



# DataSpider Servista 導入ガイド

# 目次

---

- はじめに
- 1. 冗長化構成の選定
- 2. サイジングの実施
- 3. 導入構成の選定
- 4. リポジトリDB有無の選定
- 5. 導入作業



# 1. 冗長化構成の選定

## 冗長化構成について

DataSpider Servistaの冗長化構成の構築には下記の制約事項があります。

- ✓ HA (High Availability) クラスタ構成に組み込むことでコールドスタンバイに対応します。



- ✓ フェイルオーバー発生時に、アクティブ側で処理中だったデータ連携トランザクションはフェールオーバーされません。



- ✓ **フェイルオーバーを想定した連携データリカバリ設計は個別に必要**です。  
また、DSSが使用するディレクトリはどちらのノードでも同じ構成にしなければなりません。



- 「クラスタリングガイド」を合わせてご参照ください。

## 冗長化構成の要素

DataSpider Servistaの冗長化について検討する必要がある要素は下記の通りです。

要素	No.	選択肢	説明	備考
DSSインストール先	1	ノード別	ノードごとにDSSをインストールします。	<ul style="list-style-type: none"> <li>ノード間のデータ同期が必要</li> <li><b>No.2より性能速度が速い</b></li> </ul>
	2	ファイル共有サーバ	ファイル共有サーバにDSSをインストールし、ノードから参照します。	<ul style="list-style-type: none"> <li>ノード間のデータ同期が必要ない</li> <li>No.1より性能速度が落ちる</li> </ul>
要素	No.	選択肢	説明	備考
リポジトリDB	3	有り	リポジトリDBにデータを保存します。	<ul style="list-style-type: none"> <li>リポジトリDBの機能を使える</li> <li>ノード間のデータ同期が必要ない</li> <li>No.4より性能速度が落ちる</li> </ul>
	4	無し	ローカルにデータを保存します。	<ul style="list-style-type: none"> <li>リポジトリDBの機能を使えない</li> <li>ノード間のデータ同期が必要</li> <li><b>No.3より性能速度が速い</b></li> </ul>
要素	No.	選択肢	説明	備考
ファイル共有サーバ	5	使用する	ファイル共有サーバにデータを保存します。	<ul style="list-style-type: none"> <li>ノード間のデータ同期が必要ない</li> <li>No.6より性能速度が落ちる</li> </ul>
	6	使用しない	ノードごとにデータを保存します。	<ul style="list-style-type: none"> <li>ノード間のデータの同期が必要</li> <li><b>No.5より性能速度が速い</b></li> </ul>

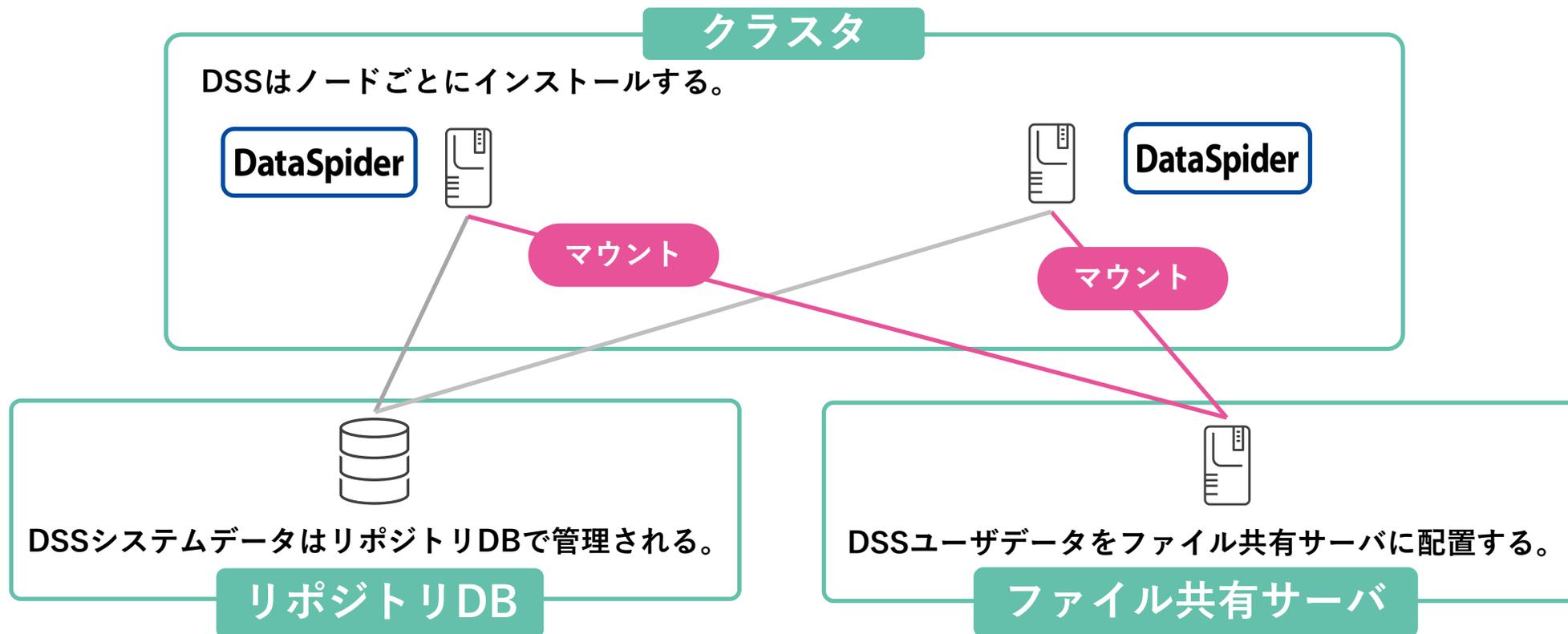
**!** リポジトリDBのメリット・デメリットについては、P.24「リポジトリDBとは」の項をご参照ください。

