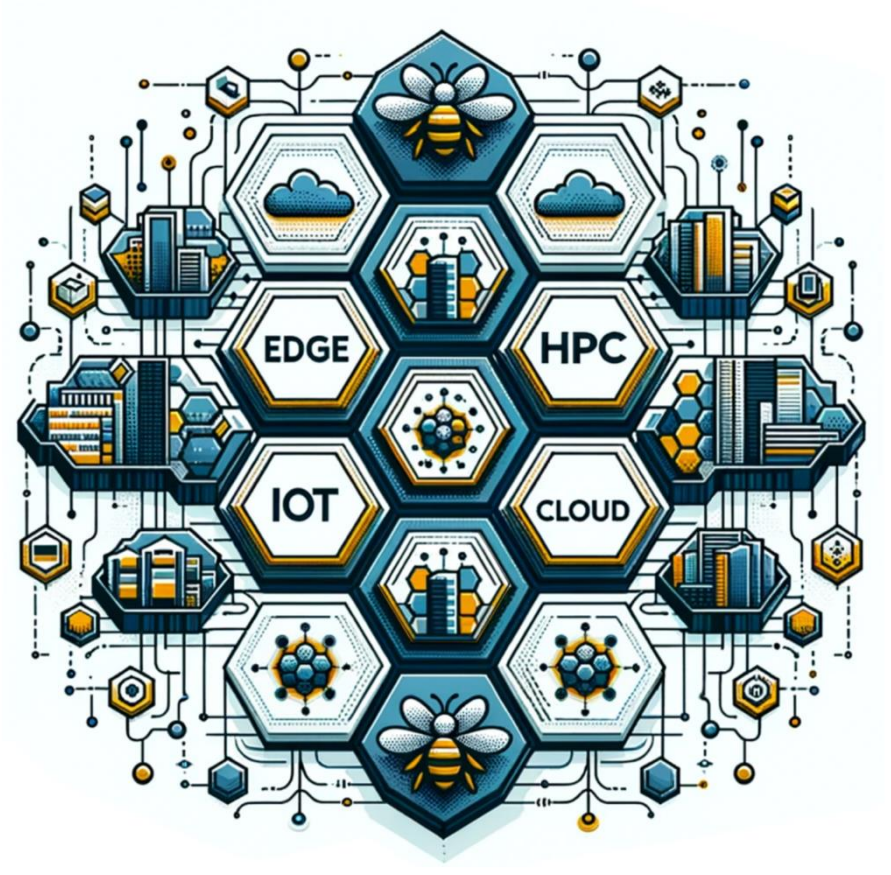


ColonyOS Tutorial

Johan Kristiansson

ColonyOS

Unleashing Computational Power Everywhere!

What is ColonyOS?

In a rapidly digitalizing world, seamless interoperability and robust large-scale computing aren't just luxuries—*they're essential.* Yet, as we shift towards decentralized and diverse computing landscapes, developing cross-platform applications becomes a daunting task. Imagine a world where AI workloads can easily be developed and run seamlessly across any platform, including Cloud, Edge, and HPC.

Welcome to ColonyOS!

ColonyOS is an open-source research project developed by RISE AB, and is used by ENCCS to foster greater High-Performance Computing (HPC) adoption. It is also used by RockSigma AB to implement a compute engine designed for seismic processing in underground mines. RockSigma AB has contributed to the development of ColonyOS.

Read more | Getting started | Contact us

Use Cases

Distributed Compute Engines

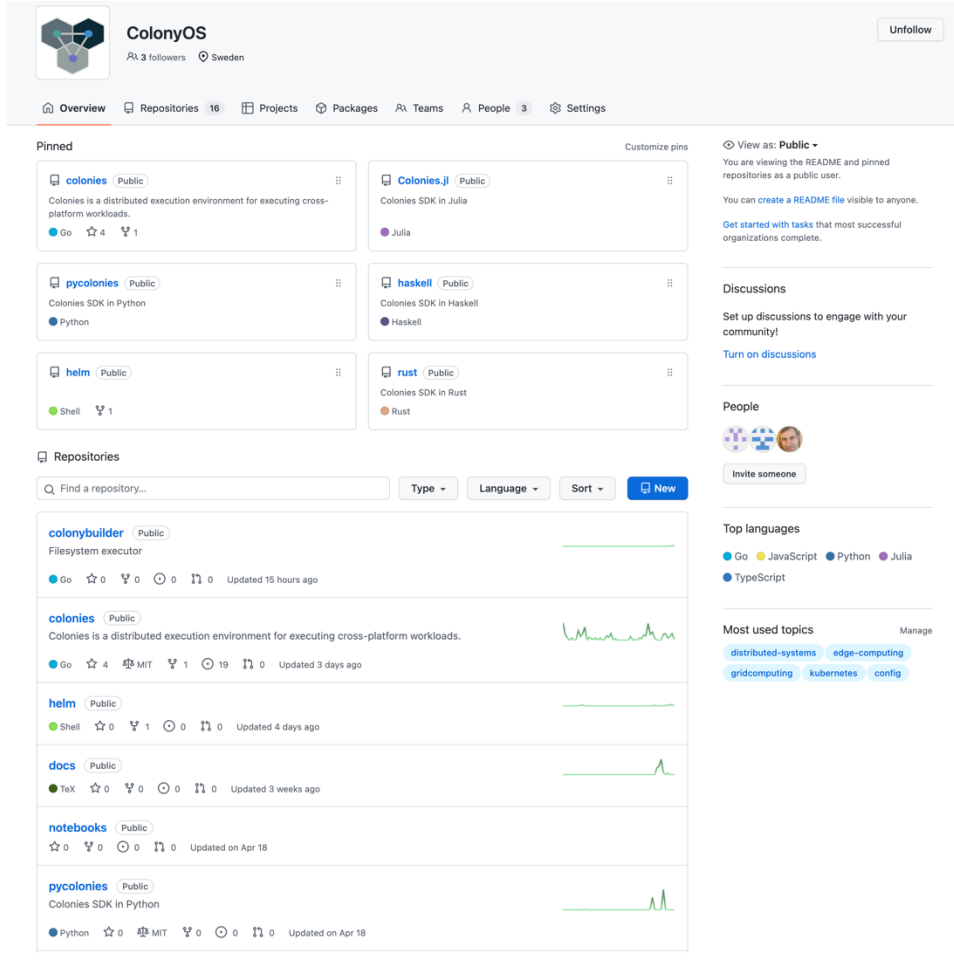
Implement distributed compute engines that optimize data processing across diverse platforms. Perform intensive computations on one platform and then effortlessly merge the

Streamlined HPC

ColonyOS offers modern APIs and cloud integration, expanding supercomputers' reach and accessibility. HPC Executors enables easy, platform-agnostic deployment of workloads, boosting global

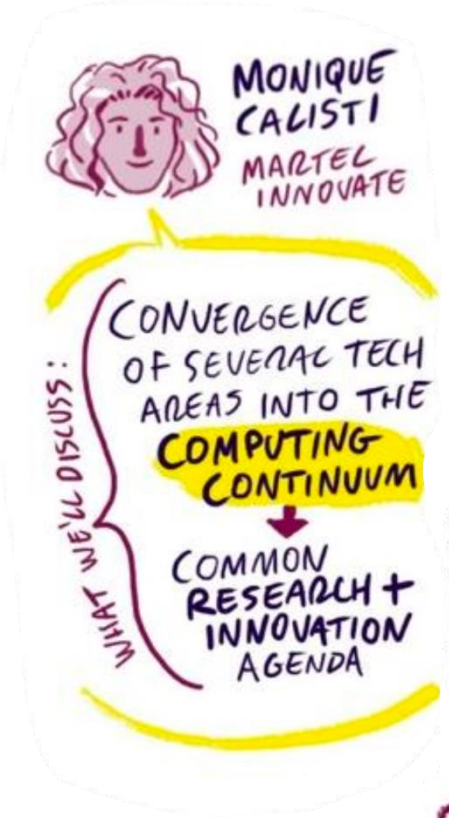
Virtual Supercomputing

Harness and combine computational power of multiple disparate computing systems, whether HPC, cloud-based infrastructures, or other computing resources, to



https://colonyos.io

https://github.com/colonyos



Continuum

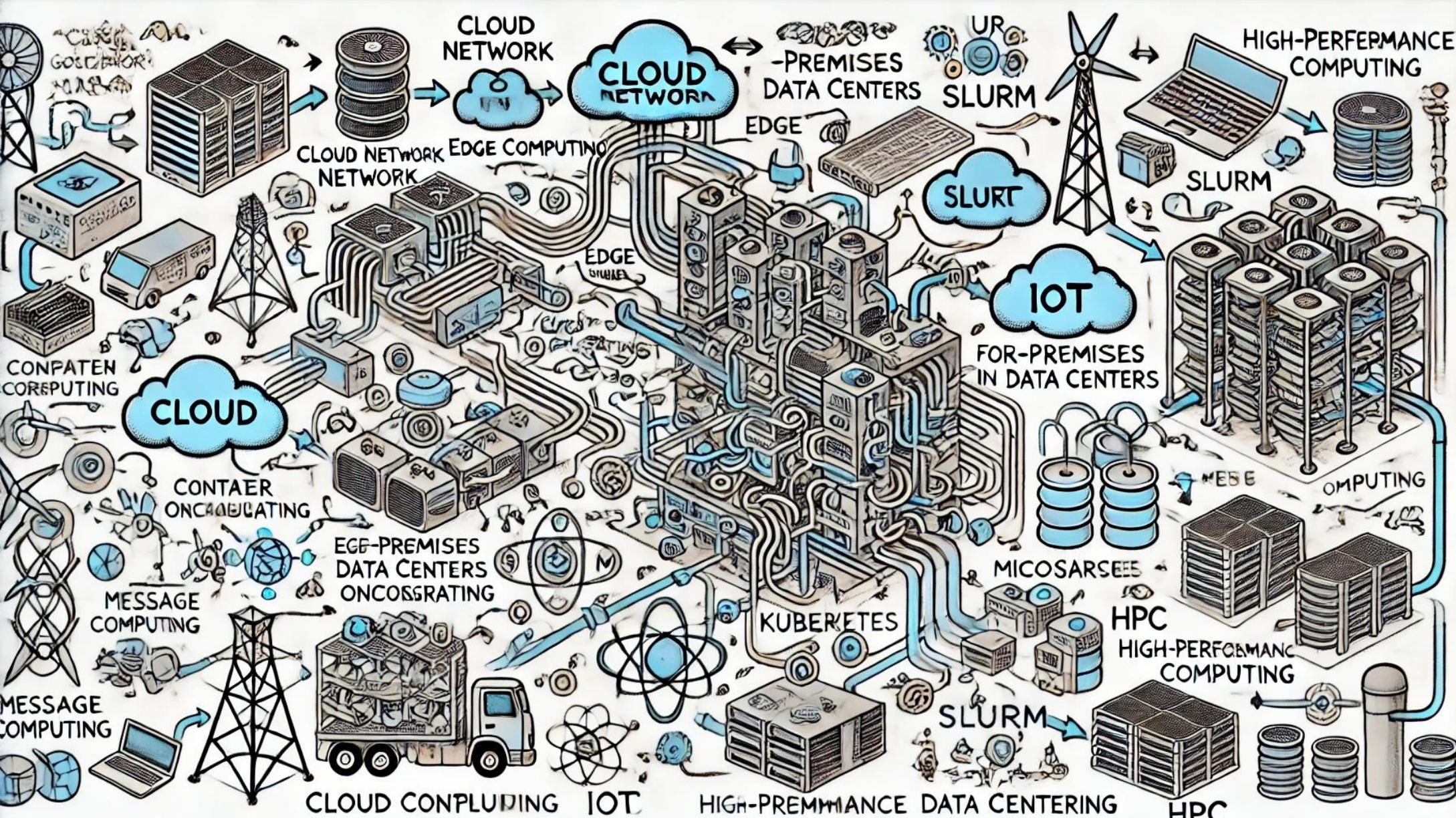
"A coherent whole characterized as a collection, sequence, or progression of values or elements varying by minute degrees"

