Introduction to MPP version of LS-DYNA®



Jason Wang 08/10/2012

Livermore Software Technology Corporation



Copyright © 2005-2011 by LIVERMORE SOFTWARE TECHNOLOGY CORPORATION

Contents

Public Domain Benchmark Tests
Cluster Tuning and Job Administration
References



Public Domain Benchmark Tests



Copyright © 2005-2011 by LIVERMORE SOFTWARE TECHNOLOGY CORPORATION

Current Benchmark Tests

- Benchmark models are provided by LSTC
 - Three car impact model (794776 elements)
 - Refined Neon model (535068 elements)
 - Car to Car crash (2448596 elements)

Benchmark results for LS-DYNA[®] and MPP-DYNA

http://www.topcrunch.org

Top Crunch was sponsored by DARPA, and the sever is placed in San Diego where Professor David Benson now maintain and sponsor the site. Users download the models together with the pfile(s).



Copyright © 2005-2011 by LIVERMORE SOFTWARE TECHNOLOGY CORPORATION

Three Car Impact Model

A van crashes into the rear of a compact car, which, inturn, crashes into a midsize car. Vehicle models created by NCAC, and assembled into the input file and provided by LSTC.



LSTC

Three Car Impact Model

 The Top Crunch project makes it possible to compare different systems since users upload results. These can then be plotted against each other.



Refined Neon Car Model

 Frontal crash with initial speed at 31.5 miles/hour, model size 535k elements, simulation length: 150 ms. Model created by National Crash Analysis Center (NCAC) at George Washington University.



Refined Neon Car Model



This model is based on NCAC minivan model and created by Dr. Makino. Supplied by Dr. Tsay, LSTC, on Feb. 13, 2006, termination time modified per John Hallquist to .120 on March 7, 2006. It is two mini vans in frontal collision.

_STC



Cluster tuning and Job Administration



Copyright © 2005-2011 by LIVERMORE SOFTWARE TECHNOLOGY CORPORATION

Cluster Tuning and Job Administration

- Cluster tuning
- Cluster IO setup
- System tuning
- Numerical Consistency
- Setup usermat Executables
- Job Administration
- Report Traceback Information



Cluster Tuning

- LS-DYNA[®] explicit is CPU bounded application
- LS-DYNA[®] implicit is CPU, memory and IO bounded appliction
- Lots of message passing activities through network
- avoid collision between IO and message passing



Cluster Tuning

- Computing nodes, ie homogeneous
- Explicit/Implicit analysis
- Memory SDDR2, SDDR3, fully populated, etc
- Interconnect, computing, IO traffic
- Local, global and master storage, etc.
- Memory distribution (master, slave nodes)
- Use real memory instead of swap space



Copyright © 2005-2011 by LIVERMORE SOFTWARE TECHNOLOGY CORPORATION

- **Computing environment** 1. Shared central file server
- 2. Diskless computing nodes





Copyright © 2005-2011 by LIVERMORE SOFTWARE TECHNOLOGY CORPORATION

System setup 2 Computing environment

- 1. Shared central file system
- 2. Local disk on computing nodes

DYNA Computing Nodes



Network Connection

dir { rmlocal local /tmp/user/\$jobid global /shared/user transfer_files }
gen { nodump nobeamout }



File Server

Global files

10

R.

Copyright © 2005-2011 by LIVERMORE SOFTWARE TECHNOLOGY CORPORATION

System setup 3

- 1. Shared central file system
- 2. Shared central storage

T/O

gen { nobeamout }

3. Local disk on computing nodes



Network Connection

dir { rmlocal local /tmp/user/\$jobid global /shared/user rep /shared/storage/user/\$jobid transfer_files }



File Server

Copyright © 2005-2011 by LIVERMORE SOFTWARE TECHNOLOGY CORPORATION

System Tuning

• Turn off hardware hyper-threading and turn on boost mode from BIOS

 cat /proc/cpuinfo 	o fgrep "cpu MH"
cpu MHz	: 1400.000

powersave -f

.



Copyright © 2005-2011 by LIVERMORE SOFTWARE TECHNOLOGY CORPORATION

System Tuning

numactl --hardware

Get correct core order

Hyperthreading on?



