Co-ordination & Harmonisation of Advanced e-Infrastructures for Research and Education Data Sharing

Coordination of intercontinental e-infrastructures

Federico Ruggieri (GARR/INFN)

e-IRG Workshop, Rome, Italy
10-11 November 2014

www.chain-project.eu
proj-office@chain-project.eu

Co-funded by the European Commission
under its 7th Framework Programme

Outline

- General Information on CHAIN-REDS
- An overview of regional e-Infrastructures
- The Vision and the Actions
- Use cases
- Conclusions

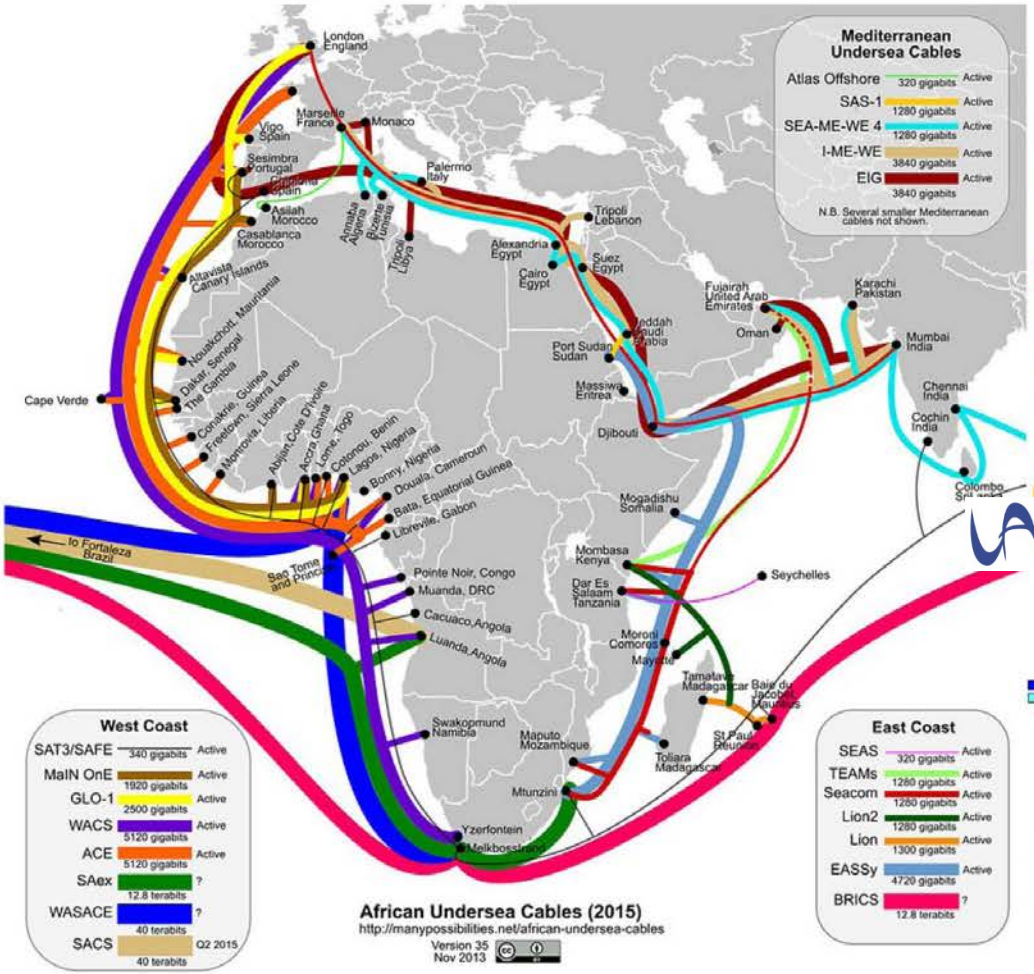
General Information

- Co-ordination & Harmonisation of Advanced eInfrastructures for Research and Education Data Sharing
- Research Infrastructures – Support Action
- Grant Agreement n. 306819
- Total Costs of € 2.3 M
- Max. EC contribution: € 1.52 M
- Start date: 1 December 2012
- Duration: 30 Months

Partners and roles

- INFN (IT) – Coordinator
- CIEMAT (ES) – WP4 Leade
- GRNET (GR) – WP3 Leader
- CESNET (CZ) – WP5 Leader
- UBUNTUNET (MW) – Africa
- CLARA (UR) – Latin America
- IHEP (CN) – China
- ASREN (DE) – Arab States
- SIGMA ORIONIS (FR) – WP2 Leader
- C-DAC (IN) – India

Africa

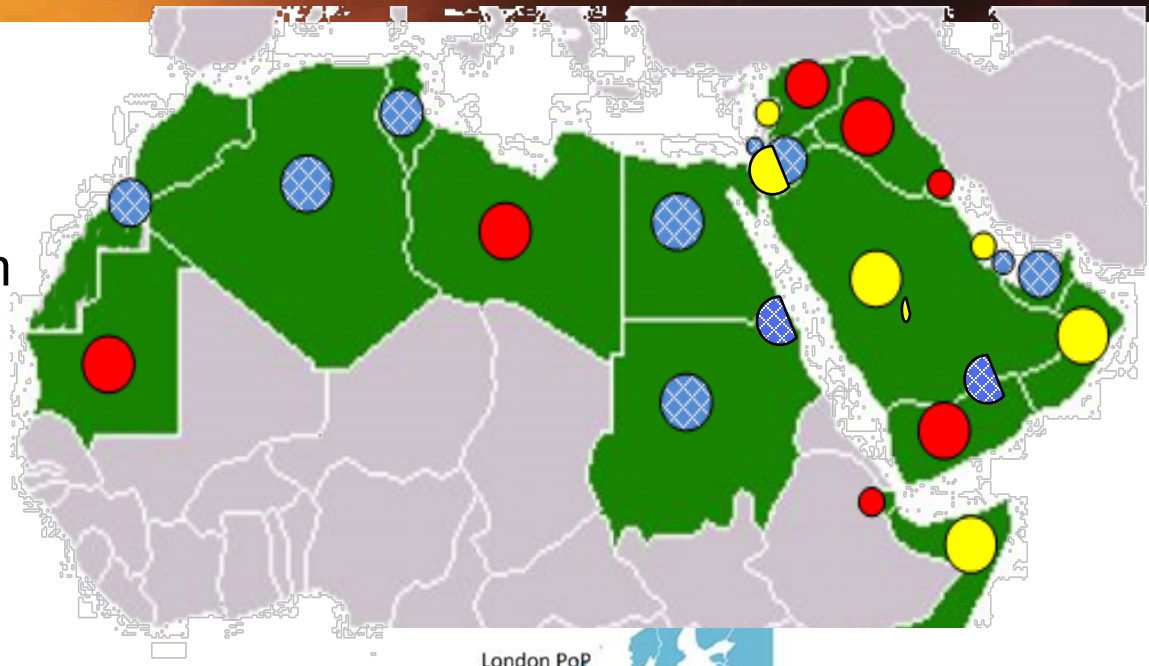


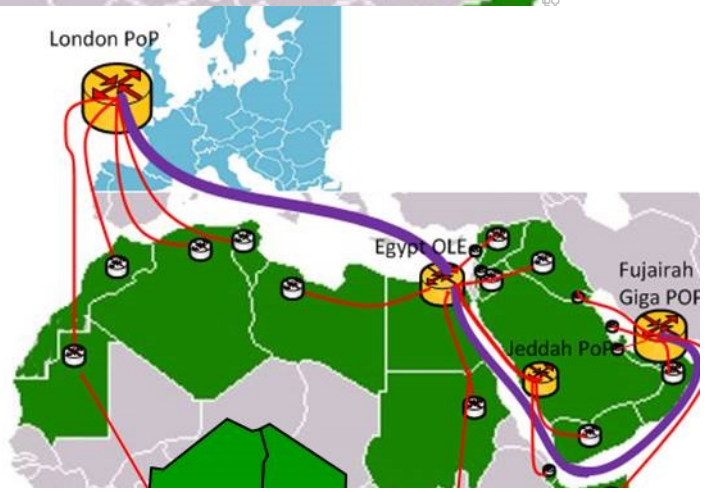
AfricaConnect project improving NREN capacity and cross-border connectivity

