

Status of HSF and the Community White Paper

Peter Elmer - Princeton University
(for the HSF startup team)

23 January, 2017
HSF Workshop at SDSC

Overview

- Welcome to the HSF workshop at SDSC!
- The main focus of this workshop is our “Community White Paper” roadmap project
- <http://hepsoftwarefoundation.org/cwp.html>



Kathryn (Katy) Huff

Assistant Vice President
Academic Programs and Technological Engineering
Kathy.huff@nysed.gov
609.462.6100 ext. 200 or 609.462.6100 ext. 201

My mission:

- Increase and promote of advanced science and technology
- Coordinate science and technical education
- Align resources for technology education

My expertise is:

- Coordinated systems of advanced science education
- Development of advanced science field projects
- Advanced science management, space and technology resource initiatives, development

A problem I'm grappling with:

- Student readiness to nuclear science careers
- Career readiness

I've got my eyes on:

- Space exploration, scientific computing.

I want to know about:

- Opportunity opportunities for high school students to participate in science simulation












Shantenu Jha

Assistant Professor, Computer Engineering
IIT Madras

www.shantenujha.com

My research:

- High-performance and distributed system architectures, and standards based middleware and software systems.

A problem I'm grappling with:

- A Systems approach to the design and fabrication of distributed systems.
- Design and architecture for distributed systems management.
- Design of Distributed Systems and Software.

I've got my eyes on:

- A Building Blocks Approach to Hadoop/MapReduce.
- Heavy Rings.



Robert Kalescky

1976 Molecular Sciences Institute, Kansas for Scientific Computing, Southern Methodist University
 Molecular Sciences Institute, Kansas for Scientific Computing, Southern Methodist University

My research:

- Computational Biophysics (Small molecules, materials)
- in silico drug discovery (protein targets, binding)

My expertise is:

- Computational chemistry methods (classical, quantum and DFT)
- Parallel programming
- GPU accelerated performance analysis

A problem I'm grappling with:

- Supporting mobility of application software stacks

I've got my eyes on:

- Improving methods and workflow to increase time-to-solution
- Machine learning methods and workflow
- Large-scale software design and support



SMU



Daniel S. (Dan) Katz

Senior Lecturer in the Institute for Software & Applications, MCS
Research Associate Professor, 2012
Research Associate Professor, 2013
University of Illinois at Urbana-Champaign
dkatz@cs.uiuc.edu or dkatz@illinois.edu

My research:

- Improving the computational efficiency of data science and engineering computing systems: how, why, and possible problems

My expertise is:

- Applications, algorithms, tool, solutions
- Program design, programming, and software development
- Software development, programming, and software development
- Software development, programming, and software development

Problems I'm grappling with:

- Scalability of data science and engineering computing systems
- Career paths for computing researchers
- Changing the academic environment

I've got my eye on & want to know more about:

- Computing systems: how, why, and possible problems

Logos:



Jim Kowalkowski

Senior Software Architect
 Director Software R&D (Cloudnet)
 Oracle Systems for Business Applications

My expertise is
 Systems architecture, design, implementation, C++, distributed computing

A problem I'm grappling with
 Designing high capacity, multi-tenant, large-scale products for CPE. Studying algorithms and software architectures for CPE, while the company grows.

I've got my eyes on
 New memory architectures. And on algorithms to be able to handle the complexity. The new distributed architectures. Distributed machine learning.

I want to know more about
 The new algorithms for distributed computing and things I get my eye on.








Amit H. Kumar

Ph.D. (University of Illinois) [Research Interests: Data Mining, Machine Learning, Knowledge Discovery, Data Science](#)

[amkumar@smu.edu.sg](#) [@amkumar](#)

My expertise is:
Data Science and High Performance Computing

A problem I'm grappling with:
How to create sampling algorithms for ATLAS (a non-Euclidean graph) in applications like development opportunities in SMU perspective that require an emphasis on SMU.

I've got my eyes on:
Growth

I want to know more about:
What is the workshop all about.



SMU SINGAPORE MANAGEMENT UNIVERSITY



David Lange

Research Lead
Department of Physics
Princeton University
D.Lange@princeton.edu

My research:
 integrable quantum systems, quantum entanglement,
 quantum chaos, quantum information
 (QFT perspective and string theory)

My expertise is:

- Event generation, detector simulation,
- Event reconstruction techniques in HEP

A problem I'm grappling with:

- Low-energy scientific output for HEP?

