

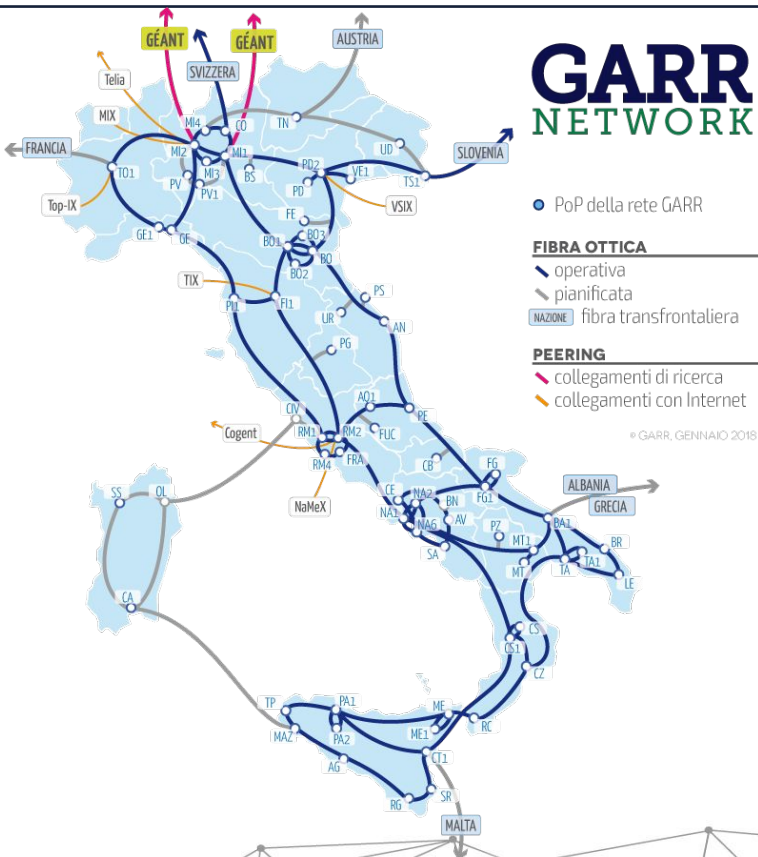
# **GARR Cloud**

Claudio Pisa - GARR

Milano, 22 ottobre 2019  
INAF ICT Workshop



# Consortium GARR



- Italian NREN (National Research and Education Network)
  - nonprofit organization
  - connected to GÉANT
- High speed network
  - 1000+ locations
  - ~4.5 million users
    - teachers, researchers, students
  - 15,000 km of optical fiber
  - 380 Petabyte Yearly Traffic



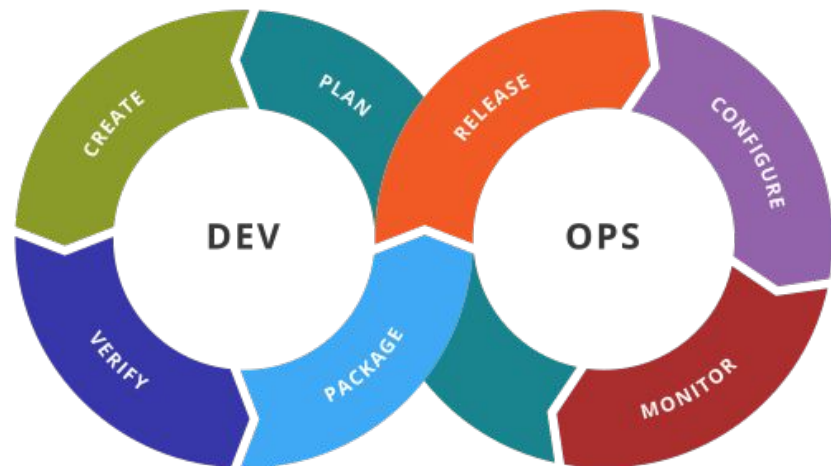
# GARR Distributed Computing and Storage Department



- Role:
  - provider of resources (“long tail of science”)
  - resource aggregator (federation)
  - embody a replicable model for storage and computing provisioning
- Goals:
  - harmonize (SSO / federation of resources)
  - build secure and open infrastructures
  - enhance the user experience

# GARR CSD - Guidelines

- Automation and Replicability
  - Infrastructure as Code
    - self-deployment
  - Open documentation
  - Continuous integration
  - Monitoring and self-healing
- Security and Privacy
  - GDPR
  - ISO 27001



# GARR Cloud Services

- Infrastructure as a Service:
  - GARR Cloud (*OpenStack*)
- Platform as a Service:
  - GARR Container Platform (*Kubernetes*)
  - Deployment as a Service (*Juju*)
- Software as a Service:
  - GARR Workplace (*OnlyOffice*)





- Based on **OpenStack**
  - the open source cloud OS for data centers
  - AWS-like services
    - EC2 → nova, EBS → cinder, S3 → swift, ...
- GARR Cloud Features:
  - Multi-region
  - IDEM authentication
  - Ceph based storage
    - replicated and erasure-coded
- Users:
  - Single Users
  - Virtual Data Centers
    - administrative delegation

Instances - OpenStack Dashboard

https://dashboard.cloud.garr.it/project/instances/

cloudusers • csd-prj1 • garr-pa1

pisa@garr.it

Project / Compute / Instances

## Instances

Instance ID =  Filter [Launch Instance](#) [Delete Instances](#) [More Actions](#)

Displaying 12 items

<input type="checkbox"/>	Instance Name	Image Name	IP Address	Flavor	Key Pair	Status	Availability Zone	Task	Power State	Time since created	Actions
<input type="checkbox"/>	testrebuild	-	192.168.1.3 Floating IPs:	c1.tiny	progresskey	Active		None	Running	1 month	<a href="#">Create Snapshot</a>
<input type="checkbox"/>	testnewglance	-	192.168.1.12 Floating IPs:	c1.tiny	progresskey	Active		None	Running	1 month, 1 week	<a href="#">Create Snapshot</a>
<input type="checkbox"/>	formazione_interna_python	-	192.168.1.19 Floating IPs:	m1.large	pcgarr126	Active		None	Running	5 months, 1 week	<a href="#">Create Snapshot</a>
<input type="checkbox"/>	smtp.cloud.garr.it	-	192.168.0.219 Floating IPs:	m1.medium	progress	Active		None	Running	6 months, 4 weeks	<a href="#">Create Snapshot</a>
<input type="checkbox"/>	onlyoffice2	-	192.168.0.198 Floating IPs:	m1.medium	attardi	Active		None	Running	9 months, 2 weeks	<a href="#">Create Snapshot</a>
<input type="checkbox"/>	clajujuplayground	-	192.168.0.194 Floating IPs:	m1.large	pcgarr126	Active		None	Running	11 months, 1 week	<a href="#">Create Snapshot</a>
<input type="checkbox"/>	grafana	-	192.168.0.175 Floating IPs:	m1.medium	pcgarr126	Active		None	Running	12 months	<a href="#">Create Snapshot</a>



Launch Instance

Details

Source

Flavor

Networks

Network Ports

Security Groups

Key Pair

Configuration

Server Groups

Scheduler Hints

Metadata

Flavors manage the sizing for the compute, memory and storage capacity of the instance.

