ViWA TP 6: Optimization of process models for HPC-systems and deployment of an IT-research data infrastructure

Leibniz Supercomputing Centre (LRZ)

Wolfgang Kurtz, Stephan Hachinger, Alexander von Ramm, Jens Weismüller, Anton Frank, Dieter Kranzlmüller

ViWA and LRZ

- Modelling- and compute-centric goals of ViWA: globally assess
  - current usage of water resources
  - agricultural yields
  - flow and trade of virtual water across borders, and identify optimization potentials
- Global simulations with agro-hydrological model PROMET:
  - test water usage and crop yields under various management scenarios
  - with high spatial (1km²) and temporal resolution

  → High-performance computing (HPC) and research data management (RDM) are needed

Key to reaching ViWA goals: simulations help to identify current land management (comparing model observables to satellite images) and possibilities for improvement

ViWA and LRZ

Specific LRZ actions in ViWA:

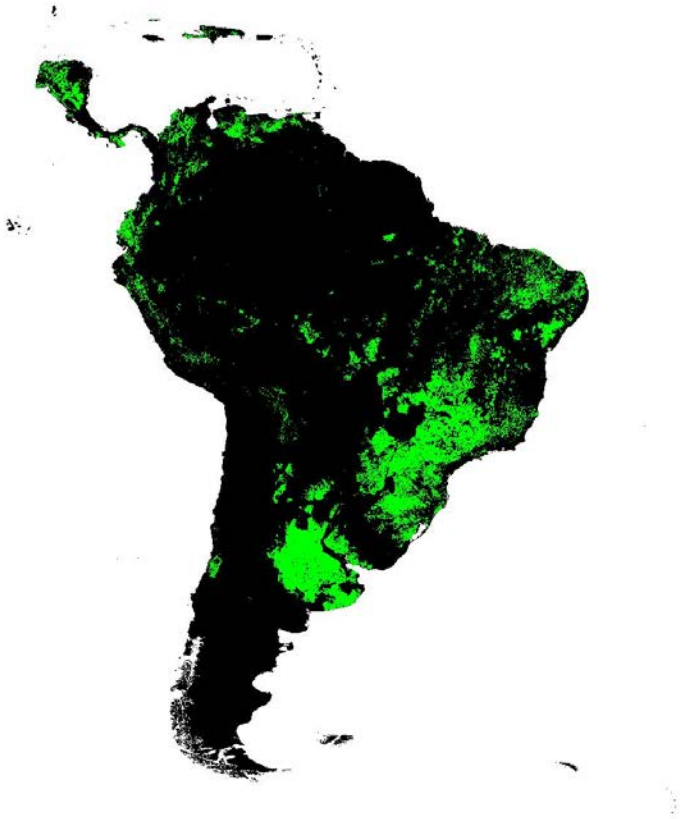
- Support/program HPC workflows for the simulations
- Devise domain decomposition methods for improved HPC efficiency
- Data management of simulation data
- Prototypical web-based data dissemination (FAIR[1] Research Data Management) and interactive data-access portal

[1]Findable, Accessible, Interoperable, Reusable – Wilkinson et al. (2016), Scientific Data 3, 160018

HPC workflows

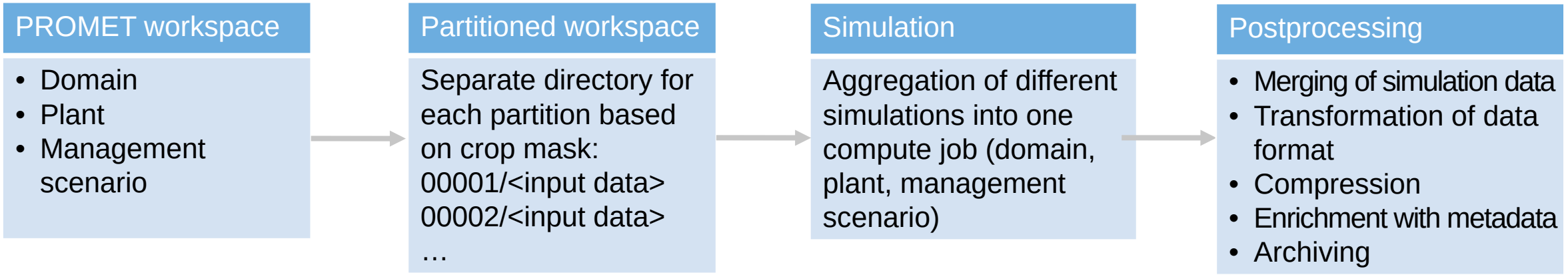
Simulation workflow

- Global ensemble simulations (resolution ~1km²) with the agro-hydrological model PROMET are run for various crop types
- Ensemble represents different management practices (e.g., irrigation, fertilization) and phenological factors
- Simulations are separated into nine geographic regions according to meteorological input data from CORDEX domains.
- Ensemble simulations are performed via job-farming on HPC infrastructure at LRZ



Simulation workflow

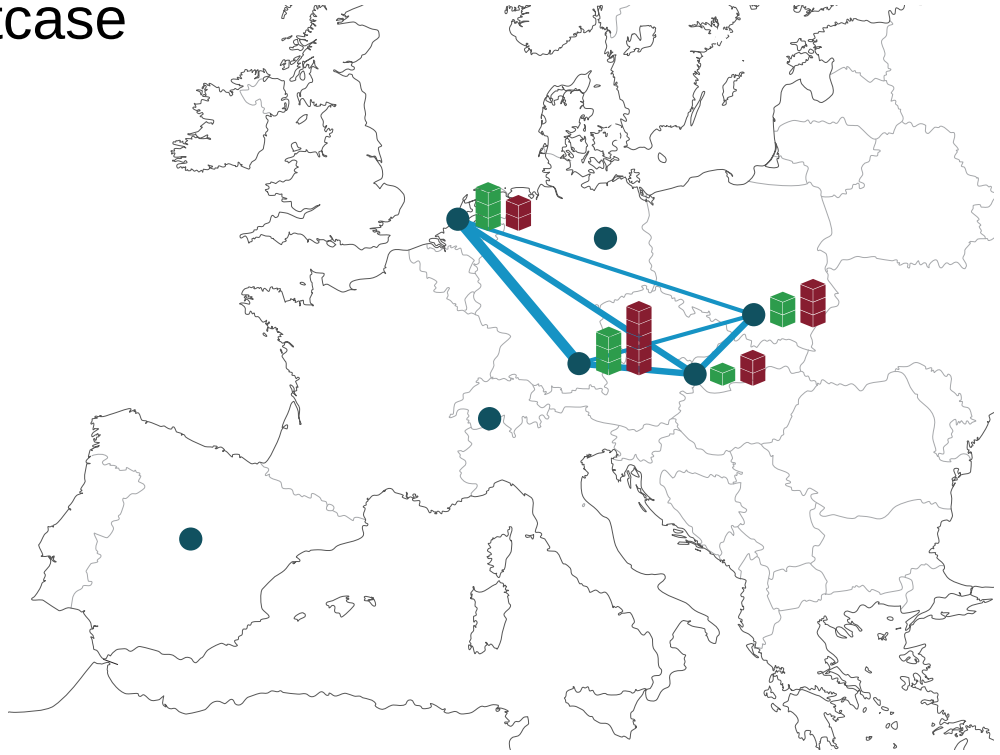
- Splitting of computational domain in subdomains via masks to distribute the workload onto HPC infrastructure
- Aggregation of different simulations into compute jobs
- Parallel merging of decomposed simulation data
- Transformation of model output from native binary format to annotated netCDF
- Storage on high-performance file system for later use (see data management section) and migration to tape archive



Collaboration with PROCESS project

Goals of PROCESS project:

- Optimal distribution of workflows across computing centres
- Modular toolbox to orchestrate workflows
- ViWA simulations as a testcase



