Crowd Simulation and Visualization

Hugo Perez, Isaac Rudomin, Eduard A Ayguade, Benjamin A Hernandez, Javier A Espinosa-Oviedo, Genoveva Vargas-Solar

Crowd Simulation and Visualization

Hugo Perez[1,2], Isaac Rudomin[1], Eduard Ayguade[1,2], Benjamin Hernandez[3],
Javier A. Espinosa-Oviedo[1,4], Genoveva Vargas-Solar[4]
[1]*Barcelona Supercomputing Center, Barcelona, Spain*
[2]*Universitat Politècnica de Catalunya BarcelonaTech, Barcelona, Spain*
[3]*Oak Ridge National Laboratory, Oak Ridge, TN, USA*
[4]French National Council on Scientific Research (CNRS), LIG-LAFMIA,
BP 72 38402 Saint-Martin d'Hères, France
hugo.perez@bsc.es

Abstract- *This paper presents a methodology to simulate and visualize crowds. Our goal is to represent the most realistic possible scenarios in a city. Due to the high demand of resources a GPU Cluster is used. We use real data from which we identify the behavior of the masses applying statistical and artificial intelligence techniques. In order to take advantge of the processing power of the GPU cluster we use the following programming models during the characters simulation: MPI, OmpSs and CUDA. We developed different visualization schemes: a) In situ, b) Streaming, c) Web. The web scheme is the most flexible, allowing to interact in real time with the simulation through a web browser. For this scheme we use WebGL and Cesium.*

Keywords: Parallel Programming Models, HPC, GPU, AI, Computer Graphics.

I. INTRODUCTION

In this era of big cities one is confronted by emergency situations caused by traffic, natural disasters or special events such as concerts, sports events and protests which require the intervention of qualified personnel in order to generate an orderly and safe urban experience, and save lives. Data management and visualization can support real-time observation, understanding the behavior of this exodus and develop security strategies. We believe that technology and particularly data management and visualization can provide tools that can help to control this complex situation: real-time observation, understanding of the behavior of the people through on-line and post-mortem analytics, real-time decision making and recommendation are some of the activities that can be supported.

Our data processing and simulation are computationally expensive and critical thus we rely on HPC infrastructure with hybrid architecture (CPUs + GPUs) to produce an efficient solution. Heterogeneous clusters provide various levels of parallelism that need to be exploited successfully to harness their computational power totally. Our particular endeavors have been focused on the design, development and analysis of crowd simulations in these systems. Our first efforts [1] combined CUDA and MPI models for in-situ crowd simulation and rendering on GPU clusters and recently [2] we have adopted OmpSs. We proposed a task-based algorithm that allows the full use of all levels of parallelism available in a heterogeneous node.

This paper presents the methodology to simulate and visualize crowds. We address the challenge of visualize on real time and predict the behaviour of individuals and groups moving and evolving within real environments that uses information harvested from different sources.

II. DEVELOPMENT

We use the GeoLife GPS trajectory dataset [3] with data of 182 users, 17,621 trajectories of *ca.* 1.2 million Km. and 48,000+ hours. *These* data is used to compute spatio-temporal people flows in real crowds to provide data driven on-line crowd simulation, enhanced with real places geometric data running on GPU and HPC. Since the data set for a given place and time is sparse, we used agent based microsimulation to complement the actual trajectories in the dataset by using all the trajectories in similar moments that are available in the dataset to derive the most probable trajectories for the simulated vehicles or pedestrians.

Figure 1 illustrates the general overview of our approach using the web visualization scheme (the details of the different schemes are described in [1]) that addresses three main problems: (i) data harvesting, (ii) crowd simulation and analytics and (iii) visualization. We use existing temporal geo located observations concerning individuals' trajectories.

We apply data analytics techniques (temporal and spatial reasoning) for computing trajectories and for identifying crowds, that is people grouped in a sufficient close spatial region that adopt a specific "behaviour" referring to four well known naïve crowd patterns: (i) casual crowd which is loosely organized and emerges spontaneously, people forming it have very little interaction at first and usually are not familiar with each other; (ii) conventional crowd results from more deliberate planning with norms that are defined and acted upon according to the situation; (iii) expressive crowd forms around an event that has an emotional appeal; and (iv) acting crowd members are actively and enthusiastically involved in doing something that is directly related to their goal.

