Data-Driven Crowd Simulation

Hugo Perez, Jorge Ramirez, Isaac Rudomin

Barcelona Supercomputing Center, Barcelona, Spain

hugo.perez@bsc.es, jorge.ramirez@bsc.es, isaac.rudomin@bsc.es

***Keywords*— HPC, CGI, Big Data, Machine Learning**

Introduction

Our objective is to simulate entire cities with the most realistic possible scenario. This kind of systems require a lot of processing power, therefore we use hybrid computer clusters with graphics cards (GPUs). GPUs allow us to accelerate calculations and visualization.

In order to fully exploit the equipment we overlap the visualization with the calculation, in addition we apply different techniques of parallelism up to four levels, these techniques allow us to execute the simulation in several nodes of the cluster making use of all available resources (CPUs and GPUs), in addition the simulation can scale to as many nodes as necessary and available in the cluster without modifying the source code. We combine different parallel programming models like MPI, OmPss and CUDA for communication and synchronization between and within the cluster nodes [1].

Some of the applications of this kind of systems are: training in emergency situations like earthquakes, hurricanes and epidemics. Planning and urban logistics in special events such as concerts, protest march, parades or daily events in large cities such as vehicular traffic. In the area of entertainment are used in video games and movies, among many others.

Development

The basis of our project is the High Performance Computing HPC, on this basis we have developed algorithms of machine learning that allow us to learn behaviors based on real data (Big Data), for example GeoLife GPS trajectory dataset released by Microsoft [2], with data of 182 users, 17,621 trajectories of ca. 1.2 million Km. and 48,000+ hours, or the history of trajectories of New York taxis during 6 years or the data that will be released by Uber soon about the movement of its passengers in different cities.

To simulate realistic urban environments we use sources such as OpenStreetMap OSM and Natural Hearth [3], [4] from them we generate 3D geometry of buildings and get data of streets, parks, points of interest, etc. To complement the data of the environment, we use Open Data, these data are published by the public administration of each city, describing aspects such as the city's infrastructure, pollution, population, among others, however, they are not standardized, therefore, each city publishes different data with its own format and structure.

The combination of different data sources such as the ones just mentioned involves a process of cleaning and extracting data before they can be useful. For this task we use Hadoop, Pig, Spark, among others.

The system can be accessed from a web browser using different types of devices such as desktop computers, tablets, mobiles, etc. For this purpose we use Websockets to transfer the final image that combines the scene with the simulation (character's motion). we receive from the client the camera position, since our system is interactive, allowing to the user manage the camera. Figure 1 shows the system architecture.

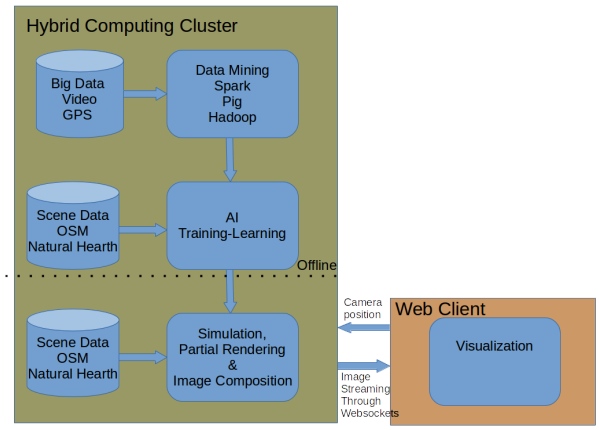


Fig. 1 System Architecture

Every new generation of embedded systems like Jetson TX2 of NVidia or Myriad 2 of Movidius-Intel has more computing resources, make them capable of compute and render part of the simulation in the client side, making more efficient procedures and communication, this concept is called Fog Computing [5]. If the client device has enough computing resources we can render the composition on the client side. For this task we use Tangram which is based on the use of OpenGL / WebGL that queries the OSM data, in the past we have also used Cesium with good results, however, to render 3D buildings, Tangram has more functionality, at the moment we were working on this task, the Cesium module in charge of rendering 3D buildings "3D Tiles" was not yet released, however, in the future will be reviewed again as it is an excellent tool.

