

B2STAGE



What is B2STAGE?

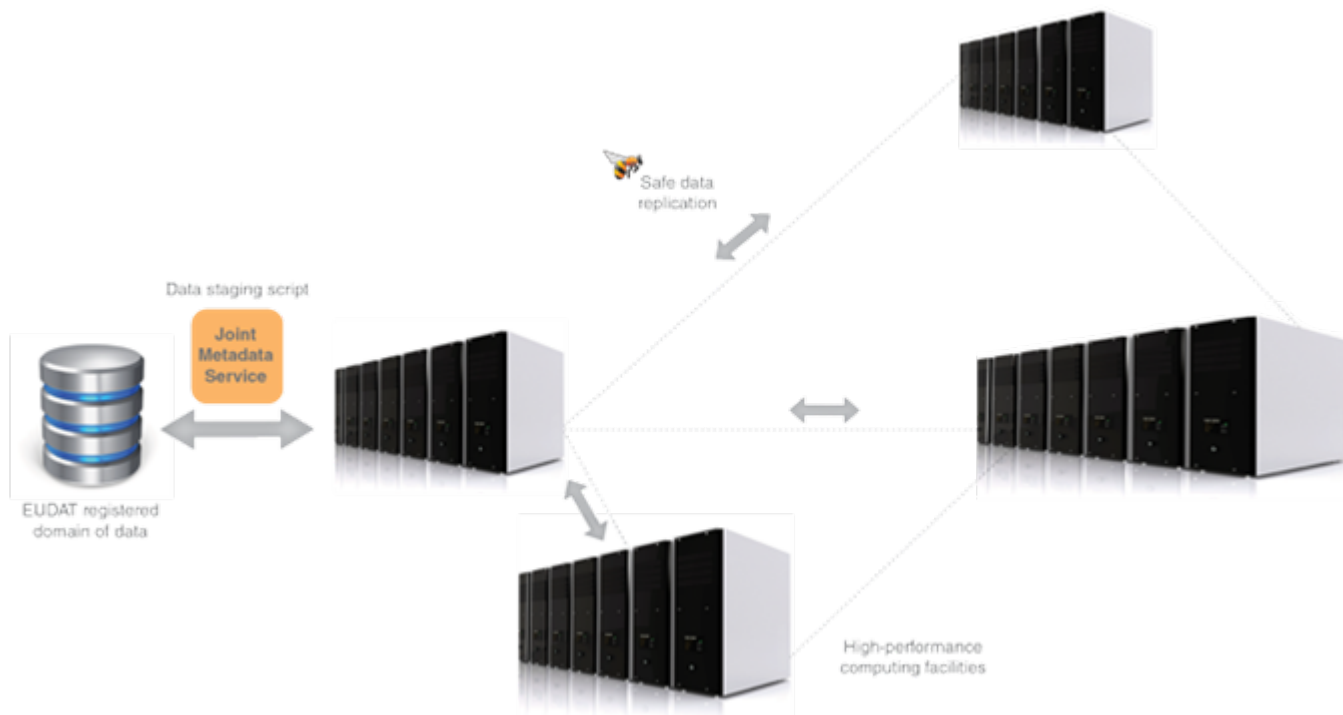
B2STAGE is a **reliable, efficient, light-weight and easy-to-use service to transfer research data sets** between EUDAT storage resources and high-performance computing (HPC) workspaces.

The service allows users to:

- Transfer large data collections from EUDAT storage facilities to external HPC facilities for processing
- In conjunction with B2SAFE, replicate community data sets, ingesting them onto EUDAT storage resources for long-term preservation
- Ingest computation results into the EUDAT infrastructure
- Access data through a RESTful HTTP interface (in progress)

B2STAGE features

- an extension of the [B2SAFE](#) and [B2FIND](#) services, which allow users to store, preserve and find data
- data-staging script facilitates staging, ingestion and retrieval of persistent identifier (PID) information of transferred data
- service available to all registered researchers and interested communities
- users negotiate access to remote HPC services in parallel
- collaboration with other infrastructures, such as the European Grid Infrastructure (EGI) and Partnership for Advanced Computing in Europe (PRACE)
- documentation, educational material and service helpdesk available to support users



Who benefits from B2STAGE?

B2STAGE is aimed at **researchers or communities**, who need:

- to access both large-scale data storage and high-performance computing systems;
- to ship data easily between the EUDAT storage resources and remote HPC facilities such as those provided by the PRACE distributed infrastructure;
- to simply ingest data onto EUDAT storage resources without setting up a full replica process.

How does it work?

The service aims at facilitating flexible data staging by offering different configuration options to support as many scenarios as possible. It combines a server component and various client interfaces.

