The infrastructure of UniNuvola

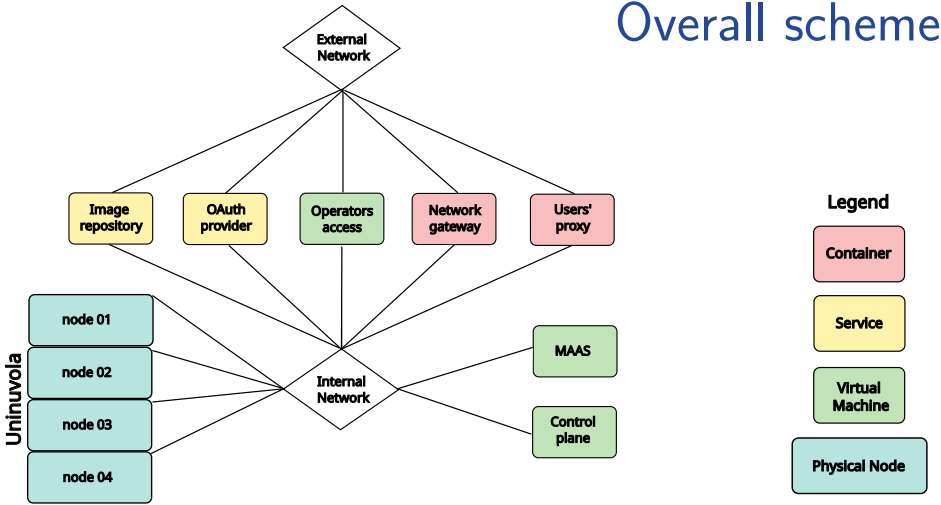
UNIVERSITÀ DEGLI STUDI
DI PERUGIA

June 26, 2024

Contents

Overall scheme

Overall scheme



The Metal

The nodes

4 Dell Power Edge R940 servers:

- 2 Intel Xeon Gold 6252N CPU
- 512 GB ECC DDR5 RAM
- 8 x 2.5" disks (2 OS + 6 CEPH)
- 16 TB storage
- 4 x 1 Gbit/s NIC
- 2 x 10 Gbit/s NIC
- Redundant power supply

The Network

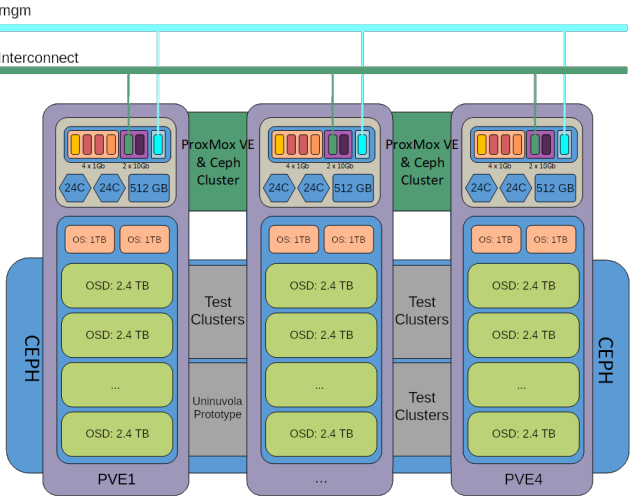
Hardware:

- 2 x Quanta LB6M - 24 port (10 Gbit/s)
- 2 x Quanta LB4M - 48 port (1 Gbit/s) + 2 port (10 Gbit/s uplink)

VLANs:

- 8: Management (ipmi, 1 Gbit/s) 10.8.0.0/16
- 9: Interconnection (10 Gbit/s) 10.9.0.0/16

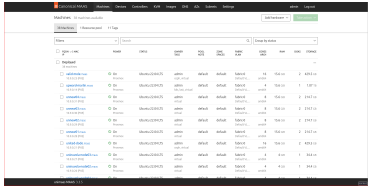
The choice of the VLANs is made to avoid conflicts with the Department's network.



