B2HANDLE for communities

About

Documentation of the EUDAT options for Persistent Identification (PID) with the use of B2HANDLE service, for site admins and community data managers.

Modified: 06 March 2017

Synopsis

B2HANDLE is a distributed service, designed to meet the fundamental research data management requirement of data persistency by maintaining opaque, globally unique persistent identifiers. This document discusses the basics of B2HANDLE and Persistent Identifiers in EUDAT.

What is B2HANDLE?

B2HANDLE is EUDAT's main Persistent Identifier (PID) service. B2HANDLE is a distributed service, designed to contribute to data persistency by maintaining opaque, globally unique persistent identifiers. PIDs are used to reliably identify and cite data objects throughout their lifecycle and they are thus a vital part of long-term data management. Moreover, data can be directly retrieved by PIDs and corresponding key metadata can be stored together with them in the so-called PID entry.

B2HANDLE is based on the Handle System, which is a reliable, redundant and scalable system built on an open architecture and designed to allow end-users to manage PIDs with ease. The B2HANDLE service encompasses management of identifier namespaces (Handle prefixes), establishment of policies and business workflows, operation of Handle servers and technical services, and a user-friendly Python library for general interaction with Handle servers and EUDAT-specific extensions.

In order to access a data object stored in EUDAT, an associated PID is needed. B2HANDLE enables EUDAT services and user communities to *assign PIDs* to different kinds of managed objects stored in the EUDAT CDI. The underlying technology is based on the Handle System, which is a reliable, redundant and scalable system built on an open architecture and mostly transparent to the end-user (users need not know the complexity of infrastructure details).

B2HANDLE service resolution is based on the Handle System which offers a very reliable resolution service. B2HANDLE forwards the user to the current location of the object. The *B2HANDLE Python library* is a client library for interaction with a Handle System server, using the native REST interface introduced in Handle System 8.

In the EUDAT ecosystem EUDAT services make use of B2HANDLE to enable data access, guarantee long lasting references to data and facilitate data publishing. B2SAFE and B2SHARE use the service to create and manage PIDs for their hosted data objects, whereas B2FIND and B2STAGE use the resolving mechanism of B2HANDLE to retrieve objects and to refer to them.

How can I benefit from B2HANDLE?

B2HANDLE provides the following end-user benefits:

- **Community applications** aiming to work with persistent identifiers or provide value-added services on top of them benefit from easier integration by using the Python library and the Handle server native interfaces.
- **Scientific users** can benefit from improved accessibility and visibility of data products they generate, including fine-grain citation and credit. Advanced options for discovering data across community boundaries or deriving new data products become available as widespread use of PIDs improves options for tracking the use of data products and tracing their provenance, independent from particular repositories.

Hosting such a service is an important undertaking. B2HANDLE can take the strain off the community, and cover the following aspects:

- **PID Service:** Services to generate PIDs for digital objects.
- **PID Resolution:** Services to guarantee resolution of the PIDs issued
- **PID Replication:** Replication of the databases of handles towards around-the-clock availability of the PID resolution
- **Global Handle Mirror Server:** Access to a mirror of the Global Handle Server in Europe
- **Service Monitoring:** Ensures service availability.

How does B2HANDLE work

- **PID service operations:** EUDAT maintains a federated B2HANDLE service for creating and delivering handles and provides mirror servers that follow mature IT service management procedures. All software components have been thoroughly tested before deployment. Updates are coordinated through documented change requests and all federation nodes are centrally monitored. The federated service is defined by policies and processes that ensure a high quality of service, while also providing an environment encouraging process and technology innovation.
- **PID record unification**: B2HANDLE manages procedures that unify the contents of Handle records. The process for this revolves around a common standard for Handle records content that all participating services agree upon. The evolution of applications may cause new elements to be recognised and included in the common standard.
- **B2HANDLE Python library**: The B2HANDLE library is developed in an open source software process. The source code is openly available under an Apache 2 software license via GitHub. The development process includes continuous integration via Jenkins and enables short release cycles.

How can I use B2HANDLE?

Assume the scenario where stakeholders from a scientific community see the need to build and maintain a trustworthy repository for research data, and, particularly, to acquire PIDs for the deposited data. With the use of PIDs, deposited data can be referenced independently from future changes in location or ownership of the repository.The repository is set up at the community institution. The B2HANDLE service can address the needs of using and hosting Persistent Identifiers for the repository.

Acquiring and hosting a prefix

The community stakeholders should first acquire a Handle prefix. A Handle prefix is acquired through EUDAT members. Prefixes are assigned to users by so-called MPAs (Multi-primary Prefix Administrators). To keep hosting of PIDs and hosting of data independent from each other, the stakeholders can approach EUDAT and ask for hosting of the prefix they just acquired. See "How to request a Production Prefix" below.

The community may specify a preferred partner for hosting the primary Handle service, or they may decide to

rely on the B2HANDLE team to make all necessary arrangements. In any case, the community does not need to get involved in the details of the set-up process as B2HANDLE takes care of all required steps. Once the process is complete, the community receives a short report with all details of their service partners within EUDAT and a primary support contact point. At this point, the prefix hosting service is operational.

How to request a Production Prefix

Prefixes are assigned to users by so-called MPAs (Multi-primary Prefix Administrators). The European Persistent Identifier Consortium consists of such MPAs of which some are also part of EUDAT and are eligible to hand out prefixes.

To **acquire a production prefix** use the EUDAT support webform to have EUDAT take care of this.

Managing PIDs

Once the prefix hosting service is operational, the repository host can use the Handle REST services to create and manage handles.

Additionally, EUDAT provides the B2HANDLE Python library - a client library for interaction with Handle servers. It offers the following features:

- Creating, reading, updating and deleting handles and handle values through the Handle System's native REST interface
- Searching over local Handles by target location (reverse look-up) or other record values
- Management of multiple location entries per record, using 10320/loc type entries
- Authentication via secret keys or client certificates
- Compatibility with Python 2.6 and Python 2.7

The B2HANDLE Python library can be installed by importing its sources from GitHub. Please refer to the B2HANDLE library homepage for more information on installation and configuration. For usage instructions, please refer to the technical documentation.

Whether through the Handle REST service ot the B2HANDLE Python library, the community benefits from a premium PID hosting service, providing stable references and added-value services without having to worry about the operational management of the PID service.

Support

Hands-on material is available from modules 7a-d on the EUDAT Training Repository.

Support for B2HANDLE is available via the EUDAT ticketing system through the webform.

If you have comments on this page, please submit them though the EUDAT ticketing system.

Document Data

DocVersion: 1.0

Authors:

Tobias Weigel, weigel@dkrz.de

Themis Zamani, themis@grnet.gr

Editors:

Christine Staiger, christine.staiger@surfsara.nl

Kostas Kavoussanakis, kavousan@epcc.ed.ac.uk

Read more