



# Zero-ETL 統合を 活用した業務データベースの ニアリアルタイム分析

濱岡 洋太

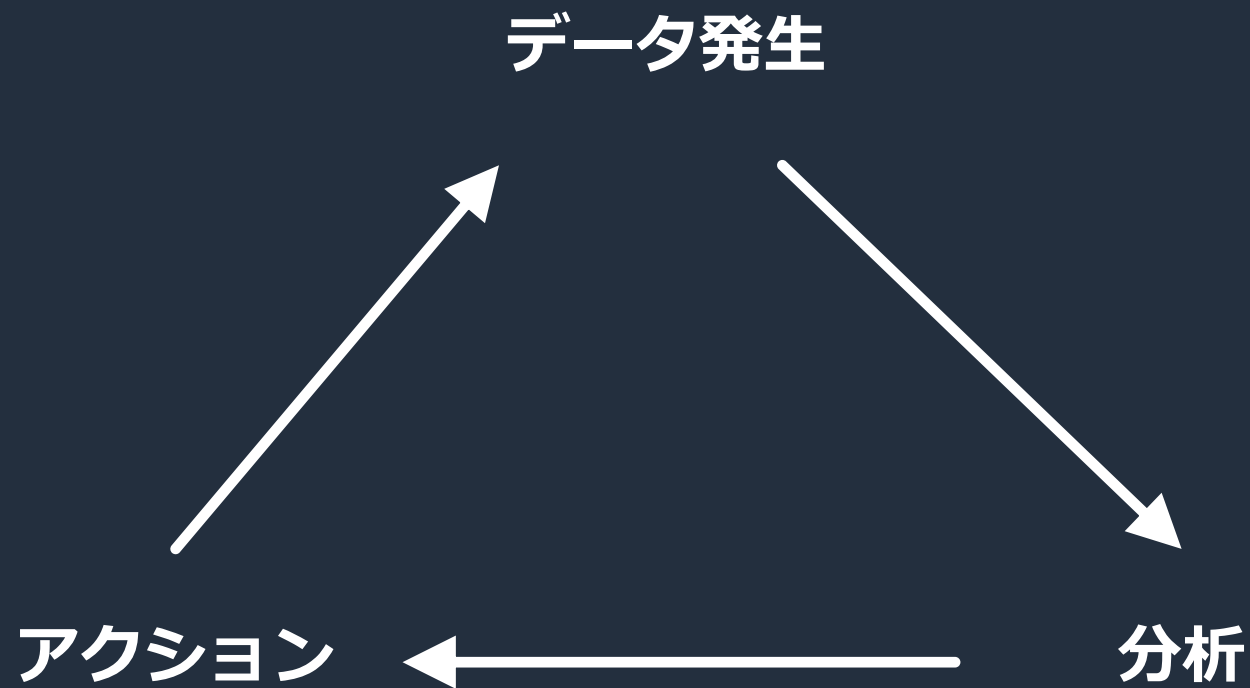
Redshift スペシャリスト  
ソリューションアーキテクト  
アマゾンウェブサービスジャパン合同会社

