



BEA AquaLogic® Enterprise Repository

Software File Identification (SFID) Guide

Version 3.0
Revised: August, 2007

Contents

- [Overview](#)
- [Enabling SFID](#)
- [Adding the SFID Element to a Type](#)
- [Adding an SFID to an Asset](#)
- [Limitations](#)
- [SFID Toolkit](#)
- [SFID Manager](#)

Overview

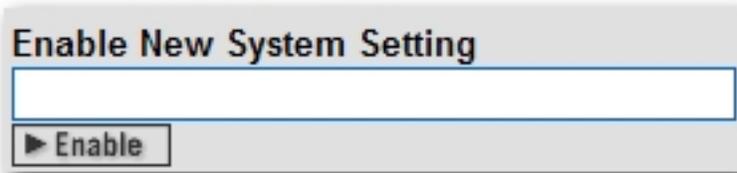
Software File Identification (SFID) provides the ability to determine asset usage independent of the manual asset **Use - Download** process within ALER. The SFID process tags selected files within an asset with a unique SFID fingerprint. This tag is then used to detect when and where an asset is used, even if the asset was acquired through means other than the **Use - Download** button in ALER. An instance of usage is recorded by ALER when tagged files within the asset are opened in a developer's IDE.

Important! SFID requires the installation of the ALER Plug-in for Eclipse, and is compatible only with Eclipse and Eclipse-based IDEs. For more information about Repository Access capabilities using Eclipse, see *Using BEA WorkSpace Studio for Repository Access* at <http://edocs.bea.com/aler/docs30/eclipse/index.html>.

Enabling SFID

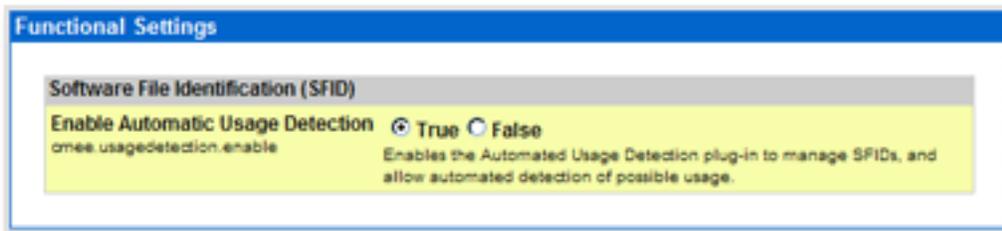
This procedure is performed on the AquaLogic Enterprise Repository (ALER) **Admin** screen.

1. Click **System Settings** in the left pane.
2. Enter `cmee.usagedetection.enable` in the **Enable New System Settings** text box.



3. Click **Enable**.

Enable Automated Usage Detection appears in the list of properties in the main pane.



4. Make sure the property is set to **True**.
5. Click **Save**.

SFID is now enabled.

- The **SFID** element can now be added to types in the **Type Manager**.
- When the **SFID** element is added to an asset type, an **SFID** section will appear on assets of that type when viewed in the **Asset Editor**. It is within that section

that files can be selected for SFID fingerprinting.

Automated Usage Detection

Please select the files that uniquely identify this asset. This information will be use...

