File Conversion Component

Installation and Setup Guide

Version: 2.2.x

Written by: CSG, Global Services
Date: February 2022

Hyland

Copyright

Table of Contents

Overview

Introduction

The File Conversion Component exposes the powerful documentation conversion capabilities of Perceptive Document Filters to Perceptive Content using the Perceptive Connect Runtime. Using the File Conversion Component, you can convert documents in Perceptive Content from hundreds of different file formats to standardized file formats such as TIF, PNG or PDF. The connector allows you to convert documents in place (such as converting a PDF received using Mail Agent to TIF), or convert them for export to the file system (such as exporting a PDF formatted document).

Prerequisites

To use the File Conversion Component, the following components are required:

- **Perceptive Content** version 7.0 or higher.
- **A license for "File Conversion Component"** This license also includes the necessary licenses for Perceptive Connect Runtime, Integration Server for Apps, the Integration Framework, and Envoy for Perceptive Content versions lower than 7.2.2.
- **Integration Server for Apps** is used by the Perceptive Connect Runtime to facilitate communicate between the connector and Perceptive Content.
- **Perceptive Connect Runtime** version 1.0.10 or higher, which provides the operating environment for the connector.
- **Perceptive Content Connector,** which provides Perceptive Content services to Connect Runtime. Version 1.2 or higher is needed for Perceptive Content 7.2.2 or higher.

For versions lower than 7.2.2:

- **Envoy** provides the connective trigger between the Workflow process and the connector.

For Linux:

- **Document Filters** will needed to be installed separately.

Because of the CPU load incurred during document conversion, it is recommended that you install the connector on a remote server, separate from your Perceptive Content server. Currently, the connector must be installed on a Windows host with at least 2GB of RAM.

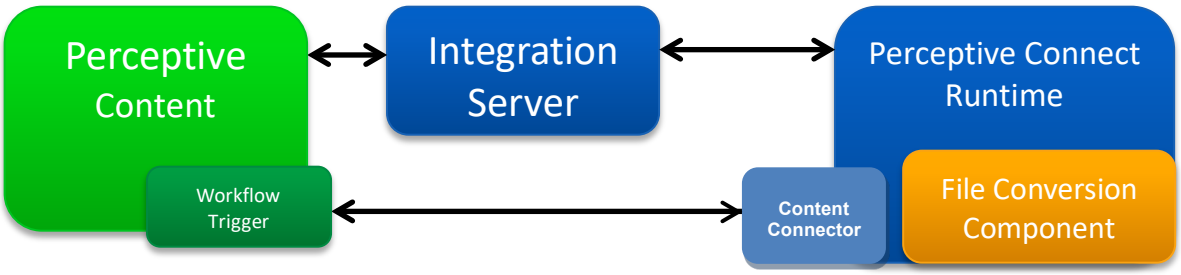
Installation Overview

To install the File Conversion Component, complete the following steps in order:

1. Install all prerequisites
2. Install and configure the connector for the File Conversion Component
3. Create a conversion or export profile
4. Setting up the Workflow trigger
5. Configure the conversion or export workflow queue process.
6. Configure the integration/connect queue mapping.

High Level System Architecture

The following diagram is a high level representation of the File Conversion Component system architecture.



Install the File Conversion Component

Install Required Prerequisites

Licenses

Work with your account executive to obtain an appropriate license(s). Install the Perceptive Content licenses before installing Connect Runtime or Integration Server.

Install Integration Server

If you do not already have an Integration Server installation, one will be required for use with Connect Runtime. You may co-locate this service on the same server as connect runtime if needed. For more information, refer to the Integration Server installation guide.

Install Perceptive Connect Runtime

Install Perceptive Connect Runtime on the remote server that will be used for conversion. It is recommended that this installation of Connect Runtime be used solely for conversion and not shared with other connectors (such as the SAP or Intelligent Capture connectors). For more information, refer to the Connect Runtime installation guide.

Note If you are upgrading the File Conversion Component to a newer version, you must restart Perceptive Connect Runtime service after the upgrade is complete.

Install the Content Connector

For more information, refer to the instructions in the Content Connector installation guide.

Per the installation instructions, configure the content connector with the applicable integration server connection information and login credentials.

Install Document Filters *(Linux only)*

Add the following 2 lines to **linuxEnvironmentFile**. This file can usually be found ../connect/conf (i.e. where PCR is installed).

