Hadoop HDFS アダプタご紹介

本資料で表記しているHadoop HDFSアダプタの正式な製品名は「DataSpider アダプタ for Apache™ Hadoop®」です。

2021年9月版

株式会社セゾンテクノロジー

対象製品 / サービス

DataSpider Servista　DataSpider Cloud

Copyright © Saison Technology Co., Ltd. All Rights Reserved.

HULFT

目次

■ **製品概要**
- Hadoop HDFS アダプタとは
- 連携イメージ
- アダプタ機能一覧

■ **各種処理と設定**
- CSVデータ読み取り処理
- CSVデータ書き込み処理
- XMLデータ読み取り処理
- XMLデータ書き込み処理
- ファイルダウンロード処理
- ファイルアップロード処理
- ファイル/ディレクトリ削除
- ファイル/ディレクトリ一覧取得処理

■ **適用パターン**
- 適用パターン：分析データの収集に
- 導入メリット

Copyright © Saison Technology Co., Ltd. All Rights Reserved.

製品概要

Copyright © Saison Technology Co., Ltd. All Rights Reserved.

Hadoop HDFS アダプタとは

分散処理フレームワークであるHadoopのファイルシステム「HDFS」に対してDataSpiderからファイルの書き込み/読み込み、アップロード/ダウンロードができるアダプタです。

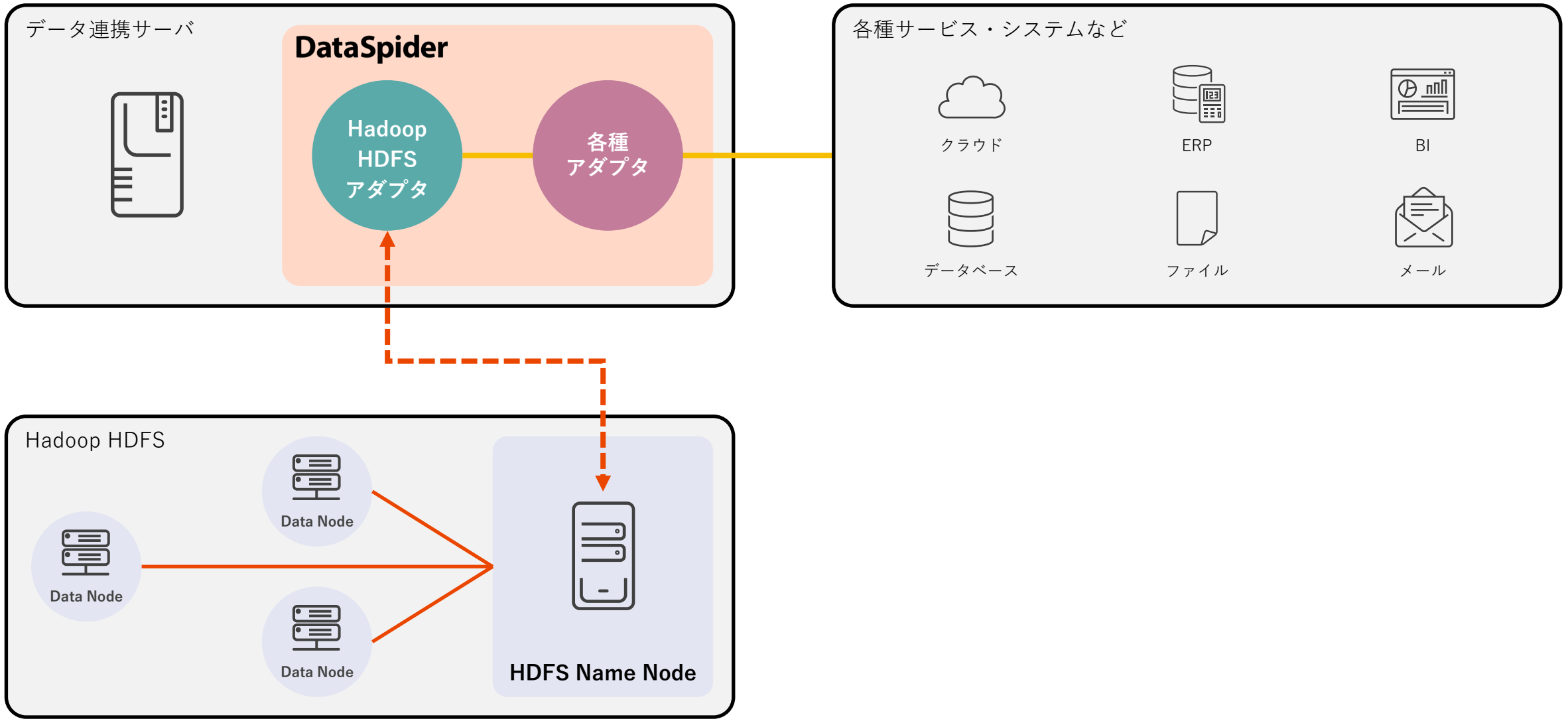
主な特徴

- HDFSに直接CSVやXML形式でファイルの書き込み/読み込みが可能
- 作成済のファイルのアップロード/ダウンロードが可能
  （アップロードするファイルのブロックサイズを指定可能）
- ファイルの一覧が取得可能
- ファイル/ディレクトリの削除が可能

※詳細は、DataSpider Servista、DataSpider Cloud 各製品のヘルプをご参照ください

Copyright © Saison Technology Co., Ltd. All Rights Reserved.

連携イメージ



Copyright © Saison Technology Co., Ltd. All Rights Reserved.

アダプタ機能一覧

○ …提供あり　－ …提供なし

種別	名称	概要説明	DataSpider Servista	DataSpider Cloud
処理	**CSVデータ読み取り処理**	HDFS(Hadoop Distributed File System)からCSV形式のファイルを読み取ります。	○	○
	CSVデータ書き込み処理	HDFS(Hadoop Distributed File System)にCSV形式でファイルを書き込みます。	○	○
	XMLデータ読み取り処理	HDFS(Hadoop Distributed File System)からXML形式のファイルを読み取ります。	○	○
	XMLデータ書き込み処理	HDFS(Hadoop Distributed File System)にXML形式でファイルを書き込みます。	○	○
	ファイルダウンロード処理	HDFS(Hadoop Distributed File System)からファイルをダウンロードします。	○	○
	ファイルアップロード処理	HDFS(Hadoop Distributed File System)にファイルをアップロードします。	○	○
	ファイル/ディレクトリ削除	HDFS(Hadoop Distributed File System)上のファイル/ディレクトリを削除します。	○	○
	ファイル/ディレクトリ一覧取得処理	HDFS(Hadoop Distributed File System)のディレクトリからファイル/ディレクトリの一覧を取得します。	○	○

Copyright © Saison Technology Co., Ltd. All Rights Reserved.

