LATVIAN JOURNAL OF PHYSICS AND TECHNICAL SCIENCES
2006, N 2

APPLICATION OF CLUSTER COMPUTING IN MATERIALS SCIENCE

A. Kuzmin

Institute of Solid State Physics, University of Latvia,
Kengaraga str. 8, Riga, LV-1063, LATVIA

Solution of many problems in materials science requires that high performance computing (HPC) be used. Therefore, a cluster computer, Latvian SuperCluster (LASC), was constructed at the Institute of Solid State Physics of the University of Latvia in 2002. The LASC is used for advanced research in the fields of quantum chemistry, solid state physics and nanomaterials. In this work we overview currently available computational technologies and exemplify their application by interpretation of x-ray absorption spectra for nano-sized ZnO.

Key words: *cluster computing, quantum chemistry, x-ray absorption spectroscopy.*

1. INTRODUCTION

High Performance Computing (HPC) is the technology used to provide solutions to the problems that require significant computational power, quick access or processing of very large amounts of data or operation across a geographically distributed network [1]. Today HPC infrastructure is a critical resource for research and development. The covered areas are related to telecommunications, weather and climate research/forecasting, financial risk analysis, car crash analysis, databases and information services, manufacturing, geophysics, computational chemistry and biology, pharmaceutics, aerospace industry, electronics and many others. The standard tasks addressed by HPC systems require normally hundreds or even thousands of processor hours to be completed. Therefore special approaches based on dedicated software and hardware components have to be used. Traditionally such problems are addressed using supercomputers.

Recently, a significant improvement in the communication network technologies and standard workstation processor speed as well as better understanding of applications and algorithms have led to emerging of a new class of systems called *clusters* or networks of workstations (NOW). Cluster-based systems are able to compete in performance with conventional, so-called "mainframe", supercomputers and have excellent price/performance ratios for special application types. In principle, the clustering technology can be used for any arbitrary group of computers, allowing homogeneous or heterogeneous systems to be built. Still better performance can be achieved by combining groups of clusters into HyperCluster or even Grid-type system [1].

Further we will give an example of the HPC cluster, called LASC (Latvian SuperCluster), which was installed at the Institute of Solid State Physics of the University of Latvia during 2002 [2]. The main goal of the LASC project is to secure the access of researchers to the HPC system, which is able to help in solving

modern physical problems by numerical simulations. The need for such installation is determined by a complexity of the tasks appearing now within fundamental and applied research projects in solid state physics and materials science, especially dealing with nanomaterials. We will also discuss computational technologies currently available at LASC and show their application to x-ray absorption spectra interpretation for nano-sized zinc oxide.

2. PRINCIPLES OF CLUSTER COMPUTING

Cluster computing refers to the technologies that allow multiple computers, called cluster nodes, to work together with the aim to solve common computing problems. The generic cluster architecture is shown in Fig. 1. Each node can be a single or a multiprocessor computer, such as a PC, a workstation or a multiprocessor server equipped with its own memory, I/O devices and an operating system. The nodes are interconnected by high-speed local area network (LAN), for example, Fast Ethernet or Gigabit Ethernet.

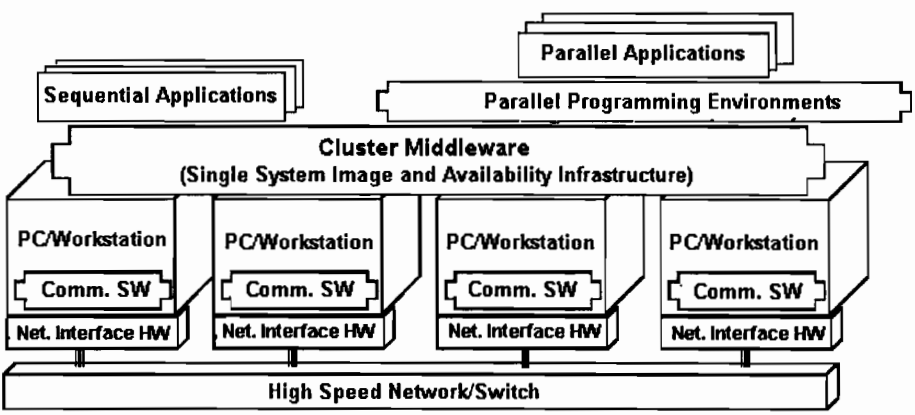

Fig. 1. Typical cluster architecture [1].

