The discovery of the Higgs boson marked an important step towards understanding the fundamental laws which underpin the universe. This elusive particle, thought to be at the very root of the existence of mass, has been the focus of experiments carried out at the Large Hadron Collider (LHC) at CERN on the French-Swiss border.

Challenges

• Crunch numbers. Process, filter and analyze petabytes of complex LHC data every year
• Needles in haystacks. Empower scientists to identify the 20 or 30 data points they need to find in millions – using advanced algorithms

Solutions

• High energy physics computing. The 9,000 servers at the heart of CERN’s WLCG are powered by roughly 90,000 cores, the majority of which are powered by Intel® technology, with more than 12,000 belonging to the Intel® Xeon® processor E5 family
• Well connected. As part of the equipment used for the interconnects, the CERN part of the computing grid uses Intel® 10 Gigabit Ethernet Network Interface Cards 82599EB to connect disparate servers and support a variety of specialized experiments

Impact

• World-class research. Scientists use the computing platform to explore mysteries surrounding the Higgs boson and the Big Bang
• More to come. Ambitious development plans for the LHC will continue to need cutting edge processing technology

Cracking nature’s code

The scientists at CERN carry out high-energy particle physics research, studying the interactions between particles accelerated around the 27km-long LHC ring and analyzing the immense quantity of data captured by sophisticated sensors. These sensors collect an extremely large volume of information, which needs to be processed by CERN’s compute infrastructure at incredibly fast rates. After massive capturing and filtering efforts, CERN stores more than 25 petabytes (or two million double-layer DVDs-worth) of data from the LHC and its four main experimental facilities each year. However, even this volume of information is small compared to the total amount of insignificant data that the physicists first must identify, using massive processing power, and then discard. The decision of what to keep and what to discard must be made in fractions of a second.

The critical role of algorithms

As two proton beams speed around the LHC, they collide with each other, creating a split second of chaos in which traces of the Higgs boson can be identified. Catching a glimpse of these fleeting interactions requires highly complex calculations, and the physicists must have the right information at their fingertips.

The first step in making these calculations is capturing the proton-proton interaction data through the sensors of LHC experiments. Specialized algorithms filter this data in real time, reducing the number of events from 40 million per second to around 100,000. However, this still leaves around 1.6 gigabytes of data per second to be reconstructed and analyzed before it is fed into WLCG, a network of over 160 data centers worldwide that carries out further refined analysis with improved calibration using the grid’s vast capacity of distributed computing resources.
Interconnection and scalability are key components of the WLCG project, enabling the collaboration of supercomputers and data centers around the globe. The Gemini Data Center in Geneva, which houses the majority of the WLCG computational power, is equipped with 9,000 servers, mostly powered by Intel Xeon family processors. This setup supports the handling of over 40 million proton beam crossings per second, with only 300 to 400 interactions deserving the attention of ATLAS physicists.

Planning ahead, CERN and Intel are implementing the latest technologies and infrastructure. Intel’s support through CERN openlab, the research and development framework created in 2001, is crucial for the success of the WLCG project. The provision of high-performance, multi-core processors and a scalable interconnectivity infrastructure is vital for the future of the LHC experiments. Intel’s technology provides the foundation for the continued advancement of particle physics research at CERN.
Lessons learned

In the world of scientific breakthroughs, there’s no time to stand still. As data gets bigger and calculations get more complex, it’s essential to have the most powerful and reliable computing resources available to keep up. CERN has created a dedicated team in CERN openlab to plan ahead, making sure the technologies deployed will meet ever-increasing demands and drive further groundbreaking discoveries like that of the Higgs boson.

The computer that changes science

The advanced computing resources that CERN openlab and Intel have made possible are driving a wide range of experiments and scientific progress beyond the search for the Higgs boson.

Frans Meijers coordinates the complex task of capturing data produced by the CMS experiment, which together with Atlas discovered the Higgs boson, determined its mass, and analyzes its subsequent decay into other particles. “In capturing our data we apply two rounds of selection. One round uses our own custom-built electronics components, and the other uses software algorithms powered by commercially available processors, such as the Intel Xeon processor E5 family,” he explains. Now that physicists have obtained the first approximate value of the mass of the Higgs boson, the future increase in the accelerator’s luminosity will be key in tackling the challenging task of determining the properties of the Higgs boson more precisely and doing searches of new particles at the higher energy after the LHC machine upgrade. The new Intel® Xeon Phi™ coprocessors (Intel® Many Integrated Core Architecture) could play a fundamental role here,” he adds. “I’m eager to put them to the test.”

