Grand Challenge: Leveraging Extreme-Scale Supercomputers for Modeling the Human Sustainability Network

Christopher D. Carothers (CS)
James Myers (CCNI)
William Wallace (ISE)
Rensselaer Polytechnic Institute
110 8th Street Troy, New York 12180
chrisc@cs.rpi.edu

Extended Abstract

Over the past 18 years, radical advances have been made in the scale and fidelity of network modeling and simulation. In 1994, we leveraged a small cluster of 8 DECstation systems over 100 Mbit Ethernet to model a PCS network consisting of 1000 base stations and 16K subscribers [3]. By 2003, multicore/hyper-threaded processors had advanced to enable a large-scale model of 250,000 packet-level, TCP flows over a real AS topology (i.e., the AT&T backbone) [15]. By 2006, the memory subsystem of commodity multicore SMPs had grown in memory capacity (e.g., 64 GB) to enable high-fidelity "slice" level, BitTorrent, peer-to-peer models composed of over 100,000 clients in a single common swarm [7, 8].

Over this same period, we observed similar advances in supercomputer systems. In 1999, IBM started the $100M Blue Gene project to build a massively parallel computing system for the purpose of investigating biological processes like protein folding [1]. The Blue Gene/L system was fielded in 2004 (initially with 16 racks) and overtook Japan's Earth Simulator as the world's fastest supercomputer with a LINPACK performance of nearly 71 teraFLOPS and grew to over 478 teraFLOPS with 104 racks by 2007. With the introduction of the Blue Gene/P system, core counts grew to 294,000 yielding a performance of over 1,000 teraFLOPs or 1 petaFLOP.

On parallel-discrete-event simulation (PDES) benchmarks, the previous mentioned PCS network model and HPC network models, these Blue Gene systems have enabled events-rates from the 100's of million to an excess of 12 billion per second [2, 9, 11]. To put this advance in computational power in prospective, the PCS network model in 1994 only achieved an event-rate of 25,000 for a configuration that was $1/10,000$ the size yielding a 480,000 times performance improvement. Additional PDES speedup results have been obtained on a Cray XT5 supercomputer system for radio signal propagation models (e.g., TLM) using even more processors [12].

As we look to current top supercomputers, Mira and Sequoia are 10 and 20 petaFLOP Blue Gene/Q systems with 768 and 1,536 TBs of RAM respectively (see www.top500.org). These systems represent another level of unprecedented computational power that could be leveraged to further improve the performance of large-scale network simulations. In early, unpublished performance benchmarking on the Blue Gene/Q system (1 rack, 1024 nodes with 65,536 MPI ranks), PDES performance grows by over a factor of 4x compared to the previous generation Blue Gene/P system. This suggests that scalable network models could compute events at a rate ranging from 50 billion to 150 billion per second depending on the per-event computational complexity. As we look toward future exascale supercomputer systems, there is the potential to increase this performance by an additional factor of 20 to 50x depending on model characteristics.

So with these new and upcoming supercomputer systems, what might be a Grand Challenge problem for Network Science that would require this scale and magnitude of computational resources? We submit that the Internet across all it's vertical and horizonal scales has become an integral piece if not the central glue of what one might call the *human sustainability network (HSN)*. The HSN is really the network of networks on which our collective daily lives depend. This includes power networks (e.g., smart grids), fresh water networks, fuel/pipeline networks, transportation networks (airplanes, trains, boats, trucks and cars), retail networks (Walmart, Target), financial networks (markets, banks, credit and insurance companies), manufacturing networks and finally social networks. Because of the coupling of all of these networks by the Internet (via wired, wireless, cellular and satellite communication systems), we are now seeing how a drought in India can lead to a power outage impacting over 600,000,000 people [10]; how errant code in a trading bot results in a stock market flash crash [14]; and how the housing bubble results in a freeze of credit markets leading to the inability of industrial giants like General Electric to finance their daily operations [13].

All these networks are interconnected at various vertical and horizonal levels and it appears that their level of interdependence will only grow in the future. Thus, the central question is how do we provide optimization and sustainability services such that we can provide some guarantees on a network's individual as well as the collective quality-of-service and in the end improve our overall human sustainability network.

With this grand challenge, there are a number of key problem areas and questions. These include (but are not limited to):

1. Treatment of Network Science as first-class "Big Science" and obtaining suitable allocations on leadership class computing systems on par with the DOE INCITE program for Computational Science.
2. Obtaining, organizing, and searching semantically related network data repositories like CAIDA.org, Data.gov, and the NSF DataNet SEAD project (`www.sead-data.net`) to create realistic topology and HSN scenario definitions.
3. Design a hierarchical HSN model that minimizes memory and network data movement to enable efficient execution on extreme-scale, supercomputer systems.
4. Identify the components of the HSN that require emulation vs. simulation (e.g., what parts could be simulated and at what degree of fidelity?).
5. Develop models for trust and security across the HSN. For example, how does a smart grid provider protect their power generation systems yet still enable remote access maintenance? What software emulations or models for simulation do we need to tackle this problem? (See [6] for the security challenges in the power grid).
6. Creation of load distribution algorithms that respond to changes in the emulated or simulated HSN event workload.
7. Creation of new modeling approaches that efficiently handle high fan-out connections that can occur in the HSN (e.g., peer-to-peer, wireless and satellite).
8. Design and implement efficient HSN optimization algorithms that execute within the overall HSN modeling framework and improve overall network functionality. An example of a potential approach is in [4].
9. Develop in-situ data analytics must be designed and are able to co-execute with the HSN simulation to reduce the amount of data that must be stored.
10. Development of uncertainty quantification methods tailored to Network Science.