# DataSpider Servistaの冗長化における推奨構成

DataSpider Servistaの冗長化における推奨構成は、「**ノード別にインストール+リポジトリ有り+ファイル共有サーバを併用**」です。

## 構成のポイント

ファイル共有サーバにデータを置くことで、可用性と保守性や性能速度を高めます。



## 2. サイジングの実施

### ■前提

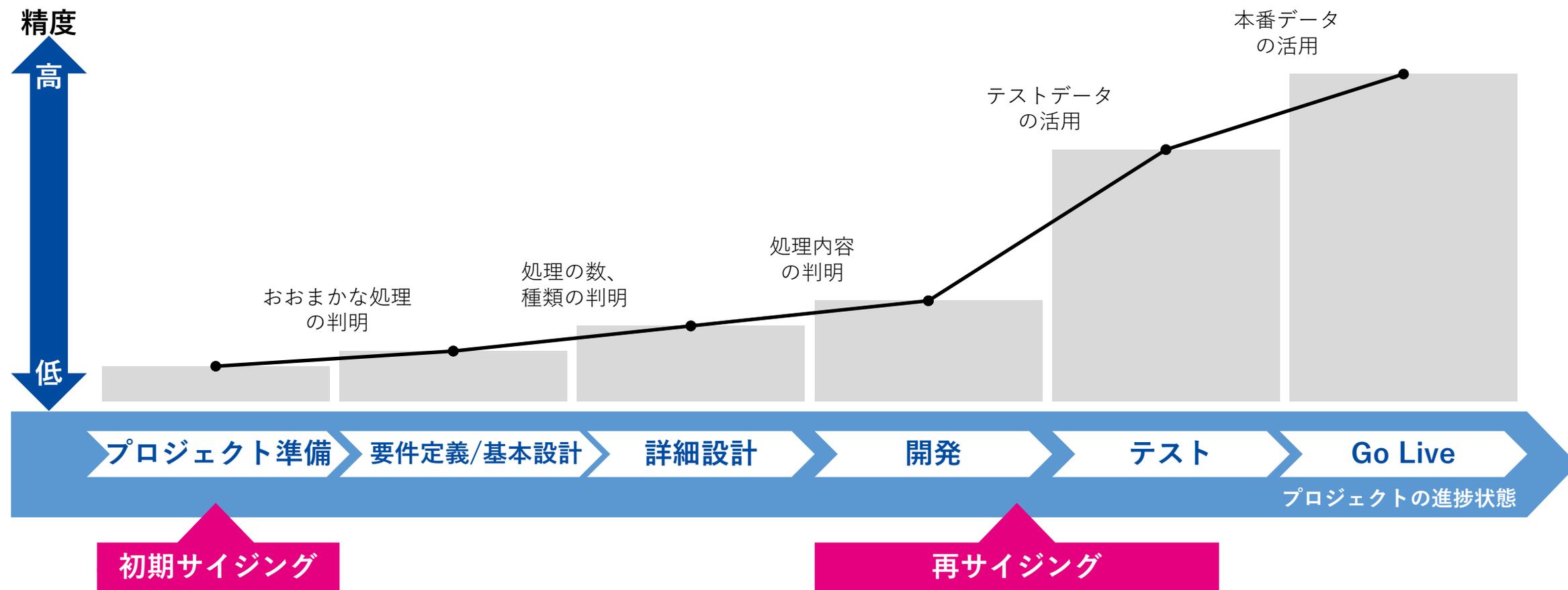
システム全体のキャパシティ・プランニングにおいて、DataSpider Servistaを運用するハードウェアを選定する際の指針と方法を参考情報としてまとめています。

