Configuration Guide | PUBLIC
2024-05-10

Quick Configuration Guide for the ILM Store

Version for Apache Hadoop

THE BEST RUN SAP

Content

1 Introduction

Apache Hadoop is an open-source software project that enables distributed processing of large data sets across clusters of commodity servers. It is designed to scale up from a single server to thousands of machines, including a very high degree of fault tolerance. Rather than relying on high-end hardware, the resiliency of these clusters comes from the software's ability to detect and handle failures at the application layer.

> **Note**
>
> This is a quick set-up guide. For a comprehensive documentation, refer to Installation & Config. Guide ILM Store (from SPS 13).

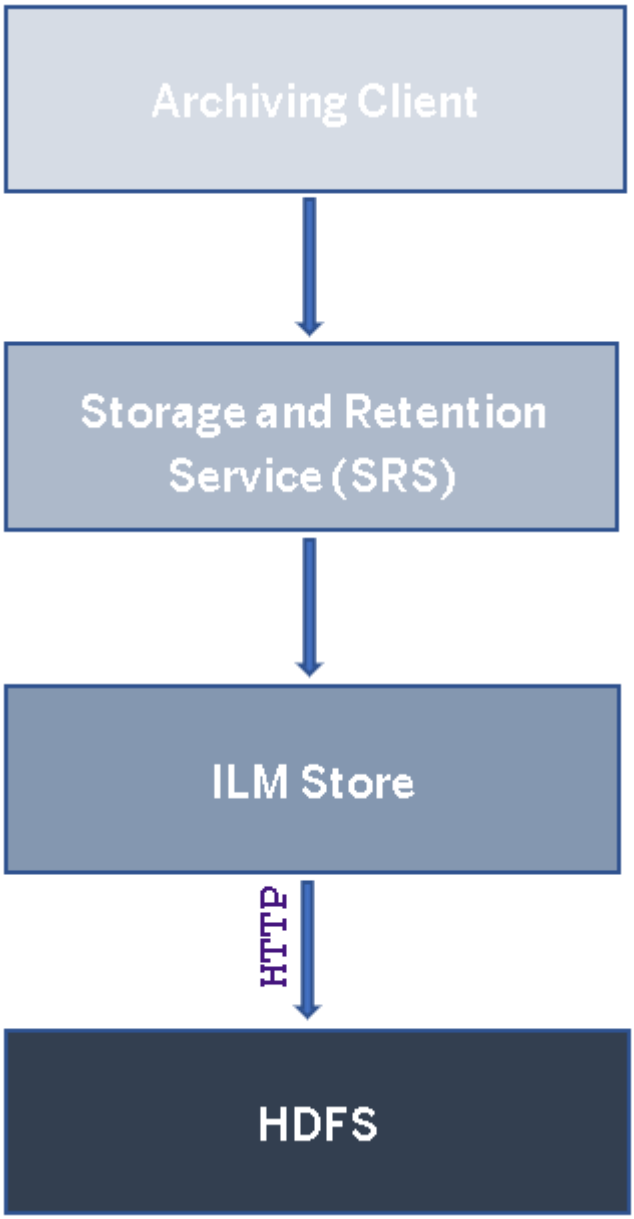


Figure 1: SAP ILM Store with Hadoop Architecture

Hadoop File System (HDFS) indicates the main difference between Hadoop and a database. While databases store data in the form of transparent tables, Hadoop operates with flat files in a directory structure. The HDFS stores files in specific directories. Therefore, the HDFS requires the file's name and directory path.

Table 1: Main URI / path components of the ILM Store and HDFS

ILM Store	Hadoop File System
Collection	Directory
Resource	File

Resources within the ILM Store are stored in specific collections that create resource URIs. These resource URIs are also used in the Hadoop File System.

Benefits:

1. Better traceablity of files in a HDFS
2. Flexible selection of resources when moving or cleaning files

You can use this document to set up the ILM Store with Hadoop.

> ⚠ Caution
>
> Once the ILM Store is set up and the operations have started, do not make any changes in the settings as this can lead to the loss of the information stored so far.

2 Prerequisites

To configure the ILM Store, ensure the following:

- You use an SAP NetWeaver system with the minimum requirements of software component *SAP_BASIS 740 Support Package 13* or *SAP_BASIS 750 Support Package 02*.
  SAP recommends using the latest version of the `SAP_BASIS` software component.
- You have activated the business function `ILM_STOR`.
- You have a productive Hadoop cluster that can be accessed from the system landscape.

> **Note**
>
> This documentation does not cover installation support for Hadoop clusters. Install and enable HttpFS or WebHDFS service on your Hadoop cluster.
>
> HttpFS and WebHDFS are HTTP services covered by the Apache Hadoop project. They provide HTTP REST API that supports all HDFS file system operations. The ILM Store's *Hadoop Connector* supports connection to both of these services. For more details about the services, refer to the official documentation of the Apache Hadoop project: Apache Hadoop .

2.1 Important SAP Notes

Make sure that you have implemented the relevant notes of component `BC-ILM-STO`.

Title	SAP Note	Comment
ILM Store - Collection of notes	2563024	This note gives a list of notes for component `BC-ILM-STO`. Refer to this note for detailed information.

3 Authorizations

Access to the ILM Store

1. Create a technical user with the authorization to access the ILM Store and to save files in the store.
2. Assign a role to the user. Use the following authorizations:

Authorization Object	Field	Value
SILMSTOR	ACTVT	*16* (Execute)
S_DATASET	FILENAME	*
	PROGRAM	CL_ILM_STOR_DATASET==========CP, RILM_STOR_PUT_WORKER
	ACTVT	*6* (Delete), *33* (Read), *34* (Write)
S_DEVELOP	OBJTYP	*TABL*
	ACTVT	*07*, *40*
S_CTS_ADMI	CTS_ADMFCT	*TABL*
S_CTS_SADM	CTS_ADMFCT	*TABL*

ILM Store Administration

To be able to install, configure, and test the store, you need a role with the authorization object SILMSTORAD with the following values of ACTVT:

1. *02* = Change
2. *07* = Activate, Generate
3. *39* = Check

4 Origin Customizing

The origin serves as the identifier of the data source. It is essential for the Store's configuration.