http://manypossibilities.net/afterfibre/

Arab States Research & Education Network

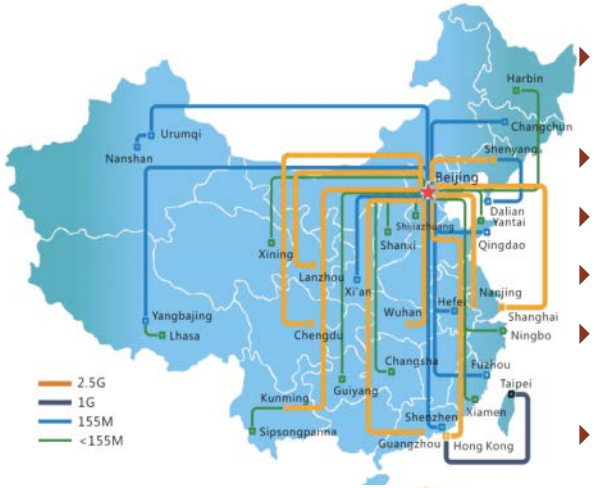
- Arab NRENs: 9 Mature, 5 Developing, 8 Starting
- Logical PoPs: London, Fujairah, Alexandria, Jeddah
- EUMEDGRID legacy continued
- 9 Grid sites (MA, JO, EG, KU, SY, DZ1,2, 3, UAE)
- Africa-Arabia ROC
- 4 HPC sites (EG, SA, UAE, QA)
- Cloud initiatives with Cloud infrastructure towards a widespread Open Source Software





ASREN
Arab States Research and Education Network

China - CSTNet, CERNET and CNGI

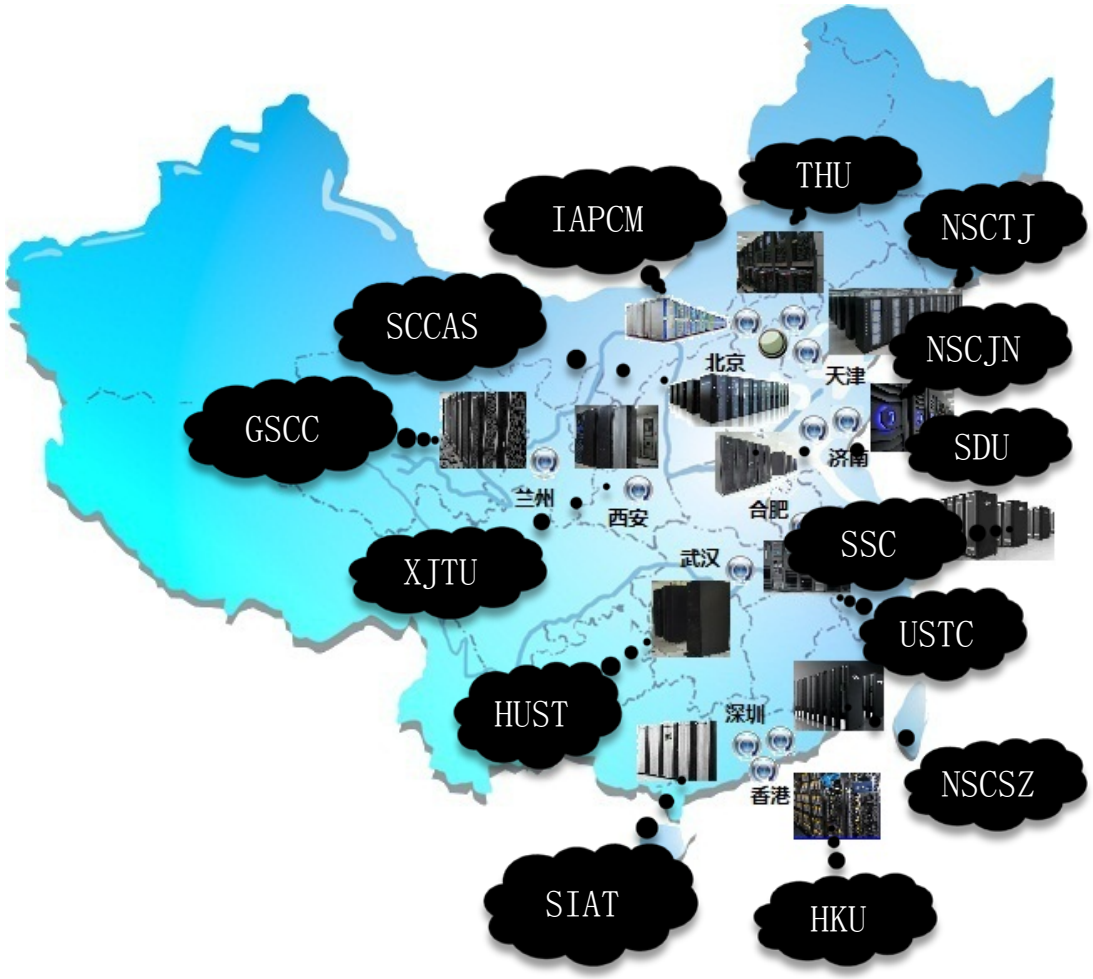


- CSTNet (China Science & Technology Network)
- S&T Network in China
- Backbone 2.5G
- Over 20 provinces
- Over 100 Research Institutes
- Over 1M End Users

- CERNET (China Education & Research NETwork)
- Biggest NREN in China
- Backbone 100G
- 31 provinces and over 200 cities
- 38 PoPs (+25)
- Over 2000 universities and institutes
- 30M (+2M) users

- CNGI: Chinese Next Generation Internet
- CNGI-6IX: International exchange point for CNGI
- Good connections with international and domestic networks

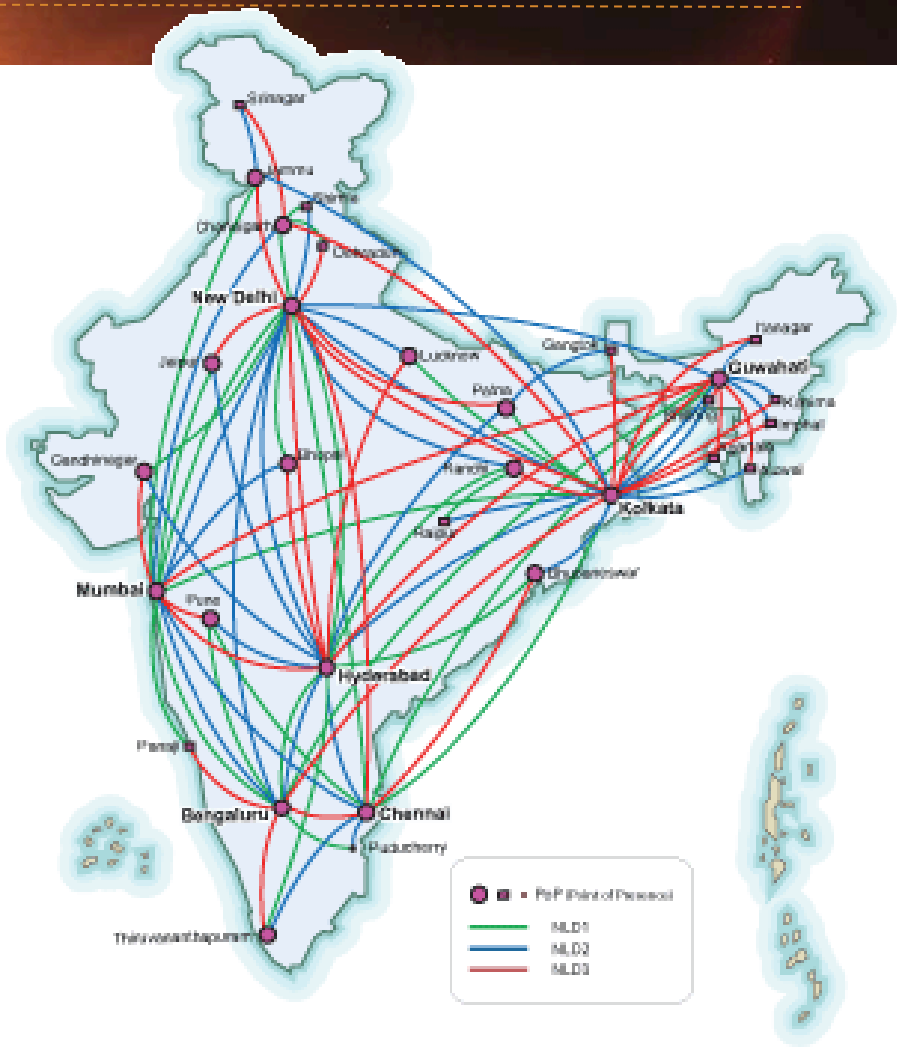
China: CNGrid



- 14 sites
- >3PF aggregated computing power
- >15PB storage
- >450 services
- >2800 users
  - China commercial Aircraft Corp; Bao Steel; automobile; institutes of CAS; universities.
- Supporting >700 projects
  - National Projects (973, 863), and those funded by National Science Foundation of China and Chinese Academy of Sciences

