



Forum TERATEC 2011

RADIOSS Parallel Performance Enhancements in Hybrid MPP

Eric LEQUINIOU Altair Director High Performance Computing

Forum TERATEC 2011

28 & 29 juin 2011





Altair Overview

Forum TERATEC 2011

\$152M Revenue

1,300 Employees





About Altair



\$100M

2008

1985 Altair Founded ...

In 1985 as a product design consulting company

Today ...

A global software and technology company focused on enterprise analytics, product development and advanced computing



Global Offices



Seattle, USA Los Angeles, US/ ∾Austin; USA Dallas, USA

> South Pacific Ocean

Detroit, USA Boston, USA Milwaukee, USA Atlanta, USA

Toronto, Canada

North Atlantic Ocean

Mexico City, Mexico Sao Paulo, Brazil Lund, Sweden Gothenburg, Sweden Leamington Spa, UK Manchester, UK Boeblingen, Germany Cologne, Germany Hanover, Germany Hanover, Germany Munich, Germany Deris, France

Sophia Antipolis, France

Torino, Italy

Milan, Italy

Moscow, Russia

Pune, India Chennal, India

Hyderabad, Ind

Bangalore, India Indian Ocean Beijing, China Shanghai, China

Tokyo, Japan Osaka, Japan

Nagoya, Japan

Seoul, Korea

Melbourne, Australia

Forum TERATEC 2011

28 & 29 juin 2011



Altair Brands & Companies







HyperWorks for Analysis & Optimization 📥 Altair



Data and Process Management

Forum TERATEC 2011

28 & 29 juin 2011



RADIOSS the HyperWorks Solver









RADIOSS Hybrid MPP Parallel version



Source : Intel Corp.



SMP Model







Constraints

- Needs shared memory
- Depends on NUMA architecture
- Limited scalability

Advantages

- Ease of use
- Good Performance for small size model



SPMD Model



Network Interconnect Advantages Node1 Node2 No need of shared memory Good locality for NUMA architecture • M_2 **M**₁ M₁ M_2 Improved scalability • Socket1 Socket2 Socket2 Socket1 Constraints High Speed interconnect preferred C_1 C_2 C_3 C_4 C_1 C_2 C_3 C_4 Model size big enough Т Т Т Т Т Т т Т Т н н н Н Н Н Н Н Н н н Н R R R R R R R R R R R R Ε Ε Ε Ε Ε Ε Ε Ε Ε Ε Ε Ε Α Α Α Α Α Α Α Α Α Α Α D D D D D D D D D D D D 1 1



Hybrid MPP Model



Network Interconnect

Node1		Node2	HMPP Advantages NodeN
M ₁	M ₂	M ₁	 Adaptable to hardware architecture - #sockets per node M₁ M₂
Socket1	Socket2	Socket1	Socket2 - NUMA memory localityocket1 Socket2 • Highest level of scalability & performance
C_1 C_2	C ₃ C ₄	C_1 C_2	C ₃ Works well with high & low speed interconnects • Scalable even for small model size
T H H R E A D D 1 2	TTHREADD12	T T H H R E A A D D 1 2	TTTTTTTTTTTHH

Forum TERATEC 2011



Hybrid MPP Parallel Version



Value proposition

"Combine the benefit of both RADIOSS parallel versions inside a unique code with enhanced performance and flexibility"

Customer Benefits

Reduce simulation cost thru maximized scalability & performance

Increase flexibility to customer's needs & hardware resources evolution

High Quality with full repeatability of SMP & SPMD numerical results

Additional cost reduction thru easiest maintenance

Implementation Challenges

Mix MPI programming and OpenMP multithreading

Keep ease of use of SMP with care of MPI installation

Keep ease of debugging of SPMD and its deterministic behavior



Elapsed (s)

Hybrid MPP Performance Results



intel

Software

NEON_1M 8 ms Intel Nehalem 2.80 GHz Scalability Study vs # SMP Threads

3500 3311 1 thread 2 threads 3000 2610 2529²⁵²² 4 threads 2500 8 threads 2000 1772 1314 1500 1301 1289 906 1000 697 675 668 473 361 358 356 500 256 246 287 ²²⁷ 203 205 ¹³⁵126 ¹⁵² ¹²⁷ 99 95 0 512 #core 32 8 16 64 128 256 14 Forum TERATEC 2011 28 & 29 juin 2011



