13th Summer School on SCIENTIFIC VISUALIZATION

Welcome!

Roberto Mucci – r.mucci@cineca.it
SuperComputing Applications and Innovation Department

CINECA

OUTLINE

- **School presentation**
- Introduction to Scientific Visualization
- Remote visualization @ Cineca

ABOUT CINECA

CINECA is a non profit Consortium, made up of 69 Italian universities, and 3 Institutions.

SCAI (SuperComputing Applications and Innovation) is the High Performance Computing department of CINECA, **the largest computing centre in Italy and one of the largest in Europe**.

The mission of SCAI is to **accelerate the scientific discovery by providing high performance computing resources**, data management and storage systems and **tools** and HPC services and **expertise** at large, aiming to develop and promote technical and scientific services related to high-performance computing for the Italian and European research community.

COURSES AND SCHOOLS

http://www.hpc.cineca.it/content/training

- Efficient use of **Molecular Dynamics** simulation applications in an HPC env.
- HPC Computer Aided Engineering
- HPC enabling of **OpenFOAM** for CFD applications
- HPC Numerical Libraries
- Introduction to **Fortran90**
- Introduction to **HPC Scientific Programming**: tools and techniques
- Introduction to Parallel Computing with **MPI** and **OpenMP**
- Introduction to Scientific and Technical Computing in **C**
- Introduction to Scientific and Technical Computing in **C++**
- Introduction to Scientific Programming using **GPGPU** and **CUDA**
- Introduction to the FERMI Blue Gene/Q, for users and developers
- **Parallel I/O** and management of large scientific data
- Programming paradigms for new hybrid architectures
- **Python** for computational science
- Tools and techniques for **massive data analysis**

23rd Summer School on PARALLEL COMPUTING

19 - 30 May, 2014 - BOLOGNA
14 - 25 July, 2014 - ROME
8 - 19 September, 2014 - MILAN

10th Advanced School on PARALLEL COMPUTING

February 10 - 14, 2014 - BOLOGNA

13th Summer School on SCIENTIFIC VISUALIZATION

YOU ARE HERE

9 - 13 June, 2014 - MILAN
13 - 17 October, 2014 - BOLOGNA

THE HARDWARE INFRASTRUCTURE

FERMI:
IBM BG/Q supercomputer composed of 10.240 PowerA2 sockets running at 1.6GHz, with 16 cores each, totaling **163.840 compute cores** and a system peak performance of **2.1 Pflops**. Is now number 12 in the **Top500**.

PLX:
IBM PLX composed of 274 compute nodes with 2 Nvidia GPU and 48GB per Compute node + 8 Fat node with 2 Nvidia GPU and 128 GB RAM per node.

EURORA (prototype):
is the world's most green supercomputer (hot water cooling system): 1st in the Green500 List of June 2013. 64 computing nodes (1024 cores) equipped with GPU.

SCHOOL AGENDA

Day1	Time	Title	Argument	Lecturers
13	9.30-10.15	School presentation and introduction to Sci Viz	**General**	Roberto Mucci
	10.15-11.15	Introduction to Python Language	**Python**	Alice Invernizzi
	11.15-11.30	coffee-Break		
	11.30-13.00	Tutorial	**Python**	Alice Invernizzi
	13.00-14.30	lunch-break		
	14.30-16.00	Basic tools for scientific visualization in Python	**Python in SciViz**	Alice Invernizzi
	16.00-17.00	Tutorial	**Python in SciViz**	Alice Invernizzi

Day2	Time	Title		Lecturers
14	9.30-10.30	Introduction to VTK	**VTK**	Stefano Perticoni - SCS
	10.30-11.15	Data Structures in VTK	**VTK**	Stefano Perticoni - SCS
	11.15-11.30	coffee-Break		
	11.30-13.00	Filtering and Rendering in VTK	**VTK**	Stefano Perticoni - SCS
	13.00-14.30	lunch-break		
	14.30-15.45	Hands-on: VTK exercises	**VTK**	Stefano Perticoni - SCS
	15.45-16.00	coffee-Break		
	16.00-17.00	Hands-on: VTK exercises	**VTK**	Stefano Perticoni - SCS

Day3	Time	Title		Lecturers
15	9.30-11.15	Introduction to GUI development using QT	**Qt GUI**	Andrea Negri - Paolo Quadrani
	11.15-11.30	coffee-Break		
	11.30-13.00	Tutorial	**Qt GUI**	Andrea Negri - Paolo Quadrani
	13.00-14.30	lunch-break		
	14.30-15.30	Introduction to Blender	**BLENDER**	Francesca Delli Ponti
	15.30-17.00	Tutorial	**BLENDER**	Francesca Delli Ponti

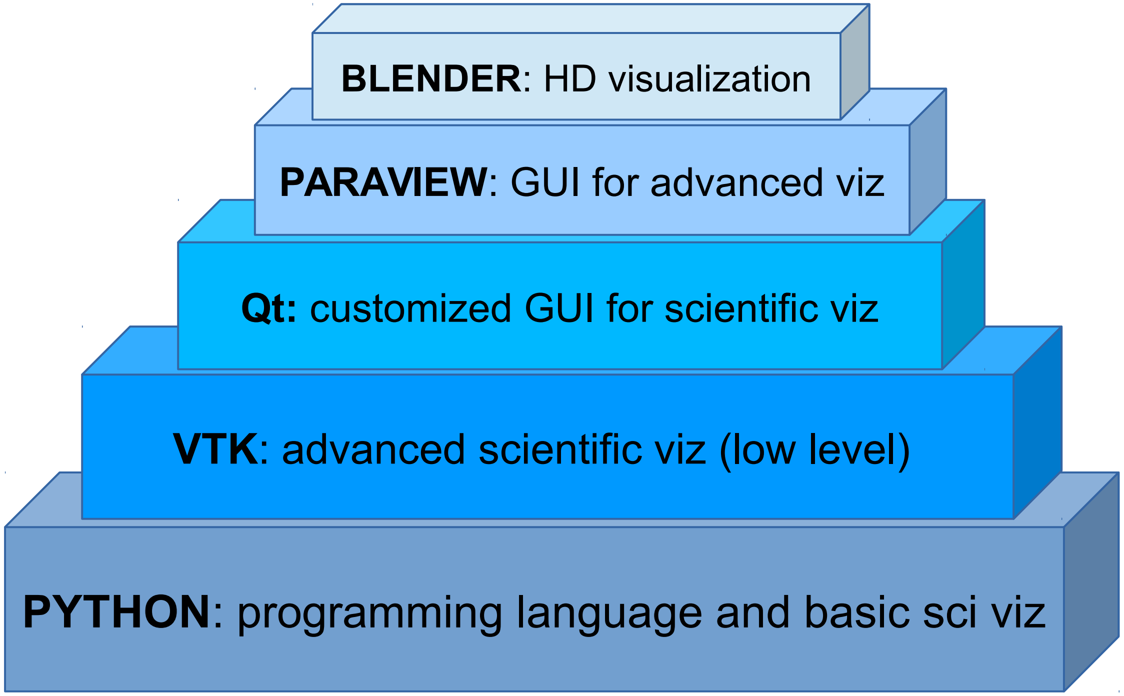
Day4	Time	Title		Lecturers
16	9.30-11.15	Introduction to Paraview GUI	**Paraview**	Ivan Spisso
	11.15-11.30	coffee-Break		
	11.30-13.00	Paraview scripting + 2 tutorial viz + scripting	**Paraview**	Ivan Spisso
	13.00-14.30	lunch-break		
	14.30-15.30	Paraview for large data viz + parallel pvserver	**Paraview**	I. Spisso
	15.30-17.00	Virtual Theatre demo	**Virtual Theatre**	Silvano Imboden

Day5	Time	Title		Lecturers
17	9.30-11.15	Paraview Summer HPC	**Paraview, BLENDER**	Massimilano Guarrasi
	11.15-11.30	coffee-Break		
	11.30-13.00	Tutorial Paraview + Blender	**Paraview, BLENDER**	Massimilano Guarrasi
	13.00-14.30	lunch-break		

BOTTOM-UP APPROCH

- From programming language (python) to high level GUI (Paraview, Blender)

- **Libraries and tools**:
  - Open-source
  - Cross-platform
  - Well documented
  - Python based



PYTHON

Why python?

