VELaSSCo

Towards the Petabyte era: post-processing and visualization for computational engineering: VELaSSCo

Miguel A. Pasenau, Javier Mora, Jorge Suit, Abel Coll,

ECCOMAS 2016, Crete, Greece
June 6th - 10th 2016

VELaSSCo

Visualization for Extremely Large-Scale Scientific Computing

Miguel A. Pasenau, Javier Mora, Jorge Suit, Abel Coll,

The University of Edinburgh School of Engineering

SINTEF

Inria informatics mathematics

Fraunhofer IGD

Atos

CIMNE

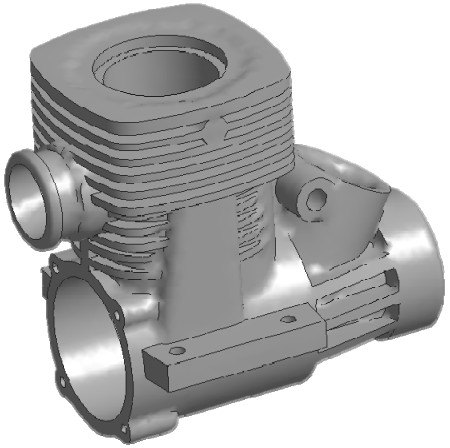
Jotne

SEVENTH FRAMEWORK PROGRAMME

ECCOMAS 2016, Crete, Greece

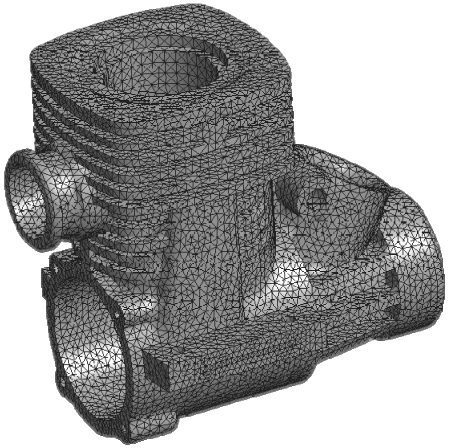
June 6th - 10th 2016

Motivation Simulation Pipeline



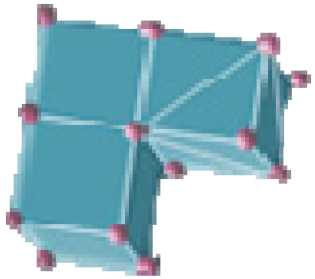
Modeling

- Not Clean Geometry
- Complex Models



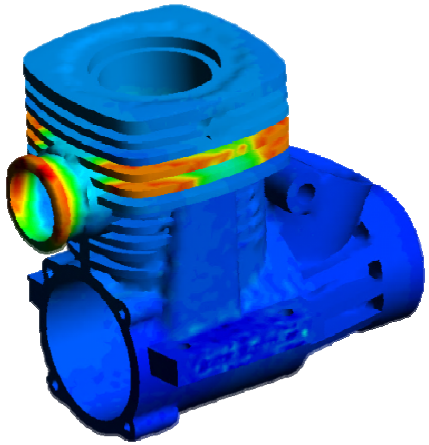
Meshing

- Robustness
- Not Scalable



Analysis

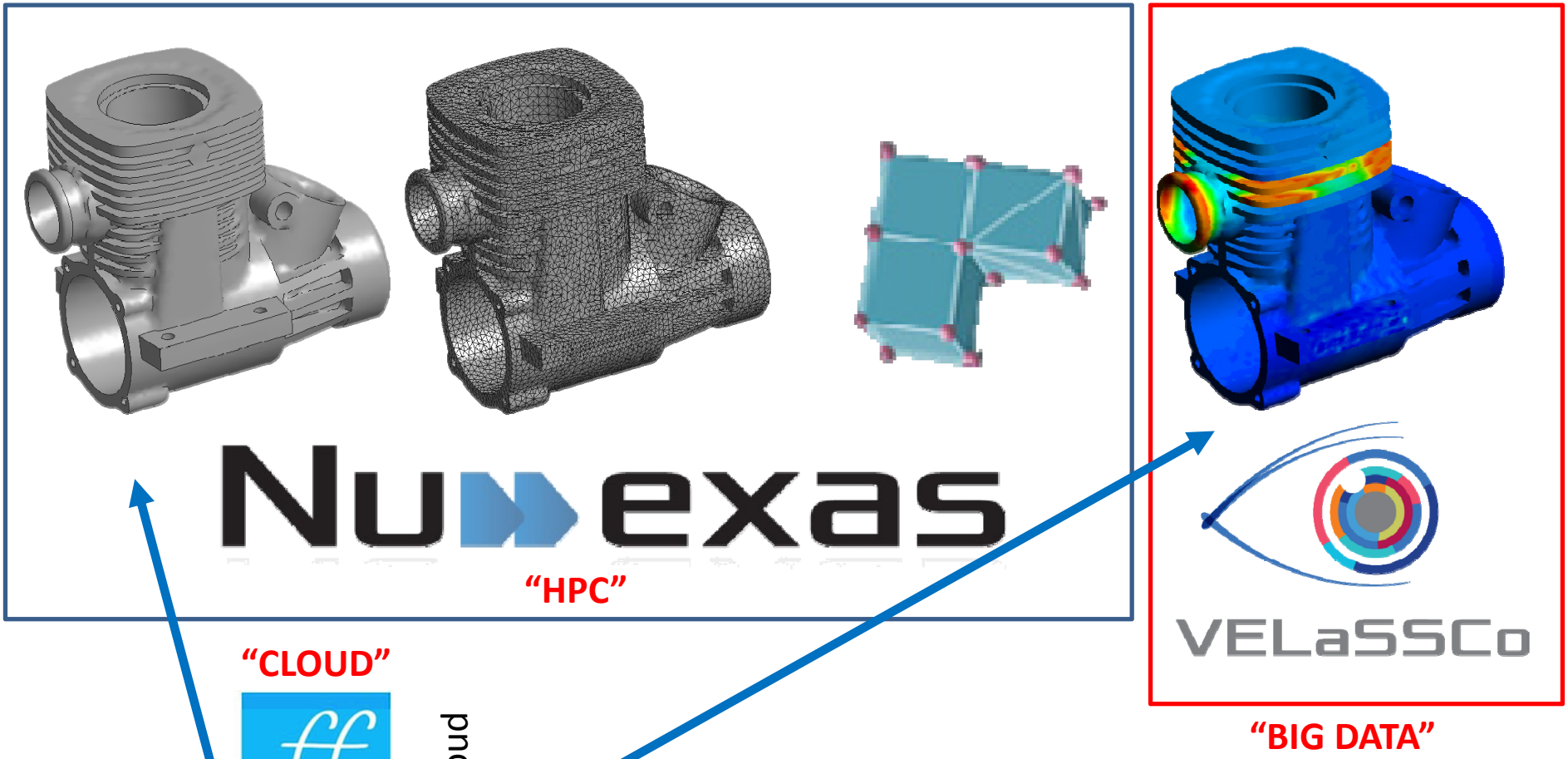
- IO
- Scalability
- Efficiency
- Complexity
- Heterogeneous Machines



Visualization

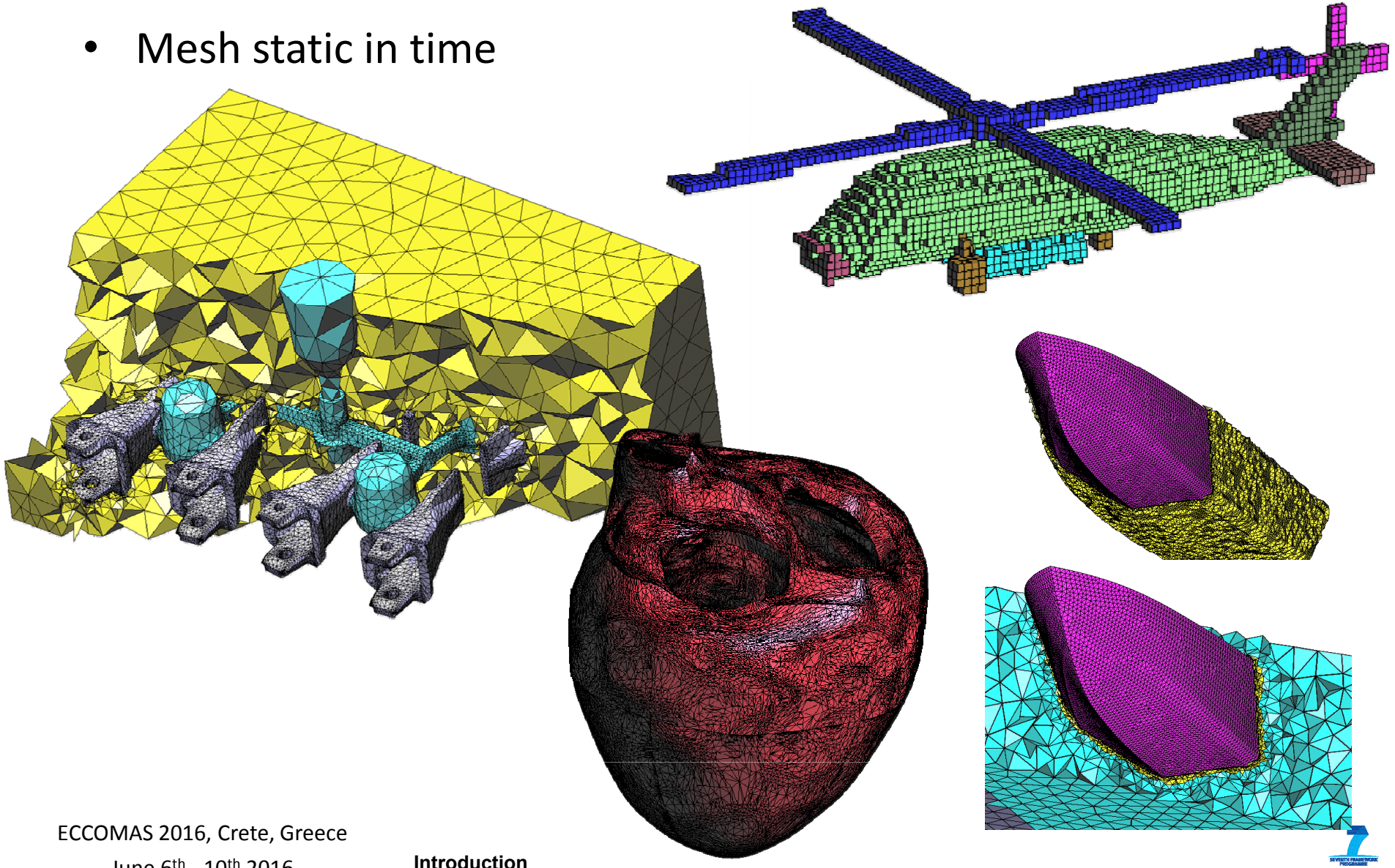
- Connection to the Servers
  - Internet
- Limited local resources
  - Small laptops, tablets, mobiles

Motivation Vision (HPC + Big Data + Cloud)



