PROCESS

PROviding Computing solutions for ExaScale ChallengeS

Towards Exascale-ready Data Service Solutions

Maximilian Höb

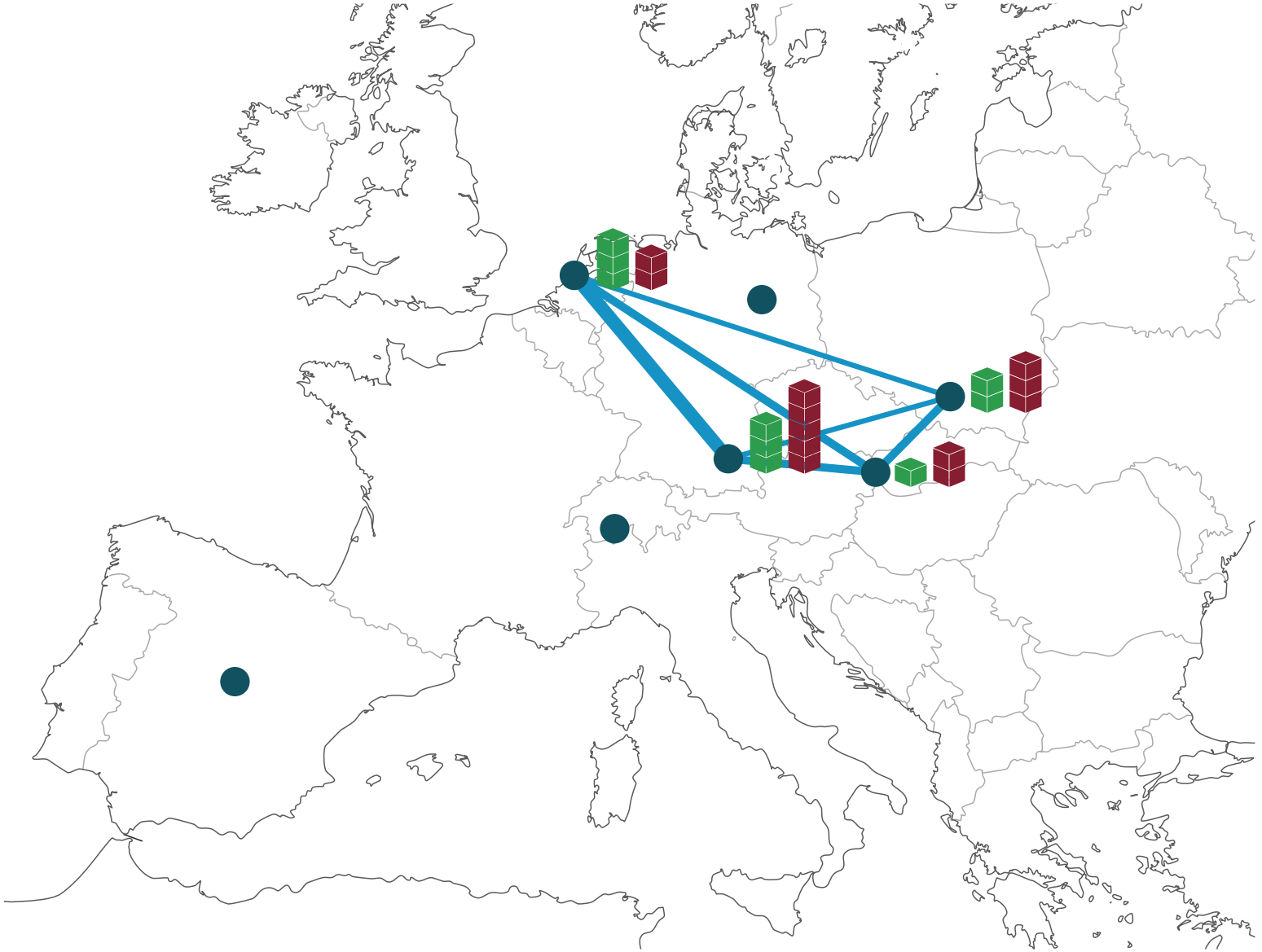
10th July 2019, Athens

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 777533.

Consortium

- Ludwig-Maximilians-Universität München, Germany
- Universiteit van Amsterdam, The Netherlands
- Stichting Netherlands eScience Center, The Netherlands
- Haute école spécialisée de Suisse occidentale, Switzerland
- Lufthansa Systems GmbH & Co. KG, Germany
- Inmark Europa SA, Spain
- Ústav informatiky, Slovenská Akadémia Vied, Slovakia
- Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie, Poland

Storage and Computing Centres



Partner's location

Storage Resources

Compute Resources

Vision of PROCESS

PROCESS will deliver a comprehensive set of **mature service prototypes and tools** specially developed to enable **extreme scale data processing** in both **scientific research and advanced industry** settings

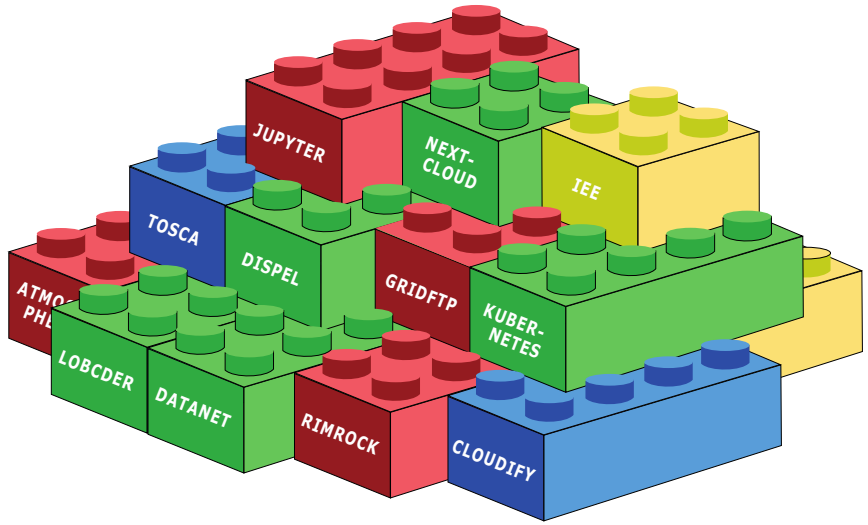
3 Principles

1. Leapfrog beyond the current state of the art
2. Ensure broad research and innovation impact
3. Support the long tail of science and broader innovation

PROCESS Concept

A user-friendly modular exascale service platform to combine data and computational services on top of European research infrastructures

SuperMUC-NG
Leibniz Supercomputing Centre Munich

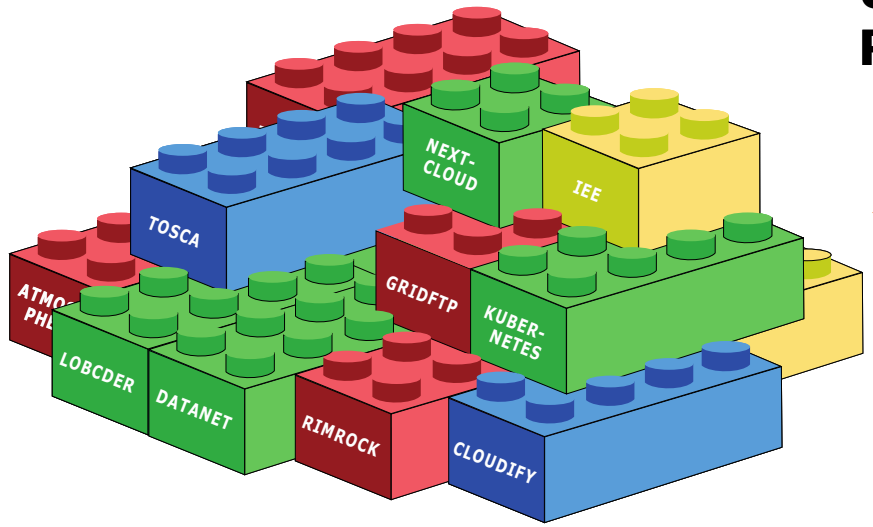


Goals of PROCESS

Mature, modular, generalizable Open Source solutions for user friendly exascale data.

