INTHEFOREST

機械学習のためのデータ処理基盤入門セミナー

株式会社INTHEFOREST　代表取締役　冨田和孝
Mail: sales@intheforest.co.jp / Tel:03-5848-2424
〒176-0023 東京都練馬区中村北1-13-13 OHD練馬ビル501

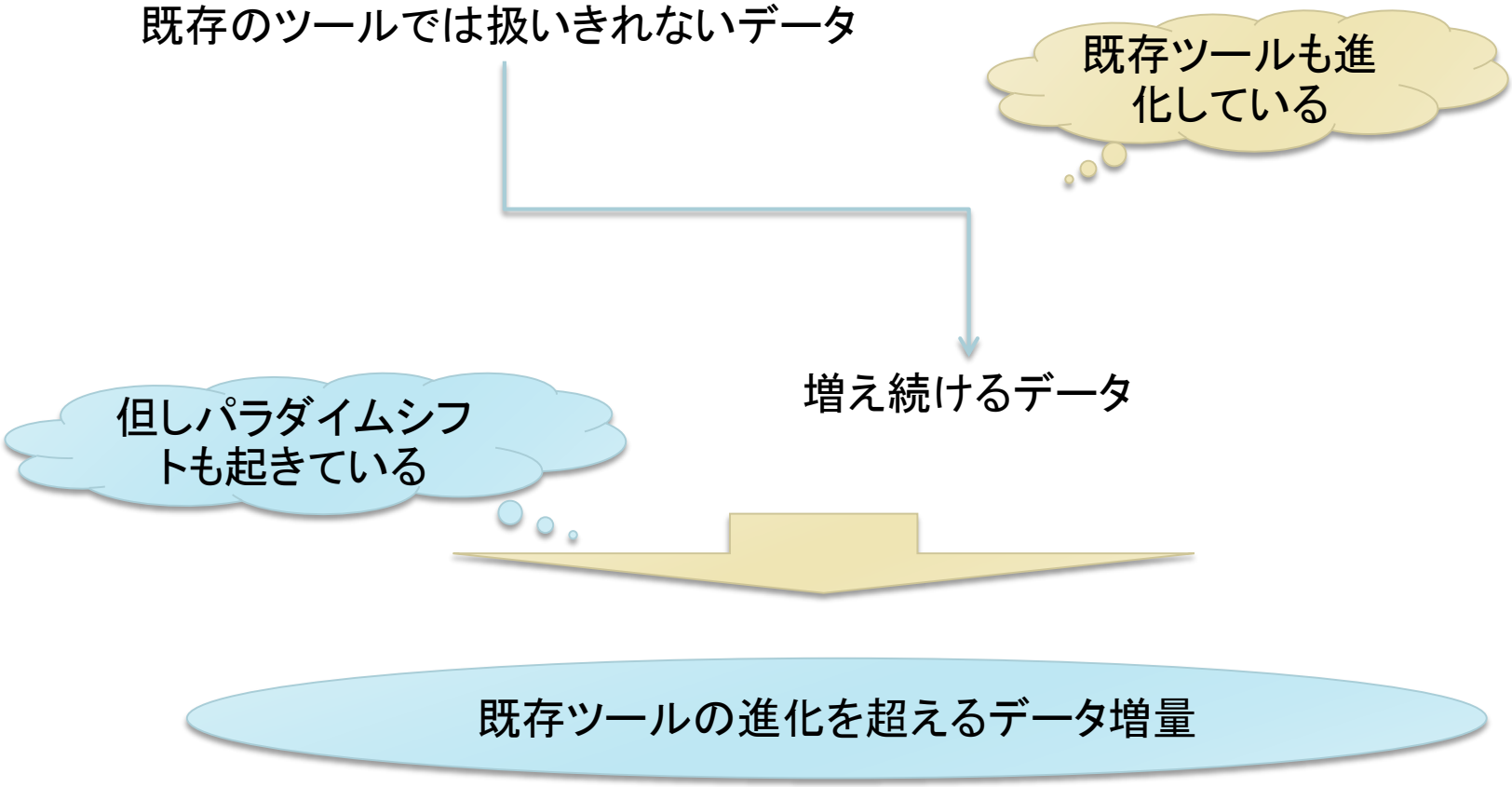
Part1
並列分散処理入門

Agenda

- イントロダクション
  - データサイズ：ペタバイトの時代
  - 並列処理とは？
  - 並列処理と分散処理
  - （参考）Pythonライブラリを利用することの問題点
- 並列分散処理入門
  - Sparkとは？
  - Sparkを用いた分散処理（hadoop+Spark）
  - 分散データベース（Cassandra）
- Q＆A

イントロダクション

データサイズ：ペタバイトの時代



データサイズ：ペタバイトの時代

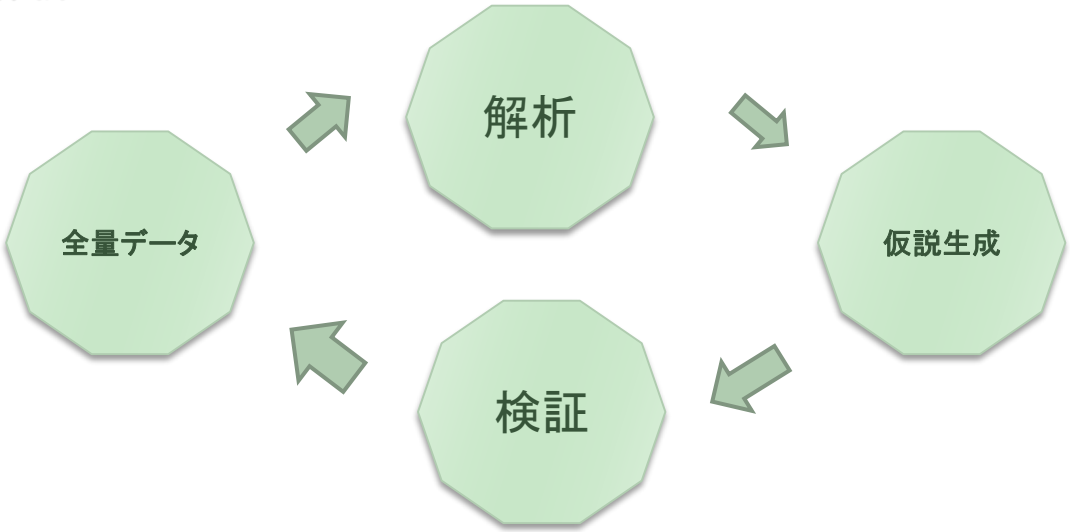
データレイク

多数のソースからのデータを元のままの多様な形式で保持する中央ストレージリポジトリ

データサイズ：ペタバイトの時代

従来のデータ解析

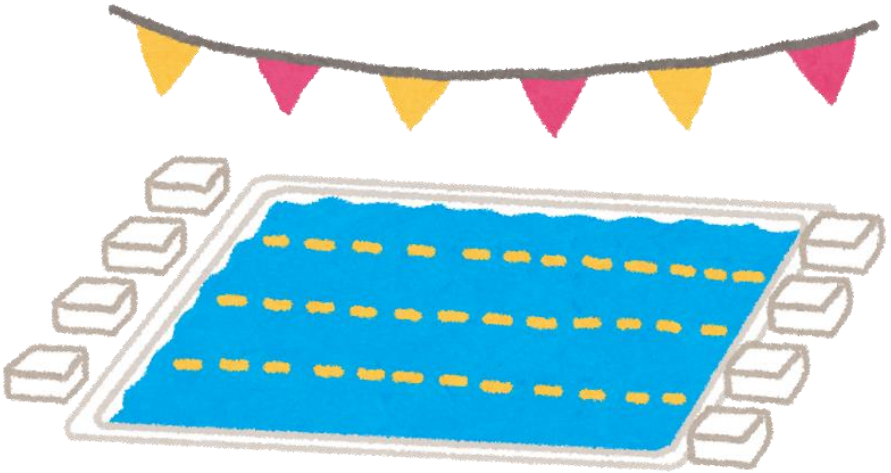
最近のデータ解析



並行処理とは

シリアルとパラレル

プールに水を入れる場合に太い一本のホース
と多数のホースとどちらが早いか



並列処理と分散処理

単独ノードと複数ノードで処理を行う場合に考えること

引用：分散システムについて語らせてくれ
https://www.slideshare.net/kumagi/ss-78765920
NTT Tech Conference #2

