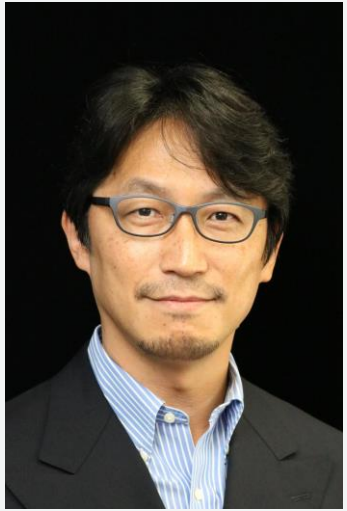
# 【AWS Black Belt Online Seminar】AWSで構築するデータレイク基盤のアーキテクチャ

アマゾン ウェブ サービス ジャパン株式会社
ソリューションアーキテクト　丹羽勝久

2018.04.24

# 自己紹介



**名前：**

丹羽 勝久（にわ かつひさ）

**所属：**

アマゾンウェブサービスジャパン株式会社
インダストリーソリューション部
ソリューションアーキテクト

**担当：**

製造公益領域のお客様担当

# 内容についての注意点

- 本資料では2018年4月24日時点のサービス内容および価格についてご説明しています。最新の情報はAWS公式ウェブサイト(http://aws.amazon.com)にてご確認ください。

- 資料作成には十分注意しておりますが、資料内の価格とAWS公式ウェブサイト記載の価格に相違があった場合、AWS公式ウェブサイトの価格を優先とさせていただきます。

- 価格は税抜表記となっています。日本居住者のお客様が東京リージョンを使用する場合、別途消費税をご請求させていただきます。

- AWS does not offer binding price quotes.  AWS pricing is publicly available and is subject to change in accordance with the AWS Customer Agreement available at http://aws.amazon.com/agreement/.  Any pricing information included in this document is provided only as an estimate of usage charges for AWS services based on certain information that you have provided.  Monthly charges will be based on your actual use of AWS services, and may vary from the estimates provided.

# Agenda

- データレイクとは何か？なぜ必要なのか？
- データレイクアーキテクチャ
- データレイクを構成するAWSのサービス
- データレイク活用事例

データレイクの話の前に….

# 「ビッグデータ」について考えてみましょう

# ビッグデータとは？

3Vs: Volume(大容量), Velocity(高更新頻度), Variety(多様)

5Vs: **3Vs** + Value(高い価値), Veracity(高い正確性)

**Wikipediaより）** https://ja.wikipedia.org/wiki/ビッグデータ

ビッグデータ（英: big data）とは、**一般的なデータ管理・処理ソフトウエアで扱うことが困難なほど巨大で複雑なデータの集合**を表す用語である。

**Gartnerの定義）** https://www.gartner.com/it-glossary/big-data/

**Big data** is **high-volume, high-velocity** and/or **high-variety** information assets that demand **cost-effective, innovative forms of information** processing that **enable enhanced insight, decision making, and process automation**.