UC#1: Exascale learning on medical image data

UC#2: Analysis of Radioastronomy Observations

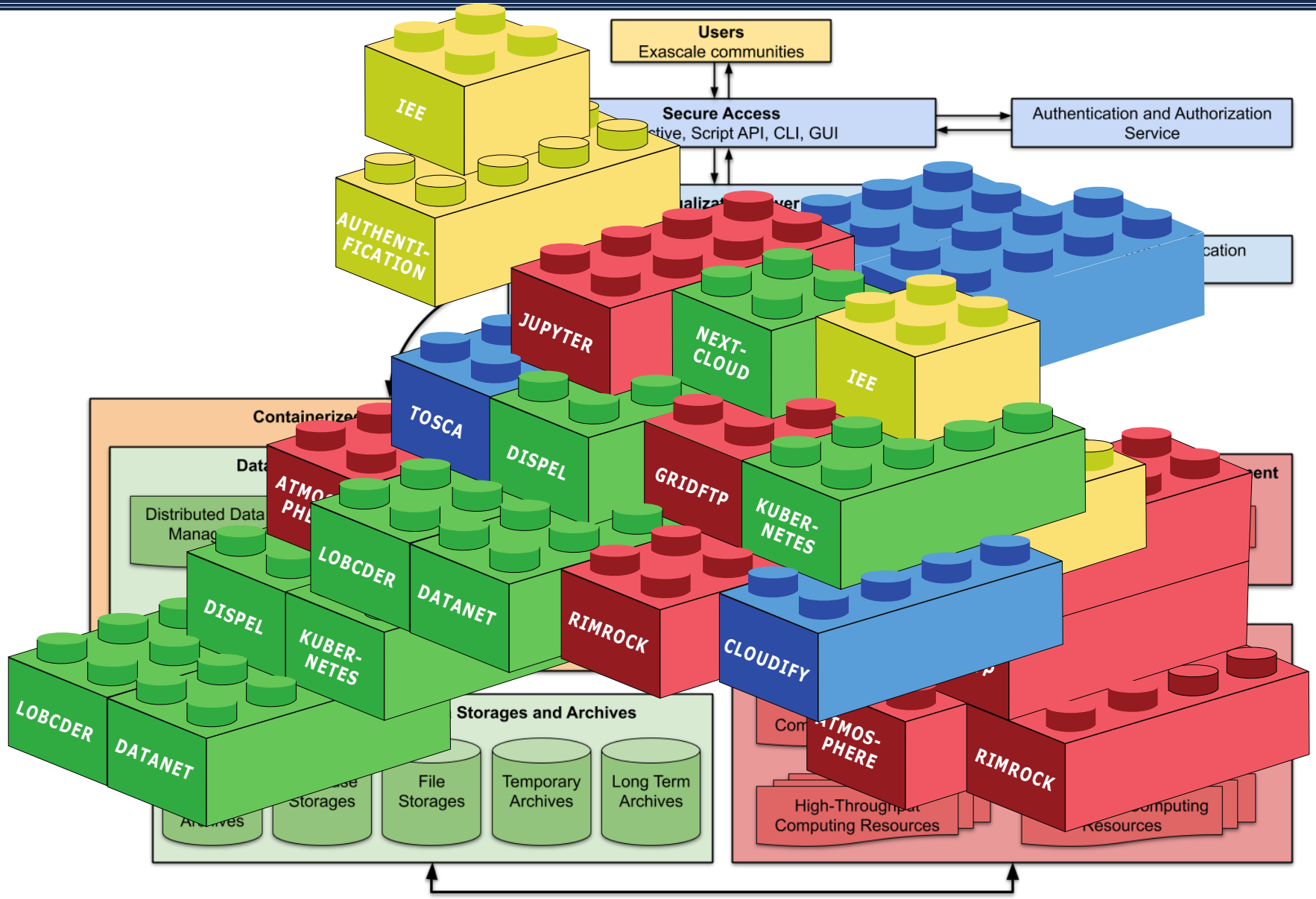


UC#3: Supporting innovation based on global disaster risk data

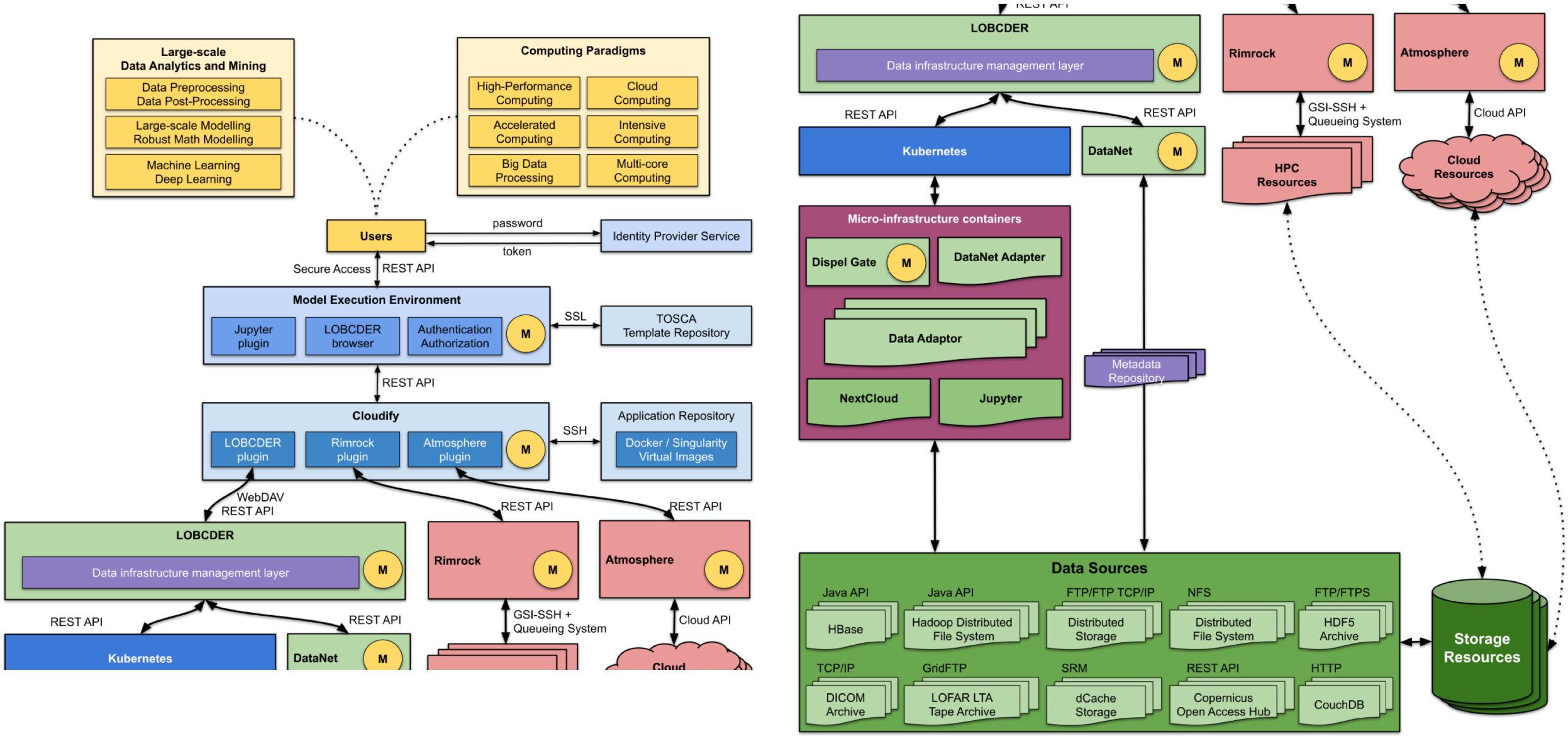
UC#5: Agro-Copernicus (correlating data between simulation and observation)

UC#4: Ancillary pricing/airline revenue management

PROCESS Architecture



PROCESS Architecture



Approach: Tiered system with a layer of virtual (data) nodes facilitating:

- data transfers,
- distributed management,
- scheduling and staging.

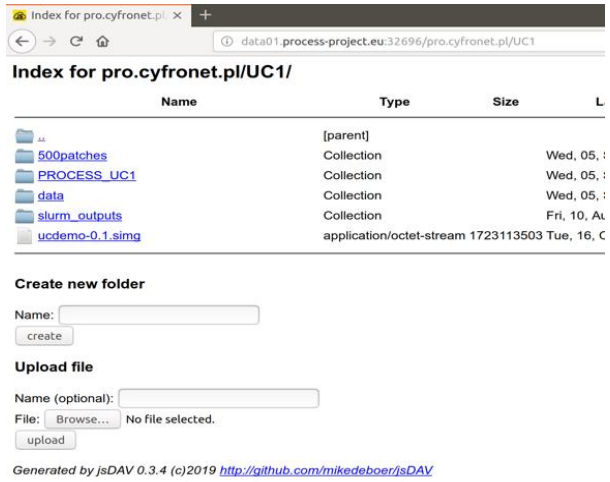
Through → **A programmable micro-infrastructure**

enable ↓

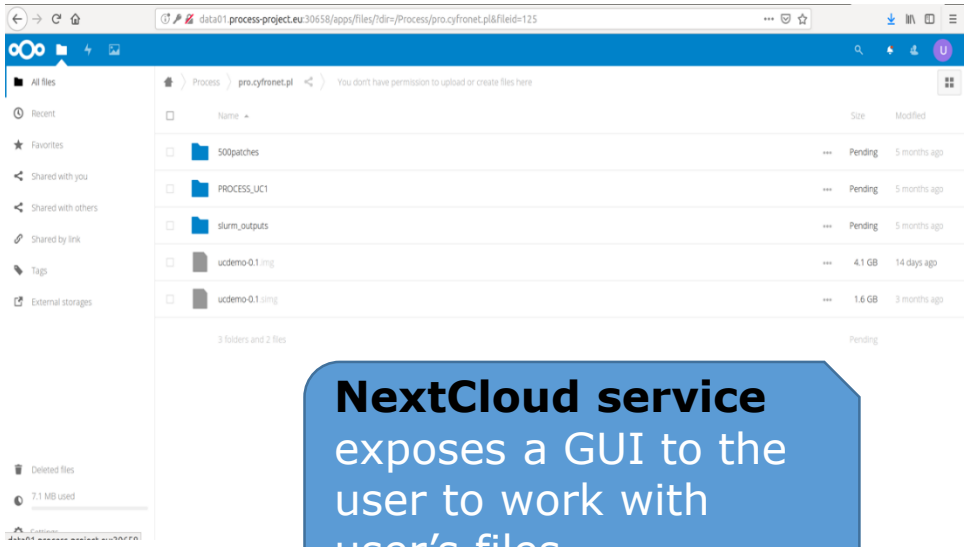
- Independent of resource providers (storage & computing)
- Work with data across distributed provider data.
- collaboration across research groups

Implementation: container-centric, orchestrated using Kubernetes.
software: https://github.com/recap/MicroInfrastructure

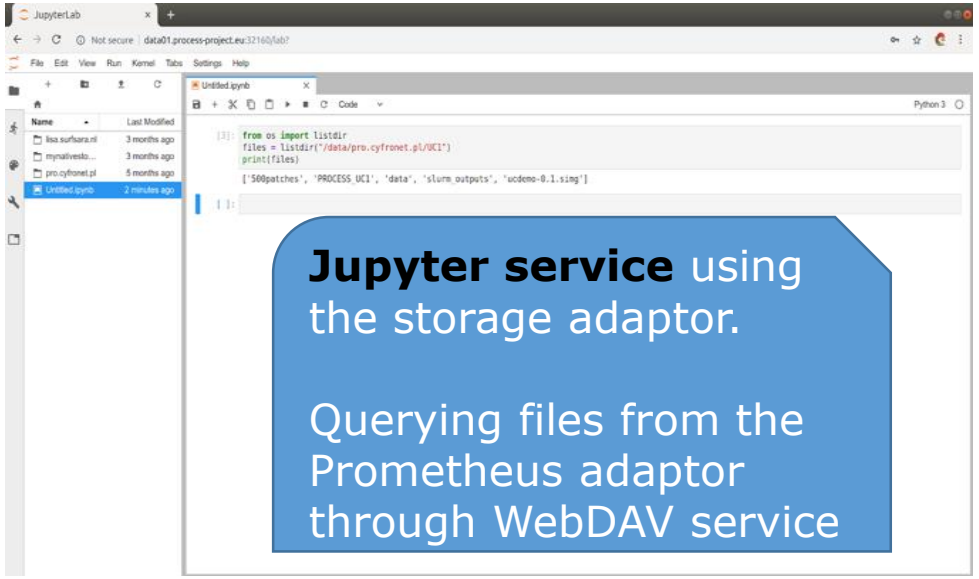
Data Management with containerized Services



WebDAV service deployed through the micro-infrastructure. Through the API the user sets the user-name and password which will protect the WebDAV point

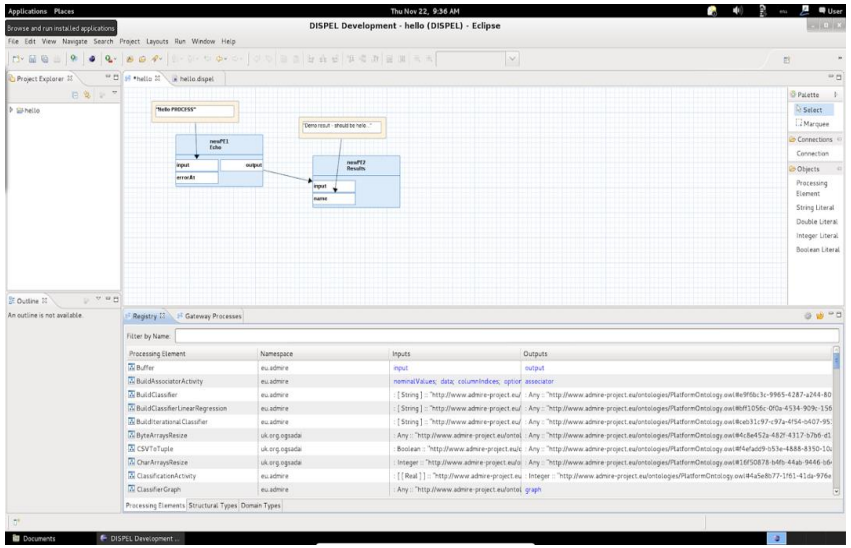


NextCloud service exposes a GUI to the user to work with user's files



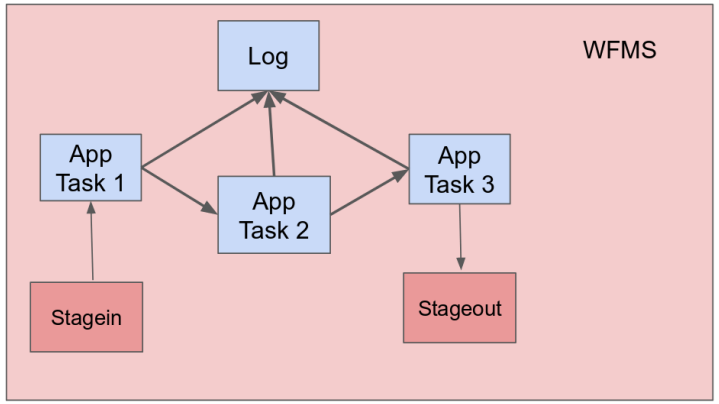
Jupyter service using the storage adaptor.