Partner‘s location

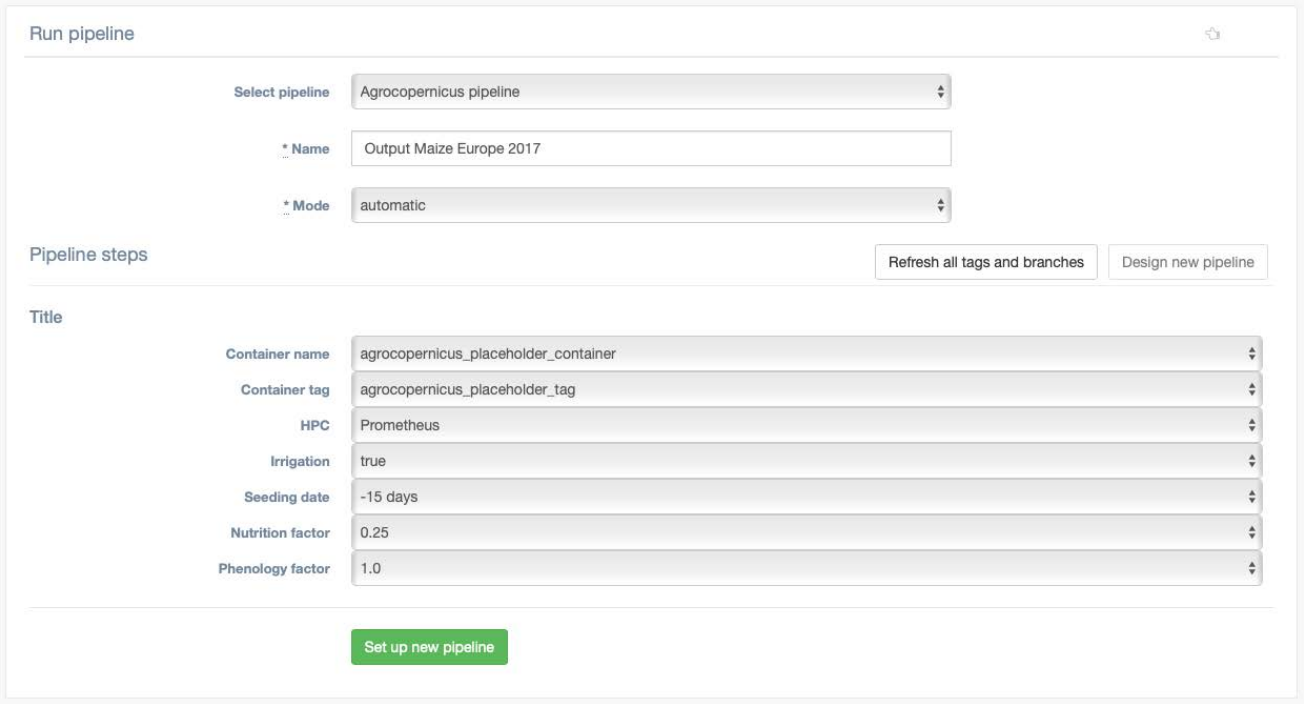
Storage Resources

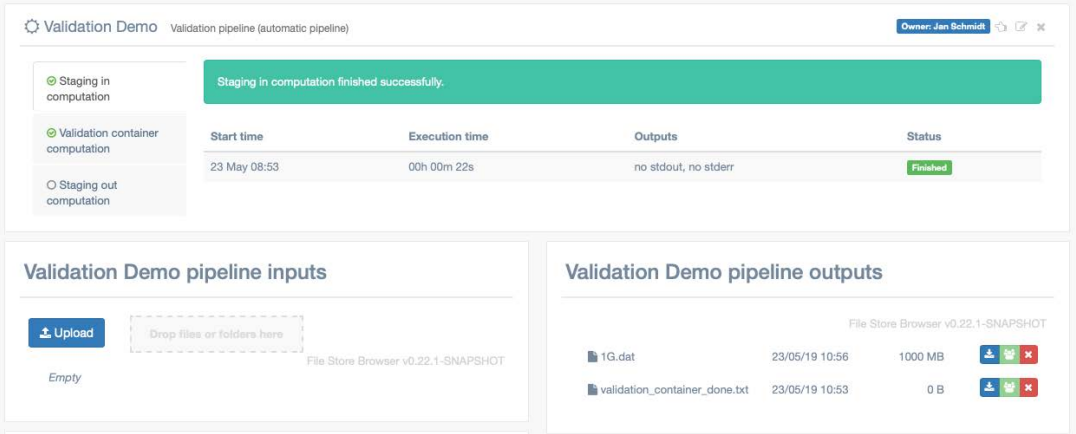
Compute Resources

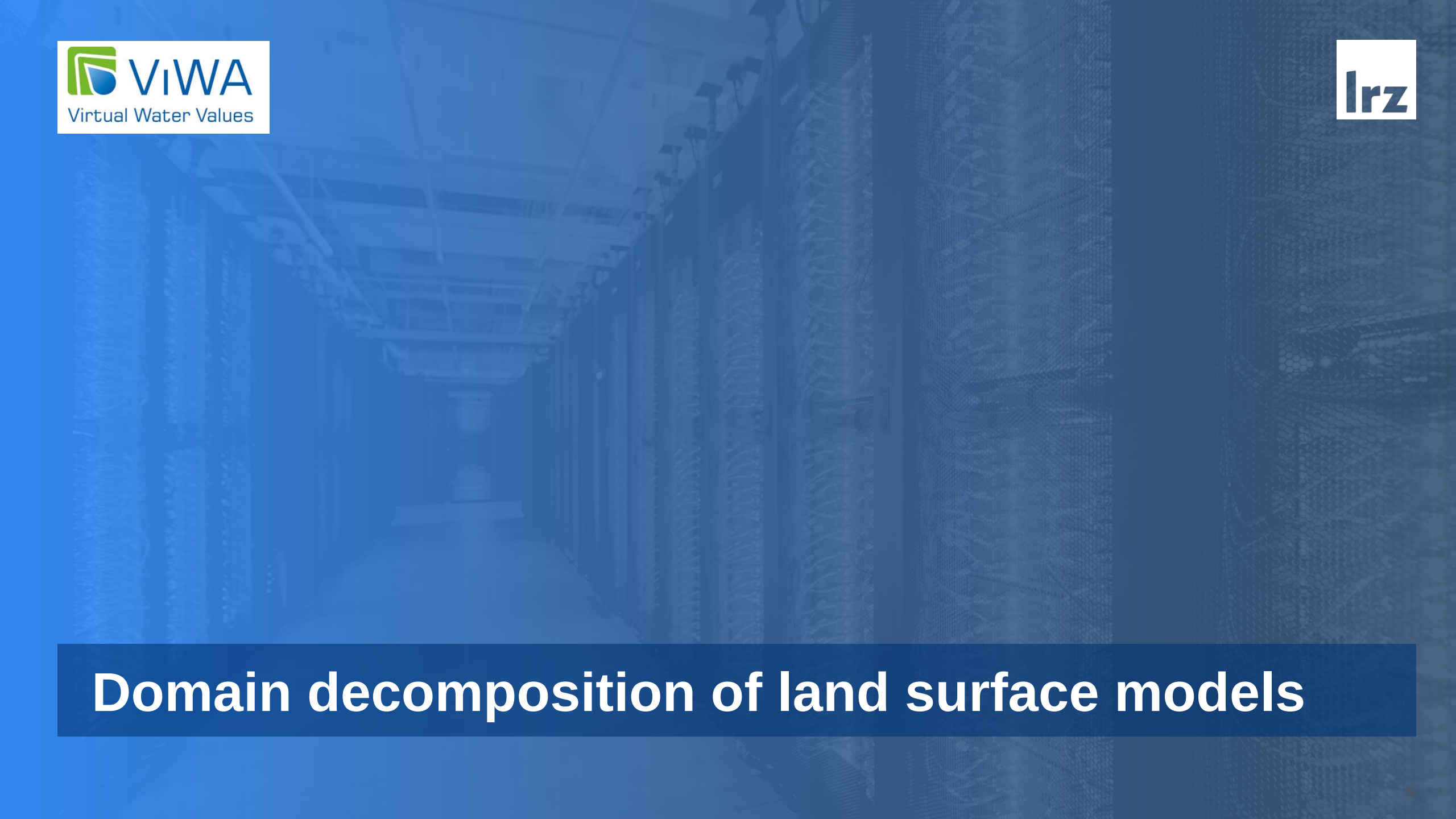
PROCESS – Example pipeline

Using experimental PROCESS services:

- ViWA simulations are launched on LRZ's LinuxCluster via web interface (automatization).
- Complexity of computational infrastructure is hidden from the user.







Domain decomposition of land surface models

Motivation

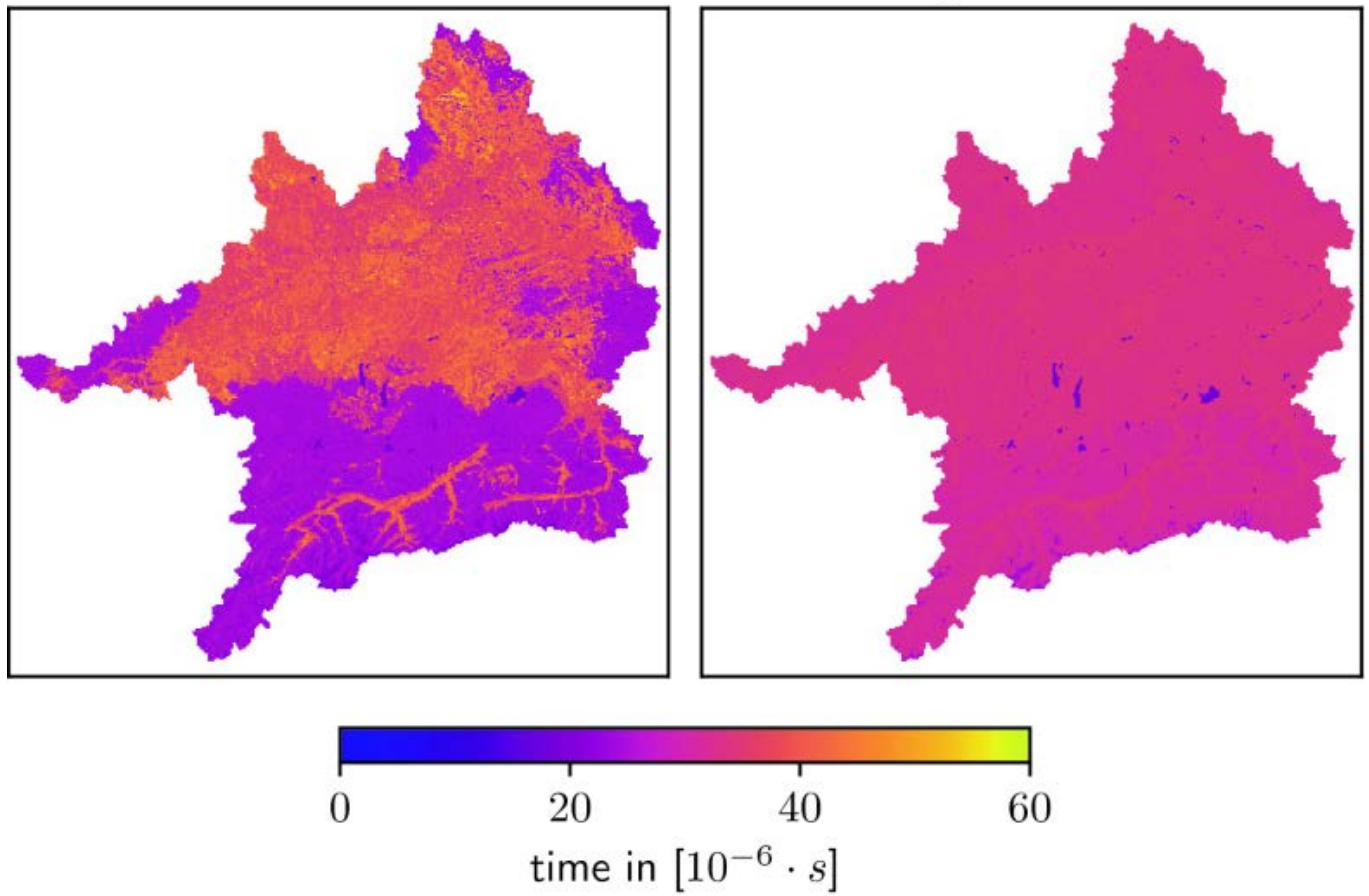
- Increasing demand for running large-scale, high-resolution land surface models
- Computational load on cell-level varies due to diurnal, seasonal variations in input data/ relevant processes and spatial heterogeneity
- Challenge: find load-balanced partitioning for running simulations on HPC systems

