Developing cutting-edge ICT systems through public-private partnership

iSGTW speaks to Alberto Di Meglio as CERN openlab enters its fifth phase

FEATURE | FEBRUARY 4, 2015 | BY ANDREW PURCELL

CERN openlab is a unique public-private partnership between CERN and leading ICT companies. Founded in 2001 to develop the innovative ICT systems needed to cope with the unprecedented computing challenges of the Large Hadron Collider (LHC), CERN openlab unites science and industry at the cutting edge of research and innovation.

At the start of this year, CERN openlab officially entered its fifth phase, which will run until the end of 2017. Now, for the first time in its history, it has extended beyond the CERN community to include other major European and international research laboratories.

iSGTW *speaks to Alberto Di Meglio, head of CERN openlab...*

"It is absolutely vital for Europe that that there is a continuous exchange of information and expertise between research and industry," says Di Meglio. "CERN openlab plays a central role in this vision." Image courtesy CERN openlab.

How does CERN openlab foster and accelerate the development of cutting-edge ICT systems?

CERN openlab is a joint collaboration between industrial companies and CERN's engineering and scientific teams. Together, we work on developing new technologies — on a time horizon of three-to-five years — and evaluate them for the LHC program.

The companies provide us with new ideas and prototypes; we then use our expertise here at CERN to examine these technologies in detail and contribute to their evolution. Technologies targeted at the needs of CERN today are likely to become mainstream products for enterprise or even consumer markets just a few years down the line.

Our collaboration with leading ICT companies relies on a virtuous cycle. We start by outlining our technical needs and we then try to find good matches between these and the technologies the companies are currently working on. We may then work together on developing prototypes or installing technologies in large-scale production-like environments for the first time. We provide the companies with detailed feedback and propose possible solutions, fixes, or enhancements where necessary.

Can you give some examples of successful systems developed through CERN openlab?

We've been working with world-leading ICT and technology companies for 12 years through CERN openlab. Many technologies have been evaluated as part of the collaboration and have now become products used in the LHC research program. There are, of course, also many technologies that have not been incorporated into the LHC research program, but these have still benefited from our testing and our feedback.

One of the very first success stories of CERN openlab came from our collaboration with Intel and HP. We worked together to develop a grid-enabled compute and storage farm called the CERN opencluster. This investigation was important to connect us to the Fermi National Accelerator Laboratory in Batavia, Illinois, US. At the time, it was so advanced that it broke the record for the rate of data-transfer over the network.

We've continued to do work that's right at the leading edge of technology since then, too. For example, we were the first to port some important grid software for use on 64-bit (rather than 32-bit) operating systems and to work on code optimization on new multicore platforms from Intel. We've worked with Siemens on some very advanced control systems, with Oracle on database scalability and replication solutions (these have since become major commercial products and are now deployed at CERN, as well as at WLCG tier 1 sites), with Huawei on cloud storage appliances, and with Yandex on applications of web data analytics to physics analysis. A good example today is our work with Rackspace on cloud federations: the software we've developed together has been fed back into the OpenStack collaboration, where it's now part of their standard distribution.

Are there other advantages to such close collaborations between public research labs and industrial companies?

Education is a major mission of CERN and CERN openlab makes a significant contribution to this. The researchers who work with us in CERN openlab are normally very young, having come straight from university. They take part in projects for two or three years and work directly with both CERN and the companies to develop new cutting-edge technologies. Over time, their training enables them to become experts in these technologies. At the end of their projects, some of the researchers stay at CERN, but most either go back to academia or go to work with commercial companies. In this way, CERN openlab is able to transfer both knowledge and ideas to the market.

We also have a program of seminars and workshops, as well as a highly successful summer student program that's open to participants from all parts of the world. Applications are currently open for this year — it's a great opportunity!

What was the driving force behind expanding CERN openlab for its fifth phase to include other European research labs?

Today, research centers in other disciplines are also starting to produce very high quantities of data at very high speed. The idea with the new phase of CERN openlab is to understand together — across disciplines — the challenges we all face and how can we collectively address them. By pooling our knowledge and expertise, we hope to build a foundation to cover the entire data-

management spectrum, from data acquisition, through simulation and analysis, and all the way to storage, compute provisioning, and networking.

The expansion of CERN openlab isn't just on the public research side: we're looking for new industrial partners, too. We're currently developing new ways for smaller companies — those with innovative, disruptive ideas — to participate.

What are the major ICT challenges at CERN you expect to be tackled during this fifth phase?

We started the process of defining the objectives for the fifth phase of CERN openlab almost a year and a half ago. We began by collecting requirements from various technical and scientific teams at both CERN and other research laboratories. Based on these, we published a whitepaper last year that describes six major areas to be addressed: data acquisition, computing platforms, data storage architectures, compute management and provisioning, networks and connectivity, and data analytics.

And how about at the other labs?

Up to a certain level, all 'big science' research follows the same basic pattern of data acquisition, analysis, computing, *etc.* However, different disciplines have different priorities in terms of their needs. Things are very centralized here in the high-energy physics community: data is generated by the experiments at CERN and then distributed across the world for analysis. By contrast, the model often used in biomedical research is almost the exact opposite of this: data is generated by thousands of distributed instruments (for example, genomic analyzers) and is then brought together for analysis. It's important to understand how the technologies developed by ICT companies can cope with such wildly varying models.

Another important difference is that we don't have to worry too much about data confidentiality in the high-energy physics community, whereas this is a paramount concern for those working in medical fields. These sorts of requirements also need to be considered when developing data-management infrastructures.

Finally, what is your vision for the future of CERN openlab?

While we're now aiming to play an important role in helping other research laboratories, supporting the LHC research program continues to be our primary goal. The new technologies we investigate in the fifth phase of CERN openlab will potentially be incorporated into the upgrade plan for the LHC during its second scheduled long-shutdown period, which will start in 2018.

It is absolutely vital for Europe that that there is a continuous exchange of information and expertise between research and industry. CERN openlab plays a central role in this vision.

Average:

Your rating: None Average: 4.3 (3 votes)

About the Author »

Andrew Purcell

Editor

Andrew Purcell is the editor of iSGTW and is based at CERN, near Geneva.

RELATED TERMS: CERN CERN openlab computing Europe high-energy physics ICT IT LHC scientific software data management systems workflow management systems physics and astronomy

Comments

ADD NEW COMMENT

Post new comment

Subject:

Comment: *

Input format

By submitting this form, you accept the Mollom privacy policy.

SAVE PREVIEW