### ■目的

システム構築する際に適切なハードウェアを選定するため、ベンチマーク結果や実例などから必要なリソースを検討することを目的としています。

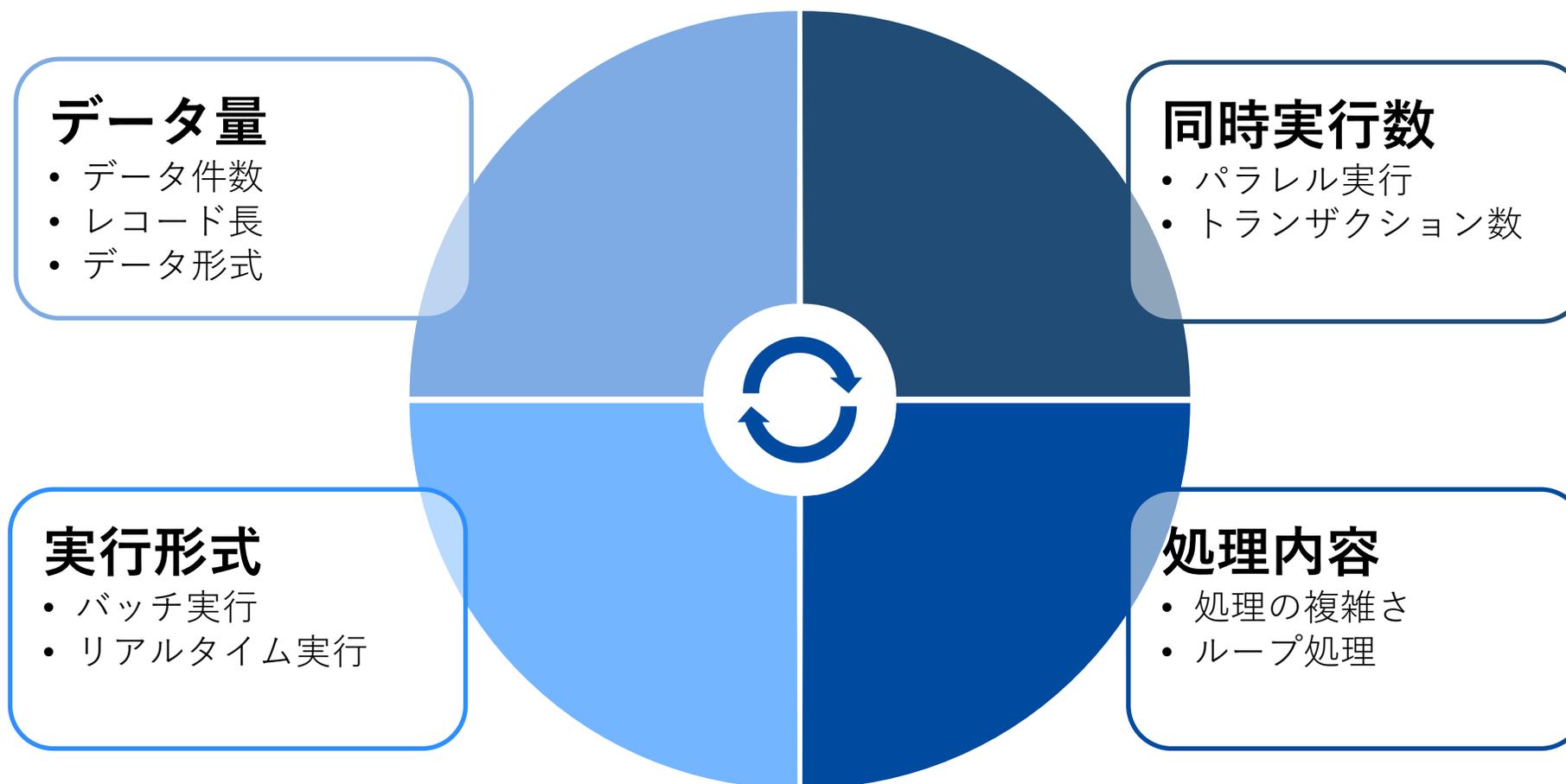
## サイジング実施のタイミングについて

サイジングはプロジェクトの進捗状態によって、精度に大きな差が出てしまいます。  
初期段階でのサイジングは精度に問題があり、ある程度の精度のぶれを考慮する必要があります。



## スケーラビリティ（拡張性）の要素

DataSpider Servistaのサイジングにおける、スケーラビリティの要素は以下のとおりです。これらの要素を確定させていくことで、サイジングの精度を上げていくことができます。



## (参考) DataSpider Servistaのサイジングの前提

本資料では DataSpider Servista のサイジングを行うにあたり、以下の内容を前提としています。

- リソース消費は線形比例する  
データ量が2倍になれば消費メモリも2倍になる。
- パフォーマンスはハードウェア（HW）の種類には依存しない  
特定のHWに依存した問題は発生しない。（製造メーカーなど）  
サーバ、デスクトップなど種類にも依存しない。
- 予測不可能な現象は無視をする  
予測不可能な現象は発生しない。
- 100%の精度は実現できない  
あくまで目安であることを意識する。

### サイジングの要素



#### CPU

性能、64bit or 32bit、個数、コア数



#### メモリ

メモリ容量、64bit or 32bit（※CPUに依存）、利用可能領域



#### ディスク

ディスク容量、性能、バックアップ

# DataSpider Servistaのサイジングの考え方

## CPUサイジングのポイント

ベンチマークをベースとした相対的な性能比較による評価が有効

- ✓ クロック数だけでは性能は計れない
  - メーカー・種類・世代・クロック数・コア数・アーキテクチャによって性能は異なる。
- ✓ 必ずしも「CPUの性能 = 処理能力」ではない
  - 但し、処理能力の指標にはなる。

## メモリサイジングのポイント

運用に支障がでない値（エラーが発生しない）を選定することを目標にする

- ✓ メモリが足りなくなると運用に支障がでる
  - OutOfMemoryエラーの発生、OSの動作が不安定になる。
- ✓ メモリ消費量は処理によって変動する
  - データ量、データ形式、処理の複雑さによって消費されるメモリ量は大きく変動する。

# DataSpider Servistaのメモリ領域について

## 💡 Javaプロセス全体のメモリ領域



### ● ヒープ領域

DataSpider サーバ上での処理に使用されており、以下の場合に大幅に消費される傾向がある。

- ✓ スクリプト上での大容量データの読み取り時
- ✓ サイズの大きなプロジェクトファイルを読み込み時

### ● Metaspace領域

DataSpider サーバが使用するJavaのクラスやメソッド情報に使用されており、以下の場合に大幅に消費される傾向がある。

- ✓ すべてのサービス上に配置されたMapperの保存情報が多い時
- ✓ デザイナから実行されたスクリプトが属するプロジェクト内のすべての Mapper の保存情報が多い時

## DataSpider Servistaに割り当てできるメモリサイズ

DataSpider Servistaは、インストールするOSに合わせて以下の2種類を用意しています。

- 32bit (x86) 版 … 32bit/64bit OSにインストール可能
- 64bit (x64) 版 … 64bit OSにインストール可能

32bit版DataSpider Servistaと64bit版DataSpider Servistaでは**割り当てできるメモリの最大値が異なる**ため、注意が必要です。



