Fermilab | U.S. DEPARTMENT OF ENERGY | Office of Science

High Energy Physics computing in the next decade

Dr. Burt Holzman, Fermi National Accelerator Laboratory

Future Online Analysis Platform Workshop

April 4, 2017

“Be Provocative”

- I am speaking about **High Energy Physics** needs
  - I’m actually an **experimental nuclear physicist** by training
- Over the next decade
  - Will we even be here in **2027**?
- Highly biased and **myopic** view of HEP

HEP Software Foundation Community White Paper:
http://hepsoftwarefoundation.org/activities/cwp.html

Fermilab

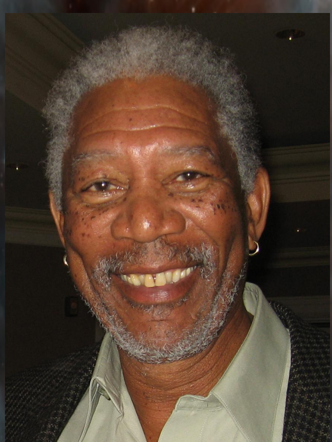
Disclaimer

Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the Fermilab, the United States Government, or any agency thereof.

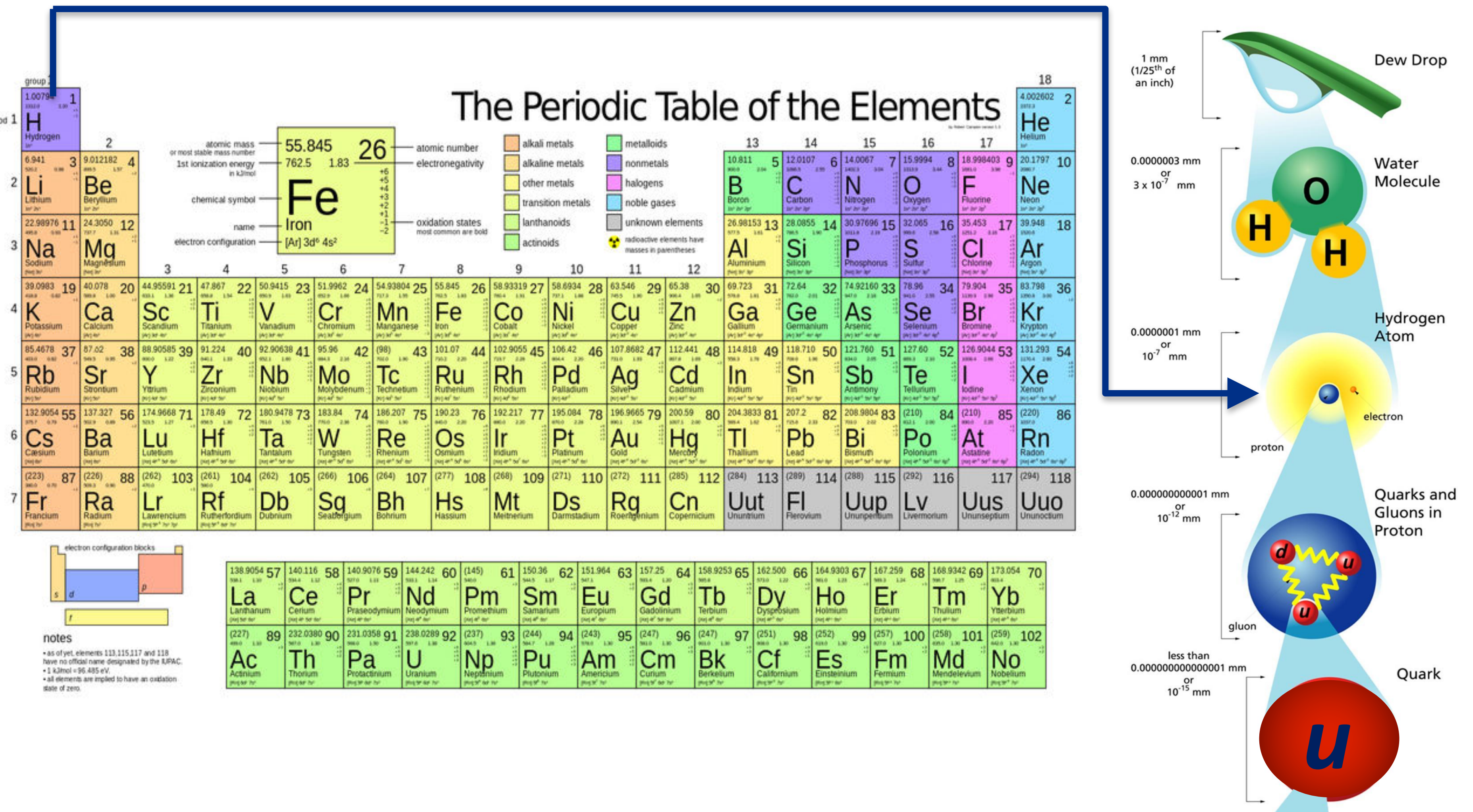
What is the Universe made of?

How did the Universe form?

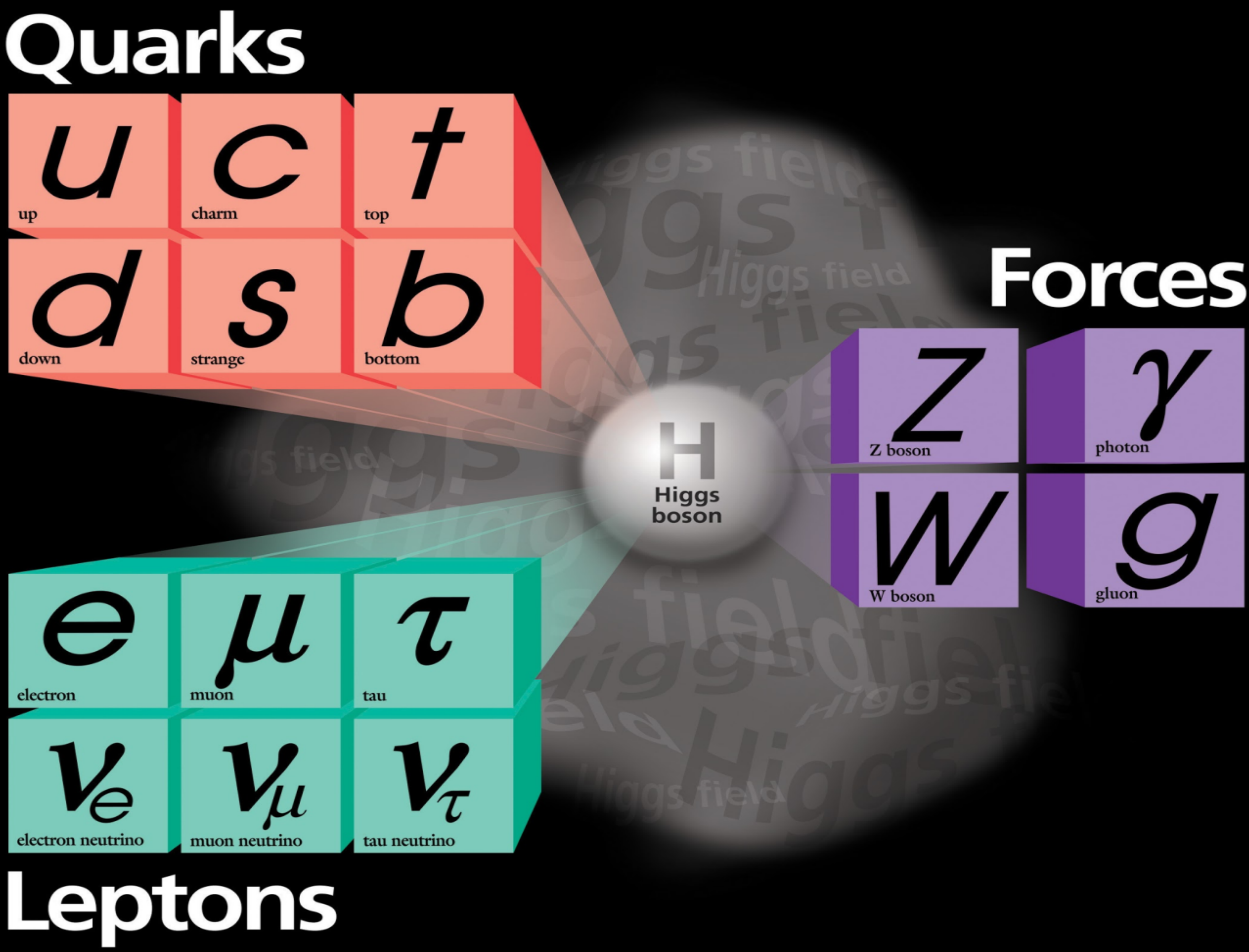
What is the future of the Universe?



"Me and Morgan Freeman" by David Sifry is licensed under CC by 2.0 (cropped)


The Periodic Table of the Elements
atomic mass
or most stable mass number
1st ionization energy in kJ/mol
chemical symbol
name
electron configuration
55.845 26
762.5 1.83
Fe
Iron
[Ar] 3d6 4s2
atomic number
electronegativity
oxidation states
most common are bold
alkali metals
alkaline metals
other metals
transition metals
lanthanoids
actinoids
metalloids
nonmetals
halogens
noble gases
unknown elements
radioactive elements have masses in parentheses
electron configuration blocks
notes
1 mm
(1/25th of an inch)
Dew Drop
0.0000003 mm
or
3 x 10-7 mm
Water Molecule
O
H
H
0.0000001 mm
or
10-7 mm
Hydrogen Atom
electron
proton
0.000000000001 mm
or
10-12 mm
Quarks and Gluons in Proton
d
u
u
gluon
less than
0.000000000000001 mm
or
10-15 mm
Quark
u

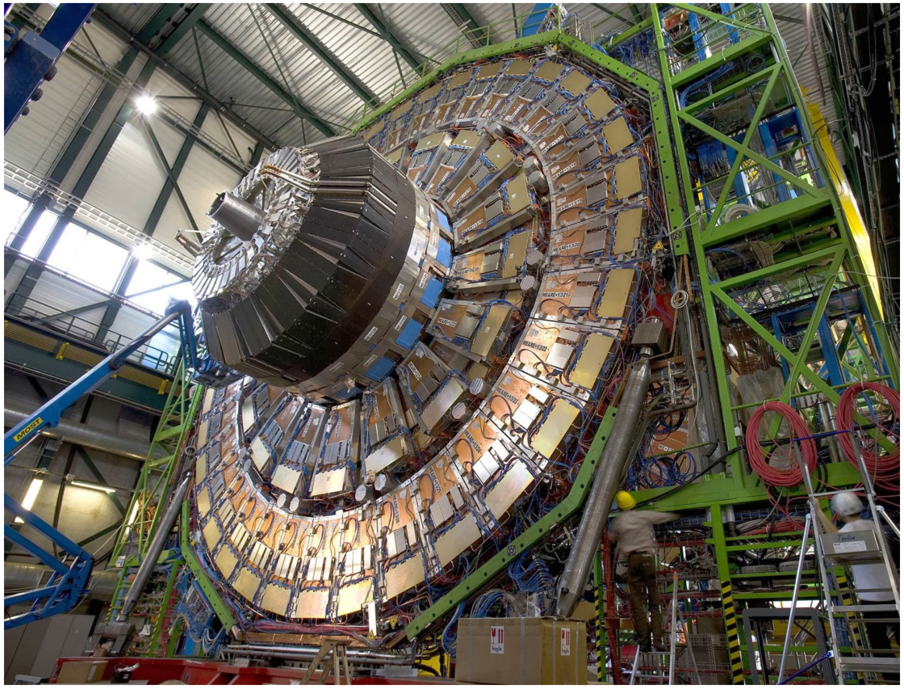
Fermilab



Quarks
u
up
c
charm
t
top
d
down
s
strange
b
bottom
Forces
Z
Z boson
γ
photon
W
W boson
g
gluon
H
Higgs
boson
Higgs field
e
electron
μ
muon
τ
tau
νe
electron neutrino
νμ
muon neutrino
ντ
tau neutrino
Leptons

