Scientific computing: an historical and inspired view

Mirco Mazzucato INFN PD

School on Open Science Cloud

Perugia

5-9 June 2017

Personal views based on direct personal experience.

The goal of SC

- Scientific Computing should made available to Scientists, for the available money the largest, most effective and easy to use computing power and storage satisfying their (different) specific needs
  - Solutions should take into account specific technical, administrative and sociological constraints
- It has been characterized since its origin by large debates
  - One at the top vs many less performant
  - Parallelism vs High Throughput
  - Commodity vs specialized hadware solutions
  - Need of middleware software specific for science
- I present here my views coming from having taken part directly to many choices as a LEP computing coordinator, LHC Computing Board Chair, Coordinator of INFN grid ….

EMI INFSO-RI-261611

The beginning : no choices

- Scientific Computing –SC starts in 60s with the second generation of very large mainframes based on transistors becoming open to general scientists
- In Italy unique Computing Center in Bologna placed at Centro Nazionale per l'Energia Nucleare (CNEN) open to scientific computing
  - No choices: in 60s Computers were too expensive to be distributed
  - Bologna was chosen for its geographical central position
- For a long time SC in Italy and elsewhere has been a sequence of IBM or Cray mainframes with growing power 704, 7090/.. -> 360
- CINECA in 1967 became in Italy the first center dedicated to Scientific Computing (agreement of 3 Universities: Bologna, Padova and Firenze) and it is also today for Supercomputing

70s and 80s: from Mainframes to PC

- In the 70s the IC continue to develop and the first microprocessor appears: Intel 8008 (F. Faggin, Physiscist of Padova was the head)
- MINI computers like DIGITAL PDP 11 diffuse IT know how around i the world and in many INFN sites
  - Used mainly for equipment control, but also for the most simple scientific computations
- With HP 9825 based on 8 bit microprocessor starts the desktop computing which will evolve in 1981 into the IBM PC
- In 80s HP/Digital/… Desktops with 32bit start to appears in all Scientists desks and for equipment control
  - a significant turning point for Scientific Computing
  - Vizualisation and statistical analysis can be done on a private system
  - Large spread out of know how
  - Time sharing, UNIX, WAN Networking (Arpanet) ,Fortran become available

Differentiation driven by HEP LEP collaborations

- In 80s mainframes like IBM 3090 were still the only reference for the large computations and also for the *computing models of HEP LEP* Collaborations developed in the middle of ‘80s
  - With LEP collaborations pass from 5-6 teams and a total of about 30 scientists to 100 teams and 500 scientists and have to cope with an explosion of data
  - The request from LEP of more friendly collaborative tool brings Tim Berners Lee supported by Robert Cailliau to develop the WWW: HTTP, HTML and Browser in Next
- New solutions required to satisfy the needs for processing and analysing LEP data. CERN mainframe already saturated in ‘89 by preliminary calibrations and preparatory work. Not money to buy another CPU
- 2 LEP teams started to think how to exploit commodity desktops based on 32 bits microprocessors for LEP data analysis
  - What matters is the ratio between the time required by the Ethernet bandwidth to move in memory from a disk a new event to process and its CPU processing time
- Industry policies and legacy was saying that it was not feasible, extensive measurements showed that 10 Mbps were enough:
- a lesson to learn!

The 1rst technological revolution

- Beginning of ‘90s two LEP pioneer computing cluster based on commodity personal Digital desktops showed a level of reliability, cpu power and network bandwith very competive with mainframes for data processing and analysis
  - The Aleph Farm showed that clusters could be used effctively for data acquisition and event filtering
  - The Delphi INFN Farm proved during a month of real tough competition in ‘91 that a Digital cluster of 12 desktops (350KSfr) was 1.5 more effective than a mainframe 3090 CPU (6MSfr) for LEP event reconstruction and simulation
- The era of Mainfraimes for HEP and in general for Scientific data processing and analysis was at the end !
- The whole Delphi processing was moved to the INFN Farm
- Differention starts: HPC/Parallel computing stay on mainframes
- Clusters and High Throghput Computing start a new tech. branch
  - Cluster Management Systems , Lan and Broadband networking architectures …

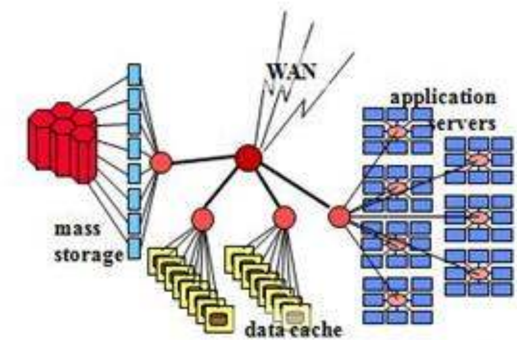
LHC: The foundation of distributed computing

