IGI Training e User Support: esperienze di collaborazione e porting di applicazioni su Grid

D.Cesini (IGI/INFN-CNAF)

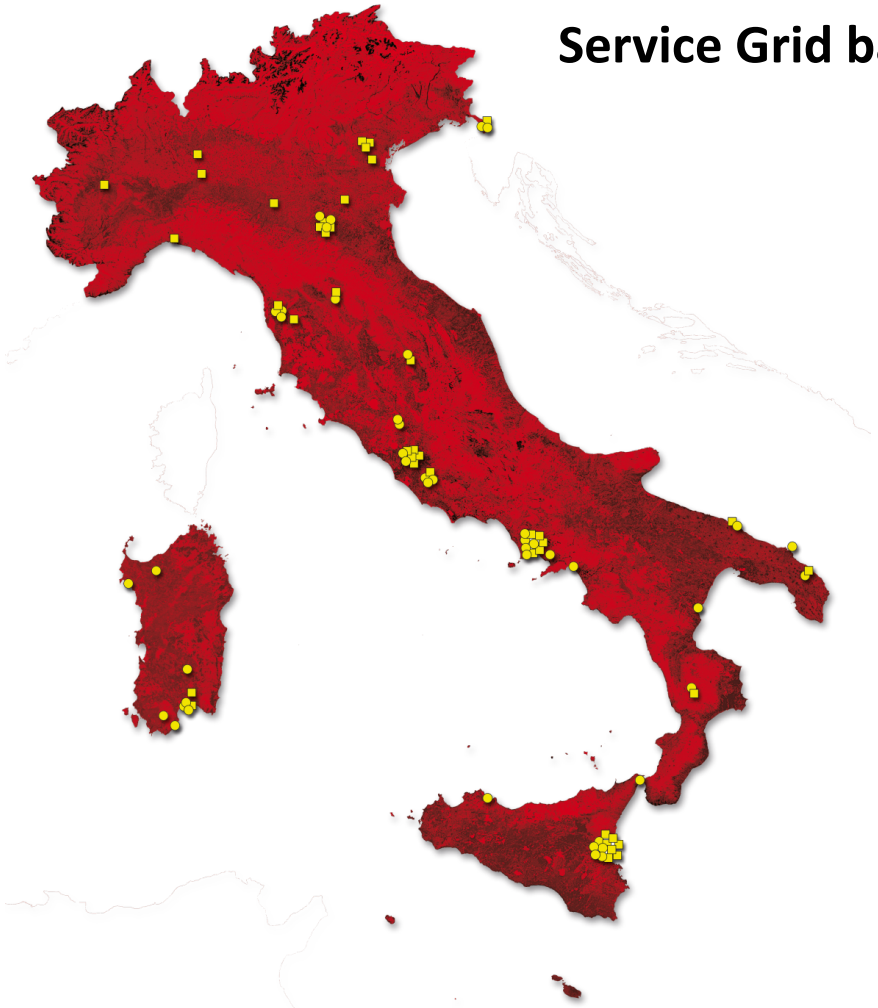
E.Giorgio (IGI/INFN-CT)

(On behalf of the IGI-FUS group)

Overview

- IGI
- IGI-FUS
- Examples of collaborations in various scientific domains
  - High Energy Physics
  - Computational Chemistry
  - Bioinformatics
  - Earth Science
  - Pharmacology
- Conclusion

The Italian Grid Infrastructure (IGI)



Service Grid based on Scientific Linux Clusters:

55	Sites
32000	Cores
20 PB	Disk Storage
10 PB	Tape Storage
1100	Users
50	VOs
260M	CPU hrs in 2012
10	Application Domains

gLite

IGI RELEASE

egi

EUROPEAN MIDDLEWARE INITIATIVE

4 Unità:

Operations

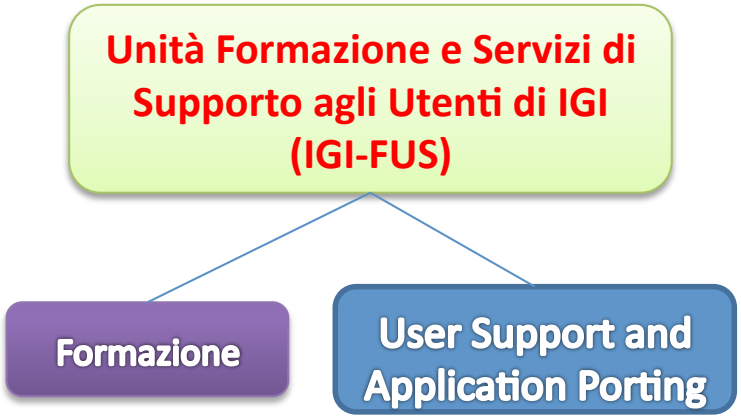
Sviluppo e progettazione middleware

Formazione e Support o agli utenti

Amministrazione e Relazioni Esterne

IGI-FUS

- Support to new user communities in Application Porting
- Consultancy in building computing models
- Organise the Italian User Forum
- User requirements gathering
- Training organisation
- Reference at European level for these activities



Directly represented in FUS User Forum communities from:
INFN, CNR, ENEA, SPACI, INAF, UNINA, UNIPG, COMPCHEM

Active also with communities not directly represented

Coordinated by a rota of a senior chosen from the User Communities:
05/2010-04/2011 – Claudio Vuerli (INAF)
05/2011-04/2012 - Antonio Laganà (UNIPG)
05/2012-04/2013 – Laura Perini (INFN)
05/2013-04/2014 – Luciano Milanesi (CNR-ITB)

~3 FTE distributed : Milano, Napoli, Bologna, Catania, Perugia, Bari
+ collaboration with staff from other Units: Operations and Middleware

Strategy for application porting

Seek, identify, and closely support new communities

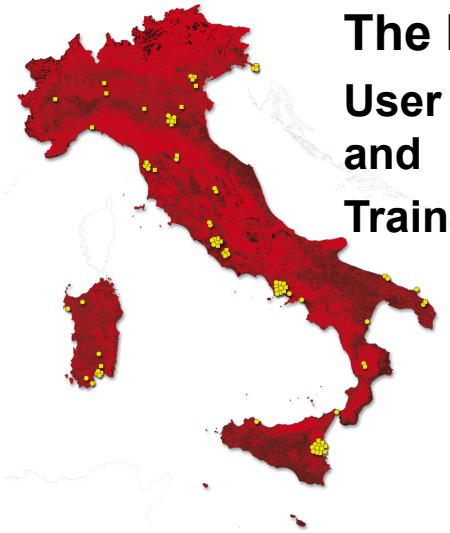
Support in building the distributed computing model

