

「Apache NiFi」ってなんだろう??

超概要のご説明なのでお気楽に~!!

データフロー活用研究会
in OSC 2019 Okinawa

自己紹介

- ◆ 椎葉 雅志 (データフロー活用研究会)
 - ◆ 本日の説明担当
 - ◆ 出身は長崎県
 - ◆ 現在は関東にてお仕事しています
 - ◆ 「日本Activitiユーザ会」、「日本Apache Camelユーザ会」でも活動中
- ◆ 松山 千尋 (データフロー活用研究会)
 - ◆ 本日のデモ担当
 - ◆ 出身は沖縄県
 - ◆ 現在は関東にてお仕事しています

Apache NiFi ? ?

- ◆ まずは読み方
あぱっち ないふあい
- ◆ ライセンスは？
Apache License Version 2.0
- ◆ なにで書かれている？
Java言語

Apache NiFi ???

- ◆ そもそもどんな製品？

データフローオーケストレーションツール

データの流れをモデリングする環境と、そのモデリング内容に則った動作をする実行環境を提供

- ◆ ブラウザ (WebベースUI) でデータフローをお絵かき
- ◆ お絵かきしたデータフローがボタンひとつで動作

Apache NiFi の歴史

- ◆ 元は米国家安全保障局 (NSA) により「Niagara Files」として開発されていたもの
- ◆ 「**Niagara Files**」 → 「NiFi」の語源
- ◆ 2014年にオープンソースソフトウェアとして公開
- ◆ 2015年にApacheのトップレベルプロジェクト (TLP) に昇格

Apache NiFi の歴史

◆ ちなみにリリース時期と頻度は以下となります

	2015	2016	2017	2018	2019
0系		▲7月 (0.2.1) ▲3月 (0.6.0) ▲9月 (0.3.0) ▲4月 (0.6.1) ▲12月 (0.4.0) ▲7月 (0.7.0) ▲6月 (0.7.4) ▲12月 (0.4.1) ▲10月 (0.7.1) ▲2月 (0.5.0) ▲2月 (0.5.1)	▲2月 (0.7.2) ▲5月 (0.7.3)		
1系			▲8月 (1.0.0) ▲5月 (1.2.0) ▲11月 (1.1.0) ▲6月 (1.3.0) ▲12月 (1.1.1) ▲10月 (1.4.0) ▲2月 (1.1.2)	▲1月 (1.5.0) ▲4月 (1.6.0) ▲6月 (1.7.0) ▲7月 (1.7.1)	▲2月 (1.9.0) ▲3月 (1.9.1) ▲4月 (1.9.2) ▲10月 (1.8.0)

Apache NiFi 使用方法

- ◆ 実現したいデータの流れ(データフロー)を Apache NiFiの部品(プロセッサ)を使ってお絵かきする
- ◆ データの流れとは、具体的なデータの入出力やそのデータの取り回しを定義すること

