US ATLAS Distributed Facilities Meeting at the University of Utah

Introduction

Not all topics relevant for the facility will be covered today

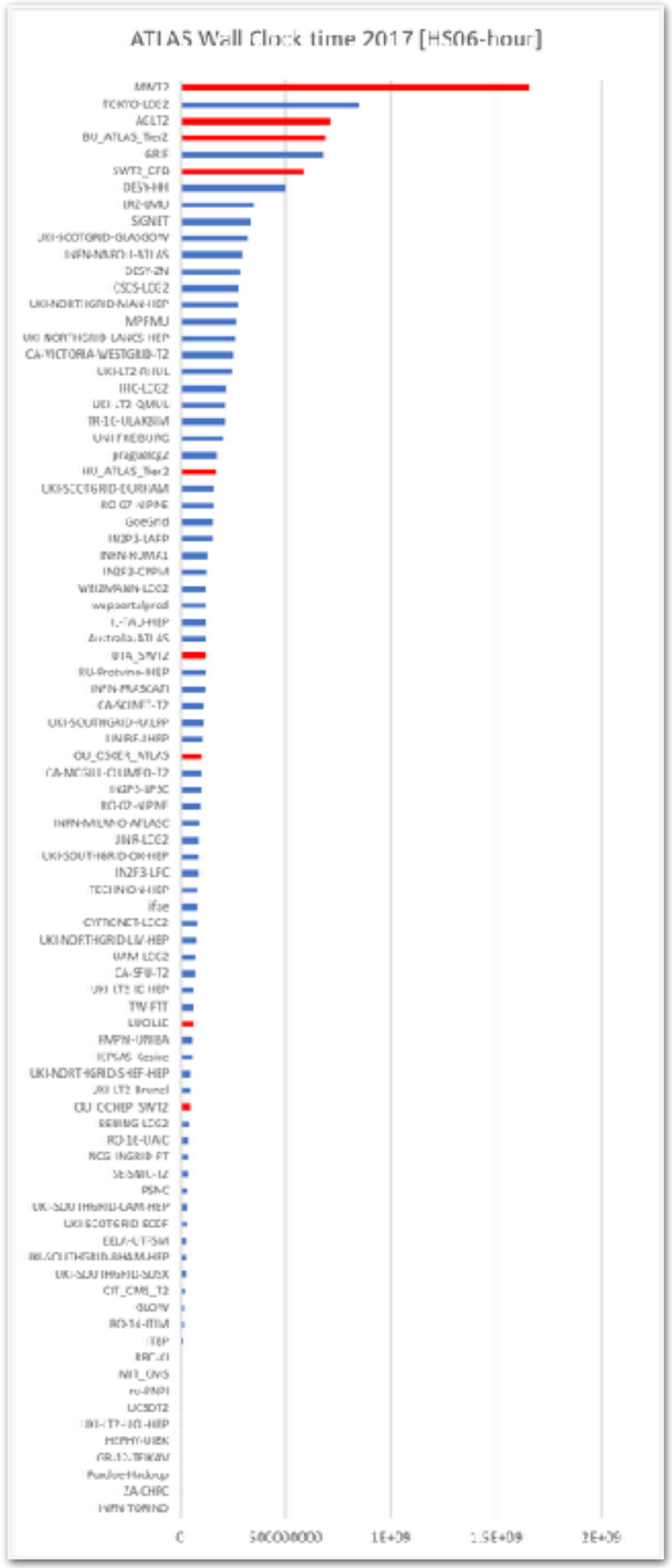
A lot of time for questions and discussions

Live notes on the indico page

Tier 2s

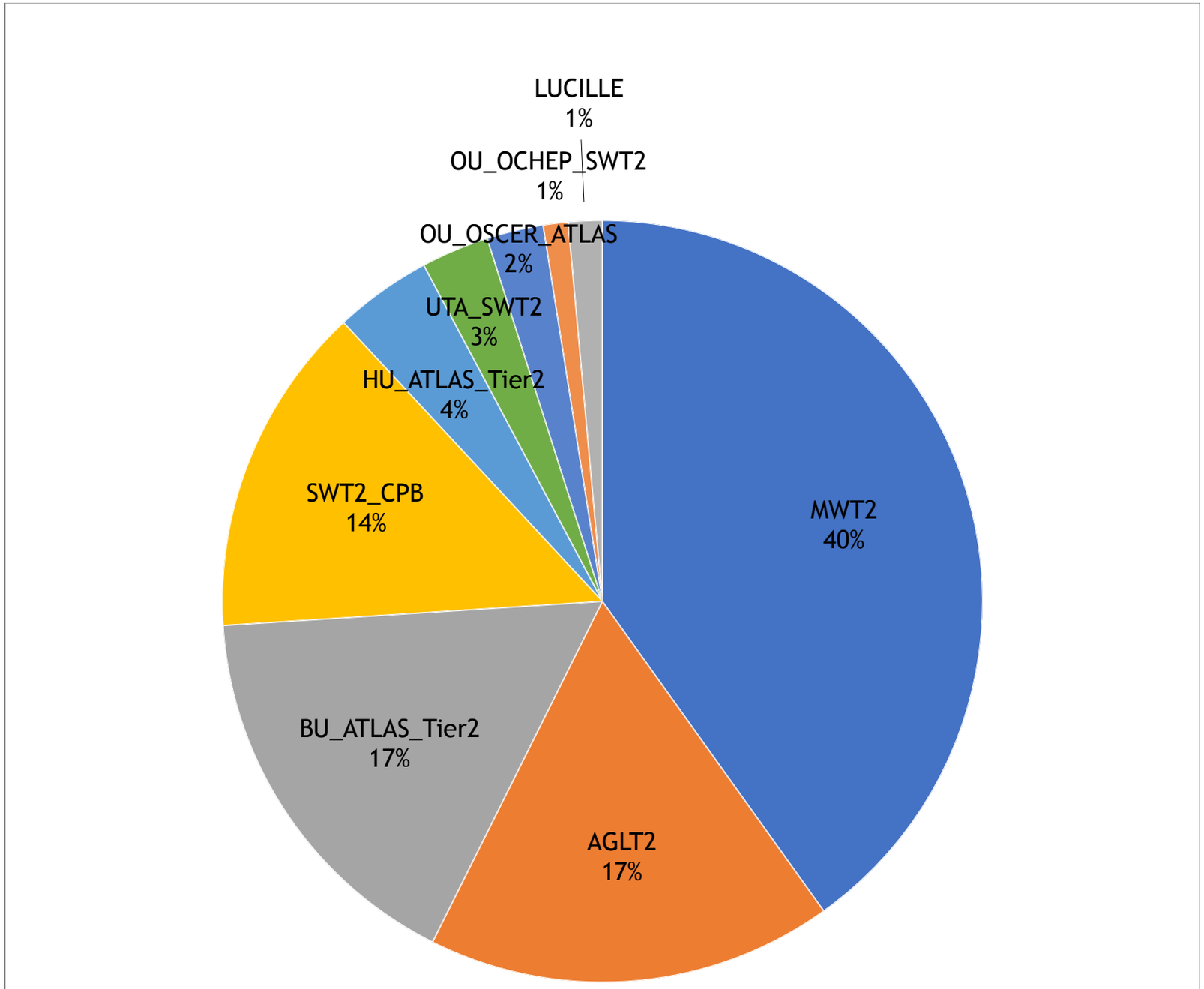
Wall clock time [HS06*h]
2017 data
86 computing sites
contributed to ATLAS
US Tier 2 sites 29.5%

