Big PanDA Experience on LCF and HPC for the ATLAS Experiment at the LHC and data intensive science

F.Barreiro, K.De, A.Klimentov, R.Mashinistov, T.Maeno, D.Oleynik, S.Panitkin, P.Svirin and J.Wells
On behalf of the ATLAS Collaboration

Grid 2018, Dubna
September 13, 2018

Outline

- Introduction
- Supercomputers (High Performance Computing - HPC and Grid (High Throughput Computing – HTC) integration challenges
- Workload and Workflow Management in ATLAS
- HPC and LCF role in ATLAS.
  - OLCF Titan
- PanDA beyond LHC
- Future challenges
- Summary

Introduction

- LHC physics with ATLAS at CERN
  - Largest experimental detector ever built
  - Higgs discovery, precision measurements of the Standard Model (including Higgs), searching for dark matter, supersymmetry, exotic particles… >700 publications
  - LHC physics program well planned out for the next 20 years
- ATLAS Computing Challenge
  - Massive scale of computing resources required
    - Currently using about 300k batch slots and 370 PB storage continuously
  - Huge collaboration – with thousands of data analyzers
  - Complex workloads and workflows
  - Computing needs grow every year with more LHC and Monte-Carlo data

The Nobel Prize in Physics 2013

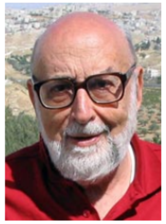


Photo: Pnicolet via Wikimedia Commons
François Englert

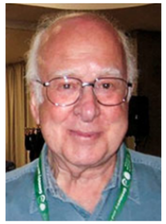


Photo: G-M Greuel via Wikimedia Commons
Peter W. Higgs

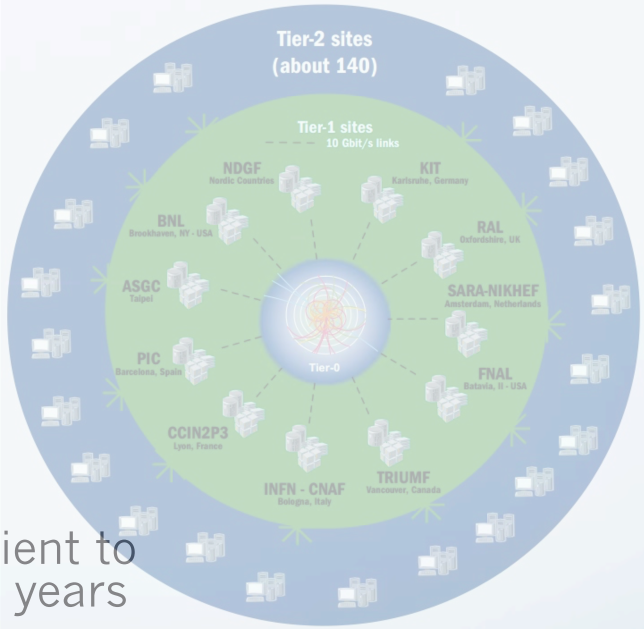
The Nobel Prize in Physics 2013 was awarded jointly to François Englert and Peter W. Higgs *"for the theoretical discovery of a mechanism that contributes to our understanding of the origin of mass of subatomic particles, and which recently was confirmed through the discovery of the predicted fundamental particle, by the ATLAS and CMS experiments at CERN's Large Hadron Collider"*

ATLAS Computing Model

- Designed to use distributed centers remotely from the start
  - No single facility of the required scale was available
  - Collaborators were scattered worldwide – early experience of using local resources proved too labor intensive
- Distributed computing model is based on four main pillars
  - WorkFlow Management System – ProdSys (Production System)
  - Workload (execution) Management System – PanDA
  - Data Management System – Rucio
  - Monitoring
- All systems are tightly integrated for efficient performance
- ATLAS data centers started out organized in Tiers, but advances in networking have eliminated strict hierarchy

ATLAS Computing Resources

- Computing infrastructure is built around WLCG
  - WLCG == The Worldwide LHC Computing Grid
  - Distributed High Throughput Clusters (HTC)
  - Currently ATLAS has access to about 150 such HTC clusters
  - While resource sizes vary greatly – they are accessed using standard grid protocols (with some regional variations)
  - Cloud computing resources are integrated into ATLAS – somewhat similar to grid sites
  - Some HPCs on grid – e.g. NorduGrid, Russia, China, …



- However, even this large pool of resources is insufficient to meet ATLAS physics goals – we realized already five years ago, after Run 1 (2013)

- Current pace of research and discovery is limited by ability of the ATLAS distributed computing facilities to generate Monte-Carlo events - "Grid luminosity limit" and to process *ALL* LHC data in quasi real-time mode

HTC + HPC

- After Run 1, ATLAS decided to add large HPCs to the mix
- Effort proved timely for highly successful Run 2 (ongoing ~x2 data), and looking ahead to Runs 3, 4 at the LHC
  - More data than expected in 2016-17 ➔ CPU shortfall
  - Expect x3-x10 shortfall by 2025 with flat WLCG budget
- ATLAS is now seriously using traditional HPC systems
  - Up to ~50% of ATLAS CPU cycles are often needed for simulations, which are well suited for HPC architectures
  - We integrated HPCs into our production, analysis and data management systems (also to monitoring and accounting) in 2016 – not used standalone
- HTC/Grid + Clouds + HPC == Truly Heterogeneous and distributed computing integrated seamlessly

The Opportunity for Supercomputer-Grid (HTC) Integration

How do we efficiently integrate supercomputing resources and distributed High Throughput Computing (HTC, or Grid) resources?

- From the perspective of large supercomputer centers, how best to integrate large capability workloads, e.g., the traditional workloads of leadership computing facilities, with the large capacity workloads emerging from, e.g., experimental and observational data?
- Workflow Management Systems (WFMS) are needed to effectively integrate experimental and observation data into our data centers.

The Opportunity for Supercomputer-Grid (HTC) Integration. Cont'd

The ATLAS experiment provides an attractive science driver, and the PanDA Workflow Management System has attractive features for capacity-capability integration

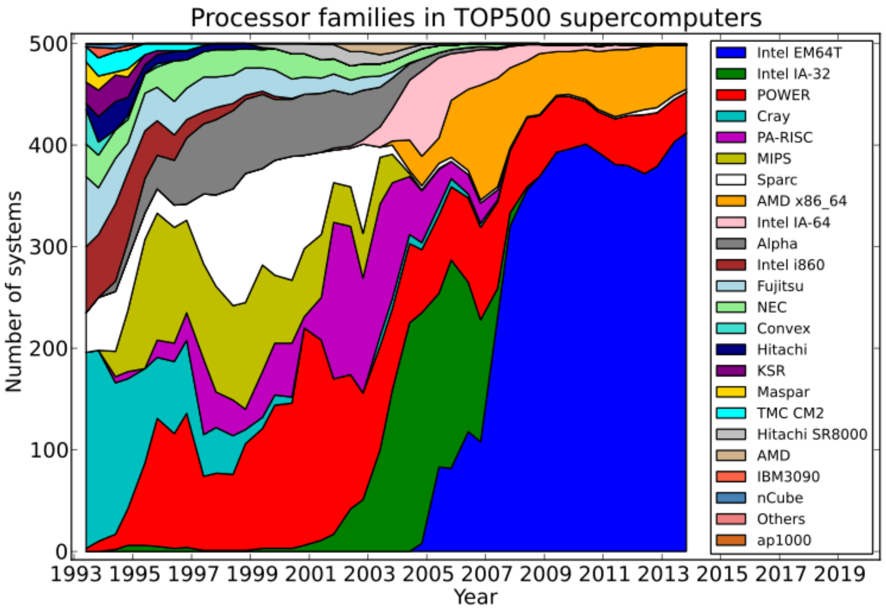
- *The Worldwide LHC Computing Grid and a leadership computing facility (LCF) are of comparable compute capacity.*
  - *WLCG ATLAS: Several 100,000's x86 compute cores*
  - *Titan: 300,000 x86 compute cores and 18,000 GPUs*

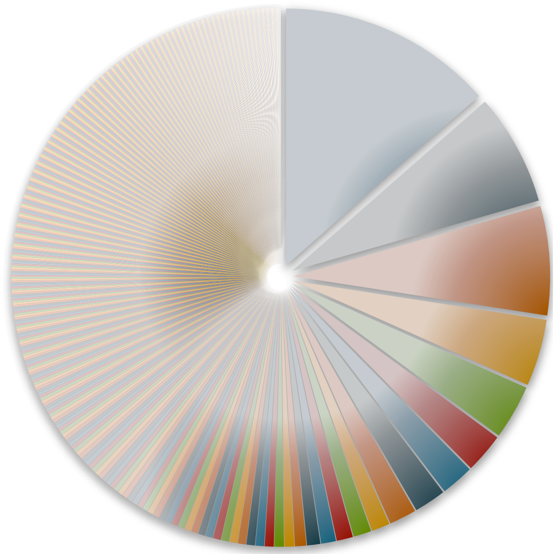
- *There is a well-defined opportunity to increase LCF utilization through backfill.*
  - *Batch scheduling prioritizing leadership-scale jobs results in ~90% utilization of available resources.*
  - *Up to 10% of Titan's cycles (~400M core hours) are available if a very large volume of capacity jobs can be run in backfill mode.*

Why is HPC Different?

- Many HPC sites do not provide grid or cloud interfaces
  - Condor-CE, ARC-CE, GCP (Google Cloud Platform), Amazon EC2, and other grid/cloud APIs make it easier to integrate distributed sites with ATLAS
- Long term storage is not available at most HPCs
  - But it is not required for MC simulations (data is transient)
- Other issues: SW installation, networking, I/O, data movement for HPC
- HPCs also require special optimizations because of their size
- PanDA now has a new component – Harvester
  - Which enables secure access through edge machine
  - Multiple edge machines for scalability
- ATLAS Event Service – another new innovation perfect for HPC

Top 500





1. Summit
2. TaihuLight
3. Sierra
4. Tianhe-2A
5. ABCI
6. Piz Daint
7. Titan
8. Sequoia
9. Trinity
10. Cori

Seymour Cray :
“supercomputer, it is hard to define, but you know it when you see it”

- *Large HPCs use a variety of architecture*
- *Half of computational power is concentrated in a small number of machines;*
- *Small HPCs use x86 architectures. Typically, these are ordinary server racks, with Infiniband interconnects. 94% of the bottom 400 of the Top 500 (including the last 130) are all x86*

In 1999...

Vs.

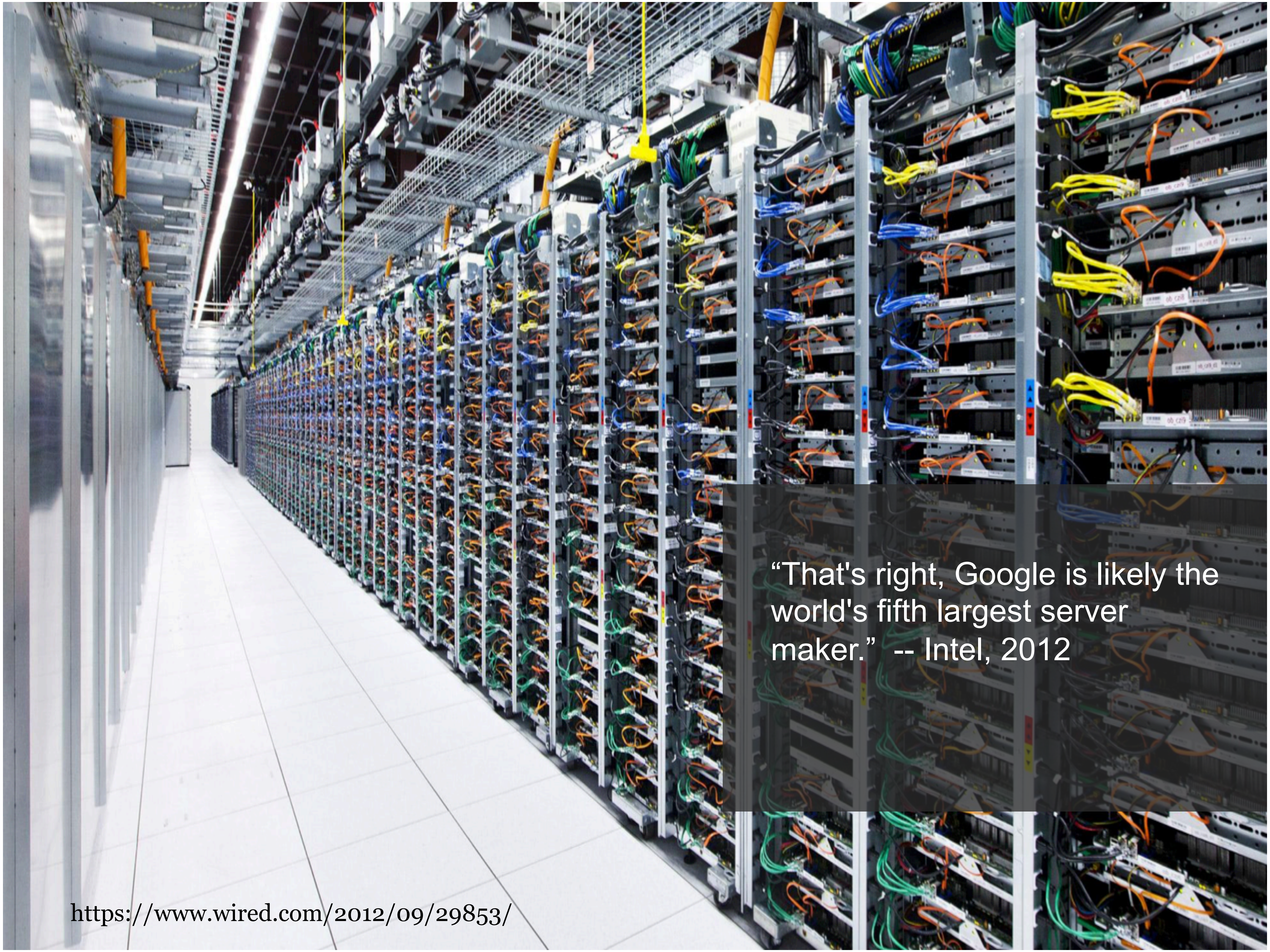
Not Difficult to Spot the Supercomputer

In 2016...

Vs.

...and the convergence isn't only visual.
GCP uses same CPU, GPU technology.

