StreamSets Data Collector 5.7.1：Dockerの展開とWebSocketパイプラインの作成

本ガイドでは、StreamSets Data Collector 5.7.1をDockerコンテナに展開する手順、およびWebSocketに接続して受信したデータをローカルに保存するパイプラインの作成方法を説明します。

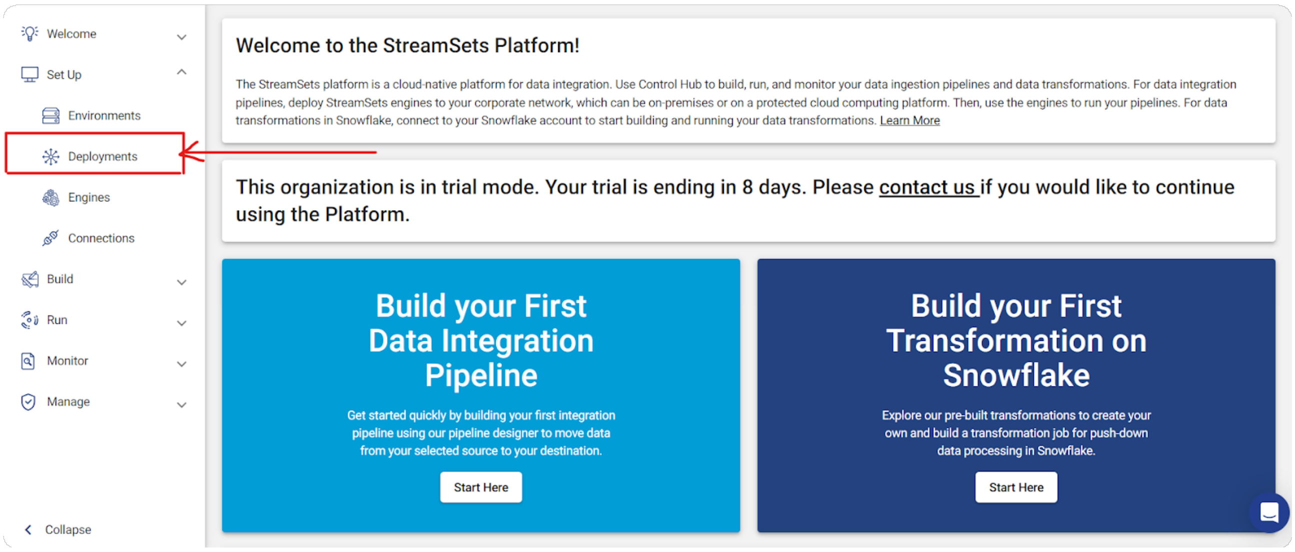
前提条件：

- お使いのマシンにDockerがインストールされていること（Dockerをダウンロードしてインストールしてください）。
- StreamSetsのアカウント。

StreamSets Data Collectorの展開

ステップ1： Data CollectorのDeployment設定

- ログイン後、サイドバーで「Deployment」を選択します。



Welcome to the StreamSets Platform!

The StreamSets platform is a cloud-native platform for data integration. Use Control Hub to build, run, and monitor your data ingestion pipelines and data transformations. For data integration pipelines, deploy StreamSets engines to your corporate network, which can be on-premises or on a protected cloud computing platform. Then, use the engines to run your pipelines. For data transformations in Snowflake, connect to your Snowflake account to start building and running your data transformations. Learn More

This organization is in trial mode. Your trial is ending in 8 days. Please contact us if you would like to continue using the Platform.

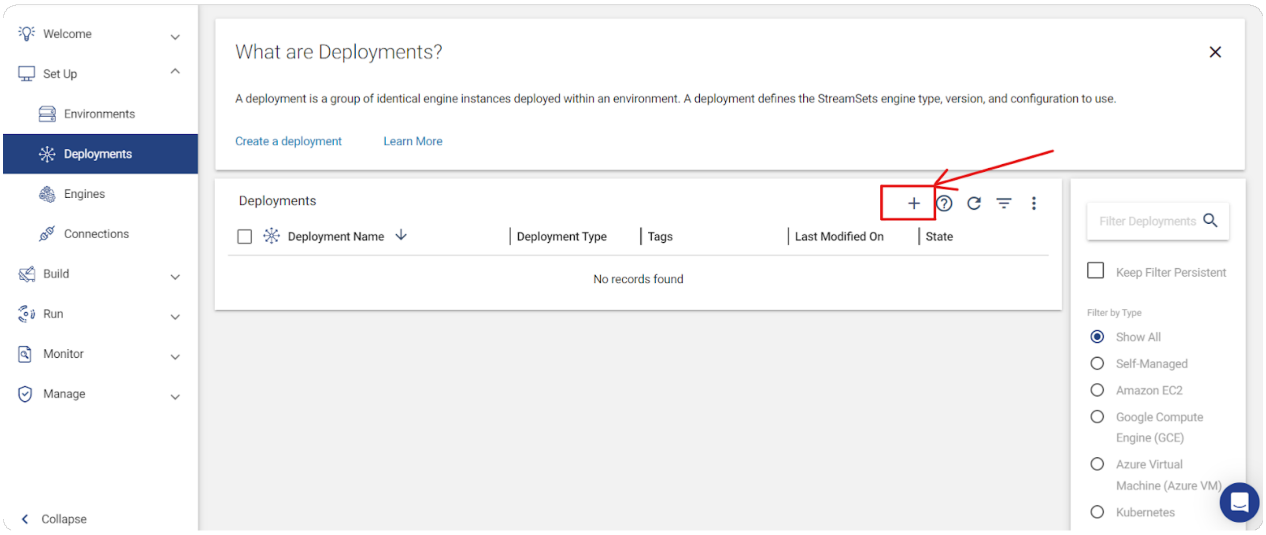
Build your First Data Integration Pipeline

Get started quickly by building your first integration pipeline using our pipeline designer to move data from your selected source to your destination.

Build your First Transformation on Snowflake

Explore our pre-built transformations to create your own and build a transformation job for push-down data processing in Snowflake.

- 以下のように「+」をクリックして新しい展開（Deployment）を開始します。



What are Deployments?

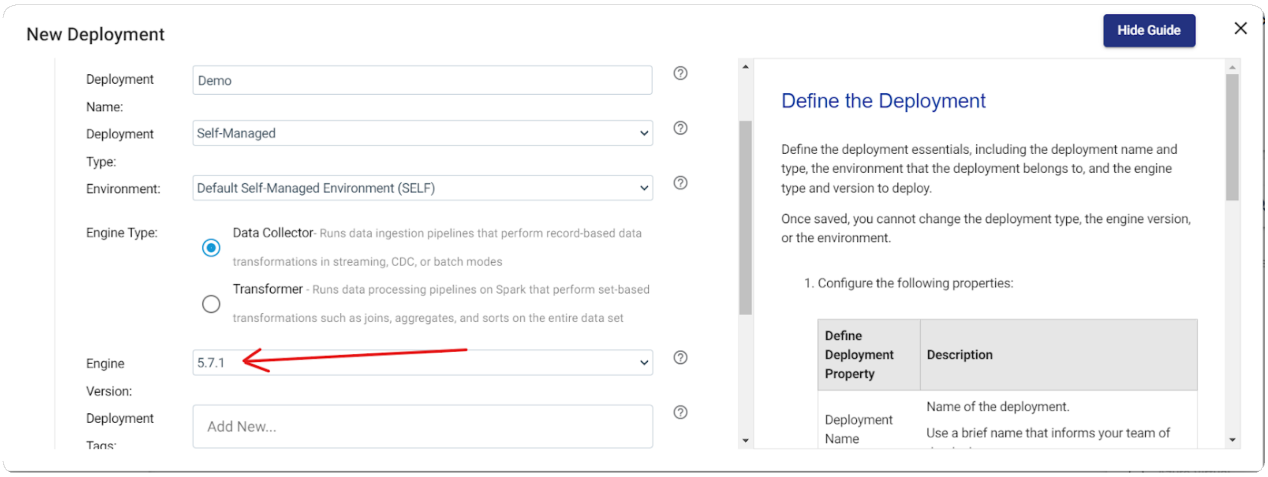
A deployment is a group of identical engine instances deployed within an environment. A deployment defines the StreamSets engine type, version, and configuration to use.

Deployments

Deployment Name	Deployment Type	Tags	Last Modified On	State

No records found

- 「Engine」で現在の安定版リリースであるバージョン5.7.1を選択し、ご使用の要件に従って定義された展開を完了します。「Save & Next」ボタンをクリックします。



New Deployment

Deployment Name: Demo

Deployment Type: Self-Managed

Environment: Default Self-Managed Environment (SELF)

Engine Type:
● Data Collector- Runs data ingestion pipelines that perform record-based data transformations in streaming, CDC, or batch modes
○ Transformer - Runs data processing pipelines on Spark that perform set-based transformations such as joins, aggregates, and sorts on the entire data set

Engine Version: 5.7.1

Deployment Tags: Add New...

Define the Deployment

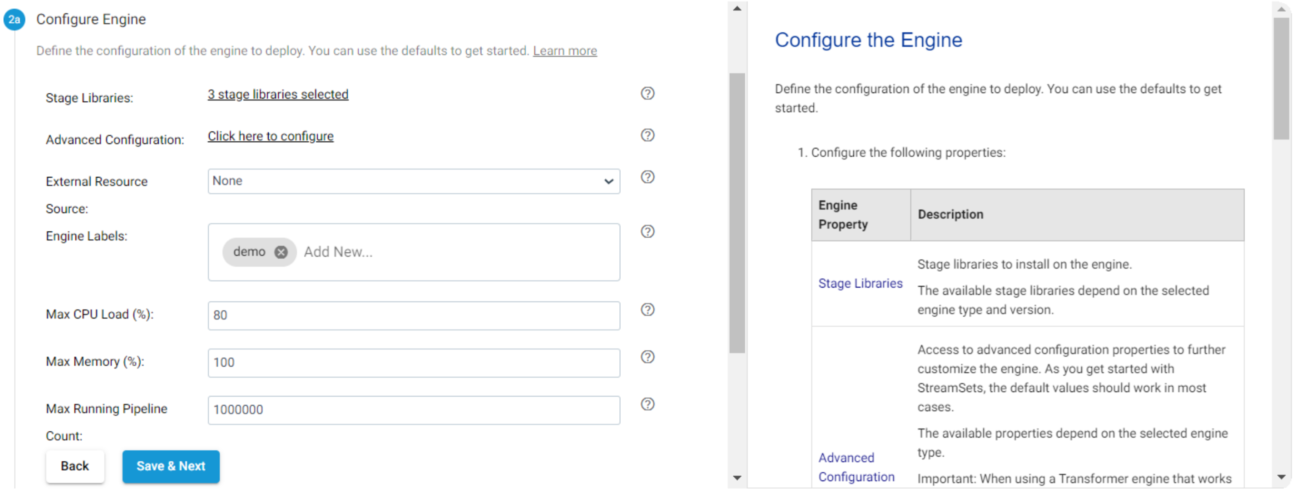
Define the deployment essentials, including the deployment name and type, the environment that the deployment belongs to, and the engine type and version to deploy.

Once saved, you cannot change the deployment type, the engine version, or the environment.

