



AWS Clean Rooms

AWS Black Belt Online Seminar

Koshiro Suzuki

Amazon Web Service Japan G.K.
Solutions Architect

2023/09

AWS Black Belt Online Seminar とは

- 「サービス別」「ソリューション別」「業種別」などのテーマに分け、アマゾン ウェブ サービス ジャパン合同会社が提供するオンラインセミナーシリーズです
- AWS の技術担当者が、AWS の各サービスやソリューションについてテーマごとに動画を公開します
- 以下の URL より、過去のセミナー含めた資料などをダウンロードすることができます
- <https://aws.amazon.com/jp/aws-jp-introduction/aws-jp-webinar-service-cut/>
- <https://www.youtube.com/playlist?list=PLzWGOASvSx6FlwIC2X1nObr1KcMCBBlqY>



ご感想は Twitter へ！ハッシュタグは以下をご利用ください
#awsblackbelt

内容についての注意点

- 本資料では 2023 年 9 月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。AWS のサービスは常にアップデートを続けているため、最新の情報は AWS 公式ウェブサイト (<https://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
- 技術的な内容に関しましては、有料の [AWS サポート窓口](#)へお問い合わせください
- 料金面でのお問い合わせに関しましては、[カスタマーサポート窓口](#)へお問い合わせください (マネジメントコンソールへのログインが必要です)

自己紹介

名前

鈴木 康士郎 (Koshiro)

所属

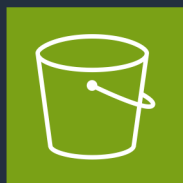
アマゾン ウェブ サービス ジャパン合同会社
技術統括本部 インターネットメディア ソリューション部
ソリューション アーキテクト

経歴

ソーシャルゲーム会社にて Tech Lead を経験した後、
モバイル向けゲーム/ライブ配信アプリの立ち上げに携わり
技術選定や開発、運用を担当



好きな AWS サービス



Amazon S3



AWS Clean Rooms



Amazon Aurora



Amazon EventBridge



本セミナーの対象者

- データクリーンルームの技術要素をキャッチアップしたい方
- 機密性の高いデータのグループ間、企業間共有に課題を抱えている方
- AWS を利用したデータコラボレーションを運用設計を検討されている方
- AWS を利用したデータ活用に興味がある方

アジェンダ

1. データ共有の価値と課題
2. AWS Clean Rooms の機能
3. AWS Clean Rooms の利用方法・パターン
4. AWS 他サービスとの連携
5. 料金・注意点
6. まとめ

アジェンダ

1. **データ共有の価値と課題**
2. AWS Clean Rooms の機能
3. AWS Clean Rooms の利用方法・パターン
4. AWS 他サービスとの連携
5. 料金・注意点
6. まとめ

課題

**AWS のお客様は
顧客パーソナライズの為に
幅広いパートナーとの
データコラボレーションを
検討しているが
さまざまな課題に直面**



増え続ける顧客動向データ

カスタマーの行動をより理解する為に、増え続けるタッチポイントやチャネルの膨大な行動データが必要

今後5年間で作成される行動データの量は、過去に作成されたデジタルデータの2倍以上の量になると言われています



データ断片化とデータサイロ

行動データは業界全体に断片化しており、より多くの顧客インサイトを得るために各企業がデータを組み合わせる試みをしています。

しかし、組み合わせる為の開発や、企業間でインタフェースの標準化の必要など、運用性と拡張性の課題があります



顧客データ保護

各企業はカスタマーのデータプライバシーを保護する為に、データの収集・保存・使用をより適切にセキュリティを担保して管理する必要があります

Source: IDC, Worldwide Global DataSphere Forecast, 2021–2025, Doc #US46410421, March 2021

背景

お客様は、データセキュリティを強化して
Rawデータを適切に保護しながらパートナーとコラボレーションする方法を模索



より良い意思決定を
行う為にパートナー
全体のデータを活用
する必要がある



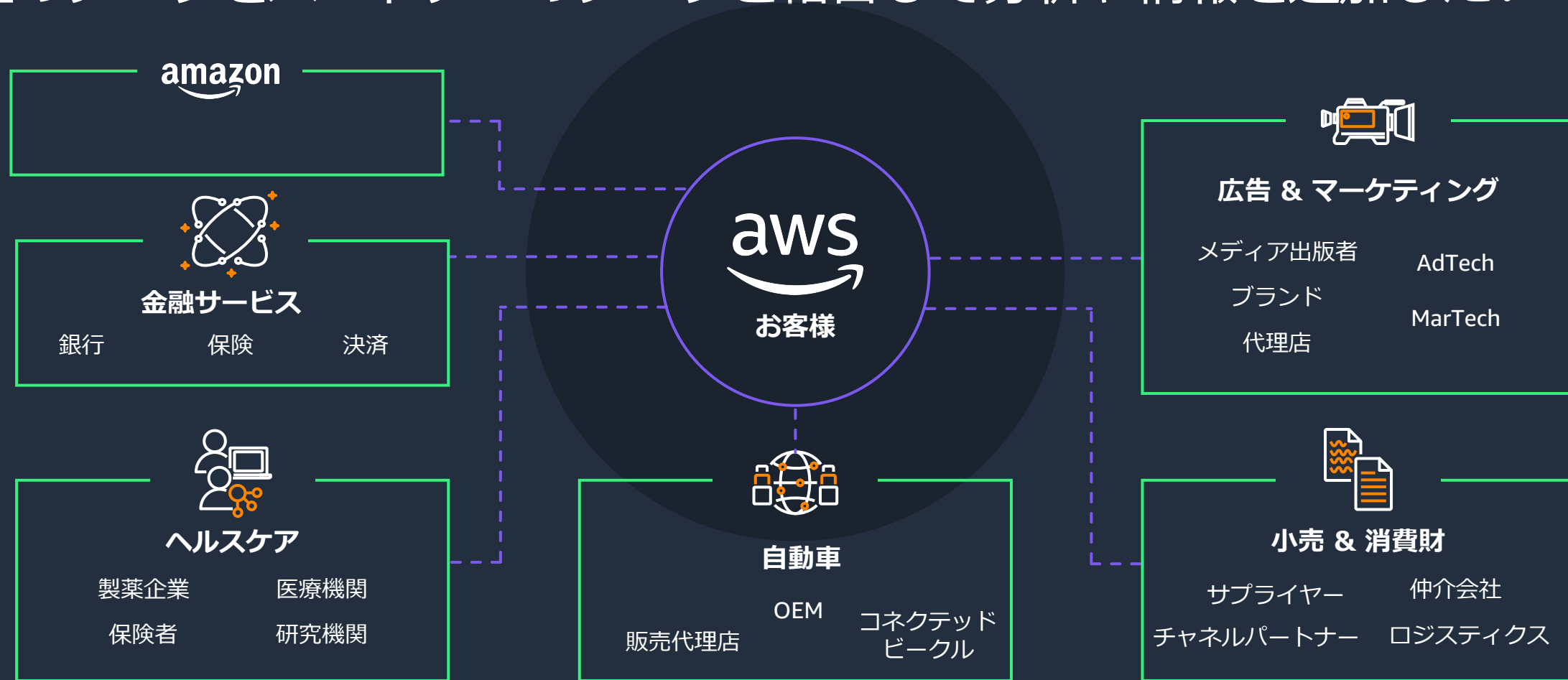
Brands, agencies, media
publishers, AdTech,
MarTech

