Introduction to LS-DYNA MPP

The Arup Campus, Blythe Gate, Blythe Valley Park, Solihull, West Midlands, B90 8AE tel: +44 (0) 121 213 3399 email: dyna.support@arup.com

2018



- Introduction to MPP LS-DYNA
- MPP decomposition methods
- Visualising decompositions
- Load balancing information
- MPP contacts
- Restart analysis



SMP (Symmetric Multi-Processing)

- Originated from the serial code
- Uses OpenMP[®] directives to split tasks into parallel threads
- Runs on computers with multiple identical cores with the cores and memory connected via a shared data bus
- <u>Consistent results with different cores</u> (consistency flag turned on!)
- Scalable up to ~ 8 CPUs

MPP (Massively Parallel Processing)

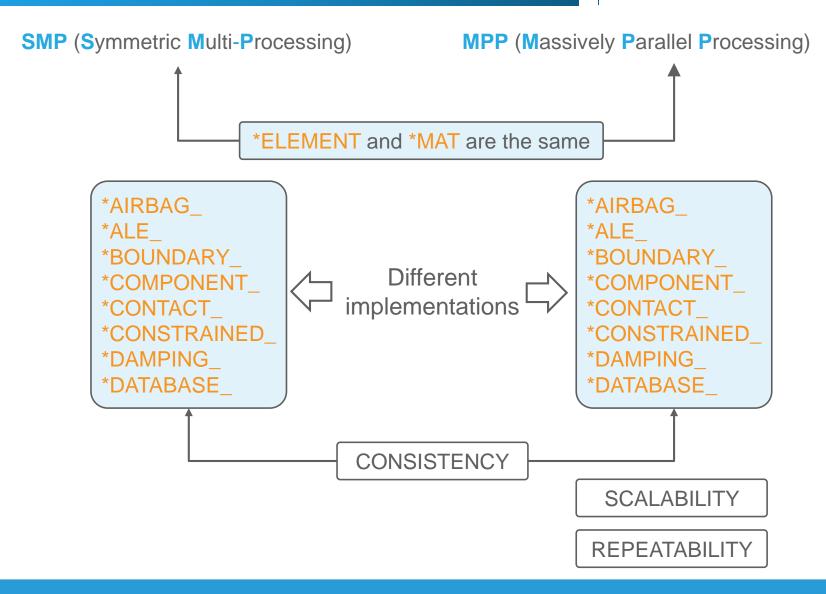
- Uses a message passing protocol to exchange information between the cores on a board or over a network
- MPP solver performs a domain decomposition of the problem...
- and distributes the sub-domains to different cores using MPI protocols for communication between the subdomains during analysis
- <u>Results change with different cores</u>
- Scalable >> 16 CPUs

$$\left[\mathsf{T}_{\mathsf{elapsed}} = \mathsf{T}_{\mathsf{cpu}} + \mathsf{T}_{\mathsf{sys}} + \mathsf{T}_{\mathsf{omp}}\right]$$

$$T_{elapsed} = T_{cpu} + T_{sys} + T_{mpp}$$









MPP domain decomposition involves dividing the model into several domains, which are done by the primary processor, and assigning each domain to a core.

The elements and nodes on the boundary of each domain transfer information to those in the other domain over a network connection using message passing protocols.

Factors that affect parallel performance:

Load Balance:

- Boundaries of the decomposed domains
- Variations between different element formulations and material models
- Treatment of contacts
- Special features used in the modelling

Communication:

- Memory/Cache system
- Interconnections
- MPI libraries
- Fortran compiler

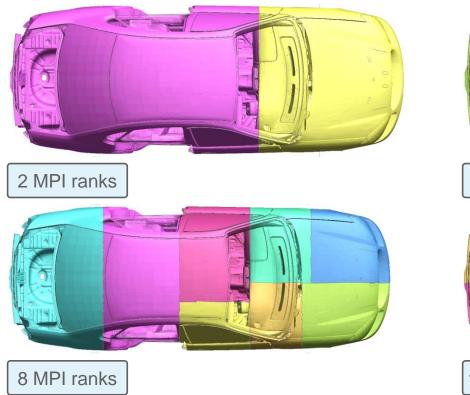


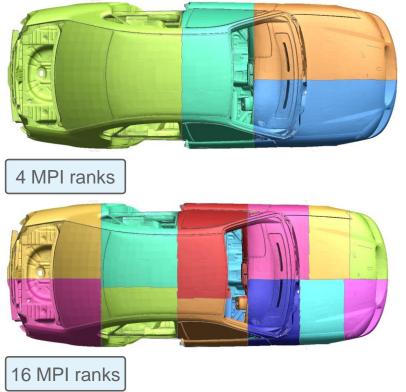
MPP Domain Decomposition

Webinar: MPP-DYNA

Recursive Coordinate Bisection (RCB)

- Recursively bisects the model about a plane (one of three axes) perpendicular to the longest dimension
- Method tends to generate cube shaped domains aligned along the coordinates axes

