

# 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

- 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 (1km<sup>2</sup>) 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

## What does a Supercomputing Centre do in ViWA? 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<sup>1</sup> Research Data Management) and interactive data-access portal

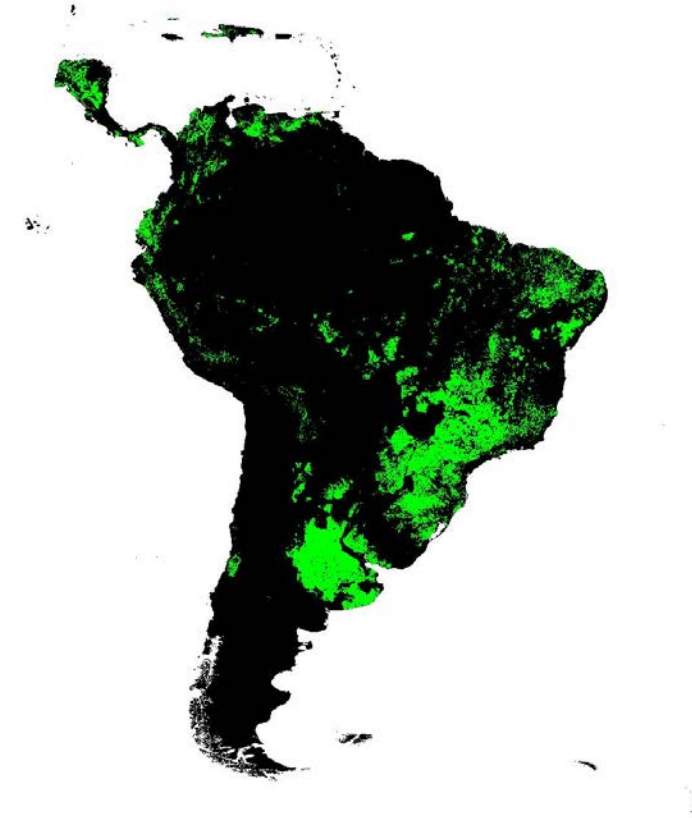
<sup>1</sup>Findable, Accessible, Interoperable, Reusable – Wilkinson et al. (2016), Scientific Data 3, 160018



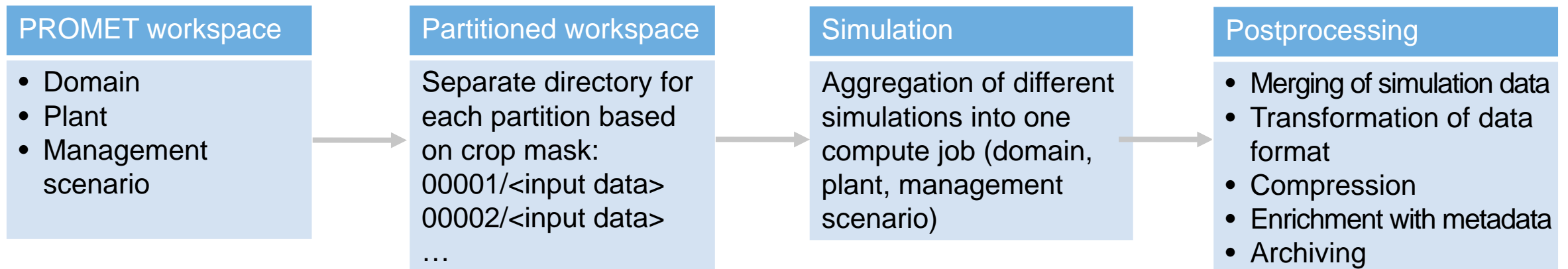
# HPC workflows

# Simulation workflow

- Global ensemble simulations (resolution  $\sim 1\text{km}^2$ ) 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



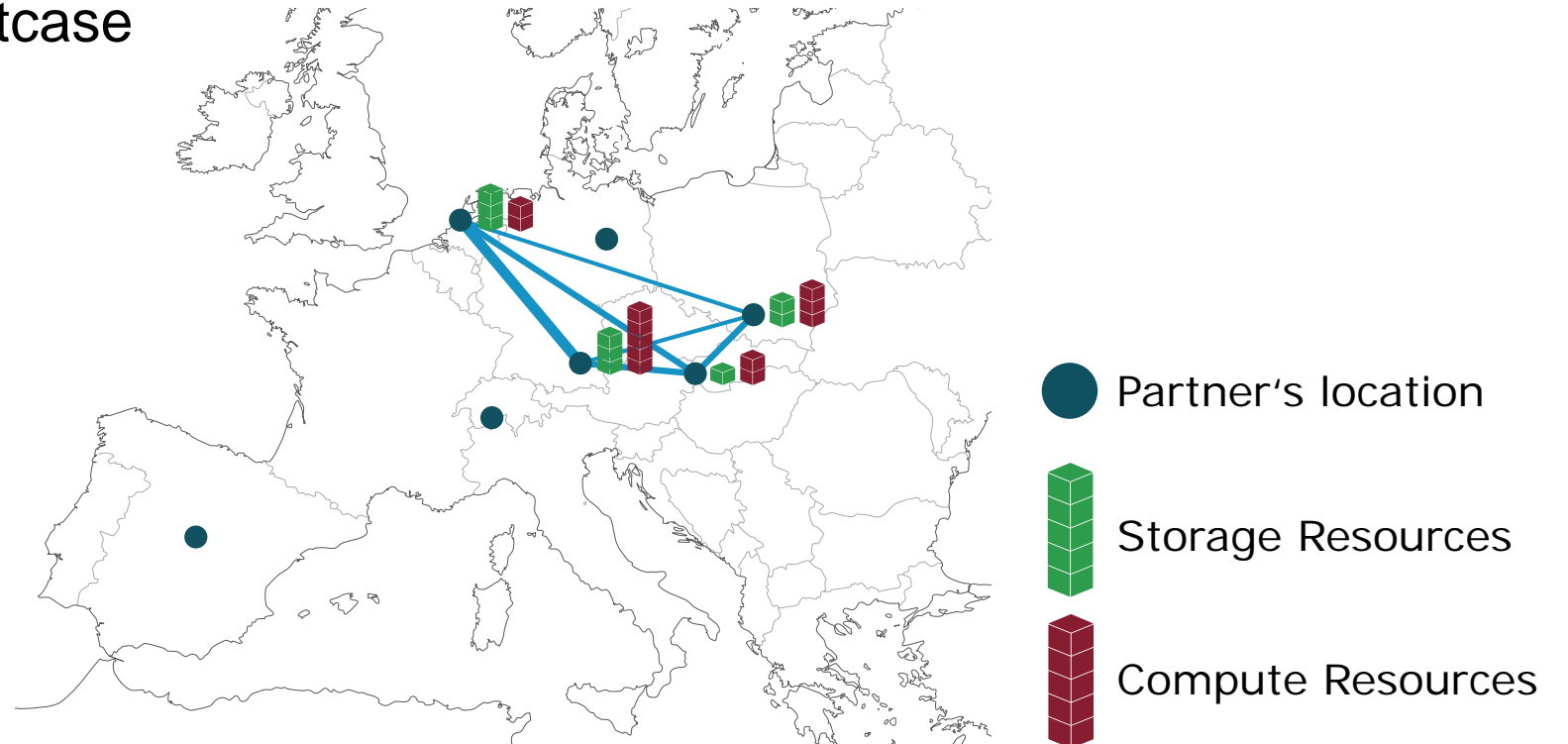
- Splitting of computational domain in subdomains via masks to distribute the workload onto HPC infrastructure
- Aggregation of different simulations into compute jobs
- Parallel merging of decomposed simulation data
- Transformation of model output from native binary format to annotated netCDF
- Storage 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



# 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.

Run pipeline

Select pipeline: Agrocopernicus pipeline

\* Name: Output Maize Europe 2017

\* Mode: automatic

Pipeline steps

Refresh all tags and branches Design new pipeline

Title

Container name: agrocopernicus\_placeholder\_container

Container tag: agrocopernicus\_placeholder\_tag

HPC: Prometheus

Irrigation: true

Seeding date: -15 days

Nutrition factor: 0.25

Phenology factor: 1.0

Set up new pipeline

Validation Demo Validation pipeline (automatic pipeline) Owner: Jan Schmidt

Staging in computation Staging in computation finished successfully.

