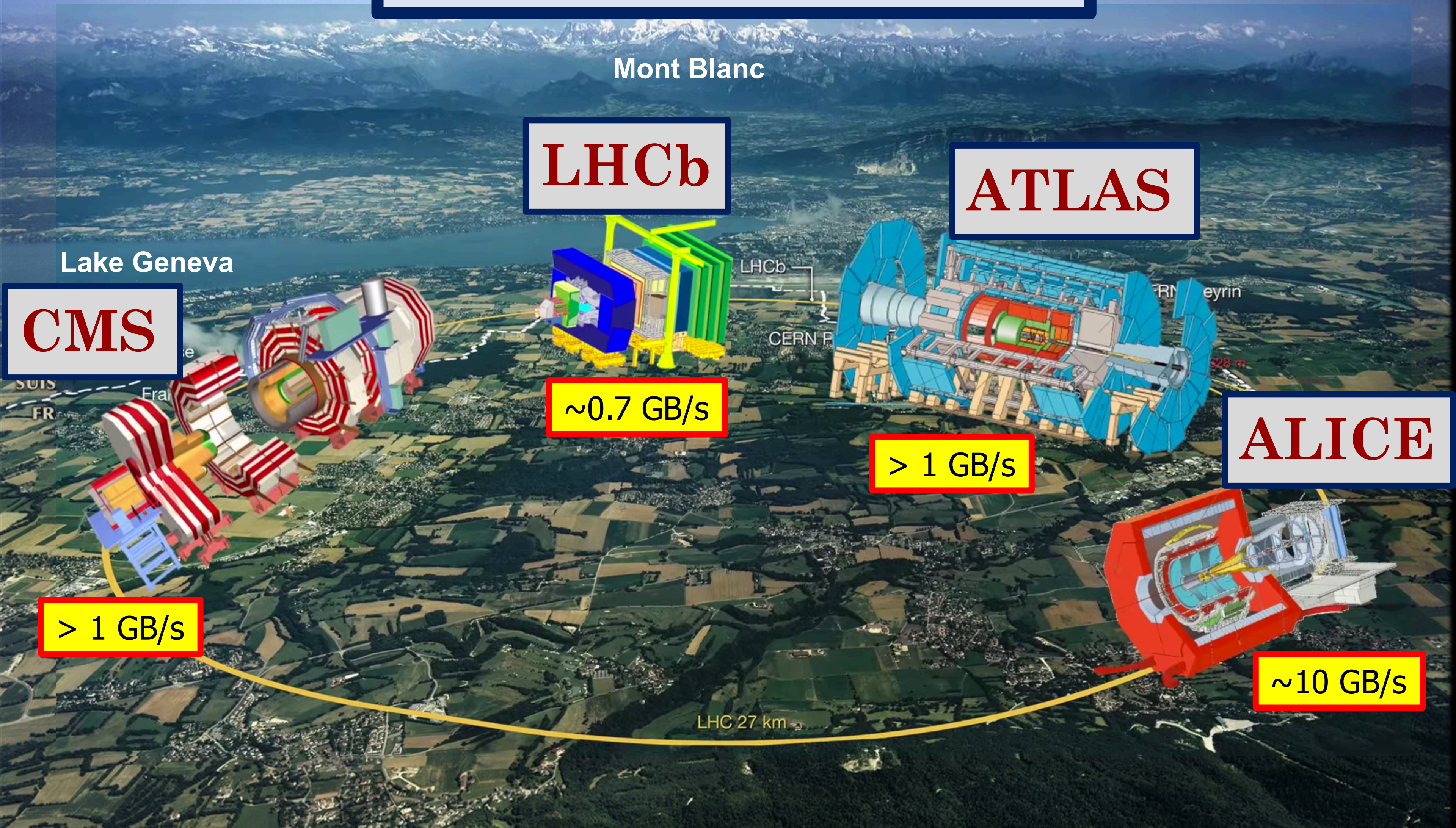
Computing Challenges & Opportunities

Mark Neubauer

University of Illinois at Urbana-Champaign

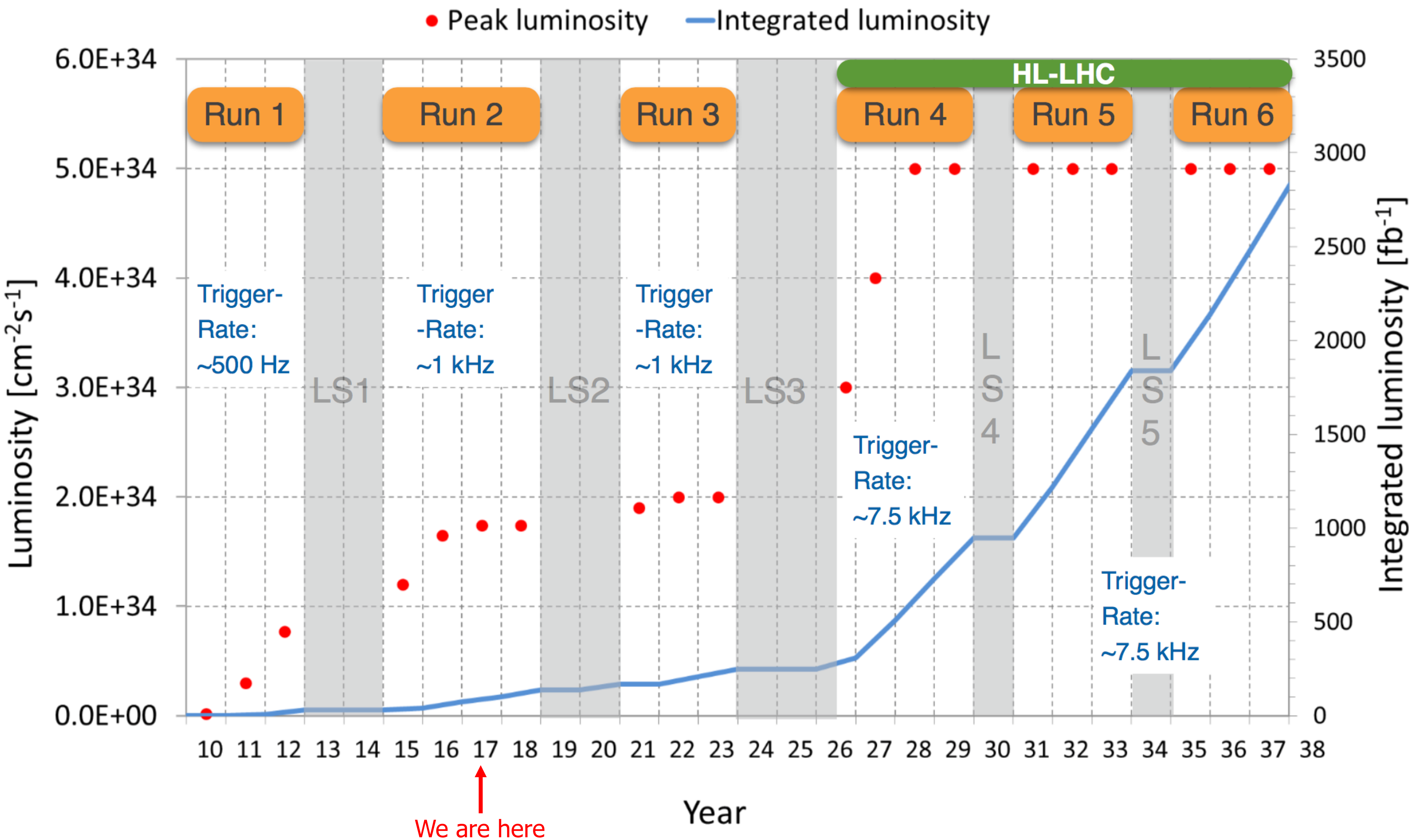
US ATLAS Physics Workshop
Argonne National Laboratory
July 28, 2017