Porting of the applications and workflow to the DCI

Build High Level Web Interface

Train users to use the new execution procedure

Disseminate the work done and success stories to find new applications



The IGI

User Support and Training

Examples of collaborations

High Energy Physics

INFN non LHC experiments

SPES (Selective Production of Exotic Species) http://web.infn.it/spes/

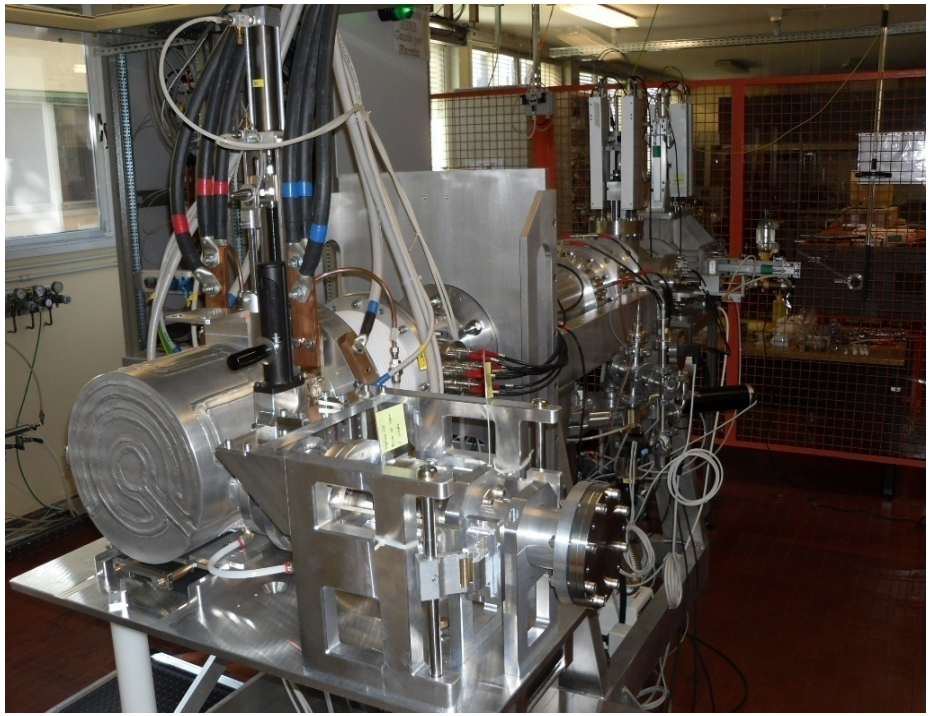
- Porting to Grid of the ANSYS suite
  - Gain experience on handling licensed software on the Grid
  - Prototype of our community support methodology
    -From first contact to high level web interface

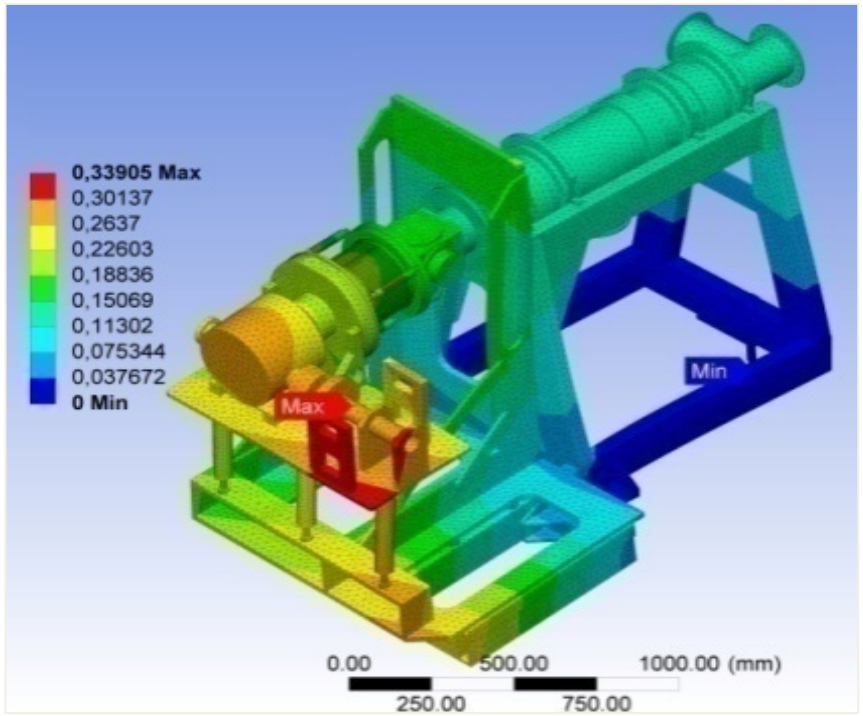
➢GERDA, ICARUS, VIRGO

- Training and support in using Grid Services

ANSYS: The scientific problem for SPES

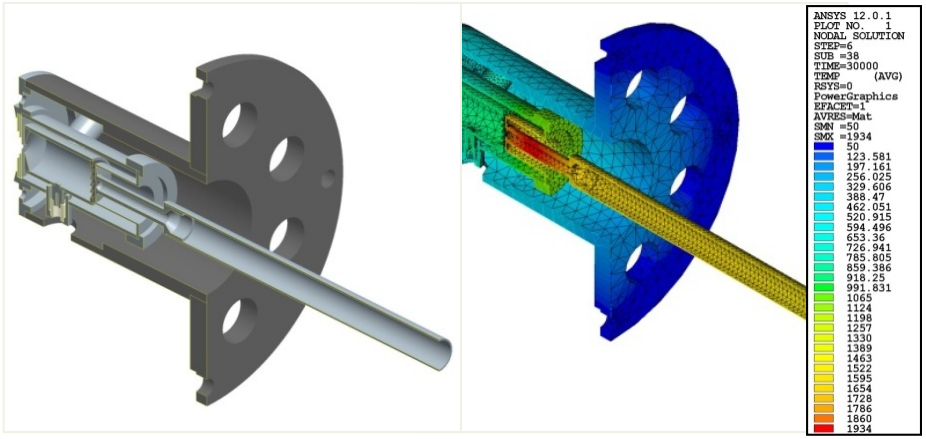
- Simulation modeling of a real experimental machinery apparatus
- Target: evaluate the stability and the mechanical strength
- Software used: ANSYS simulation suite



