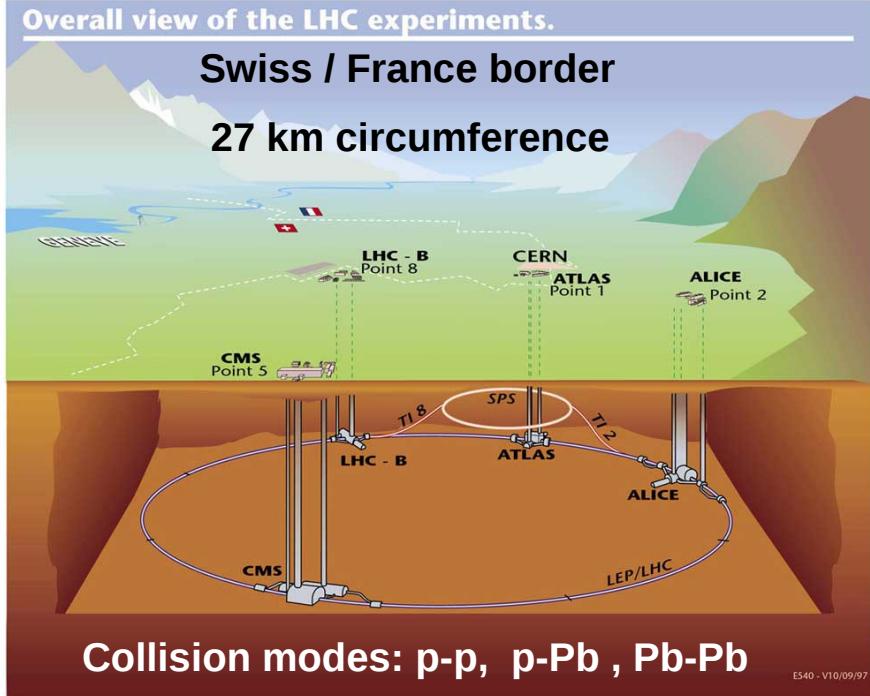
Matter substructure



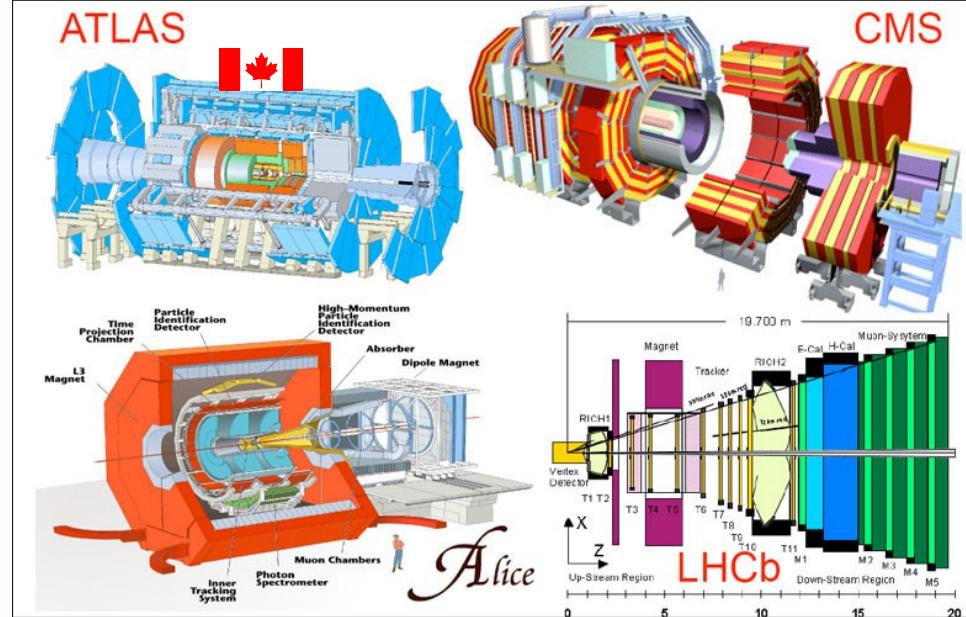
- Some of the important questions:
 - Anything else beyond the Standard Model ?
 - Nature of dark matter in the universe ?

Science Tools

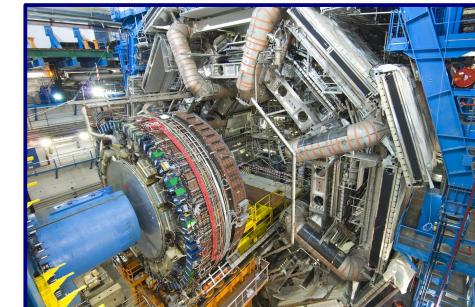
Powerful Particle Accelerator



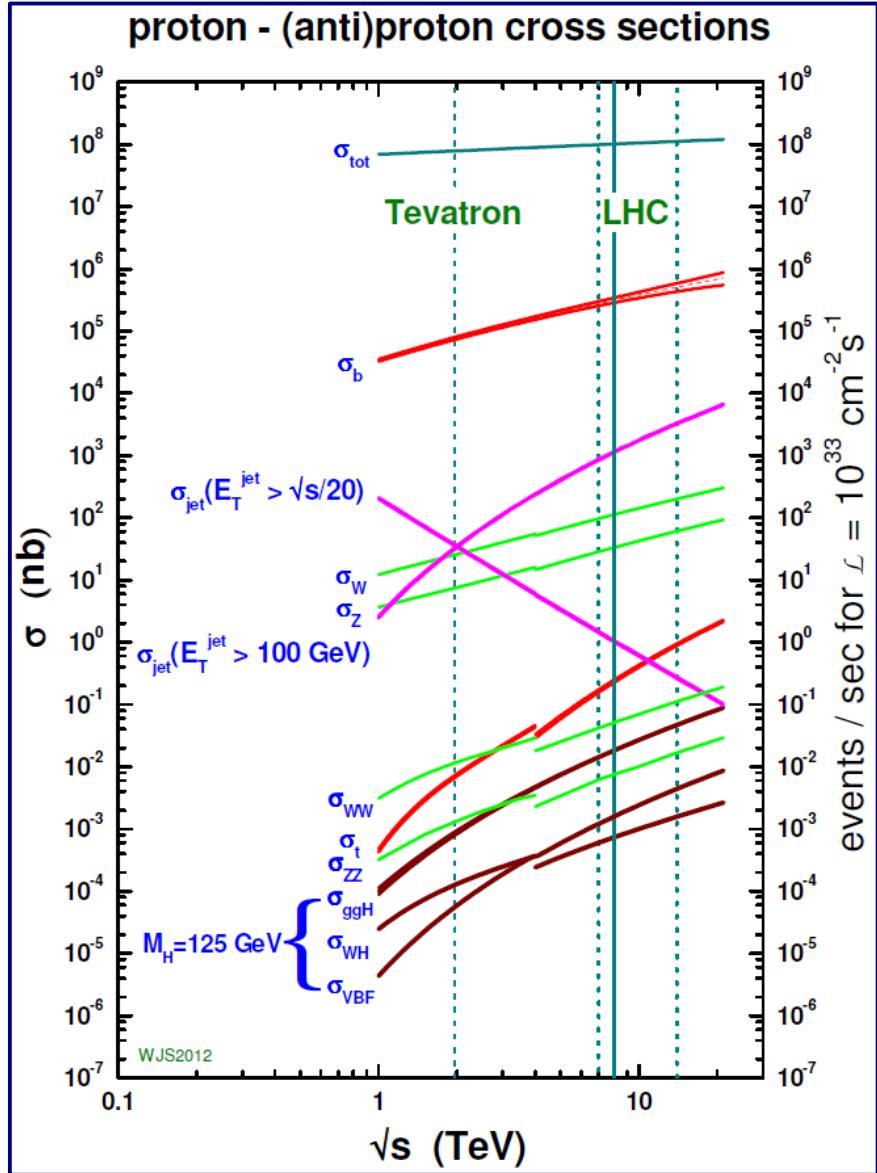
Sophisticated Detectors



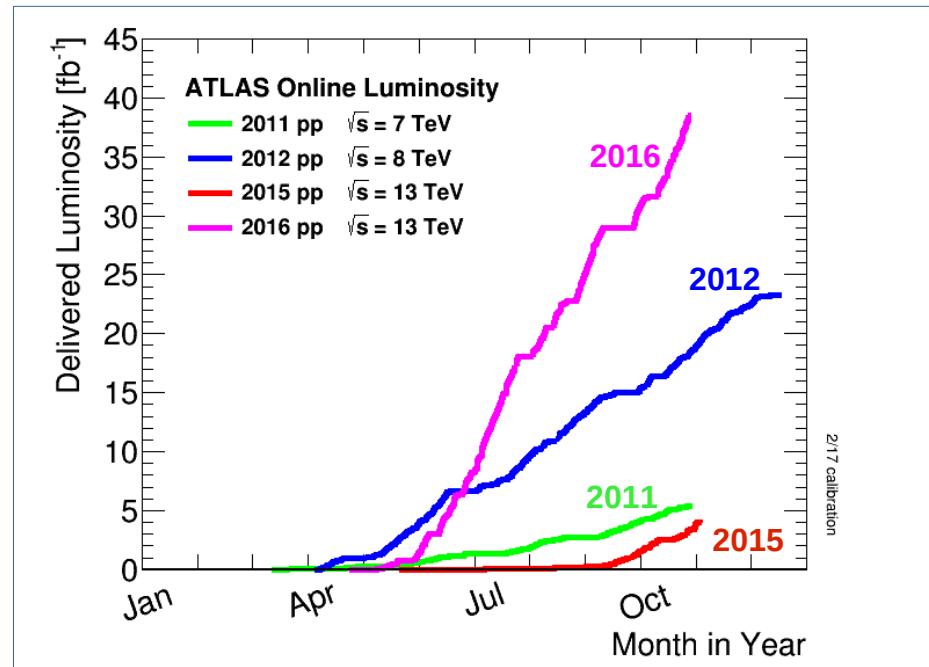
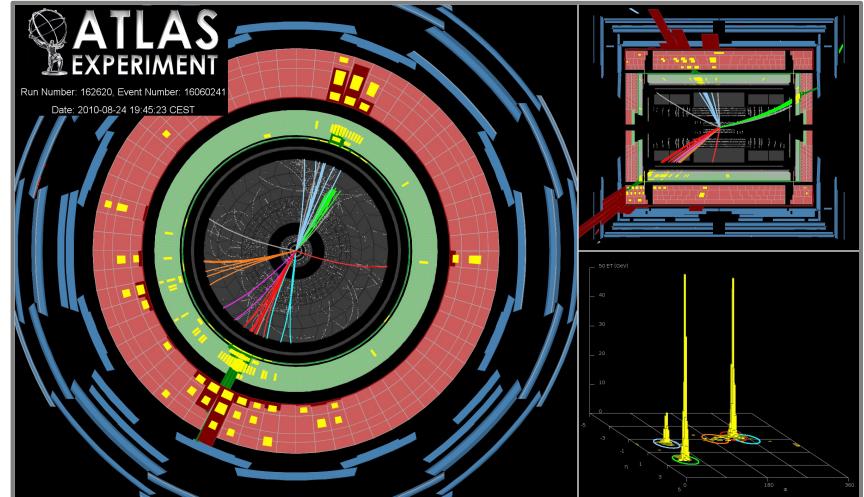
Large-scale Storage & Computing



