Computing conference in India*

Computing in High Energy and Nuclear Physics (CHEP) is a major series of conferences held at roughly 18 months interval since 1985, alternating between Europe, North America and other parts of the world. The latest meeting, CHEP06, the 15th in the series, was held recently in India. This conference was preceded by a three-day workshop on Service Challenge 4 from 10 to 12 February and a two-day school on Grid Computing on 11 and 12 February. The main theme of the conference as well as the preceding workshop was to review progress in making the Grid a powerful and reliable computing resource in time for processing LHC data. The conference also aimed to learn from the experience of experiments that are currently running, to stay in touch with GRID applications in other sciences and to have a look into the future.

This was a well attended conference with 478 delegates from all around the world. The Programme Committee considered 432 abstracts of which 32 were scheduled as plenary talks, 253 for oral presentations in eight parallel sessions and 144 as poster presentations.

The conference opened with a note that this may be the last one in this series before LHC start-up. Consequently the emphasis of readiness of the LHC machine, experiments and computing services were of major interest. In the opening talk, Joseph Engelen (CSO CERN), talked about the status of the machine and the major experiments. He confirmed that the LHC project – the machine, detectors and LCG – is well underway for physics in 2007. Schedules are tight but not impossible to meet. There is a large potential for exciting physics.

Jamie Shiers (CERN) elaborated on the readiness of the LHC computing facilities. There has been excellent progress on the three key areas addressed by the Worldwide LHC Computing Grid project, namely data-transfer tests, service availability and time to resolve problems and provisioning of resources. There is a clear plan for the remaining work needed to meet the computing challenges of the LHC and success can be achieved by a pragmatic, focused and cooperative approach.

Paris Sphicas (CERN) focused on the start-up of LHC data-taking and the state of readiness of the software of the LHC experiments. Most of the common software is in place, while much of the experiment-dependent software is either complete or has a fully functional prototype. Important tests like calibration and Grid utilization are under way. Deployment has begun in earnest and performance studies are in progress. There is still a long way to get some of the complicated analyses carried out, but all the experiments are getting closer to that goal.

Martin Purschke (BNL) talked about the experiences of the Phenix experiment at RHIC, which has LHC-era data rates. The experiment, originally designed for a data rate of 20 MB/s, eventually operated at a rate 60 times higher. Data compression and buffering of data have been found to be key ingredients in achieving this goal.

Beat Jost (CERN) illustrated design criteria and resulting architecture of future DAQ systems. There is a move toward elimination of hardware triggers. For more flexible triggering and optimal efficiency, there will be readout at bunch-crossing or train-crossing and the data will be sent out to CPU farms for event selection.

Elizabeth Sexton-Kennedy (Fermilab) talked about social and technical challenges of event processing framework. She concluded that frameworks should be judged by how well they meet the technical challenges of complexity and scalability. In addition, the framework provides a set of tools that organize a large group of developers into moving in the same direction.

Rene Brun (CERN) presented the success story of the ROOT project and its adaptation in the multi-core CPU era. Instead of pushing gigabytes of source or shared libraries to Grid working nodes, he proposed to use a Pull technique to download only the software necessary to run an application and in an incremental way. While the core ROOT software is being consolidated to be ready for LHC data-taking, prototyping work is underway for the new concept.

Peter Elmer (Princeton University) discussed distributed data management issues in high energy physics experiments. Gaining experience from the ongoing experiments, a lot of effort is going in understanding the data access pattern in the next generation of LHC experiments. He foresees several exciting data management-related work in the next 3–5 years.

Harvey Newman (CalTech) mentioned that networks used by high energy physics and other fields of data-intensive science are advancing rapidly. These fields are also learning to use long-range networks more effectively. Hybrid dark fibres are emerging as the means to rapid progress in communication speed in many countries. He emphasized that it is now important to close the digital divide to allow scientists in all world regions to take part in discoveries.

Les Robertson (CERN) talked about the LCG service focusing on where the expectations were fulfilled and where more efforts are needed to be ready for the first LHC beams. There are two Grid infrastructures, LCG and OSG, now in operation on which one would be able to complete the computing services for LHC. The main priority is to set up a reliable operation of the baseline services.

Ruth Pordes (Open Science Grid) emphasized on the interoperability aspects of different Grids in view of several Grid infrastructures all around the world today. There are many ongoing efforts towards this. Kenichi Miura (National Institute of Informatics, Japan) and Gang Chen (IHEP, Beijing) elaborated on Grid activities in Japan and China. In the National Research Grid Initiative (NAREGI) project in Japan, the primary objective is a seamless federation of heterogeneous resources. It started with computations in nanoscience and technology over Grid and expanded its scope to high energy physics, astronomy and biology. The Grid activities in China encompass a variety of disciplines like high energy physics, bioinformatics, environment science, materials science, computational chemistry, etc. The experience from LCG is driving other Grid applications in China. For Grid activities in the Asia-Pacific federation; applications, as pointed out by Simon Lin (Academica Sinica Taiwan), span

*A report on the 15th CHEP organized at the Tata Institute of Fundamental Research, Mumbai from 13 to 17 February 2006.

from high energy physics to atmospheric science to nano- and bio-applications.