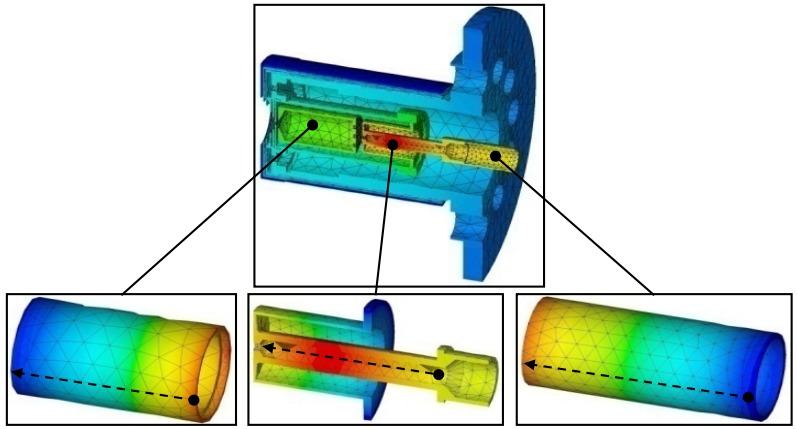


ANSYS computation for the SPES experiment

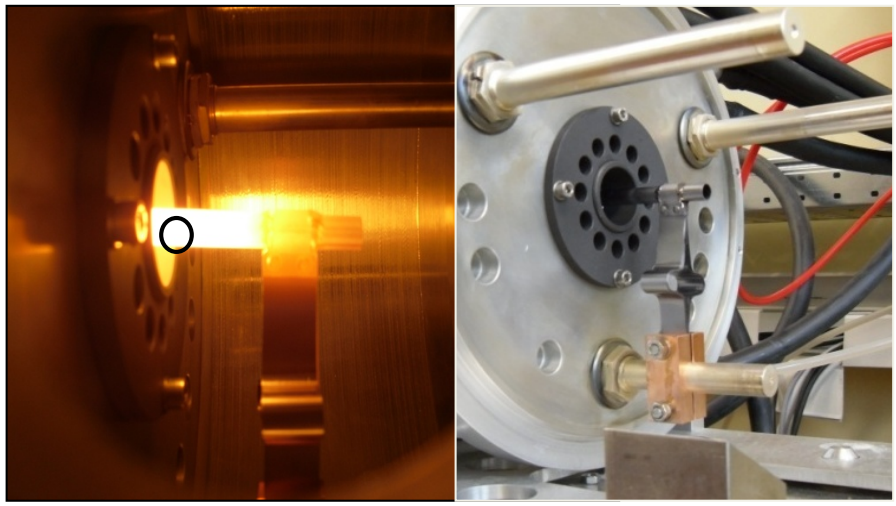
1) Elettro-termo-struttural of the (Isolde) and (Hribf) FEBIAD



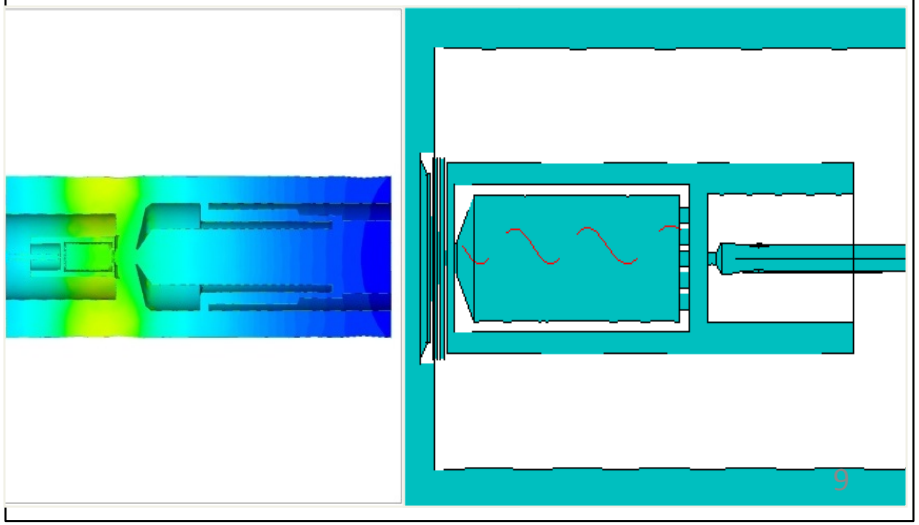
2) Comparison focused toward the new SPES FEBIAD



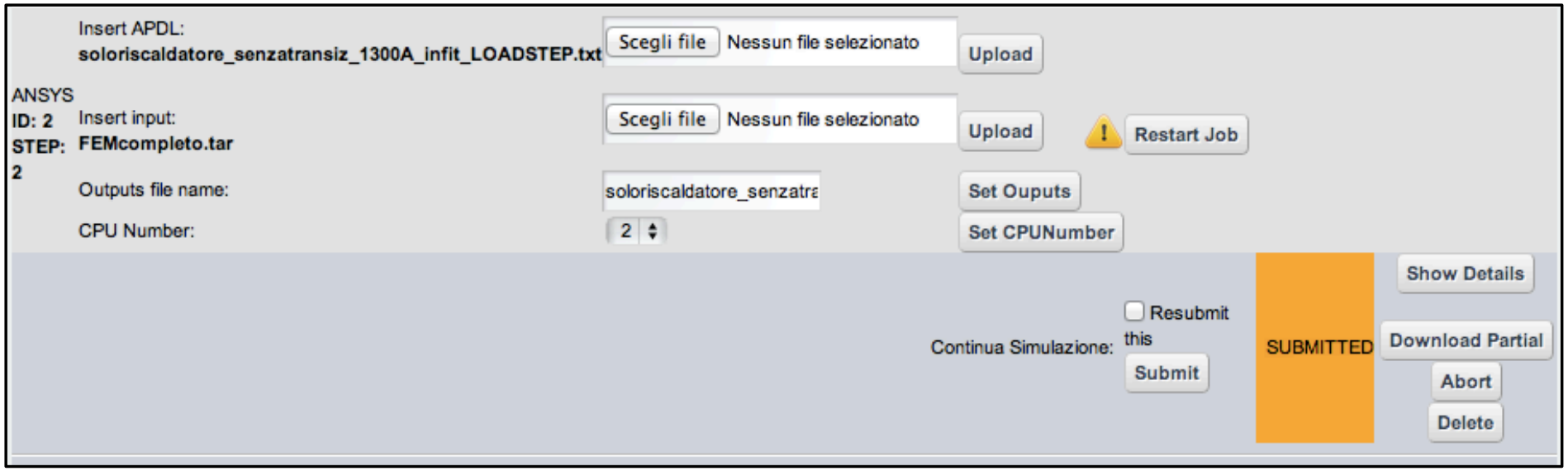
3) Validation of the FEM model



4) Study of the magnetic field using FEM



The ANSYS Portlet



Talk@EGI-tf12: https://indico.egi.eu/indico/contributionDisplay.py?contribId=40&sessionId=48&confId=1019

Workshop@INFN-LNL on December , 19th

Computational Chemistry

- Dept. Chemistry UNIPG
  - International COMPCHEM Virtual Organisation
  - Computational Chemistry Division of EUCHEM

    -Interest in developing tools to seamlessly submit workflows to DCIs and Supercomputer (**HTC/HPC interoperability**)

    -Systems to evaluate the use of the infrastructure: Quality of Services (QoS) and Quality of Users (QoU) functionalities

