

[講義資料]

Amazon Athena Federated Query とAmazon QuickSightで DynamoDBを可視化する

Akira Shimosako

Senior Solutions Architect

Amazon Web Services Japan, K. K.

自己紹介

下佐粉 昭（しもさこ あきら）

Twitter – @simosako

アマゾン ウェブ サービス ジャパン
シニアソリューションアーキテクト

著書（4人の共著）

「AWSではじめるデータレイク」



本資料について

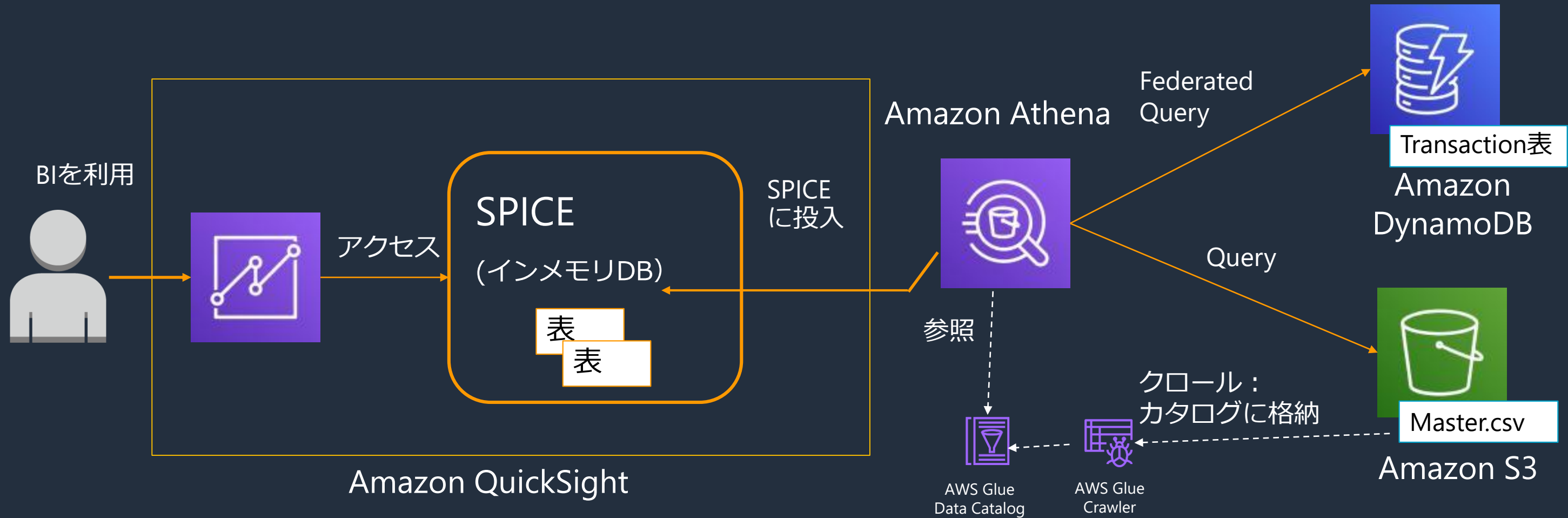
本ハンズオンでは、Amazon AthenaのFederated Query機能を利用してAmazon DynamoDBのデータをAthenaからクエリできるようにし、それをAmazon QuickSightで可視化するまでの流れを体験します

ハンズオン開始前にはQuickSightのサインアップとリージョンの確認が必要です（事前準備資料をご確認ください）

本資料ではハンズオンで利用するサービスについて簡単に説明します

ハンズオン全体像

- ① AthenaのFederated QueryでDynamoDBのデータをクエリする
- ② S3上のCSVデータとDynamoDBのデータをジョインする
- ③ クエリ結果をSPICEに格納し、QuickSightで可視化する



ハンズオンで利用するサービス

Amazon Athenaと Amazon QuickSight

Amazon Athena



Amazon S3上のデータに対して**標準 SQL** による
インタラクティブなクエリを投げて
データの分析を行うことができる
サーバレスのサービス

Amazon 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 (Prestoベース)



