Annual Report
FY2011

Research Division of AICS

October 2012

RIKEN Advanced Institute for Computational Science

Contents

I. Mission and Overview

Japan's Third Science and Technology Basic Plan (a Cabinet resolution adopted in March 2006) designated next-generation supercomputing as a key technology of national importance that should be the focus of investments in gestation periods for large-scale national projects. As a result, RIKEN along with the Ministry of Education, Culture, Sports, Science and Technology (MEXT) are putting full efforts into the development and advancement of such technology.

Under the guidance of MEXT and on the basis of the "Law on Joint Use Promotion of Large-Scale Advanced Research Facilities" (effective July 2006), RIKEN has been managing the construction of the Next-Generation Supercomputer, the "K computer." The K computer is to be a shared resource based on the above law and, with its wide accessibility, is expected to be put to a broad range of uses—from diverse fields of basic science and technology research to applications in industry. It was with this in mind that RIKEN established the Advanced Institute for Computational Science (AICS) on July 1, 2010. Through the use of its world-class supercomputer boasting 10 petaflops of computational power, and through collaboration and integration of the fields of computational science and computer science, AICS strives to create an international center of excellence dedicated to generating world-leading results. AICS's missions are:

1. Operating the K computer efficiently for users of wide research areas as well as of industries.
2. Carrying out the leading edge research of computational science and technology, and establishing itself as the COE of computational sciences in Japan.
3. Proposing the future directions of HPC in Japan.

And also, AICS's missions include:

- Promoting strong collaborations (or co-development) between computational and computer scientists, working with the core-organizations of the strategic areas identified by the Strategic Programs for Innovative Research (SPIRE),
- Raising and promoting young scientists who are strong in both computational and computer sciences
- Pursuing new concepts for the future HPC beyond the petascale (including exascale)

The research teams in Research Division of AICS are carrying out the leading edge research of computational science and technology, and also expected to provide useful software to the users of the K computer.

Today's supercomputers including the K computer employ hundreds of thousands of cores which work simultaneously and in parallel to rapidly calculate and process enormous volumes of data.

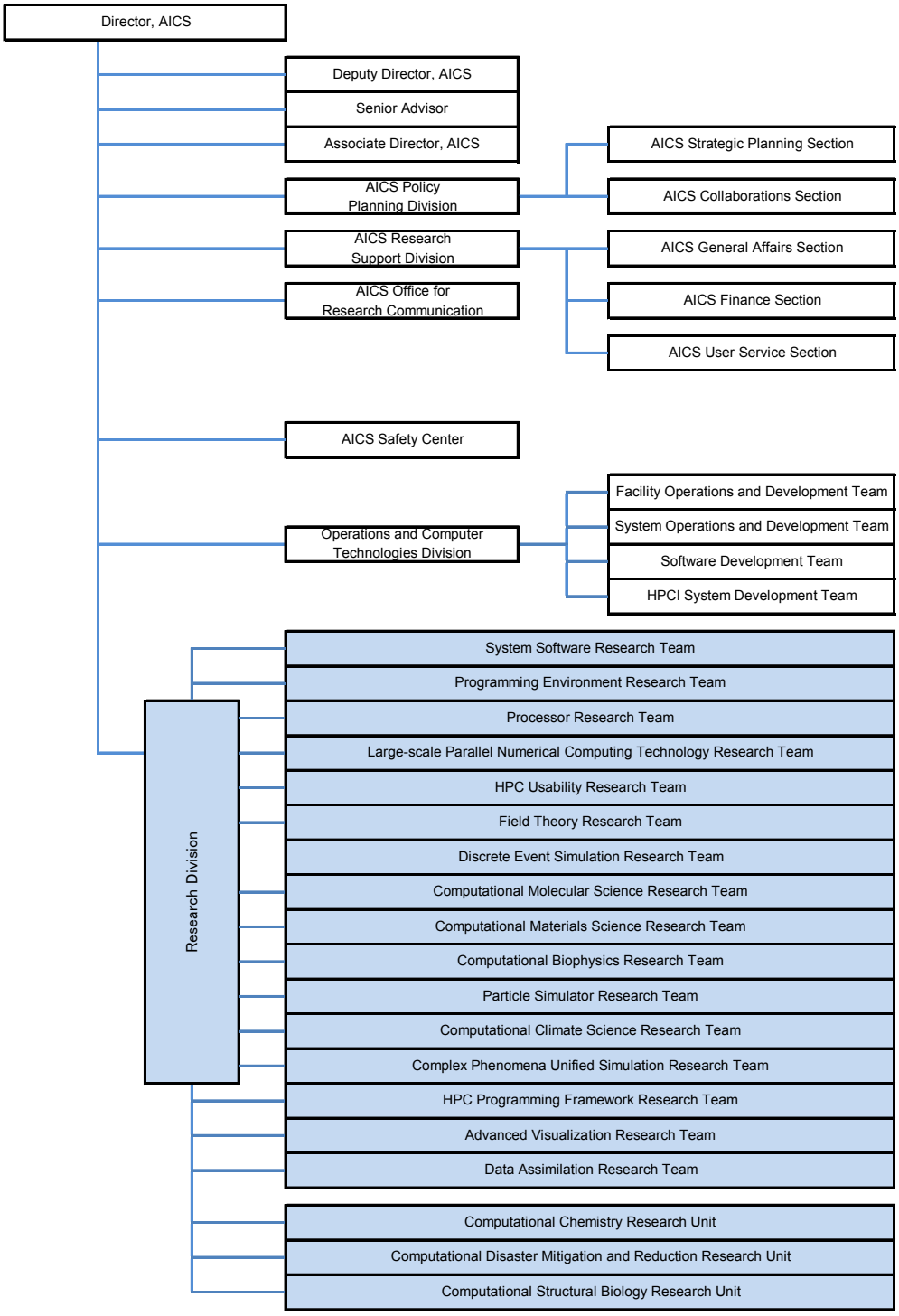
Getting a supercomputer to work at its full potential, however, requires extremely sophisticated technology.

Achieving this kind of technology is impossible without close collaboration between computational and computer science. Computational science involves the application of new methodologies of large-scale numerical simulation for research and development in physics, chemistry, biology, medical science, and engineering. Computer science, on the other hand, focuses on computer systems architecture, the algorithms that lay out the methods and procedures for problem-solving, management of computational and memory resources, systems software that allows communication among different computers, and application programs.

The AICS Research Division brings together researchers specializing in computational and computer science to merge and develop the two fields into what may eventually become a whole new field of interdisciplinary computational science that will maximize the potential of the supercomputer for major breakthroughs in science and technology.

II. Organization

The organization of AICS (as of October 2012) is schematically shown below. The research teams in computer science and computational sciences are closely integrated into the research division of AICS.



III. Reports on Research Activities

1. System Software Research Team

1.1. Team members

Yutaka Ishikawa (Team Leader)
Atsushi Hori (Researcher)
Keiji Yamamoto (Postdoctoral Researcher)
Yoshiyuki Ohno (Research Associate)
Toshihiro Konda (Research Associate)
Toyohisa Kameyama (Technical Staff)

1.2. Research Activities

The system software team focuses on the research and development of an advanced system software stack not only for the "K" computer but also for towards exa-scale computing. There are several issues in carrying out future computing. Two research categories are taken into account: i) scalable high performance libraries/middleware, such as file I/O and low-latency communication, and ii) a scalable cache-aware, power-aware, and fault-aware operating system for next-generation supercomputers based on many core architectures.

Parallel file I/O is one of the scalability issues in modern supercomputers. One of the reasons is due to heavy metadata accesses. If all processes create and write different files, the metadata server receives so many requests by all processes not only at the creation time but also at writing data to each file. Three approaches have been conducted to mitigate this issue. One approach is to introduce a file composition technique that gathers multiple data generated by an application and stores these data into one or a few files in order to reduce the number of files accessed by processes. Another approach is to provide multiple metadata server in which the requests for metadata are sent to a metadata server resolved using hash function. The third approach is to provide a smart MPI-IO implementation for applications using MPI-IO functions.

Increasing number of cores and nodes enforces strong scaling on parallel applications. Because the ratio of communication time against local computation time increases, a facility of low-latency and true overlapping communication and computation communication is desired. A communication library, integrated to the MPI library implementation in K computer, has been designed and implemented, that utilizes DMA engines of K computer. Each compute node of K computer has four DMA engines to transfer data to other nodes. If a communication library knows communication patterns in advanced, it may utilize the DMA engines. Indeed, the feature of MPI persistent communication, standardized in MPI-1.0, allows the runtime library to optimize data transfers

involved in the persistent communication using the DMA engines.

System software stack developed by our team is designed not only for special dedicated supercomputers, but also for commodity-based cluster systems used in research laboratories. The system will be expected to be used as a research vehicle for developing an exa-scale supercomputer system.

1.3. Research Results and Achievements

1.3.1. Scalable File I/O

- File Composition Library

  A file composition technique has been proposed, where lots of file data generated by an application are gathered and stored into one file or a few files. Figure 1 shows the basic concept of the file composition technique. File composition technique composes data, which looks like a file from the viewpoint of the application, into an aggregated larger file. Unlike the existing aggregation mechanisms, the proposed technique follow POSIX file I/O semantics, thus no modification of application programs are required.

  In FY2011, we developed a prototype of file composition library and evaluated the basic performance. Figure 2 shows that the file composition technique is three times faster than POSIX file I/O functions in the case that each of 128 processes creates an individual file.

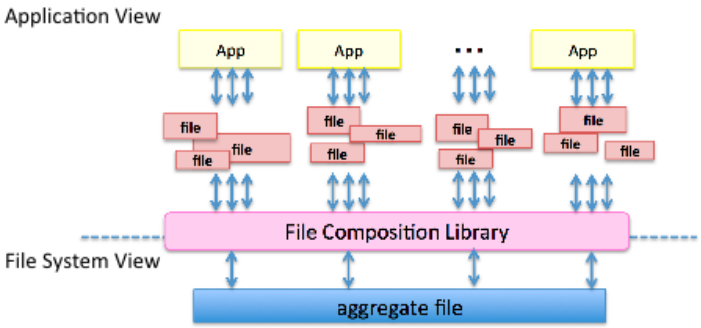


Figure 1. Concept of the file composition technique

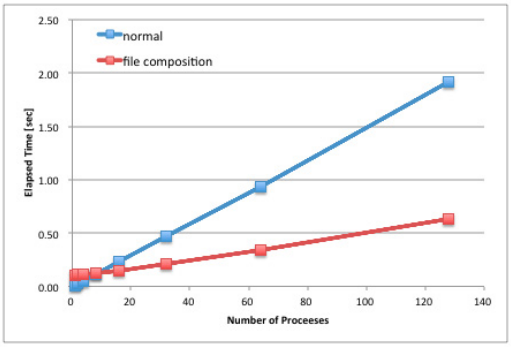


Figure 2. Time of 1 MB file creation by parallel process

- Hash-based Parallel File System

  A parallel file system used in a modern supercomputer is basically composed of MDS (Metadata Server) and OSS (Object Storage Server). A MDS handles file open, create and status operations. An OSS handles file read and write operations. Most current parallel file systems have only one MDS, and thus, such a system causes bottleneck of metadata operations requested by all compute nodes. A new scalable parallel file system based on a hash function was designed and implemented in FY2011. This system consists of multiple MDSs and OSSs as shown in Figure 3. Each MDS is responsible for metadata operations on a part of all files. The metadata set of files is determined by a hash value of the file name with path. The client determines the MDS of a file accessed in the client by a hash value of that file and path, and metadata operations for that file are sent to the MDS. The MDS informs the client to the location of OSSs for write/read operations in the client.

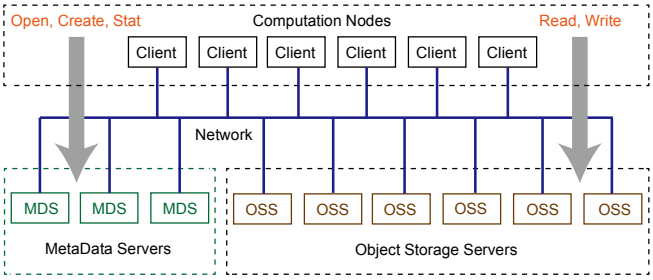


Figure 3. Hash-based Parallel File System

In FY2011, we developed only MDS for prototype file system and evaluated a metadata performance. Evaluations show that throughput of proposed file system is faster than that of current Lustre File System and has good scalability as shown in Figure 4.

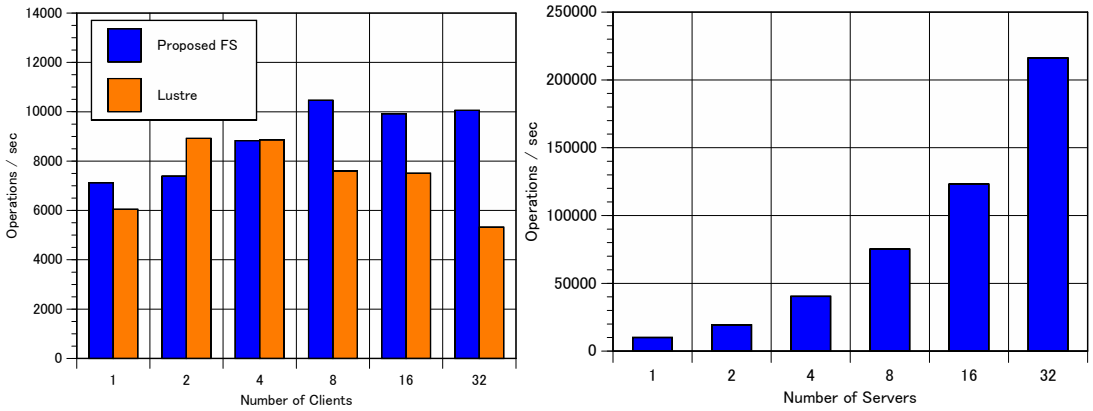


Figure 4. Throughput of the mkdir command execution

- A Smart MPI-IO Library Implementation

  The nature of highly parallelized parallel file access that consists of lots of fine grain, non-contiguous I/O requests in many cases, can degrade the I/O performance severely. To tackle this problem, a novel technique to maximize the bandwidth of the MPI-IO was proposed. This technique utilizes a ring communication topology and was implemented as an ADIO device of ROMIO, named Catwalk-ROMIO, and evaluated. The evaluation shows that Catwalk-ROMIO utilizing only one disk can exhibit comparable performance with widely-used parallel file systems, PVFS2 and Lustre, both of them utilizes several file servers and disks. As shown in Figure 5, Catwalk-ROMIO performance is almost independent from file access patterns, in contrast to the performance of parallel file systems performing only well with collective I/O shown in “Full” cases of Figure 5. Catwalk-ROMIO requires conventional TCP/IP network and only one file server. This is a quite common HPC cluster configuration. Thus, Catwalk-ROMIO is considered to be a very cost-effective MPI-IO implementation.

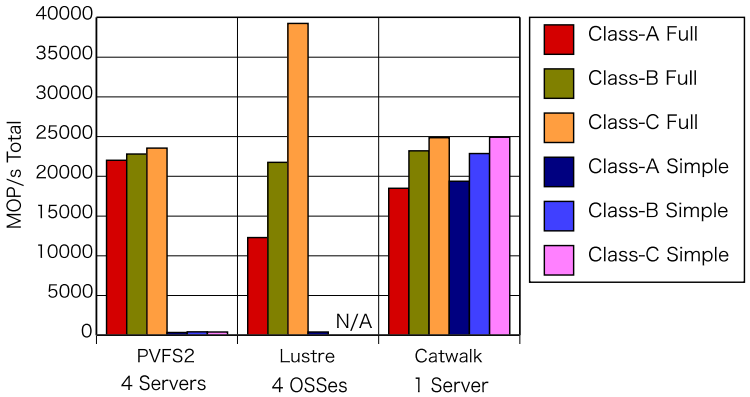


Figure 5. BT-IO Performance with PVFS2, Lustre and proposed Catwalk-ROMIO

1.3.2. Communication Library

- Persistent Remote DMA

  The implementation of persistent communication provided in MPI was reconsidered to provide low latency and true overlapping communication and computation. In the persistent communication facility, the end-points of both the sender and the receiver are set up by issuing MPI_Send_init and MPI_Send_recv primitives prior to actual communication triggered by the MPI_Start or MPI_Startall primitive. The same communication pattern is reused without reissuing the initialization. Thus, at the start of actual communications in persistent communication, the runtime system already knows all the communication patterns, i.e., peers and message sizes if both sender and receiver have issued persistent communication primitives. Such situations have a chance for the communication library to utilize four DMA engines equipped in K computer and carry out true overlapping communication and computation.