Name	Description	Generated Date
------	-------------	----------------

Add

Remove

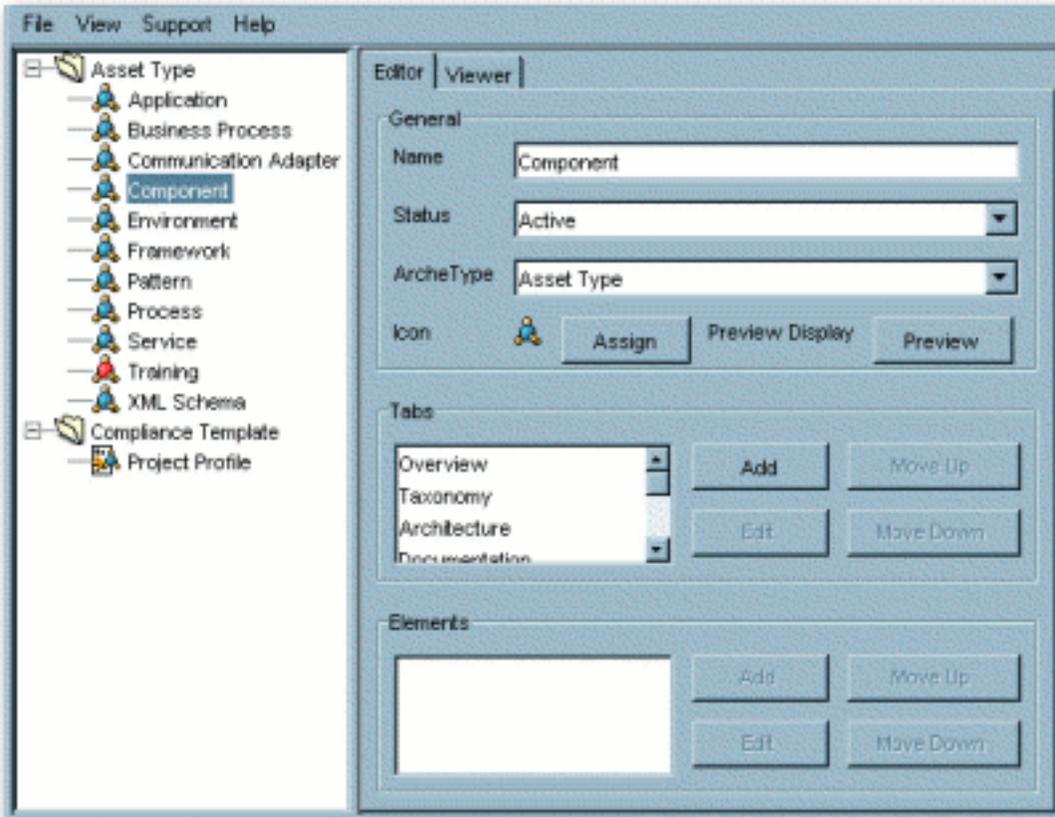
Update

- **SFID** functions are now available in the ALER OpenAPI.