並列処理と分散処理

データレイクと分散処理

増え続ける大量データとの戦い

（参考）Pythonライブラリを利用することの問題点

Pythonは並行処理に向かない

一般的に用いられるPython（Cython）には並行実行に対する制限がある

この制限のため、NumpyもPandasも並行処理時には性能を発揮できない。
※制限回避のための方法はいくつかあるがそれなりにハードルは高い

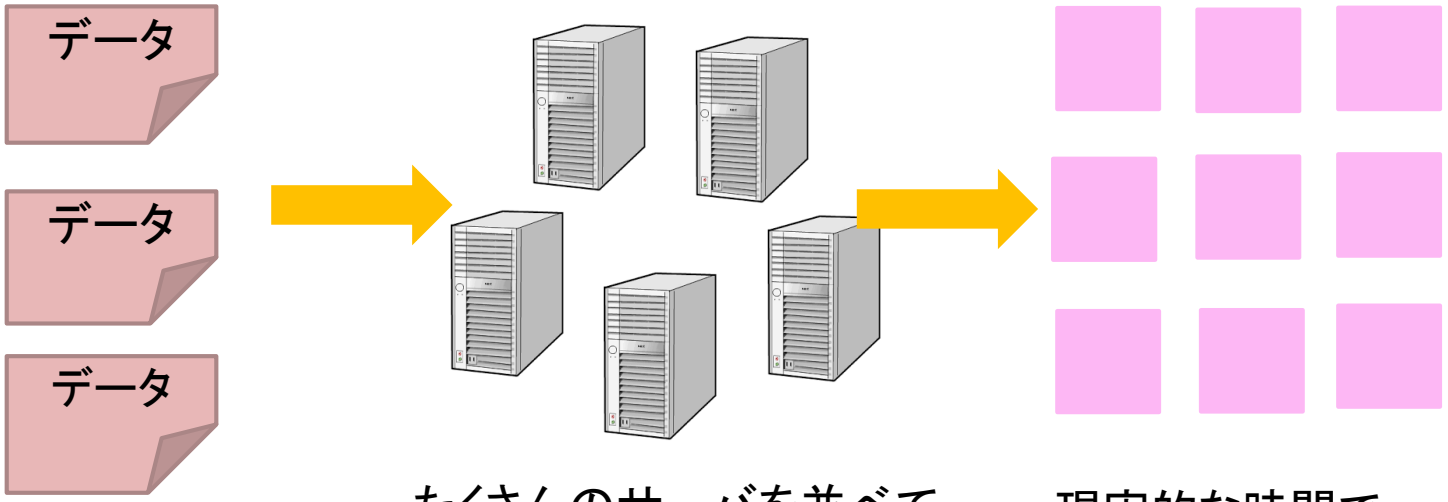
→ どんな高度なハードウェアを使っても単体のサーバーで
Pythonを使っている限り高速化は難しい

