



Davide Salomoni, INFN CNAF
(on behalf of Donatella Lucchesi)
Catania, 28/5/2014

Acknowledgments

- The content of these slides was substantially contributed by Donatella Lucchesi and Giovanni Lamanna.

What is EU-T0?

- It is a **“data research and innovation hub”**.
- EU-T0 is a Formal Association of major European research institutes and funding agencies – CERN, CIEMAT-ES, DESY-GE, IFAE-ES, IN2P3-FR, INFN-IT, KIT-GE and STFC-UK – which between them **support research projects in disciplines ranging from particle, nuclear, astro-particle physics to cosmology, astrophysics and photon science.**

Who are the EU-T0 partners?

- The EU-T0 partners are **research agencies owning large computing centers** (CCIN2P3, CNAF, GridPP, PIC, DESY-Tier2, KIT-Tier1, ... + many Tier2 national CCs) **part of the Worldwide LHC Computing Grid (WLCG) project**, having successfully implemented a distributed computing infrastructure **but also...**
- ... **supporting large Research Infrastructures (RIs)** (some in the ESFRI roadmap) **in Astroparticle Physics and Cosmology**, such as AMS, AUGER, H.E.S.S., MAGIC, CTA, FERMI, KM3Net, SKA, VIRGO/EGO and **future gravitational waves projects**, PLANCK, EUCLID, LSST, and **in photon science XFEL**, etc.
- The EU-T0 hub is built up on gathered resources currently of the order of **hundreds of thousands of processing cores**, targeting the half a million cores of computing resources in the next few years. EU-T0 archives big, heterogeneous and complex data through **storage resources of the order of some hundreds Petabytes**, which will grow up at the Exabyte scale already in the next years.

EU-T0 Position Statement



EU-T0: Position Statement

A position statement document has been approved by Signatories during a dedicated inter-agency meeting held at CERN on the 11th February 2014.

The Signatories mandate the authors to bring about the following steps:

- To expand the collaboration.
- To agree on the official roadmap for establishing the "EU-T0".
- To define a detailed work program.

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			

Why EU-T0?

- There is **considerable overlap between the research funding agencies** in Particle, Nuclear, Astroparticle Physics, Cosmology and Astrophysics.
- EU-T0 will **bring the research communities closer to each other** to support their needs and:
 - avoid fragmentation and repetitions;
 - increase cross-fertilization;
 - share standards, expertise and developments;
 - provide and share services;
 - promote outstanding CCs in Europe.

What is the EU-T0 vision?

- The EU-T0 vision is **to create a hub of knowledge and expertise** that optimizes the investment of the funding agencies in proven e-infrastructures by broadening, simplifying, and harmonizing access, driven by well-defined user requirements.

How will EU-T0 bring about its vision?

- EU-T0 will **federate several major European e-Infrastructures for research funded by the partners, into a virtual Tier-0 centre**, enabling and coordinating development of research computing services for many disciplines of physics, and potentially beyond.
- The EU-T0 data research and innovation hub will provide:
 - Federated computing infrastructure and interoperable services to support research workflows of multidisciplinary projects;
 - Software services and tools to the research communities;
 - Data management and data preservation services and solutions;
 - Governance and coordination functions across the federated centers.

What are the EU-T0 reference models?

- The EU-T0 hub is complementary to the model of private-public “research accelerator hubs - Reach” suggested by the IT working group of EIROforum.
 - This means: providing a coherent strategy and portfolio of common services, participating in the creation of a sustainable “European data infrastructure” open to all science communities.
- EU-T0 also addresses the main recommendations contained in the e-Infrastructure Reflection Group (e-IRG) 2013 white paper. This calls for a “general-purpose European e-infrastructure for research” originating from the need for a European e-infrastructures commons.

What are some of the next technical issues to be tackled?

A few examples:

- **CTA**, <https://portal.cta-observatory.org/> :
 - The pipelines will rely on a few CCs archiving O(100) PB of raw-data and need to build an efficient metadata model for data extraction and user access.
 - A worldwide community dealing with proprietary data (privileged users) and observatory products (guest observers). AAI is needed.
- **EUCLID**, <http://www.euclid-ec.org/> :
 - Implementing a computing model where data archives and data processing are distributed among national CCs producing ~500 PB data from hundreds of TB raw and with a central metadata system.
- **LSST**, <http://www.lsst.org/lsst/> :
 - 15 TB/night during 10 years; final image collection : 0.5 Exabytes; ~10 million alerts on transient event per night; catalog of ~37 billions objects.
 - Extremely Large Databases, building Petascale systems (i.e. qserv) which is a highly scalable distributed database system to handle astrophysics queries.
 - Design a coherent approach to run cross queries between them.
- **Gravitational Waves**
 - Willing to create an international network of CCs for sharing data from different antennas and combined observations.
 - Long term preservation policies for important software tools, to read, manage and analyze the data (used for important discovery).

Next steps

- Make the physical infrastructure more accessible by deploying virtualization/cloud services.
- Work with other initiatives to deploy a “single sign on” mechanism.
- Develop new services and software tools as required.
- Organize and provide repositories for sharing services, software, and data across communities.
- Harmonize security coordination – policy development and incident handling (CSIRT).
- Integrate with other forms of e-infrastructures, including capability HPC and volunteer computing.
- Create a training network to create and promote “data scientists”.
- Build on and broaden existing collaborations beyond Europe.

Pillar projects

- A few pillar projects just started:
 - The **EU-T0 data backbone**: heterogeneous storage managed in a federated way; interoperable (EUDAT-compatible); fulfilling the real-time ingestion and the archive access requirements.
 - The **EU-T0 cloud-based platform**: a hybrid distributed computing architecture, covering scheduling, virtualization, cloud computing, storage and networking, configurable for and by scientific user communities, thus serving also the long-tail of science.
 - The EU-T0 VRE and software: organization, repositories and provisioning of services, preserving data; includes new software programming test-bed provisioning across communities.
 - The EU-T0 training: “data scientist” building profile and promoting careers.
 - The EU-T0 pilot: new business model and interaction with private sector for co-developments around big-data services and cloud (Helixnebula).

Conclusions

- EU-T0 is a new path towards a hub of knowledge and expertise that optimizes the investment of the funding agencies in proven e-infrastructures and driven by well-defined user requirements (including those concerning Big Data and Open Access).
- Resources are a combination of hardware, software and skills.
- Data management and data challenges are part of our ESFRI RIs and ERFs.
- Excellence in science is based on excellent data produced by excellent facilities and managed by excellent e-infrastructures. They deserve support in Europe for frontier developments of potential transversal application.