



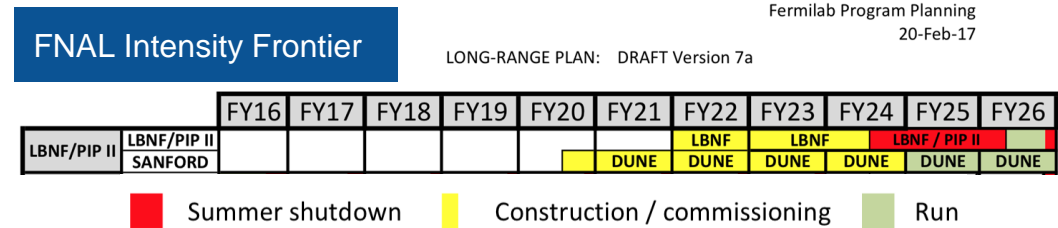
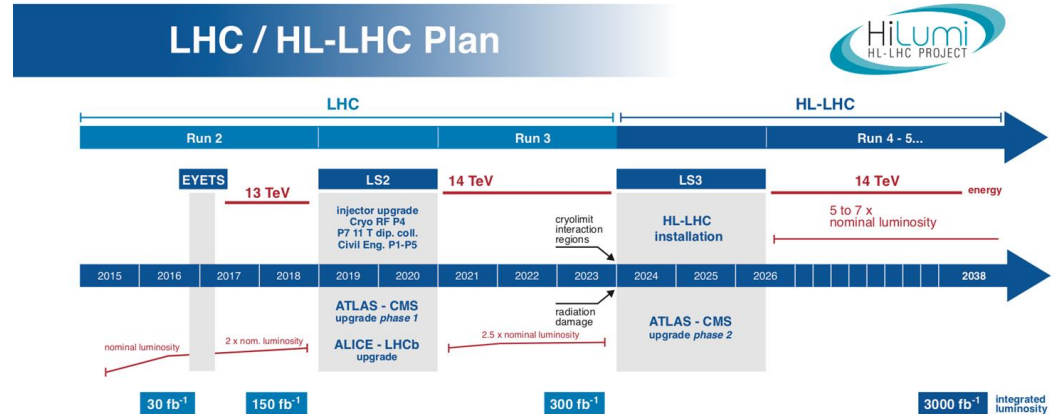
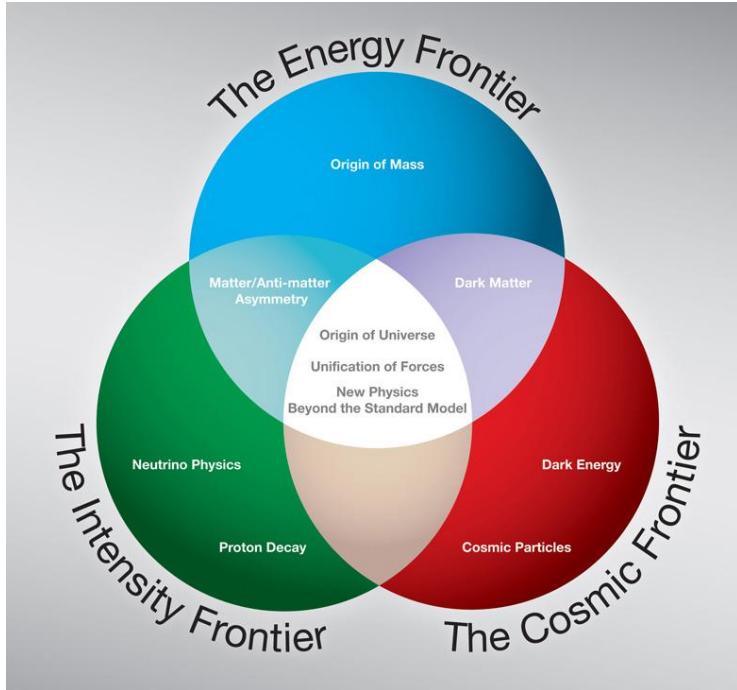
HSF Community White Paper review & outlook

Eduardo Rodrigues - University of Cincinnati
On behalf of the HSF



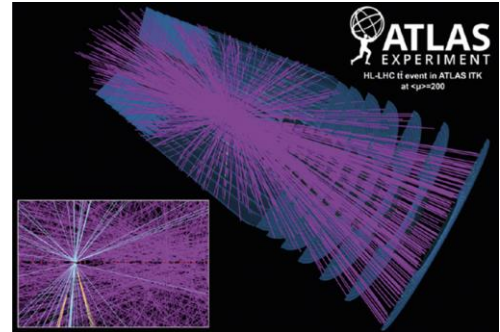
Quick recap, setting the scene ...

Experimental Particle Physics, Frontiers and Plans



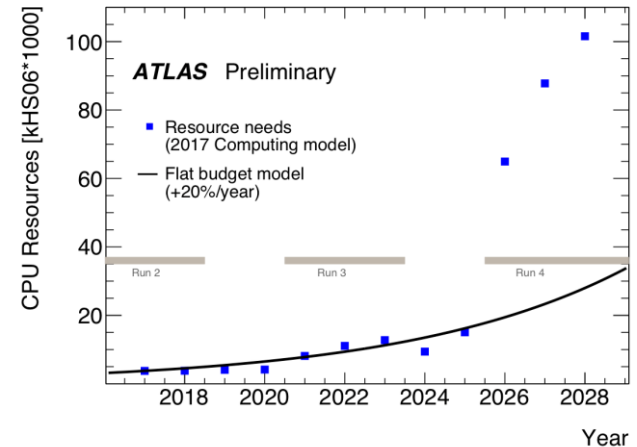
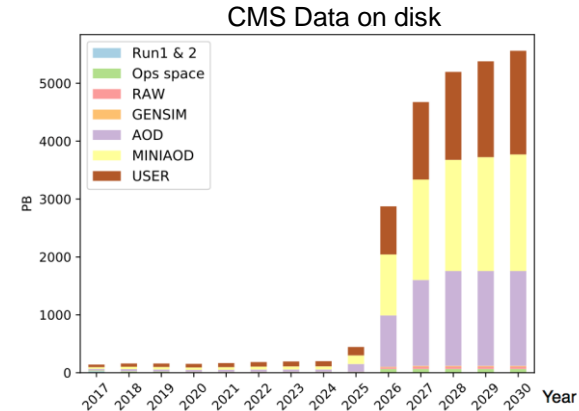
HEP Software and Computing