1. Configure the following properties:

Define Deployment Property	Description
Deployment Name	Name of the deployment. Use a brief name that informs your team of

- エンジンの設定をお使いの仕様に合わせ、「Save & Next」ボタンを選択します。



2a Configure Engine

Define the configuration of the engine to deploy. You can use the defaults to get started. Learn more

Stage Libraries: 3 stage libraries selected

Advanced Configuration: Click here to configure

External Resource Source: None

Engine Labels: demo

Max CPU Load (%): 80

Max Memory (%): 100

Max Running Pipeline Count: 1000000

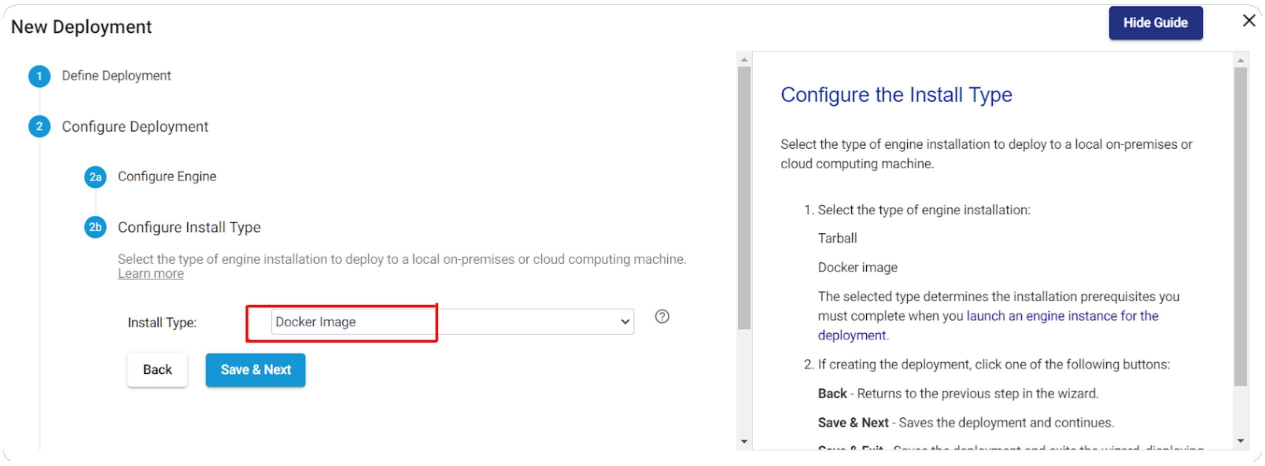
Configure the Engine

Define the configuration of the engine to deploy. You can use the defaults to get started.

1. Configure the following properties:

Engine Property	Description
Stage Libraries	Stage libraries to install on the engine. The available stage libraries depend on the selected engine type and version.
Advanced Configuration	Access to advanced configuration properties to further customize the engine. As you get started with StreamSets, the default values should work in most cases. The available properties depend on the selected engine type. Important: When using a Transformer engine that works

- インストールの種類として「Docker Image」を選択し、「Save & Next」ボタンをクリックします。



New Deployment

1 Define Deployment

2 Configure Deployment

2a Configure Engine

2b Configure Install Type

Select the type of engine installation to deploy to a local on-premises or cloud computing machine. Learn more

Install Type: Docker Image

Configure the Install Type

Select the type of engine installation to deploy to a local on-premises or cloud computing machine.

1. Select the type of engine installation:

   Tarball

   Docker image

   The selected type determines the installation prerequisites you must complete when you launch an engine instance for the deployment.

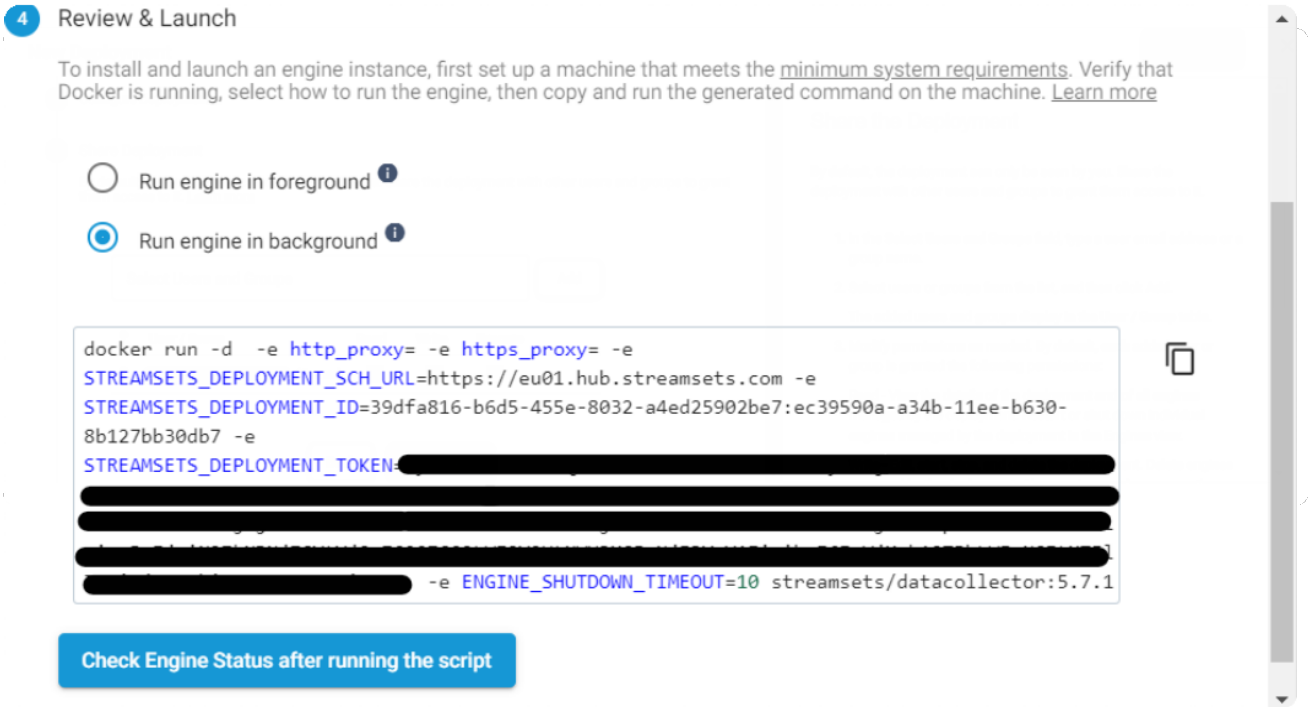
2. If creating the deployment, click one of the following buttons:

   Back - Returns to the previous step in the wizard.

   Save & Next - Saves the deployment and continues.

   Save & Exit - Saves the deployment and exits the wizard, displaying

- 展開の共有設定をカスタマイズして他のユーザーやグループにアクセスを許可します。「Save & Next」ボタンをクリックします。



4 Review & Launch

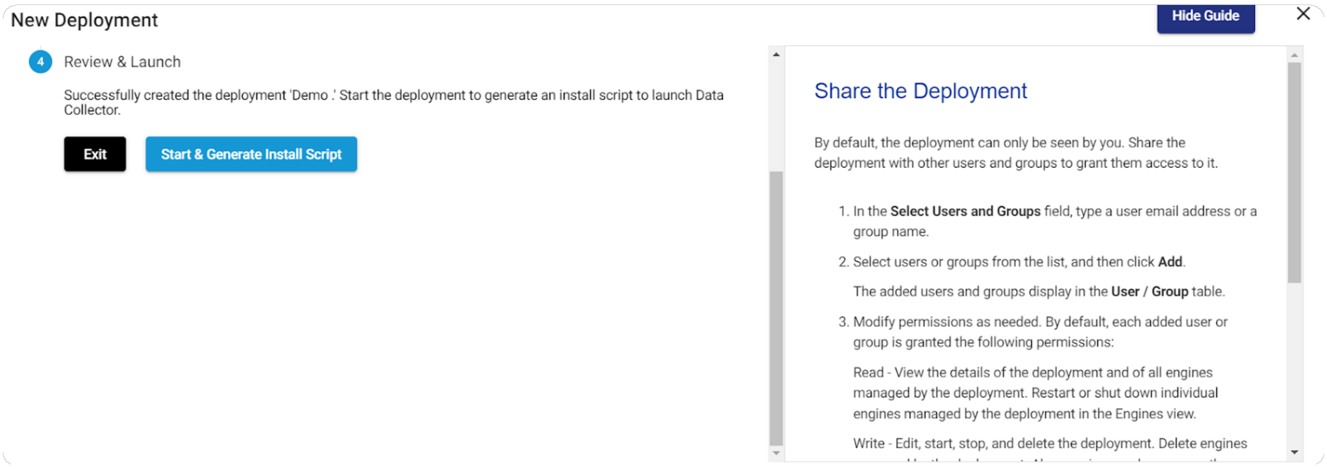
To install and launch an engine instance, first set up a machine that meets the minimum system requirements. Verify that Docker is running, select how to run the engine, then copy and run the generated command on the machine. Learn more

○ Run engine in foreground

● Run engine in background

```
docker run -d  -e http_proxy= -e https_proxy= -e
STREAMSETS_DEPLOYMENT_SCH_URL=https://eu01.hub.streamsets.com -e
STREAMSETS_DEPLOYMENT_ID=39dfa816-b6d5-455e-8032-a4ed25902be7:ec39590a-a34b-11ee-b630-
8b127bb30db7 -e
STREAMSETS_DEPLOYMENT_TOKEN=[redacted] -e ENGINE_SHUTDOWN_TIMEOUT=10 streamsets/datacollector:5.7.1
```

- 「Start & Generate Install Script 」ボタンをクリックします。



New Deployment

4 Review & Launch

Successfully created the deployment 'Demo .' Start the deployment to generate an install script to launch Data Collector.

Share the Deployment

By default, the deployment can only be seen by you. Share the deployment with other users and groups to grant them access to it.

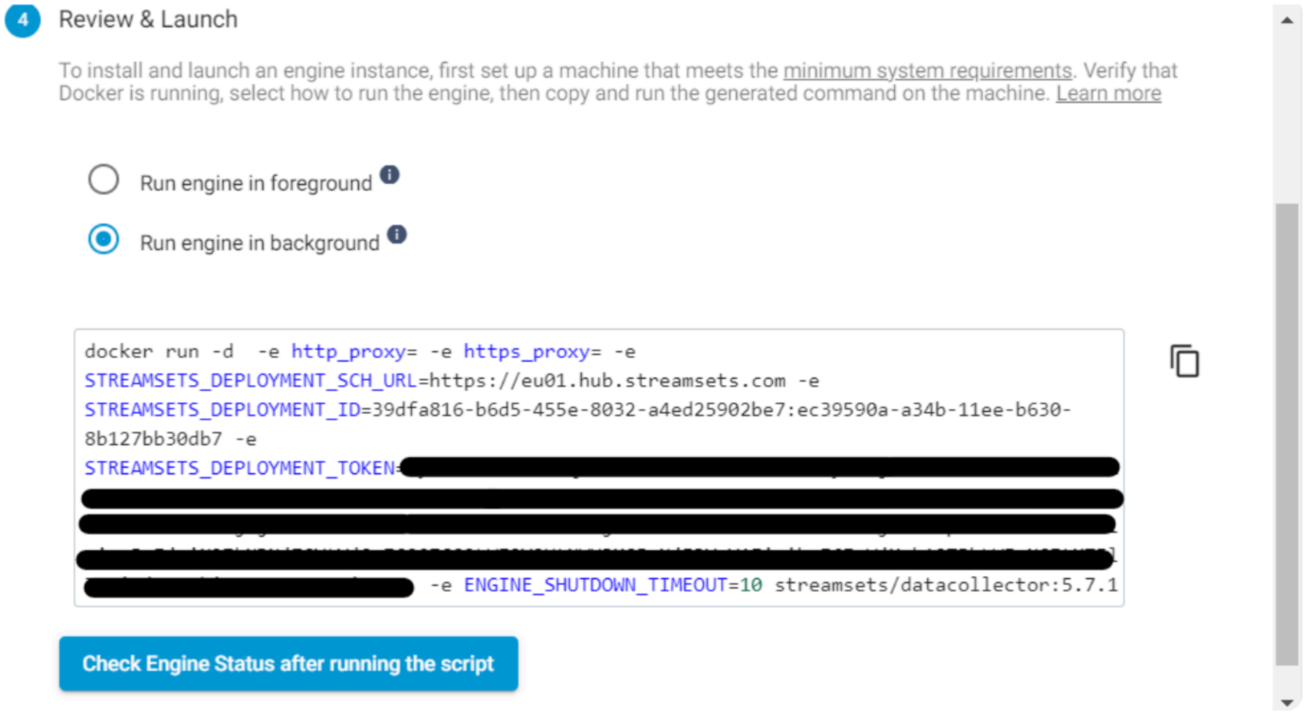
1. In the **Select Users and Groups** field, type a user email address or a group name.
2. Select users or groups from the list, and then click **Add**.