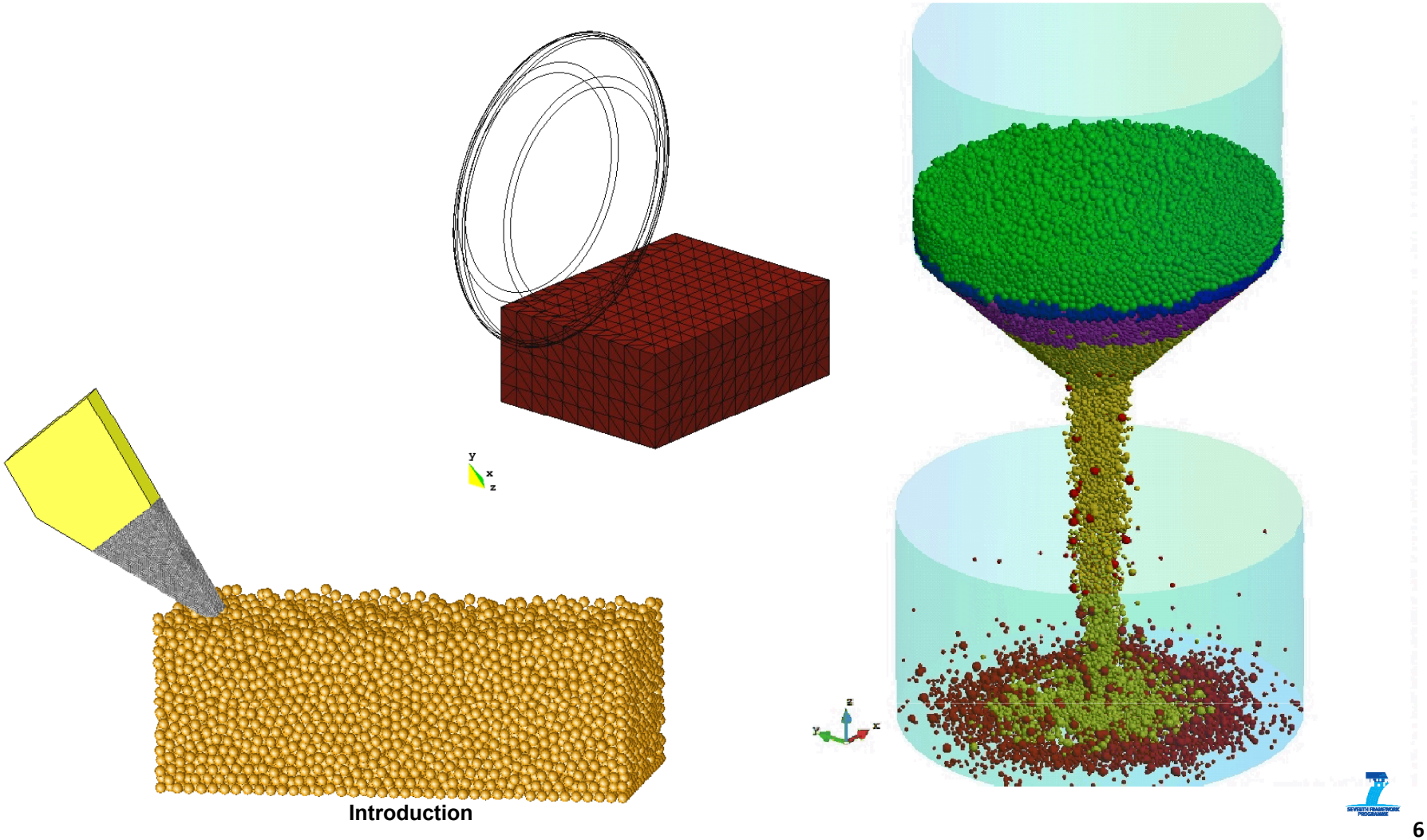
Simulation data

- Mesh static in time



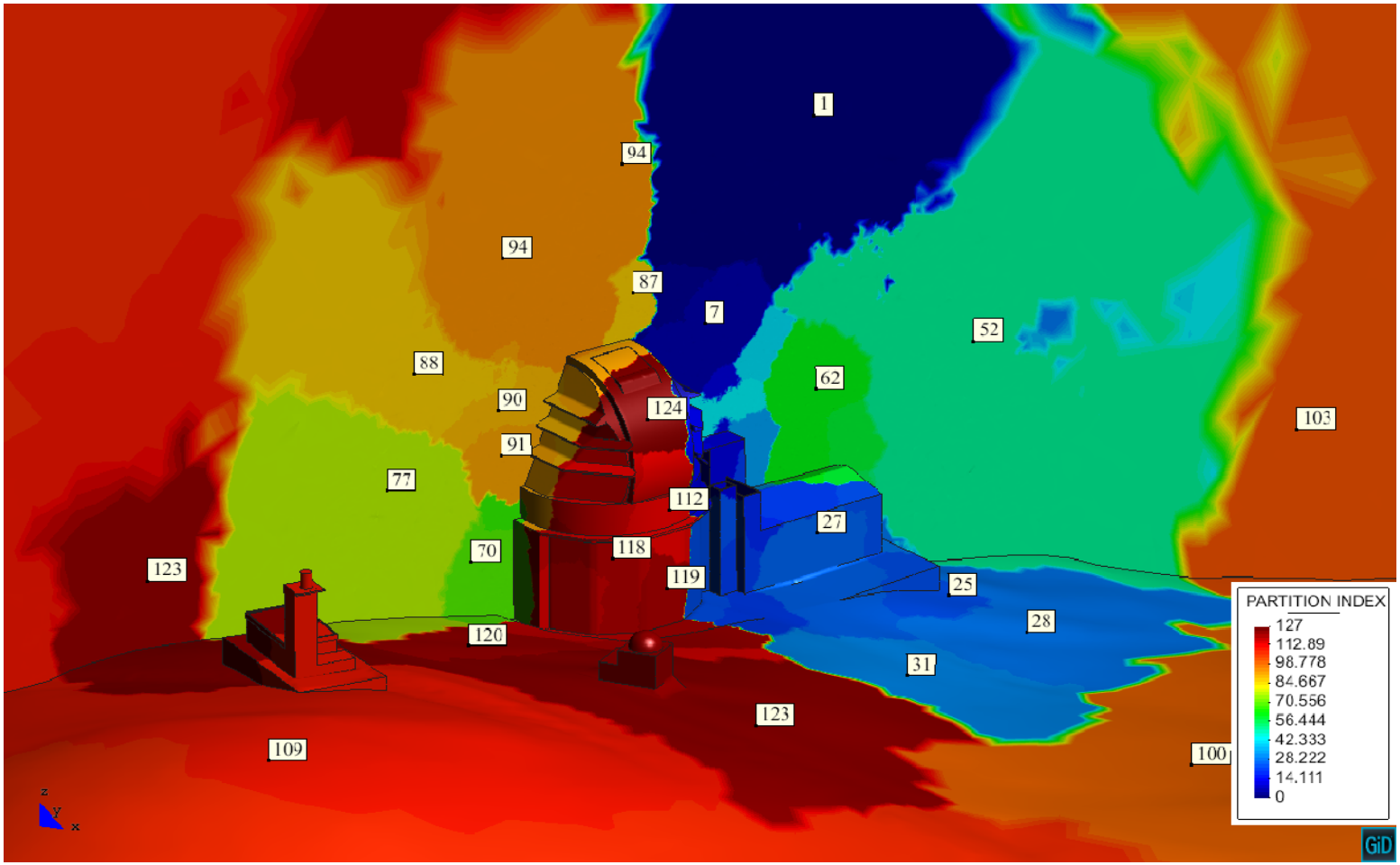
Simulation data

- Mesh evolving in time



Simulation data

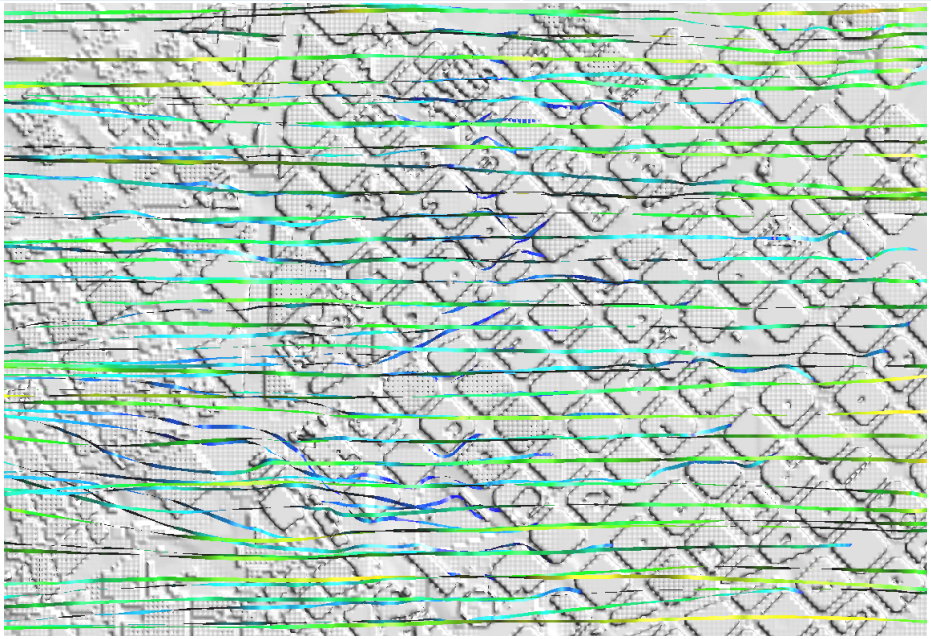
- Domain partitioning for distributed calculation



Simulation data sizes

- Barcelona domain = 64 x 64 x 1 Km

Resolution	Elements	Steps	TS size (2v+3s)	Partitions	Data size (i32,f32)	Data size (i32,f32)
16 m	20 M	5	0.3 GB	96	1.2 GB	2.3 GB
8 m	100 M	240	1.5 GB	384	230 GB	450 GB
4 m	340 M	5 (1000)	5.4 GB	3072	24 GB (2 TB)	47 GB (4 TB)
2 m	1,300 M	? 1000 ?	19 GB	?	? 8 TB ?	? 16 TB ?

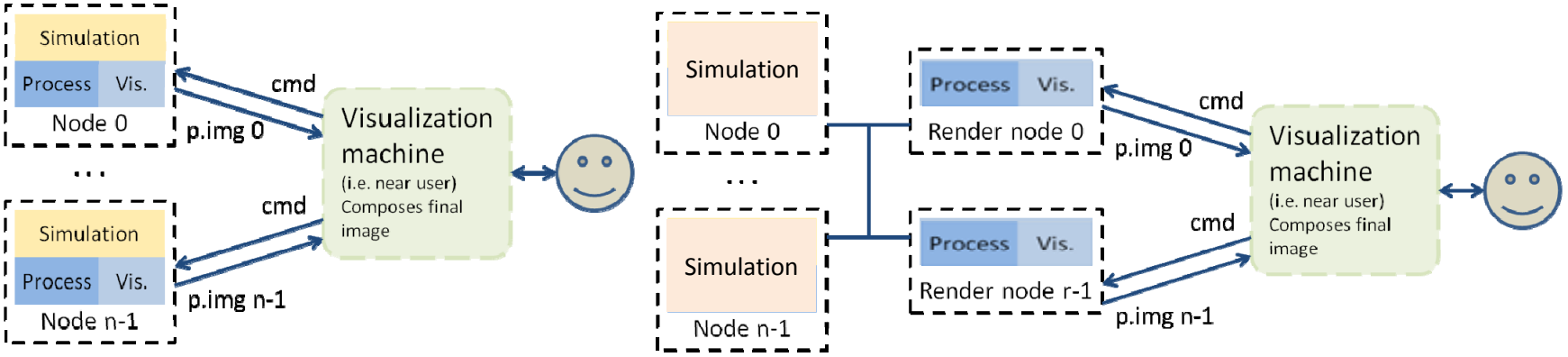


Post-processing options

- Locally:
  - download time
    - 100Melems + 240 TS, 230GB, 16h @ 4MB/s
    - 1.3Belems + 240 TS, ~2 TB, ~6d @ 4MB/s
  - Good computer (1.3Belems + vector = 68 GB RAM)

Post-processing alternatives

- Remotely: distributed rendering



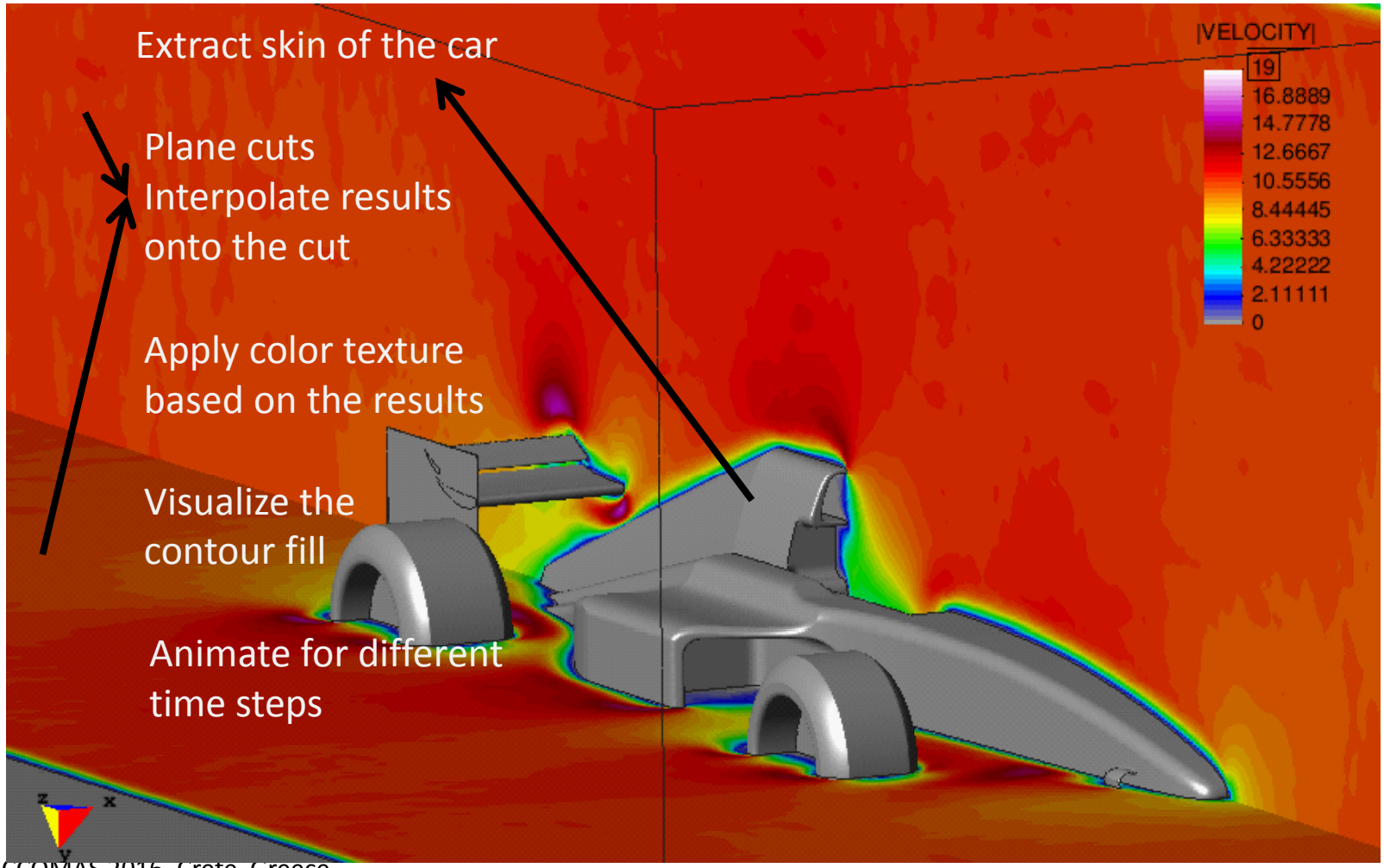
  - Good network connection (interaction depends on it)
  - Nodes with GPU, eventually dedicated nodes

Post-processing options

- In-situ post-processing:
  - 2 types:
    - In-situ image generation (simulation code generates images)
    - In-situ analysis (generates graphs, cuts or iso-surface meshes)
  - Actors placement:
    - Camera position, traveling,
    - Point sensors, cut-planes, iso-surfaces
    - At beginning or during the simulation
  - Monitors simulation
  - Reduced data storage
    - No raw data, only needed data
  - Distributed results, needs to be merged
  - No raw data → no post-process after simulation's end

Postprocessing

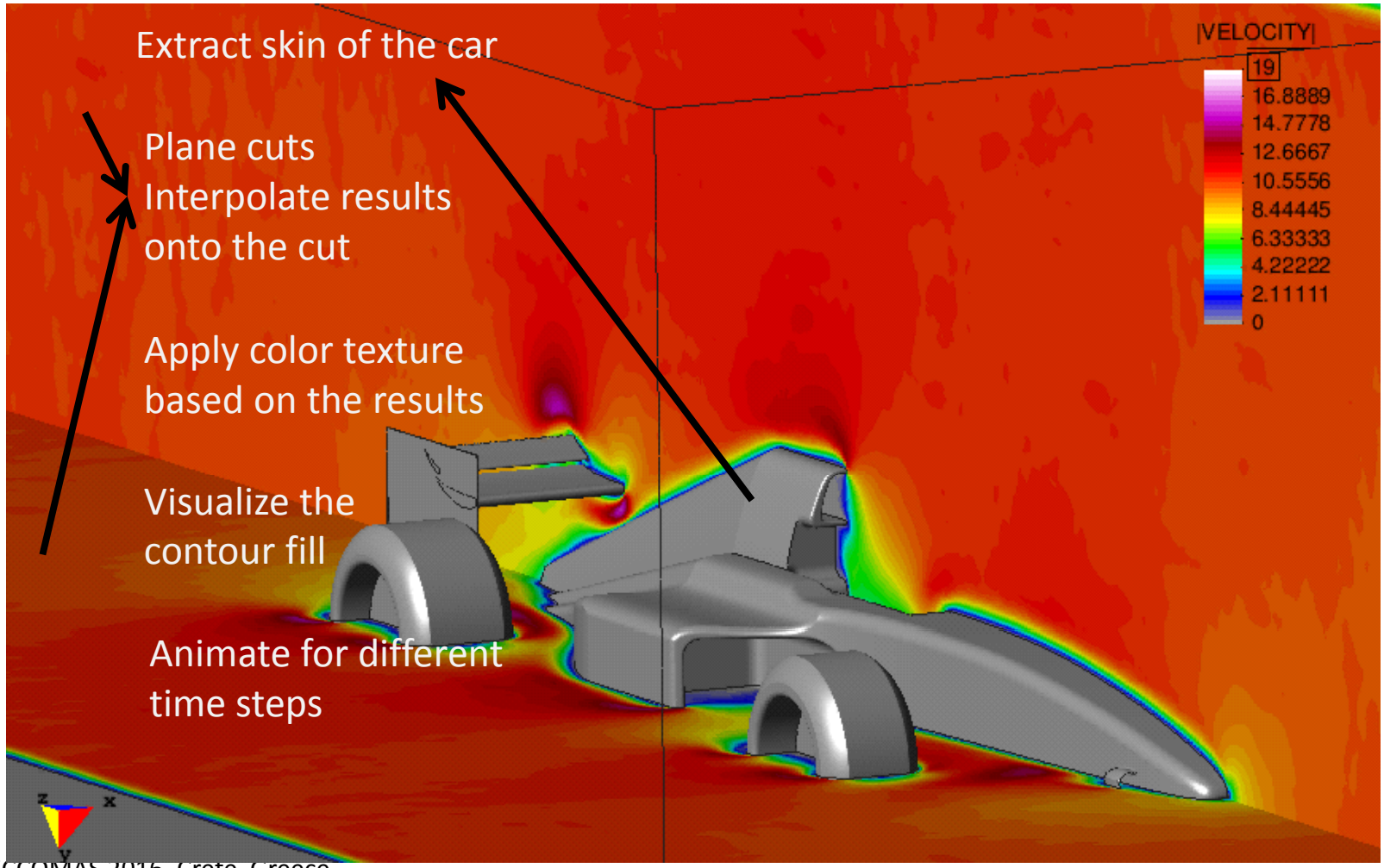
- Postprocessing operations (data analytics) and visualization of results



