Visualization for Extremely Large-Scale Scientific Computing

Dr. Abel Coll
Summary

1. Introduction
2. Motivation
3. Project objectives
4. Current status of the project
5. Next steps
6. Conclusions
Summary

1. Introduction
2. Motivation
3. Project objectives
4. Current status of the project
5. Next steps
6. Conclusions
**VELaSSCo basic information**

- **Proposal number:** 619439
- **Project Officer:** Mr. Pierre-Paul SONDAG
- **Name of the coordinating person:** Dr. Abel COLL

<table>
<thead>
<tr>
<th>Participant organisation name</th>
<th>Short name</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>International Center for Numerical Methods in Engineering</td>
<td>CIMNE</td>
<td>ES</td>
</tr>
<tr>
<td>School of Engineering. The University of Edinburgh</td>
<td>UNEDIN</td>
<td>UK</td>
</tr>
<tr>
<td>STIFTELEN SINTEF</td>
<td>SINTEF</td>
<td>NO</td>
</tr>
<tr>
<td>Institut national de recherche en informatique et en automatique</td>
<td>INRIA</td>
<td>FR</td>
</tr>
<tr>
<td>Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V. Fraunhofer-Institut für Graphische Datenverarbeitung</td>
<td>FRAUNHOFER</td>
<td>DE</td>
</tr>
<tr>
<td>Jotne EPM Technology</td>
<td>JOTNE</td>
<td>NO</td>
</tr>
<tr>
<td>Atos Spain S.A.</td>
<td>ATOS</td>
<td>ES</td>
</tr>
</tbody>
</table>

**3 years project** (2014 –2016)

393 persons x month

The total costs of VELaSSCo are **4,441,603 €**, and total funding is **3,294,000 €** (original 3,294,425 €)
Introduction

The Vision of VELaSSCo is to provide new visual analysis methods for large-scale simulations serving the petabyte era and preparing the exabyte era.

It does this by adopting Big Data tools and architectures for the engineering and scientific community and by leveraging new ways of in-situ processing for data analytics and hardware accelerated interactive visualization.
VELaSSCo consortium as a whole

Big Data Infrastructure
- Big Data Issues
  - HPC and Big Data, Handling, formatting, storage
  - Data access, extraction, reduction

Data Analytics
- Platforms
  - FEM Models
  - DEM Models
  - LB Models

Visualization Expertise
- End-user testing
  - Usability verification
  - Reactivity

End-users / Beneficiaries

Spain
- ATOS
- CIMNE

United Kingdom
- UNEDIN

Norway
- SINTEF
- JOTNE

France
- INRIA

Germany
- FRAUNHOFER

Industry

SMEs

Academia

May 2nd 2016

Pre-workshop CAxMan, CIMNE, Barcelona
Summary

1. Introduction
2. Motivation
3. Project objectives
4. Current status of the project
5. Next steps
6. Conclusions
Brief introduction to numerical simulations

VELaSSCo

Pre-processing

Calculation

Post-processing

Geometry description

Preparation of analysis data

Visualization of results

Pre and post-processor

Computer Analysis

May 2nd 2016

Pre-workshop CAxMan, CIMNE, Barcelona

Introduction
Current bottleneck

- Nowadays the huge amount of data provided by the solver in HPC cannot be stored in one single machine, so it is mandatory:
  - Distributed post-processing
  - Distributed visualization
Mesh: the main data

- Mesh static in time
Mesh: the main data

- Mesh evolving in time
Calculation: solver

- Nowadays we can run very large simulations using HPC infrastructures
- Clusters with distributed computation nodes
- Data generated in the simulation:
  - Mesh
  - Results on mesh
  for each time step! (several thousands)
(several Gb per time step, which involves several hundreds of Tb for the whole simulation)
Calculation: solver

- Domain partitioning for distributed calculation
Postprocessing

- Postprocessing operations (data analytics) and visualization of results

Extract skin of the car

Plane cuts
Interpolate results onto the cut

Apply color texture based on the results

Visualize the contour fill

Animate for different time steps
Postprocessing

- Postprocessing operations (data analytics) and visualization of results

- Extract skin of the car
- Plane cuts
- Interpolate results onto the cut
- Apply color texture based on the results
- Visualize the contour fill
- Animate for different time steps
Postprocessing

- Isosurfaces

Extract skin of the model

Create isosurface of pressure equal to 0

Animate for each time step
Postprocessing

- Isosurfaces

Extract skin of the model
Create isosurface of pressure equal to 0
Animate for each time step
Compute shadows projected onto the iso-surface
Postprocessing

- Isosurfaces

Extract skin of the model

Create isosurface of pressure equal to 0

Animate for each time step

Compute shadows projected onto the iso-surface

Build up a stereoscopic visualization
Current bottleneck

• Nowadays the **huge amount of data** provided by the solver in HPC **cannot be stored** in one single machine, so it is mandatory:
  – Distributed post-processing
  – Distributed visualization
Summary

1. Introduction
2. Motivation
3. Project objectives
4. Current status of the project
5. Next steps
6. Conclusions
Objective of VELaSSCo

• The main **objective of VELaSSCo** project is to build the VELaSSCo Platform, a system that performs distributed post-processing operations and visualization of very large simulations.