### 32bit版DataSpider Servista

OSの制限により1つのDataSpider Servistaに割り当てできるメモリは**2GB**まで  
→ JVMが使用する部分もあるため、**実質1.4～1.6GB程度**



### 64bit版DataSpider Servista

実質、搭載されている物理メモリまで割り当て可能  
※OSの種類やバージョンによって制限を受ける場合があります。

## DataSpider Servistaのメモリサイジング方法

データの容量・形式、処理の複雑さによって消費するメモリ容量は大きく変動します。また、CPUとは異なり、メモリ容量が不足するとOutOfMemoryエラーが発生し、処理が停止する恐れがあります。**メモリ不足が発生しない値をターゲットにサイジング**します。

### メモサイジング方法の検討ポイント

最もメモリ消費が見込まれるスクリプトから最大メモリ消費量を検討します。

#### STEP

1. 一番メモリ消費が見込まれるスクリプトの消費メモリを計測する
2. 1.のスクリプトと同時実行する可能性のあるスクリプトの消費メモリを計測する
3. 1.と2.の消費量の合計を算出する

#### 最大メモリ消費量

一番メモリ消費が見込まれるスクリプトの消費メモリ



同時実行する可能性のあるスクリプトの消費メモリ



オンメモリで高速な処理を実現するために、メモリは消費量の総和と捉えスループットを向上。

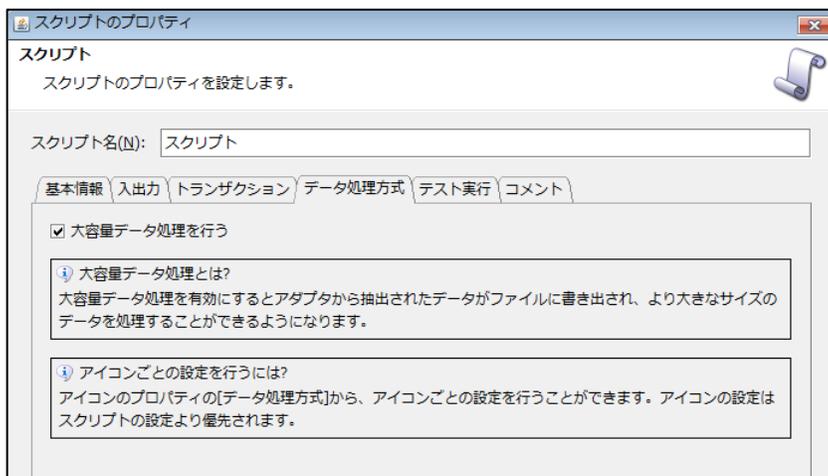
# 大容量データ処理の設定について

DataSpider Servistaに割り当て可能なメモリサイズに上限があり、ヒープサイズを十分に確保できない場合、スクリプト処理を停止させないようにする必要があります。