- Merriam webster

PRESS RELEASE | 5 December 2023 | Brussels | 8 min read

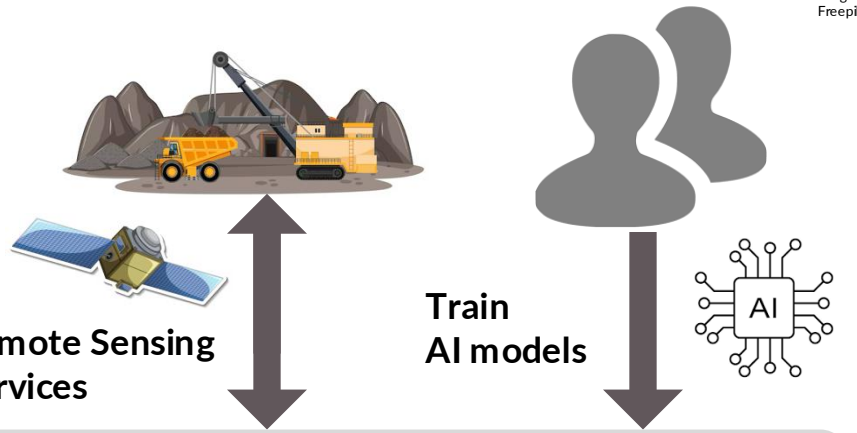
Commission approves up to €1.2 billion of State aid by seven Member States for an Important Project of Common European Interest in cloud and edge computing technologies

Source: Executive Summary - Digital Autonomy in the Computing Continuum

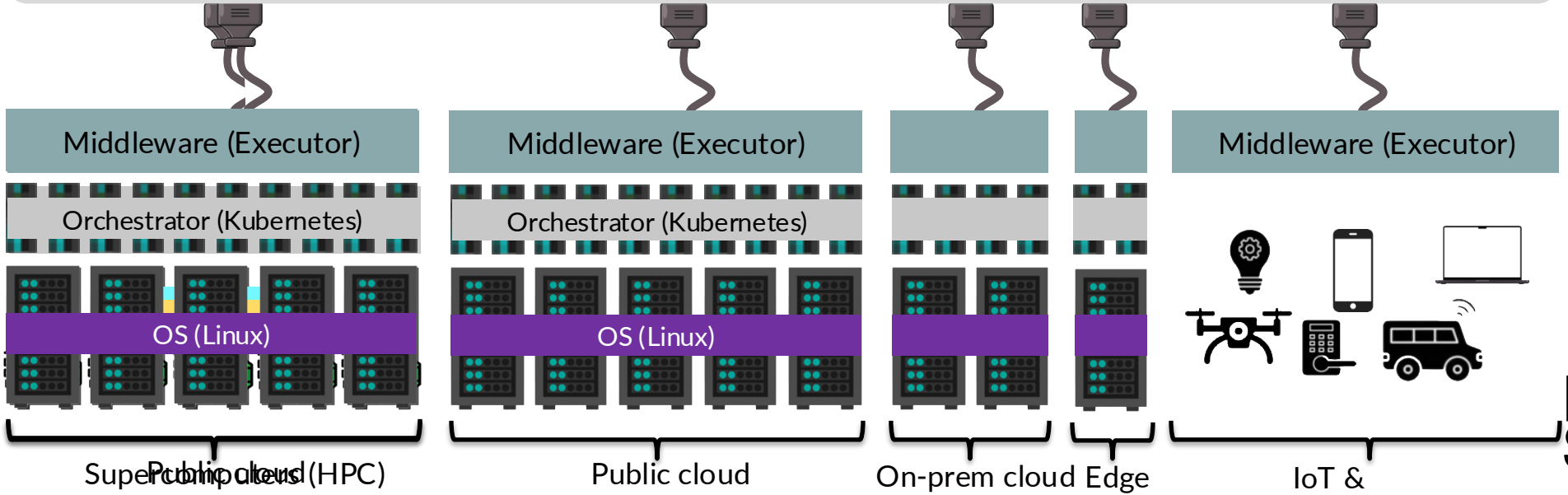

CLOUD NETWORK
CLOUD
-PREMISES DATA CENTERS
SLURM
HIGH-PERFERMANCE COMPUTING
CLOUD NETWORK NETWORK
EDGE COMPUTING
EDGE
SLURT
SLURM
EDGE
IOT
FOR-PREMISES IN DATA CENTERS
CONPATEN COREPUTING
CLOUD
CONTAER ONCAOBLATING
EGE-PREMISES DATA CENTERS ONCOSSRATING
MESSAGE COMPUTING
KUBERETES
MICOSARSEES
HPC
HIGH-PERFORMANC COMPUTING
MESSAGE COMPUTING
SLURM
CLOUD CONPLUDING
IOT
HIGH-PREMMANCE DATA CENTERING
HPC

Meta-Operating Systems

A foundation for Compute Continuums

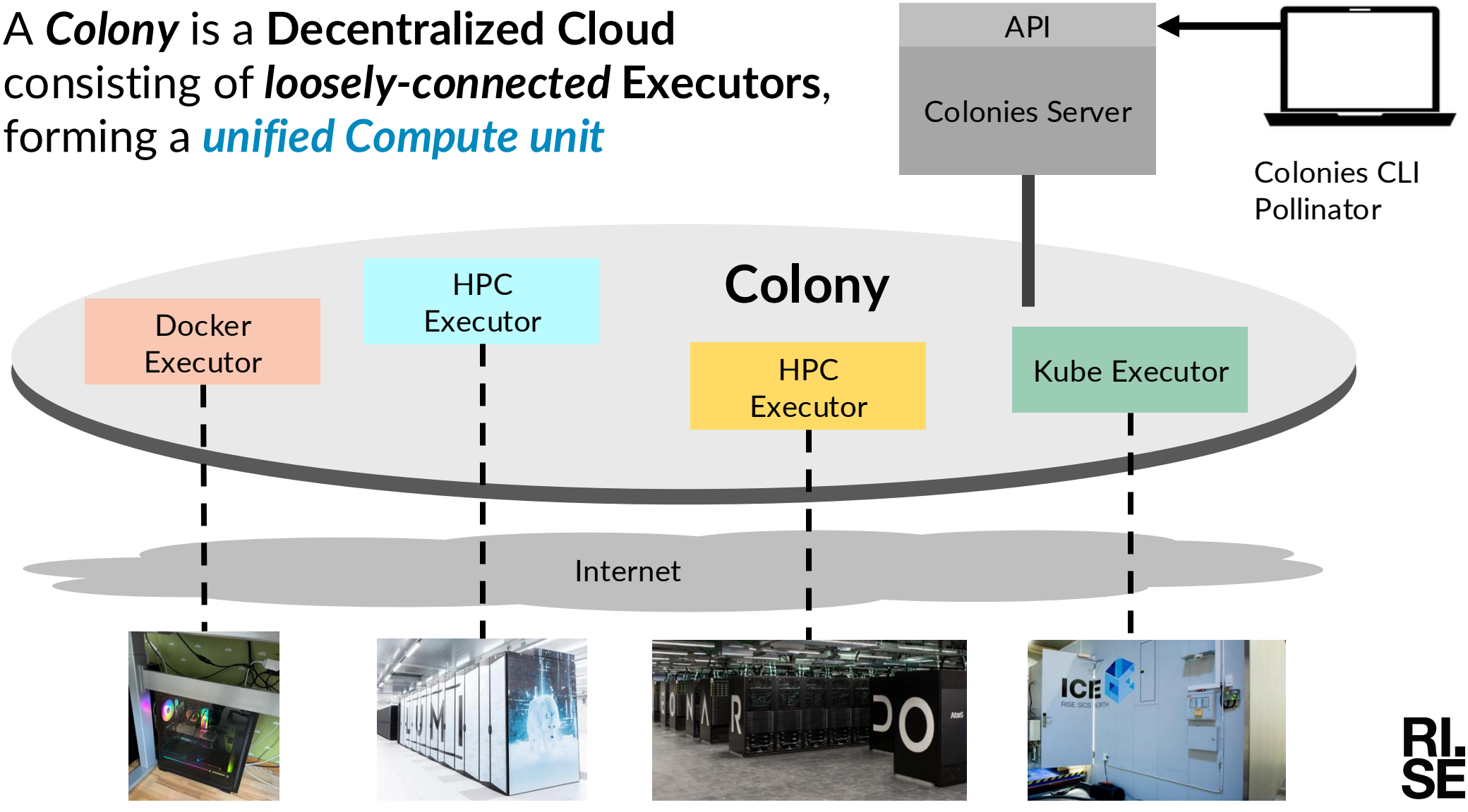


ColonyOS (Meta-Operating System)