CMS Experiment at the Large Hadron Collider

- **Protons collide** at the Large Hadron Collider **14 million times per second**
- **100 Megapixel** "camera" captures energy, position
  - **1000 times per second**
- All measurements of a collision are called an "**event**"



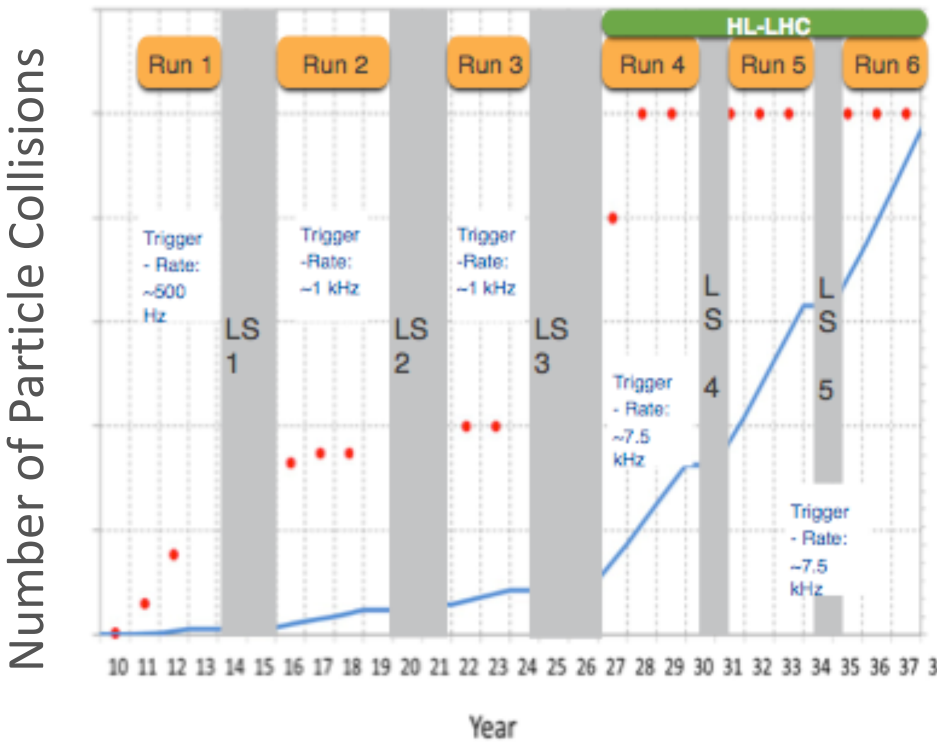
Scale of CMS computing

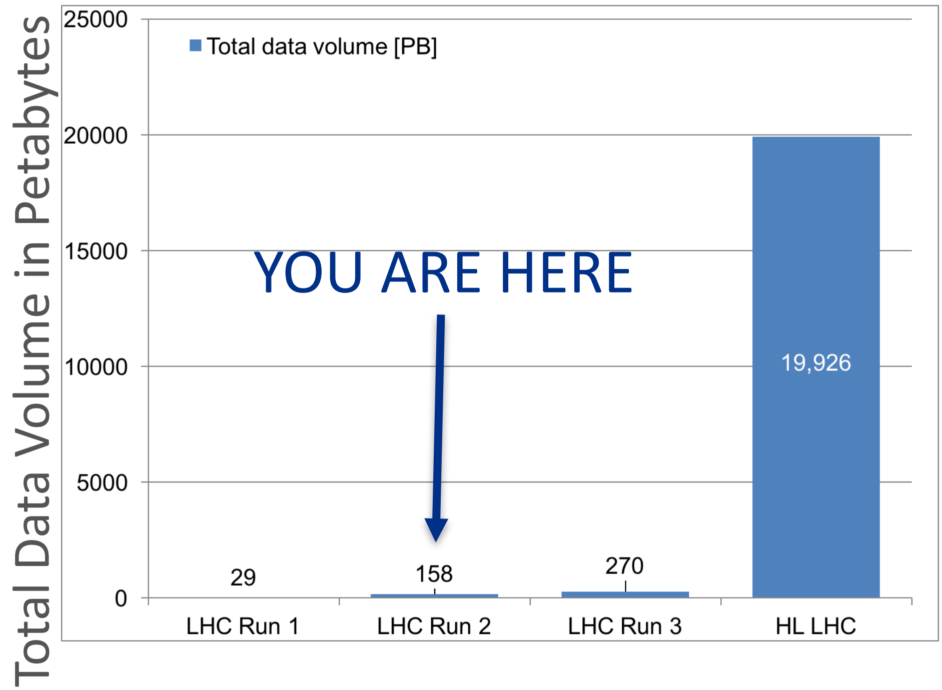
- 70+ compute clusters (Open Science Grid and Worldwide LHC Computing Grid)
  - 150,000 cores
  - ~75 Petabyte Disk
  - ~100 PB used tape space
- Strong networks connecting the individual sites
  - Weekly transfer volume between all sites: 4-6 Petabyte
  - Total LHC Trans-Atlantic network capacity: 340 Gigabits per second

High-throughput computing – data-intensive science

Open Science Grid provides tools, packaging, common ops

But that will not be enough …

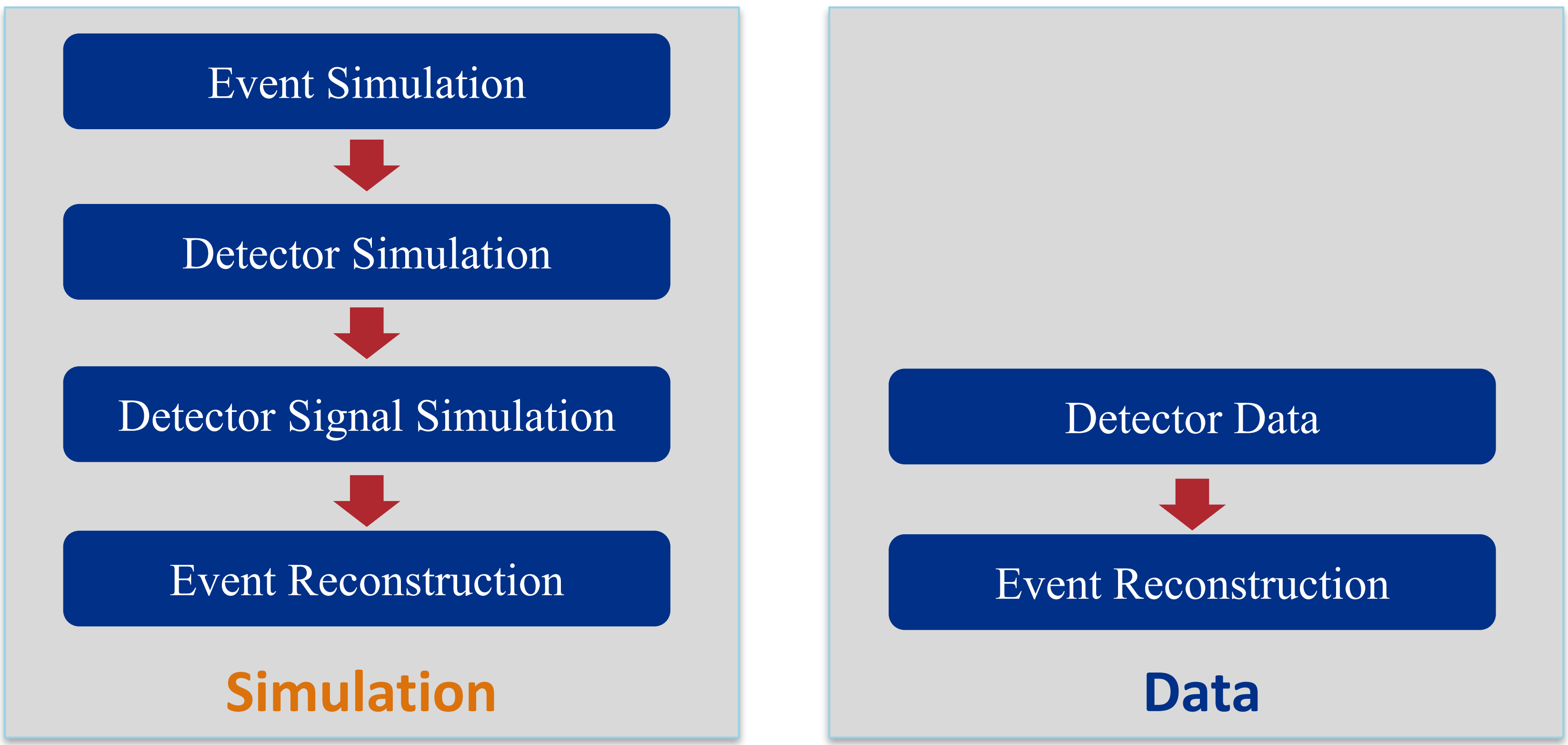




High Energy Physics computing will need **10-100x** current capacity

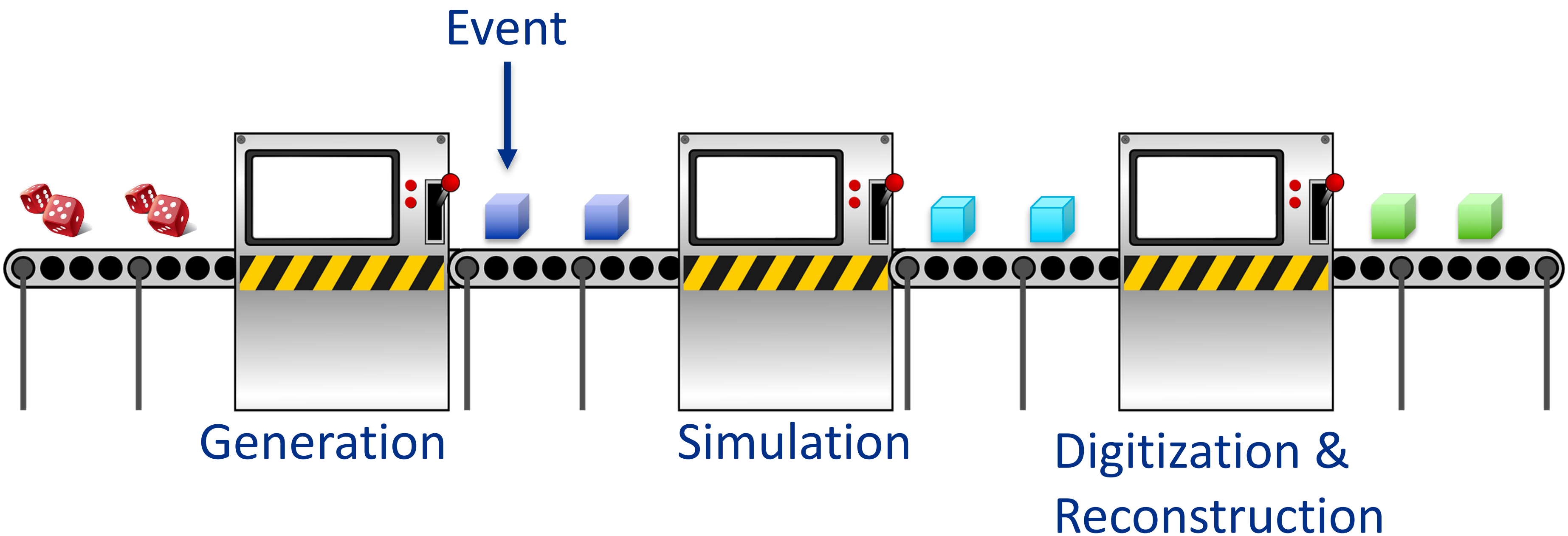
What is HEP Analysis? – "Production"