K.Kissel, Google, SMC2017 talk

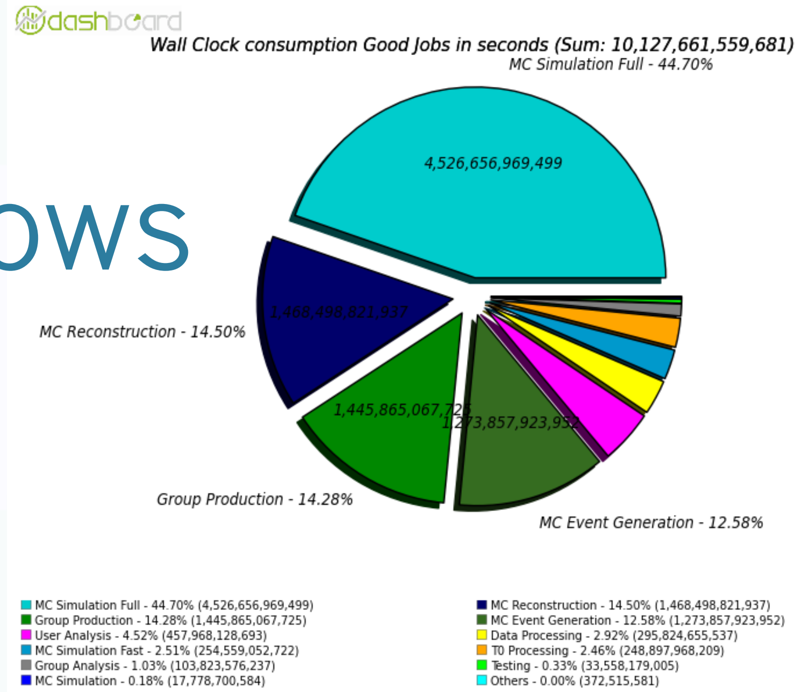

“That's right, Google is likely the world's fifth largest server maker.” -- Intel, 2012
https://www.wired.com/2012/09/29853/

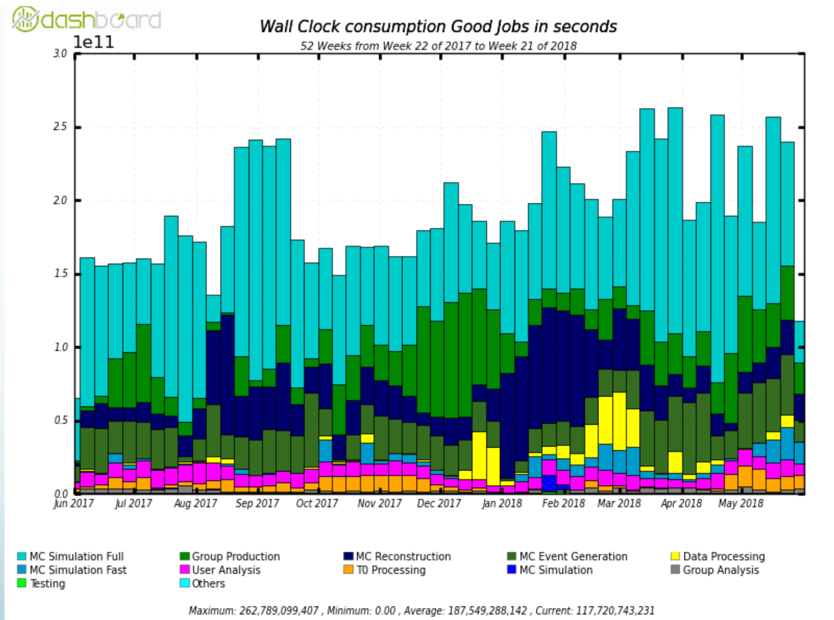
Looking Forward Looking Back

- Body of existing SMP, MPI code must be supported
- ... but new models will be required for the future
- Clouds, HPC, Exascale share many problems
- Reason to hope for common solutions

ATLAS Workflows

- ATLAS used 2.8 billion wallclock hours in the past 12 months (Jun 17-May 18)
- Mixture of workflows – some are CPU intensive, some are I/O intensive, some mixed
- About 45% is CPU intensive
- Mixture of workload changes constantly – need heterogeneous capabilities with bursts occasionally

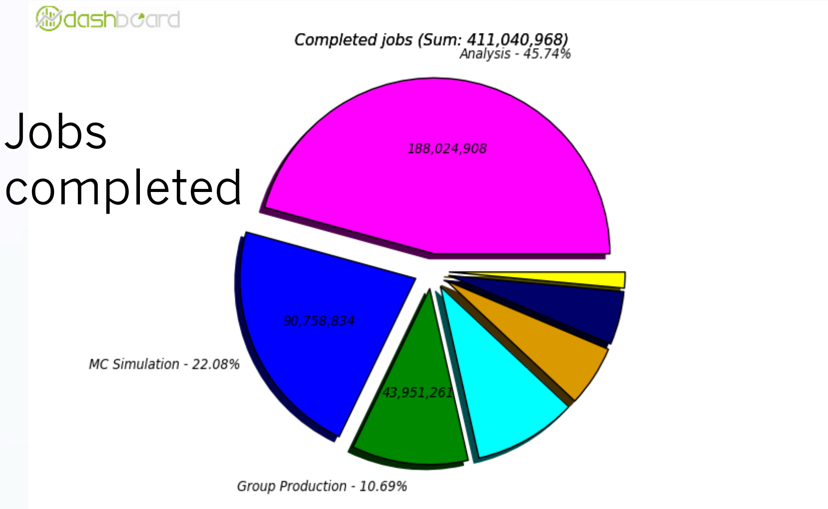




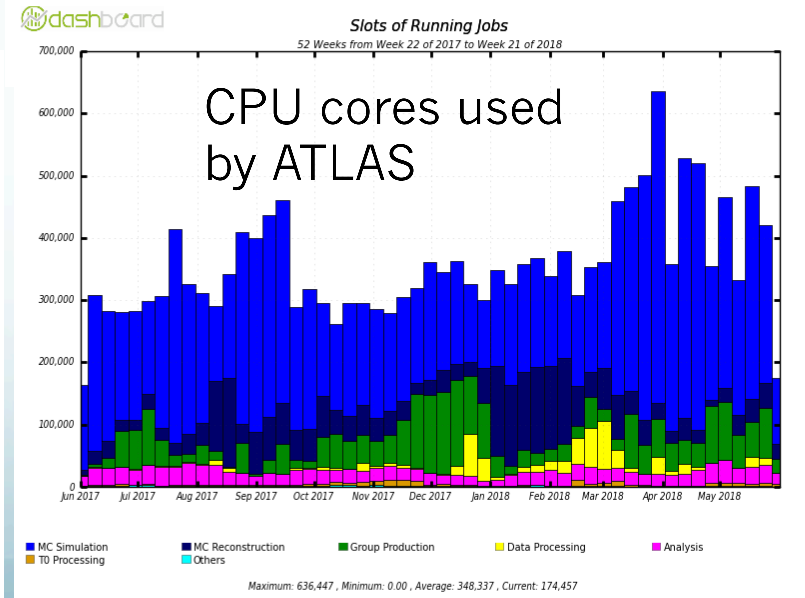
Workflow Characteristics

- User analysis and other high-IO work dominates the number of jobs submitted
- But CPU intensive workflow dominates wallclock
- Currently, ATLAS needs 300k cores, with fluctuations up to 700k, including HPC

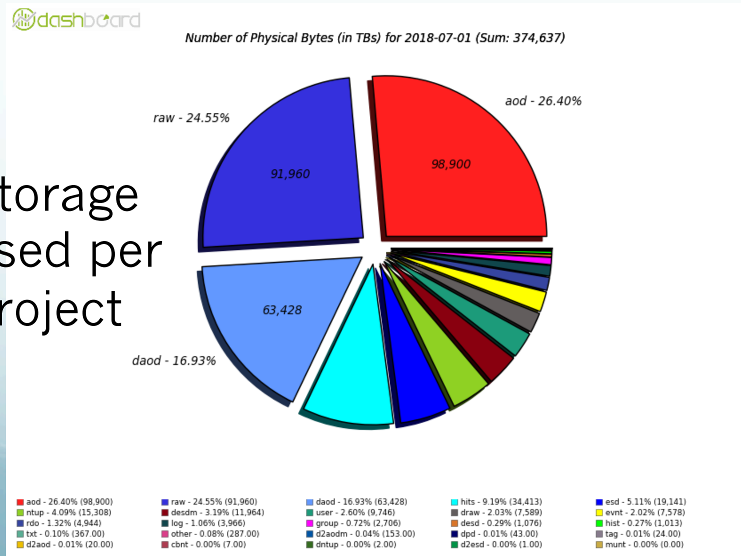
Jobs completed



CPU cores used by ATLAS

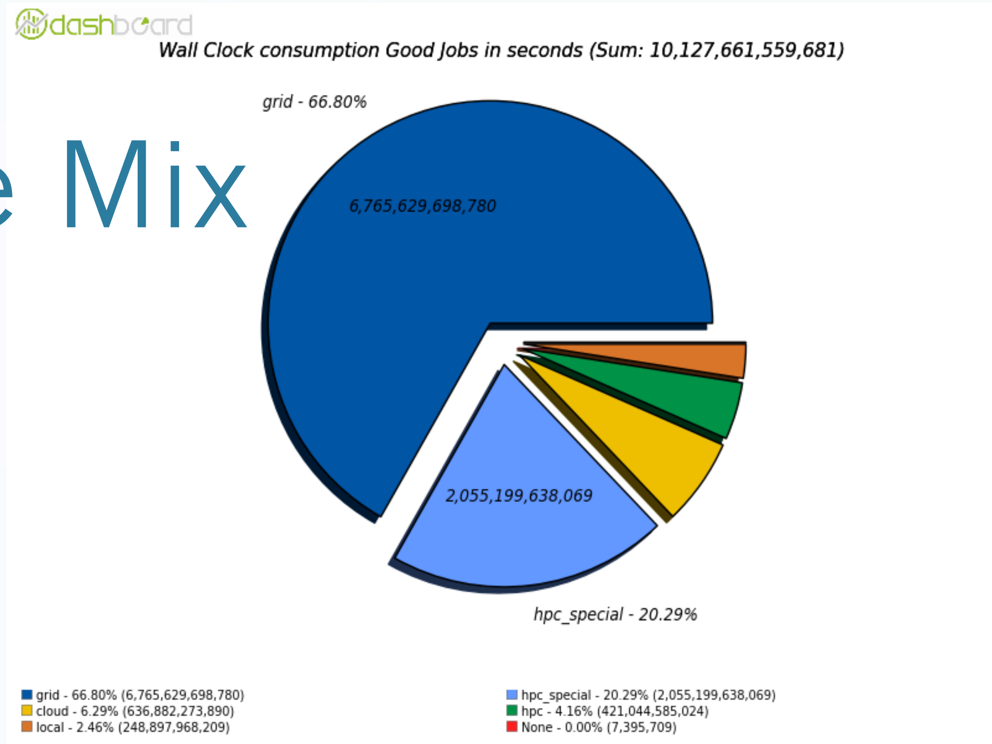


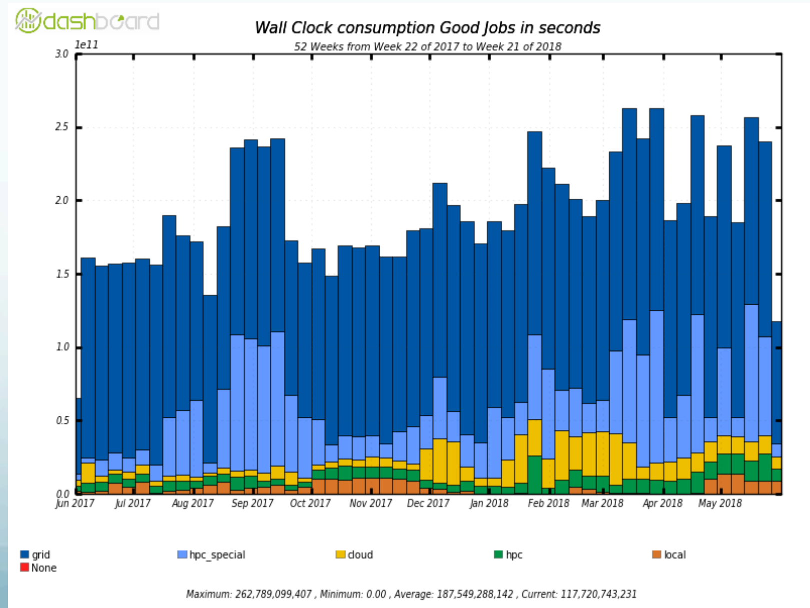
Storage used per project



CPU Resource Mix

- HPCs delivered 20% of ATLAS wallclock usage in the past 12 months
- This is a huge increase over past year – using many HPCs worldwide
- We focused on enabling all HPC centers available to ATLAS into the distributed computing system





