Snowmass On the Mississippi CSS2013

Summary from the Computing Frontier Study Group

L.A.T. Bauerdick, S.Gottlieb,
for the Computing Frontier Group

Outline

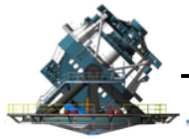
- Introduction
  - how we worked and where we stand in the process
- Computational Challenges
  - and how we address them at all frontiers
- Data Management Challenges
- Networking Challenges
- Technology Developments
- Software, Training, Careers
- Some Common Themes and Conclusions

Summer Study

- Subgroups for “user needs”
  - Each subgroup to interacted with the corresponding physics frontiers to assess the computing needs
- Subgroups for “infrastructure”
  - The infrastructure groups project computing capabilities into the future and see how the user needs map onto the trends
- The main result is a written report from each of the subgroups, and a summary
  - Heard about a DOE meeting in December to

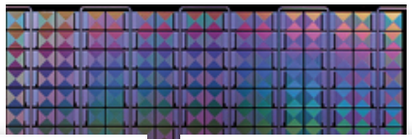
Computing Challenges at the Physics Frontiers

Cosmic Frontier

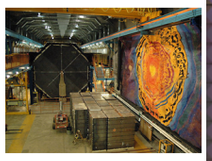


Technology developments

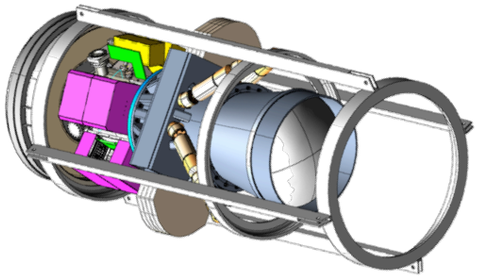
- Microwave Kinetic Inductance Detectors (MKIDs)
  - Energy resolving detectors (extended to optical and UV)
  - Resolving power: 30 < R < 150 (~5 nm resolution)
  - Coverage: 350nm – 1.3 microns
  - Count rate: few thousand counts/s
  - 32 spectral elements for uv/optical/ir photons

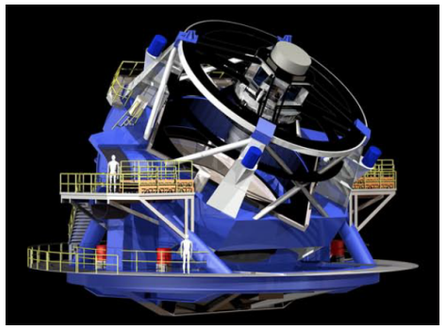


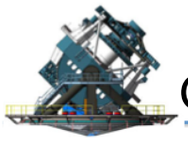
A decade of data: DES to LSST



- Wide field and deep
  - DES: 5,000 sq degrees
  - LSST: 20,000 sq degrees
- Broad range of science
  - Dark energy, dark matter
  - Transient universe
- Timeline and data
  - 2012-16 (DES)
  - 2020 – 2030 (LSST)
  - 100TB - 1PB (DES)
  - 10PB - 100 PB (LSST)

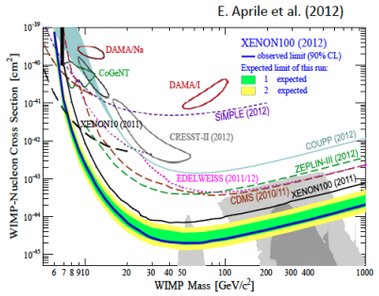


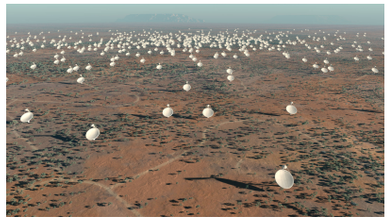




Growing volumes and complexity

- CMB and radio cosmology
  - CMB-S4 experiment's 10^{15} samples (late-2020's)
  - Murchison Wide-Field array (2013-)
    - 15.8 GB/s processed to 400 MB/s
  - Square Kilometer Array (2020+)
    - PB/s to correlators to synthesize images
    - 300-1500 PB per year storage
- Direct dark matter detection
  - Order of magnitude larger detectors
  - G2 experiments will grow to PB in size





Energy Frontier

- EF will go to very high trigger rates and more complicated events
  - we looked back 10 years to aid prediction prediction of the magnitude of changes expected from programs over 10 years
- programs suggested for EF all have the potential for another factor of 10 in trigger and 10 in complexity
  - Simulation and reconstruction might continue to scale with Moore's law as they did for LHC, but could just as easily increase much faster
  - LHC adds 25k processor cores and 34 PB a year —in 10 yrs at this rate (flat budget) the capacity would be up by 4x - 5x
- Need make better use of resources as the technology changes

	Tevatron	LHC
Trigger	50Hz	ATLAS 500Hz CMS 350Hz LHCb 2kHz
RAW Event Size	150kB	ATLAS 1.5MB CMS 0.5MB
RECO Event Size	150kB	ATLAS 2MB CMS 1MB
Reco Speed	1-2 seconds on CPU of the time	10 seconds on CPU of the time

Intensity Frontier

HEP Intensity Frontier Experiments

Outside US Taking data

List from DOE: ⟹ There are MANY

Experiment	Location	Status	Description	#US US participation	IL
Belle II *	KEK, Tsukuba, Japan	Physics run 2016	Heavy flavor physics, CP asymmetries, new matter states	10 Univ, 1 Lab	55
BES III *	IHEP, Beijing, China	Running	Precision measurements charm, charmonium, tau; search for and study new states of hadronic matter	6 Univ	26
CAPTAIN *	Los Alamos, NM, USA	R&D; Test run 2015	Cryogenic apparatus for precision tests of argon interactions with neutrinos	5 Univ, 1 Lab	20
Daya Bay	Dapeng Penisula, China	Running	Precise determination of θ_{13}	13 Univ, 2 Lab	76
Heavy Photon Search *	Jefferson Lab, Newport News, VA, USA	Physics run 2015	Search for massive vector gauge bosons which may be evidence of dark matter or explain g-2 anomaly	8 Univ, 2 Lab	47
KOTO *	J-PARC, Tokai , Japan	Running	Discover and measure $K_L \to \pi^0 \nu\nu$ to search for CP violation	3 Univ	12
LArIAT	Fermilab, Batavia, IL	R&D; Phase I 2013	LArTPC in a testbeam; develop particle ID & reconstruction	11 Univ, 3 Lab	38
LBNE	Fermilab, Batavia, IL & Homestake Mine, SD, USA	CD1 Dec 2012; First data 2023	Discover and characterize CP violation in the neutrino sector; comprehensive program to measure neutrino oscillations	48 Univ, 6 Lab	336
MicroBooNE	Fermilab, Batavia, IL, USA	Physics run 2014	Address MiniBooNE low energy excess; measure neutrino cross sections in LArTPC	15 Univ, 2 Lab	101
MINERvA	Fermilab, Batavia, IL, USA	Med. Energy Run 2013	Precise measurements of neutrino-nuclear effects and cross sections at 2-20 GeV	13 Univ, 1 Lab	48
MINOS+	Fermilab, Batavia, IL & Soudain Mine, MN, USA	NuMI start-up 2013	Search for sterile neutrinos, non-standard interactions and exotic phenomena	15 Univ, 3 Lab	53
Mu2e	Fermilab, Batavia, IL, USA	First data 2019	Charged lepton flavor violation search for $\mu N \to eN$	15 Univ, 4 Lab	106
Muon g-2	Fermilab, Batavia, IL, USA	First data 2016	Definitively measure muon anomalous magnetic moment	13 Univ, 3 Lab, 1 SBIR	75
NOvA	Fermilab, Batavia, IL & Ash River, MN, USA	Physics run 2014	Measure ν_μ-ν_e and ν_μ-ν_μ oscillations; resolve the neutrino mass hierarchy; first information about value of δ_{cp} (with T2K)	18 Univ, 2 Lab	114
ORKA	Fermilab, Batavia, IL, USA	R&D; CD0 2017+	Precision measurement of $K^+ \to \pi^+ \nu\nu$ to search for new physics	6 Univ, 2 Lab	26
Super-K	Mozumi Mine, Gifu, Japan	Running	Long-baseline neutrino oscillation with T2K, nucleon decay, supernova neutrinos, atmospheric neutrinos	7 Univ	29
T2K	J-PARC, Tokai & Mozumi Mine, Gifu, Japan	Running; Linac upgrade 2014	Measure ν_μ-ν_e and ν_μ-ν_μ oscillations; resolve the neutrino mass hierarchy; first information about value of δ_{cp} (with NOvA)	10 Univ	70
US-NA61 *	CERN, Geneva, Switzerland	Target runs 2014-15	Measure hadron production cross sections crucial for neutrino beam flux estimations needed for NOvA, LBNE	4 Univ, 1 Lab	15
US Short-Baseline Reactor	Site(s) TBD	R&D; First data 2016	Short-baseline sterile neutrino oscillation search	6 Univ, 5 Lab	28

CSS2013: Intensity Frontier Intro; July 29, 2013, H.Weerts * not explicitly surveyed 7

The Method

- We wanted to **qualitative survey the community of current and future experiments in the IF** in order to understand the computing needs but also the foreseen evolution of said needs.
- Computing liaisons and representatives for the LBNE, MicroBooNE, MINERvA, MINOS+, mu2e, g-2, NOvA, Daya Bay, IceCube, SNO+, SK, T2K, SEAQUEST collaborations all responded to the survey and provided input.
- This does not cover all experiments in all areas but we consider it a representative survey of the field.
- More input is of course welcome. Please **see/email/chat Brian Rebel and myself** over these days or the next few weeks.

