ITER: a driver for e-Science and e-Infrastructures?

Pär Strand,

Chalmers University of Technology

With input from David Coster (IPP)
ITER, ITM-TF, EUFORIA, MAPPER and ITERIS

Who am I?

Past:

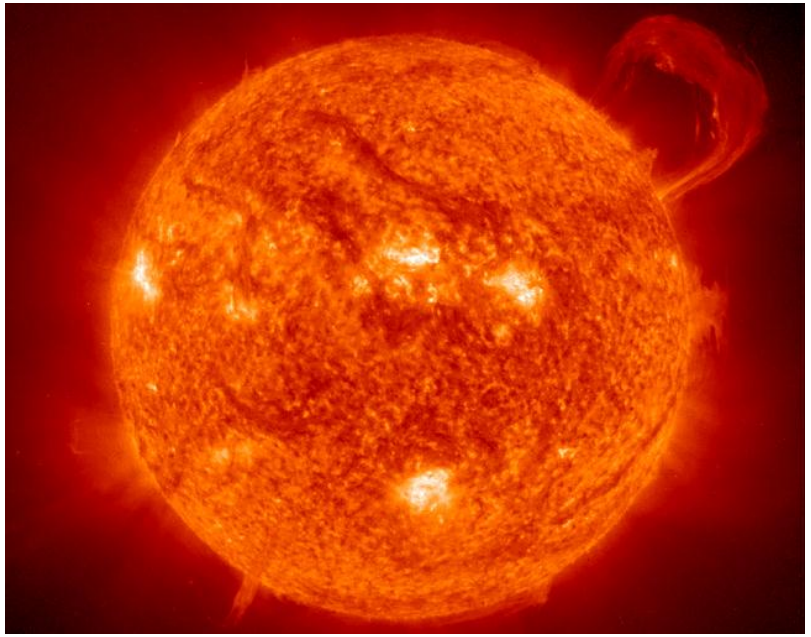
- Former Task Force Leader: Integrated Tokamak modelling task Force (2004-2010). Building the EU analysis software for ITER
- Coordinator EUFORIA: EU Fusion for ITER Applications (2008-2010). Bringing e-infrastructures to EU fusion community
- Chair: ITER Integrated Modelling Expert Group (2008-2010). Building a consensus on ITER IM needs and requirements with ITER partners.

Current:

- Director, Chalmers e-Science Centre
- Member ITERIS consortium (2010-2013): Contract for developing ITER Integrated Modelling infrastructure
- Contributor to MAPPER (2010-2013)

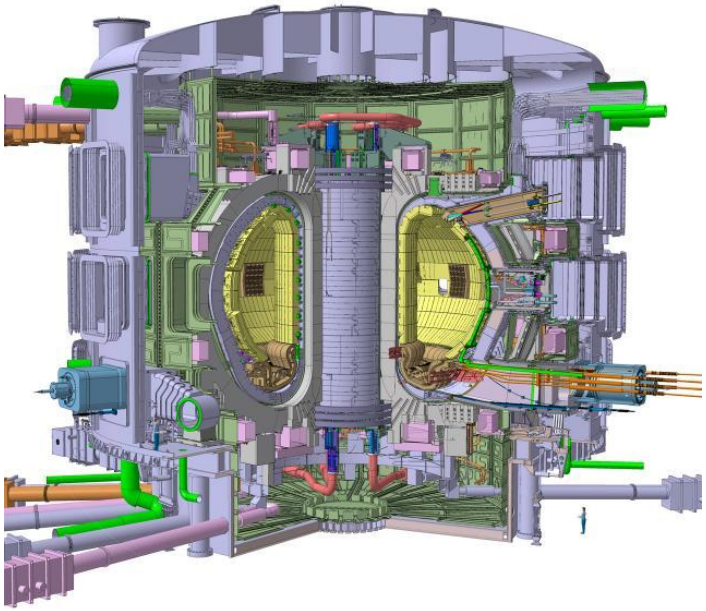
Fusion

- Energy source for the sun and other stars
- Provides a potential source of base load energy production
- Been working on this for more than 50 years
- Has turned out to be a very difficult problem



"Every time you look up at the sky, every one of those points of light is a reminder that fusion power is extractable from hydrogen and other light elements, and it is an everyday reality throughout the Milky Way Galaxy."
--- Carl Sagan, Spitzer Lecture, October 1991

"ITER aim is to demonstrate that it is possible to produce commercial energy from fusion."



