

# Development and Optimization of CICC JINR Cluster in 2008–2009

Gh. Adam<sup>1,2</sup>, S. Adam<sup>1,2</sup>, A.S. Ayriyan<sup>1</sup>, V.V. Korenkov<sup>1</sup>, V.V. Mitsyn<sup>1</sup>

e-mail: adamg@jinr.ru, <sup>1</sup>Laboratory of Information Technologies, JINR, Dubna

<sup>2</sup>DFT IFIN-HH, Magurele – Bucharest, 077125, Romania

The JINR participation in the LHC experiments and in other large-scale projects asked for a substantial increase of its networking and information resources as well as the deployment of a large volume of work toward the development of the JINR Grid-segment and its integration in the Russian Grid-infrastructure RDIG (Russian Data Intensive Grid). The implementations has been done by staff of the Laboratory of Information Technologies (LIT), which is responsible for the computing infrastructure development in JINR and hosts the Central Information and Computing Complex (CICC) of the Institute.

The LIT JINR CICC, comprising the high-efficiency computing cluster and the data storage systems, is the kernel of the resource level of the JINR information infrastructure. To reach the established targets within the WLCG project for the effective processing and analysis of the experimental data within the future LHC projects, steep increases in the performances of the CICC cluster and disk space are needed.

In this report both the extensive and intensive approaches associated to this process are highlighted. On the extensive side, rapid increase was achieved during 2007–2009 of both the installed CPU computing power (2300 kSI2K by the mid 2009) and of the disk storage capacity (500 TB by the same date). On the intensive side, optimization of the various module interconnects and of the supervising software resulted in a leading position of the CICC within the RDIG and in a significant one on the worldwide scale.

Two essential constraints have to be satisfied by the CICC configuration. First, the floating point computation runs at the CICC cluster have to accommodate requests for traditional sequential applications, parallel computing applications, as well as Grid applications launched within various virtual organizations. Second, its upgrade by new acquisitions is always subject to tough financial constraints. As a consequence, new modules are acquired from vendors offering the best price of the day, resulting in heterogeneous computing cluster and disk storage area structures, with home made implementations of the various module interconnects and of the supervising software.

The CICC computing facilities were substantially

increased during 2007–2009 by acquisitions of multicore processor modules enabled with 2 Gb/core RAM and Gigabit Ethernet (GbE) interprocessor connections inside each module. The 2007 acquisitions involved three rackmount modules – two from T-Platforms and one from Hewlett-Packard, each consisting of 20 dual-core 2.66 GHz Intel Xeon 5150 processors with one Gigabit Ethernet (GbE) interconnect (called GbE-I in Fig. 1). During 2008, three blade modules of 20 quad-core 2.66 GHz Intel Xeon E5430 processors each were acquired from SuperMicro (called GbE-II in Fig. 1). A fourth module from SuperMicro, consisting of 20 quad-core 3.0 GHz Intel Xeon X5450 processors can work in two OS defined regimes of interprocessor connection: either under GbE interconnect (called GbE-III in Fig. 1) or under InfiniBand (InfB) interconnect.

The 2009 extensions involved Supermicro Twin 40 eight-core Intel Xeon E5420 2.50GHz processors with GbE interconnect and Supermicro Superblade 10 eight-core Intel Xeon E5410 2.33GHz processors with both GbE and InfB interconnects.

In order to check the actual performance of the resulting configuration and to rise it at the state-of-the-art level, we have decided to perform independent measurements of system performance by worldwide accepted benchmarks. The characterization of computer performance is usually associated to parallel computing. Since the three classifications of the most performing computing systems, TOP500 [1], CIS TOP50 [2], and China TOP100 [3] are based on outputs obtained from the High Performance LINPACK (HPL) Benchmark [4], we decided to use the HPL benchmark as well (version 1.0a of January 20, 2004). We used an Intel C Compiler v10.1 and the Intel Math Kernel Library 10.0.

Performance measurements made in 2007 and in the first half of 2008 [5]–[7] and system exploitation evidenced the existence of bottlenecks which were identified and alleviated by parallelization of the information transfer between the different modules of the system. Low cost improvements of the hardware connections (modification of the connection of each GbE-I module to the main Backbone Ethernet switch to a four-port GbE trunk) resulted in substantial gains concerning the latency decrease and the consistent data handling to/from the associated mass storage peripherals.