- In 90s locally distributed systems had spread out in all Labs and Universities
  - Clusters had continuosly grown in dimensions
  - Large advances in Local /Wide Area Network and Cluster Mangement Tool
  - 1000-way cluster were largely available
  - All Scientific Computing models however remained central , but practical activites became naturally “manually” distributed
- **CERN LHC drived the 2nd revolution to enable effective distributed computing**

Distributed Computing - 2000's

- Giant clusters → *fabrics*
  - New level of automation required
- Geographically distributed systems
  - Computational models

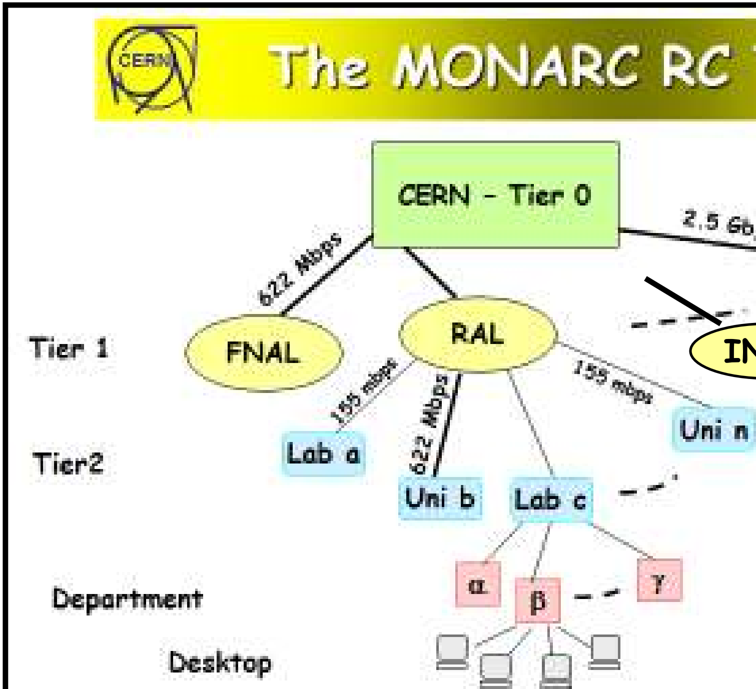
- Key areas for R&D
  - Fabric management
  - Grid middleware
  - High-performance networking
  - Grid operation



Distributed Computing Infrastructures : Not only technology but a lot of Social and Political issues)

- At the end of 90's 5 important initiatives have provided a favourable background for the foundation of a really distributed SC infrastructure driven by LHC at CERN
- The LHC Monarc Computing Model (H. Newman&L.Perini)
- The preparatory action of CERN and EU partners to prepare an EC project for the LHC computing and with US
- The launch of the eScience program in UK by T. Blair
- The action of the INFN Presidency setting up the CNTC Committee for the new technologies for LHC
- The publication of the Book "The Grid " from Ian Foster and Karl Kesselmann in summer 1999

The MONARC Multi-Tier Model (1999)



The LHC experiment's Computing Model officially Include for th 1rst time Centers outside CERN The issue became: How??

In addition for LHC Fortran was abandoned in favour of C++

1. The preparatory actions

- EU-NSF discussions on transatlantic collaboration on IT subjects
- EU-US workshop on large scientific data bases and archives in USA in September '99
- Meeting between EU and HEPCCC in November '99
- Kick-off meeting at CERN for an EC project proposal on Jan 11, 2000

2. The eScience program in UK

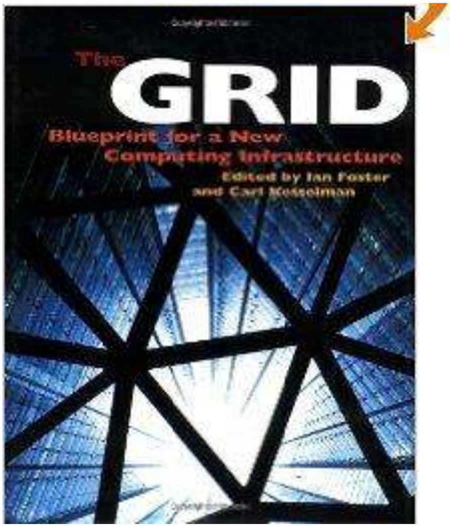
- **The e-Science** program was launched by John Taylor, the Director General of the United Kingdom's Office of Science and Technology of the Blair Cabinet in 1999
- Large funding initiative planned to start in November 2000 to create new computational tools and infrastructures to support scientific discovery.
- Particle physics took the lead with the program for the preparation of LHC

