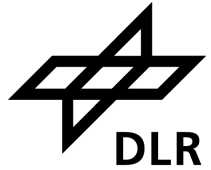


TERRABYTE ONBOARDING WORKSHOP

All you need to know to start with terrabyte



Agenda



Presentation

- What is terrabyte?
 - Collaboration DLR/LRZ
 - Hardware
- terrabyte Offerings
 - Data availability + access
 - Storage + Data transfer
 - HPC+
 - Compute Portal
 - Support
 - How to get an account

Hands-on

- Portal
 - Available apps
 - QGIS
 - Jupyter Notebooks
 - STAC Browser
 - Dask
- HPC
 - Login on the login node
 - Resource availability
 - SLURM-jobs
 - Containers

What is terrabyte?

Collaboration between DRL and LRZ

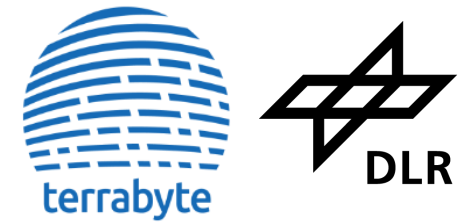


- On premise infrastructure with and at Leibniz Supercomputing Center (LRZ)
- Collaboration
 - Financing of the HPDA platform from DLR funds
 - Hosting and system operation by LRZ
 - DLR focuses on EO research, services and user support
 - LRZ focuses on IT research and operation
- Milestones
 - June 2023: Official launch of terrabyte
 - November 2024: Successful evaluation



What is terrabyte?

Hardware



Online-Storage (LRZ)

- 50 PB net
- 6 PB user space
- GPFS Ffilesystem with 360 Gbit/s Infiniband

Compute Cluster (LRZ)

- 272 CPU-Nodes with ~44.000 Intel Xeon Platinum CPU vCores und 273 TB RAM
- 48 Nvidia A 100 GPU-Nodes with 188 GPUs and 47 TB RAM + 14 TB GPU RAM



Network connection between DLR-archive (60 PB) and LRZ

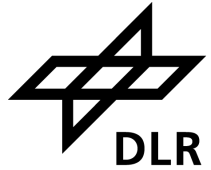
- 10 Gbit/s, upgrade to 100 Gbit/s possible

Internet connection between LRZ and CDSE, USGS, NASA

- max. 100 Gbit/s, depending on Server and Traffic)

terrabyte Offerings

An overview



Jira Service Desk



terrabyte
Support



terrabyte
Compute Portal



terrabyte
Processing services



Authentication
Register
IDM-Portal
2FA-Portal

terrabyte
Identity Mng.



terrabyte
HPC

DSS Web
Globus Online

terrabyte
Storage services

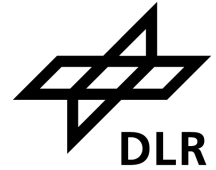


terrabyte
Data services



terrabyte Offerings

Data availability



Dataset	Current spatial coverage	Current temporal coverage	Update schema	Objective
Sentinel-1 SLC	On Demand	On Demand	On Demand	Autom. loading from D-SDA
Sentinel-1 GRD	Global	2014 - today	hourly	Global
Sentinel-1 NRB	Selected AOIs	2015 – 2023		Global
Sentinel-2 Col. 0 L1C	Europe	2018 - 2022		Available until Collection 1 is available
Sentinel-2 Col. 1 L1C*	Germany + more	2017 – 2024	On Demand	Autom. loading from D-SDA
Sentinel-2 Col. 1 L2A*	Germany + more	2015 - today	hourly	Global
Sentinel-3 OLCI L1 EFR	Global		hourly	Global
Sentinel-5p AUX, L1, L2	Global		hourly	Global
Landsat Col. 2 Level-2	Europe + more	1982 - today	hourly	Global
MODIS / VIIRS	Global			Global
Copernicus DEM	Global		yearly	Global

