

A Dynamic Data Driven Grid System for Intra-operative Image Guided Neurosurgery*

Amit Majumdar¹, Adam Birnbaum¹, Dong Ju Choi¹, Abhishek Trivedi²,
Simon K. Warfield³, Kim Baldrige^{4,1}, and Petr Krysl²

¹ San Diego Supercomputer Center, La Jolla, CA 92093, USA

² Structural Engineering Department, University of California San Diego,
La Jolla, CA 92093, USA

³ Computational Radiology Laboratory, Brigham and Women's Hospital,
Harvard Medical School, Boston, MA 11111 USA

⁴ Department of Chemistry, University of Zurich

Abstract. In the future, advanced biomechanical simulations of brain deformation during surgery will require access to multi-teraflop parallel hardware, supporting operating room infrastructure. This will allow surgeons to view images of intra-operative brain deformation within the strict time constraints of the surgical procedure - typically on the order of minutes, multiple times during a six or eight hour long surgery. In this paper we explore the grid infrastructure issues involved in scheduling, on-demand computing, data transfer and parallel finite element biomechanical simulation, which would guarantee that such a dynamic data driven real time application is actually feasible.

1 Introduction

Over the last decade, there has been tremendous progress toward fulfilling a very compelling idea: surgeons should be able to take advantage of imaging and computational technology to provide them with an enhanced ability to visualize complex hidden structures, even while operations are taking place [1-3]. Today, computer assisted image guided therapy (IGT) takes place only in research hospitals, supported by collaborative teams of clinicians, computer scientists and engineers. In this work, we explore some of the issues that must be tackled in order to fulfill the full promise of these prototype systems in the area of image-guided neurosurgery (IGNS).

The goal in neurosurgery is to provide 3D images of the brain that clearly delineate anatomical structures and tumor tissue. Key surgical challenges for neurosurgeons during tumor resection are to (1) remove as much tumor tissue as possible, (2) minimize the removal of healthy tissue, (3) avoid the disruption of critical anatomical structures, and (4) know when to stop the resection process. These challenges are compounded by the intra-operative shape deformation of the brain that happens as a result of tissue resection and retraction, injection of anesthetic agents, and loss of cerebrospinal fluid. The result is that the accuracy of the pre-operative plan diminishes steadily during the procedure.

* This research was supported in part by the NSF ITR grants CNS 0427183, 0426558, and NIH grants P41 RR13218, P01 CA67165, R01 LM0078651.

It is therefore of great importance to be able to quantify and correct for these deformations while a surgery is in progress, by dynamically updating pre-operative images in a way that allows surgeons to react to changing conditions.

In our preliminary work [3] we have developed a prototype system, residing locally at the operating room at Brigham and Women's Hospital (BWH), that integrates data acquisition, quantitative monitoring, enhanced real-time control and intra-operative surgical navigation. Images of the patient's brain are captured pre-operatively and are registered against a brain anatomy atlas. During surgery, using an intra-operative scanner, new images of the patient's brain are acquired [4] and a finite element biomechanical model is solved [5] to construct a new visualization merging pre- and intra-operative data [6]. The time constraint on this overall procedure is severe and allows only about two minutes for the finite element biomechanical simulation.

The current prototype system at BWH uses a homogeneous, linear elastic biomechanical material model of the brain, primarily since only this crude and less accurate model is able to meet the real time constraint of IGT with the available dedicated on-site compute power at BWH. However, before the system is ready for broader adoption, the accuracy and precision for broader adoption, the accuracy and precision of the simulations will be significantly improved. Since the complex biomechanical model, under development, must be solved within the real time constraint of approximately two minutes, fast access to much larger-scale computational resources such as those found at the major national supercomputer centers, is required.

That said, the desire to harness remote multi-teraflop machines to enable improved results is at odds with a fundamental requirement of the system: absolute end-to-end *time criticality*. The implications of this requirement – on-demand, low latency, high bandwidth, reliable access to multi-teraflop parallel computers from the operating room – are the subject of this work.