India - National Knowledge Network

- A state-of-the-art multi-gigabit pan-India network for providing a unified high speed network backbone for all knowledge related institutions in the country
- An ultra-high speed CORE (multiples of 10 Gbps), with a distribution layer at speeds of 1 Gbps or higher.

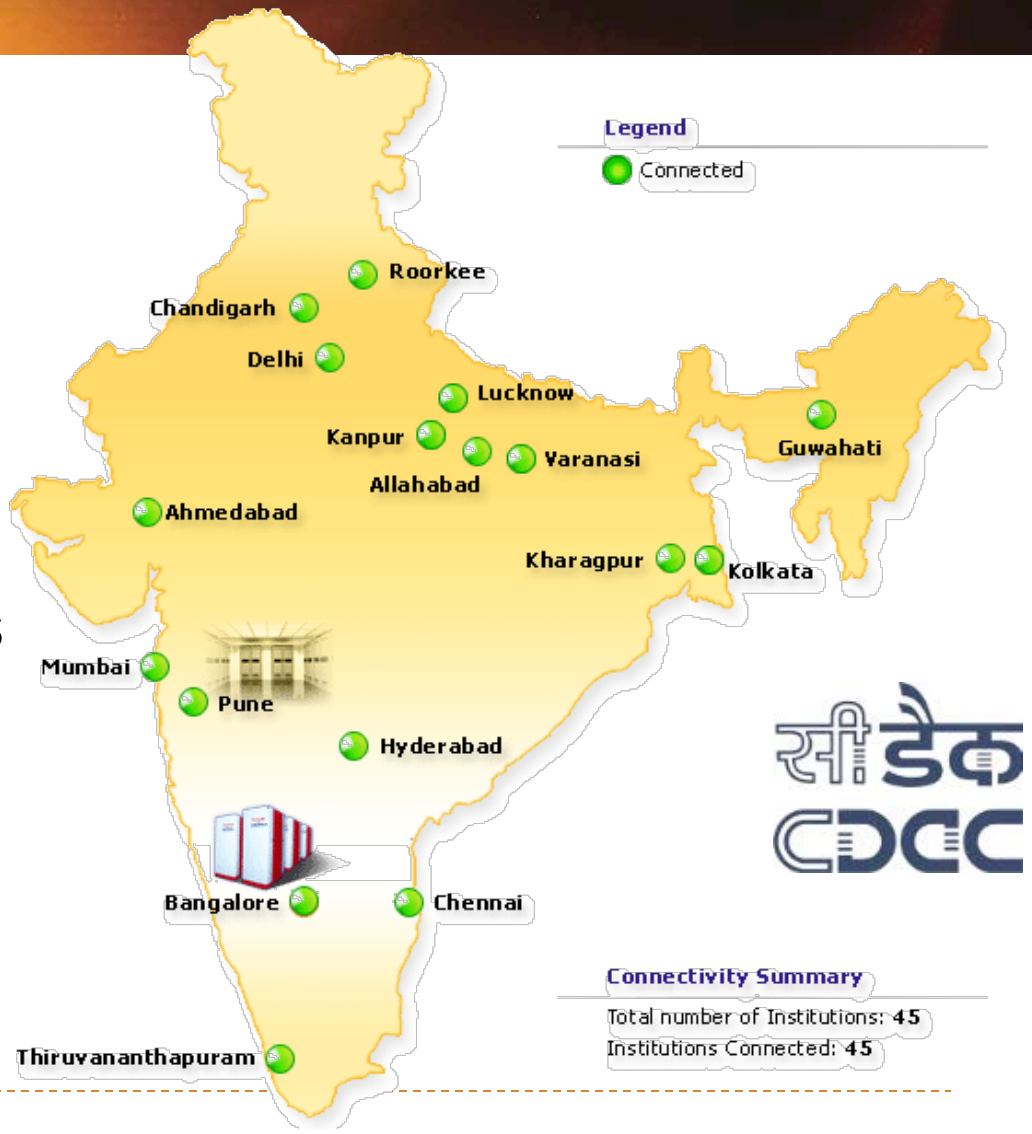


Multiple 10G Connecting all the State Capitals

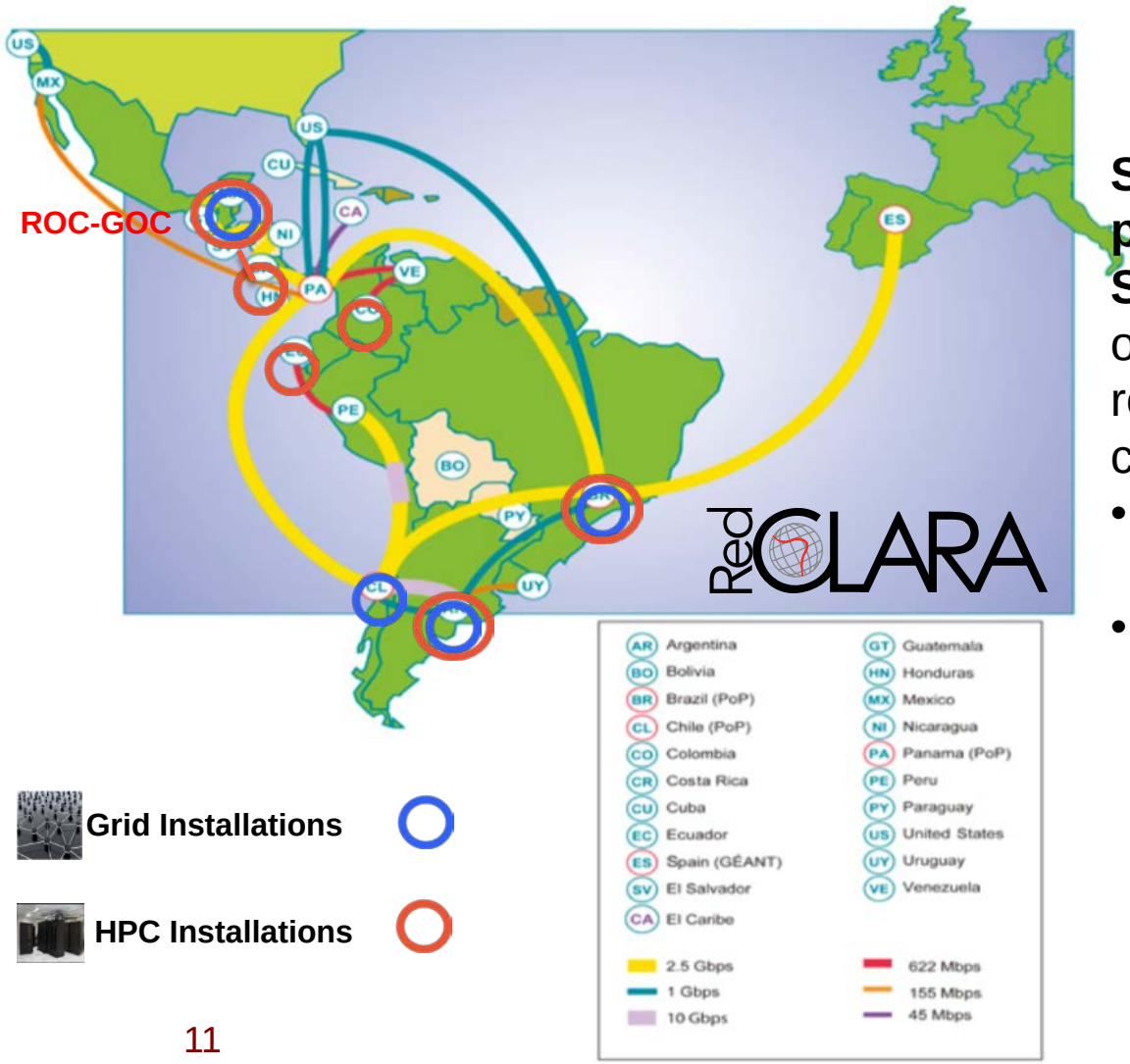
Gigabit Connectivity to all the 640 Districts