First plasma 2019, full operation 2026 (!)

Experimental facility(*):

- 10Gbit/s during discharges, 500-1000s
- 20-100PB/year

*lower bound estimates

International partners:

- Data replication several offsite repositories
- (Near) real time data streaming, inline
- modelling data to/from remote centers

Current status - construction site.

A number of elements
are being defined already now!

- Data (access and ontologies)
- Modelling infrastructure

- "Semi" remote operation
- Middleware interoperability

→ agreement on single technology (most interfaces will be centrally managed/decided!)

- Resource sharing/policies
- IPR challenging issue.
- ~3000-4000 remote participants

— **Nuclear installation(*):**
— Security
— licensing

*Generally only an issue locally for ITER.

ITER shorter term modelling needs (during construction)

- Physics design studies – modelling of critical design issues
- Implementation, integration and testing of plasma control system
- Modelling for diagnostics development
- Physics Scenario assessments and development

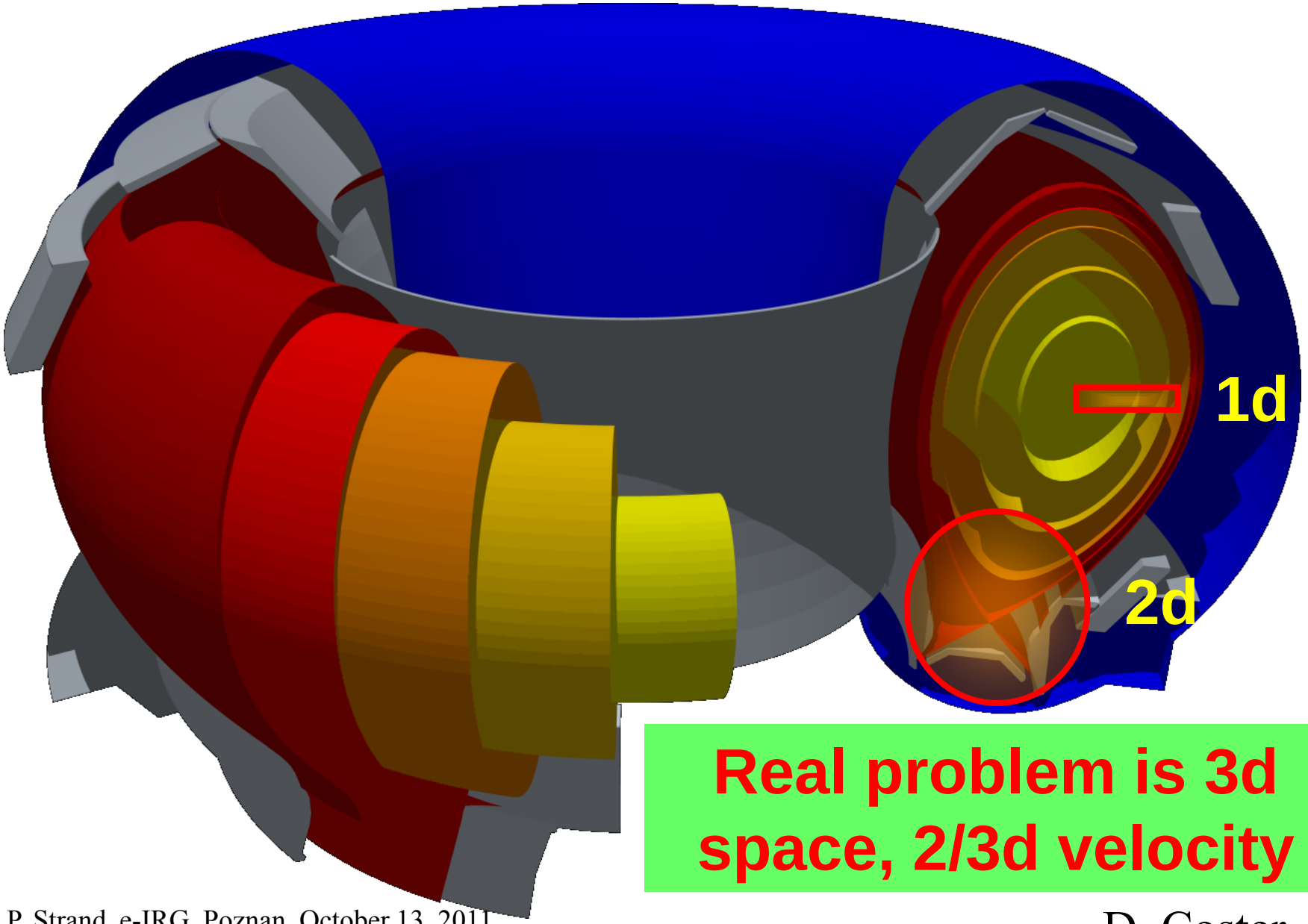
Impact through modelling (preparations and operations)