Air flow simulation around a racing car with a 100M mesh

Postprocessing

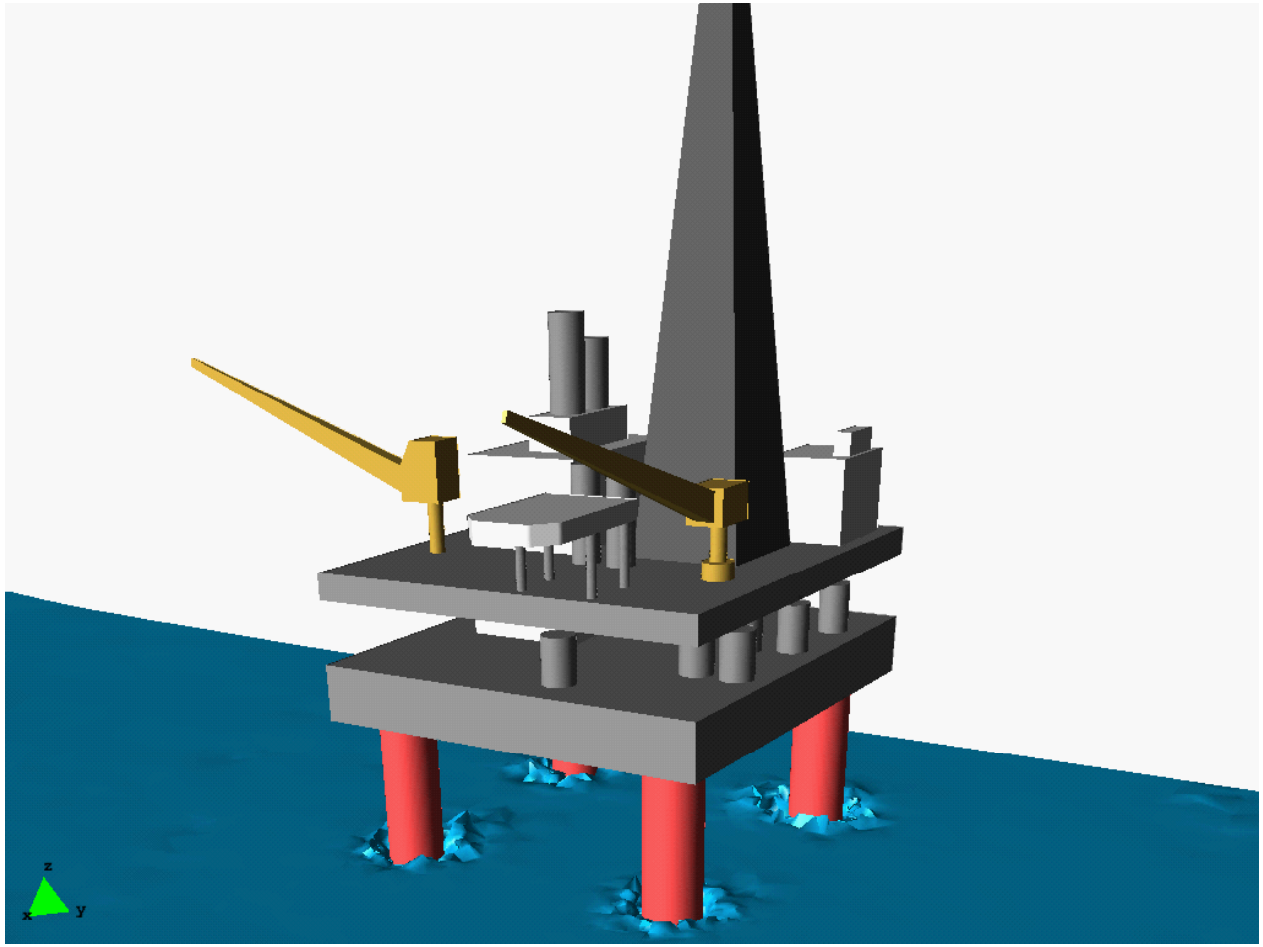
- Postprocessing operations (data analytics) and visualization of results



Air flow simulation around a racing car with a 100M mesh

Postprocessing

- Isosurfaces



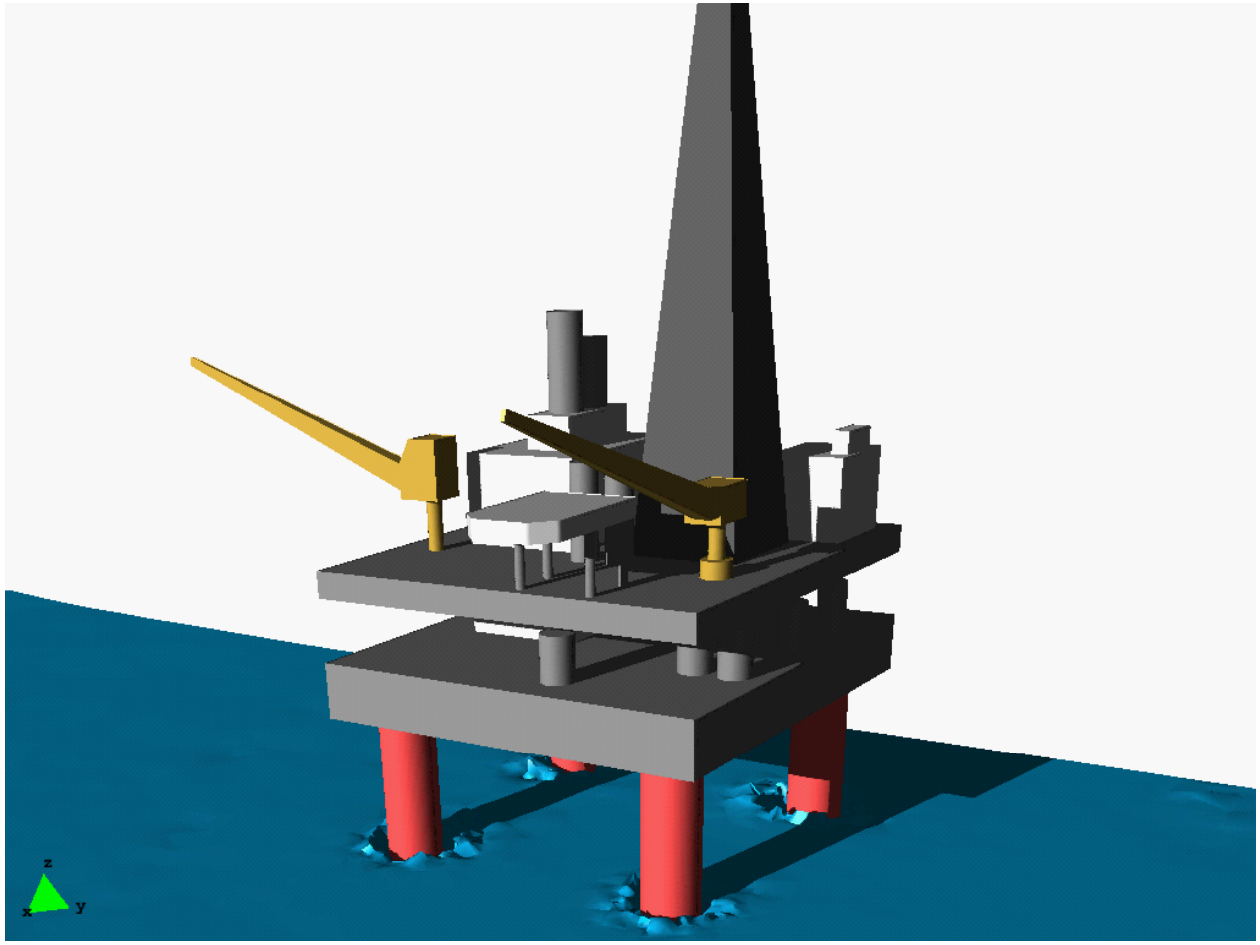
Extract skin of the model

Create isosurface of pressure equal to 0

Animate for each time step

Postprocessing

- Isosurfaces



Extract skin of the model

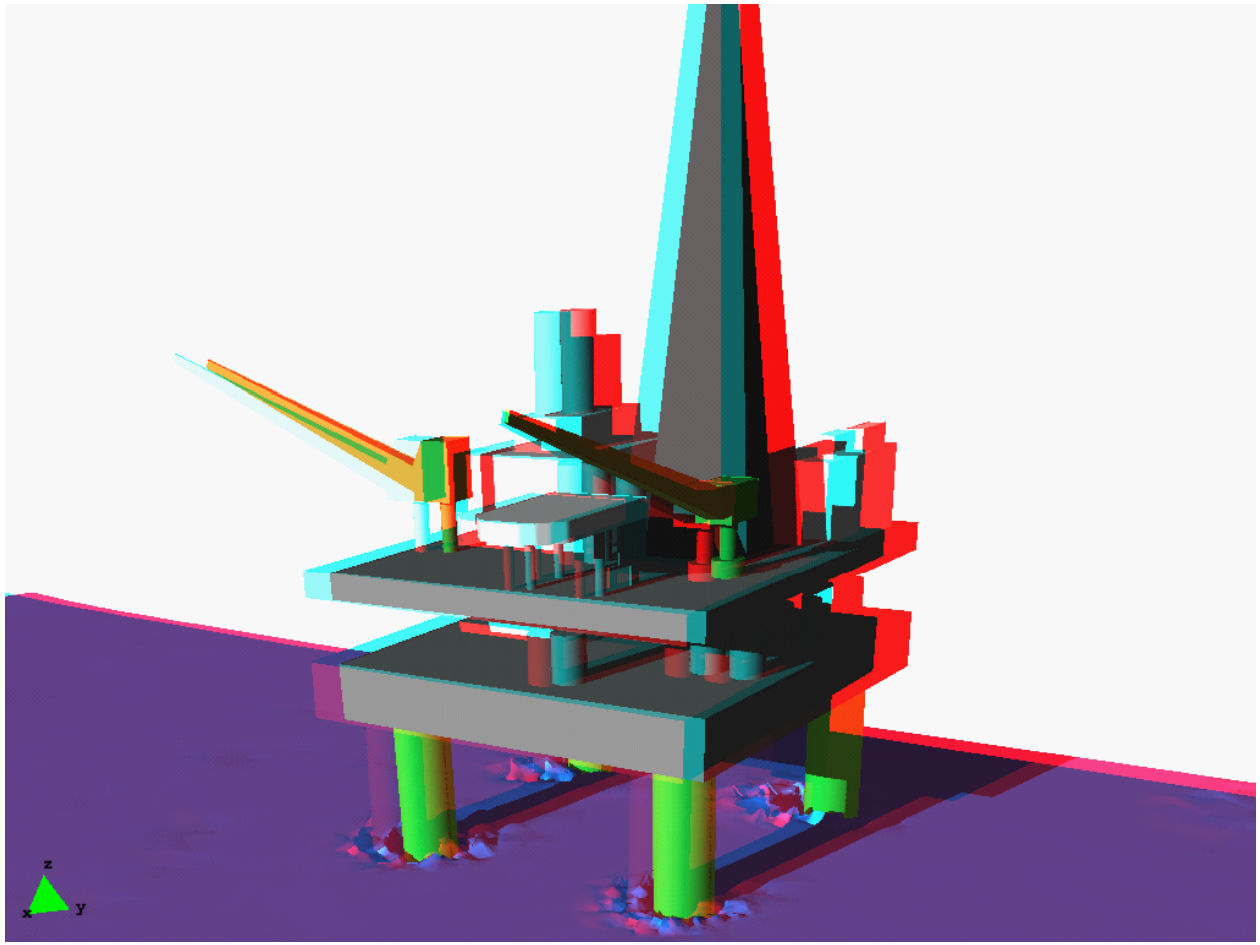
Create isosurface of pressure equal to 0

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

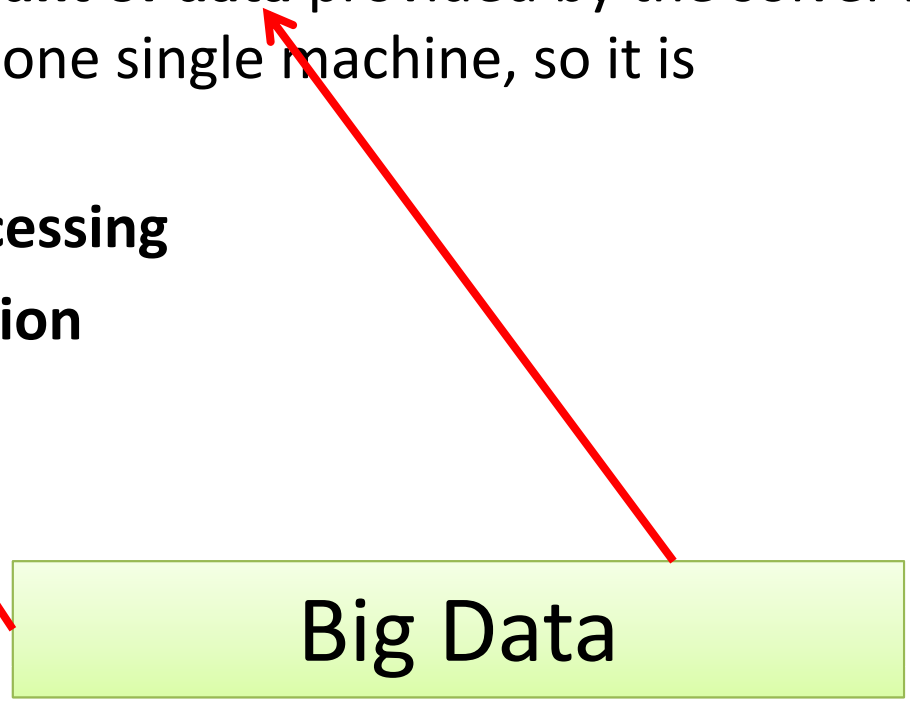
Animate for each time step

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**



VELaSSCo

*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.*

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 basic information

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

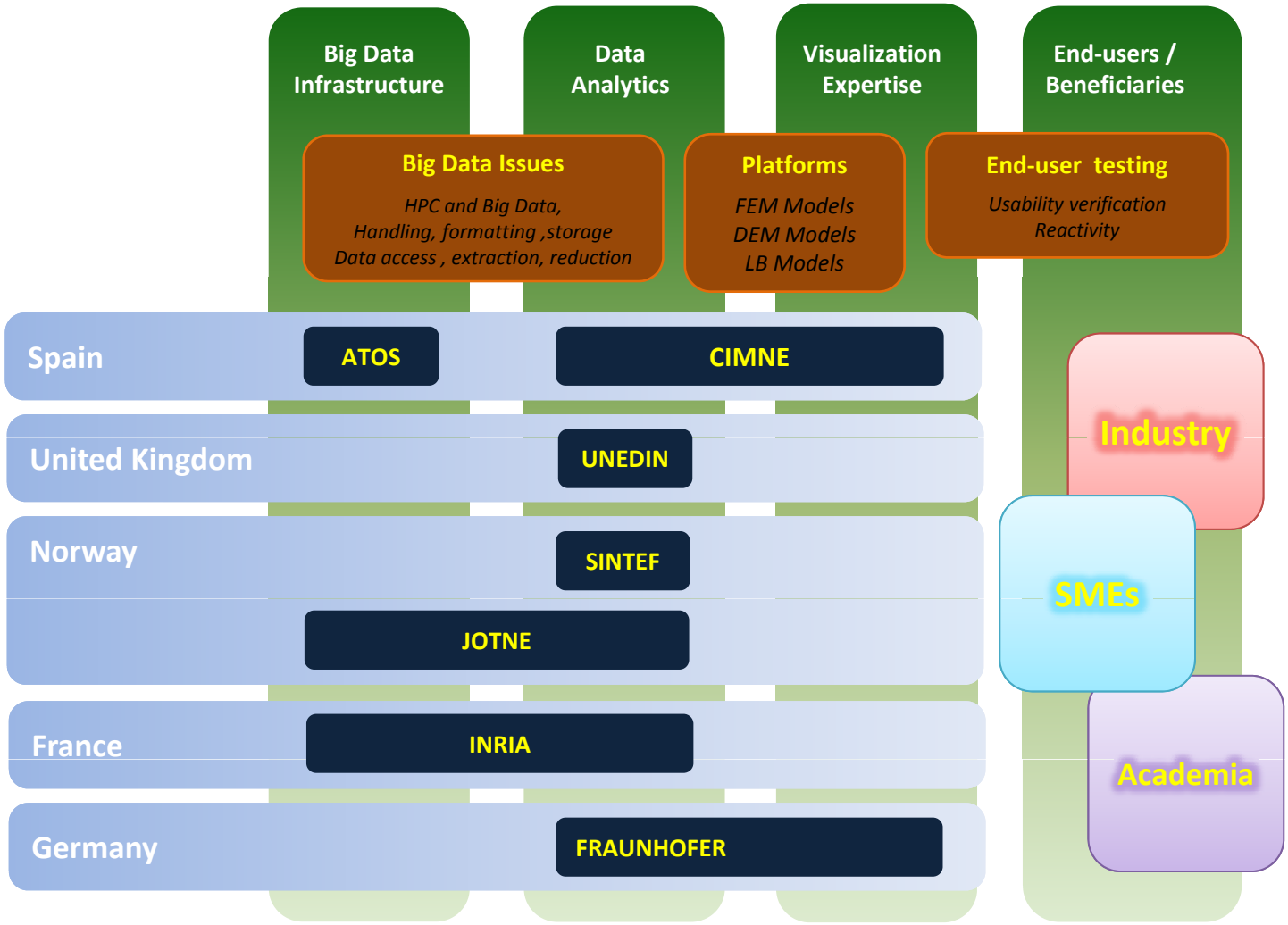
Participant organisation name	Short name	Country
International Center for Numerical Methods in Engineering	**CIMNE**	ES
School of Engineering. The University of Edinburgh	**UNEDIN**	UK
STIFTELSEN SINTEF	**SINTEF**	NO
Institut national de recherche en informatique et en automatique	**INRIA**	FR
Fraunhofer-Gesellschaft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