# アジェンダ

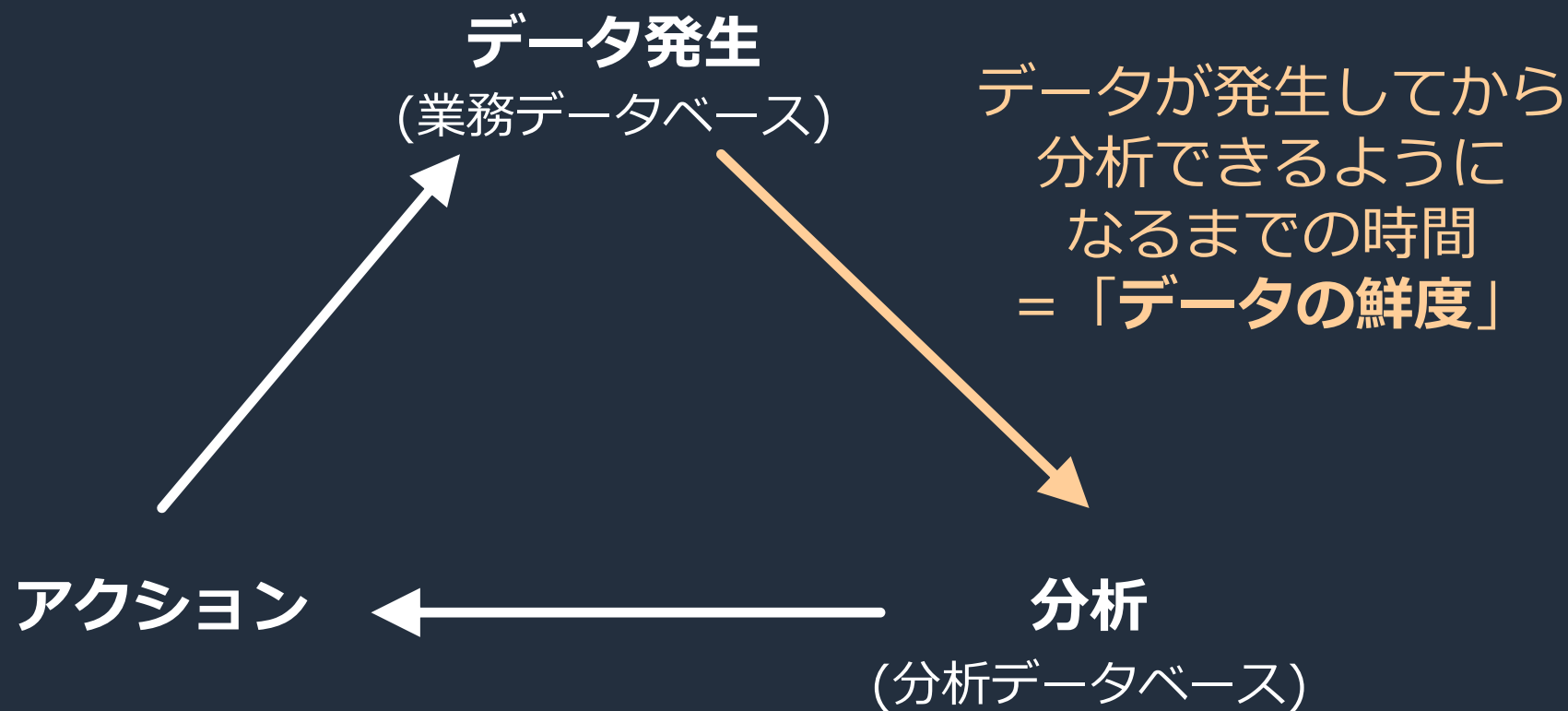
- 業務データベースを分析する際の課題
- Zero-ETL 統合と Amazon Redshift を使ったニアリアルタイム分析
- デモ

