Very Large-Scale Modeling and Simulation on High Performance Computing Systems

Robert F. Lucas & Dan M. Davis
Information Sciences Institute
University of Southern California
Marina del Rey, California 90292

David R. Pratt & Philip Amburn
Strategies, Simulation and Training Business Unit
Science Applications International Corporation
Orlando, Florida 32826

SUMMARY

The authors typically work independently, but collegially, for the U.S. Army Program Executive Office, Simulation, Training and Instrumentation (PEO STRI), the High Performance Computing Modernization Program (HPCMP) and the US Joint Forces Command (JFCOM). Under direction from those organizations, they have developed advanced methods of High Performance Computing (HPC) modeling and simulation (M&S) for battlespace analysis and training, a topic of obvious importance to the U.S. Army. These applications involve new algorithmic strategies for use in the Semi-Automated Forces (SAF) family where they utilize computers to reproduce and predict processes of interest in areas related to defense applications. The authors' research includes:

- behavior modeling of hardware, personnel and complex systems
- combat modeling on global-scale battlespaces
- real time visualization techniques
- advanced approaches and algorithms to manage massive, distributed heterogeneous data sets.

The implemented capabilities produce new and effective methods for HPC use in M&S, including:

- network structures utilizing unique software routers that scale to millions of SAF entities
- complex multi-scale adaptive systems reflecting the needs of JFCOM and PEO STRI
- new techniques for the management of data and the analysis of stochastic phenomena
- data mining and advanced analytical techniques.

They set forth approaches, methods and successes in each of these areas.

In the recent past, most military modeling and simulation codes have been run on personal computers connected by local and wide-area networks, a condition which is still common. This approach limits the number of realistically simulated entities and it constrains the sophistication of behaviors and the fidelity of physical world models. Recent research leveraged High Performance Computing to enable the expansion of simulations to represent theater-wide forces, albeit with some limitations. These advances produced the necessary increase of entity counts to better satisfy the needs of representing the entire urban environment. There is a well-defined and critical need to produce quantifiable results, which are derived from these simulations. These are needed in order to support transformation experimentation at JFCOM and training systems developed at the U.S. Army's PEO STRI. The use of HPC in simulations has proven effective at extending the capabilities of both organizations.

These techniques have resulted in the generation of very large and geographically dispersed data sets. Some of the authors' recent experiments required, utilized and integrated multiple scalable parallel processors (SPP) sites distributed across the United States. They have been hosted by the supercomputing centers at Maui (MHPCC) and at Wright-Patterson (ASC-MSRC) on a net including J9 at Suffolk, Topographic Engineering Center, Fort Belvoir, and SPAWAR San Diego.

These distributed and parallel computational assets must be harnessed by a truly scalable code in order to model futuristic sensor technologies and the complexities of the urban environment. A poorly scaling code would render this effort incapable of making use of the computational power of large Linux clusters, some of which have thousands of processors. A typical weeklong SAF simulation event, using 512 processors, generates more than two terabytes of raw data, accruing at the rate of >10GB per hour. The volume and distributed nature of this type of data poses significant challenges in developing the corresponding data-intensive applications that manage and analyze them. The authors built on lessons learned in developing simulations for earlier work at AFIT, NPS, JSIMS, PEO STRI, JFCOM and the Defense Advanced Research Projects Agency (DARPA). A next generation of Semi-Automated Forces has been developed at PEO STRI: OneSAF Objective System (OOS). Also, they have designed and developed a new simulation set-up tool set to make it easier for the analyst to modify entity behaviors, as opposed to requiring programmers to do it. A JFCOM team has designed a new data management and analysis tool, called Simulation Data Grid (SDG). It has now been developed and implemented. The design principles driving the architecture were:

- minimize network communication overhead (especially across SPPs)
  - store data near the point of generation
  - only selectively propagate the data as needed, and
- maximize the use of SPP computational resources and storage by:
  - distribute analyses across SPP sites
  - locally reduce, filter and aggregate raw data

The key implementation principle in this development was to leverage existing open standards and infrastructure from the Grid Computing initiatives. SDG services include distributed data query/analysis, data cataloging, and data gathering/slicing/distribution. It can be argued that SDG has proven to be a useful general-purpose tool for a range of simulation domains. Over all, the authors can report many successful implementations giving the Army user, analyst, trainer and trainee new capabilities, not previously possible. In each case, vast quantities of legacy code were considered to be of irreplaceable value, so care was taken to maintain compatibility with those codes.