Simulation time per cell for one particular time step

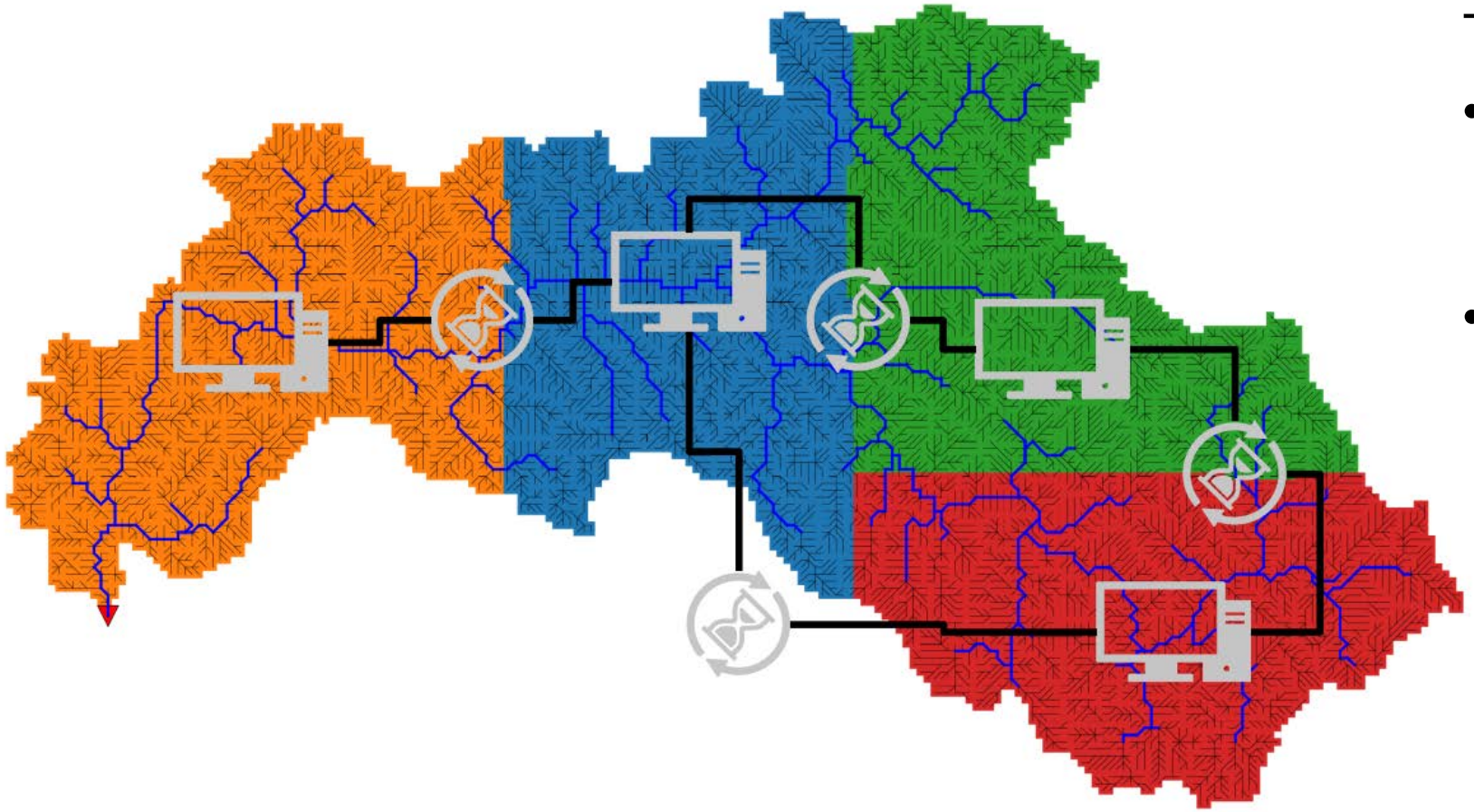
01.02-1999 09:00

Average simulation time per cell

1 year mean



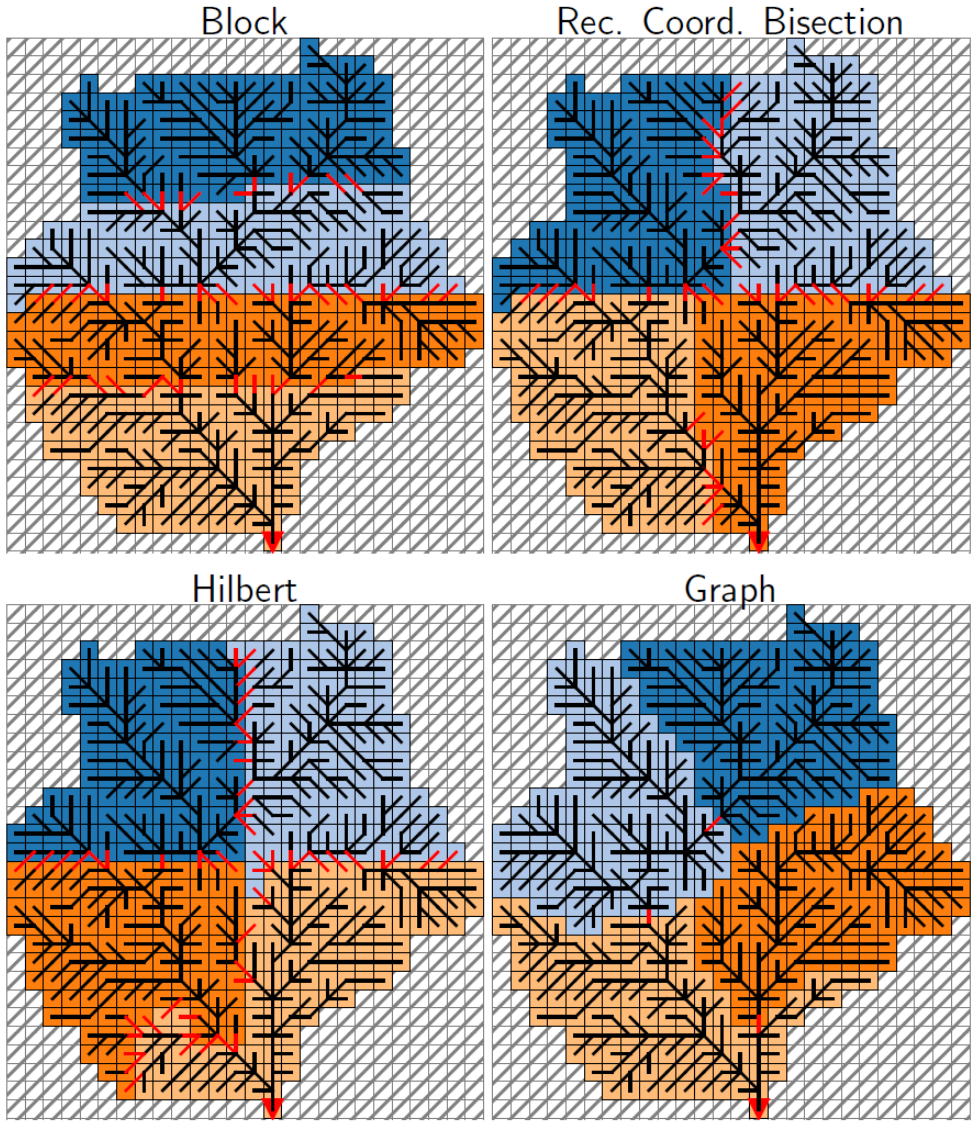
Main Challenges



Two-fold optimization:

- Find ideally sized partitions (load balancing)
- Reduce data dependencies (communication) for lateral exchange processes, e.g. river routing

Domain decomposition methods



Blockwise

- Simple ID-based scheme
- No minimization of communication

Recursive coordinate bisection

- Coordinate-based
- Communication reduction by minimizing the length of partition borders

Hilbert space-filling curve

- Coordinate-based
- Communication reduction by preserving data locality

K-way Graph-partitioning

- Graph-based
- Communication reduction by cutting the graph at smart locations

Methods

- Measure cell execution time of a benchmark simulation
  - PROMET simulation of Upper Danube catchment
  - ~76000 cells (1km²)
- Perform ‚artificial' strong scaling experiments with four domain-decomposition methods (Blockwise, Recursive coordinate bisection, Hilbert space-filling curve, Graph-based)
- Compute performance metrics:

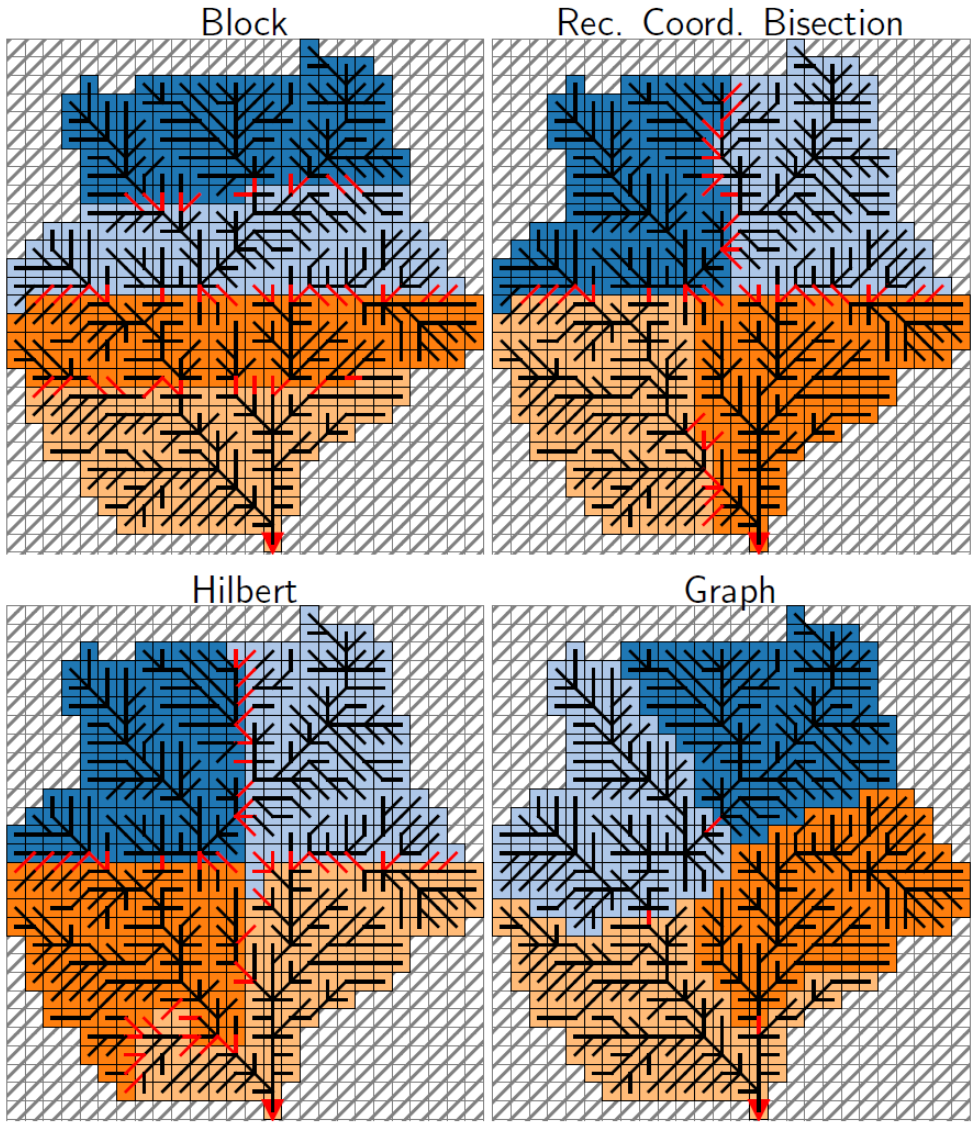
$$\text{theoretical speed up} = \frac{\text{Execution time on 1 proc.}}{\text{Execution time on n proc.}}$$

The higher the better!