ロードなしに Amazon S3 のデータを直接クエリ



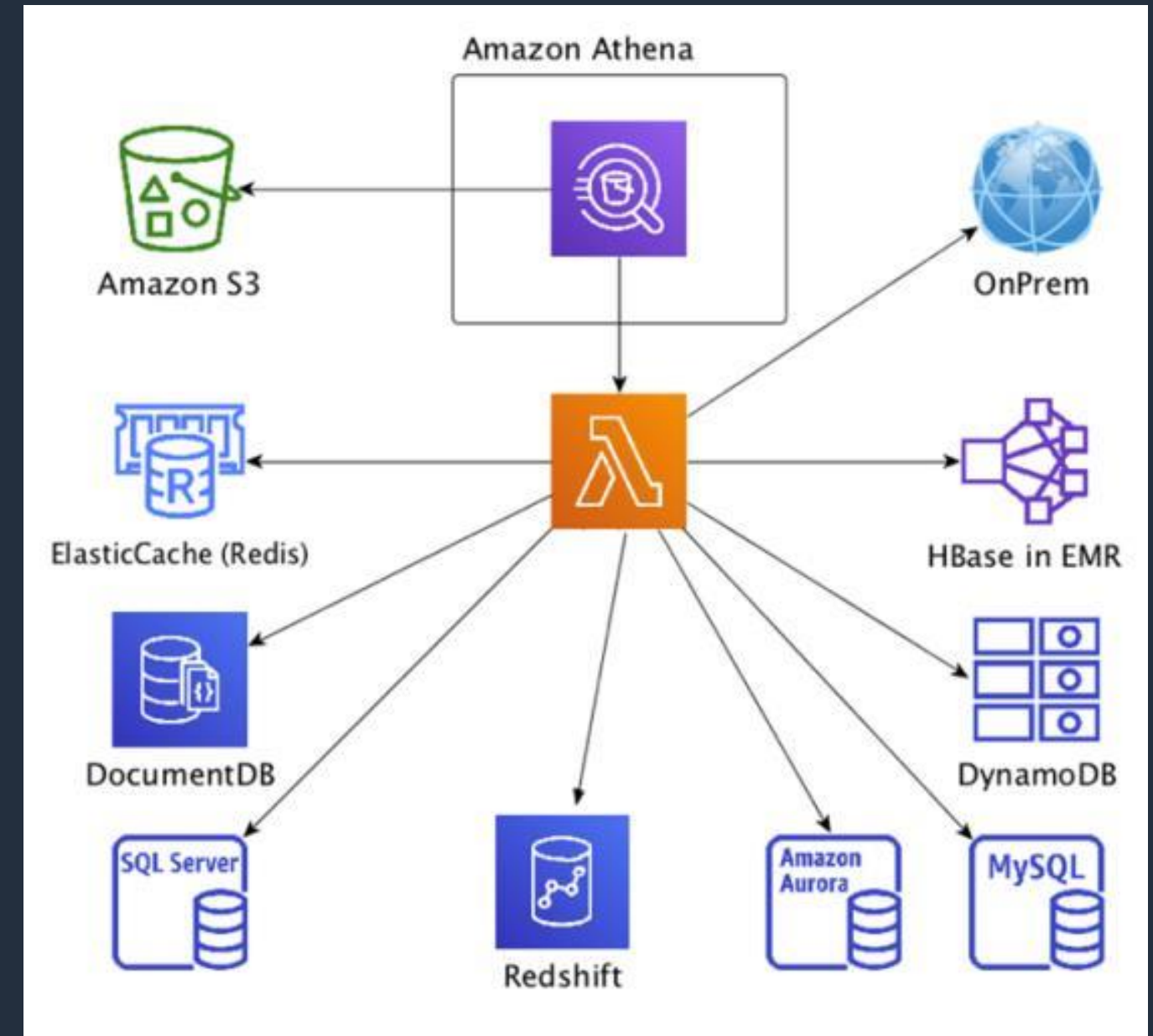
スキャンしたデータに対しての従量課金



JDBC / ODBC / API 経由で BI ツールやシステムと連携

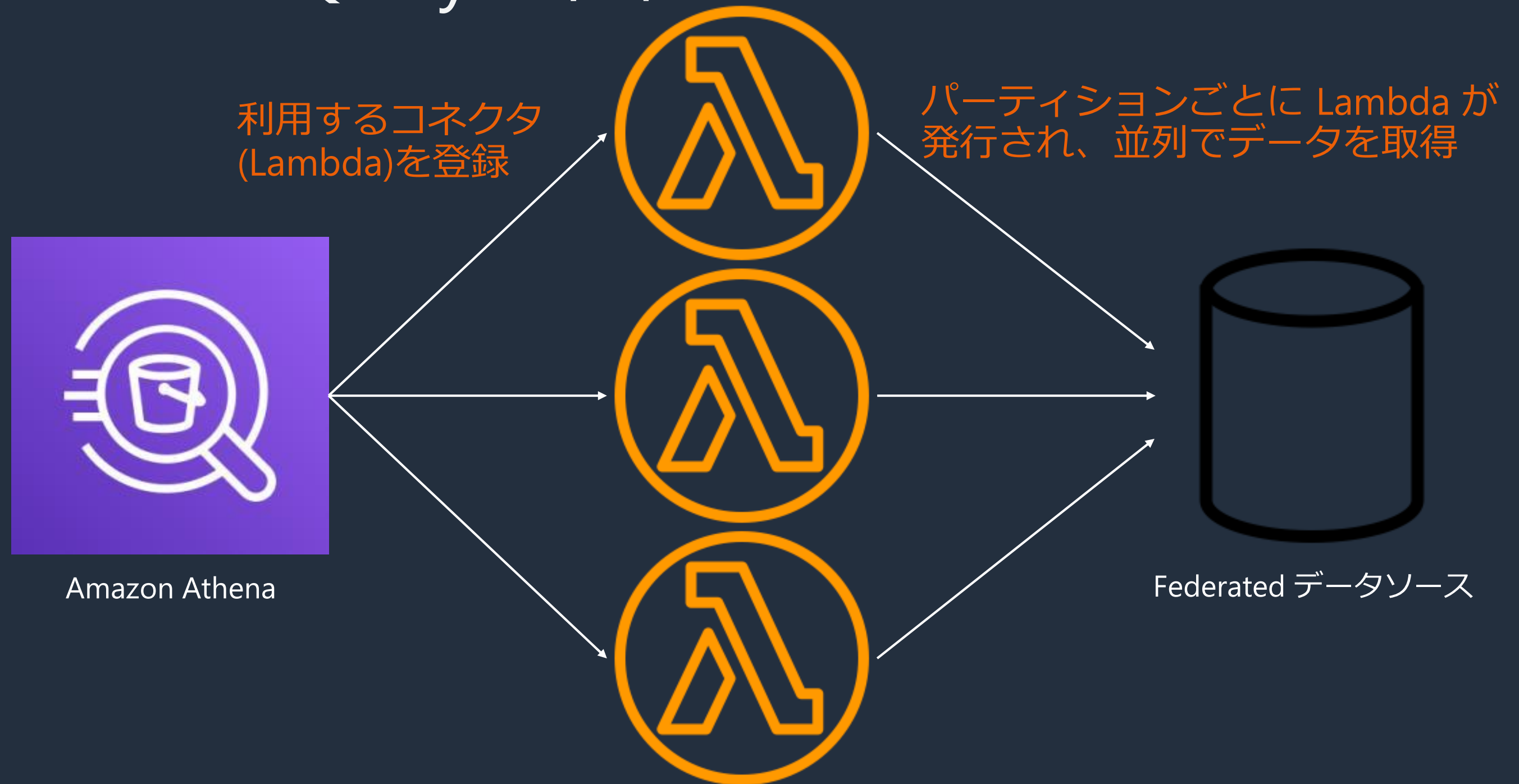
Amazon Athena Federated Query

- ❑ S3上のデータ以外に、多数のデータソースにアクセス可能にする
- ❑ 複数データソースをまたいでクエリを実行
- ❑ AWS Lambdaによる分散処理
- ❑ カスタムコネクタを作成することで、独自のデータソースにも対応可能



参考) <https://aws.amazon.com/jp/blogs/news/query-any-data-source-with-amazon-athenas-new-federated-query/>

Federated Query の仕組み



すぐにデプロイ可能なコネクタ

AWS オフィシャル

- ❑ Neptune
- ❑ ElasticSearch Service
- ❑ Timestream
- ❑ DocumentDB
- ❑ DynamoDB
- ❑ CloudWatch
- ❑ 汎用JDBC

等

AWS Serverless Application Repository
で作成コネクタの共有が可能

<https://github.com/aws-labs/aws-athena-query-federation/wiki/Available-Connectors>

Athena Federated Queryの制限事項

- Athena Query Engine v2が必要
 - 利用可能リージョンに注意
- Federated Queryで参照される表について
 - VIEWの作成が不可
 - (Connector側で自動的に表を認識するため) CREATE TABLEやDROP TABLEは利用不可
- 詳細は以下ドキュメントを参照してください
 - <https://docs.aws.amazon.com/athena/latest/ug/connect-to-a-data-source.html>

Athenaの料金

- クエリ単位：S3 のデータスキャン 1TB につき \$5
 - 最低スキャン量は 10MB （それ以下は10MBに切り上げ）
 - スキャン量の MB 未満の端数は切り上げて計算
- DDL および実行に失敗したクエリは、課金の対象外
- クエリ実行にともなって実行される Amazon S3 API リクエストや通信費用については別途発生
- Federated Queryの場合は、利用したLambdaの費用が別途発生

補足：フェデレーションとレプリケーション

どちらが良いということではなく、使い分け

フェデレーション

- データを事前コピーせず、都度アクセスする手法
- メリット：リアルタイム性
- 例：
 - Athena Federated Query
 - Redshift Federated Query
 - Redshift Spectrum
 - DB Link 等