   The added users and groups display in the **User / Group** table.
3. Modify permissions as needed. By default, each added user or group is granted the following permissions:

   Read - View the details of the deployment and of all engines managed by the deployment. Restart or shut down individual engines managed by the deployment in the Engines view.

   Write - Edit, start, stop, and delete the deployment. Delete engines

- Dockerが起動していることを確認してから、以下のコマンドをコピーしてください。このコマンドをwindows/Ubuntu/macのターミナルに貼り付けると、エンジンが起動します。



4 Review & Launch

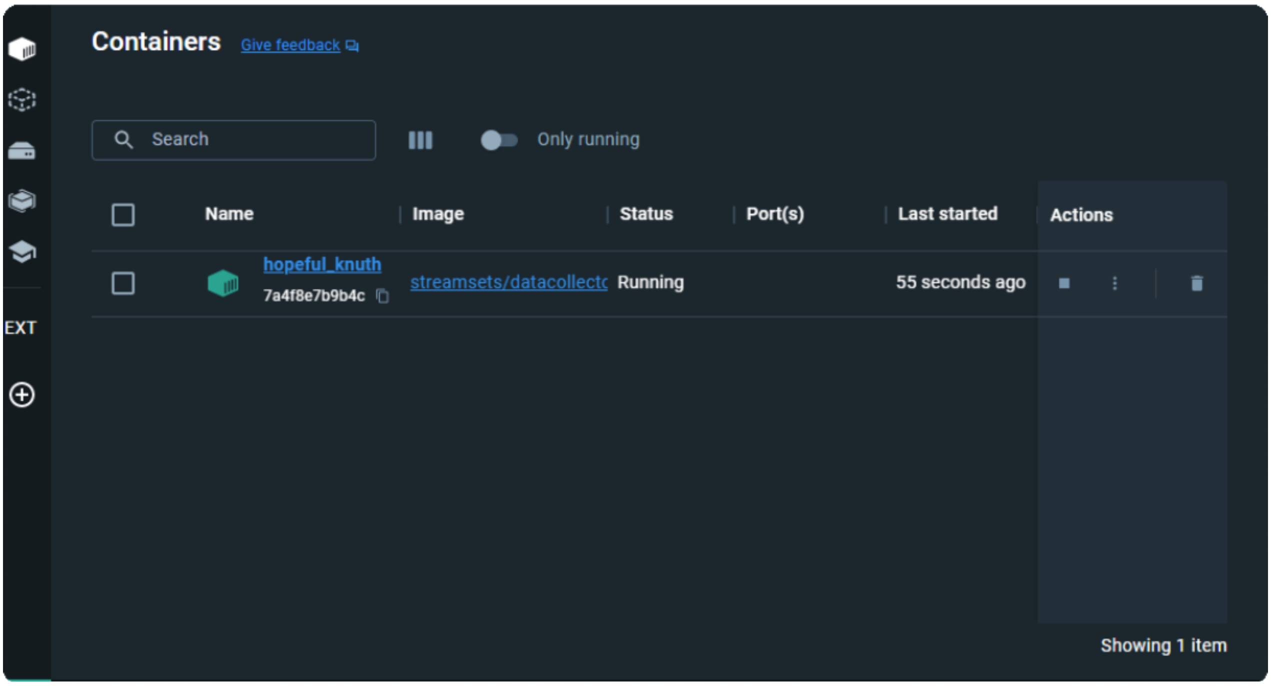
To install and launch an engine instance, first set up a machine that meets the minimum system requirements. Verify that Docker is running, select how to run the engine, then copy and run the generated command on the machine. Learn more

○ Run engine in foreground

● Run engine in background

```
docker run -d  -e http_proxy= -e https_proxy= -e
STREAMSETS_DEPLOYMENT_SCH_URL=https://eu01.hub.streamsets.com -e
STREAMSETS_DEPLOYMENT_ID=39dfa816-b6d5-455e-8032-a4ed25902be7:ec39590a-a34b-11ee-b630-
8b127bb30db7 -e
STREAMSETS_DEPLOYMENT_TOKEN=[redacted] -e ENGINE_SHUTDOWN_TIMEOUT=10 streamsets/datacollector:5.7.1
```

- コンテナがアクティブになっていることを確認します。以下のスクリーンショットの通り、Docker Desktop内で実行中のコンテナを確認してください。



Containers Give feedback

Search | Only running

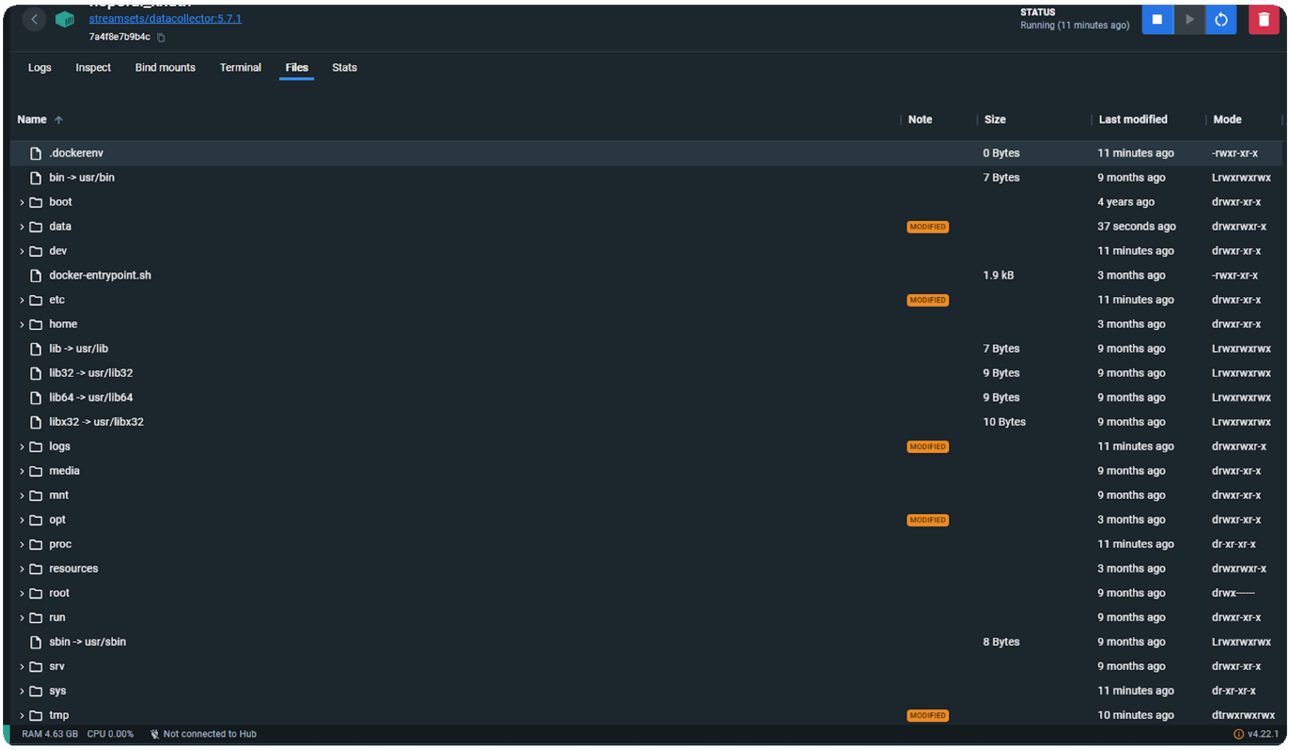
□	Name	Image	Status	Port(s)	Last started	Actions
□	hopeful_knuth 7a4f8e7b9b4c	streamsets/datacollectc	Running		55 seconds ago	■ ⋮ 🗑

Showing 1 item

ステップ2： Data Collectorの詳細設定

WhoisXML APIのWebSocketを活用するために、データコレクターの設定をカスタマイズしてください。

- ビルドされたjarファイルはこちらからダウンロードできます。ダウンロード後、Dockerコンテナ内の既存のファイルを新しく取得したファイルで置き換えます。ファイルのシームレスな置き換えには、Docker Desktopが効率的で便利です。

