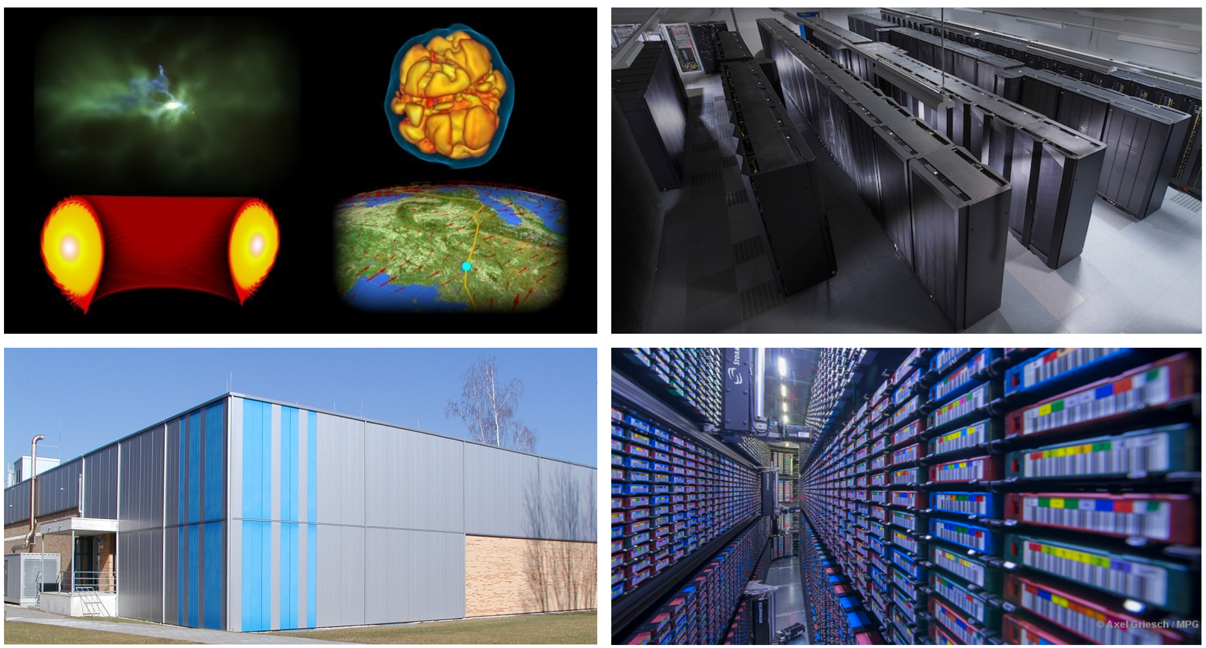
Scientific Visualization and Beyond

Services and Projects from the Max Planck Society

Klaus Reuter
Max Planck Computing and Data Facility

CECAM Workshop on *Emerging Technologies in Scientific Data Visualisation*

Pisa, April 5 2018



http://www.mpcdf.mpg.de

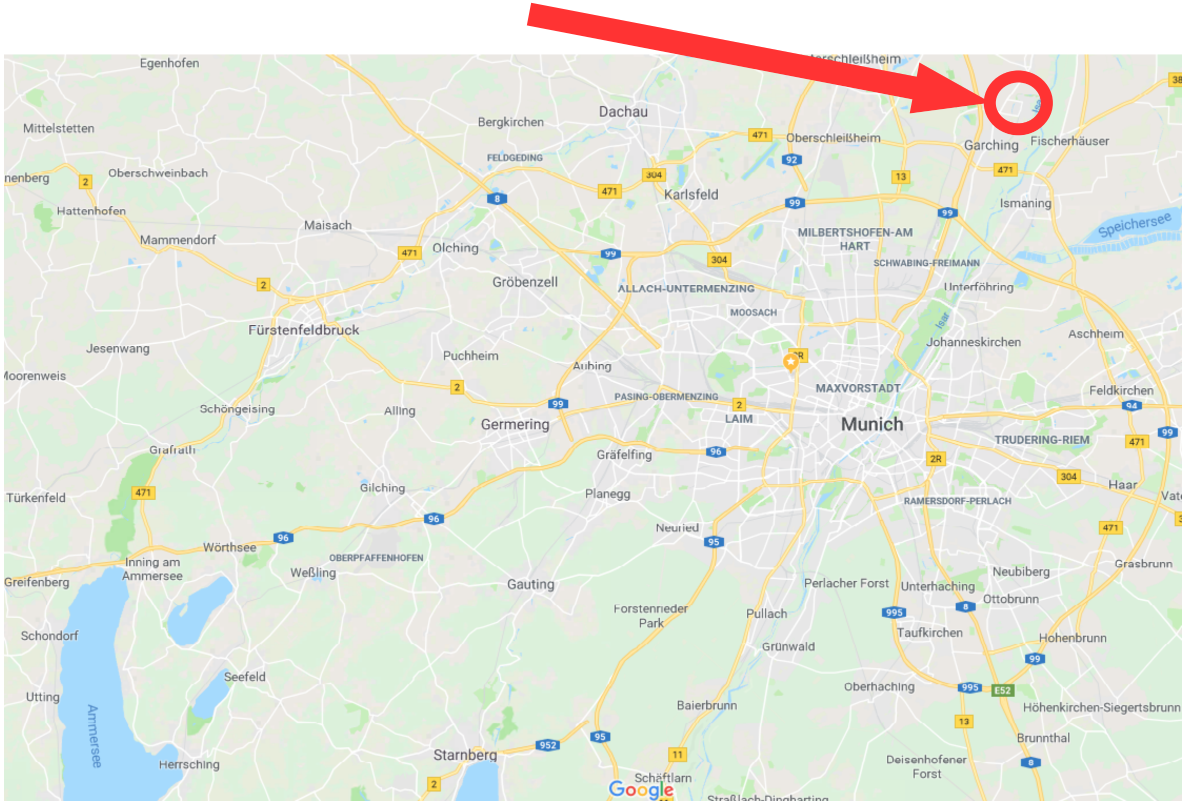
Outline

- Introduction
- Visualization Services in classical HPC environments (MPCDF)
- Virtual reality applied to large-scale turbulence simulations (MPI for Dynamics and Self-Organization, TUM, LRZ, MPCDF)
- NOMAD COE - Advanced Graphics
  - Remote Visualization Services (FHI, MPCDF)
  - High-quality visualizations (MPCDF)
- pyiron – A Python-based IDE for computational materials design (MPI for Iron Research, supported by MPCDF)
- Summary

About MPCDF

About the Max Planck Computing and Data Facility (MPCDF)

- HPC and data center of the Max Planck Society
- MPCDF (formerly known as RZG), located at the research campus in Garching near Munich, Germany



Mission of the MPCDF

Computational Application Support

- Support for the development and optimization of parallel applications for all disciplines and all institutes of the MPG

Computing and Data Services

- Managing the central HPC system(s),
  used by 45 institutes and about 55 departments
- Hosting numerous Linux clusters (50 clusters from 20 MPIs)
- Long-term data storage

Data Application Support

- Managing a broad stack of data-repository technologies
- Hosting various data repositories of the MPG
- Providing different data service layers for large-scale experiments

EU and national projects

- NOMAD CoE, EUDAT, RDA, ELPA-AEO, ...

MPCDF Organization

Management: Stefan Heinzel, Beirat

Groups:

- **data services** (*Raphael Ritz*)
- **systems and operations** (*Christian Guggenberger*)
- **application support** (*Markus Rampp*)

Cesar Allande Alvares (Computer science)
Giuseppe di Bernardo (Comp. Physics)*
Berenger Bramas (Computer science)
Michele Compostella (Comp. Physics)
Tilman Dannert (Comp. Physics)
Renate Dohmen (Comp. Physics)*
Lorenz Hüdepohl (Comp. Physics)*
Pavel Kus (Appl. Mathematics)
Rafael Lago (Computer science)
Andreas Marek (Comp. Physics)
Florian Merz (Comp. Physics), Lenovo
Werner Nagel (Comp. Physics)
Sebastian Ohlmann (Comp. Physics)
Markus Rampp (Comp. Physics)
Klaus Reuter (Comp. Physics)
Luka Stanisic (Computer science)
(*) shared with other groups

MPCDF Application Services

Basic application support

- software support, compilers, libraries, tools, ...

Optimization and development of HPC applications (in close collaboration with scientists)

- materials and bio sciences, astrophysics, plasma physics, ...
- all relevant HPC architectures (not limited to hardware operated by the MPCDF)

New HPC architectures and programming models

- assessment of hardware and software (e.g. CUDA, OpenACC, MPI-3, Coarray-FORTRAN, ...)
- porting of applications to GPUs, Intel Xeon Phi, …, ARM, …, ?

Scientific visualization

- operation of remote-visualization infrastructure (hardware, software)
- project support

High-level data application support (Group of R. Ritz)

- high-level data application support: repository services, custom web development for MPG and partners

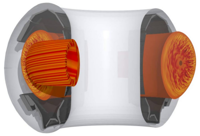
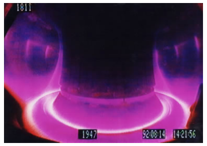
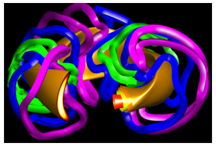
Training and consulting

- lectures, courses, tutorials (mostly MPG-internal)

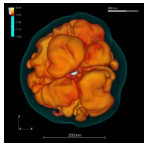
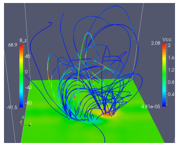
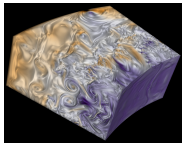
Application Projects

HPC code projects (examples)

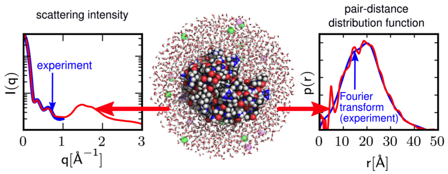
- original contributions and long-term support for development, optimization and porting of HPC codes in the Max Planck Society, e.g.
  - FHI-aims, OCTOPUS, S/PHI/nX, ESPResSo++ (materials and bio science), ELPA (eigensolver library)
  - GENE, SeLaLib/VLASOV6D, SOLPS, GPEC/IDE, VMEC (fusion research)

  - VERTEX, GOEMHD3, NSCOUETTE, MagIC, BFPS (astrophysics, comp. fluid dynamics)

  - BioEM, CAPRIQORN/CADISHI, COMPLEXES++, ... (biophysics)



Outline

- Introduction
- **Visualization Services in classical HPC environments (MPCDF)**
- Virtual reality applied to large-scale turbulence simulations (MPI for Dynamics and Self-Organization, TUM, LRZ, MPCDF)
- NOMAD COE - Advanced Graphics
  - Remote Visualization Services (FHI, MPCDF)
  - High-quality visualizations (MPCDF)
- pyiron – A Python-based IDE for computational materials design (MPI for Iron Research, supported by MPCDF)
- Summary

Scientific Visualization Services

Background: scientific visualization at the MPCDF

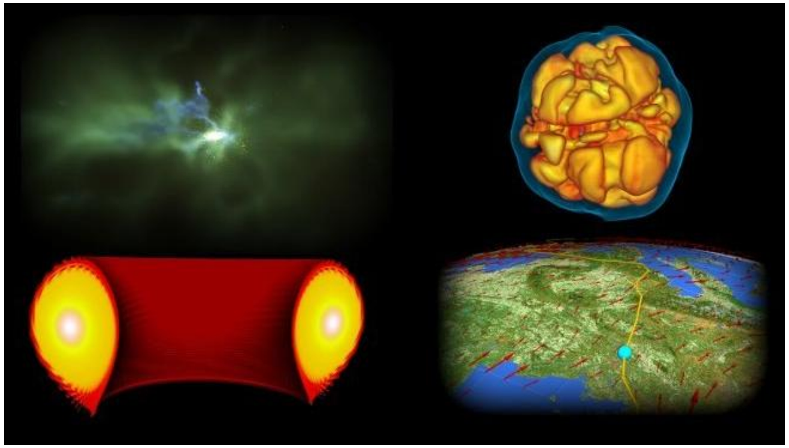
- a central visualization infrastructure and project support for the Max-Planck Society (since 2008)

Main targets: interactive, remote data exploration & analysis, presentation, publishing

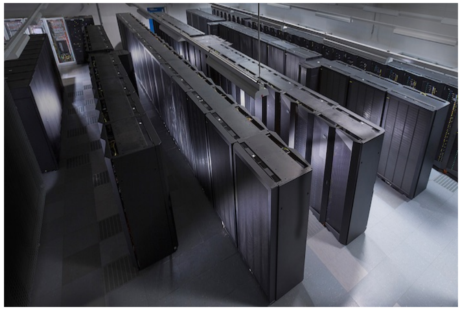
- support for adaptation and instrumentation of simulation codes
- guidance for selection, adoption and usage of analysis & visualization software
- dedicated support for individual (particularly demanding) visualization projects

Challenges

- broad range of disciplines: plasmaphysics, astrophysics, . . . , biology
- variety of simulation codes: ”home-grown”, commercial, open-source, third-party, . . .
- non-standardized, heterogeneous data structures and formats, ”legacy” analysis pipelines, . . .
- massive datasets from HPC simulations:
- massive: amount of raw data, memory requirements, complexity
- multidimensional (3D + time), multi-variate data
- ”unusual” grids: mesh-free data, special curvilinear coordinates, . . .