India - GARUDA - on NKN

- Global Access to Resources Using Distributes Architecture – National grid initiative by C-DAC
- Aggregates HPC clusters from C-DAC centres and partners
- Currently over 10000 CPUs (~ 550 teraflops) and 220 TB of storage to provide distributed data
- The GARUDA is interconnected by the NKN - a highly reliable and secure national backbone



Latin America



Grid Installations

HPC Installations

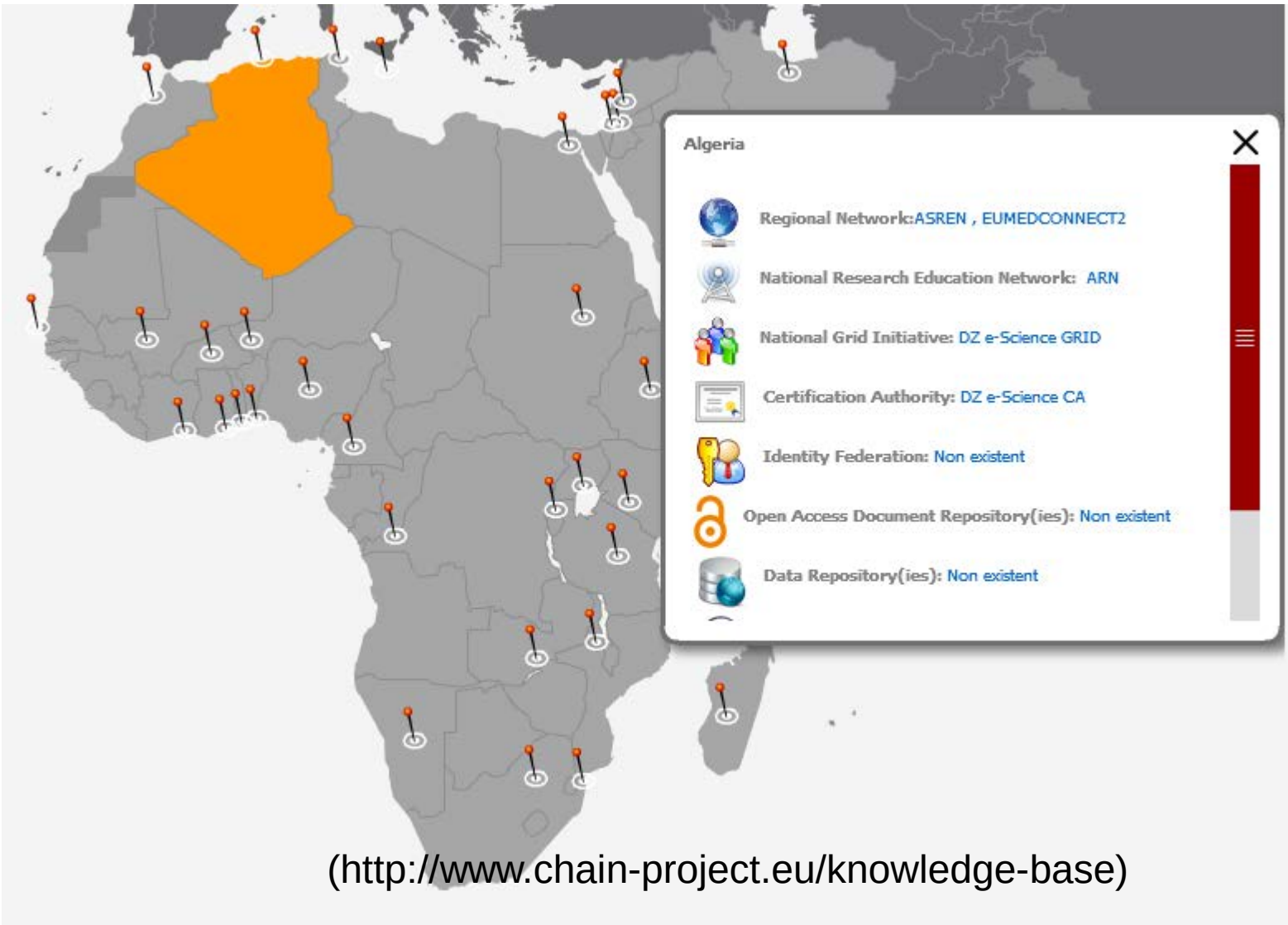
Servicio de Computo Avanzado para America Latina y el Caribe, SCALAC. It is the integration, at organizational level, of LA regional resources to provide advanced computational services. We have:

- **ROC-LA** supports **Grid** Services (HEP + Astroparticle communities)
- **Network HPC** Centres in LA
  - SINAPAD (Brazil)
  - SNCAD (Argentina)
  - CEDIA (Ecuador),
  - CENAT (Costa Rica),
  - UNAM (Mexico),
  - SC3/GridColombia (Colombia),
  - CeCalCULA (Venezuela)

The Vision

- Promote and support technological and **scientific collaboration across different e-Infrastructures established and operated in various continents** to facilitate their uptake and use by established and emerging Virtual Research Communities (VRCs) but also by single researchers
- Not only disseminate, exchange and reinforce the best practices currently adopted in Europe and other continents, but also **promote the progress of interoperability among different regional e-Infrastructures**
- Study and **define a path towards a global e-Infrastructure ecosystem** that will allow VRCs, research groups and even single researchers to access and efficiently use worldwide distributed resources

