

Oracle GoldenGateの優位性

Oracle Databaseに対するチェンジ・データ・キャプチャ（CDC）技術の比較

2021年7月、バージョン2.1
Copyright © 2021, Oracle and/or its affiliates
公開

本書の目的

本書では、Oracle GoldenGateの機能の概要と、Oracle GoldenGateがOracle Databaseに対する総合的に最適な選択肢である理由を説明します。

チェンジ・データ・キャプチャ（CDC）とデータ・レプリケーションのためのテクノロジーを選択する際は、データ・イベントがどのように取得されるかを深く理解することが不可欠です。すべてのデータベースには、データ・イベントを検出するためのオプションが1つ、または複数あり、各オプションには、理解した上で注意深く検討する必要があるトレードオフが存在します。たとえば、長年にわたりOracle Databaseには、データ・イベントを検出する方法が6つあります。

本書では、Oracle Databaseとともにデータベース・トリガー、シングルスレッドLogMiner API、XStream API、およびGoldenGateを使用する各CDCアプローチの技術的な違いについて説明します。さらに、サポートされていないサード・パーティ製ディスク・ログ・リーダーに伴うリスクについても簡単に説明します。

免責事項

本書は情報提供のみを目的としており、記載した製品機能の実装およびアップグレードの計画を支援することのみを意図しています。マテリアルやコード、機能の提供をコミットメント（確約）するものではなく、購買を決定する際の判断材料になさらないでください。本書に記載されている機能の開発、リリース、および時期については、弊社の裁量により決定されます。製品アーキテクチャの性質上、コードが大幅に不安定化するリスクなしに、本書に記載されているすべての機能を安全に含めることができない場合があります。

目次

本書の目的.....	2
免責事項	2
Oracle GoldenGate.....	4
Oracle GoldenGateの仕組み.....	5
GoldenGateトランザクションの分析構造	6
Oracle Databaseでイベントを検出するオプション.....	6
データベース・トリガーを使用した変更データ	7
Oracle Database LogMinerを使用した変更データ	8
LogMinerに伴う課題.....	8
XStreams APIを使用した変更データ	10
XStreamに関する重要な事実.....	10
Oracle GoldenGateを使用した変更データ	10
Oracle GoldenGateの一般的な優位性.....	12
サード・パーティ製ディスク・ログ・リーダーに関する注意事項.....	14
Apache Kafkaまたはオブジェクト・ストレージでのGoldenGateの使用について.....	15
サマリー：Oracle GoldenGateはOracle Databaseに最適なCDC	16

Oracle GoldenGateは、お客様とアナリストから、もっとも完全な機能、高いパフォーマンス、高い信頼性を備えたデータ統合およびデータベース・レプリケーション・ソリューションと評価されています。立証済みの異機種混合ソリューションであり、数百種類もの非Oracleデータベース、データ・ストア、クラウドの組合せと統合されます。Oracle Databaseにとって、GoldenGateはスケーラブルかつ完全で、全面的にサポートされる唯一のソリューションです。Oracle Databaseからデータ・イベントを取得するその他のフレームワークやツールは、未完全であったり、スケーラブルでなかったり、サポートされていなかったりします。たとえば、LogMinerは、機能が制限されたシングルスレッドの診断APIです。LogMiner APIに依存するサード・パーティ製のデータ統合テクノロジーは、Oracle Databaseとともに使用する場合、Oracle GoldenGateに到底及びません。

本書の主眼と目的は、Oracle Databaseからデータ・イベントを取得するための選択肢、GoldenGateの適応方法、およびGoldenGateが最適な選択肢である理由について説明することです。

Oracle GoldenGate

1990年代に創業したサンフランシスコの新興企業であるGoldenGateは、Tandem NonStopデータベースからネットワーク接続された稼働中のATM（現金自動預け払い機）で、事業継続性とデータの高可用性を実現することが当初の目的でした。

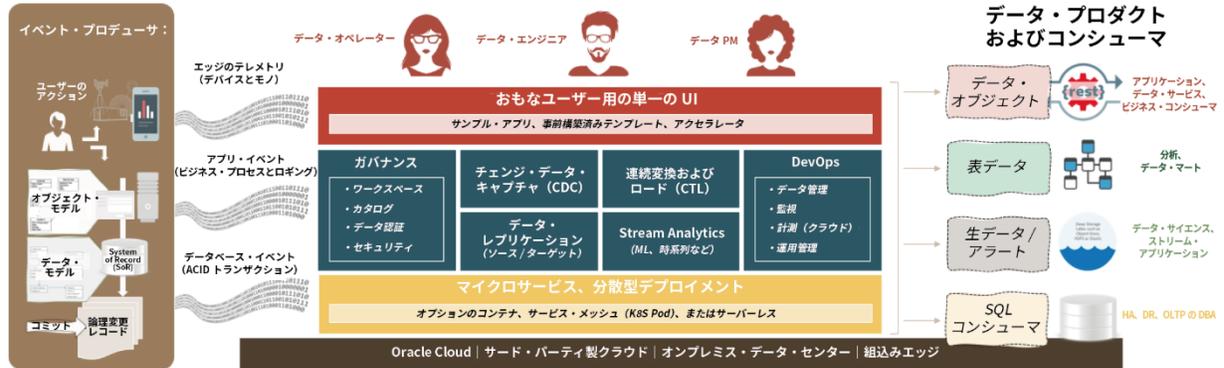
データ・イベント	データ・パイプライン	Stream Analytics
データの高可用性 <ul style="list-style-type: none"> Oracle DB/非Oracle DB 停止時間の短い移行 	リアルタイムのデータ取込み <ul style="list-style-type: none"> 忠実性の高い変更ストリーム データ・レイク/データウェアハウス 	データストリームの分析 <ul style="list-style-type: none"> 時系列分析、機械学習、地理空間、ML/AI
トランザクション・レプリケーション <ul style="list-style-type: none"> OLTP/参照データ 	ストリーム統合 <ul style="list-style-type: none"> 準備、フィルタリング、相関付け、強化 	監視、アラート、配信 <ul style="list-style-type: none"> シンプルなメッセージタイプのストリーム 複合イベント処理（CEP）
チェンジ・データ・キャプチャ <ul style="list-style-type: none"> Oracle DB/非Oracle DB サービス・メッシュでのCDC 	CTLおよびデータ統合 <ul style="list-style-type: none"> データの連続変換およびロード（CTL） 	ノーコード開発 <ul style="list-style-type: none"> 20以上の事前構築済みアクセラレータ
ガバナンス		
エンタープライズ・グレードのセキュリティ <ul style="list-style-type: none"> 証明書、SSO、暗号化など 	データ認証 <ul style="list-style-type: none"> 一貫性を比較して修正 	メタデータ・カタログ <ul style="list-style-type: none"> データ資産、系統、ポリシーなど