System Tuning AMD 6200 series

 use only the core has its own FPU and L2 cache under Platform MPI

mpirun -np 4 -cpu_bind=v,MAP_CPU:0,2,4,6
or
mpirun -np 4 -cpu_bind=v,MAP_CPU:0,8,16,24

 Use AMD Open64 AVX enabled exe http://ftp.lstc.com/beta/mpp-lsdyna/mpp971/R6.1.0 (look for exec having the string "Open64avx").



Copyright © 2005-2011 by LIVERMORE SOFTWARE TECHNOLOGY CORPORATION

Consistency

- Fix number of cores during product development
- LSTC_REDUCE *CONTROL_MPP_IO_LSTC_REDUCE
- RCBLOG
 *CONTROL_MPP_DECOMPOSITION_RCBLOG



Copyright © 2005-2011 by LIVERMORE SOFTWARE TECHNOLOGY CORPORATION

Consistency using Hybrid

There is a consistent option (ncpu=-N) in LS-DYNA SMP version. Many customers used to run their jobs with the option in SMP era, even though there is about 10-15% performance penalty with the option.

LSTC added the option into LS-DYNA Hybrid version. So customers can use the option for getting consistent numerical result. However, there is a condition here. The condition is you need to fix the number of MPI processes at first.

For example, you select 12 MPI processes, then you can run your job in this way.

mpiexec <u>**-ppn M**</u>-np 12 mpp971hyb i=input memory=xxxm memory2=xxm <u>ncpu=-N</u> p=pfile

12 cores:	12 MPI pr	ocesses x	(1 OMP	thread	(1 nodes x 12	cores)	M=12,	N=1
24 cores:	12 MPI pr	ocesses ×	(2 OMP	threads	(2 nodes x 12	cores)	M= 6,	N=2
36 cores:	12 MPI pr	ocesses ×	3 OMP	threads	(3 nodes x 12	cores)	M= 4,	N=3
48 cores:	12 MPI pr	ocesses x	4 OMP	threads	(4 nodes x 12	cores)	M= 3,	N=4
					`	,		

72 cores: 12 MPI processes x 6 OMP threads (6 nodes x 12 cores) M = 2, N = 6

Then you can get consistent results with 12c, 24c, 36c, 48c, 60c, and 72c.



Copyright © 2005-2011 by LIVERMORE SOFTWARE TECHNOLOGY CORPORATION

Setup usermat Executables

•LS-DYNA support user defined interface to allow users developing their functions

It is called "usermat" interafce

Two ways interfacing
Static -- relinking
Dynamic -- sharelib interface



Copyright © 2005-2011 by LIVERMORE SOFTWARE TECHNOLOGY CORPORATION

Setup usermat Executables static

 Download usermat package from <u>http://ftp.lstc.com/objects/</u> password: computer1

- Doenload README.first from <u>http://ftp.lstc.com/user</u> password: computer
- Add user routine in Fortran source provided, dyn21.f or dyn21b.f and use the same compiler as described in README.first
- Relink the exe using the same compiler



Copyright © 2005-2011 by LIVERMORE SOFTWARE TECHNOLOGY CORPORATION

Setup usermat Executables dynamic

 Download usermat package from (with string "sharelib") <u>http://ftp.lstc.com/objects/</u> password: computer1

- Add user routine in Fortran source provided, dyn21.f or dyn21b.f
- Use any compiler to compile the
- Use provided Makefile to create sharelib lib*.so



Copyright © 2005-2011 by LIVERMORE SOFTWARE TECHNOLOGY CORPORATION

Setup usermat Executables dynamic

 Check libraries dependency by "ldd -r mpp971_exe"

hostname [36]% ldd mpp971

 $libstdc++.so.6 => /usr/lib64/libstdc++.so.6 (0x000003861a00000) \\ libpthread.so.0 => /lib64/libpthread.so.0 (0x00000384fc00000) \\ librt.so.1 => /lib64/librt.so.1 (0x000003850000000) \\ libmpp971_d_72654.74564_usermat.so => not found \\ libmpio.so.1 => /opt/hpmpi/lib/linux_amd64/libmpio.so.1 (0x00002ac91f787000) \\ libmpi.so.1 => /opt/hpmpi/lib/linux_amd64/libmpi.so.1 (0x00002ac91f8c7000) \\ libdl.so.2 => /lib64/libdl.so.2 (0x00000384f800000) \\ libm.so.6 => /lib64/libm.so.6 (0x00000384f400000) \\ libc.so.6 => /lib64/libc.so.6 (0x00000384f000000) \\ /lib64/ld-linux-x86-64.so.2 (0x00000384ec00000) \\ libgcc_s.so.1 => /lib64/libgcc_s.so.1 (0x00000385c200000) \\$



