



Panel Discussion S2I2 Conceptualization for HL-LHC

**Frank Würthwein
OSG Executive Director
Professor of Physics
UCSD/SDSC**



The Panelists

- Mark Neubauer
 - Prof. of Physics @ UIUC
 - Co-PI of S2I2 Conceptualization Award from NSF
- Oliver Gutsche
 - Scientists @ FNAL
 - Deputy Lead of US CMS S&C Ops Project
- Peter Elmer
 - Research Scientist @ Princeton
 - Co-PI of S2I2 Conceptualization Award from NSF
- Kaushik De
 - Prof. of Physics @ UTA
 - Deputy Lead of US ATLAS S&C Ops Project

Q1 to Kaushik & Oli

- The NSF cooperative agreement that funds the US LHC S&C projects was just renewed for a third 5 year term starting with January 2017.
- What are your biggest concerns at present as you look at your next 5 years in this program? How do you think a software institute might be able to help with these challenges?

- the cornerstone of the NSF cooperative agreement is the Tier 2 program. Four Tier 2 superclusters are supported by the NSF for the next 5 years. They depend on the OSG for software packaging, integration and support. This is of course crucially important. This is 3/4th of the NSF program budget. Remainder is mostly application software and support, with a small amount for Tier 3's where OSG can also play an important role. We need OSG and the future S2I2 to play a big role in supporting the software stack that makes distributed computing possible, to make full use of our T2/T3 facilities. Big challenges: (1) Resource shortage. HPC and other opportunistic resources can help. (2) Changing middleware. OSG/S2I2 needs to work with experiments directly to make sure specific needs and opportunities are met. (3) HL-LHC challenge. Resources cannot match needs if computing model and algorithms remain static. Of course, scope of S2I2 is much larger. I only address this discussion in the context of the cooperative agreement for the next 5 years, and within the framework of OSG AHM.

Oliver Gutsche (U.S. CMS)



■ Biggest concerns of U.S. CMS

- Keep on running!

- Keep on evolving!



U.S. CMS - Keep on running!

- LHC has a 10-year track record in bringing together a large international community and enable this community to use computing resources across geographical and administrative borders in a seamless way → this is a big success by itself, OSG is a major part of this success.
- Through OSG, we expanded this community in the US and are even more successful.
- This does not come for free, there are many solutions and services that are needed to make this work → all this depends on infrastructure software and its integration and coordination with others
- The integration and maintenance of the community software stack needs to be continued to enable access to resources at the sites and forming the common infrastructure we need to do physics.
 - This can be either a renewed OSG or the NSF software institute or somebody else.



U.S. CMS - Keep on evolving!

- All the success is based on continuous evolution, developing new capabilities and updating outdated and under-performing solutions to the benefit of the community
- For HL-LHC, we need to close the resource gap
 - NSF MREFC and DOE funds new detectors need to do physics at HL-LHC
 - To harvest physics with these new detectors at the HL-LHC, we need to reach unprecedented scales and therefore need advances in application software, infrastructure software and computing resources
- For HL-LHC, we need software to exploit
 - Heterogeneity of computing resources, architectures and algorithms (clouds, HPC, ARM, data management, ...)
 - New techniques and solutions (machine learning, massive multi-threading for GPUs, KNL, ...)
 - All in a flexible way so that we can use whatever comes new in the years to come in these areas
- The benefit here is that new capabilities and solutions can be used by the whole community, so the whole is larger than the sum of the parts.
- A software institute could be a good partner to tackle these future needs.

Q2 to Mark & Pete

- What has happened so far in the S2I2 conceptualization process, and are there already emerging themes that you see coming out of it ?