Ready

In progress

Planned

* Depends on available data

terrabyte Offerings

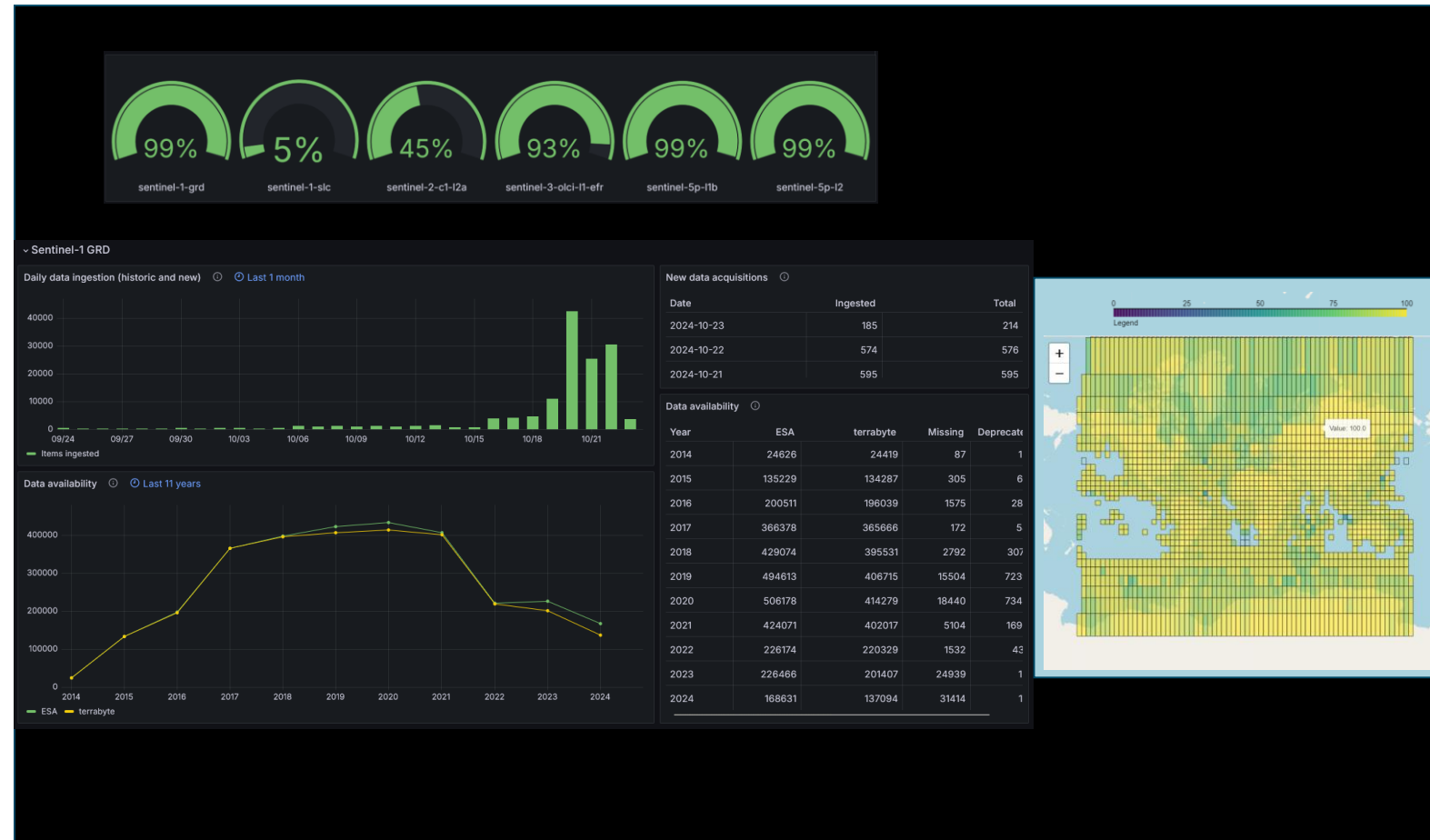
Data availability



<https://docs.terrabyte.lrz.de/datasets/terrabyte-datasets/roadmap/>

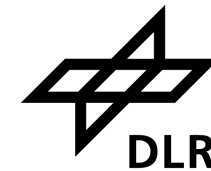
<https://docs.terrabyte.lrz.de/datasets/terrabyte-datasets/inventory/dashboards/>

Status	Dataset	DSS	STAC	Continuously updated	Due date
COMPLETED	Copernicus DEM GLO-30	Completed	Available	Static	Completed
COMPLETED	Copernicus DEM GLO-90	Completed	Available	Static	Completed
PLANNED	DESES Level-2A	Planned	Planned	Planned	
PLANNED	ERAS Land	Planned	Planned	Planned	Q4 2024 (2024-2000), Q1 2025 (1974-1999)
COMPLETED	GEDI (Germany)	Completed	Not Planned	Static	Completed
PLANNED	Landsat 1-9 Collection 2 Level-1	Planned	Planned	Planned	
DEVELOPING	Landsat 4-5 TM Collection 2 Level-2	Downloading	Available	Static	Q2 2025 (Europe available)
DEVELOPING	Landsat 7 ETM+ Collection 2 Level-2	Downloading	Available	Static	Q1 2025 (Europe available)
DEVELOPING	Landsat 8-9 OLI/TIRS Collection 2 Level-2	Downloading	Available	Updating (hourly)	Q4 2024 (Europe available)
	Meteosat - MSG.SEVIRI.L15_DISC	Planned	Not Planned	Planned	
	Meteosat - MSG.SEVIRI.L15_HRV	Planned	Not Planned	Planned	
	Meteosat - MSG.SEVIRI.L15_VISIR	Planned	Not Planned	Planned	
PLANNED	Metop Gome-2 Level-1	Planned	Not Planned		



terrabyte Offerings

Data access via STAC



<https://stac.terrabyte.lrz.de/browser>

terrabyte STAC API

API Source Share Language: English

Browse Search

Description

Curated data catalog of the terrabyte platform available at <https://stac.terrabyte.lrz.de/public/api>

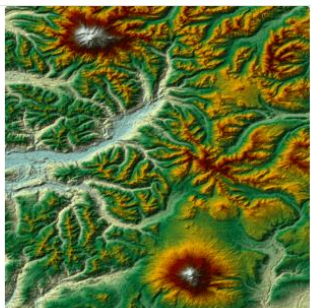
Additional Resources

- OpenAPI service description
- OpenAPI service documentation

Catalogs 20

Tiles List Ascending Descending

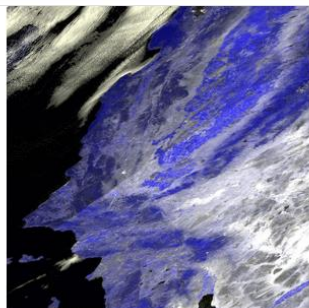
Filter catalogs by title



Copernicus DEM GLO-30

GeoTiff The Copernicus DEM is a digital surface model (DSM), which represents the surface of the Earth including buildings, infrastructure, an...

