

Oracle GoldenGate for Big Data



Oracle GoldenGate for Big Data 12c は、ソース・システムのパフォーマンスに影響を与えることなく、ビッグ・データ・システムにデータをリアルタイムにストリーミングします。Apache Hadoop、Apache HBase、Apache Hive、Apache Flume、Apache Kafka、NoSQL データベース、Elasticsearch、JDBC、Amazon Web Services、クラウド・データベース、データウェアハウスなど主要なビッグ・データ・ソリューションにリアルタイム・データを効率的に配信することにより、よりの確な洞察を引き出し、タイムリーな対処へとつなげます。

Oracle GoldenGate for Big Data

おもな機能

- 既存システムの機能を妨げることなく、トランザクション・データをリアルタイムにストリーミング
- 安全性と信頼性に優れたフォルト・トレラントなデータ配信を実現
- インストール、構成、メンテナンスが容易
- ビッグ・データ・システム、リアルタイム・メッセージ・システム、NoSQL データベース、オブジェクト・ストレージ、クラウド・データウェアハウス、オンプレミスの MPP アプライアンスにリアルタイムの変更データをストリーミング
- 変更データのストリーミング先を他のビッグ・データ・ターゲットやメッセージ・キュー（JMS、Apache Kafka、Amazon Kinesis など）に容易かつ柔軟に拡張可能
- スケーラブルで高可用性の Spark クラスタ環境で対話型ユーザー・インタフェースを使用してストリーミング・データを処理および分析

Oracle GoldenGate for Big Data 製品概要

ビジネスに対する洞察力を高め、カスタマー・エクスペリエンスを向上させるうえで、ビッグ・データ分析ソリューションはきわめて重要な役割を担っています。構造化トランザクション・データと半構造化データおよび非構造化データを組み合わせることで、データ資産の持てる価値をあますことなく活用できることは、いまや明白な事実となりつつあります。また、俊敏性、状況へのリアルタイムな対処、短い待機時間でのトランザクション・データへのアクセスがビジネスにとって不可欠なこともわかっています。

構造化データかどうかにかかわらず前日のデータを配信しても意味のない、変化の激しいオペレーションにとっては特に、待機時間の短いトランザクション・データは大きな価値を持ちます。Oracle GoldenGate for Big Data は、本番環境のソース・システムのパフォーマンスを損なうことなく、トランザクション・データをビッグ・データ・ソリューションにリアルタイムにストリーミングする包括的なプラットフォームです。これが、オペレーションの効率とカスタマー・エクスペリエンスを高めるための基盤を築きます。また、使いやすさも備えているため、IT 部門はビッグ・データ・システムにすばやく統合できます。大がかりなトレーニングや管理リソースは必要ありません。

Oracle GoldenGate は、リアルタイム・データの統合と異種データベース・レプリケーションのための実績ある製品です。分析システム、データウェアハウスおよびデータ・マートへのリアルタイム・データの統合や、クラウド・アーキテクチャへのダウンタイム・ゼロでの移行および統合が可能です。また、トランザクション・システム（オンプレミスおよびクラウド）からレポート作成をオフロードしてパフォーマンスを最大化し、アクティブ-アクティブ・データベース・レプリケーションにより信頼性の高いデータ配信と継続的な可用性を実現します。新バージョンの Oracle GoldenGate for Big Data は、Flume、HDFS、Hive および Kafka への配信を最適化して高パフォーマンスで実行し、組織が必要となるリアルタイムでのビッグ・データ分析をサポートします。また、HBase、MongoDB、Cassandra などの NoSQL データベースに対する取得と配信も最適化して高パフォーマンスで実行するため、リアルタイムでの新たなデータ分析活動や機械学習への取り組みにも役立ちます。Oracle GoldenGate for Big Data に

おもな利点

- IT 部門におけるビッグ・データ・システムとの統合作業を効率化
- ビッグ・データ分析にリアルタイム・データを活用して適時性と信頼性の高い洞察を実現
- ビジネスに対するより確かな洞察を基にオペレーションとカスタマ・エクスペリエンスを向上
- ソース・システムのオーバーヘッドを最小化して高いパフォーマンスを確保