- High level language: **easy** syntax, readable
- **Rich built-in library** (string, data type, numeric and math modules, File system access, cryptography, GUI, debugging and profiling... https://docs.python.org/2/library/)
- Can be extended with different **scientific library**
- Many application offers Python **scripting**

OUTLINE

- Introduction to **Python** (theory + practice)
- Introduction to **scientific visualization with Python**:
  - Introduction to **Numpy** array
  - Introduction to **matplotlib** module
  - Brief introduction to **Mayavi** and **mlab** module
- Demo on a real python application
- Practice (matplotlib and mayavi)

VTK

The Visualization Toolkit (VTK) is an open-source, freely available software system for 3D computer graphics, image processing and visualization.
The state-of-the-art scientific visualization library used among different application fields: medicine, CFD, astrophysics, geology.. http://www.vtk.org/VTK/resources/applications.html

OUTLINE

- Introduction: what is VTK?
- Data Structures: how is information represented in VTK?
- Filtering: how is data processed?
- Rendering: how is data visualized?
- Exercises: live coding with VTK

QT

Qt is a cross-platform application framework that is widely used for developing application software with a graphical user interface (GUI)

OUTLINE

- What is Qt?
- Introduction to PyQt
- QObjects, Connections, Event Handling
- GUI creation
- Hands-on and examples

PARAVIEW

ParaView is an open-source, multi-platform data analysis and visualization GUI application based on VTK for the visualization of 2D/3D/4D data.
ParaView is flexible enough to work with data from many areas of computational science: can read well over one hundred different file formats which cover a wide range of application domains.

The most common VTK readers, filters, renderer, viewer and writer accessible thorough a user interface.

OUTLINE

- Paraview GUI
- Paraview filters
- Scripting with Paraview
- Paraview for large date visualzation
- Hands-on

BLENDER

"Blender is a free and open source 3D animation suite. It supports the entirety of the 3D pipeline—modeling, rigging, animation, simulation, rendering, compositing and motion tracking, even video editing and game creation. […] Python scripting allows to customize the application and write specialized tools" (from http://www.blender.org/about/)

- **Photorealistic Rendering**
- **Fast Modelling**
- **Video Editing**

OUTLINE

- Introduction to Blender
- Modeling with Blender
- Hands-on:
    - User inteface and tools
    - 3D objects
    - Materials management
    - Camera animation
    - Rendering
    - Video Sequence Editor

OUTLINE

- School presentation
- **Introduction to Scientific Visualization**
- Remote visualization @ Cineca

WHAT IS SCIENTIFIC VISUALIZATION?

Scientific visualization is the practice of **producing graphics representations of scientific phenomena**. The primary goal of visualization is **insight**: to improve understanding of the data through their visual representation.

“**A picture is worth a thousand ~~words~~ numbers**”: complex idea can be conveyed with just a single still image.

1-TO ANALYSE AND EXPLORE: through the visualization it is easier to get pattern, regularity and associations. It takes less time to understand phenomena.

2- TO PRESENT AND COMUNICATE: through the visualization information can be provided briefly and efficiently

A slogan:

“Discover the unexpected, describe and explain the expected”

A TRIVIAL EXAMPLE... (1)

How many 3 are present in the picture?
(max 5 seconds...)

```
8973905709279405796297650982 94
0802808508083080280985 0-802808
567847298872t 4582020947577200
567847298872ty4582020947577200
21789843890r4557904560992721 88
8975947979028558925945 73979209
```

A TRIVIAL EXAMPLE... (2)

...try now!

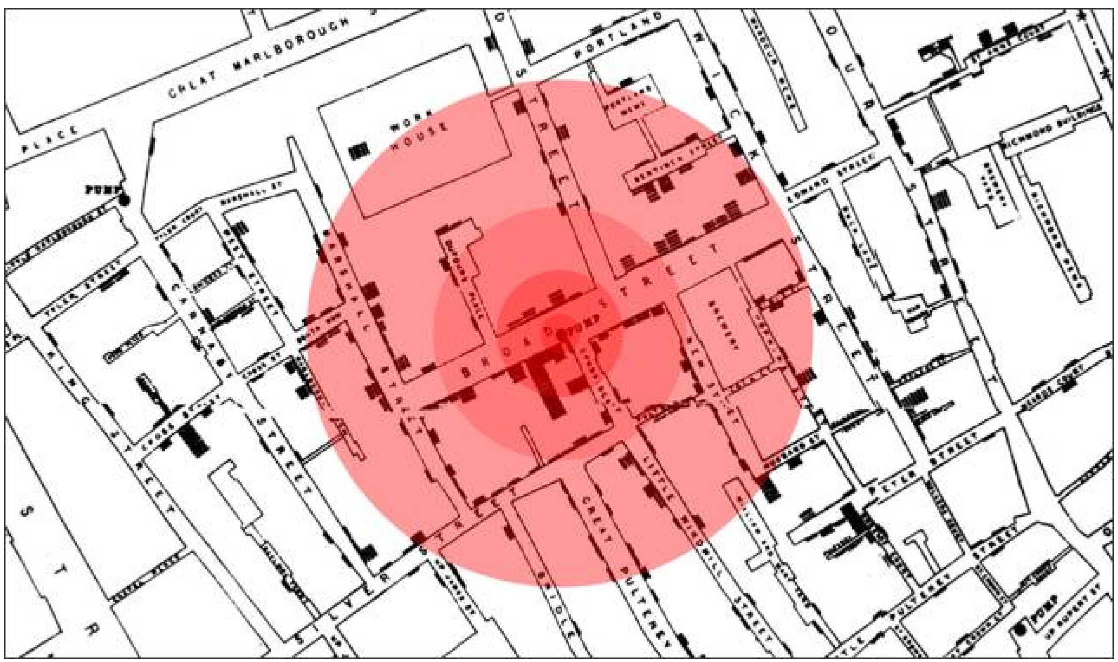
897**3**9057092794057962976509829 4
0802808508083**3**0802809850-802808
567847298872t 4582020947577200
567847298872ty4582020947577200
217898**4**3890r455790456099272188
897594797902855892594573979209

..A CONCRETE ONE

John Snow's dot distribution map showing cholera deaths in London in 1854.

1-TO ANALYSE AND EXPLORE: through the visualization of the distribution of dots on the map he identified the infected public water pump.

2- TO PRESENT AND COMUNICATE: through the map he persuaded the local council to disable the pump by removing its handle.



OBJECTIVES OF VISUALIZATON

- To **analyze** data and information
- To improve **comprehension** of phenomena and processes
- To find **new meanings** and interpretations
- To **make visible the invisible**
- To **check quality** of simulations and measures
- To **make effective presentation** of information and results

VISUALIZATION IN HPC

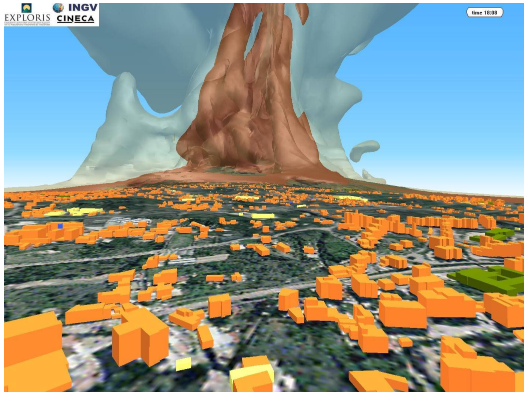
An HPC center can be seen as a **virtual laboratory** which allows users to **simulate** or **investigate** very complex phenomena, events or processes:

- **Safely** (not to cause damage, injury, or harm)
- In a **cheap** way (tests can be repeated many times)
- Too **big** to be studied in reality
- Playing with the **time** (if time variant)

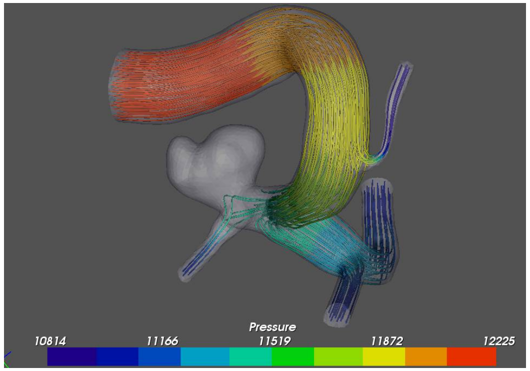
The results of the simulations is a **huge quantity on numbers** impossible to be interpreted if not visualized with different visualization techniques.