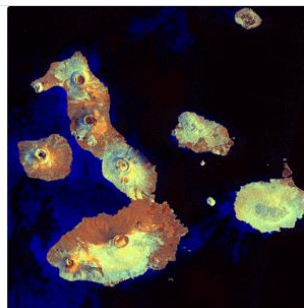
4/22/2021, 12:00:00 AM UTC



MOD09GQ.061 & MYD09GQ.061: MODIS Surface Reflectance Daily (250m)

Hdf4 HDF The MOD09GQ/MYD09GQ Version 6.1 product provides an estimate of the surface spectral reflectance of Terra/Aqua Moderate...

2/24/2000, 12:00:00 AM UTC - 6/10/2023, 12:00:00 AM UTC



Sentinel-1 SLC Level-1

This is the original Sentinel-1 SLC Level-1 dataset provided from ESA.

10/3/2014, 3:08:21 PM UTC - 6/14/2023, 3:45:08 PM UTC

```
import json
import os
from pystac_client import Client

lon1=29.327168
lat1=-11.745695
lon2=40.445137
lat2=0.985788

outdir="/dss/dsshomel/lxc0F/di46riq/Scripts/S1_ARD_SNAP/" # dir where to put scene list
indir="/dss/dsstbys01/pn56su/pn56su-dss-0008/Sentinel-1/GRD/" #main directory of input data

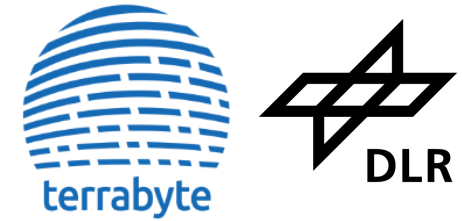
# find COP-DEM tiles that cover scene
# STAC API root URL
URL = 'https://stac.terrabyte.lrz.de/public/api'
# custom headers
headers = []
cat = Client.open(URL, headers=headers)

# limit sets the # of items per page so we can see multiple pages getting fetched
search = cat.search(
    collections = "sentinel-1-grd",
    bbox=[lon1, lat1, lon2, lat2],
    datetime="2020-01-01/2020-03-31",
    query={
        "sar:instrument_mode": {
            "eq": "IW"}
    }
)

# PySTAC ItemCollection
items = search.item_collection()
```



terabyte Offerings Storage



Basic storage

100 GB home directory

1 PB scratch space

Institute storage

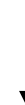
A storage container on the DSS for you or your DLR-internal project (some TB)



Contact your **terabyte institute representative**

Project storage

If your institute doesn't have a terrabyte representative or in case of a non-DLR-project

