EU-T0: data research and innovation hub

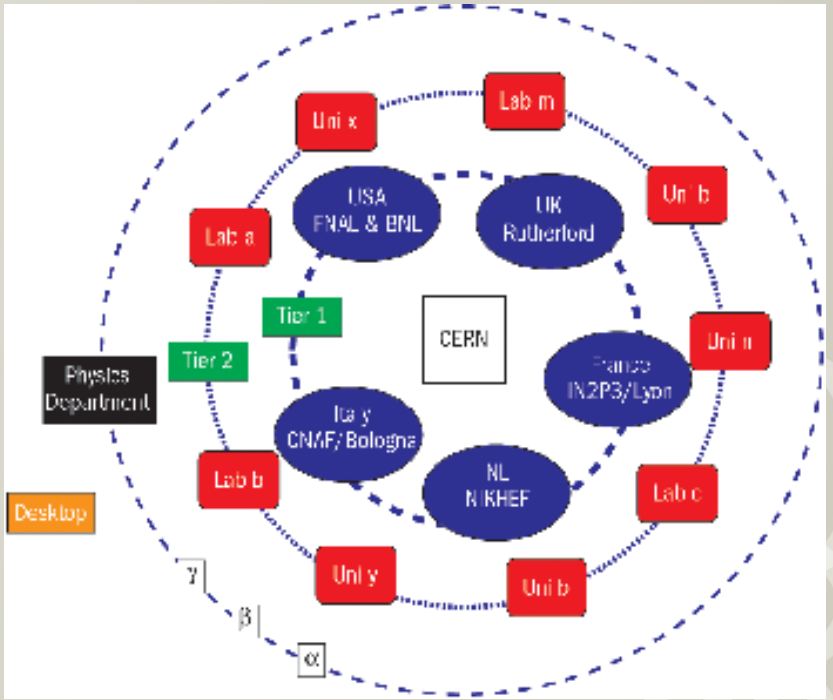
Donatella Lucchesi
Universita' & INFN Padova

A little bit of history

Before LHC

- ✔ "small" amount of data (~TB) and few users (~100)
- ✔ analysis and simulation performed using local farms.

