Amazon Aurora Deep Dive (Part-2)

Amazon Aurora で実現する高可用性構成

内山　義夫

アマゾン ウェブ サービス ジャパン 合同会社
データベース スペシャリスト ソリューション アーキテクト

内容についての注意点

- 本資料では 2022年8月23日時点のサービス内容および価格について説明しています。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ご利用される場合、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.

自己紹介

内山　義夫

データベース スペシャリスト ソリューション アーキテクト

好きなAWSサービス

- Amazon RDS
- AWS Schema Conversion Tool
- AWS Database Migration Service

Agenda

- データベースの高可用性構成で考慮すべきこと
- 一般的なオンプレミスでの構成
- Auroraでの構成例
- まとめ

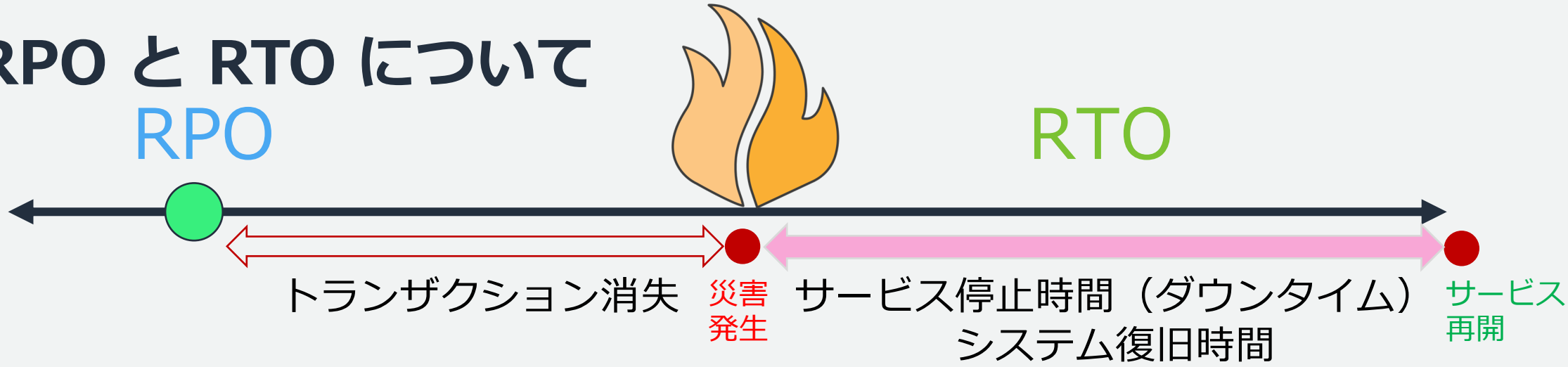
データベースの高可用性構成で考慮すべきこと

データベースの高可用性で考慮すべきこと

- **RTO: Recovery Time Objective**
  - リカバリ時間目標
  - 障害時に、どれだけの時間でリカバリを完了させるか

- **RPO: Recovery Point Objective**
  - リカバリポイント目標
  - 障害時に、どの時点のデータまでリカバリ可能か

RPO と RTO について



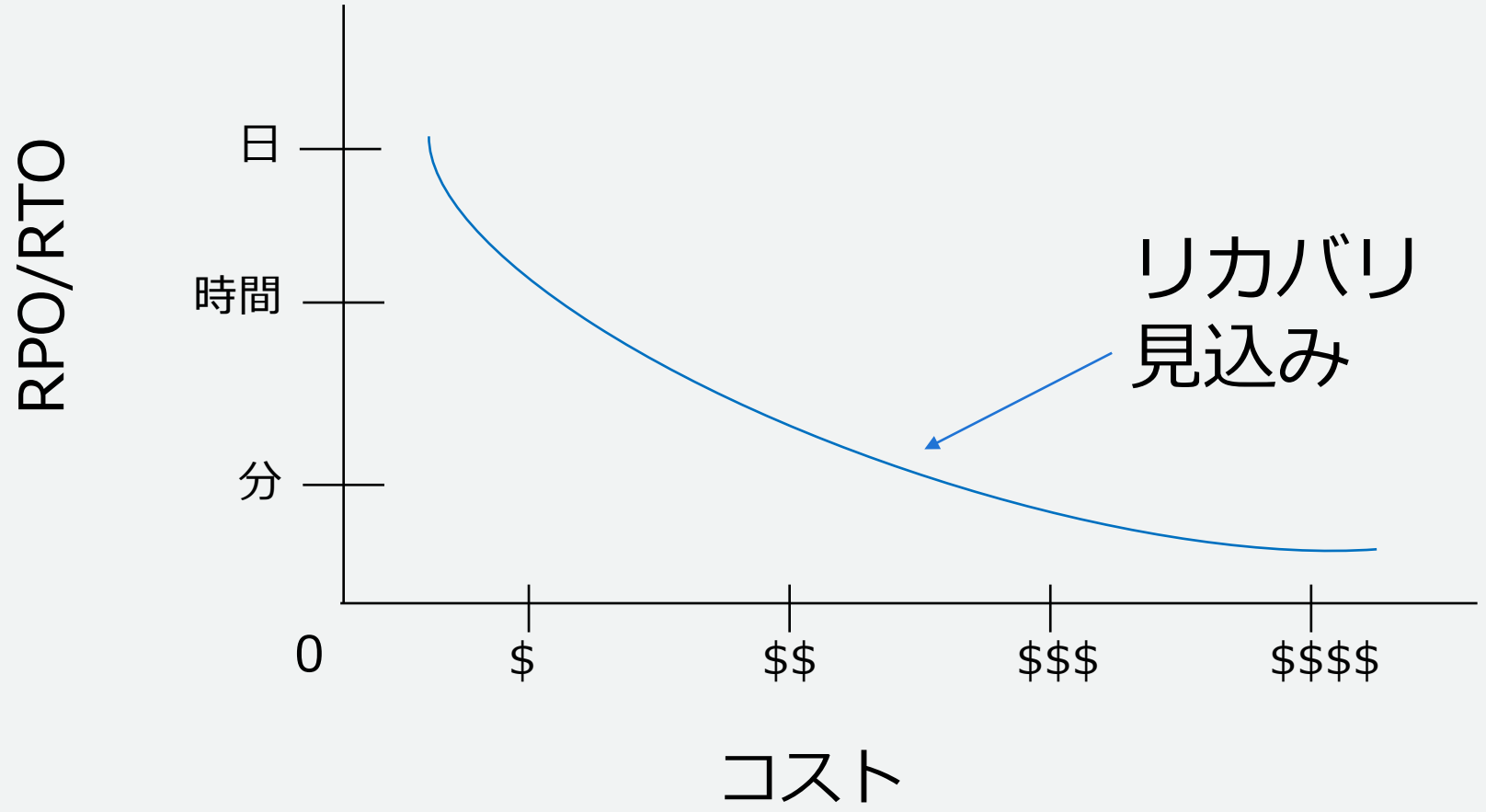
Recovery Point Objective (RPO)

- ディザスタリカバリの復旧時点の目標値バックアップ・データを取得するタイミングや頻度のこと
- どの程度のデータ損失が許容されるか？
- バックアップ・リストアの運用間隔や、データレプリケーションの技術選択に影響

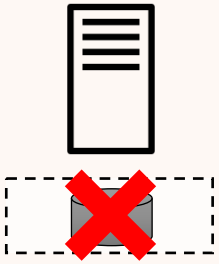
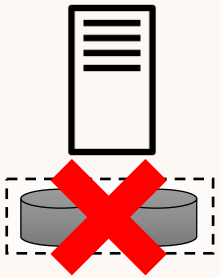
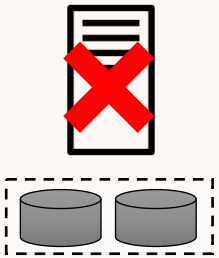
Recovery Time Objective (RTO)

- 該当サービス復旧にかけられる時間
  1分、15分、1時間、4時間、1日？
- 対象障害に応じて設定するのが一般的
- データリストアや、システム再起動等の技術選択に影響

リカバリ見込みとコスト



データベースの障害の種類

	ディスク障害	データ破損	インスタンス障害	データセンター障害	大規模障害
障害					
概要	複数のディスクで冗長化されている環境で、1本のディスクが物理的に完全に壊れた状況	複数のディスクで冗長化されている環境で、複数のディスクが同時に物理的に完全に壊れたり論理障害など、そのストレージから回復不能なデータが発生した状況	CPU故障などで、1台のコンピュータが物理的に完全に壊れた状況	電力、空調、ネットワーク障害などで、データセンター規模で障害が発生した場合。AWSの場合、AZ障害を想定。	地震などで大規模な災害が発生した場合。AWSの場合、リージョン内の全てのAZで障害が発生したことを想定。