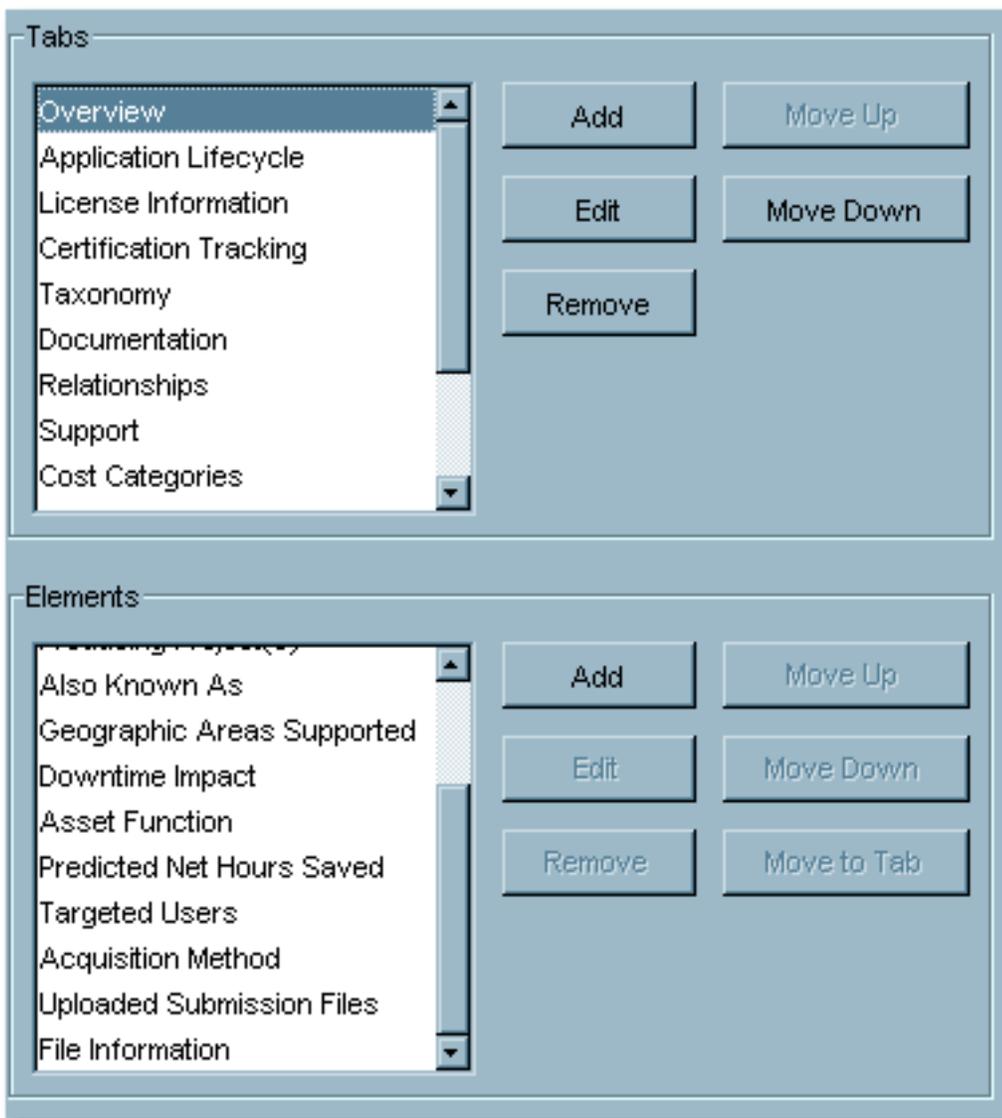
Adding the SFID Element to a Type

This procedure is performed in the **Type Manager**.

1. Select the type to which the **SFID** element is to be added.



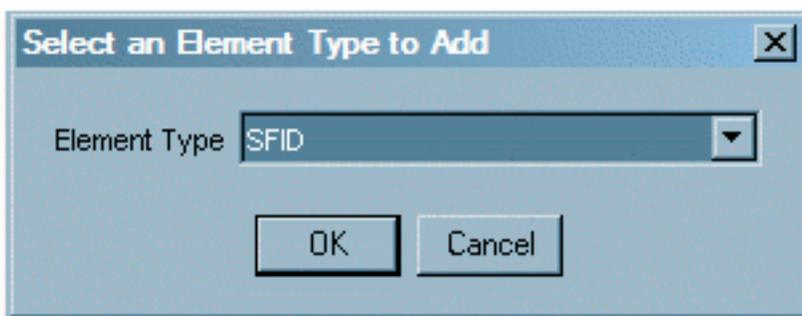
2. Click the **Editor** tab.
3. From the list in the **Tabs** section, select the tab to which to add the **SFID** element.



A list of the elements on that tab appears in the **Elements** section.

4. Click the **Add** button in the **Elements** section.

The **Select an Element Type to Add** pop-up opens.