- Coordinated production of reduced-sized datasets for use by the collaboration in analysis



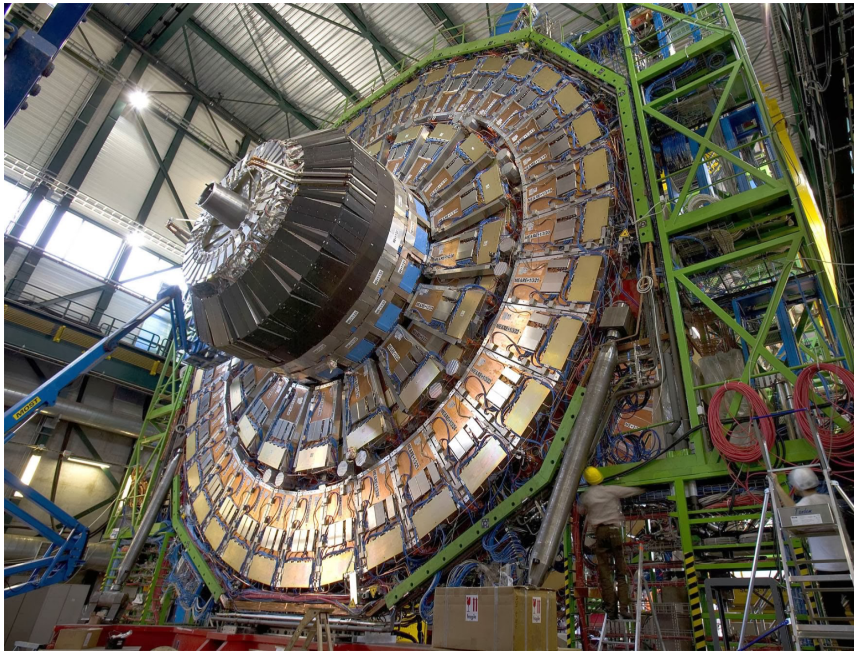
clipartfest.com

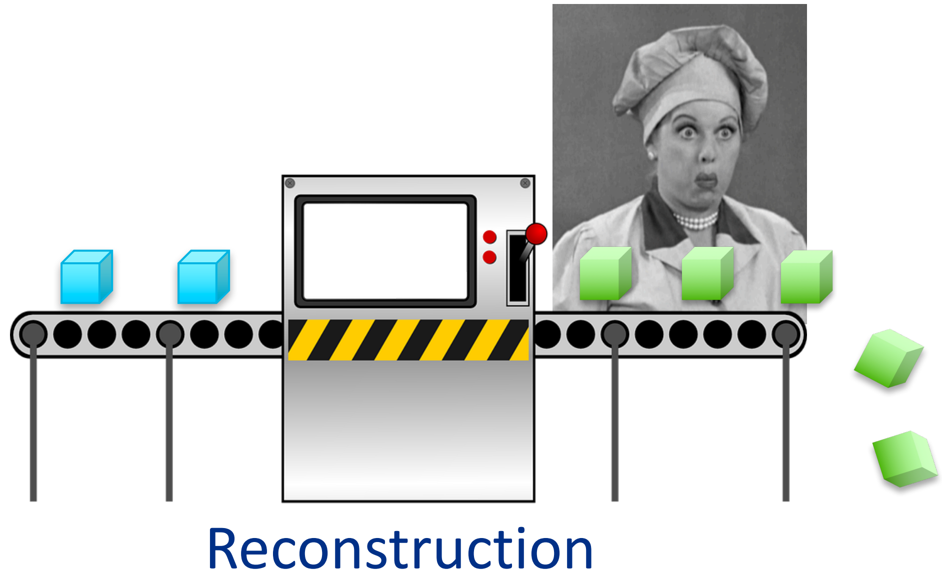
Production is "Production Line"



Fermilab

Production is essentially online





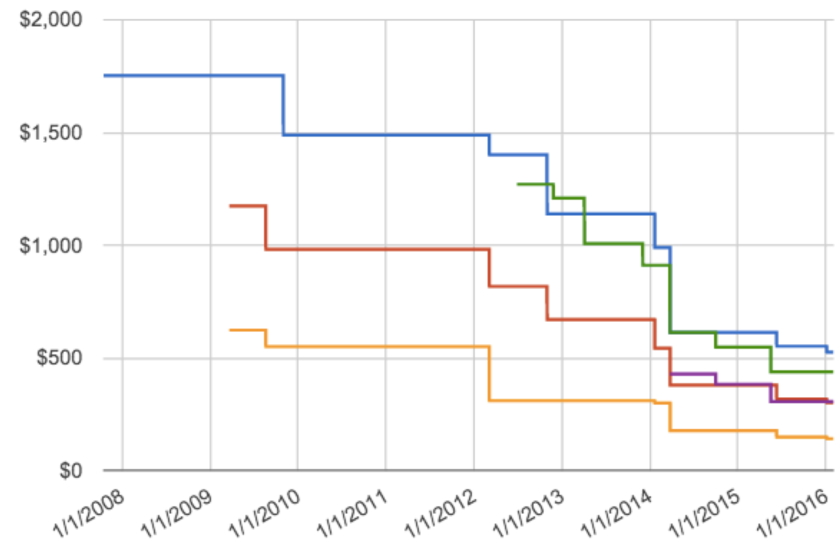
What about user analysis?

- The term analysis is overloaded
  - At a recent workshop, we spent more time arguing about **the definition of "analysis"** than on technical matters
- Chaotic and difficult to predict
  - After production, data is in the hands of analysis groups and users
- Common tools – but specific to HEP
  - Cross-experiment frameworks (*art*), Gaudi/Athena, CMSSW
  - ROOT (http://root.cern.ch)
  - Some movement towards broader tools: HDF5, R
- I'll touch briefly on analysis at the end of the talk

A Computing Model for HEP

Moore's Law (Dennard Scaling) is Dead; but

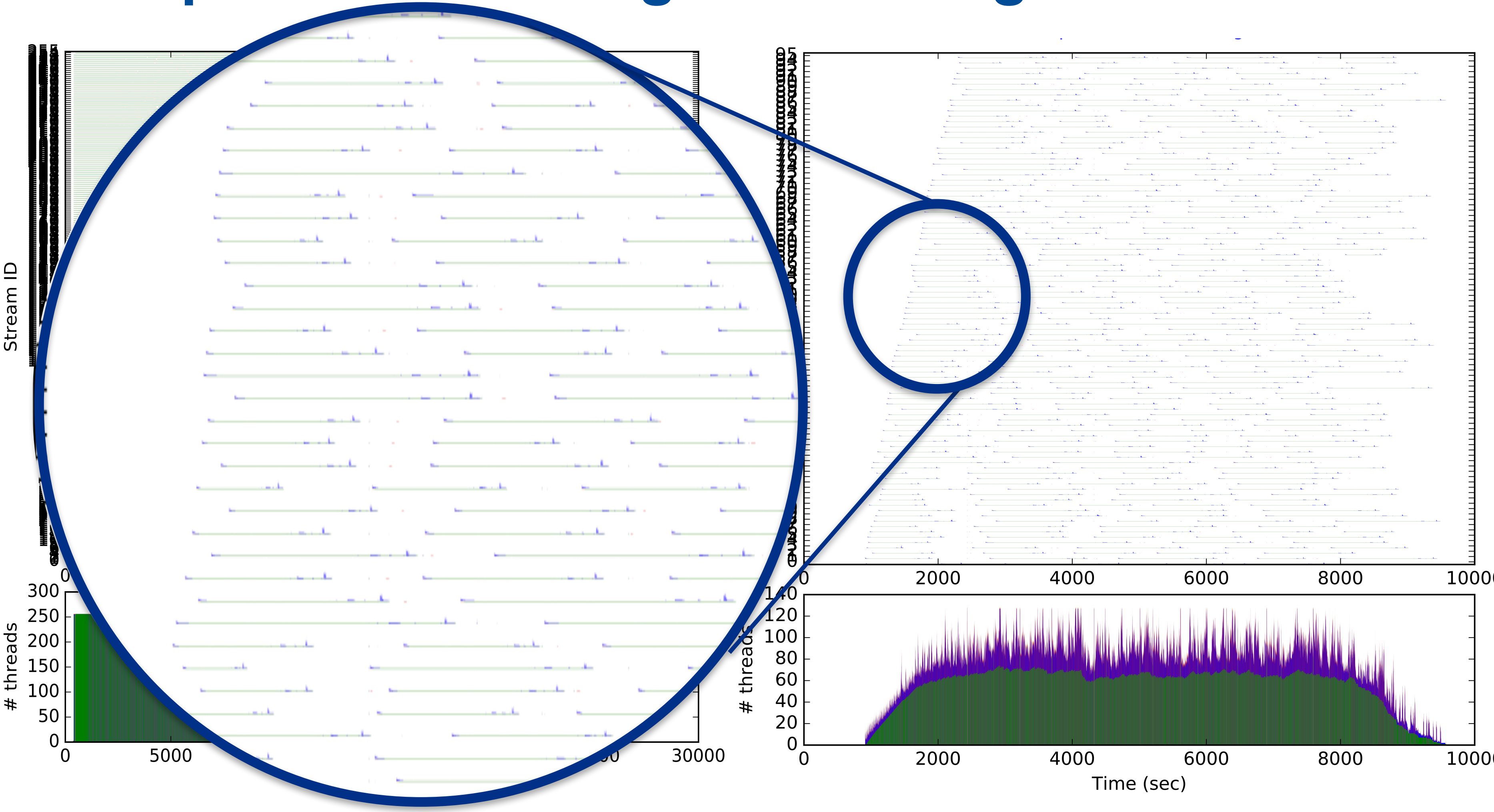
- The number of compute cores will grow to be **essentially infinite** with little cost
  - Industry: commercial cloud, IoT
  - Research: NSCI, Exascale program, Leadership Computing Facilities