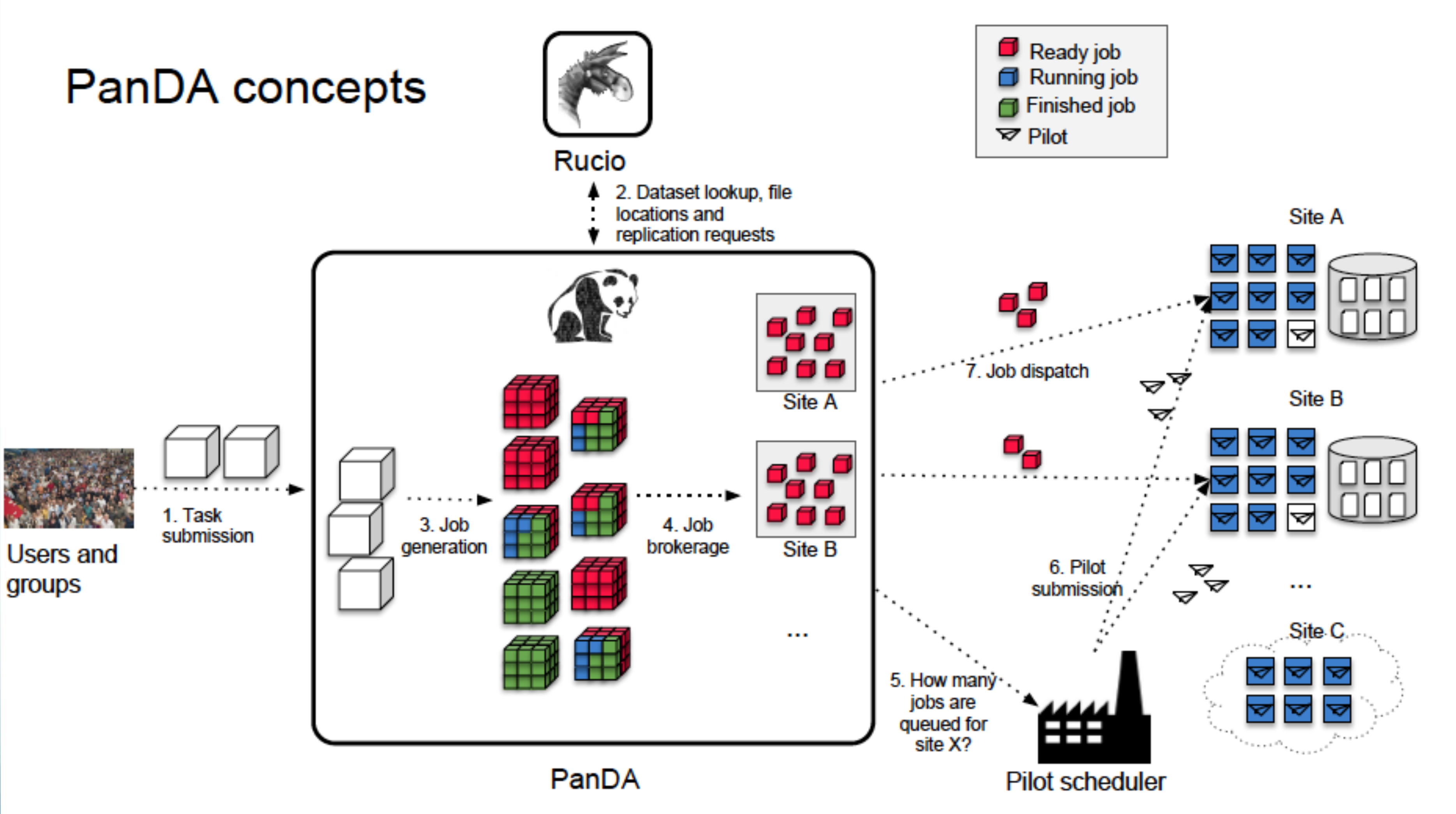
Wallclock usage by resource type

PanDA Workload Management System

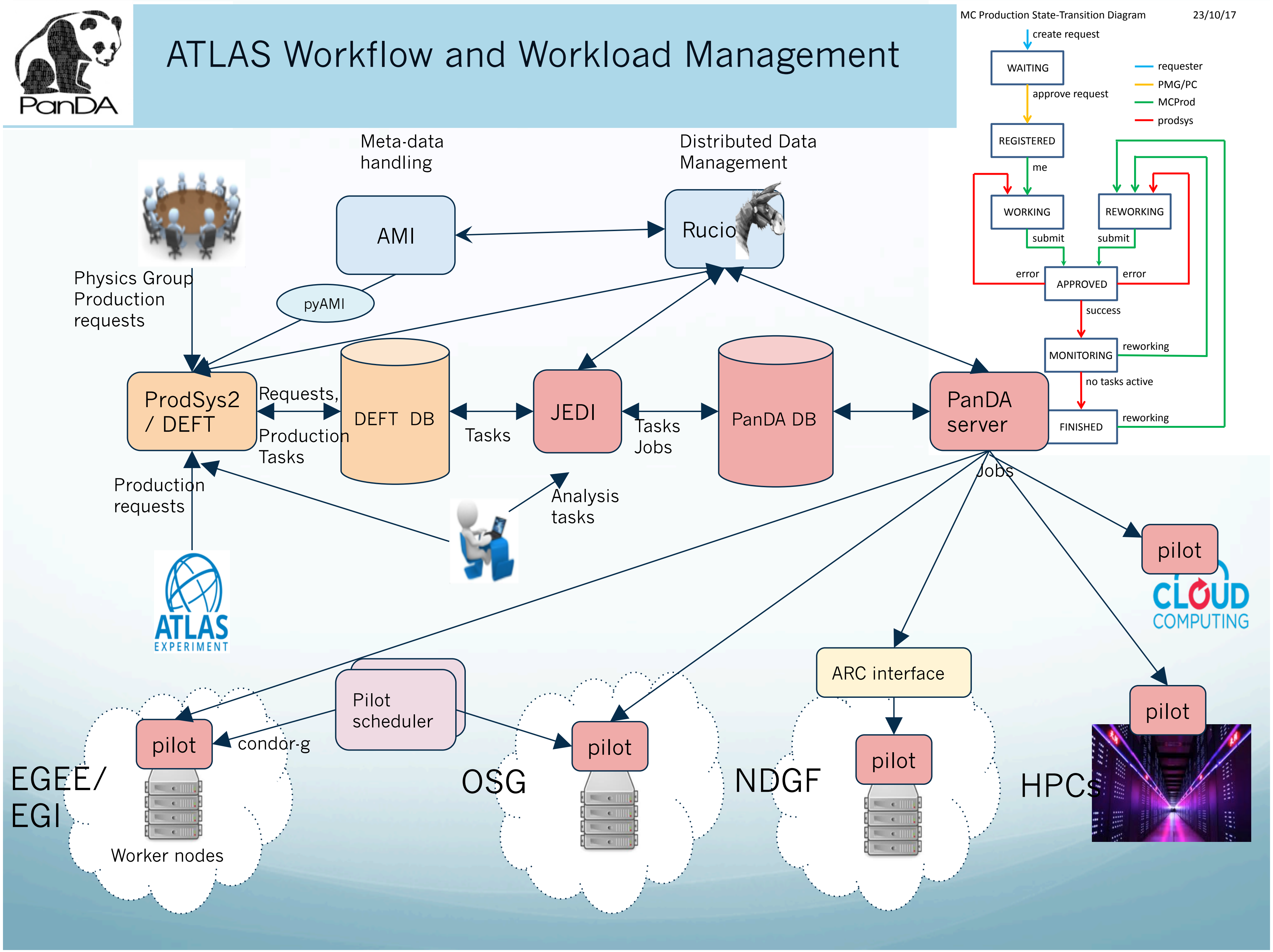
- The PanDA workload management system was developed for the ATLAS experiment at the Large Hadron Collider. A new approach to distributed computing
  - A huge hierarchy of computing centers and opportunistic resources working together
  - Main challenge – how to provide efficient automated performance
  - Auxiliary challenge – make resources easily accessible to all users

- Core ideas :
  - Make hundreds of distributed sites appear as local
    - Provide a central queue for users – similar to local batch systems
  - Reduce site related errors and reduce latency
    - Build a pilot job system – late transfer of user payloads
    - Crucial for distributed infrastructure maintained by local experts
  - Hide middleware while supporting diversity and evolution
    - PanDA interacts with middleware – users see high level workflow
  - Hide variations in infrastructure
    - PanDA presents uniform 'job' slots to user (with minimal sub-types)
    - Easy to integrate grid sites, clouds, HPC sites ...
  - Data processing, MC Production and Physics Analysis users see same PanDA system
    - Same set of distributed resources available to all users
    - Highly flexible – instantaneous control of global priorities by experiment

PanDA Workload Management System



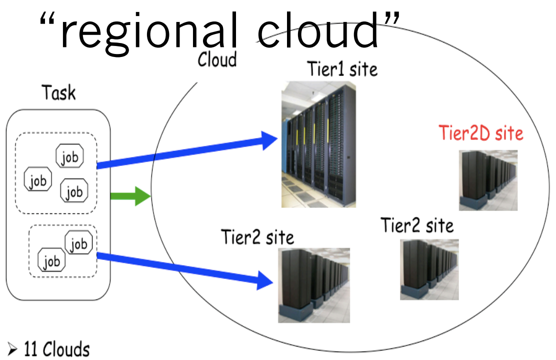
CHEP 2018 talk about new PanDA developments : "Harvester : an edge service harvesting heterogeneous resources for ATLAS", https://indico.cern.ch/event/587955/contributions/2937391/



ATLAS Workflow and Workload Management
PanDA
MC Production State-Transition Diagram
23/10/17
create request
WAITING
approve request
REGISTERED
me
WORKING
REWORKING
submit
submit
error
APPROVED
error
success
MONITORING
reworking
no tasks active
FINISHED
reworking
requester
PMG/PC
MCProd
prodsys
Meta-data handling
Distributed Data Management
AMI
Rucio
pyAMI
Physics Group Production requests
ProdSys2 / DEFT
Requests, Production Tasks
DEFT DB
Tasks
JEDI
Tasks Jobs
PanDA DB
PanDA server
Jobs
Production requests
Analysis tasks
ATLAS EXPERIMENT
pilot
CLOUD COMPUTING
ARC interface
pilot
Pilot scheduler
condor-g
pilot
pilot
pilot
EGEE/ EGI
Worker nodes
OSG
NDGF
HPCs

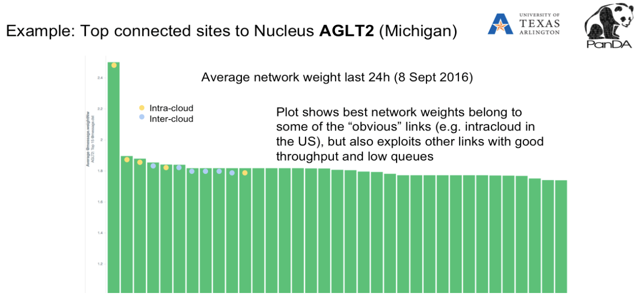
From “regional” to “world” cloud

ATLAS Computing Model 2012

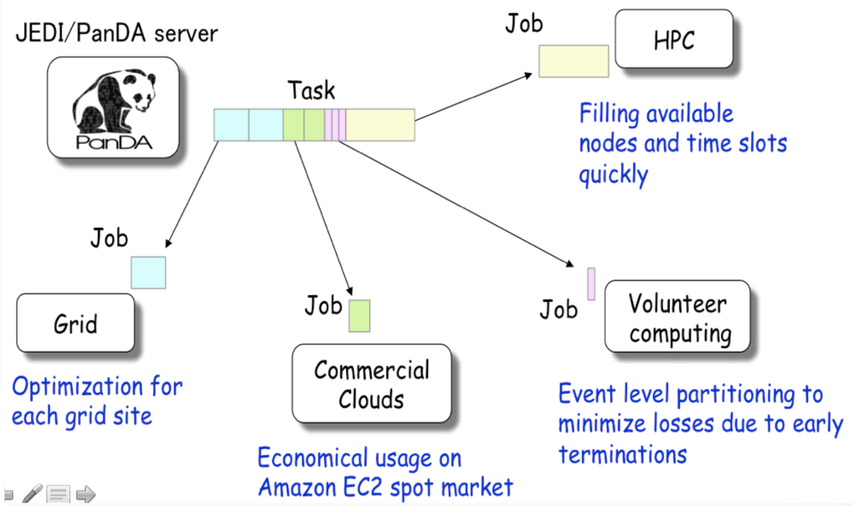


- 11 Clouds
  10 T1s + 1 T0 (CERN)
  Cloud = T1 + T2s + T2Ds (except CERN)
  T2D = multi-cloud T2 sites
- 2-16 T2s in each Cloud

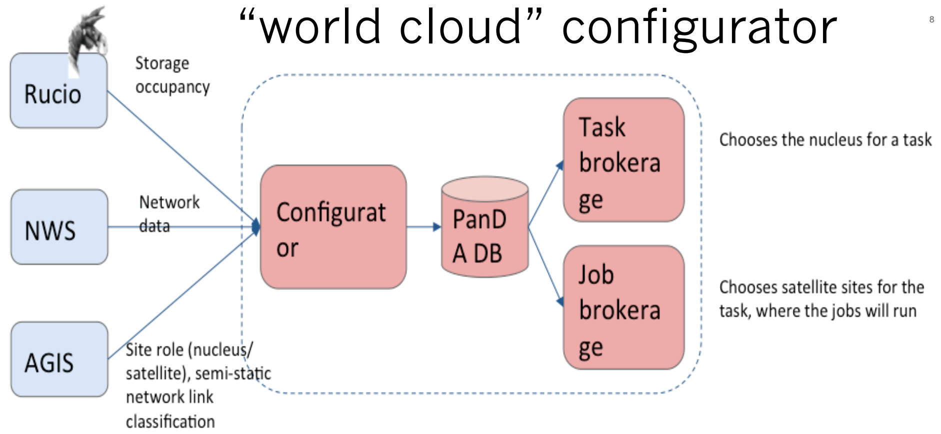
After 2012 we relaxed Tiers hierarchy and started dynamic resources configuration and dynamic workload partitioning

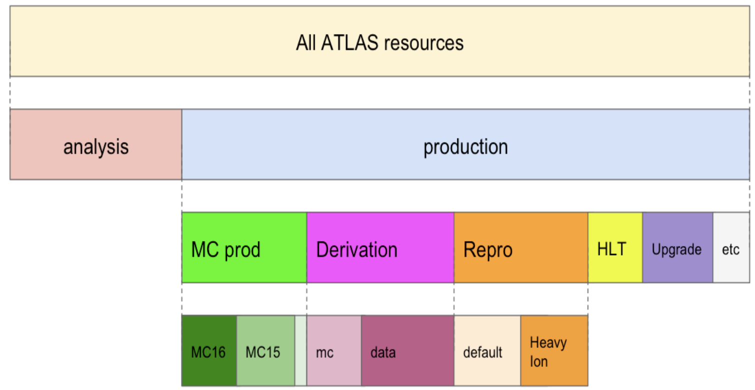


- Workload partitioning for traditional and opportunistic resources



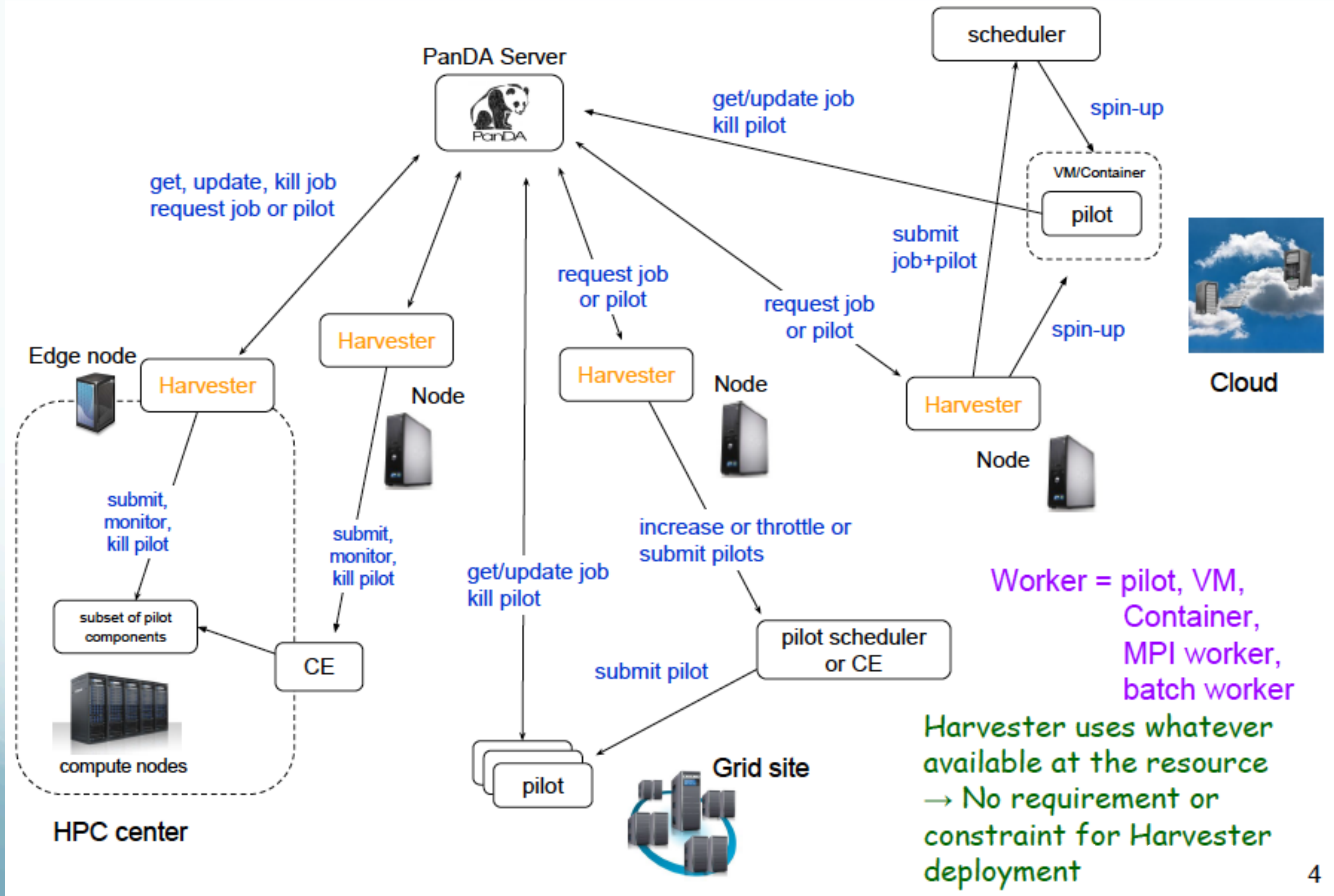
“world cloud” configurator





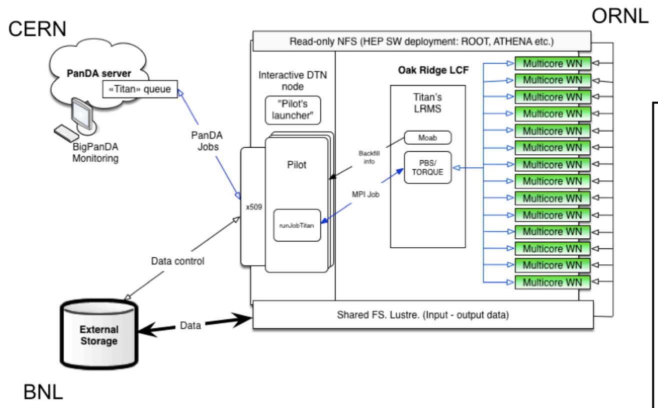
Resources allocation

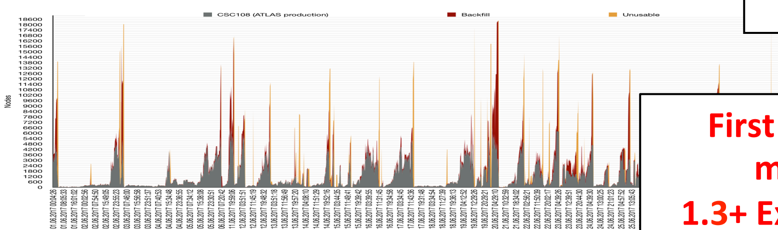
PanDA with Harvester



Workload Management. PanDA. **P**roduction **and** **D**istributed **A**nalysis System

https://twiki.cern.ch/twiki/bin/view/PanDA/PanDA





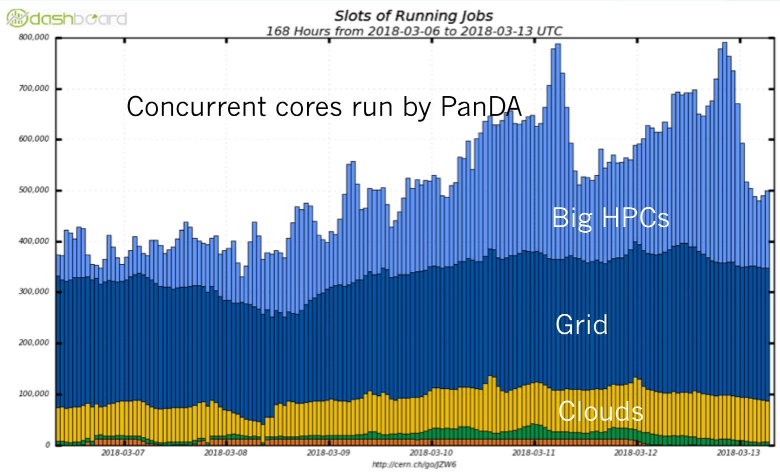
Global ATLAS operations

Up to ~800k concurrent jobs
25-30M jobs/month at >250 sites
~1400 ATLAS users

First exascale workload manager in HENP
1.3+ Exabytes processed in 2014 and in 2016-2018
Exascale scientific data processing today

