

Oracle GoldenGate Veridataを 使用したデータ整合性の確保

2020年11月

Copyright © 2020, Oracle and/or its affiliates

公開

本書の目的

このドキュメントは、もっとも効率的な方法で内容を紹介することを目的としているため、多数の重要かつ複雑な概念について簡単に触れ、それぞれについての詳細な説明は省きます。読者が製品をよく理解し、高可用性のOracle GoldenGate Veridata環境を適切に設計、実装できるように支援することがこのドキュメントの目的です。したがって、実装を適切に行う上で非常に重要な設計、ユニット・テスト、統合テストのアクティビティは、このガイドでは意図的に省かれている点に留意してください。すべての例は現状のまま提供されます。カスタマイズされた実装には、オラクルのコンサルティング・サービスが強く推奨されます。また、具体的な本番環境のドキュメントを確認する必要があります。

免責事項

本文書には、ソフトウェアや印刷物など、いかなる形式のものも含め、オラクルの独占的な所有物である占有情報が含まれます。この機密文書へのアクセスと使用は、締結および遵守に同意したOracle Software License and Service Agreementの諸条件に従うものとします。本文書と本文書に含まれる情報は、オラクルの事前の書面による同意なしに、公開、複製、再作成、またはオラクルの外部に配布することはできません。本文書は、ライセンス契約の一部ではありません。また、オラクル、オラクルの子会社または関連会社との契約に組み込むことはできません。

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下さい。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
概要	4
データ整合性維持の課題	4
移行のエラー	4
クラウドへのワークロードのリフト・アンド・シフト	5
ソースとターゲット間の相違	5
インスタンス化のエラー	5
構成のエラー	5
レプリケーションでの欠落	5
レプリケーションの待機時間	5
ユーザー・エラー	5
アプリケーション・エラー	5
データ整合性を管理するための要件	6
Oracle GoldenGate Veridata	6
Oracle GoldenGate Veridataの仕組み	7
Oracle GoldenGate Veridataのアーキテクチャ	8
Veridataサーバー	8
Vericom	8
Veridataエージェント	8
Veridata Webユーザー・インタフェース	9
顧客事例	9
課題	9
ソリューション	9
利点	10
結論	10



概要

高いデータ可用性に対する要求と、パフォーマンス低下やサービス中断を招くことなく、24時間365日のデータ・アクセスまたはそれに近いアクセスの実現を求める要求により、冗長化した分散データのコピーを持つ必要性が生じてきました。しかし、データが急増する現在の複雑なIT環境において、分散されたデータのコピー間でデータ整合性を維持することは困難であり、残念ながら、データの不一致が発生する可能性があるのが現実です。速やかに正しくないデータを見つけて対処しない限り、誤った意思決定や品質保証契約の未達成につながり、最終的には業務上、財務上、法的なリスクがもたらされます。

は、バックアップ・システム、レポート・データベース、アクティブ-アクティブ・データベース、その他の冗長化されたデータ・システムでデータ整合性を確保する際の高い信頼性を企業に提供します。本書の全体を通じて、これらのシステムをターゲット・データベースと呼びます。Veridataアプリケーションは、ソース・データベースとターゲット・データベース間で、どちらのシステムもオフラインにすることなく、必要に応じた頻度で定期的なチェックを行います。

Oracle GoldenGate Veridataは強力でありながら使いやすいソリューションを提供して、同期が取れていないデータがビジネスに悪影響を与える前にこれらを特定します。Oracle GoldenGate VeridataはOracle GoldenGateのリアルタイム・データ・レプリケーション製品と併せて導入するか、単独で導入することで、データベース間のデータ整合性を維持します。

データ整合性維持の課題

データベース間のデータ整合性の管理に役立つソリューションの要件について考察する前に、企業においてデータの不一致が発生する一般的な原因を理解する必要があります。

データの不一致が発生するのは、ターゲット・データベースがソース・データベースから逸脱している場合です。データが逸脱する程度はさまざまな要因によって異なり、意図的なものもあれば、意図的でないものもあります。

Oracle GoldenGateなどのデータを確実にレプリケーションする製品を使用した場合でも、潜在的にデータの不一致を引き起こす原因が残っています。ターゲット・データベースの目標がソース・データベースと一貫性を厳密に持つことである場合、IT部門は結果を保証するプロセスとポリシーを導入する必要があります。