- Experimental software frameworks (with some foresight and some luck) have already been refactored for **many-core architectures**

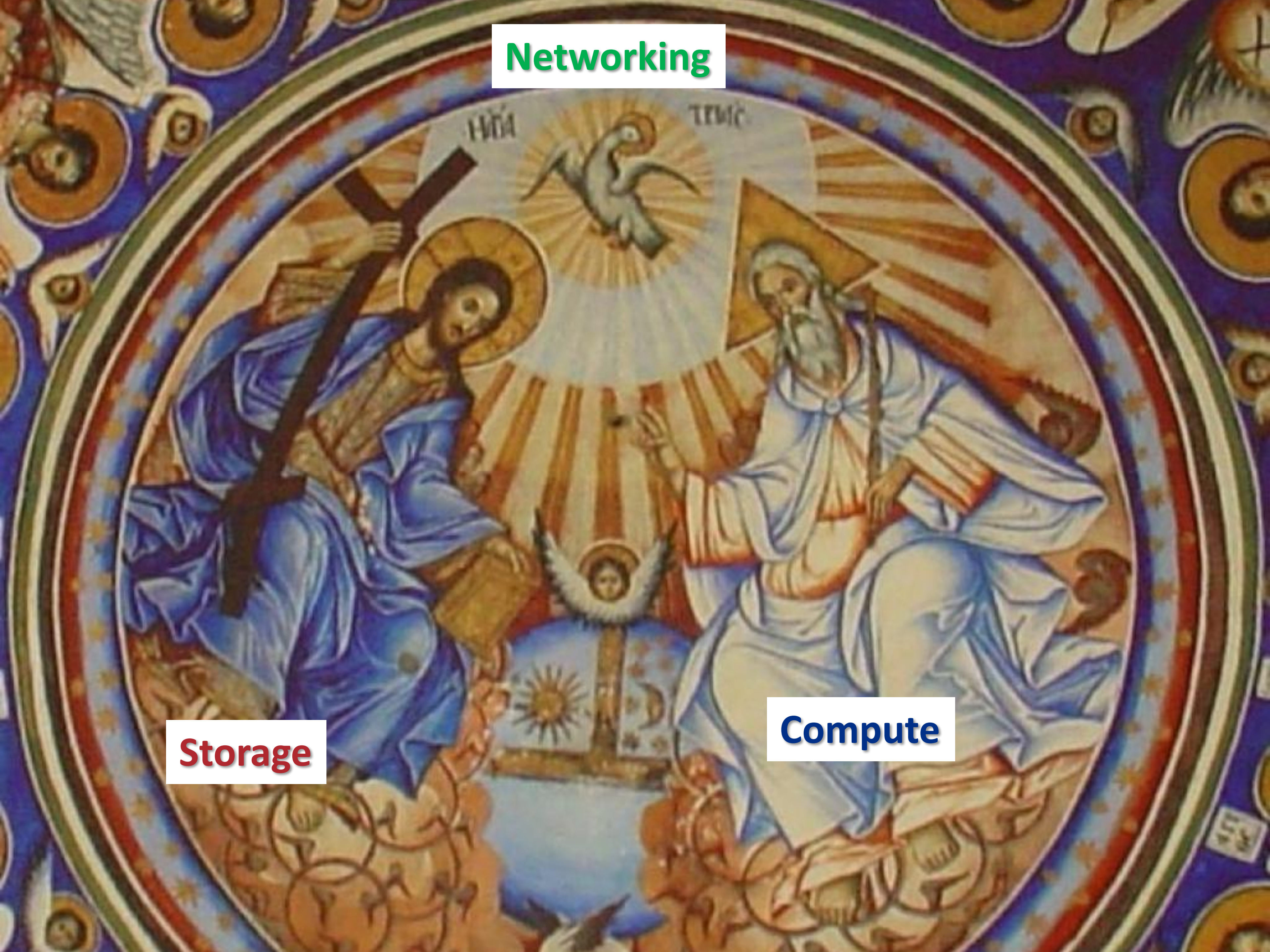
Fermilab

Example: CMS on Knights' Landing

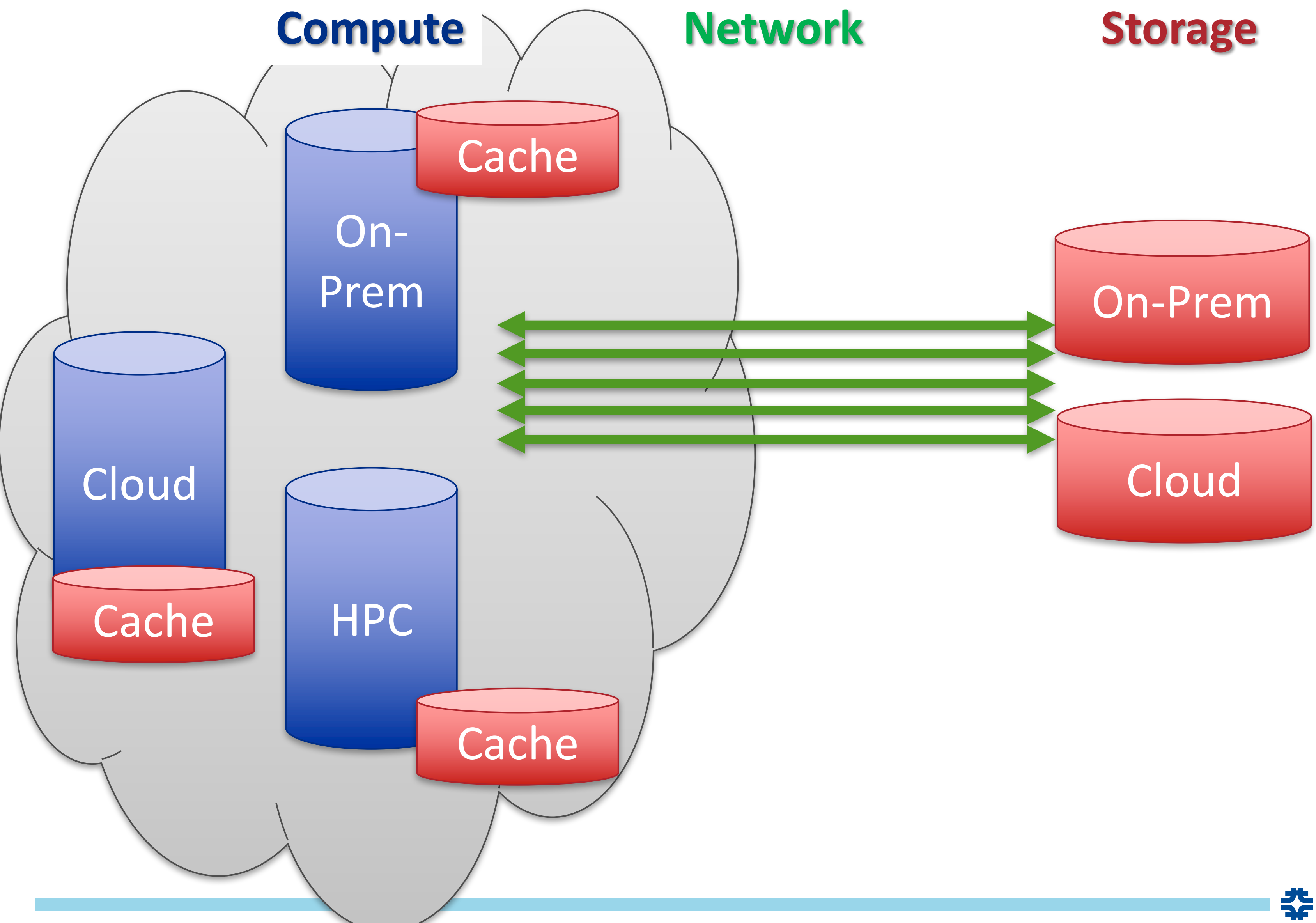


Generation + Simulation

Reconstruction

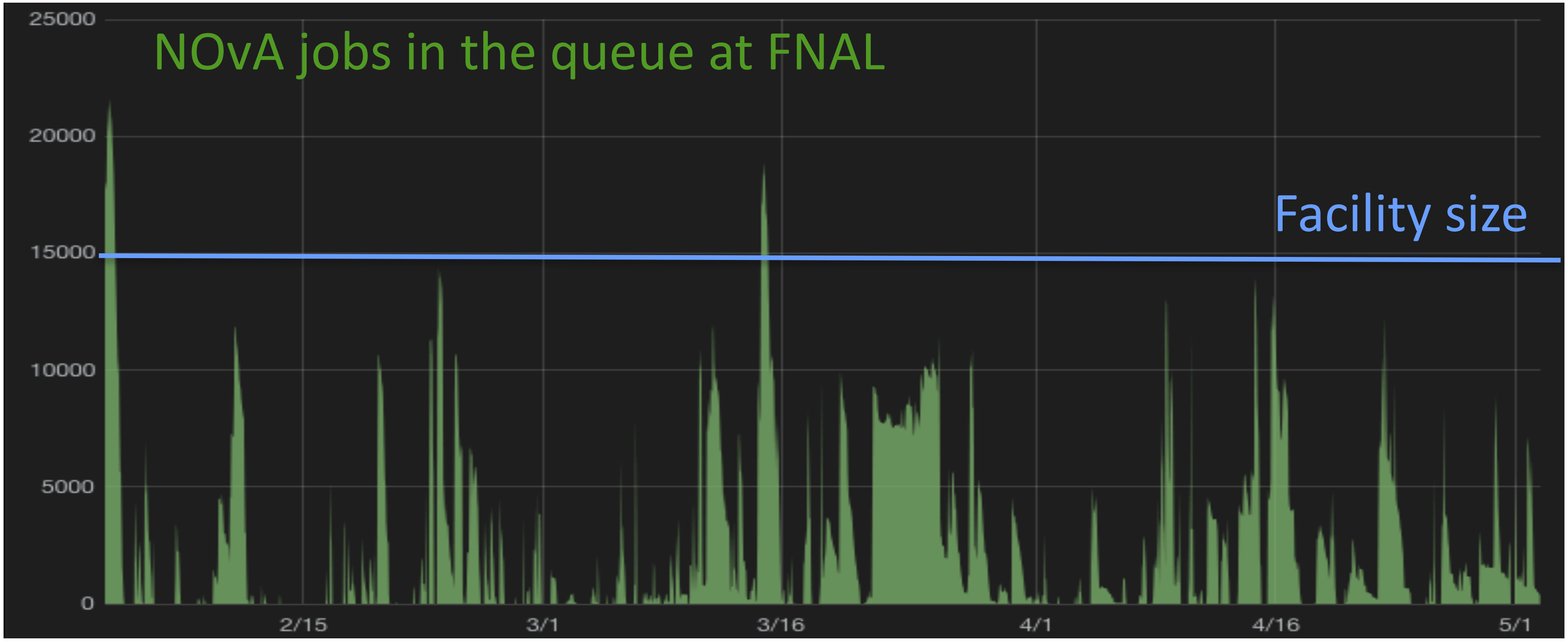

Networking
Storage
Compute

Bird's-eye View



Evolving the Computing Facility: Elasticity

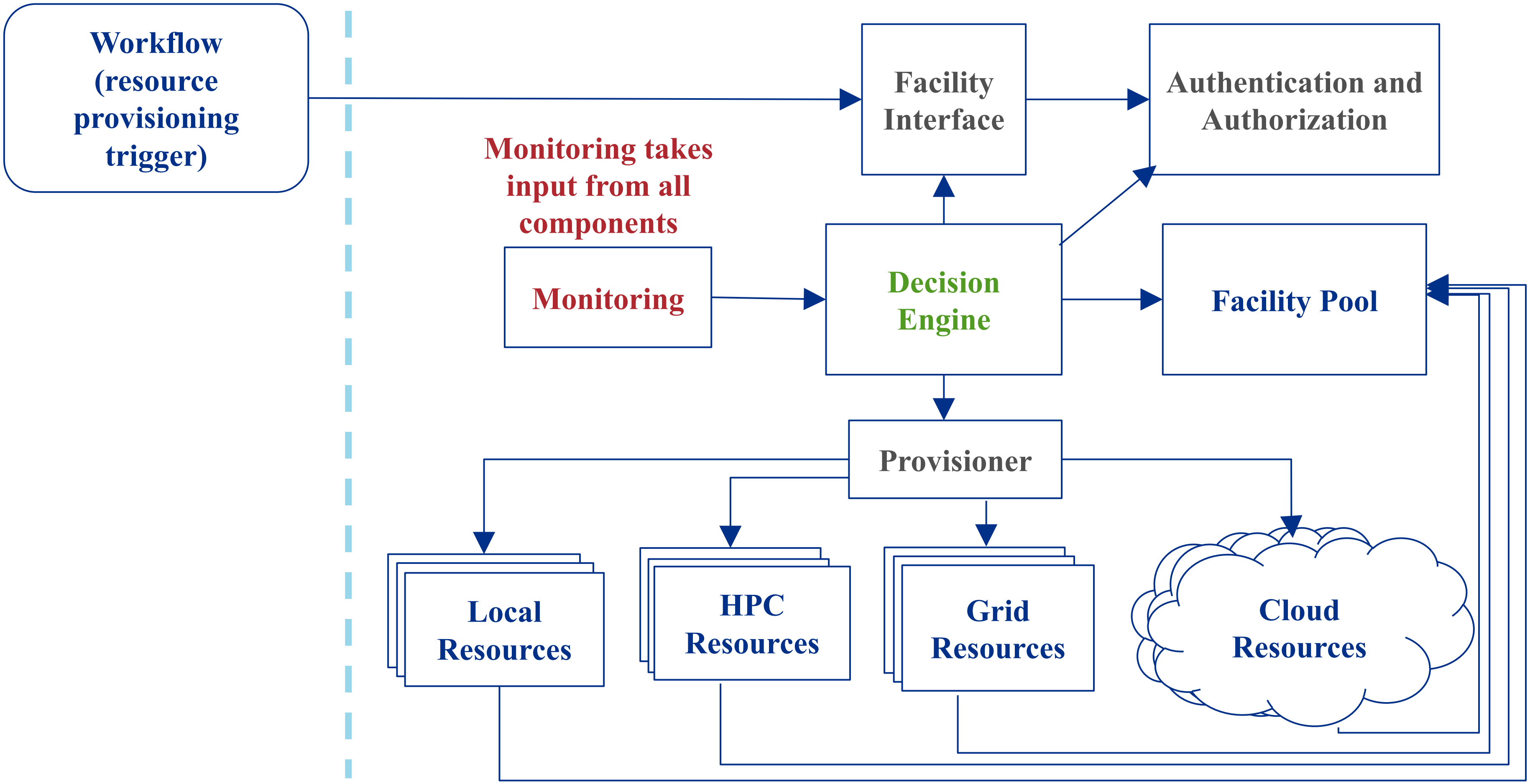
- Usage is not steady-state
- Computing schedules driven by real-world considerations (detector, accelerator, …) but also ingenuity – this is research and development of cutting-edge science