プライバシーと
知的財産を保護する
必要がある



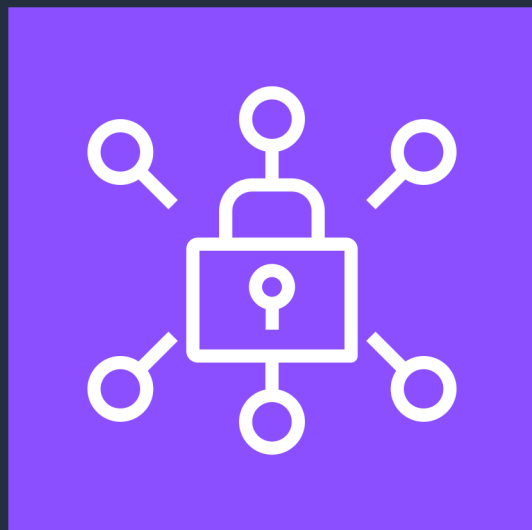
背景

データをそのまま直接コピーやデータベース自体の共有を行わずに、自社のデータをパートナーのデータと結合して分析や情報を追加したいニーズ



AWS Clean Rooms の概要

数回クリックするだけで、独自の Clean Rooms を作成
簡単にデータコラボレーションを実現



Raw データを共有・直接使用することなく、
AWS 上の何十万もの企業とコラボレーション

データセットの照合、分析、コラボレーションを安全に実施し、
インサイトの導出を実現

クリーンルームの為の幅広いアクセス制御で、基礎となるデータを
保護

柔軟性に富んだ分析ルールを使用して、ビジネスニーズに応じた
多様なクエリを実現

AWS Clean Rooms は、企業とそのビジネスパートナーが、Raw データの共有/公開なく、データセットを安全に分析し、コラボレーション可能



マルチパーティのコラボレーション

1回のコラボレーションで最大5社の参加者との共同作業が可能、複数の企業からインサイトを抽出



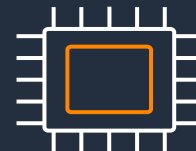
AWS上でのデータ移動は不要

直接アクセス許可を付与して Amazon S3 データを活用するため、AWS 上でのデータの移動は不要



クエリ制御とその実装

分析ルールを簡単に設定し、許可するデータ分析の種類を制限



暗号コンピューティング

データを事前に暗号化し、クエリ実行中も含め常に暗号化した状態を維持



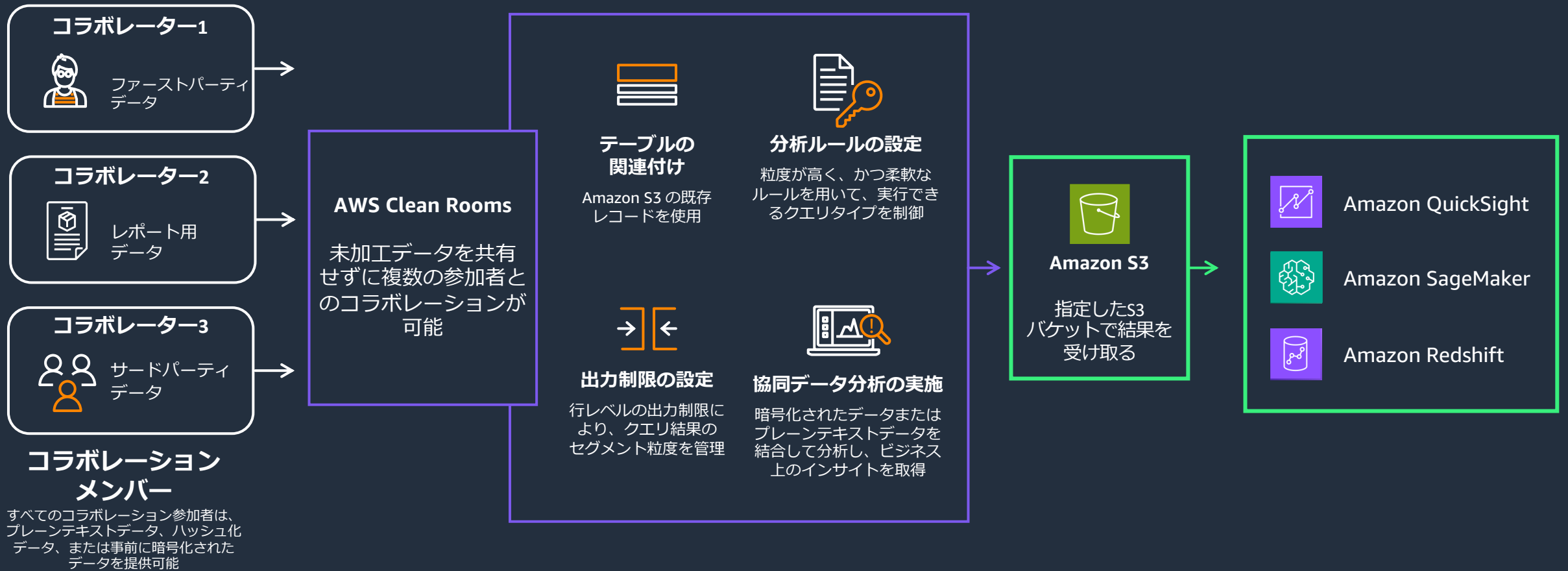
プログラム化されたアクセス

機能を自動化して既存のワークフローや製品に統合、あるいは、ホワイトラベルのクリーンルームオフリングを作成

アジェンダ

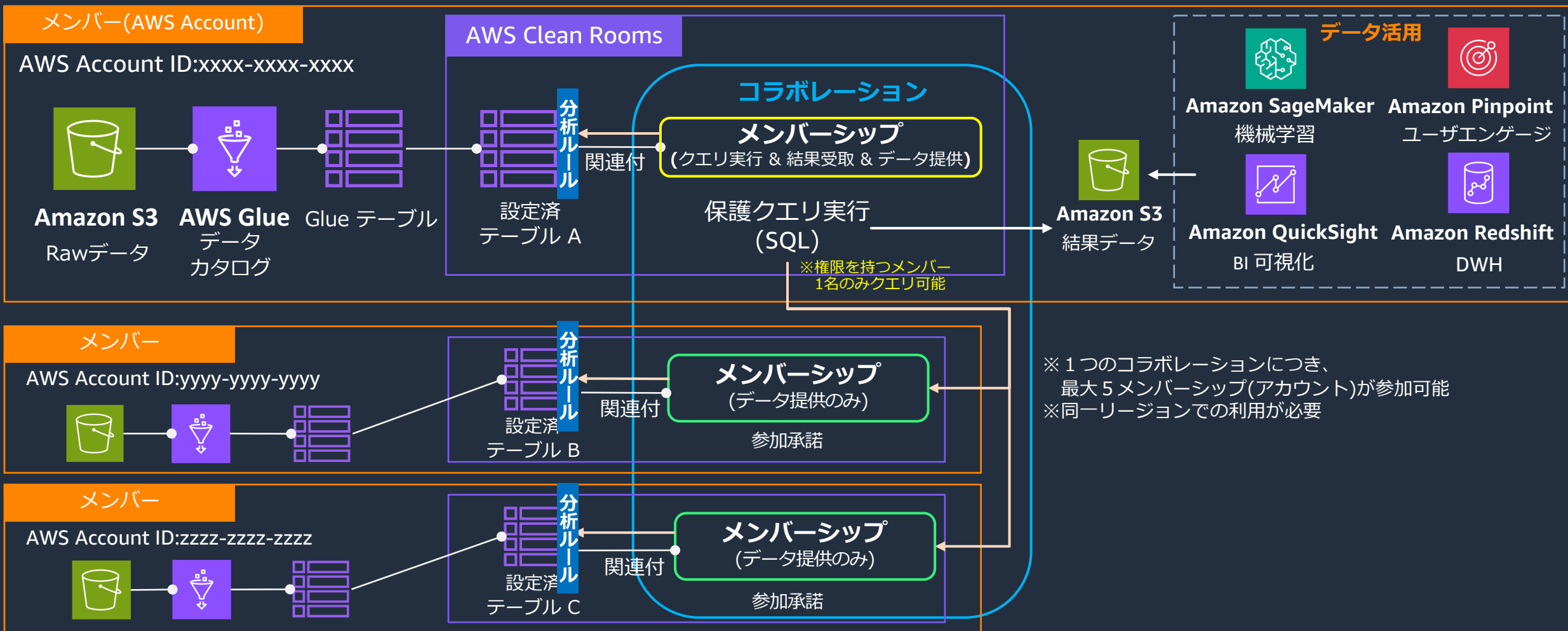
1. データ共有の価値と課題
- 2. AWS Clean Rooms の機能**
3. AWS Clean Rooms の利用方法・パターン
4. AWS 他サービスとの連携
5. 料金・注意点
6. まとめ

