Computing challenges and R&D to bridge the resource/need gap in the next 10-15 years

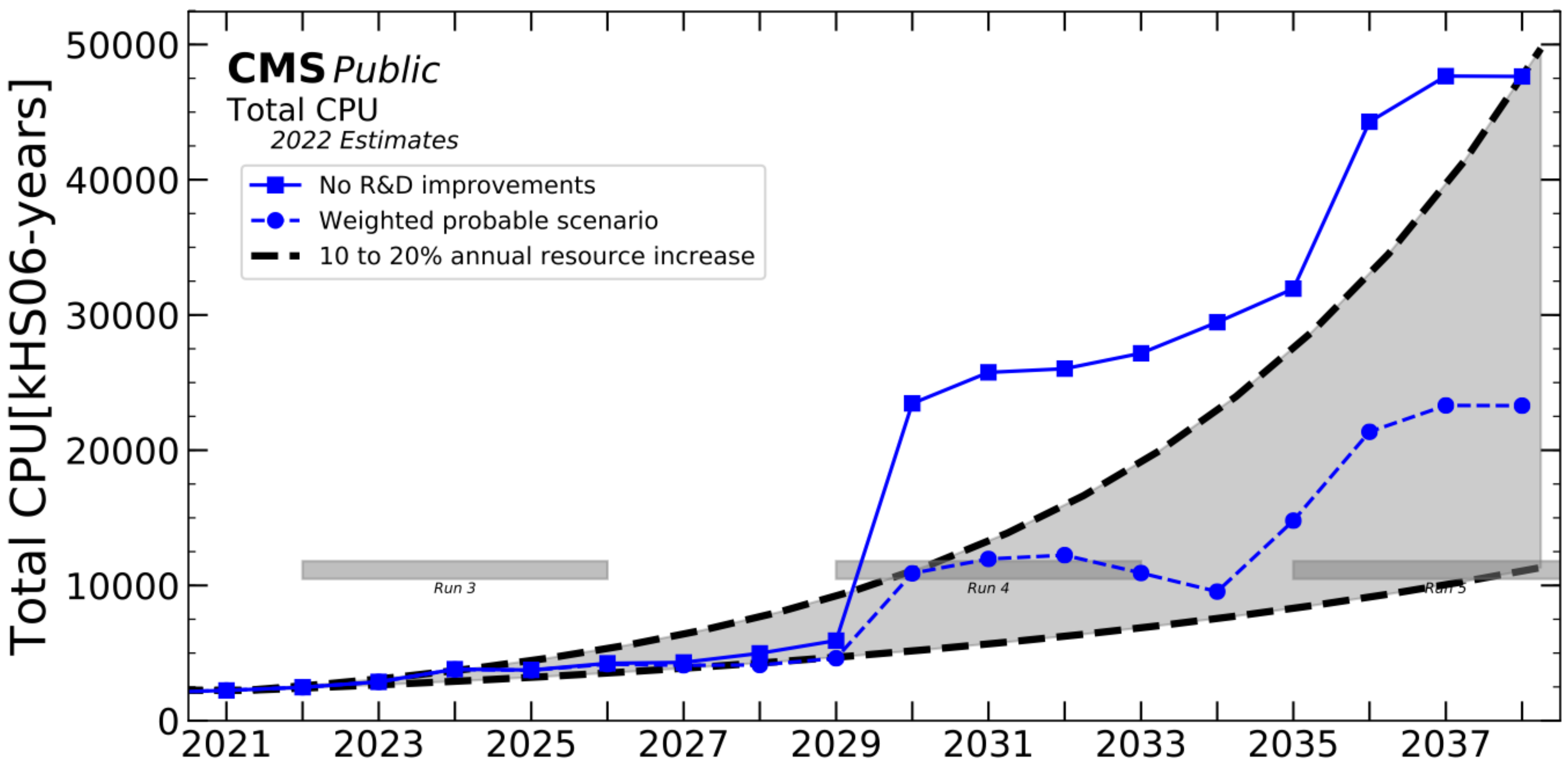
Ken Bloom
University of Nebraska-Lincoln

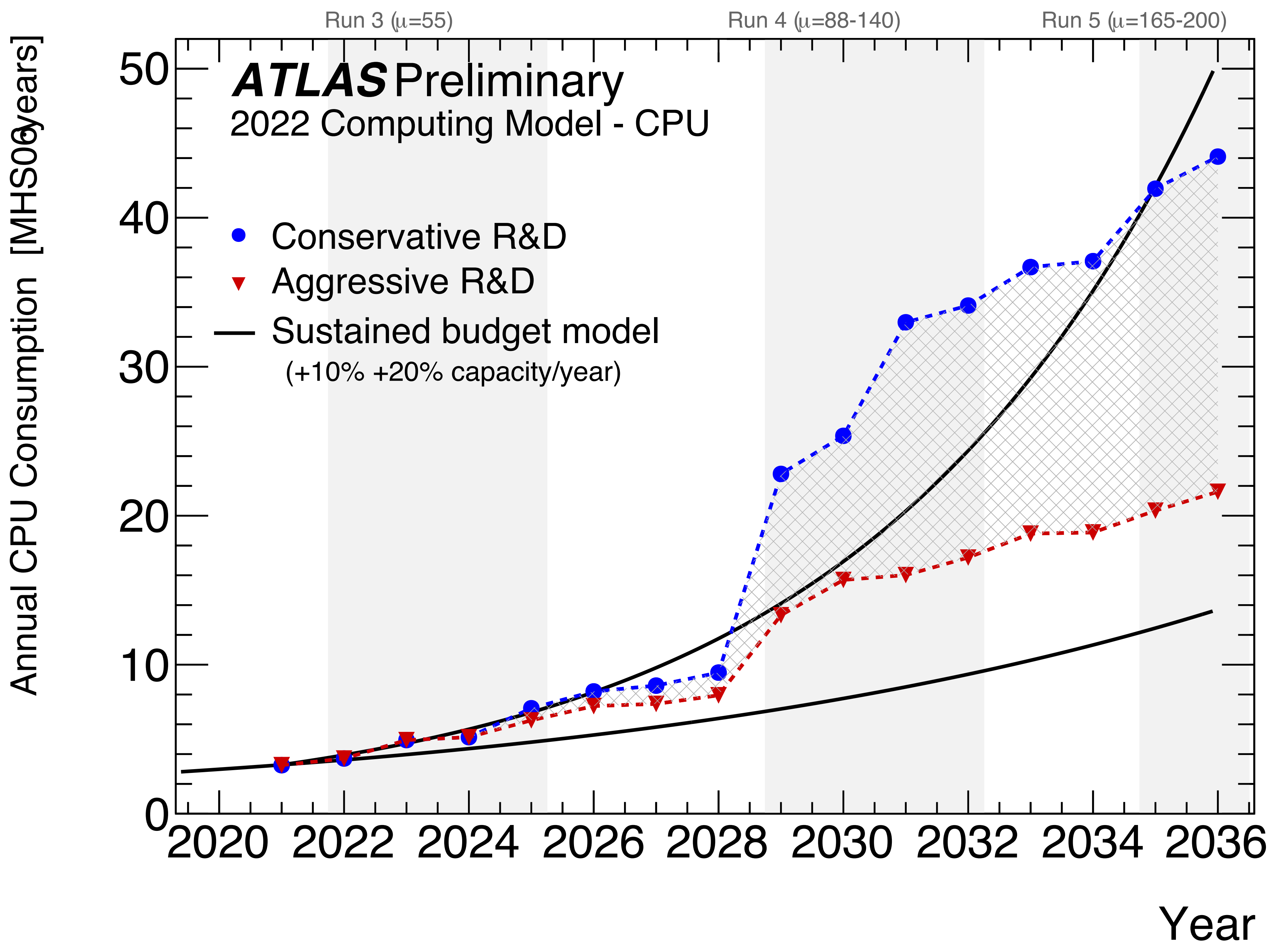
Snowmass Colloquium on the Future of Computing for HEP
23 July 2022

HL-LHC startup in 2029?

- HL-LHC will bring higher trigger rates, larger event sizes, longer processing times → greater demands on computing.
- What if this celebration is short-lived?