LHC Experiments

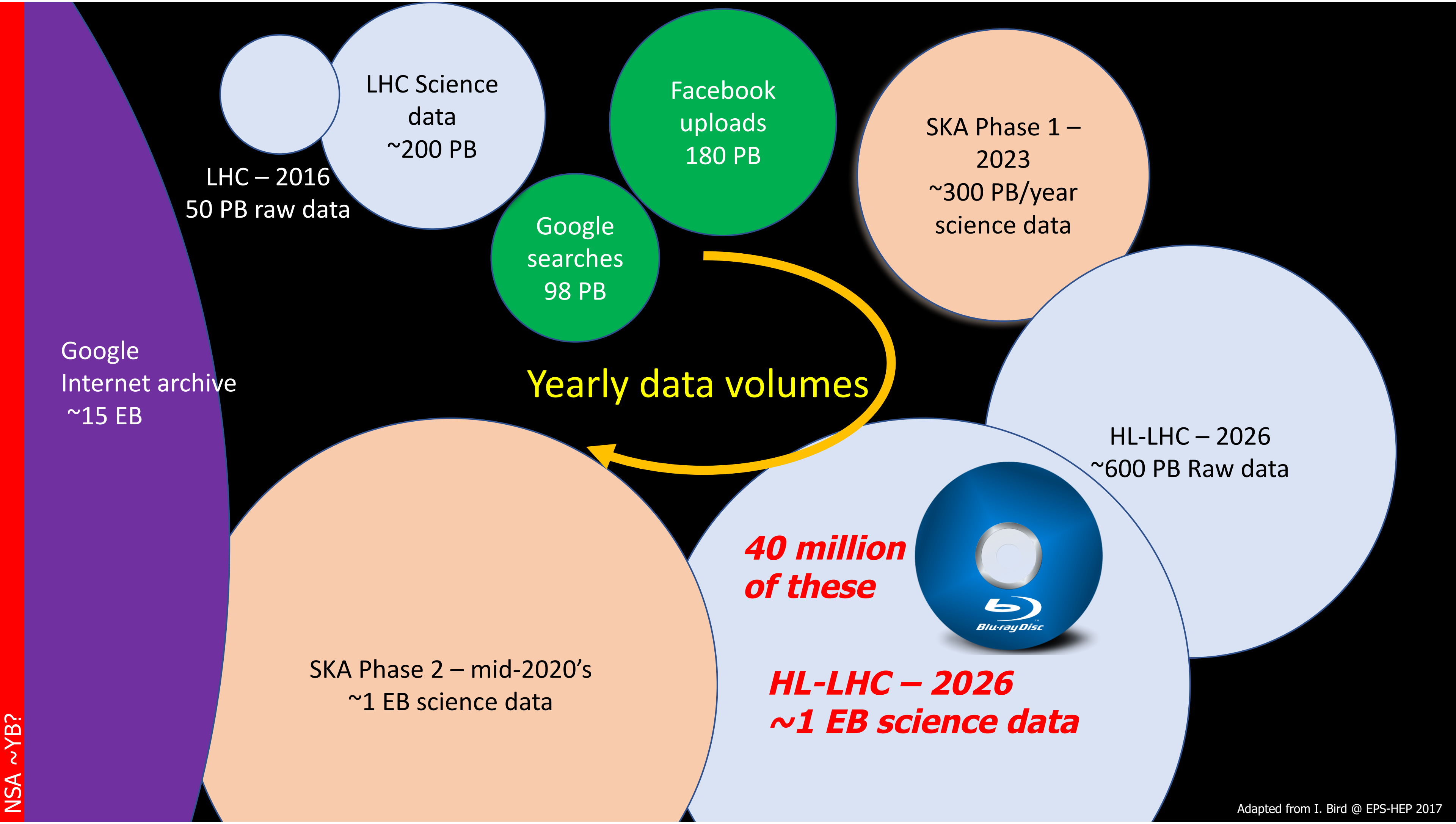


LHC Experiments generate 50 PB/year (during Run 2)