Validation container computation

Start time Execution time Outputs Status

23 May 08:53	00h 00m 22s	no stdout, no stderr	Finished
--------------	-------------	----------------------	----------

Staging out computation

Validation Demo pipeline inputs

Upload Drop files or folders here File Store Browser v0.22.1-SNAPSHOT

Empty

Validation Demo pipeline outputs

File Store Browser v0.22.1-SNAPSHOT

1G.dat	23/05/19 10:56	1000 MB	Download Share Delete
validation_container_done.txt	23/05/19 10:53	0 B	Download Share Delete



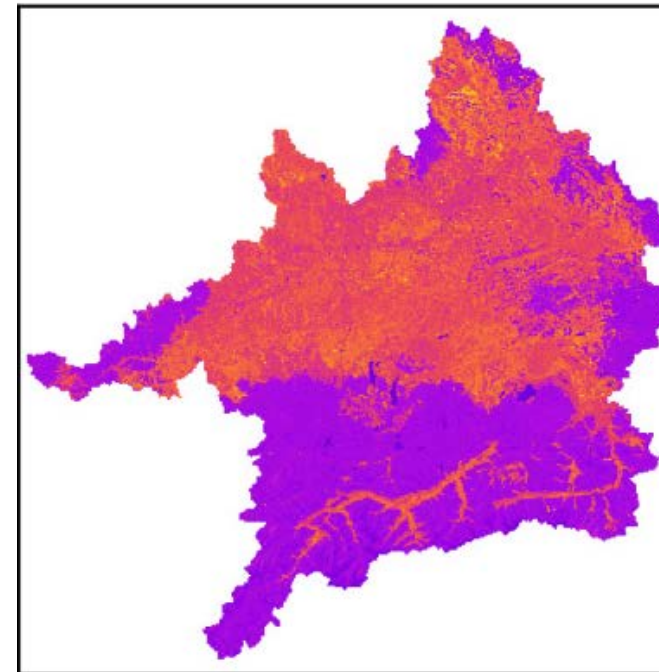
# Domain decomposition of land surface models

# 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

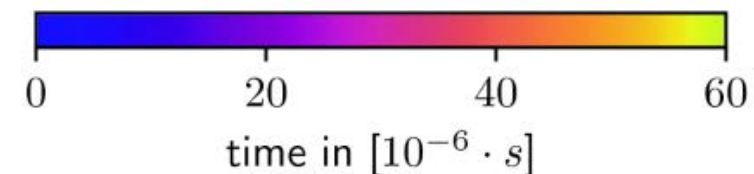
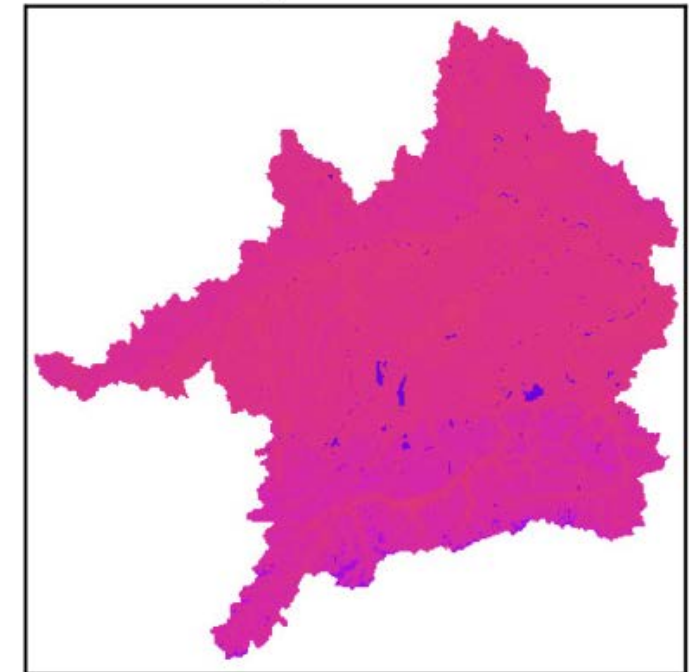
Simulation time  
per cell for one  
particular time  
step