RI.SE

A ***Colony*** is a **Decentralized Cloud** consisting of ***loosely-connected*** **Executors**, forming a ***unified Compute unit***

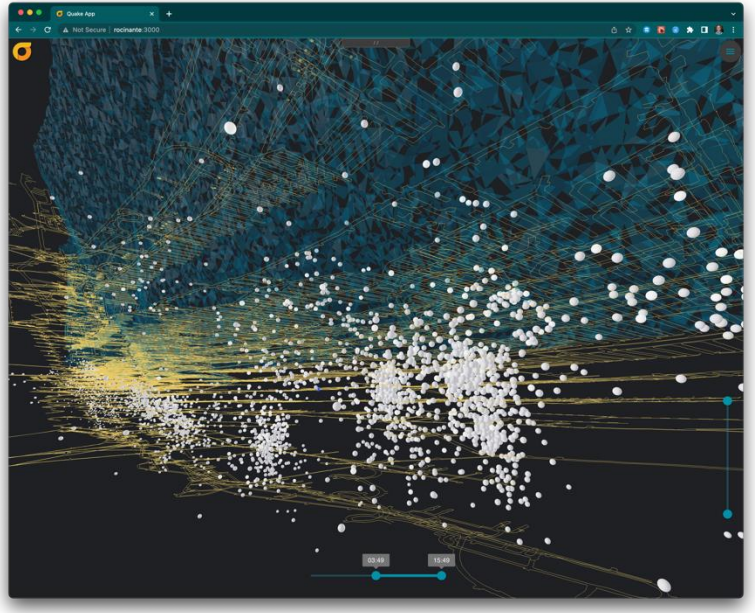


Use case
Seismic real-time data processing

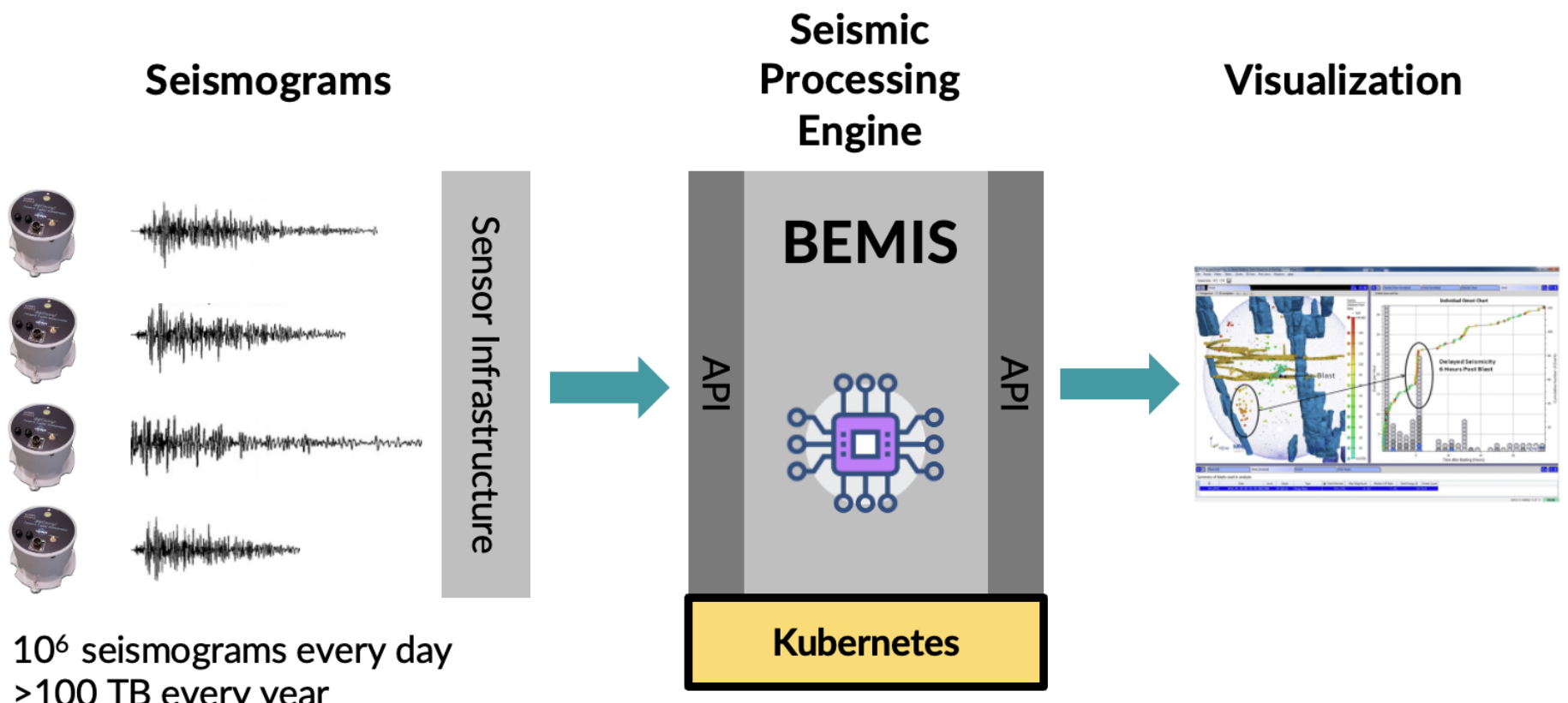
RI.
SE

RockSigma AB

- Seismic processing underground mines
- Used by LKAB to analyze seismicity and process a massive amount of data from one the largest mines in the world (Kiruna/Malmberget)
- On-preem + cloud