  A new communication protocol and an implementation for persistent communication, called PRDMA (Persistent Remote Direct Memory Access), was designed and implemented in FY 2011.

- Communication Library

  The power wall issue of current and future supercomputers is gathering attentions. The technique of user-level communication to achieve high communication performance is widely used by parallel applications, and processes are spinning-wait for the incoming messages. This spinning loop in the absence of incoming messages is simply wasting energy, and thus it increases the power consumption of a parallel computer. The proposed technique is implemented in two ways; 1) combination of spin-loop and blocking system call, and 2) combination of spin-loop and using the Intel x86 *monitor*/*mwait* synchronization instructions which put computational core into a low-power mode. Unlike the techniques using the DVFS (Dynamic Voltage and Frequency Scaling) function of CPU, our proposed technique does not sacrifice application performance but can save energy. Figure 6 shows that 7% total system power can be saved with the FT application of NAS parallel benchmarks.

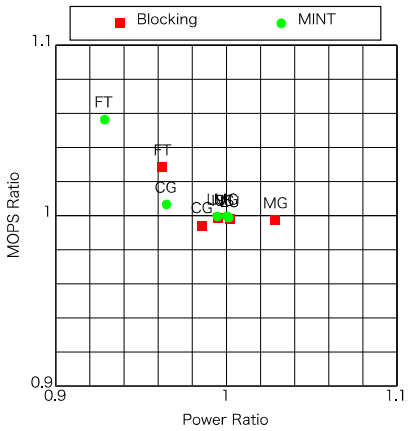


Figure 6. Power and Performance, NAS Parallel Benchmark Class B

1.3.3. Multi-Threaded Library

Towards the Exa-scale computing, hiding the latencies of memory and communication is one of crucial issues. To provide such a capability, the thread management must be fast enough in the order of sub-micro seconds. The thread library, named **Shadow Thread**, is developed to utilize Simultaneous Multi-Threading mechanism which schedules threads by hardware in a very fast way, and utilizes the *monitor* and *mwait* instructions supported by some Intel processors. Figure 7 shows that the two-phase synchronization technique combining the conventional spin-wait method and the pair of the *monitor*/*mwait* instructions can satisfy the requirement of speed and low-power consumption simultaneously. Figure 8 shows that a memory copy function using the proposed Shadow Thread library can exhibit better performance up to 20% compared with the normal *memcpy()* function.

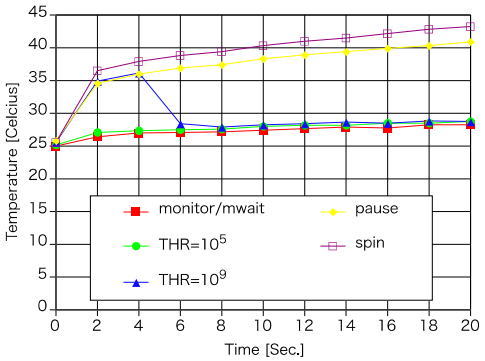


Figure 7. CPU temperature over time while synchronization

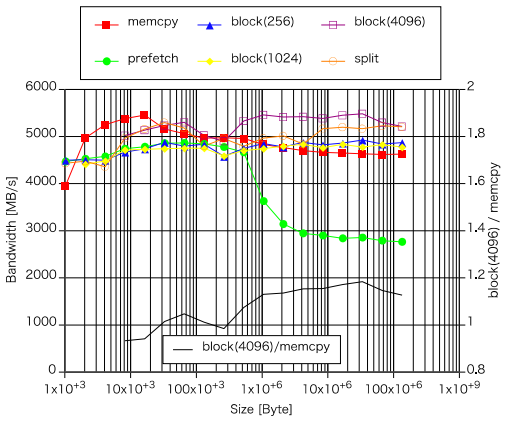


Figure 8. Multi-threaded memcpy bandwidth over region size

1.4. Schedule and Future Plan

The prototype implementation of File I/O and Communication libraries/middleware will be deployed to K computer in FY2012. Due to limited functionalities, the users will be restricted. The enhanced version of those libraries/middleware will be developed and deployed in FY 2013. In FY2013, those libraries/middleware will be deployed for users.

A scalable cache-aware, power-aware, and fault-aware operating system for next-generation supercomputers based on many core architectures are being designed and implemented by collaboration with University of Tokyo, NEC, Hitachi, and Fujitsu. The first prototype system will be distributed in FY2012 and the basic system will be available in FY2013.

1.5. Publication, Presentation and Deliverables

(1) Journal Papers

1. Atsushi Hori, Jinpil Lee, Mitsuhisa Sato, “Audit: A new synchronization API for the GET/PUT protocol”, In Journal of Parallel and Distributed Computing, 2012.

(2) Conference Papers

1. Atsushi Hori, Jinpil Lee, Mitsuhisa Sato, “Audit: New Synchronization for the GET/PUT Protocol,” In the 1st Workshop on Communication Architecture for Scalable Systems, 2011.
2. Atsushi Hori, Keiji Yamamoto, Yutaka Ishikawa, “Catwalk-ROMIO: A Cost-Effective MPI-IO,” In Proceedings of the 2011 IEEE 17th International Conference on Parallel and Distributed Systems, IEEE Computer Society, 2011.
3. Keiji Yamamoto, Atsushi Hori, Shinji Sumimoto, Yutaka Ishikawa, “Xruntime: A Seamless Runtime Environment for High Performance Computing,” In the 2011 International Workshop on Extreme Scale Computing Application Enablement - Modeling and Tools, 2011.

4. Atsushi Hori, Toyohisa Kameyama, Mitarou Namiki, Yuichi Tsujita, Yutaka Ishikawa, “Low Energy Consumption MPI Using Hardware Synchronization,” In IPSJ-SIGHPC 2011-HPC-132(7), 2011. (In Japanese)
5. Atsushi Hori, Keiji Yamamoto, Yoshiyuki Ohno, Toshihiro Konda, Toyohisa Kameyama, Yutaka Ishikawa, “A Ultra-light Thread Library Using Hardware Synchronization,” In IPSJ-SIGHPC 2011-HPC-130(6), 2011. (In Japanese)
6. Yoshiyuki Ohno, Atsushi Hori, Yutaka Ishikawa, “Proposal and preliminary evaluation of a mechanism for file I/O aggregation to one file in a parallel job,” In IPSJ-SIGHPC 2011-HPC-132(34), 2011. (In Japanese)

(3) Invited Talks

1. Yutaka Ishikawa, “HPCI and SDHPC: Towards Sustainable High Performance Computing in JAPAN,” PRAGMA21 Workshop, October, 2011.

(4) Posters and presentations

1. Atsushi Hori, Yutaka Ishikawa, “MINT: a fast and green synchronization technique,” In Proceedings of the 2011 companion on High Performance Computing Networking, Storage and Analysis Companion, ACM, 2011.
2. Keiji Yamamoto, Yoshiyuki Ohno, Atsushi Hori, Yutaka Ishikawa, “Hash-based Metadata Management for Parallel File System,” Symposium on High Performance Computing and Computational Science, 2012.
3. Keiji Yamamoto, Yoshiyuki Ohno, Atsushi Hori, Yutaka Ishikawa , “Current Issue in Parallel File System,” The 4th Forum on Data Engineering and Information Management, 2012.

2. Programming Environment Research Team

2.1. Team members

Mitsuhisa Sato (Team Leader)
Hitoshi Murai (Research Scientist)
Tetsuya Abe (Postdoctoral Researcher)
Swann Perarnau (Postdoctoral Researcher)
Tomotake Nakamura (Research Associate)
Takenori Shimosaka (Research Associate)
Masahiro Yasugi (Visiting Researcher)
Tomio Kamada (Visiting Researcher)
Tomoko Nakashima (Assistant (Concurrent))

2.2. Research Activities

The K computer system running in AICS is a massively parallel system which has a huge number of processors connected by the high-speed network. In order to exploit full potential computing power to carry out advanced computational science, efficient parallel programming is required to coordinate these processors to perform scientific computing. We conducts researches and developments on parallel programming models and language to exploit full potentials of large-scale parallelism in the K computer and increase productivity of parallel programming.

In 2011FY, in order to archive these objectives above, we carried out the following researches:

1) We have been developing a new programming language, called XcalableMP(XMP), which is originally designed in Japanese HPC language research community. Although MPI is the de facto standard for parallel programming on distributed memory systems, writing MPI programs is often a time-consuming and complicated process. XcalableMP is a directive-based language extension which allows users to develop parallel programs for distributed memory systems easily and to tune the performance by having minimal and simple notations. The specification has been designed by XcalableMP Specification Working Group which consists of members from academia and research labs to industries in Japan. In this year, we developed a prototype compiler of XMP Fortran, and deployed it to the K computer. We did a preliminary performance evaluation using a XMP version of the SCALEp code (a climate code for LES) on the K computer. We also design the interface to MPI programs in XMP, and IO supports of the XMP language. We have ported the GASnet communication library developed by LBNL/UCB for the K computer as an one-side communication layer of XMP.
2) We started the research for performance tuning tools for large-scale scientific applications running on the K computer. As a first step, we have ported the Scalasca performance turning

and analysis tool developed by JSC, Germany, to the K computer and examined the potential of this tool for the K computer.

3) We investigated methods and tools to support a correct parallel program. We proposed a light-weight XMP verification tool which helps users to verify XMP programs using descriptions of global-view programming directives in XMP.
4) The processor of the K computer has an interesting hardware mechanism called “sector cache”, which allows partition of L2 on-chip cache to optimize the locality for important data. We investigated the technique to optimize the usage of sector cache.
5) Our team supports Xcrypt, a scripting language developed by Kyoto University. We have developed an application using Xcrypt which perform an automatic tuning for Block Multi-Color Ordering Code.

We have started discussion with application developers in AICS for the collaborations on performance tuning and development of parallel programs.

2.3. Research Results and Achievements

2.3.1. Development of XcalableMP Fortran and Preliminary Performance on the K computer

In this year, we developed a prototype compiler of XMP Fortran, and deployed it to the K computer. We did a preliminary performance evaluation of XMP Fortran on the K computer. We parallelized SCALEp code by XMP Fortran. SCALEp is a kinetic core in SCALE developed by the SCALE project which develops (Parallel) Climate code for large eddy simulation under the collaboration with Climate science research team and computer science research teams. The parallelization steps of SCALEp codes are as follows:

1. Insert pragmas to specify 2D block distribution of 3D array.
2. Paralleling double nested loop by loop directives
3. Insert reflect directives for the communication periodic neighbor elements.
4. As a, we have examined the performance improvement by runtime optimization using Remote Direct Memory Access (RDMA) of K computer for neighbor communications.

Figure 1 shows the outline of XMP version of SCALEp code.

```
!$xmp nodes p(N1,N2)
!$xmp template t(IA,JA)
!$xmp distribute t(block,block) onto p

real(8) :: dens(0:KA,IA,JA)
!$xmp align (*,i,j) &
!$xmp    with t(i,j) :: dens, ...
!$xmp shadow (0,2,2) :: dens, ...

!$xmp reflect (dens(0,/periodic/2,&
!$xmp                /periodic/2), ...)

!$xmp loop (ix,jy) on t(ix,jy)
do jy = JS, JE
  do ix = IS, IE
    do kz = KS+2, KE-2
      ... dens(kz,ix+1,jy) ...
    end do
  end do
end do
```

Declarations for Node array and template

Data distribution

Neighbor comm

Loop paralization

Figure 1. The outline of SCALEp in XMP

The problem size is 512x512 in horizontal directions, 128 in vertical direction. In Figure 2, the execution time for 500 steps is shown. Note that XMP node is assigned to one node and the node program is parallelized by automatic paralleling compiler by Fujitsu.

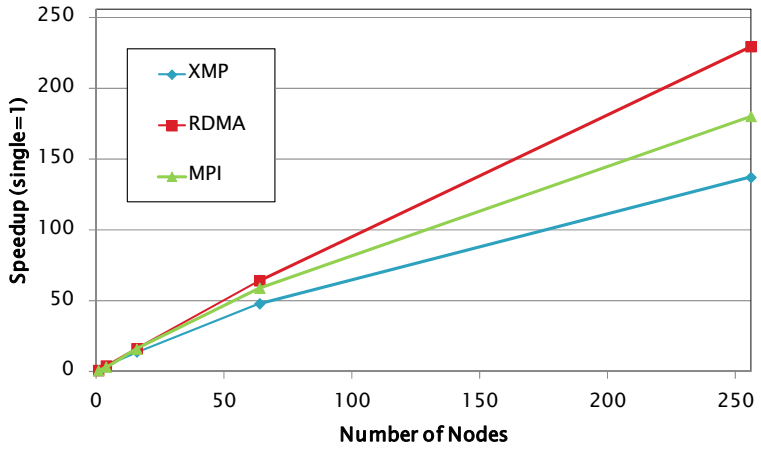


Figure 2. Performance of SCALEp code on the K computer.

We found some overhead of XMP version comparing to the hand-coded MPI version, and it can be improved significantly by the optimization using RDMA.

2.3.2. Design of the Interface of XcalableMP Programs to Parallel Numerical Libraries

Scientific parallel applications are often developed using high-performance parallel libraries. In development of scientific applications, re-writing every application programs using only XMP is unrealistic. In order to develop high performance parallel application programs easily, it is very

important to make use of existing parallel high performance libraries. We have designed an interface of XMP to parallel numerical libraries, which has the following features:

- Parallel numerical library routines can be invoked as an XMP procedure through the XMP interface procedure.
- When the parallel numerical library routine needs information on global distributed arrays, the interface extracts it from the descriptor using some query routines provided by XMP runtime library and passes it to the numerical library routine as arguments.
- The interface does not affect the behavior of numerical library routines expect for restrictions concerning the XMP specification.

We have implemented the interface between XMP and each of ScaLAPACK (Scalable Linear Algebra PACKage http://www.netlib.org/scalapack/) and MUMPS (A MUltifrontal Massively Parallel sparse direct Solver http://graal.ens-lyon.fr/MUMPS/). The example code of a XMP program using the interface to ScaLAPCK is shown in Figure 3.

```
#include “xmp.h”
int nrow=NROW,ncol=NCOL,nprow=NPROW,npcol=NPCOL;
double a[NCOL][NROW],b[NROW],ipiv[2*NROW][NPCOL];
#pragma xmp nodes p(npcol,nprow)
#pragma xmp nodes q(nprow)=p(1,1:nprow)
#pragma xmp template t(0:ncol-1,0:nrow-1)
#pragma xmp template t1(0:2*nrow-1,0:npcol-1)
#pragma xmp distribute t(block,block) onto p
#pragma xmp distribute t1(block,block) onto p
#pragma xmp align a[i][j] with t(i,j)
#pragma xmp align ipiv[i][j] with t1(i,j)
#pragma xmp align b[i] with t(*,i)

int main(int argc, char* argv[]){

  int i,j,contxt,myrow,mycol;
  int icontxt=-1,what=0;
  int nrhs=1,ia=1,ja=1,ib=1,jb=1,info;
  double a0[ncol][nrow],b0[nrow],btmp,err;
  char *order="R";
  blacs_get_(&icontxt,&what,&contxt);
  blacs_gridinit_(&contxt,order,&nprow,&npcol);
  blacs_gridinfo_(&contxt,&nprow,&npcol,&myrow,&mycol);

  for(i=0;i<ncol;i++){
    for(j=0;j<nrow;j++){
      a0[i][j] = rand()/(1.0e+10);
    }
  }

 for(j=0;j<nrow;j++){
    b0[j]=1.0;
  }

#pragma xmp loop (i,j) on t(i,j)
  for(i=0;i<ncol;i++){
    for(j=0;j<nrow;j++){
      a[i][j] = a0[i][j];
    }
  }

#pragma xmp loop on t(*,j)
  for(j=0;j<nrow;j++){
     b[j]=b0[j];
  }

ixmp_pdgesv(&nrow,&nrhs,a,&ia,&ja,xmp_desc_of(a),
ipiv,b,&ib,&jb,xmp_desc_of(b),&contxt,&info);
……
```

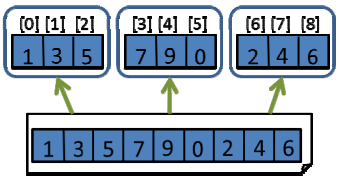
Figure 3. Example code of using the XMP interface for ScaLAPACK

2.3.3. Parallel I/O Facility of XcalableMP and Evaluation of I/O Performance

XcalableMP allows the programmer to define a distributed array over nodes to support typical data parallel programmer with global-view programming. In order to perform I/O operations of a distributed array efficiently, XMP-IO is defined a parallel I/O API for a distributed global array in the XcalableMP specification.

