2.10 Grid Computing
Physics Data Processing and ICT infrastructure

Management: dr. J. Templon (PL)

The Physics Data Processing (PDP) group, in close collaboration with the Computer Technology (CT) department, in 2014 implemented the new strategic plan focussing on scalable computing R&D and scalable multi-domain security. Important advances were made related to our data processing facilities, both in capacity ramp-up for LHC Run 2 and in broadening the user base. These facilities provide essential computing and storage resources for producing scientific results, and are also a key ingredient in validating our research and engineering work on scalable computing and multi-domain security.

Scalable infrastructure
The expected data rates and the reprocessing volume foreseen for 2015 and beyond will stress our network as well as our compute and data infrastructure. With increased CPU core counts, getting the data delivered to the systems sufficiently quickly has become the major challenge. Already the interconnect between storage and compute systems within the Nikhef site exceeds 200 Gbps. In 2014 we deployed storage arrays featuring a 40 Gbps network interface per 100 Terabyte of storage and replaced the core network equipment, co-sponsored by SURFnet and in close collaboration with the neighbouring institutes AMOLF and CWI. The new network core permitted an upgrade of the cross-connect within the NL-Tier1 (between Nikhef and SURFsara) to 100 Gbps. It also allowed Nikhef to act as a destination and transit provider for the new ATLAS Tier-1 center at the Kurchatov Institute in Moscow. Limited analysis runs by ATLAS in December 2014 demonstrated that this increased capacity can be effectively used, and that it will be needed given the expected network rates once the LHC data taking restarts in 2015. The fraction of the NL-Tier1 facility housed at Nikhef now comprises 3100 CPU cores, and 2 of the 5 Petabyte of NL-T1 disk dedicated to wLCG. Fig.1 shows the division of Nikhef CPU resources over the different projects.

Such advances in computing, data, and network performance are enabled through the national e-Infrastructure, coordinated by SURFsara and funded largely through SURF. However, that funding is insufficient to provide the capacity needed to deal with energy, luminosity, and detector upgrades of the LHC in 2015 and beyond. For the next 5 years, funding to increase capacity of the NL-Tier1 has been obtained in the NWO BIG LHC Upgrade project that was awarded in 2014 —giving us a solid basis to exploit the results of Run 2.

Despite the continuous replacement of hardware and replacement of the core routing equipment, availability and reliability the Nikhef and the wLCG NL-Tier1 remain above target at 98.5% over 2014.

Scaling the computing for Nikhef physics analyses
Also the 'stoomboot' analysis facility, used by Nikhef physicists for daily analysis tasks, has been significantly increased in size. Whereas compute capacity is now in ample supply, the way data is accessed during the analysis phase puts severe strain on storage systems —traditional data access patterns no longer suffice, leading the PDP group in collaboration with the users to experiment with various alternative file systems and the underlying disk configuration. More than ever we observed a complex interplay between low-level systems configuration at the per-disk level, the file systems used, and the way data is distributed across storage clusters. Performance analysis and subsequent tuning of the 'glusterfs' filesystem alleviated some of the pressure, but in order to fully exploit the enlarged 'stoomboot' analysis facility we foresee a need to shift more of data access logic away from traditional file systems.

But there is much more to computing than just capacity —we also work with other Nikhef groups and with CPU and network vendors including Intel, Juniper, and Aruba to stress new hardware in an environment that is markedly different from the commodity market— often identifying potential issues before they make it into products, and thus enhancing the usefulness of future products for our user communities.

Significant challenges remain in applying acceleration techniques like massive multi-core computing to actual physics codes —partly due to the data sizes and throughput requirements involved. While we work on adapting such cores to use massive multi-processing, such as on the Intel Xeon Phi, and while also CPU vendors are addressing the issue of data throughput from general-purpose cores to co-processing systems, we foresee this as a major area of activity for the PDP and CT groups in the years to come.

Beyond a single institute
Whilst having computing capacity at a single location is nice, all our research is based on collaboration. Being able to transgress (electronically) institutional, organisational, and political boundaries is therefore essential, and the Scalable Multi-Domain Security (SMDS) activity of the PDP group addresses the complex authentication and authorisation issues of global collaboration. This area is strongly driven by research beyond traditional high-energy physics, has gained much European and global interest, and now appears to have gathered enough

