Complex Workflows between HPC and Grid applications

Francisco Castejón (francisco.castejon@ciemat.es)
(Leader of Fusion Cluster)
Antonio Gómez (antonio.gomez@ciemat.es)

CIEMAT, Spain

GOBIERNO DE ESPAÑA MINISTERIO DE CIENCIA E INNOVACIÓN

Ciemat
Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

EGEE-III Final Review, 23-24 June, 2010

e-infrastructure

www.eu-egee.org

Collaborators

Michal Owsiak, Tomasz Zok, Marcin Plociennik

What we want to show

- The feasibility of mixed HPC-Grid scientific workflows.
- Building blocks for complex fusion modeling workflows (to be used by EFDA -European Fusion Development Agreement- community).
- Developments essential for fusion community that could be reused by other communities.
- Results of a pilot project among EGEE-DEISA-EUFORIA.
  - Develop heterogeneous workflows for fusion community.
  - Put workflows in production using real computing infrastructures.
  - Workflows between grid-grid, grid-HPC and HPC-HPC applications.

Why?

- **To help fusion scientists by enhancing the modeling capabilities for ITER sized plasmas.**
- **To promote innovative aspects:**
  - Dynamic coupling of codes and applications on a set of heterogeneous platforms into a single framework through a workflow engine.
- **To foster a collaboration environment between EUFORIA, DEISA and EGEE: Building Community.**
- **DEMO activities funded by EGEE.**