In a large-scale parallel system, some applications must often perform I/O for a large data efficiently. These I/O intensive operations may cause a serious problem in a large scale parallel system. To solve the I/O problem, parallel IO libraries such as MPI-IO and pNetCDF has been developed.

We have designed the XMP-IO using MPI-IO as an underlying parallel I/O library. The XMP interface to obtain the information of the distributed array described in the program is used to perform parallel I/O operations by MPI-IO. As other parallel I/O interface, we also designed a XMP library interface to a parallel library for PNetCDF which is a commonly used file format. Figure 4 shows an example code using XMP-IO and XMP pNetCDF interface and these MPI counterparts.



(a) File IO example (upper: node, lower: file)

XMP program

```
int a[9];
#pragma xmp nodes p(3)
#pragma xmp template t(0:8)
#pragma xmp distribute t(block) onto p
#pragma xmp align a[i] with t(i)

xmp_file_t  fp;     // file handle
xmp_range_t  ra;   // range of array: ra
xmp_desc_t  desc;  // info of the specified array
desc = xmp_desc_of(a);
xmp_set_range(ra, 0, 0, 3, 1);

fp = xmp_fopen_all(fname, “r”);
xmp_file_set_view_all(fp, 0, desc, ra);
xmp_fread_darray_all(fp, desc, ra);
xmp_fclose_all(fp);
```

Translate

MPI program

```
int a[3];
gsize[0] = 9;
dist[0] = MPI_DISTRIBUTE_BLOCK;
darg[0] = MPI_DISTRIBUTE_DFLT_DARG;
psize[0] = 3;

MPI_File_open(comm, fname, mode, info, &fh);
MPI_Type_create_darray(3, rank, 1, gsize, dist,  psize,
MPI_ORDER_C, MPI_INTEGER, &type);
MPI_Type_commit(&type);
MPI_File_set_view(fh, 0, MPI_INTEGER, type, “native”,
 info);
MPI_File_read_all(fh, a, 3, MPI_INTEGER, &status);
MPI_File_close(&fh);
```

(b) XMP-IO and equivalent MPI program

XMP program using our library interface

```
int a[9];
xmp_desc_t  desc;
desc = xmp_desc_of(a);

xmpnc_open(comm, fname, NC_READ, info, &id);
xmpnc_inq_varid(fid, “data”, &vid);
xmpnc_get_varaw_int_all(id, vid, desc);
xmpnc_close(id);
```

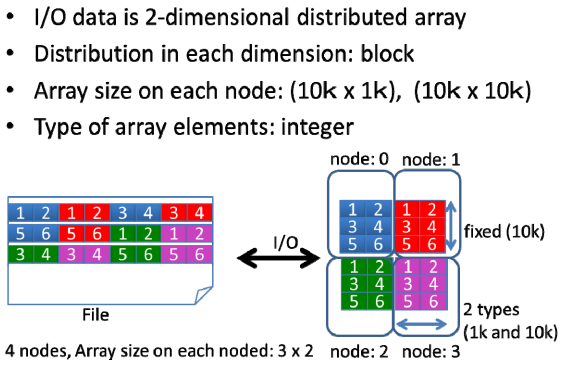
Translate

MPI program using PnetCDF

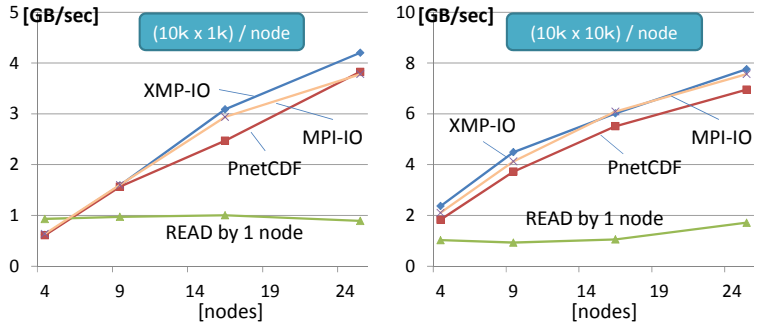
```
int a[3];

ncmpi_open(comm, fname, NC_READ, info, &id);
ncmpi_inq_varid(ncid, “data”, &vid);
start[0] = rank * 3;
count[0] = 3;
ncmpi_put_vara_int(id, vid, start, count, &a[0]);
ncmpi_close(id);
```

(c) XMP pNetCDF interface and equivalent MPI with PNetCDF call

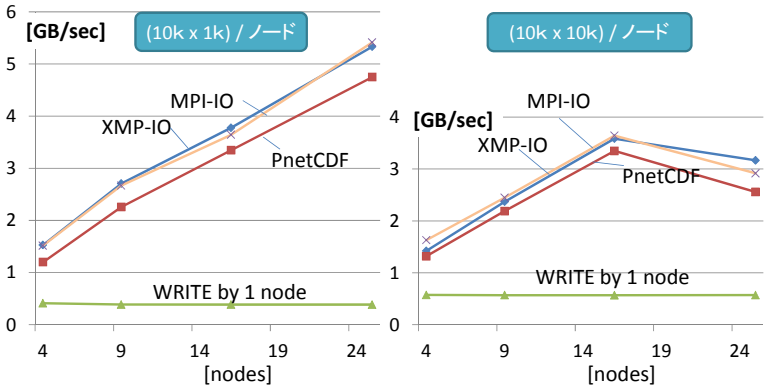
Figure 4. XMP-IO and XMP netCDF



(a) Experimental Setting



(b) Read performance



(c) Write performance

Figure 5. Performance results of XMP-IO and XMP pNetCDF interface

In Figure 5, the experimental results of these parallel I/O in XcalableMP are shown. We measured the IO performance for data with size 10kB x 1KB and 10kB and 10kB respectively. We found that these parallel I/O interfaces provide a good parallel I/O performance on the Lustre file system with reasonable programming cost.

2.4. Schedule and Future Plan

The major goal of researches in next financial year (2012) is to finish the full implementation of XcalableMP C and Fortran for the K computer, and release them to the users. With this release, we will provide XMP-IO and XMP interface to some high-performance parallel libraries written in MPI.

As for the performance tuning tool for the K computer, we will perform several case studies on real scientific applications using Scalasca as well as providing Scalasca to general users. Through these case studies, we will extend it for valuable performance analysis in the K computer.

We are also continuing the researches on the optimization for “sector cache” and parallel program verification tools for XMP program and MPI programs.

2.5. Publication, Presentation and Deliverables

(1) Journal Papers

- None

(2) Conference Papers

1. Yonezawa, Tadashi Watanabe, Mitsuo Yokokawa, Mitsuhisa Sato, and Kimihiko Hirao Advanced Institute for Computational Science (AICS), "Japanese National High-Performance Computing Research Institute and its 10-petaflops supercomputer "K"", Proceeding of SC11, SC '11 State of the Practice Reports , 2011.

[not refereed, in Japanese]

1. Tomotake Nakamura and Mitsuhisa Sato, "Parallel I/O Facility of PGAS Programming Language XcalableMP and Evaluation of I/O Performance on the Lustre File System", IPSJ SIG Technical Report (in Japanese), Vol. 2011-HPC-132 No. 36, pp. 1-9, 2011.
2. Hitoshi Murai and Mitsuhisa Sato, "Extensions of XcalableMP for User-Defined Distribution of Data and Loops", IPSJ SIG Technical Report (in Japanese), Vol. 2011-HPC-130, No. 59, pp. 1-9, 2011.
3. T. Shimosaka, H. Murai and M. Sato, "A Design of MPI Parallel Library Interface of Parallel Programming Language XcalableMP", IPSJ SIG Technical Reports (in Japanese), Vol.2011-HPC-130 (55), pp.1--8, July, 2011.
4. T. Abe, T. Hiraishi, Y. Miyake, T. Iwashita and H. Nakashima, "Job-level parallel executions for satisfying distributed constraints". Summer United Workshops on Parallel, Distributed and Cooperative Processing, IPSJ SIG Technical Reports (in Japanese), 2011-HPC-130(59), pp.1--8, July, 2011.
5. T. Abe and M. Sato, "Directive-based source code checking for verification-oblivious programming in high performance computing" (in Japanese). Summer Programming Symposium, pp.1--6, September, 2011.

(3) Invited Talks (From April 2011 to March 2012)

1. Mitsuhisa Sato, "The K Computer Project and Research on Parallel Programming Languages", PGAS 2011: Fifth Conference on Partitioned Global Address Space Programming Models, October 18, 2011, Galveston Island, Texas, USA.
2. Mitsuhisa Sato, "Challenges of programming environment and tools for peta-scale computers -- programming environment researches for the K computer ---," Invited Presentation, 4th Workshop on Productivity and Performance (PROPER 2011) Tools for HPC Application Development EuroPar 2011 Conference, Bordeaux/France, August 30th 2011.
3. Mitsuhisa Sato, "Perspective of HPC Development for Computational Science in Japan", HPC in Asia Workshop, ISC 2011.
4. Mitsuhisa Sato, "The K Computer and XcalableMP Parallel Language Project: Towards a Programming Environment for Peta-scale Computing", Joint ENCORE & PEPPHER Workshop on Programmability and Portability for Emerging Architectures (EPoPPEA), Jan 24, 2012, Paris, France. (in conjunction with the HiPEAC'12 conference)

(4) Posters and presentations

1. Takenori Shimosaka, Hitoshi Murai, Mitsuhisa Sato, “A Design of the Interface of XcalableMP Programs to Parallel Numerical Libraries”, the 2nd AICS International Symposium, Poster, March 2012.
2. Tomotake NAKAMURA and Mitsuhisa SATO, “Paralle I/O Facility of PGAS Programming Language XcalableMP and Evaluation of I/O Performance on Parallel File System Lustre”, the 2nd AICS International Symposium, Poster, March 2012.

(5) Patents and Deliverables

1. XcalableMP compiler ver. 0.5 for the K computer
2. Scalasca performance analysis tool for the K computer (test version)
3. GASnet one-sided communication Library for the K computer (test version)

3. Processor Research Team

3.1. Team members

Makoto Taiji (Team Leader)

Gentaro Morimoto (Research Scientist)

3.2. Research Activities

The aim of the processor research team is to create a future basis of high-performance processors for scientific simulations, as well as to explore the processor performance of the K computer.

In future high performance computing, we have to tackle with millions or more parallel operation units to extend the performance. However, many applications require acceleration while keeping the problem size, i. e. the strong scaling, and they can often be parallelized up to thousands of core, not to millions. To achieve better strong scaling, we have to decrease the cost of parallelization by improving the latency in everywhere – network, main memory, and processors. For this, we will try to develop the platform of System-on-Chip (SoC) based accelerators. It consists of general-purpose processor cores, memories, network units and computing accelerators on the single chip. By such integration we aim to achieve the ultimate performance for selected applications.

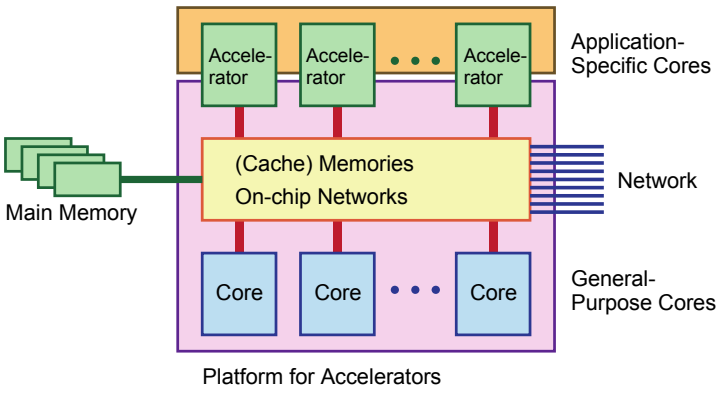


Figure 1. Diagram of platform for accelerators.

In addition to the researches on future computing platform, we will contribute to accelerate the application performance on the K computer. The processor of K computer, SPARC64 VIIIfx, has several special features for high-performance computing called VISIMPACT and HPC-ACE. We will explore to extract its power for several applications based on our experience on the processor architecture.

3.3. Research Results and Achievements

3.3.1. Platform of accelerators

In this year we have designed the MDGRAPE-4 SoC in RIKEN QBiC (Quantitative Biology Center). From the viewpoint of the platform of SoC based accelerator, we can use as the MDGRAPE-4 SoC as the basis. It has 64 processor units, 64 dedicated pipelines for molecular dynamics force calculation, main memories, and network units for 3-dimensional torus network. By replacing the dedicated pipelines we can use the design as the platform of accelerators.

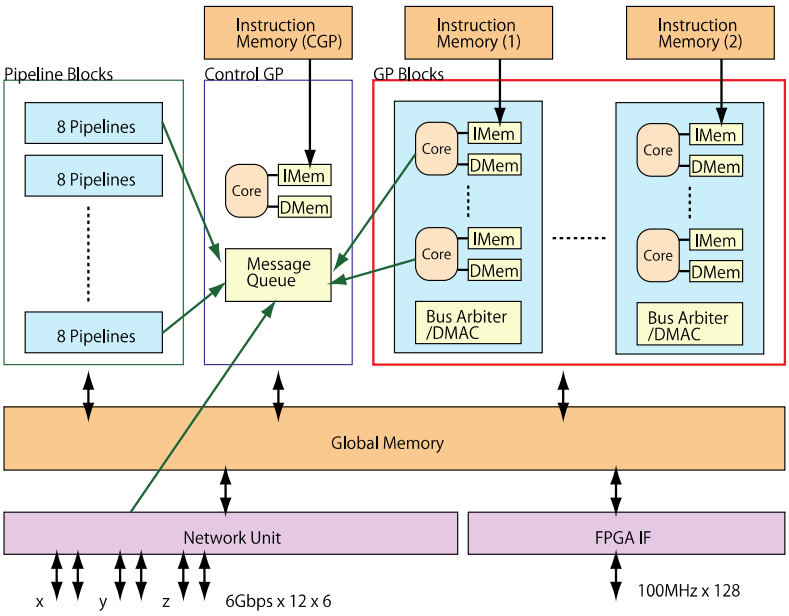


Figure 2. Block diagram of MDGRAPE-4 SoC.

3.3.2. Application Optimization on K computer

For application optimization we have started the optimization of the molecular dynamics core code.

3.4. Schedule and Future Plan

In the next year, we will finish the design of the MDGRAPE-4 SoC in RIKEN QBiC. In the team we will implement the part of the MDGRAPE-4 SoC as the platform of accelerators. We will also develop simulation environments for such system at RTL-level and behavior level. We will continue the optimization of MD core and the other codes for the K computer.

4. Field Theory Research Team

4.1. Team members

Yoshinobu Kuramashi (Team Leader)
Yoshifumi Nakamura (Research Scientist)
Xia-Yong Jin (Postdoctoral Researcher)
Ken-Ichi Ishikawa (Visiting Scientist)
Takeshi Yamazaki (Visiting Scientist)
Shinji Takeda (Visiting Scientist)

4.2. Research Activities

Our research field is physics of elementary particles and nuclei, which tries to answer questions in history of mankind: What is the smallest component of matter and what is the most fundamental interactions? This research subject is related to the early universe and the nucleosynthesis through Big Bang cosmology. Another important aspect is quantum properties, which play an essential role in the world of elementary particles and nuclei as well as in the material physics at the atomic or molecular level. We investigate nonperturbative properties of elementary particles and nuclei through numerical simulations with the use of lattice QCD (Quantum ChromoDynamics). The research is performed in collaboration with applied mathematicians, who are experts in developing and improving algorithms, and computer scientists responsible for research and development of software and hardware systems.

Lattice QCD is one of the most advanced case in quantum sciences: Interactions between quarks, which are elementary particles known to date, are described by QCD formulated with the quantum field theory. We currently focus on two research subjects: (1) QCD at finite temperature and finite density. We try to understand the early universe and the inside of neutron star by investigating the phase structure and the equation of state. (2) First principle calculation of nuclei based on QCD. Nuclei are bound states of protons and neutrons which consist of three quarks. We investigate the hierarchical structure of nuclei through the direct construction of nuclei in terms of quarks.