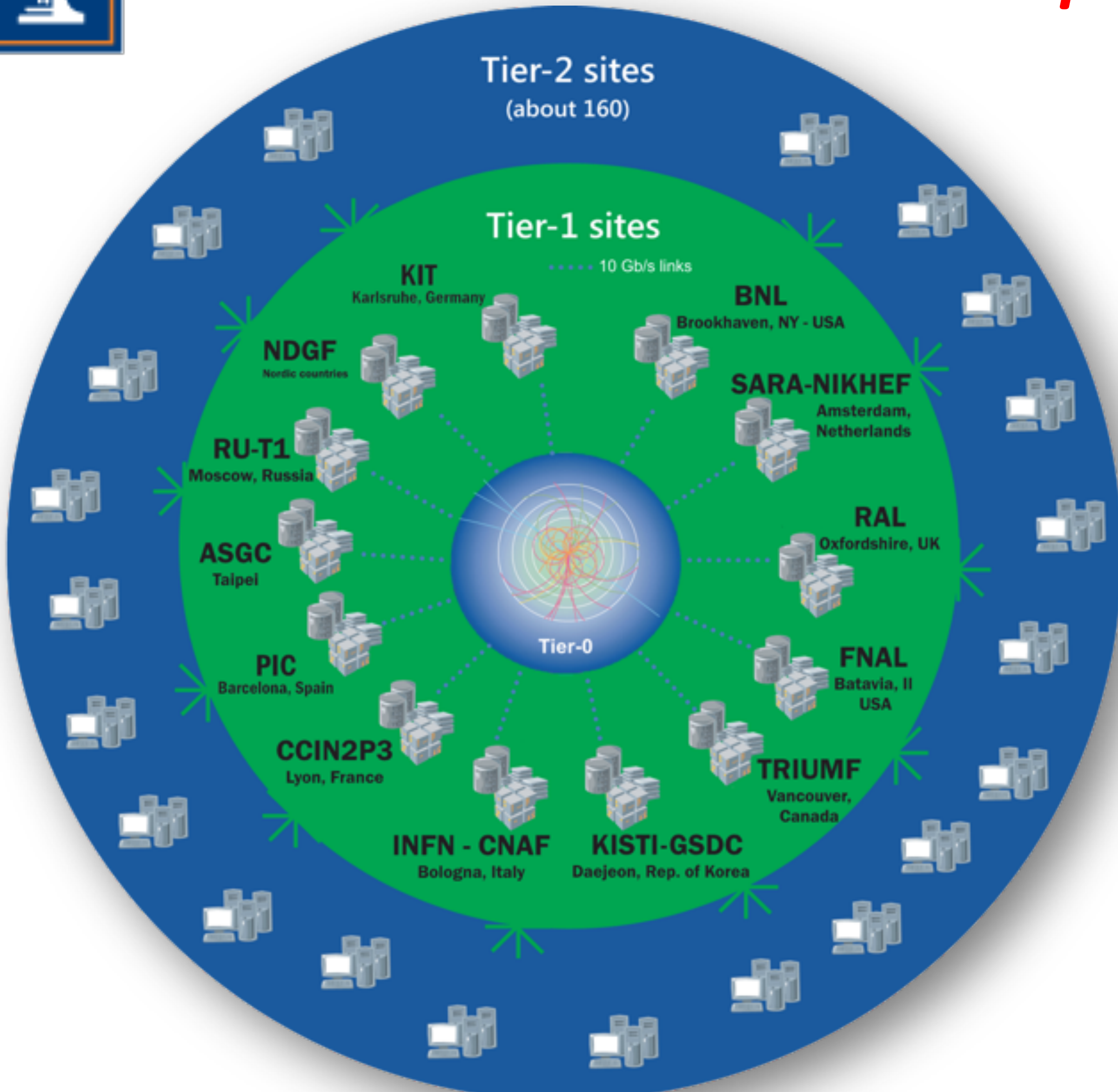
LHC Schedule



LHC as Exascale Science



Worldwide LHC Computing Grid



Tier-0 (CERN and Hungary):
- Data recording
- Event reconstruction
- Data distribution

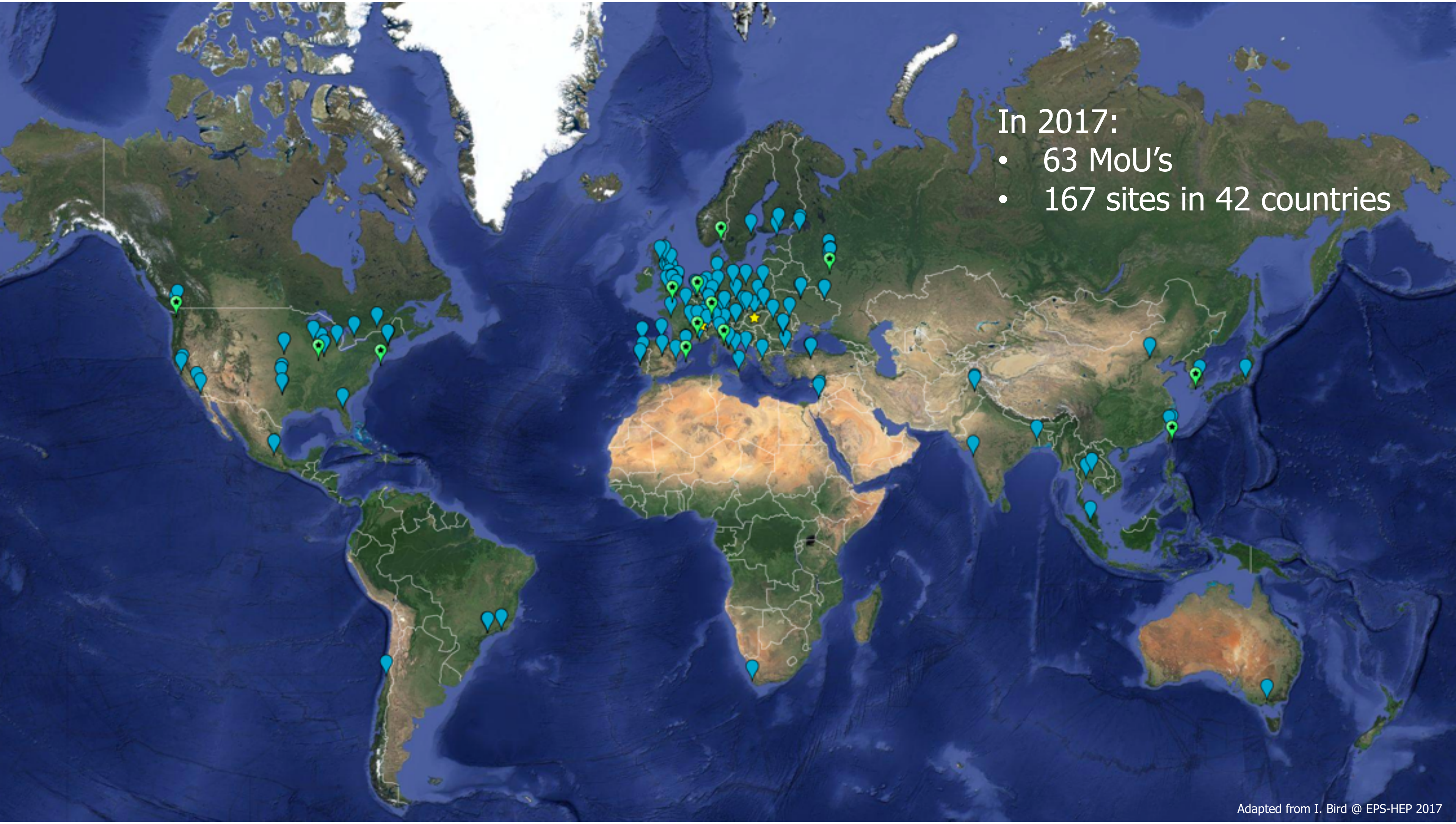
Tier-1 sites:
- Permanent data storage
- Event re-processing
- Data analysis

Tier-2 sites:
- Simulation
- End-user analysis

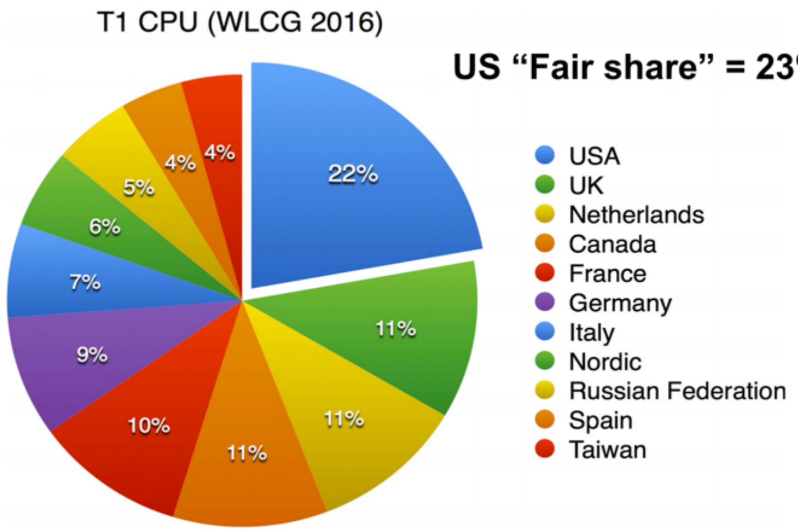
Important Factoids:
- 167 sites, 42 countries
- ~750k CPU cores
- ~1 EB of storage
- > 2 million jobs/day
- 10-100 Gbps links