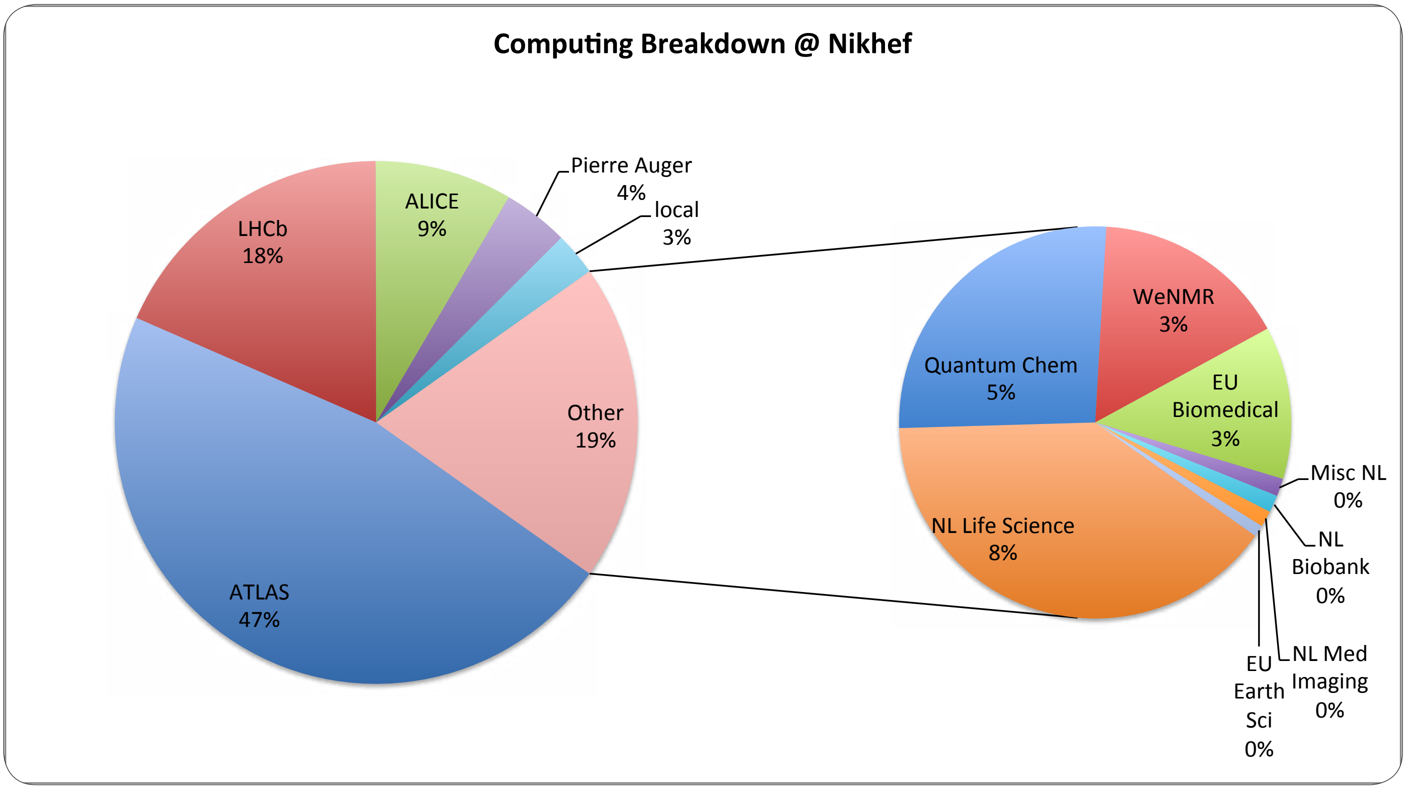

Figure 1. Division of Nikhef CPU resources over different projects in- and outside Nikhef.

momentum to make a major breakthrough in the next couple of years. For the last decade the LHC community has managed to deal with relatively complex technology such as user-controlled public key infrastructure because of the then-still-niche need to use command-line applications and non-web technology. Most other research domains have been able to work largely in a web-only environment and did not have to deal with large data volumes. This is now changing rapidly, an all major ESFRI infrastructures have aligned their requirement for global federation and the use of non-web single sign-on in the 'Federated Identity Management for Research' (FIM4R) activity, that formed also the basis for the FIM activities in the Research Data Alliance RDA. In 2014 the Nikhef SMDS team, with its background in policy coordination and software engineering for community-based site access control, joined four European project initiatives to further the role of federated identity management specifically for cross-national collaborations like those in sub-atomic physics, and for the non-web technologies that are critical for data-intensive research. Of these, the European Grid Infrastructure (EGI) ENGAGE initiative and the Authentication and Authorization for Research Collaborations (AARC) initiative, coordinated by the Geant Association (TERENA), have been targeted at dedicated calls in the EU Horizon 2020 programme. Nikhef is a leading partner for (authentication) policy development and involved in software engineering.

Meanwhile, our site access control middleware continues to see world-wide adoption, with increasing deployment in the US as part of Open Science Grid supporting the US LHC community. Our role in authentication coordination and the Interoperable Global Trust Federation warrants a seat for Nikhef on the Geant Association's TERENA Technical Committee, overseeing its Technical Programme, and engaged in the Phase-I testing of the pan-European TCS Certificate Service.