The HL-LHC resource gap

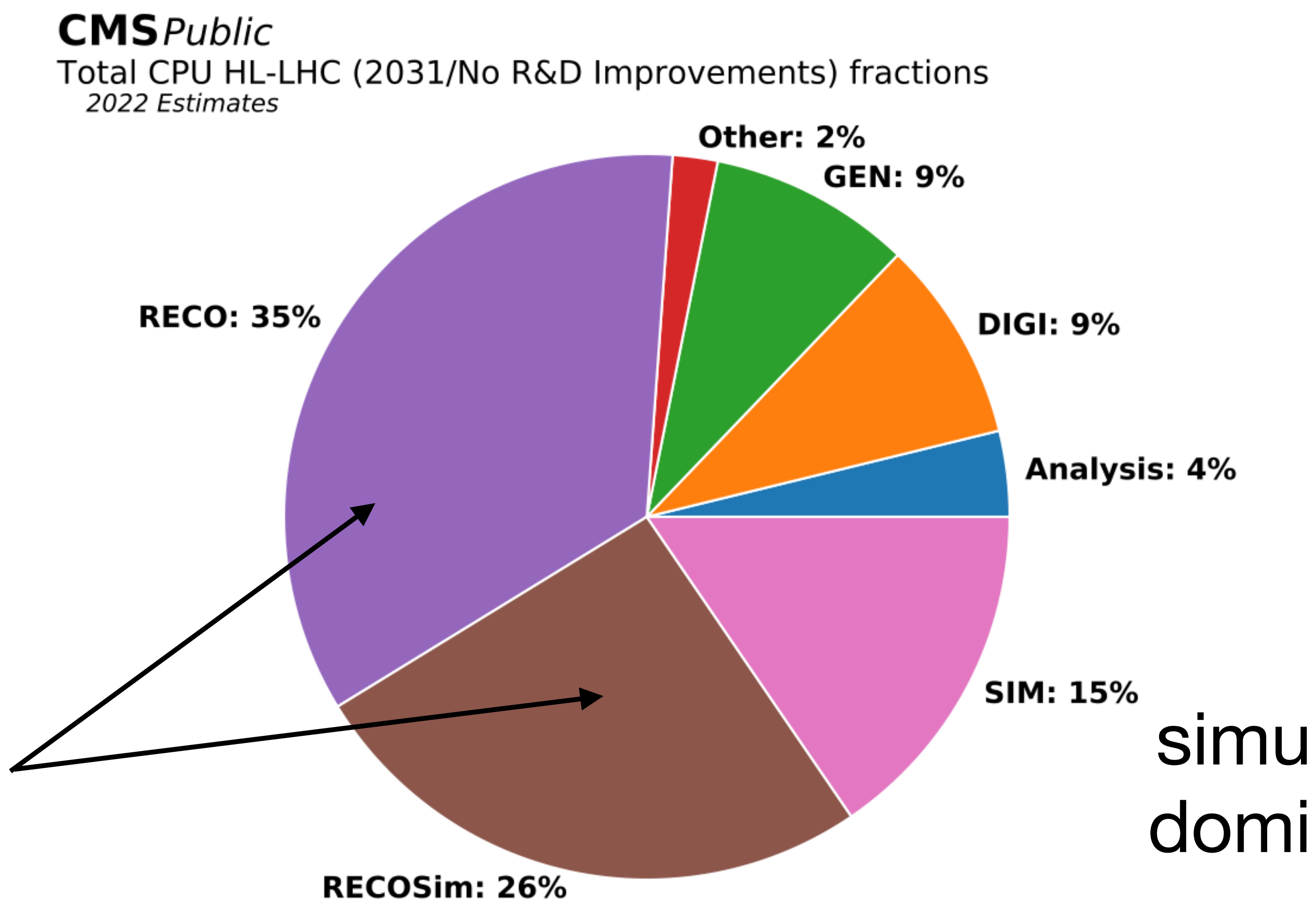




- This is processing, similar concerns for disk and tape resources.
- A potential gap between the available computing resources and experiment needs is a tangible threat to the HL-LHC physics program.
- But this risk can be reduced through sustained R&D that leverages technical advances in software and computing.

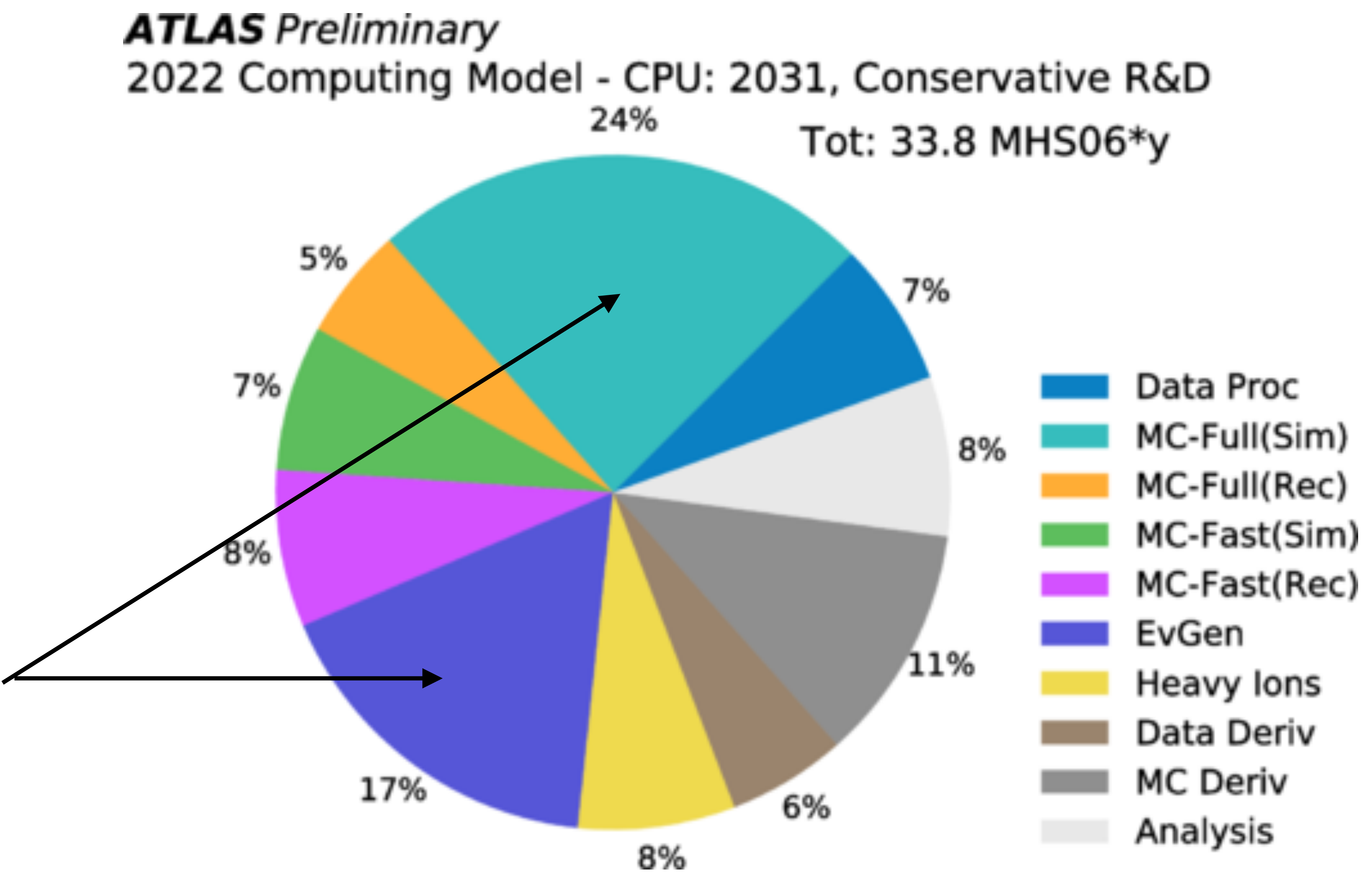
What drives the needs?

