A VOSpace deployment: interoperability and integration in big infrastructure

S. Bertocco

Scope

Provide our users with a local data storage and computation infrastructure compliant with a world wide Virtual Observatory vision

- astronomical standards based

  to be widely accessible especially with astronomical tools

- interoperable with other similar storage services

  to increase the accessible sets of data

- integrated *with other* big software infrastructures

CANFAR Infrastructure

CANFAR, the Canadian Advanced Network for Astronomical Research combines:

- the Canadian national research network (CANARIE),
- cloud processing and storage resources (Compute Canada)
- an astronomy data center (Canadian Astronomy Data Center – CADC)
  - hosting a very large data set
  - specialized in data mining, data processing, data distribution and data transferring
  - providing a lot of sophisticated tools to support and enhance the research efforts of Canadian and international astronomers

CANFAR Infrastructure

CANFAR

www.canfar.net/en/resources/services/

Apps | JavaScript Tutoria | New Tab | JavaScript & jQue | Federated AAI Co | Catania: da una fir | batchata dance w | Astrochannel 2.0

CANFAR

Nodes | Resources | Documentation | About

Expertise
Services
Open Source

Services

Improve this page

Get started with CANFAR!

OpenStack Cloud
Run your own virtual machines on Compute Canada cloud

Default is 10 VMs sharing 20 CPUs, 50GB RAM, 1TB disk and 1 public IP

User Documentation
Go to service portal

Digital Object Identifiers
Set a Digital Object Identifier for your data

User Documentation
Go to service portal

Batch Processing
Access large resources for batch processing on the cloud

Up to 16 CPUs, 120GB RAM per VM and up to 2,000 VMs

User Documentation
Reference API
Go to service portal

Storage
Manage your own large storage for astronomy data

Default is 500GB, can accomodate up to 100TB per project

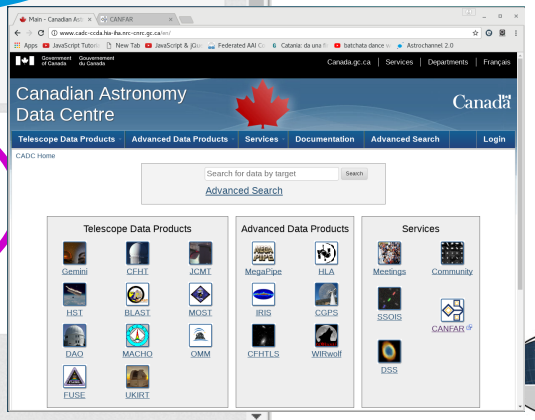
User Documentation
Reference API
Go to service portal