3. The CNTC in Italy

- At the end of ’90s the INFN Presidency took the initiative of setting up a Committee (CNTC) for developping the new IT technologies and infrastructures necessary for the LHC Computing
- All INFN IT top experts collected in a unique board to discuss and plan a general solution for the LHC computing
- The INFN Grid project was proposed by CNTC as the first grid project in Europe at the end of ‘99 and approved for an initial funding by the INFN Board of Directors in February 2000

EMI INFSO-RI-261611

Spring '99 :The Grid Book



The Grid: Blueprint for a New Computing Infrastructure

Ian Foster (Editor), Carl Kesselman (Editor)

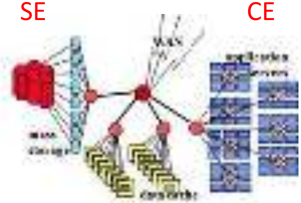
-> A. Ghiselli proposed this book to me (President CNTC) in July '99

EMI INFSO-RI-261611

The open issues in LHC SC in ’99

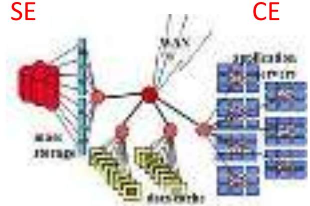
- Data Centers based on all sort of different legacy tools
  - *Batch Systems (Fork, LSF, PBS, Condor, NQE.....)*
  - *Storage Systems (all sort of combinations of different tools)*
  - *Tape storage systems (many solutions)*
- Need for Security , delegation to access different services...
- Centers Directors not willing to change their legacy solutions and the current way user were working
- Nothing was commercially available and it was not easy !

