The role of e-Infrastructure for the preservation of cultural heritage

Federico Ruggieri
Head of Distributed Computing and Storage Department of GARR
Outline

- GARR in a nutshell
- Cultural Heritage and Humanities communities in GARR
- e-Infrastructures services for CH
- International Cooperation
- Conclusions
GARR in a nutshell

- Italian R&E Network
- Providing:
  - High-bandwidth transparent symmetric connectivity
  - Advanced services
  - E-Infrastructure support
- to the Italian Scientific, Academic, Education and Cultural community

- Non-for-profit association founded by CNR, ENEA, INFN and Fondazione Crui (on behalf of Italian Universities) under the aegis of the Ministry of Education & Research
The GARR-X network

- More than 8,500 km of GARR owned fibers
  - ~6,500 km of backbone
  - ~2,000 km of access links
- Over 500 user sites interconnected
- 500 Gbps aggregated backbone bandwidth
- 100 Gbps IP capacity to GÉANT (yes: 100!)
- Cross border fibers with ARNES, SWITCH and UoM.
- Several 10G e2e provided by GÉANT
The user community/1

- Founding members: CNR, INFN, ENEA + Fondazione CRUI representing ~90 public and private universities in Italy.
- Other connected organizations: scientific institutes, laboratories, observatories & other facilities, supercomputing centers, research hospitals, Cultural institutions, libraries, museums and schools.
- Special projects with Ministry of Health, Ministry of Cultural Heritage and Ministry of Education & Research.

[logos of various organizations]
The user community/2

- GARR community is **the whole R&E world**
National and Governmental Libraries

Biblioteca Nazionale Centrale di Firenze
Biblioteca Nazionale Braidense
Biblioteca Riccardiana
Biblioteca Nazionale Marciana
Biblioteca Nazionale di Bari
Biblioteca Marucelliana

22/09/2014 DCH-RP Final Conference - Rome, Italy
GARR-X Progress

- Started on 1 July 2013
- About 3,900 km new fibres backbone
- About 2,200 km new fibres for the access of users
- 24 Points of Presence (PoP)
- Up to 100 Gbps links
- **New Distributed Computing Infrastructure with:**
  - More than 10 PB disk storage
  - More than 8,000 Virtual CPUs
Geographically distributed architecture

Internal network
40-100 Gbps

GARR-X
The GRID Model

Modello based on resource sharing: optimal for omogeneous Virtual Communities. Lacks a model for managing costs and payment of services.

Every site offers and consumes resources in a long term balance.

Resource sharing is limited by the administrative domain of the site that decides the priorities.

The coordinating body has difficulties in having its costs recognised and shared.

22/09/2014

DCH-RP Final Conference - Rome, Italy
The Cloud Model

Based on a clear business model: resources are paid by the users; no compensation mechanism foreseen.

The provider sends invoices to the customer with pre-determined costs parameters and SLA.

Resources are made available by the provider and the usage is regulated by a commercial contract.
GARR manages a core of HW & SW resources that are complemented by other resources made available by the community in a federated model.

Bilateral agreements between GARR and sites define the debit/credit balance.

Users see a coherently managed infrastructure with clear costs.

Resources are managed by GARR with uniform usage policies.
New Services

- **GARRbox** – Personal Cloud storage based on ownCloud

- **Cloud IdP** – A full fledged Identity Provider pre-configured to be connected to IDEM and eduGAIN

- **WebConference** – Web conference system based on AdobeConnect

- **BigStorage** – Data archives and repositories

- **PaaS** (Platform as a Service) or **IaaS** (Infrastructure as a Service)
The solution: not only technological

- Developed under the agreement between Ministry of Health, 57 Organizations (research hospitals and health institutes), and GARR
- Ready to use IDM and IDP platform, hiding tech complexity
- Platform is designed to satisfy IDEM and eduGAIN policy requirements
- Tutoring the Organization on a simplified joining procedure to Fill and Sign all the necessary documents
- Conformant to REFEDS Discovery Guide
e-Infrastructure Services for Cultural Heritage

- **Storage**: Archives, Libraries, Repositories, Replicas, Backup, Data Movement, Disaster Recovery, Long Term Data Preservation

- **SaaS**: Authentication & Authorisation, Web Conference, Video Streaming, Web Portals

- **PaaS**: Science Gateways, Grid Computing, Data Curation

- **IaaS**: Cloud Computing
International Cooperation

- Standards and related services
  - Persistent Identifiers (PID) for Data Objects
  - Metadata retrieval and schema (e.g. OAI-PMH, Dublin Core)
  - Digital preservation (e.g. OAIS model)

- Coordination and harmonisation of e-Infrastructures
- Interoperations and interoperability
Knowledge Base of Data Repositories

www.chain-project.eu/knowledge-base
KB of Document Repositories

www.chain-project.eu/knowledge-base
Conclusions

- European e-Infrastructures are rapidly evolving to provide more services and GARR has a state of the art national infrastructure interconnected with GEANT and the rest of the world
- Cultural Heritage communities can build on top of it both using the new services and proposing new requirements
- The long term stability of those infrastructures is granted by the NRENs
- International cooperation in DCH should aim also at providing agreements on standards and harmonised requirements to e-Infrastructures
Thank you!

http://www.garr.it

http://www.garr.tv

info@garr.it