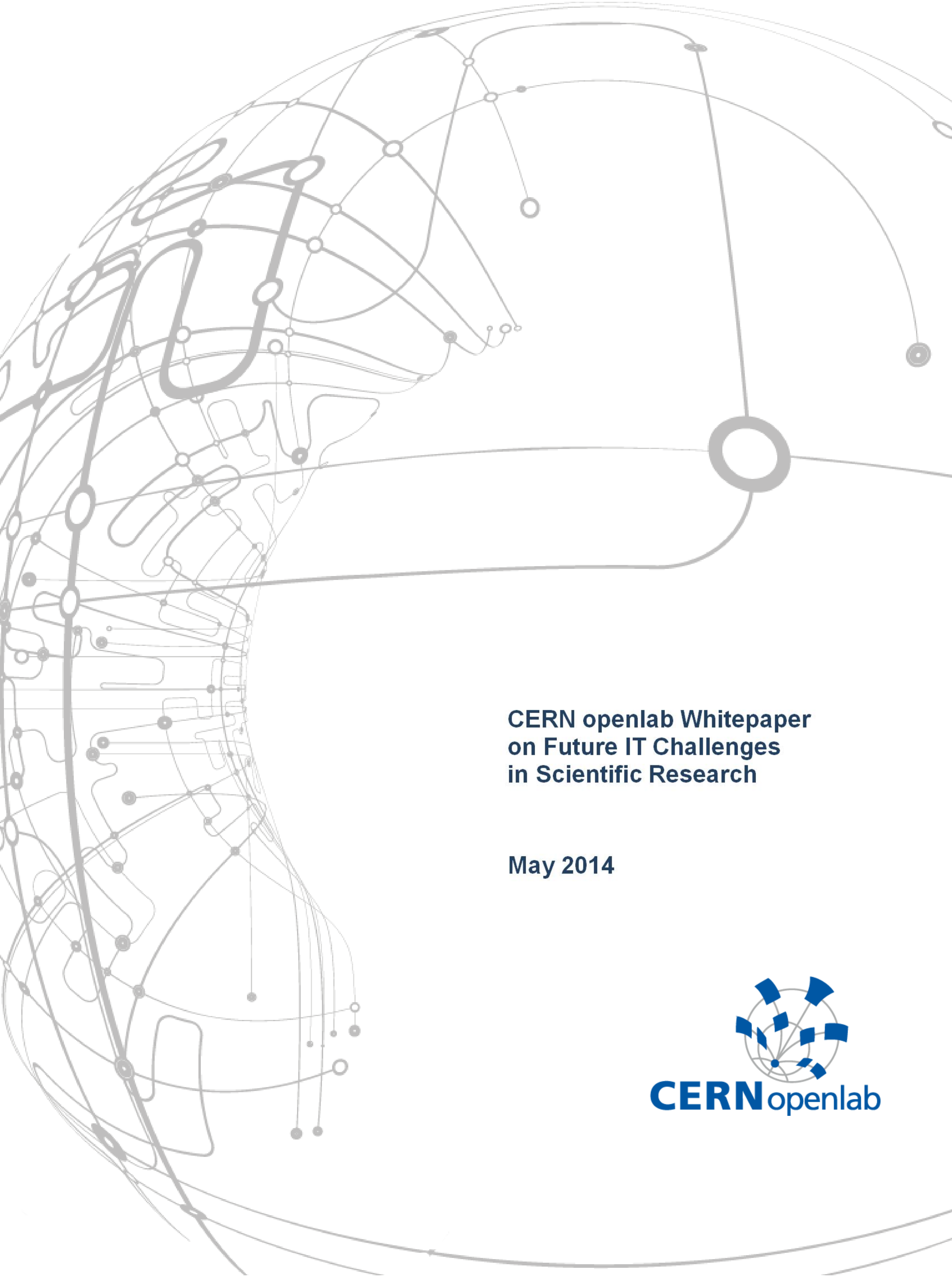


CERN openlab Whitepaper on Future IT Challenges in Scientific Research

May 2014

V. 1.3
Publication prepared by CERN openlab on behalf of the contributors.
DOI: http://dx.doi.org/10.5281/zenodo.8765

This whitepaper is released under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Photos: Shutterstock (cover, p. 38), CERN Photo Service (first five images at the top of p. 6), EMBL (from the top, images 6, 7, 8 p. 6), ESA (image 9 p. 6), ESA/ATG medialab (image 10 p. 6), ESA/ROB (image 11 p. 6), ESA/S. Corvaja (image 12 p 6), ESRF (image 13 p. 6), Ecliptique (image 14 p. 6), ILL/P. Ginter (image 15 p. 6), Human Brain Project (images 16, 17, 18 at the bottom of p. 6), Veronika McQuade p. 14 and p. 48, Roger Claus p. 18 and 26, Kaustav Ghose p. 22, Sylvain Fasel p. 32.

Contributors

Salim ANSARI. European Space Agency
Olof BARRING, CERN, IT Department
Tim BELL, CERN, IT Department
Ian BIRD, CERN, IT Department
Nenad BUNCIC, EPFL - Human Brain Project
Predrag BUNCIC, CERN, ALICE
Federico CARMINATI, CERN, PH Department
Tony CASS, CERN, IT Department
Sebastien CEUTERICKX, CERN, IT Department
Brice COPY, CERN, EN Department
Eva DAFONTE PÉREZ, CERN, IT Department
Marek DENIS, CERN, IT Department
Alberto DI MEGLIO, CERN openlab
Dirk DUELLMANN, CERN, IT Department
Ian FISK, FNAL, CMS
Mélissa GAILLARD, CERN openlab
Philippe GAYET, CERN, EN Department
Ioannis GEORGOPOULOS, CERN openlab
Maria GIRONE, CERN, CMS
Piotr GOLONKA, CERN, EN department
Manuel GONZALEZ BERGES, CERN, EN Department
Eric GRANCHER, CERN, IT Department
Kristina GUNNE, CERN openlab
Johannes GUTLEBER, CERN, Future Circular Colliders
Frédéric Hemmer, CERN, IT Department
Sverre JARP, CERN openlab
Bob JONES, CERN openlab
Sami KAMA, Southern Methodist University (US), ATLAS
Massimo LAMANNA, CERN, IT Department
Bruno LEBAYLE, ESRF
Edoardo MARTELLI, CERN, IT Department
Manuel MARTIN MARQUEZ, CERN, IT Department
Frans MEIJERS, CERN, CMS
Niko NEUFELD, CERN, LHCb
Steven NEWHOUSE, EMBL-EBI
Andrzej NOWAK CERN openlab
Alberto PACE, CERN, IT Department
Bernd PANZER-STEINDEL, CERN, IT Department
Jean-François PERRIN, ILL
Andrew PURCELL, CERN, IT Department
Fons RADEMAKERS, CERN, PH Department
Graeme Andrew STEWART, University of Glasgow (GB), ATLAS
Filippo TILARO, CERN, EN Department
Axel VOITIER, CERN, EN Department

CERN openlab has also collected input from the existing CERN openlab industrial sponsors. The views expressed in this document are those of the authors and do not necessarily reflect the view of their organisations and/or affiliates.

Executive summary

CERN openlab is a unique public-private partnership between CERN and leading IT companies. It was created in 2001 in support of the ambitious computing and data management goals set by the construction of the Large Hadron Collider (LHC) and detectors. Building on more than 10 years of ground-breaking work, CERN openlab continues to address the key topics in the CERN scientific and technical programme driven by the planned LHC upgrade activities spanning the next 20 years. The underlying principle behind the successful history of CERN openlab is the mutual benefit that CERN and the openlab industrial partners derive from the collaboration. CERN gets early access to new technologies and the companies have the unique possibility of testing their upcoming products on CERN's very challenging IT infrastructure.

In the past several years, scientific research has seen a dramatic rise in the amount and rate of production of data collected by instruments, detectors, or sensors. The LHC detectors at CERN produce a staggering 1 PB of data per second, a figure bound to increase during the next LHC run starting in 2015 and even more in future generations of colliders being discussed as part of initiatives like the Future Circular Collider (FCC) study. New international research infrastructures are being deployed by international and European laboratories like ESA, EMBL, or ILL and are expected to produce comparable—or even greater—amounts of data in diverse scientific domains, such as neurology, radio astronomy or genetics, produced by satellite imaging devices, high-performance genomic sequencers, neutron diffractometers or x-ray antennas.

For the years to come, the main challenge in scientific research include collecting and analysing such amounts of data to find evidence of new scientific discoveries; performing accurate and efficient simulations of the instruments and the underlying physical or biochemical processes; providing manageable, cost-effective, secure, large-scale computing infrastructures; and sharing the data across hundreds of research institutes and thousands of scientists and engineers.

The production of such quantities of data in many different formats, the availability of newer multi-core computing platforms, and the increasing need to cross scientific domain boundaries (e.g. in areas like High-Energy Physics, or HEP, for medical applications) require the emergence of new professional profiles. It is vital that new generations of scientists and engineers are formed with adequate skills and expertise in modern parallel programming, statistical methods, data analysis, efficient resource utilisation, and a broader understanding of the possible connections across seemingly separate knowledge fields.

The ever-increasing usage of the World Wide Web, created at CERN, and the advent of data-intensive consumer-oriented services have started generating and moving quantities of data in the order of hundreds of PB each month. Technologies that today are at the bleeding edge of research will be commodity items tomorrow. The continuous collaboration between the research infrastructures and IT companies is therefore more critical than ever in ensuring that scientific objectives and technological roadmaps are aligned and the required technical expertise is available. CERN openlab plays an important role in this endeavour, setting goals and providing opportunities for collaboration, technical expertise and educational programs. This model can be successfully extended to newer scientific and technical areas with the participation of new major laboratories and research projects.

In order to define the long-term technological context in which joint research activities can take place in the next five years, the CERN IT Department, CERN openlab, and a number of European laboratories and projects, such as ESA, EMBL-EBI, ILL, ESRF, and the Human

Brain Project, have started defining ambitious challenges covering the most crucial needs of IT infrastructures. This process was started in 2013 with the organisation of collaborative events and workshops that have led to the definition of six major areas of investigation and a substantial number of specific use cases in different scientific and technological subjects. The identified areas, or challenges, are data acquisition, computing platforms, data storage architectures, compute provisioning and management, networks and communication, and data analytics.

Data acquisition is where instruments meet IT systems. The increasing amounts and rates of data require more sophisticated and flexible means to collect, filter, and store scientifically relevant data via high speed network links without losing potentially valuable information. Computing systems must be rapidly reconfigured or repurposed to take into account changes in theories and algorithms or exploit idle cycles. Costs and complexity must be reduced by replacing custom electronics with high-performance commodity processors and efficient software.

Once scientifically valuable data is available, it must be supplemented with reliable simulation data and undergo further processing. Throughput can only be increased nowadays by exploiting multi-core platforms or new general-purpose graphical processors, but existing software must be optimised or even redesigned to do that. In turn, this requires parallel programming skills that are not yet readily available and must be formed in the new generations of scientists and engineers.

The data produced by the instruments and the additional information generated by processing it must be made available to increasingly large scientific communities and must be preserved and be accessible for a very long time, ideally forever. Reliable, efficient and cost-effective data storage architectures must be designed to accommodate different usage scenarios and quality-of-service levels based on the many different needs of the user community.

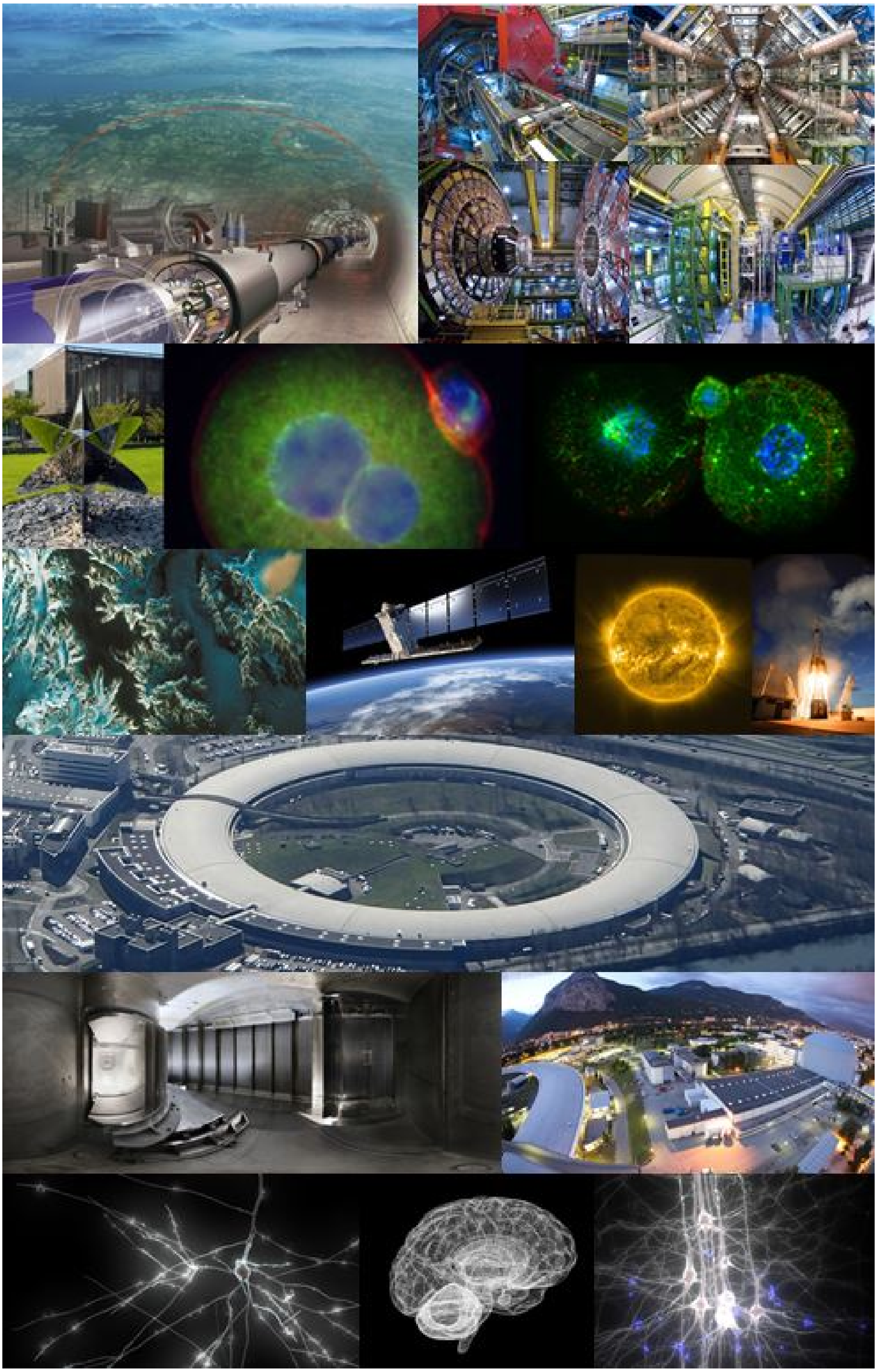
The provisioning and management of the computing infrastructures must become a standard, reliable, and flexible service to fit the needs of large and small communities alike, with measurable and enforceable quality-of-service levels. Increasingly high numbers of computing nodes have to be managed without proportional increases of human resources. This requires automation, virtualisation and efficient policies to scale out when and as needed. At the same time, access to resources within and across different scientific infrastructures must be made secure and transparent to allow collaborations without overlying excessive technical constraints.