SE CE



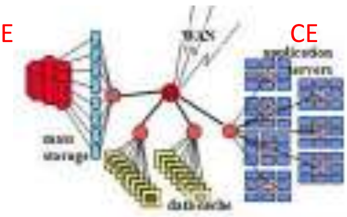
Center A

SE CE



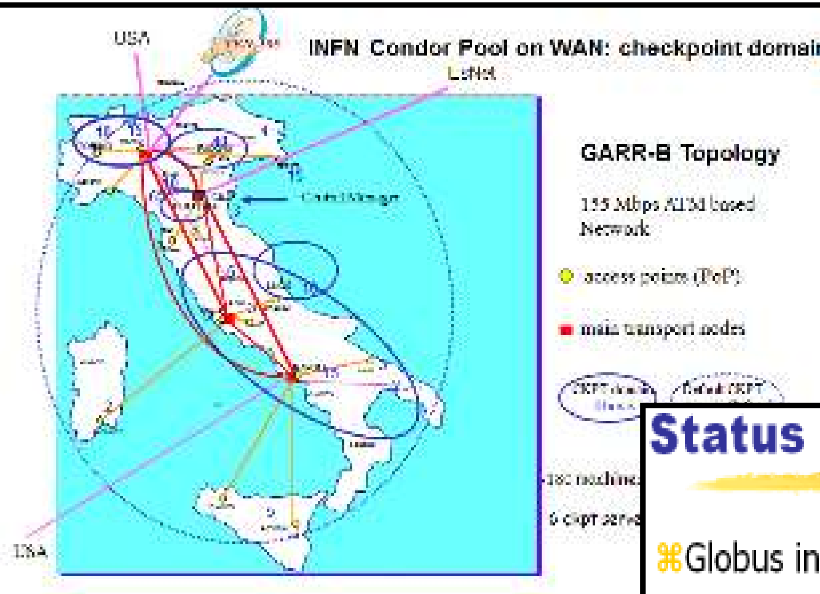
Center B

SE CE

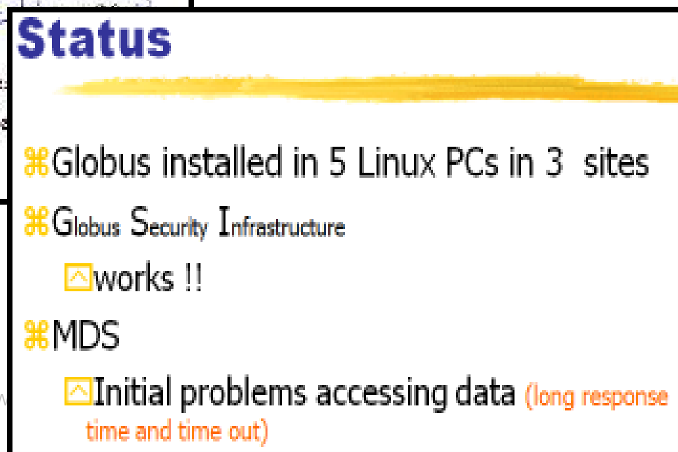


Center C

INFN unique pioneer experience on Condor Pool



In late '90s first hands-on experience in INFN on Distributed Computing: Development of Condor on WAN
A. Ghiselli (CNAF)



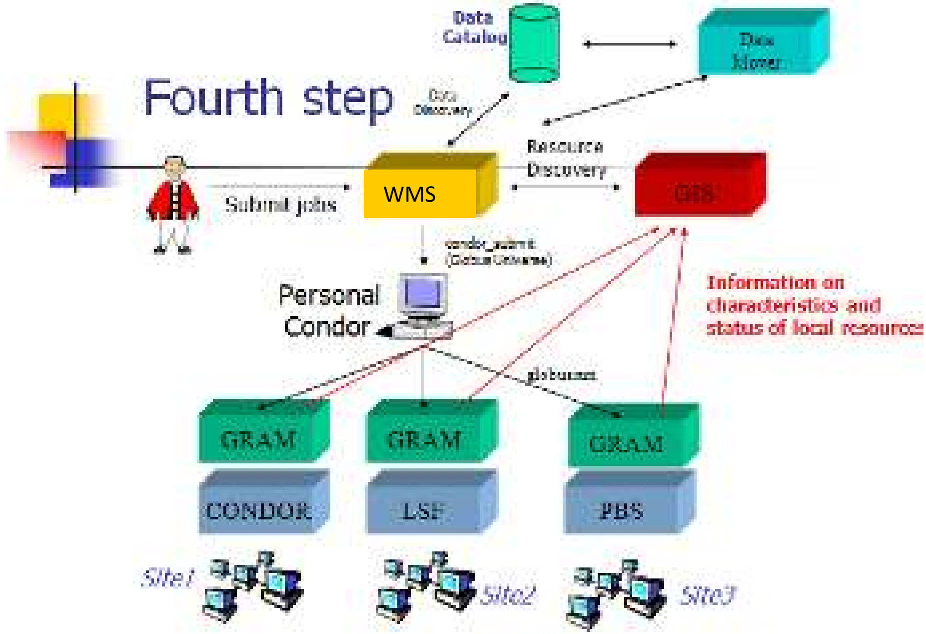
Launch of Globus tests in INFN in September 1999
-> Look for a multi-proprietary solution

The start of DCI : CHEP 2000 in Padova

- LHC with its regional centre architecture (MONARC) becomes the driver for the establishment of a new type of Computational and Storage Distributed Infrastructure in Europe and Worldwide
- Focus was on Scientific Computing needs but all problems were addressed having in mind a broader view
- Ian Foster and Karl Kesselman invited by me at the last moment came on Saturday and Sunday after the end of the conference to explain Grid and Globus
  - 100 people remained there during the whole week end
- Confirmed the choice for an EC project

  -> Launch of the DataGrid proposal
- Start Globus test program coordinated by CERN

6/2000: INFN proposal for Datagrid architecture



M. Sgaravatto Giugno 2000

All main components were there! We could start to devleop

The DataGrid EC Project (10 M€)

- DataGrid in 2001 made available the money to develop the Software services for the first "prototype" Distributed Compute&stoarge infrastructure in the world
- Capable of:
  - Managing large amount of distributed data
  - Providing high throughput distributed computing
  - Managing both local computing fabrics and wide-area GRID infrastructure
- New services developed to support distributed computing
  - Growth of the Open source community in Europe aiming at producing general solutions valid for all Science and beyond
- Every step was proved with real testbed demonstrators
  - Very clear guidance from communities :LHC, ESA , Bio

The Global Grid Forum: Standards

- In October 2001 after the approval of Datagrid international agreement to create the Global Grid Forum (mainly Research a Academia people) having the aim of defining world-wide standards
- Transformed in Open Grid Forum in 2006 (+industry)
- Balanced management between EU and US
- "The GGF working groups are investigating best practices for the design and interoperation of distributed systems, and the development of recommendations regarding the implementation of grid software"
- Since then some successes like SRM protocol for storage access or GLUE for the information model

