EXPLOITING ACCELERATOR-BASED HPC FOR ARMY APPLICATIONS

James Ross
High Performance Technologies, Inc (HPTi)
Computational Scientist

Edward Carmack
David Richie
Song Park, Brian Henz and Dale Shires

HPTi
Brown Deer Technology
U.S. Army Research Lab
OUTLINE

- Motivation
- Ongoing Investigation
  - Investigation of Algorithms
  - Octree Algorithm
- Ballistic Threat Simulation
  - Early Prototype
  - Quadtree Search Algorithm
  - Application-Specific Ray Tracer
- Combined Simulation + Visualization
- Results
MOTIVATION

The power of supercomputing in the hands of the Army warfighter
ONGOING INVESTIGATION

- Ongoing 2+ year investigation of GPGPU technology
  - Collaboration/support from BDT and HPTi
- Spanned 3 generations of processor architectures
- Investigation includes:
  - Hardware (Nvidia and AMD/ATI)
  - Programming environments (CUDA, Brook+, OpenCL)
  - Software/algorithm analysis, design, optimization
- Objectives:
  - Investigate hardware performance for Army relevant HPC applications
  - Develop approaches to software design and optimization
  - Develop in-house expertise with GPGPU technology
  - Leverage expertise of government, industry and academic partners
Investigation of Algorithms

- Computational kernels investigated by ARL across range of Army HPC applications:
  - Encryption
  - Image registration
  - Monte Carlo
  - N-Body dynamics
  - Seismic
  - Ballistics
  - Ray tracing
  - Radar image processing
  - Electromagnetics
OCTREE ALGORITHM

- Octree used to represent the recursive bisection of space in 3 dimensions
- Algorithms using octree require tree traversal techniques
- Accelerating data structure for 3D spatial search
  - Application to ray tracing
- Octree partitions 3D space
BALLISTIC THREAT SIMULATIONS

- **Problem:** for a given set of known threats within an urban environment, determine the threat probability at every location

- **Depends on:**
  - 3D polygon representation of the environment
  - Line-of-sight paths
  - Specific ballistic models
  - Dependent on specific ballistic models

- **Applications to research and training**
  - Requires user interaction with the calculation
BALLISTIC THREAT SIMULATIONS

- Components of the calculation
  - First-hit ray tracing to compute line-of-sight / distance
  - Ballistic model(s) for hit probability
  - Accelerating data structures and tree search algorithms
    - Choose quadtree – maps well to 2D cityscapes
    - Adapt octree algorithms from earlier work for use with quadtree
INITIAL PROTOTYPE

- Constructed initial prototype using OpenCL to perform the threat probability calculation on a GPU
- Probability mapped/associated to each polygon in the 3D map
- Algorithm:
  - For each threat, identify polygons with line-of-sight path
  - For each such polygon, apply ballistic model to determine probability of a ballistic hit
- Key component is the quadtree search for ray-polygon hit calculation
- Results post-processed for visualization using Paraview/VTK
QUADTREE SEARCH ALGORITHM

- Quadtree pre-processed on CPU and sent to GPU
- Each cell has an associated start and final index into triangle list
- Performance improvements can be obtained by moving tree data to local memory (not triangle list)

<table>
<thead>
<tr>
<th>Processor / Method</th>
<th>Execution Time (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU / linear</td>
<td>323</td>
</tr>
<tr>
<td>CPU / quadtree</td>
<td>34</td>
</tr>
<tr>
<td>GPU / quadtree</td>
<td>3.4</td>
</tr>
</tbody>
</table>
VISUALIZATION-DRIVEN CALCULATIONS

- Issues with the initial prototype:
  - Partial triangle occlusion leads to false/imprecise probability
  - Calculation performed for non-visible locations that may or may not be of interest.