Various operating systems, including Linux, Unix and Windows, can be used to operate the nodes. However, in order that the cluster is able to pool its computing resources, special cluster-enabled applications must be written using clustering libraries, such as MPI (Message Passing Interface) and PVM (Parallel Virtual Machine), or a system level *middleware* [1] should be used. Using MPI or PVM programmers design applications that can span across entire cluster computing resources rather than to be confined to the resources of a single machine. For many applications, MPI and PVM allow computing problems to be solved at a rate that scales almost linearly in relation to the number of processors in the cluster. There are many implementations of the clustering libraries available, and the one used at the LASC is MPICH-MPI.

However, many potential users of the cluster technologies would like to have some kind of performance benefit using standard (sequential) applications. A reason for that can be the absence of access to the application source code or impossibility of the application parallelization due to the algorithm specifics. To partially solve this problem, a rather new technology, called openMosix [3], can be used. It allows standard applications to take advantage of clustering without being rewritten

or even recompiled. OpenMosix is an "add-on" to the standard Linux kernel, which uses adaptive load-balancing techniques and allows process running at one node in the cluster to migrate transparently to another node where they can execute faster. However, openMosix cannot execute a single process on multiple physical processors (CPU) simultaneously, therefore openMosix will not be able to speed up a single process/program, except to migrate it to a node where it can perform most efficiently. At the same time, openMosix can migrate the most standard Linux processes between nodes and, thus, allows for extremely scalable parallel execution at the process level. Besides, if an application forks many so-called "child" processes then openMosix will be able to migrate each of these processes to an appropriate node in the cluster. Thus, openMosix provides a number of benefits over traditional multiprocessor systems. Moreover, openMosix allows the creation of automatically configurable cluster with dynamic architecture having a variable number of computational nodes. This allows for a temporary increase in the cluster computation power using idle laboratory computers.

3. LATVIAN SUPERCLUSTER (LASC)

In this section we will briefly describe a cluster computer, Latvian Super-Cluster (LASC) (Fig. 2), located at the Institute of Solid State Physics (ISSP) of the University of Latvia [2]. Currently, the LASC is used for advanced research in the fields of quantum chemistry, solid state physics and nanomaterials.

The LASC is running under RedHat Linux operating system with open-Mosix extension. Currently it consists of 12 double-processor nodes: one front-end node playing the role of the fileserver, and 11 computational nodes, which are interconnected by two local area networks with 100 Mbps and 1 Gbps speeds. The total computational resources available to the users are 22 CPU, having the theoretical peak performance of about 48 Gflops, about 44 GB of RAM and 3 TB of the total hard disk space. The cluster is isolated from the Internet by a firewall allowing secured connection for authorized users.