各種処理と設定

DataSpider Servista	DataSpider Cloud
○	○

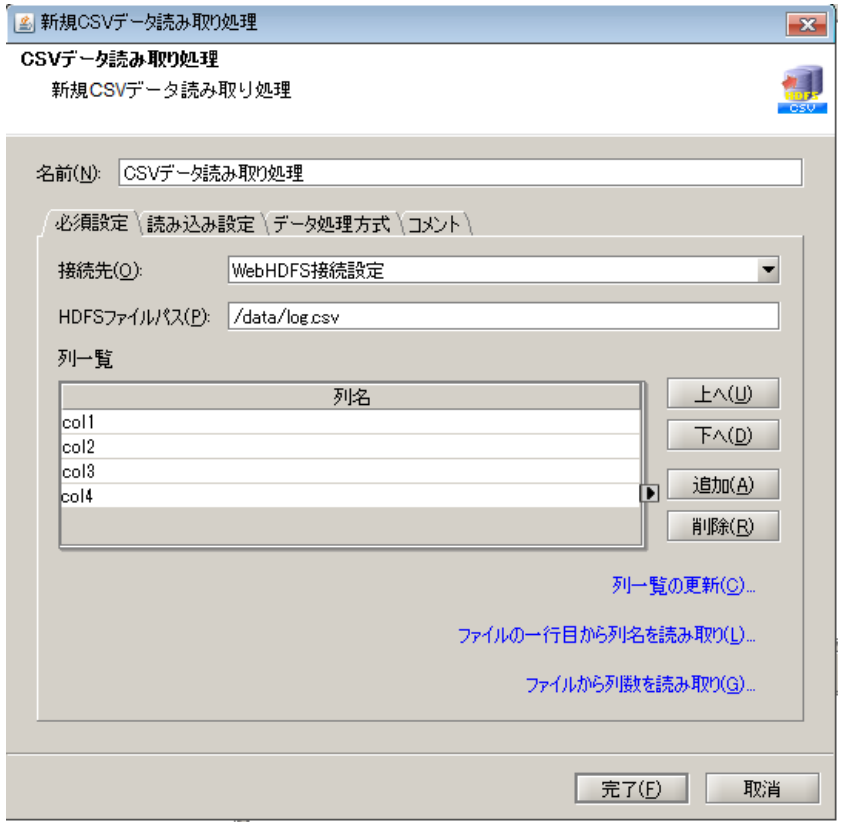
○ …提供あり ／ ― …提供なし

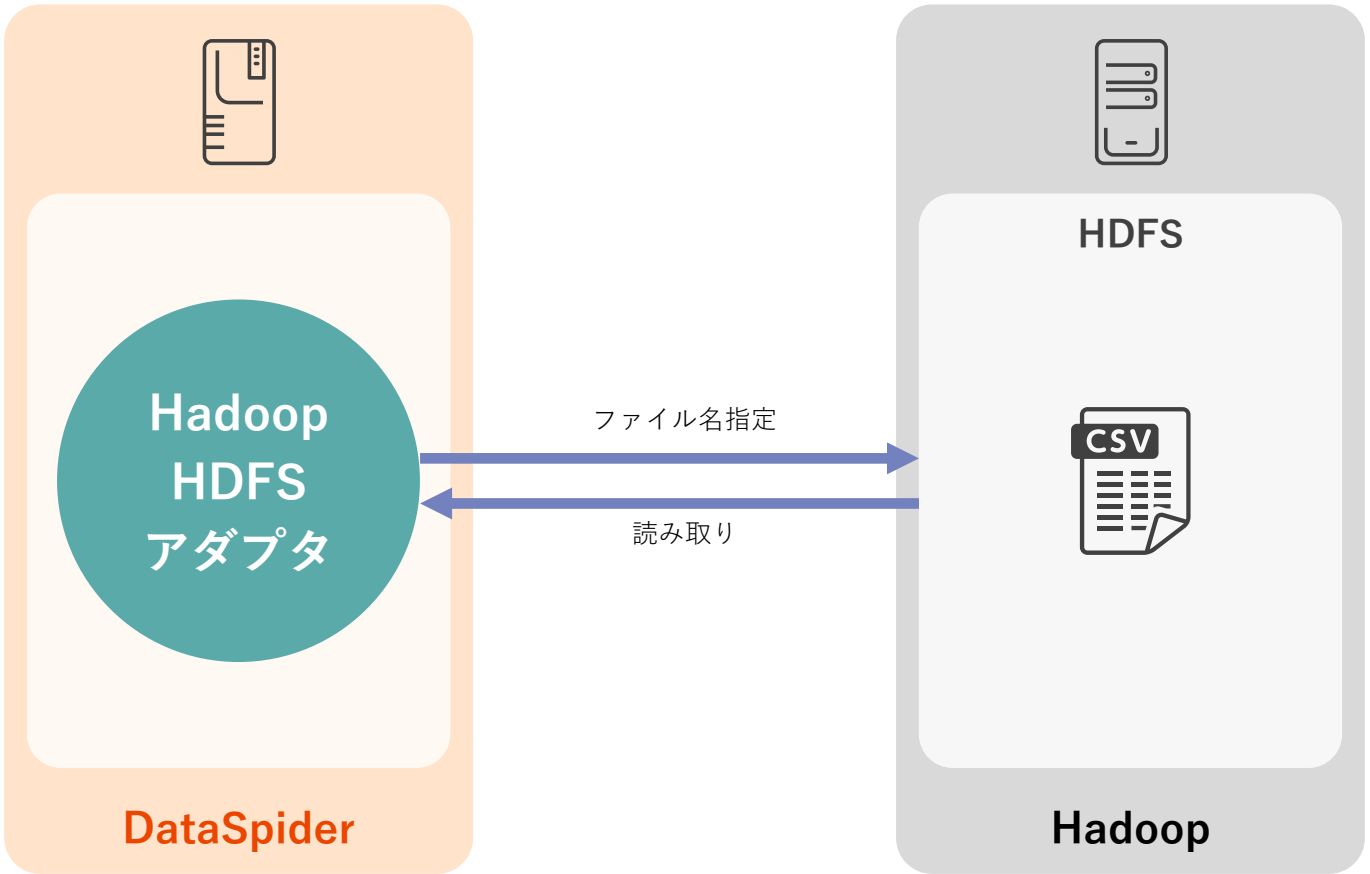
※当資料のスクリーンショットは、DataSpider Servistaを使用して取得しています。

Copyright © Saison Technology Co., Ltd. All Rights Reserved.

