EXE – PowerFlow – Execution Guidelines for running applications in aggregated environment using ScaleMP's vSMP Foundation

Overview

PowerFlow's simulation phase is a multi-process application that uses MPI for inter-process communication. HP-MPI has been set as the default MPI for the PowerFlow application. In addition upcoming versions of PowerFlow would support Intel-MPI/MPICH2 as well.

Guidelines for running PowerFlow with Intel-MPI / MPICH2 would be available as soon as the new version of PowerFlow is released.

Running PowerFlow with ramfs

Since PowerFlor jobs perform a lot of scratch IO, running with a ramfs may yield shorter runtimes. In order to run with ramfs, perform the following:

1) Make sure that /ramfs is mounted, and if not mount it as follows:

```
sudo mkdir /ramfs

sudo mount /dev/ram -t ramfs /ramfs -o noatime

sudo chmod 777 /ramfs
```

2) Copy the input into /ramfs and run from there, so the '*.lgi' file would be created in /ramfs

Separating between the discretesizer and Simulation

The discretesizer part is not intended to run on a large number of cores; while the simulation part is. In order to run it optimally, once should run the discretesizer phase separately with up to 8 cores, and bound this part of the job to a single board, by using:

```
numactl –cpunodebind=2,3 [run script]
```

The above example would bind the process to the 2nd board.

After the discretesizer phase is complete, run the simulation part using MPI as described below.

Number of Ranks

As PowerFlow always creates one additional control process; when the target is to launch it using N cores, please specify 'N-1' as the number of MPI ranks to use.

Running PowerFlow with HP-MPI

HP-MPI has a built-in mechanism for assigning MPI processes to specific CPUs. Process placement is controlled by environment variables named **MPI_BIND_MAP** and **MPIRUN_OPTIONS**. When these variables are not set, process placement will not be performed.

Environment variables – HP-MPI

If you are running Fluent with HPMPI, you should set the following environment variables prior to running Fluent to yield the optimal performance:

```
export MPI_BIND_MAP=0,1,2,3,4,5,6,7  (For example)

export MPIRUN_OPTIONS="-cpu_bind=map_cpu,v"

export HPMP_FRAGSIZE=131072

export MPI_SHMEMCNTL=16,24000000,4000000
```

MPI_BIND_MAP specifies a list of CPUs to which MPI ranks will be bound. You should replace the list above with a list of integers, zero to #cpus-1.
For more information on HP-MPI CPU affinity settings, refer to the HP-MPI user's guide available from "http://docs.hp.com/en/B6060-96022/B6060-96022.pdf".

Sample script for HP-MPI

```
#!/bin/sh


export PATH=/opt/powerflow/4.2c/bin:$PATH
export LM_LICENSE_FILE=27007@<hostname>
export NCPUS=`grep -i processor /proc/cpuinfo | wc -l`
export MPI_BIND_MAP=`seq -f "%g," 0 $((NCPUS-1)) | tr -d "\n" | sed -e 's/,$//'`
export MPIRUN_OPTIONS="-cpu_bind=map_cpu,v"
export HPMP_FRAGSIZE=131072
export MPI_SHMEMCNTL=16,24000000,4000000

# Always, use Nproc-1 processes as PowerFlow creates one additional control process.
np=15

taskset 0xff exaqsub -nprocs $np external_case_2.cdi
```

Running PowerFlow with MPICH2 tuned for vSMP

Would be available as soon as the new version of PowerFlow is released.