LD_LIBRARY_PATH='/opt/docfilters/bin/linux/intel-64'

ISYS_FONTS='/opt/docfilters/bin/linux/intel-64/fonts'

Note: In this file you are telling PCR where it can find the fonts and the Document Filters library (i.e .so) files by setting two environment variables.

Once this file has been updated, restart PCR via command line.

Windows Note: Document Filters should not already exist on the server you are installing File Conversion Component.

Install and Configure the File Conversion Component

To install and configure the File Conversion Component, complete the following steps.

1. Open a web browser and complete one of the following actions.
   - If on the Perceptive Connect Runtime (PCR) server, type http://localhost:<port of PCR>.
   - If on another server, type http://<<PCR server name>>:<<port>>.
2. Click **Perceptive Connect Runtime**.
3. In the User Name and Password fields, type the appropriate credentials. This is typically **admin** for username and **admin** for password.
4. Click **Install a Connector**.
5. Drag and drop the **DocumentFiltersService-<*version*>.jar file** into the dotted outlined box.
6. When the upload is complete, click **Accept**.
7. To configure the basic settings of the connector, complete the following substeps.
   1. Under **MANAGE**, click **Configure**.
   2. Under **File Conversion Component**, click **Configure**.
   3. In the **Workers Thread** field, type the desired number of worker threads. This **should not exceed your licensed number of cores**.
   4. In the **Task Timeout** field, type the time in seconds that you want the system to wait before it terminates a conversion or export. The default is 60, but you can set it to 90 or 120 if the volume of documents is very large.
   5. Click **Save**.
8. To configure and connect Integration Server, complete the following substeps.
   1. Under **Configure**, click **Connection Manager**.
   2. In the **Connection Provider Target** list, select the correct target. This typically matches the installed version of Integration Server.
   3. In the **User name** and **Password** fields, type the credentials you use to connect to Perceptive Content.

      Important This user should be a manager user in Perceptive Content. We also recommend that you do not change this user's password. You must repeat this step each time you change the password. If a manager user is not an option, then use a Service User Account.

4. Click **Save**.
5. Under **Configuration**, locate and open the Connection Provider Target configuration you selected in step 2. For example, Integration Server x.x Connection.
6. In the **Integration Server URL** field, type the appropriate Integration Server URL.
7. Click **Save**.

Note If you are upgrading the File Conversion Component to a newer version, you must restart Perceptive Connect Runtime service after the upgrade is complete.

Create a Conversion or Export profile

Conversion and Export profiles controls the conversion process and output format. For example, the profile determines if the document will be converted to TIF or PDF and at what resolution. Conversion profiles are used for converting a document in-place. For example, converting a document that has a PDF page to a document that has only TIF pages. Export profiles are used for converting a document to an output directory on the file system, such as converting a document with all TIF pages to a PDF file placed on the server.

You will later create one workflow process for each conversion or export profile. This will allow you to route a document to one queue to convert it to B&W tiff, to another queue to convert it to color PNG, or to another queue to export as a PDF file.

To create a new profile, complete the following steps.

1. From the **Perceptive Connect Runtime** menu, under **MANAGE**, select **Configure** or **View Configuration** depending on your version of PCR.
2. Click **+**. Depending on your use case, this button is located next to **FCC Convert Profiles** or **FCC Export Profiles**.
3. In the **FCC Convert Profiles** configuration window, complete the following substeps.
    1. In the **Unique Name** field, type a unique name. Make note of this name.
    2. In the **Split Results** list, select the option for how you want to divide the resulting documents.
    3. In the **Update Mode** list, select the option for how you want to update existing documents.
    4. Select **Backup Original** if you want the system to create a backup of the original document.
    5. In the **Backup Drawer** field, type the name of the drawer of where you want to place the backup documents.
    6. In the **Backup Queue** field, type the name of the archival or delete queue where you want the system to route the backup documents.
    7. In the **Output Format** list, select the conversion output type.
    8. In the **DPI** field, type the numeric value that you want t to use when rendering the document.
    9. In the **TIFF Compression** list, select the appropriate compression for the TIF output.
    10. In the **TIFF Colorspace** list, select the appropriate colorspace for the TIF output.
    11. In the **Convert Word Docs With** list, select the library you want to use for converting Word documents.
    12. In the **Convert Excel Files With** list, select the library you want to use for converting Excel files.
    13. In the **Convert PowerPoint Files With** list, select the library you want to use for converting

PowerPoint files.