3 years project (2014 –2016)

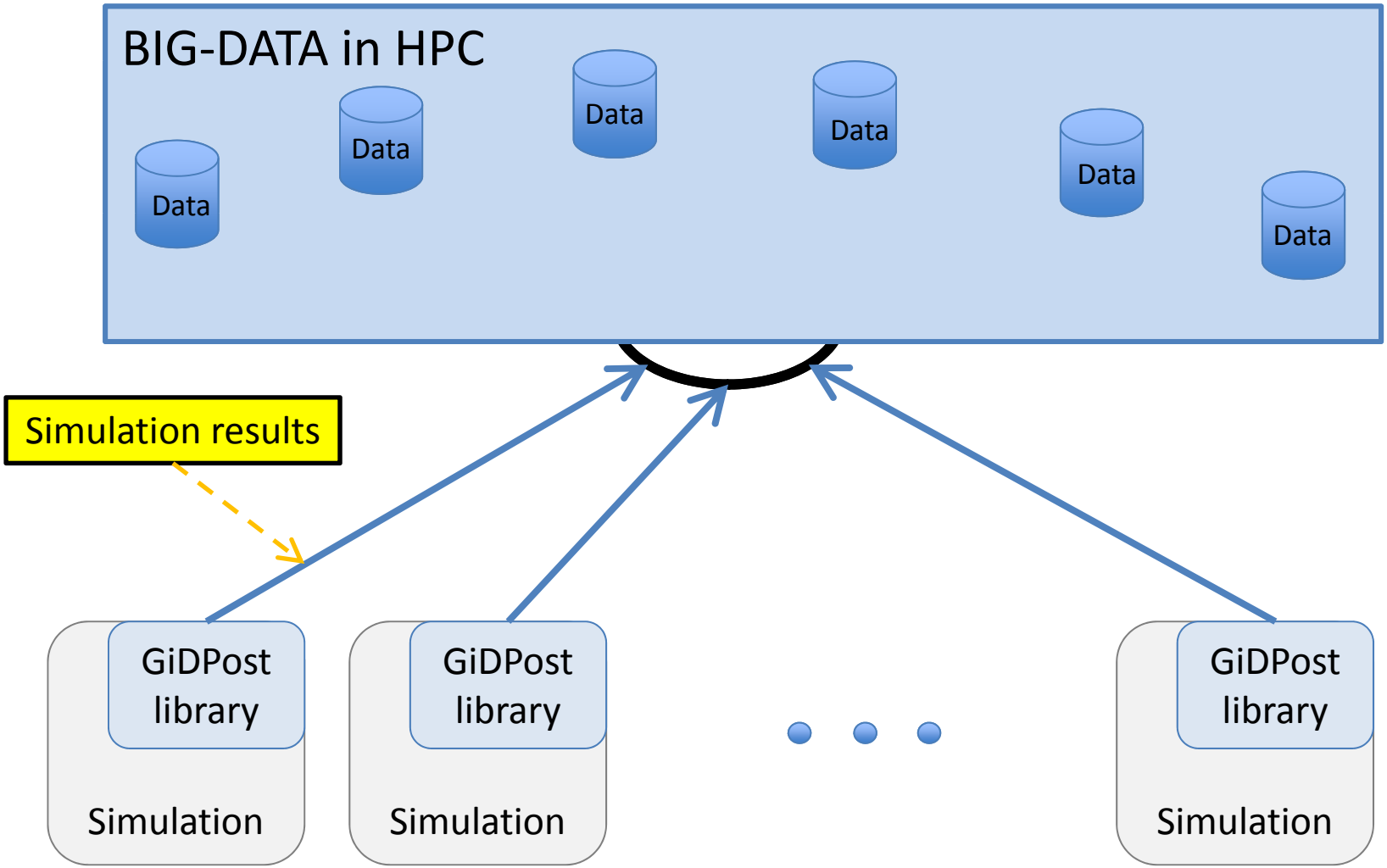
393 persons x month

The total costs of VELaSSCo are **4.441.603 €**, and total funding is **3.294.000 €** (original 3.294.425 €)

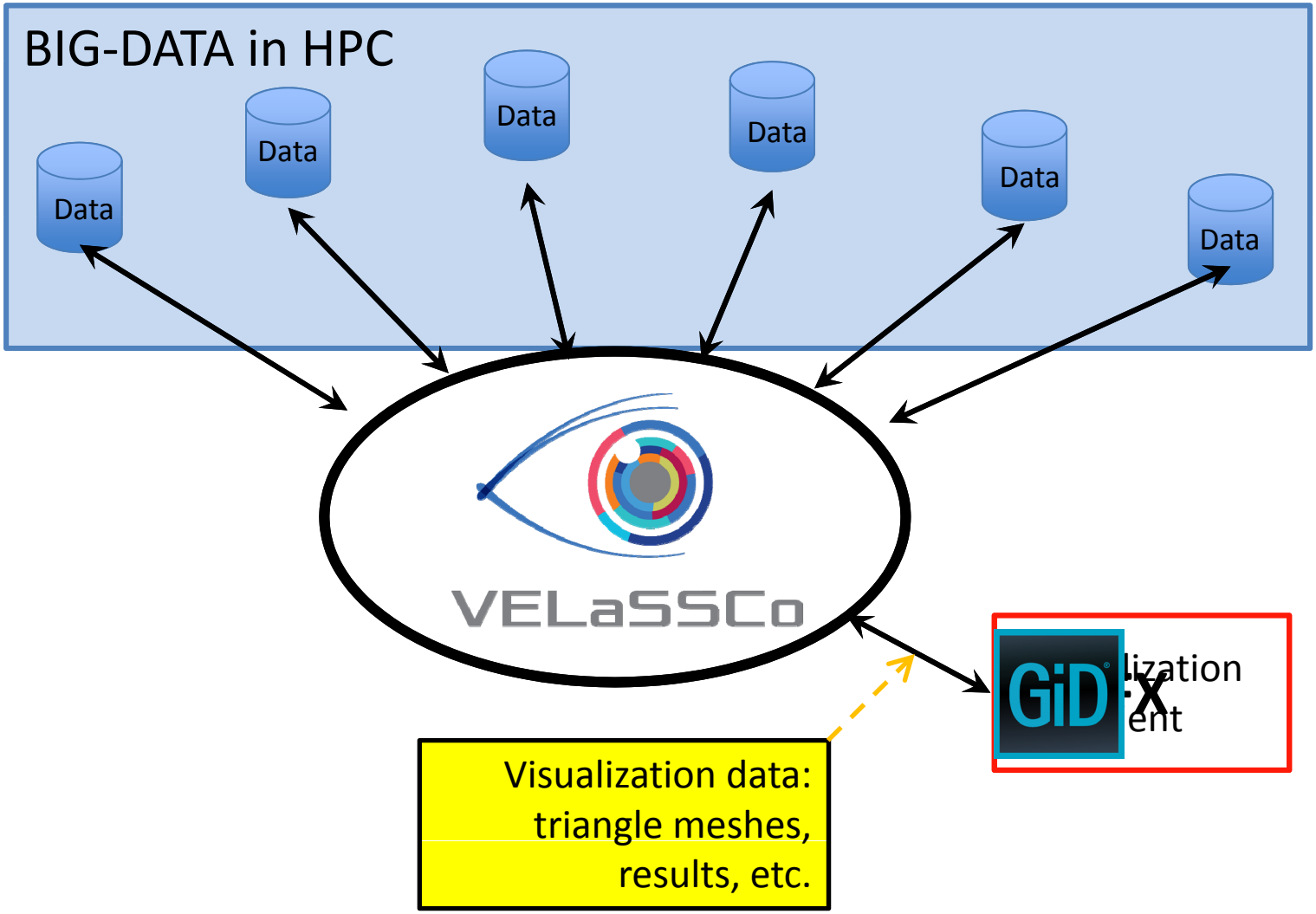
VELaSSCo consortium as a whole



VELaSSCo Platform



VELaSSCo Platform

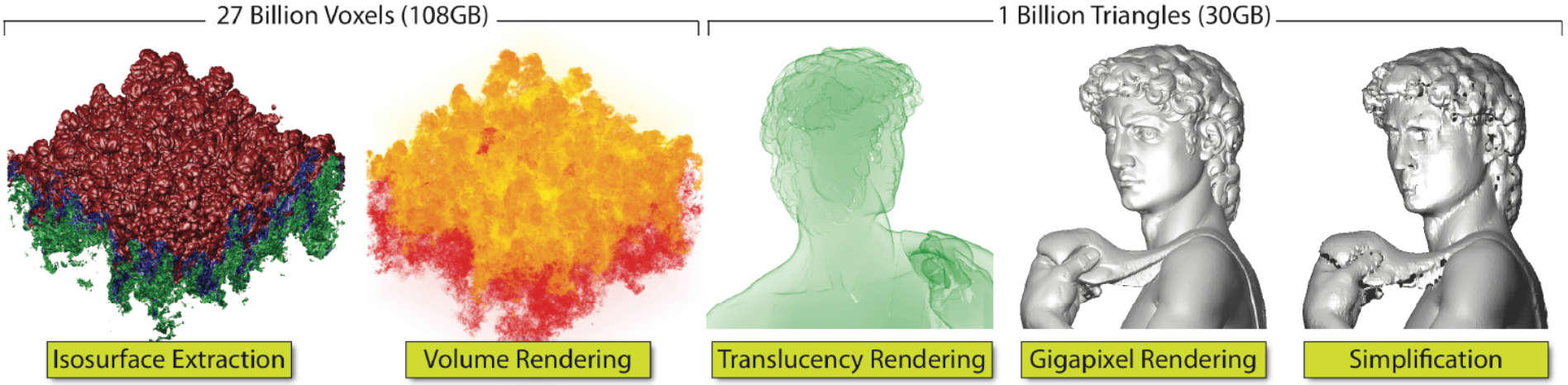


VELaSSCo architecture

- Two BigData frameworks:
  - OpenSource: using Hadoop, HDFS, Hbase, YARN
  - CloseSource: using JOTNE's EDM
- Visualization client communication using Thrift
- Data ingestion using Apache flume:
  - From files or running simulation (using GiDPost library)
- Runs on an HPC with:
  - Local storage on some nodes: Vnodes
  - Dedicated queue for Vnodes
  - Vnodes can also be used to calculate when idle

Why Hadoop?

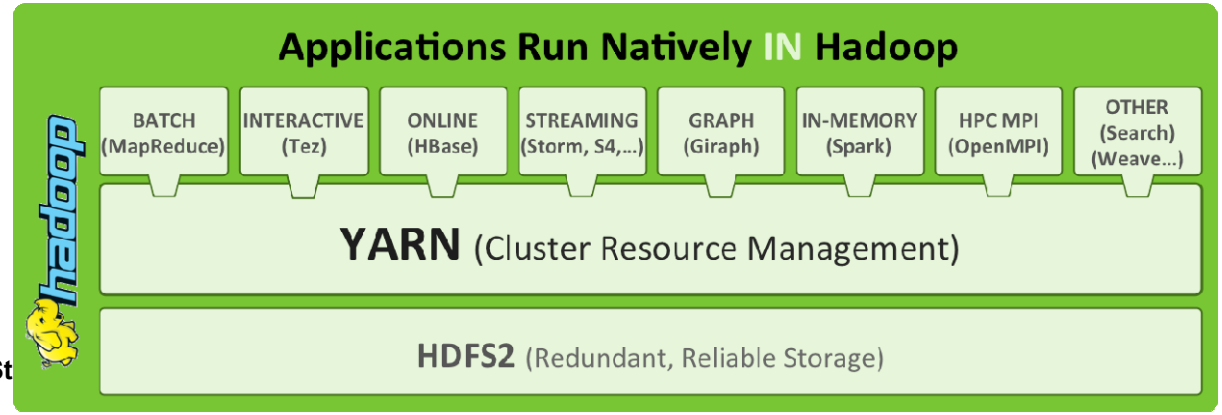
- H. T. Vo et al., “Parallel visualization on large clusters using MapReduce”, in *Proceedings of the IEEE Symposium on Large-Scale Data Analysis and Visualization*, October 2011, pp. 81–88, DOI 10.1109/LDAV.2011.6092321.