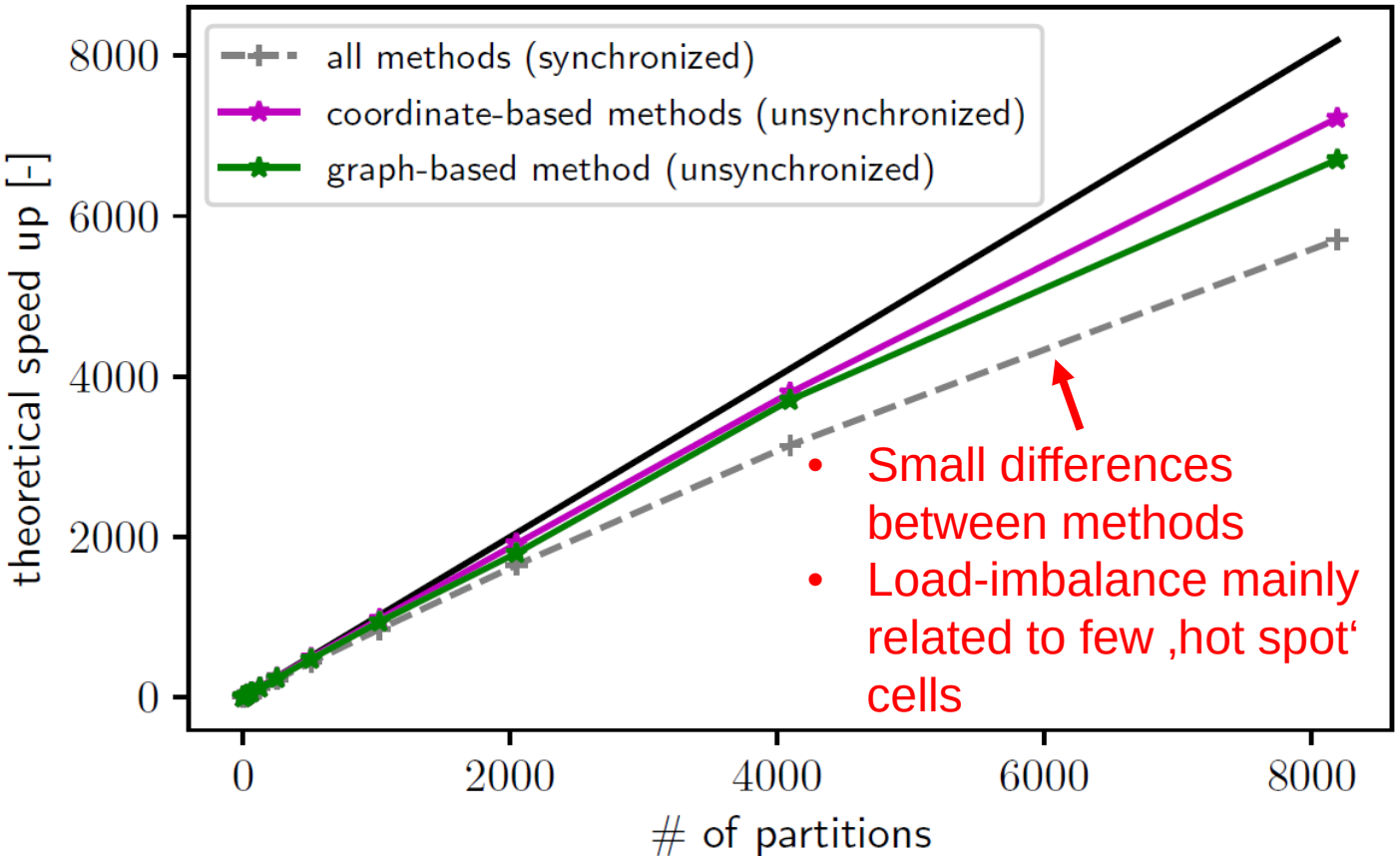
Number of cut flow-paths

The lower the better!

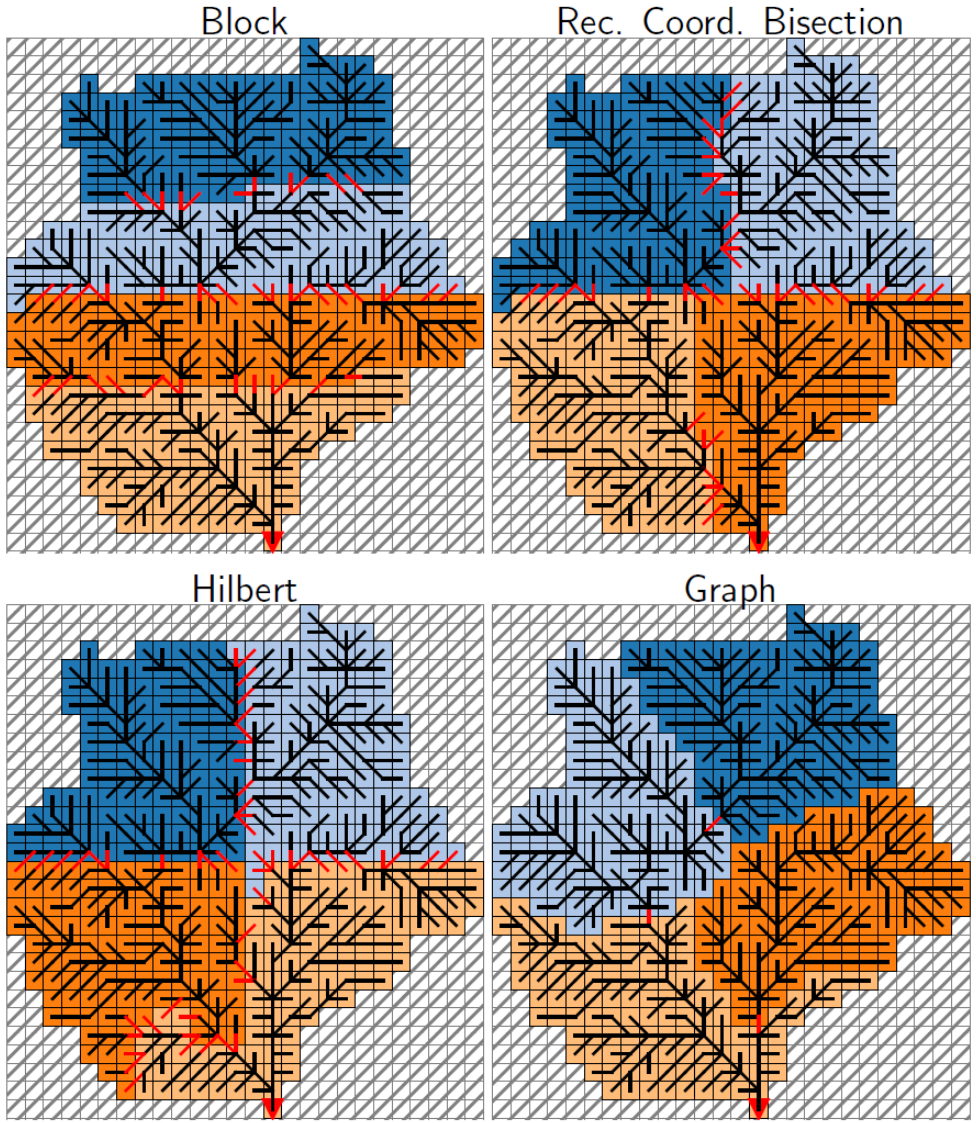
Theoretical peak speedup

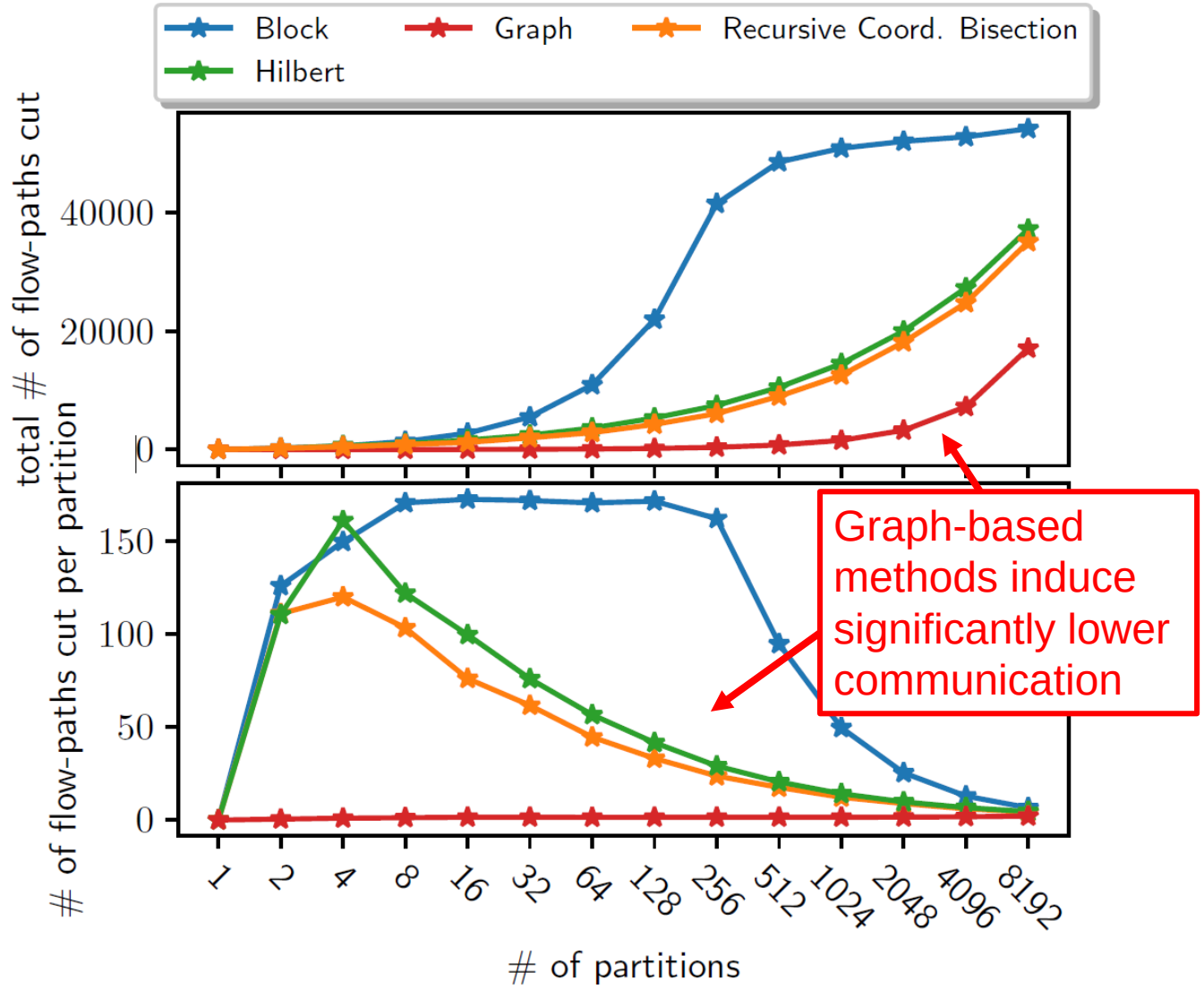


- With re-infiltration (synchronized):
  Drop in speedup due to load imbalances
- No re-infiltration (unsynchronized):
  Graph-based 6% worse at balancing the loads