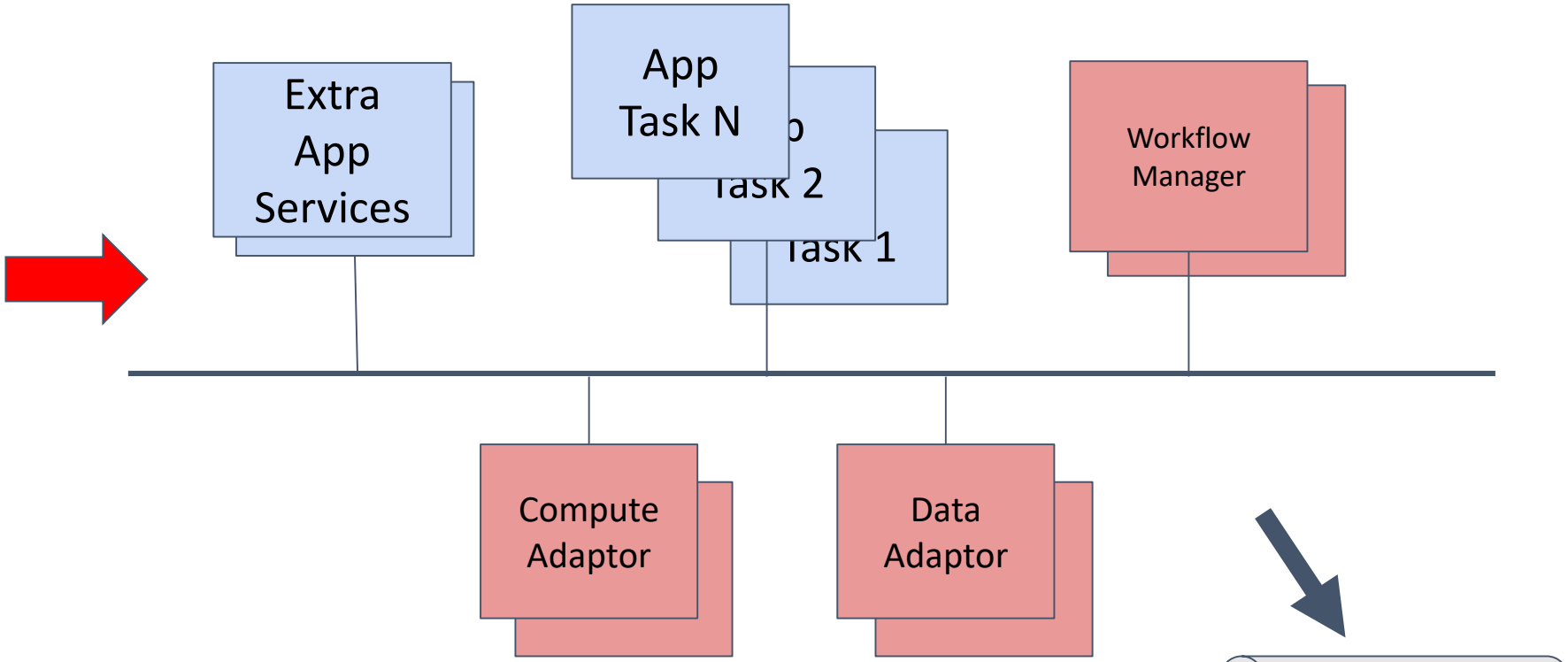
Querying files from the Prometheus adaptor through WebDAV service

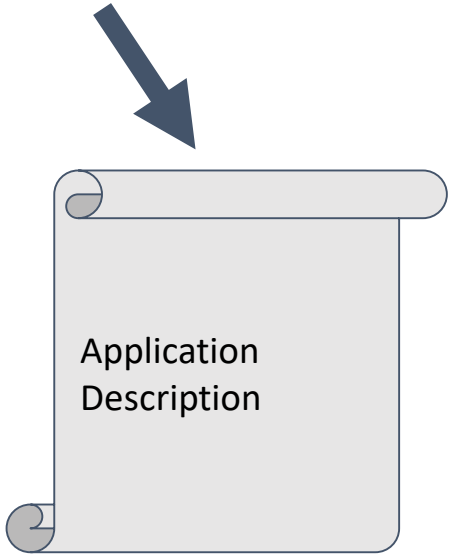


DISPEL graphical authoring and execution environment based on Eclipse

Workflow to infrastructure



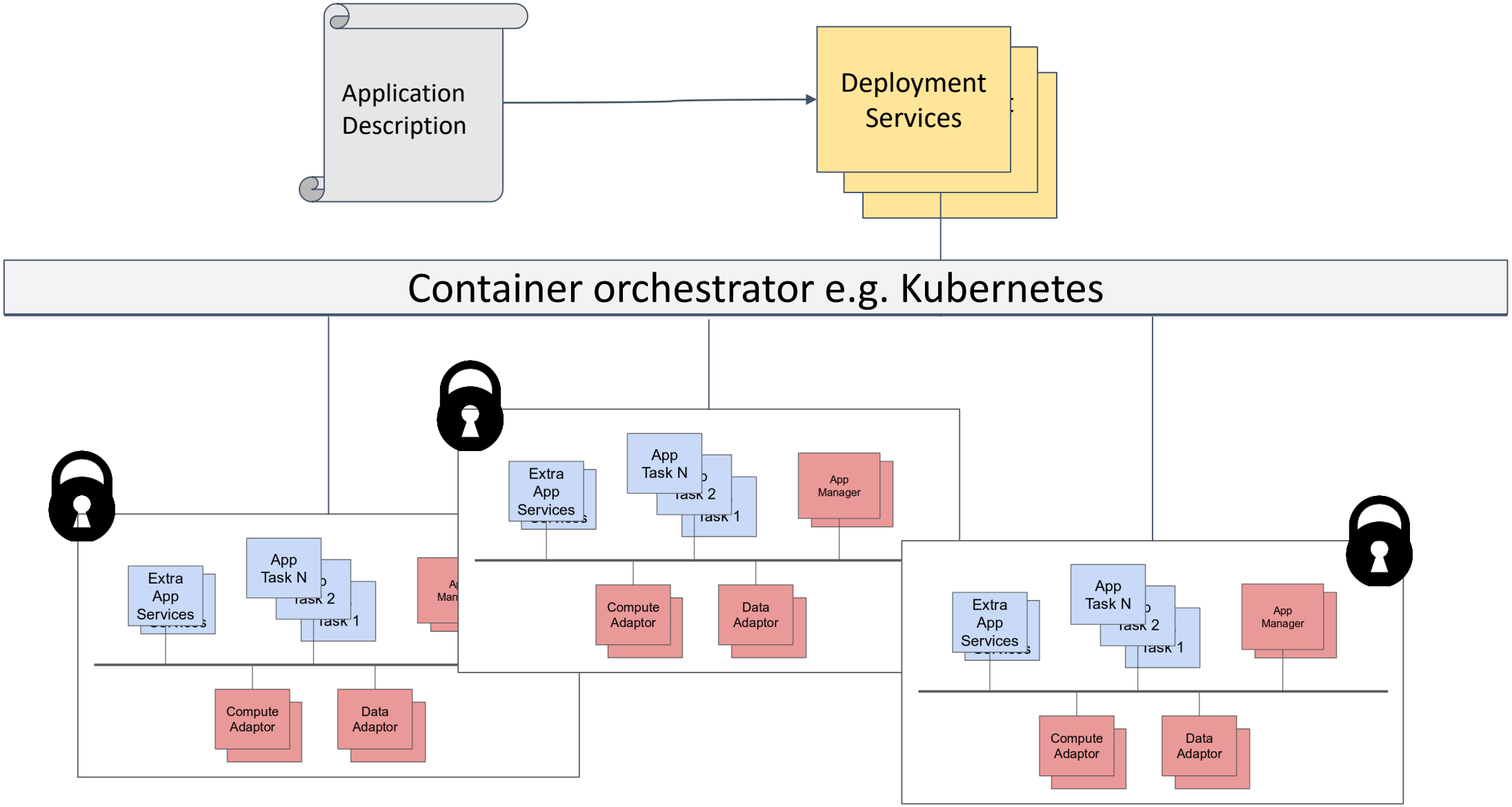




Application logic containers

Application specific management containers

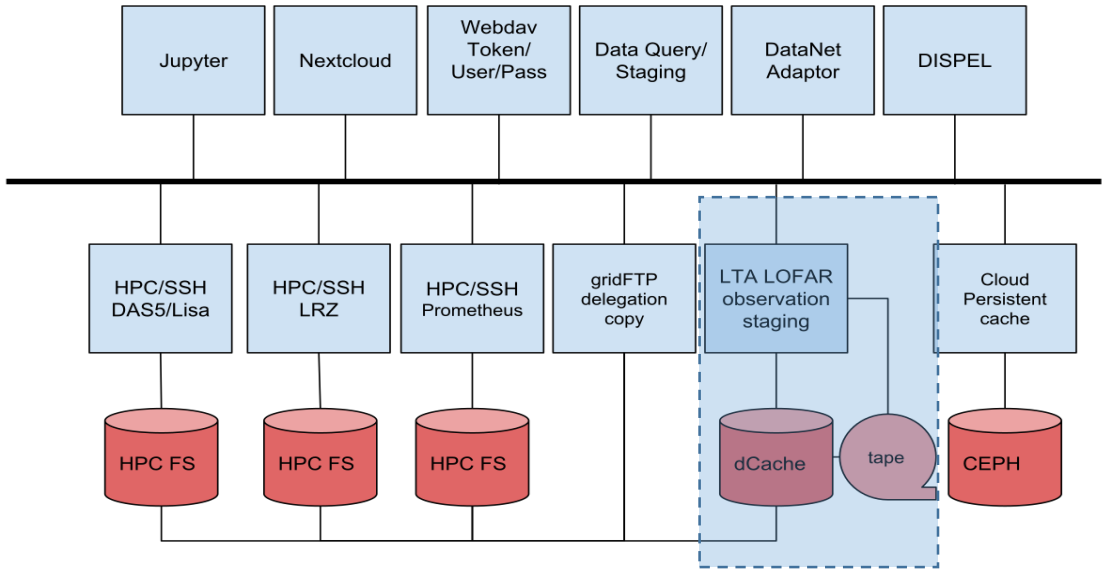
Deploy application infrastructure

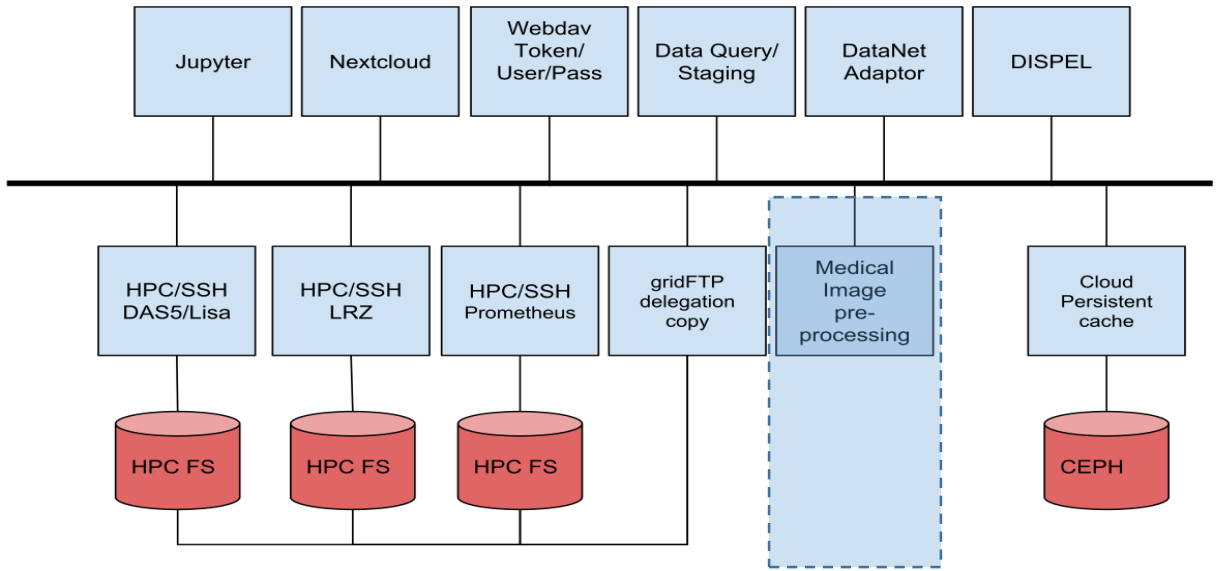


Data infrastructure

PROCESS Data infrastructure including data adaptors for UC#1 and UC#2:

- **Reuse** of container adaptors **across use cases**
- Ability to **add new application specific container** adaptors





PROCESS

PROviding Computing solutions for ExaScale ChallengeS

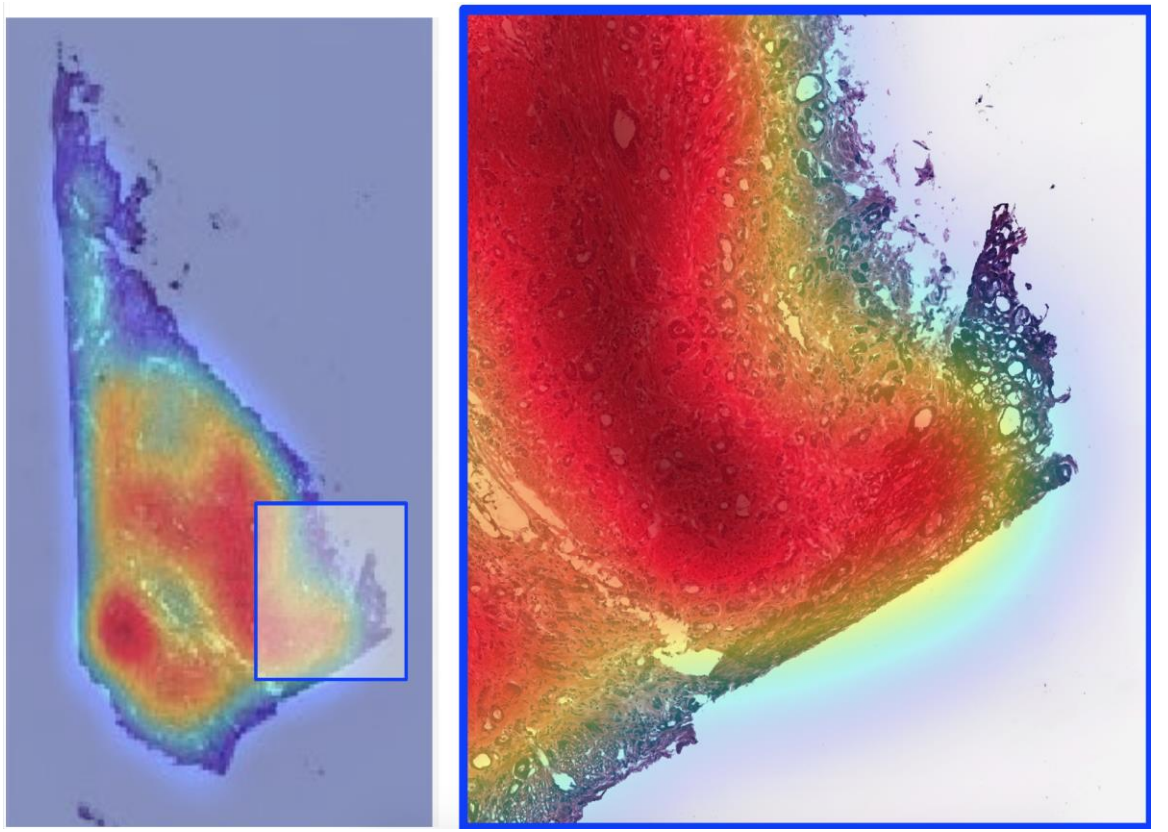
Use Case 1

Machine Learning in Medical Imaging

Hes·so

Haute école spécialisée

de Suisse occidentale, Switzerland



Use Case Scenario and Objectives

- Use of machine learning to analyse large histopathology images (>100,000x100,000 pixels)
  - Cut into small patches for treatment
  - Mainly for cancer care to highlight regions of interest
- Use of standard tools such as Keras, Tensorflow, … for Deep learning.

- Adapt the machine learning tools to large data centres and make them scale to improve the amount of training data and thus improve the quality of the models
  - Histopathology data is produced in massive quantities constantly
- Use a safe environment for possibly confidential data
- Have a simple user interface to test new pipelines

PROviding Computing solutions for ExaScale ChallengeS

Use Case 2

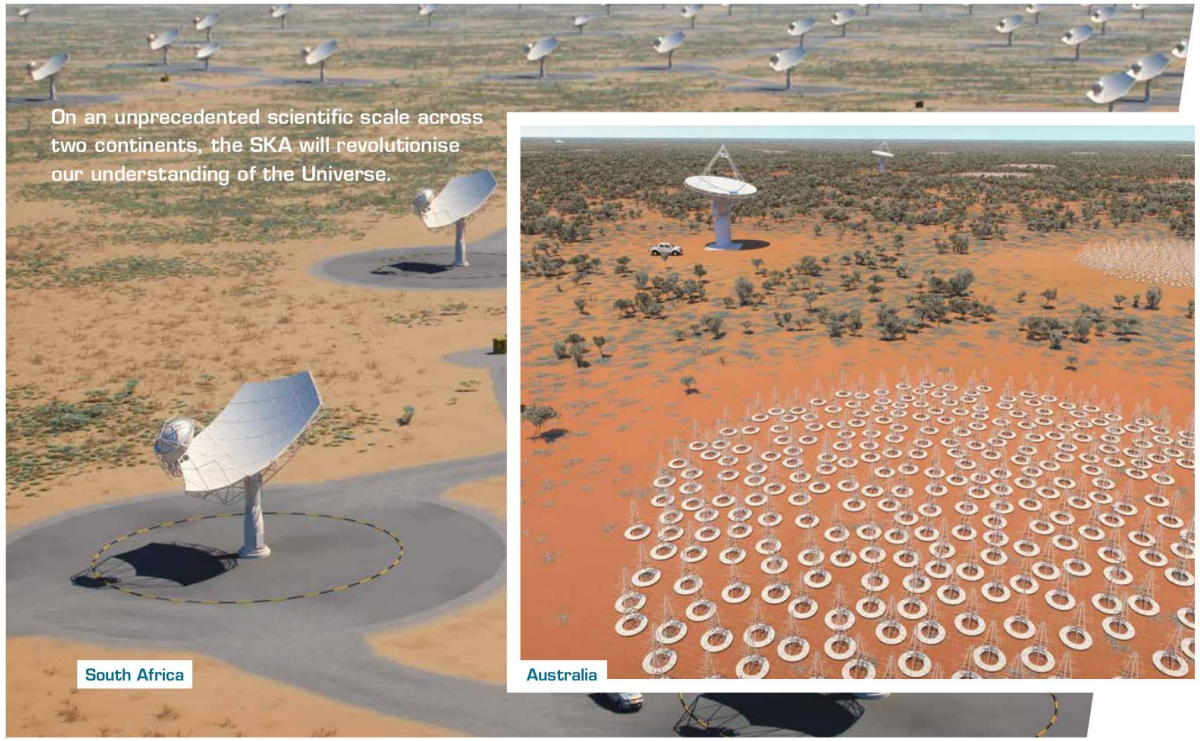
Analysis of Radioastronomy Observations LOFAR / SKA

Stichting Netherlands eScience Center, The Netherlands

Analysis of Radioastronomy Observations

Images courtesy of:
ASTRON

LOFAR: Low Frequency Array radio telescope – is a “distributed software telescope” consisting of ~88.000 antennas in ~51 stations scattered over Europe. It produces up to **35 TB/h of intermediate data** (visibilities) which is stored for further analysis.

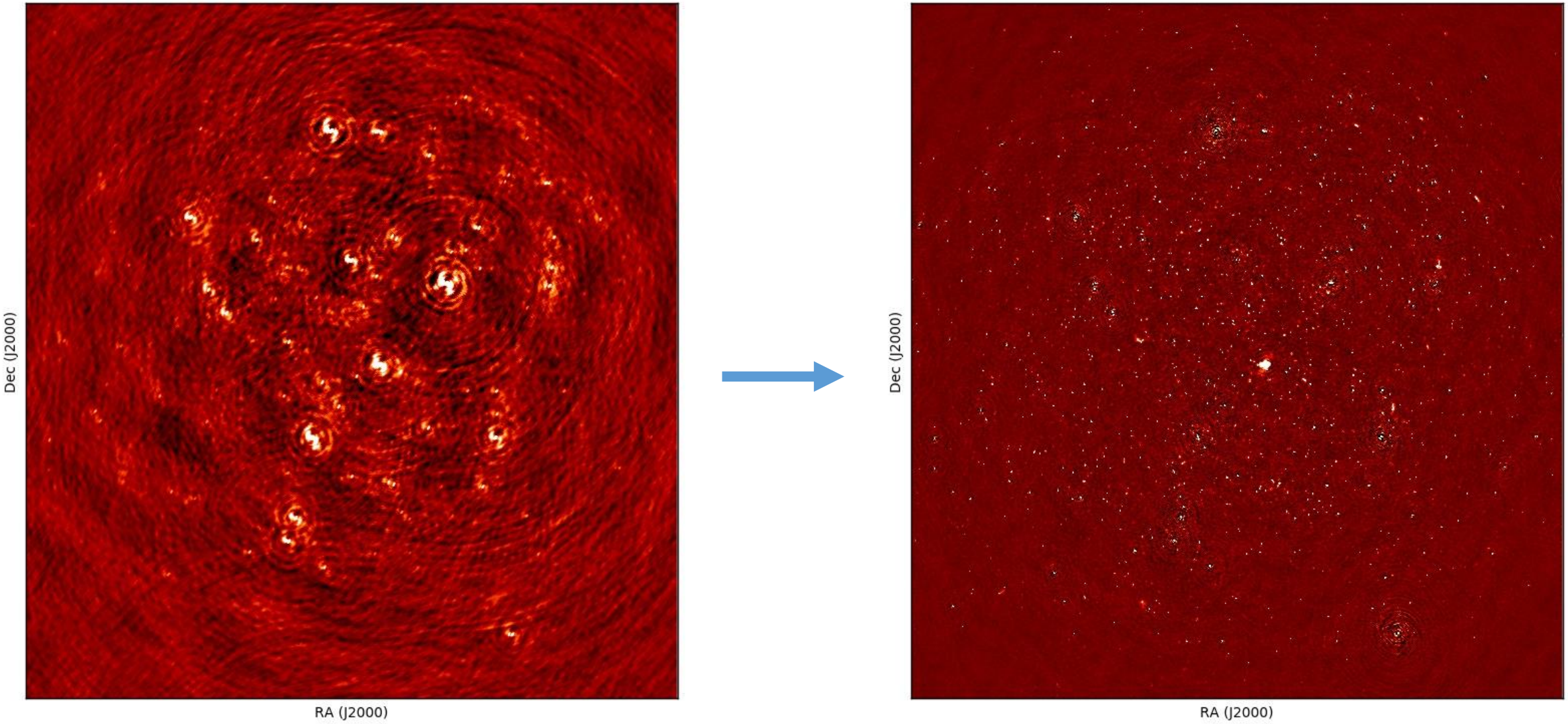


SKA: Square Kilometer Array (Operational in 2022+)

