ICSC and Terabit

Federated Infrastructure and AAI

ICSC
Centro Nazionale di Ricerca in HPC,
Big Data and Quantum Computing

terabit

ICSC unified infrastructure needs

- Support to **High Performance Computing (HPC)** and **High Throughput Computing (HTC)** batch workflows
- Capability to **store** and **elaborate** data providing continuum of HPC and **Big Data** solutions for the ICSC users
- Provide federated access to **Interactive, Storage** and **Archival Services**
- Support to **Machine Learning (ML)** and **Deep Learning (DL)** applications
- **High Level Abstraction** view separating architecture and HW details
- **Autonomy of scientific communities** exploiting technologies and tools so as to implement **workflows** and analyse data in a ***Distributed HPC Cloud model***

Towards a unified infrastructure both at national and international level

- Increase **user experience** in accessing distributed resources through **federated solutions**, at the same time to simplify the way to instantiate and use distributed resources.
- Support the availability of **heterogeneous** resources, which should include e.g. **CPUs**, **GPUs** and **storage**, also providing different service levels
- Flexible service development across multiple sites and multiple Clouds
- **Interoperability** with other infrastructures and solutions. These include for example other **key National Initiatives** such as **TeRABIT**, which aims to integrate and improve the 3 major digital infrastructure in Italy (**GARR-T**,**PRACE-Italy**,**HPC-BD-AI**)

Main available services

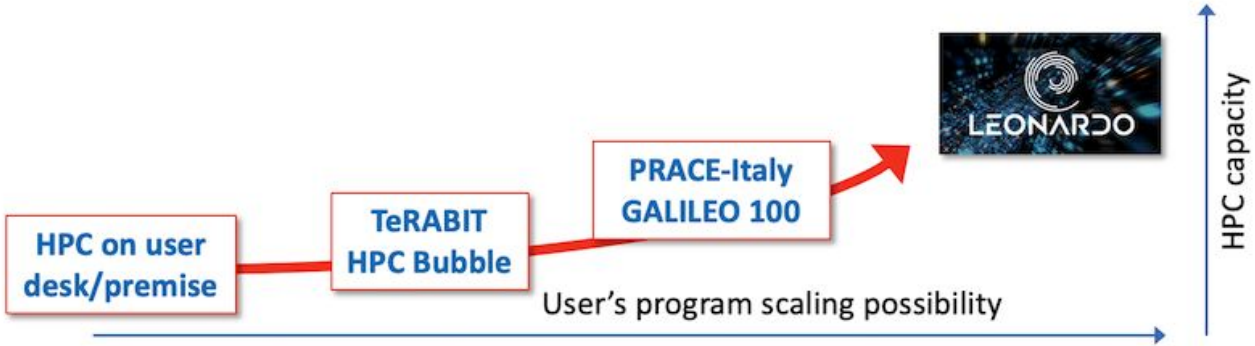
- **Scale-out**
  - Used to perform Simulation, Data processing, etc. (**HPC/HTC**)
- **Interactive and Cloud Computing**
  - **Post-processing**, **Visualization**, **Data Analytics**, **Deep Learning**, etc.
  - **Openstack** Cloud service is included (available both at **CINECA** and **INFN**)
  - **Interactive Computing** based on technologies as: SluRM, UNICORE, JupyterHub at CINECA, HTCondor and WLCG, INDIGO-DataCloud at INFN
- **Active Data Repository**
  - **Massive I/O**
  - Parallel File systems (es. **GPFS**)
- **Archive Data Repository**
  - **Long term** archiving
  - **Object storage** (**S3 DDN / SWIFT**)

Advanced Enabling services

- **AAI and federation**
  - HPC/HTC sites federation (CINECA, INFN, etc..)
  - CINECA **IdP** is integrated in the **FENIX AAI**
- **Data management**
  - **FTS3** Data Transfer
- **Operations**
- **Production services**
  - Tools and Framework
  - Virtual Machines
  - **PaaS**

Hardware infrastructure (INFN/CINECA/GARR/…)

- Available HW resources
  - HPC/HTC
  - Scale out + interactive
  - Storage
  - Networking



ICSC and Terabit

Authentication and Authorization Infrastructure

AAI Main goals

- Authorizing the access to services and resources of each provider (e.g. **CINECA**, **INFN** and **TeRABIT**) in a way to possibly **extend** the federation in the future including more sites
- Provide **seamless access** to services and resources in the federation
- Users previously registered on local sites are authorized to access resources **without registering twice**
- A user, registered to a federated IdP, could be authorized to consume resources and access them providing same credentials, if he/she has been previously assigned to a project and provided with **credits** or **budget**