23% the canonical share



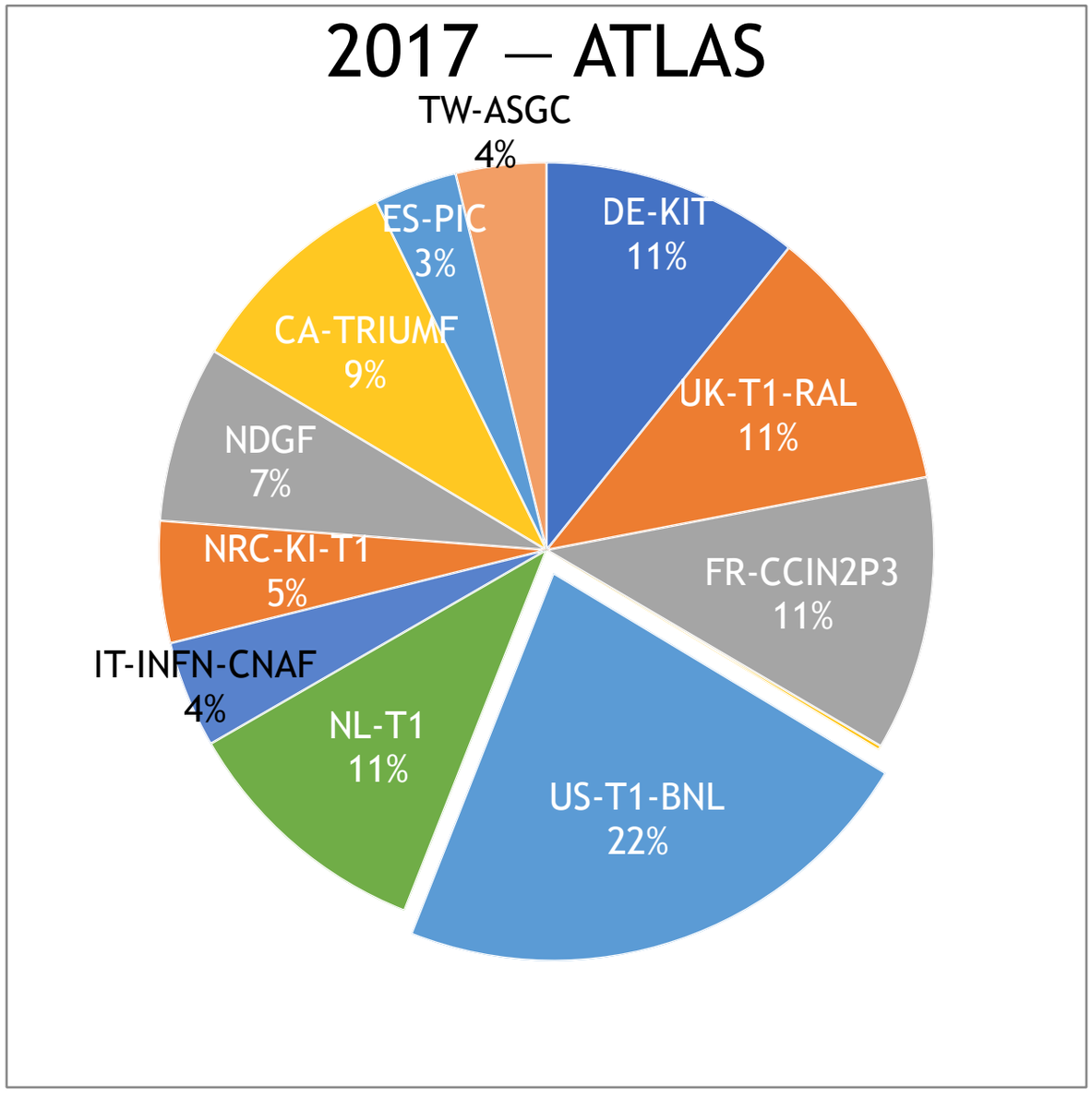
https://accounting.egi.eu/wlcg/tier2/normelap_processors/SITE/VO/2017/1/2017/12/lhc/onlyinfrajobs/

2017 wall clock by Tier 2



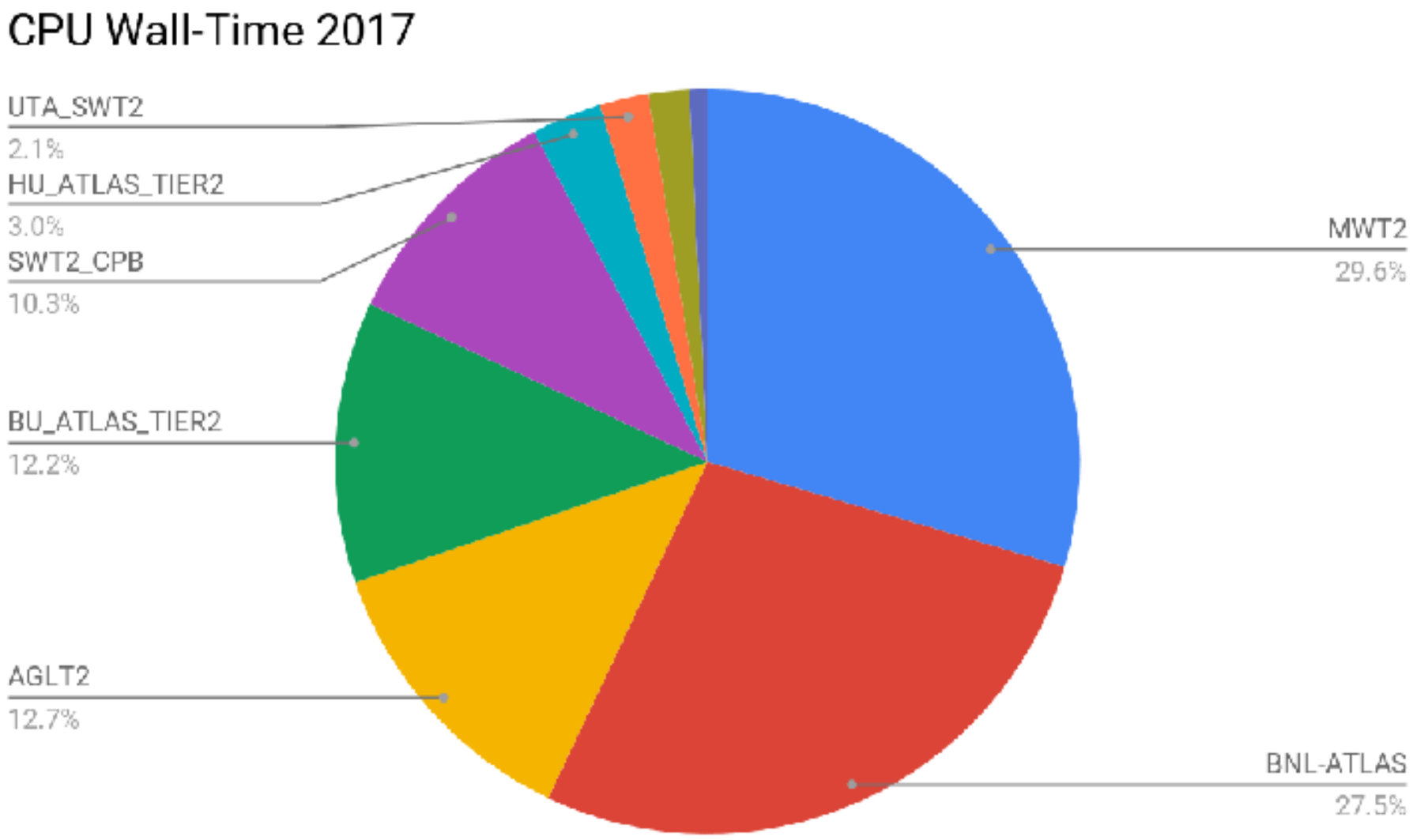
Tier 1s

Wall clock time [HS06*h]
2017 data
US Tier 1 22%



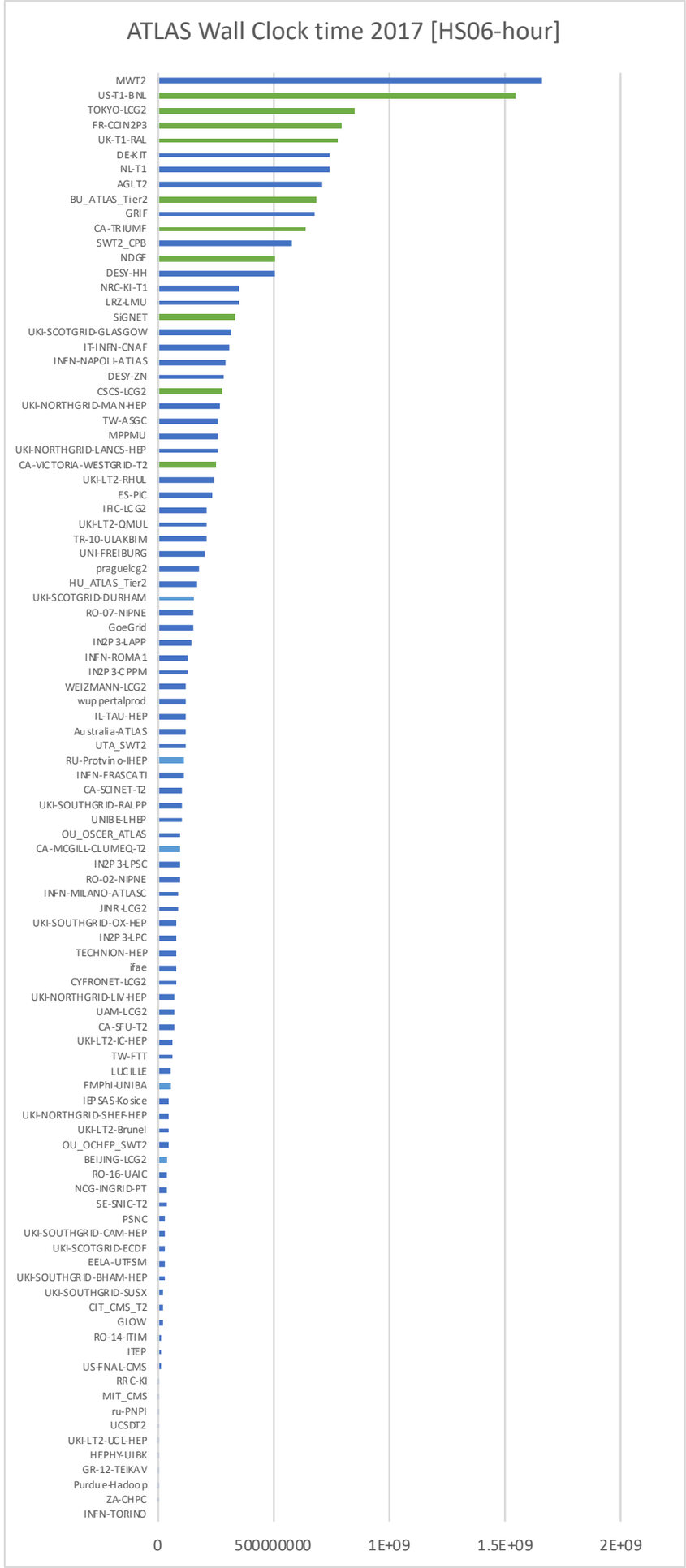
https://accounting.egi.eu/wlcg/tier1/normelap_processors/TIER1/VO/2017/1/2017/12/lhc/onlyinfrajobs/

