…ye who enter here…

Pär Strand,

Chalmers University of Technology

Drawn from experiences with
EFDA, ITM-TF, EUFORIA, MAPPER and ITERIS and others
Reused materials from all projects – thanks!

- Title is a “tongue in cheek” observation of
  - the challenges in modelling fusion plasmas
  - some of the frustration in (successfully) bringing in a new (sub)community into the EU e-infrastructures world
  - Getting people from very different backgrounds to jointly advance capacities way beyond state of art
- Hopefully there will be some take home observations along the way

Outline

- Fusion in general
- Challenge of predictive modelling of fusion reactors
- Integrated Fusion modelling in EU
  - EFDA integrated tokamak modelling Task Force
  - EUFORIA – EU Fusion for ITER Applictions
  - MAPPER – Multiscale application on EU e-infrastructures
  - ITERIS
- Lessons learned: Sociology, policies, sustainability and other challenges
- Summary

Nomenclature

If I mention

- "GRID", I generally mean High Throughput Computing on a gLite based (or similar) middleware
- "HPC" I generally mean High Performance computing as on a CRAY (or similar) somewhere in a single computing centre

Non-standard nomenclature in this community but what it has grown to mean within the fusion community…

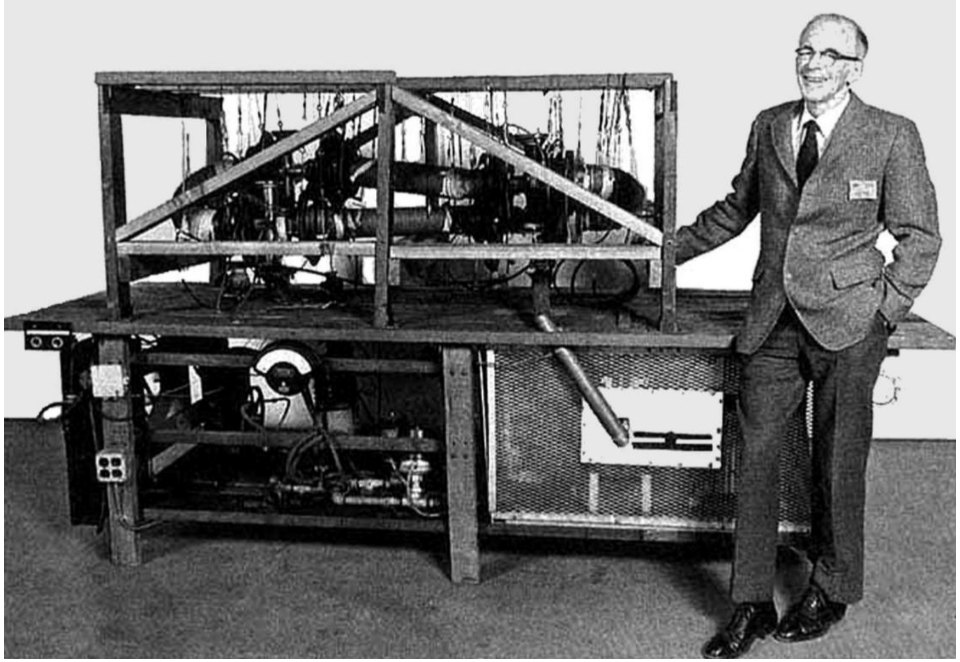
Who am I?

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)

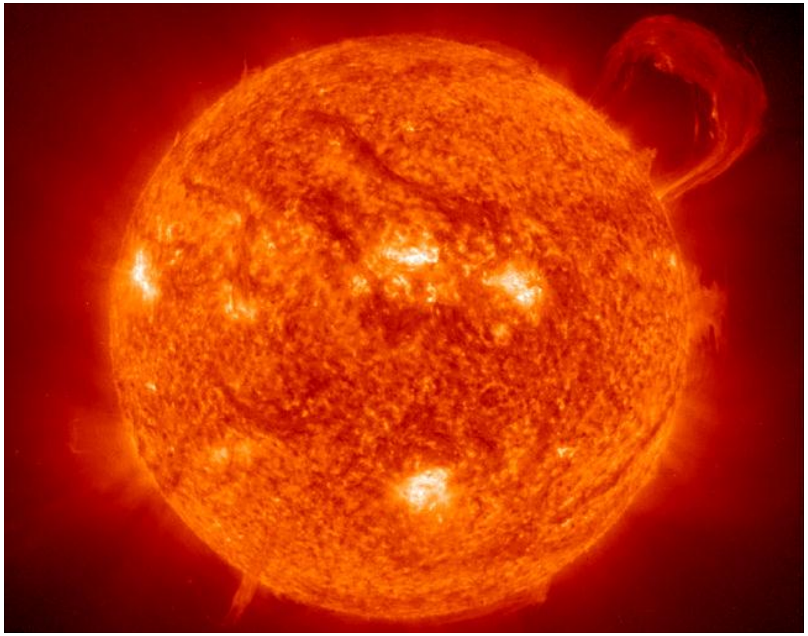


GENERAL INTRODUCTION TO FUSION ENERGY

P. Strand, e-science 2011, December 5-8, Stockholm

Fusion

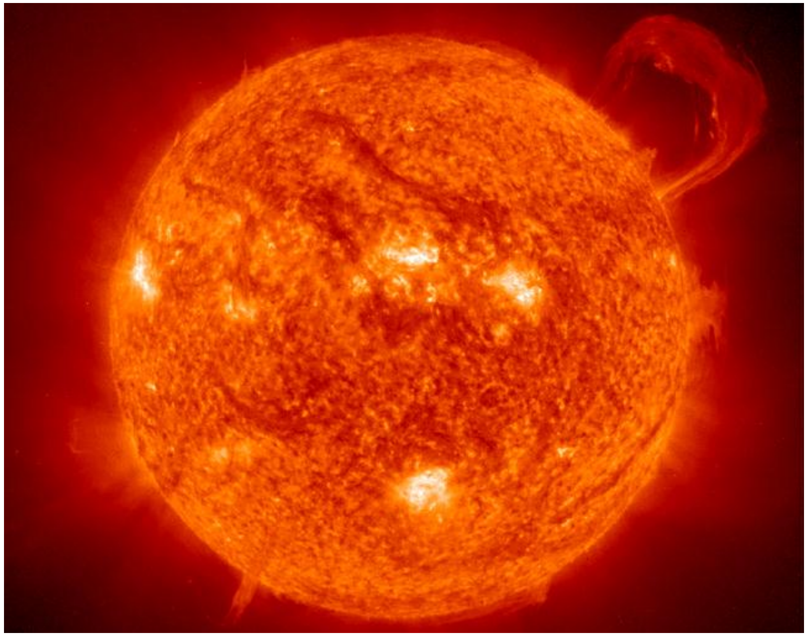
- Energy source for the sun and other stars
- Provides a potential source of base load energy production
- Been working on this for more than 50 years
- Has turned out to be a very difficult problem



"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

Fusion

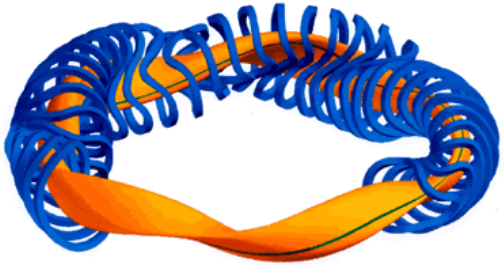
- Fuel cycle in sun requires extreme time and length scales
  - Not available in labs
- Need to find other alternatives with sufficiently large fusion triple product: $nT\tau$

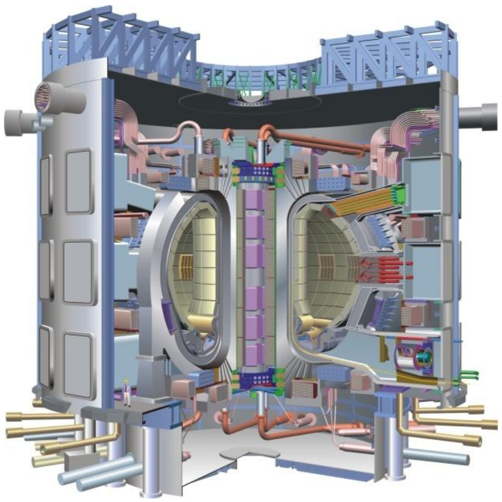


"Every time you look up at the sky, every one of those points of light is a reminder that fusion power is extractable from hydrogen and other light elements, and it is an everyday reality throughout the Milky Way Galaxy."
--- Carl Sagan, Spitzer Lecture, October 1991

Fusion

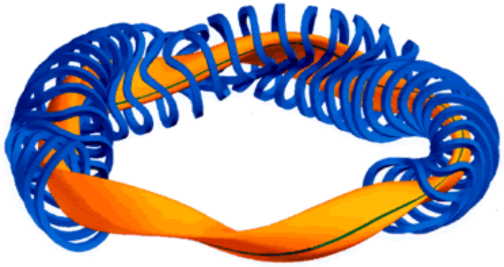
- Two main lines of research
  - Inertial confinement
    - Implosion of small pellets
    - NIF at LLNL
  - Magnetic confinement
    - Two main type of configurations studied:
      - Stellarator – W7X
        - Currently under construction in Greifswald in Germany
        - Steady state device
      - Tokamak – ITER
        - Under construction in Cadarache in France
        - Inductive, pulsed device

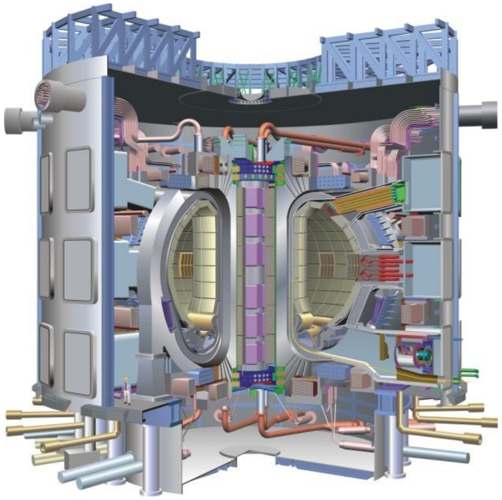




Fusion

- Two main lines of research
  - Inertial confinement
    - Implosion of small pellets
    - NIF at LLNL
  - Magnetic confinement
    - Two main type of configurations studied:
      - Stellarator – W7X
        - Currently under construction in Greifswald in Germany
        - Steady state device
      - Tokamak – ITER
        - Under construction in Cadarache in France
        - Inductive, pulsed device





D + T → ^{4}He (3.561) + n(14.029)