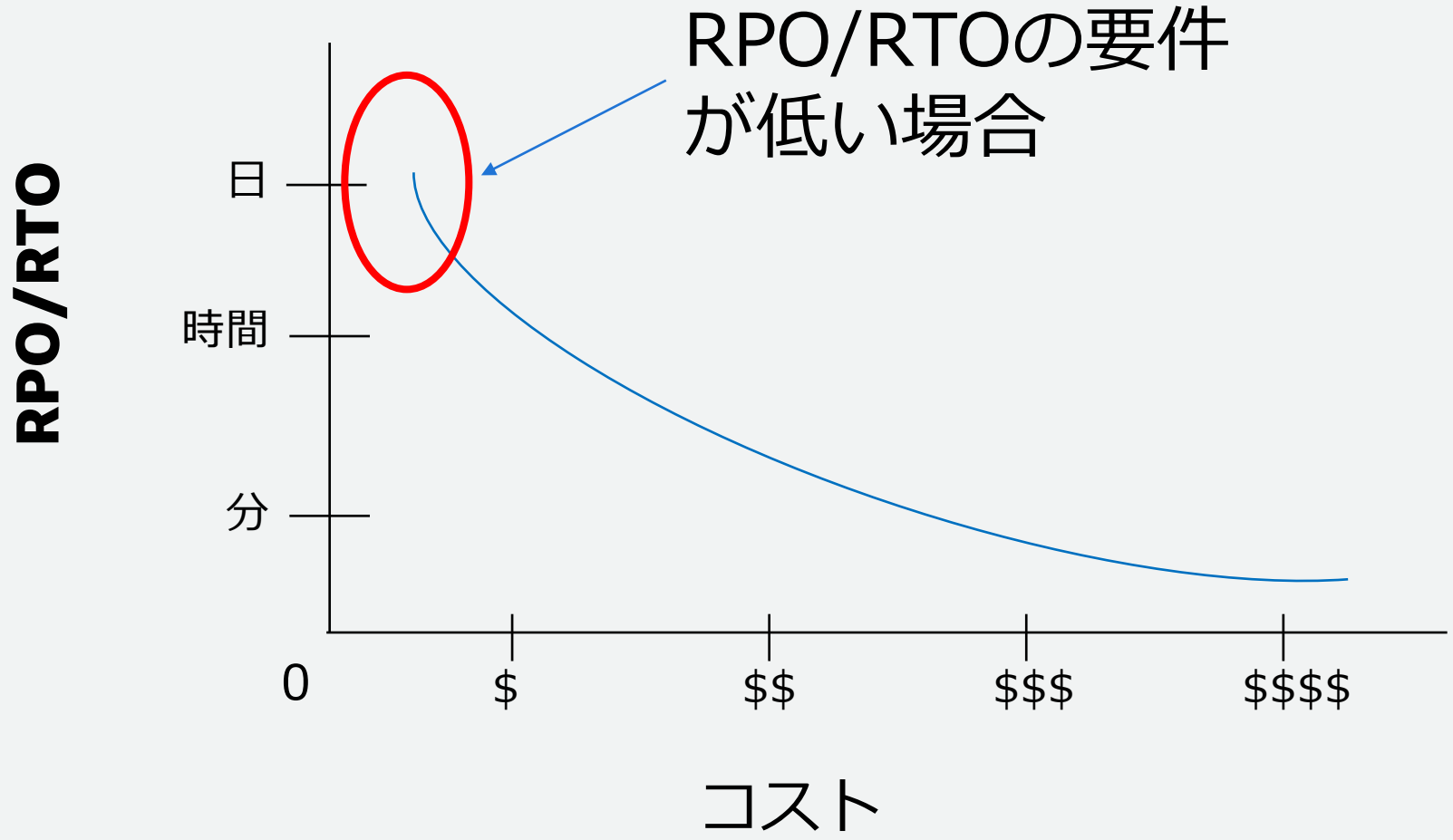
一般的なオンプレミスでの構成

一般的なオンプレミスでの構成

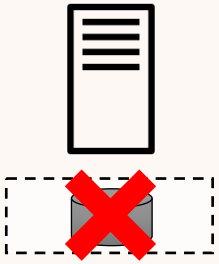
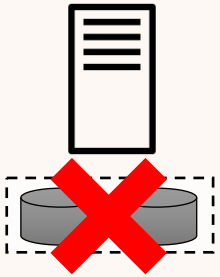
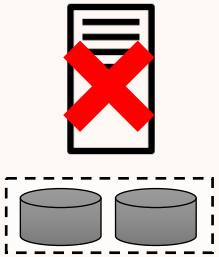
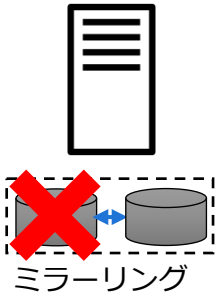
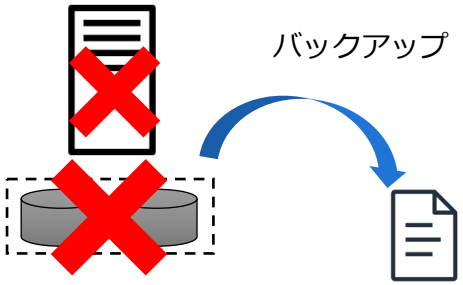
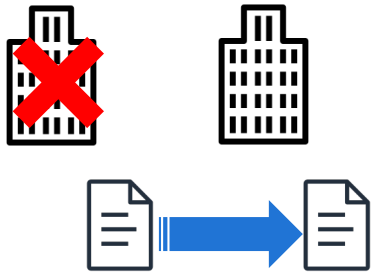
- **RPO/RTOの要件が低いシステム**
  - 自社システムなどユーザー数が少なく復旧できれば問題ないようなシステム

- **RPO/RTOの要件が高いシステム**
  - ミッションクリティカルなシステム

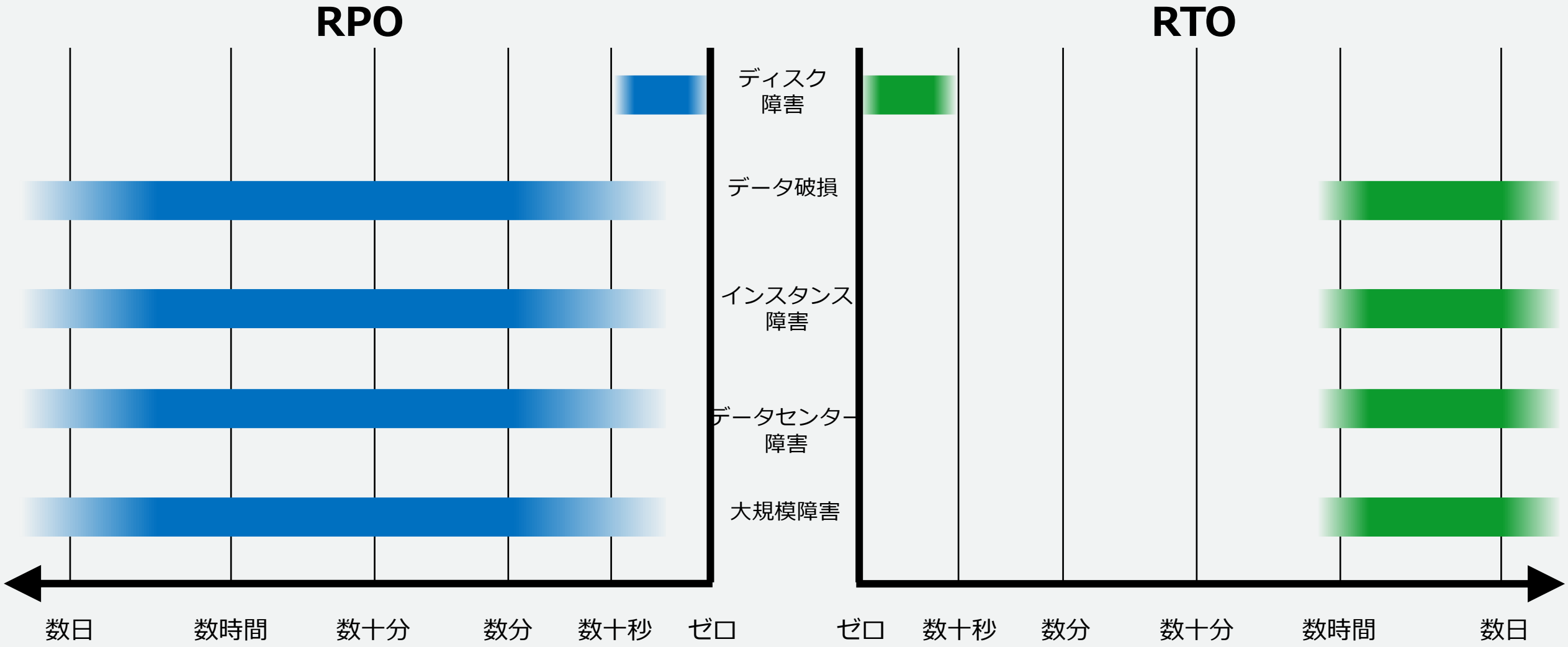
リカバリ見込みとコスト



RPO/RTO要件が低い場合の構成例（オンプレミス）

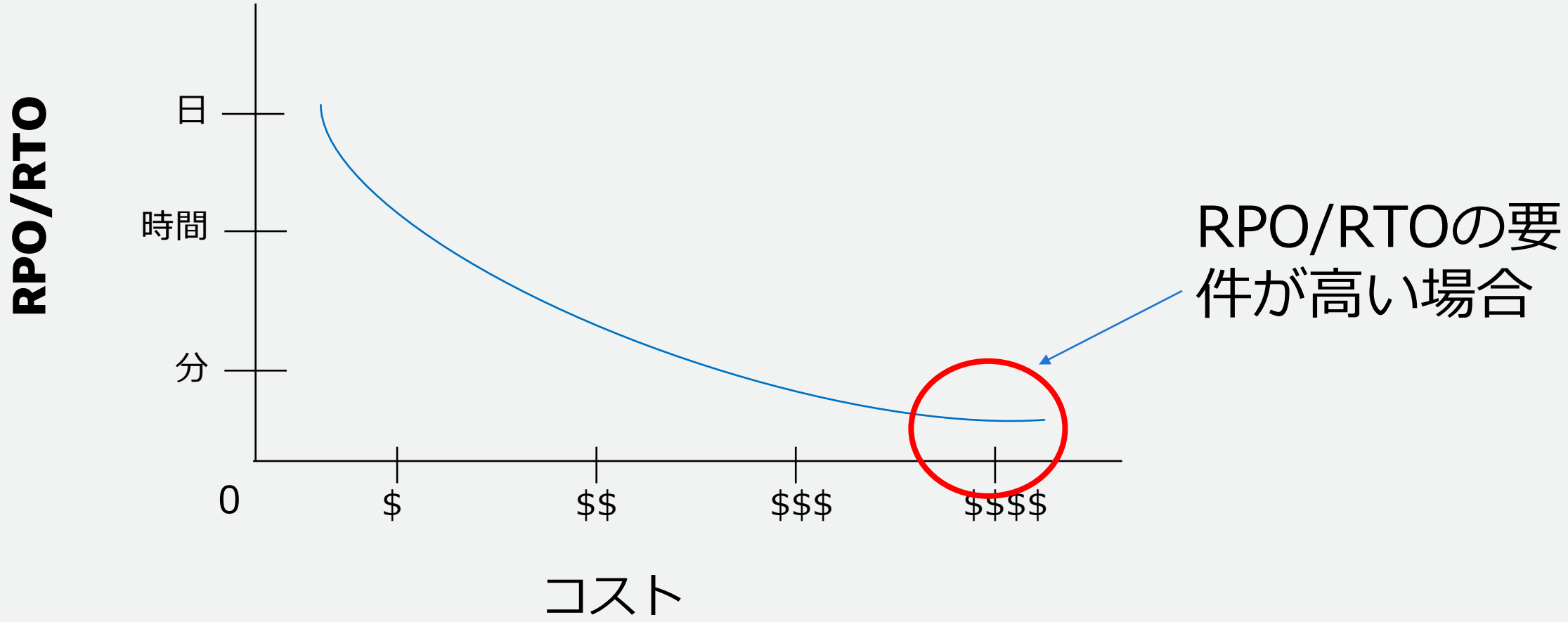
	ディスク障害	データ破損	インスタンス障害	データセンター障害	大規模障害
障害					
構成例					
概要	ディスクのミラーリングによる冗長化	データベースのバックアップの取得		データベースのバックアップを別のDCに転送	データベースのバックアップを遠隔地にあるDCに転送

要件が低い構成例のRPO/RTO

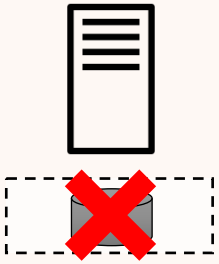
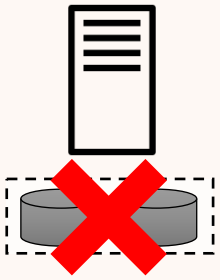
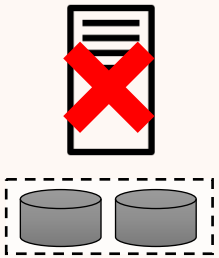
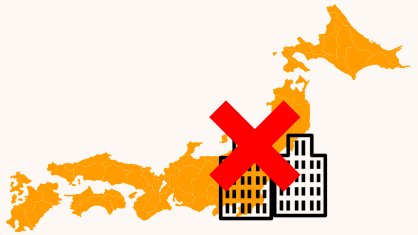
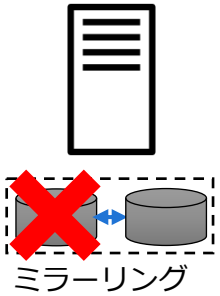
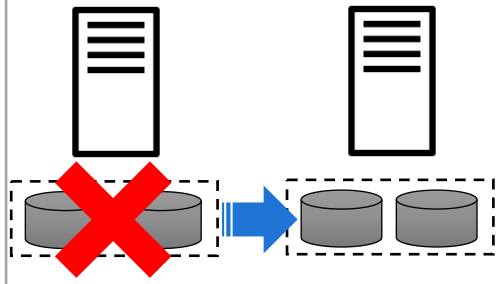
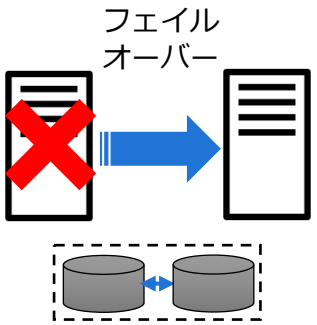
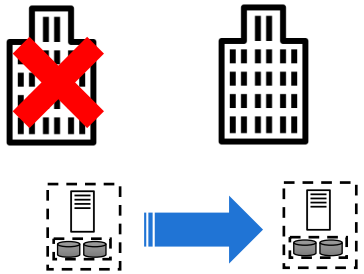
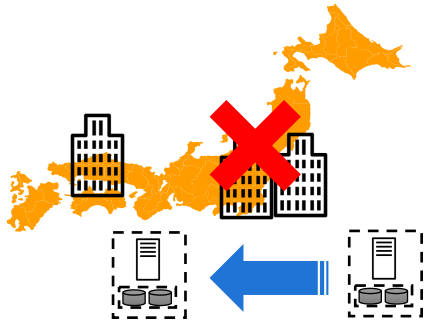