# 業務データベースを 分析する際の課題

# データ分析のサイクル

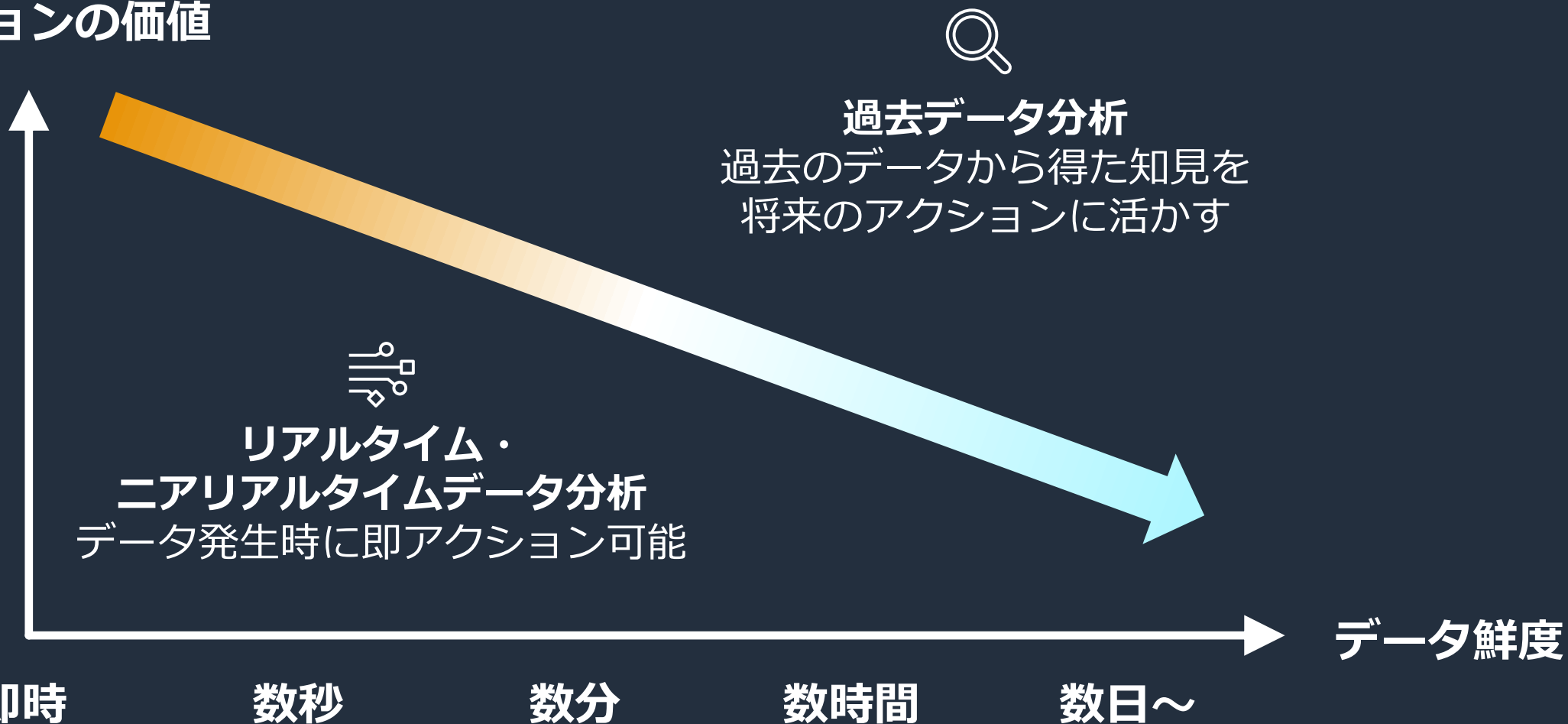


# データ分析のサイクル



# データ鮮度はビジネス価値に直結

## アクションの価値



# リアルタイムな分析のユースケース

## モニタリング



売上分析



製造ライン分析



ゲームの  
リーダーボード

## 異常検知



顧客の離反防止



不正検知



システム  
異常検知

## 最適化



マーケティング  
最適化



在庫最適化



パーソナ  
ライゼーション

# 業務データベースと分析データベースの違い

## 業務データベース

業務アプリケーションの  
データストア

トランザクション処理  
(OLTP)

行指向

## 分析データベース

データ分析

大量データの集計、集約、  
フィルタ処理 (OLAP)