データフローをお絵かきする？

例)ファイル(CSV形式)を読み込み、読み込んだデータをJSON形式に変換し別ファイル名として書き込む。

データの流れ(処理のイメージ)を簡単にお絵かきすると



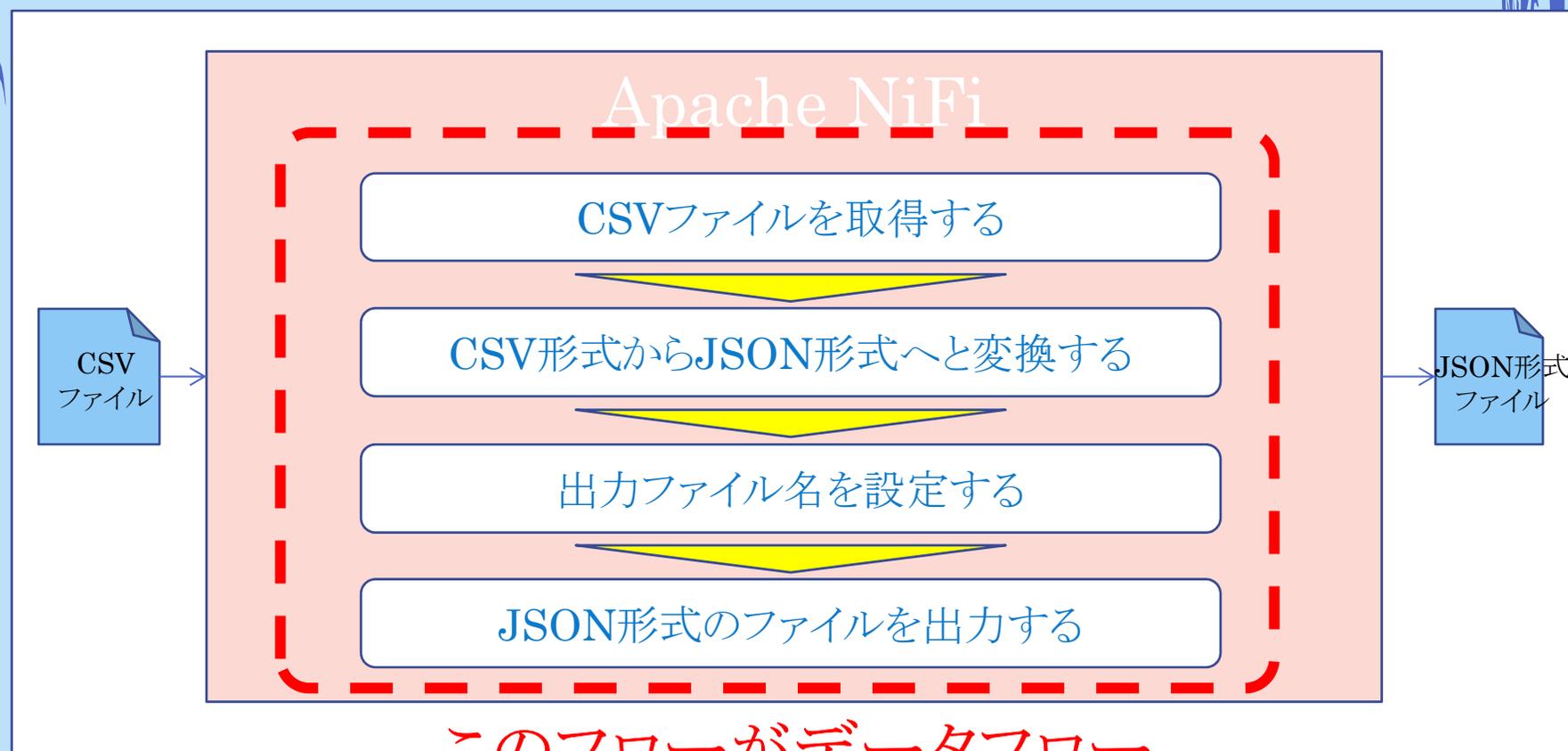
データフローをお絵かきする？

もう少し処理(手順)をイメージし、お絵かきすると



データフローをお絵かきする？

処理の手順(データの流れ)っぽいものが出てきました。



このフローがデータフロー

(かなり拡大解釈してます。本来は取り扱うデータにもっと着目します)