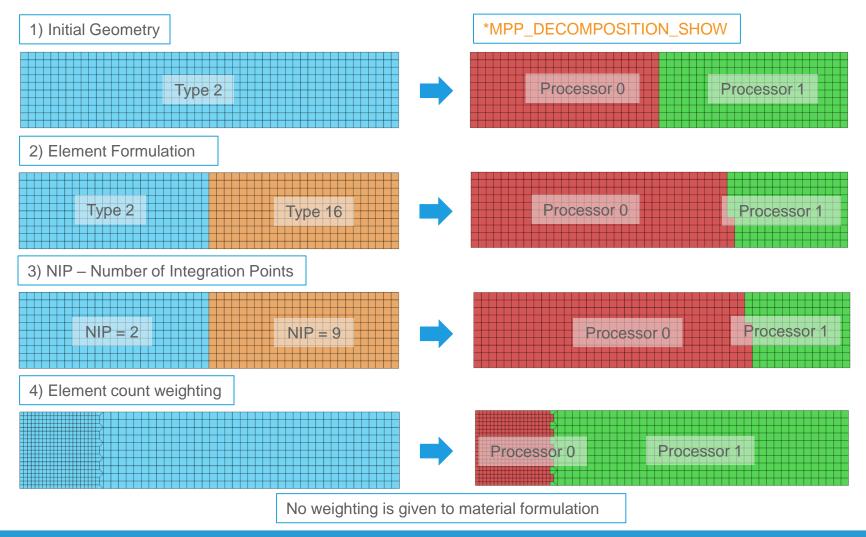






Decomposition methods in LS-DYNA

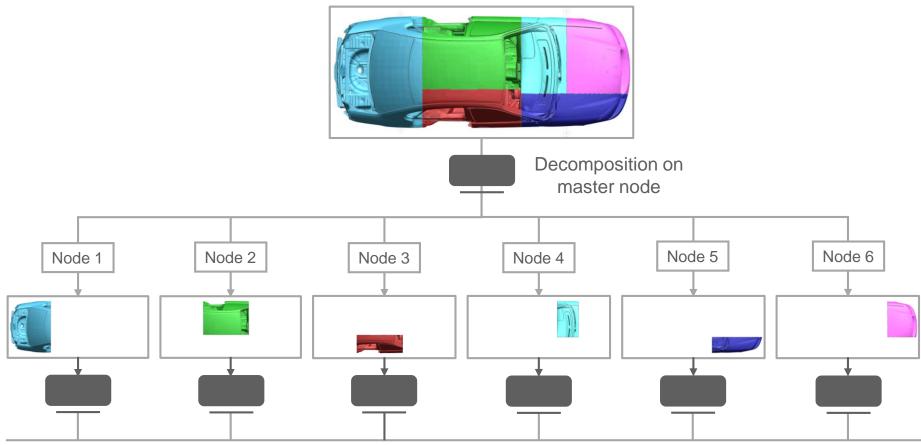
To improve the load balance LS-DYNA has some additional built-in intelligence





MPP Domain Decomposition

LS-DYNA Input Deck



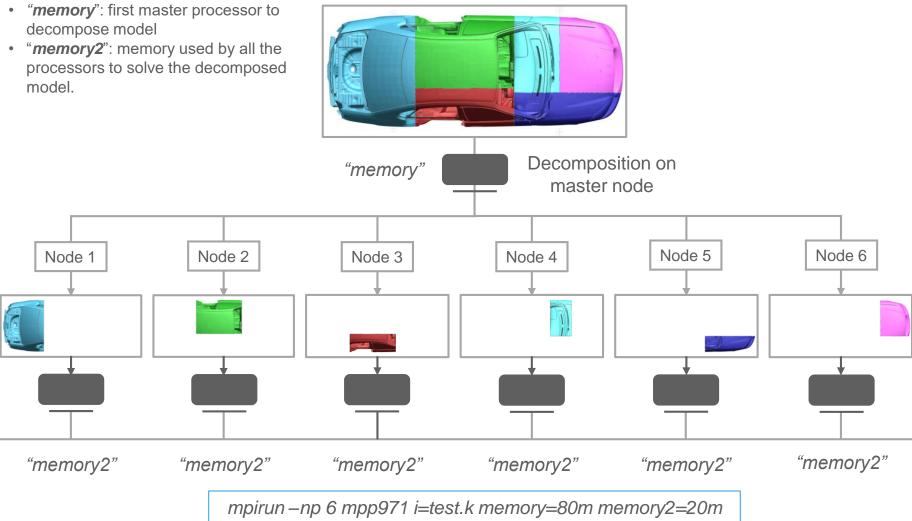


Memory Settings for MPP LS-DYNA

Webinar: MPP-DYNA

Two memory options for MPP LS-DYNA:

LS-DYNA Input Deck

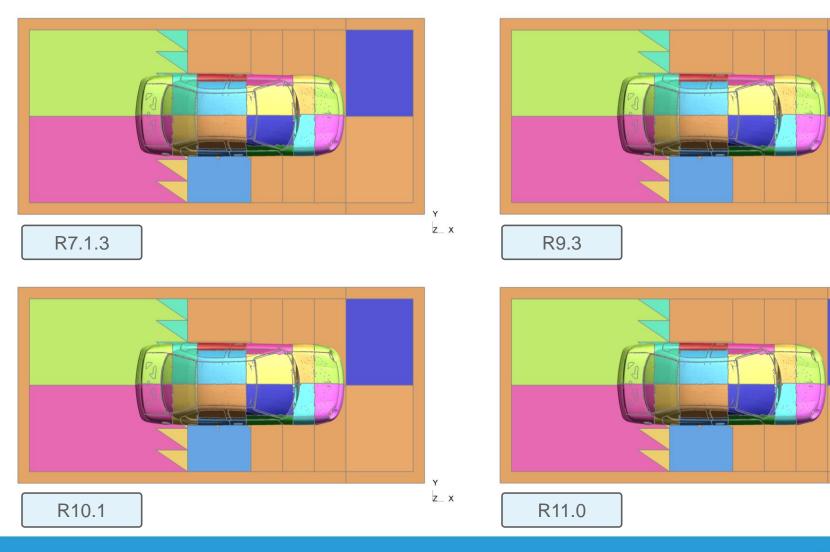




Decomposition in LS-DYNA

Webinar: MPP-DYNA

Note: default rcb decomposition method pretty consistent from one LS-DYNA solver to another