Large Amount of data needed



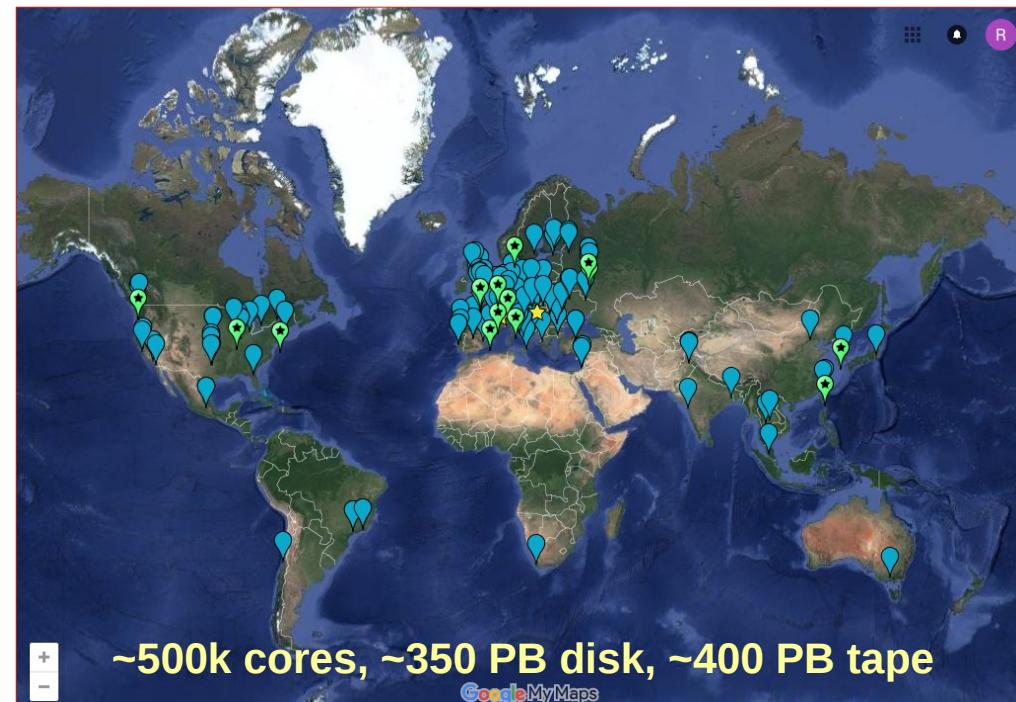
<http://www.hep.phy.cam.ac.uk/~wjs/plots/plots.html>



Worldwide LHC Computing Grid

- Global collaboration of 167 computing centres in 42 countries
- MOU between parties participating in the project
 - CERN (Tier-0)
 - Tier-1 centres
 - Tier-2 centres
- Description of roles:
 - Baseline services & up-time
 - Minimum Service Levels
 - Response times
 - Grid operations
 - Pledged resources

<http://wlcg.web.cern.ch>



- Additional capacity being deployed for 2017 LHC operations



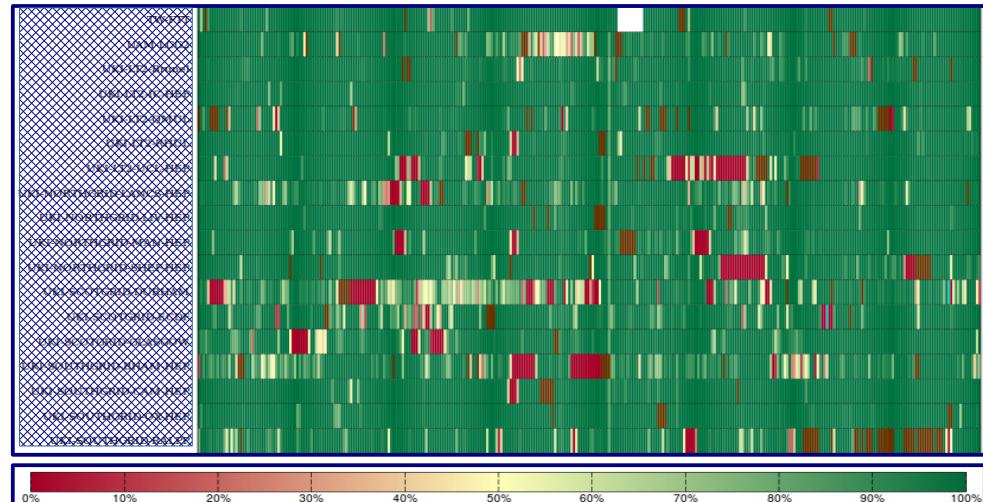
Tier-1: TRIUMF (Vancouver)

Tier-2: Compute Canada centres (Victoria, SFU, Toronto, McGill)

Key Challenges

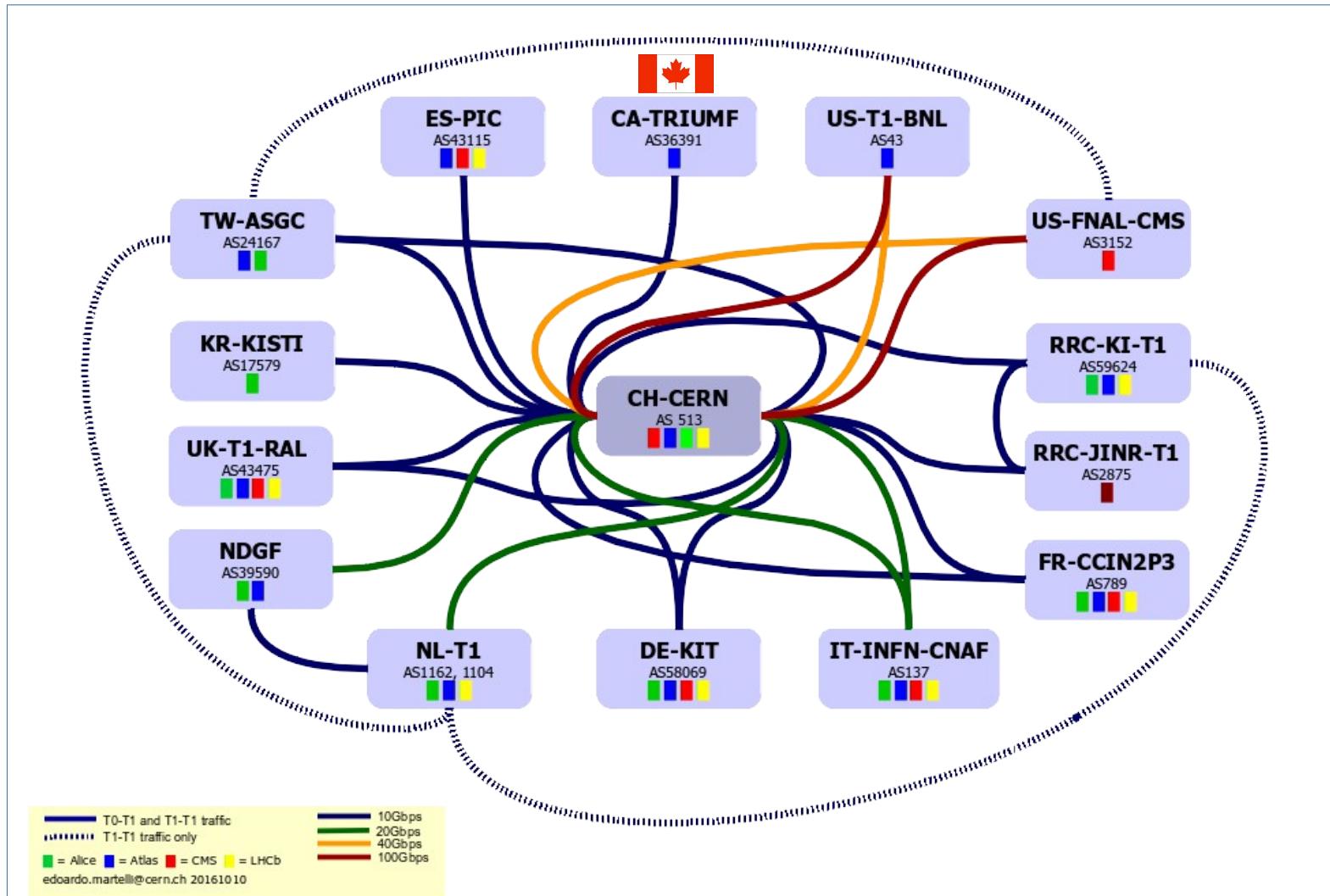
- Computing infrastructure is critical to the scientific output and success of the overall LHC experimental program:
 - Middleware (baseline services) & Interoperability
 - Availability & Reliability of each site
 - Performance & Scalability of each site
 - Distributed computing operations: workload and data management systems
 - Security
 - Network performance
 - Pledged resources
- Optimize physics output & costs

Site Usability Monitor (*sample of 18 sites*)



