

Coupled Storage System for Efficient Management of Self-Describing Data Formats

Michael Kuhn, Kira Duwe

michael.kuhn@ovgu.de, kira.duwe@ovgu.de

<https://cosemos.de/>



OTTO VON GUERICKE
UNIVERSITÄT
MAGDEBURG

PROJECT DESCRIPTION

- CoSEMoS's goal is to rethink the **architecture of storage systems**
 - DFG project to improve performance and data management
 - Built upon JULEA: Modern C11 code, available as open source
 - Currently in its second year with a funding period of 2019–2022
- JULEA provides a **flexible storage framework**
 - Contains necessary building blocks for storage systems
 - Facilitates rapid prototyping and evaluation
- Runs in **user space** and has **few dependencies**
 - Kernel code increases complexity and fragility significantly
 - Possible to use on clusters without root access

PROBLEM STATEMENT

- **Self-describing data formats (SDDFs)** widely used to exchange data
 - Structural information is encoded in the files themselves
 - Files can be accessed and interpreted without prior knowledge
- 1. **Weak treatment of different types of metadata**
 - Two different types of metadata
 - **File system metadata** is stored on the metadata servers
 - **File metadata** (for example, attributes or additional annotations) is stored within SDDF files on the data servers
 - Strict separation of metadata leads to inefficient file access
- 2. **Static I/O semantics**
 - Strict consistency and coherence semantics due to POSIX
 - Static approaches are unable to satisfy all requirements
- 3. **Inefficient data placement**
 - Hierarchical structuring of different hardware is used
 - Data movement across storage tiers is an expensive operation
 - Hardware is available, new approaches need to be developed

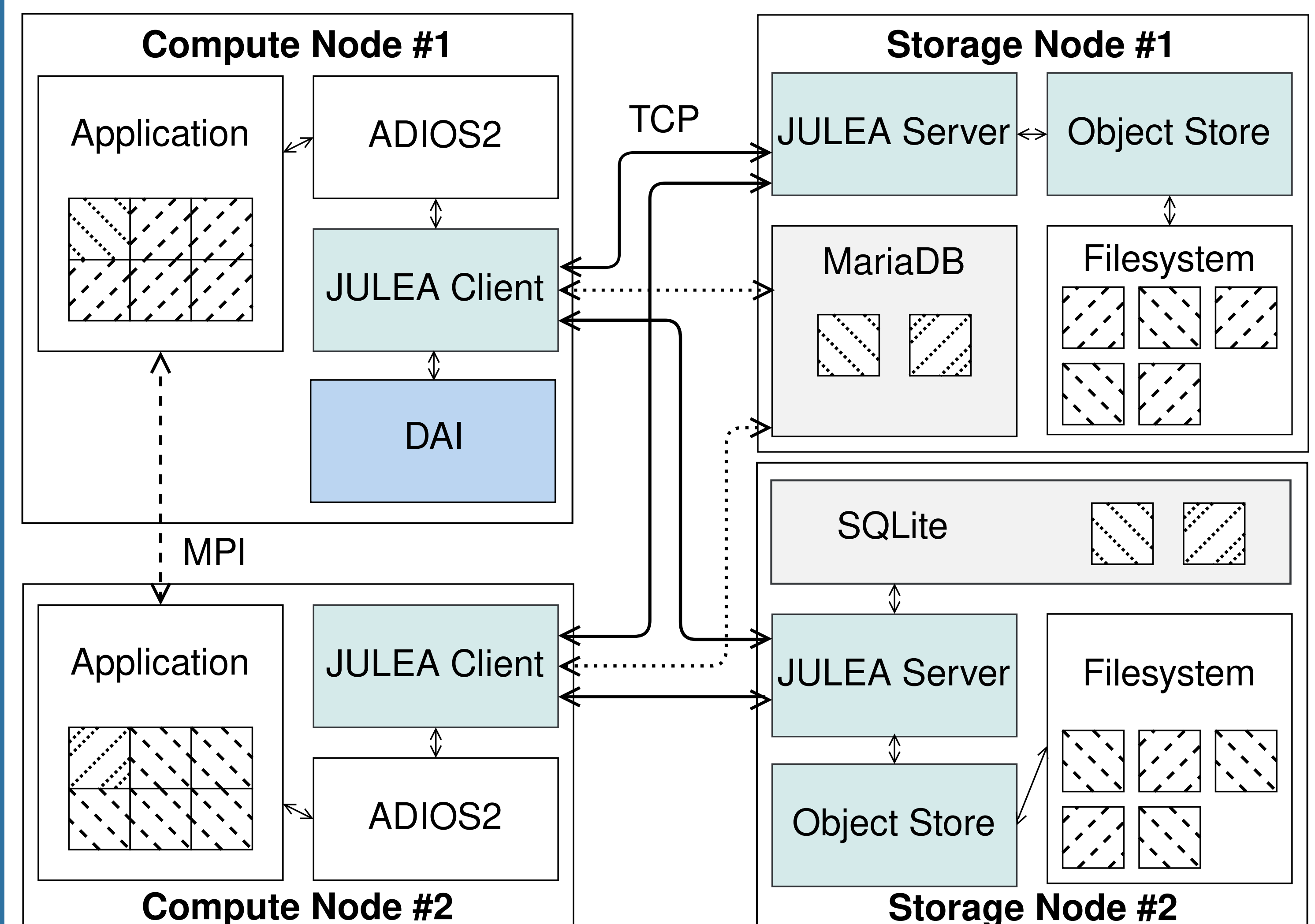
OBJECTIVES

1. **Global metadata management**
 - Closely **couple storage system and self-describing data formats**
 - All metadata handled by metadata servers
 - Optimize metadata accesses using database technologies
 - Novel data management approaches via a data analysis interface
 - Query file metadata across multiple files in a unified way
2. **Adaptable I/O semantics**
 - Possible to dynamically adapt semantics
 - Provide appropriate interfaces for applications and libraries
 - Specify **requirements regarding atomicity, consistency etc.**
3. **Intelligent storage selection**
 - Use structural information for **informed data placement decisions**
 - Improve performance by optimizing data placement
 - Different parts of self-describing files can be put on different tiers