次の項では、潜在的なデータ不一致の原因のいくつかについて説明します。

移行のエラー

レプリケーションを開始する前のターゲット・データベースの初期ロードの促進には、さまざまな種類の移行ツールが使用されます。

移行ツールとレプリケーション製品でデータ処理の構成が異なると、データの不一致につながる場合があります。

たとえば、列の値が未知の場合、移行ツールが '?' を使用する一方で、レプリケーション製品は 'null' を使用場合があります。また、移行の実行時に確認できないオープン・トランザクションがある可能性もあり、ターゲット・データベース上のデータ欠落の原因となります。

クラウドへのワークロードのリフト・アンド・シフト

世界中でクラウドが普及するにつれ、オンプレミスからクラウドへのデータベース・ワークロードのリフト・アンド・シフトが、現在のIT分野における必須事項となっています。Oracle GoldenGateは、このワークロードの移行を支援します。オンプレミスとクラウド・データセンターの間でのデータ整合性とデータ・コンプライアンスを確保することが、ハイブリッド・クラウド・シナリオでの最大の課題です。

ソースとターゲット間の相違

エンコードやロケール、エンディアンネス、データベース・バージョンなどの構成がソース・データベースとターゲット・データベース間で異なる場合、移行中やレプリケーション中に微妙な相違が発生する可能性があります。たとえば、互換性のないキャラクタ・セットや日付/時間の形式と範囲により、エラーがターゲット・データベースに紛れ込む場合があります。

インスタンス化のエラー

移行またはレプリケーションを開始する前に、正しいスキーマと制約を使用してターゲット・データベースをインスタンス化する必要があります。この処理に失敗すると、ソースとターゲットの同期が取れていない状態になります。

一例としては、主キーや一意キーの設定に失敗すると行の重複が発生します。ソース・データベースに重複行がない場合でも、移行ジョブがエラーなしで終了したという保証はないため、重複行が作成される可能性があります。その他のインスタンス化のエラーには、ジョブやスクリプト、トリガーの不適切な移行があります。そのためデータが誤って変更される場合があります。

構成のエラー

レプリケーション製品の不適切な構成や意図せぬ構成が不一致を引き起こすこともあります。この種の不一致は、レプリケーション製品から見ると構成どおりであるため、レプリケーション・ログには表示されません。このため、QAテストでのこの問題の検出ができない場合があります。

たとえば、レプリケーションの観点から見て問題のない構成でも、データ品質の観点からは容認できないものがあります。例を挙げると、Oracle GoldenGateを使用する場合、DBAはパラメータ'REPLACEBADCHAR'を使用できます。このパラメータは印刷不可能な文字を、より適切な文字で置き換えます。これによりレプリケーションを問題なく続行できますが、適切な承認を得ていない限り、企業のデータ品質上は容認できない可能性があります。

もう1つの例として、Oracle GoldenGateは、ターゲット・データベースにトランザクションを適用するときに、トリガーとカスケード削除を許可しないオプションを提供しています。これらのオプションを使用しないか、または不適切に使用することでも、無効なデータが生成される可能性があります。

レプリケーションでの欠落

ソース・データベースとターゲット・データベース間でレプリケーションが有効化されており、正常に機能していても、ソースに挿入されたデータがレプリケートされないケースがあります。たとえば、データを一括挿入する場合に、通常データベース内で使用されるオプション（例：OracleデータベースのNOLOGGING）は、レプリケーション・システムによるこのデータの取得を妨げます。

レプリケーションの待機時間

非同期レプリケーションの場合、ソース・データベースが変更され、これらの変更がターゲットに配信されるまでの間に短い遅延が発生します。ただし、最大待機時間要件を満たさない場合、品質保証契約レベルやデータ・コンプライアンス要件の違反につながる可能性があります。

ユーザー・エラー

ターゲット・データベースは、ソース・データベースの間合せ処理をオフロードするために作成されることが多いですが、こうすることで、ソース・データベース上で稼働するアプリケーションに影響を与えることなく、さまざまな運用レポートを作成できます。

しかし、使用するテクノロジーによっては、ターゲット・データベースの読取りだけでなく書込みが可能になっている場合があります。この場合、ITポリシーが適用されていても、ユーザーやDBAが誤って、または悪意を持ってデータを変更することができます。