Demo

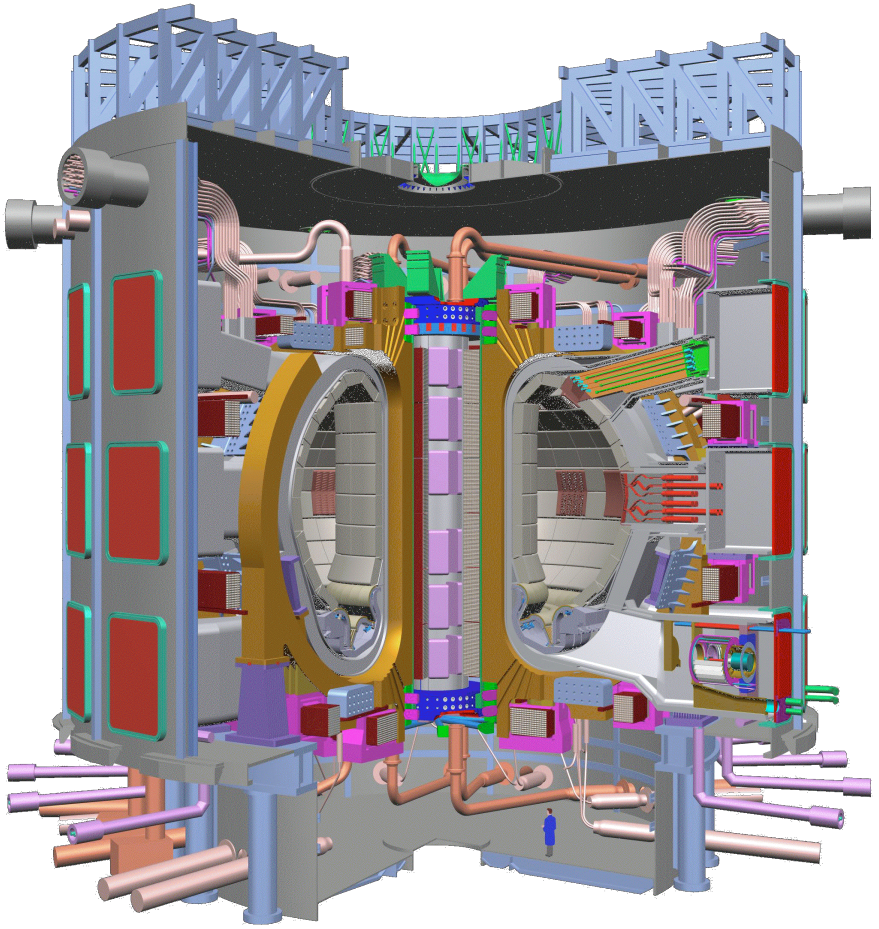
Scientific Gateways

Building Workflows

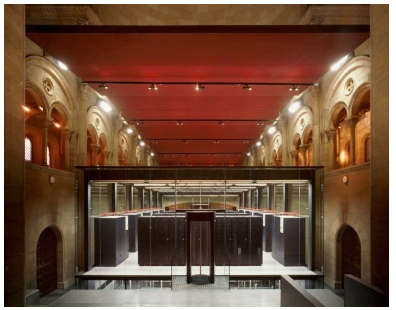
ITER

TARGET: Fusion community

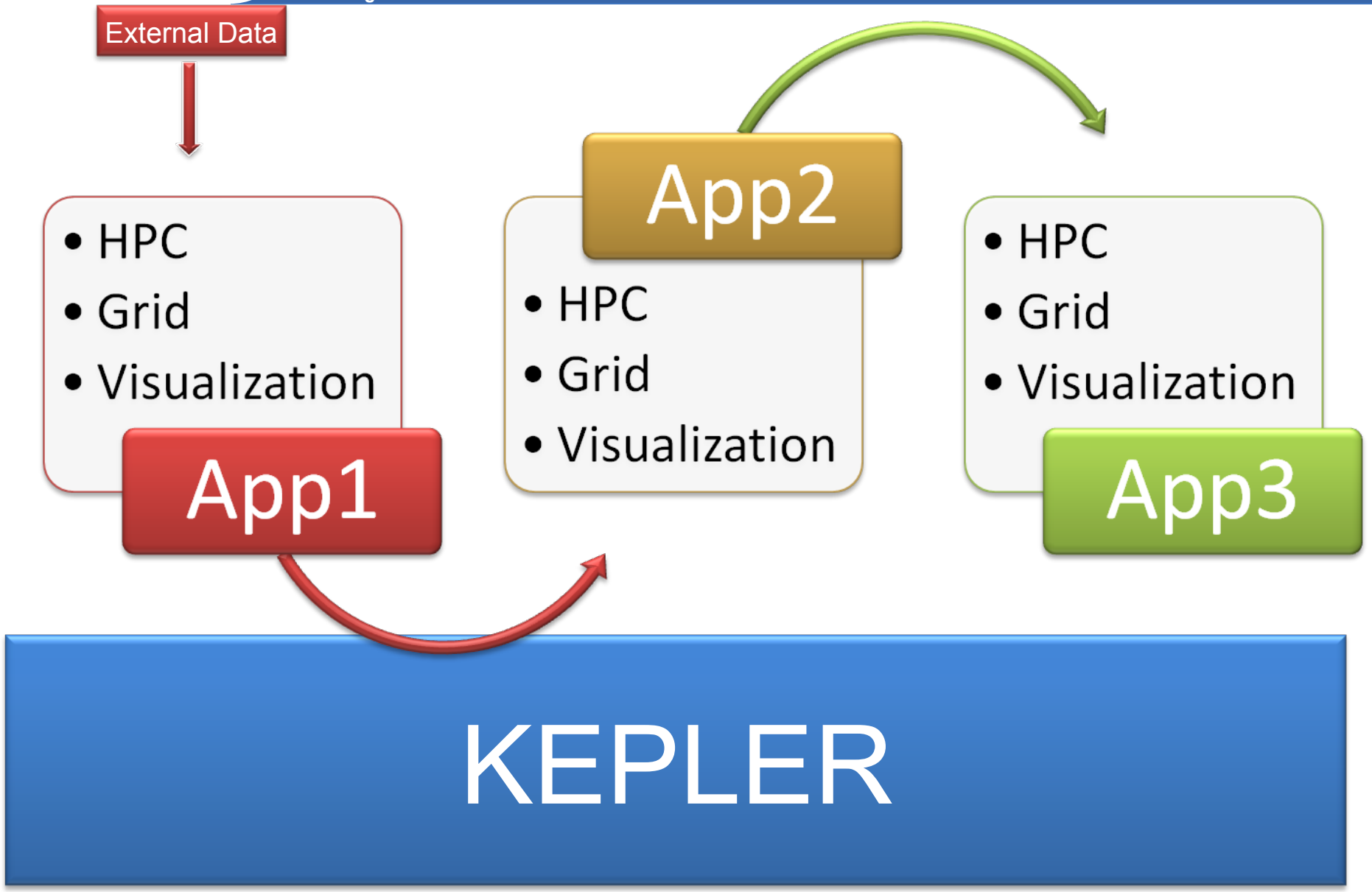
- **Next generation of energy production: social and environmentally acceptable**
- **It still has many open issues: plasma confinement & material science.**
- **It becomes not only a scientific problem but also a large scale computational problem where distributed infrastructures and HPC are required:**
  - Experimental data processing.
  - Experimental scenario development.
  - Theory.