RockSigma AB

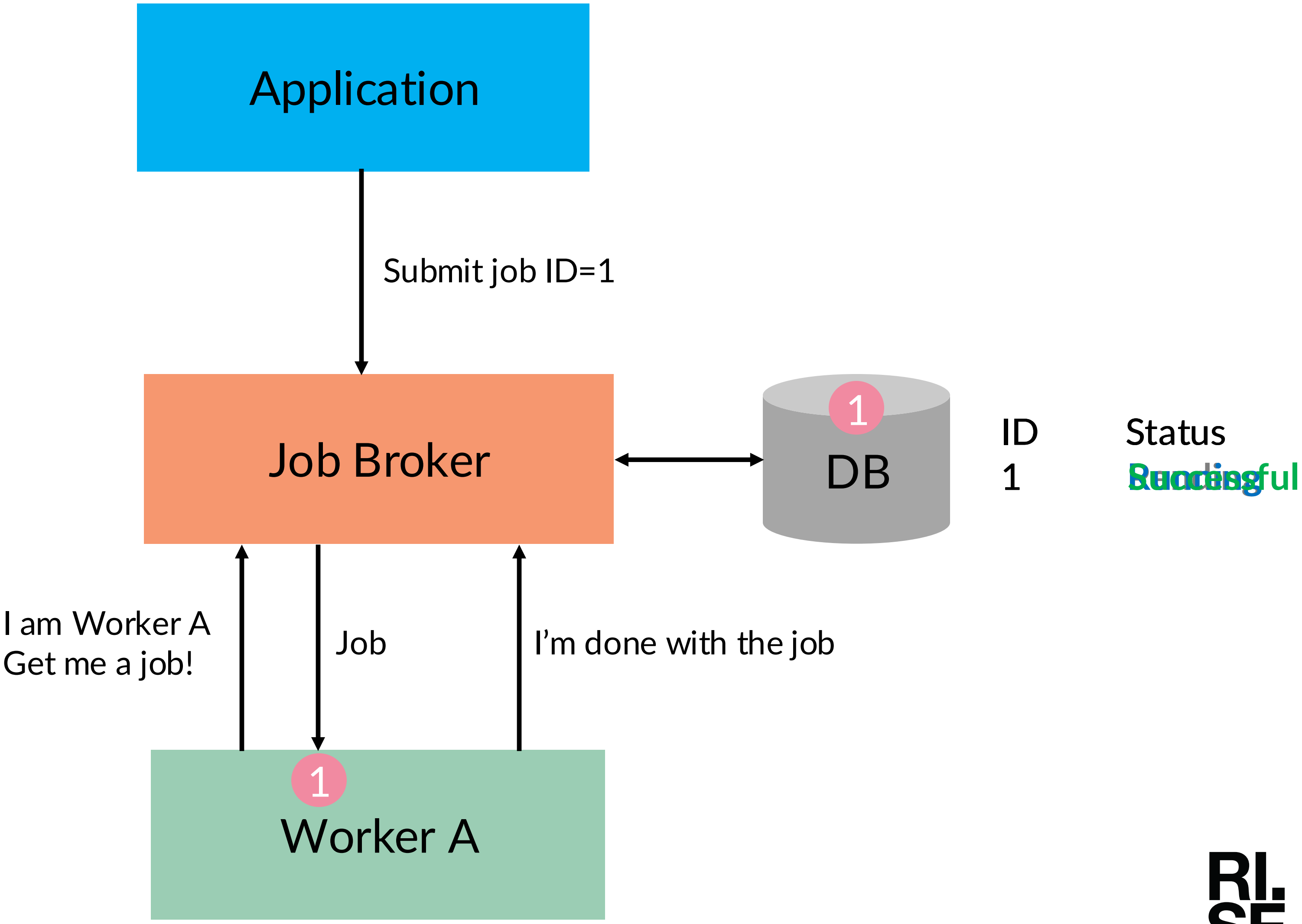


Challenges

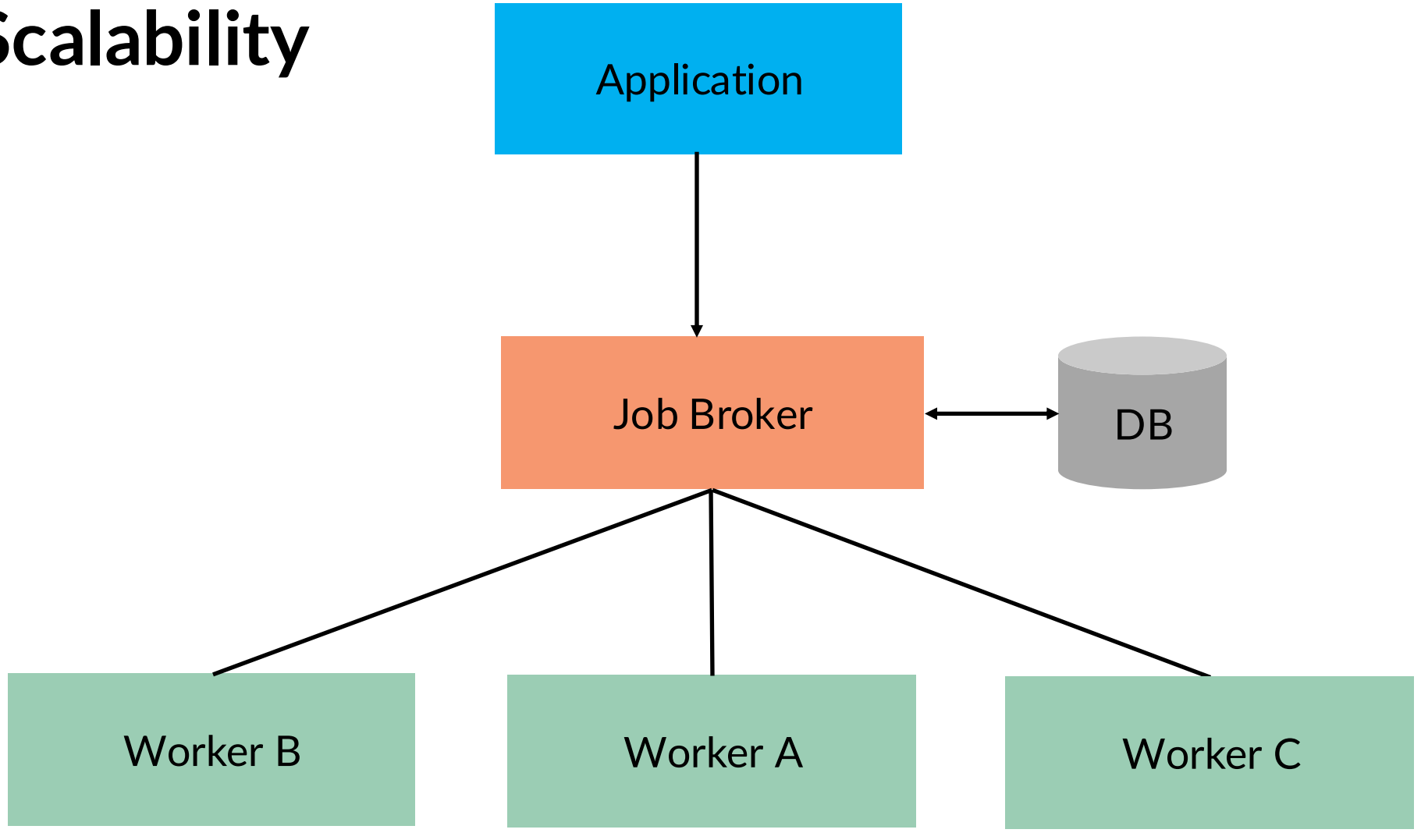
- Starting K8s pods can take long time!
- Julia code takes an eternity to start! **We need worker pods!**
- How to deploy and upgrade workers?
- What happens if worker is killed while processing a task?
- Scalability? Cloud bursting?
- Easy to deploy & manage
- Minimalistic

RI.
SE

Job broker

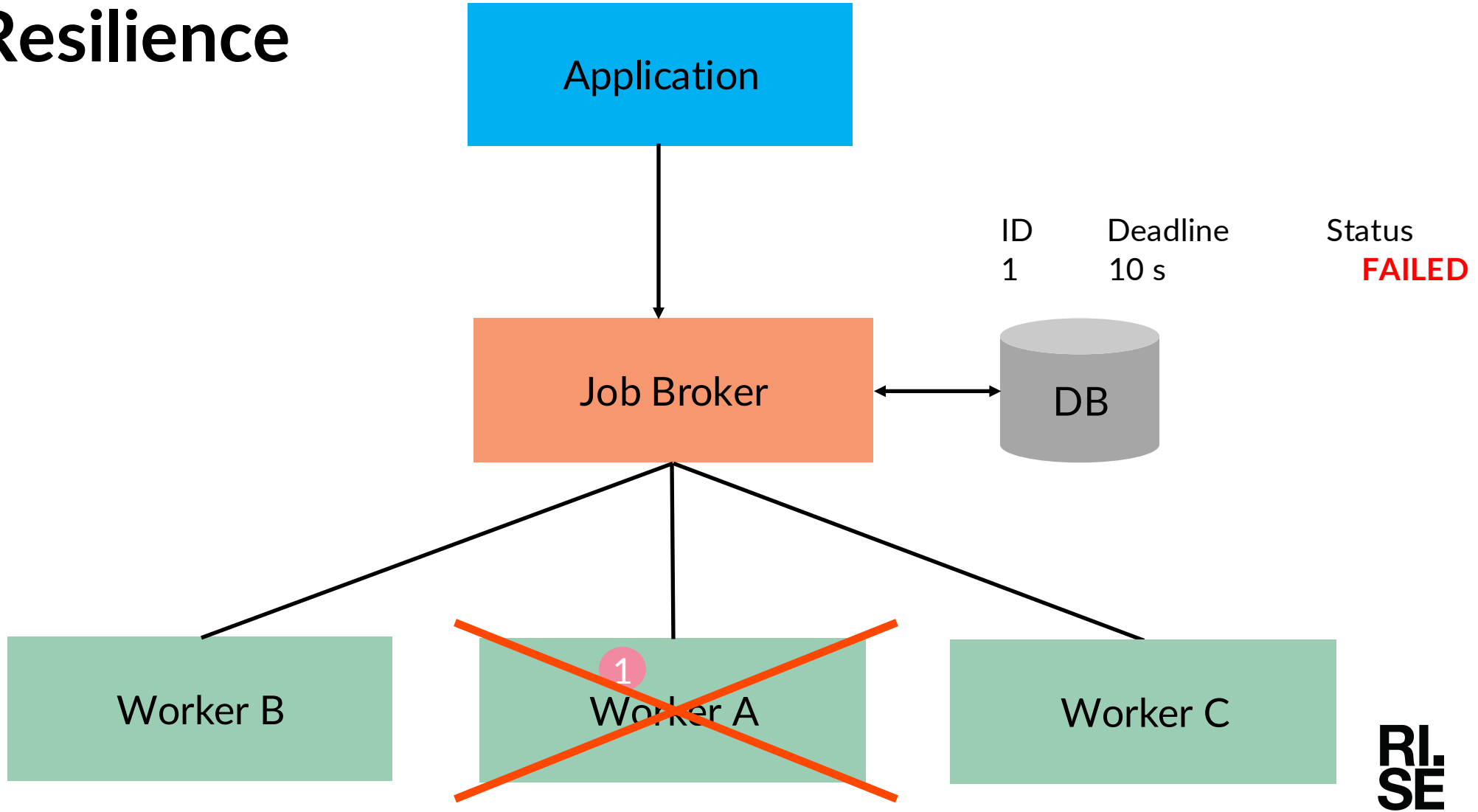

Application
Submit job ID=1
Job Broker
1
DB
ID
Status
1
Successful
I am Worker A
Get me a job!
Job
I'm done with the job
1
Worker A

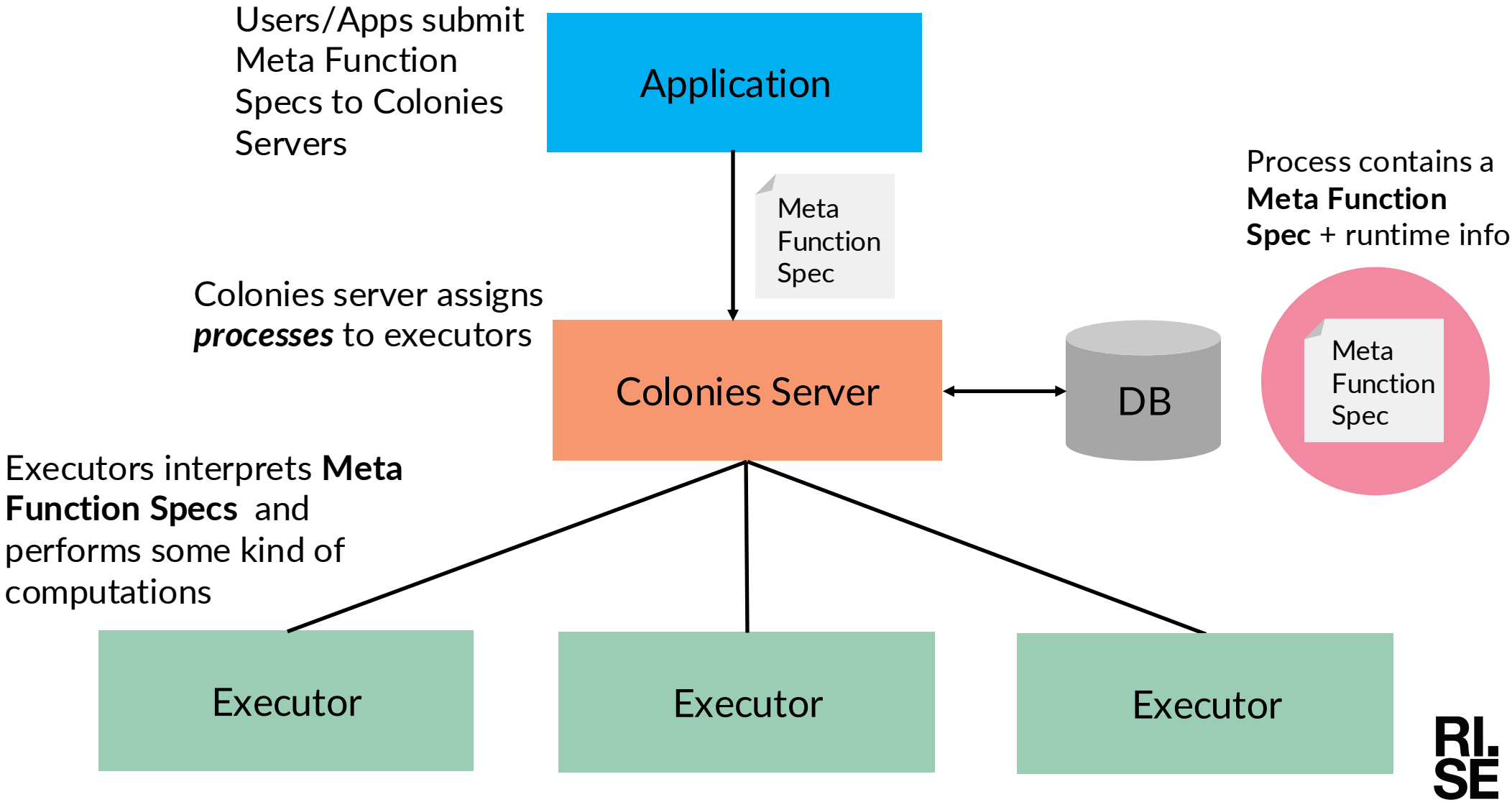
Scalability



RI.
SE

Resilience


Application
ID
1
Deadline
10 s
Status
FAILED
Job Broker
DB
1
Worker B
Worker A
Worker C