Adapted from I. Bird @ EPS-HEP 2017

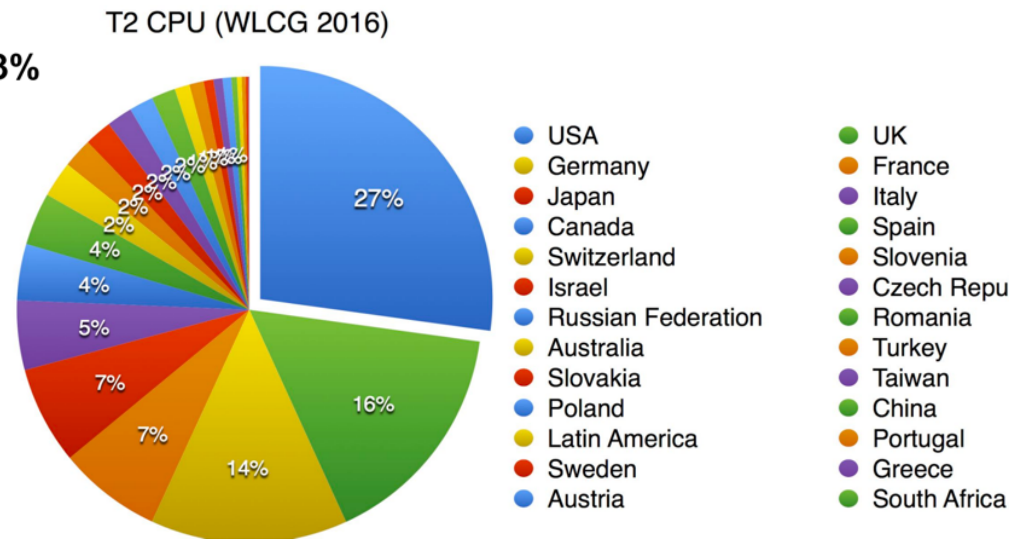
Global Computing for Science



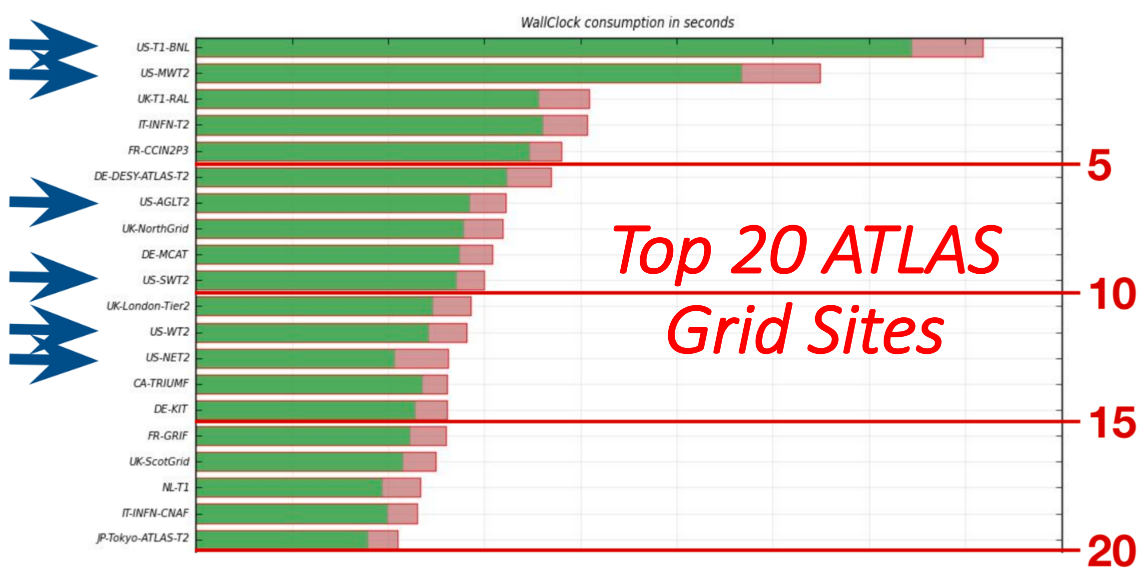
US ATLAS Cloud



US "Fair share" = 23%

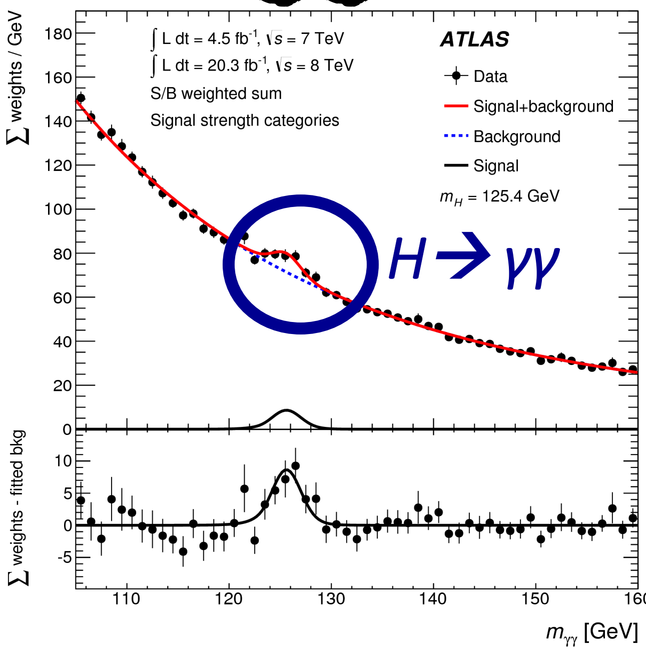


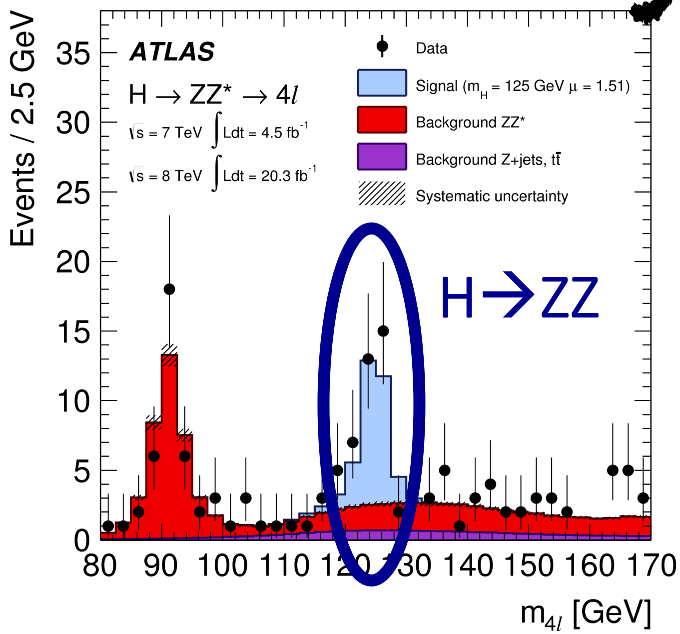
US facilities continue to be the lead contributor to ATLAS Computing

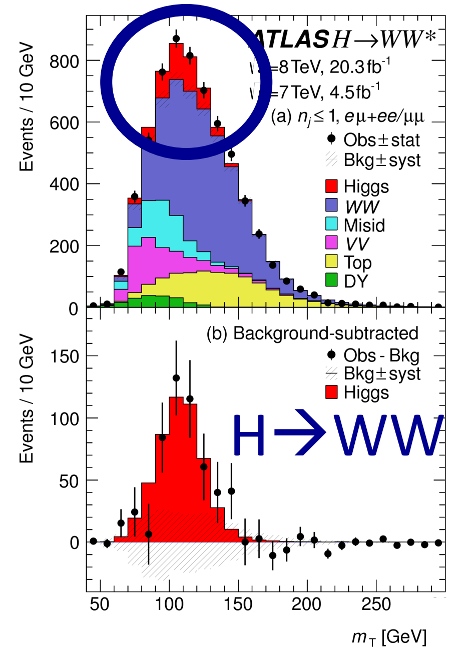


Its all about the Science

Higgs Boson Discovery! (2012)







