

Orchestrating complex deployments with the PaaS Orchestrator

Ibergrid 2019

Santiago de Compostela, Spain September 2019

MARICA ANTONACCI, ALESSANDRO COSTANTINI

marica.antonacci@ba.infn.it, alessandro.costantini@cnaf.infn.it

Istituto Nazionale di Fisica Nucleare (INFN)



DEEP-Hybrid-DataCloud is funded by the Horizon 2020 Framework Programm of the European union under grant agreement number 77743





- PaaS layer
- Deployment workflow
- Orchestrator architecture
- Usage scenarios
- APIs and tools



- The PaaS Orchestrator is based on the developments carried out during the INDIGO-DataCloud project
 - advanced features and important enhancements are being implemented in the framework of three projects: DEEP-Hybrid DataCloud, eXtreme-DataCloud and EOSC-Hub
- It allows to coordinate the **provisioning** of *virtualized* compute and storage resources on different Cloud Management Frameworks (like OpenStack, OpenNebula, AWS, etc.) and the **deployment** of dockerized services and jobs on Mesos clusters.
- The PaaS orchestrator features advanced **federation** and **scheduling** capabilities ensuring the transparent access to heterogeneous cloud environments and the selection of the best resource providers based on criteria like user's SLAs, services availability and data location

INDIGO Platform as a Service Layer





The deployment workflow



- The Orchestrator receives the deployment request (TOSCA template)
- The Orchestrator collects all the information needed to deploy the virtual infra/service/job consuming others PaaS µServices APIs:
 - **SLAM Service**: get the prioritized list of SLAs per user/group;
 - **Configuration Management DB**: get the the capabilities of the underlying IaaS platforms;
 - **Data Management Service**: get the status of the data files and storage resources needed by the service/application
 - *Monitoring Service*: get the IaaS services availability and their metrics;
 - **CloudProviderRanker Service** (Rule Engine): sort the list of sites on the basis of configurable rules;
- The orchestrator delegates the deployment to **IM**, **Mesos or QCG-Computing** based on the TOSCA template and the list of sites.
- Cross-site deployments are also possible.

PaaS Orchestrator Architecture





Scenario I: deployment of Virtual Infrastructure





Use case: frontend + elastic batch system





Scenario II: Deployment of managed services/jobs





- The PaaS orchestrator interacts with:
 - Marathon to deploy, monitor and scale
 Long-Running services, ensuring that
 they are always up and running.
 - Chronos to run user applications (jobs), taking care of fetching input data, handling dependencies among jobs, rescheduling failed jobs.
- Marathon and Chronos are two powerful frameworks that can be deployed on top of a Mesos Cluster.
- **Mesos** is able to manage cluster resources (cpu, mem) providing *isolation* and *sharing* across distributed applications (frameworks)

Use-case: execution of batch-like jobs





Data-aware Scheduling

0

Ο

Ο

Based on the data location \bigcirc

Use-case: deployment of a long-running service





Long-running service: generic template



tosca_definitions_version: tosca_simple_yaml_1_0 imports: - indigo_custom_types: https://raw.githubusercontent.com/indigo-dc/tosca-types/master/custom_types.yaml description: > TOSCA examples for specifying Marathon applications to enable the specification of long-running services in INDIGO. topology_template: inputs: cpus: type: float description: Amount of CPUs for this service required: yes default: 1.0 mem: type: scalar-unit.size description: Amount of Memory for this service required: yes default: 1 GB docker_image: type: string description: Docker image to be used to run the container application required: yes default: "" port: type: integer description: service port (exposed by the docker container) required: yes default: 8080 data_path: type: string description: container path for persistent data required: no default: "/data"

outputs: endpoint:

value: { concat: [{ get_attribute : [marathon-app, load_balancer_ips, 0] }, ':', { get_attribute : [docker_runtime, host, publish_ports, 0, target] }] }



Deep Learning prediction modules included in the DEEP-HybridDataCloud Open Catalog (https://marketplace.deep-hybrid-datacloud.eu/) can be deployed through a TOSCA template

- the DEEPaaS API is deployed as long running service on Mesos cluster
- the API can be accessed from the web browser and used to make predictions





• Elastic Galaxy Cluster

 a Galaxy portal is automatically deployed from TOSCA and configured to use a SLURM elastic cluster

• Elastic Mesos Cluster

- a complete HA Mesos cluster with Chronos/Marathon framework is automatically deployed from a TOSCA template
- Jupyter with K8s Cluster
- HTCondor cluster on Mesos (DODAS)
- Big-data Analysis Cluster (Spark on Mesos)
- Deep Learning training/prediction services (DEEPaaS)



Advanced features

Deployment retry strategy



- The Orchestrator implements a trial-and-error mechanism that allows to re-schedule the deployment on the next available cloud provider from the list of candidate sites.
- Example: deployment fails because of exceeding the quota on the chosen site





The problem: user applications need access to sensitive data (e.g. password for dbs, storage service credentials, etc.)