14. In the **Convert XML With** list, select the library you want to use for converting XML files.
15. In the **DocFilters Parameters** field, type the additional parameters that you want to pass to Document Filters.
16. Select **Extract Text** if you want to extract text from texted-based document files.
17. In the **Component group** list, select the Perceptive Connect component group of this profile.
18. In the **File Extensions to Ignore** field, type the extensions of the files that you do not want to convert.

4. In the **FCC Export Profiles** configuration window, complete the following substeps.
    1. In the **Unique Name** field, type a unique name. Make note of this name.
    2. In the **Split Results** list, select the option for how you want to divide the resulting documents.
    3. In the **Export Filename Template** field, type a filename template for the exported document.
    4. In the **Export Directory** field, type the path to the export directory.
    5. In the **Output Format** list, select the conversion output type.
    6. In the **DPI** field, type the numeric value that you want t to use when rendering the document.
    7. In the **TIFF Compression** list, select the appropriate compression for the TIF output.
    8. In the **TIFF Colorspace** list, select the appropriate colorspace for the TIF output.
    9. In the **Convert Word Docs With** list, select the library you want to use for converting Word documents.
    10. In the **Convert Excel Files With** list, select the library you want to use for converting Excel files.
    11. In the **Convert PowerPoint Files With** list, select the library you want to use for converting PowerPoint files.
    12. In the **Convert XML With** list, select the library you want to use for converting XML files.
    13. In the **DocFilters Parameters** field, type the additional parameters that you want to pass to Document Filters.
    14. Select **Extract Text** if you want to extract text from texted-based document files.
    15. In the **Component group** list, select the Perceptive Connect component group of this profile.
    16. In the **File Extensions to Ignore** field, type the extensions of the files that you do not want to convert.

Set up the Workflow Trigger

To set up the Workflow trigger, complete the following steps.

7.2.2 and newer:

1. Open <*drive letter*>:\\inserver\etc\inserverWorkflow.ini.
2. Set **connect.uri** to: http://<<PCR server name>>:<<port>>/rs/workflowTrigger?_wadl

7.2.2 and older:

1. Open the **Perceptive Content Management Console**.

2. Select **Envoy Services** and click **New**.
3. In the **Envoy Services** dialog box, complete the following substeps.
    1. In the **Name** field, type an appropriate name.
    2. In the **Description** field, type any additional information.
    3. In the URL field, type http://<<PCR server name>>:<<port>>/ws/workflowTrigger?wsdl.
    4. Click **Next** and then **Finish**.

Recommendation: The File Conversion Component process is a workflow process. Therefore you need to ensure that the InserverWorkflow users defined has at a minimum read/write/read-execute access to the messages.log file.

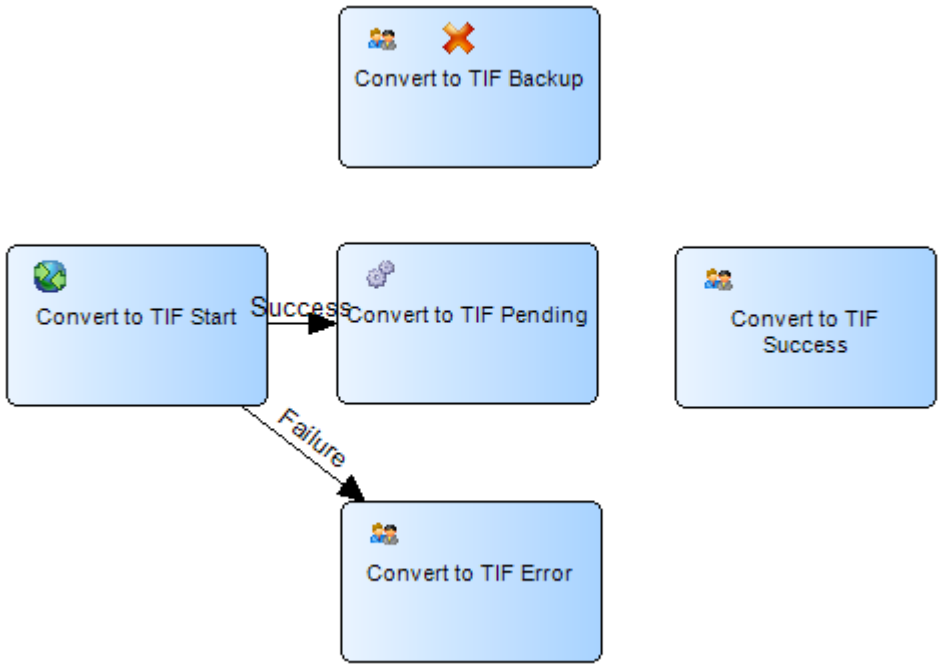
Create the Conversion or Export Workflow Process