Allocated

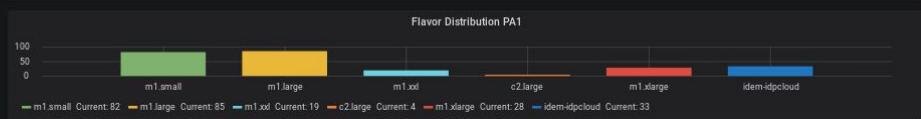
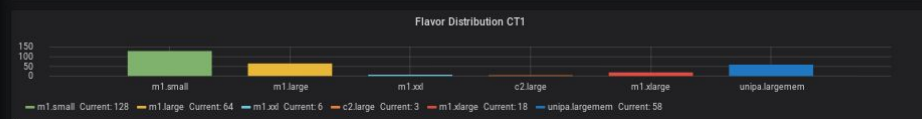
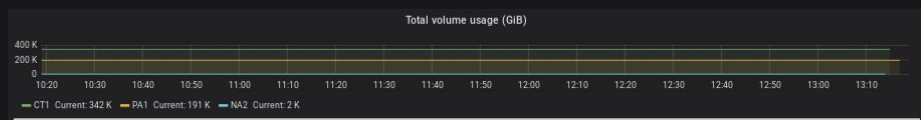
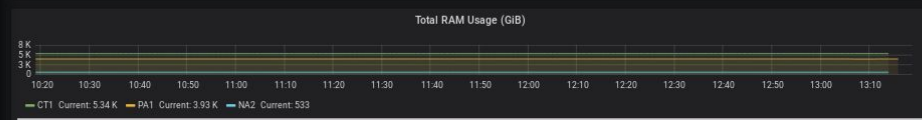
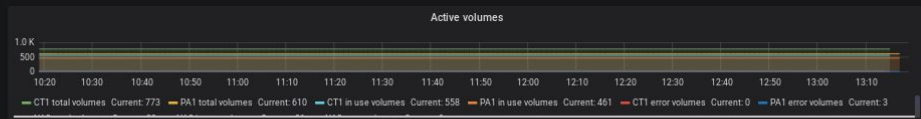
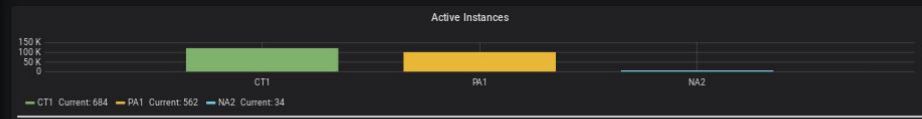
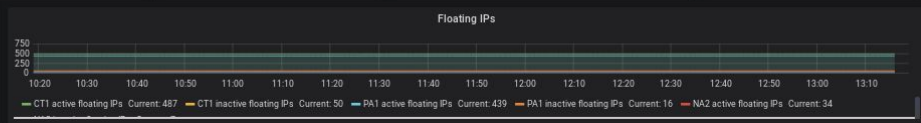
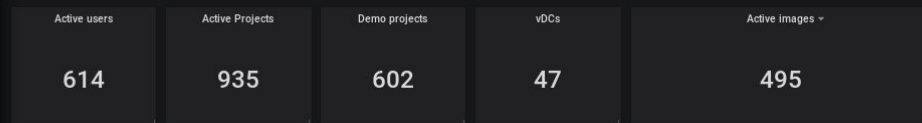
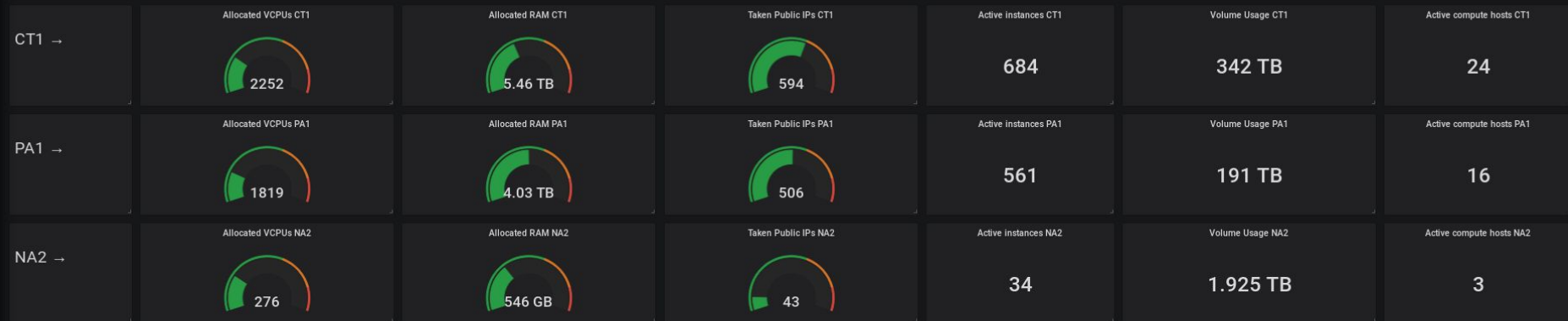
Name	VCPUS	RAM	Total Disk	Root Disk	Ephemeral Disk	Public
Select an item from Available items below						

Available 26

Click here for filters.

Name	VCPUS	RAM	Total Disk	Root Disk	Ephemeral Disk	Public
> c1.tiny	1	1 GB	10 GB	10 GB	0 GB	No
> m1.small	1	2 GB	20 GB	20 GB	0 GB	No
> test-donotuse	1	2 GB	0 GB	0 GB	0 GB	No
> c1.small	2	2 GB	20 GB	20 GB	0 GB	No
> m1.medium	2	4 GB	40 GB	40 GB	0 GB	No
> c1.medium	4	4 GB	40 GB	40 GB	0 GB	No
> d1.small	1	6 GB	20 GB	20 GB	0 GB	No
> c1.large	8	8 GB	40 GB	40 GB	0 GB	No
> m1.large	4	8 GB	80 GB	80 GB	0 GB	No
> m2.small	2	8 GB	40 GB	40 GB	0 GB	No



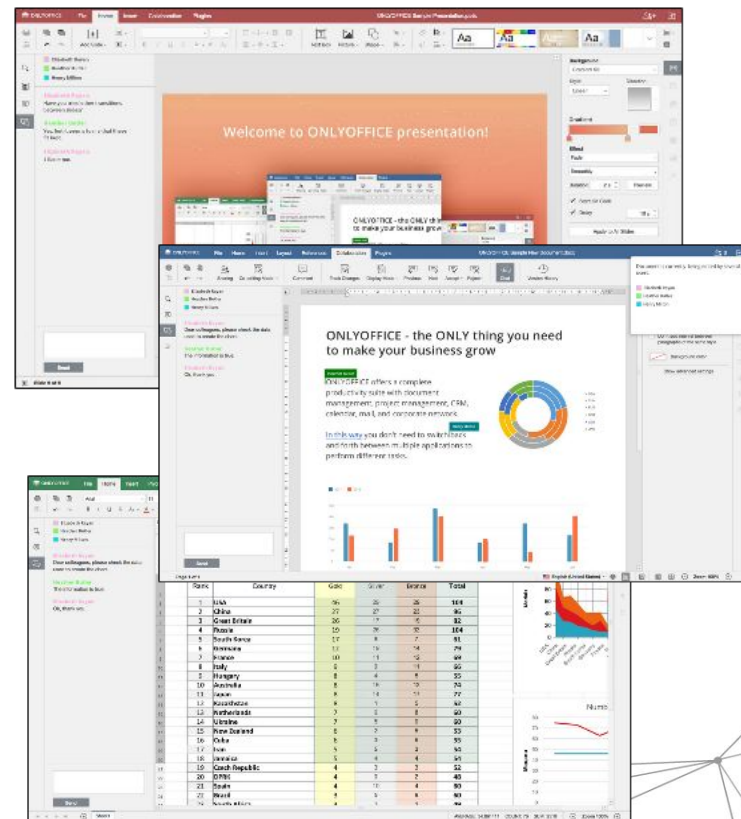




- Environment for automating deployment, scaling, and management of containerized applications
  - based on **Kubernetes**
- Cluster
  - Deployed on bare metal
  - Provides GPU resources
  - Ceph based storage
- Helm support
- Multitenant
  - Unified authentication with OpenStack

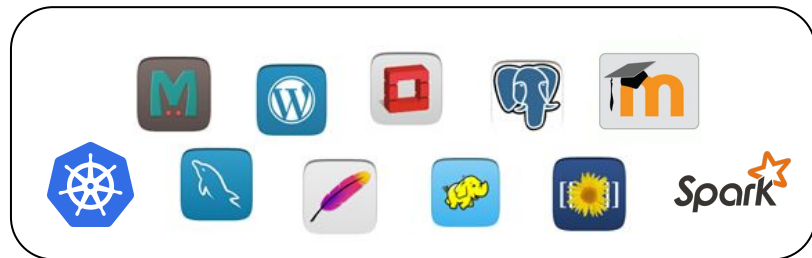
# GARR Workplace

- A web-based office and collaborative suite hosted on the GARR cloud
  - based on OnlyOffice
- Features:
  - Document editor
    - highly compatible with popular document formats
  - Collaborative tools
    - e-mail, calendar, wiki, chat, etc.
  - CRM
  - Project management
  - Pluggable cloud storage
    - e.g. GARRBox