2013 Nobel prize in Physics to Peter Higgs and Francois Englert

A new era in particle physics. The discovery of a Higgs boson with mass 125 GeV opens up a new window to search for beyond-the-SM physics

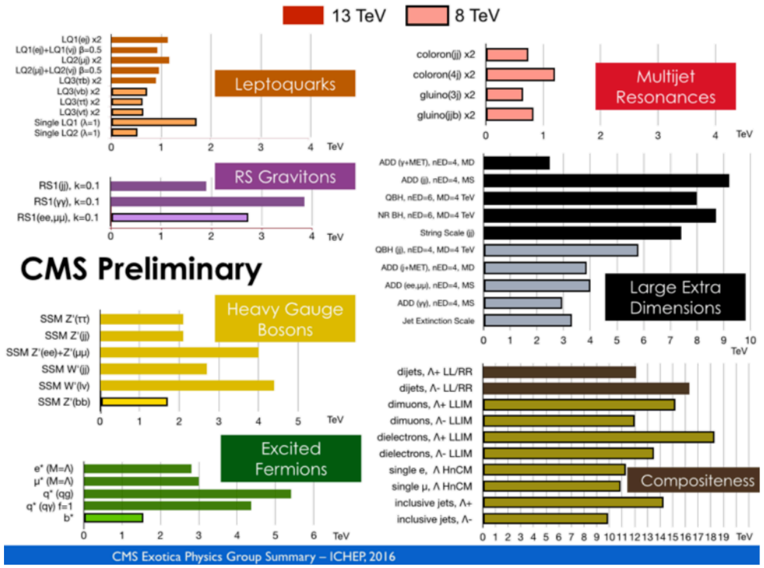
Its all about the Science

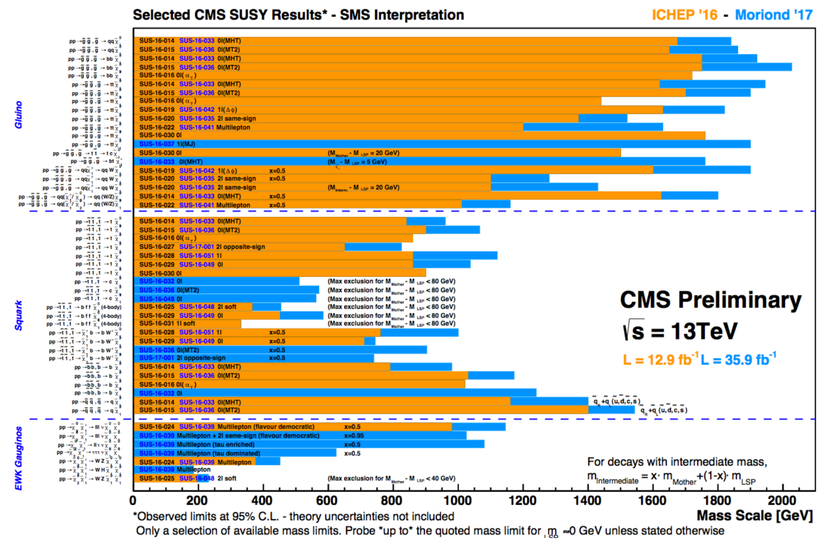
Needle in a Haystack of Needles

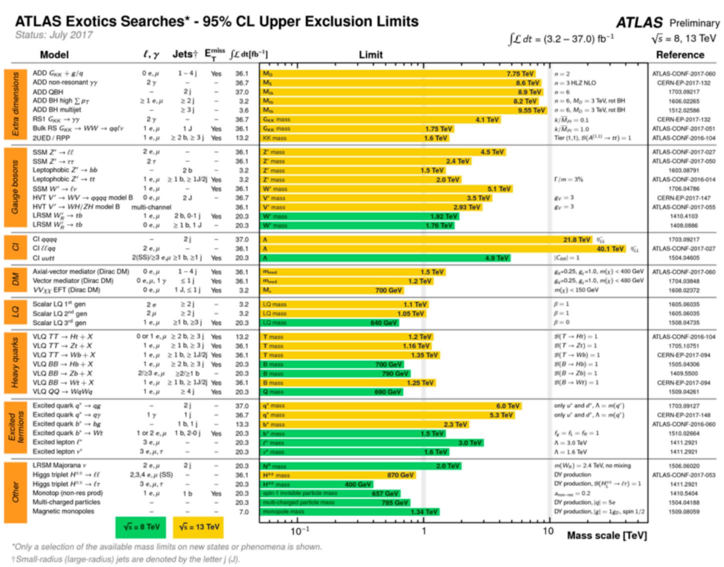
- The Higgs boson discovery was based on analysis of 1 quadrillion proton-proton collisions!
  - 2 million Higgs bosons produced ($7000/Higgs)

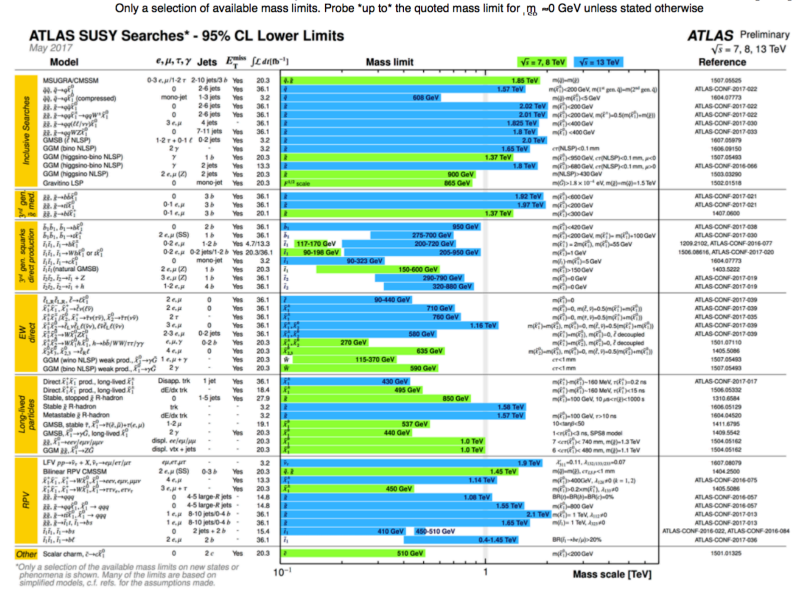
- The vast majority look virtually the same as less interesting processes
  - Only a few really stand out (e.g. $H \to ZZ \to \mu\mu\mu\mu$)

Its all about the Science!