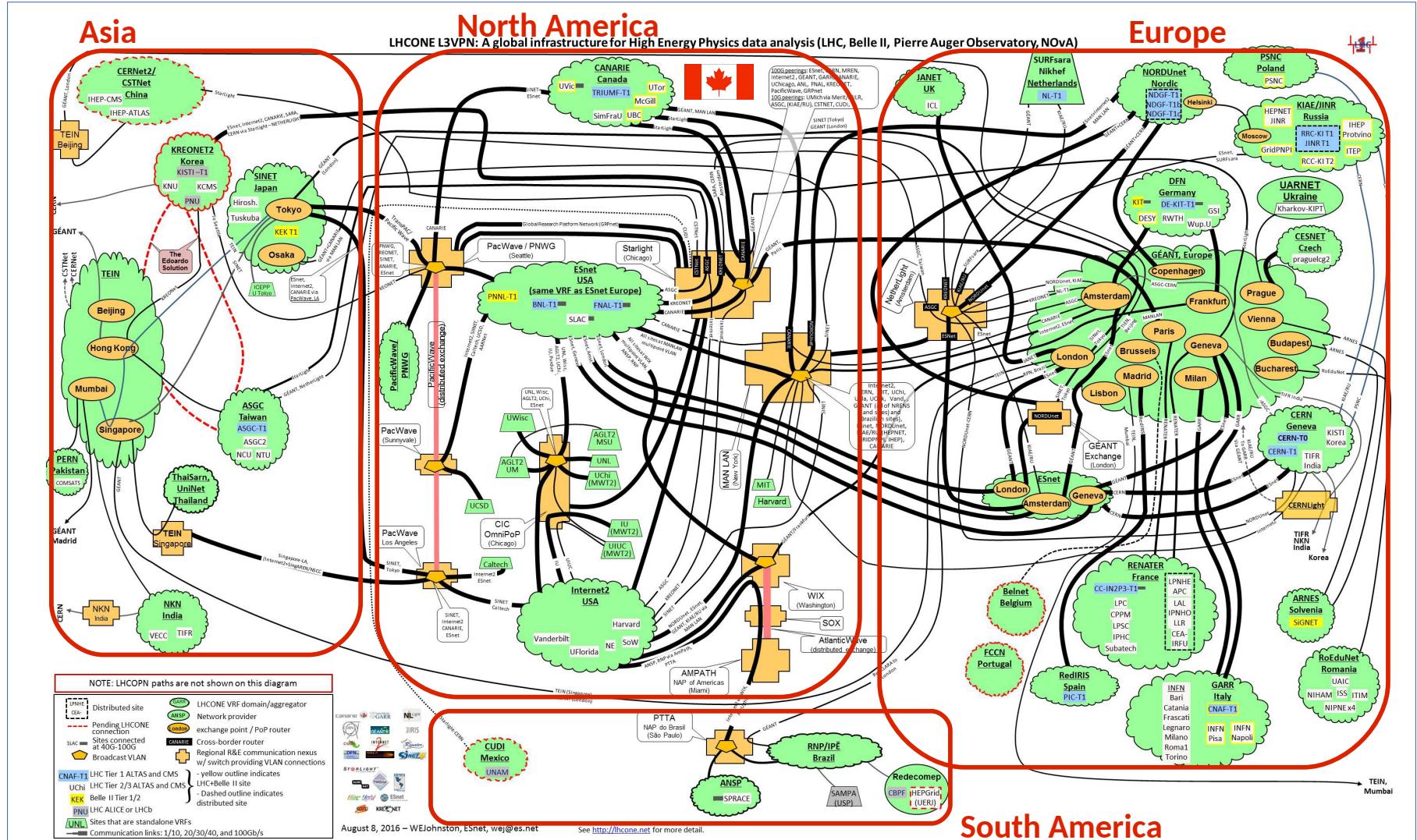
Network Connectivity (I)

- LHCOPN: LHC Optical Private Network, only for Tier-0 & Tier-1 centres



Network Connectivity (II)

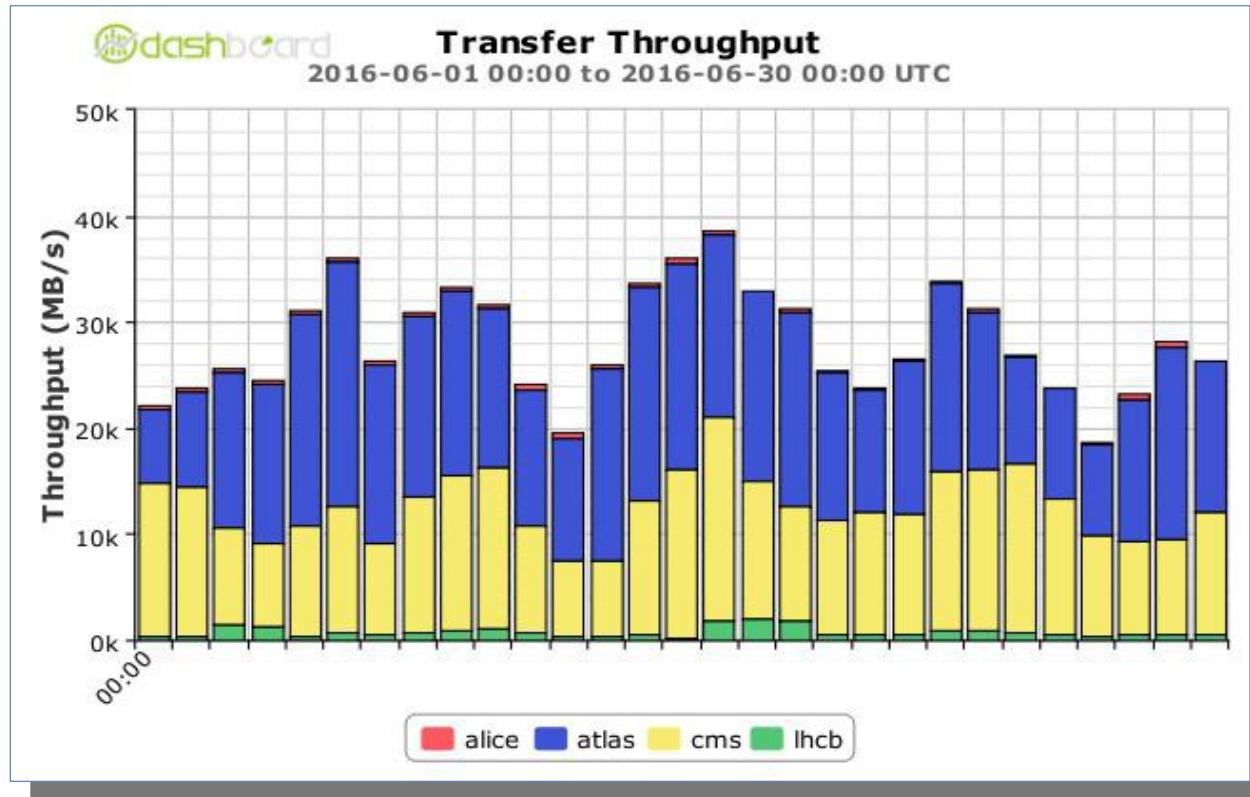
- LHCONE: LHC Open Network Environment (for Tier-2 centres + T0 / T1)



Global Network Traffic

- In 2016, more than 800 Petabytes moved across the computing grid (from all LHC experiments combined).

~35 GB/s peaks of global WAN transfers



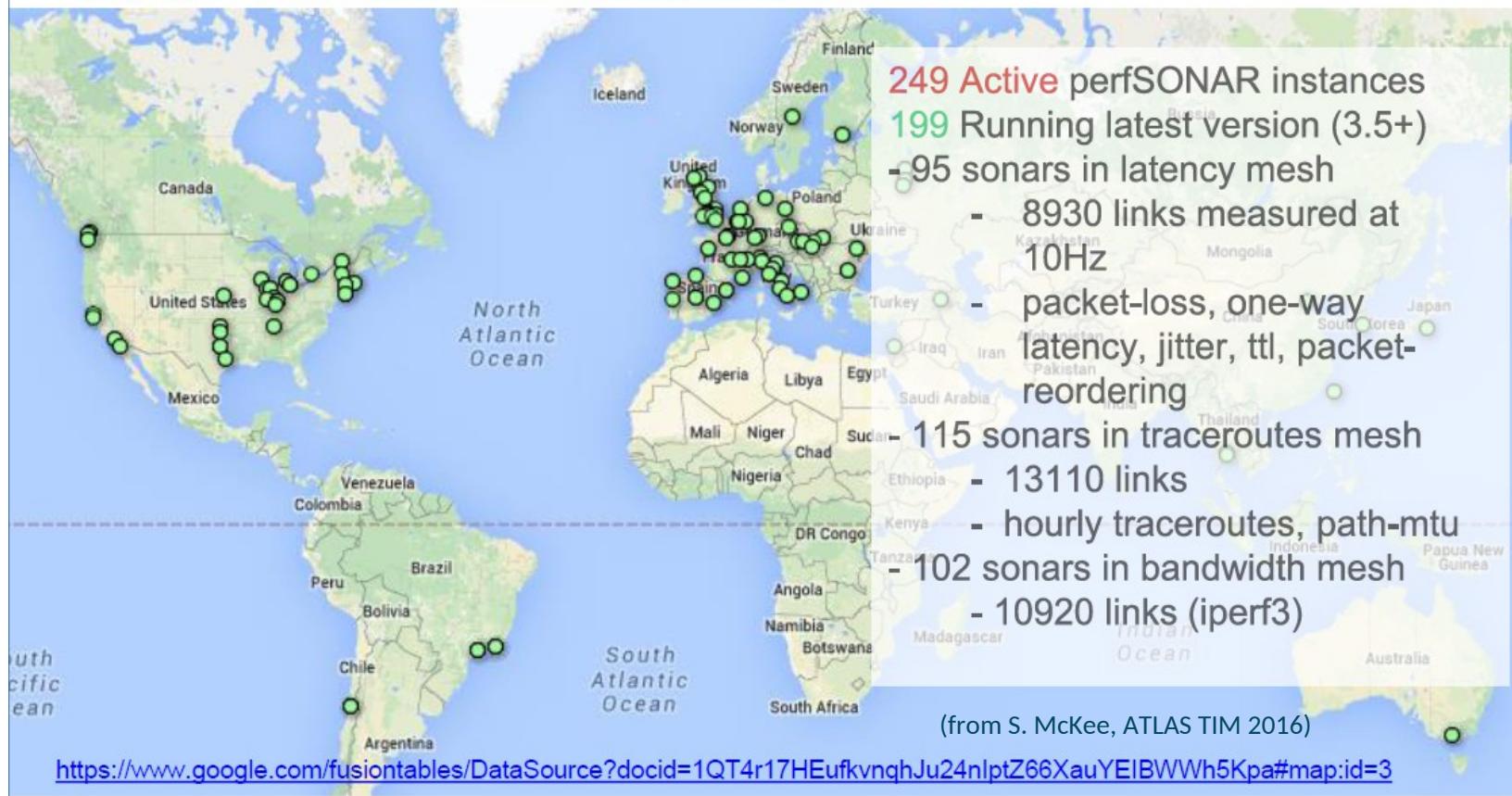
Shown for period
06 / 16

(from I. Bird, HSF Workshop, San Diego, 01/2017)

Network Monitoring (I)

- End-to-end network problems difficult to diagnose: multi-domain
- perfSONAR-PS deployment (for bandwidth & latency): two instances at each site