We want to thank the people that took the time to give well thought answers this survey

Computing Model

- We found a **high degree of commonality among** the various experiments' **computing models** despite large differences in type of data analyzed, the scale of processing, or the specific workflows followed.
- The model is summarized as a traditional **event driven analysis and Monte Carlo simulation using centralized data storage** that are **distributed to independent analysis jobs running in parallel on grid computing clusters**. Peak usage can be 10x than planned usage.
- **For large computing facilities** such a Fermilab, it is useful to **design a set of scalable solutions** corresponding to these patterns, with associated toolkits that would allow access and monitoring. Provisioning an experiment or changing a computing model would then correspond to adjusting the scales in the appropriate processing units.
- **Computing should be made transparent to the user**, such that non-experts can perform any reasonable portion of the data handling and simulation. Moreover, all experiments would like to see **computing become more distributed across sites**. Users without a home lab or large institution require **equal access to dedicated resources**.

- ✦ Diversity of experiments
- ✦ Survey still live: click here
- ✦ convergence on a ~common computing model

Accelerator Science

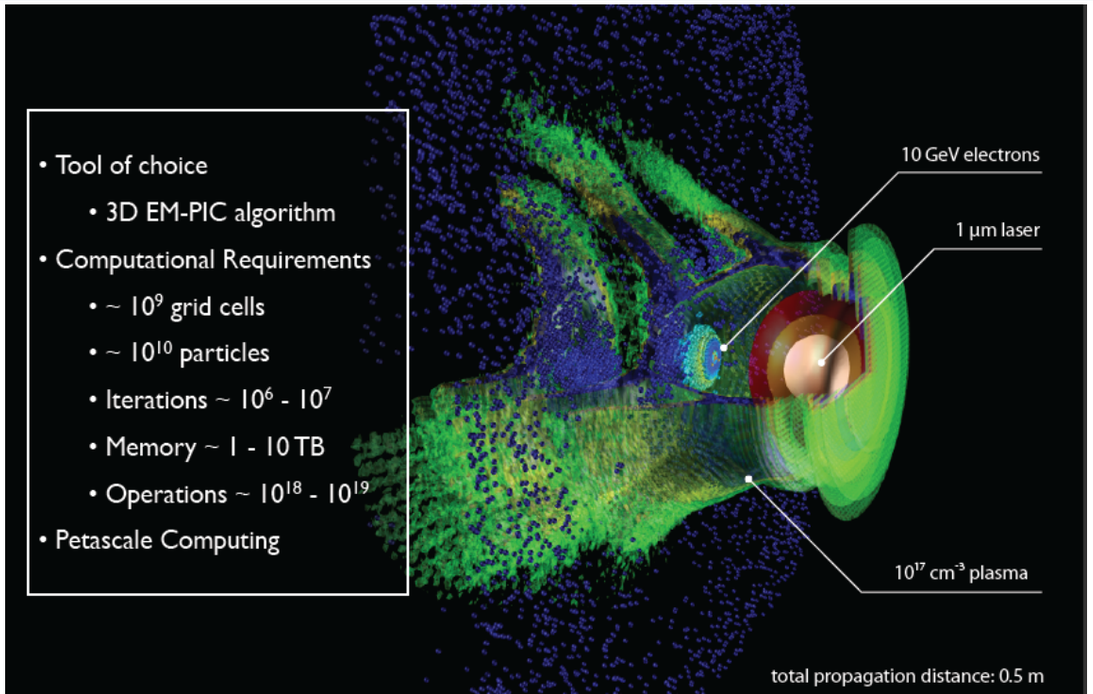
Accelerator Science Drivers

- New techniques and technologies
  - Optimize, evolve concepts, design accelerator facilities based on new concepts
- Maximize performance of "conventional" techniques and technologies
  - Optimize operational parameters, understand dynamics (manipulation and control of beams in full 6D phase space)
- Desirable outcome: achieve higher gradients for energy frontier applications, minimize losses for intensity frontier applications

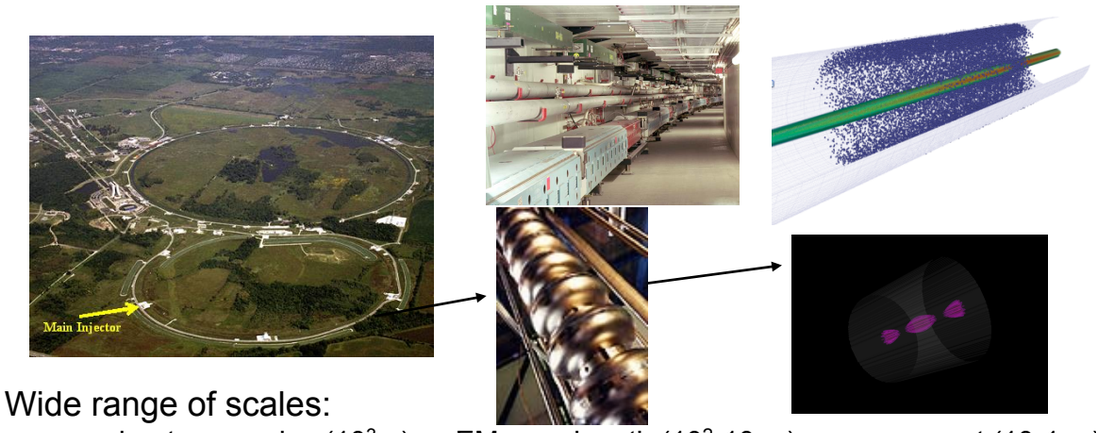
Summary of Requirements

- Intensity Frontier accelerator needs
  - beam loss characterization and control
  - control room feedback : need fast turnaround of simulations
  - direct comparison between beam diagnostics detectors and simulated data (need development of new tools)
- Energy Frontier accelerator needs
  - beam stability characterization
  - ability to produce end-to-end designs
  - control room feedback: fast turnaround and fast analysis of massive data (requires development of new, automated, analysis tools)
  - new physics model capabilities (e.g., radiation and scattering)
  - better numerical models (less numerical noise) - identification of most suited numerical techniques
- All Frontiers
  - integrated / multi physics modeling: improve algorithms (integrate more physics in the model), massive computing resources
  - common interfaces, geometry and parameter description, job submission, from tablet to supercomputer.

Example:LWFA multi-scale physics



Example: High-Intensity Proton Drivers



- Wide range of scales:
  - accelerator complex (10^3m) → EM wavelength (10^2-10 m) → component (10-1 m) → particle bunch (10^{-3} m)
  - Need to correctly model intensity dependent effects and the accelerator lattice elements (fields, apertures), to identify and mitigate potential problems due to instabilities that increase beam loss; thousands of elements, millions of revolutions
  - **Calculating 1e-5 losses at 1% requires modeling 1e9 particles, interacting with each-other and the structures around them at every step of the simulation**

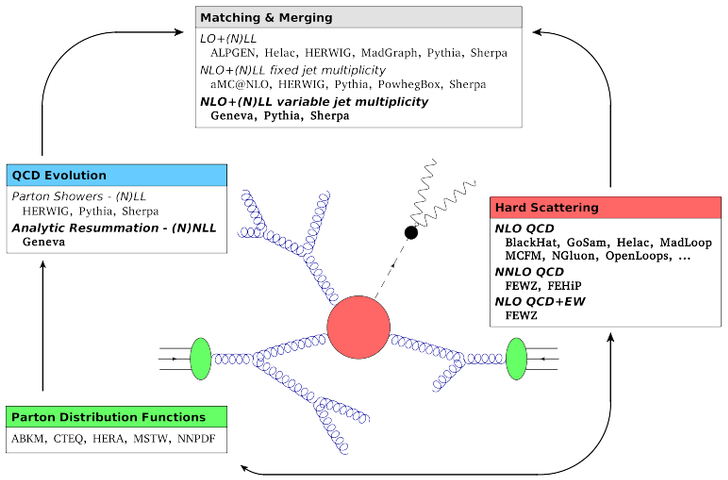
Lattice Field Theory

Summary bullet 1: Scientific goals

- The scientific impact of many future experimental measurements at the energy and intensity frontiers hinge on reliable Standard-Model predictions on the same time scale as the experiments and with commensurate uncertainties. Many of these predictions require nonperturbative hadronic matrix elements that can only be computed numerically with lattice-QCD. The U.S. lattice-QCD community is well-versed in the plans and needs of the experimental high-energy program over the next decade, and will continue to pursue the necessary supporting theoretical calculations. Some of the highest priorities are improving calculations of hadronic matrix elements involving quark-flavor-changing transitions which are needed to interpret rare kaon decay experiments, improving calculations of the quark masses m_c and m_b and the strong coupling α_s which contribute significant parametric uncertainties to Higgs branching fractions, calculating the nucleon axial form factor which is needed to improve determinations of neutrino-nucleon cross sections relevant experiments such as LBNE, calculating the light- and strange-quark contents of nucleon which are needed to make model predictions for the $\mu \to e$ conversion rate at the Mu2e experiment (as well as to interpret dark-matter detection experiments in which the dark-matter particle scatters off a nucleus), and calculating the hadronic light-by-light contribution to muon $g-2$ which is needed to solidify and improve the Standard-Model prediction and interpret the upcoming measurement as a search for new physics.