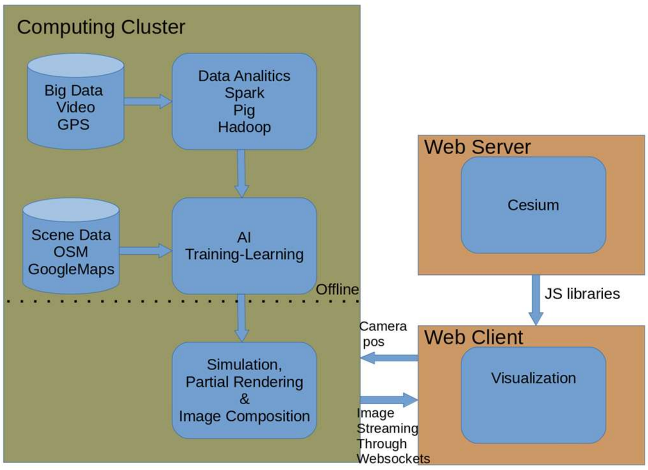

Fig. 1. General overview

Simulation and visualization is based on the work described in [2, 4]. However for clarity sake, we describe the basic charactiristics: Processing these tasks in parallel within a cluster, requires tiling and stencil computations. First, the navigation space (from now on it will be called the World) and information for all the agents is divided into zones. Each zone is assigned to a node which in turn is divided into sub-zones. Then, the tasks performed in each sub-zone can be executed in parallel by either a CPU or GPU. Stencil computations are performed on an inter and intra-node basis. A step by step description of the algorithm is included:

Step 1: Navigation space is discretized using a grid; then the resultant grid is discretized into zones. Each node will compute each zone.
Step 2: Divide each zone into tiles (sub-zones). A CPU or GPU will compute each sub-zone. Each CPU or GPU stores their corresponding tile of the world.
Step 3: Set up the communication topology between zones and sub-zones.
Step 4: Exchange borders (stencils).
Step 4a: (Intra-node) Exchange the occupied cells in the borders between sub-zones
Step 4b: (Internode) Exchange the occupied cells in the borders between zones
Step 5: Update position for each agent in parallel
Step 6: Agents' Information Exchange.
Step 6a: (Intra-node) Exchange agents' information that crossed a border and moved to another sub-zone
Step 6b: (Internode) Exchange agents' information that crossed a border and moved to another zone .

Simulated vehicles and pedestrians will use heatmaps derived from the dataset in order to follow the most popular routes. The details of these modifications are not within the scope of this paper and will be published later.

In general, the principle consists in mapping human perception of the space stemming from cameras and expressed in geographical coordinates (latitude, longitude), for example, into pixels. For instance, as shown in figure 2 "*give me the GPS coordinates of the users evolving in Beijing ordered by time*". Once this query has been evaluated by the appropriate data processing infrastructure (in the work presented here PigLatin [5] execution environment), results are transformed into the appropriate format. Textures and maps are retrieved in order to create the 3D space where individuals' movements will be visualized (simulated) according to the observed information.´

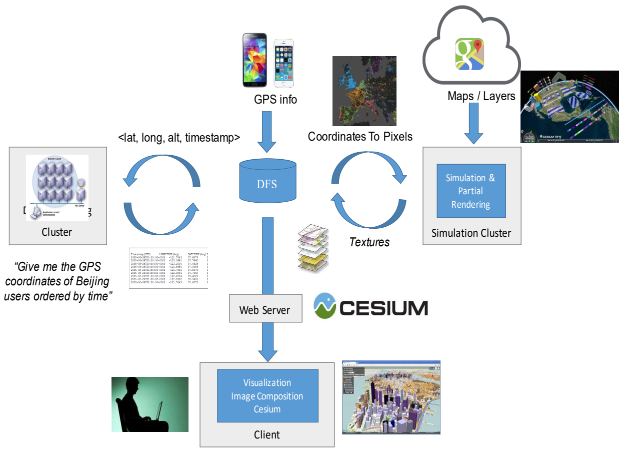

Fig. 2. Visualization process of individuals movement within urban 3D spaces.

III. CONCLUSIONS AND FUTURE WORK

Crowd sourced location data is used to compute spatio-temporal people flows in real crowds. We combine both to provide data driven on-line crowd simulation, enhanced with real places geometric data. This paper presented the general approach for simulating crowd behaviour and thereby supporting individuals' and crowd behaviour in public spaces. The main contribution is combining location based data collections previously harvested together with online geo-tagged data for visualizing crowds at different levels of precision and detail, according to access control and privacy constraints. Our data processing and simulation process are computationally expensive and critical thus we rely on HPC and GPU infrastructures for producing an efficient solution.

As future work we will use automatic learning techniques (deep learning) so that the system can “react” to events and simulate a synthetic behaviour of the crowd.

ACKNOWLEDGMENT

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] Hernandez, B., Perez, H., Isaac, R., Ruiz, S., DeGyves, O., Toledo, L.: *Simulating and visualizing real-time crowds on gpu clusters*. *Computacion y Sistemas* 18 (2014) 651–664

[2] Perez, H., Hernandez, B., Rudomin, I., Ayguade, E.: *Task-based crowd simulation for heterogeneous architectures*. In: *Innovative Research and Applications in Next-Generation High Performance Computing*. Qusay F. Hassan (Mansoura University, Egypt). IGI Global 2016.

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

[4] Perez, H., Hernandez, B., Rudomin, I., Ayguade, E.: Scaling Crowd Simulations in a GPU Accelerated Cluster. Springer 2016 (In progress)

[5] Olston, C., Reed, B., Srivastava, U., Kumar, R., Tomkins, A.: Pig latin: *a not-so-foreign language for data processing*. In: Proc. of the 2008 ACM SIGMOD Int. Conference on Management of data (SIGMOD’08). ACM Press, Vancouver, Canada (2008).