図1：Oracle GoldenGateプラットフォームの機能

現在、Oracle GoldenGateは、オラクルとオラクル以外のデータソースを扱う統合機能および分析機能を備えた成熟したプラットフォームへと拡張されました。トランザクションの送信が保証されるという信頼性のおかげで、GoldenGateは、Oracle MAA（Maximum Availability Architecture）の‘プラチナ層’サービス・レベルを支えるテクノロジーになることができました。このような信頼性は、NonStop、SQL Server、DB2 iSeries、メインフレーム、その他のサポート対象データ・ストアをはじめとする非Oracleデータベースに対しても提供されるようになってきました。何千ものグローバルな銀行、小売業者、通信事業者、医療組織などが、Oracle GoldenGateという信頼性のある基盤上で業務系データ・プラットフォームを実行しています。

GoldenGateは、データ・イベントを検知し、非常に短い待機時間でネットワーク全域に転送できるリアルタイム・データ・レプリケーション・プラットフォームです。GoldenGateテクノロジーは、業務系データベース、停止時間の少ないデータ移行、オンラインのマルチアクティブ・データ・ストア、クラウドへのリアルタイムのデータ取込み、データ・レイク、データウェアハウスなどの地理的なシャーディングに使用されます。2015年以来、GoldenGateはポリグロット・ビッグ・データとNoSQLデータ・パイロードにますます重点を置くようになり、ネイティブなマイクロサービスの‘サービスとして’のデプロイメント向けに完全にリファクタリングされています。

図2：Oracle GoldenGateプラットフォームの論理的アーキテクチャと主要コンポーネント



2018年にGoldenGateプラットフォームは、順序付けされたデータ処理をナノ秒スケールに保ちながら、1秒あたり数十億イベントにまでスケールする堅牢な複合イベント処理 (CEP) コア・エンジンを備えたStream Analyticsとデータ・パイプラインを追加しました。このイベント・エンジンは、非常に強力なセマンティックを変換や分析で使用し、大規模パラレル・プロセス (MPP) のためにオープンソースのApache Spark上で実行できます。

GoldenGateはこれらの新機能を使用して、高価値のデータ・プロダクトをデータ・コンシューマに直接提供します。これまでは、GoldenGateは主に待機時間の短い生データをデータ・パイプラインやその他のデータ・プロダクトに提供するために使用されていました。現在は、生データ・イベントをプッシュできるほか、高価値のデータ・プロダクトを提供できます。

Oracle GoldenGateの仕組み

GoldenGateのCDC機能は、基本的にデータ・イベントを検出し、送信するために使用されます。データ・イベントには、通常、データベースDML (データ操作言語) とDDL (データ定義言語) が含まれます。GoldenGateはOracle Databaseでプロシージャ・イベントも検知および送信できるため、データベースを同期し続けるのに役立ちます。

GoldenGateは、データ・イベントの信頼されるプロバイダとなるように、一から設計されています。また、Oracle MAAのユースケースのほか、DW (データウェアハウス)、データ・レイク、データ・ストリーム (Kafkaベース) などのミッション・クリティカルなシナリオにおいても信頼性を発揮します。

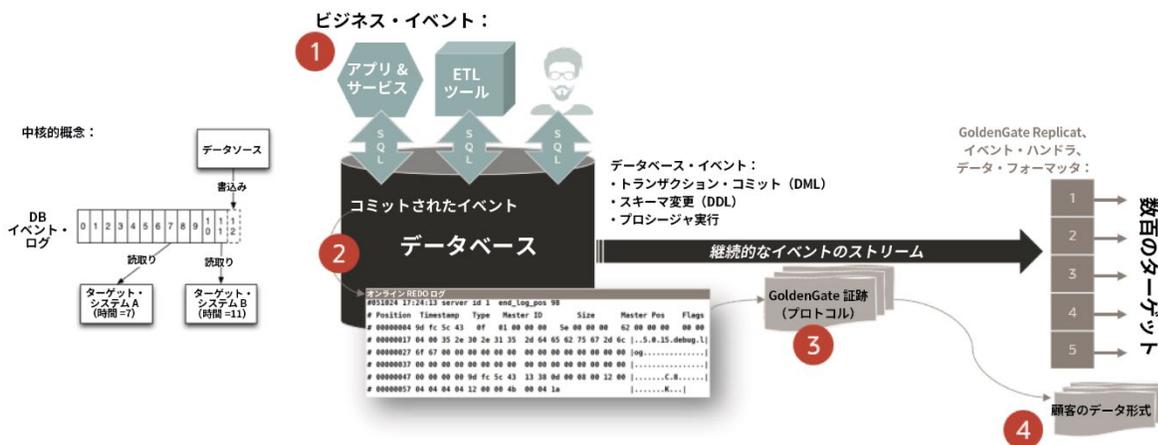


図3：Oracle GoldenGateにおける変更の検出と伝播

GoldenGateトランザクションの分析構造

データベース・ログ・イベントが取得されると、証跡と呼ばれるGoldenGateの標準台帳に変換されます。サポートされるソースやプラットフォームの数百もの組合せに対するデータ・イベント・ログはすべて、この標準的な単一の証跡形式に変換されます。これが、GG Distributionマイクロサービス経由で配布される証跡です。各証跡には、1つまたは複数の‘パス’があり、証跡を使用する側が、パスへのアクセスを独自に開始できます。あるいは、証跡を下流のReceiverマイクロサービスにプッシュするようにGoldenGateを構成することもできます。