- High Energy Physics programme supported thanks to a vast investment in software
 - Estimated to be around 50M+ lines of C++
 - Which would cost more than 500M\$ to develop commercially
- It is a **critical part of our physics production pipeline**, from triggering all the way to analysis and final plots, as well as simulation
- LHC experiments use about 600k CPU cores every hour of every day and have around 400 PB of data stored on disk and 600 PB on tape
 - We are in the exabyte era already
- This is a huge and *ongoing* cost in hardware and human effort
- HL-LHC brings a huge challenge to software and computing in support for our ongoing physics programme
 - Both event rate and complexity rise (event pile-up of ~ 200)
 - Increased amount of data requires to revise/evolve our computing and data management approaches



SW&C Challenges for HL-LHC

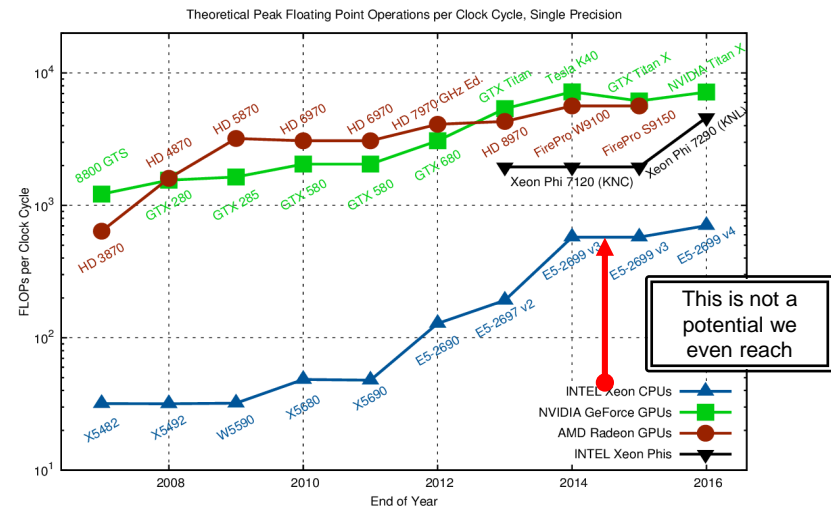
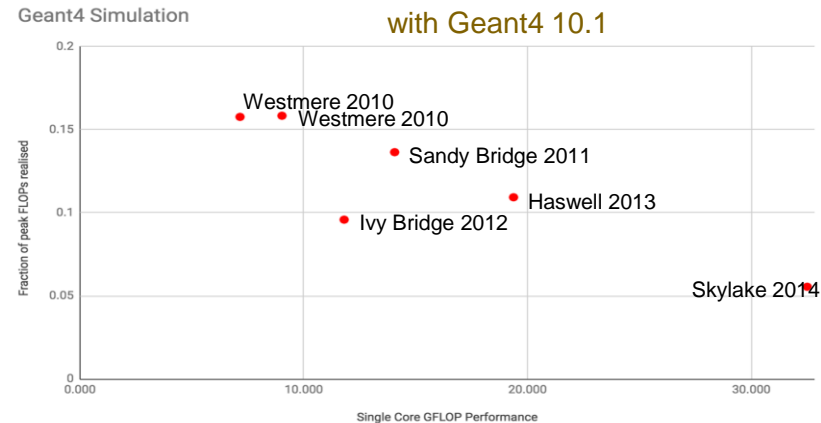
- High-Luminosity LHC is far from being a solved problem for software and computing
 - Hadron colliders (HE-LHC, FCC-hh) bring a massive data rate and complexity problem
 - Not just simple extrapolations of Run 2 software and computing
 - Anyway, naive extrapolations from today are not affordable
 - Also, resources needed would hugely exceed those from technology evolution alone
- Whatever the future, we pass through HL-LHC on the way
- *HEP Software Foundation Community White Paper maps out that path*



HEP code performance versus technology evolution

- Fraction of the potential floating point performance we use has been dropping over time
 - Growing performance gap on modern architectures
- HEP software typically executes one instruction at a time (per thread)
 - Usage of full potential of processors (internal parallelism) requires major SW re-engineering (but rewriting everything is not an option)
 - Co-processors like GPUs are of little use until the problem has been solved ...
- Accelerators have a different model
 - We are far from using available FLOPs
- *We have to adapt to maintain our ability to use processors effectively*