列指向

主な用途

得意な処理

エンジンの特徴



# 業務データベースのデータを分析する際の課題

業務データベース



データ連携



データ鮮度の低下

ETL 処理の構築・運用

分析データベース



→素早く、簡単にデータを連携したい

# Zero-ETL 統合と Amazon Redshift を使った ニアリアルタイム分析

# AWS が提供するデータウェアハウス: Amazon Redshift

## あらゆるデータを分析



- 業務データベースやデータレイクを統合分析
- 一貫したセキュリティとガバナンスを実現

## 優れたパフォーマンス



- 他のクラウドデータウェアハウスと比較し最大3倍のコストパフォーマンス
- 機械学習による自動最適化

## 低いコスト



- 使った分だけのお支払い
- 予測可能なコストによりスモールスタートが可能

# Amazon Aurora Zero-ETL 統合

## ETL パイプラインなしでニアリアルタイム連携

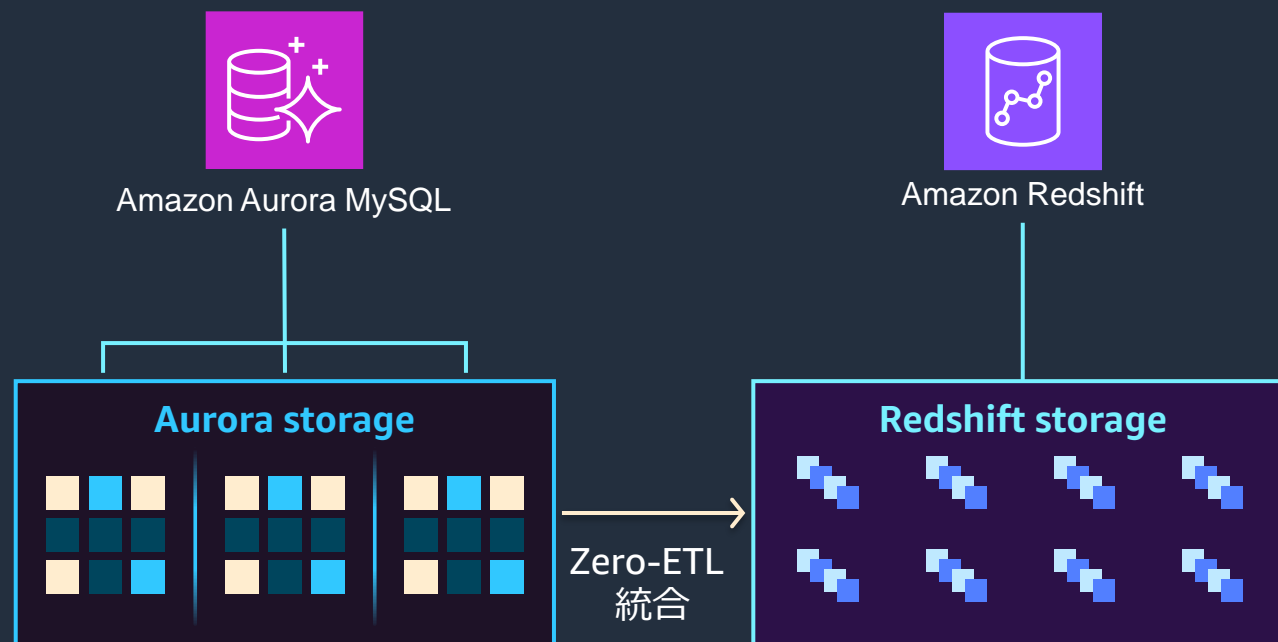