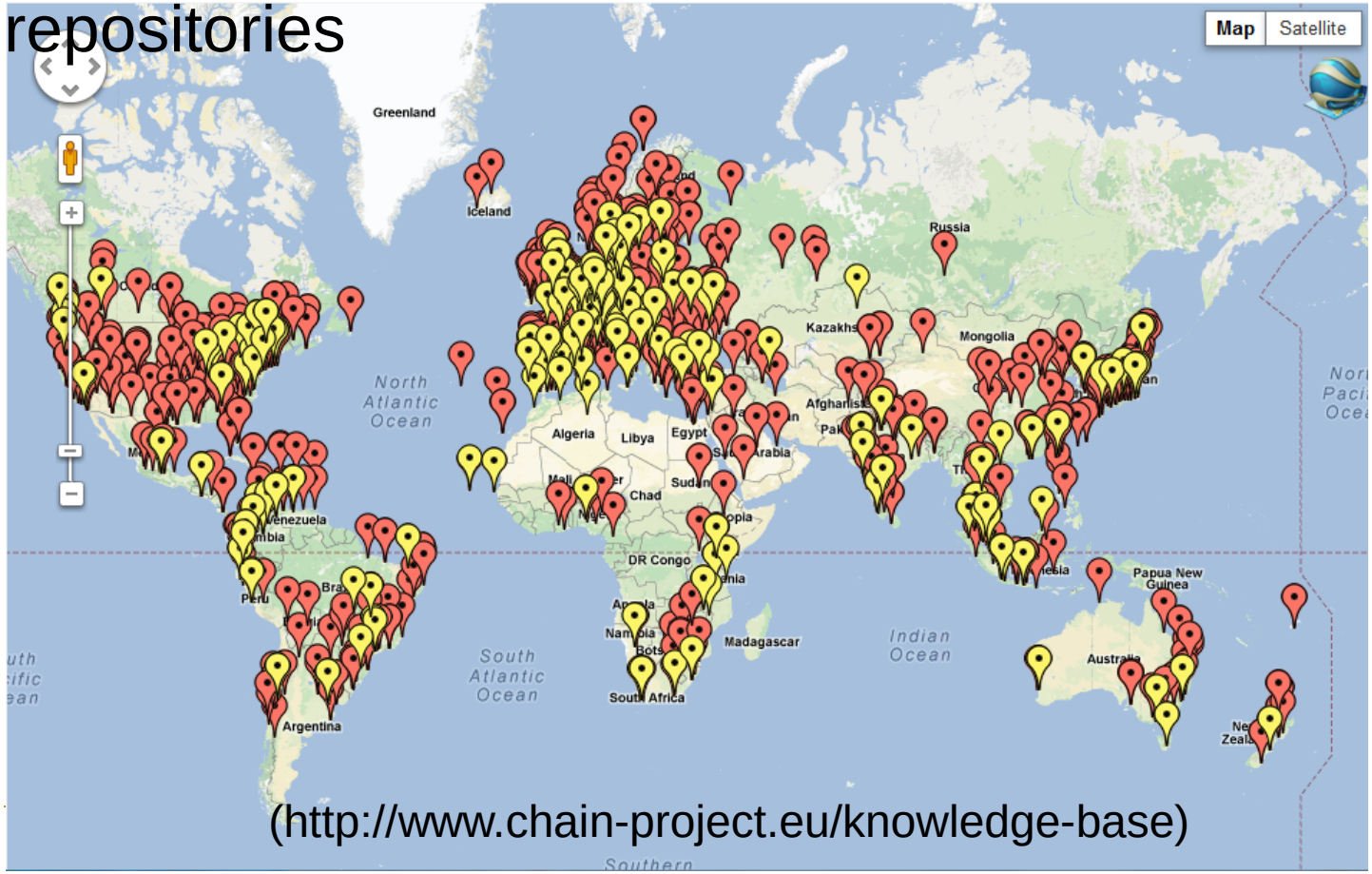
The Knowledge Base



(http://www.chain-project.eu/knowledge-base)

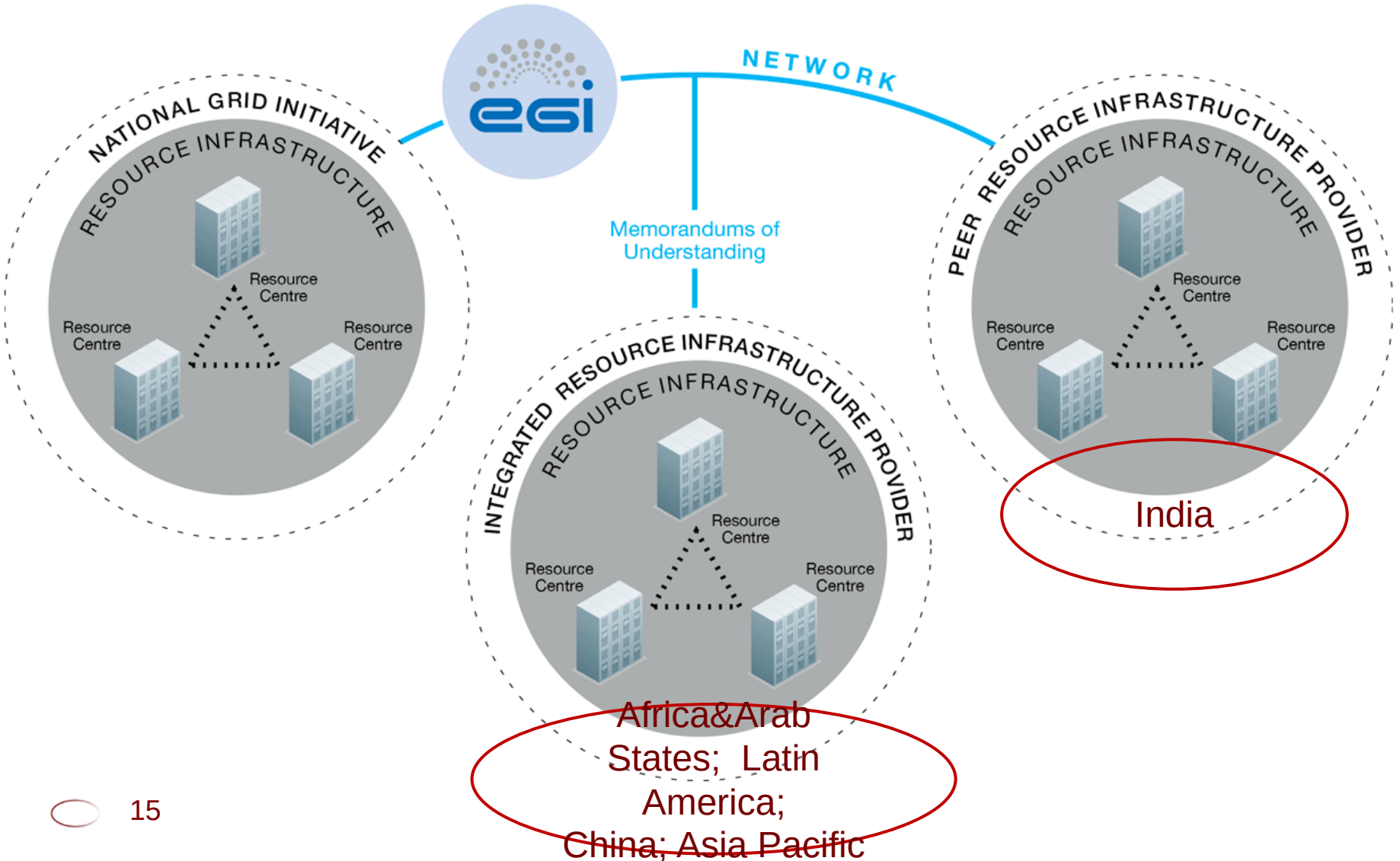
The KB of Open Access Repositories

- The Knowledge Base of e-Infrastructures has been extended to open access data document and (recently) educational repositories