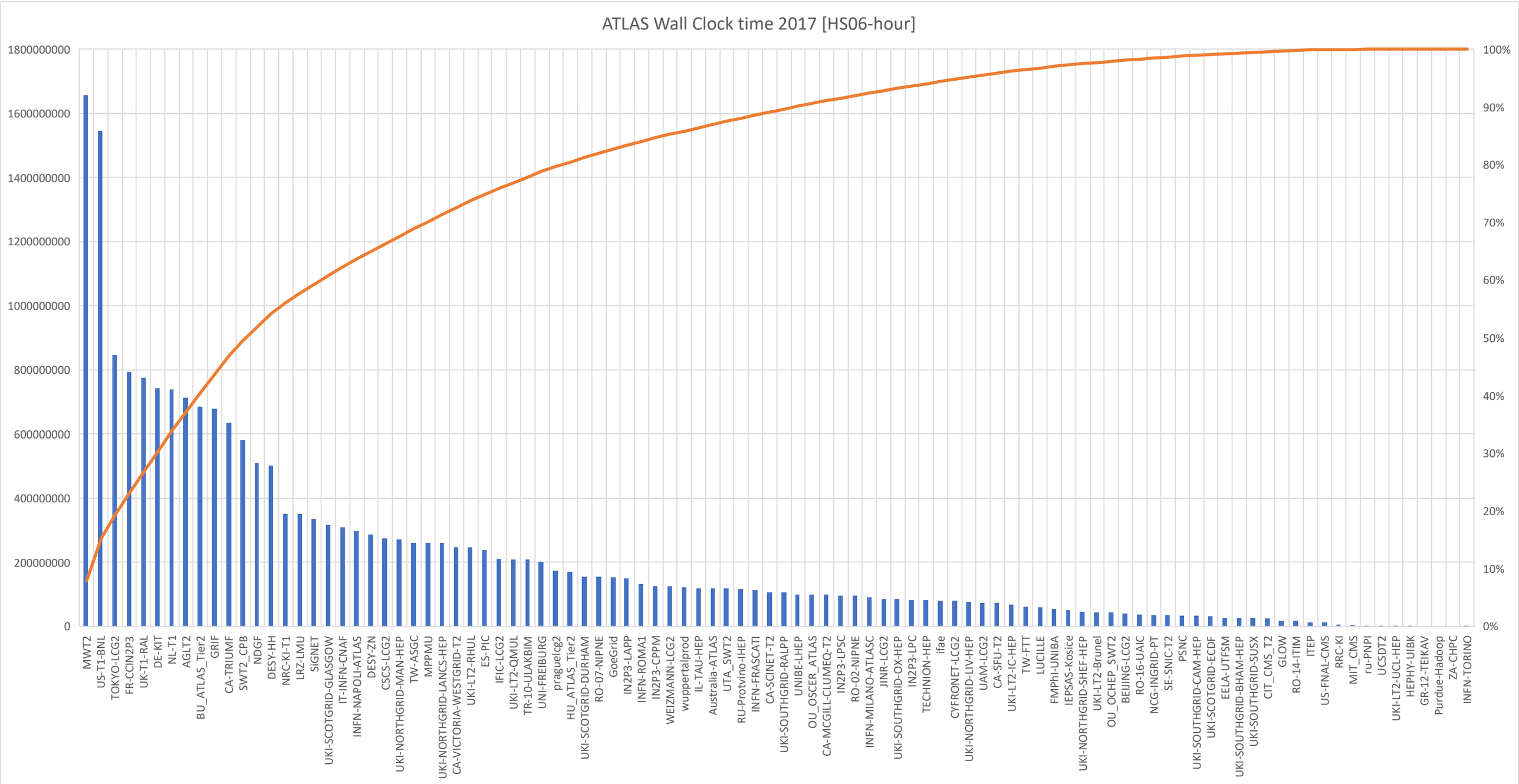
US Tier 1 & Tier 2s



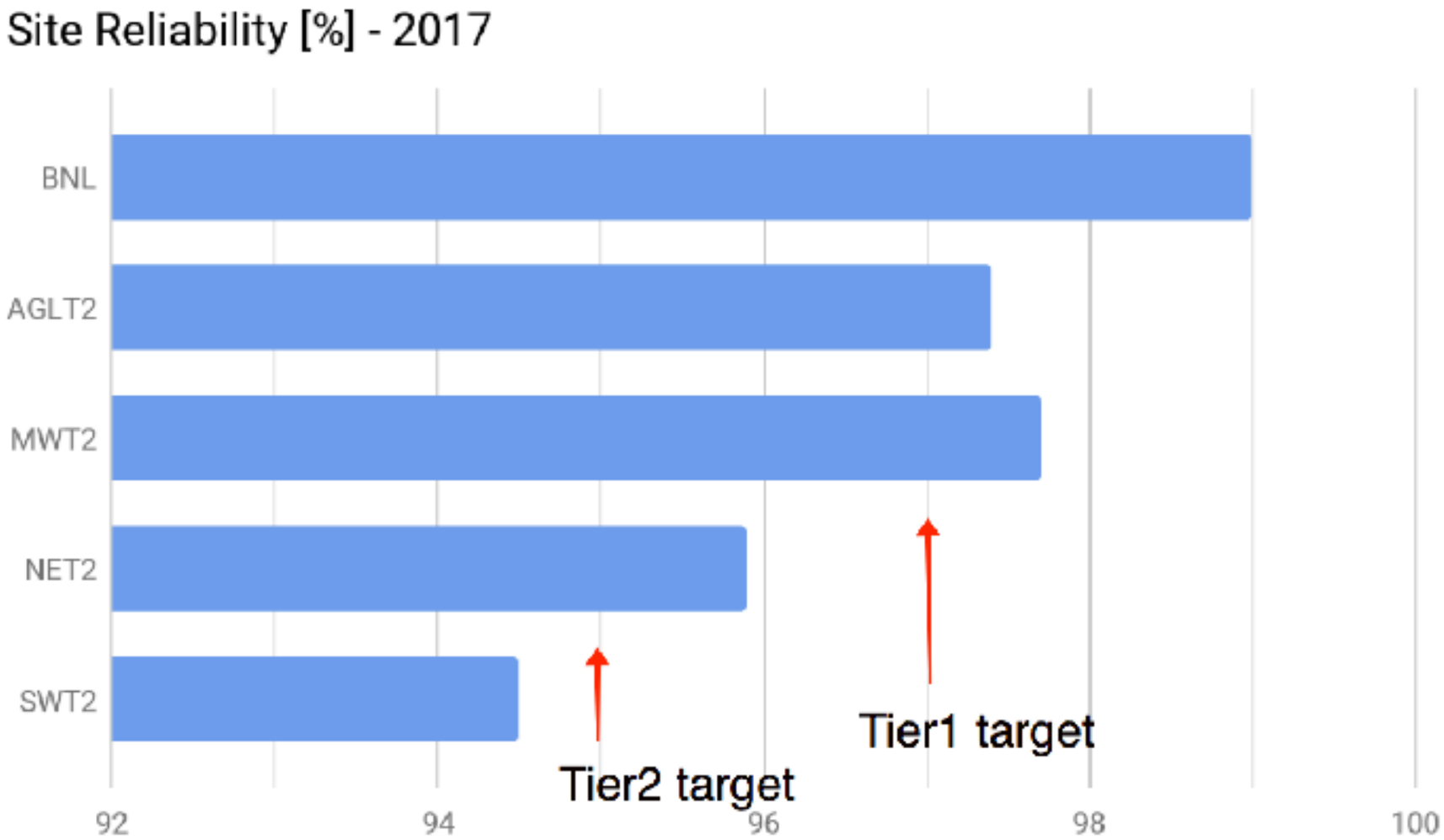
All ATLAS Tier 1s & Tier 2s

Wall clock time [HS06*h]
2017 data
98 computing sites
contributed to ATLAS
Tier 1 sites in green