- Safe and optimal ITER operation will rely on a high degree on physics modelling and simulation
  - Not funded directly by ITER - modelling capacity derived from partner programmes (EU, US, JP, CHINA, RU, INDIA, S. Korea)
  - ITER modelling very challenging from computational point of view – will likely require heterogenuous resources!

Competitiveness

- ITER Experimental time allocated through competitive proposals
  - Modelling integral and essential component in proposal process!
  - Pan-European structure needed to compete with national programme structures in US and JP in particular
  - High end modelling leads to scientific edge

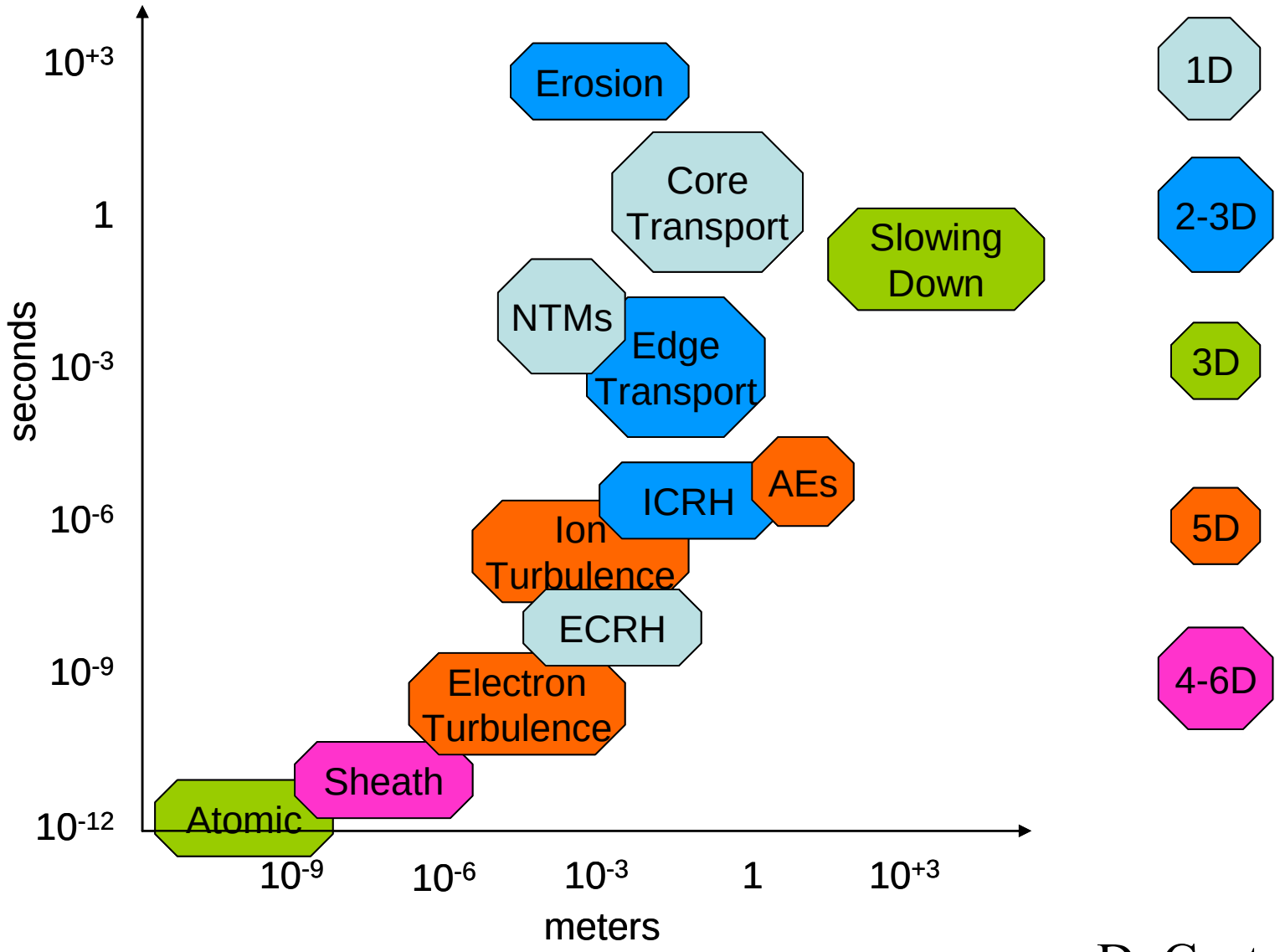
Simulations



Real problem is 3d space, 2/3d velocity

D. Coster

Models describing the plasma vary in complexity

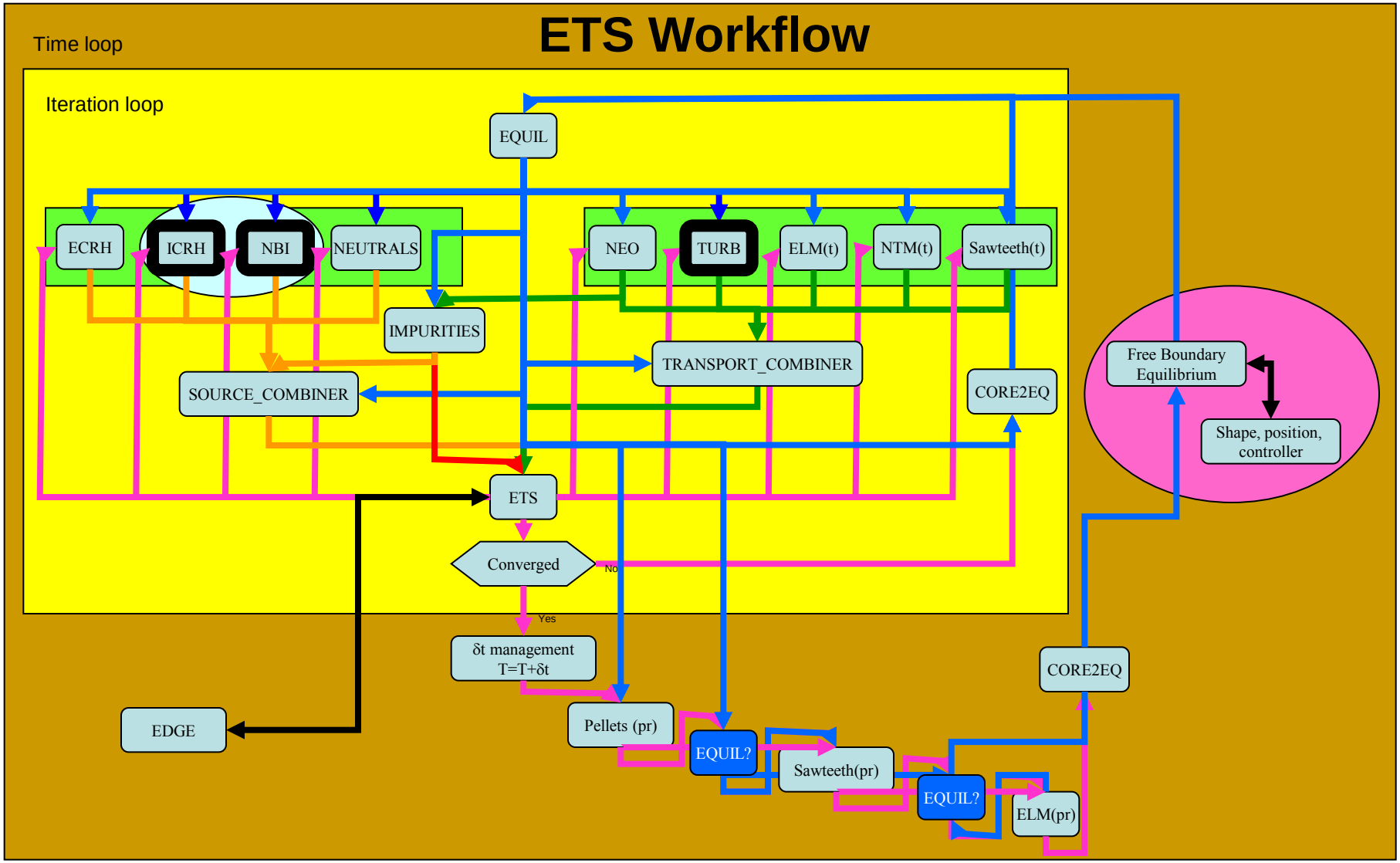


D. Coster

Several Use Cases and application structures in parallel

- **Experimental analysis Chain –**
  - Loosely coupled physics modules set up to analyse experimental data - DAG structures
- **Predictive modelling**
  - Plasma evolution on transport timescales