• To address this objective, VELaSSCo brings together **Simulation** and **Big Data**.
VELaSSCo Platform

**Objectives**

- **Visualization**
- **Client**
- **Data**
- **Data**
- **Data**
- **Data**
- **Data**

**Postprocessing operations**
(Data Analytics)

- **Model reduction**
- **& GPU friendly**

**Visualization client**

**GiD**

**iFX**

Pre-workshop CAxMan, CIMNE, Barcelona
WPs distribution

VELaSSCo

WP7: VELaSSCo Project Management

WP6: VELaSSCo Dissemination and Exploitation

WP2: Data formatting, handling and storage

WP3: Data analytics for engineering simulations

WP4: High performance visualization/ scalable visual analysis

WP5: Usability and Effectiveness Evaluation

WP1: Steering on target users and general VELaSSCo’s requirements
Some numbers…

• Data sizes we are currently working with at first prototype level

**Fluidized bed DEM example:**
- 12,000 particles / time-step
- 3,000 contacts (p2p and p2w) / time-step
- 40,000 time-steps
- Mass, volume and velocity vector, and force vector for contacts

**Telescope FEM example:**
- 4 M nodes, 24 M tetrahedra
- 19 time steps
- Partition index, pressure and velocity (5 doubles)
Some numbers…

- Data sizes we already have for the **final prototype** of VELaSSCo Platform
  - Barcelona model:
    - \(~100\text{M Tetrahedra}\) (4 m resolution), \(~384\) subdomains
    - \(~340\) time-steps
    - \(227\) GB disk space

Simulations done in the framework of Numexas EU funded project
Barcelona model

Mesh of 6,000,000 elements (8 m resolution)
Barcelona model

400M elements correspond to a resolution of 4m.
Barcelona model

400M elements correspond to a resolution of 4m.
Barcelona model

400M elements correspond to a resolution of 4m.
Some numbers…

• Data sizes we are expecting in the next future
  • Barcelona model: ~1300M Tetrahedra (2 m resolution)
Summary

1. Introduction
2. Motivation
3. Project objectives
4. Current status of the project
5. Next steps
6. Conclusions
# Project progress

## Current status of the project

### WP 1: Steering on target users and general VELaSSCo's requirements
- Task 1.1: Steering on target users, end-user requirements
- Task 1.2: System architecture and technical requirements
- Task 1.3: Usability criteria and tests, questionnaires definition

### WP 2: Data formatting, handling and storage
- Task 2.1: Management of huge and complex existing datasets available
- Task 2.2: Specification of the Big Data architecture
- Task 2.3: Data storage design for the HPC cloud infrastructure for engineering analysis
- Task 2.4: Move the source data to the HPC cloud infrastructure via a petabyte size data
- Task 2.5: Data conversion algorithms/tools to convert the data to engineering friendly
- Task 2.6: Selection and adaptation of scalable data formats/access for interaction of

### WP 3: Data analytics for engineering simulations
- Task 3.1: Implementation of the infrastructure to access the big data and the interface
- Task 3.2: Devel. of algo. in the 'batch layer': multi-resolution, coarsening, coordinates
- Task 3.3: Devel. of distributed db system that efficiently executes simple users' query
- Task 3.4: Data analytics for particle-based simulations: define and implement DEM solver
- Task 3.5: Design and implementation of complex queries for FEM, DEM and DEM continu
- Task 3.6: Steering simulation and quality control of output data: validity and completeness
- Task 3.7: Detection/identification of specific data patterns and other novel techniques

### WP 4: Visualisation and real time data interrogation
- Task 4.1: Scalable GPU-based interactive visualization methods
- Task 4.2: GPU-driven representations of result sets
- Task 4.3: Interaction metaphors to express queries on visualization features
- Task 4.4: GPU-enabled interactive visual analysis

### WP 5: Usability and Effectiveness Evaluation
- Task 5.1: Evaluation methodology
- Task 5.2: Verification of system architecture
- Task 5.3: Verification of algorithms and implementations
- Task 5.4: Effectiveness evaluation: reactivity, real-time data access/visualisation for e
- Task 5.5: Usability evaluation: testing human factors

### WP 6: VELaSSCo Dissemination and Exploitation
- Task 6.1: VELaSSCo Dissemination
- Task 6.2: VELaSSCo Exploitation Plan

### WP 7: VELaSSCo Project Management
- Task 7.1: Administrative project management
- Task 7.2: Scientific project management

---

*34*
User Panel

• We have created a user panel (around 50 members):
  – Collect requirements
  – Perform evaluations of the platform
  – Periodically updated about the status of the project

Join our user panel at www.velassco.eu.
Big Data architecture

- The Open Source:

  - VELaSSCo.Engine.Layer (YARN)
  - VELaSSCo.Data.Layer
  - HadoopAbstractFileSystem
  - HBase
  - Phoenix
  - Hive
  - FS

Results / data flow

Queries flow

QueryManager Thrift API

StorageModule Thrift API

Hbase & co. Thrift API

Current status of the project
Architecture (updated 21.10.2015)

- The Close Source version:
First prototype of VELaSSCo Platform
First prototype of VELaSSCo Platform
Summary

1. Introduction
2. Motivation
3. Project objectives
4. Current status of the project
5. Next steps
6. Conclusions
Next steps

• Final prototype to be ready by September 2016
• Evaluation event to be held on October/November 2016
• Integration as a product in some partners strategy
• Take advantage of some ‘opened doors’…
VELaSSCo Platform

HPC

Visualization client

Next steps
Summary

1. Introduction
2. Motivation
3. Project objectives
4. Current status of the project
5. Next steps
6. Conclusions
Conclusions

• We have reached the first prototype of VELaSSCo platform obtaining a prof of concept of it
• We are learning a lot from putting together BigData and Simulation worlds
• We have promising results in order to change the paradigm of postprocessing and visualization for numerical simulations
Thank you for your attention

Project coordinator – Dr. Abel Coll

www.velassco.eu