Networks across different research centres worldwide must evolve to support the envisioned agile and community-driven infrastructures. Optimisation of data transfers—taking into account user priorities, data criticality and the need for seamless operations across distributed sites—requires new software-based approaches to network architecture design.

Finally, as data becomes too vast and diverse for human beings to understand at a glance, new sophisticated techniques become fundamental to identifying valuable information from noise, finding emerging patterns, helping take proactive decisions, and making new scientific discoveries at an unprecedented pace.

This whitepaper is the result of many discussions among IT experts and scientists and is here proposed as a first step in creating a vision and a strategy for identifying and addressing the big data challenges in scientific research. It will be used as the basis for defining the upcoming CERN openlab V phase and shaping the collaboration among CERN and other European laboratories, international scientific projects and leading IT companies. We welcome contributions, comments and active participation in this endeavour to push the limits even further in support of many more years of outstanding scientific discoveries.

Table of contents



Introduction

CERN openlab[1] is a unique public-private partnership between CERN and leading IT companies. Its mission is to accelerate the development of cutting-edge solutions for the worldwide LHC[2] community and collaborating research institutes. Within this framework, CERN provides access to its engineering experience and its complex IT infrastructure, in some cases extended to research institutes worldwide. Testing in demanding and extreme environments pushes technology to its limits, providing the IT industry partners with valuable feedback on their products, while allowing CERN and the collaborating research laboratories to assess the merits of new technologies in their early stages of development for possible future use. ‘You make it, we break it’ is the CERN openlab motto.

CERN openlab was created in 2001 to cope with the unprecedented computing challenges of the LHC. It has been organised into successive three-year phases. CERN openlab I (2003-2005) focused on the development of an advanced prototype called ‘opencluster’[3]. CERN openlab II (2006–2008) addressed a range of domains, including platforms, databases, grid, security, and networking. The third phase (2009-2011) capitalised on and extended the work carried out in CERN openlab II and tackled new areas with the creation of the Automation and Controls Competence Centre.

CERN openlab is currently in its fourth phase (2012-2014), which addresses topics crucial to the CERN scientific programme, such as cloud computing, business analytics, the next generation of hardware, and security for the myriad networked devices. Current partners are Huawei, Intel, Oracle, and Siemens. Rackspace is a CERN openlab contributor and Yandex an associate. The combined knowledge and dedication of the engineers from CERN and the companies have produced exceptional results leading to significant innovation in many areas.

On the 10th and 11th of December 2013, the CERN openlab IT Challenges Workshop[4] took place in Thoiry, France. The objective of the event was to define the key IT challenges to be addressed in the upcoming CERN openlab V phase, in support of the ambitious research programme of the collaborating research institutes and upgrade schedule of the LHC programme. For the first time, the new phase preparation was extended beyond the physics community, bringing together more than 40 participants from CERN, the IT industry, the LHC experiments, the Human Brain Project, and inter-governmental scientific research organisations EMBL-EBI, ESA, ESRF and ILL. Such participation of laboratories belonging to other research disciplines is desired in CERN openlab V as science becomes more multidisciplinary and cooperation between sciences is critical to foster innovation. The following six challenges were considered as relevant R&D domains to be investigated: data acquisition (online), computing platforms (offline), data storage architectures, compute management and provisioning, networks and connectivity, and data analytics. Use cases for these six challenges were identified and elaborated during the workshop. This document is the summary of the ideas shared during this workshop and follow-up discussions that will pave the way for future collaborations in CERN openlab V.

---

[1] http://www.cern.ch/openlab
[2] http://home.web.cern.ch/topics/large-hadron-collider
[3] http://openlab-mu-internal.web.cern.ch/openlab-mu-internal/10_openlab-I/opencluster/default.asp
[4] http://indico.cern.ch/e/openlab-challenges

Collaborating laboratories and projects

CERN and the LHC experiments

CERN's Large Hadron Collider (LHC) is the world's most powerful particle accelerator and also the largest and most complex scientific instrument ever built. Located in a 27 km-long circular tunnel buried 50-150 m below ground, it accelerates particles to more than 99.9% of the speed of light, to energies never before reached. Some 9600 superconducting magnets operating at just 1.9 degrees above absolute zero (-271.3 ˚C), which is colder than outer space, provide the very strong magnetic fields needed to keep the particles on the right orbit. Four very large detectors (ALICE, ATLAS, CMS, LHCb)—comparable to huge high-resolution 100-megapixel 3D cameras—record up to 600 million times per second the 'mini big bangs' created by collision of the particle beams. In the last weeks of the LHC's first run, the remarkable figure of 100 PB of data stored in the CERN mass-storage systems was surpassed. This data volume is roughly equivalent to 700 years of full HD-quality movies. The Nobel Prize in Physics was awarded jointly to François Englert and Peter W. Higgs in 2013 for their theoretical discovery of the Higgs boson, the existence of which was confirmed by the ATLAS and CMS experiments during the LHC's first run. However, the Standard Model only describes the 4% of the known universe. What is the other 96% made of? Will we see a unification of forces at the high energies of the LHC? Why is gravity so weak? Why is there more matter than antimatter in the universe? Is there more exotic physics waiting to be discovered at higher energies? Will we discover evidence for a theory called supersymmetry at the LHC? Can the LHC shed light on dark matter?

To unlock such mysteries, the LHC needs to be operated at higher energies and collision rates. To keep up with the challenge, the LHC and detectors are currently being upgraded. This shutdown period, known as LS1, started in February last year and will continue until early 2015. At this time, the LHC will resume operation at higher energies until another similar stop takes place in 2018. This second shutdown, known as LS2, is expected to last just over a year and will see a further increase in the rate of collisions achieved by the LHC. A significant amount of effort will also be invested to ensure that data acquisition can keep up with the increased particle collision rates. Importantly, not only does this mean that more data will be produced by the experiments, requiring increased storage and improved systems interconnects, but it also means that the 'trigger' systems used to select interesting collisions for further analysis will need to be upgraded. With CERN openlab V set to run over the three years immediately prior to LS2, there is an excellent opportunity to collaborate on producing the solutions that will be implemented at CERN during the LS2 phase.

EMBL-EBI

The European Molecular Biology Laboratory outstation the European Bioinformatics Institute (EMBL-EBI[5]) has the largest collection of freely available life science experiment data in Europe. This public data is available for download, browsing online and can be analysed through both interactive and programmatic web services. EMBL-EBI provides an extensive user-training programme based around its services, which is used to help drive their development. Groups within EMBL-EBI are both service providers and consumers of services to support their own research activities. EMBL-EBI hosts the hub of ELIXIR[6], the European life sciences infrastructure for biological information that will bring together member states

---

[5] http://www.ebi.ac.uk/

[6] http://www.elixir-europe.org/

from across Europe to provide the research infrastructure needed for a variety of life-sciences disciplines.

The computing and storage infrastructure needed to support these diverse services is distributed across on-site and off-site data centres. This distributed architecture provides the capacity to reliably and resiliently deliver these public services to millions of unique users each year. The sites are linked through the academic and research networks within the UK (JANET), Europe (GÉANT) and internationally to provide the general and dedicated bandwidth needed by their users and collaborators. The ability to deliver services in partnership with other service providers within Europe is growing in importance as the ELXIR life science infrastructure starts its construction phase.

ESA

The European Space Agency (ESA[7]) is an international organisation composed of 20 European member states, with Canada as an associate member. ESA's purpose is to advance Europe's vision in space research and technology. There are six main programmes dealing with space science, Earth observation, human spaceflight, telecommunications, navigation and launchers. The mandatory programme of ESA is science, to which all member states contribute.

Data acquisition and dissemination make up the bulk of the work carried out in the domains space science and Earth observation. In the next five years, ESA spacecraft will acquire about 25 PB of Earth-observation data, while the space-science spacecraft will enhance our understanding in the Universe by acquiring data in all major wavelength regions. Challenges, such as the performance of networks, connectivity, resource management, and data analytics are relevant to how ESA serves its scientific and user communities.

In addition, and based on ESA's technical (for spacecraft development) and operational capabilities, there are several data centres across ESA establishments. Some are highly configured for crucial space operations and data acquisition, while others are more flexible based on the needs and requirements of projects. In these areas, and due to the duration of the development of any space mission, which may require up to 10 years, data simulation and processing become crucial. Data storage and virtualisation in the context of data centre technology are very relevant in this area.

Finally, upcoming IT technologies such as big data and predictive analytics are becoming more and more a necessity in order to better design spacecraft and improve data dissemination performance.

ESRF

The European Synchrotron Radiation Facility (ESRF[8]) is the most powerful synchrotron radiation source in Europe. Each year, several thousand researchers travel to Grenoble, where they work in a first-class scientific environment to conduct exciting experiments at the cutting edge of modern science.

---

[7] http://www.esa.int/ESA

[8] http://www.esrf.eu/

A synchrotron is a stadium-sized machine that produces many beams of bright X-ray light. Each beam is guided through a set of lenses and instruments called a beamline, where the X-rays illuminate and interact with samples of material being studied. Many countries operate synchrotrons—there are 10 in Europe—but only four worldwide are similar in design and power to ESRF.

At more than 40 specialised experimental stations on the beamlines, physicists work side by side with chemists and materials scientists. Biologists, medical doctors, meteorologists, geophysicists and archaeologists have become regular users. Companies also send researchers, notably in the fields of pharmaceuticals, consumer products, petrochemicals, and microelectronics.

Today, ESRF has reached the peak of its Upgrade Programme Phase I, with the first upgraded beamlines open for users and many more to follow by the end of 2015. The Upgrade Programme Phase I, from 2009 to 2015, is due to deliver:

- Eight beamline upgrade projects with capabilities unique in the world, comprising 11 different beamlines with 15 independently operable end stations.
- Complete or partial refurbishment of the remaining ESRF beamlines, to maintain them at world-class level.
- Improvements to the X-ray source to maintain world-leading beam availability, stability, and brilliance.
- New state-of-the-art instrumentation, driven by the needs of the new beamlines, notably in X-ray mirror engineering, diamond technologies, nano-focussing optics, pixel detectors, on-line data analysis and high-rate data collection.
- New buildings for long beamlines, support laboratories, and offices.

The second phase of the ESRF Upgrade Programme will cover new developments from 2015 to 2018. Accelerators and X-ray source will be pushed to limits beyond the most optimistic expectations at the time of their design, with a brilliance fifty or one hundred times higher than possible today.

The main challenges for IT at ESRF will be to:

- Develop solutions in terms of bandwidth, storage capacity, and computational capacity, dedicated to specific detectors/beamlines to guarantee bandwidth for fast data writing and simultaneous reading for on-line data analysis.
- Explore new concepts for central storage systems for increased bandwidth in line with the expected data flow from the beamlines and the requirements for fast data analysis with massively parallel clusters.
- Explore new co-processor architectures for potential gains in price, performance, and power consumption.

In parallel to these development activities, the primary mission has to be fulfilled:

- Provide enough central storage capacity to maintain, and in some cases even extend, the current data retention policy for visitor data and in-house research data.
- Provide computing capacity for on-site and remote data analysis of large datasets that are difficult to transport.
- Keep the data communication network and backup infrastructure ‘state of the art’.

ILL

The Institut Laue-Langevin (ILL[9]) is an international research centre at the leading edge of neutron science and technology. It operates the most intense neutron source in the world, providing beams of neutrons to a suite of 40 high-performance instruments. Over 750 experiments are completed every year, in fields including magnetism, superconductivity, materials engineering, and the study of liquids, colloids and biological substances.