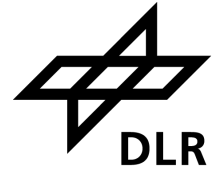


Apply for a terrabyte project

<https://docs.terabyte.lrz.de/services/project-management/apply-project/>

terabyte Offerings

Data transfer

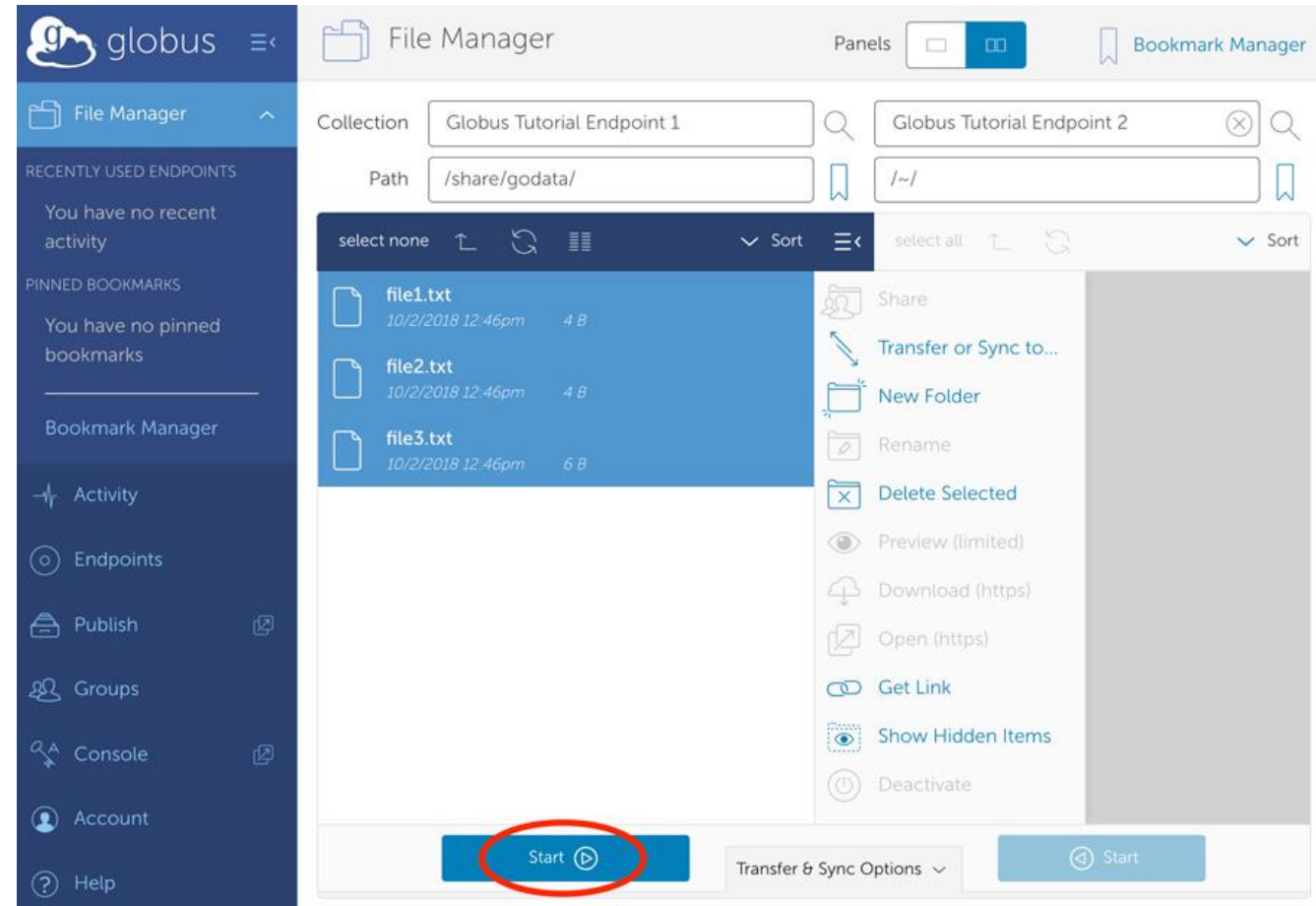


- Fast parallel data transfer between terrabyte and external systems
- Data can be shared with non-terabyte users
- Access via web GUI or CLI (automation, scripting)

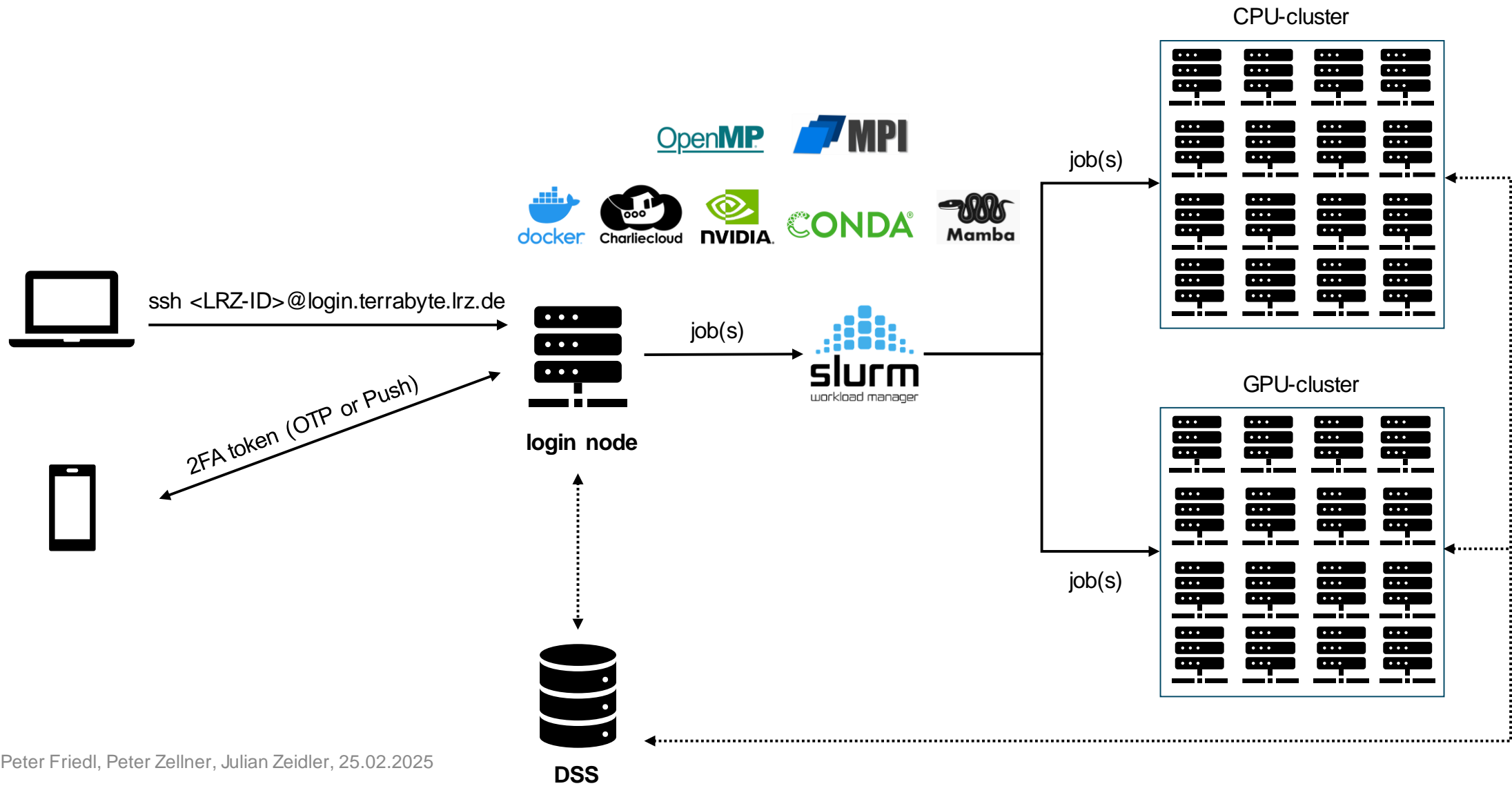
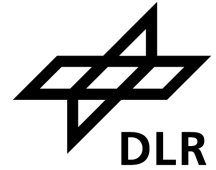
```
ubuntu@terabyte-vm:~$ globus endpoint search 'DLR'
```