The solution: introduce a secrets manager, **Hashicorp Vault**, in the PaaS architecture to store the sensitive information safely.

The Orchestrator writes the secrets in Vault on behalf of the user, then the Marathon plugin retrieves the secrets from Vault and makes them available to the application container (via environment variables)





GPU scheduling and Integration with HPC



- The PaaS Orchestrator supports the deployment of virtual machines and containers that need to access specialised hardware devices, namely GPUs, to provide the processing power required by tasks like Machine Learning algorithms
 - the GPU requirements (num, vendor, model) can be specified in the TOSCA template
 - the Orchestrator automatically selects the sites/services that provide the needed capabilities (flavors, gpu support)
- The Orchestrator includes a plugin for **submitting jobs to HPC** facilities
 - exploits the QCG-Computing service (PSNC) that exposes REST APIs to submit jobs to the underlying batch systems

Support for hybrid deployments of elastic clusters



Scenario I:

• exploits L2 network provided by the sites

Scenario II:

 dedicated private nets are automatically provisioned





The PaaS Orchestrator has been enhanced to:

• schedule the processing jobs near the data

The PaaS Orchestrator is being extended in order to:

- Integrate a data management policy engine (QoS and Data Life Cycle)
 - move data between distributed storages
 - \circ $\,$ specify different QoS for replicas
- Support workflows for data pre-processing at ingestion





• Create a deployment:

- POST request to /deployments parameters:
 - template: string containing a TOSCA YAML-formatted template
 - parameters: the input parameters of the deployment (map of strings)

• Get deployment details:

- GET request to /deployments:
 - curl 'http://localhost:8080/deployments/<uuid>'

• Delete deployment:

- DELETE request
 - curl 'http://localhost:8080/deployments/<uuid>'
- Documentation: http://indigo-dc.github.io/orchestrator/restdocs/#overview

Orchent: The Orchestrator CLI



export ORCHENT_TOKEN=<your access token>
export ORCHENT_URL=<orchestrator_url>

usage: orchent <command> [<args> ...]

Commands:

help [<command>...]
Show help.

depls

list all deployments

depshow <uuid>

show a specific deployment

depcreate [<flags>] <template> <parameter>
 create a new deployment

```
depupdate [<flags>] <uuid> <template> <parameter>
update the given deployment
```

deptemplate <uuid>
 show the template of the given deployment

depdel <uuid>
 delete a given deployment

Installation guide: https://indigo-dc.gitbooks.io/orchent/content/admin.html User guide: https://indigo-dc.gitbooks.io/orchent/content/user.html

The Orchestrator Dashboard



Simple graphical user interface for the Orchestrator

PaaS	Orchestrator Dashboard Deployr	ments SLAs Settings	Marica Antonacci 🝷	Authentication via INDIGO IAM
	Search			PaaS Orchestrator Dashboard Deployments SLAs Settings 🍙 Marica Antonacci
1 Select deploymentype	t	DisVis	Galaxy Elastic Cluster	2 Configure input parameters Description: TOSCA template for deploying an instance for Kepler
	Dariah Data Repository	Kepler TOSCA template for deploying an instance for Kepler Configure	Galaxy Elastic Cidster - Enxir Italy Galaxy	Input Values Advanced number_cpus 1 number of cpus required for the instance
	Galaxy - Elixir Italy Galaxy	Mesos Cluster	Kepler batch job	memory_size 1 GB ram memory required for the instance 3 Submit © Cancel Submit deployment request