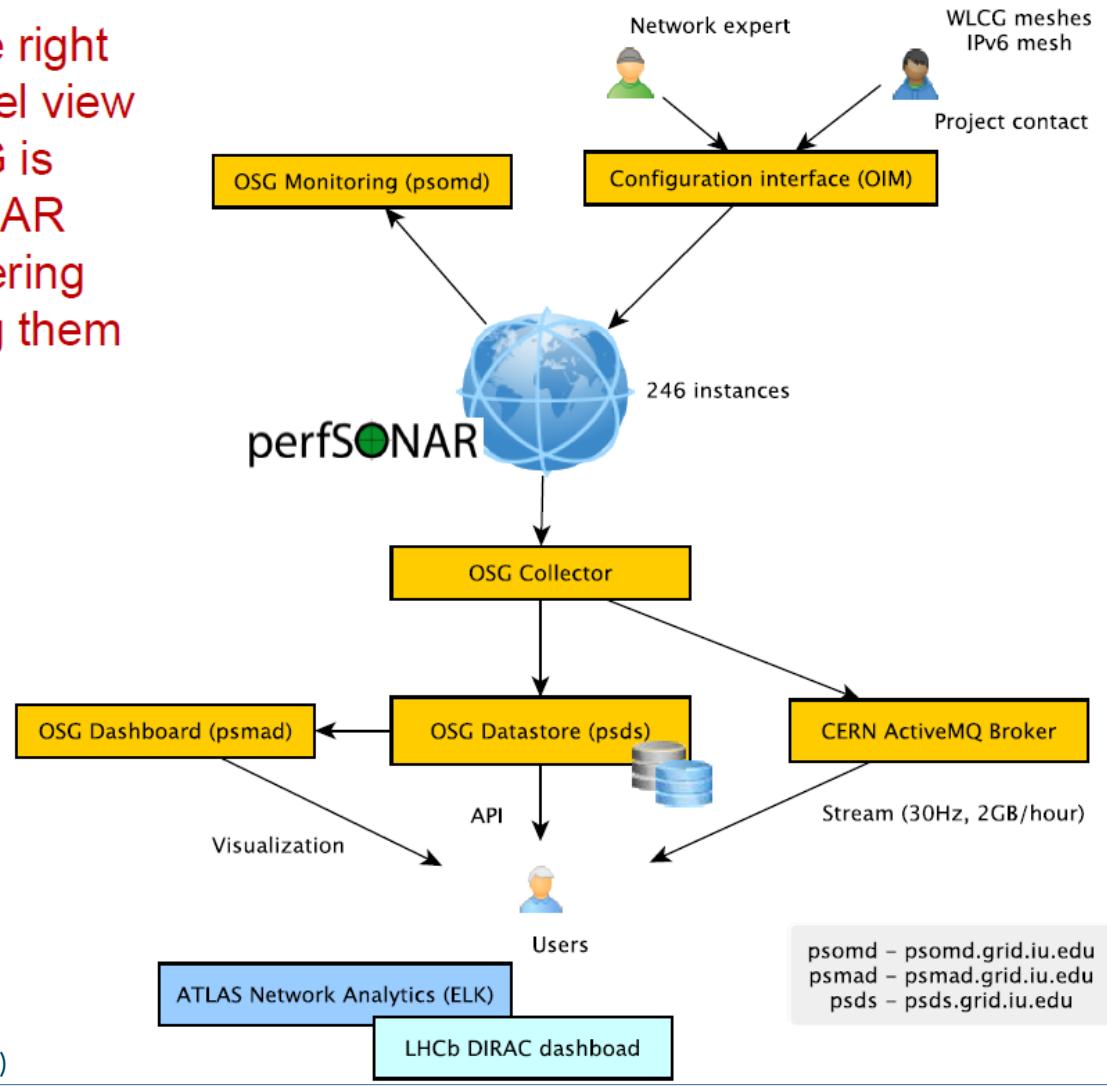
http://grid-monitoring.cern.ch/perfsonar_report.txt for stats



Network Monitoring (II)

- perfSONAR-PS configuration, data collection and dashboard

The diagram on the right provides a high-level view of how WLCG/OSG is managing perfSONAR deployments, gathering metrics and making them available for use.



Network Monitoring (III)

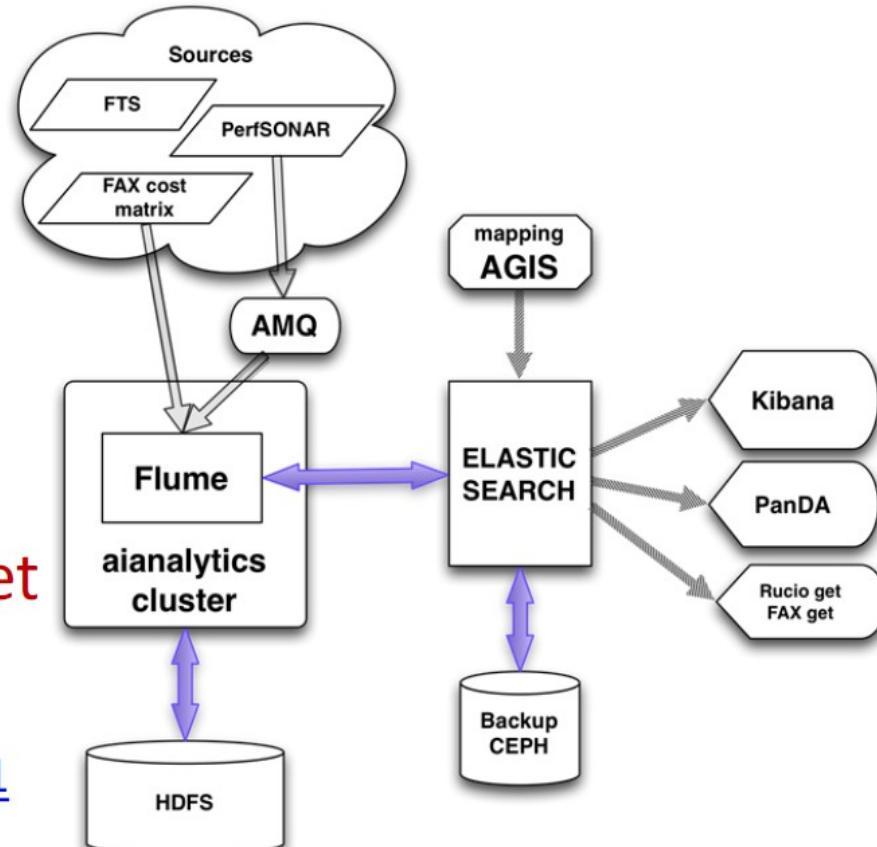
- Data analytics tools using different sources: cost matrix and throughput predictions to optimize work flows (for ATLAS)

- Diagram shows the flow

- End-to-end+perfSONAR data both available to jointly analyze

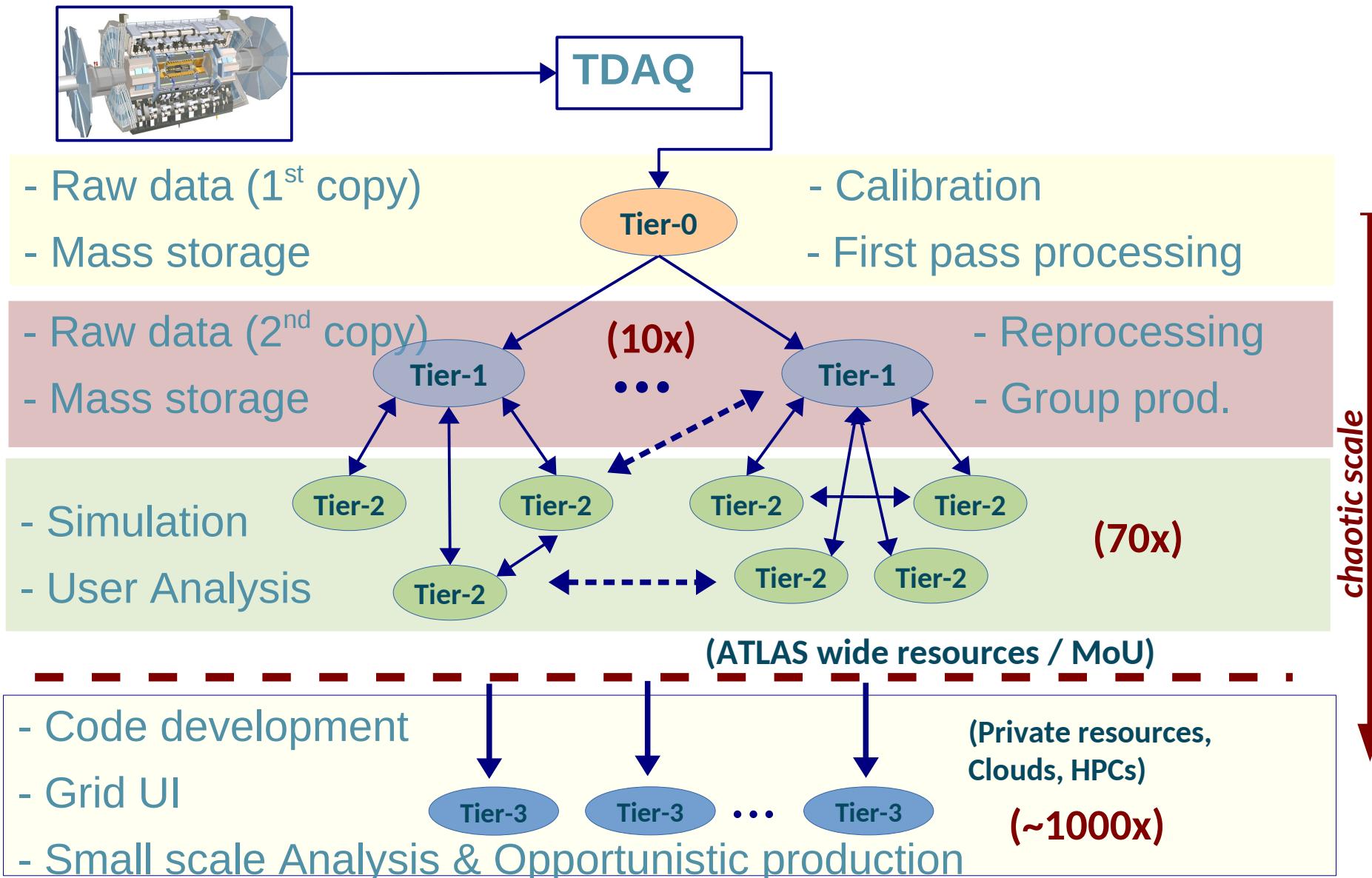
- Kibana can be used to get customized views