Perturbative QCD

Broad impact of Perturbative QCD on collider physics



- interpreting LHC data requires accurate theoretical predictions
- complex SM backgrounds call for sophisticated calculational tools
- higher order QCD(+EW) corrections mandatory

This effort could greatly benefit from:

- unified environment for calculations/data exchange
- adequate computational means to provide accurate th
  pace and in a format useful to experimental analyses
- extensive computational resources to explore new tech

As pQCD component of the Computing Frontier we have set:

- **Short term goals**
  - provide collider experiments with state-of-the-art theoretical predictions;
  - make this process automated/fast/efficient;
  - facilitate progress of new ideas and techniques for cutting-edge calculations (NLO with high multiplicity; NNLO).
- **Long term goals**
  - take advantage of new large-scale computing facilities and existing computer-science knowledge;
  - work in closer contact with computing community to benefit from pioneering new ideas (GPU, Intel Phi, programmable networks, ...).

We have explored available options and provided some proofs of concept

U.S. Computational and Data Infrastructure for HEP

U.S. HEP relies on a well-developed Compute and Data Infrastructure

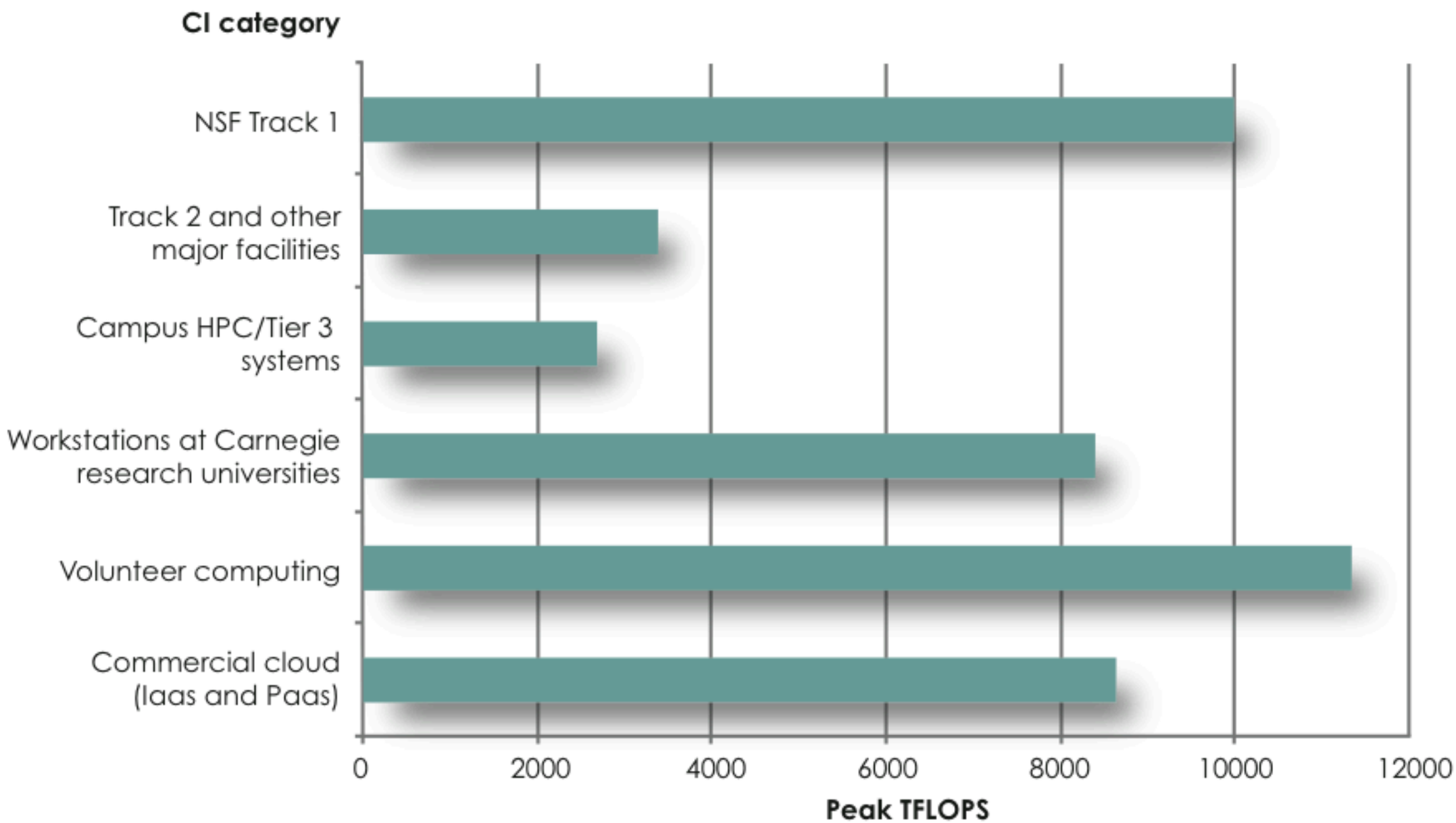
- National HPC Resources
  - a planned cyber-infrastructure, based on “program needs”
  - planned for, funded and built around **High Performance Computing** Centers to provide computing and storage resources
    - the NSF XD program, DOE Leadership-class facilities
  - provide the “glue” across institutions (e.g. user accounting)
    - for XD program: the XSEDE project
  - run an allocation process to satisfy computing needs of PIs
    - example: XD XRAC process
    - DOE allocations, USQCD, ...

- Computing Grids
  - independently funded computing resources for science projects like LHC, HEP groups, labs, campuses
  - Open Science Grid makes it into a grid infrastructure by forming a consortium of resource providers, science projects, campuses etc
  - OSG services provide the “glue” to enable distributed **High Throughput Computing** across sites
  - enable sharing of resources across stakeholders (VOs), and enable PIs to “opportunistically use” otherwise unused resources

Complementary Approaches!

Open Science Grid

Computing Resources Available to Science and Engineering in U.S. (3/2011)



Based on: Welch, V.; Sheppard, R.; Lingwall, M.J.; Stewart, C. A. 2011. Current structure and past history of US cyberinfrastructure (data set and figures). hdl.handle.net/2022/13136