ID	Owner	Display Name
d249dc36-8eca-11ee-be03-c52a29481bea	di46riq@lrz.de	DLR
e82dc6b4-81fb-4d96-a115-3f8d9943817c	barnes@contractor.usgs.gov	DLR IC Collection

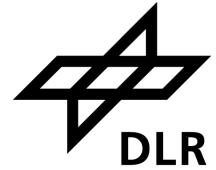
```
ubuntu@terabyte-vm:~$ globus transfer $dss_ep:/dss/dssfs02/pn56su/pn56su-dss-0007/ $dlr_ep
```



terrabyte Offerings HPC +



Terrabyte Offerings Portal



<https://portal.terrabyte.lrz.de>

{ terrabyte::Portal } Apps ▾ Files ▾ Jobs ▾ Clusters ▾ Data ▾ Desktop Apps ▾ Development ▾ My Interactive Sessions

About us Privacy policy Help ▾ Logged in as di46riq Log Out

integrated, single access point for all of your HPC resources.

Pinned Apps A featured subset of all available apps

Data

- Data Transfer
- STAC Browser

Desktop Apps

- QGIS
- Remote Desktop
- Remote Desktop Compute

Development

- Code Server for VS Code
- Jupyter Notebook
- RStudio Server

Files

- Home Directory System Installed App

Message of the Day

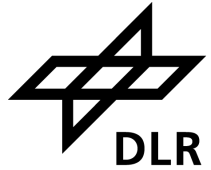
03.06.2024 :: Open OnDemand Update and Cluster mode
The Portal has been upgraded to the latest Open OnDemand version, 3.1.4. The cluster mode functionality is also enabled to improve site availability.

11.04.2024 :: Remote Desktop on Compute Node
The Remote Desktop Compute provides users with a remote desktop session on a compute node, complete with a fully supported module system. This enables built-in GUI applications such as QGIS, ParaView, TecPlot 360, ANSYS/Fluent ... to be launched conveniently.

10.04.2024 :: UI Update
The Code Server for VSCode and RStudio Server have been updated with an advanced UI, offering users greater flexibility in customizing computational resources for interactive sessions.

Terrabyte Offerings

Support: Institute Representatives

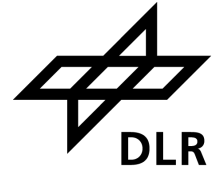


Institute	HPDA terrabyte representatives	
DFD	Julian Zeidler julian.zeidler@dlr.de +49 8153 28 1229	Dennis Kaiser dennis.kaiser@dlr.de +49 8153 28 2239
IHR	Philipp Posovszky philipp.posovszky@dlr.de +49 8153 28 2393	Marc Jäger marc.jaeger@dlr.de +49 8153 28 2802
IMF	Conrad Albrecht conrad.albrecht@dlr.de +49 8153 28 1861	Vytautas Jancauskas vytautas.jancauskas@dlr.de +49 8153 28 3730
IPA	PA-IT pa-it@dlr.de +49 8153 28 3528	Björn Brötz Bjoern.Broetz@dlr.de +49 8153 28 3554
RM	Dominik Winkelbauer Dominik.Winkelbauer@dlr.de +49 8153 28 3805	
SF FF	Rafal Broda rafal.broda@dlr.de +49 2203 601 1597	

Terrabyte Offerings

Support: Service Desk

Jira Service Management



<https://servicedesk.terrabyte.lrz.de>

servicedesk@terrabyte.lrz.de

Welcome to the terrabyte Service Desk

Search for help resources

Welcome to the terrabyte Service Desk
Please choose a category to open a ticket.

Contact us about

- General**
General questions regarding terrabyte
- HPC Slurm**
Usage of HPC, Slurm or other HPC-software, HPC infrastructure incident
- terrabyte Portal**
Usage of the terrabyte portal and its applications, terrabyte portal incident
- STAC API**
Usage of the STAC API, STAC API incident
- Curated data**
General questions regarding curated data, Curated data incident
- Feedback**

Jira Your work ▾ Projects ▾ Filters ▾ Dashboards ▾ Teams ▾ Plans ▾ Assets ▾ Apps ▾ Create

