Social Simulations Accelerated: Large-Scale Agent-Based Modeling on a GPU Cluster

Peter Wittek and Xavier Rubio-Campillo

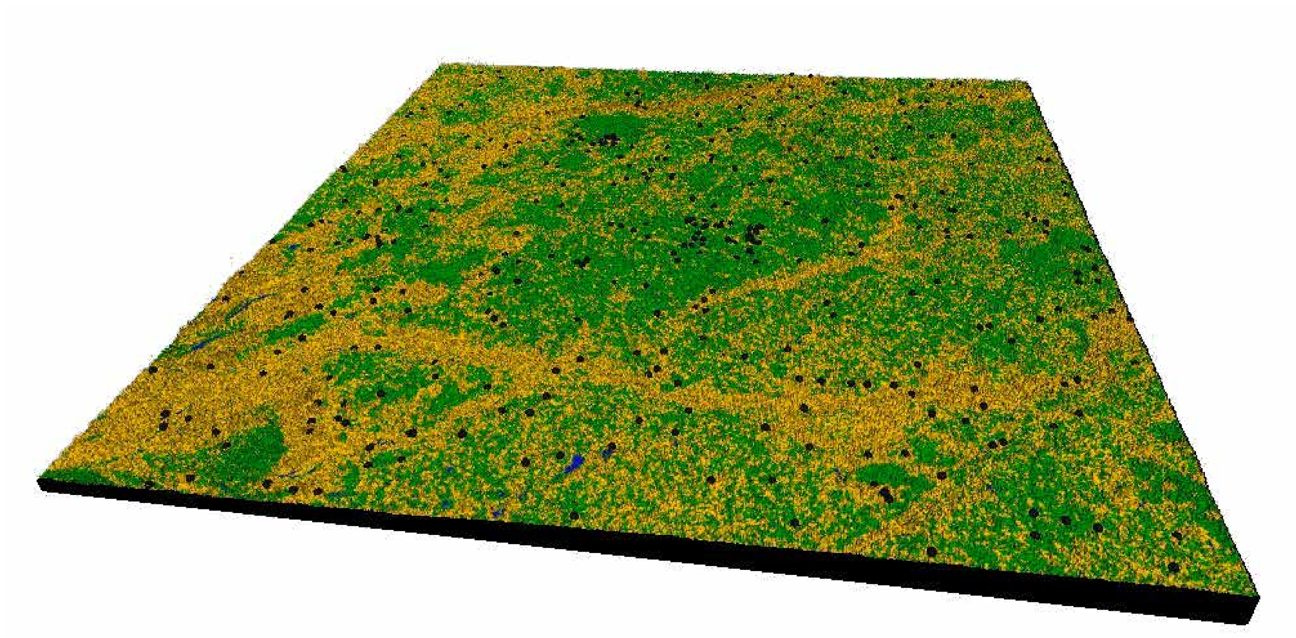
peterwittek@acm.org, xavier.rubio@bsc.es