並列分散処理入門

Sparkとは

Apache Sparkとは

MapReduceに限らず、DAG(有向非循環グラフ)型で
柔軟に並列分散処理を実行できるエンジン



大量のデータ

たくさんのサーバを並べて
並列分散

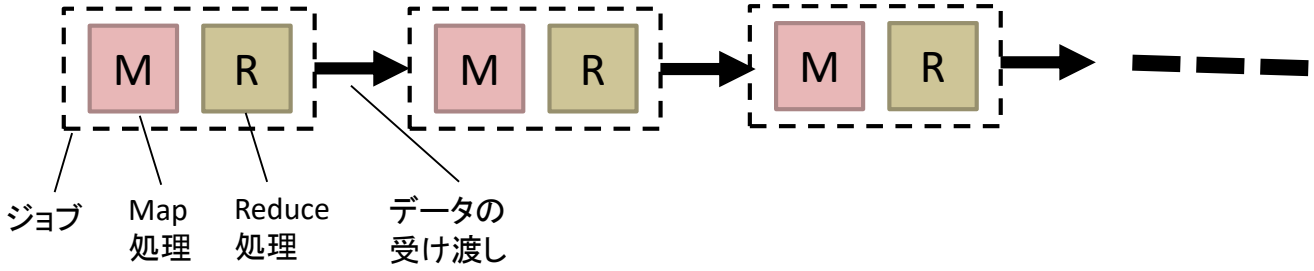
現実的な時間で
目的の処理結果を得る

データ管理には向かないが、データ処理に向いている

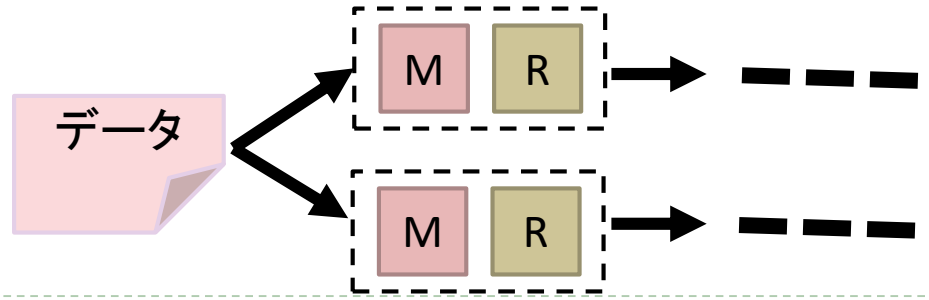
MapReduceフレームワークの課題

Hadoopが普及にするにつれて、次のような場合の
MapReduceフレームワークの処理効率が課題になってきた

①ジョブが多段に構成される場合

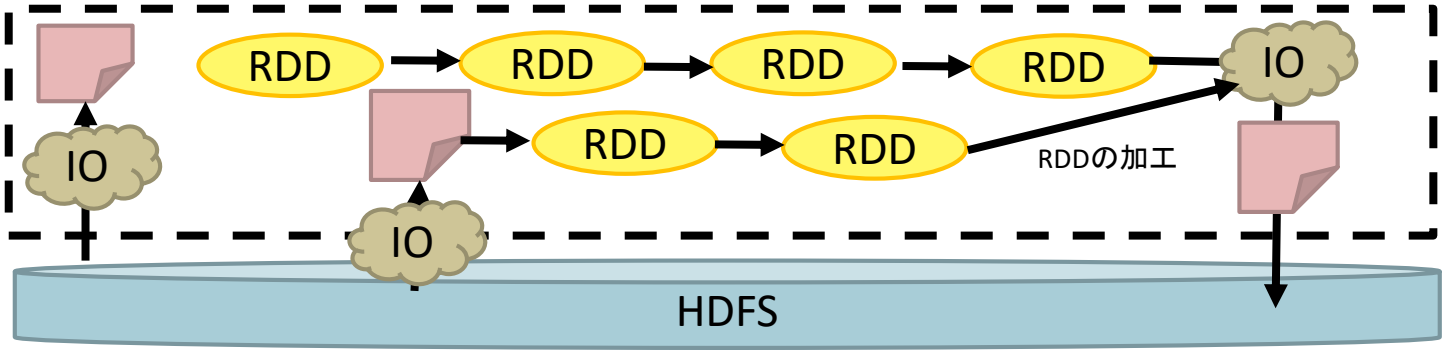


②複数のジョブで何度も同じデータを利用する場合

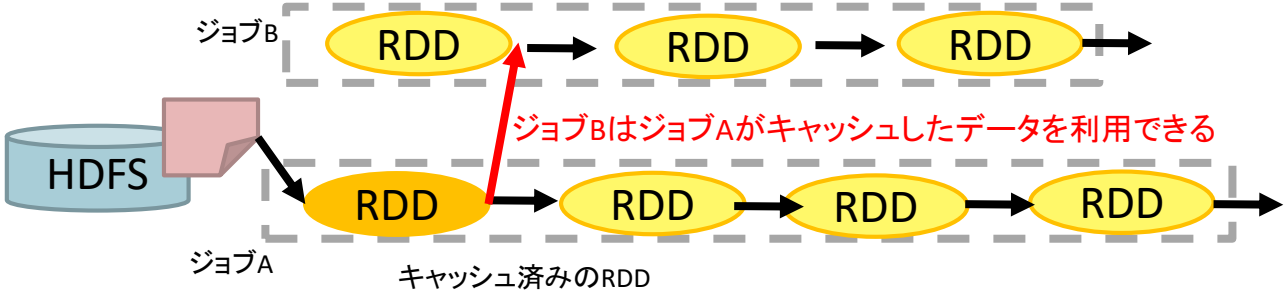


Sparkでは

①ジョブが多段に構成される場合
　複雑な処理を少ないジョブ数で実現できる

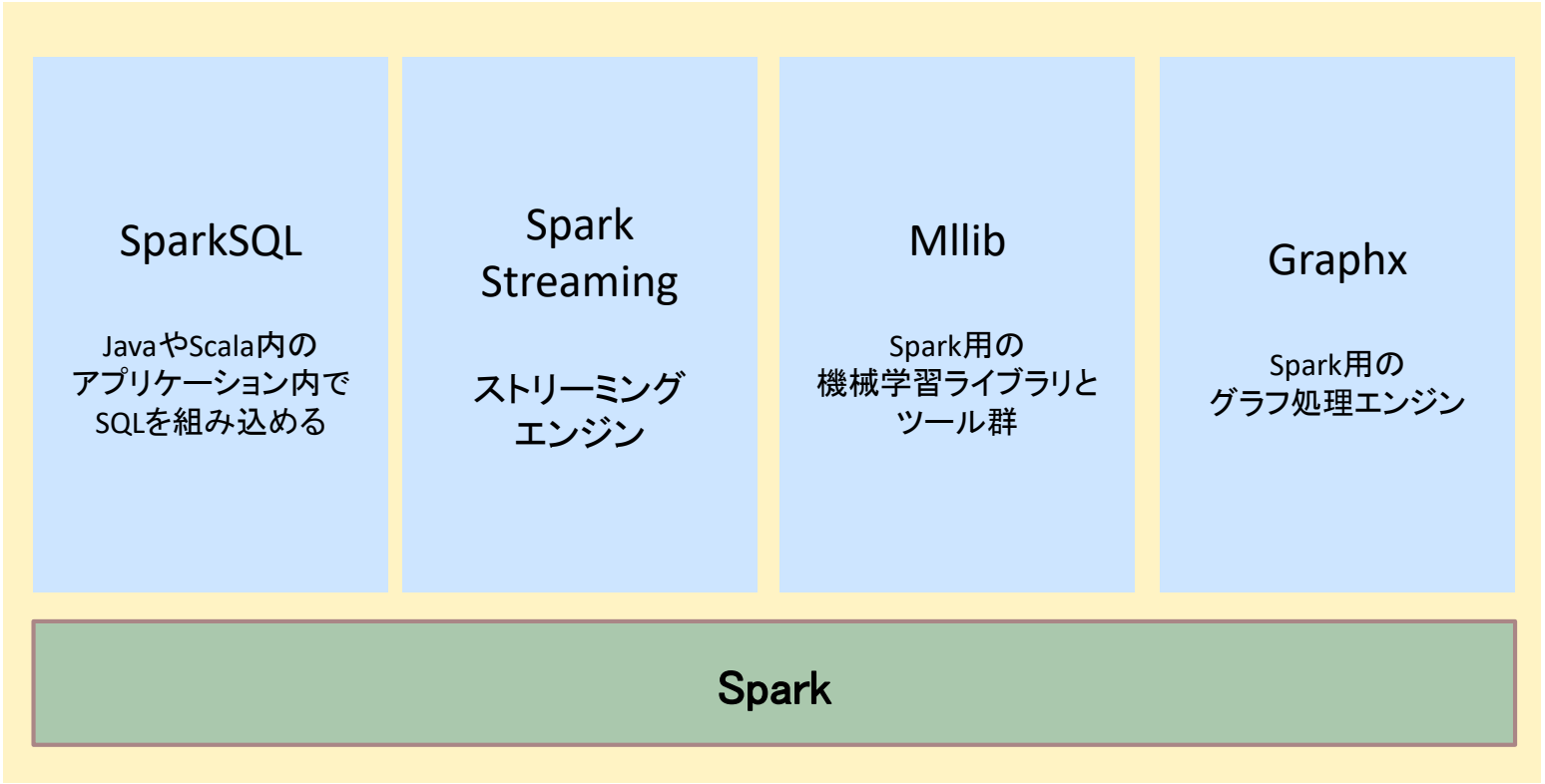


②複数のジョブで何度も同じデータを利用する場合
　何度も利用するRDDは複数のサーバのメモリに分割してキャッシュできる

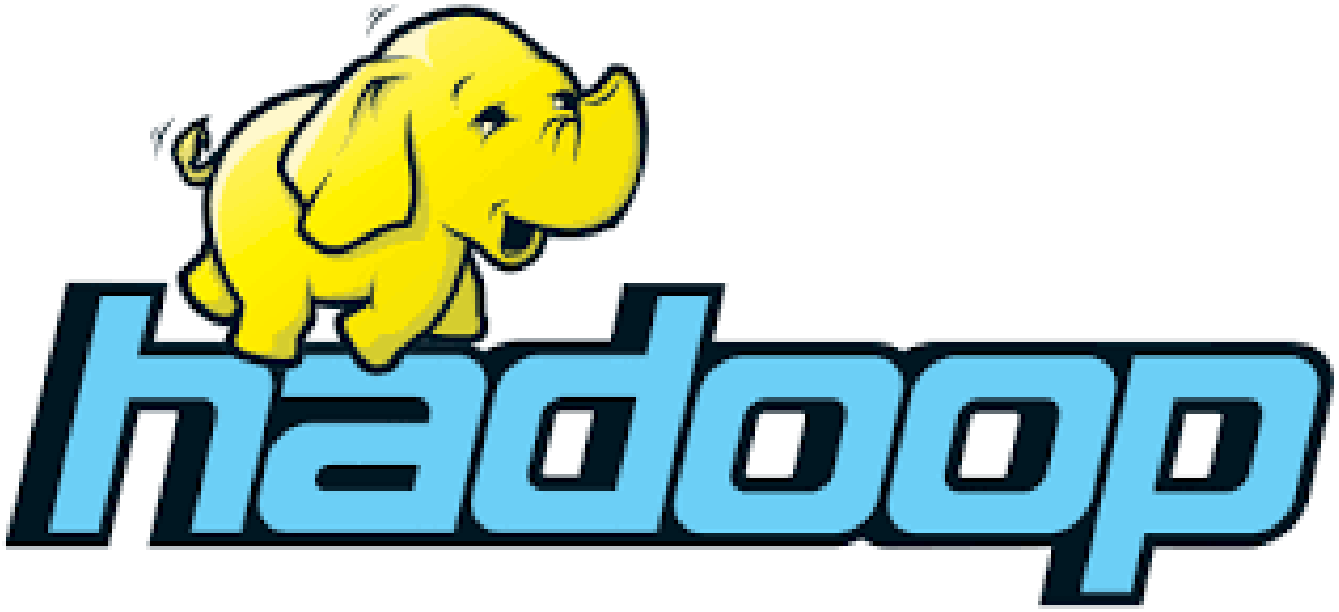


Sparkとは

Sparkは大量データをスキャンするのが得意



Hadoopとは



Hadoopとは

Hadoopとは、大規模データの蓄積・分析を
分散処理技術によって実現する
オープンソースのミドルウェアです。

- 複数のIAサーバを束ねて
  一つの大きな処理システムとして利用

- 特に大量データの格納・処理に最適化
  HDDは今でも実質は 80MB/sec 程度が限界
  この問題を解決するために並列分散処理を活用

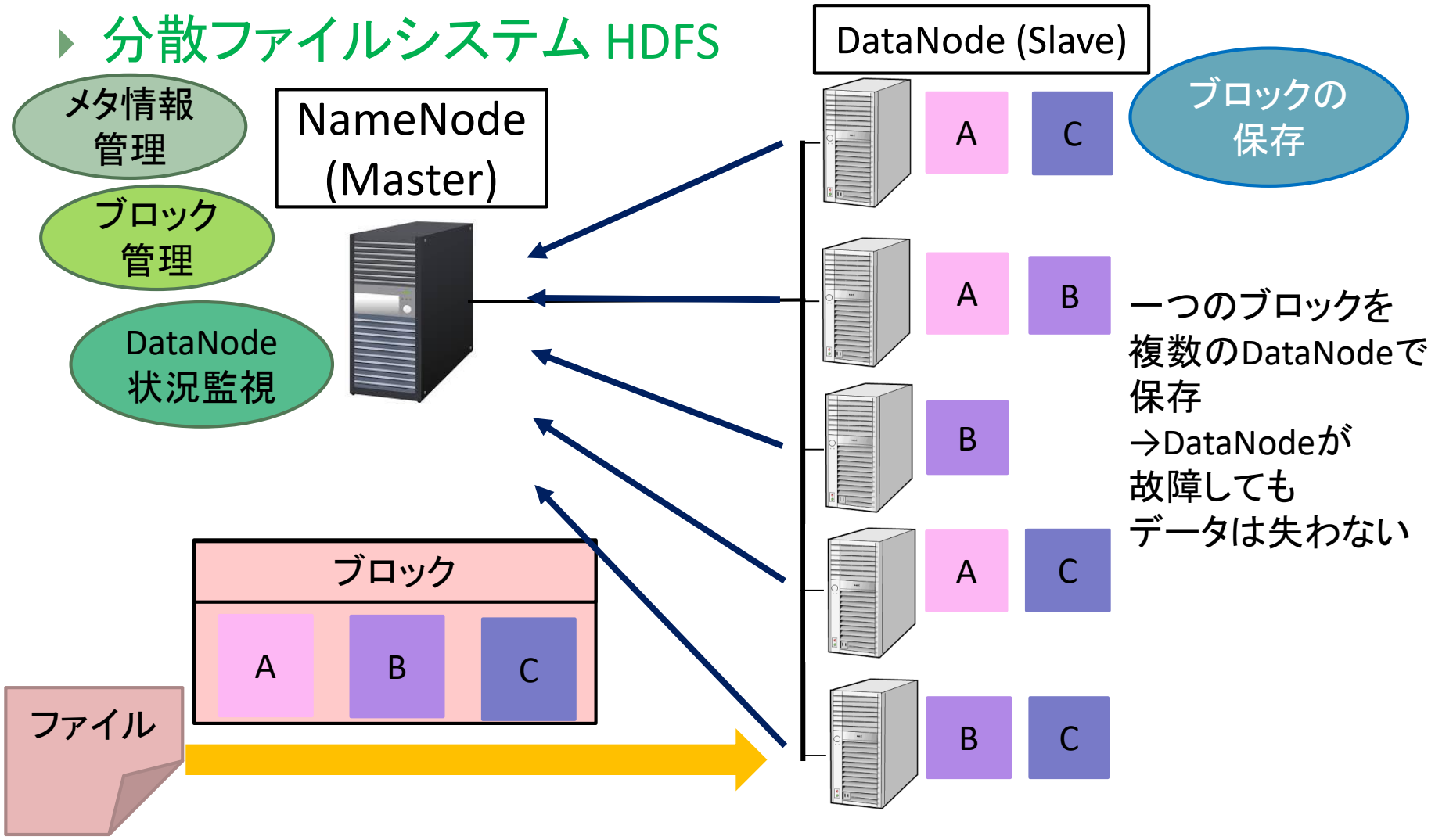
Hadoopとは

データを複数のサーバに分割して格納
利用時には、複数のサーバから、
それぞれのデータを同時に読み込む

- HDDの台数分だけのスループットを確保
  1台だと 80MB/sec 程度でも
  1000台だと 80GB/sec のスループット
  5TBのデータを読み込みも62.5秒で実現

