~~Independent Studies IV~~
CERN Spring Campus 2015

Nuno Ricardo Gago Pinto

Activities Report

Abstract—This report is intended to describe in summary the activities performed during the 2015 CERN Spring Campus with the main focus on the latest technologies and industry trends in IT. This event has brought together experts from CERN, to share on a very open, simple and direct approach in a program of scientific and technological dissemination and cultural exchange.

Index Terms—(CERN, soft skills, industry trends, technology).

The document mostly describes CERN, not the SPRING CAMPUS activity!

1 Introduction

THE European Organization for Nuclear Research, known as CERN (derived from the name "Conseil Européen pour la Recherche Nucléaire") is a European research organization that operates the largest particle physics laboratory in the world. Established in 1954, the organization is based in a northwest suburb of Geneva on the Franco–Swiss border, and has 21 European member states [1].
In the 2015 CERN Spring Campus was schools dedicated to Information Technology and Computing, over 3 intensive days, this event has brought together experts from around European Organization for Nuclear Research (CERN), to meet with future engineers, like me, and scientists in a program of scientific and technological dissemination and cultural exchange.
The core program consisted of 25 hours of sessions, delivered in English from specialists in their respective fields, covering topics such as:

- Information Technology
- Team work
- Science and society
- Entrepreneurship
- Project Management
- Writing a Curriculum Vitae
- Job Interview Techniques

Along with the lectures, the campus included 2 social events where speakers and students shared time together in a very open and friendly environment, opening a door to the students to know more about CERN or ask questions related to the school or even the complex scientific subject presented.

2 What CERN is?

At the moment CERN's main function is to provide the particle accelerators and other infrastructure needed for high-energy physics research.
As a result, numerous experiments have been constructed at CERN as a result of international collaborations, which in 2013 had 2.513 staff

~~Figure 1.~~ European Organization for Nuclear Research (CERN)

• *Nuno Ricardo Gago Pinto, nr. 38202,*
E-mail: nuno.gago.pinto@tecnico.ulisboa.pt,
Instituto Superior Técnico, Universidade de Lisboa.

Manuscript received June 06, 2015.

	ACTIVITY						DOCUMENT						
(1.0) Excellent	Object ×2	Opt ×1	Exec ×4	Summ ×.5	Concl ×.5	SCORE	Struct ×.25	Ortog ×.25	Exec ×4	Form ×.25	Titles ×.5	File ×.5	SCORE
(0.8) Very Good													
(0.6) Good	0.4	0.4	0.5	0.4	0.4		0.8	0.6	0.4	1.0	0.8	0.6	
(0.4) Fair													
(0.2) Weak													

members, and hosted some 12.313 fellows, associates, apprentices as well as visiting scientists and engineers representing 608 universities and research facilities and 113 nationalities. In the following Figure 2 is shown the member present in CERN.
CERN is also the place where the World Wide

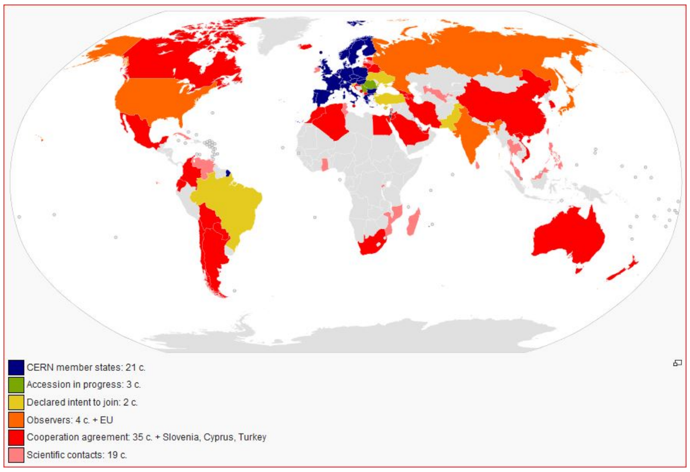

Figure 2. CERN International Relations.