※一般的な傾向であり、約束されたものではありません

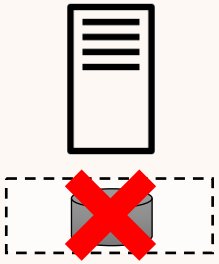
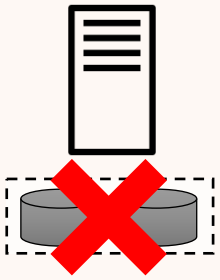
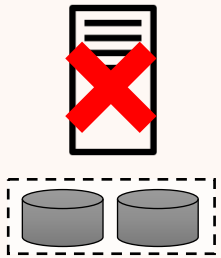
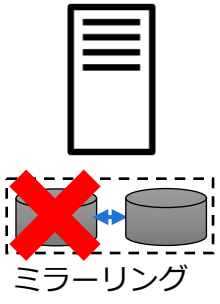
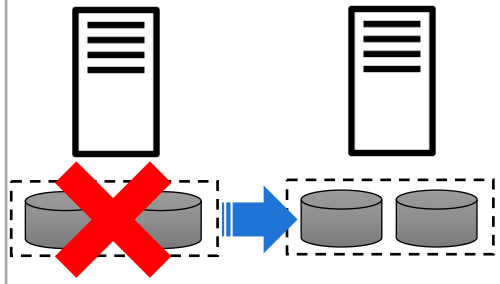
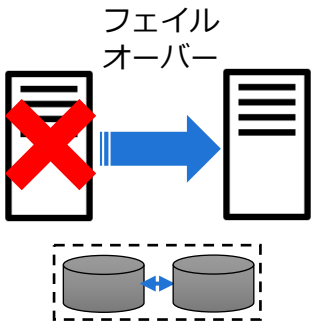
RPO/RTOの要件が高い場合



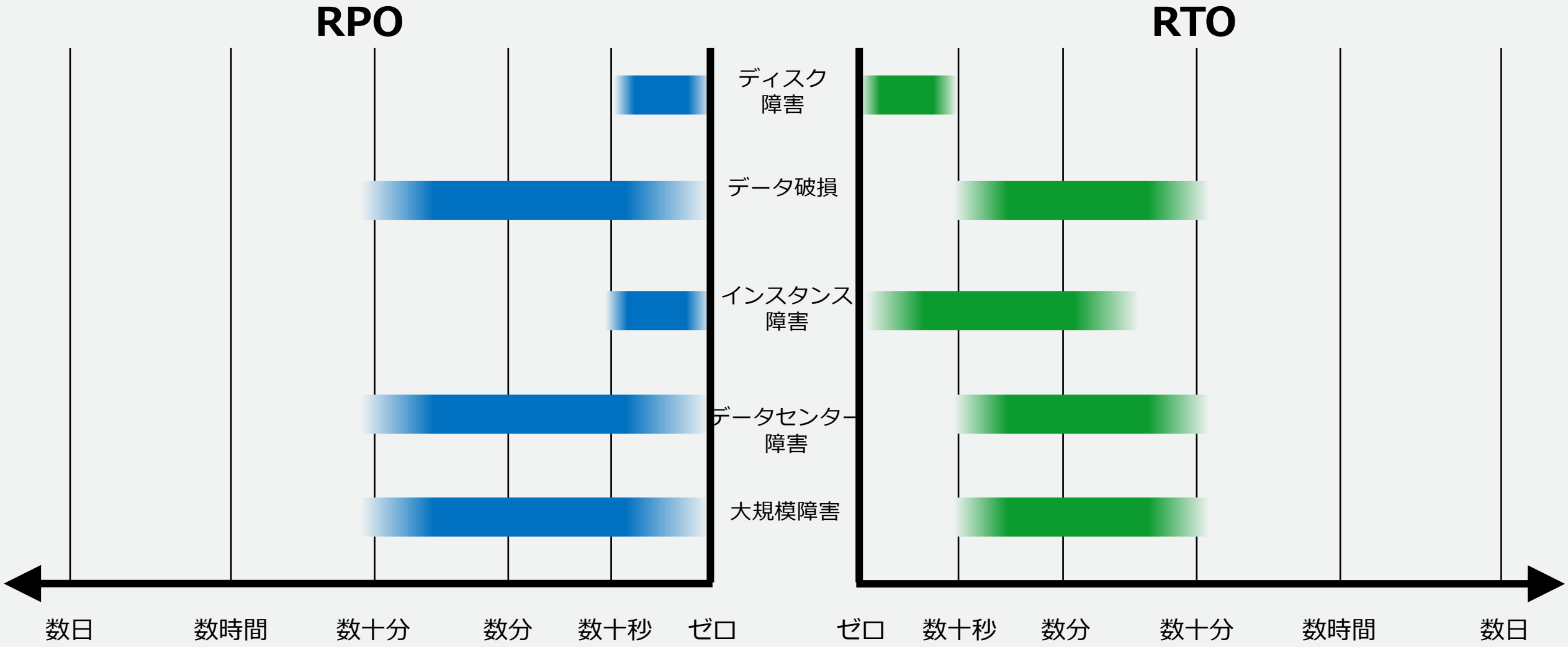
RPO/RTO要件が高い場合の構成例（オンプレミス）

	ディスク障害	データ破損	インスタンス障害	データセンター障害	大規模障害
障害					
構成例					
概要	ディスクのミラーリングによる冗長化	レプリケーション等のデータ連携によるサーバー/データ冗長化	共有ディスクを使用したサーバー冗長化	DC間のデータ連携による冗長化	遠隔地にあるDC間のデータ連携による冗長化

RPO/RTO要件が高い場合の構成例（オンプレミス）

	ディスク障害	データ破損	インスタンス障害	データセンター障害	大規模障害
障害					
構成例				データセンター障害や 大規模障害も考慮した システムは少ない （バックアップデータを転送等）	
概要	ディスクのミラーリングによる冗長化	レプリケーション等のデータ連携によるサーバー/データ冗長化	共有ディスクを使用したサーバー冗長化	DC間のデータ連携による冗長化	遠隔地にあるDC間のデータ連携による冗長化

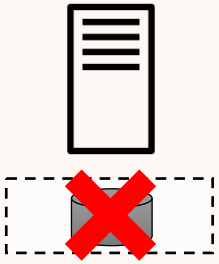
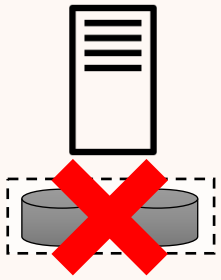
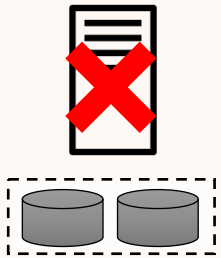
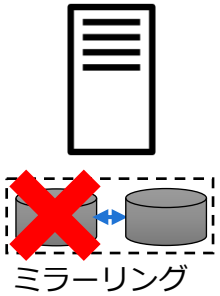
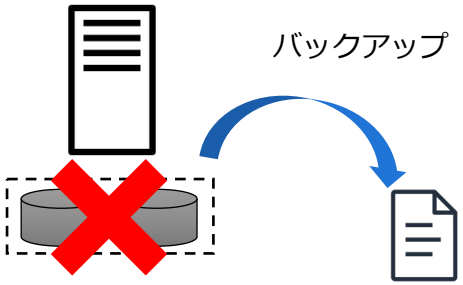
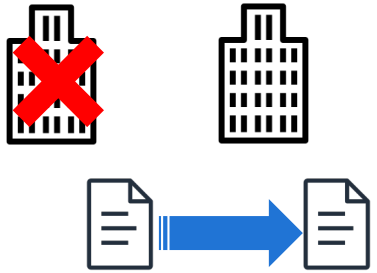
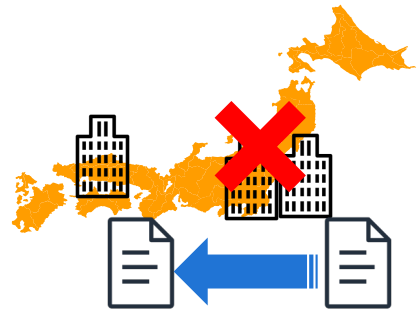
要件が高い構成例のRPO/RTO



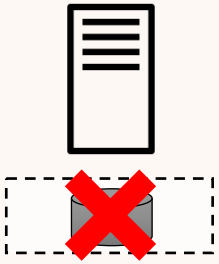
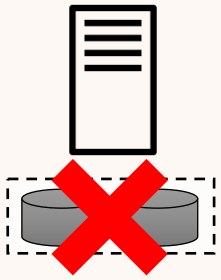
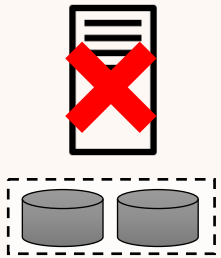
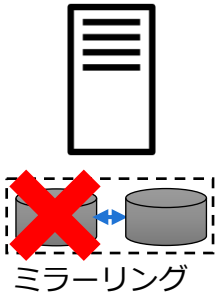
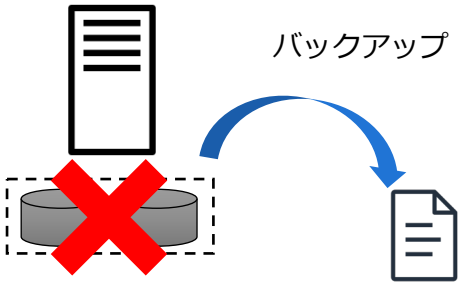
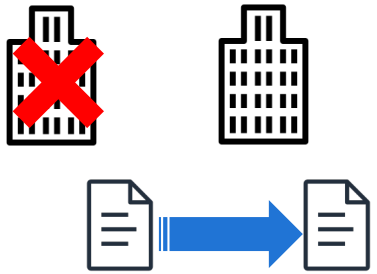
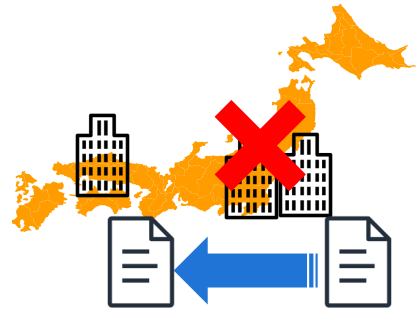
※一般的な傾向であり、約束されたものではありません