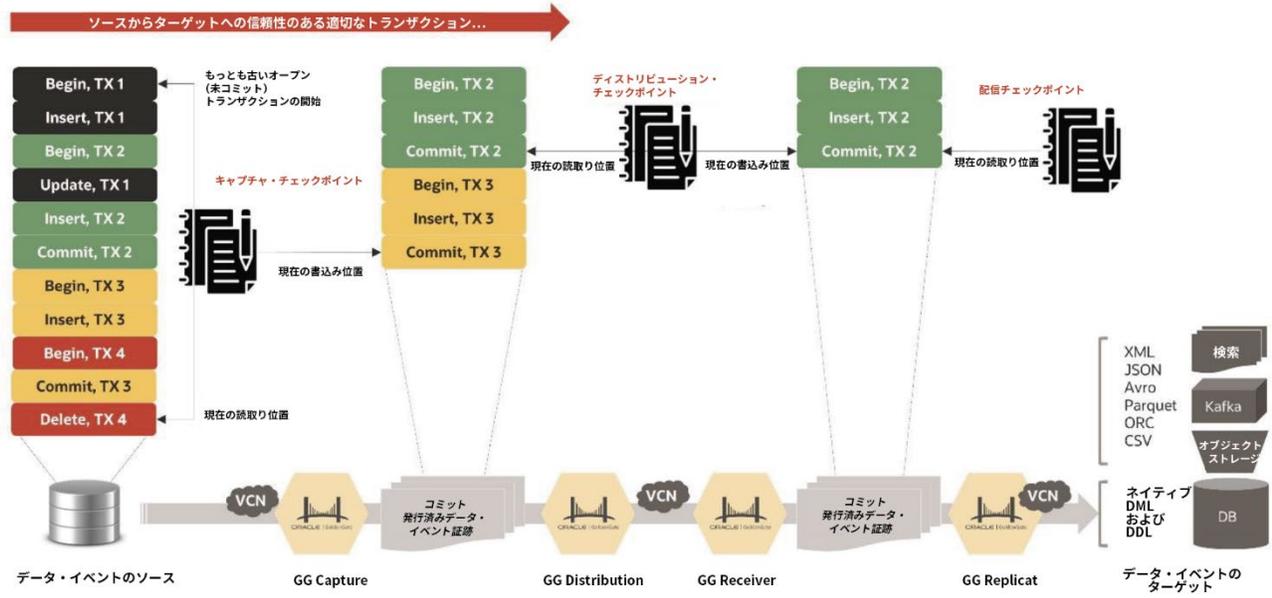


図4：ACIプロパティを分散アーキテクチャに保存するGoldenGateトランザクションの分析構造

GoldenGateは、データベース・ログを発行するときのソースでの操作をすべて認識しており、種類の異なる各データベースは、異なる方法で操作を処理します。原則として、GoldenGateはデータ・イベントがソース・システムでコミットされるとすぐに、データ・イベントを発行します。データベース・システムでは、多くの操作を単一のコミットにまとめることができます。実際、ある程度長い時間実行されるトランザクションでは、単一のトランザクションの一環として数百万のオブジェクトが変更される場合があります。そのようなトランザクションは完了まで数時間かかり、数千の小規模な未コミット操作がインターリーブされています。GoldenGateは、このようなトランザクションがネットワーク全体を移動するときにグループ化し、分離し、トランザクションの一貫性を維持します。

Oracle Databaseでイベントを検出するオプション

長年にわたりOracle Databaseには、チェンジ・データ・キャプチャ（CDC）で利用できる多くのAPIがあります。現在は、以下に示す少数の技術が、オラクルのエコシステムの戦略的要素として残っています。

- **データベース・トリガー**は、保管されたPL/SQLブロックであり、表に関連付けられ、INSERT、UPDATE、DELETEなどのDML文が発行された時に実行されます。トリガーを使用してイベントを検出するには、スキーマとアプリケーションに関する詳しい知識が必要であり、これは従業員の異動や質の悪いドキュメントの原因となる傾向にあります。この極めて侵襲的な戦略は、不完全であるとともに、多大な実行時オーバーヘッドを招く可能性があります。言うまでもなく、多くのETLベンダーは、データベース・トリガーを使用したCDCを引き続きサポートします（ドキュメント：https://docs.oracle.com/cd/F19136_01/lnpls/CREATE-TRIGGER-statement.html#GUID-AF9E33F1-64D1-4382-A6A4-EC33C36F237B）。
- **Oracle LogMiner**はOracle Databaseのコンポーネントです、このコンポーネントを使用すると、ユーザーは、SQLインタフェースを介してオンラインおよびアーカイブのREDOログ・ファイルを問い合わせることができます。LogMinerは、追加ライセンスなしで無償で使用できるデータベースAPIであるため、多くのETLベンダーやレプリケーション・ベンダーに人気のオプションです。LogMinerがアクティブ化されていると、Oracle Databaseの更新パフォーマンスがある程度顕著に低下する場合があります。

(ドキュメント：https://docs.oracle.com/cd/F19136_01/sutil/oracle-logminer-utility.html#GUID-3417B738-374C-4EE3-B15C-3A66E01AE2B5)

- **Oracle XStream**は、クライアント・アプリケーションがOracleデータベースとの間でデータ変更を送受信できるようにするOracle Databaseのコンポーネントとアプリケーション・プログラミング・インタフェース (API) で構成されます。XStreamはOracle GoldenGateとともにライセンス供与され、Oracle Databaseからの高速CDCに対応するために、ISVアプリケーション・パートナーと他社製レプリケーション・ツールによって使用されます (ドキュメント：https://docs.oracle.com/cd/F19136_01/xstrm/introduction-to-xstream.html#GUID-5939CB6C-8BA9-4594-8F96-B0453D246722)。
- **Oracle GoldenGate**は、最適なパフォーマンスと極めて豊富な機能により、Oracle Databaseの変更を検出し、トランザクションをレプリケートします。最新のマイクロサービス・アーキテクチャを備えたGoldenGateは、業務系のITシステム、アプリケーション、および分析データ・プラットフォームをモダナイズするための理想的なソリューションです (ドキュメント：<https://docs.oracle.com/en/middleware/goldengate/core/21.1/index.html>)。

本書では、Oracle Databaseの変更を検出する以下の非戦略的な代替方法についても説明します。

- **サポート対象外のディスク・ログ・リーダー**は、ディスクベースのストレージからOracle Databaseログを直接マイニングするサード・パーティの専用ツールです。これらのアプローチでは、Oracle DatabaseのデータファイルとAPIがリバース・エンジニアリングされています。オラクルによってサポートされておらず、セキュリティ上の重大なリスクをもたらすアプローチであり、オラクルのライセンス同意書に違反する可能性があります。本書の後半で、この技術に伴うリスクについて説明します。

Oracle GoldenGateは、Oracle DatabaseでのCDCの使用とレプリケーションにおいて推奨されるアプローチです。300を超えるソースとターゲットの組合せをサポートしており、すべてのバージョンとパッチ・レベルのOracle Databaseに対応しています。GoldenGateは、ログ・イベントを検出するとき、およびOracle Databaseからログ・レコードを読み取るときに、緊密に統合された独自のログ・キャプチャAPIを使用するため、もっとも侵襲性が低く、極めて高速かつスケーラブルなソリューションです。

データベース・トリガーを使用した変更データ

トリガーの使用は、表の数が少ない場合やトランザクション量が多くないユースケースでは、DMLイベントの検出に許容される方法です。ただし、各トリガーは、データベースに新しいオブジェクトを配置し、トリガーされる各イベントのストアド・プロシージャを基本的に常時実行しています。そのため、トリガーを大規模に使用すると、アプリケーションとデータベースに相当なオーバーヘッドが生じる可能性があります。ほとんどのハイエンド・ツールはトリガーを使用しませんが、一部の主流のETLツールでは、依然としてETLフローの一環としてトリガーを設定および構成する開発者向けのツールが提供されています。このようなETLツールは、データベースのパフォーマンスに悪影響を及ぼす可能性があることを認識しておく必要があります。