2 Grid Infrastructure

The application described in this work sets requirements for "Grid" infrastructure [7,8], which would enable large-scale, remote access to distributed computational and data resources. The central focus of the original Grid infrastructure was the creation of standard protocols to enable uniform access to heterogeneous hardware platforms [9], embodied in the Globus toolkit [10]. Over the last several years, this strategy has shifted to one of integration with emerging enterprise distributed computing software systems, notably Web Services [11]. While progress has been made in these areas, the advances have generally been in the reduction of complexity in coordinating heterogeneous resources, perhaps at the cost of an increase in complexity of system administration and application development. In that context, the type of infrastructure development in the application described in this work may be considered as a driver of grid requirements.

A decade from now, we envision a worldwide network of computationally-assisted neurosurgery operating rooms (CAN-ORs). These CAN-ORs will include some combination of local compute clusters, as in our prototype at BWH, as well as remote, national-scale, shared computational infrastructure. The appeal of using local cluster

infrastructure is clear, as it sidesteps risks associated with contention for the remote shared resources. Unfortunately, a pure dependence on local resources sacrifices the greater precision and accuracy promised by the improved biomechanical models, discussed below, that require more substantial computational power. We are therefore attempting to delineate the broader infrastructure problems that must be solved in order to enable time-critical applications such as described here for the CAN-Ors.

The critical requirement of this application is the tight end-to-end time constraint. This requirement in turn places constraints on the performance and reliability of networks and the performance of computational resources. For large shared supercomputers, the dominant component of end-to-end performance is often that of queue wait time. Common supercomputing wisdom indicates that this queue delay depends on system policies, as well as on the size of the requested compute allocation. For example, it is common for systems to be configured to favor requests for large numbers of CPUs, but to also be able to "squeeze in" or "backfill" [12] jobs that require only a few CPUs and are short in duration.

The Teragrid [13] is an NSF funded grid infrastructure across multiple academic and research sites within the US. We measured queue delays on two of the Teragrid clusters over a three day period, for requests of 5 minutes of wallclock time, from 2 to 64 CPUs. We submitted only a single job at a time; if the job had not started within the critical timeframe of 10 minutes, the submission was terminated, and the next request processed. The purpose of this measurement was to illustrate the likelihood that the finite element simulations would be able to run on a Teragrid cluster under our stringent time constraints. In all, we submitted 313 jobs to the NCSA Teragrid cluster, and 332 to the SDSC Teragrid cluster, which represents from 50 to 56 jobs of each size, on each cluster.

Figure 1 shows the percentage of tasks that advanced through the queue successfully, and started to execute before the 10-minute time limit elapsed. As one might predict, both clusters show a decreasing likelihood of success with increasing number of requested CPUs. This decline is more dramatic for the NCSA machine, which is larger and thus currently more in demand than the SDSC cluster. Figure 2 shows the mean total queue delay for the tasks that did make it through the queue successfully. Again, there appears to be a direct relationship between the size of request and the length of the queue delay, although it would also appear that the two clusters exhibit different performance profiles. The behavior of queuing systems clearly merits further study, including more rigorous statistical characterization based on much larger data sets.

These results just illustrated depict clearly the challenge this application faces. Even for very modest resource requests, the likelihood that CPUs will become available within the critical time period is very low. One possible alternate solution would be to simply submit each job to many distributed resources. We prototyped such a "flooding" scheduler, that submits a single job to many resources; as soon as the job runs on one resource, it is removed from all of the other queues. This approach has worked well in our early tests and demonstrations. However, this technique would be disastrous if used on a large scale, as it would undermine the ability for batch schedulers to make reasonable choices. Furthermore, the issue of data transfer would become even more onerous if we had even less ability to predict where the job would run. What is really required is a scheduler that can infer the likelihood of success on various resources, again a subject for future work.