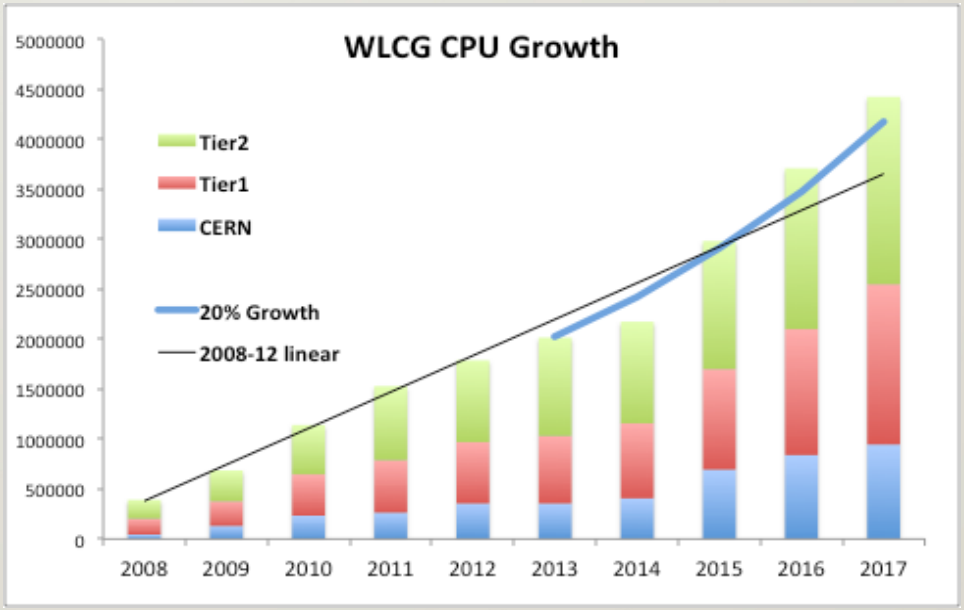
LHC run 1 era

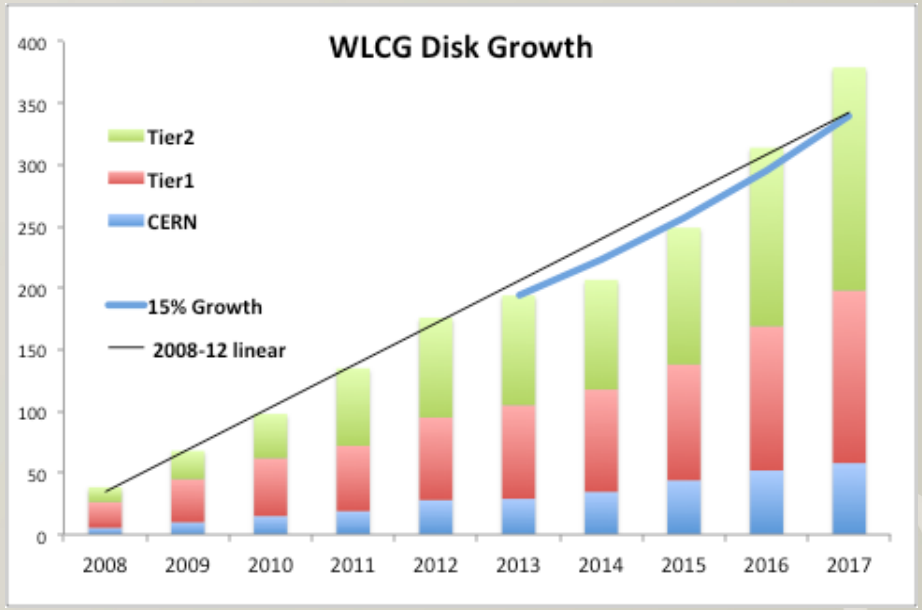


- data pre-placed on Tiers, network too expensive and not fast enough
- GRID = glue among sites
- WLGC = the orchestrator
- EGI (European Grid Infrastructure) with NGI (National Grid Infrast.) supposed to become the e-infrastructure for any science
- Hardware capacity increased 2x each year at the same price. (Moore's & Kryder's law)

Today: Expected Resources Needs

- LHC:

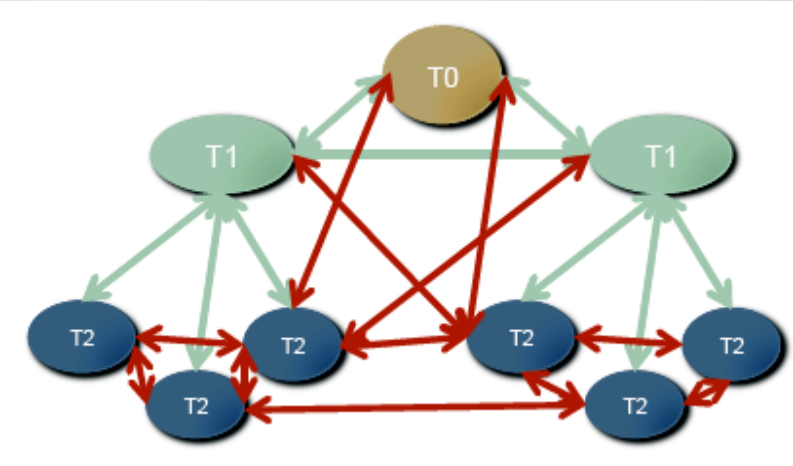




- APPEC (AstroParticle Physics European Consortium) experiments are entering the same game

Today: Technology Evolution

High speed network already changed LHC computing model



- Everybody talks to everybody
- Tier-0 is distributed
  (CERN - Wigner center)

