Urban Reconnaissance Planning: Discovering New Applications with Tactical High Performance Computing

Song Jun Park, Stephen Allen, and Dale Shires
U.S. Army Research Laboratory
Aberdeen Proving Ground, MD, USA
{song.j.park.civ,stephen.allen43.ctr,dale.r.shires.civ}@mail.mil

Abstract— **High performance computing (HPC) has been a key enabler for discovery in science, engineering, and business. HPC being a catalyst for innovation, new capabilities are anticipated for the concept of tactical HPC. The goal of tactical HPC project is to advance computational capabilities of individual Soldiers at the tactical edge by leveraging mobile HPC systems. Since tactical edge is complex, dynamic, and uncertain, significant opportunities exist for real-time computing. As an example of enabling new capabilities at the tactical edge, compute-intensive urban reconnaissance planning application is presented. With GPU accelerator augmented systems, this computationally expensive algorithm was solved at interactive speeds. We demonstrate enhancing situational awareness and tactical intelligence by deploying HPC-in-a-box technology to the tactical edge.**

Keywords—Tactical HPC; situational awareness; GPU accelerators

I. Introduction

Computing technologies have been a key enabler for the military and society in general. High performance computing (HPC) in particular has enabled aircraft design, weather prediction, materials modeling, and drug discovery to name but a few. Even everyday tasks such as voice search and social networking are made possible by HPC infrastructures. In similar manner, HPC can have a significant impact to Soldiers in tactical environments.

By definition, tactical environments are complex, dynamic, and uncertain in nature. Complexity and uncertainty characteristics in tactical settings increase the computational requirement of potential algorithms that aims to enhance tactical intelligence. Unavoidable dynamic changes in tactical environments further compress the processing time requirement. Here, on-demand high performance computing is needed in addition to a strategic supercomputing that rely on reach back to a supercomputing center.

The vision of tactical HPC is to empower individual Soldiers with mobile supercomputing. The concept is to deploy high-performance systems to the tactical edge by packing as much computational power as it can be contained within a vehicle. To maximize tactical HPC's raw compute capacity, parallel and heterogeneous computing framework was adopted for constructing a HPC-in-a-box system.

A mobile HPC in tactical operations opens possibilities for addressing a new class of computationally challenging applications. High computing capacity in the field allows for solving compute-intensive algorithms previously reserved for supercomputing facilities. Example application areas include situational awareness, data analytics, decision making, and prediction modeling. In this respect, tactical HPC can have a broad impact in solving a wide range of computationally expensive applications in tactical environments.

As a concrete example of a newly enabled capability with tactical HPC, the urban reconnaissance planning application is presented. The application illustrates a demonstration of computing situational awareness at interactive speeds at the tactical edge. The example conveys how an access to a mobile HPC augments urban tactical intelligence amid a mission by deriving threat calculations in a 3D environment. To achieve an interactive execution time, accelerator equipped heterogeneous systems are evaluated for performance. With the demise of Dennard scaling, advancement in computing power calls for alternative and unconventional architectures. Since the application called for a high requirement of floating-point operations in non-recursive ray tracing algorithm, massively parallel GPU architectures offered an ideal mapping.

II. Tactical HPC System

Computing is in the midst of power limited era [1]. Across the spectrum, from mobile devices to supercomputers, power-efficiency is the central and primary issue. The initial target application area for tactical HPC was floating-point intensive algorithms, and hence GPU technology proved optimal for maximizing computational power. Advancing compute capacity in power-efficient manner mandates leveraging accelerators such as GPUs [2]. GPUs trade off speed of a single thread operation for a collective throughput of many threads. Originally designed for driving displays, GPUs can effectively process SIMD structured algorithms in parallel. In a sense, graphics cards have a dual-purpose where it can be viewed as a specialized hardware for floating-point processing.