この例では、条件を使用して、可能な4つのトリガー文のどれが実行されたかを判断するDMLトリガーを作成しています。

```
CREATE OR REPLACE TRIGGER t BEFORE
    INSERT OR
    UPDATE OF salary, department_id OR
    DELETE
    ON employees
BEGIN
CASE
    WHEN INSERTING THEN
        DBMS_OUTPUT.PUT_LINE('Inserting');
    WHEN UPDATING('salary') THEN
        DBMS_OUTPUT.PUT_LINE('Updating salary');
```

```
WHEN UPDATING('department_id') THEN
    DBMS_OUTPUT.PUT_LINE('Updating department ID');
WHEN DELETING THEN
    DBMS_OUTPUT.PUT_LINE('Deleting')
;
END CASE;
```

図5：DMLトリガーの例

Oracle Database LogMinerを使用した変更データ

これまで、オラクルはサポートと診断のユースケースのために、LogMinerを作成し、保守していました。多くのサード・パーティ・ベンダーがチェンジ・データ・キャプチャでこのAPIを使用していますが、そのようなユースケースはAPIの当初の目的ではありませんでした。

データベース・リカバリ操作が、リレーショナル・データベースのACID（原子性、一貫性、整合性、永続性）プロパティに準拠して実行されるようにするために、オラクルのREDOログは、ユーザー・データまたはデータベース・ディクショナリに加えられたすべての変更を取得します。

LogMinerは、Oracle SupportがREDOログ・ファイルのコンテンツを確認するために使用する、Oracleデータベース内に組み込まれたユーティリティです。監査ツールや診断ツールであるとともに、データ分析ツールでもあります。LogMinerはDBAとして、PL/SQLプロシージャおよび関数を使用して、REDOログ・ファイル内の変更レコードを見つけます。

LogMinerのおもな機能の一部を以下に示します。

- アプリケーション・レベルで作成されたエラーなど、データベースに対する論理的な破損が開始された場合の追跡機能。時間ベースまたは変更ベースのリカバリーをいつ開始するかを把握できるよう、エラーがいつ作成されたかを正確に知ることが重要です（ドキュメント：<https://docs.oracle.com/en/database/oracle/oracle-database/21/sutil/oracle-logminer-utility.html#GUID-CF064432-57A0-4891-ABE5-800DF327615A>）。
- 表を初期状態に戻すために、一連のリバースSQL文を使用して、きめ細かいリカバリーを実行します。
- 傾向分析機能を使用したデータベース・チューニング。これまでのディスク・アクセス統計を大局的に把握するために、更新と挿入をもっとも受け取った表を見極めます。
- トランザクション一貫性を確認します。DDLとDMLだけでなく、実行の順序や文を実行したユーザーを追跡できるためです。

LogMinerはOracle Database向けの強力なユーティリティですが、LogMinerのユースケースを本来の目的以上に拡大しようとすると、いくつかの課題が発生することを認識しておく必要もあります。

LogMinerに伴う課題

以下のセクションに、LogMinerを直接使用するお客様、およびLogMinerを基盤とするサード・パーティ製CDCツールを使用するお客様の全員が認識すべき潜在的な課題の一部を挙げます。

- LogMinerは、データベースREDOログの診断ツールとして設計されました。REDOログの継続的なマイニングに使用すると、最適なパフォーマンスは得られません。
- LogMinerはシングルスレッドであり、REDOレコードの検索中のオーバーヘッドが低くなるように設計されていません。
- LogMinerは、現行のシステム変更番号（SCN）に到達すると停止します。全表スキャン（FTS）操作と同じく、ログの先頭から新規リクエストが開始されます（ドキュメント：<https://docs.oracle.com/en/database/oracle/oracle-database/21/sutil/oracle-logminer-utility.html#GUID-319446A8-6FEC-42CE-A6A4-582CA65377CF>）。
- CDCにLogMinerを使用した場合にロールバックまたは部分ロールバックを行うと、厄介な状況になる場合やエラーが起こりやすくなる場合があります。アプリケーション・セッションが強制終了した場合（突然終了するなど）、トランザクションの失敗に対してロールバックが発生します。トランザクションがキャンセルされるか失敗に終わると、他のトランザクションが進行できるようにするために、データベースはこのトランザクションが実行した未コミット操作をクリーンアップする必要があります。このクリーンアップには、未コミット操作のロールバックが含まれます。

LogMinerから見ると、ロールバック文はSQL_UNDOではなくSQL_REDOとして自動的に報告されます。ロールバックするSQLでは、UNDO SQLは生成されず、ロールバック・フラグが設定されます。

- Oracle Database 12.2より、LogMinerのcontinuous_mineオプションは非推奨となり、Oracle Database 19.1以降は使用できなくなりました。continuous_mineを使用するサード・パーティ製ツールはいずれも、19c以降のデータベースでは機能しません。詳細、および廃止された全機能の確認については、ドキュメントを参照してください（ドキュメント：https://docs.oracle.com/cd/F19136_01/upgrd/behavior-changes-deprecated-desupport-oracle-database.html#GUID-3BAFD95E-4D00-4F0F-BC80-6064F497F878）。
- DB 12.2以降、次に挙げるデータタイプはサポートされないSQL_REDOとSQL_UNDOを生成します。JSON、OSON、BFILE、ネストした表、ネストした表を使用するオブジェクト、Identity列を使用する表、時制有効性列、PKREF列、長い識別子、PKOID列、ネストした表属性とスタンドアロンのネストした表列（データタイプは、Oracle Database 12.2以降のバージョンでDBMS_ROLLINGと互換性があります）。

LogMinerベースの独自クライアントを実装している各ベンダーには、おそらく固有の制約と責任があるということは重要な事実であり、必ず理解する必要があります。LogMinerを使用することで、サービス・レベル（SLA）とKPIに関して一定レベルの統一性や一般的な期待が保証されるわけではありません。

たとえば、2021年の夏にLogMinerを使用してCDCのサポートを開始したあるサード・パーティ・ベンダーには、次のような制約がありました。

- 100 GBを超える表はバックフィルできない。
- オラクルのマルチテナント・アーキテクチャ（CDB/PDB）がサポートされない。
- Oracle Autonomous Databaseがサポートされない。
- 主キーがない表からのイベントには、コンシューマ側で更新を実行するために必要な情報が含まれない。
- イベント・サイズが3 MBに制限される。
- 索引構成表（IOT）がサポートされない。
- BFILEタイプの列では、ファイルへのパスのみがレプリケートされ、ファイルのコンテンツはレプリケートされない。
- データタイプがANYDATA、BLOB、CLOB、LONG/LONG RAW、NCLOB、UDT、UROWID、XMLTYPEの列はサポートされず、NULL値に置き換えられる。
- Oracle 11gでは、データタイプがANYDATAまたはUDTの列がある表はサポートされず、表全体がレプリケートされない。
- Oracle Label Security（OLS）がレプリケートされない。
- スキーマが変更された場合、古いスキーマが引き続き適用され、一方で新しいスキーマの一部のイベントが読み取られる場合がある。
- ソース・スキーマへのすべての変更を自動的に検出できるわけではなく、そのような場合、データ破損が発生する場合がある。以下のスキーマ変更により、データ破損が引き起こされる場合がある。または下流イベントの処理が失敗する場合がある。
 - 列のドロップ
 - 表の中央への列の追加
 - 列のデータタイプの変更
 - 列の順序の変更
 - 表のドロップ（新たなデータを追加して同じ表を再作成する場合も該当）
 - 表の切捨て
- ビューのレプリケーションがサポートされない。
- ストリームの実行中に作成されたマテリアライズド・ビューが自動的にバックフィルされない。

