How to Use Fujitsu New Supercomputers of Information and Communication Center (ITC) of Nagoya University

June 1, 2009

Tatsuki Ogino

Three kinds of New Supercomputers of Information and Communication Center (ITC) of Nagoya University were partly begun to operate on March 18, 2009 and will be fully operated from October 1, 2009. They are composed of scalar-parallel supercomputers, Fujitsu M9000 (S1), HX600 (S2) and FX1 (S3) as follows,

S1 (SPARC Enterprise M9000): 3 node x 128 core
    Host name: sp1.cc.nagoya-u.ac.jp
    Large shared memory
    3 nodes x 128 cores (1 node is FX1's front-end)
    Per node: Performance: 1.28TFlops, memory: 1TB

S2 or Application server: HX600 64 node x 4 cpu x 4 core (Opteron)
                          HX600 160 node x 4 cpu x 4 core (from Oct. 1, 2009)
    Host name: sp2.cc.nagoya-u.ac.jp
    Cluster type computer, node uses shared memory
    160 nodes x 16 cores
    Per node: Performance: 160GFlops, memory: 64GB

S3: FX1 256 node x 4 core (SPARC)
    FX1 768 node x 4 core (from Oct. 1, 2009)
    Large distributed memory computer in connection with Next Generation
    Supercomputer
    768 nodes x 4 cores
    Per node: Performance: 40GFlops, memory: 32GB,
    Memory bandwidth: 40GB/s

There, S1 is shared memory machine and also front-end processor to use S3 (FX1). You need to connect with S1 to use S1 and S3 and you need to connect with S2 to use S1. At that time, you need to do "login" by SSH. FX1 corresponds to
    ssh a41456a@sp1.cc.nagoya-u.ac.jp

ssh a41456a@sp2.cc.nagoya-u.ac.jp

The previous files in a large capacity disc move to
　　/large/a41456a/
where a41456a is user ID.
You can use "jstat" as well as "qstat" in order to check the status jobs, and use qdel
(qdel –l job_number) to cancel a job. Moreover, you need to use "sftp" in place of "ftp"
and it becomes slower. Binary format is big endian for S1 and S2, and little endian for
S2.

(A) How to use M9000 and FX1
First please login S1 (M9000), then you can use S1 (M9000) and S3 (FX1). It is
available batch Job only for FX1.

　　ssh a41456a@sp1.cc.nagoya-u.ac.jp

Memory per core is Maximum 7 GBfor FX.
　 FX1　　1 node = 4 cores
　　　　Maximum memory for node　　　28GB
　　　　Maximum memory for core　　　　7GB

It is usually recommended to use auto-parallel tool like thread or OpenMP for node(=4
cores) in FX1.

Example of Compile and Execution
One need to compile by S1 (M9000) to make execution file mearthd3dd2n016.fx1, then
execute the code with 64 cores (16 process parallel and 4 thread parallel) by FX1.

mpifrt mearthd3dd2n016.f –o mearthd3dd2n016.fx1 –Kimpact –Z mpilist
cp mearthd3dd2n016.fx1 progmpi
qsub mpiex_fx0064s4.sh

se000% more mpiex_fx0064s4.sh
@$-q f64 –lp 4 –lP 16 –eo –o pexecmpi0064s4.out

@$-lm 8.0gb -cp 1:00:00
cd ./gridtest2/
mpiexec -n 16 ./progmpi

Moreover, you can use prefetch with the option, -Kimpact -Kprefetch_model=FX1, which is best option to obtain the fastest execution file.
mpifrt prog.f -o prog.fx1 -Kimpact -Kprefetch_model=FX1 -Z mpilist

One cannot use f-Kimpac for flat-MPI.
mpifrt progmpi712bb4a.f -o progmpi64 -Kprefetch_model=FX1 -Z mpilist
qsub mpiex_fx0064s1.sh

se000% more mpiex_fx0064s1.sh
@$-q f64 -lp 1 -lP 64 -eo -o pexecmpi0064s1.out
@$-lm 7.0gb -cp 24:00:00
cd ./vpp05a/mearthb3/
mpiexec -n 64 ./progmpi64

(B) How to use HX600
Firstly, you need to connect with S2 (HX600)

ssh a41456a@sp2.cc.nagoya-u.ac.jp

Compile and Execution of single cpu job.
frt -o prog prog.f
qsub exeh16.sh

se000% more exeh16.sh
@$-q h16 -lp 16 -eo -o sexec016.out
@$-lm 8.0gb -cp 1:00:00
setenv   parallel 16
cd ./mhdta/hx600/
./prog

Compile and Execution of MPI Parallel Program

128 core of 32 nodes with 4 cores = 4 threads

```
mpifrt progmpi.f -o progmpi -Kparallel -Z mpilist
qsub mpiex_fx0128s4.sh
```

```
se000% more mpiex_fx0128s4.sh
#   @$-q h128 -lp 4 -lP 32 -eo -o pexecmpi.out
#   @$-lm 12.0gb -cp 1:00:00
cd ./gridtest2/hx600/
mpiexec -n 32 ./progmpi
```

(C) How to use M9000

S1 (M9000) is upgrade version of PRIMEPOWER HPC2500 and so you can use S1 as
  you used to use HPC2500. You need usually to use S1 for data analysis and
  visualization by using simulation data which obtained by FX1. You can also S1 by
  TSS as HPC2500.

Compile and Execution of job by TSS in S1 (M9000)

```
frt prog.f -o prog
prog
```

Compile and Execution of MPI Parallel Program as Batch Job

```
mpifrt progmpi.f -o progmpi -Kparallel -Z mpilist
qsub mpiex_m64s4.sh
```

```
se000% more mpiex_m0064s4.sh
#   @$-q m64 -lp 4 -lP 16 -eo -o pexecmpi0064s4.out
#   @$-lm 10.0gb -cp 1:00:00
cd ./gridtest2/m9000/
mpiexec -n 16 ./progmpi
```

_____

Homepage to use New Supercomputer System of ITC, Nagoya University is located at
   the following URLs.

Documents are Japanese, however you can understand the important information.
Service of ITC

http://www2.itc.nagoya-u.ac.jp/center/index.html

Document to use new system

http://www2.itc.nagoya-u.ac.jp/sys_riyou/manual.htm
_____