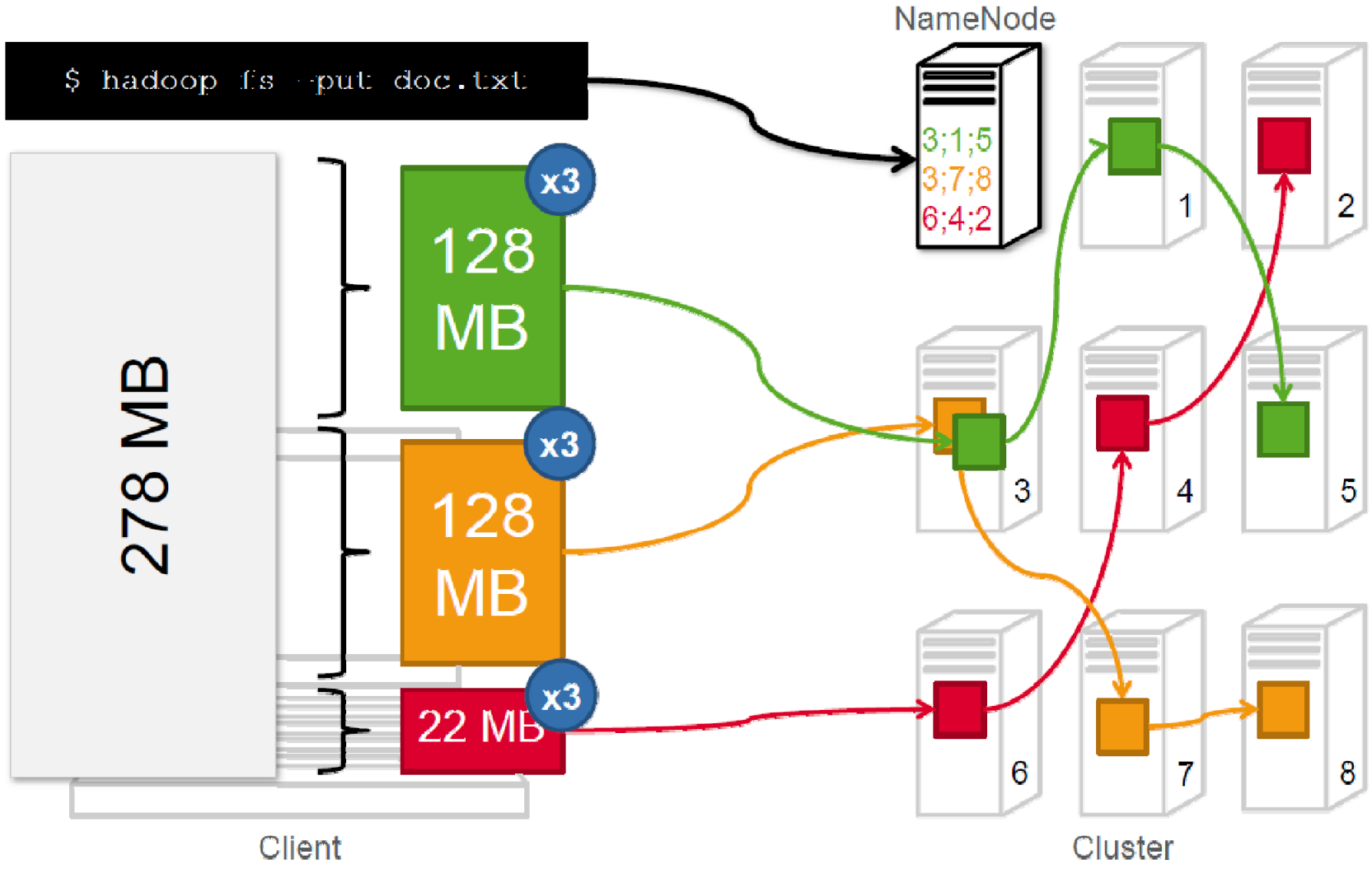
- J. A. Stuart et al., “Multi-GPU volume rendering using MapReduce,” in *1st International Workshop on MapReduce and its Applications*, June 2010.
- Da-Qi Ren et al., “Distributed Large Scale Mesh Simplification with MapReduce and MPI in 3-D Finite Element Electromagnetics with Tetrahedra”,

Hadoop

- "Apache Hadoop is an **open-source** software framework written in **Java** for **distributed storage** and distributed **processing** of very large data sets on computer clusters built from commodity hardware." (Wikipedia)
- Core blocks:
  - HDFS: Hadoop Distributed File System
  - YARN: Yet Another (computing) Resource Manager
  - MapReduce: programming model
  - Common: tool-set, scripts for extensions
- Extensions:
  - Hbase: distributed (table) database, Hive: SQL-like access
  - Spark, Flink: fast data analysis
- Scalable 10,000 nodes
- Uses local storage on each data node



HDFS



From "Hadoop in a Nutshell"
http://www.metafinanz.de/sites/default/files/Hadoop_in_a_Nutshell.pdf

Hbase in VELaSSCo

- Table (column) oriented distributed data store, over HDFS
- 3 tables:
  - List of models

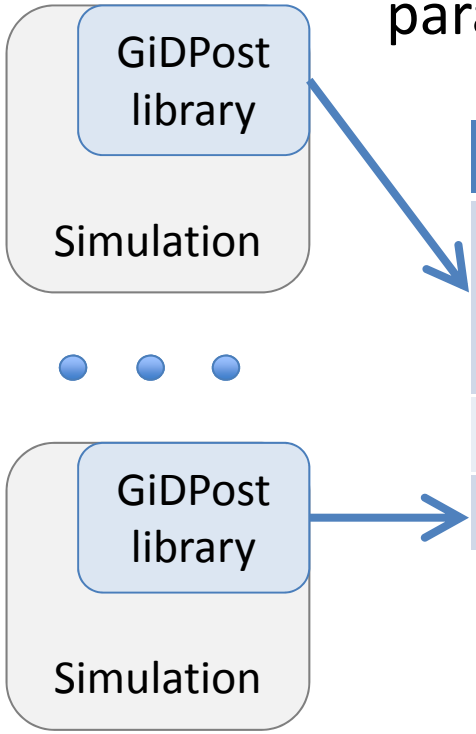
Rowkey	Properties
ModelID_1	nm: “car”, fp: “/home/…car123”
ModelID_2	nm: “bcn”, fp: “/…/bcn100M”

  - Simulations metadata: faster retrieval of mesh and results information

Rowkey	M	R
ModelID+Analysis+TS_1	m01_nm: “body” m01_et: “tetrahedral” … m02_nm: “fluid” …	r01_nm: “Velocity” r01_rt: “Vector” … r02_nm: “Distance” …
ModelID+Analysis+TS_2	…	…

Hbase in VELaSSCo

- 3 tables: list of models, simulation metadata, simulation results
  - Simulation results: one row per partition to avoid blocking, parallel execution of jobs

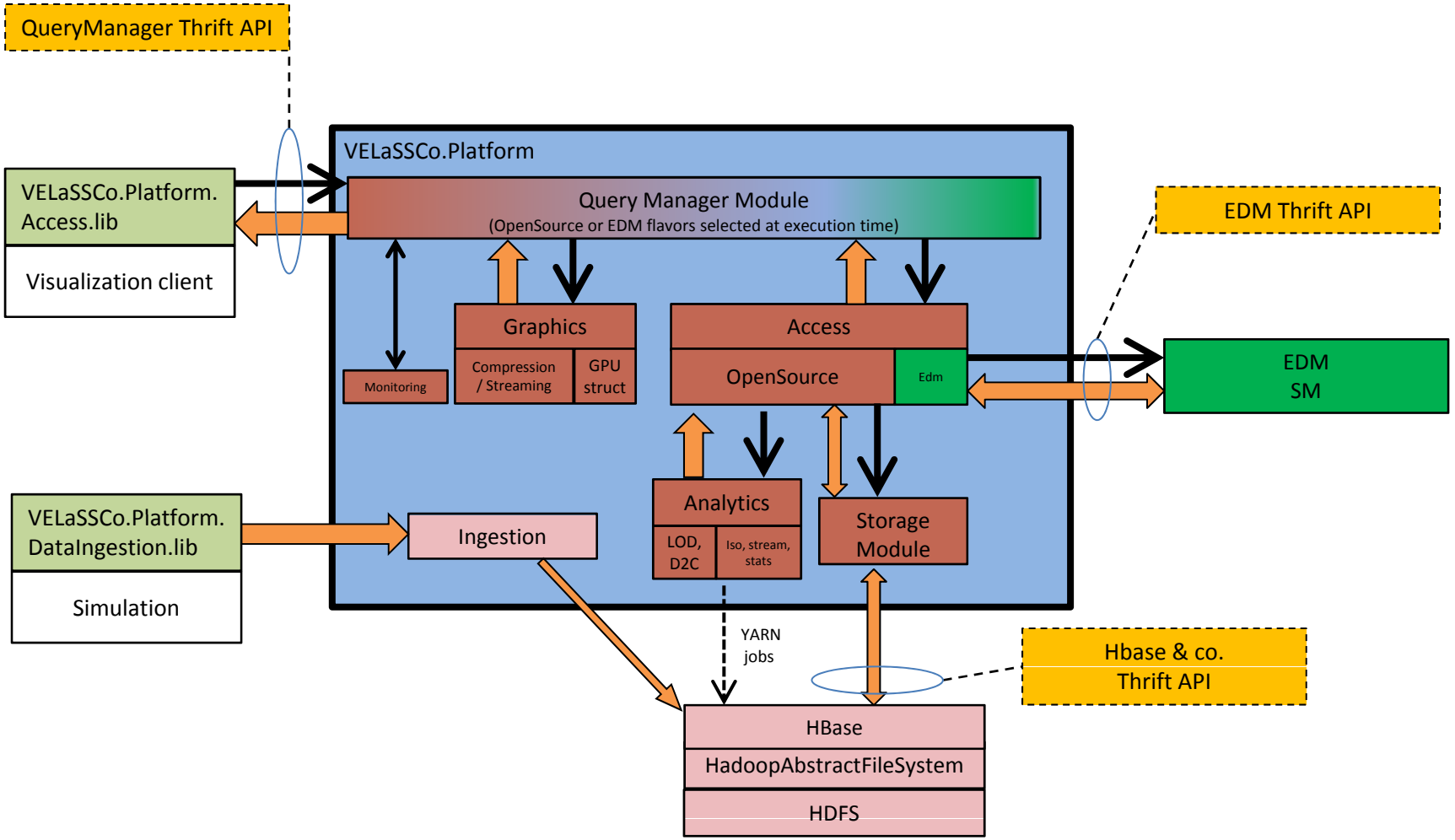


Rowkey	M	R
ModelID+An+TS_1+Part1	c01_1: xyz, c01_2: xyz, … e01_1: n1n2n3n4, …	r01vl_1: VxVyVz r01vl_2: VxVyVz … r02vl_1: d
ModelID+An+TS_1+PartX	…	…
ModelID+An+TS_1+PartN	…	…

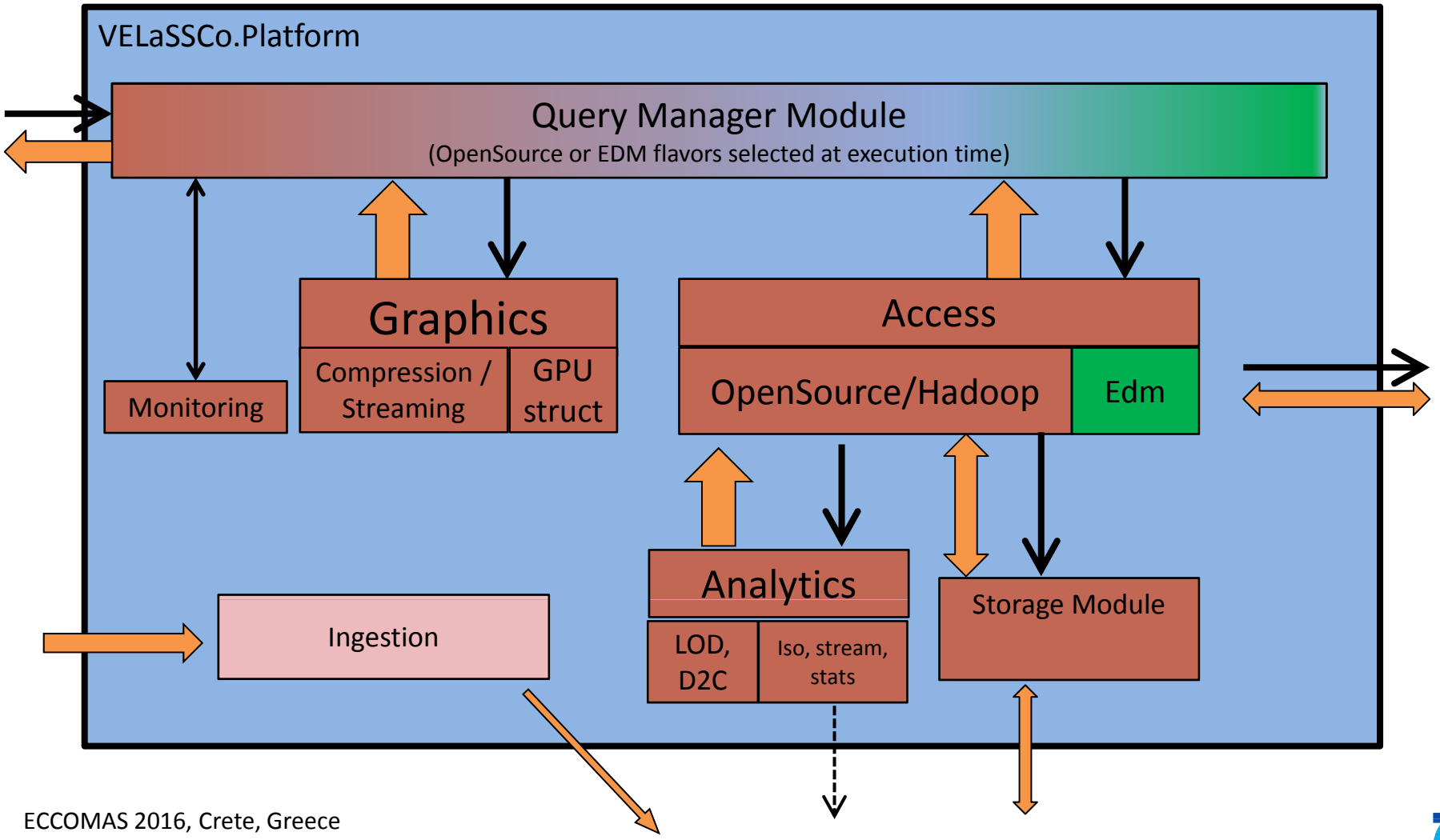
MapReduce: getBoundaryMesh

- INPUT: From Simulations_Data table, get rows with
  row-key = ModelID+An+TS1+* value= M:e*
- MAP: Calculate partial boundaries
  output list of unique triangles: key=t1,t2,t3 value=1
- Shuffle/Sort (done by yarn): Merges repeated triangles together
  triangles: key=t1,t2,t3 list of value = 1, 1, 1, 1
- REDUCE: Drop repeated triangles
  output unique triangles (t1,t2,t3)

Architecture (April 2016)

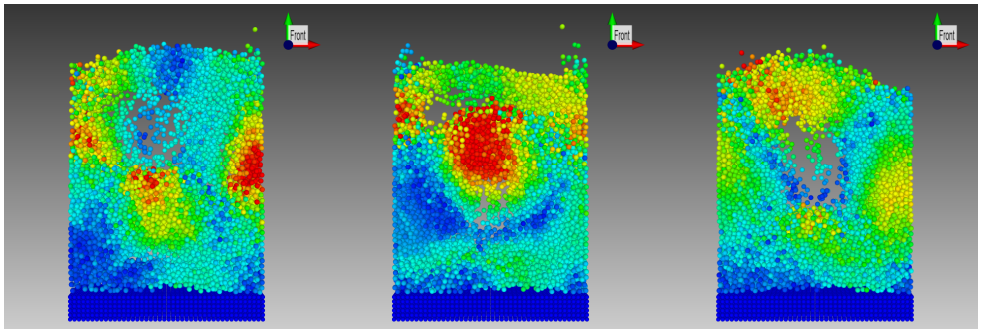


VELaSSCo platform



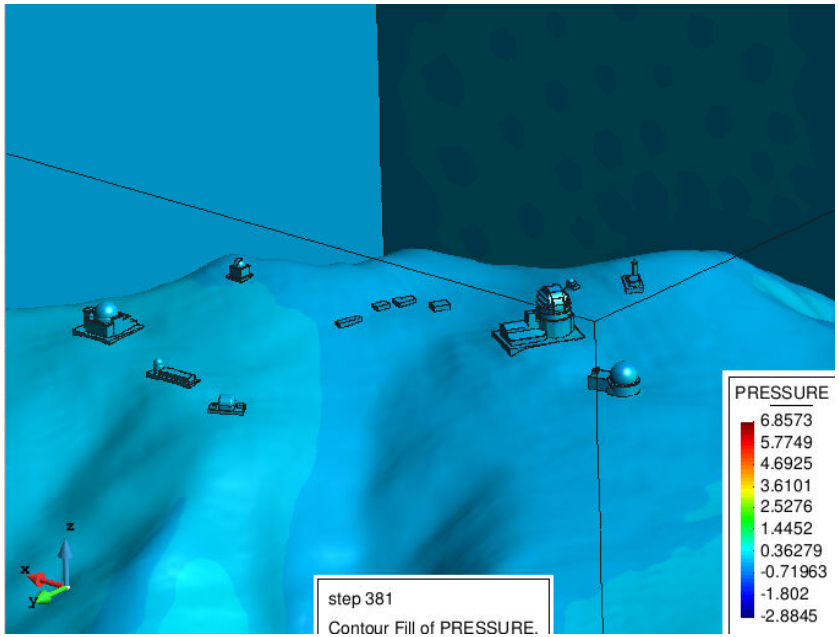
Some numbers…

- Data sizes we are currently working with at **first prototype level** (March 2016)