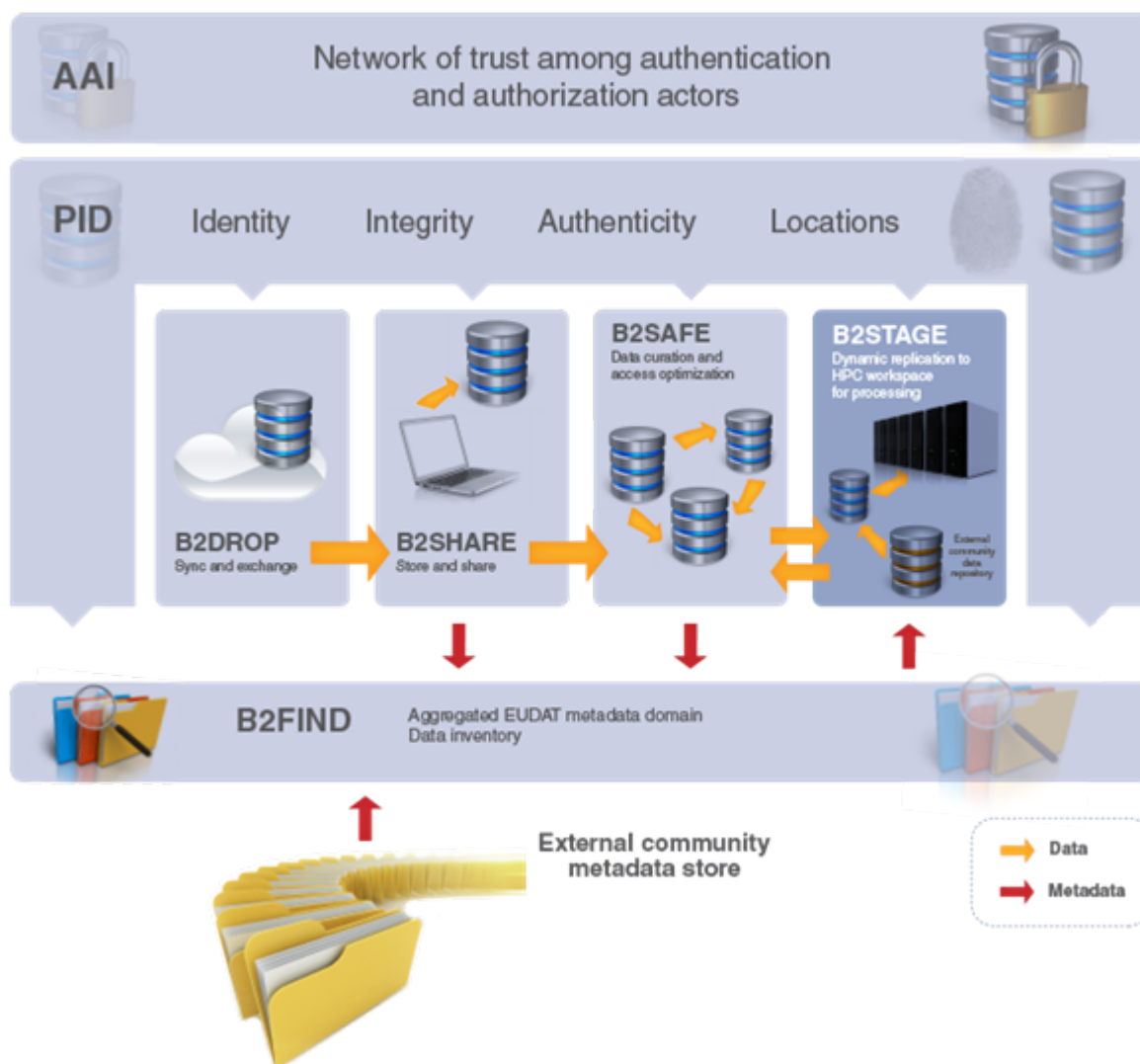
The staging functionality is an extension of the [iRODS](#) system with a GridFTP interface using a component being developed for this purpose - the GridFTP Data Storage Interface for iRODS (<https://github.com/EUDAT-B2STAGE/B2STAGE-GridFTP>). Therefore, the transfer of data is performed through a reliable, high-performance protocol - the GridFTP. On the client side, users have a rich selection of alternatives. In principle, any existing client supporting the GridFTP protocol can be used - [globus-url-copy](#), [GlobusOnLine](#), [UberFTP](#), and the [XSEDE-EUDAT File Manager](#) which also supports a wider range of transfer protocols (i.e. GridFTP, FTP, native iRODS, etc.). All proposed solutions offer the same core functionalities and can be used seamlessly. EUDAT also provides a script to facilitate the integration of B2STAGE within existing community solutions, such as web portal, workflow engine, etc. The script, named **Data Staging Script (DSS)**, as well as providing common data staging functionalities, also allows the retrieval of PIDs assigned to ingested data.

How can you use B2STAGE?

EUDAT offers B2STAGE to all registered researchers and interested communities enabling them to make use of the service for stage data out of EUDAT, and ingest any computational result back. Access to remote HPC facilities should be negotiated and arranged by individual users in parallel. To help researchers to use the B2STAGE service, EUDAT offers documentation, educational material (such as screencasts) and a service helpdesk. To find out more details about this service, please contact the EUDAT B2STAGE team at the address below.

For more information please email: eutat-datastaging@postit.csc.fi

B2STAGE in the B2 Service Suite



[Read more](#)