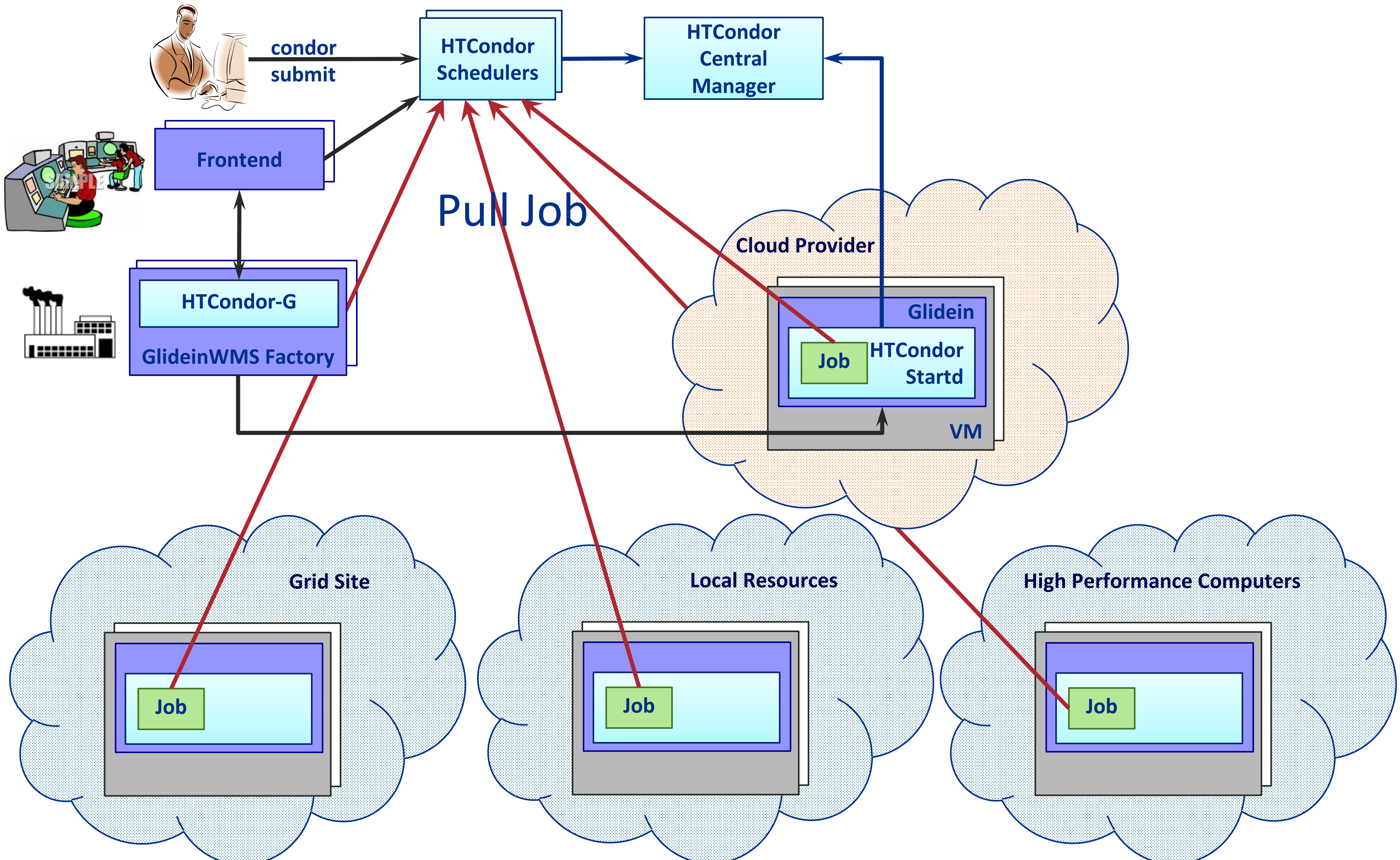
Inside an elastic facility: "HEPCloud"

- Vision Statement
  - HEPCloud is envisioned as a **portal** to an ecosystem of diverse computing resources, commercial or academic
  - Provides "complete solutions" to users, with agreed upon levels of service
  - The Facility routes to local or remote resources based on **workflow requirements**, cost, and efficiency of accessing various resources
  - **Manages allocations of users** to target compute engines
- Pilot project to explore feasibility, capability of HEPCloud
  - Goal of moving into production during CY18
  - Seed money provided by industry
- Don't blame for the name, I didn't name it