Fluidized bed DEM example:

- 12.000 particles / time-step
- 3000 contacts (p2p and p2w) / time-step
- 40.000 time-steps
- Mass, volume and velocity vector, and force vector for contacts

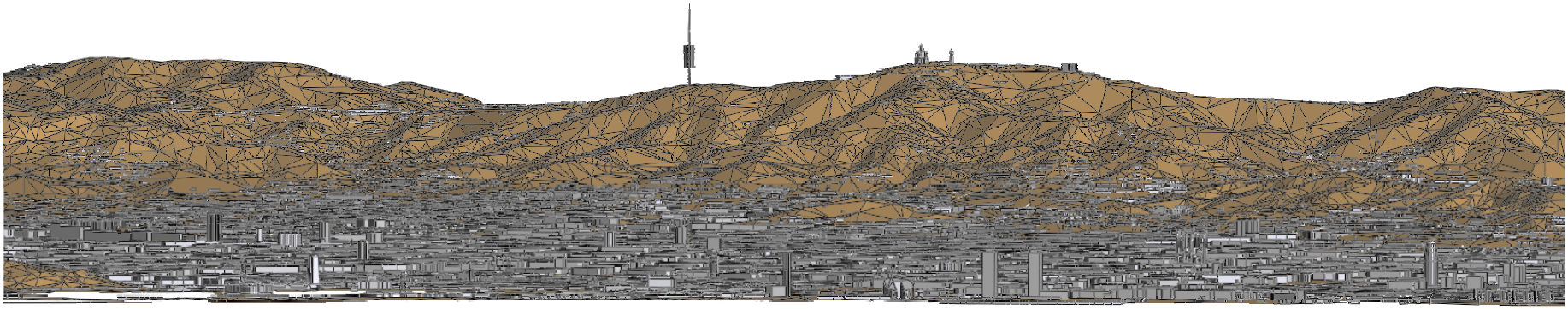


Telescope FEM example:

- 4 M nodes, 28 M tetrahedra
- 19 time steps
- Partition index, pressure and velocity

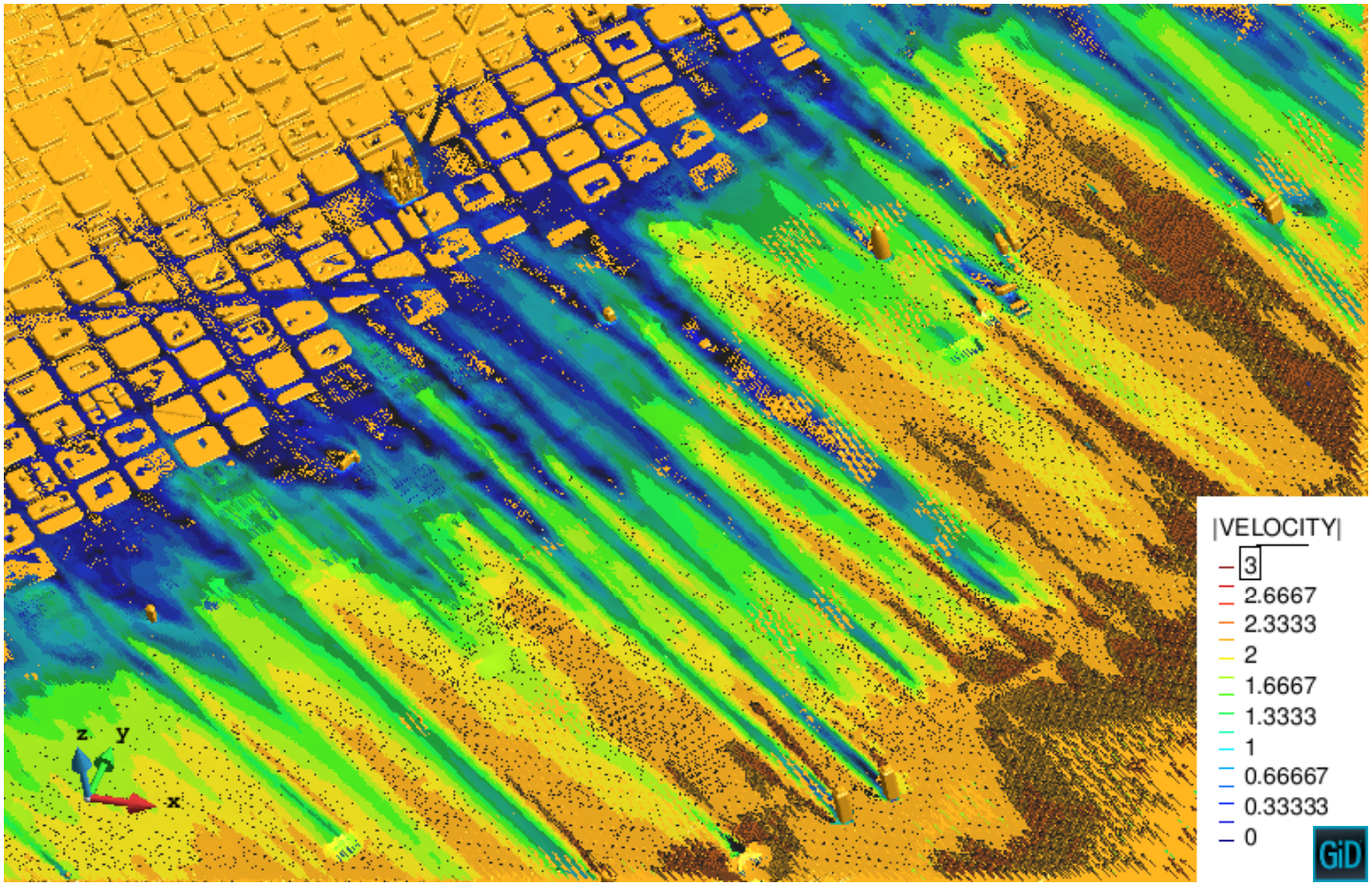
Some numbers…

- Data sizes we have for the **final prototype** of VELaSSCo Platform
  - Barcelona model (domain of 64 x 64 x 1 Km):
    - ~**400M Tetrahedra** (4 m resolution), ~3000 sub-domains
    - ~1000 time-steps
    - 2 TB disk space



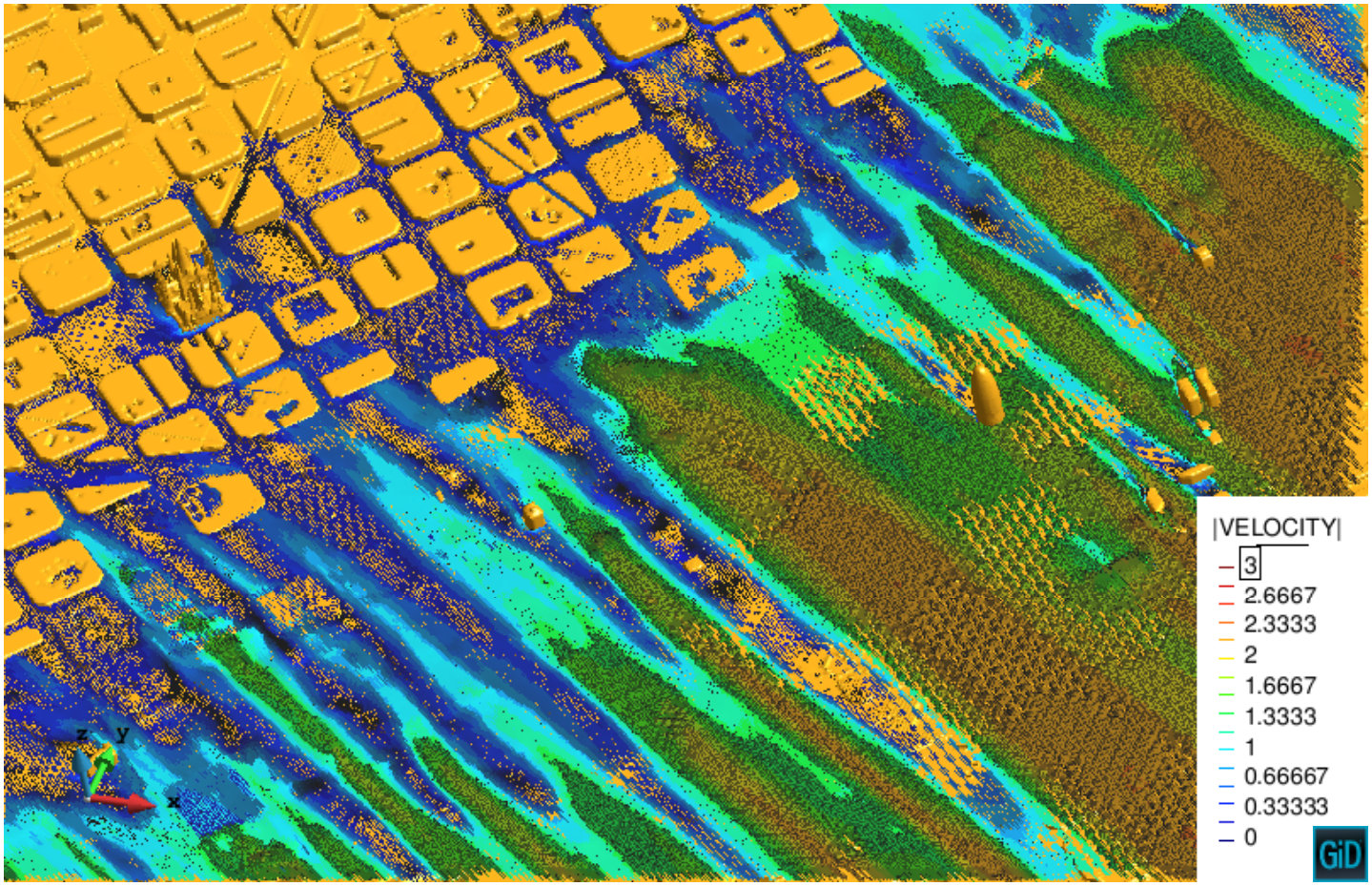
Simulations done in the framework of Numexas EU funded project

Barcelona model



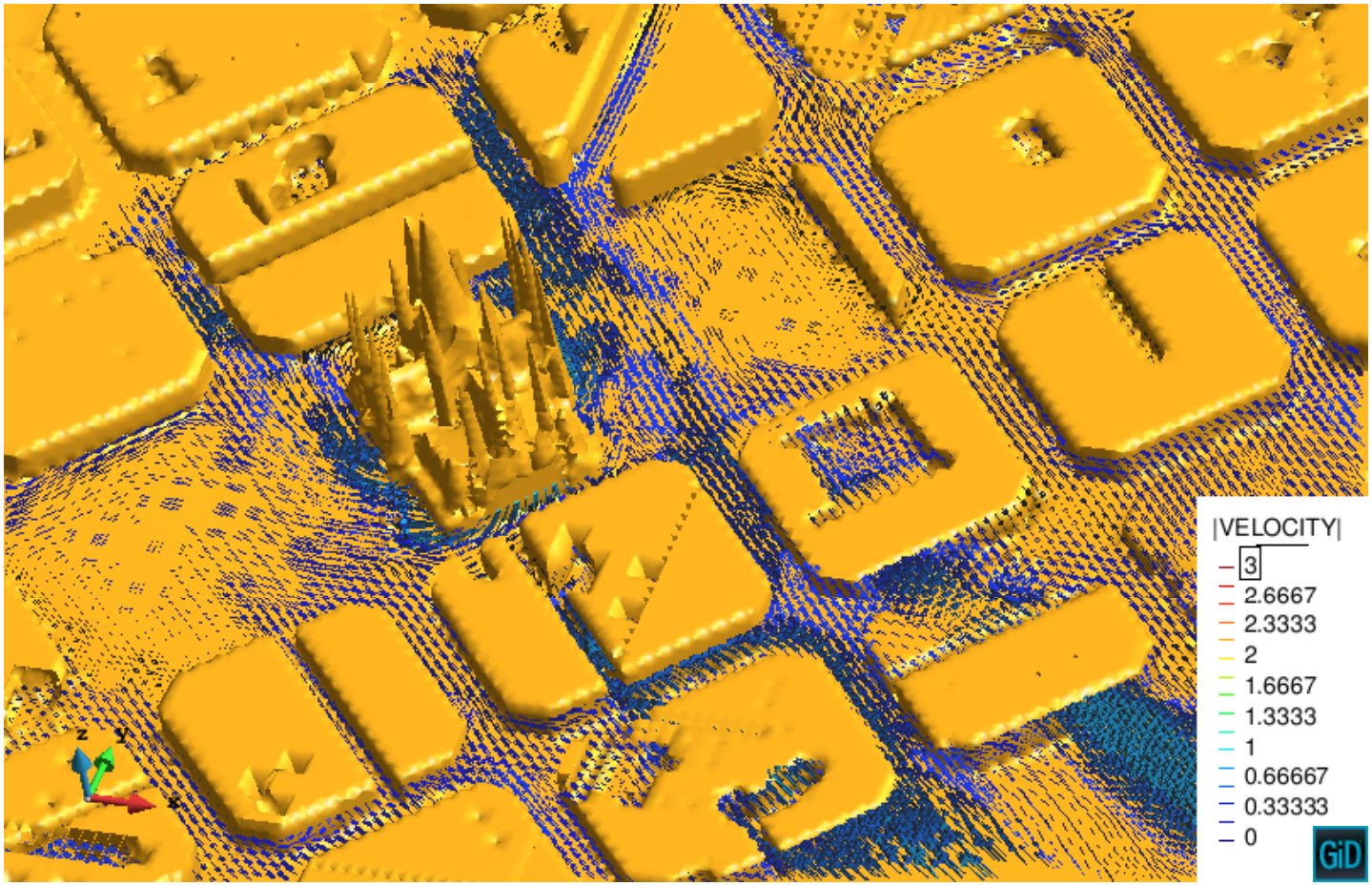
100 M elements correspond to a resolution of 8m.