A workflow process within content is used to submit documents for conversion or export. You can create multiple workflow processes for various conversion and export needs that each map to a different profile. To start, create a workflow process like the one pictured below.

Note: In 7.2.2 and newer, **Convert to TIF Start** queue is a **Connect** queue instead of an **Integration** queue.

Documents are first placed in the **Convert to TIF Start** queue, are automatically moved to the **Convert to TIF Pending** and then move to the **Convert to TIF Complete** after conversion finishes. If backups are enabled, the **Convert to TIF Backup** queue can be used to automatically remove backups after a certain period of time.

The following diagram is an example of a File Conversion Component workflow.



Configure Convert to TIF Start

To configure Convert to TIF Start, complete the following steps.

1. Create all other queues in the workflow first.
2. Drag a new **Connect** onto the Canvas, or **Integration** queue for versions lower than 7.2.2.
3. Double click the queue to open it.
4. In the Queue Properties dialog box, complete the following substeps.
    1. In the Name field, type a name for the queue.
    2. Make note of the **Queue ID**. You will need this later.
    3. For **Success Action**, select the **pending queue**. This should **NOT** be the **Convert to TIF Success** queue. You will assign this queue in the next mapping step.
    4. For **Failure Action**, select the error queue.
    5. Versions 7.2.2 and earlier. For **Service Operation Name**, select the Envoy Service name you created previously.

Configure the Integration/Connect Queue Mapping

To map the Integration/Connect queue to the conversion or export profile, complete the following steps.

1. Open the **Perceptive Connect Runtime Web** dashboard.
2. Click **Create a Channel**.
3. Optional. In the **Channel Name** field, type a name for the channel.
4. Select **Integration ASQ Trigger** as the trigger.
5. In the **Queue ID** field, type the queue ID of the Integration/Connect queue you previously noted.
6. Click **Continue**.
7. Select **ConvertAction** action for in-place conversions or **ExportAction** action for conversion for output.
8. Remove all the current XML from **Inputs**.
9. Copy the appropriate XML from one of the following examples and paste it into **Inputs**, substituting just **(1)** the queue name of your success queue (e.g. Convert to TIFF Success), and **(2)** the profile name of the conversion or export profile you created previously (e.g. BWTIF). **Do not update** any other parts of the mapping.

    Conversion Mapping example

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<c:inputs xmlns:c="http://www.perceptivesoftware.com/pif/mapping"
xmlns:ns2="http://www.perceptivesoftware.com/pie/mapping/sample"
xmlns:in="http://www.perceptivesoftware.com/pif/imagenow">
    <c:parameter>
        <c:name>ConversionProfileName</c:name>
        <c:literal>CONVERSION PROFILE NAME</c:literal>
    </c:parameter>
    <c:parameter>
        <c:name>PendingQueue</c:name>
        <c:trigger>SuccessQueueName</c:trigger>
    </c:parameter>
```

```
    <c:parameter>
        <c:name>SuccessQueue</c:name>
        <c:literal>SUCCESS QUEUE NAME</c:literal>
    </c:parameter>
    <c:parameter>
        <c:name>ErrorQueue</c:name>
        <c:trigger>FailureQueueName</c:trigger>
    </c:parameter>
    <c:parameter>
        <c:name>WorkflowItemID</c:name>
        <c:trigger>WorkflowItemId</c:trigger>
    </c:parameter>
</c:inputs>
```

Export Mapping example

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<c:inputs xmlns:ns2="http://www.perceptivesoftware.com/pie/mapping/sample"
xmlns:in="http://www.perceptivesoftware.com/pif/imagenow"
xmlns:c="http://www.perceptivesoftware.com/pif/mapping">
    <c:parameter>
        <c:name>ExportProfileName</c:name>
        <c:literal>EXPORT PROFILE NAME</c:literal>
    </c:parameter>
    <c:parameter>
        <c:name>PendingQueue</c:name>
        <c:trigger>SuccessQueueName</c:trigger>
    </c:parameter>
    <c:parameter>
        <c:name>SuccessQueue</c:name>
        <c:literal>SUCCESS QUEUE NAME</c:literal>
    </c:parameter>
    <c:parameter>
        <c:name>ErrorQueue</c:name>
        <c:trigger>FailureQueueName</c:trigger>
    </c:parameter>
    <c:parameter>
        <c:name>WorkflowItemID</c:name>
        <c:trigger>WorkflowItemId</c:trigger>
    </c:parameter>
</c:inputs>
```