Private tbar event test in the ATLAS detector with Geant4 10.1



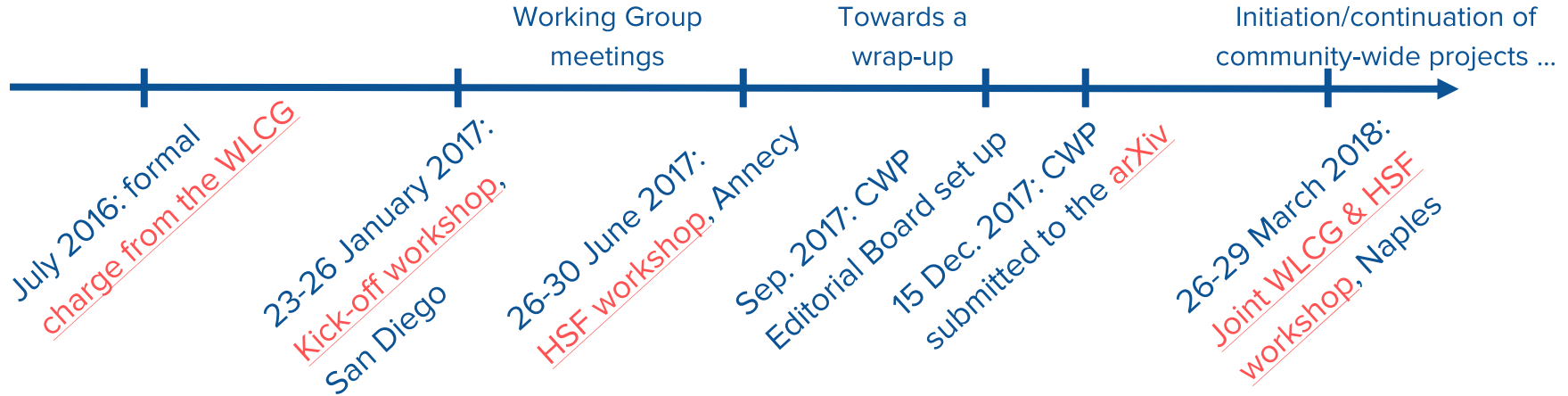
The HSF and the CWP process

HEP Software Foundation (HSF)



- The LHC experiments, Belle II and DUNE face similar challenges
 - HEP software must evolve to meet these challenges
 - Need to exploit all the expertise available, inside and outside our community, for parallelisation
 - New approaches needed to overcome limitations in today's code
- Cannot afford any more duplicated efforts
 - Each experiment often has its own solution for almost everything (framework, reconstruction algorithms, ...)
- HSF already started with a number of workshops and working groups on common topics (packaging, licensing, analysis)
- HSF established in 2015 to facilitate coordination and common efforts in software and computing across HEP in general
 - Our philosophy is bottom up, a.k.a. *do-ocracy*
- Get in touch: hsf-coordination@googlegroups.com

Community White Paper Process



Involved ~100 people, mainly US and EU



Involved ~100 people, mainly US and EU



Involving ~200 people

A Roadmap for HEP Software and Computing R&D for the 2020s

HEP Software Foundation¹

ABSTRACT: Particle physics has an ambitious and broad experimental programme for the coming decades. This programme requires large investments in detector hardware, either to build new facilities and experiments, or to upgrade existing ones. Similarly, it requires commensurate investment in the R&D of software to acquire, manage, process, and analyse the shear amounts of data to be recorded. In planning for the HL-LHC in particular, it is critical that all of the collaborating stakeholders agree on the software goals and priorities, and that the efforts complement each other. In this spirit, this white paper describes the R&D activities required to prepare for this software upgrade.

Community White Paper

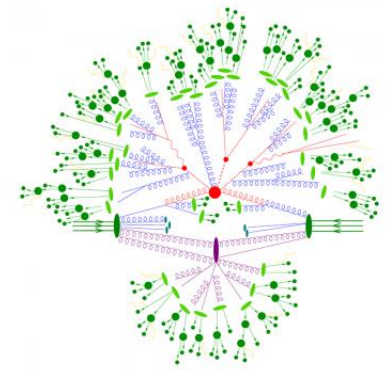
- Main resource page with links to topical papers, status of their publication to arXiv, updates on related activities, presentations about the CWP:
 - <http://hepsoftwarefoundation.org/activities/cwp.html>

“A Roadmap for HEP Software and Computing R&D for the 2020s”

- 70 page [document](#)
- **13 topical sections** summarising R&D in a variety of technical areas for HEP Software and Computing
 - Almost all major domains of HEP SW&C are covered
- **1 section on Training and Careers**
- **~300 authors** (signers) from 120 HEP-related institutions
 - Signing policy: sign the document if you agree with the main observations and conclusions
 - Remains open up till end of this workshop ... hurry up !
- Final version of the document published on [arXiv: 1712.06982](#) on 20 December (as announced in Oct. !)
 - Ongoing process for publication in the *Computing and Software for Big Science* journal
- CWP supported by several [topical WG papers](#) giving significant additional details, being pushed to arXiv

Contents	
1	Introduction 2
2	Software and Computing Challenges 5
3	Programme of Work 11
3.1	Physics Generators 11
3.2	Detector Simulation 15
3.3	Software Trigger and Event Reconstruction 23
3.4	Data Analysis and Interpretation 27
3.5	Machine Learning 31
3.6	Data Organisation, Management and Access 36
3.7	Facilities and Distributed Computing 41
3.8	Data-Flow Processing Framework 44
3.9	Conditions Data 47
3.10	Visualisation 50
3.11	Software Development, Deployment, Validation and Verification 53
3.12	Data and Software Preservation 57
3.13	Security 60
4	Training and Careers 65
4.1	Training Challenges 65
4.2	Possible Directions for Training 66
4.3	Career Support and Recognition 68
5	Conclusions 68
Appendix A List of Workshops 71	
Appendix B Glossary 73	
References 79	

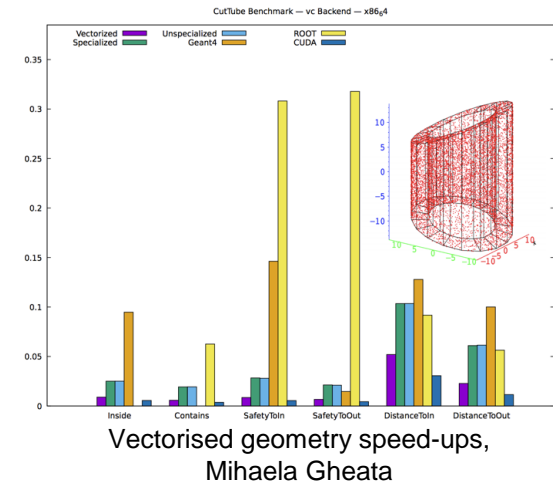
Physics Event Generators



- Physics event generation starts our simulation chain to enable comparisons with detector events
 - Depending on the precision requested, CPU for event generation ranges from modest to huge
 - At Next-to-Leading Order (NLO) precision used today, CPU consumption can become important
 - Study of rare processes at the HL-LHC will require the more demanding NNLO for more analyses
- Generators are written by the theory community
 - Need expert help and long-term associations to achieve code optimisation
 - Even basic multi-thread safety is problematic for many older, but still heavily used, generators
 - Ongoing maintenance of tools like HepMC, LHAPDF, Rivet is required and needs rewarded
- Writing this section was the result of intense contacts between HEP experts and the main players in the generator community
 - Several actions identified to foster the collaboration