Barcelona model



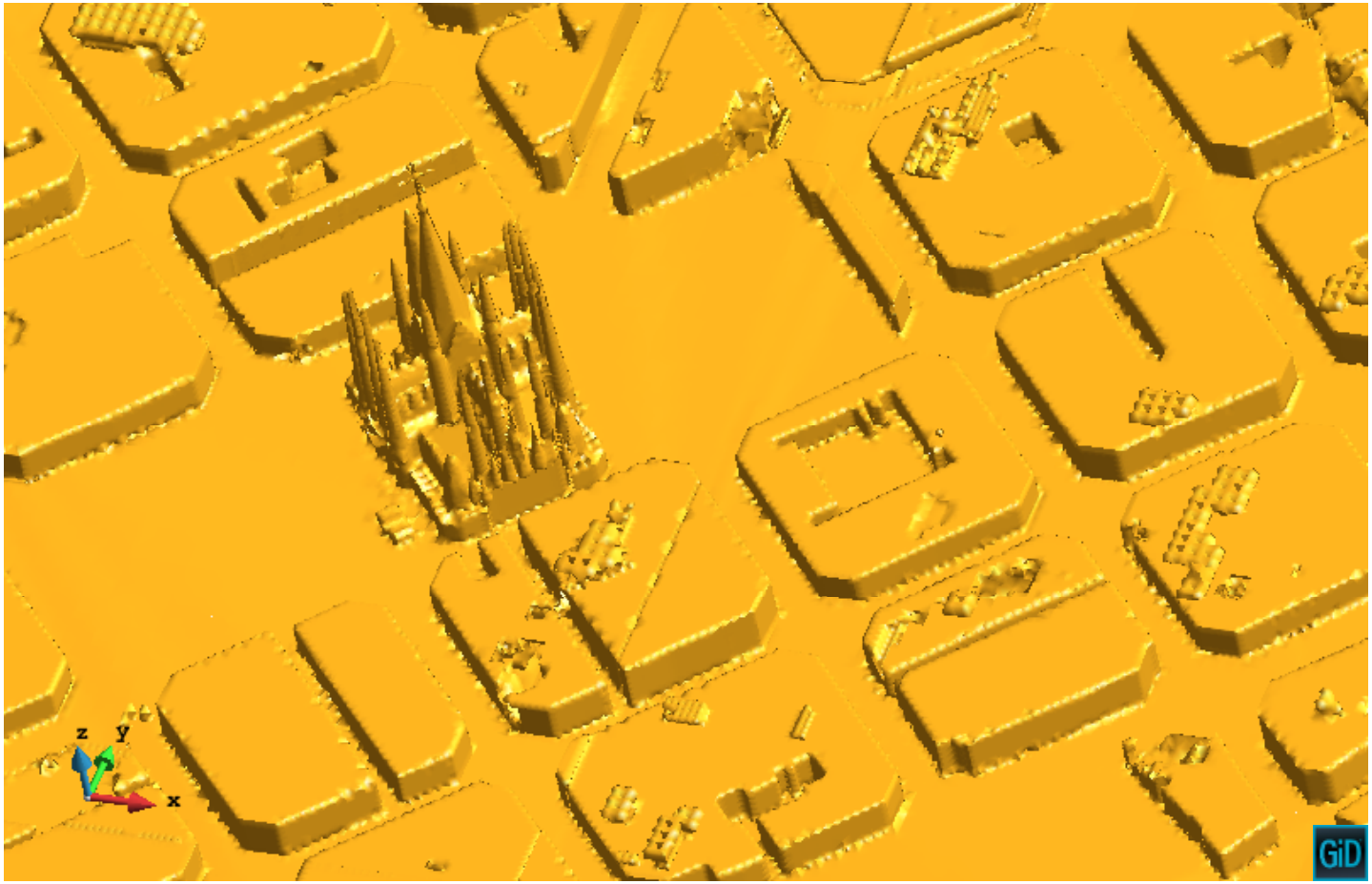
100 M elements correspond to a resolution of 8m.

Barcelona model



100 M elements correspond to a resolution of 8m.

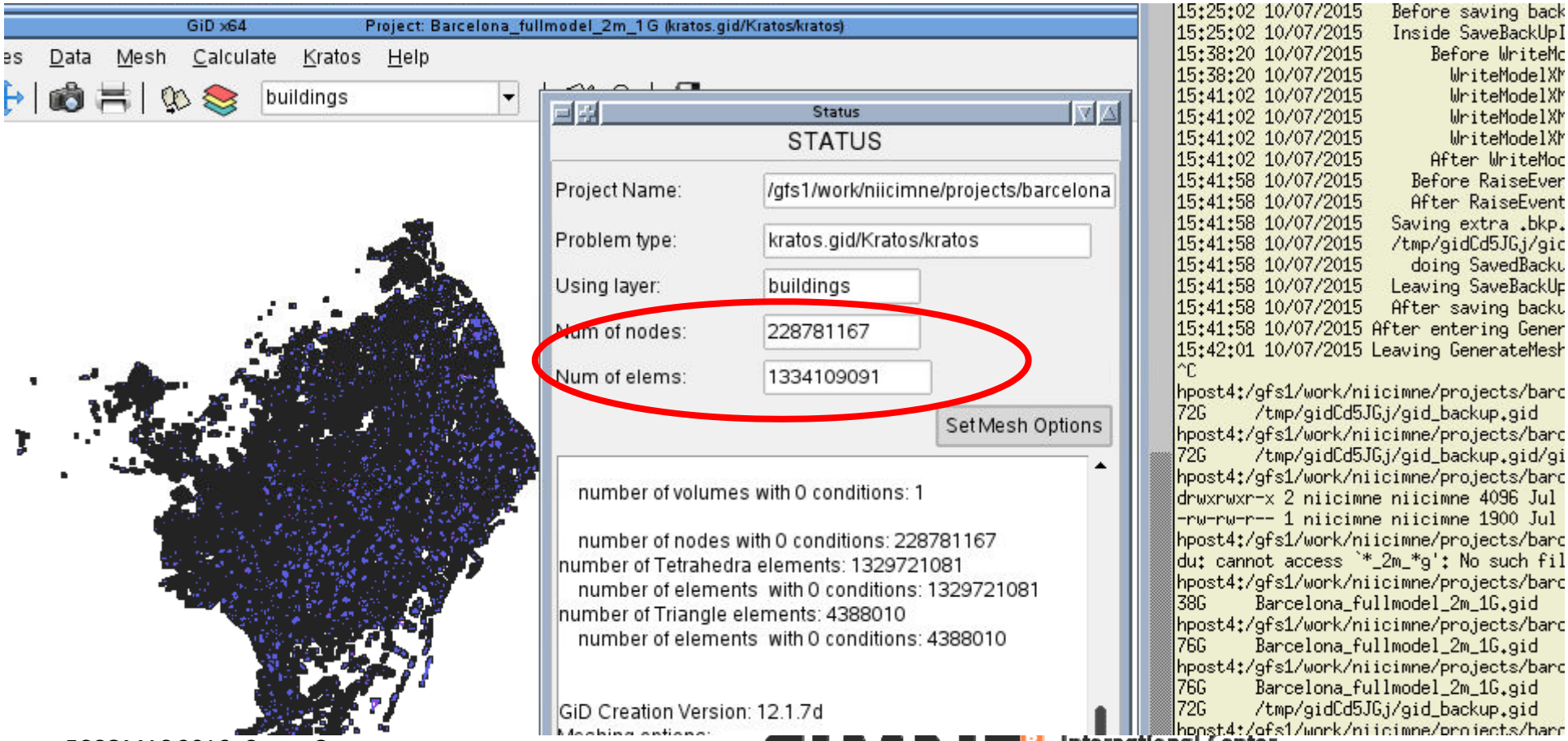
Barcelona model



340 M 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)

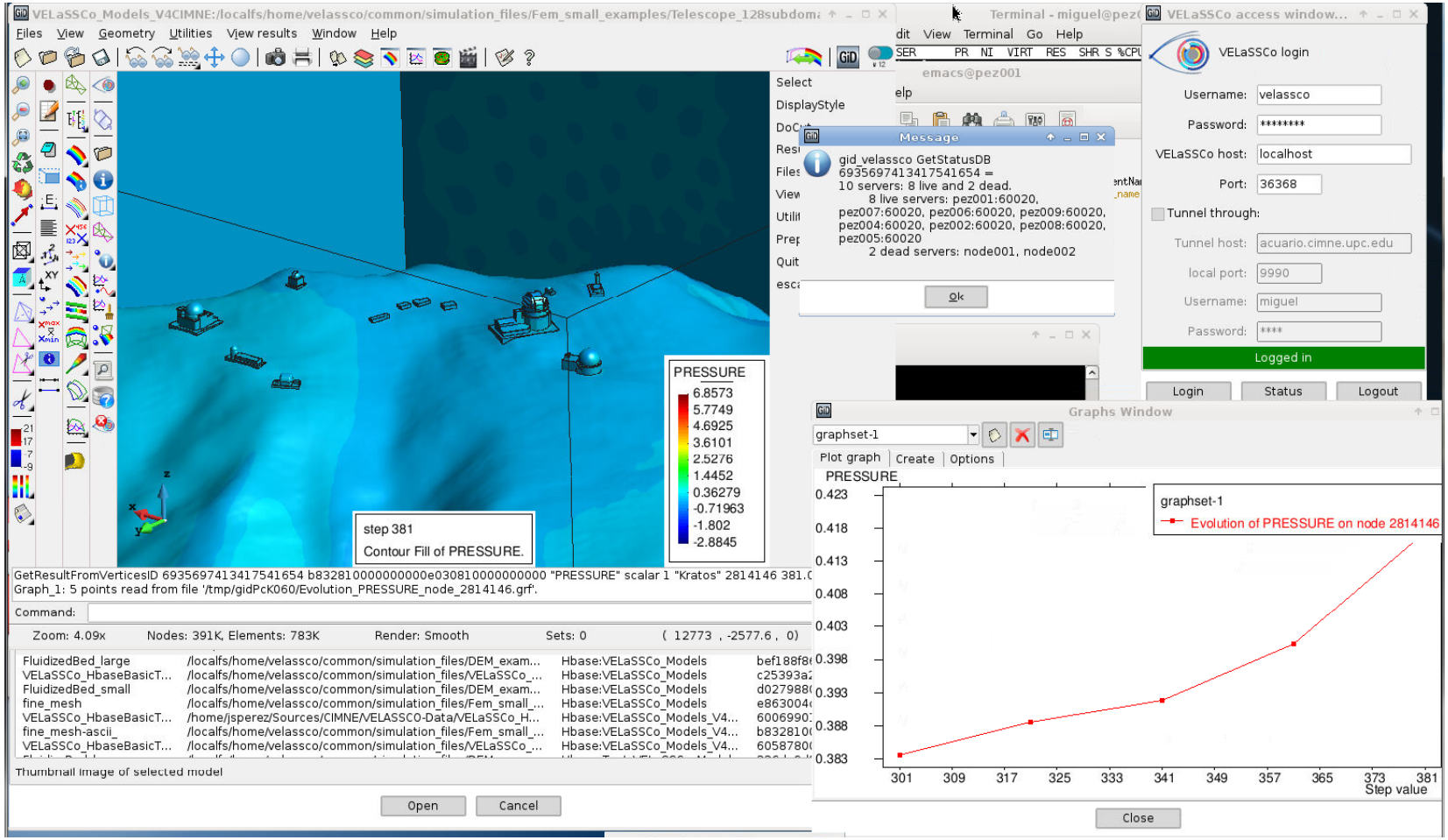


VELaSSCo functionalities

- Session connection:
  - User Authentication
  - List of models for model selection with information summary:
    - Name, number of nodes, elements, list of analyses, time steps, and results properties (name, type, …)
- Model view:
  - Get DEM particles, mesh triangles
  - Extract the bounding box of the model, skin of a mesh of tetrahedrons
- Results view & analysis:
  - Given a list of nodes (vertices), get the results values for these nodes
  - Result evolution on nodes graphs
  - Discrete to continuum transformation