Auroraでの構成例

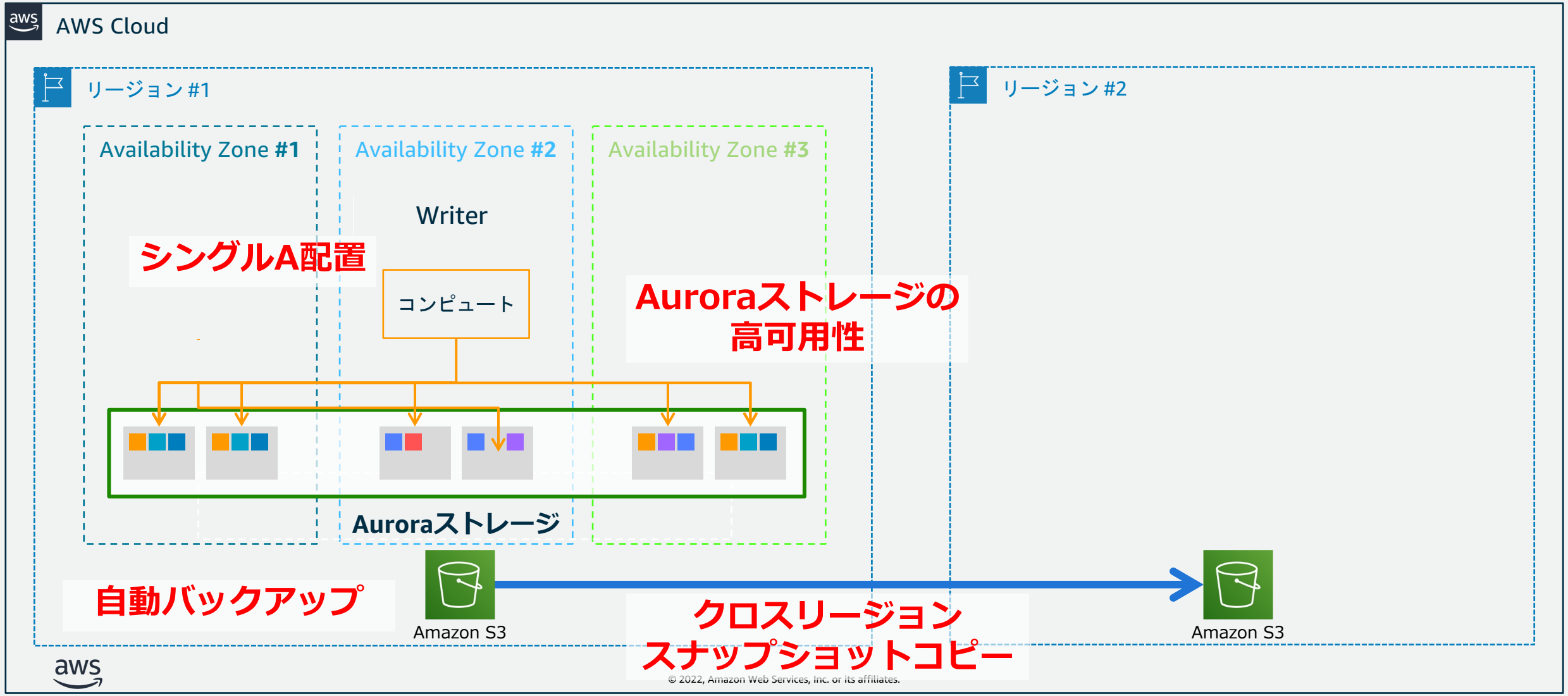
RPO/RTO要件が低い場合の構成例（オンプレミス）

	ディスク障害	データ破損	インスタンス障害	データセンター障害	大規模障害
障害					
構成例					

RPO/RTO要件が低い場合の構成例（Aurora）

	ディスク障害	データ破損	インスタンス障害	データセンター障害	大規模障害
障害					
構成例					
Aurora 構成	**Aurora ストレージの 高可用性**	**Auroraストレージ の高可用性/ 自動バックアップ**	**シングルAZ配置**		**クロスリージョン スナップショット コピー**

RPO/RTO要件が低い場合のAurora構成例



ディスク障害	データ破損	インスタンス障害	データセンター障害	大規模障害
Auroraストレージの高可用性	Auroraストレージの高可用性 自動バックアップ	シングルAZ配置		クロスリージョンスナップショットコピー

Aurora ストレージの高可用性

データベース用に設計された
専用の **log-structured 分散ストレージシステム**

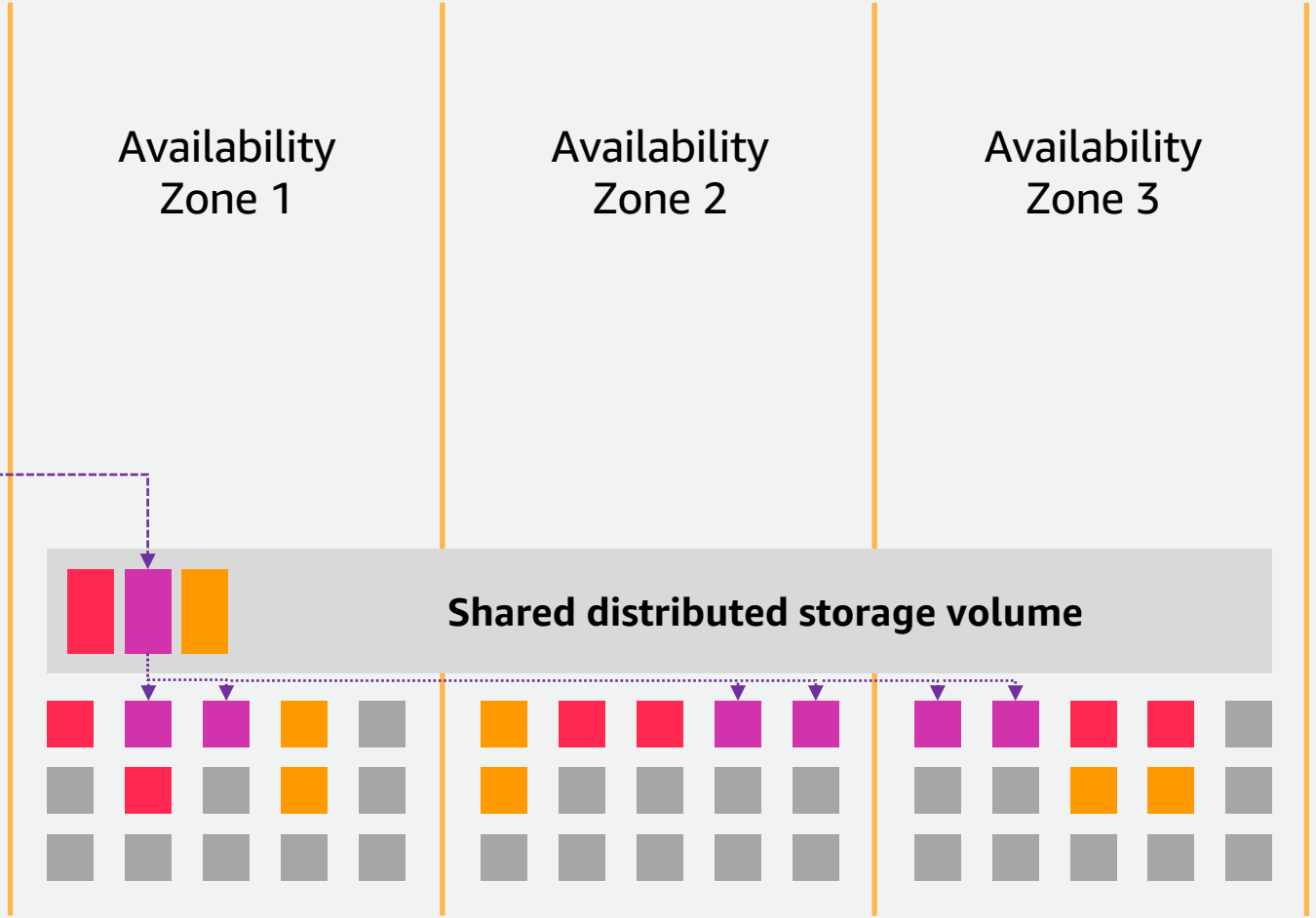
3つの異なるアベイラビリティゾーンに分散された
数百のストレージノードにストライピングされた
ストレージボリューム

データは 10GB の プロテクショングループ の
単位で管理される。自動的に最大 128TB まで拡張

AZ+1 の障害から保護するためにデータを
各アベイラビリティーゾーンに2つのコピー、
リージョン内で計6つのコピー

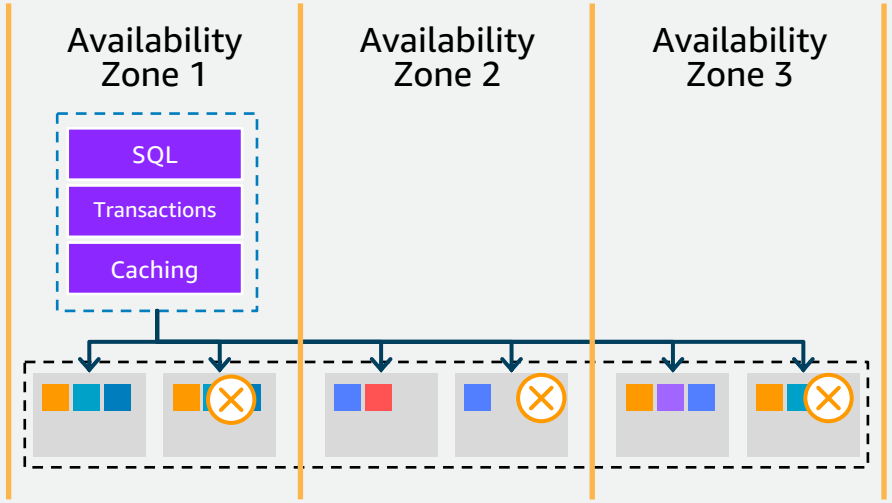
データは継続的に S3 へバックアップ

詳細：
https://docs.aws.amazon.com/ja_jp/AmazonRDS/latest/AuroraUserGuide/Aurora.Overview.StorageReliability.html

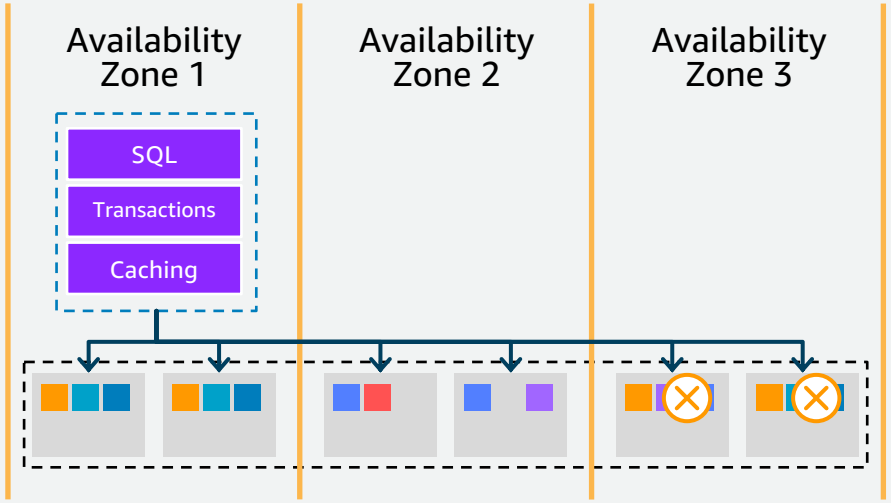


ディスク障害	データ破損	インスタンス障害	データセンター障害	大規模障害
Auroraストレージの高可用性	Auroraストレージの高可用性 自動バックアップ	シングルAZ配置		クロスリージョンスナップショットコピー