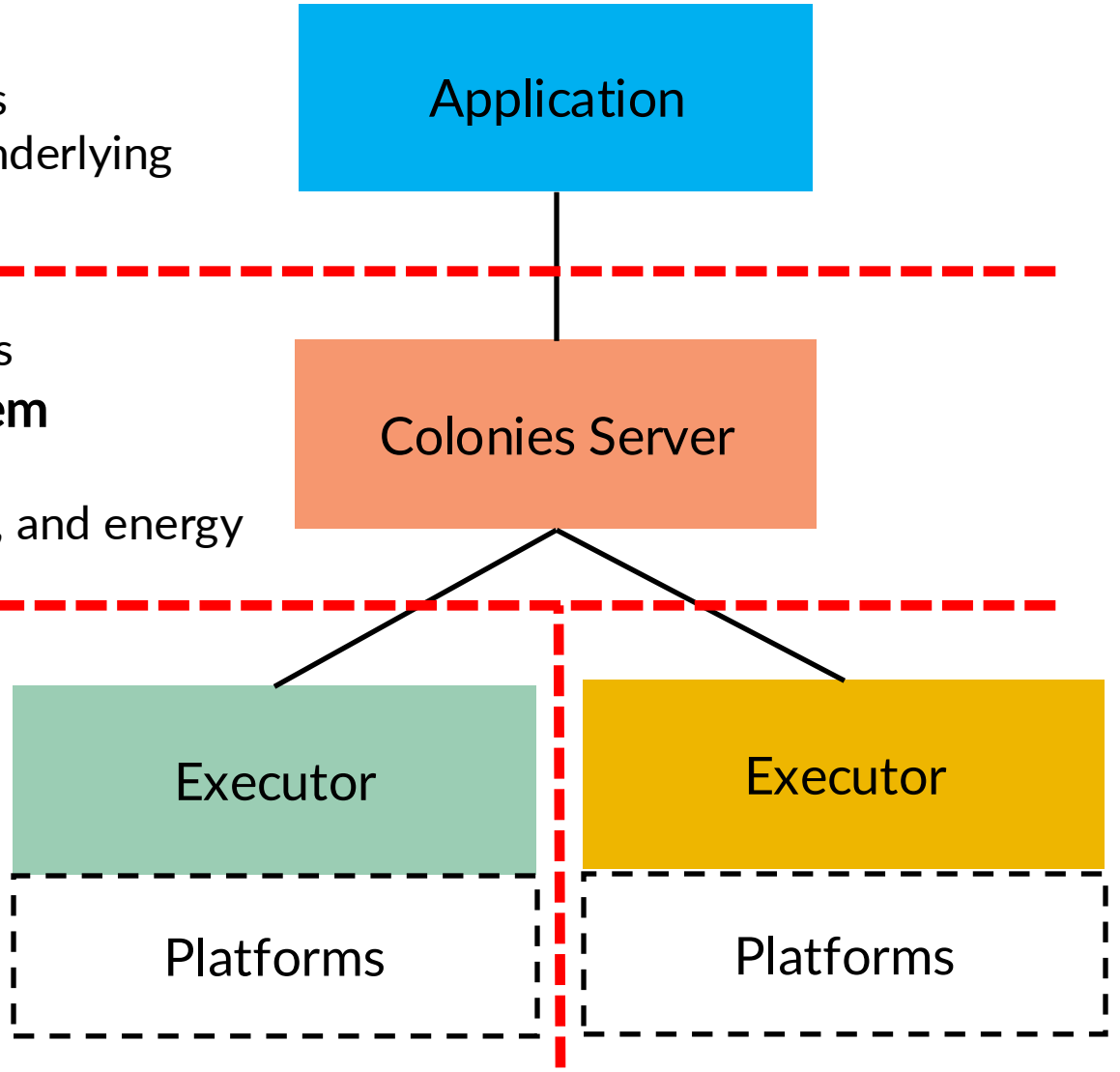

Users/Apps submit Meta Function Specs to Colonies Servers
Application
Meta Function Spec
Process contains a **Meta Function Spec** + runtime info
Colonies server assigns ***processes*** to executors
Colonies Server
DB
Meta Function Spec
Executors interprets **Meta Function Specs** and performs some kind of computations
Executor
Executor
Executor

Separation of concerns

- Users describe meta function calls
- Do not need to understand the underlying platforms

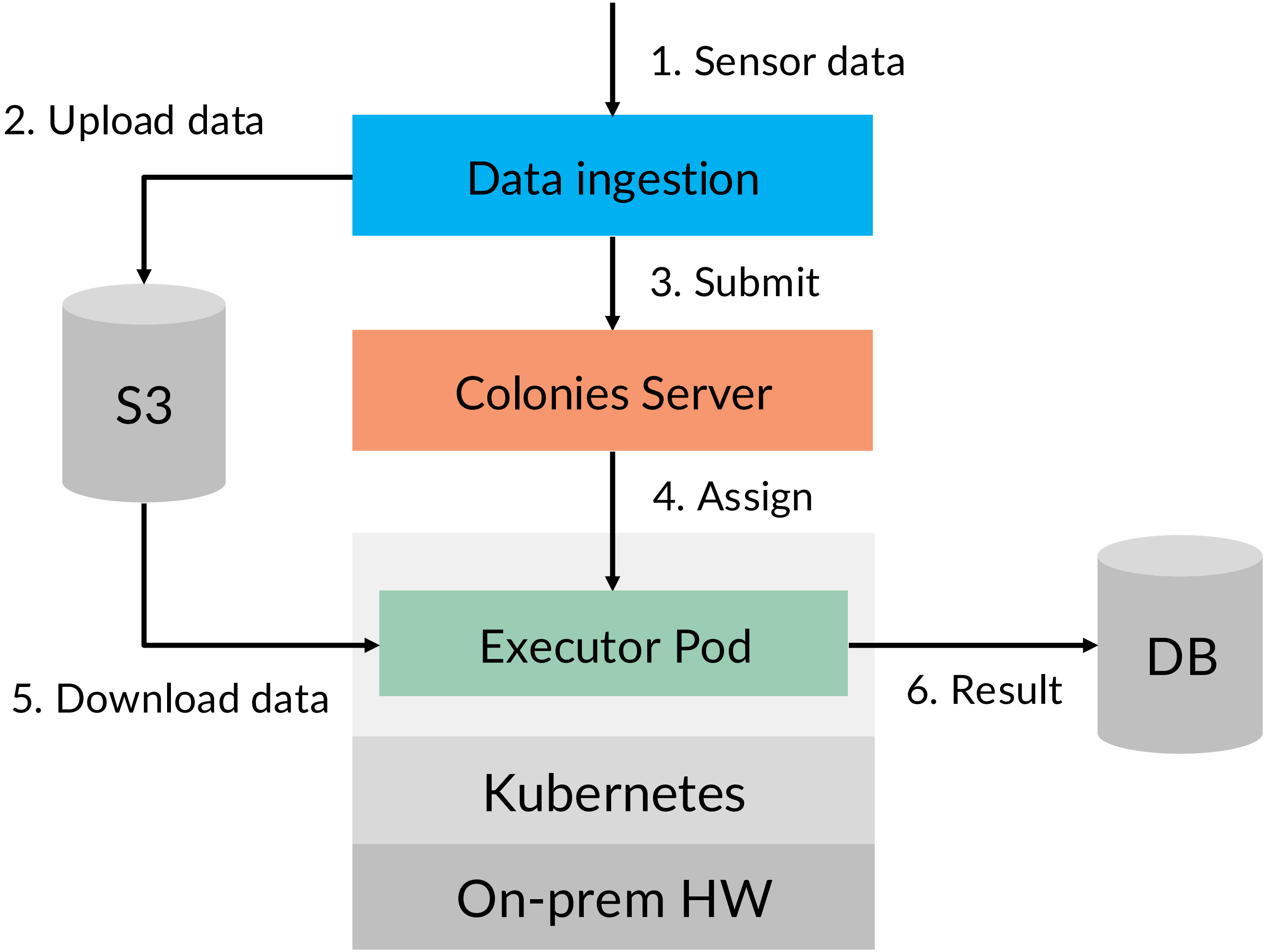
- Abstracts away complex platforms
- Enables a **loosely coupled system**
- Dynamic allocation of resources
- Optimize performance, scalability, and energy

- Executors are microservices designed to execute specific functions
- Integrate with other platforms
- DevOps teams responsible for different Executors
- Reside anywhere on the Internet