Y Z X

Y Z X

pfile_neon_refined_revised:
general { nodump nofull nod3dump nofail nobeamout }
decomp { sy 2 }
dir { local /local _ shm _ dir/neon _ refined _ revised }

- pfile contains MPP specific parameters that effect the execution of the program
- The file is split into sections, with several options in each section:
 - directory, decomposition, contact, general
- The file is case insensitive and free format input
- Can be used as a separate file or via the *CONTROL_MPP_PFILE keyword
- Full list of options can be found in Appendix O and *CONTROL_MPP card of LS-DYNA Keyword User's Manual Vol. I

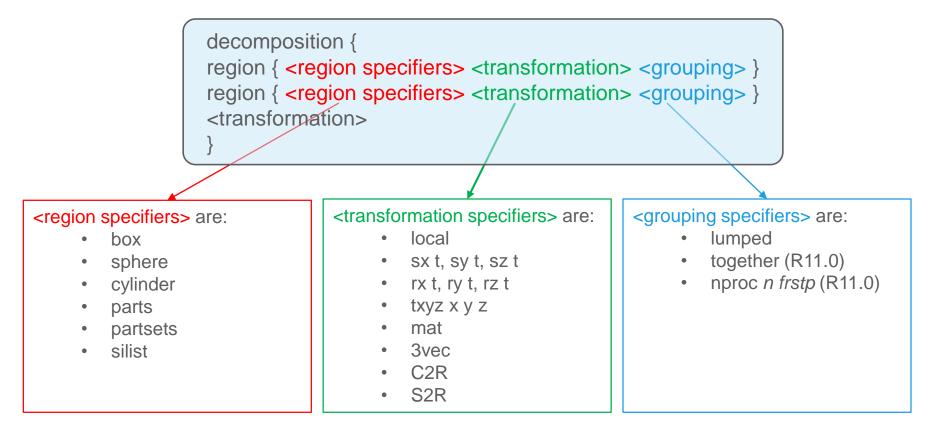
Additional comments:

- Check the output of PFILE directives in the <u>.otf/d3hsp file</u> useful when learning pfile syntax
- From R10.1 use of parameters (defined via *PARAMETER) in *CONTROL_MPP_PFILE



Special Decompositions

- If the default decomposition algorithm is not desired, it is possible for the user to provide a set of coordinate transformation functions which are applied to the model before it is decomposed (Appendix O: LS-DYNA MPP User Guide).
- **General form** for a special decomposition would look like this in pfile:



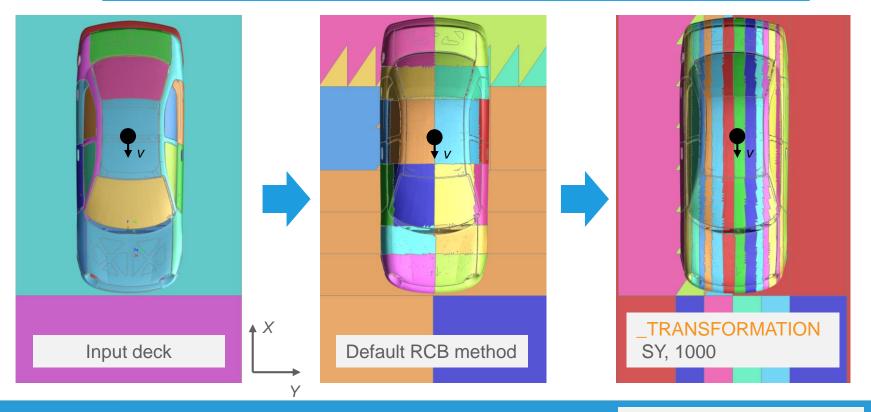


Special Decompositions

 If the default decomposition algorithm is not desired, it is possible for the user to provide a set of coordinate transformation functions which are applied to the model before it is decomposed (Appendix O: LS-DYNA MPP User Guide).

*CONTROL_MPP_DECOMPOSITION_TRANSFORMATION

Purpose: specifies transformations to apply to modify the decomposition



http://www.d3view.com/2006/09/simulation-modeldecomposition-using-recursive-bisection-method-rcb/

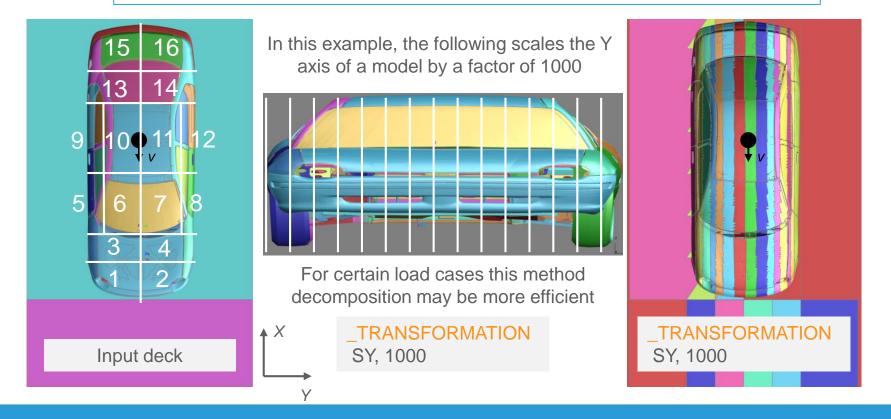


Special Decompositions

 If the default decomposition algorithm is not desired, it is possible for the user to provide a set of coordinate transformation functions which are applied to the model before it is decomposed (Appendix O: LS-DYNA MPP User Guide).