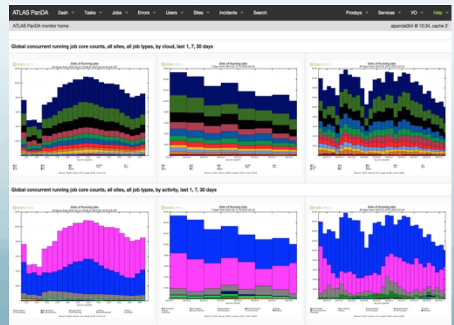
PanDA Brief Story

2005: Initiated for US ATLAS (BNL and UTA)
2006: Support for analysis
2008: Adopted ATLAS-wide
2009: First use beyond ATLAS
2011: Dynamic data caching based on usage and demand
2012: ASCR/HEP BigPanDA project
2014: Network-aware brokerage
2014 : Job Execution and Definition I/F (JEDI) adds complex task management and fine grained dynamic job management
2014: JEDI- based Event Service
2014:megaPanDA project supported by RF Ministry of Science and Education
2015: New ATLAS Production System, based on PanDA/JEDI
2015 :Manage Heterogeneous Computing Resources
2016: DOE ASCR BigPanDA@Titan project
2016:PanDA for bioinformatics
2017:COMPASS adopted PanDA , NICA (JINR)
PanDA beyond HEP : BlueBrain, IceCube, LQCD



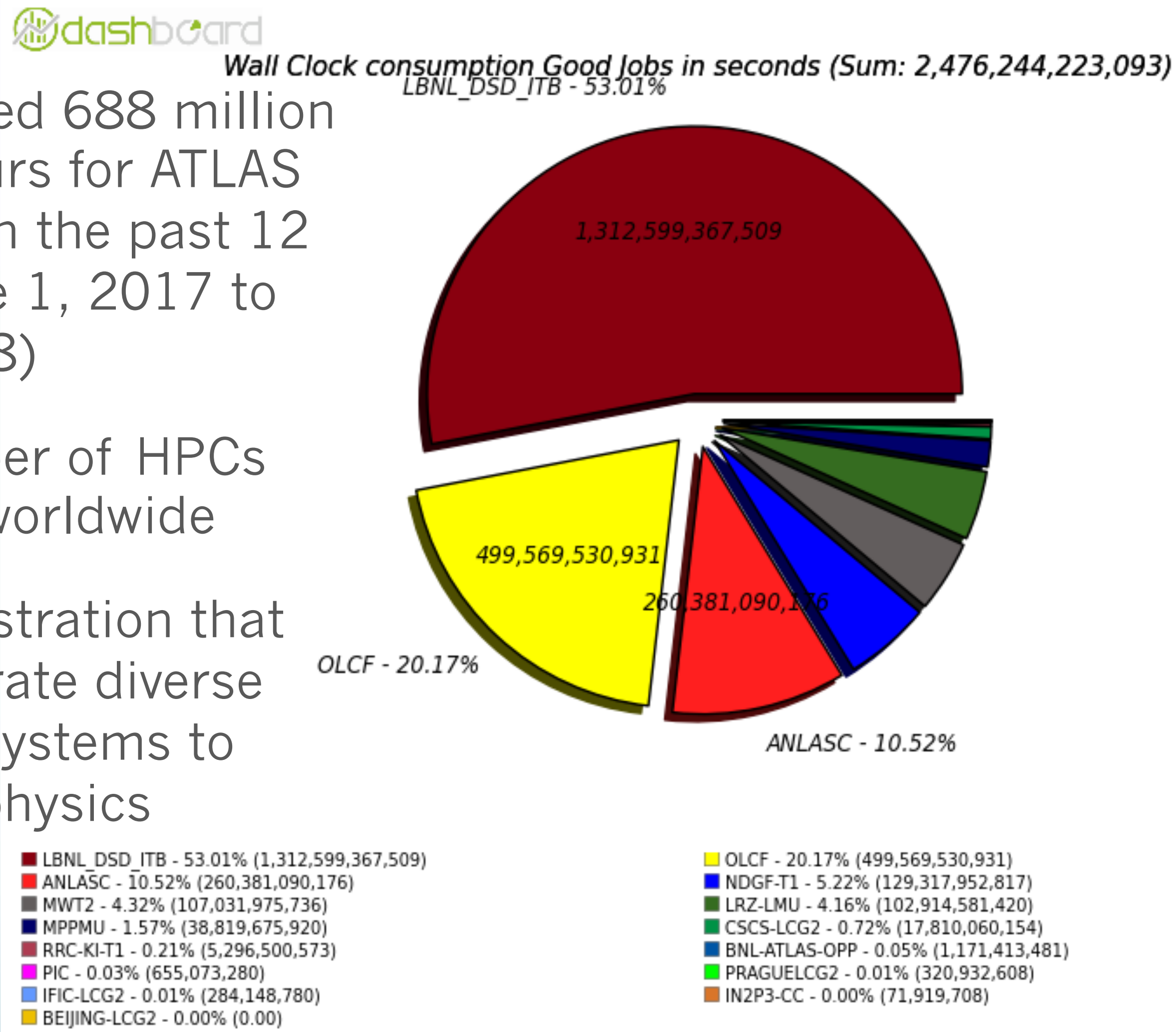
BigPanDA Monitor

http://bigpanda.cern.ch/



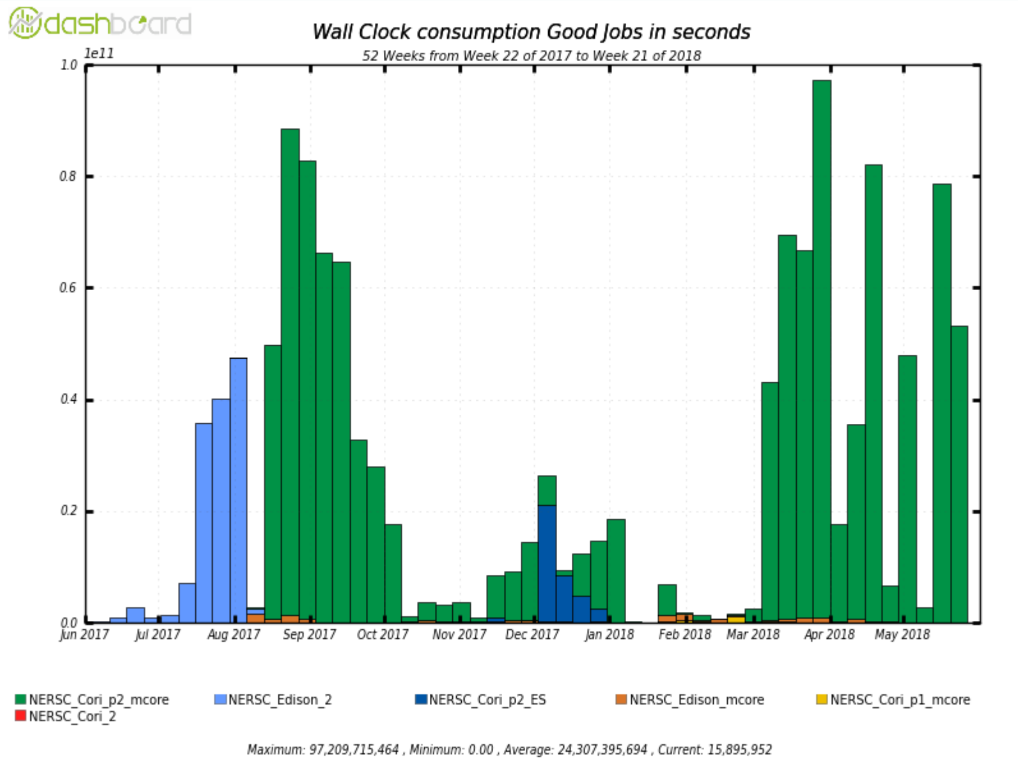
Diversity of HPC Sites

- HPCs delivered 688 million wallclock hours for ATLAS simulations in the past 12 months (June 1, 2017 to May 31, 2018)
- A large number of HPCs contributed worldwide
- Clear demonstration that we can integrate diverse mix of HPC systems to enable LHC physics



NERSC

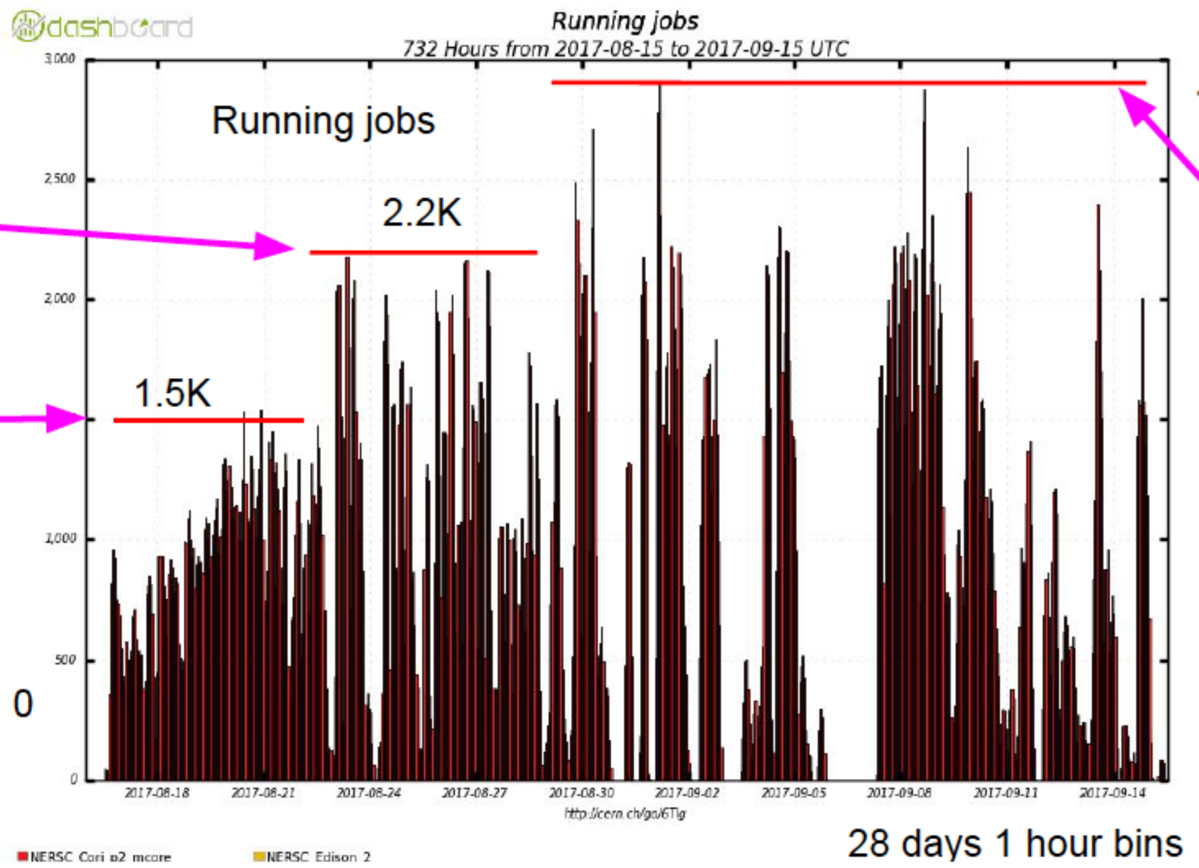
- Largest HPC contributor to ATLAS walltime usage
  - Allocations through ALCC and DOE HEP
  - Most of the hours are utilized on CORI (some on EDISON)
  - Multiple modes of job submissions – through CE, using Event Service, using Harvester
  - Active R&D on containers, data caching
  - Postdoc (NESAP award) working on multinode I/O
  - Moved from demonstration to full fledged production mode in past 12 months



Effect of using shifter containers at NERSC

Containers and output to Shared file system

Running in Shared File system only - at limit



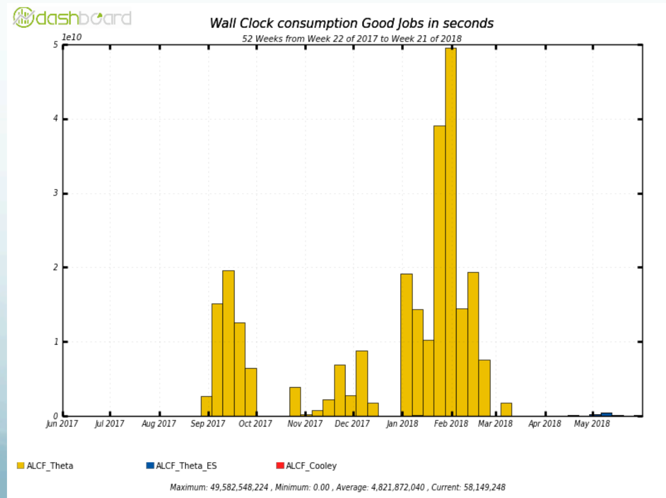
Containers and loopback filesystem for output

Very Bursty usage

Fast ramp up

Argonne LCFs

- MIRA first HPC to provide large CPU cycles to ATLAS
  - Used for event generator workflows ported to PowerPC
  - Not fully integrated with ATLAS Production System
- THETA is now fully integrated using ALCC allocations
  - Commissioning is focused on Event Service using Harvester – already third largest HPC contributor
- New early science (ESP) award on Aurora for R&D



Nordic HPCs

- NorduGrid (NDGF) enabled early adoption of HPCs
  - Through use of ARC-cache - using NDGF storage
  - Integrated with ACT – ARC front-end for PanDA
  - Using common services for grid and HPC resources
- Currently the following sites in use by ATLAS:
  - Abel, UIO, University of Oslo
  - Abisko, HPC2N Umea, Sweden
  - Tetralith, NSC, Linkoping, Sweden
  - Looking forward to new systems soon: Kebnekaise and Tetralith (and others)