アプリケーション・エラー

ターゲット・データベースを使用するアプリケーションが、不完全なロジックやアプリケーション・アップグレードによってデータを変更するおそれがあります。また、現在レプリケーションが稼働しており、データの整合性が取れていても、ターゲット・データベースを使用する新しいアプリケーションが開発され、将来のある時点でデータが変更されるかもしれません。

データ整合性を管理するための要件

企業全体で一貫したデータを管理するための課題について十分に理解したところで、これらの課題に対処するだけでなく、IT組織にシームレスに適応するためのおもな要件について考察しましょう。データベース間でデータ整合性を管理するためのテクノロジー要件は次のとおりです。

- 高速で影響の少ないデータ比較
- 異種データベースのサポート
- 大量データの処理機能
- データ比較を管理するための柔軟なオプション
- 常にデータが変更される実行中データベースのサポート
- 最小限の負担
- 連続レプリケーションでの変更データのみと比較
- 自動および手動パーティション化による大規模表の比較
- 監査目的のデータ比較レポート
- ソース・システムとターゲット・システムの停止時間なし
- データの非一貫性を識別する機能
- ハードウェア・リソースとネットワーク・リソースに対する少ない影響
- 一貫性のないデータに関する詳細で実践的なレポート
- 各種ロールおよびアクセス・レベルに対応する柔軟なレポート
- データ・セキュリティ
- 使用、理解、構成、デプロイ、診断における容易さ

Oracle GoldenGate Veridata

Oracle GoldenGate Veridataは、負担を最小限に抑えた、高パフォーマンスの使いやすい製品であり、企業全体でのデータ整合性の管理を支援します。この製品は、大量データを含むデータベースの比較や常にデータが変化するデータベースの比較など、データ整合性を確保するさまざまなシナリオで使用できます。

Oracle GoldenGate Veridataは、ユーザー・インタフェースとさまざまなコマンドライン・ユーティリティを介した容易な構成を実現し、柔軟な比較オプションを提供して、異種データベースをサポートします。Javaテクノロジーをベースに構築されたOracle GoldenGate Veridataは、データ使用時のグローバリゼーションとセキュリティのサポートを提供します。

Oracle GoldenGate Veridataは、Oracle GoldenGateやその他の同等製品とシームレスに連携してその機能を補完し、データ整合性を確保します。最新の製品情報については、Oracle GoldenGate VeridataのWebサイトを参照してください。

Oracle GoldenGate Veridataの仕組み

Veridataのアーキテクチャ、インストール、構成の詳細に進む前に、Veridataの仕組みについて簡単に確認しましょう。

オール・オア・ナッシングのアプローチとは異なり、Veridataは比較対象を指定し、比較方法を柔軟に選択できる機能を提供するため、関連するデータのみを処理し、関係のある相違（同期していない行とも呼ぶ）のみを選択できます。

初期の比較（または行ハッシュ）ステップでは、問合せを使用してソース・テーブルとターゲット・テーブルから行が取得されます。ソース・データベースとターゲット・データベースの種類が異なる場合、正確な比較を行うため、列から標準化されたデータ型形式への変換が実行されます。

Veridataはデフォルトで、すべての主キー列を値ごとに比較し、キーでないすべての列にはハッシュ値を使用することで、行を比較します。ハッシュ値の計算に使用される一意のデジタル署名は、比較用にネットワーク経由で転送されるデータのサイズを縮小しながら、2つの行が同じ列値を含むか、別の列値を含むかを決定するための、完全ではないものの信頼性の高い効率的なメカニズムを提供します。

完全な確実性をもって同期が取れていない行を検出するには、キー以外の列でハッシュを使用せずに列ごとに比較するように、Oracle GoldenGate Veridataを構成します。すべての列を比較すると、列の数に比例して処理パフォーマンスが低下し、ネットワーク使用量も増加します。この方法はベスト・プラクティスとしては推奨されていません。

ライブ・レプリケーション環境では、Veridataが初期比較を完了した後で、同期されていないと思われる行がmaybe out-of-sync (MOOS) キューに格納されます。この曖昧さが生じるのは、レプリケーションと比較が同時に実行されているため、ソースで取得された実行中トランザクションがターゲットにはまだ適用されていない結果として、相違が発生している可能性があるからです。