  - Heterogeneous computing and physics coupling requirements – iterative and complex interactions between physics modules varying time scales and dimensionalities
- **First principles modelling**
  - petascale towards exa-scales for full integration.

European Transport Solver



Impact on e-infrastructure

- ITER Modelling Framework (IMAS) shall be operational well before ITER Operation: first prototype needed end of 2013 for starting to test Plasma Control System software and algorithms
- IMAS shall accompany Operation and Research over the ITER lifespan (~ 30 years total)
  - Changes in computer/software technologies
  - Changes in physics understanding and methods to solve physics problems
  - → The IMAS shall be flexible and extensible, both in terms of physics components and Infrastructure
- A prototype IMAS Infrastructure/framework technology has to be chosen shortly (beginning 2012)
- Its structure shall allow for future evolutions and possible changes of technologies + inclusion of distributed resources

→ Do NOT underestimate inertia – likely the infrastructure that we need to link in our e-infrastructures into.

Current Status - Players

- EURATOM (EFDA), Fusion 4 Energy (F4E EU procurement agency for ITER) , ITER+ ITER partners, INFSO, National and regional funding agencies.

- Modelling/Modelling Infrastructure
  - EUFORIA (INFSO, ended December 2010 –pushing e-infrastructures for fusion)
  - Integrated Tokamak Modelling Task Force (Euratom) (ongoing – 60ppy/year physics)
  - Mapper (Multiscale simulation software, fusion one of application areas)
  - ITERIS (EU Framework contract on IM infrastructure w ITER – 2010-2013)

- Computing Hardware (mainly used for single applications NOT Integrated modelling)
  - EGI and PRACE, local and national access
  - HPC-FF (100TF fusion part of JUROPA until end 2013)
  - IFERC computer(1.3PF user access from 1 April 2012 to 2016)

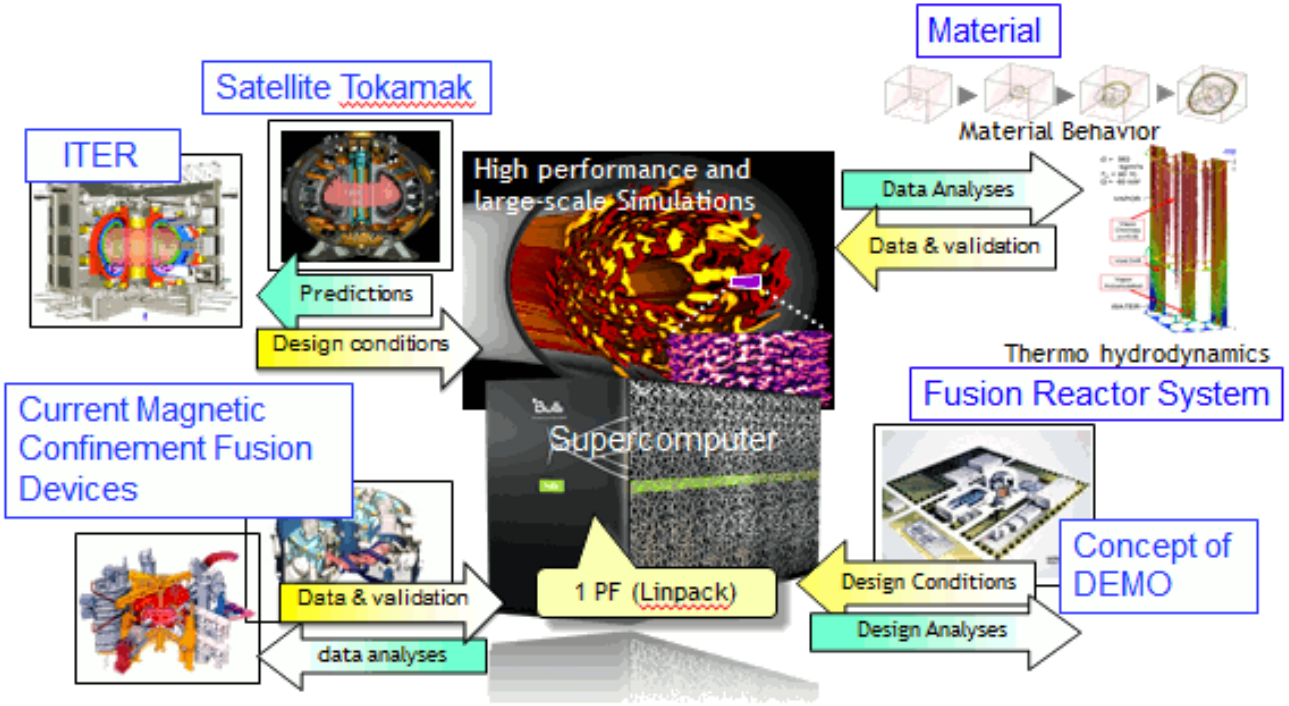
- Storage and data access
  - No real activities except for local developments at ITER CODAC (ITER lead)
  - ITERIS will make first recommendations on simulation data by end this year.
  - EUDAT? Very minor – keep in touch activity only!!!