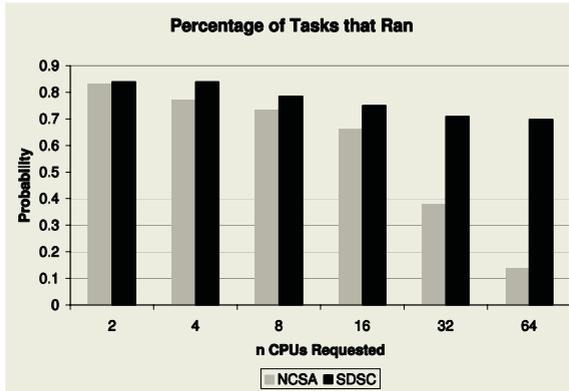


Fig. 1. Percentage of submitted tasks that ran, as a function of CPUs requested

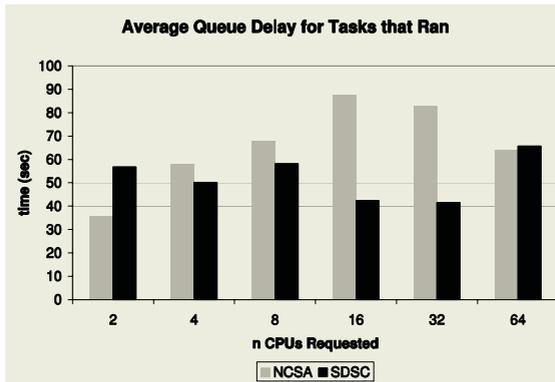


Fig. 2. Average queue delay for tasks that began running within 10 minutes

2.1 Network and Data Transfer

Access to remote supercomputers implies the need for reliable transfer of input and output data sets. Of course the most desired solution would be to use a dedicated network with guarantee Quality of Service (QoS), however the reality is that the network will be shared, so our system will ultimately need to detect and react to a variety of changing conditions. Also, as we envision a scenario where there are 100s of CAN-ORs sending time critical requests for simulation to say 10s of supercomputer centers across the nation, the expectation of a dedicated network may not be a realistic solution. We have performed preliminary tests of network speed for transfer of 20MB file (this is the size of the volumetric brain deformation result from the current finite element model) between SDSC's TeraGrid [13] machine and a machine residing inside the firewall at the BWH hospital. Note that BWH, as with every hospital that must protect patient privacy, is behind a firewall, that in this case only allows a single port of entry to the internal network.

We have investigated also Storage Resource Broker (SRB) [14] as a means to transfer and share data between BWH and SDSC. The SRB is client-server middleware that provides a uniform interface for connecting to heterogeneous data resources over a network. The BWH has acquired a SRB rack as a part of their research collaboration with the Biomedical Informatics Research Network (BIRN) [15]. The SRB server at BWH exists outside the firewall and enables a direct feed of data onto the BWH SRB from the SDSC TeraGrid cluster; this means that a machine inside the firewall must retrieve the data from the BWH SRB server.

We also measured file transfer rates using the `globus-url-copy` [10] which demonstrates slower performance going into BWH compared to coming out of BWH. This is again due to the firewall, which allows only one port of entry into BWH but allows multiple ports to go out of BWH firewall. The following table compares the results for the transfer of a 20 MB file using `globus-url-copy`, SRB and SCP (with two different sets of command line options.)

Table 1. Time to Transfer 20 MB file

Transfer Direction	globus-url-copy	SRB	scp	scp -C
TG to BWH	50	49	68	31
BWH to TG	9	12	40	30

3 Advanced Biomechanical Model Development

Our finite element simulation model, currently under development, is based on the conforming hierarchical adaptive refinement method (CHARMS) [16,17]. Inspired by the theory of wavelets, this refinement method produces globally compatible meshes by construction. The framework based on CHARMS avoids mesh compatibility problems, and provides extended support for object oriented design. CHARMS based solver FAMULS [18] provides an ideal choice for time critical applications. The present objective is to replicate results produced by the existing crude biomechanical finite element solver, currently in use by the prototype system at BWH, and prove the effectiveness of the new solver, called FAMULS which has the additional capability of adaptive mesh refinement. Figure 3 shows simulation of a deformation case with the FAMULS elastic solver using adaptive mesh refinement which added additional tetrahedra in the interior to reduce error. The color indicates deformation in the direction of y axis. This case reproduces the deformation that was originally calculated by the existing crude finite element model.