HPC Visualization Infrastructure



Remote visualization directly on the HPC system

why centralizing visualization ?

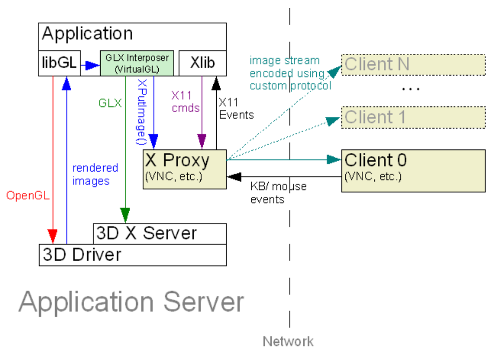
- huge amounts of output data produced by HPC simulations
- transfer of raw data for local analysis & visualisation no more possible
- even dumping the RAM is becoming prohibitive due to I/O constraints
  → in-situ visualisation
- visualisation requires HPC-like resources (specialized hardware, housing, . . .)
- requires substantial expertise on methods, software, . . ., sustainability

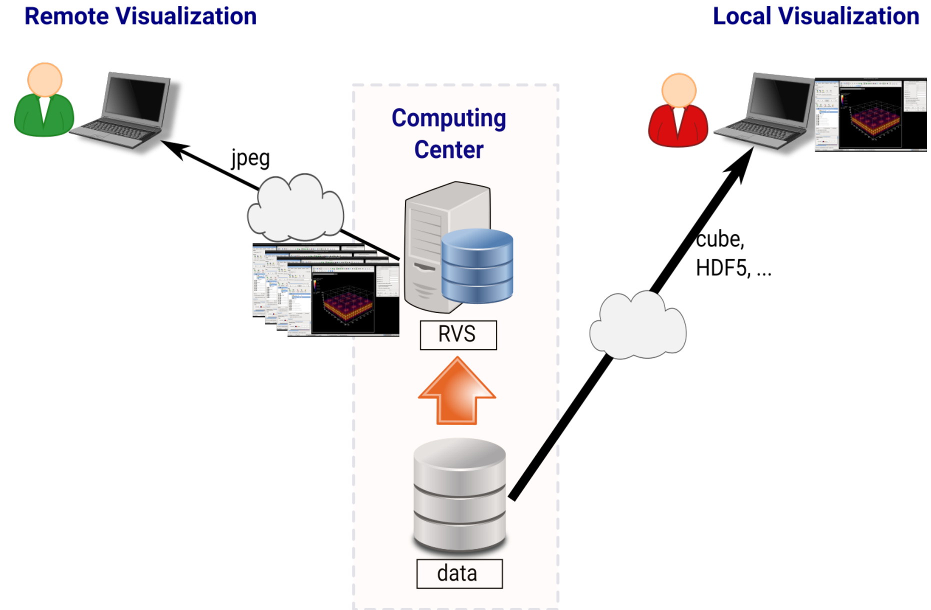
→a necessity for a HPC centre rather than an optional service

enabling technology

- “server-side” rendering on GPUs
- efficient and *transparent* remote rendering solution via WAN: VirtualGL/TurboVNC
- issues: trans-continental latency

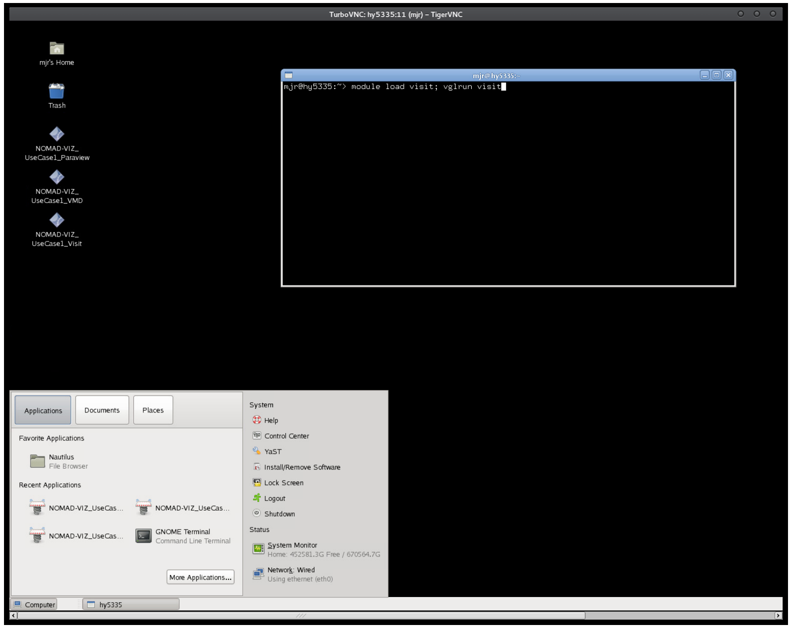
Remote Visualization Technology





Technology for remote visualisation (TurboVNC/VirtualGL):

- proven open-source solution, deployed by many HPC centres (e.g. BSC, CSC, LRZ, MPCDF)
- application-agnostic “remote desktop” with support for hardware-accelerated for 3D graphics
- user's experience: remote Linux desktop with optimization options for network bandwidth, latency, quality of rendering
- transparent use of visualization resources and applications (look-and-feel like local desktop):
  ~>vglrun <executable>



Visualization Tools

Overview of software tools

- IDL, Matlab, python/matplotlib, … for 1D and 2D plots (+time)

  → automated, quantitative analysis with lots of data processing (powerful languages)

- VisIT and Paraview for >2D data (+time)

  → interactive exploration

  → quantitative analysis

  → publication-quality plots, movies

VisIT, Paraview provide complementary functionality, not a replacement for IDL, Matlab, python/matplotlib, ...

- VisIT or Paraview?
  - primarily a matter of taste: very similar functionality, free software, well supported, …
  - Paraview may look and feel more "natural"
  - VisIT has its roots in astrophysics (and some built-in strengths in this area)
- Others?
  - VAPOR, VOREEN (optimized for special purposes), …, + commercial tools (AVIZO)

Visualization Expertise

Example projects

- scientific domains:
  - plasmaphysics, astrophysics, CFD, molecular dynamics, biology, ...
- data structures/grids:
  - regular: cartesian, polar (2D, 3D), block-structured ("Yin-Yan")
  - irregular: (mapped) point clouds
- data sizes, dimensions:
  - up to 4096^3 (Cartesian), $1000 \times 180 \times 360$ (polar), $2048 \times 769 \times 1153$ (cylindrical)
  - up to 10^6 particles in 3D, 10^7 nodes in 3D unstructured mesh
  - all: multi-variable (scalar, vector), time-dependent
  - see also: http://www.rzg.mpg.de/services/visualisation/scientificdata/projects
- tools: parallel HDF5 (+XDMF), VisIT, Paraview

Aims

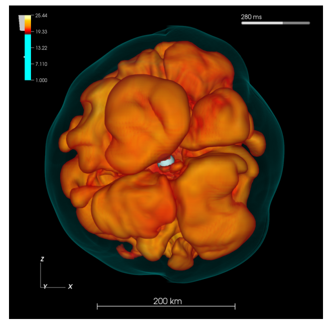
- sketch results & experiences from real-world visualisation projects. (visualisation team's & scientific user's perspective)
- *what can be done? are the tools worth considering at all for my research?*

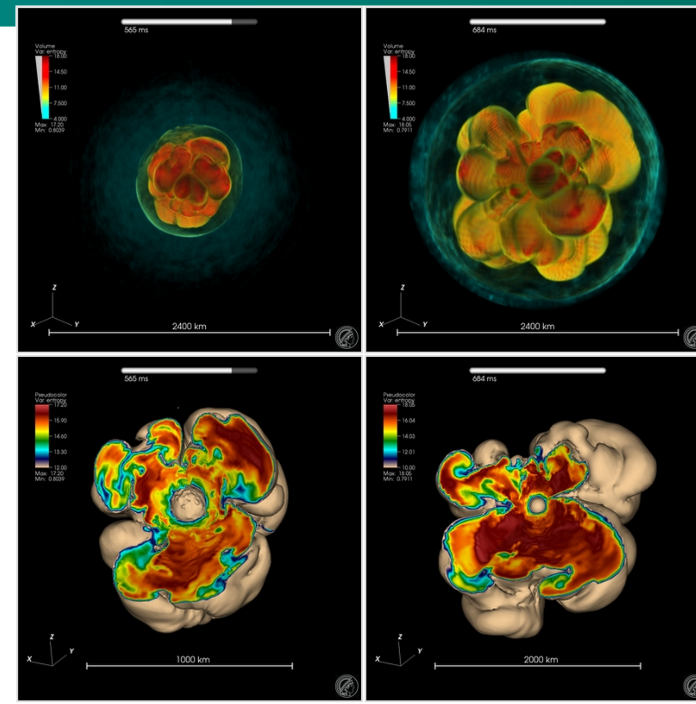
Example 1: Core-collapse Supernova

Simulations by Th. Janka et al. (MPA)

- neutrino-driven explosions of massive stars from first principles
- simulation code: VERTEX (3D, time-dependent radiation hydrodynamics with detailed microphysics)-first 3D simulations of long-term evolution
- code writes HDF5 and XDMF
- spiral mode discovered

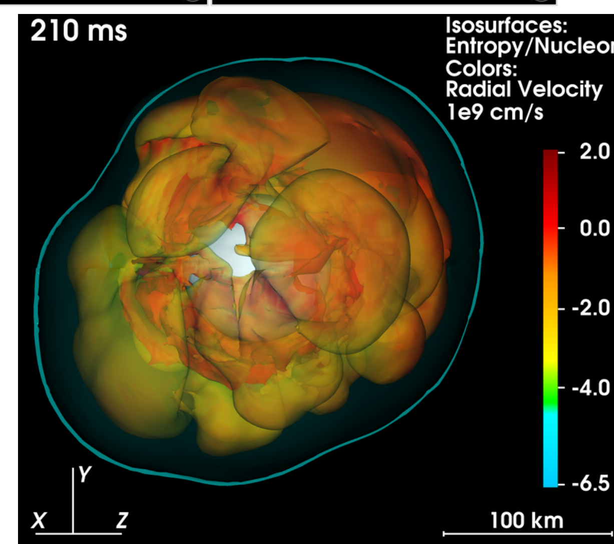
  with the help of 3D visualization





Visualisation approach (E. Erastova, M. Rampp, MPCDF)