<http://cl-analytics.mwt2.org:5601>



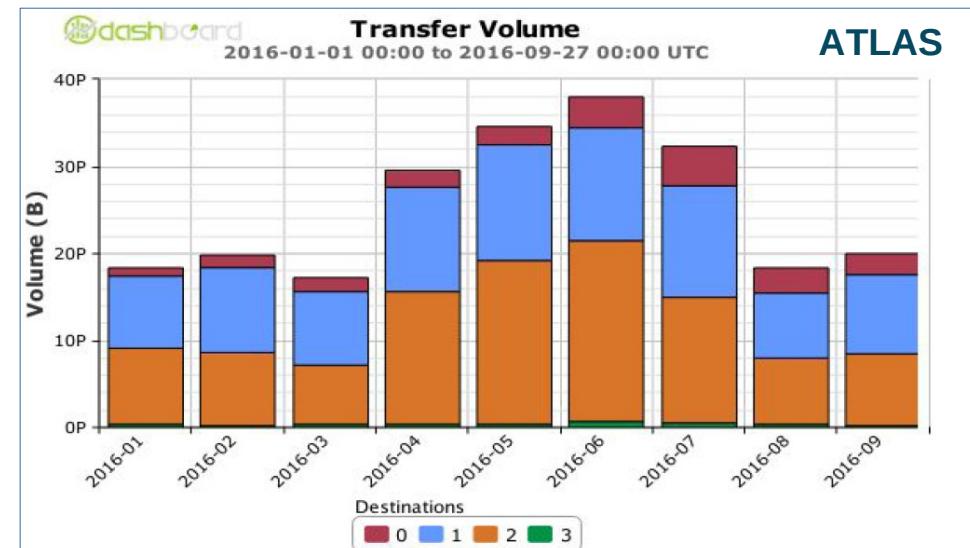
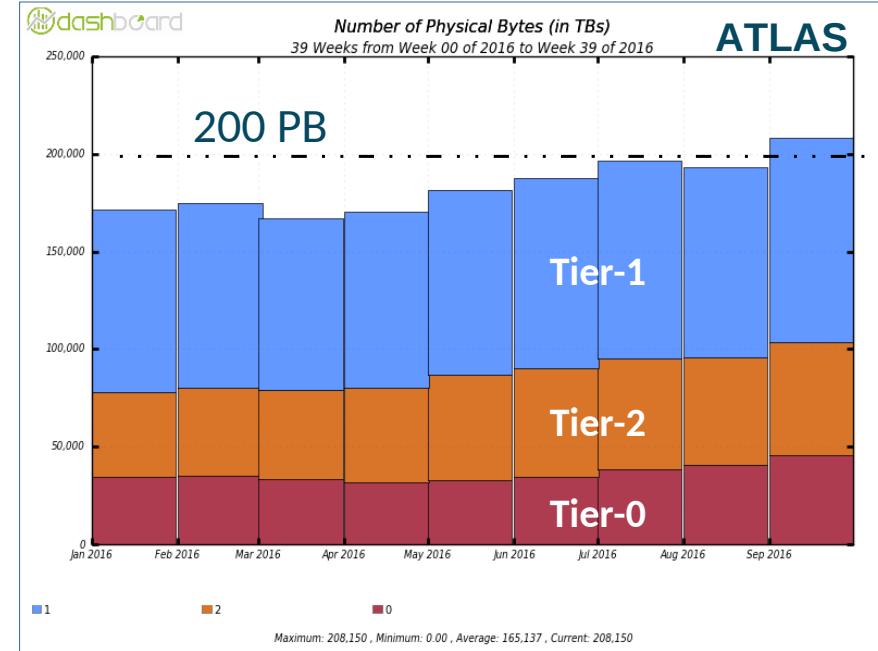
(from S. McKee, ATLAS TIM 2016)

ATLAS Distributed Computing



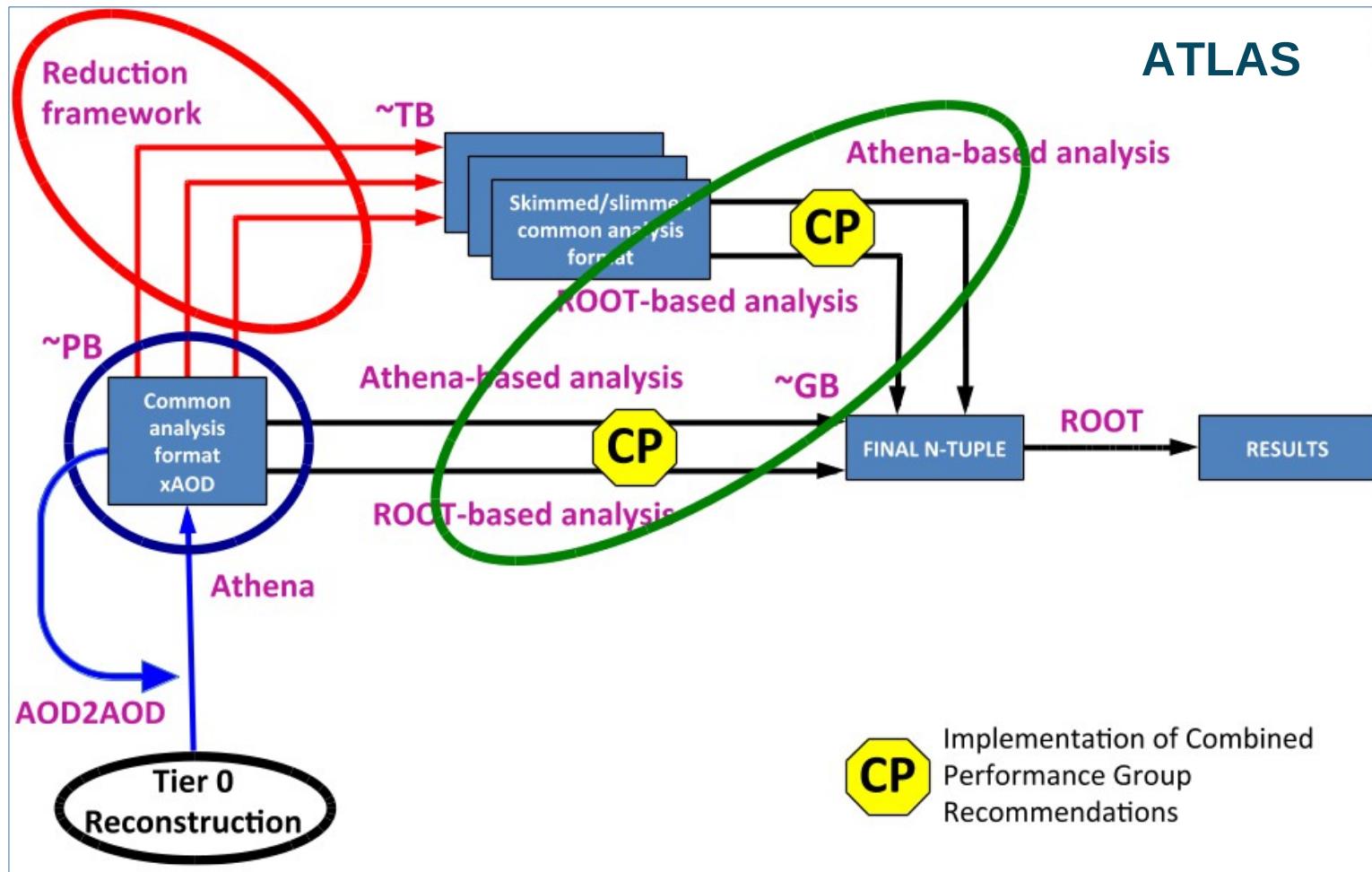
Data Management System

- ATLAS DDM system (Rucio):
 - datasets cataloging
 - replication & lifetime policies
 - primary vs secondary
 - disk vs tape
 - quotas & deletions
 - transfer protocols
 - hundreds of “storage elements”
- File Transfer Service (FTS)
 - 20-40 PB monthly (WAN)
 - network saturation seen
- Constant pressure on disk resources



Data Reduction & Derivation

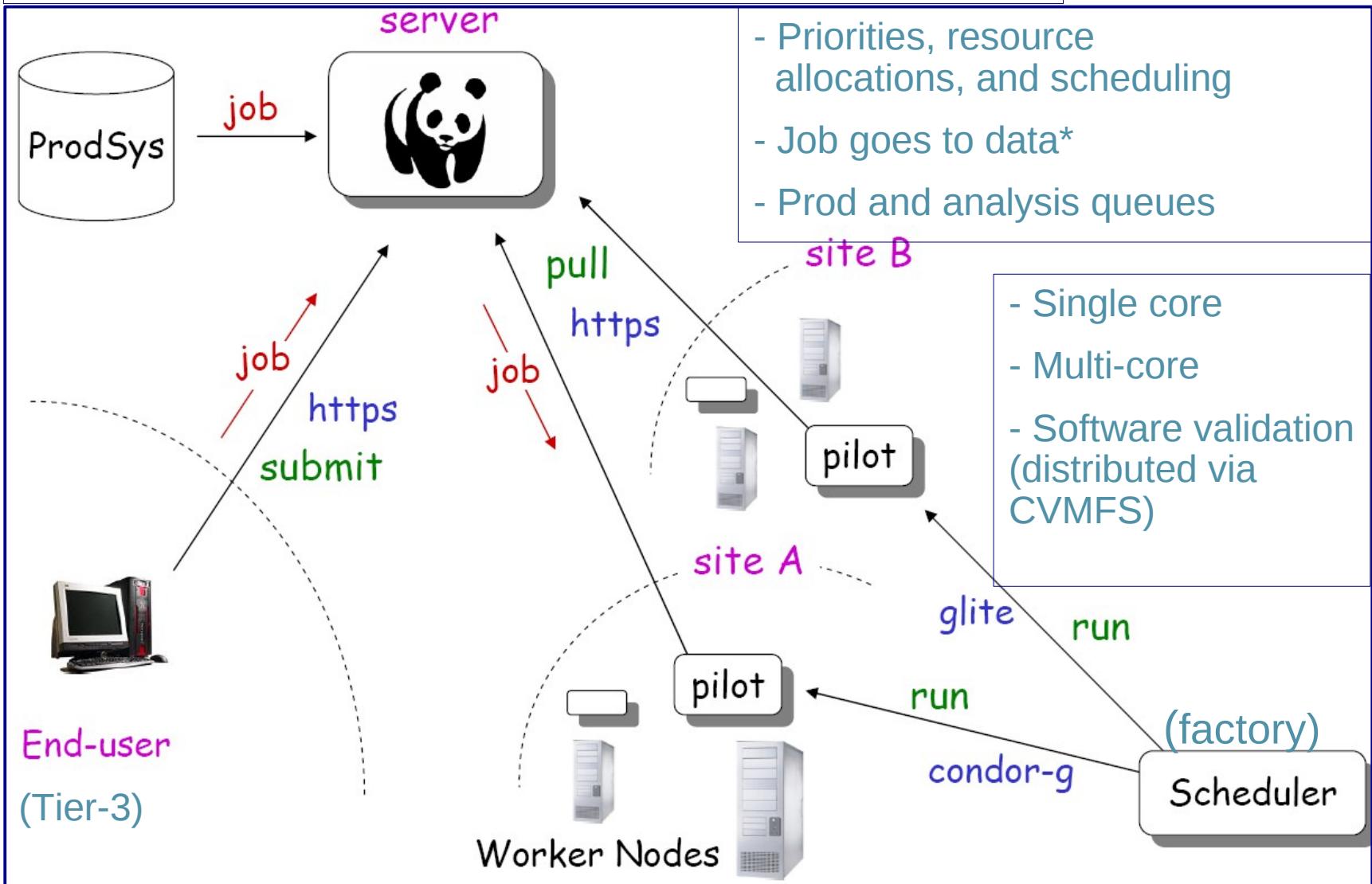
- Reduction & derivation framework for more efficient use of computing resources; less chaotic and more coherent approach for physics analysis.



