IT behind ESA Datalabs

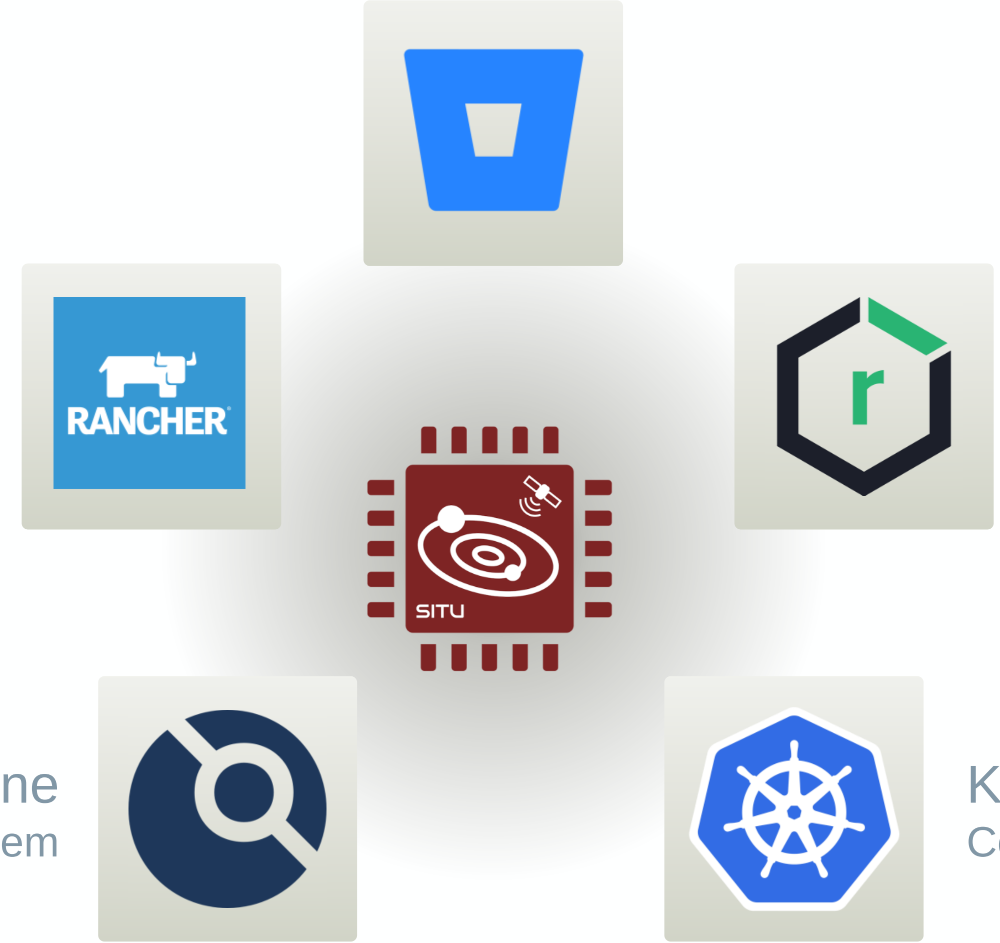
Overview

IT Platforms: Container Management, Source Code Management, Artifacts Management, CI/CD