AAI: Leading Principles

- Authorize users to **different classes of services** and **resources** made available by the centres from the **local site IdPs**
- **Independence of each site** and each provider by the federation
- Already registered users are considered members of the federation by default
- Multiple accounts must be associated to a **single unique identity**
- Sites are free to choose their own **policy**, even if they should comply with already existing common standard protocols and technologies (**SAML**, **OIDC**, etc..)
- Each center will maintain own version of the **registration procedure**
- ICSC will support federated authentication and authorization policies, through the adoption of industry-grade, **open standard solutions**

The AAI today

- CINECA provides user authentication through **Keycloak** IdP (sso.hpc.cineca.it) for CINECA users to local sites services (e.g. Openstack)
- CINECA is also part of **FENIX**, a federated e-infrastructure made of the major 5 Supercomputing centers in Europe (CINECA,JSC,CSCS,CEA,BSC) and more recently CSC. CINECA IdP is federated with the **FENIX Proxy** (**Géant**)
- The INFN **DataCloud** components use the OpenID Connect and OAuth protocols through **INDIGO-IAM**, available both as a general IdP Proxy for the entire infrastructure, and as an “AAI-as-a-Service” solution that can be self-instantiated through the INFN DataCloud Service Catalogue.

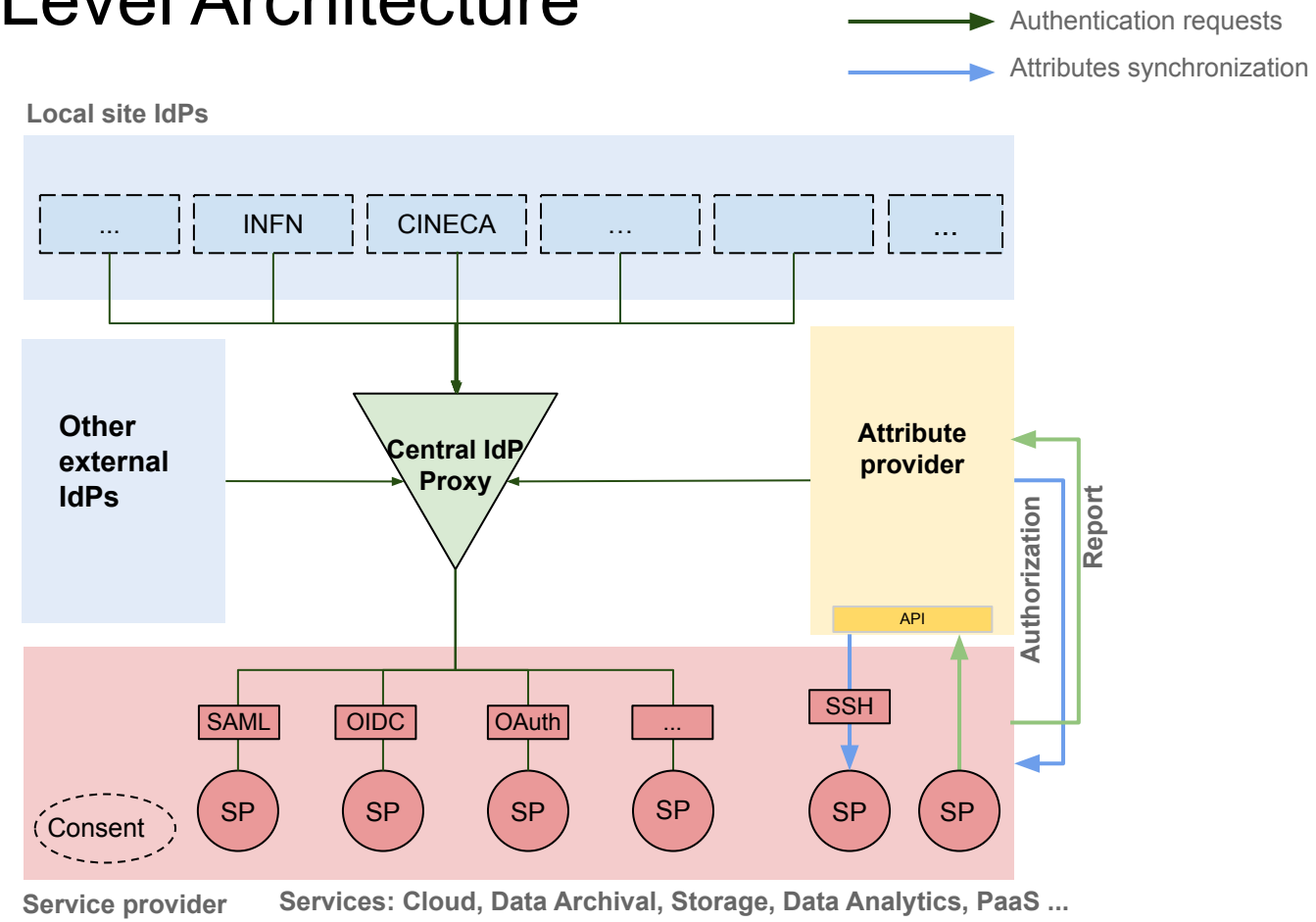
ICEI/Fenix

- Federation of the 5 major computing centers: CINECA, INFN, CSCS, BSC, CEA
- Federated access to services provided by HPC sites (Cloud, Interactive Computing, Storage, ...)
- Identity federation and services standardization
- Centralized solution for managing authorization of users (**FURMS**)
- Technologies: SATOSA (Geant Proxy) and Keycloak (site IdPs)
- Now migrating to **MyAccessId**, a new IdM layer that will provide more flexibility and simplify contractual obligations among partners

