



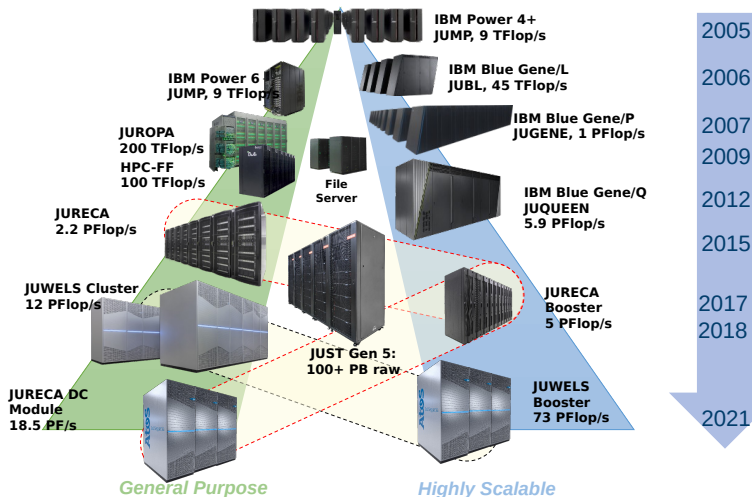
RDM FOR DATA FROM HPC

06.04.2022 | Sander Apweiler | JSC

1. HPC Systems
2. Stagesystems
3. Transfertools
4. B2SHARE
5. DataPub
6. Jülich DATA

HPC Systems

HPC SYSTEMS - THE ROAD ...



HPC SYSTEMS - ... TO THE PRESENT

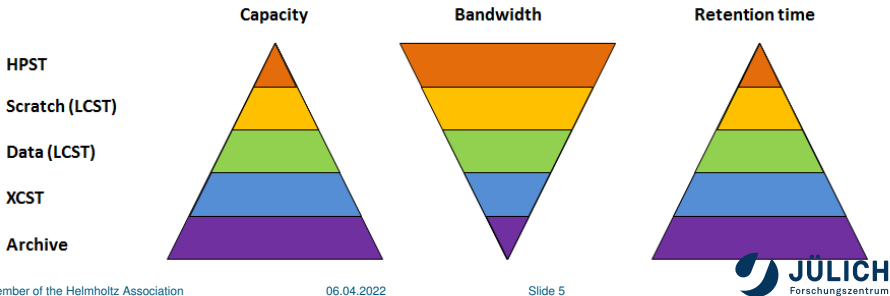
	JUWELS		JURECA	
	Cluster	Booster	Data Centric	Booster
Peak Petaflops/s	12,3	73	18,5	5
Nodes	2567	936	768	1640
Accelerators	224 GPUs	3744 GPUs	768 GPUs	1640 KNLs

- Beside of the large HPC clusters
 - Prototype as a blueprint for modular super computers: DEEP
 - Small cluster for user from campus or specific communities: HDF-ML, JUZEA1, JUAMS
 - Cluster/cloud hybrid: JUSUF
 - Data access cluster: JUDAC
 - Others: HDF-cloud

Storagesystems

STORAGESYSTEMS

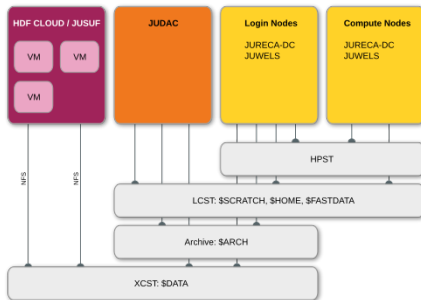
- High Performance Storage Tier (HPST):
NVMe based Storage, 110 servers, 1100 drives
- Large Capacity Storage Tier (LCST):
5th Gen. of JUST, 22 building blocks, 7500 x 10TB discs
- Extended Capacity Storage Tier (XCST):
7 GPFS Building Blocks using 7000 discs (10TB, 12TB, 16TB)
- Archive: Tape storage (Backup + GPFS&TSM-HSM),
1 building block + 3 tape libraries, \approx 34500 tapes