Setup usermat Executables dynamic

 Setup LD_LIBRARY_PATH point the usermat library and try ldd -r mpp971_exe again

Setenv LD_LIBRARY_PATH /the_path_of/lib:\$ LD_LIBRARY_PATH

hostname [36]% ldd mpp971

libstdc++.so.6 => /usr/lib64/libstdc++.so.6 (0x000003861a00000) libpthread.so.0 => /lib64/libpthread.so.0 (0x00000384fc00000) librt.so.1 => /lib64/librt.so.1 (0x000003850000000) libmpp971_d_72654.74564_usermat.so => ./libmpp971_d_72654.74564_usermat.so (0x00002b5473e78000)



Copyright © 2005-2011 by LIVERMORE SOFTWARE TECHNOLOGY CORPORATION

Job Administration

•setenv LSTC_MEMORY auto auto: for explicit analysis disable it if running implicit

•setenv LSTC_RSH your_remote_shell

•submit job with "jobid=\$JOB_NAME"



Copyright © 2005-2011 by LIVERMORE SOFTWARE TECHNOLOGY CORPORATION

Job Administration

"out of range velocity"

Please add following command line option for all jobs

mpirun -np 4 mpp971_exe i=input ... checknan=1



Debugging

Find the "nodelist" file which contains the list of hostname used for the current job.
 node001
 node002
 etc

- The error messages from MPP-DYNA can be different from LS-DYNA[®]
- To locate an error one often has to search each of the messag files mes#### in order to find any information. These files are written for each processor.
- The code will trap the segmentation violation (SEGV) and output the rank number. One could rerun the job and attach the debugger to the running thread and get the trace back map. This usually gives good information for changing input.

gdb path_to_mpp_code/mpp971 PID

- > continue
- SEGV
- > where



Copyright © 2005-2011 by LIVERMORE SOFTWARE TECHNOLOGY CORPORATION

Debugging

*** Error Memory is set 1235165 words short Current memory size 5000000 Increase the memory size by one of the following where #### is the number of words requested:
1) On the command line set - memory=####
2) In the input file define memory with *KEYWORD
i.e., *KEYWORD #### or *KEYWORD memory=####

 The memory unit is in WORD. For single precision is 4 Bytes/word and for double precision is 8 Bytes/word.

- LS-DYNA[®] explicit uses real memory to store all data. However, the amount of static memory requested is controlled by "memory=" option and the amount of dynamic memory is adjusted automatically.
- Please use "top" command to check the available memory in the system and you *DO NOT* want your job using swap space



Type of executable

under user account:

- ifort101, ifort120, pgi105, open64
- Intel, AMD SSE
- Intel avx
- AMD avx







Copyright © 2005-2011 by LIVERMORE SOFTWARE TECHNOLOGY CORPORATION

References MPI supported by LS-DYNA/MPP®

- MPICH 1.2.6 www.mcs.anl.gov/research/projects/mpi/mpich1
- MPICH 2-1.2.1 <u>www.mcs.anl.gov/research/projects/mpich2</u>
- OpenMPI 1.2.5 www.open-mpi/org
- Intel-MPI
 - HP-MPI user@ftp.lstc.com/mpp-dyna/HP-MPI
- Platform-MPI
- user@ftp.lstc.com/mpp-dyna/Platform_mpi



Copyright © 2005-2011 by LIVERMORE SOFTWARE TECHNOLOGY CORPORATION

References LS-DYNA/MPP ® Executables

- Official releases <u>user@ftp.lstc.com/mpp-dyna/ls971</u> computer
 - Beta tests beta@ftp.lstc.com/mpp-dyna keyboard



References

- General information
 Appendix L in LS-DYNA 970 User Manual
- CPU time depending on decomposition:
 Conf. 2000: 17.1-17.40
- Performance:Conf 2000: 18.29-18.44



Copyright © 2005-2011 by LIVERMORE SOFTWARE TECHNOLOGY CORPORATION

References

/Wainscott et al, 98/ Brian Wainscott, Jason Wang and Nielen Stander, LSTC.
 Page N 1-10. Proceedings of 5th International LS-DYNA User Conferences, 1998.

Galbraith et al, 2002/