- data: (1000×180×360) zones on non-uniform, polar grid
- ≈ 1000 output files (time steps)
- VisIt: pseudo-color plots for data exploration and quantitative analysis
- VisIt: combined volume renderings for HQ movies
- alternative technique: multiple, semi-transparent iso-surfaces



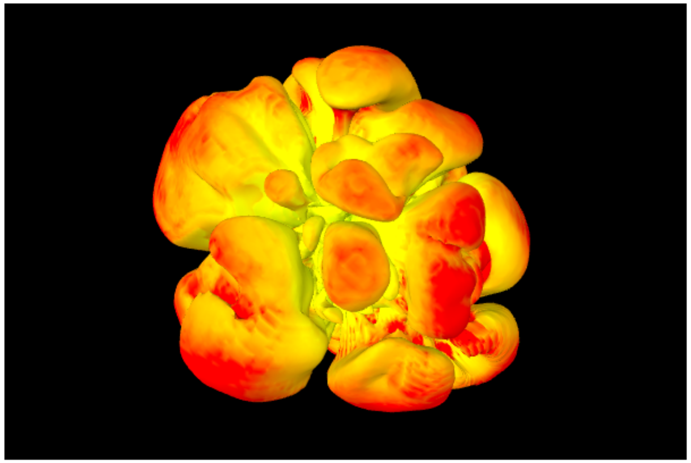
Melson, et al., Astrophysical J. Lett. 808, L42 (2015)

MPCDF

Supernova cont'd: Interactive graphics

Interactive graphics with X3DOM

- supplements publishing of simulation results, e.g., by APJ (http://iopscience.iop.org/0004-637X/793/2/127/media)
- 3D data format and object model (http://www.x3dom.org/)
- X3D(OM) file export supported by Paraview, VisIT (>= 2.10)
- controls: mouse, zoom, +custom interaction
- HTML5, no browser plugin required



Visualization by E. Erastova (MPCDF)
Simulations by H.-Th. Janka et al. (MPA)

.x3d file export

File Edit View Help

```
    <html>
        <head>
            <title>450 ms</title>
            <script type='text/javascript' src='http://www.x3dom.org/download/x3dom.js'> </script>
            <link rel='stylesheet' type='text/css' href='http://www.x3dom.org/download/x3dom.css'></link>
        </head>
        <body>
            <p>450 ms
<br/>
Powered by <a href="http://www.x3dom.org">X3DOM</a>
<br/>
</p>

    <x3d width='600px' height='400px'>
    <scene>
    <inline url="450-1.x3d"> </inline>
    </scene>
    </x3d>
        </body>
    </html>
```

.x3d file reference

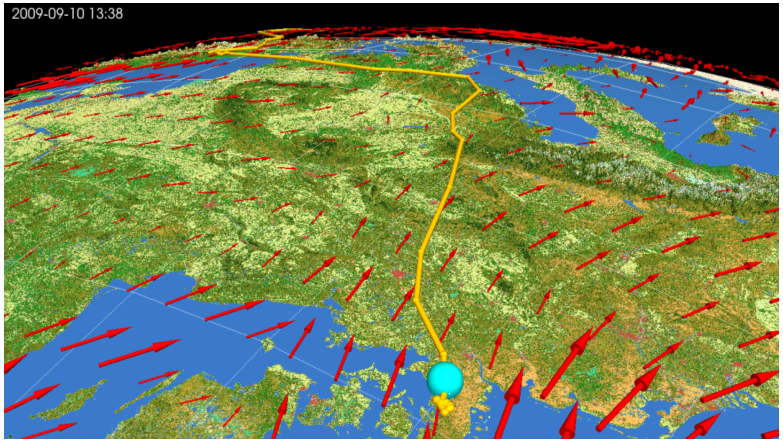
```
<?xml version="1.0" encoding ="UTF-8"?>

<X3D profile="Immersive" version="3.0">
  <head>
    <meta name="filename" content="/u/fruwl/Desktop/450-1.x3d"/>
    <meta name="generator" content="Visualization ToolKit X3D exporter v0.9.1"/>
    <meta name="numberofelements" content="65"/>
  </head>
  <Scene>
    <Background skyColor="0 0 0"/>
    <Viewpoint fieldOfView="0.523599" position="-1.88614e+08 2.40873e+08 -5.17449e+07" description="Default View" orientation="-0.502423 0.577924 0.643098 -3.20849"
centerOfRotation="370850 -810850 1.6746e+06"/>
    <NavigationInfo type='"EXAMINE" "FLY" "ANY"' speed="4" headlight="true"/>
    <DirectionalLight ambientIntensity="1" intensity="0" color="1 1 1"/>
    <Transform DEF="ROOT" translation="0 0 0">
      <DirectionalLight direction="-0.111619 -0.766044 -0.633022" color="1 0.97232 0.90222" intensity="0.75" on="true"/>
      <DirectionalLight direction="0.0449435 0.965926 -0.254887" color="0.90824 0.93314 1" intensity="0.25" on="true"/>
      <DirectionalLight direction="-0.939693 0 0.34202" color="0.9998 0.9998 0.9998" intensity="0.214286" on="true"/>
      <DirectionalLight direction="0.939693 0 0.34202" color="0.9998 0.9998 0.9998" intensity="0.214286" on="true"/>
      <Transform translation="0 0 0" rotation="0 0 1 -0" scale="1 1 1">
        <Shape>
          <Appearance>
            <Material ambientIntensity="0" emissiveColor="0 0 0" diffuseColor="1 1 1" specularColor="0 0 0" shininess="0.78125" transparency="0"/>
          </Appearance>
          <IndexedFaceSet solid="false" colorPerVertex="true" normalPerVertex="true" coordIndex="
            0 1 2 3 -1
            1 4 2 -1
            5 6 1 0 -1
            7 4 1 8 -1
            8 1 6 -1
            9 0 3 10 -1
            9 11 5 0 -1
            12 13 14 -1
            13 15 16 14 -1
            12 14 17 18 -1
            14 16 19 -1
            17 14 19 20 -1
            21 22 23 24 -1
            22 21 25 26 -1
            27 28 25 21 -1
            21 24 29 27 -1
            30 29 24 31 -1
```

Example 2: Geospatial data & bird migration

Data by M. Wikelski (MPI f. Ornithology)

- observational data
- a bird’s (gull) track correlated with wind data
- topography, earth's magnetic field, …
- time-dependent data



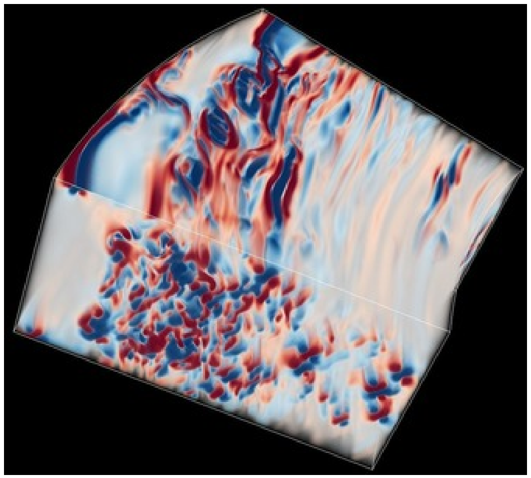
movie presented by M. Wikelski at general assembly of the MPG, 2012

Visualisation (K. Reuter, MPCDF & K. Safi, MPI-Orn.)

- ParaView (support for importing geo data), lots of Python scripting
- tedious generation and adaptation of camera movement (use Blender instead?)

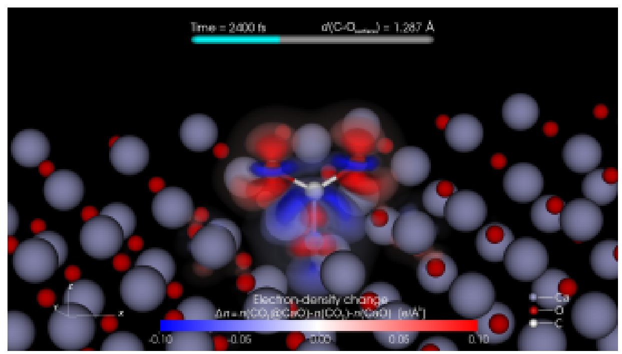
adapted for wall-projection in the “*henn house*” (visitors and media center in Radolfzell at lake Bodensee)

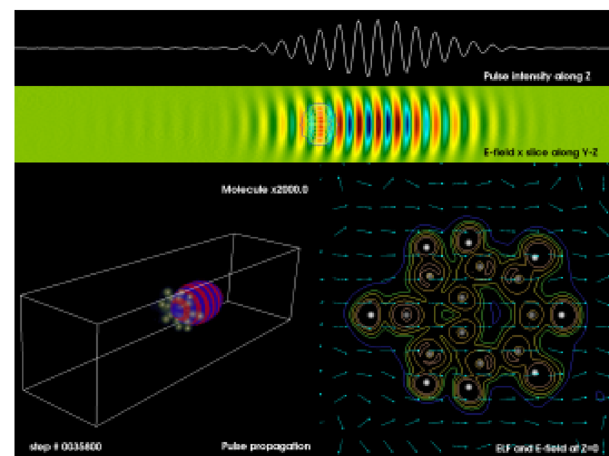
More MPCDF project examples ...



(3) DNS of turbulence in quasi-Keplerian flows

Simulations: M. Avila et al. (U. Bremen)
Visualization: M. Rampp (MPCDF)





(4) Ab-initio molecular simulations

Simulations: S. Levchenko (FHI), H. Appel (Max-Planck-Inst. f. Structue and Dynamics of Matter)
Visualization: M. Compostella & M. Rampp (MPCDF)

→ *NOMAD*

Outline

- Introduction
- Visualization Services in classical HPC environments (MPCDF)
- **Virtual reality applied to large-scale turbulence simulations**
  (MPI for Dynamics and Self-Organization, TUM, LRZ, MPCDF)
- NOMAD COE - Advanced Graphics
  - Remote Visualization Services (FHI, MPCDF)
  - High-quality visualizations (MPCDF)
- pyiron – A Python-based IDE for computational materials design
  (MPI for Iron Research, supported by MPCDF)
- Summary

(5) Turbulence Visualization

- Visualization of DNS of Turbulence: A *Rollercoaster Experience*

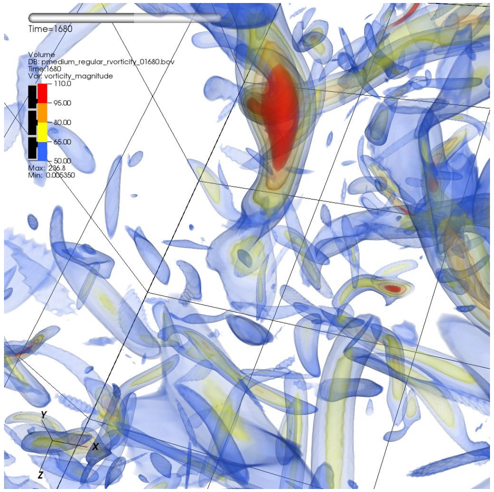
  (Matthias Albert[1], Michael Wilczek[2], Cristian Lalescu[2], Markus Rampp[1])