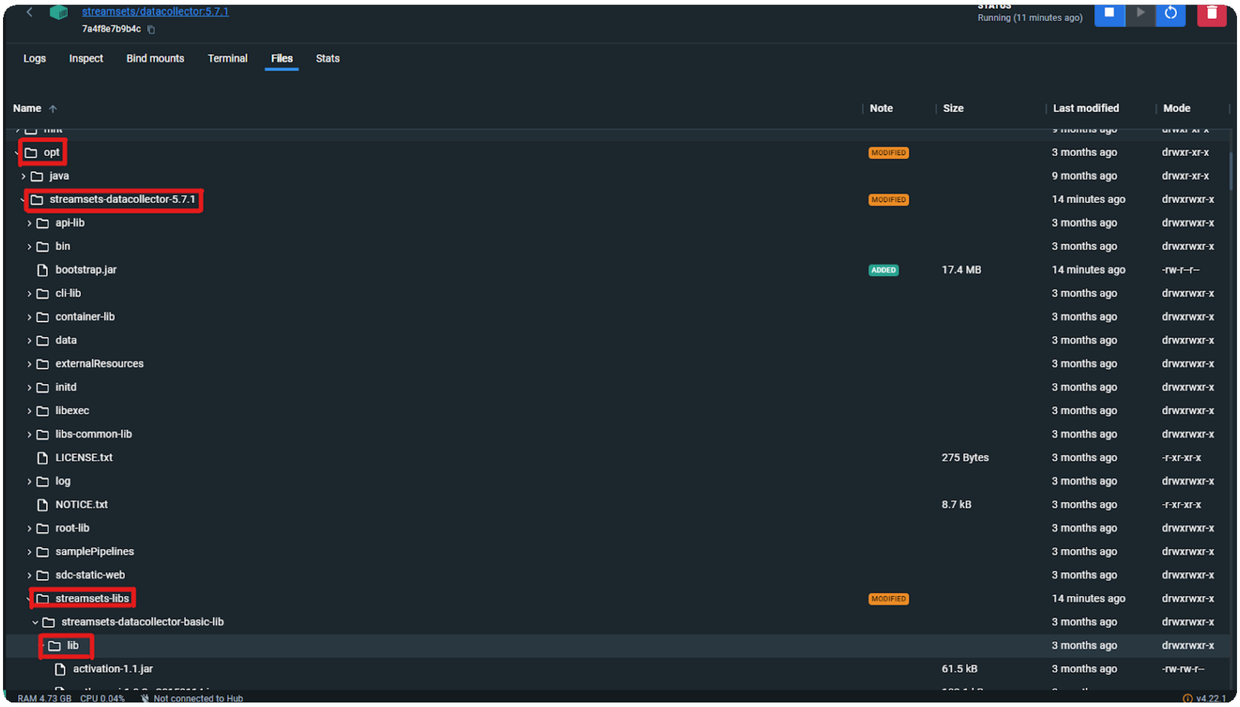


streamsets/datacollector:5.7.1
7a4f8e7b9b4c

Name	Note	Size	Last modified	Mode
.dockerenv		0 Bytes	11 minutes ago	-rwxr-xr-x
bin -> usr/bin		7 Bytes	9 months ago	Lrwxrwxrwx
boot			4 years ago	drwxr-xr-x
data	MODIFIED		37 seconds ago	drwxrwxr-x
dev			11 minutes ago	drwxr-xr-x
docker-entrypoint.sh		1.9 kB	3 months ago	-rwxr-xr-x
etc	MODIFIED		11 minutes ago	drwxr-xr-x
home			3 months ago	drwxr-xr-x
lib -> usr/lib		7 Bytes	9 months ago	Lrwxrwxrwx
lib32 -> usr/lib32		9 Bytes	9 months ago	Lrwxrwxrwx
lib64 -> usr/lib64		9 Bytes	9 months ago	Lrwxrwxrwx
libx32 -> usr/libx32		10 Bytes	9 months ago	Lrwxrwxrwx
logs	MODIFIED		11 minutes ago	drwxrwxr-x
media			9 months ago	drwxr-xr-x
mnt			9 months ago	drwxr-xr-x
opt	MODIFIED		3 months ago	drwxr-xr-x
proc			11 minutes ago	dr-xr-xr-x
resources			3 months ago	drwxrwxr-x
root			9 months ago	drwx------
run			9 months ago	drwxr-xr-x
sbin -> usr/sbin		8 Bytes	9 months ago	Lrwxrwxrwx
srv			9 months ago	drwxr-xr-x
sys			11 minutes ago	dr-xr-xr-x
tmp	MODIFIED		10 minutes ago	dtrwxrwxrwx

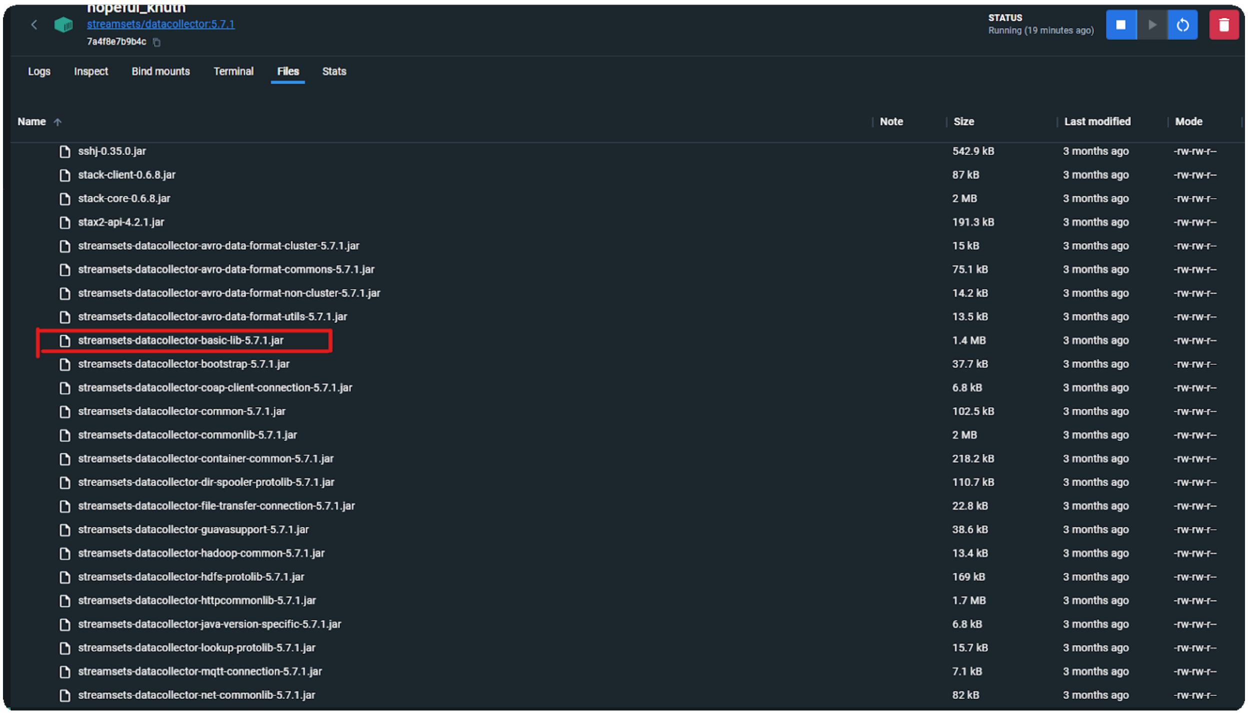
RAM 4.63 GB CPU 0.00% Not connected to Hub v4.22.1

- ディレクトリ /opt/streamsets-datacollector-5.7.1/streamsets-libs/streamsets-datacollector-basic-lib/lib に移動します。



Name	Note	Size	Last modified	Mode
opt	MODIFIED		3 months ago	drwxr-xr-x
java			9 months ago	drwxr-xr-x
streamsets-datacollector-5.7.1	MODIFIED		14 minutes ago	drwxrwxr-x
api-lib			3 months ago	drwxrwxr-x
bin			3 months ago	drwxrwxr-x
bootstrap.jar	ADDED	17.4 MB	14 minutes ago	-rw-r--r--
cli-lib			3 months ago	drwxrwxr-x
container-lib			3 months ago	drwxrwxr-x
data			3 months ago	drwxrwxr-x
externalResources			3 months ago	drwxrwxr-x
initd			3 months ago	drwxrwxr-x
libexec			3 months ago	drwxrwxr-x
libs-common-lib			3 months ago	drwxrwxr-x
LICENSE.txt		275 Bytes	3 months ago	-r-xr-xr-x
log			3 months ago	drwxrwxr-x
NOTICE.txt		8.7 kB	3 months ago	-r-xr-xr-x
root-lib			3 months ago	drwxrwxr-x
samplePipelines			3 months ago	drwxrwxr-x
sdc-static-web			3 months ago	drwxrwxr-x
streamsets-libs	MODIFIED		14 minutes ago	drwxrwxr-x
streamsets-datacollector-basic-lib			3 months ago	drwxrwxr-x
lib			3 months ago	drwxrwxr-x
activation-1.1.jar		61.5 kB	3 months ago	-rw-rw-r--

- このディレクトリに、「streamsets-datacollector-basic-lib-5.7.1.jar」というjarファイルがあるはずです。



Name	Note	Size	Last modified	Mode
sshj-0.35.0.jar		542.9 kB	3 months ago	-rw-rw-r--
stack-client-0.6.8.jar		87 kB	3 months ago	-rw-rw-r--
stack-core-0.6.8.jar		2 MB	3 months ago	-rw-rw-r--
stax2-api-4.2.1.jar		191.3 kB	3 months ago	-rw-rw-r--
streamsets-datacollector-avro-data-format-cluster-5.7.1.jar		15 kB	3 months ago	-rw-rw-r--
streamsets-datacollector-avro-data-format-commons-5.7.1.jar		75.1 kB	3 months ago	-rw-rw-r--
streamsets-datacollector-avro-data-format-non-cluster-5.7.1.jar		14.2 kB	3 months ago	-rw-rw-r--
streamsets-datacollector-avro-data-format-utils-5.7.1.jar		13.5 kB	3 months ago	-rw-rw-r--
streamsets-datacollector-basic-lib-5.7.1.jar		1.4 MB	3 months ago	-rw-rw-r--
streamsets-datacollector-bootstrap-5.7.1.jar		37.7 kB	3 months ago	-rw-rw-r--
streamsets-datacollector-coap-client-connection-5.7.1.jar		6.8 kB	3 months ago	-rw-rw-r--
streamsets-datacollector-common-5.7.1.jar		102.5 kB	3 months ago	-rw-rw-r--
streamsets-datacollector-commonlib-5.7.1.jar		2 MB	3 months ago	-rw-rw-r--
streamsets-datacollector-container-common-5.7.1.jar		218.2 kB	3 months ago	-rw-rw-r--
streamsets-datacollector-dir-spooler-protolib-5.7.1.jar		110.7 kB	3 months ago	-rw-rw-r--
streamsets-datacollector-file-transfer-connection-5.7.1.jar		22.8 kB	3 months ago	-rw-rw-r--
streamsets-datacollector-guavasupport-5.7.1.jar		38.6 kB	3 months ago	-rw-rw-r--
streamsets-datacollector-hadoop-common-5.7.1.jar		13.4 kB	3 months ago	-rw-rw-r--
streamsets-datacollector-hdfs-protolib-5.7.1.jar		169 kB	3 months ago	-rw-rw-r--
streamsets-datacollector-httpcommonlib-5.7.1.jar		1.7 MB	3 months ago	-rw-rw-r--
streamsets-datacollector-java-version-specific-5.7.1.jar		6.8 kB	3 months ago	-rw-rw-r--
streamsets-datacollector-lookup-protolib-5.7.1.jar		15.7 kB	3 months ago	-rw-rw-r--
streamsets-datacollector-mqtt-connection-5.7.1.jar		7.1 kB	3 months ago	-rw-rw-r--
streamsets-datacollector-net-commonlib-5.7.1.jar		82 kB	3 months ago	-rw-rw-r--

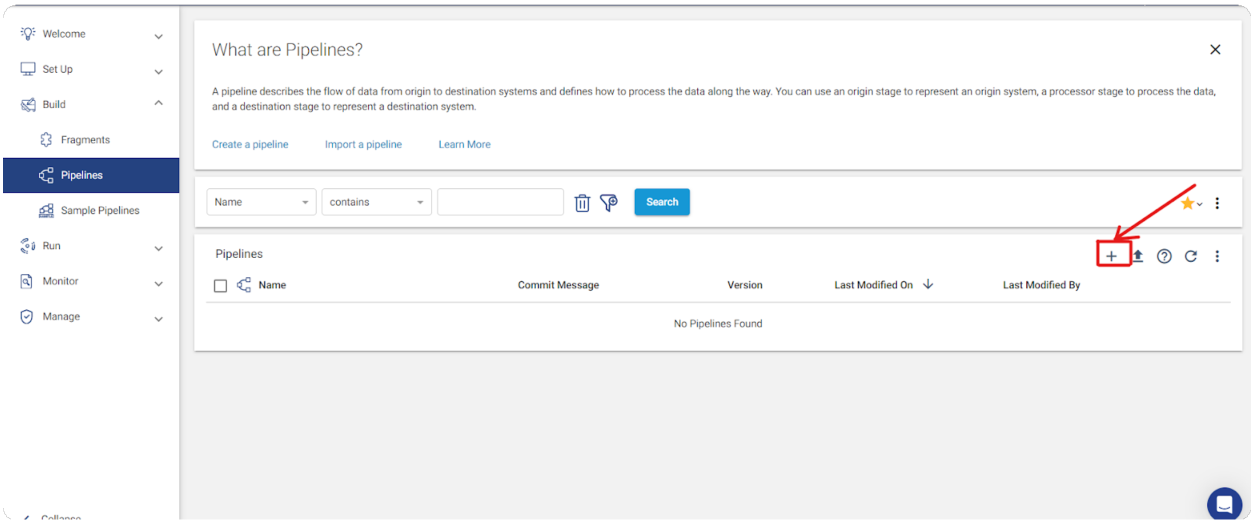
- 既存のファイルをダウンロードしたファイルで置き換えます。または、ダウンロードしたファイルをこの場所にドラッグすれば、簡単に直接置き換えることができます。

- Data Collector/Containerを再起動します。

パイプラインの作成

ステップ3：パイプラインのセットアップ

- StreamSets UIでサイドバーから「Build」を選択し、次に「Pipelines」を選択します。以下のように「+」ボタンをクリックしてパイプラインを開始します。



What are Pipelines?