CADC Data Collections
CADC Data Discovery and Access

User Documentation
Reference API
Go to service portal

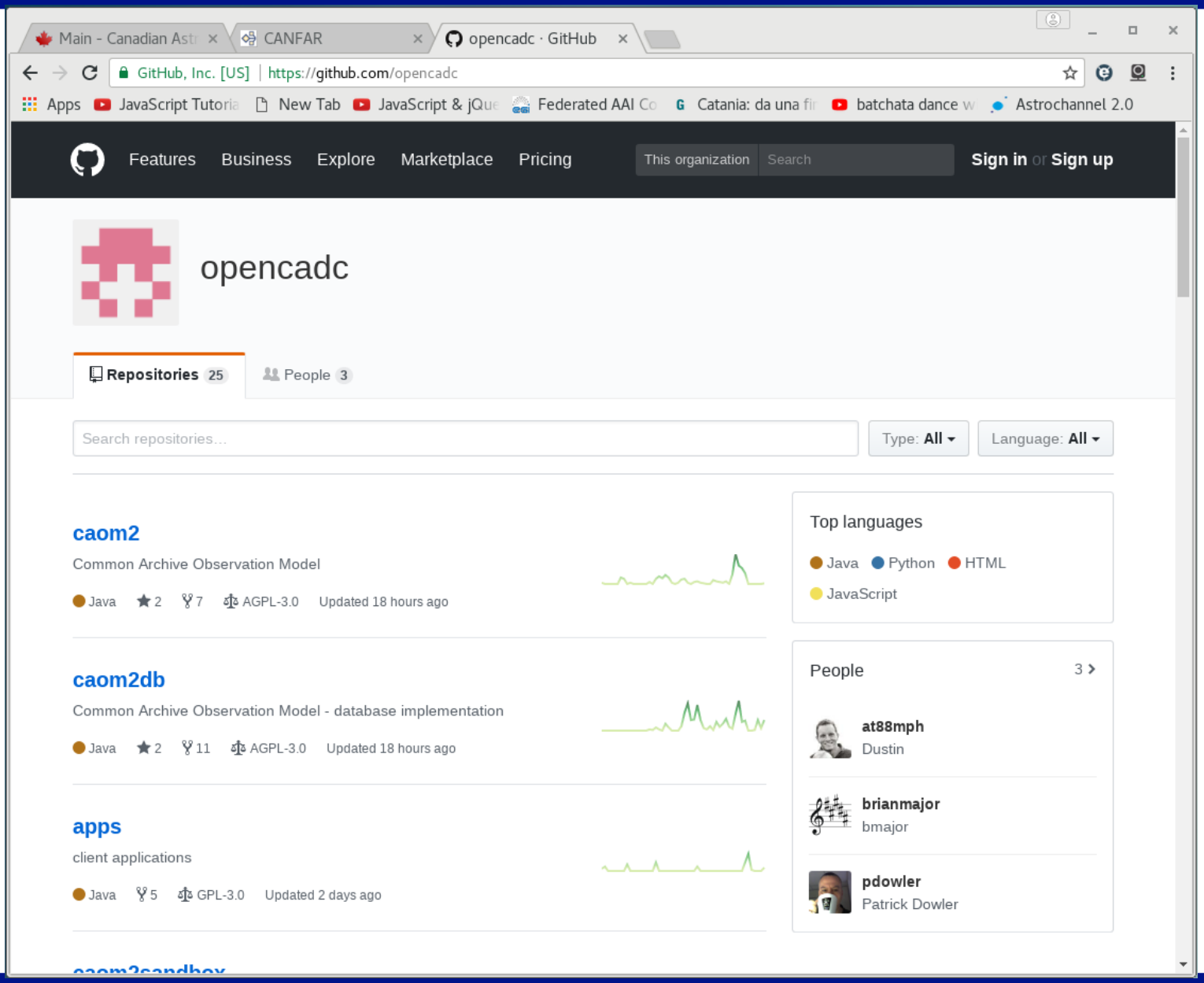
Group Management
Manage access permission to your data or data located with the Storage service

User Documentation
Reference API
Go to service portal



IVOA

OpenCADC Repository



IVOA Standards

➡ https://github.com/opencadc **CADC**

Modules:

- vos VOSpace standard implementation
- ac Access Control (including GMS)
- cdp Credential Delegation Protocol implementation
- reg Registry Interface implementation (including VOSI)
- uws Universal Worker Service Pattern implementation
- core core utilities and logging

➡ IVOA Standards and recommendations based
(http://ivoa.net/)

Storage and VOSpace IVOA recommendation

VOSpace recommendation:

"VOSpace is the IVOA interface to distributed storage. It specifies how VO agents and applications can use network attached data stores to persist and exchange data in a standard way.

A VOSpace web service is an access point for a distributed storage network. Through this access point, a client can:

- add or delete data objects in a tree data structure
- manipulate metadata for the data objects
- obtain URIs through which the content of the data objects can be accessed

VOSpace does not define how the data is stored or transferred, only the control messages to gain access. Thus, the VOSpace interface can readily be added to an existing storage system.

When we speak of "a VOSpace", we mean the arrangement of data accessible through one particular VOSpace service."