01.02-1999 09:00



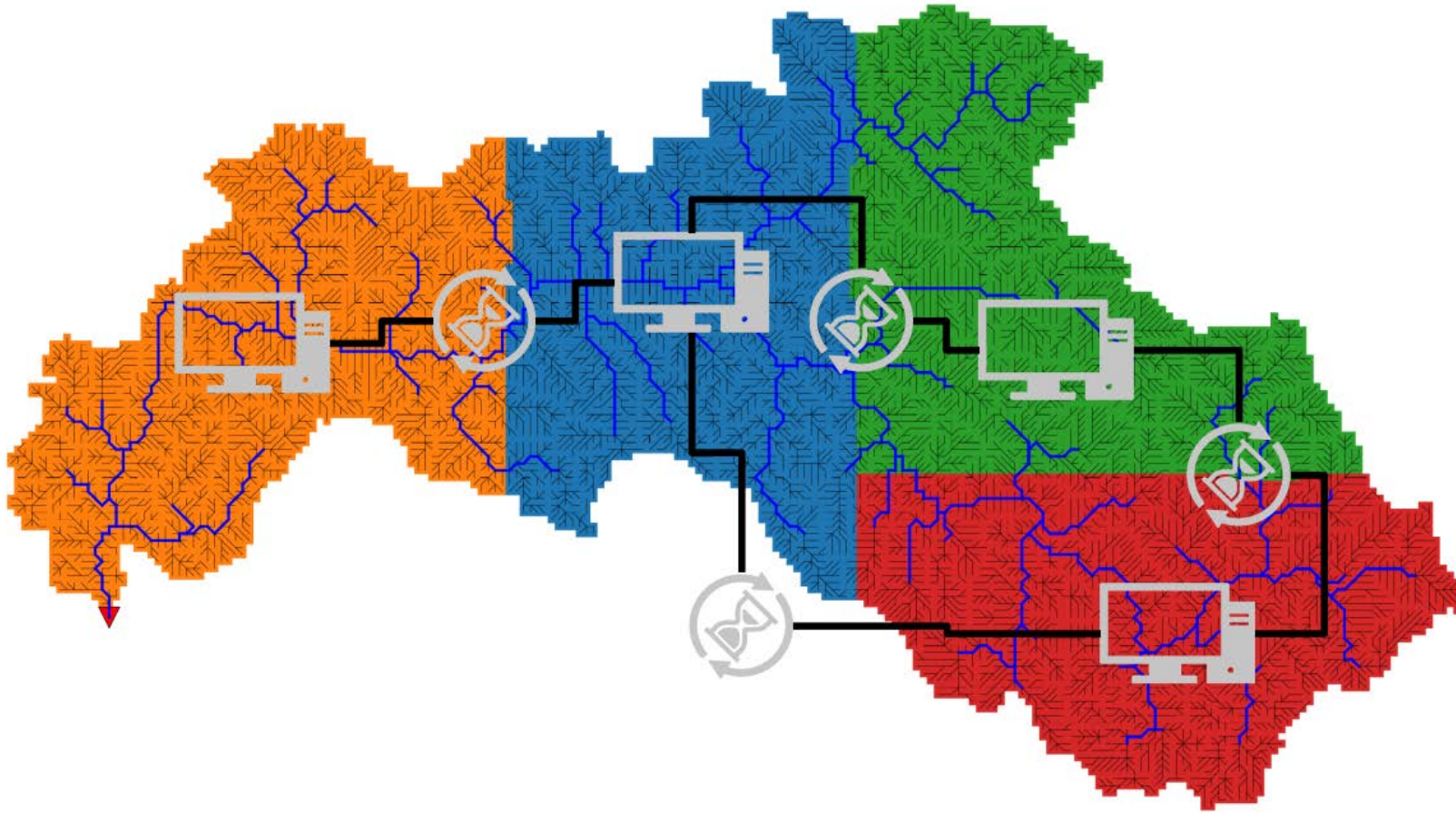
Average  
simulation time  
per cell

1 year mean





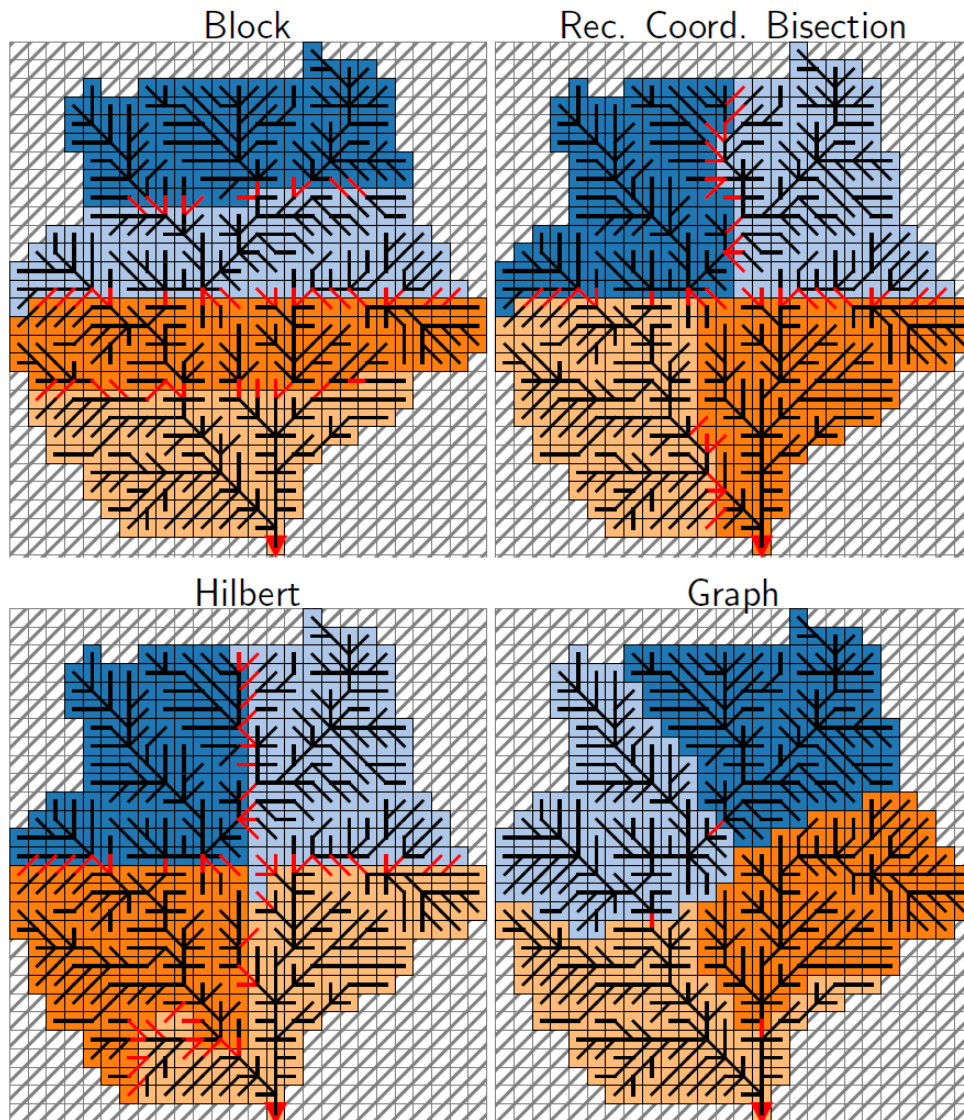
# Main Challenges



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



- Measure cell execution time of a benchmark simulation
  - PROMET simulation of Upper Danube catchment
  - ~76000 cells (1km<sup>2</sup>)
- Perform ,artificial‘ strong scaling experiments with four domain-decomposition methods (Blockwise, Recursive coordinate bisection, Hilbert space-filling curve, Graph-based)
- Compute performance metrics:

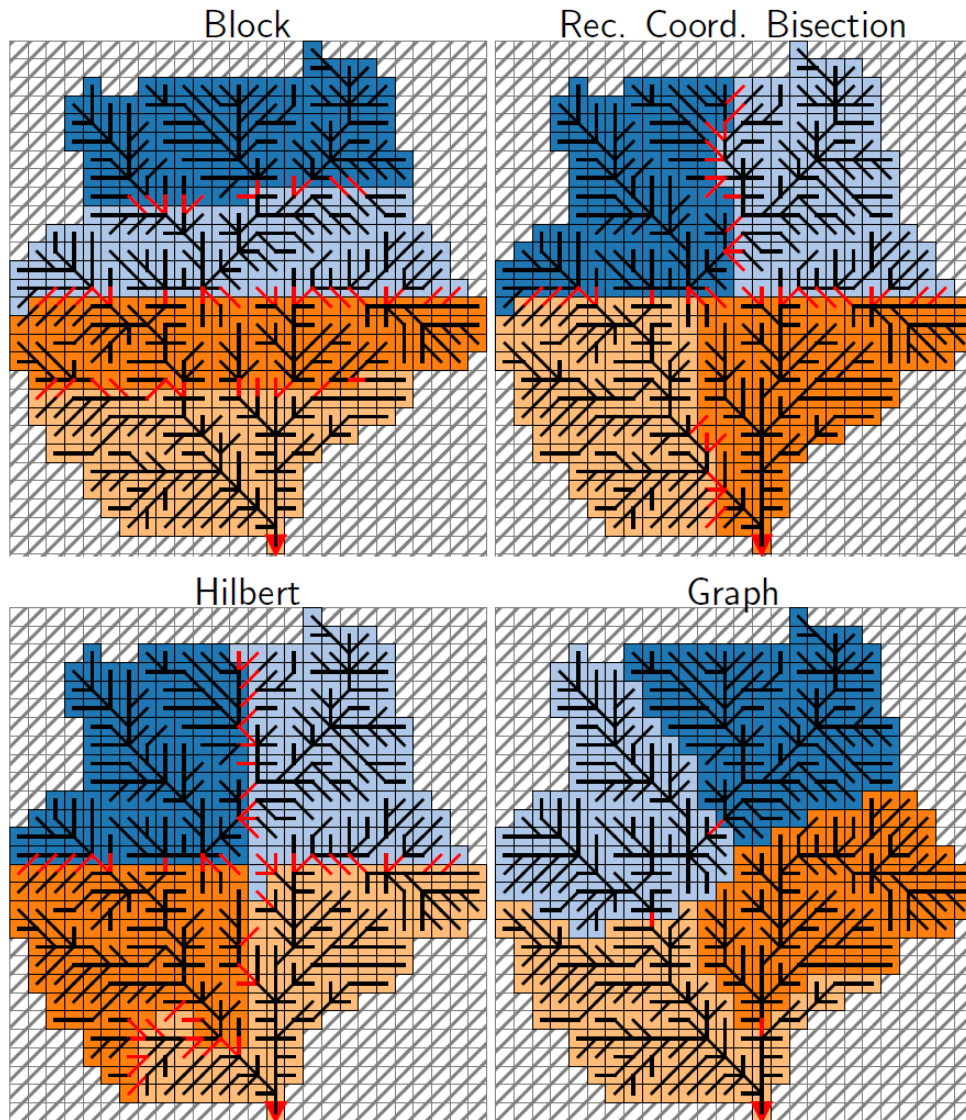
$$\text{theoretical speed up} = \frac{\text{Execution time on 1 proc.}}{\text{Execution time on } n \text{ proc.}}$$

← The higher the better!