# ビッグデータ分析：論理的推論の２つの手法

- **演繹法**
  - ➢ 一般論やルールに観察事項を加えて、必然的な結論を導く思考方法で三段論法とも言われる

例）「人間はいつか死ぬ」→「ソクラテスは人間である」→**「ソクラテスはいつか死ぬ」**

- **帰納法**
  - 帰納法は、多くの観察事項（事実）から類似点をまとめ上げることで結論を引き出すという論法

例）

- 「人であるソクラテスは死んだ」
- 「人であるプラトンは死んだ」
- 「人であるアリストテレスは死んだ」

→ **「したがって人は全て死ぬ」**

機械学習や高度統計分析のアプローチ

# ビッグデータ分析の分類

トップダウンアプローチ（演繹的）

①分析の要件、目的を明確化
→ 経営分析、財務分析、売上分析など

②データモデリング、ディメンジョン設計

③データ収集、変換、ロード (ETL)
→ どのシステムからデータを収集するか？

④分析実行、レポート生成

**データウェアハウス、データマート**
SQL主体で基本100%の正確性で
事実把握 （Schema on Write）

**機械学習／深層学習など**
Java, Scala, R, Pythonなどが主体
予測分析、傾向分析、推論導出
（Schema on Read）

⑤分析結果の可視化、レポート

④データ形式変換、分析の実行

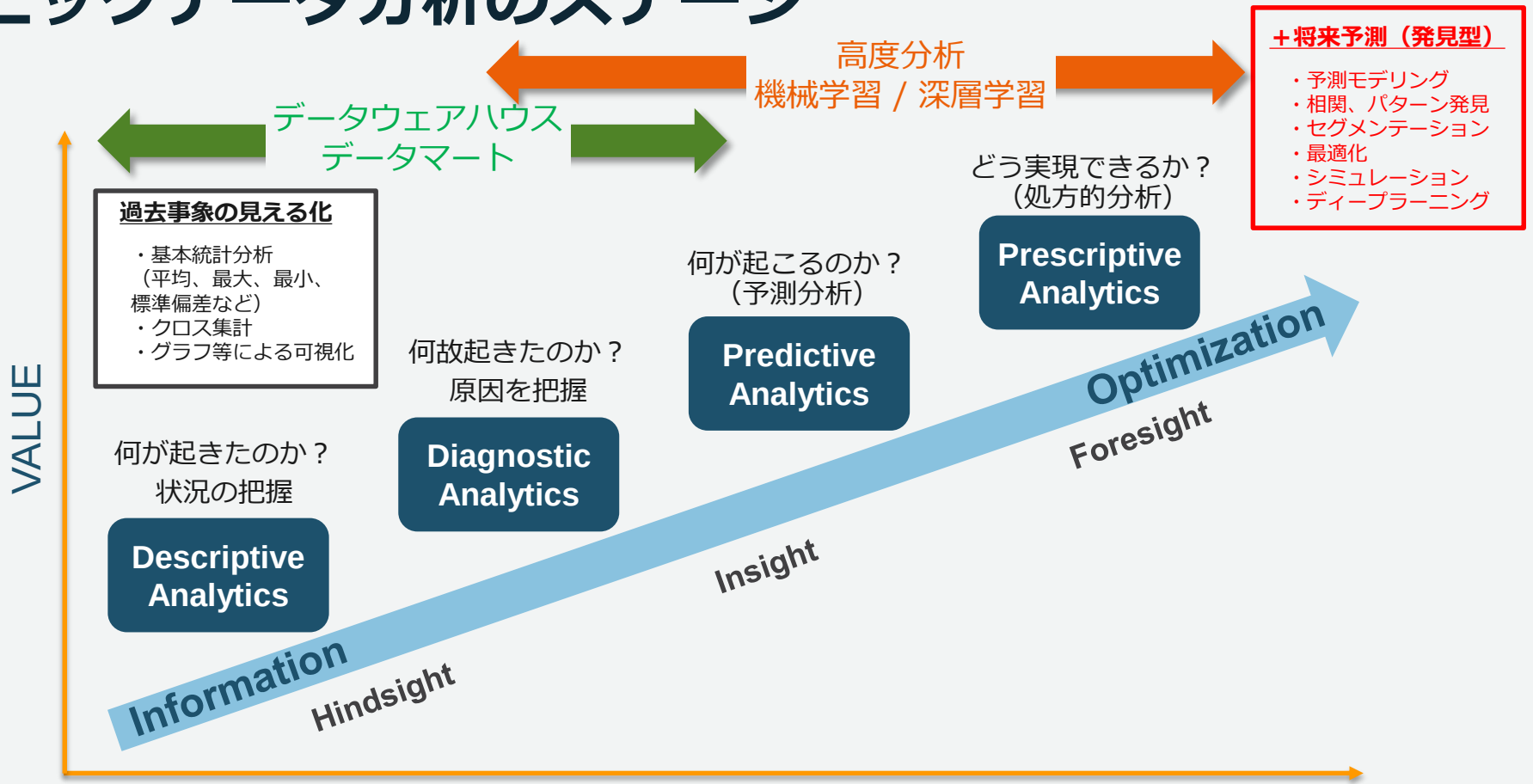
③求められる分析に対する仮説化、理論化

②収集されているデータを観察、パターン化

①要件にかかわらず関連データ収集

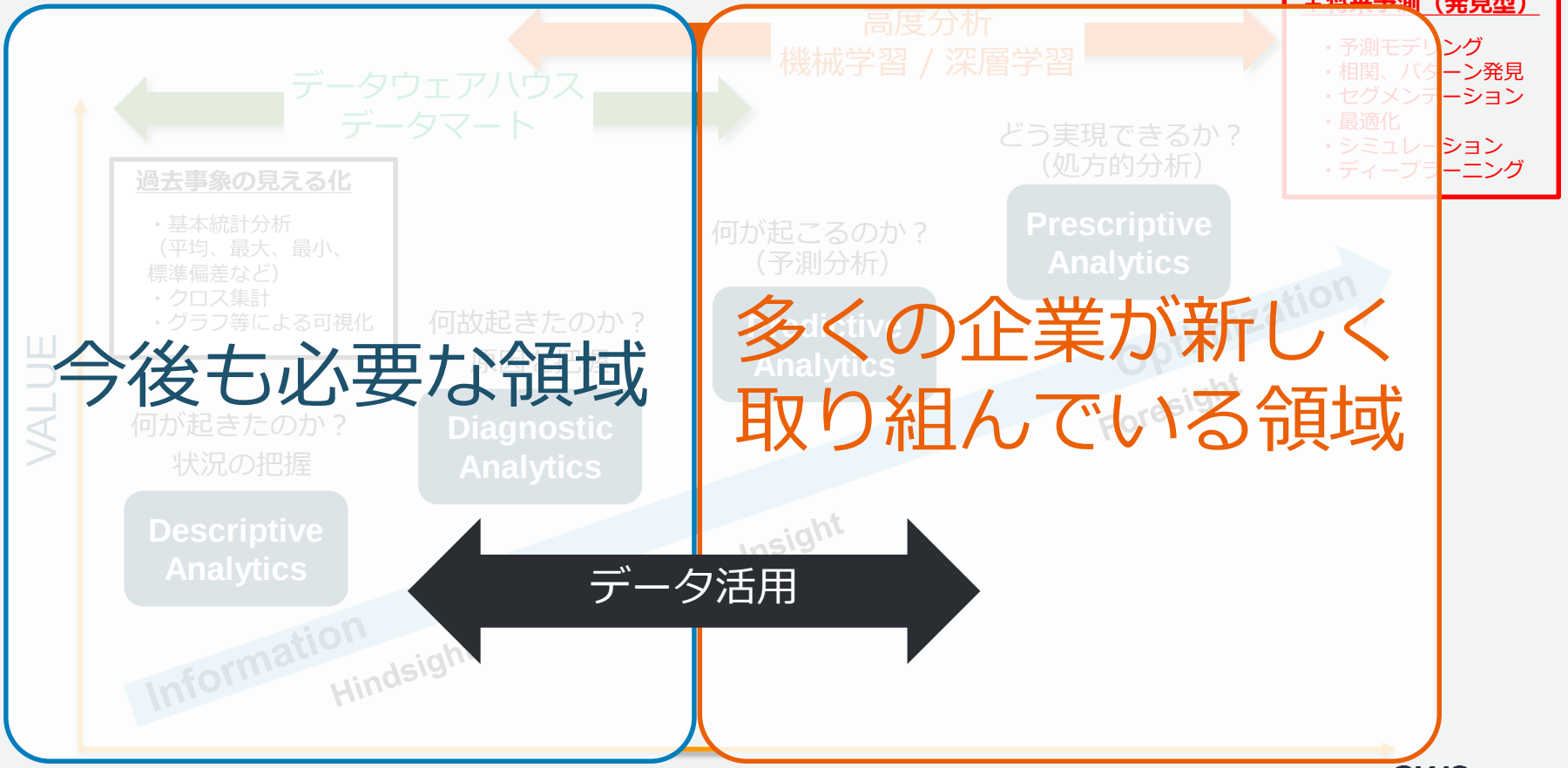
ボトムアップアプローチ（帰納的）

# ビッグデータ分析のステージ




Source: **Gartner**

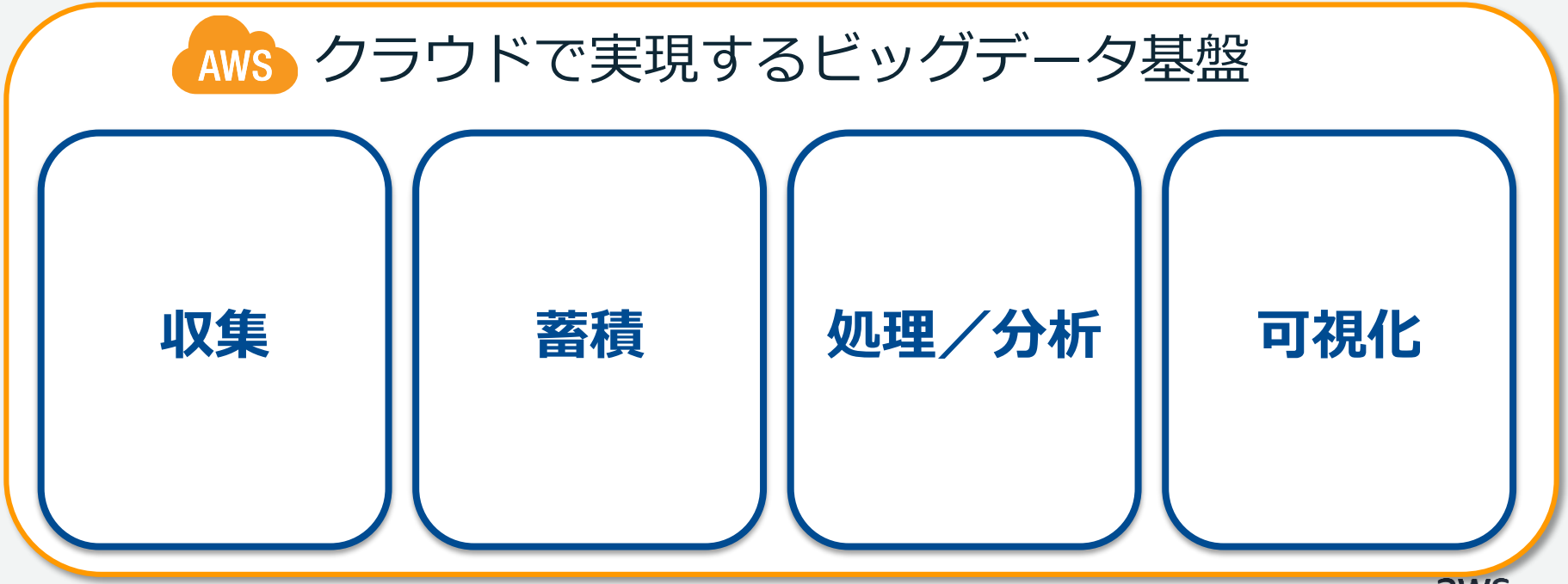
# ビッグデータ分析のステージ




Source: **Gartner**

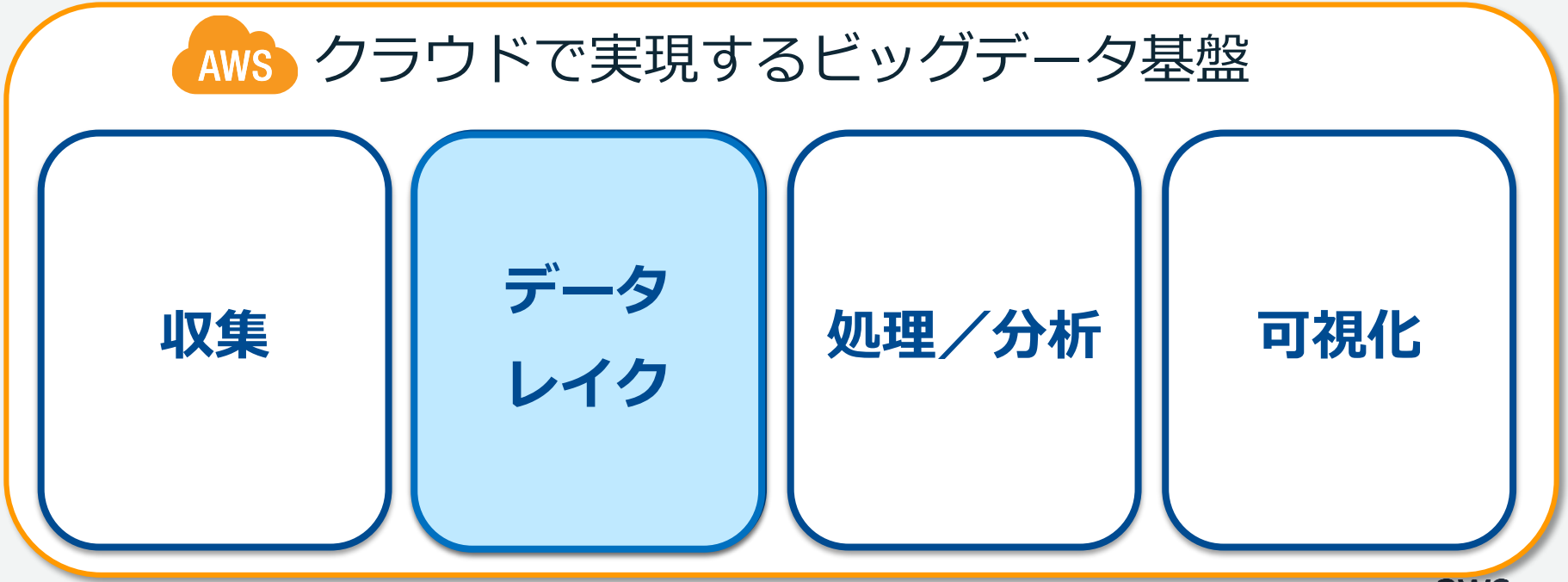
# ビッグデータ基盤とは？

- **データ活用サイクル全体を管理できる**新しいツール群で構成
- **技術的にもコスト的にも実現可能な基盤**である必要あり