List your deployments



PaaS Orchestrator Dashboard Deployments	SLAs Settings					🍙 Marica Antonacci 👻
My deployments					P Refresh	+ New deployment
Show 10 \$ entries					Search	:
Deployment uuid	∿ Status	₩	Creation time	∕∿	Deployed At	$^{∧}$ Actions $^{∧}$
11e9ddcc-74e1-e98e-b691-067adb74d617	CREATE_COMPL	ETE	2019-09-23T06:36+0000		provider-RECAS-BARI	🛅 Delete 🔽
11e9dc71-fe4a-a242-b691-067adb74d617	CREATE_COMPL	ETE	2019-09-21T13:16+0000		provider-RECAS-BARI	■ Details Q Show template
11e9d915-526c-bfb4-8967-067adb74d617	CREATE_COMPL	ETE	2019-09-17T06:35+0000		provider-RECAS-BARI	Log
11e9aeaa-8fea-48c9-83c8-067adb74d617	CREATE_COMPL	ETE	2019-07-25T07:05+0000		provider-RECAS-BARI	🛅 Delete 🔻
11e9ac63-6f8e-095d-9841-067adb74d617	CREATE_COMPL	ETE	2019-07-22T09:31+0000		provider-RECAS-BARI	🛅 Delete 🔻
11e9a7b2-172a-5c41-8c92-067adb74d617	DELETE_FAILED		2019-07-16T10:11+0000		provider-IFCA-LCG2	🛅 Delete 🔻
Showing 1 to 6 of 6 entries						Previous 1 Next

Get deployment details and outputs



🍙 Marica Antonacci 🝷 PaaS Orchestrator Dashboard Deployments SLAs Settings My deployments + New deployment C Refresh Deployment details × Show 10
entries Search: Overview Outputs 4 ΛV Deployment uuid **Deployed At** ∿∿ Actions UUID: 11e9ac63-6f8e-095d-9841-067adb74d617 11e9ddcc-74e1-e98e-b691-067adb74d617 provider-RECAS-BARI 🛅 Delete 💌 STATUS: CREATE_COMPLETE CREATED AT: 2019-07-22T09:31+0000 Details provider-RECAS-BARI 11e9dc71-fe4a-a242-b691-067adb74d617 UPDATED AT: 2019-07-22T09:32+0000 Q Show template 🖹 Log DEPLOYED AT: provider-RECAS-BARI 11e9d915-526c-bfb4-8967-067adb74d617 provider-RECAS-BARI Close 11e9aeaa-8fea-48c9-83c8-067adb74d617 provider-RECAS-BARI Delete CREATE_COMPLETE 11e9ac63-6f8e-095d-9841-067adb74d617 2019-07-22T09:31+0000 provider-RECAS-BARI 🛅 Delete < 11e9a7b2-172a-5c41-8c92-067adb74d617 DELETE_FAILED 2019-07-16T10:11+0000 provider-IFCA-LCG2 面 Delete Previous Next Showing 1 to 6 of 6 entries

View the deployment log



PaaS Orchestrator Dashboard Deployments SLAs Settings

🍙 Marica Antonacci 🝷

Deployment log

♂ Refresh ← Back

2019-09-23 06:37:05.617885: Select master VM 2019-09-23 06:37:05.617607: Wait master VM to boot 2019-09-23 06:37:50.635409: Vait master VM to have the SSH active. 2019-09-23 06:37:50.835489: Creating and copying Ansible playbook files 2019-09-23 06:37:51.765238: Galaxy role indigo-dc.zabbix-agent detected setting to install. 2019-09-23 06:37:51.765432: Performing preliminary steps to configure Ansible. 2019-09-23 06:37:52.395987: Configure Ansible in the master VM. 2019-09-23 06:42:08.735649: Ansible successfully configured in the master VM. 2019-09-23 06:42:16.319374: Copying YAML, hosts and inventory files. VM 0: Contextualization agent output processed successfullyGenerate and copy the ssh key

Sleeping 0 secs.

Launch task: wait_all_ssh Waiting SSH access to VM: 90.147.75.27 Testing SSH access to VM: 172.30.99.81:22 Remote access to VM: 90.147.75.27 Open! Changing the IP 172.30.99.81 for 90.147.75.27 in config files. Task wait_all_ssh finished successfully Process finished Contextualization agent output processed successfullyGenerate and copy the ssh key Sleeping 0 secs. Launch task: basic Waiting SSH access to VM: 90.147.75.27 Testing SSH access to VM: 172.30.99.81:22 Remote access to VM: 90.147.75.27 Open! Requiretty successfully removed Install indigo-dc.zabbix-agent with ansible-galaxy. Galaxy depencies file: [{src: indigo-dc.zabbix-agent}]

Call Ansible

Useful for debugging purposes.

You can also download the log file clicking on the 'download' button at the end of the page.