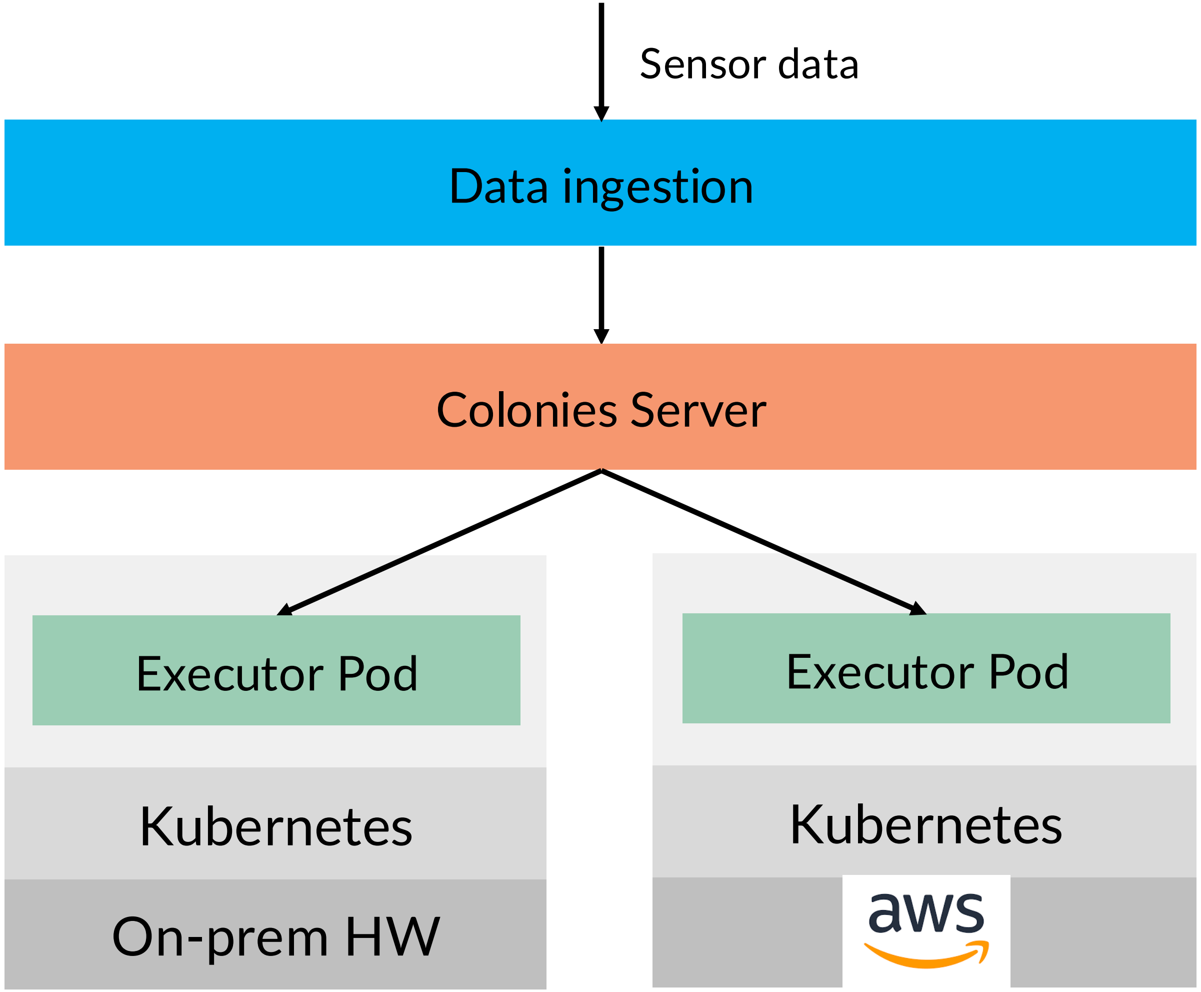


RI. SE

A Seismic Processing Engine



A Seismic Processing Engine



Use case
Cloud-HPC Contiunuum

RI.
SE

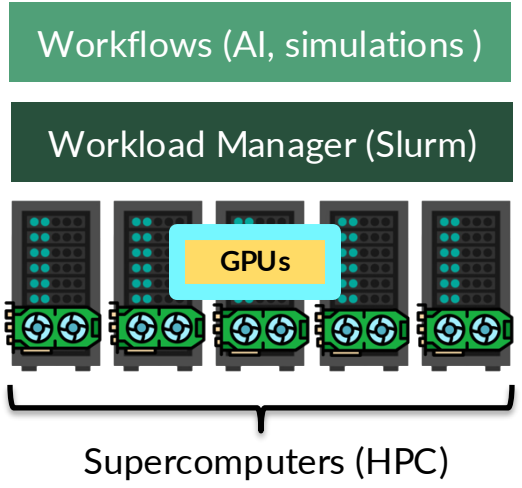
Supercomputers in the Europe Union

EuroHPC Joint Undertaking has a total budget of ~€7 billion for 2021-2027

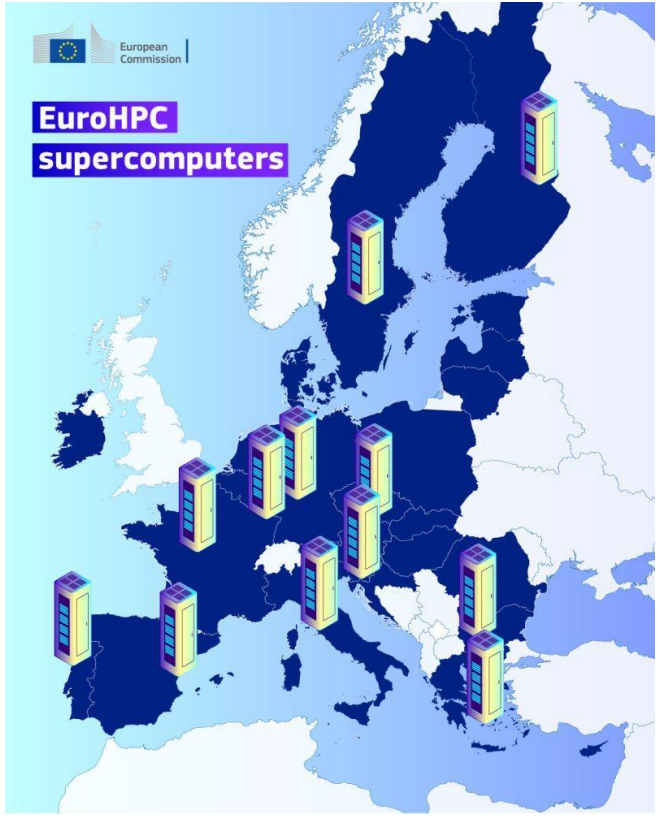
Challenges

- Mainly used by research communities
- Few enterprise users (prefers cloud-based solutions)
- Traditional software stack (but not so easy to use)

- Also, Quantum Computers



LUMI pre-exascale EuroHPC supercomputer
(539 petaflops)



Challenges with the HPC

From a cloud developer perspective …

- **User experience**
  - Complex login process: SSH to a login node
  - Setting up tunnels
  - Mastering Slurm jobs
  - When will my job run? Someone kill my job?
- **Data management**
  - Determining data storage locations
  - Manual data transfers can be time-consuming and error-prone
- **Integration issues**
  - Connecting HPC systems with cloud to streamline workflows?
  - No APIs? Lack of automation tools (GitOps/CI/CD)
  - Multi-factor authentication
  - Sometimes no Internet access on compute nodes

RI.
SE

Ideal for **scientific workflows, large-scale simulations, complex engineering computations**, and tasks requiring extensive computational power and high data throughput

Ideal for **development, testing, and small-scale experimentation**. Suited for prototyping, debugging, and tasks that require immediate, hands-on access to computational resources

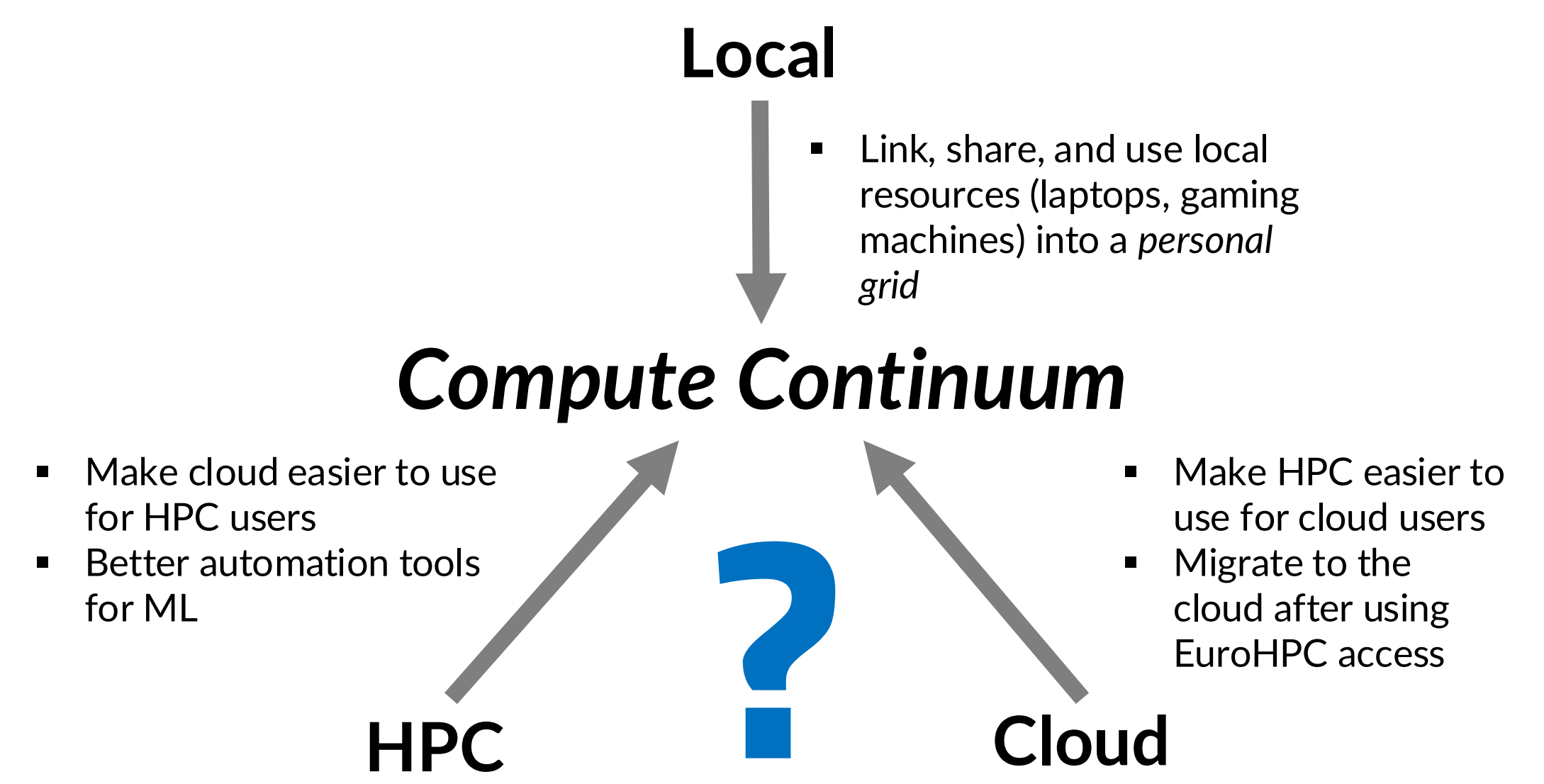
Ideal for **data storage, big data processing, machine learning, and production environments**. Optimized for scalable, distributed web services, and cost-effective resource management across global infrastructures