## 大容量データ処理のポイント

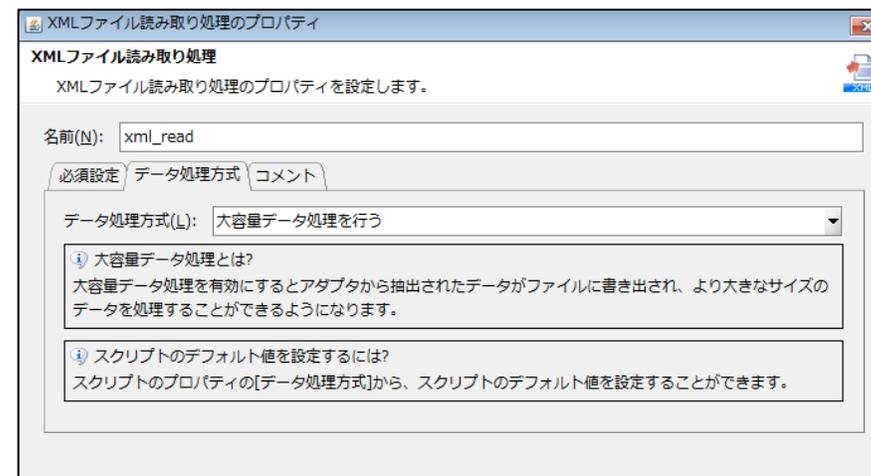
「①スクリプト全体設定」もしくは「②コンポーネント個別設定」を有効化します。

### スクリプト全体設定



or

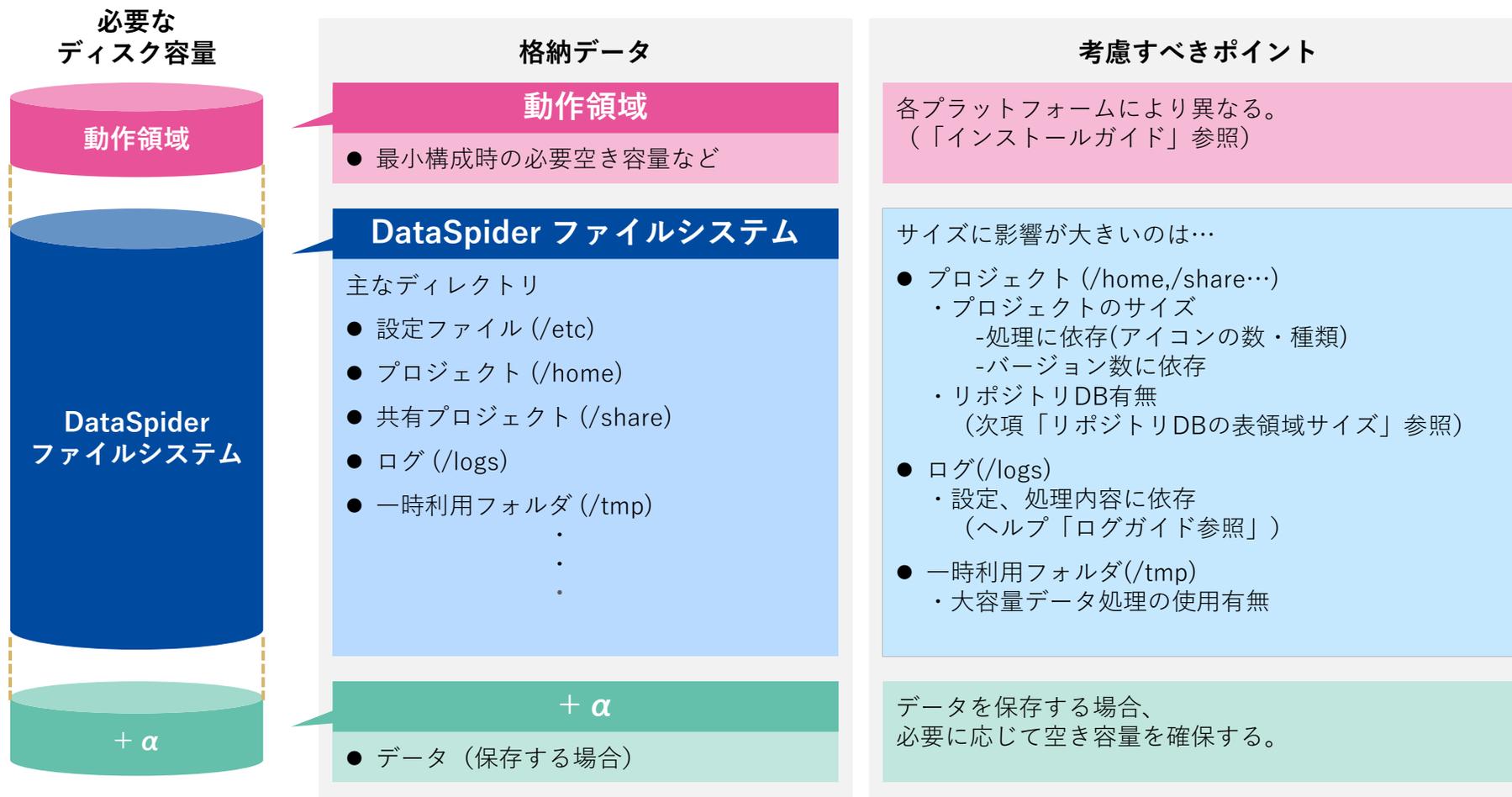
### コンポーネント個別設定



ディスクに書き出しながら処理するため、ファイルのI/Oが多発し、実行速度の低下が発生する場合があります。

# ディスクサイジングの考え方

DataSpider Servista におけるディスクサイジングの考え方は以下のとおりです。



# DataSpider Servista リポジトリDBの表領域サイズ

リポジトリDBには、ユーザ情報、プロジェクトの他、DataSpiderファイルシステム情報（※1）、DataSpider Servistaの設定情報（※2）などが格納されます。

（※1）ディレクトリ構造やファイル構成、アクセス権、実データを含む

（※2）グローバルリソース、トリガーなどの各種設定ファイル

## リポジトリDB格納対象

「/etc」 「/home」 「/share」 ディレクトリ以下



リポジトリDBのサイズに大きな影響を与えるのは「/home」「/share」以下のプロジェクトです。プロジェクトサイズは、要件による影響が非常に大きいので実環境での評価検証を推奨します。

プロジェクトサイズは、基本的にスクリプトで使用しているアイコンの種類や数に依存します。

また、各種Mapperで、どのようなロジックアイコンを使用するかによってもサイズは異なるため、要件による影響が大きく、標準的な目安を示すことは難しいといえます。実環境で必要とされるサイズについては、要件に応じたスクリプトを実際に作成したあと、プロジェクトをダウンロードしたサイズを目安としていただくことを推奨します。

## 3. 導入環境の選定

## 導入環境の選定

### ■ 単一環境と複合環境

DataSpider Servistaでは、単一環境（1つのDataSpiderServer）で開発～本番までを行う方式と、環境を複数（開発～検証～本番など）に分けて使用する複合環境の2種類の方式を選択できます。

### ■ クラウド環境とオンプレミス環境

DataSpiderServerをクラウド環境とオンプレミス環境、どちらに配置するかを選択します。

単一環境と複合環境のそれぞれのメリット、デメリットは以下のとおりです。

	メリット	デメリット
単一環境	<ul style="list-style-type: none"> <li>● サーバ1台で済むため、サーバ自体の運用負荷を軽減できる。</li> </ul>	<ul style="list-style-type: none"> <li>● DataSpider Servistaの運用が複雑になる。</li> <li>● 本番稼働後、追加開発や検証・再現調査などが、本番の処理に影響を与える恐れがある。</li> </ul>
複合環境	<ul style="list-style-type: none"> <li>● DataSpider Servistaの運用がシンプルになる。</li> <li>● 本番と切り離して追加開発・検証・再現調査を実施することが可能。</li> </ul>	<ul style="list-style-type: none"> <li>● 環境ごとに複数のサーバ、ライセンスを用意する必要がある。</li> </ul>