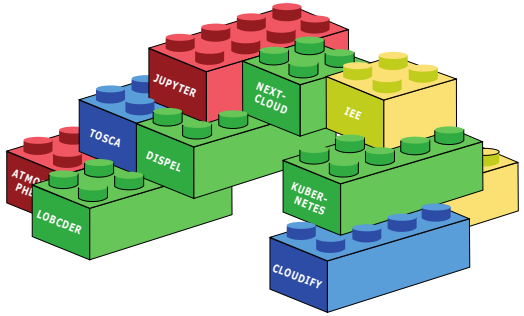
130K ~ 1M (LOFAR-style) antenna in Australia + 200 ~ 2000 dishes in South Africa. Wider frequency range and higher sensitivity and survey speed than existing telescopes.
Zettabytes/year **raw** data: **130~300PB/year of correlated data**

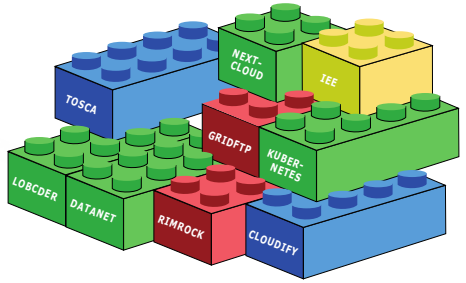
Huge data and processing problem

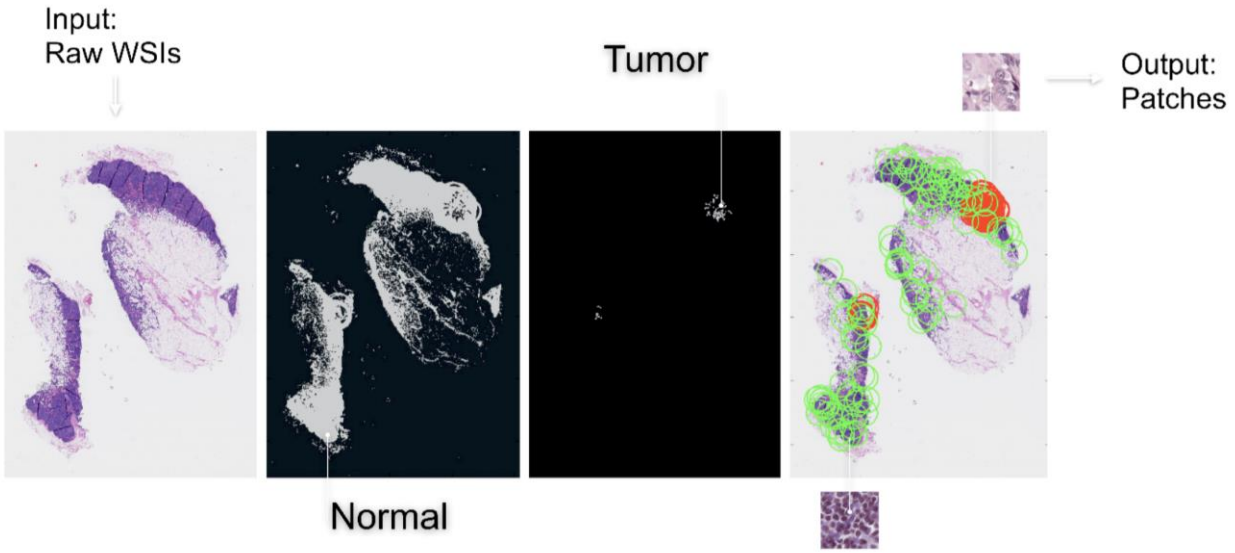
Effect of Processing

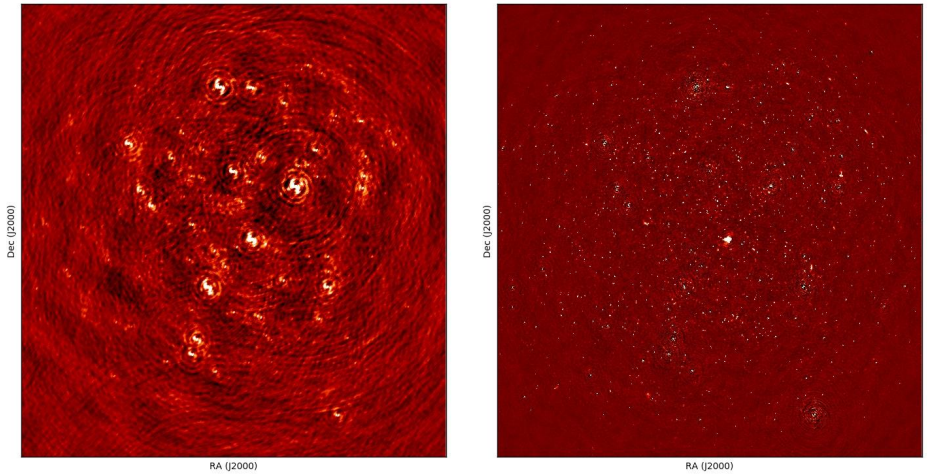


UCs prototypes based on modular services

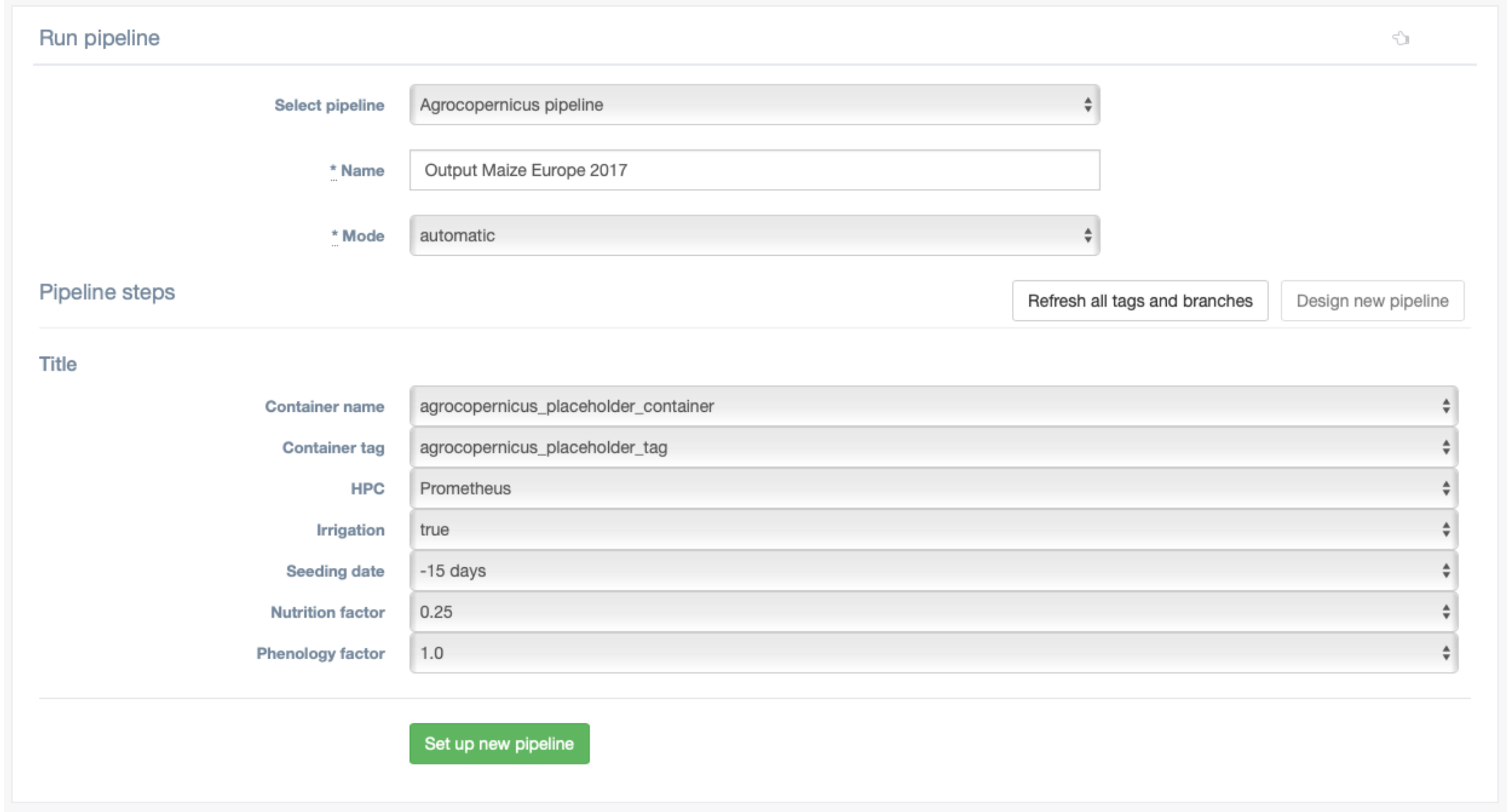




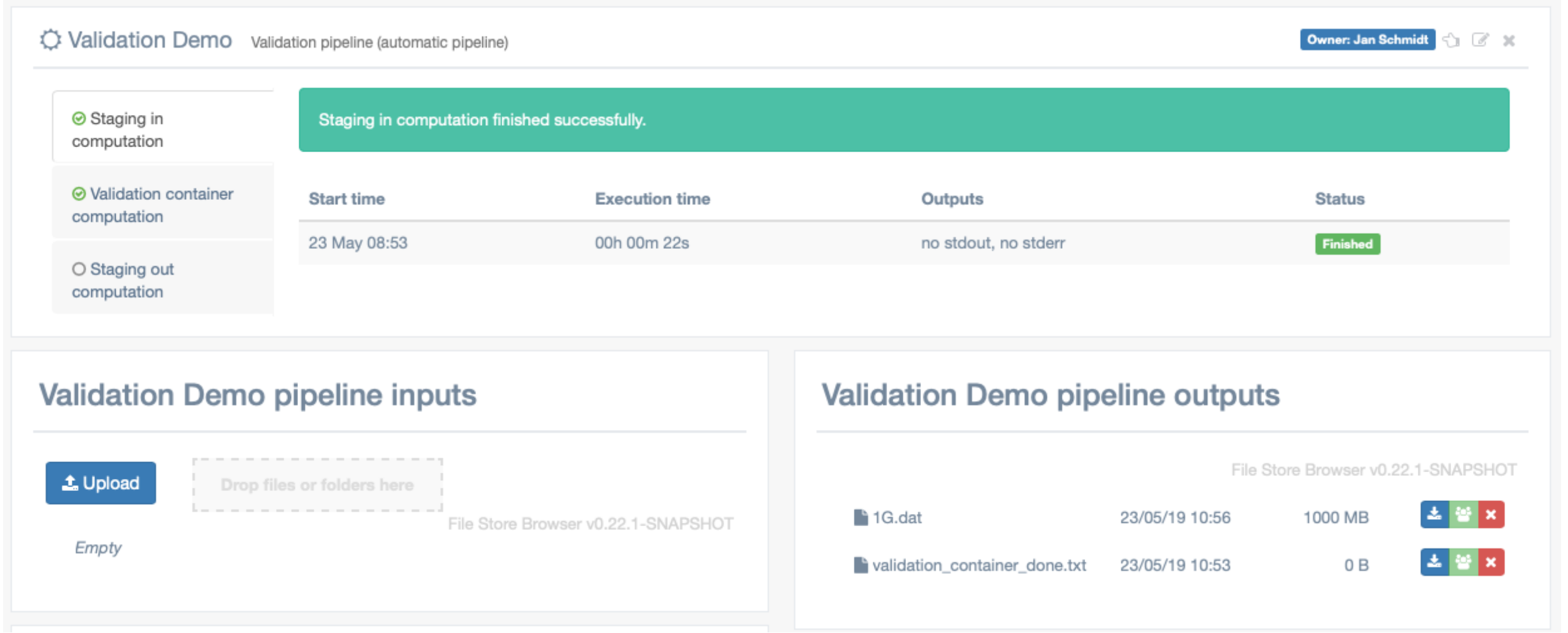




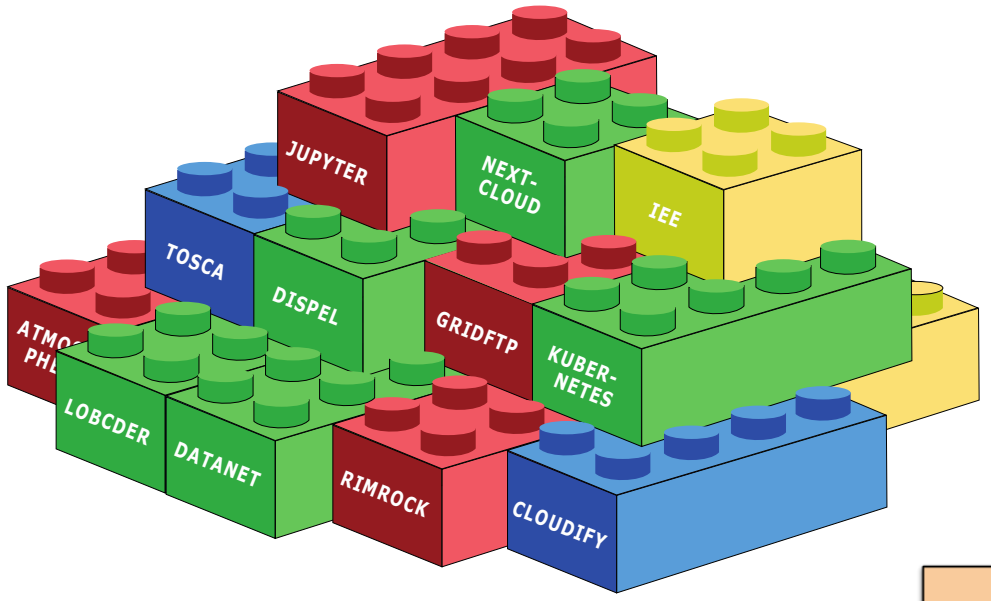
Pipeline and Workflow Configuration Portal