Ashok Jhunjhunwala (IIT Madras) talked about connectivity issues in India. Internet technology will have a huge impact on the lives of rural communities in developing countries, if there is a big enough vision. In India, one foresees 50 million broadband connections by 2010 and this will double per capita GDP in rural India. One needs lots of innovations to meet this effort and in particular, uninterrupted power supply will be the bottleneck for this.

Piergiorgio Cerello (INFN) talked about role of Grid in medical applications. Rajiv Gavai (TIFR) addressed computing challenges in Lattice QCD. Mathai Joseph (Tata Research Development and Design Centre) elaborated on the software development process and computing challenges for astronomy experiments like the Square Kilometer Array.

There were several exciting presentations from IT industries from all over the world. Tony Hey (Microsoft Corporation) talked about the capability to access, move, manipulate and mine data in collaborative science applications. He emphasized the continuing trend towards decentralized, networked resources and usage of internet and open access contributions for promoting a global scientific knowledge base. Alan Gara (IBM T. J. Watson Research Centre) talked about the role of high performance computing in varieties of physics applications and scalability issues in supercomputing. He emphasized that architectural innovation is critical to continue performance scaling. David Axmark (MySQL AB) gave an insight into the making of a scalable database system. He was in favour of free software quoting software patents as a threat to software innovation. Giving access to the source code, the database system becomes more robust as well as secure. Lalitesh Kathragadda (Google India) talked about Google's approach to organize the entire world's information and to make it universally accessible and useful. He emphasized on their in-house solution for data storage (Google File System), for running jobs on pools of machines (Global Work Queue), and for simplification of large-scale data processing (MapReduce). Anirban Chakrabarti (Infosys Technologies) elaborated on various themes of Grid applications starting from middleware to specific applications where Infosys is focusing its current research activities.

A key attraction of the conference was the visit of the President of India, A. P. J. Abdul Kalam, who gave the Valedictory address of the conference. He summarized the missions for CHEP06: computing in particle physics, mission for space research and particle research, and mission for energy. He encouraged CERN–DAE collaboration on Grid activities and emphasized on enlarging the Grid activities to Knowledge Grid, E-Governance Grid, Health Care Grid and PURA Grid connecting one billion people all over India.

Another key attraction of CHEP06 was setting up a 622 Mbps link to USA through Japan, over which several application jobs were demonstrated. This was achieved through a collaboration of CalTech, Internet-2, World Bank and CDAC-India. This effort drew attention of the funding authorities of India, who could appreciate the role of high bandwidth network for research and education.

The final session of CHEP06 was a panel discussion on digital divide issue where the panelists addressed the importance of bridging the gap among different countries in the world and also how this has been addressed in Russia, Brazil, etc. The panellists, S. Ramakrishnan, Harvey Newman, Viatcheslav Ilin, Alberto Santoro and A. S. Kolalskar tried to formulate possible ways of addressing this problem for nations where low bandwidth connectivity still prevails.

The conference concluded with the coconveners Sunanda Banerjee and Atul Gurtu thanking all the participants and all those who worked hard to make the conference successful. The next conference in this series will be held in Victoria, Canada during September 2007.

The pre-conference workshop on Service Challenges, organized by Jamie Shiers, was well attended (155 participants), and devoted to the primary activities of the WLCG (Worldwide LHC Computing Grid) Service Challenge programme in the run-up to the full WLCG production service that will commence in October 2006. The topics concerned were data and storage management, WLCG services in general and finally computing models of the experiments and their plans for the validation of their offline frameworks using the WLCG services during the remainder of 2006. The workshop was highly interactive, with many profitable discussions and breakout sessions. The first day of the workshop was spent on storage management focusing on the question of interoperability between different Storage Resource Manager implementations as well as the deployment issues. Key issues that were discussed during the second day were the service-level targets defined in the WLCG Memorandum of Understanding, how they could be met and how they would be measured. The third day was spent discussing impact of the computing models of different LHC experiments on the service requirements at a multi-disciplinary Tier1 or Tier2 site.

The pre-conference tutorial programme was organized by Francois Fluckiger in two days with six lecture series on fundamentals of Grid technology, cryptography and security protocols, fundamentals of networking QoS, security from theory to implementation, pragmatic software engineering and databases. 55 participants took part in the school and the experience has been found to be extremely positive.

The conference was followed by a one-day workshop organized jointly by TIFR and CDAC, focusing on ways of improving the connectivity issues in India. This workshop was attended by people from Knowledge Commission in India and the Government ministries on information technology. A memorandum of understanding was signed between Internet-2, CDAC and ERNET for promoting high bandwidth connectivity in India.

Sunanda Banerjee, Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400 005, India
e-mail: sunand@tifr.res.in