- Connectivity
  - Later needs: 2020? F4E, Geant, NRENs (needed when? Who is the driver here?)

- Distributed access and policies.
  - ???

International Fusion Energy Research Centre (IFERC)

- DEMO Design and R&D Coordination Centre
- ITER Remote Experimentation Centre
- Computational Simulation Centre (CSC)
  - **4410 nodes (**16 cores with 58 memory).
  - Peak performance of 1.30 PF,available memory of 256 TB.
  - Available for users (EU + JP) April 2012

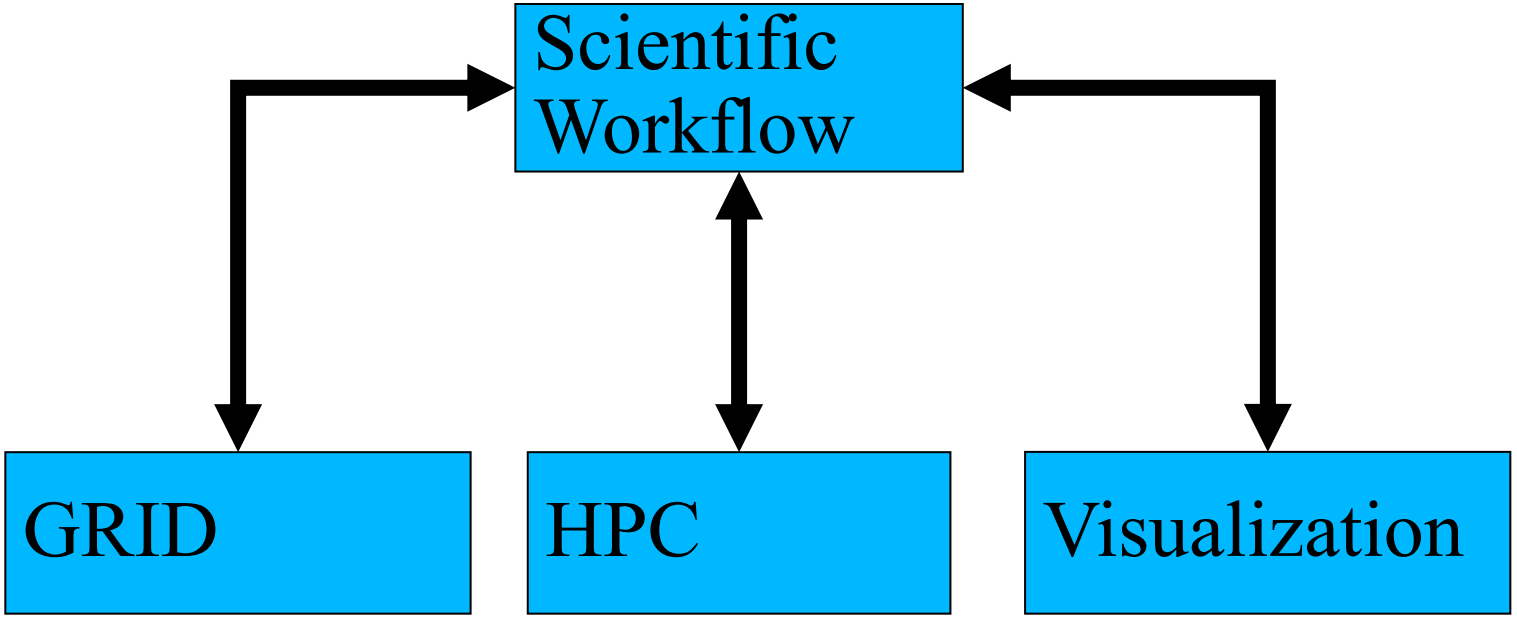


www.iferc.org

EUFORIA Activities

- Satisfied EUFORIA user community
  - 550 training days provided, More than 50 publications from users
  - 10 million HPC hours provided
- Complex workflows established across range of application scenarios/types (Grid serial, Grid parallel, parameter scan, HPC, …)
- Significant parallel performance improvement in high impact fusion codes. Continued in EFDA (HPC-FF), PRACE and CRESTA projects.
- Workflows providing transparent and distributed access to Grid, HPC, and Cloud resources. Hiding infrastructure from users.
  - Including EGEE-EUFORIA-DEISA pilot project – TRANSPARENT ACCESS over infrastructure boundaries
  - Partially continued through MAPPER activities
- Extensive uptake in fusion community (and strong interest from ITER) of EUFORIA developed visualisation and access tools

Developing a new paradigm for fusion computing



- Building on e-infrastructure tools, middleware and installations
- Integrating tools and physics models together with a ”fusion simulation ontology”
- (At least initially) building on fusion de facto standards for data access and communication
- Largely accepted by ITER as a model fro framework implementation

Joint taskforce between MAPPER, EGI, and PRACE

- Collaborate with EGI and PRACE to introduce new capabilities and policies onto e-Infrastructures
- Deliver new application tools, problem solving environments and services to meet end-users needs
- Work closely with various end-users communities (involved directly in MAPPER) to perform distributed multiscale simulations and complex experiments

LESSONS LEARNED

- A number of technical issues/developments
- ”General purpose European infrastructure” is a complex issue. In particular, domain specific demands on minimum common resources vary significantly between application areas.
  - Only small subset of EGEE grid usable by memory hungry EUFORIA applications → EuFORIA maintained its own resources
  - Middleware(s)! Wishing for