VISUALIZATION IN HPC: examples

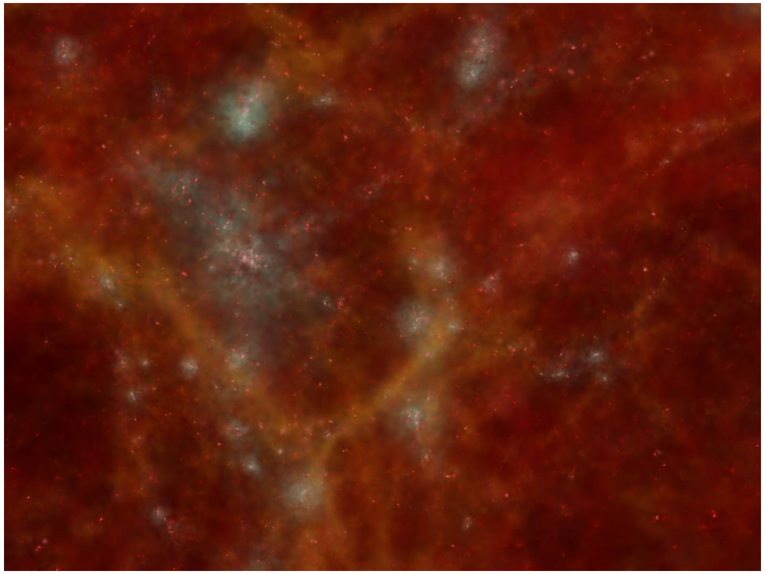
Safely: simulate the effects of the eruption of a volcano.
EXPLORIS:
http://www.cineca.it/it/progetti/exploris



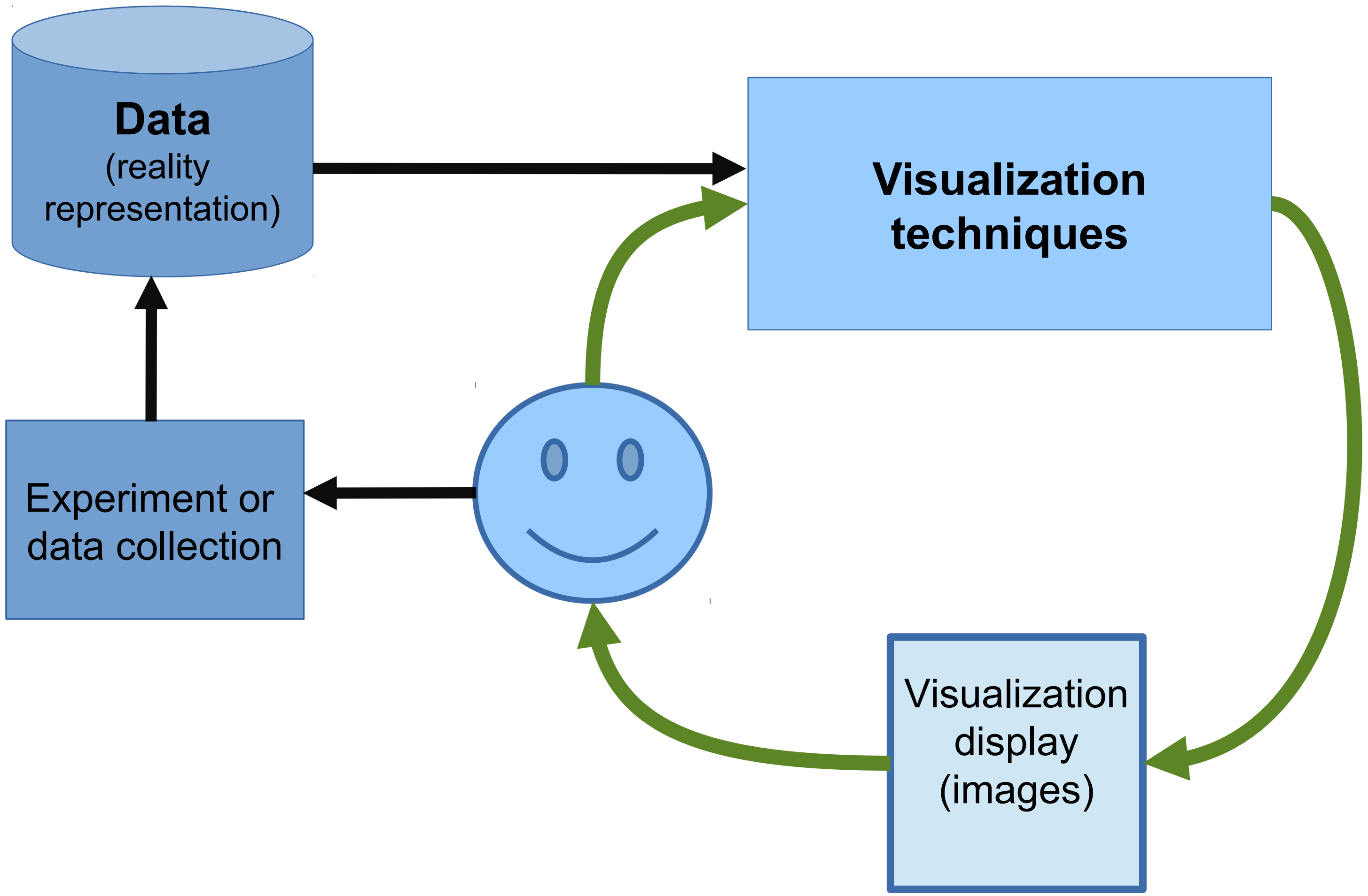
Safely: predictive simulations on cerebral aneurysm.
ANEURIST:
http://www.aneurist.org/