Communication-inducing cutting of flow-paths





Data management

Data management for ViWA global simulation data

- ViWA‘s high-resolution simulations are performed with PROMET on LRZ‘s HPC System SuperMUC-NG (currently rank #13 in world top 500).
- Primary model output is stored on SuperMUC-NG‘s parallel file systems and tape archive in PROMET‘s native binary output format.
- Datasets are grouped into different
    - management scenarios, e.g. different settings for irrigation, plant phenology or fertilization
    - plant types
    - spatial domains, complying with meteorological input data (CORDEX domains)

Data management for ViWA global simulation data

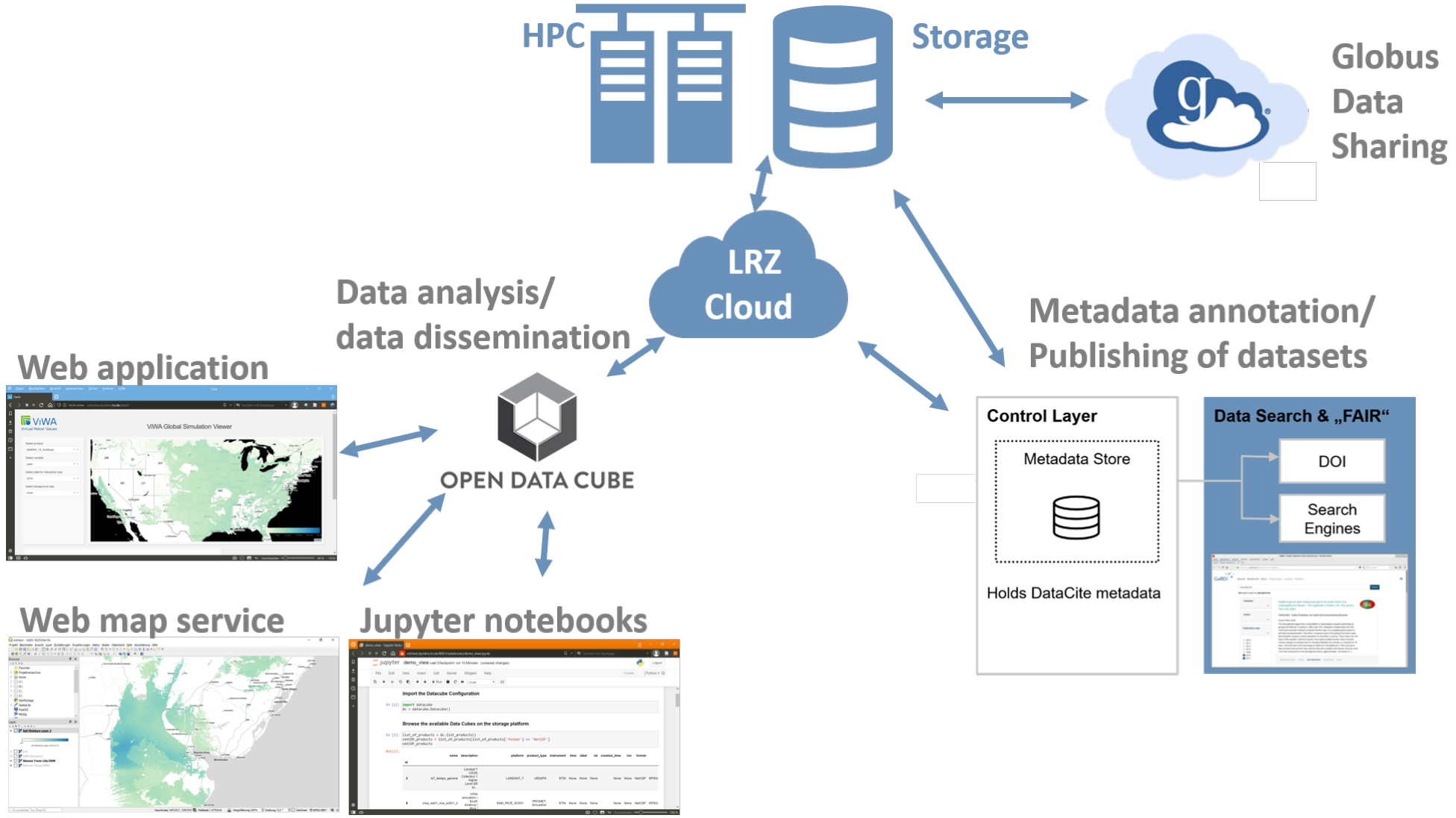
Goals of data management in ViWA:

- Support data processing and data analysis for the project members
- Disseminate simulation data to the scientific community and interested stakeholders
- Enrich output with metadata to support FAIR principles (cf. Slides 2/3)

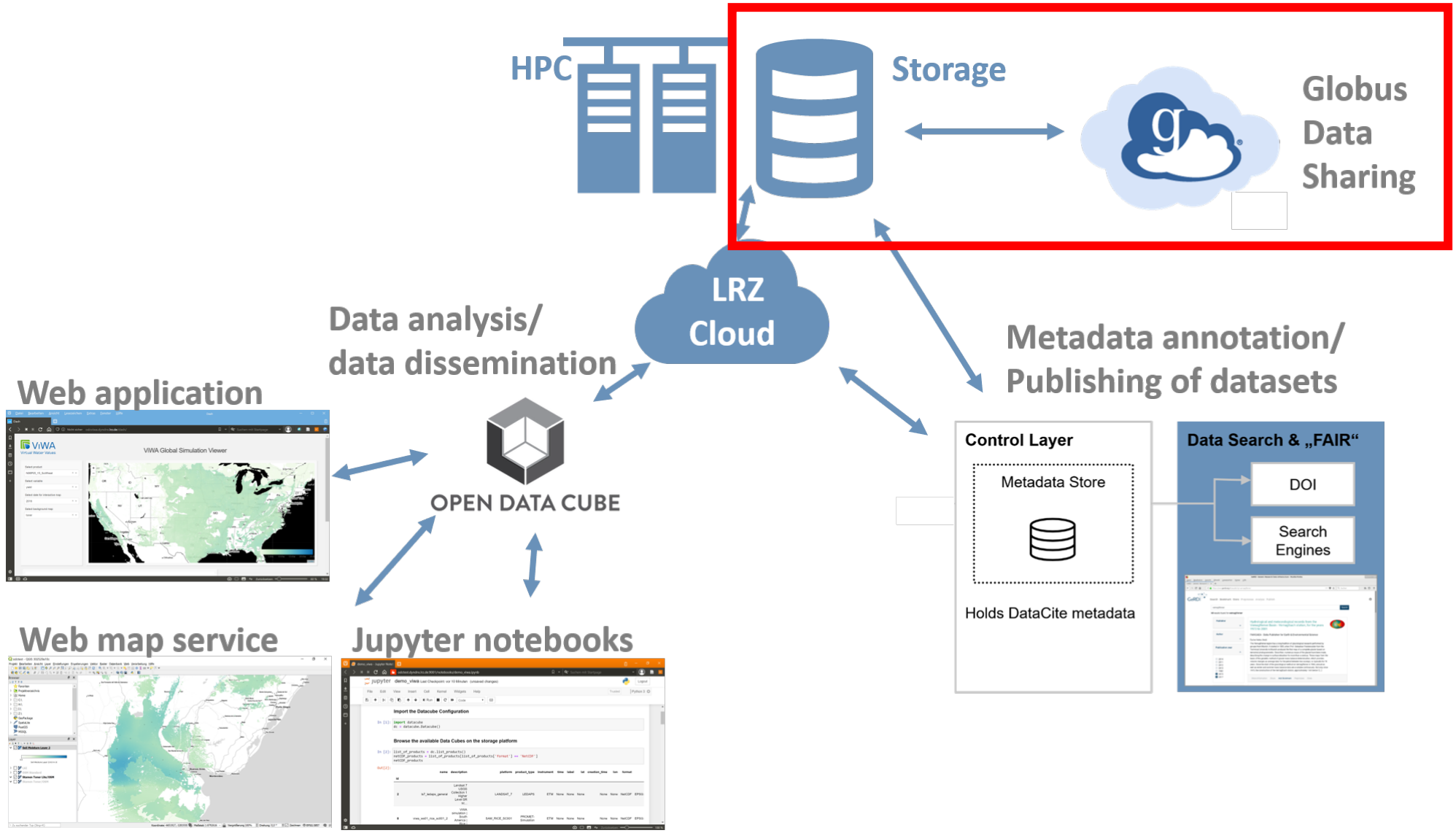
Approach:

- Provide assistance for handling data on HPC systems
- Transform simulation data to standard exchange formats (netCDF, CF conventions)
- Establish higher-level data management/ data processing capabilites for different users by using Open Data Cube and Cloud technologies
- Enable automated metadata enrichment for publishing of ViWA output in data catalogues