Web (WWW) was first implemented. The main site at Meyrin has a large computer facility containing powerful data processing facilities, primarily for experimental-data analysis; because of the need to make these facilities available to researchers elsewhere, it has historically been a major wide area networking hub.
Prior to the Web's development, CERN had pioneered the introduction of Internet technology, beginning in the early 1980s. A short history of this period can be found at CERN.ch. More recently, CERN has become a facility for the development of grid computing, hosting projects including the Enabling Grids for E-sciencE (EGEE) and Large Hadron Collider (LHC) Computing Grid. It also hosts the CERN Internet Exchange Point (CIXP), one of the two main internet exchange points in Switzerland.

2.1 Current complex

Everything starts by the basic element of Hydrogen particles. Then CERN operates a network of six accelerators and a decelerator. Each machine in the chain increases the energy of particle beams before delivering them to experiments or to the next more powerful accelerator. Currently active machines are:

- Two linear accelerators generate low energy particles. Linac2 accelerates protons to 50 MeV for injection into the Proton Synchrotron Booster (PSB), and Linac3 provides heavy ions at 4.2 MeV/u for injection into the Low Energy Ion Ring (LEIR).

- The Proton Synchrotron Booster increases the energy of particles generated by the proton linear accelerator before they are transferred to the other accelerators.

- The Low Energy Ion Ring (LEIR) accelerates the ions from the ion linear accelerator, before transferring them to the Proton Synchrotron (PS). This accelerator was commissioned in 2005, after having been reconfigured from the previous Low Energy Antiproton Ring (LEAR).

- The 28 GeV Proton Synchrotron (PS), built during 1954—1959 and still operating as a feeder to the more powerful SPS.

- The Super Proton Synchrotron (SPS), a circular accelerator with a diameter of 2 kilometres built in a tunnel, which started operation in 1976. It was designed to deliver an energy of 300 GeV and was gradually upgraded to 450 GeV. As well as having its own beamlines for fixed-target experiments (currently COMPASS and NA62), it has been operated as a proton–antiproton collider (the SppS collider), and for accelerating high energy electrons and positrons which were injected into the Large Electron–Positron Collider (LEP). Since 2008, it has been used to inject protons and heavy ions into the LHC.

- The On-Line Isotope Mass Separator (ISOLDE), which is used to study unstable nuclei. The radioactive ions are produced by the impact of protons at an energy of 1.0–1.4 GeV from the Proton Synchrotron Booster. It was first commissioned in 1967 and was rebuilt with major upgrades in 1974 and 1992.

- The Antiproton Decelerator (AD), which reduces the velocity of antiprotons to about 10% of the speed of light for research of antimatter.

- The Compact Linear Collider Test Facility, which studies feasibility for the future normal conducting linear collider project.

- The CERN Control Centre combines control rooms for the laboratory's accelerators, the cryogenic distribution system and the technical infrastructure. It holds 39 operation stations for four different areas – the LHC, the SPS, the PS complex and the technical infrastructure.

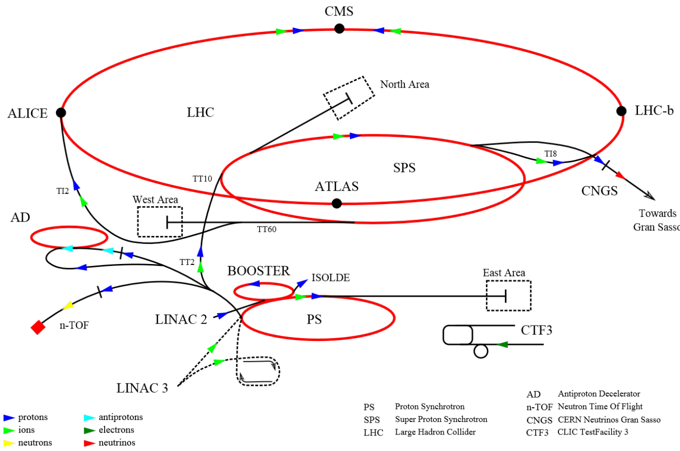

Figure 3. Map of the CERN accelerator complex.

