PROviding Computing solutions for ExaScale ChallengeS

Towards Exascale-ready Data Service Solutions

Maximilian Höb

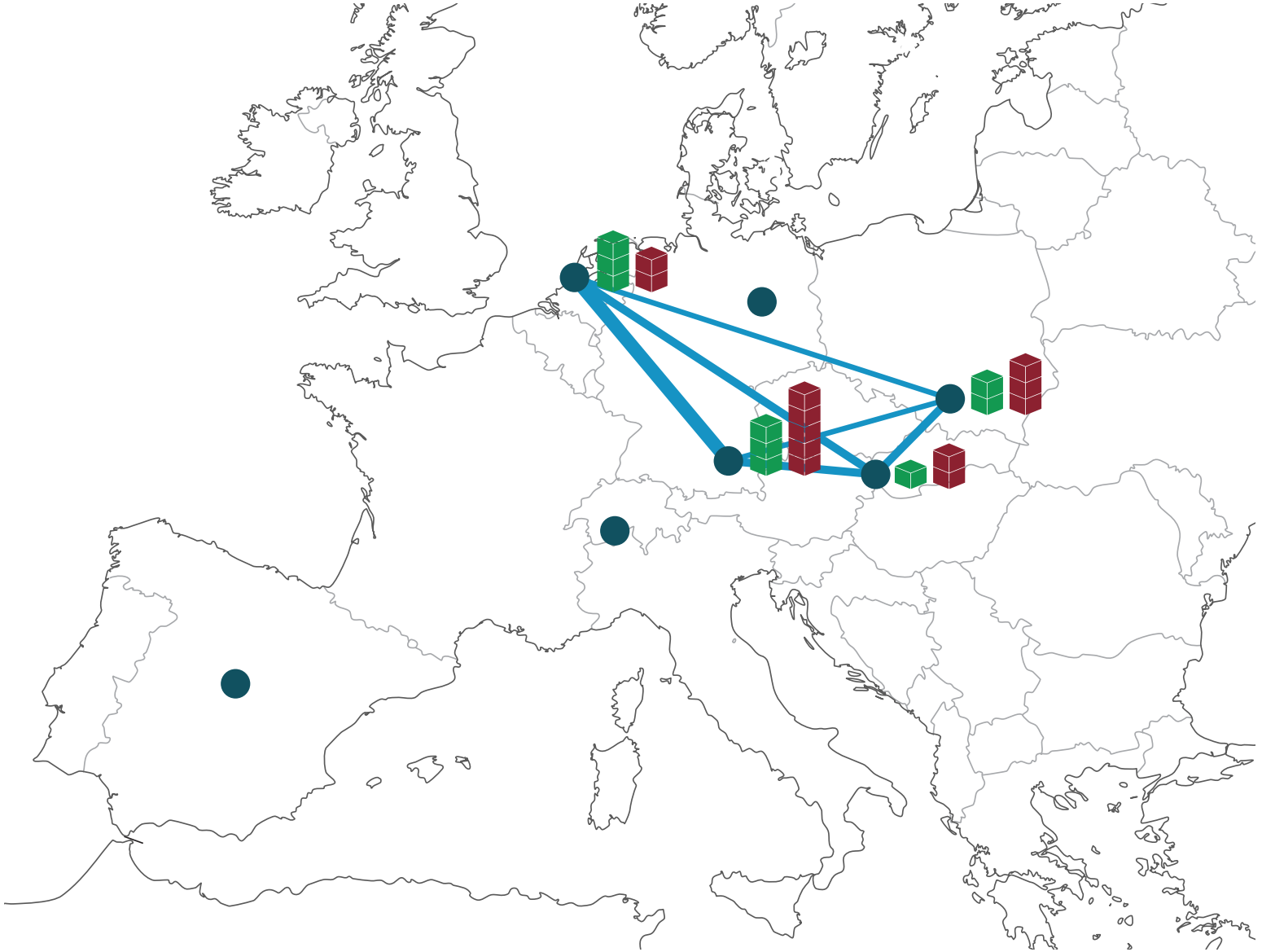
October 22, 2019

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



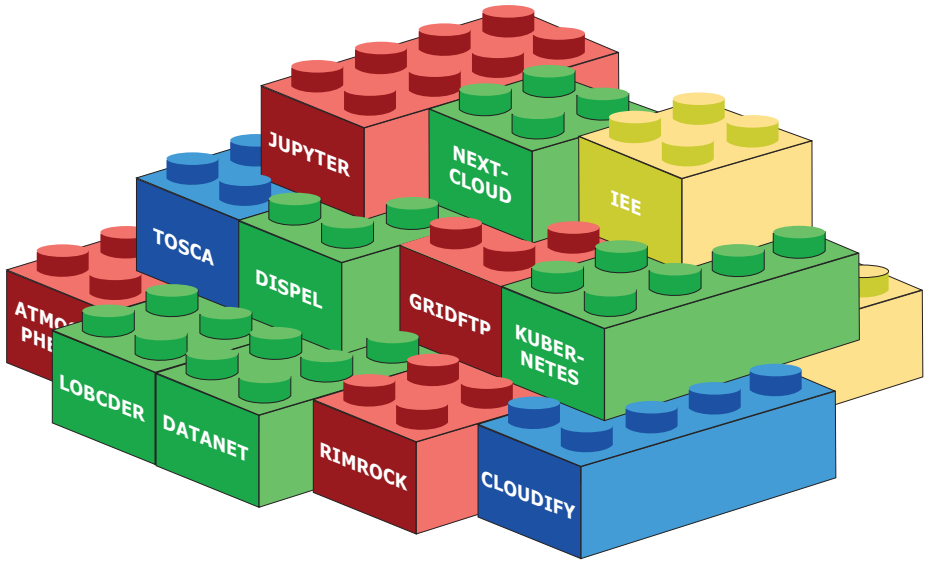
Partner's location