*CONTROL_MPP_DECOMPOSITION_TRANSFORMATION

Purpose: specifies transformations to apply to modify the decomposition





*CONTROL_MPP_PFILE decomp {sy 200.00}

*CONTROL_MPP_PFILE decomp { silist 1 sy 200.00}

SILIST = *CONTROL_MPP_DECOMPOSITION_CONTACT_DISTRIBUTE



sy 200.0: Scale the current y coordinates by 200.00

11 minute run time on 16 CPU

silist 1: All elements involved in a contact interface 1 are included in the region.

7 minute run time on 16 CPU



Problem:

MPP decomposition is based on averaging the computational across the processors. If a
model has been modified or refined, the cost profile will change and model will decompose
in a different way. This may change numerical results, particularly for sensitive models that
exhibit material/element failure. In such models, it would be difficult to distinguish between
'real' changes due to design updates, and changes due to the code.

RCBLOG keyword: *CONTROL_MPP_DECOMPOSITION_RCBLOG

pfile:

decomposition{ rcblog file_rcblog}

In the first job run, LS-DYNA will store all the cut information and also retain all other options in the pfile into 'file_rcblog'.

In subsequent runs, replace <u>p=pfile</u> to <u>p=file_rcblog</u> and LS-DYNA will decompose the model based on the preserved cut lines.



Decomposition of element domains can be visualised by:

- *CONTROL_MPP_DECOMPOSITION_SHOW
 - Outputs one plot state then terminates the analysis.
 - Each part correspond to the group of solids, shells, beams, thick shells, or SPH particles assigned to a particular processor
- *CONTROL_MPP_DECOMPOSITION_OUTDECOMP
 - Does not terminate the analysis early. Instructs LS-DYNA to output a settings file that contours elements according to processor ID.
 - ITYPE EQ.1: database in LS-PrePost format:

decomp_parts.lsprepost (binary)

• ITYPE EQ.2: database in animator format:

decomp_parts.ses (ASCII)

 When ITYPE EQ. 1, the elements assigned to any particular core can be viewed and animate by LS-PrePost by (1) reading the d3plot data, and then (2) selecting Models > Views > MPP > Load > decomp_parts.lsprepost

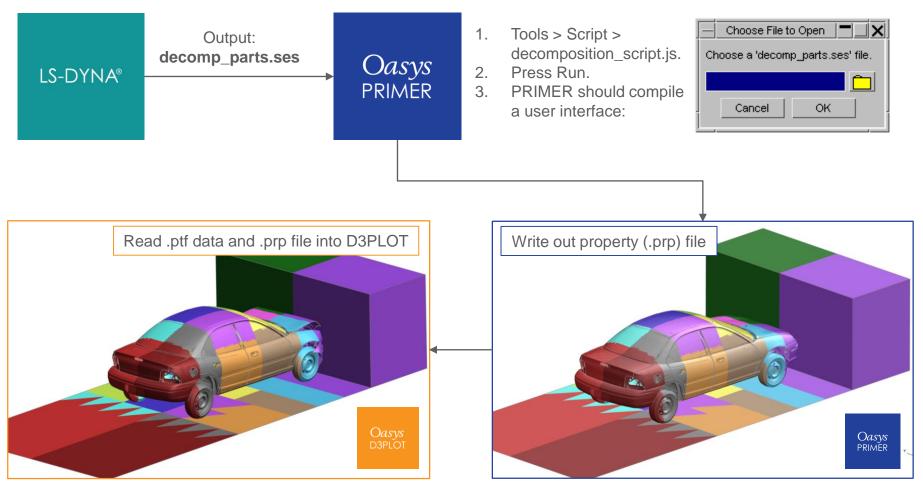
Command in partition file (pfile): OUTDECOMP ITYPE



Webinar: MPP-DYNA

*CONTROL_MPP DECOMPOSITION_OUTDECOMP (ITYPE EQ. 2)

Run LS-DYNA job





Information during execution

Whenever you have an LS-DYNA job running on a cluster node, you can access that node using SSH. One advantage of this is that it allows you to look at the processor core and memory usage reported by that specific system (compute node). This can useful for troubleshooting and for other purposes.

Here is an MPP LS-DYNA job running on our remote cluster vdgcls01:

vdgcls01 gmohamed 101% qstat -u gmohamed | grep neon.refined
223865.vdgcls01.global gmohamed dyna neon.refined.rev 11314 2 16 -- 2476:40:3 R 00:00:25

From that output, we can see that the job ID number is 223865:

```
vdgcls01 gmohamed 106% qstat -f 223865 | grep host
    exec_host = atrnode29/0-15
    submit_host = vdgcls01.global.arup.com
```

The LS-DYNA job is running on node atrnode29 and we've requested 16 (0-15) cores on that node. If we want to look at the processor core and memory utilisation on these nodes:

vdgcls01 gmohamed 108% ssh atrnode29 vdgcls01 gmohamed 108% top -u gmohamed





Information during execution

Use the 'top' command to check the available memory in the system

Tasks: 600 tota Cpu(s): 66.7%us	up 6 1, , 0	5 da 17 r .1%s	ys, 23:17, unning, 583 y, 0.0%ni,	1 user, load av sleeping, 0 st 33.2%id, 0.0%wa	rerage: 15.44, 9.90, 7.51 copped, 0 zombie 1, 0.0%hi, 0.0%si, 0.0%st				
Mem: 132030628k total, 93124324k used, 38906304k free, 162844k buffers									
PID USER	PR	NI	VIRT RES	SHR S %CPU %MEM	TIME+ COMMAND				
11603 gmohamed				17m R 100.0 1.4					
11604 gmohamed					3:24.08 ls-dyna_mpp_s_R				
11607 gmohamed					3:23.74 ls-dyna_mpp_s_R				
11610 gmohamed				12m R 100.0 1.4					
					3:23.71 ls-dvna_mpp_s_R				
11605 gmohamed				12m R 99.4 80.3					
11606 gmohamed				13m R 99.4 1.3					
11608 gmohamed 11609 gmohamed				13m R 99.4 1.3 13m R 99.4 1.3	3:24.08 ls-dyna_mpp_s_R 3:23.95 ls-dyna mpp s R				
11611 gmohamed				13m R 99.4 1.3	3:23.73 ls-dyna mpp s R				
11612 gmohamed				12m R 99.4 1.3					
11613 gmohamed				13m R 99.4 1.4	3:23.82 ls-dyna mpp s R				
11614 gmohamed				13m R 99.4 1.4	3:23.96 ls-dyna mpp s R				
11616 gmohamed				12m R 99.4 1.3	3:24.05 ls-dyna mpp s R				
11617 gmohamed				13m R 99.4 1.4	3:23.69 ls-dyna mpp s R				
11618 gmohamed				13m R 99.4 1.4	3:23.92 ls-dyna_mpp_s_R				



Information during execution

Use the 'top' command to check the available memory in the system.

You DO NOT want your job using swap space.

Often characterised by a job appear to be hanging for prolonged periods of time.

See d3hsp/otf extract below:

23149 t 2.5000E-02	08E-06 write d3p	lot file	10/26/18 11:2	23:16			
node number 7	7348 d	eleted at time 2	.50020E-02	shell element	7120	failed at time	2.5002E-02
shell element	7154	failed at time	2.5002E-02	shell element	7201	failed at time	2.5016E-02
shell element	7458	failed at time	2.5032E-02	shell element	1480	failed at time	2.5037E-02
shell element	1504	failed at time	2.5041E-02	shell element	7288	failed at time	2.5048E-02
shell element	9232	failed at time	2.5051E-02	shell element	7107	failed at time	2.5059E-02
shell element 2	25408	failed at time	2.5068E-02	shell element	9259	failed at time	2.5070E-02
shell element	7227	failed at time	2.5077E-02	shell element	20330	failed at time	2.5078E-02
shell element	9181	failed at time	2.5095E-02	shell element	7100	failed at time	2.5096E-02
shell element	7103	failed at time	2.5096E-02	shell element	7108	failed at time	2.5096E-02
shell element	7125	failed at time	2.5096E-02	shell element	7130	failed at time	2.5097E-02
shell element	7255	failed at time	2.5098E-02	shell element	7277	failed at time	2.5101E-02
shell element	7310	failed at time	2.5101E-02	shell element	7122	failed at time	2.5104E-02
shell element 2	25792	failed at time	2.5107E-02	shell element	1526	failed at time	2.5112E-02
24075 t 2.6000E-02 d	dt 1.0	8E-06 write d3pl	ot file	10/26/18 12:24	4:11		



Timing information CPU(seconds)	%CPU	Clock(seconds)	%Clock	
Keyword Processing 1.9453E+00	0.04	1.9692E+00	0.04	
MPP Decomposition 2.5644E+00	0.05	2.6758E+00	0.05	
Init Proc 9.8150E-01	0.02	9.8608E-01	0.02	
Decomposition 1.1864E+00	0.02	1.1884E+00	0.02	
Translation 3.9653E-01	0.01	5.0131E-01	0.01	
Initialization 9.8191E-01	0.02	1.0404E+00	0.02	
Init Proc Phase 1 6.9641E-01	0.01	7.1895E-01	0.01	
Init Proc Phase 2 6.5679E-02	0.00	8.5949E-02	0.00	
Element processing 4.4645E+02	8.41	4.4728E+02	8.40	
Shells 4.4602E+02	8.40	4.4672E+02	8.39	
Binary databases 5.7473E+01	1.08	5.7575E+01	1.08	
ASCII database 1.2730E+00	0.02	1.5664E+00	0.03	
Contact algorithm 7.8894E+02	14.86	7.9040E+02	14.85	Run time data taken from the d3hsp file as shown:
Interf. ID 1 4.8851E+02	9.20	4.8940E+02	9.20	
Interf. ID 2 1.3022E+02	2.45	1.3048E+02	2.45	
Interf. ID 3000 3.8772E+01	0.73	3.8860E+01	0.73	Element processing - material/element calculation
Interf. ID 3001 8.4251E+01	1.59	8.4395E+01	1.59	
Interf. ID 3002 4.4200E+01	0.83	4.4283E+01	0.83	Contact algorithm - all contacts in the model
Particle Algorithm 3.8919E+03	73.28	3.8994E+03	73.27	contact argorithm arr contacts in the moder
P-structure 2.5996E+03←	48.95	2.6048E+03	48.95	
Particle collision . 1.2921E+03	24.33	1.2945E+03	24.32	- P-structure - particle-to-fabric collision
Rigid Bodies 1.3857E+01	0.26	1.3838E+01	0.26	calculation
Time step size 1.6579E+01	0.31	1.6568E+01	0.31	
Rigid wall 1.0820E-01	0.00	9.6764E-02		
Group force file 1.1662E-01	0.00	1.3421E-01	0.00	- Particle collision - particle-to-particle
Others 4.8768E+01	0.92	4.8825E+01	0.92	collision calculation
Misc. 1 1.2468E+01	0.23	1.2602E+01	0.24	
Misc. 2 4.0869E-01	0.01	6.0941E-01	0.01	
Misc. 3 1.7122E+01	0.32	1.7227E+01	0.32	T o t a l s - total analysis run time (clock time)
Misc. 4 9.9219E+00	0.19	9.9172E+00	0.19	
Totals 5.3108E+03	100.00	5.3218E+03	100.00	
Problem time = 1.0000E-01				
Problem cycle = 111112	_			Elapsed Time
Total CPU time = 5311 secon	ds (1 hours 28 minu	tes 31 🧲	
seconds)				
	nanose	econds		
Clock time per zone cycle= 397	nanose	econds		