Workload Management System

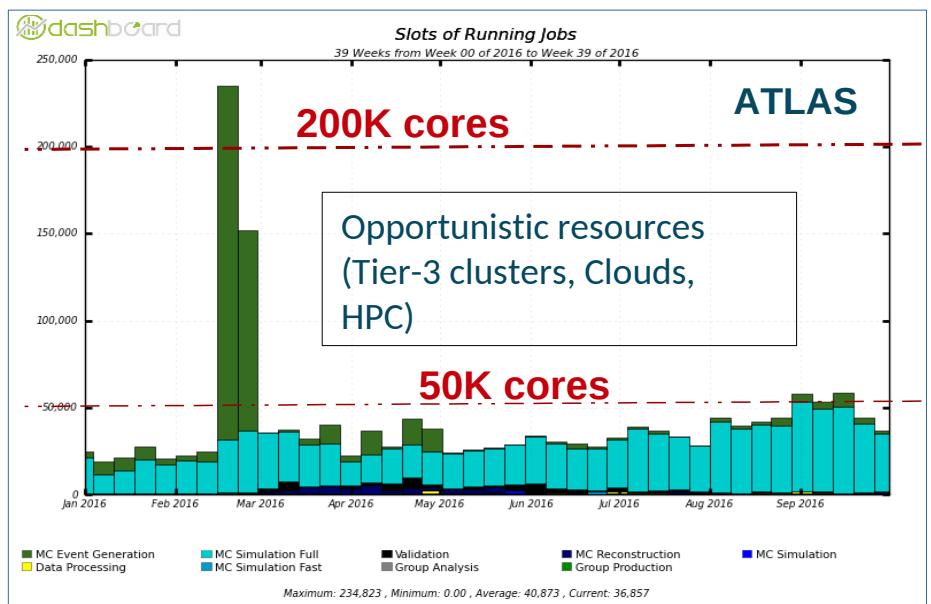
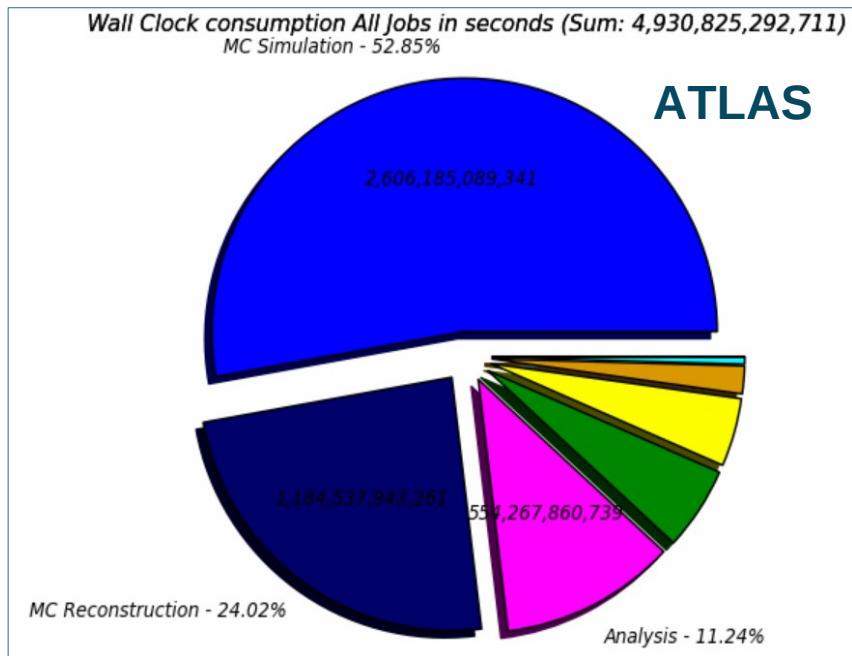
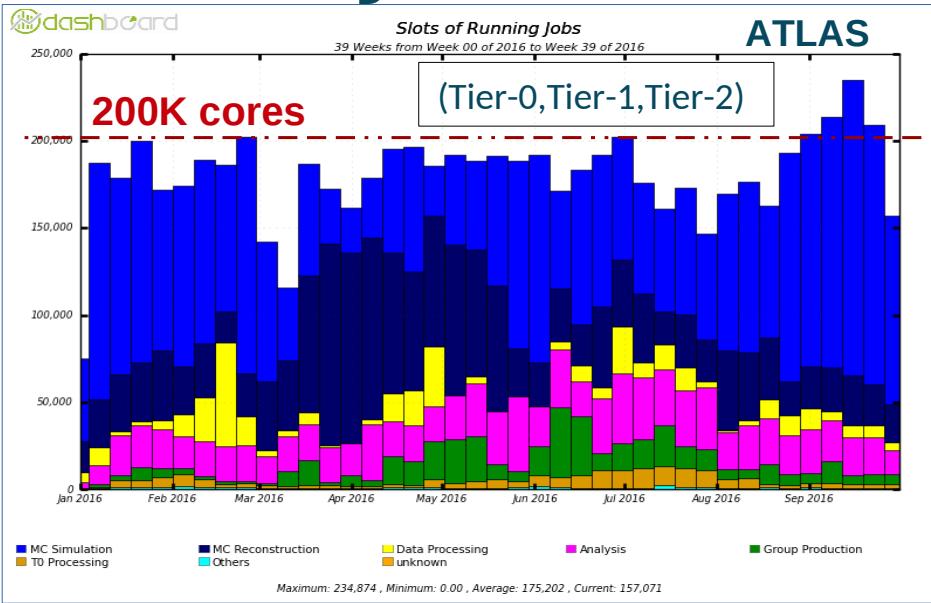
Production And Distributed Analysis (PANDA)

ATLAS



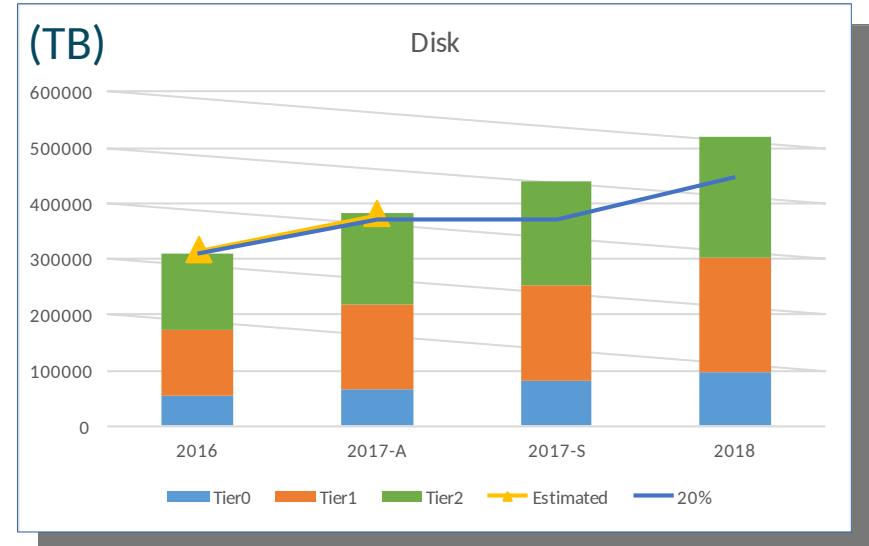
Production & Analysis

- Overall, smooth ATLAS distributed computing operations (ADC)
- For ~10 months in 2016:
 - 254 million jobs completed
 - 57 million cpu-days consumed
 - 89% production
 - 11% user analysis

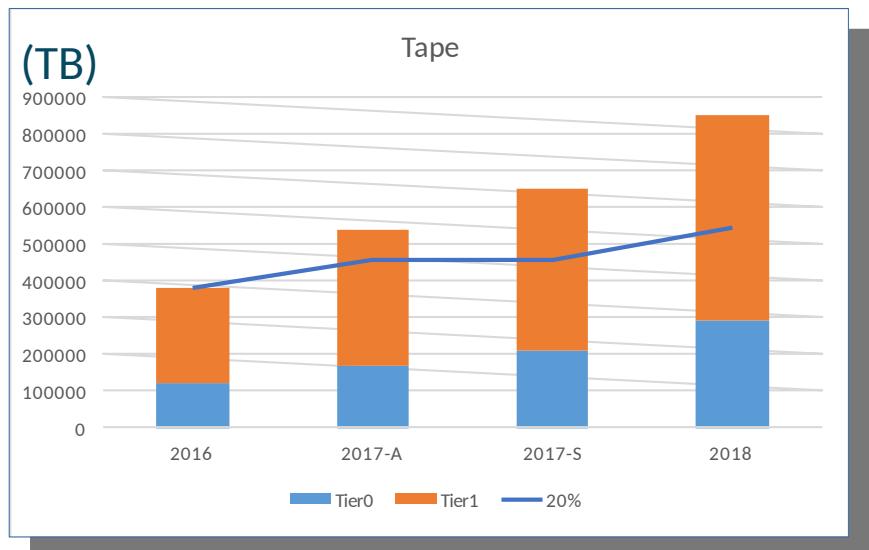
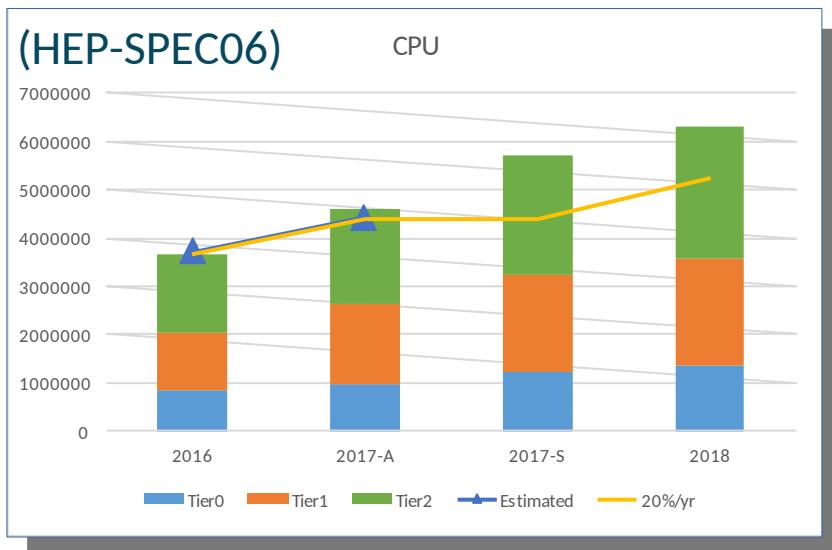


Future Outlook (I)

- Substantial increase due to excellent LHC performance in 2016
- Reviewed by the WLCG Computing Resources Scrutiny Group and approved by CERN Resources Review Board.
- For 2018 (all LHC experiments):
 - Disk: 520 PB
 - Tape: 840 PB
 - CPU: 6,200,000 HEP-SPEC06 (1 core ~ 14 HS06, benchmark unit)