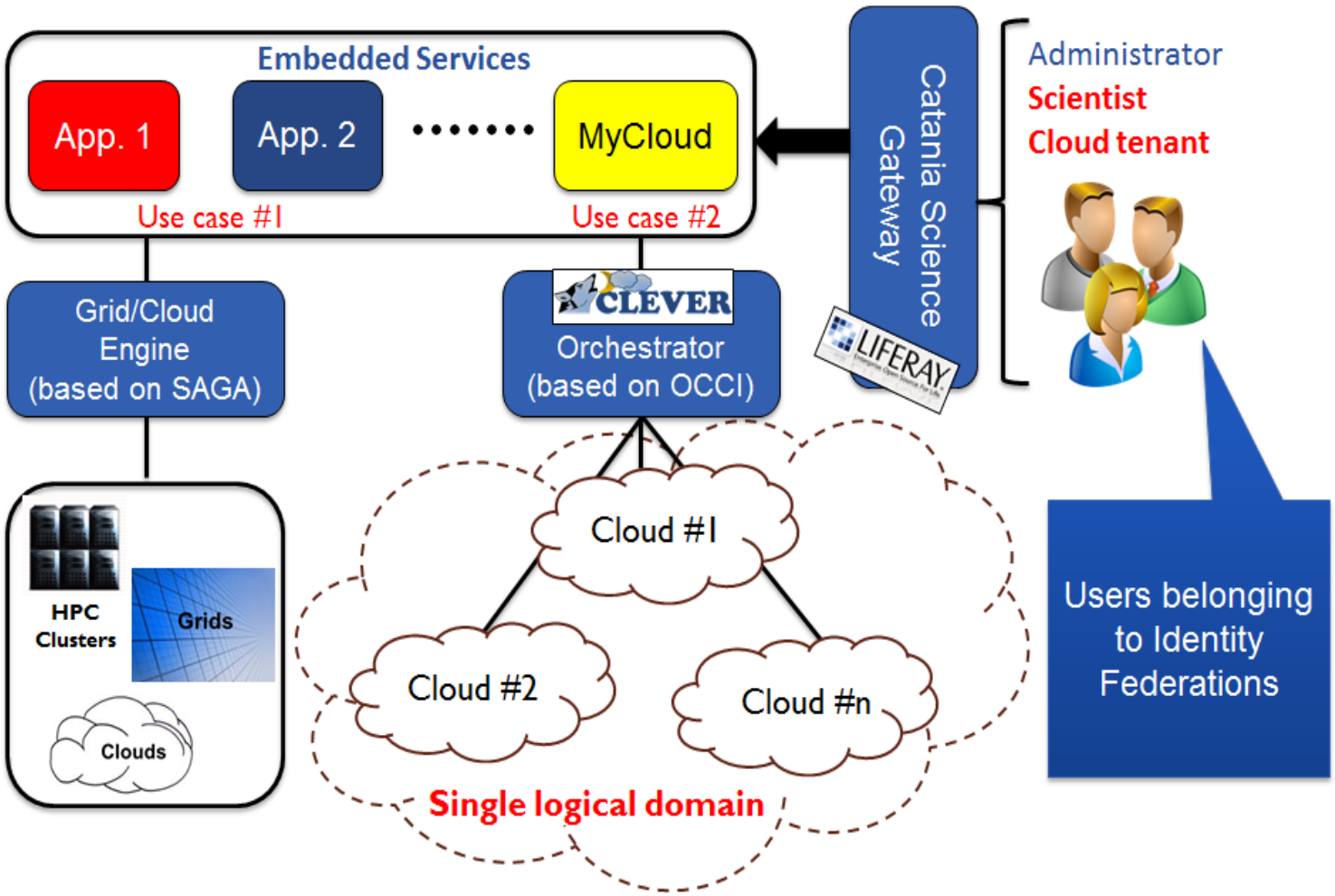


(http://www.chain-project.eu/knowledge-base)