Hadoopとは

- 分散ファイルシステム HDFS



Hadoopとは

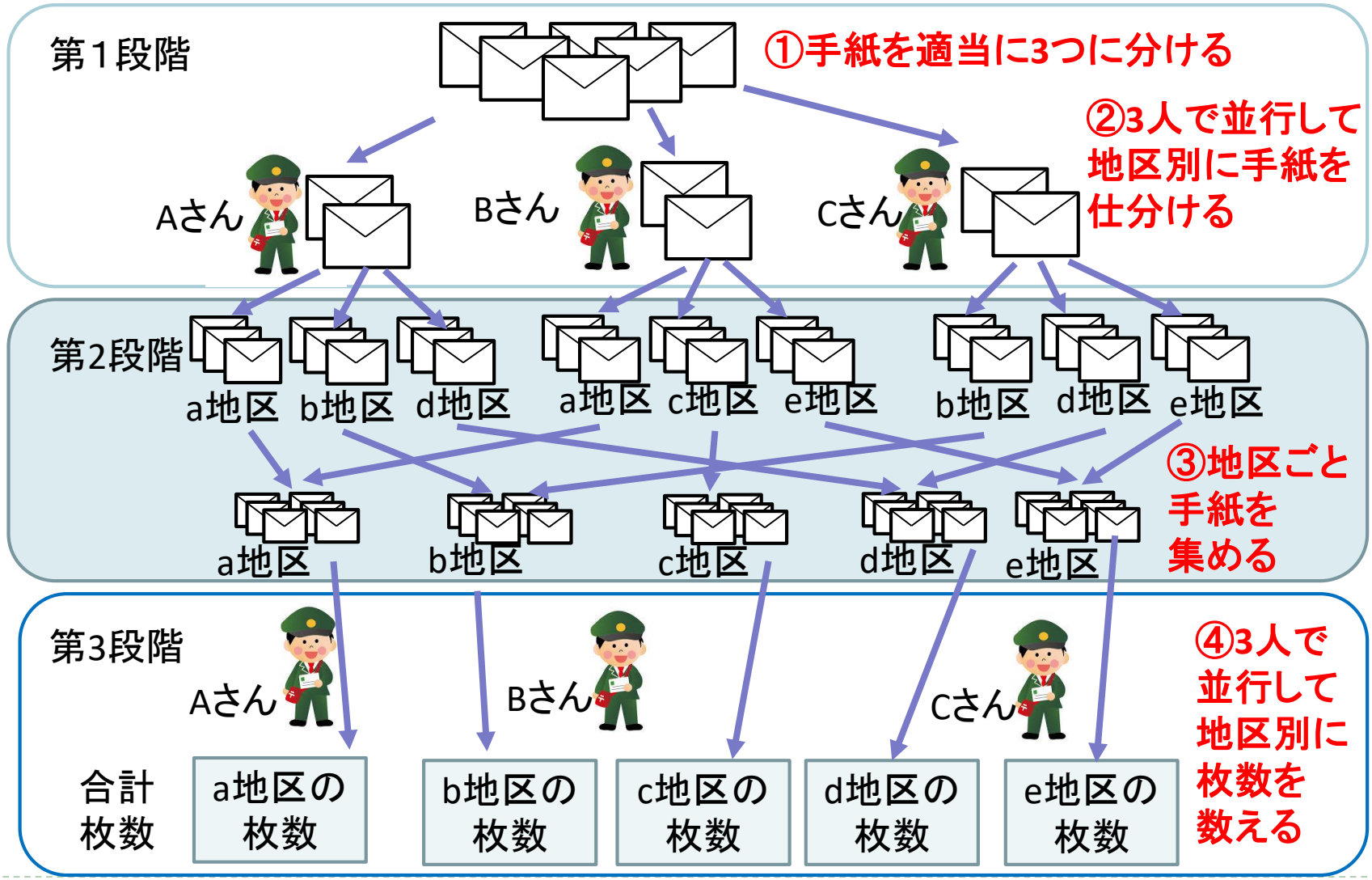
- 並列分散処理フレームワーク

  MapReduce Framework

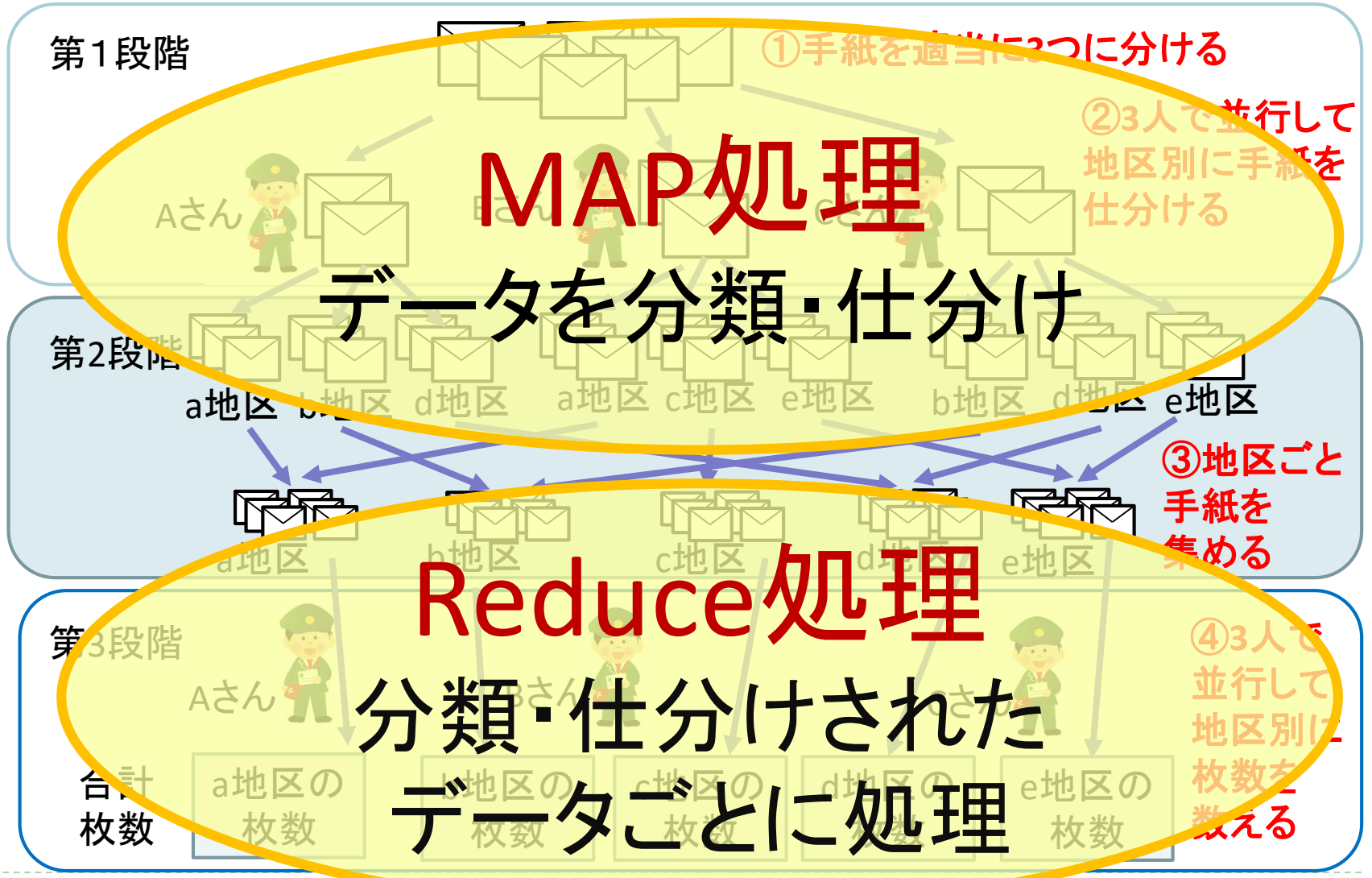
  アルゴリズムMapReduceを実現する

- Map処理、Reduce処理のみを指定すれば
  あとはフレームワークが並列分散処理を実現
- ノード数を増やせば、基本スケール
- サーバが故障しても、ジョブは実行される

Hadoopとは –MapReduce(アルゴリズム)

