Development of Transferable Materials Information and Knowledge Base for Computational Materials Science

Group Representative

Shuhei Ohnishi   CAMP (Collaborative Activities for Materials Science Programs) Group*

Authors

Shuhei Ohnishi   NEC Fundamental and Environmental Research Laboratories

Satoru Shingu   NEC System Technologies, Ltd.

＊CAMP group members in the ES project

NEC Corp., Fundamental and Environmental Research Labs.

Toyota Central R&D Labs., INC

Sumitomo Chemical Co. Ltd., Tsukuba Research Labs.

Toshiba Corp., R&D Center,

Mitsubishi Heavy Industries, Ltd., Advanced Technology Research Center

Asahi Glass Co. Ltd., Research Center

FDK Corp., R&D Div.

Sony Corp., Research Center

Fuji Research Institute Corp.

ANCL, Inc.

NEC Informatic Systems, Ltd.

NEC System Technologies, Ltd.

The multi-industry CAMP program on the Earth Simulator (ES) is focused on dealing with various kinds of nanostructure materials of which calculated results are to be used as integrated information of comprehensive skills and knowledge base on computational materials science. We plan to implement two simulation codes for non-periodic and periodic systems based on the first principles density functional theory (DFT). In the first year of our project, we have tuned the non-periodic cluster calculation code by the standard linear combination of atomic orbitals method with the norm-conserving pseudopotential scheme (LCAO-PS). The LCAO-PS code was optimized by both inter-node and intra-node parallelization. We obtain 99.95% parallelization ratio and 64% efficiency using 176 CPUs and 1080 CPUs by this hybrid parallelization in the ES for the Pt_{135} cluster case. Electronic structures of Pt and Au clusters of nano-size are discussed.

Keywords: nanostructure materials, the first principles DFT, metal cluster, hybrid code

1. Overview of the project

The CAMP project (Collaborative Activities for Materials Science Program) is a multi-industry activity of cooperation in research on materials-modeling subjects in which computational scientists have broadly overlapping interests in materials simulation using various computers. Collaborations have mainly been devoted to develop standard free software and user-interfaces for information exchange for more than fifteen years by applying the first principles calculations based on the density functional theory (DFT) which has been established as one of the most powerful tools in understanding material properties theoretically. Recent rapid technological development in nanostructure systems, however, requires us not only to simulate but to design systems using computers. Large scale parallel computations by the Earth Simulator (ES) can provide us an innovative powerful technology in research on nanostructure materials, which are one of the most important systems in present industrial applications. Since we can deal with various nanostructure systems of several hundred atoms within a very fast turnaround time, we can build up materials information from calculated results. The goal of our project in the ES is to create the integrated knowledge base by exploiting calculated nanostructure material databases; transferable atoms and molecules picked up from calculated output-data are useful enough to use as the input data for other calculations. We plan to develop two type codes in the ES during FY2003–2005. One is for the non-periodic nanostructure

system of molecules and clusters by the standard linear combination of atomic orbitals method using the norm-conserving pseudopotential (LCAO-PS). The other one is for the periodic system. The non-periodic LCAO-PS code has been tuned for the ES in FY2003.

2. Tuning of the LCAO-PS code for the ES

A multi-center problem in the non-periodic system is the most intractable problems in simulating nanostructure materials; atomic basis functions and singularities by the nucleus Coulomb potentials are localized in the atomic site. The pseudopotential method makes it feasible get rid of the Coulomb singularity numerically. We have to solve the single particle Schodinger equation of which effective potential field is unique functional of the electron charge density distribution $\rho(\vec{r})$, so called the Khon-Sham equation, and the Poisson equation to provide Coulomb potential field to the Khon-Sham equation in the DFT. Our method in the LCAO-PS code is completely based on the multi-center numerical integration scheme for both equations.