4.1 Administrative Customizing

1. Start transaction `ILM_STOR_ADM_CUST`, or go to *SAP NetWeaver Customizing*, and choose *Application Server* > *Basis Services* > *Information Lifecycle Management* > *ILM Store* > *Define Settings for Administrative Customizing* > .
2. Click *Create*.
3. Enter the *Client* and the *Logical File Name*.
4. Specify a name and a description for the new *Administrative Origin*.

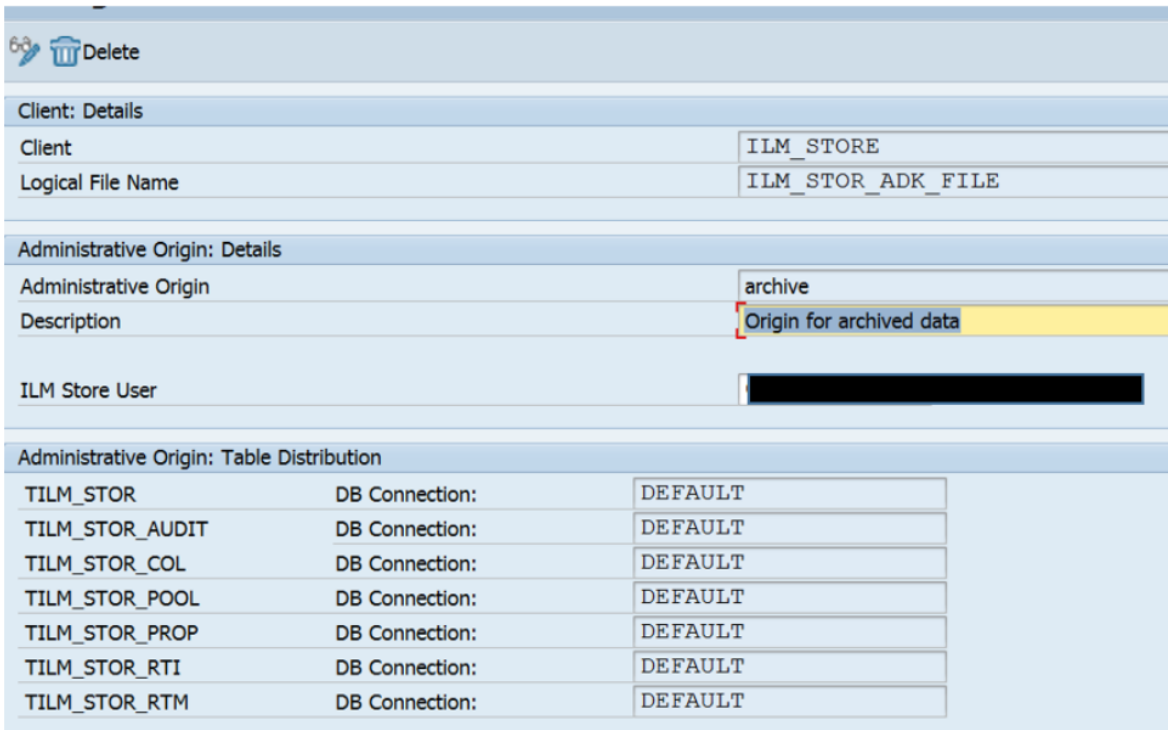


Figure 2: Customizing in Transaction ILM_STOR_ADM_CUST

5. Click *Add Operational Origin*.
6. In the pop-up *Create Operational Origin*, enter the *Operational Origin* name and a *Description*.
7. In the *DB Connection* field for `TILM_STOR_BLOB`, enter **`DEFAULT`**.

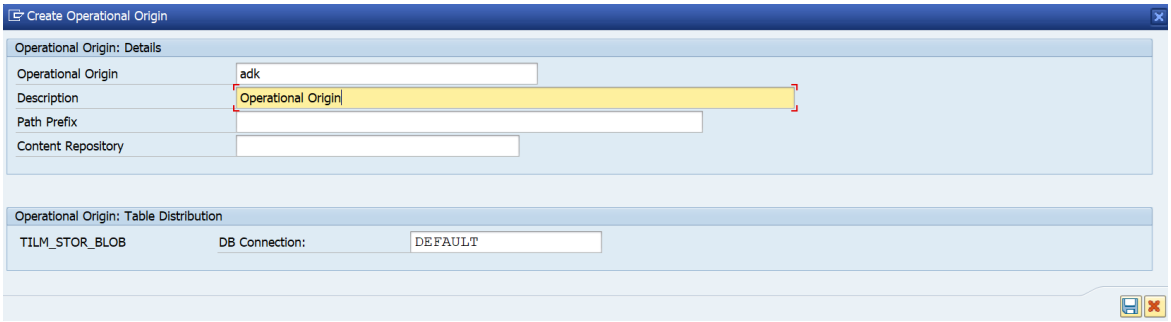


Figure 3: Customizing of the Origin DB Connection in Transaction ILM_STOR_ADM_CUST

4.2 Operational Customizing

To maintain additional properties, start transaction `ILM_STOR_OPR_CUST`, or go to the *SAP NetWeaver Customizing*, choose *Application Server* > *Basis Services* > *Information Lifecycle Management* > *ILM Store* > *Define Settings for Operational Customizing*.

To maintain additional properties, enter the client maintained in *Administrative Origin* and click *Execute*.