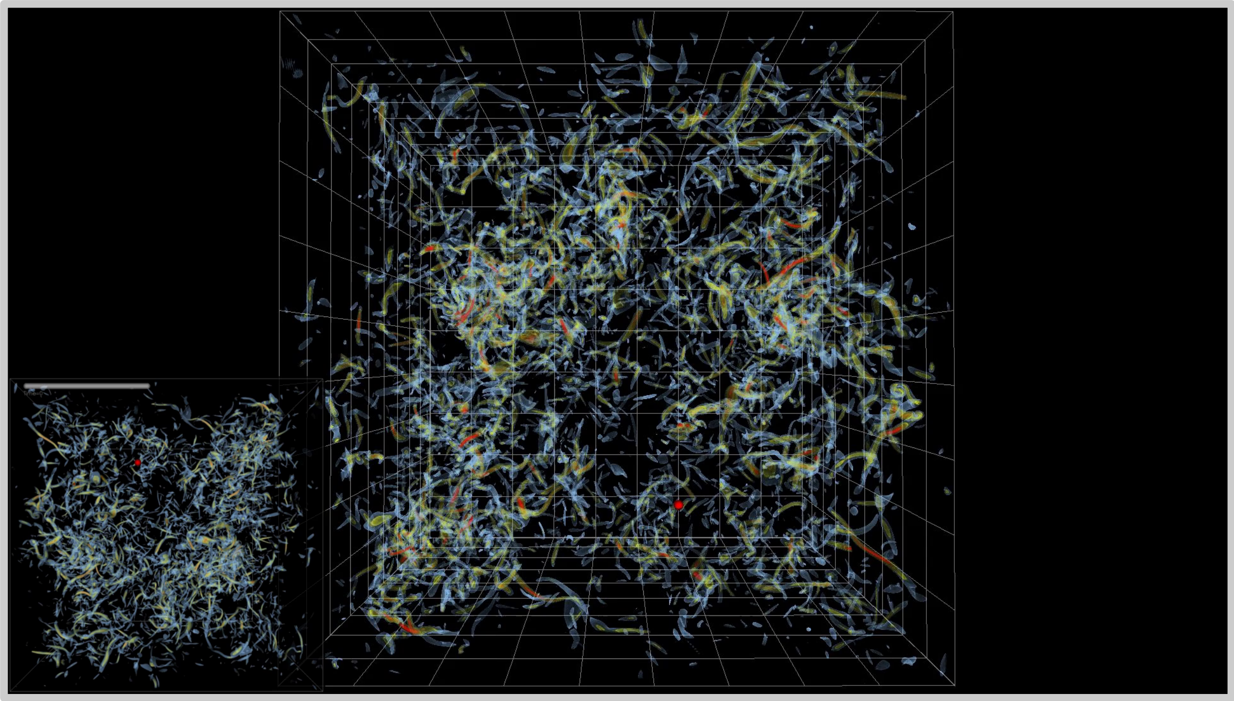
- HPC Simulation Code: BFPS
  developed at MPI of Dynamics and Self-Organization[2] and MPCDF[1]

- Data:
  - HDF5 + XDMF
  - time-dependent, 3-D scalar field (vorticity): 128^3, 576^3 ,…, 4096^3
  - trajectories of tracer particles advected with the flow

- Tools and Technique:
  - VisIt and Python Scripting

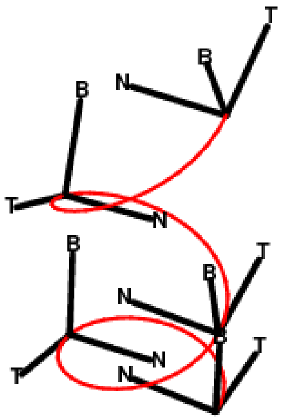
… let's follow a tracer particle on its ride with the turbulent flow

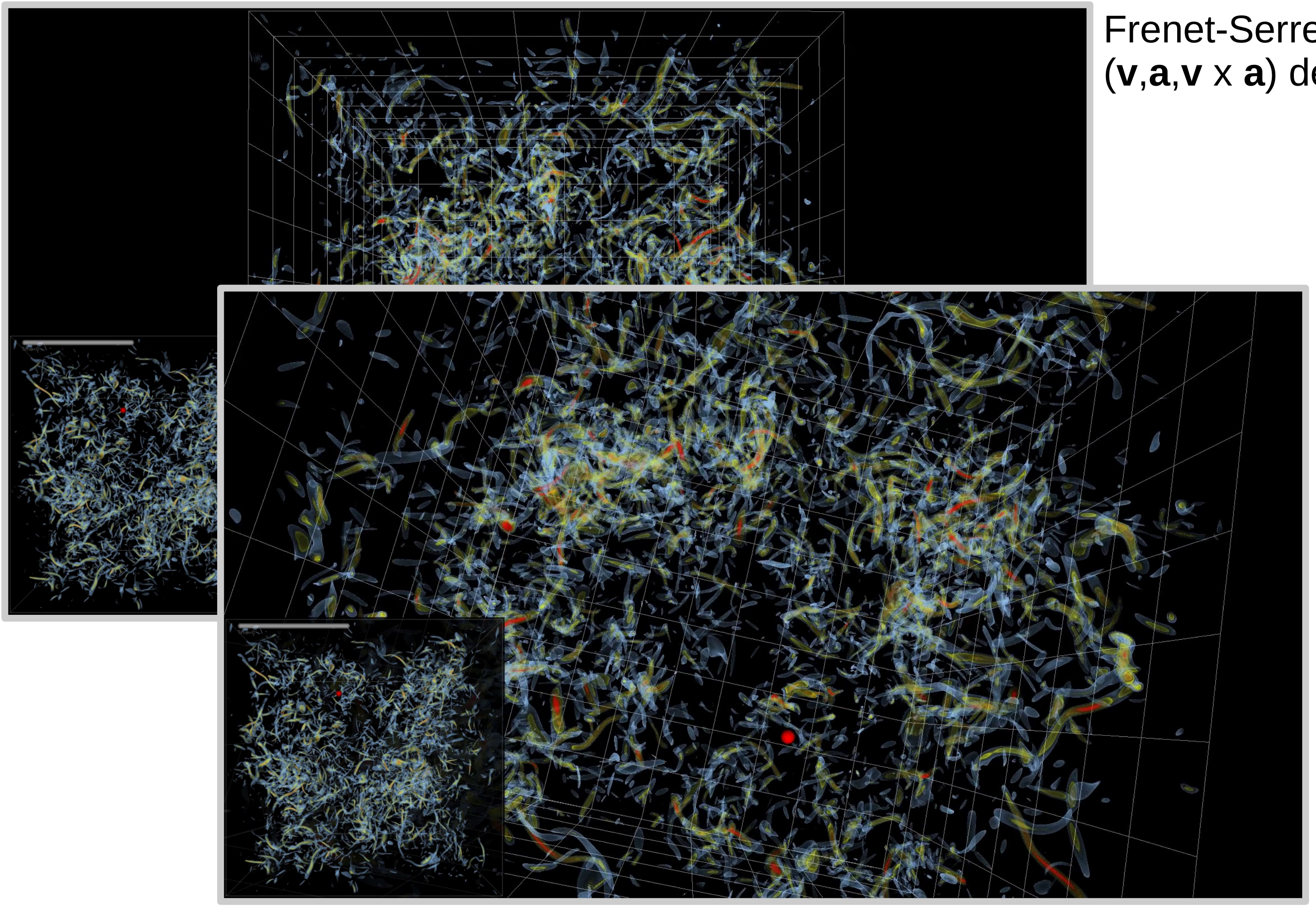




Frenet-Serret Frame

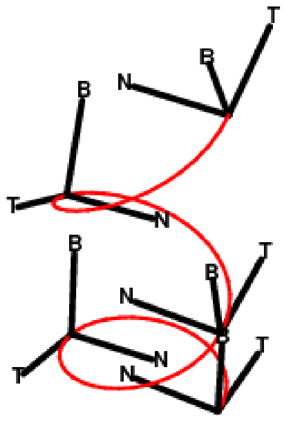
(**v**,**a**,**v** x **a**) defines the camera

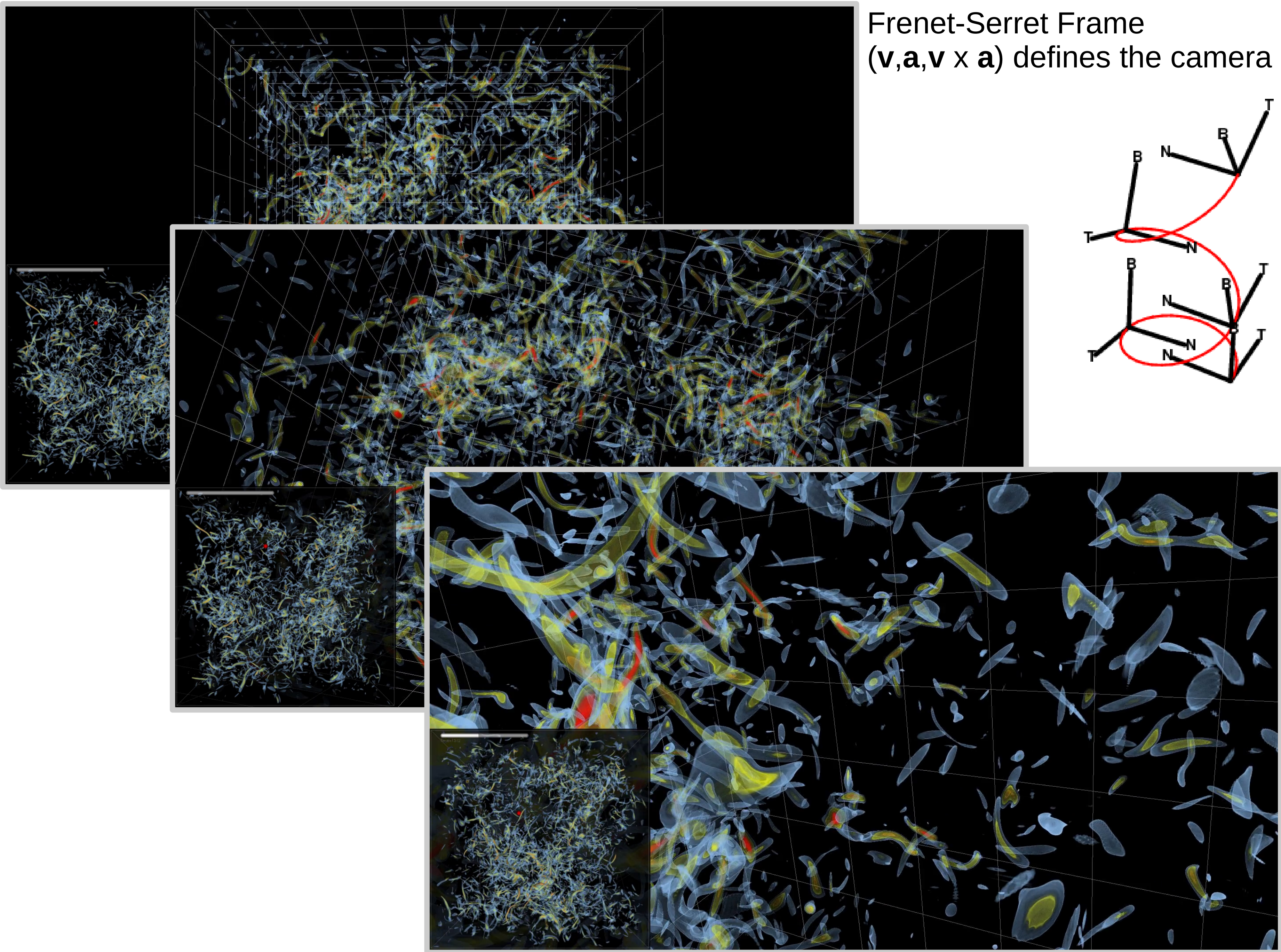




Frenet-Serret Frame

(**v**,**a**,**v** x **a**) defines the camera





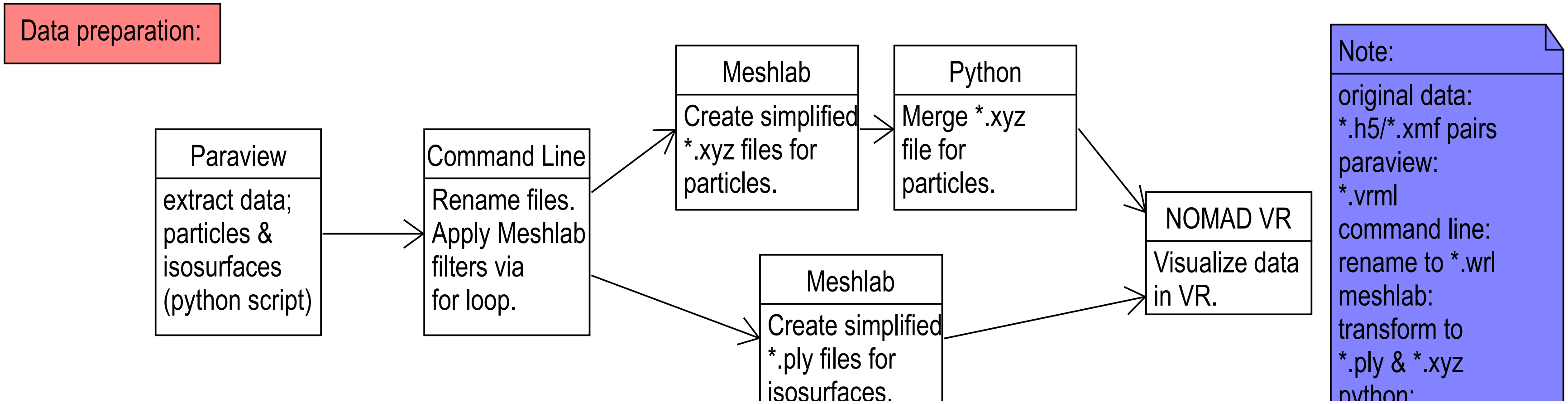
Frenet-Serret Frame
(**v**,**a**,**v** x **a**) defines the camera