2.2 Large Hadron Collider

Most of the activities at CERN currently involve operating the LHC, and the experiments for it. The LHC represents a large-scale, worldwide scientific cooperation project.

The LHC is the world's largest and most powerful particle accelerator. It first started up on 10 September 2008, and remains the latest addition to CERN's accelerator complex. The LHC consists of a 27-kilometre ring of superconducting magnets with a number of accelerating structures to boost the energy of the particles along the way.

Inside the accelerator, two high-energy particle beams travel at close to the speed of light before they are made to collide. The beams travel in opposite directions in separate beam pipes – two tubes kept at ultrahigh vacuum. They are guided around the accelerator ring by a strong magnetic field maintained by superconducting electromagnets. The electromagnets are built from coils of special electric cable that operates in a superconducting state, efficiently conducting electricity without resistance or loss of energy. This requires chilling the magnets to -271.3°C – a temperature colder than outer space. For this reason, much of the accelerator is connected to a distribution system of liquid helium, which cools the magnets, as well as to other supply services.

Thousands of magnets of different varieties and sizes are used to direct the beams around the accelerator. These include 1232 dipole magnets 15 metres in length which bend the beams, and 392 quadrupole magnets, each 5–7 metres long, which focus the beams.

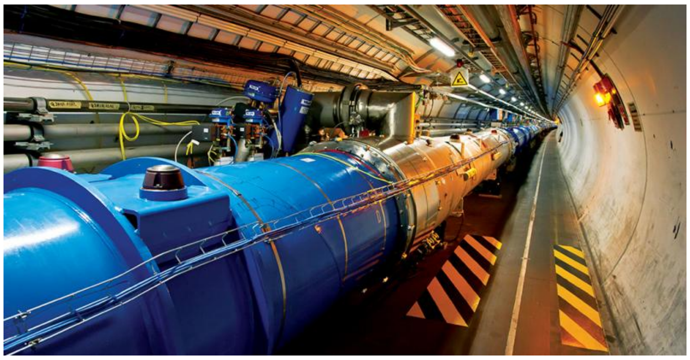

Figure 4. The Large Hadron Collider.

Just prior to collision, another type of magnet is used to ”squeeze” the particles closer together to increase the chances of collisions. The particles are so tiny that the task of making them collide is akin to firing two needles 10 kilometres apart with such precision that they meet halfway.

All the controls for the accelerator, its services and technical infrastructure are housed under one roof at the CERN Control Centre. From here, the beams inside the LHC are made to collide at four locations around the accelerator ring, corresponding to the positions of four particle detectors – ATLAS, CMS, ALICE and LHCb.

The LHC tunnel is located 100 metres underground, in the region between the Geneva International Airport and the nearby Jura mountains. The 27 km circumference circular tunnel was previously occupied by the Large Electron-Positron Collider (LEP) which was shut down in November 2000. CERN's existing PS/SPS accelerator complexes will be used to pre-accelerate protons which will then be injected into the LHC.

Seven experiments (CMS, ATLAS, LHCb, MoEDAL, TOTEM, LHC-forward and ALICE) will be performed on the collider; each of them will study particle collisions from a different aspect, and with different technologies. Construction for these experiments required an extraordinary engineering effort. For example, a special crane was rented from Belgium to lower pieces of the CMS detector into its underground cavern, since each piece weighed nearly 2,000 tons. The first of the approximately 5,000 magnets necessary for construction was lowered down a special shaft at 13:00 GMT on 7 March 2005.

The LHC has begun to generate vast quantities of data, which CERN streams to laboratories around the world for distributed processing (making use of a specialized grid infrastructure, the LHC Computing Grid). During April 2005, a trial successfully streamed 600 MB/s to seven different sites across the world.

The initial particle beams were injected into the LHC August 2008. The first attempt to circulate a beam through the entire LHC was at 8:28 GMT on 10 September 2008, but the system failed because of a faulty magnet connection, and it was stopped for repairs on 19 September 2008.