The TOP500 list ranks the world's most powerful supercomputing systems by measuring LINPACK [3] benchmark performance. The ranking list is produced bi-annually in June and November. At the bottom of the list, a system with LINPACK performance of 206 TFLOPS in double-precision and peak of 251 TFLOPS in double-precision qualified for the November 2015 TOP500 list [4]. To compare a potential computational power of a tactical HPC system to the low-end TOP500 system, a peak performance value can be estimated for a multi-GPU machine with NVIDIA Tesla K80. Tesla K80 GPU is rated at 8.74 TFLOPS in single-precision [5]. Suppose that eight Tesla K80 cards are housed in a 4U server, then this server would possess a theoretical peak of 70 TFLOPS in single-precision. If four of these servers can be installed in a vehicle, then the aggregate peak performance would be in the 280 TFLOPS range. Albeit single-precision, for applications acceptable in 32-bit precision, a collective multi-GPU servers in a mobile platform is competitive to a system at the bottom of the 2015 TOP500 list. In other words, a lower-precision supercomputer or tactical HPC nodes can be placed in an MRAP vehicle. Physical dimension requirement for four 4U multi-GPU servers can be estimated to be 28 inches in height, 19 inches in width, and 30 inches in depth. Fig. 1 depicts the tactical HPC concept in an urban environment.

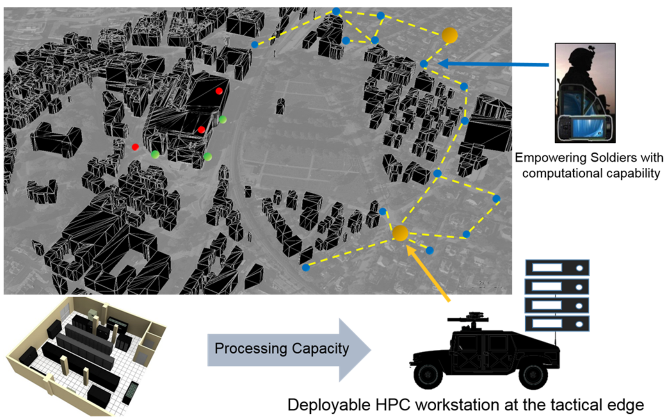

Figure 1. Tactical HPC concept illustration.

Although heterogeneous hardware provide HPC capacity, a challenge exists in programming to extract performance. Currently, heterogeneous computing lacks a strong software foundation. Diversity in architectures for heterogeneous computing (ARM, GPU, Xeon Phi, FPGA, etc.) further increases programming models and frameworks for software development. With increasing number of APIs, it is difficult and not realistic for a programmer to be an expert in every programming languages. Yet, performance is directly related to understanding optimization techniques and architecture details specific to each architecture [6].

Innovations in both hardware and software need to be considered as miniaturization of semiconductor manufacturing process collides with the fundamental law of physics. Projections indicate the fabrication process to reach the limit by as early as year 2020, at which point, further advancement in computing is going to rely on specialized architectures and software optimizations. Moving away from the dependence on shrinking of transistor feature size will likely stimulate and spur innovative solutions.

III. Reconnaissance Planning Application

A situational awareness application was developed to showcase the potential capabilities of a tactical HPC. Reconnaissance planning application demonstrates a novel capability for individual Soldiers operating at the tactical edge, providing threat and surveillance map data at interactive speeds. The core of the application contains non-recursive ray tracing and geospatial optimization algorithms.

The first stage of the application computes the ballistic hit probabilities for hostile entities. This can be viewed as a full 3D line-of-sight (LOS) calculations with ballistic hit probability. The algorithm is computing ray-triangle intersections for LOS of a hostile entity and applying a 5th order polynomial for a hit probability. LOS is computed for the 3D scene, although the threat map is visualized on the ground plane because of the potential reconnaissance locations. Fig. 2 shows the ballistic threat risk map as shaded in red, which denotes the surfaces the hostile entity can target. Similarly, surveillance observability map, denoted by the green overlay, is generated using the non-recursive ray tracing algorithm as shown in Fig. 3. Non-recursive ray tracing algorithm possesses a large data parallelism and benefits from parallel architectures.

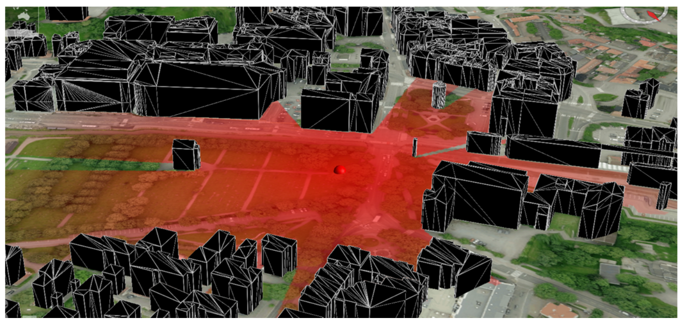

Figure 2. Ballistic threat risk map shaded in red for the red node.