Storage Service

VOSpace:

- VOSpace front-end

https://github.com/opencadc/vos

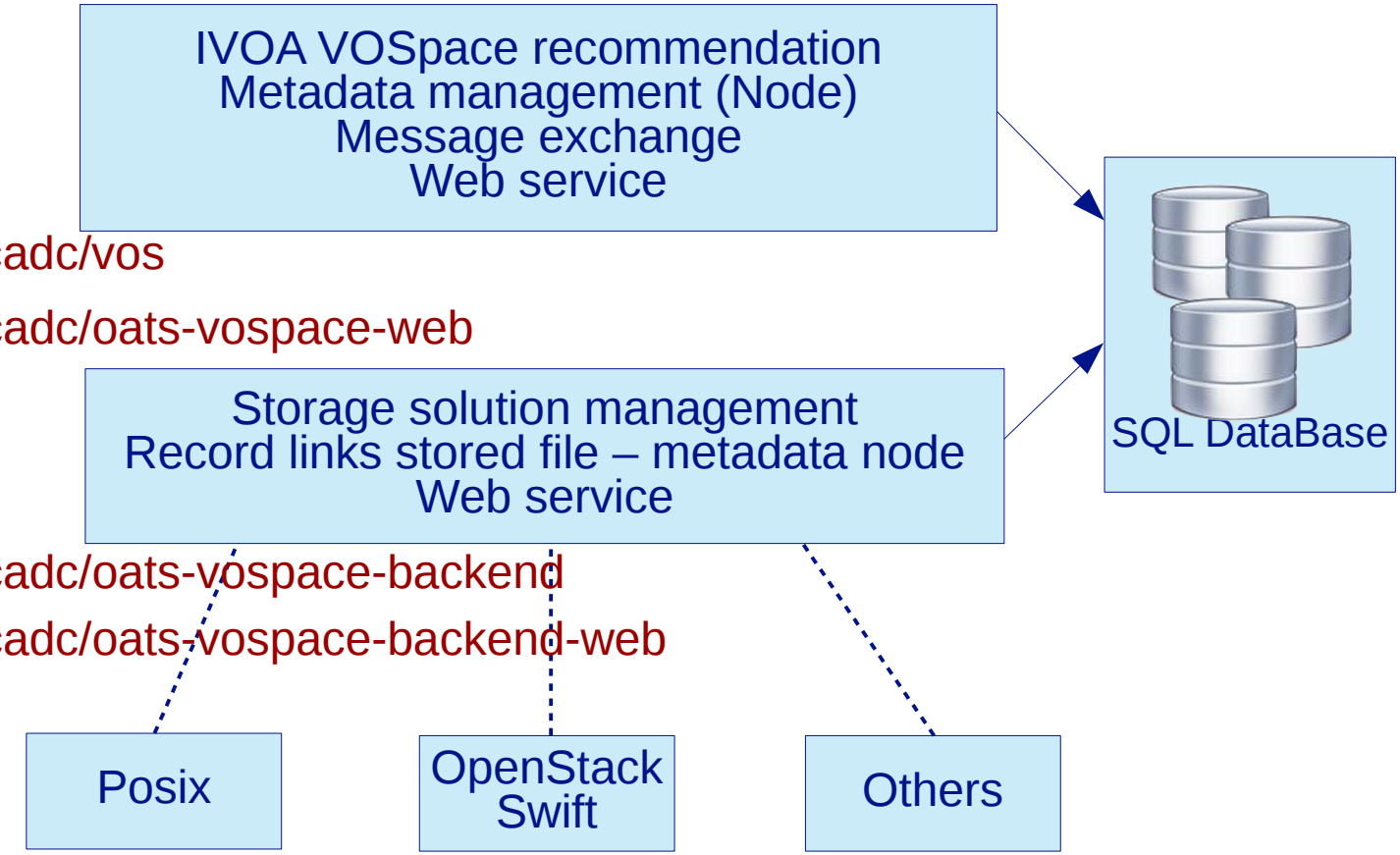
https://github.com/oats-cadc/oats-vospace-web