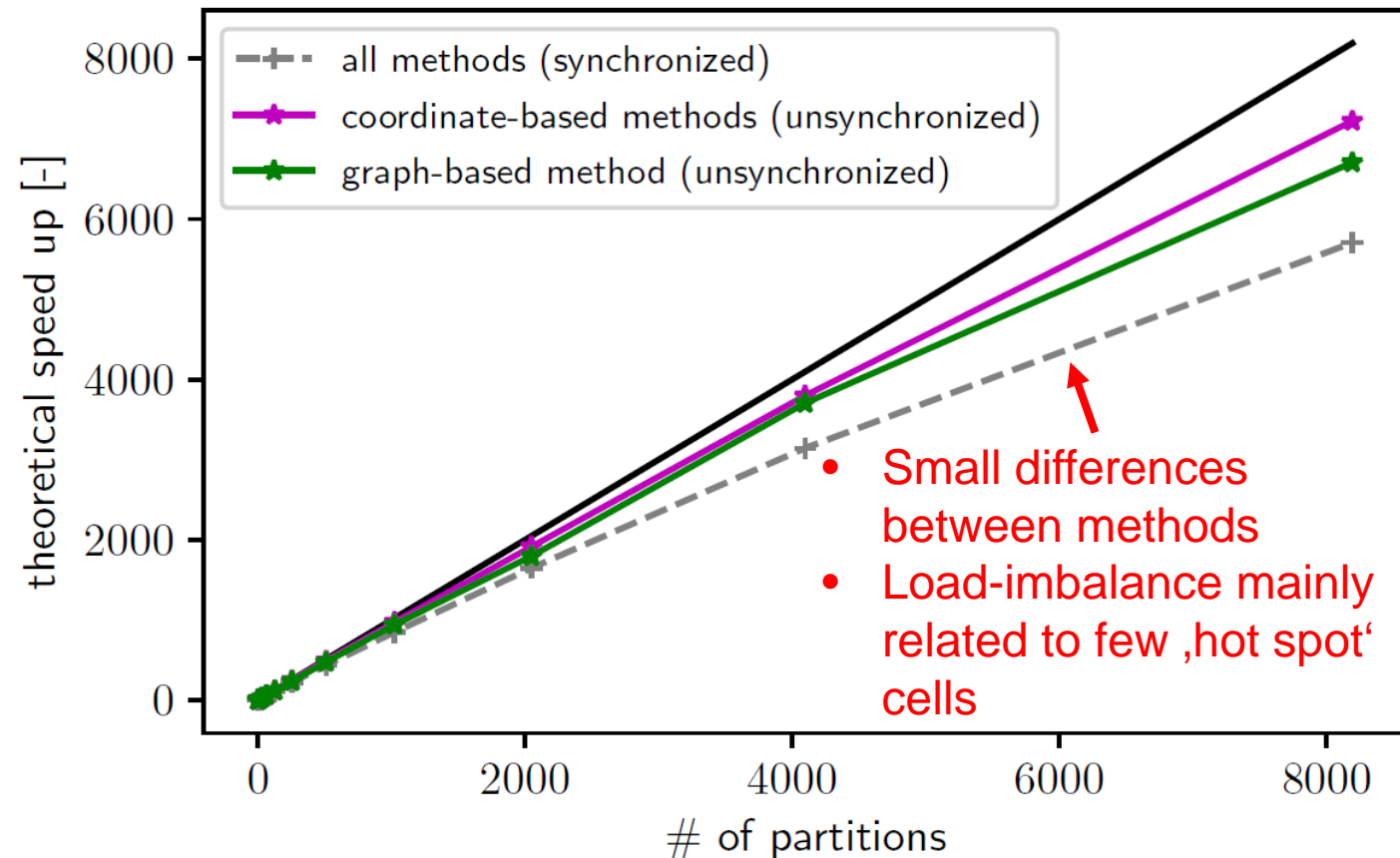
Number of cut flow-paths

← The lower the better!

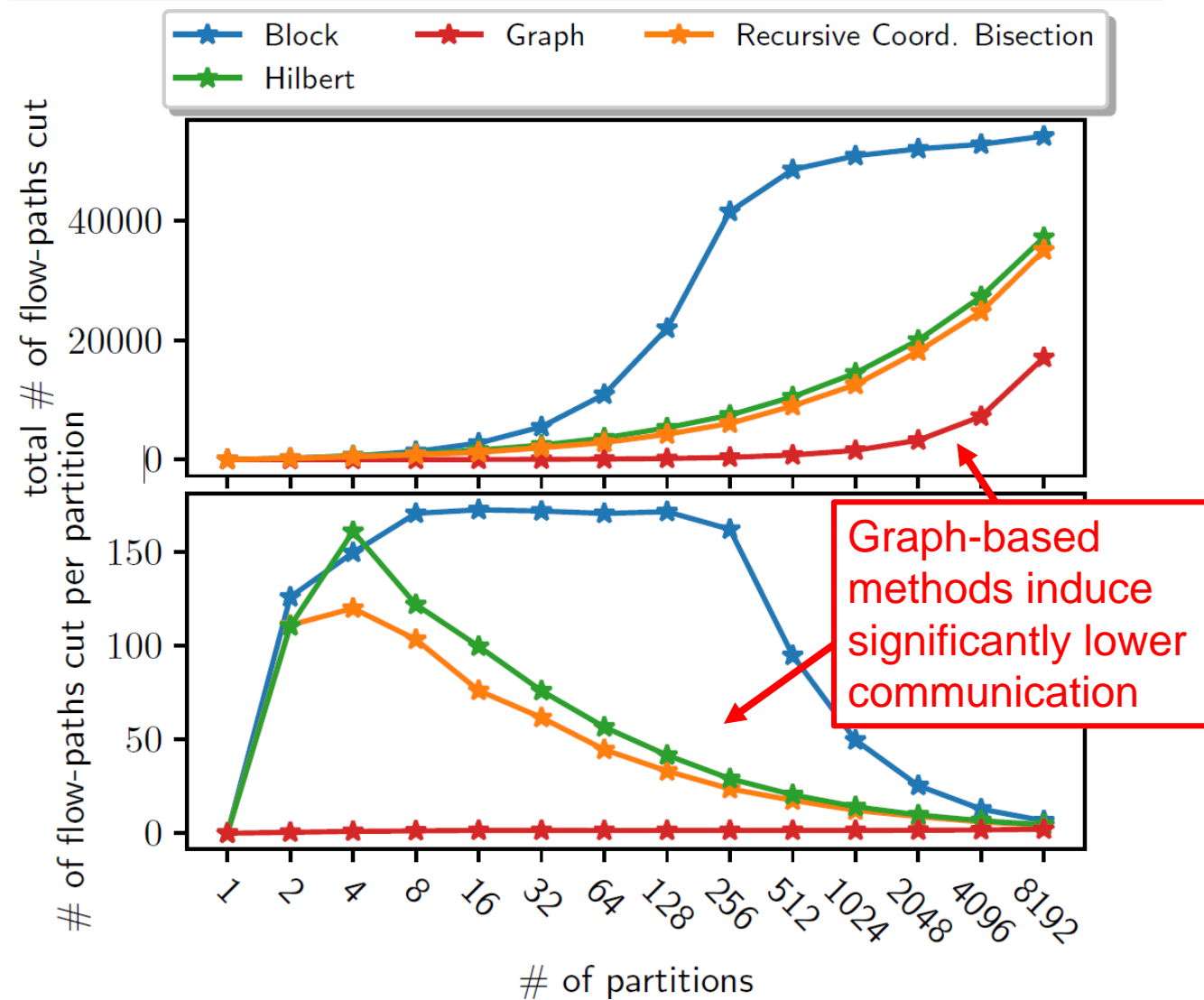
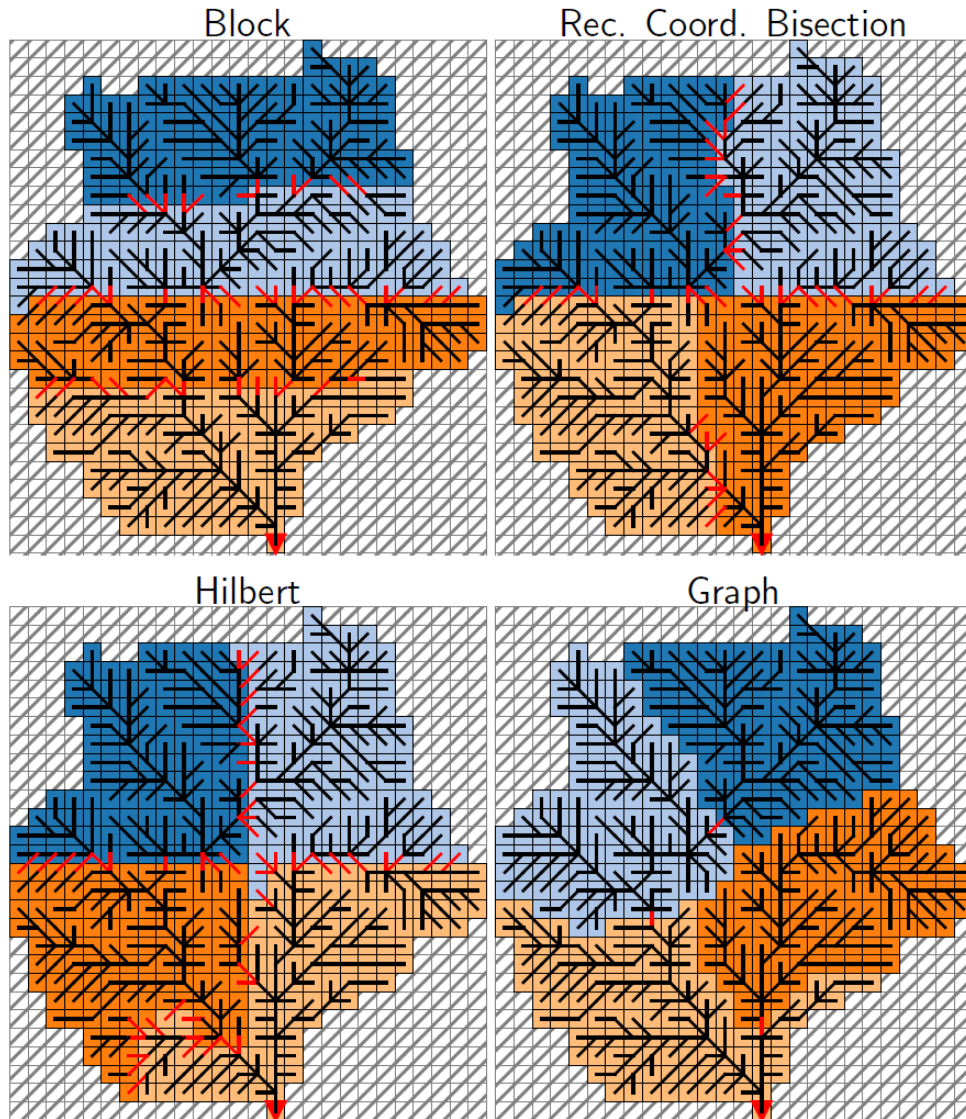
# Theoretical peak speedup