We divide a physical quantity $F(\vec{r})$ to the particle site $\vec{R}_\alpha$ by multiplying localized weight function $w_\alpha(\vec{r}_\alpha)$ as follows,

$$F(\vec{r}) = \sum_\alpha F_\alpha(\vec{r}_\alpha) \qquad F_\alpha(\vec{r}_\alpha) = w_\alpha(\vec{r}_\alpha)F(\vec{r}) \quad where \quad \vec{r}_\alpha \equiv \vec{r} - \vec{R}_\alpha$$

$w_\alpha(\vec{r}_\alpha)$ is normalized as

$$\sum_\alpha \omega_\alpha(\vec{r}_\alpha) = 1 \qquad for \quad all \quad \vec{r}_\alpha$$

$$\iiint F(\vec{r})d\vec{r} = \sum_{\alpha=1}^{N} I_\alpha \qquad I_\alpha \equiv \iiint F_\alpha(\vec{r}_\alpha)d\vec{r}_\alpha$$

The multi-center numerical integration is decomposed to the one center integration problem for which efficient numerical grids are applied in terms of radial and spherical summations with typically 40×146 grid points. Total number of grid points are $N \times 40 \times 146$, where N is the number of atoms. The Coulomb potential $V_C(\vec{r})$ is also given by the sum of contribution from each site α by dividing the electron charge density $\rho(\vec{r}) = \sum_\alpha \rho_\alpha(\vec{r}_\alpha)$. Coulomb potential is given by the one dimensional numerical integration by decomposing $\rho_\alpha(\vec{r}_\alpha)$ into the sum of radial and spherical harmonics parts. There are several methods of this kind charge density decomposition. We selected the fitting method because of keeping high vector performance this FY2003.

Our present code had already been completely tuned to be able to have a high vector performance by implementing all integration parts into the deepest do-loop; the vector ratio is normally 99.5%.

Typical example for $F(\vec{r})$ is the overlap matrix of the Hamiltonian $\hat{H}$, which depends on atomic sites α, β and quantum numbers and number of radial basis functions of atomic kinds represented by κ, κ'.

$$F_{\beta\kappa'\alpha\kappa}(\vec{r}) = \varphi_{\beta\kappa'}(\vec{r})\hat{H}\varphi_{\alpha\kappa}(\vec{r})$$

Complicated quantum indexes require us deep do-loops operations. All physical quantities are stored on total grid points so that standard molecular orbitals method require huge amount of data which is normally dealt with using the temporary file. Almost computation time is spent by this file I/O. The high performance in the ES was obtained by dealing with whole data on memory. The present code, however, has a memory restriction to simulate nanostructure materials of several hundred atoms of especially heavy and noble atomic elements like Pt due to the 16GB memory limit per node. By keeping information of other atoms in each MPI process, we performed the hybrid type parallelization so as to use CPUs as possible as we can. Outer operation for atomic sites is dealt with MPI for inter-node parallelization
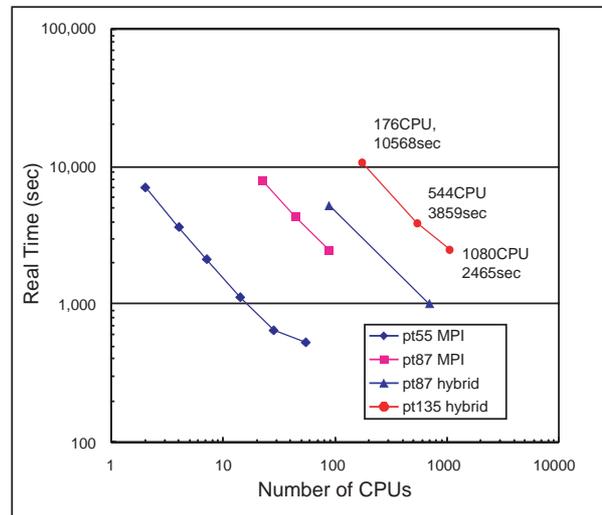
Fig. 1-1 Elapse time by MPI and hybrid parallelization of LCAO-PS code for various nano-size Pt clusters.
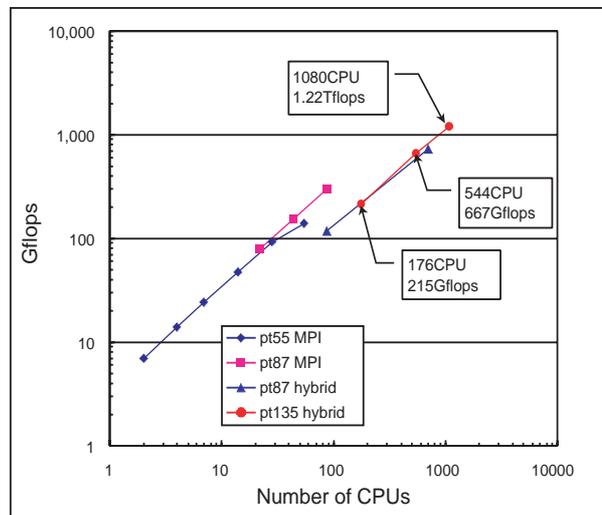
Fig. 1-2 Parallel performance by MPI and hybrid parallelization of LCAO-PS code for various nano-size Pt clusters.