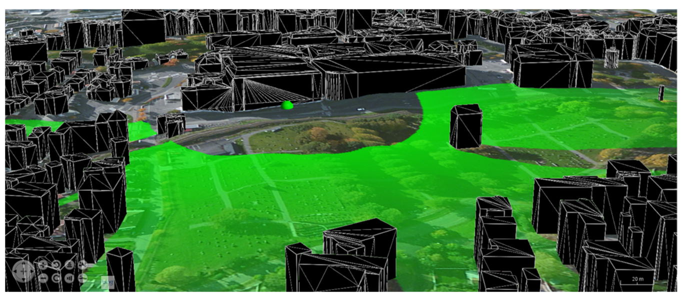

Figure 3. Surveillance observability map indicating areas with line-of-sight to the green node.

Once a ballistic threat risk map and a surveillance observability map are computed, optimal positioning of blue force is calculated by geospatial optimization. The objective is to minimize the risk of ballistic threat while maximizing the visual quality to surveillance targets. Markov chain Monte Carlo is employed to search for the best positioning of blue force soldiers. Observe the following scenario where a reconnaissance mission is to perform a surveillance of two entry points (green nodes) given two hostile guards in an urban environment. Fig. 4 illustrates the computed optimal blue force locations given two hostile threats (red nodes) and two surveillance targets (green nodes).

For a scenario consisting of four hostile threats and four surveillance targets, execution time at interactive speeds were achieved by leveraging multiple discrete GPUs. GPUs provided cost-effective, enhanced floating-point capacity to computing systems. This CPU and GPU setup of heterogeneous architectures made power-efficient computational platforms that can serve as a tactical HPC.

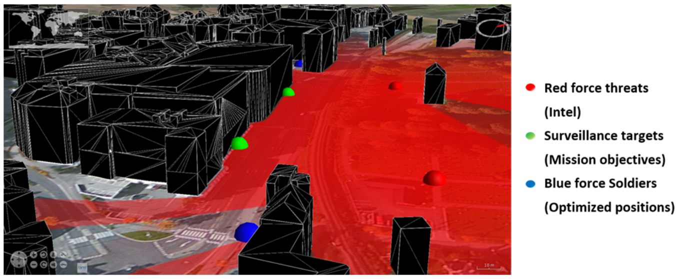

Figure 4. Reconnaissance planning of blue force for two hostile threats and two surveillance targets.

As manifested in the reconnaissance planning application, we believe tactical HPC creates opportunities for discovering new applications that can now be computed in real-time and dynamically at the tactical edge.

IV. VISUALIZATIONS

Tactical HPC enables individual Soldiers operating at the tactical edge with computational capabilities. To Soldiers, visualization techniques assist in conveying information and interface usability. We describe the visualization layer in this section.

For the user interface of the urban reconnaissance planning application, a virtual globe developed by NASA called, World Wind, was employed for ease of use input and output. World Wind is open source and cross platform for interactive visualizations of 3D globe. World Wind is available with the NASA Open Source Agreement license.

The implementation of the first person view feature adds animation to the demonstration that enhances understanding of the application. After selecting the first person view from the context menu, a JDialog box will appear on top of the screen. This new window shows the world from the view of the selected entity as portrayed in Fig. 5. The view is labeled with the name of the entity and the type of entity (Watch Point, Shooter, and Watcher). The view for enemy soldiers (red nodes) pans horizontally from side to side with the center of the arc being a watch point node. This gives a 180° view of the area surrounding the node. The view for watch points is a stationary camera facing enemy soldiers. The view for the watchers tilts left and right to simulate a running motion. The views are also dynamic, meaning that as changes occur within the main World Wind globe, they will be reflected in the first person view.

An animation was developed to indicate the invocation and busy status of the ray-tracing-based ray-to-plane intersection calculations. For each ballistic threat unit, the animation simulates how the server-side code determines the hostiles' field of view as shown in Fig. 6. The animation stops when the server has sent the results back. Once the results are returned from the server, a blue node will be placed on the map at the location the code has deemed as the optimal solution to the scenario. The number of blue nodes returned is formulated to have a one-to-one correlation with the number of watch points on the map. The returned blue nodes cannot be moved from their initial location and do not have a context menu.