CPU

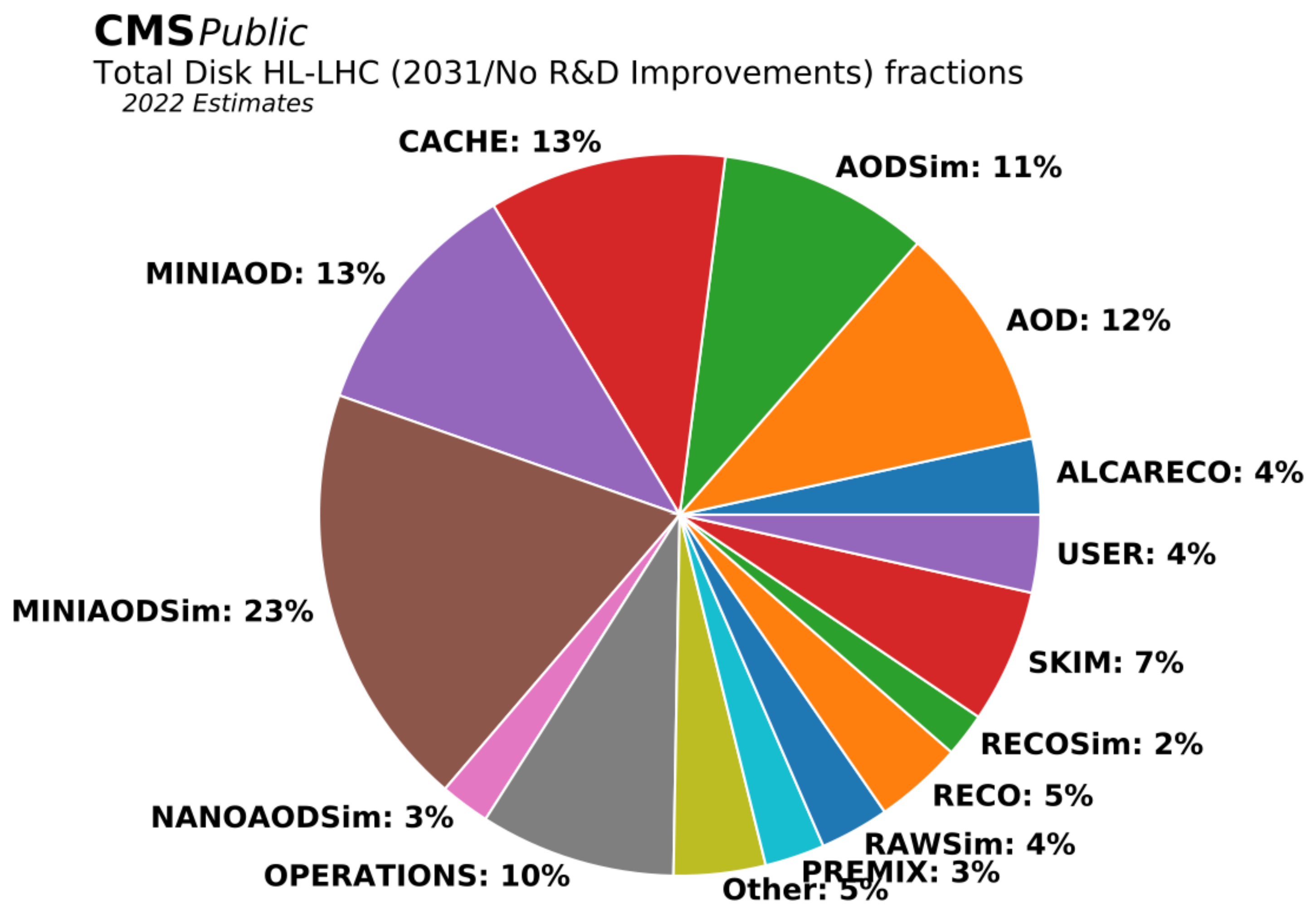

CMS *Public*
Total CPU HL-LHC (2031/No R&D Improvements) fractions
2022 Estimates

reconstruction dominated

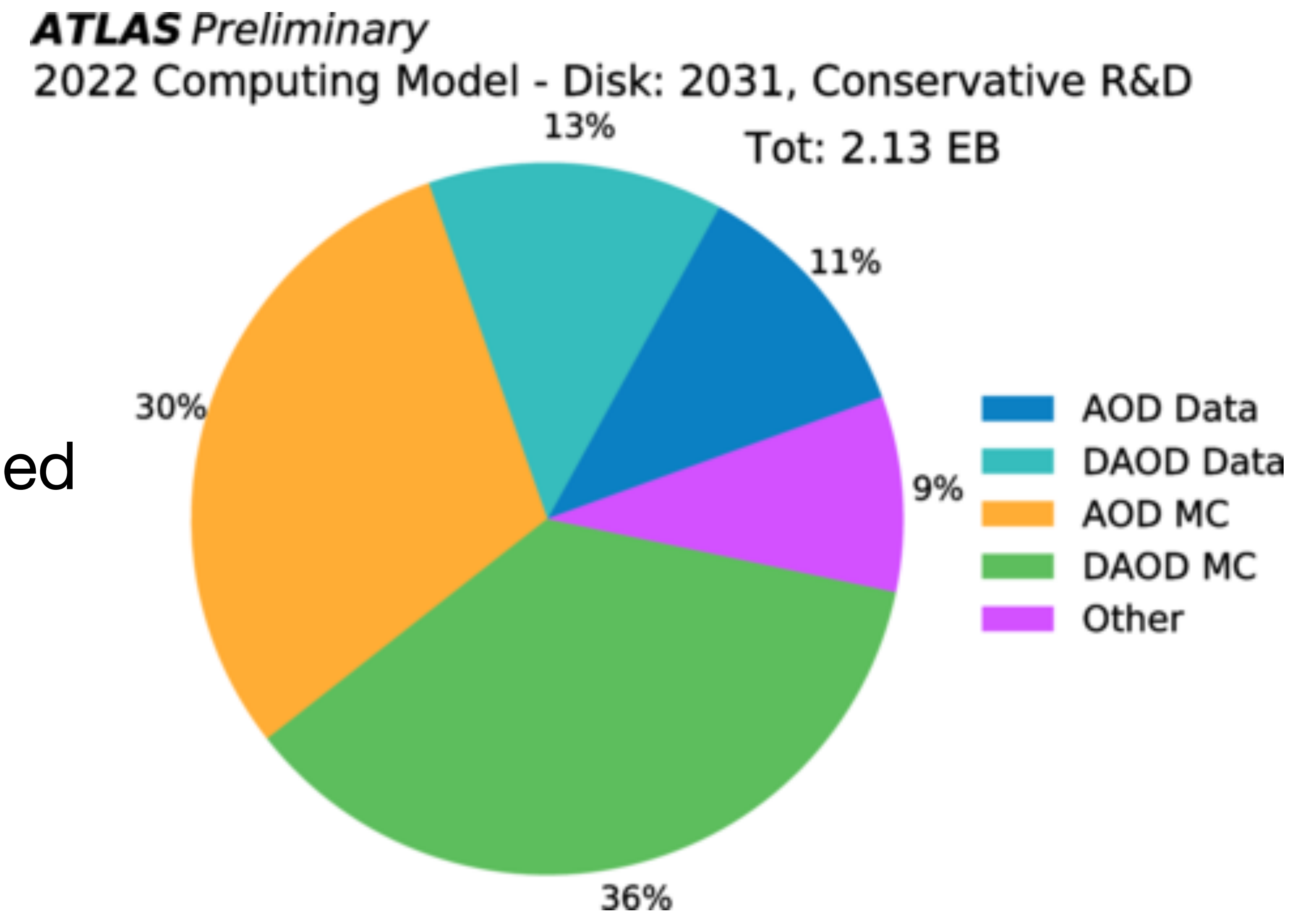
simulation dominated


ATLAS *Preliminary*
2022 Computing Model - CPU: 2031, Conservative R&D
Tot: 33.8 MHS06*y

Disk

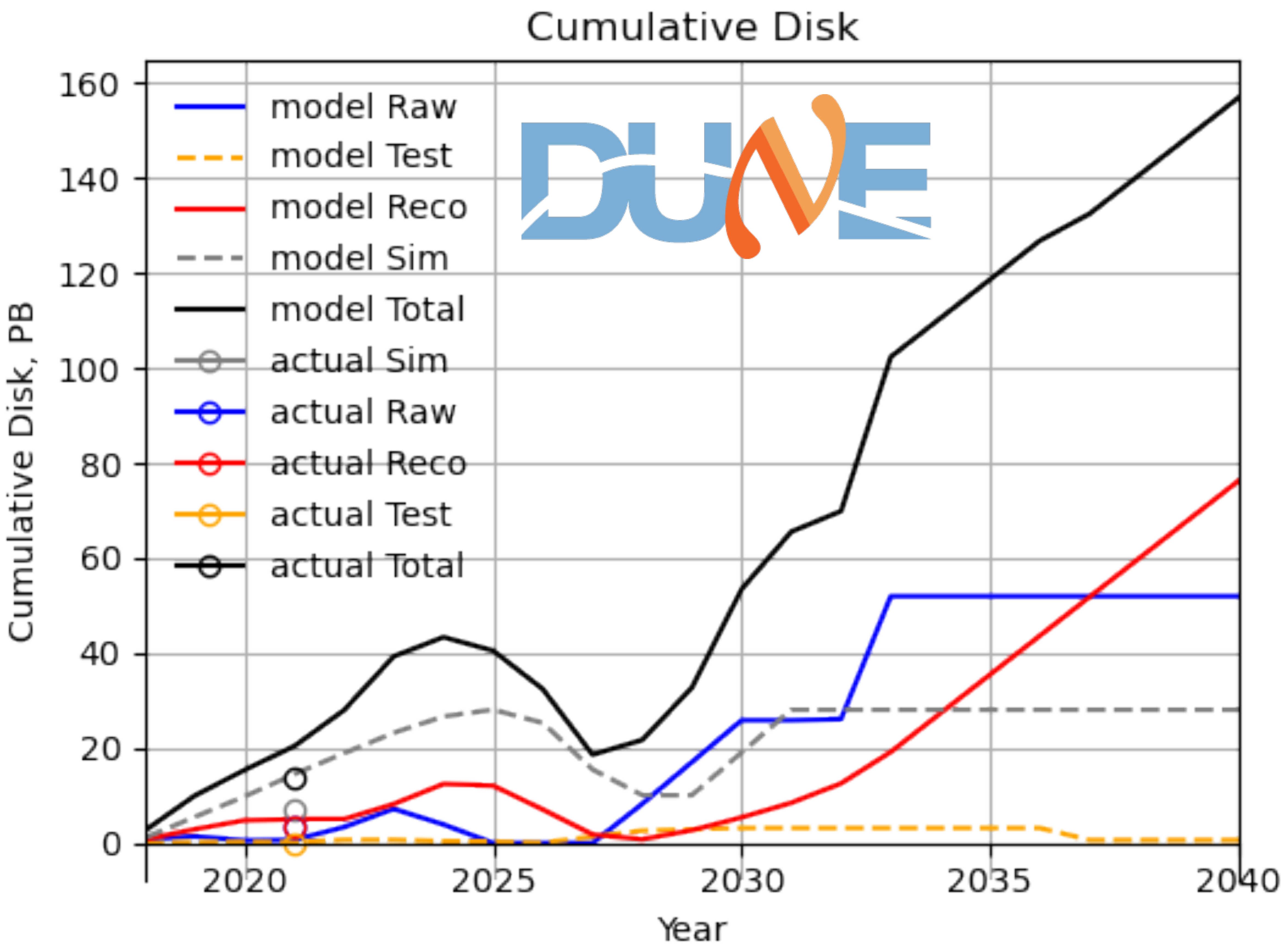

CMS *Public*
Total Disk HL-LHC (2031/No R&D Improvements) fractions
2022 Estimates

complicated story


ATLAS *Preliminary*
2022 Computing Model - Disk: 2031, Conservative R&D
Tot: 2.13 EB

Other projects' resource gaps?