- ✓ データ鮮度向上
- ✓ 簡単にデータを連携

## 高いパフォーマンス

- ✓ 数秒～数十秒以内に同期
- ✓ サイジング・チューニング 不要

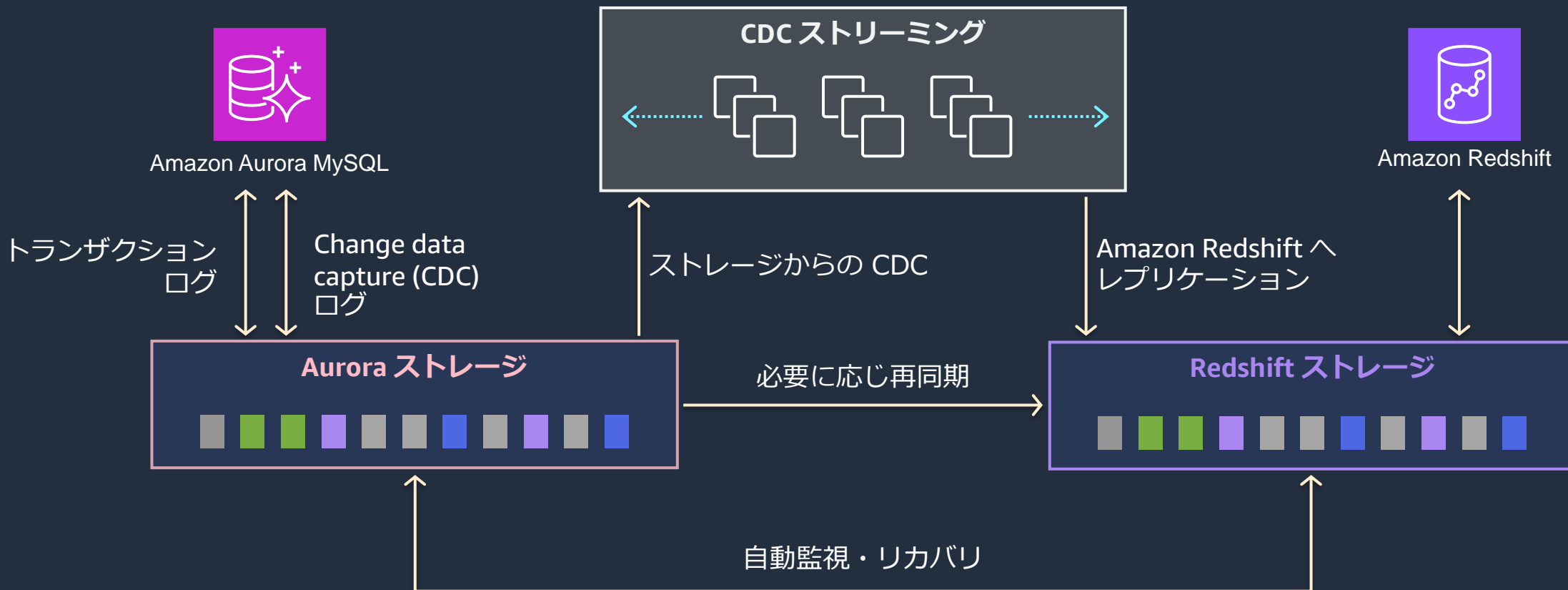
## 対応データソース

- ✓ Amazon Aurora MySQL 3.03.1 (MySQL 8.0.26 互換) 以降



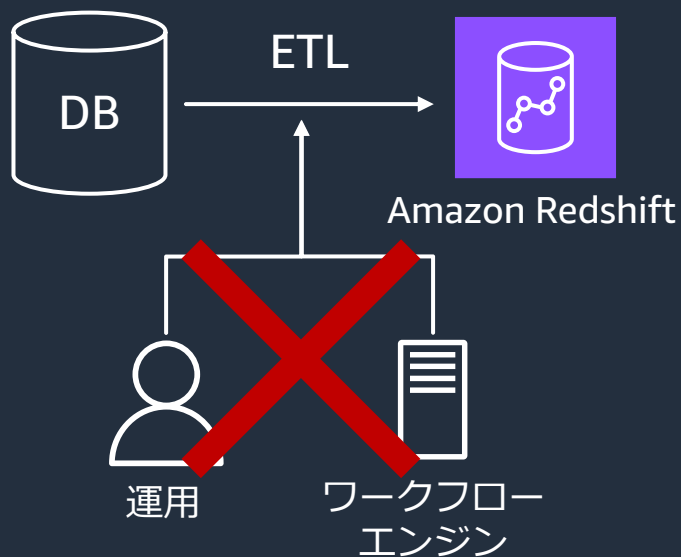
# Amazon Aurora Zero-ETL 統合

## ストレージレイヤーでの統合

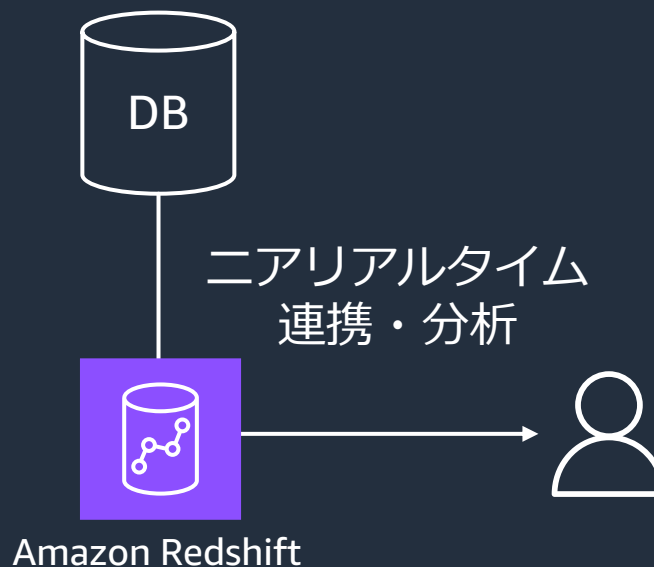


# Zero-ETL を使うメリット

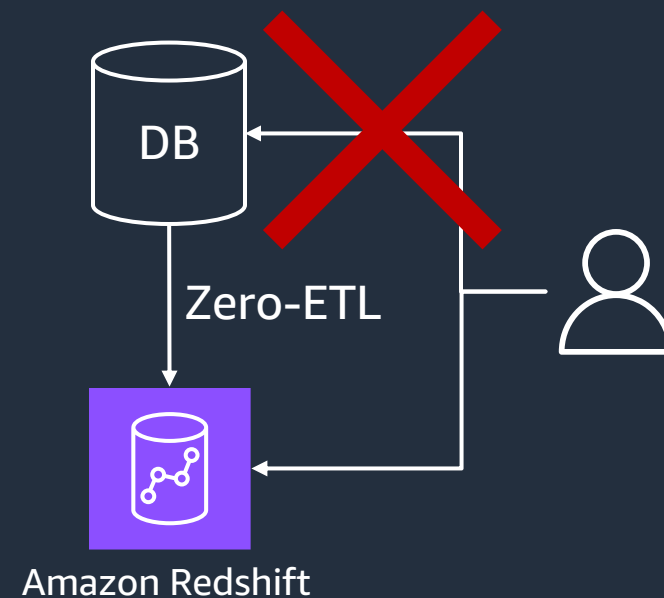
## ETL の構築・運用軽減



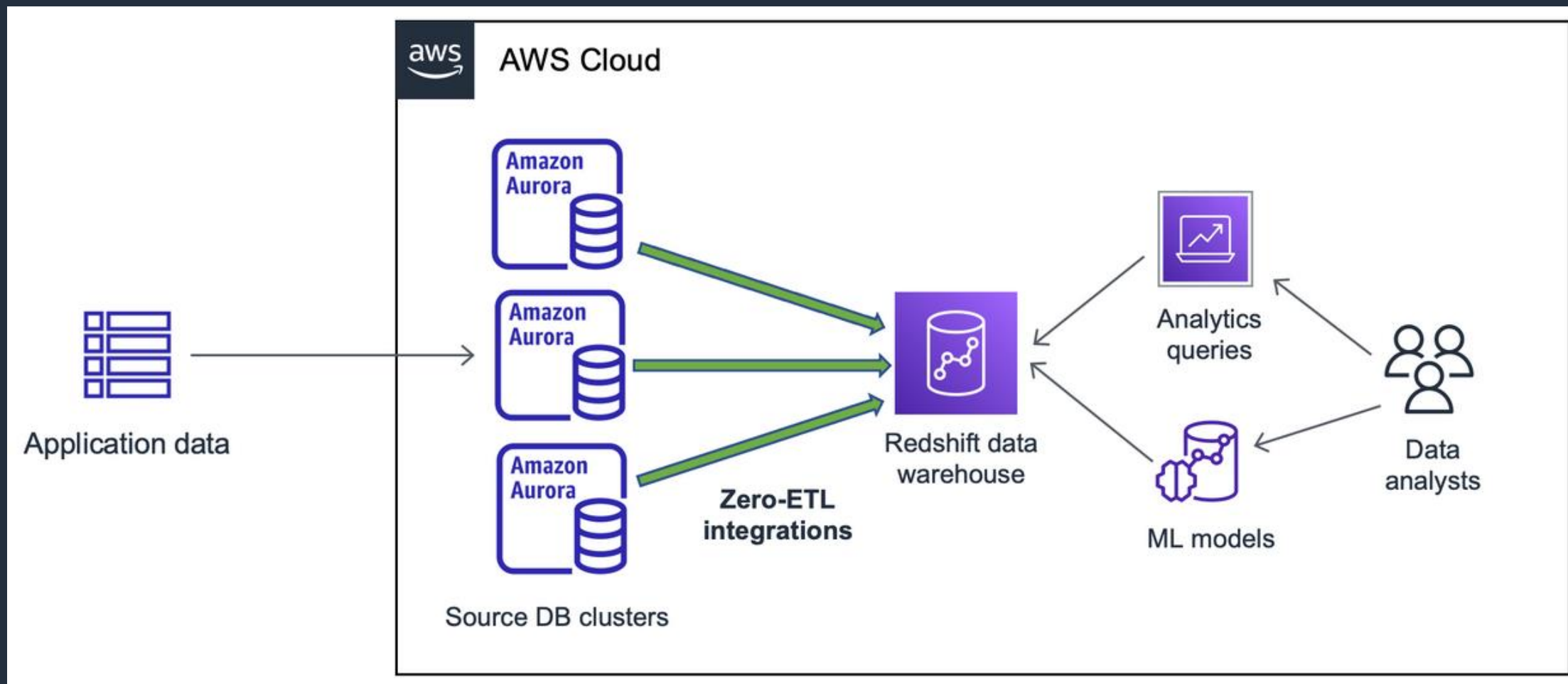
