

EU-TO
DATA RESEARCH AND INNOVATION HUB

Giovanni Lamanna

LAPP - Laboratoire d'Annecy-le-Vieux de Physique des Particules,
Université de Savoie, CNRS/IN2P3, Annecy-le-Vieux, France



HEP Software Meeting, CERN 3 April 2014

EU-T0: Integrated Data Management Infrastructures for Science and Technology

EU-T0 is a Data Research and Innovation Hub: a federated European Tier 0 data-management and computing center, implementing a point of connection and coordination among major national e-infrastructures.

The EU-T0 collaboration aims for: development of modern data management services and solutions, deployment and operation of the federated computing infrastructure and interoperable services to support research workflows, improving networking capability and software development in support of multidisciplinary research projects.



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:

- Approach other European funding agencies to expand the collaboration.
- Agreement on the official roadmap for establishing the “EU-T0”, the terms of the agreement among parties and the kick-off of the federation.
- Definition of a detailed work program around a series of simultaneous projects which will be submitted to a series of H2020 calls.

**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			

New major projects currently under design and/or construction in Astrophysics, Astro-Particle Physics and Cosmology are new challenging platforms of frontier developments in e-Science, e.g. CTA, SKA, EUCLID, LSST, VIRGO-LIGO and new GW projects.

A data throughput of the same order of magnitude and larger than LHC, with important data-management issues.

These projects encompass a larger research community bringing together High Energy Physics and Astrophysics agencies.

ISSUES UNDER DEVELOPMENT AND INVESTIGATION IN AP PROJETS:

- Big Data management: e.g. towards new DB technologies, distributed and federated archives, data mining, ...
- New computing models: e.g. distributed cloud infrastructures, new private-public paradigms, sustainable e-infrastructures, ...
- New software and middleware: e.g. parallel programming and MC simulations, new computing and storage technologies, ...
- Data management: e.g. open data access, scientific frameworks and user services, certification protocols, ..

CONSEQUENCES :

Potential shared developments and common deliverables/services.

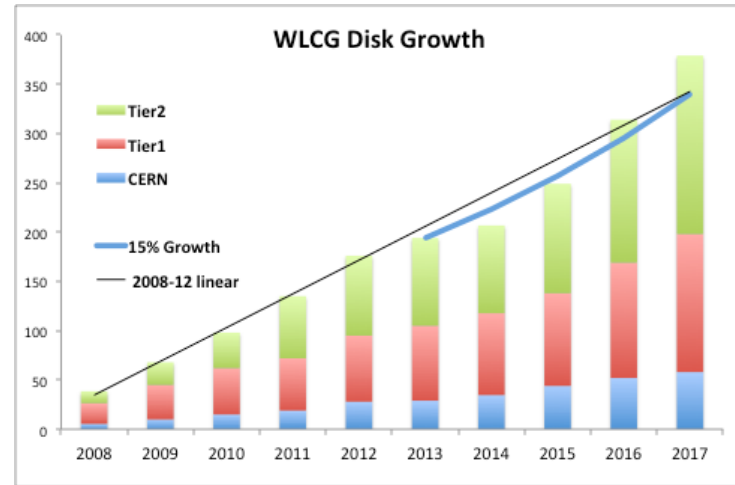
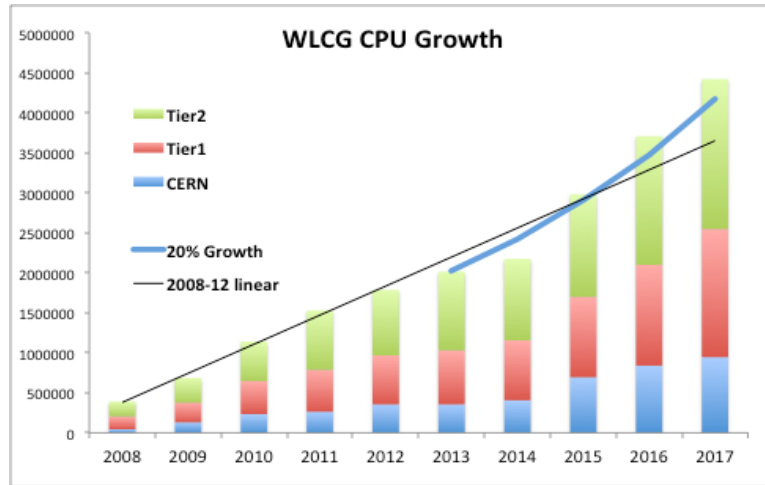
SOME DIRECTIONS :

- Creating an e-infrastructure “Centre of excellence”
- Participation to “e-Science” initiatives
- Promote world-wide cooperation on computing and software
- Training network



LHC experiments Computing Models changed thanks to the high speed network availability towards an almost "democratic" Tiers system.

LHC throughput is already doubling in 2015



LHC resources needs in next years goes beyond what technology evolutions can provide:

CPU: +25%/year

Disk: +20%/year

Moreover in order to exploit technological improvements we need to evolve the code and framework at least to:

- fully use many-cores devices
- move data efficiently in the cloud paradigm
- solve the high speed disk access

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.

DATA ANALYSIS

Activities:

- incubation of forefront software developments;
- collaborative exchanges and common developments of new parallel software for scientific data analysis;
- exploring and adapting new distributed software framework to specific scientific cases;
- adaptation of analysis software to HPC infrastructures;
- GPU software programming incubator and consulting;
- cooperative development and updating of new Monte Carlo simulation radiation-matter interactions software for multi-domain application and compliant with new hardware (e.g. Geant V);
- development and provision of libraries, services and tools for high-level data products visualization and statistical analysis.
- ...

IC-INFRASTRUCTURES

Activities:

- ...
- coherent resource provision for the development and benchmarking of new software on a representative range of candidate computing architectures and new distributed software framework;
- ...

DATA ACCESS

Activities:

- ...
- provision of services for concurrent and collaborative software development and analysis tools (e.g. repositories, build and test systems, compilers, etc.);
- ...

The EU-T0 collaboration prepares projects to be submitted to some H2020 EC calls. In some cases “software proposals” fits in, e.g. Virtual Research Environment, Integrating activities, Cluster of ESFRI projects.

EU-T0 is defining a path for a long term and sustainable funding scheme.

EU-T0 would aim at contributing to the coordination of European and multidisciplinary partnership in the context of an international HEP-Soft. Initiative (discussed today in this workshop).

EU-T0 is a genuine initiative to federate the Institutes , the Computing Centres and the large research community involved in Particle, Nuclear, Astro-Particle Physics, Cosmology and Astrophysics.

EU-T0 is our domain-based e-infrastructure to support our (large) community and provide multidisciplinary services.

EU-T0 will promote the HEP-Soft initiative including a larger community and look forward to build up a cooperative platform to share development and training.

EU-T0 will organize funding requests in which the HEP(++) Soft. goals will be also considered. Joint and integrated pan-European efforts within EU-T0 are welcome... we are the same people, the same Institutes!