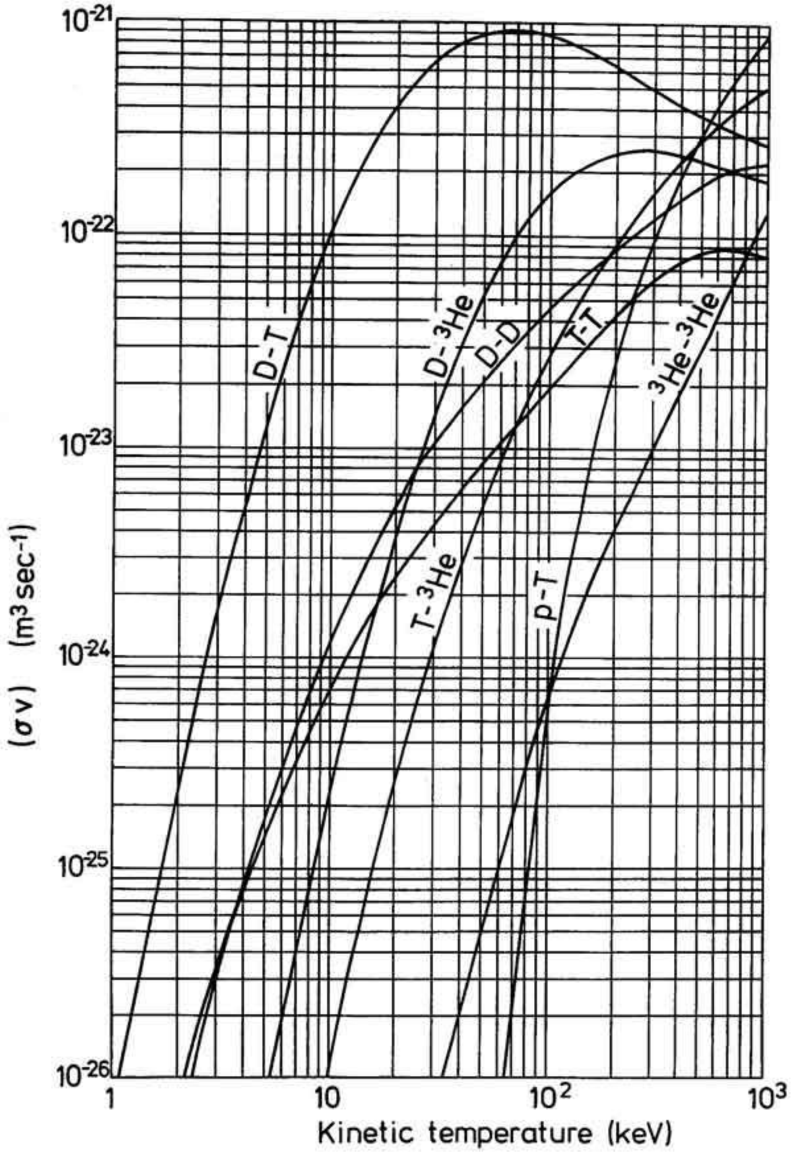
D + D → T(1.011) + p(3.022) (50%)

D + D → ^{3}He(0.820) + n(2.449) (50%)

p + T → ^{3}He + n − 0.764

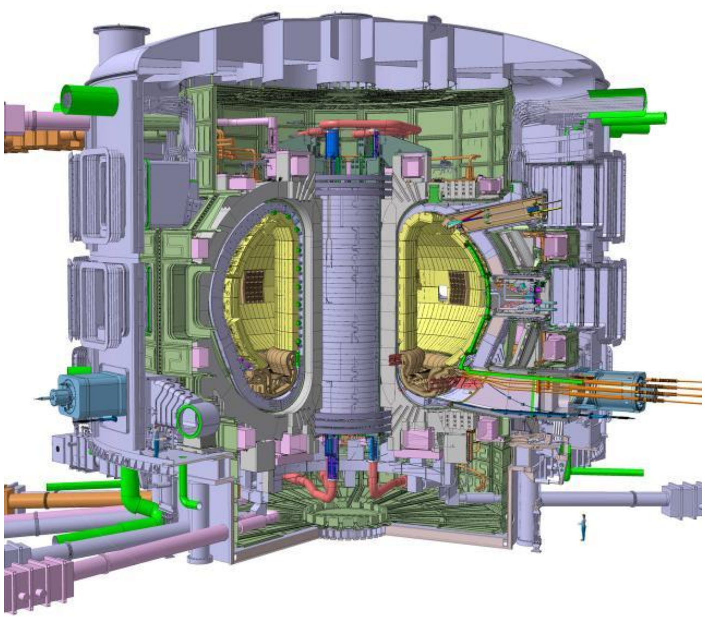
T + T → ^{4}He + 2n + 11.332

D + ^{3}He → ^{4}He(3.712) + p(14.641)



Sun: 10 Million degrees, Fusion: 100 Million degrees (in vicinity of materials)

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

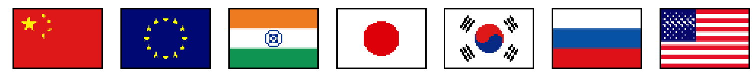


First plasma 2019, full operation 2026 (!). Make or break time for fusion

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

- "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

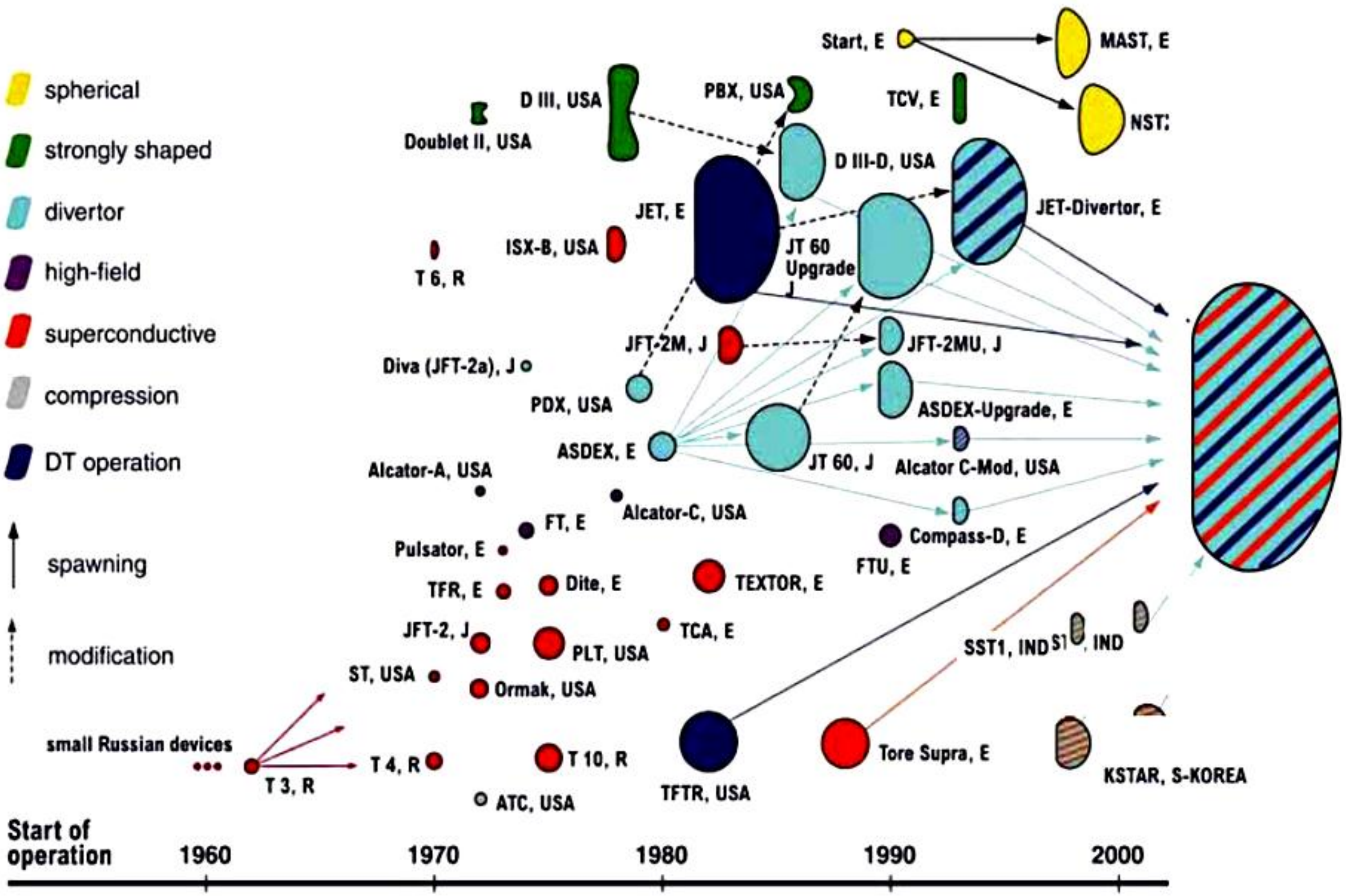
Current status - construction site.

A number of elements are being defined already now!

- Data (access and ontologies)
- Modelling infrastructure

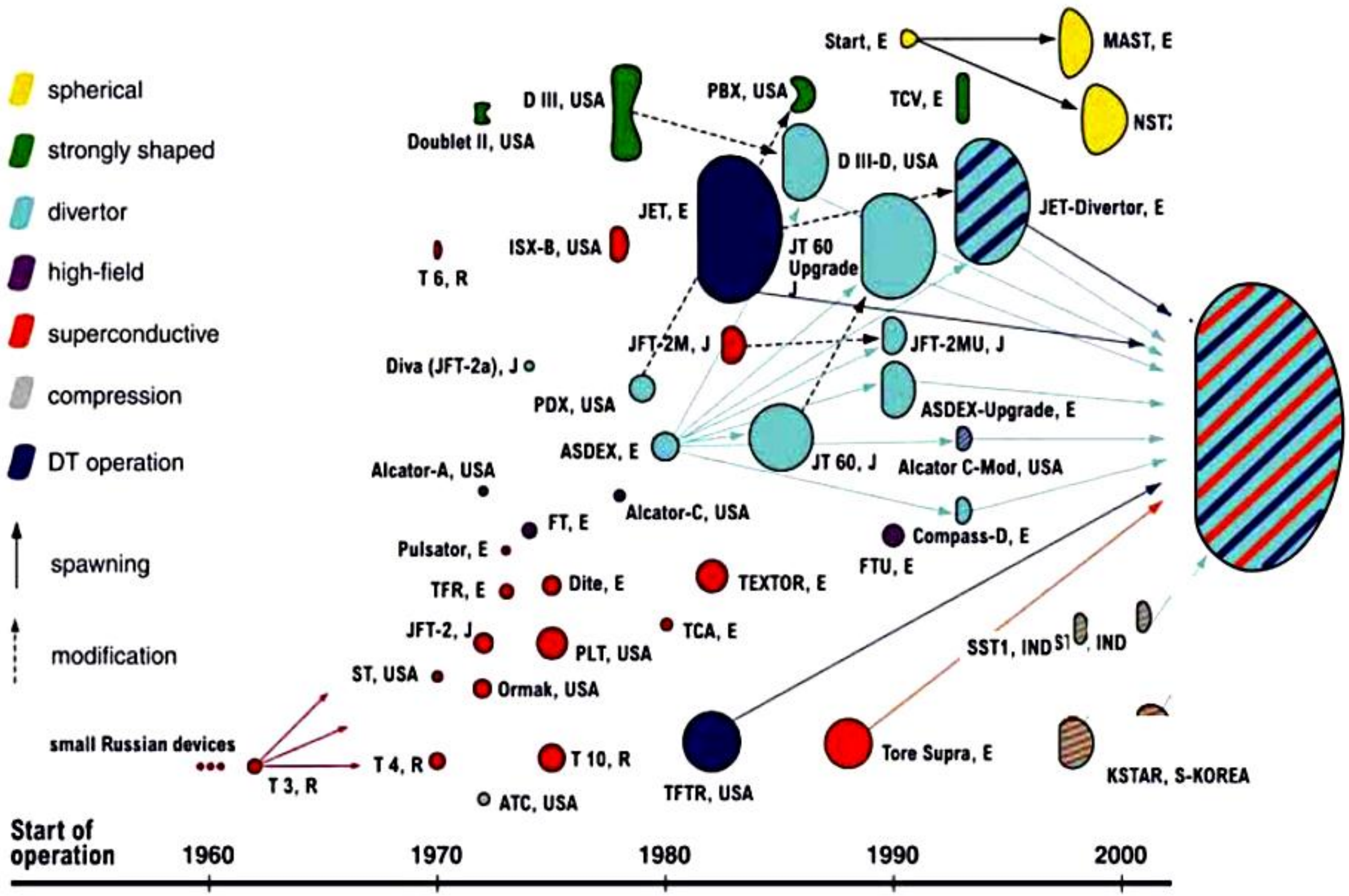
- **Nuclear installation(*):**
- Security
- licensing

*Generally only an issue locally for ITER.