LogMinerは、引き続き診断のユースケースに重点を置きます。LogMinerを使用するサード・パーティ製CDCツールを選択するお客様は、そのフレームワークの既知の制約を踏まえて、慎重に取組みを進める必要があります。トランザクションが1秒あたり数百を超えるオンライン・ワークロードのユースケースでは、待機時間、メモリ、スケーラビリティの問題が発生する可能性があります。データ量の少ない小規模なデータベースでは、LogMinerは単純なワークロードに“十分に対応できる”かもしれませんが。しかしながら、多くの場合、LogMinerはさらなるサード・パーティ製ツールと組み合わせて使用され、そのようなサード・パーティ製ツールが、ソース・データベースでライフサイクル問題や不必要な負荷をさらに引き起こす可能性があります。

LogMinerを使用したサード・パーティ製CDCツールの利用者はすべて、ソース・データベース自体への影響とソース・データベースに課される負荷を慎重に評価し、サポートされるデータタイプの範囲、機能上の制約、およびセキュリティの懸念事項がビジネスに支障をきたさないことを確認する必要があります。

XStreams APIを使用した変更データ

XStreamは、Oracle Databaseの戦略的要素であり、トランザクション・ストリームをデータベースの“内部”または“外部”にレプリケートできる低レベルAPIを提供します。XStreamの目的は、論理変更レコード（LCRには、ソースからターゲットに送信されるDMLおよび/またはDDLが含まれます）を暗示的または明示的なストリームに挿入する機能やLCRを抽出する機能をクライアント・アプリケーションに提供することです。XStream APIは、パワーユーザーや技術パートナー、およびOracle Databaseを直接使用する高パフォーマンスのCDC、ストリーム、レプリケーションを備えたサード・パーティ製ツールを支援することを目的としています。

XStreamは、*XStream Out*と*XStream In*という2つの主要機能で構成されています。XStream Outでは、Oracleデータベースに加えられたデータ変更を他のシステムと共有できるようにするOracle DatabaseのコンポーネントとAPIが提供されます。XStream Outは、REDOログからデータ操作言語（DML）とデータ定義言語（DDL）の両方の変更を取得し、それらの変更をAPIを使用するクライアント・アプリケーションに送信できます（ドキュメント：https://docs.oracle.com/cd/F19136_01/xstrm/introduction-to-xstream.html#GUID-5939CB6C-8BA9-4594-8F96-B0453D246722）。

一般的に、XStream APIは、Oracle Databaseの内部および外部にレプリケートする高速トランザクション向けの非常に強力な最新のアクセス・レイヤーです。ただし、認識しておくべき重要な事項がいくつかあります。

XStreamに関する重要な事実

- XStream APIを使用するには、XStreamを使用するすべてのデータベースでGoldenGateの**ライセンス**が必要です。
- XStreamを使用するツールの**自動化とライフサイクル**は保証されません。XStreamは低レベルAPIのため、それよりも高レベルのツール（サード・パーティ製CDCツールなど）は、ユーザー、接続、ディクショナリ、トランザクション、およびサプリメンタル・ロギングやチェックポイント・セマンティックと連携するには、独自のソフトウェア自動化機能とライフサイクル管理機能を実装する必要があります。オラクル以外のベンダーが提供する個々のツールは、品質と機能が一律ではない可能性があります。
- XStream APIとともに使用する任意のデータベース向けの**リカバリ・セマンティック**。どのようなソース・データベースやターゲット・データベースも、さまざまな理由で障害が発生する可能性があるため、データベースに接続され、XStreamを使用するサード・パーティ・クライアントは、再起動を正常に処理するための独自のロジックを実装する必要があります。このチェックポイントとリカバリが正しく行われていないと、サード・パーティ製CDCツールは一部のレコードを省略したり、レコードを重複して書き込んだり、そうでない場合はレプリケートされたデータ・ストアで一貫性のないデータセットを生成したりする可能性があります。このような状態は、Oracleデータベースから非リレーショナルのターゲット（Apache Kafkaやオブジェクト・ストレージなど）にトランザクションをレプリケートする場合でも重要です。
- **サード・パーティ製ツールの一貫性**はXStreamによって保証されません。サード・パーティ製ツールはXStream APIにとって単なるクライアントにすぎないため、高品質のCDCを保証する魔法のような方法は存在しません。実装方法はベンダーによって異なり、各オープンソース・フレームワークがXStream上に実装される方法も全く異なる可能性があります。ベンダーのXStreamの使用は一律ではないという問題に注意してください。Oracle Databaseのお客様が、XStreamを使用するさまざまなCDCツールを実装している場合でさえも、それぞれの機能、品質、結果はツールごとに異なります。

Oracle GoldenGateを使用した変更データ

Oracle GoldenGateは、高度に最適化された独自のネイティブAPIを利用して、Oracle Databaseと緊密に統合されます。GoldenGate APIはデータベースのコア内に深く統合されているため、高速データ・イベント用に最適化され、データベースの新機

能を利用可能になった時点で活用できます。GoldenGateは、オラクルのMaximum Availability Architecture (MAA) プラチナ層に対して認定された唯一のレプリケーション・フレームワークです (ドキュメント:

<https://www.oracle.com/a/tech/docs/maa-goldengate-hub.pdf>)。

上述のOracle MAA技術資料に記載されているように、GoldenGateは通常は‘ハブ’構成で実行されます。これは、単一ポイントのプロジェクトの場合や、GoldenGateコンポーネントの管理を簡素化する際に有益です。