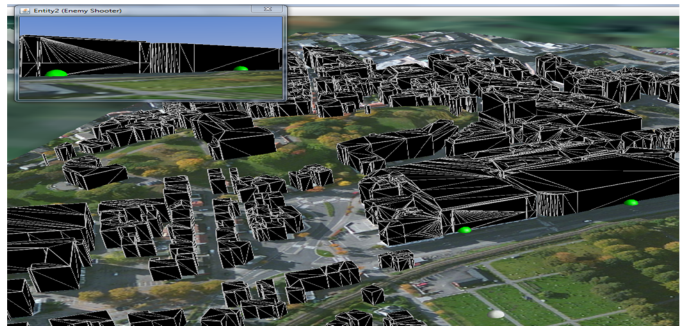

Figure 5. The first person view window of the white node shown at the upper-left corner.

After execution of the Run command, any watcher nodes placed on the map prior the execution will be capable of the utilizing the 'Run A* Path' option from their respective context menus. Choosing this option will cause the node to automatically traverse the map from its starting position to the predetermined destination using the route calculated by the server as shown in Fig. 7. Because World Wind requires mouse motion within the map to move nodes, the program takes control of the mouse to create fluent movement to drive the animation. This causes the user to lose control over mouse movements while the node runs its path. Once the node reaches its destination, mouse functionality will return to normal.

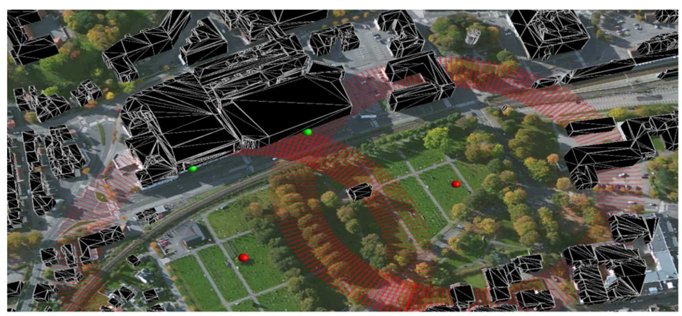

Figure 6. The ray-tracing animation uses red color rays that start at the enemy soldier's position and fan out in all directions interacting with the terrain and the buildings in the area.

The first person view of the watcher node operates differently while running the A* option. Since mouse functions are inactive during a run, the first person view must be activated before the 'Run A* Path' option is selected from the context menu. While the node traverses the path, the first person view has its normal functionality, but when the node reaches its destination the camera will begin to pan left and right in similar fashion to the first person view of enemy shooters. A comprehensive detail on user interface and usage is documented in [7].

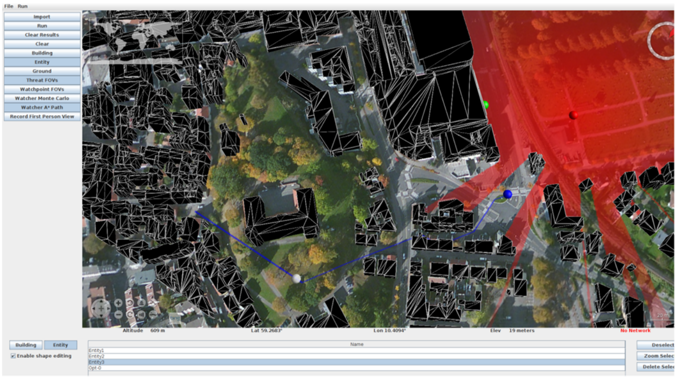

Figure 7. The blue path line shows the route generated by the server to get the watcher from its initial position to the derived optimal point. The path uses the A* algorithm to determine the fastest route to the goal.

V. Related work

The U.S. Marine Corps has been looking to enable tactical forces to have the same capabilities found in Garrison. The Corps is attempting to bring a private cloud to the tactical edge [8]. However, the emphasis of this work is focused on integration opposed to new application capabilities and HPC.

DARPA's Squad X Core Technologies [9] is a program that addresses ensuring dismounted infantry squads to maintain tactical superiority over potential adversaries. Squad X is a broad project where a tactical HPC can assist in advancing squad's information superiority.

DARPA project Content-based Mobile Edge Networking (CBMEN) attempts to design a network for sharing critical information in remote locations [10]. The challenge of real-time information sharing at the tactical edge is the lack of network infrastructure such as cell towers. Communication infrastructure and protocols are being addressed by this effort, which can connect distributed tactical HPC systems.