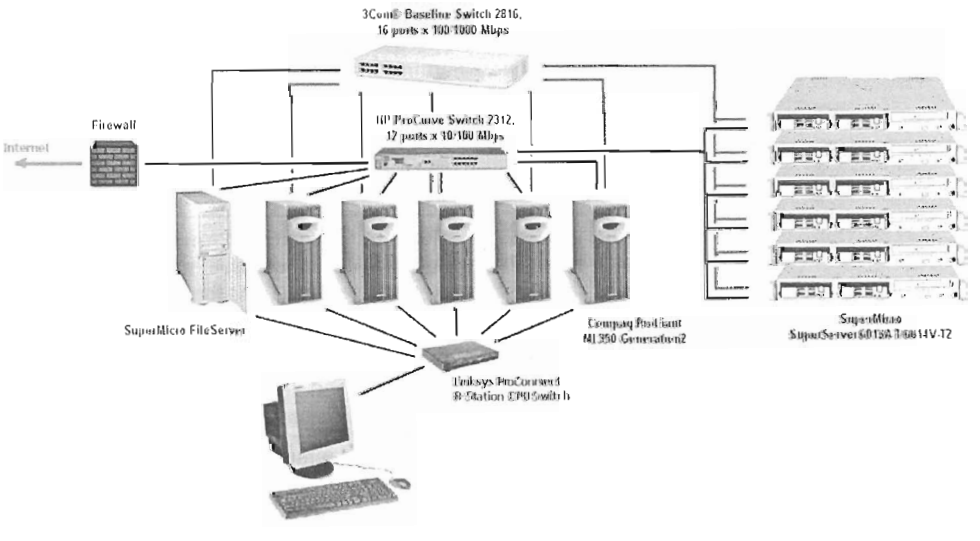

Fig. 2. Latvian SuperCluster (LASC) [2] (for details see http://www.cfi.lu.lv/lasc).

3. HPC APPLICATION IN MATERIALS SCIENCE

The current use of HPC LASC system at the ISSP is dedicated to several projects: quantum chemical modelling of advanced materials, *ab initio* modelling of x-ray absorption spectra, simulation of the electronic paramagnetic (spinal) resonance (EPR or ESR) spectra and simulation of the structure of organic monolayers on gold surface. These activities involve about 20 researchers/students from five ISSP laboratories and have resulted in more than 40 publications during three years of the LASC operation.

While each researcher is free to use its own simulation software, a number of well- known codes is available on LASC for the common use (see Table 1). They allow performing *ab initio* and semi-empirical modelling of 0D/1D/2D/3D materials and their properties at different levels of approximation. Further, possible interpretation of the Zn K-edge x-ray absorption spectra from ZnO nanorods by *ab initio* FDMNES code [4] is given.

Table 1

Scientific software available on LASC
(for details see [5])

Code name	Application description
Crystal 98/2003	The program computes the electronic structure of periodic materials within the Hartree-Fock, density functional or various hybrid approximations.
VASP	A package for performing *ab initio* quantum-mechanical molecular dynamics (MD) using pseudopotentials and a planewave basis set.
Abinit	*Ab initio* code for molecules and solids, within Density Functional Theory (DFT), using pseudopotentials and a planewave basis.
DeFT	Density Functional Theory (DFT) code for molecules.
PWscf	A set of programs for electronic structure calculations within Density-Functional Theory and Density-Functional Perturbation Theory, using a planewave basis set and pseudopotentials.
LmtART	The full-potential linear-muffin-tin-orbital (FP-LMTO) program to perform band structure, total energy and force calculations within the methods DFT methods.
SPR-KKR	A spin polarized relativistic Korringa-Kohn-Rostoker code for calculating solid state properties.
FDMNES	The program calculates the spectra of different spectroscopies related to the real or virtual absorption of x-ray in material using the Finite Difference Method (FDM) and the Green functions formalism (multiple scattering) with a muffin-tin potential.
FEFF	An automated program for *ab initio* multiple scattering calculations of x-ray Absorption Fine Structure (XAFS) and x-ray Absorption Near-Edge Structure (XANES) spectra for clusters of atoms.
GULP	The General Utility Lattice Program is a semi-empirical code performing a variety of tasks relating to 3D solids, as, for example, interatomic potential fitting, energy minimisation and phonon calculations.
MOLDY	A general-purpose Molecular Dynamics simulation program.