CSVデータ読み取り処理

HDFS(Hadoop Distributed File System)からCSV形式のファイルを読み取ります。

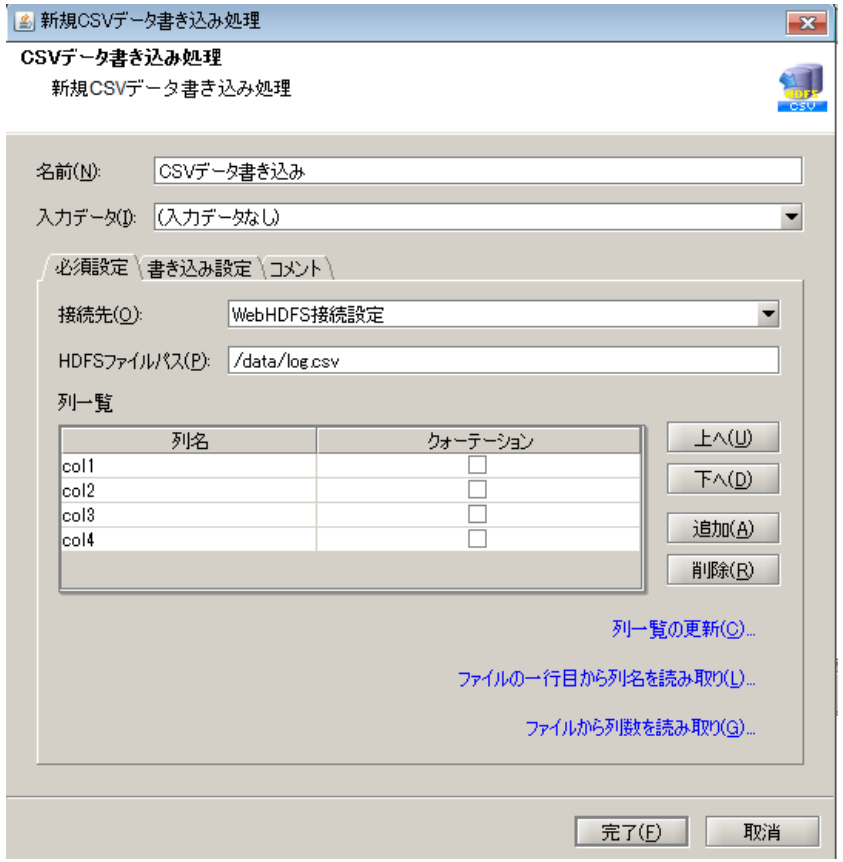


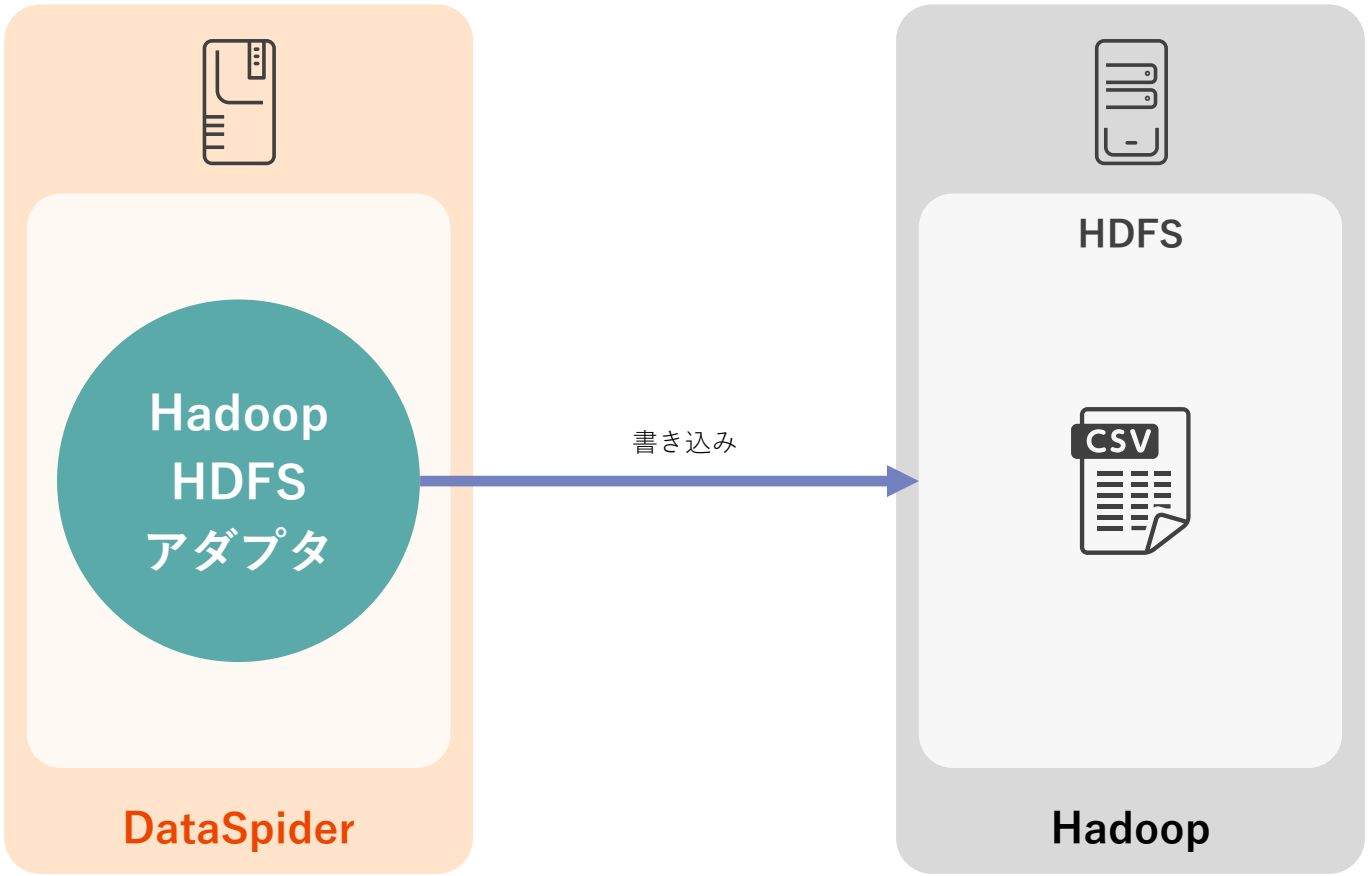


Copyright © Saison Technology Co., Ltd. All Rights Reserved.