Multi-GPU systems has been utilized to accelerate computational fluid dynamics [11]. Researchers used two workstation with four GPU cards in each node. An order magnitude speed up is observed for the eight GPU cluster. Similarly, we expect heterogeneous systems to enable tactical intelligence computations near real-time.

Many of the autonomous vehicles under development consist of a processing system onboard. These projects have been leveraging GPUs where GPU computing has been instrumental in deep neural networks [12] and computer vision [13]. Our approach aims to maximize the onboard computational power for solving complex situational awareness problems.

VI. Conclusion and future work

Geospatial optimization for reconnaissance mission planning was computed at interactive speed with tactical HPC. Urban reconnaissance planning is but a one example of tactical HPC enhancing situational awareness for the Soldiers at the tactical edge. Tactical HPC project is focusing on discovering new and enabling capability for a deployed mobile HPC. As HPC touches our daily lives through search engines, weather, and social networking, tactical HPC can similarly have a transformative impact to military operations. We envision that tactical HPC will significantly augment Soldier's tactical intelligence in remote locations.

As for future work, we plan to extend the static setup to account for dynamic movements in red force or threats. Adding dynamic realism would further increase the computational requirement supporting the utility and motive of a tactical HPC. Moreover, machine learning is an area of promising research that can significantly impact tactical operations. The field of machine learning has potential to assist and advise Soldiers in tactical environments with above human classification and maximum reward decision making.

References

[1] R. Hameed, W. Qadeer, M. Wachs, O. Azizi, A. Solomatnikov, B. C. Lee*, et al.*, "Understanding sources of inefficiency in general-purpose chips," *SIGARCH Comput. Archit. News,* vol. 38, pp. 37-47, 2010.

[2] S. Huang, S. Xiao, and W. Feng, "On the energy efficiency of graphics processing units for scientific computing," in *Parallel & Distributed Processing, 2009. IPDPS 2009. IEEE International Symposium on*, 2009, pp. 1-8.

[3] J. J. Dongarra, P. Luszczek, and A. Petitet, "The LINPACK Benchmark: past, present and future," *Concurrency and Computation: Practice and Experience,* vol. 15, pp. 803-820, 2003.

[4] (2015). *Top500 List - November 2015*. Available: http://www.top500.org/list/2015/11/?page=5

[5] R. Smith. (2014). *NVIDIA Launches Tesla K80, GK210 GPU*. Available: http://www.anandtech.com/show/8729/nvidia-launches-tesla-k80-gk210-gpu

[6] N. Satish, C. Kim, J. Chhugani, H. Saito, R. Krishnaiyer, M. Smelyanskiy*, et al.*, "Can traditional programming bridge the Ninja performance gap for parallel computing applications?," presented at the Proceedings of the 39th Annual International Symposium on Computer Architecture, Portland, Oregon, 2012.

[7] S.J. Park, S. Allen, and D. Shires, "Visualization Development of the Ballistic Threat Geospatial Optimization," U.S. Army Research Laboratory, 2015.

[8] R. K. Ackerman. *Marines Bring the Cloud to the Tactical Edge* Available: http://www.afcea.org/content/?q=Article-marines-bring-cloud-tactical-edge

[9] *Squad X Core Technologies Takes First Steps toward Improving Capabilities for Dismounted Soldiers and Marines*. Available: http://www.darpa.mil/news-events/2015-12-10

[10] (2013). *DARPA evaluates content-based mobile edge networking software*. Available: http://www.army-technology.com/news/newsdarpa-evaluates-mobile-edge

[11] M. Griebel and P. Zaspel, "A multi-GPU accelerated solver for the three-dimensional two-phase incompressible Navier-Stokes equations," *Computer Science - Research and Development,* vol. 25, pp. 65-73, 2010.

[12] B. H. Adam Coates, Tao Wang, David J. Wu, Andrew Y. Ng, and Bryan Catanzaro, "Deep learning with COTS HPC systems," in *International Conference on Machine Learning*, Atlanta, Georgia, 2013.

[13] K. Pulli, A. Baksheev, K. Kornyakov, and V. Eruhimov, "Real-time computer vision with OpenCV," *Commun. ACM,* vol. 55, pp. 61-69, 2012.