自動修復、耐障害性



Read availability



Write availability

データは6つのノードすべてに非同期・並列で書き込み

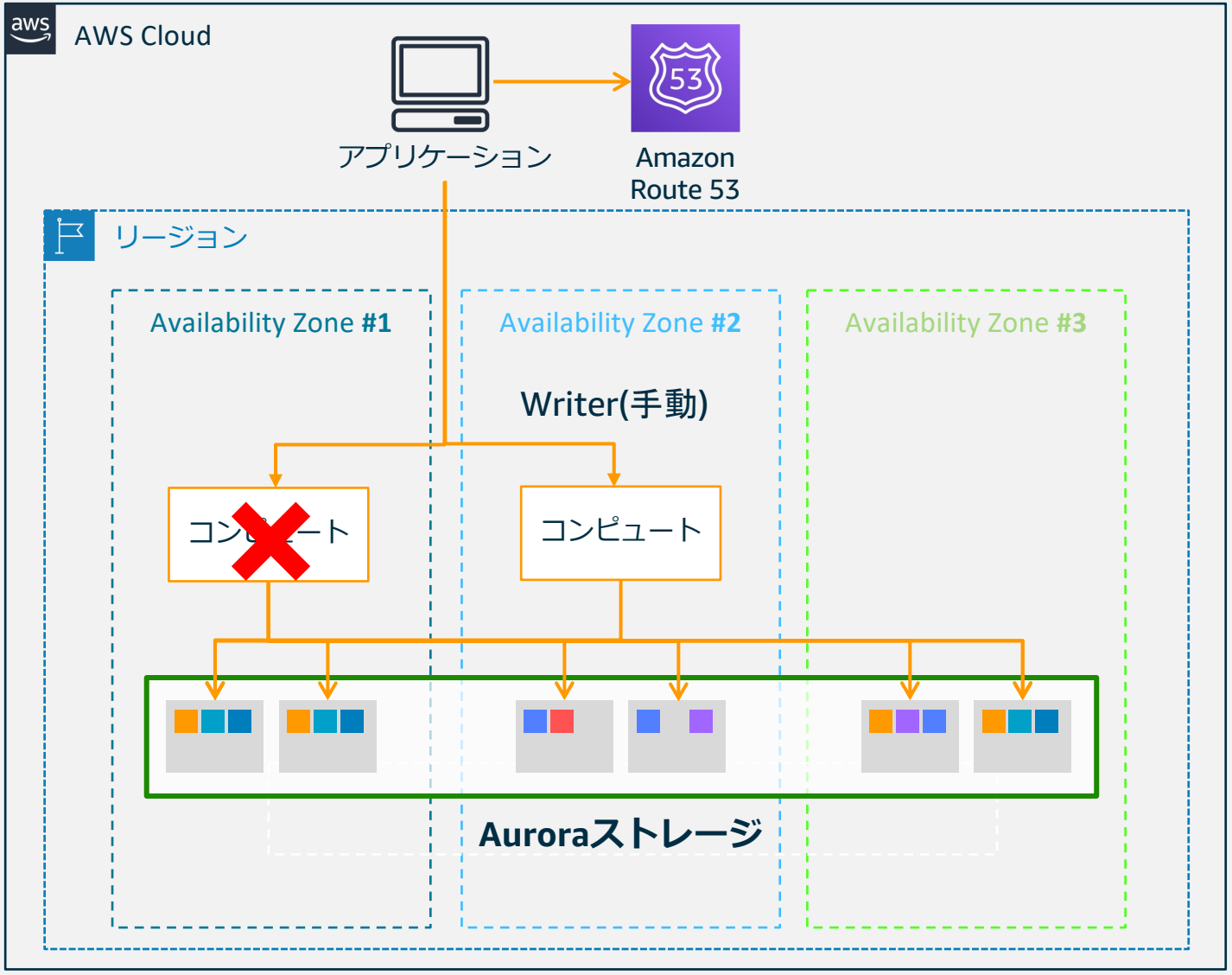
書き込みには 4/6 ノードのクォーラム、読み込みには 3/6 クォーラムが必要

各ノード間のデータの欠損、破損は P2P のゴシッププロトコルで確認、修復される

詳細： https://aws.amazon.com/jp/blogs/news/amazon-aurora-under-the-hood-quorum-and-correlated-failure/

ディスク障害	データ破損	インスタンス障害	データセンター障害	大規模障害
Auroraストレージの高可用性	Auroraストレージの高可用性 自動バックアップ	シングルAZ配置		クロスリージョンスナップショットコピー

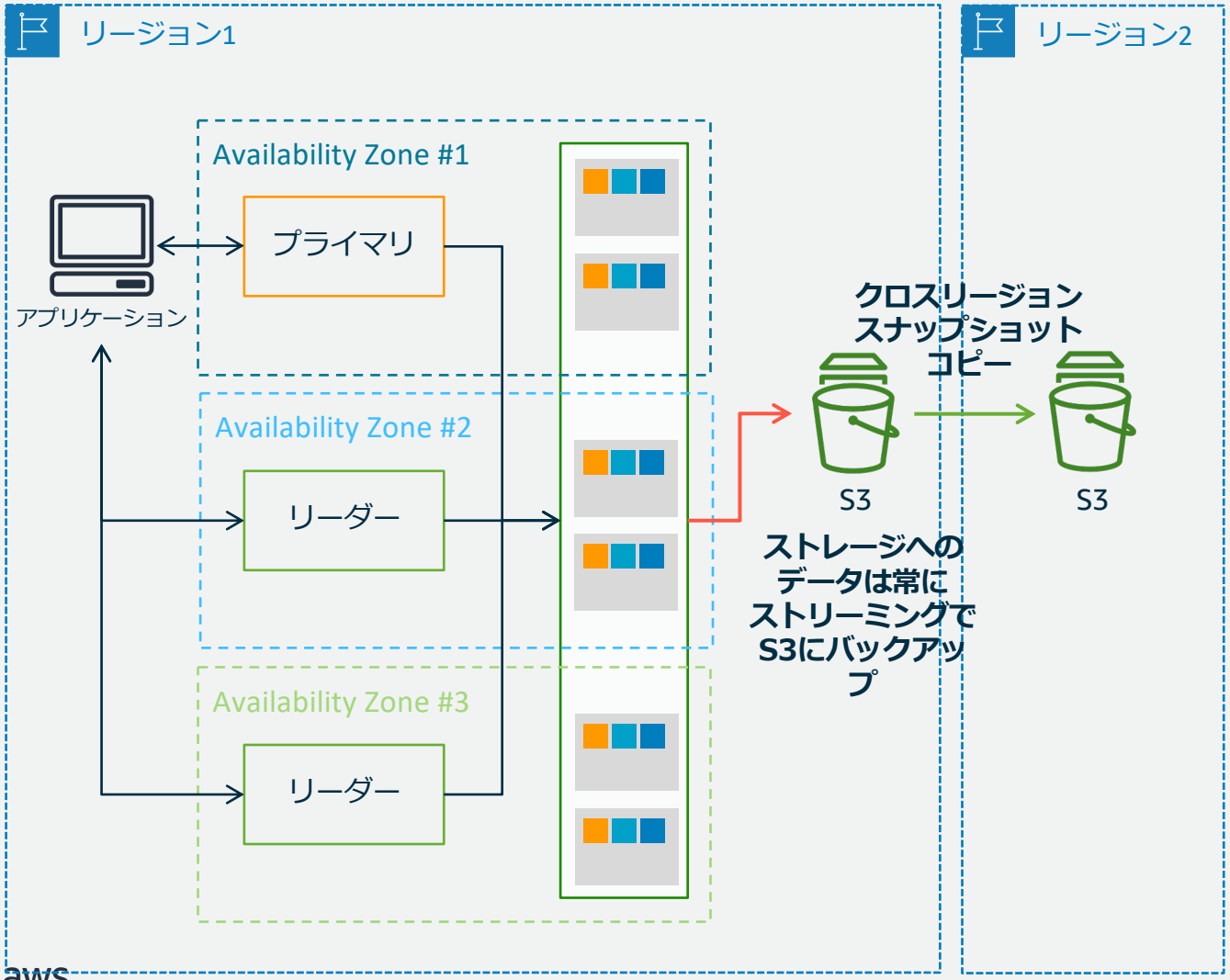
シングルAZ配置(Amazon Aurora)



- シングルAZ配置は、１つのAZにインスタンスが作成されている状態
- Auroraストレージは３つのAZに６つのコピー
- インスタンスに障害が発生した時やAZ障害が発生した時、同一AZにインスタンス作成を試みる(10分未満)
- AZ障害が発生した場合は手動で起動する必要がある

ディスク障害	データ破損	インスタンス障害	データセンター障害	大規模障害
Auroraストレージの高可用性	Auroraストレージの高可用性 自動バックアップ	シングルAZ配置		クロスリージョンスナップショットコピー

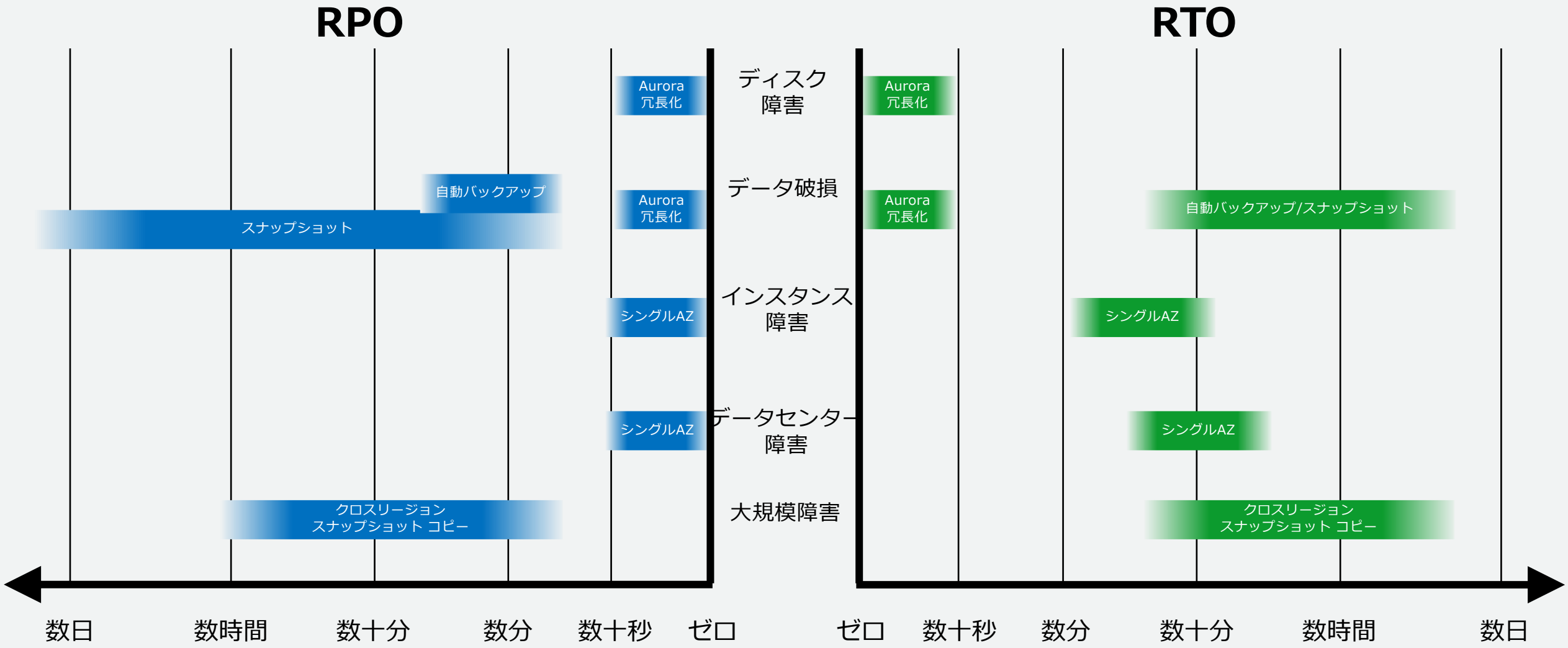
自動バックアップ&クロスリージョンスナップショットコピー



- 2つの選択肢–自動バックアップ(必須)、手動スナップショット(任意)
- Auroraは継続的にバックアップをS3に保存
- Auroraのバックアップは差分で取得されるためバックアップ取得期間の任意の時点にリカバリすることが可能
- Auroraのバックトラック機能を使うことでさらに高速に任意の時点にデータベースを戻すことも可能（MySQLのみ）
- バックアップはストリーミングで実行されバックアップ処理がデータベースのパフォーマンスに影響を与える事はない
- スナップショットを別リージョンにコピーしたり別アカウントと共有ができる