Interoperations

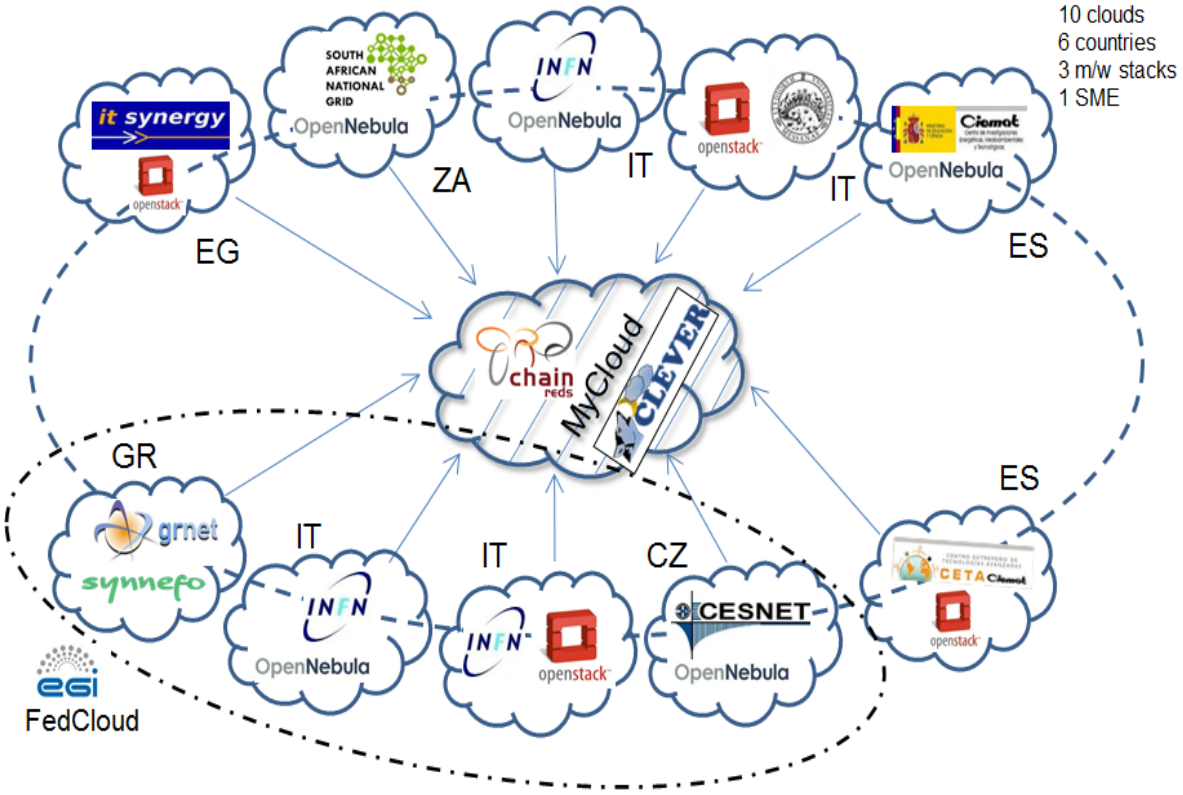


Science Gateway and Standards: SAGA, OCCI, CDMI

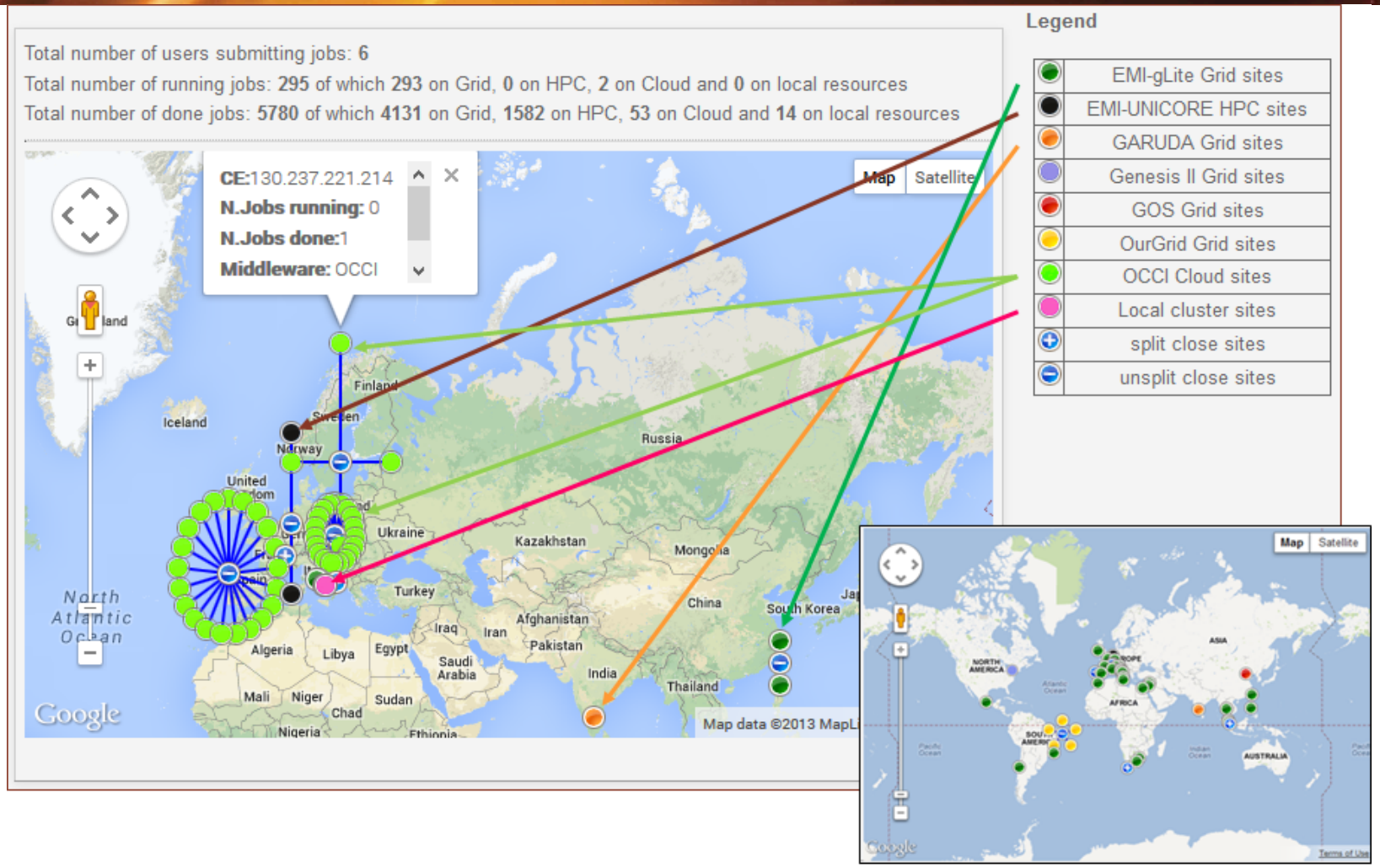


Clouds for R&E - Federation demo

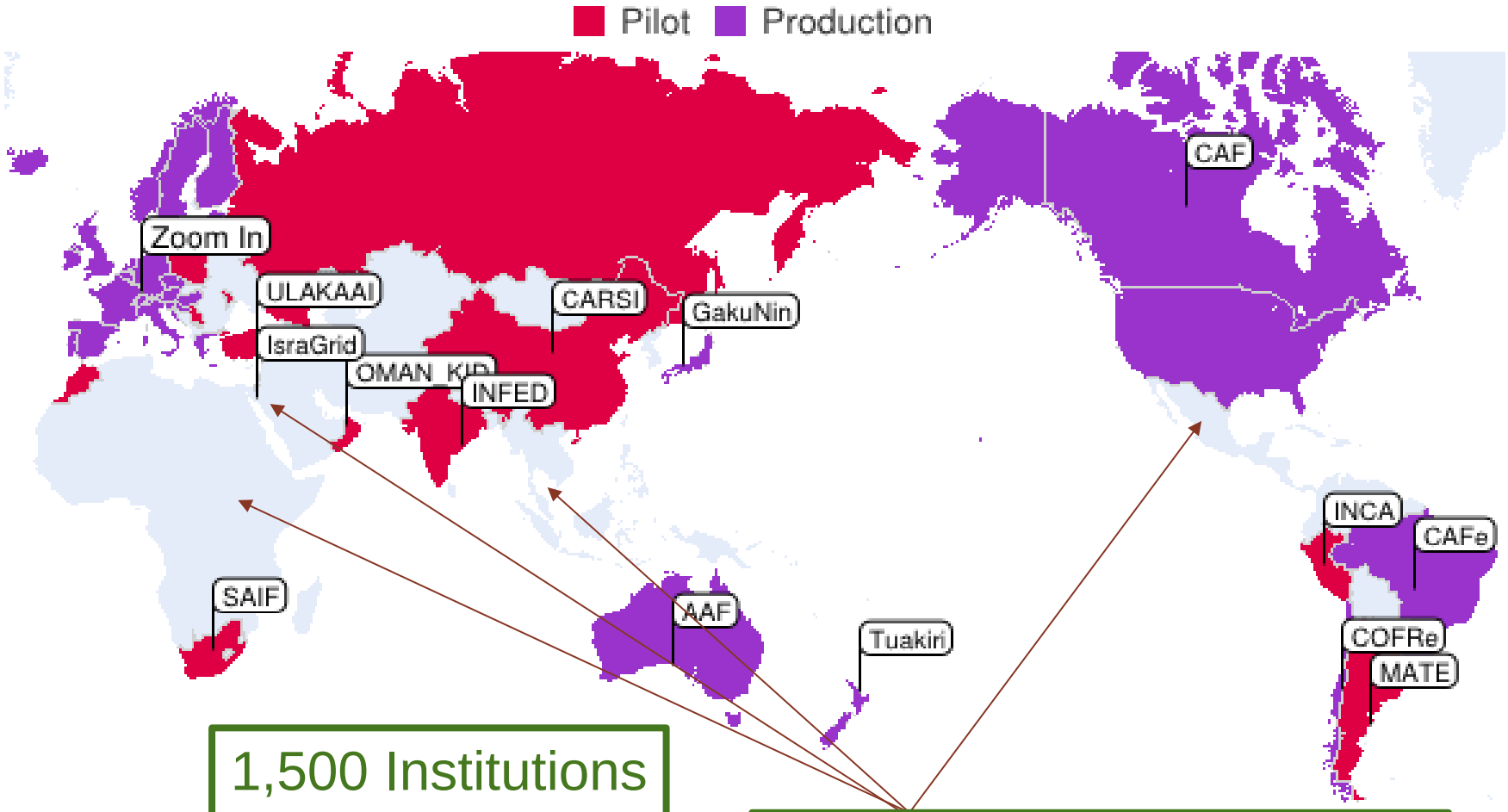
- Baseline example for cloud federation: EGI Federated Cloud Task Force
  - OCCI and CDMI standards as a prerequisite for cloud federation
  - Orchestration system, based on CLEVER, linking several Cloud providers
- Cloud federation demo
  - As an application accessible by the CHAIN-REDS Science Gateway



Clouds for R&E - Federation demo



Federated Identities



1,500 Institutions
2,100 Services
>17 million people

A lot of work still to be done

Federated Identities – The CHAIN-REDS impact

Countries where CHAIN-REDS has supported the deployment of federated services

Use cases

- APHRC (Africa)
  - Societal health and well-being domain
- LAGO (LA)
  - Astroparticle domain
  - DART based
- GROMACS (India)
  - Molecular dynamics domain
- ABINIT (Arab region)
  - *Ab initio* calculations on quantum chemistry and Physics of materials domains
- TreeThreader (China)
  - Life processes at the molecular level

The African Population Health Research Centre – www.aphrc.org

- "APHRC is a ***leading pan-African research institution*** headquartered in Nairobi, Kenya, that conducts high quality ***policy-relevant*** research on **population, health, education, urbanization** and related development issues across Africa."

- "Over the last decade, APHRC has contributed to a better Africa through quality research, and ***results-based strategic communications and policy engagement.***"

