



[News](#) [HPC Hardware](#) [HPC Software](#) [Industry Segments](#) [White Papers](#) [Resources](#) [Special](#)

Sign up for our newsletter and get the latest HPC news and analysis.

Email Address

AdChoices

[Home](#) » [HPC Hardware](#) » Addressing Computing Challenges at CERN openlab

Addressing Computing Challenges at CERN openlab

July 22, 2018 by [staff](#) [Leave a Comment](#)

In this special guest feature from [Scientific Computing World](#), Robert Roe speaks with Dr Maria Girone, Chief Technology Officer at CERN openlab ahead of her keynote presentation at [ISC 2018](#)

As the CERN organization prepares for 'high-luminosity' experiments in 2026, the organization faces the significant challenge of supplying a computing infrastructure that can handle the huge amount of data generated. To overcome this, CERN openlab is working on a mixture of novel new approaches to data handling and the use of commercial cloud companies that can help expand the capacity available to CERN researchers.



Maria Girone is CTO of CERN openlab

The challenge of creating the largest particle accelerator is now complete but there is another challenge – harnessing all of the data produced through experimentation. This will become even greater when the 'high-luminosity' LHC experiments begin in 2026.

The demands of capturing, storing, and processing the large volumes of data generated by the LHC experiments may be as much as 50 to 100 times larger than today, with storage needs expected to be in the order of exabytes.

FEATURED JOB

**Software and Hardware
Architects
University of
Southern Denmark
(SDU)
Odense**

[Learn More »](#)

Other Jobs

[High-Performance
Computing \(HPC\)
Computational Scientist](#)

[HPC Systems
Administrator](#)

[Fellowships for the
Master in High
Performance
Computing](#)

[See all Jobs](#) | [Post a Job](#)

CERN is working to tackle many of these challenges together with ICT industry leaders, through a public-private partnership known as '[CERN openlab](#)'. Maria Girone, chief technology officer at CERN openlab, discuss the work being done to prep CERN's computing infrastructure for future LHC experiments ahead of her keynote at ISC High Performance in Frankfurt in July.

SCW: What are the scale of the ICT

challenges faced by CERN?

Maria Girone: CERN is home to the Large Hadron Collider (LHC), the world's largest and most powerful particle accelerator. Built in a 27km-long tunnel, about 100m underground the Franco-Swiss border, the LHC helps scientists to unlock the secrets of the universe.

The particles within the LHC are made to collide at close to the speed of light. This gives the physicists clues about how the particles interact, and provides insights into the fundamental laws of nature.

Within the LHC, there are up to one billion particle collisions per second. Custom hardware ASICs filter this down to approximately 100,000 collision events, which are then sent for digital reconstruction. More detailed algorithms whittle this down to around 100 'events of interest' per second.

In 2017, this process resulted in 40 petabytes of data being sent to the main CERN data centre, which is where archiving and processing takes place. This data centre – together with its remote extension in Hungary – hosts 230,000 processor cores and 15,000 servers. Today over 230 petabytes are permanently archived on tape.

Researchers at CERN have developed custom-built disk and tape systems that can scale to huge capacity and are capable of delivering data at a rate of petabytes per day.

SCW: What is the approach to handling the data produced from LHC?

Maria Girone: The Worldwide LHC Computing Grid (WLCG) is used to store, distribute, and analyse this enormous volume of data. The WLCG is the largest collection of computing resources ever assembled for a single scientific endeavour. It consists of more than 800,000 processor cores, with 400 petabytes on disk and 500 petabytes on tape. There are around 170 sites in the WLCG, located across all continents except Antarctica.

The networks of the WLCG move roughly two petabytes of data daily. The WLCG differs from more traditional high-performance computing centres in that it is much more distributed, with applications well adapted for parallelisation and for running independently.

Multiple LHC experiments have also carried out tests to examine the feasibility of expanding into commercial cloud resources for processing data. I recently led demonstrations that showed we could double the processing resources available to a single experiment for bursts.

SCW: *What is CERN openlab?*

Maria Girone: CERN openlab is a public-private partnership between CERN and leading ICT companies. Our mission is to accelerate the development of cutting-edge ICT solutions for the worldwide LHC community – as well as for wider scientific research.



Through CERN openlab, CERN provides access to its complex ICT infrastructure and its engineering experience. Testing in CERN's demanding environment provides the ICT industry collaborators with valuable feedback on their products, while enabling CERN to assess the merits of new technologies in their early stages of development for possible future use. The framework also offers a neutral ground for carrying out advanced research-and-development activities with multiple companies.