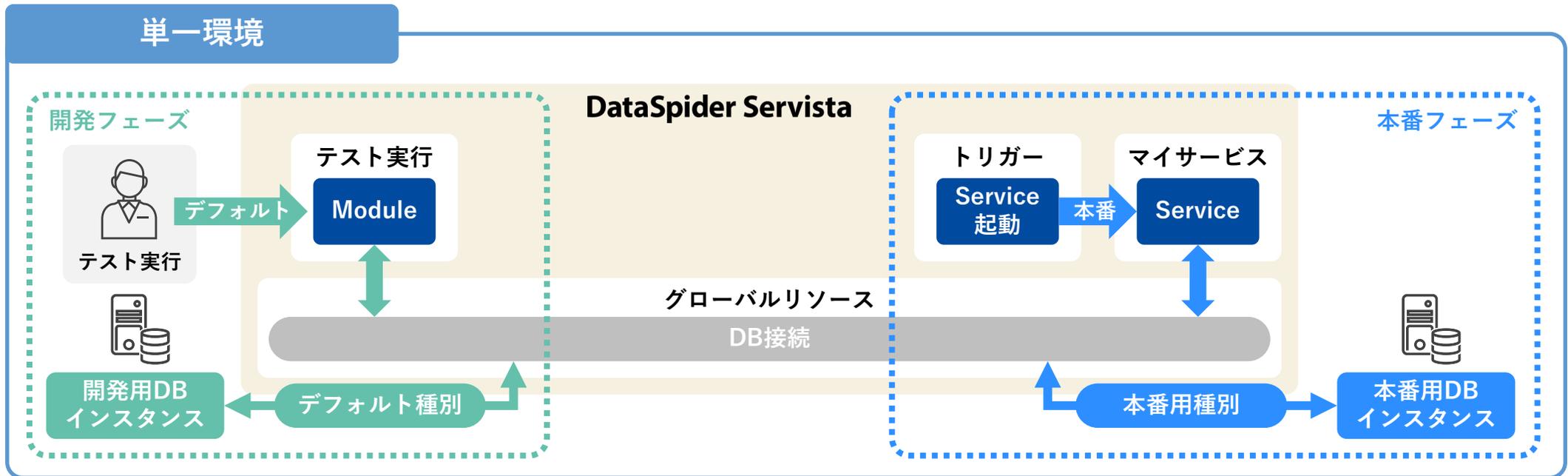
DataSpider Servistaで作成したスクリプトに十分なテスト環境を確保し、DataSpider Servistaの運用負荷を下げるために、複合環境を推奨します。

# 単一環境の運用について

## 単一環境の運用ポイント

DataSpider Servistaの「種別」機能を活用して運用します。

スクリプトの実行までを開発フェーズとし、サービスの実行（トリガー実行）を本番フェーズとして2種類の種別を使い分けるといった例になります。



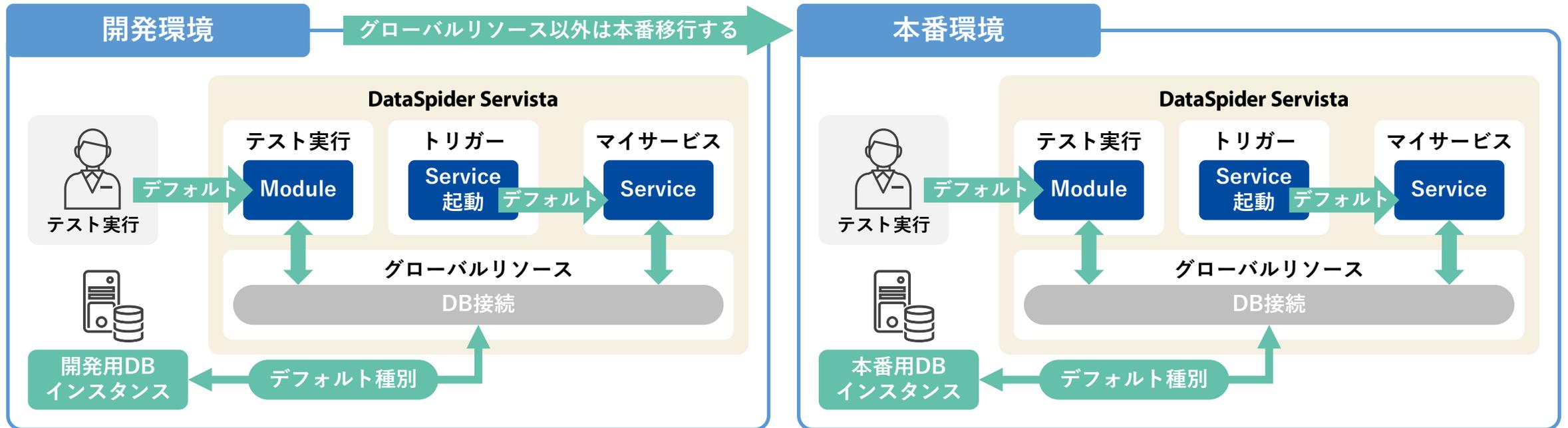
3フェーズ（開発>検証>本番）の場合は、3種類の種別が必要となるため、トリガー設定およびサービスを検証用と本番用と2元管理するなどの工夫が必要になります。

# 複合環境の運用について

## 復号環境の運用ポイント

個々の環境に同一名称のグローバルリソース設定を静的に配置し、環境個々に管理する

スクリプトに変更を加えず、各フェーズの移行が可能になります。  
この場合、**グローバルリソースは個々の環境で個別に管理し、移行対象から外す**というルールになります。