ATLAS Wall Clock time 2017 [HS06-hour]
1800000000
1600000000
1400000000
1200000000
1000000000
800000000
600000000
400000000
200000000
0
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
MWT2
US-T1-BNL
TOKYO-LCG2
FR-CCIN2P3
UK-T1-RAL
DE-KIT
NL-T1
AGLT2
BU_ATLAS_Tier2
GRIF
CA-TRIUMF
SWT2_CPB
NDGF
DESY-HH
NRC-KI-T1
LRZ-LMU
SiGNET
UKI-SCOTGRID-GLASGOW
IT-INFN-CNAF
INFN-NAPOLI-ATLAS
DESY-ZN
CSCS-LCG2
UKI-NORTHGRID-MAN-HEP
TW-ASGC
MPPMU
UKI-NORTHGRID-LANCS-HEP
CA-VICTORIA-WESTGRID-T2
UKI-LT2-RHUL
ES-PIC
IFIC-LCG2
UKI-LT2-QMUL
TR-10-ULAKBIM
UNI-FREIBURG
praguelcg2
HU_ATLAS_Tier2
UKI-SCOTGRID-DURHAM
RO-07-NIPNE
GoeGrid
IN2P3-LAPP
INFN-ROMA1
IN2P3-CPPM
WEIZMANN-LCG2
wuppertalprod
IL-TAU-HEP
Australia-ATLAS
UTA_SWT2
RU-Protvino-IHEP
INFN-FRASCATI
CA-SCINET-T2
UKI-SOUTHGRID-RALPP
UNIBE-LHEP
OU_OSCER_ATLAS
CA-MCGILL-CLUMEQ-T2
IN2P3-LPSC
RO-02-NIPNE
INFN-MILANO-ATLASC
JINR-LCG2
UKI-SOUTHGRID-OX-HEP
IN2P3-LPC
TECHNION-HEP
ifae
CYFRONET-LCG2
UKI-NORTHGRID-LIV-HEP
UAM-LCG2
CA-SFU-T2
UKI-LT2-IC-HEP
TW-FTT
LUCILLE
FMPhI-UNIBA
IEPSAS-Kosice
UKI-NORTHGRID-SHEF-HEP
UKI-LT2-Brunel
OU_OCHEP_SWT2
BEIJING-LCG2
RO-16-UAIC
NCG-INGRID-PT
SE-SNIC-T2
PSNC
UKI-SOUTHGRID-CAM-HEP
UKI-SCOTGRID-ECDF
EELA-UTFSM
UKI-SOUTHGRID-BHAM-HEP
UKI-SOUTHGRID-SUSX
CIT_CMS_T2
GLOW
RO-14-ITIM
ITEP
US-FNAL-CMS
RRC-KI
MIT_CMS
ru-PNPI
UCSDT2
UKI-LT2-UCL-HEP
HEPHY-UIBK
GR-12-TEIKAV
Purdue-Hadoop
ZA-CHPC
INFN-TORINO

Reliability of US sites



“Is there a US squad team ?”

–X from ADC

Communication

- Please use our mailing list and bi-weekly meetings as the first place for question and discussion

- Do not overload ADC operation

Capacity & pledges

Site	**Available Disk [TB]**	Pledges [TB] 2017	Pledges [TB] 2018	**Available CPU [kHS06]**	Pledges [kHS06] 2017	Pledges [kHS06] 2018
Tier-1	**16,500**	15,600	17,000	**156**	212	218
AGLT2	**6,440**	4,300	4,700	**96**	58	59
MWT2	**6,370**	6,400	7,000	**196**	86	89
NET2	**4,580**	4,300	4,700	**111**	58	59
SWT2	**4,880**	4,300	4,700	**129**	58	59
US ATLAS Tier2s	**22,270**	19,300	21,100	**532**	260	266
US ATLAS Facility	**38,770**	34,900	38,100	**688**	472	484

Retirement profile for next years needed

Reminder : WLCG year starts April 1st

FY18 milestones

https://docs.google.com/document/d/1gk5Hc-WTmdjCveYtGQ3Hy_iVXGIU2IbnVut6pYtMXCw/edit?usp=sharing

Q1

- Upgrade analytics platform @ MWT2
- Operation optimisation / accounting on Leadership HPCs (likely Q1 through Q4)
    - Develop a plan for an US ATLAS Leadership HPC operations team. This principally includes the DOE centers: NERSC (Cori), ALCF (Theta) and ORNL (Titan).
    - Standardize solutions for software distribution where possible including thin-container approaches.
    - Standardize data delivery where possible, including a Globus-based distribution service (developed and hosted by ANL)
    - Automatic benchmarking and consistent accounting on various HPCs
- T3 capacity increase at BNL
- SLAC validation as ACF
- Characterisation of ML analysis platform (HW, SW)

Q2

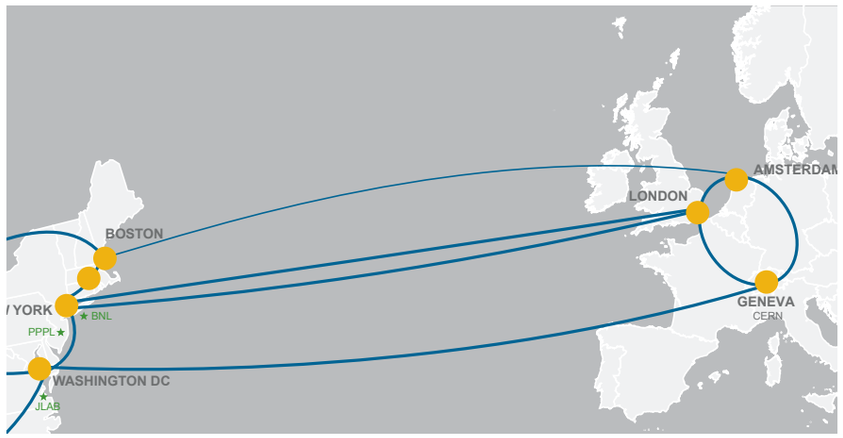
- Jupyter notebook available for test at BNL T3, T2
- Deploy Xcache at every T2
- Software distribution on NSF HPC Centers
- Distributed Ceph system testbed plan
- NET2 storage migration to Ceph

Q3

- Develop ML analysis platform (HW, SW)
  - Plan is to build a ML/AI platform for general use by the ATLAS community
- Network monitoring
  - Integrated Maddash & analytics dashboard
- Distributed Ceph system testbed deployment
- WLCG 2018 pledges delivered

Q4

- T1, T2 sites migrated to CentOS 7
    - ATLAS plans (https://twiki.cern.ch/twiki/bin/view/AtlasComputing/CentOS7Readiness
- Singularity is deployed on all sites
    - ATLAS plans
    - US: https://twiki.cern.ch/twiki/bin/viewauth/AtlasComputing/ContainersInUScloud
    - ADC: https://twiki.cern.ch/twiki/bin/view/AtlasComputing/ADCContainersDeployment



ESnet European Funding Model

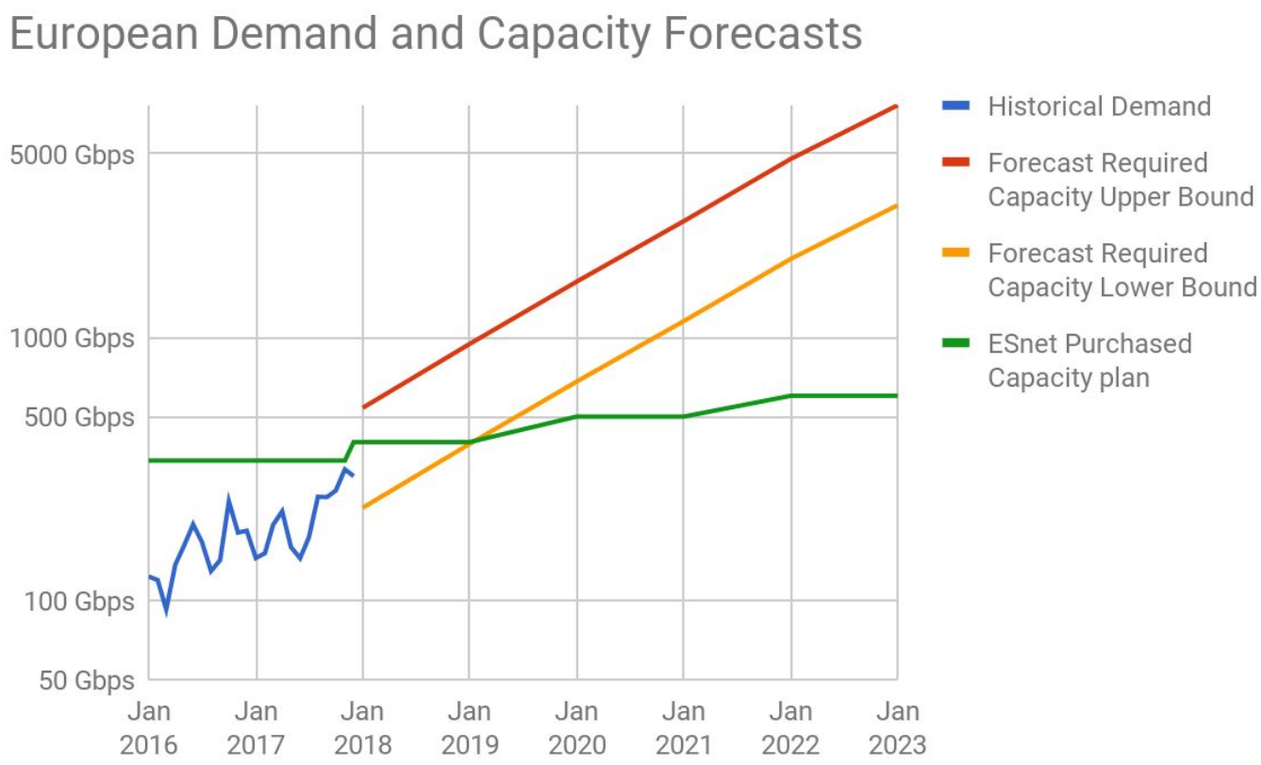
Overview for LHC Experiment Coordinators Meeting