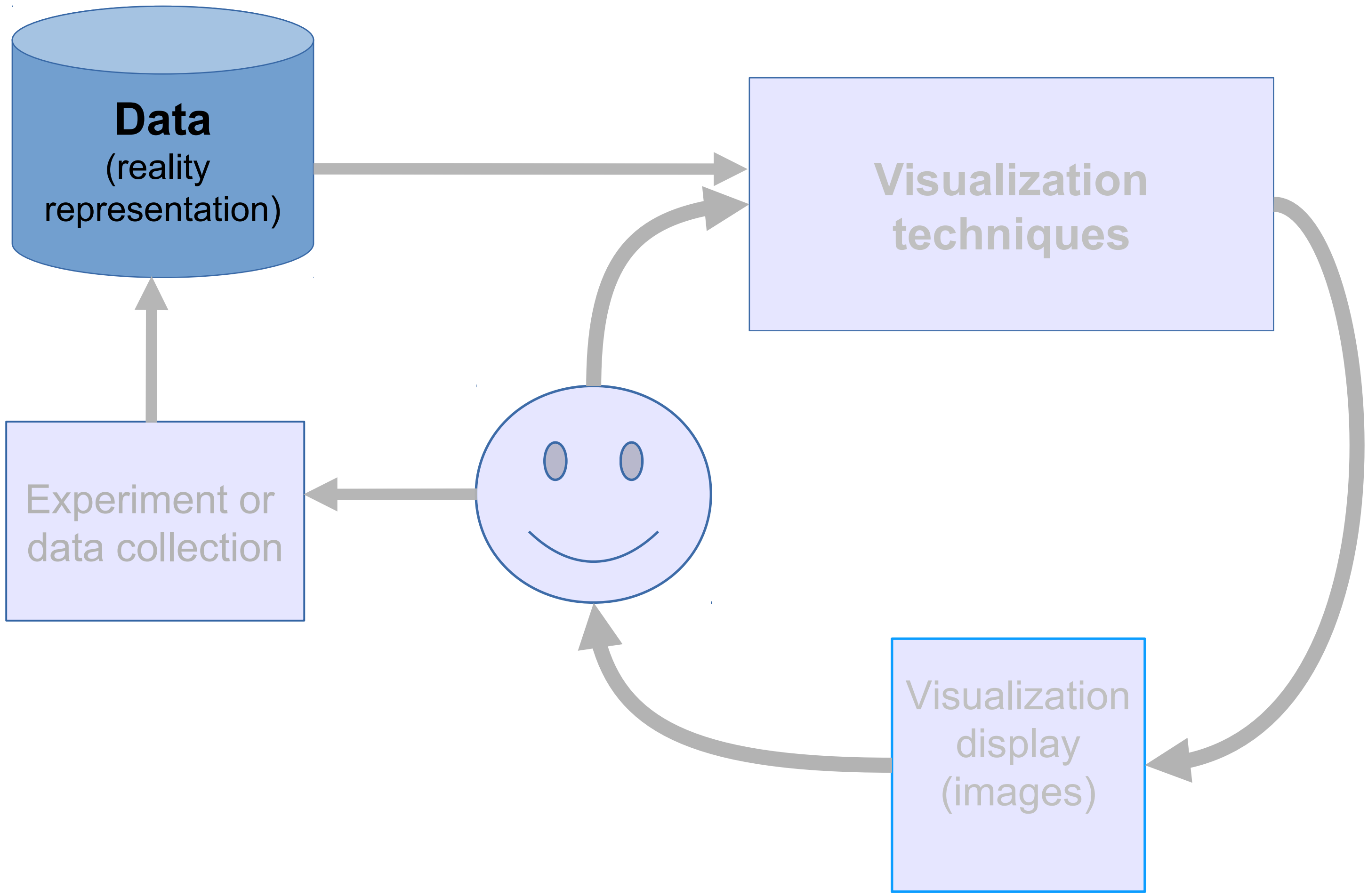
Too big: Big Bang simulation.
BIG BANG:
http://www.cineca.it/it/progetti/big-bang-documentario



