

RDO and Tripleo: exploit ansible to customize both the Undercloud and the Overcloud

A journey through the installation of a hyperconverged OpenStack cluster

Franca Debole Andrea Dell'Amico

CNR, ISTI InfraScience laboratory

franca.debole@isti.cnr.it andrea.dellamico@isti.cnr.it

CS3 2023 Barcellona

6-8 March 2023



Overview

- 1 Where we started
 - SaaS Oriented Research projects
 - An old infrastructure looking to the future
- 2 So, OpenStack
 - Why OpenStack
 - How Tripleo works
- 3 And now
 - Our setup
 - Our contribution
- 4 The outcome
 - Our OpenStack on premise
 - Headaches
- 5 The future
 - Federation? Still OpenStack?

Where we started
So, OpenStack
And now
The outcome
The future

SaaS Oriented Research projects

An old infrastructure looking to the future

D-Net and D4Science

D-Net is a research group that develops the core of the *OpenAIRE* <https://www.openaire.eu> services. Big data for real, we host the development infrastructure and services.

D4Science <https://www.d4science.org/> is a Data Infrastructure connecting +15.000 scientists in +50 countries, integrating +50 heterogeneous data providers, executing +55,000 models & algorithms/month. It hosts +175 Virtual Research Environments (VREs) to serve the biological, ecological, environmental, social mining, culture heritage, and statistical communities world-wide.



Where we started
So, OpenStack
And now
The outcome
The future

SaaS Oriented Research projects
An old infrastructure looking to the future

The needs of a large computing and storage infrastructure

The old computing and storage infrastructure is based on Xen and ATA o Ethernet:

- requires a lot of manual effort to setup machines and services for the users;
- the services made available to the institute and to the research partners require a significative amount of computing and storage resources.

OBJECTIVE: have research services available in a simple way, fast, secure eccetera.



Where we started
So, OpenStack
And now
The outcome
The future

SaaS Oriented Research projects
An old infrastructure looking to the future

What now: on premise OpenStack

We wanted flexibility, a set of APIs that could be also used out of our institute, and the possibility to federate our resources.



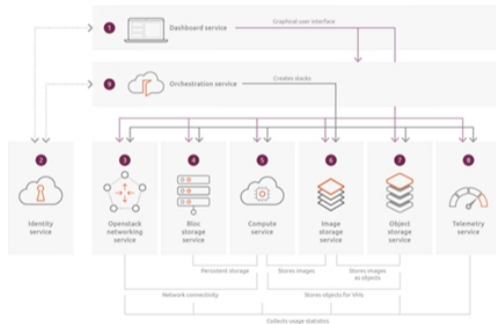
A lot of services: Rstudio, JupyterHub, TagME, Geoserver/Geonetwork.



Where we started
So, OpenStack
And now
The outcome
The future

Why OpenStack
How Tripleo works

So, OpenStack ...



We chose a CentOS base distribution for the baremetal and RDO
<https://www.rdoproject.org/> was a natural choice.



and TripleO

TripleO is a set of tools for the deployment and management of OpenStack which utilizes the *RDO* repositories.



- *TripleO* Stable versions: *Wallaby*, then *Zed*
- Uses a mix of Puppet and Ansible
- Ceph can be installed either using `ceph-Ansible` or `cephadm`
- Everything is containerized



Tripleo

TripleO requires a dedicated machine, called *undercloud*. It installs a all-in-one OpenStack into it, and then uses that OpenStack installation to deploy the main (*overcloud*) OpenStack.

The *undercloud* uses `ironic` to install the physical machines, and then configure the OpenStack services.

To do that, we must tell the *undercloud* which machines to install, what roles (Controller, Compute, Storage, HCI, etc.)

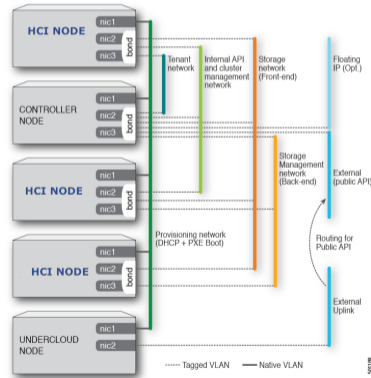


Where we started
So, OpenStack
And now
The outcome
The future

Our setup
Our contribution

The architecture

- Hyperconverged configuration (HCI), to simplify the hardware choices
- Separate networks for different needs:
 - VLANs, and 6 network interfaces on each server
 - Bonding on the Linux side, LACP on the switches