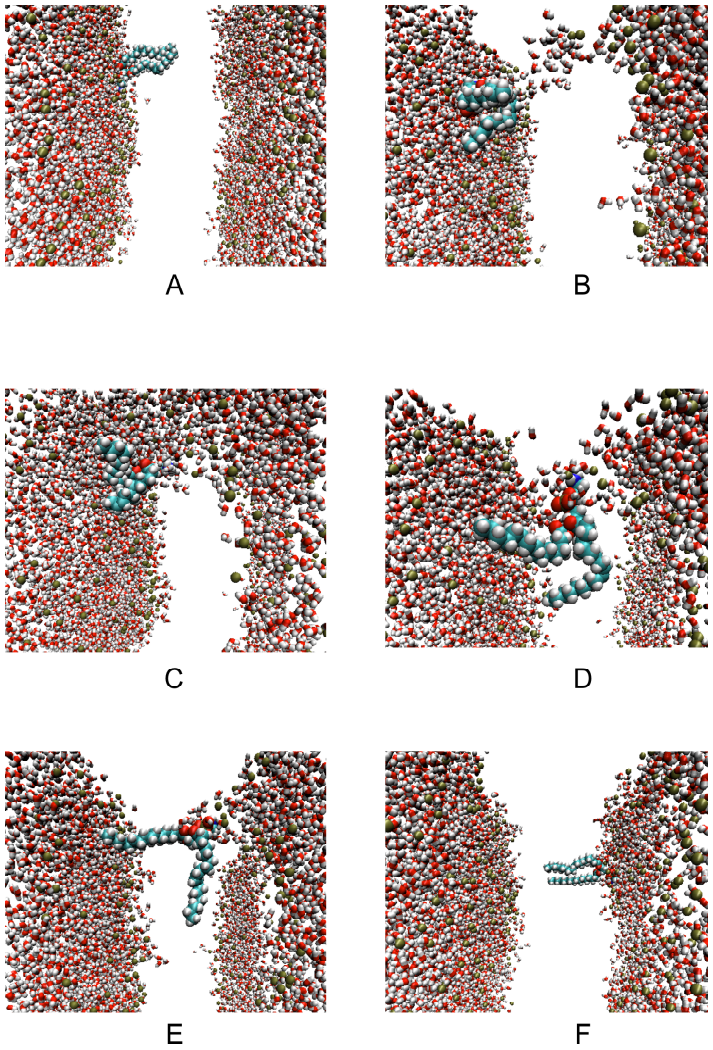
    - Porting of various applications: DL_PLOY, GROMACS, NAMD, VENUS-C…
  - Dept. of Chemistry, Univ. Torino

    - Porting of a **licensed** application: **CRYSTAL** (http://www.istm.cnr.it/csrsrc/sw_crystal.html)

- CNR-ISOF (Bologna) - http://www.isof.cnr.it/
  - Collaboration with the DUCK project (Grid in Emilia Romagna http://www.comput-er.it/)
  - Molecular dynamics applications (NAMD) and ab-initio simulations (GAUSSIAN)

- Dept. Chemistry Sassari University
  - Molecular dynamics applications (CP2K) and ab-initio simulations (GAUSSIAN)
  - Simulation of the diffusion of small molecules in nanoporous materials (zeolites)

Molecular Dynamics



Dynamics of a lipid bilayer induced by electric fields:

These processes are the basis of important properties and functions of cell membranes

Lipid translocation in a bilayer can be driven by switching on and off an electric field

In collaboration with:

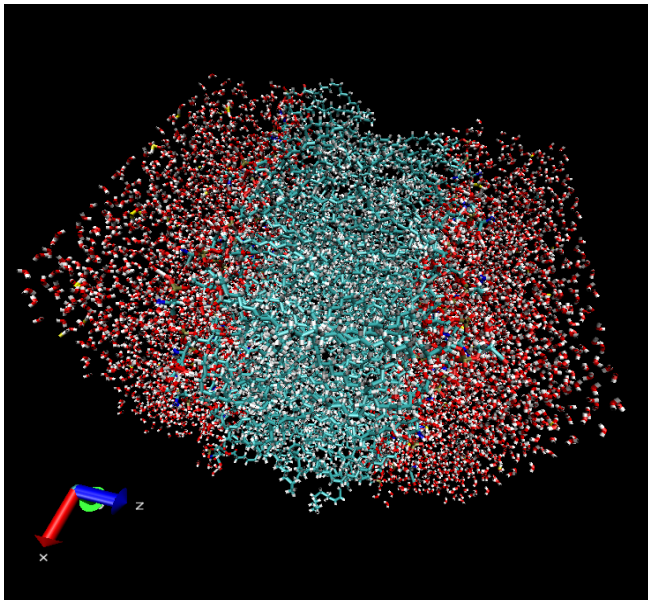
ISOF-CNR (CNR- Institute for the OrganicSynthesis and Photoreactivity, Bologna Unit)

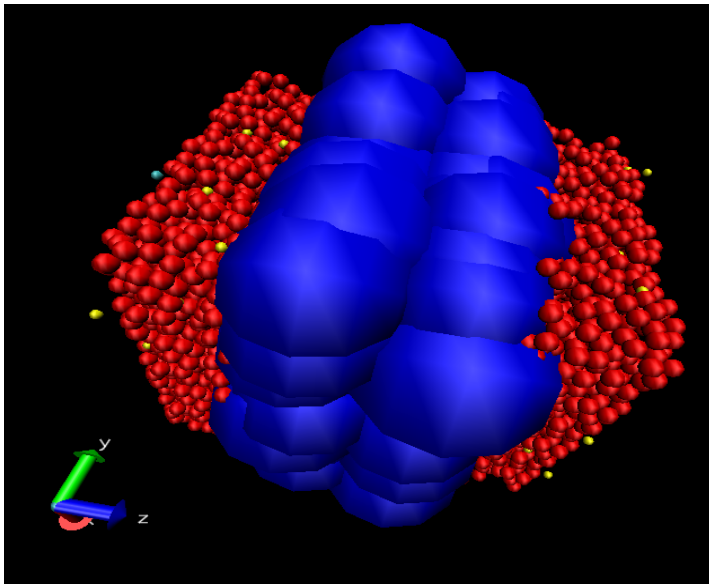
DUCK (Computing in Emilia Romagna)

Authors and ref.: A. Venturini and F. Zerbetto, Phys. Chem. Chem. Phys. 2011, DOI: 10.1039/c1cp20322c

NAMD @ CNR-ISOF

- Long dynamical simulation of very big molecules
- Typically one of this molecule contains 35k atoms and 25ns of simulated time were required
- Non embarrassingly parallel computation on Grid (18 days on a 32HT core machine for each simulation)
- Implemented on Grid using MPI and checkpointing