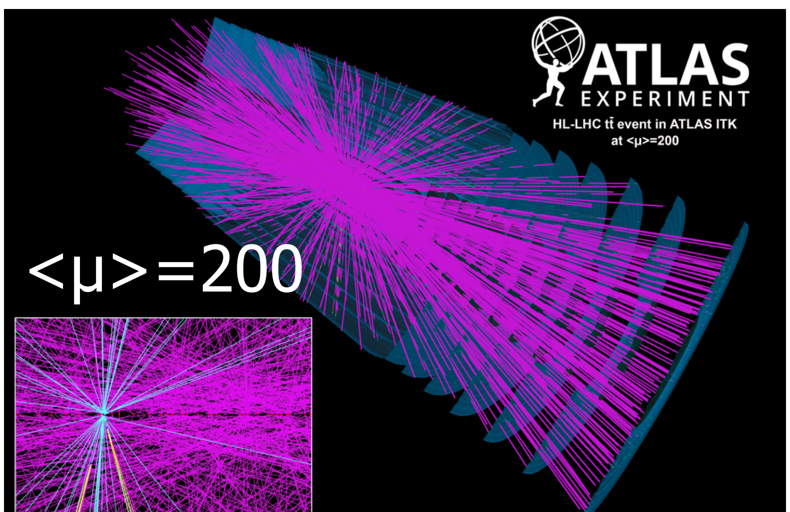


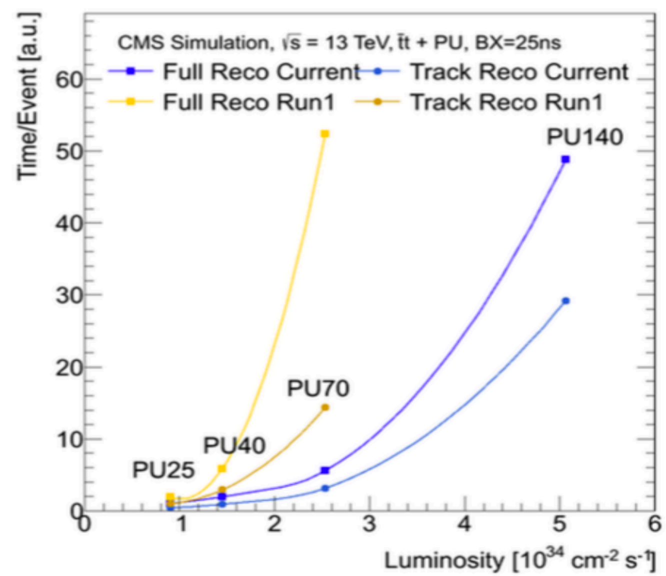


Future LHC Computing Challenges

Resource (CPU/Storage) Wall

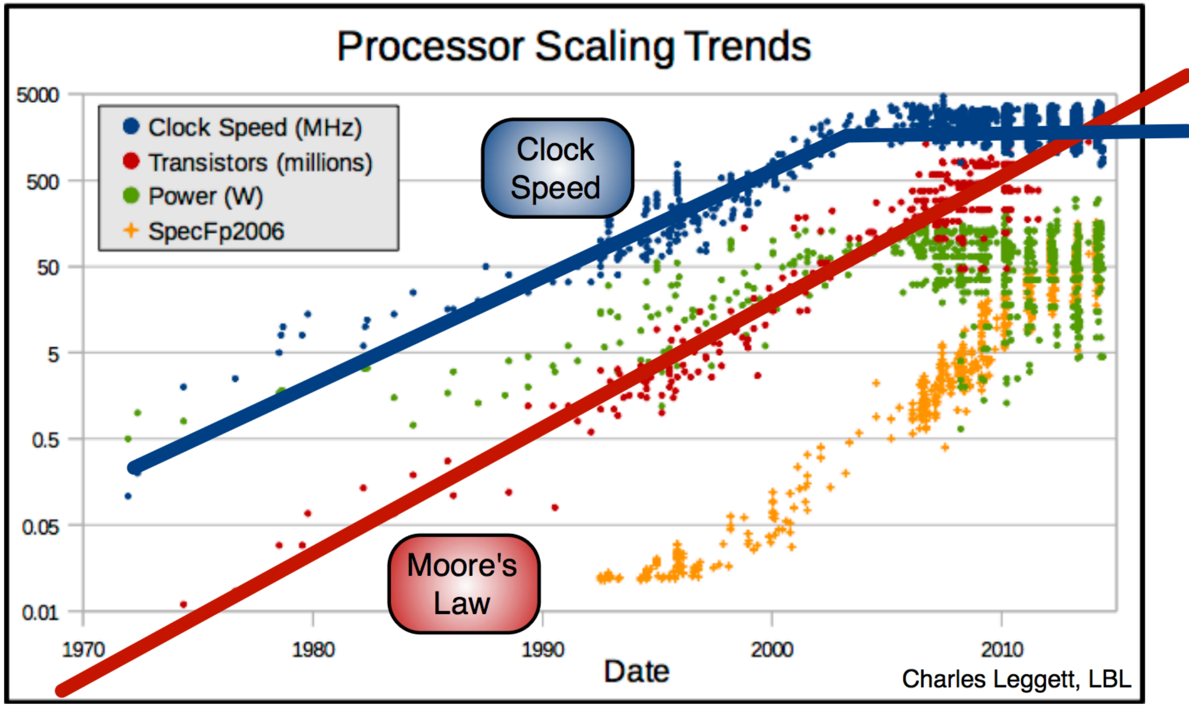
- Shortfall of CPU and storage (disk & tape)
  - True for Run 3, but the real trouble comes in Run 4 (HL-LHC) where the projected needs are ~×10 larger than what is realistic from funding levels (e.g. flat) and gains from hardware technology (~20% / year)
    - ➢ Raw data volume increases exponentially and with it processing and analysis load





CPU Processor Evolution

- Moore's law continues to deliver increases in transistor density
  - Doubling time is lengthening
  - IBM recently demonstrated 5nm wafer fabrication