• **TOSCA Templates**

- Use-cases templates: <u>https://github.com/indigo-dc/tosca-templates</u>
- Example templates: <u>https://github.com/indigo-dc/tosca-</u> <u>types/tree/master/examples</u>

• Ansible Roles

• Ansible Galaxy: <u>https://galaxy.ansible.com/indigo-dc/</u>

• Docker images

• Docker hub:

https://hub.docker.com/u/indigodatacloudapps/dashboard/



Building and Deploying Complex Applications with Alien4Cloud

Ibergrid 2019 Santiago de Compostela, Spain September 2019

> Marica Antonacci, Alessandro Costantini on behalf

Andy S. Alic, Miguel Caballer, Germán Moltó asalic@upv.es, micafer1@upv.es, gmolto@dsic.upv.es

Universitat Politècnica de València (UPV)





Building/composing TOSCA Topologie

- Web app for Composition/management of TOSCA topologies
 - Scratch or Existing Template
- Open Source under Apache 2.0
 - o <u>https://github.com/alien4cloud/alien4cloud</u>
- Very low barrier for non-experts
 - Topologies with minimal TOSCA knowledge
 - Easy deployment/interaction with orchestrator deploying/managing the actual infrastructure
- Java REST backend (SpringBoot) / HTML5 frontend (AngularJS)
- Extensible via Plugins
 - Easy to add new orchestrators to create topologies built/composed via the UI



Building/composing TOSCA Topologie

- Extension of the Existing Version
 - DEEP fork @ <u>https://github.com/indigo-dc/alien4cloud</u>
 - Improvements like outputs handling, TOSCA functions parsing,
- Implementation of a plugin to connect to the IndigoDC Orchestrator
 - Orchestrator @ <u>https://github.com/indigo-dc/orchestrator</u>
 - O Plugin @ <u>https://github.com/indigo-dc/alien4cloud-deep</u>
 - Repo with Dockerized A4C, plugin, TOSCA normative types, and DEEP custom types
- Support for Oauth2
 - Via A4C dependency @ <u>https://github.com/indigo-dc/spring-social-oidc</u>

Building/composing Alien4Cloud UI - Configure Orchestrator

	🖵 Applications 🛛 🚳 Ca	atalog 🎤 Administration		🔒 admin 进 🗸
i ormati				▲ Unlock
on	Global configurat	tion		
ploym ents	Naming policy of depl (application.name + '-'	loyment + environment.name).replaceA	ll('[^\w\]', '_')	0
tion	Driver configurat	tion		
ation(s)	client	tld none	•	
ifacts	clientSecr	ret none		
	tokenEndpoi	int https://iam.deep-hybrid-	dat	
	clientScope	es openid profile email offl	n	
	orchestratorEndpoi	int https://deep-paas.cloud	ba	
	iamHo	https://iam.deep-hybrid-	dat	
	us	ser none		

Building/composing Alien4Cloud UI - Create New Users



Applications 🗞 Catalog 🕫	Administration			🛔 admin 🚔 🗸
Users Groups	New User			
New User	Username	Jsername		
	Password	Password Confirm	iroups	Roles
a	- First Name	First Name		• ADMIN &
	Last Name	Last Name		ê
	Email	Email		
	Roles	ADMIN APPLICATIONS_MANAGER ARCHITECT COMPONENTS_BROWSER COMPONENTS_MANAGER Create Create Cancel		

Building/composing Alien4Cloud UI - Components



ب	Applications 🚳 Catalog 🎤	Administration				🖰 admir	
ponents							
Manage	archives Browse components	Browse topologies Brows	se policies				
Search						Q	
	PersistentVo PersistentVolumeClaimVolumeSource represents a reference to a PersistentVolumeClaim in the same namespace. More info: https://kubernetes.io/docs/concepts	PersistentVo An extension of PersistentVolumeClaimSource that allows to specify a storageClass. If no volume that match this claim is found in the PersistentVolume pool. a volume	VolumeBase 2	DockerExtVol 3	BlockStorage @ The TOSCA BlockStorage node currently represents a server-local block storage device (i.e., not shared) offering evenly sized blocks of data from which raw storage volumes can be		
	rg.alien4cloud.kubernetes.api	org.alien4cloud.kubernetes.api	e org.alien4cloud.kubernetes.api	docker-types	tosca-normative-types		
	2.0.0 -	2.0.0 -	2.0.0 -	2.0.0 -	1.0.0-ALIEN20 -		
	Compute (a) The TOSCA Compute node represents one or more real or virtual processors of software applications or services along with other essential local resources. Collectively, the resources the	Application (2) The TOSCA Container Application node represents an application that requires Container-level virtualization technology.	کر Node کر Docker	DockerContai The TOSCA Container Application Docker node represents an application running in a Docker container. Properties defined in the node will be interpreted as the entrypoint's	Runtime @ The TOSCA Container Runtime node represents operating system-level virtualization technology used to run multiple application services on a single Compute host.		
	tosca-normative-types	tosca-normative-types	indigo-types	docker-types	tosca-normative-types		
	1.0.0-ALIEN20 -	1.0.0-ALIEN20 -	1.0.0 •	2.0.0 •	1.0.0-ALIEN20 -		
	DBMS @ The TOSCA DBMS node represents a typical relational, SQL	MySQL MusqL	Database a The TOSCA Database node represents a logical database that	MySQL	LoadBalancer @ The TOSCA Load Balancer node represents logical		

