The Grid: Computing without bound

Grid computing is the act of sharing tasks over multiple computers. The tasks can range from data storage to complex calculations and can be spread over large geographical distances.

If you think of ten men pulling on a rope to lift a load, that is parallel computing. If ten men have ten ropes and are lifting ten different loads from one place to consolidate at another place, that would be distributed computing.

Towards the end of the last century, the discovery of World-Wide-Web allowed several hundred scientists in Particle Physics experiments to have access to scientific data. This technology has revolutionised very fast the telecommunications and networking applications. It has changed our lifestyle!

The Grid is the information highway which is changing the style of computing in all branches of science, be it High Energy Physics, Biology or Meteorology. It allows immensely complex and huge data volume to be analysed in faster and more efficient way that not a single or a handful of individuals can provide. The whole community pools in their resources together and, necessarily, acts in a coordinated way.

The GRID technology allows one to share computing power and resources in terms of disk storage, databases and software applications among centres spread all over the world. Also, importantly, a Grid facility is built with commercial hardware and open-source software providing inexpensive access to high end computational capabilities on demand.

When you plug in a refrigerator in your house, you do not bother about where from the power comes, similarly a scientist using the grid facility does not have to worry to gather and organise the computational resources he needs for his work.

The Computing GRID is developed in the context of Large Hadron Collider (LHC) experiments, which are about to start collecting real data. The high energy hadron collisions at LHC facility will be studied by several experiments which will record and archive data at an unprecedented scale to resolve fundamental mysteries in subnuclear matter. To pinpoint the smallest fragments of the universe we have to build the biggest machine in the world and to recreate the first millionths of a second after the creation of the universe, we have to focus energy at an awesome scale. The complexity of the experiments result in data volume which cannot be stored and accessed with usual technology of today. The LHC Computing Project is meant for building and maintaining a data storage and analysis infrastructure for the entire high energy physics community that will use the LHC for more than a decade.

Indian scientists are taking part in various aspects of LHC from its early days of inception. TIFR is host to a Tier2 Grid computing facility for the CMS experiment at LHC and provides a reasonable fraction of the required computing resource for about 3000 scientists-strong CMS collaboration. In particular this Tier2 facility is meant to serve the Indian community in the experiment to reap the harvest of LHC experiment in par with the scientists world-wide. Stay tuned!

Text prepared for TIFR Open Day, 15.11.2009