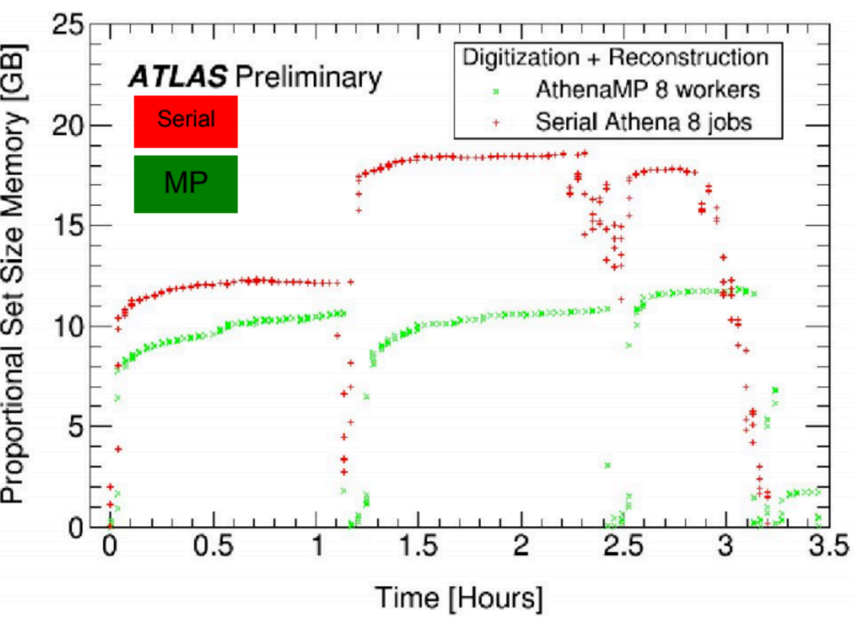


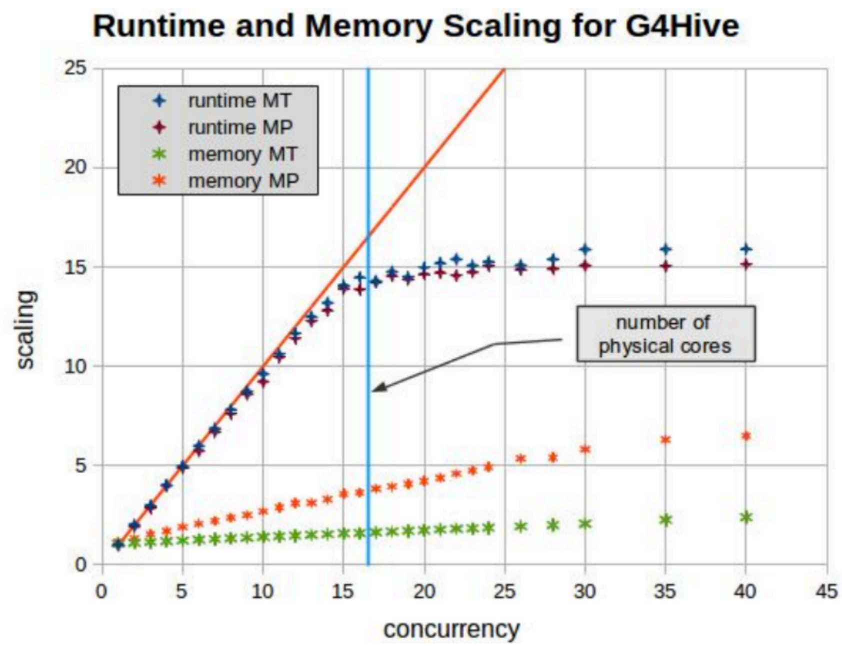
- Clock-speed scaling crashed around 2006
  - No longer able to ramp the clock speed as process size shrinks
  - Leakage currents become an important source of power consumption
  - Basically stuck at 3 GHz from the underlying Wm^{-2} limit ("power wall")

G. Stewart @ EPS-HEP 2017

Memory Wall

- Memory consumption is a major issue in LHC
  - Complicated detectors with millions of channels, large field maps, precision geometry and material maps
  - In many-core architectures, memory/core is at a premium
    - ➢ Early on, just ran multiple independent jobs on the servers
    - ➢ Multi-processing → Multi-threading





Adapted from S. Campana @ CHEP 2016

Development Wall

- Advances in hardware technologies alone will not get us to where we need to get to for HL-LHC
- We will need a serious, science-driven campaign of software R&D over the next 5 years to ready to exploit the physics from the HL-LHC running. This will require:
  - new ideas and new approaches
  - additional funding and additional people
  - a dedication to software sustainability through the lifetime of HL-LHC
- It is a challenge to get postdocs and students interested, trained and productive in taking on challenging S&C projects
  - This touches on many issues, including professional development in the field and recognition for key software contributions to physics results

Opportunities

(along with associated challenges)

Sound familiar? Too soon?

There are two ways of constructing a software design: One way is to make it so simple that there are obviously no deficiencies, and the other way is to make it so complicated that there are no obvious deficiencies. The first method is far more difficult.

- C.A.R. Hoare

HEP Software Foundation

- The ***HEP Software Foundation*** (***HSF***) was founded in 2015 as a means for organizing our community to address the software challenges of future projects like the HL-HLC

- The HSF has the following objectives:
  - Catalyze new common projects
  - Promote commonality and collaboration in new developments to make the most of limited resources
  - Provide a framework for attracting effort and support to S&C common projects (and new resources!)
  - Provide a structure to set priorities and goals for the work

- The HSF is a ***HEP community effort***, open enough to form the basis for collaboration with other sciences

Community Building and Roadmap

- DOE/HEP: Snowmass P5 (computing) and HEP-FCE reports, followed up by the HEP-CCE Initiative
- NSF: S2I2-HEP Conceptualization Project (awarded 2016)
  - Conceptualization of a ***Scientific Software Innovation Institute*** (S^2I^2) where U.S. university-based researchers can play an important role in key software infrastructure efforts that will complement those led by U.S. national laboratory-based researchers and international collaborators
  - PIs: Elmer (Princeton/CMS), Neubauer (UIUC/ATLAS), Sokoloff (Cincinnati/LHCb)
  - Kick-off meeting in Dec 2016 at University of Illinois / NCSA
- HSF Community White Paper (CWP)
  - A process by which a roadmap document in the form of a ***Community White Paper*** (CWP) is produced which aims to broadly identify the elements of computing infrastructure and software R&D required to realize the full scientific potential during the HL-LHC era
  - Charged by WLCG, viewed by NSF as a roadmap for HL-LHC computing
  - Kick-off meeting in Jan 2017 at UCSD

HSF Community White Paper

Areas of focus were identified, which formed the basis for CWP Working groups (active shown):

- Software Trigger and Event Reconstruction
- Machine Learning ← See talk by Prasanna Balaprakash
- Data Access, Organization and Management
- Software Development, Deployment and Validation/Verification
- Data Analysis and Interpretation
- Conditions Database
- Simulation
- Data and Software Preservation ← See talk by Lukas Heinrich
- Event Processing Frameworks
- Physics Generators
- Workflow and Resource Management
- Visualization
- Computing Models, Facilities, and Distributed Computing

Charge to the CWP WGs

Each CWP WG should identify and prioritize the software investments in R&D required to:

1) achieve improvements in software efficiency, scalability and performance and to make use of the advances in CPU, storage and network technologies

2) enable new approaches to computing and software that could radically extend the physics reach of the detectors

3) ensure the long term sustainability of the software through the lifetime of the HL-LHC

Practical questions to CWP WGs

- How will the proposed activities empower HEP physicists to get the most physics out of the experiments during the HL-LHC era? What new physics capabilities might these bring?
- What are the proposed R&D activities over the next 5 years toward these applications? How will the software be deployed by the experiments and sustained for the duration of the HL-LHC?
- What are the primary future applications you see in this WG area? What is the likely impact that the techniques and applications will have on overcoming the challenges of the HL-LHC era?
- Are there risks associated with going in the direction of each of the proposed ideas/R&D? What are the associated costs and is the development and implementation realistic in this regard?

Heterogenous Resources

- In order to close the resource gap, we will need to utilize all resources at our disposal
  - Great progress using HPCs! (mostly for simulation)
    - Event granularily (Event Service) very important here
    - Need multi-year reliability of allocations
  - Commercial & Institutional Clouds and Clusters
  - Modern and evolving architectures (GPUs, FPGAs..)
- Key challenges include
  - Making these heterogeneous resources look not so
  - Accommodating changes on resources we do not control (technical and financial)

Machine Learning (ML)

- Offers great promise to re-think nearly every approach to our experimental program
- ML WG has been very active and many ideas have been put into the WP
  - Particle identification, Event classification, Simulation GANs, for sustainable MEM, …
- Key challenges include
  - How does this change our required computing resources?
  - How can we build bridges to industry-standard tools?
  - How can we efficiently collaborate with CS? Industry?