AWS Clean Rooms の仕組み



AWS Clean Rooms の詳細説明

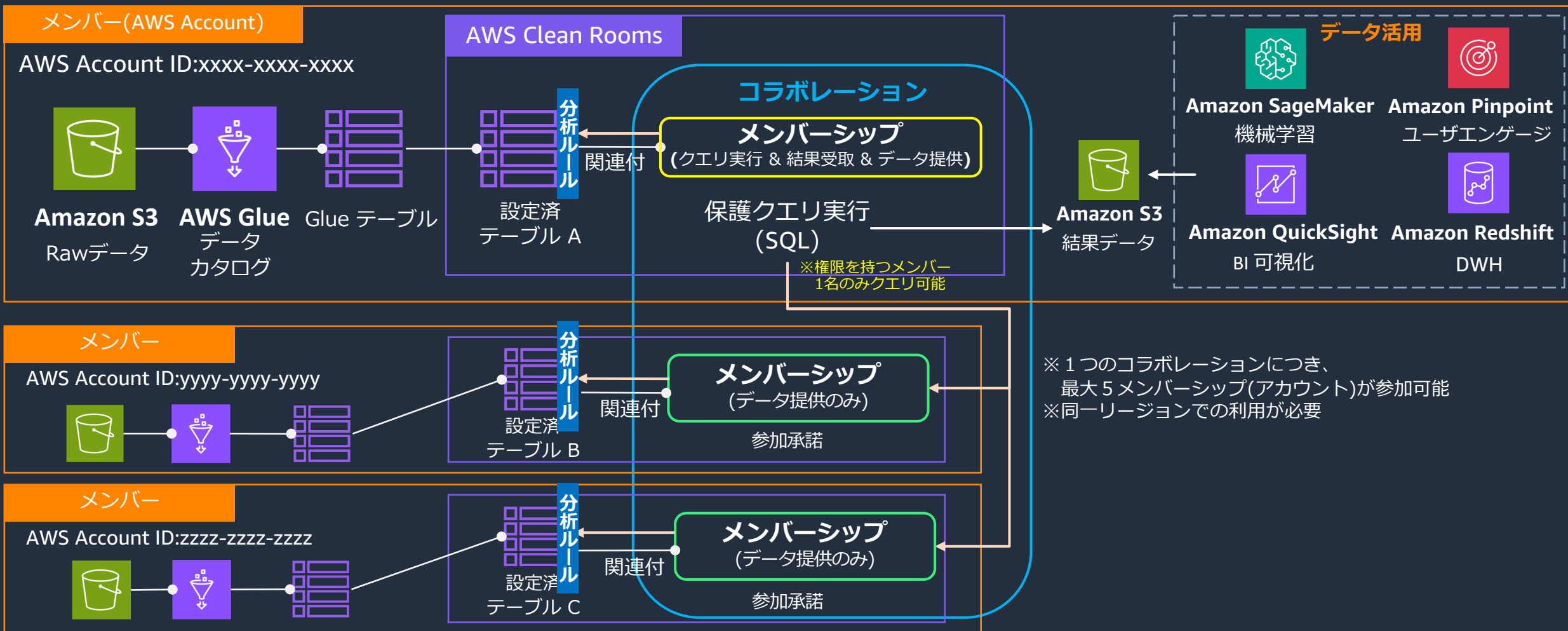
自社データのアクセス範囲・クエリ条件を管理して安全に他社に提供



※ 1つのコラボレーションにつき、
最大5メンバーシップ(アカウント)が参加可能
※ 同一リージョンでの利用が必要

AWS Clean Rooms の詳細説明

自社データのアクセス範囲・クエリ条件を管理して安全に他社に提供

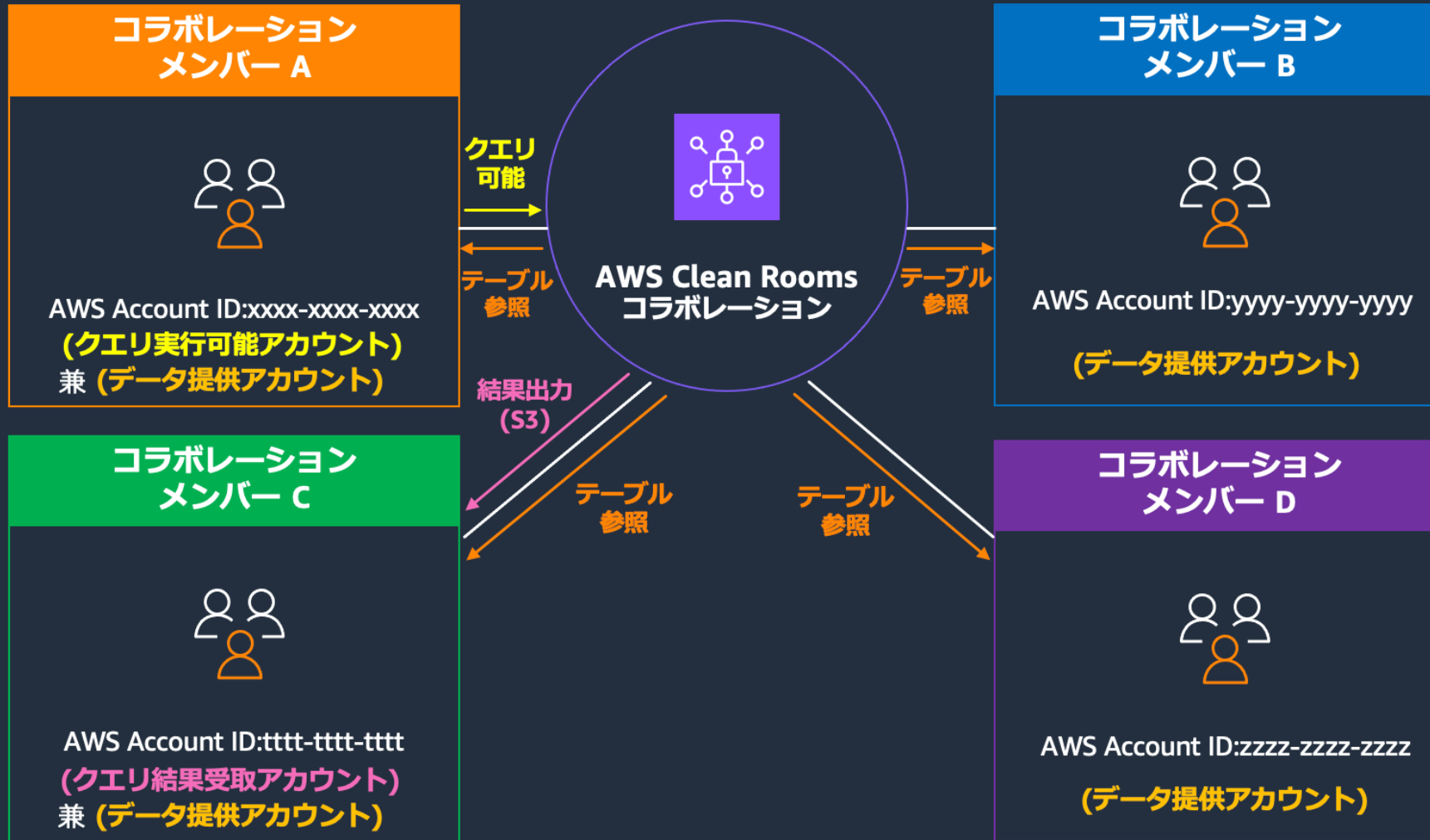


※ 1つのコラボレーションにつき、
最大5メンバーシップ(アカウント)が参加可能
※ 同一リージョンでの利用が必要

AWS Clean Rooms : コラボレーション

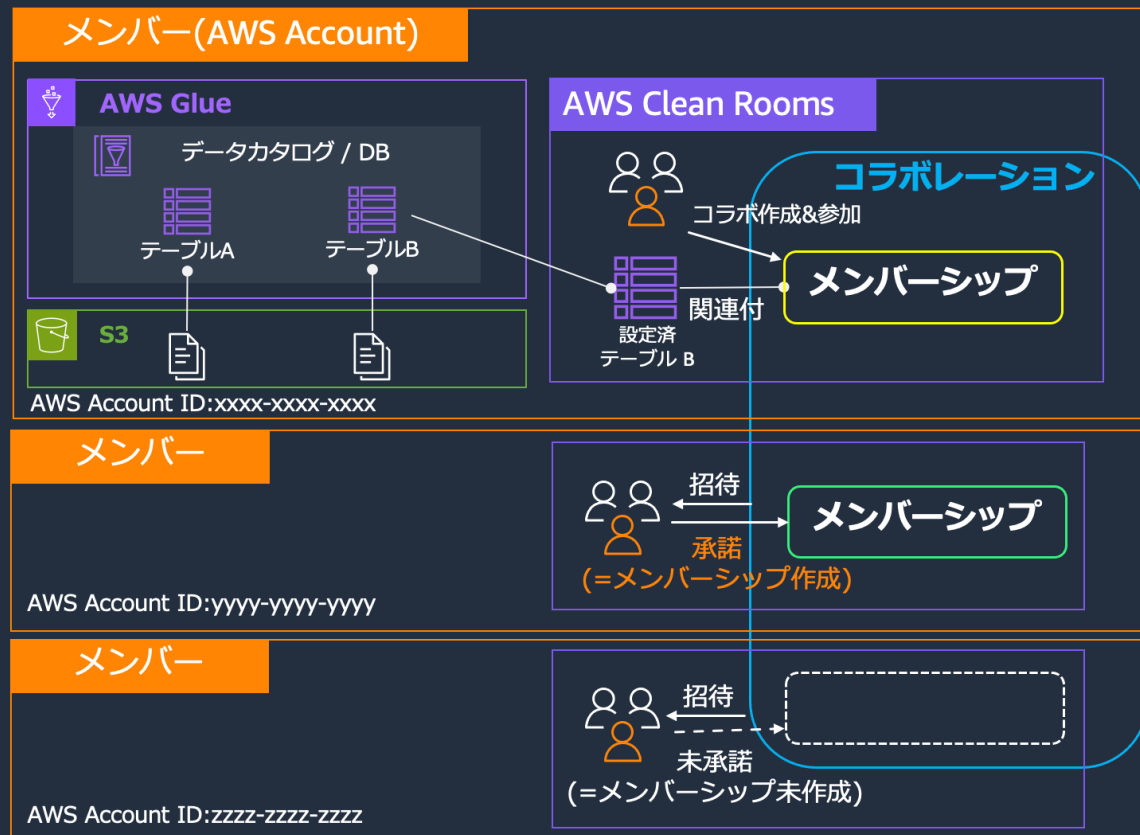
- データコラボレーションにおける安全な論理境界
- コラボレーション作成時にメンバーの招待と役割、ログ出力の有無を指定
- クエリ実行可能なメンバーが 1 アカウント、クエリ結果受け取りメンバーが 1 アカウント、他はデータ提供者で構成
- コラボレーションあたり最大 5 アカウントが参加可能（上限緩和申請は可能）

AWS Clean Rooms : コラボレーション



AWS Clean Rooms : コラボレーションメンバー

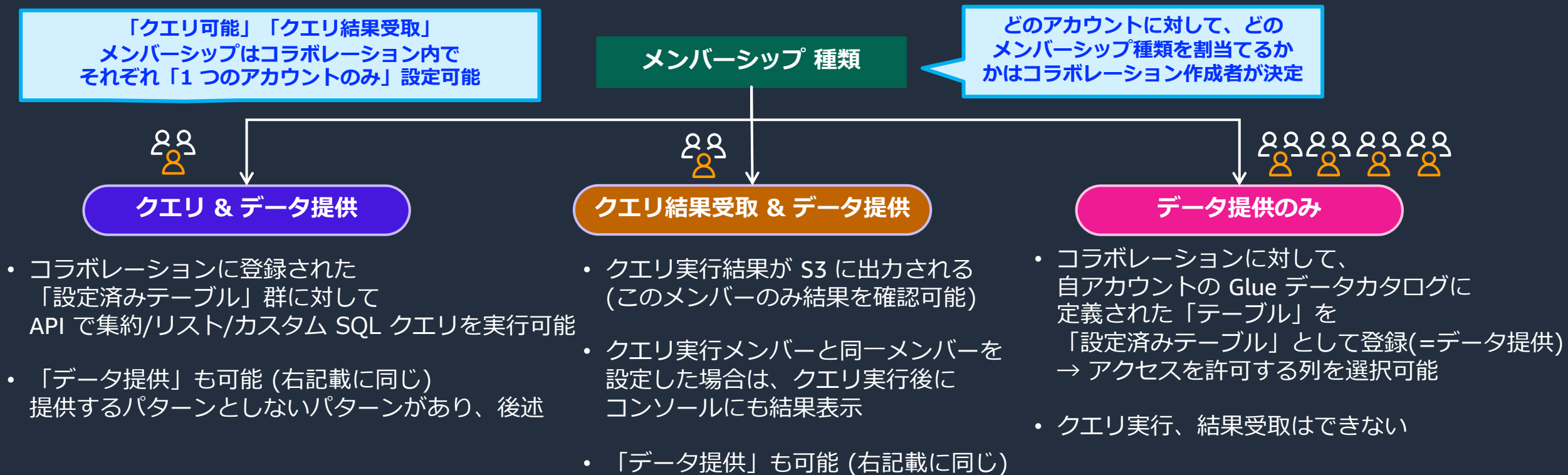
- 招待されたコラボレーションにメンバーシップを作成することで参加
- 各メンバーのリソースはメンバーシップに紐づき、メンバーの任意タイミングで削除・追加が可能



AWS Clean Rooms : メンバーシップ種別

コラボレーション内のメンバーシップは3種類

「クエリ実行/クエリ結果受取」可能な種別をメンバーに割り当てる



※クエリ可能なメンバーシップ、クエリ結果受取メンバーシップには、コラボレーション作成アカウント以外でも割り当て可能

AWS Clean Rooms : メンバーシップ (招待受領側)

- 参加可能タブより確認。詳細よりメンバーシップの種類も確認可能。

AWS Clean Rooms > コラボレーション

コラボレーション 3
設定済みのテーブル

最終使用日 (1)

test2

最終使用日
3週間前

アクティブメンバーシップあり **参加可能 3** 利用できなくなりました

参加可能なコラボレーション (3)

名前	メンバーステータス	コラボレーションクリエイター	コラボレーションが作成されました
B社購買ユーザの広告接触における興味関心分析	招待済み	6[redacted]1 (A社)	2023年7月23日 17:20 (UTC+09)
test	招待済み	[redacted]	[redacted]
test-collaboration-001	招待済み	[redacted]	[redacted]