- VOSpace back-end

https://github.com/oats-cadc/oats-vospace-backend

https://github.com/oats-cadc/oats-vospace-backend-web

- Physical storage



https://github.com/oats-cadc/oats-vospace-backend-developers-guide

Access Control

Access permissions stored in VOSpace database

VOSpace access policy based on group membership

Access Control Service (https://github.com/opencadc/ac)

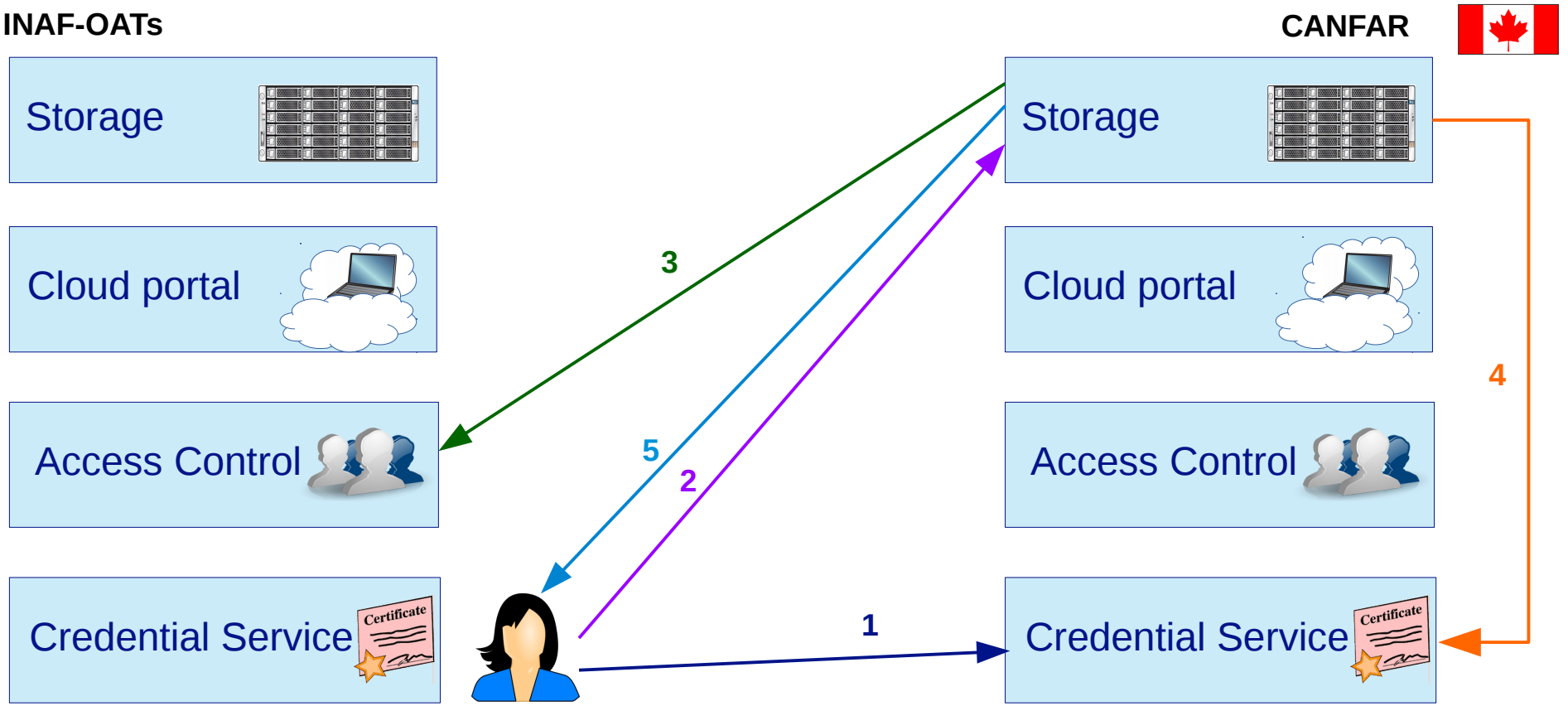
- Manage users
- Manage groups
- Has info about group memberships
- Manage more user's identities:
  - username/password
  - cookies
  - numeric
  - X.509 certificates

Credential Delegation Protocol IVOA recommendation

The credential delegation protocol allows a client program to delegate a user's credentials to a service such that the service may make requests of other services in the name of that user.