The high-flux research reactor produces neutrons through fission in a specially designed, compact-core fuel element. Neutron moderators cool the neutrons to useful wavelengths, which are then directed at a suite of instruments and used to probe the structure and behaviour of many forms of matter by elastic and inelastic neutron scattering, and to probe the fundamental physical properties of the neutron. Nothing goes to waste: Fission products and gamma rays produced by nuclear reactions in the reactor core are also used by specialised instruments, which form an important part of the instrument suite.

An ambitious modernisation programme (2001-2014) was launched in 2000, through the design of new neutron infrastructure and the introduction of new instruments and instrument upgrades. The first phase has already resulted in 17-fold gains in performance. The second phase started in 2008 and comprises the building of five new instruments, the upgrade of four others, and the installation of three new neutron guides. The reactor is currently shut down. During this period, a number of major projects have been scheduled. The main operations planned during the shutdown are: the installation of a new instrument (ThALES), the major upgrade of another four instruments (SuperADAM, D16, D22, IN15), and the replacement of two beam tubes (H13 and IH3, plus almost all of H5).

The IT Department not only provides IT support and solutions, but also networking, archiving, and curation for the large dataset acquired from scientific experiments carried out since 1973. It also provides analysis infrastructure. Since the publication of the ILL data policy in 2011, data is publically available to the scientific communities following a three-year period when the dataset is exclusively available to the initial experimenters.

Until recently, visiting scientists were able to easily transfer their data to their home laboratory for further analysis (using hard drives or standard network transfer protocol). Other scientists used locally available infrastructure for processing their data. Nowadays, with the recent growth of the volume of experimental data generated at ILL, transporting data to a home facility is no longer feasible. Providing a modern solution for the data analysis has become a paramount objective for the IT Department.

ILL needs to improve its analysis facility by providing more capacity, flexibility, greater performance, security, user friendliness, and interaction with other capacity providers. Hopefully cloud technologies are now mature enough to help ILL achieve this goal.

The Human Brain Project

The Human Brain Project[10] is a ten-year research project, coordinated by EPFL[11] in Lausanne (Switzerland) and funded by the European Commission, with a goal of laying the foundations

---

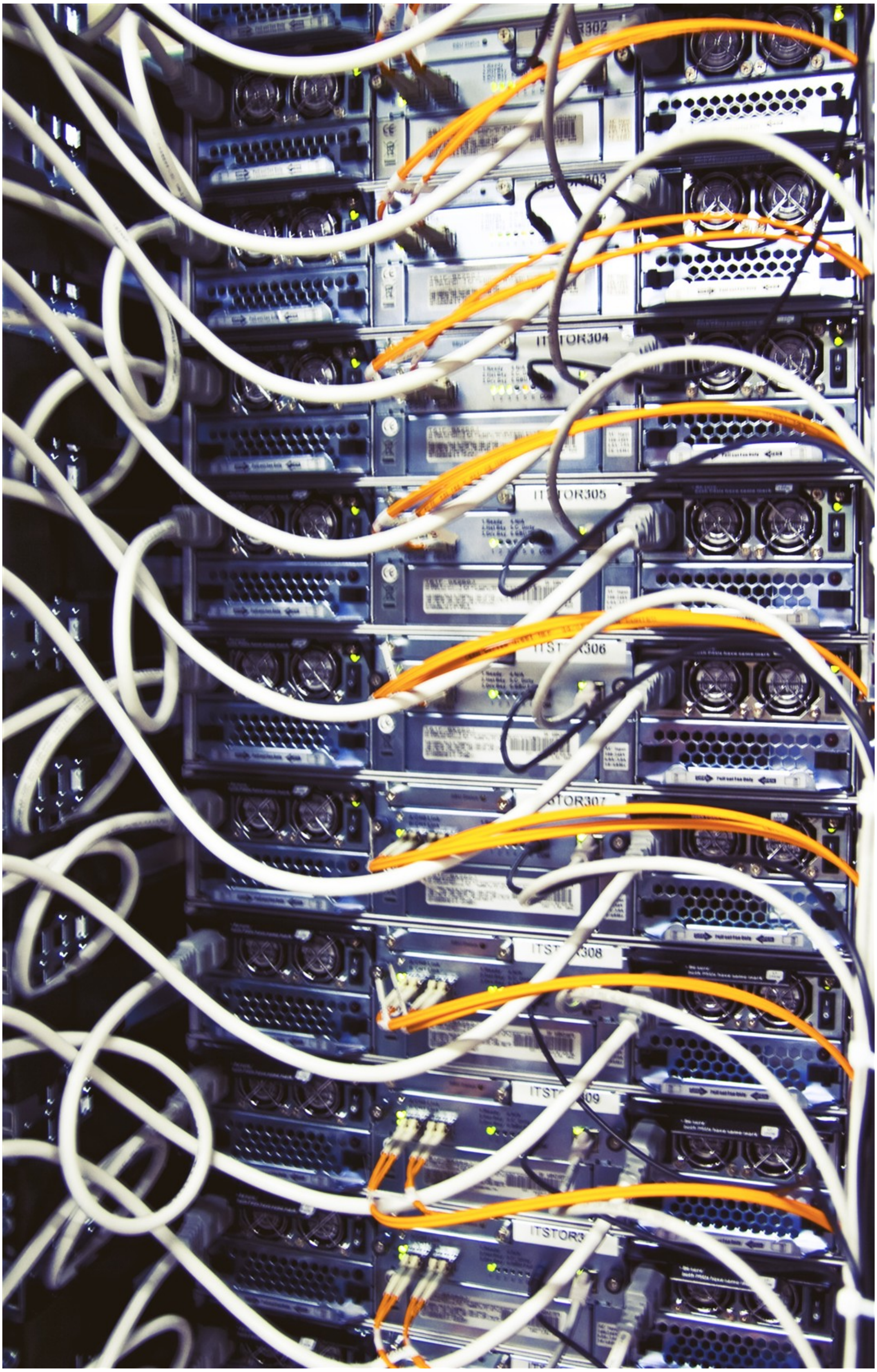
[9] https://www.ill.eu/

[10] https://www.humanbrainproject.eu/

[11] http://www.epfl.ch/

for a new approach to brain research. The fields of neuroscience, medicine, and IT each have important roles to play in addressing this challenge, but the knowledge and data that each is generating are currently very fragmented. The Human Brain Project is driving integration of these different contributions and catalysing a community effort to achieve a new understanding of the brain, new treatments for brain disease, and new brain-like computing technologies.

To support this effort, the Human Brain Project is creating an integrated system of IT platforms, offering services to neuroscientists, clinical researchers, IT developers, and roboticists. These platforms are supported by other subprojects that are focused on filling critical gaps in physical brain data and our theoretical understanding of the structure and functioning of the brain. The origins of the Human Brain Project lie in the convergence between neuroscience, medicine and IT. This convergence can be seen, for example, in the application of computer technology to help construct accurate brain models from limited sets of physical data. It is clear that future progress in neuroscience and medicine will be increasingly dependent on IT.



Challenge 1: Data acquisition (online)

Existing and emerging large-scale research projects are producing increasingly high amounts of data at faster and faster rates. The quantity and rates at which the data is produced is expected to increase as new technology and instruments are deployed. Projects in different scientific disciplines use a wide variety of instruments, including sensors, detectors (such as the LHC experiments' detectors at CERN), high-throughput genome sequencers, X-ray free-electron lasers, satellite imaging devices, and radio telescopes or antennas. While the different instruments have specialised capabilities for their particular field of application, they all have in common the need to reliably transform physical or chemical processes into digital data and the need to support complex chains of systems to filter, store, and analyse the data in real-time.

The data produced by large scientific instruments, or by large numbers of smaller instruments, is often generated in extremely large amounts. However, not all of this data is useful for research. In fact, in many cases, only a small fraction of this data may be of interest: Much of the data may, for example, contain information that is already known. Even when all the data is potentially useful, the rates at which it is produced prevent it from all being stored on existing data storage systems.

A prime example of this comes from CERN's LHC, which produces millions of particle collisions every second in each of its detectors, thus generating approximately 1 PB of data per second. None of today's computing systems are capable of recording such rates, so sophisticated selection systems must be used to filter the data. This filtering process takes place in two stages:

- The first, known as 'the level-1 trigger', is based on custom hardware and selects just one in 10,000 collision events. It is an extremely fast and wholly automatic process that looks for simple signs of interesting physics, such as particles with large amounts of energy or in unusual combinations.
- Following this, the 'high-level trigger' analyses in more detail the events selected by the level-1 trigger and selects 1% of the events for storage and further analysis. To achieve this, tens of thousands of computer processor cores work in unison to assimilate and synchronise information from different parts of the particle detectors to recreate the entire collision event.

Even after such drastic data reduction, the LHC's four big experiments (ALICE, ATLAS, CMS, and LHCb) together store over 25 PB of data per year. This is set to increase significantly as the LHC is upgraded to enable an even higher rate of particle collisions.

Many major research laboratories and scientific projects are today experiencing an exponential increase of their data sets and hence have similar needs in at least three areas: online (real-time) data filtering and processing; high-bandwidth networking; and data transfer and storage.

Use case 1: Online data filtering and processing

Online (real-time) data filtering (triggers) of large quantities of high-rate data is traditionally done using specialised electronics. While hardware-based trigger systems are certainly rapid (the decision time is just a few microseconds), there are a number of disadvantages. Interesting events or physical processes can be erroneously filtered out due to the relative simplicity of hardware systems. Upgrades to the system are expensive, and maintenance can only be carried out by experts. If the knowledge of the physical events changes (new theories are developed) it is not easy to reprogram the systems without costly delays or even the need to stop the instruments. At CERN, for example, each sub-detector within the LHC experiments' particle detectors has its own solution in terms of the signals it uses to decide whether or not to filter out collisions. Whereas all experiments have a common link from the detectors into the data acquisition systems using a common interface, the triggers themselves are highly customised by each experiment.

The CERN LHC experiments now have plans to replace the existing hardware-based level-1 triggers with systems based on commodity hardware processors and co-processors running more standard (or even shared) software. Switching to a software-based system for the level-1 trigger would allow flexible pattern recognition, although it would be important to achieve very low latency. An efficient interface to the detector hardware would also be paramount. Finally, the code-base would need to be made ready for multi/many-core computing environments and the compute would need to be optimised in terms of cost, power, cooling, etc.

Use case 2: High-bandwidth (TB/s) networking for data acquisition at LHC

The data produced by separate instruments directly or after an initial filtering must usually be collected together to reconstruct the physical process through which it was generated. In the LHC experiments, events accepted by the level-1 trigger are filtered using the high-level trigger, which gathers together information from various readout units on the particle detectors to recreate the collision event. It is important that fast networking links are in place to bring this data together quickly. The planned upgrades of the LHC experiments will require higher bandwidth networks to cope with the increasing data rates resulting from the increase in machine luminosity. We expect the network requirements for each experiment to approach 100 TB/s. Other 'big science' projects expected to come on stream in the coming decade also have the need to cope with the handling of significant data rates and their reduction to a rate that is sustainable by the offline infrastructure.

Upgrades carried out by the experiments during the LHC's LS1 phase will bring major changes to the data acquisition systems of ALICE and LHCb. Both experiments will require multi-Tbit/s local area networks to collect data from multiple data sources into one of many compute nodes to perform software compression (ALICE) and filtering (ALICE and LHCb) of the data. These networks need to be very high-bandwidth and cost-effective due to the large number of links (about a thousand 100 Gbit/s links will be needed for LHCb). Often, data will need to be transferred at rates of multiple TB per second, and this will have to be done reliably and cost effectively. In turn, this network will have to be integrated closely and efficiently with the compute resources, be they based on classical, single-, or dual-core CPUs (increasingly less) or many/multi-core platforms (increasingly more). Equally, it is key that multiple network technologies should be able to seamlessly co-exist in the same, integrated fabric, providing an end-to-end solution from online processing to data analysis.