メンバーシップ作成にて参加完了

B社購買ユーザの広告接触における興味関心分析

B社購買データとA社広告接触データコラボレーションにおけるセグメント分析

概要

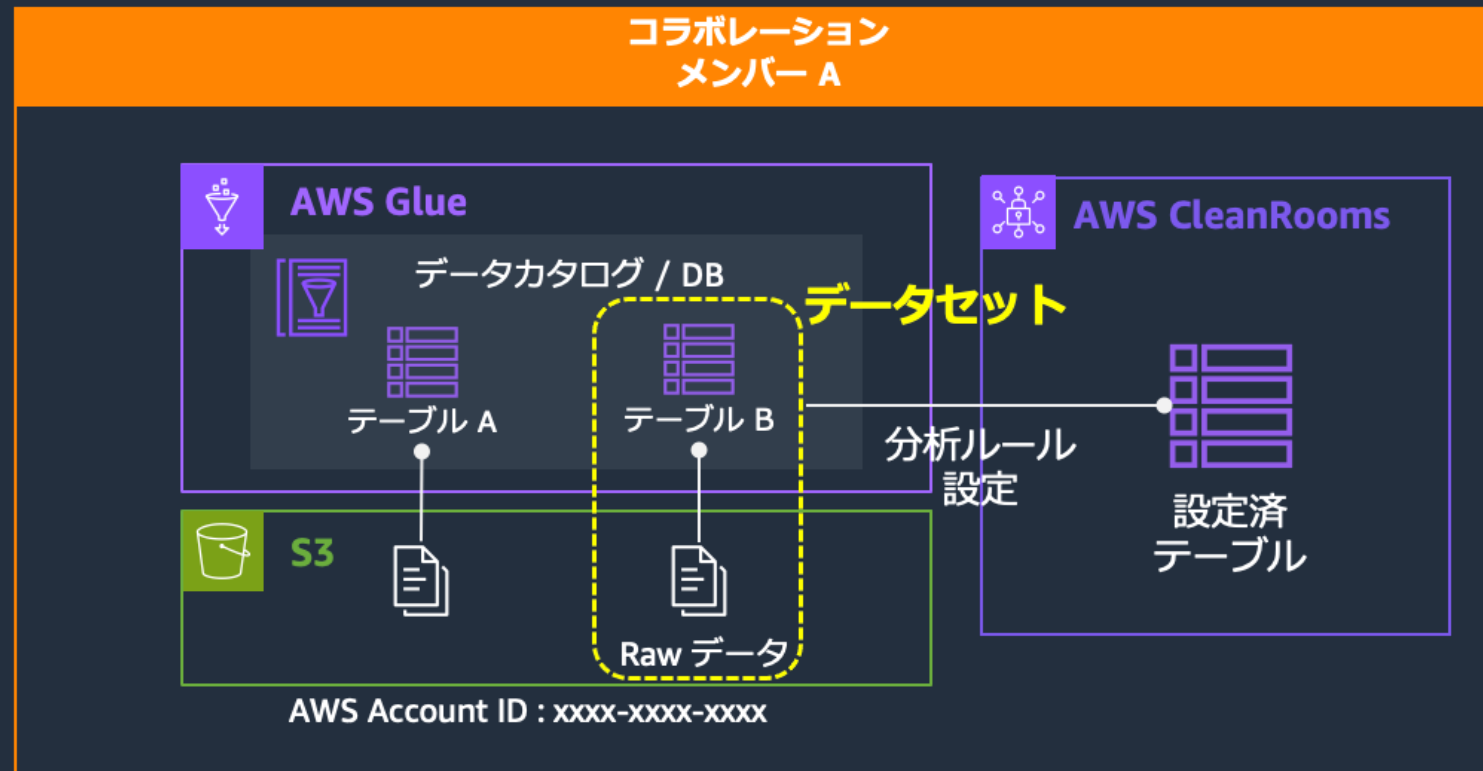
メンバーステータス 招待済み	コラボレーションクリエイター 6[redacted]1 (A社)	コラボレーション ARN arn:aws:cleanrooms:ap-northeast-1:[redacted]:collaboration/[redacted]	暗号コンピューティング サポートされていません
自身のメンバー能力 データを閲覧	メンバー 1/2人がアクティブ	クエリログ記録 サポートされています	

メンバーシップを作成



AWS Clean Rooms : 利用するデータ(データセット)

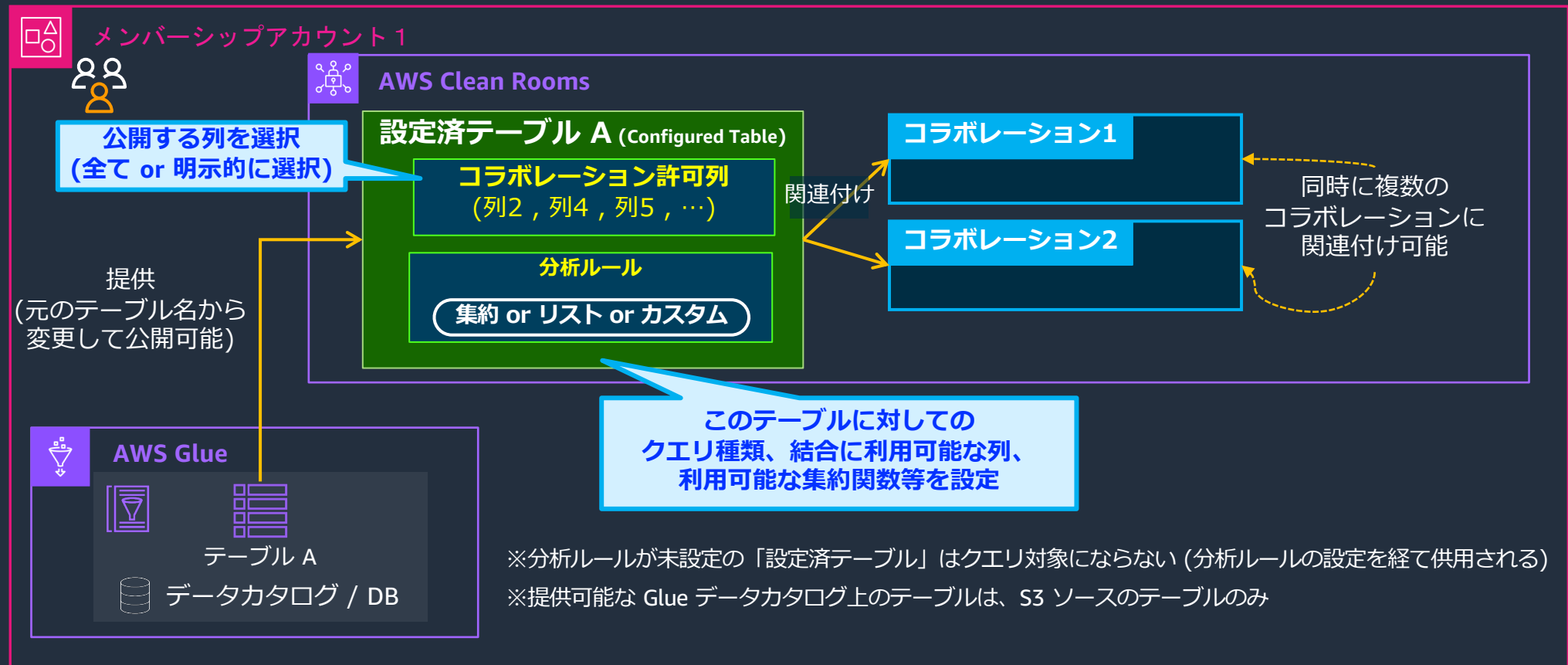
- Glue データカタログに登録されているテーブルを利用
※S3 データ(CSV / Parquet / iceberg (preview))が利用可能
- AWS Clean Rooms の機能を使ってRowデータを安全にコラボレーション可能に



分析ルール

AWS Clean Rooms : 設定済テーブル & 分析ルール

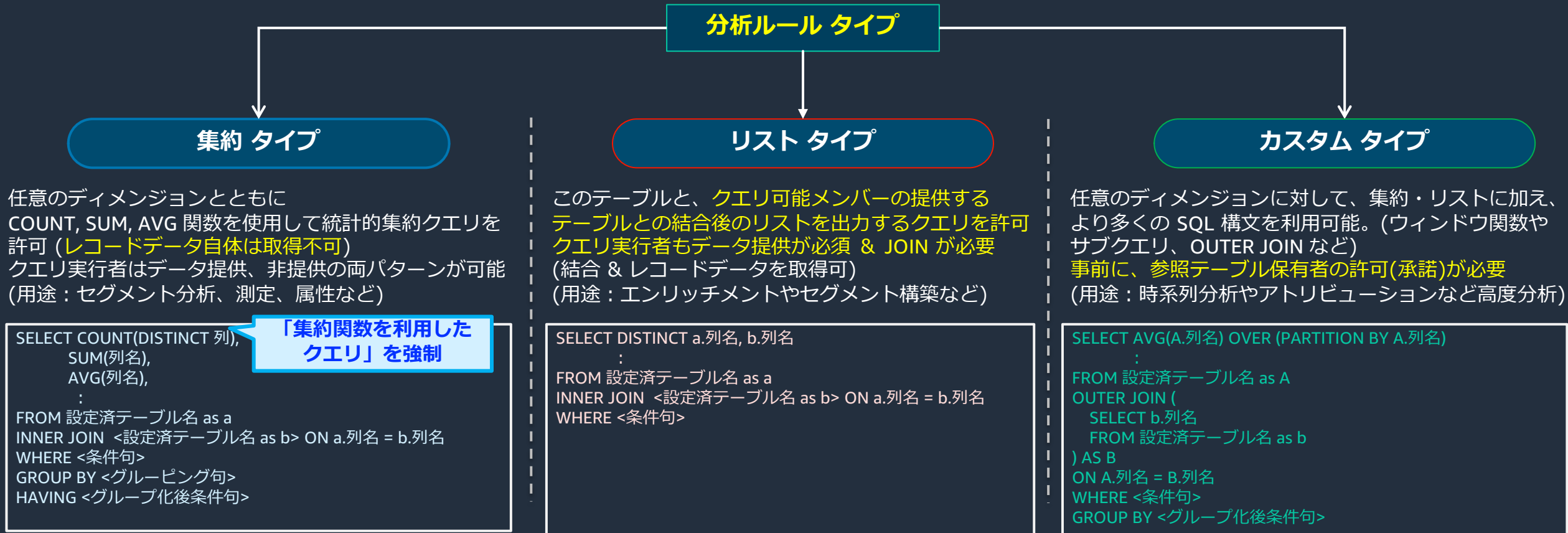
メンバーは、自アカウントの Glue データカタログ内テーブル(1 つ以上)を「列を選択」&「分析ルール」を設定し、コラボレーションに提供する



AWS Clean Rooms : 分析ルール (集約 or リスト or カスタム)

設定済テーブルのクエリは 必ず SELECT 文 として実行

→ 分析ルールは「どのような結合・どのようなクエリを許可するか」を設定



AWS Clean Rooms : 分析ルール (集約タイプ)

「集約タイプ」の分析ルールの場合、
集約関数での結果のみ得られるデータセットとして提供

分析ルール

集約タイプ

集約関数コントロール

※1 つ以上設定必須

列名

+

#	使用可能な集約関数
1	COUNT
2	COUNT DISTINCT
3	SUM
4	SUM DISTINCT
5	AVG

結合コントロール

単独クエリを許可する
(単独クエリ / 結合 両方 OK)

or

単独クエリを許可しない
(クエリメンバー所有テーブルとの
INNER JOIN のみ可能)

結合可能 列名 (0 個以上の列を指定)
(※ INNER JOIN 句内でのみ使用)

リストコントロール (結果出力列)

列名 (1 個以上の列を指定)

集約制約 (クエリ結果の最小集計閾値。最低1 つ指定, COUNT DISTINCT 値)

列名 + 条件値 → (COUNT DISTINCT(列) >= 値)

スカラー関数コントロール (このテーブルに利用出来るスカラー関数を指定。許可関数はディメンジョン列・集計列の両方で利用可能になる)

Clean Rooms で使用可能な全関数を許可

or

許可する関数をリストで指定

or

なし (スカラー関数の使用禁止)

AWS Clean Rooms : 分析ルール (リストタイプ)

「リストタイプ」の分析ルールの場合、
行データも結果セットとして得られるデータセットとして提供

分析ルール

リストタイプ

結合コントロール (結合可能列) ※少なくとも1 つ以上必須

列名 (1 個以上の列を指定)

(※指定した列は INNER JOIN 句で結合列として利用可能)