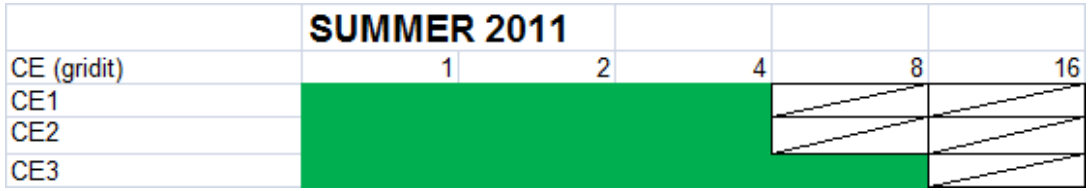




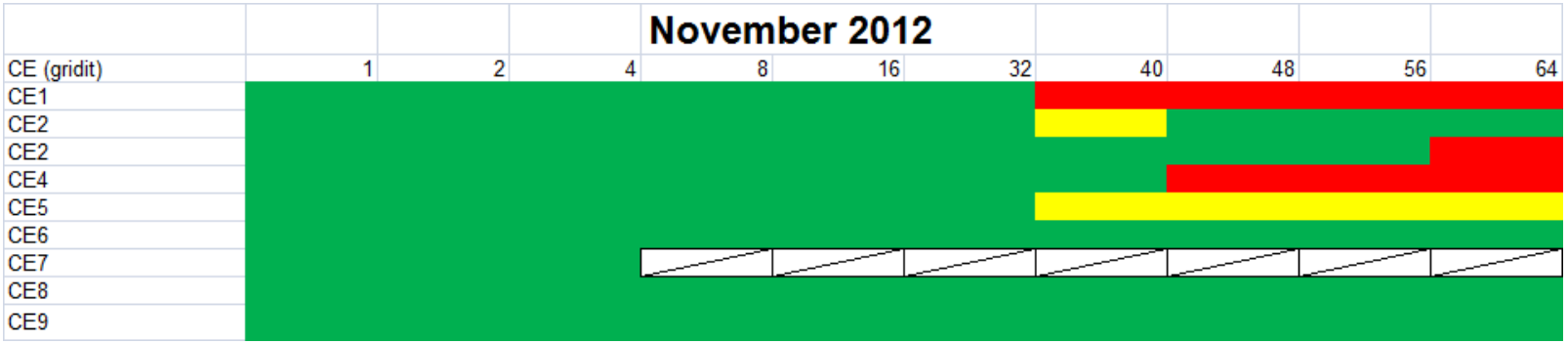
- ~ 19 days on Grid per simulation with **N=50, CPU_NUM = 32**

MPI/Multicore IGI Working Group

- Multidisciplinary WG to improve the HPC-MPI support within IGI
- Deployment issues analyzed
- Standardization of sw environment and published information
- https://wiki.italiangrid.it/twiki/bin/view/MPI/



NAMD APOA1 test using GIRDIT VO hosting new users and users without a specific VO



BioComputing

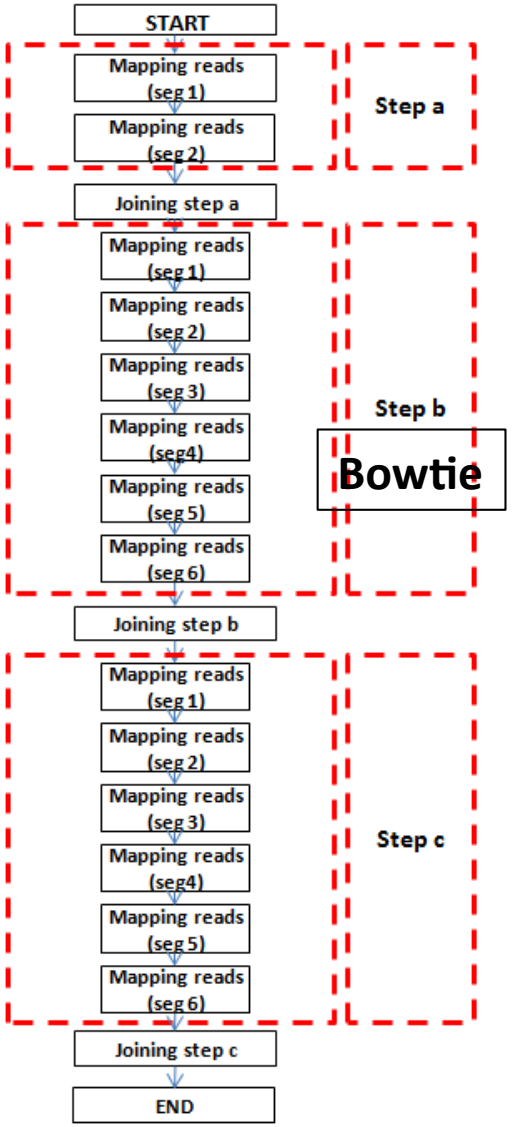
- **Collaboration within DUCK:**
  - **Biocomputing group UNIBO (http://www.biocomp.unibo.it/)**
  - **Centro interdipartimentale ricerche sul cancro Giorgio Prodi Bologna (http://www.circbologna.it/)**
  - **Centro Interdipartimentale di Ricerche Genomiche – CeIRG Modena (http://www.unimore.it/ateneo/centro.html?ID=147)**
    - Protein Annotation and Next Generation Sequencing (NGS)
    - De Novo Assembly

- **Istituto Superiore Mario Boella Torino (http://www.ismb.it)**
- **IS4AC Research Unit e EDA group (politecnico TO)**
  - NGS and genomics in Cloud/Grid (TopHat application)