Survey of Who Provides Resources that Researchers Utilize

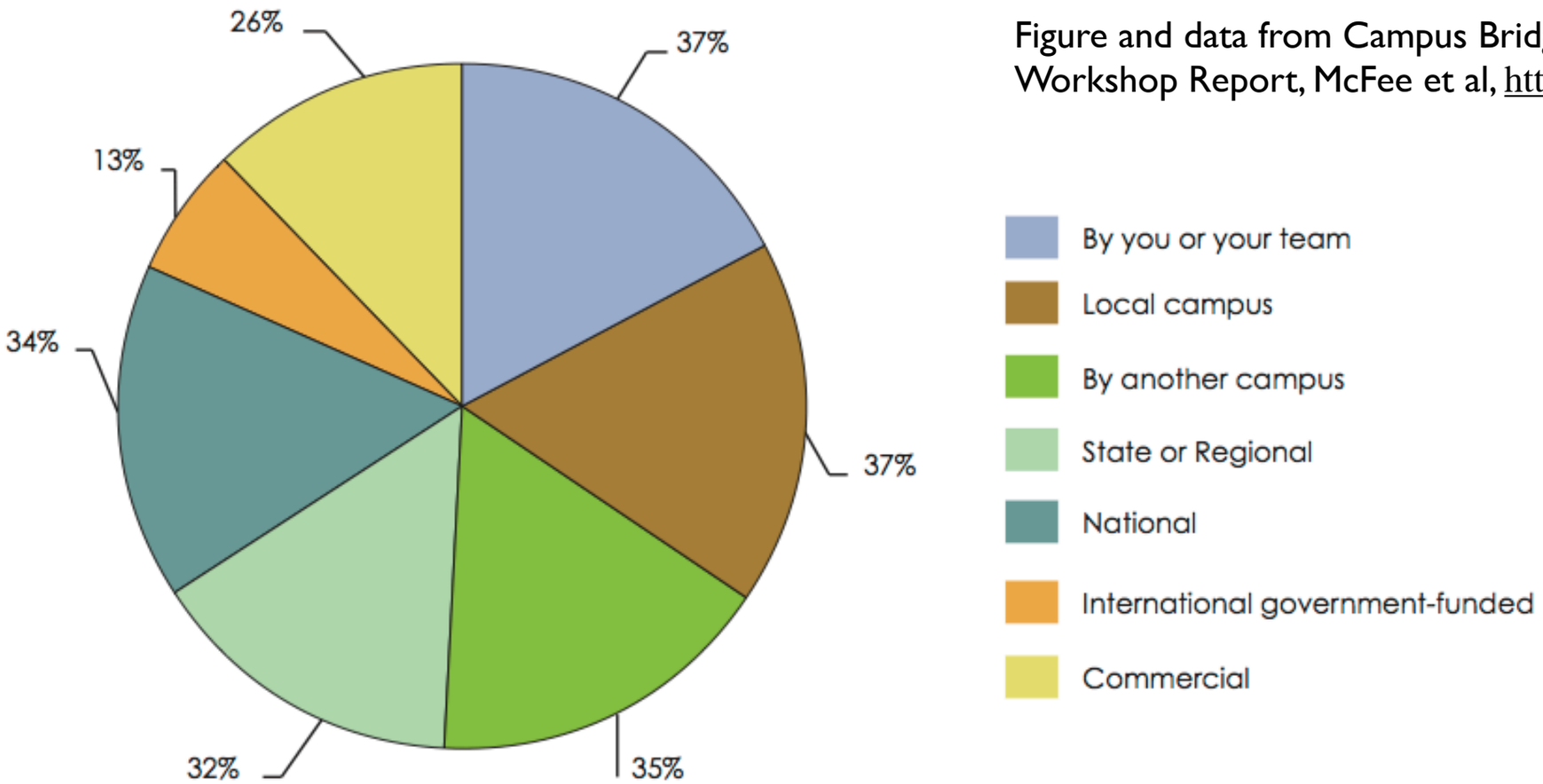
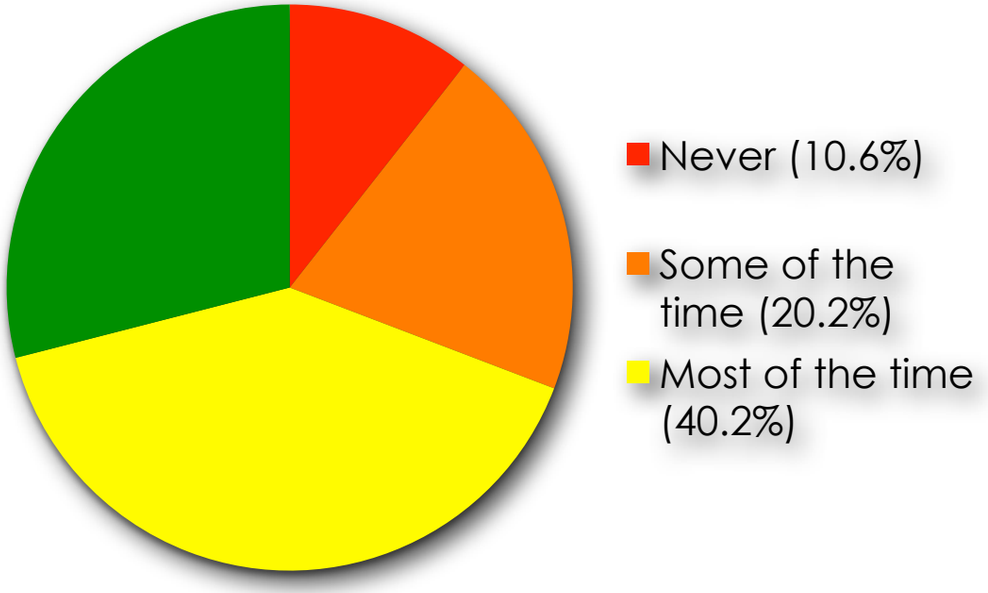


Figure and data from Campus Bridging: Software and Service Issues Workshop Report, McFee et al, http://hdl.handle.net/2022/13070



★ Responses to asking if researchers had sufficient access to cyber infrastructure resources

C.A. Stewart et al: Survey of cyberinfrastructure needs and interests of NSF-funded principal investigators. 2011. hdl.handle.net/2022/9917

Computational Needs: Cosmic Frontier

- ★ Computational resources will have to grow to match the largely increasing associated data rates (makes for data intensive compute needs!)
- ★ Require new computational models for distributed computing (including many-core systems)
- ★ Infrastructure for data analytics applicable to large and small scale experiments will need to grow over the next decade (with an emphasis on sustainable software, not just build-you-own)
  - ✦ Data Archiving and Serving: data archives, databases, and facilities for post-analysis are becoming a pressing concern
  - ✦ Archives now mostly used to “download”, development of powerful, easy-to-use remote analysis tools
- ★ Simulations (cosmological and instrument) will play a critical role in evaluating and interpreting the capabilities of current and planned experiments
- ★ Also need for instrument simulation
- ★ (add table?)

Computational Needs: Energy Frontier

- Driven by Trigger rate, event size, and reconstruction time
  - all have increase by a factor of ~10 over past 10 yrs, leading to processing capacity up by a factor of 30
  - The programs suggested for energy frontier all have the potential for another factor of 10 in trigger and 10 in complexity
  - Simulation and reconstruction might continue to scale with Moore's law as they did in the past, but could just as easily increase much faster
- transformative computing technologies help with these challenges
  - sharing, on-demand resource provisioning, opportunistic resources
    - smooth out peaky computing needs, improve turn-around of "campaigns"
  - commercial clouds are not (yet) competitive, but the clouds are here
    - e.g. use of HLT farms, use of Vodaphone cloud opportunistically for the LHC, etc
- more efficient code, high-performance low-power hardware (e.g.GPUs)
  - significant re-engineering effort under way, long-term program
- More selective application of high-CPU processes
  - e.g. archive high trigger rates streams and only selective reconstruction

Computational Needs: Intensity Frontier

- ★ Computational demands of IF experiments and IF R&D modest compared to those of EF experiments. However the needs are NOT insignificant.
- ★ Efficient use of available grid resources has had/could have a huge impact on IF experiments and IF R&D
- ★ All efforts will benefit from dedicated transparent access to grid resources
  - ✦ and this is a strategy for the Fermilab-based experiments and others
- ★ US participation in international IF efforts uses a combination grid resources based mainly outside of the US and smaller local clusters.
- ★ Peak usage can be 10x of the planned usage.
- ★ It was widely noted that the lack of dedicated US resources has a detrimental impact on the science. Dedicated grid resources for the intensity frontier (in the form of intensity frontier VO?) would have the largest impact on our international efforts.
- ★ Found a high degree of commonality among the various experiments' computing models despite large differences in type of data analyzed, the scale of processing, or the specific workflows followed

Computational Needs: Field Theory LQCD and pQCD

- Lattice QCD simulations require parallel machines
  - Codes are floating point intensive, limited by memory bandwidth, network latency and bandwidth, implemented using message passing (MPI)
- computationally intensive: generate gauge-field config. ensembles
  - Capability Hardware, jobs each use 10K to 100K+ processors
    - USQCD uses DOE Leadership Computing Facilities (DOE ASCR funded)
    - Also NERSC, LLNL, NSF XSEDE, Japan (RIKEN BNL), UK (UKQCD BlueGene/Q)
    - allocations e.g. in 2013 290M CPUh at ANL, 140M CPUh at ORNL, among largest at LCFs
- "Measurements" repeat a calculation with all members of an ensemble
  - Capacity Hardware (now using > 50% of Flops)
    - USQCD has dedicated LQCD systems and support personnel at BNL, FNAL, JLab (including currently 812 GPUs)
- USQCD has submitted a follow-on 5 year proposal
  - 2.1 —> 12.5 B CPUh capacity-class,
    2.8 —> 17.8B CPUh capability-class in 2015..2019
- pQCD computational needs 100k-1M hours for dedicated measurements
  - for verifications of event generation libraries for NLO and NNLO
  - otherwise computational needs contained in experiment requests

Computational Needs: Accelerator Science

- Multi-physics modeling necessary to advance accelerator science.
  - Requires frameworks to support advanced workflows
  - Efficient utilization of large computing resources, HPC in many cases
- Evolving technologies (light-weight CPU plus accelerators) require R&D and could result in significant changes
  - Advanced algorithmic research underway, will require continuing support
  - Programmatic coordination necessary to efficiently utilize resources
  - Opportunity for multi-scale, multi-physics modeling and “near-real-time” feedback, if could be used efficiently
  - Intensity frontier machines would like “control room feedback” capabilities
- Current strategy is to abstract and parameterize data structures so that are portable and enable efficient flow of data to a large number of processing units in order to maintain performance.
  - Already ported subset of solvers and PIC infrastructure on the GPU
  - Evaluate current approach, develop workflow tools and frameworks
  - Investigate new algorithms and approaches

HEP Computational Infrastructure (to be edited)

- Distributed High-Throughput (Grid) Computing
  - The Worldwide LHC Computing Grid (WLCG) and OSG in US
  - HTC workflow model is working well for experiments
  - improvements can be implemented in an evolutionary manner
- High Performance Computing
  - HPC used and required by a number of projects
  - Adapt to ongoing and future architectural changes
    - diversity of complex nodes, memory/core, communication bottlenecks, multi-level memory hierarchy, power restrictions, --
  - new programming models -- how to rewrite large code bases?
- Cloud Providers
  - Commercial clouds still too costly to replace dedicated resources
  - Clouds have the ability to quickly surge resources to address larger problems
  - Expect existing resources to move to cloud interfaces
  - Significant gaps and challenges exist in managing virtual environments, workflows, data, cyber-security, and other areas.
  - There are efforts by traditional HPC platforms to combine the flexibility of cloud models with the performance of HPC systems.