- 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



- 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)

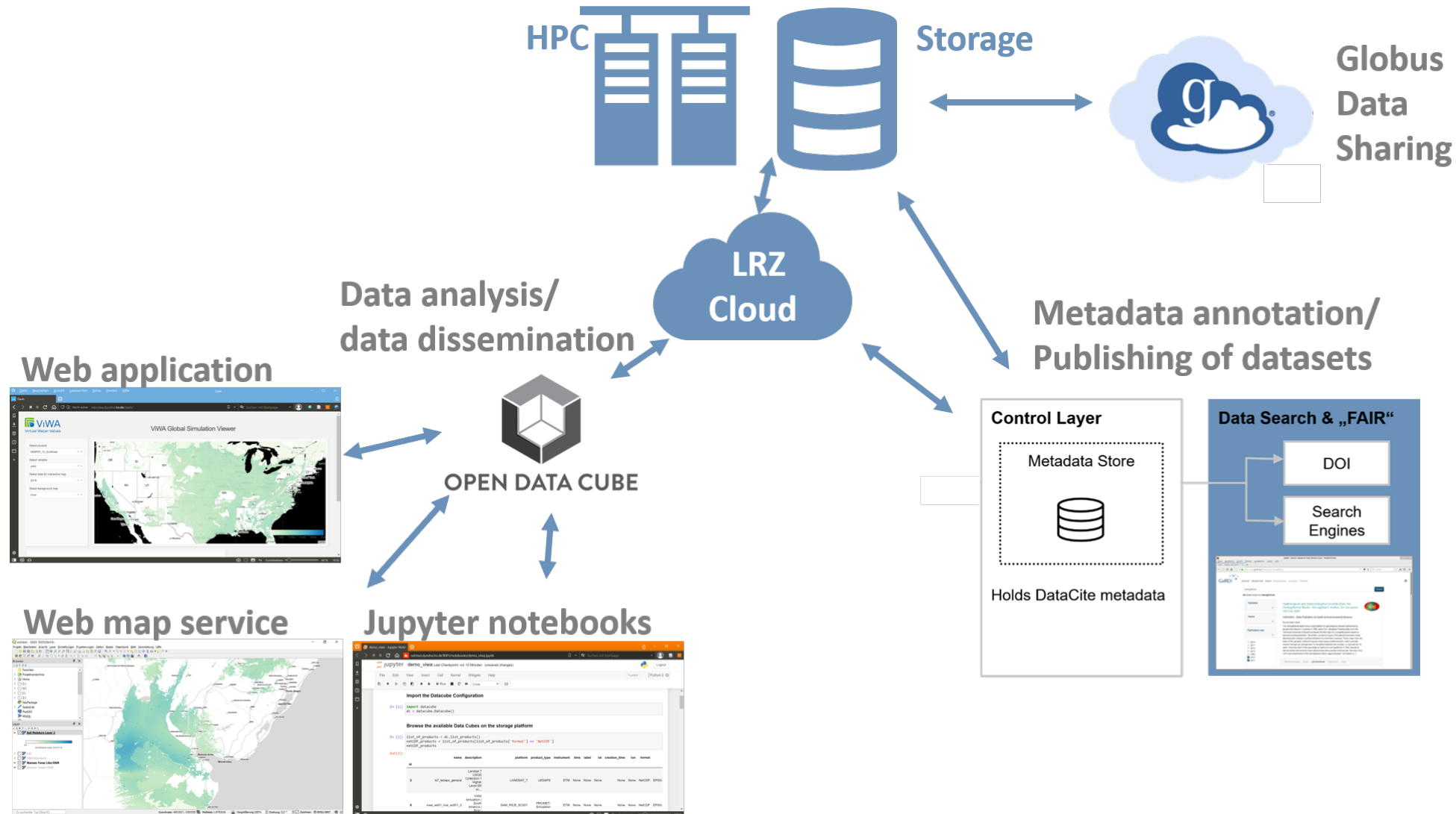
## Goals of data management in ViWA:

- 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)

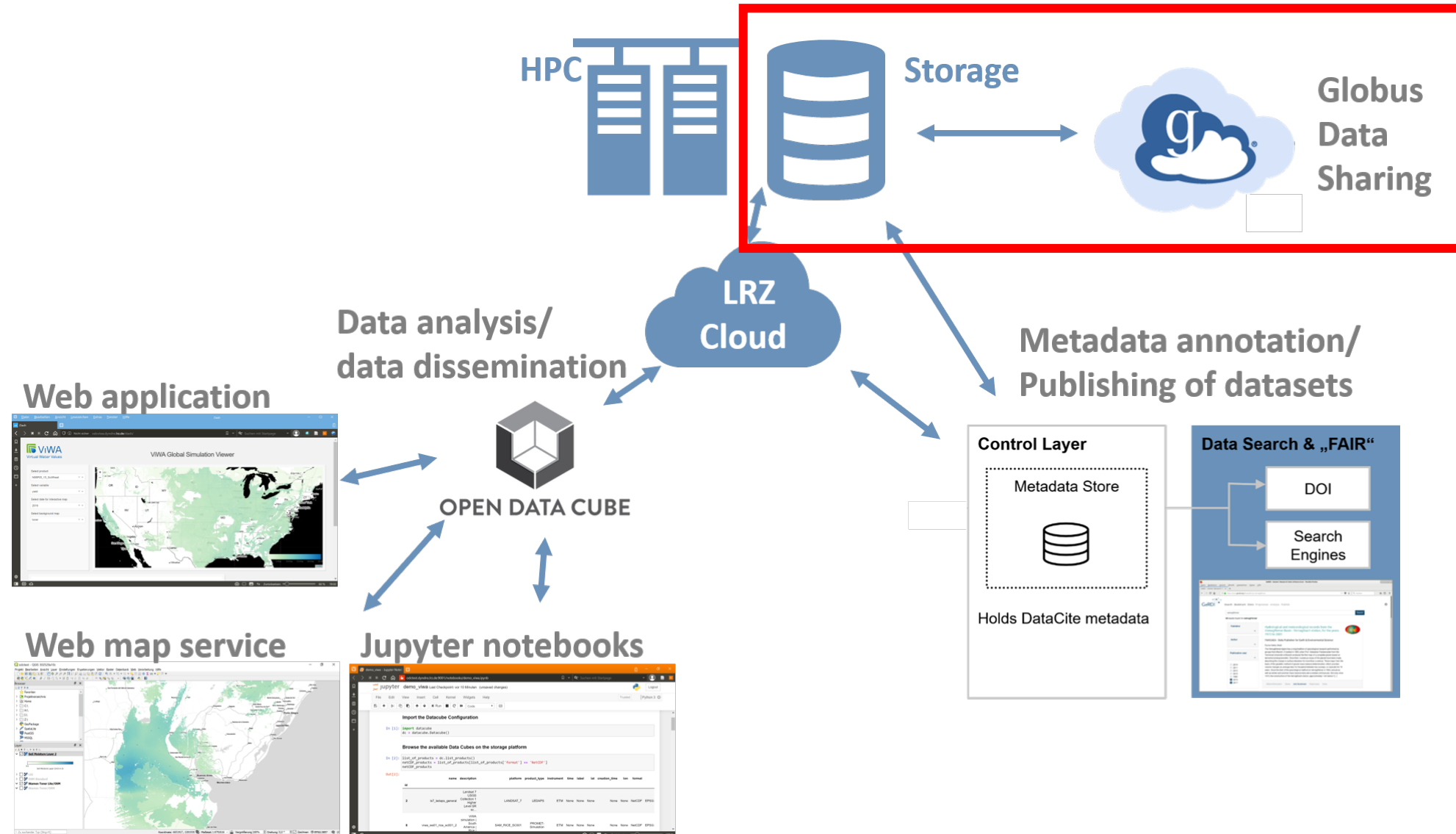
## Approach:

- 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 capabilities 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

