CWP Status

Michel Jouvin (CNRS/LAL)
LHCOPN/ONE Meeting, KEK
October 17 2017

hepsoftwarefoundation.org

- **The HSF started in 2014 to build a collaboration around SW across HEP**
  - HEP SW must evolve to meet the challenges posed by new experiments
  - The computing landscape is evolving rapidly; we need to exploit all the expertise available in our community, and outside it, to meet the challenges
  - The free-lunch form of Moore's Law is over, and we can't buy our way out of the problem with hardware
  - The affordable way to do that is to do it collaboratively
- **HSF Goal: facilitate coordination, share expertise and promote common efforts**
  - Raise awareness of existing software and solutions
  - Promote commonality and collaboration in new developments
  - Aid developers and users in creating, discovering, using and sustaining common software
  - Support training career development for software and computing specialists
  - Provide a structure for the community to set priorities and goals
- To follow the HSF activities, look at the web site!

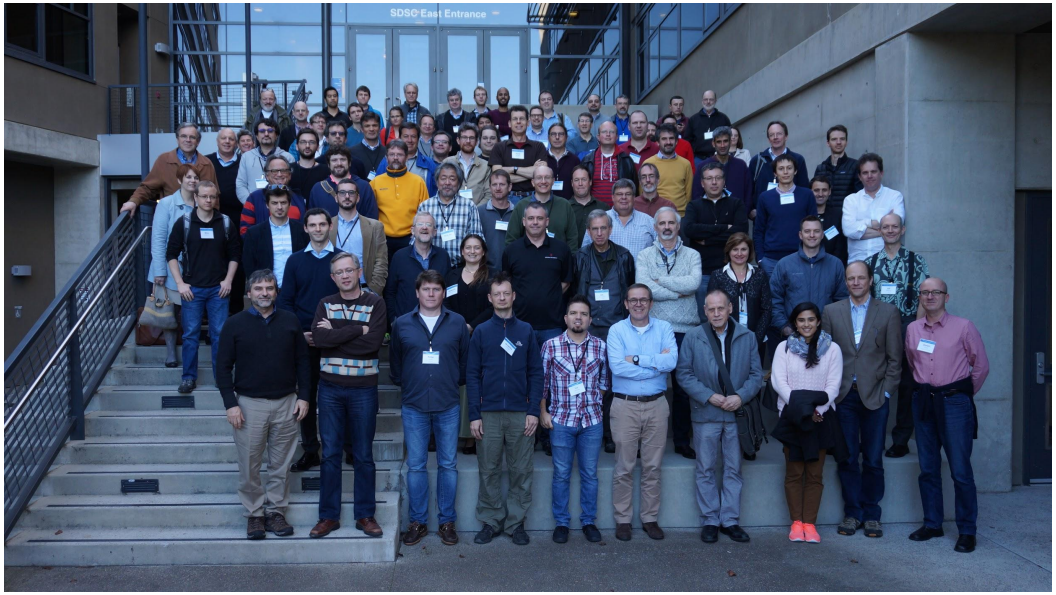
- 2016: proposal of a **Community White Paper (CWP)** to describe a global vision for software and computing for the HL-LHC era and HEP in the 2020s

- **The CWP will identify and prioritize the software research and development investments required:**
  - to achieve improvements in software efficiency, scalability and performance and to make use of the advances in CPU, storage and network technologies
  - to enable new approaches to computing and software that could radically extend the physics reach of the detectors
  - to ensure the long term sustainability of the software through the lifetime of the HL-LHC

- **The HEP Software Foundation (HSF) is engaging the HEP community to produce the CWP via a community process**
  - Initiated as an HL-LHC planning process
  - Aiming for a broader participation (LHC, neutrino program, Belle II, linear collider, FCC...)
  - The resulting roadmap will be used for the HL-LHC computing TDR and other strategic plans

CWP Kick-Off Workshop in San Diego (23.-26. January)

- ~110 participants, mainly US + CERN
  - Unfortunately very few Europeans from outside CERN
- 2.5 days of parallel topical WG meetings
  - Agenda : http://indico.cern.ch/event/570249/timetable/#all
  - From infrastructure to reconstruction and analysis, through simulation, data management…
  - Notes from (almost) all WG discussions in the WG Google Docs, summary slides in the agenda