10. Click **Save Inputs**.
11. Click **Validate Inputs**.
12. Click **Enable Channel** to activate channel. The system completes the setup and any document placed in the Start queue is now converted/exported based on the configured profile.

Configure Content Access

There are several areas where user access settings should be thought through and planned before **production** deployments are put into service. Below are some minimal requirements to consider. Within the Perceptive Content Management Console ensure the following permissions are set. The user and/or group to focus on is the same user defined previously in the Connection Manager within PCR. This is the user that will utilize **Integration Server** to engage Perceptive Content through the File Conversion Component processs.

- User has “Allow” access to the File Conversion Component process default drawer(s):
    - Content => Open
    - Documents => Page Delete
- (Optional if using Backups) User has “Allow” access to the File Conversion Component process Backup Drawer:
    - Content => Open
    - Content => Create/Append
- User has “Allow” access to the File Conversion Compnent process document type(s):
    - Document => Open
    - Document => Page Delete

Appendix

Advanced Topics

Native Conversion

Document Filters is capable of rendering hundreds of different file types using its own, well-established rendering engine. Some, however, might find small variations in the rendering output compared to the rendering of the native application unacceptable to a given solution. For this reason, native conversion is also available for Microsoft Office file formats.

Native conversion uses Microsoft Office APIs to convert the source document first to PDF, then uses Document Filters to convert the PDF to the output format. Native conversion will increase conversion time by about 20%, but will improve the fidelity to match what a document might look like when printed from Office.

To enable native conversion, change the converting application from Document Filters to the appropriate Microsoft Office application in the conversion or export profile.

File Type Exclusions

There may be instances where you would like to ignore specific files types during the conversion process. To do this add the file extension in the section labeled “File Extensions to Ignore”. The plus and minus buttons to the right of this configuration option will allow you to add or subject file extensions.

High Availability

Both Integration Server and Connect Runtime are web applications that can be load balanced in an high availability configuration. Ensure that, in an Active-Active environment, Integration Server is also pointed at the load-balanced endpoint for the Perceptive Content server.

High Volume

For high volume conversions consider getting FCC_FeedScript.js from the SDE team. This will prevent the FCC start queue from getting too backed up. Configure **maxItemsToFeedFCC** (in the script) to the average number of items FCC can process (for this customer) in 1 minute and then add 1.

DFParams Values

Multiple DFParams can be set as part of the conversion process so that additional Document Filters conversion options not specified in the default configuration can be supplied. These values should be equivalent to the Open Document Options and Open Document Flags specified in the Perceptive Document Filters Implementation Guide. If multiple values are to be provided they should be delimited by a semicolon.

Troubleshooting

For move in-depth troubleshooting tips and tricks review the Knowledgebase article titled “File Conversion Component Technical Overview”.

- **When configuring Envoy, I get this error: "The remote service identified by the URI cannot be opened.”**

  This could be caused by any number of issues related to connectivity between the Content server and Connect Runtime. Can you access the same URL using a browser? If so, make sure the Content service account is a member of the Administrators group of the Content server and restart all Content services.

  If you cannot access the URL through a browser, check that, in the Connect Runtime configuration, you have properly configured the Content connection with an accessible Integration Server URL, username and password.

- **A documented was routed to the error queue. How do I find out why?**

  When conversion or export of a document encounters an error, details of the error are logged to the workflow item routing history. If the reason is “Web Service call failed,” the Content server could not contact the Connect Runtime server. For any other error, you will find additional details in the Connect Runtime log, found in:

  {PCRInstallationDirectory}/logs/pif.all.log

- **I see in the log file that the user was unauthorized. What do I need to check to correct this?**

  Double check the account you used in the Connection Manager within Connect Runtime. This user should be able to view documents within the Perceptive Content system.