EMI INFSO-RI-261611

The components of DataGrid infrastructure

- Access API
- Security Framework
  - Authentication, Authorization, Policies, Delegation
- Workload Management
  - Resource Broker and job submission
  - General Compute interfaces to local systems
- Data Management
  - Metadata Catalog and File&Replica Catalog
  - General Storage interface to local systems
  - Data Mover
- Information, Application Monitoring and Accounting
- Everything had to be devloped and it has been !
- INFN provided about half of the required services

EMI INFSO-RI-261611

Information Modelling and Resource Discovery

- While it was developing WMS in DataGrid, INFN + CERN started the DataTag (2002) EC project with CERN and in collaboration with C. Kesselman of Globus
- The goal was to look at interoperability
- INFN defined the grid information model as the cornerstone for information discovery across multiple infrastructures
- S. Andreozzi was co-author of the GLUE standard specification at the Global Grid Forum
  - *GLUE is still today the foundation of Information discovery, one of the EGI core infrastructure services*
- One example of successful specifications of GGF

The Storage Resource Manager-SRM

- The interface to Storage took more time than CI Globus GRAM-Condor and it came only in 2006/7 thanks to a very long period of test of SRM v2.2 interface implementation
  - Space reservation
  - Direct access to reserved storage space
  - Space management policies
- SRM V2.2 implemented for:
- BeStMan (LBNL), CASTOR (CERN and RAL), dCache (DESY and FNAL), DPM (CERN), and StoRM (INFN and ICTP).
- Globus GRAM replaced by INFN CREAM still in use doday as Compute Interface of EGI

The VO Membership Service

- The Virtual Organization Membership Service is today a cornerstone of EGI and OSG (US) security infrastructures and allows authorized access to resources based on user membership
- Inspiring principle : Centers managers remain always in control of their resources
- Alternative to Globus CAS: resource allocated to VO once
- The service which implements the VO membership management (VOMS) was developed by INFN and is deployed across all EGI resource infrastructures
- Nice example with STORM and ARGUS of evolution from an initial national development to a wider usage

2002: The LHC GRID Comp.Project

- Fostered by CERN and LHC Computing coordinators with th objctive of establishing a worldwide unique collaborative Grid platform for the LHC DCI
- 2002-2005 – LHC Computing Grid – LCG project
  - Deploying the results of DataGrid providing a production facility for LHC experiments
  - Clear European leadership more and more unacceptable for US
- 2005/6 US impose for LCG a change of goal
- Not obligations to develop and adopt a unique Grid platform
- WLCG as a LHC DCI should be composed of independent national Grids
  - US started the OSG project to develop US Grid services

2002. The Grid.it project

- Grid attracts the interest of the Italian Ministry and other Italian Research communities
- The Grid.it national project (funded by MIUR for a cost of 12 M€) was financed to develop a national Grid infrastructure in Italy
  - INFN, CNR, Astrophysics, Geophysics ,Universities, Compute Science…
  - Software development, Prototyping of operations centre services, User requirement driven
  - Extension of service functionality for non HEP sciences
    - *VO membership service (VOMS), SRM(StoRM) for parallel file systems, authorization framework **(ARGUS)***

2003. Start of EGEE operation model

- The success of DataGrid and Datatag open the way to the EGEE series (about100 M€ from EC in 6 years)
  - About 20 M€ to Italy, mostly INFN
- Definition and implementation of the European Grid tiered operations model
  - Local support delegated to Regional Operations Centres (ROCs)
  - Central support (at CERN initially became distributed)
- DCI managment is distributed
  - C. Vistoli becomes Coordinator of the EGEE ROCs
  - Central incident Management system in Karlsruhe
  - Middleware coordination to INFN: Grandi /Giacomini