Despite these manifold successes, substantial major advances in computing capability are required by the U.S. Army for future battlespace modeling and simulation. These are needed to achieve the requisite increases in the fidelity of both entity behaviors and physical phenomena in order to faithfully represent military operations. Fortunately, these advances are being actively considered and fostered. The DARPA High Productivity Computing Systems (HPCS) Program is developing a new generation of economically viable, high performance computing systems for supporting the U.S. national security missions. HPCS systems are not only being designed to run faster and more efficiently, but also to reduce code-development costs, facilitate ports amongst machines, and enhance recoveries from system failures. Proposed HPCS systems feature hundreds of thousands of processors and shared global address spaces, all made accessible by new generations of programming languages.

The authors will discuss the potential that these systems will deliver to the Army: incorporation of C4ISR system models (including wireless signals propagation, router properties, other networking infrastructure), fully modeled structures, dynamic terrain, realistic weather, complex cognitive representations, entity learning, and first order physics-based models of local events. There will be a reduction in the constraints on models of sensors and on human behavior such as Political, Military, Economic, Social, Infrastructure, and Information (PMESII). HPCS systems should become available early in the next decade. A roadmap is presented for developing software to exploit these systems and to better simulate modern urban warfare, combat systems, and global scale battlespaces.

AUTHORS

Robert F. Lucas is the Director of the Computational Sciences Division of the University of Southern California's Information Sciences Institute (ISI). He manages research in large-scale simulations, computer architecture, VLSI, compilers and other software tools. He has been the principal investigator on the JESPP project since its inception in 2002. Prior to joining ISI, he was the Head of the High Performance Computing Research Department for the National Energy Research Scientific Computing Center at Lawrence Berkeley National Laboratory, the Deputy Director of DARPA's Information Technology Office, a member of the staff at the Institute for Defense Analyses' Center for Computing Sciences and an engineer at Hughes Aircraft. Dr. Lucas received BS, MS, and PhD degrees in Electrical Engineering from Stanford University in 1980, 1983, and 1988.

David R. Pratt is Chief Scientist and a Fellow at Science Applications International Corporations' Applied Software Systems Engineering Technology Group. As the Group's Technical Lead, he coordinates the internal research in modeling and simulation, high performance computing and the training of tri-athletes. Prior to joining SAIC he was Technical Director of the Joint Simulation System (JSIMS) Joint Project Office. He saw active duty as an officer in the US Marine Corps. Dr. Pratt received a BSE in Electrical Engineering from Duke University and a Ph.D. and M.S. in Computer Science from the Naval Postgraduate School.

Philip Amburn currently is employed by SAIC as the Productivity Enhancement and Technology Transfer (PET) on-site professional for Forces Modeling and Simulation at Wright-Patterson AFB OH. His research interests are constructive and virtual simulation, interactive 3D graphics, and visualization. LtCol Amburn taught computer science at the Air Force Institute of Technology up until the time of his retirement in 2000. Dr. Amburn received his Ph.D. degree in Computer Science from the University of North Carolina, Chapel Hill, his MSCS degree in from the Air Force Institute of Technology, and a BS degree in Physics from Kansas State Teachers College.

Dan M. Davis is the Director, JESPP Project, Information Sciences Institute (ISI), University of Southern California, and has been active in large-scale distributed simulations for the DoD for more than a decade. While he was the Assistant Director of the Center for Advanced Computing Research at Caltech, he managed Synthetic Forces Express, a major simulation project. Prior to that, he was a Software Engineer on the All Source Analysis System project at the Jet Propulsion Laboratory and on a classified project for Martin Marietta. An active duty Marine Cryptologist, he retired as a Commander, USNR. He received a B.A. and a J.D., both from the University of Colorado in Boulder.