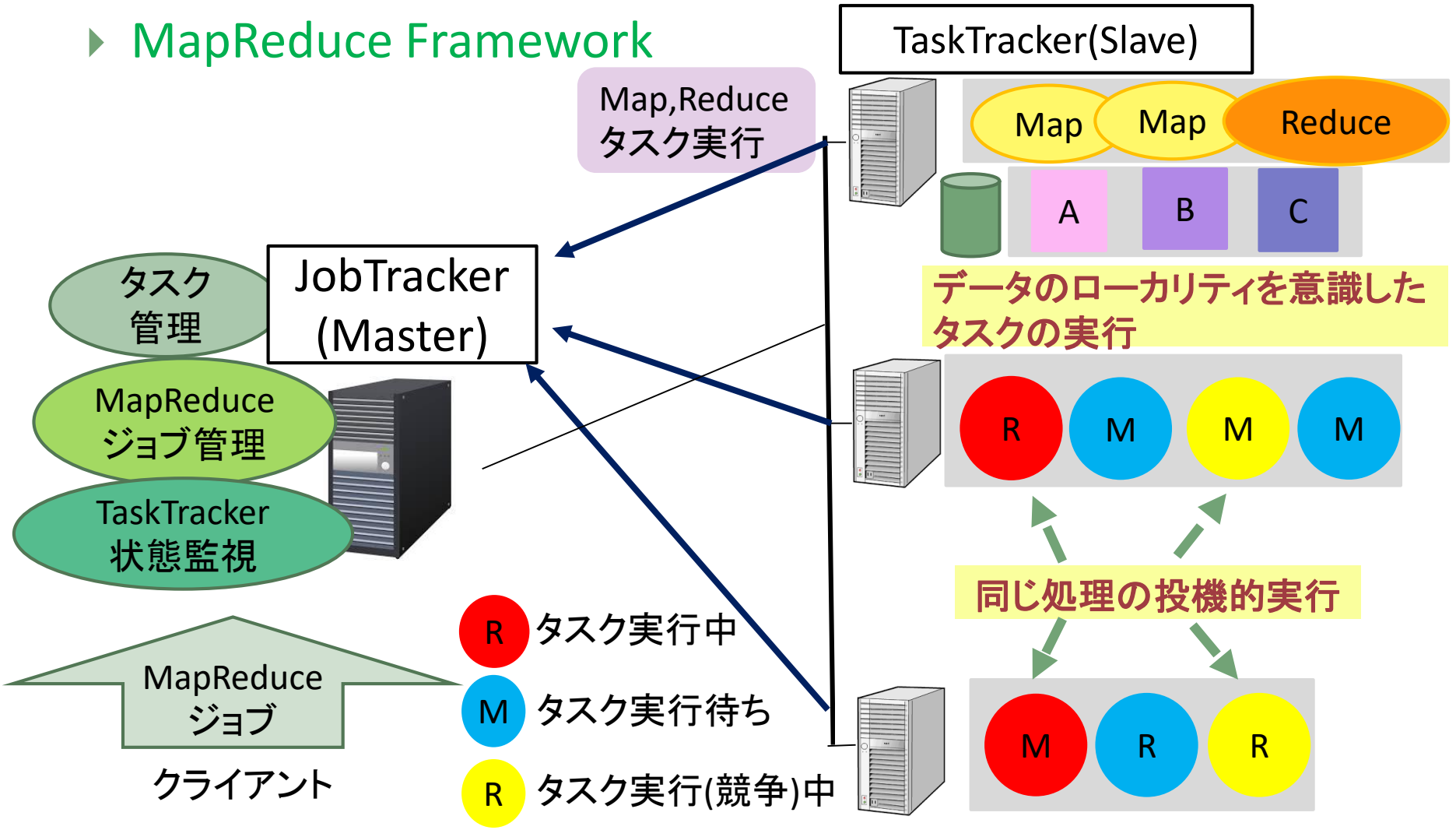


Hadoopとは –MapReduce(アルゴリズム)