CWP Goals

- 1) to achieve improvements in software efficiency, scalability and performance and to make use of the advances in CPU, storage and network technologies
- 2) to enable new approaches to computing and software that could radically extend the physics reach of the detectors
- 3) to ensure the long term sustainability of the software through the lifetime of the HL-LHC

Past Events:

- 23-26 Jan, 2017 - HEP Software Foundation Workshop
 - *University of California at San Diego / San Diego Supercomputer Center*
 - Indico page
- 7-9 Dec, 2016 - S2I2 HEP/CS Workshop
 - *University of Illinois at Urbana-Champaign*
 - Event page, Indico page
 - Workshop summary report (pdf)

<http://s2i2-hep.org>



Events planned

Upcoming Events:

- 8 Mar, 2017 - S2I2-HEP/OSG/USCMS/USAtlas Panel at OSG All Hands Meeting
 - [Indico page](#)
- 9 Mar, 2017 - Software Triggers and Event Reconstruction WG meeting
 - A CWP session at the [Connecting The Dots workshop](#)
 - [Indico page](#)
- 20-22 Mar, 2017 - IML Topical Machine Learning Workshop
 - *CERN*
 - The workshop includes a CWP session on Machine Learning
 - [Indico page](#)
- 28-30 Mar, 2017 - CWP Visualization Workshop
 - *CERN (and Vidyo)*
 - [Indico page](#)
- early May, 2017 - 2nd S2I2 HEP/CS Workshop (exact dates TBD)
 - *Princeton (TBC)*
- 22-24 May, 2017 - HEP Analysis Ecosystem Retreat
 - *Amsterdam (TBC)*
 - [Indico page](#)
 - [Workshop proposal](#)
- 5-6 Jun, 2017 (TBC) - CWP Event Processing Frameworks Workshop (TBC)
 - *FNAL*
 - The workshop is just prior to the FNAL 50th Anniversary and User Meeting
- 26-30 Jun, 2017 - HEP Software Foundation Workshop
 - *LAPP (Annecy)*
 - [Indico page](#)

<http://s2i2-hep.org>

March 8th, 2017

Q3 to all 4 panelists

- Close to 1/3 of the OSG effort goes into the software life cycle management of an integrated middleware stack that then gets deployed by the two S&C projects at their compute and storage facilities across the US and latin america. This lifecycle includes evaluation, integration, performance benchmarking, adoption into orphanage if necessary, replacement of functionality by a different software, retirement from stack.
- The community thus has seen this practiced for the last 10 years or so with major pieces of software like e.g. globus GRAM going through the entire cycle.
- In addition, each experiment has its own software lifecycle management process for its application software, also with a long history.
- **How do you see the software lifecycle being managed in a future software institute? How do you see it interact with the software lifecycle management in the experiments? In OSG?**



Kaushik De (US ATLAS)