データフローをお絵かきする？

個々で行う処理ができました。



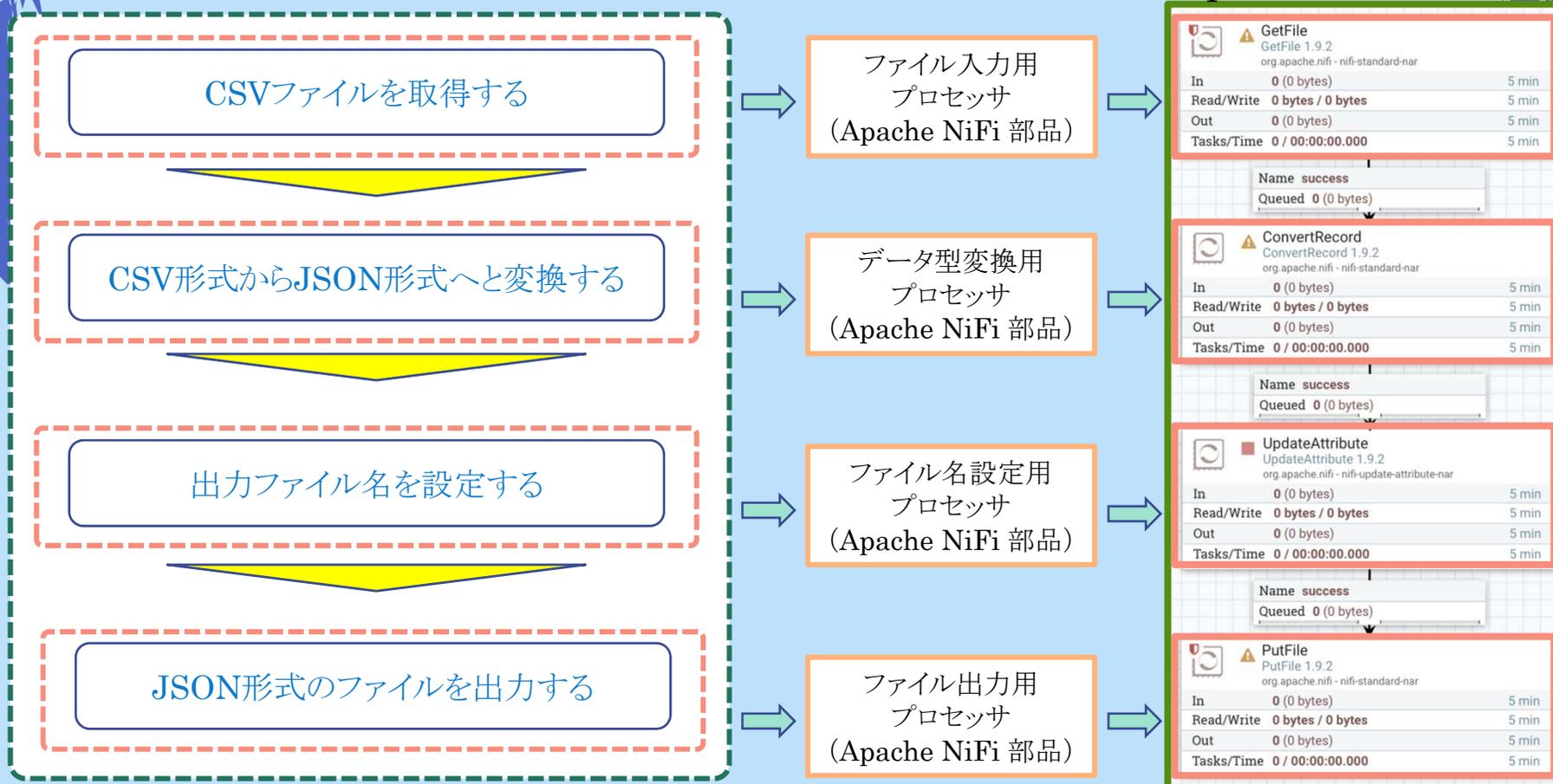
この個々の囲みの単位がプロセッサ (NiFi部品)

まとめると

- ◆ プロセッサを利用して、データフローを構築する

やりたいこと

実際のApache NiFiの開発画面



プロセッサにはどんなものがある？

NiFiにはデータ入出力用プロセッサが豊富！！

RDBMS

Cloud

No SQL

Big Data

OS
Command

Log/
File/
Search

MOM
(Mesaging Queue)