Hadoopとは

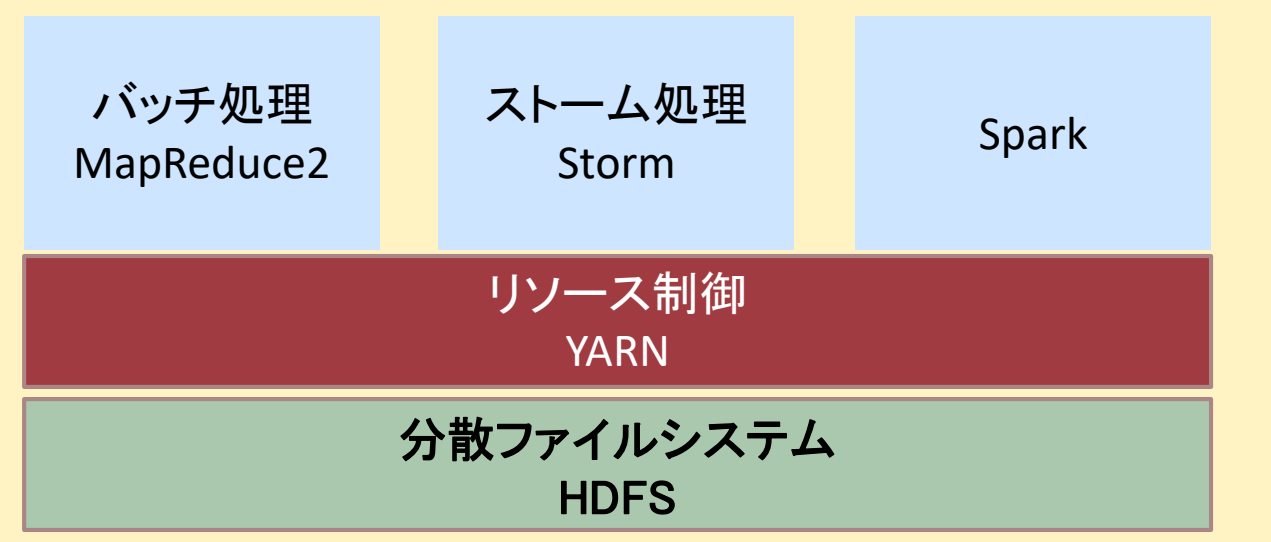
- MapReduce Framework



Hadoopとは

- YARNの導入

  並列分散処理フレームワークから
  リソース管理のレイヤを切り出された



Cassandraとは

Cassandraとは

Apache Cassandraとは

Apache Cassandraとは、

今までビッグデータ解析ビジネスで幅広く活用されてきた**Hadoop**

（大規模データの蓄積・分析を分散処理技術によって実現するオープンソースのミドルウェア）

を用いた**DB**に替わり、北米では飛躍的にその地位を占め始めているスケーラブルな

オープンソースの非リレーショナルデータベースで、継続的可用性、

リニアなスケールパフォーマンス、シンプルな操作性、

複数のデータセンターやクラウド利用領域にわたる容易なデータ分散を

実現できる**DB**です。

Cassandraとは

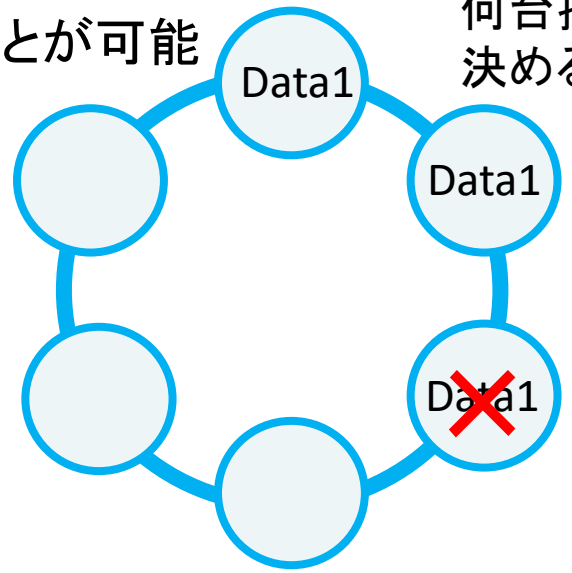
- マスターという概念がなく、
  すべてのノードが完全に同じ機能を持つ

マシンが故障しても
特定のマシンが特定の機能を
持っているわけではないので、
他のデータのあるマシンが
同じ作業を、全く問題なく処理することが可能

↓

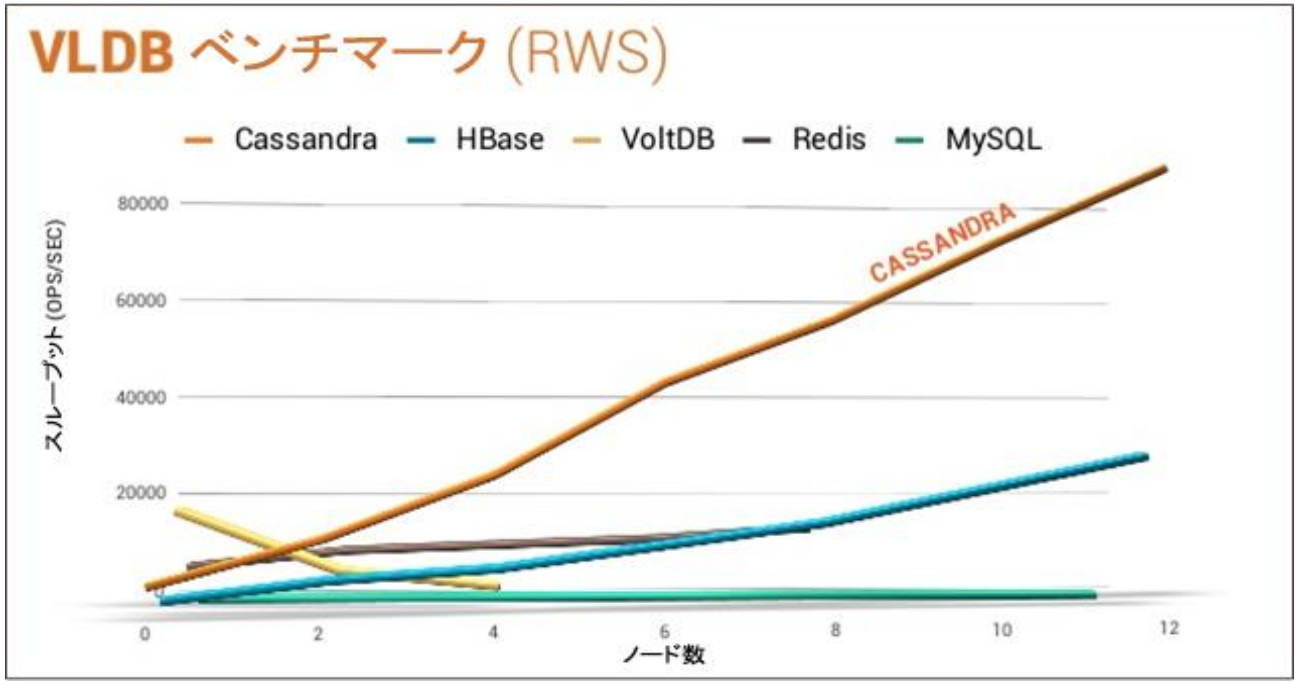
単一障害点がない

レプリケーション
(データのコピー)は
何台持たせるかを
決めることができる



Cassandraとは

- 無償かつオープンソース
- 非常にパフォーマンスが高く、適切なユースケースでは、ほぼ線形で水平にスケーリング



Cassandraユーザー事例

- Apache Cassandraのユーザー事例としては？

  日本国内、海外含めると実に多岐多様な業種/分野で使用されています。

ストリーミングメディア	工場	映像	電子商取引
キャリア	銀行	自動車	SNS
検索エンジン	旅行	出版業	CRM

Cassandra向きな事例

- 『モノ』のインターネット用アプリケーション：
  数多くの異なる場所に分散設置された端末やセンサー、
  その他類似の機器から受信する膨大な量の
  高速データ処理に最適として稼働中。

- ユーザーアクティビティのトラッキングとモニタリング：
  数多くのメディア並びにエンターテイメント企業で、
  ユーザーが映画や音楽、ウェブサイト、
  オンラインアプリケーションを利用する際のアクティビティを、
  Apache Cassandraを使用したトラッキングしモニターとして
  稼働しています。

Cassandra向きな事例

- メッセージング：
  Apache Cassandraは非常に数多くの携帯電話ならびに
  メッセージングプロバイダ用アプリケーションのデータベース
  として活用されています。

- ソーシャルメディアの解析ならびにレコメンドエンジン：
  多くのオンライン会社、ウェブサイト、
  ソーシャルメディアプロバイダが、Apache Cassandraを使用して、
  データの収集や分析、また顧客に対し分析結果とレコメンドの
  提供を行なっています。
  代表的な例として『Facebook』などがあります。

Cassandra向きな事例

- 時系列ベースアプリケーション:
  Apache Cassandraは高速書き込み能力、
  ワイド・ロー設計、リードオンリー列を備えているため、
  時系列ベースのアプリケーションに
  最適なDBとして活用されています。

Part2
リソース管理・並列分散処理
OSS紹介

Agenda

- なぜ新しいアーキテクチャが必要か？
- テクノロジーレイヤと各OSSの位置付け
- 各OSSの使い分けのポイント、特徴や強み・弱み

Cassandraの普遍性

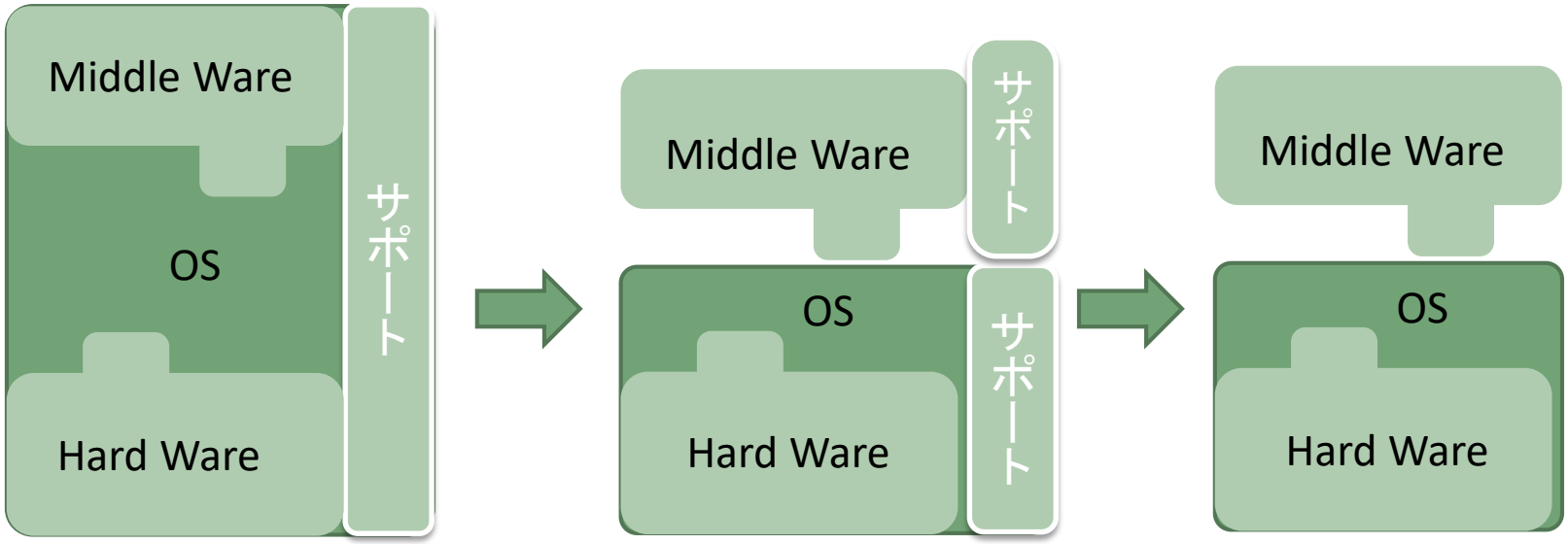
IoT向けデータベースとしてのCassandraの特徴

- 書込みに強い。
  - 書込み先が分散化されているので同時多数書込みに強い（秒間100万書込み等）
  - 結果整合性による柔軟な書込み精度を選択可能
- 解析ツールとの親和性
  - 多彩なドライバ（ODBC、JDBC、PHP、Ruby、Perl等）
  - Apache Hadoop、Apache Spark、Presto等の多彩な解析ツールを利用可能
- マルチベンダー
  - Windows、Linux、各種クラウド、JVMが稼働すれば使用する事が出来ます。Windowsでの採用実績もあります。

なぜ新しいアーキテクチャが必要か？

システム構築の変遷

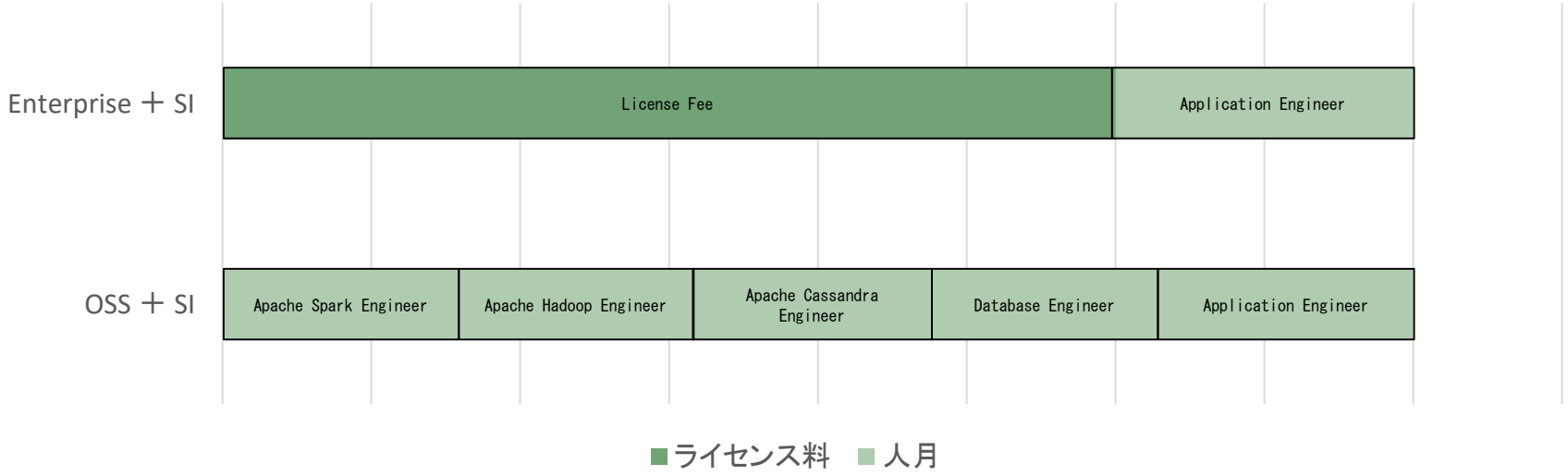
EnterpriseからOSSへ



なぜ新しいアーキテクチャが必要か？

OSSとコストパフォーマンス

OSSを使用しても安定稼働させる為には精通したEngineerの確保が必須。
その為、総額のコストはEnterprise Production を用いてもさほど費用が変わらない場合が多い