Successful numerical simulations heavily depend on an increase of computer performance by improving algorithms and computational techniques. However, we now face a tough problem that the trend of computer architecture becomes large-scale hierarchical parallel structures consisting of tens of thousands of nodes which individually have increasing number of cores in CPU and arithmetic accelerators with even higher degree of parallelism: We need to develop a new type of algorithms and computational techniques, which should be different from the conventional ones, to achieve better computer performance. For optimized use of K computer our research team aims at (1) developing a Monte Carlo algorithm to simulate physical system with negative weight effectively

and (2) improving iterative methods to solve large system of linear equations. These technical development and improvement are carried out in the research of physics of elementary particles and nuclei based on lattice QCD.

4.3. Research Results and Achievements

4.3.1. QCD at finite temperature and finite density

Establishing the QCD phase diagram spanned by the temperature T and the quark chemical potentialμin a quantitative way is an important task of lattice QCD. The Monte Carlo simulation technique, which has been successfully applied to the finite temperature phase transition studies in lattice QCD, cannot be directly applied to the finite density case due to the complexity of the quark determinant for finiteμ. We investigate the phase of the quark determinant with finite chemical potential in lattice QCD using both analytic and numerical methods. Applying the winding expansion and the hopping parameter expansion to the logarithm of the determinant, we show that the absolute value of the phase has an upper bound that grows with the spatial volume but decreases exponentially with an increase in the temporal extent of the lattice. This analytic but approximate result is confirmed with a numerical study in 4 flavor QCD in which the phase is calculated exactly. Figure 1 shows histogram of the phaseθat the chemical potential aμ=0.2 with a the lattice spacing on the 63×NT lattice in 4 flavor QCD and NT dependence of the phaseθas a function of μ. We observe that the phaseθis suppressed as the temperature T=1/NT decreases with the chemical potentialμfixed. The winding expansion is useful to study the high density region of the QCD phase diagram at low temperatures.

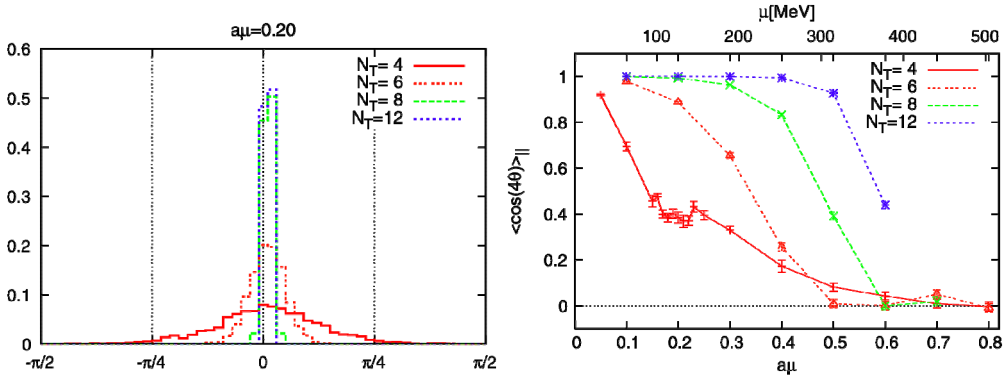


Figure 1. Histogram of the phaseθat the chemical potential *a*μ=0.2 with various N_T (left) and N_T dependence of the phaseθas a function of μ(right).

4.3.2. Nuclei in lattice QCD

In 2010 we succeeded in a direct construction of the 4He and 3He nuclei from quarks and gluons in lattice QCD for the first time. Calculations were carried out at a rather heavy degenerate up- and down-quark mass corresponding to mπ=0.8 GeV in quenched QCD to control statistical errors in the Monte Carlo evaluation of the helium Green's functions. As a next step we address the issue of bound state in the two-nucleon system at mπ=0.8 GeV in quenched QCD. To distinguish a bound state from an attractive scattering state, we investigate the spatial volume dependence of the energy difference between the ground state and the free two-nucleon state by changing the spatial extent of the lattice from 3.1 fm to 12.3 fm. In Fig. 2 we plot the spatial volume dependence of the energy differenceΔEL as a function 1/L3 with L the spatial extent. A finite energy difference left in the infinite spatial volume limit leads us to the conclusion that the measured ground states for not only spin triplet (3S1) but also singlet (1S0) channels are bound. Furthermore the existence of the bound state is confirmed by investigating the properties of the energy for the first excited state obtained by a 2×2 diagonalization method.

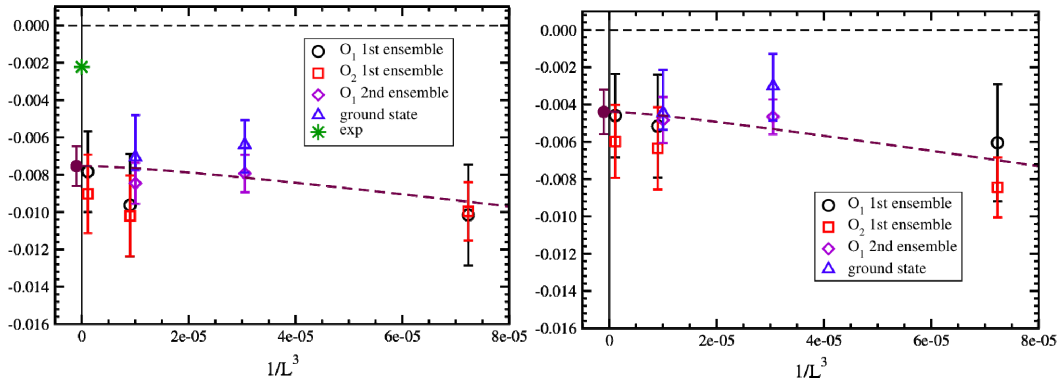


Figure 2. Spatial volume dependence of ΔEL in GeV units for the 3S1 (left) and 1S0 (right) channels. Open symbols represent the results obtained on the L3 boxes employing a couple of alternative calculational methods. Extrapolated results to the infinite spatial volume limit (filled circle) and experimental values (star) are also presented.

4.3.3. Development of algorithms and computational techniques

We consider to solve the linear systems with multiple right-hand sides expressed as AX=B, where A is an N×N matrix and X, B are N×L matrices with L the number of multiple right-hand side vectors. Various fields in computational sciences face this type of problem. In lattice QCD simulations, for example, one of the most time consuming part is to solve the Wilson-Dirac equation with the multiple right-hand sides, where A is an N×N complex sparse non-Hermitian matrix and X, B are N×L complex matrices with N the number of four dimensional space-time sites multiplied by 12. We aim at reducing the computational cost with the block Krylov subspace method which makes convergence faster than the non-blocked method with the aid of better search vectors

generated from wider Krylov subspace enlarged by the number of multiple right-hand side vectors. We improve the block BiCGSTAB algorithm with the QR decomposition. Figure 3 shows a representative case for residual norm as a function of number of iterations for the modified block BiCGSTAB. We observe an important feature that the number of iterations required for convergence decreases as the block size L is increased. The numerical results are summarized in Table 1, where the fourth and fifth columns are the number of matrix-vector multiplication (NMVM) and its gain factor, respectively. We find that the gain factor for time is about twice larger than that for NMVM. This is another important advantage of the block methods which allow us an effective usage of cache.

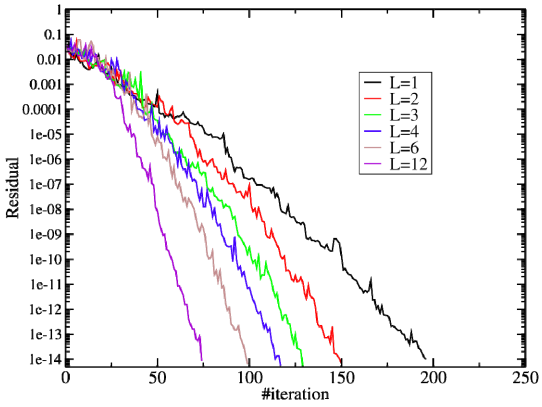


Figure 3. Representative case for residual norm as a function of number of iteration with L=1, 2, 3, 4, 6, 12 on a $32^3 \times 64$ lattice.

Table 1. L dependence for time, gain factor for time, number of matrix-vector multiplication and its gain factor on a $32^3 \times 64$ lattice. Central values are given for gain factors.

$L \times 12/L$	Time [s]	T (gain)	NMVM	NM (gain)
1×12	3827 (755)	1	17146 (3326)	1
2×6	2066 (224)	1.9	12942 (1379)	1.3
3×4	1619 (129)	2.4	10652 (832)	1.6
4×3	1145 (99)	3.3	9343 (835)	1.8
6×2	1040 (87)	3.7	7888 (663)	2.2
12×1	705 (70)	5.4	6106 (633)	2.8

4.4. Schedule and Future Plan

4.4.1. QCD at finite temperature and finite density

Before exploring the phase structure in 2+1 flavor QCD, we plan to investigate the 4 flavor case, where the first order phase transition is confirmed with zero chemical potential and is also expected with the non-zero chemical potential. To pin down the order of the phase transition we have to carry out the finite size scaling analysis by changing the spatial volume systematically.

4.4.2. Nuclei in lattice QCD

The existence of the bound state in the $^1S_0$ channel looks odd from the experimental point of view. However, there are two major systematic errors in our calculation: rather heavy quark mass corresponding to m_π=0.8 GeV and the quenched approximation. We expect that the bound state in the $^1S_0$ channel vanishes at some lighter quark mass with the inclusion of the dynamical quark effects. To confirm this scenario we are now carrying out a simulation of 2+1 flavor QCD with a reduced quark mass.

4.4.3. Development of algorithms and computational techniques

We have demonstrated the effectiveness of the block BiCGSTAB with the QR decomposition: remarkable cost reduction at large L thanks to smaller number of iterations and efficient cache usage. We are going to optimize the code for the K computer and make a numerical test to check the arithmetic performance and the scalability.

4.5. Publication, Presentation and Deliverables

(1) Journal Papers

1. S. Takeda, Y. Kuramashi, A. Ukawa, “Phase of quark determinant in lattice QCD with finite chemical potential”, Physical Review D 85 (2012) 096008.
2. T. Yamazaki, Y. Kuramashi, A. Ukawa, “Two-nucleon bound states in quenched lattice QCD”, Physical Review D 84 (2011) 054506.
3. Y. Nakamura, K.-I. Ishikawa, Y. Kuramashi, T. Sakurai, H. Tadano, “Modified block BiCGStab for lattice QCD”, Computer Physics Communications 183 (2012) 34.

(2) Conference Papers

1. S. Takeda, Y. Kuramashi, A. Ukawa, “On the phase of quark determinant in lattice QCD with finite chemical potential”, Proceedings of Science (Lattice 2011) 144.
2. Takeshi Yamazaki for PACS-CS Collaboration, “Bound state of two-nucleon systems in quenched lattice QCD”, Proceedings of Science (Lattice 2011) 147.
3. Y. Nakamura, K.-I. Ishikawa, Y. Kuramashi, T. Sakurai, H. Tadano, “Modified block BiCGStab for lattice QCD”, Proceedings of Science (Lattice 2011) 042.

(3) Invited Talks

1. Yoshinobu Kuramashi, “Nuclei in lattice QCD”, Perspective in Nuclear Physics with Large Scale Computing (RIKEN, Wako, Saitama, January 24-25, 2012).

2. Yoshinobu Kuramashi, “Large scale simulations of the femtometer world with lattice QCD”, The 4th Forum on Data Engineering and Information Management, deim2012 (Seaside Hotel MAIKO VILLA KOBE, Kobe, March 3-5, 2012).

(4) Posters and presentations

1. Yoshifumi Nakamura, “Modified block BiCGStab for lattice QCD”, The XXIX International Symposium on Lattice Field Theory, Lattice 2011 (Squaw Valley, Lake Tahoe, CA, USA, July 11-16, 2011).
2. Yoshifumi Nakamura, “Modified block BiCGStab for lattice QCD”, 2nd AICS International Symposium (AICS, Kobe, Hyogo, March 1-2, 2012).
3. Xiao-Yong Jin, “Exploring phases of QCD on the lattice”, 2nd AICS International Symposium (AICS, Kobe, Hyogo, March 1-2, 2012).
4. Takeshi Yamazaki, “Bound state of two-nucleon systems in quenched lattice QCD”, The XXIX International Symposium on Lattice Field Theory, Lattice 2011 (Squaw Valley, Lake Tahoe, CA, USA, July 11-16, 2011).
5. Takeshi Yamazaki, “Bound state in two-nucleon systems with lattice QCD”, The Physics Society in Japan 2011 Fall Meeting (Hirosaki University, Hirosaki, Aomori, September 16-19, 2011).
6. Takeshi Yamazaki, “Bound states of multi-nucleon systems in lattice QCD”, 2nd AICS International Symposium (AICS, Kobe, Hyogo, March 1-2, 2012).
7. Takeshi Yamazaki, “Calculation of light nuclei in lattice QCD”, HPCI Strategic Program Field 5 “The origin of matter and the universe” Symposium (Akihabara Convention Hall, Akihabara, Tokyo, March 7-8, 2012).
8. Shinji Takeda “Running coupling constant of ten-flavor QCD with the Schrödinger functional method”, Symposium on Lattice Gauge Theory (Wuppertal, Germany, May 2, 2011).
9. Shinji Takeda, “On the phase of quark determinant in lattice QCD with finite chemical potential”, The XXIX International Symposium on Lattice Field Theory, Lattice 2011 (Squaw Valley, Lake Tahoe, CA, USA, July 11-16, 2011).
10. Shinji Takeda, “Complex phase in QCD with finite chemical potential”, Progress of Computational Physics in Particle Physics, Nuclear Physics and Astrophysics (Nemunosato, Mie, December 3-5, 2011).

(5) Patents and Deliverables

- None

5. Computational Molecular Science Research Team

5.1. Team members

Takahito Nakajima (Team Leader)
Yoshinobu Akinaga (Research Scientist)
Michio Katoda (Postdoctoral Researcher)
Yutaka Nakatsuka (Postdoctoral Researcher)
Yusuke Ootani (Postdoctoral Researcher)
Muneaki Kamiya (Visiting Researcher)

5.2. Research Activities

The primary goal of the present project is to develop our own theoretical and computational molecular theory in order to perform accurate first-principle calculations on large-size and complicated molecular systems including nano- and bio-materials. The project involves the novel development of theory, algorithm, and software with the collaborative use of the K computer. We are currently developing our own theoretical and computational approaches for molecular calculations including large-scale molecular theory, accurate molecular theory, and relativistic molecular theory. We are also developing a new software called "NTChem". The quantum chemical software forms a basis for material and biological science. However, the development of the quantum chemical software in Japan falls far short of that in Western communities. Thus, we decided to develop a comprehensive new software of ab initio quantum chemistry made in Japan from scratch. The "NTChem" program package includes our own developing theoretical approaches and is a research product of our work to develop new and better theoretical approaches. With help of the K computer, the "NTChem" suite of programs is expected to be a useful tool in various computational studies for large and complicated molecular systems. By developing our own theory and software on the K computer, we will also perform the novel computational applications for realistic molecular systems. We hope to lead the way toward a new frontier of computational molecular science.

5.3. Research Results and Achievements

5.3.1. Development of two-component relativistic coupled-cluster algorithm

Relativity is one of the indispensable ingredients in molecular electronic-structure calculations for atomic and molecular systems involving heavy-elements. The coupled-cluster (CC) method, which is one of the most successful approaches among single-reference electron correlation methods, has also been generalized into relativistic ones by several research groups. In this work, we developed

two-component relativistic CC programs considering the scalar-relativity and the spin-orbit interaction. The heart of a CC program is the multiplication of CC amplitudes to intermediates which are either bare Hamiltonian elements or Hamiltonian elements contracted to other CC amplitudes. To enable the development of the relativistic CC program with an arbitrary order excitation, we followed the automatic code-generation scheme by Kállay *et al*. Our CC code generator is capable to generate the CC program at a given order with or without spin-orbit interaction. The program imports a two-component spin-orbit SCF wavefunction calculated with the NTChem program package. In addition to the ground-state correlation energy, the program is capable to compute electronic excitation energies by diagonalizing the CC Jacobian.