CSVデータ書き込み処理

HDFS(Hadoop Distributed File System)にCSV形式でファイルを書き込みます。

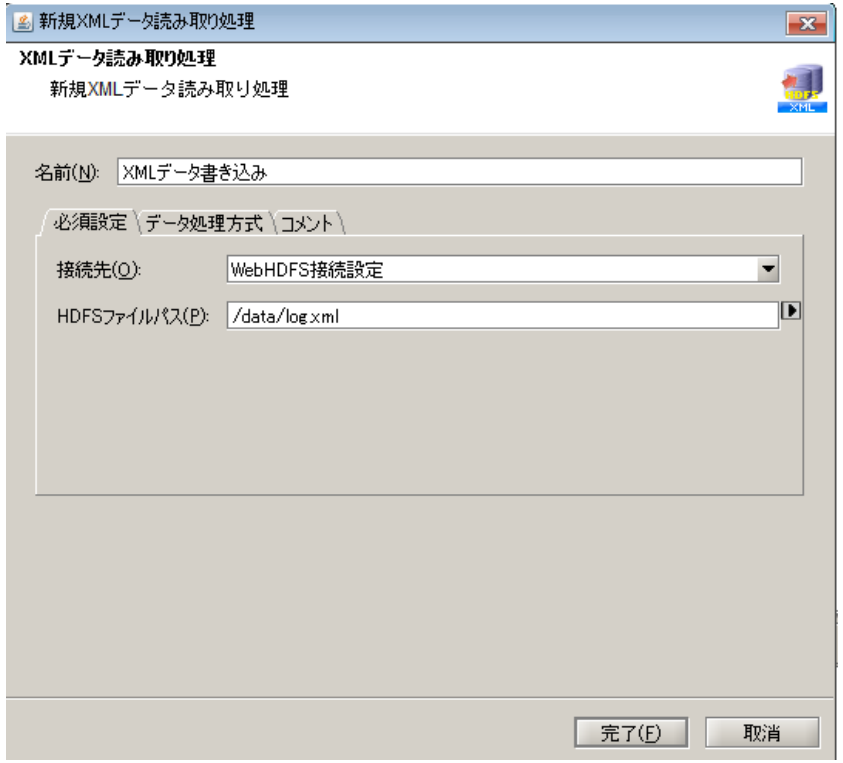


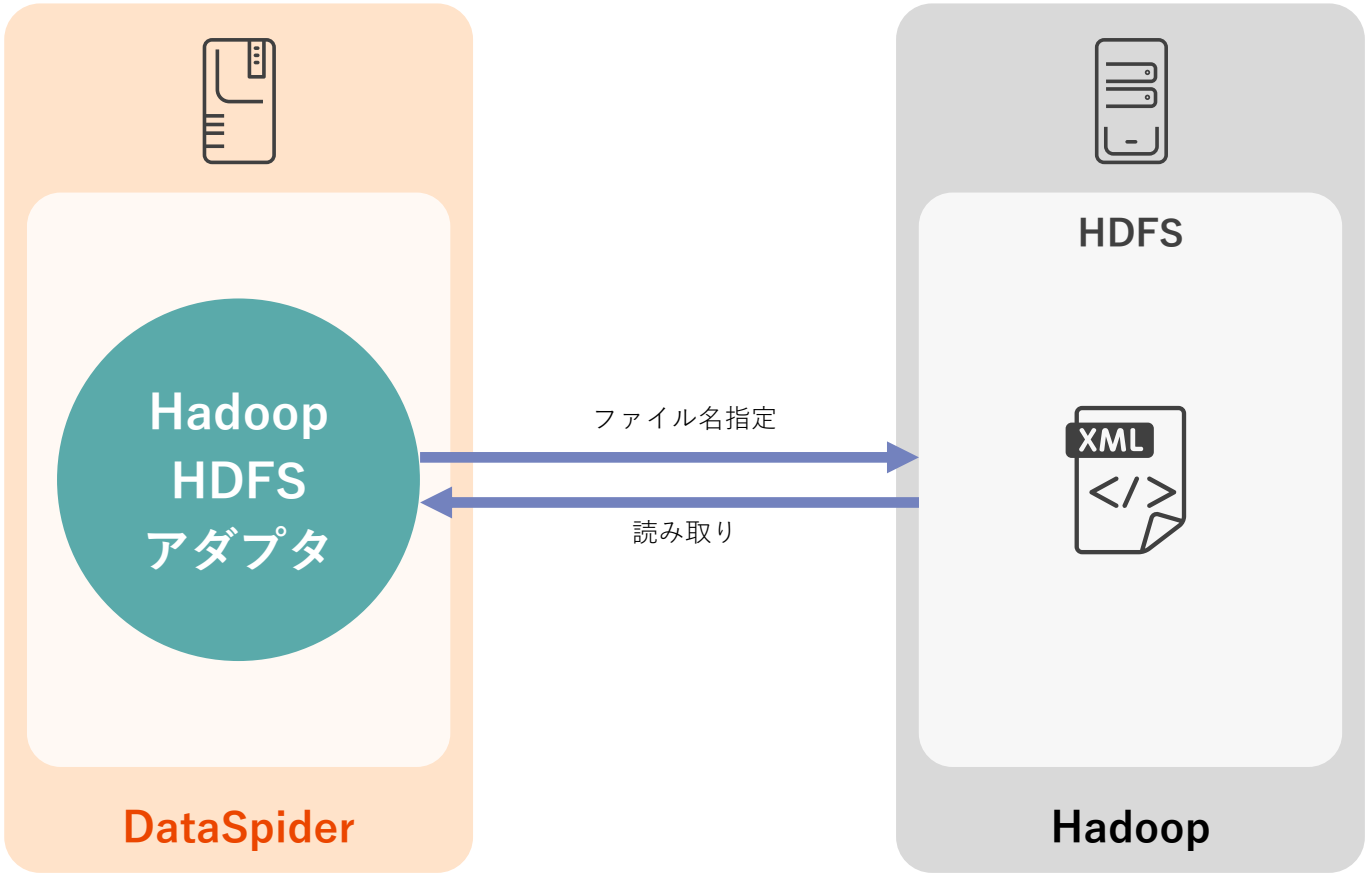


Copyright © Saison Technology Co., Ltd. All Rights Reserved.