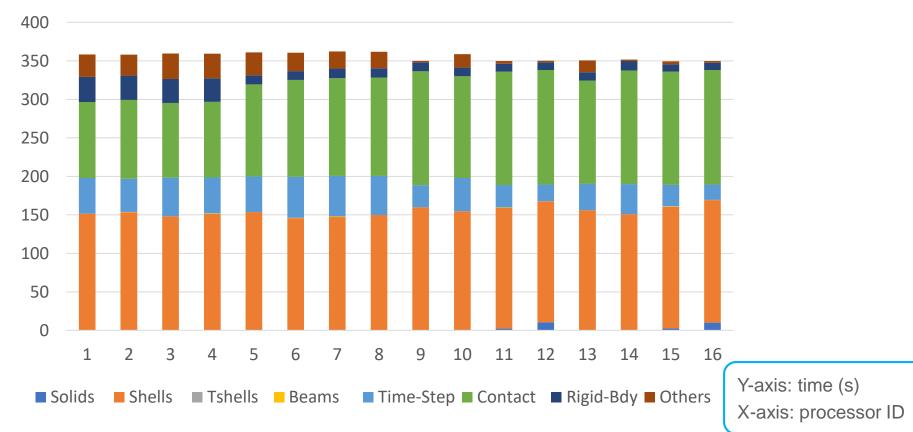
Load-balancing information

Webinar: MPP-DYNA

Oasvs Ltd

Information after execution

• Processor load balance can be found in the 'load_profile.csv' file



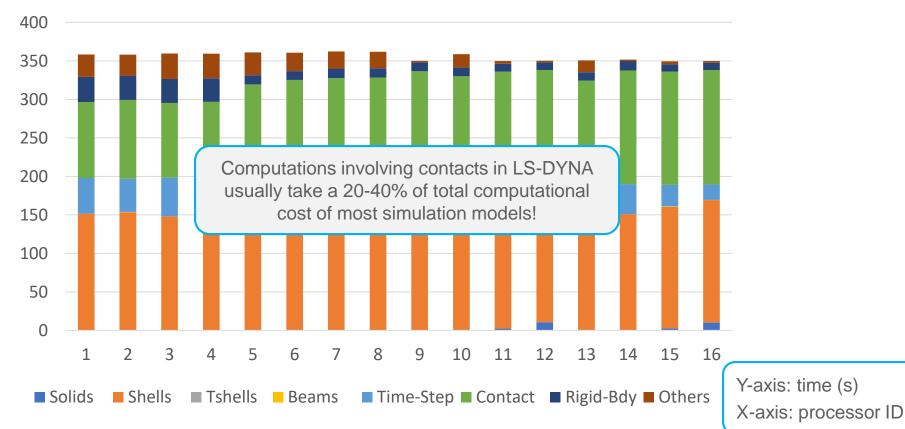
NCPU = 16

Load-balancing information

Webinar: MPP-DYNA

Information after execution

• Processor load balance can be found in the 'load_profile.csv' file



Slide 23

NCPU = 16

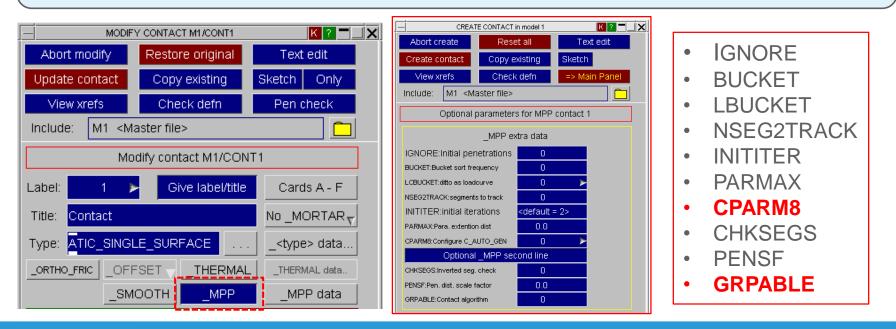


MPP Contact Algorithms

In simple terms, contacts comes down to the problem of comparing a single node *N* with a single segment *S* and applying forces as necessary to ensure *N* does not pass through *S*.

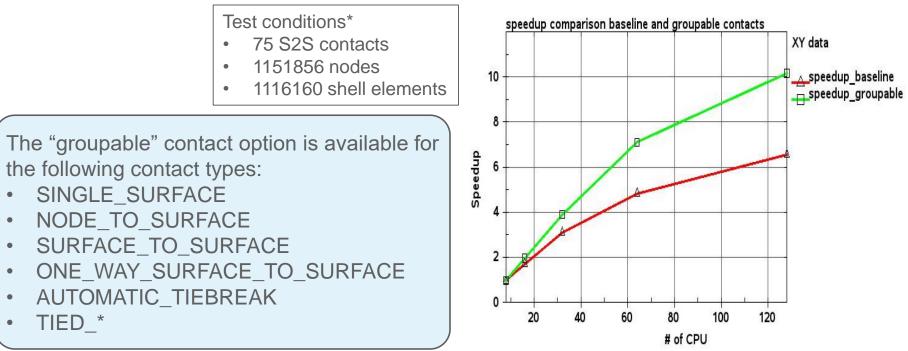
Problem:

- Complex models may contain several contact interfaces.
- The decomposition of such models can result in an uneven distribution of these contacts among the available processors.
- Some processors may have many contacts to handle, and others may have none. This variability can cause inefficiencies which adversely impact scalability.