4.3 Configuring Storage Connections

Maintain the following entries in transaction `SM30` for table `TILM_STOR_CUS`.

Origin	Namespace	Property	Value
adk	DB	`DBCON.TILM_STOR_BLOB`	`HADOOP_CONNECTION_ID`
adk	DB	`CONSYS.HADOOP_CONNECTION_ID`	`HADOOP`

4.4 Routing Table Configuration

1. In transaction `SM30`, enter the table name `TILM_STOR_O_ROUT`.
2. Create the following entry:
   SAP System ID: **`<System ID>`**
   Client: **`<Client>`**
   Data Source: **`<Your operational origin>`**

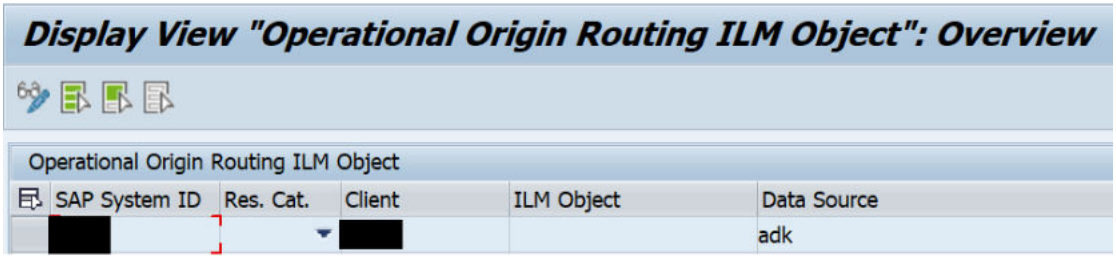


Figure 4: Routing Table Entries

4.5 Class Factory Customizing

1. In transaction `SM30`, enter the table name `TILMSTOR_CF`.
2. Create an entry with the following values.

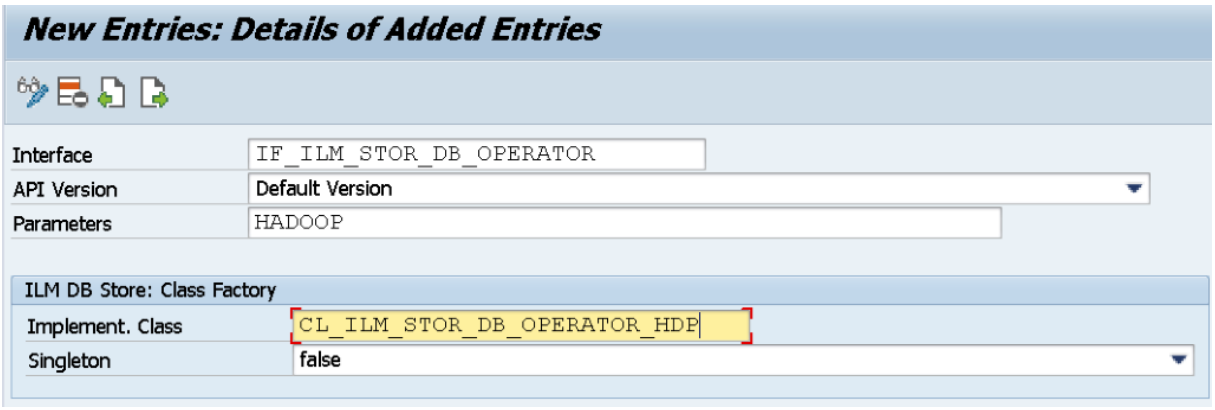


Figure 5: Class Factory Table Entries

5 Publishing the ILM Store

5.1 Create the ICF Node for the ILM Store

1. Start transaction `SICF` and create a new service under node `ILM`. Enter the service name in the *Service Name* field.

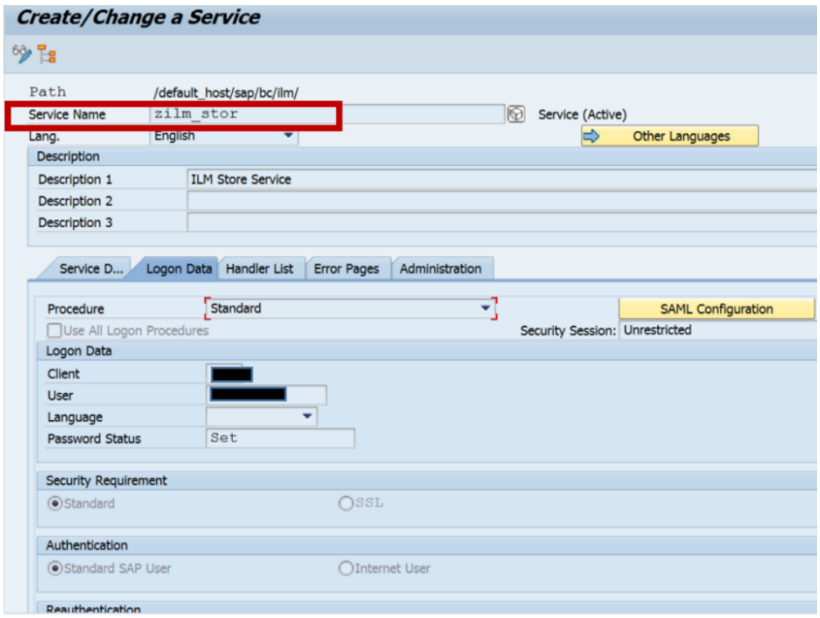


Figure 6: SICF Node Creation for the ILM Store

2. In the *Logon Data* tab, enter a user with the authorization to access the ILM Store. For more information, see Authorizations [page 8].
3. On the *Handler List* tab, enter `CL_ILM_STOR_WD_REQUEST_HANDLER` as *Handler*.
4. Activate the service.