ITER has developed from a long sequence of experiments BUT is the first tokamak where modelling is set to play an important role in both design and exploitation.

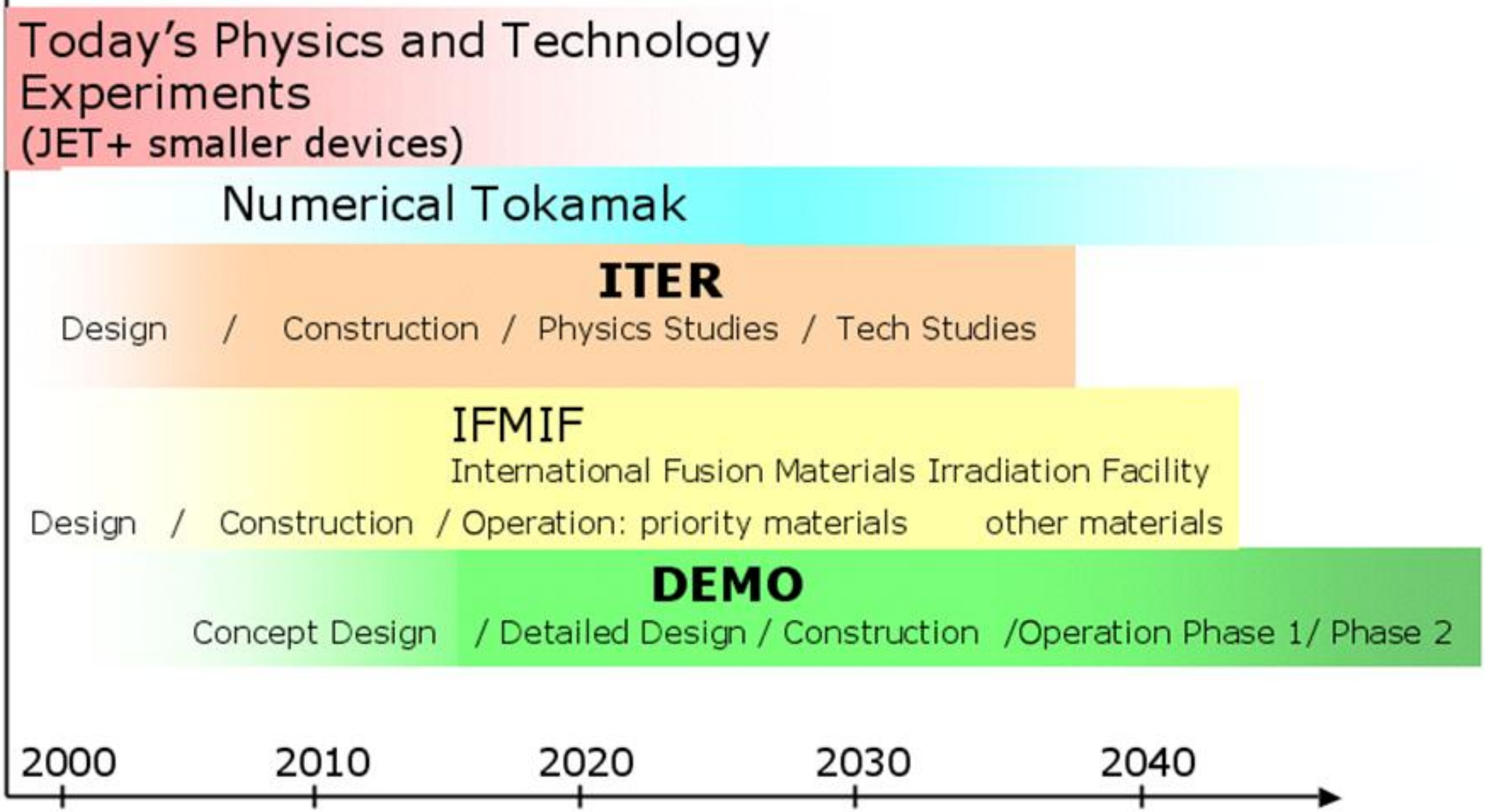
This is even more true for the next stage - a first DEMO reactor.



Bold step with a lot of foresight: US and slightly later EU (2003) – launch programmes to secure a predictive modelling capability in time for ITER. Planned <u>10-15 year activity!</u>

Idea driven by the progress in computing and physics.

Roadmap – Fast Track



The "fast track" to fusion energy – hinges on several parallel developments

"Numerical Tokamak" – a comprehensive simulation package with predictive capabilities.

"May your trails be crooked, winding, lonesome, dangerous, leading to the most amazing view. May your mountains rise into and above the clouds."

Edward Abbey

CHALLENGE OF PREDICTIVE MODELLING OF FUSION REACTORS

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)

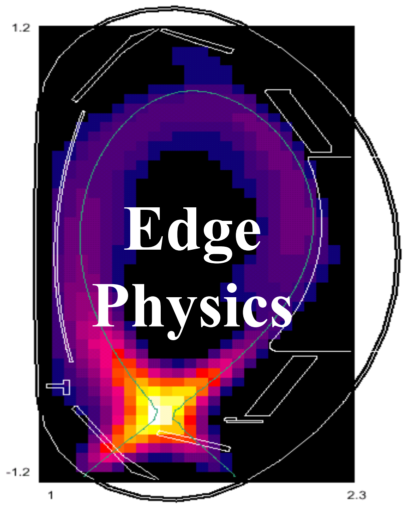
- Safe and optimal ITER operation will rely on a high degree on physics modelling and simulation
  - Not funded directly by ITER - modelling capacity derived from partner programmes (EU, US, JP, CHINA, RU, INDIA, S. Korea)
  - ITER modelling very challenging from computational point of view – will require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

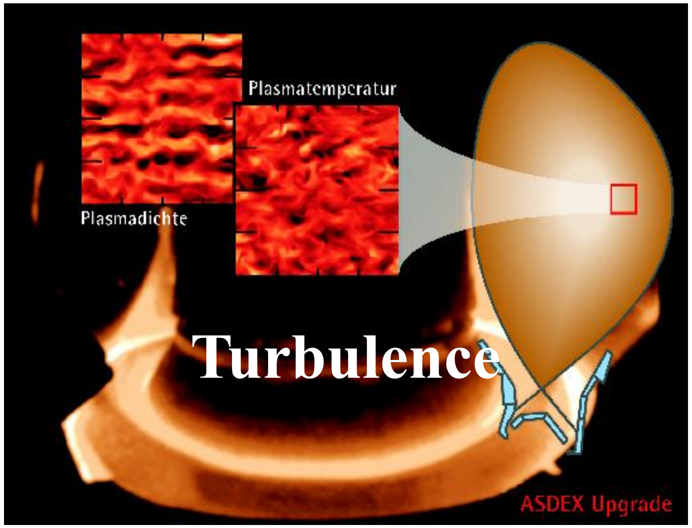
Competitiveness

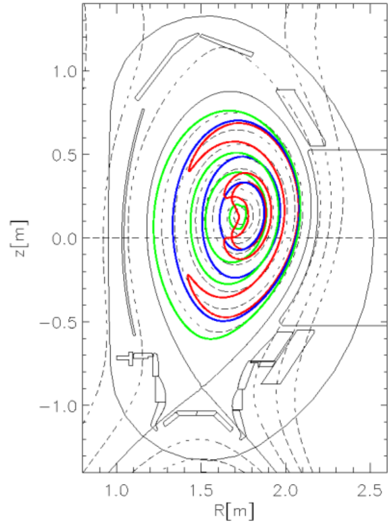
- ITER Experimental time allocated through competitive proposals
  - Modelling integral and essential component in proposal process!
  - Pan-European structure needed to compete with national programme structures in US and JP in particular
  - High end modelling leads to scientific edge