Recent Estimates (Panzer CERN 2013) of hardware capacity increasing:

CPU: +25%/year

Disk: +20%/year

But:

- change experiments software to exploit *many-cores* devices
- high speed disk access still to solve
- develop data movement in the *cloud* framework

Today Facts

- Tier1's and Tier2's infrastructure works very well
- Computing centers are coping with technology changes and experiments computing models evolution in a not fully coordinated way
- NGI model proposed by EGI never took off in Europe
- HEP computing alone is not a strong competitor for H2020

- CERN with EIROforum worked out papers for Research Accelerator Hub(s). Presented in different places including WLCG-POB. Proposals were not well received neither by EC nor by major FA
- EGI-2 proposal has few millions specific funding in H2020, but EGI will have no sustainable future without radically changing its model

So, what to do?

The idea:

Federate major computing and data process centers of Particle, Nuclear, Astro-Particle Physics, Cosmology and Astrophysics into a integrated distributed infrastructure, a virtual European Tier0 data and computing center around which all other national centers revolve and from which all concerned national e-infrastructures radiate.

The uniqueness of *our* infrastructure is the foundation stone.

Starting EU-T0 Federation

February 11th

Position statement signed by

IN2P3

INFN

STFC-UK

DESY-DE

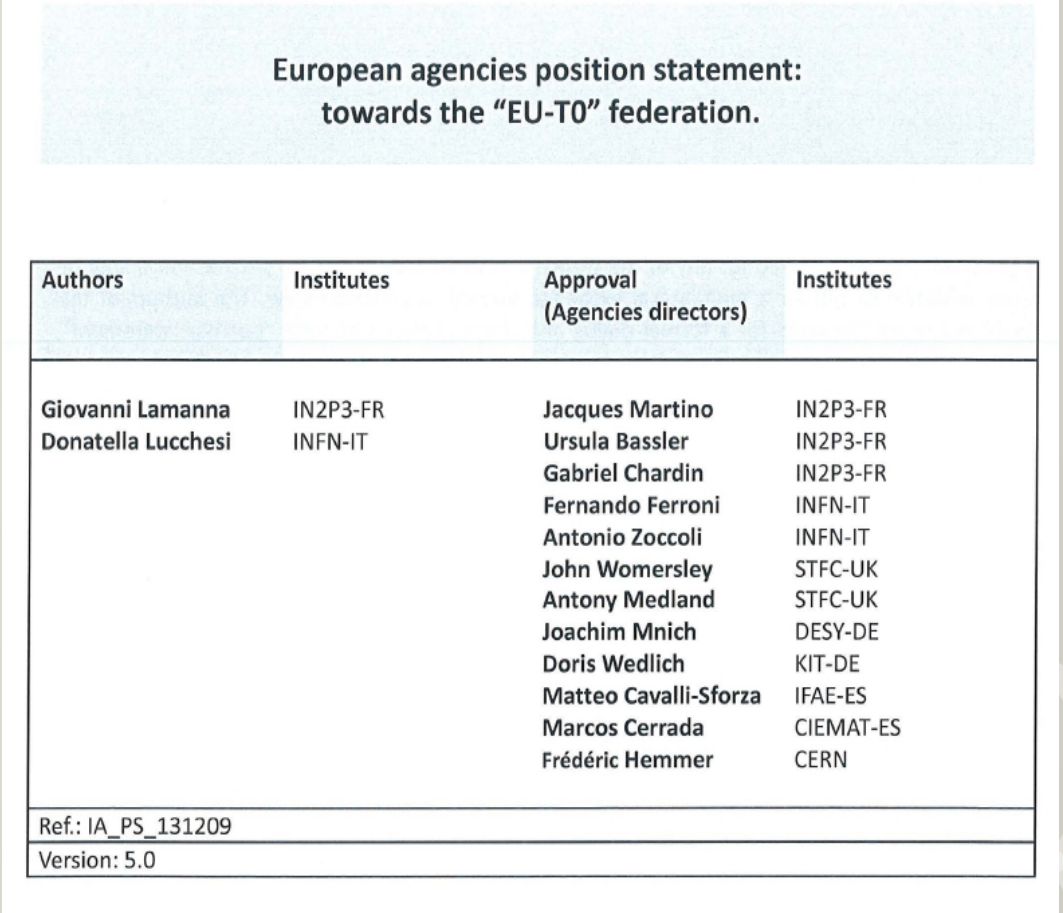
KIT-DE

IFAE-ES

CIEMAT-ES

CERN

NeIC (Nordic e-Infrastructure Collaboration) asked to join



European agencies position statement:
towards the "EU-T0" federation.

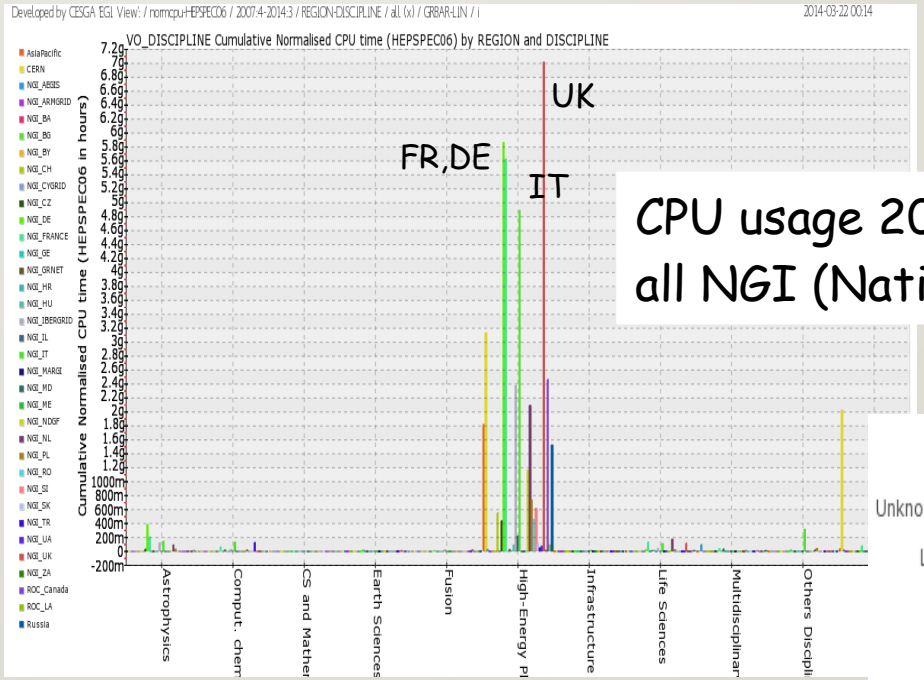
Authors	Institutes	Approval (Agencies directors)	Institutes
Giovanni Lamanna	IN2P3-FR	Jacques Martino	IN2P3-FR
Donatella Lucchesi	INFN-IT	Ursula Bassler	IN2P3-FR
		Gabriel Chardin	IN2P3-FR
		Fernando Ferroni	INFN-IT
		Antonio Zoccoli	INFN-IT
		John Womersley	STFC-UK
		Antony Medland	STFC-UK
		Joachim Mnich	DESY-DE
		Doris Wedlich	KIT-DE
		Matteo Cavalli-Sforza	IFAE-ES
		Marcos Cerrada	CIEMAT-ES
		Frédéric Hemmer	CERN
Ref.: IA_PS_131209			
Version: 5.0			

Where we are:

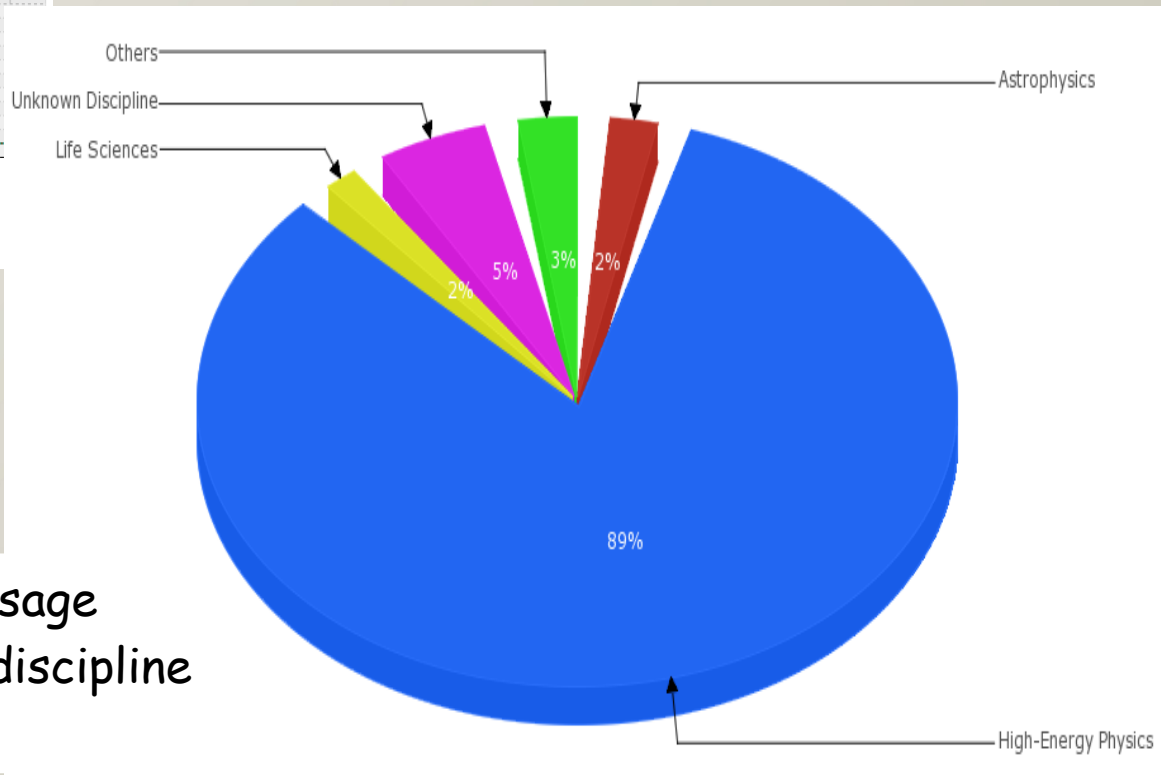
- Collaboration Agreement almost ready to submit to legal offices
- Statement of Interest document first draft out

Why only Particle, Nuclear, Astro-Particle Physics, Cosmology and Astrophysics?

The centers are mainly WLCG T1 and T2



CPU usage 2007-2014 by discipline, all NGI (National Grid Infrastructures)



Share of CPU usage 2007-2014 by discipline all NGI

Analysis of the *old* model

SIENA (Standard and Interoperability for Einfrastructure implemeNtation initiAtive http://www.sienainitiative.eu) report on *Roadmap on Distributed Computing Infrastructure for e-Science and Beyond in Europe*
in 2012 writes on NGI/EGI:

- funding beyond 2014 is critical, sustainable funding scheme not defined yet
- lack of resources ownership
- high fragmentation of resource providers → high operation costs

New Trend in Europe

New or old European Research Infrastructures secure computing resources funding from FA:

- ELIXIR (Life science) identified nodes in the consortium
- LifeWatch (Earth science) has IT research center
- CLARIN (Arts, humanities and social science) has certified centers

The trend is not to have global initiative anymore but virtual hubs federating major computing centers to offer resources and services