5.2 Create the RFC Destination

Create two RFC connections as listed below.

5.2.1 Connection to HDFS

1. Start transaction SM59.
2. Create a new HTTP connection to the external server (type **G**).
3. Maintain the target host and service number (port) where the Hadoop HTTP Service is initialized and running.
4. Give the path prefix referring to the home directory that refers to the user to be invoked for the connection. It has the following syntax:
   /webhdfs/v1/<hdp_usr_home_path>
   - */webhdfs/v1* – This is the mandatory path prefix defined by the HttpFS and WebHDFS APIs.
   - *<hdp_usr_home_path>* – This is the home directory path of the target Hadoop user. Each Hadoop user has their own home directory. In UNIX type systems, for example, this is typically */user/*. The user is the owner of the directory and has higher access privileges.

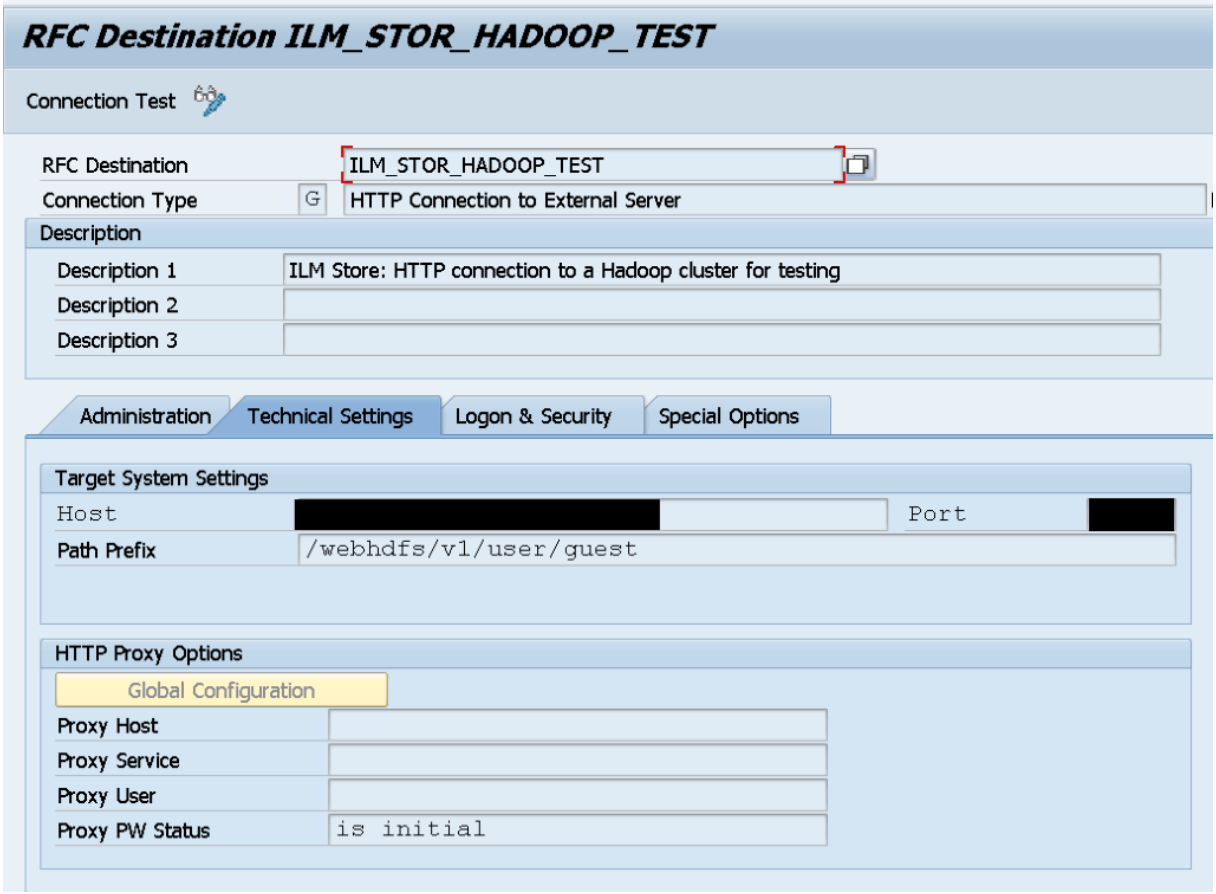


Figure 7: RFC Connection to HDFS in Transaction SM59

5.2.2 Connection to the ILM Store

1. Start transaction SM59.
2. Create a new HTTP connection to the external server (type **G**).

3. In the *Technical Settings* tab, maintain the below values:
    - Target host
    - Service number (port) corresponding to your system.
    - Path prefix: insert the service path you have defined using transaction `SICF` in Create the ICF Node for the ILM Store [page 12]. Example path: **`/sap/bc/ilm/zilm_stor`**.
4. Save your changes and perform a connection test.

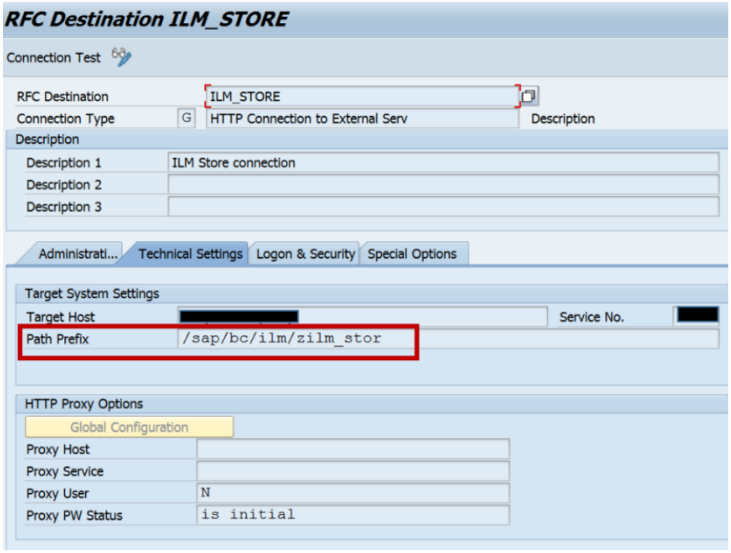


Figure 8: RFC Connection to ILM Store in Transaction SM59

> ⓘ Note
>
> In case you are running the SRS and the ILM Store on two different systems, set up the RFC destination in the SRS system using the same procedure. The RFC destination in the ILM Store system is for testing purposes.