XMLデータ読み取り処理

HDFS(Hadoop Distributed File System)からXML形式のファイルを読み取ります。

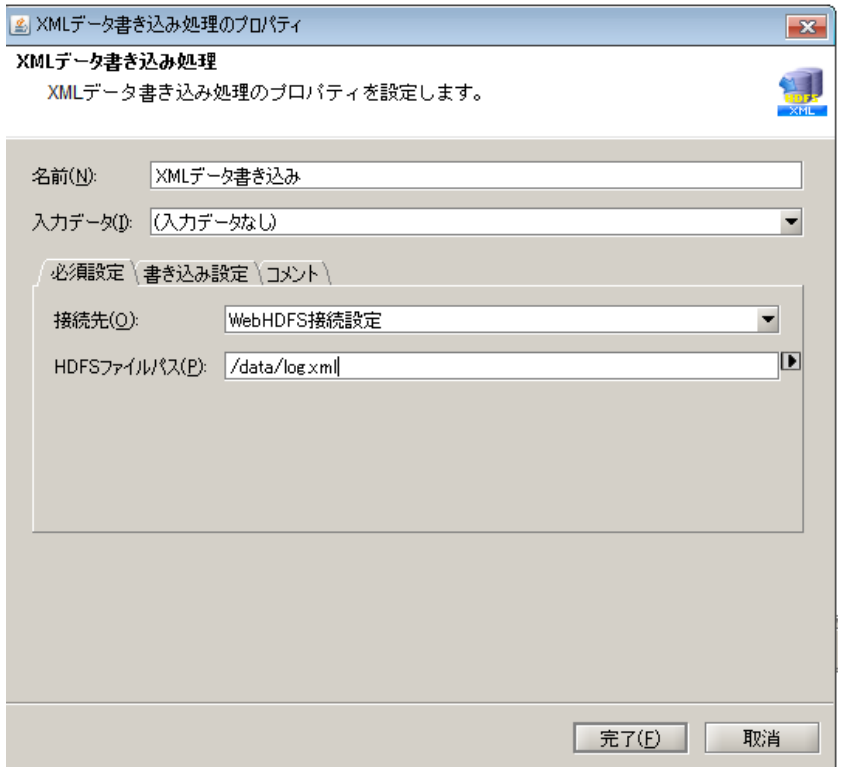


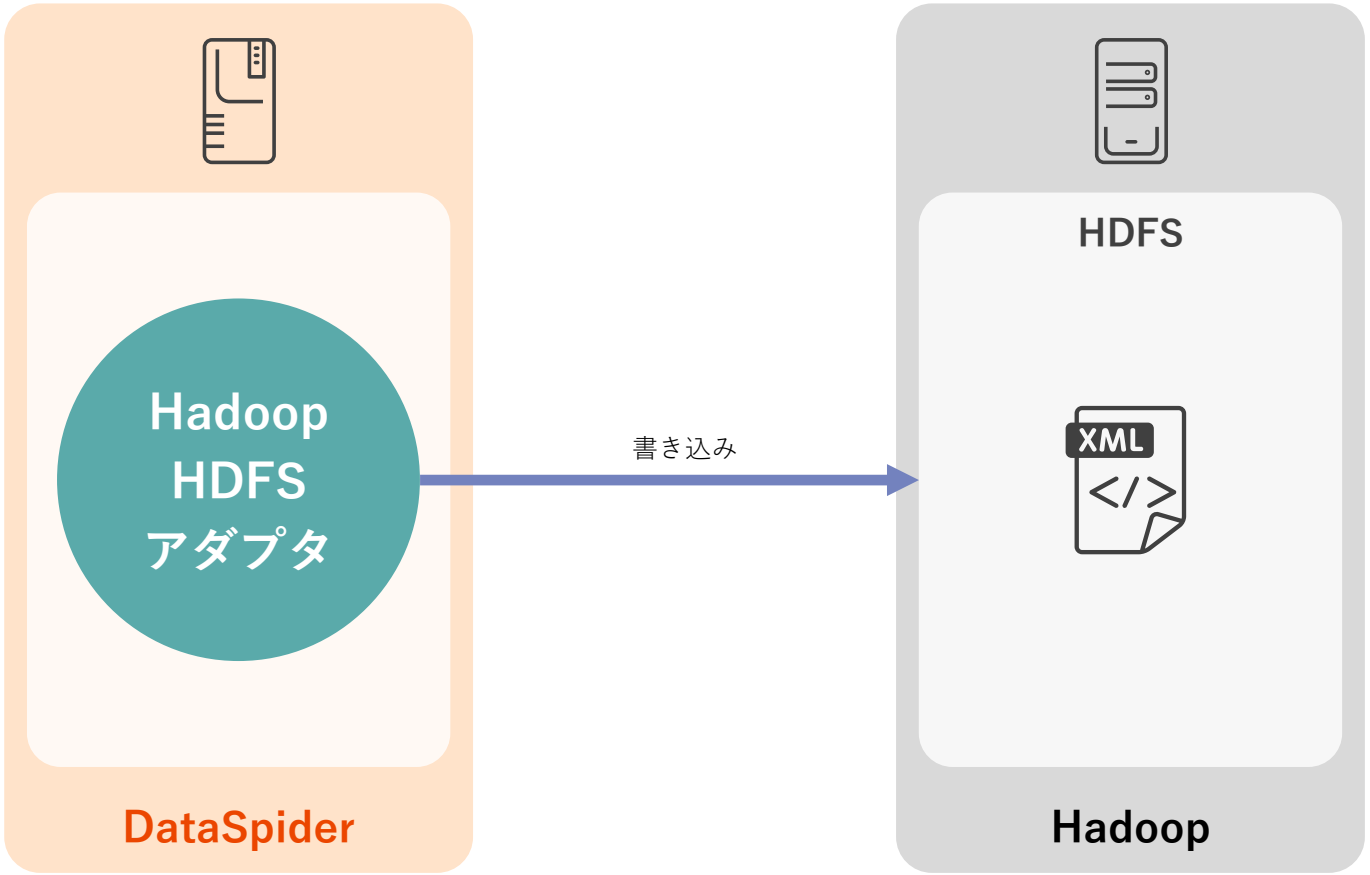


Copyright © Saison Technology Co., Ltd. All Rights Reserved.

XMLデータ書き込み処理

HDFS(Hadoop Distributed File System)にXML形式でファイルを書き込みます。

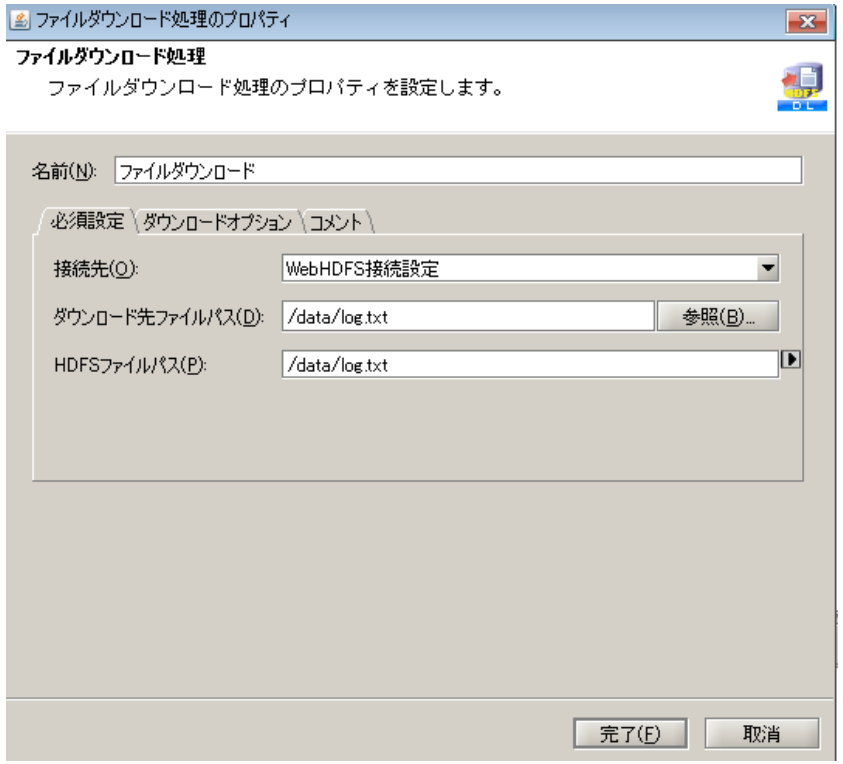


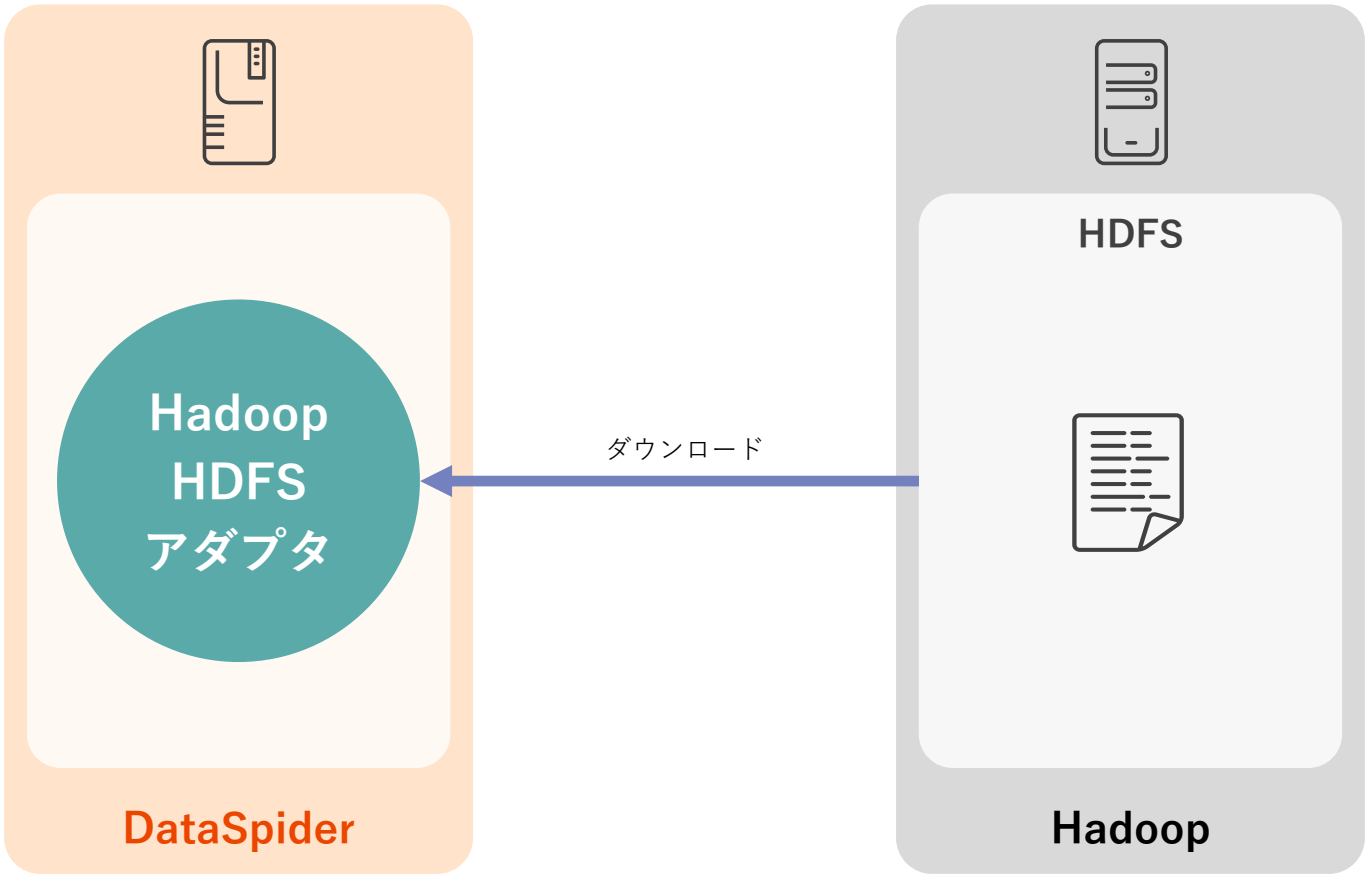


Copyright © Saison Technology Co., Ltd. All Rights Reserved.