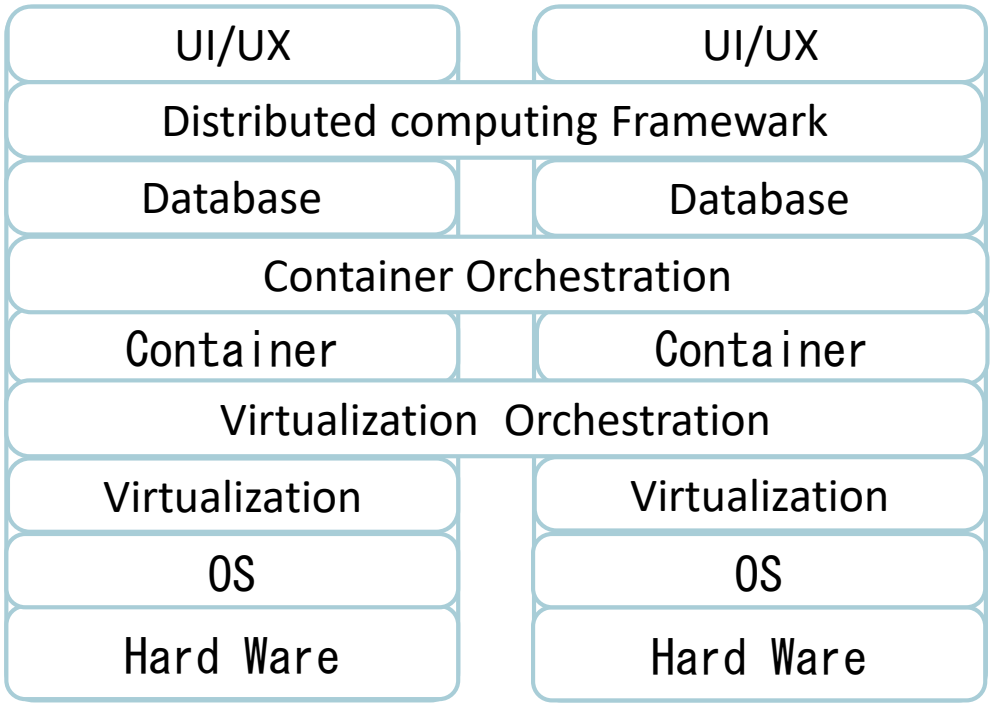
なぜ新しいアーキテクチャが必要か？

システム適用への最適化と費用対効果の精査

- システム化の目的の明確化
- 機能要件の精査
- システムの費用対効果の明確化
- 適切なサイジング

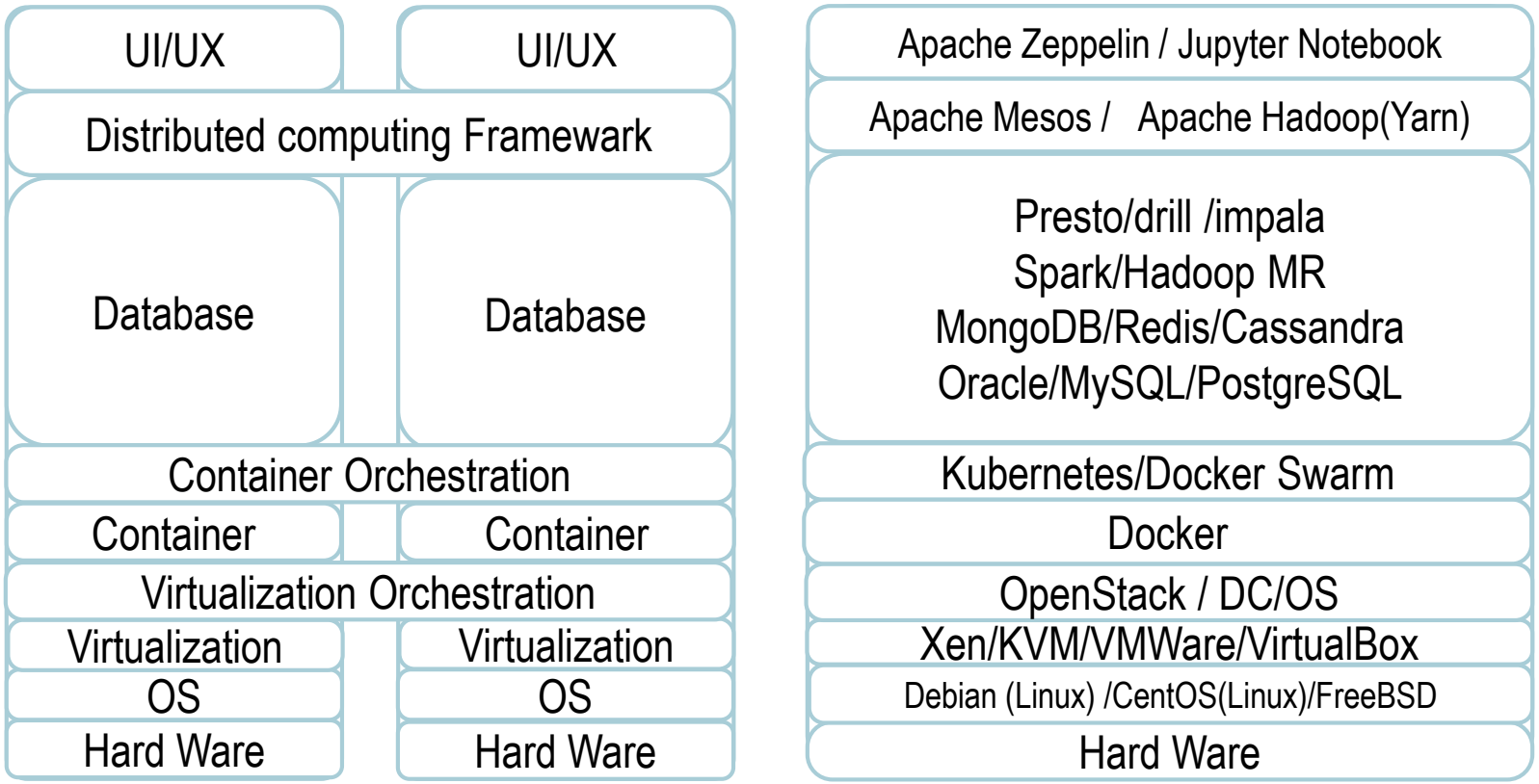
テクノロジーレイヤと各OSSの位置付け

基盤とレイヤー



テクノロジーレイヤと各OSSの位置付け

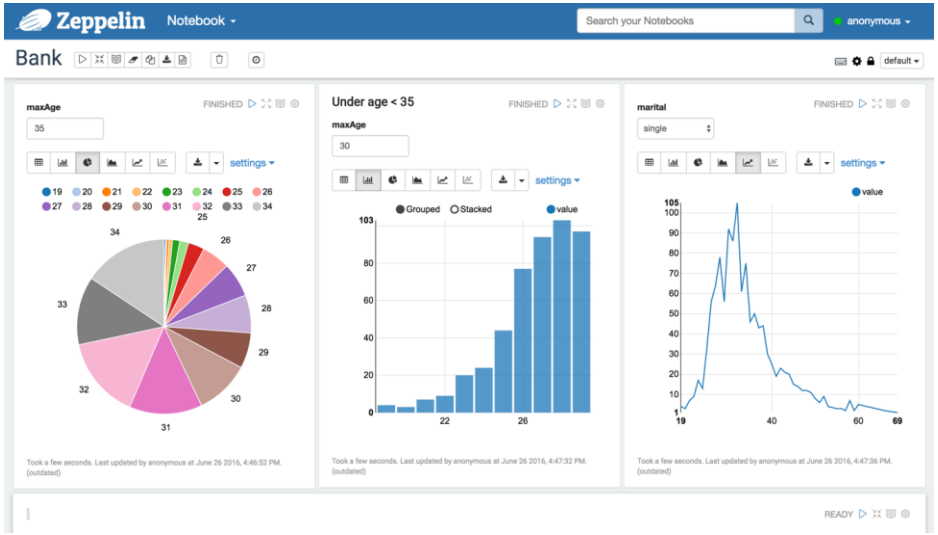
レイヤーとアプリケーション



UI/UX

Apache　Zeppelin

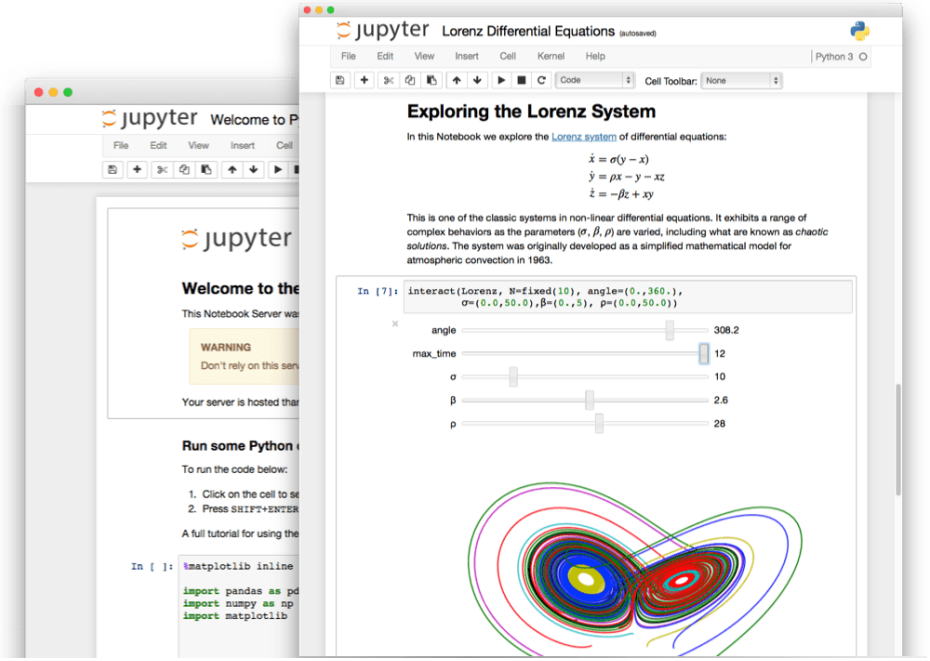
QUERY Interface/グラフ化Engine



UI/UX

Jupyter Notebook Produced by The Jupyter Project

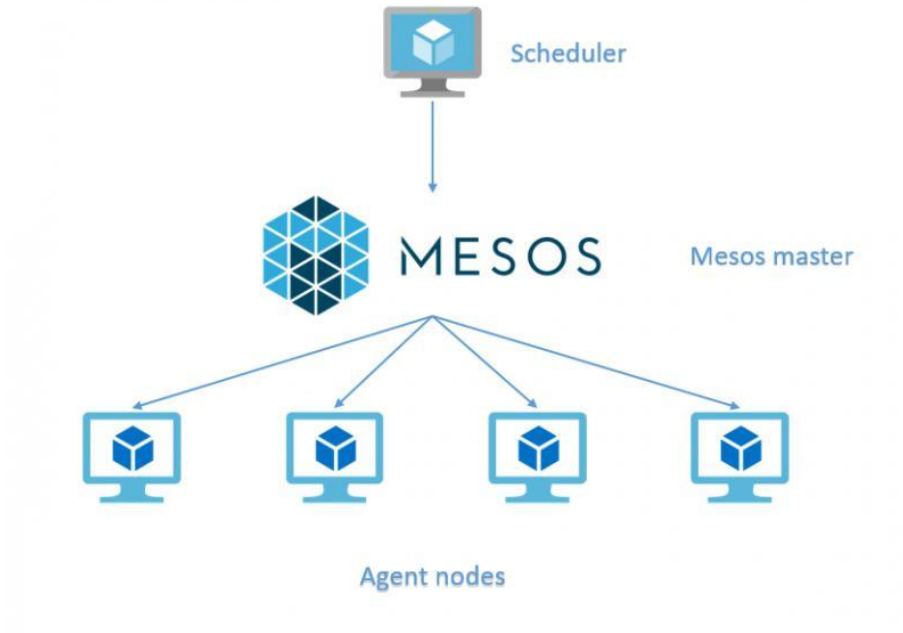
QUERY Interface/グラフ化Engine



Distributed computing Framework

Apache Mesos

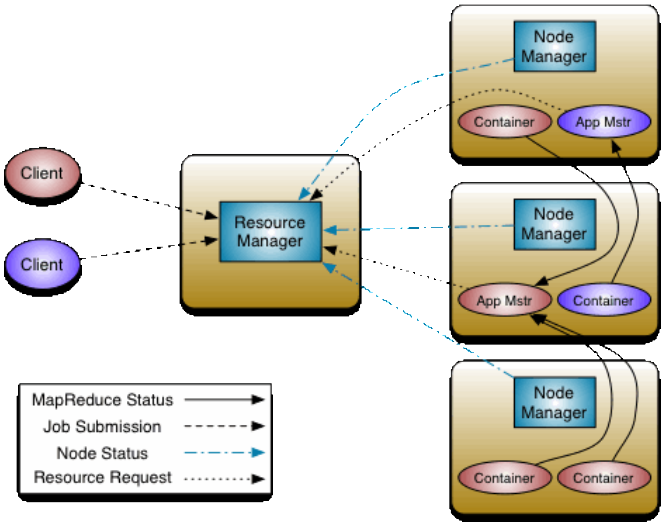
分散システムカーネル



Distributed computing Framework

Apache Hadoop Yarn

ジョブスケジューリングとクラスタ資源管理のためのフレームワーク

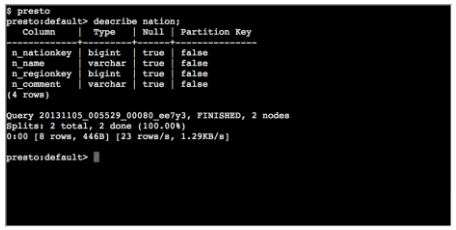


Database

Interactive Query Execute Engine

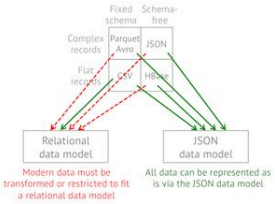
presto

producted by Facebook



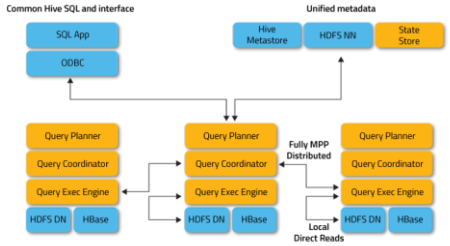
インタラクティブな分析クエリを実行するためのオープンソースの分散型SQLクエリエンジン

APACHE DRILL



Hadoop、NoSQL、クラウドストレージ向けのスキーマフリーSQLクエリエンジン

Apache Impala



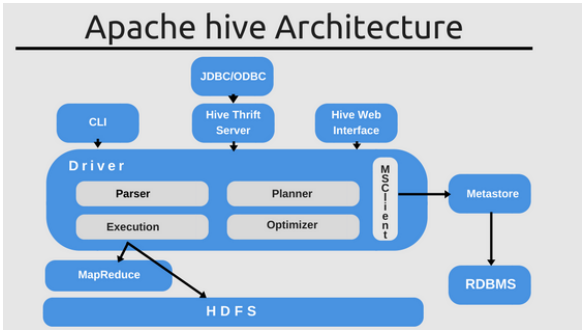
Apache Hadoop用のオープンソースのネイティブ分析データベース

Database

Query Execute Engine

構造化データを扱うためのApache Sparkのモジュール

Apache Hadoop Hive



Apache Hadoop用のオープンソースのQuery実行エンジン