レプリケーション

- データを事前コピーしておく手法
- メリット：性能、可用性の確保が容易
- 例：
 - Glue (ETL)
 - Glue Elastic Views
 - DMS
 - マテリアライズドビュー 等



Amazon QuickSight

サーバーレスで提供され、スケーラブル
多くのユーザで利用可能な料金体系
多様なアクセス手段



サーバーレスで大規模対応

管理するサーバーなし、自動スケールアウト
高速インメモリデータベース内蔵



全社員で利用可能な料金体系

利用した分だけのReader料金
使わなければゼロ円



多様なアクセス手段

アプリケーションに容易に組み込み
メール配信、モバイルアプリからのアクセスに対応

直感的な可視化を実現するBIサービス



- 簡単な操作で可視化を実現
- ブラウザだけで利用可能
- 数クリックで共有
- フィルタ、ドリルダウン、ユーザ入力等の機能
- テーマによる見た目のカスタマイズ

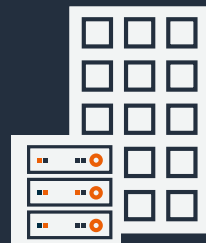


多様なデータソースに接続

- AWS内のサービスとの統合だけでなく、オンプレミスのデータソースやサードパーティーのビジネスアプリケーション（SaaS）にも対応
- Amazon AthenaのFederated Query経由で、下記以外のデータソースにも対応可能

オンプレミス

オンプレミスにセキュアに接続。
S3やEXCELファイルのアップロードに対応



- Excel
- CSV
- Teradata
- MySQL
- SQL Server
- PostgreSQL
- Oracle

AWSクラウド

ConnectS3上のデータやVPC内のデータベースにセキュアに接続



- Redshift
- RDS
- S3
- Athena
- Aurora
- Teradata
- MySQL
- Oracle
- SQL Server
- Timestream
- Amazon Elasticsearch Service
- Presto
- Spark
- Postgre SQL
- MariaDB
- Snowflake
- IoT Analytics

SaaS

3rdパーティーのSaaSに接続



- Salesforce
- Square
- Adobe Analytics
- Jira
- ServiceNow
- Twitter
- Github

インメモリデータベース：SPICE

- S3上のファイル等、非RDBのデータをSPICEに取り込むことで分析可能にする
- 高頻度でアクセスされるデータをSPICEに取り込むことで、データベースの負荷低減
- SPICEに定期的に読み込む設定が可能



MLインサイト 内蔵の機械学習ベースのインサイト

専門家不要で使える
インサイト（洞察）機能を提供

1. MLベースの異常検知：異常値を自動発見し、報告
2. MLベースの予測：過去の値から将来を予測
3. 自動ナラティブ：分かりやすい文章で分析結果を提供
4. QuickSight Q：ユーザの言葉を理解して可視化 (preview)



Amazon QuickSightの料金体系

Enterprise EditionのReaderライセンスは、使った分だけの利用
使わなければゼロ円。最大でも\$5/ユーザ/月の上限

Author

Reader

ダッシュボードを作成して公開

\$18

/ ユーザー / 月

年単位契約

\$24 /ユーザー/ 月（月単位契約）

作成済みダッシュボードを閲覧

\$0.30 最大で

/ セッション*

\$5

/ ユーザー / 月

*1 セッション = ログインから 30 分間

※この他、ユーザ認証不要で利用するためのセッションキャパシティライセンスもあります

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

QuickSightを無料で利用する

Free Trial（試用期間）

Standard EditionやEnterprise Editionを60日間無料で試用可能

正確には2回の支払い締め日を迎えるまで利用可能

SPICEを10GB利用可能

試用期間中のユーザ数は合計4名まで

Free Tier（無料枠）

無料で期限無しに利用し続けることが可能

SPICEを1GB利用可能

無料枠で利用可能なユーザ数は1アカウントあたり1名のみ

SPICEが不足した場合は利用料金を払うことで容量を追加可能

参考情報

Amazon Athenaホームページ

<https://aws.amazon.com/jp/athena/>

Athena Federated Query 解説

<https://aws.amazon.com/jp/blogs/news/query-any-data-source-with-amazon-athenas-new-federated-query/>

Athena Federated Queryのコネクター

<https://github.com/aws-labs/aws-athena-query-federation/wiki/Available-Connectors>

Amazon QuickSightホームページ

<https://aws.amazon.com/jp/quicksight/>

Amazon QuickSight ギャラリー（埋め込まれたダッシュボードのサンプルが確認できます）

<https://aws.amazon.com/jp/quicksight/gallery/>

Amazon QuickSightハンズオン資料（QuickSightの操作方法のハンズオン）

<https://aws.amazon.com/jp/blogs/news/amazon-quicksight-handson-202006/>

内容についての注意点

- 本資料では2021年3月26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おります。日本居住者のお客様には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.