Optimization of data center performance in V.Lashkaryov Institute of Semiconductor Physics NAS of Ukraine for the physicochemical diagnostics of materials properties

V.P. Bryksa, O.O Mishenko, V.O Panchenko

[1]*V. Lashkaryov Institute of Semiconductor Physics NAS of Ukraine, 41 Nauky pr., 03028 Kyiv, Ukraine*

bryksa@isp.kiev.ua

Abstract. *In connection with enormous adoption of highly efficient calculations with the use of multiprocessor and multinode computational clusters in applied physical-chemical material analysis the necessity of development the special-purpose datawarehousing and data processing centers become of decisive importance. Such centers must be round-the-clock user-accessible which implies adherence of standards concerning safety of data storage and durability. The presented report is devoted to the evaluation of data center capabilities and to performed optimization of computational cluster productivity in ISP NAS of Ukraine.*

Keywords

Data center, Tier 1-4 levels, HPC cluster, AMD Opteron

1 Introduction

The range of problems is caused by the necessity of servicing the large clusters which must be round-the-clock one and safety under maintenance. This in turn depends both on the engineering of infrastructure in the server rooms or in data centers and on the communication channels to support the remote user access to the computational cluster resources. The infrastructure must guarantee the usability of system administration and possibility of computation power increasing both for standalone cluster and for the whole data center. The data centers classification may be realized using a several features, e.g. according to a technical compliance (ANSI TIA-942, GOSTs of 34 series), according to available computer resources or according to the reliability category (Tier 1-4 levels).

In addition for the research purposes it is convenient to perform classification according to the kinds of considered scientific problems and installed on the data center servers software which determines the efficiency of computer resources using. At the other hand, the productivity of data center depends on a lot of architecture features of servers and communication hardware for interserver data exchange. It is important correctly develop the cluster architecture so that access channels to the distributed file system or to the data storage of cluster were balanced as well as possible. When the complexity of investigated scientific and technical tasks and the number of users is rise then the problem of computational clusters scaling appears. The productivity often may be increased by an increasing the number of servers/processors or by the replacement of server hardware with the new type one. As a rule, the hardware became obsolete during 3-5 years which sometimes make impractical the service and usage of such hardware in the working cluster systems located on the data center area. Therefore estimating of servicing or upgrading the data center together with evaluation of multiprocessor codes productivities are of crucial importance.

2 Operation of data center in ISP NAS of Ukraine

The experimental data center of Tier 2 level with the total number of server sites 252U was designed and commissioning as a result of collaboration of ISP NAS of Ukraine, TOV “Entry” [1] together with regional agency of AMD company in Ukraine. There is the balanced system of ventilation operating with the air recirculation and cooling in the data center room. This system has two operating modes (summer-winter) and is designed for an optimized removing of heat excess from the server and communication hardware of about 60 kW.

The stable temperature of +21±0.5°C is maintained in the room. The air inflowing proceeds through the variable geometry inlets built-in the raised floor while the air outflowing pass through the upper zone of the room and the variable geometry inlets as well. To support the high level of engineering infrastructure the data center of ISP is supplied with electric power according to the first category with the system of uninterruptible power supplies, system of automatic fire-fighting and safeguard system with secure control of access. For the safe communication channel assurance two independent optical inputs from the different internet providers working according to the BGP4 protocol are used in the data center.

At present cluster located at the data center area in ISP of NASU and based on five servers each of 20 AMD Opteron quad-core processors and Socket F bus type is available. Interserver data exchange is realized with high-speed switch Infiniband Flextronics DDR. Peak productivity of cluster was estimated to be as large as ~736GFLOPs. In addition the data warehouse server with the storage capacity of ~16TB is connected to the cluster (Fig. 1). Operation system is based on the free GNU/Linux software and Debian 6.0 package optimized for AMD64 architecture.

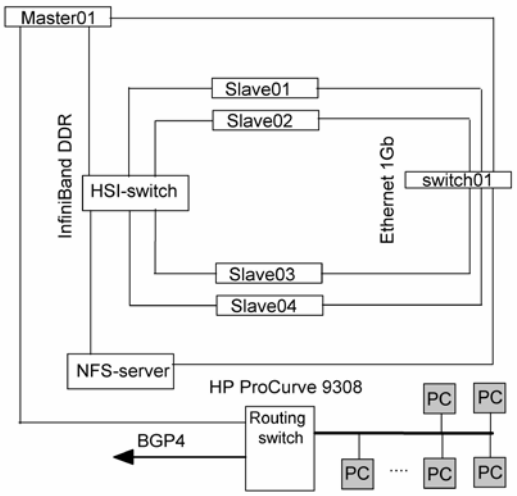

Fig. 1. Map of cluster in ISP NAS of Ukraine. Master01 – the main host (Tyan Transport FT48 B4985). Slave01-04 – auxiliary hosts (Tyan Transport GT26 B4987). NFS-server – data storage server (Tyan Thunder h2000M S3992-E). HIS-switch – high-speed data switch (Infiniband Flextronics DDR F-X430046). Switch01 – gigabit management switch (HP ProCurve Switch 2810-24G). PC – local network personal computers.