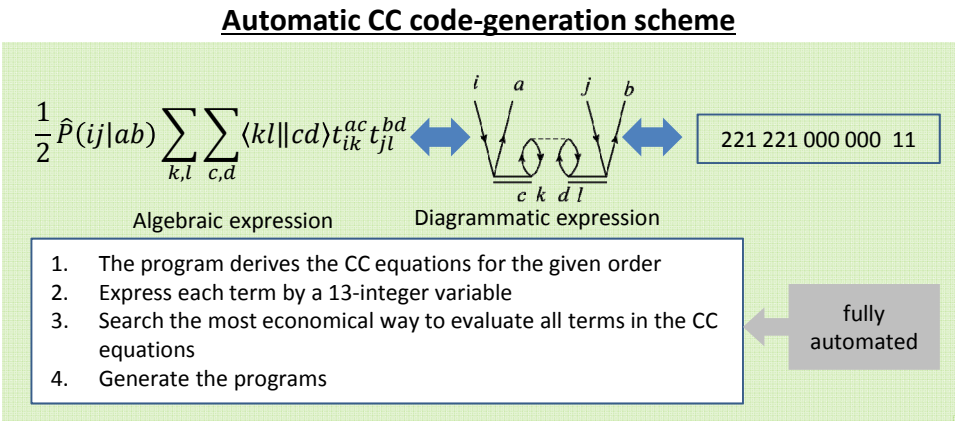


Figure. Scheme of the automatic CC code generation.

5.3.2. Development of efficient MPI/Open-MP hybrid parallel RI-MP2 algorithm

Second-order Møller–Plesset perturbation theory (MP2) is the simplest but robust electron correlation method to account for the non-covalent interactions that play important roles in the chemical phenomena of nano and biological molecules. However, the computational costs of MP2 calculations scale $O(N^5)$ with respect to the size of molecules (N), and practical applications are limited to molecules of moderate size. To make the MP2 calculations applicable to the large nano and biological molecules, the development of efficient computational techniques is desired. We have developed the parallel resolution-of-identity (RI) MP2 (RI-MP2) algorithm that is suitable for the computations of large molecule on the massively parallel supercomputers such as the K computer. In the previous parallel algorithms, the occupied orbitals are distributed to processors for the computations of two-electron integrals. However, the number of occupied orbitals is generally much smaller than that of the total molecular orbitals, so that the load balancing problems occur in the cases of the massively parallel computations. We have improved the algorithm to use the virtual orbitals for the parallel task distribution in order to make use of more large number of CPU cores with the efficient load balancing. The parallel RI-MP2 algorithm is also designed for the efficient parallel calculations by reducing the overheads of I/O and network communication. We have

implemented the new parallel RI-MP2 algorithm into the NTChem program. In this implementation, we also have performed MPI/Open-MP hybrid parallelization for the efficient usage of the memories and the network devices in the multi-core CPU architectures. We are developing the parallel RI-MP2 analytical energy gradient program enable to perform the geometry optimization calculations and ab initio molecular dynamic simulations of large molecules on the massively parallel supercomputers.

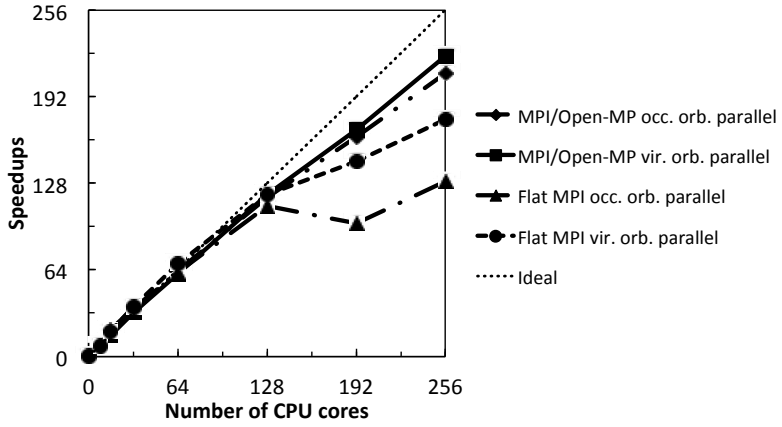


Figure. Speedups of parallel RI-MP2 calculations of taxol on the RICC massively parallel cluster

5.3.3. Development of the relativistic diffusion Monte Carlo algorithm

We have developed a relativistic extension of the quantum Monte Carlo (QMC) method which is an electron-correlation method suitable for massively parallel environment for general molecules. As the extension of the previously developed zeroth-order regular approximation variational Monte Carlo (ZORA-VMC) method, the approximate Green's function of the scalar relativistic version of the ZORA Hamiltonian and the relativistic diffusion Monte Carlo (DMC) method were derived. The ZORA-DMC method is implemented in our NTChem program package. The ZORA-DMC method can simultaneously treat the scalar relativistic effect on the total energy and the same amount of electron-correlation effects as the nonrelativistic DMC (e.g. about 96% and 91% of the estimated "exact" electron-correlation effects for Ne and Ar atom, respectively). The dissociation energy of the CuH molecule was evaluated with the ZORA-DMC calculation and the result was in the reasonable agreement with the experimental value. In this seminal year, the task parallelization has been implemented and the code achieved near 100% of the parallel efficiency with 8192 cores in the RICC system of RIKEN, Wako. Further code improvements will be performed to achieve the more flexible parallelization and the higher performance on the K computer.

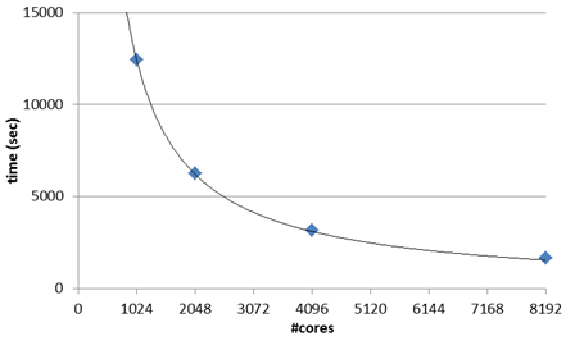


Figure. Strong scaling of R4QMC

5.3.4. Development of time-dependent density functional theory within the Tamm-Dancoff Approximation

Time-dependent density functional theory (TDDFT) is widely used for excited-state calculations of various molecules. Including the electron correlation through the exchange–correlation functional, TDDFT is known to give highly accurate results in spite of the simplicity and the low-computational cost. Recently, the Tamm–Dancoff approximation (TDA), which neglects the de-excitation effect in TDDFT and gives a similar formulation to the configuration interaction singles (CIS) method, has been applied to the evaluation of nonadiabatic coupling elements from the approximate TDDFT/TDA excited-state wave functions. In this work, we have implemented the TDDFT/TDA method in the NTChem programs by using the Liu-Davidson trial-vector algorithm. This implementation permits the calculations of excitation energy for both closed- and open-shell systems. The scaling of the computational cost of a TDDFT/TDA calculation per state with respect to the system size is the same as that for a Kohn-Sham calculation for the ground state. Since standard TDDFT and TDDFT/TDA exchange–correlation functionals generally underestimate excitation energies, particularly for Rydberg and charge-transfer states, we have also implemented range-separated functionals, with which charge-transfer excitation energies are generally improved. The required higher-order functional derivatives of these exchange correlation functionals have been derived and implemented into efficient computer codes with the aid of a newly-developed computerized symbolic algebra system.

5.3.5. Theoretical analysis of chemical reactions on (111) surfaces of noble metals

Adsorption of alkanethiols on noble metal surfaces has been one of the topics of interest in the field of surface science for more than 15 years. Lots of works, both experimental and theoretical, have been devoted to gain atomistic insight into the adsorption of alkanethiols and subsequent surface chemical reactions. It is well known that methanethiol (CH_3SH), the simplest alkanethiol, adsorbs molecularly to the on-top site on Au(111) surface, while methanethiolate (CH_3S) prefers the position

near the bridge site. Several computational works attempted to identify the reaction pathway of S-H dissociation of CH_3SH on Au and Cu surfaces using the NEB (Nudged Elastic Band) method. Despite that such calculations provide information which cannot be obtained experimentally, such as transition-state structures, the number of applications to surface reactions is still limited. We applied density-functional theory (DFT) to analyze the adsorption and S-H dissociation reaction of methanethiol on (111) surfaces of noble metals (Cu, Ag, and Au). Our primary purposes are: (1) identification of molecular and dissociative adsorption states, and (2) the determination of S-H reaction pathways. We used DFT with the PBE exchange-correlation functional, and the Kohn-Sham orbitals were constructed with plane-waves and pseudopotentials. The (111) surface was modeled by a four-layer slab, in which each layer contains 12 metal atoms. When optimizing adsorption structures, the atomic coordinates of the bottom layer were fixed at experimental values. We identified molecular and dissociative adsorption states of methanethiol on (111) surfaces of Cu, Ag, and Au. Moreover, we determined the reaction pathways for S-H dissociation on Cu and Au (111) surfaces using the NEB scheme. The calculated reaction pathways reveal a high energy barrier for S-H dissociation on Au(111) surface, in contrast to Cu(111) surface where the reaction is exothermic and can take place more easily than on Au(111) surface.

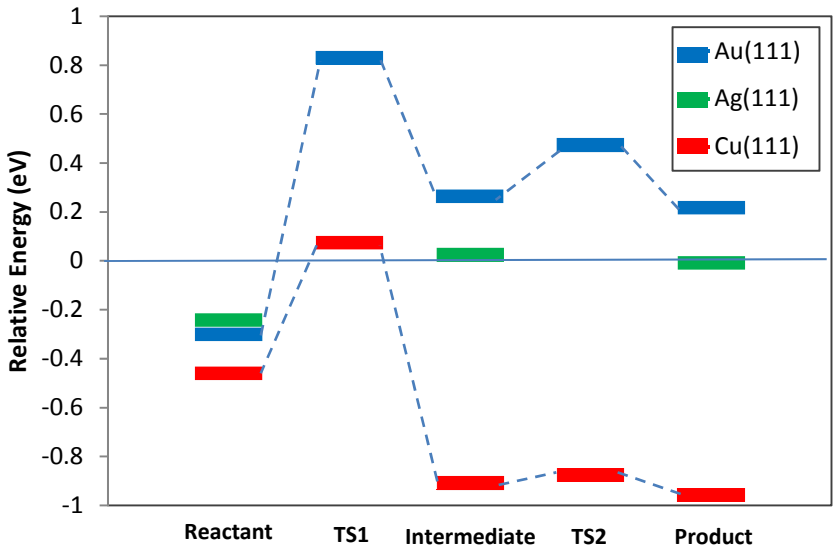


Figure. Relative energies along the methanethiol dissociation pathway on Cu, Ag, and Au (111) surfaces

5.3.6. Theoretical investigation of novel chemical reaction within supramolecular assembly

We investigated the novel chemical reaction which takes place in a supramolecule. It is reported that the aza-Cope rearrangement of allyl enammonium cation is accelerated when it is encapsulated in the cavity of M_4L_6 ($M=Ga^{3+}$, L=1,5-bis(2',3'-dihydroxybenzamido) naphthalene)) (Fig.). We investigated the catalytic activity of M_4L_6 by means of electronic-structure calculation. The geometry optimization was performed to locate of equilibrium structures. The reaction path for an isolated enammonium molecule was determined by means of The intrinsic reaction coordinate calculation. The Nudged Elastic Band (NEB) method was used to determine the reaction coordinate. The NEB method is a powerful tool for calculation of reaction paths of large molecular systems. By comparing the reaction path of an isolated enammonium with that of the encapsulated enammonium, we explored the catalytic activities of M_4L_6.

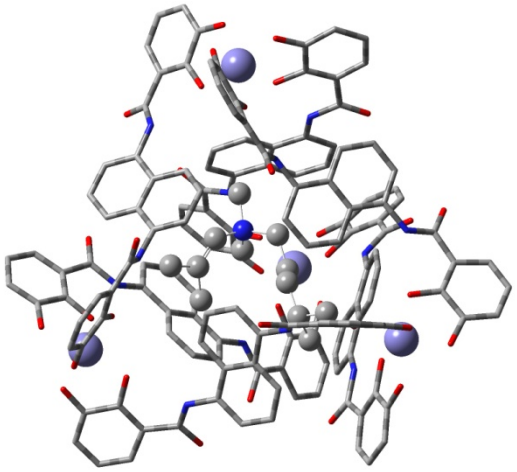


Figure. M_4L_6($M=Ga^{3+}$,L=bis(2,3-dihydroxybenzamido)naphthalene)

5.4. Schedule and Future Plan

In the next year, we will optimize the NTChem suite of programs on the K computer. Especially, we will highly parallelize the NTChem program on the K computer. We will also extend the functionalities of the NTChem program in order to explore the chemical reaction in the large and complicated molecular systems efficiently and accurately.

5.5. Publication, Presentation and Deliverables

(1) Journal Papers

1. Y. Kuwahara, K. Nishizawa, T. Nakajima, T. Kamegawa, K. Mori, H. Yamashita, "Enhanced Catalytic Activity on Titanosilicate Molecular Sieves Controlled by Cation-π Interactions", J. Am. Chem. Soc. 133, 12462–12465 (2011).
2. T. Yoshizawa, T. Nakajima, "Second-order generalized unrestricted Møller–Plesset perturbation theory for the spin–orbit part of zero-field splitting tensors", Chem. Phys. Lett. 515, 296–301 (2011).
3. W. Mizukami, T. Nakajima, K. Hirao, T. Yanai, "A dual-level approach to four-component relativistic density-functional theory", Chem. Phys. Lett. 508, 177–181 (2011).

(2) Review Articles and Books

1. T. Nakajima, K. Hirao, "The Douglas–Kroll–Hess approach", Chem. Rev. 112, 385–402 (2012).
2. T. Nakajima, Y. Nakatsuka, "Relativistic quantum Monte Carlo method", in Practical Aspects of Computational Chemistry I: An Overview of the Last Two Decades and Current Trends, edited by J. Leszczynski, M. K. Shukla, H. de Rode (Springer-Verlag), 293–317 (2012).
3. Takahito Nakajima, "New Frontier for Molecular Orbital Calculation", Kagaku, 66, 24–25 (2011). (in Japanese)

(3) Invited Talks

1. Takahito Nakajima, "Relativistic effect in Chemistry", "Winter School for Quantum Chemistry", Okazaki 20 Dec. 2011. (in Japanese)
2. Takahito Nakajima, "Theoretical Molecular Science on K Computer", 82th Kinki Kagaku Kyokai Reikai, Osaka, 11 Oct. 2011. (in Japanese)
3. Takahito Nakajima, "Theoretical Molecular Science in AICS", 2nd TCCI Symposium, Kobe, 12 Aug. 2011. (in Japanese)
4. Takahito Nakajima, "Theoretical Molecular Science and K Computer", 4th Progress Quantum Chemistry Symposium, Kyoto, 30 Apr. 2011. (in Japanese)
5. Takahito Nakajima, "New Generation Molecular Science", 4th Progress Quantum Chemistry Symposium, Kyoto, 7 Apr. 2011. (in Japanese)

(4) Posters and presentations

1. Y. Akinaga and T. Nakajima, "Theoretical Study on Adsorption and Reaction of Methanethiol on Nobel Surface" 92th Annual Meeting of Chemical Society of Japan, Yokohama, 25 Mar. 2012. (in Japanese)
2. Y. Ootani, Akinaga and T. Nakajima, "Theoretical Study on Aza-Cope Reaction inside Supramolecule" 92th Annual Meeting of Chemical Society of Japan, Yokohama, 25 Mar. 2012. (in Japanese)
3. Y. Nakatsuka, T. Nakajima, "Relativistic Quantum Monte Carlo in chemistry", 2nd AICS International Symposium, Kobe, 1 Mar. 2012. (Poster)
4. M. Katouda, T. Nakajima, S. Nagase, "Development of efficient computational methods for second-order Møller–Plesset perturbation calculation of extended systems", 2nd AICS International Symposium, Kobe, 1 Mar. 2012. (Poster)
5. Y. Ootani, Y. Akinaga, T. Nakajima, "Theoretical investigation of the aza-Cope rearrangement in a supramolecular assembly", 2nd AICS International Symposium, Kobe, 1 Mar. 2012. (Poster)

6. Y. Akinaga, T. Nakajima, “Methanethiolate on noble metal surfaces: Determination of reaction paths”, The 14th Asian Workshop on First-Principle Electronic Structure Calculations, Tokyo, 1 Nov. 2011. (Poster)
7. Takahito Nakajima, “EPR g tensor Calculation Based on Two-component Relativistic Molecular Theory”, Annual Meeting of Molecular Science, Sapporo, 20 Sep. 2011. (in Japanese)
8. Y. Nakatsuka and T. Nakajima, “Relativistic Diffusion Monte Carlo Method”, Annual Meeting of Molecular Science, Sapporo, 20 Sep. 2011. (in Japanese)
9. Y. Nakatsuka and T. Nakajima, “Relativistic Diffusion Monte Carlo Method”, 14th Annual Meeting of Theoretical Chemistry, Okayama, 12-14 May 2011. (in Japanese)
10. T. Yoshizawa and T. Nakajima, “Zero-field Splitting Tensor Calculation with Two-component Relativistic Approach”, 14th Annual Meeting of Theoretical Chemistry, Okayama, 12-14 May 2011. (in Japanese)

6. Computational Materials Science Research Team

6.1. Team members

Seiji Yunoki (Team Leader)
Yuichi Otsuka (Research Scientist)
Atsushi Yamamoto (Postdoctoral Researcher)

6.2. Research Activities

Solid state materials exhibit different properties ranging from metals, semi-conductors, insulators, magnets, and superconductors. A wide variety of these properties of materials are determined by quantum mechanical behaviors of a large number ($\sim 10^{23}$) of electrons. Thus, in order to understand materials properties, predict new behaviors, and design new materials, we must understand correctly the quantum mechanical motion of electrons that are described by Schrodinger equation. However, simulating many electron systems is computationally highly demanding and requires a massively parallel computer system such as the K computer in AICS.