The LHC resumed operation on 20 November 2009 by successfully circulating two beams, each with an energy of 3.5 trillion electron volts. The challenge for the engineers was then to try to line up the two beams so that they smashed into each other. This is like "firing two needles across the Atlantic and getting them to hit each other" according to the LHC's main engineer Steve Myers, director for accelerators and technology at the Swiss laboratory.

At 1200 BST on 30 March 2010 the LHC successfully smashed two proton particle beams traveling with 3.5 TeV (trillion electron volts) of energy, resulting in a 7 TeV event. However, this was just the start what was needed for the expected discovery of the Higgs boson. When the 7 TeV experimental period ended, the LHC revved to 8 TeV (4 TeV acceleration in both directions) during March 2012, and soon began particle collisions at that rate. In early 2013 the LHC was deactivated for a two-year maintenance period, to strengthen the huge magnets inside the accelerator. Eventually it will attempt to create 14 TeV events. In July 2012, CERN scientists announced the discovery of a new sub-atomic particle that was possibly the much sought after Higgs boson believed to be essential for formation of the Universe. In March 2013, CERN announced that the measurements performed on the newly found particle allowed it to conclude that this is a Higgs boson.

On 5 April 2015 and after two years of maintenance and consolidation, the LHC restarted for a second run. Proton beams successfully circulated in the 27-kilometer ring in both directions. The first ramp to the record-breaking energy of 6.5 TeV was performed on 10 April 2015. The volume of data produced at the LHC presents a considerable processing challenge.

Particles collide at high energies inside CERN's detectors, creating new particles that decay in complex ways as they move through layers of subdetectors. The subdetectors register each particle's passage and microprocessors convert the particles' paths and energies into electrical signals, combining the information to create a digital summary of the "collision event". The raw data per event is around one million bytes (1 Mb), produced at a rate of about 600 million events per second.

The Worldwide LHC Computing Grid tackles this mountain of data in a two-stage process. First, it runs dedicated algorithms written by physicists to reduce the number of events and select those considered interesting. Analysis can focus on the most important data - that which could bring new physics measurements or discoveries.

3 The human part

CERN believes that scientific curiosity is an essential part of being human and motivates to share the excitement and enthusiasm of discovery with a wide a public as possible. CERN shows what people can achieve when they come together for a common purpose. The basic technology of the LHC was impossible when the project was conceived, after all today it runs routinely. Such grand challenges are just a normal part of life at CERN, and wants to bring all our expertise and share it with the community. [?].

3.1 Campus

During the 3 days the CERN has been introduced and LHC has been presented in detail, but it was really interesting to notice how a very complex physical and IT infrastructure has been break down to a simple and enthusiastic approach which lead to an increase of the scientific curiosity.
The volume of data produced at the LHC presents a considerable processing challenge. Particles collide at high energies inside CERN's detectors, creating new particles that decay in complex ways as they move through layers of sub-detectors. The sub-detectors register each particle's passage and microprocessors convert the particles' paths and energies into electrical signals, combining the information to create a digital summary of the "collision event". The raw data per event is around one million bytes (1 Mb), produced at a rate of about 600 million events per second.
The Worldwide LHC Computing Grid tackles this mountain of data in a two-stage process. First, it runs dedicated algorithms written by physicists to reduce the number of events and select those considered interesting. Analysis can focus on the most important data - that which could bring new physics measurements or discoveries.
In the first stage of the selection, the number of events is filtered from the 600 million or so per second picked up by detectors to 100,000 per second sent for digital reconstruction. In a second stage, more specialized algorithms further process the data, leaving only 100 or 200 events of interest per second. This raw data is recorded onto servers at the CERN Data Center at a rate around 1.5 CDs per second (approximately 1050 megabytes per second). Physicists belonging to worldwide collaborations work continuously to improve detector-calibration methods, and to refine processing algorithms to detect ever more interesting events.
In order to cover these requirements the campus presentation covered how to over comes using the latest advances on:

- Programming languages, mainly based on Java, PHP, and object oriented.
- Virtualisation environments and cloud, for both processing capacity and data storing.
- The importance of soft skills, not on how to right a CV and how to behave for a job interview, but also good practices at the work environment for a higher productivity.

