

GSA Connector setup

- Installing the Adaptor
 - Fileshare
 - Sharepoint 2013
- The GSA Mock
- Running a connector

Setup a GSA Adaptor

Links to the Google Adaptors: <http://googlegsa.github.io/adaptor/index.html>

Deploying a FileSystem Adaptor: <http://static.googleusercontent.com/DeployingtheConnectorforFileSystems.pdf>

Deploying the Database Adaptor: <http://static.googleusercontent.com/DeployingtheConnectorforDatabases.pdf>

Installing the Adaptor

If you have an error (Installer reports "Windows error 2 occurred while loading Java VM") when launching the installer, try launching it with a command line: `fs-install-4.1.1.exe LAX_VM "C:\Program Files\Java\jre1.8.0_121\bin\java.exe"`

I have setup a file share and a sharepoint2013 connectors on:

Already setup connector:
vm-dev-gsa-conn

Fileshare

My test config for fileshare :

```
GSA hostname: localhost
UNC path: \\corp.coveo.com\dfs\QBC\Temporary\jpcroteau
Behavioral controls:
  Did not allow the adaptor to crawl hidden folder and files
  Did not allow indexing of folders
  Check the don't read ACLs of file shares
URLs: default value
Preserve the last access... : IF_ALLOWED
Restrict to... accessed recently?: No ...
Restrict to... modified recently?: No ...
Adaptor Port: 5978
Dashboard Port: 5979
Maximum Java Heap size (in megabytes): 1024
Do not checked start adaptor after installation finishes
```

The complete adaptor-config.properties (Default path: "C:\Program Files\GSA FS Adaptor\adaptor-config.properties"):

TestingProperties

```
# This file is UTF-8 encoded.
gsa.hostname=localhost
gsa.version=7.6.50-0
filesystemadaptor.src=\\corp.coveo.com\dfs\QBC\Temporary\jpcroteau
# one or more UNC path(s)
filesystemadaptor.src.separator=;
# the separator between UNC paths
server.port=5978
# retriever port on the local machine (for this instance of the adaptor).
server.dashboardPort=5979
```

```

# port on the local machine for accessing adaptor's dashboard.
filesystemadaptor.crawlHiddenFiles=false
# if set to true, this will crawl hidden files as well as hidden folders.
filesystemadaptor.indexFolders=false
# if set to true, this will create and index an HTML document consisting only of links.
filesystemadaptor.searchResultsLinkToRepository=true
# if set to true, results link to the repository. If false, to the adaptor (which then serves the content).
filesystemadaptor.preserveLastAccessTime=IF_ALLOWED
# valid values: ALWAYS, IF_ALLOWED, and NEVER
filesystemadaptor.directoryCacheSize=50000
# used to identify which directories are hidden or not hidden.
# a cache of 50,000 entries typically consumes 10-15 megabytes of RAM.
filesystemadaptor.statusUpdateIntervalMinutes=150000
# How often to update the Dashboard display of each file system's status.
filesystemadaptor.skipShareAccessControl=true
# If true, sends no ACLs. If false, one ACL sent for each file.
#
# Only one of filesystemadaptor.lastAccessedDate or
# filesystemadaptor.lastAccessedDays may be specified.
# filesystemadaptor.lastAccessedDate=2015-01-01
# filesystemadaptor.lastAccessedDays=365
#
# Only one of filesystemadaptor.lastModifiedDate or
# filesystemadaptor.lastModifiedDays may be specified.
# filesystemadaptor.lastModifiedDate=2015-01-01
# filesystemadaptor.lastModifiedDays=365
#
# Optional values -- for meanings of specific properties, please consult
# https://googlegsa.github.io/librarydocsreleased/4.0.3/lib/com/google/enterprise/adaptor/Config.html
#
# filesystemadaptor.supportedAccounts=BUILTIN\Administrators,\\Everyone,
# BUILTIN\Users,BUILTIN\Guest,NT AUTHORITY\INTERACTIVE,
# NT AUTHORITY\Authenticated Users
# filesystemadaptor.builtinGroupPrefix=BUILTIN\
# adaptor.incrementalPollPeriodSecs=300
# adaptor.namespace=Default
# gsa.acceptsDocControlsHeader=true
# adaptor.docContentTimeoutSec=180
# adaptor.docHeaderTimeoutSecs=30
feed.archiveDirectory=C:\\temp\\
# feed.crawlImmediatelyBitEnabled=false
# feed.maxUrls=5000
# feed.name=
# feed.noRecrawlBitEnabled=false
# gsa.version=7.0.14-114
# gsa.614FeedWorkaroundEnabled=false
# gsa.70AuthMethodWorkaroundEnabled=false
# gsa.characterEncoding=UTF-8
# gsa.samlEntityId=http://google.com/enterprise/gsa/security-manager
# gsa.scoringType=content
# journal.reducedMem=true
# server.docIdPath=/doc/
server.fullAccessHosts=wks-000299, vm-dev-gsa-connector
# server.hostname=
# server.keyAlias=adaptor
# server.maxWorkerThreads=16
# server.queueCapacity=160
# server.reverseProxyPort=5678
# server.reverseProxyProtocol=http
# server.samlEntityId=http://google.com/enterprise/gsa/adaptor
# server.secure=false
# server.useCompression=true