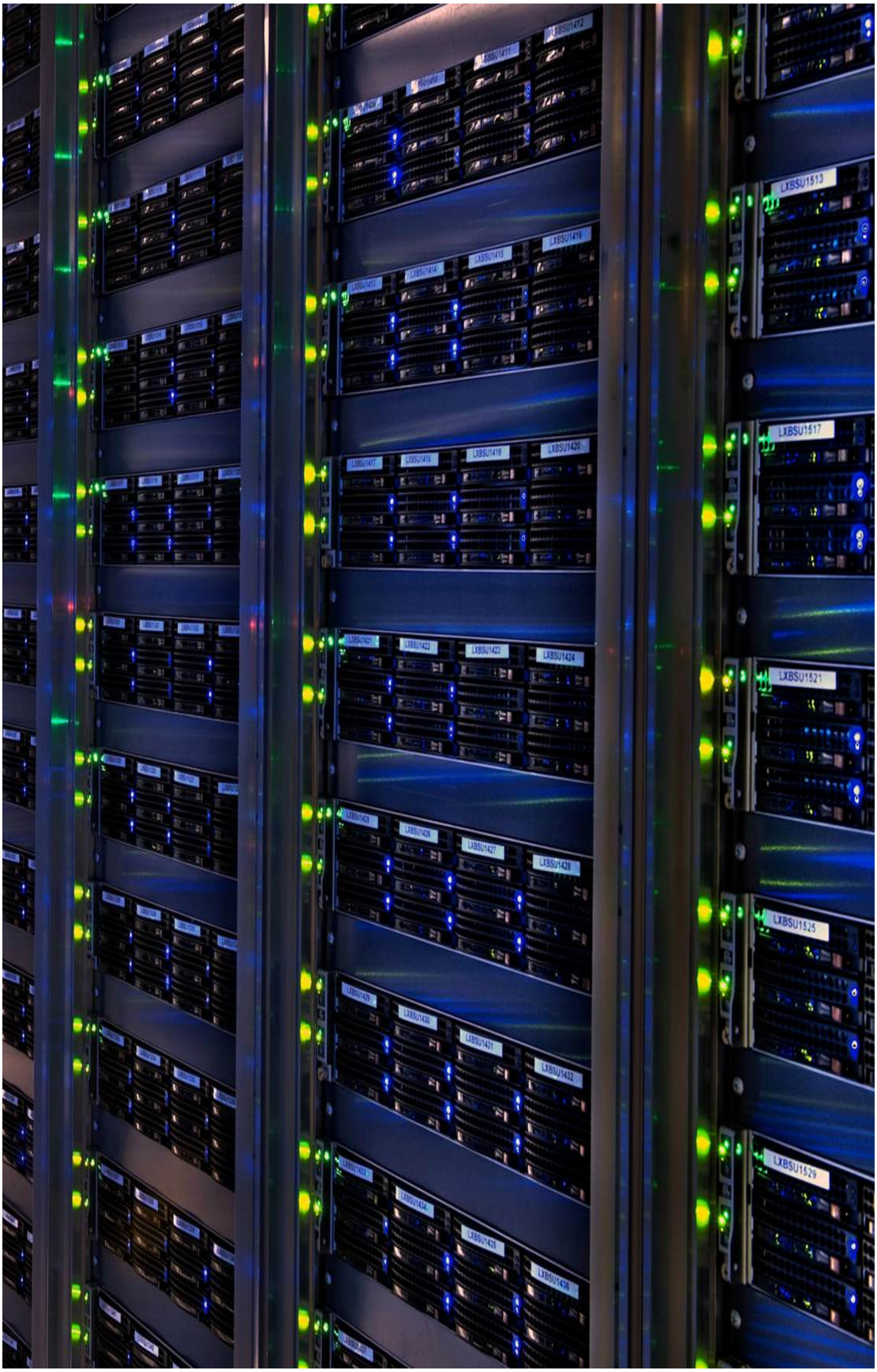
Whilst data-centre oriented Ethernet is clearly an option, other more light-weight technologies, such as those used in the HPC domain (e.g. InfiniBand), are an interesting alternative, provided they give a cost advantage. Traditional data acquisition networks have used large core-router style devices. Modern high-bandwidth single-chip devices open up the possibility to build sophisticated topologies (Clos, fat-tree, etc.) specifically adapted to the rather special (unidirectional, bursty) data-flow of the LHC experiments. Having such a technology available will facilitate the move to even more use of commercial off-the-shelf

(COTS) solutions than is already the case, shifting resources from custom-built hardware to COTS solutions provided by openlab partners.

The expected outcome and impact for stakeholders is that success in these projects will be a clear landmark for the upgrade efforts of ATLAS and CMS, which are due several years later for the post-LS3 upgrades. These experiments will then also see massively increasing data acquisition needs.

Use case 3: Data transfer and storage

Online storage is typically transient storage. Data needs to be buffered and consolidated before it can be sent to the computing systems for processing and to permanent storage. Moreover, data verification is usually done on a file basis. This results in requirements where files are usually written only once, but potentially read multiple times (at least twice). For example, the requirements of the upgraded LHC experiments are dozens of parallel streams with an aggregated write-throughput of up to 20 GB/s. The integrated volume is large and expected to reach the PB-scale. A special challenge comes from the necessity to flexibly repurpose the computing farms to perform data aggregation and reconstruction (tasks typical of the online systems) or data re-processing, simulation and analysis workloads (tasks typical of offline systems taking data from storage rather than the data acquisition systems). With the LHC experiments, for instance, the large high-level trigger computing farms can be valuably used by the LHC collaboration for offline tasks when not needed for online tasks. Since the duty-cycle of the LHC has so far been far below 50%, this rapid and flexible reconfiguration of the systems becomes critical. This requires quick retiring of jobs and flexible management of data to accommodate the different workloads, while ensuring high-priority is given to the online data processing. This leads to much increased demands on storage due to many more parallel streams (up to several thousands of reads) and more aggregated throughput. Distributed architectures and high-performance file systems are clearly an option, but they must provide a global namespace to simplify the management of the facility by a very small crew of administrators.



Challenge 2: Computing platforms (offline)

The success of existing and future scientific and experimental programmes depends among other factors on an efficient exploitation of the recent and future advances in computing technology.

Computing power is needed for a number of very important tasks. The data produced by scientific detectors and instruments—and possibly filtered by the data acquisition systems—needs to be processed and analysed to understand whether it represent a significant event or physical phenomenon, or to produce aggregated, consolidated, or derived information. The data must also be reprocessed when new algorithms are defined or the instruments are recalibrated, and simulations of the instruments' behaviours and properties have to be performed to provide comparative numbers against which the real data can be measured.

Different computing models are currently used by research infrastructures depending on the types of data and analysis to be made. Computations in biomolecular analysis, neurology or fluid dynamics are traditionally done on specialised supercomputers, providing high performance and close interaction among computing jobs (HPC). HEP analysis is traditionally done on large, distributed computing infrastructures of commodity hardware providing high throughput of job execution for long-running, independent computations (high-throughput computing, HTC). In recent years, the Grid has been a successful implementation of HTC, although cloud computing is now gradually increasing its presence in scientific research as a more flexible and potentially cost-effective solution.

Continuous advances in computing mean, however, that code that was written 10 or 20 years ago is today no longer able to properly exploit the new features of modern platforms (e.g. the presence of multiple cores in the same CPU). Existing code must be optimised using various techniques like vectorisation, possibly using automated features of modern compilers and optimisation tools. At the same time, new code must be written taking into account the new hardware, which means that developers need to have appropriate skills in multi-threaded programming and be familiar with modern programming models and compilers.

In the past several years, a number of initiatives have started in various research programmes to optimise existing code on new architectures. CERN openlab has itself conducted pioneering work in optimising HEP code on new platforms. New international programmes like, for example, the Square Kilometre Array (SKA[12]) or the Human Brain Project have fully recognised the need to train scientists and engineers to be able to combine scientific knowledge with modern programming techniques.

In order to fully exploit the benefits of the new computing platforms, a number of important activities, therefore, need to take place. The continuous process of evaluation, benchmarking and optimisation of the computing platforms has to continue to provide benefits to both users and vendors, as CERN openlab has successfully demonstrated over the past years. Existing software needs to be revised, optimised or completely redesigned to fully exploit the performance gains provided by newer multi-core platforms, fast co-processors, and graphical processors. A close collaboration between scientists and experimentalists on one side and computing experts from academia and hardware vendors on the other side has to be established in order to generate the necessary knowledge transfer and the creation of new skills and competencies. The following use cases exemplify the tasks defined above, mainly

---

[12] https://www.skatelescope.org/

drawing from the experience of the computing and code development activities in the HEP community.

Use case 1: Continuous benchmarking and evaluation of hardware platforms and software tools

A solid frame of reference is indispensable when evaluating computing systems in terms of their performance, energy consumption, and thermal characteristics. Such studies allow for significant optimisations and are of utmost importance when making choices in the computing domain. Most large-scale research projects require increasing computing capacity and performance. Scalability, seamless throughput, and intelligent power optimisation are paramount to the provision of a highly-efficient and cost-effective computing infrastructure.

As new platforms are developed it becomes critical to be able to benchmark and evaluate them as early as possible in their development cycle for two important reasons: First-hand experience on newer platforms enables timely decisions to be taken when planning for future upgrades or expansions of the computing infrastructure. Equally, early access allows experts to provide essential feedback to the hardware manufacturers and make sure that the expected increases in performance and efficiency are indeed delivered. Evaluation needs to be done on different classes of platforms, from high-performance server systems to desktop systems where new microarchitectures are often first introduced to new classes of low-power or ‘high performance per watt’ ratios to evaluate their usability in future ‘green’ infrastructures.

Advanced performance and scalability studies would not be possible without the proper support software. Compilers, performance tuners, correctness- and memory-checking applications are essential tools in the standard development cycle of complex applications, especially on recent multi-core platforms. Testing and evaluation of such tools on production and experimental platforms and prompt interaction with hardware and tools manufacturers not only helps ensure the availability of efficient tools, but also makes them available and usable to the larger community of scientific developers.

Use case 2: Design and optimisation of simulation and analysis software

Simulation and analysis is where most of the CPU time is spent for HEP computing. Therefore, it is of utmost importance that these codes make optimal use of the advanced capabilities of modern CPUs. Most modern CPUs contain vector pipelines and multiple cores and existing codes need to be vectorised and parallelised to make the best use of these new features. The Geant V project is targeting the development of new simulation software with the aim to make optimal use of these specific CPU features. A lot of prototyping has already been undertaken and an operational prototype has been assembled. Based on this, a lot of further research and testing is needed to understand if optimal usage is being achieved. This is also an ideal environment to deploy the latest tuning and profiling tools developed for these new CPUs. Moreover, such simulation is widely used beyond HEP. Radiation treatment planning, design and optimisation of medical instruments, design of all instruments and processes using ionising radiation, radiation protection in nuclear power installations, and evaluation of the effects of ionising radiation in space flights are only some of the examples where simulation is relevant for research, industry and society at large. The availability of a fast and reliable radiation transport programme could change these fields entirely, enabling the next step in design optimisation to be taken.

What has been said for simulation is also true for the data analysis and big data applications being used in HEP. Early studies have shown that restructuring the analysis codes to use large vectors can lead to important speedups. A project to redesign the core analysis engine of the ROOT data analysis system would have a large impact as ROOT, like Geant, is very widely used in HEP and in many other sciences.

Use case 3: Investigation of applications of GPUs and co-processors

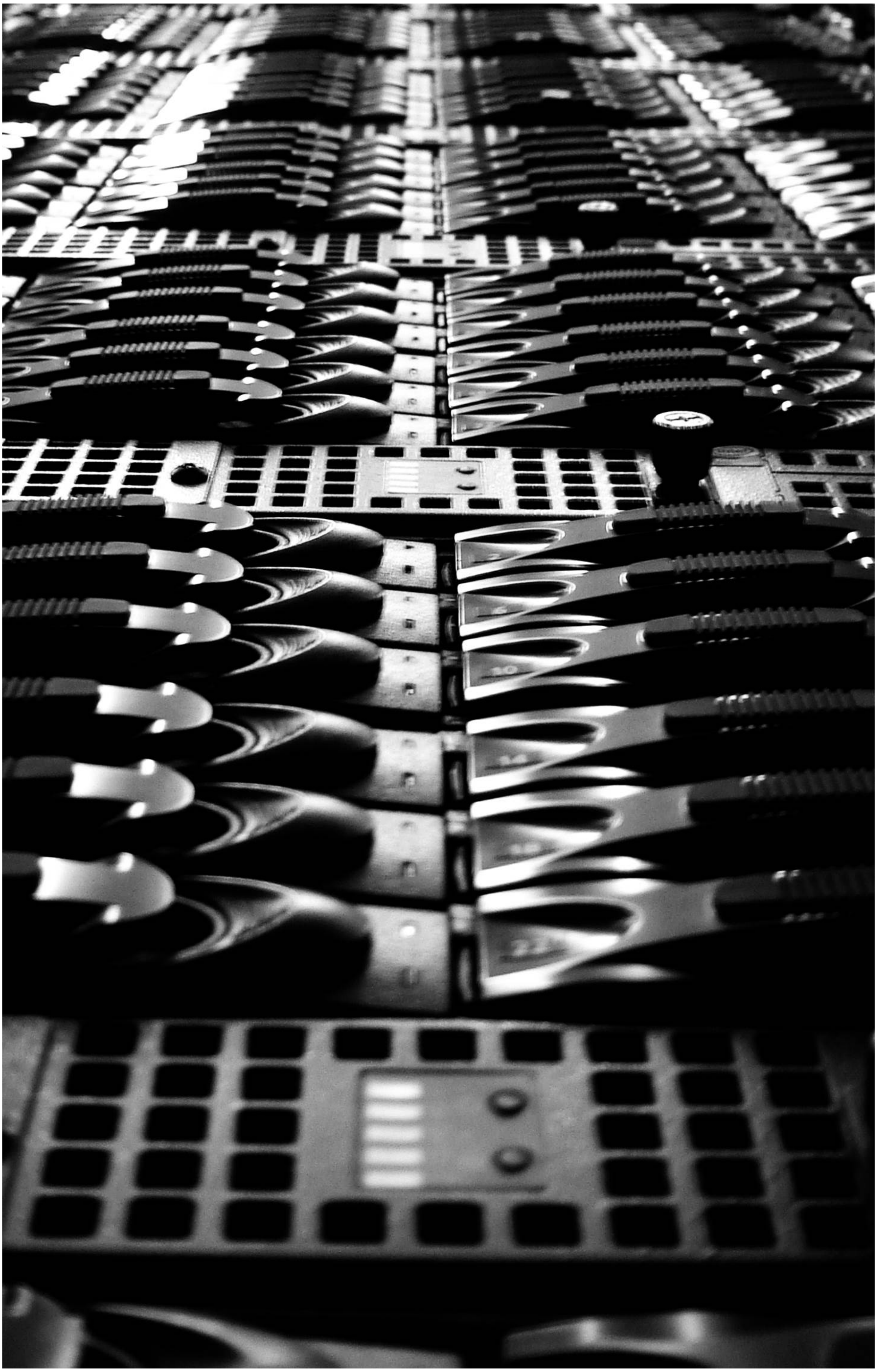
Graphic processor units (GPUs) and co-processors (e.g. Intel Xeon-PHI) are generally not available in offline computing platforms used for simulation, reconstruction and analysis. However, these processors will very likely be available in the dedicated high-level trigger farms used for online event filtering. As these farms will also be used for offline event processing during accelerator down times, it is important to research event processing algorithms that can make good use of the data parallel processing features of the GPUs and co-processors, as this can massively speedup the event processing. The algorithms can be implemented using open standard languages like OpenCL or OpenMP, to be CPU and GPU independent, or in a vendor-specific language if that brings clear performance benefits.