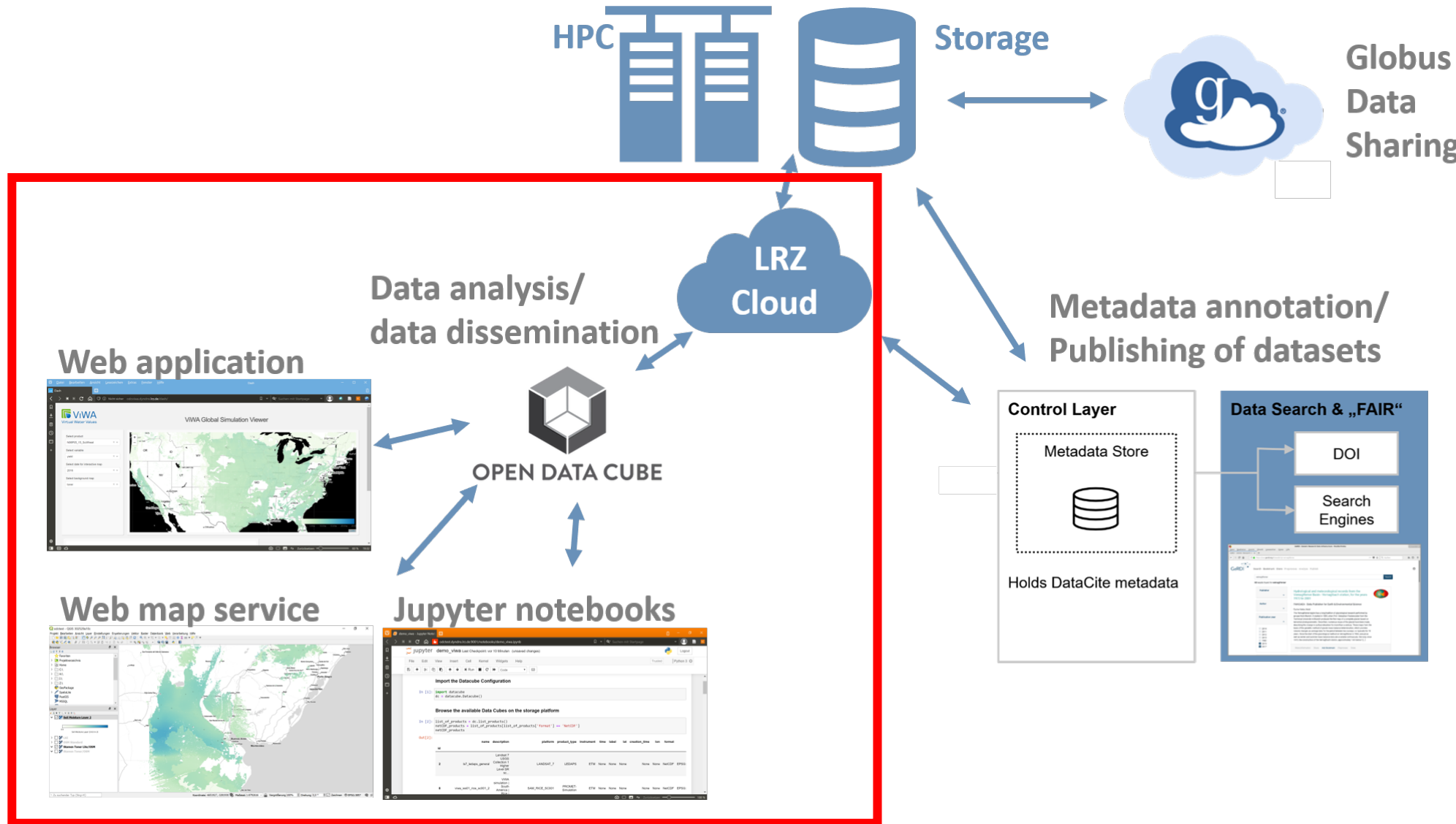




## Steps:

- 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



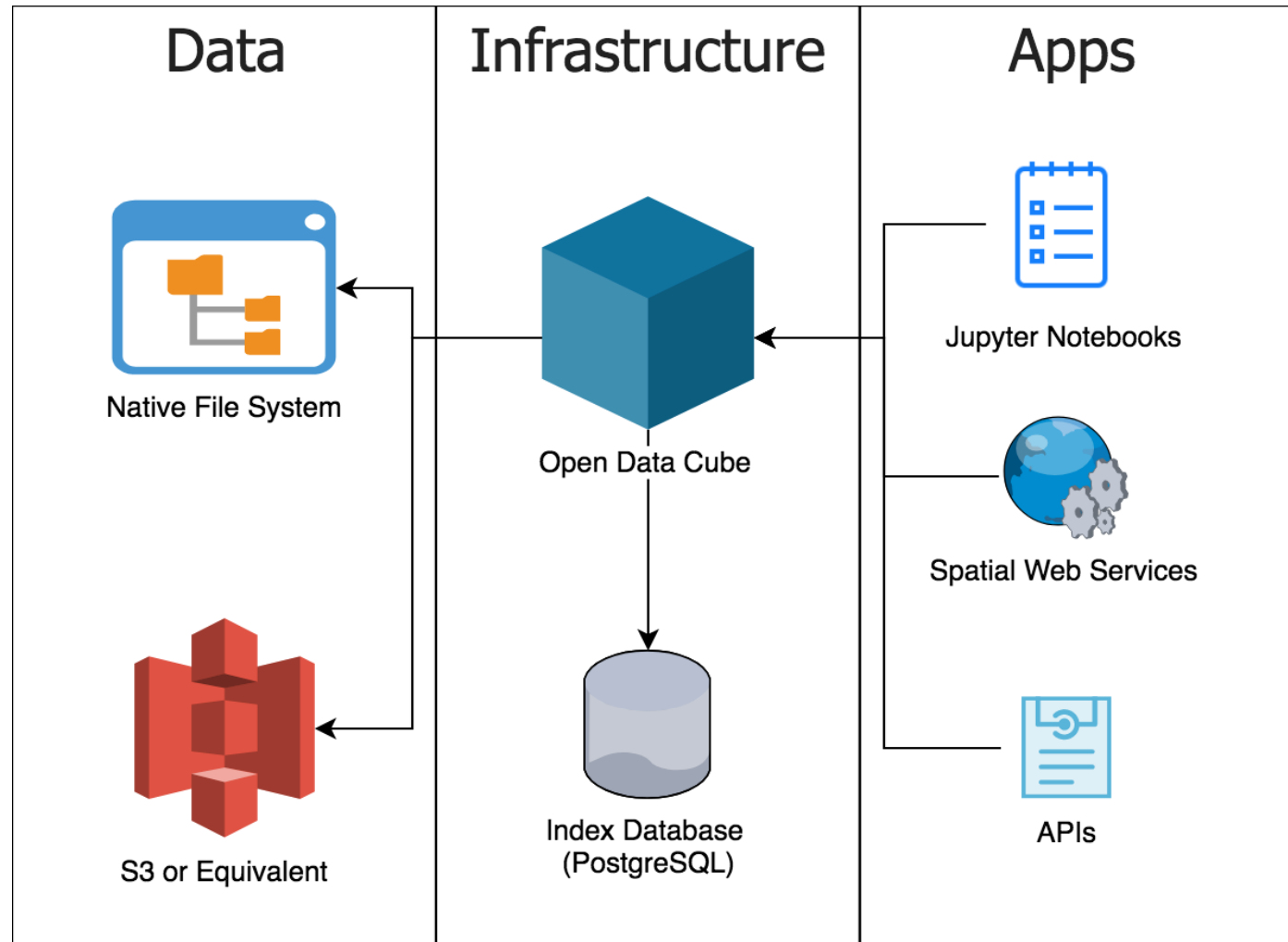
# Data access via Open Data Cube and Web applications (Dash)

- 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  $10^{\circ} \times 10^{\circ}$  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-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 interface for ViWA Global Simulation Viewer.