- Idea: let the visualization drive the calculation
  - For each pixel in rendered image, cast a ray from the camera into the scene geometry
  - The cast a ray to each threat accessible via a line-of-sight path
  - Apply the ballistic model to determine the hit probability to be displayed
**VISUALIZATION-DRIVEN CALCULATION**

- **How the image was formed:**
  - Map copied to GPU
  - Threat probability calculated on GPU
  - Result copied to host
  - Visualization using VTK/Paraview

- **Proof-of-concept for simulation**

- **How the image was formed:**
  - Map copied to GPU
  - Elevation map generated on GPU
  - Threat probability calculated on GPU
  - Combined bitmap generated on GPU
  - Bitmap copied to host

- **Proof-of-concept for combined simulation and visualization**
COMBINED SIMULATION + VISUALIZATION

- Traditional HPC built upon data generation through computational simulation, with visualization as a post-processing step
- GPU-compute capability allows the possibility to tightly couple simulation with visualization
  - Mirrors the OpenCL/OpenGL buffer sharing mechanisms
- Visualization of simulation results can be performed entirely on the GPU
- Combine simulation + visualization opens up interesting applications of GPU-based HPC
DYNAMIC SCENARIO DEMO

- Dynamic scenario demonstration:
  - Shooter moves along a fixed path
  - Hit probability calculated each frame
  - 4 seconds per frame - ray-traced
  - Bitmaps copied back to host
  - Sequenced into simple MPEG

- These initial proof-of-concept demonstrations lead to current work investigating OpenCL/OpenGL buffer sharing for entire simulation + visualization on GPU
INTEGRATION WITH USER INTERFACE

- Demonstrations integrating the ballistic threat simulations with external user interfaces
  - Interactive performance using Google Maps
  - Cross-platform, browser-based API for portability (Android, iOS, PC)
  - Scenario and model selection using simple controls
ONGOING WORK

- Focus on small, powerful workstation-class systems to be placed in critical locations requiring performance
- Exploiting CL/GL buffer sharing
- Tightly coupled simulation and visualization
- Remote access from low-power smart phones and other devices
- Scenario and model selection request
SUMMARY

- ARL investigating use of heterogeneous CPU/GPU platforms
- Application-specific first-hit ray tracer for training in urban environments
- Accelerator data structures to obtain high performance on GPU systems
- Prototype demonstrated combined computation and visualization
  - Cuts out post-processing of simulation data
  - Visualization-driven calculation
- Cross-platform capabilities on diverse, heterogeneous CPU/GPU architectures
- Remote access, scenario generation, and graphical display on mobile platforms
Disclaimer & Attribution

The information presented in this document is for informational purposes only and may contain technical inaccuracies, omissions and typographical errors.

The information contained herein is subject to change and may be rendered inaccurate for many reasons, including but not limited to product and roadmap changes, component and motherboard version changes, new model and/or product releases, product differences between differing manufacturers, software changes, BIOS flashes, firmware upgrades, or the like. There is no obligation to update or otherwise correct or revise this information. However, we reserve the right to revise this information and to make changes from time to time to the content hereof without obligation to notify any person of such revisions or changes.

NO REPRESENTATIONS OR WARRANTIES ARE MADE WITH RESPECT TO THE CONTENTS HEREOF AND NO RESPONSIBILITY IS ASSUMED FOR ANY INACCURACIES, ERRORS OR OMISSIONS THAT MAY APPEAR IN THIS INFORMATION.

ALL IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR ANY PARTICULAR PURPOSE ARE EXPRESSLY DISCLAIMED. IN NO EVENT WILL ANY LIABILITY TO ANY PERSON BE INCURRED FOR ANY DIRECT, INDIRECT, SPECIAL OR OTHER CONSEQUENTIAL DAMAGES ARISING FROM THE USE OF ANY INFORMATION CONTAINED HEREIN, EVEN IF EXPRESSLY ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

AMD, the AMD arrow logo, and combinations thereof are trademarks of Advanced Micro Devices, Inc. All other names used in this presentation are for informational purposes only and may be trademarks of their respective owners.

The contents of this presentation were provided by individual(s) and/or company listed on the title page. The information and opinions presented in this presentation may not represent AMD’s positions, strategies or opinions. Unless explicitly stated, AMD is not responsible for the content herein and no endorsements are implied.