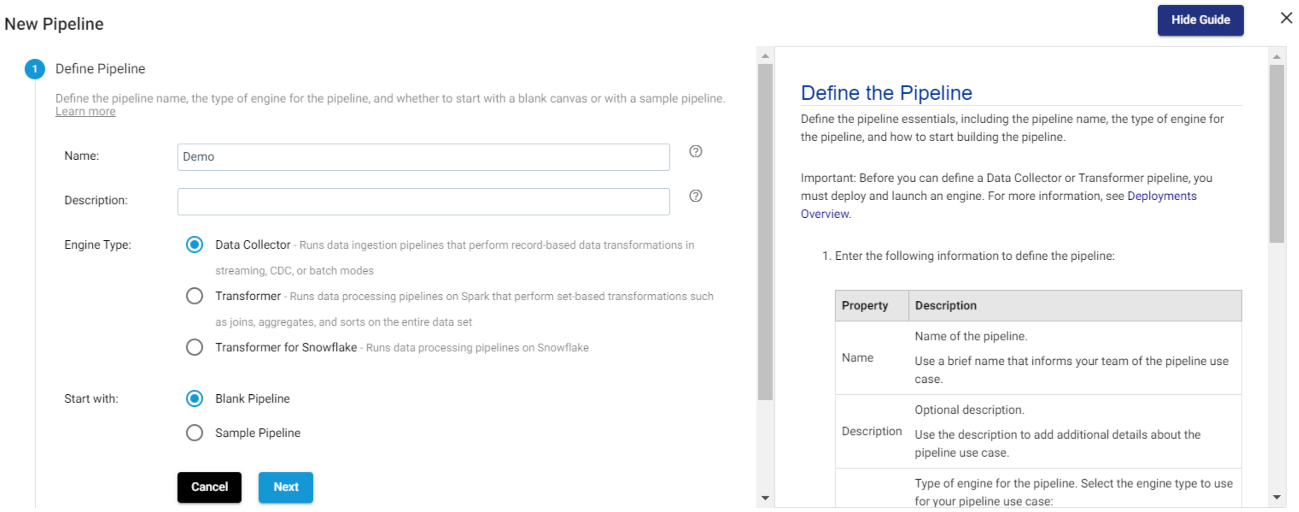
A pipeline describes the flow of data from origin to destination systems and defines how to process the data along the way. You can use an origin stage to represent an origin system, a processor stage to process the data, and a destination stage to represent a destination system.

Create a pipeline　Import a pipeline　Learn More

Pipelines

□ Name	Commit Message	Version	Last Modified On	Last Modified By
No Pipelines Found				

- 新しいパイプラインをお使いの仕様に合わせてカスタマイズし、「Next」ボタンをクリックして進みます。



New Pipeline

1 Define Pipeline

Define the pipeline name, the type of engine for the pipeline, and whether to start with a blank canvas or with a sample pipeline. Learn more

Name: Demo

Description:

Engine Type:
● Data Collector - Runs data ingestion pipelines that perform record-based data transformations in streaming, CDC, or batch modes
○ Transformer - Runs data processing pipelines on Spark that perform set-based transformations such as joins, aggregates, and sorts on the entire data set
○ Transformer for Snowflake - Runs data processing pipelines on Snowflake

Start with:
● Blank Pipeline
○ Sample Pipeline

Define the Pipeline

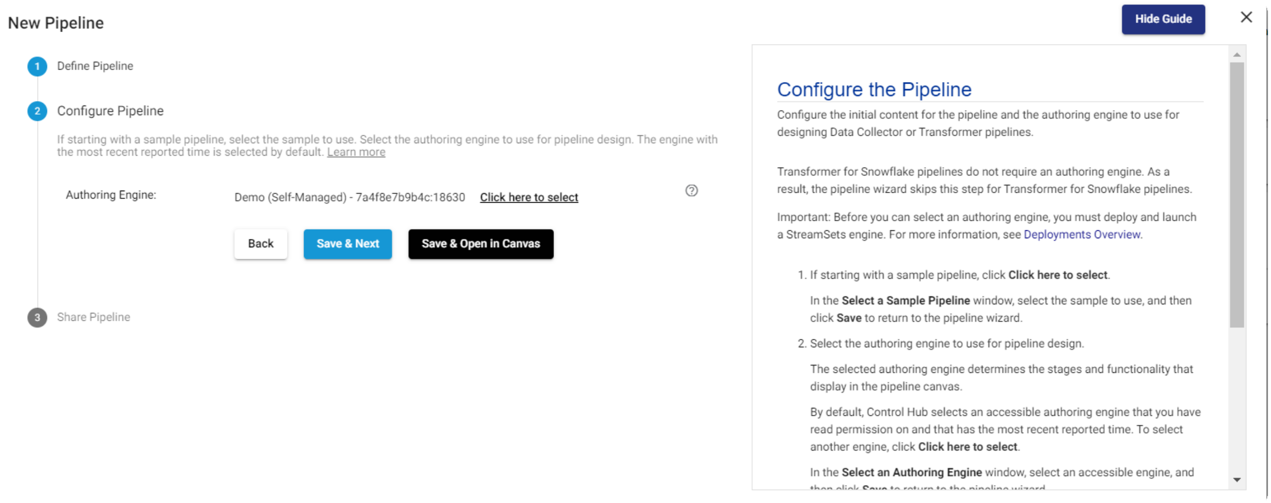
Define the pipeline essentials, including the pipeline name, the type of engine for the pipeline, and how to start building the pipeline.

Important: Before you can define a Data Collector or Transformer pipeline, you must deploy and launch an engine. For more information, see Deployments Overview.

1. Enter the following information to define the pipeline:

Property	Description
Name	Name of the pipeline. Use a brief name that informs your team of the pipeline use case.
Description	Optional description. Use the description to add additional details about the pipeline use case.
	Type of engine for the pipeline. Select the engine type to use for your pipeline use case:

- パイプラインの設定を調整して指定のデータコレクタを選択し、「Save & Open in Canvas」ボタンをクリックします。



New Pipeline

1 Define Pipeline

2 Configure Pipeline

If starting with a sample pipeline, select the sample to use. Select the authoring engine to use for pipeline design. The engine with the most recent reported time is selected by default. Learn more

Authoring Engine: Demo (Self-Managed) - 7a4f8e7b9b4c:18630 Click here to select

Configure the Pipeline

Configure the initial content for the pipeline and the authoring engine to use for designing Data Collector or Transformer pipelines.

Transformer for Snowflake pipelines do not require an authoring engine. As a result, the pipeline wizard skips this step for Transformer for Snowflake pipelines.

Important: Before you can select an authoring engine, you must deploy and launch a StreamSets engine. For more information, see Deployments Overview.

1. If starting with a sample pipeline, click **Click here to select**.

   In the **Select a Sample Pipeline** window, select the sample to use, and then click **Save** to return to the pipeline wizard.

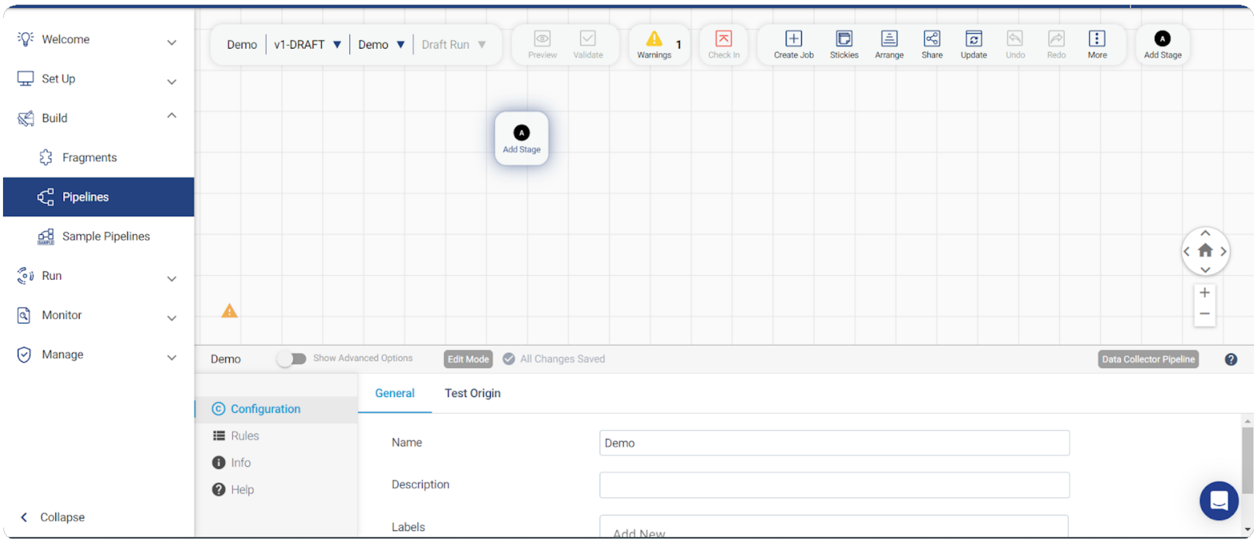
2. Select the authoring engine to use for pipeline design.

   The selected authoring engine determines the stages and functionality that display in the pipeline canvas.

   By default, Control Hub selects an accessible authoring engine that you have read permission on and that has the most recent reported time. To select another engine, click **Click here to select**.

   In the **Select an Authoring Engine** window, select an accessible engine, and then click **Save** to return to the pipeline wizard.