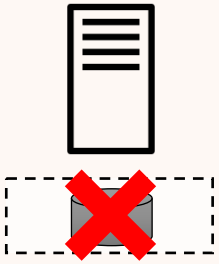
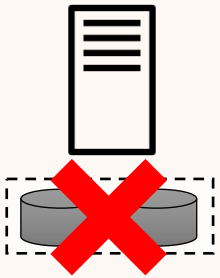
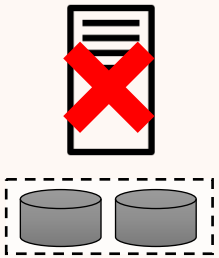
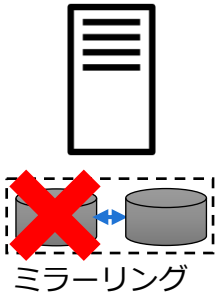
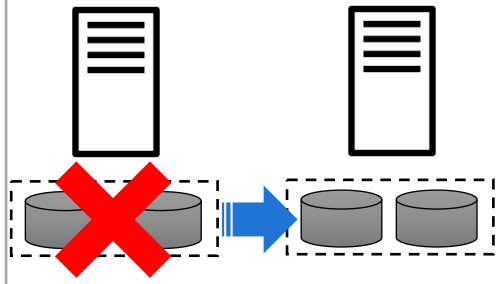
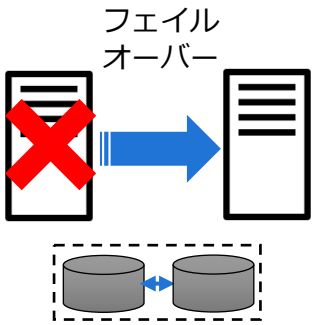
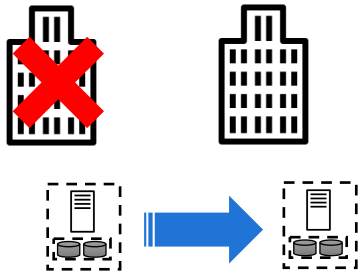
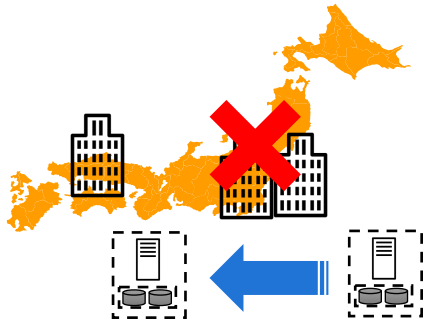
* SLAを満たせるかデータベースのRTO/RPOを検証で確認することが重要

要件が低い構成例のRPO/RTO（Aurora）

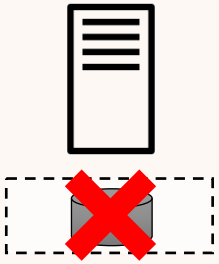
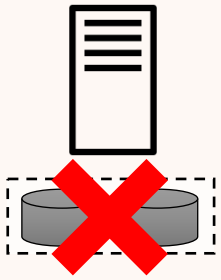
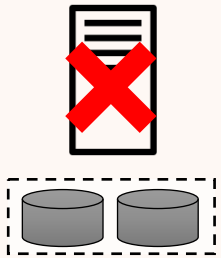
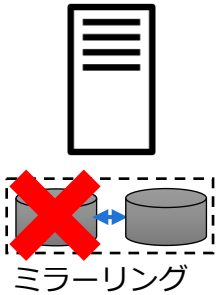
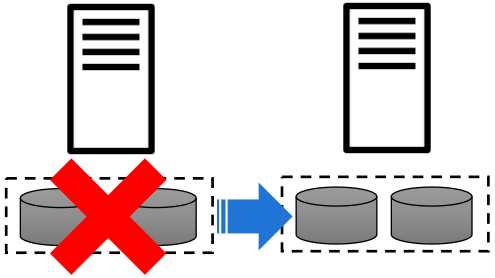
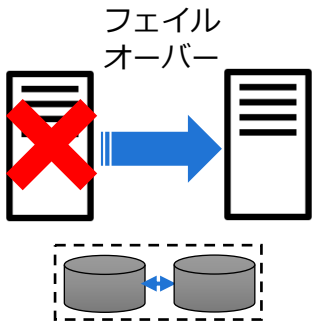
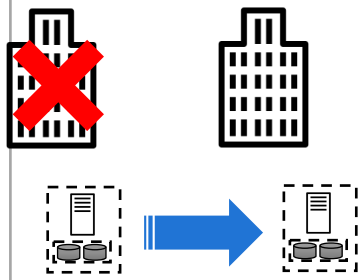
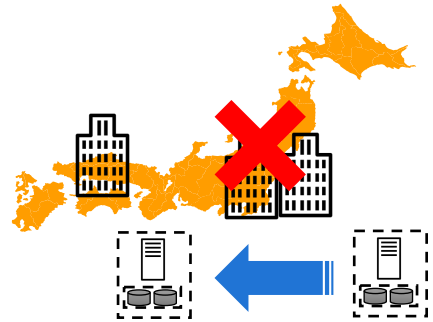