(5) Turbulence Visualization in VR

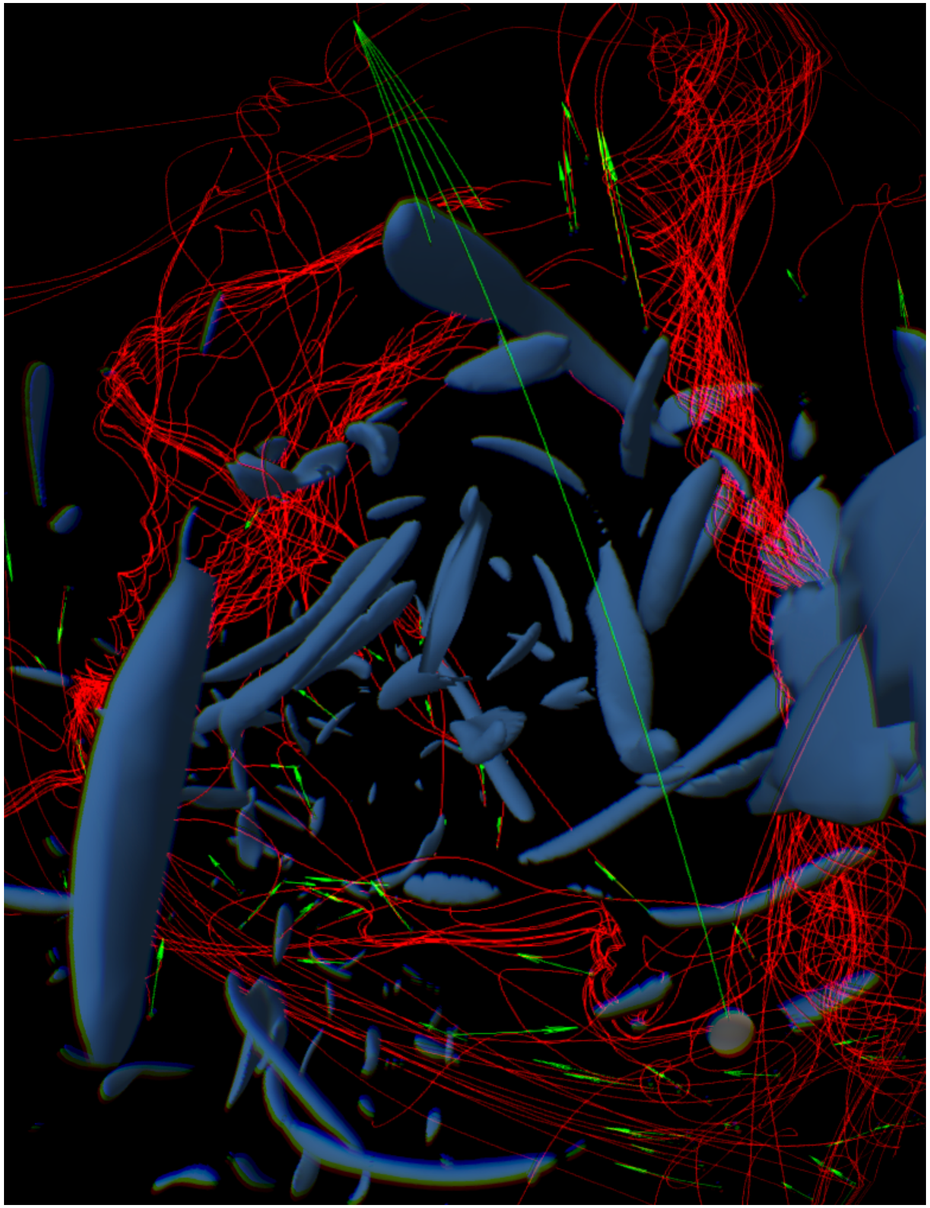
- Turbulence Visualization in Virtual Reality
  (M. Albert, Batchelor project; TUM, LRZ, MPCDF)
- VR technology fits well with the Lagrangian approach
- Extension of the software *NOMAD VR* (→ Rubén García-Hernández, LRZ)
- New developments:
  - data preprocessing (isosurfaces)
  - particle picker (VR ←→ VisIt)
  - trajectory visualization
  - haptic feedback
  - challenge: preprocessing and handling of massive HPC simulation data,
    solved by a separate thread that streams data from an SSD through the GPU