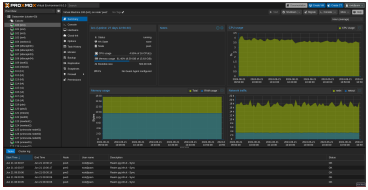
The infrastructure of UniNuvola

Metal and VM provisioning

Canonical MAAS



Proxmox VE



Metal

- inventory management
- Hardware provisioning
- ipmi management

VMs

- VM commissioning and deployment
- OS installation

Networking

External networking elements

Two external networking containers have been implemented:

uninuvolagw

Network gateway for the 10.9.0.0/16 subnet

```
#!/bin/sh

echo 1 > /proc/sys/net/ipv4/ip_forward

# SNAT
/sbin/iptables -t nat -I POSTROUTING 1 -s 10.9.0.0/16 ! -d 10.9.0.0/16 -j SNAT --to-source 141.250.2.8

# Clamping
/sbin/iptables -t filter -I FORWARD 1 -p tcp --tcp-flags SYN,RST SYN  -j TCPMSS --clamp-mss-to-pmtu

# Keep state
/sbin/iptables -t filter -N keep_state
/sbin/iptables -t filter -F keep_state
/sbin/iptables -t filter -A keep_state -m conntrack --ctstate INVALID -j DROP
/sbin/iptables -t filter -A keep_state -m conntrack --ctstate ESTABLISHED,RELATED -j ACCEPT

# Forwarded rules
/sbin/iptables -t filter -P FORWARD ACCEPT
/sbin/iptables -t filter -A FORWARD -j keep_state
/sbin/iptables -t filter -A FORWARD -j matrix
#/usr/sbin/dhcpd -cf /etc/dhcpd.conf eth1

while true; do
        sleep 1000
done
```

uninuvolaproxy

Reverse proxy for the uninuvola services

```
server {
        listen                          443 ssl;
        server_name                     uninuvola.flsgeo.unipg.it;

        ssl_certificate                 /etc/ssl/certs/nginx/uninuvola.flsgeo.unipg.it/fullchain.pem;
        ssl_certificate_key             /etc/ssl/certs/nginx/uninuvola.flsgeo.unipg.it/privkey.pem;
        add_header                      Strict-Transport-Security "max-age=31536000";

        location / {
                proxy_pass                      https://10.9.3.1/;
                proxy_buffering                 off;
                proxy_set_header Host           $http_host;
                proxy_set_header X-Real-IP      $remote_addr;
                proxy_set_header X-Forwarded-For        $proxy_add_x_forwarded_for;
                proxy_set_header X-Forwarded-Proto      $scheme;
                proxy_http_version              1.1;
                proxy_set_header Upgrade        $http_upgrade;
                proxy_set_header Connection     "Upgrade";
                client_max_body_size            2G;
        }
}
```

Operator's access

The uniaccess VM has been created for the operators. Its main features are:

- Bastion host for the operators, allowing access to the internal network
- Build and deploy UniNuvola images
- repository for the UniNuvola recipes and automation scripts

Authentication

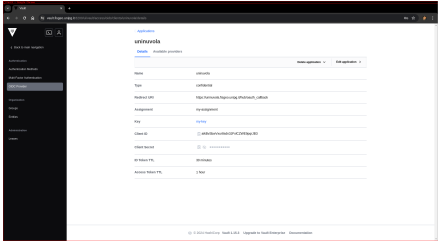
LDAP

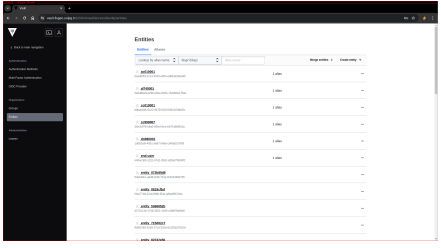
```
$ ./matricoler.py
transform["Rossi-Mario"]="Mario-Rossi"
to_import_cns["Rossi-Mario"]=["User"]
extra_filter["Rossi-Mario"]={'matricolaDipendente': "123456"}
extra_account_services["Rossi-Mario"]=['ppim','uninuvola']

$ ./authuniv_importer.py
Mario Rossi->Mario Rossi
Created - cn=Mario Rossi,ou=people,dc=priv - {'perugiarole': 'User', 'givenname': 'Mario', 'sn': 'Rossi'}

Created - cn=Mario Rossi,ou=users,dc=priv - {'PerugiaRole': ['User'], 'PerugiaExpire': ['20263'],
    'cn': ['Mario Rossi'], 'PerugiaActivation': ['19898'], 'objectClass': ['inetOrgPerson',
    'posixAccount', 'top', 'PerugiaObject'], 'userPassword': ['{SSHA}xxx'], 'uidNumber': ['2571'],
    'PerugiaOwner': ['Mario Rossi'], 'gidNumber': ['1000'], 'sn': ['Rossi'], 'homeDirectory':
    ['/home/mm123345'], 'PerugiaGroups': ['account', 'ppim', 'uninuvola'], 'givenName': ['Mario'],
    'PerugiaType': ['UNIPG'], 'uid': ['mm123345']}
```

The Physics and Geology Department has its own LDAP server that inherits the University LDAP and add some specific authorizations tags.
The Uninuvola project uses this LDAP server and a dedicated tag for the users.