※一般的な傾向であり、約束されたものではありません

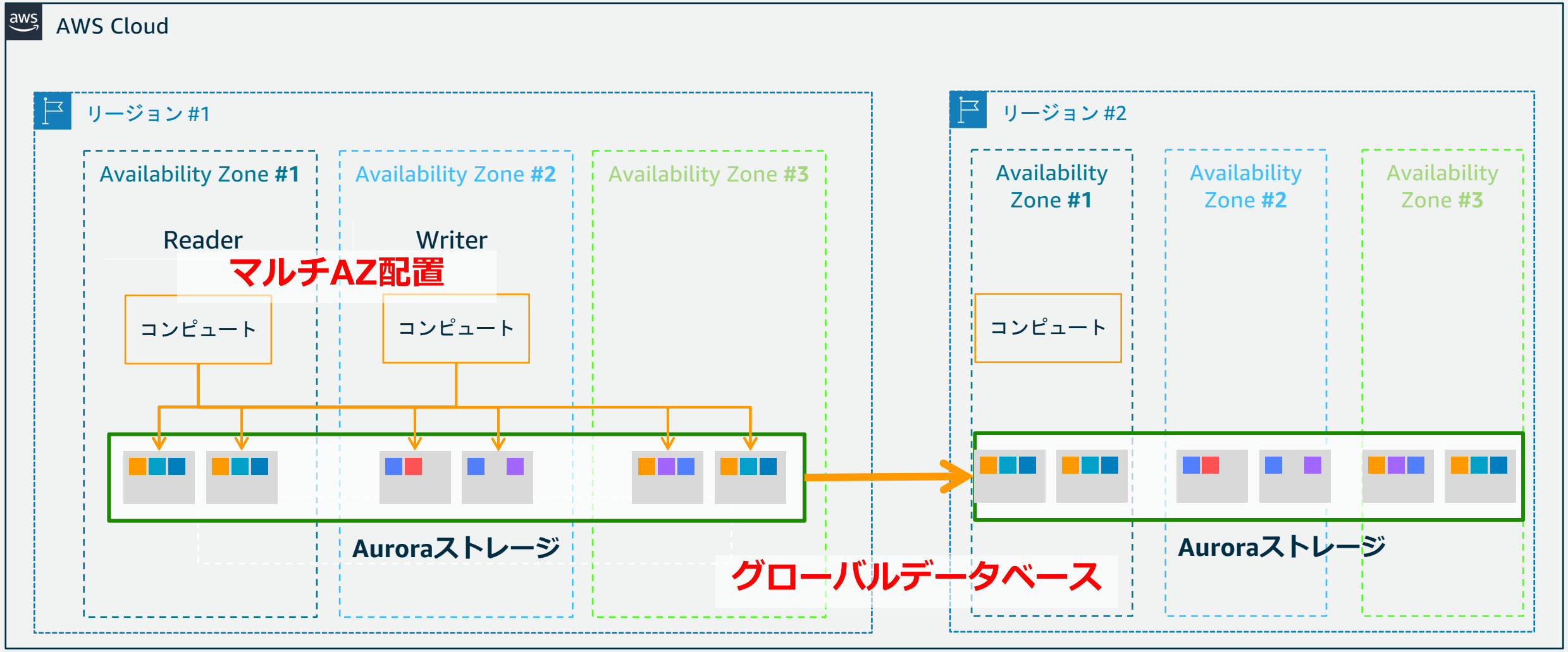
RPO/RTO要件が高い場合の構成例（オンプレミス）

	ディスク障害	データ破損	インスタンス障害	データセンター障害	大規模障害
障害					
構成例					

RPO/RTO要件が高い場合の構成例（Aurora）

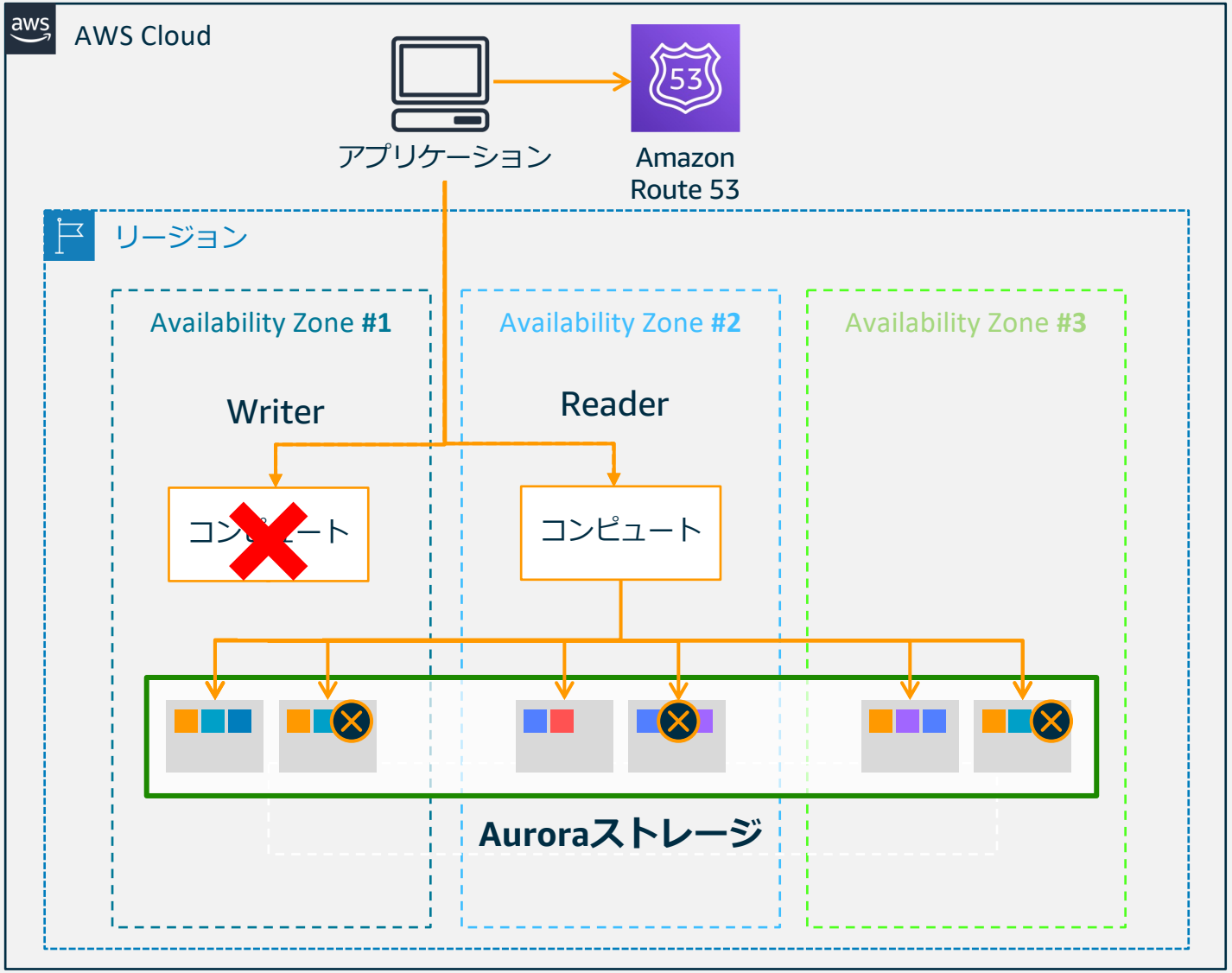
	ディスク障害	データ破損	インスタンス障害	データセンター障害	大規模障害
障害					
構成例					
Aurora 構成	**Aurora ストレージの 高可用性**	**Auroraストレージの 高可用性/ 自動バックアップ/ グローバル データベース**	**マルチAZ配置**		**グローバル データベース**

RPO/RTO要件が高い場合のAurora構成例



ディスク障害	データ破損	インスタンス障害	データセンター障害	大規模障害
Auroraストレージの高可用性	Auroraストレージの高可用性/自動バックアップ/グローバル データベース	マルチ AZ配置		グローバルデータベース

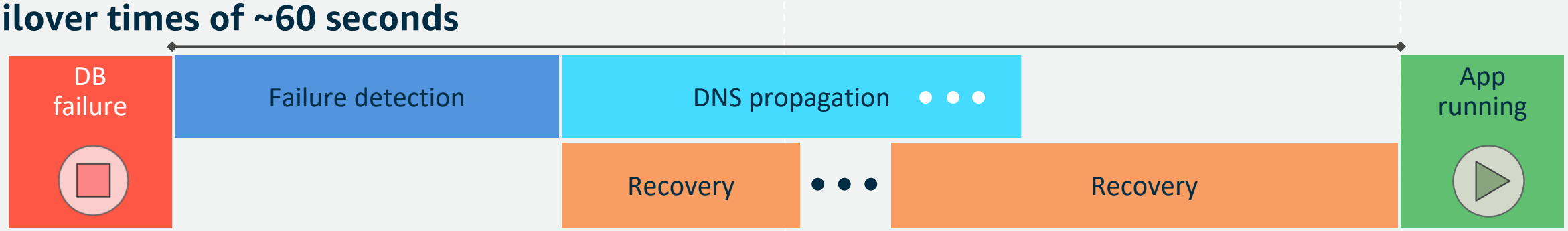
マルチAZ配置(Amazon Aurora)



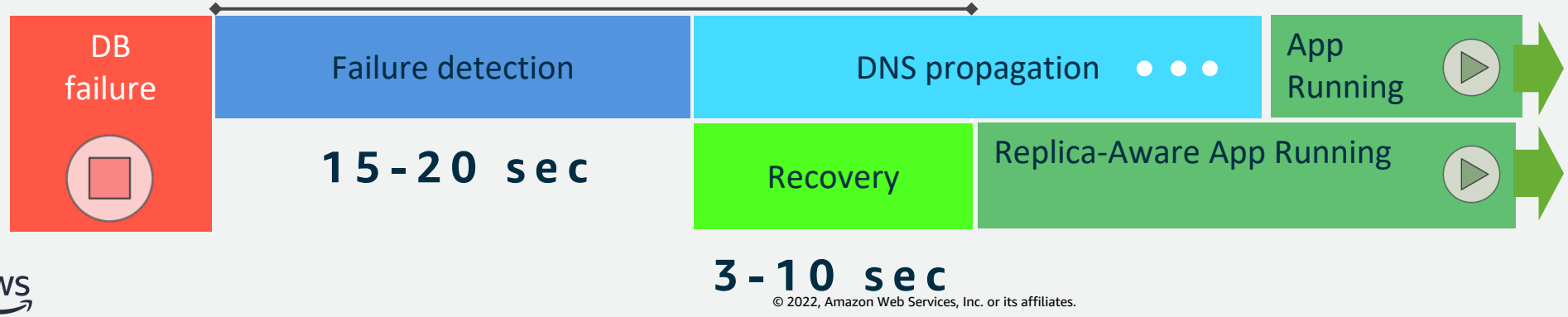
- マルチAZ構成は、データベースにエンタープライズグレードの高可用性ソリューションを提供
- 各インスタンスはAuroraストレージを共有
- インスタンスは外部のオブザーバーによって監視され、クォーラムに対する合意を維持
- フェイルオーバーは自動または Amazon RDS API経由で実行
- エンドポイントのDNSレコードは Route53により実施
- データベースの問題に限らずAWSの基盤の問題も検知
- 1クリックで有効化

高速でより予測可能なフェイルオーバー時間

Amazon RDS for PostgreSQL is good: failover times of ~60 seconds



Amazon Aurora is better: failover times < 30 seconds



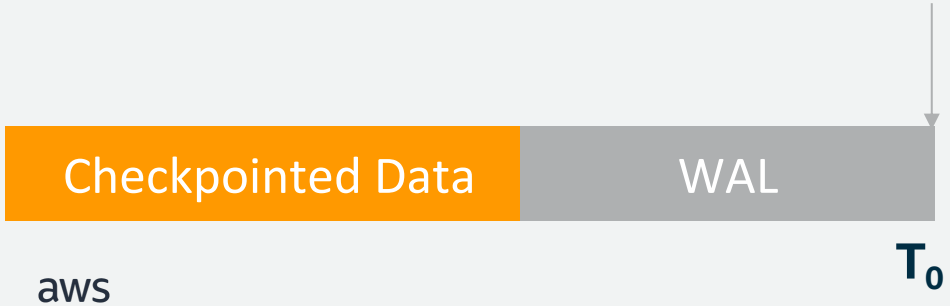
高速なリカバリー

既存のデータベース

最後のチェックポイントからログを適用していく

PostgreSQL ではシングルスレッドなため適用完了までの時間が増加

T0 でクラッシュが発生すると最後のチェックポイントからのログを適用する必要がある

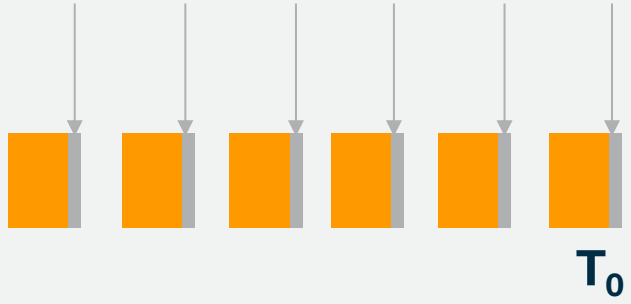


Amazon Aurora

リカバリ中か否か関係なく、ストレージノードでは継続的にログレコードを再生(No Checkpoint)

並列、分散、非同期で行われる

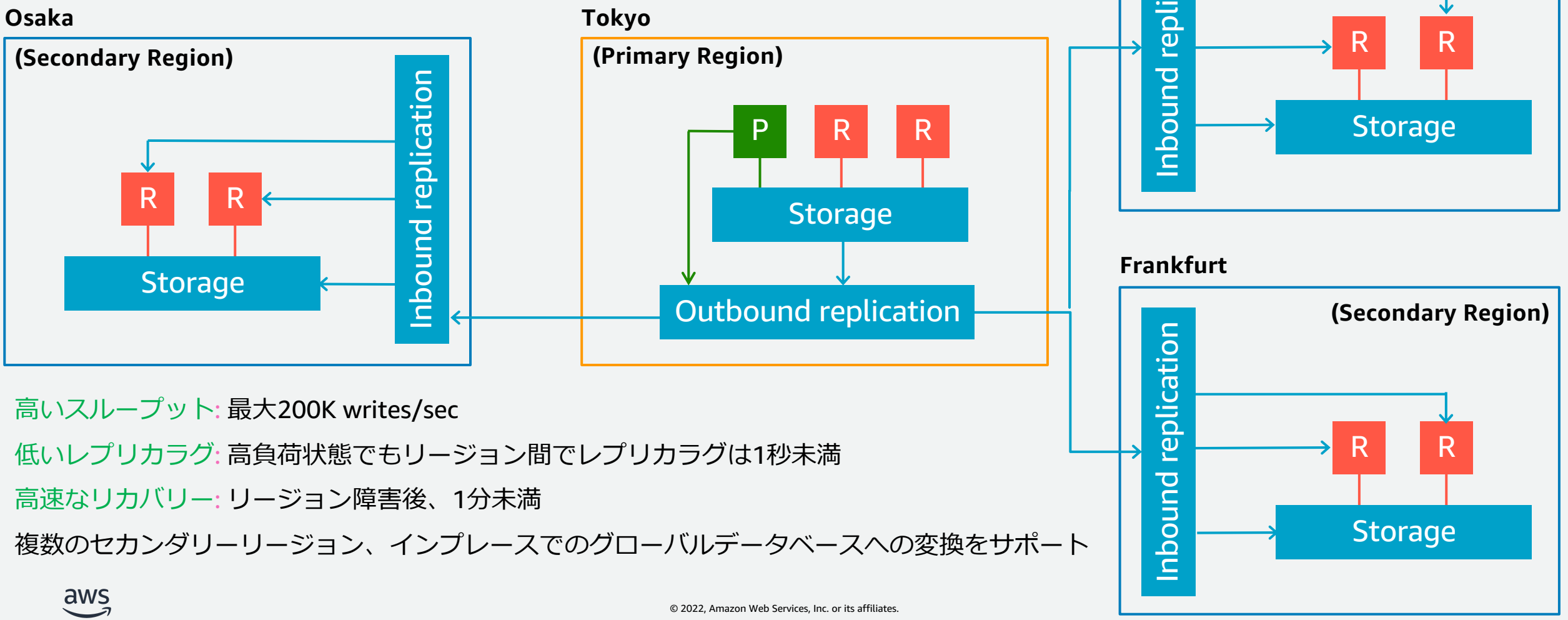
T0 でクラッシュが発生するとredo を並列で分散して非同期でログの適用を行う



ディスク障害	データ破損	インスタンス障害	データセンター障害	大規模障害
Auroraストレージの高可用性	Auroraストレージの高可用性/自動バックアップ/グローバル データベース	マルチ AZ配置		グローバルデータベース