Storage Resources

Compute Resources

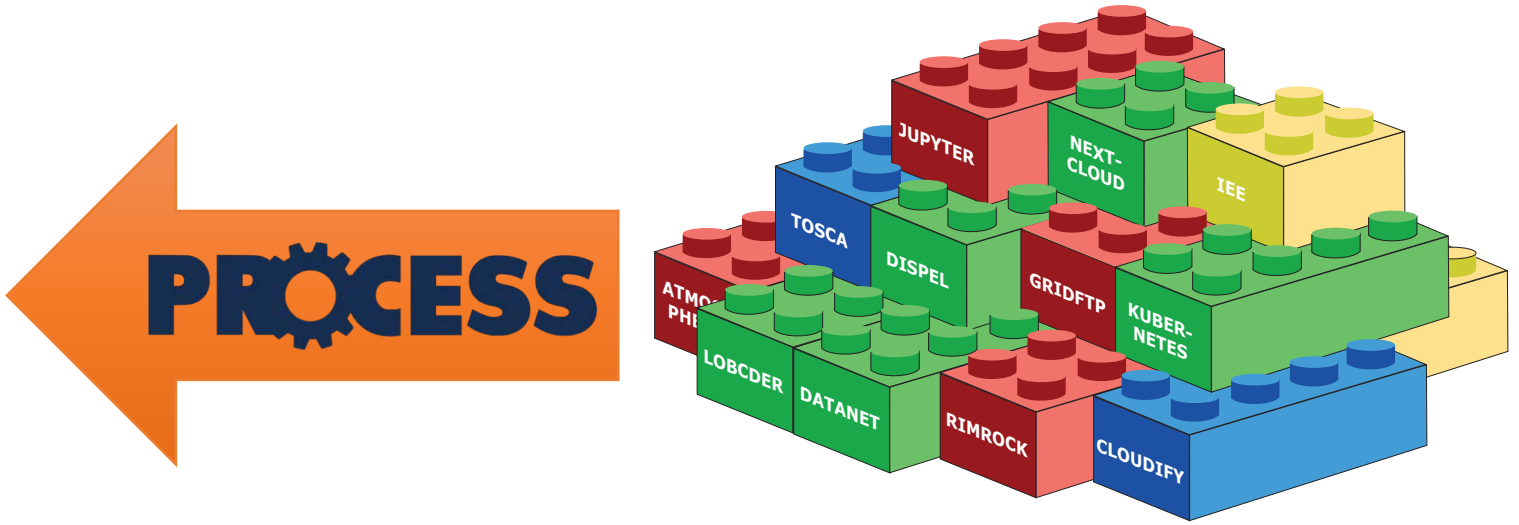
Vision of PROCESS

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.



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

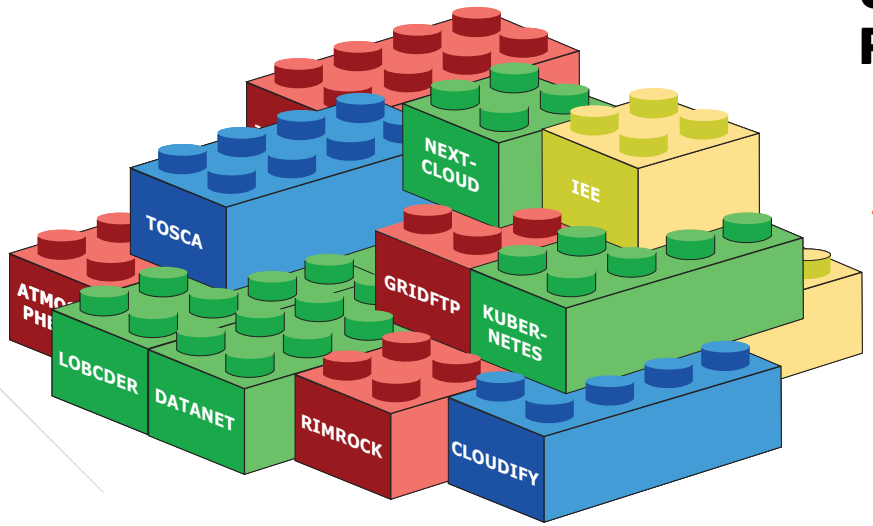
SuperMUC-NG
Leibniz Supercomputing Centre Munich