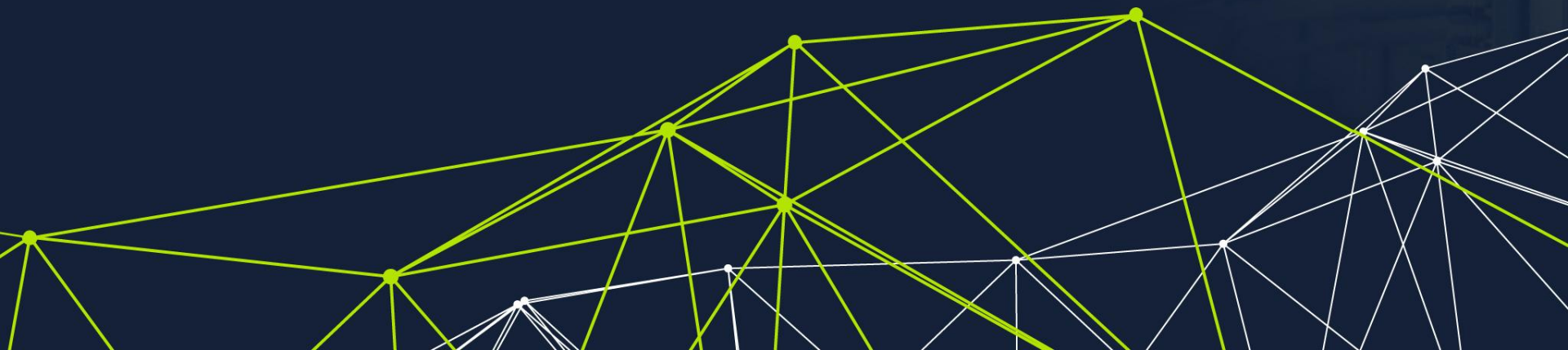
# DaaS (Deployment as a Service)

- A graphical user interface to install applications on the GARR Cloud Platform
- Applications:
  - Hadoop
  - Spark
  - Kubernetes
  - MediaWiki
  - Moodle
  - ...
- Application catalogue at <https://jaas.ai/>



<https://daas-pa.cloud.garr.it>  
<https://daas-ct.cloud.garr.it>

# GARR Cloud technologies



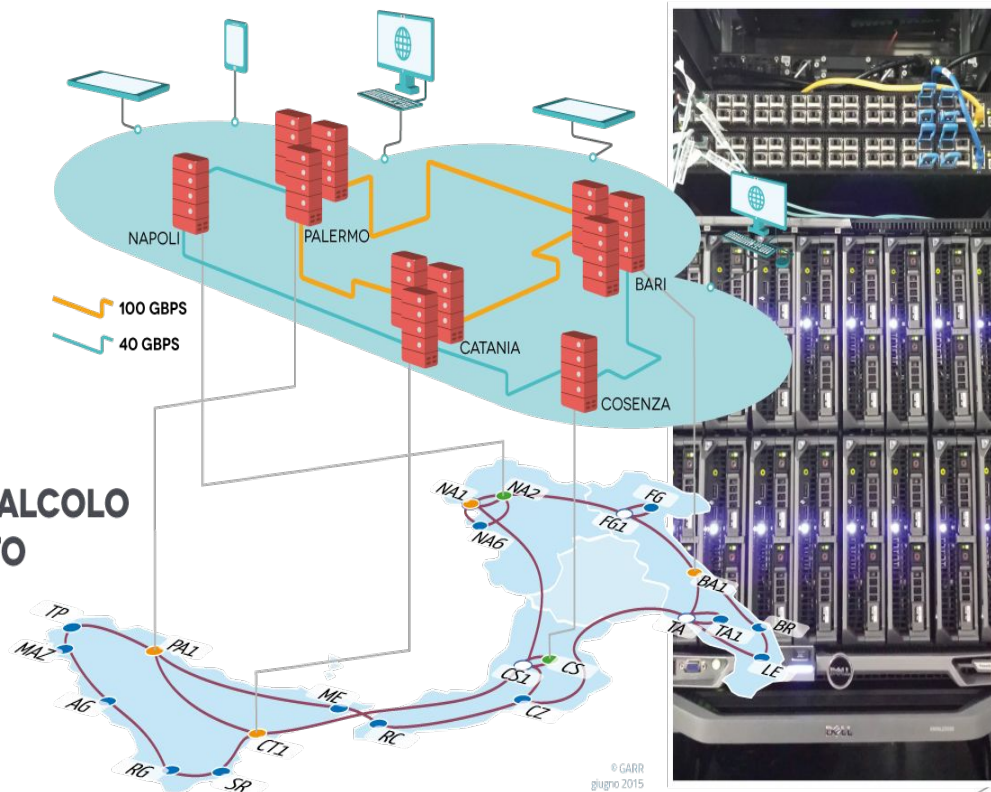
# GARR Computing and Storage Infrastructure