Fusion community



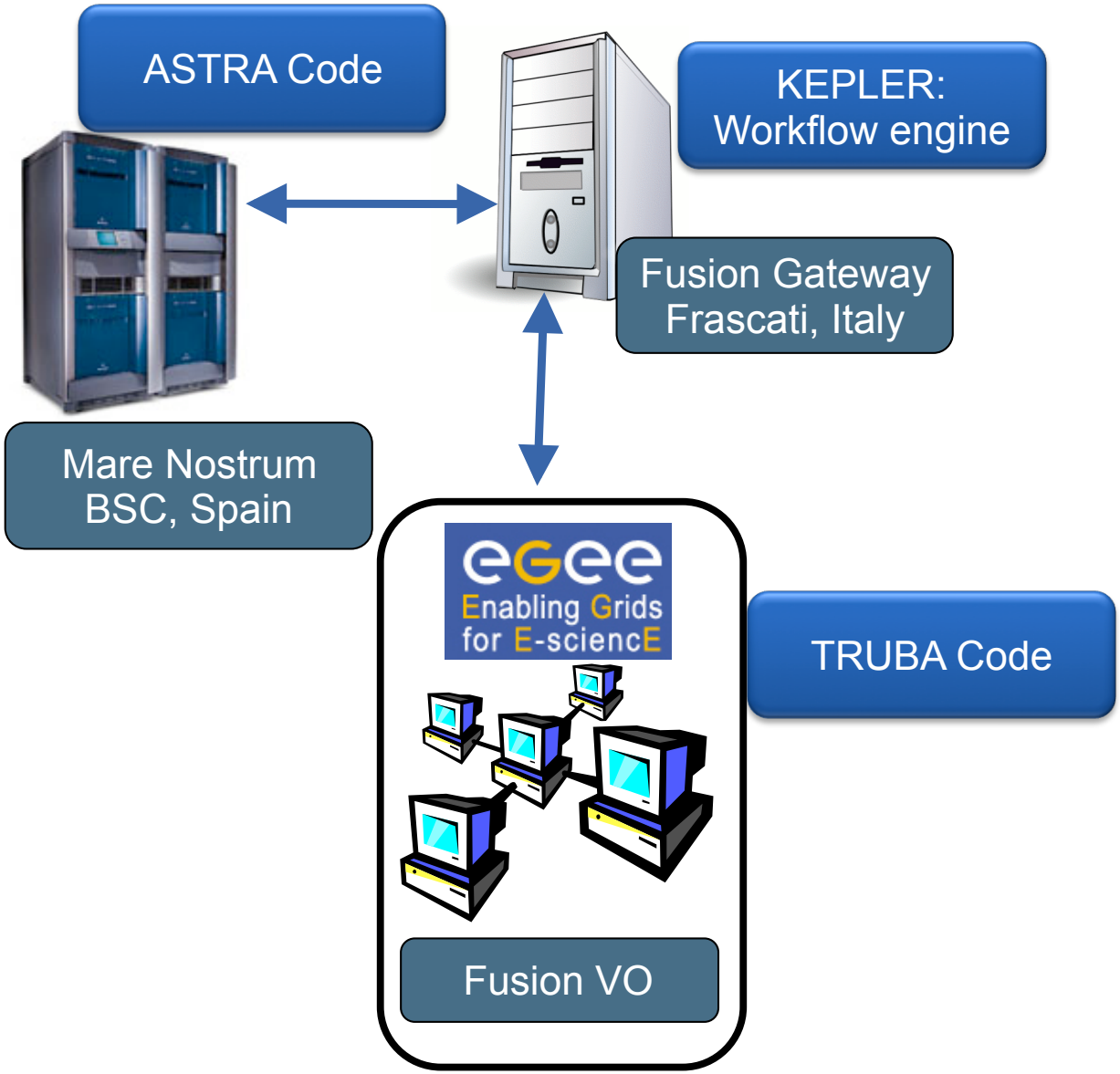
- Using top computational environments.
- Wide range of applications: serial, MPI, shared-memory, …
- Complex experiments: Necessity of connecting different models (applications) → WORKFLOWS
- Several applications running and exchanging data in different infrastructures.
- Necessity for an easy and widely known environment.
- EGEEIII: Grid used by fusion community. Distributed applications suitable for the grid. Relevant scientific output.