For Andrzej Nowak, a researcher working alongside Jarpl in CERN openlab, the relationship between CERN and Intel has a unique impact on his research. “We have the opportunity to access Intel’s exclusive know-how and, in turn, Intel can achieve a deeper and more direct relationship with an environment that is very demanding from a technological and computing standpoint, where innovative ideas circulate,” he says. One of the main lines of Intel research that CERN openlab is exploring is the integration of silicon and photonics. This new technology promises to increase the applicability of optical systems in chips, ultimately providing a qualitative jump in the transfer of data – like that generated by the sensors of LHC experiments – to the digital systems that have to filter and analyze it. In 2010, Intel introduced a photonic component for connections of 50 gigabits per second and now it is on the brink of reaching transfer speeds measured in terabits. CERN’s interest in this technology has also inspired it to collaborate with Intel in offering advanced training to young scientists through programs such as the Intel-CERN European Doctorate Industrial Program (ICE-DIP). This is integrated with FP7, the Seventh Framework Programme for Research in Europe, established by the European Commission.

ICE-DIP trainees will focus on the techniques needed for acquiring and processing many terabits per second, using and expanding the most innovative concepts available in the information and communications industry today. The program will explore new, untested ideas such as silicon photonics for network links in harsh operational conditions, tight integration of reconfigurable logic with standard computing, and new approaches to data acquisition.

The mysteries of the universe tackled using 260,000 cores

Ian Bird, who works with CERN’s IT Department as a project leader for the WLCG, says he’s convinced of the need to work even more closely with CERN openlab and Intel in the future to optimize the code that runs on the 260,000 combined cores of this immense distributed architecture. “A solid relationship with an industry leader like Intel is important,” stresses Bird, “not only so we can make the most of the resources currently available to us, but also in terms of future development.”

In agreement with Bird is Niko Neufeld, who works on the online team for the Large Hadron Collider beauty (LHCb) experiment, which is investigating the mystery of antimatter. He explains: “The Big Bang Theory hypothesizes that when the universe
was formed, the quantity of matter created should have been equal to that of antimatter. However, these two incompatible states did not cancel each other out, as the creation of the universe goes to prove. So where did the antimatter go? To answer this question, the LHCb researchers require computational efforts ranging across a vast number of compute nodes able to work simultaneously. For this reason, Neufeld’s team of researchers works extensively on software parallelization, studying the opportunities provided by Intel Xeon Phi coprocessor devices and their Infiniband® connectivity for the interconnection of the compute nodes. They see these technologies as essential support for the research project’s most complex calculations.

Like Lewis Carroll’s famous character, the A Large Ion Collider Experiment (ALICE) passes through the looking glass of visible matter to get back to conditions of density and temperature very similar to those of the Big Bang. Where the other experiments concentrate on observing the interactions among individual protons, ALICE uses lead nuclei, which are much heavier. These nuclei collide and form a plasma, a gas made up of quarks and gluons, the particles that make up protons and neutrons. “Observing this plasma in the ALICE detector,” explains Pierre Vande Vyvre, ALICE project leader for data acquisition, “generates an enormous volume of data to analyze – up to 16 gigabytes a second.” This is a significant computational challenge which will be even more grueling after the accelerator’s second upgrade scheduled for 2018, requiring up to 100 times more events to be captured in real time. The ALICE researchers worked with the CERN openlab team on software optimization to maximize the potential benefits coming from code parallelization and vectorization technology to be tested with the new Intel Xeon Phi coprocessor. In these investigations into the origin of the universe, computational power is crucial, but the role of the scientists – as indispensable mediators between the theoretical aspects and the analysis of experimental events and mathematical models – is, of course, paramount.

“CERN is at the forefront of IT industry with its thought leadership in using the latest computing technologies to achieve scientific breakthroughs,” says Steve Pawlowski, Intel senior fellow, Datacenter and Connected Systems Group, chief technology officer and general manager, Pathfinding. “Its need for more performance and increasingly efficient high-performance computers reflects the hunger to reach further in solving some of the world’s biggest mysteries like the existence of the Higgs boson. CERN’s insatiable hunger for more compute capacity took every advantage of Moore’s Law and new generations of processing technology that Intel has created. Going forward, as we see some significant changes in how we architect the next generation of high-performance computer systems, the relationship with CERN openlab will be even more important for us in making sure the technology we develop delivers as expected.”

By contributing both technology and expertise to support the team at CERN, Intel is continuing to pursue its goals of overcoming the physical and technological limits of supercomputing.

“CERN is at the forefront of IT industry with its thought leadership in using the latest computing technologies to achieve scientific breakthroughs.”

Steve Pawlowski, Intel Senior Fellow, Datacenter and Connected Systems Group, Chief Technology Officer and General Manager, Pathfinding

Find the solution that’s right for your organization. Contact your Intel representative, visit Intel’s Business Success Stories for IT Managers (www.intel.co.uk/ftcasesstudies) or explore the Intel.co.uk IT Center (www.intel.co.uk/itcenter).