# ビッグデータ基盤とは？

- **データ活用サイクル全体を管理できる**新しいツール群で構成
- **技術的にもコスト的にも実現可能な基盤**である必要あり



# 統制の取れていないビッグデータの取組みでは

System of Records
構造化データ

- 経理・購買・人事 (ERP)
- CRM／顧客管理
- 生産情報、SCM、ロジ

*etc.*

フラット
ファイル

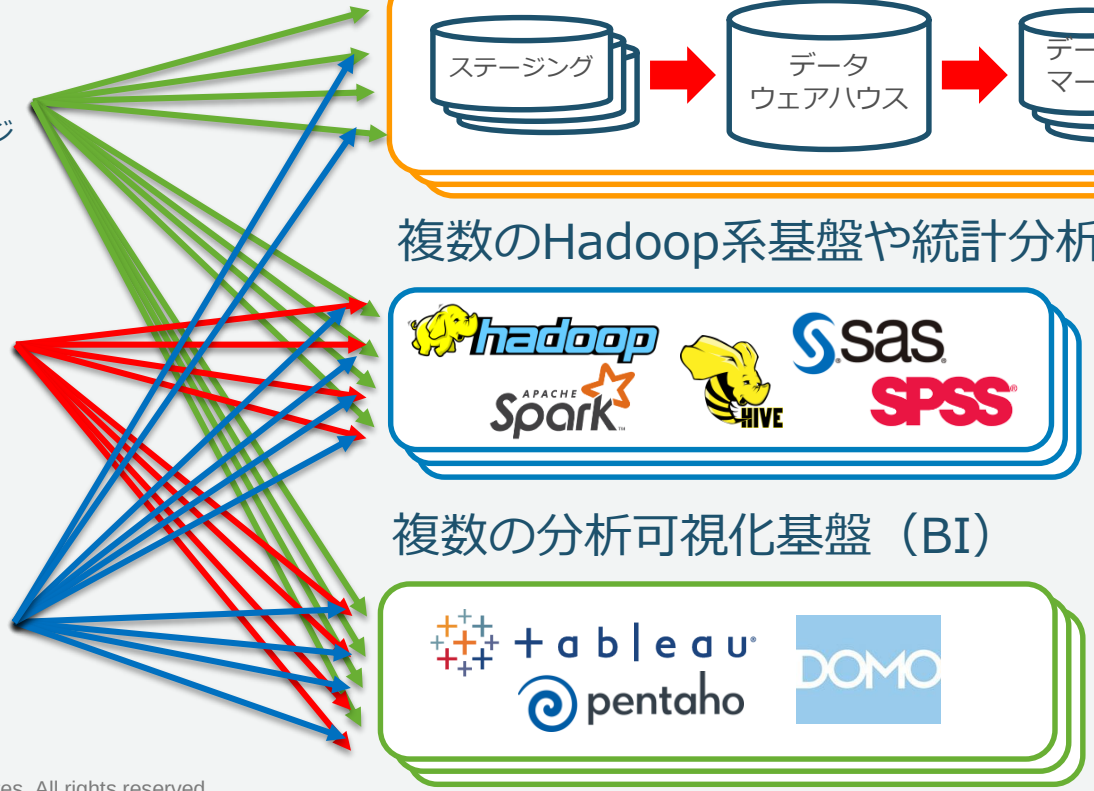
- ログファイル
- ストリームデータ
- レポート
- 監視カメラ映像

*etc.*

外部ソース
オープン
データ

- デモグラフィック
- 経済指標・為替
- 気象情報
- 地理情報
- ソーシャル

*etc.*



複数のデータウェアハウス基盤

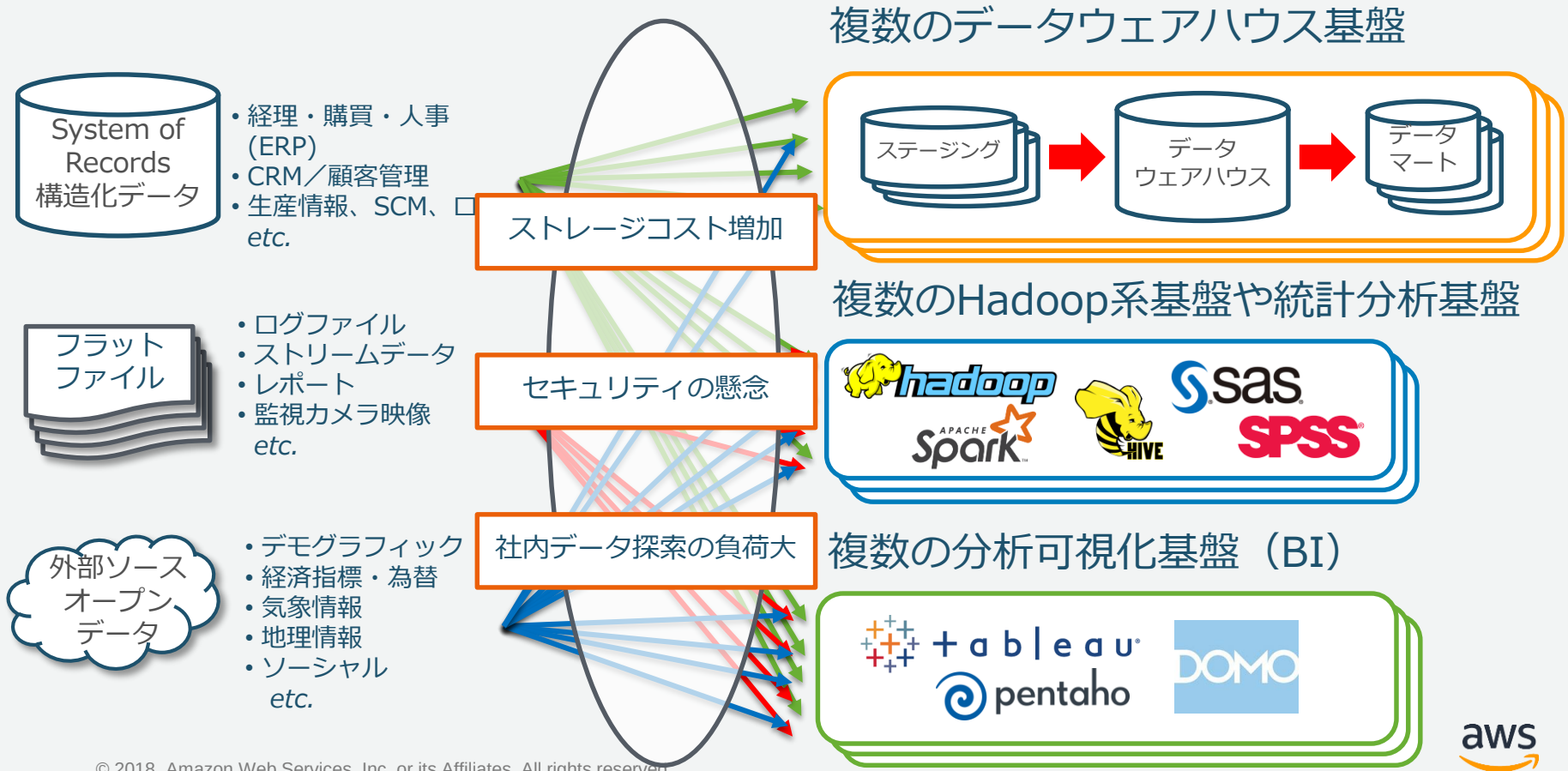



複数のHadoop系基盤や統計分析基盤

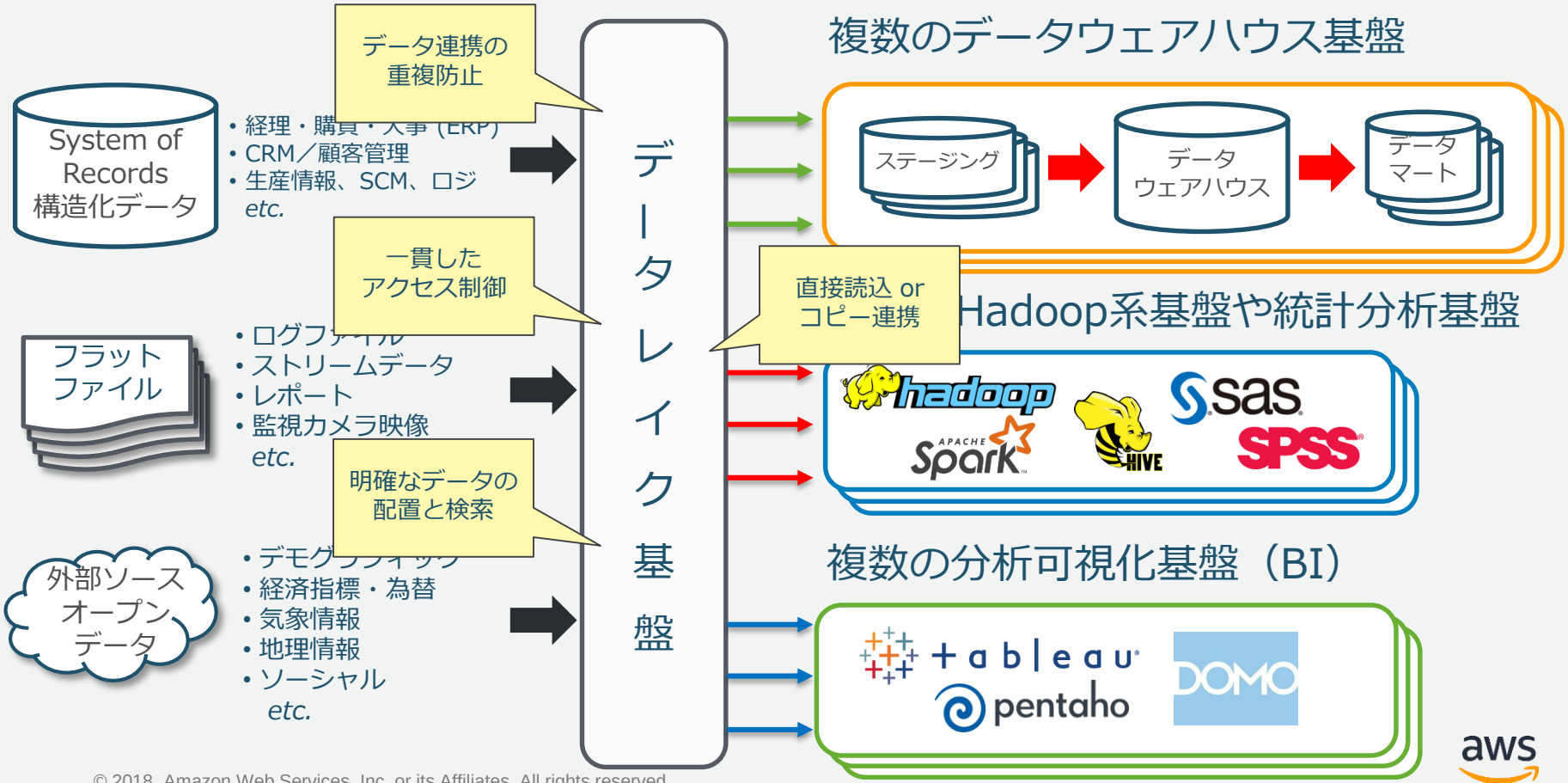



複数の分析可視化基盤（BI）

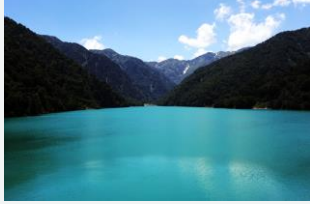
# 統制の取れていないビッグデータの取組みでは

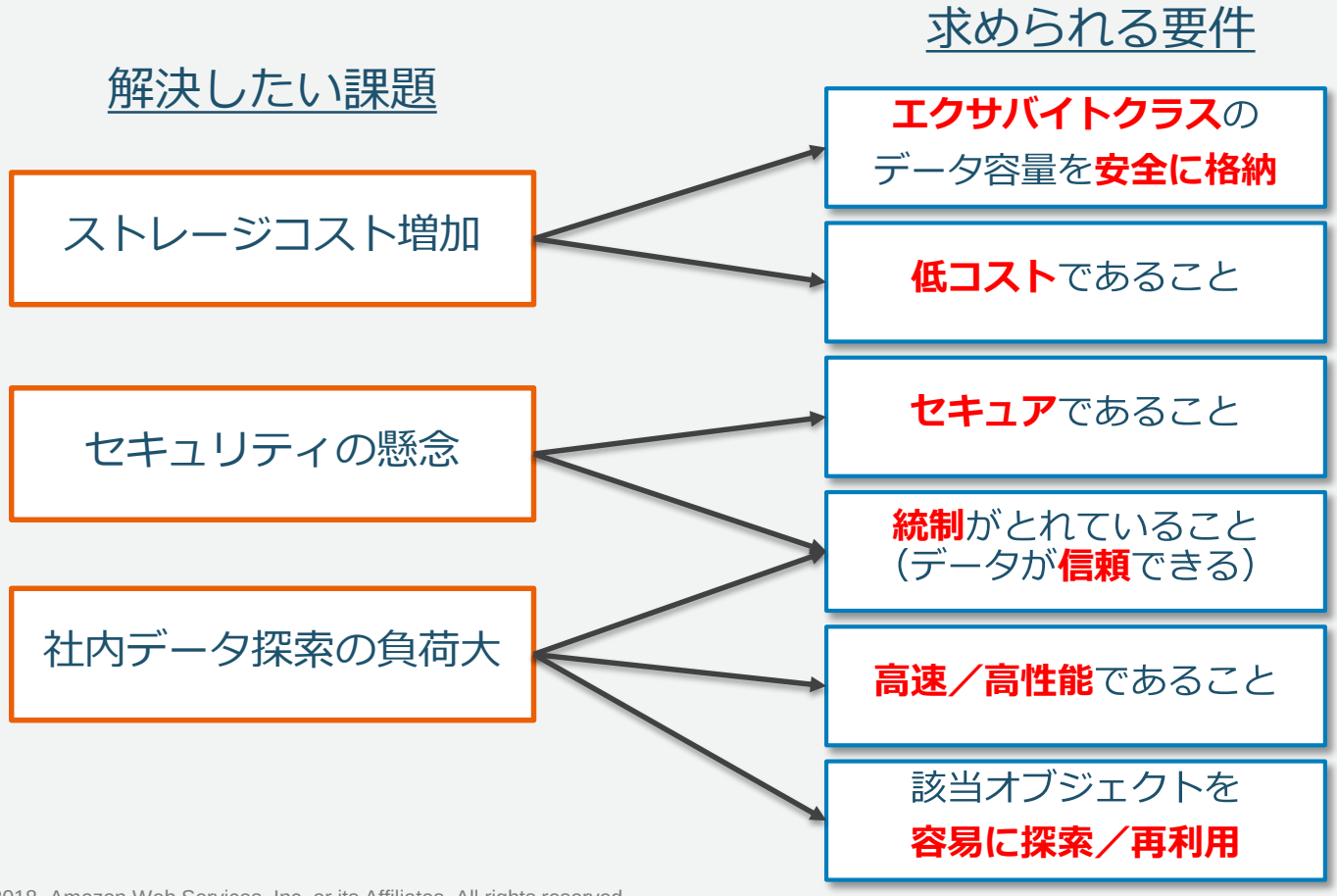


# データレイクを採用したビッグデータ処理



# データレイク基盤の求められる要件

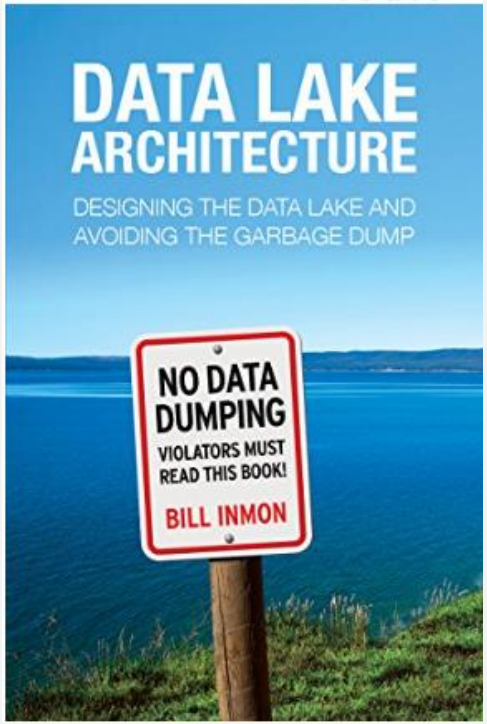




# Agenda

- データレイクとは何か？なぜ必要なのか？
- データレイクアーキテクチャ
- データレイクを構成するAWSのサービス
- データレイク活用事例

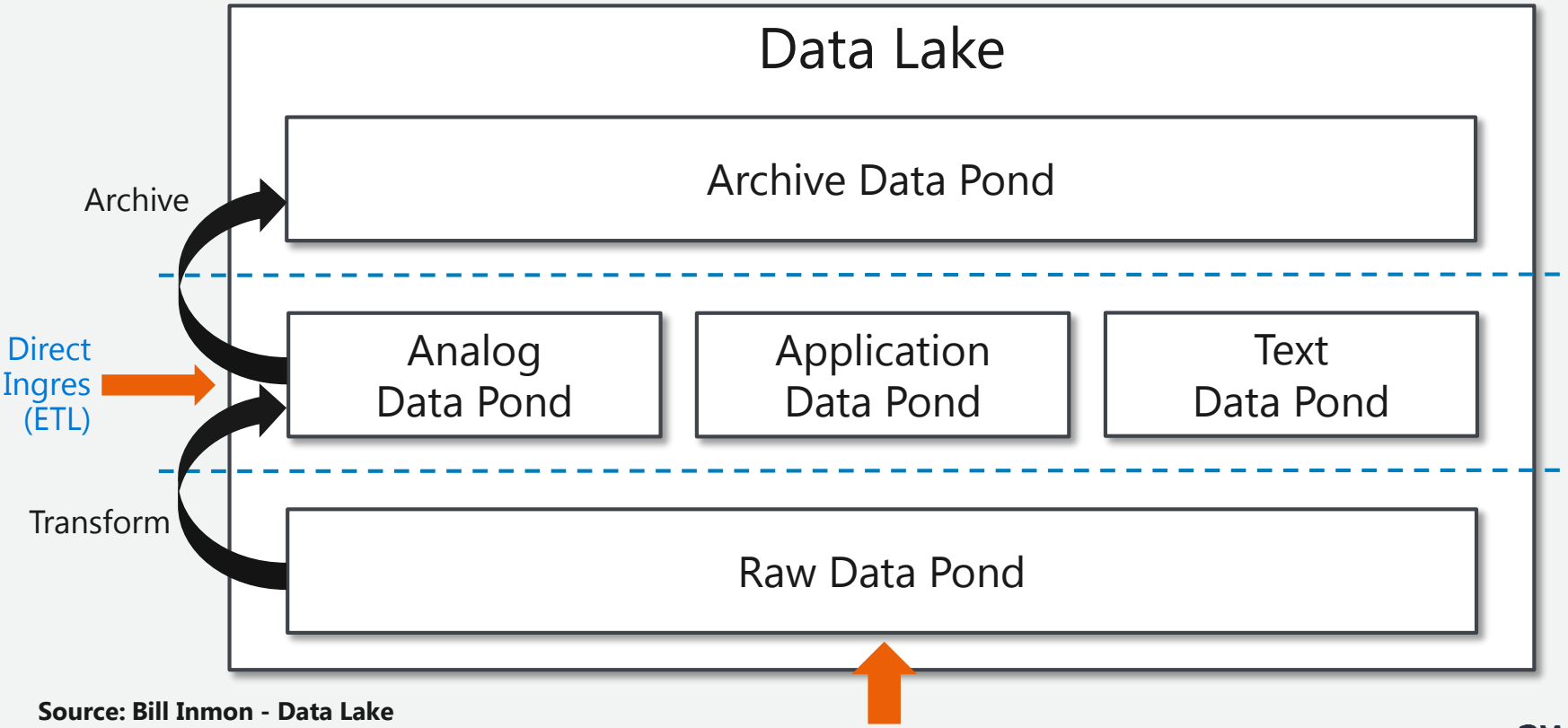
# ご参考）データレイク アーキテクチャ




**Data Lake Architecture**
**Designing the Data Lake and Avoiding the Garbage Dump**

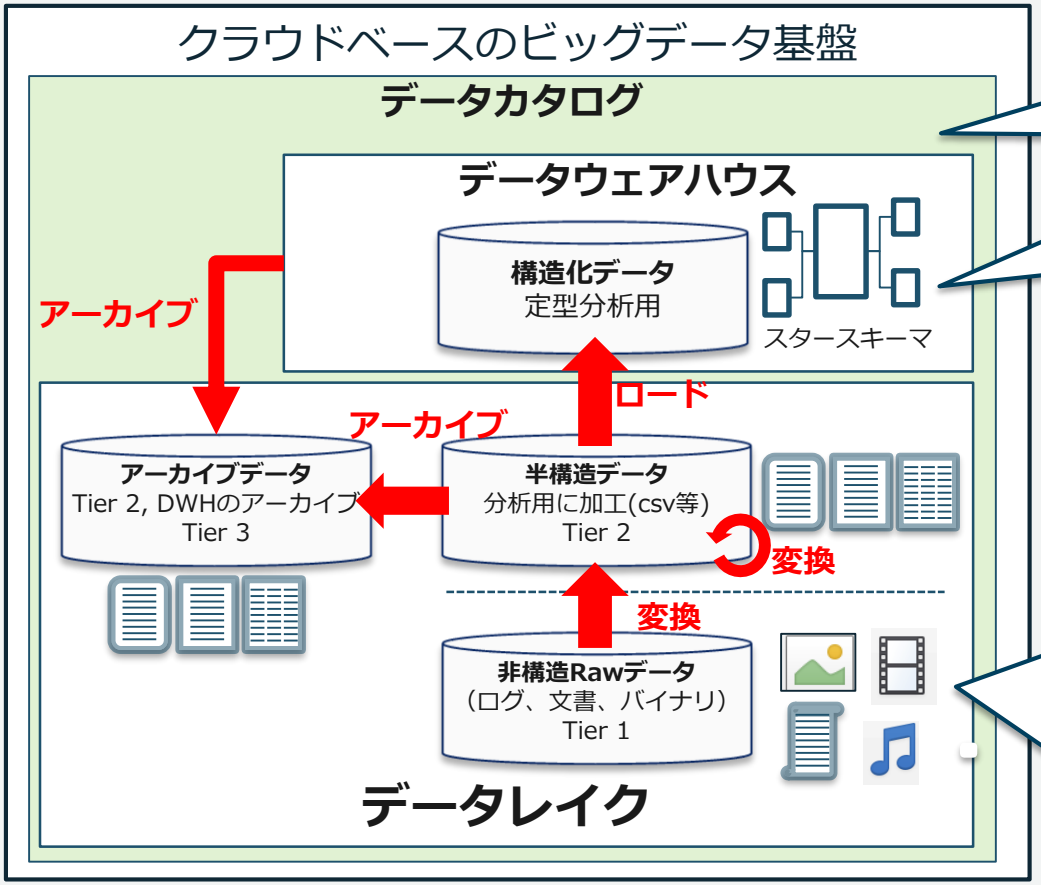
著者： Bill Inmon
データウェアハウスの提唱者とされていて、Ralph Kimballと並んで、データウェアハウスアーキテクチャの第一人者

