

PIETRO MANDREOLI

Consortium
GARR

THE ITALIAN
EDUCATION
& RESEARCH
NETWORK

BORSISTI DAY 2018



Portabilità su GARR Cloud di **Laniakea** : un servizio Galaxy on-demand basato su tecnologia INDIGO- Datacloud

GIORNATA DI INCONTRO
BORSE DI STUDIO GARR
"ORIO CARLINI"
6 DICEMBRE 2018 ROMA

Roma 6 dicembre
Borsisti day 2018



Galaxy : vantaggi, caratteristiche, componenti

- Workflow manager bioinformatico
- Consente di facilitare l'interazione con gli strumenti computazionali e la gestione ordinata di grandi quantità di dati biologici

VANTAGGI:	CARATTERISTICHE:	COMPONENTI:
<ul style="list-style-type: none">• Favorisce la riproducibilità• Consente la condivisione dei dati e dei workflow• Permette il lavoro di gruppo	<ul style="list-style-type: none">• Ricca documentazione• Ampia comunità di utilizzatori• Semplice da utilizzare• Compatibilità con *nix system	<ul style="list-style-type: none">• Piattaforma software• Galaxy Tool Shed



GIORNATA DI INCONTRO BORSE DI STUDIO GARR "ORIO CARLINI" GIOVEDI' 6 DICEMBRE 2018 - ROMA



The screenshot shows a web browser window displaying the Galaxy web interface. The browser's address bar shows the URL 90.147.75.87/galaxy/. The page title is "Galaxy / covacs". The main content area displays a "Hello, Galaxy is running!" message with two buttons: "Configuring Galaxy »" and "Installing Tools »". Below this, there is a section for "Take an interactive tour:" with buttons for "Galaxy UI", "History", and "Scratchbook". The page also includes a "History" sidebar on the right, which is currently empty. The left sidebar contains a list of tool categories such as "Get Data", "Send Data", "Collection Operations", "Text Manipulation", "Convert Formats", "Filter and Sort", "Join, Subtract and Group", "Fetch Alignments/Sequences", "Operate on Genomic Intervals", "Statistics", "Graph/Display Data", "Phenotype Association", "BED Tools", "pre-alignment", "DNA-alignment", "picard-tools", "vcf-tools", "COVACS", and "Workflows".



Galaxy : problematiche

- Sistema complesso
- Installazione complessa

AMBIENTE DI PRODUZIONE GALAXY:

- SO → CentOS 7
- Database engine → PostgreSQL
- Web server → NGINX
- Interfaccia server/servizi web → uWSGI
- FTP server → Proftpd





GALAXY: installazione su infrastrutture differenti, vantaggi e svantaggi.

SERVER PUBBLICI	SERVER PRIVATO	CLOUD COMMERCIALE
VANTAGGI: <ul style="list-style-type: none">• Facile accesso	VANTAGGI: <ul style="list-style-type: none">• Facile personalizzazione• Sicurezza dei dati• Risorse consistenti	VANTAGGI: <ul style="list-style-type: none">• Infrastruttura hardware mantenuta dal provider• Facile personalizzazione• Risorse consistenti
SVANTAGGI: <ul style="list-style-type: none">• Risorse limitate• Difficile personalizzazione• Problemi sicurezza dati sensibili	SVANTAGGI: <ul style="list-style-type: none">• Mantenimento infrastruttura hardware	SVANTAGGI: <ul style="list-style-type: none">• Costoso• Problemi di sicurezza sui dati sensibili



Laniakea: Galaxy as a cloud service

- Servizio per la creazione automatica di macchine virtuali contenenti un ambiente di produzione Galaxy
- Utilizza i componenti del catalogo software di INDIGO-DataCloud

VANTAGGI:

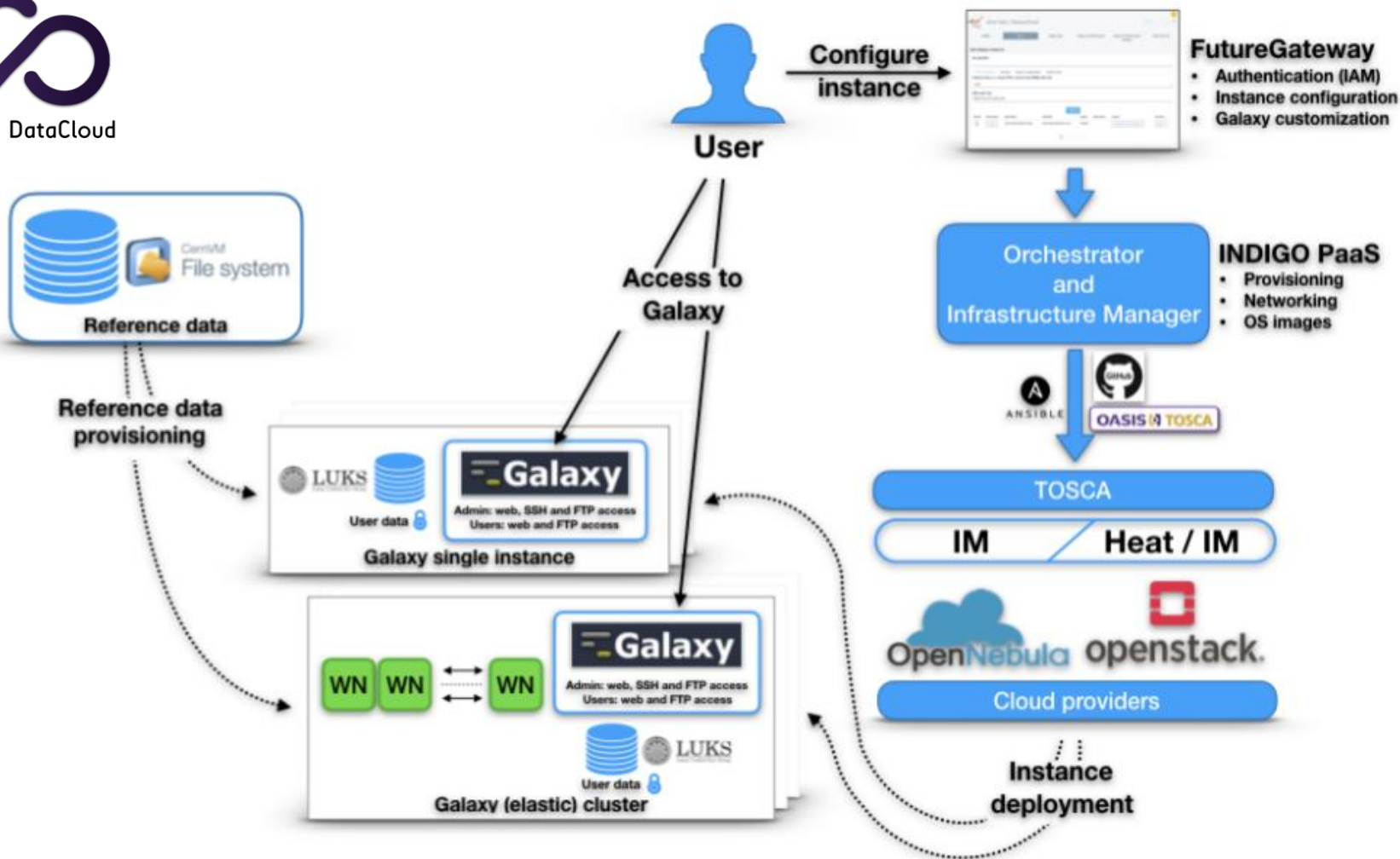
- Permette, a chi non dispone di risorse IT di accedere a questo strumento
- Accesso a elevata capacità computazionale



ARCHITETTURA DEL SERVIZIO



INDIGO - DataCloud





FutureGateway: opzioni

Virtual hardware

Galaxy

Instance flavor, i.e. CPUs, memory size (RAM), root disk size

medium

SSH public key

Paste here your public key

Enable encryption

plain

Volume storage size

50 GB

Virtual hardware

Galaxy

Galaxy version (release_18.05 recommended)

[release_18.05](#)

Instance description (Galaxy brand)