The current solver is based on the small strain isotropic elastic principle. However, in order to expand the capacity of the system to predict brain deformation ahead in time beyond the deformation shown by the intraoperative 3D volumetric MRI, the accuracy and precision of the simulations must be improved. We are working to replace the current biomechanical model with an anisotropic non-homogeneous visco

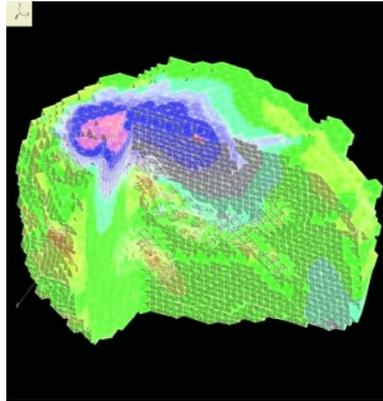


Fig. 3. Mesh cutoff showing displacement contour in color

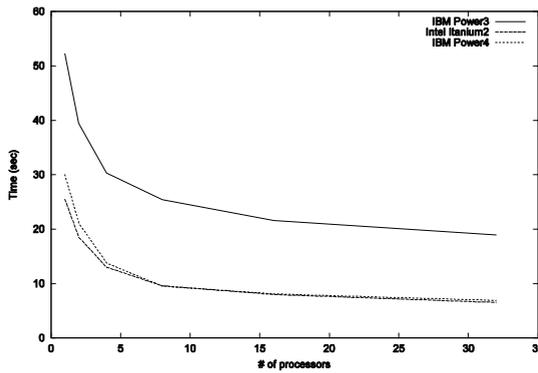


Fig. 4. Performance of the Current Linear Elastic Finite Element Model

elastic and visco plastic, scalable, nonlinear model which will give more accurate predictions. Since this complex model still has to meet the real time constraint of neurosurgery, it requires fast access to much larger computational resources than those typically located in a hospital operating room. This motivates our efforts to create a novel grid architecture for real-time data driven simulations during image guided neurosurgery.

In order to plan our deployment onto a heterogeneous Grid platform, we have started testing performance of the current linear elastic biomechanical model on various parallel computers. We have ported the existing parallel code to the IBM Power3 (8-way node; 375 Mhz; 4 GB memory/node), IBM Power4 (8-way node; 1.5 Ghz; 16 GB memory/node), and Intel Itanium2 clusters (2-way node; 1.5 Ghz; 4 GB memory/node), exemplifying typical heterogeneous platforms that would be available for grid computing involving our application.

Figure 4 shows computation time required for the current linear elastic biomechanical simulation model on these three different parallel machines. In this simulation the total number of meshes was 43584, defining 214035 tetrahedral elements.

The Figure clearly indicates the importance of focusing on improved scalability as we develop the improved model, and further directs our attention to machines with fast internal network fabrics.

4 Conclusions

In the long term, the contribution we are attempting to make is to clarify the requirements for Grid infrastructure to support time-critical medical applications such as IGNS. This infrastructure, which may consist perhaps of hundreds of operating rooms and tens of computational resource centers, will be based on improved networking and software infrastructure. In this paper, we have reported our initial progress, and have discussed some areas in which further progress is required.

The described research involves development and deployment of an integrated and practical grid architecture for focused, on demand computer-assisted neurosurgery. The initial working prototype system currently operating at BWH, using a homogeneous, linear elastic biomechanical material model of the brain, is being improved both in terms of algorithmic enhancements as well as grid support infrastructure for more reliable processing and response time.