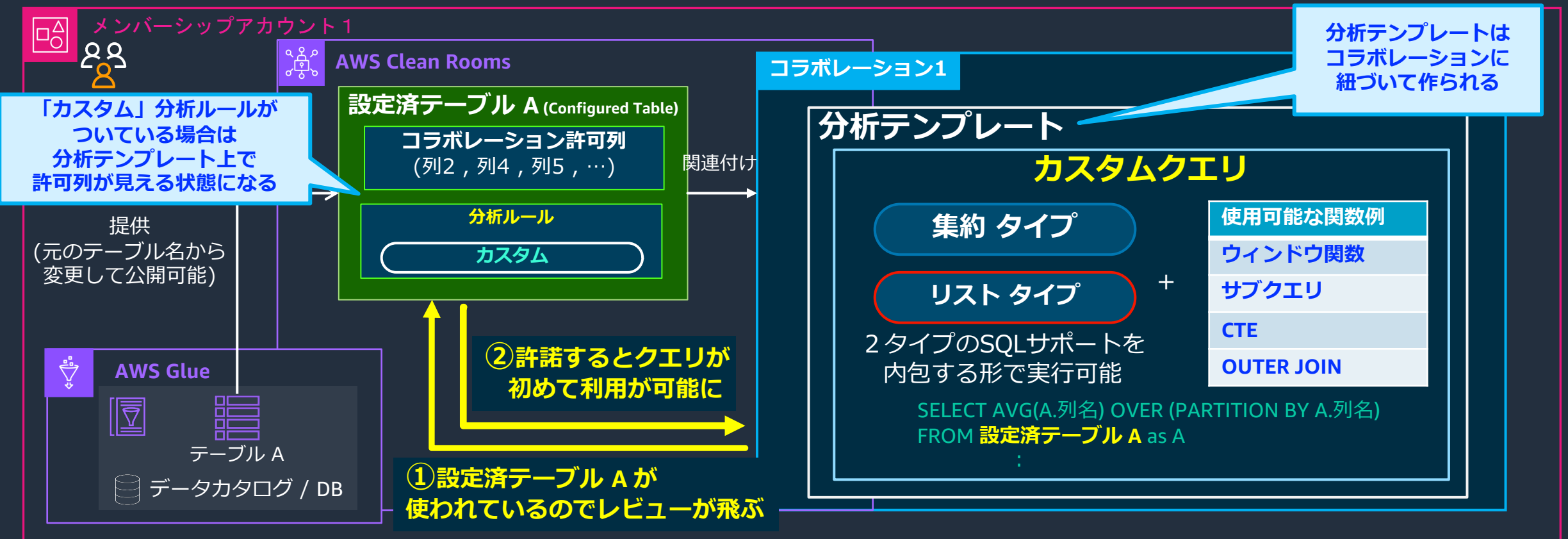
リストコントロール (結果出力列) ※オプション

列名 (0 個以上の列を指定)

(※指定しない場合は、全ての列を結果出力する)

AWS Clean Rooms : 分析ルール (カスタムタイプ) の動作概要

「カスタムタイプ」の分析ルールの場合、
分析テンプレート内に事前にカスタムクエリを作成し、
参照するデータ所有メンバーの利用許可(レビュー承諾)が必要



AWS Clean Rooms : 分析ルール (カスタムタイプ)

「カスタムタイプ」の分析ルールの場合、
集約タイプやリストタイプのような SQL 構造の制約がない
加えて、追加の構文、関数がサポートされている。

分析ルール

カスタムタイプ

コントロールタイプ (テーブルのデータを利用する際の許可)

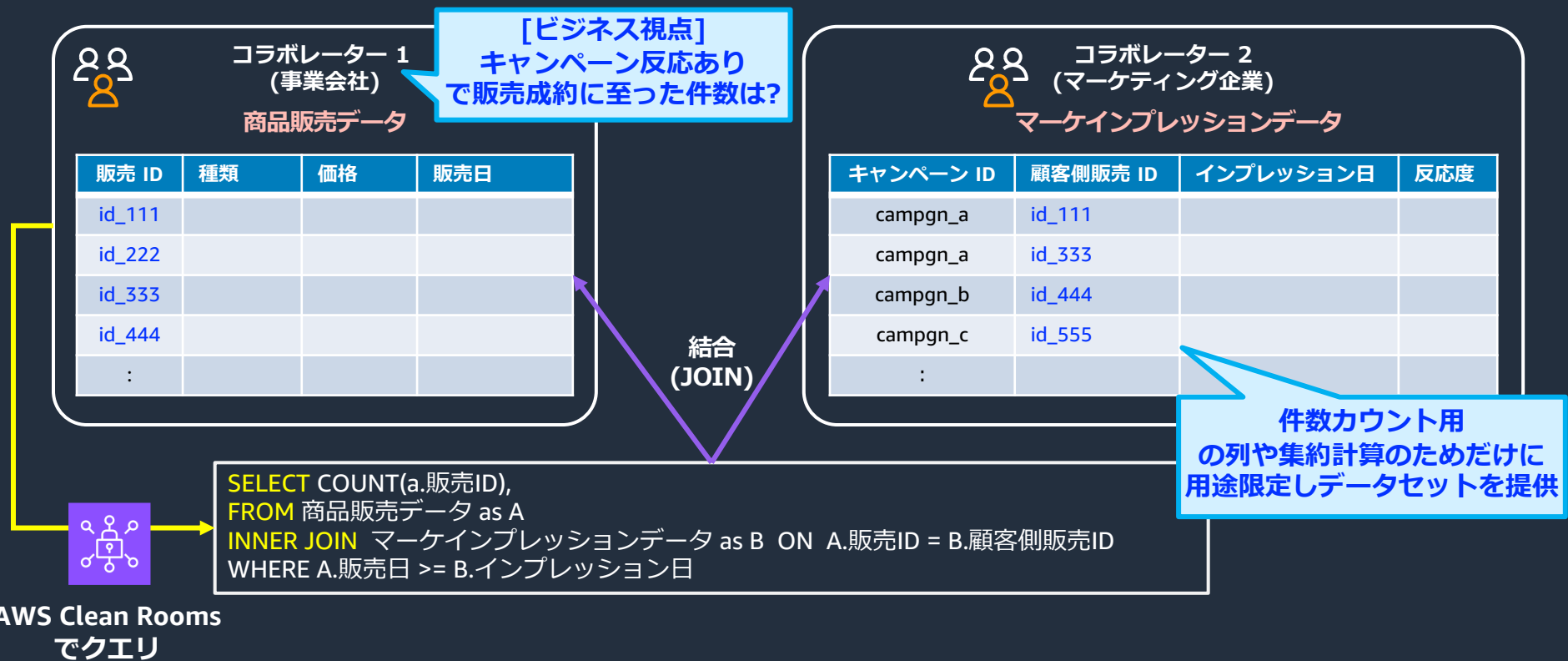
分析テンプレートにてテーブルデータを利用するクエリの場合はレビューを受け取る
(※所有テーブルが利用されるカスタムクエリは必ずレビューを受け取る。許諾後に初めて利用可能)

or

特定のメンバーが作成したカスタムクエリを全て許諾
(※指定したメンバーが作成したカスタムクエリはレビュー無しで利用可能)

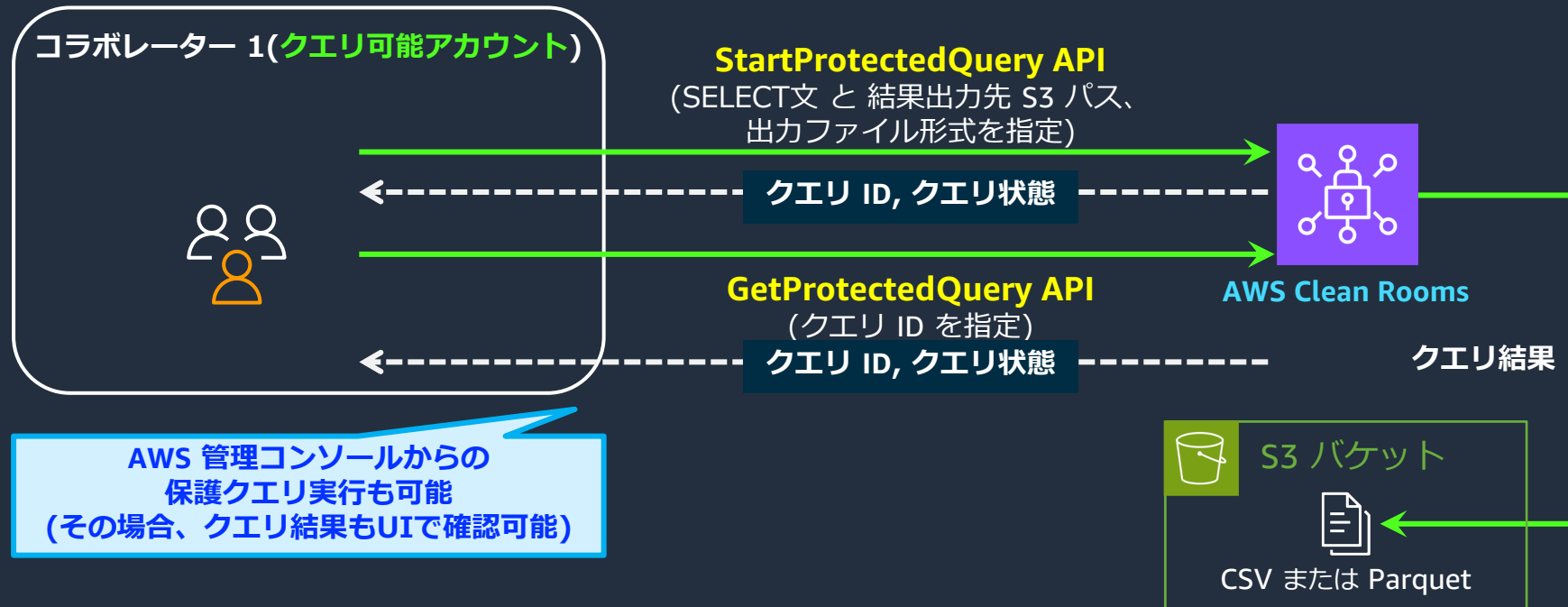
AWS Clean Rooms : 保護クエリ(Protected Query)

自社 および 他社が提供したデータセット群に対して、JOIN を伴う SQL (SELECT 文)を実行し、データ集計処理や情報を付加する



AWS Clean Rooms : 保護クエリ(API からの実行方法)

保護クエリ(SELECT 文)は、StartProtectedQuery API を利用して実行
→ クエリは非同期で実行され、結果は指定 S3 バケット*に出力



※クエリ結果受取メンバーの S3 バケットのパスを指定。

AWS Clean Rooms : 保護クエリ(サポート SQL 関数)

AWS Clean Rooms がサポートしているSQL 関数は下記を参照
(※SQL コマンドとしては SELECT 文のみ)

https://docs.aws.amazon.com/ja_jp/clean-rooms/latest/sql-reference/sql-functions-topic.html