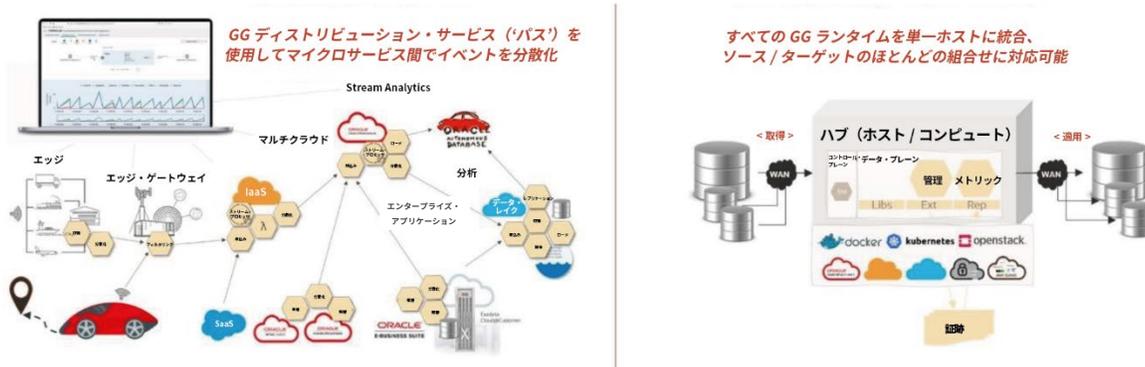


図6：Oracle GoldenGateはハブとして、またはマイクロサービス・ベースのデータ・メッシュとしてデプロイされる

Oracle GoldenGateは、大規模企業、分散環境、またはマルチクラウドのレプリケーション・プロジェクトで、メッシュとしても実行できます。データ・メッシュのアプローチを採用すると、GoldenGateディストリビューション・サービスを使用して、マイクロサービス・レイヤー/アプリケーション間でイベントを分散できます。

Oracle GoldenGateの一般的な優位性

- Oracle GoldenGateは、Oracle Database（ロギング、REDO、バックアップおよびリカバリ・レイヤー、セキュリティ、管理機能、Oracle Data Guard、Oracle Database Vault、Oracle Autonomous Databaseなど）と深く統合されています。
- Oracle GoldenGateは、もっとも高い安定性、パフォーマンス、信頼性を実現するオラクルのCDC/レプリケーション・ソリューションです。
- 高パフォーマンス：Oracle GoldenGateは、データ移動の待機時間を1秒未満に抑えることができ、トランザクション・データの取得、ルーティング、変換、送信の負荷が低いのが特徴です。
- GoldenGateは、中核となるOracle Databaseの新規および新興のイノベーション（セキュリティ、機能、データタイプなど）に対応する唯一のCDC/レプリケーション・ツールです。
- GoldenGateは、従来型のOLTPデータ・レプリケーションや高可用性（一方向、双方向、Peer-to-Peerなど）から、データ・レイク取込み、マルチクラウド取込み、SaaSアプリケーション・レプリケーション、メッセージング・レプリケーションに至るまでのさまざまなユースケースに対応しています。過去5年間でGoldenGateは、データ・パイプライン、データ変換、時系列分析、予測分析、地理空間、リアルタイム分析などのユースケースで、ストリーム処理パターンへと拡張されています。
- Oracle Databaseからの取得の場合、特にOracle Real Application Clusters（Oracle RAC）やOracle Automated Storage Management（Oracle ASM）などのオプションを扱う場合、Oracle GoldenGateはより容易に使用できます。

さらにGoldenGateは、Oracle Exadata、Oracle Exadata Cloud Service、およびOracle Exadata Cloud at Customerに対して認定された唯一のCDC/レプリケーション・テクノロジーです。

- GoldenGateは、オラクルのMaximum Availability Architectureのプラチナ層に対して認定された唯一のレプリケーション・ツールであるため、もっとも厳格なRPO（リカバリ・ポイント目標）とRTO（リカバリ時間目標）のシナリオでテストされています。オラクルのお客様は、極めて価値の高い情報や機密情報をGoldenGateにゆだねています。

運用の視点から見ると、GoldenGateには、データベース管理者、システム管理者、およびDevOpsチームを支援するコマンドラインとWebのグラフィカル・ユーザー・インタフェースの両方が用意されています。GoldenGateマイクロサービスは以下の利点を提供します。

- シンプルな操作性のWebインタフェース
- 高速かつ容易なインストールまたはプロビジョニング
- どこでも実行できるAdminClient経由でアクセス可能なコマンドライン
- DevOps向けのRESTful API
- 組み込みの監視機能の提供（ただし、独自の監視機能にも容易に適応可能）

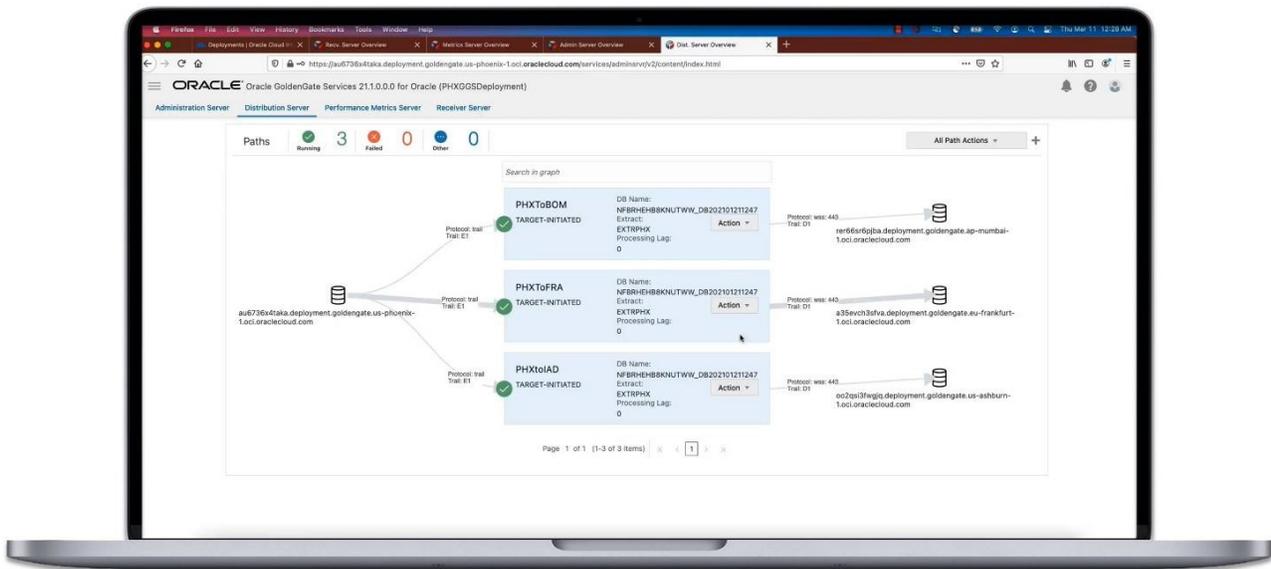


図7：Oracle GoldenGateのシンプルで使いやすいユーザー・インタフェース（オンプレミスまたはクラウドネイティブ）

銀行および金融機関のお客様は、極めて厳格なセキュリティ・フレームワークを常に必要としてきました。GoldenGateは以下のセキュリティ機能を提供します。

- ディスク上およびワイヤ・プロトコルでデータを暗号化できます。
- 暗号化プロファイルにより、Oracle Key Vaultからマスター鍵を取得するために使用される構成情報がカプセル化されます。
- DMZ環境とプライベート・クラウド環境をネイティブにサポートします。
- すべてのバージョンのデータベースでオラクルの透過的データ暗号化（TDE）とオラクルの表領域暗号化（TSE）をサポートします（専用のディスク・リーダーを使用するサード・パーティ・ベンダーはこれらをサポートできません）。
- あらゆる分散サービスで豊富な証明書をサポートします（mTLS 1.3のサポートを含みます）。
- APIは、クロス・サイト・リクエスト・フォージェリ（CSRF）認証によって保護されます。CSRFトークンベースの保護の実装はデフォルトの構成です。