「種別」を使用するとフェーズ（開発>検証>本番）毎に手動設定が必要となるのを防ぎます。

## 4. リポジトリDB有無の選定

## リポジトリDBとは

RDB（リレーショナルデータベース）内に設定したリポジトリ領域にて、サービスやユーザ情報、各種設定データを管理する機構です。

### リポジトリDBのポイント

ユーザ情報やファイルのアクセスコントロール機能を使用する場合には必要です。

リポジトリDB	メリット	デメリット
有り	<ul style="list-style-type: none"> <li>● ユーザ、グループが作成可能</li> <li>● チーム開発機能が使用可能</li> <li>● アクセスコントロール機能が使用可能</li> <li>● アプリケーションとデータの分離が可能</li> </ul>	<ul style="list-style-type: none"> <li>● リポジトリ用DBの準備</li> <li>● リポジトリ用DBのメンテナンス要 (バックアップ、バージョンアップなど)</li> <li>● ローカルアクセスに比べて遅い (リポジトリDBが遠隔地にあるとより顕著)</li> </ul>
無し	<ul style="list-style-type: none"> <li>● シンプルな運用が可能</li> <li>● ローカル（OSファイルシステム）アクセスのため早い</li> </ul>	<ul style="list-style-type: none"> <li>● ユーザ、グループが作成不可 (rootユーザ/rootグループのみ)</li> <li>● アクセスコントロール機能が使用不可</li> </ul>



複数人で開発する場合はリポジトリDBの使用が推奨です。1～2人の少人数の場合は非推奨です。

※ リポジトリDB有り/無しの動作については[マニュアル](#)をご確認ください。

## 注意事項

### ■ リポジトリDBに関する注意事項

- リポジトリDB有りから無しへの移行はできません。（無しから有りは可能です。）
- リポジトリDBは、1つのDataSpiderServerにつき、1つ用意してください。  
複数のDataSpiderServerで同一のリポジトリDBに接続することはできません。
- リポジトリDBとして使用するデータベースのインスタンスは、リポジトリDB専用として使用してください。

### ■ リポジトリDBの設定方法

- リポジトリDBの設定は、以下の2とおりの方法があります。

DataSpider Servistaのインストール時に  
インストーラで設定する

or

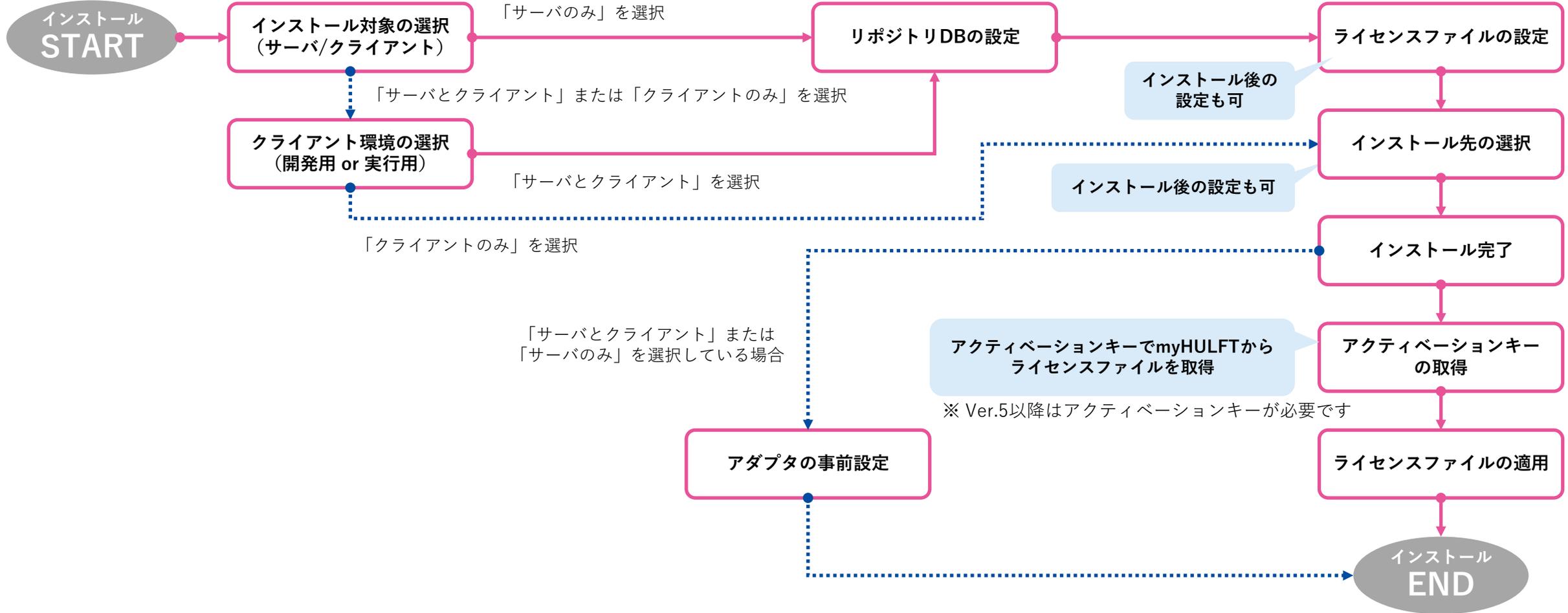
DataSpider Servistaのインストール後、  
DataSpider Studioのコントロールパネルで設定する