サポート対象の SQL 関数 (一部抜粋)		
AVG	ABS	CAST
COUNT / COUNT DISTINCT	CEILING	LOWER
SUM / SUM DISTINCT	FLOOR	UPPER
日付・時刻データでの連結演算子	LN	RTRIM
CURRENT_DATE	LOG	COALESCE
GETDATE	ROUND	TRUNC
Window関数	RANK	JSON_PARSE

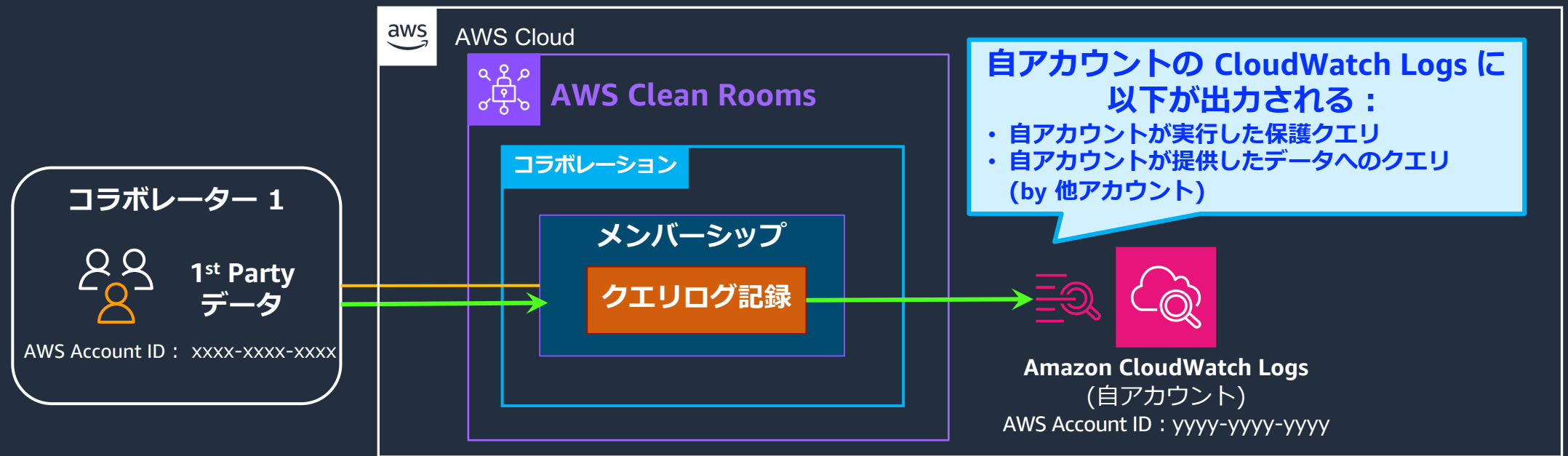
※扱えるデータ型や関数、変換ルール、構文や制約の詳細について、上記ドキュメントに記載されている為、ご利用時にご参照ください

ログ出力

AWS Clean Rooms : クエリログ記録

コラボレーション作成時にクエリログ記録を有効にした場合
各メンバーシップの作成時に「クエリログ」の有効化が可能

→ 自アカウントの CloudWatch Logs に AWS Clean Rooms の関連クエリログが保存



#	ログ種類	ロググループ名	ログストリーム名
1	クエリログ	/aws/cleanrooms/<collaboration-id>	YYYY/MM (例 : 「2023/03」)

暗号化コンピューティング

AWS Clean Rooms - 暗号化コンピューティング機能

メンバー間での事前の秘密鍵共有を前提に、保護クエリにおいて
保管・転送・処理中 を通じて常に暗号化された状態でデータ処理も可能



オープンソースの クライアントを利用

GitHub で公開され
無料で利用できる
暗号化クライアントツール
(<https://github.com/aws/c3r>)



Java SDK or CLI

当社の Java SDK または
使いやすいコマンドライン
インターフェースと統合



セキュアな マルチパーティ コンピューティング

安全なマルチパーティコン
ピューティング (SMPC) を
活用して、暗号化された
データを計算

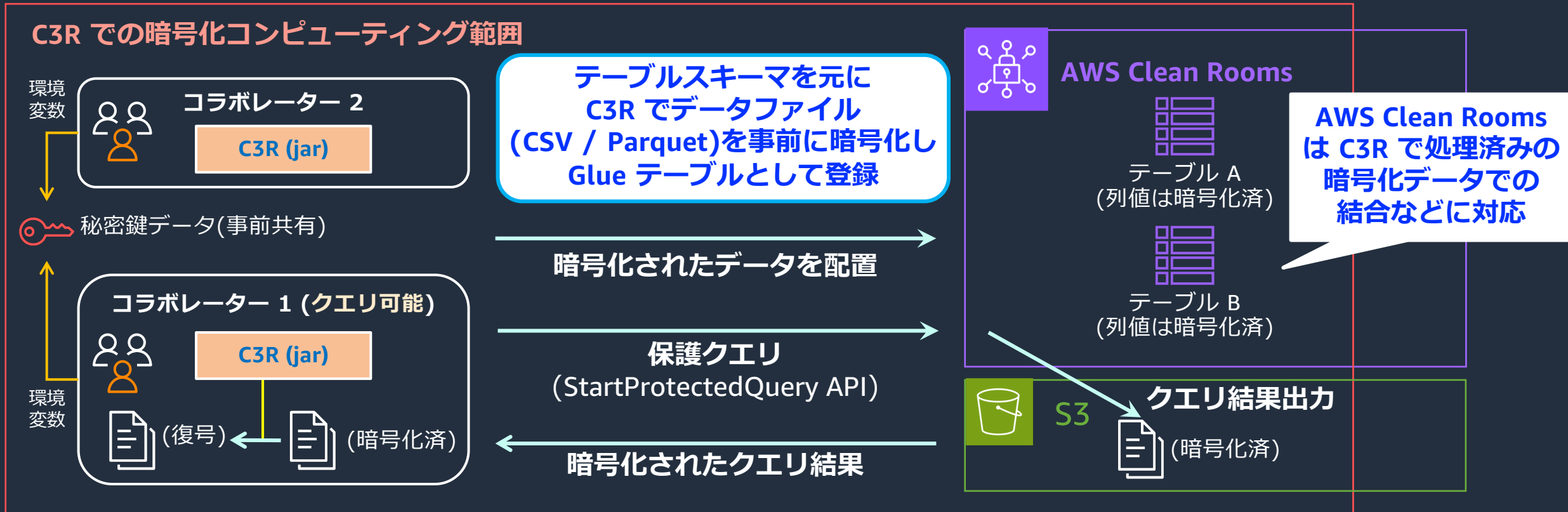


サポート データ形式

CSV または Parquet
データファイルのサポート

AWS Clean Rooms - 暗号化コンピューティング機能

OSS Java クライアント(C3R : Cryptographic Computing for Clean Rooms) を利用することで End to End で暗号化された状態での処理が可能