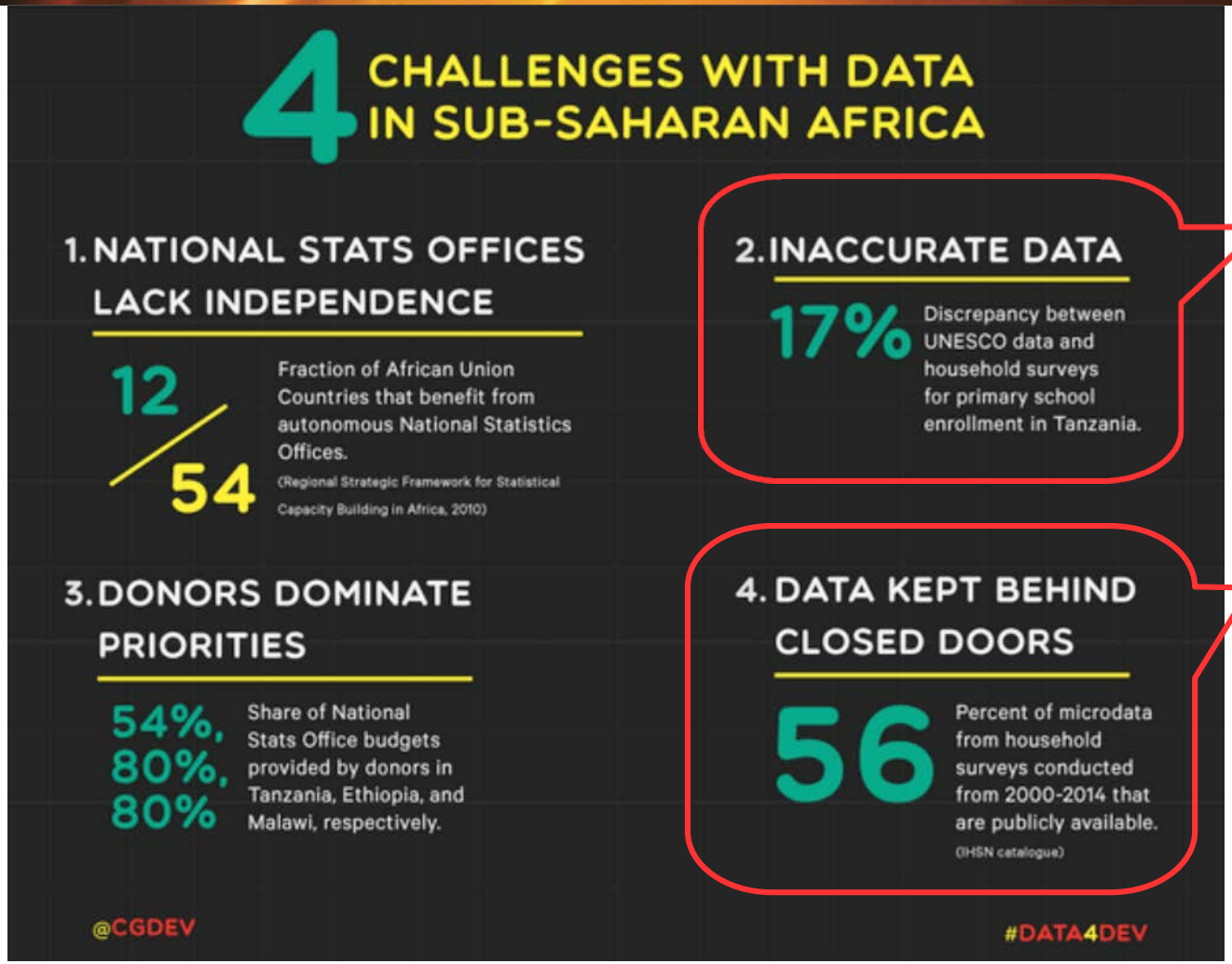
APHRC and Open Data in Africa

- APHRC mandate: better societies and government through the use of quality data

- Donor and mission require that data is open, publicly available and accountable

- *#Data4Dev* is a hot topic:
  - African needs a data revolution; http://aphrc.org/blog/you-say-you-want-a-data-revolution
  - And there is a plan: http://www.cgdev.org/blog/delivering-data-revolution-sub-saharan-africa

#Data4Dev – challenges in Sub-Saharan Africa



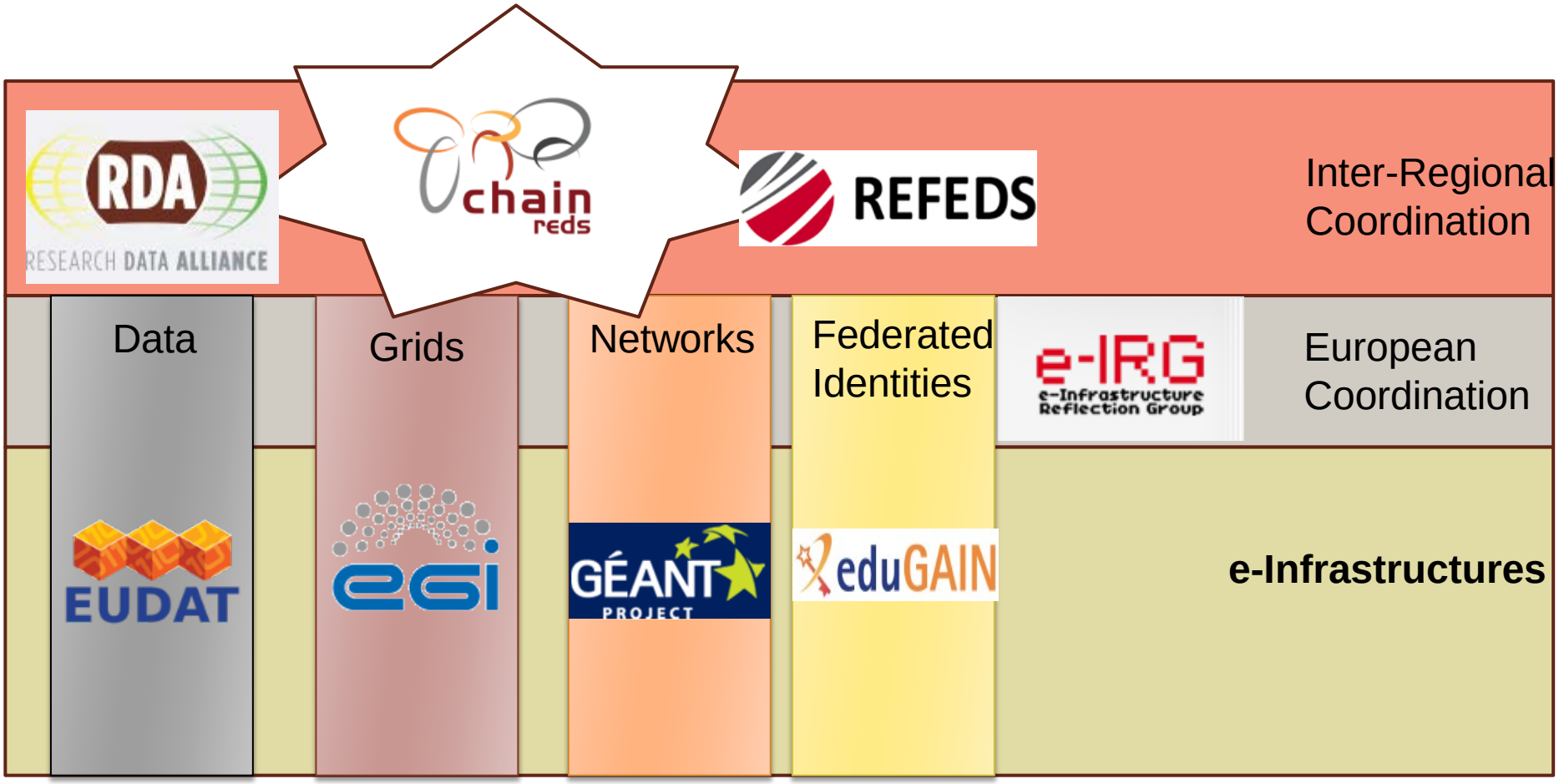
Reproducibility

Accessibility

What CHAIN-REDS is doing for APHRC ?

- **Support provided:**
  - **Consultancy** on the benefit of issuing persistent identifiers
  - **Access** to EPIC **PID** issuer
  - **Reference** implementation of **PID** issuer for existing data sets
- **Activities:**
  - **MoU** signed with APHRC: “Promotion of **standards** to **widen** and **simplify** the access to e-Infrastructures worldwide and federation of Big Data archives across continents”
  - **Standards and interoperability**: Ongoing technical investigation to determine feasibility of integrating IHSN with semantic search engine
  - **Availability** (APHRC already have a microdata portal)
    - http://aphrc.org/catalog/microdata/index.php/catalog
  - **Discoverability and impact:** Included in CHAIN-REDS Knowledge Base, issued PIDs to data sets

Coordination & Harmonisation



Conclusions

- CHAIN-REDS Vision has been implemented in a series of actions to promote coordinated activities among regional e-Infrastructures
- Interoperation, Science Gateways, Identity Federations and standards are promoted to support intercontinental VRCs activities
- Successful demos have been deployed for Federated Clouds and DART
- Five Use Cases are supported to show the successful implementation of the project's vision
- The project is currently filling the gap of stable Inter-regional Coordination: a Plan will be proposed in the final part of the activity
- Are peer-to-peer MoUs between regional e-Infrastructures enough ?
- Contributions from e-IRG are needed to propose a realistic and stable Inter-regional coordination activity for the future

Co-ordination & Harmonisation of Advanced e-Infrastructures for Research and Education Data Sharing

Thank you !

www.chain-project.eu

proj-office@chain-project.eu

Co-funded by the European Commission
under its 7th Framework Programme

e-infrastructure