Building/composing Alien4Cloud UI - New TOSCA Topology



9 -	Applications 🙈 Catalog 🔑 Administration	nin 🕒 🗸
	New Application	
Sea		T
No a	Name	
	TestApp	
	Archive name (Id)	
	TestApp	
I	Description This is a test application	
	Initialize topology from Topology template Scratch	
	Creates a single new empty topology version.	
	Create	

Building/composing Alien4Cloud UI - GUI composition



	environment/477cafbd-702e-4ce4-9b5f-d1783902043a/archiv ▷♡
Applications 🚓 Catalog 🗲 Administration	🕹 admin 👙
TestApp Environment Topology Editor (0.1.0-SNAPSH01) Image: Save Download Dundo C Redo Premote Pull Push	
Topology Workflow	Compute Type: Compute ♥ Properties
Inputs and variables Kepler Kepler	Image: server orchestrator.clou Image: server Image: server Image: server 10051 Image: server Image: server
Archive content	Image: server Linux 668c875e- ℃ Image: server 9a39-4d Image: server
opology alidatio n	
History	Ax ➡ min_instances 1 ♪
	A ⊕ default_instan 1 ♂ 0

Building/composing Alien4Cloud UI - Text composition



÷	🖵 Applications 🛛 🚓 Catalog 🖌	Administration		🛎 admin 🚇 🗸
TestAp	op Environment Topology	Editor (0.1.0-SNAPSH	OT)	
🖺 Sa	ave 🛓 Download 🏷 Undo 🧲 Redo 🦹 Rer	mote I Pull 1 Push	< > /	、
Topology Workflow s Inputs and Variables Archive content Topology validatio n	■.git ☐ readme.txt ☐ topology.yml	C 1 1 2 3 4 5 6 7 7 8 9 9 10 11 12 3 4 5 6 7 7 8 9 9 9 10 11 12 13 3 14 15 16 16 17 7 8 9 9 9 9 10 11 12 3 3 4 4 5 6 6 7 7 8 9 9 9 10 11 12 13 14 15 15 16 16 17 18 19 19 19 201 201 201 201 201 201 201 201	Save topology.yml tosca_definitions_version: alien_dsl_2_0_0 metadata: template_name: TestApp template_aratmo: admin description: "" imports: tosca-normative-types:1.0.0-ALIEN20 indigo-types:1.0.0 topology_template: node_templates: Compute: metadata: adc_edit_x: 208 adc_edit_x: "189" type: tosca.nodes.indigo.Compute properties: zabbix_server_metadata: "Linux 668c875e-9a39-4dc0-a710-17c41376c1e capabilities: capabilities:	Θ"

Building/composing Alien4Cloud deploy



Environment	Info		
oyment init			
are next deployment 0	0.1.0-SNAPSHOT Manage current deplo	byment 0.1.0-SNAPSHOT Deployment His	story
Undeploy			
Reminder of curre	nt deployment		
Application :	TestApp	Version :	0.1.0-SNAPSHOT
Environment :	Environment	Environment type :	OTHER
🗞 Ressources insta	lled		
Compute	(tosca.nodes.indigo.Co	mpute)	
Kepler	(tosca.nodes.indigo.Ke	pler)	
Output properties			
No output properties	s or attributes		
Deployed at	Jun 18, 2018 11:23:32 PM		
Duration	a few seconds		



Thank You

All names, logos, and brands are property of their respective owners in the European Union member states and/or other countries. All company, product and service names used on this presentation can be considered fair use for teaching purposes.