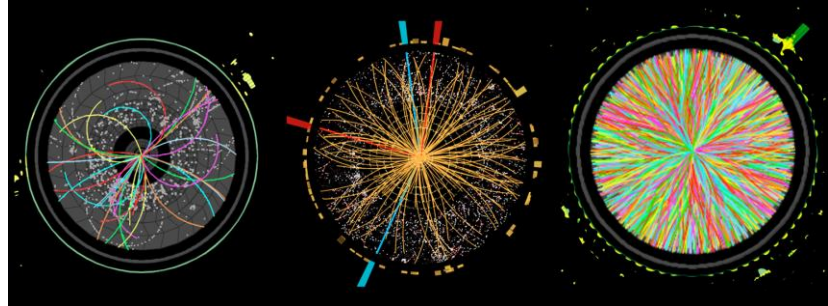
Detector Simulation

- Simulating our detectors consumes huge resources today
 - Remains a vital area for **HL-LHC and intensity frontier experiments** in particular
- Main R&D topics
 - **Improved physics models** for higher precision at higher energies (HL-LHC and then FCC)
 - Adapting to **new, in particular vectorised, computing architectures**
 - **Fast simulation** - develop a common toolkit for tuning and validation
 - **Geometry modelling**: easier and efficient modelling of complex detectors
- **All main experts and frameworks contributed** to a 50 page detailed review of the detector simulation challenges and required R&D actions
 - **CWP brought a more consistent view and workplan** between the different projects, in particular Geant4 and GeantV



→ See plenary session Mon. 16h

Software Trigger and Event Reconstruction



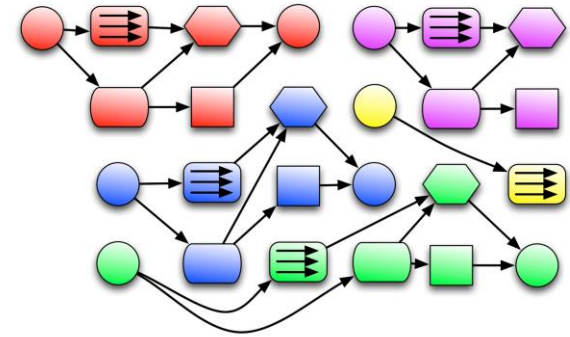
- Move to software triggers is already a key part of the program for LHCb and ALICE already in Run 3
 - 'Real time analysis' increases signal rates and can make computing much more efficient (storage and CPU)
- Main R&D topics
 - Controlling charged **particle tracking resource consumption** and **maintaining performance at very high pile-up** (of 200 or even 1000!)
 - Detector design itself has a big impact (e.g., timing detectors, track triggers)
 - Improved use of **new computing architectures**, extending the use of GPGPUs and possibly FPGAs
 - **Robust validation techniques when information will be discarded**
- Section rather HL-LHC specific
 - Trigger is highly dependent on the kind of experiments
 - One area where experience of Run 1 + Run 2 gives a better idea of challenges

Data Analysis and Interpretation

- HEP analysis currently dominated by many cycles of data reduction
 - Goal: reduce the input to an analysis down to a manageable quantity that can be cycled over quickly on ~laptop scale resources
 - Key metric is 'time to insight'
 - ROOT playing a central role
- Main R&D topics
 - Can we benefit, and how, from the latest techniques/tools in data analysis developed/used outside HEP: skimming/slimming cycles consume large resources and can be inefficient
 - For this, need ways to **seamlessly interoperate** between their data formats and ROOT
 - New analysis facilities: can **interactive data analysis clusters** be set up in our budget constraint
- Area with a strong potential for collaborating with the non-HEP world
 - In particular, links to be strengthened with the data science community

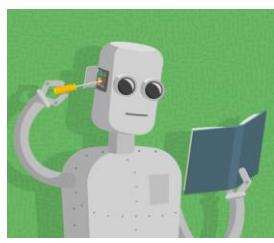
Data Processing Frameworks

- Experiment **software frameworks provide the scaffolding for algorithmic code**
 - Currently there are many implementations of frameworks, with some (limited) sharing between experiments (e.g. ATLAS/LHCb Gaudi)
 - Ongoing efforts in all these frameworks to support concurrency
- Main R&D topics
 - **Adaptation to new hardware**, optimising efficiency and throughput
 - Incorporation of external **(co)processing resources**, such as GPGPUs
 - **Interface with workload management** to deal with the inhomogeneity of processing resources
- General agreement that it is an area for consolidation in the future
 - Reasons for so many frameworks are not really related to experiment specificities...
 - But also the hardest component to change in the experiment SW stack: need to identify the best approaches to promote commonalities, e.g., in underlying components



Event processing
framework schematic
(colours are events,
boxes algorithms)

Machine Learning



- Not a challenge for HEP, per se, rather a tool set providing possible solutions
 - ML techniques applicable to many different areas
- Neural networks and Boosted Decision Trees have been used in HEP for a long time
 - e.g., particle identification algorithms
- The field has been significantly enhanced by new techniques (Deep Neural Networks), enhanced training methods, and community-supported (Python) packages
 - Very good at dealing with noisy data and huge parameter spaces
 - A lot of interest from our community in these new techniques, in multiple fields
- Main R&D topics
 - **Speeding up** computationally intensive pieces of our workflows (fast simulation, tracking)
 - **Enhancing physics reach** by classifying better than our current techniques
 - Improving **data compression** by learning and retaining only salient features
 - **Anomaly detection** for detector and computing operations
- **Good links with the broader Machine Learning and Data Science communities** required
 - Required to match the efforts and expertises needed to make effective use of these techniques