UNIVERSITY OF BORÅS
SCIENCE FOR THE PROFESSIONS

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Key points

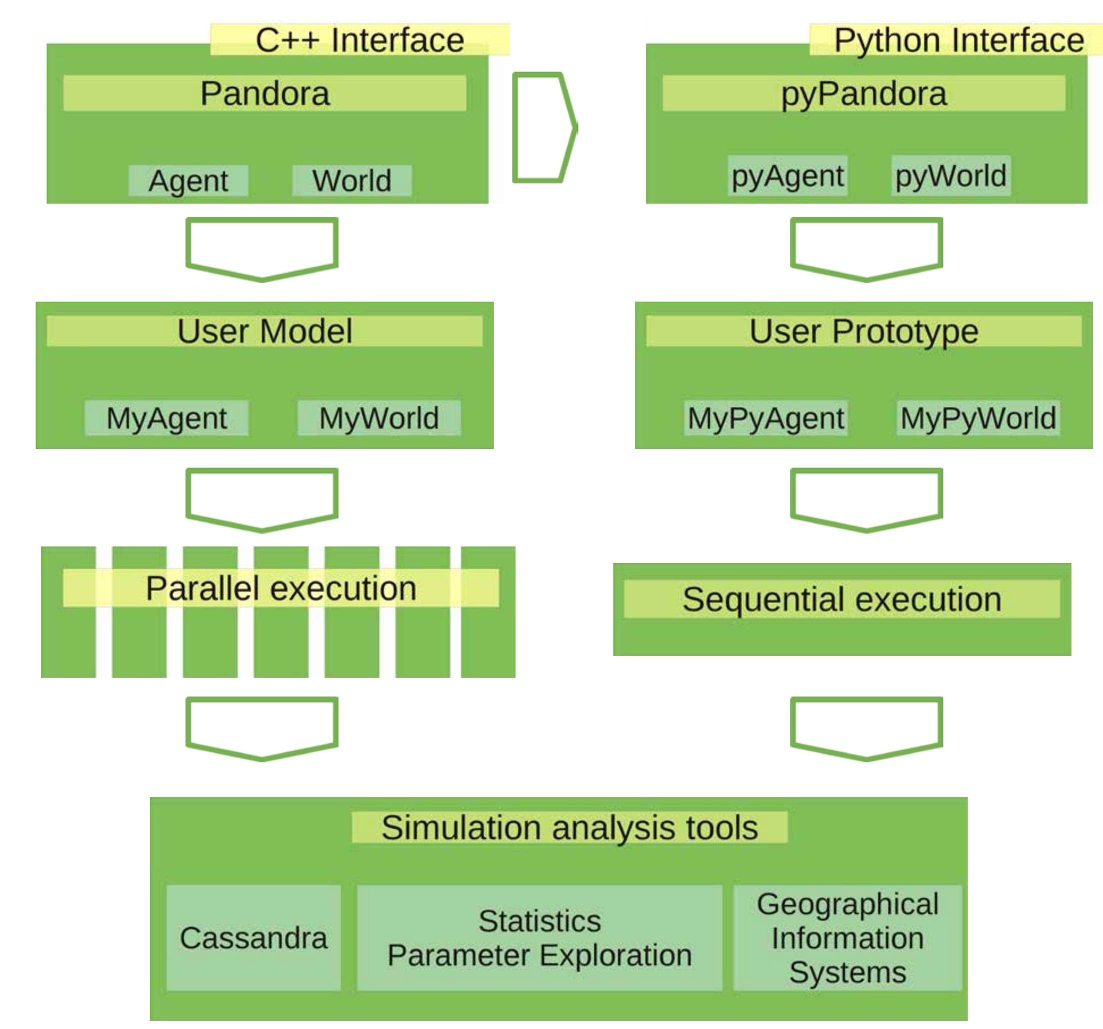
- Extending Pandora, a library for agent-based modeling for social scientists;
- Critical region of code accelerated by up to 7x, including memory transfer;
- Relying on Thrust alone;
- Considerable rework of data structures was necessary.



Agents are modeled over real terrain, such as the arid planes of Gujarat

High-Performance Agent-Based Modeling

Recent years have seen a growing number of projects in Humanities and Social Sciences that use computer simulation as their main research tool. An agent-based model (ABM) defines the behavior of any entity of a system that involves decision-making processes known as agents. These agents interact within a controlled environment that can be a real landscape with **geographical features** like vegetation and transport systems. The generation of emergent properties that arose from the definition of individual agents can include both quantitative and qualitative concepts, combining behavior aspects and data. Thus, the explanation provided by an ABM is closer to **how knowledge is acquired in social sciences**. Only **high-performance computing** resources are capable of dealing with large simulation scenarios containing agents with artificial intelligence and high computing costs. Distributing the workload of an ABM execution is a complicated task as the system is intrinsically communication intensive between the various components of the simulation. Agents need to gather knowledge from their environment, as well as from other agents in order to execute their decision-making processes. Once this phase is completed, there is a possibility that the agents will modify the environment and so will the other agents as well. These mechanics are translated to sharing several layers of environmental data and agents. **Pandora** is a novel open-source framework designed to accomplish a simulation environment for social scientists [2]. It provides a C++ and Python environment that splits the workload of a simulation across computer nodes using MPI, and also uses multiple cores relying on OpenMP.