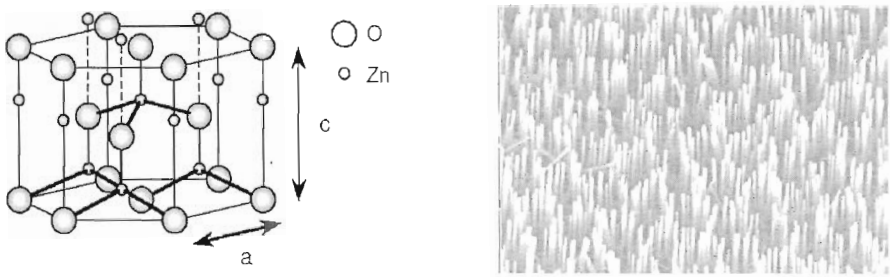

Fig. 3. Wurtzite zinc oxide structure (left) and aligned ZnO nanorods/nanowires epitaxially grown on Si substrate in the presence of catalyst [5] (right)

Wurtzite zinc oxide has a hexagonal structure (space group $C6mc$) with lattice parameters a=3.2495 Å and c=5.2069 Å (Fig. 3). The structure of ZnO can be simply described as a number of alternating planes composed of tetrahedrally coordinated O^{2-} and Zn^{2+} ions stacked alternately along the c-axis. The tetrahedral coordination in ZnO results in a noncentral symmetric structure and, consequently, in piezoelectricity and pyroelectricity. ZnO has a rich family of nanostructures, such as nanocombs, nanorings, nanobelts, nanohelixes/nanosprings, nanowires and nanocages, which can be synthesized under specific growth conditions [6]. Among these, the ZnO nanorods have attracted significant interest owing to their potential applications for short-wavelength light emitting devices [7]. ZnO nanorods can be prepared using several techniques, such as the vapour–liquid–solid (VLS) approach with metal catalyst (Au or Sn) [6] or the electrodeposition procedure [8]. Depending on the preparation conditions, nanorods of different (several tens to thousands of nanometers) height and diameter can be obtained. Knowledge of the electronic structure of the nanorods is crucial for understanding the basic physics underlying their applications.

Recently, polarization and angle-dependent x-ray absorption near-edge structure (XANES) studies have been performed for highly oriented ZnO microrods and nanorods [9-12]. Note that XANES is sensitive to the local atomic and electronic structure around the absorbing centre [13], thus being an optimal experimental technique to probe nanostructures. Typical Zn K-edge XANESs in ZnO nanorods are shown in Fig. 4.

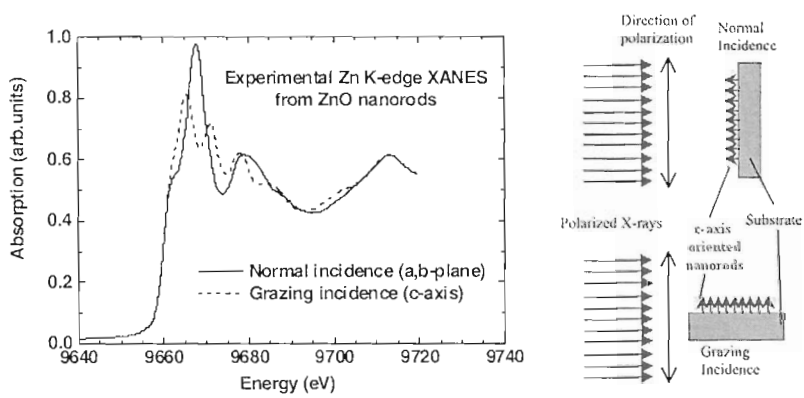

Fig. 4. Left panel: experimental Zn K-edge XANES in ZnO nanorods (from [10]) grown on Si substrate in the presence of catalyst and oriented in the [0001] direction. Right panel: schematic drawing of sample orientation relative to a polarized synchrotron radiation beam

Interpretation of the XANES spectra requires intensive computer simulations, which can be easily parallelized or equally distributed among several CPUs [14]. Therefore this task is a proper candidate for the use on a cluster computer such as LASC. The XANES signal $\mu(\omega)$ can be described using the Fermi's Golden rule [14]

$$\mu(\omega) \propto \sum_f \left|\langle i|d|f\rangle\right|^2 \delta(h\omega + E_i - E_f).$$

Here $h\omega + E_i \equiv E$ is the photoelectron energy, d is the coupling to x-ray field ($d = \hat{\varepsilon} \cdot \vec{r}$ in the dipole approximation), E_i is the energy of the core level (1s(Zn) in our case) and the sum is taken over unoccupied final states of energy E_f. Usually, a standard quasi-particle model is used in which the final states $|f\rangle$ are calculated in the presence of an appropriately screened core-hole, and all other many-body effects and inelastic losses are treated using a complex energy-dependent self-energy or optical potential [14].