IT Infrastructures: Computing, Storage, Networking

IT Security: Context, Security Measures

IT Platforms



S2E2 Bitbucket
Code & configuration repository

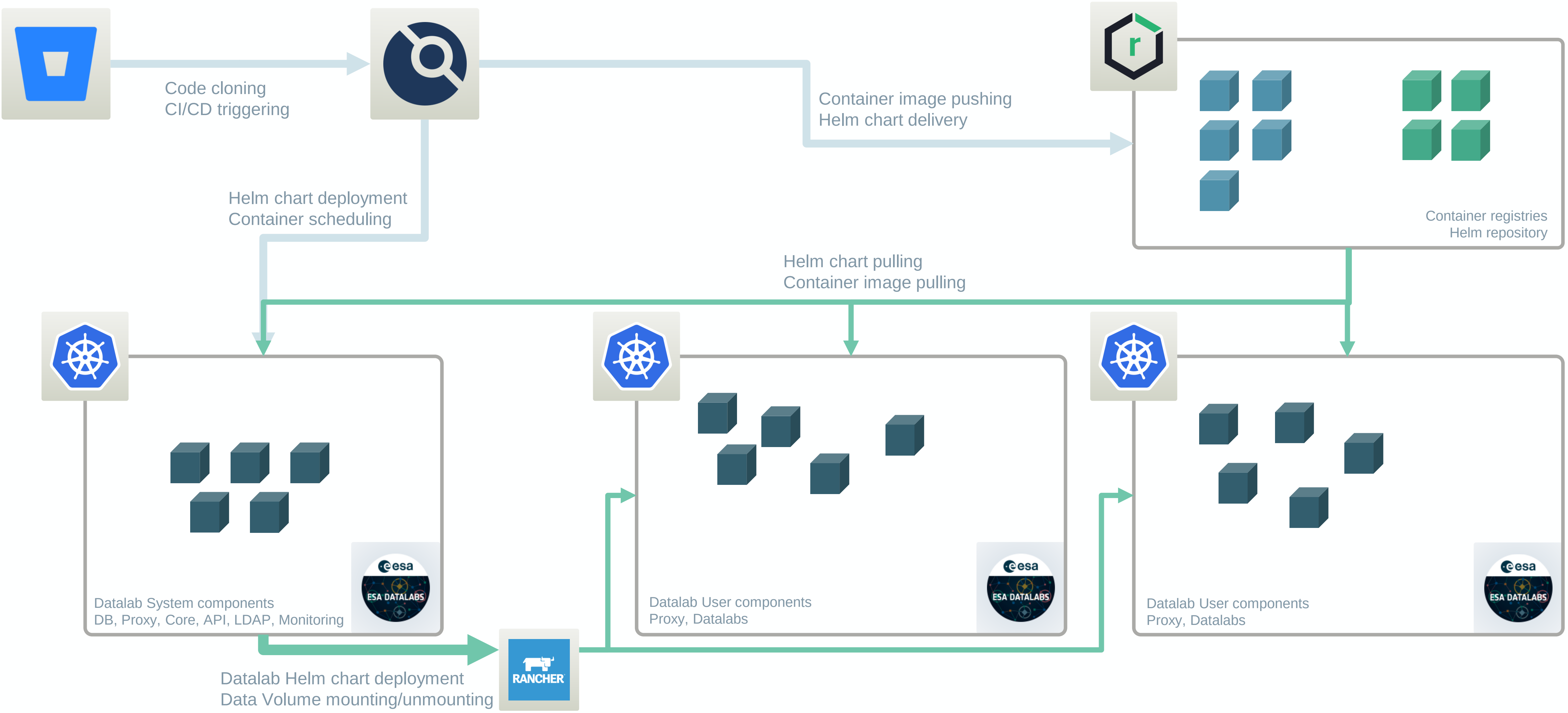
Rancher
Kubernetes management System

Sonatype Nexus
Container Registry & Package repository

Drone
CI/CD management System

Kubernetes
Container Orchestrator

IT Platforms


Code cloning
CI/CD triggering

Container image pushing
Helm chart delivery

Helm chart deployment
Container scheduling

Container registries
Helm repository

Helm chart pulling
Container image pulling

Datalab System components
DB, Proxy, Core, API, LDAP, Monitoring

Datalab User components
Proxy, Datalabs

Datalab User components
Proxy, Datalabs

Datalab Helm chart deployment
Data Volume mounting/unmounting

IT Infrastructures

In ESA Datalabs, we are running several environments (DEV, E2E, PRE, PRO):

Computing:

- Virtualized environments (DEV, E2E, PRE):
  - OS: CentOS 7
  - ~ 88 vCPUs: Intel(R) Xeon(R) CPU E5-2680 v3 @ 2.50GHz
  - Memory (type, amount): 256 GB or RAM
- Production environment:
  - OS: Red Hat Enterprise Linux 8.6
  - 2 main nodes with 48 cores on 2 Intel(R) Xeon(R) Gold 6226 CPU @ 2.70GHz, each.
  - Memory on each node: 512 GB or RAM
- AI environment:
  - OS: Ubuntu 18.04.6 LTS
  - CPUs: 1 node with 2 sockets of 64 Cores AMD Rome (256 cores with Hyperthreading).
  - GPUs: 8 A100 with 40 GB of VRAM (1024 Tensor cores), each.
  - Memory: 1024 GB of RAM

Networking:

- Service network: DMZ specific network for containers, with 10 Gbps of bandwidth.
- Storage network: Internal network for Science Storage traffic, with 10 Gbps of bandwidth.
- ESA Datalabs network: Specific ESA Datalabs setup with Load Balancing in different levels, exposing the services to Internet, with 10 Gbps of bandwidth.

Storage:

- NAS Space shared: 521 TBs, from Archives of, e.g., XMM, Integral, JWST, Planck, Hubble, Solar Orbiter missions and ESA Datalabs persistent areas.
- IOPS: 100.000 in one volume on average, but we can have until 7M IOPS with SSD or 500.000 IOPS with HDD
- NetApp ONTAP 9.8P5

2.5 TB RAM
440 Cores
8 GPUs
320 GB VRAM

3 networks
10 Gbps

0.5 PB
500K IOPS

IT Infrastructures

Slide 2 on infrastructure

IT Security – Current status

Background

- SCI-S missions/projects starting to use containerization and cloud native application in their developments.
- It creates significant challenges in securing these containerized applications and their entire lifecycle.

Activities performed

- Created a SCI-S container security policy (Currently under Review).
- Performed a trade-off analysis of container security solutions (Currently under Review).

Next activities (coming soon)

- Procure the selected SCI-S container security solution.
- Integrate this SCI-S container security solution into the SCI-S containerized and cloud native applications progressively.

IT Security- Security Measures

Container Registry

CI/CD

- Scanning:
  - Container Image vulnerabilities.
  - Configuration defects.
  - Embedded malware.
  - Embedded clear text secrets.
  - Untrusted container images.

Container runtime

- Container runtime vulnerability scan.
- Ingress/Egress container network visibility.
- Insecure container runtime configurations.
- Rogue containers detection.
- Admission controller.

Security Integration *(coming soon)*

Questions & Comments