AAI: Proposal

- In order to integrate existing services and to create a new federated infrastructure we propose to identify two main components:
  - A **Central Proxy service** which simply redirects users authentication requests and provide a web page for General Policy acceptance
  - An **Attribute Provider** as central attribute source (for the authorization aspects)
- We will adopt standard protocols like e.g. **OpenID Connect (OIDC) or SAML**
- The user profile will be validated and all the required attributes must be provided
- Sites will provide services to users and communities independently by the Central Proxy Idp
- Authorization to resources will be managed according to local policies

AAI: High Level Architecture



AAI: Main components

- **Central Proxy IdP**
  - Redirect authentication requests
  - Identify and authenticate users (also with no affiliation)
  - User profile validation
  - General Access Policy management
- **Attribute Provider**
  - Manages projects budget and groups for the users
  - Manage eventually local site policies
  - Reports and statistics

AAI: User identity

For each user accessing the infrastructure the first time will be created a new profile on the Central Proxy IdP with a minimal set of attributes associated, these include:

- A unique **opaque identifier**
- A **username** associated to the user

For each user a new profile is created, in this way for example:

1. The user is able to accept the ICSC general policy
2. His/her activities are monitored
3. His/her identity can be associated to a project
4. A **LoA** can be associated to him/her

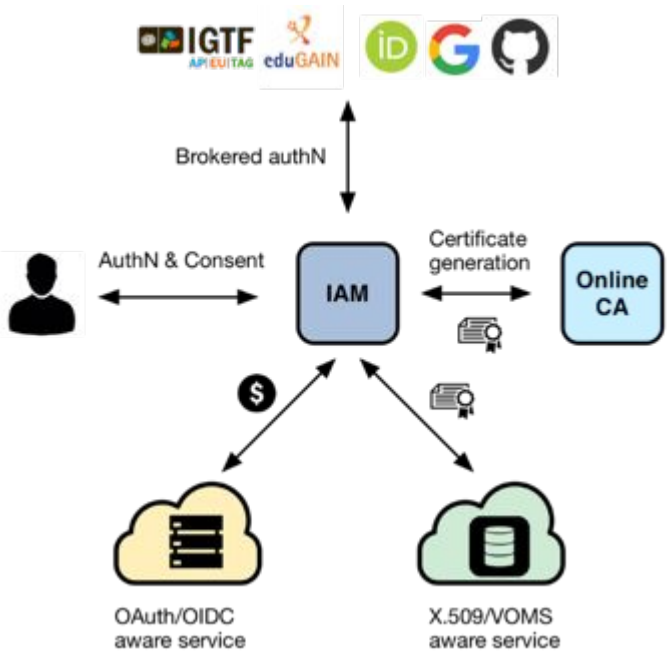
INDIGO Identity and Access Management service

Indigo IAM can manage authentication in a flexible way, including account linking…

Easy the integration of different services adopting standard protocols (OAuth, OpenIDConnect, …)

Manage also authorization?

The Indigo IAM could be a candidate to implement a Centralized Proxy solution



AAI: How to manage credits and budget

- Possible solutions:
  - Resources made available as **credits** for the users
  - Some **evaluation process** must be defined in order to associate credits to the users
  - Manage the authorization attributes as well as credits and budget in a **centralized** way
    - In this case we need to decide how and where to implement this functionality

AAI: Project management proposal

- In order to access resources and services the user must be associated to a project by the PI as a collaborator (storage quote and compute)
- Project credits and budget must be defined
- The project list must be added to the authorization service by a technical committee and approved

Work plan

- **Phase 1 (months 1-6) PoC TeRABIT**
  - Each site must provide an IdP to federate
  - First **stepwise** integration between the two major core centres (CINECA and INFN)
    - CINECA provides **Keycloak** IdP endpoint, e.g. to initiate federation with **Indigo IAM**
  - Provide federated access to users of both sites
  - All users are provided with **budget** e.g. from the local sites web portal
- **Phase 2**
  - Implement a central Authorization service to manage budget and credits to propagate on sites (could be always Indigo IAM or a different one)
  - Federate other resources and services as part of the infrastructure
  - We finalize all the configuration including managing **LoA**, **account linking, token introspection** and so on…

People involved: CINECA, INFN, GARR, …

Questions?