Our future plans include the continued development of the advanced biomechanical parallel finite element model based on FAMULS. The parallel scaling of this advanced model will also be improved such that the simulation can be performed within the defined time constraint on a multi-teraflop parallel machine. With respect to grid-computing issues, our objectives include the rigorous characterization of the queue wait time of various classes of jobs on many production clusters, as well as issues related to network traffic performance to these geographically distributed clusters from BWH. These two metrics will allow us to predict, with certain confidence, the overall end-to-end time needed for this time critical application of IGNS.

References

1. D.L.G. Hill, J.V. Hajnal, D. Rueckert, S.M. Smith, T.Hartkens, and K.McLeish. A Dynamic Brain Atlas. *IEEE Trans. Medical Imaging*, 18, Issue 8:712--721, 1999.
2. Y. Kawasaki, F. Ino, Y. Mizutani, N. Fujimoto, T. Sasama, Y. Sato, N. Sugano, S. Tamura, and K. Hagihara. High-Performance Computing Service Over the Internet for Intra-operative Image Processing. *IEEE Transaction on Information Technology in Biomedicine*, 8, No. 1:36--46, 2004.
3. S.K. Warfield, F. Talos, A. Tei, A. Bharatha, A. Nabavi, M. Ferrant, P.M. Black, F.A. Jolesz, and R. Kikinis. Real-Time Registration of Volumetric Brain MRI by Biomechanical Simulation of Deformation During Image-Guided Neurosurgery. *Journal of Computing and Visualization in Science*, 5:3--11, 2002.
4. P.M. Black, T. Morairty, E. Alexandar, P. Stieg, E. J. Woodard, P. L. Gleason, C.H. Martin, R. Kikinis, R.B. Schwartz, and F. A. Jolesz. The Development and Implementation of Intra-operative MRI and its Neurosurgical Applications, *Neurosurgery*, 41:831--842, 1997.
5. S.K. Warfield, F. Jolesz, and R. Kikinis. A High Performance Computing Approach to the Registration of Medical Imaging Data. *Parallel Computing*, 24:1345--1368, 1998.

6. M. Ferrant, A Nabavi, M. Macq, F. A. Jolesz, R. Kikinis, and S.K. Warfield. Registration of 3D Intraoperative MR Images of the Brain Using a Finite Element Biomechanical Model. *IEEE Transactions on Medical Imaging*, 20, Issues 12:1384--1397, 2001.
7. Ian Foster and Carl Kesselman, editors. *The grid: blueprint for a new computing infrastructure*. Morgan Kaufmann Publishers Inc., 1999.
8. Fran Berman, Geoffrey Fox, and Anthony J. G. Hey. *Grid Computing: Making the Global Infrastructure a Reality*. John Wiley & Sons, Inc., 2003.
9. Ian Foster, Carl Kesselman, and Steven Tuecke. The anatomy of the Grid: Enabling scalable virtual organizations. *Lecture Notes in Computer Science*, 2150:1--??, 2001.
10. <http://www.globus.org>.
11. Foster, C. Kesselman, J. Nick, and S. Tuecke. The physiology of the grid: An open grid services architecture for distributed systems integration, 2002.
12. D. Lifka. The anl/ibm sp scheduling system. In D.G. Feitelson and L. Rudolph, editors, *Job Scheduling Strategies for Parallel Processing*, volume 949 of *Lecture Notes in Computer Science*, pages 295--303. Springer-Verlag, 1995.
13. <http://www.teragrid.org>.
14. <http://www.sdsc.edu/srb/>.
15. Biomedical Informatics Research Network (BIRN), <http://www.nbirn.net>.
16. P. Krysl, A. Trivedi, and B. Zhu. Object Oriented Hierarchical Mesh Refinement with CHARMS. *International Journal of Numerical Methods in Engineering*, Vol. 60, Issue 8:1401--1424, 2004.
17. J. D. Mcnamara. Health Monitoring of Rail Road Tracks by Elastic Wave Based Monitoring Technique. *Ph.D Thesis*, Structural Engineering Department, University of California San Diego, 2004.
18. <http://hogwarts.ucsd.edu/~pkrysl/software.html>.