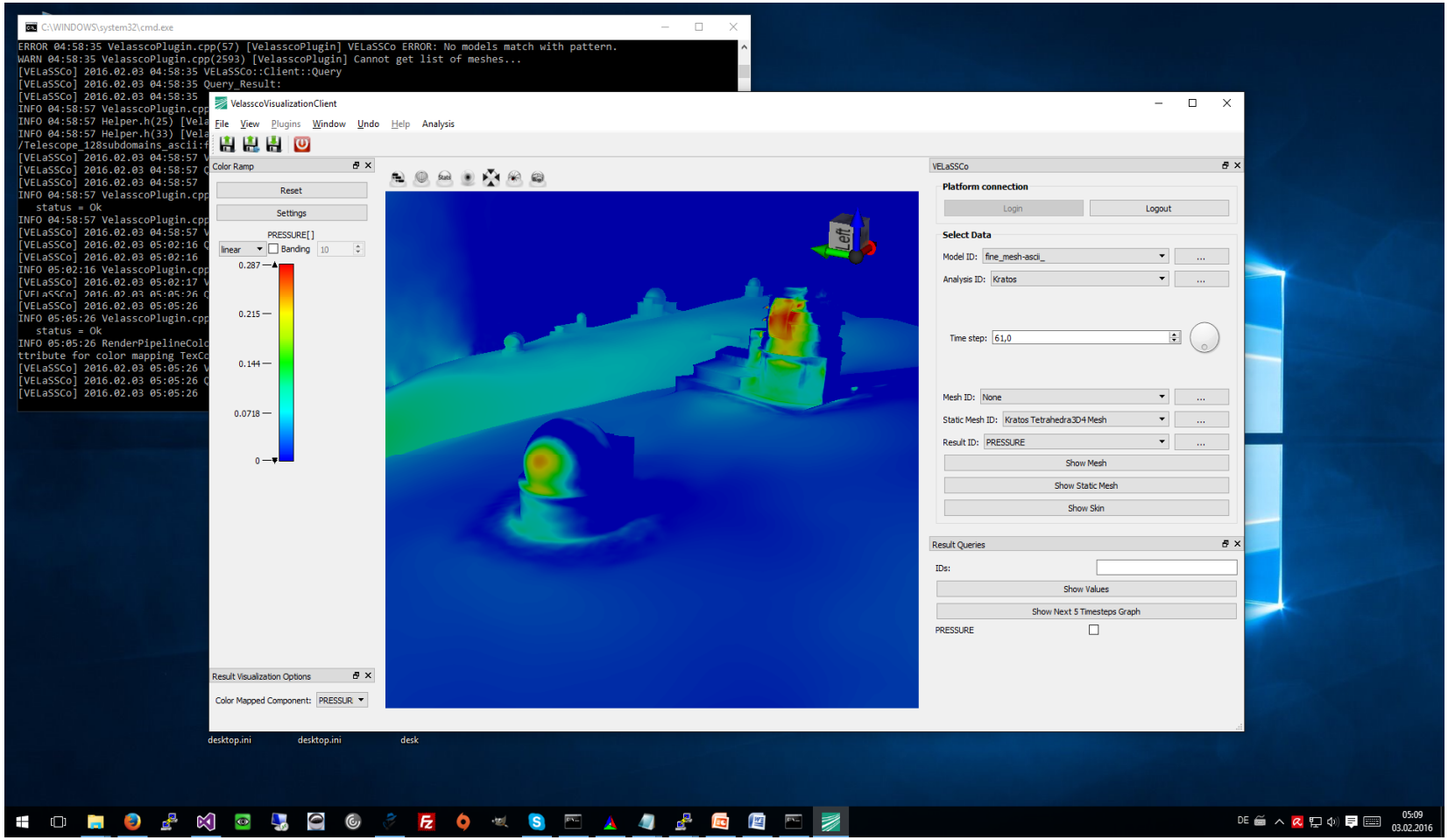
First prototype of VELaSSCo Platform

- GiD-VELaSSCo plug-in



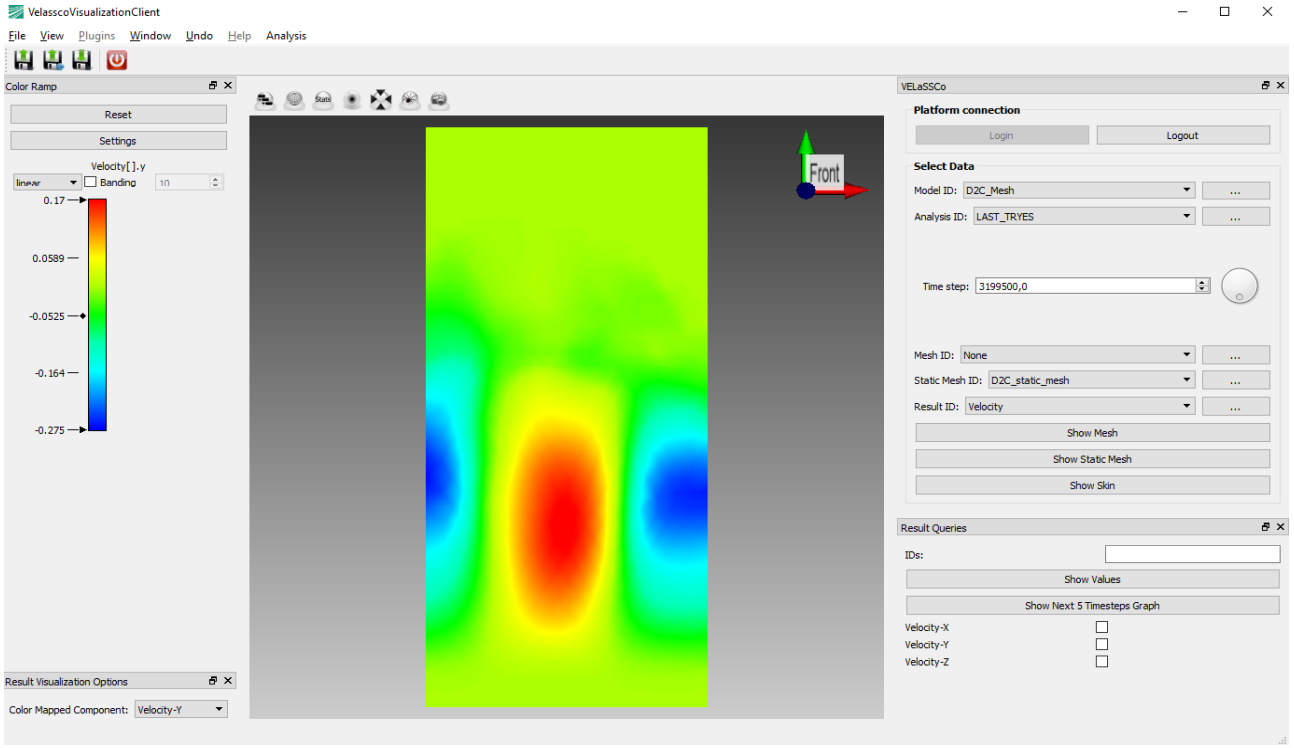
First prototype of VELaSSCo Platform

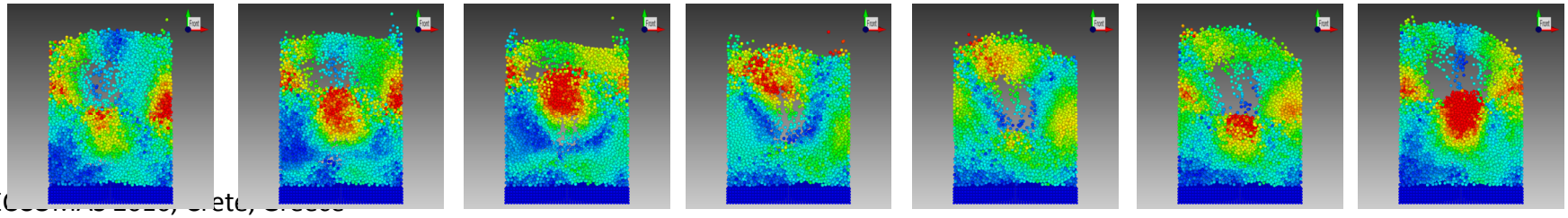
- Fraunhofer's iFX visualization client



First prototype of VELaSSCo Platform

- Fraunhofer's iFX visualization client





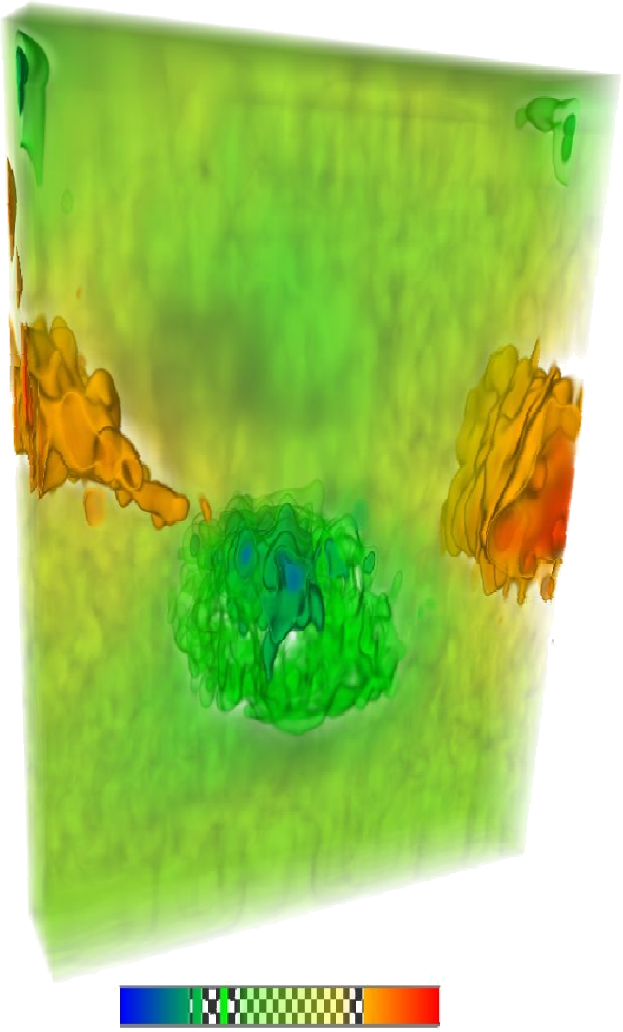
Next steps

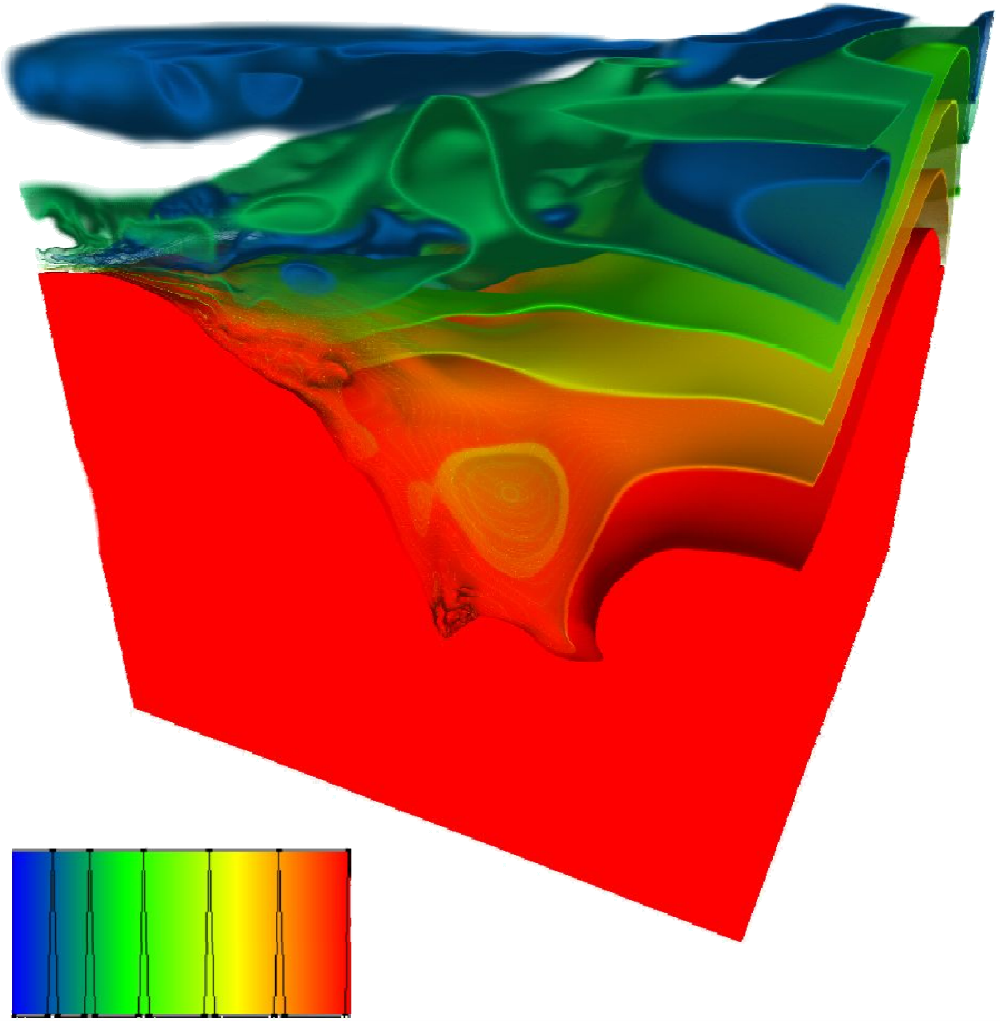
- Final prototype to be ready by September 2016
  - Functionalities: Iso-surfaces, cut planes, streamlines, results statistics, spatial integration
  - Multi-resolution: mesh simplification, LR-Bspline representations,
  - Workflow: animation and navigation = streaming of data
  - Life data injection from a running simulation

- Evaluation event to be held on October/November 2016

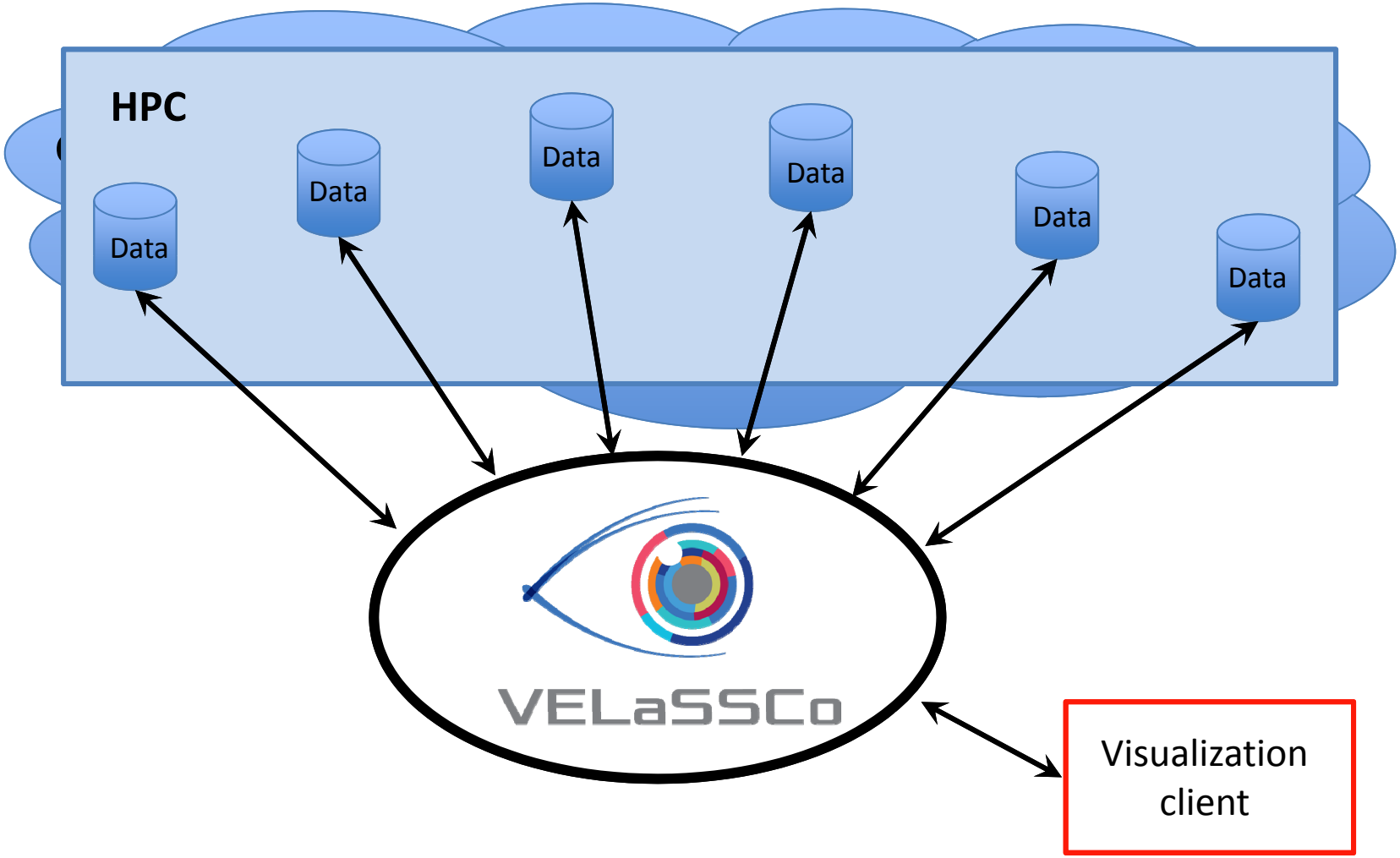
- Integration as a product in some partners strategy

Locally Refined (LR) B-splines representation





VELaSSCo Platform



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 post-processing and visualization for numerical simulations

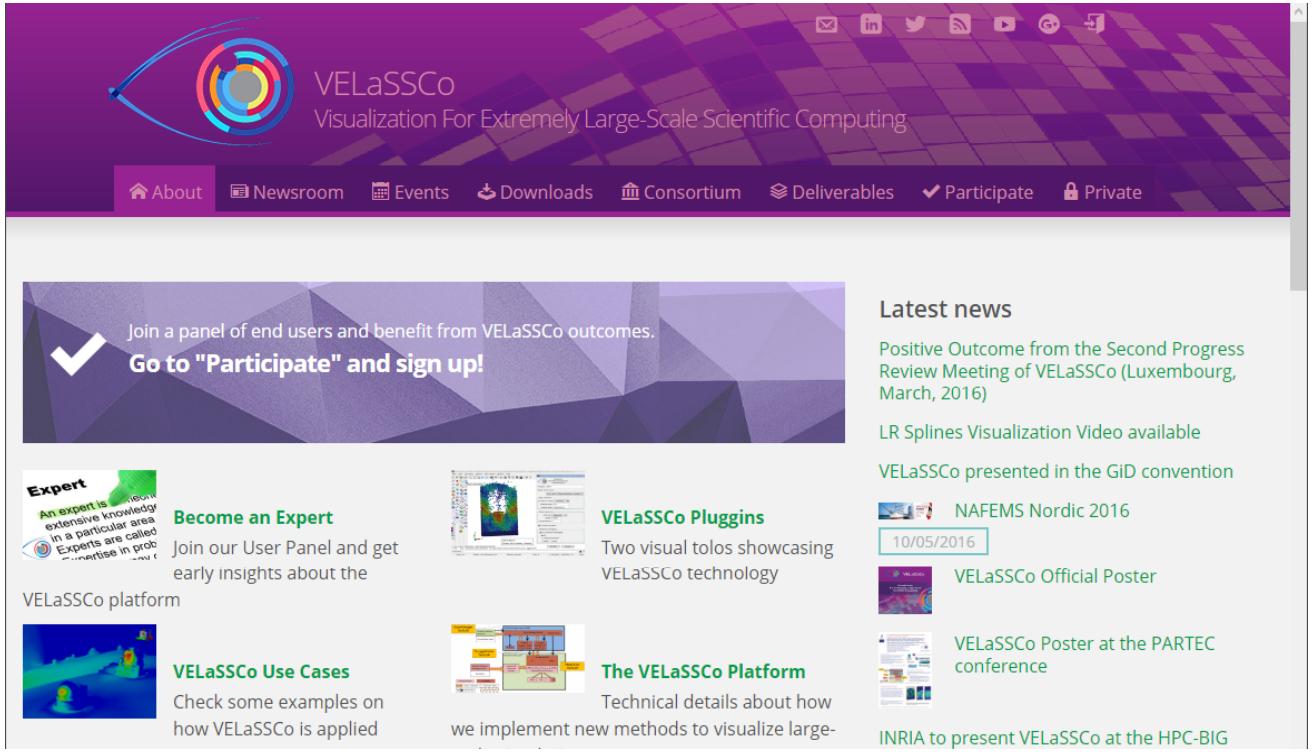
Lessons learnt

- Hadoop is not easy
- Big-data expertise is high in demand
- Telescope model (24m) is a crossing example:
  - “too big” for modest computers
  - “too little” for Hadoop (requires tuning)
- Evolving framework
- Data explosion: 5-6 factor (+ redundancy)
- Some limitations when programing in C++
- Yarn: not suitable for “interactive” queries.

VELaSSCo

Thank you for your attention

www.velassco.eu



ECCOMAS 2016, Crete, Greece

June 6th - 10th 2016