This was the beginning for many other topical workshops

June 2017: (almost) concluding HSF Workshop at LAPP/Annecy (26.-30. June)

90 participants:

- US: 48 (8 FNAL)
- CERN: 20 (7 EP/SFT)
- France: 14 (7 LAPP)
- Italy: 3
- UK: 2
- Germany: 2
- Switzerland: 1

Organization:

- Indico Agenda
- Monday: Introduction+Status
- Tuesday: Parallel Sessions of WGs
- Wednesday: Plenaries+WGs
- Thursday: Plenaries+WGs
- Friday: Closeout

Working Groups active during the CWP process

- SW Trigger and Reconstruction
- Machine Learning
- Data Access, Organization and Management
- Software Development, Deployment and Verification/Validation
- Data Analysis and Interpretation
- Conditions Database
- Data and Software Preservation
- Event Processing Frameworks
- Physics Generators
- Simulation
- Workflow and Resource Management
- Visualization
- Facilities and Distributed Computing
- Careers, Staffing and Training

No WG dedicated to network:but network discussed in several WGs!

Full list of all working groups and their working documents:
http://hepsoftwarefoundation.org/activities/cwp.html

WGs produced a "CWP chapter" ready for review

- Generally a Google Doc open to comments
- Ranging from 10 to 50 pages! Long documents augmented by a shorter executive summary document
- See http://hepsoftwarefoundation.org/activities/cwp.html for details/updates

Documents contain the analysis of the challenges and the identification of areas for innovation + a roadmap of R&D actions for the next 5 years

- Prioritized based on the potential impact
- Milestones for the 2 main timescales:
  - 3 years (time of HL-LHC computing TDRs): deciding what is worth implementing
  - 5 years: implementation phase targeting ability to test/use during Run3

It is a community WP

- Good diversity of people in most (but not all) WGs, with non-LHC participants
- **Feedback is crucial to ensure these papers represent the community view**
- Chapters open to signature by the community

Deadline for the global CWP: end of November

- Global vision out of the individual documents rather than a summary
- Refer to the individual topical papers for each area details
- Advanced draft expected by the end of October: see the HSF web site

Besides the authors, there will be the possibility for everyone to sign the final CWP

Members of the editorial board so far (see HSF site)

- Predrag Buncic (CERN)
- Simone Campana (CERN)
- Peter Elmer (Princeton)
- John Harvey (CERN)
- Maria Girone (CERN Openlab)
- Michel Jouvin (LAL Orsay)
- Mark Neubauer (U.Illinois Urbana-Champaign)
- Stefan Roiser (CERN)
- Liz Sexton-Kennedy (FNAL)
- Mike Sokoloff (U.Cincinnati)
- Graeme Stewart (CERN)
- Jean-Roch Vlimant (Caltech)

Network not discussed as a separate topic

- Learnt so far that network is a dependable resource, historically under-estimated. Data federation is now everywhere in LHC experiments.
- General assumption is that network capacity will continue to increase and that we can rely on it for more network-centric and network-aware workflows
- Some worries that LHC experiments are no longer the only very demanding users: may have to learn how to efficiently share the available resources

How can new advanced technologies benefit to our use cases?

- In particular SDN and NDN
- Already some R&D projects going on: CWP roadmap may put them better in perspective and increase their importance
- Decreasing border between WAN and LAN
- Strengthening the links between network community and HEP would help: the WG proposal presented yesterday by Marian Babik may be a contribution

Chapters particularly relevant to the network community

- Facilities and distributed computing: analysis of current infrastructure strengths/weaknesses, explicitation of the tradeoffs and identification of R&D areas that will help to build the HL-LHC era infrastructure. “Data-lake” idea: large, potentially federated, data facilities used by compute resources around the world. Cross-cutting topics with many other WGs. 50 pages!
- Data management, organization and access: mainly focused on data organization versus data access patterns and the impact/potential of new storage technologies.
- Data Analysis and Interpretation: new analysis models, analysis facility

Network community feedback is very welcome

- The CWP is a unique attempt to get the community thinks together at its future
  - Will help define the requirements for the future and discuss with the resource providers
  - Final CWP expected in November (but a draft should be available by end of October), with topical papers complementing it with detailed analysis and proposals for each area
- The CWP is an open, bottom-up process
  - Significant community engagement: ~250 people involved in the WG discussions
  - **This is still time to provide feedback**
- CWP information and update are on the HEP web site