<Team>

Production Operation

<Data Pipeline>

Run Book

**Prepared for: ML Operations Team**

For the Delivery of

Integration of <Data Pipeline>

Contact

**AI Data Integration Team**

**Matt Grant:** [**Matthew.Grant@cotiviti.com**](mailto:matthew.grant@cotiviti.com)

Table of Contents

[Project Overview 2](#_Toc1882438284)

[Project Operation and Run Detail 3](#_Toc105149094)

[Raw File Source 3](#_Toc2012611486)

[Output Target 4](#_Toc1043264221)

[Pipeline 5](#_Toc1288060153)

[Adhoc Processing of Client Files 6](#_Toc1796464181)

[Pull new client files 6](#_Toc747671610)

[Run pipeline E2E with new ingestion date 7](#_Toc1713059781)

[Monitoring 7](#_Toc1027695195)

[Status Check 7](#_Toc1741905329)

[Email Notification 7](#_Toc1989309498)

[Nile Pipeline Runs Monitor 7](#_Toc1106426115)

[Hue Job Browser 8](#_Toc906598247)

[Troubleshooting 8](#_Toc1588013329)

# Project Overview

BCBSNC sends tba institutional, tba professional, history institutional, and history professional claims for both Amisys and Facets in GDF V2.5 format. All files are pgp encrypted .txt files that are received from the client on transfer2, and then pushed automatically to B2B. Nile then pulls the data from B2B into HDFS, processes the data at different intervals based on the dataset into CDF V2.5, and pushes PPM TBA/History Claims data to SFTP location (T1).

The data receipt from the client and the pipeline scheduling is provided below

# Project Operation and Run Detail

## Raw File Source

To do: Write some information about raw file sources such as entities, filenames, scheduled receiving (time, daily/weekly/monthly), ect.

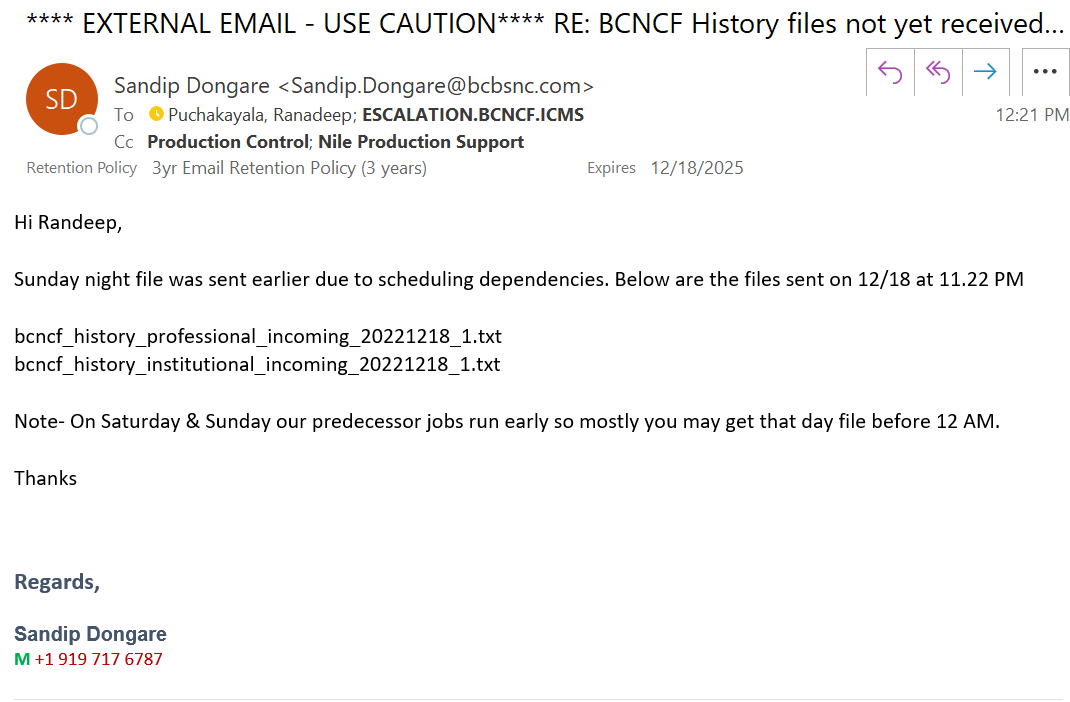
<Name 1> (Production)

|  |  |
| --- | --- |
| Entity | File |
|  |  |
|  |  |
|  |  |
|  |  |

<Name 2> (Production)

|  |  |
| --- | --- |
| Entity | File |
|  |  |
|  |  |
|  |  |
|  |  |

To do: Include a sample of emails that can be send to recipients. (see example below)