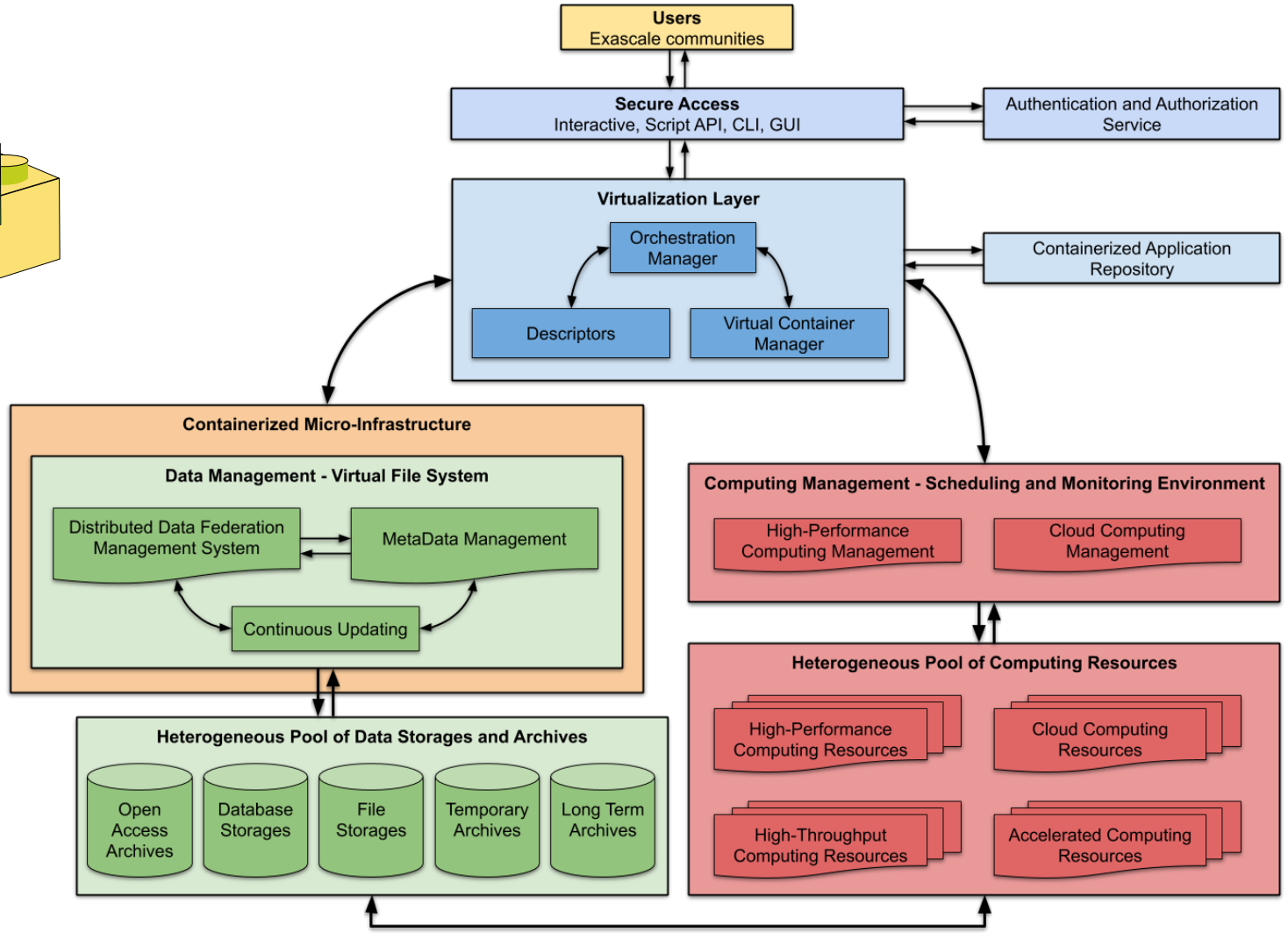


Pipeline Deployment and Output

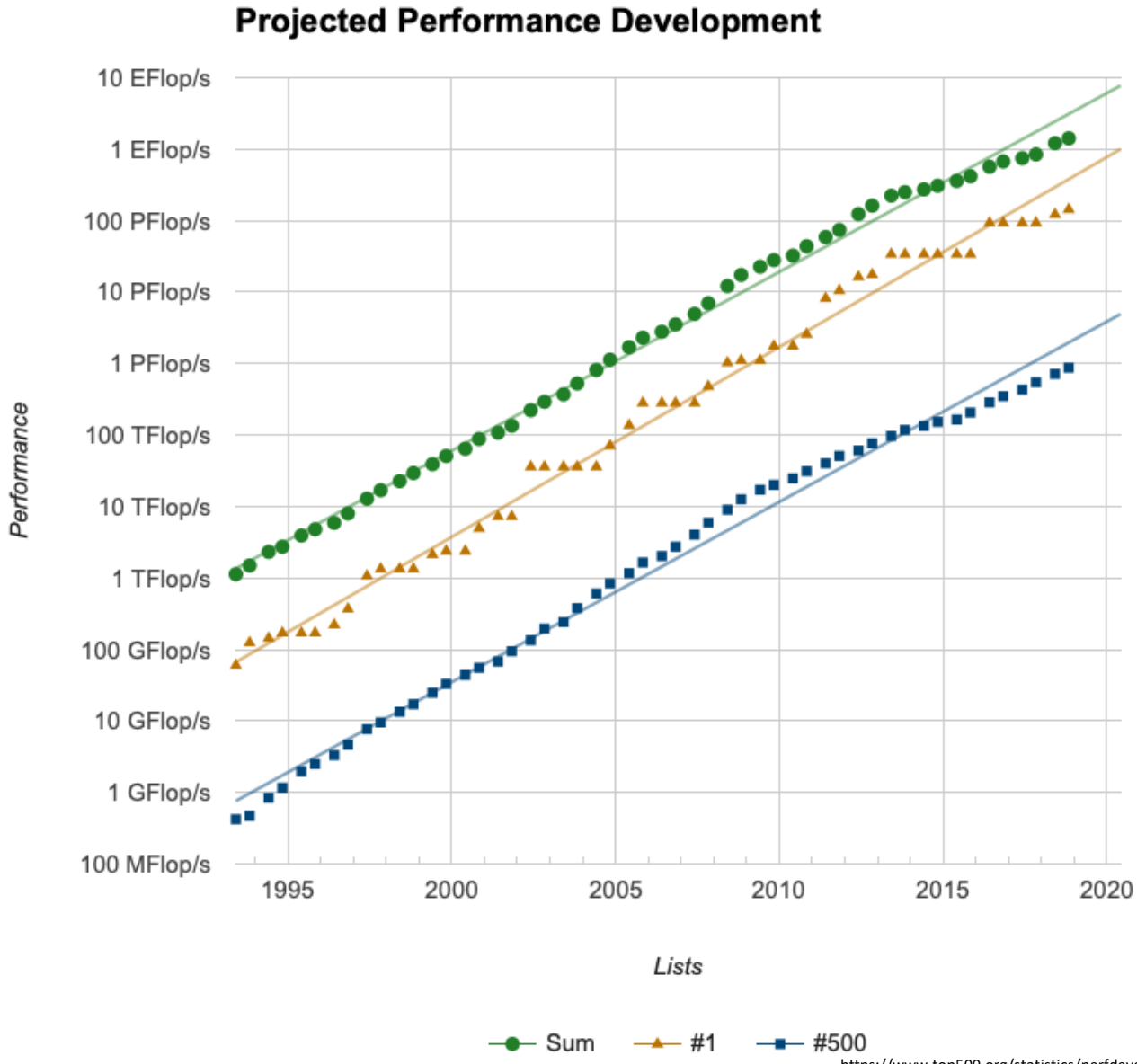


Towards an Exascale-ready Solutions

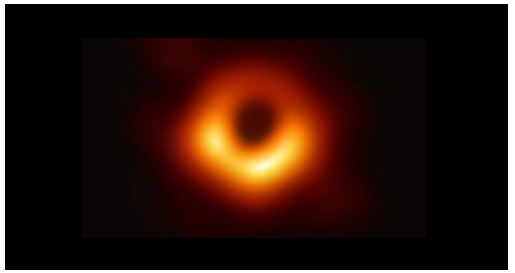


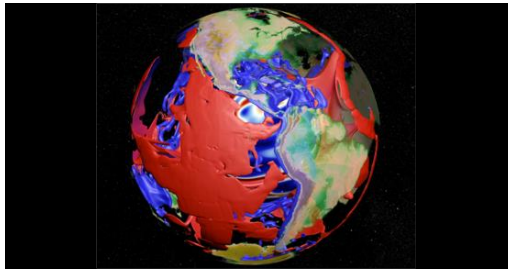


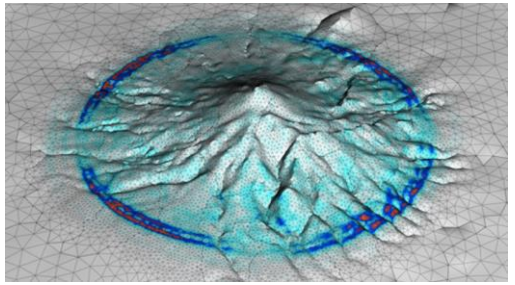
Enabling Exascale

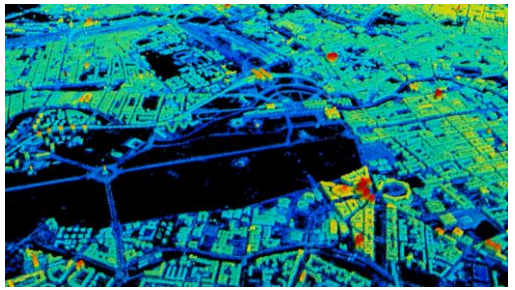


https://www.top500.org/statistics/perfdevel/









PROCESS

PROviding Computing solutions for ExaScale ChallengeS

Maximilian Höb

hoeb@mnm-team.org

eScience Workshop

Platform-driven e-Infrastructure Innovations