6 Maintenance of Hadoop Preferences in ILM Store

To ensure successful operation and data transfer between the ILM Store and Hadoop, maintain the connection parameters, the authentication, and the access methods.

6.1 Hadoop Connection Customizing

To configure the Hadoop connection, do the following:

1. Start transaction `SM30`, and enter the table name `TILM_STOR_HDP_C`.
2. Maintain the *Connection ID* in alignment with the name in `TILM_STOR_CUS` in Configuring Storage Connections [page 10].
3. Create an entry with the following values.

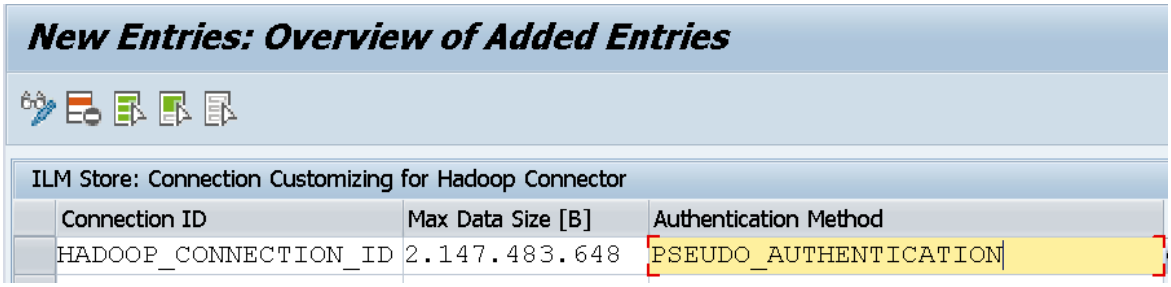


Figure 9: Hadoop Connection Table Entries

The Customizing table contains the following properties:

- *Max Data Size*: Maximum amount of data (in bytes) transferrable by an HTTP request. The value depends on the HTTP service setup in Hadoop.
- *Authentication Method*: Authentication method used for user access to Hadoop. The value depends on the HTTP service setup in Hadoop. You can choose between *Pseudo Authentication* and *No Authentication*.

6.2 Access Customizing

The *No Authentication* method requires no further maintenance. The *Pseudo Authentication* method requires the following additional configuration steps.

1. Start transaction `SM30`, and enter the table name `TILM_STOR_HDP_PU`.
2. Create an entry with the following values.

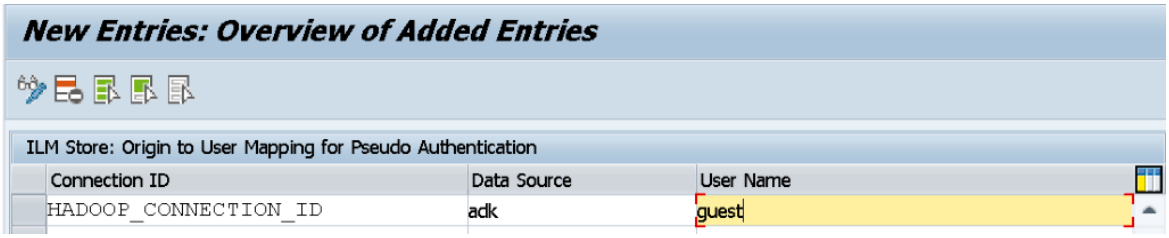


Figure 10: Origin to User Mapping Table Entries

This entry maps the origin (data source)-to-Hadoop user that is supposed to perform the connection and data transfer. The Hadoop user must exist in the Hadoop system.

6.3 Hadoop Origin-Dependent Connection Customizing

To configure the Hadoop origin-dependent connection, do the following:

1. Start transaction `SM30`, and enter the table name `TILM_STOR_HDP_CO`.
2. Maintain the *Connection ID* in alignment with the database connection name maintained in `TILM_STOR_CUS`.
3. In the *RFC Destination* field, maintain the RFC Connection to HDFS [page 13].

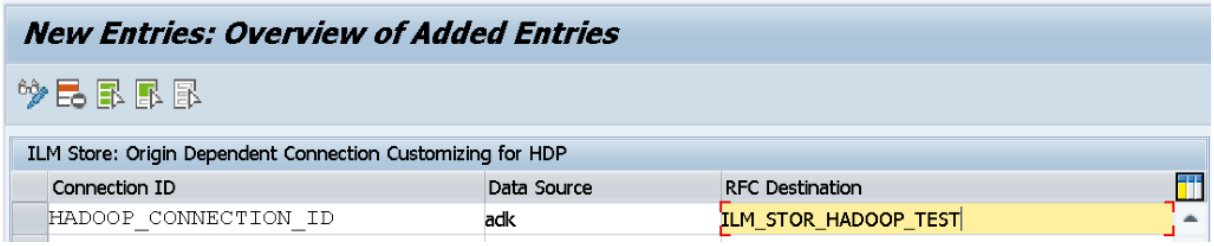


Figure 11: Hadoop Origin-Dependent Connection Table Entries

7 Storage and Retention Service

The *Storage and Retention Service* (SRS) is needed for the storing of ILM-enabled archive files in the ILM Store.