Multi-faceted physics

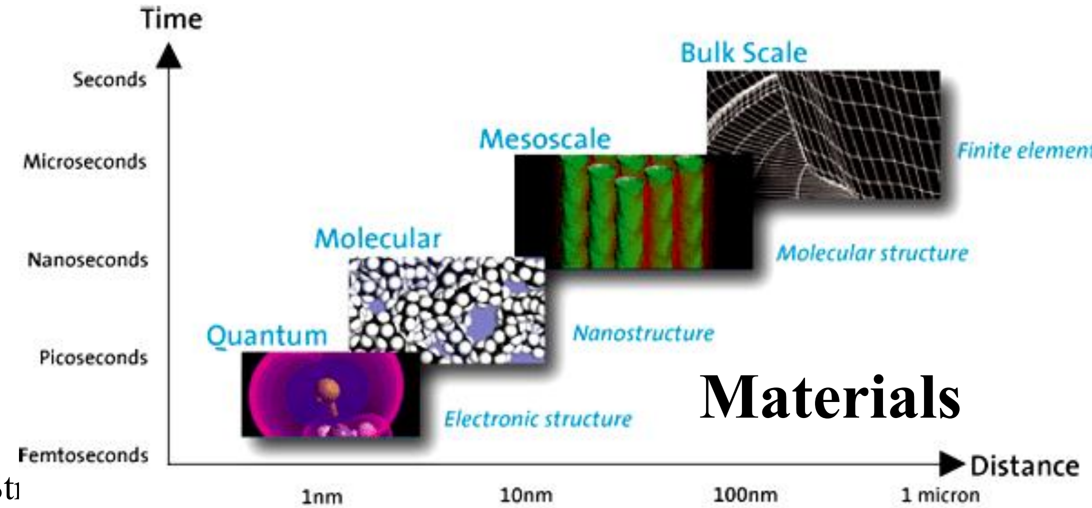
Sibylle Gunter, IPP

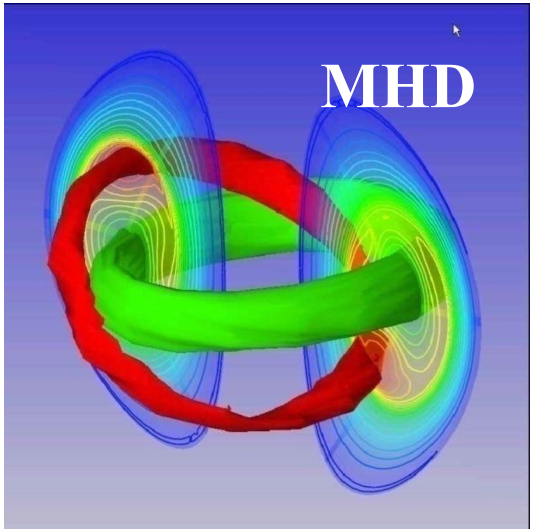






Fast Particles



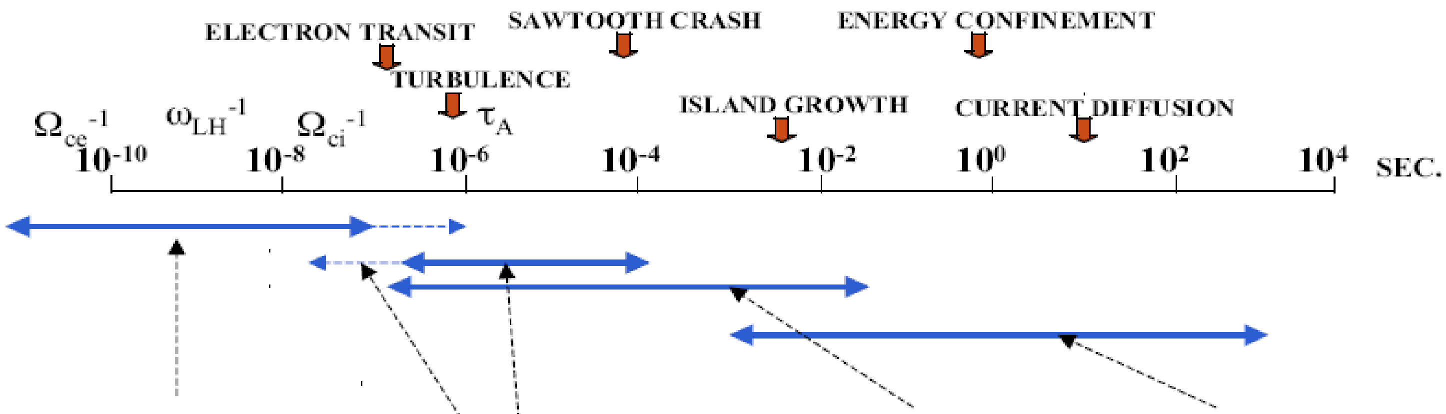


Maxwell-Boltzmann system

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t}, \qquad \nabla \cdot \mathbf{B} = 0, \qquad \mathbf{J}(\mathbf{x},t) = \sum_a \mathbf{J}_a(\mathbf{x},t) = \sum_a q_a \int d^3\mathbf{v}\,\mathbf{v} f_a(\mathbf{x},\mathbf{v},t),$$

$$\frac{\partial f_a}{\partial t} + \mathbf{v}\cdot\nabla f_a + \frac{q_a}{m_a}\left[\mathbf{E} + \mathbf{v}\times\mathbf{B}\right]\cdot\nabla_{\mathbf{v}} f_a = C(f_a)$$

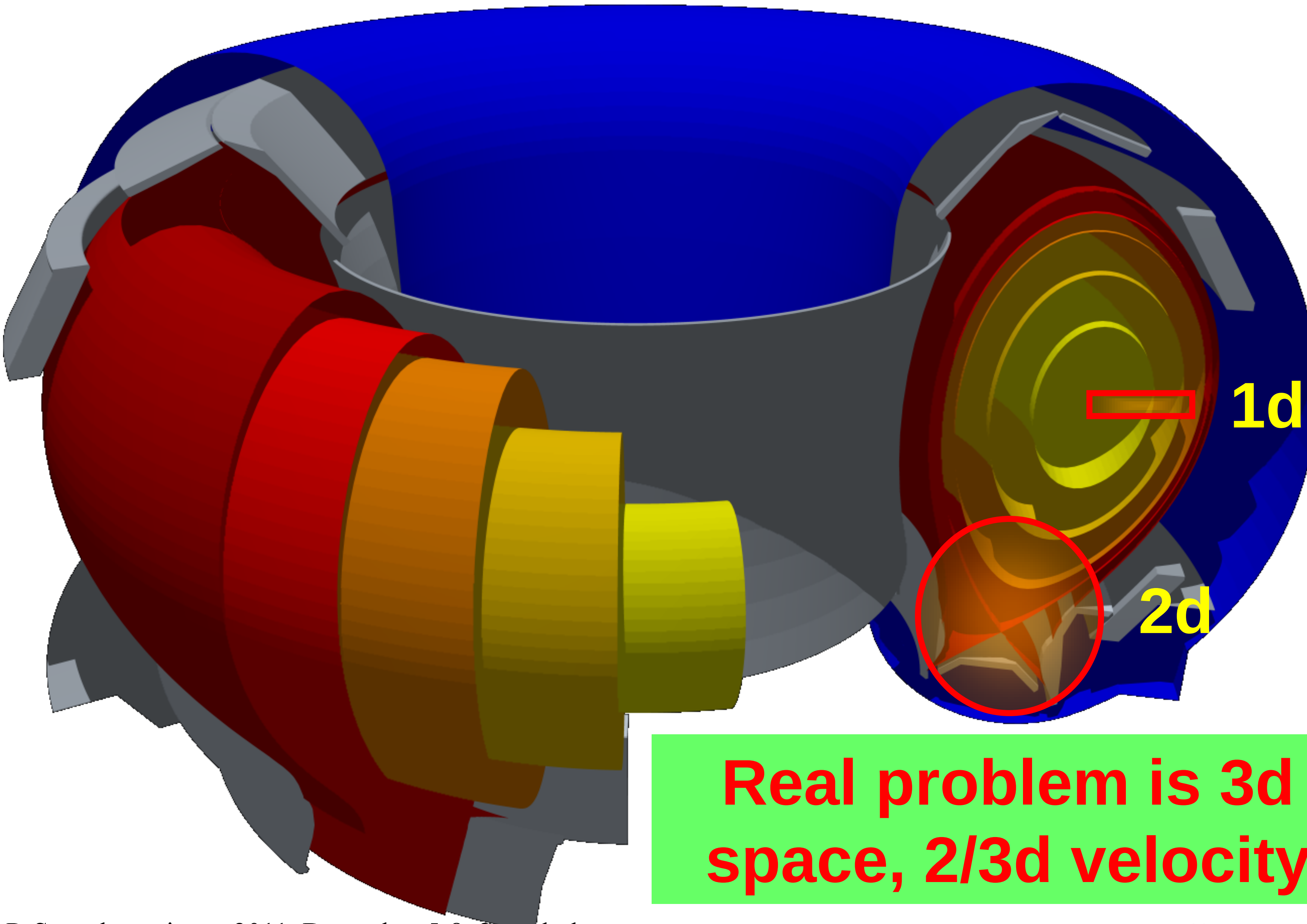
$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \qquad \nabla \cdot \mathbf{E} = \rho/\varepsilon_0, \qquad \rho(\mathbf{x},t) = \sum_a \rho_a(\mathbf{x},t) = \sum_a q_a \int d^3\mathbf{v}\, f_a(\mathbf{x},\mathbf{v},t).$$



Single frequency and prescribed plasma background	Neglect displacement current, average over gyroangle, (some) with electrons	Neglect displacement current, integrate over velocity space, neglect electron inertia	Neglect displacement current, integrate over velocity space, average over surfaces, neglect ion & electron inertia
RF Codes	Gyrokinetics Codes	Extended MHD Codes	Transport Codes
wave-heating and current-drive	turbulent transport	device scale stability	discharge time-scale