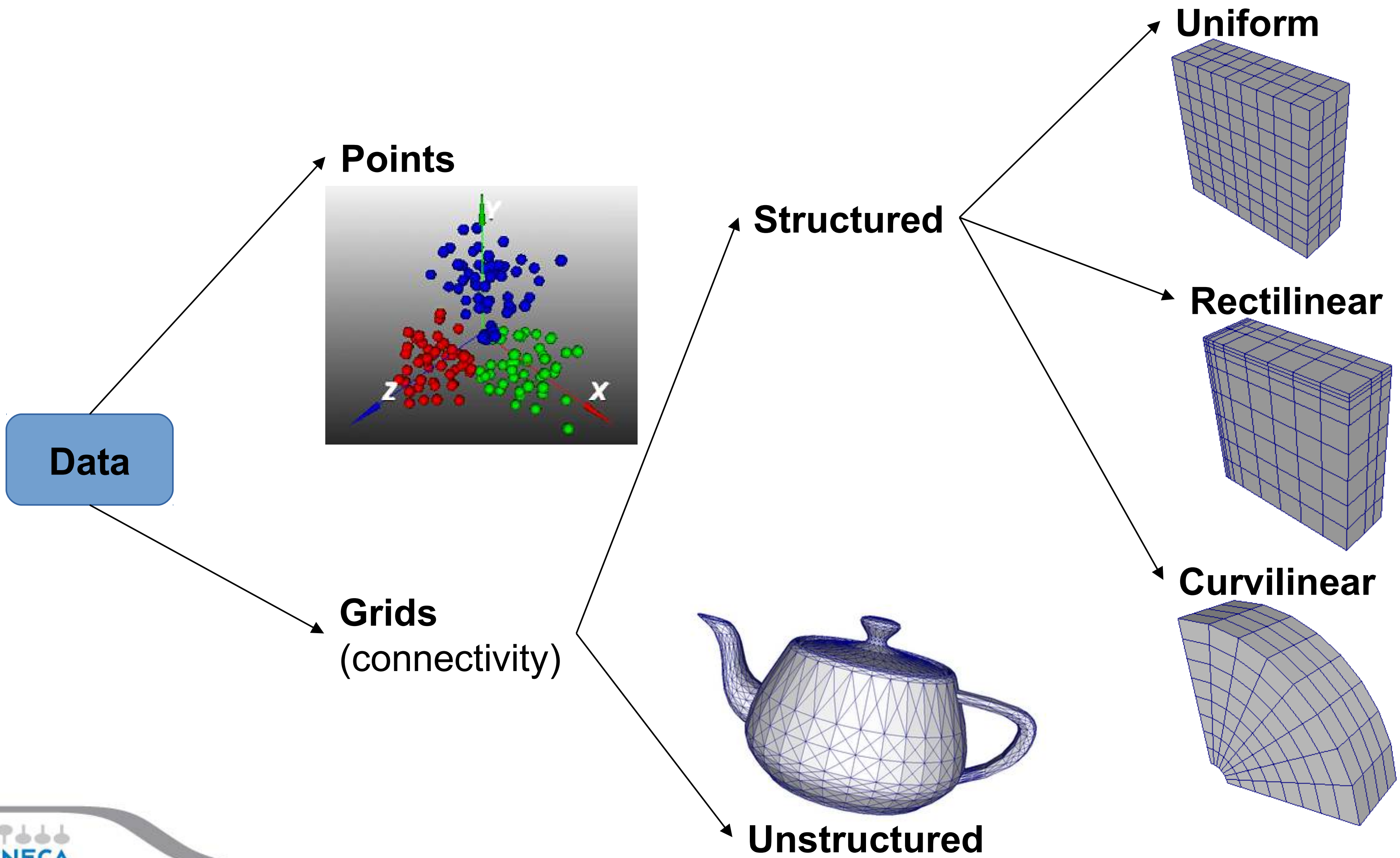
A VISUALIZATION MODEL



A VISUALIZATION MODEL

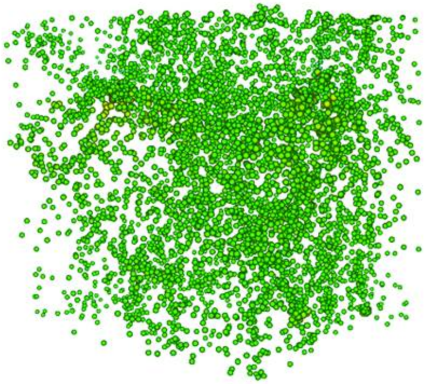


DATA: geometrical structures

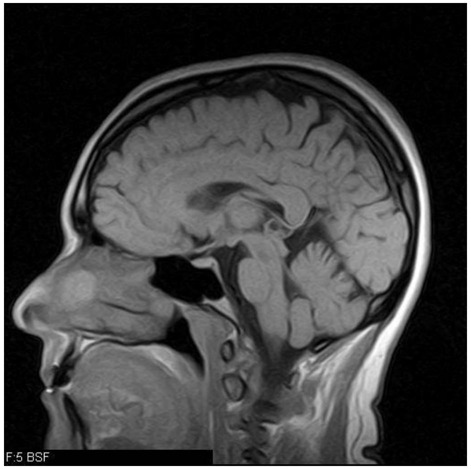


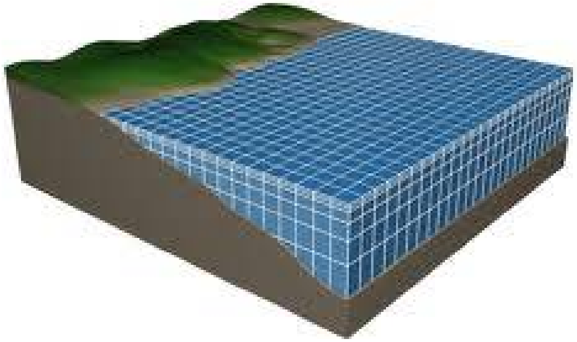
DATA TYPES: examples

Points: Particle simulations

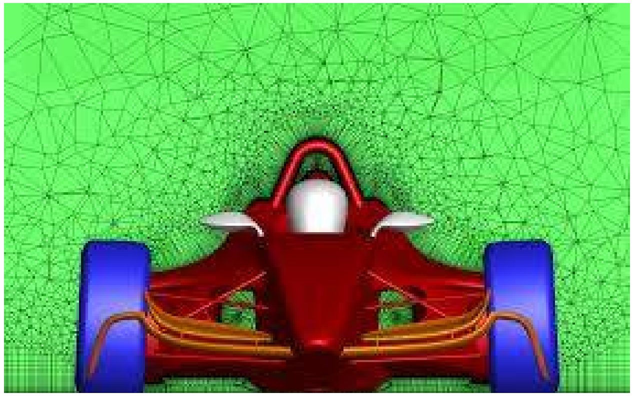


Rectilinear grid: medical, oceanographic

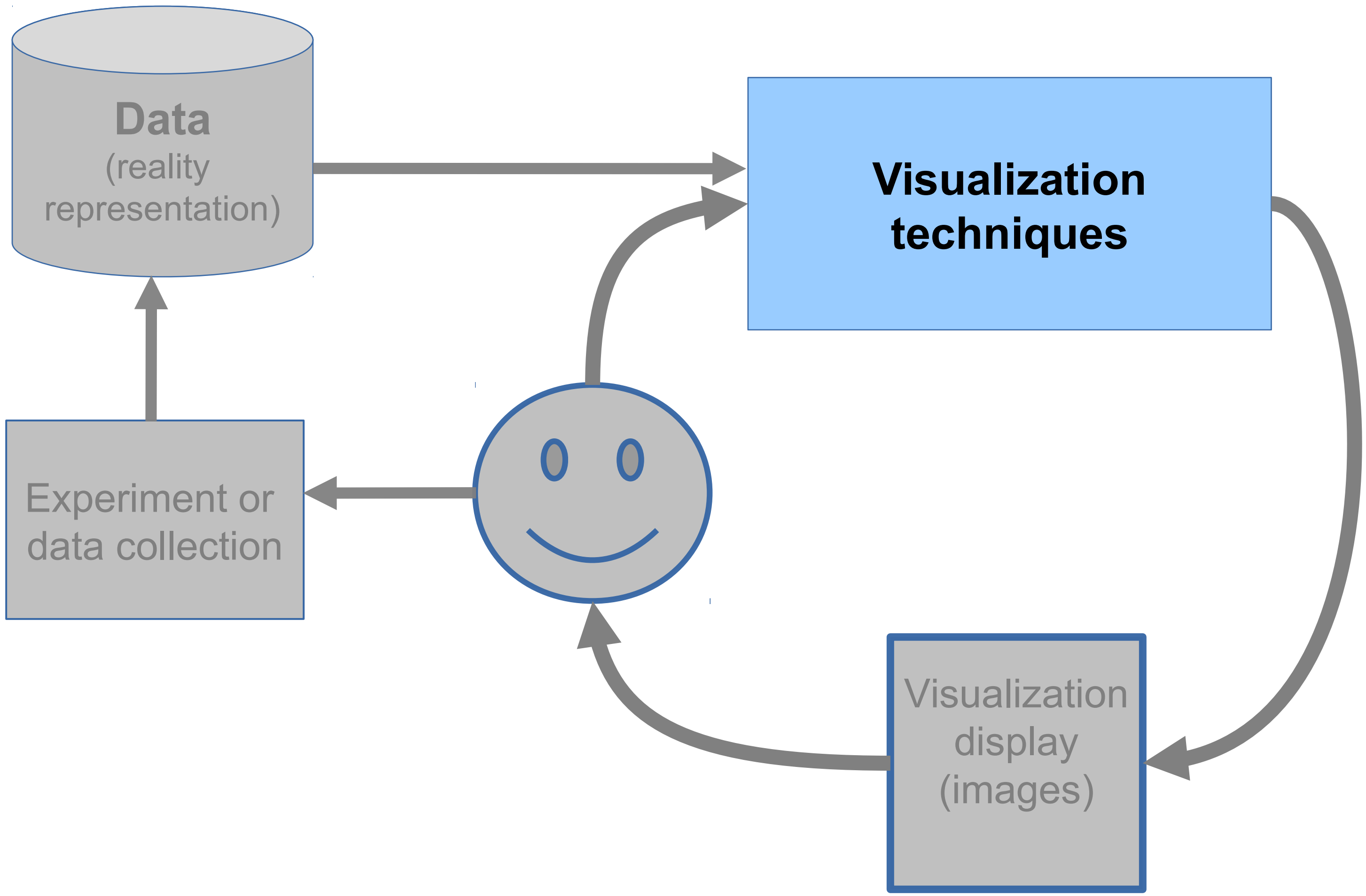




Unstructured grid: CFD, aerodynamics



A VISUALIZATION MODEL

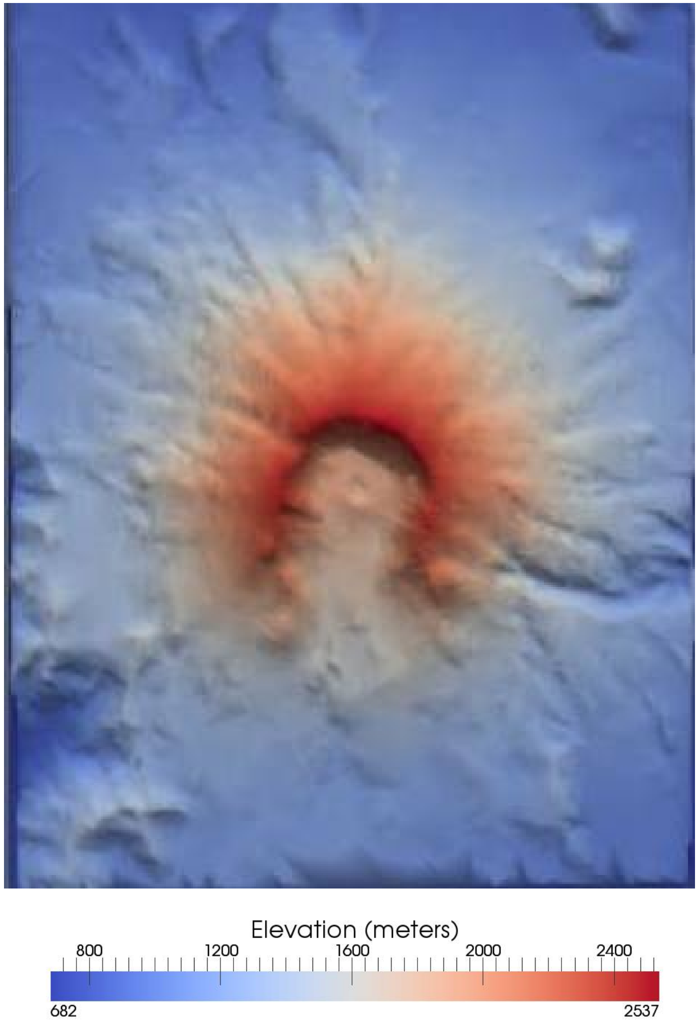


PSEUDOCOLOR MAPPING

Data type: SCALAR
Domain: 2D, 3D

Process:
- Map scalar data to a color table (colormap)

Utility:
- To investigate range of data (temperature, pressure, elevation..)
- Fast and great for Error diagnostic and Visual Validation