STORAGESYSTEMS

Filesystems:

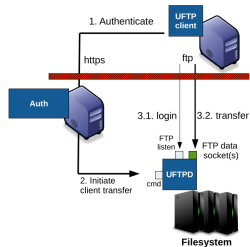
- \$HOME: user's home directory inside GPFS
- \$SCRATCH: temporary storage location for applications with large size and I/O demands, deleted after 90 days
- \$DATA: large capacity for storing and sharing data
- \$FASTDATA: storage for large projects in collaboration with JSC
- \$ARCH: storage for all files not in use for a longer time



Transfertools

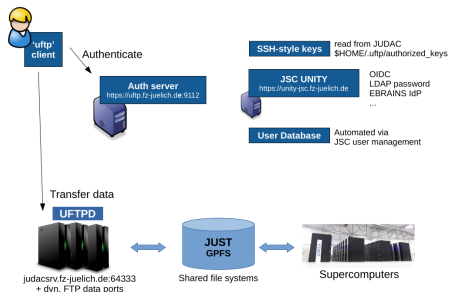
TRANSFERTOOLS - UFTP

- „UNICORE FTP“ \Rightarrow it grew out of the need for a high-performance data transfer for the UNICORE federated middleware
- Client/server file transfer system, based on FTP, with flexible, federated authentication
- Extended with
 - Compression and/or encryption of data stream(s)
 - Multiple TCP streams per data connection
 - Bandwidth management (both server and client)
- UFTPD Server supports partial file reads/writes
- UFTPD creates multiple sessions in parallel for scalability
- Multiple UFTPD backends are possible (round-robin)



TRANSFERTOOLS - UFTP

- Share data with other users via UFTP
- Command line tool with well known commands like: cp, ls, ...
- Authentication via
 - ssh-style keys (recommended, at JSC it must differ from ssh key)
 - OAuth tokens
 - username/password
- Available at JSC, HLRS and LRZ



Full presentation about UFTP at

DOI:10.23728/b2share.83b63c4bf4674e638711ef9bddb93bfb or

PID: 11304/7d62b0f3-02a3-4d82-aa47-436fa9778cda

B2SHARE

B2SHARE

- Developed by EUDAT CDI
- Provided by JSC as community instance
- URL: <https://b2share.fz-juelich.de>
- Audience: researchers from Forschungszentrum Jülich and related communities
- Quota: 50GB per file; 200GB per deposit
- Listed as repository on re3data
- Accepted repository for Springer publication
- Harvested by B2FIND

GO TO EUDAT WEBSITE

Search records for...

HELP COMMUNITIES UPLOAD CONTACT

Store and publish your research data

Search in public datasets or register as a user to upload and publish your data!

Create record Create a new record

Latest records

Segmentation Models for 3D neuroanatomy: segmentation
26 Nov 2022 by Ina Schab, Edgar David Braker, Martin
With this work, we provide a 3D whole brain segmentation model that outperforms the 2D approach MedialSurfaces. Furthermore, we explore the efficacy of using...
segmenting and its parameters. This re

Global, high-resolution mapping of tropospheric ozone – explainable machine learning and impact of uncertainties – Research Data
24 Feb 2022 by International, China, Switzerland, France, Brazil, Australia, Portugal, Austria, Sweden, Spain
This ozone code contains all methods that is being used in ozone mapping project. In addition, it contains scripts to run both explainable AI methods and methods used to study the impact of uncertainty.

Global_Aerosol_DPP_profile_reanalysis_from_MERRA-2_v01.1
21 Jun 2022 by Copernicus, Belgium, Virginia, The Gambia, Netherlands, India, France, Mexico, Iran
Global distribution of monthly aerosol optical properties (production, single scattering albedo and asymmetry parameter) created from the MSA-MERRA-2.3 hourly reanalysis profiles of aerosol loading.