In this work we have used the real-space full-multiple-scattering (FMS) approach, as implemented in the FDMNES code [4], to calculate Zn K-edge XANES signals for bulk polycrystalline ZnO and oriented ZnO nanorods. The calculations were performed within the dipole approximation (1s→np transitions) for clusters of up to 8 Å radius having the structure of ZnO crystal and centred at a zinc ion (the absorber). The cluster potential was of the muffin-tin type, with 10% overlap for the nearest atoms. The exchange-correlation effects were approximated by the complex energy-dependent Hedin-Lundqvist potential. The calculated spectra were intentionally not broadened to keep the fine structure information (shown in Fig. 5).

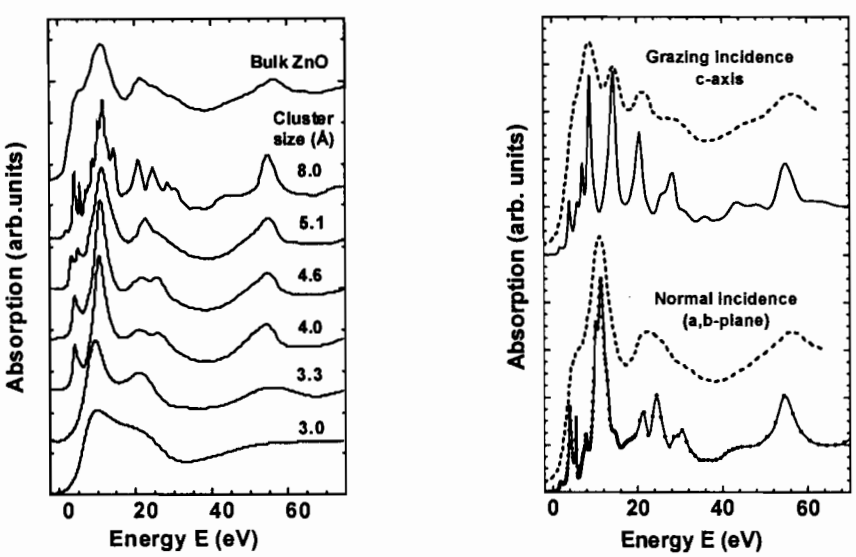

Fig. 5. Left panel: Comparison between experimental (upper curve) and FMS calculated Zn K-edge XANES signals in bulk polycrystalline ZnO for different cluster sizes around zinc atom. Right panel: Comparison between experimental (dashed curves, taken from [10]) and FMS calculated Zn K-edge XANES signals for oriented ZnO nanorods. The energy origin E=0 eV corresponds to the continuum threshold.

Our first set of the Zn K-edge XANES simulations (left panel in Fig. 5) for bulk polycrystalline ZnO gives good agreement between theory and experiment, allowing all peaks in the experimental signal to be interpreted. The main peaks of the experimental XANES are reproduced already for the cluster of 4 Å radius containing only 27 atoms of the first 3 coordination shells. The nearest 4 oxygen atoms at 1.95–1.99 Å are responsible for broad peaks observed in the cluster with a size of 3.0 Å. The contribution from the second coordination shell composed of 12 Zn atoms at 3.21–3.25 Å makes the peaks in Fig. 5 at 10, 20 and 55 eV more resolved. Finally, the 10 oxygen atoms of the third shell at ~3.8 Å give origin to the shoulders at 4, 27 and 48 eV. Thus, the Zn K-edge XANES probes mainly the nearest environment around zinc atoms. However, a cluster with a larger size of ~ 8 Å and composed of 177 atoms is required to reproduce the whole fine structure observed. This size is an indication of the photoelectron mean free path in ZnO and thus allows estimating the XANES sensitivity to the size effects in the case of nanocrystalline oxide. Note that ZnO nanorods in [10] were ~250 nm long and ~45 nm in diameter, so that their angle-dependent XANES spectra (Fig. 4) can be accurately described using our bulk model. This conclusion is also supported by weak dependence of the experimental Zn K-edge XANES on the ZnO nanorod diameter [12]. The results of the calculations performed for a cluster size of 8 Å, taking into account the synchrotron beam polarization and sample orientation, are shown in Fig. 5 (right panel). They are in very good agreement with experiment, indicating the ability of modern theory [14] to predict and explain the XANES signals from nanosized zinc oxides.