3.1.1 Programming Languages

The four main component layers of the Worldwide LHC Computing Grid (WLCG) are physics software, middle-ware, hardware and networking.

Physics software
WLCG computer centers are made up of multi-petabyte storage systems and computing clusters with thousands of nodes connected by high-speed networks. They need software tools that go beyond what is commercially available to satisfy the changing demands of the high-energy-physics community.
The physics software on the Grid includes

programs such as ROOT, a set of object-oriented core libraries used by all the LHC experiments; POOL, a framework that provides storage for event data; and other software for modeling the generation, propagation and interactions of elementary particles. Grid projects supply much of the software that manages distribution and access, as well as job submission, user authentication and authorization. They also supply software known collectively as ”middle-ware”.

Middle-ware

Although the Grid depends on the computer and communications networks of the underlying internet, novel software allows users to access computers distributed across the network. This software is called ”middle-ware” because it sits between the operating systems of the computers and the physics-applications software that can solves a user’s particular problem.

The most important middle-ware stacks in the WLCG are the European Middle-ware Initiative (link is external), which combines key middle-ware providers ARC (link is external), gLite, UNICORE (link is external) and dCache (link is external); Globus Toolkit (link is external) developed by the Globus Alliance; OMII (link is external) from the Open Middleware Infrastructure Institute; and Virtual Data Toolkit (link is external).

Hardware

Each Grid center manages a large collection of computers and storage systems. Installing and regularly upgrading the necessary software manually is labor intensive, so management systems such as Quattor (developed at CERN) automate these services. They ensure that the correct software is installed from the operating system all the way to the experiment-specific physics libraries, and make this information available to the overall Grid scheduling system, which decides which centers are available to run a particular job.

Each of the 11 Tier 1 centers also maintains disk and tape servers, which need to be upgraded regularly. These centers use specialized storage tools – such as the dCache system developed at the Deutsches Elektronen Synchrotron (DESY) laboratory in Germany, the ENSTORE system at Fermilab in the US or the CERN advanced storage system (CASTOR) developed at CERN – to allow access to data for simulation and analysis independent of the medium (tape or disk) that the information is stored on.

Networking

The Grid file-transfer service (link is external), developed by the Enabling Grids for E-science projects, manages the exchange of information between WLCG centers. The file-transfer service has been tailored to support the special needs of grid computing, including authentication and confidentiality features, reliability and fault tolerance, and third-party and partial-file transfer.

Optical-fiber links working at 10 gigabits per second connect CERN to each of the Tier 1 centers around the world. This dedicated high-bandwidth network is called the LHC Optical Private Network (LHCOPN).

3.1.2 Virtualization and Cloud

The Worldwide LHC Computing Grid (WLCG) is a global collaboration of computer centers. It was launched in 2002 to provide a resource to store, distribute and analyses the 15 petabytes (15 million gigabytes) of data generated every year by the LHC.

In 1999, when work began on the design of a computing system for LHC data analysis, it rapidly became clear that the required computing power was far beyond the funding capacity available to CERN. On the other hand, most of the laboratories and universities collaborating on the LHC had access to national or regional computing facilities.

These were integrated into a single LHC computing service – the Grid – in 2002. It now links thousands of computers and storage systems in over 170 centers across 41 countries. These computer centers are arranged in ”Tiers”, and together serve a community of over 8000 physicists with near real-time access to LHC data. The Grid gives users the power to process, analyze and in some cases to store LHC data.

The WLCG is the world’s largest computing grid. It is based on two main grids – the

European Grid Infrastructure (link is external) in Europe, and Open Science Grid (link is external) in the US – but has many associated regional and national grids (such as TWGrid (link is external) in Taiwan and EU-IndiaGrid (link is external), which supports grid infrastructures across Europe and Asia).
This grid-based infrastructure is the most effective solution to the data-analysis challenge of the LHC, offering many advantages over a centralized system. Multiple copies of data can be kept at different sites, ensuring access for all scientists independent of geographical location; there is no single point of failure; computer centers in multiple time zones ease round-the-clock monitoring and the availability of expert support; and resources can be distributed across the world, for funding and sociological reasons.