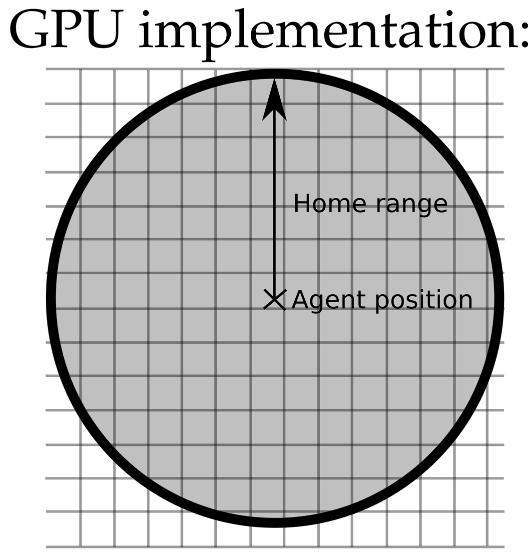
Implementation workflow of an ABM using Pandora

GPU Acceleration

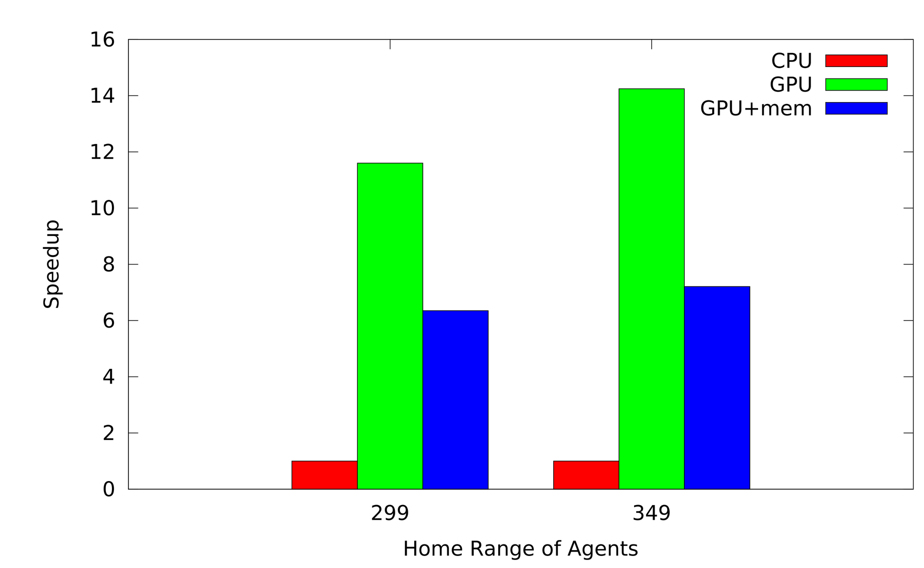
Inspired by [1], we identified the critical performance bottleneck in the simulations. This was the update of knowledge of the individual agents; a part that is parallelized by OpenMP in the original implementation. This part of the algorithm has completely irregular data access patterns.

Each agent has a copy of the simulated world, and is able to make decision based on local knowledge, a subset of the world. This local knowledge is accessible via a linear array describing essentially a circular region around the agent. The GPU implementation:

- Replaces the linear array containing irregular memory access patterns by a rectangular stencil.
- Keeps a copy of the simulated world in the device memory for each agent.
- Does a 2D reduction based on the rectangular stencil from the agents copy of the world in device memory.
- Reduction is achieved via Thrust iterators, custom kernels were not written.



The speed of the critical region was benchmarked on a Tesla C2050 GPU against an Intel Xeon E5620 CPU.



Speedup of critical region on a single GPU

Future Work

Boosting the speed of the critical region is just one step towards a more thorough acceleration. Ongoing work includes transferring the entire agent model to the GPU to further reduce latency and improve memory access patterns. Ideally all agents could update their knowledge on the GPU simultaneously.

References

[1] B. Aaby, K. Perumalla, and S. Seal. Efficient simulation of agent-based models on multi-GPU and multi-core clusters. In *Proceedings of SIMUTOOLs-10*, 2010.

[2] P. Wittek and X. Rubio-Campillo. Scalable agent-based modelling with cloud HPC resources for social simulations. In *Proceedings of CloudCom-12*, pages 355–362, 2012.

Funding

The second author is supported by the SimulPast Project (CSD2010-00034), funded by the CONSOLIDER-INGENIO2010 program of the Ministry of Science and Innovation – Spain. This work is also supported by AWS in Education Machine Learning Grant award.