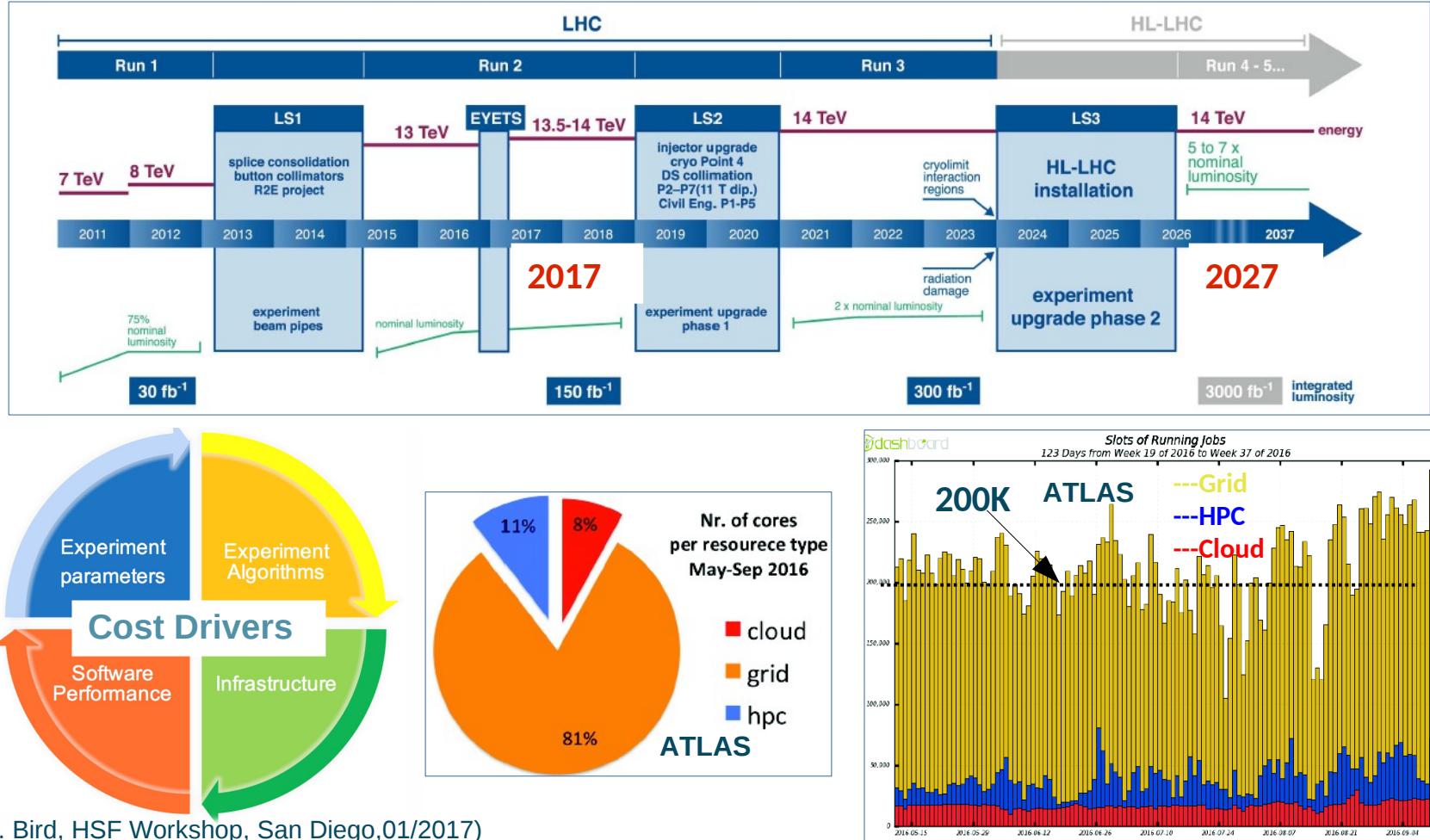


(from I. Bird, HSF Workshop, San Diego, 01/2017)



Future Outlook (II)

- LHC program has a very long time horizon
- Computing Resources requirements will increase by x10 for storage and by x50 for cpu (for HL-LHC run).



Extra / Additional Material

“Canadian” region / operations



Tier-1



ROC



Tier-2 West Federation

Tier-2 East Federation

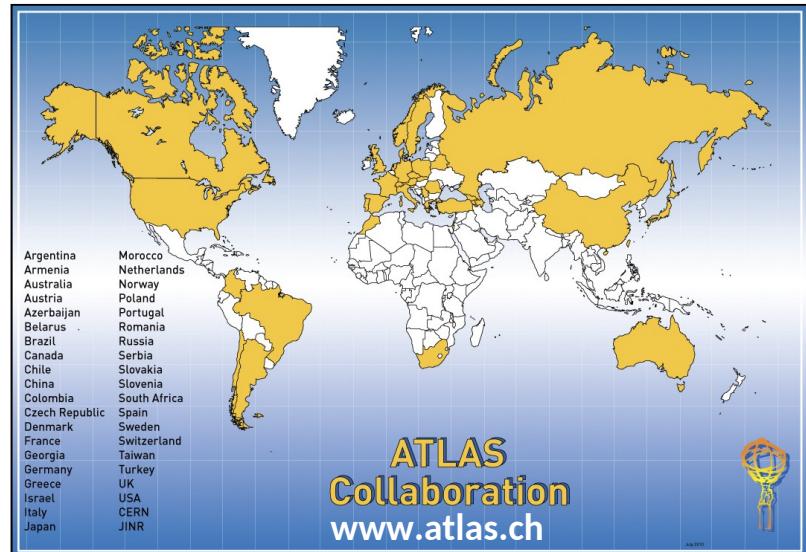


Australian Tier-2

* Until 04 2013

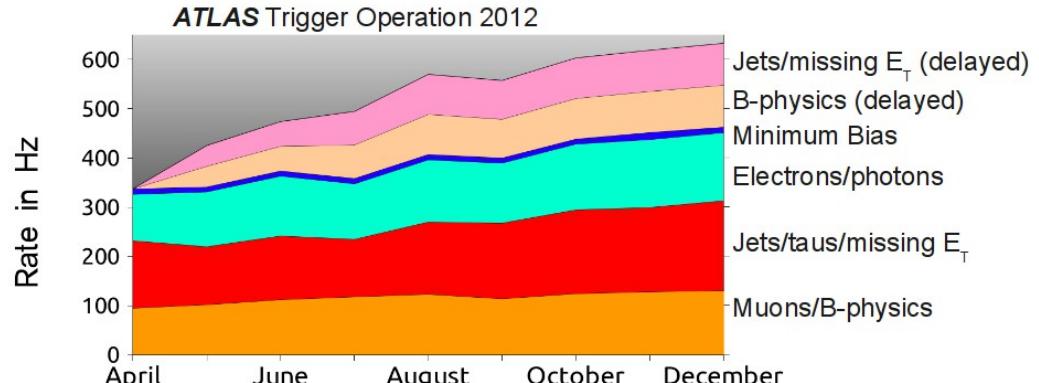
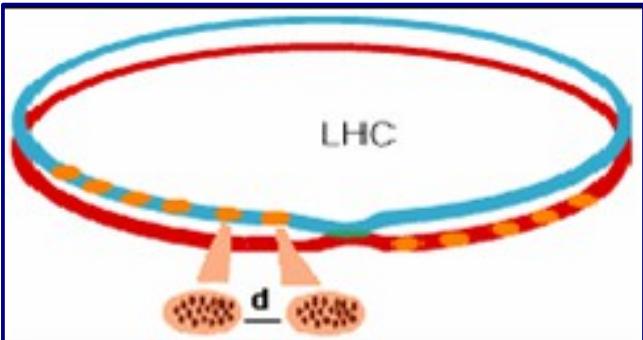
ATLAS experiment

- Explores the fundamental nature of matter and the forces that shape our universe by studying proton-proton collisions at very high energy at the Large Hadron Collider (LHC) in Geneva, Switzerland.(+heavy ions)
- Largest collaborative effort ever attempted in the physical sciences:
 - 38 countries , 177 institutions (universities & labs)
 - 3000 active scientists (1800 PhD's, 1200 students)
- Very broad, rich and high impact scientific program:
 - Nobel Prize-winning Higgs boson discovery in 2012 (Phys. Lett. B 716 (2012) 1-29, with more than 4,300 citations)

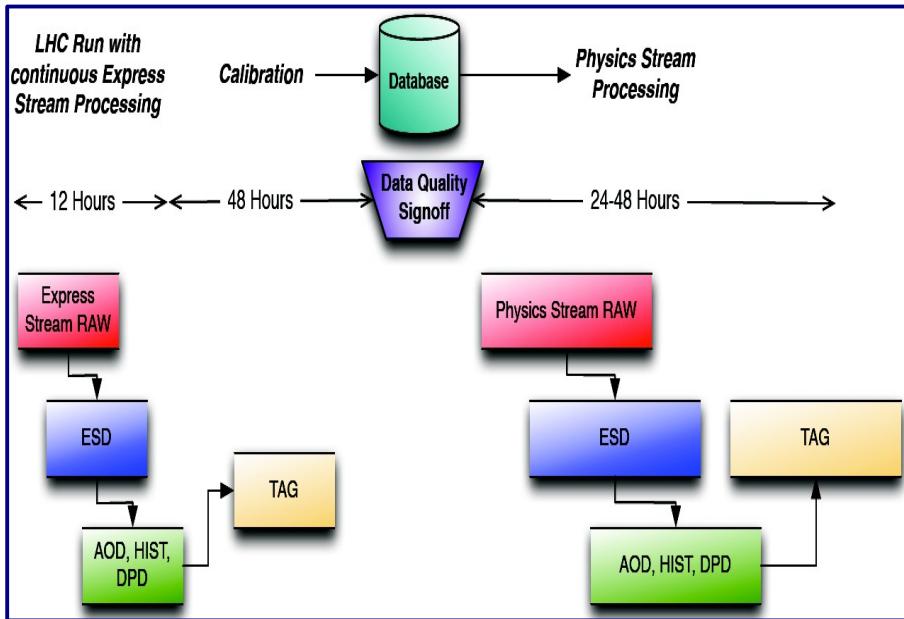
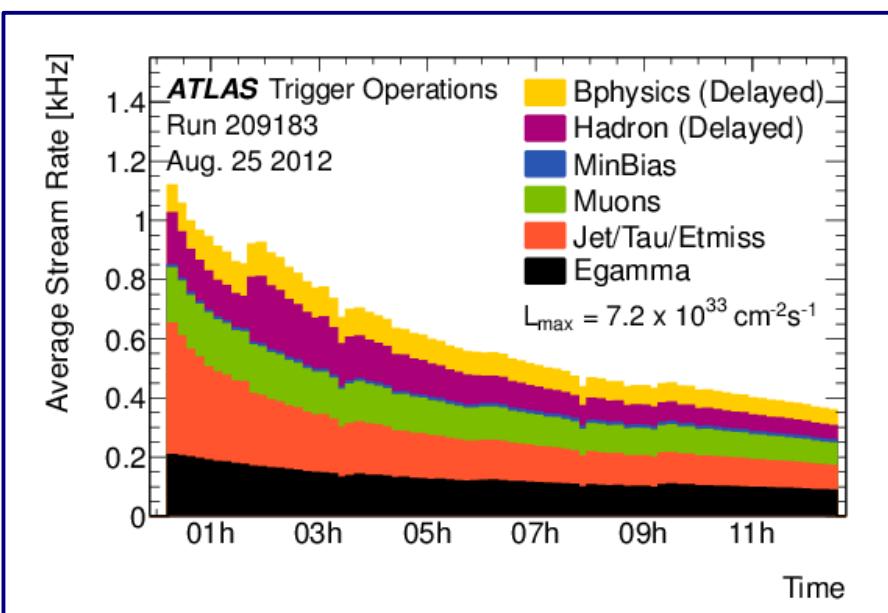


