LS-DYNA Installation Best Practices

HPC ADVISORY COUNCIL
NETWORK OF EXPERTISE
BEST PRACTICES

1. Introduction:

The following best practices document is provided for running LS-DYNA with Mellanox HPC-X Software Toolkit.

2. Application Description:

One of the most demanding applications of automotive design is crash simulation (full-frontal, offset-frontal, angle-frontal, side-impact, rear-impact and more). Crash simulations, while performed very early in the development process, are validated very late in the development process once the vehicle is completely built. The more sophisticated and complex the simulation, the more parts and details can be analyzed. Automotive makers increase their dependency for car crash simulations throughout the design process while reducing the need for real prototypes, thus achieving faster time to market with less cost associated with the design phase.

LS-DYNA is a general purpose structural and fluid analysis simulation software package capable of simulating complex real world problems. It is widely used in the automotive industry for crashworthiness, occupant safety and metal forming and also for aerospace, military and defense and consumer products.

3. Version Information:

Download Mellanox HPC-X Software Toolkit (latest public version: HPC-X v1.5): http://www.mellanox.com/products/hpcx

Alternatively, the latest Mellanox HPC-X Software Toolkit v1.6 can be downloaded at: http://bgate.mellanox.com/products/hpcx/beta/v1.6

Download LS-DYNA: http://www.lstc.com/download/ls-dyna

Please use this executable for HPC-X:

ls-dyna_mpp_s_r8_0_0_95359_x64_redhat54_ifort131_sse2_openmpi183.tar.gz

Download LS-DYNA "neon_refined_revised" and "3 Vehicle Collision" benchmarks at: http://topcrunch.org/benchmark_problems.sfe

4. Prerequisites:

- Dell PowerEdge R730 32-node (896-core) "Thor" cluster
- Dual-Socket 14-Core Intel E5-2697v3 @ 2.60 GHz CPUs
- Memory: 64GB memory, DDR4 2133 MHz
- BIOS Settings
  - » Power Management to Maximum Performance
  - » Memory Snoop Mode: Home Snoop
  - » Turbo enabled
- OS: RHEL 6.5
- MLNX_OFED_LINUX-3.2-1.0.1.1 InfiniBand SW stack
- Hard Drives: 2x 1TB 7.2 RPM SATA 2.5" on RAID 1
- Mellanox ConnectX-4 EDR 100Gb/s InfiniBand Adapters
- Mellanox Switch-IB SB7700 36-port EDR 100Gb/s InfiniBand Switch
- MPI:
  - » hpcx-v1.5.370-icc-MLNX_OFED_LINUX-3.2-1.0.1.1-redhat6.5-x86_64
  - » Application:
  - » LS-DYNA 8.0.0 (builds 95359, 95610), Single Precision
- Benchmarks:
  - » 3 Vehicle Collision
  - » Neon refined revised

5. Installation:

5.1 Installation for Mellanox HPC-X Toolkit

Download and extract HPC-X tar ball on each compute node according to the README. The following is an installation overview:

1. Install HPC-X

a) Extract tarball and set environment variable to HPC-X location

```
% tar zxvf hpcx.tar

% cd hpcx

% export HPCX_HOME=$PWD
```

b) Load KNEM (High-Performance Intra-Node Communication Device)

Please run following commands on all cluster nodes to enable KNEM intra-node device:

```
# insmod $HPCX_HOME/knem/lib/modules/$(uname -r)/knem.ko
```

```
# chmod 666 /dev/knem
```

5.2 Changes to make LS-DYNA work

Since HPC-X is based on Open MPI version 1.10.x, and LS-DYNA is built and linked with an older Open MPI version 1.8, the following steps are needed for LS-DYNA to work together with HPC-X:

```
$ cd /opt/hpcx-v1.5.370-icc-MLNX_OFED_LINUX-3.2-1.0.1.1-redhat6.5-x86_64/ompi-v1.10/lib

$ ln -s libmpi_usempif08.so libmpi_usempif08.so.0

$ ln -s libmpi_usempi_ignore_tkr.so libmpi_usempi_ignore_tkr.so.0

$ ln -s libmpi_mpifh.so libmpi_mpifh.so.2

$ ln -s libmpi.so libmpi.so.1
```