Last Updated: 22-January-2018

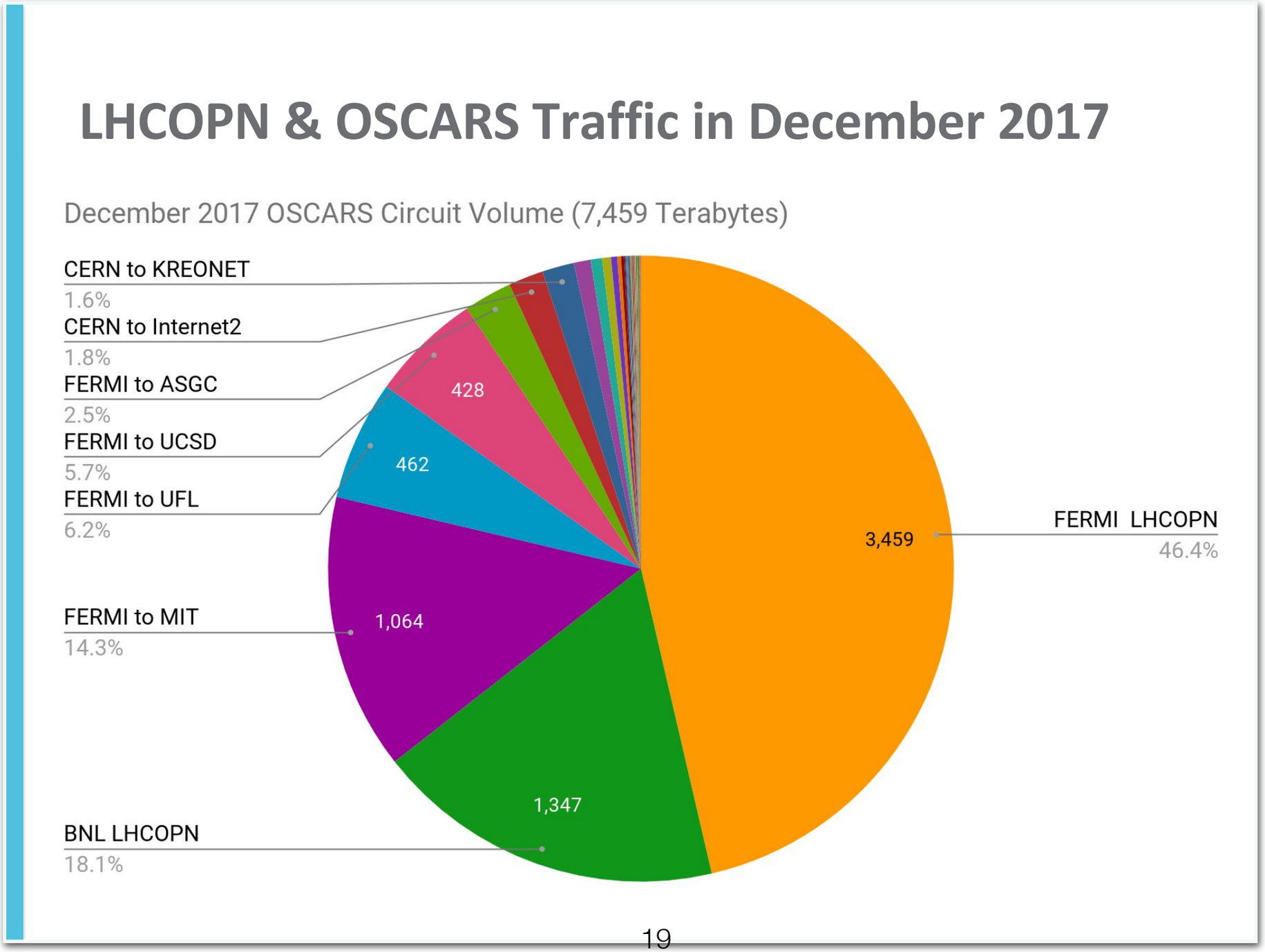
ESnet's initial transatlantic networking infrastructure was commissioned in late 2014, and was funded by both the ESnet and HEP Programs. HEP funding is $6M thru FY18. The initial transatlantic capacity of 340Gbps, has been augmented to 400Gbps by the end of 2017. Beyond FY18, ESnet's multi-year budget plan allocates ~$2M annually for transatlantic capacity, including growth. This amount of funding should allow regular increases in capacity in the future as transatlantic circuits and european circuit costs continue to drop. However, we note that this funding level may not support projected capacity requirements in the future (see table below).

Network

- Was our best friend up to now
- Need to be considered as a true resource and not an externally granted resource in futire



Something to be understood

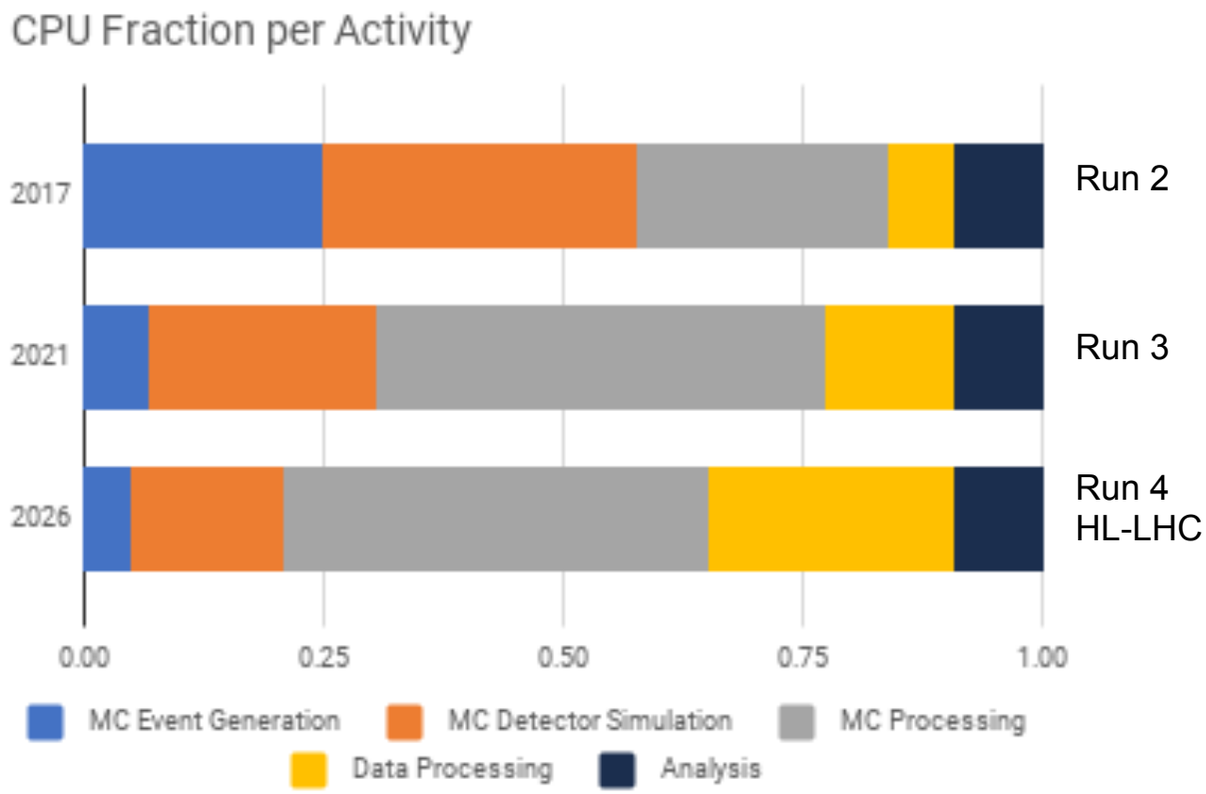


Allocation on DOE HPC

US ATLAS Computing Needs and Plans Leading up to HL-LHC

P. Calafiura, K. De, E. Lançon

- Document submitted to DOE (June 2017)
- Extrapolation of current ATLAS computing model up to 2030!
- Low I/O workflows only are suitable for (current) HPCs
- Today ~50% of ATLAS CPU consumption from low I/O workflows
  - Event generation
  - Simulation
- ATLAS allocation request on DOE HPCs for 75% low I/O workflows
- **Caveat** : ATLAS CPU needs & usage have always been higher than computing model
  - ~20% more at Tier-1s
  - ~100% more at Tier-2s