Distributed and HPC Computing in HEP:

CMS + ATLAS in 2012:

~1.5 Billion Hours

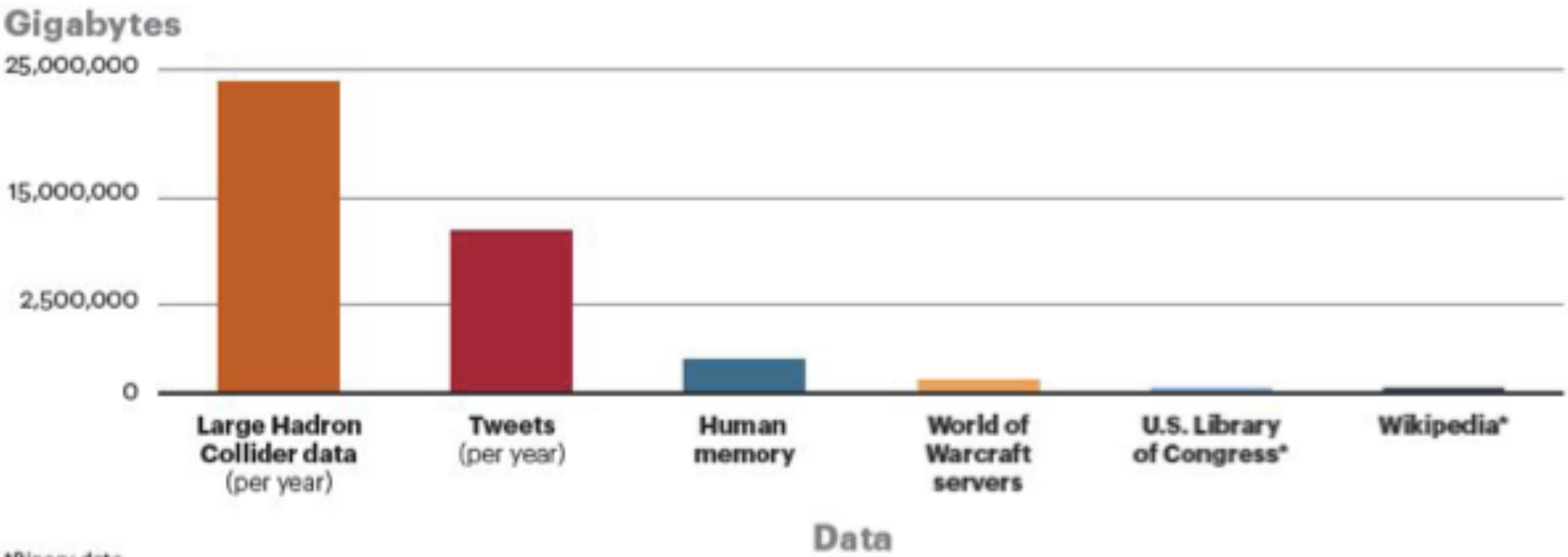
National DOE & NSF HPC Centers 2013:

~1.4 Billion Hours

The Big Data Frontier

Figure

The LHC collects about 25 million gigabytes of data per year



*Binary data

Note: All numbers are approximate.

Source: "Particle Physics Tames Big Data," Leah Hesla, *Symmetry*, 1 August 2012

Data Management: Cosmic Frontier (to be edited)

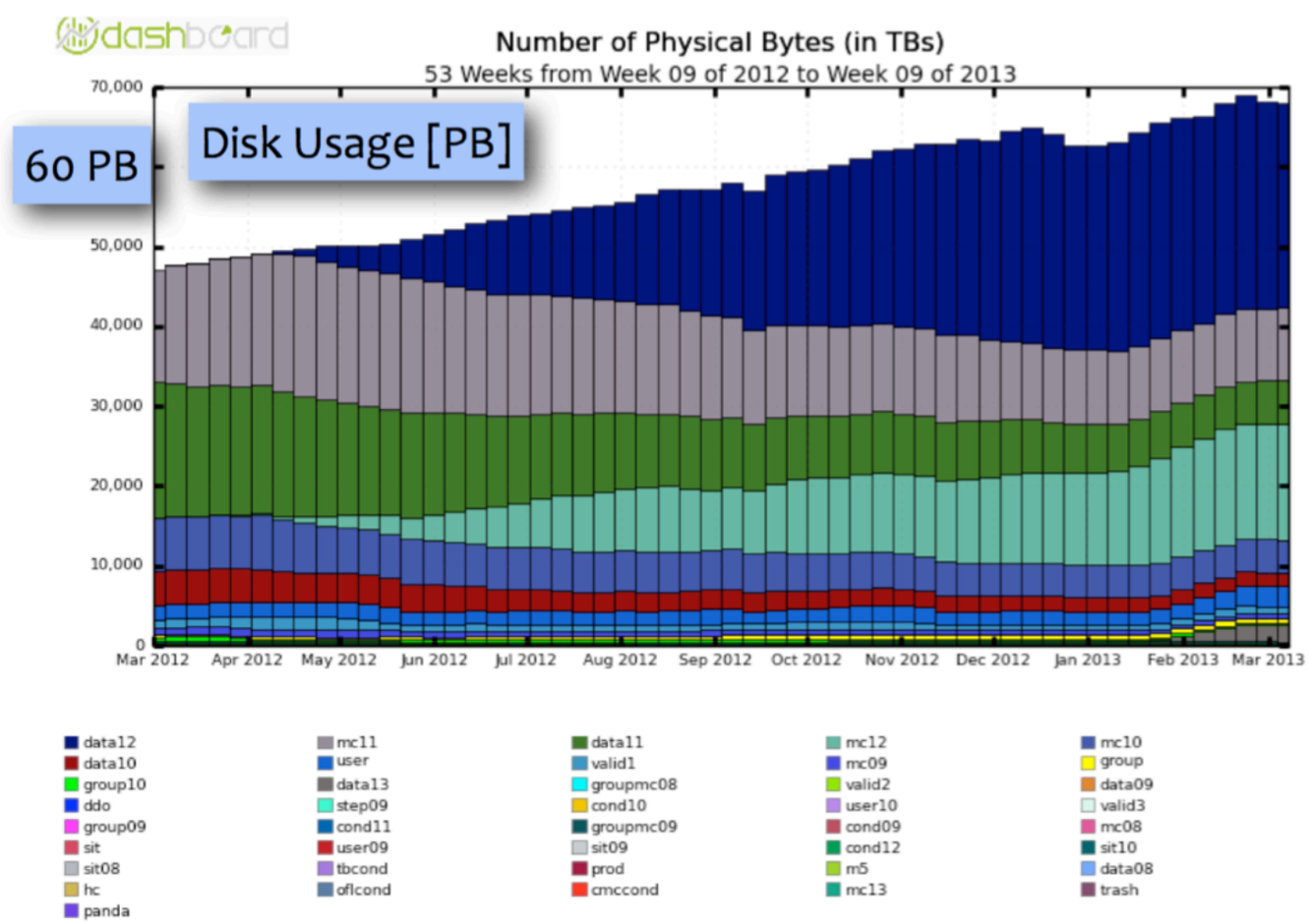
- Cosmic Frontier (add table?)
  - There is a continued growth in data from Cosmic Frontiers experiments (currently exceed 1 PB, 50 PB in 10 years, 400 PB per year in 10-20 years)
  - The cosmic frontier presents several faces, each presenting its own challenges for data management and storage: terrestrial sky survey telescopes, terrestrial radio telescopes, HEP- detector-in-space telescopes, and large-scale simulations.
  - Sky Surveys
  - SDSS pioneered the use of innovative database technology to make its data maximally useful to scientists. This approach continues with LSST, notably the development of a multi-petabyte scalable object catalog database that is capable of rapid response to complex queries. The data management needs of the sky surveys – handling image catalogs and object catalogs – appear very different from those of experimental HEP, but nevertheless, the baseline LSST object catalog employs HEP's xrootd technology in the key role of providing a switchyard between MYSQL front ends and thousands of MYSQL backend servers. LSST's 3.2 gigapixel camera will produce 15 terabytes per night, building up to over 100 petabytes of images and 20 petabytes of catalog database during the first ten years.
  - Although the basic data-access technology to make LSST science achievable has already been demonstrated, it is certain that a vigorous LSST science community will want to attempt many scientific studies that will be poorly served without major additional developments. Not all LSST science will be possible using only the object catalog database. In particular, studies such as those for dark energy effects, of particular interest to the HEP community, are likely to require reprocessing of the LSST image data on HEP analysis facilities. The model for funding and executing these studies is not yet clear.
  - The Dark Energy Survey (DES) can be considered a precursor to LSST, taking data with a 0.6 gigapixel camera for five years from 2012 culminating in a petabyte dataset