The computational materials science research team conducts researches on developing new simulation methods/codes to search for a new quantum state of matter and to design new materials, specially in strongly correlated materials. Strongly correlated materials are a class of systems where Coulomb interactions between electrons are larger than the electron kinetic energy. These materials include various transition metal oxides such as high-Tc cuprate superconductors, colossal magneto resistive manganites, and organic conductive materials. The electronic states of weakly correlated materials such as semiconductors, where Coulomb interactions between electrons are only perturbative, can be well described by a single Slater determinant (one-particle approximation), and thus the standard density functional theory with an appropriate approximation works the best for weakly correlated materials. Instead, it is well known that the one-particle description is a poor approximation (and very often breaks down) for strongly correlated materials, and thus the standard method based on the density functional theory cannot be simply applied. Therefore, lots of efforts have been focused currently on developing new numerical methods to simulate strongly correlated materials.

The computational materials science research team focuses mainly on the following projects:

1) We develop a quantum Monte Carlo (QMC) method for a Hubbard-type model with more than 10,000 electrons where there is no sign problem.
2) We develop a density matrix renormalization group (DMRG) method to simulate two-dimensional Hubbard-type and quantum spin models for clusters as large as 16x16 and both at zero and at finite temperatures.
3) We develop a dynamical DMRG method and a time-dependent DMRG method to calculate

dynamical quantities as well as to simulate non-equilibrium quantum dynamics.

4) We offer highly efficient routines to perform in the K computer matrix-vector and matrix-matrix operations which are necessary in quantum many-body simulations such as the ones listed above.

With these newly developing methods, we conduct fundamental researches in strongly correlated quantum systems including various 3d transition metal oxides, low-dimensional organic materials, quantum spin systems, and cold atoms in optical lattices. We also plan to develop a first-principles simulation for strongly correlated materials.

6.3. Research Results and Achievements

6.3.1. QMC using the auxiliary field Monte Carlo technique

Among many QMC techniques to simulate interacting many electron systems like Hubbard model, we have developed the auxiliary field Monte Carlo method at zero temperature, which was originally proposed by J. E. Hirsch in 1985, by highly optimizing the code for the K computer. Describing one Slater determinant of many electrons in a single node (i.e., a Slater determinant not distributed to multiple nodes), we were able to use up to 24576 nodes with quite high efficiency of about 50% of the peak performance.

We have applied this method to the two-dimensional single-band Hubbard model, one of the simplest models to describe high-T_c cuprate superconductors and carbon sheets (graphene). Previously, the auxiliary field Monte Carlo method has been applied extensively to Hubbard-type models, and the largest lattice size reached was about N=600 sites by a German group in 2010. Using the K computer, we have reached up to N=2596 sites, about 4 times larger than the previous study. Note that the computational complexity scales like N^3. Thus it could have not been possible without the K computer.

Using this state-of-the-art QMC method, we have stared re-examined the ground state phase diagram of the half-filled Hubbard model on the honeycomb lattice, a model of graphene, as a function of the on-site Coulomb repulsion. This study is motivated by very recent finding of a possible spin liquid phase in the phase diagram of this model. This finding has surprised many researchers in this field because it is widely believed that a stable spin liquid appears most likely in geometrically frustrated systems and the Hubbard model on the honeycomb lattice is not frustrated. We are hoping to establish the grand state phase diagram of this model and resolve this issue of spin liquid.

6.3.2. DMRG

We have developed dynamical DMRG to calculate the one-particle excitation spectra for strongly correlated quantum systems. DMRG is known to be the best numerical method for one-dimensional quantum systems with extremely high precision. Extending DMRG to higher spatial dimensions is not trivial since the computational cost increases exponentially. To achieve high performance in the K computer, we have optimized the code using both MPI (and MPI2) and OpenMP. This improvement was necessary to simulate strongly correlated quantum systems in higher spatial dimensions even for quasi one-dimensional systems like several-leg ladders.

We have applied the dynamical DMRG to simulate cold fermionic atoms in quasi one-dimensional optical lattices with site dependent interactions. Depending the sign of local interactions, we have found a variety of phases at half-filling, including a possible superfluid phase. We have also studied the one-particle excitation spectra to characterize the nature of different phases.

6.4. Schedule and Future Plan

In the next academic year, we will conduct the following research.

6.4.1. QMC for a large number of interacting electrons

What prevents us from reaching even large lattice sites up to 10,000 sites in the K computer is the memory capacity available in each node. To simulate larger sites, we must distribute a Slater determinant into several nodes. So far we have used only a single node to describe a Slater determinant. In the next academic year, we are planning to develop a simulation code (using the auxiliary field Monte Carlo technique) in which a Slater determinant is distributed into several (4-6) nodes. A large size simulation is essential particularly when we want to study the phase boundary of a continuous quantum phase transition where the order parameter becomes inevitably small, thus requiring a careful finite-size scaling analysis.

6.4.2. DMRG for two-dimensional quantum systems

To simulate two-dimensional quantum systems, we will develop highly parallelized DMRG method at zero temperatures as well as finite temperatures. We have already developed highly parallelized DMRG code for one-dimensional systems, and in this academic year we will extend this to higher spatial-dimensions. Using this method, we are planning to study the ground state phase diagrams as well as various excitation spectra, including one-particle excitation spectra, linear and non-linear optical conductivities, and magnetic excitation spectra, for geometrically frustrated quantum spin systems and Hubbard-type interacting electrons.

6.4.3. Time-dependent DMRG to simulate quantum dynamics in strongly correlated systems

Time-dependent DMRG (t-DMRG) is the only reliable method thus far to simulate non-equilibrium quantum dynamics in strongly correlated systems. Thus, it is crucial to develop a highly efficient code of t-DMRG in the K computer. We have been developing a t-DMRG for the K computer, and we will continue doing it in this academic year. The main goal in this academic year is to improve the performance much better both for a single node and for a moderate number of nodes, before trying massively parallel simulations. For this purpose, we will take a simple model such as Hubbard-like models for cold atoms in optical lattices.

6.5. Publication, Presentation and Deliverables

(1) Papers

1. S. Zhang, S. Yamagiwa, M. Okumura, and S. Yunoki, "Performance Acceleration of Kernel Polynomial Method Applying Graphics Processing Units", IPDPS/APDCM 11, p. 564-571 (2011).
2. Q. Zhang, S. Hikino, and S. Yunoki, "First-principles study of the spin-mixing conductance in $Pt/Ni_{81}Fe_{19}$ junctions", Appl. Phys. Lett. 99, 172105 (2011).
3. H. Watanabe, T. Shirakawa, and S. Yunoki, "Spin-orbit-induced Mott insulator in 5d electron systems", J. Phys. Soc. Jpn. 80, SB006 (2011).
4. T. Shirakawa, H. Watanabe, and S. Yunoki, "Microscopic study of electronic and magnetic properties for Ir oxide", J. Phys. Soc. Jpn. 80, SB010 (2011).
5. T. Saitou, A. Koga, and A. Yamamoto, “Metal-insulator transition in optical lattice system with site-dependent interactions”, arXiv:1204.4188 (2012).

(2) Invited Talks

1. S. Yunoki, "Microscopic study of spin-orbit-induced Mott insulator in Ir oxides ", Nanjing University (China) (Octobor, 2011).
2. S. Yunoki, "Microscopic study of spin-orbit-induced Mott insulator in Ir oxides ", SISSA (Italy) (February, 2012).

(3) Posters and presentations

1. S. Zhang, S. Yamagiwa, M. Okumura, and S. Yunoki, "Performance Acceleration of Kernel Polynomial Method Applying Graphics Processing Units", 13th Workshop on Advances in Parallel and Distributed Computational Models (APDCM2011), Anchorage (USA) (May, 2011).
2. H. Watanabe, T. Shirakawa, and S. Yunoki, "Variational Monte Carlo study for superconductivity in multi-orbital systems", The 26th International Conference on Low Temperature Physics, Beijing (China) (August, 2011).

3. Q. Zhang and S. Yunoki, "Magnetic Properties and Improper Ferroelectricity in $LaFeO_3/LaCrO_3$ Superlattices", The 26th International Conference on Low Temperature Physics, Beijing (China) (August, 2011).
4. Q. Zhang, G. Chen, X.-G. Gong, and S. Yunoki, "d^0 Ferromagnetic Surface in HfO_2", The 26th International Conference on Low Temperature Physics, Beijing (China) (August, 2011).
5. V. Badaut, T. Shirakawa, and S. Yunoki, "A Haldane-Anderson model study for the iron spin and charge state in Myoglobin", The 26th International Conference on Low Temperature Physics, Beijing (China) (August, 2011).
6. A. Yamamoto, S. Yamada, and M. Machida, "A novel quantum phase of one-dimensional fermionic optical lattices with spatially alternating interactions" (24aBA-6), The 67th Annual Meeting of the Physical Society of Japan, Osaka (Japan) (March, 2012).
7. T. Saitou, A, Koga, and A. Yamamoto, "Metal-insulator transition of fermionic optical lattices with spatially modulated on-site interactions" (24pAD-2), The 67th Annual Meeting of the Physical Society of Japan, Osaka (Japan) (March, 2012).
8. S. Yamada, J, Igarashi, M. Okumura, H. Onishi, A. Yamamoto, and M. Machida, "A DMRG study of superfluid correlations in a repulsive Hubbard model in a confined potential" (27aBG-13), The 67th Annual Meeting of the Physical Society of Japan, Osaka (Japan) (March, 2012).

(4) Patents and Deliverables

- None

7. Computational Biophysics Research Team

7.1. Team members

Yuji Sugita (Team Leader)
Jaewoon Jung (Research Scientist)
Ryuhei Harada (Postdoctoral Researcher)
Yasuhiro Matsunaga (RIKEN Special Postdoctoral Researcher)
Naoyuki Miyashita (Research Scientist (Concurrent))*
Takaharu Mori (RIKEN Special Postdoctoral Researcher (Concurrent)*
Hiromi Kano (Assistant (Concurrent))*
* The main affiliation of these people is Laboratory for Biomolecular Function Simulation, Computational Biology Research Core, RIKEN Quantitative Biology Center.

7.2. Research Activities

Recent advances in structural biology, atomic structures of macromolecules can be determined by X-ray crystallography and nuclear magnetic resonance (NMR). These coordinates are stored in protein data bank (PDB) and are utilized for academic researches or industrial usages like drug design. This information is usually quite useful to understand molecular mechanisms underlying protein stability, large-scale conformational changes, ligand-receptor binding, and enzymatic reactions. However, due to the complexity of the structures, the static structural information is, in some cases, not enough to understand biological phenomena. Some of the proteins show large conformational changes during their functional cycles.

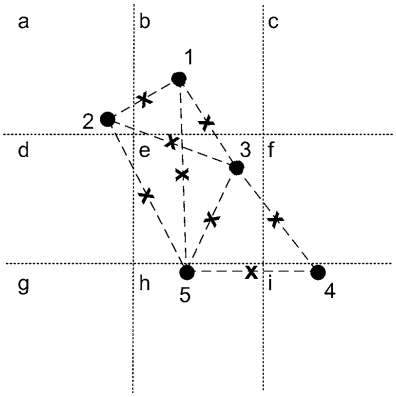
Computer simulations based on molecular dynamics (MD) method using the macromolecular structures now become important research tools in biological sciences. By performing the MD simulations in supercomputer, we can simulate molecular motions of proteins, which occur on the time scale between nsec and μsec. However, much longer simulations (from msec to sec) are highly desired to simulate whole processes of most of biological phenomena.

In our team, we aim to develop software for MD simulations to use K computer most efficiently. We also try to introduce novel algorithms or models into the software. Our software allows us not only to simulate dynamical properties of macromolecules, but also to provide thermodynamic quantities at physiological conditions. Currently, we focus on the parallelization of MD simulation code, the development of replica-exchange molecular dynamics (REMD) method or reaction-path search method and data-assimilation algorithm, and large-scale MD simulations of biological systems in cellular environment.

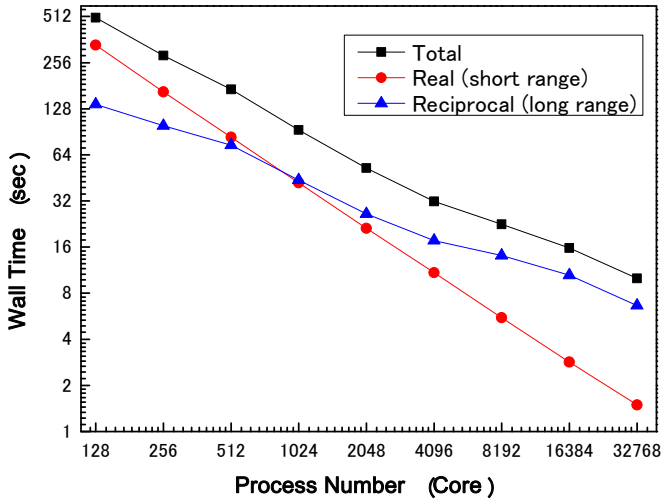
7.3. Research Results and Achievements

7.3.1. Parallelization of Molecular dynamics simulation codes

In SPDYN, we use the midpoint method as a basic scheme. In the midpoint method, two particles interact on a particular box if and only if the midpoint of the segment connecting them falls within the region of space associated with that box.

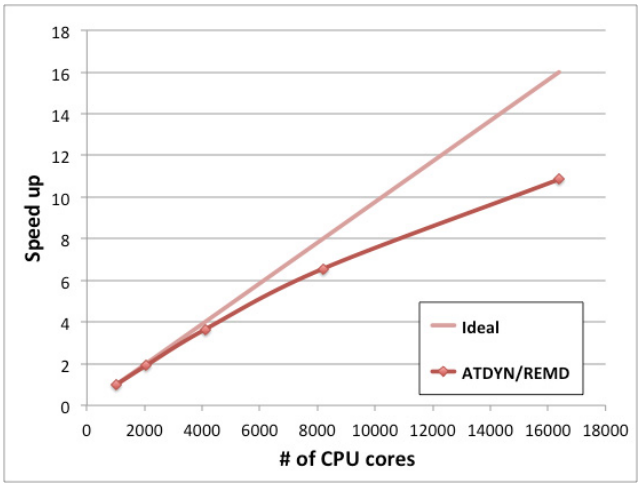


In this figure, each pair of particles separated by a distance less than R (cutoff distance) is connected by a dashed line segment, with "x" at its center lying in the box which will compute the interaction of that pair. We apply this scheme not only non-bonded but also bonded interactions. With this scheme, we could get a good scalability for short-range interactions. As for the long range electrostatic intreactions, smooth particle mesh Ewald method is assigined, in which real part is calculated with cutoff (short-range) and reciprocal part uses 3D FFT. For the efficient parallelization of reciprocal part calculation, volume decomposition scheme with MPI_Allgather is used for FFT.