At HL-LHC, because of pile-up, event (**real** and **simulation** data) reconstruction dominates (~70%) CPU budget

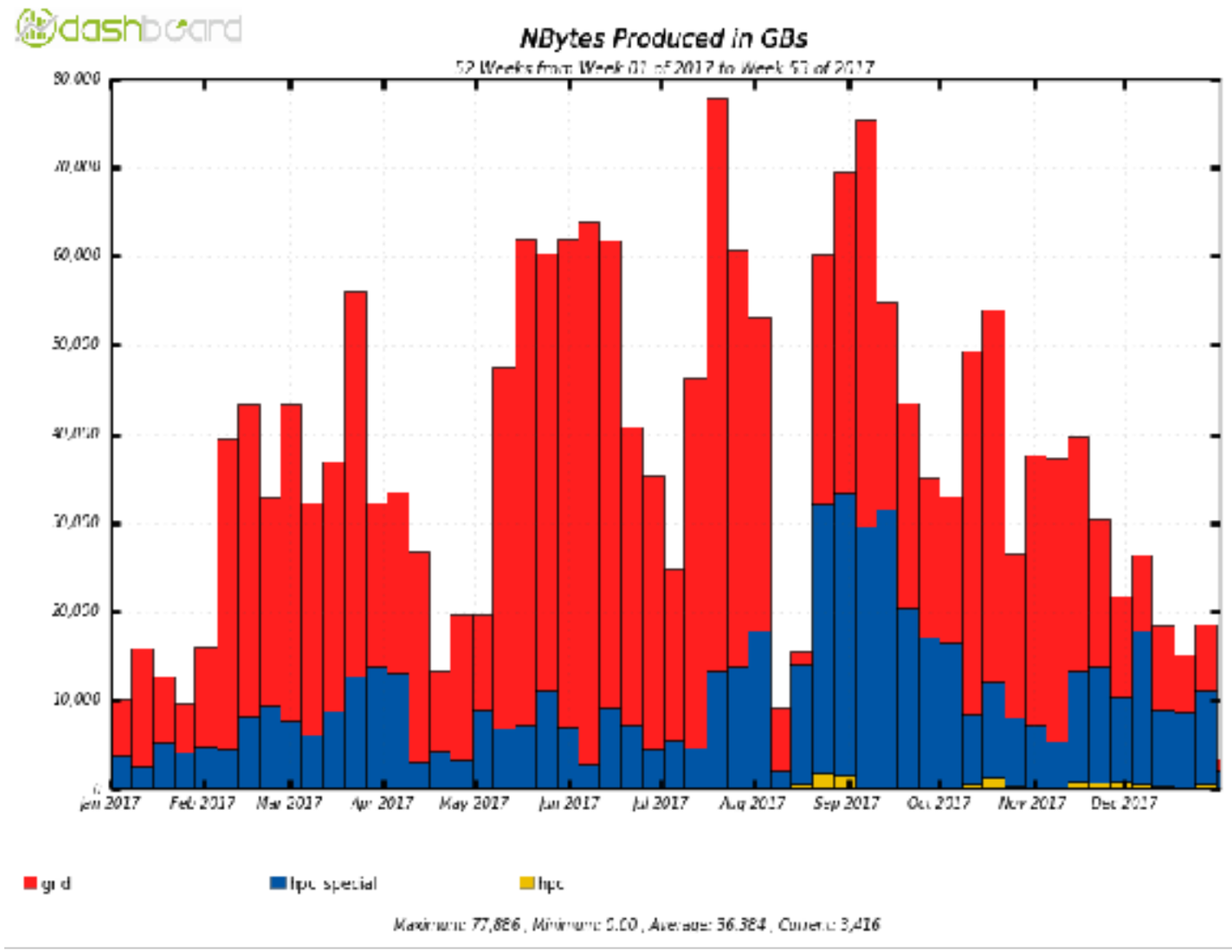
US ATLAS HPC CPU Allocation Request (Millions of core-hours)												
2018	**2019**	**2020**	**2021**	**2022**	**2023**	**2024**	**2025**	**2026**	**2027**	**2028**	**2029**	**2030**
240	410	550	400	400	400	650	1140	2930	2930	2930	2930	2930

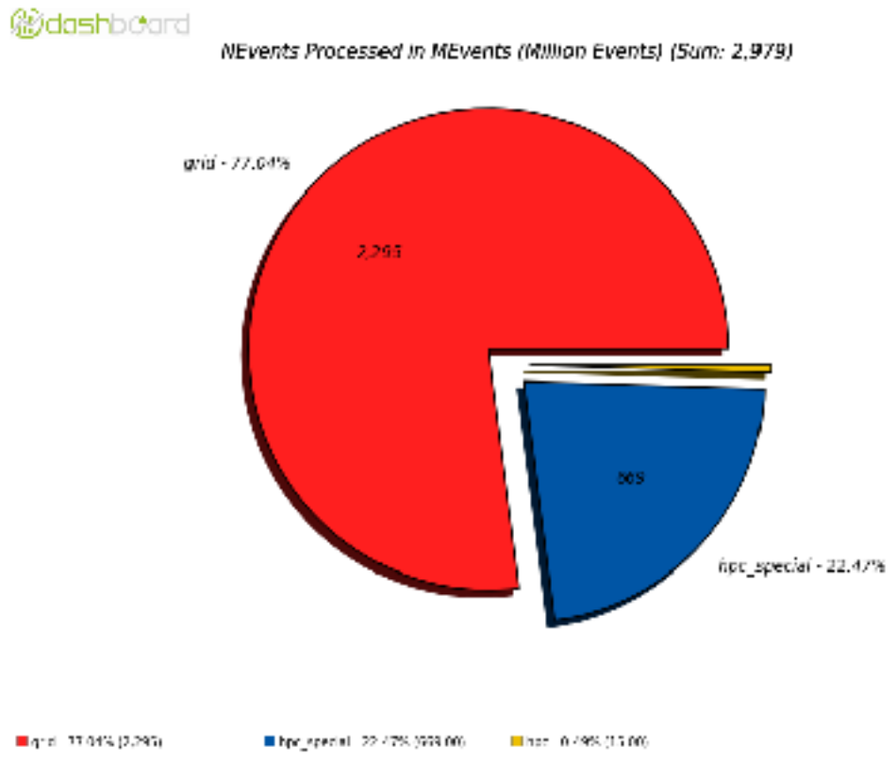
Table 4: US ATLAS HPC CPU Allocation Request

We need HPCs in the future

How are we using them today ?
CPU accounting unclear, used # of simulated events for accounting