    - Compact and maintainable,
    - Scalable and extensible
    - Robust and reliable
    - Easy to use…and replaceable
  - Authentication (single sign over multiple infrastructures)

Lessons Learned: SERVICES vs USERS

- Even if services are available access to (HPC) resources may be too restrictive due to policy issues.
- Advanced or novel access patterns rapidly emerge as you allow the application needs to take the central place – not the service itself. (Generic issue not only for fusion)
  - Ability for advance co-reserve of resources
  - Launch emergency simulations
  - Consistent interfaces for federated access
  - Access to back end nodes: steering, visualisation
  - Data integration from multiple sources
- "e-infrastructures are impotent without proper training!"
- Cultural divide: We have already invented all wheels necessary (services)…OR…The clothes don't fit please retailor them (applications)

Opportunities/challenges!

- Network connectivity – global scale but still some times away
- Governance models (several EU agencies, number of international partners, ITER IO…,)
- New, enhanced role of modelling and analysis
  - Integral with machine exploitation
  - Extreme range of resource needs (from smaller local ITER resources to PRACE level installations in ITER partners… and beyond), multiscale → heterogenuous needs!
  - Complex range of interdependent tools required for even basic understanding level - workflow organization ~100 interacting apps.
- Data access and storage (distributed exploitation!)
- Data provenance and QA
- Large international user base – and ”ownership”
- Thematically well aligned with e-infrastructures scope and possibly strong need for connectivity, but,
  - HOW TO BRING IT ALL TOGETHER TO A SUCCES STORY FOR SCIENCE DRIVEN e-INFRASTRUCTURES?

Summary

- A number of prototyping activities are ongoing or being finalized.
- Early days still, BUT
  - Some elements are already being defined or settled now
  - Largely relating to the local infrastructure
- Potential areas for e-infrastructure input/impact
  - Local access not sufficient
  - Distributed computing resources and modelling/analysis landscape
  - Global user base and data sharing
  - Challenge is to put ITER in the hands of the users
- ITER: a driver for e-Science and e-Infrastructures? Yes – but we need to get involved now!
- Time to influence/review/input from e-infrastructure point of view - ITERIS project sensible point of contact.