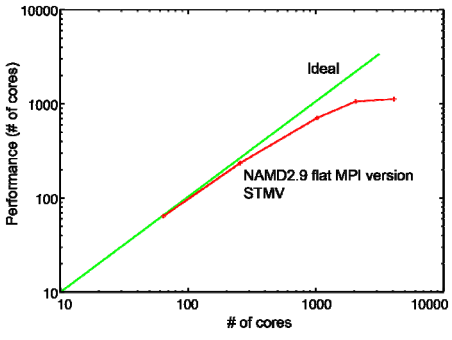


Parallel performance of virus (1,000,000 atoms) with SPDYN using K computers. Wall time is checked for 1000 steps MD calculations.

In the ATDYN program, computation of the energy and forces acting on atoms are parallelized for the number of interactions in each term of the potential energy function. We introduced not only an all-atom model but also coarse-grained model (Go model) in ATDYN. We also implemented a standard molecular dynamics method and the replica-exchange molecular dynamics method (REMD). REMD is one of the efficient sampling methods for biomolecular systems. In this method, some replicas of the system are prepared, where each replica has different parameters. During the REMD simulation, such parameters are exchanged between the neighboring replicas at certain intervals. REMD can sample structures of biomolecules in a wide range of the configurational space without getting trapped in local-minimum free-energy states. The following figure shows a parallel performance for the system containing 92,224 atoms using REMD method (128 replicas) with ATDYN/REMD on K computers. We found that the computational speed is scalable even if we use more than 10,000 CPU cores.



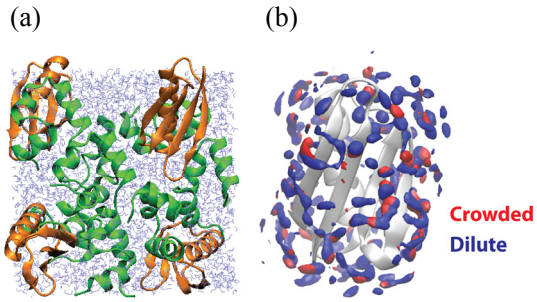
NAMD2, which have been developed by Theoretical and Computational Biophysics Group in University of Illinoi, is a highly parallelized molecular dynamics code designed for high-performance simulation of large biomolecular system, and it is widely used in biophysics field. Comparison of performance between GENESIS and NAMD2 is useful not only for the developers of GENESIS but also for many NAMD2 users in K computer. We have installed the flat MPI version of NAMD2.9 in K computer. Though it is based on Charm++ parallel objects, it has been compiled using MPI. The figure shows the parallel performance of virus, which has 1M atoms, with NAMD2.9 flat MPI version using K computer. We are now trying to compile the program using hybrid parallel scheme based on pthread/MPI library in collaboration with Dr. Kamada at



Kobe University.

7.3.2. Large-scale molecular dynamics simulations for biological systems

As future applications in terms of GENESIS with Replica-Exchange Molecular Dynamics (REMD) on K computer, we are planning to perform large-scale molecular dynamics simulations for biological systems such as crowded cellular environments. As a first step to simulate actual cellular environments, we investigated simple crowding systems consisting of protein G and villin headpiece sub-domain (Figure (a)). The results of crowding simulations showed significantly differences between crowded and dilute conditions. For instance, the distributions of water were perturbed due to crowding (Figure (b)), indicating that crowding effects hydration under crowded cellular environments.



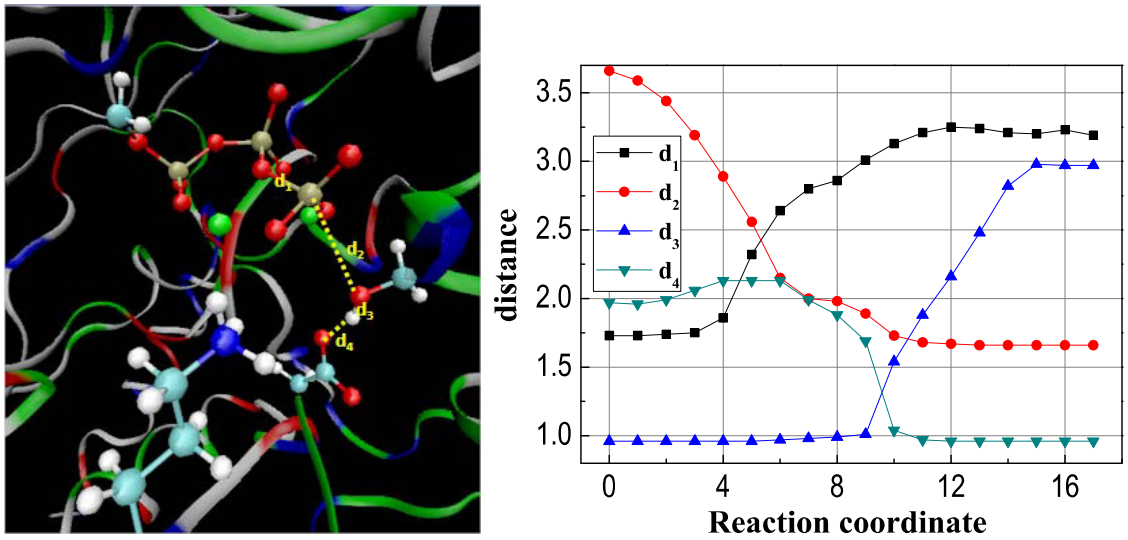
(a) A protein crowding system consisting of two proteins, protein G (brown) and villin headpiece sub-domain (green).

(b) Perturbations of three dimensional distributions of water due to crowding around protein G (white). Each color corresponds to the distributions of water under crowded (red) and dilute (blue) conditions.

7.3.3. Development of the replica-path optimization

The nudged elastic band (NEB) and string methods are widely used to obtain the minimum-energy path of chemical reactions and phase transitions. In these methods, however, it is difficult to define an accurate Lagrangian to generate the conservative forces, resulting in slow convergence. On the other hand, the constrained optimization with locally updated planes scheme (CO-LUP) defines the target function properly, although the method does have problems of inaccurate estimation of reactions and inappropriate accumulation of images around the energy minimum. We introduce three modifications into CO-LUP to overcome these problems: (1) An improved tangent estimation of the reaction path, which is used in the NEB method, (2) Redistribution of images using an energy-weighted interpolation before updating local tangents, and (3) Reduction of the number of constraints, in particular translation/rotation constraints, for improved convergence. The present method benefits from a micro-iteration scheme for protein environments in QM/MM optimization.

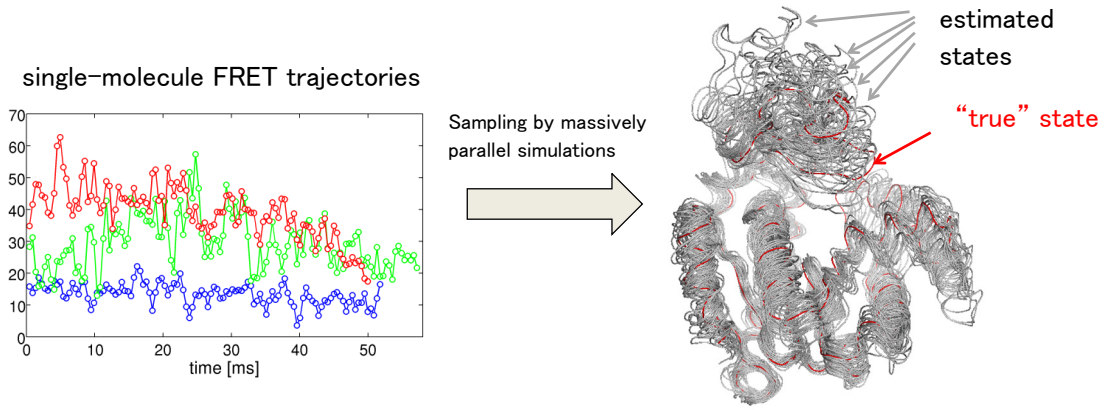
We test the method on the isomerization of alanine dipeptide. We also apply the method for defining the reaction paths of the rearrangement reaction catalyzed by chorismate mutase (CM), and of the phosphoryl transfer reaction catalyzed by cAMP-dependent protein kinase. In both cases, the results are consistent with previous QM/MM calculations.



Changes of key structural parameters of PKA during the reaction process.

7.3.4. Data assimilation algorithm

Förster resonance energy transfer (FRET) is a powerful method for determining various mechanistic properties of biomolecules. However, except for the simple cases of well-separated, low-noise, and limited to only a few states, it is usually difficult to determine the conformational states and the transition rates between them from FRET trajectories. The general state space hidden Markov models have opened a possibility to analyze such a complex time series. With the help of the recent advances in computer power and sampling techniques, it has become possible to determine the hidden states of the high-dimensional models even from low-dimensional information (Figure (a)). As future applications with GENESIS and K computer for FRET trajectories, we are developing a sampling scheme utilizing the sequential Monte Carlo filtering (particle filtering (Figure (b))). As a first demonstration of the algorithm, we have applied the method to the high-resolution single-molecule FRET trajectories (J.A. Hanson *et al.*, *PNAS* 2007) of the domain movements in adenylate kinase.



A schematic picture of the sequential Monte Carlo estimation of the high-dimensional states (i.e., protein conformations in this case) from one-dimensional FRET trajectories. The protein conformations represented by gray lines indicate the estimated conformations, and the red line corresponds to the “true” conformation.

7.4. Schedule and Future Plan

The major goal of research in next financial year (2012) is to finish the parallelization of our MD code and perform large-scale MD simulations of biomolecules in cellular environment using K computer. Before the large-scale applications, we need to perform test calculations on several different molecules under various conditions. By the end of FY2013, we are planning to open the source code of our MD program for academic researchers as well as industrial users under the license of GPL. We also continue to develop the MD code by introducing enhanced conformational sampling techniques like the generalized-ensemble method and path-optimization method. These methods are in particular useful for the free-energy calculations of biomolecules.

We are also planning to develop QM/MM molecular dynamics module in our code in collaboration with Dr. Nakajima’s team at RIKEN AICS. By combining with parallelized QM code developed by Dr. Nakajima’s group, we can perform highly parallelized QM/MM molecular dynamics simulations or QM/MM free-energy calculations of biomolecules or other molecular systems.

7.5. Publication, Presentation and Deliverables

(1) Journal Papers

1. R. Harada, Y. Sugita, and M. Feig, “Protein crowding affects hydration structure and dynamics”, Journal of American Chemical Society, 134, 482-4849 (2012).
2. Y. Matsunaga, H. Fujisaki, T. Terada, T. Furuta, K. Moritsugu, and A. Kidera, “Minimum Free Energy Path of Ligand-Induced Transition in Adenylate Kinase”, PLoS Computational Biology 8, e1002555 (2012).

(2) Conference Papers

- None

(3) Invited Talks (From April 2011 to March 2012)

1. R. Harada, Y. Sugita and M. Feig, “Protein crowding affects hydration structure and dynamics”, The 4th Applied Systems Biology Workshop at RIKEN AICS, November 7, 2011.
2. Y. Matsunaga, “Basics and advances of molecular dynamics simulation of biomolecules”, Bio super-computing summer school 2011 in Awaji, September 26-27, 2011.
3. Y. Matsunaga, “Sampling of conformational transitions in proteins: string method and data assimilation”, IMS workshop at The Institute of Statistical Mathematics, March 7, 2012.

(4) Posters and presentations

1. J. Jung, S. Re, Y. Sugita, and S. Ten-no, “New implementation of the reaction path determination in the QM/MM method”, The 49th Annual meeting of the Biophysical Society of Japan in Himeji, September 16-18 (2011).
2. R. Harada, Y. Sugita and M. Feig, “Protein crowding affects hydration structure and dynamics”, The 25th annual meeting of the molecular simulation society of Japan at Tokyo Institute of Technology, December 6, 2011.
3. Y. Matsunaga, H. Fujisaki, T. Terada, and A. Kidera, “Minimum Free Energy Path of Ligand-Induced Transition in Adenylate Kinase”, The 25th annual meeting of the molecular simulation society of Japan at Tokyo Institute of Technology, December 6, 2011
4. Y. Matsunaga, H. Fujisaki, T. Terada, and A. Kidera, “Conformational Transition Pathways of Adenylate Kinase Explored by the String Method”, Biophysical Society 56th Annual Meeting in San Diego, USA, February 24-29, 2012.

(5) Patents and Deliverables

- None

8. Computational Climate Science Research Team

8.1. Team members

Hirofumi Tomita (Team Leader)
Shin-ichi Iga (Research Scientist)
Yoshiaki Miyamoto (Postdoctoral Researcher)
Tatsuya Seiki (Postdoctoral Researcher)
Hisashi Yashiro (Postdoctoral Researcher)[1]
Ryuji Yoshida (Research Associate)[2]
1: joined us from the October in 2011
2: joined us from the September 2011 as a member of G8 Research Councils Initiative "ICOMEX"

8.2. Research Activities

In climate research area, the simulation has become one of the most vital technique as well as theory and observation. Since the IPCC started to report environmental assessment in 1990s and tackled to the global warming issue, many climate projections by various climate models have been performed. However, from restrictions of computer resources, the reliability issue of models still remains and the climate model do not become in the mature stage now.

In this research team, a pioneering research, which suggests the direction of future climate simulation, will be conducted. In order to raise the reliability of climate model more than the current status, we aims to construct a climate model based on more theoretically physical principles.

Such a model needs tremendously large computer resources. Therefore, it is necessary to design the model to pull out the capability of computers as much as possible. Recent development of supercomputers has a remarkable progress, however, new numerical techniques may be needed under the collaboration of hardware research and software engineering for effective use of them on the future HPC, including K computer.

For the above research achievement, our team is cooperating with the computational scientists in other fields and computer scientists. We enhance the research and development including effective techniques and make a next-generation climate model. Now, establishment of the above basic and infrastructure research on K Computer is required. This kind of research also leads to post K computer or subsequent ones.

In order to archive the team mission above, we have started three projects in this fiscal year:

1) Construction of a new library for climate study:
   We have proposed the subject "Estimation of different results by many numerical techniques and their combination" as a synergetic research to MEXT in the last autumn during the

discussion with the Strategic 5 fields. In this project, our team committed the provision of numerical libraries for climate study and other related fields, which contribute to K users.

2) Feasibility study of Exa-scale computing using the general circulation model:
The project of G8 Research Councils Initiative "ICOMEX" has been started from the last autumn. Through this project, a part of our team does the feasibility study of Exa-scale computing by the global cloud resolving model.

3) Grand challenge run for sub-km horizontal resolution run by global cloud-resolving model:
A part of our team commits the contribution to the 3rd Strategic Field Project. In this project, the global cloud resolving model NICAM is a main simulation code. Especially, we are responsible to get high computational performance on K computer.

8.3. Research Results and Achievements

8.3.1. A new library for climate study on K computer

In FY2011, we conducted a feasibility study how efficiently the current numerical schemes do run on the K computer. For this purpose, we have organized a new multi-disciplinary team "Team SCALE" with computer scientists in AICS. Through this activity, we developed a new meteorological program code, which targets to K computer and also directs to the future meteorological science. This program is a kind of Large Eddy Simulation (LES) code. Since the required resolution is much higher than Cloud-Resolving Model (CRM), the adaptation of such a technique is still limited to a narrow domain. However, once Peta-scale computers such as K computer are available, the technique will be more popular in the wider domain. In addition, it is very useful to develop the LES code on K computer not only for the meteorological field but also for other fields in which their research are achieved by the computational fluid dynamics.

For the meteorological application, the time integration method has a wide variety; full explicit scheme, full implicit scheme, and horizontally explicit and vertically implicit scheme. We started to compare the various time integration methods. First, we employed the full explicit scheme, because it gives a physical and computational reference solution. Figure 1 shows a snapshot result from the planetary boundary layer test case. Comparing with the results published already, we checked that our result has a reasonable statistics for the atmospheric turbulence.

For the microphysics, the single moment bulk method by the Kessler(1969) and a double moment bulk method developed by Seiki and Nakajima (2012) are implemented. Table 1 gives a typical computational performance on K computer. In this table, we show the elapse time. The sustained performance is over 10% in the dynamical part. Figure 2 shows the scalability up to 24576 nodes. This code has a reasonable weak-scaling.

