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

Overview

StarCD is a multi-process application that uses MPI for inter-process communication. HP-MPI has been set as the default MPI for the StarCD application. In addition, StarCD supports MPICH1 as well.
Running StarCD with HP-MPI on the aggregation platform would result in good performance. For achieving even better performance (especially for smaller jobs), using MPICH1 tuned for vSMP Foundation (a.k.a. vSMPICH) may yield a performance improvement of ~ 10%.

The following execution guidelines are relevant for StarCD versions 3 and 4.

Running StarCD 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

You should set the following environment variables prior to running 'starlaunch' to yield the optimal performance:

```
export STARFLAGS="-mpi=hp"
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"*.

Handling Excessive System-Calls

HP-MPI performs excessive system calls to sched_yield, which are unnecessary when running with vSMP Foundation; hence running with a pre-load library which eliminates such calls would result in better performance.
The pre-load library (libnoyield.so) can be downloaded from ScaleMP's support portal (http://support.scalemp.com) under Libraries / System Calls Pre-Load Libraries.
In the run-script for StarCD please add the pre-load library by adding:

```
export LD_PRELOAD=[full path to libnoyield.so]:$LD_PRELOAD
```

Running StarCD with Ram-Drive

StarCD may perform better if you copy your input files to a directory in a RAMFS mount, and 'cd' there before running starlaunch. This is especially useful for large-scale jobs.
In order to run with Ram-Drive, perform the following

```
sudo mkdir –p /ramfs
sudo mount -t ramfs ramfs /ramfs -o noatime
sudo chmod 777 /ramfs
mkdir /ramfs/starcd
cp *.ccm PROINIT *.echo param.prp *.mdl /ramfs/starcd
cp *.info parm.inc *.prob *.run *.geom  /ramfs/starcd
cd /ramfs/starcd
starlaunch
```

After the run terminates, you will need to copy the output files back to a disk-backed file system; **files in RAMFS do not survive system reboots**.

Sample script for StarCD with HP-MPI and Ram-Drive

```
export STARFLAGS="-mpi=hp"
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
export LD_PRELOAD=`pwd`/libnoyield.so :$LD_PRELOAD

sudo mkdir –p /ramfs
sudo mount -t ramfs ramfs /ramfs -o noatime
sudo chmod 777 /ramfs
mkdir /ramfs/starcd
cp *.ccm PROINIT *.echo param.prp *.mdl /ramfs/starcd
cp *.info parm.inc *.prob *.run *.geom  /ramfs/starcd
cd /ramfs/starcd
starlaunch
```

Running StarCD with MPICH1 tuned for vSMP Foundation

StarCD is provided with support for MPICH1 (as well as other MPI implementations). In order to configure StarCD to work with MPICH1 tuned for vSMP Foundation, perform the following:

When installing STARCD, be sure to select the MPICH library, and omit LAM

StarCD uses MPICH1 version 1.2.4 which gets installed into the following folder:

<starcd-install-dir>/MPICH/1.2.4/linux64_2.4-gcc_3.2.2-glibc_2.2.5-dso/ch_shmem

In order to clear that folder and install MPICH1 turned for vSMP Foundation (provided in mpich1-1.2.7-vSMP0.1_20090424.tgz) run the following script:

```
export INS_DIR=<starcd-install-dir>/MPICH/1.2.4/linux64_2.4-gcc_3.2.2-glibc_2.2.5-dso/ch_shmem
echo "Removing MPICH-1.2.4"
rm -rf  $INS_DIR
echo “Installing MPICH1 1.2.7 tuned for vSMP Foundation”
sudo tar xfz mpich1-1.2.7-vSMP0.1_20090424.tgz -C  /
echo “Pointing starCD to use new MPICH1”
ln -sf  /opt/vsmpich1 $INS_DIR
```

Once the above is done, MPICH1 has been successfully replaced.

Environment variables – MPICH1

```
export VSMP_PLACEMENT=PACKED
export VSMP_MEM_PIN=YES
```

Sample script for MPICH1 tuned for vSMP

```
export STARFLAGS="-mpi=mpich"
export VSMP_PLACEMENT=PACKED
export VSMP_MEM_PIN=YES

sudo mkdir –p /ramfs
sudo mount -t ramfs ramfs /ramfs -o noatime
sudo chmod 777 /ramfs
mkdir /ramfs/starcd
cp *.ccm PROINIT *.echo param.prp *.mdl /ramfs/starcd
cp *.info parm.inc *.prob *.run *.geom  /ramfs/starcd
cd /ramfs/starcd
starlaunch
```