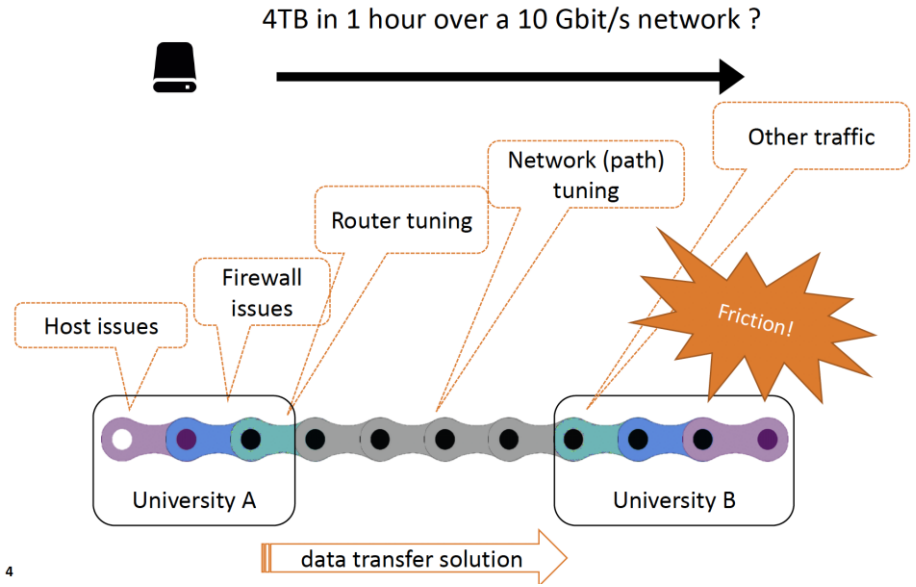
September 24, 2019, San Diego, USA

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 777533.

Data Transfer Nodes

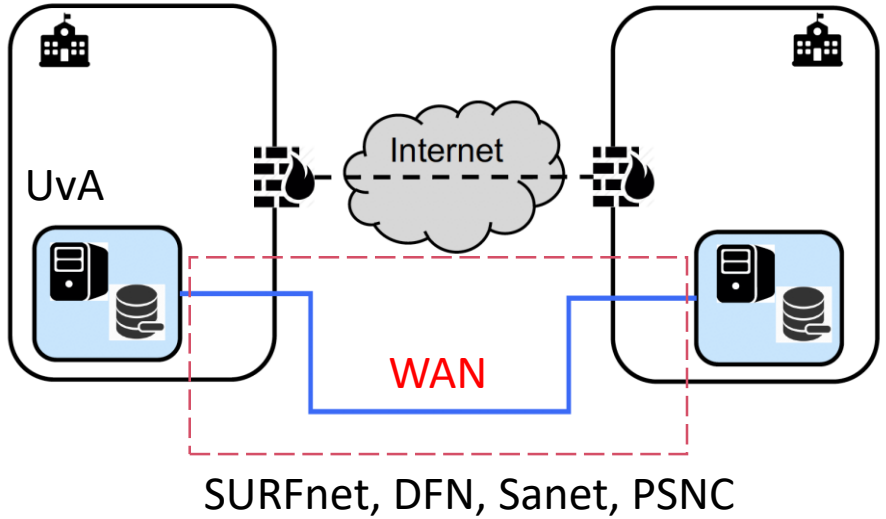
- Data Transfer Nodes (DTN)
  - Needed for optimal data transfers between different data centres

- Problems
  - Limitation of TCP
  - Firewalls „are evil“

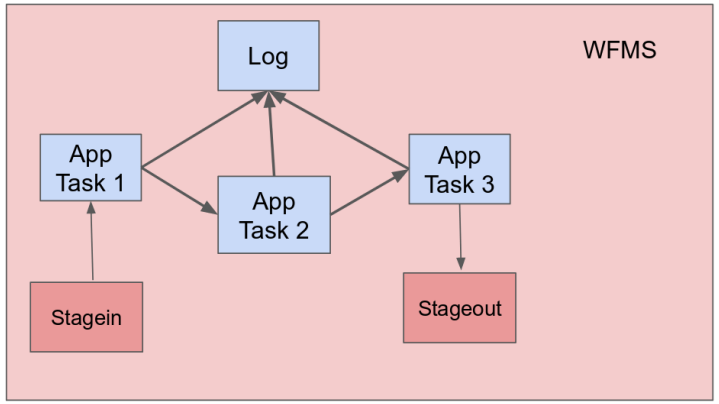


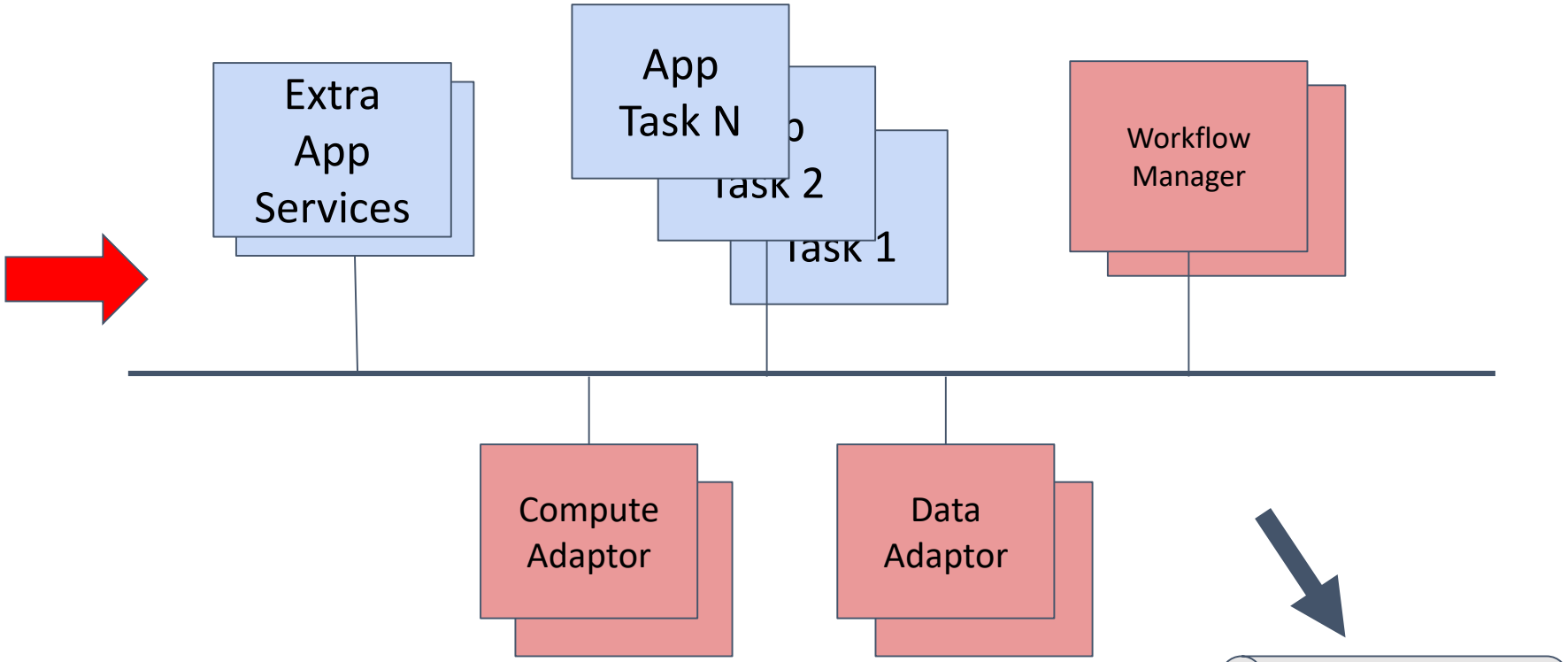
Source: Peter Hinrich SURFnet “Problems with data transfers“

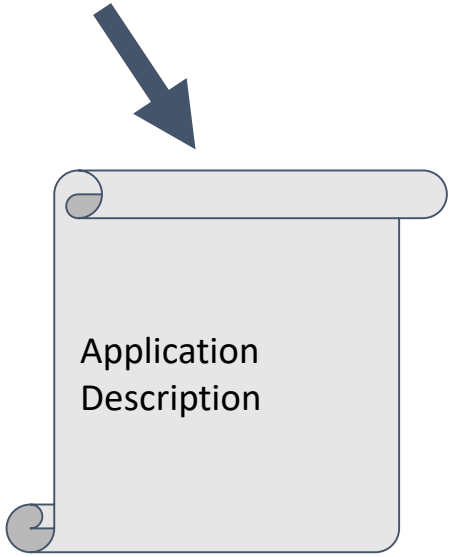
Dedicated, optimized Data Transfer Nodes (DTN)