Computational cluster is used in the fundamental scientific investigations as well as in applied fields such as optoelectronics, micro- and nanoelectronics. Intensive developments of application of the numerical methods of computer tomography are now in progress [2] as well as quantum mechanics calculations of semiconductor materials [3] and Monte-Carlo modeling of semiconductor devices operating in strong electric and magnetic fields.

Data center serve the number of modern diagnostic laboratories in ISP – atomic force microscopy, confocal scanning spectroscopy of Raman scattering, high-resolution X-ray structural analysis and precise electrophysical measurements for diagnostics of physicochemical parameters of device structures [4]. A large data arrays occurring during complex investigation of such structures and realization the nondestructive methods of a volume diagnostics of device parameters is the reason for using clusters with the productivity more then 100 TFLOPs. Such server parameters at present are available for new AMD Opteron processors of 6100 series supporting the Socket G34 bus. With the use of such servers one may evaluate a possibility of cluster design with peak productivity more than 10 TFLOPS which will occupy area about 20U at the data center site.

3 Productivity analysis of AMD Opteron based clusters

The quad-core AMD Opteron processors of 8000 series enter the market in 2007 and were designed mainly for multiprocessors server systems with SMP architecture supporting DDR2 type of memory (maximum 256 Gb per server). By the use of high speed connection Hyper Transport of 2.x specification the data exchange between processors buses (Socket F) as well as between the chip sets of data-in/data-out for the peripheral devices, e.g. for PCIx bus controller or controllers of LAN, SAS devices and hard disks arrays are realized. Peak productivity of Socket F based four-processors system as well as of recently announced new Socket G34 bus for AMD Opteron processors of 6100 series appeared in 2012 is estimated in Table 1.

Tab.1. Analysis of parameters and peak productivity of AMD Opteron processors based four-processors servers. For calculation of server productivity the frequency of 2.3 GHz is supposed

	***4-CPU* Systems**	
	Total Cores/CPU	Productivity, GFLOPS
Socket F	**16/4**	**147,2**
	24/6	**220,8**
Memory DDR2, GB	**256**	
Hyper Transport, GT/s	**2,0×3-channels**	
Socket G34	**48/12**	**441,6**
	64/16	**588,8**
Memory DDR3, GB	**768**	
Hyper Transport, GT/s	**6,4×4-channels**	

Productivity of AMD Opteron based cluster with Socket F bus is not additive value of server number. The cluster productivity losses are mainly caused by the features of interserver connection for data exchange, e.g. connection type (LAN, SAN, WAN), by the size of data burst sequentially transferred according to connection protocol (MTU) as well as by sizes of FIFO bus registers (PCIx) of both bridge motherboard chip sets and communication controllers. Incorrect balancing of these parameters may lead to considerable looses of cluster productivity though one use the most fast systems for interserver data communication such as Infiniband QDR or FDR.

It is difficult to suggest reliable model of cluster productivity dependence versus its real parameters. Therefore it is more rational to perform the computational cluster optimization using testing technique for known software codes with monitoring of main parameters of individual cluster units. The *ab-initio* quantum mechanical codes ORCA 2.9.1 and Quantum Espresso 4.3.2 which use the OpenMPI parallelization technique and GCC 4.4.5 compiler were selected for this purpose. Performed in the data center of ISP NAS of Ukraine productivity testing of computational cluster build of five servers and Infiniband DDR interserver data communication shows that decreasing of productivity may be up to 20-30% of peak value.

4 Conclusion

The data center of Tier 2 level in ISP NAS of Ukraine which is used for investigations in the field of technology of semiconductor materials was presented. The analysis of data center engineering architecture features and working test were performed. The analysis of cluster productivity built of series 8000 AMD Opteron processors and depending on the type of interserver communication for data exchange was carried out. The problems of cluster productivity dependencies on its main parameters were discussed.

References

[1] http://www.entry.kiev.ua/.
[2] V.V. Strelchuk, V.P. Bryksa, K.A. Avramenko, M.Ya. Valakh, A.E. Belyaev, Yu.I. Mazur, M.E. Ware, E.A. DeCuir Jr., and G.J. Salamo: Confocal Raman depth-scanning spectroscopic study of phonon-plasmon modes in GaN epilayers. J. Appl. Phys., 109:123528-123536, 2011.
[3] T.T. Petrenko, T.L. Petrenko, Jahn-Teller centers in semiconductors as a challenge for first-principles methodologies. Analysis of failure of the supercell approximation for “shallow boron” impurity in SiC. International workshop “Advanced EPR for material and solar energy research”, Helmholtz-Zentrum Berlin fur Materialein und Energie, 13-14 October 2011, Berlin.
[4] Diagnostic center for collective use “Diagnostic of Semiconductor Materials, Structures and Devices”. http://www.microscopy.org.ua/.