Credential Delegation Service
https://github.com/opencadc/cdp
https://github.com/oats-cadc/oats-cred-web

Interoperability



1) INAF-OATs user Bertocco delegates her x509 credentials to CANFAR Credential Service

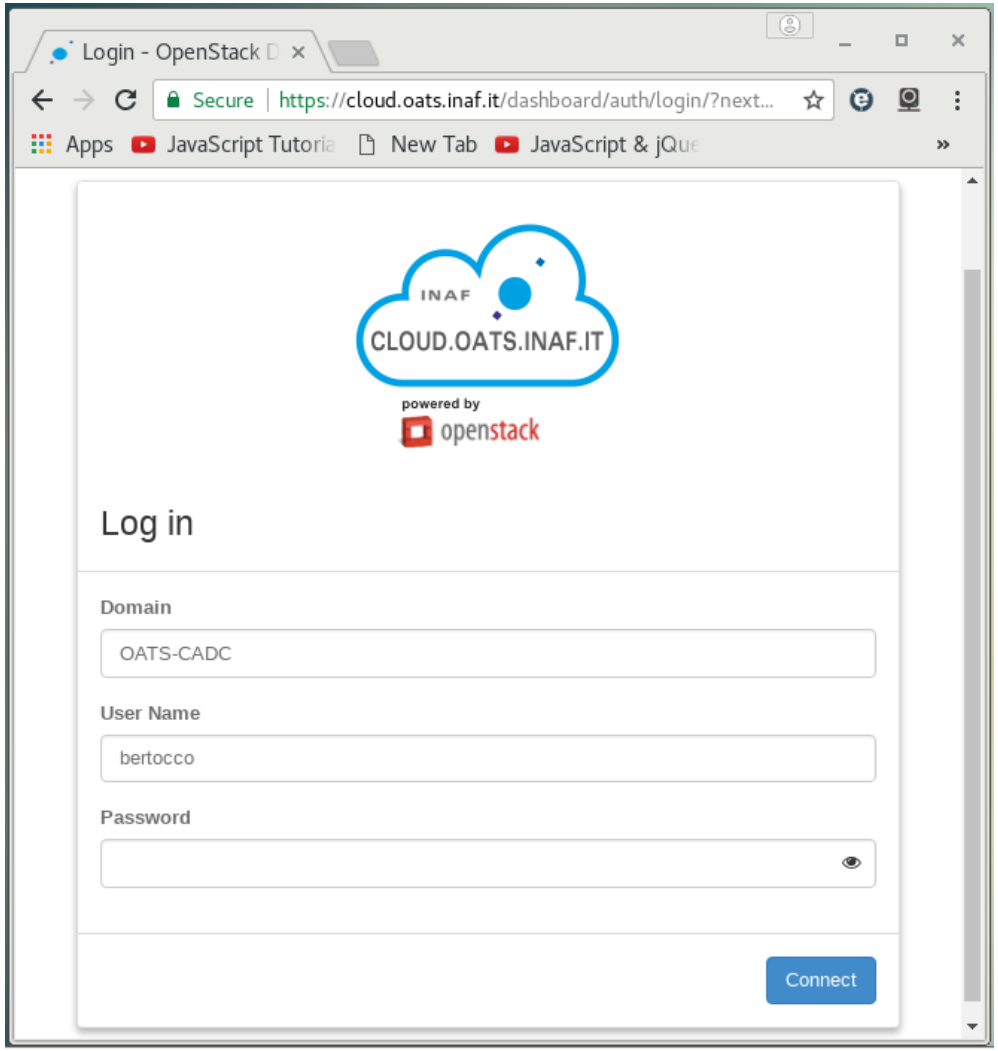
2) user Bertocco asks for data of her INAF-OATs group to CANFAR storage service