ファイルダウンロード処理

HDFS(Hadoop Distributed File System)からファイルをダウンロードします。

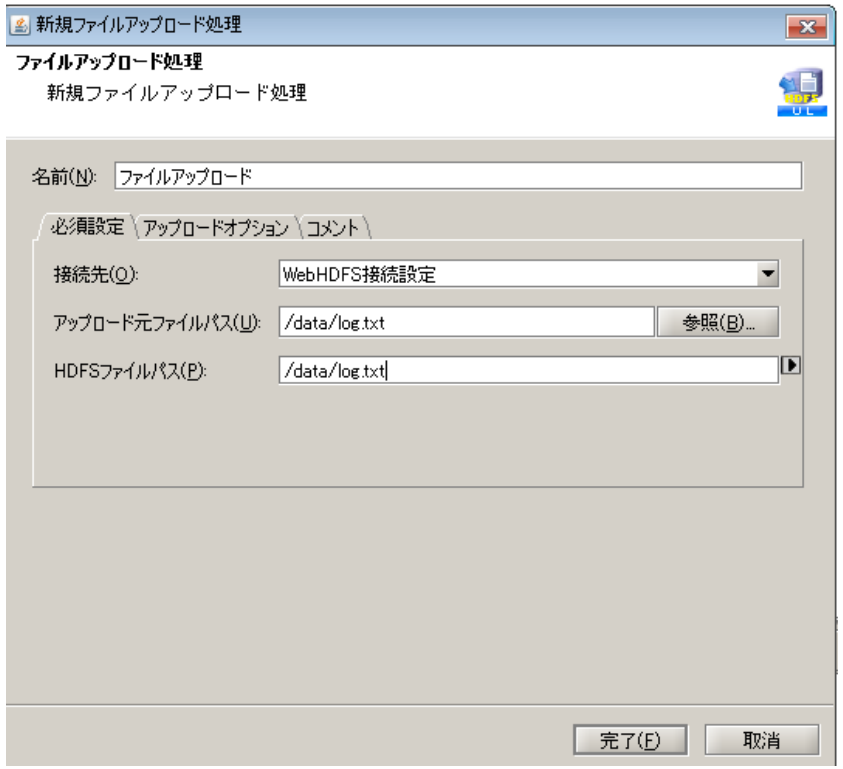


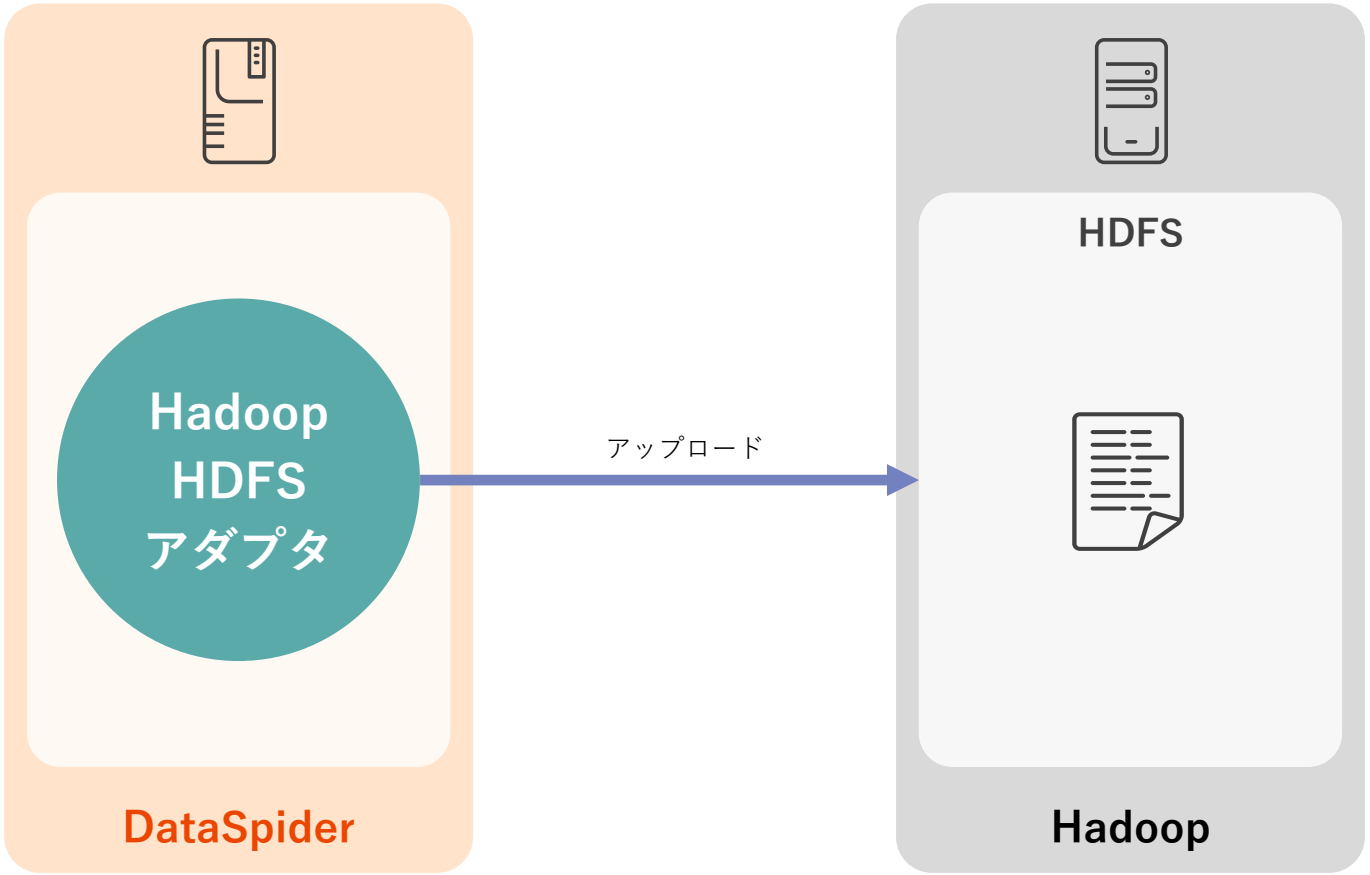


Copyright © Saison Technology Co., Ltd. All Rights Reserved.