Database

Library for Distributed System

Apache Spark

分散環境においてのデータ構造化ライブラリ＋ツール

Apache Hadoop

分散環境においての各種ツール群

Database

NoSQL

Apache Cassandra

P2P型分散データベース

MongoDB

Prodacted by MongoDB

非構造データをそのまま扱えるドキュメントデータベース

Redis

Prodacted by RedisLabs

永続化機能を備えたインメモリデータベース

Database

RDBMS

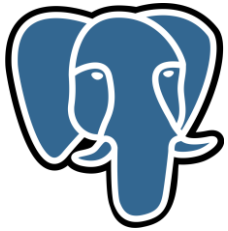
MySQL

Prodacted by Oracle

オープンソースで公開されているRDBMS

PostgreSQL

Prodacted by The PostgreSQL Global Development Group



30年以上続くOSSのRDBMS

Container Orchestration

Kubernetes

Prodacted by Cloud Native Computing Foundation

Docker Swarm

Prodacted by Docker Inc.

Googleが開発したDocker向けコンテナオーケストレーションツール

Docker社がKubernetes対抗として開発したDocker向けコンテナオーケストレーションツール

Container

Docker

Prodacted by Docker Inc.

Docker社が開発したLXC用統合ツール
chroot、cgroup、Namespaceを基本とする環境作成ツール

Virtualization Orchestration

Open Stack

Prodacted by OpenStack Foundation.

米Rackspace社が開発した仮想環境用統合管理ツール。PrivateCloudを作成することが可能

Virtualization Orchestration

DC/OS

Prodacted by Mesosphere

米Mesosphere社が開発したApache Mesosをベースとした統合リソース管理システム
Datacentar as a Serviceを標榜する。

Virtualization

Prodacted by citrix

OSSとして開発された
最初の仮想環境

Linuxのカーネルに組
み込まれた仮想環境

データ基盤・分散システム
構築・運用・サポートなど
お気軽にお問合せください

株式会社INTHEFOREST　担当 冨田・高木
sales@intheforest.co.jp