## INFRASTRUTTURA DI CALCOLO E STORAGE DISTRIBUITO

📍 5 siti distribuiti

📄 8.448 virtual CPU

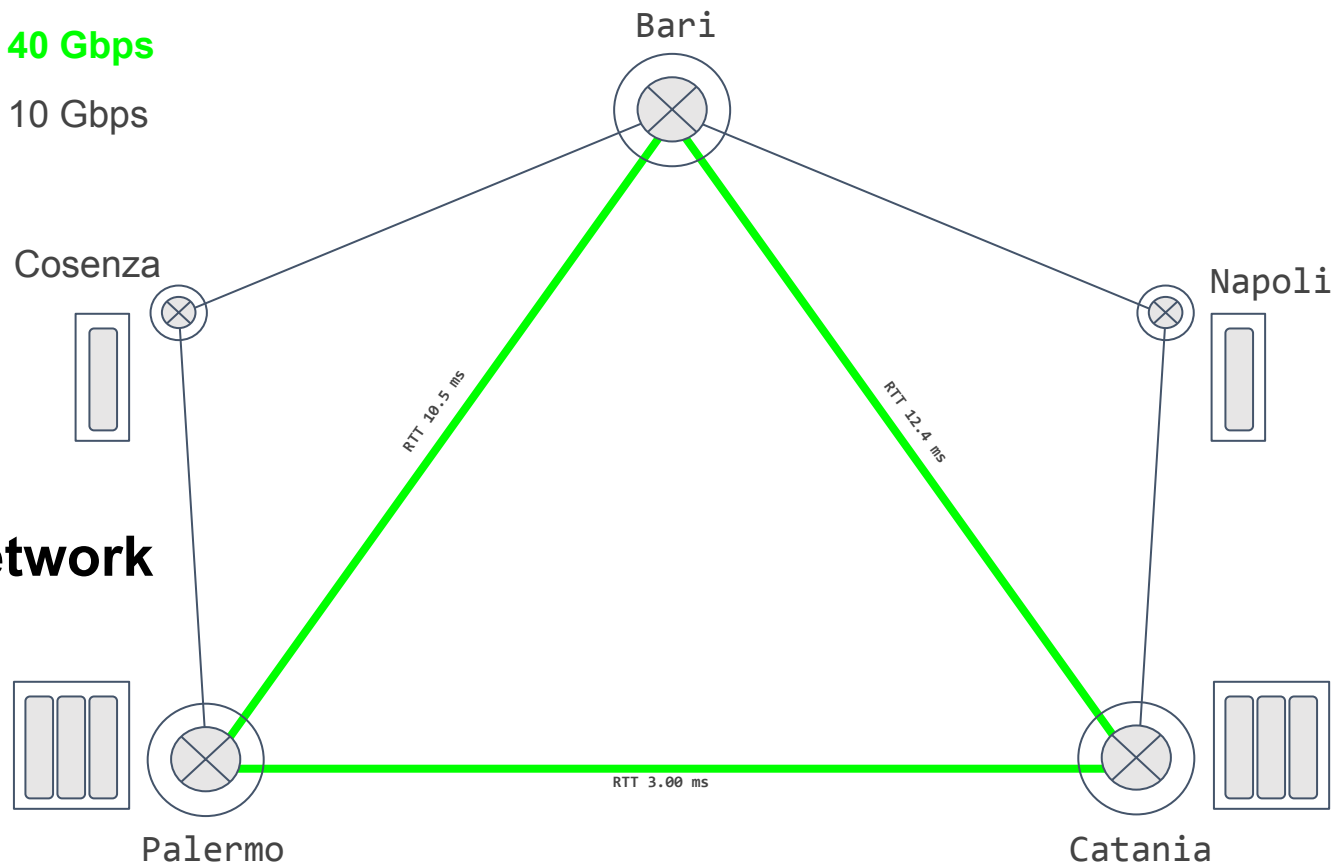
💻 10 PB spazio storage



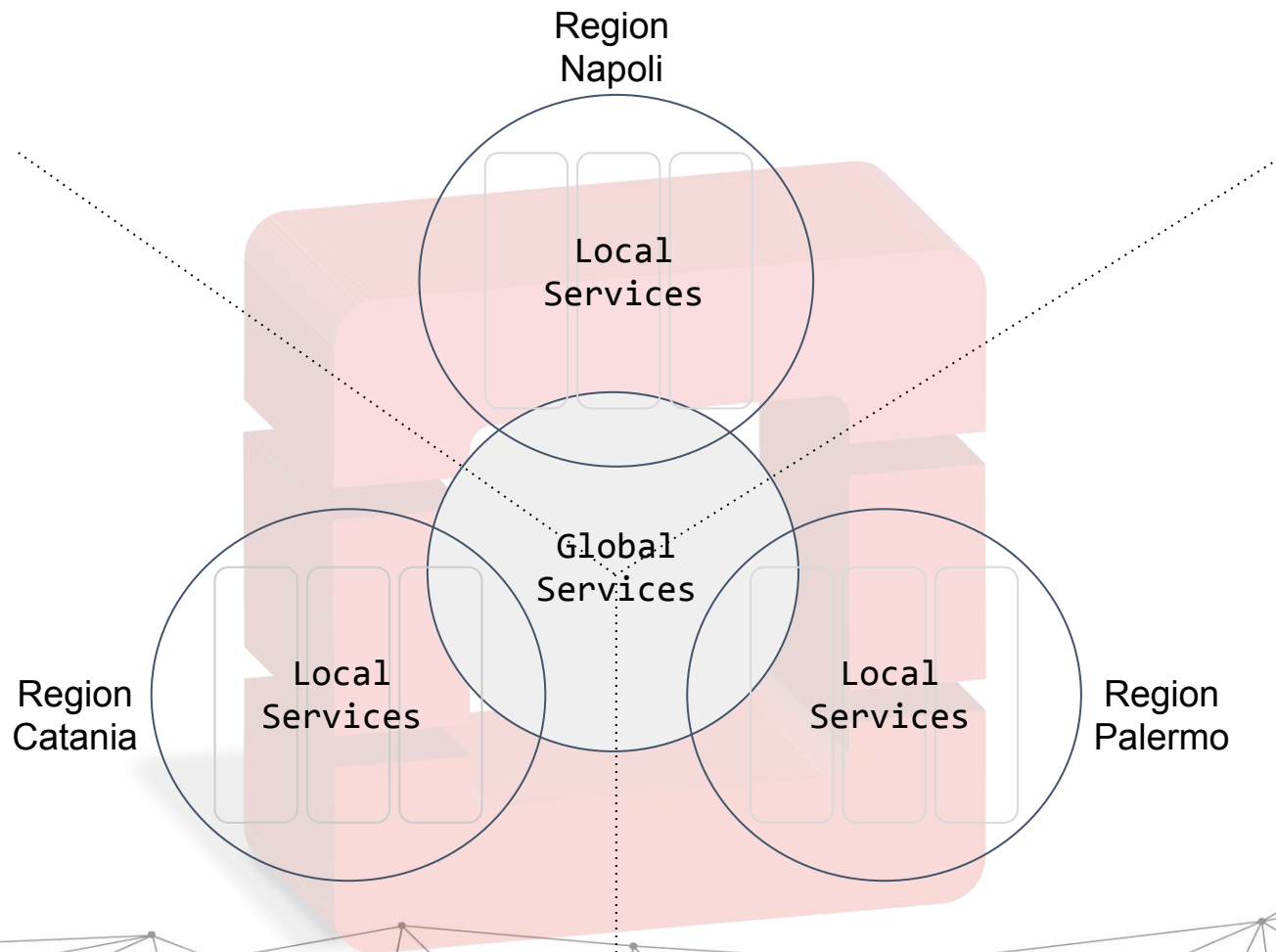
40 Gbps

10 Gbps

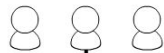
Network







# Internet



- CLI clients(nova, cinder, neutron and so on)
- Cloud management tools
- GUI tools

HTTP(S)