# データレイクの論理アーキテクチャ




**Source: Bill Inmon - Data Lake Architecture**

# クラウドベースのビッグデータ基盤 ～ データレイク




**データカタログ：**
- データレイク、データウェアハウスなどのデータオブジェクトのカタログ

**データウェアハウス：**
- 大量の構造化データを高速に集約／分析処理を実行

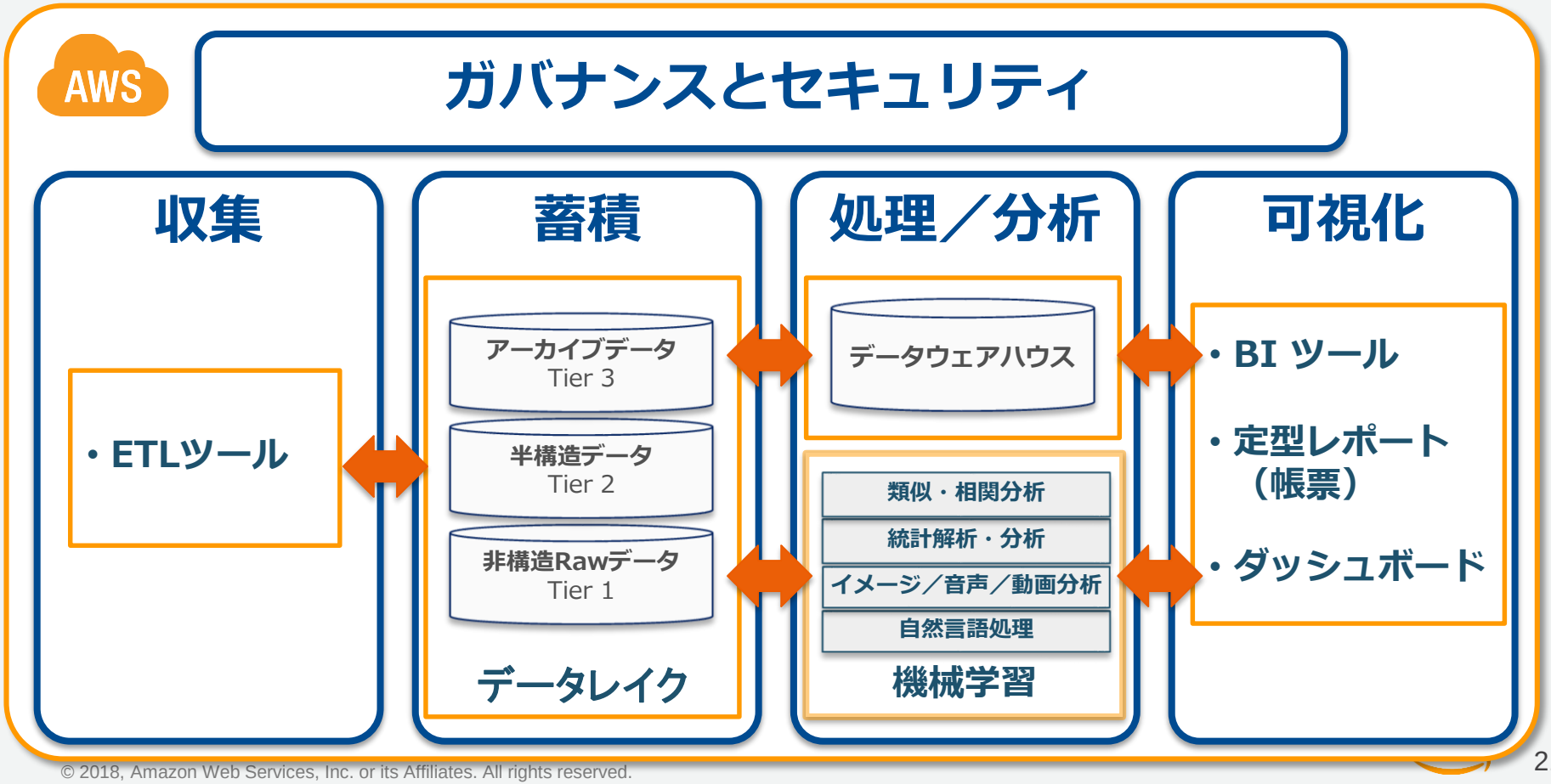
**データレイク：** ※前ページの論理アーキテクチャ踏襲

第1層（Tier 1）
- 非構造データ／Rawデータ

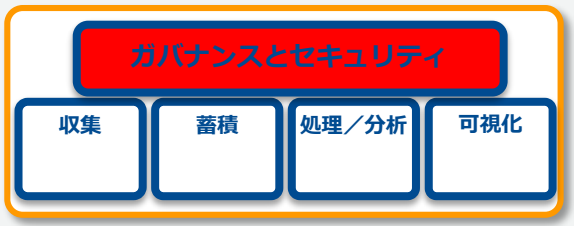
第2層（Tier 2）
- 半構造データ（CSV, Parquetなど）
- Hadoop/Spark、機械学習などから直接利用される

第3層（Tier 3）
- 半構造形式のアーカイブデータ

# ビッグデータ基盤



# ガバナンスとセキュリティ




- ガバナンス
  - データカタログ
  - データスチュワード
- セキュリティ
  - 認証／認可
  - データの暗号化
  - 監査

# ガバナンス：データレイクを「ゴミ溜め」にしないために

## データスチュワード

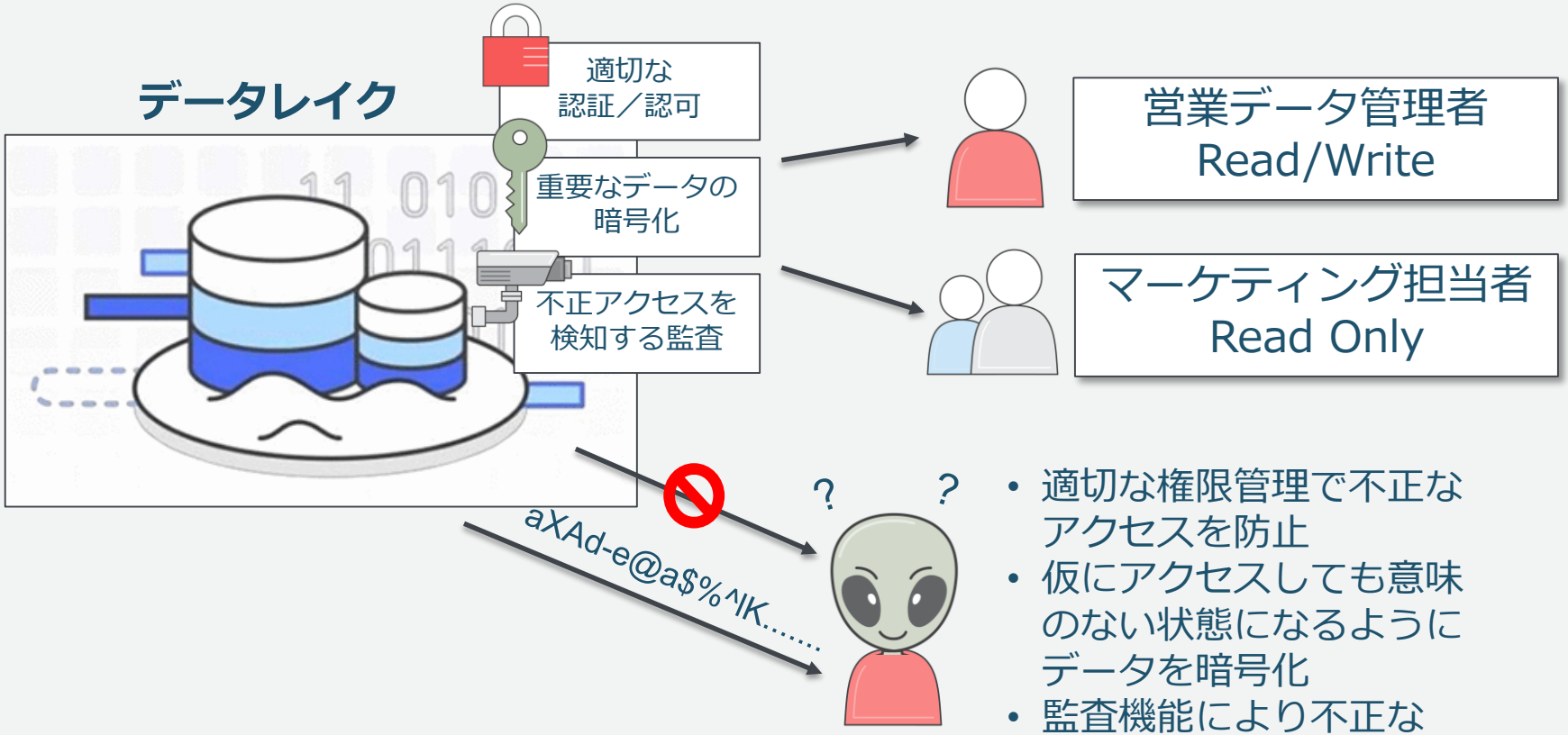
**組織内のデータの統制を保つ特殊な任務を担当**

- データの内容（コンテンツ）、メタデータを共に管理
- メタデータ内のデータエレメントの定義を明確にする
- データカタログ内のエレメント間で矛盾を発生させない
- データレイク、周辺データのアクセス権限管理を統括

- データの健全性を保つ
- データ連携の定義管理
- 利用者サポート

## データレイク

# セキュリティ：データレイクから重要な情報を漏洩させない




- 適切な権限管理で不正なアクセスを防止
- 仮にアクセスしても意味のない状態になるようにデータを暗号化
- 監査機能により不正なアクセス挙動を検知

# Agenda

- データレイクとは何か？なぜ必要なのか？
- データレイクアーキテクチャ
- データレイクを構成するAWSのサービス
- データレイク活用事例

# AWSサービス：ビッグデータ領域

## セキュリティとガバナンス

AWS CloudHSM
AWS Directory Service
AWS KMS
AWS Shield
IAM
Amazon Cloud Directory
Amazon Inspector
Amazon Macie
データカタログ
AWS Glue

## 収集

AWS IoT
Amazon Kinesis Data Firehose
AWS Snowball
AWS DMS
Amazon Kinesis Data Stream
AWS Glue

## 保存

Amazon S3
Amazon Glacier
Amazon DynamoDB

## 分析

Amazon EMR
Amazon Elasticsearch
Amazon Redshift
Amazon SageMaker
Amazon Athena
Amazon Kinesis Data Analytics

## 可視化

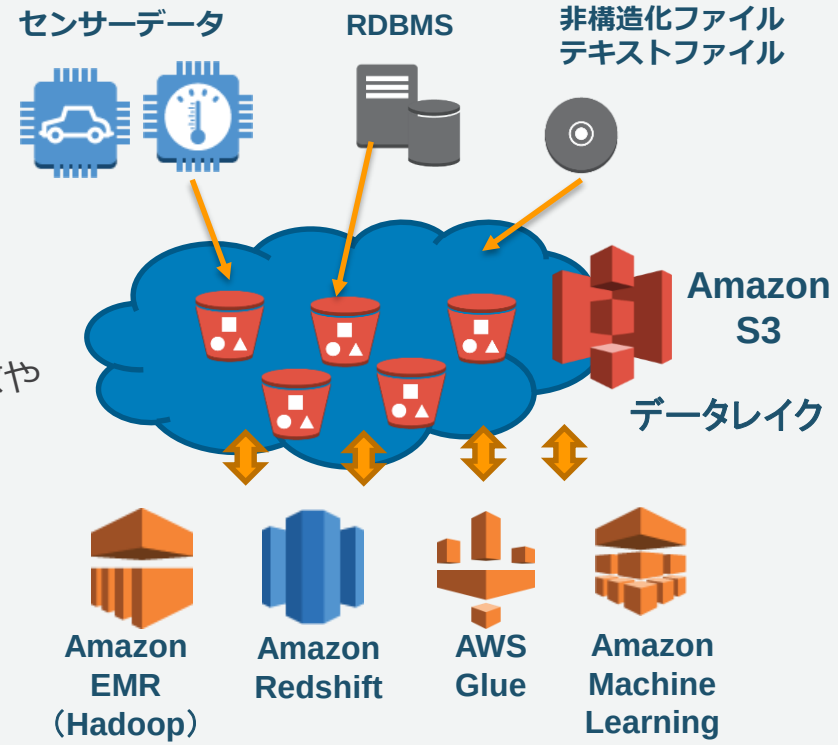
Amazon QuickSight


# Amazon S3

ビッグデータを安全に格納する
データレイク基盤の中心

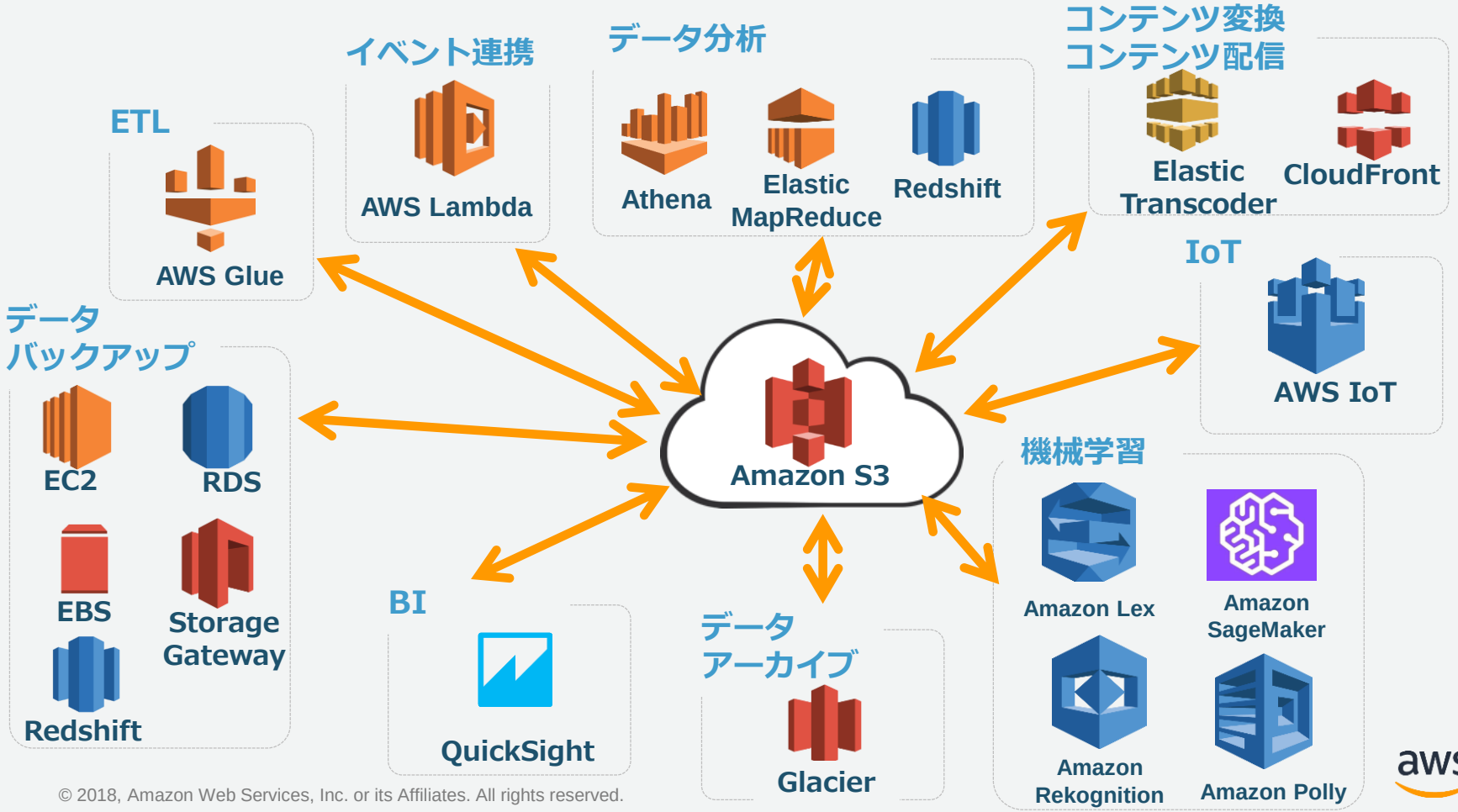
# Amazon S3によるデータレイク実現のメリット

- 大容量
  - 上限無し：サイジング不要
  - 様々な種類のデータも格納可能
- 強固なセキュリティ
  - IAMを利用した高い認証・認可レベル
  - 暗号化設定可能
- 安定した性能
  - データ容量に依存しない性能（サーバ台数、媒体本数やRAID、RAIDコントローラを考える必要がない）
- 高信頼性
  - 高い耐久性：99.999999999% ※複数AZにコピー
  - バージョニング、リージョン間コピー設定可能
- 安価： 例）10TBの保存で約2.1万円/月**
  - $0.025/GB/月*（スタンダード）
  - $0.019/GB/月*（標準-低頻度アクセス）