Conditions Data

- Growth of alignment and calibration data is usually linear in time
 - Per se, this does not represent a major problem for the HL-LHC
- Opportunities to use modern distributed techniques to solve this problem efficiently and scalably
 - Cacheable blobs accessed via REpresentational State Transfer (REST)
 - CVMFS + Files
 - Git
- Area where one can refocus current effort towards common software solutions ...
... should improve quality and economise overall effort
 - Actions already started thanks to the CWP

Visualisation

- Many software products developed for event visualisation
 - Part of the framework, with full access to event and geometry data
 - Standalone as a lightweight solution
- New technologies for rendering displays exist, e.g., WebGL from within a browser
 - We need to be adaptable to new technologies
- A new joint R&D project starting between CMS and ATLAS exploring advanced imaging techniques from bio-med community
- In general, agreed that community would benefit from commonality in tools and event data formats
 - Sustainability of all experiment-specific vis. Tools is a challenge

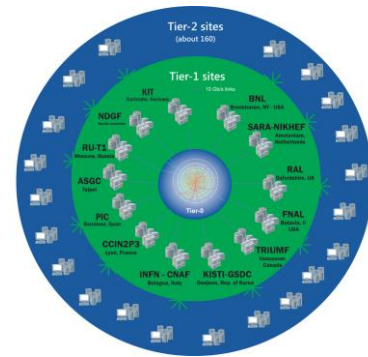
→ See plenary session Tue. 9h

Data Management, Organisation & Access



- Data storage costs are a major driver for LHC physics today
 - HL-LHC will bring a step change in the quantity of data being acquired by ATLAS and CMS
 - LHCb already there for this year's run 3
- Main R&D topics
 - Adapt to new needs driven by changing algorithms and data processing needs, e.g., **fast access to training datasets** for Machine Learning, **high granularity access** to event data, **rapid high throughput** access for a future analysis facilities...
 - **Consolidate** storage access interfaces and protocols
 - **Efficient hierarchical access** to data, from high latency tape and medium latency network
- HEP should be able to benefit from advances in industry standards and implementations, such as Apache Spark-like clusters
 - Not a drop-in replacement for our solutions: what they do is not exactly like what we do (structured access to complex data)

Facilities and Distributed Computing



- Storage and computing today are provided overwhelmingly from WLCG resources
 - Expected to continue for HL-LHC, but to be strongly influenced by developments in commodity infrastructure as a service, e.g. (commercial) Cloud Computing
- Main R&D topic: understand the **effective costs** involved in delivering computing
 - Activity started during autumn 2017: a WLCG WG formed and meeting fortnightly
 - This needs to be sensitive to regional variations in funding and direct and indirect costs like resources “beyond the pledge”, power and human resources
 - Full model unfeasible: only need a model good enough to answer questions about future investments, e.g. network or CPU vs. storage investments, impact of more concentrated storage resources (aka “data lake”)...
- Strengthen **links to other big data sciences** (SKA) and computing science
 - Common solutions where possible: cannot expect sites to deploy per-experiment technologies
 - Understand how to benefit from Software-Defined Networks or Storage

Security Matters



- **We have a large infrastructure** that is an important resource for us
 - **Protecting it is necessary** for both our work and for our reputation
- **HEP cannot live alone: cooperation with others is a requirement**
 - Already exists: HEP as a structured community often a driver in common efforts
 - Need to evolve to integrate new requirements from new big science communities and from new legislation, e.g; EU Data protection
 - Facing new threats: requires better intelligence sharing for threat monitoring and response
- **Main R&D topics**
 - Trust and policy, in particular **integration of commercial resources and hybrid clouds** that we don't control
 - Operational security: R&E Forum for incident response, broaden regional participation
 - Authentication and Authorisation, e.g. **generic authentication services** (e.g., eduGAIN) help users and are **easier than X.509 certificates**

→ See || session Tue. 13h30

Data, Software and Analysis Preservation



- We seem to be doing well compared to other fields
 - Already involved in cross-community efforts (e.g., Research Data Alliance forum)
- Challenge is both to physically preserve bits and to preserve knowledge
 - Knowledge preservation is very challenging
 - [DPHEP](#) has looked into both
- Need a workflow that can capture complexity of experiment production flows
 - Experiment production workflows vary in significant details
 - Variety of different steps are undertaken at the analysis stage, even within experiments
 - Technology developments that can help are, e.g., containers
- CERN [Analysis Preservation Portal](#) forms a good basis for further work
 - Needs to have a low barrier for entry for analysts
 - Can provide an immediate benefit in knowledge transmission within an experiment

Software Development



- Experiments have modernised their SW development models a lot recently
 - Suite of tools adopted from open-source community, which facilitate collaboration and knowledge sharing
 - source code migrated to **Git**, social coding (**GitLab/GitHub**) coupled with **Continuous Integration**, **CMake** becoming standardised way to build
 - Continuous integration and code review become natural and much easier than before
 - Pull/Merge request workflows help a great deal with code quality
 - More tools and expertise sharing will benefit the community - (re-)establish the development forum
- The more commonality in the tools and techniques, the more training we can share
 - Using new tools requires **investing in training** for the community
 - Recently put in practice by ALICE and LHCb, using LHCb **StarterKit** material
 - This helps with preservation and propagation of knowledge
- Our environment is becoming more complex: input required from physicists *whose concerns are not primarily in software*
 - **Sustainability** of these contributions is extremely important
 - We should become better at publication and citation of work to help this, e.g. using new tools such as **Zenodo** to trivially obtain DOIs





Training and Careers

- To address the technical challenges, we need to **raise the SW&Computing expertise in our community**
 - Investment in SW critical to match HL-LHC requirements with a “flat-budget” scenario
 - Using new tools requires **investing in training** for the community
 - This provides preservation and propagation of knowledge
 - A lot of the training we need to do is generic, but some is quite specific (to HEP or experiment)
 - We should also encourage appropriate training at the undergraduate and graduate school level
 - Sharing between experiments is still an exception: **training must become a first class activity**
 - ALICE and LHCb recently did this in practice using the [StarterKit](#) material
- Historically, **many different profiles involved in HEP computing**, from physicists, PhDs to real SW&Computing experts
 - Required by the cutting-edge challenges we face that demand all expertises to collaborate
 - No way to “outsource” the challenging problems to a few experts
 - But career path is often a challenge as well
- **Recognition of contribution** of our specialists in their careers is **extremely important**
 - Need for appropriate career path for software experts as much as any other technical discipline
 - We should also improve our publication and citation record (weak in some areas) and explore new avenues, e.g., [Zenodo](#)