## ニアリアルタイムでの分析



## DB への負荷軽減



# Zero-ETL により複数のデータソースを1箇所で分析



# デモ



# デモシナリオ

Webアプリ

業務データベース

分析データベース

可視化



- ポイント
- ・設定の手軽さ
  - ・データの鮮度

## 制限事項

- 一部のサポートされていないデータ型 (BINARY, BIT, JSON 等) を含むテーブルは同期されず、コンソールにワーニングが表示される
- Zero-ETL 統合の作成と管理は AWS コンソールからのみ可能
- 異なるリージョンでの Zero-ETL 統合はサポートされていない
- Redshift DC2, DS2 ノードタイプはターゲットとしてサポートされていない
- 同期するテーブルを選択することはできないため、Aurora 内の全テーブルが同期される

※ プレビュー中の機能のため、GA (一般利用開始) 時に変更の可能性あり

[https://docs.aws.amazon.com/ja\\_jp/redshift/latest/mgmt/zero-etl.reqs-lims.html](https://docs.aws.amazon.com/ja_jp/redshift/latest/mgmt/zero-etl.reqs-lims.html)

## よくある質問

Q. binlog を有効化することで Aurora のパフォーマンスに影響はありますか？

A. 拡張 binlog を利用するため、binlog を有効化することによる Aurora への負荷影響は最小限となります。

Q. Zero-ETL を有効化することで Redshift のパフォーマンスに影響はありますか？

A. データの同期はストレージレイヤで行われるため、Redshift のコンピュータリソースへの影響は最小限です。

Q. Aurora MySQL 以外のデータソースは対応していますか？

A. 対応データソースは Aurora MySQL 3系以降のみとなります。



**Thank you!**