TOAR II data gathering event, 21 Jun 2022
21 Jun 2022 by Schuster, Isenroth, A., Schröder, S., Selke, N., Bombard, M.
The Tropospheric Ozone Assessment Report (TOAR) is an initiative of the International Global Atmospheric Chemistry (IGAC) project. TOAR-II is the second phase of TOAR. It builds on the successful comp.

Global_Aerosol_DPP_profile_reanalysis_from_MERRA-2_v01.1
21 Jun 2022
See Global_Aerosol_DPP_profile_reanalysis_from_MERRA-2_v01.1 for full description.

Global_Aerosol_DPP_profile_reanalysis_from_MERRA-2_v01.1
21 Jun 2022
See Global_Aerosol_DPP_profile_reanalysis_from_MERRA-2_v01.1 for full description.

Global_Aerosol_DPP_profile_reanalysis_from_MERRA-2_v01.1
21 Jun 2022
See Global_Aerosol_DPP_profile_reanalysis_from_MERRA-2_v01.1 for full description.

Archived material from the TOAR II Manuscript Writing Event, Nov 16 - 18, 2021
30 Nov 2021 by Cooper, Oleson, B., Schmidt, Martin G., Chatterjee, Lisa, Kanya, Hogg, Morita, Rasmus, Raut, Iran, von Schroederwasser, Julia, Segal, Rodrigo, Sarika, Bhatia, Venkatesh, Heine, Zhong, Gu
The Tropospheric Ozone Assessment Report (TOAR) is an initiative of the International Global Atmospheric Chemistry (IGAC) project. TOAR-II is the second phase of TOAR. It builds on the successful comp.

GO TO EUDAT WEBSITE

Search records for...

HELP COMMUNITIES UPLOAD CONTACT

Records

Title: Select the schema and enter the deposit title

Community

Aalto
 B2SHARE
 CLARIN
 DASH
 ESCAT
 EUDAT
 EUDAT
 EURO-CORDEX
 GWF
 HANSEN
 LTK
 NEM
 RDA
 TOAR

Create Draft Record

You can also update the data in an existing record by creating a new version of that record. Search for the 'Create new version' button on the record's page.

GO TO EUDAT WEBSITE

Search records for...

HELP COMMUNITIES UPLOAD CONTACT

Records

Editing draft version Delete draft

Add files

Drop files here, or click to select files

Add B2DROP files

Basic fields

Community*

Title* Select the schema and enter the deposit title add del clear

Description

Type* add del clear

Creators add del clear

Open Access* True

Embargo Date add del clear

License add del clear

URL add del clear

Shortcuts add del clear

Keywords add del clear

Contact Email

Publication Date

Show Form details

Submit draft for publication

When the draft is published it will be assigned a PID and a DOI, making it publicly visible. Please note that the published record's files can no longer be modified by its owner.

This publication will get the following DOI: 10.34733/242406482423242115214601917 add del clear

The draft is up to date add del clear

Discard Changes

B2SHARE

- Attaches DOI to each deposit and PID to each deposit and all files
- Supports multiple versions of a deposit
- Supports individual (community) metadata schemas
- Usage with web UI or REST API
- Command line tool for basic usage available
- Link to other data storage is possible, if the data is stored somewhere else
- More information: <https://go.fzj.de/B2SHARE>

DataPub

DATAPUB

- URL: <https://datapub.fz-juelich.de/>
- Solution for files which do not fit in B2SHARE
- Audience: researchers from Forschungszentrum Jülich
- Listed as repository on re3data
- Stores only data but no metadata and does not assign identifiers
- Accessible via https
- Datarepository can be managed by single users or groups
- Usage must be requested

Jülich DATA

JÜLICH DATA

- URL: <https://data.fz-juelich.de/>
- Based on Dataverse
- Service provided by ZB (central library) as central solution
- Connected via OpenID Connect to Helmholtz AAI
- Provides DOIs
- Storage for metadata and small files (up to 2GB)
- Link to other storages with URIs
- Storage tiering if institutes provide their own S3 storage
- Individual metadata schemas are possible