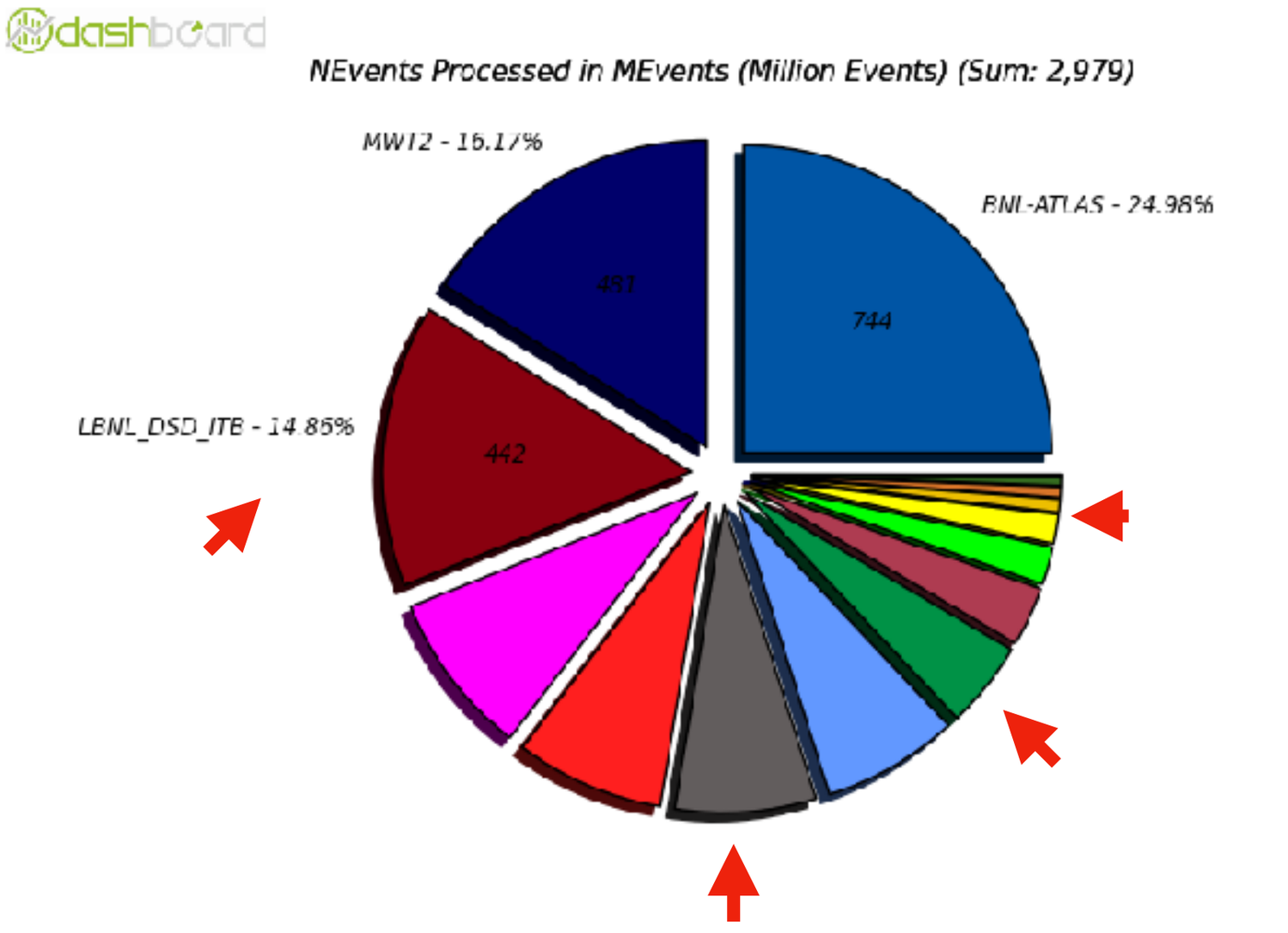
MCfull in the US





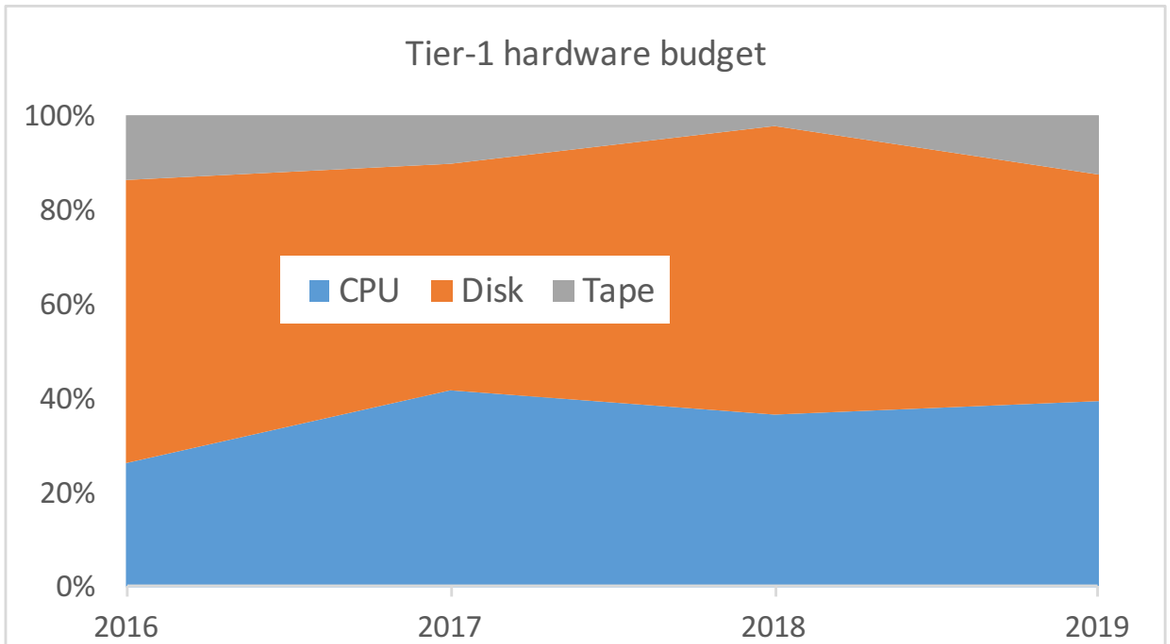
HPCs ~23% of simulation produced in the US
And HPCs do ~only full simulation

MCfull in the US



Options for addressing storage cost

- Use cheaper hardware
- Avoid duplication of data
  - Caching
  - Distributed storage



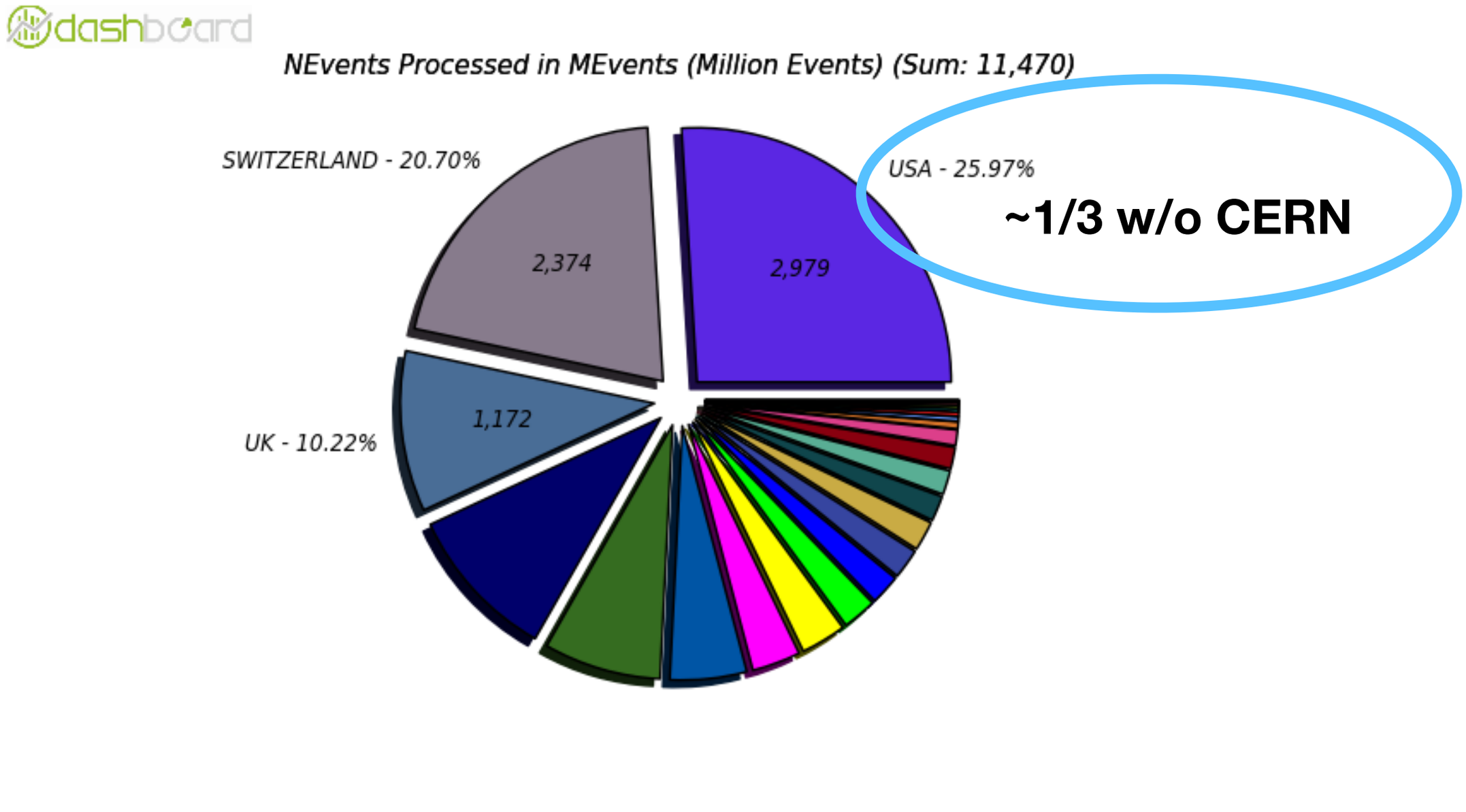
Disk storage is the main driver of Tier-1 hardware equipment cost (tape robot not considered)