**Navigation and Controls:**

- Menu: Datei, Bearbeiten, Ansicht, Lesezeichen, Extras, Fenster, Hilfe
- Dashboard: Dash
- Search: Suchen mit Startpage
- Map from ODC WMS service (indicated by a red arrow)

**Data Selection Panel (indicated by a red arrow):**

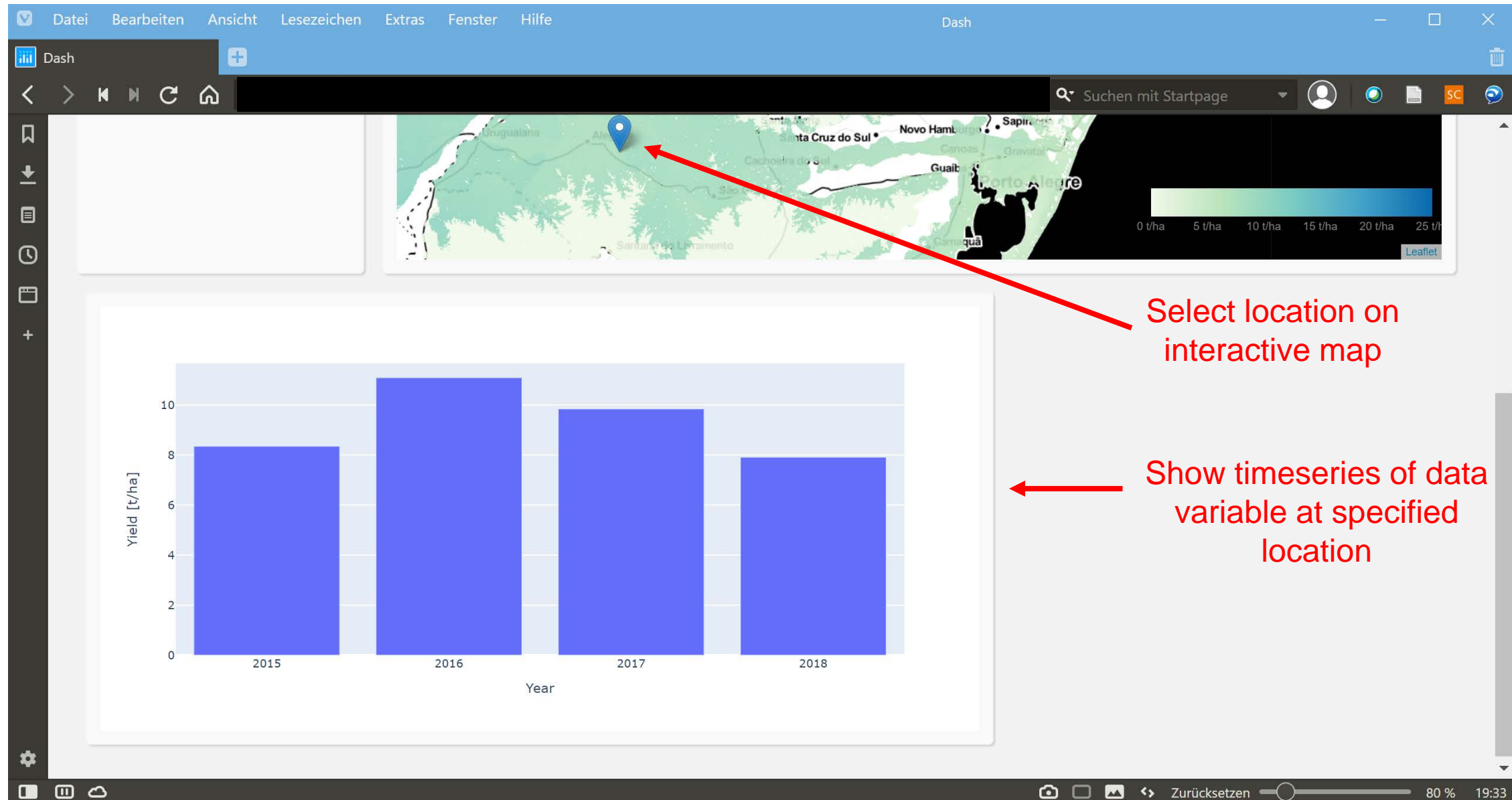
- Select product: N06P05\_15\_SuWheat
- Select variable: yield
- Select date for interactive map: 2016
- Select background map: toner

**Map Display:**

ViWA Global Simulation Viewer

The map shows the United States and Mexico, with a color scale indicating yield values (0 t/ha to 25 t/ha). Major cities and states are labeled. A red arrow points to the map area, indicating it is sourced from the ODC WMS service.

# Web application for viewing of ViWA data products



# Web application for retrieving of ViWA data products

The screenshot shows the 'ViWA Global Simulation Extractor' web application. The interface includes a top navigation bar with menu items: Datei, Bearbeiten, Ansicht, Lesezeichen, Extras, Fenster, and Hilfe. Below this is a browser address bar showing 'Suchen mit Startpage'. The main content area is divided into a left sidebar and a central map area.

**Left Sidebar:**

- Select product:** A dropdown menu showing 'N06P05\_15\_SuWheat'.
- Select variables:** A list of checkboxes with the following items checked: cropfailure, gdd, growingdays, precsum, and yield.
- Select upper left coordinates:** Two input fields containing '-50' and '-15'.
- Select dx/dy:** Two input fields containing '8' and '8'.
- Select time range:** Two date input fields showing '01/01/2015' and '12/31/2018', separated by a right-pointing arrow.
- START EXTRACTION:** A button located below the time range selection.

**Central Map Area:**

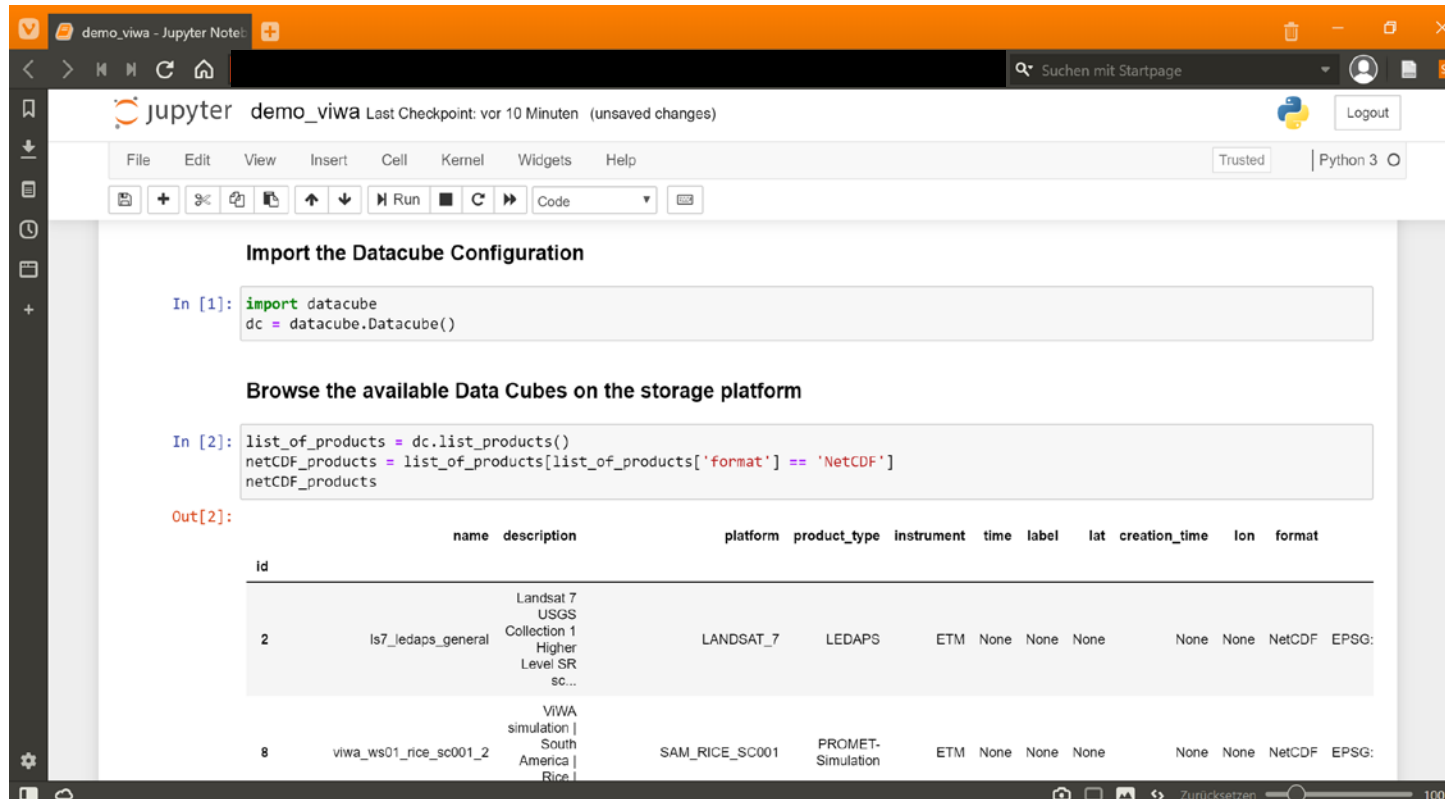
- A map of South America, specifically focusing on Brazil, with labels for various cities including La Paz, Santa Cruz, Brasília, Goiânia, Belo Horizonte, Rio de Janeiro, São Paulo, Curitiba, Asunción, Ciudad del Este, and Porto Alegre.
- A blue rectangular selection box is drawn on the map, covering a region in central Brazil.
- Below the map, there are two dropdown menus: the first is labeled 'toner' and the second is labeled '2015'.
- At the bottom of the map area, a list of generated file names is shown, with the first one being [N06P05\\_15\\_SuWheat -50 -15 8x8degree 2015-01-01 2018-12-31.nc](#).