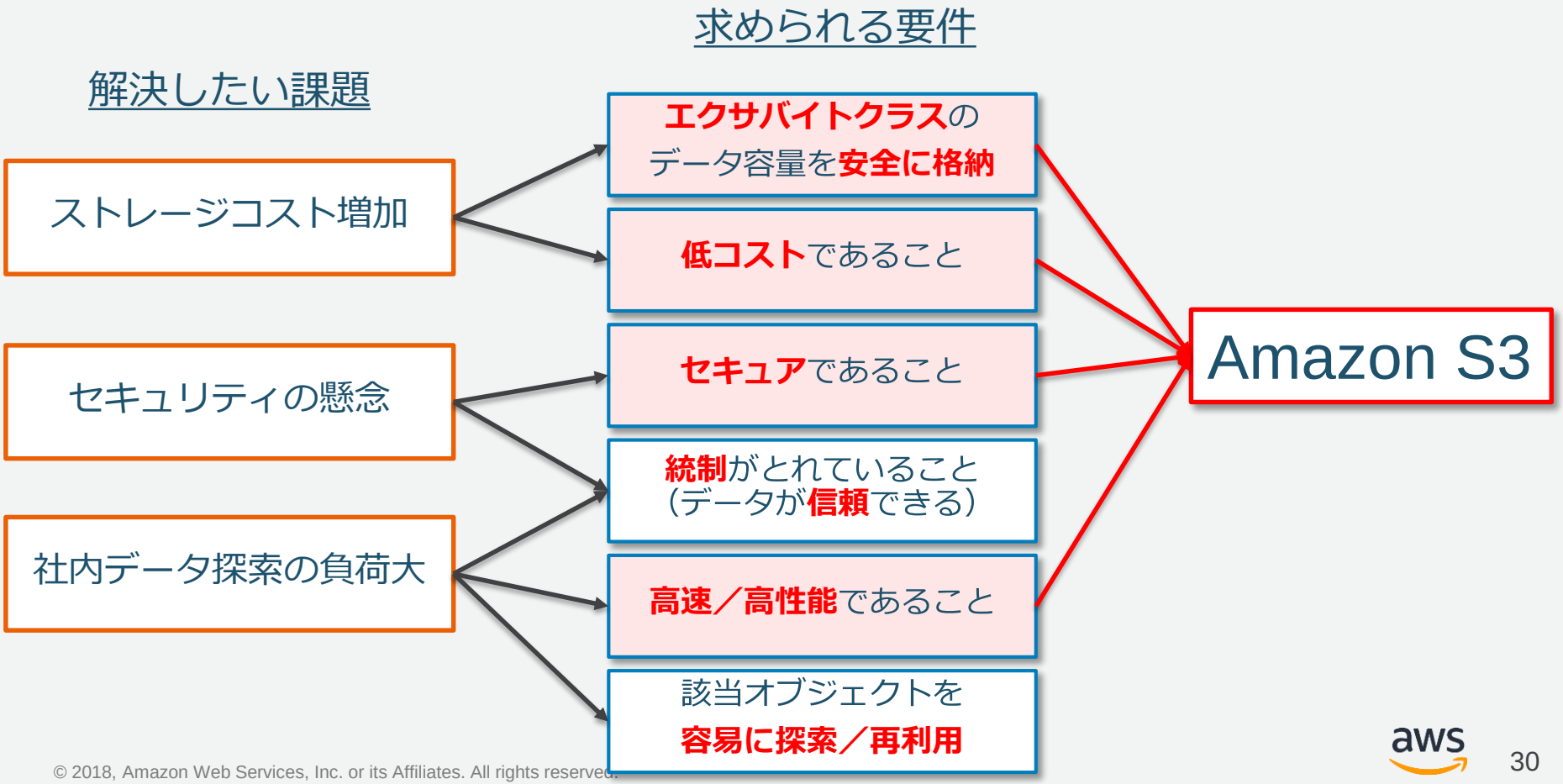


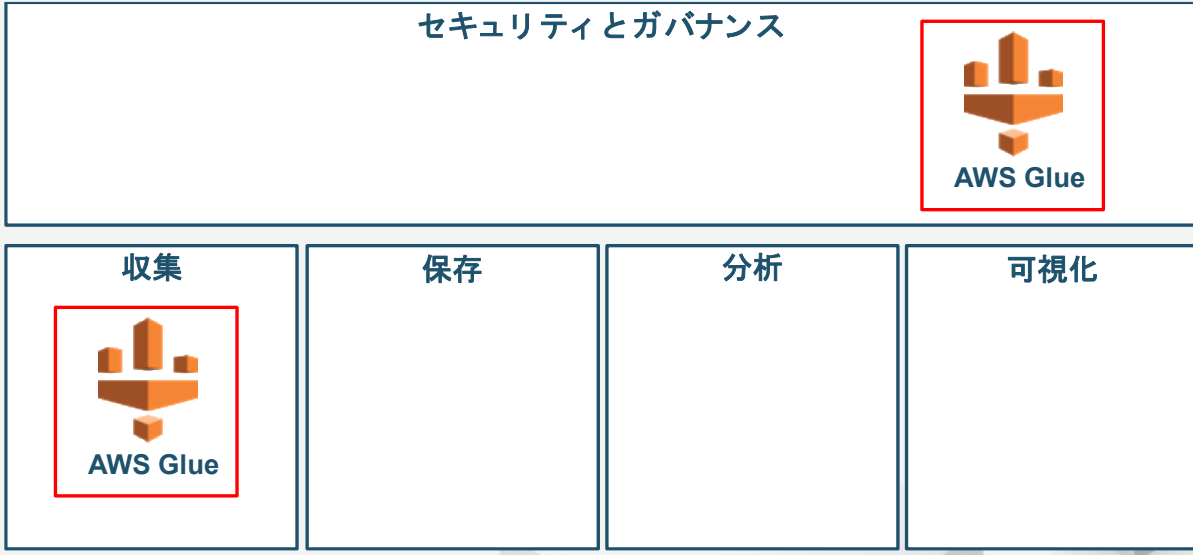

* 費用は2018年4月時点での東京リージョンでの価格です
** 1USドル = 110円で、標準-低頻度アクセスでの試算

# Amazon S3：様々なAWSのサービスと連携



# データレイク基盤の求められる要件の多くをS3でカバー

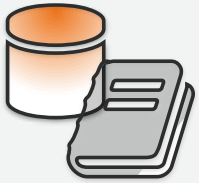





# AWS Glue

データレイク基盤、分析基盤への
ETL、プリプロセス処理を行い
連携したデータをカタログ化する

# AWS Glueの構成要素



**データカタログ**

- Hiveメタストア互換のデータソース用メタデータリポジトリ
- テーブル、データ型、パーティションフォーマットを推測するためにソースを**クロールしてカタログ自動生成**
- **Redshift Spectrum, Athena, EMRのカタログとしても利用・連携可能**

**ジョブオーサリング**

- ETL処理のためのPythonコード(PySpark)、またはScalaを生成
- 任意のIDE（統合開発環境）ツールでコードを作成可能

**オーケストレーション**

- **分散処理**でジョブを実行
- **サーバレス** - 利用したリソース分だけの支払い

# AWS Glue データカタログ (Data catalog)

表のメタデータをHiveメタストアで管理

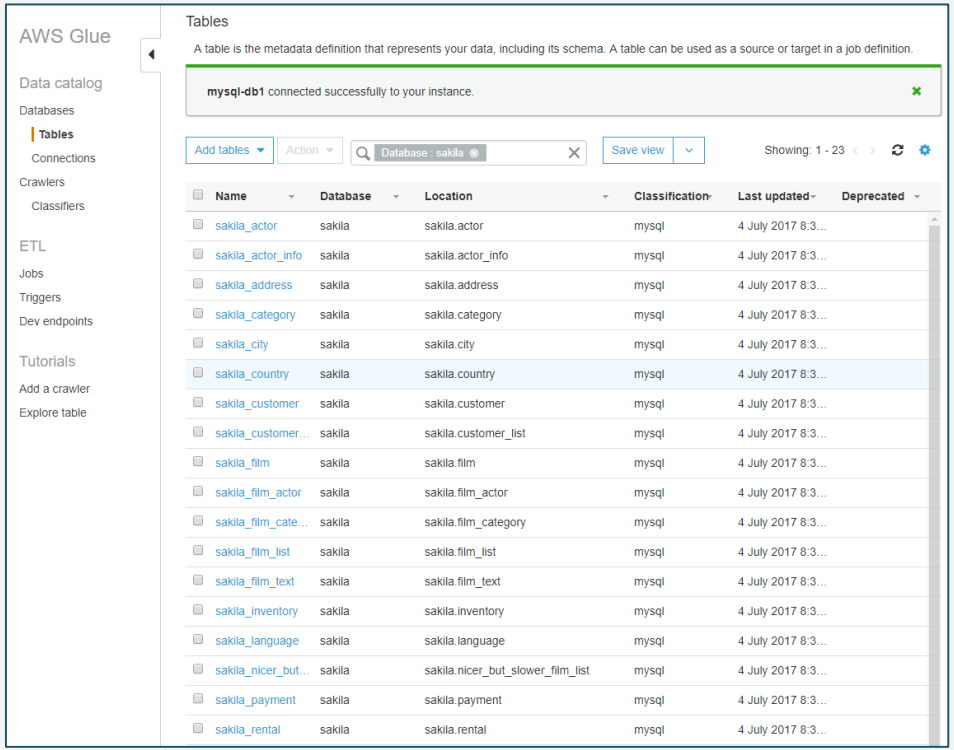
メタデータ

- **列・プロパティ・型**
- **データロケーション（URI）**
- **更新情報　　等**

**クローラー**による自動チェックと登録

- **Hiveパーティションを認識**し登録を自動化

```
/mydata
    /year=2017
        /month=11/...
```