ファイルアップロード処理

HDFS(Hadoop Distributed File System)にファイルをアップロードします。

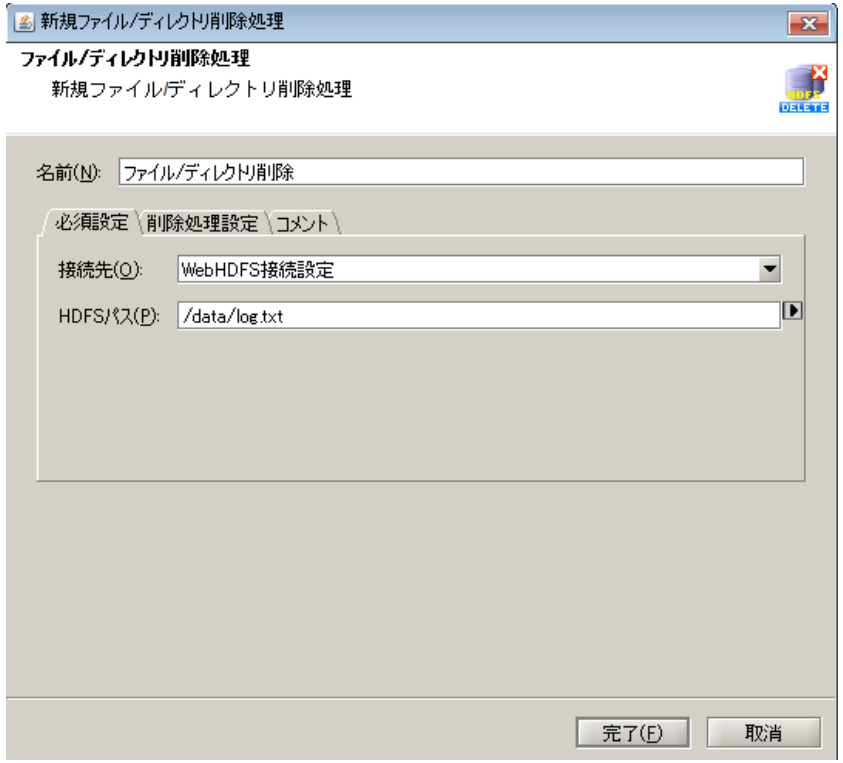


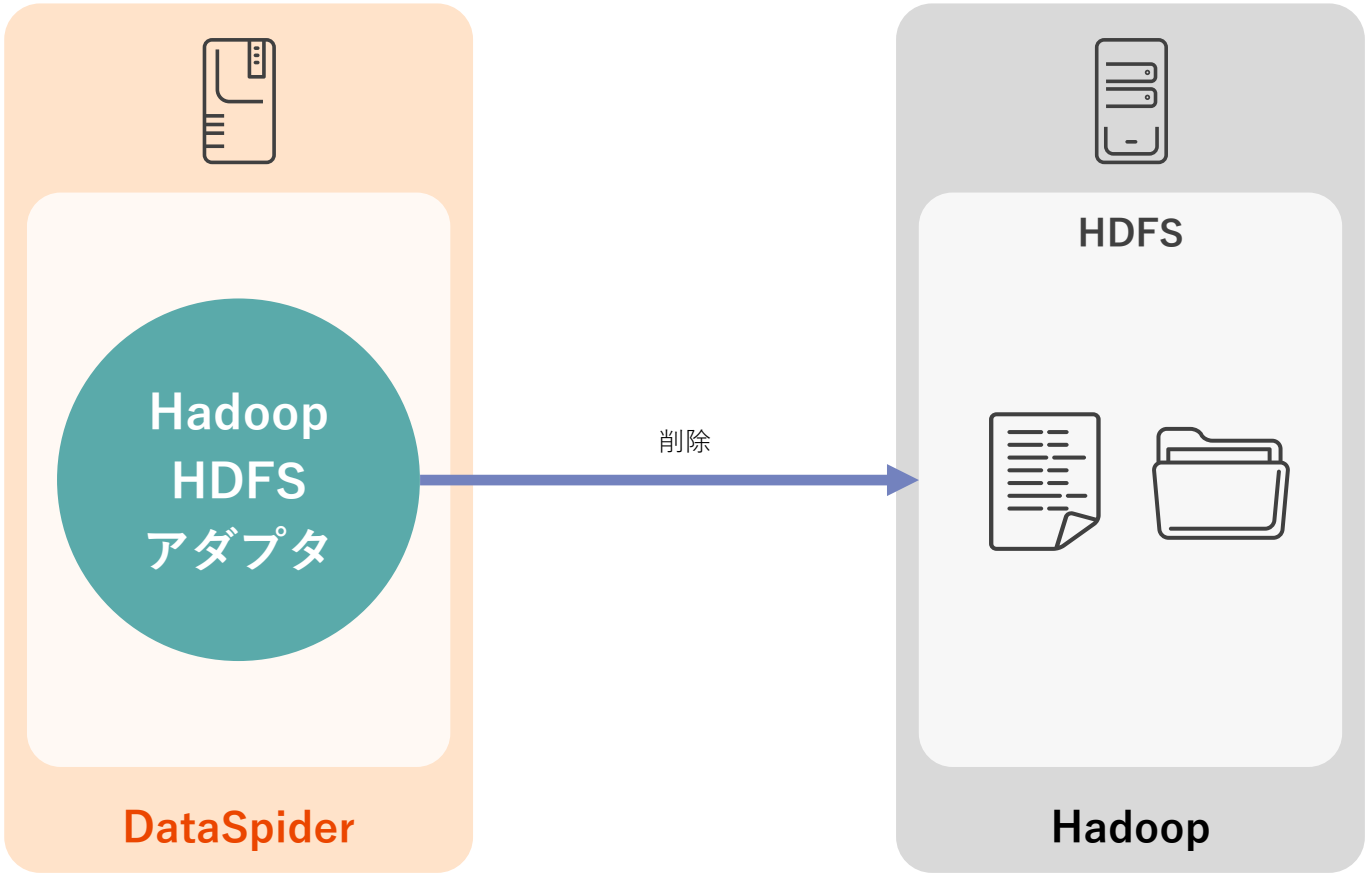


Copyright © Saison Technology Co., Ltd. All Rights Reserved.