and middle parts are dealt with the intra-node multi-task treatments. Figures 1-1 and 1-2 show the elapse time and parallel performance of LCAO-PS code for nanostructure Pt clusters, respectively. The parallelization ratio is 99.95% and the parallelization efficiency is 64% evaluated from elapse time using 176 CPUs and 1080 CPUs for Pt_{135} after tuning to the ES of the LCAO-PS code by the hybrid type parallelization.

3. Nano-size clusters

Figure 2 indicates a typical number of atoms of the nano-size cluster. Nanostructure of about 2.4 nm in the fcc structure is the critical size that the volume of surface region is comparable with the bulk region in which each atom has full nearest neighbors. Larger size clusters than this critical scale will have surface region similar to the surface of the bulk crystal which should be dealt with the periodic simulation code. Since experimental researchers have been explor-

ing about 2 nm size Pt clusters to obtain most efficient catalytic properties empirically, they may need to pay attention to nanostructure morphology. Calculated nanostructure databases will be very useful in materials design in nanotechnology. We calculated Pt and Au clusters as a typical example of nanostructure material as shown in Figures 3 and 4. Electronic structures of Pt_{135} are very different from those of Au_{135}; Pt cluster is almost metallic while Au one has large HOMO-LUMO gap. The first principles calculation of Pt systems is one of the most difficult ones in terms of the self-consistent calculation. The ES provides us results almost 100 times faster than those from standard laboratories with normal computer facilities. There exists another symmetry problem in nanostructure materials. Figure 5 shows total energy curves of Pt_{55} clusters. In small clusters, metal clusters prefer to have an icosahedral symmetry. The cluster morphology becomes very important in nanostructure materials.
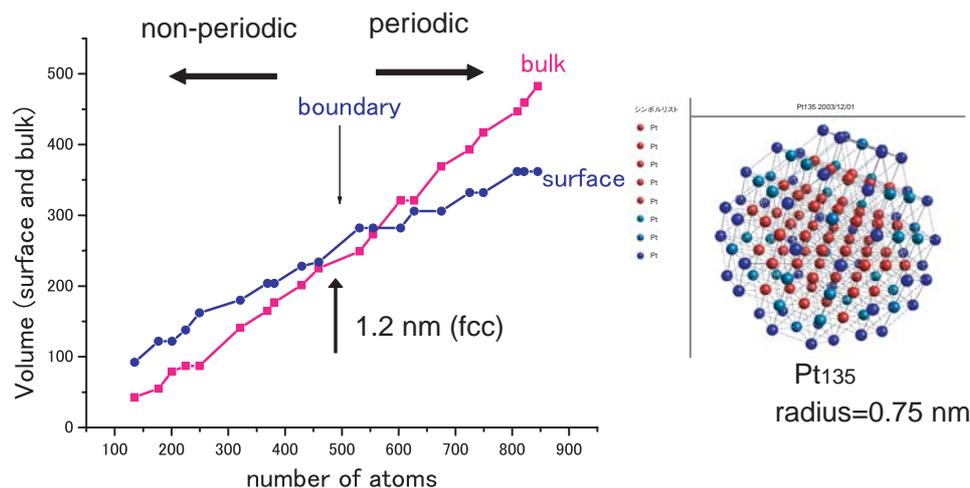
Fig. 2 Critical scale of nanostructure of fcc cluster. 1.2 nm is the radius of the cluster for the boundary.
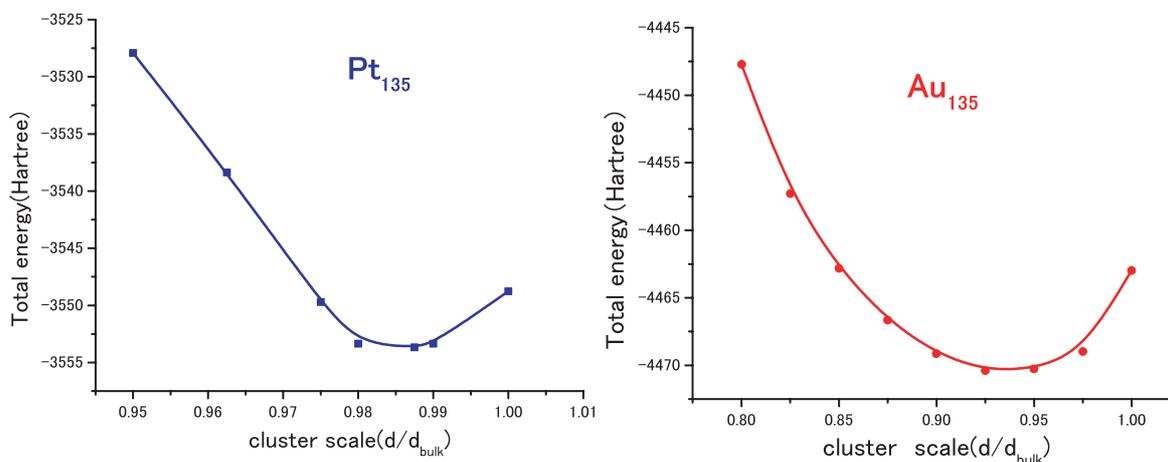
Fig. 3 Optimized cluster sizes of Pt_{135}, 1.5% contraction, and Au_{135}, 7% contraction, of fcc nanosturucture.
