Scalability of LS-DYNA on SGI Systems

Roger Chu and **Guangye Li**
SGI
1600 Amphitheatre Parkway
Mountain View, CA 94043
Chu@corp.sgi.com
Gli@houst.sgi.com

ABSTRACT

In parallel computation, the scalability of the application software is critical. It is especially important when large number of processors are used. In this paper, we present the scalability results of LS-DYNA on the SGI multiprocessor computer systems. Furthermore, since MPP-DYNA is a domain decomposition based software, data partitioning algorithms play an important role in the scalability of the code. We will show in this paper that for some car crash models, special data partitioning techniques may improve the scalability significantly.

INTRODUCTION

With the dramatic increase of the price-performance of modern high performance computers, the model sizes of car crash and metal stamping simulations have been constantly increasing in recent years. Nowadays, a 200,000-element model is a common case. For large models, highly scalable computers and software play crucial rolls in reducing turnaround time for simulations. The traditional symmetric multiprocessor (SMP) systems and shared memory programming models have difficulties to scale on large number of processors. The combination of the domain decomposition based MPP LS-DYNA and the highly scalable SGI ccNUMA system is an ideal solution for reducing simulation turnaround time.

In this paper, we will first briefly introduce the SGI ccNUMA architecture and the MPP LS-DYNA software. We will also present some performance results of both SMP and MPP LS-DYNA on the SGI systems. Finally, we show the performance improvement of MPP LS-DYNA by using variations of domain decomposition techniques.

THE HIGHLY SCALABLE ccNUMA ARCHITECTURE AND MPP LS-DYNA

SGI has offered multiple generations of symmetric multiprocessor (SMP) systems based on the MIPS microprocessors such as Challenge and Power Challenge. The cache-coherent globally addressable memory architecture of these SMP systems has provided a convenient programming environment for efficient execution of both parallel and throughput based workloads. To meet the demand of high scalability, ease of programming and low cost, SGI introduced the cache-coherent non-uniform memory access (ccNUMA) multiprocessor systems – SGI Origin 2000. The Origin system employs distributed shared memory (DSM), with cache coherence maintained via a directory-based protocol. It was designed from ground up as a multiprocessor capable of scaling to both small and large processor counts without bandwidth, latency, or cost cliffs. The Origin system consists of up to 512 nodes interconnected by a scalable Craylink network. Each node consists of one or two MIPS processors with 8 MB secondary cache and up to 4 GB of memory. For details about the SGI Origin 2000, please refer to J. Laudon and D. Lenoski. The next generation of the SGI Origin system coming to the market later this year will improve the memory bandwidth, latency and the scalability to an even higher level.

Unlike the SMP version of LS-DYNA which exploits parallelism in loop level, the MPP LS-DYNA uses the domain decomposition technique, i.e., the mesh of a model is partitioned into sub-domains and, in most cases, each domain is assigned to one processor. Processors exchange boundary data by using explicit communication tools, specifically, MPI, the Message Passing Interface. On distributed memory computers, because most of data is truly local, and each processor has plenty of computation work to do before communicating to other processors, the MPP version may scale to much large number of processors. In theory, the parallel speedup should be less than linear. However, on cache-based computer systems such as the Origin 2000, a super-linear speedup effect may happen on certain number of processor (See Lin, 1998). This is due to the fact that with the increase of processors, the sizes of sub-domains and, therefore, the local arrays, are getting smaller. At some point, the local arrays may all fit into the cache, which results in a performance jump.

In domain decomposition applications, to balance the workload and to minimize communications between processors are two crucial factors for scalability. These two goals may be reached by choosing right data partitioning techniques. For LS-DYNA, the simple Recursive Coordinate Bisection (RCB) algorithm may provide satisfactory data partitioning results. According to our experience, the built-in RCB decomposer with default options provides good scalability on systems with 16 processors or less for most of models. However, on large number of processors, the scalability may deteriorate dramatically due to the smaller computation and communication ratio and unbalanced workload caused mainly by contacts. Fortunately, this situation may be significantly improved by choosing right data partitioning techniques.