To use SRS for managing ILM stores in, you need to activate it in the application system.

7.1 Activation of SRS

To activate SRS, you can choose between the following options:

- Activate the SRS that runs locally on the application system.
- Activate the SRS that runs on a separate (remote) system. To facilitate this, establish an HTTP connection between the relevant systems.

For more information, refer to the *SAP Help Portal* under *SAP Information Lifecycle Management* > *Making SAP ILM Available* > *Providing Stores for SAP ILM* > *Configuring the Service for the Control of ILM Stores* > *Configuring Storage and Retention Service for ILM Stores*.

7.2 Creation of the ILM Store in the SRS Administration

To enter stores that are available for the SRS in transaction `ILMSTOREADM`, create a new entry with the following values:

1. *ILM Store*: **`<identifying name for the store>`**
2. *Description*: **`<descriptive text>`**
3. *HTTP Connection*: **`<previously created RFC destination>`**

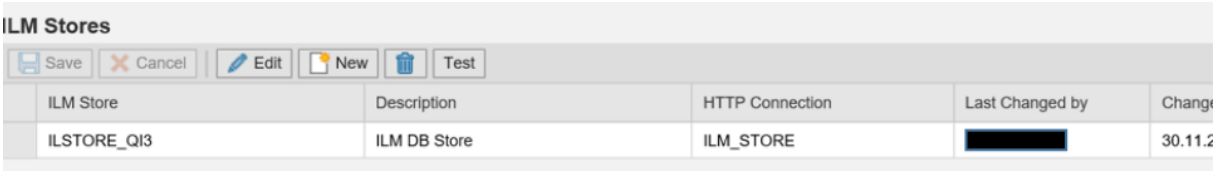


Figure 12: Creation of Store in ILMSTOREADM

This store can further be used in the ILM rule maintenance in transaction `IRMPOL`. For more information, refer to the *SAP Help Portal* under *SAP Information Lifecycle Management* > *Using ILM Retention Management in the Application System* > *Editing ILM Policies* > *Editing Retention Rules*.

8 Security

There are three main security tasks in the ILM Store:

1. Ensure that data is only accessible to its owners.
2. Ensure the confidentiality of data during HTTP communication.
3. Ensure the integrity of data.

8.1 Authentication

The Hadoop Connector offers two authentication methods:

1. `NO_AUTHENTICATION`
2. `PSEUDO_AUTHENTICATION`

Since Hadoop is an open-source software, you can implement other authentication protocols. To enable this for the SAP application as well, you can implement your own authentication methods. These methods must be registered in the table `TILM_STOR_HDP_A`, and a corresponding implementation of the authentication BAdI must be created in the enhancement spot `ILM_STOR_HDP_CONNECTIVITY`. The ILM Store's Hadoop Connector also supports the standard HTTP Basic authentication. User credentials are maintained within the corresponding RFC destination.

8.2 Apache KNOX

The Apache Knox Gateway is a `REST` API Gateway for interacting with Hadoop clusters. The Knox Gateway provides a single access point for all `REST` interactions with Hadoop clusters. It supports HTTP Basic authentication. You can maintain the HTTP Basic authentication within the RFC destination. Use the authentication method `NO_AUTHENTICATION` for this purpose.

8.3 HttpFS over HTTPS

HTTPS can be enabled on the level of RFC destination. To determine whether an SSL certificate shall be used, you can use transaction `SM59`. Before you do so, make sure you have a correct setup of the security certificate in transaction `STRUST`.

9 Testing the ILM Store

After completing the setup, test the ILM Store according to the following guidelines.

