New utilities and tools for using and monitoring ENEA-GRID environment

Status monitoring and networking
While data are sent and received by means of internet transport protocols like UDP and TCP, a remote user or application can start its processes on a host of the ENEA GRID by using several authentication protocols like rsh, ssh etc. Therefore it is important to test both network performance and authentication procedures. We present some of the developed tools: the rsh test, the UDP test and the TCP/IP bandwidth test.

These MPI implementations excell in one or more areas, while OpenMPI tries to collect them.

Parallel programming in Matlab
Matlab (MatLAB®) environment provides an environment for numeric elaboration and an interpreted programming language widely used in science and technology. The package is equipped with specialized libraries (toolboxes) which can solve problems in several fields. For some years the implementation of parallel programming in Matlab environment has been under study. Mathworks, the producer of Matlab, distributes toolboxes with this target, but there are also third party projects with the same aim. Among them are the “MatlabMPI” and “pMatlab” libraries developed by Lincoln Laboratories of the Massachusetts Institute of technology.

Since all ENEA-GRID’s hosts access to the same file system, we place a job launcher: its task is to check the host’s platform from which application was launched, select the proper binary and run it on the initial host. In a grid environment, jobs should be under control of resource manager and thus job launchers give to programmers the choice to submit applications in interactive or batch mode. Fig. 2.4 shows an example: regardless of serial or parallel jobs, we can integrate many different applications in a single interface to grid’s hosts. Moreover such an interface can be implemented.

Serial and parallel launchers
Grid computing enables development of large scientific applications using computational resources not available at a single site. Designing and implementing grid application often require to pool together resources that belong to different distributed computing systems. Grid computing enables development of large scientific applications using computational resources not available at a single site. Designing and implementing grid application often require to pool together resources that belong to different distributed computing systems.

Performance tests
HPL is a software package that solves a dense linear system in double precision (64 bits) arithmetic on distributed-memory computers. It provides a testing and timing program to quantify the accuracy of the obtained solution as well as the time it took to compute it. We have used HPL to measure the performance of two kind of ENEA-GRID’s Linux cluster 16 nodes with 2 Dual Core AMD Opteron per node, interconnected via Infiniband (IB) and 1 SMP node Power 4 with 32 CPUs. Fig. 4.1 and fig. 4.2 show the collected data. We report the following main parameters:

1. Memory: it is related to the dimension N of the matrix. Growing values of N allow a better utilization of memory hierarchy, from local to level 1 cache memory. Moreover, since HPL uses BLAS library, the algorithm can use Level 3 BLAS due to the two-dimensional Block-Cyclic Distribution Data.

2. Peak: it is the most significant benchmark’s result since it gives the number of floating point operation per second that the system can perform. We report it as a function of memory utilization (or N).

3. HPL: it is a software package that solves a dense linear system in double precision (64 bits) arithmetic on distributed-memory computers. It provides a testing and timing program to quantify the accuracy of the obtained solution as well as the time it took to compute it. We have used HPL to measure the performance of two kind of ENEA-GRID’s Linux cluster 16 nodes with 2 Dual Core AMD Opteron per node, interconnected via Infiniband (IB) and 1 SMP node Power 4 with 32 CPUs. Fig. 4.1 and fig. 4.2 show the collected data. We report the following main parameters:

1. Memory: it is related to the dimension N of the matrix. Growing values of N allow a better utilization of memory hierarchy, from local to level 1 cache memory. Moreover, since HPL uses BLAS library, the algorithm can use Level 3 BLAS due to the two-dimensional Block-Cyclic Distribution Data.

2. Peak: it is the most significant benchmark’s result since it gives the number of floating point operation per second that the system can perform. We report it as a function of memory utilization (or N).