Use case 4: Development of expertise and skills in multi-core parallel programming

Although parallel programming is nothing new, the challenges due to the constraints in indefinite frequency scaling and the advent of many-core and multi-core platforms, as well as specialised platforms like GPGPUs, have introduced a shift in how computing systems and software must be designed to keep improving performance and efficiency. Many large scientific research projects require increasing computational power which cannot be achieved by adding more and ever faster CPUs both from an overall cost and a power-consumption point of view. As mentioned earlier in this paper, simulation and analysis are fundamental activities in HEP research, radio-astronomy, computational biology, neurosciences, environmental modelling, and many other disciplines. It has been noted that a large part of the software in use today has not been designed to exploit multi-core platforms and will have to be adapted or most likely redesigned. However, there is still a widespread lack of expertise in multi-core platforms, the related support tools (like compilers or profilers), and the parallel programming techniques required to write reliable applications for them.

Young scientists and engineers are rarely taught the necessary programming skills for such complex environments and usually only as part of advanced computer science courses. Furthermore, considerable understanding of the scientific or technical subject matter is often required to design domain-specific applications. Equally, advanced mathematical knowledge may be necessary to fully understand how, for example, synchronisation or vectorisation algorithms work. The lack of skills and experience in this area may have considerable impact on how scientific computing evolves in the coming years, as well as how educational and training programs from universities and industrial companies may evolve.

Focused collaboration, therefore, becomes fundamental among scientific research laboratories, academia, and vendors of hardware and software. This is vital in ensuring that university curricula are properly updated and new professional profiles are formed. Public-private partnership models must be explored whereby the laboratories provide the actual science cases, the vendors provide consultancy and expert training, and universities provide the formal curricula and accreditation necessary to ensure that the new generation of engineers have the required knowledge and experience as soon as possible in their professional careers.



Challenge 3: Data storage architectures

The storage and management of LHC data is one of the most crucial and demanding activities in the LHC computing infrastructure at CERN and also at the many collaborating sites within the Worldwide LHC Computing Grid (WLCG). Every year, the four large-scale LHC experiments create tens of PBs of data, which need to be reliably stored for analysis in the CERN Data Centre and many partner sites in the WLCG. Today, most physics data is still stored with custom storage solutions, which have been developed for this purpose within the HEP community. As the user demands are increasing in data volume and aggregated speed of data access, CERN and its partner institutes are continuously investigating new technological solutions to provide their user communities with more scalable and efficient storage solutions. At the same time, CERN closely follows the larger market trends on the commercial side and continuously evaluates new solutions for the physics use cases, so as to be ready for their adoption as soon as they have matured sufficiently for deployment at large scale.

The recently emerged cloud storage architecture and its implementations may provide scalable and potentially more cost effective alternatives. Native cloud storage systems, such as the Amazon Simple Storage Service (S3), are typically based on a distributed key-value store, and divide the storage namespace up into independent units called buckets. This partitioning increases scalability by insuring that access to one area (bucket) is unaffected by the activity in other parts of the distributed storage system. In addition, the internal replication and distribution of data over different storage components provides intrinsic fault-tolerance and additional read performance: multiple data copies are available to correct storage media failures and to serve multiple concurrent clients.

On the larger scale (across site boundaries), the HTTP-based S3 protocol has become a *de facto* standard among many commercial and open-source storage products. Hence, it may become an important integration technology for consistent data access and exchange between science applications and a larger group of sites. One of the advantages of S3 is that the decision either to operate a private storage cloud or use commercial cloud services would still be left to the site, based on its size and local cost evaluation.

Similarly, EMBL-EBI continues to see exponential increases in its archived data due to the growth of activity in the life-sciences community, with storage volumes doubling approximately every 12 months. Work is in progress to investigate a variety of storage solutions, including long-term tape-based archive, local non-replicated storage and geo-replicated storage, so that the data can be matched to the most appropriate storage solution.

Use case 1: Evaluation of cloud storage for science use cases

Most cloud storage providers today offer simple solutions with limited flexibility. In general, users can buy storage space for agreed period of times and at a more or less fixed cost per MB. Space and cost are the main parameters on which a choice can be made. However, many applications—especially scientific ones—often need flexibility that goes beyond mere availability of space. Flexibility may also be required, for instance, in terms of the quality of service provided by the cloud storage. Parameters such as reading and writing bandwidth, reliability, and single or multi-user access optimisation have to be taken into account when designing the data architecture of a scientific application. The possibility of defining different quality-of-service levels based on general requirements (e.g. cheap even if slow, or fast and expensive) is considered of strategic importance in the future provisioning of data services. The ultimate goal would be to be able to arbitrarily select which storage parameters to optimise for any given user application.

Use case 2: End-to-end implementation of operational procedures

Archival and long-term data storage strategy makes extensive use of tape systems due to their reliability and cost-effectiveness. The current archival system at CERN is based on the CERN Advanced STORage manager (CASTOR[13]) system and makes use of a tape back-end and a pool of data staging disks.

An important area of investigation is the end-to-end implementation of operational procedures, particularly in the area of data integrity and protection. New tape drives enable the achievement of end-to-end integrity and improved reliability via record-level CRC-based integrity validation.

The recently defined T10-PI protocol allows for end-to-end integrity checking by ensuring that data is validated as it moves through the data path from the application to the storage. T10-PI is being implemented in both tape and disk systems by various vendors. An investigation of full T10 PI end-to-end data protection across the whole of CASTOR, including both the tape and the disk layers, is considered of high interest as increases in the channel data rates also increase the chance of data errors.

Use case 3: NoSQL solutions for big data

As the volume, variety, and rates of data produced by scientific research grows, the need to provide scalable and cost-effective solutions to process the data closely following the user application requirements becomes critical. Scientific data comes in many different forms, from structured data from computations and text in papers or presentations, to digital measurements and images from satellites or graphical visualisations.

New data architectures should include support for NoSQL engines and other suitable technologies. They also need to be capable of handling the increasing variety and quantity of data. This means they will need to support data versioning, dynamic schemas, integration of data from different sources, complex or hierarchical data types, and so on.

Use case 4: Scalable namespaces and catalogues

Data files produced by scientific applications as part of large scientific collaborations are made available to the community in multiple copies. This provides redundancy and improves speed of access.

Locating files across massively distributed data infrastructures requires scalable namespaces and catalogues as integral parts of data architectures. The data associated with the data files, or metadata, must not only provide 'physical' information, such as the number of replicas or the location of a file, but also 'contextual', domain-specific information about aspects like provenance, usage, purpose, associations with other digital objects or within a hierarchy of objects, and so on. In addition, the definition of common vocabularies and a common approach to namespaces and metadata across different disciplines would enable reuse and sharing of the data.

---

[13] http://castor.web.cern.ch/

Challenge 4: Compute management and provisioning

European scientific research has benefited in the past several years from the increasing availability of computing and data infrastructures that have provided unprecedented capabilities for large-scale distributed scientific initiatives. A number of major projects and endeavours, such as EGI[14], PRACE[15], WLCG[16], OSG[17] (in the USA), and others, have been established to share the ever-growing amount of computational and storage resources. This collaborative effort has involved hundreds of participating research organisations, academic institutes, and commercial companies. The major outcome was a number of active production infrastructures providing services to many research communities, such as HEP, life sciences, material science, astronomy, computational chemistry, environmental science, humanities, and more.

One of the first implementations of massively distributed computing for scientific research was the Grid. Grid computing essentially combines computers from multiple administrative domains to solve single, but independently parallelisable tasks. Compute provisioning and management in grid computing is an extension and abstraction of the traditional concept of batch computing, whereby individual computing nodes are managed by a master system allocating tasks to the nodes. The distribution and allocation of computing and data tasks across computing sites is done using specialised middleware services responsible for different functions, such as authentication and authorisation, workload management, data transfers, logging, etc.

As the use of virtualisation has become a more and more viable and efficient solution for instantiating computing nodes, the concept of ‘the Cloud’ or cloud computing has gradually established itself as a more efficient and cost-effective solution to scientific computing. Although grid and cloud computing have many similarities, cloud computing differs in a number of important aspects for both providers and users. Compute provisioning and management in cloud computing can make better use of virtualisation and automation thanks to an increasing number of standard tools and services, which are supported both commercially and at a community level. This, in turn, allows computing sites to provide increasing amounts of resources, faster and more reliably. Cloud computing shifts the focus from pure resource provisioning (IaaS) to service provisioning, allowing the combination of different elements into higher level platforms (PaaS) or applications (SaaS), often tailored to specific user requirements. Finally, better defined costing models for cloud computing allow service providers and users a way to establish usage and provisioning contracts.

CERN, as infrastructure and service provider for the HEP community, has been very actively involved in grid and cloud computing since the early days. The WLCG provides computing and data services to LHC experiments across European and international research and computing sites. CERN provides 15% of the WLCG resources and needs to continually support the growing computing and data handling requirements of the LHC and its major experiments as the LHC technical and scientific programme evolves. In 2013, CERN inaugurated its data centre extension at the Wigner Research Centre for Physics (close to

---

[14] https://www.egi.eu/
[15] http://www.prace-ri.eu/
[16] http://wlcg.web.cern.ch/
[17] http://www.opensciencegrid.org/

Budapest, Hungary). CERN needs to manage the increasing amount of resources and requests across the two sites with a standard, cost-effective, and scalable provisioning and management system, as well as a fixed workforce. CERN launched a production service for the physicists in July 2013 based on the OpenStack platform.

Use case 1: Data analysis facility

Many scientific analysis applications have highly sophisticated algorithms and very specific purposes. The code is often maintained by a small company or a specific academic team. These applications can age rapidly, such that maintenance on the latest operating system is difficult and keeping old hardware running is not sustainable.

A typical use case of this is at ILL, where there are around 90 legacy applications, many of which require interactive access and are subject to network latency issues. These applications are accessed by thousands of users with confidential data sets that must be kept private during the embargo period.

In addition, the significant increase in the volume of data produced by scientific facilities—often referred to as 'big data'—highlights the urgent need for adding capacity, flexibility and user friendliness to the local data analysis facility where data is also archived. The previously dominant model at scientific facilities, whereby users would simply carry their data back to their home laboratories for treatment, is no longer sufficient and should be replaced.

The EMBL-EBI Embassy Cloud offers a similar facility based on the principle of moving a researcher's analysis activity to a remote infrastructure based at EMBL-EBI that hosts the public and managed data sets that they wish to use within the analysis. This removes the increasingly costly need (in terms of time and expense) for the researcher to download and establish the public data sets and services locally that they need for their analysis pipeline. The Embassy Cloud model is being developed by EMBL-EBI to support future activities in the life-sciences community and is expected to gain in adoption as the size of the databases in this community continue to grow.

The use case can therefore be characterised by the following needs:

- to preserve legacy applications for an extended period of time beyond the lifetime of individual hardware
- to be able to easily replicate the analysis with additional datasets
- to access the application remotely with performance close to the local desktop
- to provide access to public and confidential datasets on the remote infrastructure
- to maintain high levels of data confidentiality

Use case 2: Secure data federations

Within the Human Brain Project at EPFL, there is a need to securely share data across multiple providers distributed around Europe. A federated identity management system is a pre-requisite to this use case to ensure that the data access is consistently managed and securely protected. With a large number of data sets, high quality metadata is vital to ensure that they can be found, especially by other researchers and in the scope of data preservation.

The falling cost of genomic sequencing is providing the opportunity for genetic analysis to become a routine diagnostics tool. Such personalised medicine has the potential to improve the targeting of drugs to an individual's specific medical and physical condition. However, for such an approach to demonstrate benefit, research needs to be undertaken with sequence data that is associated with information that could potentially identify the patient. Federation of such data, even if pseudo-anonymised, needs to be undertaken securely and aligned with the consent given for the initial collection, as well as remaining inside the legal jurisdiction it was collected in.

EMBL-EBI hosts the European Genome-phenome Archive (EGA)[18] on behalf of a range of data providers and is looking to securely federate similar archives within the ELIXIR research infrastructure. The EMBL-EBI Embassy Cloud is one approach that avoids moving sensitive data sets across national boundaries to where researchers wish to undertake their analysis, by allowing authorised researchers to bring their analysis to the sensitive datasets.

This use case can be characterised by the following needs:

- to securely share data across multiple sites with strong and consistent access control
- to rapidly identify the correct data set given metadata queries

Use case 3: Remote management of analysis facility

Medical labs, hospitals and clinics generally do not have local operators or system managers to run clouds on site. However, in view of patient confidentiality, there is a requirement for the data to remain at the location.