Veridataはconfirm-out-of-sync (COOS) ステップと呼ばれる確認ステップで、変化する環境での行のステータスを確認することで正確な結果を提供します。デフォルトで、確認プロセスは初期比較ステップと並列のスレッド内で実行されますが、各行の確認は、指定されたレプリケーション待機時間のしきい値に達してから実行されます。このステップの最後で、MOOS キューに含まれる行が次のように評価されます。

- 実行中 - 初期比較ステップの後に行が更新されたため、Veridataは行の同期が取れていないことを確認できませんでした。
- 同期 - Veridataは行が同期されていることを確認しました。
- 永続的に非同期 - Veridataは、初期比較ステップ後も引き続き、行の同期が取れていないことを確認しました。

ジョブが完了すると、Oracle GoldenGate VeridataのWebインターフェースを使用するか、直接ファイルを参照することで、比較レポートと非同期レポートを確認できます。

Oracle GoldenGate Veridataのアーキテクチャ

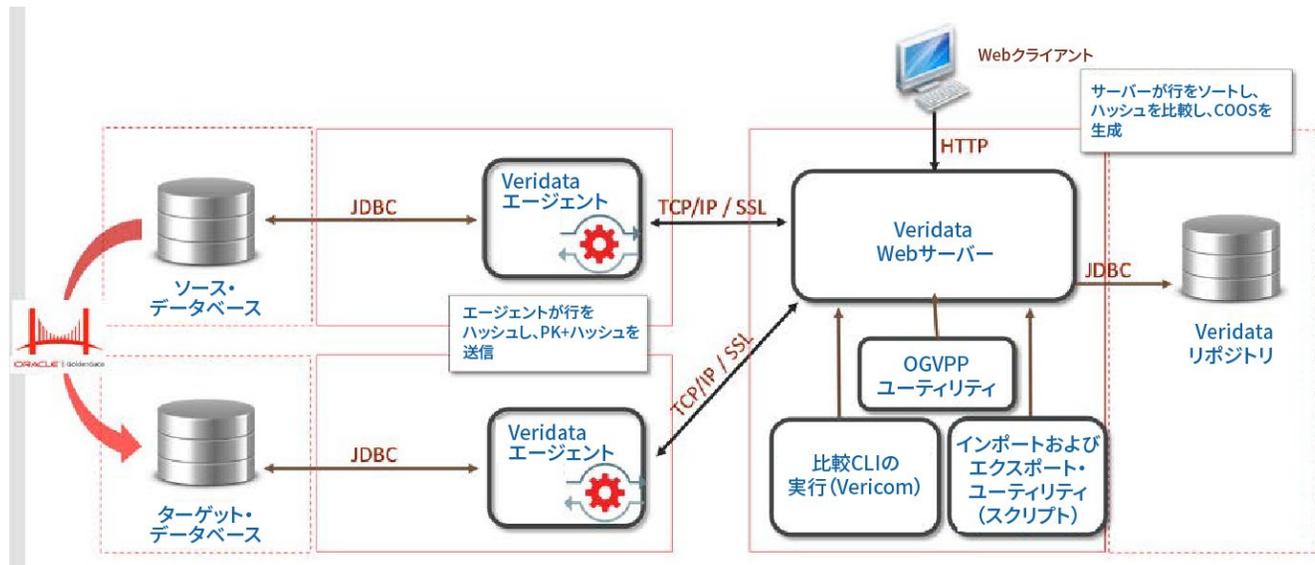


図1：GoldenGate Veridataのアーキテクチャ

上のアーキテクチャ図は、各種のVeridataコンポーネントに対する一般的なアーキテクチャ設定を示しています。矢印は通信が開始された方向を示していますが、通信はすべて双方向です。Veridataコンポーネントは、データベースをホストしているマシンと同じマシン上、または別のマシン上に導入できます。Veridataリポジトリには構成の詳細情報が格納されますが、セキュリティ上の目的から、ユーザー・データは格納されません。VeridataエージェントとVeridataサーバーの間の通信はSSL通信により保護されます。Veridataの各種ユーティリティは、データ比較ジョブの実行や、さまざまなリポジトリ間での構成のエクスポートおよびインポートに使用できます。