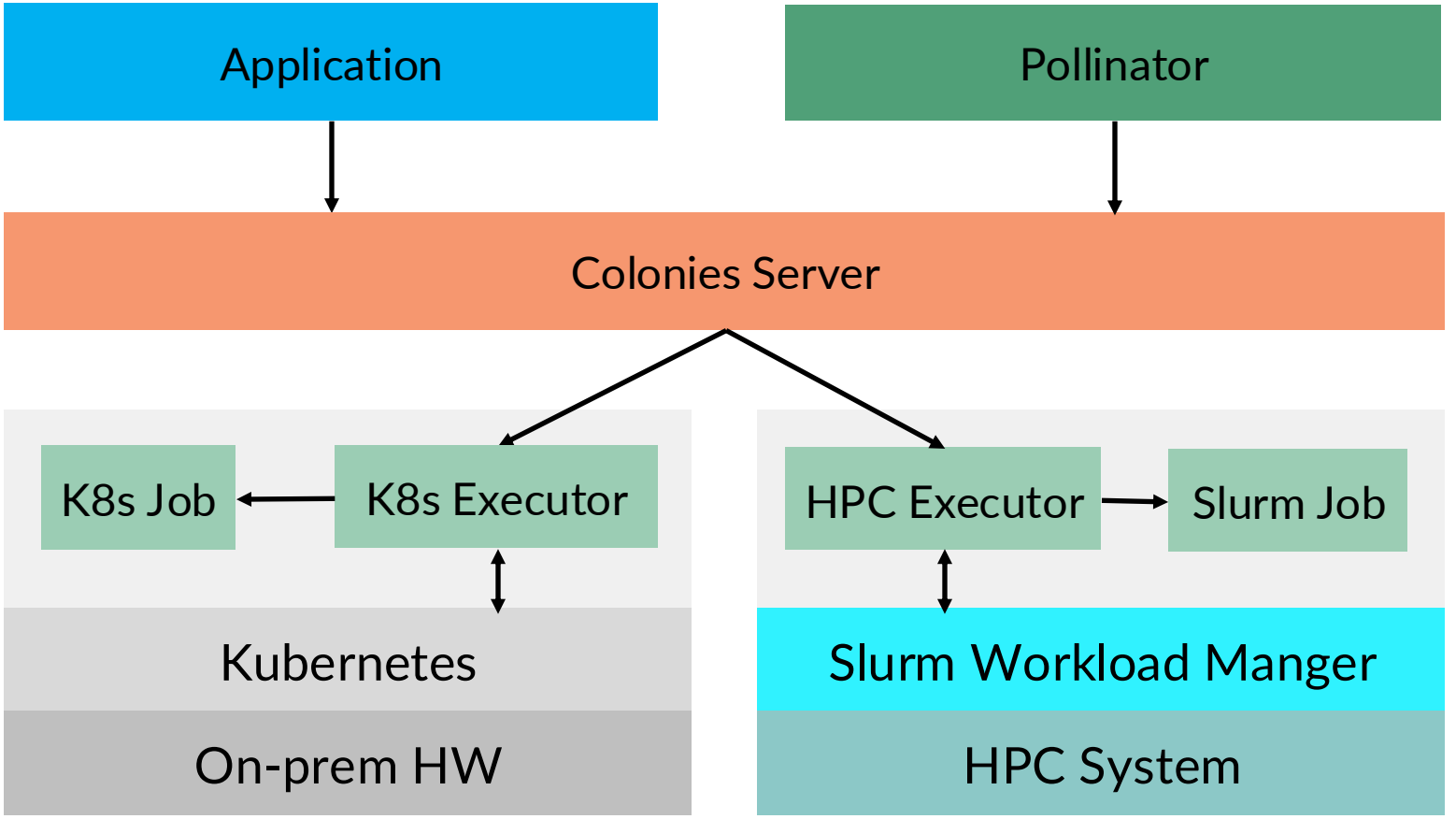
HPC ***Local*** ***Cloud***

Azure

aws

ORACLE CLOUD


Local
Link, share, and use local resources (laptops, gaming machines) into a *personal grid*
Compute Continuum
Make cloud easier to use for HPC users
Better automation tools for ML
Make HPC easier to use for cloud users
Migrate to the cloud after using EuroHPC access
HPC
?
Cloud

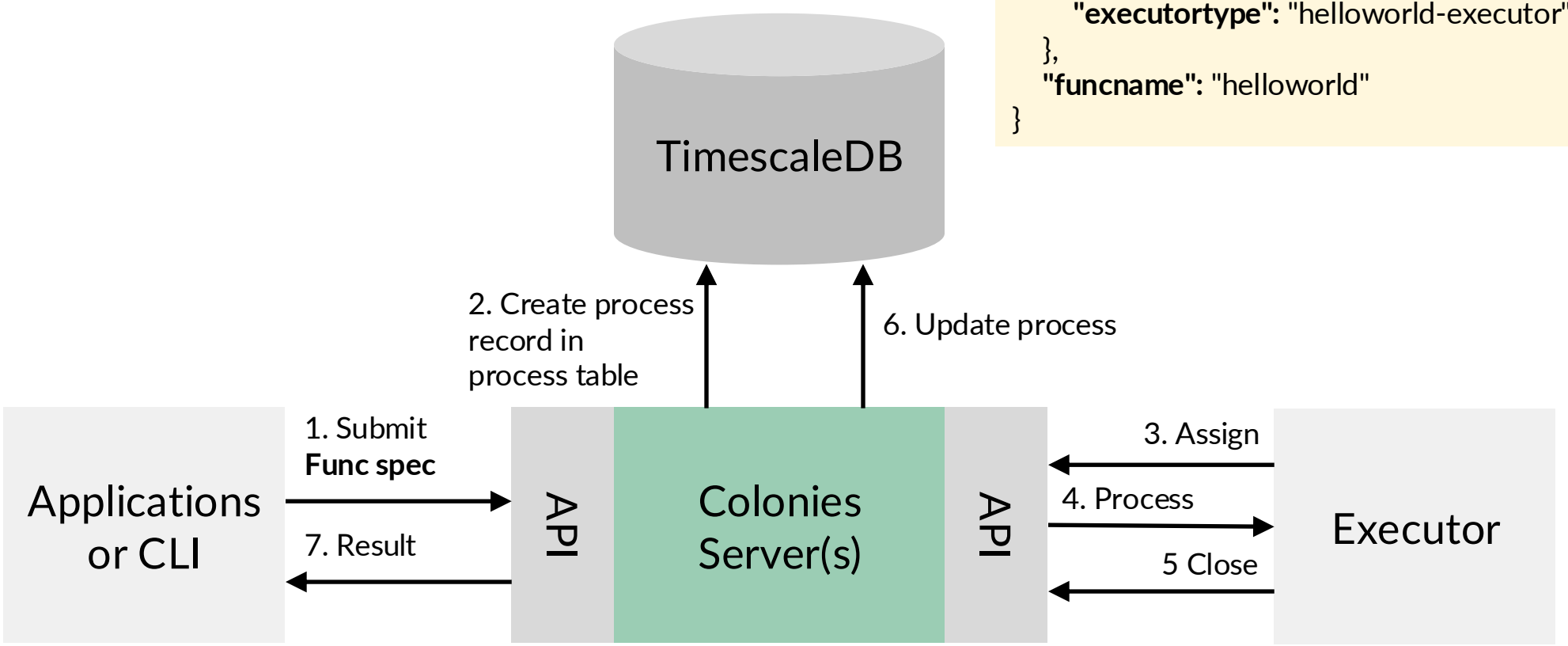

Application
Pollinator
Colonies Server
K8s Job
K8s Executor
HPC Executor
Slurm Job
Kubernetes
Slurm Workload Manger
On-prem HW
HPC System

RI.
SE

Submitting jobs

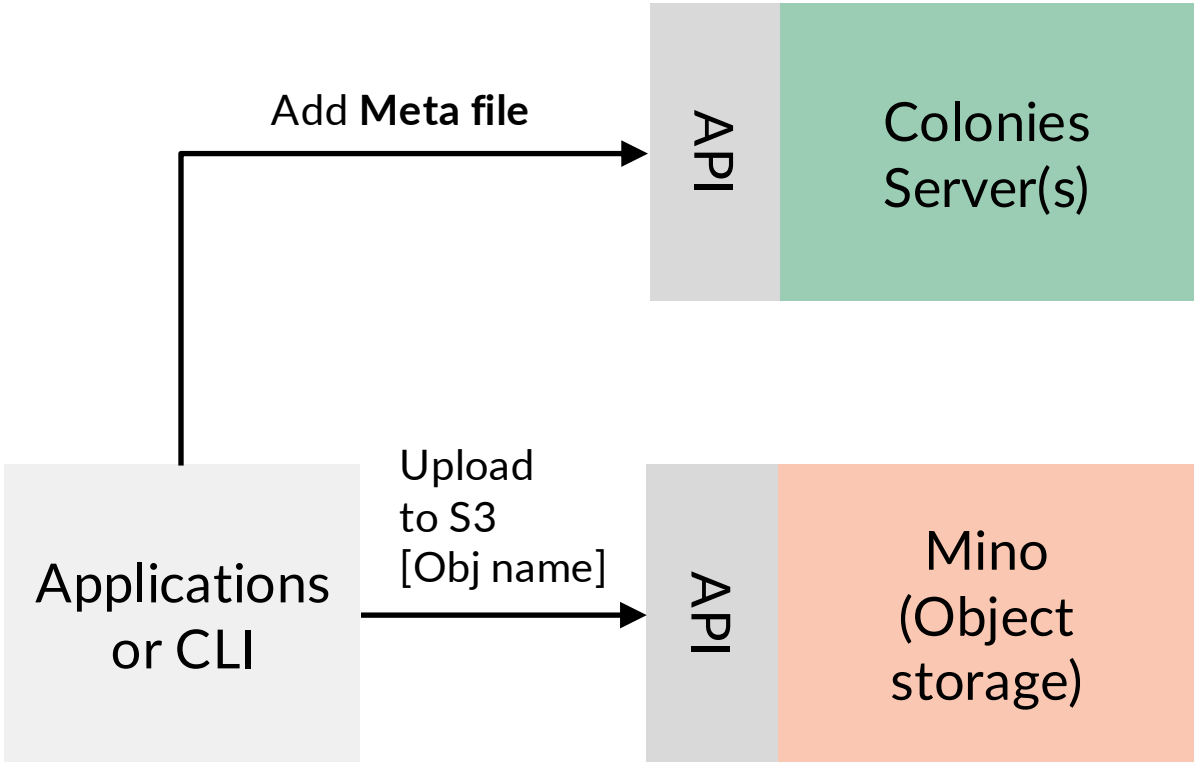
Func spec JSON

```
{
    "conditions": {
        "executortype": "helloworld-executor"
    },
    "funcname": "helloworld"
}
```