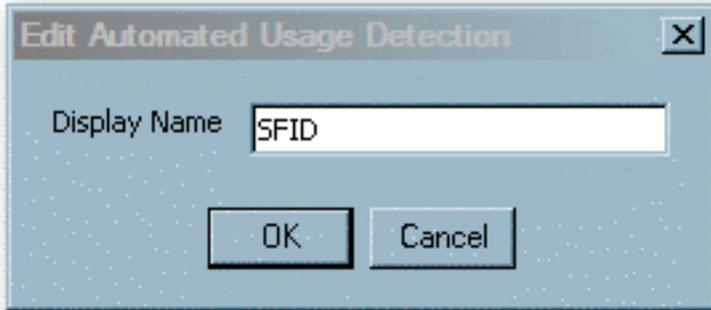


5. Select **SFID** from the drop-down menu.

6. Click **OK**.

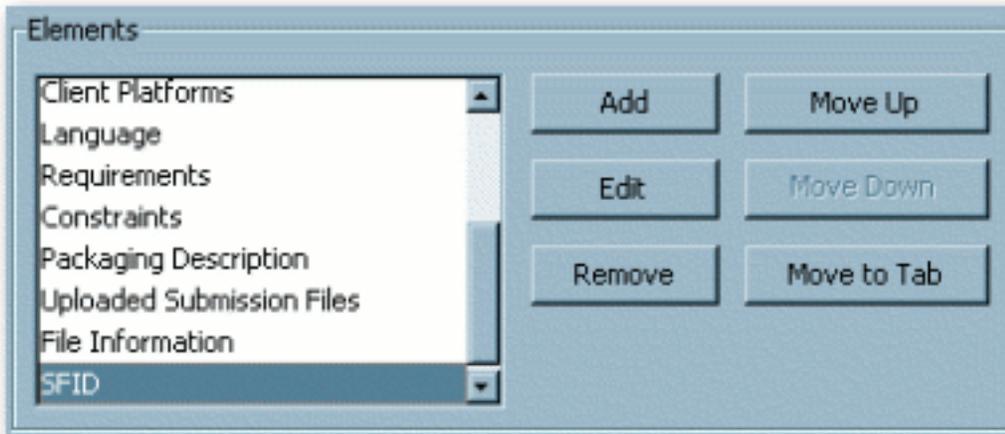
The **Select an Element Type to Add** pop-up closes, and the **Edit Automated Usage Detection** pop-up appears.

7. Enter a name for the **SFID** element.



8. Click **OK**.

The new element is listed in the **Elements** section.



9. Use the **File -> Save** menu to save the modified asset type.

10. Exit the **Type Manager**.

11. In the **Asset Editor**, open an asset of the type to which the **SFID** element was added.

The **SFID** section now appears on the tab to which it was assigned.

Automated Usage Detection

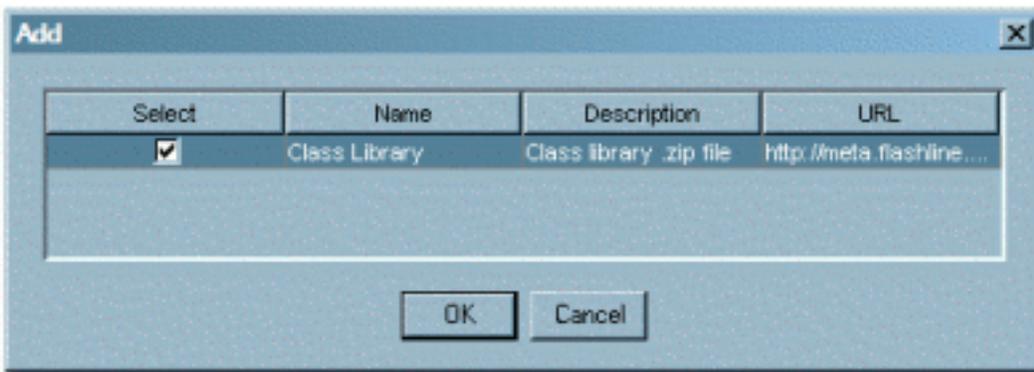
Please select the files that uniquely identify this asset. This information will be use...

Name	Description	Generated Date	Add
			Remove
			Update

Adding an SFID to an Asset

This procedure is performed in the **Asset Editor**.

1. Select the asset to which an SFID is to be added.
2. Click the **Add** button in the **Automated Usage Detection** section.
3. Select the files to be fingerprinted from the list in the **Add** dialog box.



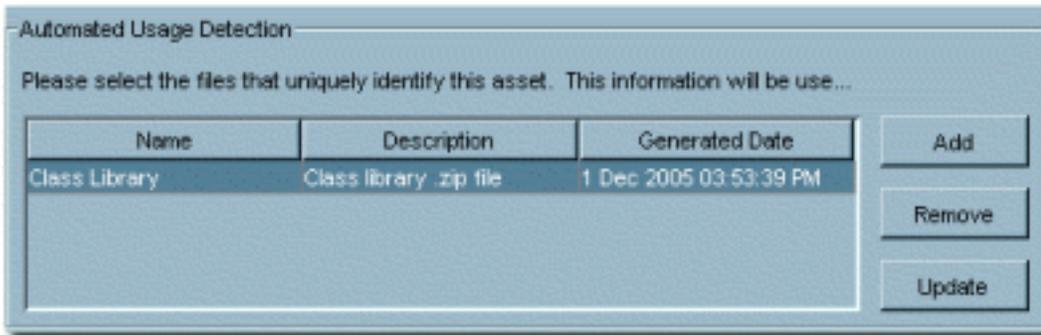
4. Click **OK**.