HEPCloud Architecture

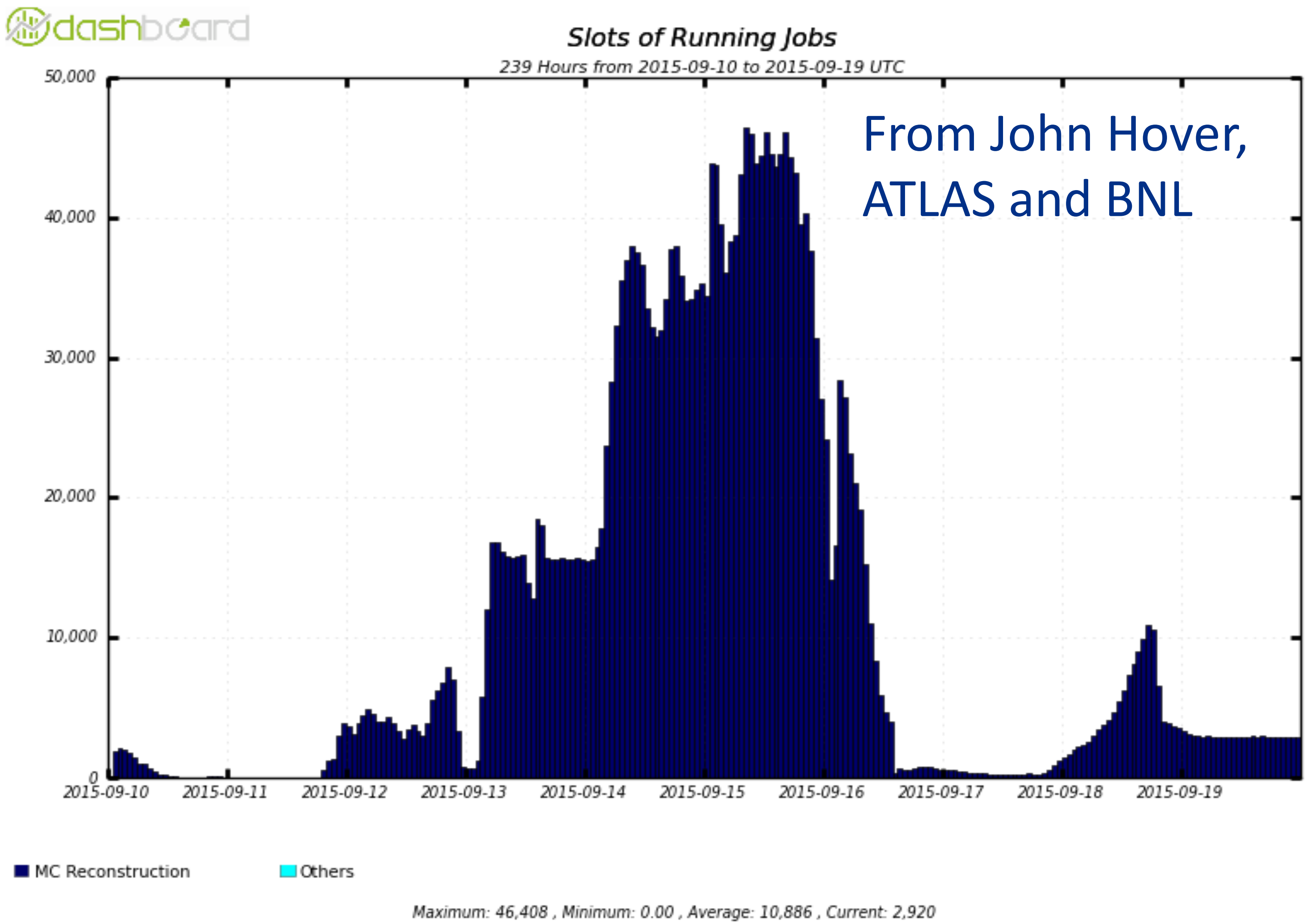


HEPCloud – one pilot-based implementation

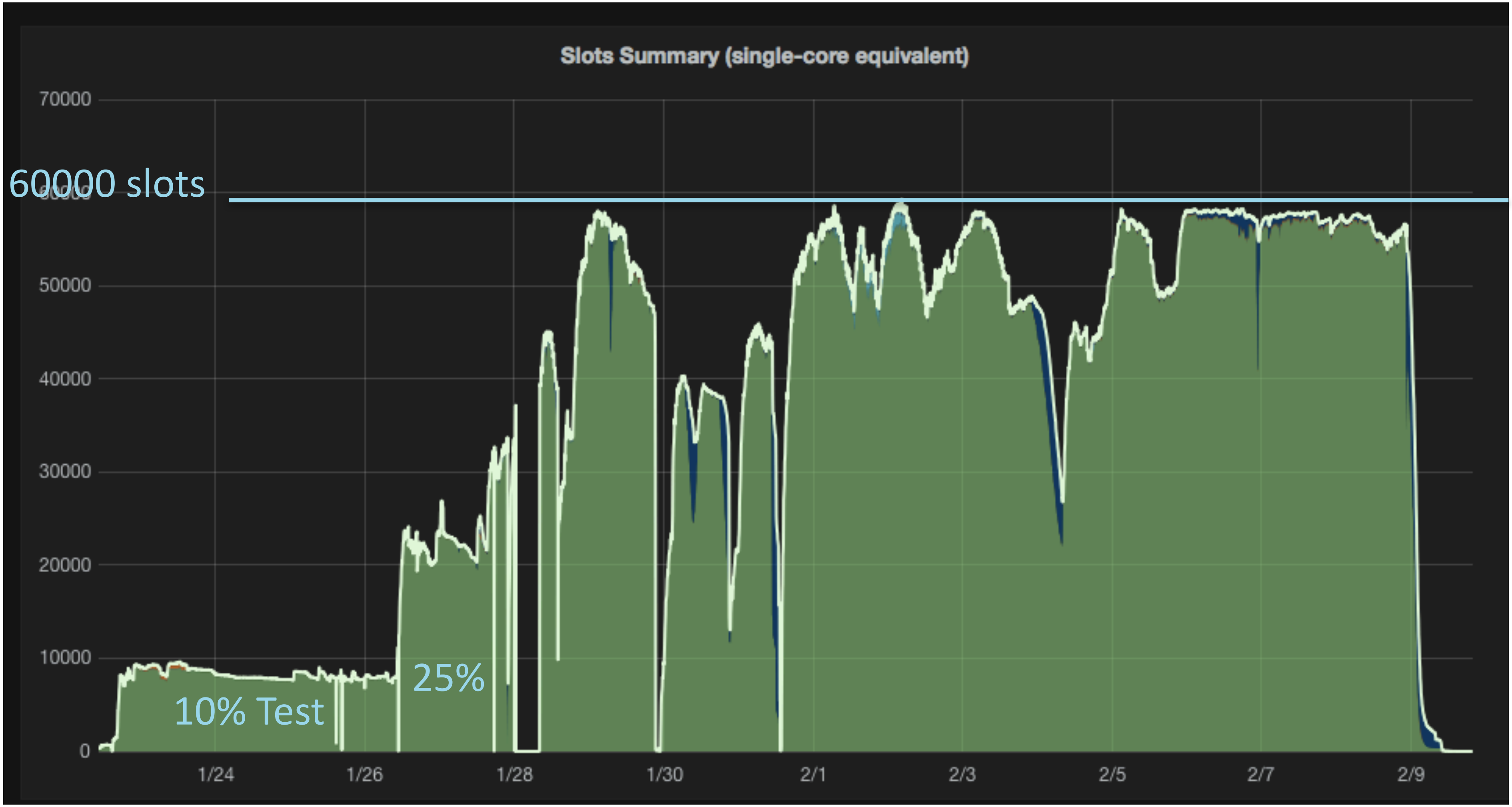


Fermilab

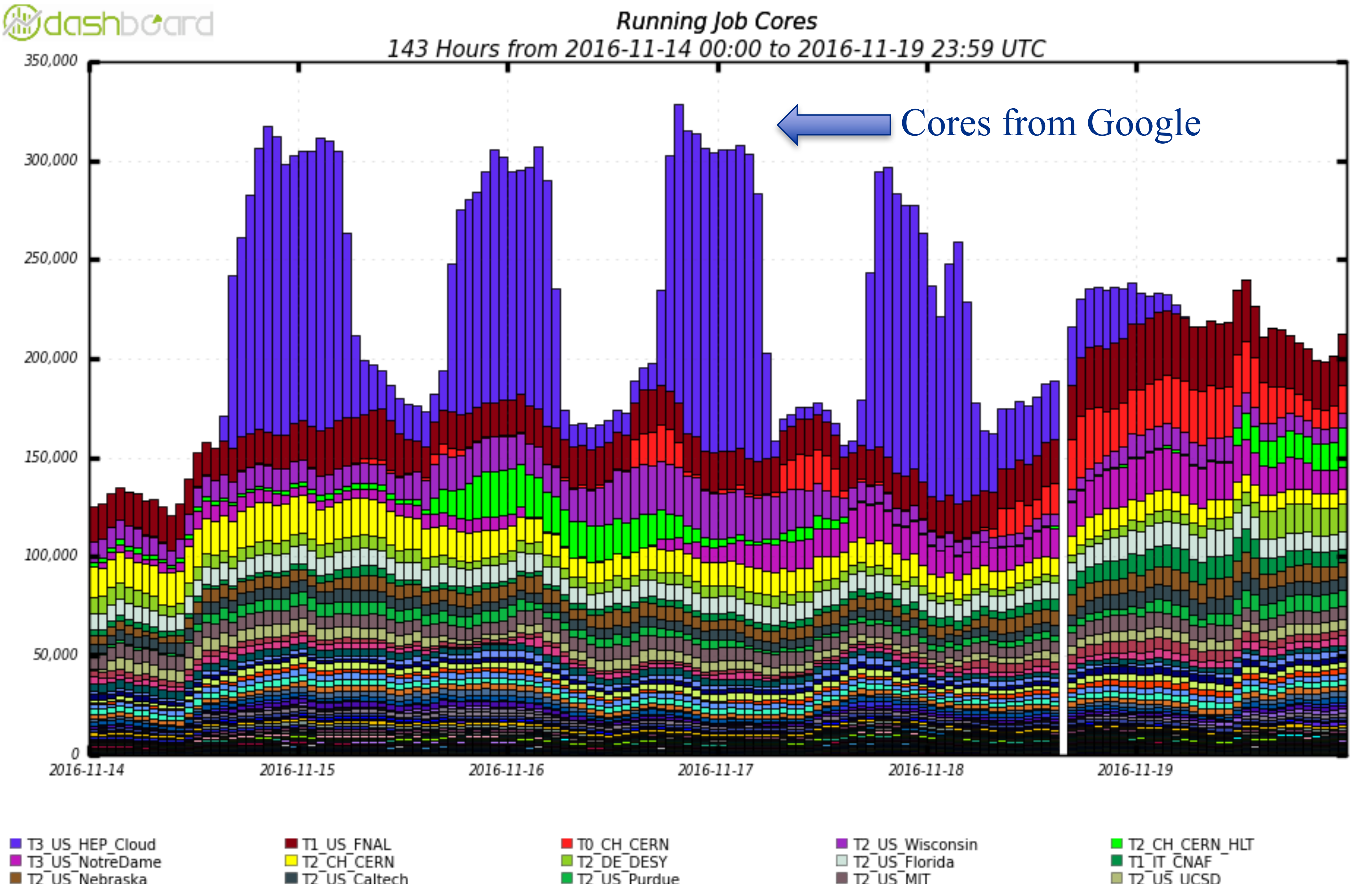
Reaching 45k slots on AWS via elastic provisioning