<Name 1> (MVP)

|  |  |
| --- | --- |
| Entity | File |
|  |  |
|  |  |
|  |  |
|  |  |

<Name 2> (MVP)

|  |  |
| --- | --- |
| Entity | File |
|  |  |
|  |  |
|  |  |
|  |  |

Production

**Name 1**:

**<Pipeline>**

Frequency:

Number of files:

Format of file:

Spooler:

Spooler CRON (UTC):

For example:

History

Frequency: Weekly

Number of files: 2 .txt files (institutional and professional) we will receive daily, Monday-Sunday at 2 AM ET

Format of file: bcnca\_history.\*.txt.pgp

Spooler: bcbsnc-amisys-history

Spooler CRON (UTC): 2/15 \* \* \* \*

MVP:

**Name 1**:

**<Pipeline>**

Frequency:

Number of files:

Format of file:

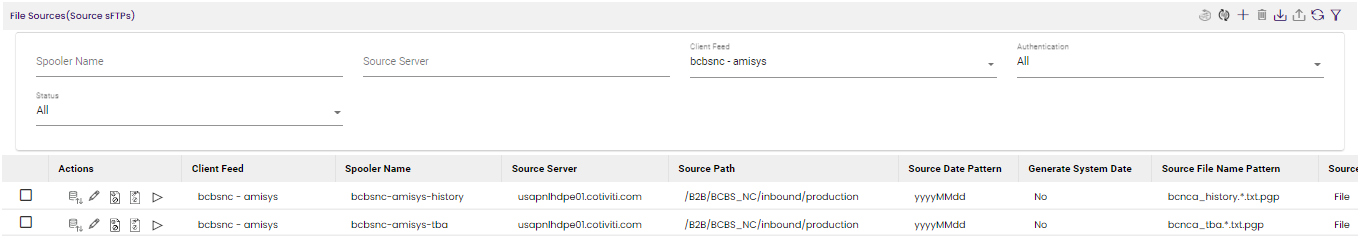
Spooler:

Spooler CRON (UTC):

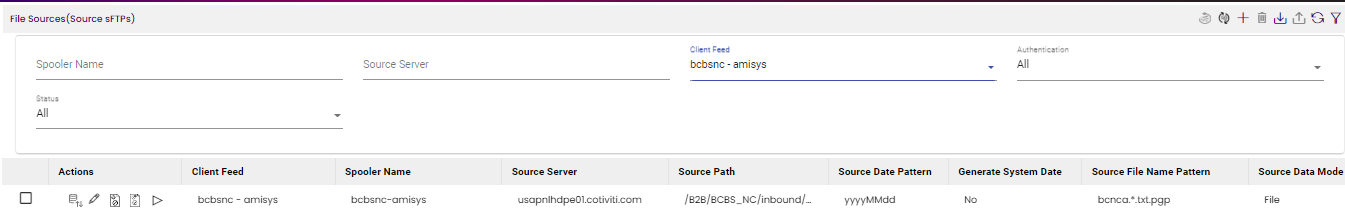
To do: Include some snapshots of tools that can be used for this. (see examples below)

Function Access: Hamburger > Site Administration > File Sources

**Production**:



**QC**:



Credentials for the Oscar transfer1 account are below

URL: https://transfer1.cotiviti.com/

username: BCBS\_NC

pwd: given upon request

Credentials for the Oscar transfer2 account are below

URL: https://transfer2.cotiviti.com/

username: BCBS\_NC

pwd: given upon request

## Output Target

To do: Write some information about output target such as data/file location, server, path, etc. (see examples below)

**Production**

**Amisys**:

File Location name : bcbsnc-amysys-T1-nile\_prod

File server: transfer1.cotiviti.com

File path : /inbound/nile\_production

**Facets**:

File Location name : bcbsnc-facets-T1-nile\_prod

File server: transfer1.cotiviti.com

File path : /inbound/nile\_production

**QC**

**Amisys**:

File Location name : bcbsnc-amysys-T1-nile\_prod

File server: transfer1.cotiviti.com

File path : /inbound/nile\_production

**UAT Facets**:

File Location name : bcbsnc-uat-T1-prod

File server: transfer1.cotiviti.com

File path : /inbound/nile\_production

## Pipeline

To do: Write some information about pipelines details here including some snapshots from UI software/tools (see examples below)

**Pipeline Details:**

**Production:**

1. **Name:**   
   **Pipeline Name:** ppm-bcbsnc-amisys-tba\_institutional\_claims\_GDF2\_5\_CDF2\_5

**Version:** V1.9

**Id:** 2022-02-18-110892468

**Description:** Process daily tba institutional claims files of PPM and pushes cdf to T1

**UTC cron expression:** **1/10 \* \* \* \***

**2. Name:**   
 **Pipeline Name:**  ppm-bcbsnc-amisys-tba\_professional\_claims\_GDF2\_5\_CDF2\_5

**Version:** V1.11

**Id:** 2022-02-18-111127854

1. **Description:** Process daily tba professional claims files of PPM and pushes cdf to T1