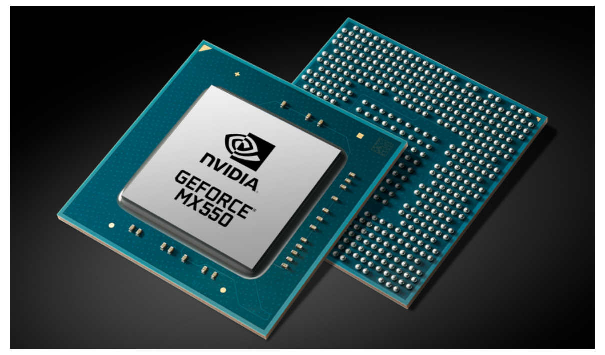
- Early estimates of DUNE computing resource needs becoming available.
  - In general, an order of magnitude smaller than HL-LHC experiment needs.
  - Specific concern: Treatment of large events from supernovae.
- DUNE has yet to baseline the computing budget, cannot yet know if there is a similar resource gap.
- Lattice QCD: need resources 10x (!) faster than planned exascale machines.
- Small experiments ≠ small computing problems.
  - As we will see, some resource gaps don't scale with project size.



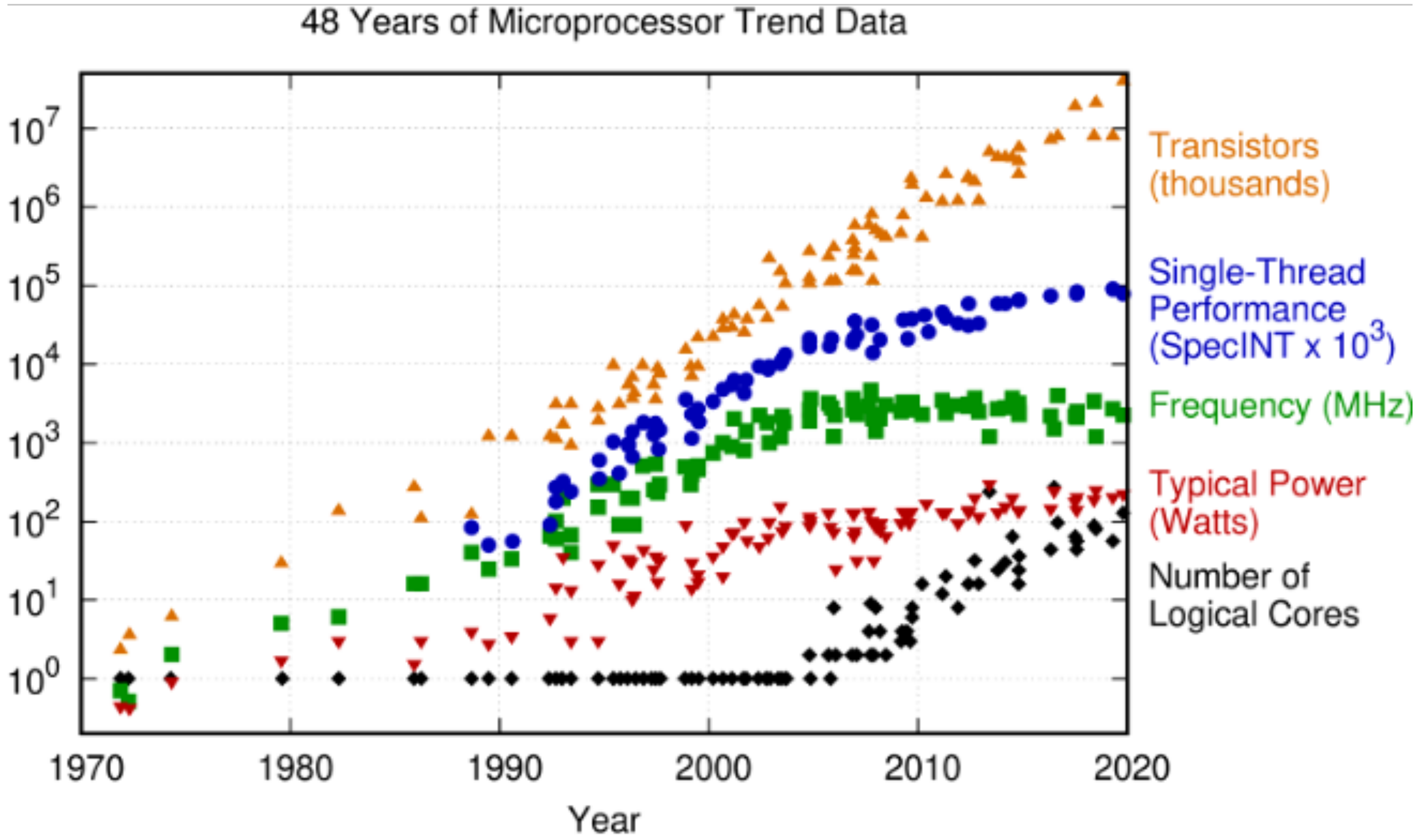
Technology evolution: heterogeneity

- Reaching the end of ~two decades of stability and uniformity in processor architecture (x86).
    - “Accelerators”: GPUs, tensor processors, FPGAs
    - Can be more power efficient than CPUs, many many small processing cores, becoming more common at computing centers.
    - Requires specialized programming techniques, re-casting of existing software, expert support.
- Reaching the end of an era of grid computing in which all computing centers looked alike.
    - Greater diversity of processing and storage systems.
    - “High performance computing” (HPC) centers are funded and operated from outside the HEP program and provide significant “free” resources, with further growth expected.
    - But each HPC center is configured differently, may have specific participation rules, and requires expert support for usage.

intel inside +

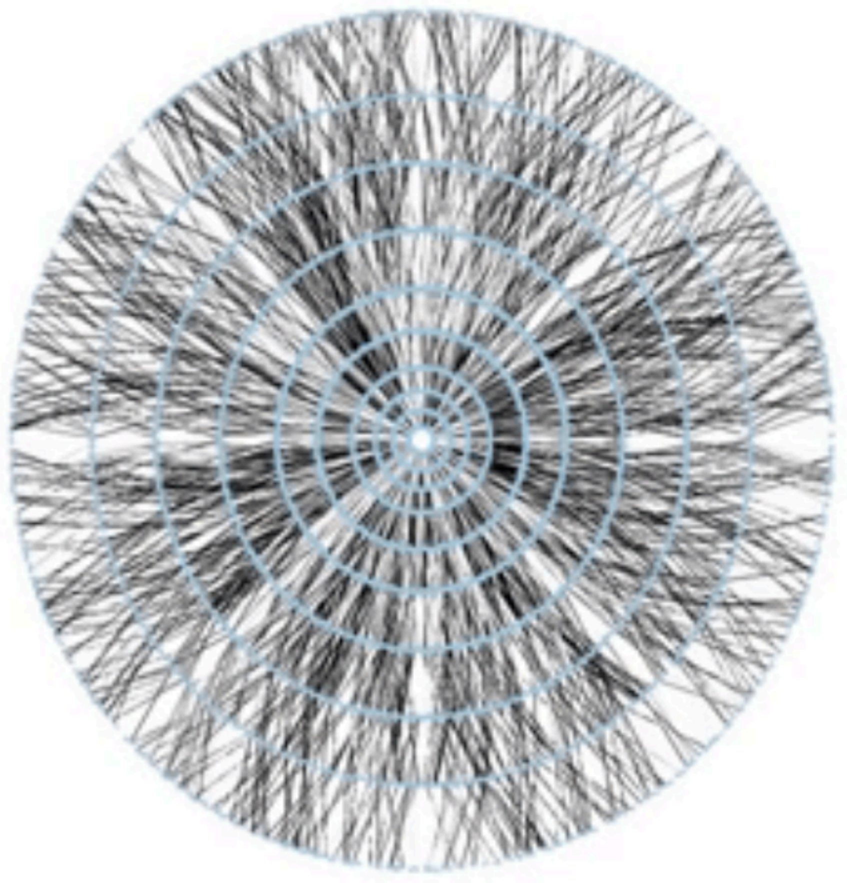


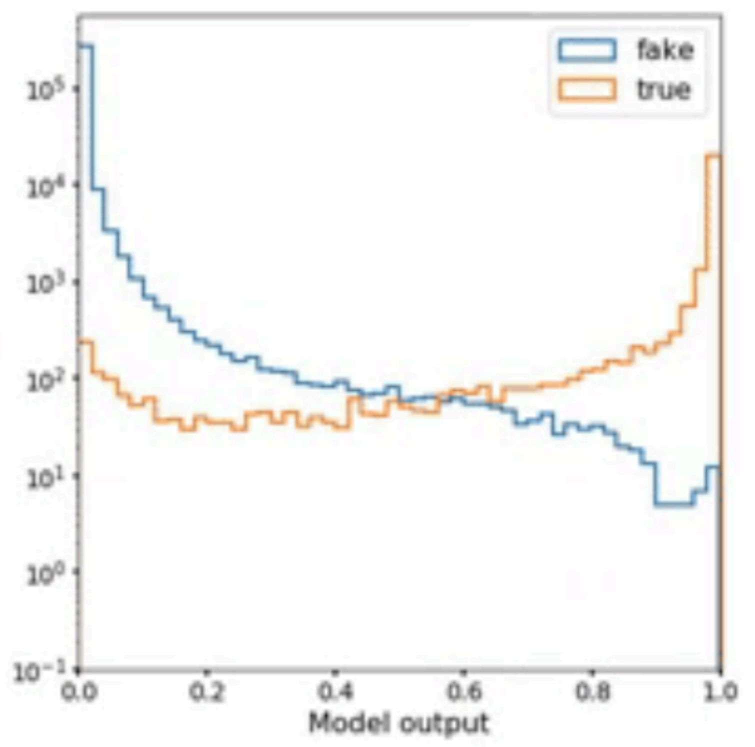
+ ....