**Annotations:**

- A red arrow points from the text 'Data selection' to the 'START EXTRACTION' button.
- A red arrow points from the text 'Selection shown on interactive map' to the blue selection box on the map.
- A red arrow points from the text 'Selection download (netCDF file)' to the first file name in the list at the bottom.

# Jupyter notebook for data analysis

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



The screenshot shows a Jupyter Notebook window titled "demo\_viwa - Jupyter Note". The notebook has two input cells. The first cell, labeled "In [1]:", contains the code to import the Datacube module and create a Datacube object. The second cell, labeled "In [2]:", contains the code to list products and filter for NetCDF format. The output of the second cell, labeled "Out[2]:", is a table of products.

```
In [1]: import datacube
dc = datacube.Datacube()

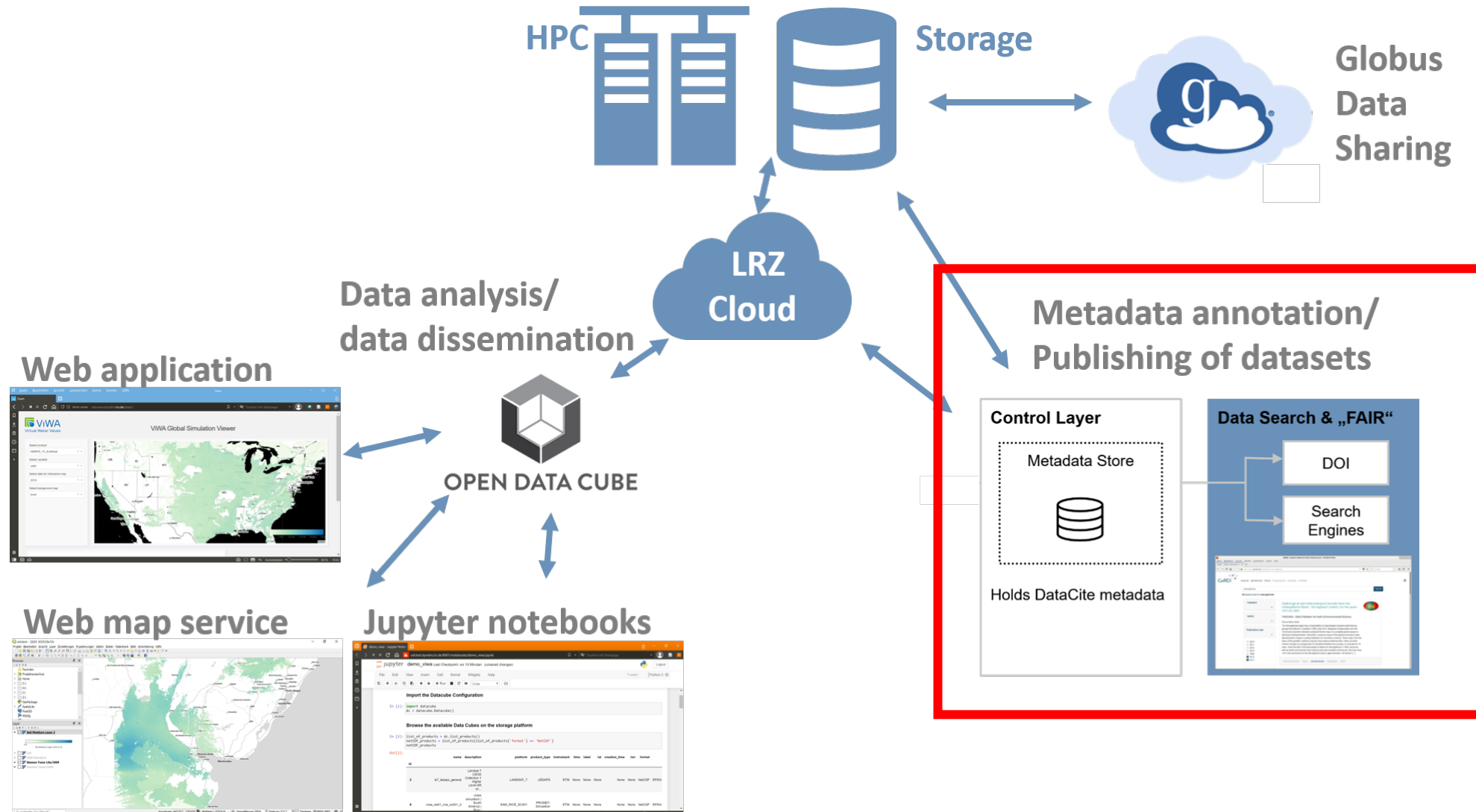
In [2]: list_of_products = dc.list_products()
netCDF_products = list_of_products[list_of_products['format'] == 'NetCDF']
netCDF_products
```

Out[2]:

id	name	description	platform	product_type	instrument	time	label	lat	creation_time	lon	format
2	ls7_ledaps_general	Landsat 7 USGS Collection 1 Higher Level SR sc...	LANDSAT_7	LEDAPS	ETM	None	None	None	None	None	NetCDF EPSG:
8	viwa_ws01_rice_sc001_2	VIWA simulation   South America   Rice I	SAM_RICE_SC001	PROMET- Simulation	ETM	None	None	None	None	None	NetCDF EPSG:



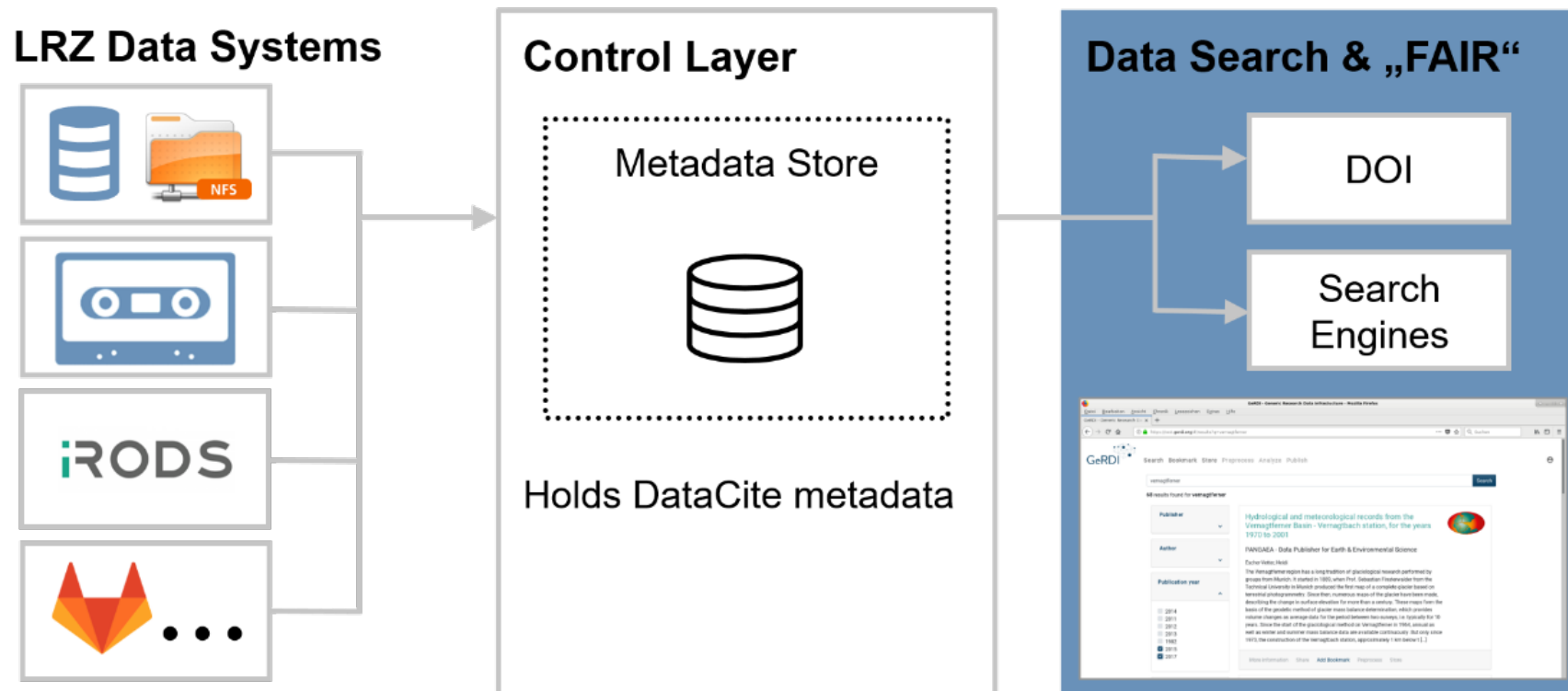
# Data management architecture for ViWA simulation data



- 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



- 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