The system downloads the selected files, calculates the SFID, and then deletes the local (downloaded) copies of the files. A warning message is displayed if fingerprinting for any of the selected files is unsuccessful.

Update an SFID file

This procedure is performed in the **Asset Editor**.

1. Select the SFID file to be updated from the list in the **Automated Usage Detection** section.



2. Click the **Update** button.

The **Perform Operation** window opens, indicating the progress of the update.

Remove an SFID File

This procedure is performed in the **Asset Editor**.

1. Select the SFID file to be removed from the list in the **Automated Usage Detection** section.
2. Click the **Remove** button.

The file is removed from the list.

Limitations

SFID System Limitations

- The file must be accessible to generate the SFID.
 - The client machine/user must be provided with a path (URI) to the file. The URI must be accessible from the client system, without the need for user intervention, such as requests for authentication information).
- SFIDs are associated with files (1:1) which are associated with assets (n:n)
- SFIDs can be added to files through the **Asset Editor** or the **OpenAPI**.

Limitations of the Asset Editor (SFID generation)

- An SFID can be associated only with extractable files.
- An SFID can be generated only for files that are accessible to the client.

Limitations of the ANT Integration (SFID comparison)

- SFIDs can be generated for ALL accessible files (for comparison).

BEA AquaLogic™ Enterprise Repository

Software File Identification (SFID) Tool Kit

Overview

The **SFID Toolkit** provides a framework that allows third party developers to develop applications that integrate with AquaLogic Enterprise Repository's OAPI subsystem for the purpose of submitting SFID Candidates. The SFID toolkit provides an application that will create an SFID Candidate for each file in a specified directory and then submit those SFID Candidates to the AquaLogic Enterprise Repository system for storage and analysis.

Notes on MD5 hashing of text files

- A file is considered a text file if it contains no characters above ASCII code 128.
- The content of a text file is normalized before the MD5 hash is computed.
 - All CR, LF, and Tab chars are stripped.
 - All text is converted to lower case.

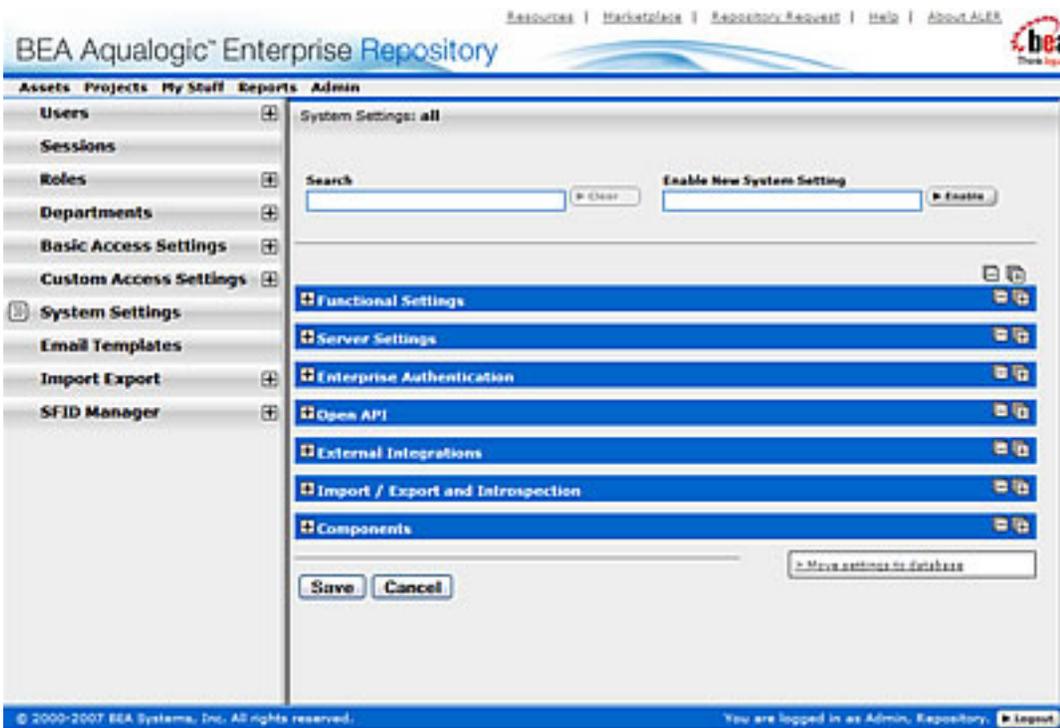
Installation / Configuration

Enable Harvesting of Unrecognized SFIDs

This procedure is performed on the AquaLogic Enterprise Repository **Admin** screen.