(5) Turbulence Visualization in VR

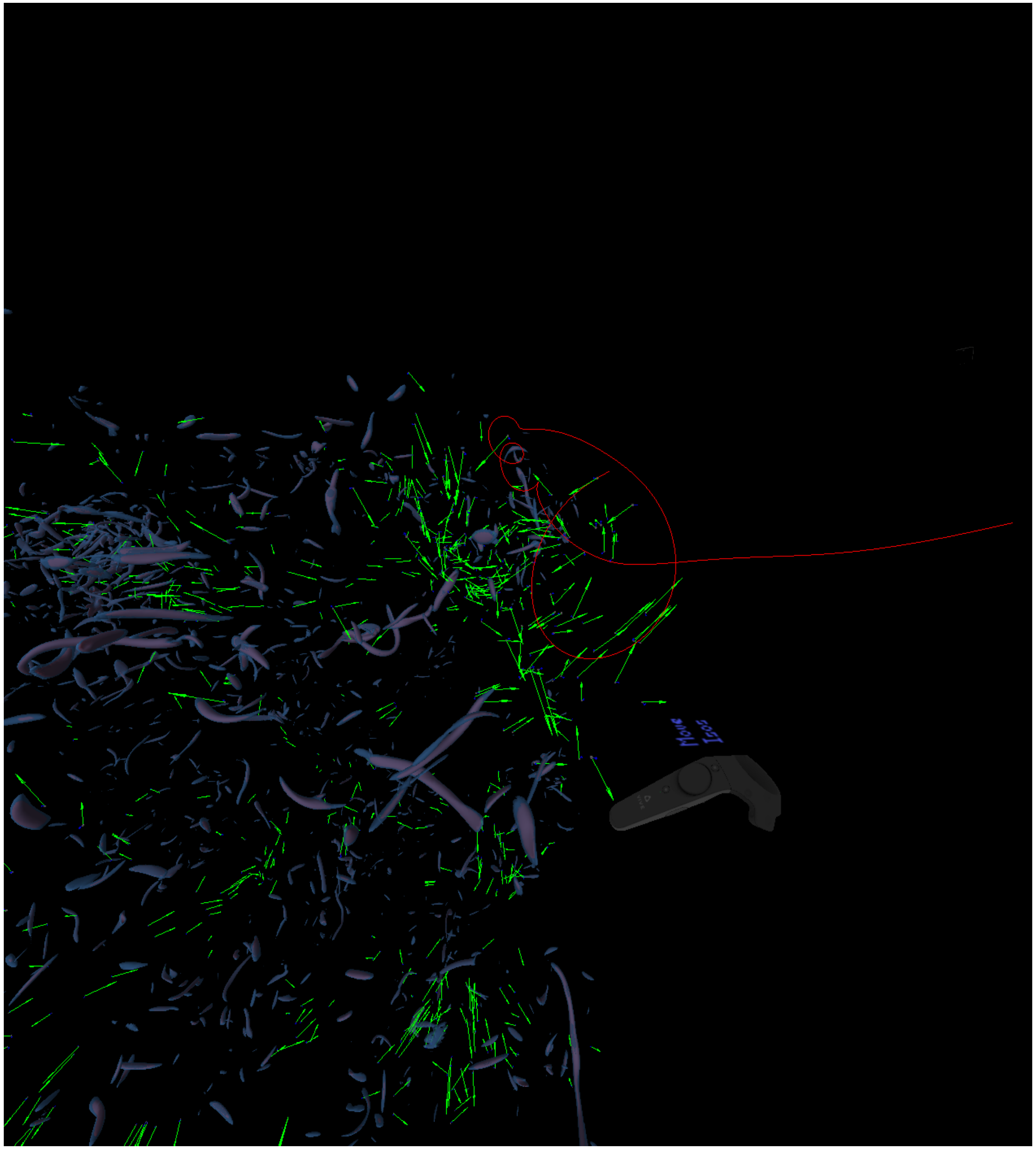
- complex data preprocessing pipeline



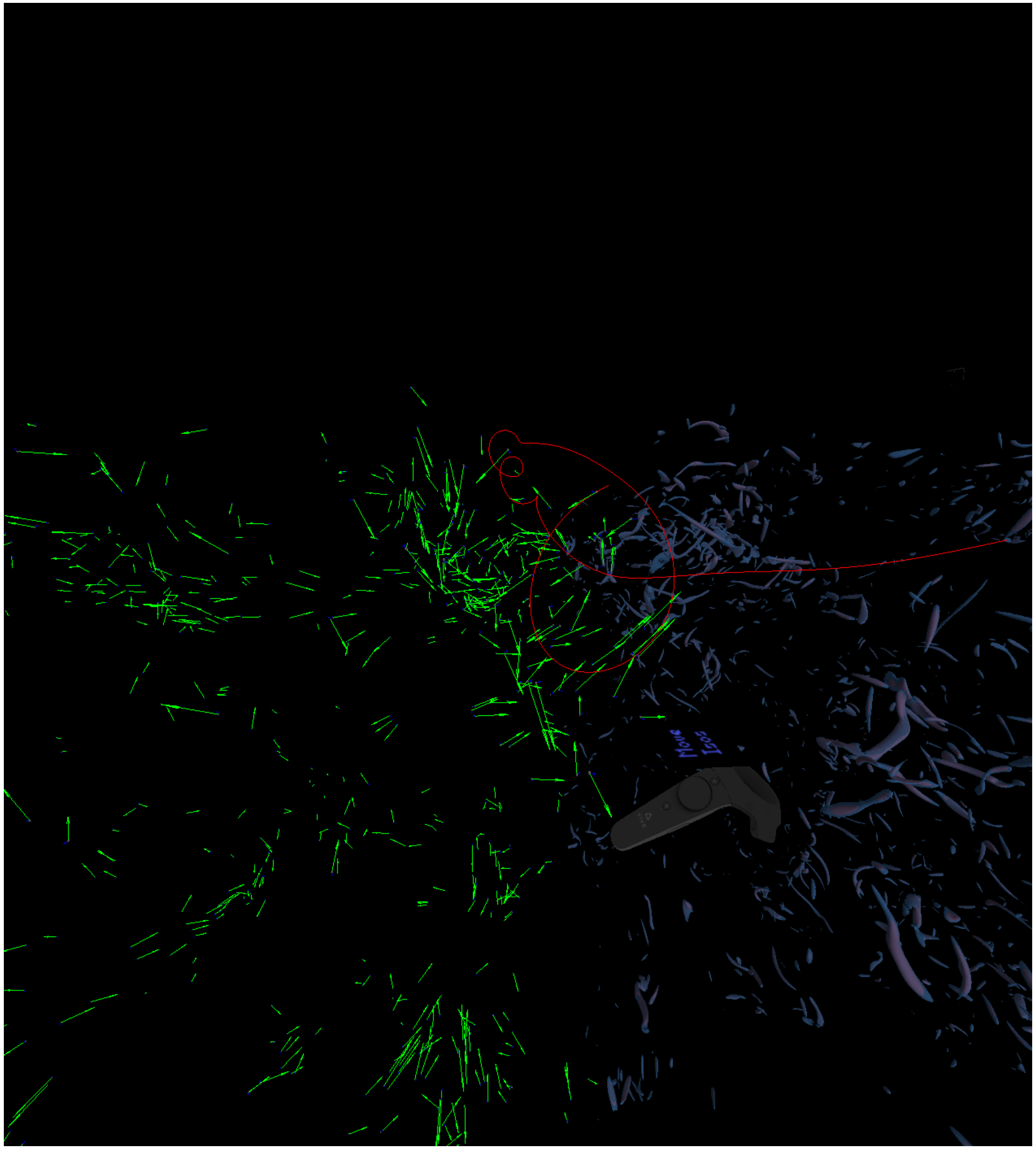
(5) Turbulence Visualization in VR



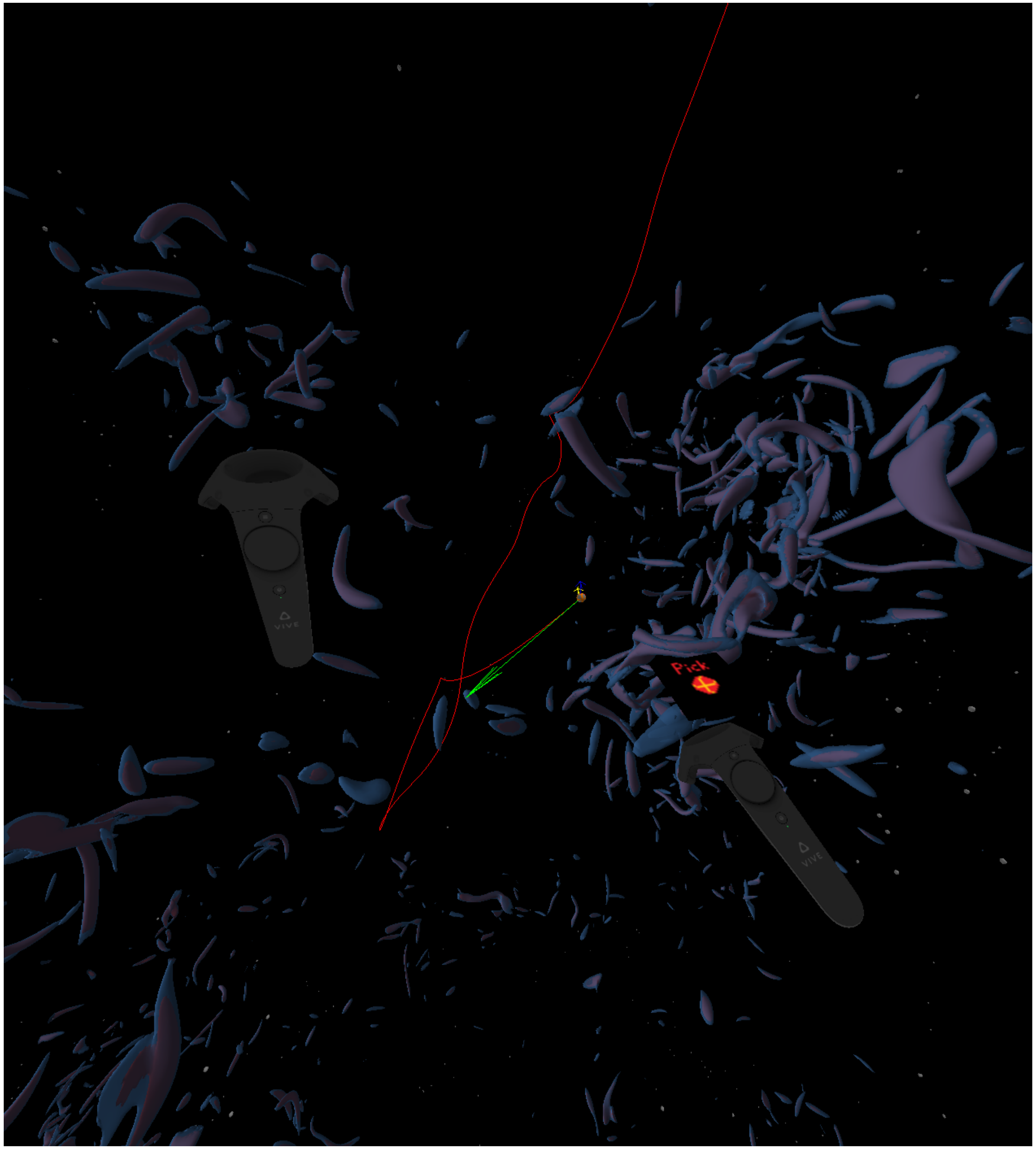
(5) Turbulence Visualization in VR



(5) Turbulence Visualization in VR



(5) Turbulence Visualization in VR



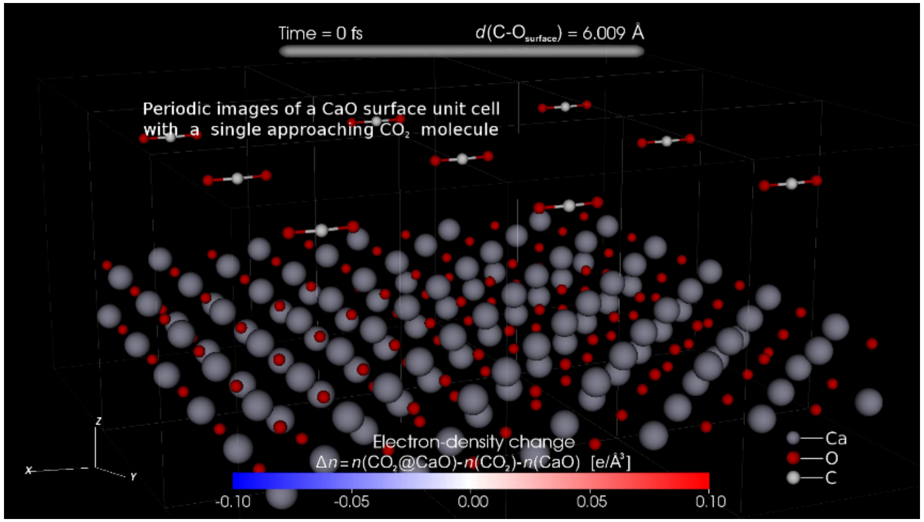
Outline

- Introduction
- Visualization Services in classical HPC environments (MPCDF)
- Virtual reality applied to large-scale turbulence simulations (MPI for Dynamics and Self-Organization, TUM, LRZ, MPCDF)
- **NOMAD COE - Advanced Graphics**
  - Remote Visualization Services (FHI, MPCDF)
  - High-quality visualizations (MPCDF)
- pyiron – A Python-based IDE for computational materials design (MPI for Iron Research, supported by MPCDF)
- Summary

NOMAD COE – Advanced Graphics

Goals

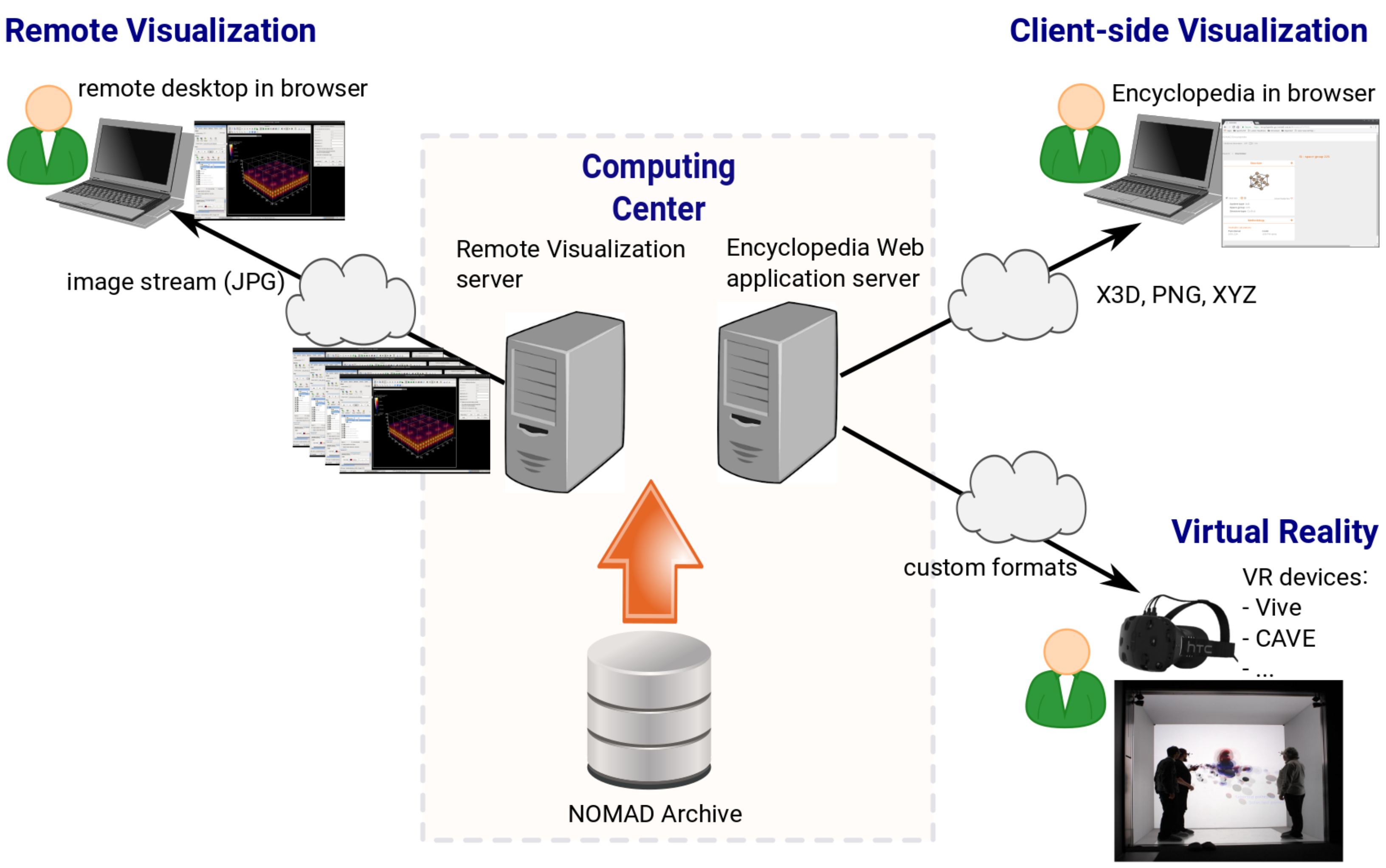
- develop & apply advanced methodologies
- adopt new technologies: VR [R. García-Hernández (LRZ)],
  browser-based remote visualization [M. Compostella (MPCDF), et al.]
- produce high quality material for training, dissemination, education;
  example movie: reaction of a CO2 molecule with a CaO surface



