

SharePoint 2010 FAQ and Troubleshooting

FAQs

Specific

Is incremental indexing supported by the connector?

Yes. This is done using a snapshot file which is created when doing a full crawl. It saves the state of every crawled item and on the next incremental crawl it'll use this file to compare and get any adds, updates or deletes. A new snapshot file is created after an incremental crawl to save the changes done.

Why do I need to set up permissions at Web Application level?

Those permissions ("Full Read") are required so the connector can fetch additional information from site collections. This information is required for incremental indexing.

If setting permissions at Web Application level is not suitable to your environment, consider setting site collection administrator rights to the crawler account on each site collection you want to crawl. (This will work without Web App permissions, but has to be set on every site collection.)

Why is the connector crawling the whole site collection? Even if the start URL is from a sub-site?

This usually happens when you are trying to access your site using an IP address instead of hostname. The traffic is redirected to the top level site by IIS or a router, so the connector receives the wrong information from SharePoint. To fix, simply use the hostname of your SharePoint front end.

Why is does an incremental crawl last as long as a full crawl?

In Aspire 2.0 and 2.0.1, the Sharepoint 2010 Connector performs incremental crawls based on snapshot files, which are meant to match the exact documents that have been indexed by the connector to the search engine. On an incremental crawl, the connector fully crawls the Sharepoint content the same way as a full crawl, but it only indexes the modified, new or deleted documents during that crawl.

Since Aspire 2.0.2 the Sharepoint 2010 Connector uses the change token to get only the changes instead of using the snapshot file.

For a discussion on crawling, see [here](#).

General

Why does an incremental crawl last as long as a full crawl?

Some connectors perform incremental crawls based on snapshot entries, which are meant to match the exact documents that have been indexed by the connector to the search engine. On an incremental crawl, the connector fully crawls the repository the same way as a full crawl, but it only indexes the modified, new or deleted documents during that crawl.

For a discussion on crawling, see [Full & Incremental Crawls](#).

Save your content source before creating or editing another one

Failing to save a content source before creating or editing another content source can result in an error.

```
ERROR [aspire]: Exception received attempting to get execute component command com.searchtechnologies.aspire.services.AspireException: Unable to find content source
```

Save the initial content source before creating or working on another.

My connector keeps the same status "Running" and is not doing anything

After a crawl has finished, the connector status may not be updated correctly.

To confirm this, do the following:

1. In RoboMongo, go to your connector database (like: *aspire-nameOfYourConnector*).
2. Open the "Status" collection and perform the following query:

```
db.getCollection('status').find({}).limit(1).sort({$natural:-1})
```

Key	Value	Type
(1) ObjectId("5964a5e19ff5542988bb9f7c")	{ 18 fields }	Object
_id	ObjectId("5964a5e19ff5542988bb9f7c")	ObjectId
connectorSource	{ 12 fields }	Object
@action	start	String
@actionProperties	full	String
@crawlId	0	String
@normalizedCSName	IBM_Connections	String
displayName	IBM Connections	String
@scheduler	AspireSystemScheduler	String
@scheduleId	0	String
@jobNumber	1	String
@sourceId	IBM_Connections	String
@actionType	manual	String
@dbId	0	String
crawlStart	1499768289022	Int64
crawlStatus	S	String
processDeletes	none	String
processingDeletesStatus	finished	String
crawlEnd	1499770238669	Int64

3, Edit the entry and set the status to "S" (Completed).

```
{
  "_id": ObjectId("5964a5e19ff5542988bb9f7c"),
  "connectorSource": {
    "IMServer": "https://ws8-ibm5.ga.local/",
    "IMUser": "wasadmin",
    "IMPassword": "encrypted:9D927FC87FB8745A7A68BD76DC96A96B",
    "useLTPA": "false",
    "IMLoginUri": null,
    "pageSize": "100",
    "extractACL": "true",
    "ldapComponent": "/Ldap_Cache",
    "ldapUserGUID": "myGUIDValue",
    "ldapUserName": "sam",
    "crawlByApps": "false",
    "withLimitedAccess": "false"
  },
  "action": "start",
  "actionProperties": "full",
  "crawlId": "0",
  "normalizedCSName": "IBM_Connections",
  "displayName": "IBM Connections",
  "scheduler": "AspireSystemScheduler",
  "scheduleId": "0",
  "jobNumber": "1",
  "sourceId": "IBM_Connections",
  "actionType": "manual",
  "dbId": "0",
  "crawlStart": NumberLong(1499768289022),
  "crawlStatus": "S",
  "processDeletes": "none",
  "processingDeletesStatus": "finished",
  "crawlEnd": NumberLong(1499770238669)
}
```

Note: To see the full options of "Status" values, see [MongoDB Collection Status](#).