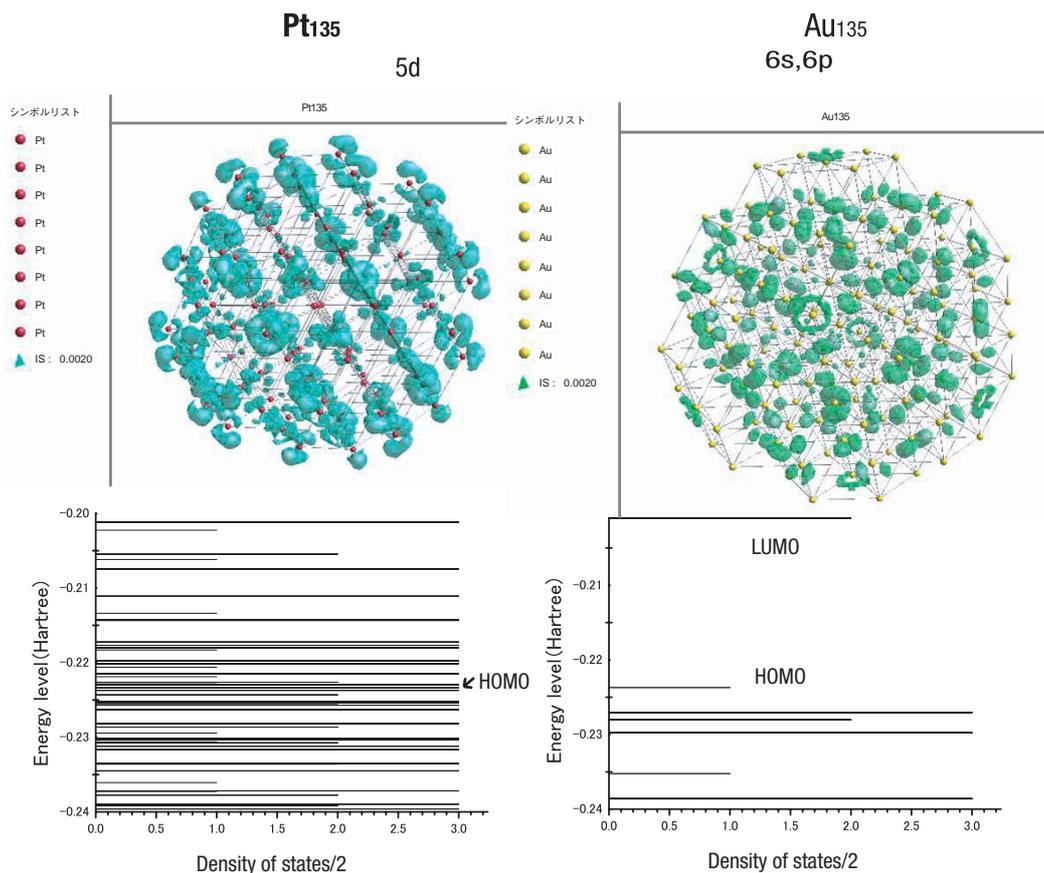
Fig. 4 Energy levels and charge density distribution of highest occupied molecular orbital (HOMO) of Pt_{135} and Au_{135}. HOMO-LUMO energy gaps are less than 0.01 eV for Pt_{135} and 0.6 eV for Au_{135}. HOMO states comprise 5d states for Pt_{135} and 6s and 6p states for Au_{135}.
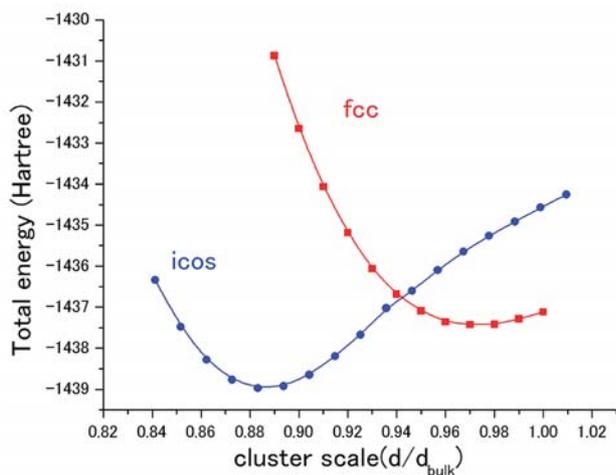
Fig. 5 Energy curves of Pt_{55} clusters of fcc and icosahedral symmetry. d_{bulk} in the icosahedron is given by the nearest neighbor interatomic distance of twelve vertices.

4. Future Work

In this year, we have mainly focused on optimizing the LCAO-PS code in the ES in order to have high parallel performance. The main difficulty was the memory restriction in the single node. We, therefore, will develop a new numerical scheme to adapt the present situation to be able to deal with larger clusters of various kinds, like oxides, alloys, metal in water molecules, and metals with hydrogen etc.. As for the periodic system, we will develop a parallel code in the ES. Calculated databases will be processed to produce transferable databases.

154