Data management architecture for ViWA simulation data



Data management architecture for ViWA simulation data

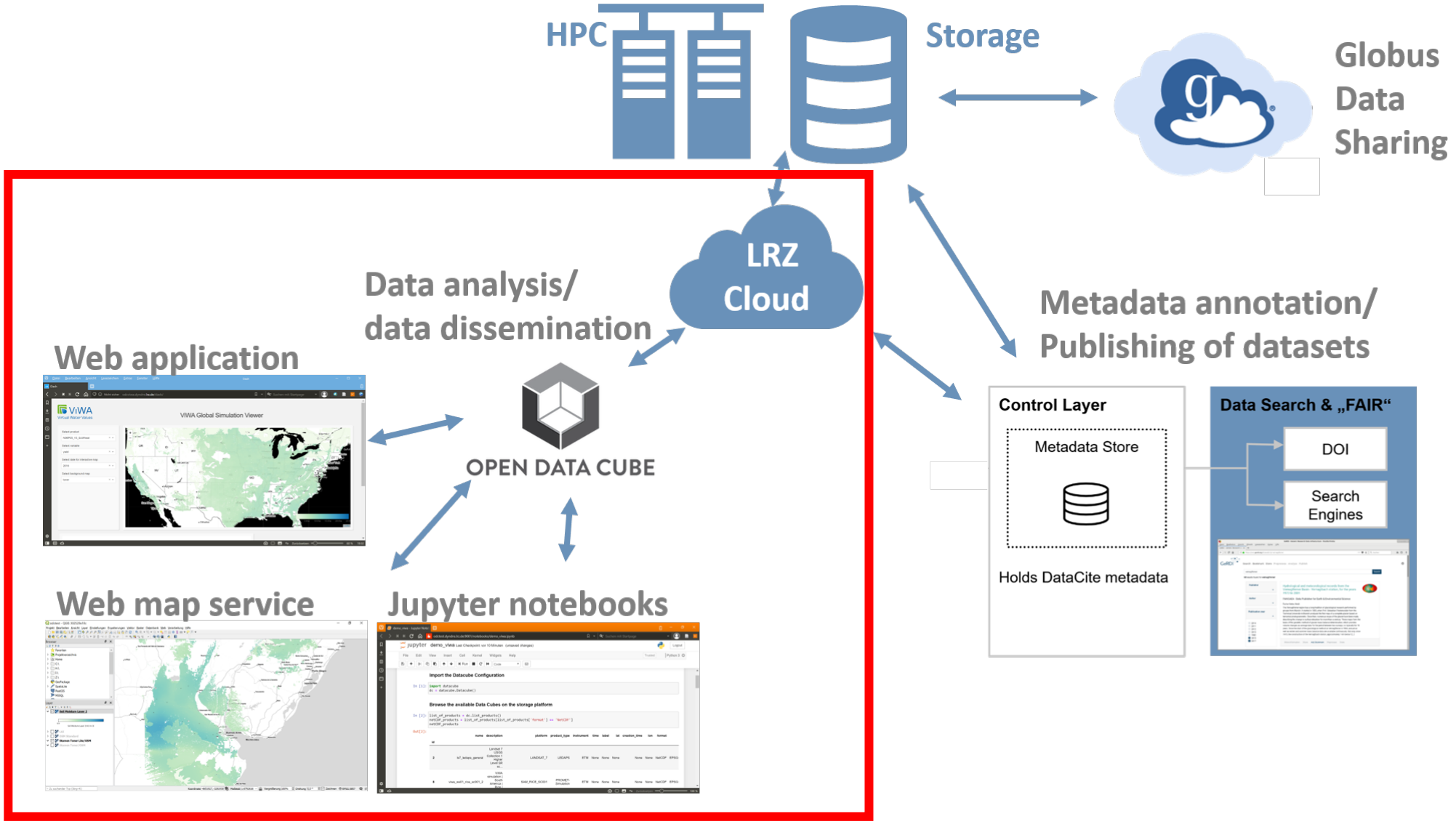


Data handling on HPC systems

Steps:

- Converting PROMET output (native binary output format) to annotated netCDF files following CF conventions
- Storing data on SuperMUC-NG‘s parallel file systems (Data Science Storage) and archiving on tapes
- Transferring data to Data Science Storage (DSS), accessible from different LRZ compute systems (e.g. LRZ Cloud)
- Sharing of data on DSS with external users via Globus services

Data management architecture for ViWA simulation data



Data access via Open Data Cube and Web applications (Dash)

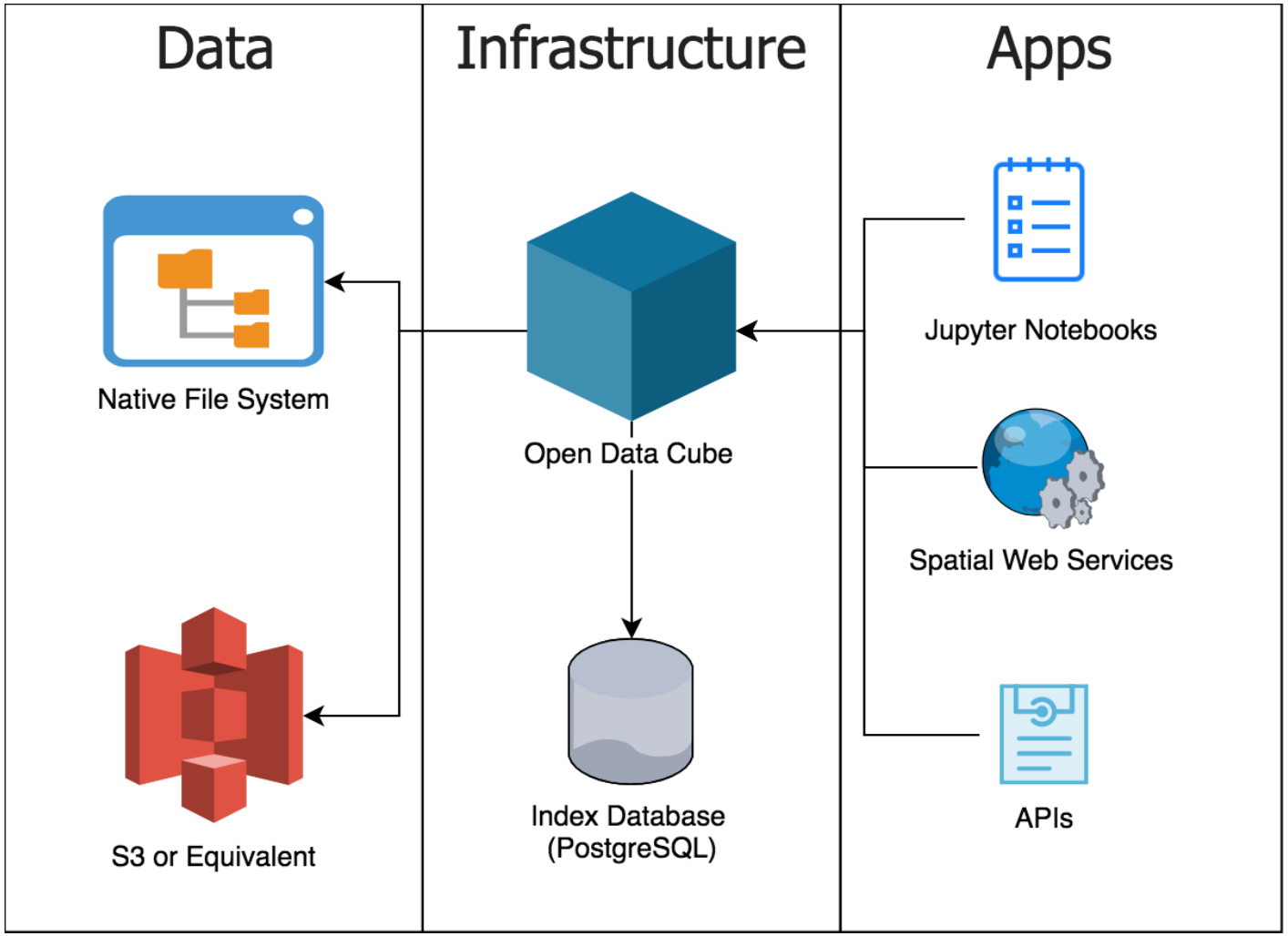
- Open Data Cube (ODC) framework (see next slide) has been set up for managing netCDF simulation datasets:
  - ODC framework installation in LRZ‘s Compute Cloud
  - Preparation scripts for ingesting ViWA simulation data into ODC framework and definition of metadata for ODC data catalogue
  - Tiling of ViWA datasets into 10°x10° tiles for faster data access

- On top of ODC and its API, data analysis services have been established:
  - JupyterHub framework for detailed analysis (primarily for project members)
  - Web applications (based on Dash) for viewing and retrieving data (primarily for 3rd parties)

Open Data Cube framework - Background

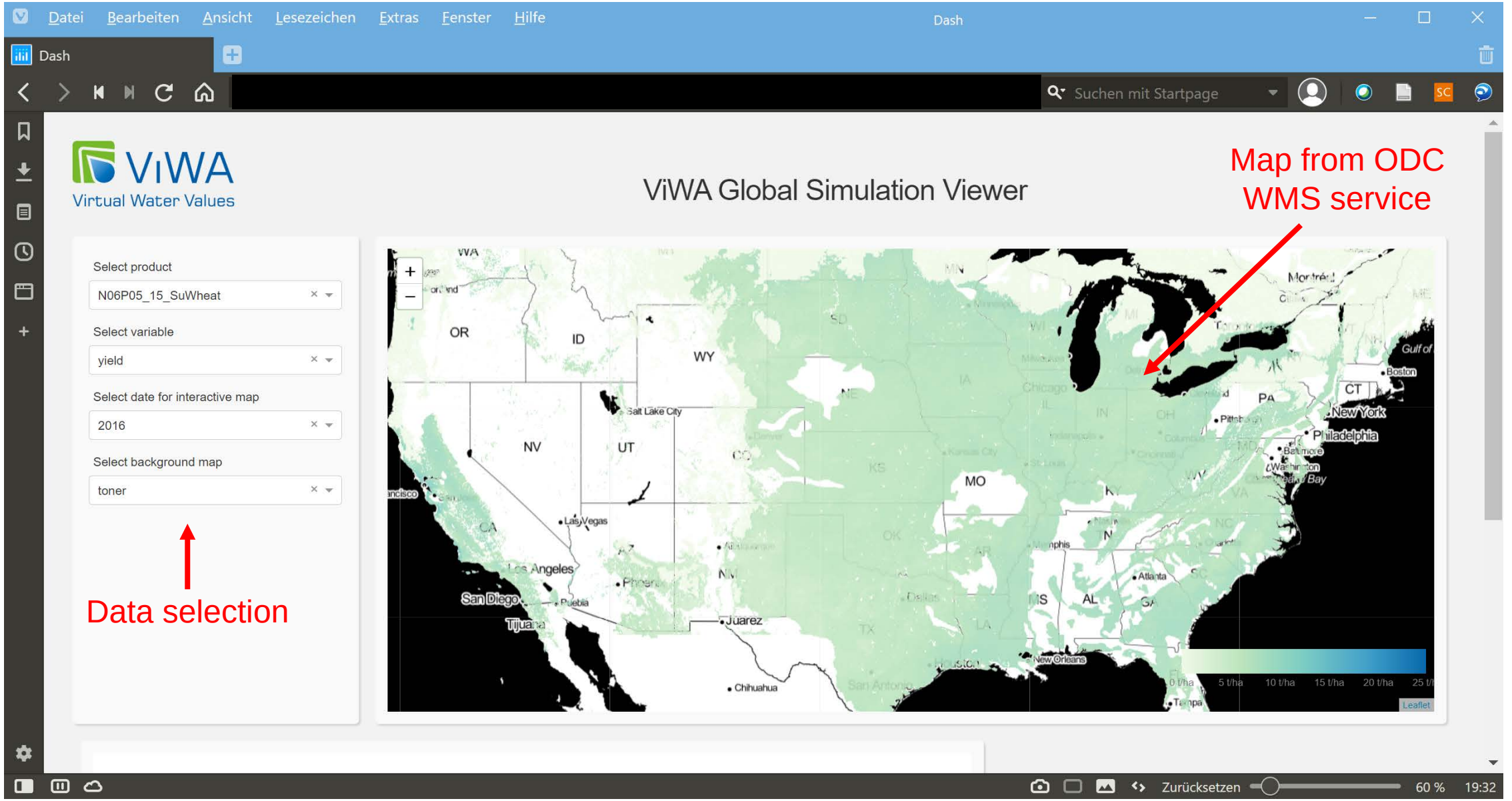
- Open-source project initiated by Geoscience Australia
- Primarily developed for data management and analysis of remote sensing data
- Core component manages data bases (PostgreSQL) for data products and metadata
- Data storage is file based (netCDF, geoTIFF)
- Functionality for searching, retrieving and transforming (e.g. tiling, reprojecting) datasets is provided
- Definition and ingestion of datasets implemented via yaml configuration files
- Additional service components available (e.g. WMS service, data explorer, Jupyter environment)