Reaching 60k slots on AWS with FNAL HEPCloud



Doubling CMS's compute reach @ Google via FNAL HEPCloud

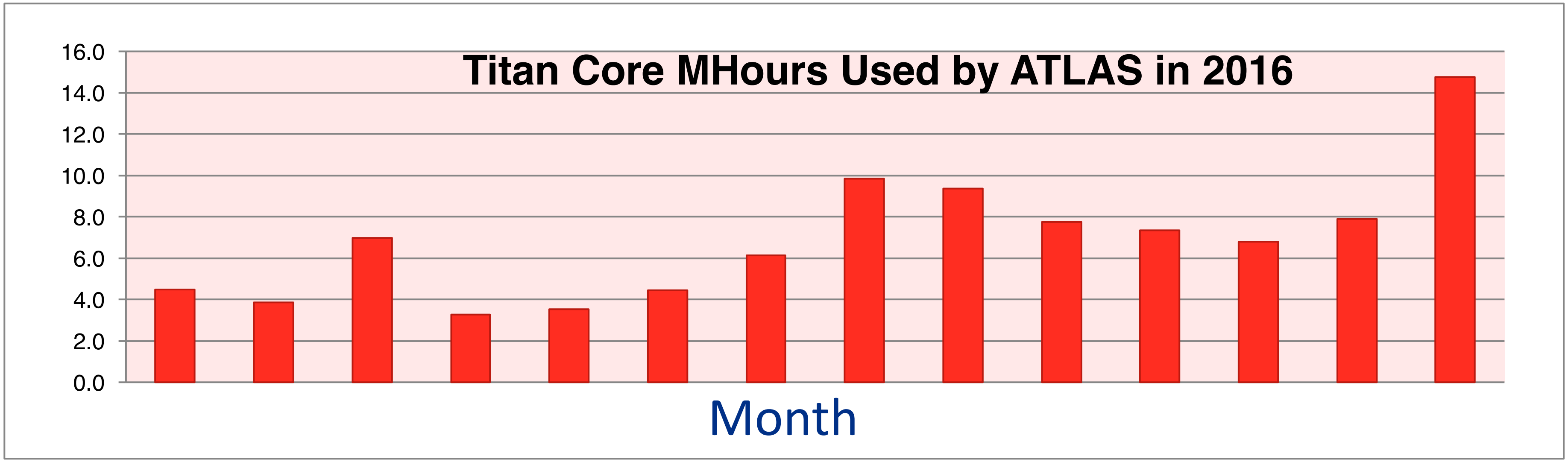


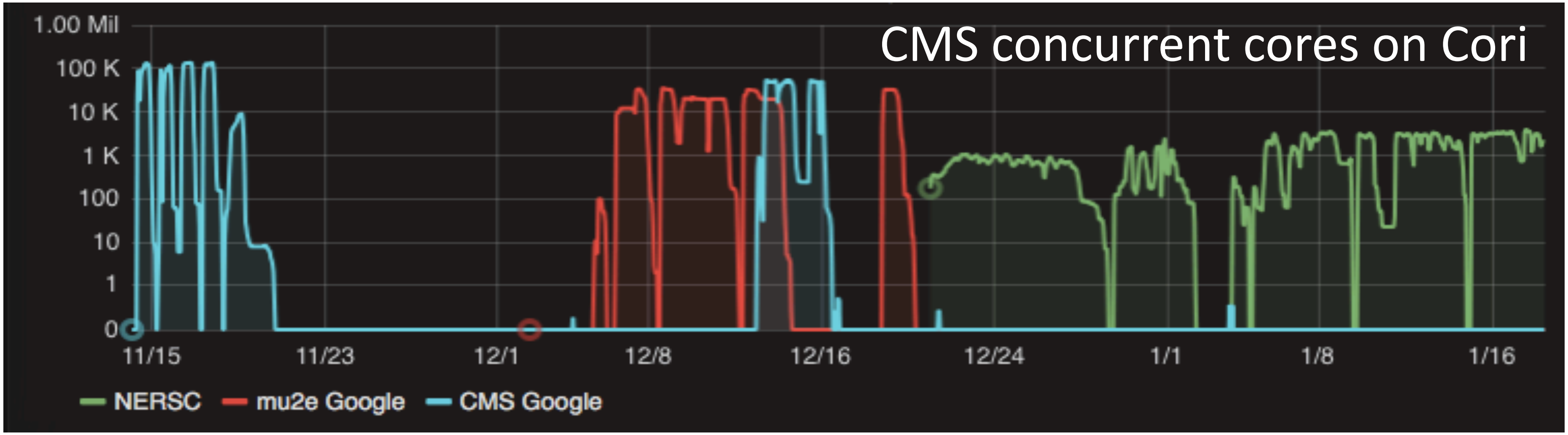
Future Directions for
NSF ADVANCED COMPUTING INFRASTRUCTURE
to Support U.S. Science and Engineering in 2017–2020

To support data-driven science, advanced computing hardware and software systems will need adequate data capabilities, in most cases more than is currently provided. Some research will need large-scale data-centric systems with data-handling capabilities that are quite different from traditional high-performance computing systems. For example, data analytics often requires that data reside on disk for extended periods. Several factors suggest that meeting these needs will require one or more large investments, rather than just multiple small projects, including the following: (1) the scale of the largest problems, (2) the opportunities for new science when disparate data sets are colocated, and (3) the cost efficiencies that come from consolidating facilities. **Indeed, the growth in data-driven science suggests that investments will ultimately be needed on a scale comparable to those that support modeling and simulation**. At the very least, the systems should be better balanced for data (input/ output and perhaps memory size), thereby allowing the same systems to be used for different problems without needing to double the size of the resources. As data play a growing role in scientific discovery, long-term data management will become an important aspect of all planning for advanced computing. A partnership with a commercial cloud provider could provide access to larger systems than NSF could afford to deploy on its own. Of course, even as it moves to provide better support for data- driven research, NSF cannot neglect simulation and modeling research.

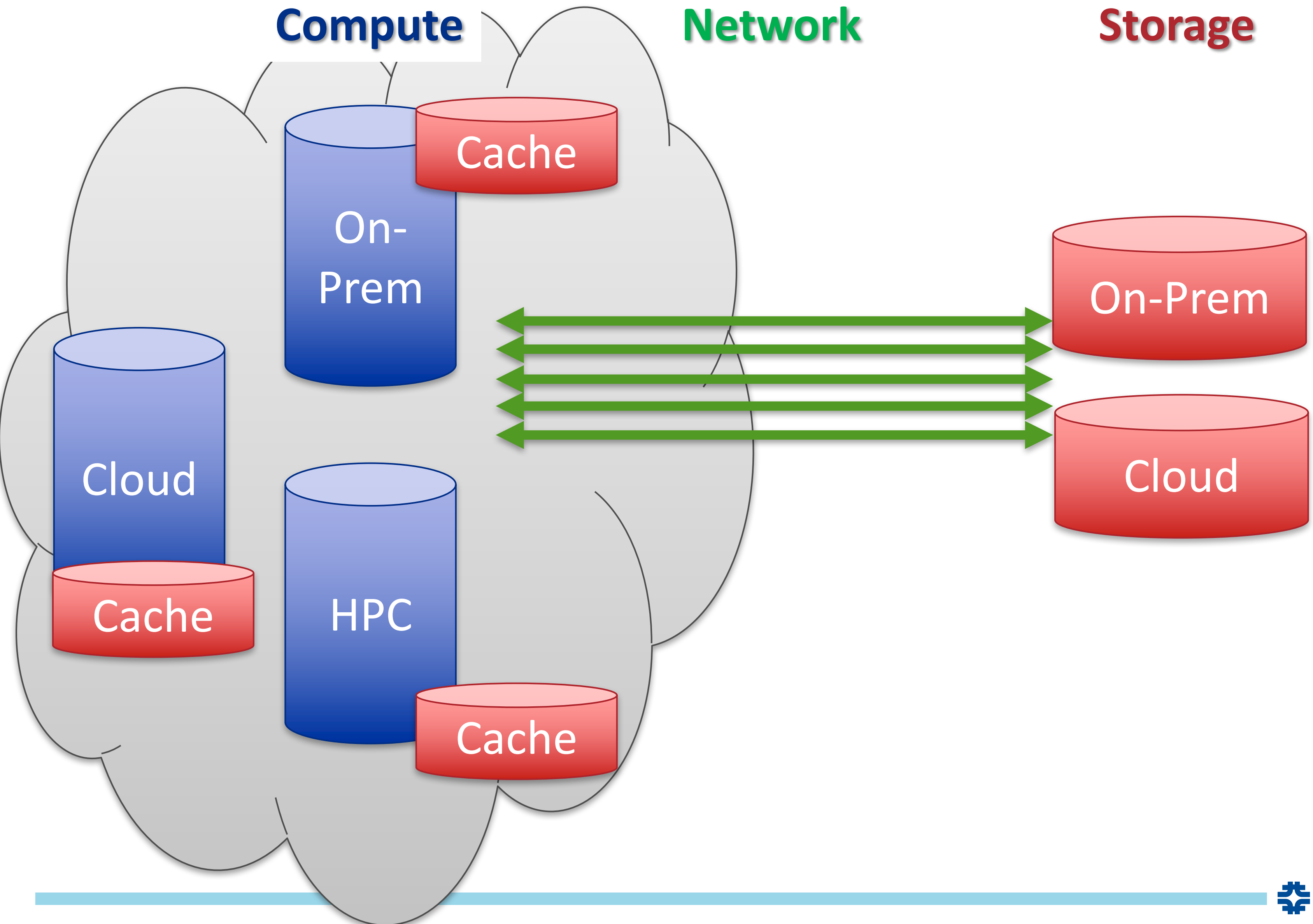
Recommendation 2.2. NSF should (a) provide one or more systems for applications that require a single, large, tightly coupled parallel computer and **(b) broaden the accessibility and utility of these large-scale platforms by allocating high-throughput as well as high-performance workflows to them.**

HEP use of DOE HPC facilities – two examples





Bird's-eye View



Fermilab

Workflow and Data Management

- The community has built experiment-agnostic workflow/provisioning tools based on pilot paradigms
  - Single-experiment (PanDa)
  - General purpose (glideinWMS, Dirac)

> Vaandering's Law: Any discussion of workflow management evolves into a discussion on data management

- Data management solutions
  - Do bookkeeping but also data placement/transfers
  - Often tightly integrated with experiment (DQ2, PhEDEx, …)
  - General-purpose solutions are advancing (FIFE, Rucio, …)

One aside on networking

- We (HEP scientists) treat networking not as a first-class resource
  but as a utility
- We try to ensure we have enough capacity (with added safety factors)

- We only care about the size of the pipe when it gets clogged
- And then we desperately call the experts **(1-800-ESNet?)**

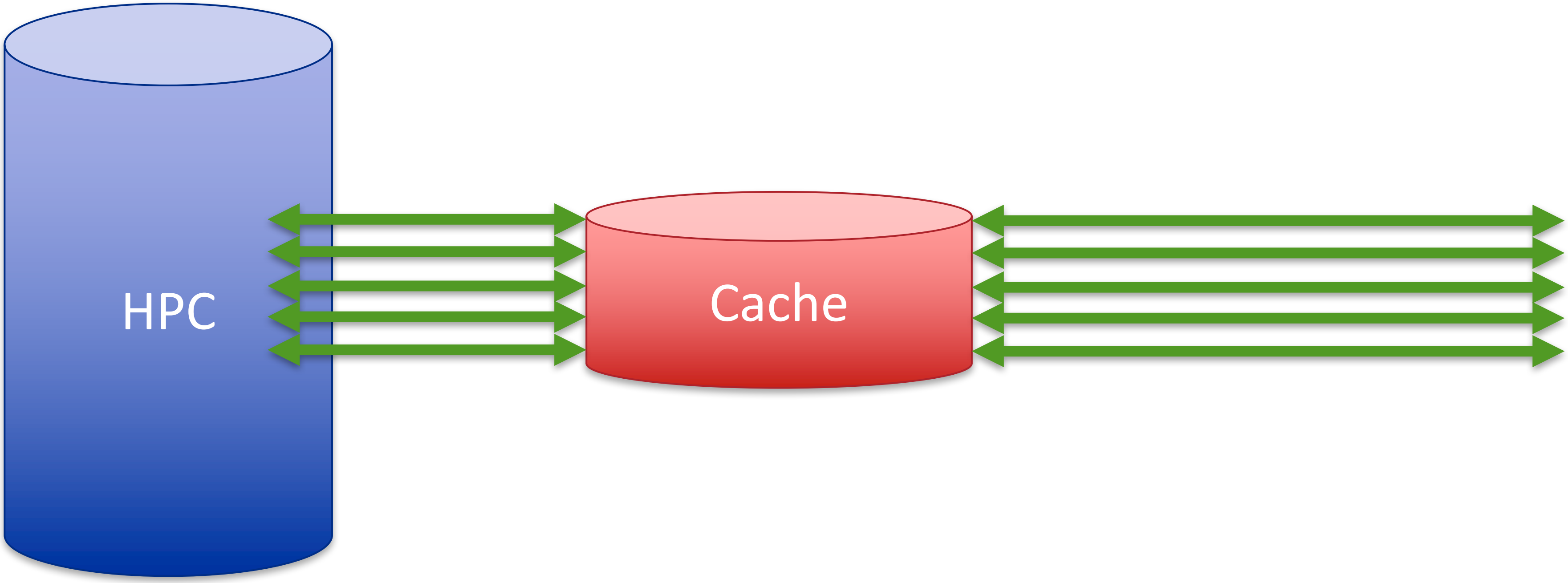
Challenge: distributed storage management

- Domain scientists should only manage **data silos**
- General solutions for data placement and transfer are a step forward, but

- A storage management system should handle
  - Managing **data caches** close to compute silos
  - Triggering **data transfers** and placement
  - Apportioning network bandwidth (with latency guarantees) – treating **network as a first class resource**

- This imposes requirements on the application side
  - Need to request **network** the same way we request cores, memory, local disk
- Could be rolled into “smart” networks?