Horizon  
OpenStack Dashboard

heat-api heat-api-cfn

Queue

heat-engine

OpenStack Orchestration

Database LDAP

keystone-all

Optional

OpenStack Identity Service

ceilometer-collector

ceilometer-agent-notification

ceilometer database

ceilometer-agent-compute

Queue

ceilometer-agent-central

ceilometer-api

ceilometer-agent-evaluator

Log or HTTP callback

ceilometer-agent-notifier

OpenStack Telemetry

ironic-api

Queue

ironic-conductor

ironic database

drivers

OpenStack Bare Metal Service

swift-proxy-server

swift-object-server

swift-account-server

swift-container-server

Account database

Object database

Container database

Openstack Object Storage

trove-api

Queue

Trove Database

trove-taskmanager

trove-conductor

OpenStack Database Service

nova-api

nova-scheduler

nova-console

Nova database

Queue

nova-cert

nova-conductor

nova-consoleauth

nova-compute

Guest agent

Instance

Hypervisor

OpenStack Compute

cinder-api

Queue

Cinder database

cinder-volume

Volume provider

cinder-scheduler

OpenStack Block Storage

glance-api

glance store

Glance database

glance-registry

OpenStack Image service

neutron-server

Neutron L2 agent \*

neutron-l3-agent \*

neutron-dhcp-agent

Neutron database

Neutron 3rd party plugin

OpenStack Networking

Optional, depends on plugin \*

sahara-all

Queue

sahara database

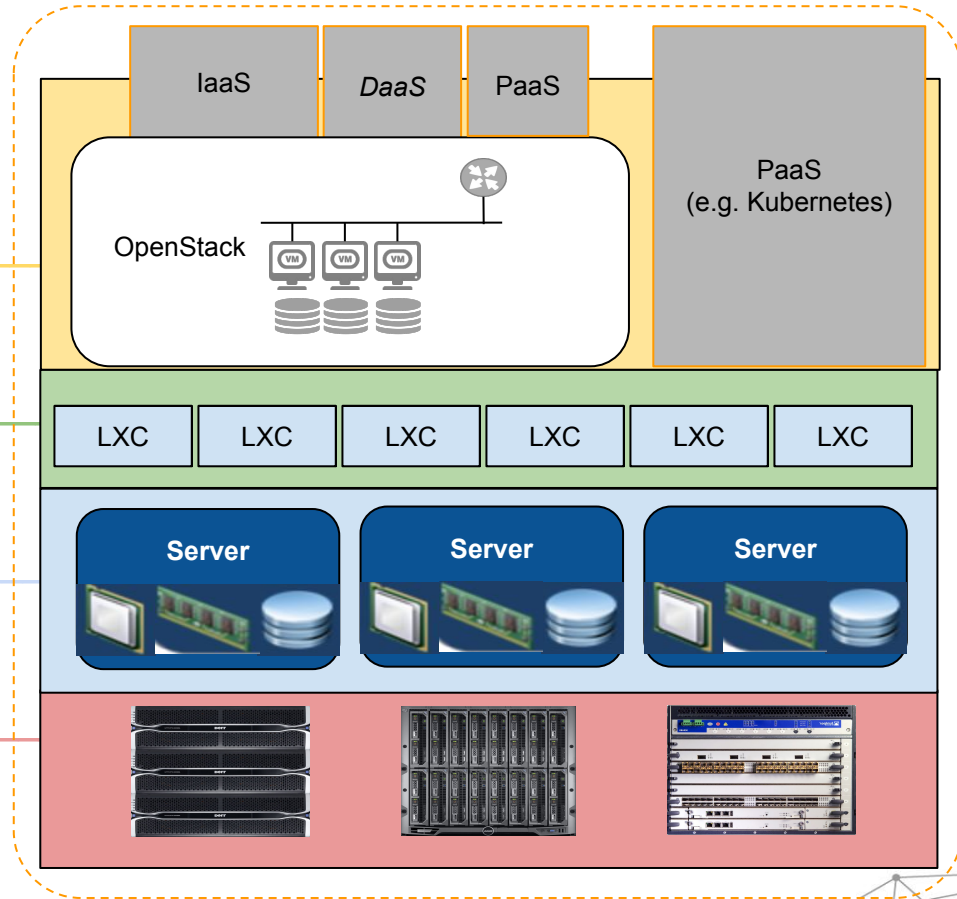
OpenStack Data Processing

1. Application Services

2. Infrastructure *Virtualization*

3. Operating System

4. Physical resources

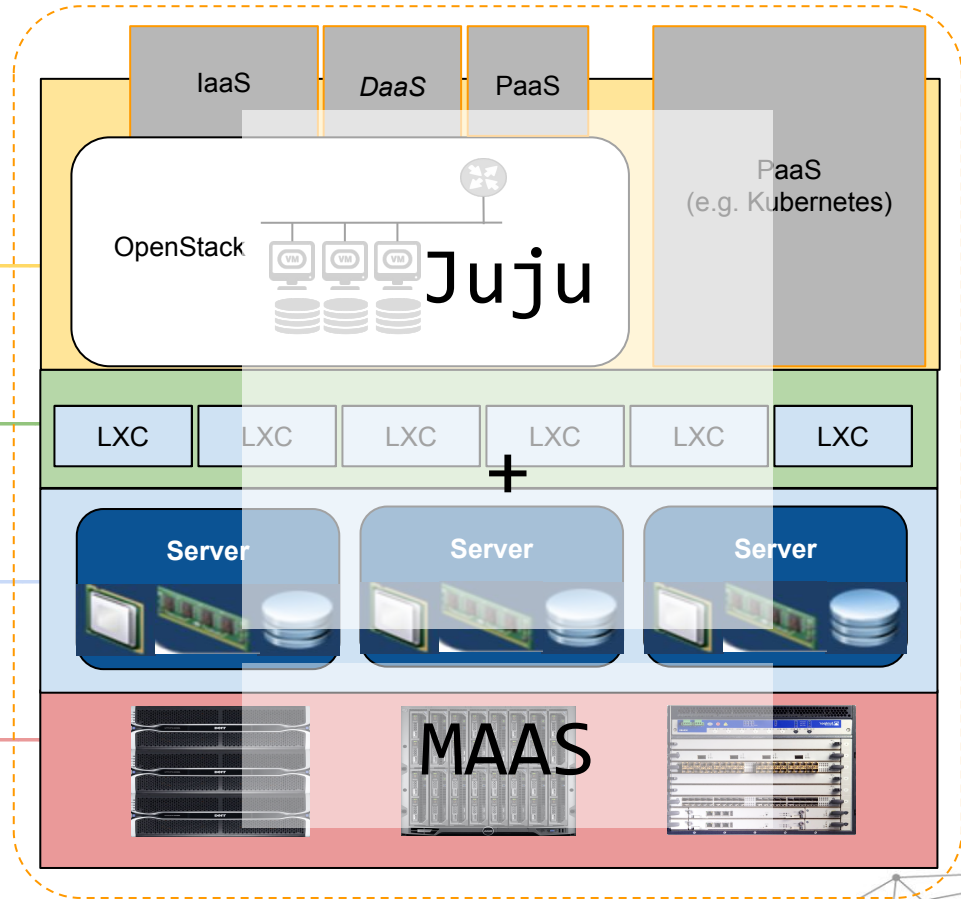


1. Application Services

2. Infrastructure *Virtualization*

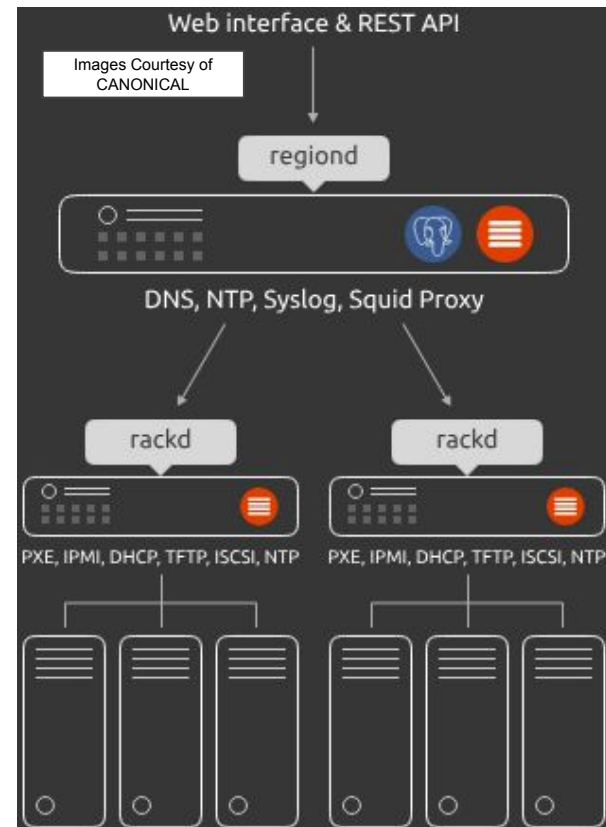
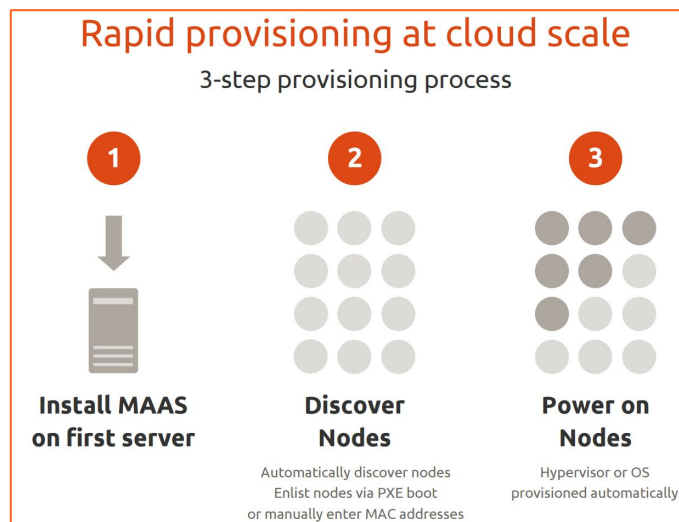
3. Operating System

4. Physical resources



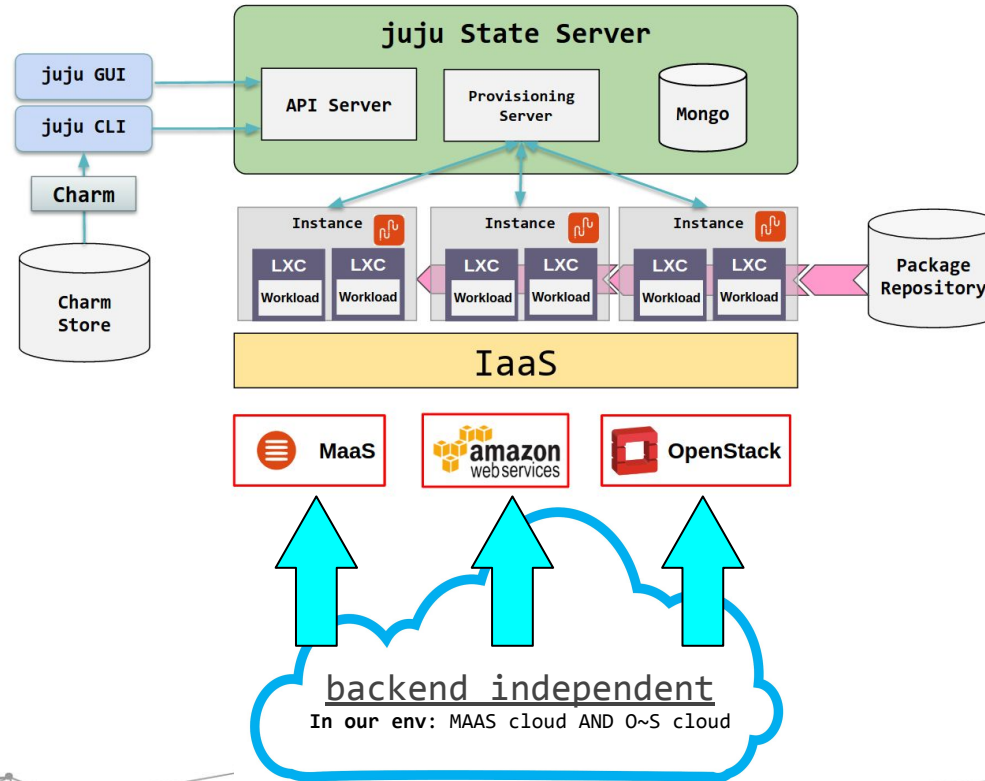
# Metal As A Service (MAAS)