ファイル/ディレクトリ削除

HDFS(Hadoop Distributed File System)上のファイル/ディレクトリを削除します。

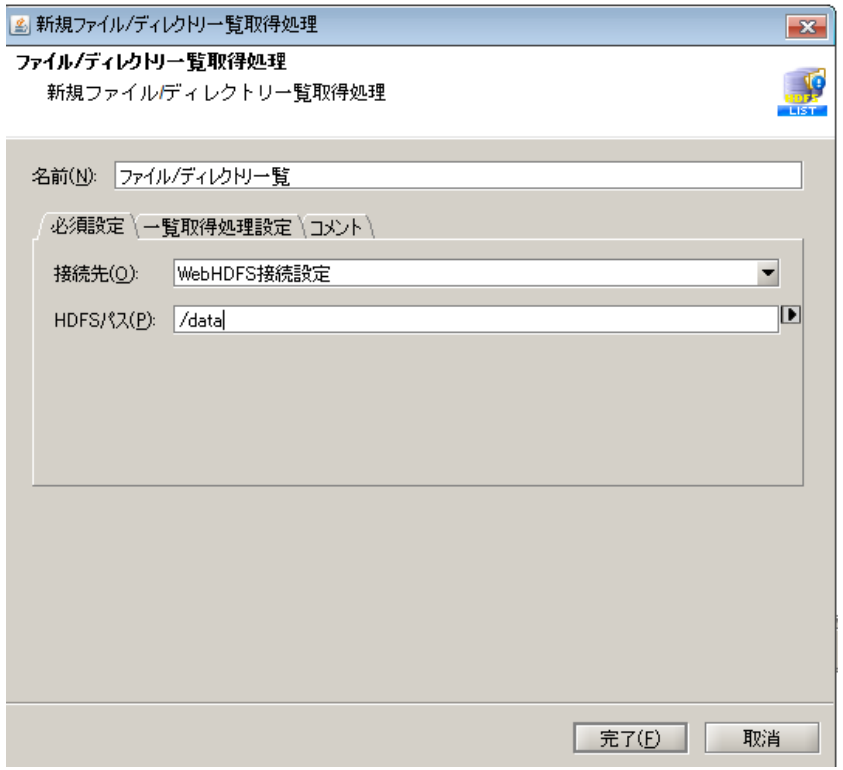


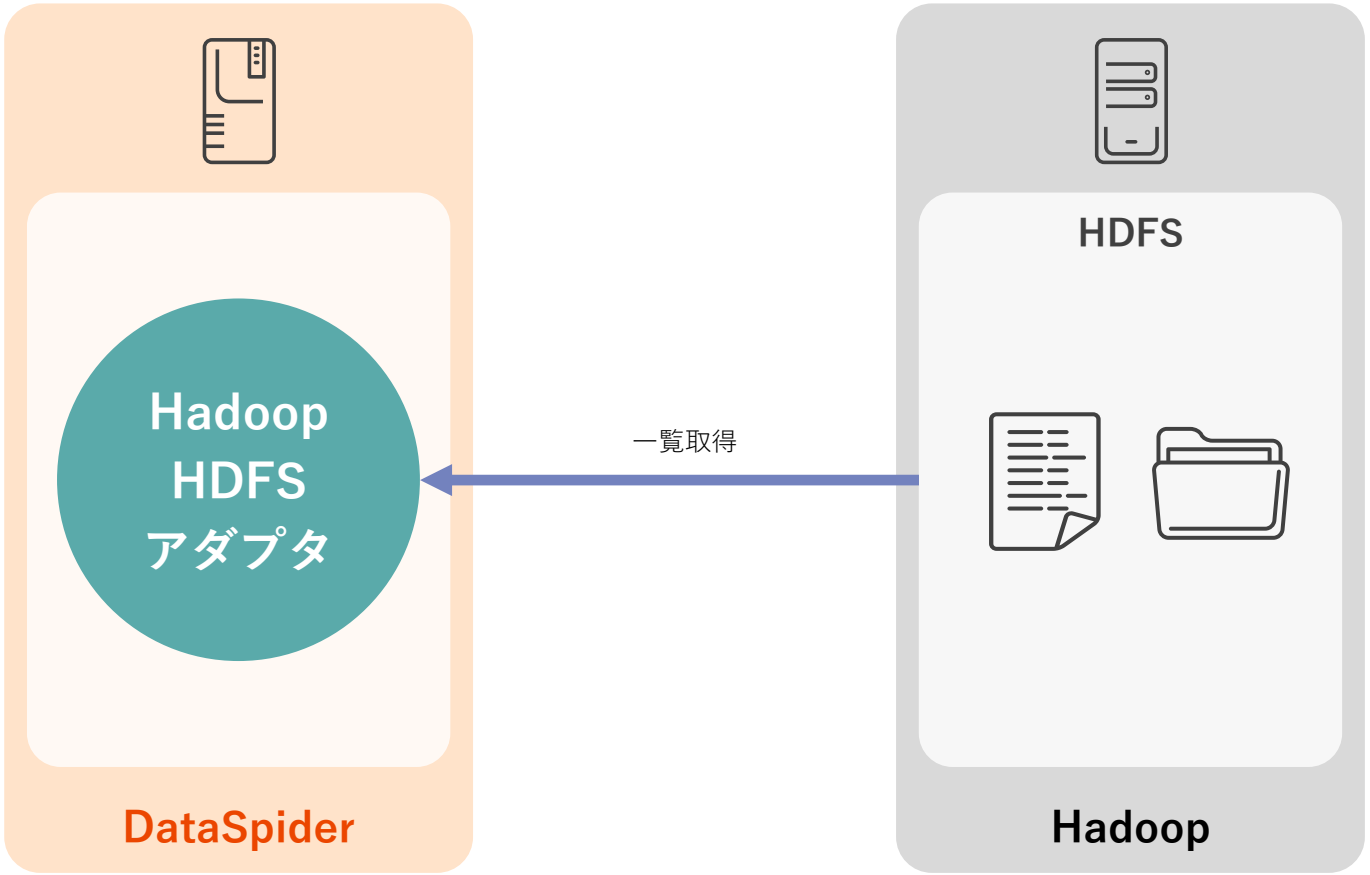


Copyright © Saison Technology Co., Ltd. All Rights Reserved.

ファイル/ディレクトリー覧取得処理

HDFS(Hadoop Distributed File System)のディレクトリからファイル/ディレクトリの一覧を取得します。



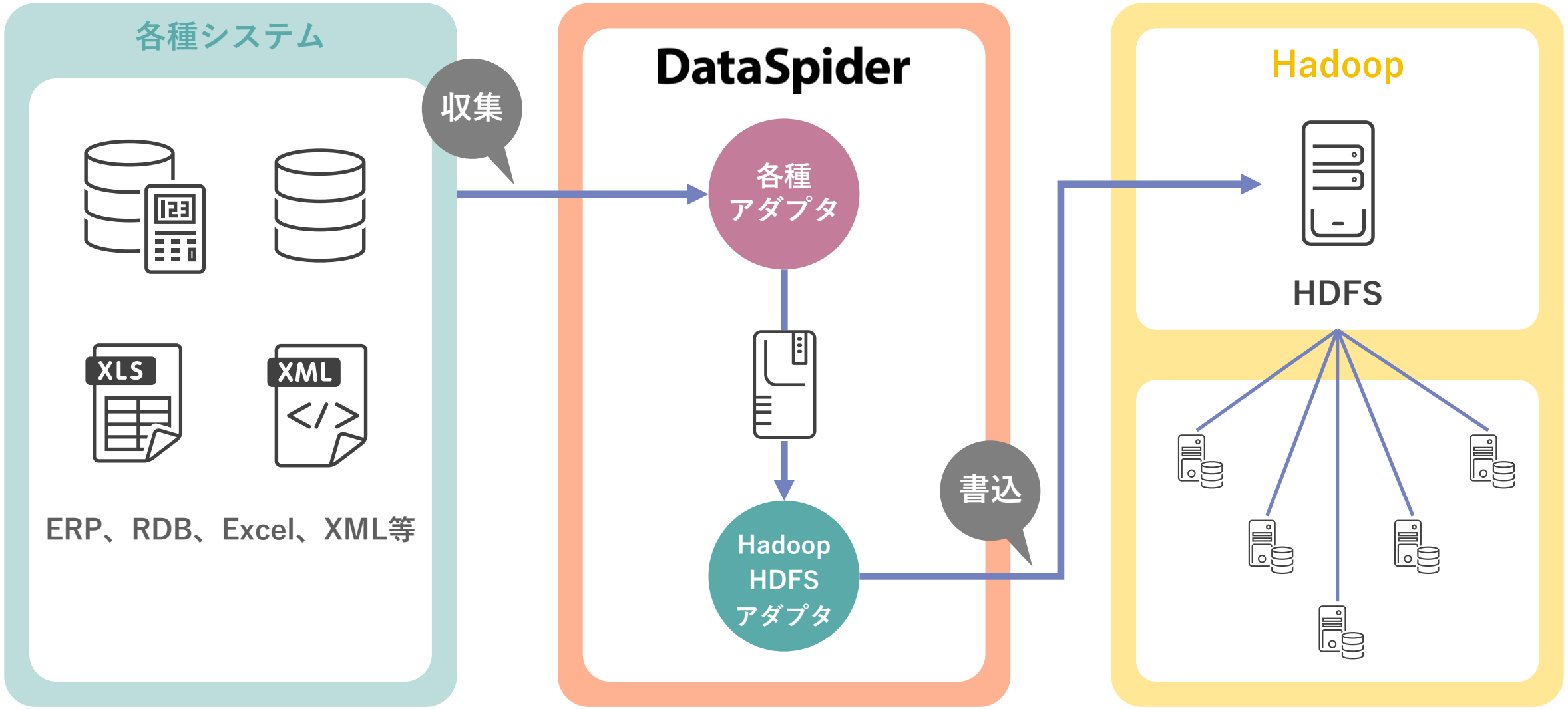


Copyright © Saison Technology Co., Ltd. All Rights Reserved.

適用パターン

Copyright © Saison Technology Co., Ltd. All Rights Reserved.

適用パターン：分析データの収集に



Copyright © Saison Technology Co., Ltd. All Rights Reserved.

導入メリット

CSV、XML形式で直接HDFSとやりとり可能

中間データを出力せずに連携可能

様々なリソースからデータ収集処理が作成可能

豊富なアダプタで連携可能

ノンプログラミングで連携可能

メンテナンスコスト削減

Copyright © Saison Technology Co., Ltd. All Rights Reserved.

HULFT

Move knowledge. Move markets.