D. Post, S. Jardin & D. Batchelor, DOE

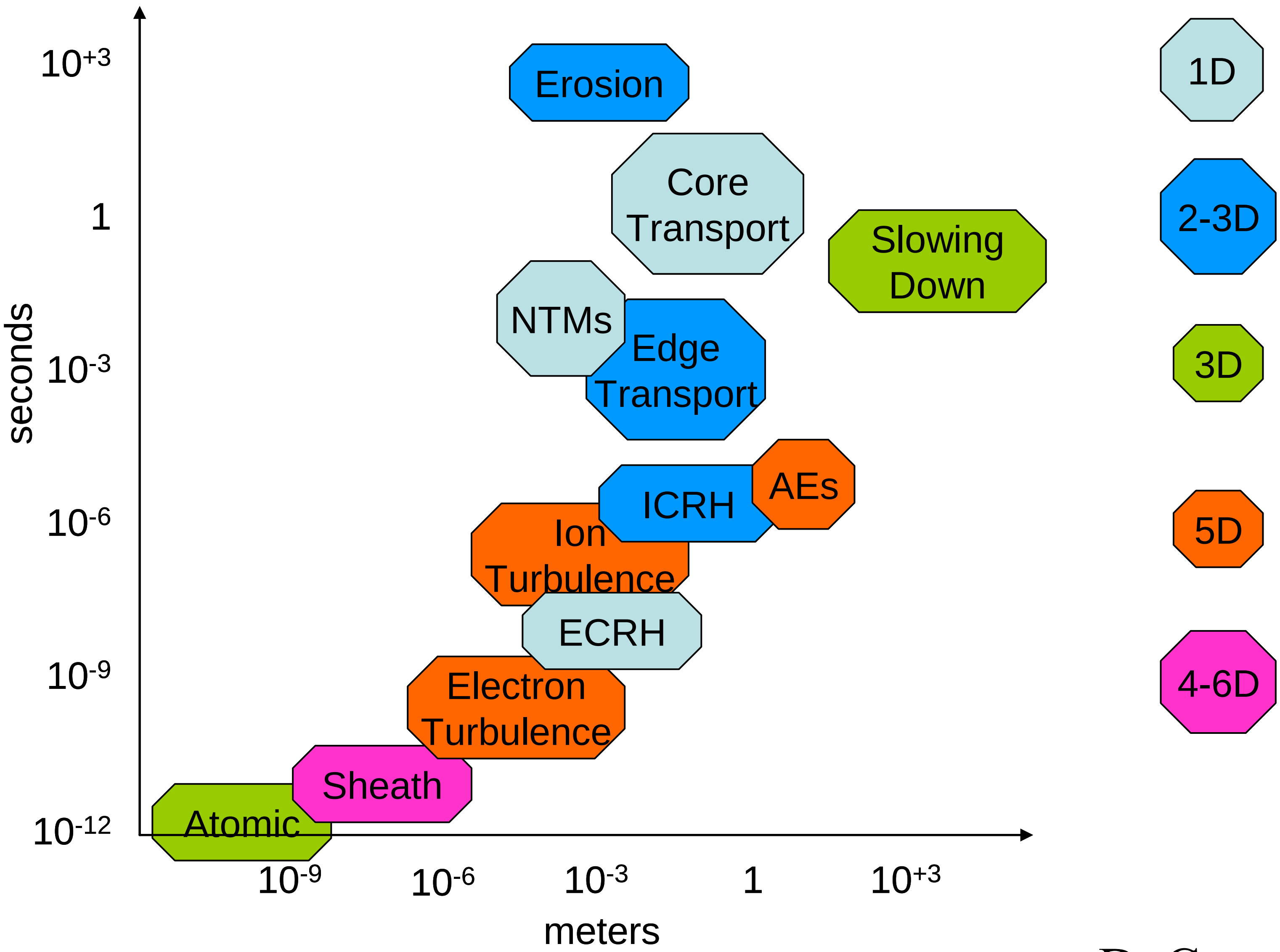
Simulations



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

D. Coster

Models describing the plasma vary in complexity

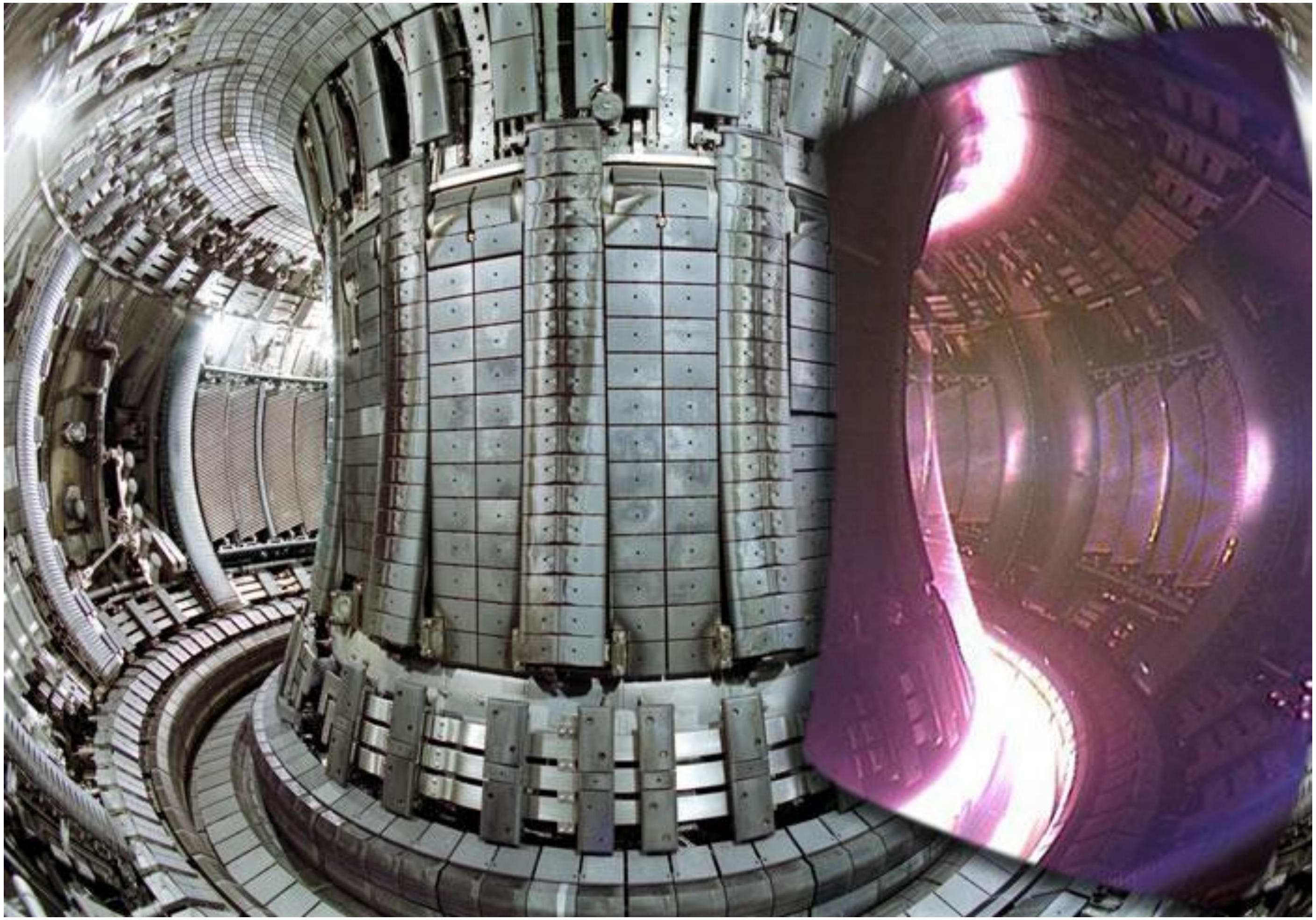


D. Coster

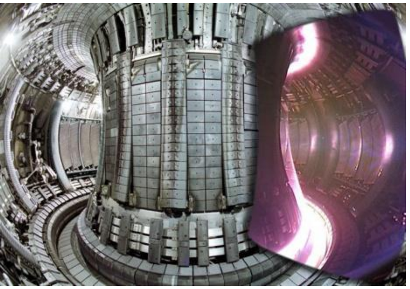
Opportunities/challenges!

- New, enhanced role of modelling and analysis
  - Integral with machine (ITER) 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!)
- Network connectivity – global scale but still some times away
- Governance models (several EU agencies, number of international partners, ITER IO…,)
- 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?

JET Pulse #64159 (EFDA-JET 2005)



CHALMERS
UNIVERSITY OF TECHNOLOGY



JET Pulse #64159 (EFDA-JET 2005)

1. **Plasma Breakdown**
2. **Current Ramp up**
3. **Flattop**
4. **External heating**
5. **Current Ramp down**

ITER will be run for 100 times longer, not all timescales follow

"Software is like entropy. It is difficult to grasp, weighs nothing, and obeys the second law of thermodynamics; i.e. it always increases."

Norman Ralph Augustine

INTEGRATED FUSION MODELLING IN EU

EU INTEGRATED MODELLING

EFDA Task Force
EUROPEAN FUSION DEVELOPMENT AGREEMENT
INTEGRATED TOKAMAK MODELLING

Preparatory Phase

EURATOM general support

Development Phase

EURATOM funding
Task Agreements

Deployment phase

2004 Expert Working Groups

Code Platform, Data Coord.+ 5 IM Projects

2006

ITM server EFDA-CSU, Garching

ITER Scenario Modelling Working Group (ISM)

2008

GATEWAY

HPC-FF

2010

EUFORIA

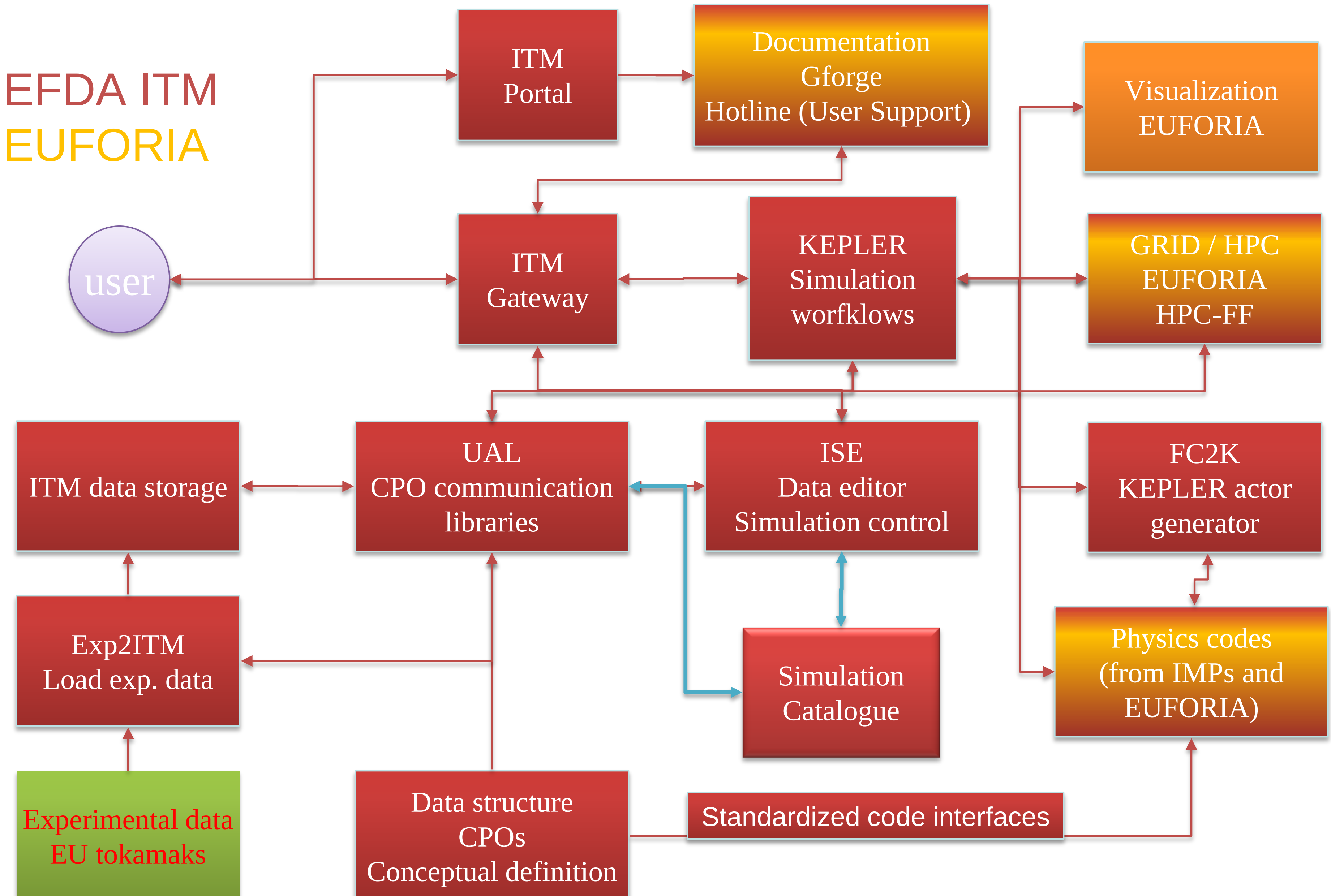
MAPPER

ITER Framework services contract CEA, CRPP, Chalmers, Areva

EFDA GOTiT Goal Oriented Training in Theory

CHALMERS
UNIVERSITY OF TECHNOLOGY

Platform Overview

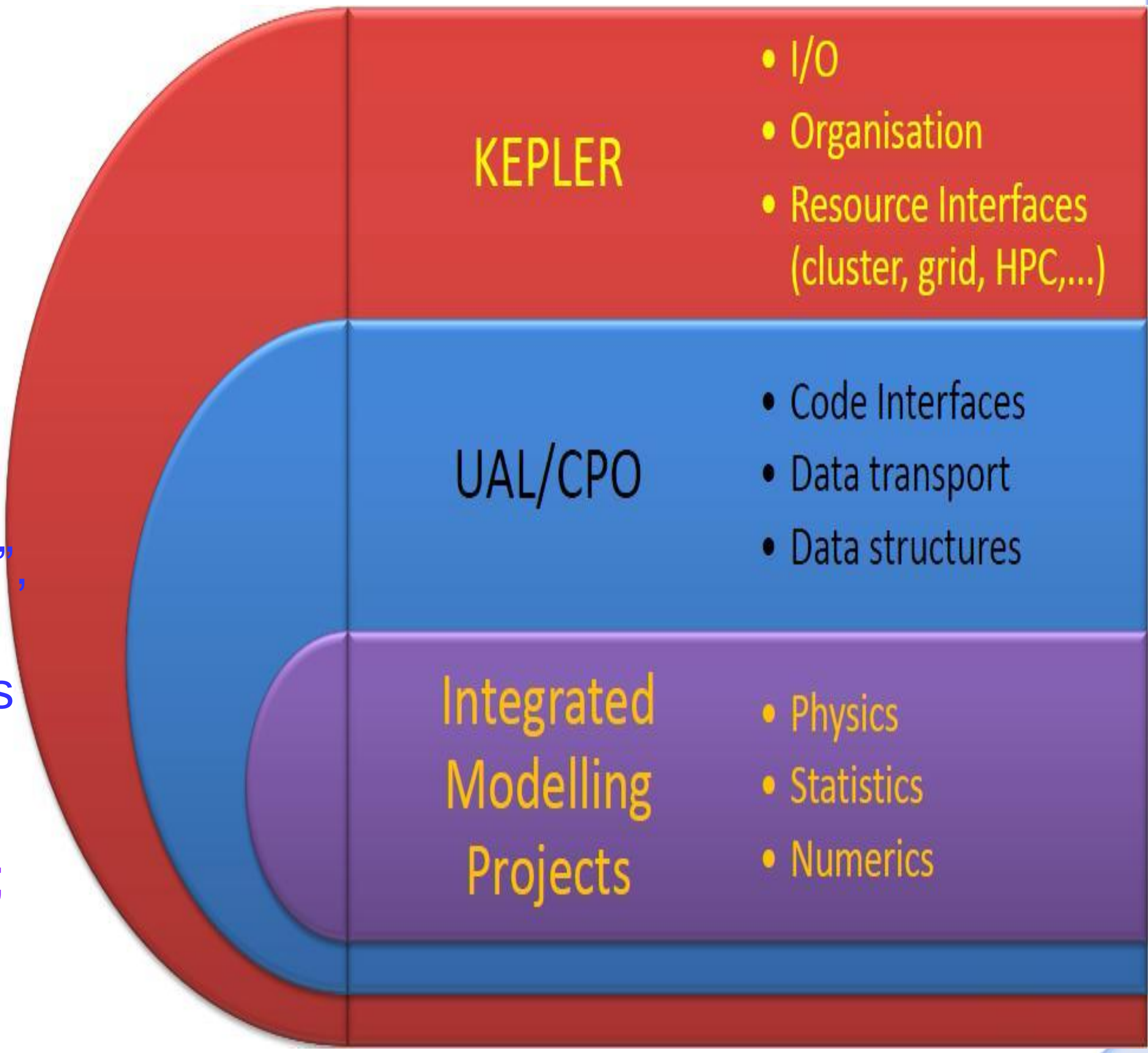


ITM-TF IM framework: layered and modular

Workflow orchestration layer: Kepler graphical system provides transparent representation of physics interactions and related data flows