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

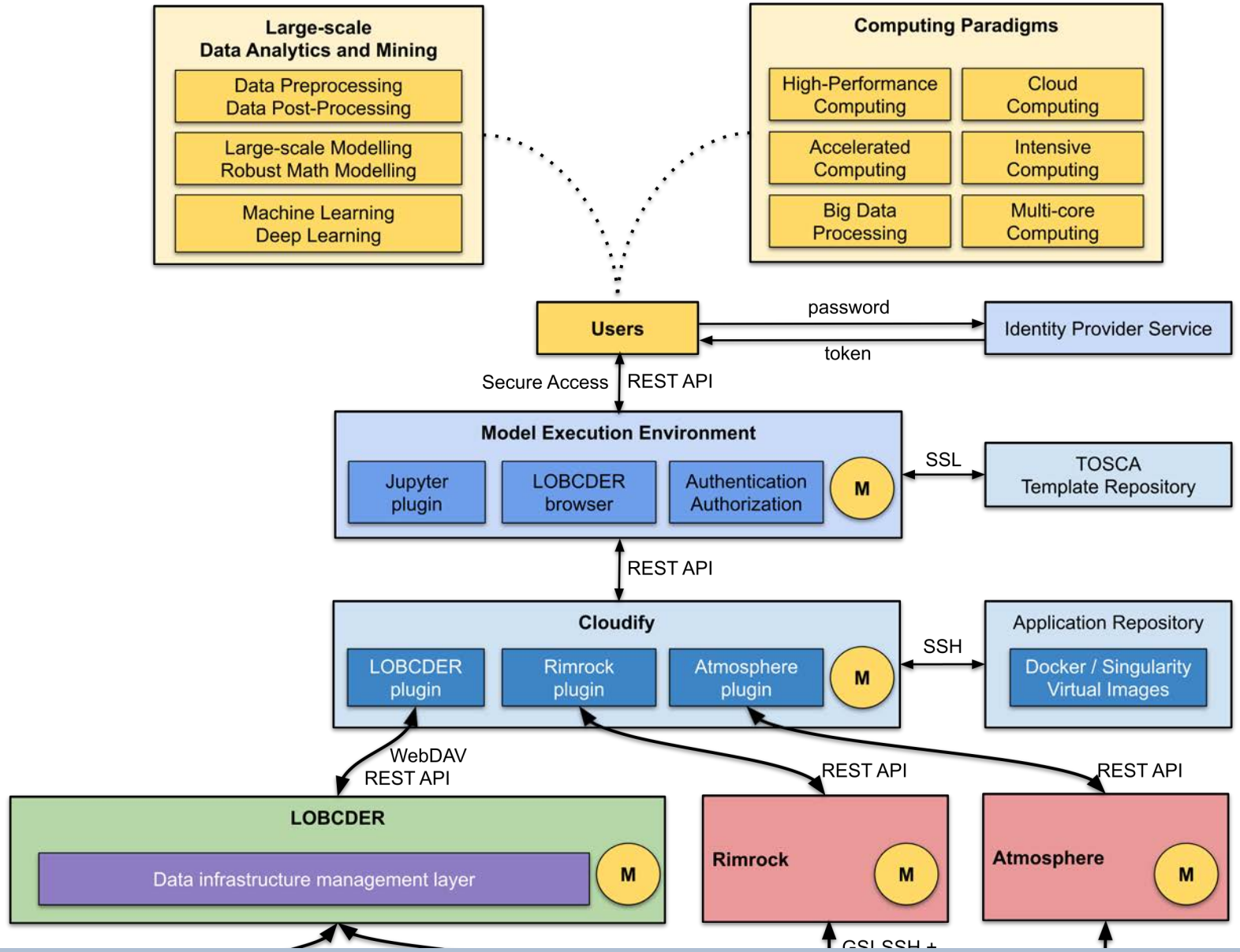


UC#3: Supporting innovation based on global disaster risk data

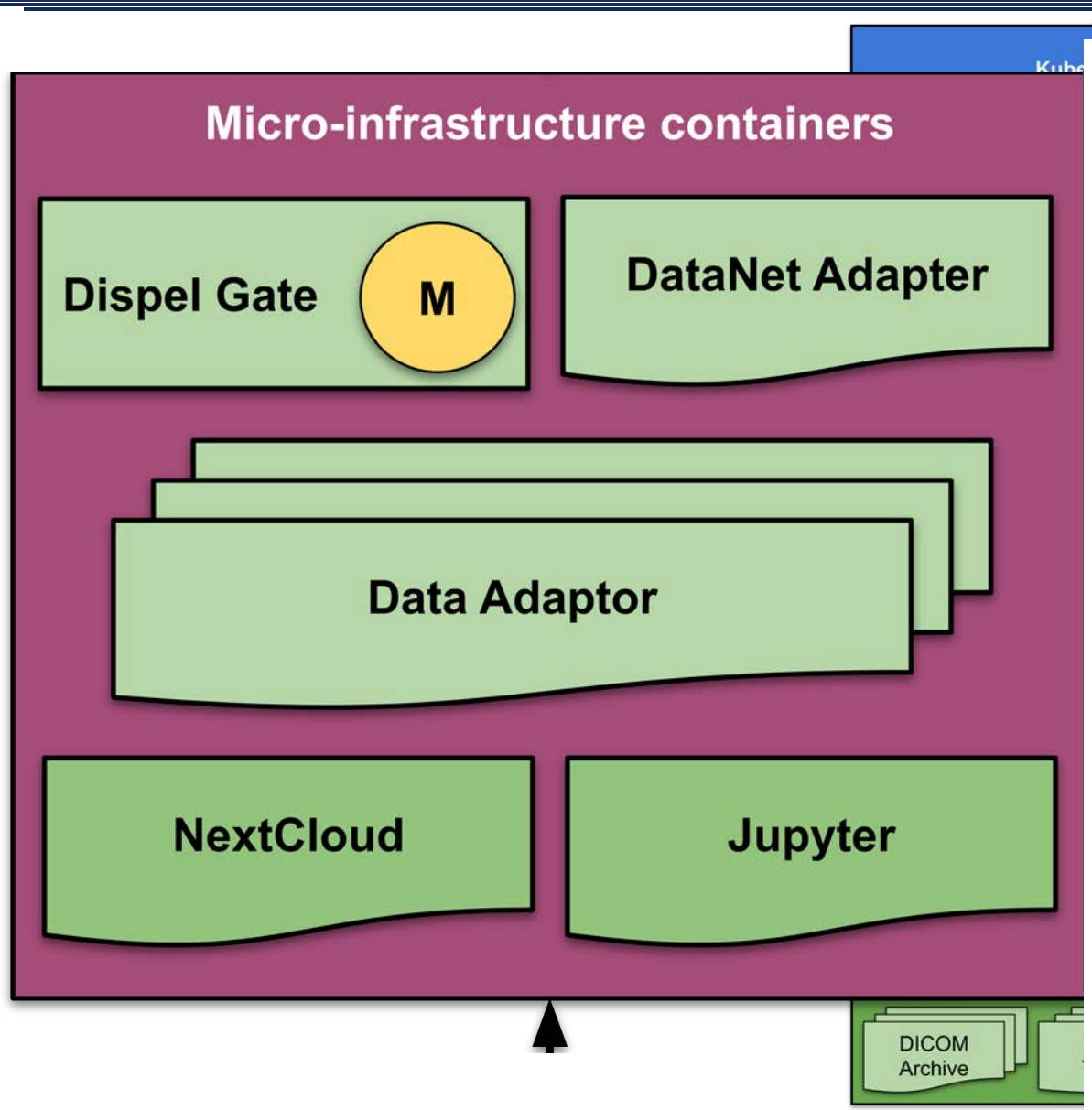
UC#5: Agro-Copernicus (correlating data between simulation and observation)