Note that the calculation of one XANES spectrum for the cluster size of 8 Å takes about 25 hours of computations on single XEON 2.4GHz CPU. Therefore, the use of such HPC system as LASC is essential to address these problems.

4. CONCLUSIONS

The Latvian SuperCluster (LASC) system forms the core of high-performance computing at the Institute of Solid State Physics of the University of Latvia. It provides scientists with computational resources that make it possible to use modern theoretical methods and to maintain a competitive level of research in the field of materials science. Besides, the LASC plays the important role in verification of new computational technologies, thus creating a background for future development. The cluster-type architecture of LASC ensures its adequate reliability, simple expandability in the future and relatively easy administration. An example of LASC application to x-ray absorption spectra analysis in nanocrystalline ZnO has been demonstrated.

REFERENCES

1. Buyya R. (ed.) (1999) *High Performance Cluster Computing: Architectures and Systems* New York: Prentice Hall 881.
2. Kuzmin A. (2003) *Computer Modelling & New Technologies* **7** 7–15.
3. Sloan J.D. (2004) *High Performance Linux Clusters with OSCAR, Rocks OpenMosix, and MPI* Sebastopol: O'Reilly 361.
4. Joly Y. (2001) *Phys. Rev. B* **63**, 125120:1–10.
5. See http://www.cfi.lu.lv/lasc.
6. Wang Z.L. (2004) *J. Phys.: Condens. Matter* **16** R829–R858.

7. Huang M.H., et al. (2001) *Science* **292** 1897–1899.
8. Cembrero J., Elmanouni A., Hartiti B., Mollar M., Marí B. (2004) *Thin Solid Films* **451–452** 198–202.
9. Guo J.H., et al. (2002) *J. Phys.: Condens. Matter* **14** 6969–6974.
10. Chiou J.W., et al. (2004) *Appl. Phys. Lett.* **84** 3462–3464.
11. Lee E. Y. M., Tran N., Russell J., Lamb R. N. (2004) *J. Appl. Phys.* **92** 2996-2999.
12. Chiou J.W., *et al.* (2004) *Appl. Phys. Lett.* **85** 3220–3222.
13. Aksenov V.L., Kuzmin A.Y., Purans J., Tyutyunnikov S.I. (2001) *Phys. Part. Nucl.* **32** 675–707.
14. Rehr J.J., Ankudinov A.L. (2005) *Coord. Chem. Rev.* **249** 131–140.

KLASTERU APRĒĶINU PIELIETOŠANA MATERIĀLU ZINĀTNĒ

A. Kuzmins

K o p s a v i l k u m s

Materiālu zinātnē daudzu problēmu risināšanā jāizmanto augstas veiktspējas skaitļošana (AVS). Tradicionāli šādas problēmas risina, izmantojot superdatorus. Taču PC-tipa datoru darbības ātruma pēdējais milzīgais pieaugums atver relatīvi lētu un pielietojamu AVS risinājumu, izmantojot klasteru tehnoloģijas. Tāpēc Latvijas Universitātes Cietvielu fizikas institūtā 2002. gadā tika izveidots klasteru dators – Latvijas SuperKlasteris, kas tiek izmantots perspektīviem pētījumiem kvantu ķīmijas, cietvielu fizikas un nanomateriālu jomā. Šajā darbā aplūkotas patlaban pieejamās skaitļošanas tehnoloģijas un parādīta to izmantošana nano-ZnO rentgenabsorbcijas spektru interpretācijā.