Workflow to infrastructure



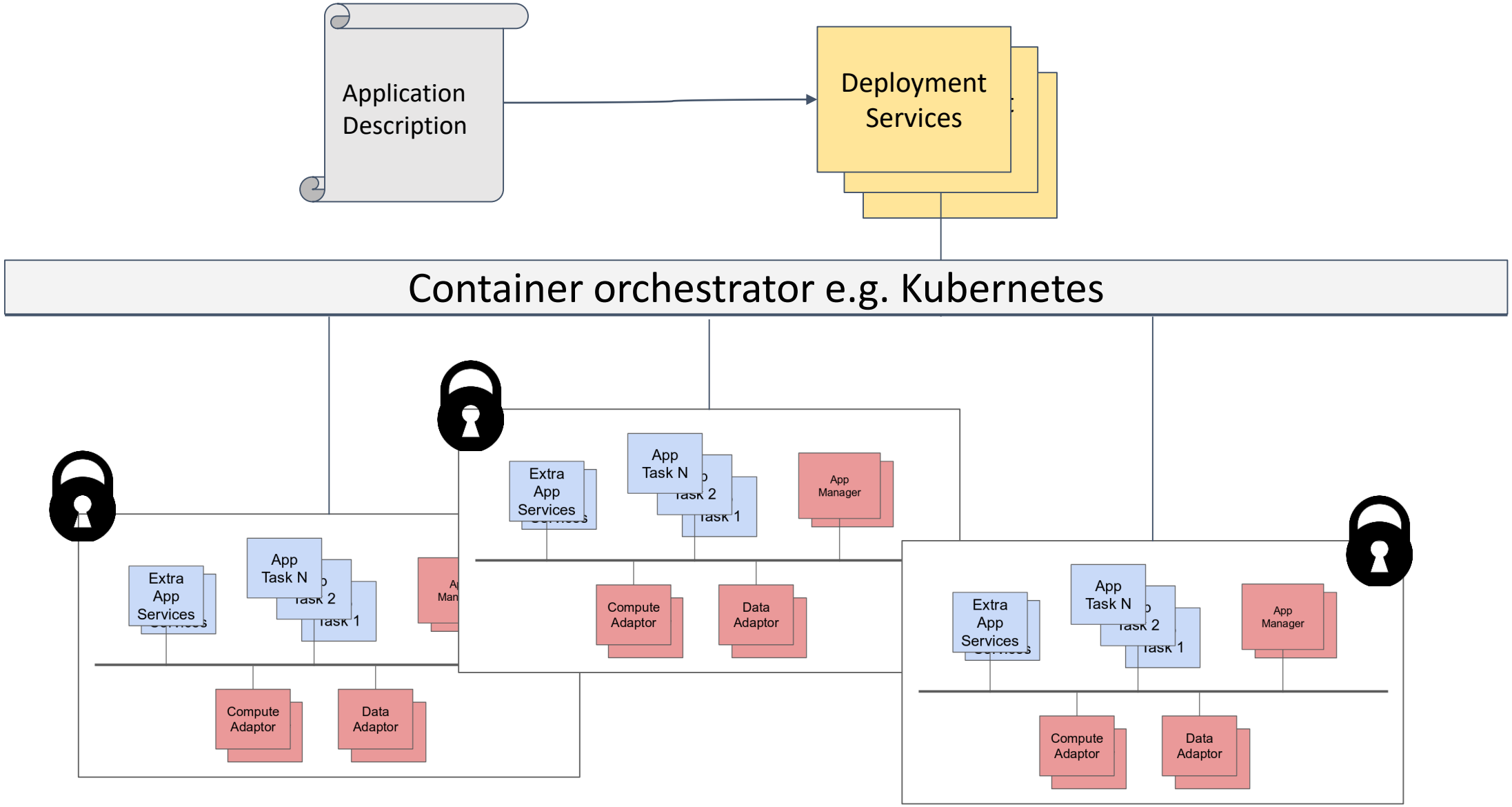




Application logic containers

Application specific management containers

Deploy application infrastructure



Some containerized services

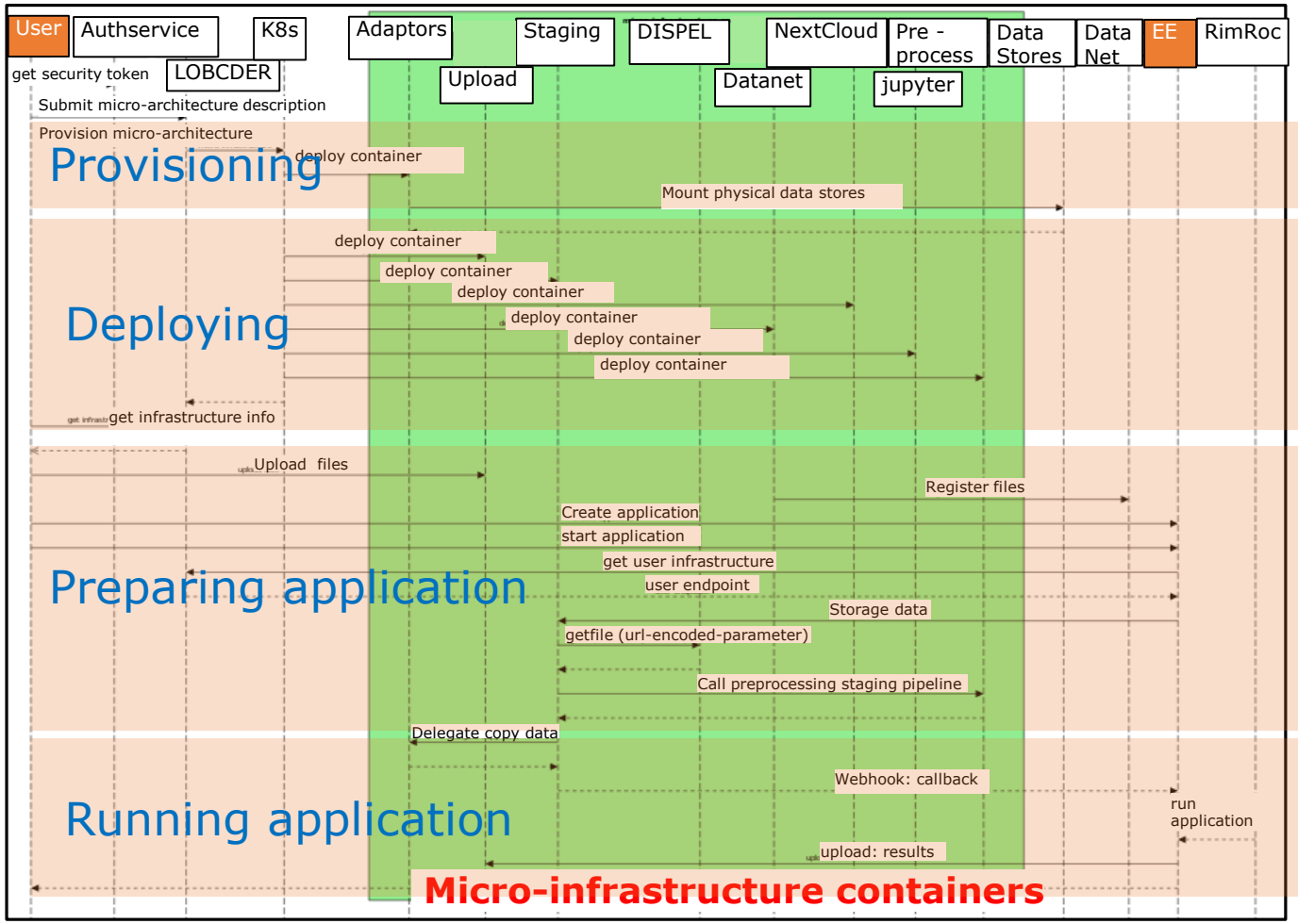
- Data-staging service (in progress)
  - Batch system for data transfers between sites
  - Minimize data copies e.g. JIT data transfers
  - Containerize protocol handlers (adaptors)
- Data adaptors (in progress)
  - SCP to SCP
  - gridFTP to gridFTP
  - FTS3 to FTS3
  - Define a common container interface
- Compute offload (to do)
  - Compute scheduler to decide where to run processing; on the application infrastructure or offload to an HPC site.

Evolution of Data Management

From: **Functional** Architecture → **Technology** selection → **Architecture** design

TO: specification of the interaction of the different services

Containers	Description
WebDAV container	Protected a public WebDav entry point.
Token-based WebDAV	meant for access by computing services.
DataNet-adaptor	performs operations on metadata
Staging service	stage data just-in-time on the HPC file systems.
DISPEL	access to data (pre)processing environment.
Jupyter container	access data through Jupyter notebook.
NextCloud	view data in Dropbox fashion.



Workflow of the Use Case

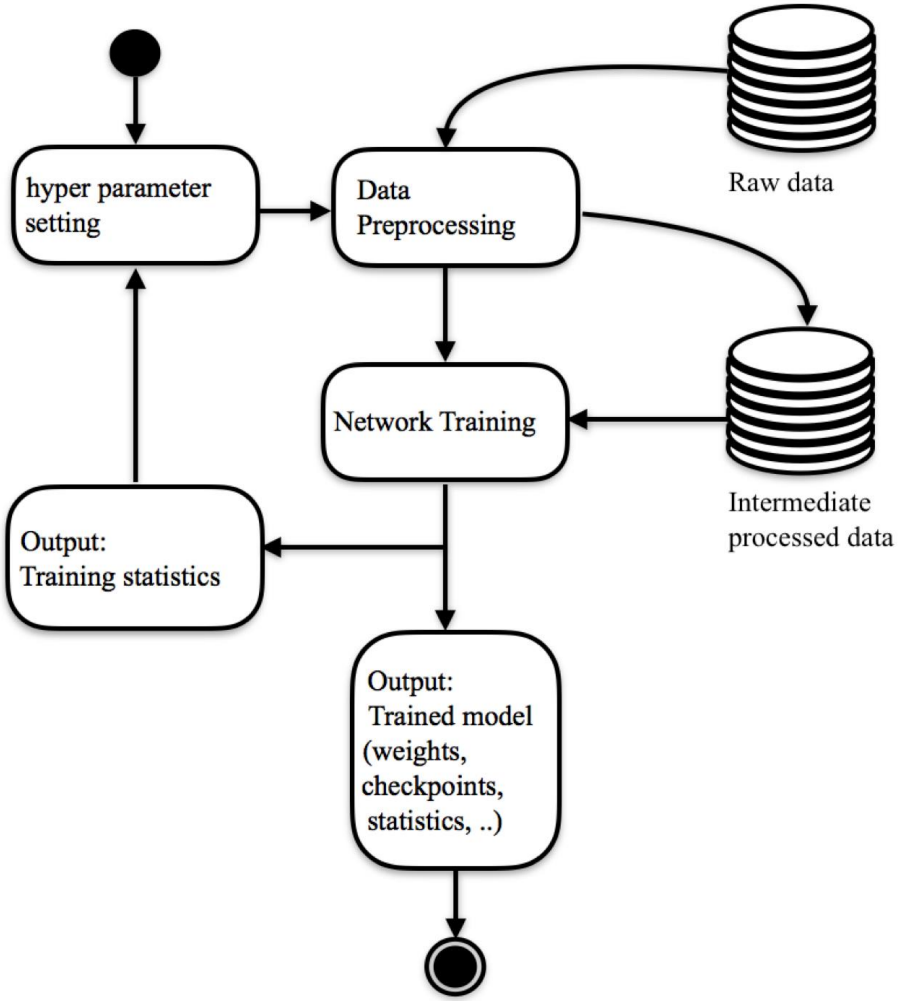


Figure 3: Network Training workflow. The solid dot represents the starting point in the workflow and the dot sorrounded by a circle represents the ending point.

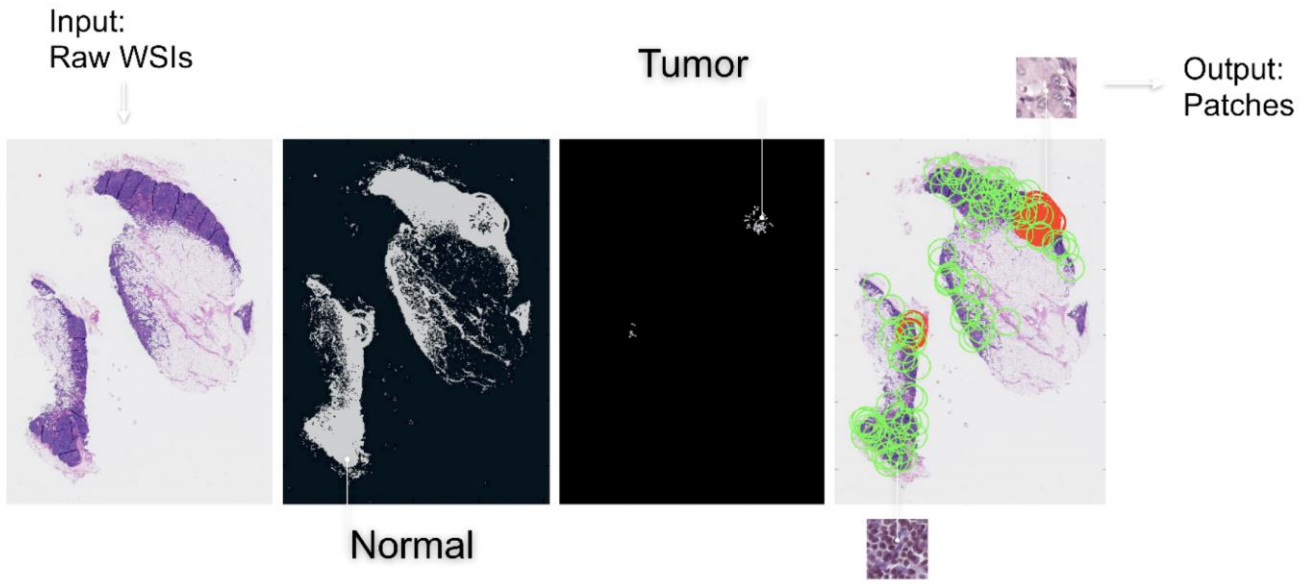


Figure 1: WSIs preprocessing pipeline: Normal tissue and tumor tissue masks are extracted and high-resolution patches are sampled from the selected regions.