WLCG tier-2 data centres often have limited resources and skills to perform complex system administration tasks.

A cloud resource can be geographically distributed but centrally managed (within the constraints of allowing access to the hypervisor remotely). A distributed cloud could provide a balance between data privacy and cost-effective management. A regular hardware maintenance contract, similar to a photocopier contract, could ensure that the hardware is kept in a working state.

This use case can therefore be characterised by the following needs:

- to securely protect sensitive data according to policies
- to provide local compute resources to process the data
- to enable remote management to avoid the need for dedicated highly trained staff

Use case 4: Provisioning research clouds at scale

Organisations are requiring more significant compute resources to support their research programs. At the same time, there is a need to maximise the efficiency of the use of these resources and the workforce required to run them.

As clouds expand, problems are arising concerning the measurement of usage at any given time, the identification of areas for efficiency improvements, the opportunistic use of spare resources, capacity planning, and the fair sharing of resources according to appropriate priorities. In growing the clouds, the ease of use for researchers must also be maintained.

---

[18] www.ebi.ac.uk/ega/

This use case can, therefore, be characterised by the following needs:

- to allow clouds to be scaled without an associated increase in the workforce cost
- to maintain simple, seamless access to resources in a familiar way
- to rapidly identify efficiency improvements and workloads which can exploit the spare resources
- to integrate resources from public and private-sector providers within the restrictions specified by the user

Use case 5: Hardware support for large-scale desktop computing

The concept of exploiting the idle time of desktop computers to contribute voluntary computing power to research initiatives is not a new one. It has been used on many occasions[19] in the past to run large quantities of small, largely independent computing jobs. This model is generally known as volunteer or desktop computing.

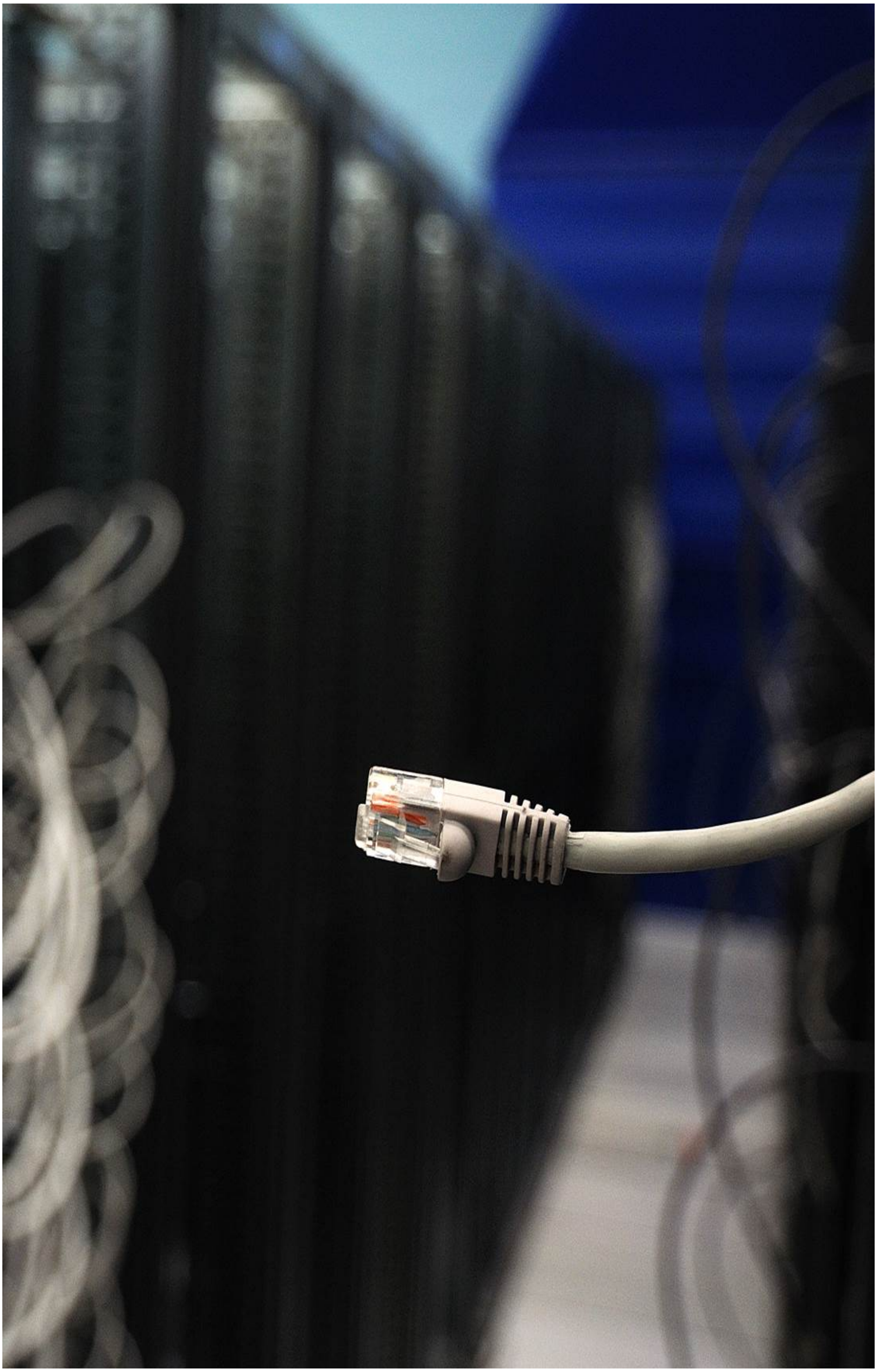
The current approach is based on the use of specialised client-server programs[20] to manage the distribution of jobs and the collection of results. Recently, such programs have been extended with the possibility of distributing self-contained, preconfigured virtual machines to the host computers. The main difficulties with this model are the inherent unpredictability of the host desktop environment, the need to install and maintain client software on the hosts, and the security and confidentiality risks associated with running programs and distributing data on unmanaged PCs.

However, many organisations have large installed bases of desktop computers in their infrastructures that are often idle for significant fractions of the day or during the night. Most of these desktop computers often have enough resources (memory, disk, and processor cores) to run many tasks at the same time. The possibility of having hardware-level support for deploying virtual machines and using advanced virtualisation and trusted computing features could make this model viable as a means of building private computing infrastructures to run lower-priority, highly parallelisable tasks on potentially very large numbers of processor cores. This model could be suitable for many classes of data analytics applications.

---

[19] http://setiathome.ssl.berkeley.edu/, http://lhcathome.web.cern.ch/, or http://folding.stanford.edu/home/ among many examples

[20] http://boinc.berkeley.edu/ is the most used and known



Challenge 5: Networks and connectivity

Most of the CERN infrastructure is controlled and managed over a pervasive IP network. Safety and access controls for the accelerator complex use communication channels over IP, with robots even being used to carry out remote inspections of dangerous areas and giving feedback over Wi-Fi and GSM IP networks. In total, there are over 50,000 devices connected to the CERN network, which is comprised of roughly 5,000 km of optical fibre out of a total of 40,000 km for all networks on CERN sites. This whole network is operated and monitored by the CERN Network Operation Centre.

Networking also plays a vital role in data acquisition from the experiments. Collision data flows from the detectors to the first level of filter systems at a staggering 3 PB/s. Good networking is, of course, also paramount to the success of the WLCG. WLCG uses a tier structure with the CERN data centre as tier-0. CERN sends out data to each of the 12 major data centres around the world that form the first level, or tier-1, via optical-fibre links working at multiples of 10 Gb/s. Each tier-1 site is then linked to a number of tier-2 sites, usually located in the same geographical region.

The large tier-1 data centres are linked to CERN, as well as to one another, via the LHC Optical Private Network. Meanwhile, the LHC Open Network Environment is facilitating evolution of the WLCG model by enabling flattening of the strict tier-1-tier-2-tier-3 hierarchy model, so that any site may connect with any other as demand necessitates. It is thanks to improvements in networking over recent years that CERN has now been able to expand the capacity of its tier-0 data centre with a remote facility in Hungary. The Wigner Research Centre for Physics was inaugurated last year and has approximately 500 servers, 20,000 computing cores, and 5.5 PB of storage already operational. The dedicated and redundant 100 Gbit/s circuits connecting the two sites are functional since February 2013 and are among the first transnational links at this distance.

In the biological and life sciences area, the ELIXIR research infrastructure is tasked with providing a European infrastructure for biological information that enables researchers from all disciplines to benefit from the rapidly growing store of information about living systems. Within such an infrastructure it will become necessary for data to be contributed from sources across the community and downloaded for analysis across Europe. It is envisaged that the major data contributors and data providers will need to be linked to Europe's high-performance networks and their international peers. To provide a reliable and resilient service it is expected that service instances will be replicated and load-balanced.

Use case 1: Ability to migrate public IP address between sites

Increasingly, IT services must be continuously available. Although, in principle, a 'well-designed' computing service will gracefully recover from interruptions, many services (perhaps most in the scientific environment) will not try to re-establish a network connection if there is any failure. It is, therefore, desirable for the underlying network infrastructure to support service continuity. The transparent migration of a public IP address from one site to another would facilitate this.

Both hardware and software issues can lead to the requirement to intervene on a system at a time that is not convenient for application services. The ability to migrate a public IP address would allow application services to be moved to other hardware without dropping any active client connection. In most cases it is sufficient to relocate application services to another server in the subnet (or broadcast domain). For complete generality, however (e.g. a planned major infrastructure outage), there is a need to migrate services between different subnets.

The expected outcome and impact for stakeholders is a successful delivery of the ability to migrate a public IP address everywhere in a routed network. This would enhance the ability of research organisations to deliver continuously available services.

For industrial partners, it is interesting to note that the ability to migrate a public IP address between independent network domains was a proposed feature of the IPv6 specification, but has never been widely supported. A successful demonstration of such a migration would likely attract wide interest.

Use case 2: Cross-site, on-demand VLANs

Today, with data transfers sharing a common network connection between research centres or between home institutes and a research hub, control and management of transfers by individual researchers can be difficult. However, with dynamically created, on-demand VLANs, individual transfers would be isolated and traceable and it would be possible to implement fine-grained quality-of-service policies.

The possibility of dynamically creating VLANs across a multi-domain network by a programmatic interface would allow network-aware user applications to take advantage of the available connectivity. Specific examples include applications that need a private environment shared between a few clients distributed around the world; a server that - for example, for security reasons - only allows communication between a tightly restricted set of clients; or a data transfer that needs special treatment in terms of quality of service or bandwidth.

The expected outcome and impact for stakeholders is that intelligent applications could help in optimising network utilisation, especially of the expensive, long-distance legs.

Use case 3: Intelligent bandwidth optimisation

Most research institutes are interconnected with multiple network paths and with a rather static configuration of load between the different paths—frequently either simply load-balanced or with one link active and another in standby mode. Delivery of an intelligent bandwidth optimisation capability would allow a more dynamic and flexible routing of traffic over available links according to a wide range of metrics.

Recent servers have the capabilities of creating single data flows that can take a large percentage of the available bandwidth of a given link (so called ‘elephant flows’). Imagine two sites connected by three equal capacity links where a server starts sending elephant flows. Today, the router connected to the three links will apply basic load-balancing algorithms that cannot take into account the size of the stream. This may cause congestion over one link, whilst the other two remain idle. Intelligent bandwidth optimisation would enable a network to recognise the bandwidth demand of certain data flows, be aware of its own status, and use this information to dynamically move flows from between links for optimum overall performance.

The expected outcome and impact for stakeholders is that an intelligent network that learns what traffic it is carrying and where the ‘pain points’ are could deliver improved service to users and optimise overall costs.

Use case 4: Latency-tolerant wide-area NFS

Individual, high-speed data transfers would be a benefit to many researchers. In most cases, researchers are interested in ensembles of files and wish to replicate tree-like file hierarchies at many sites. AFS[21], first developed in the 1980s, was a promising development but the technology is now old and not effective for exchanging large volumes of data across long-distance links. Similarly, although the CernVM File System with its exploitation of HTTP caches is highly effective as a way of hierarchically distributing read-only data, it cannot address the need for true bidirectional data exchange. A modern, efficient implementation of an AFS-like global file system is required.

With such technology, members of worldwide research collaborations could—whether located at an institute anywhere in the world or at home—make files available to fellow collaborators simply by writing them to the file system. The existence of the file would be made known 'instantly', but actual transfer of the file to another site could be scheduled 'lazily' or be done through transfer from dedicated nodes in a CDN-like model. Policies should enable the automatic creation of replica copies for redundancy and availability purposes.

The expected outcome and impact for stakeholders is that such a facility would clearly ease the life of researchers, with policy-based configuration of file systems replacing active management of file transfers and file catalogues.

Use case 5: Layer-3 roaming across Wi-Fi and mobile telephony data services

Large organisations such as CERN and other EIROForum members generally have extensive Wi-Fi coverage - and may even have enabled layer-3 roaming - within buildings, but cannot guarantee network connection continuity for users moving between buildings on their campuses. Although extension of the Wi-Fi network to cover outdoor areas might be possible, a more attractive alternative could be to link Wi-Fi with today's ubiquitous mobile telephony-based data services to provide a seamless roaming capability.

The expected outcome and impact for stakeholders is that successful delivery of such a capability would greatly benefit mobile staff on research campuses. It would also be of interest to any other organisation with geographically distributed Wi-Fi services.