Data Management: Cosmic Frontier (to be edited)

- Arrays of radio telescopes can present a data-volume challenge comparable with that posed by energy frontier hadron collider experiments. The most extreme example now being planned is the European-led Square Kilometre Array (SKA) project that expects to complete its Phase I system in 2020. SKA will feed petabytes/s to correlators that will synthesize images in real time, producing a reduced persistent dataset on the scale of 300 to 1500 petabytes per year. These volumes can only be realized if considerable evolution of computing and storage costs happens by the time SKA data flows. Although SKA currently has no US involvement, it presents a concretely planned example of the technologies and data-related challenges that will certainly be faced by US scientists involved in projects in the same timeframe.
- Todays example of the SKA concept is the Murchison Wide-Field array where a raw 15.8 gigabytes/s is processed to a produce a stored 400 MB/s.
- Simulation provides our only way to perform “experimental cosmology” since only one universe is observable. Simulation also plays a vital role in understanding all aspects of astrophysics, such as supernovae, for which only very limited observation data can be collected for each occurrence. Finally, simulation is needed for the design of observational programs and for their detailed technical elements.
- Already today, post-processing of simulation data presents a major data-intensive computing challenge, requiring data management, large-scale databases and tools for data analytics. Some of today’s pain relates to the much more ready availability of national resources for computation than those for data management and analysis: “we can easily generate many petabytes from simulations and have [almost] no place to store them and analyze them”

Data Management: Energy Frontier

Atlas Data on Disk, across 11 Tier-1 Centers



CMS Data in the Fermilab Tape Library

	Custodial [TB]		Non-Custodial [TB]		Total [TB]	
	Last week	Total	Last week	Total	Last week	Total
Data	143.39	9048.89	86.96	1854.74	230.35	10903.63
MC	0.69	6831.85	0.00	416.03	0.69	7247.88
Total	144.08	15880.73	86.96	2270.77	231.05	18151.51

- **Current** LHC data sizes
  - examples
    - Atlas: 70 PB on disk, world-wide
    - CMS 18.2 PB on tape, at Fermilab
- **Future** increases
  - an estimate for 2021
    - ~130 PBytes detector data
    - ~350 PBytes simulated data
    - ~270 PBytes US "data library" (across Tier-1/2 centers)
    - CMS estimates for #events:
      - 30B data events, 46B simulated events

Data Management: Energy Frontier

- Cost consideration dictate the data rate: storage and processing cost
  - make choices and set priorities about which type of events we can collect and what data analyses we can follow through on, based on how much computing and storage resources we can afford
  - BTW, Science at EF lepton colliders is unlikely to be constrained by data management and storage issues
- LHC now treats the all the data written to persistent storage equally through the production phases
  - active dataset defined by the online trigger w/ ~100ms-level decision time
- Instead stream large fraction of data for storage on tape only
    - for just a penny you can store almost 1,000 CMS raw events on tape!
  - no further reconstruction, distribution unless physics case arose
    - exploits cost hierarchy for storage, also in view of slowed disk price decay
  - progressively pair down the active dataset with understanding and time

- LHC-style distributed data management is major success, but costly
  - distribute and serve the data much more flexibly and dynamically

Data Management: Energy Frontier

- We cannot afford another factor of 100 increase in storage, so we need to find ways of being more efficient in the use of the space
- The move to remote data access over the network (without performance penalties) is transformative
  - release requirement of data locality (enables more cost-effective clouds)
  - allows flexible, dynamic data placement, federation of existing data stores
  - possible centralization of cost-effective data stores and archival facilities
  - data federations already deployed as a first step, but work is needed!
- Data Management moving to “content delivery networks”
  - Data Management resources that deliver data on demand
  - data to be cached and replicated and intelligent about the placement
  - this brings large demands on network connectivity
  - a 10k core cluster (probably typical for 2020) would require 10Gb/s networking for organized processing, analysis would require 100Gb/s

Data Management: Intensity Frontier

- Largest data set: Lepton colliders in intensity frontier "factory mode"
  - also run up against the cost of storage, but the physics of lepton collisions is relatively clean and recording all events relevant to the targeted physics has proved possible in the past and is a realistic expectation for the future.
  - The Belle II TDR estimates a data rate to persistent storage of 0.4 to 1.8 gigabytes/s, which is comparable to LHC Run-2 rates but without the need to discard data with significant physics content.
- Most of the many other intensity frontier experiments do not individually challenge storage capabilities, but there is a recognition that data management (and workflow management) is often inefficient and burdensome.
  - Most experiments find it hard to escape from the comfort and constriction of limiting all their data-intensive work to a single site – normally Fermilab.
  - The statement "all international efforts would benefit from an LHC-like model" was made and should probably be interpreted as a need for LHC-like data management functions at a much lower cost and complexity than the current LHC system

Data Management: The Challenge of Data Longevity

- What to do with 100s of PB of data over 10s of years?
  - Irreplaceable resource, should be preserved, some how, for the future
    - in the past, e.g. the entire data for a LEP experiment fits on a desktop hard drive
    - still, just 1 of 4 LEP experiments actually demonstrated ability to reanalyze data
- With respect to data preservation and open availability, the LHC experiments are actively developing appropriate policies
  - The intensity frontier community does not have a plan yet, but recognizes that the issue exists across the frontiers
  - data preservation and open availability is relatively simple to achieve once policies have been decided, but requires funding to do it
  - A first U.S. project for EF, linking to Biology, Astrophysics, Digital Curation
- For the Cosmic Frontier, images, and tabular object catalogs of sky surveys and other image-based astronomy are readily intelligible by other scientists and even the general public
- More information is in the report

Computations and Data

- Federated data is well-matched to distributed high-throughput computing on grids and clouds
  - requires networking and last-mile problems to be solved (e.g. science DMZ)
- Coupling the HPC centers to large data sets
  - current systems like at Argonne and Oakridge can be configured for data-centric analysis but they are not specifically designed for this task, nor will next generation systems
  - ALCF, NERSC, and OLCF are arguing for the Virtual Data Facility (VDF) concept as their preferred mode for data storage/analysis
  - HEP experiments will heavily benefit from the development of a data handling system that is easily distributed and has transparent access for the user. The example of the tiered computing used by the LHC experiments is a good basis for developing the model for the IF. It is expected that any solution developed would provide access to the data as well as methods for submitting jobs to the grid.

