1 Grid & Large-Scale Computing

Management: dr. J.A. Templon (PL)
ing. W. Heubers

For the grid facilities at Nikhef, 2013 was a year of business mostly as usual, the notable exception being a marked increase in the amount of non-HEP computing work performed on our grid cluster. Given the close of the BiG Grid project as well as the european middleware projects (EMI and IGE), which had funded much of our development work, 2013 gave our group an opportunity to refine the focus of our research activities.

Facilities

The grid facilities at Nikhef continued to run smoothly (99% availability) this year. Usage was also high, with only 2.4% of the computing power going unused. Fig. 1 shows how the computing power of the Nikhef grid cluster was used in 2013. In May, new capacity was installed, a large fraction of which replaced out-of-support hardware purchased in 2009; the rest (purchased via a subsidy from the national e-infrastructure operated by SURF) was used to expand capacity available to non-HEP experiments, whose usage of the facility expanded to 13% in 2013. A further expansion is expected soon, as LOFAR data processing on the grid begins in earnest early next year.

The grid group took increasing responsibility for the "stoomboot" cluster, used mostly by Nikhef physicists for data analysis tasks. Main activities here were improving the scheduling, taking into account the wide range of activities for which stoomboot is used, and in improving access to data via both the network (mostly to grid storage at Nikhef and SARA) and via the `glusterfs` distributed file system.

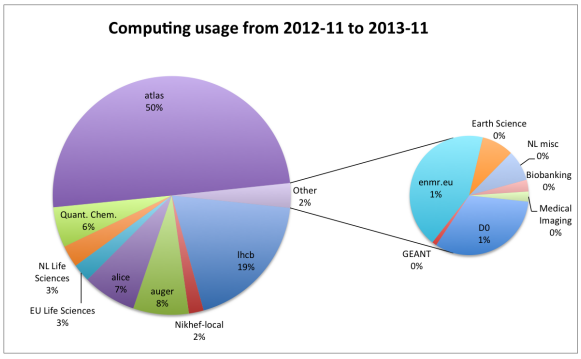

Figure 1.1: *Usage of computing resources on the Nikhef grid cluster in 2013.*

Activities in support of Users

Activities regarding the Nikhef stoomboot cluster have already been mentioned above. Most of the effort went into understanding and correcting stability problems with the `glusterfs` distributed file system, via which much of the data was being analyzed. Work on improving the *performance* is still in progress.

Work on *multiprocessing* (for lack of a better term) continues to increase. Pilot studies on GPU use were undertaken in the context of the KM3Net and ATLAS experiments. Exploratory work on more traditional multiprocessing (parallelization across multiple cores and machines) was performed in the context of solid-state physics research undertaken in support of new-detector development, understanding how to optimize deployment and production for computing work using the VASP electronic structure code.

The first jobs from the Virgo (gravitational waves) group have been sighted on the Nikhef grid cluster, following on from discussions last year. Virgo intends to run their general-relativity simulation codes on the grid. Collaboration with the Nikhef XENON group on data management continues.

CHEP 2013

Amsterdam was host to the conference on Computing in High Energy Physics (CHEP), organized by our group (conference chair D. Groep). The conference was held in the Beurs van Berlage and had as its main theme, computing in the Multicore and Big-data era.

The conference was well attended (479 participants) and extremely well received by the attendees, despite the fact that many colleagues from US national laboraties (such as FNAL) were unable to attend due to the US government shutdown.

Computer Science Research

Daniela Remenska's PhD project (our group collaborating with prof. dr. Henri Bal at the Free University of Amsterdam) concerns using formal methods to verify soundness of design in distributed systems. Her studies are carried out on LHCb's DIRAC system, on which their computing operations is based. Currently her research focuses on forging a bidirectional link between the UML diagrams used by software engineers to specify software design, and the formal-method model representation of the software, in which verification can

Figure 1.2: *The CHEP 2013 conference in the Beurs van Berlage in Amsterdam.*

be carried out. Such a link will allow systems to be verified during the design stage, without the burden of constructing yet another representation of the program (in the formal-methods language).

National e-Infrastructure

This year, the SURF foundation assumed responsibility for the dutch national e-science infrastructure (formerly operated by the BiG Grid project co-directed by Nikhef). One of the most visible changes was a major overhaul of the procedures for submission and review of requests for computing resources on the infrastructure. The resulting simplifications resulted in a marked increase in the number of active users on the infrastructure; as discussed above this is immediately apparent from the usage statistics of our own grid cluster, which is operated as part of the dutch national infrastructure.

Nikhef is also involved in the NetherLands e-Science Center (NLESC), another element of the national e-infrastructure. The NLeSC consists mainly of a team of e-Science engineers who work on a wide variety of projects across many disciplines, the common feature being the enabling of new research via use of software, computing, storage, and network infrastructures. Nikhef contributes a member of the e-Science "Integrator" team, a group of senior scientists deeply involved in computing for diverse fields such as High Energy Physics, Astronomy, Cognition, Genetics, and Water Management. This group is sufficiently broad *and* deep to significantly influence the national scientific computing agenda.

Middleware

Our middleware products are in active use by grid projects around the world. Active development on these products is decreasing, due to both the level of maturity reached by the software, as well as the end of the european middleware projects that have funded these developments for the last ten years. Nikhef will continue to maintain these products as long as "relevant customers" (for example, grid infrastructures for physics research) continue to use them.

The national e-infrastructure funds development work at Nikhef as well, mainly in the area of *scalable multi-domain security*. One can think of this as an academic equivalent to "sign on using your Google credentials" — in our case, it's "sign on using your institutional credentials", the idea being that when you are employed as a scientist at a research institute in the Netherlands, this status should grant you immediate access to a variety of electronic resources, without having to first "register" and "create an account". Nikhef is active in this area both as a policy maker (how secure do institutional credentials need to be before you can trust them?) and as technology producer, constructing software enabling cross-institute collaborations based on the technology described above. Via funding provided by the European Grid Infrastructure (EGI), Nikhef also assumes primary responsibility for these policy tasks at the European level, along with major responsibility for coordination of the cyber security of the EGI.

Other Activities

Given sufficient funding, we plan to engage with the KM3Net and LHCb groups at Nikhef surrounding optimization of their trigger software, the expectation being in both cases that there is much to gain from restructuring of the code and/or pursuing a move to newer hardware. There are also plans for a significant expansion of the stoomboot cluster, both in the sense of more job slots for the current activities, as well as a segment dedicated to parallel workflows such as those being run by the Nikhef Detector R&D group. Data management work will continue with the XENON group, who have also indicated interest in our GPU research. Finally we are closely following the (inter)national developments aimed at the Horizon2020 EU program, with the aim of making significant contributions, for example in the area of Identity Federations, an area both of Nikhef expertise and specific EU focus, given that most ESFRI projects are requesting this technology.