- **ATLAS-Canada:** Alberta, British Columbia, Carleton, McGill, Montréal, Simon Fraser, Toronto, TRIUMF, Victoria, York
- (38 faculties, 26 postdoctoral researchers, 66 graduate students, 25 undergrads/year)
- The Canadian government has invested over \$175 million

Streams & Data Quality



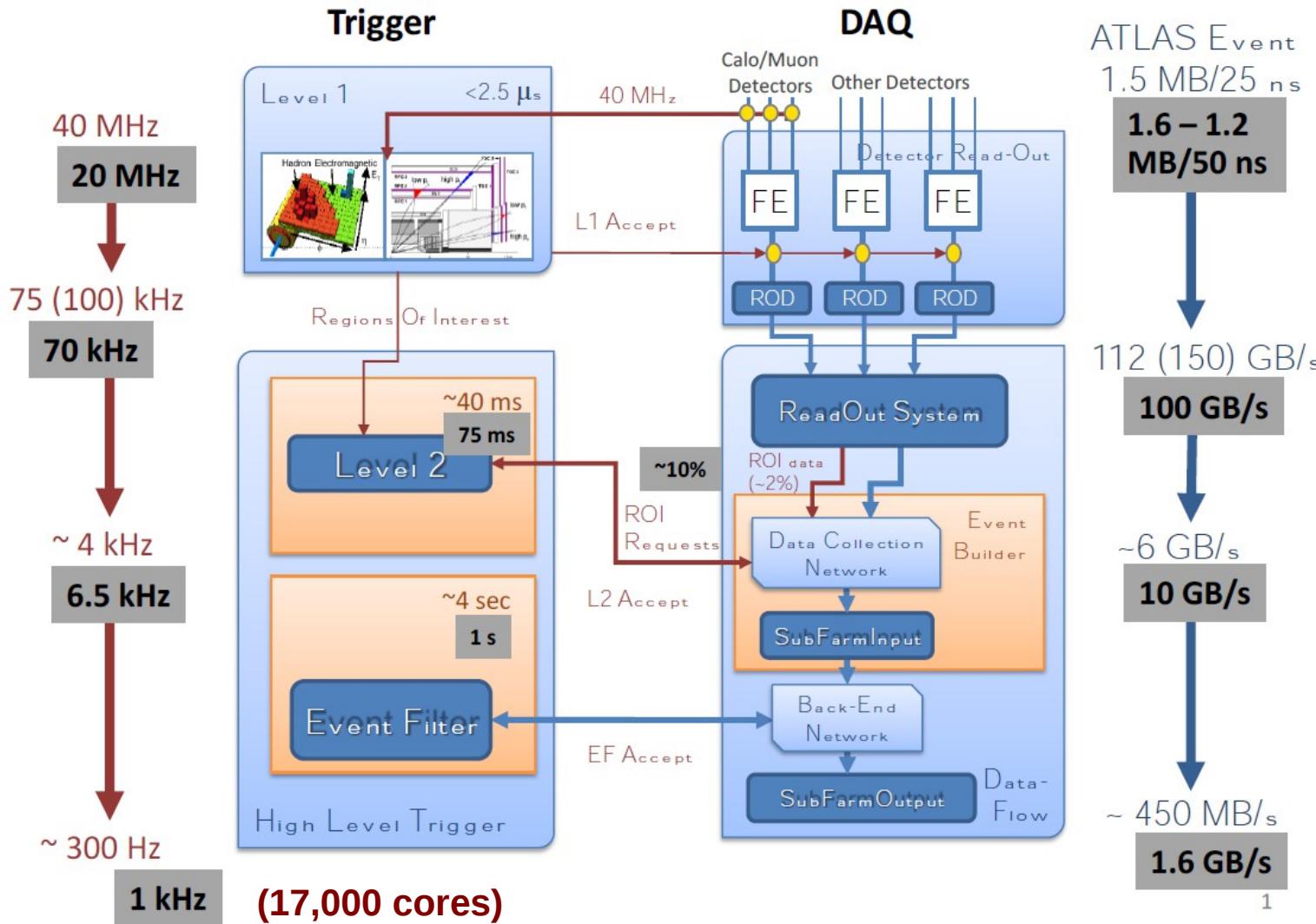
- 1380 bunches per beam
- 100 billion protons per bunch



Trigger & Data Acquisition

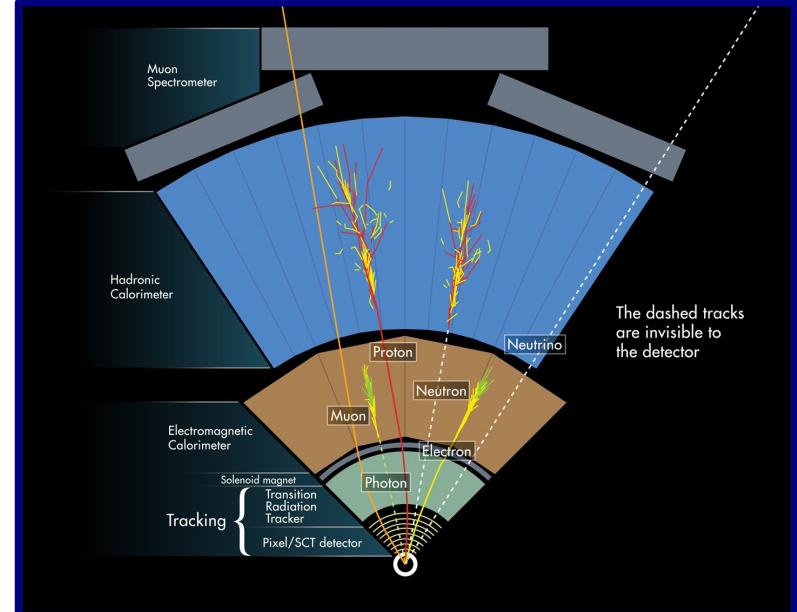
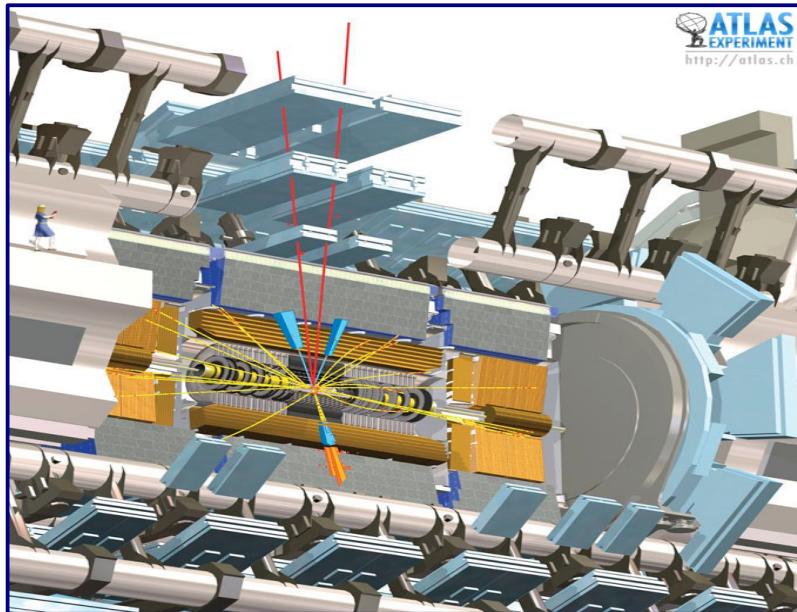
2012

TDAQ in 2012



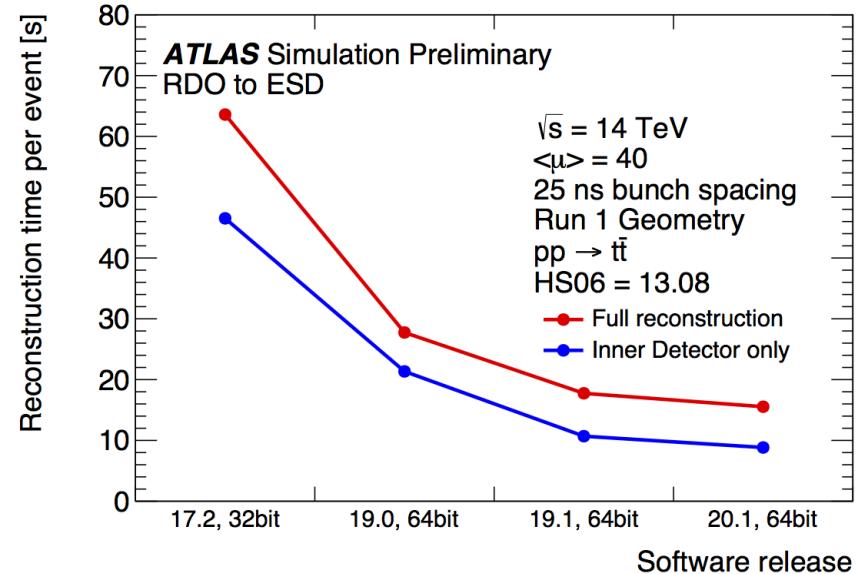
Simulation framework

- Several billion simulated proton-proton collisions are needed
- Essential for analysis: model comparison with data
- Monte Carlo generators: various physics processes resulting from proton-proton collisions
- ATLAS detector response: Geant4 program
- Full vs Fast simulation: tracking, calorimetry
- Simulated data goes through same data processing chain as real data to produce derived data for analysis

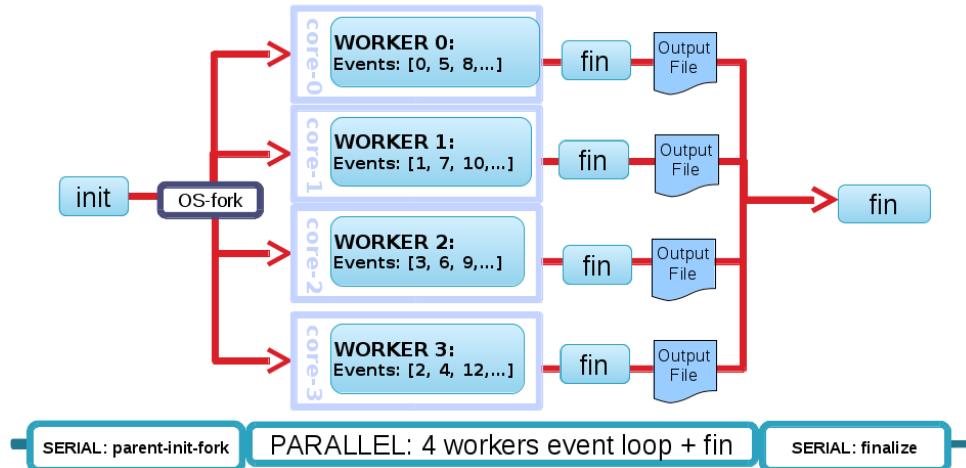


Software performance

- Large development effort during LHC long shutdown phase 1 (2013-2014)
- Improved code for event reconstruction tasks
- Move to Multi-threaded processing (memory reduction)



Schematic View of ATLAS AthenaMP



ATLAS Preliminary. Memory Profile of MC Reconstruction