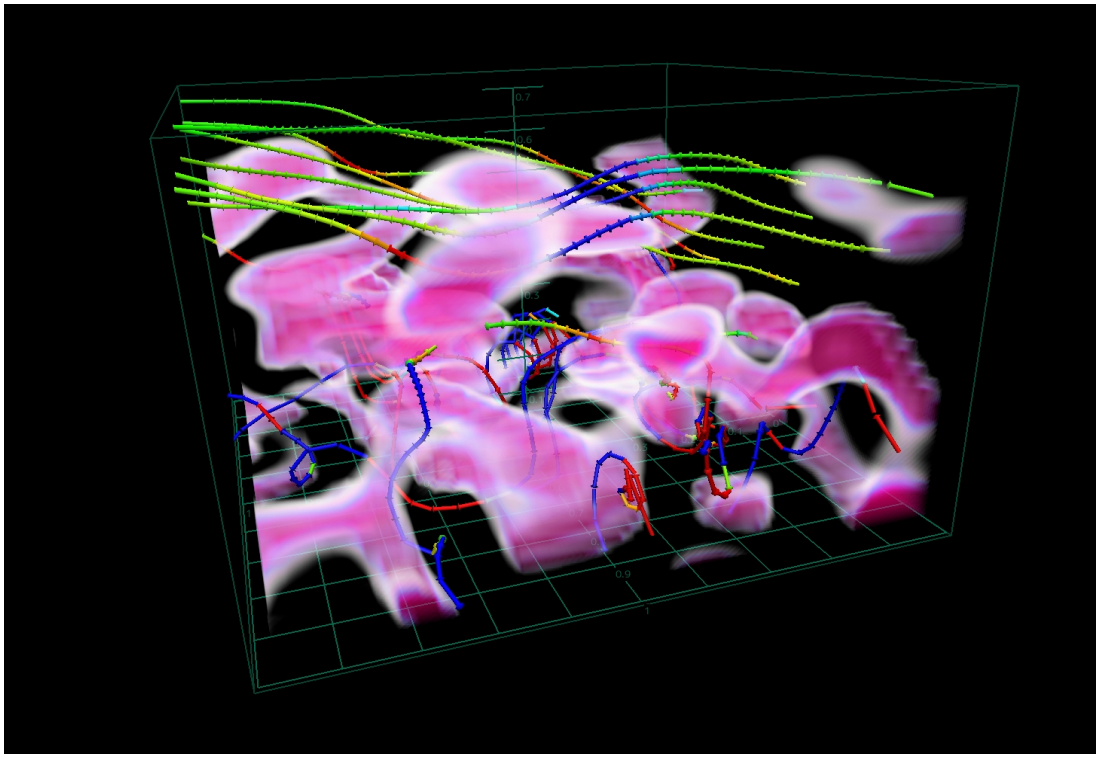


Figure 1. A snapshot of planetary boundary layer simulation with the 25m isotropic resolution. The isosurface and vector indicates the anomaly of potential temperature form the environment and velocity vector, respectively.

Table 1. Computational performance of SCALE-LES: The array size and time step are (kmax,imax,jmax)=(1256,32,32) and tmax = 300, respectively. The CTL is the case without turbulence(TB) and microphysics schemes (MP). RDMA is the case that uses RDMA functions directly in the communication. w/kessler and w/NDW6 are the cases with cloud microphysics of single and double moment bulk scheme.

[sec]	CTL	RDMA	w/ Kessler	w/ NDW6 orig	w/ NDW6 tuning
Main loop	110.6	102.5	112.7	124.9	116.3
Dynamics	109.5	101.5	108.7	108.7	108.7
TB	0.0	0.0	0.8	1.0	0.9
MP	0.0	0.0	1.6	13.7	5.4
COMM	15.9	4.7	15.7	15.7	15.7

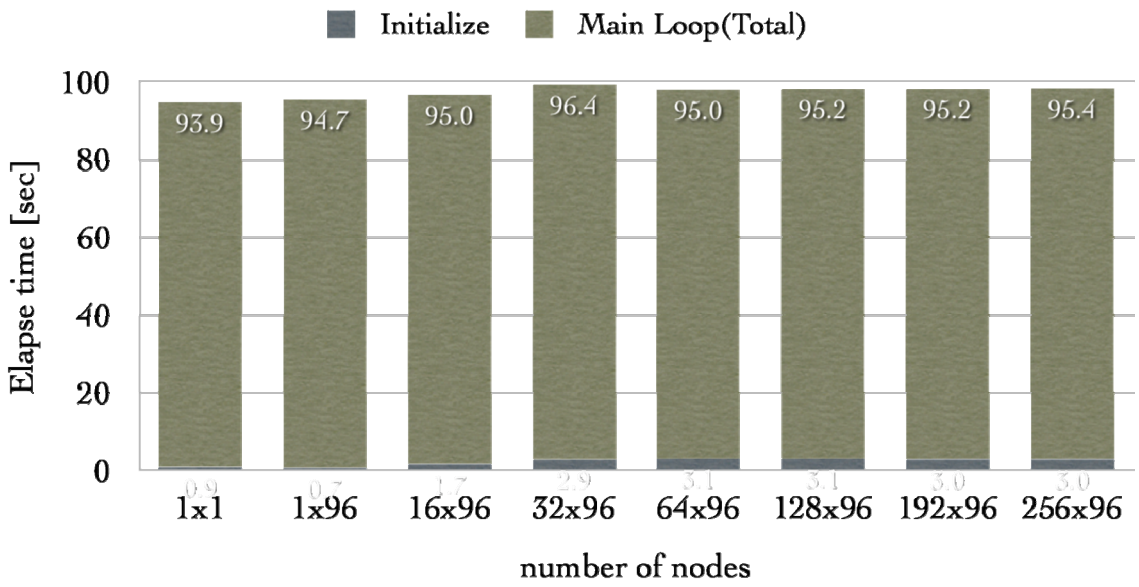


Figure 2. The scalability (weak-scaling) of SCALE-LES on K Computer.

8.3.2. ICOMEX project

This project is, especially, focusing on the computational efficiency and possibility of popular icosahedral models, NICAM, ICON, and MPAS. Since the kick-off meeting on the October 2011, we have been concentrating to prepare the environment for model inter-comparison.

As the first step, we have compared the physical results from NICAM and ICON. Figure 3 shows the results of Held & Suarez Test Case for both of models. As a first step, the reasonable results are obtained. For example, the intensity of jet decrease with increase of resolution, the cold region in the upper atmosphere on the tropic becomes clear. After the scalability problem will be resolved, we will systematically increase resolution and discussed the model convergence for both of models.

Another test case is Jablonowski & Williamson Test Case, which is a deterministic baroclinic test case. In the original paper, the equilibrium state is formulated by eta-coordinate in the vertical direction. We are now changing this condition to one by z-coordinate. T.

8.3.3. Grand challenge run by NICAM

In FY2011, we have computationally tuned NICAM with AICS Operations and Computer Technologies Division so as to efficiently conduct it on K Computer. We have also prepared to conduct a grand challenge simulation, which tremendously increases the resolution by 800m horizontal grid and 96 vertical layers. We mainly have made efforts of reduction of the memory load/store. As a result, we got 7% sustained performance in the configuration of grand challenge run.

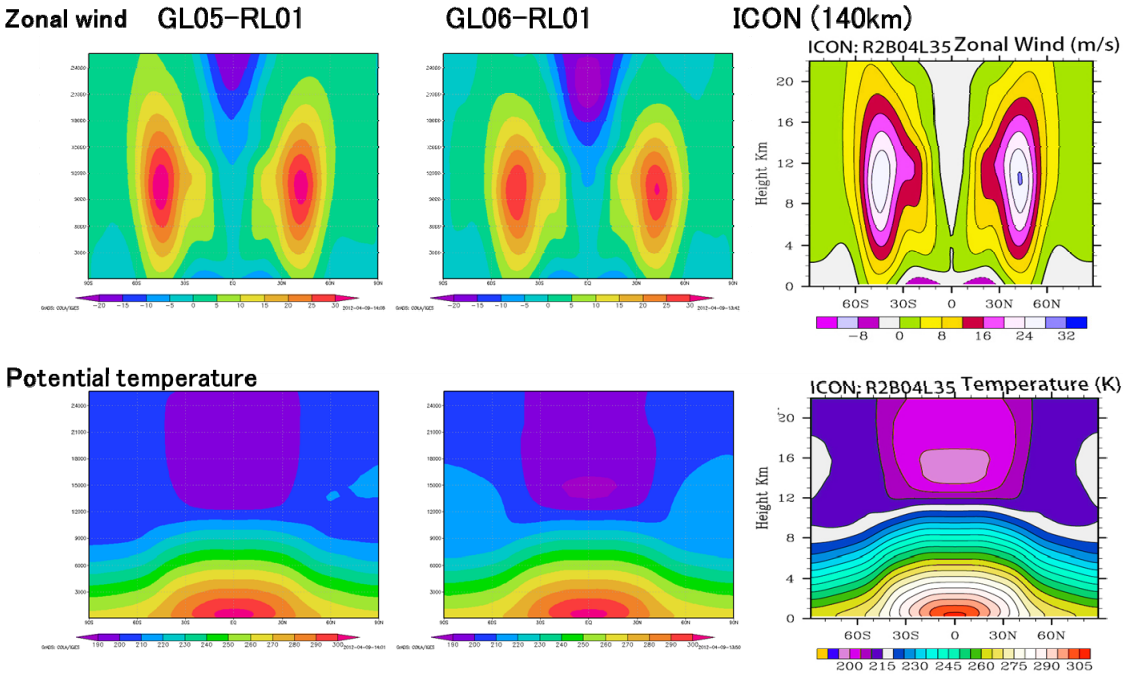


Figure 3. The inter-comparison for Held & Suarez Test Case. Climatology of zonal wind and temperature.

8.4. Schedule and Future Plan

In the next year, we will continue the feasibility study for the fast numerical library on K computer. Especially, we will give an insight into what kind of the time integration method is promising on K computer from the viewpoints of computational and physical performances. Further, we will go ahead with the construction of numerical library, considering the I/O and communication. The first version will be released by the end of FY2013.

In the ICOMEX, we also continue to do model inter-comparison, adding MPAS as soon as its code will be provided from the U.K. team. Until the end of the next fiscal year, the measures of computational and physical performance for NICAM, ICON, and MPAS will be finished and such information will be provided to other country teams (German, France, U.K., and U.S.).

We also will conduct the grand challenge run by NICAM with sub-km horizontal resolution, cooperating with the 3rd Strategic Field Project.

8.5. Publication, Presentation and Deliverables

(1) Journal Papers

1. S. Iga, H. Tomita, Y. Tsushima, M. Satoh, "Sensitivity of upper tropospheric ice clouds and their impacts on the Hadley circulation using a global cloud-system resolving model.", J. Clim., 24, 2666-2679. doi: 10.1175/2010JCLI3472.1 , 2011.

2. H. Yashiro, K. Sudo, S. Yonemura and M. Takigawa, "The impact of soil uptake on the global distribution of molecular hydrogen: chemical transport model simulation", Atmos. Chem. Phys., 11, 6701–6719, doi:10.5194/acp-11-6701-201, 2011.
3. A. T. Noda, K. Oouchi, M. Satoh, H. Tomita, "Quantitative assessment of diurnal variation of tropical convection simulated by a global nonhydrostatic model without cumulus parameterization", J. Clim., accepted, 2012.
4. M. Satoh, S. Iga, H. Tomita, Y. Tsushima, A. T. Noda, "Response of upper clouds due to global warming tested by a global atmospheric model with explicit cloud processes", J. Clim., 25, 2178-2191, 2012.
5. M. Yoshizaki, K. Yasunaga, S.-I. Iga, M. Satoh, T. Nasuno, A. T. Noda, H. Tomita , "Why do super clusters and Madden Julian Oscillation exist over the equatorial region? " , SOLA, 8, 33-36, doi:10.2151/sola.2012-009, 2012.
6. M. Satoh, K. Oouchi, T. Nasuno, H. Taniguchi, Y. Yamada, H. Tomita, C. Kodama, J. Kinter III, D. Achuthavarier, J. Manganello, B. Cash, T. Jung, T. Palmer, N. Wedi, "The Intra-Seasonal Oscillation and its control of tropical cyclones simulated by high-resolution global atmospheric models", Clim. Dyn., accepted, DOI 10.1007/s00382-011-1235-6, 2011.
7. Y. Yamada, K. Oouchi, M. Satoh, A. T. Noda and H. Tomita, "Sensitivity of tropical cyclones to large-scale environment in a global non-hydrostatic model with explicit cloud microphysics", Nova Science Publishers, Inc., (Eds. K. Oouchi and H. Fudeyasu), accepted, 2011.
8. K. Oouchi, H. Taniguchi, T. Nasuno, M. Satoh, H. Tomita, Y. Yamada, M. Ikeda, R. Shirooka, H. Yamada, K. Yoneyama, "A prototype quasi real-time intra-seasonal forecasting of tropical convection over the warm pool region: a new challenge of global cloud-system-resolving model for a field campaign", Nova Science Publishers, Inc., (Eds. K. Oouchi and H. Fudeyasu), accepted, 2011.
9. P. A. Dirmeyer, B. A. Cash, J. L. Kinter III, T. Jung, L. Marx, M. Satoh, C. Stan, H. Tomita, P. Towers, N. Wedi, D. Achuthavarier, J. M. Adams, E. L. Altshuler, B. Huang, E. K. Jin and J. Manganello, "Simulating the diurnal cycle of rainfall in global climate models: Resolution versus parameterization", Clim. Dyn., in press, DOI10.1007/s00382-011-1127-9, 2011.
10. M. Satoh, S. Iga, H. Tomita, Y. Tsushima, "Response of upper clouds due to global warming tested by a global atmospheric model with explicit cloud processes", J. Climate, accepted, 2011.
11. Y. Niwa, H. Tomita, M. Satoh, R. Imasu, "A Three-Dimensional Icosahedral Grid Advection Scheme Preserving monotonicity and Consistency with Continuity for Atmospheric Tracer Transport", J. Meteor. Soc. Japan, 89, 255-268, 2011.

(2) Conference Papers

1. H. Tomita, M. Satoh, H. Miura and Y. Niwa, "Current status of nonhydrostatic modeling at NICAM", ECMWF Workshop on Non-hydrostatic Modelling, page 171-182, ECMWF, 2011.

(3) Invited Talks

1. H. Tomita, "Research of tropical meteorology and climate using K computer", Singapore-Japan High Performance Computer Workshop, Singapore, 27-28 Feb.2012.
2. H. Tomita, "The current NICAM dynamical core and future plan for its improvement", Computing in Atmospheric Sciences 2011 (CAS2K11), Annecy, Sep. 11-14, 2011.

(4) Posters and presentations

1. H. Tomita, "Development of moist-LES model in RIKEN/AICS", The 14th international specialist meeting on the next generation models on climate change and sustainability for advanced high performance computing facilities Hawaii, 12-15 Mar., 2012.
2. H. Tomita, "Present Status of K Computer and activity toward the Exa-Scale Computing", The 14th international specialist meeting on the next generation models on climate change and sustainability for advanced high performance computing facilities Hawaii, 12-15 Mar., 2012.
3. H. Tomita, "Scale library : toward the global LES", Open Science Conference on WRCP, Denver, 24-28 Oct. 2011.
4. H. Tomita, "Recent outcomes from NICAM", 6th EU-Japan Workshop on Climate Change Research, Brussels, 10-11 Oct. 2011.
5. S. Iga, "Further simulation of aqua-planet experiment using NICAM", CFMIP/GCSS/EUCLIPSE Meeting on Cloud Processes and Climate Feedbacks, Exeter, 6-10 June, 2011.
6. S. Iga, M. Tsugawa, H. Tomita, "A generalization and combination of map projection on a sphere and its application to global grid system", Workshop on Numerical Methods for Scale Interactions 2011, Hamburg, 21-23 Sep. 2011.
7. Y. Miyamoto, and T. Takemi, "A transition mechanism for the spontaneous axisymmetric intensification of tropical cyclones", 5th Korea-Japan-China Joint Conference on Meteorology, Busan, Korea, 2011.

(5) Patents and Deliverables

- None

RIKEN Advanced Institute for Computational Science
7-1-26, Minatojima-minami-machi,
Chuo-ku, Kobe, Hyogo 650-0047, JAPAN
TEL : +81-(0)78-940-5555
FAX : +81-(0)78-304-4956
http://www.aics.riken.jp/

RIKEN Advanced Institute for Computational Science

RIKEN 2012-065