Vault

HashiCorp Vault is used as an OAuth2 authenticator back-end for the LDAP.
The users can authenticate themselves using the University credentials.



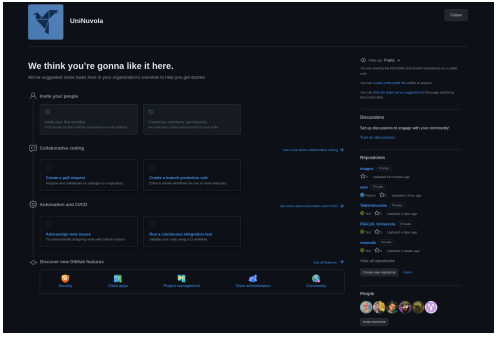


Repositories

github

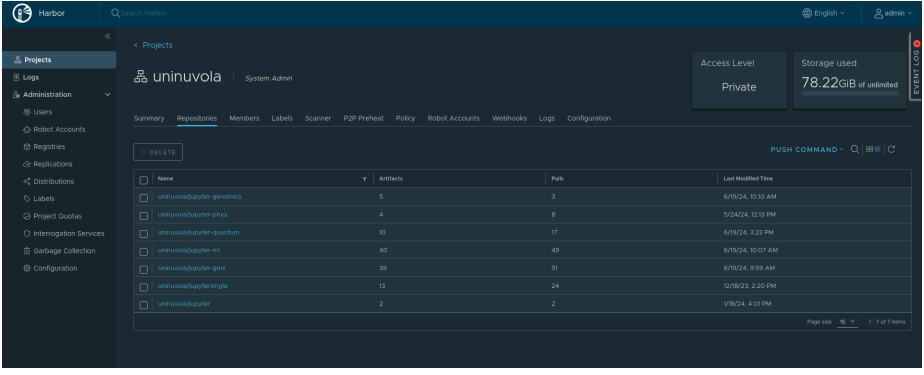
To keep track of the project, a GitHub Organization has been created.
The repositories are divided into:

- **ansible**: the ansible playbooks for the Kubernetes cluster
- **manuals**: the users' manuals
- **internal docs**: the internal documentation for the developers
- **images**: the recipes for the container and VM Creation
- **web**: the code for the uninuvola portal



harbor

The PhysGeo Harbor registry is used to store the images created for the Uninuvola project, that contains both the images for the scientific applications and the images for the infrastructure.



Services

Services

The Uninuvola users will have access to services provided via container technology.

A container is a standard unit of software that packages up code and all its dependencies so the application runs quickly and reliably from one computing environment to another.

All Uninuvola services are containerized and managed by Kubernetes.

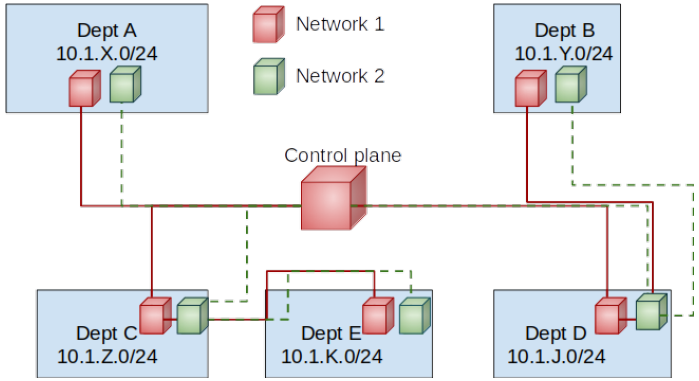
Kubernetes

Among the different installation and management solutions, the Uninuvola kubernetes cluster implements the following:

- **Ansible**: Uninuvola setup has been automated using Ansible. Multiple k8s control planes with HA are supported.
- **Kube-ovn***: the network solution for Kubernetes
- **Rook Ceph***: the distributed storage solution
- **Kubevirt***: the virtualization extension for Kubernetes

Kube-OVN

A good network solution is essential for reliable, efficient and secure infrastructure.



https://www.kube-ovn.io/

- **OVN** is implemented through *kube-OVN*.
- different underlay networks, mapped to different VLANs, eventually spanning multiple Depts.
- the new **Conca Backbone** will be used for the interconnection.