- Discovers and deploys physical servers
- Allocates resources to match requirements
- Undeploys servers when they are no longer needed
- Allows cross datacenter provisioning



- Deployment, configuration and management cloud-native tool
- Free and open source
- Exposes a high-level declarative language
- Supports different backends:
  - OpenStack
  - Kubernetes
  - Amazon AWS
  - Microsoft Azure
  - Google CGE
  - VMWare
  - LXD
  - MAAS

# Juju Architecture

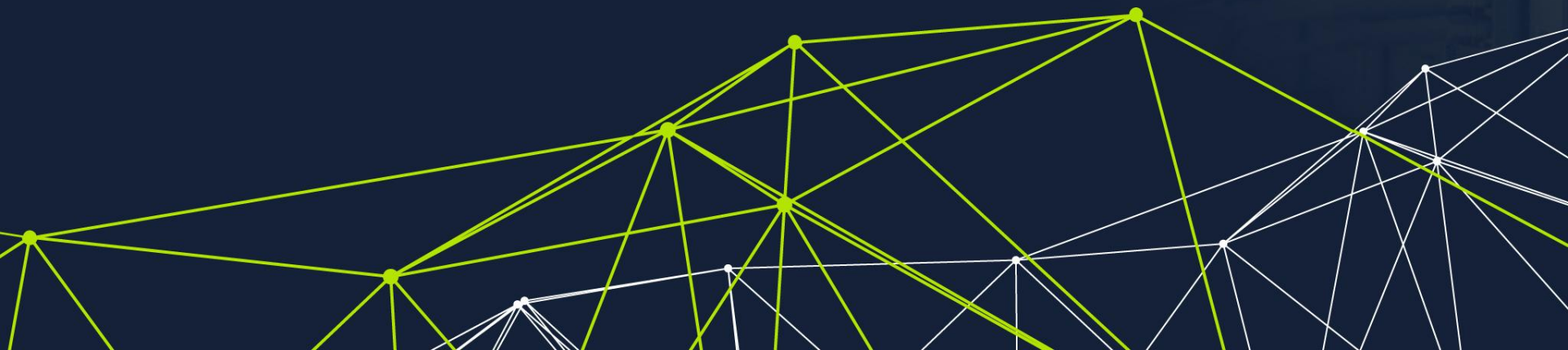




# Open Infrastructures

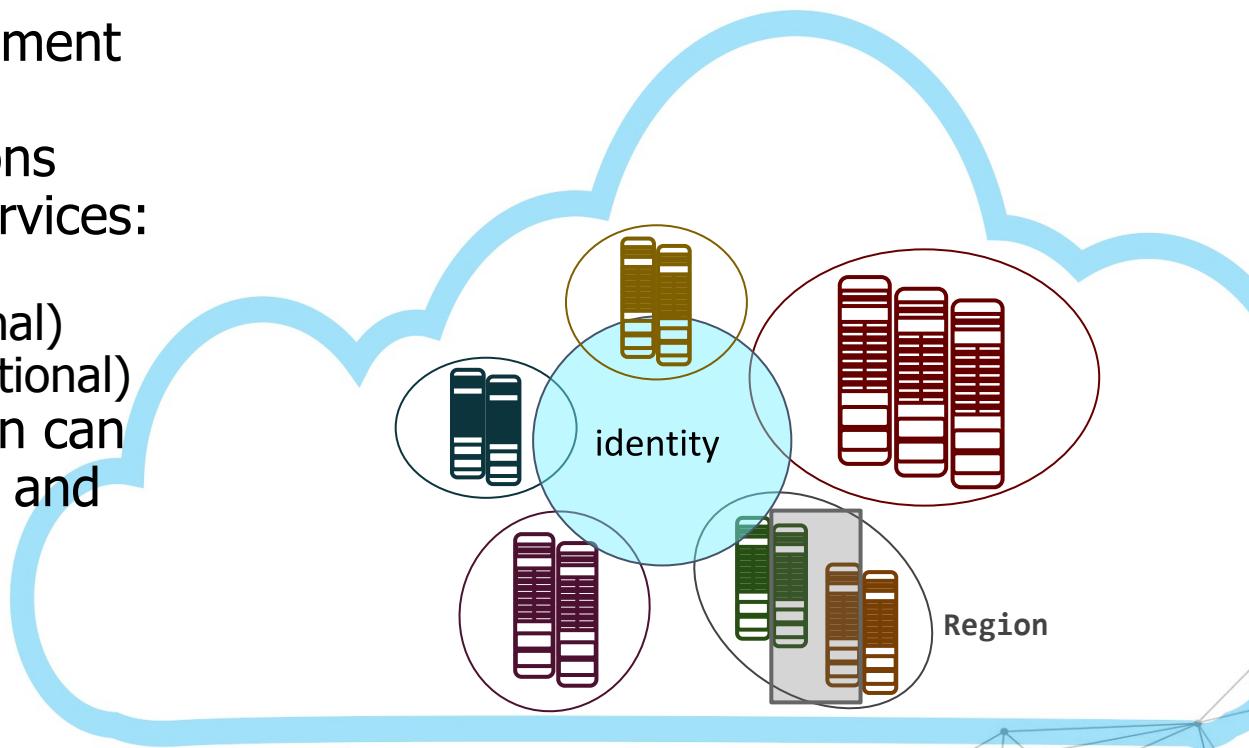
- Open Documentation
  - replicable model
  - Infrastructure as (open source) Code
- Open Source / Free Software contributions:
  - OpenStack Horizon
  - k8s-keystone-auth
  - Juju charms: ceph, keystone SAML/SSO, default gw, moodle...
- Open Federation
  - members of the GARR community can federate their computing resources

# GARR Cloud Federation



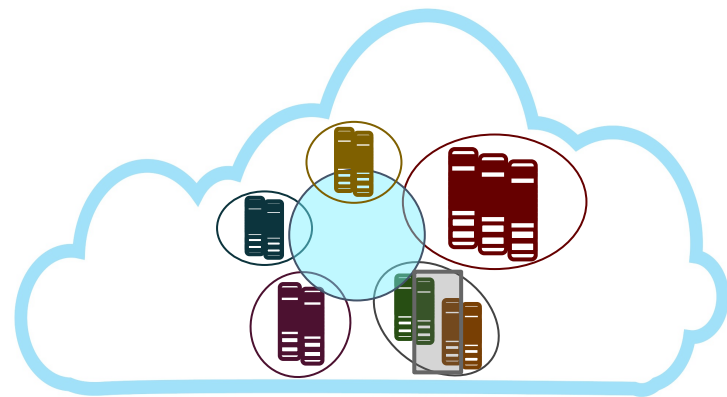
# Multi-region/multi-domain Federation Model

- Region: separate deployment of OpenStack
  - linked to other regions through common services:
    - identity service
    - dashboard (optional)
    - image service (optional)
  - nodes inside a region can be logically grouped and reserved to specific projects