Veridataサーバー

VeridataサーバーのコンポーネントはWebサーバーと同じマシンに配置されるため、これらの1組をVeridataサーバーと呼びます。ユーザーはWebブラウザを使用してVeridataサーバーに接続し、Veridataジョブと関連オブジェクトを構成できます。この構成は、データベース内に存在するリポジトリに格納されます。必要なVeridataメタデータを構成すると、ユーザーは初期比較を開始してレポートを表示できます。

Vericom

Vericomはスタンドアロン・ユーティリティであり、Veridataジョブの実行に使用されます。Vericomはリポジトリの構成情報を使用して、Veridataジョブの詳細を把握します。Vericomを起動するには、コマンドラインから適切なオプションを指定します。Vericomは比較ジョブのスケジュール設定にも使用できます。VeridataサーバーとVericomは通常、ソース・データベースやターゲット・データベースとは異なるマシン上にインストールされます。

Veridataエージェント

Veridataエージェントは通常、ソース・データベースやターゲット・データベースのインスタンスが実行されるマシン上に配置されます。構成ファイルを使用して構成されるエージェントは、データベースから行をフェッチし、VeridataサーバーまたはVericomプロセスにこの行を転送します。Javaベースのエージェントのみがサポートされます。Cエージェントは、NonStopプラットフォームで稼働するNonStop SQL/MPデータベースとEnscribeデータベースに対してのみサポートされます。

Veridata Webユーザー・インタフェース

Veridata Webユーザー・インタフェースを使用すると、Veridataのアーティファクトを構成し、Veridataジョブを開始してレポートを表示できます。次のスクリーンショットに、Veridata Webユーザー・インタフェースを示します。

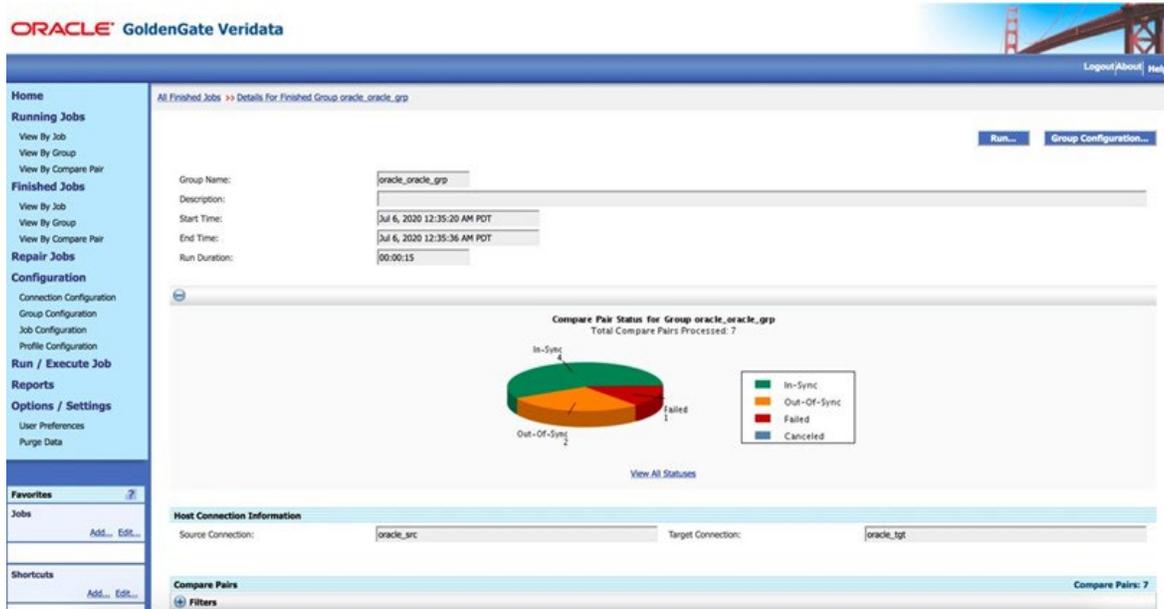


図2：GoldenGate Veridataのダッシュボード

顧客事例

Thomson Reutersは、企業や専門家向けにインテリジェント情報を提供する世界有数のプロバイダです。業界の専門知識と革新的なテクノロジーを組み合わせ、ニュース、法的報告書、分析などの情報を、金融、法律、税務、会計、科学、メディア市場の意思決定者に提供しています。