External Data
• HPC
• Grid
• Visualization
App1
App2
• HPC
• Grid
• Visualization
• HPC
• Grid
• Visualization
App3
KEPLER

KEPLER

- **Flexible workflow engine.**
- **It enables the communication with UNICORE and gLite.**
- **It permits establishing complex workflows:**
  - Grid- Grid
  - HPC-HPC
  - Grid-HPC
- **The applications are actors launched by Kepler.**
- **Kepler is a de-facto standard in fusion community.**

System Architecture

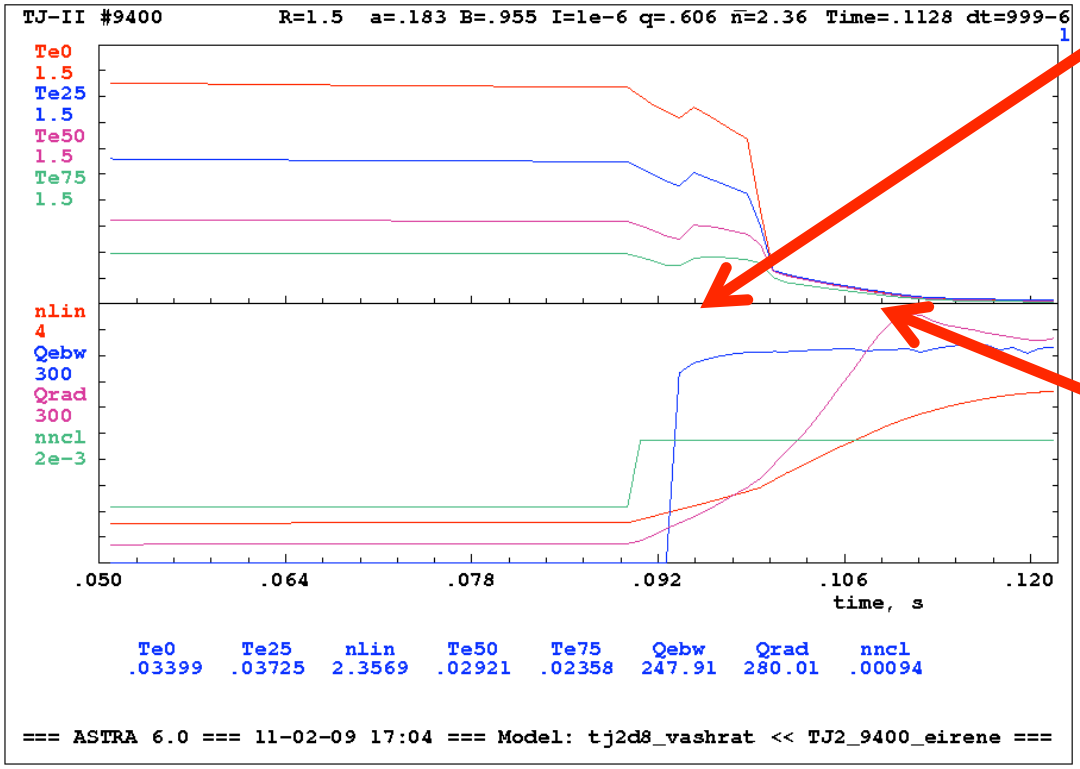


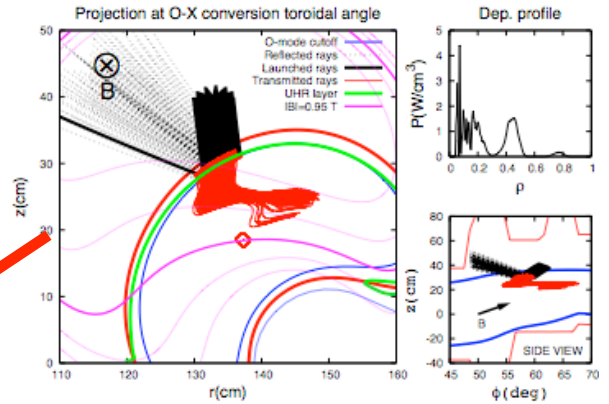
- KEPLER LAUNCHES THE DIFFERENT ACTORS AND ORGANISE THE WORKFLOW.
- ONE ACTOR (ASTRA) RUNNING ON HPC AND THE OTHER (TRUBA) ON THE GRID.