2006. The National Grid Infrastructures

National Grid Infrastructures develop all around Europe

The Italian Grid Infrastructure is a mature production computing infrastructure

Including INFNGrid resources and operated by the Grid Operations team at CNAF

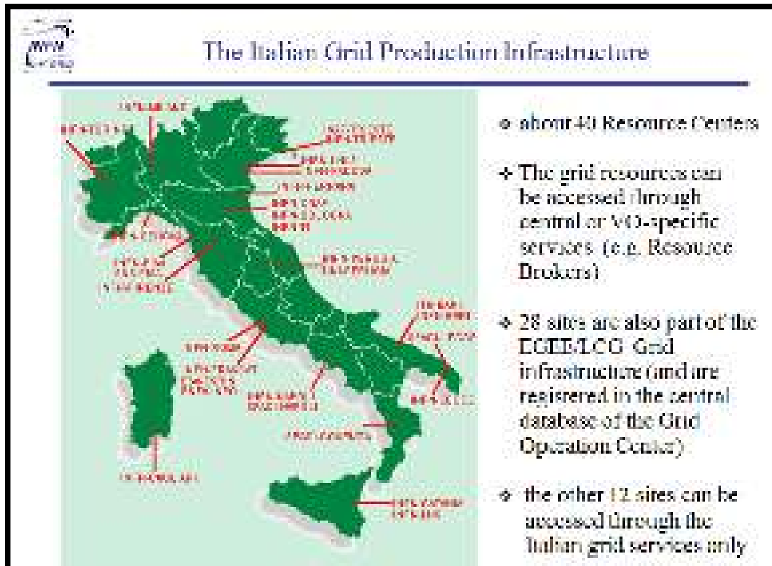
Grid monitoring, accounting, user and operations support, on duty support teams

Multiple operations services at CNAF

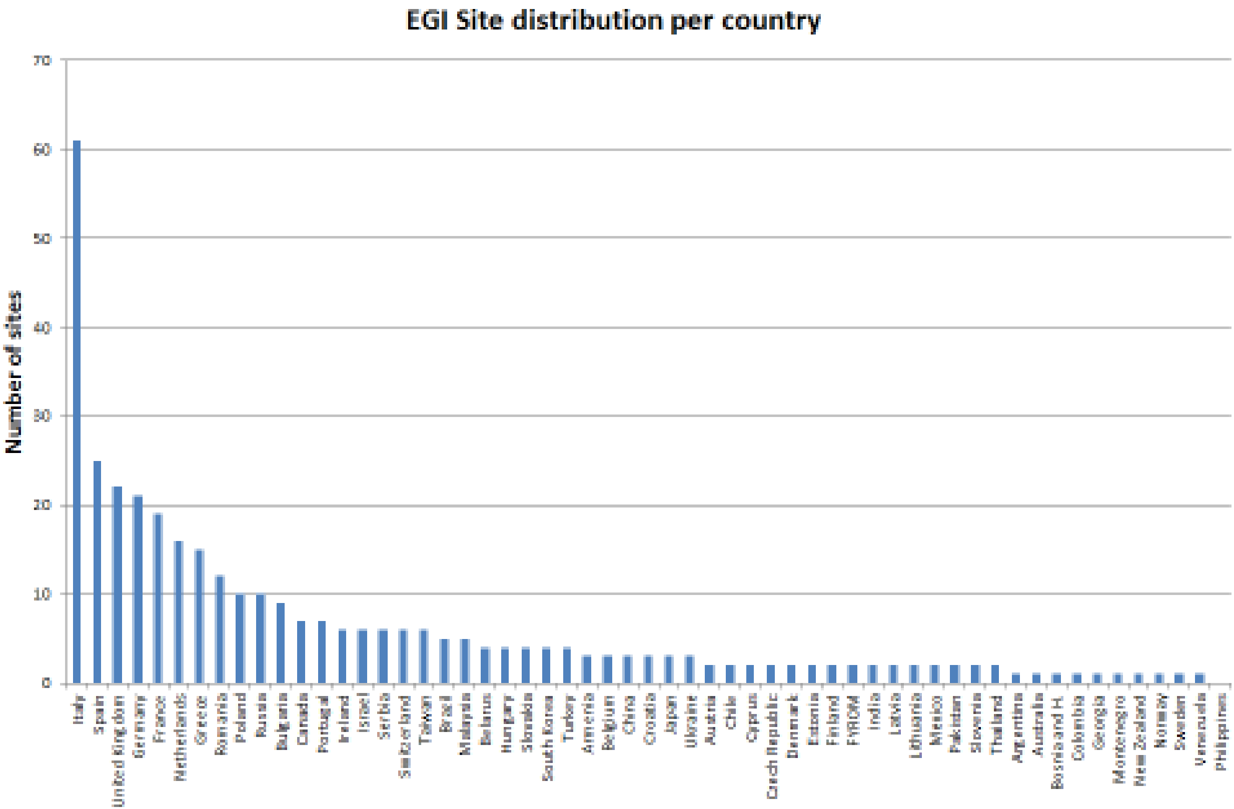
23 multidisciplinary VOs

40 resource centres

28 EGEE sites



Italy always with the largest n. sites



EMI INFSO-RI-261611

2001-2013: The florishing of projects

…..but not real progress in architecture and services compared to Datagrid/EGEE I