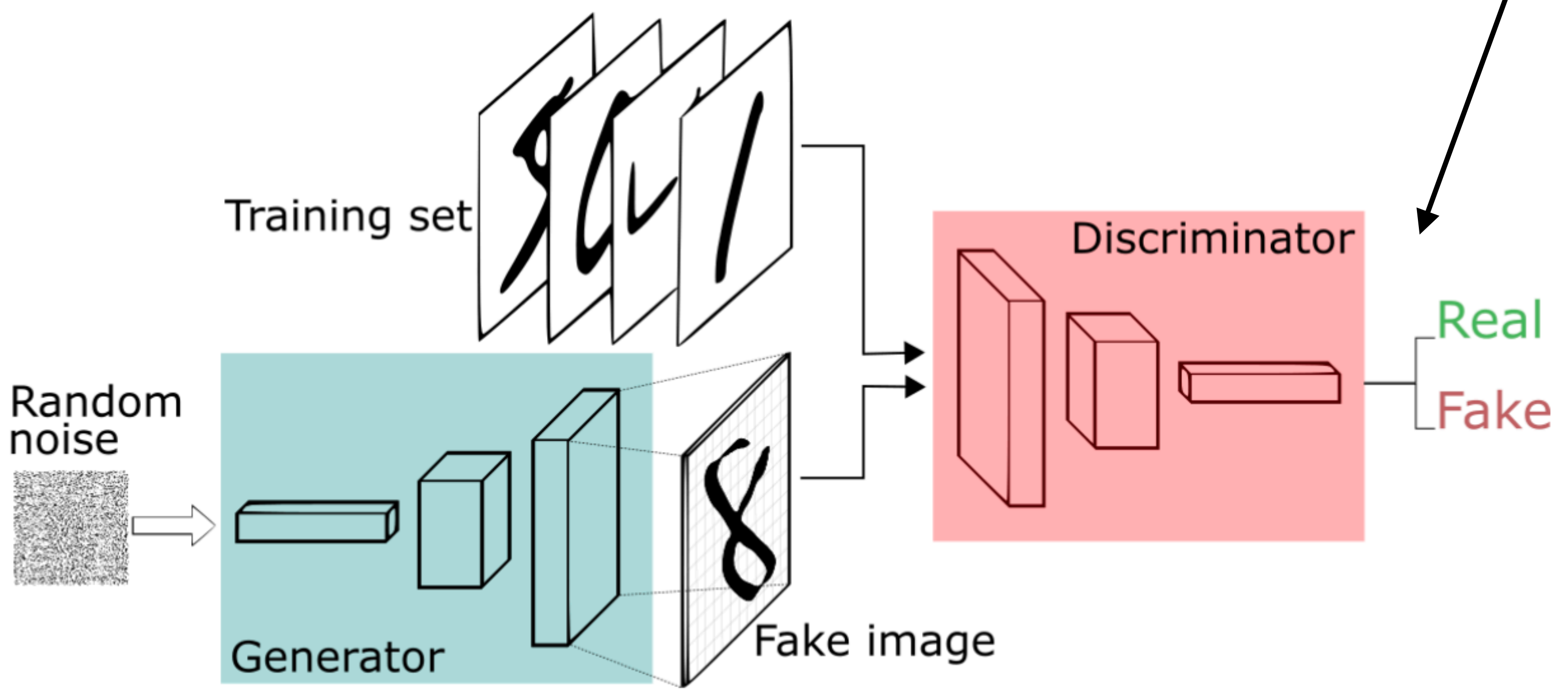

48 Years of Microprocessor Trend Data

Original data up to the year 2010 collected and plotted by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond, and C. Batten
New plot and data collected for 2010-2019 by K. Rupp

Technology evolution: machine learning



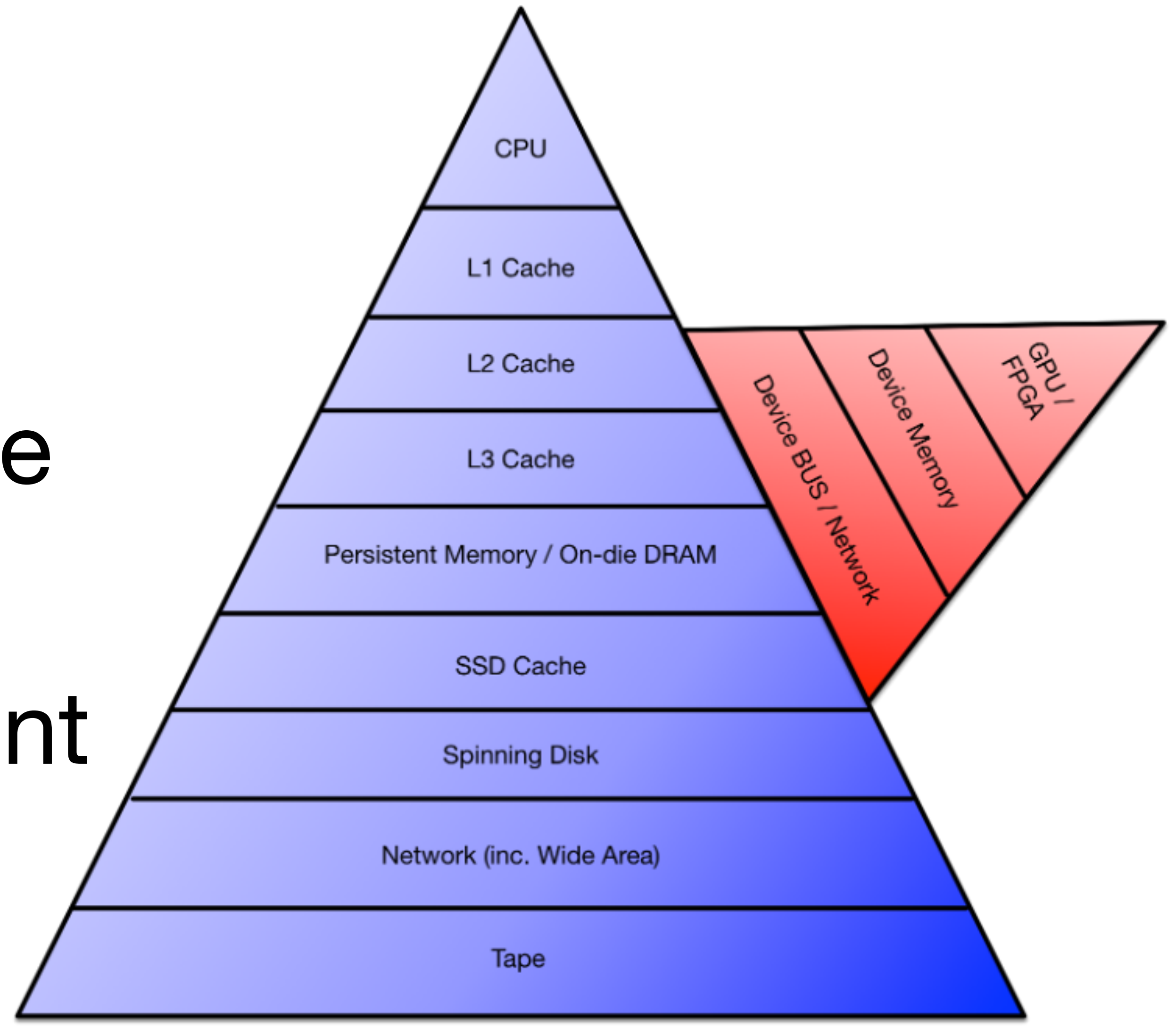




- Insightful plenary talks about ML on Monday.
- HEP has made use of multivariate techniques for decades, but recent advances in machine learning have made these approaches even more powerful and provide new opportunities.
- ML tools are particularly well matched with accelerator processors.
- Funding agencies are investing in ML and physicists are excited about using these new tools.
- Expecting that ML will play a growing role across all reconstruction and (fast) simulation algorithms; planning assumes incremental reductions in resource needs.