Search

terrabyte_SM Service project

Queues

Service requests

Incidents

OPERATIONS

- Services
- Alerts
- On-call

KNOWLEDGE

- Knowledge base
- Reports

CHANNELS & PEOPLE

- Channels
- Customers
- Invite team

SHORTCUTS

- Add shortcut
- Project settings
- Customer notification logs
- Give feedback

Projects / terrabyte_SM / All tickets

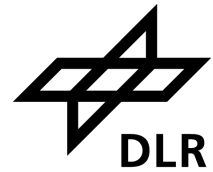
Resolved tickets

Search for issues

25 issues

T	Key	Summary	Reporter	Assignee	Status	Created
<input type="checkbox"/>	TS-386	External LRZ-ServiceDesk Incident207662: [#HPDA-DLR]: SSH Connection Refused, fail2ban	Jonas Jäger	Peter.Friedl	RESOLVED	11/Dec/23
<input type="checkbox"/>	TS-385	Connection refused for ssh-connections from workstation	Jonas Jäger	Peter.Friedl	RESOLVED	11/Dec/23
<input type="checkbox"/>	TS-384	login, Connection timed out	Timothy.Kodikara@dlr.de	Peter.Friedl	RESOLVED	09/Dec/23
<input type="checkbox"/>	TS-383	Maintenance of Terrabyte Cluser Over ?	Sandeep.Jangir@dlr.de	John Truckenbrodt	RESOLVED	06/Dec/23
<input type="checkbox"/>	TS-382	ssh not working	Thorsten Seehaus	Julian Zeidler	RESOLVED	05/Dec/23
<input type="checkbox"/>	TS-379	Adding a computer vision dataset as data-request	Wout Boerdijk	Unassigned	RESOLVED	30/Nov/23
<input type="checkbox"/>	TS-377	Austritt Mitarbeiter	Sven.Pruefer@dlr.de	Peter.Friedl	RESOLVED	29/Nov/23
<input type="checkbox"/>	TS-376	Fwd: RE: Meeting terrabyte account	Timothy.Kodikara@dlr.de	Jonas Eberle	RESOLVED	28/Nov/23
<input type="checkbox"/>	TS-375	Globus File Transfer: Setup issues	Jonas Jäger	Peter.Friedl	RESOLVED	28/Nov/23
<input type="checkbox"/>	TS-373	Berechtigung für HPDA	Thorsten Seehaus	Peter.Friedl	RESOLVED	24/Nov/23
<input type="checkbox"/>	TS-371	Error Report	Sheikh Mastura Mastura	Unassigned	RESOLVED	22/Nov/23
<input type="checkbox"/>	TS-369	Available ports	Sheikh Mastura Mastura	Julian Zeidler	RESOLVED	21/Nov/23
<input type="checkbox"/>	TS-365	Re: AW: AW: Request for a storage container external sharing	Timothy.Kodikara@dlr.de	Jonas Eberle	RESOLVED	17/Nov/23
<input type="checkbox"/>	TS-360	Video of Workshop on 13.11.2023	maximilian.ulmer@dlr.de	Peter.Friedl	RESOLVED	16/Nov/23
<input type="checkbox"/>	TS-359	Terrabyte 2FA	Thorsten Andresen	Peter.Friedl	RESOLVED	16/Nov/23
<input type="checkbox"/>	TS-358	Neue Python Version auf LRZ Terrabyte Cluster	Jonas Jäger	John Truckenbrodt	RESOLVED	14/Nov/23
<input type="checkbox"/>	TS-357	terrabyte: some questions	Seyed Majid Azimi	Peter.Friedl	RESOLVED	13/Nov/23
<input type="checkbox"/>	TS-356	1 terrabyte moderator request(s) waiting	terrabyte-bounces@lists.lrz.de	Peter.Friedl	RESOLVED	13/Nov/23
<input type="checkbox"/>	TS-355	1 terrabyte moderator request(s) waiting	terrabyte-bounces@lists.lrz.de	Peter.Friedl	RESOLVED	13/Nov/23

Terrabyte Offerings Support: Forum



<https://support.terrabyte.lrz.de>

all categories ▾ all tags ▾ **Categories** Latest New (2) Unread (1) Top

Category Topics Latest

- Announcements** 4 / month Maintenance completed 4d
We will post updates and downtimes (e.g., for maintenance) in this category.
User Settings/Preferences in forum cannot be saved 8d
Maintenance interruptions from 04.12.2023 - 07.12.2023 8d
- General** 1 / month Welcome to our community! Jul 8
Create topics here that don't fit into any other existing category.
Disable quick tips 8d
Errors starting jobs 8d
- terrabyte Portal** 2 / month Cannot start qgis through the portal 2h
Discussions and questions about the terrabyte portal available unter <https://portal.terrabyte.lrz.de>
Failed to submit Jupyter Notebook session via terrabyte Portal 8d
Earth System Modelling Framework on Terrabyte 15d
- HPC Slurm** 4 Failed linked ssh connection to jupyter lab node Sep 21
Discussions and questions about the usage of the terrabyte HPC Slurm environment.
CommandNotFoundError: Your shell has not been properly confi... Sep 14
Extending the home directory Sep 5
- STAC API** 4 Filter for Sentinel-2 tiles does not work in STAC browser 21h
Please ask questions about the use of the STAC API available at <https://stac.terrabyte.lrz.de/public/api>.
Creating my own STAC collections 4d
STAC sentinel 2 band names possible error 28d
- Curated data** 1 / month Roadmap for Sentinel-1 and 2 data over Antarctica 20h
Discussions and questions about the curated data provided on terrabyte.
S5P RPRO Data with 2 different versions Aug 31
For new data requests, please see the other category!
Difference between the Copernicus-DEM on terrabyte vs. Google E... Aug 11
- Data requests** 3 / month VIIRS/JPSS1 Snow Cover Daily L3 Global 375m SIN Grid, Version 2 11d
If you need data, which should be provided as part of the curated data catalog, please open a topic under this category. Please always search for existing topics!
VIIRS/NPP Snow Cover Daily L3 Global 375m SIN Grid, Version 1 11d
SA-1B Computer Vision Dataset 19d
- Feedback** 2 / month Visualization of cluster utilization 6d
General discussion about the terrabyte platform and its services, how it works, and how we can improve it.
Accessibility issues since 2FA introduction Oct 18

Creating my own STAC collections

STAC API di76qjr



This is the first time Sandro Groth has posted — let's welcome them to our community!