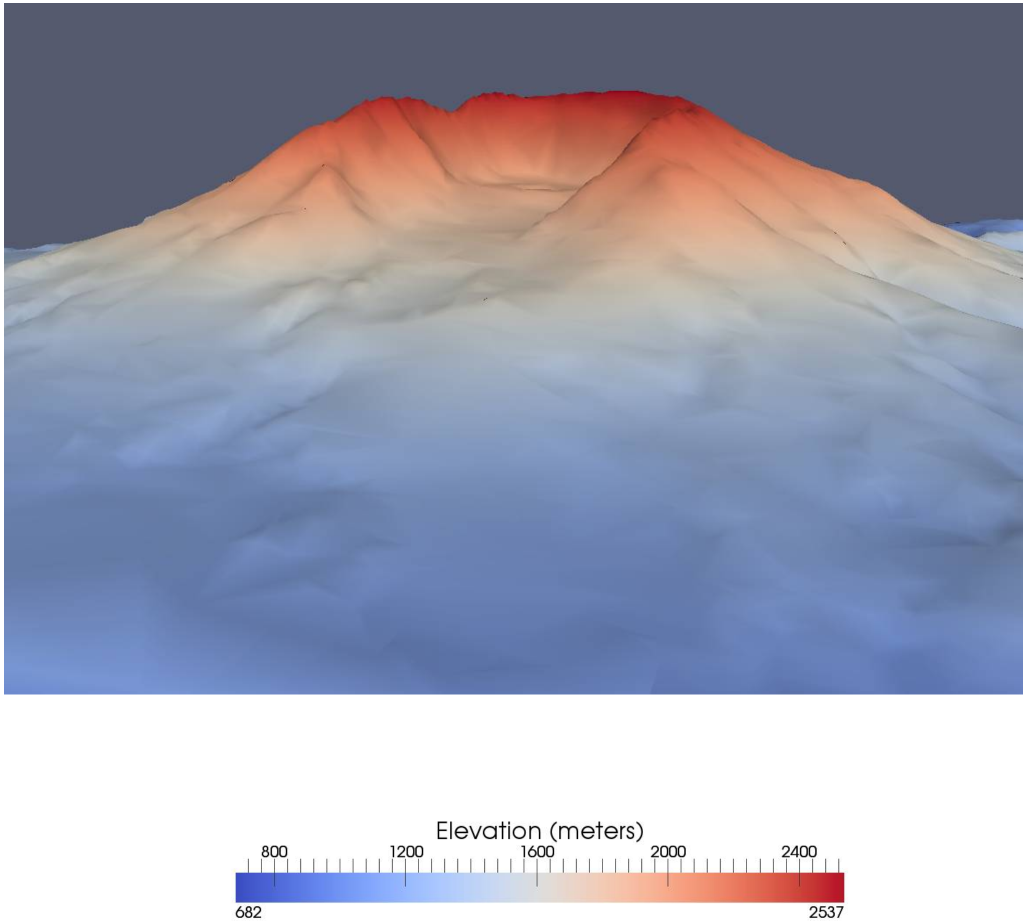


SURFACE VIEW

Data type: SCALAR
Domain: 2D

Process:
• Scalar values are used as Z component (height)

Utility:
- 2D representation becomes 3D
- Good for geographic data
- To quickly understand the different intensity of the scalar values

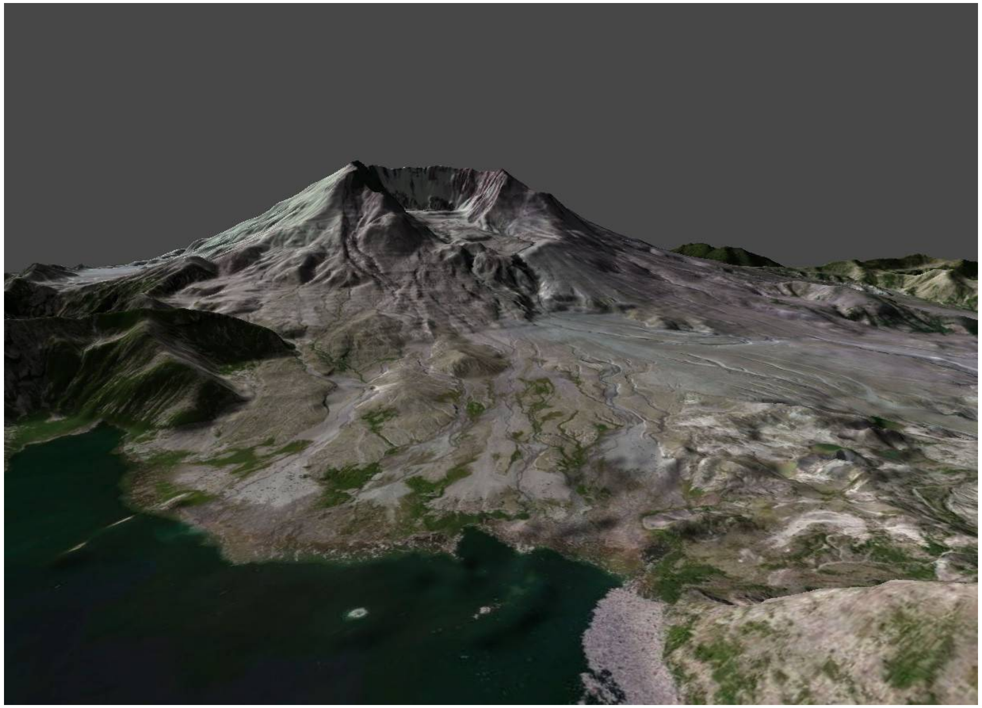


TEXTURE MAPPING

Data type: SCALAR
Domain: 2D, 3D

Process:
• Apply a 2D image on a surface specifying the correspondence among some points of the image and some points of the surface

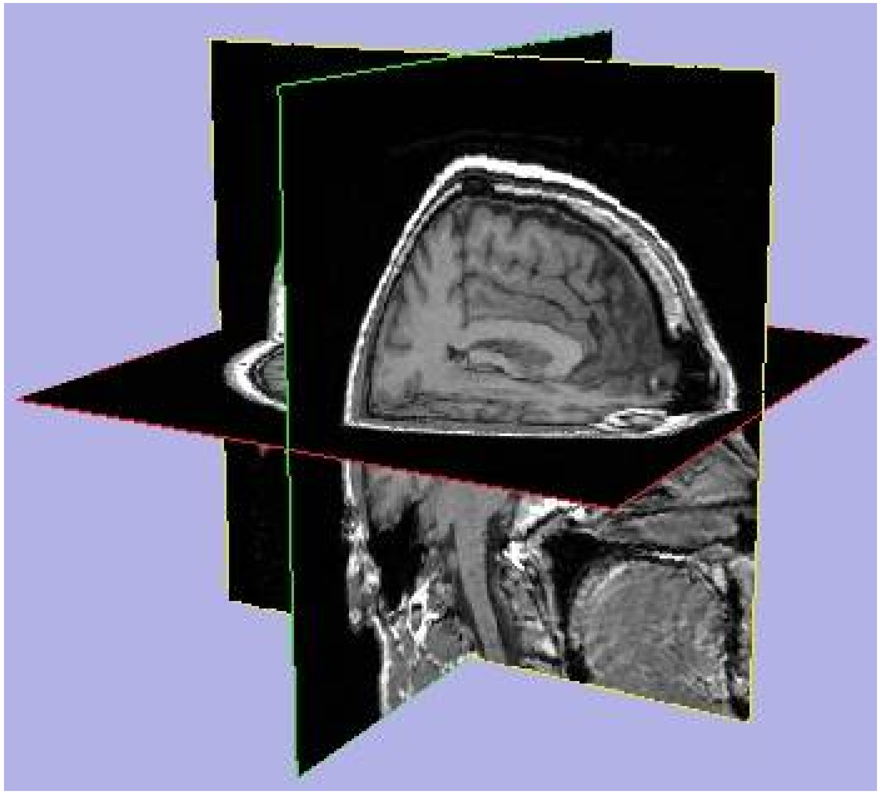
Utility:
- Contextualize the visualization
- Give details
- Realistic visualization



SLICING

Data type: SCALAR
Domain: 3D

Process:
• Define a cutting surface that cuts the 3D data: the intersection of the plane with the data is visualized in 2D.