Auroraグローバルデータベース

高速な災害対策と拡張されたデータローカリティー



高いスループット: 最大200K writes/sec

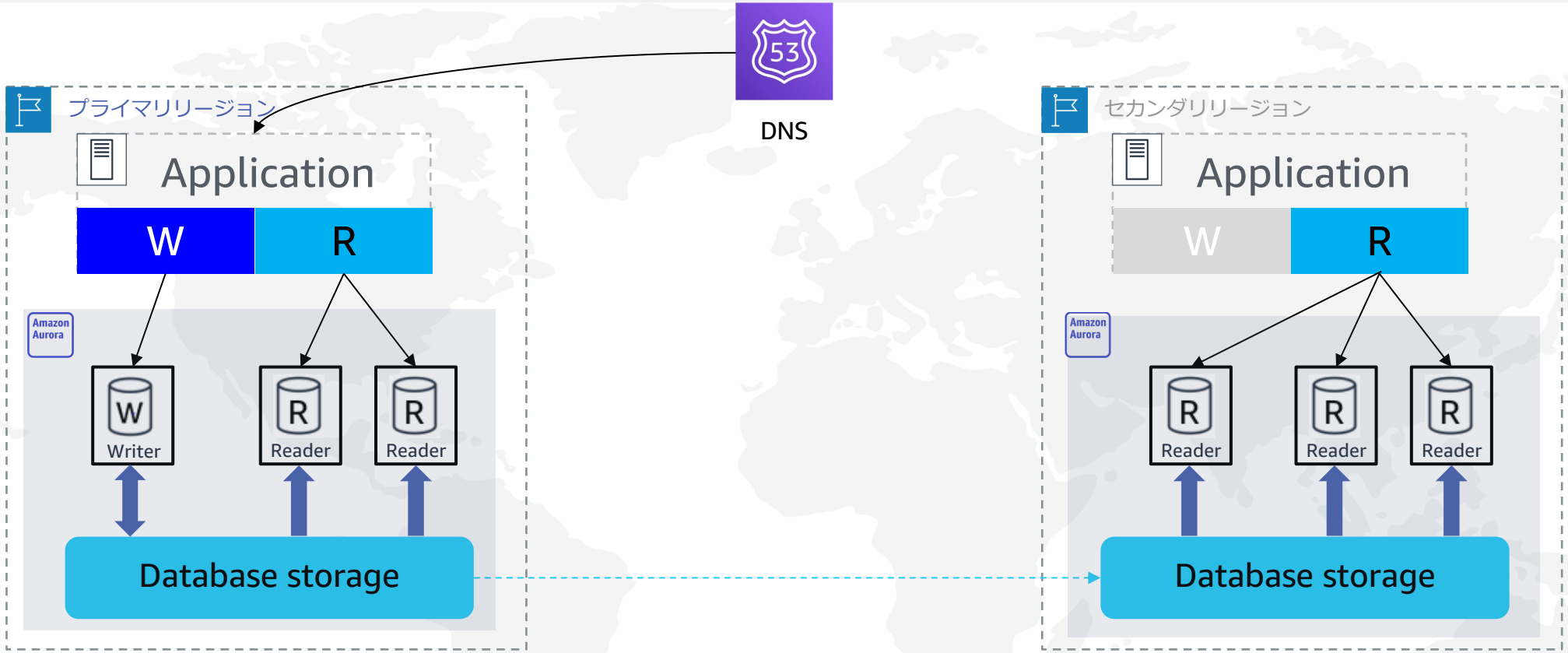
低いレプリカラグ: 高負荷状態でもリージョン間でレプリカラグは1秒未満

高速なリカバリー: リージョン障害後、1分未満

複数のセカンダリーリージョン、インプレースでのグローバルデータベースへの変換をサポート

ディスク障害	データ破損	インスタンス障害	データセンター障害	大規模障害
Aurora ストレージの高可用性	Auroraストレージの高可用性/自動バックアップ/グローバル データベース	マルチ AZ配置		グローバル データベース

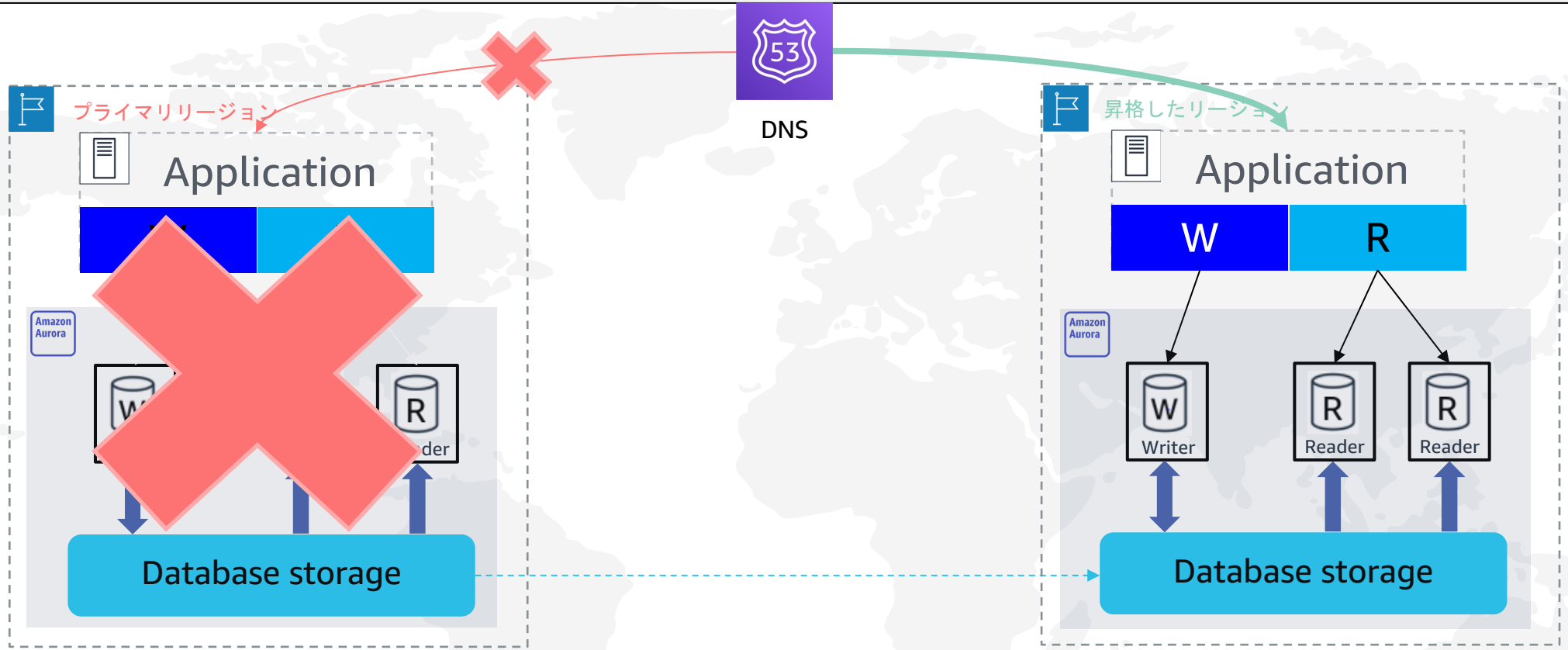
Auroraグローバルデータベース



アプリケーションは災害からの復旧と高速なローカルリードを
提供するためそれぞれのリージョンに設定されています。

ディスク障害	データ破損	インスタンス障害	データセンター障害	大規模障害
Auroraストレージの高可用性	Auroraストレージの高可用性/自動バックアップ/グローバル データベース	マルチ AZ配置		グローバルデータベース

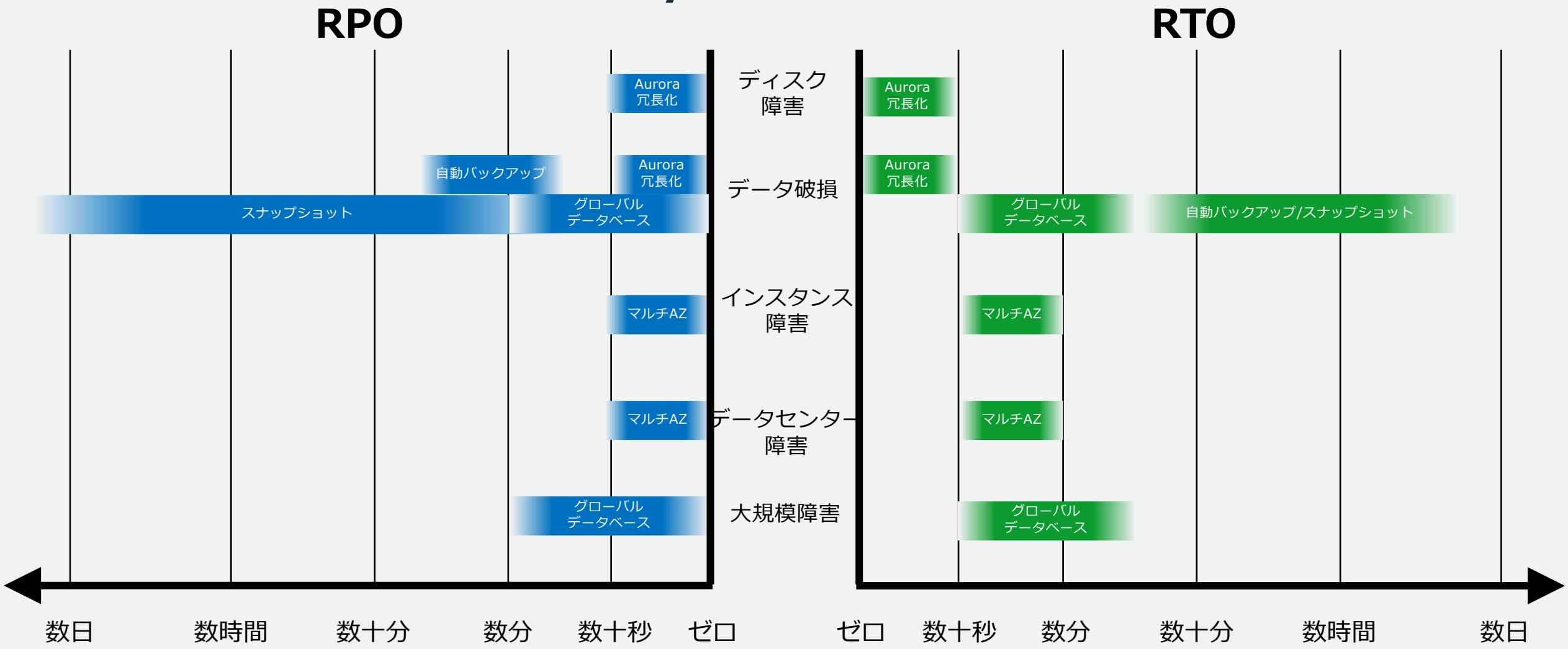
Auroraグローバルデータベース



セカンダリリージョンでアプリケーションを起動してトラフィックのルートを開始し（DNS経由）、

書き込み可能なWriterを1分以内に昇格することができます。

要件が高い構成例のRPO/RTO（Aurora）



※一般的な傾向であり、約束されたものではありません

まとめ

まとめ

- データベースを構成する際は、RTO/RPOや障害の種類、コストを考慮する必要があります。
- Auroraでは、高可用性要件に合わせて最適なアーキテクチャが構成可能です。
- オンプレミスでは想定していなかったデータセンター障害や大規模障害も考慮したアーキテクチャを検討することをお勧めします。

Thank you!