SCW: *CERN openlab was first initiated 2001. How has its role changed over this time?*

Maria Girone: Since its establishment in 2001, CERN openlab has been organised into successive three-year phases. Working in this manner provides us with regular opportunities to assess our progress and to ensure that we are always working to tackle the ICT challenges most relevant to the LHC research community.

Our first phase focused specifically on the development of an advanced computing-cluster prototype, but since then our work has expanded to incorporate a much wider range of domains, including work related to grid computing, networking, virtualisation, industrial control systems, new computing architectures, and much more.

Our fifth three-year phase came to a close at the end of 2017. This phase featured around 100 R&D projects; these tackled ambitious challenges covering the most critical needs of ICT infrastructures in domains such as data acquisition, computing platforms, data storage architectures, compute provisioning and management, networks and communication, and data analytics.

During our fifth phase, CERN openlab also grew to include more collaborating companies, thus enabling a wider range of ICT challenges to be addressed. For the first time, other research institutes also joined CERN openlab in our endeavour to accelerate the development of cutting-edge ICT solutions for science. Working together with other laboratories to tackle common ICT challenges is both highly useful and helps to ensure maximum relevancy of our work.

As well as this technical work, CERN openlab also carries out training and educational activities. This is an area that has really grown over the past decade and a half. For example each year we run a nine-week programme for about 40 students. Bachelor's and master's students specialising in subjects related to computer science come to CERN to work on cutting-edge projects with our collaborators. These projects often feature hands-on practice with the latest ICT solutions.

SCW: *As CERN prepares for 'the High-Luminosity LHC' in 2026 how will the computing requirements change?*

Maria Girone: CERN openlab's new three-year phase, which started at the beginning of 2018, aims to address many of the ICT challenges posed by the High-Luminosity LHC (HL-LHC). Last September, we published a white paper highlighting many of the key challenge areas that we believe are ripe for tackling together with our industry collaborators.

A series of planned upgrades to the LHC will result in the HL-LHC coming online in around 2026. This will crank up the performance of the LHC significantly and will increase the potential for discoveries. The higher the luminosity, the more collisions, and the more data the experiments can gather.

An increased rate of collision events means that digital reconstruction of collisions becomes significantly more complex. At the same time, the LHC experiments plan to employ new, more flexible filtering systems that will collect a greater number of events.

This will drive a huge increase in computing needs. Using current software, hardware, and analysis techniques, the estimated computing capacity required would be around 50 to 100 times higher than today. Data storage needs are expected to be in the order of exabytes by this time. Technology advances over the next seven to 10 years will likely yield an improvement of approximately a factor 10 in both the amount of processing and storage available at the same cost, but will still leave a significant resource gap. The solution to this is likely to be found in a series of hard-fought improvements and technology migrations – each contributing small advances – rather than a single revolutionary change.

In order to achieve the required scale, we will have to make use of a mix of new hardware architectures and new computing techniques. We do not have existing proof of concepts, we do have examples from industry related to improvements in data analytics, and techniques like machine learning, that we think can help. High-performance hardware architectures like GPUs, FPGAs, and the next generation of CPUs also offer the potential to dramatically improve some applications.

Tying this heterogeneous system together and equipping it with the latest software techniques is going to be a huge R&D challenge over the next few years. Innovation is therefore vital; we believe that working together with leading ICT companies through CERN openlab can play a very important role in helping us to overcome the challenges we face.

This story appears here as part of a [cross-publishing agreement](#) with [Scientific Computing World](#)

[Sign up for our insideHPC Newsletter](#)



Related Content:

- [Video: Computing Challenges at the Large Hadron Collider](#)
- [Maria Girone from CERN openlab to Keynote at ISC 2018](#)
- [Why Storage Matters to Scientists](#)
- [SKA and CERN Sign Big Data Agreement](#)
- [Why the HPC Industry will Converge on Europe at ISC 2016](#)

Filed Under: [Events](#), [High Performance Analytics](#), [HPC Hardware](#), [HPC Software](#), [Industry Perspective](#), [Industry Segments](#), [News](#), [Research / Education](#), [Resources](#), [Storage](#) Tagged With: [CERN](#), [CERN openlab](#)

Leave a Comment

 Name * Email * Website

Notify me of follow-up comments by email.

Notify me of new posts by email.

Resource Links:

Post your job openings to the insideHPC job board. - [More Info Here](#)

Download the latest HPC white papers for free from the insideHPC White Paper Library. - [More Info Here](#)



[About insideHPC](#)

[Contact](#)

[Advertise with insideHPC](#)

Copyright © 2018