Use case 6: Intelligent, self-adjusting and self-healing Wi-Fi networks

With the explosion in use of, and reliance on, mobile devices, Wi-Fi networks are expected to be as reliable, performant and efficient as wired connections. However, ensuring optimum Wi-Fi services for large buildings is difficult today and will only become more challenging still with the advent of higher densities of Wi-Fi-enabled devices and the wider deployment of 802.11ac technology[22]. Competition for radio spectrum and medium access are major factors: adding more access points can often make the situation worse, not better, as channel coverage areas overlap and spectrum bandwidths are saturated. The radio frequency environment is also dynamic; statically optimised radio frequency plans cannot always cope with the sudden

---

[21] Andrew File System, http://www.openafs.org/

[22] http://standards.ieee.org/findstds/standard/802.11ac-2013.html

changes that occur on wireless LANs. Access point failure and the need for higher bandwidth, higher density of coverage, or both are difficult to predict and often impossible to respond to in real time. Similarly, large auditoriums need many access points to support audiences of 500 or more, many of whom may have two active devices, or even more. Yet, for much of the time, most of these access points will be unused, wasting power and, potentially, interfering with the coverage of access points nearby. In short, today, a modern Wi-Fi infrastructure requires real-time flexibility.

In an intelligent, self-adjusting and self-healing Wi-Fi network, access points are ideally installed at regular intervals and communicate amongst themselves to adjust 802.11 parameters, in order to provide an optimal service in any situation. These over-provisioned access points would probe the network, self-activating or deactivating as necessary and raising alerts to signal any need for human intervention (e.g. to install additional hardware).

Use case 7: Wireless networks for data acquisition

A major issue in the design of particle physics experiments is the routing of information out from detectors at the heart of the experiment. Copper cables and fibre-optic links both add material to the detector and can absorb or deflect particles as they travel outwards. Wi-Fi—or, more generally, radio-frequency links—would remove this problem, in addition to eliminating—or at least reducing—cabling costs. To cope with the volumes of data involved—today equivalent to the simultaneous streaming of several very high speed 4K videos—the wireless network would have to operate, in a high-radiation, high magnetic field environment, at bandwidths exceeding 100Gbit/s and without losing a single frame. Although this use case refers to the CERN data acquisition systems as an example of high-rate, harsh environments, it has more general applications in any situation where the removal of cabled connections without loss of performance could provide higher flexibility and reduced costs of deployment and maintenance.

Use case 8: Secure, high-bandwidth mobile communication technology

Despite, or perhaps because of, the communication improvements delivered by the switch to digital technologies (e.g. Terrestrial Trunked Radio), police forces, fire and rescue services, and other civil protection agencies are looking for secure, high-bandwidth communications technologies to improve their operational capabilities and effectiveness. For example, use of multiple live video streams between the field and command centre would greatly improve remote live assistance and enable post-response analysis of incidents, with the possibility of keeping videos as evidence. Fire and rescue services could also benefit from enhanced data capability to remotely receive safety procedures via augmented reality, thus enabling safer and more rapid intervention in unfamiliar environments.

Whilst there are ongoing discussions regarding the allocation of dedicated Long-Term Evolution frequencies for such purposes, the underlying needs might be better addressed by 5G technologies. The unique environment at CERN presents many opportunities for the exploration of innovative ideas concerning the applicability of such technologies for critical communication applications. These opportunities include CERN's 350-terminal advanced Terrestrial Trunked Radio network, the organisation's 2G/3G mobile network covering over 65km^2, and its multi-national fire and rescue service—as well as significant IT, networking, and telecommunication expertise. Additionally, the underground areas at CERN offer the possibility of broadcasting a wide range of radio frequencies over a large area with no risk of perturbing public mobile services and are also one of the few locations where radiation resistance can be evaluated in a real-world environment.


12
15
16
17
18
19
22
23
24
25
26
29
30
3

Challenge 6: Data analytics

During the past decades, CERN and other internatonal research laboratories have been gathering not only enormous amounts of scientific data, but also very large quantities of systems-monitoring data from their instruments. Curating, enriching and managing this data would enable its exploitation. Added value could be obtained in terms of increasing knowledge of the engineering systems, enabling better delivery of services to the scientific community, and helping appropriate decisions to be taken during the lifecycles of the systems and instruments.

The investigation of state-of-the-art data analytics at CERN in this sector has been carried out by interacting with the various CERN groups and by organising dedicated workshops with the participation of CERN engineering teams and representatives of other international laboratories. The outcome of the workshops has been the definition of the major challenges in this domain and a set of relevant data analytics use cases. A general interest in having a common data analytics platform has also been identified.

The main challenges in data analytics for scientific and engineering applications involve technological, integration and educational aspects. This area is rather new in the R&D activities of CERN openlab and it is therefore described in extended detail.

Technology challenges

Near-real-time processing

The challenges of near-real-time processing are focused on the ability of processing large amounts of data (GBs per second) with low latency (in the order of seconds) coming from different sources and domains. The tools and techniques should be flexible enough to import and apply the knowledge inferred from the data (batch analysis), or the pre-existent (human) knowledge of the systems. They may include models, pattern definitions and matching, thresholds, and—most importantly—the ability to trigger actions based on the discovery of complex events within the streaming data.

Due to the critical nature of the services the different systems offer, and the huge data volumes they produce and manage, it is important to consider some mandatory aspects, such as scalability, fault-tolerance, and the ability to guarantee that all the streaming data produced is processed and analysed. In addition to the aforementioned factors, the near-real-time processing system needs to integrate the different domains and analysis technologies that currently exist at CERN. Therefore, it is vital to support a wide range of data analysis tools and programming languages.

Batch processing (including predictive analytics)

Batch processing is meant to analyse the data coming from the repositories, learn from the past and apply that knowledge to better understand current and future systems. The main challenge is to mine and analyse huge amounts of structured and unstructured data coming from various repositories, relational databases models, and NoSQL.

Most of the near-real-time analysis methods that exist today are based on our knowledge about the underlying systems. However, there are patterns or correlations among large quantities of data that are not immediately visible to humans and that could be discovered by other means, like machine-learning tools. This would provide much more reliable real-time or predictive information. Accordingly, it represents one of the main challenges for both the present and the future. Linear and non-linear modelling, classical statistical tests, complex time-series analysis and forecasting, classification, and clustering are just some of many techniques used with this purpose.

Nowadays, the trend in industry is to spread the data computation to multiple nodes while performing the analysis as close to the data as possible. This is totally inline with CERN's requirements in terms of analytics. Hence, concepts such as data locality and in-database analytics will play an important role, although we will investigate other solutions.

An expected consequence of this type of analysis is the possibility of improving not only the measurable efficiency of the systems, but also the relative user perception. ESA, for example, is investigating ways of improving the user perception of its IT systems by implementing special tools to monitor users' behavioral patterns. This entails deep analysis of all applications, including study of their usage through log files and other means. In the long run, it is envisaged that such predictive analytics techniques can not only improve the efficiency of the systems but also help idenfy areas where improvements or investments should be focused.

Data repositories

One of the main objectives and challenge for any data analytics framework is to integrate the existing data repositories, rather than replace them. In an environment as specialised and heterogeneous as CERN's the previous statement is even more valid.

The extract-transform-load (ETL) processes, required for loading the data from the repositories while standardising and reshaping it to make it suitable for complex analyses, is itself a challenging task, given the variety of formats and the nature of the data gathered and stored.

In addition to the integration with the existing data repositories, the upcoming requirements lead to the need to evaluate a potential general solution, capable of storing huge amounts (hundreds of TBs) of structured and unstructured data safely and for the long term.

Integration challenges ('data analytics as a service')

The three main areas defined in the previous section can be grouped together to provide a self-contained solution for 'data analytics as a service', which is CERN's main goal for the long term.

The platform should be a set of technologies for every layer tool, since it is clear from the requirements collected that there is no 'one-size-fits-all' solution. This set should be encapsulated in a common framework capable of easily transferring data between the layers and the tools, so as to let the analyses be performed with the most appropriate solutions.

The data analytics platform should be accessible for some of the use cases to external institutions willing to cooperate. It should, therefore, use open and well-defined standards for exchanging the data. This is also important in terms of future support.

Educational challenges

Training the next generation of engineers/employees, disseminating results, and outreach to new audiences are key goals of CERN openlab. This is particularly challenging in the case of data analytics, since many domains of expertise are involved and therefore need to be covered.

The challenges described so far can be exemplified by a number of specific use cases identified from the experience of the CERN LHC engineering and control teams, CERN IT teams, and equivalent teams in other laboratories or projects.

Use case 1: The CERN Accelerator Logging Service

The logging service stores data of close to one million pre-defined signals coming from heterogeneous sources. These signals range from data related to core infrastructure such as electricity, to industrial data such as cryogenics and vacuum, to beam-related data such as beam positions, currents, beam losses, etc.

The logging service provides access to logged data for more than 700 registered individuals, over 100 registered custom applications from around CERN, and even off-site access for purposes such as the CERN Neutrinos to Gran Sasso experiment. Data extraction is performed via a custom tool and an extraction application programming interface, which can extract time series data from multiple data sources simultaneously. Currently, there are two different points of interest in data analysis: the improvement of the CERN Accelerator Logging Service infrastructures using complex event processing and the possibility of enabling users to perform custom data analysis.

In this context, the main focus is on accelerator operations, to make common data analysis use cases easier, to provide simple and fast access to data, and to save and share analysis results. The key idea to follow is to perform analysis as close to data as possible, and the main challenge is the offline (batch) analysis performance. The ability to offer different interfaces for different users is important for some of the user community; working with a structured native query language would be a great improvement.

Use case 2: CERN industrial control systems

The Industrial Controls and Engineering Group in the CERN Engineering Department develops solutions and provides support in the domain of large and medium scale industrial control systems, as well as laboratory control systems. Currently, they support five major installations: ALICE, CMS, ATLAS, LHCb and the accelerator complex, each of them with hundreds of logical boards and millions of parameters. As a result, a huge amount of data is acquired and stored in collaboration with the control groups of the accelerator and the experiments (for example the Accelerator Logging Services described above).

In terms of data analytics, the main goal for the Industrial Controls and Engineering Group is to develop a data analytics framework. It should be a common solution for data analytics needs in the different equipment groups, such as cryogenics, gas, vacuum, machine protection, etc. The main challenge, taken by the Controls Group, is to integrate the data analytics functionalities with the current CERN control system (mainly through the SIMATIC WinCC OA SCADA system) and expose it as a service. Thus, each group could use it to perform its own custom analysis based on the knowledge of the system experts. To achieve this, several use cases are being explored in the fields of control system health, threshold learning for alarms, and root cause analysis. We have to face different issues for both batch and near-real-time analysis: data heterogeneity, data access (i.e. sensible or protected information), data synchronisation, different data source formats, data classification, data completeness, etc.

The challenges on the infrastructure to achieve 'analysis as a service' lie in the ability to succeed in the following three aspects:

- Near-real-time analysis is vital for any of the analysis that needs to run in continuous mode to generate alarms, commands, or reports, as well as live visualisation.
- It is necessary to perform batch analysis on historical datasets in order to learn from and visualise their content; for development and training, it is also important to memorise the results of previous analysis.
- The whole system should be scalable and fault tolerant, while remaining easy to use for domain-experts.

Use case 3: IT monitoring

The historical motivation of the IT Monitoring project resides in several independent monitoring activities carried on over the years in the IT field. All employ a similar overall approach and have similar limitations, but use different tool-chains. Moreover, given that high-level CERN IT services have largely interdependent workloads with obvious common characteristics (accelerator run or not, experiment data taking or not, etc.), the combination and correlation of data from different services becomes necessary to have a thorough understanding of the end-to-end systems chain. In this context, understanding performance is becoming more important, and requires more combined data and complex analysis. It was, therefore, necessary to find a shared architecture and shared tool-chain components.

The sub-use cases in this field include real-time analytics for taking automated decisions, dashboards and interactive analytics, and data mining. In the future, the list of analytics use cases should grow, enabled by technology. In this context, the main objective for the IT Monitoring project is to offer 'analytics as a service' for computing services.

Analytical tools can also study the behaviour of users regarding dissection of a session in terms of functionalities used in a system and can help to optimise and make systems more user-friendly. Such information can be established from a set of correlated logs (if multiple systems are in use), as well as request tracker data from service desks providing first-line support.

Use case 4: Intelligent data placement

EMBL-EBI archives data that is continually being updated for many life-sciences research communities. For these communities to undertake their research they need access to the latest data on their local resources. To support the diverse data analysis that will take place within ELIXIR, the ability to 'push' data from a provider to major analysis centres, or for the major analysis centres to 'pull' the required data set from a nearby source, becomes a critical capability.

In the ATLAS experiment, the Distributed Data Management project (DDM) manages the experiment's data transfers on the WLCG. The use cases include trace mining (user interactions with DMM), popularity (used for deciding which data to delete), accounting (reports on data contents), and log file aggregation. Currently, popularity is used only in an automated deletion decision, but the next steps are to use the popularity to make new copies of datasets (i.e. forecasts about future dataset popularity; decisions about how many datasets to delete and where; decisions on where to replicate new copies for the different datasets, etc.).

The CMS experiment also makes considerable use of the distributed grid resources for the storage and offline analysis of the collected data. The current data management model is workforce intensive and results in inefficient usage of disk space. Projections for the second run of LHC imply a factor-of-six increase in required computing resources, and it is therefore necessary to optimise the usage of the current resources. CMS already has a data popularity service in production, which works by monitoring the patterns of usage of accessed data samples and can provide automatic identification of obsolete replicas per site. The next steps are to extract further knowledge from the monitoring data in order to implement effective data placement.