```

Once the install is complete:

Need to run the batch file "run" as administrator in C:\Program Files\GSA FS Adaptor

Sharepoint 2013

My config for the SharePoint 2013 connector:

▼ Adaptor-config.properties de Sharepoint

```
# This file is UTF-8 encoded.
# gsa.hostname=wks-000299
gsa.hostname=localhost
gsa.version=7.6.50-0

# sharepoint.server=https://vm-dev-sp2013:27443/
sharepoint.server=http://vm-dev-sp2013:20817/
# Above should specify the actual host name or IP address.
sharepoint.siteCollectionsToInclude=
# Limit our results to only the site collections specified. If blank, all site
# collections will be indexed.
sharepoint.username=devcoveo\devuser1
sharepoint.password=Test4sk8board

# Windows: use the above two lines if you want to override the currently-logged
# in user. Otherwise, the currently-logged in user is used.
# Linux: The two fields must be specified.

server.port=5678
# retriever port on the local machine (for this instance of the adaptor).
server.dashboardPort=5679
# port on the local machine for accessing adaptor's dashboard.
#
# Optional values -- for meanings of specific properties, please consult
# https://googlegsa.github.io/librarydocsreleased/4.1.0/lib/com/google/enterprise/adaptor/Config.html
#

# sharepoint.xmlValidation=false
# sharepoint.maxIndexableSize=2097152
# sharepoint.sts.endpoint=
# sharepoint.sts.realm=
# sharepoint.sts.trustLocation=
# sharepoint.sts.login=
# sharepoint.useLiveAuthentication=false
# adaptor.userAgent=
# adaptor.maxRedirectsToFollow=20
# adaptor.lenientUrlRulesAndCustomRedirect=true

# gsa.acceptsDocControlsHeader=true
# adaptor.fullListingSchedule=0 3 * * *
# adaptor.incrementalPollPeriodSecs=900
# adaptor.docContentTimeoutSec=180
# adaptor.docHeaderTimeoutSecs=30
# adaptor.pushDocIdsOnStartup=true
# docId.isUrl=false
feed.archiveDirectory=C:\\temp\\
# feed.crawlImmediatelyBitEnabled=false
# feed.maxUrls=5000
# feed.name=my_feed_name
# feed.noRecrawlBitEnabled=false
# gsa.version=7.0.14-114
# gsa.614FeedWorkaroundEnabled=false
# gsa.70AuthMethodWorkaroundEnabled=false
# gsa.characterEncoding=UTF-8
# gsa.samlEntityId=http://google.com/enterprise/gsa/security-manager
# gsa.scoringType=content
# journal.reducedMem=true
# server.docIdPath=/doc/
server.fullAccessHosts=wks-000299, vm-dev-gsa-connector
# 192.168.56.1
# ,wks-000299,wks-000299.corp.coveo.com,127.0.0.0,127.0.0.1,localhost,127.0.0.2,127.0.0.255
# server.hostname=localhost
# server.keyAlias=adaptor
# server.maxWorkerThreads=16
# server.queueCapacity=160
# server.reverseProxyPort=5678
# server.reverseProxyProtocol=http
# server.samlEntityId=http://google.com/enterprise/gsa/adaptor
```

```
# server.secure=false  
# server.useCompression=true
```

The GSA Mock

The mock is now also included in the [GSA connector wrapper POC](#).

You can find a GSA mock on the VM: vm-dev-gsa-conn. The mock is located at: C:\GSA\GSAMock

With this tool, any login information will lead you to the dashboard of the connector. (With my test config: <http://localhost:5979/dashboard/>)

The connector simply do a POST request to the provided GSA hostname on the port 8000.

The reponse must have a status code equal to 200 (OK) and contain the following body. The value of the "Auth" can be different. :

```
Authentication Success!  
Auth=737db7e7e42aac47e75223fb85dd3c03
```

The connector need to think that the GSA version is 7.2.0.0 or higher to send the groups. So you need to add : "gsa.version=7.6.50-0" dans le fichier de config du connecteur.

An adaptor, by default, will deny all document accesses, except from the GSA. (Source: <https://googlegsa.github.io/librarydocsreleased/4.0.3/lib/overview-summary.html>)

Running a connector

To launch the connectors, run the two following commands in a cmd:

- The file share connector: "C:\Program Files\GSA FS Adaptor\run.bat"
- The Sharepoint connector: "C:\Program Files\GSA SP Adaptor\run.bat"

If you want to access the dashboards of the connectors you will need to run the small console program that mock GSA for the connection(C:\GSA\GSAMock.exe). The connector wrapper is also mocking GSA for the connection to the dashboard.

To launch the two connectors and the wrapper, simply launch the start.bat at C:\GSA\ "C:\GSA\start.bat"

You need to run the connectors in administrator.

I did encounter a problem where the connection to the connector was refused and to solve the problem I need to add this line to my host file : vm-dev-gsa-conn vm-dev-gsa-connector.dev.coveo.com

I do not know if it is a network problem or the connector suddenly bock me.