I've got my eye on:

- The end, but someone to support application-developers, integrate analysis of incoming applications in HEP (or at least CERN)

I want to know more about:

- Synergies of HEP techniques with astronomical community-developed tools and applications






David Leary

[Enter Research Proposal \(closed\)](#)
[High Level Overview of Research](#)
[Abstracts of research on Change.org](#)
[@learydavis](#)

My expertise is:

- Linux system administration
- HTCondor, @Cache
- Large computing clusters

A problem I'm grappling with:

- Integrating the Atlas workflow into Bioconductor

I've got my eyes on:

- I want to learn more about:

My research:

System administrator for National T1 Z







Miron Livny

Professor of Management and Computer Science
University of Massachusetts
Amherst, MA

My research:
Interconnected High Throughput Computing

I sent mail about this yesterday. The slides to edit are at <http://bit.ly/2jhqpu5>. If you have problems talk to me.

HEP Software Foundation - Timeline

- Apr 2014 - HEP Software Collaboration meeting
- Jan 2015 - HEP Software Foundation workshop (2 days, SLAC)
- Apr 2015 - HEP Software Foundation session (0.5 day, Okinawa, CHEP2015)
- May 2016 - HEP Software Foundation workshop (3 days, LAL-Orsay)
- Jan 2017 - HEP Software Foundation workshop (3.5 days, SDSC/UCSD)

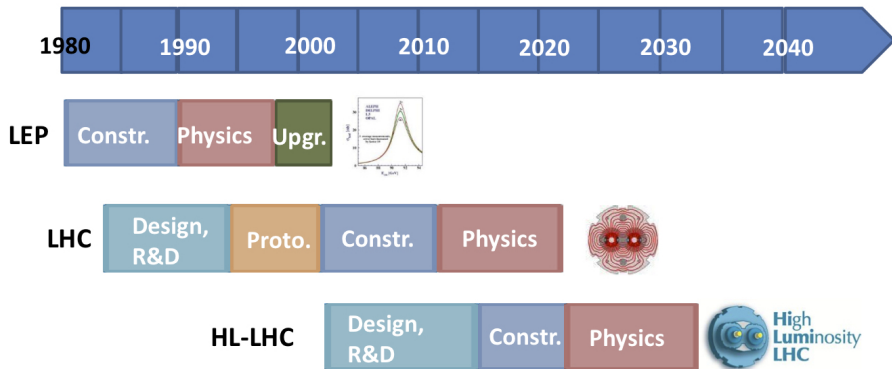
HEP Software Foundation (HSF)

The HSF (<http://hepsoftwarefoundation.org>) was created in early 2015 as a means for organizing our community to address the software challenges of future projects such as the HL-LHC. The HSF has the following objectives:



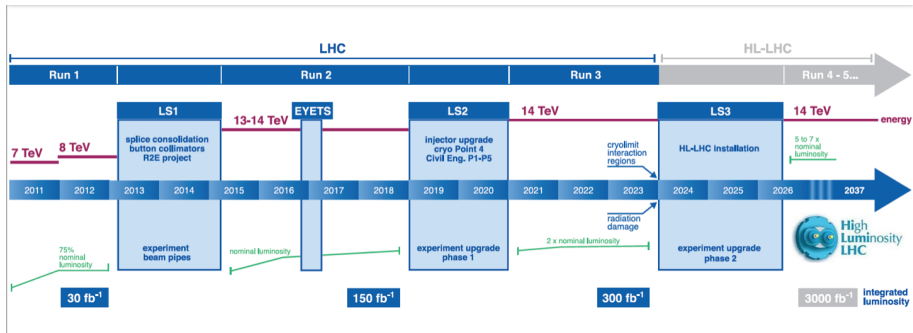
- Catalyze new common projects
- Promote commonality and collaboration in new developments to make the most of limited resources
- Provide a framework for attracting effort and support to S&C common projects (new resources!)
- Provide a structure to set priorities and goals for the work

CERN Accelerator Timeline



Various concepts also exist for subsequent machines.

Plans for upgrading the LHC and Experiment Detectors



A Software “Upgrade” for HL-LHC and 2020s HEP?

Looking forward to the next 10 years, we see a number of challenges for HEP software and computing:

- **Scale:** The HL-LHC will integrate 100 times the current data, with significantly increased data (pileup) and detector complexity.
- **Performance/cost:** Estimates of computing needs run faster than Moore's Law by factors of 3-30
- **Technology/Market evolution:** the return of heterogeneity; technology change will also make it challenging to exploit Moore's Law without software evolution.
- **Sustainability:** Most of the current software, which defines our capabilities, was designed 15-20 years ago: there are many software sustainability challenges.

Why Software? Software is *the* Cyberinfrastructure



Computer hardware is a consumable.
Software is what we keep, and invest in, over time.