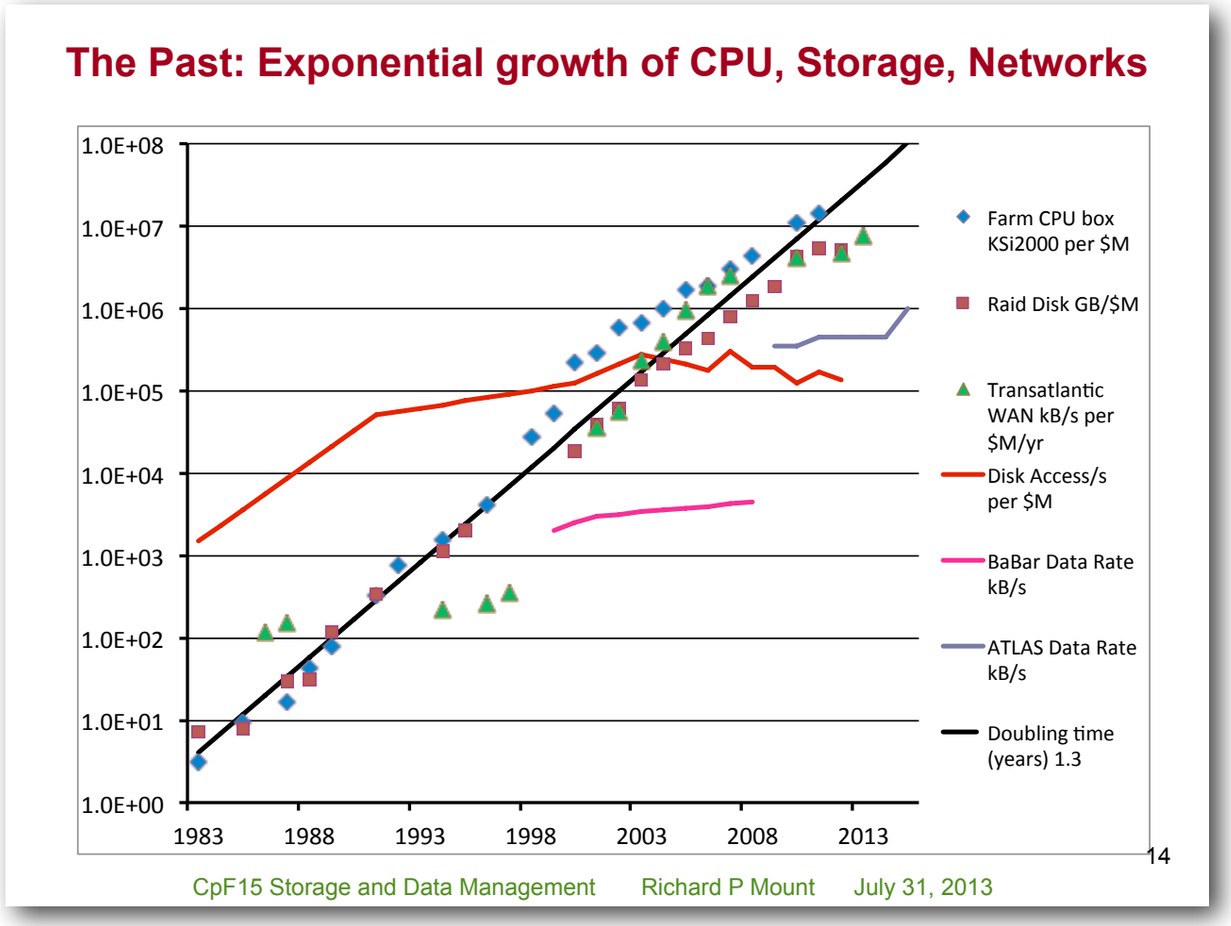
Networking

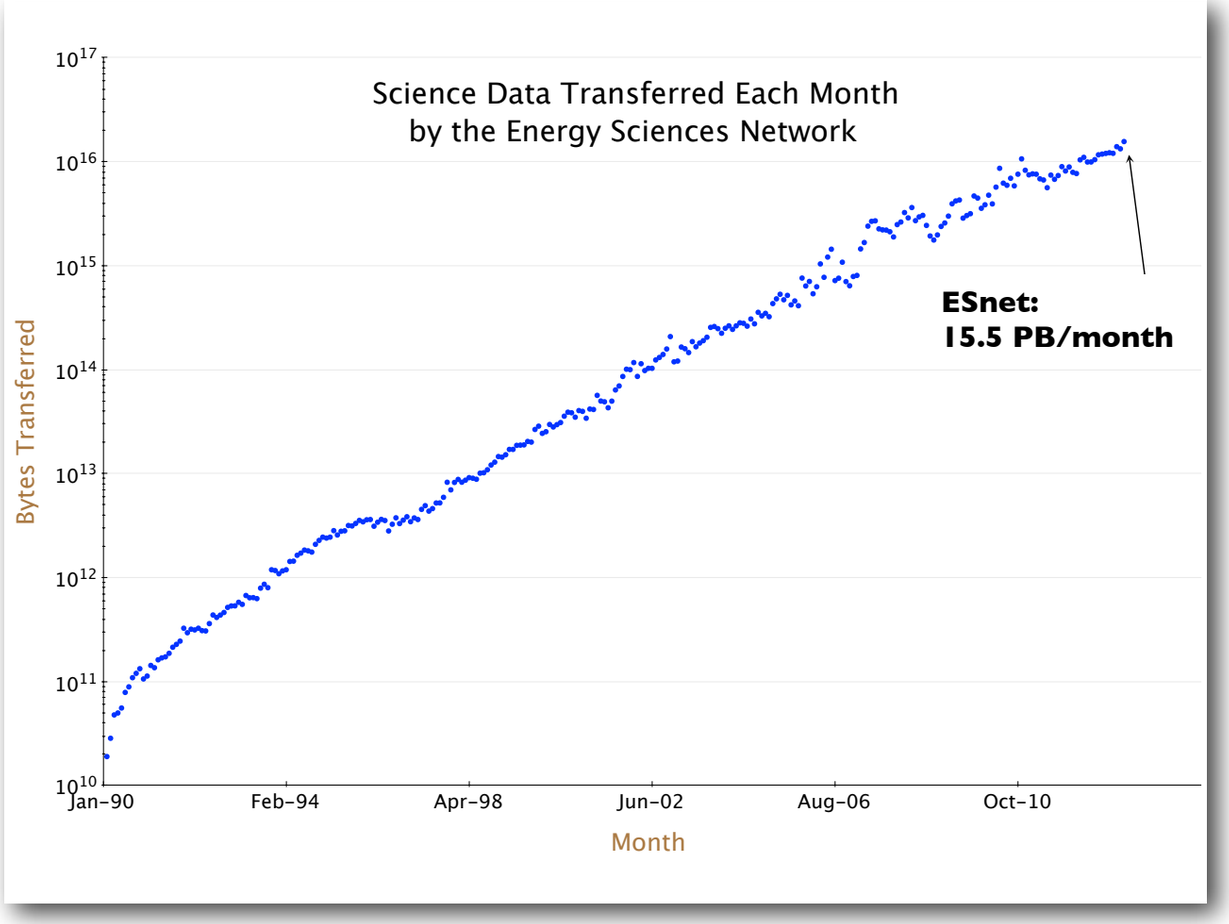
- All Frontiers depend on the availability of reliable, high-bandwidth, feature-rich computer networks for interconnecting instruments and computing centers, globally.
- Most HEP-related data is transported by National Research and Education Networks (NRENs), supplemented by infrastructures dedicated to specific projects
  - NRENs differ from commercial network providers; they are optimized for transporting massive data flows from large-scale scientific collaborations.
  - NRENs offer advanced capabilities (such as multi-domain bandwidth guarantees) that are not generally available commercially.
- Although HEP was a pioneer in exploiting international research networks, other science disciplines are making a similar transition now.
  - NRENs will be challenged as a result, and must be adequately motivated, resourced, and engaged.
- HEP's objectives through 2020 require basic and applied Network Research
  - Need support and growth for research in areas relevant to networking science.
  - Recent investments have been too small, using up dividends from prior research
  - Translate research results into operational practices is critical, but poorly funded

Networking

- There are no fundamental technical barriers associated with transporting 10x more traffic in 4 years; However, even at that traffic level and certainly beyond it, basic and applied research is necessary
  - to develop cost-efficient architectures
  - manage complexity
  - exploit programmability and other emerging network paradigms
  - assure that networks and applications become more tightly integrated
- Expectations for network performance need to be raised significantly, so that collaborations do not design workflows around a historical impression of what is possible, to allow the transformational changes addressed at yesterday's colloq.
  - The large gap between peak and average transfer rates must be closed.
  - Campuses must deploy high performance Local Area Network Infrastructure
- Whether the overall cost of networking remains stable over the next decade depends on the declining cost-curve for optical components, as well as the price of trans-Atlantic capacity
  - but there is a broad market for both.
  - In general, it's much cheaper to transport data than to store it

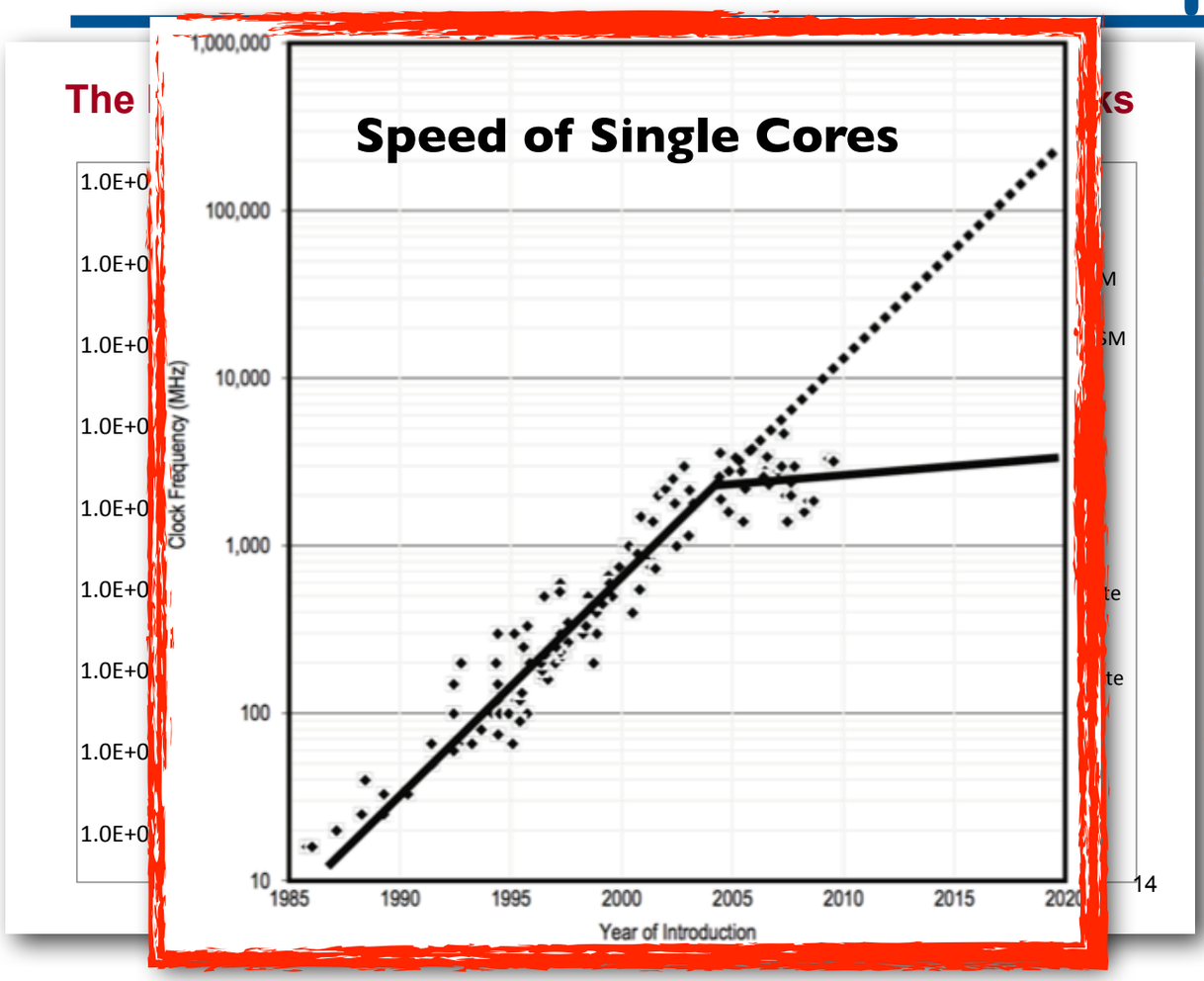
Technology Challenges after Decades of Exponential Growth