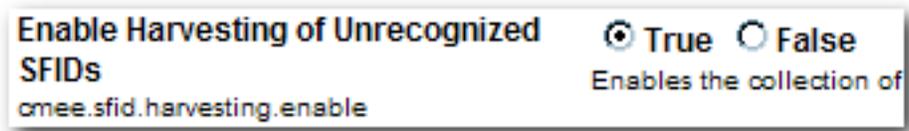
1. Click **System Settings** in the left pane.

The **System Settings** section opens in the main pane.



2. Enter `cmee.sfid.harvesting.enable` in the **Enable New System Setting** text box.
3. Click the **Enable** button.

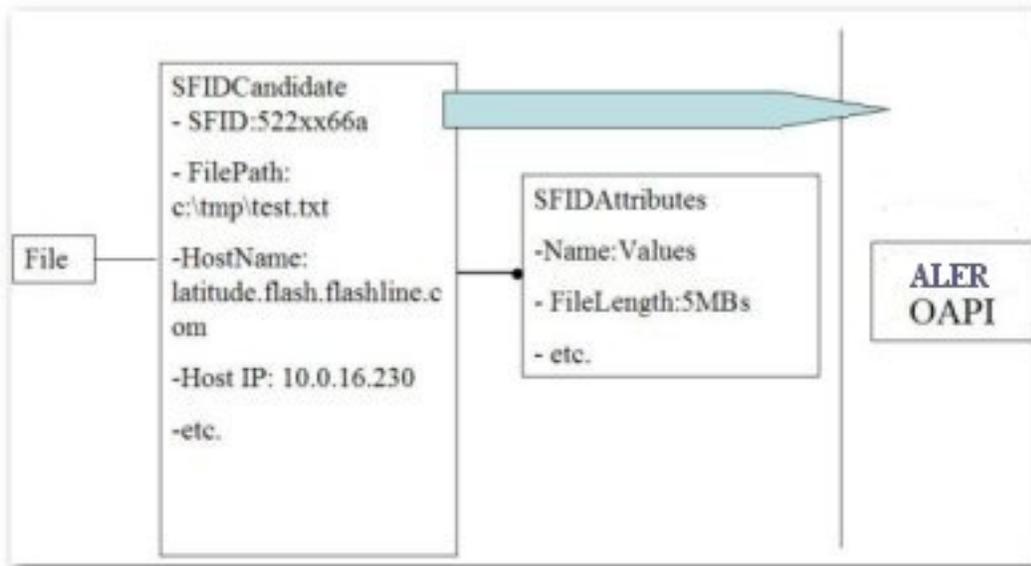
The **Enable Harvesting of Unrecognized SFIDs** property appears in the list of system settings.



4. Make sure property is set to **True**.
5. Click **Save**.

The property is now enabled.

Object Model



Description of Sample Directory Crawler

- Iterate through all the files and folders of a specified folder.
 - Instantiate an `SFIDCandidate` for each candidate
 - Populate general metadata such as path and sfid.
 - For each `SFIDCandidate`, create `SFIDAttributes` that describe arbitrary data to be captured, such as Eclipse Workspace, etc.
- Invoke the AquaLogic Enterprise Repository Open API method `sfidCandidateSubmit`, passing an array of `SFIDCandidate` objects.
- AlerUtil Object
 - Instantiate a `AlerUtil` object to connect to the AquaLogic Enterprise Repository Open API.

```
String lUrl = "http://localhost:9080/cmee/services/FlashlineRegistry";
String lUserName = "username";
String lPassword = "password";
lAlerUtil = new AlerUtil(lUrl, lUserName, lPassword);
```

- Use the `AlerUtil` object to invoke a AquaLogic Enterprise Repository Open API method.

```
lAlerUtil.getFlashlineRegistry().sfidCandidateSubmit  
(lAlerUtil.getAuthToken(), lSfidCandidates);
```

- The `AlerUtil` object provides a series of convenience methods.