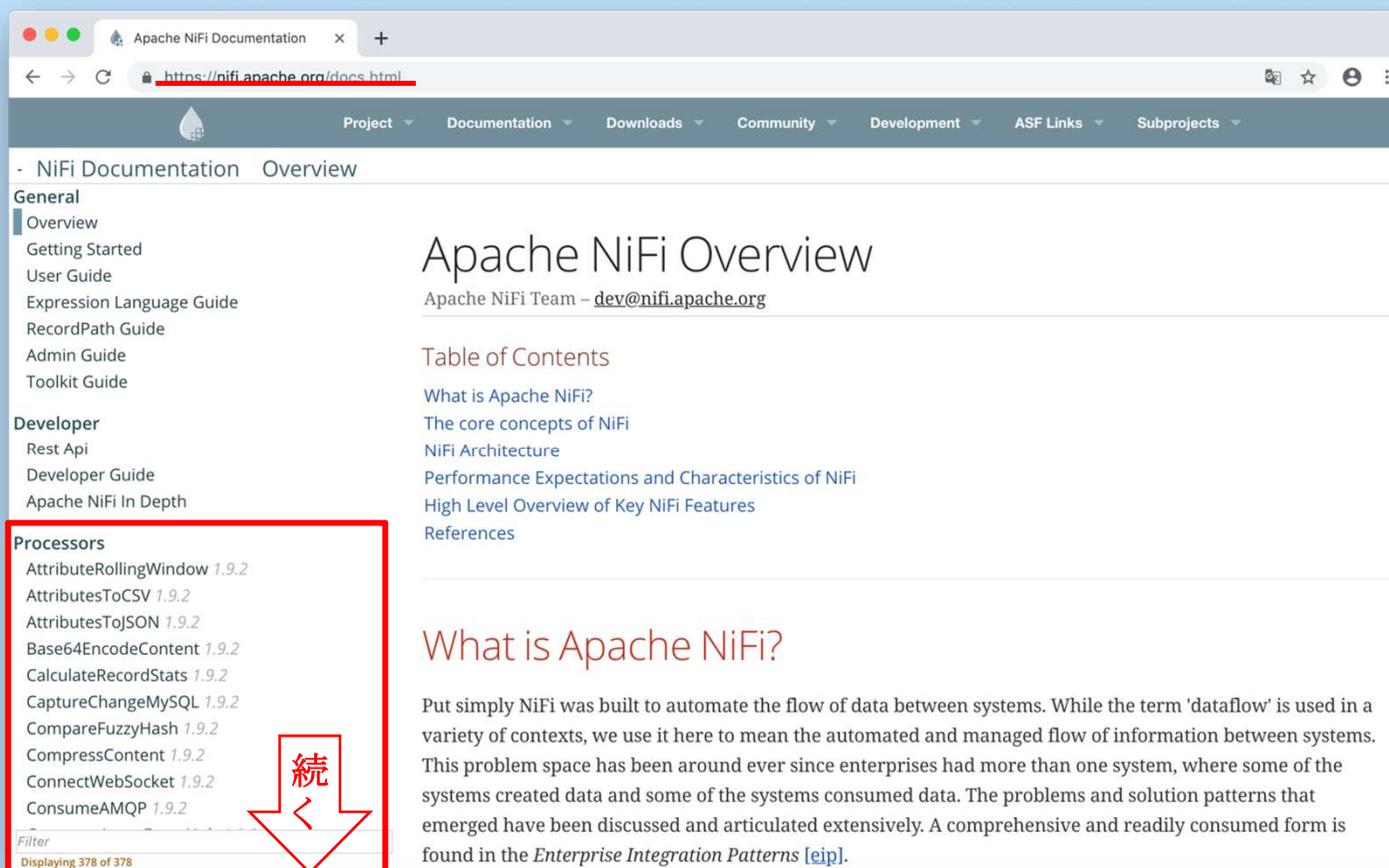
Web

HTTP,FTP,
SFTP,TCP,
UDP

プロセッサにはどんなものがある？

プロセッサ一覧を知りたい方はドキュメントを参照願います。

<<https://nifi.apache.org/docs.html>>



The screenshot shows the Apache NiFi Documentation website. The 'Processors' section is highlighted with a red box, and a red arrow with the Japanese character '続' (Continue) points to the 'What is Apache NiFi?' section.

Apache NiFi Documentation Overview

General

- Overview
- Getting Started
- User Guide
- Expression Language Guide
- RecordPath Guide
- Admin Guide
- Toolkit Guide

Developer

- Rest Api
- Developer Guide
- Apache NiFi In Depth

Processors

- AttributeRollingWindow 1.9.2
- AttributesToCSV 1.9.2
- AttributesToJSON 1.9.2
- Base64EncodeContent 1.9.2
- CalculateRecordStats 1.9.2
- CaptureChangeMySQL 1.9.2
- CompareFuzzyHash 1.9.2
- CompressContent 1.9.2
- ConnectWebSocket 1.9.2
- ConsumeAMQP 1.9.2