RI. SE

Meta-filesystem (ColonyFS)

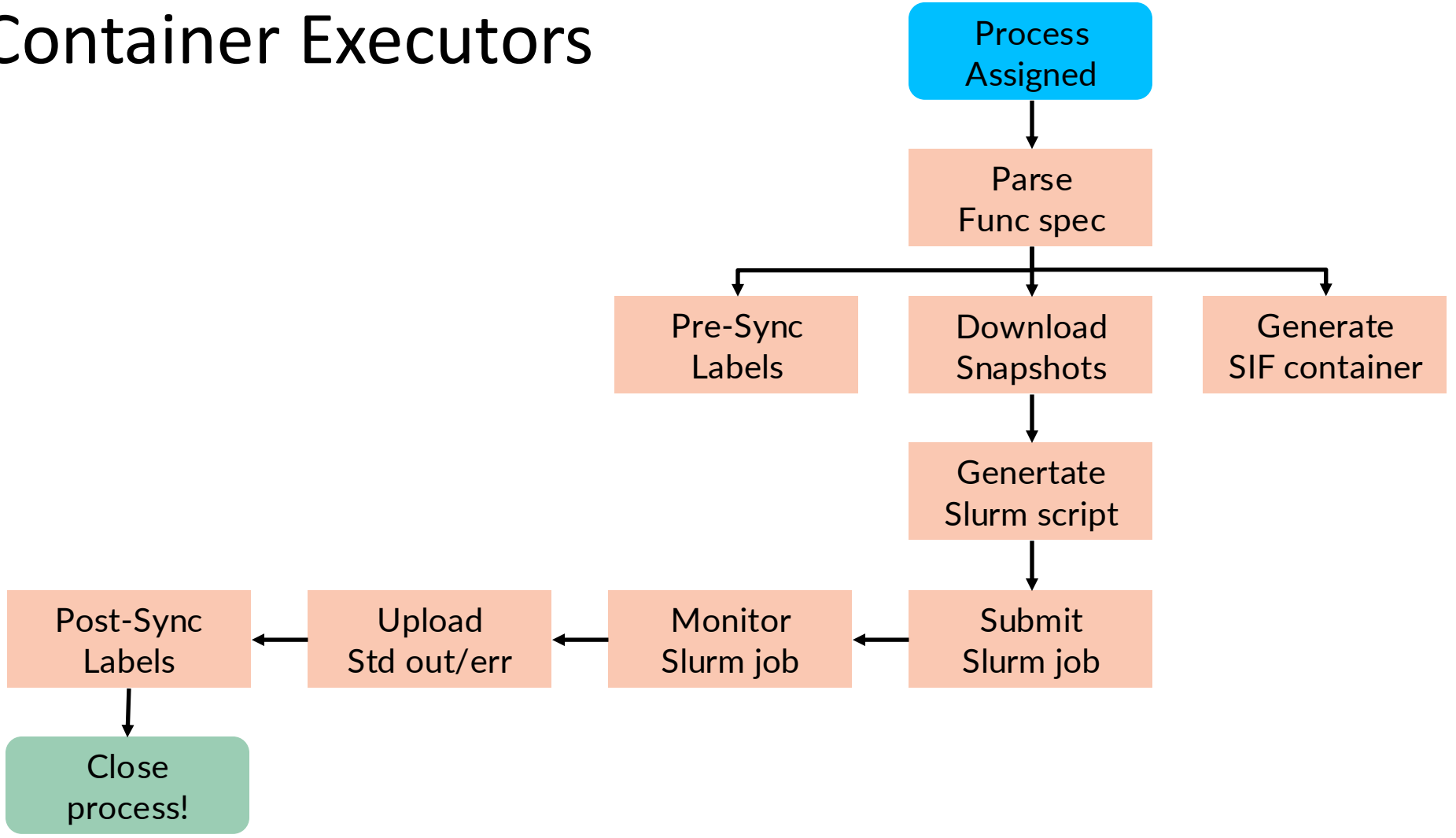


Meta file record

- File name
- Colony name
- Label
- Latest revision ID
- List of revisions, *where each revision:*
  - ID
  - Checksum
  - Version
  - Reference
    - Protocol
      - S3
        - S3 Endpoint
        - S3 Object
        - S3 Bucket

RI. SE

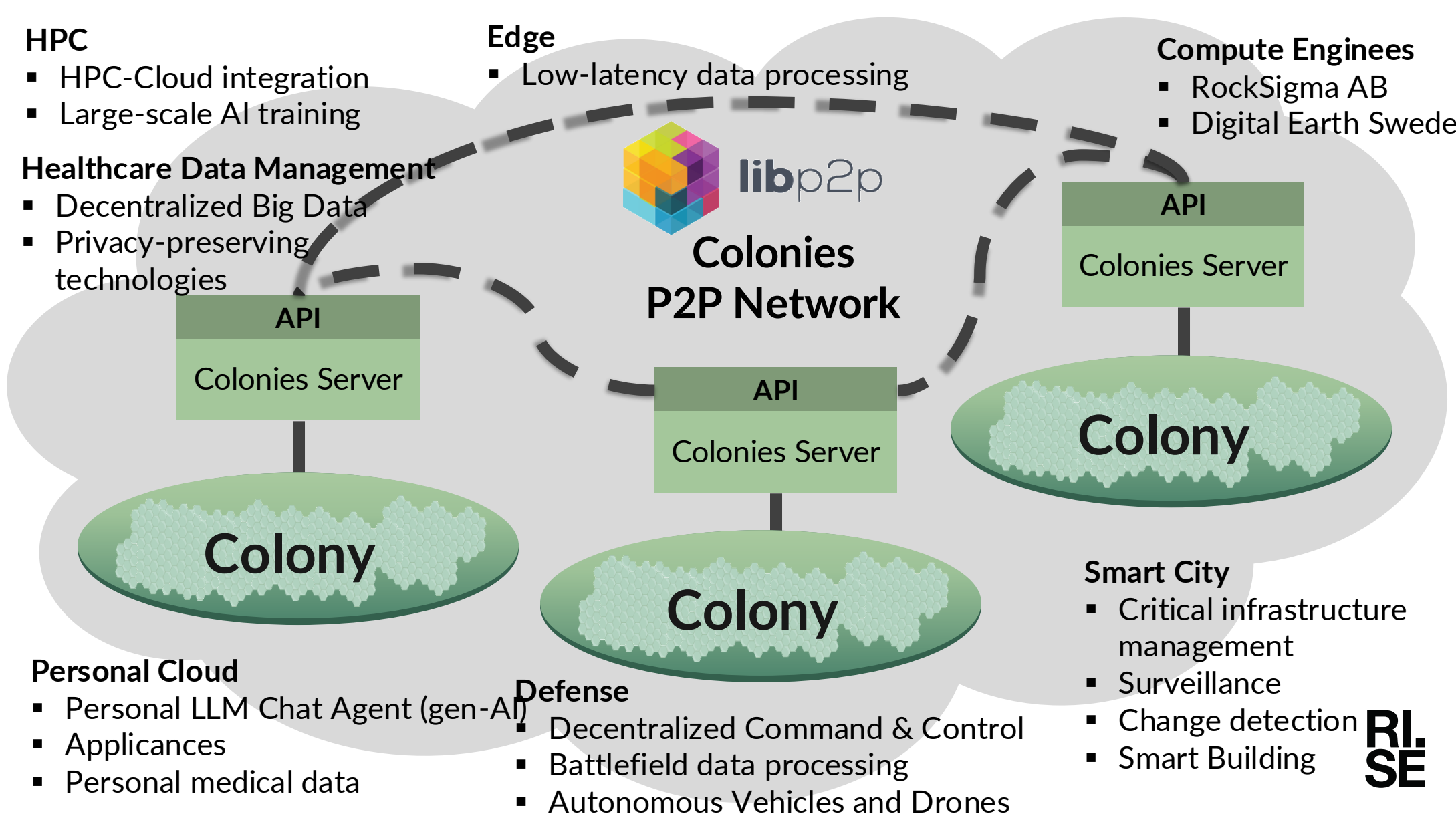
Container Executors



RI. SE

Future?

RI.
SE


HPC
HPC-Cloud integration
Large-scale AI training
Healthcare Data Management
Decentralized Big Data
Privacy-preserving technologies
Edge
Low-latency data processing
Compute Enginees
RockSigma AB
Digital Earth Swede
libp2p
Colonies P2P Network
API
Colonies Server
API
Colonies Server
API
Colonies Server
Colony
Colony
Colony
Personal Cloud
Personal LLM Chat Agent (gen-AI)
Applicances
Personal medical data
Defense
Decentralized Command & Control
Battlefield data processing
Autonomous Vehicles and Drones
Smart City
Critical infrastructure management
Surveillance
Change detection
Smart Building
RI. SE

Future of Computing?

2020- Compute Continuum

Next 10+ years Pervasive Computing

- **Pervasive Computing:** Seamless integration of computing into daily life, operating intelligently and unobtrusively in the background
- **Decentralized Cloud:** Enhances privacy, security, and resilience by distributing data and processing across a network of nodes
- **Autonomic Computing:** Functions like a nervous system, automatically managing and optimizing itself with minimal human intervention
- **Seamless AI integration:** AI becomes an invisible, integral part of everyday life