※通信の暗号化(TLS)とは別に、送信されるデータそのものや、Glue テーブルの列値が暗号化されている状態を実現

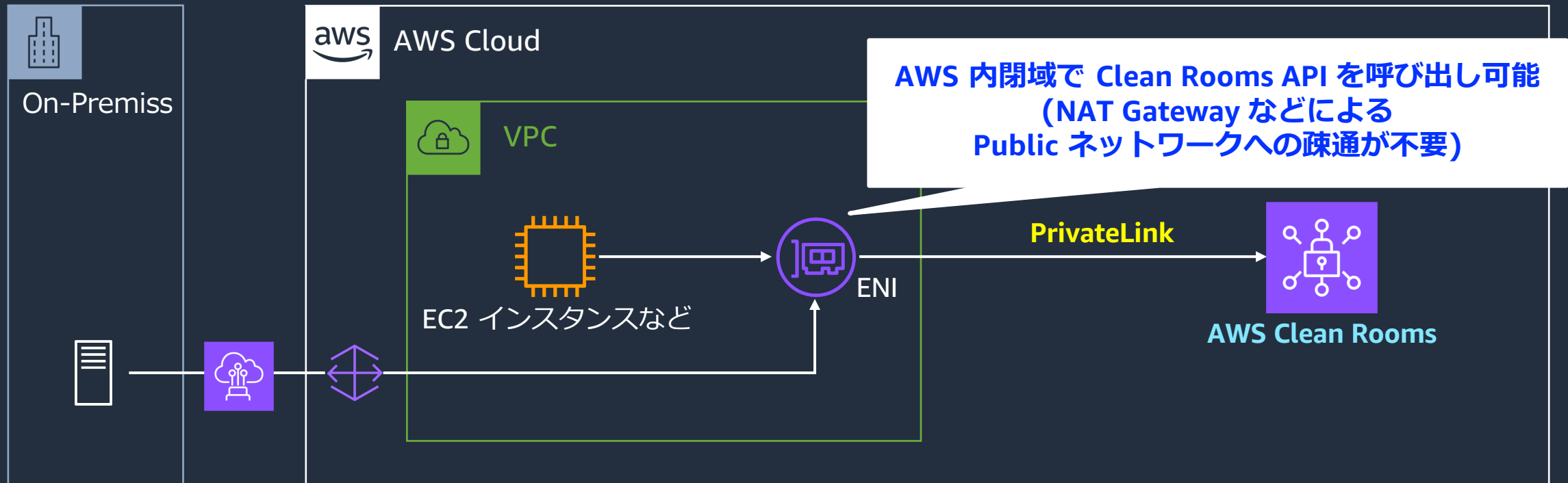
※C3R ユーティリティは Java11 以降 で動作



AWS Clean Rooms - AWS PrivateLink 対応

AWS Clean Rooms は AWS PrivateLink に対応
→ ユーザー VPC から AWS 内閉域で API を呼び出し可能

エンドポイントサービス名 :
`com.amazonaws.<region>.cleanrooms`



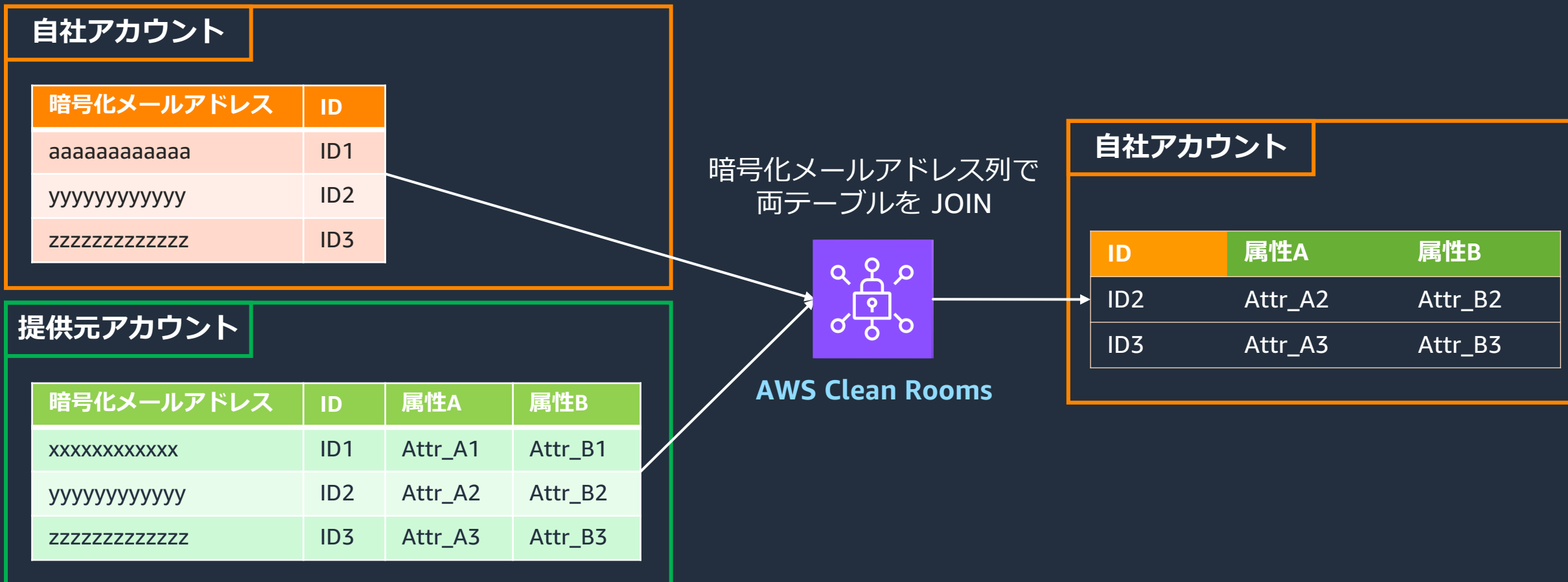
※VPC エンドポイントポリシーには未対応
(VPC エンドポイントでのセキュリティグループの利用による制御は可能)

アジェンダ

1. データ共有の価値と課題
2. AWS Clean Rooms の機能
3. **AWS Clean Rooms の利用方法・パターン**
4. AWS 他サービスとの連携
5. 料金・注意点
6. まとめ

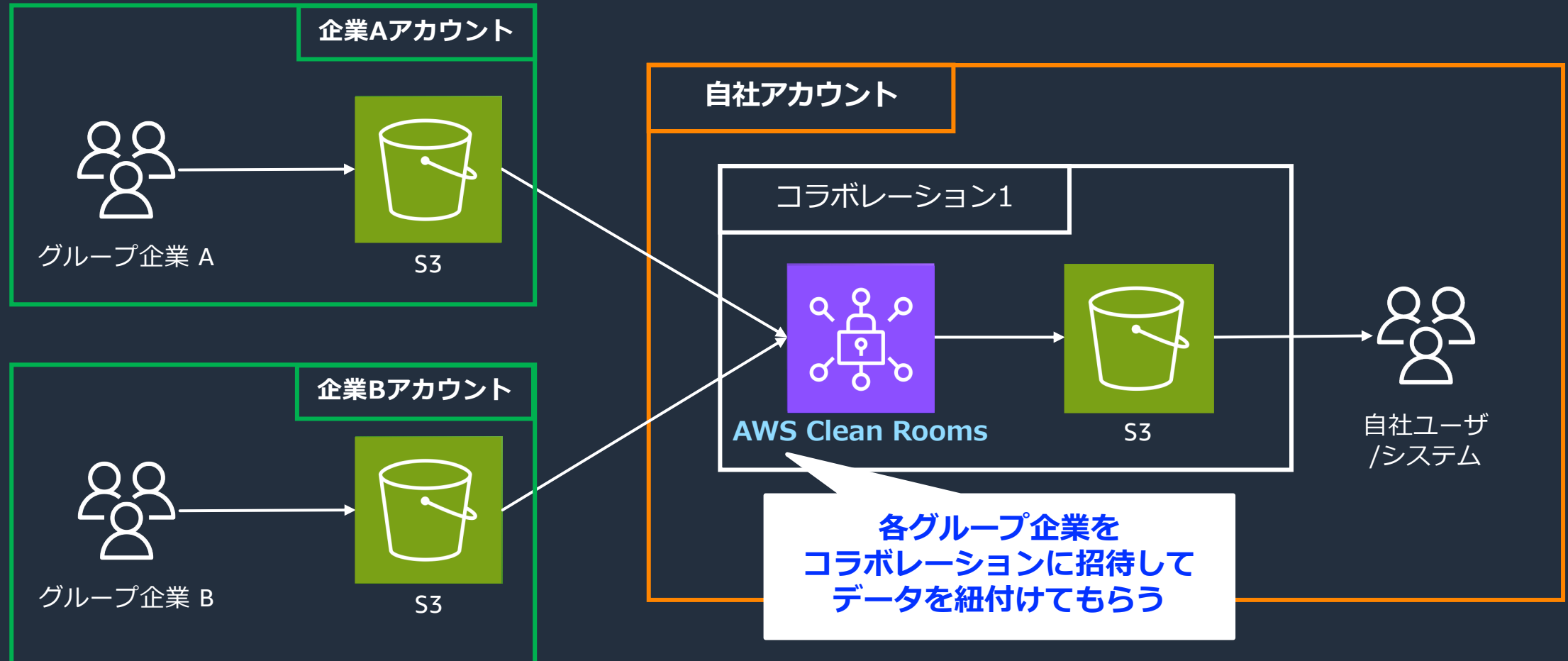
AWS Clean Rooms : ユースケース(基本)

提供元のデータから、自社のデータに関係あるデータだけを抽出・結合



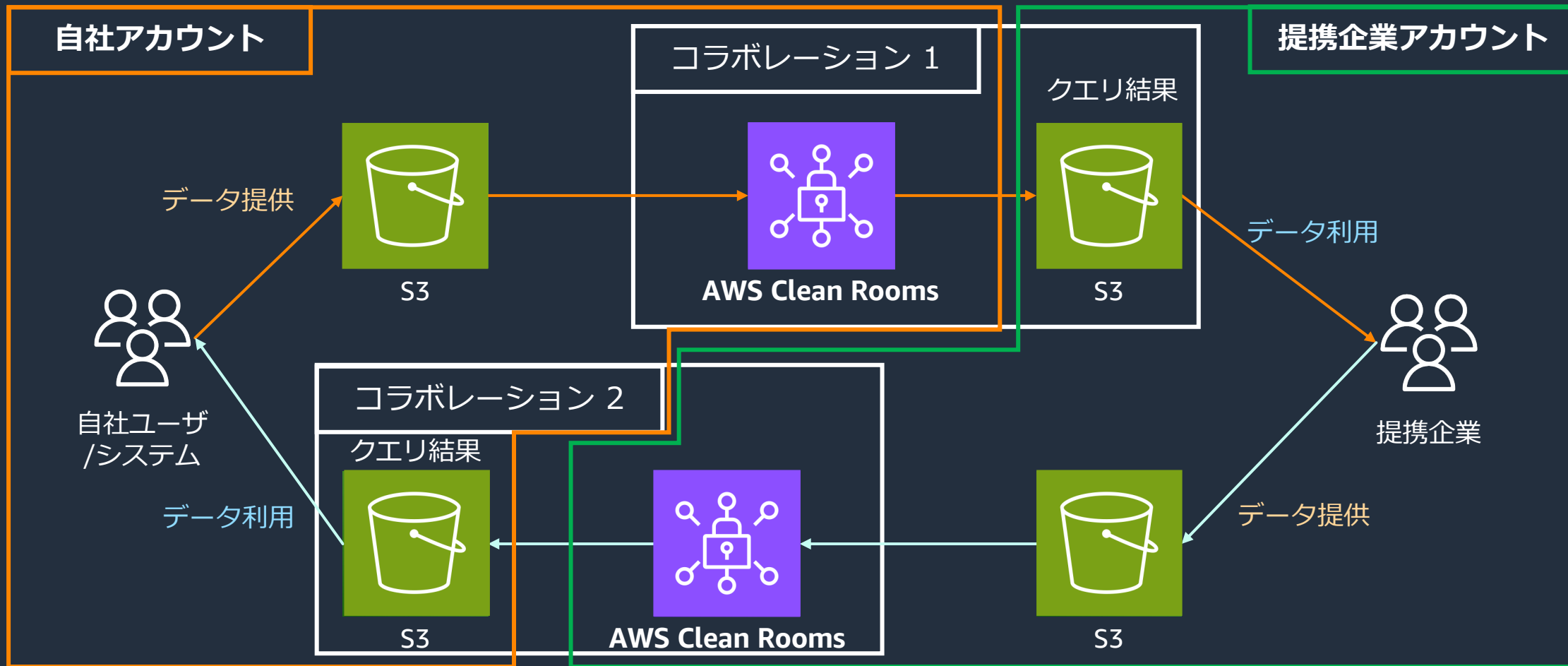
AWS Clean Rooms : ユースケース(グループ内連携)

グループ企業の各データをコンプライアンスを遵守した形で取得



AWS Clean Rooms : ユースケース(業務提携)

お互いに必要な情報を、コンプライアンスに遵守した形で提供・取得

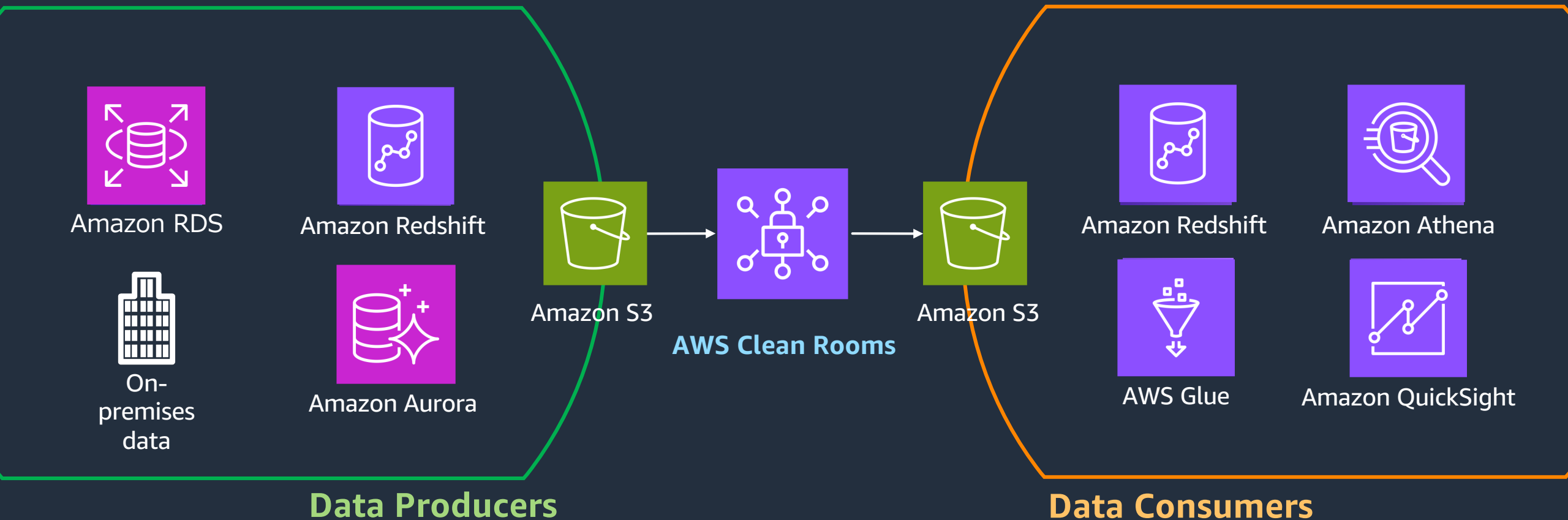


アジェンダ

1. データ共有の価値と課題
2. AWS Clean Rooms の機能
3. AWS Clean Rooms の利用方法・パターン
4. AWS 他サービスとの連携
5. 料金・注意点
6. まとめ