3) CANFAR storage service checks the group affiliation of the user in the INAF-OATs group management service

4) CANFAR storage service gets the user's delegated credentials from the CANFAR Credential Delegation Service to be able to make calls to each other service on behalf of the initial user

5) CANFAR storage service returns data to the INAF-OATs user Bertocco

OATs-INAF Cloud Site



OATs-INAF hosts a cloud site:

- OpenStack Mitaka based
- Storage:
  - 10TB VM storage (cinder)
  - 50 TB user's data storage
- Authentication:
  Keystone
  - backend ldap
  - Plus keystone-voms module

EGI FedCloud Compliance & interoperability

Cloud Management Framework: **OpenStack**

(compliant with EGI federated cloud architecture)

Common Authentication and interoperability:

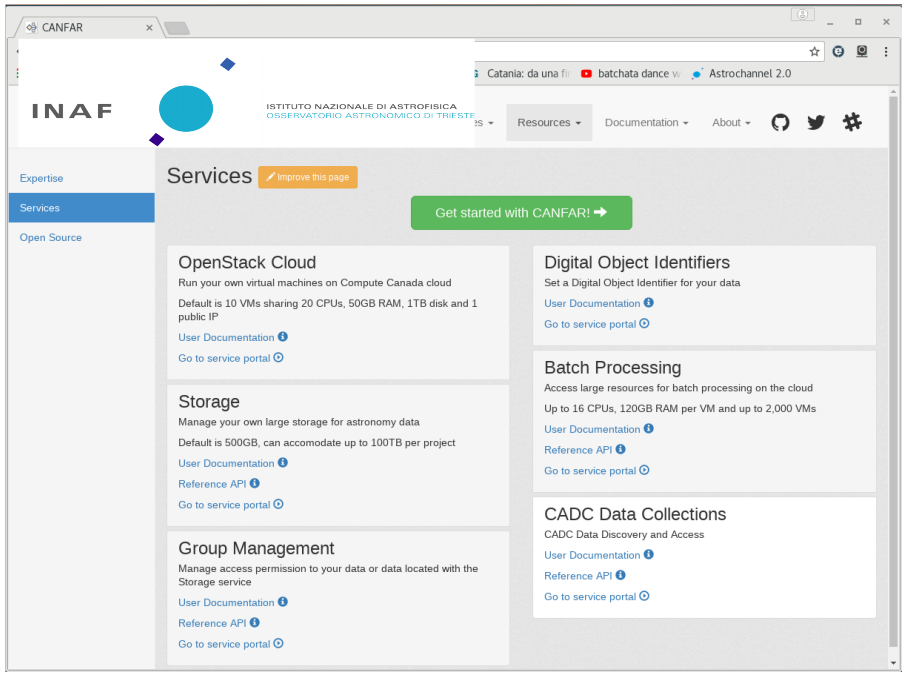
- Using VOMS-proxy
  - Request a cloud authorization token
  - Connect to the OpenStack console
  - Manage virtual machines
  - Authenticate in VOSpace and other IVOA base services
- Using username/password
  - log-in into VMs
  - log-in into OATs cloud portal
  - authenticate in VOSpace and other IVOA base services

Summarizing: wath we have

- Services IVOA recommendation based
  - VOSpace, Access Control and Credential
  - Interoperable with CANFAR Services (VOSpace, Credential Service, Registry)
- Cloud resources
- EGI federation compliant and accessible by EGI users (X.509)
- Cloud access to data stored in VOSpace (both OATs and CANFAR hosted)

Concluding: wath we need

- An integrated approach to our services exploitation



- A common authentication (will be provided by EOSC Pilot ?) and authorization model (will it be an IVOA recommendation on group-based auth?)

Thanks!