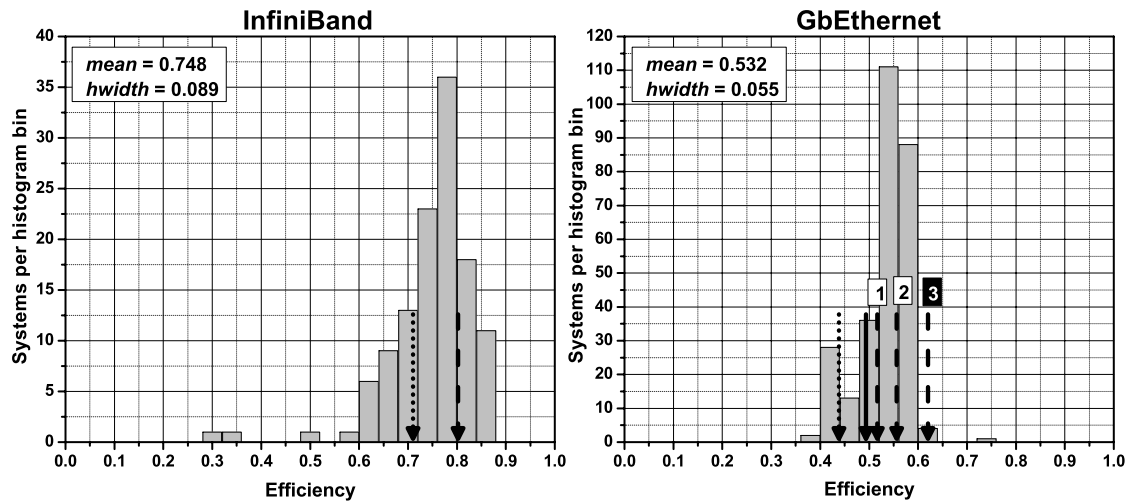


Figure 1: Performance of CICC JINR in 2008 [8]: results of optimization implementations and comparison with the histogram representations of the June 2008 issue of the TOP500 data for the InfB and GbE interconnects. *InfiniBand*: Dotted arrow reproduces the data from [7]. Interrupted-line arrow points to the newly measured performance. *Gigabit Ethernet*: Dotted arrow points to the previously reported data in [5, 6]. Solid line arrow points to the overall 560-core heterogeneous structure. Interrupted-line arrows point respectively to: 1 – GbE-I; 2 – GbE-II; 3 – GbE-III.

The increase of the RAM/core also resulted in sensible performance gains too. As a result of the implemented optimizations, the system worked efficiently for all three abovementioned categories of jobs.

Within the Russian Data Intensive Grid (RDIG) consortium comprising, besides the CICC JINR, 14 Russian computing centres, our cluster covered a sizeable part of the RDIG share to the LHC projects since 2007.

Performance measurements of the upgraded 2008 configuration [8] showed relative figures at the level of the best results reported in the June 2008 TOP500 edition of the most performing computers in the world (Figure 1).

The 2009 upgrades of both the computer cluster and the disk area were heavily based on the use of the lessons learned from the 2007–2008 system developments.

## References

- [1] <http://www.top500.org/>
- [2] <http://www.supercomputers.ru/>
- [3] <http://www.samss.org.cn/2007-China-HPC-top100-20071110-eng.htm>
- [4] A. Petitet, R.C. Whaley, J. Dongarra, A. Cleary, *HPL - A Portable Imple-*

*mentation of the High-Performance Linpack Benchmark for Distributed-Memory Computers*, <http://www.netlib.org/benchmark/hpl/>, 2004.

- [5] Gh. Adam, S. Adam, A. Ayriyan, E. Dushanov, E. Hayryan, V. Korenkov, A. Lutsenko, V. Mitsyn, T. Sapozhnikova, A. Sapozhnikov, O. Streltsova, F. Buzatu, M. Dulea, I. Vasile, A. Sima, C. Visan, J. Busa, I. Pokorny, *Performance assessment of the SIMFAP parallel cluster at IFIN-HH Bucharest*, *Romanian Journ. Phys.*, **53** (2008) 665-677.
- [6] A. Ayriyan, Gh. Adam, S. Adam, E. Dushanov, V. Korenkov, A. Lutsenko, V. Mitsyn, O. Streltsova, *Performance assessment of JINR CICC supercomputer*, *Proceedings of XII Scientific Conference of JINR Young Scientists and Specialists*, JINR Dubna, ISBN 978-5-9751-0045-0 (2008), pp. 71-74.
- [7] Gh. Adam, S. Adam, A. Ayriyan, V. Korenkov, V. Mitsyn, M. Dulea, I. Vasile, *Consistent performance assessment of multicore computer systems*, *Romanian Journ. Phys.*, **53** (2008) 985-992.
- [8] A. Ayriyan, Gh. Adam, S. Adam, V. Korenkov, A. Lutsenko, V. Mitsyn, *CICC JINR Cluster 2008 Performance Improvement*, communication POS(ACAT08)054, in *Proceedings of the International Conference "XII Advanced Computing and Analysis Techniques in Physics Research"*, November 3-7 2008, Erice, Italy, *Proceedings of Science*, SISSA, Trieste, 2009.