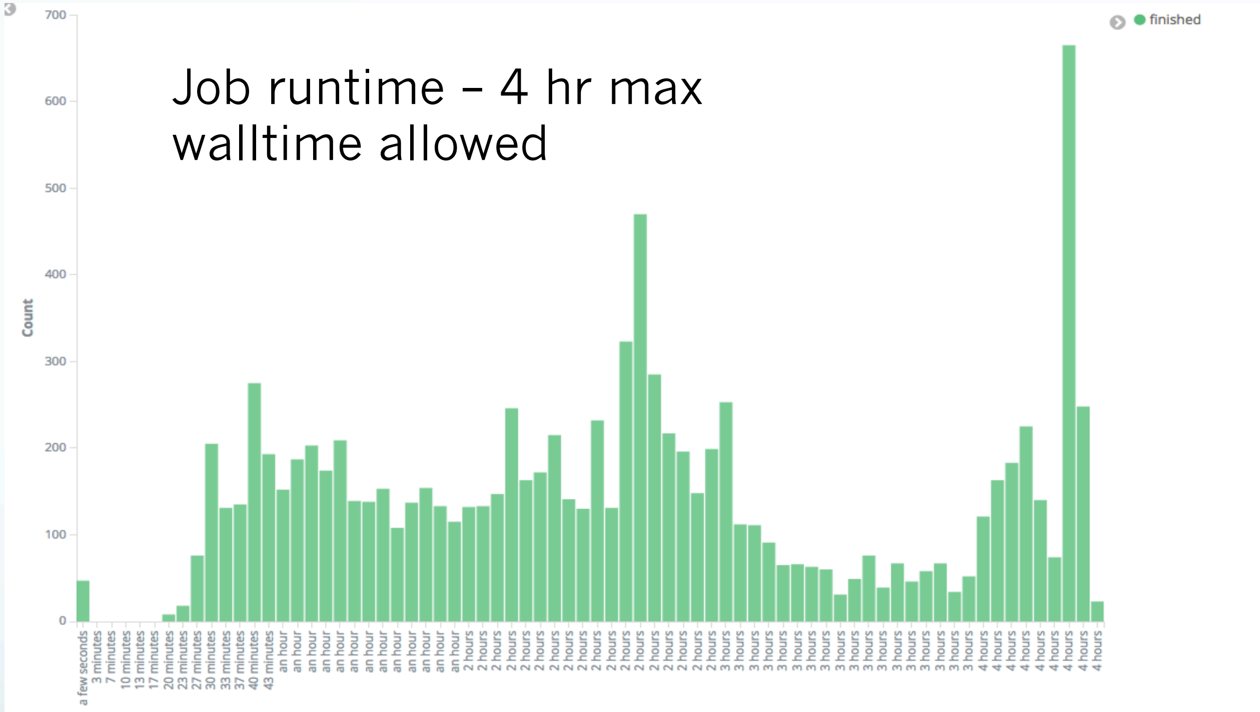
U.S. NSF HPC's

- Many HPC's available – part of XSEDE ecosystem
- All accessible through ATLAS PanDA system
  - ATLAS Connect technology by US Midwest T2 (MWT2)
  - Use MWT2 storage endpoint
  - Job submissions through ssh access
  - Variety of solutions to deliver software: Docker/Shifter at BlueWaters, Singularity on Stampede, Odessey and UI; Stratum-R copies of CVMFS elsewhere
- Allocations are available for big HPCs, but also backfill and pre-emptible queues are used
  - Reached maximum ~48k cores in May 2018

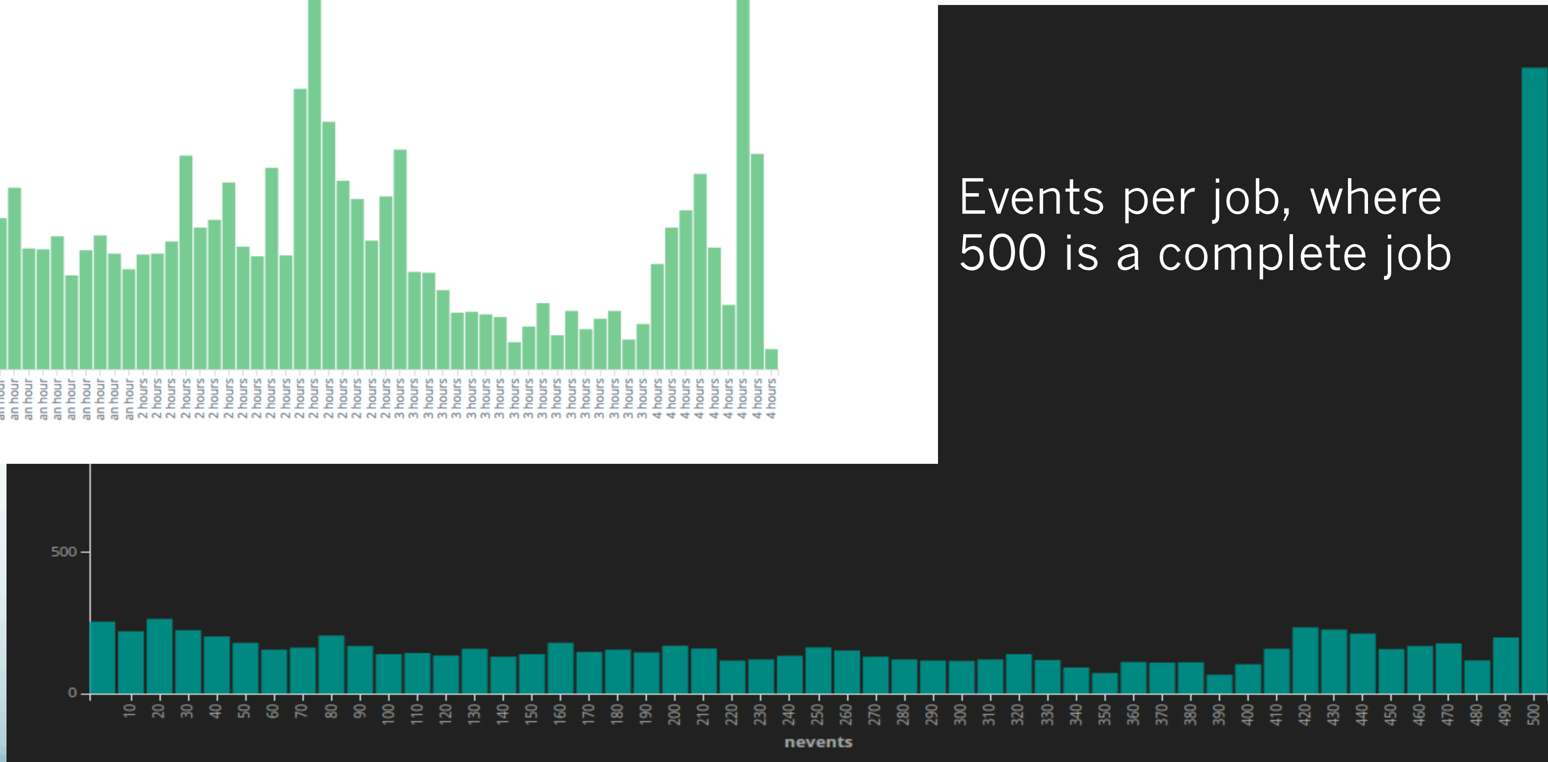
Munich HPCs

- SuperMUC at LRZ, Munich
  - Interfaced to ATLAS via ARC-CE (and ARC cache)
    - Tricks for SLES11, no CVMFS or outbound IP (using Parrot from cctools)
  - Running in single-node preempt mode: 300 node max allowed
  - ATLAS EventService: ARC-CE stores produced events after job ends or after preemption
- Hydra/Draco at MPPMU, Munich
  - ARC-CE handles stage-in/out of data
    - /cvmfs link on WNs to shared FS replica. Outbound IP for conditions data.
  - Can run all workloads: limited to sim due to shared FS I/O load
- Next generation LRZ HPC in Autumn
  - Planning container service (likely Shifter) to ease operations

Preemption on Munich



Job runtime – 4 hr max walltime allowed



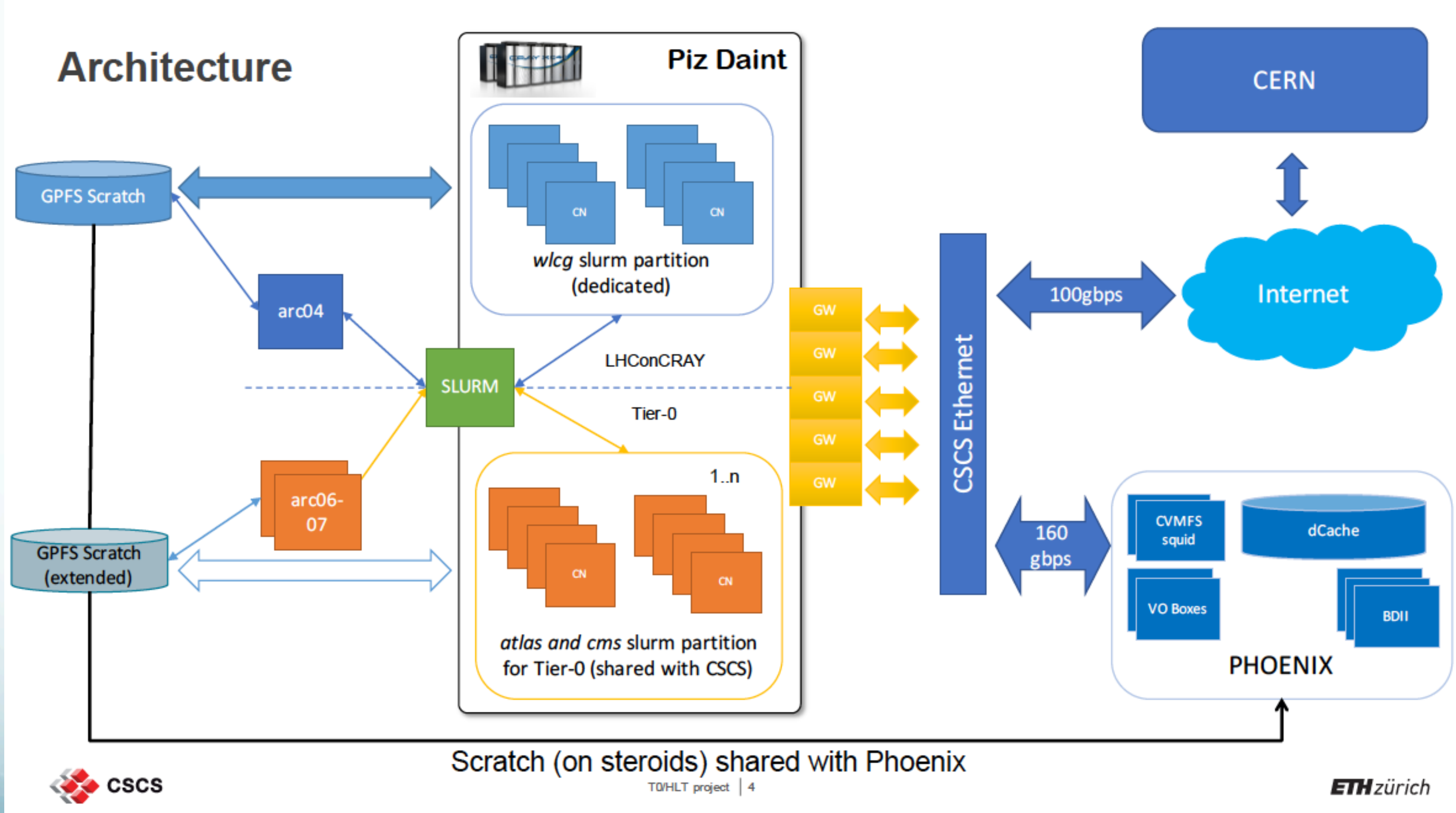
Events per job, where 500 is a complete job

ATLAS EXPERIMENT

CSCS (Piz Daint)

- ATLAS CSCS Tier 2 is now integrated with Piz Daint
  - First use of HPC as a WLCG Tier 2 Center
  - Using ARC-CE and ARC cache: HPC == HTC
  - Running successfully now, older PHOENIX will retire
- ATLAS and CSCS successfully launched R&D and PoC projects
  - For deeper integration of CSCS with ATLAS computing
  - Testing Tier-0 spillover (raw data processing) at CSCS
  - Use of CSCS for high level trigger (HLT) reprocessing
  - Future R&D for HL-LHC

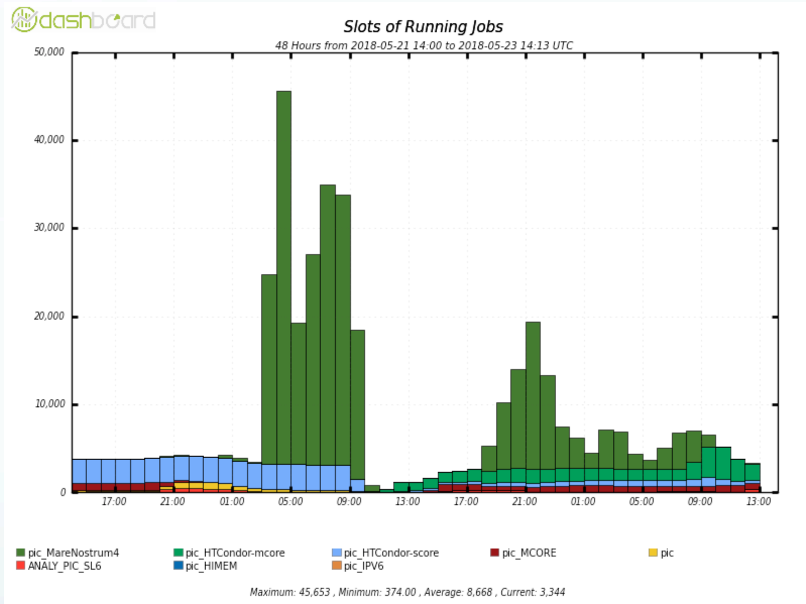
CSCS Tier-0 Spillover PoC



ATLAS EXPERIMENT

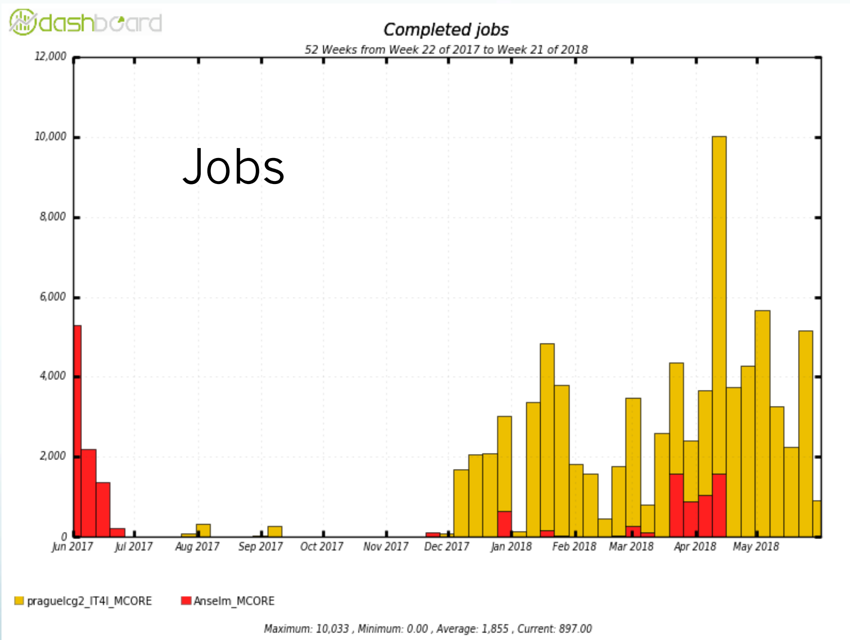
MareNostrum

- Recently MareNostrum in Barcelona was added to ATLAS
  - Served by ARC-CE at PIC (Spain Tier 1)
  - Using fat singularity container (no internet from WN)
  - 1 job per machine (48 cores)
- It is becoming easy to add new HPC resources



Czech IT4I

- Anselm – integrated first
  - Uses same solution as Titan
  - Proof-of-concept – maximum 200 nodes available
- Salomon – using ACT
  - Similar to Munich solution
- Both will be migrated to Harvester in the near future



Key contributors : J.Chudoba, D.Oleynik, A.Petrosyan,

Titan

- First large scale LCF integrated into ATLAS
  - Through the BigPanDA project funded by DOE-ASCR
  - Team leaders: Alexei Klimentov (BNL), Jack Wells (ORNL), Shantenu Jha (Rutgers University), Kaushik De (University of Texas at Arlington)
- 300 million TITAN core hours in past 12 months
  - Both backfill usage and ALCC allocation
- Successful story:
  - More details in talk by Danila Oleynik at this conference