UC#4: Ancillary pricing/airline revenue management

PROCESS Architecture 1/2



PROCESS Architecture 2/2



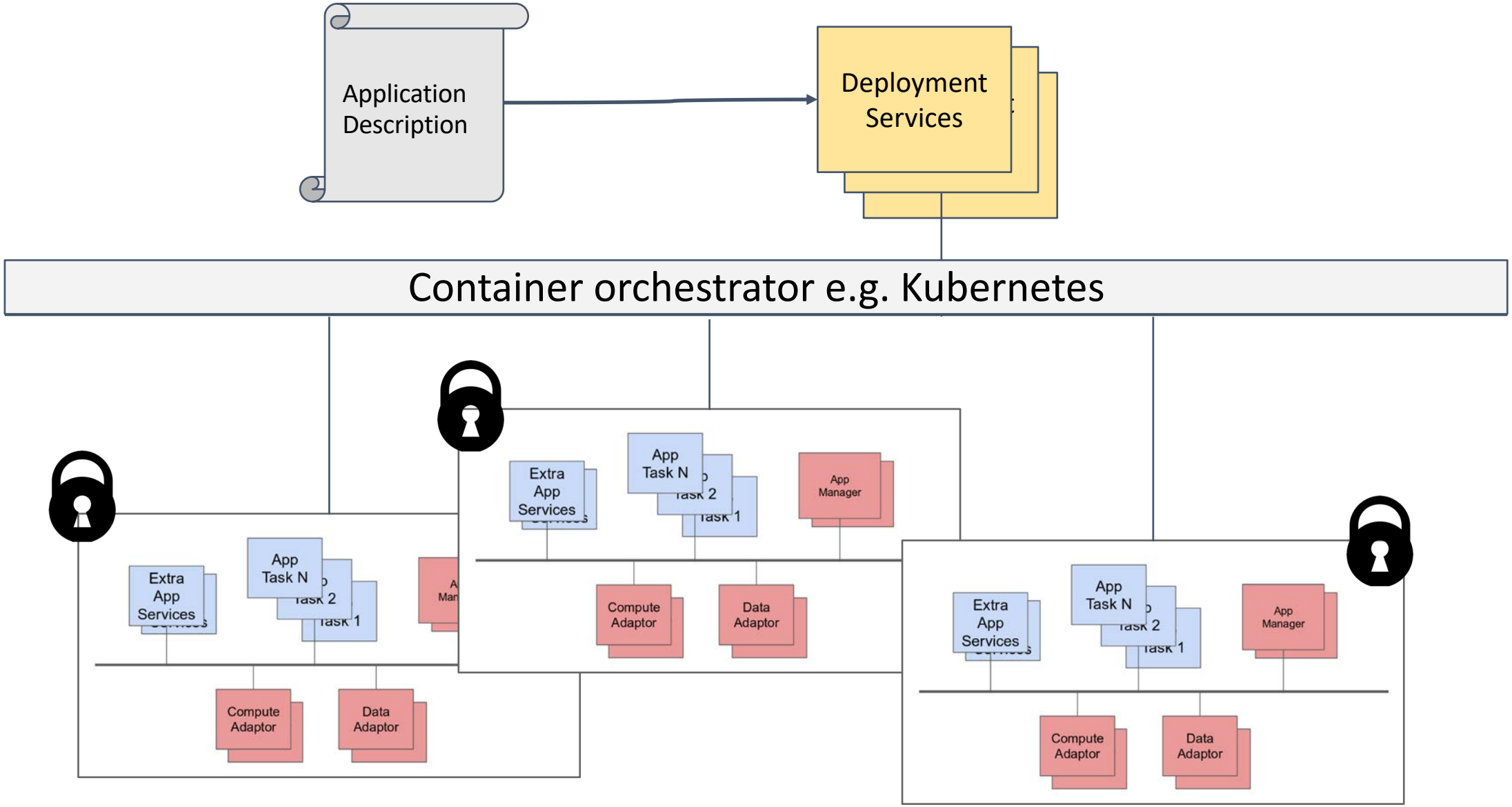
Approach: Tiered system with a layer of virtual (data) nodes facilitating:

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

Through → **A programmable micro-infrastructure**

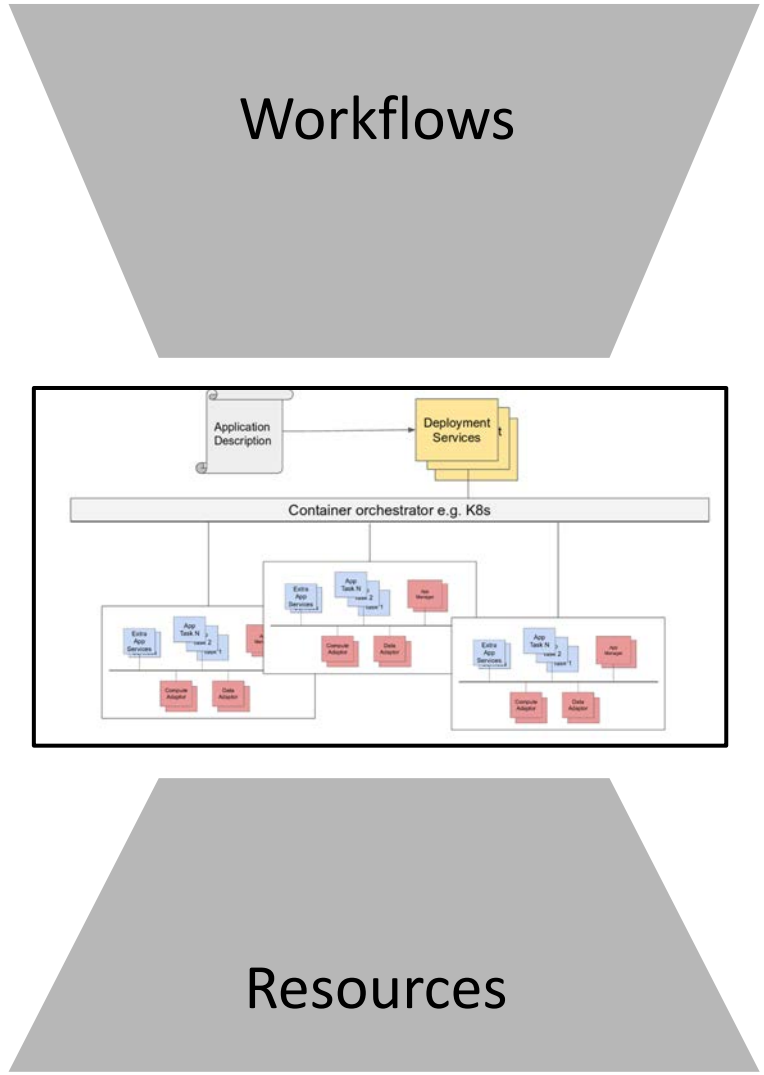
- Independent
- Scalable
- Distributed

Deploy application infrastructure

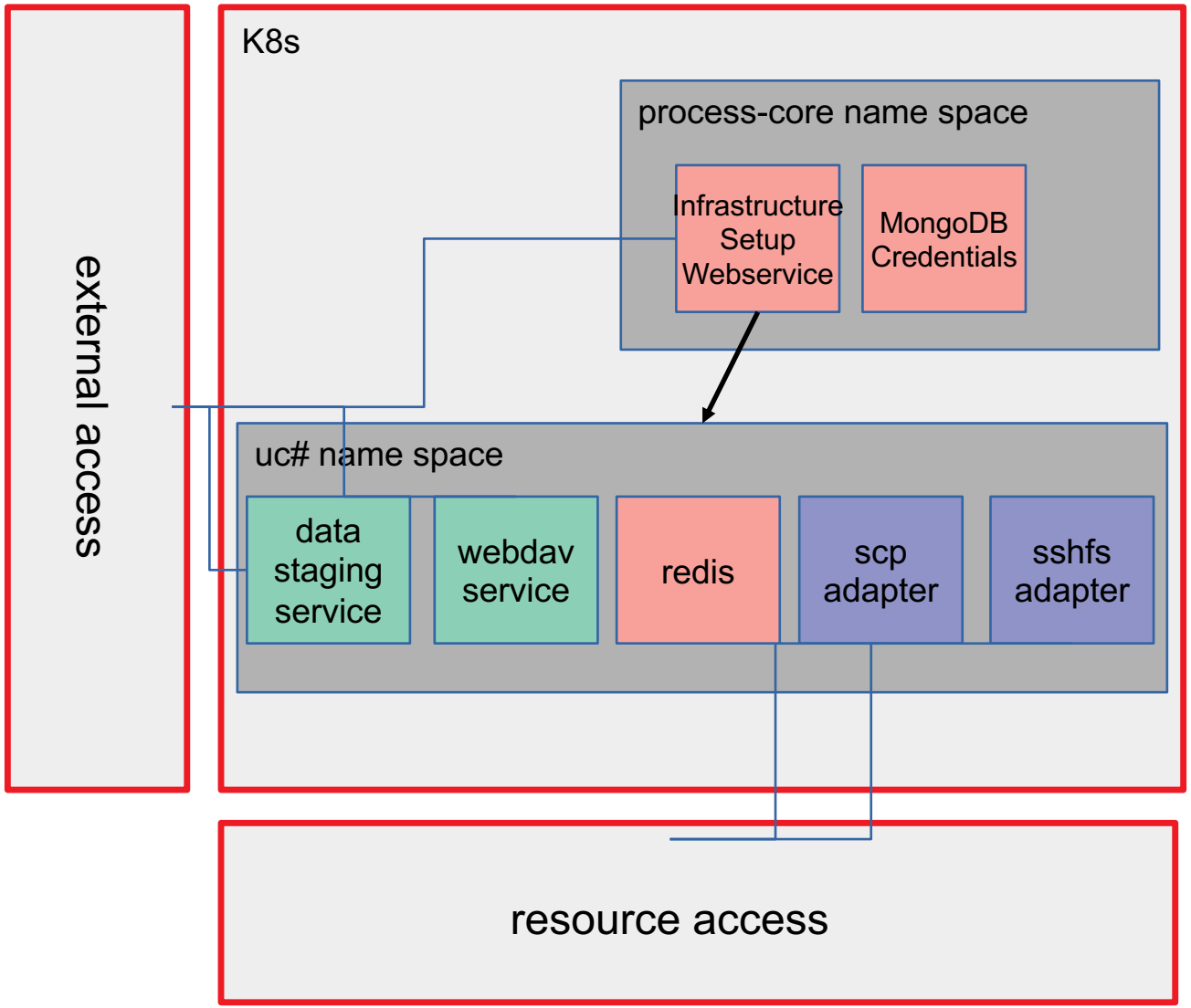


Data Service Middleware

- **Middleware** becomes a programmable distributed hyper-converged infrastructure.
- **Applications** define their own micro-infrastructure.
- Containers allow applications to be **portable** and **scalable** between different physical infrastructures.
- New **generation** of middleware exploits virtualized compute, storage and network.