# GARR Cloud Federation

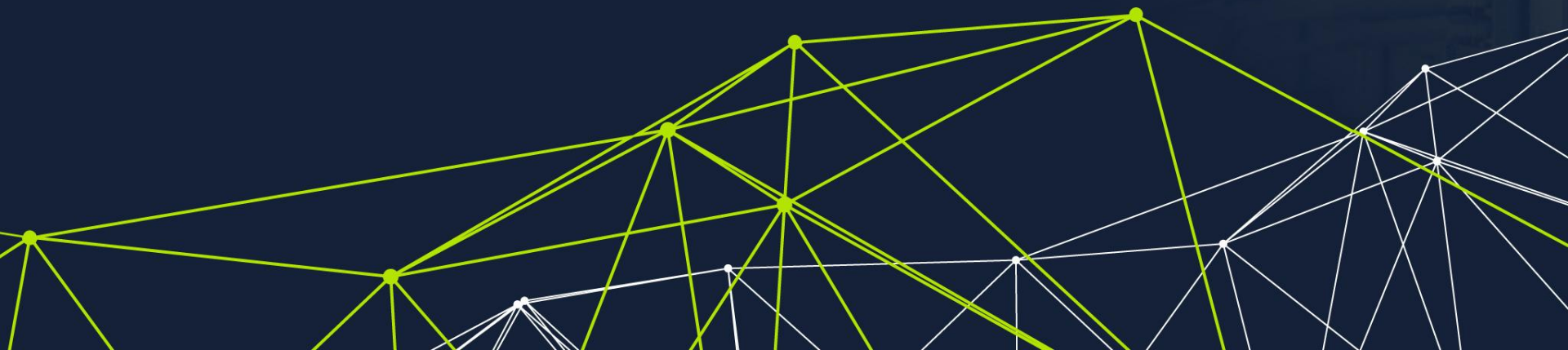
- Simple Federation recipe
  - Reference architecture with use cases
  - git repository and knowledge base
    - <https://git.garr.it/cloud/federation/>
    - <https://cloud.garr.it/doc/federation/>
  - Federated authentication
  - Administrative delegation



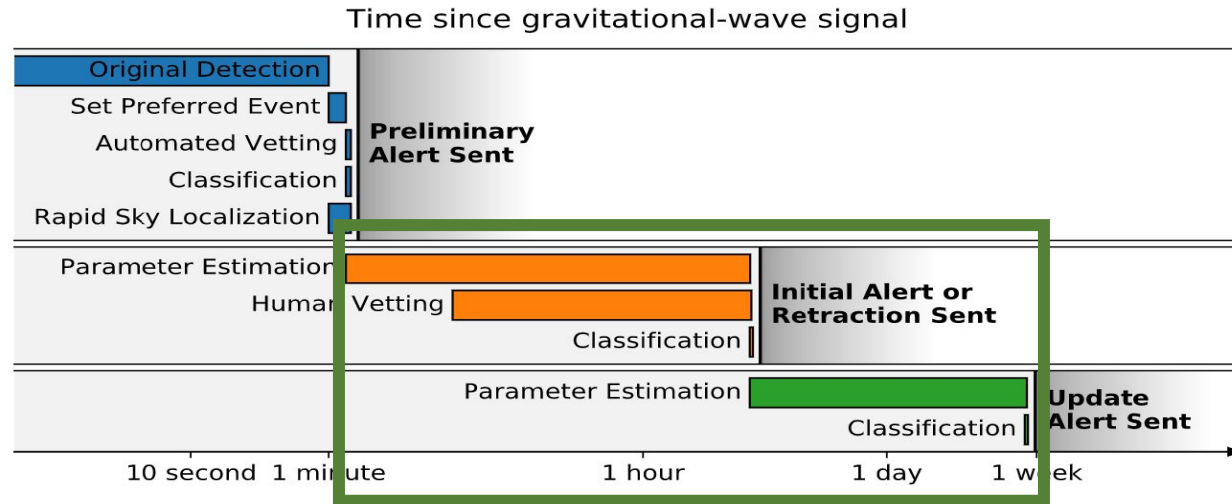
# GARR Cloud Federation Features

- Low latency, Terabit network
- Federated authentication
- **“The Cloud in your datacenter”**
  - Data Sovereignty
    - keep control on every data processing and communication stage
  - Know How Fostering
  - Operational Ability
    - certifiable data and metadata confidentiality, which enables collaboration with several external organizations
  - Legal Evolution Tracking
    - Perimetro di Sicurezza Cibernetica nazionale (Italia 19/9/19)
    - Clarifyng Lawfull Overseas Use of Data Act (CLOUD Act) (USA 23/3/18)
    - 中华人民共和国网络安全法 (Cybersecurity Law of the People’s Republic of China) (Cina 7/11/16)

# GARR Cloud Use Cases



# Virgo - H2020 Asterics - Gravitational Wave alert system



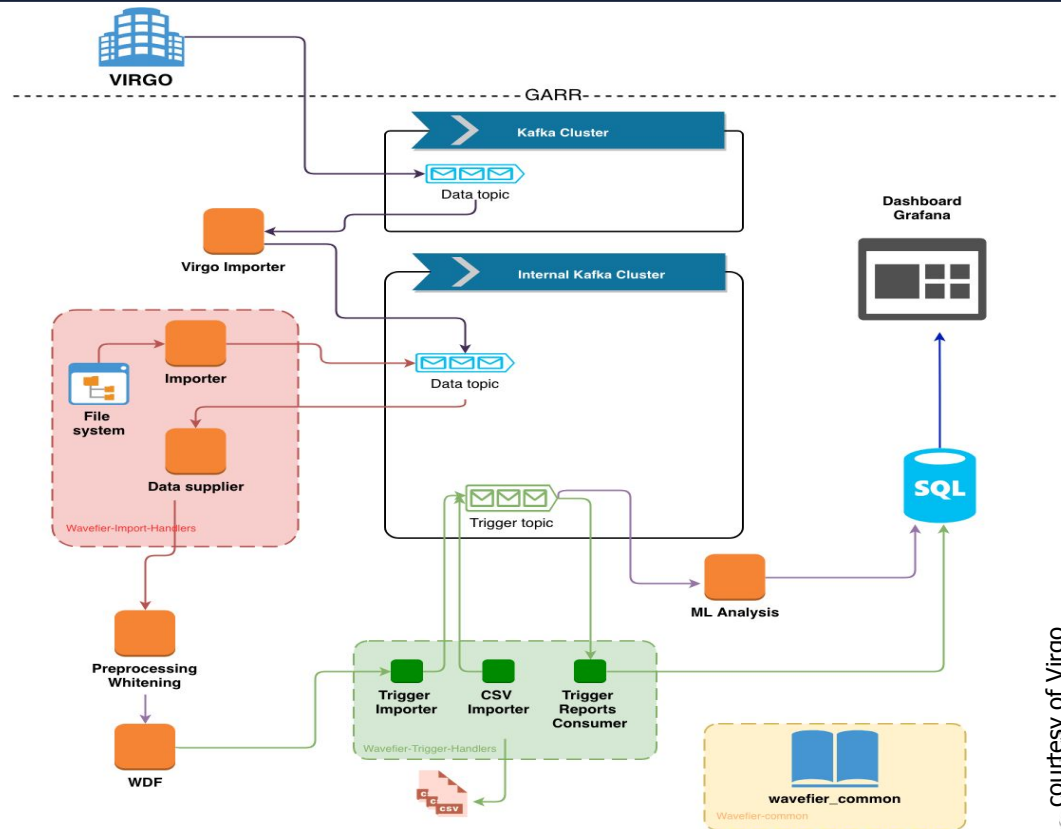
<https://emfollow.docs.ligo.org/userguide/index.html>



# Virgo - H2020 Asterics - Wavefier Gravitational Wave Detection over the GARR Container Platform



Wavefier:  
real time pipeline for the detection of transient  
signals and their automatic classification based  
on big data / machine learning



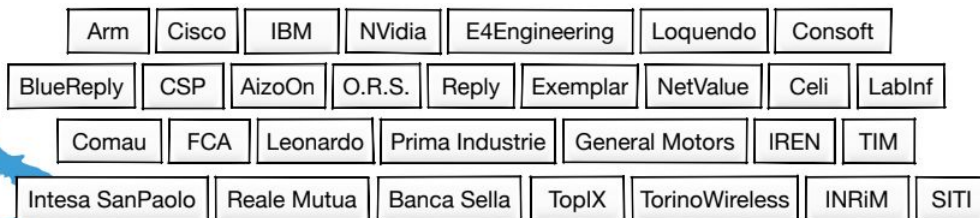
# INGV - Seismic e-Infrastructure

- Seismic measurement, alerting and analysis infrastructure
- Federated cloud architecture, based on the GARR Cloud model
  - Work in progress