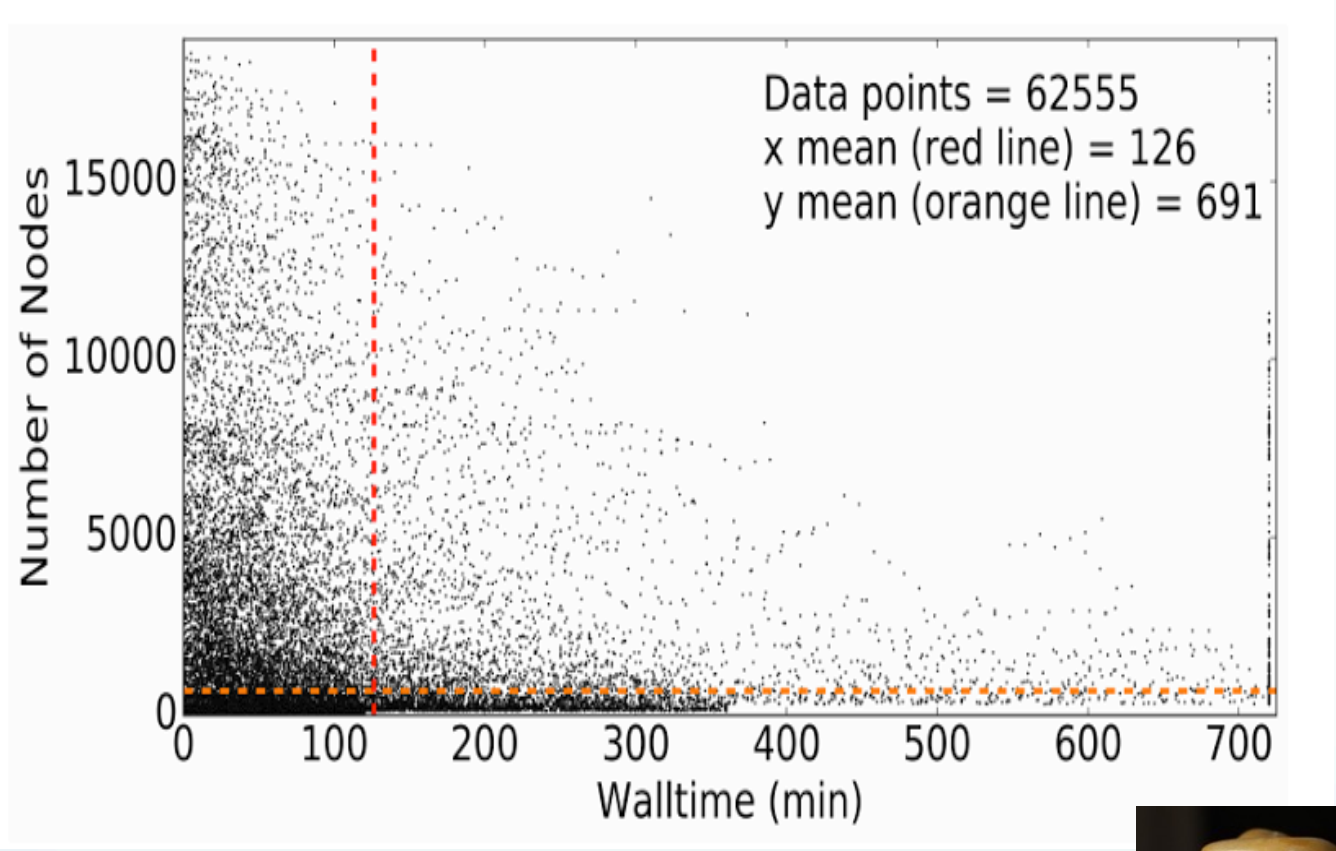
Titan (OLCF) : a Leadership Computing Facility (LCF)

- Collaborative USA DOE Office of Science user-facility program at ORNL and ANL
- Mission: Provide the computational and data resources required to solve the most challenging problems.
- 2-centers/2-architectures to address diverse and growing computational needs of the scientific community
- Highly competitive user allocation programs (INCITE, ALCC).
- Projects receive 10x to 100x more resource than at other generally available centers.
- LCF centers partner with users to enable science & engineering breakthroughs (Liaisons, Catalysts).

The Worldwide LHC Computing Grid and a leadership computing facility are of comparable compute capacity.

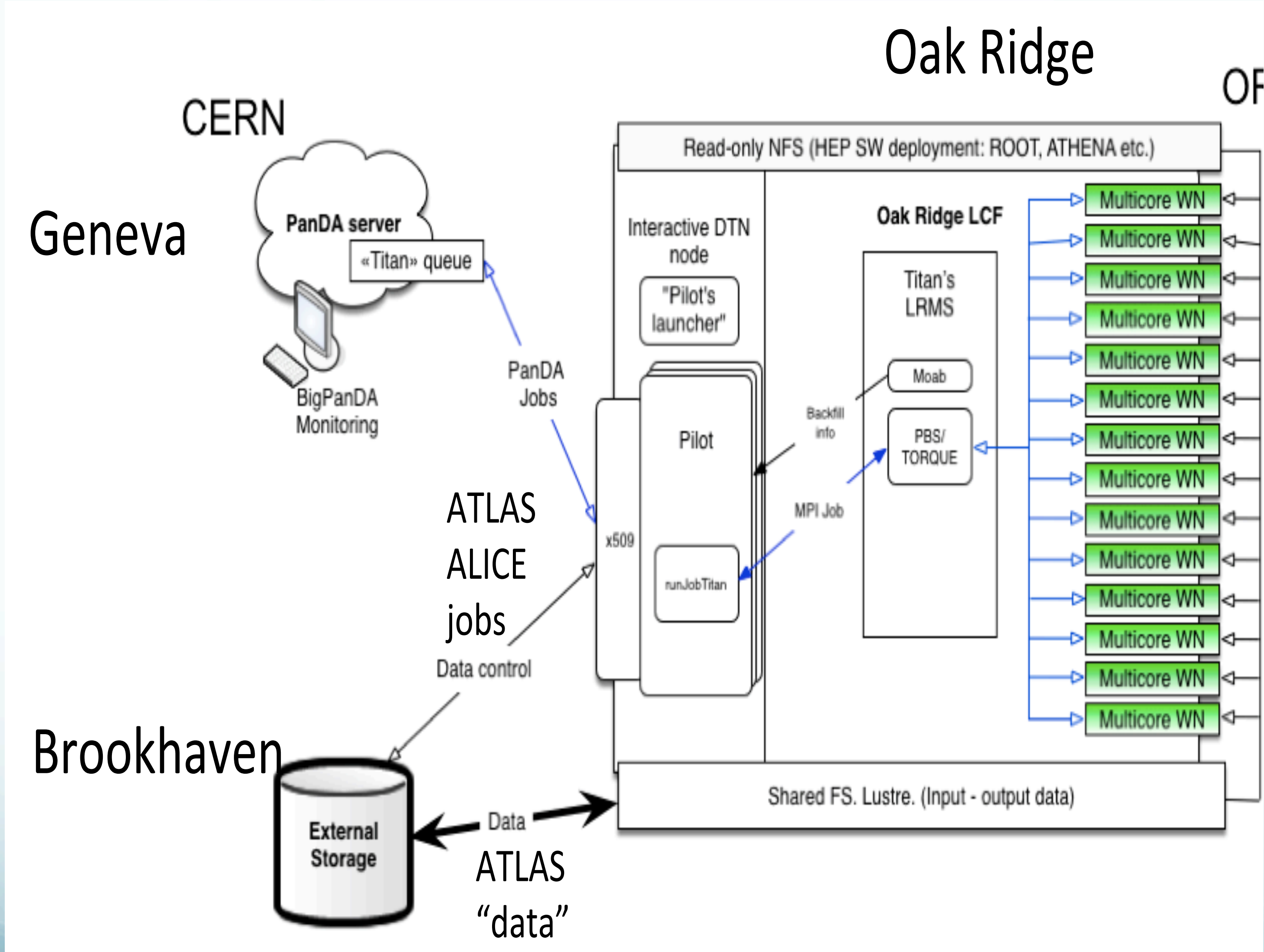
- **WLCG: 750,000 x86 compute cores (ATLAS 300K x86 cores)**
- **Titan: 300,000 x86 compute cores and 18,000 GPUs**

OLCF. Understanding Backfill Slot Availability



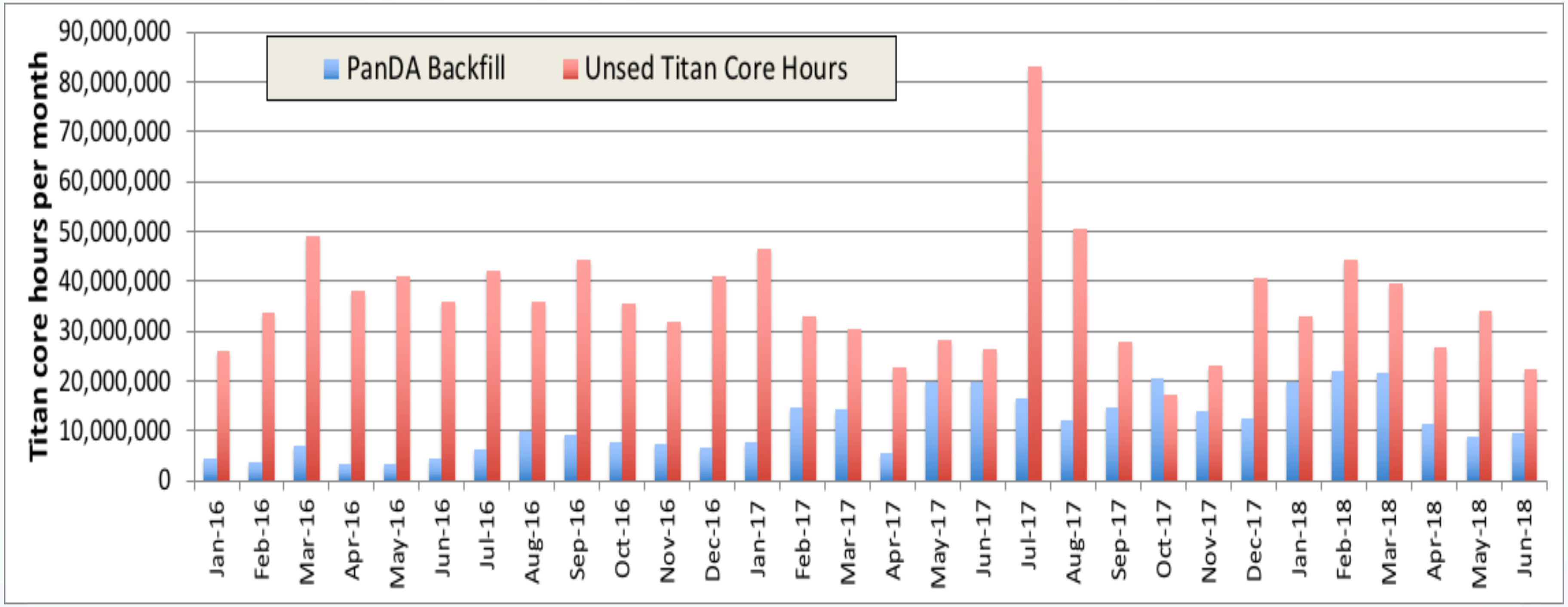
- Mean Backfill availability: 691 worker nodes for 126 minutes.
- Up to 15K nodes for 30-100 minutes
- Large margin of optimization

OLCF Titan Integration with ATLAS Computing



D. Oleynik, S. Panitkin, M. Turilli, A. Angius, S. Oral, K. De, A. Klimentov, J. C. Wells and S. Jha, ”High-Throughput Computing on High-Performance Platforms: A Case Study”, IEEE e-Science (2017) available as: https://arxiv.org/abs/1704.00978