- It is generally recommended to have as few contact definitions as possible convert multiple contact definitions to a single surface with force transducers (*CONTACT_FORCE_TRANSDUCERS) and 'FTALL = 1' in *CONTROL_CONTACT.
- Where this approach is not desired by the user, the <u>_GROUPABLE</u> contact option is available.
- With this option turned on, contact definitions are internally combined in LS-DYNA to reduce communication.



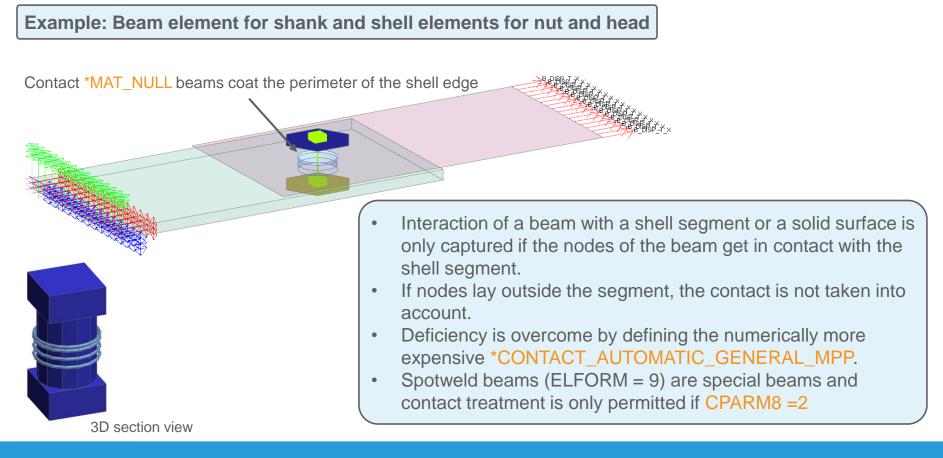


MPP Contacts: CPARM8

Webinar: MPP-DYNA

CPARM8:

- 1. Exclude beam to beam contact from the same part ID. This is for *CONTACT_AUTOMATIC_GENERAL.
- 2. Consider Spotweld beams in contact.





Parameters are set in a hierarchy:

- 1. *CONTROL_CONTACT sets overall defaults,
- 2. *CONTACT_... sets parameters for that contact overrides 1 above,
- 3. *PART_CONTACT sets parameters for that part in the contact overrides 1 & 2 above.

Important parameters are:

- Contact thickness and scaling factors
- Penalty stiffness scale factors or SOFT options
- Friction values
- Treatment of initial penetrations with **IGNORE** option
- Note: LS-DYNA will use the MIN(*SECTION_SHELL thickness, 0.4*element_length) for all parts unless *CONTROL_CONTACT SSTHK (=1) or *PART_CONTACT OPTT is activated.

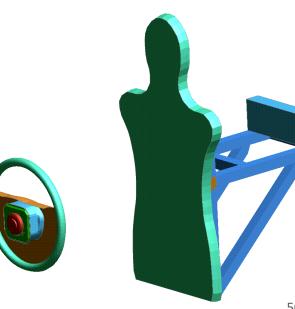




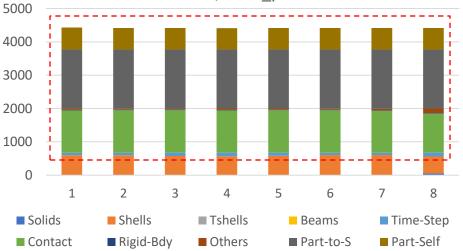


Notes on Airbags

D3PLOT: DAB CPM06



- In airbag analysis, the particle-to-structure collision calculation is the largest consumer of CPU.
 - Most affected by interconnect bottleneck
- Tasks that require a lot of processor communication have the worst scalability:
 - Particle-to-structure
 - Particle-to-particle contact



NCPU = 16, 'load profile.csv'

 Deployment and correlation of airbags can be sensitive to decomposition:
 *CONTROL_MPP_DECOMPOSITION_BAGREF
 *CONTROL_MPP_DECOMPOSITION_PARTS
 *CONTROL_MPP_DECOMPOSITION_PARTS_ DISTRITBUTE

> Oasys Ltd The Software House of ARUP

Notes on Airbag: Recommendations

- Users should benchmark their cluster performance and use the fastest processors per node (PPN) where possible.
- Run time of multiple airbags can be reduced by separating the CPU allocated to each CPM airbag using NPROC and FRSTP:
 - *CONTROL_MPP_DECOMPOSITION_ARRANGE_PARTS_{OPTION}
 - Users should try different ratios of NPROC based on the airbag requirements.
- To improve consistency, users should ensure airbag models are developed using a certain number of processors (NPROC), which is fixed throughout the life of the airbag model.

Decomposition procedure for a model with two CPM bags

- 1. Run model with airbag 1 active and airbag 2 removed.
 - Make a note of total time spent on CPM airbag 1 activity.
- 2. Run model with airbag 2 active and airbag 1 removed.
 - Make a note of total time spent on CPM airbag 2 activity.
- 3. Use NPROC to distribute the total number of CPUs between the two airbags in the same ratio as the time spent on each bag individually.