Sandro Groth di93wec

Aug 21

Hi all,

I'm wondering if someone could help me out with this:

We are currently trying to setup a surface water segmentation algorithm based on Sentinel-1/2 data that should run on terrabyte. The data is stored as COGs (for each scene) but also aggregated to tiles to create reference water products. In order to publish the data to the outside world in the future and to keep track on what was already processed, we want to organize the data in a STAC collection.

Currently, this happens on our DFD-GZS internal GeoFarm STAC API instance, but when processing is moved to terrabyte, this API is not accessible anymore, therefore my question:

Is it possible to create my own STAC collection at stac.terrabyte.lrz.de and populate it with items or do you know a better way to deal with the situation described above?

Additionally, would it be possible to sync items that are created within terrabyte with collections hosted elsewhere?

Assigned topic to di76qjr

1 Reply ▾

♡ 🔗 ⋮ ↩ Reply

created	last reply	7	53	4	1	1		▼
Aug 21	4d	replies	views	users	like	link		



Jonas Eberle di76qjr

Aug 22

Hi @di93wec,

thanks for your questions.



Sandro Groth:

↕ ↑

Is it possible to create my own STAC collection at stac.terrabyte.lrz.de and populate it with items or do you know a better way to deal with the situation described above?

Terrabyte Offerings

Support: Documentation



<https://docs.terrabyte.lrz.de>

- About terrabyte
- Architecture >
- Roadmap
- Citation
- Contact >

Home > About terrabyte

About terrabyte

What is terrabyte?

terrabyte is a cooperation between [German Aerospace Center \(DLR\)](#) and [Leibniz Supercomputing Center \(LRZ\)](#) to provide first class information technology for researchers comprising dedicated storage for EO data and large-scale CPU and NVIDIA GPUs computing resources for global decadal data analyses. Objective of this cooperation is the development of the **operational platform** terrabyte as well as **joint research** and **development** in the area of EO data processing. LRZ hosts and operates the infrastructure as well as basic services (e.g. user authentication, HPC Slurm, web-based computing portal) whereas DLR is responsible for data management, EO-specific web services, and user support.

terrabyte is an innovative High-Performance Data Analytics (HPDA) platform operated by the DLR and the Leibniz Supercomputing Center (LRZ), an institution of the Bavarian Academy of Sciences and Humanities. The platform provides researchers with efficient access to Earth Observation data, a powerful processing environment, and practical tools for data analysis. It was set up from 2019 and put into operation in summer 2023. The high-performance computers and the data science storage (online data storage) are operated and managed by the LRZ, while the Earth observation data and the central software interfaces and analysis tools for researchers are developed and operated by DLR.

terrabyte Access

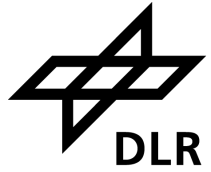
DLR users and scientists in Bavaria can access terrabyte using their LRZ account. A large part of the satellite data is pre-processed (known as Analysis Ready Data, ARD) and can be used in pre-installed applications. This enables terrabyte to perform a variety of geoscientific analyses that were previously only possible on commercial platforms due to the large

What is terrabyte?

- terrabyte Access
- terrabyte Offerings
 - Infrastructure
 - Services
 - Datasets
 - Support
- Contact

terrabyte Offerings

How to get an account



DLR employees

DLR cooperation chairs

**External DLR-partners
for EO projects with
DLR involvement**

**Non-EO DLR
Projects**



Get LRZ-ID via self-service <https://register.terrabyte.lrz.de>

terrabyte Offerings

How to get an account



<https://register.terrabyte.lrz.de/>

DFN-AAI-eduGAIN 

Organisation auswählen

Um auf den Dienst **register.terrabyte.lrz.de** zuzugreifen, wählen oder suchen Sie bitte die Organisation, der Sie angehören.

Auswahl für die laufende Webbrowser Sitzung speichern.

Auswahl permanent speichern und diesen Schritt von jetzt an überspringen.

[DFN-AAI](#) [Impressum](#) [Datenschutz](#)

Betrieben mit SWITCHwayf.

[Contact](#) | [Imprint](#) | [Privacy Policy](#)




 **Leibniz Supercomputing Centre**
of the Bavarian Academy of Sciences and Humanities

[DOCUMENTATION](#) [SERVICEDESK](#)

Die LRZ-Kennung  wurde erstellt.