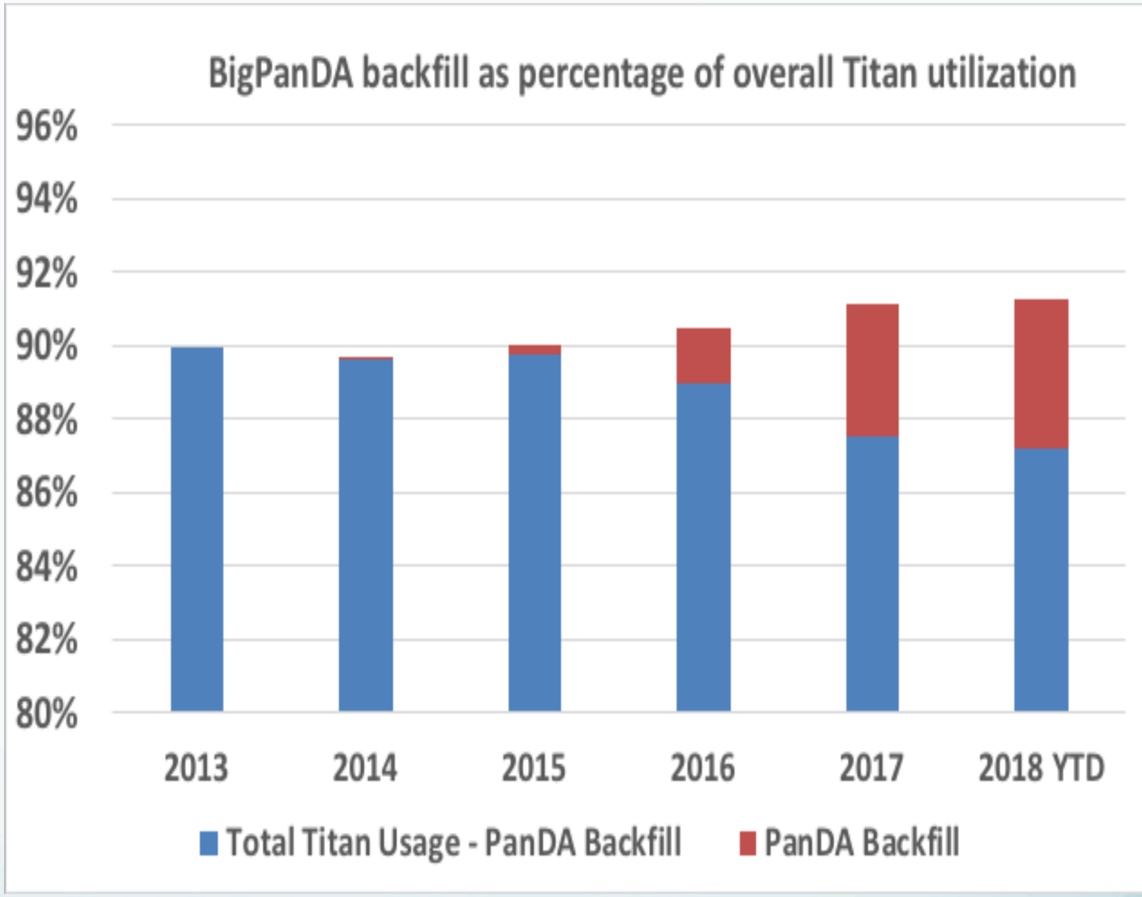
Operational Demo: Scaling

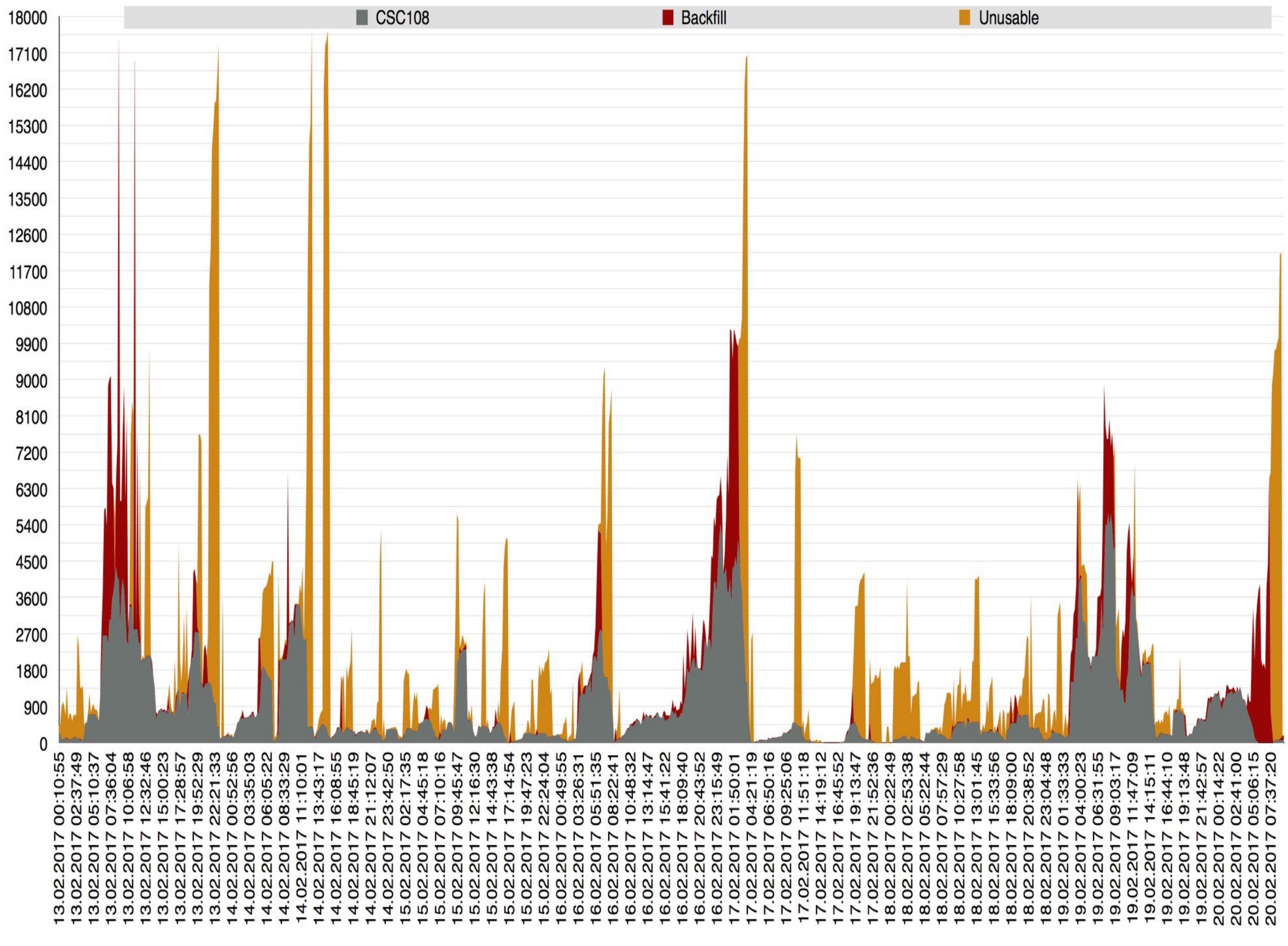


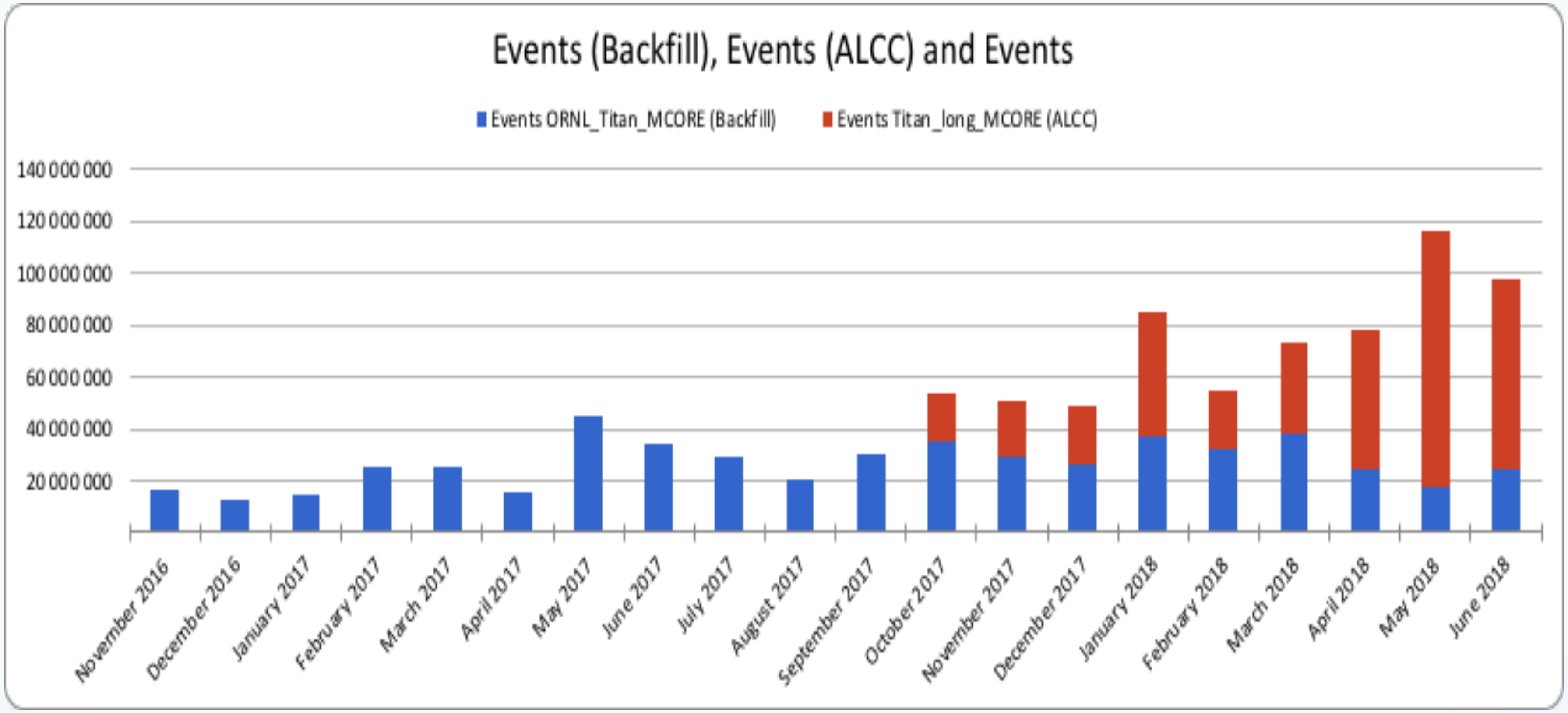
- Consumed 340 Million Titan core hours from January 2016 to present
  - This is 2.9 percent of total available time on Titan over this period
- Remaining used backfill slots are often too short or too small for assigned ATLAS payloads

Operational Demo: Scaling Up Active Backfill

- Increased Titan's utilization by ~2 percent over historical trends
  - May have displaced ~ 2 percent of Titan's traditional throughput
  - Impact on throughput of traditional work loads is under evaluation
- Preemption of PanDA payloads to be evaluated
  - Check pointing needed
  - Will benefit from Event Generator's ability to save incremental results.



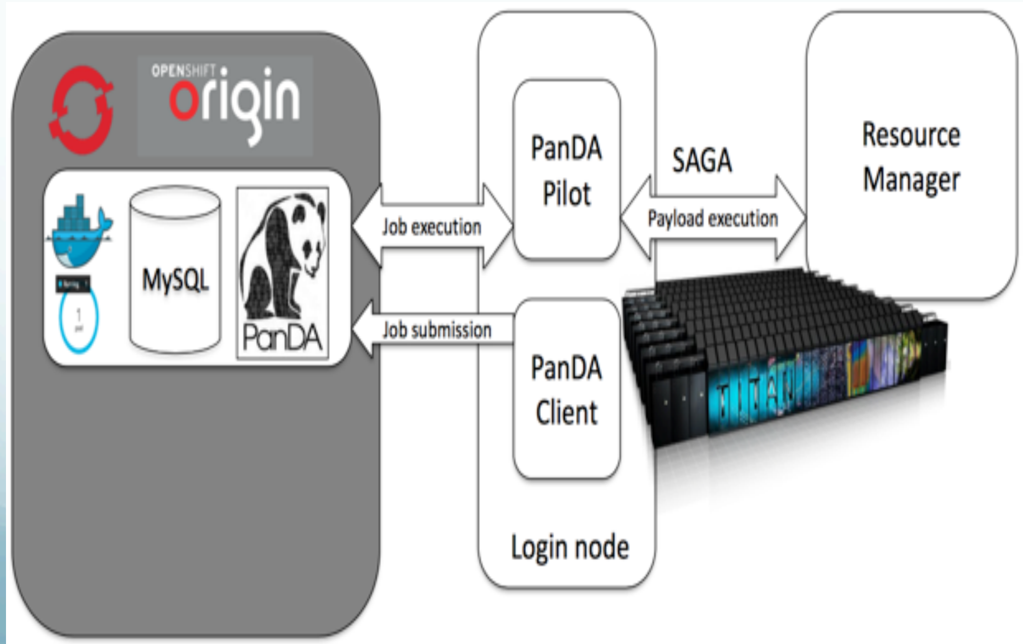

CSC108
Backfill
Unusable
18000
17100
16200
15300
14400
13500
12600
11700
10800
9900
9000
8100
7200
6300
5400
4500
3600
2700
1800
900
0
13.02.2017 00:10:55
13.02.2017 02:37:49
13.02.2017 05:10:37
13.02.2017 07:36:04
13.02.2017 10:06:58
13.02.2017 12:32:46
13.02.2017 15:00:23
13.02.2017 17:28:57
13.02.2017 19:52:29
13.02.2017 22:21:33
14.02.2017 00:52:56
14.02.2017 03:35:03
14.02.2017 06:05:22
14.02.2017 08:33:29
14.02.2017 11:10:01
14.02.2017 13:43:17
14.02.2017 16:08:55
14.02.2017 18:45:19
14.02.2017 21:12:07
14.02.2017 23:42:50
15.02.2017 02:17:35
15.02.2017 04:45:18
15.02.2017 07:10:16
15.02.2017 09:45:47
15.02.2017 12:16:30
15.02.2017 14:43:38
15.02.2017 17:14:54
15.02.2017 19:47:23
15.02.2017 22:24:04
16.02.2017 00:49:55
16.02.2017 03:26:31
16.02.2017 05:51:35
16.02.2017 08:22:41
16.02.2017 10:48:32
16.02.2017 13:14:47
16.02.2017 15:41:22
16.02.2017 18:09:40
16.02.2017 20:43:52
16.02.2017 23:15:49
17.02.2017 01:50:01
17.02.2017 04:21:19
17.02.2017 06:50:16
17.02.2017 09:25:06
17.02.2017 11:51:18
17.02.2017 14:19:12
17.02.2017 16:45:52
17.02.2017 19:13:47
17.02.2017 21:52:36
18.02.2017 00:22:49
18.02.2017 02:53:38
18.02.2017 05:22:44
18.02.2017 07:57:29
18.02.2017 10:27:58
18.02.2017 13:01:45
18.02.2017 15:33:56
18.02.2017 18:09:00
18.02.2017 20:38:52
18.02.2017 23:04:48
19.02.2017 01:33:33
19.02.2017 04:00:23
19.02.2017 06:31:55
19.02.2017 09:03:17
19.02.2017 11:47:09
19.02.2017 14:15:11
19.02.2017 16:44:10
19.02.2017 19:13:48
19.02.2017 21:42:57
20.02.2017 00:14:22
20.02.2017 02:41:00
20.02.2017 05:06:15
20.02.2017 07:37:20



- Since Nov. 2016, 508 million ATLAS events computed via backfill
- Since Oct. 2017, 395 million ATLAS events computed via "normal" batch queue
  - Increases in batch queue event generation beginning in Feb. 2018 show the impact of Harvester

PanDA Server at OLCF: Broad application across domains

- In March 2017 a new PanDA server instance has been set up at ORNL to serve various experiments. This installation the first at OLCF to demonstrate application of a container cluster management and orchestration system, Red Hat OpenShift Origin.
  - We are looking forward for experimenting with GPU payloads with PanDA provided by ATLAS
- OpenShift, when fully in production, will give OLCF users the ability to deploy and manage infrastructure services
  - https://www.olcf.ornl.gov/2017/06/05/olcf-testing-new-platform-for-scientific-workflows/



Key Contributors:
Jason Kincl
Ruslan Mashinistov

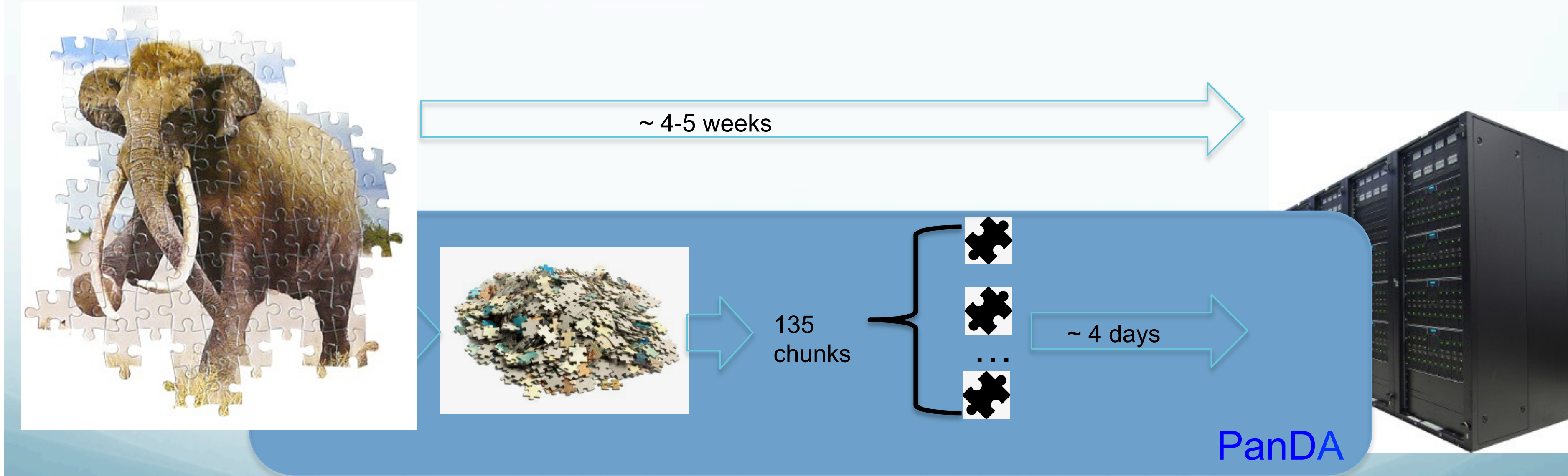
PanDA Server at OLCF: PanDA WMS beyond HEP

- Biology / Genomics: Center for Bioenergy Innovation at ORNL
- Molecular Dynamics: Prof. K. Nam (U. Texas-Arlington)
- IceCube Experiment
- Blue Brain Project (BBP), EPFL
- LSST/DESC (Large Synoptic Survey Telescope) project, Dark Energy Science Collaboration
- LQCD, US QCD Project
- nEDM (neutron Electric Dipole Moment Experiment, ORNL