- **`getAuthToken`**

Obtain a valid `authToken` for the for current username and password. If the `authToken` is expired, obtain a new `authToken`. Otherwise, a `UserValidationException` is thrown.

- **`calculatesfidValue`**

Calculate the MD5 hash for a file or path and return the hash as a string.

- **`sfidCandidateListToArray`**

Convert a list of `SFIDCandidates` to an array of `SFIDCandidates`.

- **`sfidAttributeListToArray`**

Convert a list of `SFIDAttributes` to an array of `SFIDAttributes`.

- **`sfidAttributeValueListToArray`**

Convert a list of `SFIDAttributeValues` to an array of `SFIDAttributeValues`.

- **`createCandidate(File)`**

Create a standard `SFIDCandidate` for a file. `SFIDCandidate` fields `filepath`, `hostip`, `hostname`, and `sfid` will be pre-populated. An `SFIDAttribute` named `__Filelength__` will be attached to the `SFIDCandidate` with the value of length of the file.

- **`createCandidate(File, Map)`**

Create a standard `SFIDCandidate` for a file as specified by the method `createCandidate(File)`. Each `(String)key/(String)value` in the map will attach an additional `SFIDAttribute` to the `SFIDCandidate` whose name is `key` and whose value is `value`.

- **createCandidate(File, List)**

Create a standard `SFIDCandidate` for a file as specified by the method `createCandidate(File)`. Attach each `SFIDAttribute` passed in the list to the `SFIDCandidate`.

- **sendCandidate**

Send a single `SFIDCandidate` for harvesting.

- **sendCandidates**

Send a list of `SFIDCandidates` for harvesting.

SFID Manager

Overview

The **SFID Manager** provides an easy way to view SFID candidates harvested by AquaLogic Enterprise Repository.

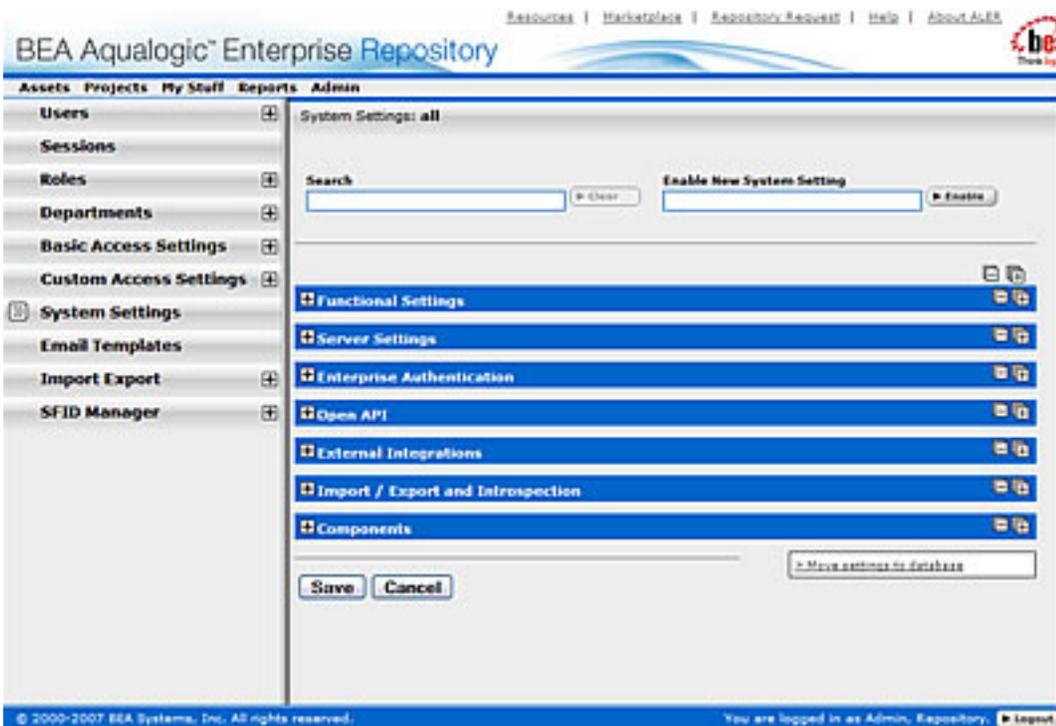
Use of the **SFID Manager** requires System Administrator privileges.