# Università di Torino

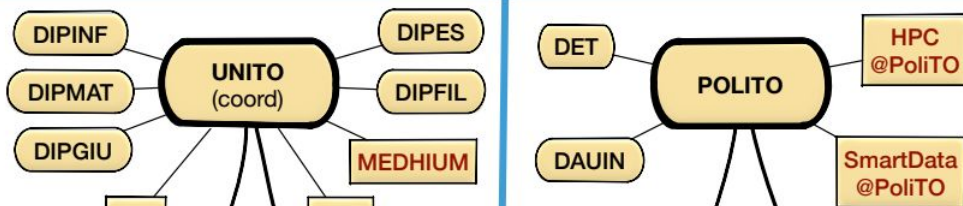
Supported  
by



Technological  
Partner



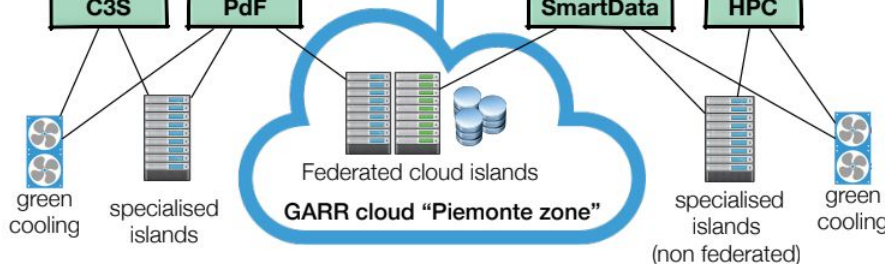
Universities  
and  
Departments



Data Centers



Hardware



AI-in-demand platform

## Facts

- INFRA-P call Nov. 2017
- Ranked 1st on ~30 submitted projects
- Kick-off mid apr 2018
- 4.5M€ funding
- 2 partners
- 8 associated partners
- Coord. M. Aldinucci
- Many industrial stakeholders

# Politecnico di Torino - GARR Cloud OpenStack Region

- One of the first GARR Cloud federated regions
- Used for:
  - IaaS - user self service
  - Cloud native application development space
  - Experiment cross-region highly available services
  - Data backup
    - right now 60TB/day of deduplicated and encrypted data

**CLOUD**

**1,500+**

CPU cores

**400 TB**

Storage

**9 TB**

RAM



**openstack®**

# ISTI CNR - D4Science + OpenAire

- D4Science.org
  - Integrated technologies that provide elastic access and usage of data and data-management capabilities
    - Virtual Research Environments (VRE)
    - Data discovery, accessing, analysis, and transformation
- OpenAIRE
  - A pan-European (and global) network for Open Science
- Infrastructure:
  - 500+ VMs, ~100 of them hosted by the GARR cloud
  - 75% of them are production services (they must work)
  - Migration, in 2018, from Xen/Aoe to OpenStack/Ceph

- GARR Cloud used for alignment-free sequence comparison
  - Map/Reduce
    - Spark
    - Hadoop
- GARR Cloud credited in scientific publications

Ferraro Petrillo U., Roscigno G., Cattaneo G., Giancarlo R. "FASTdoop: a versatile and efficient library for the input of FASTA and FASTQ files for MapReduce Hadoop bioinformatics applications.", *Bioinformatics*. 2017 May 15;33(10):1575-1577. doi: 10.1093/bioinformatics/btx010.

Ferraro Petrillo U., Roscigno G., Cattaneo G., Giancarlo R., "Informational and linguistic analysis of large genomic sequence collections via efficient Hadoop cluster algorithms.", *Bioinformatics*. 2018 Jun 1;34(11):1826-1833. doi: 10.1093/bioinformatics/bty018.

Invited Talk, Raffaele Giancarlo, "Methods, tools & platforms for Personalized Medicine in the Big Data Era", NETTAB 2017 Workshop, Palermo, October 16-18, 2017



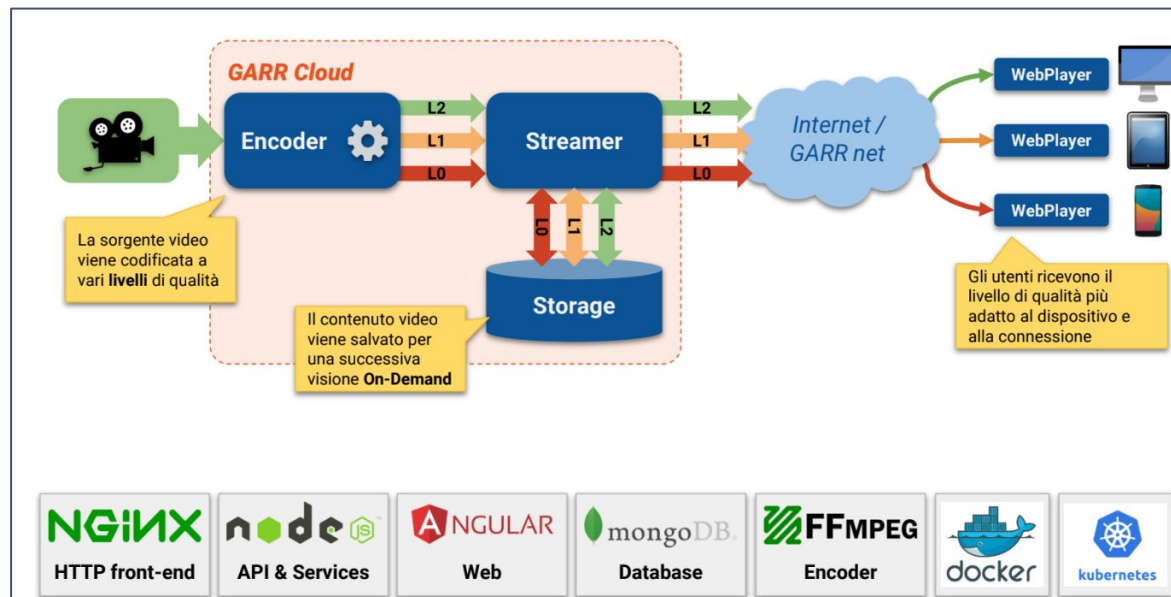
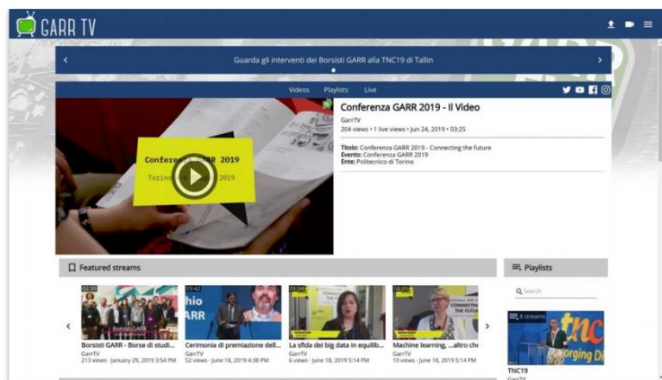
# IEO - Bioinformatics

- container-based bioinformatic research applications
  - reproducible results
- GARR Cloud advantages
  - different research groups can access the same infrastructure
  - dedicated network access/VLANs
  - data doesn't need to cross different networks





- Video streaming platform - <https://www.garr.tv>
  - adaptive video encoding
  - deployed with Kubernetes



# GARR IdP in the Cloud

- Federated identity as a service
  - IDEM and eduGAIN compliant
  - Targeted to small research and education organizations
  - Manage user digital identities without the need to manage also the technological aspects

- Online since 2001
- 200 TB in two network locations
- Backup in the GARR Cloud

# EAPConnect

---

- East European NRENs
- Replicating the GARR Cloud model

# HACK the CLOUD

2019

Milano, 5-7 novembre

Universita' degli Studi di Milano-Bicocca



# Grazie

**Claudio Pisa**

GARR - Dipartimento CSD

<https://cloud.garr.it>

<mailto:cloud-support@garr.it>