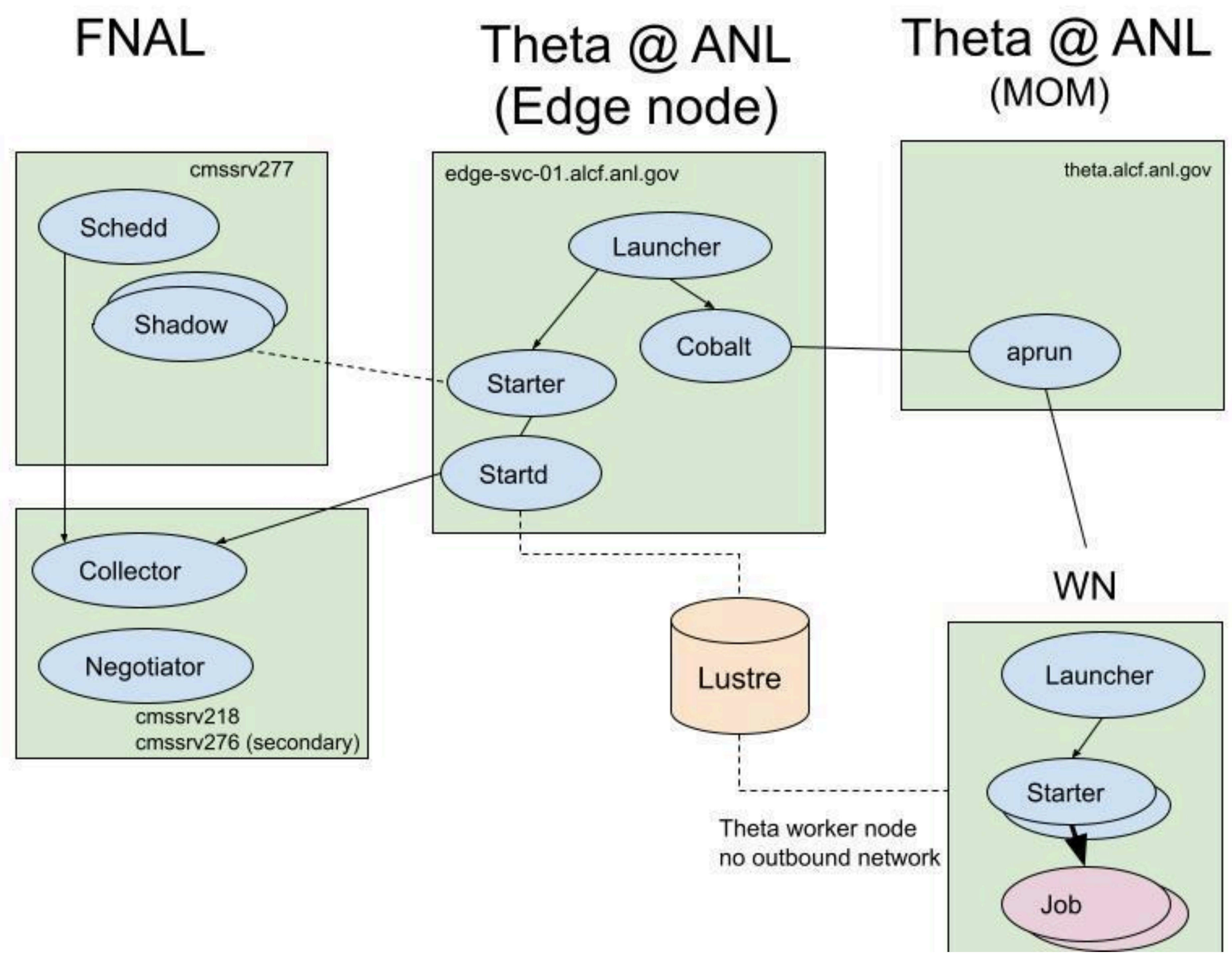
Using accelerators

- Accelerators have great promise but need to be used efficiently to maximize event throughput. Many considerations:
  - Not all required algorithms are accelerator-friendly → need to balance between accelerator and CPU.
  - Expensive to reformat data and transfer it to accelerator, most efficient to do so for many sets of data → need to parallel process different events.
  - Don’t want to leave the CPU idle while accelerator is working → need to parallel process different parts of the event.
- Addressing this requires R&D into adopting *multi-threaded processing frameworks* that can optimally manage the distribution of work across processors.
- And: every accelerator is different, requiring different programming approaches! Don’t know what hardware you might get until run time.
- *Portability libraries* are key to reducing effort needed to enable heterogeneous processors.



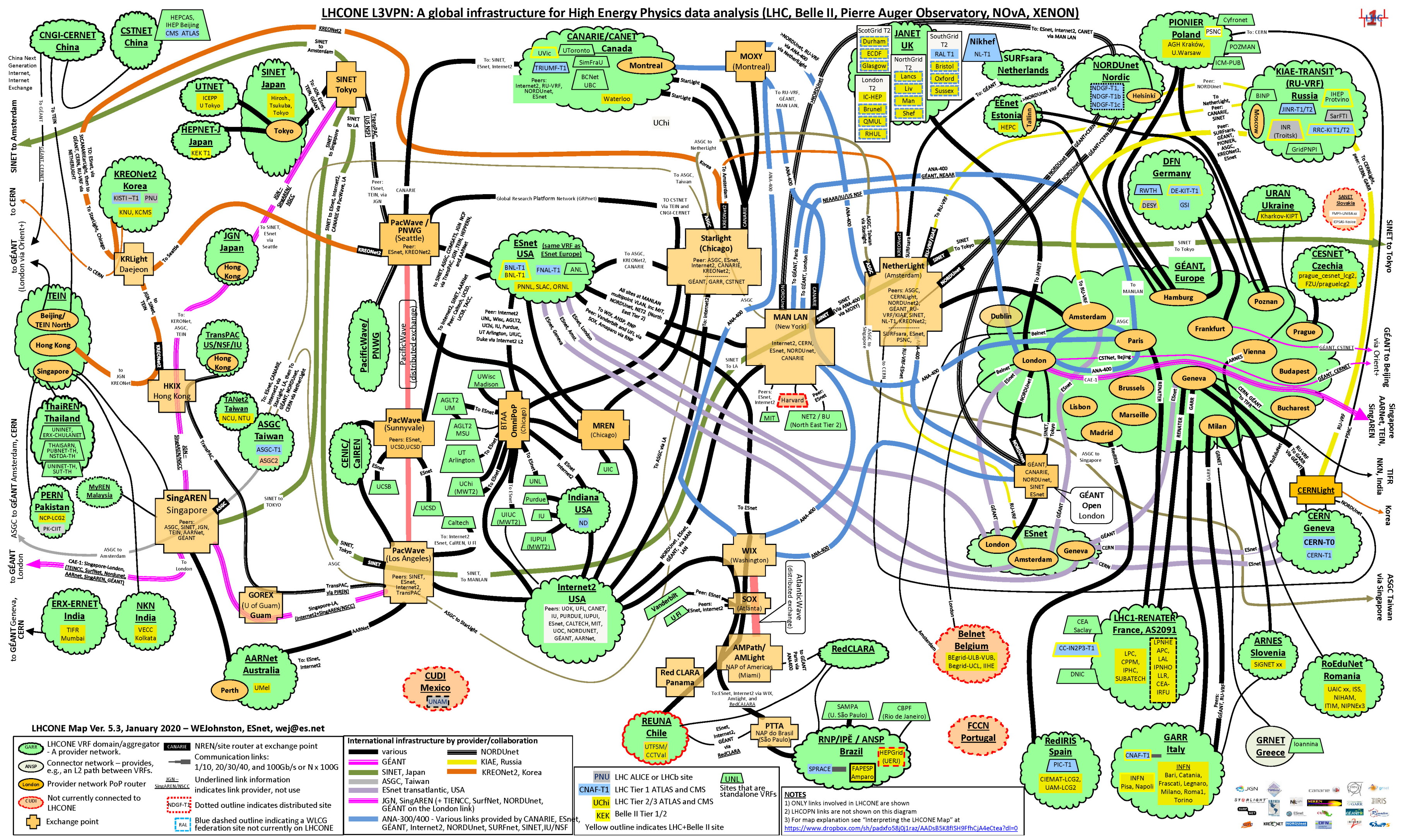
Using HPCs

- Emerging exascale high-performance computing centers are new resources that can help close the gap.
- Such centers are already in use by the LHC experiments, but there are technical/sociological challenges in using them maximally efficiently:
    - Job submission infrastructure needs to scale up to high levels, ≥1M simultaneous jobs, or be restructured.
    - Need system flexibility and operational ease to handle resources that are not uniform in configuration or availability.
        - HEP has limited (at best) input on HPC system design.
        - Some HPC allocations might be for a very specific period of time, or very specific scientific purposes.
        - Different HPCs might require different models for integration with an experiment's workflow management systems.
        - Some HPCs have much more limited network connectivity than a typical grid computing site → tunnels, edge services.
        - Great variety of processor architectures → portability again.
- All issues need ongoing development and integration efforts.



Some other topics

- Wide-area network is as fundamental to distributed computing as processing and storage.
  - Significant uncertainties about future bandwidth constraints.
  - Many R&D projects around smarter use of networks, better integration into workflow systems, improved monitoring.
- Raw detector data event size is an important driver of tape storage needs → store semi-processed data with lossy compression?
- Greater use of fast storage (SSDs) for storage-less sites, data caches within the network, high-throughput analysis.
- Simulation R&D: Potentially disruptive ML-based R&D, more use of parameterization.