Microinfrastructure Architecture



UC external API containers

LOBCDER core containers

UC resource adapters

PROCESS

PROviding Computing solutions for ExaScale ChallengeS

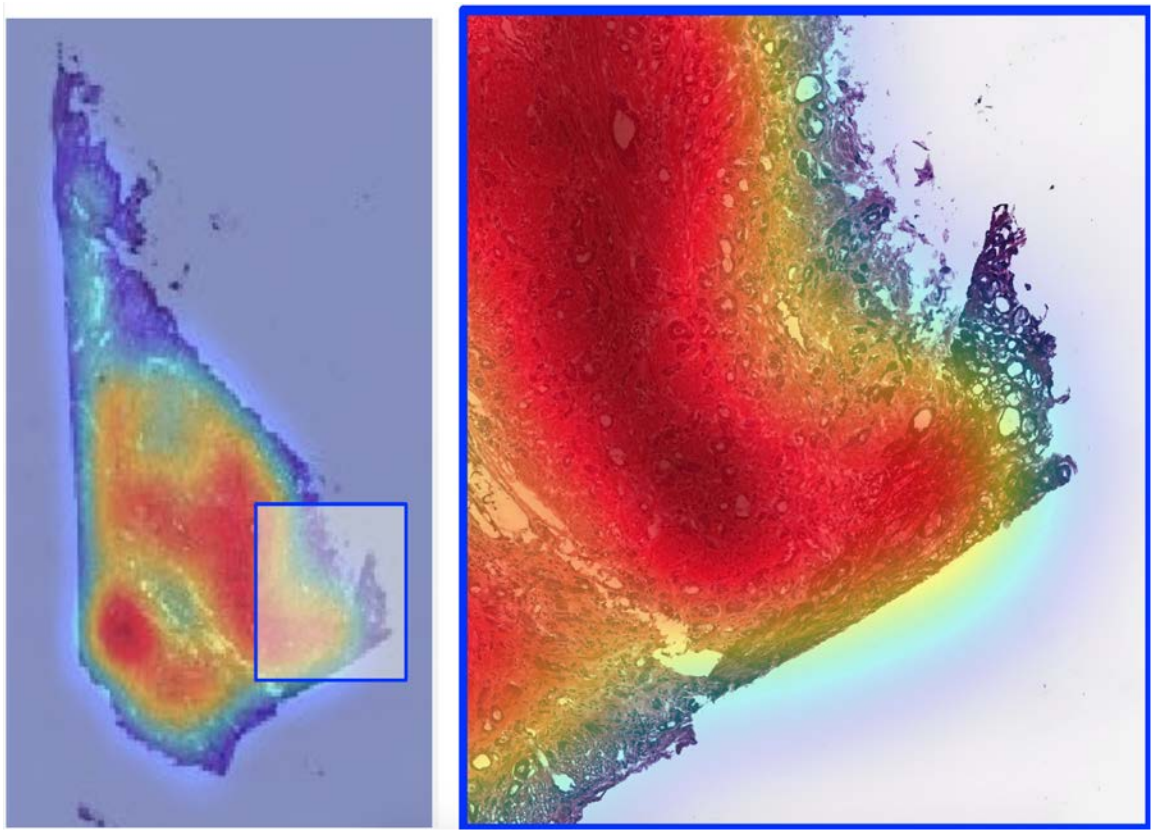
Use Case 1

Machine Learning in Medical Imaging

Hes·so

Haute école spécialisée

de Suisse occidentale, Switzerland



PROviding Computing solutions for ExaScale ChallengeS

Use Case 2

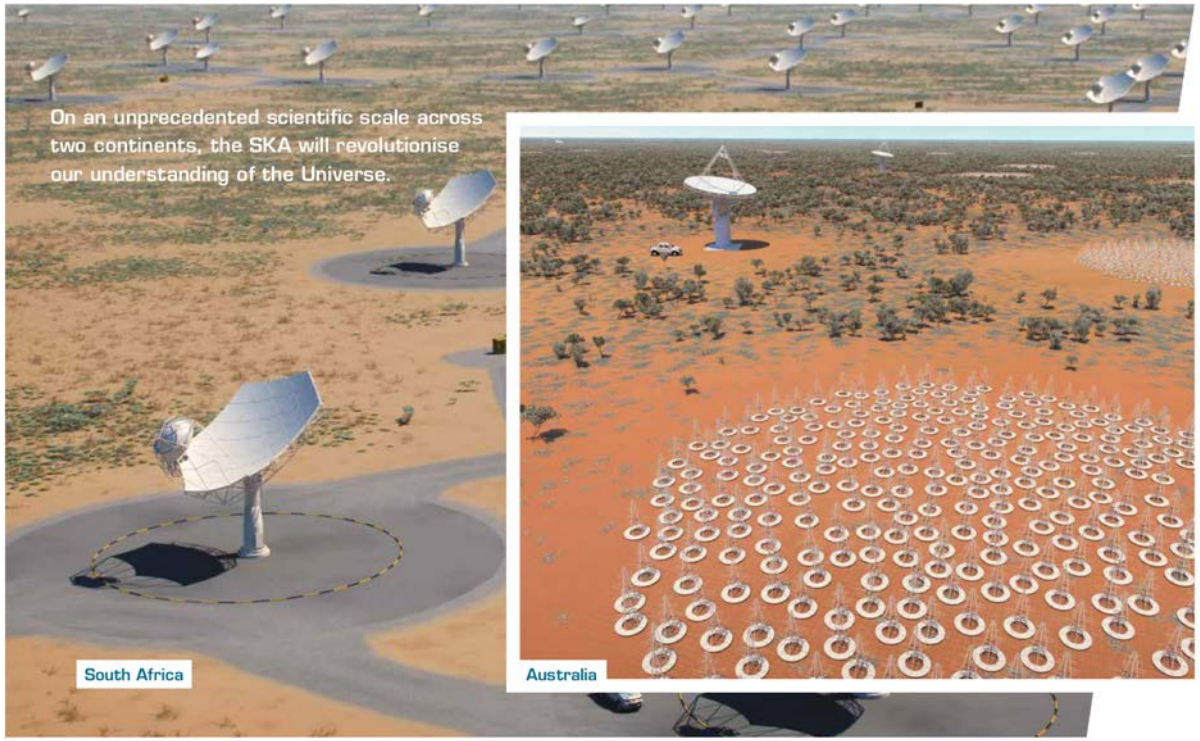
Analysis of Radioastronomy Observations LOFAR / SKA