サード・パーティ製ディスク・ログ・リーダーに関する注意事項

ここ何年かの間で、一部のベンダーは、Oracle Databaseログを直接リバース・エンジニアリングしてデータ変更を検索する専用のソフトウェアを作成することを選択しました。この戦略を選択したベンダーは、LogMinerやXStreamといったサポート対象のAPIやユーティリティを回避してきました。このような戦略により、データ、データ・セキュリティ、および継続的なサポートがリスクにさらされています。

これらの専用ログ・パーサーには、以下のような欠点があります。

1. **サポートされておらず、文書化されていない**：オラクルは、REDOログ/アーカイブ・ログのコンテンツや構造を文書化していません。これらのログをリバース・エンジニアリングするサード・パーティ製ツールの使用はサポートされておらず、Oracle Master Agreement (pt.8) およびOracle Software License Agreement (pt.9) に違反します。
2. **著しい機能制限**：これらのリーダーは、リバース・エンジニアリングのディスク・ログであるため、データベース・エンジンの実行内容のうち、ほんの一部のセマンティック（意味）を参照しているにすぎません。そのため、Exadata、データ圧縮（EHCC）、表の暗号化、セキュリティ鍵へのアクセス、データタイプの不一致などの非互換性を数多く引き起こします。バイナリ・ログ・リーダーを使用するベンダーは、データを処理するために、データベースでこれらの機能を無効にすることをお客様に強いることとなります。
3. **ガバナンスとセキュリティに関するエンドユーザーのリスク**：データベース・ログにホストが直接アクセスする必要があり、ユーザーによって提供された暗号化鍵を使用できる管理ツールは、SOC2、Basel、BCBS239をはじめとする数多くのポリシーを遵守する上で、著しいリスクをもたらします。
4. **ライフサイクルの統合の欠如**：ミッション・クリティカルなシステムで、サード・パーティのバイナリ・ログ・パーサーを使用してフェイルオーバーと再起動を自動化するOracle MAAの構成がサポートまたは保証されません。
5. **計画外停止のリスク**：これはサポート対象外の統合のため、オラクルはログの形式、暗号化、鍵のボルトなどを任意の時点で変更する場合があります。その結果、データベースのパッチが適用されるときに計画外停止や複雑なロールバックが発生する可能性があります。
6. **クラウド・データベースの使用に関する問題**：オラクルのマネージド・クラウド・データベース（Oracle Database Cloud Service、Oracle Autonomous Database、Oracle Exadata Cloud Service、Oracle Exadata Cloud at Customerなど）がサポートされない場合があります。特に、オラクルのマネージド・クラウド・ホストからのサード・パーティ製バイナリ・ログ・パーサーのインストールはサポートされません。さらに、Exadataのクラウド・システムには、暗号化された表が必要であり、Oracle Database 19.1以降では、バイナリ・ログ・パーサーはこの暗号化された表を読み取ることができません。
7. **平凡なパフォーマンスとスケーラビリティ**：GoldenGateのOracleデータベース・エンジンへのネイティブ・アクセスとは異なり、REDOログへのディスクベースの直接アクセスでは、並列処理と計算能力の使用において、不必要なI/O問題と制約が発生します。ネイティブに統合されたGoldenGate Extractを使用した場合、実際のワークロードは2~5倍高速に実行され、3~10倍のデータ量を処理できます。これらの問題は、Oracle ASMストレージにリモートからアクセスした場合はさらに悪化します。
8. **所有権と制限事項**：Oracle Master Agreementでは、ソフトウェアのデータ構造をリバース・エンジニアリングすること、またはそれを許可することを明示的に禁じています。次のリンク先でOnline Transactional Oracle Master Agreementの3.4項をご覧ください：<https://www.oracle.com/a/ocom/docs/lic-online-toma-us-eng-v091120.pdf>
9. **データベース・ソフトウェアのライセンス**：Oracle Software License Agreementでも、ソフトウェアのデータ構造をリバース・エンジニアリングすること、またはそれを許可することを明示的に禁じています。次のリンク先でD項の3つ目の箇条書きをご覧ください：

<https://www.oracle.com/us/corporate/contracts/olsa-services/olsa-renewals-en-us-v053012-1867431.pdf>

Apache Kafkaまたはオブジェクト・ストレージでのGoldenGateの使用について

GoldenGateは、Apache Kafkaおよびオブジェクト・ストアとの間でデータ・レプリケーションを行うための素晴らしいプラットフォームです。

Oracle GoldenGateは、Kafkaをサポートする初めてのCDC/レプリケーション・ツールです。Kafkaを考案したLinkedInチームは、長年にわたりKafkaの実装でGoldenGateを十分に活用してきました。世界でもっとも大規模な一部のKafka実装では、毎日GoldenGateに依存しながら、ペタバイト規模の変更データをイベント・ストリームに取り込んでいます。GoldenGateのお客様は、ソリューションが、世界でもっとも要求の厳しいミッション・クリティカルなオラクルからKafkaへのユースケースにまでスケールリングされることを確信できます。



図8：CDC、Oracle Data Integration、Stream AnalyticsでのGoldenGateの使用

大まかな目的は、データベース・イベントをソース・データベースからKafkaにストリームすることです。ただしこれを行うには、トランザクションの整合性を保つためにトランザクションの境界線（DMLおよびDDL）を保存することが不可欠です。これが重要なのは、Kafkaなどのメッセージ・システムには、ACID（原子性、一貫性、整合性、永続性）プロパティがないためです。GoldenGateのターゲットがACIDに対応したストア（リレーショナル・データベース、Apache Hive、MongoDBなど）の場合、データ・イベントは、ターゲット・データ・ストアに書き込まれ、マージされるときに、確実に正しく処理されます。そのため、ACID対応のデータ・ストアでは、信頼性のある‘同期されたデータ’がもたらす大きな利点をユーザーは享受できます。

GoldenGateがデータ・イベントを非ACIDストア（Apache Kafka、オブジェクト・ストレージ、検索ベースの索引など）に書き込んでいる場合、GoldenGateは通常、ペイロードをさらなるメタデータで‘飾る’という追加の措置を講じる必要が出てきます。そうすることで、データ・イベントのコンシューマは、ロールバック操作の実行が必要な場合に、データ・イベントを適切に再構築し、スキーマの変更を記録し、リーダの位置を解読できるようになります。