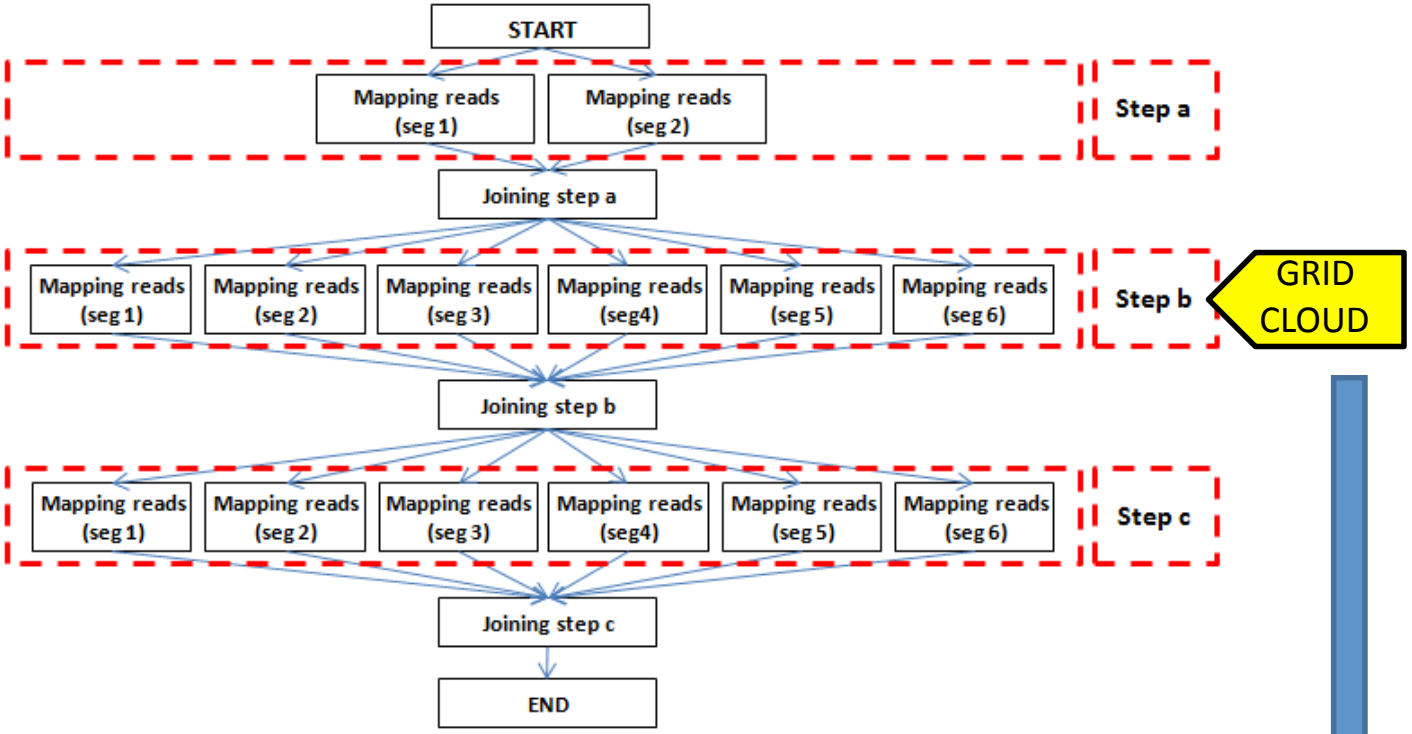
- **Support to the Elixir ESFRI project through an EGI Virtual Team**
  - **https://wiki.egi.eu/wiki/VT_ELIXIR**
  - Interaction with the Italian component just started

TopHat@Boella Optimization

Original version



Distributed Version

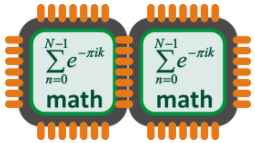


Genome alignments with a reference very efficient on Grid

Earth Science

➢ INGV-BO

- Porting of the modeling framework for oceanographic research **NEMO** (http://www.nemo-ocean.eu/)
- Parallel run on the Grid
- High Level Interface created
- Collaboration within **DUCK**

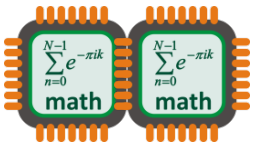


➢ INGV-ROMA && INFN

- Support to the creation of the computing/ data management for the EMSO ESFRI Project (http://www.emso-eu.org)
- DIOeD project submitted to FP7 for the creation of a data management infrastructure for Earth Observation Data (5MEuro, 5 Research Institutes, 5 SMEs)
- Collaboration with the IDIPOS project

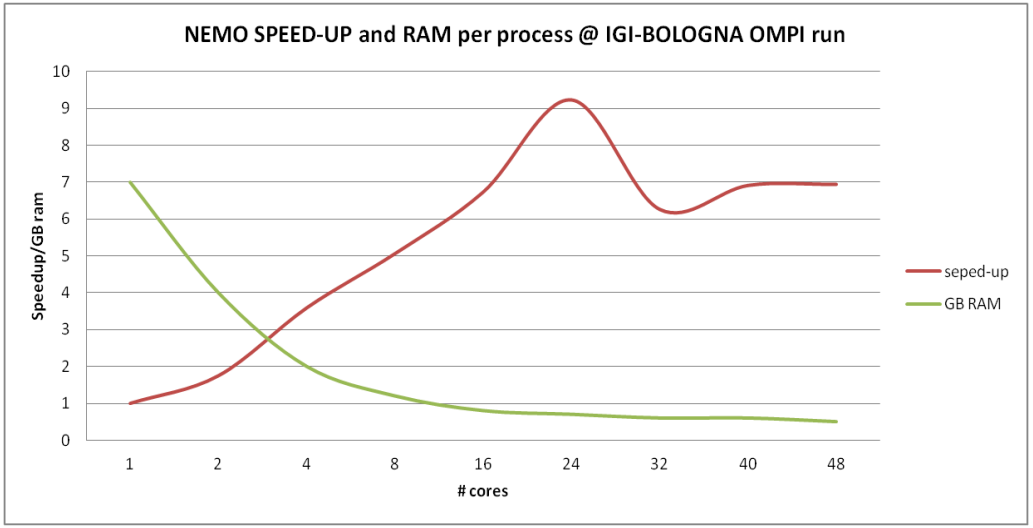
➢ CNR-ISAC Bologna

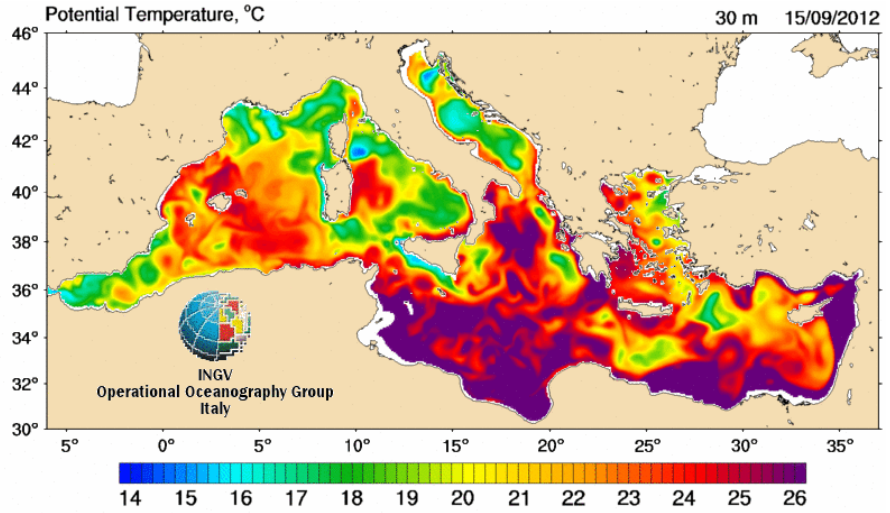
- Tuning of a parallel climate model (GLOBO - http://www.isac.cnr.it/dinamica/projects/forecasts/globo.html)
- Parallel and scalar runs combined to optimize throughput