Coupling layer: communication through standardized data, CPO "*Consistent Physical Objects*", transferred via UAL supporting several languages

Modelling tools layer: physics modules unchanged; transparent integration via wrapping tool & automatic generation of Kepler actor



ITM-TF IM framework: e-infrastructure challenges

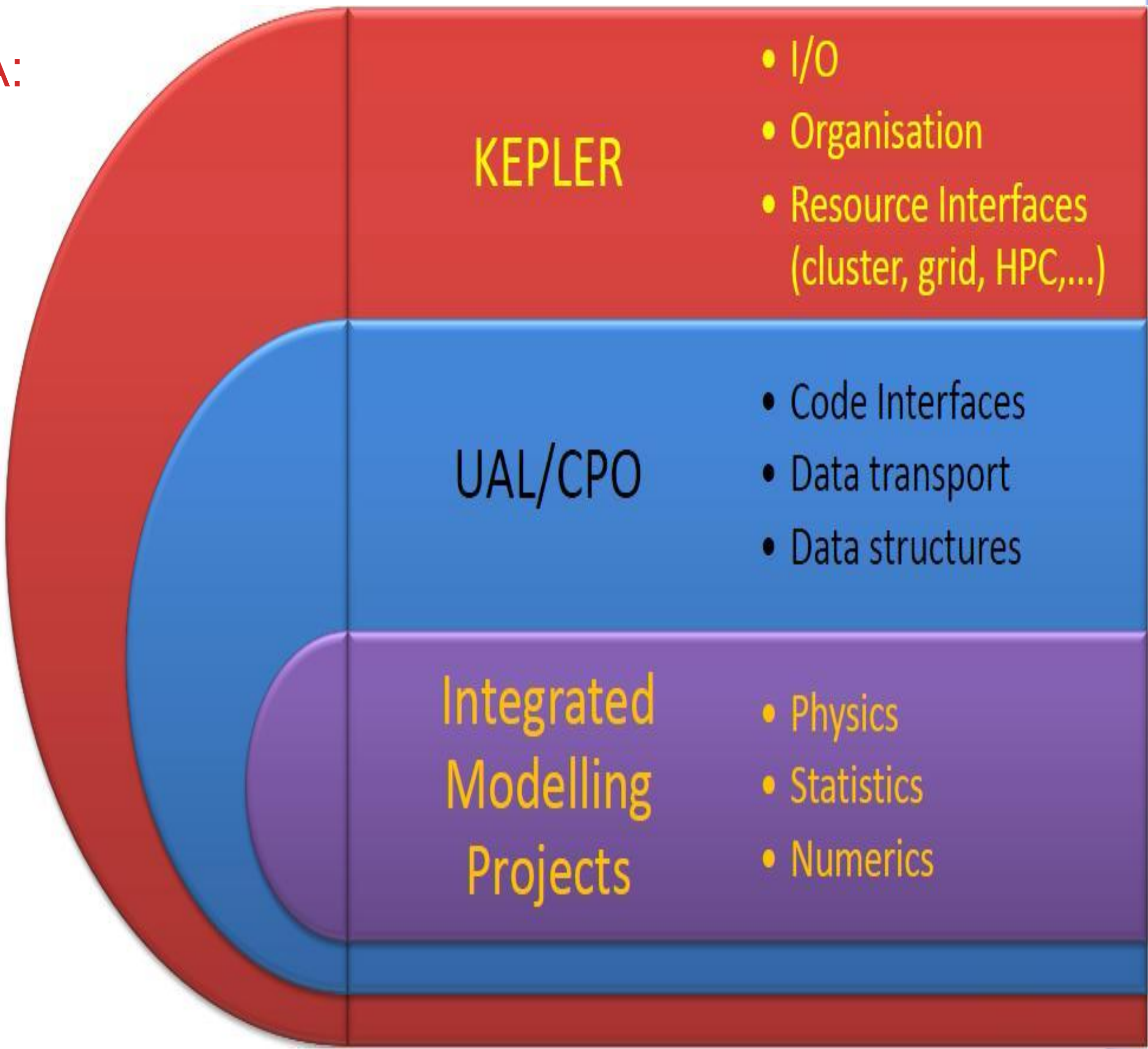
Workflow integration a challenge but one of the neater outcomes of EUFORIA: grid and HPC access part of Kepler distributions.

Communication layer using central databases (TCPIP) or in memory transfers –not files

Physics modules rather than codes challenge for standard Grid tools and workflows

EUFORIA provides generic resource interfaces

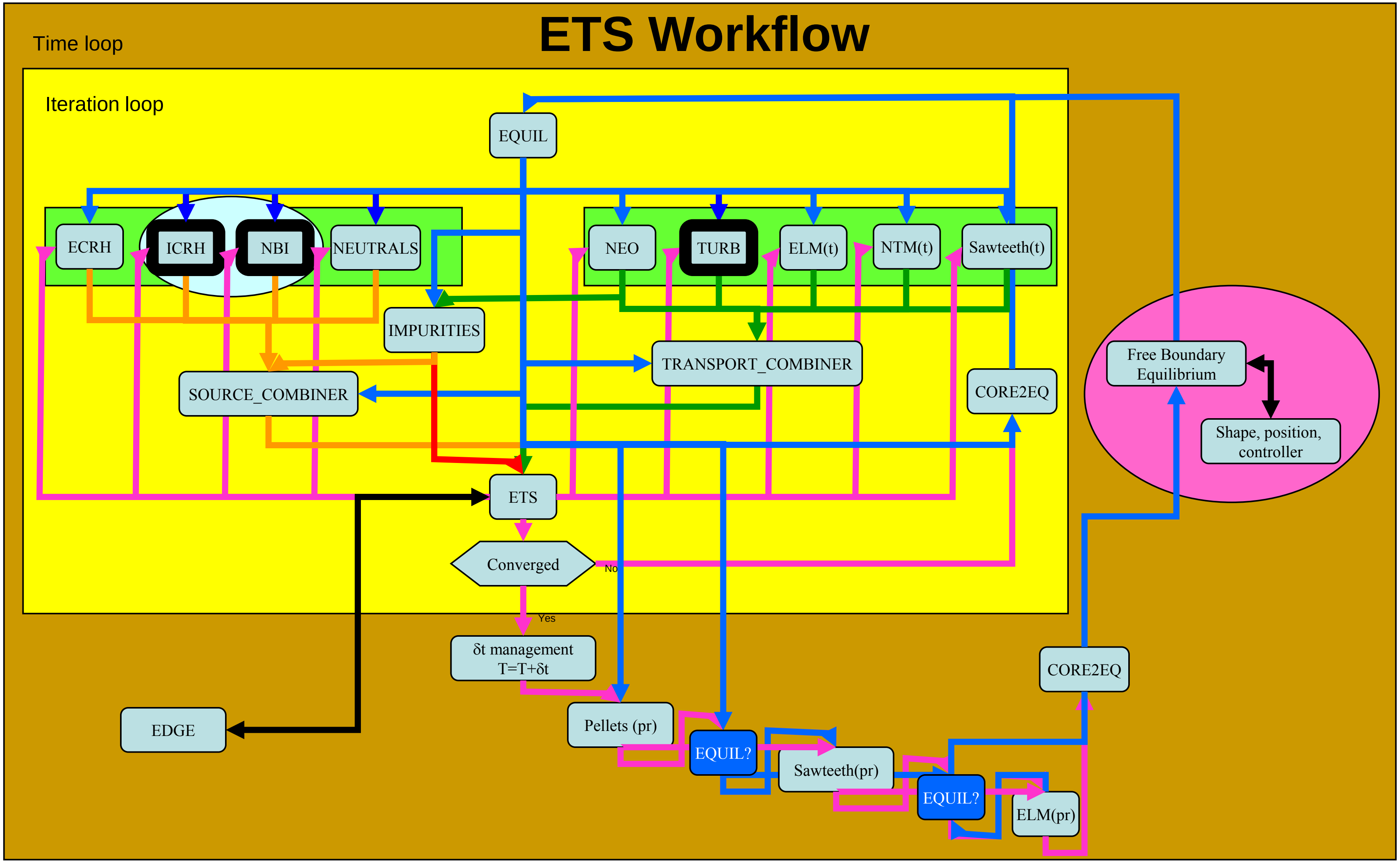
MAPPER addressing distributed data transfer layers



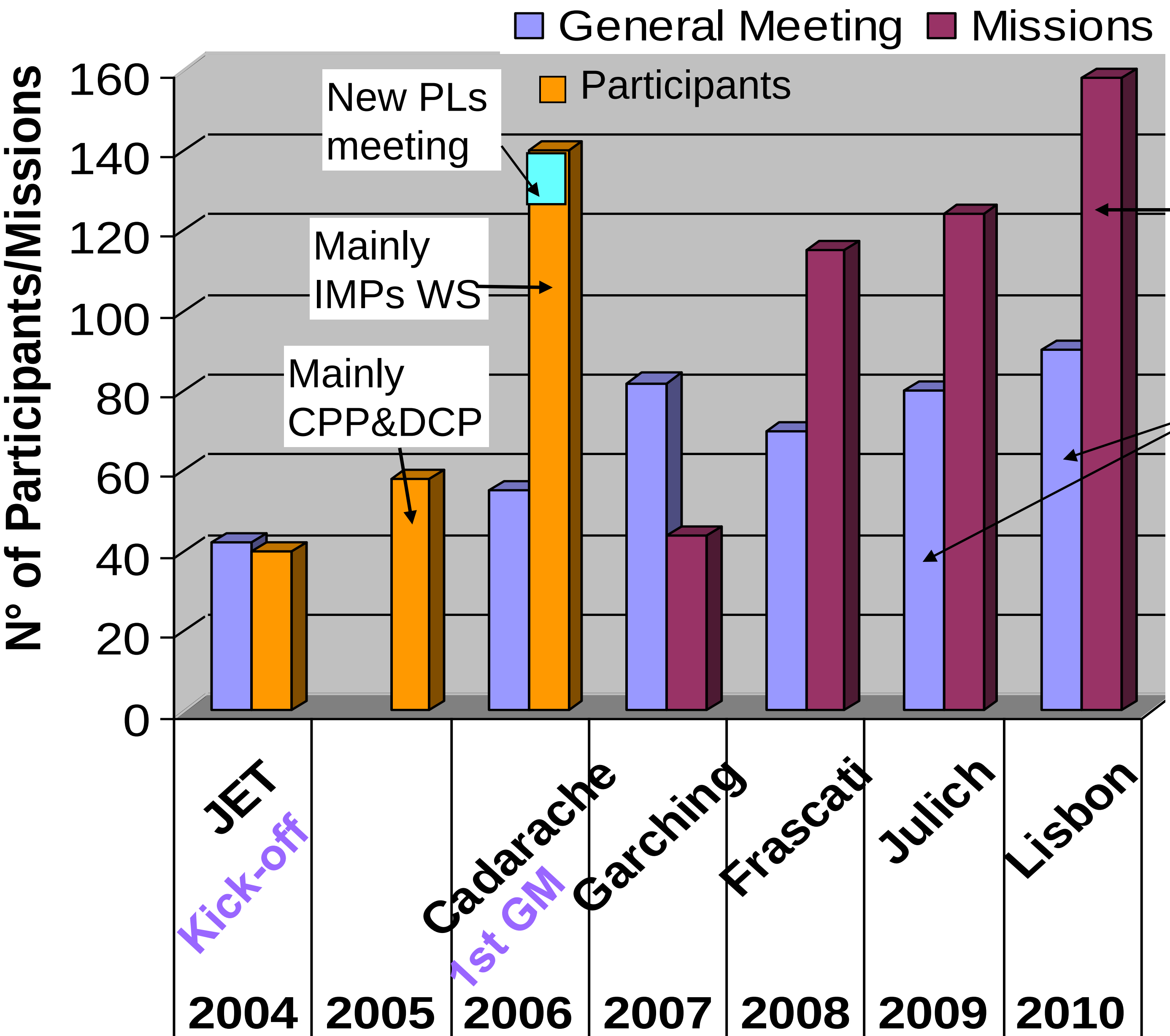
Several Use Cases and application structures in parallel