LHCONE L3VPN: A global infrastructure for High Energy Physics data analysis (LHC, Belle II, Pierre Auger Observatory, NOvA, XENON)
LHCONE Map Ver. 5.3, January 2020 – WEJohnston, ESnet, wej@es.net

Evolution > revolution

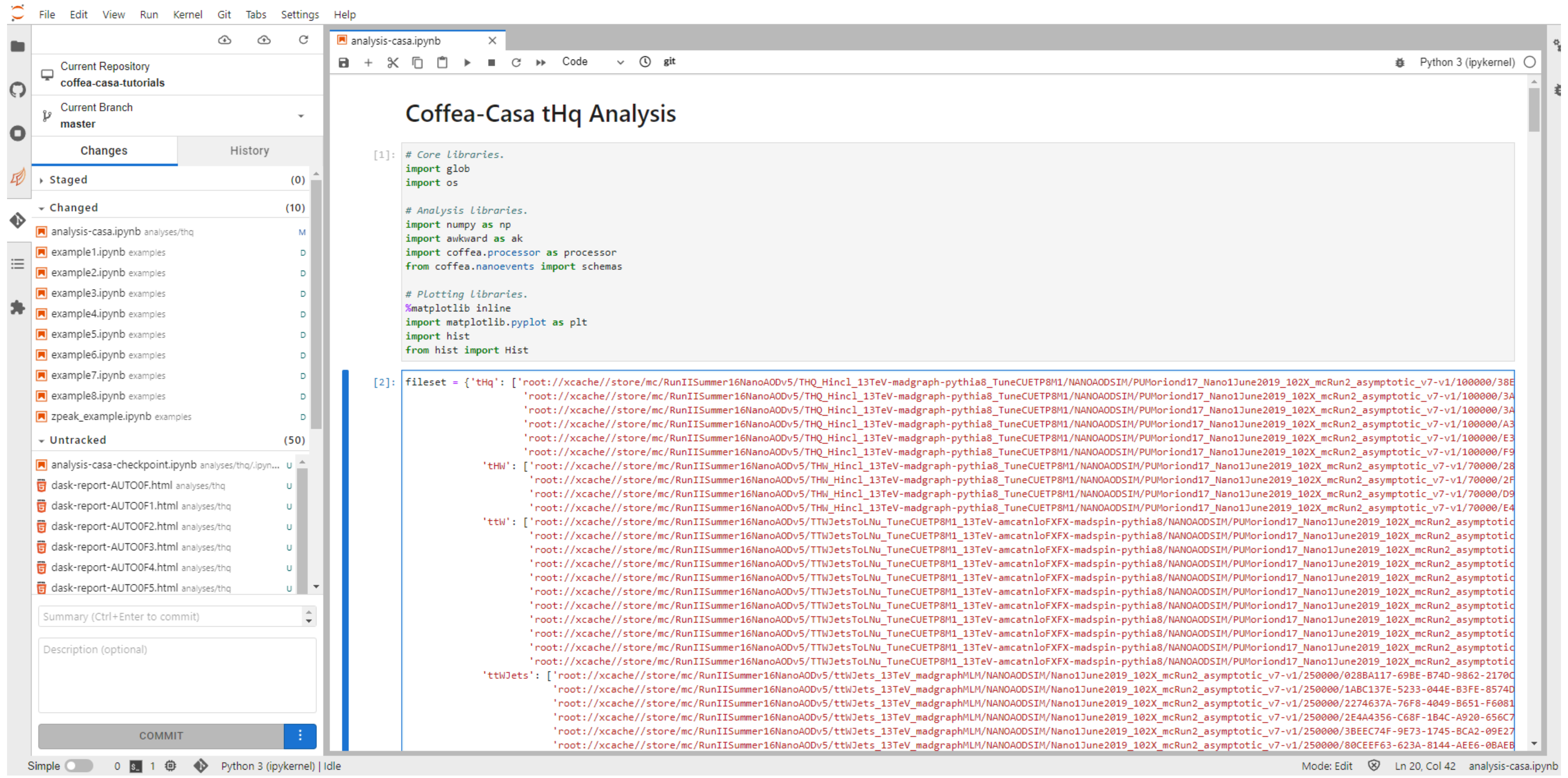
- Exploiting new technology is important, but significant reduction of the resource/needs gap will come through continuing evolution of existing software, e.g.:
  - Steady effort to improve code performance.
  - Develop more compact data formats and motivate physicists to use them.
  - Optimize detector-specific simulation and improve Geant4 performance.
  - Improve parameterized simulation and promote its use.
  - Optimize cuts/parameters for track reconstruction.
  - Implement production versions of R&D prototypes.
- These approaches have yielded significant improvements in the past.
- All of this requires a sustained level of effort on “maintenance and operations”-like tasks.

Effort/expertise gap

- Continual improvement required to advance software and computing systems requires significant effort, even before considering additional effort for R&D that could have an impact on reducing future computing needs.
  - → We must invest in both.
- Currently a limited number of people with the requisite expertise.
  - → Need to invest here too, in education and training.
  - Many ongoing efforts, need to get people engaged.
  - Might be an even more critical issue for smaller experiments.
- Software and computing lacks diversity, even more so than other areas of our field.
- We also need to provide career paths for software/computing professionals, much like we do now for engineers.

Usability gap

- Physicists depend on fast, easy, user-friendly access to data analysis resources, *at scale,* without the barriers that can be posed by complex software and computing tools.
  - Let physicists be physicists!
  - Number of physicists needed to analyze the data can't scale with the amount of data.
- Developing software ecosystems and analysis facilities that provide an easy path to data analysis.


File
Edit
View
Run
Kernel
Git
Tabs
Settings
Help
Current Repository
coffea-casa-tutorials
Current Branch
master
Changes
History
Staged (0)
Changed (10)
analysis-casa.ipynb analyses/thq M
example1.ipynb examples D
example2.ipynb examples D
example3.ipynb examples D
example4.ipynb examples D
example5.ipynb examples D
example6.ipynb examples D
example7.ipynb examples D
example8.ipynb examples D
zpeak_example.ipynb examples D
Untracked (50)
analysis-casa-checkpoint.ipynb analyses/thq/.ipyn... U
dask-report-AUTO0F.html analyses/thq U
dask-report-AUTO0F1.html analyses/thq U
dask-report-AUTO0F2.html analyses/thq U
dask-report-AUTO0F3.html analyses/thq U
dask-report-AUTO0F4.html analyses/thq U
dask-report-AUTO0F5.html analyses/thq U
Summary (Ctrl+Enter to commit)
Description (optional)
COMMIT
Simple
0
1
Python 3 (ipykernel) | Idle
analysis-casa.ipynb
Code
git
Python 3 (ipykernel)
Mode: Edit
Ln 20, Col 42
analysis-casa.ipynb