- 以下のようなユーザーインターフェース（UI）が表示されます。

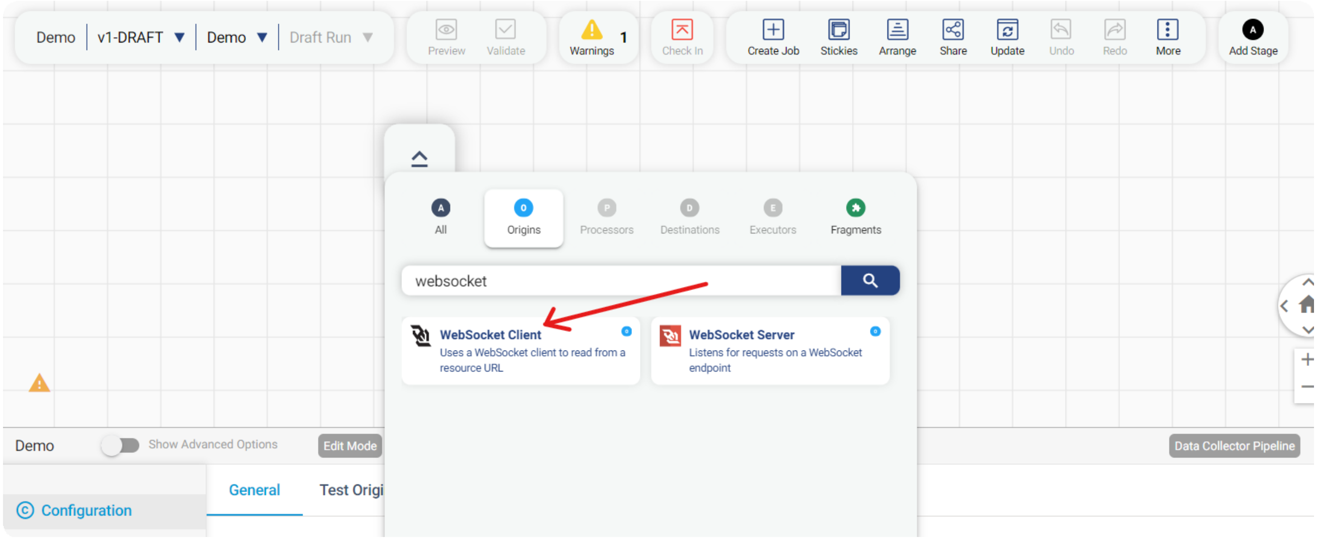


Name: Demo

Description:

Labels: Add New

- 「Add Stage」ボタンをクリックし、「WebSocket」を検索して「WebSocket Client」を選択します。



websocket

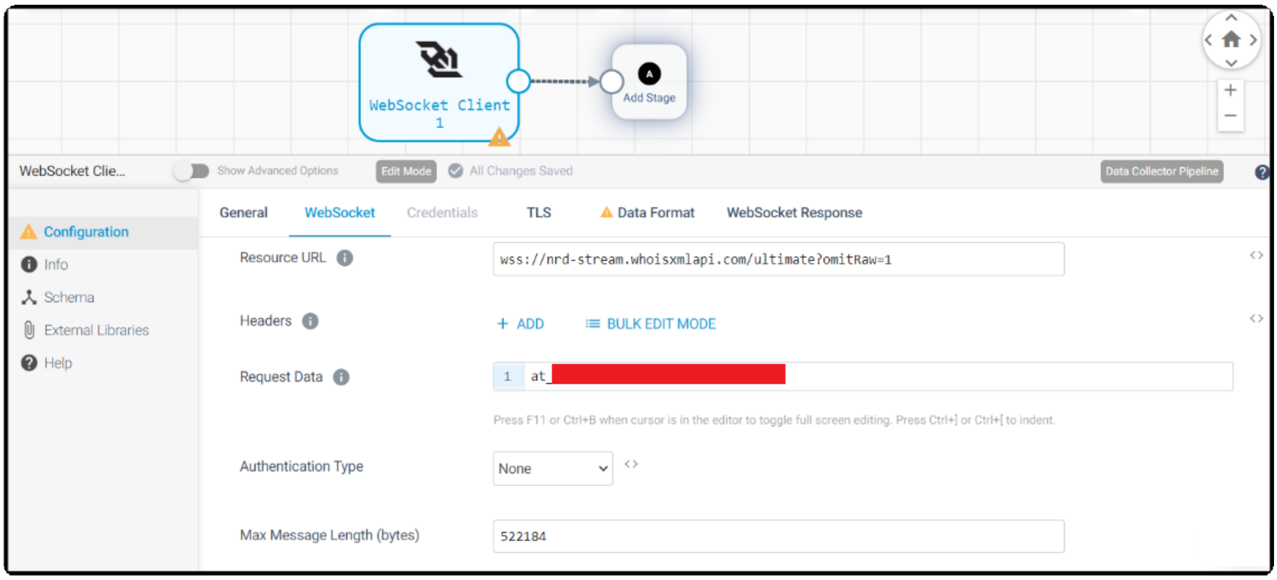
WebSocket Client
Uses a WebSocket client to read from a resource URL

WebSocket Server
Listens for requests on a WebSocket endpoint

WebSocketステージを選択し、適宜設定します：

WebSocketの設定

- Resource URLを入力
- Request Dataを入力（WHOIS APIキーが入ります）
- Max Message Length (bytes) として最小522184 を入力



Resource URL: wss://nrd-stream.whoisxmlapi.com/ultimate?omitRaw=1

Headers

Request Data: at[redacted]

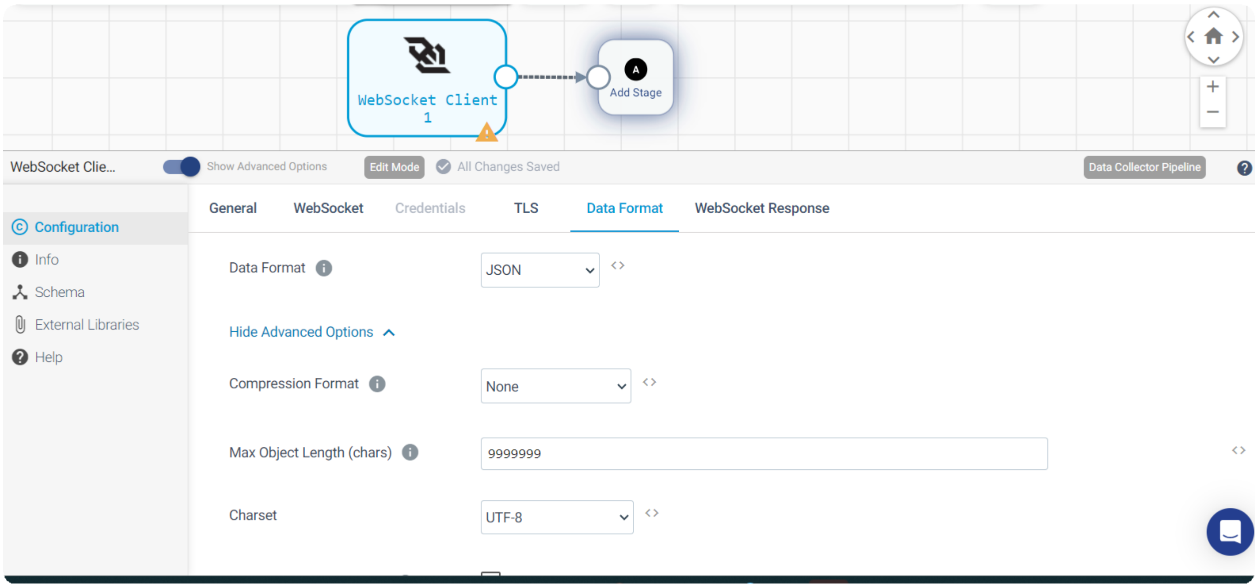
Press F11 or Ctrl+B when cursor is in the editor to toggle full screen editing. Press Ctrl+] or Ctrl+[to indent.