- 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



Bringing people together



200-250 persons
60FTE

Stronger participation to larger Code Camps
117 this year

Participation to General Meeting ~ const
75 this year

Collaborative TF wide activity increasing
(1-1 missions < 10)

2008-2010

14 member Institutes

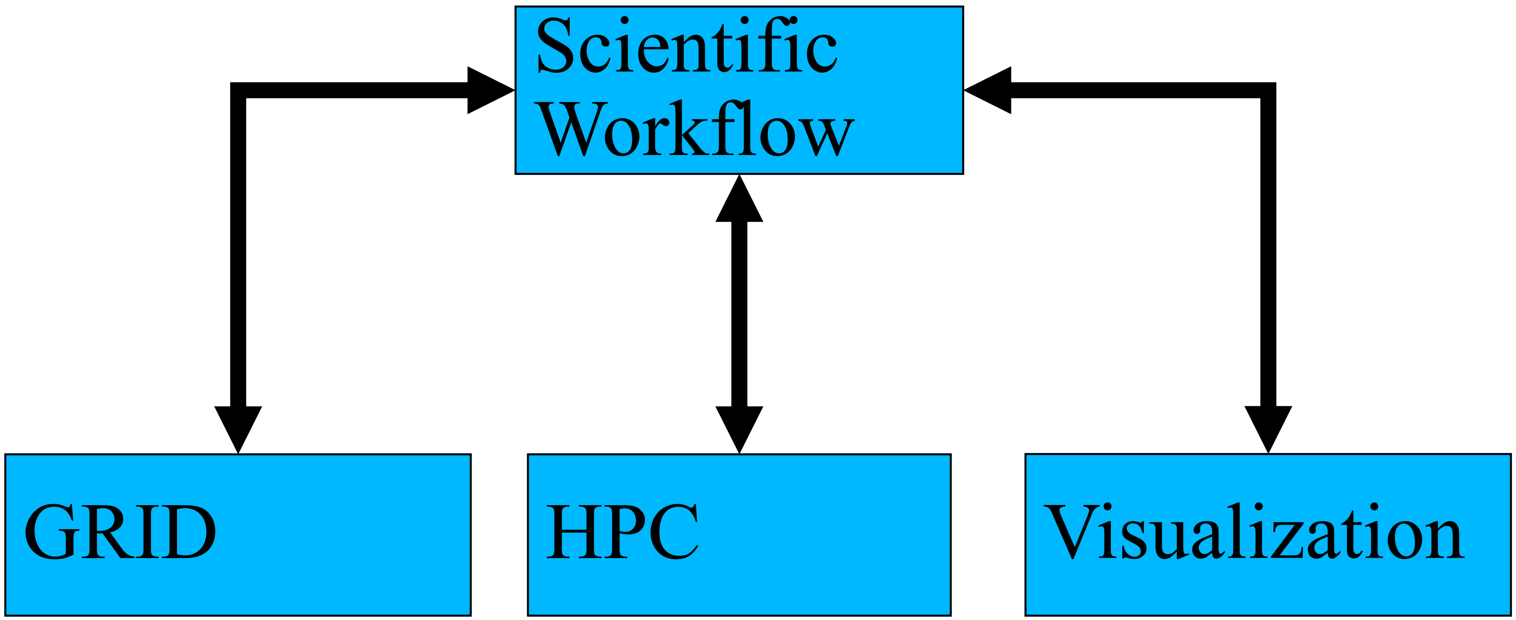
522pms covering

- **Management**
- **Training**
- **Dissemination**
- **Grid and HPC infra-structure & support**
- **Code adaptation & Optimization (grid, HPC,[cloud])**
- **Workflows**
- **Visualization**

EUFORIA Activities

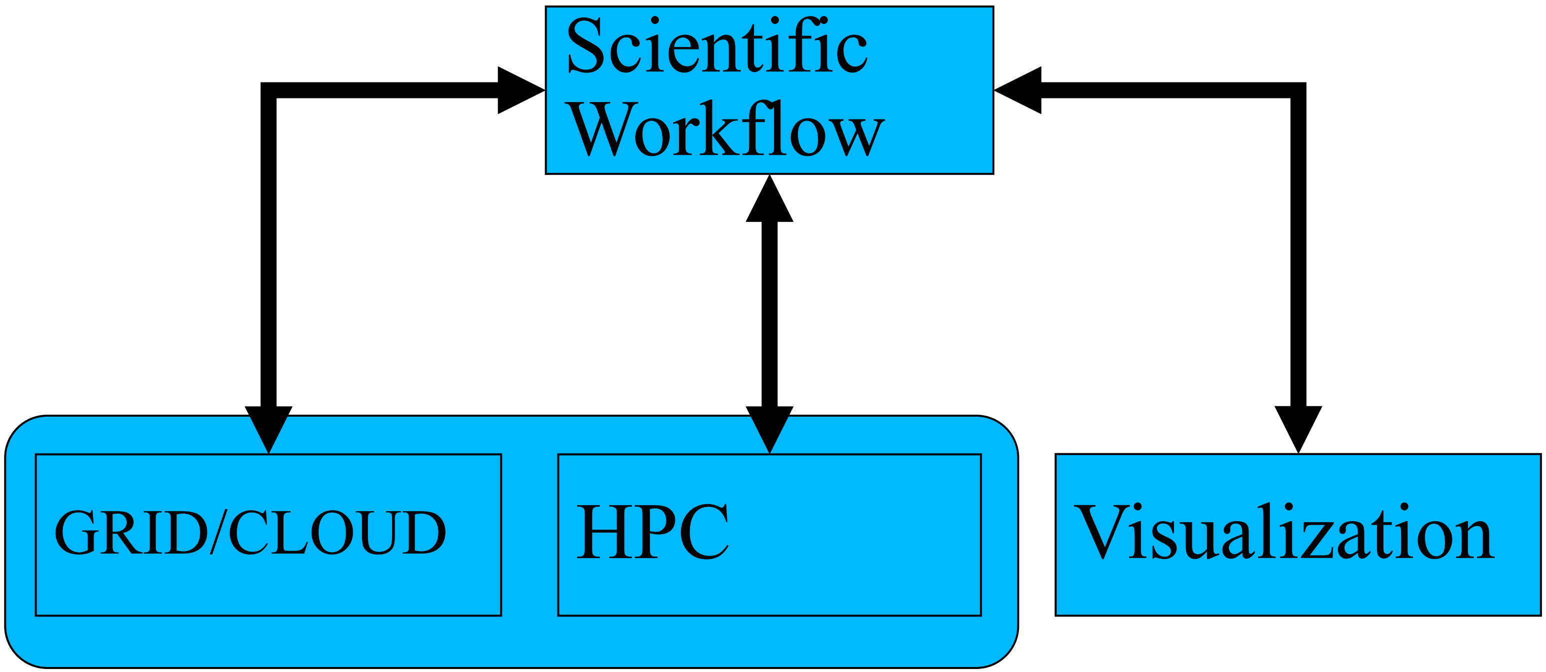
- Satisfied EUFORIA user community
  - 550 training days provided, More than 50 publications from users
  - 10 million HPC hours provided
- Complex workflows established across range of application scenarios/types (Grid serial, Grid parallel, parameter scan, HPC, …)
- Significant parallel performance 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



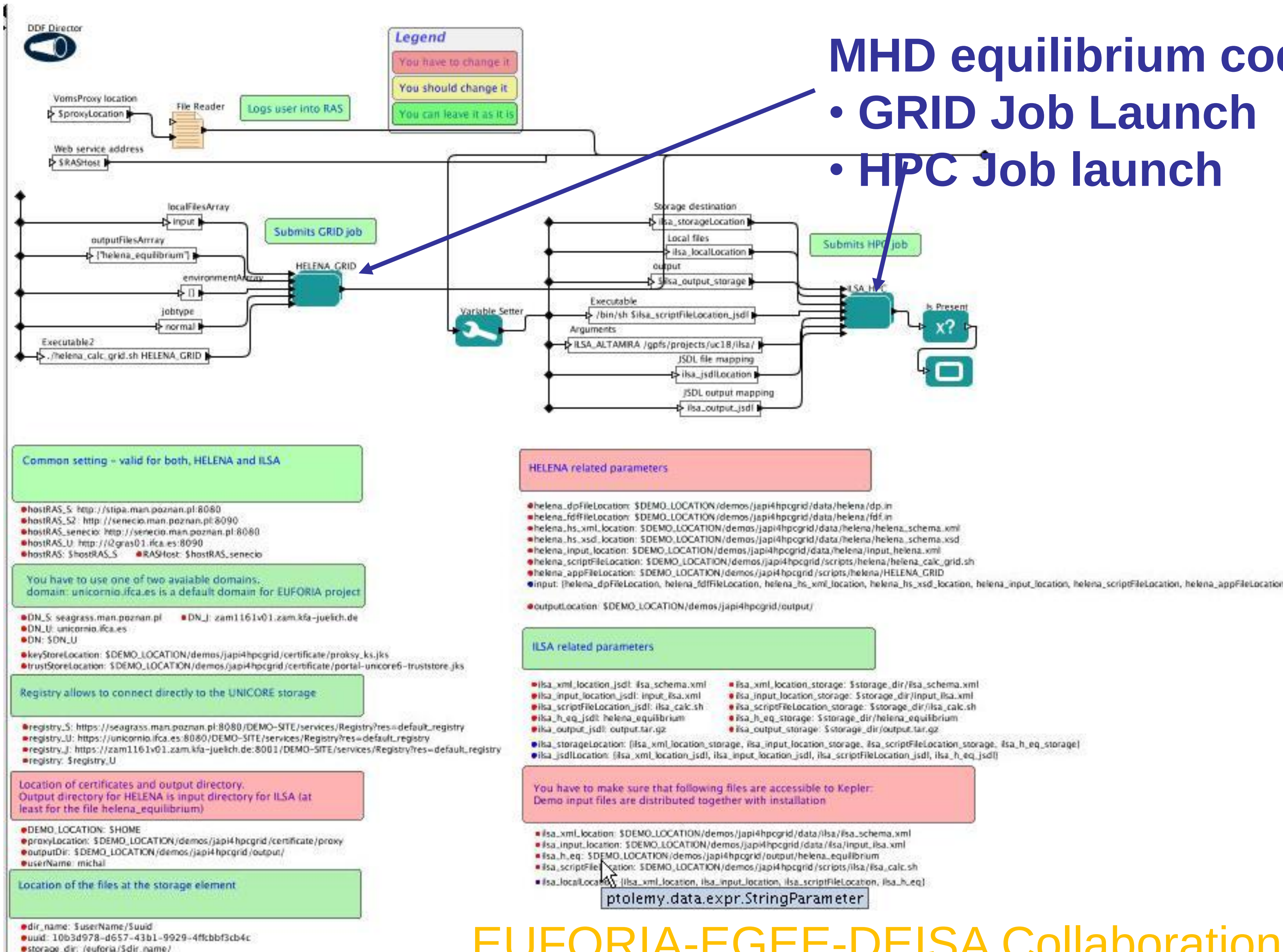
- Building on <u>e-infrastructure</u> tools, middleware and installations
- Integrating tools and physics models together with a <u>”fusion simulation ontology”</u>
- At least initially building on fusion de facto standards for data access and communication