たとえば、GoldenGateユーザーは通常、以下のような種類のメタデータをデータ・イベントで確認したいと考えています。

- イメージの前および後のトランザクション
- ソースの操作の種類（挿入、更新、削除、切捨て、作成、変更、ドロップ）
ソースのコミット順序番号（SCN、LSN、RBAなど）
- ソースのコミット・タイムスタンプ
- ソースのメタデータ（表や列の定義など）
- GoldenGateターゲットの証跡順序番号

- GoldenGateターゲットの証跡RBA
- メタデータのバージョン
- ソース・データベースの名前と種類
- ソースの操作のログ位置（利用可能な場合）
- ソースのトランザクションのユーザーと名前（利用可能な場合）

GoldenGateは、トランザクションをデータ・ストアとプラットフォームの数百もの異なる組合せに送信できます。特にKafkaでは、GoldenGateの出力形式として、CSV、固定長、XML、JSON、Avro、Parquet、ORCといったさまざまな非リレーショナル・ペイロードがサポートされます。高度なファイル・タイプに、高レベルなスキーマの進化やリレーショナルに類似した他の属性（KafkaレジストリのAvro、メタストア・カタログのParquetなど）を組み込む試みがますます進んでいます。しかしながら、これらのペイロード形式はいずれも、ACIDに対応したDBの豊富な機能とセマンティックに対応できるとは言えません。

結局のところ、非ACIDターゲットでは、開発者には独立性と一貫性をつなぎ合わせる以外に道はありませんが、これを簡素化する上で、GoldenGateメタデータが大きな役割を果たします。Stream AnalyticsでGoldenGateを選択すれば、GoldenGateメタデータがデフォルトですぐにサポートされる組込みパターンを利用できます（部分的または完全なサプリメント・ロギングなど）。これにより、ストリーム処理で一貫性と独立性を実現するベスト・プラクティスがさらに簡素化され、'作り込まれます'。このような包括的なストリーム処理とストリーム分析のユースケースは、動的でリアルタイムの最新データ・プラットフォームにとってますます重要になっています。

サマリー：Oracle GoldenGateはOracle Databaseに最適なCDC

オラクルのエンジニアリング・チームがOracle Databaseに対する総合的に最適なCDCおよびレプリケーション・ソリューションも提供できる非常に優れた能力を有していることは、驚くべきことではありません。GoldenGateは、もっとも完全な機能、高いパフォーマンス、高い信頼性を備えており、Oracle Databaseのほか、数百ものデータ・プラットフォームにも対応した信頼できるデータ統合ソリューションであることが、お客様とアナリストによって幾度となく認められています。

本書では、Oracleデータベースと連携するサード・パーティのCDCテクノロジーにおける数多くの制限事項と考慮事項を説明しました。以下の表に、これらのCDCアプローチをまとめています。

	GoldenGate	XStream	LogMiner	トリガー	ディスク・リーダー
[最適] ●●●○ [不適]					
開発者のエクスペリエンス					
ローコード開発	●	○	○	○	サード・パーティ
マイクロサービスAPIとメッシュのデプロイメント	●	○	○	○	○
組込みの監視機能とメトリック	●	○	○	○	サード・パーティ
柔軟なネイティブ・クラウド・サービス	●	Oracle ADB	Oracle ADB	Oracle ADB	○
Oracle Database に対する堅固性					
もっとも高いパフォーマンス（取得とレプリケーション）	●	●	○	○	ストレージに結合
リモート取込み/適用	●	●	●	●	●
データタイプのサポート	●	●	○	●	○
データベースのHA/リカバリ機能	●	●	●	○	○
Oracle MAA に対する認定	●	○	○	○	○
マルチテナント、コンテナ、プラグラブルDBのサポート	●	●	○	●	○
オラクルの圧縮機能のサポート	●	●	●	●	○
セキュリティ（クラウドDB、Oracle TDE、ウォレット、ポルトなど）	●	●	●	●	古いDBのみ
異種データ・エコシステム					
非 Oracle Database	●	○	○	○	リーダーごとに異なる
メインフレーム（DB2/z、Tandem など）	●	○	○	○	リーダーごとに異なる
ビッグ・データ、NoSQL、Kafka など	●	○	○	○	リーダーごとに異なる
マルチクラウド（ホスティングおよびデータ・ストアのサポート）	●	○	○	○	リーダーごとに異なる
その他					
サポート対象（オラクルがSRの発行を受け入れる）	対象	対象	対象	対象	非対象
ライセンス	GoldenGate	GoldenGate	Oracle DB	Oracle DB	ライセンス許可なし

図9：Oracle GoldenGateと代替のチェンジ・キャプチャAPIおよび各アプローチとの比較

GoldenGateに匹敵するサード・パーティのテクノロジーはありません。

この分析が、ビジネスでの使用に適したレプリケーション・ソリューションを選択し、チェンジ・データ・キャプチャのストリーム要件に対処する手助けになることを願っています。

Connect with us

+1.800.ORACLE1までご連絡いただくか、[oracle.com](https://www.oracle.com)をご覧ください。北米以外の地域では、[oracle.com/contact](https://www.oracle.com/contact)で最寄りの営業所をご確認いただけます。

 blogs.oracle.com/mysql

 facebook.com/mysql

 twitter.com/mysql

著者：Oracle Product Development、Ionut Bruma

Copyright © 2021, Oracle and/or its affiliates. All rights reserved. 本文書は情報提供のみを目的として提供されており、ここに記載されている内容は予告なく変更されることがあります。本文書は、その内容に誤りがないことを保証するものではなく、また、口頭による明示的保証や法律による黙示的保証を含め、商品性ないし特定目的適合性に関する黙示的保証および条件などのいかなる保証および条件も提供するものではありません。オラクルは本文書に関するいかなる法的責任も明確に否認し、本文書によって直接的または間接的に確立される契約義務はないものとします。本文書はオラクルの書面による許可を前もって得ることなく、いかなる目的のためにも、電子または印刷を含むいかなる形式や手段によっても再作成または送信することはできません。

本デバイスは、連邦通信委員会のルールに基づいた認可を未取得です。認可を受けるまでは、このデバイスの販売またはリースを提案することも、このデバイスを販売またはリースすることもありません。

OracleおよびJavaはOracleおよびその子会社、関連会社の登録商標です。その他の名称はそれぞれの子会社の商標です。

IntelおよびIntel XeonはIntel Corporationの商標または登録商標です。すべてのSPARC商標はライセンスに基づいて使用されるSPARC International, Inc.の商標または登録商標です。AMD、Opteron、AMDロゴおよびAMD Opteronロゴは、Advanced Micro Devicesの商標または登録商標です。UNIXは、The Open Groupの登録商標です。0120

免責事項：データシートにこの免責事項の記載が必要かどうか分からない場合は、収益認識方針を参照してください。ホワイト・ペーパーの内容と免責事項の要件についてさらに質問がある場合は、REVREC_US@oracle.com宛てに電子メールでご連絡ください。