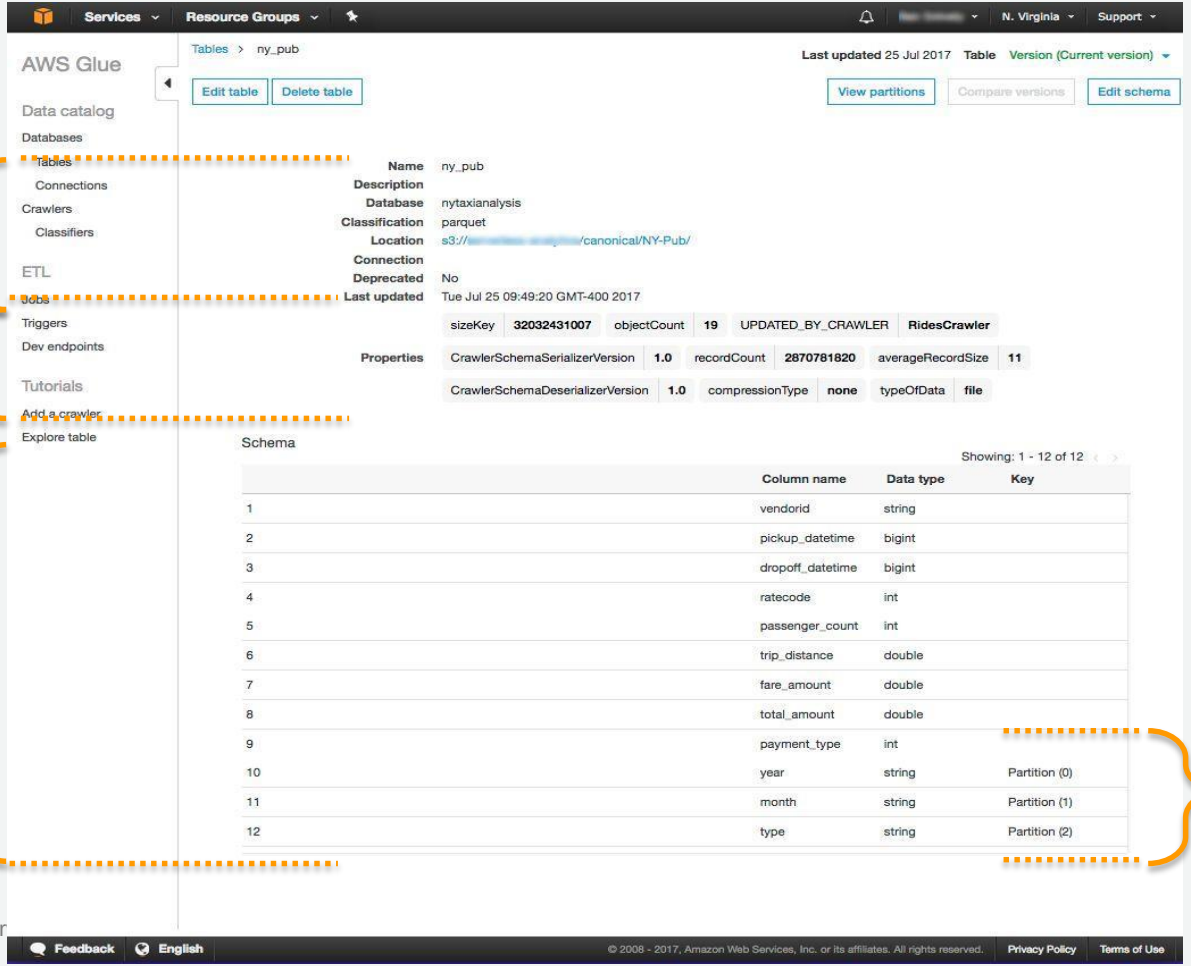


# データカタログのメタデータ内容の例

テーブルの
主要情報

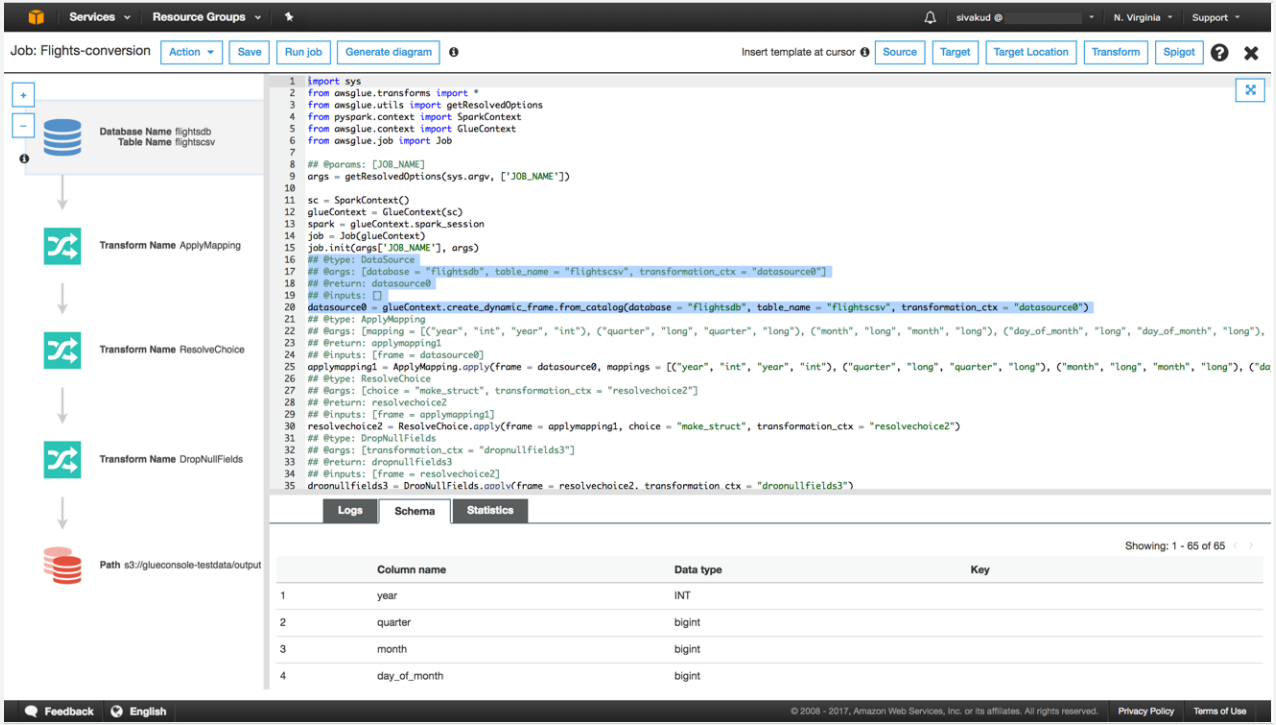
テーブルの
プロパティ

テーブル
スキーマ



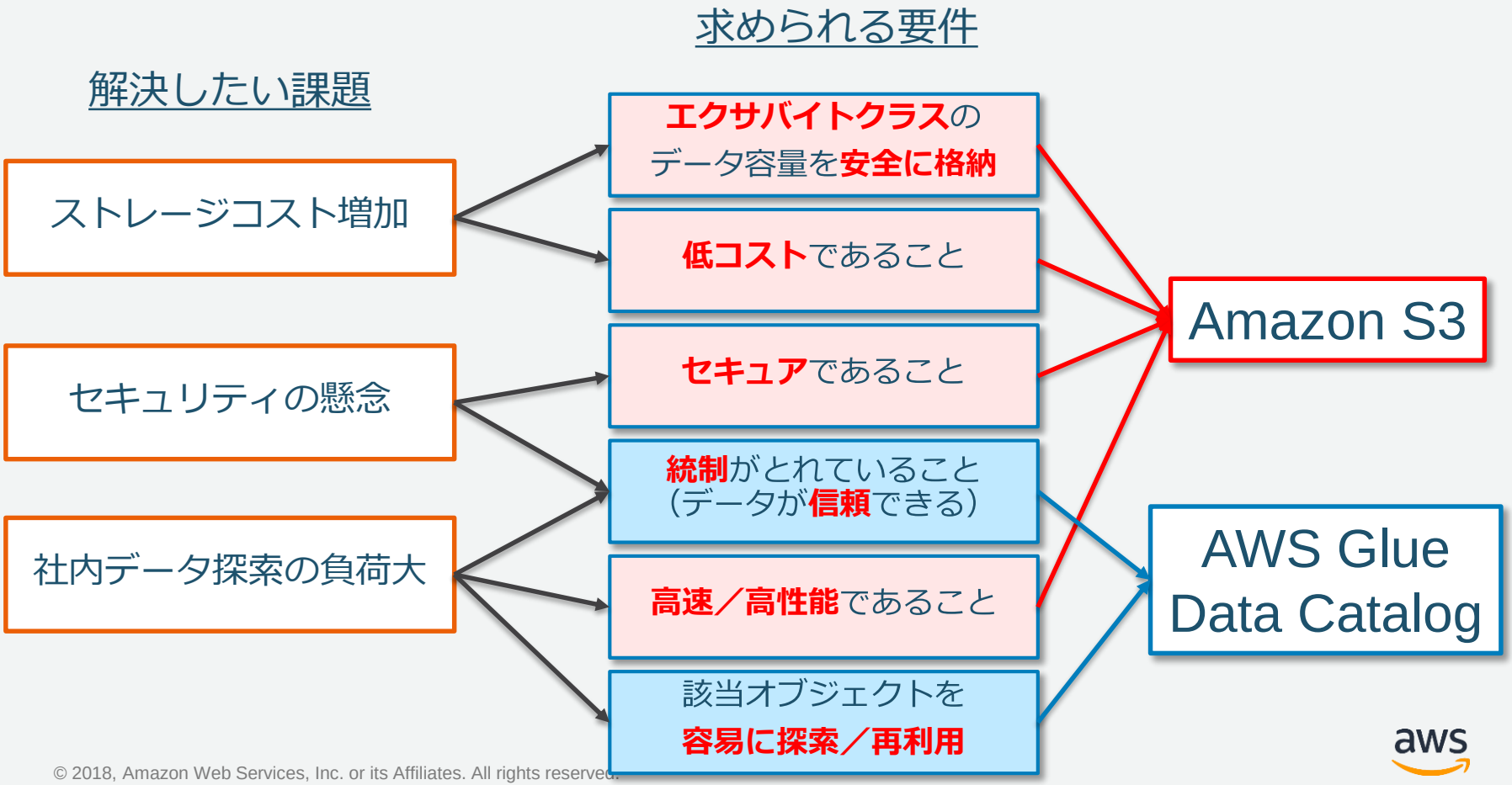

テーブル
パーティション

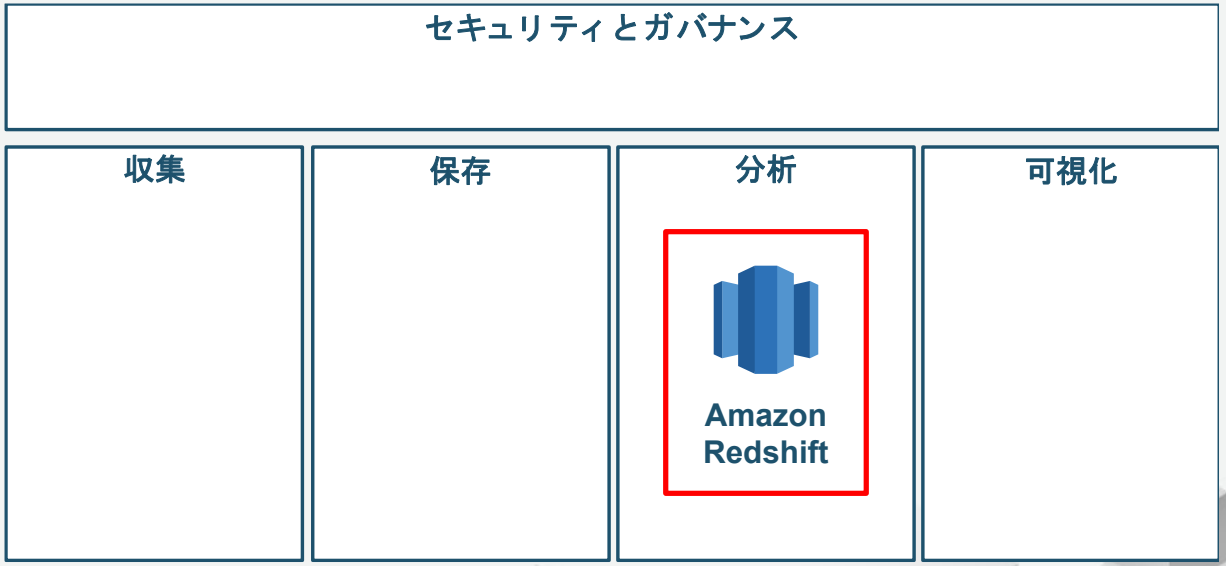
# ジョブの生成：Glue ETLによるデータ変換の自動化




- サーバ管理不要（サーバレス）
- リポジトリとクローラー機能
- ジョブの雛形が生成され、独自の変換処理をPySpark、またはScalaで記述可能

# データレイク基盤の求められる要件の多くを S3 + Glueで実現



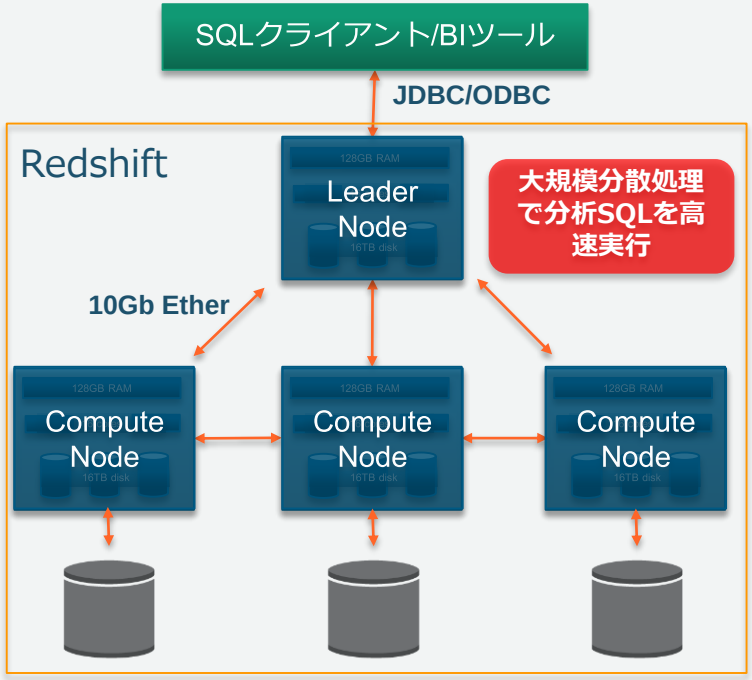



# Amazon Redshift

大量の構造化データを
高速に分析・集計する
データウェアハウス基盤

# Amazon Redshift

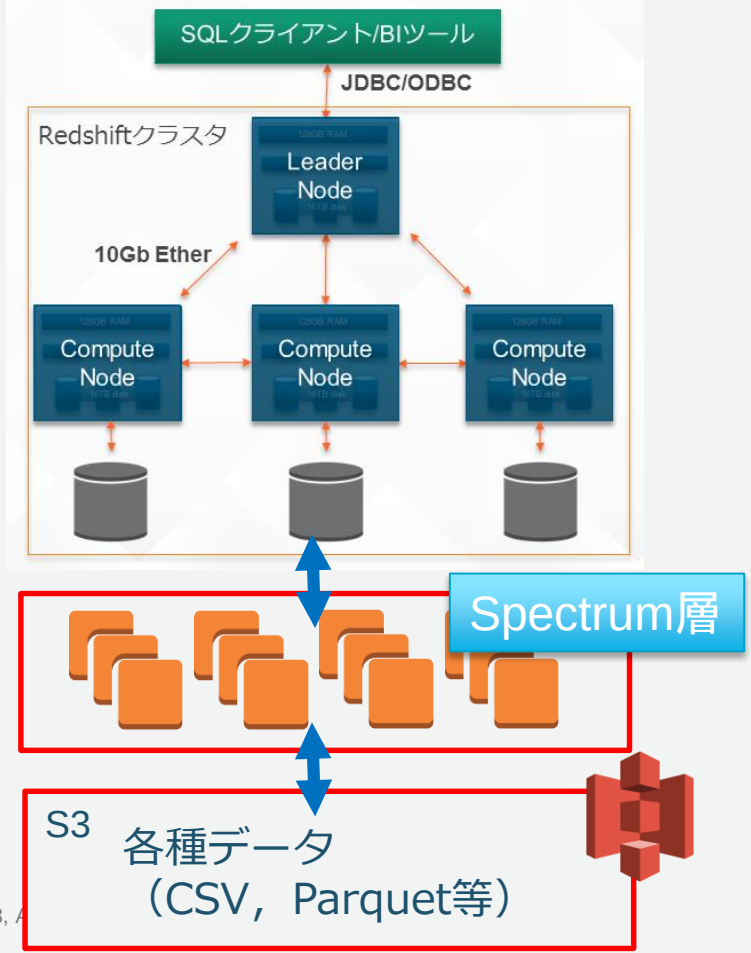
## フルマネージドのデータウェアハウスサービス




### 特徴

- データサイズ：最大2PBまで拡張可能
- 超並列(MPP)、カラムナ型DBエンジンによる高速SQL処理
- スケールアウト可能。最大128台
- PostgreSQLとの互換性
- 使った分だけの利用料金。従来のデータウェアハウスの1/10のコストで実現
- 管理機能がビルトイン