EU-T0 Initiative

A strategy focusing on a list of main scientific and technological priorities:

- homogeneous evolution of current DCI computing model;
- shared archival and data access/analysis services, together with tools and applications for scientific data analysis;
- federating interdisciplinary modern software developments;
- issues, models, policies and services for data preservation;
- interfaces with the private sector aiming to be a "pilot" public/private data management and processing system;
- overtaking High Performance and High Throughput Computing paradigms;
- preparation of "data scientists" aimed to lead the major changes in e-Science

EU-T0 at work

The list of priorities already matchs the H2020 calls

In progress:

- establishment of working groups to write projects depending on national interest
- workshop @CERN the week of April 28th to setup H2020 calls
- web site http://www.eu-t0.eu/

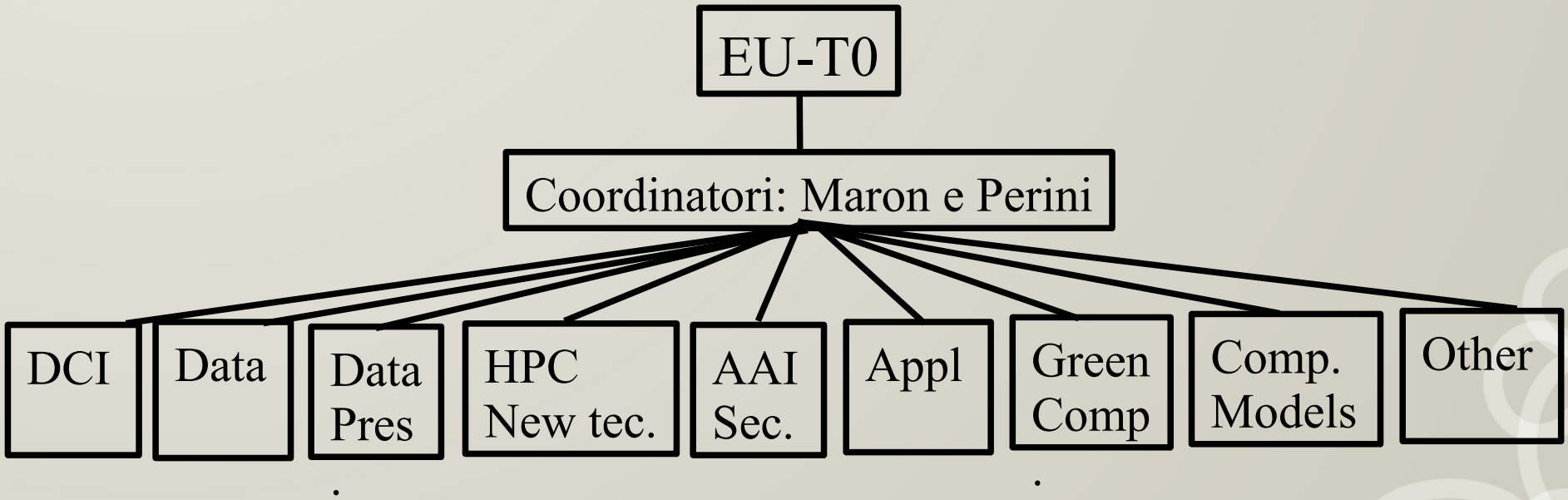
Open issues (not a complete list) :

- Relation with other already existing e-infrastructures, in particular with EGI
- Get CERN spontaneous collaboration
- Collaboration with other hubs
- Collaboration with industry

EU-T0 at work in Italy

INFN:

- Laura Perini and Gaetano Maron are the EU-T0 coordinators
- Working groups definition is in progress
- Several present and future computing activities, in particular related to H2020 calls, could be coordinated within this project



very preliminary organization scheme

EU-TO at work in Italy

Open issues in INFN:

- Relation with other already existing e-infrastructures, in particular with IGI within INFN
- Collaboration with other entities
  - already started the discussion with INAF
  - local and EU entities
- Collaboration with industry