PERFORMANCE RESULTS

In this section, we present performance comparisons of the SMP LS-DYNA and MPP LS-DYNA as well as comparisons of different domain decomposition techniques on the SGI Origin 2000 systems.

1. Models. We used two frontal impact models in our experiments: The small one has about 35,000 elements. The large one is the Neon model provided by NCAC. It has 322 components, consisting 285,720 nodes, 267,895 shell elements, 67 beam elements and 2,920 solid elements. Lin [1] used the Neon model to show HP system's performance for 5 milliseconds simulation. We used 30 milliseconds simulation in our experiments considering the fact that in the first 5 milliseconds of simulation, there are essentially no contacts and that contacts are crucial in testing scalability of parallel computers.
2. Systems. We used two types of Origin 2000 systems: Origin 2000 with MIPS R12000 processors at the clock speed of 300 MHz and Origin 2000 with 400 MHz MIPS R12000. Both systems have 8 MB secondary cache per processor. These two systems use the same interconnection architecture.
3. LS-DYNA versions. We used the ls950c for the SMP version and used mpp940.2 (October 1999) for the MPP version.

In the following tables, NCPU stands for the number of processors used in the simulation, and Seconds stands for the elapsed time for the simulation in seconds. The term of speed-up is defined as the ratio of the elapsed time on one processor and that on NCPU processors. Table 1 compares the SMP LS-DYNA performance on the two Origin 2000 systems using the small model. Table 2 shows the MPP version performance using the same model as in Table 1. From these two tables, one may see that on a single processor, the MPP version is a bit slower than the SMP version. However, the MPP version is significantly faster than the SMP version on 4 and more processors. Table 3 compares the performance of the MPP version for the neon model using two different domain decomposition techniques. The machine used in this experiment was the Origin 2000 with 300 MHz R12000. The domain decomposer used in the first column was RCB with default parameters. In the second column, the decomposer was RCB with scale 0.0 on the X direction i.e. the partition was on the Y and Z plane only. One may see from Table 3 that the MPP version achieves much better performance on large number of processors by using the modified RCB decomposition. This is due to the fact that the modified RCB provides better data partitioning with more even distribution of workloads and less communications between processors.

Table 1. Performance of the SMP version, frontal crash, 35K elements

NCPU	Origin 2000 (300 MHz)		Origin 2000 (400 MHz)	
	Seconds	Speed-up	Seconds	Speed-up
1	7138	1.00	5544	1.00
2	4207	1.70	3321	1.67
4	2667	2.68	2011	2.76
8	1800	3.97	1446	3.83

Table 2. Performance of the MPP version, frontal crash, 35K elements

NCPU	Origin 2000 (300 MHz)		Origin 2000 (400 MHz)	
	Seconds	Speed-up	Seconds	Speed-up
1	8001	1.00	6208	1.00
2	4175	1.92	3299	1.88
4	1909	4.19	1486	4.18
8	1005	7.96	821	7.56

Table 3. Performance of the MPP version on Origin 2000 (300 MHz),
Neon model, 267K elements, 30 millisecond simulation

NCPU	Default RCB		Modified RCB	
	Seconds	Speed-up	Seconds	Speed-up
1	107332	1.00	107332	1.00
2	58137	1.85	53147	2.02
4	27311	3.93	24016	4.47
8	13959	7.69	11945	8.99
16	7342	14.64	6689	16.05
32	4812	22.31	3405	31.52
64	3060	35.08	2453	43.76

SUMMARY

We studied the performance of both the SMP version and the MPP version of LS-DYNA on two car crash simulation models. Our results show that the MPP version may scale to large number of processors on the SGI ccNUMA computer systems. Furthermore, our study showed that the scalability of the MPP version may be further boosted by carefully choosing the right domain decomposition techniques.

REFERENCES

1. Lin, Yih-Yih, 1998. *Crash Simulation on Parallel Multiprocessors*. 5th International LS-DYNA Users Conference 1998, September 21 and 22, 1998, Southfield, Michigan.
2. Lovett, Tom, and Russell, Clapp, 1996. *The SGI Origin: A ccNUMA Highly Scalable Server*. Silicon Graphics, Inc. 2011 North Shoreline Boulevard, Mountain View, California.