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

Workflow + fusion codes on GRID and HPC



MHD equilibrium codes
- GRID Job Launch
- HPC Job launch

EUFORIA-EGEE-DEISA Collaboration

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

1st EU review selected two apps on MAPPER e-Infrastructure (EGI and PRACE resources)

Taskforce established

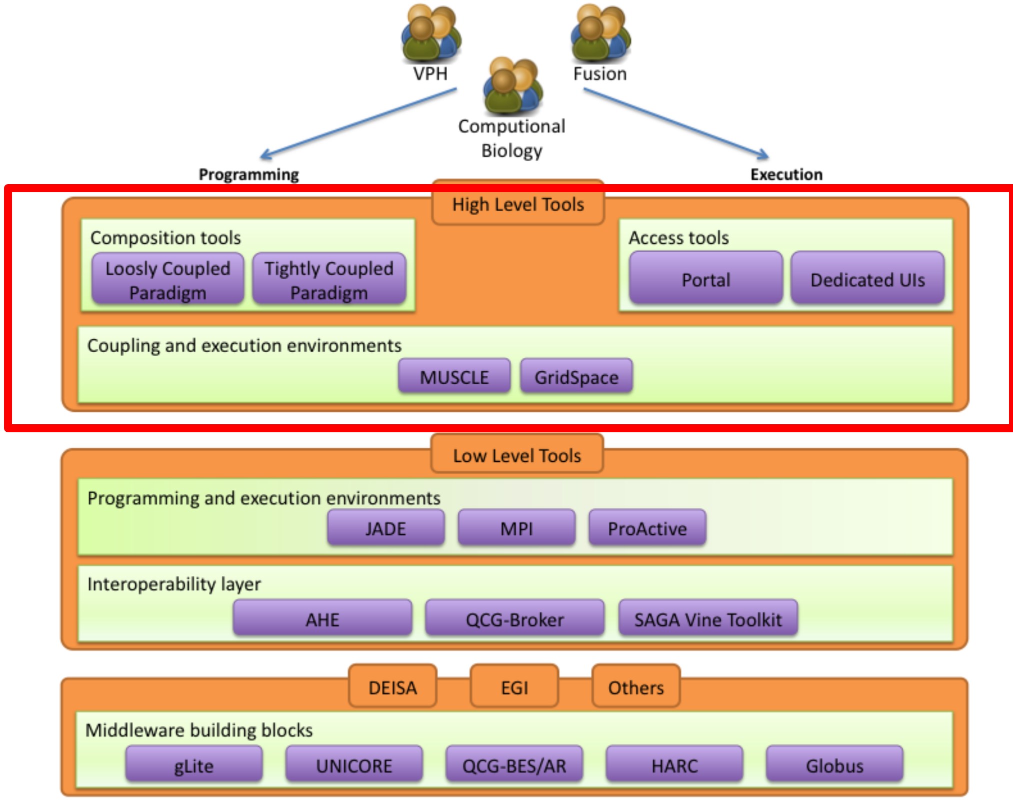
MoU signed

1st evaluation

2011 05 06 08 09 11 … 2012 … 2013

High level tools: objectives

- Design and implement an environment for composing multiscale simulations from single scale models
  - encapsulated as scientific software components
  - distributed in various e-infrastructures
  - supporting loosely coupled and tightly coupled paradigm
- Support composition of simulation models:
  - using scripting approach
  - by reusable “in-silico” experiments
- Allow interaction between software components from different e-Infrastructures in a hybrid way.
- Measure efficiency of the tools

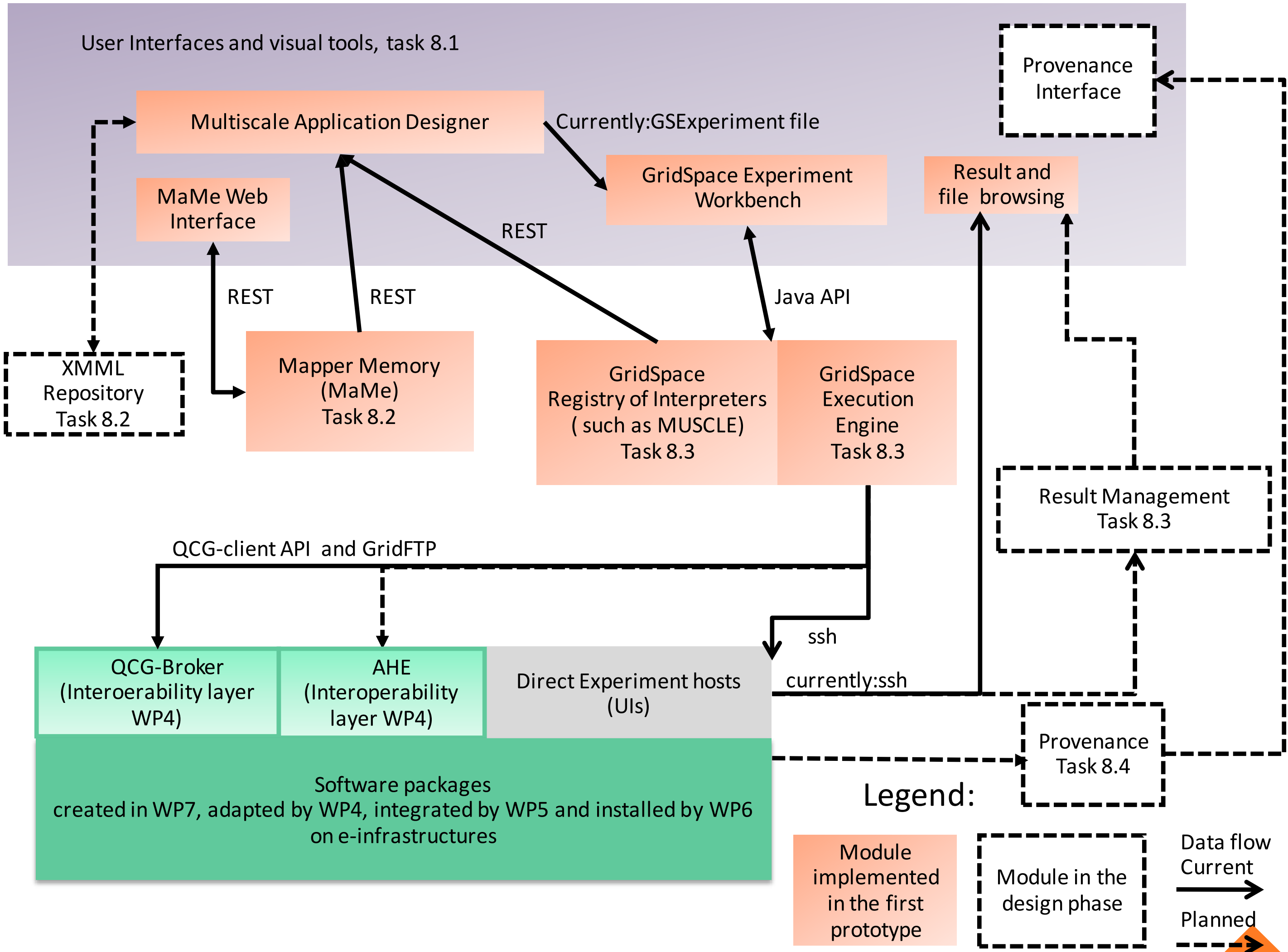


Requirements analysis

- Focus on multiscale applications that are described as a set of connected, but independent single scale modules and mappers (converters)
- Support describing such applications in uniform (standardized) way to:
  - analyze application behavior
  - support switching between different versions of the modules with the same scale and functionality
  - support building different multiscale applications from the same modules (reusability)
- Support computationally intensive simulation modules
  - requiring HPC or Grid resources
  - often implemented as parallel programs
- Support tight (with loop), loose (without loop) and hybrid (both) connection modes

Overview of tools

- **MAPPER Memory (MaMe) -** a semantics-aware persistence store to record metadata about models and scales
- **Multiscale Application Designer (MAD) -** visual composition tool transforming high level MML description into executable experiment
- **GridSpace Experiment Workbench (EW) -** execution and result management of experiments on e-infrastructures via interoperability layers (AHE, QCG)



"The way I see it, if you want the rainbow, you gotta put up with the rain."

Dolly Parton

LESSONS LEARNED

ITER 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.

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
  - USER acceptance a challenge as learning thresholds are high!
  - Authentication (single sign over multiple infrastructures)

LESSONS LEARNED

Sociology:

- EFDA provides a legal framework between 27 pan-european partners – note this is largely in-kind or voluntary activity
  - Long lived activity that have traversed most sociological strains
    - North-South mentalities (yes it is a real issue but is manageable through time and joint understanding of language)
    - National Laboratories (or experiments) vs Universities – or rather line management control structures
  - Despite strong and clear legal framework (all participants share and have access to all tools)
    - IPR remains a major issue
- e-infrastructure contingents vs physics providers
  - Both will in general need to adapt to each other but can be a painful excercise

Lessons Learned Sustainability

- All larger scale projects tend to have strong requirements to sustain and exploit tools rather than the knowledge gained!
  - Is this the right thing to promote? Is not the knowledge gained and the development path to new and improved tools more important?
  - Risk to promote a plethora of tools not for their usefulness but to preserve the notion of investment well spent.
    - Worse case scenario: poorly designed, bloated, tools being maintained, forcing a lot of overhead structures and new tools to ameliorate deficiencies to be invented and reinvented again and again.
  - Continuous improvement hampered or even blocked by design?

Lessons Learned: SERVICES vs USERS

- Even if services are available access to (mainly 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
- Cultural divide: Security and integrety of services on one hand and usefulness and availability on the other

Summary

- A number of prototyping activities are ongoing or being finalized.
- Early days still for ITER, 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 (federated resources)
  - Challenge is to put ITER modelling and data in the hands of the users
- Time to influence/review/input from e-infrastructure point of view - ITERIS project sensible point of contact.
- User emabracement of the thought of e-science is MUCH higher than actual implementation, or own investment, interest!