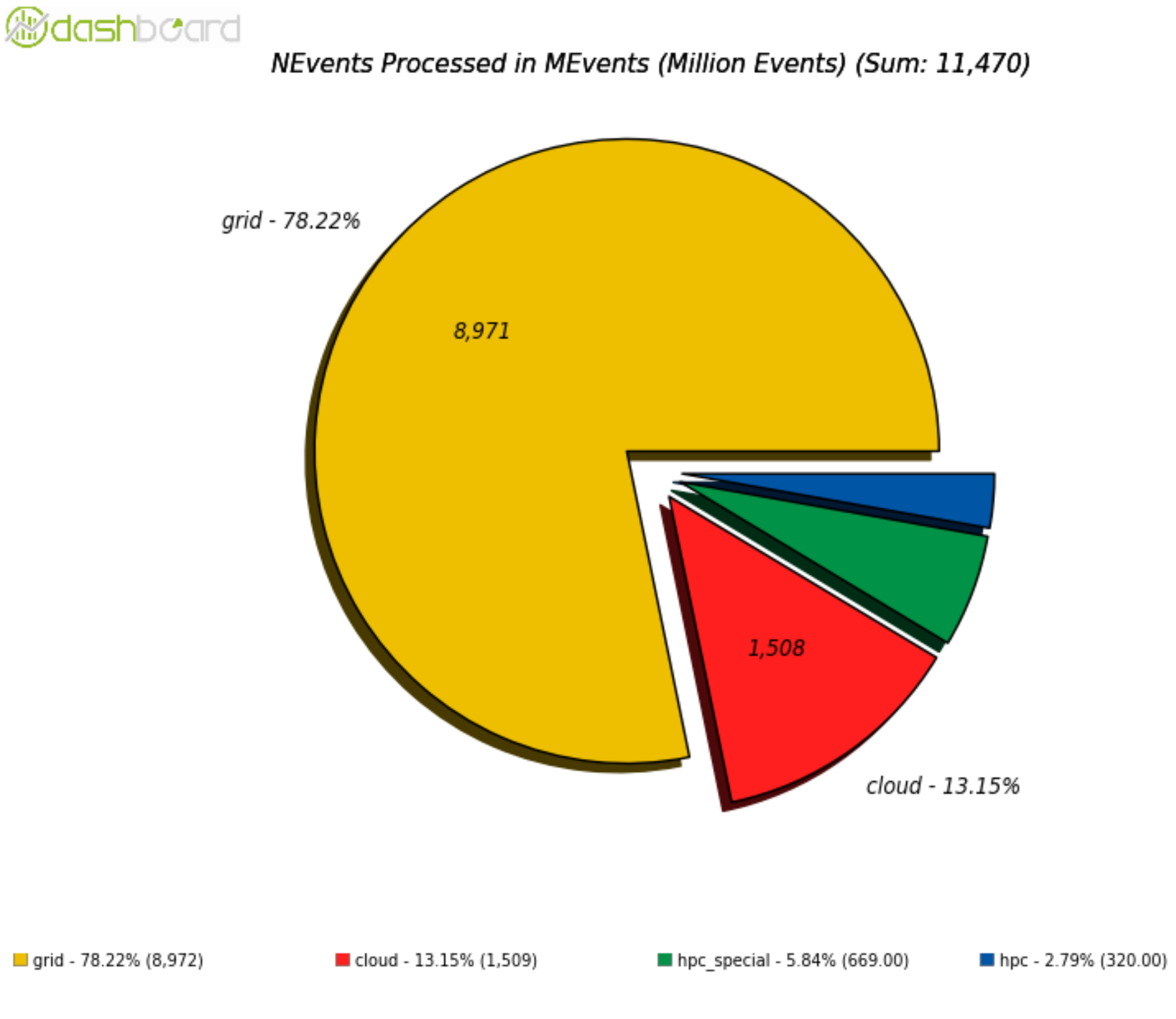
Looking forward

- Evolution of the facility : Please look at the document
- Integration of HPCs
- Distributed storage
- Caching & optimization of network usage

Spares

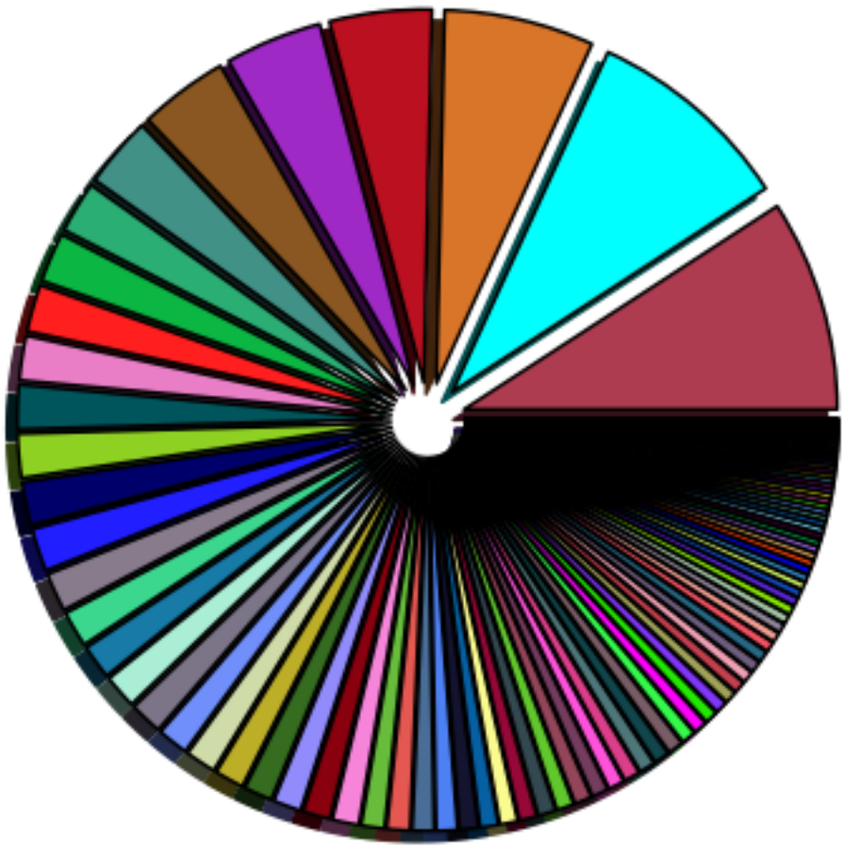
MCfull — all ATLAS sites




dashboard
NEvents Processed in MEvents (Million Events) (Sum: 11,470)
grid - 78.22%
8,971
1,508
cloud - 13.15%
grid - 78.22% (8,972)
cloud - 13.15% (1,509)
hpc_special - 5.84% (669.00)
hpc - 2.79% (320.00)

dashboard

NEvents Processed in MEvents (Million Events) (Sum: 11,470)

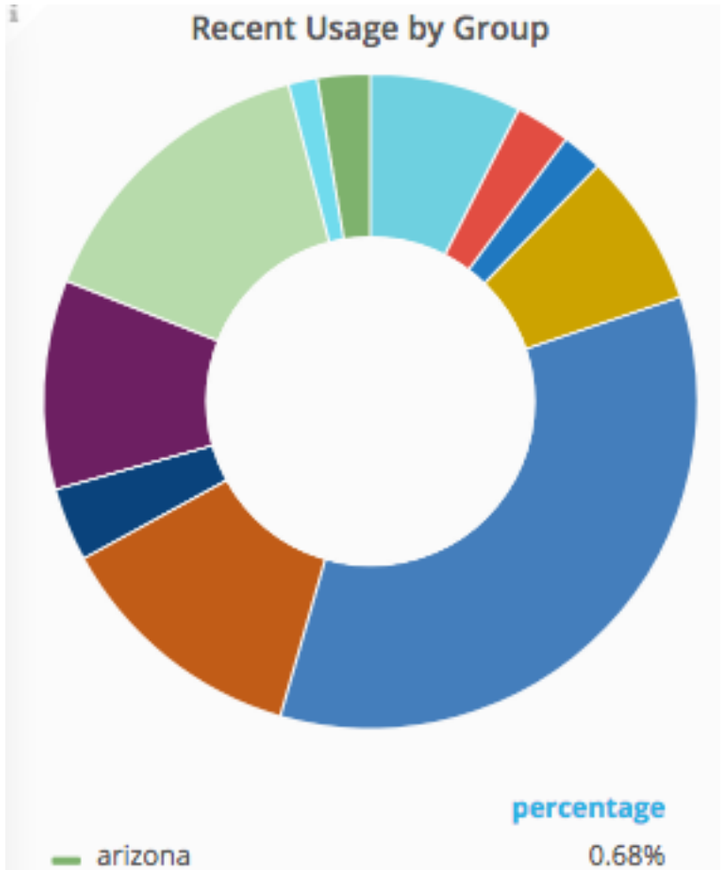


- CERN-P1 - 9.16% (1,051)
- CERN-PROD - 9.00% (1,033)
- BNL-ATLAS - 6.49% (744.00)
- IN2P3-CC - 4.42% (506.00)
- MWT2 - 4.20% (482.00)
- LBNL_DSD_ITB - 3.86% (443.00)
- TOKYO-LCG2 - 3.08% (354.00)
- SWT2_CPB - 2.24% (257.00)
- LRZ-LMU - 2.14% (246.00)
- AGLT2 - 2.03% (233.00)
- OLCF - 1.92% (220.00)
- IAAS - 1.91% (219.00)
- DESY-HH - 1.89% (217.00)
- BU_ATLAS_TIER2 - 1.86% (213.00)
- RAL-LCG2-ECHO - 1.85% (212.00)
- FZK-LCG2 - 1.67% (191.00)

Analysis facility at BNL

- Interactive nodes + batch farm (decommissioned Tier-1 equipment)
- ~100 users from a dozen of institutions

Groups past 12 months



Batch farm