It is very clear the Internet has become an integral part of socio-economic fabric of the world and it is influenced by and influences the real-world. This inter-network dependence suggests there are deep, fundamental challenges governing its continued and future design, implementation and operation. By leveraging current and future leadership class supercomputer systems, we can potentially gain greater insight into the behavior of the HSN and design better approaches that insure its sustainability and overall performance.

References

[1] N. R. Adiga and et al., "An Overview of the Blue Gene/L Supercomputer", In *Proceedings of the ACM/IEEE Conference on Supercomputing*, November 2002, pp. 1–22.

[2] D. W. Bauer Jr., C. D. Carothers, and A. Holder. "Scalable Time Warp on Blue Gene supercomputers.", In *Proceedings of the 2009 ACM/IEEE/SCS 23rd Workshop on Principles of Advanced and Distributed Simulation*, pages 35–44, Washington, DC, USA, 2009. IEEE Computer Society.

[3] C. D. Carothers, R. M Fujimoto, Y-B Lin and P. England, "Distributed Simulation of Large-Scale PCS Networks,", In *Proceedings of the Second International Workshop on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems*, pages 2–6, 1994.

[4] B. Cavdaroglu, E. Hammel, J. E. Mitchell, T. C. Sharkey, W. A. Wallace, Integrating Restoration and Scheduling Decisions for Disrupted Interdependent Infrastructure Systems, Annals of Operations Research, In Print. (Online first, URL: http://dx.doi.org/10.1007/s10479-011-0959-3), 2011.

[5] A. Holder and C. D. Carothers, "Analysis of Time Warp on a 32,768 Processor IBM Blue Gene/l Supercomputer," In *2008 Proceedings European Modeling and Simulation Symposium (EMSS)*, 2008.

[6] GAO Report #11-117, "Electricity Grid Modernization: Progress Being Made on Cybersecurity Guidelines, but Key Challenges Remain to be Addressed", `http://www.gao.gov/products/GAO-11-117`.

[7] R. LaFortune, **C. D. Carothers**, W. D. Simth, J. Czechowski and X. Wang, "Simulating Large-Scale P2P Assisted Video Streaming", In *Proceedings of the Hawaii International Conference on System Sciences (HICSS-42)*, Waikoloa, Big Island, Hawaii, January 2009.

[8] R. LaFortune, C. D. Carothers, W. D. Smith and M. Hartman, “An Abstract Internet Topology Model for Simulating Peer-to-Peer Content Distribution”, In *Proceedings of the ACM/IEEE/SCS 21th Workshop on Principles of Advanced and Distributed Simulation (PADS ’07)*, pages 152–162, June 2007.

[9] N. Liu and C. D. Carothers. “Modeling billion-node torus networks using massively parallel discrete-event simulation”, In *Proceedings of the 2011 IEEE Workshop on Principles of Advanced and Distributed Simulation*, PADS ’11, pages 1–8, Washington, DC, USA, 2011. IEEE Computer Society.

[10] J. Romero, “Lack of Rain a Leading Cause of Indian Grid Collapse”, IEEE Spectrum, Energy-Wise Blog, July 31, 2012, `http://spectrum.ieee.org/energywise/energy/the-smarter-grid/disappointing-monsoon-season-wreaks-havoc-with-indias-grid`

[11] K. S. Perumalla, “Scaling Time Warp-based Discrete Event Execution to 10^4 Processors on a Blue Gene Supercomputer,” In *Proceedings of the 4^{th} International Conference on Computing Frontiers*, 2007, pp. 69–76.

[12] S. K. Seal and K. S. Perumalla, “Reversible Parallel Discrete-Event Formulation of a TLM-Based Radio Signal Propagation Model”, *ACM Transactions on Modeling and Computing Simulation*, Volume 22, Number 1, 23 pages, 2012.

[13] P. Schweizer, *Throw Them All Out*, Houghton Mifflin Harcount Publishing Company, New York, 2011.

[14] E. Ullman, “Errant Code? It’s Not Just a Bug”, New York Times, The Opinion Pages, August 8th, 2012, `http://www.nytimes.com/2012/08/09/opinion/after-knight-capital-new-code-for-trades.html`

[15] G. Yaun, C. D. Carothers, and S. Kalyanaraman, “Large-Scale TCP Models Using Optimistic Parallel Simulation”, In *Proceedings of the 17^{th} ACM/IEEE/SCS Workshop on Parallel and Distributed Simulation (PADS ’03)*, page 153– 161, June 2003.