Hybrid MPP Recommendations



To handle high number of nodes & cores, use Hybrid MPP

Hybrid becomes interesting when pure MPI scalability drops

Hybrid allows to scale to thousands of cores

Recommended mix

One SPMD domain per CPU socket

As many SMP threads per SPMD domain than cores per socket

Total number of processes \leq number of (logical) cores

Importance of memory & process placement in MPI (Intel MPI example)

KMP_AFFINITY=compact

I_MPI_PIN_DOMAIN=omp





Hybrid MPP Case Studies



Case Study 1 The 5 minutes breakthrough



Challenge Be the first FE code able to simulate a million element frontal crash in less than 5 minutes (explicit computation) Combine several features & innovations Hybrid MPP Advanced Mass Scaling Extended Single Precision Intel MPI & OpenMP cluster tools Intel Nehalem processors

http://www.prnewswire.com/news-releases/altair-achieves-breakthrough-time-savings-in-crash-simulations-68838927.html



Case Study 1 The 5 minutes breakthrough



Neon 1Million Element - 80 ms

Platform			Intel Supermicro X8DTN	
			128 node system	m
	(Intel)		Dual socket x 4	core
	Software		Intel Xeon 5560	@ 2.8 GHz
			18GB 1066 DDR3 memory	
Hybrid		128 SPMD x 8 SMP		
Single Precision			1024 cores in total	
AMS		Objective	10µs	
			Mean Time St	tep 5.5µs
		14514 cycles		
Performance		5 minutes (294s)		
Efficiency		3.5x w/o AMS		
			10x w/o Hyt	orid & AMS





Case Study 2 15 Million Element Benchmark



Challenge

Test the ability of RADIOSS to handle input deck of more than 15 million elements for explicit simulation of car crash against offset deformable barrier

Combine different constraints

Efficient pre/post processing tools

Access to large compute resources with state of the art processors

Ability to scale up to several thousands cores

Turn around time must stay reasonable (ie less than one night)



Case Study 2 15 Million Element Benchmark



15 Million Element - 100 ms						
Platf	orm	Intel CRT Data Center in Dupont WA				
		Intel SR1600UR 144 nodes system				
(s	(intol®)	Dual socket x 6 core				
		Intel Xeon X5570 @ 2.93 GHz				
	Software	24GB 1333 DDR3 memory				
		1-GbE, QDR IB Mellanox				
		RHEL5 update 3				
Hybrid		144 SPMD x 12 SMP				
Single Precision		1728 cores in total				
DT/N	ODA/CST	Mean Time Step 0.2µs				
		502107 cycles				
Performance		11,4 hours (41181s)				









Case Study 3 Multi GPUs Hybrid MPP



Challenge

Demonstrate possibility to use Hybrid MPP model with Multi GPUs

Combine different challenges

Porting RADIOSS implicit iterative solver on GPU

- 3 parallelization models working together
 - MPI on top
 - OpenMP (CPU)
 - OpenCL (GPU)

Speedup calculation by using multi GPUs

 RADIOSS Presentation at AMD Fusion Developer Summit at Seattle



Hybrid MPP Computing with multi GPUs Altair





400K Elt Bench with M=2M NZ=64M



- Dual GPUs using 4 SMP threads per MPI delivers best performance, with a speedup of near 16 times versus single core CPU
- Encouraging results on multi GPUs to pursue
- Demonstration of Hybrid MPI/OpenMP/GPU

Forum TERATEC 2011

Ter@tec



Conclusion



Hybrid MPP is a SQR solution Speed with increased scalability & performance Quality of RADIOSS Solver Reliability with full repeatability Hybrid MPP is the right solution Ready for future hardware evolution toward multi/many cores Run today simulations faster than ever Run more accurate simulations one couldn't afford before Hybrid MPP is available now within HW11







RADIOSS The Power To Solve