# Redshift Spectrum




- Redshiftが拡張され、S3上に置いたファイルを外部テーブルとして直接参照可能に
- Spectrum層により、S3上のファイルを高速処理
- ローカルディスク上のデータと組み合わせてSQLでクエリ可能
- 多様なファイルフォーマットに対応
  - CSV, TSV, Parquet, SequenceFile, RCFile, ORC, RegexSerDe 等

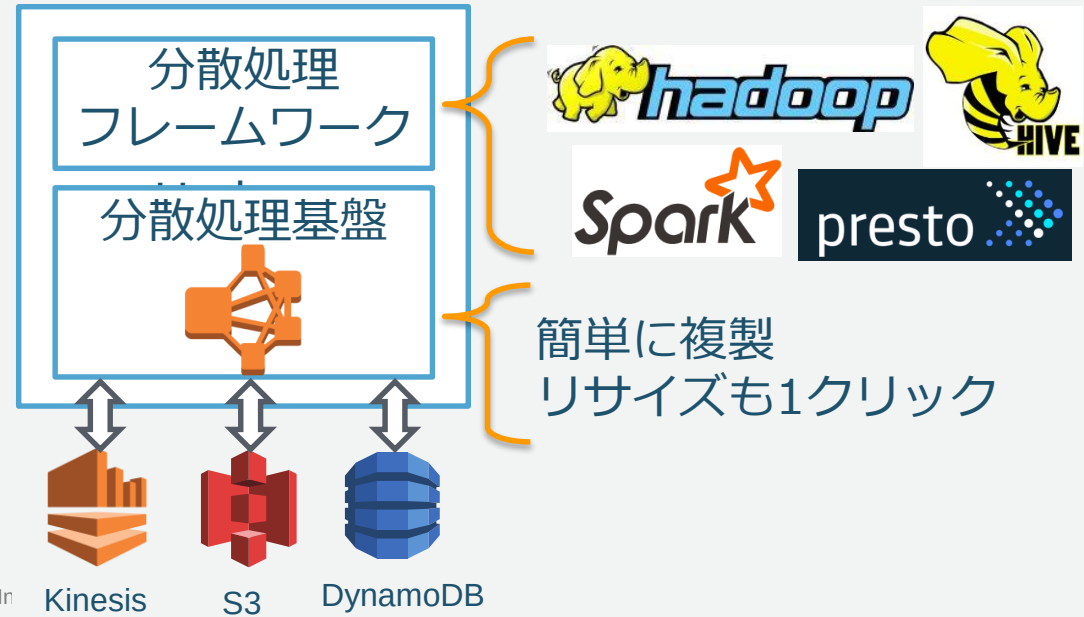

# Amazon EMR

データレイク基盤にアクセスして
ログ分析、データ変換、機械学習、
シミュレーションなどを高速実行する

# Amazon Elastic MapReduce（EMR）

Hadoop/Sparkなどの大規模分散処理環境を数クリックで構築

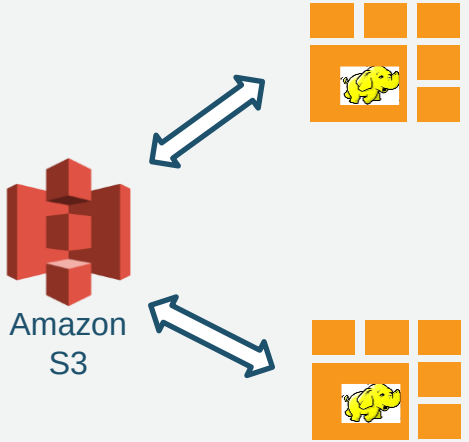
- 簡単スタート：数クリックでセットアップ完了
- 低コスト：従量課金、必用な時間だけクラスターを稼働



# EMRFS: S3をHDFSの様に扱う

“s3://”と指定するだけでHDFSと同様にS3にアクセス

- 計算ノードとストレージを分離できる
  - コスト面でもメリット大
- クラスタのシャットダウンが可能
  - クラスタを消してもデータをロストしない
- 複数クラスタ間でデータ共有が簡単
- データの高い耐久性（S3）



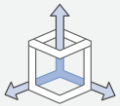

データレイクに直接並列でアクセスすることが可能


# Amazon Athena

データレイクに直接アクセスし
SQLによってインタラクティブに
検索分析実行する

# Athena の特徴

サーバレスでインフラ管理の必要なし



大規模データに対しても高速なクエリ

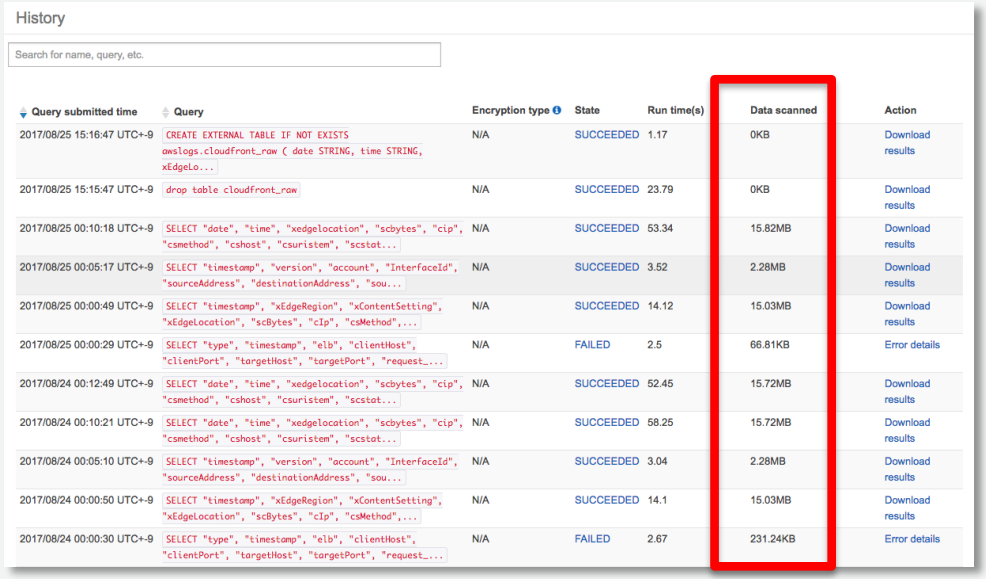
事前のデータロードなしに S3 に直接クエリ

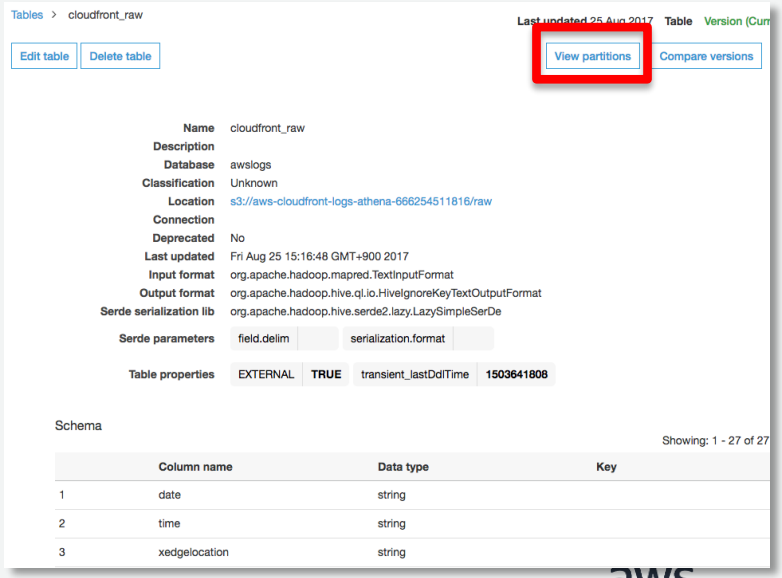
スキャンしたデータに対しての従量課金

JDBC 経由で BI ツールから直接クエリ

# マネジメントコンソールから簡単にアクセス

- 利用開始から数分でクエリを投げられる
- クエリごとのスキャンデータ量を確認可能
- Glue データカタログでテーブルスキーマやパーティションを確認



# Amazon Athena

S3に保存したファイルを直接クエリー。管理は不要

```
select
  class
  , sex
  , total_cnt
  , survived_cnt
  , round(cast(survived_cnt as double)/cast(total_cnt as double), 2) as survival_rate
from (
  select
    class
    , sex
    , count(survived) as total_cnt
    , sum(survived) as survived_cnt
  from
    titanic_db.titanic
  where
    class != '*'
  group by
    class
    , sex
)
order by survival_rate desc;
```

Run Query | Save As | Format Query | New Query | (Run time: 1.9 seconds, Data scanned: 25.43KB)

Results

| | class | sex | total_cnt | survived_cnt |
|---|---|---|---|---|
| 1 | 1st | female | 143 | 134 |
| 2 | 2nd | female | 107 | 94 |
| 3 | 3rd | female | 212 | 80 |
| 4 | 1st | male | 179 | 59 |

S3上に保存したファイルにSQLを実行可能に

PrestoベースでANSI SQL対応

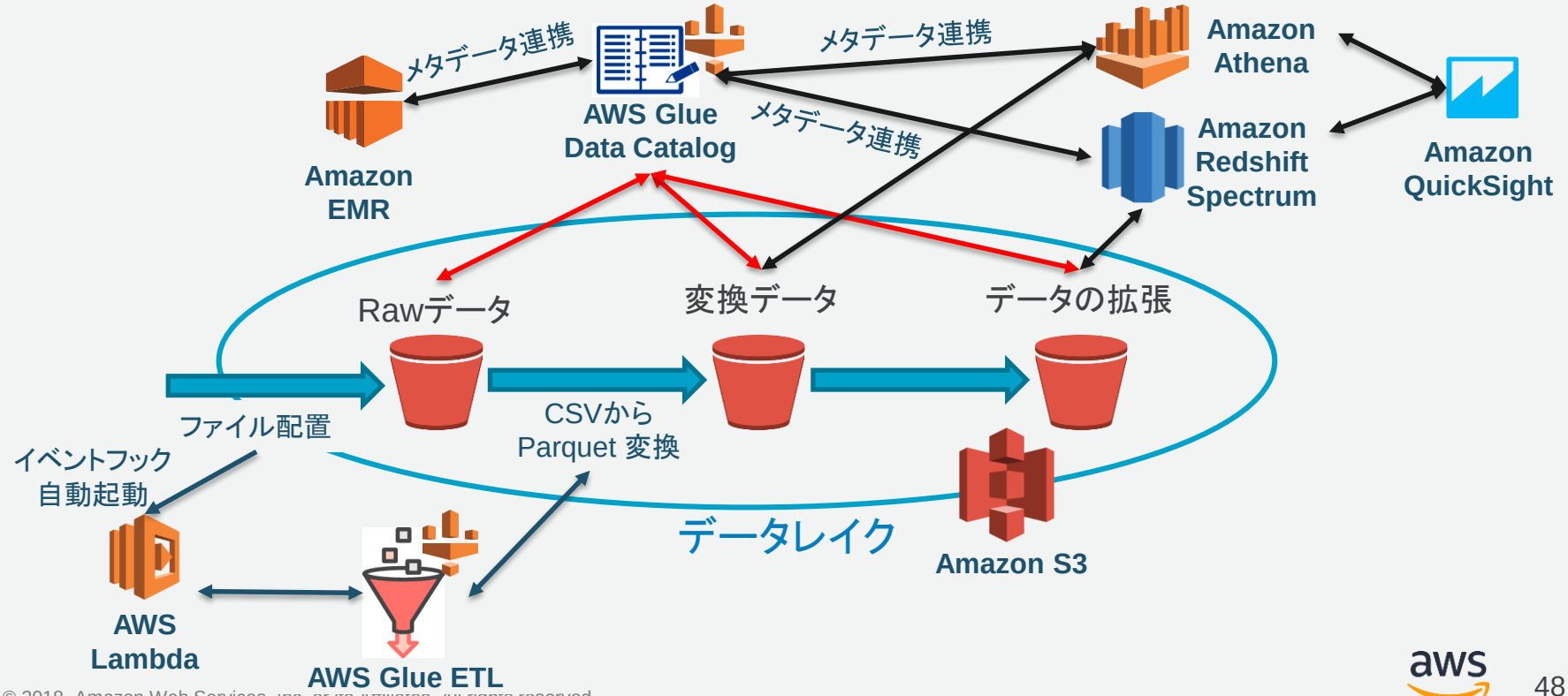
サーバ管理は不要でファイルを置いてSQLを書くだけ

# データ分析プラットフォーム比較

| | Amazon Redshift | Amazon EMR | Amazon Athena |
|---|---|---|---|
| 分類 | 大規模データ処理に特化したマネージドRDBMS | Hadoop/Spark等分散処理フレームワークのマネージド環境 | SQL特化のSQL処理環境（Prestoベース）。サーバ管理が不要 |
| 処理系 | SQL（PostgreSQLと互換性） | アプリケーションに依存：Hive、Spark等<br>Presto等でSQLでの分析も可能 | SQL（Prestoベース） |
| スケールアウト | 可能（手動）<br>※Spectrum層は自動 | 可能（手動） | **可能（自動）** |
| ストレージの種類 | **高速なローカルストレージ+S3上データへのクエリ** | HDFS、もしくは**EMRFSでS3上のデータを直接アクセス（ロードなし）** | **S3上のデータを直接アクセス** |
| ノードとストレージの分離 | ノードとストレージは同時に増加・削減。**S3上のデータはノードと関係なく増減可能** | **ストレージに影響を与えずにノード増減が可能（EMRFSの場合）** | **ストレージとノードは分離され、ユーザはノード管理自体不要** |
| 最大処理サイズ | 2PB＋S3（上限無し） | S3（上限無し）※EMRFSの場合 | S3（上限無し） |
| 運用管理 | **運用管理に必要な機能がビルトイン** | 分散処理系＋OSSアプリ環境を容易に導入可能 | **運用が不要** |
| 分析ツール、ETL等 | **JDBC/ODBC + SQLプッシュバック等BIアプリがネイティブレベルで対応** | JDBC/ODBC経由もしくはHiveメタストア経由で多くの環境がサポート | JDBC経由で多くの環境がサポート。API呼び出しも可能 |