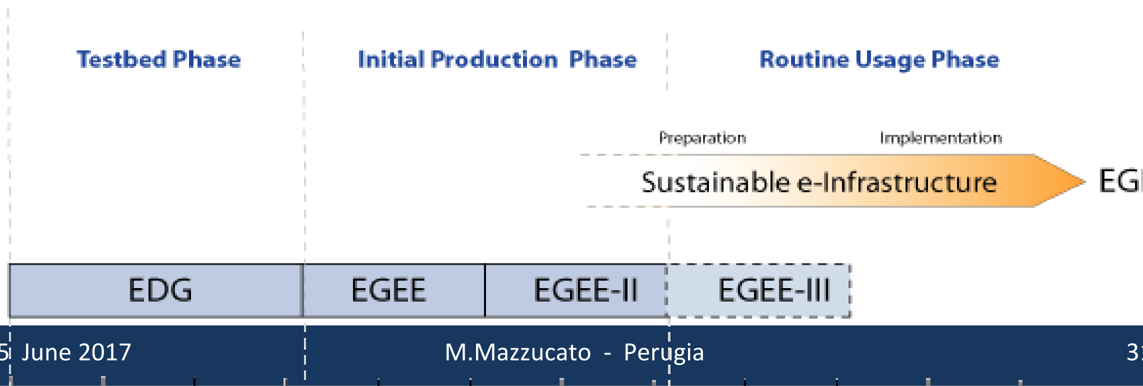
2010: The review of the EGEE Model

- **The success of EGEE:**
- Has collected together effectively many Research Communities, infrastructure providers and Software developers
- Has defined the European Grid operation model: a world success
  - Local support delegated to ROCs
  - Central support (CERN originally, later distributed)
- Work was effectivly distributed around Europe
- The EGEE Distributed Compute Infrastructure was working in a superb way with more than 99% constant efficiency, **...but**
- The European Physics community and LHC were keeping steering the software development process for 6 years
- Little attention and effort dedicated to the specific needs of Biomed, ESA and other communities (Not yet well organized......)

EMI INFSO-RI-261611

The EC push

- Prepare for permanent sustainable Grid infrastructure
  - Ensure a reliable and adaptive support for all sciences
  - Independent of short project funding cycles
  - DCI managed by a sustainable collaboration of National Grid Initiatives like NRENs manage Geant
  - ...but NRENs just buy bandwidth and services from Telcos while DCIs are made by services developed for Science



EMI INFSO-RI-261611

EGEE to EGI ->The loss of innovation rush

- The European infrastructure operation and management,
  - >EGI Inspire project
- The middleware development -> EMI project
- The Research Communities ; Many independent projects
- **separate from each other with the advent of EGI**
- More difficult to keep innovation going
- EGI is managed by new “bureaucratic” organizations (NGIs) more confortable with the deployment of th same legacy solutions than new developments
  - *Little attentions to new needs*
- The European Physics communities inside WLCG loose their reference project , become completely disorganized and dominated by US
  - Not influence on the EGI management
  - No more steering on WLCG roadmaps

The EMI software project

- Started in 2010 to continue to provide middleware components to EGI satisfying the general prescription of open distributions for DCIs
- **The goal**
- Attempt to armonize the European middleware to guarantee interoperability and consolidate an open software community in Europe to make the next steps
- Move in new technologies and introduce new components to expand the user base and to make acess easier

EMI INFSO-RI-261611

The crisis of EGI and EMI in 2013

- EGI is managed by new organizations (NGIs)
- EMI is managed by software development teams
- Scientific Communities not part of either two
- After 3 years 93.7 percent of EGI usage was still coming from LHC
- Innovation limited to formal little improvements
- The architecture of the EMI middleware was still the one defined in 2001
- No real serious discussion over 3 years on how to move forward and tackle open issues
- No more mobilization/generalization of national efforts