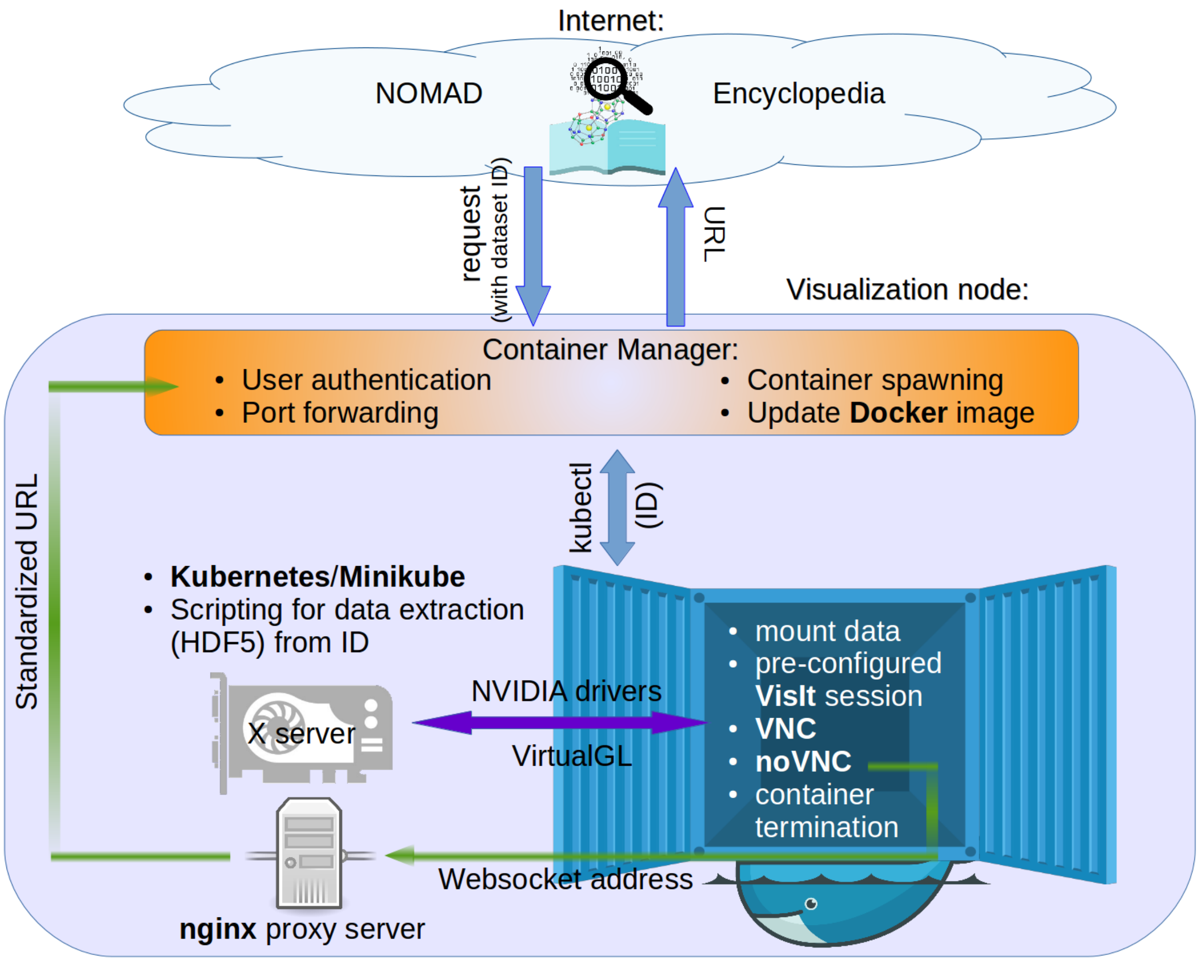
[M. Compostella (MPCDF)]

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 676580.

NOMAD Visualization Strategy



NOMAD container-based Remote Visualization



→ demo at https://labdev-nomad.esc.rzg.mpg.de/remotevis/vnc.html

Upcoming: NOMAD SUMMER 2018

- A hands-on course on tools for novel-materials discovery
- Scope: Data-driven research meets materials science and engineering: Making Big Data of materials comprehensible.
- September 24 - 27, 2018, Lausanne, Switzerland
- Registration deadline: May 2, 2018
- Topics
  1. data repositories, archives and metadata
  2. advanced graphics (M. Rampp, MPCDF)
  3. nomad encyclopedia
  4. preparation and analysis of high-throughput simulations
  5. big-data analytics

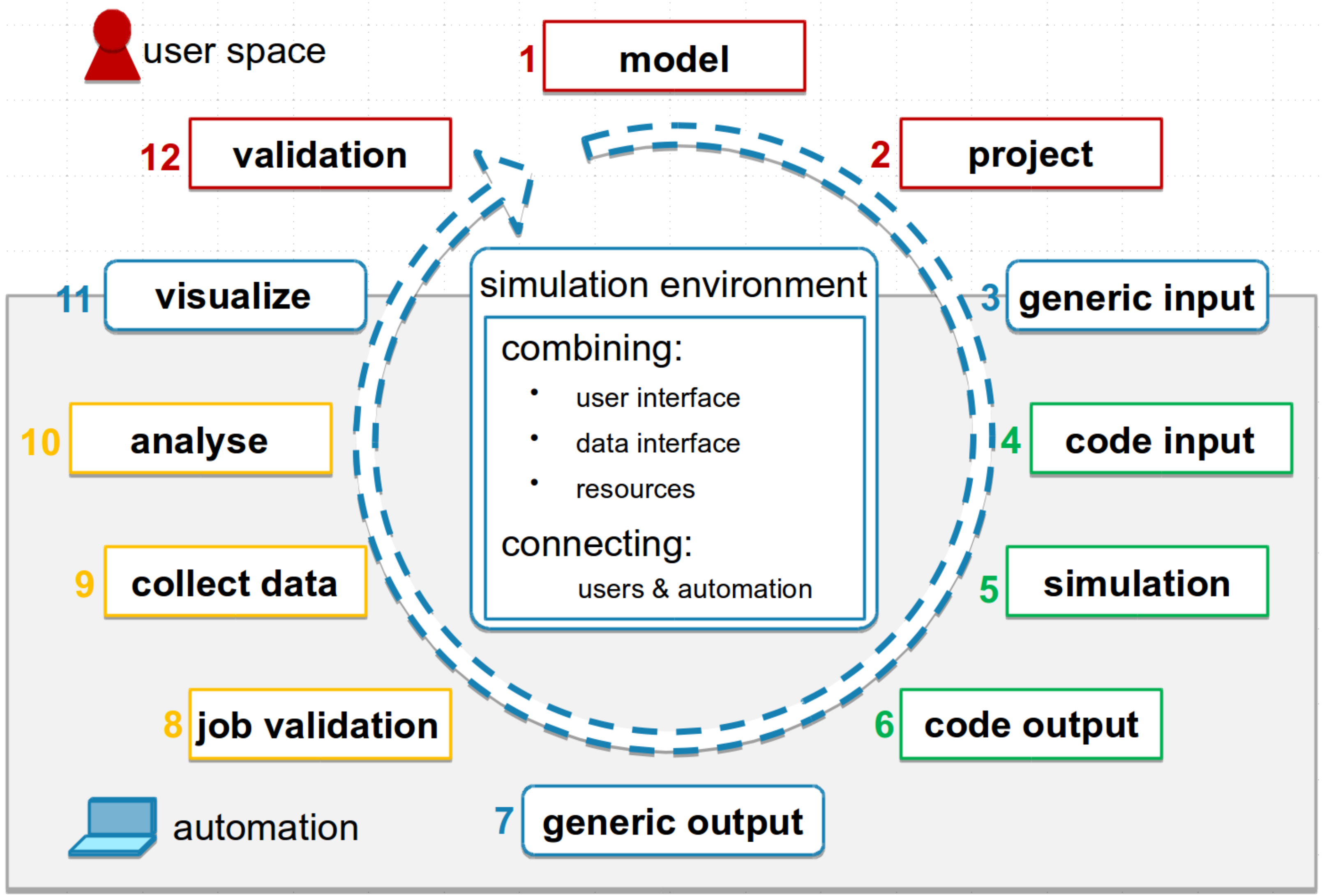
Outline

- Introduction
- Visualization Services in classical HPC environments (MPCDF)
- Virtual reality applied to large-scale turbulence simulations (MPI for Dynamics and Self-Organization, TUM, LRZ, MPCDF)
- NOMAD COE - Advanced Graphics
  - Remote Visualization Services (FHI, MPCDF)
  - High-quality visualizations (MPCDF)
- **pyiron – A Python-based IDE for computational materials design** (MPI for Iron Research, supported by MPCDF)
- Summary

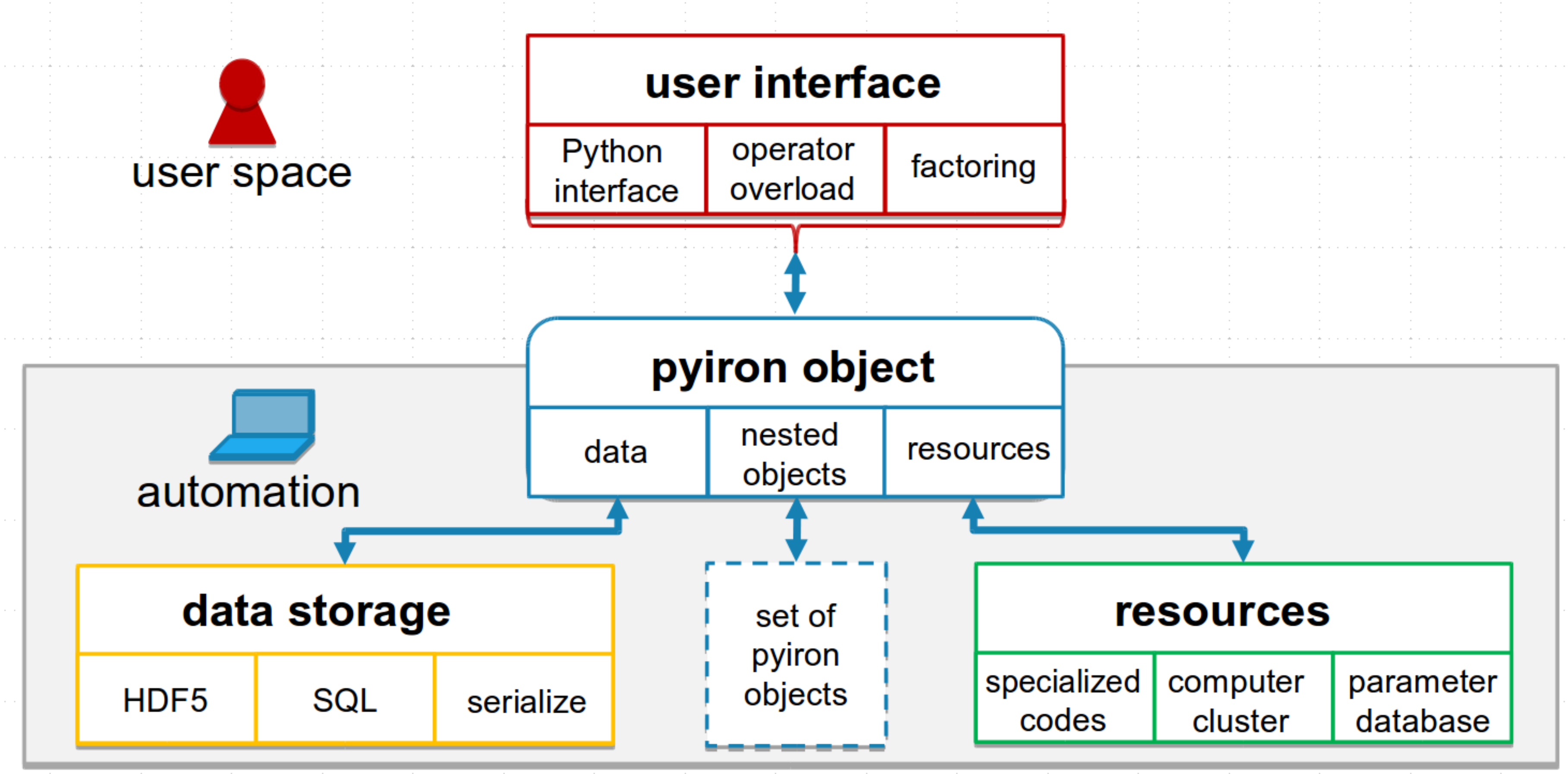
pyiron

- pyiron is a Python based IDE for computational materials design (**Py**thon + MPI for **Iron** Research)
- developed at the MPI for Iron Research by Jörg Neugebauer, Jan Janssen, et al.
- pyiron enables researchers to automate simulation protocols
  - handle high complexity: Physics, IO formats, simulation codes, HPC, ...
  - conserve expertise: inconsistency, documentation, ...
- pyiron user interface is based on Jupyter notebooks
  - simulation, data analysis, publication-quality visualization and plotting in a single document