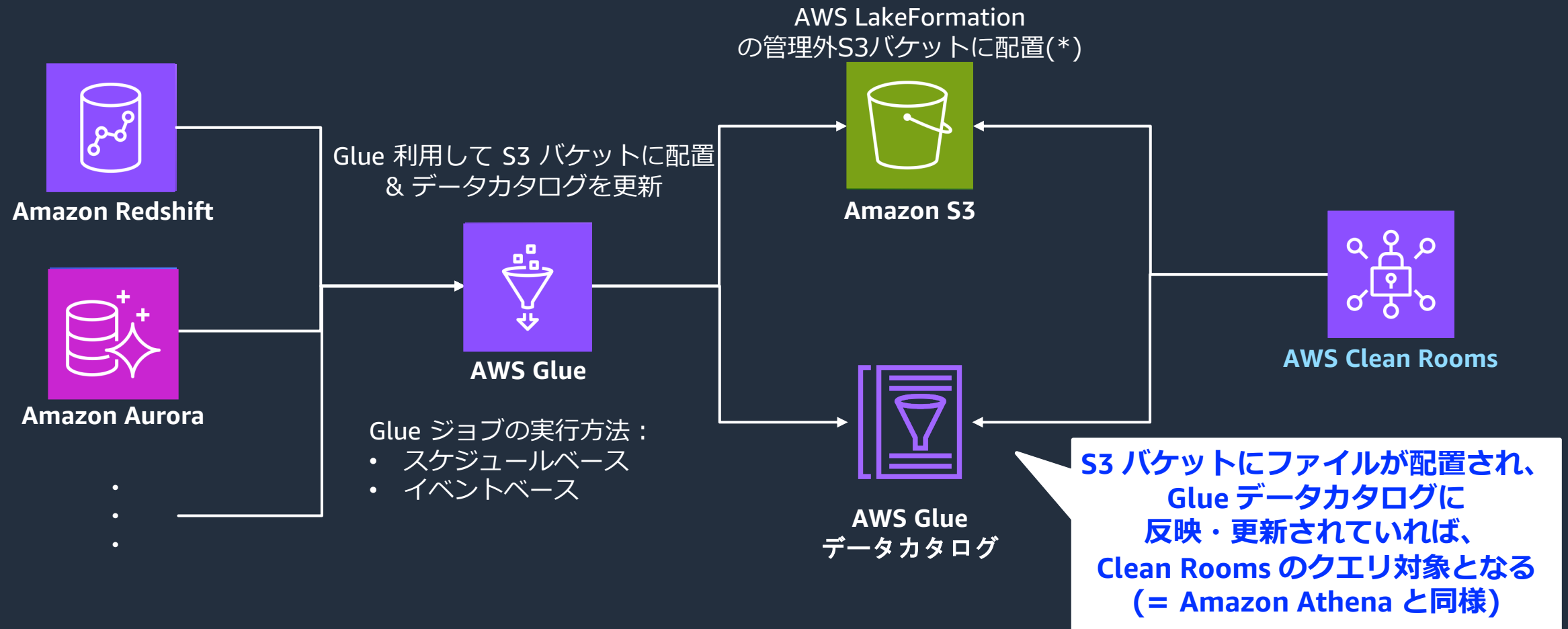
AWS Clean Rooms - アナリティクスサービス連携

Amazon S3 を中心に、AWS アナリティクスサービスと連携可能



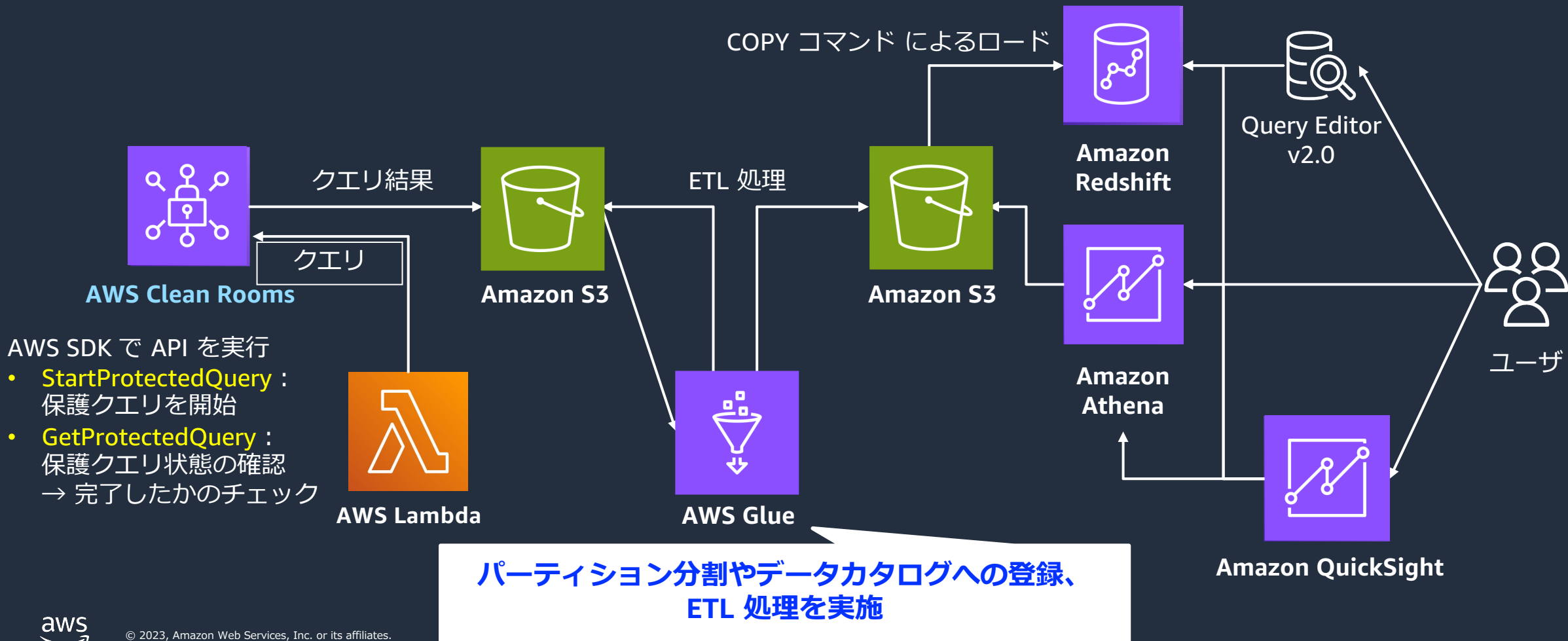
AWS Clean Rooms - 設定済テーブルへの自動連携

AWS Glue を利用してオリジナルのデータソースからデータを抽出
→ Glue データカタログの更新とあわせて S3 バケットに配置



AWS Clean Rooms からのデータ自動取得 & 活用

システム連携用途の場合、AWS API で 保護クエリ を実行
→ 結果ファイルは Glue ジョブでファイル分割やデータカタログ登録が便利



アジェンダ

1. データ共有の価値と課題
2. AWS Clean Rooms の機能
3. AWS Clean Rooms の利用方法・パターン
4. AWS 他サービスとの連携
5. 料金・注意点
6. まとめ

AWS Clean Rooms - サービス料金

※資料作成時点の東京リージョン料金を記載

CRPU (Clean Rooms Processing Unit) を単位とした
自アカウントのクエリ実行でのコンピューティング能力に基づく料金

AWS Clean Rooms サービス料金

①. 自アカウントのクエリ実行に対する消費 CRPU

集約、リスト分析ルールの場合
0.865 USD / CRPU-時間

カスタム分析ルールの場合
2.63 USD / CRPU-時間

※60 秒未満の CRPU 消費は60 秒に切上げ
※クエリに応じて自動的に CRPU はスケーリング
(クエリ群に対して **32 CRPU** がデフォルト容量として割り当て)

関連サービス料金

Amazon S3

+

Amazon CloudWatch Logs

※クエリログ有効化時

+

...

+

コラボレーションに参加 & データ提供のみであれば、AWS Clean Rooms の料金は発生しない



© 2023, Amazon Web Services, Inc. or its affiliates.

<https://aws.amazon.com/jp/clean-rooms/pricing/>

AWS Clean Rooms - 主なクォータ

AWS Clean Rooms の利用時は、下記クォータの確認

<https://docs.aws.amazon.com/clean-rooms/latest/userguide/quotas.html>

#	クォータ	クォータ値	上限緩和申請可否
1	コラボレーション毎の招待可能なメンバー最大数	5	Yes
2	アカウント毎のアクティブなメンバーシップ最大数	100	Yes
3	アカウント毎コラボレーション作成可能最大数	10	Yes
4	アカウント毎の設定済テーブル最大数	60	Yes
5	メンバーシップ毎のテーブル関連付け最大数	25	Yes
6	メンバーシップ毎の同時実行可能な最大クエリ数	5	Yes
7	設定済テーブル毎の許可対象リストに指定可能な最大列数	100	Yes
8	保護クエリ毎の設定済テーブル数	15	Yes
9	分析ルールの最大 JSON サイズ	100 KB	Yes
10	クエリ文の最大サイズ	16 KB	Yes
11	クエリの最大実行時間 (タイムアウト発生までの時間)	4 時間	Yes

アジェンダ

1. データ共有の価値と課題
2. AWS Clean Rooms の機能
3. AWS Clean Rooms の利用方法・パターン
4. AWS 他サービスとの連携
5. 料金・注意点
6. まとめ

まとめ

- AWS Clean Rooms を使ったデータクリーンルームの構築について、サービス概要と機能のご紹介しました
- 機密性の高いデータを安全にコラボレーションする為の方法として暗号化や分析ルールの機能詳細についてご紹介しました
- これからデータコラボレーションを運用設計される方に向けて、具体的な利用パターン、他サービスとの連携についてご紹介しました



Thank you!