- Major shift in the nature of processors
  - single sequential applications has roughly stalled due to limits on power consumption
- We have been living in a temporary period of “multicore”, but even this cannot last due to power constraints
- Rotating disk will suffer marked slowdown capacity/cost.
  - Computing models must attempt to optimize roles of tape, rotating disk, solid-state storage, networking and CPU

Technology Challenges after Decades of Exponential Growth



Disks – from Per Brashers/DDN

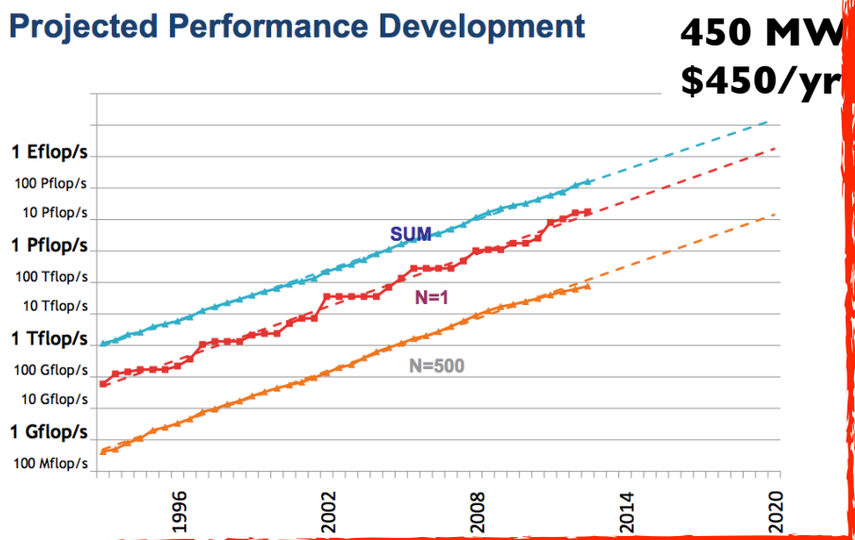
SLAC

- The area of a "bit" in current products is close to the limit where what is written will remain magnetically stable.
- New technologies to make the "bits" more stable are on the horizon:
  - "Shingled Recording" Not easily re-writable

  - Heat Assisted Magnetic Recording (HAMR)
  - [Laid-out-in-advance] Bit Patterned Recording
- None of these looks good for the near future.

NERSC Revolution in Energy Efficiency Needed

Even though energy efficiency is increasing, today's top supercomputer (N=1) uses ~9 MW or roughly $9M/year to operate. Even if we could build a working exaflop computer today, it would use about 450 MW and cost $450M/year to pay for power.



- Major shift in the nature of processors
  - single sequential applications has roughly stalled due to limits on power consumption
- We have been living in a temporary period of "multicore", but even this cannot last due to power constraints
- Rotating disk will suffer marked slowdown capacity/cost.
  - Computing models must attempt to optimize roles of tape, rotating disk, solid-state storage, networking and CPU

Technology Issues

- Advances in adapting key software tools to exploit multi-threading and GPU environments will be beneficial across frontiers.
- Large scale calculations are going to be limited by the energy consumption of the computer. Writing efficient codes is likely to become more difficult as we move to more exotic processors like GPUs or the Xeon Phi. It is not clear that one can abstract the details of the hardware in such a way that a single code can be written for both those targets.
- Large scale simulations in the cosmic frontier and lattice gauge theory are probably fairly similar
- Many of the components required to support virtual data already exist in the data and workflow management software of the largest experiments. The rigorous provenance recording required to support the virtual data concept would also benefit data preservation.
- Computing model implementations should be flexible enough to adapt to a wide range of relative costs of the key elements of HEP computing. In preparing for Run 3, the LHC program should seriously consider virtual data as a way to accommodate scenarios where storage for derived and simulated data becomes relatively very costly

Software and Training

- Three main themes or goals, and a number of recommendations
  - maximize the scientific productivity in an era of reduced resources
    - use software development strategies and staffing models that will result in products that are generally useful for the **wider HEP community**
  - evolving technology especially with respect to computer processors
    - develop, evolve software that will perform with efficiency in future computing systems
  - increasingly complex software environments and computing systems
    - insure that developers and users have the training needed to create, maintain, use
- Some of the recommendations
  - Significant investments in software to adapt to the evolution of computing processors: R&D into techniques, and as reengineering “upgrades”
  - Allow flexible, reliable funding of software experts to facilitate transfer of software and sharing of technical expertise between projects
  - Facilitate code sharing: open-source licensing, publicly-readable repositories
  - Include software i/s, frameworks, and detector-related applications early in project reviews, integrate software professionals with scientists

All Frontiers Agree: Need for Career Paths and Training

- ★ Use certification to document expertise and encourage learning new skills
- ★ Encourage training as a continuing experimental activity
- ★ Use mentors to spread scientific software development standards
- ★ Involve computing professionals in training of scientific domain experts
- ★ Use online media to share training
- ★ Use workbooks and wikis as evolving, interactive software documentation
- ★ Provide young scientists with opportunities to learn computing and software skills that are marketable for non-academic jobs
- ★ Training and career paths (including tenure stream) for researchers who work at the forefront of computation techniques and science is critical

Conclusions on Computation

- ★ HEP has large experience in both Distributed High-Throughput computing (experiment program) and High-Performance computing (mostly theory/simulation/modeling)
- ★ we are good at "collecting together" compute resources from where ever we can get them, from all tiers of computing including commercial
- ★ this brings significant challenges with data management and access
  - ✦ which the field actively works on
- ★ emerging network capabilities and data access technologies improve our ability to use resources independent of location, over the network
  - ✦ this requires more R&D
- ★ enables a large spectrum of provisioning resources: dedicated facilities, universities, opportunistic use, commercial clouds, leadership-class HPC,...
  - ✦ funding and allocation models: PI/allocation oriented vs research community oriented
  - ✦ supporting IT infrastructure provided by HPC center or labs, and those provided by "consortium" of collaborating institutions like OSG
  - ✦ emerging experiment programs might consider a mix to fulfill demands
- ★ with the need for more parallelization the complexity of systems continues to increase: frameworks, workload management, physics code

Conclusions on Data and Networking

- have to learn to do more with less. This requires being more flexible and perhaps tolerating higher levels of risk
- computing dedicated to projects needs to be solidly in place, such arrangements are possible and have been made in the past, through dedicated funding and through allocations at the centers
- current situation won't suffice for the next 10 years and continued investment will be needed to maintain our leadership in the face of the increase in data on all Frontiers
- The growth in data drives continued investment in data management, data access methods, networking
  - Continued evolution will be needed in order to take advantage of new network capabilities, ensure efficiency and robustness of the global data federations, and contain the level of effort needed for operations

Computing is at a great Starting Point for Moving into the HEP Future

- We have established and well-working computing models
  - the different frontiers are at some level separate in terms of facilities
    - but we are identifying many commonalities in terms problems and approaches
  - by coming together we are mapping out a good way to go forward
- HEP success has always been tied to advances in computing
  - like LHC computing being enabled by networks and the Grid
- HEP still drives areas of the technology, and certainly the collaborative space
  - distributed computing requires collaboration and partnerships, between sites, science communities, with computer scientists, between funding agencies etc
- push for using new technologies and approaches that are transformative
  - like parallelization and multi-core, virtualization, GPUs etc
  - for sure we'll see things over the coming years we have not yet thought of
- industry caught up to us and in cases surpassed us
  - might reassure us that with hard work computing won't be a road block
  - look at how industry can help and when this becomes feasible
- More commonality and community planning is needed for future computing systems in HEP

Collaboration and Partnerships

- Computing is an integral part of HEP science and is predominantly funded via the operations budgets for the major experiments and projects.
- The separate Computational HEP program funds computational science research via partnership initiatives, such as SciDAC and ASCR projects, the Open Science Grid, some NSF PIF grants, and a selection of community tools and pilot projects that cut across HEP programs.
- It optimizes a nominal budget by leveraging external and internal partnerships.

- add something about the international, inter-institutional, inter-disciplinary nature

Thanks!

to all sub-conveners and observers, to all participants in and contributors to the Computing Frontier summer study, and to the organizers of an exciting and inspiring Snowmass Meeting here in Minneapolis