Master Direction: Avoid HEP-specific Solutions

- HEP faced many computing challenges before other communities and has developed over the decades a lot of community-specific solutions
 - Mainly for good reasons!
 - Several HEP-tools adopted by some other communities, e.g. GEANT4 and ROOT, and WLCG itself was a model/driver for large-scale computing adopted by some other disciplines (e.g., EGI)
- But the world changed: other scientific communities and industry facing some similar challenges and HEP must be able to benefit from them
 - Machine learning, distributed analysis, distributed infrastructure (e.g., clouds...)
- Does not mean that we have drop-in replacements for our solutions
 - Challenge: find the proper integration between our community tools and the available technologies outside, maintain the necessary backward compatibility/continuity
 - As illustrated in CWP chapters, not one single approach for every topic: several paths for moving in this direction are part of the roadmap

Future beyond the CWP

The CWP - an Important Milestone for HEP SW&C

- **The Community White Paper process is **concluded** and has been a **success****
 - Engaged 250+ people and produced 300+ pages of detailed descriptions in many areas
 - A real & major step forward compared to situation before the CWP, thanks to fruitful discussions
 - Not a shopping list of all the possible ideas
- **But the **CWP is a milestone, not a final step****
 - R&D programme proposed in each area should serve as the **basis for future work**
 - Links fostered between the people involved in the SW&C of the major HEP experiments
 - Concrete paths identified to **move towards more common solutions** in HEP and to benefit from **solutions developed outside the community**
- **Each experiment must build its own prioritized R&D programme out of the CWP**
 - Priorities are different between all experiments: not facing the same challenges at the same time or scale: not possible to have a prioritized R&D program at the HEP level
 - On each topic relevant to several experiments, **must work together**

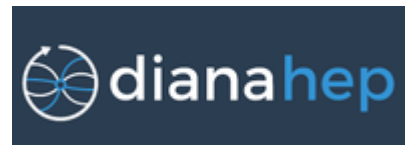
Advancing from here



- CWP laid down a roadmap for the future
- It identifies main areas we need to **invest** in for future for our **Software Upgrade** identified: we need to turn them into concrete actions
 - Explicit goal of this workshop !
 - HL-LHC is a driver: the Run 1 and Run 2 experience helps to better identify the challenges
 - **Must be inclusive of the whole HEP community**: better links with Intensity Frontier and Belle II
- **HSF, with its bottom-up approach, has proved its worth in delivering this CWP**
 - Managed to build a community consensus: not an easy and usual process in HEP
 - The challenges are formidable, working together will be the most efficacious way to succeed
 - Now a recognized organization to **spread knowledge** of new initiatives, to **encourage collaboration** and to **monitor progress**
- **Organisations and funding agencies support is required** for marshalling and refocusing the R&D efforts, and helping to **attract new investment** in critical areas
 - Career path of the needed experts is of critical importance for the medium/long term

Advancing from here - examples of initiatives

- Initiatives exist already to invest in future of software:
- DIANA-HEP project around Data-Intensive ANALysis
 - US project, with team ~10 people
 - Ongoing since 3-ish years
 - Strong links with various communities/experiments/HSF
- CERN's R&D programme on Experimental Technologies
 - Recent call for R&D proposals to gather ideas from the whole of the HEP software and computing community
 - 2 sessions of lightning talks [1, 2] took place recently
 - [1st workshop](#) on March 16th with reports from Software WG
- <Place yours here ...>



C-W-P : a personal interpretation / outlook

- **Can We Project (onto the future)?**

- We just did, in some sense, with the CWP ;-)

- **Community Work Packages**

- HSF helps here already, e.g. with its WGs on packaging & licensing
- We might need more ? To be discussed at this workshop, e.g. thumbs up on training ...
- To make best use of community expertise ... and help/funnel links to communities outside HEP?

- **Community-Wide Projects**

- “Look-elsewhere effect” - move out of exp.-specific solutions when possible (tradition is sticky)
- E.g. Gaudi (common framework), common toolset(s), common training (StarterKit)
- Definitely need many of these. This workshop will hopefully serve (also) as an incubator

- **In short:**

Propose



R&D

+

**Work
Deployment**



**Conclude !
Enable great science**

Back-up slides

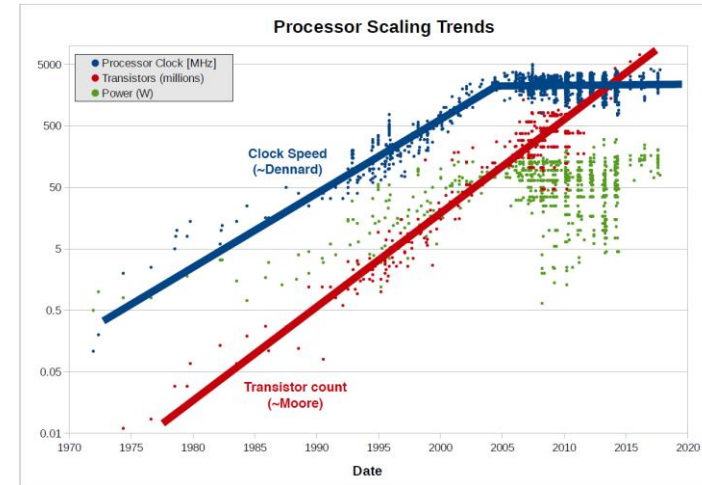
Back-up slides

Technology evolution - processors

- Moore's Law continues to deliver increases in transistor density
 - Doubling time is lengthening
 - With a flat budget, Moore's lawish improvements are the real maximum we can expect on HW side
- Clock speed increases stopped around 2006
 - No longer possible to ramp the clock speed as process size shrinks (Dennard scaling failed)
- So we are basically stuck at $\sim 3\text{GHz}$ clocks from the underlying Wm^{-2} limit
 - This is the *Power Wall*
 - Limits the capabilities of serial processing