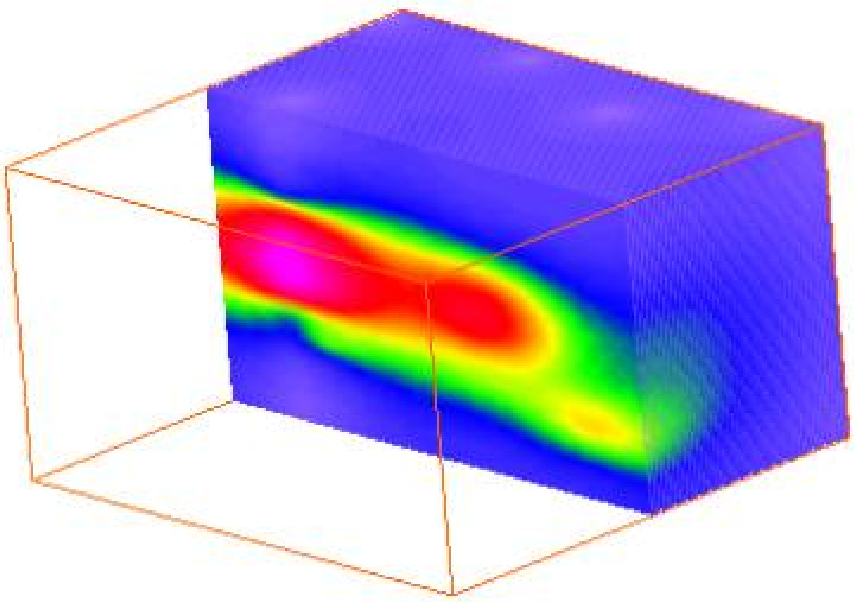
Utility:
- Investigate the scalar values inside the volume
- Give an inner view of a 3D object

CROPPING (CLIPPING)

Data type: SCALAR
Domain: 2D, 3D

Process:
• Define a cutting surface that cuts the 3D data: returns everything inside the cutting plane

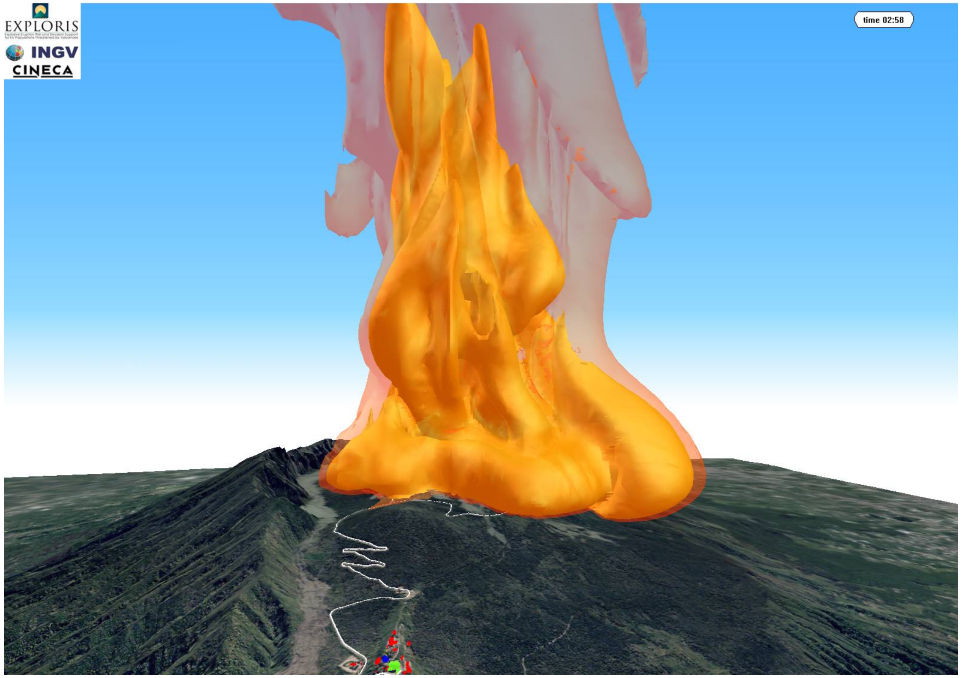
Utility:
• Remove part of the dataset



ISOSURFACE (ISOLINE)

Data type: SCALAR
Domain: 2D, 3D

Process:
• A surface that represents points of a constant value (e.g. pressure, temperature, velocity, density) within a scalar volume. It is a line in a 2D domain (isoline).



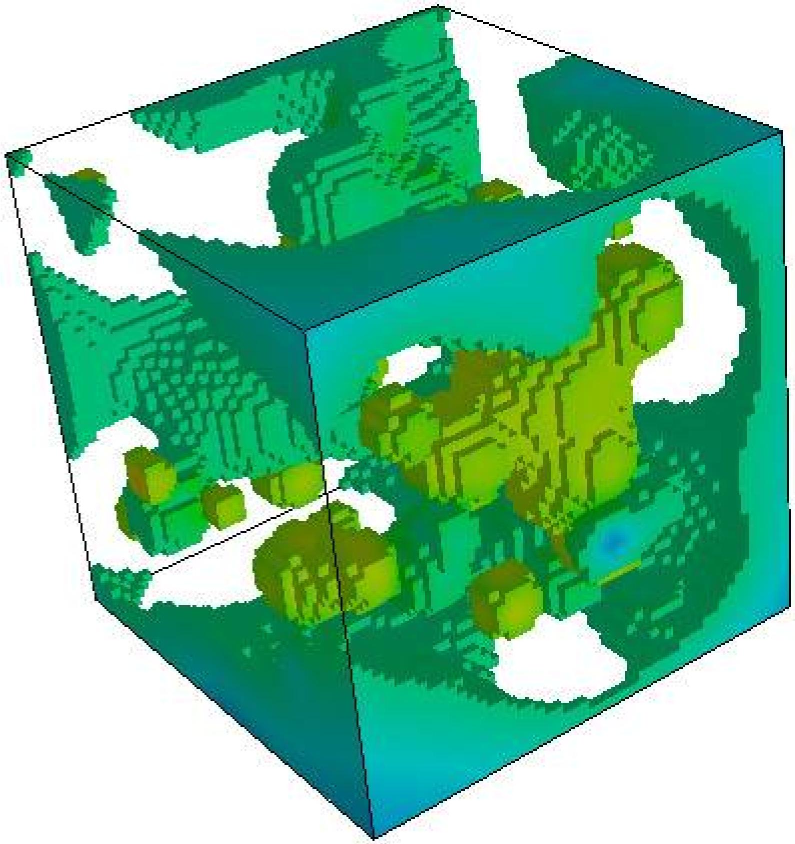
Utility:
- Identify how scalars with constant value are distributed (temperature, pressure..)

THRESHOLD

Data type: SCALAR
Domain: 2D, 3D

Process:
Visualize only scalar values higher (lower) of a defined value, or inside an interval of values

Utility:
- Data filtering
- Emphasize part of the data
- Remove unused data



VOLUME RENDERING

Data type: SCALAR
Domain: 3D

Process:
- Volume rendering does not use intermediate surface representations
- Computing 2-D projections of a colored semitransparent volume (typically a 3D scalar field). Need to define the opacity and color of every voxel.(jellyfish effect)

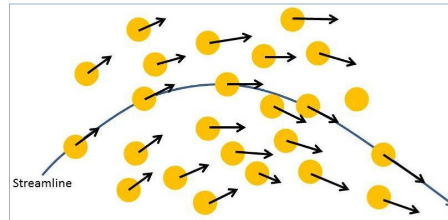
Utility:
- Look at the 3D data set as a whole
- Investigate interior/density of scalar volumetric data