ELIXIR-ITALY

Galaxy administrator mail address

admin@elixir-italy.org

Galaxy flavor

galaxy-CoVaCS

Reference data repository

elixir-italy.galaxy.refdata



LUKS
Linux Unified Key Setup

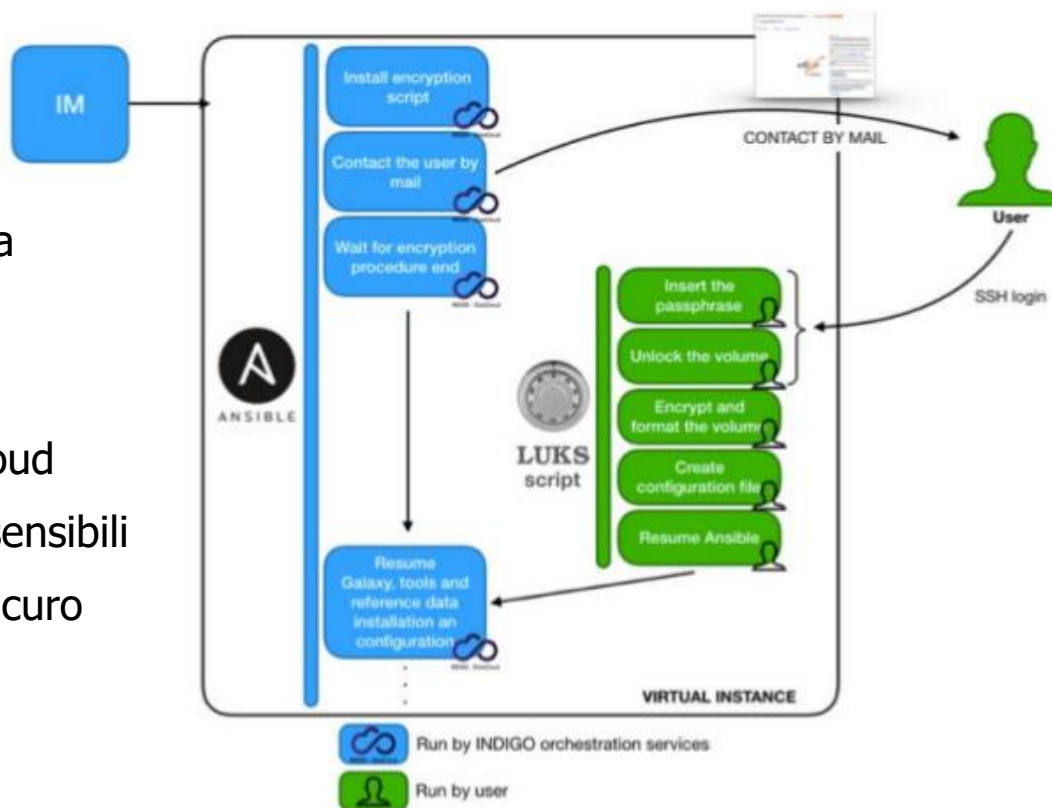
: Criptazione dello storage

COMPONENTI:

- Linux Unified Key Setup

CARATTERISTICHE:

- Criptazione dei dati in scrittura
- Decriptazione file in lettura
- I dati non sono accessibili dall'amministratore del sito cloud
- Possibilità di processare dati sensibili genetici umani in modo più sicuro





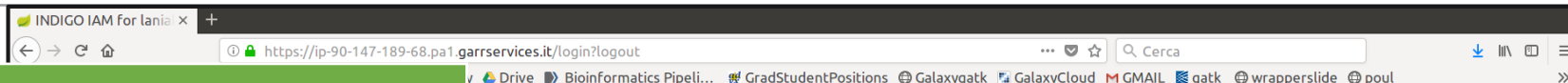
IAM: Identity and Access Management Service e Openid connect

IAM

Servizio di autenticazione e autorizzazione (AAI)

FUNZIONI:

- Autenticazione
- Gestione delle sessioni
- Iscrizione
- Gestione attributi e identità
- OpenID connect





Docker e Ansible



Docker

Software che permette l'automazione di applicazioni all'interno di container

CARATTERISTICHE:

- Coesistenza di più container sulla stessa istanza
- Servizio Docker Hub, repository per la condivisione di applicazioni "dockerizzate"

Ansible

Software che permette di automatizzare le procedure di configurazione e gestione sui sistemi unix-like

CARATTERISTICHE:

- Non necessita di installazione di agent sugli host
- Utilizzo della sintassi YAML
- Possibilità di utilizzo dei ruoli
- Servizio Ansible Galaxy per la condivisione e il riutilizzo dei ruoli



IAM deployment parte 1

Deployment su tenant di GARR cloud di 2 virtual machine:

- Vm_master
- Vm_IAM

Installazione e configurazione su vm_master dei software:

- Ansible

Creazione e ottenimento :

- Certificati Let's Encrypt
- Id client e id secret di Google



IAM deployment parte 2

vm_master:

- Importazione del ruolo per l'installazione di Docker sull'host
- Importazione e modifica del ruolo per l'installazione di IAM

Successivamente al deployment:

- Test delle funzioni base del servizio IAM: iscrizione, autenticazione di un utente di test.