Community White Paper (CWP)

- The CWP will identify and prioritise 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 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 so far)

Monday Plenary Session 1

08:00	Registration and Continental Breakfast	
	<i>SDSC Auditorium</i>	08:00 - 08:30
	Workshop Welcome	
	<i>SDSC Auditorium</i>	08:30 - 08:35
	Status of HSF and Community White Paper Plans	<i>Dr. Peter Elmer</i>
	<i>SDSC Auditorium</i>	08:35 - 09:00
09:00	LHC Future Challenges - Physics	<i>Mark Neubauer</i>
	<i>SDSC Auditorium</i>	09:00 - 09:20
	LHC Future Challenges - Software/Computing	<i>Ian Bird</i>
	<i>SDSC Auditorium</i>	09:20 - 09:35
	Challenges for the Fermilab neutrino and muon programs	<i>Robert Kutschke</i>
	<i>SDSC Auditorium</i>	09:35 - 09:50
	Software plans and challenges for Linear Collider	<i>Frank-Dieter Gaede</i>
	<i>SDSC Auditorium</i>	09:50 - 10:00
10:00	Future software challenges at Jefferson Lab	<i>Amber Boehnlein</i>
	<i>SDSC Auditorium</i>	10:00 - 10:15
	Google Summer of Code (GSoC) and the HSF	<i>Sergei Gleyzer</i>
	<i>SDSC Auditorium</i>	10:15 - 10:20

Monday Plenary Session 2

11:00	HEP experiments and software status in China <i>SDSC Auditorium</i>	<i>Prof. Wenjing Wu</i> 11:00 - 11:15
	CERN Openlab <i>SDSC Auditorium</i>	<i>Maria Girone</i> 11:15 - 11:30
	NSF View <i>SDSC Auditorium</i>	<i>Dr. Bogdan Mihaila</i> 11:30 - 11:50
12:00	DOE View <i>SDSC Auditorium</i>	<i>Abid Patwa</i> 11:50 - 12:10
	Other agencies view (survey) <i>SDSC Auditorium</i>	<i>Michel Jouvin</i> 12:10 - 12:30

Monday Plenary Session 3

14:00

LIGO Software and Computing	<i>Peter Couvares</i>
<i>SDSC Auditorium</i>	13:30 - 14:00
CERN Cloud and Facilities Outlook	<i>Helge Meinhard et al.</i>
<i>SDSC Auditorium</i>	14:00 - 14:20
The Fermilab HEPCloud, or How to add 240 TFlops in an hour or two	<i>Oliver Gutsche</i>
<i>SDSC Auditorium</i>	14:20 - 14:40
Technology overview	<i>Helge Meinhard</i>
<i>SDSC Auditorium</i>	14:40 - 15:00

Monday Plenary Session 4

16:00

OSS Licensing (TBC)	<i>Aaron Sauers</i>
<i>SDSC Auditorium</i>	15:30 - 15:45
Citation and Reproducibility in Software	<i>Daniel S. Katz</i>
<i>SDSC Auditorium</i>	15:45 - 16:15
Survey of Contributed White Papers	<i>Ian Fisk et al.</i>
<i>SDSC Auditorium</i>	16:15 - 17:00

Goals of the SDSC HSF Workshop

- Each Working Group (WG) should aim to produce two things by the end of the workshop.
 - ① **A WG charge, specifying the scope, challenges and questions to be answered**
 - ② **A plan for how the WG will work in the next 6 months: specific input to collect; videoconference meetings, workshops or workshop sessions planned, etc.**
- The aim at this workshop is discussion focused on those two items, and not necessarily beginning the full process of gathering information.
- Most of the action will take place in the parallel sessions.

Two Bonus Questions for the WG Parallel Sessions

- What connections and/or impact would meeting the challenges of this WG have on the global computing models?
- How can the CWP process be best organized to include the goals of each WG?

Panels

- Maria Girone (CERN Openlab) has organized three topical panels (one on Tue., two Wed.):
 - Panel: Compute Architectures, platforms and software performance - moderator David Abdurachmanov
 - Panel: Data Centres and Facilities Technologies (Clouds, Networks, Storage) - moderator Oliver Gutsche
 - Panel: Applying Machine Learning in HEP - moderator Jim Pivarski
- The panels consist of several people from our community, plus participation from industry