Need for a new tech. revolution

- Grid: Made by Science Communities
  - Very powerful but long, steep learning curve.
  - Provided SSO for DCIs but based only on X.509 certificates.
  - Difficult to use for real-time analysis, visualization, provisioning of complex virtual environments.
  - Storage management normally at the file/block level, not as distributed objects.
- Cloud: Made by Industries (Amazon) -> Openstack-> Private Cloud
- Large open source developments independent from Science
  - IaaS very powerful, simple access to resources, flexible VM, elastic
  - but simple-minded scheduling policies block large adoption by SC Centers
  - Federated Cloud not existing.
  - SSO need to be redeveloped , Attribute Authorities, DCI wide Policies
  - Complexity in the exploitation of distributed Cloud resources due to multiple IaaS and/or PaaS solutions.
  - Commercial vs Openstack
  - Tools for dynamic data handling and preservation
  - Framework for Big Data Analisis (Mesos, Hadoop, Spark…)

The INFN Indigo Data Cloud project

- In 2013 INFN proposed to EC that Europe should face actively the new technological revolution where Data Centers services will move:
- from Batch Systems and Grids
- to IaaS/PaaS and Federated Clouds
- There was a urgent need for Europe to put together:
  - The best teams of software developers
  - The most active research communities
  - The cloud infrastrcture providers (EGI ; Helix Nebula)
- To design and develop the next generation of general open cloud services enabling a DCI for Analysis of Big Data
- The Indigo Data Cloud project (12M€) lead by D. Salomoni was then approved

The INFN Cloud projects for PAs

- MIUR has funded the PRISMA and OCP project (2013) for Smart Cities and Social Communities (27.5 Meuro)
  - Collaboration of INFN, Universities and IT industries
  - INFN (Bari) has lead the initial PRISMA developments of an open source cloud platform for a federated infrastructure for PAs
- Challenge very similar as for LHC in '90s
  - Thousands of small PA distributed centers offering services -> Very expensive
- The new 3 year Digital Agenda plan of Agid and Digital Team (Piacentini) forseees a rationalization with a reduced number of strategic poles (Tier1,2,3) offering cloud services to small PAs
- Many common requirements between Public Administrations and Research for a Federated Distributed Infrastructure
  - Federated Security Framework
  - Standard templates for PaaS
  - Monitoring and Accounting

The Open City Platform project

ATI Marche

Toscana/E.R

OCP: MIUR project for Industrial Research

Duration : 42months (1/1/2014-> 30/6/2017)

Personnl: about 100 units

PM: more than 2000

Budget : 11.949.448,89 euro

MIUR Contribution : 10.688.786,90€

Cofinancing: 1.260.662,00 euro

Start from results of **MIUR PON PRISMA**and thosm of INFN Cloud and Marche Cloud Collaborat with INDIGO-DataCloud (12 M€), coordinato da INFN (D.Salomoni)

PRISMA e OCP have invested ~ 35 M€ to make available a cloudd platform enabling the stsablishment of a fedrated DCI for PAs Data Centers

Not only technology, but a lot of social and political challenges

Europe: Future challenges

- After 16 years need to put in place a new general vision for the commodity Distributed Computing and Data infrastructure for Science in Europe and Worldwide :
- Re-establish effective working relations between Users, SW providers, Centers and mobilization of national efforts
- Make a next step in providing a simple and transparent access to the DCIs using cloud
- Re-involve users directly driving developments
- Consolidate the open source development community in Europe and connect it world-wide
- Mak a true Open Science Cloud

Conclusions

With LHC and Grids Europe got a world-wide leadership in DCIs

Open Source Cloud was born in Europe, as WWW, but we lost the opportunity

Federated Cloud DCI do not exists

Europe could use OSC to

retake the lead both for SC and PAs

EMI INFSO-RI-261611

OCP in sintesi

- Tool automatici basati sui prodotti open source Puppet e Foreman per l' Installazione e la gestione di OpenStack con: HA, Monitoring integrato, Fault Tolerance, Disaster Recovery
- PaaS per l'automazione dell'attivazione di servizi su IaaS a livello di piattaforma eliminando la complessità di gestione dell'infrastruttura IaaS e delle configurazioni e fornitura as a service di un set di servizi .
- Piattaforma per il monitoring dei servizi infrastrutturali e applicativi completamente integrata
- Framework di sicurezza di valenza generale integrato con SPID, con Attribute Authority e Policyas service per garantire single sign on, interoperabilità e gestione di policies federata
- In corso di rilascio finale piattaforma per l'analisi del Big Data basata sull'integrazione di Mesos+Hadoop+Myriad + tools di Data Analytics
- App Store e insieme di servizi applicativi attivabili da questo sulla piattaforma OCP con un click
- Citizen Market Plac come intrfaccia unica verso il cittadino
- Il core della piattaforma è basato su soluzioni Open Source, su API e Standard che sono o stanno diventando leader di mercato per garantire stabilità nel tempo degli investimenti