# データレイクと各サービスの連動イメージ

AWSブログ： AWS GlueとAmazon S3を使用してデータレイクの基礎を構築する
https://aws.amazon.com/jp/blogs/news/build-a-data-lake-foundation-with-aws-glue-and-amazon-s3/



# Agenda

- データレイクとは何か？なぜ必要なのか？
- データレイクアーキテクチャ
- データレイクを構成するAWSのサービス
- データレイク活用事例

# Amazon.comの事例

本章は以下で公開されているre:invent 2017発表資料を元に作成しています
"A Look Under the Hood – How Amazon.com Uses AWS Services for Analytics at Massive Scale" (ABD329)
https://www.youtube.com/watch?v=lWK96BNx3Qk
https://www.slideshare.net/AmazonWebServices/a-look-under-the-hood-how-amazoncom-uses-aws-services-for-analytics-at-massive-scale-abd329-reinvent-2017

# Amazon.comのデータウェアハウス

多数の系列会社・部門で
分析・レポートに利用される
データウェアハウス

多様で膨大なデータ

- コアデータ：圧縮後で
  5PB+
- 全ストレージ：35PB+
- 100億行+ロード/日
- 20万クエリ/日

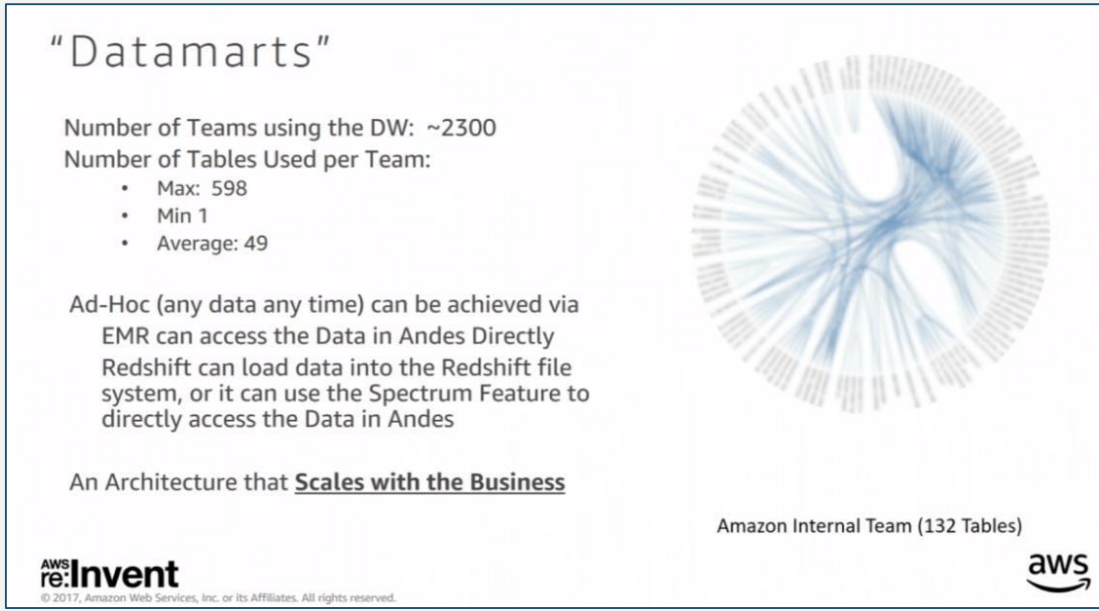
# データウェアハウスの利用状況と特長

データウェアハウス利用：2,300チーム

チームごとの表の数
- 最小1～最大598
- 平均：49

どんなチームも部分的に必要な
データを活用するため、
全部入のDWHは無駄が多い

分析環境にはスケーラビリティが
必須要件

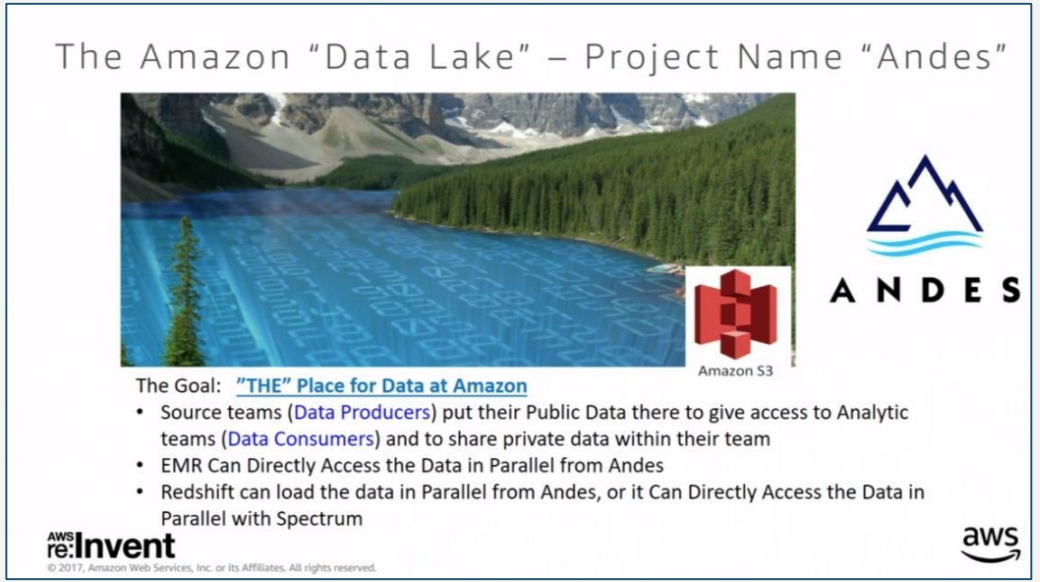


# Amazonのデータレイクプロジェクト "ANDES"

- アマゾンのデータを集めておく「THE（正式な）」場所

  "THE" Place for Data at Amazon

- S3がデータレイクの核




**ポイント）**

- データ提供元のチームはデータレイクにデータを配置し、分析チームにアクセス権限を付与
- EMR（Hadoop系処理）は、Andesから**並列で直接データレイクにアクセス可能**
- RedshiftはAndesから**並列でアクセスしてデータをロード**、またはSpectrumにより**並列で直接アクセス可能**

# 課題と解決：
# 調整からセルフサービスへ

[課題]

- どこにデータがあるか分からない
  - ✓ どこに有るか問い合わせる必要がある
- データを利用可能にするまでが大変
  - 自分達の環境にコピーしてもらうには、データ転送やETL等の仕組みの調整と構築するための申請（予算化）が必要
- データ鮮度が不明
  - データの更新頻度の調整に開発が必要

[解決]

- 利用者が検索出来るようにする
- 利用者が必要なデータを自分の領域に「登録(Subscribe)」出来るようにする
- 利用者が更新頻度の調整等を設定・変更出来るようにする

# セルフサービスを実現するための仕組み

- **収集と後続処理が分離**され、収集側がデータレイクにデータを置くだけで良い構成
- セルフサービスを促進するために**「発見」「登録」層を導入**

# データレイクポータルによる「発見」の実現

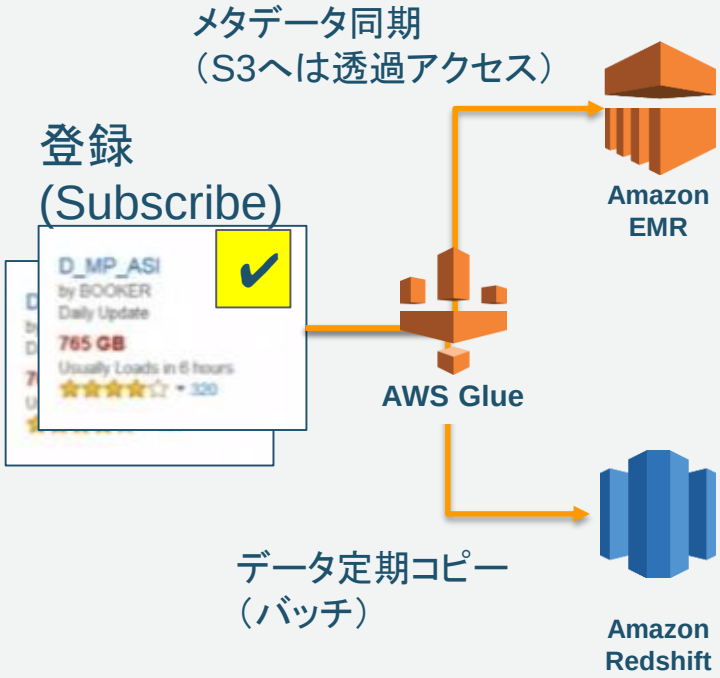
# 「登録」によるセルフサービスのデータ取得

## 関係者との調整や開発不要

欲しいデータを発見したら、登録（Subscribe）する

登録時にはデータの行き先を指定（自部門のRedshiftやEMR等）

自動的にコピー処理やメタデータ同期処理が構築され、定期的にフレッシュなデータが供給される

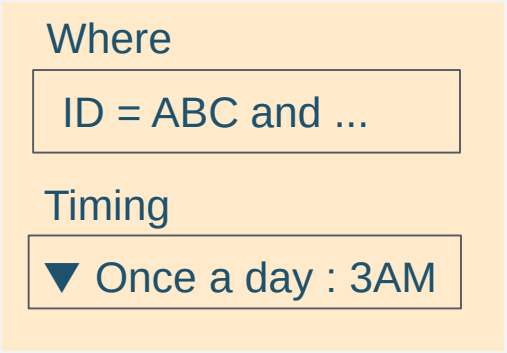


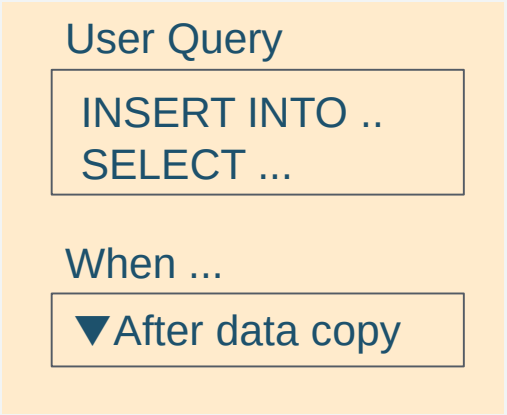
# 「登録」モデルの実現とその効果

ユーザ調整が可能な設計
- 必要なデータ範囲や頻度を設定可能
- 独自のクエリを登録して、データ連携のタイミングで自動実行

自動的なバリデーション（表行数チェック、スキーマチェック等）

誰が何を使っているか把握できるため、データの削除や変更時にも影響範囲が把握できるという効果あり



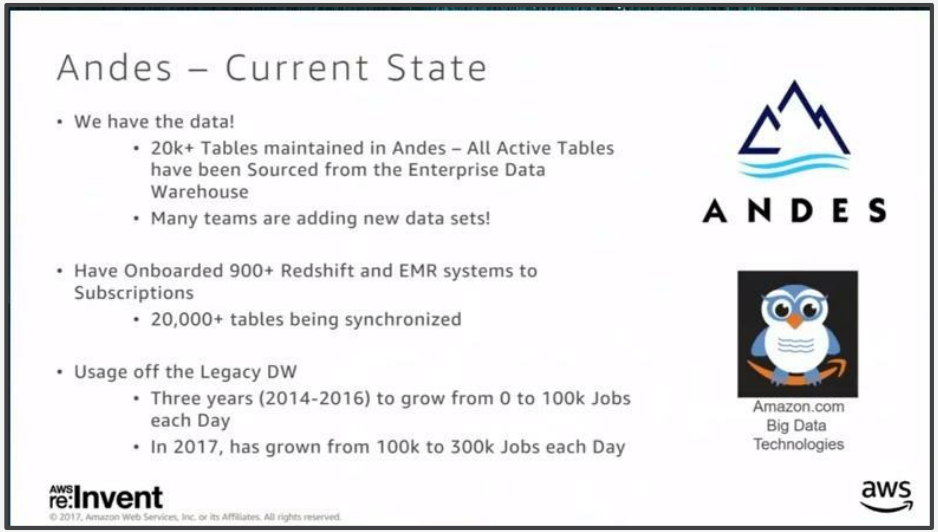
# 本章のまとめ: Amazon.comのデータレイク

## 徹底したセルフサービス化でデータ活用を促進

- 小売の経験を活かしたカタログ機能
  - ✓ メタ情報定義、利用者の評価など
- データレイクからの連携処理を手続き型での処理で自動化

## Andesの現在の状態：

- データレイクに20,000表があり同期されている
- データ連携先はRedshiftとEMRで計900+クラスター



# まとめ

# 本セッション全体のまとめ

- データレイクは、複数のデータウェアハウス、高度分析環境（機械学習など）が存在する場合の無駄、重複を最小化
- データレイクをゴミ溜めにしないためには
  - ✓ データスチュワード／データカタログ／セキュリティ
- AWSのサービスで構築するデータレイク
  - ✓ 代表的なサービス：S3 / Glue / Redshift / Athena / EMR
- Amazon.comの事例
  - ✓ Discover / Subscription を採用

# 参考資料

- Amazon S3
  - http://aws.amazon.com/jp/s3/
- AWS Glue
  - http://aws.amazon.com/jp/glue/
- Amazon EMR
  - http://aws.amazon.com/jp/emr/
- Amazon Athena
  - http://aws.amazon.com/jp/athena/
- Amazon Web Services ブログ：「AWS Glue と Amazon S3 を使用してデータレイクの基礎を構築する」
  - http://aws.amazon.com/jp/blogs/news/build-a-data-lake-foundation-with-aws-glue-and-amazon-s3/

# オンラインセミナー資料の配置場所

AWS クラウドサービス活用資料集

- https://aws.amazon.com/jp/aws-jp-introduction/


Amazon Web Services ブログ

- 最新の情報、セミナー中のQ&A等が掲載されています。
- https://aws.amazon.com/jp/blogs/news/

# 公式Twitter/Facebook
# AWSの最新情報をお届けします

@awscloud_jp


もしくは
http://on.fb.me/1vR8yWm

最新技術情報、イベント情報、お役立ち情報、
お得なキャンペーン情報などを日々更新しています！