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

PROCESS and EOSC

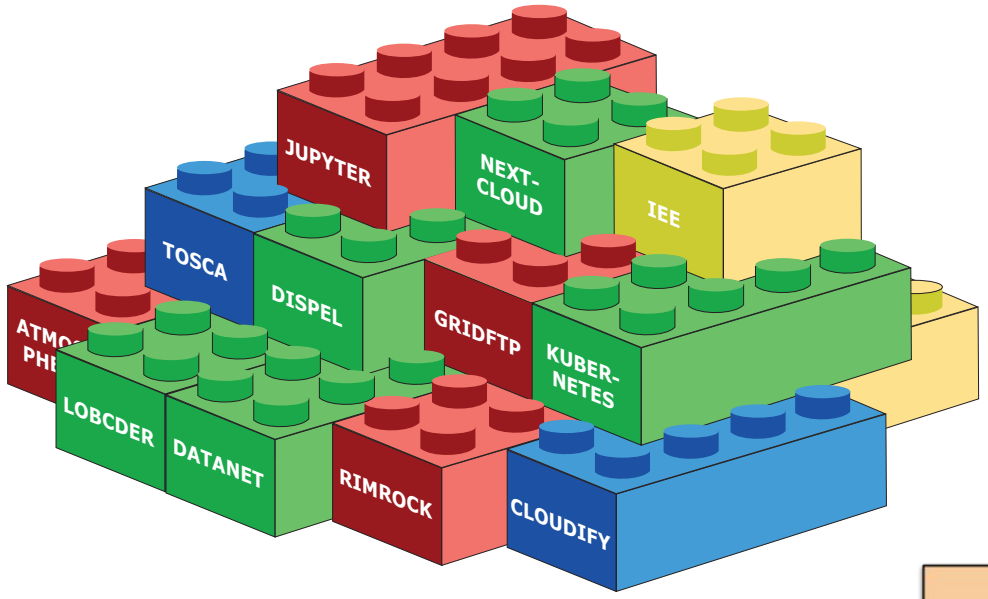
- Modules focusing on Data Service
  - Data Preprocessing
  - Data Transfer and Storage
  - Data Stage-In and –Out

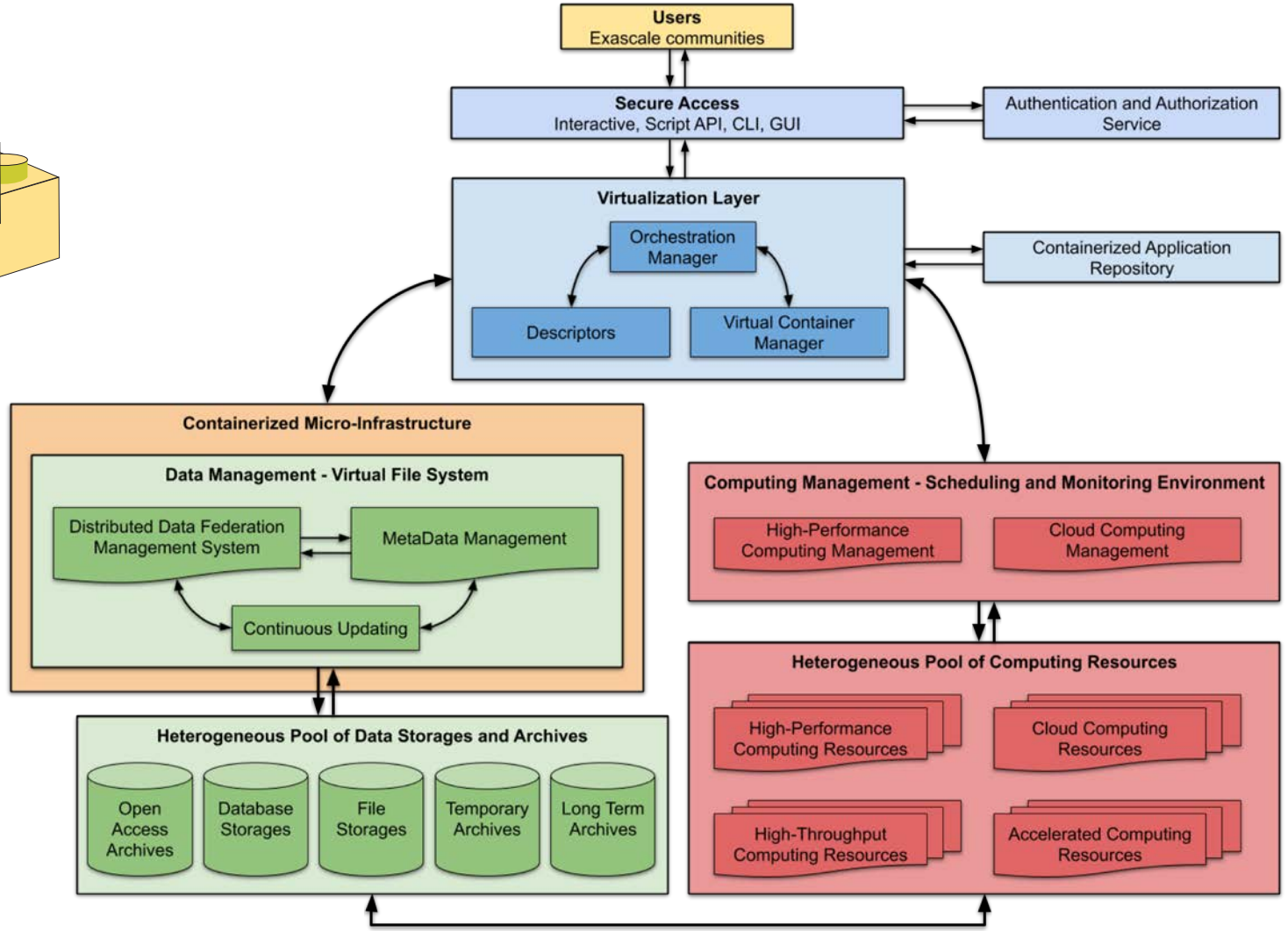
- Workflow-Deployment
  - on heterogenous HPC and Cloud resources