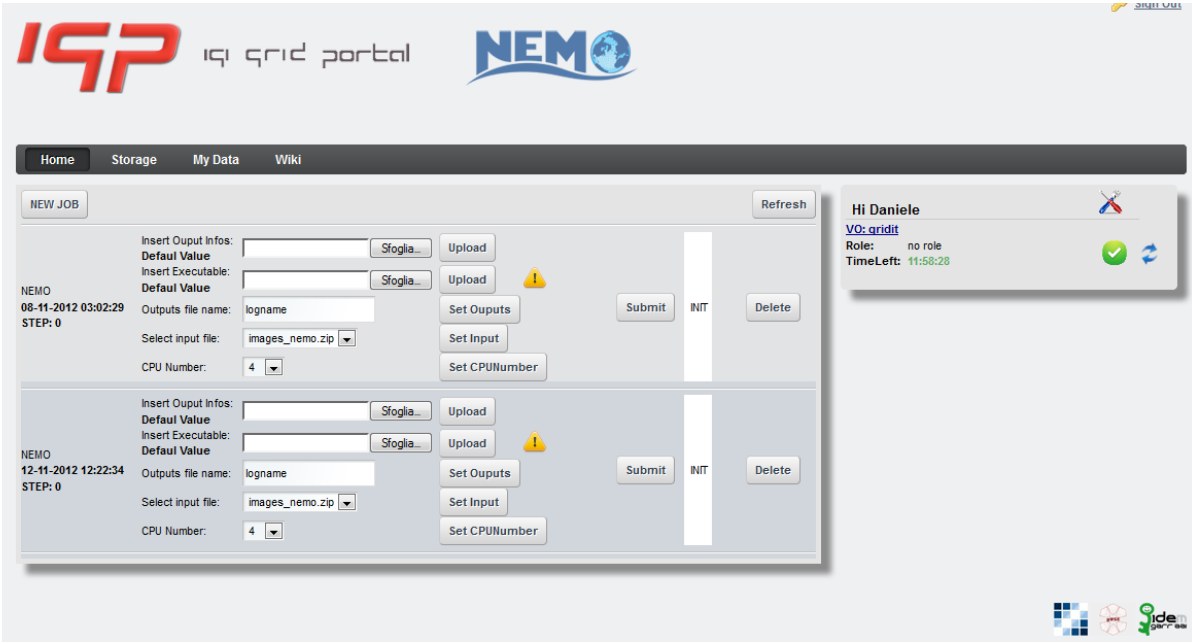


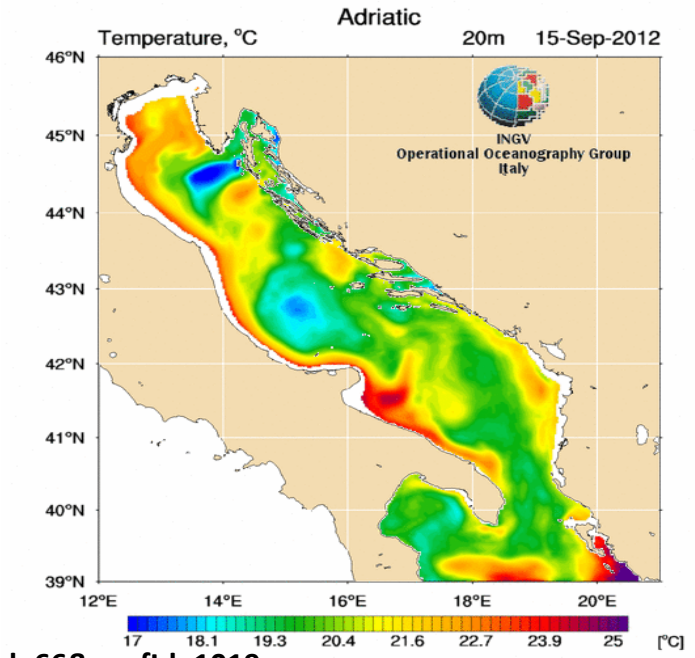
Ocean Modeling@INGV-BO

http://www.nemo-ocean.eu/

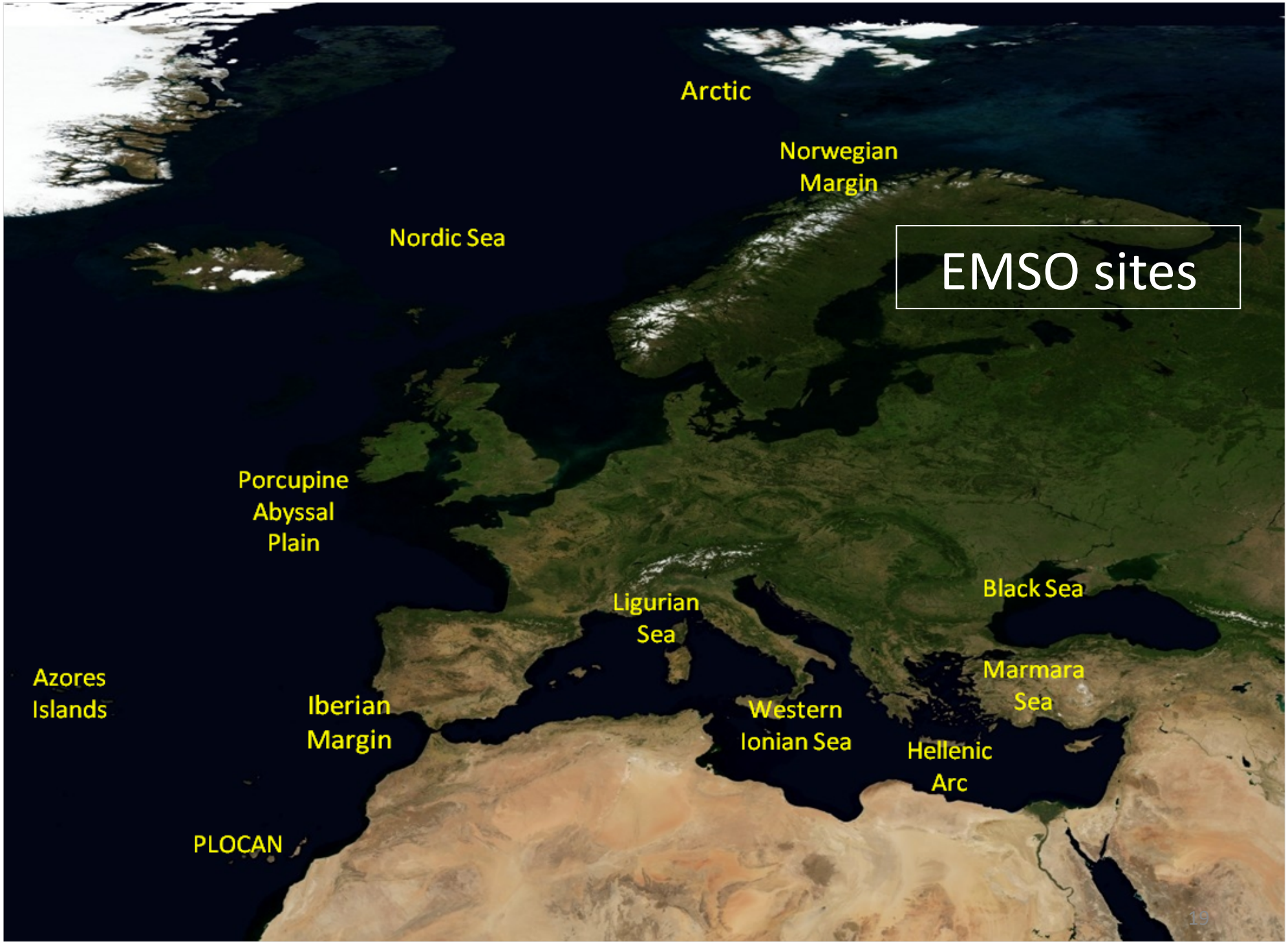








TALK@EGI-TF12: https://indico.egi.eu/indico/contributionDisplay.py?contribId=41&sessionId=66&confId=1019