5.3 Specifying pfile for domain decomposition

User can specify their choice for domain decomposition. The following is the pfile used in this benchmark example.

pfile_3cars:

```
gen { nodump nofull nod3dump nofail nobeamout }

decomp { sy 2 }

dir { local / local_shm_dir/3cars }
```

pfile_neon_refined_revised:

```
general { nodump nofull nod3dump nofail nobeamout }

decomp { sy 2 }

dir { local /local_shm_dir/neon_refined_revised }
```

6. Running LS-DYNA with HPC-X:

To run with HPC-X, source the compiler and MPI runtime settings:

```
$ source /opt/intel/compilers_and_libraries/linux/bin/compilervars.sh intel64

$ module use /opt/hpcx-v1.5.370-icc-MLNX_OFED_LINUX-3.2-2.0.0.0-redhat6.5-x86_64/modulefiles

$ module load hpcx
```

The following command run the neon_refined_revised benchmark using HPC-X on 336 cores:

```
mpirun -hostfile /home/lstc/logs-nodes/nodes.14316 -np 336 -cpu-set 0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27 -report-bindings -display-map -mca btl_sm_use_knem 1 -x MXM_SHM_KCOPY_MODE=knem -x MALLOC_MMAP_MAX_=0 -x MALLOC_TRIM_THRESHOLD_=-1 -mca coll_fca_enable 0 -mca coll_hcoll_enable 0 -mca pml yalla -mca mtl_mxm_np 0 -x MXM_TLS=ud,shm,self -x MXM_SHM_RNDV_THRESH=32768 -mca btl openib,sm,self -mca btl_openib_if_include mlx5_0:1 -x MXM_RDMA_PORTS=mlx5_0:1 -mca rmaps_base_dist_hca mlx5_0:1 -x fca_ib_dev_name=mlx5_0 -x HCOLL_MAIN_IB=mlx5_0:1 -x HCOLL_IB_IF_INCLUDE=mlx5_0:1 -mca rmaps_base_mapping_policy slot  /dev/shm/lstc/ls-dyna_mpp_s_r8_0_0_95359_x64_redhat54_ifort131_sse2_openmpi183 i=/dev/shm/lstc/neon_refined_revised/neon.refined.rev01.k memory=400m p=/dev/shm/lstc/neon_refined_revised/pfile_pak memory2=400m  ncpu=336
```

This command explicitly calls out the option for the mlx5_0 device, which is the ConnectX-4 HCA.

The following explains the options used in the command. CPU binding and mapping options. Mapping is based on the locality of the mlx5_0 HCA:

```
-cpu-set 0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27

-mca rmaps_base_mapping_policy slot

-mca rmaps_base_dist_hca mlx5_0:1
```

Process placement and core binding mapping:

```
-report-bindings -display-map
```

For enabling KNEM support for intranode communications:

```
-mca btl_sm_use_knem 1 -x MXM_SHM_KCOPY_MODE=knem
```

For FCA and FCA 3.0+ (HCOLL) support:

```
-mca coll_fca_enable 0 -mca coll_hcoll_enable 0
```

Tuning flags for system memory management:

```
-x MALLOC_MMAP_MAX_=0 -x MALLOC_TRIM_THRESHOLD_=-1
```

Explicitly enabling MXM with yalla PML in HPC-X (which is the default for HPC-X.

```
-mca pml yalla -mca mtl_mxm_np 0 -x MXM_TLS=ud,shm,self -x MXM_SHM_RNDV_THRESH=32768 -mca btl openib,sm,self -mca btl_openib_if_include mlx5_0:1 -x MXM_RDMA_PORTS=mlx5_0:1
```

Since LS-DYNA creates and destroys many MPI

communicators, use the following flags to enable context caching when using HCOLL:

```
-x HCOLL _ CONTEXT _ CACHE _ ENABLE=1
-x HCOLL _ POLLING _ LEVEL=1
```

HPC ADVISORY COUNCIL
NETWORK OF EXPERTISE

350 Oakmead Pkwy, Sunnyvale, CA 94085
Tel: 408-970-3400 • Fax: 408-970-3403
www.hpcadvisorycouncil.com