➔ Usable as Service Modules within EOSC

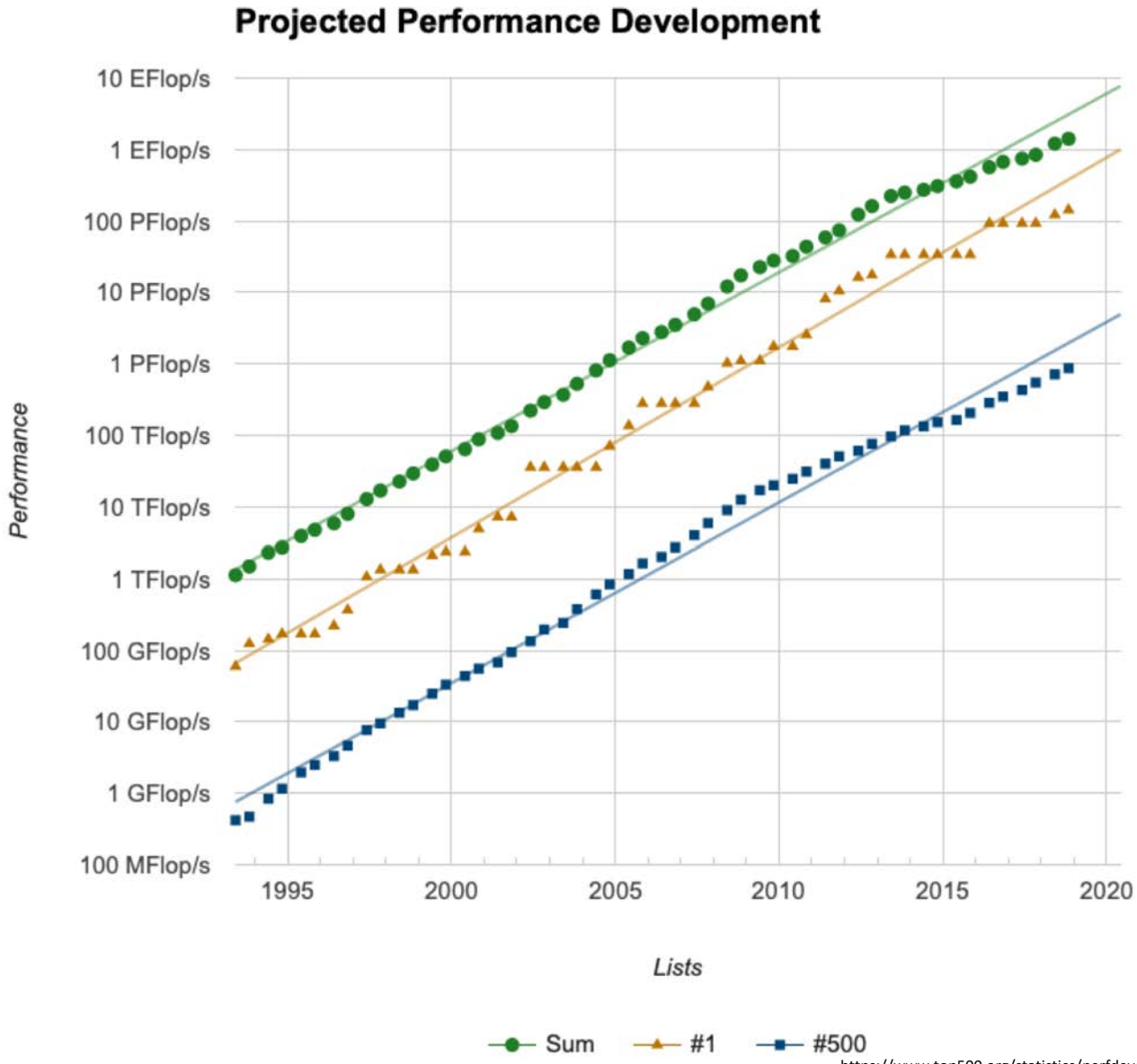
➔ Requires a modular execution environment and workflow management

Towards Exascale-ready Solutions

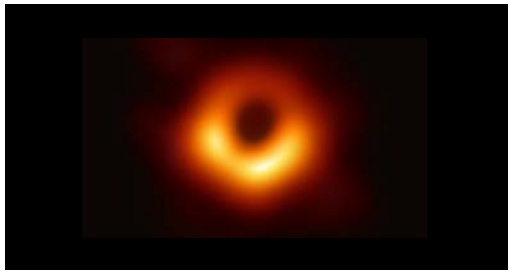


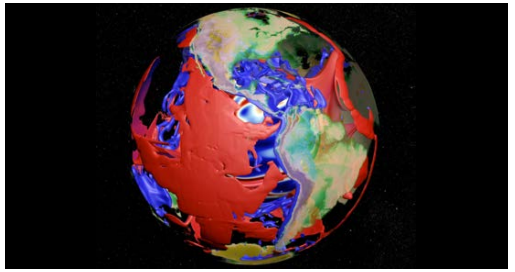


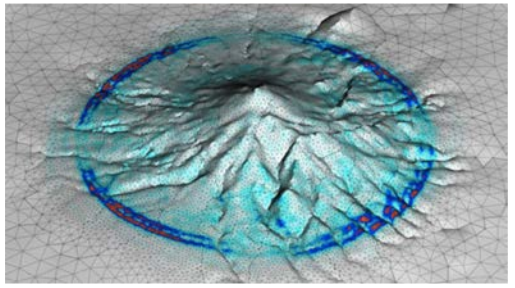
Enabling Exascale

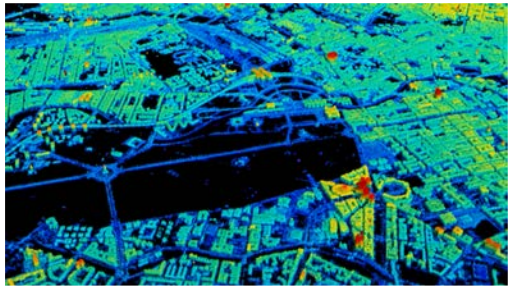


https://www.top500.org/statistics/perfdevel/









PROCESS

PROviding Computing solutions for ExaScale ChallengeS

Maximilian Höb

hoeb@mnm-team.org

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