**UTC cron expression:** **1/10 \* \* \* \***

**QC:**

1. **Name:**   
   **Pipeline Name:** ppm-bcbsnc-amisys-tba\_institutional\_claims\_GDF2\_5\_CDF2\_5

**Version:** V1.10

**Id:** 2022-02-18-110892468

**Description:** Process daily tba institutional claims files of PPM and pushes cdf to T1

**UTC cron expression:** **5/30 \* \* \* \***

**2. Name:**   
 **Pipeline Name:**  ppm-bcbsnc-amisys-tba\_professional\_claims\_GDF2\_5\_CDF2\_5

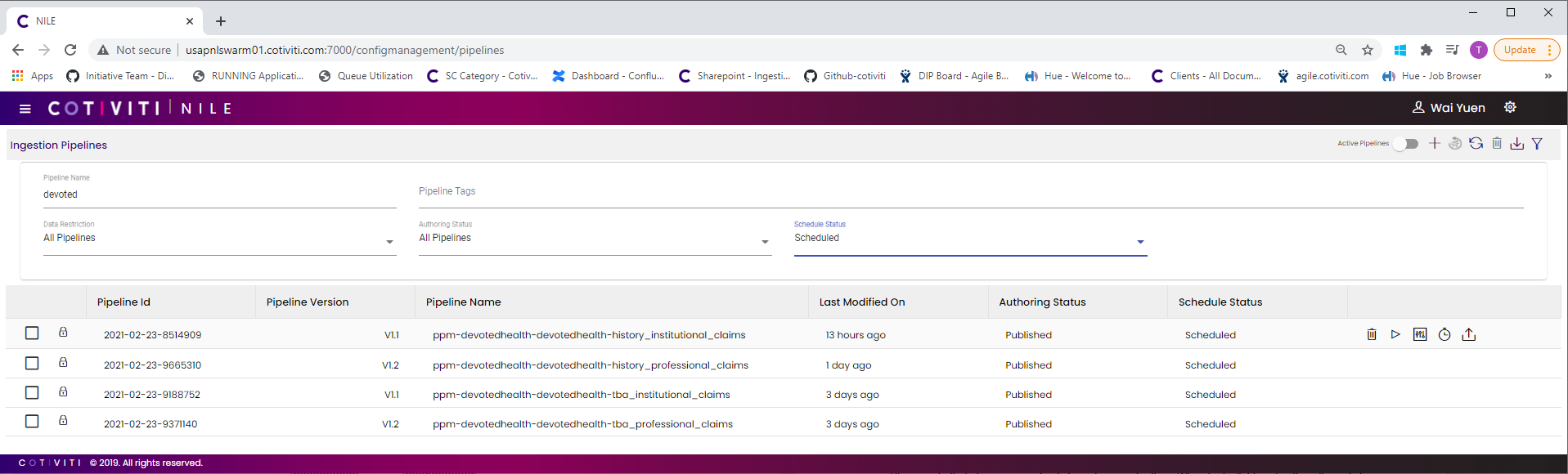
**Version:** V1.11

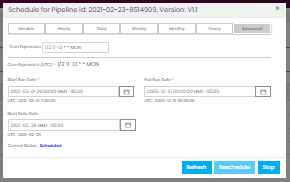
**Id:** 2022-02-18-111127854

1. **Description:** Process daily tba professional claims files of PPM and pushes cdf to T1

**UTC cron expression:** **5/30 \* \* \* \***

Click on the “Schedule Pipeline Runs” icon to set/stop schedule.





## Adhoc Processing of Client Files

# 

# Monitoring

To do: Write some details about what you want these to be monitored (via email notification, monitoring UI, or Hue Job browser, etc.)

## Status Check

# Troubleshooting

To do: Write some details about what need to be checked at Level 1 Production support and include the resolution. (see example below)

To troubleshoot a failed job, go to the corresponding Oozie workflow (refer to Monitoring section #3 above). Drill down (click on Tasks tab then External id link of task with ERROR status) to the bottom level failed task for stdout and stderr.

|  |  |  |
| --- | --- | --- |
|  | Fail Reason | Resolution |
| 1 | Hadoop cluster services down or resources inadequate, including out of disk space | Send out email to [hadoop.administration@cotiviti.com](mailto:hadoop.administration@cotiviti.com) to report system down. After system is back up and running, rerun pipeline manually. |
| 2 | Platform command failed, usually due to postgres database unavailability | Manually rerun once. If problem persists, then post in the DI-Nile-PlatformSupport Teams chat group providing the application id of the failed task. After confirming issue resolved, rerun pipeline manually. |
| 3 | Error: 400 bad request. More than 2000 jobs are triggered because of which any new job which is triggered gives an error. | DI-Nile-PlatformSupport Team will clear the queue and helps in executing jobs |