Rook Ceph

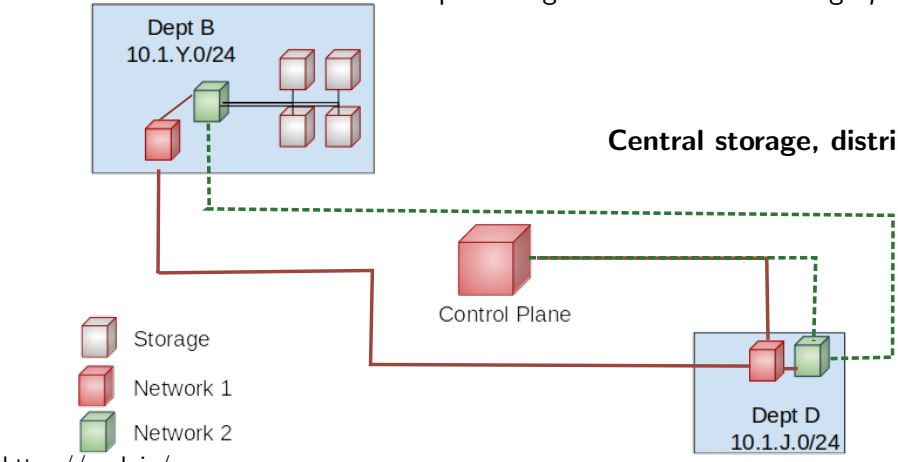
Rook is an open-source operator for managing Ceph storage on Kubernetes. It automates the deployment, configuration, and management of Ceph clusters.
Key features of Rook Ceph include:

- Dynamic provisioning of Ceph storage resources such as pools, block devices, and file systems.
- Support for various storage backends, including local disks.
- Scalability to handle large-scale storage deployments.

In the present Uninuvola prototype, Rook Ceph uses virtual disks for storage, but in the future, it will be extended to use distributed physical disks without downtime.

Rook Ceph

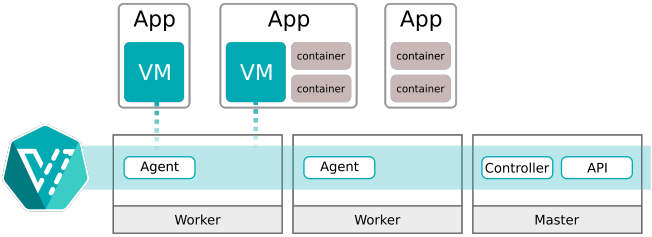
CEPH distributes data across multiple storage devices to achieve high *performances*.



https://ceph.io/

Kubevirt

Kubevirt provides new features for the virtualisation functionalities to Kubernetes.



A KubeVirt VM is a Pod running a KVM instance in a container.
KubeVirt allows unique VM states and tracks and schedule Pods across nodes when migrating it.

https://kubevirt.io/

JupyterHub on top of Kubernetes

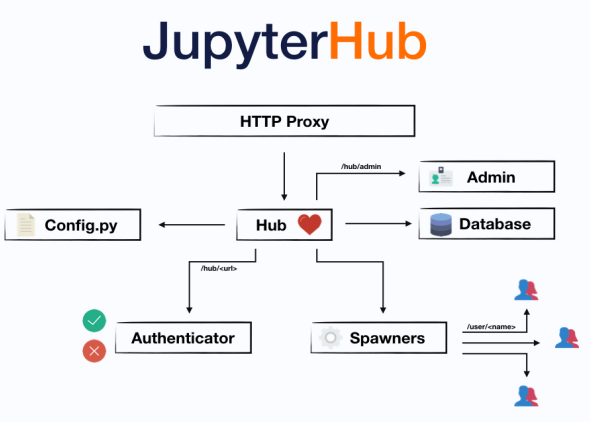
JupyterHub has been chosen as computing environment manager.

Users:

- Built-in image and resource selector
- Notebooks and terminal interface

Administrators:

- Easy to configure and maintain
- Quick implementations of add-ons



https://jupyterhub.readthedocs.io/en/stable/
https://z2jh.jupyter.org/en/stable/

Ongoing steps and Outlooks

Conclusions:

Ongoing steps

- Definition of a queuing system (*kueue*).
- Creation of a **dashboard** for the virtual machines
- **Uninuvola-GPU** addition of GPU nodes

Outlooks

- All materials will be made available *via* publications and repositories.