Slide 29



Webinar: MPP-DYNA

Airbag 1

Airbag 2





*CONTROL_MPP_DECOMPOSITION_ARRANGE_PARTS_OPTION

Purpose: Allow users to distribute certain part(s) to all processors or to isolate certain part(s) in a single processor. This keyword supports multiple entries. Each entry is to be processed as a separate region for decomposition.

When this keyword is part of an included file and the LOCAL option is given, the decomposition will be done in the coordinate system of the included file, which be different from the global system, if the file is included using the *INCLUDE_TRANSFORM keyword

Card 1	1	2	3	4	VARIABLE	DESCRIPTION
Variable	ID	TYPE	NPROC	FRSTP	ID	Part ID/Part set ID
Туре	I	I	I	I	TYPE	EQ.0: Part ID to be distributed to all processorsEQ.1: Part Set ID to be distributed to all processors
Default	none	none	None	None		EQ.10: Part ID to be lumped into one processor
					NPROC	EQ.11: Part Set ID to be lumped into one processor. Used only for TYPE equal to 0 or 1. Evenly distributed Part ID/Part set ID to NPROC of processors.
					FRSTP	Used only for TYPE equal to 0 or 1. Starting MPP rank ID.



MPP General Guidelines

- For consistency, use *CONTROL_MPP_IO_LSTC_REDUCE and _RCBLOG
- Merge small contact definitions into big ones (and use *CONTACT_FORCE_TRANSDUCERS for output).
- To improve load balancing, contact definitions should be equally distributed to all processors, use _GROUPABLE to handle smaller contact definitions.
- Distribute large contact area evenly among processor via pfile:
 - decomp { silist 1,2,3}
 - or in input deck

*CONTROL_MPP_DECOMPOSITION_CONTACT_DISTRIBUTE

- To isolate any given contact to a single processor, use
 - *CONTROL_MPP_DECOMPOSITION_CONTACT_ISOLATE
- SMP and MPP contact algorithms are implemented differently:
 - The use of *CONTACT_AUTOMATIC_SURFACE_TO_SURFACE with SOFT=2 will be the contact that will give most similar results between MPP-DYNA and SMP.



- An MPP restart requires binary restart files called d3dump## and d3full##, where ## is a number
- Binary dump files contain a complete record of the model (stress, strain, deformation etc.) at a particular point in time
- There are three classes of restarts:
 - Simple restart: No changes to input deck. mpirun mpp971 -np 16 r=d3dump09
 - Small restart: Few small changes permitted, e.g. change termination time mpirun mpp971 -np 16 i=small.key r=d3dump09
 - Small restart file required
 - Full restart: Make significant changes to model. mpirun mpp971 -np 16 i=full.key n=d3full09
 - Full restart input deck required
 - *STRESS_INITIALISATION keyword required



Restart dump files

d3dump files:

- Permanent restart files which **accumulate** throughout the analysis.
- By default, binary d3dump file written at normal termination or crash of a run.
- Output frequency controlled with *DATABASE_BINARY_D3DUMP
 - A new restart file is created after each interval, CYL, thus a family of dump files is created and numbered sequentially, e.g. d3dump01, d3dump02 etc.

These files can be quite large and care should be taken with the d3dump files not to create too many.

Can be suppressed in MPP LS-DYNA:

- *CONTROL_MPP_IO_NOD3DUMP,
- *CONTROL_MPP_IO_NODUMP,
- *CONTROL_MPP_IO_NOFULL.





Sense switches allow you to control the behaviour of an analysis while it is running.

To activate one of the sense switch options a text file called *jobname*.kil needs to be created, containing one line of text, which is the sense switch you want to use:

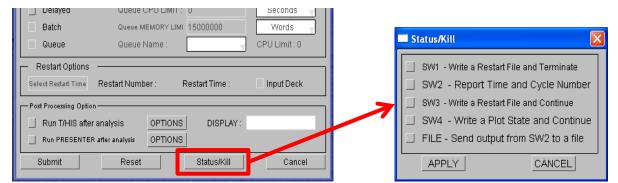
Slide 34

e.g. sw1. (include the dot)

Some of the options are:

- sw1. A restart file is written and LS-DYNA terminates
- sw2. LS-DYNA responds with time & cycle info
- sw3. A restart file is written and LS-DYNA continues
- sw4. A plot state is written and LS-DYNA continues
- swa. Flush ASCII file buffers

The *jobname*.kil file can also be created using the Oasys SHELL:





When restarting an LS-DYNA job:

- Use the same LS-DYNA executable as in the run that produced the dump file.
- Use same numbers of CPU's as in the run that produced the dump file.
- Use the same memory as in the run that produced the binary dump files.
- Run the analysis in the same directory.





- A Short course of LS-DYNA/MPP[®], Jason Wang, 2010
- MPP Contact: Options and Recommendations, Brian Wainscott, 2015
- Appendix O: LS-DYNA MPP User Guide
- D3view blog (<u>www.d3view.com</u>), Suri Bala
- http://ftp.lstc.com/anonymous/outgoing/jday/restart.pdf



ARUP

www.arup.com/dyna

For more information please contact the following:

UK:
Arup
The Arup Campus
Blythe Valley Park
Solihull, West Midlands
B90 8AE
UK
T +44 (0)121 213 3399
F +44 (0)121 213 3302
dyna.support@arup.com

China:

Arup 39/F-41/F Huai Hai Plaza Huai Hai Road (M) Shanghai China 200031

T +86 21 6126 2875 F +86 21 6126 2882 china.support@arup.com

India:

nHance Engineering Solutions Pvt. Ltd (Arup)

Ananth Info Park HiTec City Madhapur Hyderabad - 500081 India T +91 (0) 40 44369797 / 8 india.support@arup.com