Agenda for Tuesday/Wednesday/Thursday

Switch to Indico: <http://indico.cern.ch/event/570249/timetable/>

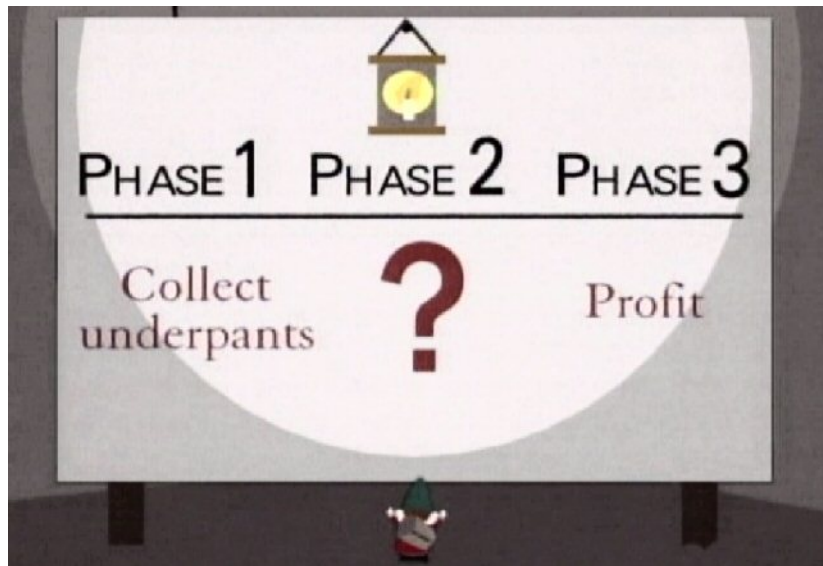
CWP process - contributed whitepapers

- The CWP process envisions contributions in the form of short white papers
- Contributions on all related topics are welcome.
- We made a specific call for white paper contributions regarding “Computing Models, Facilities, and Distributed Computing”
- A number of white papers have been received:
- <http://hepsoftwarefoundation.org/cwp-whitepapers.html>
- We will review what is there later today, so if you still have something to send, please send it this morning!
- We expect further contributions over the next 6 months.

CWP as roadmap

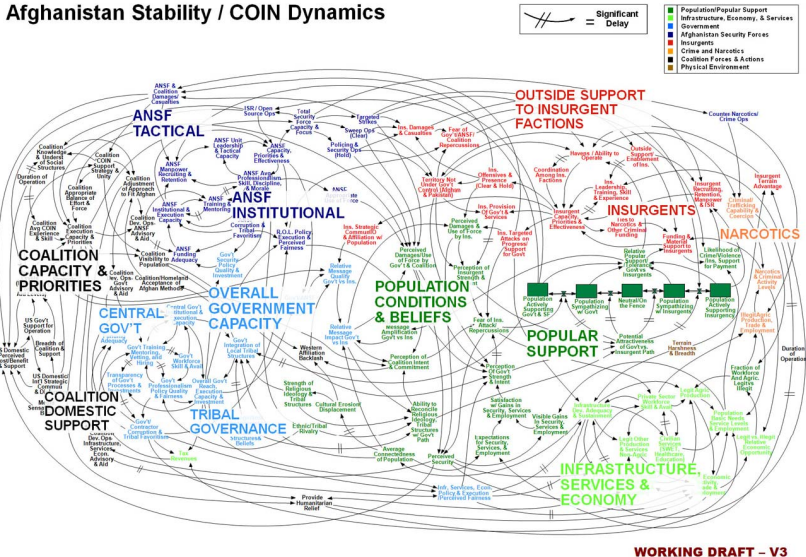
- The goal in 6 months is to produce “Roadmap for HEP Software and Computing R&D for the 2020s” and respond both to the WLCG charge:
- `http://hepsoftwarefoundation.org/assets/CWP-Charge-HSF.pdf`
- as the HSF goal of engaging more widely the HEP community (to everyone’s mutual benefit).
- We need to be clear about defining priorities for the projects we need, how we will be organized to accomplish the projects and how these things fit in a timeline (for the mid-2020s/HL-LHC era, but which runs through LHC Run 3 and other things in between)
- This is not a full specification of computing models for all HEP experiments. For HL-LHC (CMS/Atlas), for example, that will happen in 2019-2020 in the computing TDRs.

This plan is too simple



This plan is too complex

Afghanistan Stability / COIN Dynamics



WORKING DRAFT - V3

The rest of the CWP process

- From the WG plans produced in this meeting we will define the intermediate topical workshops and meetings which are required
- We are aiming for a final general CWP workshop at the end of June or beginning of July. Several locations in Europe might be possible. We hope to pick one specific tentative date by the end of this workshop.

Summary

- We have a significant investment in software, it embodies the core of our intellectual property and the real cyberinfrastructure.
- Significant challenges of scale, performance, technology and long term sustainability exist as we face the HEP projects of the 2020s.
- The workshop is the beginning of our “Community White Paper” process, a roadmap for HEP Software and Computing for the 2020s
- Let's see what we can do this week!