Open Data Cube framework – Basic architecture

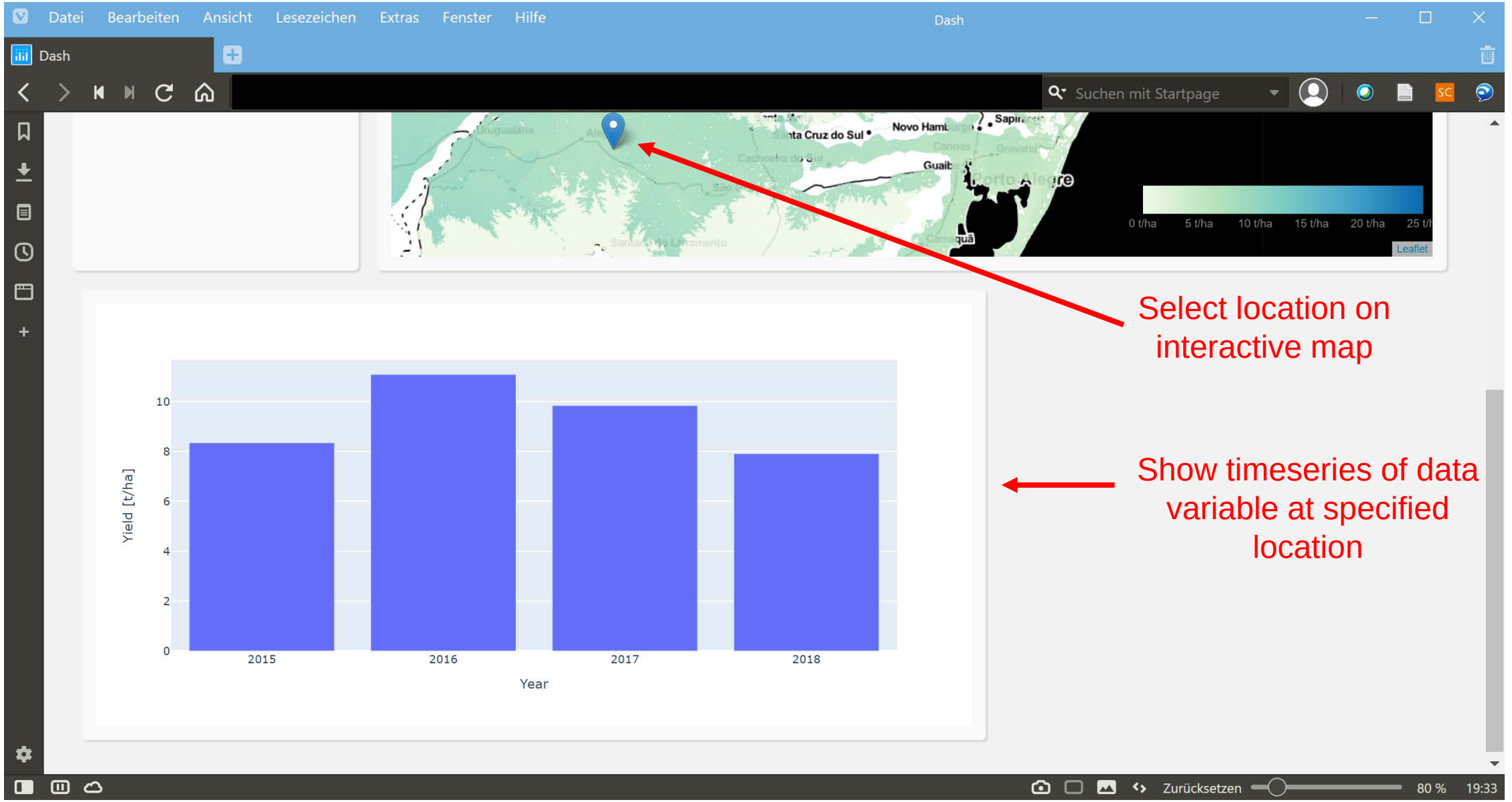


https://medium.com/opendatacube/what-is-open-data-cube-805af60820d7

Web application for viewing of ViWA data products



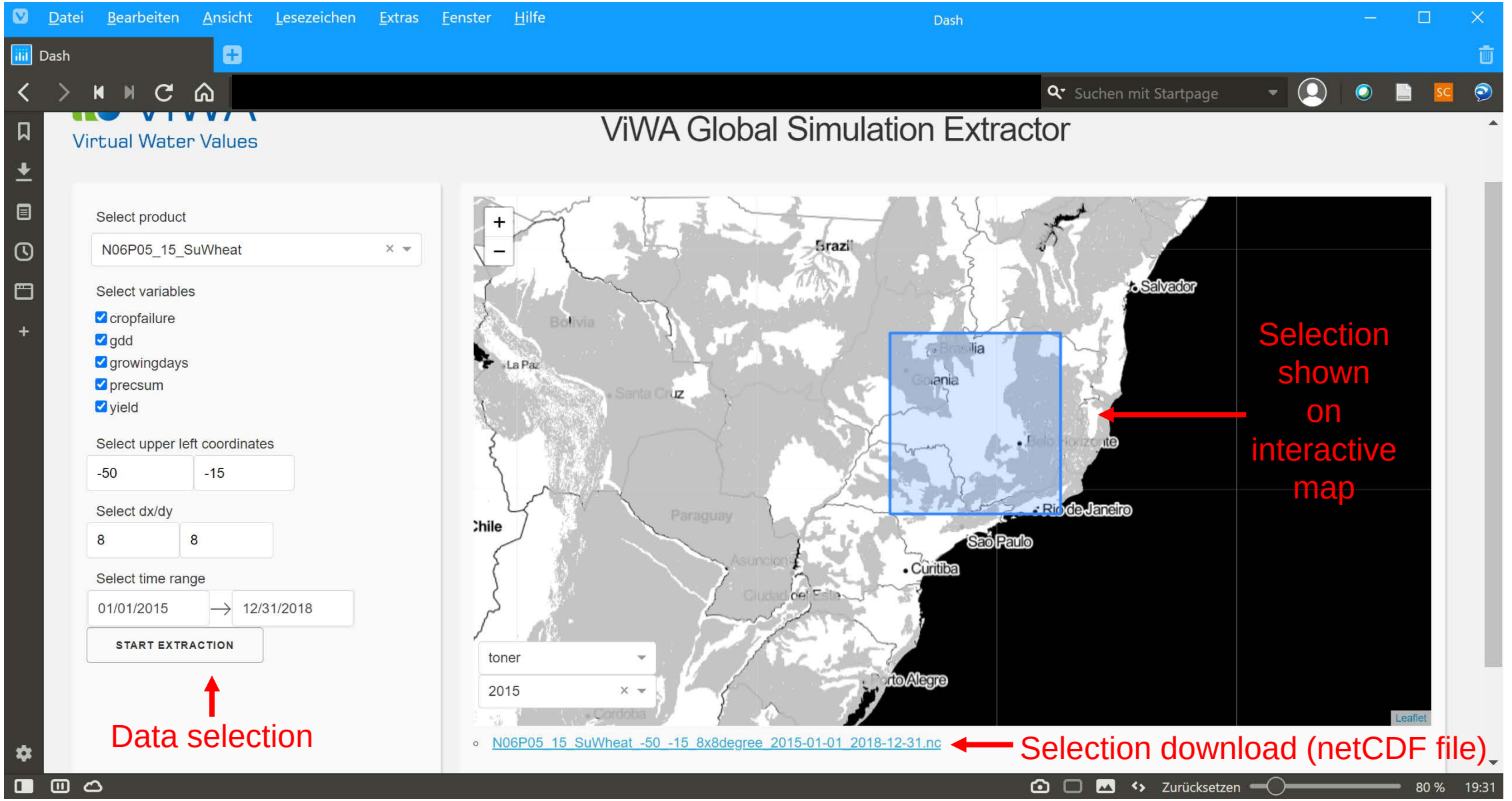
Web application for viewing of ViWA data products



Select location on interactive map

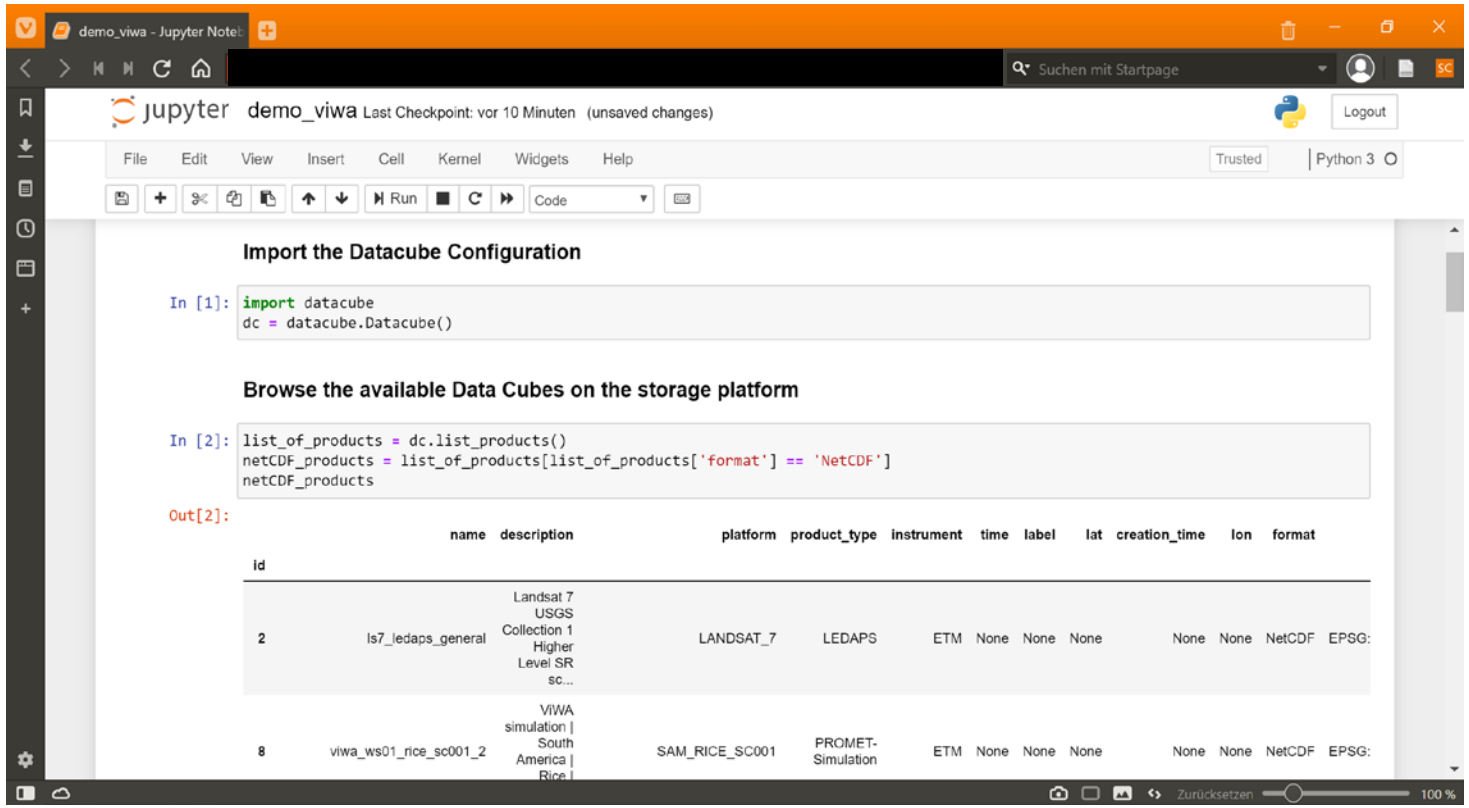
Show timeseries of data variable at specified location