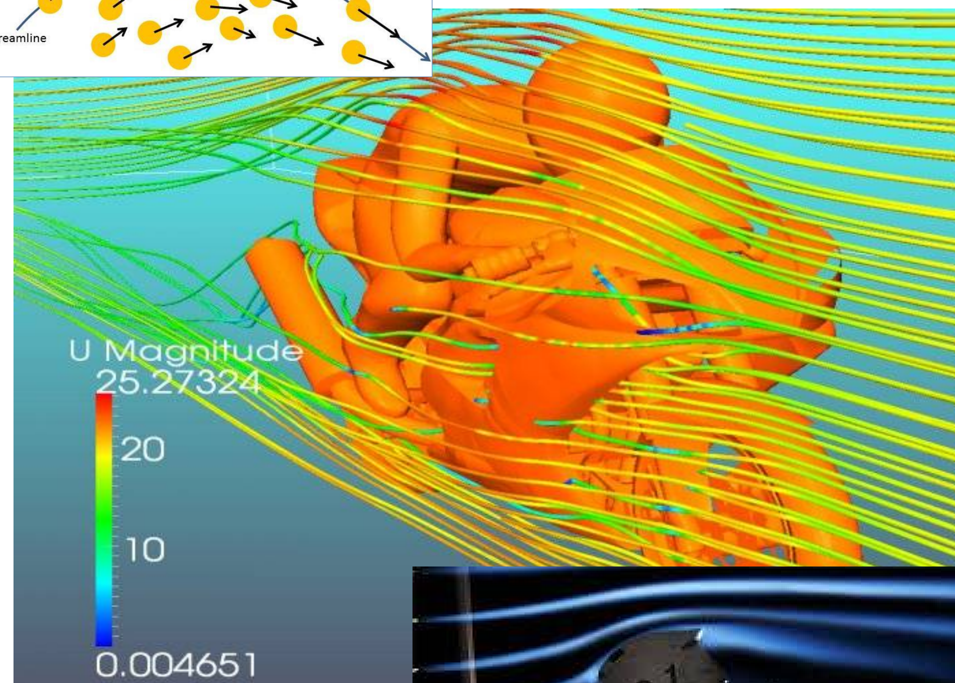
STREAMLINES

Data type: VECTORIAL
Domain: 2D, 3D

Process:
- A streamline is a path traced out by a massless particle as it moves with the flow .
- Velocity is tangent to streamline at every point

Utility:
- Investigate nature of flow (fluid/aero dynamics)

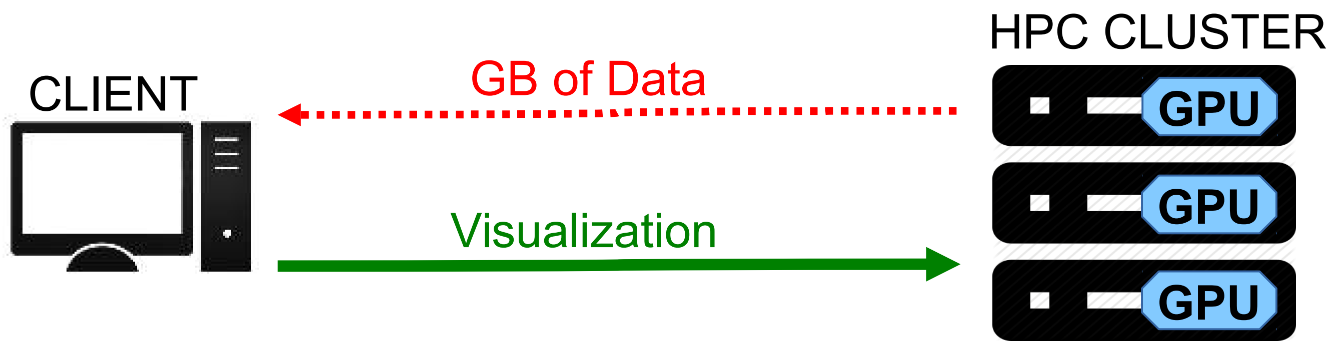




OUTLINE

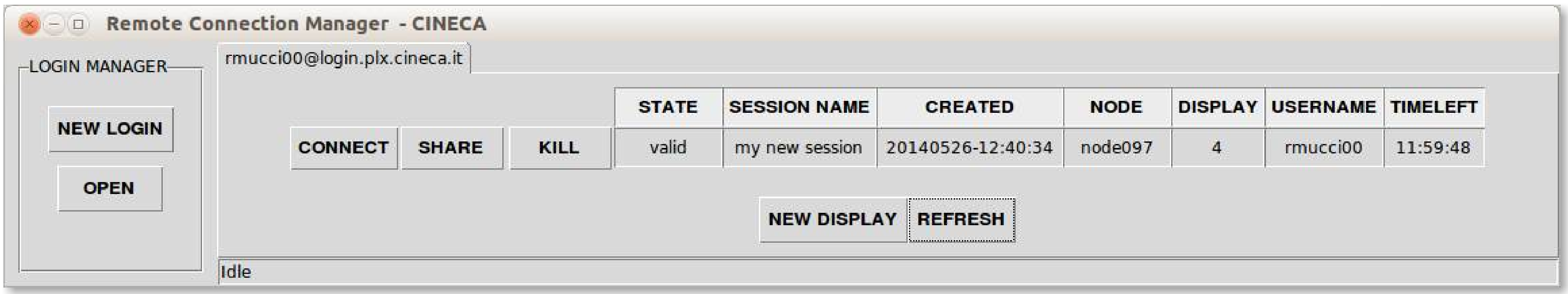
- School presentation
- Introduction to Scientific Visualization
- **Remote visualization @ Cineca**

REMOTE VIZ OBJECTIVE



- Allow users to **performe visualization** e post-processing activities on **HPC machines** (GPU)
- **Avoid transferring** of GB of data produced on Cineca HPC systems
- **Simplify** operations to create and manage remote displays
- Give the possibility to **share** the same display among different users

RCM: Remote Connection Manager



- Cross platform client/server **GUI application**
- Automates operations to **setting up a remote connection** to the Cineca clusters
- **Simplify the management** of the remote dislpays
- Integration of existing open-source technologies: **TurboVNC** and **VirtualGL**

RCM: Remote Connection Manager

TurboVNC

Free remote control software package that support VirtualGL.
TurboVNC performs very well on high-latency, low-bandwidth networks. More info at http://www.turbovnc.org/

VirtualGL

Open source toolkit that gives any Unix or Linux remote display software the ability to **run OpenGL applications with full 3D hardware acceleration**. It optimizes user experience of remote 3D applications by rendering on remote GPU while streaming only the 2D result images.
More info at http://virtualgl.org/

REMOTE VIZ INFRASTRUCTURE

- **PLX GPU cluster**
  - **2 RVN nodes**:
    - 2 CPU IBM X5570 Intel(R) Xeon(R), 12 Core
    - 128 GB of RAM
    - 2 NVIDIA Tesla M2070
  - **1 Big mem node**:
    - 512 GB of RAM
    - 1 GPU NVIDIA Quadro 6000

- **PICO cluster**
  - **2 Viz nodes** each with:
    - 2 CPU Intel XEON E5 2670 v2, 10 core, @2.5Ghz
    - 128 GB RAM (8*16GB DDR3@1866Mhz)
    - 2 GPU NVIDIA K40
  - **2 Fat nodes** each with:
    - 2 CPU Intel XEON E5 2650 v2, 8 core, @2.6 GHz
    - 128GB RAM (16*8GB DDR3@1600 Mhz)
    - 1 GPU NVIDIA K20

PRE-INSTALLED SOFTWARE

The open-source softwares (no license needed to use them) are reported in bold, the other softwares are under license
To use them, refer to your project academic/industrial or ask the support (superc@cineca.it)

- Abaqus
- Ansys (Mechanical and Fluent)
- Comsol
- hyperstudy
- **Paraview** version 3.14, 3.98, 4.0.1, 4.1
- Pointwise version 17.0
- Tecplot version 2012R1
- **Vaa3D**
- **vmd**
- **Blender**

RCM user documentation: http://www.hpc.cineca.it/content/remote-visualization

PRACTICAL INFORMATION

- School stuff repository: https://hpc-forge.cineca.it/files/Visualization_School/public/2014/BOLOGNA/
- PC password: **corsi_2013!** (Generally windows 7 will be used)
- Lunch is provided by Cineca (restaurant is about 400 m from here).
- Each student will be given a two week access to the Cineca's supercomputing resources.

AND THEN…

ENJOY THE LESSONS!!