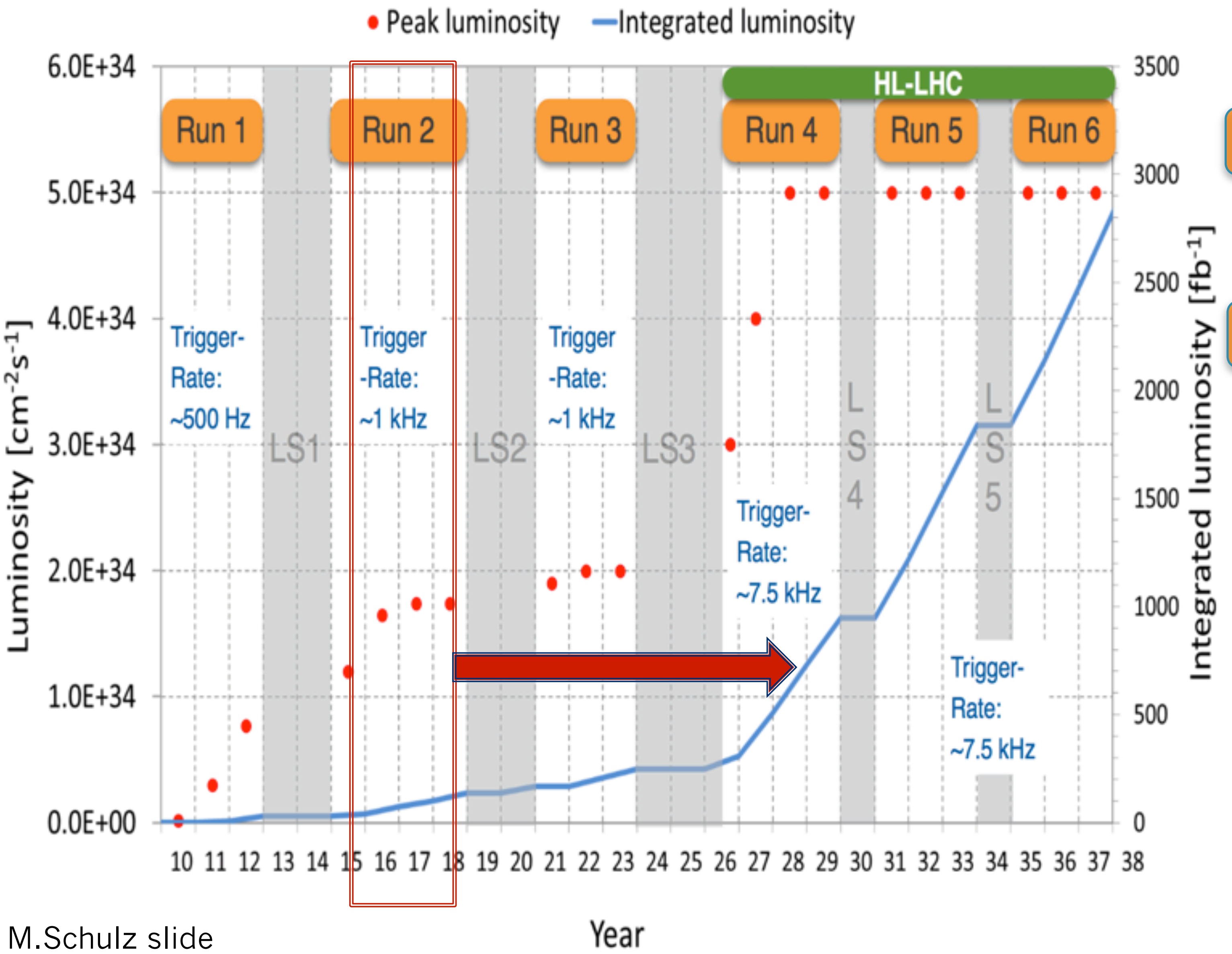
See D.Oleynik talk for more details

BigPanDA in Genomics. RF mega-grant project (PanDA server @ NRC KI)

- At NRC KI PALEOMIX pipeline was adapted to run on local supercomputer resources powered by PanDA.
- It was used to map mammoth DNA on the African elephant reference genome
- Using software tools developed initially for HEP and Grid reduced genomics payload execution time for Mammoths DNA sample from weeks to days.



Towards High Luminosity LHC



Run 3 — ALICE, LHCb upgrades

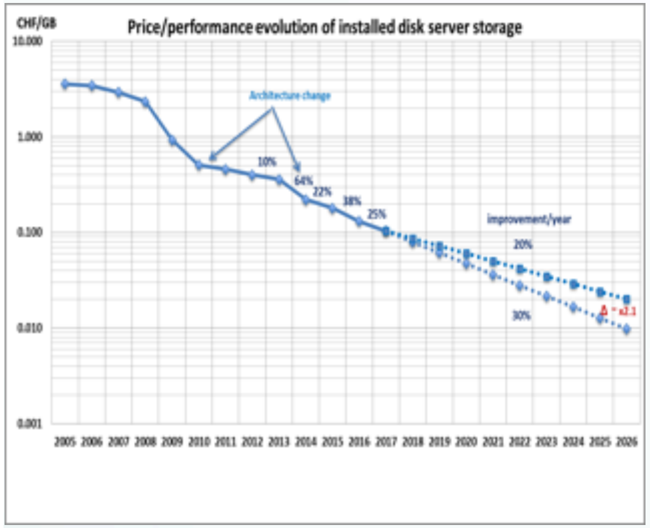
Run 4 — ATLAS, CMS upgrades

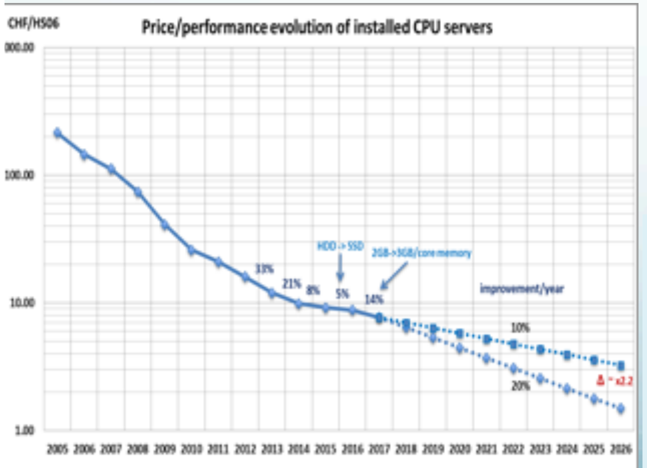
M.Schulz slide

Future Challenges

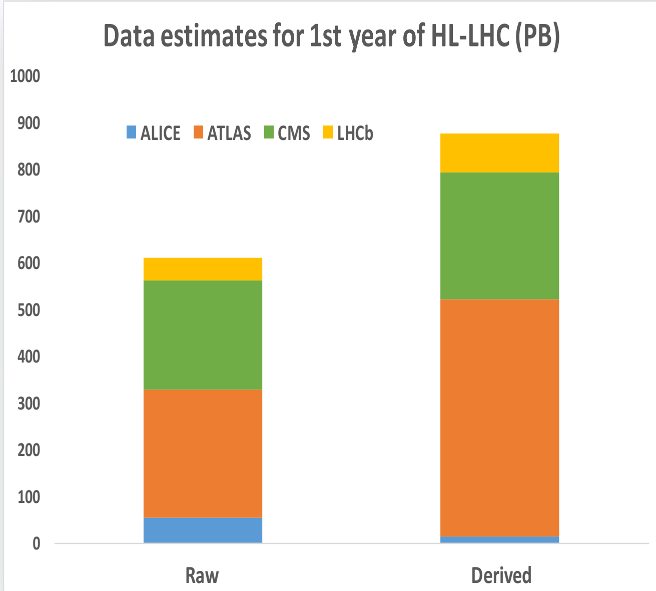
2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	……	2030?

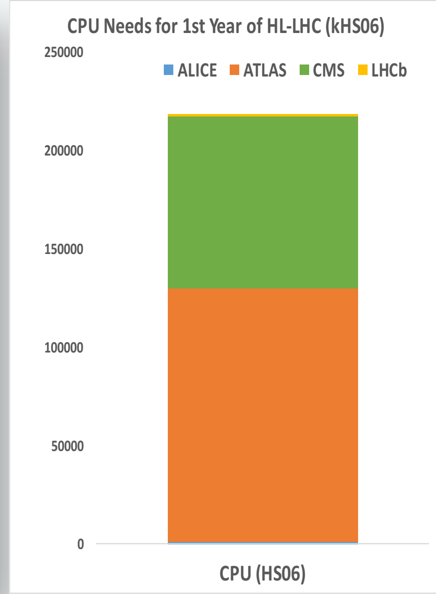
First run | LS1 | Second run | LS2 | Third run | LS3 | HL-LHC | FCC?





I.Bird slide





Data:
- RAW 2016: 50 PB → 2027: 600 PB
- Derived (1 copy): 2016: 80 PB → 2027: 900 PB

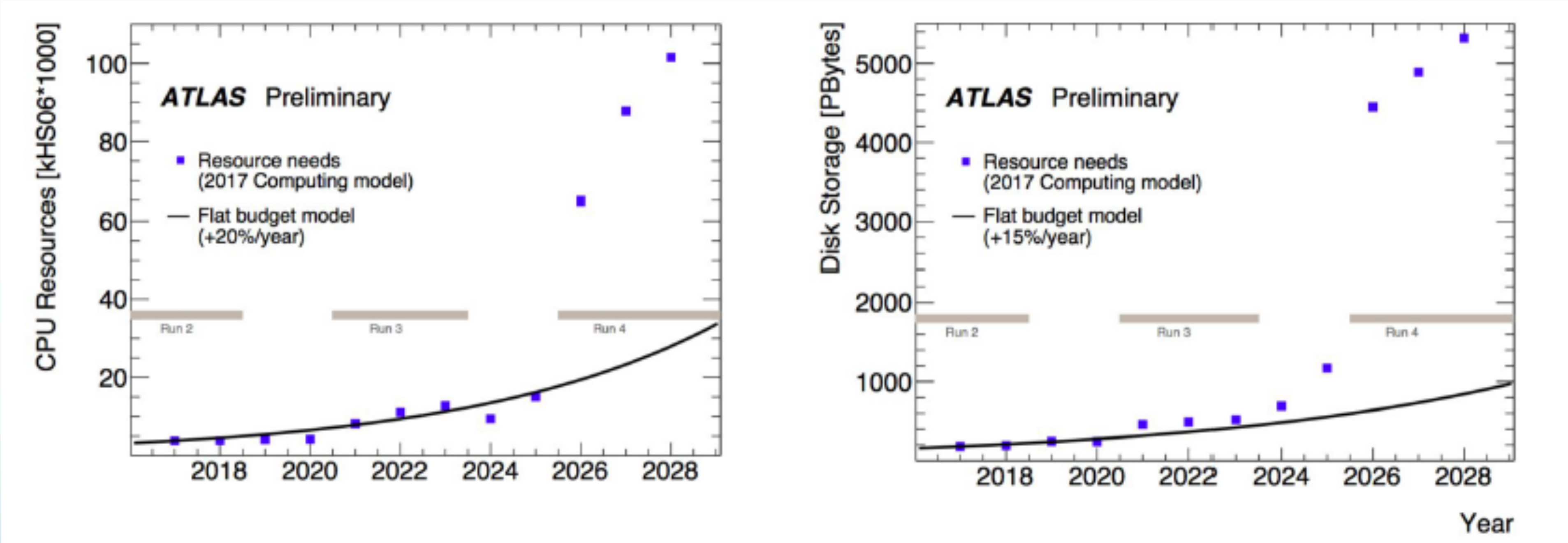
CPU:
- x60 from 2016

- RAW data volume for LHC increases exponentially and with it processing and analysis load
- Technology at ~20%/year will bring x6-10 in 10-11 years

Estimates of resource needs at HL-LHC x10 above what is realistic to expect from technology with reasonably constant cost

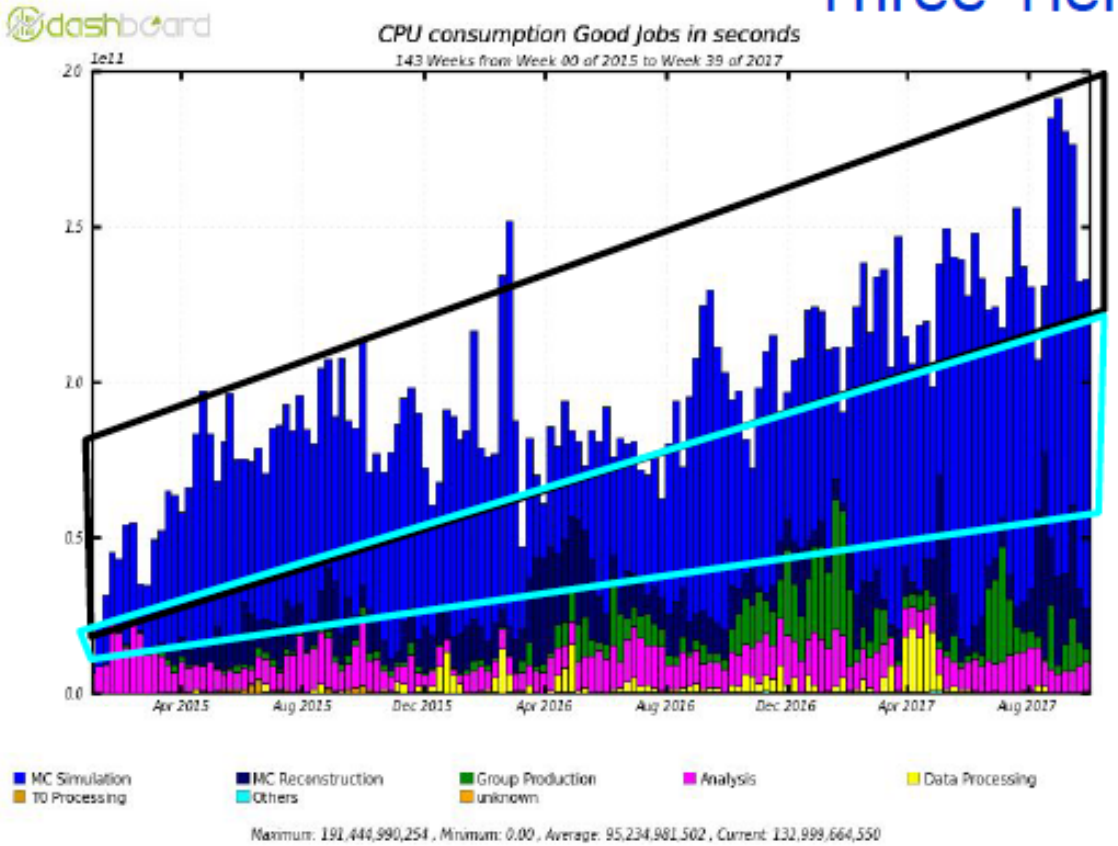
ATLAS EXPERIMENT

Challenges Ahead



ATLAS CPU Consumption Since 2015

Three Tiers:



Utilization spikes. May be cost effective to offload to on-demand resources

CPU-intensive (simulation)
Well suited for HPCs and Clouds

Data Intensive Processing
Best run on grid

P. Calafiura

HTC+HPC+Cloud – potential sharing

Conclusion

- Over the past 5 years, the ATLAS experiment at the LHC collaborated with many large HPC sites for full integration into ATLAS distributed computing
- Success of this effort is seen on physics results
  - From ongoing Run 2 at the LHC
  - In the large number of LCF and HPC sites used for processing
  - In the new software innovations – Containerization, Harvester, Event Service, ATLAS release build on HPC
  - HPCs needed for future challenges at the LHC
  - ATLAS is ready for future Exascale machines

Acknowledgements

- Thanks to members of the ATLAS Collaboration
- Special thanks for contributions by: Doug Benjamin, Paolo Calafiura, Taylor Childers, Pablo Fernandez, Andrej Filipcic, Rob Gardner, Wen Guan, Andreu Pacheco Pages, Kevin Kissel, Artem Petrosyan, Gianfranco Sciacca, Rodney Walker, Torre Wenaus, Wei Yang
- Thanks to LIT JINR members contributed to BigPanDA and PanDA projects
- The BigPanDA project was funded by the U. S. Department of Energy, Office of Science, High Energy Physics and Advanced Scientific Computing under Contracts No. DE-AC02-98CH10886 and DE-AC02-06CH11357
- The presented research work at NRC KI was funded by the Russian Ministry of High Education and Science under Contract No. 14.Z50.31.0024