Laniakea on RECAS : flavour e test

- Test del servizio su RECAS.
- Development e test dei Galaxy flavour:
 - GDC somatic Variants
 - CoVacs (Consensus Variant Calling System)
- Sottomissione dell'articolo in preprint.

Laniakea: an open solution to provide Galaxy "on-demand" instances over heterogeneous cloud infrastructures.

Marco Antonio Tangaro¹, Giacinto Donvito², Marica Antonacci², Matteo Chiara³, Pietro Mandreoli³, Graziano Pesole^{1,4}, Federico Zambelli^{1,3}

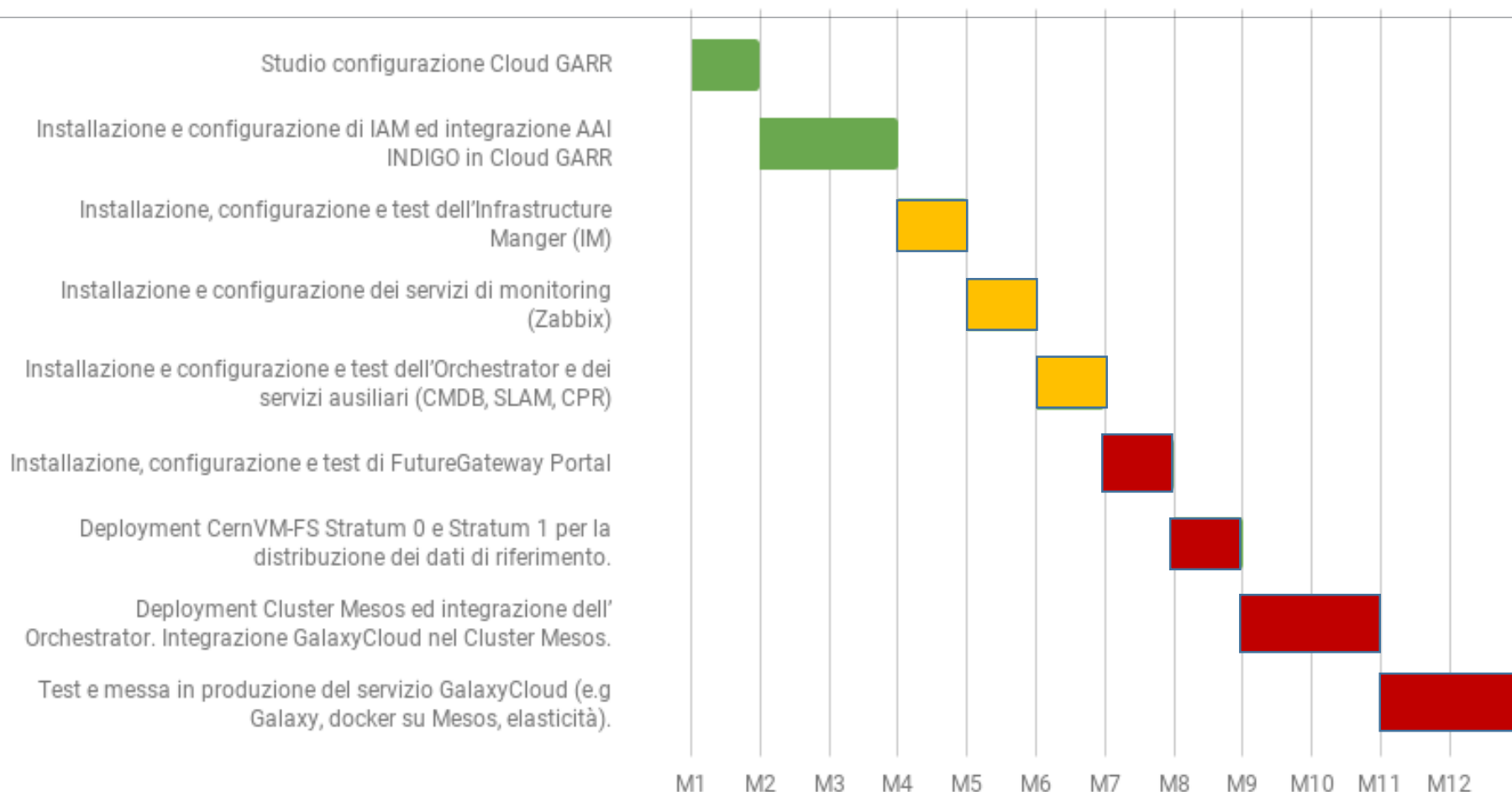


<https://github.com/Laniakea-elixir-it>

<https://github.com/pmandreoli>



Timeline



Grazie