Conclusion and Future Enhancement

The research areas covered by the project have made significant progress in recent years: HPC, CGI, Machine

Learning and Big Data.

In the area of high-performance computing and visualization, technology is constantly evolving, this implies the adaptation of the system to new computing architectures and the corresponding programming models that are used to manage it, in order to exploit resources to the maximum. As future work we plan to explore the use of OmpSs Cluster that would allow us to perform the communication within the cluster in a simpler and more efficient way, Vulkan that is the successor of OpenGL handles multithreaded rendering, so it may have a better interaction with OmpSs that will allow us to launch several threads with visualization tasks within a GPU. We also will try WebGL2 to perform visualization tasks on the client device.

Machine learning can be applied to many different tasks, for example: developing varied characters combining features of a dataset, animating characters realistically with movements learned from video data, among others. We will explore different possibilities.

To improve the interface we are developing a version in virtual reality that will allow the user to have a total immersion in the simulation and interact in a natural way using devices such as Oculus Rift and Leapmotion.

We will test different embedded systems that improve the simulation and allow greater functionality.

A. ACKNOWLEDGEMENT

This research was partially supported by: CONACyT doctoral fellowship 285730, BSC-CNS Severo Ochoa program (SEV-2011-00067), CUDA Center of Excellence at BSC, the Spanish Ministry of Economy and Competitivity under contract TIN2012-34557, and the SGR programme (2014-SGR-1051) of the Catalan Government.

.

References

[1] Hugo Pérez, Benjamín Hernández, Isaac Rudomin, Eduard Ayguade. Scaling Crowd Simulations in a GPU Accelerated Cluster. High Performance Computer Applications: 6th International Conference, Mexico City, Mexico, Revised Selected Papers. Springer 2016.

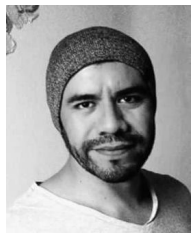
[2] Zheng, Y., Xie, X., Ma, W.-Y.: GeoLife: A Collaborative Social Networking Service among User, location and trajectory. IEEE Data Engineering Bulletin. 33, (2010).

[3] Menzel, J. Robert and Middelberg, Sven and Trettner, Philip and Jonas, Bastian and Kobbelt, Leif. City Reconstruction and Visualization from Public Data Sources. Eurographics Workshop on Urban Data Modelling and Visualisation 2016.

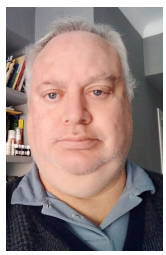
[4] Ari Seff, Jianxiong Xiao. Learning from Maps: Visual Common Sense for Autonomous Driving. Computer Vision and Pattern Recognition 2016.

[5] Shanhe Yi, A Cheng Li, A Qun Li. "A Survey of Fog Computing: Concepts, Applications and Issues". Proceedings of the 2015 Workshop on Mobile Big Data.

Author biography



Hugo Perez got his B.S. degree in Electronic Engineering from National University in Mexico (UNAM). He got his M.Sc. degree in Computers Architecture, Networks and Systems from Universitat Politècnica de Catalunya BarcelonaTech. Currently he is working in the eXtreme Computing Group at the Barcelona Supercomputing Center as PhD student. His research project entitled "Crowd Simulation and Visualization" which aims to represent the most realistic possible scenarios in a city, these kind of systems allow: urban planning, simulating disasters, simulate epidemics, among other applications. The project combines differents areas research such as: High Performance Computing, Parallel Programming Models, Computer Graphics, Big Data, Machine Learning, Virtual Reality, between others.



Isaac Rudomin is a senior researcher at the Barcelona Supercomputer Center, which he joined in 2012. His focus is on crowd rendering and simulation including generating, simulating, animating, and rendering large and varied crowds using GPUs in consumer-level machines and in HPC heterogeneous clusters with GPUs. Previously, Isaac was on the faculty at Tecnologico de Monterrey Campus Estado de Mexico (from 1990 to 2012). He finished his Ph.D. at the University of Pennsylvania under Norman Badler on the topic of cloth modeling.