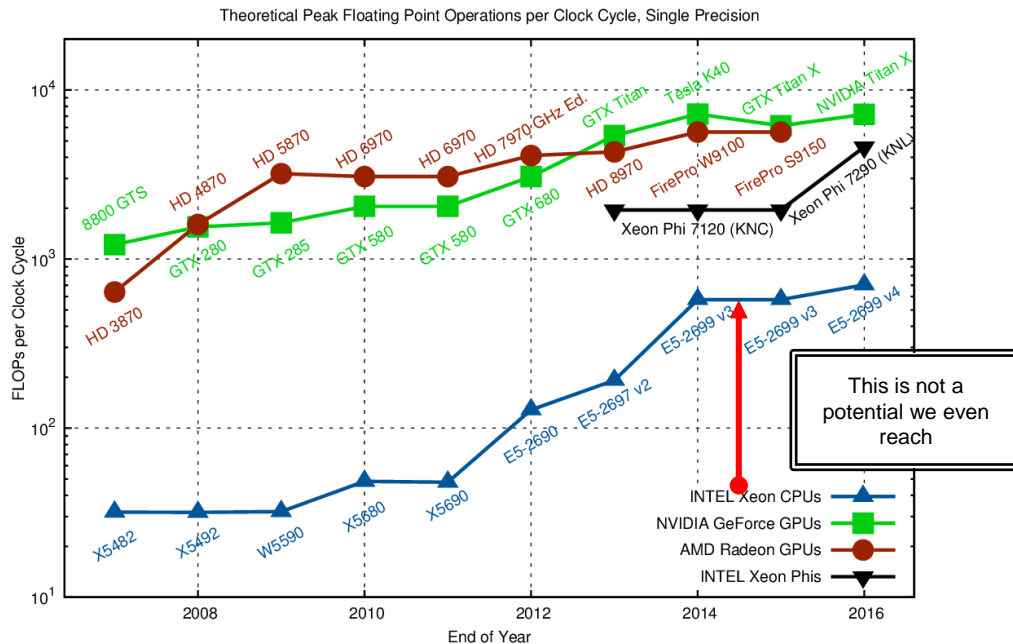
- HEP software typically executes one instruction at a time (per thread)

- Usage of full potential of processors (internal parallelism) requires major SW re-engineering (but rewriting everything is not an option)
- Co-processors like GPUs are of little use until the problem has been solved ...



Technology evolution - accelerators

- Most of the CPU die goes to things other than doing maths
 - Even CPU vector registers are hard for us to exploit
- Accelerators have a different model
 - Many cores, high floating point throughput, but lose a lot of 'ease of use'
- We have to adapt to maintain our ability to use processors effectively



Other Technology Trends

● Memory

- DRAM improvements now modest
- Overall, memory 'landscape' becomes more complex
 - Memory/storage boundary blurring

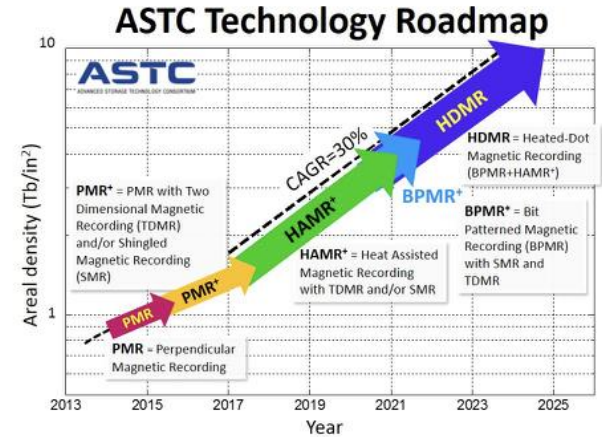
● Storage

- Spinning disk capacity keeps climbing
 - Time to read and cost improves, but slowly
- SSDs can read much faster, but price remains too high for bulk storage
- Tape remains cheap to buy, slow to access with few companies left, O(1)

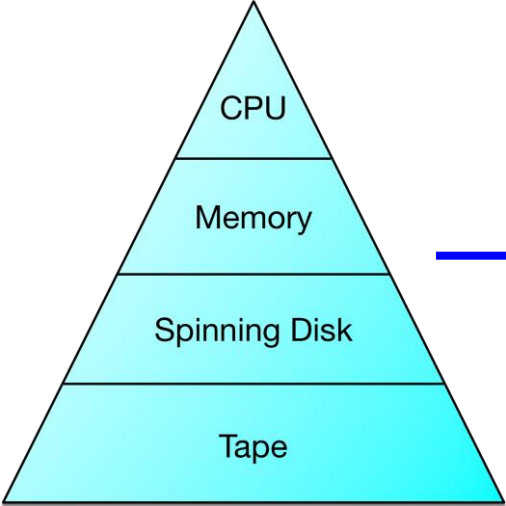
● Networks

- Capacity increases expected to continue, latency will not change
- Next generation networks offer capability to open channels between sites on demand
 - Useful, but an additional complexity

