GRID-technology and ALMA

One of the tasks of the ALMA Regional Centres (ARCs) will be to offer the ALMA-users support for the elaboration of the data. This forces us to think about the computing resources that have to be set up at the ARCs, and how to make the best possible use of existing resources.

ALMA produces data cubes with dimensions in the range from 50 Gbyte to 1 Terabyte. From this fact alone it is clear that for jobs of average complexity the typical desktop computers will not be sufficient to elaborate the data and that it will be necessary to set up specialized systems and/or to use the available Grid nodes. In fact, if we look at the evolution of computers over the past years we see that the CPU have by now reached their maximum frequency and the manufacturers have not succeeded to exceed the barrier of 4 GHz. Instead, manifacturers have become oriented towards internal parallelization, such as dual Core, quad Core systems, that do not, however, give the same performance as one CPU with double or quadruple speed. On the other hand, the increase in storage capacity of discs does not seem to slow down and we may expect relatively economic Raid – NAS* systems with a capacity of tens or hundreds of Terabyte to be available on the market soon. Furthermore the speed of the computer buses is increasing, which allows network connections of 10 Gbit/sec.

A resource that we intend to implement for the ARC therefore consists of a classic cluster with several computers connected in a high speed network that can use Raid-NAS systems with a capacity of tens of Terabyte. While at present these systems are used extensively for theoretical calculations and in nuclear physics, they are not often used for typical astronomical and radioastronomical data reduction applications, also due to the limitations of existing Network File Systems. The new generation of parallel network file systems (Lustre, GPFS etc) seems to be in a position to provide the required I/O disk performance. We therefore plan to test this Network File System in astronomical applications, such as CASA, in order to create our own cluster.

The users of the Italian ARC can already access the Virtual Organization (INAF –VO) of the Grid. In this VO area two medium-sized (some hundreds of CPU) clusters are available, while for specific projects arrangements can be made for the use of the large clusters that are available in the national Grid (Tier-1 and Tier-2 with thousands of CPU).

The Grid is still very under-developed for much of the astronomical applications. Presently it allows one to run a batch job that does not demand a particularly complex environment and that can take advantage from a "rough parallelism" like a pipeline for data reduction of hundreds or thousands of independent images. The language for the management of the Grid applications is similar to the Job Control Language of IBM-computers of the 1970's-1980's except that a "broker" will select the network node on which to run the job, transferring data and programs to this node.

But the Grid offers a computing power that cannot be ignored and creates a technological spin-off that can be used also outside the Grid itself. An example of this is the study of the integration of clusters, through the development of libraries for the parallelization of computer codes, network parallel file systems, and the techniques of virtualization and migration of applications. This last technique in particular promises to overcome one of the major limitations of the Grid, which is the inflexibility of the software environment on which the Grid is based. When an astronomical application is launched on the Grid it must

indicate which libraries it uses (fitsio, nag etc), and it can therefore run only on those nodes that have these resources, or the user must also send the required libraries and compile them *in loco*, at the node. With the present trend in software development to use dynamic rather than static compilation the problem becomes more critical because this increases the need to identify the nodes of the Grid that use those specific versions of the operating system that have compatible basic libraries.

With the virtualization and the migration of the applications the user could launch one application from his own work station, in other words he can migrate the entire "image" (kernel of the operating system, environment, libraries, and data) to the Grid to execute the job. When the job is finished it can return the original work station.

The virtualization and migration are some of the tasks on which is work is being done in the Grid-community that can lead to an extensive use of the Grid for the reduction of radioastronomical data. This obviously requires also an effort in order to make our applications ready for the Grid and to evaluate which tasks can benefit from the parallelization of the software codes.

For the Italian ARC we plan to create a small (tens of CPU) cluster on which to test, and eventually to optimize the application software for ALMA. This cluster will offer its users tens of Terabyte of disc space and a high-speed link (10 Gbit/sec) with the research network. The users will use the cluster in a "standard way", with access via 'ssh' or an X-windows sessions, but the cluster will also be made an integral part of the INAF Grid to be used as a test tool and for the preparation of applications that can eventually run on the major nodes of the Grid, offering the users the necessary support.

* Raid-Nas (**Redundant Array of Inexpensive Driver – Network Attached Storage**) typically is a system that can hold tens of disks and share the date space as a standard network file system via a local network. Other, more expensive solutions are the Raid-SAN (Storage Area Network) where the data space is connected to computers via specialised channels (fibre channel etc.)

For more information on what GRID is, see http://it.wikipedia.org/wiki/Grid (in Italian), or http://en.wikipedia.org/wiki/Grid_computing (in English).

October 10, 2007

Mauro Nanni
INAF – Istituto di Radioastronomia