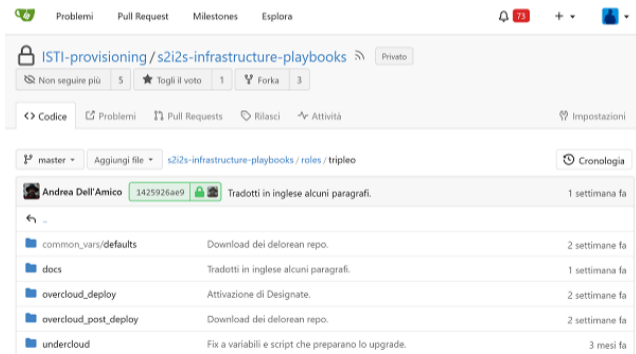


Where we started
So, OpenStack
And now
The outcome
The future

Our setup
Our contribution

Our contribution: Ansible to help on the installation of Undercloud and the Overcloud

We use *Ansible*, <https://www.ansible.com> to distribute the configuration files required by the undercloud, the overcloud, and for some tasks after the overcloud deployment completed successfully.



The screenshot shows a GitHub repository page for 'ISTI-provisioning / s2i2s-infrastructure-playbooks'. The repository is private and has 5 non-followers, 1 star, and 3 forks. The current branch is 'master'. The file list includes:

File Name	Description	Last Modified
..		
common_vars/defaults	Download dei delorean repo.	2 settimane fa
docs	Tradotti in inglese alcuni paragrafi.	1 settimana fa
overcloud_deploy	Attivazione di Designate.	2 settimane fa
overcloud_post_deploy	Download dei delorean repo.	2 settimane fa
undercloud	Fix a variabili e script che preparano lo upgrade.	3 mesi fa

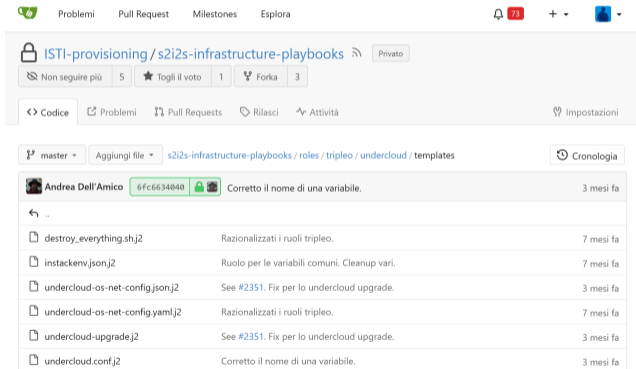
Where we started
So, OpenStack
And now
The outcome
The future

Our setup
Our contribution

Preparing for Undercloud deployment

The *undercloud* wants the list of the baremetal nodes to extract information from them

- Network interfaces configuration (other than the main one)
- Undercloud main configuration file
- List of the target nodes, with their BMC credentials



The screenshot shows a GitHub repository page for 'ISTI-provisioning/s2i2s-infrastructure-playbooks'. The page is in Italian and shows a commit history table. The table has columns for the commit author, commit ID, commit message, and the time since the commit. The commit history is as follows:

Author	Commit ID	Message	Time
Andrea Dell'Amico	6fc6634848	Corretto il nome di una variabile.	3 mesi fa
..
	destroy_everything.sh.j2	Razionalizzati i ruoli tripleo.	7 mesi fa
	instackenv.json.j2	Ruolo per le variabili comuni. Cleanup vari.	7 mesi fa
	undercloud-os-net-config.json.j2	See #2351. Fix per lo undercloud upgrade.	3 mesi fa
	undercloud-os-net-config.yaml.j2	Razionalizzati i ruoli tripleo.	7 mesi fa
	undercloud-upgrade.j2	See #2351. Fix per lo undercloud upgrade.	3 mesi fa
	undercloud.conf.j2	Corretto il nome di una variabile.	3 mesi fa

Preparing for Overcloud deployment

The *overcloud* deployment requires a lot of information, that we provide populating template files.