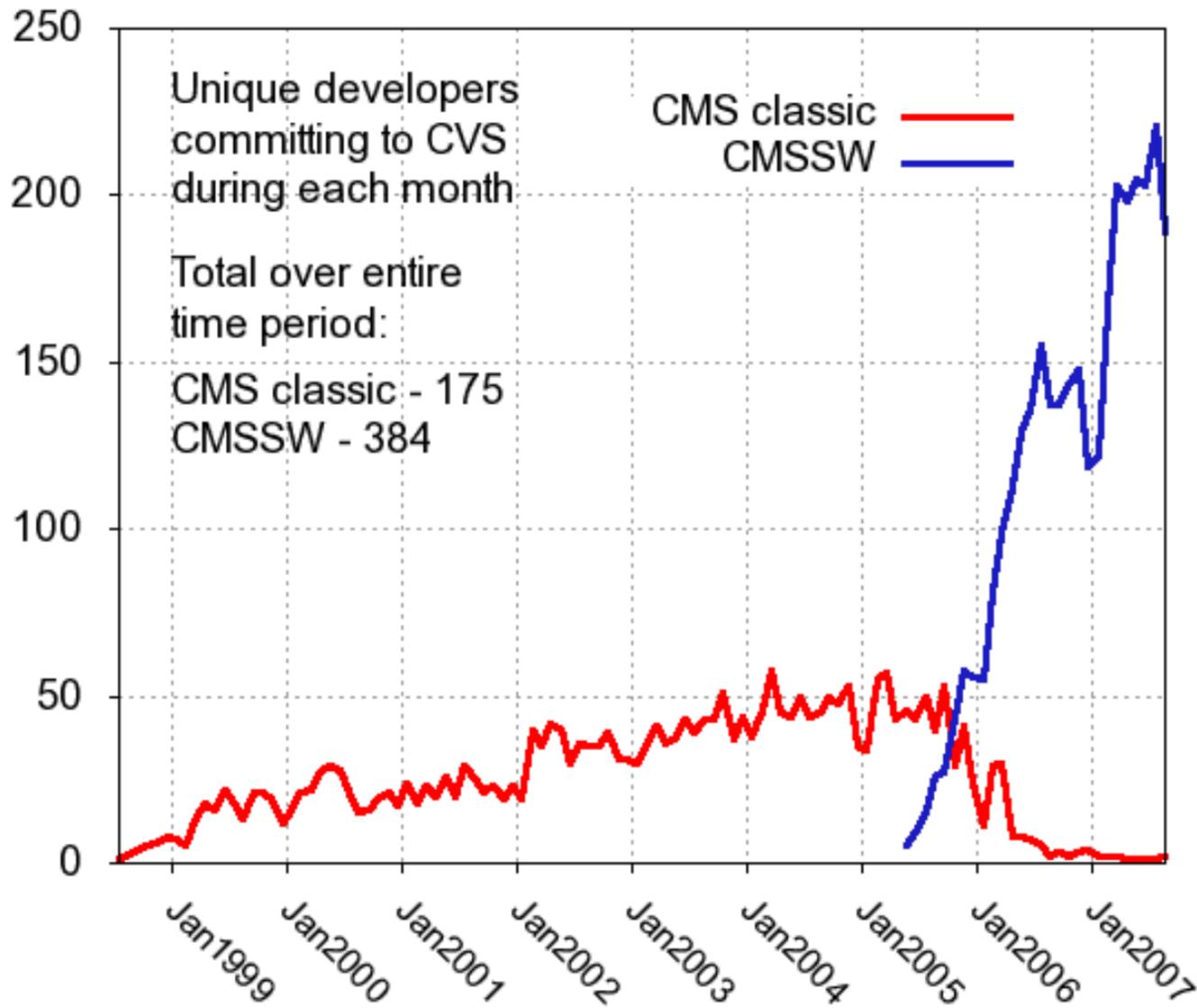
- S2I2 is in conceptualization phase, so nothing is set in stone. We can learn from the past 10 years of OSG experience. And from the experience of LHC experiments during Run 1 and 2. The S&C activities in the LHC experiments are very different from the S&C model imagined before the experiment started. A lot of evolution in thinking, infrastructure, capabilities... have been driven by realities. We should not start S2I2 from scratch, but informed by the real past experience. There are many challenges. We need well maintained, stable, sustainable ... software to meet the physics challenges of Run 3 and the HL-LHC. Run 3 is not expected to be a big challenge - but we can use it as test bed for future software. S2I2 is defined to be “long-term hubs of excellence in software infrastructure and technologies” - both OSG and the LHC experiments have over a decade of experience.

- I hope a software institute would have a structured stakeholder process with requirement gathering and a transparent prioritization process to serve the whole community (kind of like OSG).
- The experiments would be stakeholders and their needs and requirements would be heard in the process.
- National and international coordination is key.
- In a collaboration (for example CMS) this partnership is well defined:
 - Guiding principle: “Your concerns are our concerns” → institutes/groups come together working on the same goal
- To tackle the HL-LHC problems in larger groups spanning experiments / collaborations / projects, we need a similar setup to work on the same goals and to make this a sustainable success.