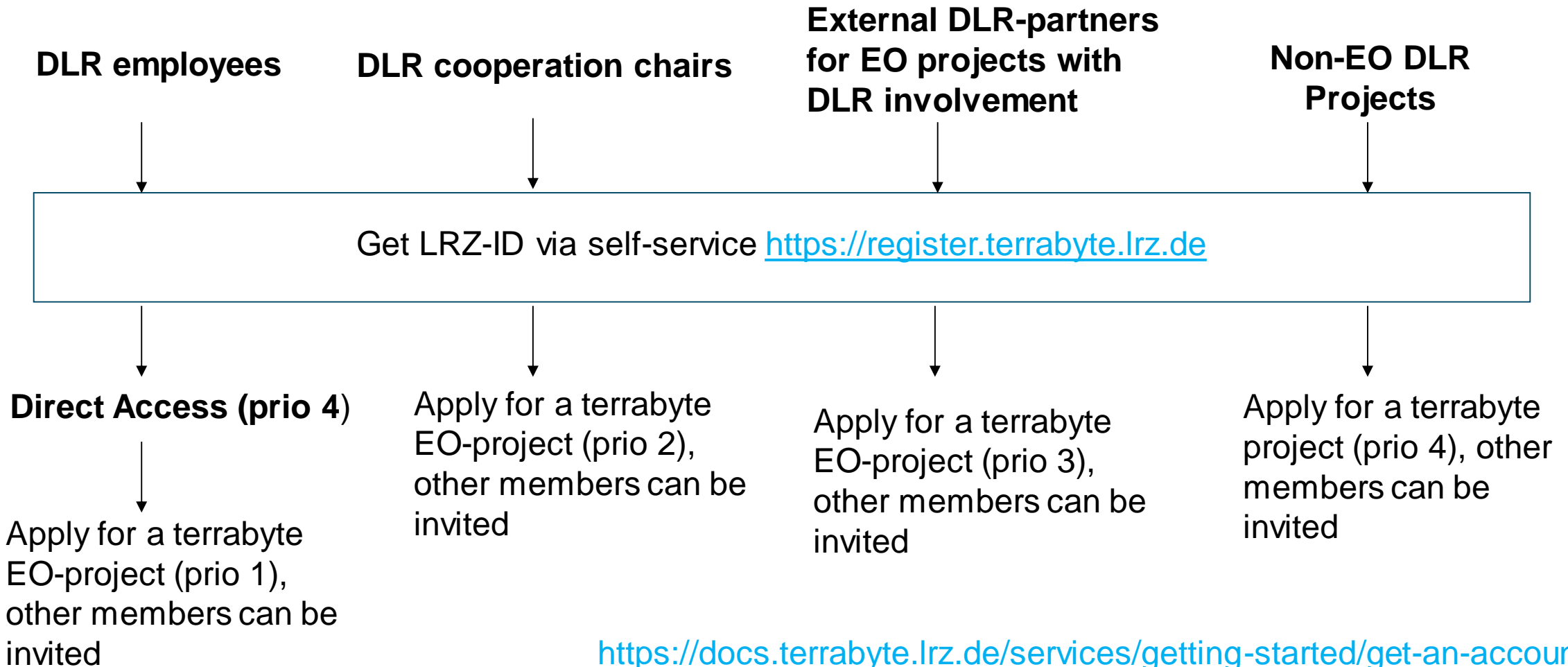
Sie müssen nun Ihr Konto im IDM-Portal aktivieren, indem Sie ein persönliches Kennwort vergeben.

Sollten Sie den Vorgang später fortsetzen wollen, notieren Sie sich bitte dieses vorübergehende Startpasswort: 

Sie werden es zusammen mit der Kennung brauchen, um sich am [IDM-Portal](#) anzumelden.

terabyte Offerings

How to get an account

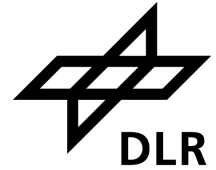


<https://docs.terabyte.lrz.de/services/getting-started/get-an-account/>

<https://docs.terabyte.lrz.de/services/project-management/apply-project/>

terrabyte Offerings

2FA registration



<https://simmfa.sim.lrz.de>

Get at least 2 tokens,
whereas 1 should be
independent from your
mobile phone!



Get the Authenticator App for
Android.



Get the Authenticator App for iOS.

The screenshot shows the 'Neuen Token ausrollen' (Register New Token) page in the privacyIDEA web interface. The browser address bar shows 'https://simmfa.sim.lrz.de/#/token/enroll/'. The page has a navigation menu on the left with 'Alle Token', 'Token ausrollen' (selected), 'Token zuweisen', and 'Help about Tokentypes'. The main content area is titled 'Neuen Token ausrollen' and features a dropdown menu for 'TCTP: Zeitbasiertes Einmalpasswort.' which is currently selected. Below this, there is a list of token types with descriptions: E-Mail, HOTP, IndexedSecret, PPR, PUSP, Fragenkatalog, SSH Public Key, TAN, TOTP (highlighted), WebAuthn, and Yubikey AES Mode. There are two buttons for downloading the Authenticator App for Android and iOS. Under 'Tokendaten', the checkbox 'OTP-Schlüssel auf dem Server erzeugen' is checked, and 'Ausrollen mit Zwei-Schritt und der privacyIDEA Authenticator App' is unchecked. The 'OTP-Länge' is set to 6, 'Zeitschritt' to 30 seconds, and 'Hash-Algorithmus' to sha1. There are input fields for 'Beschreibung' (containing 'Ein paar nette Worte...') and 'PIN' (with 'Type a password' and 'Passwort wiederholen' sub-fields). A 'Token ausrollen' button is at the bottom right.