Hands-on: NPB-MZ-MPI / BT

VI-HPS Team

Tutorial exercise objectives

- Familiarize with usage of Score-P, Cube, Scalasca & Vampir
  - Complementary tools' capabilities & interoperability
- Prepare to apply tools productively to *your* application(s)
- Exercise is based on a small portable benchmark code
  - Unlikely to have significant optimization opportunities

Local setup

- Load environment modules
  - Required for each shell session

```
% module load intel/16.0.3
% module load mxm/3.3.3002
% module load fca/2.5.2393
% module load bullxmpi_mlx/bullxmpi_mlx-1.2.8.3
```

- Copy & extract the tutorial sources

```
% cp /work/kg0166/PATworkshop2016/NPB3.3-MZ-MPI.tar.gz .
% tar zxvf NPB3.3-MZ-MPI.tar.gz
% cd NPB3.3-MZ-MPI
```

NPB-MZ-MPI suite

- The NAS Parallel Benchmark suite (MPI+OpenMP version)
  - Available from http://www.nas.nasa.gov/Software/NPB
  - 3 benchmarks in Fortran77
  - Configurable for various sizes & classes
- Move into the NPB3.3-MZ-MPI root directory

```
% ls
bin/    common/  jobscript/  Makefile  README.install   SP-MZ/
BT-MZ/  config/  LU-MZ/      README    README.tutorial  sys/
```

- Subdirectories contain source code for each benchmark
  - Plus additional configuration and common code
- The provided distribution has already been configured for the tutorial, such that it is ready to "make" one or more of the benchmarks and install them into a (tool-specific) "bin" subdirectory

Building an NPB-MZ-MPI benchmark

```
% make
  ========================================
  =      NAS PARALLEL BENCHMARKS 3.3     =
  =      MPI+OpenMP Multi-Zone Versions  =
  =      F77                             =
  ========================================

  To make a NAS multi-zone benchmark type

         make <benchmark-name> CLASS=<class> NPROCS=<nprocs>

  where <benchmark-name> is "bt-mz", "lu-mz", or "sp-mz"
        <class>          is "S", "W", "A" through "F"
        <nprocs>         is number of processes

   [...]

 ************************************************************
 * Custom build configuration is specified in config/make.def  *
 * Suggested tutorial exercise configuration for Mistral:      *
 *       make bt-mz CLASS=B NPROCS=4                           *
 ************************************************************
```

- Type "make" for instructions

Building an NPB-MZ-MPI benchmark

```
% make bt-mz CLASS=B NPROCS=4
make[1]: Entering directory `BT-MZ'
make[2]: Entering directory `sys'
icc  -o setparams setparams.c -lm
make[2]: Leaving directory `sys'
../sys/setparams bt-mz 4 B
make[2]: Entering directory `../BT-MZ'
mpif77 -c  -O3 -qopenmp       bt.f
                                    […]
mpif77 -c  -O3 -qopenmp       mpi_setup.f
cd ../common;  mpif77 -c  -O3 -qopenmp        print_results.f
cd ../common;  mpif77 -c  -O3 -qopenmp        timers.f
mpif77 -O3 -qopenmp   -o ../bin/bt-mz_B.4 bt.o
 initialize.o exact_solution.o exact_rhs.o set_constants.o adi.o
 rhs.o zone_setup.o x_solve.o y_solve.o  exch_qbc.o solve_subs.o
 z_solve.o add.o error.o verify.o mpi_setup.o ../common/print_results.o
 ../common/timers.o
make[2]: Leaving directory `BT-MZ'
Built executable ../bin/bt-mz_B.4
make[1]: Leaving directory `BT-MZ'
```

- Specify the benchmark configuration
  - benchmark name: **bt-mz**, lu-mz, sp-mz
  - the number of MPI processes: NPROCS=**4**
  - the benchmark class (S, W, A, B, C, D, E): CLASS=**B**

Shortcut: % **make suite**

NPB-MZ-MPI / BT (Block Tridiagonal Solver)

- What does it do?
  - Solves a discretized version of the unsteady, compressible Navier-Stokes equations in three spatial dimensions
  - Performs 200 time-steps on a regular 3-dimensional grid
- Implemented in 20 or so Fortran77 source modules

- Uses MPI & OpenMP in combination
  - Proposed hands-on setup on Mistral:
    - 2 compute nodes of "compute" partition
    - 4 MPI processes with 12 OpenMP threads each
  - bt-mz_B.4 should run in around 11 seconds

NPB-MZ-MPI / BT reference execution

```
% cd bin
% cp ../jobscript/mistral/reference.sbatch .
% vim reference.sbatch
% sbatch reference.sbatch
% less mzmpibt.o<job_id>
 NAS Parallel Benchmarks (NPB3.3-MZ-MPI) - BT-MZ MPI+OpenMP Benchmark
 Number of zones:  16 x  16
 Iterations:  200   dt:   0.000300
 Number of active processes:     4
 Total number of threads:       48   ( 12.0 threads/process)

 Time step    1
 Time step   20
  [...]
 Time step  180
 Time step  200
 Verification Successful

 BT-MZ Benchmark Completed.
 Time in seconds = 10.79
```

- Copy jobscript
- Set project account
- Launch as a hybrid MPI+OpenMP application

Hint: save the benchmark output (or note the run time) to be able to refer to it later

Tutorial exercise steps

- Edit config/make.def to adjust build configuration
  - Modify specification of compiler/linker: MPIF77
  - See next slide for details
- Make clean and build new tool-specific executable

```
% make clean
% make bt-mz CLASS=B NPROCS=4
Built executable ../bin.$(TOOL)/bt-mz_B.4
```

- Change to the directory containing the new executable before running it with the desired tool configuration

```
% cd bin.$(TOOL)
% cp ../jobscript/mistral/$(TOOL).sbatch .
% vim $(TOOL).sbatch
% sbatch $(TOOL).sbatch
```

NPB-MZ-MPI / BT: config/make.def

```
#                SITE- AND/OR PLATFORM-SPECIFIC DEFINITIONS.
#
#---------------------------------------------------------------------

#---------------------------------------------------------------------
# Configured for generic MPI with INTEL compiler
#---------------------------------------------------------------------
#OPENMP  = -fopenmp      # GCC compiler
OPENMP = -qopenmp        # Intel compiler

...
#---------------------------------------------------------------------
# The Fortran compiler used for MPI programs
#---------------------------------------------------------------------
MPIF77 = mpif77  # Intel compiler

# Alternative variant to perform instrumentation
#MPIF77 = scorep --user mpif77

# PREP is a generic preposition macro for instrumentation preparation
#MPIF77 = $(PREP) mpif77
...
```

Default (no instrumentation)

Hint: uncomment a compiler wrapper to do instrumentation