pyiron – IDE for simulation protocols



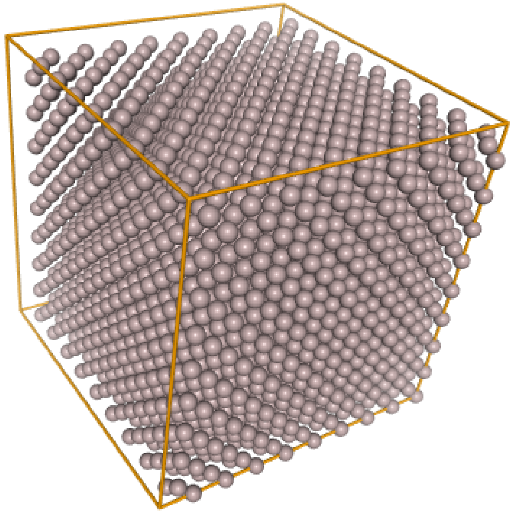
pyiron internals



Create an atomic structure

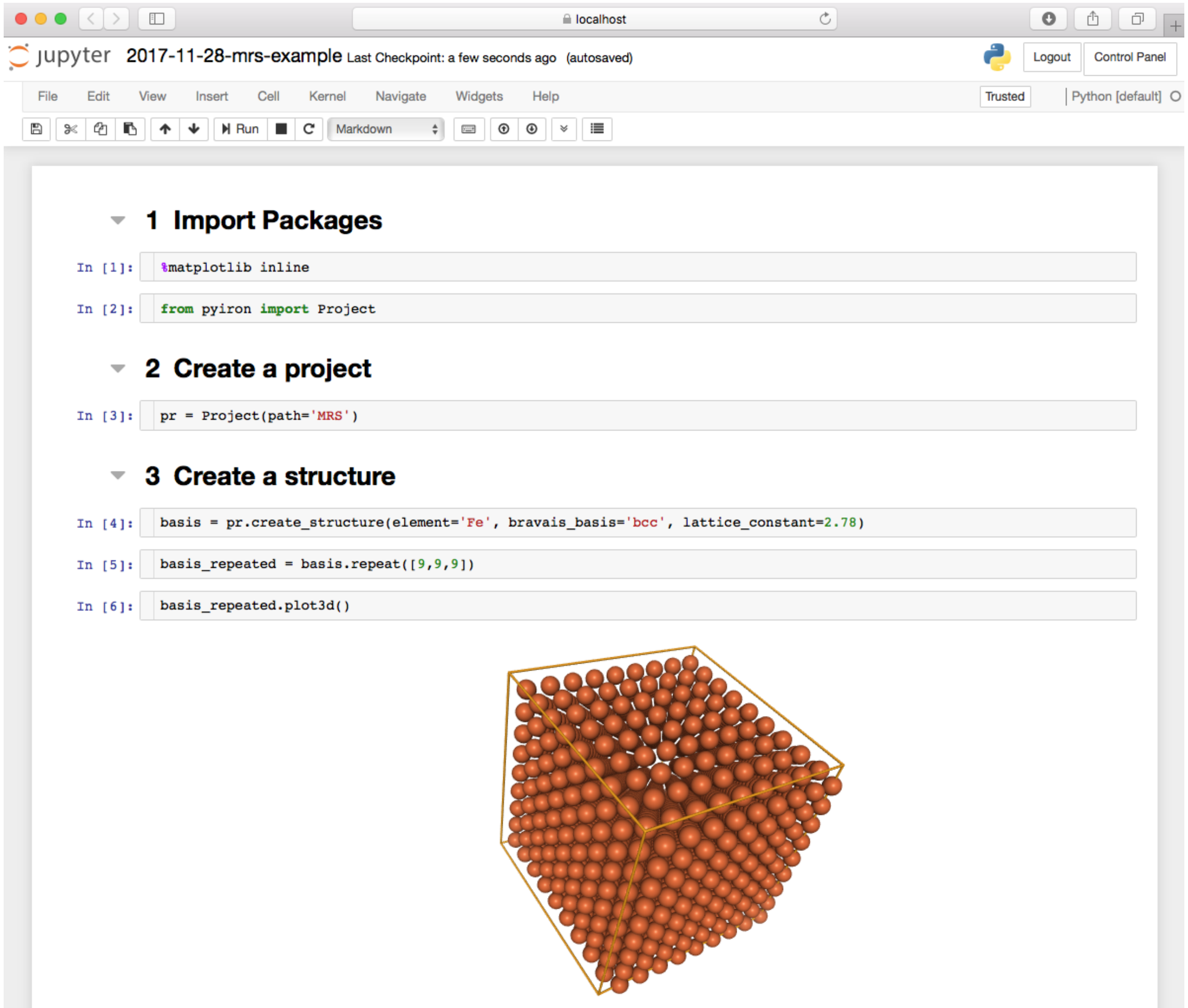
In [6]:
```
basis = pr.create_structure(element="Al",
                             bravais_basis="fcc",
                             lattice_constant=lat)
```

In [7]:
```
basis_repeated = basis.repeat([9, 9, 9])
basis_repeated.plot3d()
```



MPCDF

pyiron IDE for material science



pyiron IDE for material science

4 Setup a VASP calculation

In [7]:
```
ham_vasp = pr.create_job(pr.job_type.Vasp, 'vasp')
```

In [8]:
```
ham_vasp.structure = basis
```

5 Setup Murnaghan

In [9]:
```
murn = pr.create_job(pr.job_type.Murnaghan, 'murn')
```

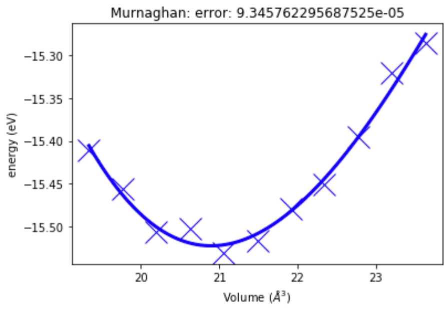
In [10]:
```
murn.ref_job = ham_vasp
```

6 Run Murnaghan

In [11]:
```
murn.run()
```

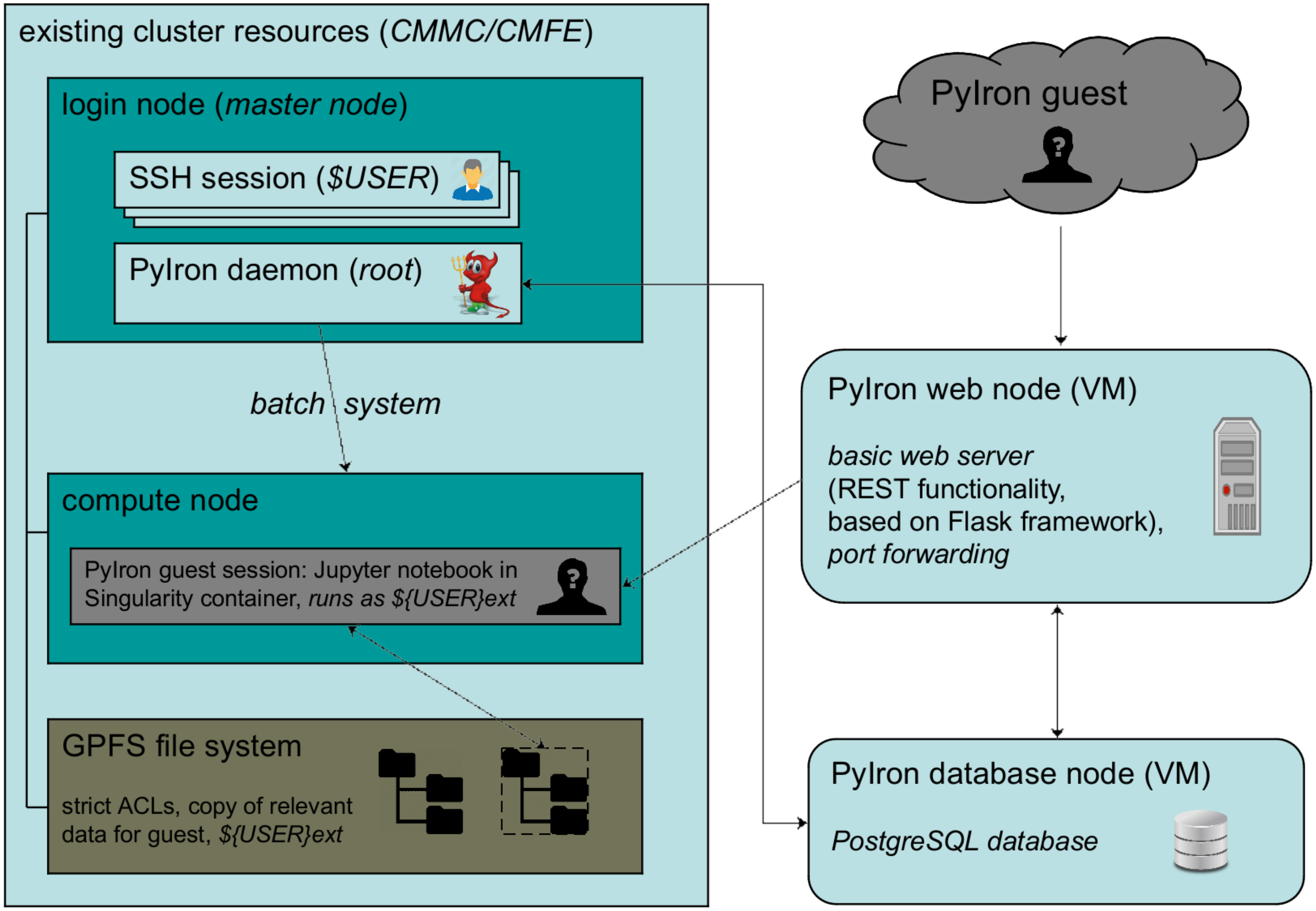
7 Analyse Murnaghan

In [12]:
```
murn.plot()
```



In [ ]:

enabling guest sessions for pyiron



MPCDF proposal, work in progress

pyiron – Availability

- Documentation

  http://pyiron.org

- Upcoming paper on arXiv.org

- Installation e.g. via 'conda' for Anaconda Python

```
conda config --add channels conda-forge
conda config --add channels pyiron
conda install pyiron
```

- Docker image (→ live demo)

```
docker pull pyiron/pyiron
docker run -i -t -p 8888:8888 pyiron/pyiron /bin/bash -c
"/opt/conda/bin/jupyter lab --notebook-dir=/opt/notebooks --ip='*' --
port=8888 --no-browser --allow-root"
```

Summary – Q&A?

- Visualization Services in classical HPC environments (MPCDF)
- Virtual reality applied to large-scale turbulence simulations (MPI for Dynamics and Self-Organization, TUM, LRZ, MPCDF)
- NOMAD COE - Advanced Graphics
  - Remote Visualization Services (FHI, MPCDF)
  - High-quality visualizations (MPCDF)
- pyiron – A Python-based IDE for computational materials design (MPI for Iron Research, supported by MPCDF)