Web application for retrieving of ViWA data products

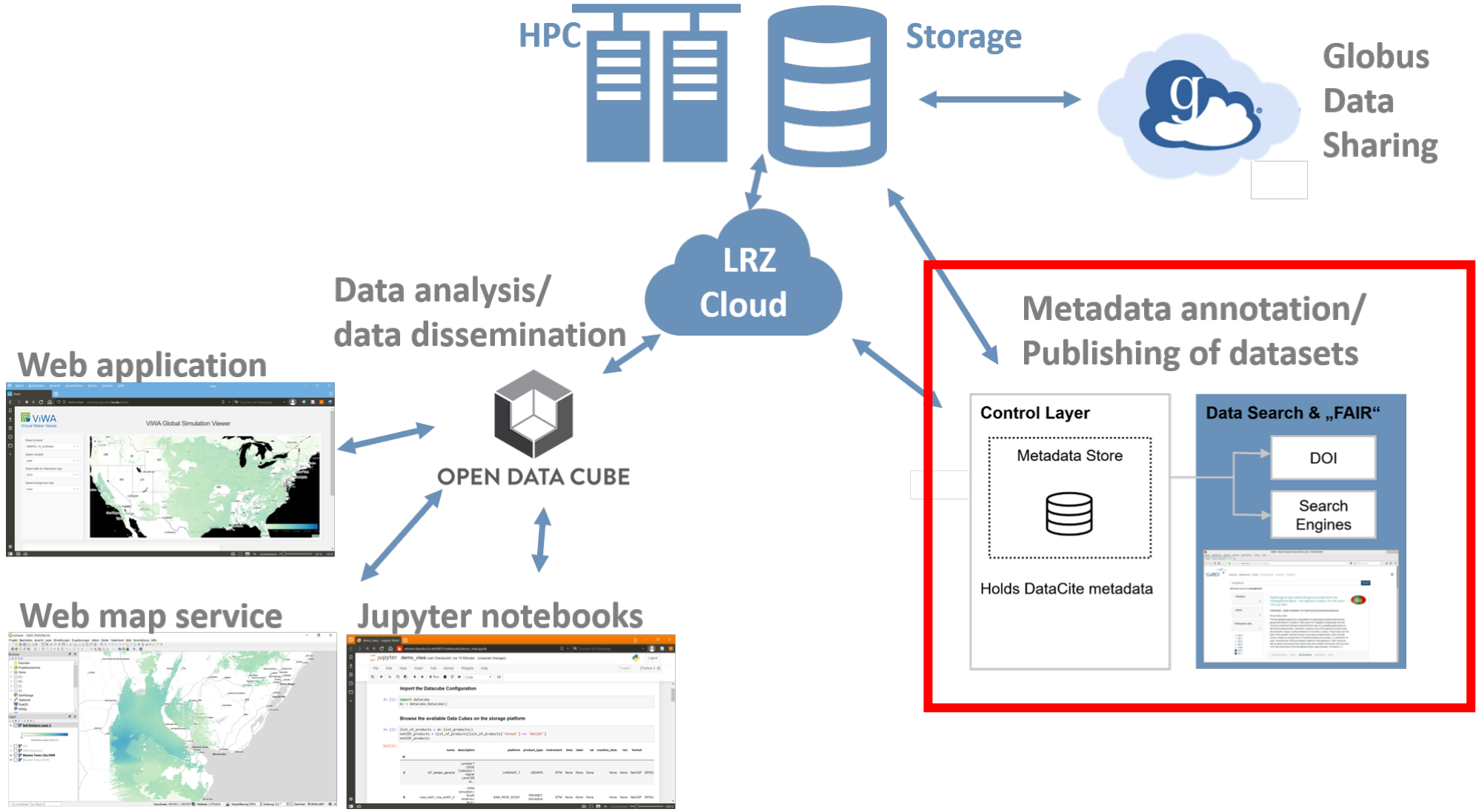


Jupyter notebook for data analysis

- Jupyter: interactive data processing via web browser
- ODC core component provides API to interact with ingested data products



Data management architecture for ViWA simulation data

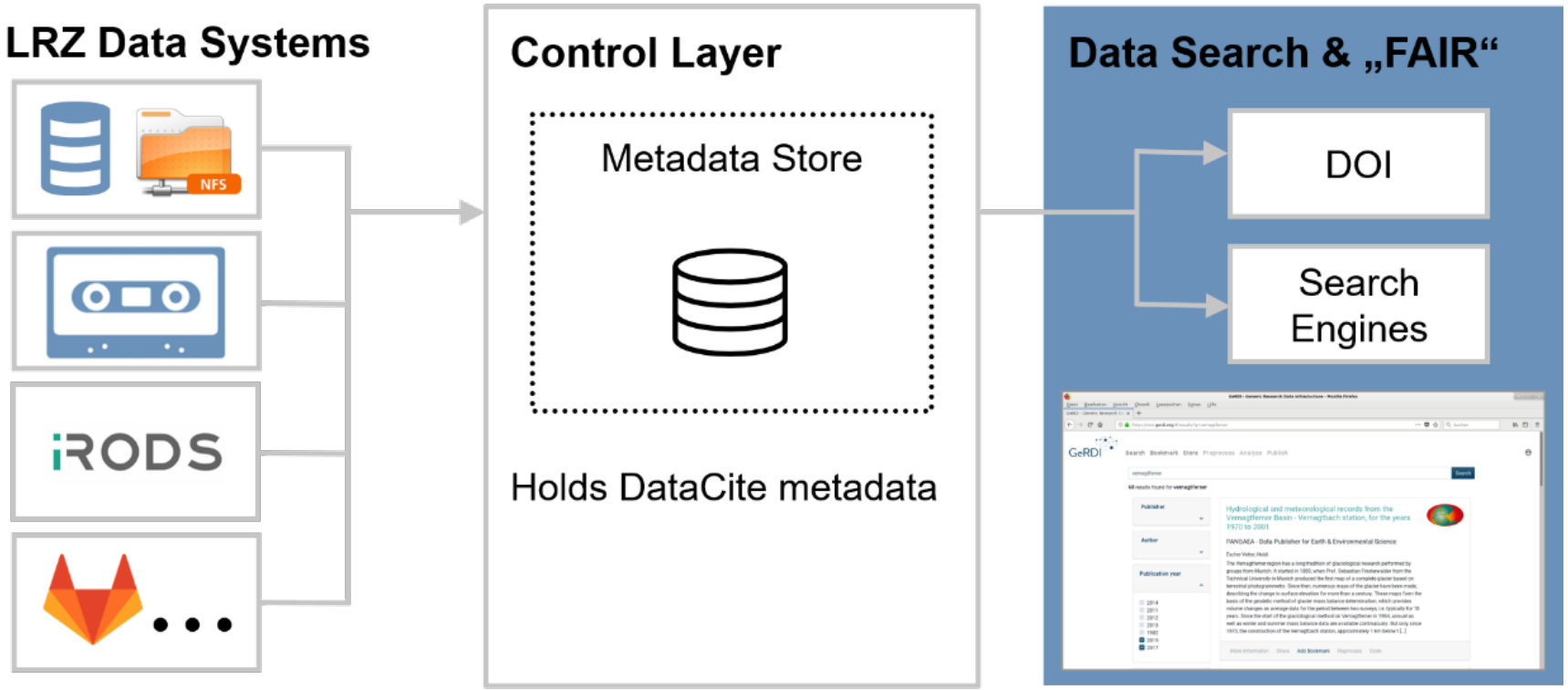


Metadata enrichment for ViWA‘s HPC simulation data

- FAIR principles: Findable, Accessible, Interoperable, Reusable
- LRZ works on a service for FAIR management of large simulation datasets, including
  - semi-automatic enrichment with metadata (covering DataCite standard and more),
  - metadata store,
  - assignment of persistent identifiers (e.g. Digital Object Identifiers - DOIs),
  - data product description web-pages (DOI landing pages), and
  - dissemination to data search portals (e.g. EUDAT-B2FIND, Google Dataset Search).
- Benefits for ViWA as pilot use case of LRZ service:
  - DOIs
  - ViWA data findable on scientific data portals
  - automatic persistence checks

Metadata enrichment - LTDS

- Enrichment with metadata covering DataCite scheme + ViWA special attributes (e.g. geolocation); assignment of DOIs:
  - Common attributes of all ViWA/PROMET data products written in configuration files
  - System scans data, augments common attributes by subset attributes (e.g. crop type), and automatically assigns DOIs and metadata to each subset



Summary

Summary

- Workflows for running ViWA ensemble simulations with PROMET were established on LRZ‘s HPC infrastructure
- Graph-based partitioning schemes require less communication overhead for lateral processes
- Simulation data are available in annotated netCDF format on LRZ‘s storage systems (and can be shared via Globus)
- Metadata for simulation products can be published to scientific search engines
- Subsets of the simulation data can be viewed, analysed and downloaded via web interface based on Open Data Cube technology

Acknowledgements

ViWA is a collaborative project of the funding program "Global Resource Water (GROW)" in the framework program FONA (Research for Sustainability) of the German Ministry for Education and Research (BMBF).

SPONSORED BY THE

Federal Ministry of Education and Research