Some Other CWP WG Highlights

(just a sampling of the excellent CWP WG efforts!)

- **Data and Software Preservation**
  - Increasing focus on re-usability of analysis software
  - Approaches to make analysis preservable with little or no effort on the part of users
- **Data Analysis and Interpretation**
  - Leveraging industry-standard tools in Data Science
    - Raise profile and support for python and other language for anlaysis
    - E.g. thinking about how optimize DS-standard data formats for HEP workflows
  - Declarative languages and and query-based analysis
- **Visualization**
  - Leverage industry software for improved rendering
  - Develop visualization tools that are modular and are
    - Client-server (e.g. browser) based
    - Distributed, collaborative and immersive

CWP Status and Plans

- The process has been open and inclusive, involving computing coordinators for the experiments, workshops and engaging with many outside of LHC experiments (e.g. IF, industry, theory)
- Each WG is producing a document (WP), eventually to be posted on the archive (arXiv). Vast majority are ~complete
- There will be a single summary document drawn from each WG's WP. This will be the CWP. It will be posted to arXiv.
  - The Editorial Board for the CWP is being formed
  - An opportunity will be made for those not active in the process to sign-on to the CWP
- Our hope was to have the WG WPs final by the end of July, with the CWP near-final at the end of August. With many people away on vacation, this is slipping a bit (but not much)

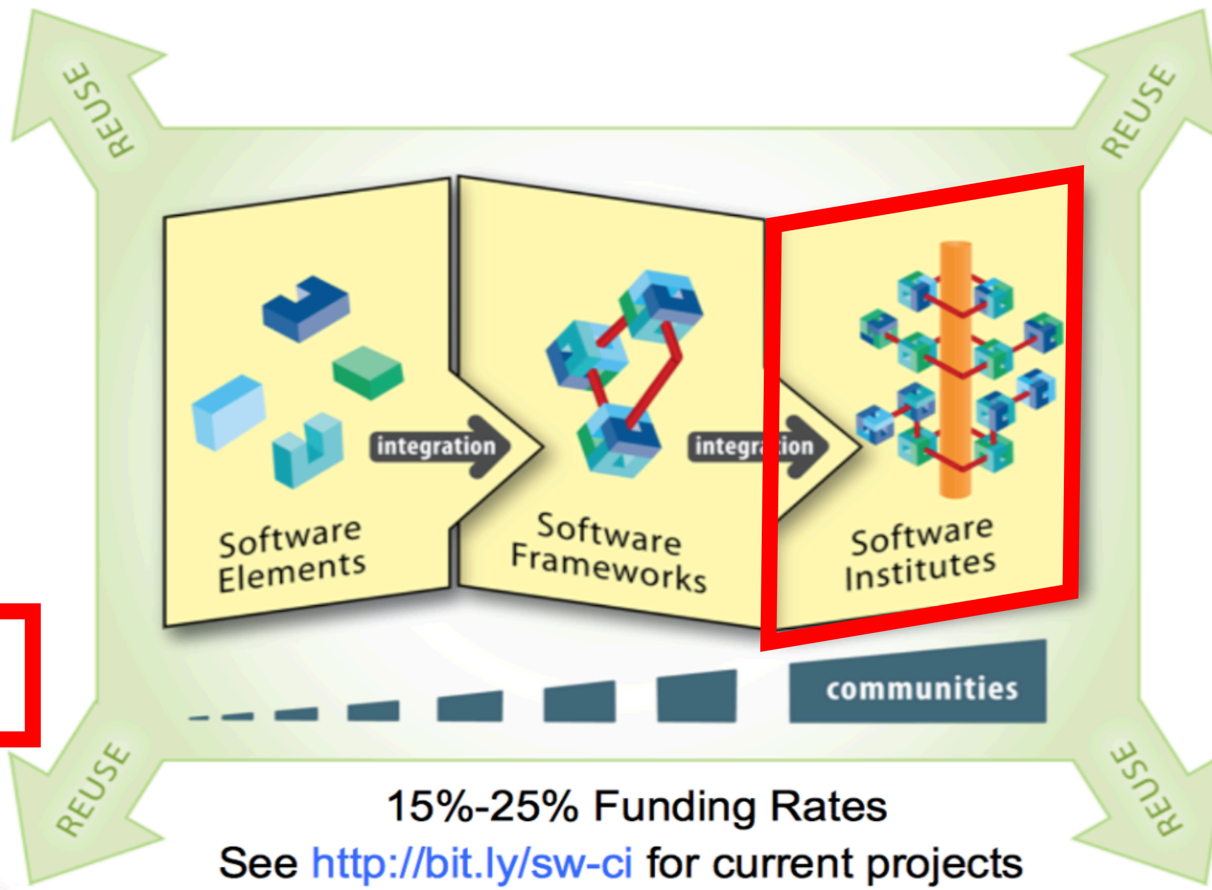
S2I2-HEP

ACI Flagship - Software Infrastructure for Sustained Innovation (SI2)

Elements: $500K/3 years

Frameworks: $1M/ year 3-5 years

Institutes: $3-$5m/ year 5-10 years



15%-25% Funding Rates

See http://bit.ly/sw-ci for current projects

S2I2-HEP Status

- We submitted a conceptualization proposal to NSF last August 2015: "***Conceptualization of an S2I2 Institute for High Energy Physics***" (link). http://s2i2-hep.org/
  - Awarded ***July 2016***
- An important element is to foster partnerships between HEP and Computer Scientists on key topics
- The main deliverable for the Conceptualization is a *Strategic Plan (SP)* for the Institute, along with the CWP
  - The NSF asked for a nearly complete version by the end of October (2017)
- The SP is informed by the outcomes of the CWP. Initial focus areas will be a subset of the proposed R&D from the CWP, considering US-specific elements (interests, expertise, etc), priorities, synergies, commonality, etc.

S2I2-HEP @ ACAT 2017

There will be an S2I2-HEP Workshop in Seattle in conjunction with ACAT2017

- 23 Aug Wednesday (afternoon) - Strategic Areas of Focus and Priorities (Discussion)
- 24 Aug Thursday (afternoon) - Institute Organization and Processes (Discussion), in parallel to ACAT 2017 parallel sessions
- 25 Aug Friday (afternoon) – SP Writing and Review Sessions
- 26 Aug Saturday (afternoon) – SP Writing and Review Sessions

Our objective will be to come away at the end with an early rough draft of text for the S2I2-HEP Strategic Plan describing what the U.S. university community could do with such an Institute to meet the challenges of the HL-LHC era.

Some Parting thoughts

- There are real ***challenges*** for S&C during the HL-LHC era. New approaches and resources needed!
- We should also think about novel approaches that bring new ***opportunities*** to extend our physics reach
- We should not assume that the mixture of physics analysis we are doing now will be the same for HL-LHC
- Guiding principles going forward should be:
  - Projects should be collaborative across experiments from day one
    - We should ask: "*What makes experiment X special as compared to experiment Y*?"
  - We should better engage with industry and CS, and be willing to listen
    - We should ask: "*What make HEP special as compared with CS/Industry/DomainX*?"
- ***We are making huge investments in detector upgrades for HL-LHC. We need robust R&D for a S&C upgrade to realize the full physics potential from the HL-LHC***