Filter

Displaying 378 of 378

Apache NiFi Overview

Apache NiFi Team – dev@nifi.apache.org

Table of Contents

- What is Apache NiFi?
- The core concepts of NiFi
- NiFi Architecture
- Performance Expectations and Characteristics of NiFi
- High Level Overview of Key NiFi Features
- References

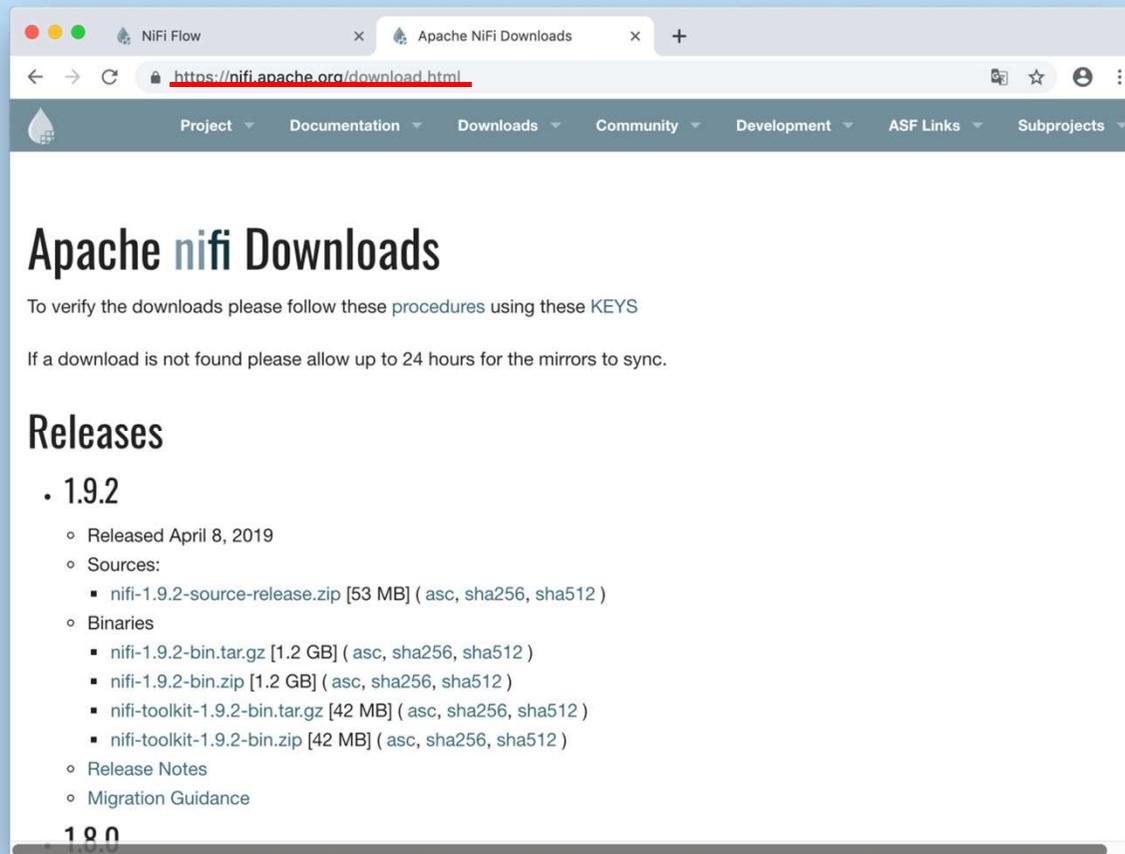
What is Apache NiFi?

Put simply NiFi was built to automate the flow of data between systems. While the term 'dataflow' is used in a variety of contexts, we use it here to mean the automated and managed flow of information between systems. This problem space has been around ever since enterprises had more than one system, where some of the systems created data and some of the systems consumed data. The problems and solution patterns that emerged have been discussed and articulated extensively. A comprehensive and readily consumed form is found in the *Enterprise Integration Patterns* [eip].