Controllers and Compute nodes

- Hosts roles
- Enabled services and their configuration
- Controller nodes IP addresses and hostnames
- Special images configuration (machines with GPUs, for example)

Ceph configuration

- Ceph installation technology (*ceph-ansible* vs *cephadm*)
- Ceph services
- List of storage nodes
- List of OSDs for each nodes
- HCI setups: reserve RAM and CPU



After the deploy

When the *overcloud deploy* completes successfully, a lot of work has still to be done [customization]

- Keystone OIDC: transparent association between users and projects
- Public endpoints (HAPROXY): certificates managed by Letsencrypt
- Creation of a basic set of flavours
- Upload a set of ready to use Linux distribution images
- Setup a backup procedure for both the undercloud and the controller nodes
- External Prometheus to gather metrics from the baremetal nodes



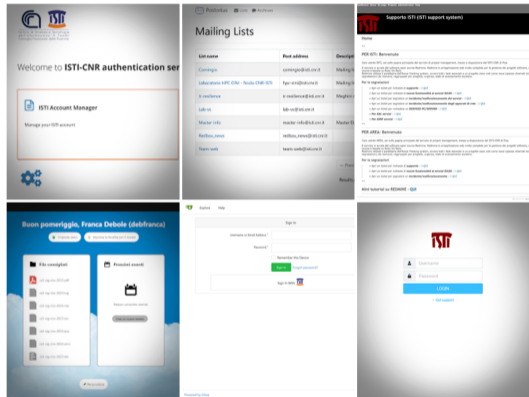
What went well \implies hypervisors

Hostname	Type	VCPUs (used)	VCPUs (total)	RAM (used)	RAM (total)
istl-cloud-computehciad8640-0.istl.cnr.it	OEMU	0	56	78GB	376GB
istl-cloud-computehciad8640-1.istl.cnr.it	OEMU	0	56	78GB	376GB
istl-cloud-computehciad8640-2.istl.cnr.it	OEMU	0	56	78GB	376GB
istl-cloud-computehciad8640-3.istl.cnr.it	OEMU	0	56	78GB	376GB
istl-cloud-computehciad8640-4.istl.cnr.it	OEMU	0	56	78GB	376GB
istl-cloud-computehciad8640-5.istl.cnr.it	OEMU	0	56	78GB	376GB
istl-cloud-computehciad8740gpu-0.istl.cnr.it	OEMU	0	96	91GB	754GB
istl-cloud-computehciad8740gpu-1.istl.cnr.it	OEMU	0	96	91GB	754GB
istl-cloud-computehciad8740gpu-2.istl.cnr.it	OEMU	0	96	91GB	754GB
istl-cloud-computehciad8740sd-0.istl.cnr.it	OEMU	0	72	88.5GB	314.3GB
istl-cloud-computehciad8740sd-1.istl.cnr.it	OEMU	0	72	88.5GB	314.3GB
istl-cloud-computehciad8740sd-2.istl.cnr.it	OEMU	0	72	88.5GB	314.3GB

Instance Name	Image Name	IP Address	Size	Key Pair	Status	Availability Zone	Task	Power State	Time since created	Actions
Enter your full name here	Genomics course setup v2	144.173.115.130	m1.xlarge	konstgenomicshuld	Shutoff	nova	None	Shut Down	3 days, 26 hours	Start Instance
MRS_Genomic_Medicine	unix_jyflco	144.173.115.114	m1.xlarge	krw1_genomicmedicine1	Active	nova	None	Running	3 weeks, 1 day	Create Snapshot
Vivien's instance	Unix workshop 2016	144.173.115.115	m1.xsmall	Vivien@biggpes	Shutoff	nova	None	Shut Down	1 month, 3 weeks	Start Instance
istl_unix_python	unix_python	192.168.0.8 Floating IP: 144.173.115.153	m1.medium	ocean	Active	nova	None	Running	2 months, 3 weeks	Create Snapshot
istl_unixperl	-	192.168.0.7 Floating IP: 144.173.115.173	m1.medium	ocean	Active	nova	None	Running	2 months, 3 weeks	Create Snapshot

Hypervisor run and going, dashboard to allow is quite easy self-service on baremetal.

What went well \implies data services/backbone services



Researchers and technicians enjoy it 😊

What went wrong

Almost *everything*

- Documentation: often obsolete or missing or confusing
- Code bugs, each time our configuration did not match the setups tested by the developers (example: bonding with vlans)
- The use of both `Puppet` and `Ansible` makes the troubleshooting more complicated than it should be
- The upgrade procedure only works under specific conditions (tested: Victoria \Rightarrow Wallaby)



Next steps

- Integrate a multitenant Kubernetes cluster (OKD <https://www.okd.io?>)
- Share
 - Publish our ansible roles and documentation (almost ready)
 - Federation of our cluster with other institutions (INFN, CERN 😊)?
 - Contribute back to the OpenStack community
 - Share/collaborate with the GÉANT community 😊

The End

Questions? Comments?