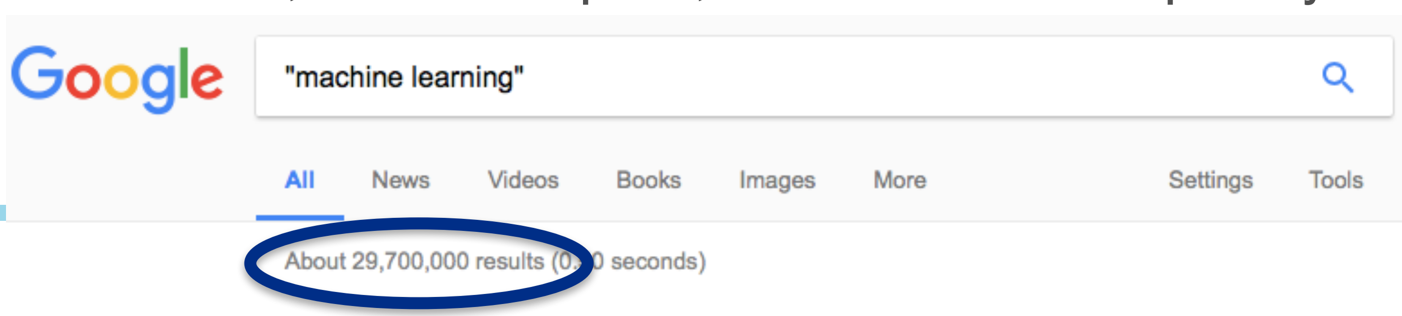
Storage and Networking at the Facility



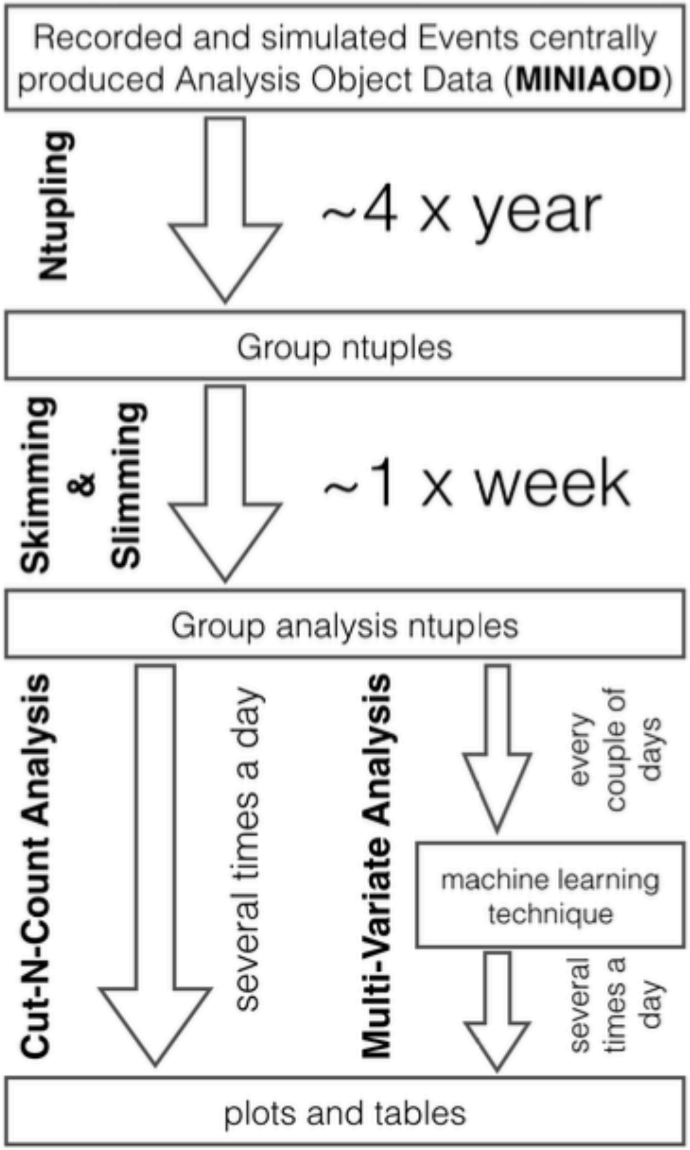
- Distributed, reliable, performant storage cache that **scales to the size of the computational resource**
- Good networking between the cache and compute workers
- Wide networks between the cache and ESNet

”New” analysis methodologies

- Machine Learning
  - ”Cybernetics” (1940s), ”Artificial Neural Networks” (1990s)
    - Marginal recognition in HEP
      “Ten years ago there was a great deal of excitation about NN in HEP. Many people felt that a revolution in HEP computing was in the works. It is perhaps a bit disappointing today to see that although NN have gained some measure of acceptance, the networks making their way into physics papers are still simple […] and often quite small ones at that.”
      - *Neural networks in high energy physics: a ten year perspective.”* Denby, B. CPC 199 (1999)
- Industry has caught up and surpassed us
  - More data, more compute, more model complexity

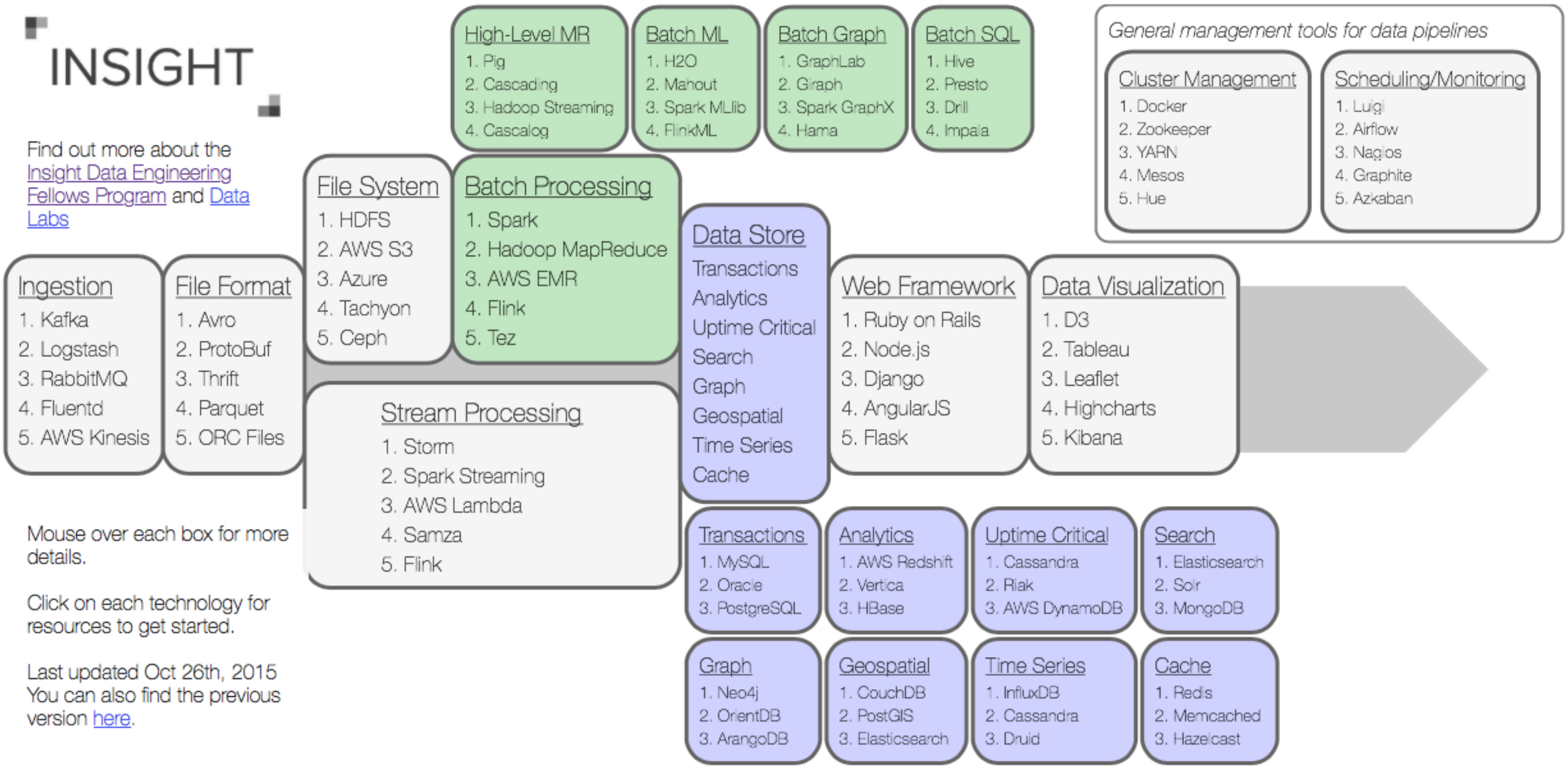


Analysis – also a conveyor belt



- New toolkits and systems collectively called "**Big Data**" technologies have emerged to support the analysis of petabyte and exabyte datasets in industry.
- Three-fold goal in HEP using Big Data
  - Reduce time to physics
  - Educate grad students/postdocs in data science skills
  - Leverage advances in toolkits outside of HEP

Rich landscape of Big Data tools



Summary and Needs

- Data, Data, Data!
  - We need facilities to support high throughput computing as a key science stakeholder
  - We need general-purpose robust storage and cache management solutions across distributed networks
  - We need smart networks to interconnect the two as we treat the ”network as a resource”
- Partnerships in industry will help: “data science” is a field
  - ”New” analysis methodologies
  - Co-design: begin optimizing hardware and software for each other

2017 is Fermilab's 50th Anniversary!

Visit http://50.fnal.gov/ for anniversary-related events and content

- **June 7**: 50th Anniversary Symposium
- **June 15**: Social media birthday celebration
- **September 23**: Public Open House and Innovation Fair

…and much more!

Fermilab