Apache NiFiを使ってみよう！！

1. Apache NiFiのダウンロード

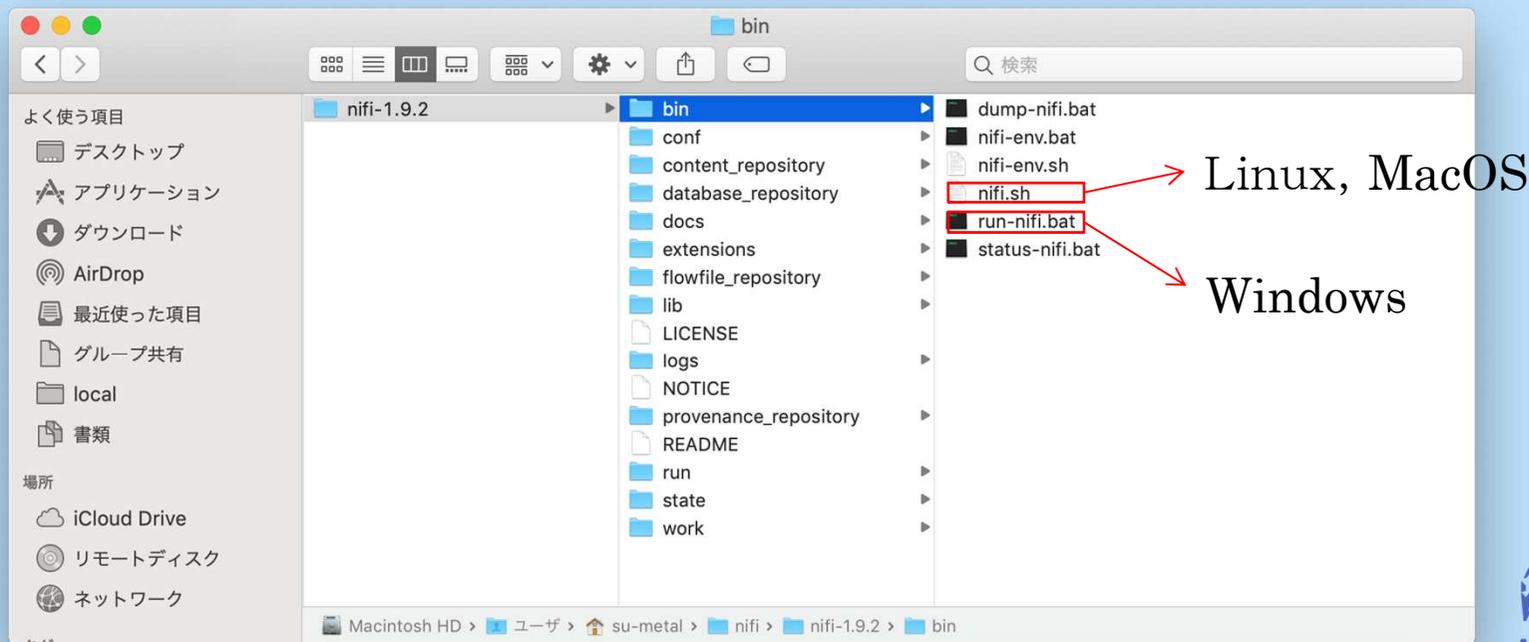
<<https://nifi.apache.org/download.html>>



The screenshot shows a web browser window with two tabs: 'NIFI Flow' and 'Apache NiFi Downloads'. The address bar shows the URL <https://nifi.apache.org/download.html>. The page content includes a navigation menu with items like 'Project', 'Documentation', 'Downloads', 'Community', 'Development', 'ASF Links', and 'Subprojects'. The main heading is 'Apache nifi Downloads'. Below this, there is a note: 'To verify the downloads please follow these procedures using these KEYS' and 'If a download is not found please allow up to 24 hours for the mirrors to sync.' The 'Releases' section is expanded to show version '1.9.2', which was released on April 8, 2019. Under 'Sources', there are links for 'nifi-1.9.2-source-release.zip [53 MB]' and 'nifi-1.9.2-bin.tar.gz [1.2 GB]'. Under 'Binaries', there are links for 'nifi-1.9.2-bin.zip [1.2 GB]', 'nifi-toolkit-1.9.2-bin.tar.gz [42 MB]', and 'nifi-toolkit-1.9.2-bin.zip [42 MB]'. At the bottom, there are links for 'Release Notes' and 'Migration Guidance'. The version '1.9.0' is partially visible at the bottom of the list.