Enabling the the SFID Manager

This procedure is performed on the **Admin** screen in AquaLogic Enterprise Repository.

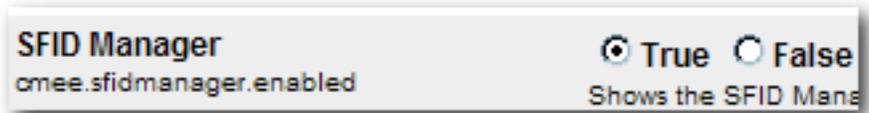
1. Click **System Settings** in the left pane.

The **System Settings** section opens in the main pane.



2. Enter `cmee.sfid.harvesting.enable` in the **Enable New System Setting** text box.
3. Click the **Enable** button.

SFID Manager is displayed in the list of system settings.



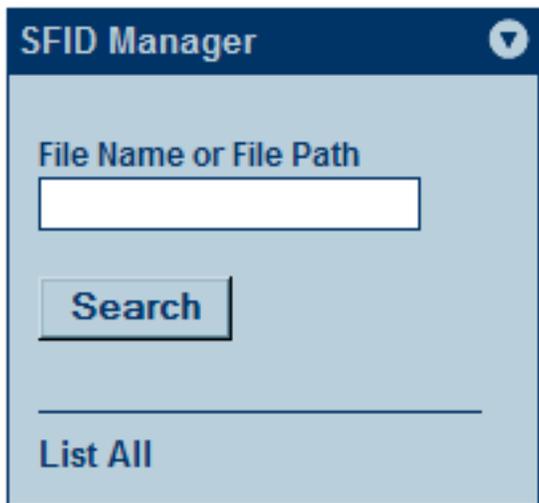
4. Make sure the **True** radio button is selected.
5. Click **Save**.
6. Click **Admin** in the ALER menu bar to refresh the **Admin** screen.

The **SFID Manager** section now appears in the left pane.

Using the SFID Manager

This procedure is performed on the **Admin** screen in AquaLogic Enterprise Repository.

1. Click **SFID Manager** in the left pane.
2. Click **List All** in the **SFID Manager** section to display a list of available SFID candidates.



Be advised that the number of available candidates affects the time it takes for this operation to complete.

SFID candidates can also be located using the **SFID Manager** search function.

Identified candidates will be listed in the upper section of the main pane.

<u>Name</u>	<u>#</u>	<u>SFID</u>
BillPayment-ideal.wsdl	6	ef89ff1d1e09dffe4529f0ebdd082fc0
Shell.java	4	aad1e63a213a770cddd906ee875ef3d8
Taco.java	6	f0f95effad81248d2f53fe9491a2c5da

The number of time the listed SFID has been detected (**Times Seen**) is indicated in the **#** column.

3. Click any listed candidate to display extended information in the bottom section of the main pane.

SFID Candidate Properties
Display Name: BillPayment-ide
Generated By: Smith, John
Generated Date: 2006-05-25 10:
Updated Date: 2006-05-25 13:
File Path: C:\workspaces
Hostname IP: 10.0.16.218
#Times Seen: 6
Source: Eclipse 3.x

The displayed information consists of a core set of attributes common to all candidates:

- **Display Name**
 - The name of the candidate. Generally, the *simple* file name.
- **Generated By**
 - The username of the user that caused the candidate to be generated.
- **Generated Date**

- The date on which the candidate was created.
- **Updated Date**
 - The last time that the candidate record was modified in the database.
- **File Path**
 - The absolute path to the file on the machine from which the candidate was harvested.
- **Hostname/IP**
 - The hostname or IP address of the machine from which the candidate was harvested
- **# Times Seen**
 - The number of times this candidate has been viewed in this context. This metric gives an indication of the frequency of file use, which provides some indication of the value of the file in this context.
- **Source**
 - A description of the context (e.g., Eclipse, Ant, etc.)

The display also includes an extended set of attributes that will vary based on the context from which the candidate was harvested.