Authentication Type: None

Max Message Length (bytes): 522184

Data Formatの設定

- Data FormatとしてJSONを選択
- Max Object Length (chars) として9999999 を入力（必要に応じて変更できます）



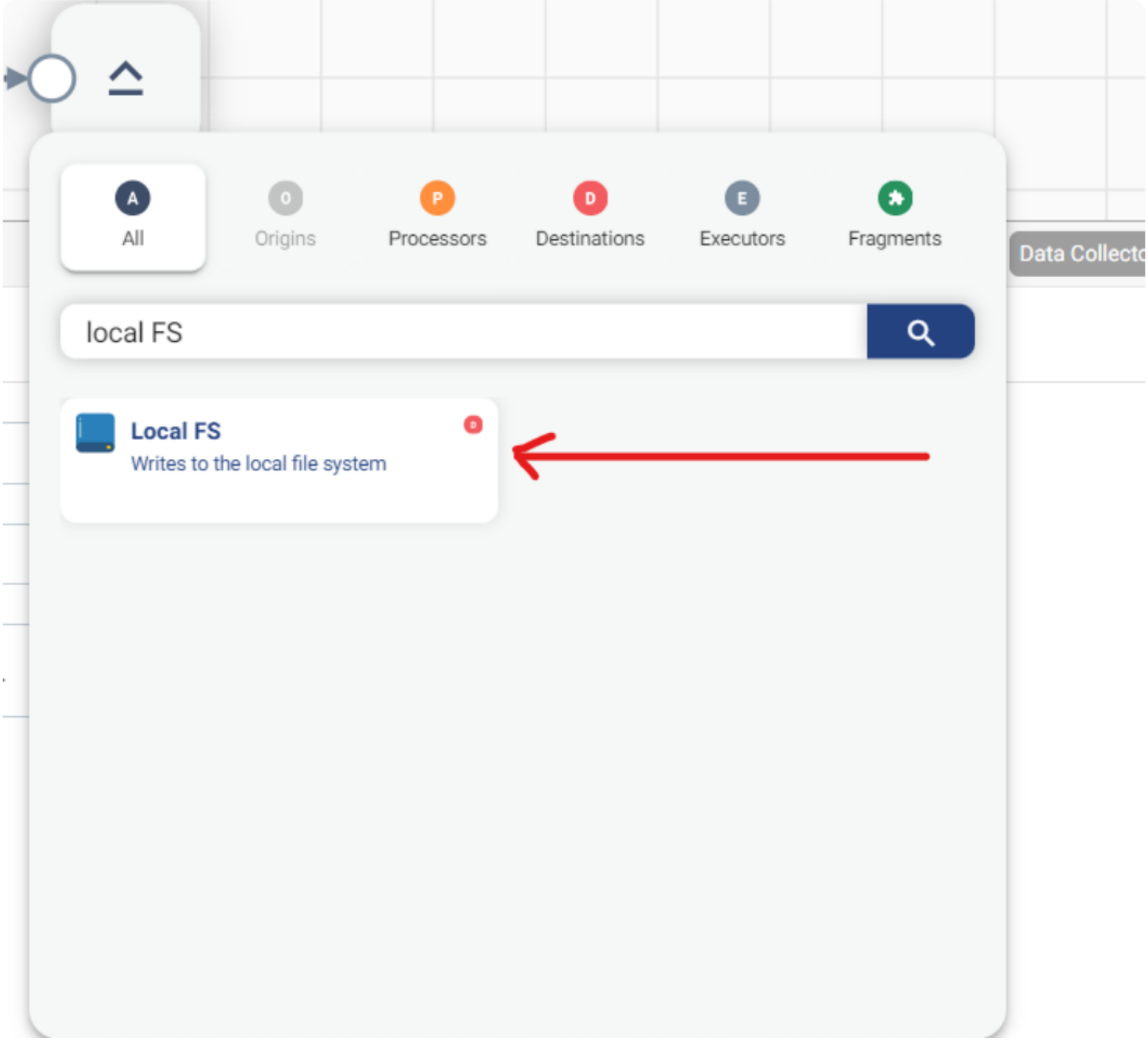
Data Format: JSON

Compression Format: None

Max Object Length (chars): 9999999

Charset: UTF-8

- UIの「Add Stage」ボタンでステージを追加し、「Local FS」ステージを選択します。



local FS

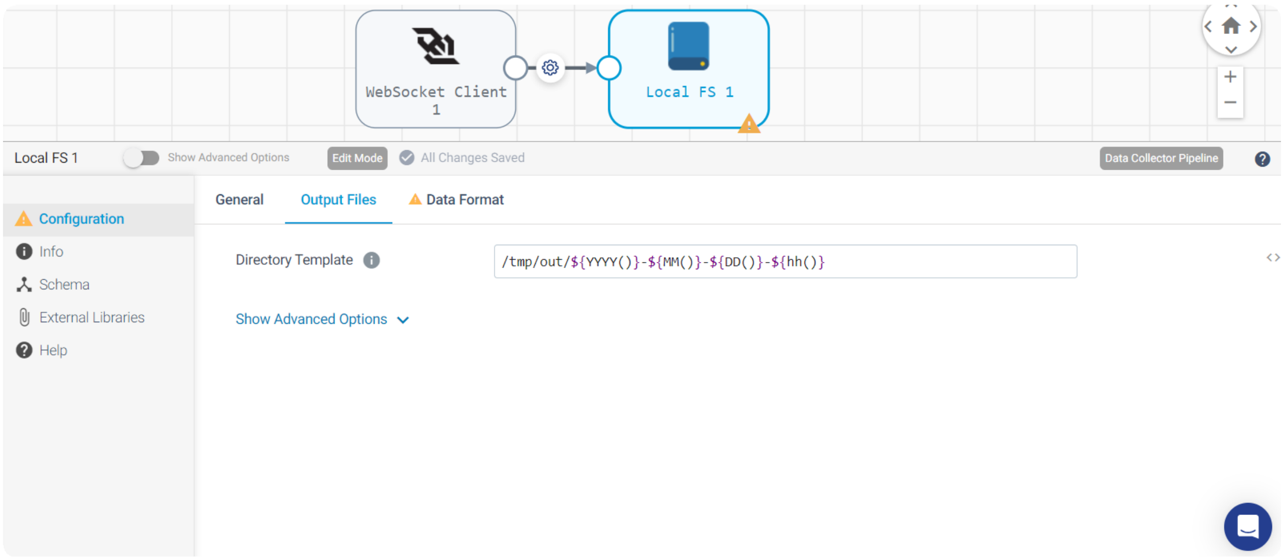
Local FS

Writes to the local file system

Local FSの設定

「Local FS」ステージを選択し、要件に従って設定します。

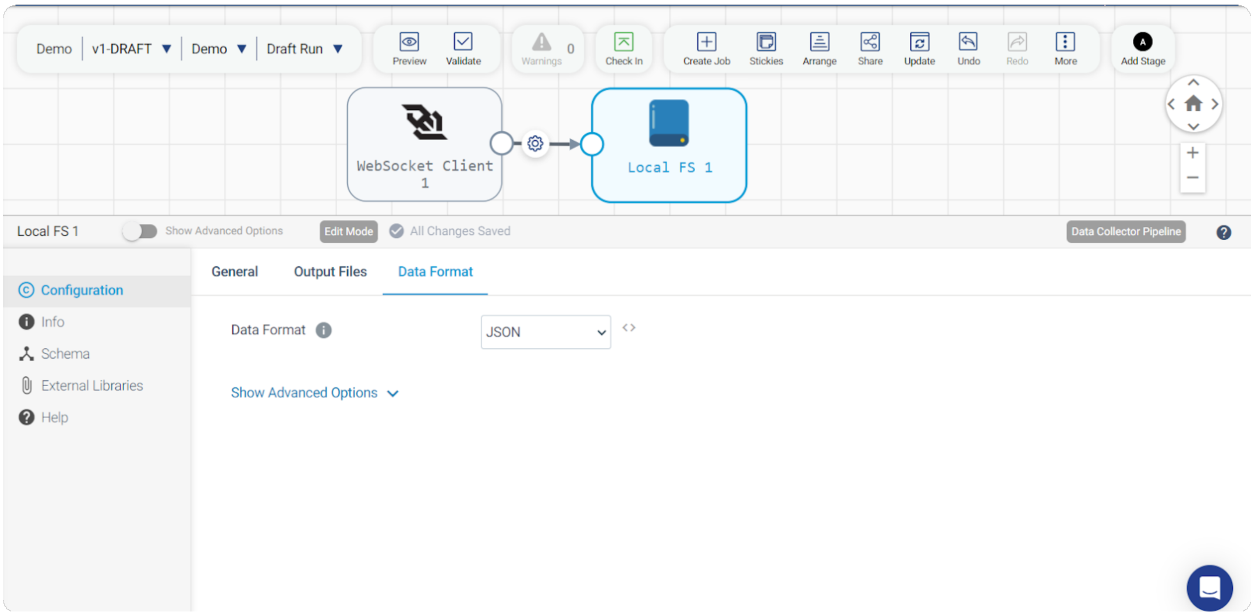
- 希望する出力ファイルの場所を「Directory Template」に入力します。



Directory Template: /tmp/out/${YYYY()}-${MM()}-${DD()}-${hh()}

要件に応じて「Data Format」に必要な設定を入力します。

- 「Data Format」としてJSONを選択します。

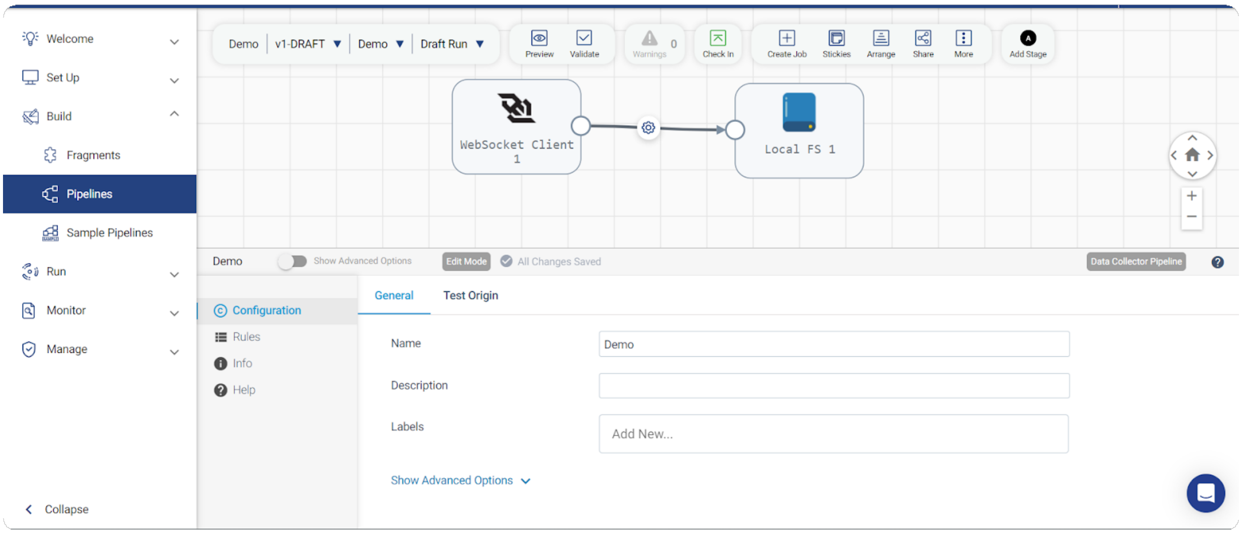


Data Format: JSON

設定後、「Validate」ボタンをクリックしてパイプラインを検証し、エラーを特定して修正しま

す。パイプラインの最終状態は、以下の例のようになります。

最終ステップ



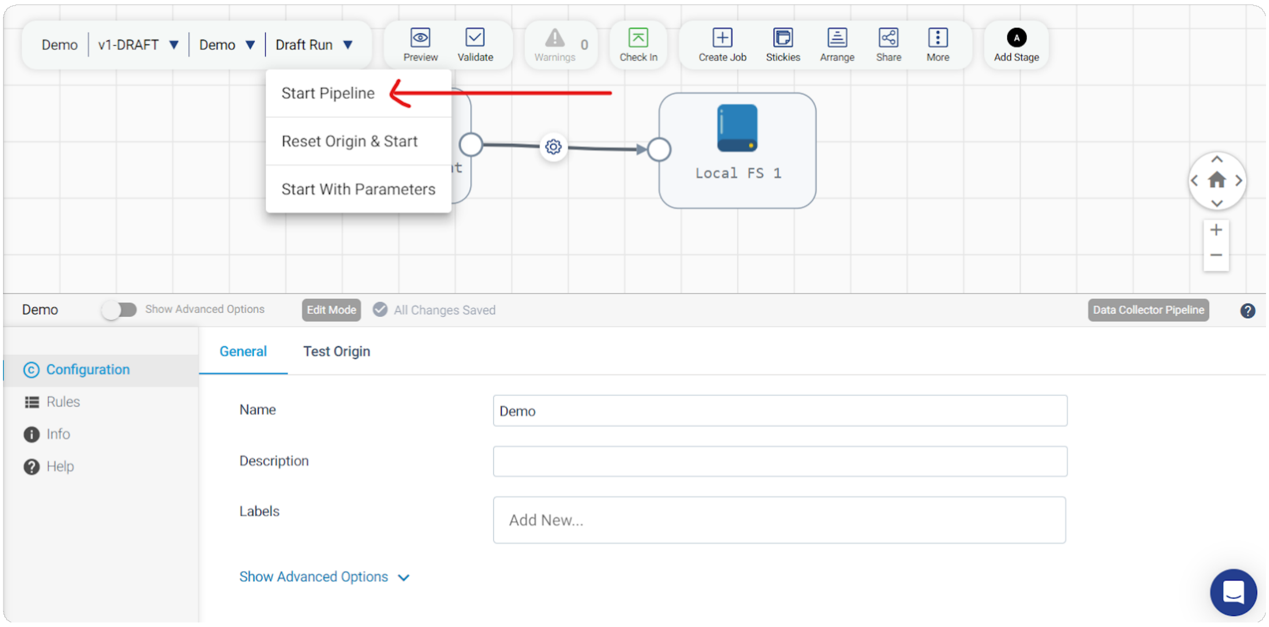
Name: Demo

Description:

Labels: Add New...