Apache NiFiを試してみよう！！

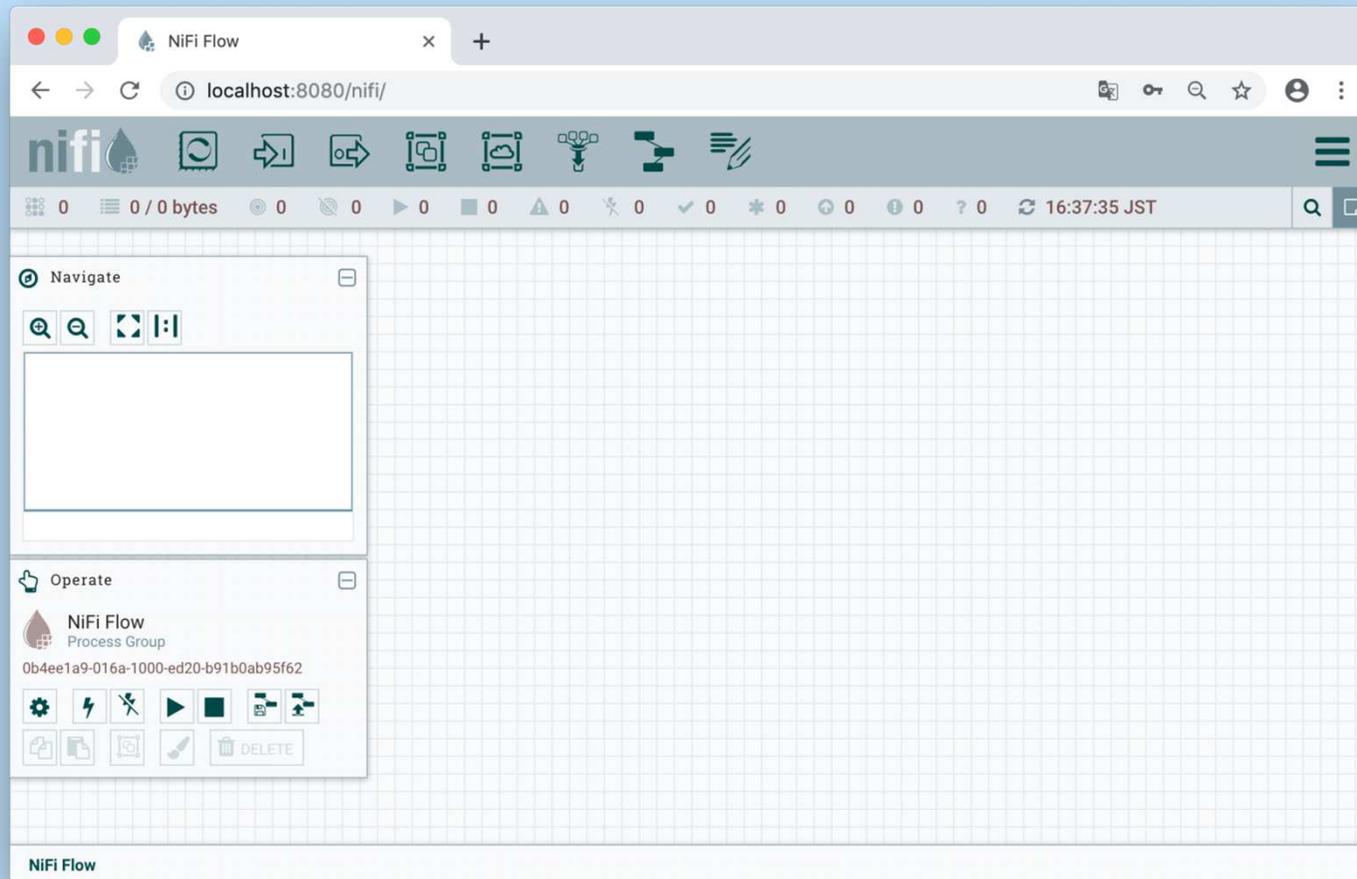
2. ダウンロード資料の (zip or tar.gz) 解凍
3. NiFiの起動



Apache NiFiを使ってみよう！！

4. ブラウザからNiFiにアクセス

<<http://localhost:8080/nifi/>>



Apache NiFi で データフローを作ってみよう

Demo