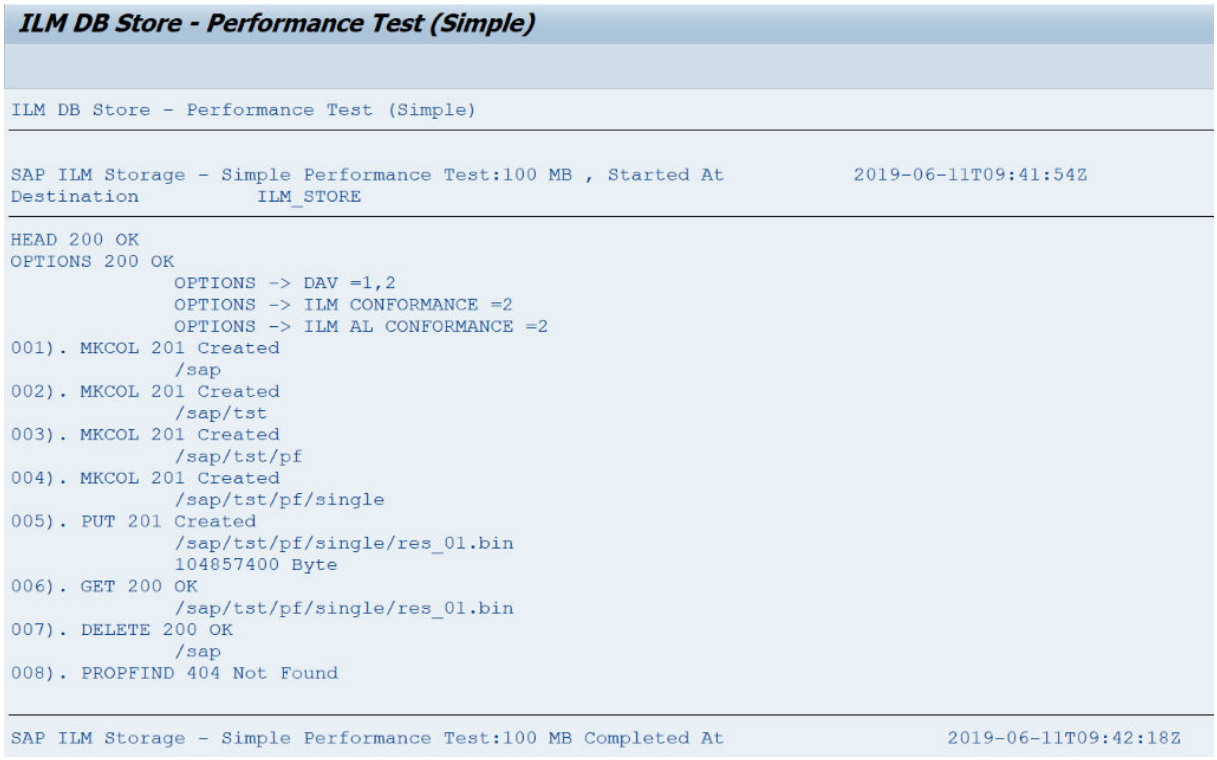
9.1 Origin for Test Purposes

To test the ILM Store configurations, you can use a test origin archeb.

1. Start transaction `SM30`, and enter table `TILM_STOR_CUS`.
2. Create new entries for the origin archeb by copying all the entries of your origin (for example *adk* and *archive*).
   It is mandatory to use the test reports to check the ILM Store functionality.

9.2 Test Reports

To test the ILM Store functionality, you can use the report `RILM_STOR_TEST_PF_SINGLE`. In the report selection, provide the RFC destination (created in Configuring Storage Connections [page 10]), and execute.

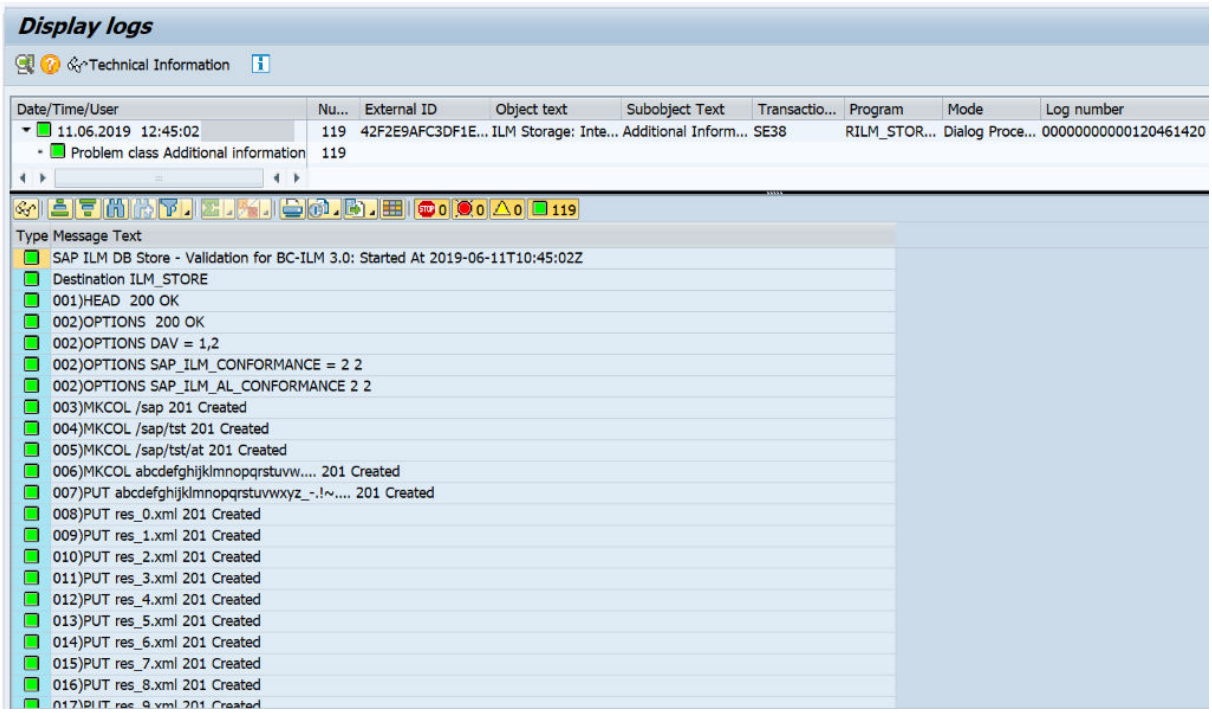


ILM DB Store - Performance Test (Simple)

```
ILM DB Store - Performance Test (Simple)

SAP ILM Storage - Simple Performance Test:100 MB , Started At          2019-06-11T09:41:54Z
Destination          ILM_STORE

HEAD 200 OK
OPTIONS 200 OK
              OPTIONS -> DAV =1,2
              OPTIONS -> ILM CONFORMANCE =2
              OPTIONS -> ILM AL CONFORMANCE =2
001). MKCOL 201 Created
              /sap
002). MKCOL 201 Created
              /sap/tst
003). MKCOL 201 Created
              /sap/tst/pf
004). MKCOL 201 Created
              /sap/tst/pf/single
005). PUT 201 Created
              /sap/tst/pf/single/res_01.bin
              104857400 Byte
006). GET 200 OK
              /sap/tst/pf/single/res_01.bin
007). DELETE 200 OK
              /sap
008). PROPFIND 404 Not Found

SAP ILM Storage - Simple Performance Test:100 MB Completed At                2019-06-11T09:42:18Z
```

Figure 13: Expected Output of the Report RILM_STOR_TEST_PF_SINGLE

To check various data constellations and operations as defined in the specification `BC-ILM` 3.1., you can use the report `RILM_STOR_TEST_AT`. In the report selection, provide the RFC destination (created in Connection to the ILM Store [page 13]), and execute.