List of projects where INFN is involved (thanks to C. Vistoli)

Progetti Nazionali:

- Progetti legati all'attività INFN
- DHTCS-it
- ReCaS - PON Ricerca e Competitività 2007-2013
- Prin-Stoa

Progetti Interdisciplinari (PON smart cities)

- Prisma (PiattafoRme cloud Interoperabili per SMArt-government)
- OCP (Open City Platform)
- Cagliari 2020

Trasferimento tecnologico

- Mcloud (concluso)

Infrastrutturali

- EGI-Inspire (termine 31/12/2014)
- WLCG

List of projects where INFN is involved (thanks to C. Vistoli)

Progetti Interdisciplinari e/o extraeuropei

- agINFRA (data infrastructure for agriculture)
- Chain-Reds(Coordination and harmonisation of e-infrastructure for research and data sharing
- DCH-RP (Digital Cultural Heritage Roadmap for Preservation)
- BioVel (Biodiversity Virtual E-Laboratory)

Collaborazioni in ambito IGI

- SPES INFN (Selective Production of Exotic Species)
- Chimica computazionale: Uni. Pg, Uni. To, CNR-ISOF
- Scienze ambientali: EMSO (European Multidisciplinary Seafloor Observatory) (esfri), DRHIM (Distributed Research Infrastructure for Hydro-Meteorology) (FP7 proj.) CIMA(Centro Monitoraggio Ambiente)
- Bionformatica: CNR-ITB, Uni. Bo

Partecipazione JRU

- Elixir (European life science infrastructure for Biological Information)
- Life Watch : in corso di definizione

To conclude

- ✔ Computing is one of the major ingredients to be successful in our Research
- ✔ Based on the successful infrastructure built for LHC and on the experience gained so far INFN and IN2P3 are proposing a new model: the EU-T0 federation
- ✔ EU-T0 is in line with the European trend of having science specialized hubs
- ✔ EU-T0 is setting up working groups to preparing H2020 calls and to start the collaboration among centers
- ✔ EU-T0, if successful, can become one of the future Research Infrastructure

BACKUP

Matching H2020 Calls

Call H2020-EINFRA-2014-1

- EINFRA 2-2014 – e-Infrastructure for Open Access (15/4/2014)

Call H2020-EINFRA-2014-2

- EINFRA 1-2014 – Managing, preserving and computing with big research data (2/9/2014)
- EINFRA 3-2014 – Towards global data e-infrastructure Research Data Alliance (2/9/2014) taking part at a project aiming to:
  (2) active participation of European stakeholders (organisations and individual experts) in RDA and leadership initiatives in strategic working group activities;
- EINFRA 4-2014 – Pan-European HPC infrastructure and services (2/9/2014)
- EINFRA 6-2014 – Network of HPC Competence Centres for SME (2/9/2014)
- EINFRA 7-2014 – Provision of core services across e-infrastructures (2/9/2014)

Matching H2020 Calls

Call H2020-EINFRA-2015-1

- EINFRA 5-2015 – Centres of Excellence for computing applications (14/1/2015)
- EINFRA 9-2015 – e-Infrastructures for virtual research environments (VRE) (14/1/2015)

Call H2020-INFRASUPP-2014-2

- INFRASUPP 7-2014 - e-Infrastructure policy development and international cooperation (2/9/2014)

Call H2020-INFRASUPP-2014-2

- INFRASUPP 4-2015 - New professions and skills for e-infrastructures (14/1/2015)

Matching H2020 Calls

Call H2020-ICT-2014-1

- ICT 07-2014 Advanced Cloud Infrastructures and Services (23/4/2014)
- ICT 15-2014 Big data and Open Data Innovation and take-up (23/4/2014)

Not possible to participate as EU-T0, but if there is any international project involving at least one of our Agencies, it would be a great opportunity to join as EU-T0.