

Publish to HBase How to configure

On this page:

- [Step 1. Launch Aspire and open the Content Source Management Page](#)
- [Step 2. Add a new Content Source](#)
- [Step 3. Add a new Publish to HBase to the Workflow](#)
 - [Step 3a. Specify Publisher Information](#)
 - [Configuration type](#)
 - [Step 3b. HBase settings file](#)

Step 1. Launch Aspire and open the Content Source Management Page

Before launching Aspire, you need to change the `felix.properties` file and add these lines if the Kerberos authentication is going to be used:

```
# To append packages to the default set of exported system packages,
# set this value.
org.osgi.framework.system.packages.extra=\
...
sun.security.krb5, \
com.sun.security.auth.callback

# The following property makes specified packages from the class path
# available to all bundles. You should avoid using this property.
org.osgi.framework.bootdelegation=\
...
javax.security.sasl, \
sun.security.krb5
```

Launch Aspire (if it's not already running). See:

- [Launch Control](#)
- Browse to: <http://localhost:50505>. For details on using the Aspire Content Source Management page, please refer to [Admin UI](#).

Step 2. Add a new Content Source

For this step, please follow the step from the Configuration Tutorial of the connector of your choice, please refer to [Connector list](#).

Step 3. Add a new Publish to HBase to the Workflow

To add a Publish to HBase, drag from the **Publish to HBase** rule from the *Workflow Library* and drop to the *Workflow Tree* where you want to add it.

This will automatically open the *Publish to HBase* window for the configuration of the publisher.

Step 3a. Specify Publisher Information

In the *Publish to HBase* window, specify the connection information to publish to the .

1. **Name:** Enter the source name to use for publishing the document.
2. **Namespace prefix:** Enter the prefix of the name space.

? Unknown Attachment

Aspire Publish To HBase Application

Coordinates: com.snowflakecomputing:aspire-app-publish-to-hbase:1.2-SNAPSHOT

System Name: Aspire Publish To HBase Application

Name: Aspire Publish To HBase Application

Namespace prefix: aspire, ☒ Create Namespace

Configuration type

☒ Use Settings file

HBase configuration file path: S:\aspire\home\config\hbaseSettings.xml ☒

☐ Do not use Settings file

☒ Clear database before full crawl

Update Cancel

3. **Create Namespaces:** Select the check box if the publisher should attempt to create the namespaces.

Configuration type

1. **Use Settings file:**
 - **HBase configuration file path:** Enter the path that contains the HBase configuration file.
2. **Do not use Settings File**
 - **Username:** Kerberos User with the permissions to publish to HBase.
 - **Keytab:** Path to the Keytab file to use.
 - **Add resources:** Select the check box to pass Hadoop resources files (hbase-site.xml) to the publisher.
 - **Resource file:** Hadoop resource file path
 - **Add properties:** Check if you are going to pass specific Hadoop properties to the publisher.
 - **Name:** Hadoop property name
 - **Value:** Hadoop property value
3. **Clean database before full crawl**
4. **Debug:** Select the check box to enable debugging messages.

The screenshot shows a configuration form titled "Do not use Settings File". It contains the following fields and controls:

- Username:** A text input field with the value "clusteruser@STDEV.LOCAL".
- Keytab:** A text input field with the value "config/kerberos/clusteruser.keytab".
- Add resources:** A checkbox that is checked.
- Add New:** A button next to the "Add resources" checkbox.
- Resource file:** A text input field with the value "config/kerberos/conf/hbase-site.xml".
- Add properties:** A button at the bottom right.

Step 3b. HBase settings file

If used, the HBase settings file has the following structure:

```
<settings>
  <properties>
    <property name="hbase.zookeeper.quorum">10.0.0.114</property>
  </properties>
  <configDir>config\kerberos\conf</configDir>
  <security>
    <kerberos >
      <user>hbase/it5452-cdh01.novalocal@STDEV.LOCAL</user>
      <path>config\kerberos\hbase.keytab</path>
    </kerberos>
  </security>
</settings>
```

1. **Properties:** Put any Hadoop property required
2. **Config Dir:** Path where the Hadoop Resources files are located
3. **Username:** Kerberos user with the permissions to publish to HBase
4. **Keytab:** Path to the Keytab file to use

Once you've clicked on the *Add* button, it will take a moment for Aspire to download all of the necessary components (the Jar files) from the Maven repository and load them into Aspire. Once that's done, the publisher will appear in the Workflow Tree.



For details on using the Workflow section, please refer to [Workflow](#) introduction.