Arctic
Norwegian Margin
Nordic Sea
EMSO sites
Porcupine Abyssal Plain
Ligurian Sea
Black Sea
Azores Islands
Iberian Margin
Western Ionian Sea
Marmara Sea
Hellenic Arc
PLOCAN

Western Ionian Sea (offshore Catania)

2 ESFRI infrastructures: KM3NeT & EMSO

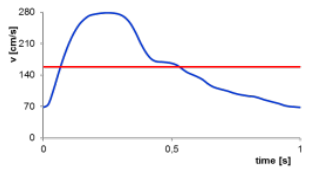
TALK@EGI-TF12:https://indico.egi.eu/indico/contributionDisplay.py?contribId=42&sessionId=66&confId=1019

Pharmacology

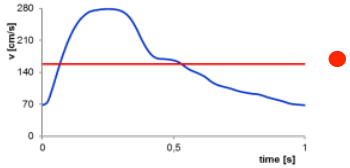
➢ Pharmacological Institute Mario Negri (http://www.marionegri.it/mn/it/index.html)

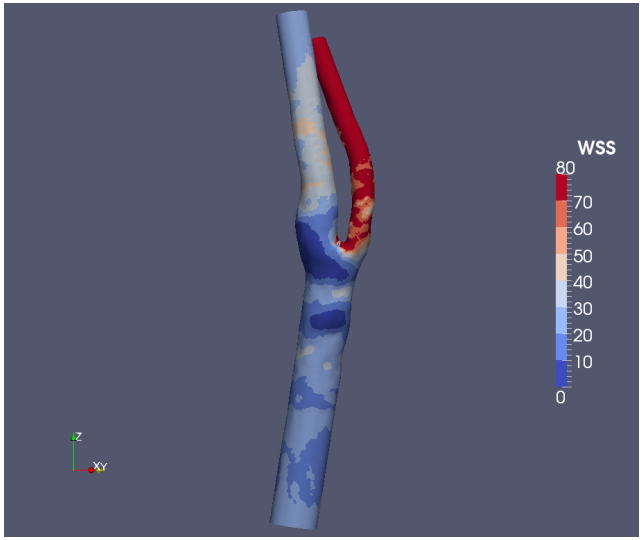
- Simulation on Carotid Bifurcation with OPENFOAM
- Full rendering of the proliferation process at the cell population level during the response to anticancer treatments with MATLAB
- Strong collaboration with INFN-CATANIA

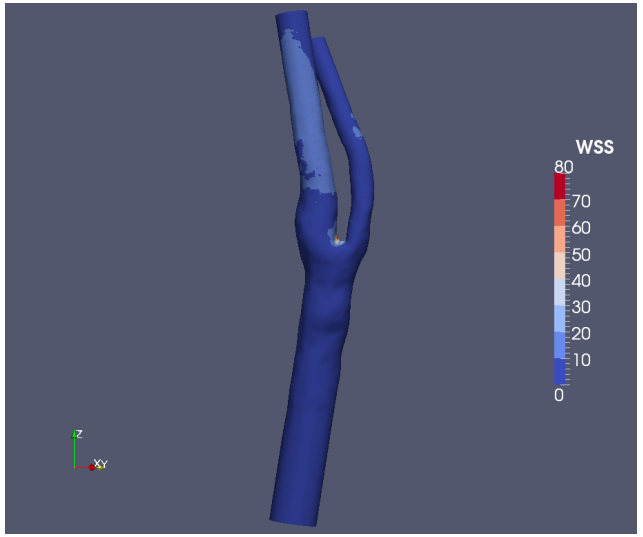
Peak systolic WSS



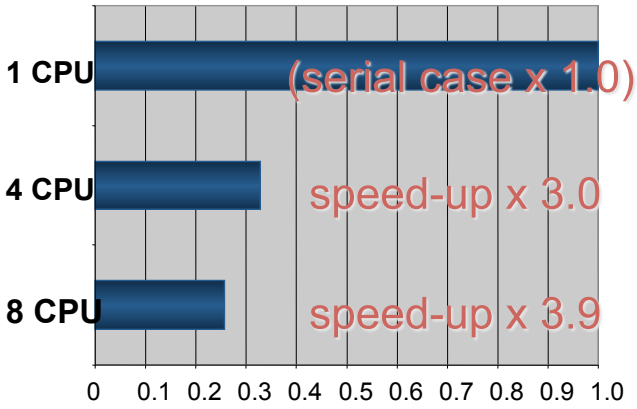
Diastole WSS







NGI_IT infrastructure has been successfully set up and used to efficiently run parallel CFD simulations in large arteries



Talk at EGI-TF12: https://indico.egi.eu/indico/contributionDisplay.py?contribId=33&sessionId=55&confId=1019

Other Collaborations

- **INAF**
  - **EGI Virtual Team to support the ESFRI CTA project**
    - (http://www.cta-observatory.org/)
- **ENEA (cresco)**
- **UNINA (scope)**

All interested in HTC/HPC interoperability tools

- **Workflow splitting on different infrastructures**
- **Unloading of low latency cluster from scalar jobs**

Conclusion

- Many of the supported communities had, are having, or can have advantages in using the IGI infrastructure
  - Even with just an opportunistic usage
  - In particular genome sequencing application

- Support to parallel codes and small HPC is now improved on IGI
  - There is still a lot to be done
    - Increase MPI support at sites, deploy more low latency networks
    - Add GPU, ManyCore Accelerators support

- Transparent interoperation between Grids and Supercomputing is of interest for various communities in different scientific domains

- It is important to support new communities in porting their applications and in creating the computing models since the very beginning
  - To minimize the learning curve
  - To maximize the exploitation of the infrastructure capabilities

References

- https://www.italiangrid.it/
- https://wiki.italiangrid.it/twiki/bin/view/UserSupport/WebHome
- https://portal.italiangrid.it/
- https://wiki.italiangrid.it/twiki/bin/view/MPI
- https://www.egi.eu/
- http://www.eu-emi.eu/

Contact: user-support@italiangrid.it