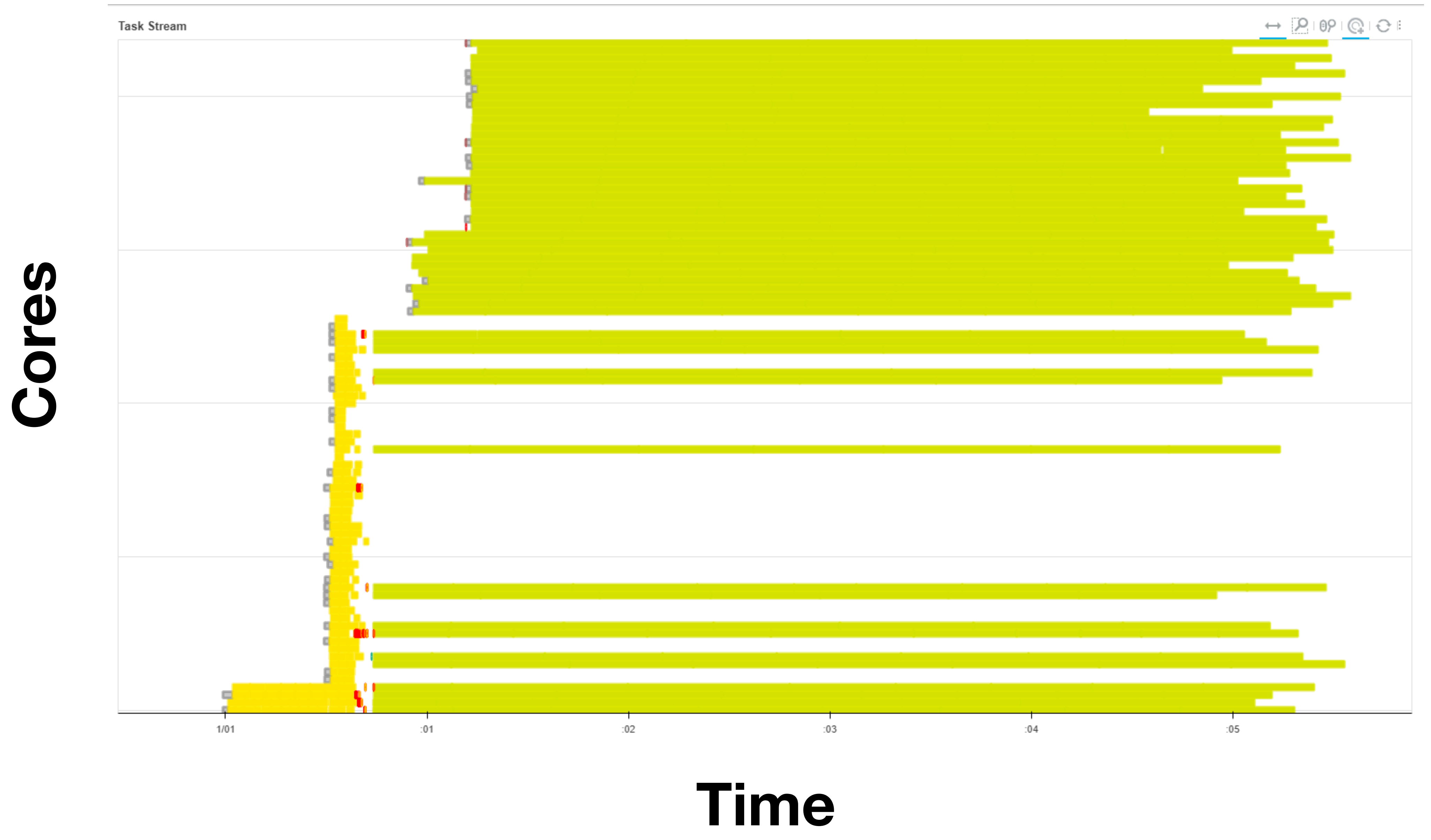
Coffea-Casa tHq Analysis

```
[1]: # Core libraries.
     import glob
     import os

     # Analysis libraries.
     import numpy as np
     import awkward as ak
     import coffea.processor as processor
     from coffea.nanoevents import schemas

     # Plotting libraries.
     %matplotlib inline
     import matplotlib.pyplot as plt
     import hist
     from hist import Hist

[2]: fileset = {'tHq': ['root://xcache//store/mc/RunIISummer16NanoAODv5/THQ_Hincl_13TeV-madgraph-pythia8_TuneCUETP8M1/NANOAODSIM/PUMoriond17_Nano1June2019_102X_mcRun2_asymptotic_v7-v1/100000/38E
                        'root://xcache//store/mc/RunIISummer16NanoAODv5/THQ_Hincl_13TeV-madgraph-pythia8_TuneCUETP8M1/NANOAODSIM/PUMoriond17_Nano1June2019_102X_mcRun2_asymptotic_v7-v1/100000/3A
                        'root://xcache//store/mc/RunIISummer16NanoAODv5/THQ_Hincl_13TeV-madgraph-pythia8_TuneCUETP8M1/NANOAODSIM/PUMoriond17_Nano1June2019_102X_mcRun2_asymptotic_v7-v1/100000/3A
                        'root://xcache//store/mc/RunIISummer16NanoAODv5/THQ_Hincl_13TeV-madgraph-pythia8_TuneCUETP8M1/NANOAODSIM/PUMoriond17_Nano1June2019_102X_mcRun2_asymptotic_v7-v1/100000/A3
                        'root://xcache//store/mc/RunIISummer16NanoAODv5/THQ_Hincl_13TeV-madgraph-pythia8_TuneCUETP8M1/NANOAODSIM/PUMoriond17_Nano1June2019_102X_mcRun2_asymptotic_v7-v1/100000/E3
                        'root://xcache//store/mc/RunIISummer16NanoAODv5/THQ_Hincl_13TeV-madgraph-pythia8_TuneCUETP8M1/NANOAODSIM/PUMoriond17_Nano1June2019_102X_mcRun2_asymptotic_v7-v1/100000/F9
                 'tHW': ['root://xcache//store/mc/RunIISummer16NanoAODv5/THW_Hincl_13TeV-madgraph-pythia8_TuneCUETP8M1/NANOAODSIM/PUMoriond17_Nano1June2019_102X_mcRun2_asymptotic_v7-v1/70000/28
                         'root://xcache//store/mc/RunIISummer16NanoAODv5/THW_Hincl_13TeV-madgraph-pythia8_TuneCUETP8M1/NANOAODSIM/PUMoriond17_Nano1June2019_102X_mcRun2_asymptotic_v7-v1/70000/2F
                         'root://xcache//store/mc/RunIISummer16NanoAODv5/THW_Hincl_13TeV-madgraph-pythia8_TuneCUETP8M1/NANOAODSIM/PUMoriond17_Nano1June2019_102X_mcRun2_asymptotic_v7-v1/70000/D9
                         'root://xcache//store/mc/RunIISummer16NanoAODv5/THW_Hincl_13TeV-madgraph-pythia8_TuneCUETP8M1/NANOAODSIM/PUMoriond17_Nano1June2019_102X_mcRun2_asymptotic_v7-v1/70000/E4
                 'ttW': ['root://xcache//store/mc/RunIISummer16NanoAODv5/TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8/NANOAODSIM/PUMoriond17_Nano1June2019_102X_mcRun2_asymptotic
                         'root://xcache//store/mc/RunIISummer16NanoAODv5/TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8/NANOAODSIM/PUMoriond17_Nano1June2019_102X_mcRun2_asymptotic
                         'root://xcache//store/mc/RunIISummer16NanoAODv5/TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8/NANOAODSIM/PUMoriond17_Nano1June2019_102X_mcRun2_asymptotic
                         'root://xcache//store/mc/RunIISummer16NanoAODv5/TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8/NANOAODSIM/PUMoriond17_Nano1June2019_102X_mcRun2_asymptotic
                         'root://xcache//store/mc/RunIISummer16NanoAODv5/TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8/NANOAODSIM/PUMoriond17_Nano1June2019_102X_mcRun2_asymptotic
                         'root://xcache//store/mc/RunIISummer16NanoAODv5/TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8/NANOAODSIM/PUMoriond17_Nano1June2019_102X_mcRun2_asymptotic
                         'root://xcache//store/mc/RunIISummer16NanoAODv5/TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8/NANOAODSIM/PUMoriond17_Nano1June2019_102X_mcRun2_asymptotic
                         'root://xcache//store/mc/RunIISummer16NanoAODv5/TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8/NANOAODSIM/PUMoriond17_Nano1June2019_102X_mcRun2_asymptotic
                         'root://xcache//store/mc/RunIISummer16NanoAODv5/TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8/NANOAODSIM/PUMoriond17_Nano1June2019_102X_mcRun2_asymptotic
                         'root://xcache//store/mc/RunIISummer16NanoAODv5/TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8/NANOAODSIM/PUMoriond17_Nano1June2019_102X_mcRun2_asymptotic
                         'root://xcache//store/mc/RunIISummer16NanoAODv5/TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8/NANOAODSIM/PUMoriond17_Nano1June2019_102X_mcRun2_asymptotic
                 'ttWJets': ['root://xcache//store/mc/RunIISummer16NanoAODv5/ttWJets_13TeV_madgraphMLM/NANOAODSIM/Nano1June2019_102X_mcRun2_asymptotic_v7-v1/250000/028BA117-69BE-B74D-9862-2170C
                             'root://xcache//store/mc/RunIISummer16NanoAODv5/ttWJets_13TeV_madgraphMLM/NANOAODSIM/Nano1June2019_102X_mcRun2_asymptotic_v7-v1/250000/1ABC137E-5233-044E-B3FE-8574D
                             'root://xcache//store/mc/RunIISummer16NanoAODv5/ttWJets_13TeV_madgraphMLM/NANOAODSIM/Nano1June2019_102X_mcRun2_asymptotic_v7-v1/250000/2274637A-76F8-4049-B651-F6081
                             'root://xcache//store/mc/RunIISummer16NanoAODv5/ttWJets_13TeV_madgraphMLM/NANOAODSIM/Nano1June2019_102X_mcRun2_asymptotic_v7-v1/250000/2E4A4356-C68F-1B4C-A920-656C7
                             'root://xcache//store/mc/RunIISummer16NanoAODv5/ttWJets_13TeV_madgraphMLM/NANOAODSIM/Nano1June2019_102X_mcRun2_asymptotic_v7-v1/250000/3BEEC74F-9E73-1745-BCA2-09E27
                             'root://xcache//store/mc/RunIISummer16NanoAODv5/ttWJets_13TeV_madgraphMLM/NANOAODSIM/Nano1June2019_102X_mcRun2_asymptotic_v7-v1/250000/80CEEF63-623A-8144-AEE6-0BAEB
```



Closing our gaps

- The gap between computing needs and resources is a real threat to the future particle physics program.
- This gap can be reduced and perhaps eliminated through R&D efforts already underway that take advantage of new technologies.
- Many promising pathways, but they require sustained investments of time and effort — including effort for implementation and maintenance of R&D outcomes.
- Not our only gaps: expertise, career paths, usability.
- There are opportunities for you to get involved and make a difference!

- Thanks to: Mat Adamec, Tulika Bose, Paolo Calafiura, Daniel Elvira, Steve Gottlieb, Heather Gray, Oliver Gutsche, Dirk Hufnagel, Mike Kirby, David Lange, James Letts, David Mason, Ben Nachman, Danilo Piparo, Heidi Schellman, Liz Sexton-Kennedy
- This is my first presentation in 16:9 mode!