Use case 5: Network monitoring

Existing and emerging research infrastructures are largely distributed across tens of hundreds of different resource sites. The sites provide both computing and data storage resources and are connected via the existing academic networks (such as GÉANT in Europe) or commodity commercial networks.

The WLCG has chosen to deploy an infrastructure for network performance monitoring, which constantly monitors the network ‘health’ (bandwidth, latency, route and ping across the sites). Similar approaches are also being taken elsewhere, such as with the resource network managed by the ELIXIR initiative for the life sciences community.

While a lot of data has been recorded, making some sense out of it is not obvious: measurements span different time periods, they measure different parameters (while all related to network), and they might be affected by other measurements and/or events. Some of the questions to be addressed are: during a bandwidth test, was there any known activity in the same link? If there is an unstable link, does this appear as degraded performance somewhere at the same time? If a loss of performance is registered in some network link, is there also a network problem and where?

To make sense out of the data and respond to these questions, a data analytics approach to understand correlations in both time and topology is required. A possible tactic is to analyse the existing data and mine the information, looking for known issues in the past to learn the patterns. Then, make the system predictive, identifying issues as they appear, before the users notice them.

Use case 6: The CERN Advanced Storage Manager (CASTOR)

CASTOR is the mass storage solution of CERN, including LHC data. It is a hierarchical storage system based on both disks and tapes (12,000 disks, 30,000 tapes), which hold more than 100 PB of data.

This infrastructure generates a lot of monitoring data: up to 20 GB per day (~100 million messages), totalling ~10 TB per year. This data is stored in a long-term repository for auditing, error recovery, and historical studies (e.g. usage of protocols). It is processed live for display and simple online analysis (time series and histograms) in a dashboard.

The current system does not cover two important topics: an ‘expert system’ for spotting interesting time series out of large monitoring datasets (so as to avoid time-consuming eye inspection) and an early warning system to forecast potential dangerous situations, such as overloads.

Use case 7: ESA and serendipity in the data archives

The archives in astrophysics are ever growing as more data becomes available about the universe that surrounds us. Astrophysics itself has become a multi-disciplinary domain, where the wavelength region becomes the driving factor. As much attention is focused on specialised fields in those domains, be they stellar or extra-galactic astronomy, little interaction occurs between the disciplines.

By using sky coordinates as the common paradigm, one can envisage 'intelligent' bots doing much of the reasercher's work in scanning the archives to collect relevant information in a particular field. Such automated bots would present their results only when called upon and only focus on a problem at hand (e.g. "give me serendipitous objects in the X-ray range lying around the Crab Nebula, since an unexplained region of hot gas may have an effect on the infra-red region I am studying…"). The bot may be further refined to extract only very good quality data from all X-ray missions or for a given time period.

Use case 8: Analytics and modelling for availability improvement in the CERN Future Circular Collider

The Future Circular Collider (FCC[23]) study develops options for potential high-energy frontier circular colliders at CERN for the post-LHC era. This conceptual design study is targeted to be available by the end of 2017, in time for the next update of the European Strategy for Particle Physics.

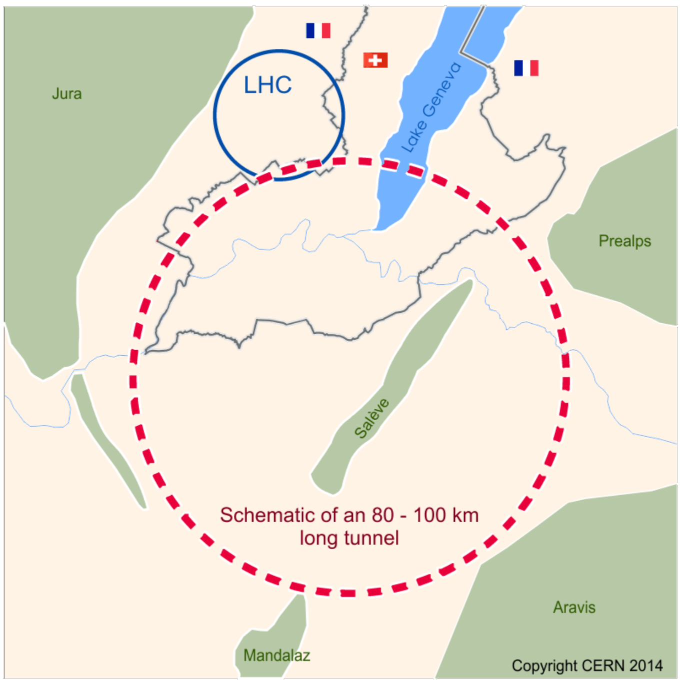


Figure 1: Schematic of future circular collider

---

[23] http://press.web.cern.ch/press-releases/2014/02/cern-prepares-its-long-term-future

The LHC and the studied 100-km-long FCC options are large-scale 'Systems of Systems'. In-time data analytics can contribute to improve availability and operational efficiency of accelerator installations by timely detecting deviations from expected behaviour that eventually may lead to faults. *In-silico* near-real-time modelling of the accelerator system and its infrastructure services would further improve early-warning capabilities, permit preventive maintenance and enable co-scheduling of fault-prevention interventions. Such analysis can be based on laws and correlations derived from large quantities of monitoring data using data analytics tools.

Outcomes of these assessments may also result in applications for industrial installations. Examples include (but are not limited to) large-scale manufacturing, processing and chemical plants, railway transportation systems, and utilities networks.

For this study, real-world use cases taken from LHC accelerator operation shall serve as the basis for developing formal data analytics scenarios. This will lead to the creation of a catalogue of functional and performance requirements on data analytics infrastructures for systems consisting of large numbers of components and diverse data sources. Consequently, a reference data analytics ecosystem to be established at CERN could be used to train operators in data analytics techniques to test malfunction prediction hypotheses.

This investigation should foresee the following activities:

- Developing a requirements catalogue for data analytics ecosystem.
- Conducting training in use of end-user analytics systems.
- Formulating analytics use cases, analysis of LHC operation data, formulation of subsystem and component fault dependencies and testing of hypothesis.
- Carrying out gap analysis of the emerging data analytics ecosystem, with respect to future needs and performances.
- Developing a requirements catalogue for 'system of systems' simulation infrastructure.
- Assessing potential of future *in-silico* modelling and predictive data analytics capabilities for improvement of collider operation availability and scheduling of preventive maintenance.

Use case 9: Data analytics on scientific articles

Two interesting challenges for scientific publications and datasets management systems are *information* extraction and discovery and *knowledge* extraction and discovery. The first case concerns the automated extraction of information about authors, references, key words, etc. The second case concerns semantic analysis of text, enabling identification of the main field, key words that do not appear in the text, sentiment of references, etc. This would provide the ability to join and correlate concepts from different domains and publications.

A number of initiatives exist where similar challenges could be best addressed:

> **INSPIRE:** CERN, DESY, Fermilab and SLAC have built the next-generation HEP information system, INSPIRE. It combines the successful SPIRES database content, with the Invenio digital library technology developed at CERN. INSPIRE represents a natural evolution of scholarly communication, built on successful community-based information systems, and provides a vision for information management in other fields of science. The information extraction algorithms of INSPIRE are currently

tuned for the HEP domain, so the challenge is to develop more inclusive algorithms that extend it to a wider range of domains[24].

ZENODO: Using the same code base as INSPIRE, a service for scientists across all domains has been developed, so as to capture 'the long tail' of science[25].

ORCID: ORCID provides a persistent digital identifier that distinguishes you from every other researcher and, through integration in key research workflows such as manuscript and grant submission, supports automated linkages between you and your professional activities, ensuring that your work is recognised[26].

Extending INSPIRE, ZENODO, ORCID and other similar systems with features and analytics capabilities would open interesting paths of collaboration, possibly across multiple disciplines. Overcoming the challenges mentioned will be extremely valuable for many research and scientific organisations and projects, such as CERN, ESA and the Human Brain Project.

As an example, the Human Brain Project seeks to mine large numbers of scientific papers and extract structured information from them (e.g. concentration of proteins in different cell types). There are also international efforts to publish neuroscientific data in structured databases and repositories. However, a much larger part of information and knowledge is embedded in the bodies of scientific publications and is inaccessible for automatic consumption. This is of particular concern as the the number of neuroscientific publications per year is growing and the absolute numbers make it impossible to harvest the information manually. Therefore, the Human Brain Project is investing in novel search and extraction mechanisms to both guide the attention of researchers and to automate the extraction process completely. Numerous informatics challenges are involved in this task: classical informatics for large-scale text search and mining, creation and curation of ontologies, information extraction, ranking of information and confidence levels, provenance, and reasoning. All of this has to be combined with the domain knowledge and domain-specific validation.

Use case 10: Administrative information systems

The CERN Administrative Information Systems (AIS) group is responsible for providing administrative applications for CERN and therefore handles huge amounts of data that need to be exploited by the system users for different purposes. Among the maintained applications are the corporate ones like the financial system Qualiac, the stores system Baan, and the human resources system Oracle HR. These corporate back-office systems are complemented by a number of in-house-developed reporting and analysis solutions for the wider CERN population that have been implemented over the past 20 years. A selection of these is given below:

- CERN Expenditure Tracking application (CET) for budget tracking by the different budget holders at CERN and more detailed reports for the finance department.
- Human Resource Toolkit (HRT) for personnel reporting for the supervisors at various levels.
- Management Data Layer (MDL) – a 'data warehouse' used as base for various

---

[24] http://inspirehep.net/
[25] http://zenodo.org/
[26] http://orcid.org/

business objects reports.
- Activity Planning Tool (APT) for project resources planning.

The AIS group started the ‘AIS-20’ project in 2012. One of its objectives is to streamline how business-critical information is provided to users. On the technical level, this requires the unification of different databases and extractions by building a unique, common AIS data warehouse. This new data warehouse acts as a main data source for the aforementioned reporting/analysis solutions to be replaced by a new common reporting tool.

Populating this common data warehouse presents several technical challenges, which need to be considered. These include:

- Being able to refresh all or parts of the data warehouse within increasingly diminishing time windows available for complex extraction processes.
- Live (or near-real-time) refresh of certain selected data.
- Integration with a messaging-based enterprise service bus (ESB) system.
- Making the data available using a bi-temporal model, while maintaining good query performance and minimising the impact on the amount of stored information. In the bi-temporal model, one time dimension comes from the business (e.g. contractual dates) and the other one is purely technical (it indicates when data was effectively part of the data warehouse and allows writing queries using a ‘show-data-as-of’ date. The goal is to be able to answer business questions like ‘what has been visible on the principal financial dashboard of the organisation on any given date in the past?’.

The goal of the project is to continue elaborating technical solutions for the given problems, to demonstrate the feasibility of the presented solutions and to provide a production-ready partial implementation as part of a new, fully managed AIS data warehouse push- and pull-based data population process.

About CERN openlab

CERN openlab is a unique public-private partnership between CERN and leading IT companies. Its mission is to accelerate the development of cutting-edge solutions to be used by the worldwide LHC community.

Within this framework, CERN provides access to its complex IT infrastructure and its engineering experience, in some cases extended to collaborating institutes worldwide. Testing in CERN's demanding environment provides the IT industry partners with valuable feedback on their products while allowing CERN to assess the merits of new technologies in their early stages of development for possible future use. This framework also offers a neutral ground for carrying out advanced R&D with more than one company.

CERN openlab was created in 2001 and is now in its fourth phase (2012-2014). This phase addresses topics crucial to the CERN scientific programme, such as cloud computing, business analytics, the next generation of hardware, and security for the myriads of networks devices. Phase V will start in 2015.

CERN openlab Industry Members 2014

Partners

HUAWEI

intel

ORACLE

SIEMENS

Contributor

rackspace
the open cloud company

Associate

Yandex

Additional information

Related events:

- IT requirements for the next generation research infrastructures workshop. CERN, 1st February 2013.
  Presentations and video of the event available at:
  http://indico.cern.ch/event/212402/
- CERN openlab Workshop on IT Challenges in Scientific Research. CERN, 10-11 December 2013
  Presentations and additional material available at:
  http://indico.cern.ch/e/openlab-challenges
- CERN openlab Data Analytics Workshop. CERN, 20 February 2014
  Presentations and additional material available at:
  http://indico.cern.ch/event/289770/

Related documents:

- EIROforum IT working group (2013). e-Infrastructure for the 21st century. Publication prepared by CERN on behalf of EIROforum IT Working Group.
  http://zenodo.org/record/7592
- Realising the full potential of research data: common challenges in data management, sharing and integration across scientific disciplines. Field, Laurence; Suhr, Stephanie; Ison, Jon; Los, Wouter; Wittenburg, Peter; Broeder, Daan; Hardisty, Alex; Repo, Susanna; Jenkinson, Andy
  http://zenodo.org/record/7636
- CERN openlab annual reports
  www.cern.ch/openlab/resources/annual-reports
- CERN openlab brochure & guiding principles
  www.cern.ch/openlab/becoming-sponsor

www.cern.ch/openlab

EXECUTIVE CONTACT

Bob Jones, Head of CERN openlab

bob.jones@cern.ch

TECHNICAL CONTACT

Alberto Di Meglio, CERN openlab Chief Technology Officer

alberto.di.meglio@cern.ch

COMMUNICATION CONTACT

Mélissa Gaillard, CERN openlab Communications Officer

melissa.gaillard@cern.ch

CERN openlab

www.cern.ch/openlab