Using the Grid
With more than 8000 LHC physicists across the four main experiments – ALICE, ATLAS, CMS and LHCb – actively accessing and analyzing data in near real-time, the computing system designed to handle the data has to be very flexible.
WLCG provides seamless access to computing resources which include data storage capacity, processing power, sensors, visualization tools and more. Users make job requests from one of the many entry points into the system. A job request can be almost anything – storage, processing capacity, or availability of analysis software, for example. The computing Grid establishes the identity of the user, checks their credentials, and searches for available sites that can provide the resources requested. Users do not have to worry about where the computing resources are coming from – they can tap into the Grid's computing power and access storage on demand.
Tier 0 of the Grid runs around one million jobs per day. Peak data-transfer rates of 10 gigabytes per second – the equivalent of two full DVDs of data per second – are not unusual.

3.1.3 Soft Skills

At CERN the project are executed using the Agile methodology [2]. Agile software development is a group of software development methods in which requirements and solutions evolve through collaboration between self-organizing, cross-functional teams. It promotes adaptive planning, evolutionary development, early delivery, continuous improvement, and encourages rapid and flexible response to change.
Scrum is the most popular way of introducing Agility due to its simplicity and flexibility. Because of this popularity, many organizations claim to be "doing Scrum" but aren't doing anything close to Scrum's actual definition. Scrum emphasizes empirical feedback, team self management, and striving to build properly tested product increments within short iterations. Doing Scrum as it's actually defined usually comes into conflict with existing habits at established non-Agile organizations.
Scrum has only three roles: Product Owner, Team, and Scrum Master. These are described in detail by the Scrum Training Series. The responsibilities of the traditional project manager role are split up among these three Scrum roles. Scrum has five meetings: Backlog Grooming (aka Backlog Refinement), Sprint Planning, Daily Scrum (aka 15-minute standup), the Sprint Review Meeting, and the Sprint Retrospective Meeting.

4 Conclusion

Based on the description provided is clear the CERN has provided an wonderful presentation in a very friendly environment using extensive scientific terms and thoughts [3].
The presenters are very competent on their areas but at the same time the message and the way to present very complex and scientific information was done using a very accessible speech and very enthusiastic leaving in our mind the seed for new developments.
I consider the presentations on the campus excellent for my personal development and

possibilities to develop and grow on new areas.

Acknowledgments

The author would like to thank his family which often is left for second place to commit all the time lines and schedules required. Specially the two little kids, Duarte 5 years and Alice 4 months, that have not all possibilities to push the patience to the limits...

References

[1] European Organization for Nuclear Research (CERN). (2015) `http//www.cern.ch`.
[2] Agile Methodology. (2015) `http://agilemethodology.org/`.
[3] N. Pinto, *2015 CERN spring campus*, Instituto Superior Técnico, June 2015.

Nuno Ricardo Pinto Degree in Electrical and Computer Engineering since 2001 and currently Master's student in Information Systems and Computer Engineering at c IST in Lisbon.
Subject Matter Expert in Telecommunications intelligent network services, including enterprise solutions with a strong representation internationally

Appendix
Statements of Execution

Statements of Execution for the Activity Performed.

- CERN -
SPRING
CAMPUS

SECOND EDITION
1st – 3rd April 2015
PORTUGAL

CERN SPRING CAMPUS DIPLOMA

This certifies that

Nuno PINTO

has completed the course of study during the 2015 CERN Spring Campus.
The program consists of 25 hours of lectures held over 3 days.

CERN

TÉCNICO LISBOA

The 2015 CERN Spring Campus was jointly organized by the European Organization for Nuclear Research (CERN), Geneva, Switzerland and The Instituto Superior Técnico, Lisbon, Portugal.

Derek Mathieson

Derek Mathieson
Director, CERN Spring Campus
8th April 2015