My connector is not providing group expansion results

Make sure your connector has a manual scheduler configured for Group Expansion.

Add New

Scheduled: Manually
 Action: Start
 Crawl: Cache Groups

1, Go to the Aspire [debug console](#), and look for the respective scheduler (in the fourth table: Aspire Application Scheduler).

Aspire Application Scheduler:

Scheduler	enabled				
Name	Schedule	Last run	Next run	Status	
IBM_Connections:1	manual	never	disabled	disabled	detail run
IBM_Connections:2	manual	never	disabled	disabled	detail run

2. If you are unsure which scheduler is for Group Expansion, you can check the Schedule Detail.

- You can identify it with the value: cacheGroups

Schedule Detail
[X]

ID	1
Schedule ID	aspire.AspireSystemScheduler.1
Name	IBM_Connections:1
Schedule	manual
Source ID	IBM_Connections
Event	start
Properties	cacheGroup
Enabled	false
Singleton	true

Job

```

<doc action="start" actionProperties="cacheGroups" crawlId="0" normalizedCSName="IBM_Connections">
  <connectorSource>
    <IBMServer>https://us8-ibm5.qa.local/</IBMServer>
    <IBMUser>usadmin</IBMUser>
    <IBMPasswd>encrypted:90927FC87FB8745A7A68BD76DC96A96B</IBMPasswd>
    <useLTPA>false</useLTPA>
    <IBMLoginUrl>
    <pageSize>100</pageSize>
    <extractACL>true</extractACL>
    <ldapComponent>/Ldap_Cache</ldapComponent>
    <ldapUserGUID>myGUIDValue</ldapUserGUID>
    <ldapResourceName>/LdapResourceName
  
```

3.To run the Group Expansion process, click **Run**.

Aspire Application Scheduler:

Scheduler	enabled				
Name	Schedule	Last run	Next run	Status	
IBM_Connections:1	manual	never	disabled	disabled	detail run
IBM_Connections:2	manual	never	disabled	disabled	detail run

Troubleshooting

The connector fails with 401 unauthorized and this stack trace:

```
AspireException(aspire-sharepoint-scanner.scanFailed):
```

```

com.searchtechnologies.aspire.services.AspireException: An error occurred on SharePoint Scanner. Detail: Server
was unable to process request. ---> Attempted to perform an unauthorized operation..
    at com.searchtechnologies.aspire.components.SharePointScanner.process(SharePointScanner.java:211)
    at com.searchtechnologies.aspire.application.JobHandler.runNested(JobHandler.java:120)
    at com.searchtechnologies.aspire.application.PipelineManagerImpl.process(PipelineManagerImpl.java:196)
    at com.searchtechnologies.aspire.application.JobHandler.localJobRouteProcess(JobHandler.java:354)
    at com.searchtechnologies.aspire.application.JobHandler.runNested(JobHandler.java:249)
    at com.searchtechnologies.aspire.application.JobHandler.run(JobHandler.java:64)
    at java.util.concurrent.ThreadPoolExecutor$Worker.runTask(Unknown Source)
    at java.util.concurrent.ThreadPoolExecutor$Worker.run(Unknown Source)
    at java.lang.Thread.run(Unknown Source)
Caused by: javax.xml.ws.soap.SOAPFaultException: Server was unable to process request. ---> Attempted to perform
an unauthorized operation.
    at com.sun.xml.internal.ws.fault.SOAP11Fault.getProtocolException(Unknown Source)
    at com.sun.xml.internal.ws.fault.SOAPFaultBuilder.createException(Unknown Source)
    at com.sun.xml.internal.ws.client.sei.SyncMethodHandler.invoke(Unknown Source)
    at com.sun.xml.internal.ws.client.sei.SyncMethodHandler.invoke(Unknown Source)
    at com.sun.xml.internal.ws.client.sei.SEIStub.invoke(Unknown Source)
    at $Proxy39.getContent(Unknown Source)
    at com.searchtechnologies.aspire.components.SharePointScanner.getCurrentChangeToken(SharePointScanner.
java:508)
    at com.searchtechnologies.aspire.components.SharePointScanner.scan(SharePointScanner.java:436)
    at com.searchtechnologies.aspire.components.SharePointScanner.process(SharePointScanner.java:202)
    ... 8 more

```

This means that the service account used to access SharePoint content doesn't have enough permissions. In particular, you need to grant FULL READ permissions at Web Application level to the account if you need to proceed. Check above on the required permissions to crawl SharePoint.

Check [prerequisites](#) section for more details.

The connectors keeps failing with 401 unauthorized, even when the user has enough permissions.

Make sure the user has enough permissions. You can use the 'Check Permissions' option on the SharePoint site that you want to crawl.