ポジットリDBはDataSpider Servistaのインストール時だけでなく、  
インストール後に設定することも可能です。

## 5. 導入作業

# DataSpider Servistaインストール・フロー



 インストール作業はインストールガイドをご参照のうえ、行ってください。

# DataSpider Servistaインストール前の準備

## DataSpider Servistaのインストール前に行う5つのステップ



クラスタソフト



リポジトリDB



ライセンスファイル



インストール環境



セキュリティ対策

- |   |                                 |  |
|---|---------------------------------|--|
| 1 | クラスタソフトの準備                      | <ul style="list-style-type: none"> <li>✓ DataSpider Servistaを監視するクラスタソフトを準備します。<br/>※冗長化構成をしない場合には省略可</li> </ul>   |
| 2 | リポジトリDBの準備                      | <ul style="list-style-type: none"> <li>✓ リポジトリDBとして使用するデータベースに専用のインスタンスを作成します。<br/>※リポジトリDBを使用しない場合は省略可<br/>※DataSpiderServerのインストール後に設定することも可能</li> </ul>   |
| 3 | DataSpider Servistaライセンスファイルの準備 | <ul style="list-style-type: none"> <li>✓ インストール時に指定するDataSpider Servistaのライセンスファイルを準備します。</li> </ul>   |
| 4 | DataSpider Servistaインストール先環境の準備 | <ul style="list-style-type: none"> <li>✓ DataSpider Servistaをクラウド環境に構築する場合は、クラウドインスタンスにDataSpider ServistaがサポートするOSを準備します。<br/>DataSpider Servistaをオンプレミス環境に構築する場合は、実筐体上もしくは仮想環境上にDSSがサポートするOSを準備します。</li> </ul> |
| 5 | セキュリティ対策の検討                     | <ul style="list-style-type: none"> <li>✓ DataSpiderServer や JDBC Proxy Serverなど、外部アクセスを受け付けるアプリケーションについて、セキュリティ対策を検討します。</li> </ul>  |

## アダプタの事前設定

DataSpiderServerのインストール時に、アダプタのインストールも同時に行われます。

ただし、一部アダプタについては別途ライブラリのインストールや事前設定が必要になる場合があります。

使用するアダプタが該当していないかを精査する必要があります。

### ■ ライブラリのインストールなど事前設定が必要なアダプタ

No.	アダプタ名	No.	アダプタ名	No.	アダプタ名
1	DB2 アダプタ	6	HCL Domino アダプタ	11	JMS アダプタ
2	MySQL アダプタ	7	HCL Notes アダプタ	12	Amazon Web Services アダプタ
3	Oracle アダプタ	8	SAP アダプタ	13	Microsoft Azure アダプタ
4	SQL Server アダプタ	9	Tableau アダプタ ※1	14	Salesforceアダプタ ※2
5	JDBC アダプタ	10	List Creator アダプタ		

※1 Version5.0よりライブラリのインストールは必要なくなりました。

※2 Version5.0よりコンポーネントがインストーラに非同梱になりました。



使用するアダプタが上記一覧に含まれている場合は、インストールガイドを参照のうえ、事前設定内容を確認してください。

## DataSpiderServerとDataSpider Studio

DataSpider Servistaはクライアント・サーバモデルで動作しています。

DataSpiderServerの主な役割は、処理の実行です。またDataSpider Servistaでは、クライアントのことを主にDataSpider Studio（以下、Studio）と呼びます。Studioの主な役割は、統合開発環境およびDataSpiderServerに実行指示を行うインターフェースです。



インストーラは同一です。環境が同じ場合は一度のインストールでOKですが、環境が別の場合はそれぞれインストールが必要です。

## セキュリティ対策

DataSpider Servista をより安全にご使用いただくためには、十分なセキュリティ対策が不可欠となります。

DataSpiderServer や JDBC Proxy Server など、外部からのアクセスを受け付けるアプリケーションに対しては、十分なセキュリティ対策の実施を強く推奨いたします。

ここでは代表的なセキュリティ脅威と、その対策の一例をご紹介します。

### ウイルス対策

ウイルス対策アプリケーションを利用して、ウイルス感染などのセキュリティ脅威からサーバを保護します。

- ウイルス定義ファイルを常に最新版に更新する
- 定期的にウイルススキャンを実行し、ウイルスを検知・駆除する

### ネットワークセキュリティ

外部からの不正なアクセスによる、情報漏えいや改ざんといった攻撃からサーバを保護します。

- サーバにアクセス可能なIPアドレスを制限する
- IDS/IPSを設置し、不正な通信を検知・遮断する
- Webアプリケーションファイアウォールを設置して、脆弱性を悪用した攻撃から防御する

### ログの収集・監視

予兆検知によりセキュリティ事故を未然に防ぎ、発生した場合の事後調査を円滑にします。

- ログ監視ツールを導入し、パケットログやイベントログ（syslog）、Tomcatのログファイルなどを監視する
- DataSpider Servista のアクセスログ機能を有効にし、監査証跡を残す

### 運用面でのセキュリティ

適切な運用により攻撃者からの不正行為を抑止し、サーバを保護します。

- OS や DataSpider Servista の管理者権限を不必要に与えない
- 定期的にパスワードを変更する



< 免責条項 >

本資料の内容は、資料作成時点の当社の判断に基づいて作成されているものであり、今後予告なしに変更されることがあります。よって本資料使用の結果生じたいかなる損害についても、当社は一切責任を負いません。

また、本資料の無断での複製、転送等を行わないようお願いいたします。

なお、本資料に記載されている会社名、製品名は各社の商標または登録商標です。