課題

Thomson Reutersは、同社を代表する法的リサーチ製品のデータ整合性確保を含む、データ管理に関するいくつかの課題に直面していました。これらの法的リサーチ製品は、マルチマスターのマルチデータベース構成で稼働しています。可用性と信頼性の高いサービスを提供することで、ユーザーの信頼を確保するという目標を持つThomson Reutersには、弁護士、法律事務員、法務書記、裁判官を含むシステム・ユーザーが、訴訟事例と判例に関して最新かつ正確な情報を確実に得られるようにする必要がありました。

また、一貫したユーザー・エクスペリエンスを提供し、サービス・レベルの低下に付随する収益損失を回避するため、停止時間を発生させることなく、各種のデータベース間でデータの不一致を識別し、再同期を行うソリューションを使用しなければなりませんでした。

Thomson ReutersはOracle GoldenGate Veridata 11gを導入したことで、データベースの可用性を妨げることなく、データの非同期を素早く自動的に特定できるようになりました。影響を受けるデータを特定するために、複雑な問合せを実行したり、多くの人的リソースを注ぎ込んだりする必要がなくなり、何週間もの労力を節約できたケースもありました。

このソリューションにより、使用しているOracle Database環境で非同期状態が発生した場合の停止時間の最小化と平均リカバリ時間の短縮が可能になりました。Oracle GoldenGate Veridataを通じて、サービス・レベルとユーザーの信頼の確保に欠かせないデータベースのオンライン状態と同期を維持するプロセスが容易になりました。Thomson ReutersのITチームは、Veridataの非同期XMLレポート・ファイルを解析し、同期の取れていない行に的を絞り、自動的にソースを修正するPL/SQLプロシージャを作成することで、Oracle GoldenGate Veridataの価値を高め、時間とリソースをさらに節約することに成功しました。

利点

Oracle GoldenGate Veridataを使用することで、Thomson Reutersはサイト間でデータを手動コピーし、長時間にわたってサイトを停止して別のサイトを利用した再インスタンス化を行う必要性がなくなりました。これにより、システム可用性が向上し、エラーのリスクが軽減され、顧客満足度が向上しました。

Thomson Reutersは、IBM DB2を含むサードパーティ・データベースからOracle Database 11gへの移行時に、Oracle GoldenGate Veridataを利用して潜在的なエラーとデータの欠落を効率的に検出することで、プロセスを合理化しながらデータ損失の可能性を回避しました。

拡張ソリューションの高度な同期機能のおかげで、データベースの再インスタンス化が不要になったため、6カ月にわたる移行期間中に数百時間もの停止時間が解消されました。この機能は、夜間処理中に7200万行以上のデータを処理します。

また、最後になりますが、このソリューションによって、事象解決後にデータベース・サービスを再開する前に、本番環境とテスト環境のデータベース状態をチェックする能力が向上しました。

結論

ハイブリッド・クラウドをサポートする現在の複雑なIT環境では、残念ながら、データの不一致が発生する可能性を避けられないのが現実です。一致しないデータを見つけて対処しない限り、誤った意思決定や品質保証契約の未達成につながり、最終的には業務上、財務上、法的なリスクがもたらされます。

Oracle GoldenGate Veridataは強力でありながら使いやすいソリューションを提供して、同期の取れていないデータがビジネスに悪影響を与える前にこれらを特定します。Oracle GoldenGateとOracle GoldenGate Veridataを併せて導入することで、データ整合性が検証されたリアルタイムのデータ統合および継続的な可用性ソリューションを実現できます。

世界中でクラウドが普及するにつれ、オンプレミスからクラウドへのデータベース・ワークロードのリフト・アンド・シフトが、現在のIT分野における必須事項となっています。Oracle GoldenGateは、このワークロードの移行を支援します。GoldenGate Veridata製品により、オンプレミスとクラウド・データセンターの間でのデータ整合性とデータ・コンプライアンスが確保されます。

Connect with us

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

 blogs.oracle.com

 facebook.com/oracle

 twitter.com/oracle

Copyright © 2020, 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：本文書は情報提供のみを目的としています。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下さい。オラクル製品に関して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。