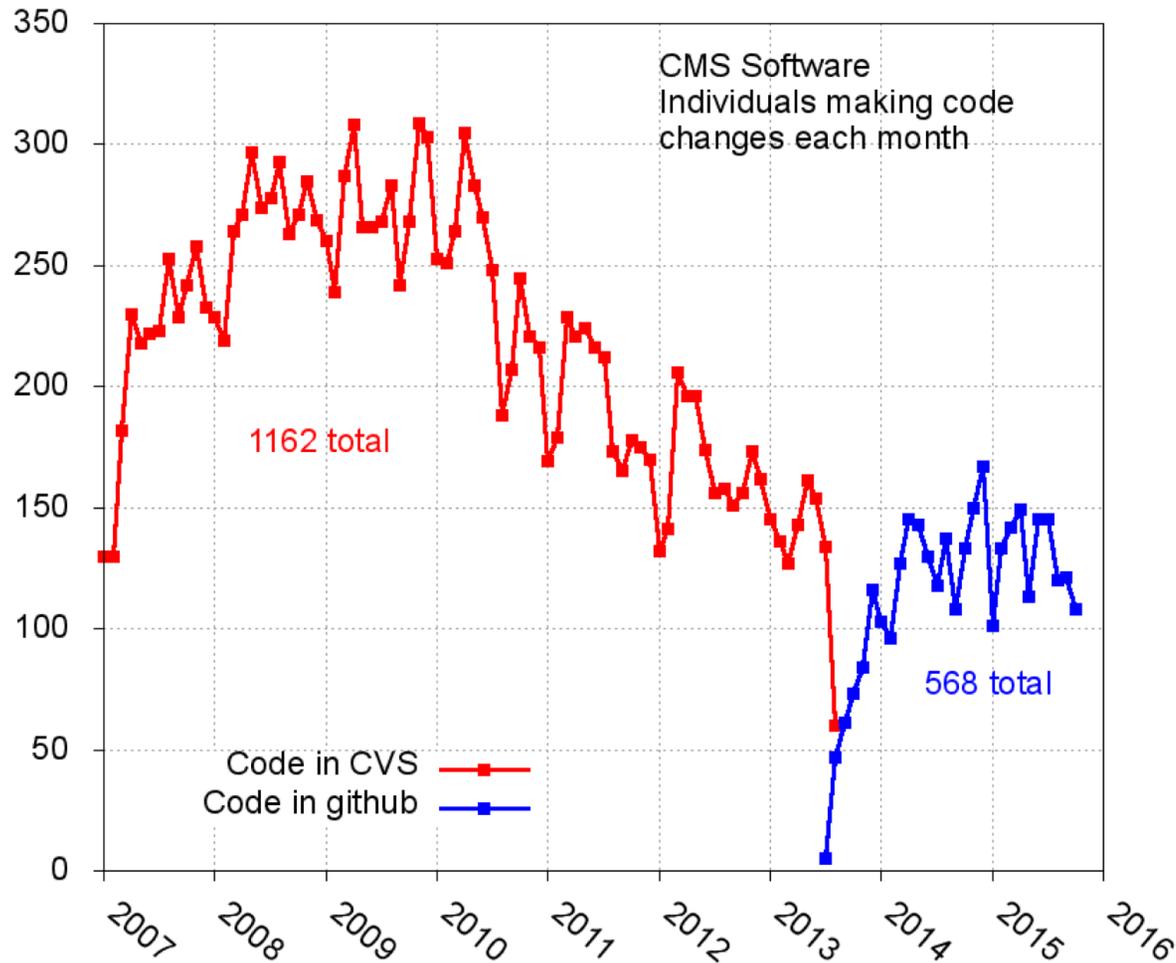
- As FKW has identified, managing a software lifecycle process will be an important aspect of any eventual software institute
- US-Atlas (Atlas), US-CMS (CMS) and OSG have all learned to do this in their own domains. We know the basic elements of how to do this.
- We have to build on and extend our (extensive) existing experience



Software Process



Software Process



Migrating CMS software from CVS to github