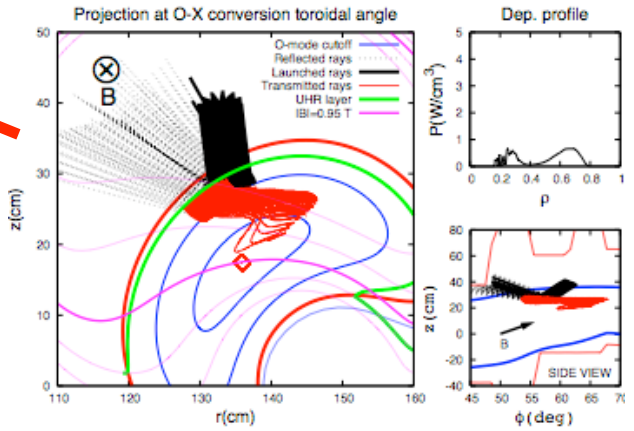
Live DEMO

Example of Workflow

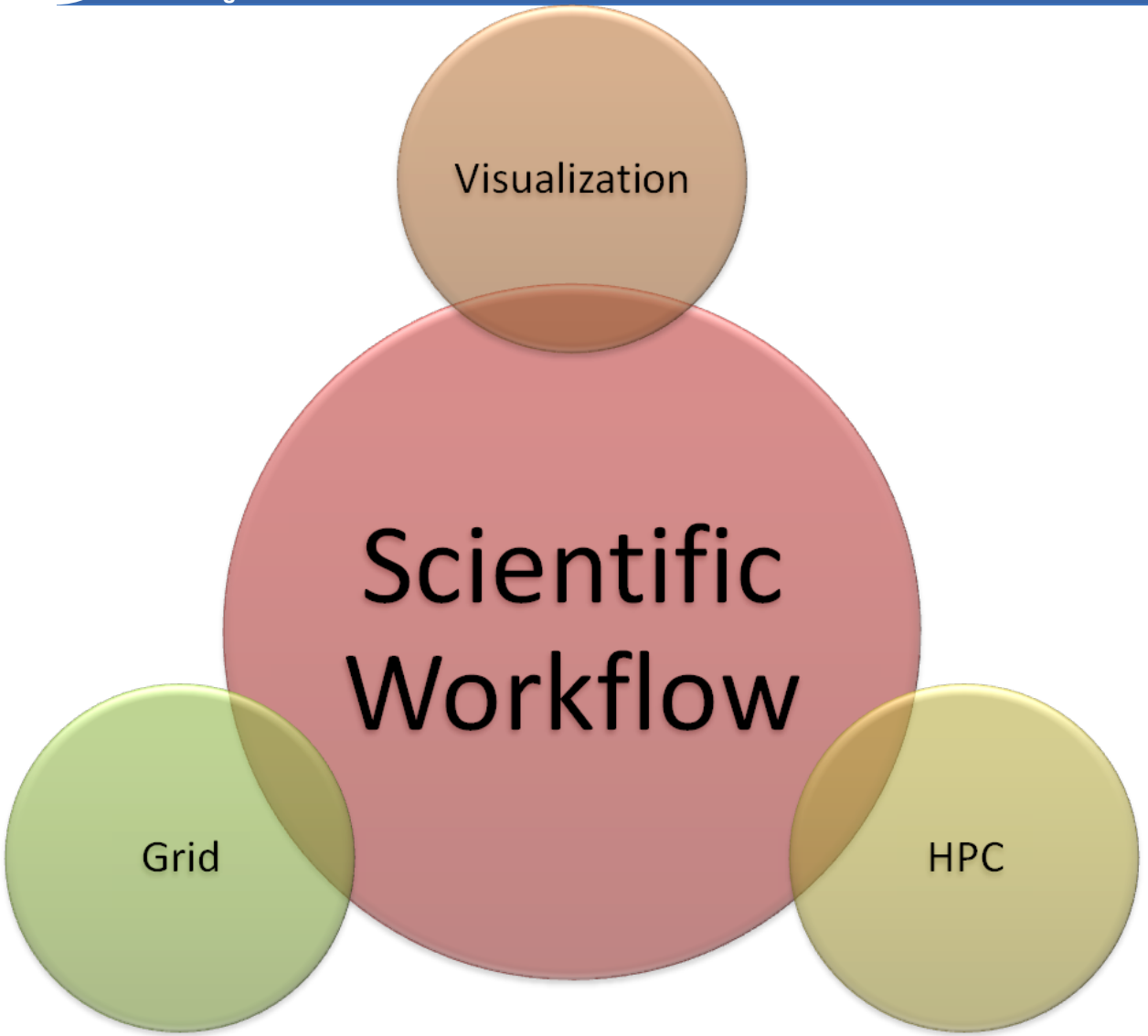
Plasma (n, T) →Plasma Evolution (ASTRA-HPC)→
Heating properties (TRUBA-grid)→
Plasma Evolution (ASTRA-HPC)→
Heating properties (TRUBA-grid)→…







CONCLUSION: A NEW COMPUTING SCHEME READY



CONCLUSIONS

- **After EGEE Projects, Fusion has become a HUC of the grid.**
- **More than ten fusion applications running on the grid:**
  - Covering different fusion research topics.→ IMPORTANT FOR THE DIVERSITY OF WFS.
  - Using several parallelization strategies.
- **New computing paradigm in fusion that establishes workflows of heterogenous applications running on different architectures.**
- **Relevant scientific results on the grid: 13 fusion papers in peer reviewed journals, including a PhD thesis.**
- **The results of the workflow here presented have been published in:**

Á. Cappa, D. López-Bruna, F. Castejón, et al. "Calculated evolution of the Electron Bernstein Wave heating deposition profile under NBI conditions in TJ-II plasmas" Accepted in Contributions to Plasma Physics, 2010.