
Scalable Platforms

Summary

S2I2 HEP/CS Workshop
May 1-3, 2017, Princeton, University



Rob Gardner
University of Chicago

Doug Thain
University of Notre Dame

Scalable Platforms

- For our purposes, we used this definition:

"Scalable platforms aggregate resources from multiple machines to improve performance and increase capacity of processing and data"

Guidance for lightning talks



- What aspects of scalable platforms are of most interest? (**briefly** describe focus area)
- What are the big or new ideas?
- What are indicated R&D paths?
- What would you like to see in an Software Institute?
 - What sorts of CS-HEP collaborations are most useful?
 - How should activities of the Institute relate to the computing programs of HEP experiments?

Lightning Talks

PHY-ATLAS: Corraling Heterogeneous Systems Speaker: Gordon Watts (University of Washington (US))

PHY-CMS: Distributed Scalable Platforms Issues and Approaches Speaker: Harvey Newman (California Institute of Technology (US))

PHY-ATLAS: Services at the Edge Speaker: Robert William Gardner Jr (University of Chicago (US))

CS: Towards 1000x with Heterogeneous, Programmable Hardware Datacenter Speaker: Anton Burtsev (University of California - Irvine)

CS: Non-determinism in applications at the exascale: impact on debugging and numerical reproducibility Speaker: Michela Taufer (University of Delaware)

CS: Exploiting node-level parallelism at exascale Speaker: Prof. Sunita Chandrasekaran (University of Delaware)

CS-Astro: Missing Abstractions Speaker: Daniel S. Katz (University of Illinois)

PHY-ATLAS: Do we need distributed computing? Speaker: Kaushik De (University of Texas at Arlington (US))

PHY-Astro: Developing non-LHC software and our Services-based infrastructure dream Speaker: Christopher Tunnell (Enrico Fermi Institute-University of Chicago-Unknown)

CS: Fundamental Problems of Distributed Systems Speaker: Douglas Thain

IND: Learnings from Industry: Tooling, Datasets, Productivity, & Software Quality @SoundCloud Speaker: Meghan Kane (SoundCloud)

Themes (c.f. [live notes](#))

- Declarative abstractions & missing abstractions
- CS-HEP terminology mappings (again)
- Edge capabilities including SDN to aggregate resources
- Future systems, challenge of parallelism and latency
- Reproducibility of non-deterministic code at exascale
- Directive-based software to exploit heterogeneous architectures
- Questioning fundamental assumptions about the distributed computing model
- Broader impacts, re-use perspectives from non-LHC experiments
- Services model w/ building blocks with clear bindings
- Lessons from industry - importance of staff education as a continuous activity "up-skilling"
- Helping CS collaborate - benchmark code, datasets
- Institute as meeting point between CS research and physics systems at scale
- Role for students at both MS and Ph.D. levels; fellowships

Questions (4/53) that came up

- How do we express HEP computing processes in CS language?
 - "skimming, slimming, thinning" and "production & analysis"
 - RAW, ESD, AOD, xAOD, NTUP, ...
- How do we generalize HEP and non-HEP solutions?
- How to express computing model abstractions to expose tradeoffs
- How to continue to collaborate?
 - No interest in "refresh workshops" --> Need to maintain dialog
 - Need context to establish credibility to try out ideas, create realistic prototypes informed by experiment needs
 - Through training - e.g. hackathons w/ real codes on real infrastructure

Charge questions (ran out of time!)

Are there places where the HEP language to describe a problem or system doesn't match how a CS person would describe the same problem?

Yes. Ex: Production/Analysis, Slimming/Skinning/Thinning

Which things does the HEP community want to do, does not know how, and believes that Computer Scientists may be able to figure out?

- Novel datacenter architectures
- Exploiting accelerator devices

How do these problems map to CS research questions?

The database community could have approaches to deal with declarative languages.

Are there CS workshops/meetings on these questions which HEP people should consider attending?

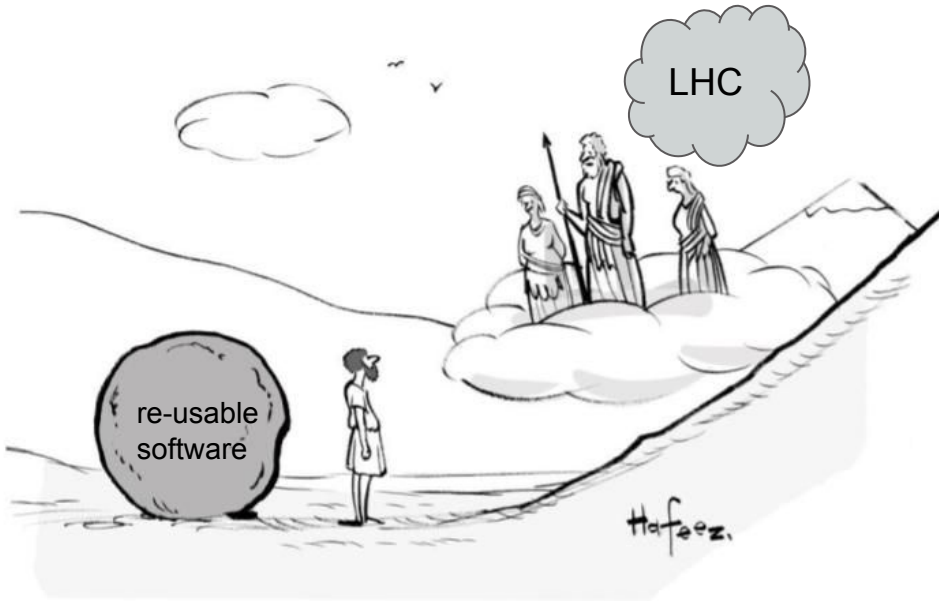
<https://craft-conf.com/>

<http://events.linuxfoundation.org/events/containercon/program/about>

A place that has a large concentration of influencers at an exciting event in software development

- Events taking place around NIPs if NIPs is too expensive
- Events taking place around Apple's WWDC

Chris Tunnel (Xenon1T)



Thanks to
contributors and
participants

"Professional stone-rollers would have it to the top in two days. For a beginner like you it'll take about a week."