パイプラインの実行

UIで「Draft Run」、「Start Pipeline」を選択してパイプラインを実行します。

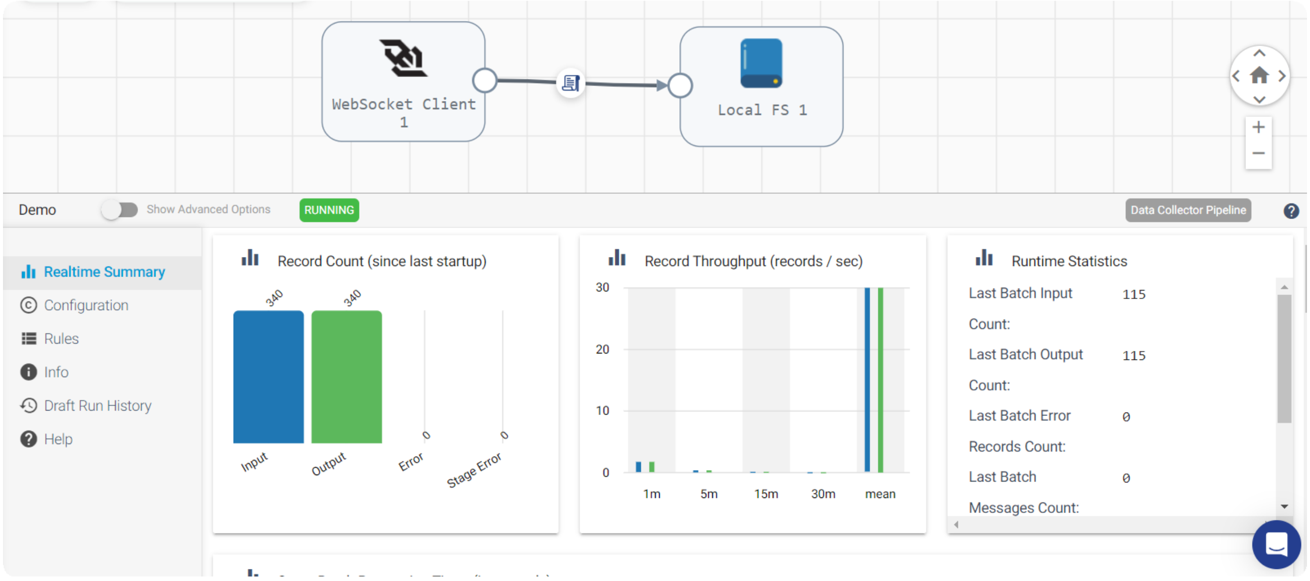


Name: Demo

Description:

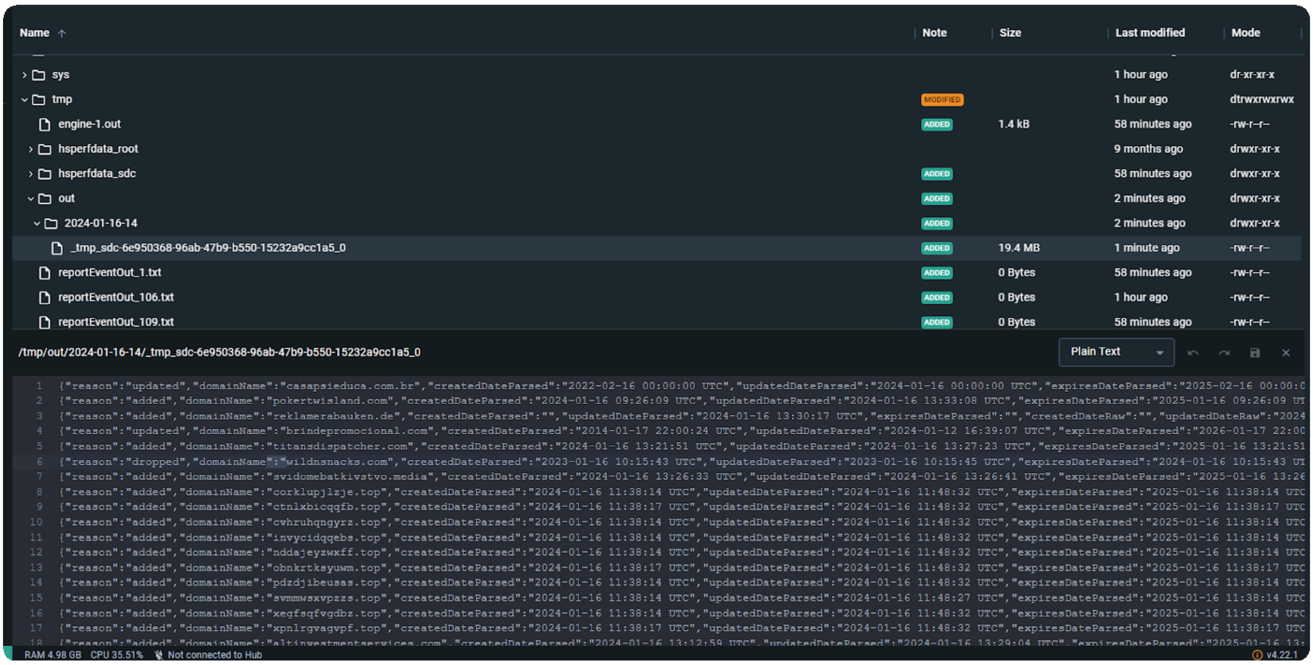
Labels: Add New...

上記の手順が完了した後、パイプラインを起動すると、下図のようなUIが表示されます。



Last Batch Input Count:	115
Last Batch Output Count:	115
Last Batch Error Records Count:	0
Last Batch Messages Count:	0

Docker コンテナ内では、WebSocket から取得したデータを含むファイルが指定したディレクトリに作成されていることを確認できます。



Name ↑	Note	Size	Last modified	Mode
› sys			1 hour ago	dr-xr-xr-x
⌄ tmp	MODIFIED		1 hour ago	dtrwxrwxrwx
engine-1.out	ADDED	1.4 kB	58 minutes ago	-rw-r--r--
› hsperfdata_root			9 months ago	drwxr-xr-x
› hsperfdata_sdc	ADDED		58 minutes ago	drwxr-xr-x
⌄ out	ADDED		2 minutes ago	drwxr-xr-x
⌄ 2024-01-16-14	ADDED		2 minutes ago	drwxr-xr-x
_tmp_sdc-6e950368-96ab-47b9-b550-15232a9cc1a5_0	ADDED	19.4 MB	1 minute ago	-rw-r--r--
reportEventOut_1.txt	ADDED	0 Bytes	58 minutes ago	-rw-r--r--
reportEventOut_106.txt	ADDED	0 Bytes	1 hour ago	-rw-r--r--
reportEventOut_109.txt	ADDED	0 Bytes	58 minutes ago	-rw-r--r--

/tmp/out/2024-01-16-14/_tmp_sdc-6e950368-96ab-47b9-b550-15232a9cc1a5_0

```
{"reason":"updated","domainName":"casapsieduca.com.br","createdDateParsed":"2022-02-16 00:00:00 UTC","updatedDateParsed":"2024-01-16 00:00:00 UTC","expiresDateParsed":"2025-02-16 00:00:0
{"reason":"added","domainName":"pokertwisland.com","createdDateParsed":"2024-01-16 09:26:09 UTC","updatedDateParsed":"2024-01-16 13:33:08 UTC","expiresDateParsed":"2025-01-16 09:26:09 UT
{"reason":"added","domainName":"reklamerabauken.de","createdDateParsed":"","updatedDateParsed":"2024-01-16 13:30:17 UTC","expiresDateParsed":"","createdDateRaw":"","updatedDateRaw":"2024
{"reason":"updated","domainName":"brindepromocional.com","createdDateParsed":"2014-01-17 22:00:24 UTC","updatedDateParsed":"2024-01-12 16:39:07 UTC","expiresDateParsed":"2026-01-17 22:00
{"reason":"added","domainName":"titansdispatcher.com","createdDateParsed":"2024-01-16 13:21:51 UTC","updatedDateParsed":"2024-01-16 13:27:23 UTC","expiresDateParsed":"2025-01-16 13:21:51
{"reason":"dropped","domainName":"wildnsnacks.com","createdDateParsed":"2023-01-16 10:15:43 UTC","updatedDateParsed":"2023-01-16 10:15:45 UTC","expiresDateParsed":"2024-01-16 10:15:43 UT
{"reason":"added","domainName":"svidomebatkivstvo.media","createdDateParsed":"2024-01-16 13:26:33 UTC","updatedDateParsed":"2024-01-16 13:26:41 UTC","expiresDateParsed":"2025-01-16 13:26
{"reason":"added","domainName":"corklupjlzje.top","createdDateParsed":"2024-01-16 11:38:14 UTC","updatedDateParsed":"2024-01-16 11:48:32 UTC","expiresDateParsed":"2025-01-16 11:38:14 UTC
{"reason":"added","domainName":"ctnlxbicqqfb.top","createdDateParsed":"2024-01-16 11:38:17 UTC","updatedDateParsed":"2024-01-16 11:48:32 UTC","expiresDateParsed":"2025-01-16 11:38:17 UTC
{"reason":"added","domainName":"cvhruhqngyrz.top","createdDateParsed":"2024-01-16 11:38:14 UTC","updatedDateParsed":"2024-01-16 11:48:32 UTC","expiresDateParsed":"2025-01-16 11:38:14 UTC
{"reason":"added","domainName":"invycidqqebs.top","createdDateParsed":"2024-01-16 11:38:14 UTC","updatedDateParsed":"2024-01-16 11:48:32 UTC","expiresDateParsed":"2025-01-16 11:38:14 UTC
{"reason":"added","domainName":"nddajeyzwxff.top","createdDateParsed":"2024-01-16 11:38:14 UTC","updatedDateParsed":"2024-01-16 11:48:32 UTC","expiresDateParsed":"2025-01-16 11:38:14 UTC
{"reason":"added","domainName":"obnkrtksyuwm.top","createdDateParsed":"2024-01-16 11:38:17 UTC","updatedDateParsed":"2024-01-16 11:48:32 UTC","expiresDateParsed":"2025-01-16 11:38:17 UTC
{"reason":"added","domainName":"pdzdjibeusas.top","createdDateParsed":"2024-01-16 11:38:14 UTC","updatedDateParsed":"2024-01-16 11:48:32 UTC","expiresDateParsed":"2025-01-16 11:38:14 UTC
{"reason":"added","domainName":"svmmwsxvpzzs.top","createdDateParsed":"2024-01-16 11:38:14 UTC","updatedDateParsed":"2024-01-16 11:48:27 UTC","expiresDateParsed":"2025-01-16 11:38:14 UTC
{"reason":"added","domainName":"xegfsqfvgdbz.top","createdDateParsed":"2024-01-16 11:38:14 UTC","updatedDateParsed":"2024-01-16 11:48:32 UTC","expiresDateParsed":"2025-01-16 11:38:14 UTC
{"reason":"added","domainName":"xpnlrgvagvpf.top","createdDateParsed":"2024-01-16 11:38:17 UTC","updatedDateParsed":"2024-01-16 11:48:32 UTC","expiresDateParsed":"2025-01-16 11:38:17 UTC
{"reason":"added","domainName":"altinvestmentservices.com","createdDateParsed":"2024-01-16 13:12:59 UTC","updatedDateParsed":"2024-01-16 13:29:04 UTC","expiresDateParsed":"2025-01-16 13:
```

RAM 4.98 GB CPU 35.51% Not connected to Hub v4.22.1

まとめ

本ガイドでは、Dockerを使用してStreamSets UIでパイプラインをセットアップおよび実行する手順を概説しました。WebSocket Clientの設定からLocal FSステージの定義に至る各ステップを踏むことで、最終的にデータ処理のパイプラインを構築できます。検証ステップはパイプラインの整合性を保証するもので、実行に成功すると、UIがDockerコンテナ内の指定されたディレクトリ内に出力ファイルを表示します。本ガイドで示した手順により、合理化されたデータ処理パイプラインの作成、設定、実行を正常に行い、効果的なデータ統合および管理を実現できます。