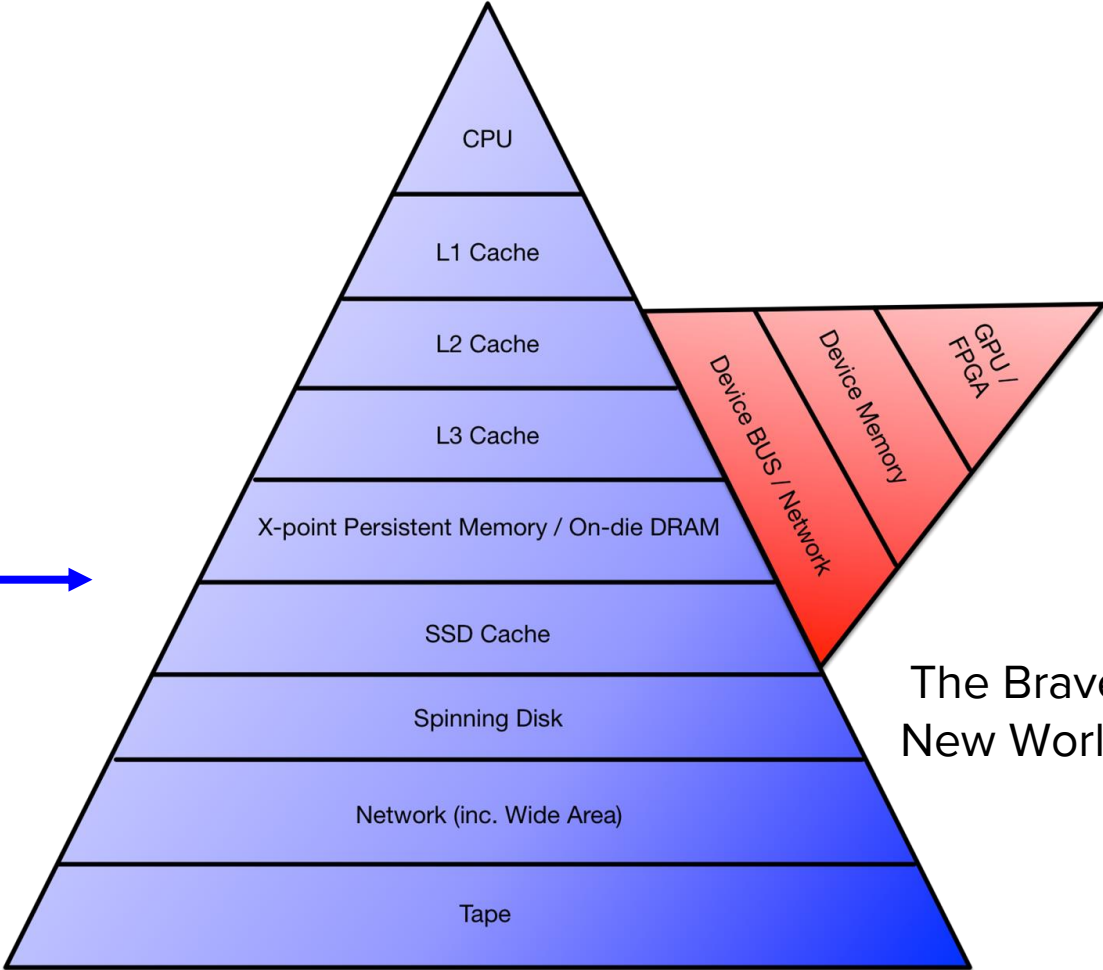
- Note: Game changer technologies might appear, but we cannot count on them



Meaning...



The Good Old Days



The Brave New World

Drivers of Technology Evolution

- Low power devices
 - Driven by mobile technology and Internet of Things
- Data centre processing
 - Extremely large clusters running fairly specialist applications
- Machine learning
 - New silicon devices specialised for training machine learning algorithms, particularly low precision calculations
- Exascale computing
 - Not in itself general purpose, but poses many technical problems whose solutions can be general
- Energy efficiency is a driver for all of these developments
 - Specialist processors would be designed for very specific tasks
 - Chips would be unable to power all transistors at once: dark silicon is unlit when not used



CWP - Making a Roadmap for the Future

- Editorial Board set up in Sep. 2017, with the aim of encompassing the breadth of our community
 - Wide regional/experimental representation
- First draft released Oct. 20, second draft Nov. 17
- These drafts elicited a **substantial response from the community**, leading to many improvements
 - **100s of commenters**, some sections deeply rewritten
- Final version of the document published [arXiv: 1712.06982](https://arxiv.org/abs/1712.06982) on 20 December (as announced in Oct. !)
 - Ongoing process for publication in the *Computing and Software for Big Science* journal
- Several [topical WG papers](#) giving significant additional details are being pushed to arXiv
- Main resource page with links to topical papers, status of their publication to arXiv, updates on related activities, presentations about the CWP:
 - <http://hepsoftwarefoundation.org/activities/cwp.html>

Editorial Board

- Predrag Buncic (CERN) - ALICE contact
- Simone Campana (CERN) - ATLAS contact
- *Peter Elmer (Princeton)*
- *John Harvey (CERN)*
- *Benedikt Hegner (CERN)*
- Frank Gaede (DESY) - Linear Collider contact
- Maria Girono (CERN Openlab)
- Roger Jones (Lancaster University) - UK contact
- *Michel Jouvin (CNRS/LAL)*
- Rob Kutschke (FNAL) - FNAL experiments contact
- *David Lange (Princeton)*
- Dario Menasce (INFN-Milano) - INFN contact
- Mark Neubauer (U.Illinois Urbana-Champaign)
- *Eduardo Rodrigues (University of Cincinnati)*
- Stefan Roiser (CERN) - LHCb contact
- *Liz Sexton-Kennedy (FNAL) - CMS contact*
- Mike Sokoloff (University of Cincinnati)
- *Graeme Stewart (CERN, HSF)*
- Jean-Roch Vlimant (Caltech)

(italic: core team aka "ghost writers")

Looking further into the future ...

Example of CERN's call for R&D proposals

- Software Working Group (among others)

- Open process

- Gather ideas from the whole of the HEP software and computing community
- Ensure alignment with developments outside CERN EP
- ~ 100 people on the [mailing list](#)

- Lightning Talks

- Two sessions of lightning talks [1, 2] - open to anyone to propose a topic
- Total of 28 short talks presented and discussed
 - Speakers from CERN EP and beyond

- [Core group](#)

- Formed to distill these ideas and guide towards R&D proposals
- 15 people including 2 HSF coordination team representatives (LHC Exps, CLIC, FCC, SFT, CERN IT)

- [1st workshop](#) on March 16th with reports from all WGs