Display logs

Technical Information

Date/Time/User	Nu...	External ID	Object text	Subobject Text	Transactio...	Program	Mode	Log number
11.06.2019 12:45:02	119	42F2E9AFC3DF1E...	ILM Storage: Inte...	Additional Inform...	SE38	RILM_STOR...	Dialog Proce...	00000000000120461420
Problem class Additional information	119							

Type	Message Text
	SAP ILM DB Store - Validation for BC-ILM 3.0: Started At 2019-06-11T10:45:02Z
	Destination ILM_STORE
	001)HEAD 200 OK
	002)OPTIONS 200 OK
	002)OPTIONS DAV = 1,2
	002)OPTIONS SAP_ILM_CONFORMANCE = 2 2
	002)OPTIONS SAP_ILM_AL_CONFORMANCE 2 2
	003)MKCOL /sap 201 Created
	004)MKCOL /sap/tst 201 Created
	005)MKCOL /sap/tst/at 201 Created
	006)MKCOL abcdefghijklmnopqrstuvw.... 201 Created
	007)PUT abcdefghijklmnopqrstuvwxyz_-.!~.... 201 Created
	008)PUT res_0.xml 201 Created
	009)PUT res_1.xml 201 Created
	010)PUT res_2.xml 201 Created
	011)PUT res_3.xml 201 Created
	012)PUT res_4.xml 201 Created
	013)PUT res_5.xml 201 Created
	014)PUT res_6.xml 201 Created
	015)PUT res_7.xml 201 Created
	016)PUT res_8.xml 201 Created

Figure 14: Expected Output of the Report RILM_STOR_TEST_AT

To clear all table entries, you can use the report `RILM_STOR_TEST_CLEAR`.

To test the Hadoop communication, you can use the report `RILM_STOR_TEST_HADOOP`. In the report selection, provide the *DB Connection Name* and *Custom Connection Type* (created in Connection to the ILM Store [page 13]), and execute.

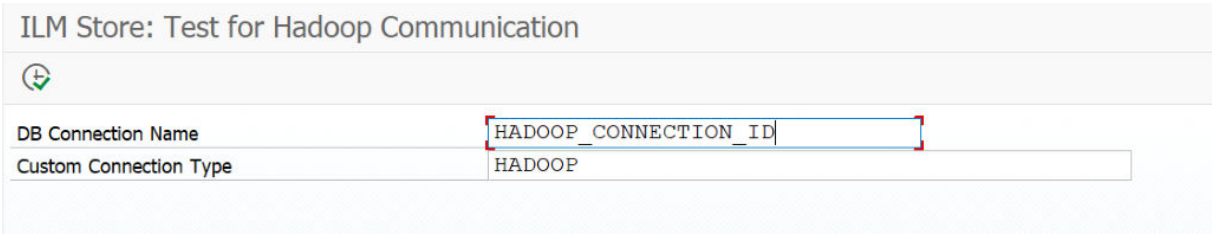


Figure 15: Selection Screen for the Report RILM_STOR_TEST_HADOOP

Figure 16: Expected Output of the Report RILM_STOR_TEST_HADOOP

9.3 Application Logs

The application log object for the ILM Store is `ILM_STOR`. To access the logs of all operations performed in the store, start transaction `SLG1` and enter the object as `ILM_STOR`.

9.4 Troubleshooting

If you run into issues, access the troubleshooting blog to find a list of common issues.

For further support, raise an incident for the application component `BC-ILM-STO`.

10 References

References	Link
ILM Store on the SAP Help Portal	Link
ILM Store Guide	Installation and Configuration Guide for the ILM Store
Hadoop Configuration Guide	Hadoop Connector Configuration Guide

Important Disclaimers and Legal Information

Hyperlinks

Some links are classified by an icon and/or a mouseover text. These links provide additional information.
About the icons:

- Links with the icon  : You are entering a Web site that is not hosted by SAP. By using such links, you agree (unless expressly stated otherwise in your agreements with SAP) to this:
  - The content of the linked-to site is not SAP documentation. You may not infer any product claims against SAP based on this information.
  - SAP does not agree or disagree with the content on the linked-to site, nor does SAP warrant the availability and correctness. SAP shall not be liable for any damages caused by the use of such content unless damages have been caused by SAP's gross negligence or willful misconduct.
- Links with the icon  : You are leaving the documentation for that particular SAP product or service and are entering an SAP-hosted Web site. By using such links, you agree that (unless expressly stated otherwise in your agreements with SAP) you may not infer any product claims against SAP based on this information.

Videos Hosted on External Platforms

Some videos may point to third-party video hosting platforms. SAP cannot guarantee the future availability of videos stored on these platforms. Furthermore, any advertisements or other content hosted on these platforms (for example, suggested videos or by navigating to other videos hosted on the same site), are not within the control or responsibility of SAP.

Beta and Other Experimental Features

Experimental features are not part of the officially delivered scope that SAP guarantees for future releases. This means that experimental features may be changed by SAP at any time for any reason without notice. Experimental features are not for productive use. You may not demonstrate, test, examine, evaluate or otherwise use the experimental features in a live operating environment or with data that has not been sufficiently backed up.
The purpose of experimental features is to get feedback early on, allowing customers and partners to influence the future product accordingly. By providing your feedback (e.g. in the SAP Community), you accept that intellectual property rights of the contributions or derivative works shall remain the exclusive property of SAP.

Example Code

Any software coding and/or code snippets are examples. They are not for productive use. The example code is only intended to better explain and visualize the syntax and phrasing rules. SAP does not warrant the correctness and completeness of the example code. SAP shall not be liable for errors or damages caused by the use of example code unless damages have been caused by SAP's gross negligence or willful misconduct.

Bias-Free Language

SAP supports a culture of diversity and inclusion. Whenever possible, we use unbiased language in our documentation to refer to people of all cultures, ethnicities, genders, and abilities.

www.sap.com/contactsap

THE BEST RUN SAP