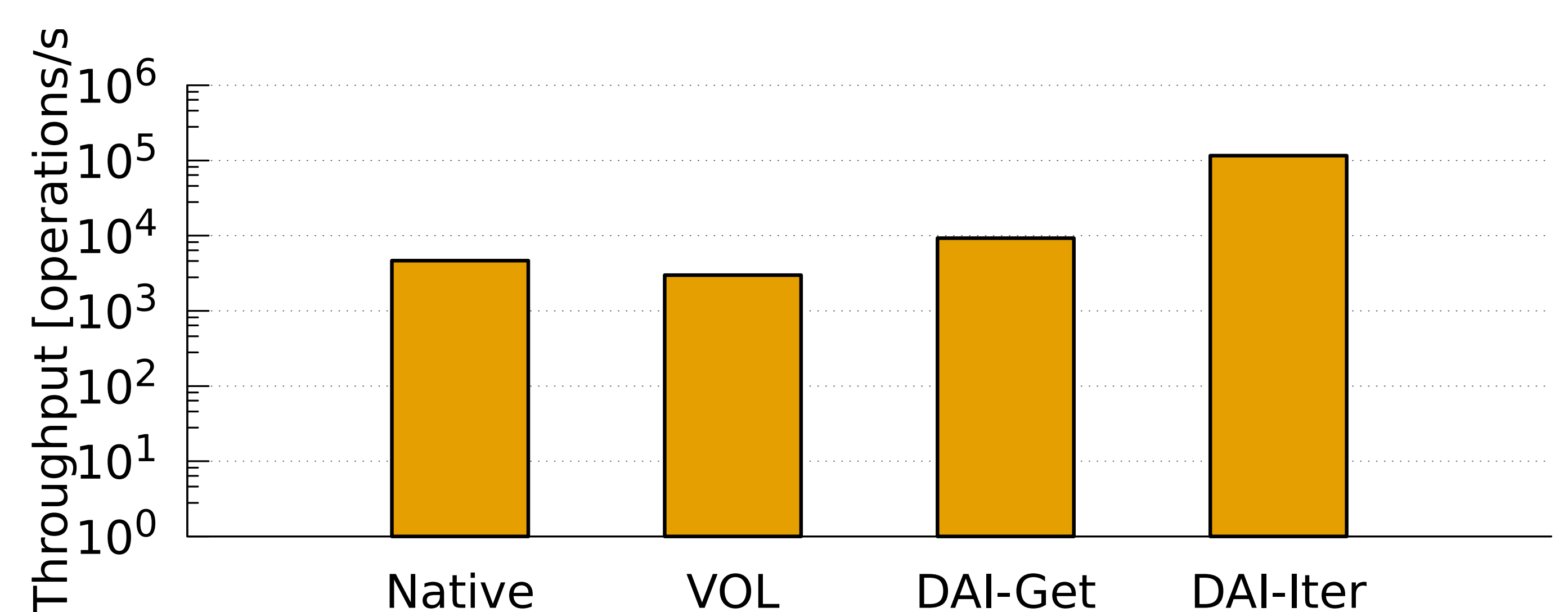
PARTNERS

- German Climate Computing Center (Prof. Dr. Thomas Ludwig)
- Intel (Johann Lombardi)
- Max Planck Institute for Meteorology (Uwe Schulzweida)

ARCHITECTURE AND EVALUATION



- JULEA (**turquoise**), database (**gray**) and data analysis interface (**blue**)



- Use case: Reading all attributes from an HDF5 file
 - Native HDF5 and JULEA VOL plugin are inefficient due to API overhead
 - DAI-Get: Fetch attributes separately, DAI-Iter: Iterate over attributes
- DAI-Get and DAI-Iter approaches offer **speedups of 2 and 25**, respectively

WORK PACKAGES

- WP1: **Application Interface**
 - T1.1 SDDF Interface ✓
 - T1.2 Application Requirements and Semantics ⌘
 - T1.3 Data Analysis Interface ⚙
- WP2: **Storage Tier Selector and Global Metadata Manager**
 - T2.1 Database Backend ✓
 - T2.2 Database Client ✓
 - T2.3 Metadata Backend Selection ⌘
 - T2.4 Data Storage Tiering ⚙
- WP3: **Evaluation and Dissemination**
 - T3.1 Compatibility Tests ⌘
 - T3.2 Case Study ⌘
 - T3.3 Workshop Organization ⚙

ACKNOWLEDGEMENTS AND LINKS

CoSEMoS is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 417705296. All results are being published on the project website and integrated into JULEA:

- <https://cosemos.de/>
- <https://github.com/julea-io>