Oracle GoldenGate for Big Data

Oracle GoldenGate for Big Data は、リアルタイム・データをビッグ・データ環境にストリーミングし、ビジネスに対する洞察の質と適時性を高めます。オペレーションの効率とカスタマ・エクスペリエンスを高めるための基盤を築きます。

関連製品

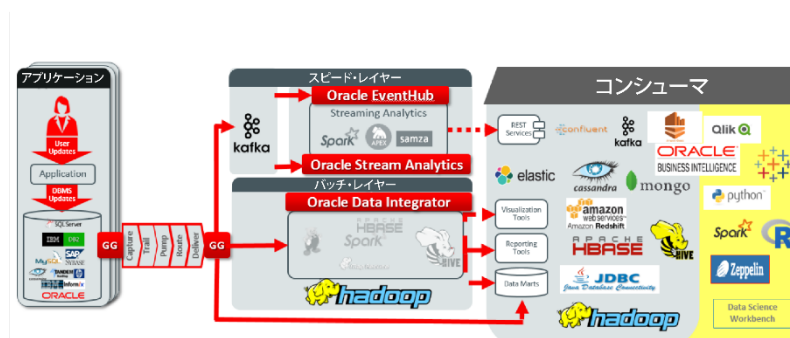
Oracle GoldenGate for Big Data をサポートおよび補完する製品は次のとおりです。

- Oracle GoldenGate
- Oracle GoldenGate Studio
- Oracle GoldenGate Veridata
- Oracle Management Pack for Oracle GoldenGate
- Oracle Stream Analytics
- Oracle GoldenGate Cloud Service
- Oracle Data Integration Platform Cloud

は Oracle GoldenGate for Java および JDBC が含まれているため、Amazon Web Services、Oracle NoSQL、Elasticsearch、SAP HANA、IBM PureData System for Analytics、Greenplum といった他のビッグ・データ・システムへの統合も容易です。

Oracle GoldenGate Big Data リアルタイム・データ・ストリーミング・プラットフォームでは、ビッグ・データの貯水池（ビッグ・データ・レイク）を本番システムに合わせて最新の状態に保つこともできます。また、より高度な分析ツール・セットへ引き渡す前に、待機時間の短いデータに対してアドホック探索を実行したり体系化や拡充を行えます。

Oracle GoldenGate for Big Data は、Oracle Data Integrator および Oracle Stream Analytics とともにオラクルのビッグ・データ統合製品を構成する重要なコンポーネントです。オラクルは、ビッグ・データの統合とオラクルのビッグ・データ・アーキテクチャを連携させることで、Hadoop、NoSQL、クラウドの各テクノロジーの利点であるパフォーマンスとリアルタイム性を活かせるようにしています。



Oracle GoldenGate for Big Data は、主要なビッグ・データ・ターゲットへのリアルタイムのデータ取得とストリーミングをサポートしています。Oracle GoldenGate for Big Data と他のオラクル製品を組み合わせ、ビッグ・データの機械学習や分析のニーズを満たすエンド・ツー・エンドのプラットフォームを構築できます。

HDFS および Hive 用の Oracle GoldenGate ハンドラ

Hadoop 分散ファイル・システム (HDFS) は、汎用サーバーからなる大規模なクラスターにまたがってスケーラビリティと信頼性に優れたデータ・ストレージを提供する Java ベースのファイル・システムです。Apache Hadoop のコアは、HDFS、MapReduce および YARN で形成されています。これは、Microsoft Azure HDInsight、Cloudera Platform、Hortonworks Data Platform、MapR Platform といった商用ベンダー製品でも同じです。Oracle GoldenGate for Big Data を使用して HDFS に変更データをストリーミングすることにより、下流アプリケーションが Hadoop 上でさらにデータをネイティブに処理できるようになるため、ビジネスに対するより確かな洞察を引き出し、総所有コストを削減できます。一方、Hive は、サマライゼーションや分析を実行する、Hadoop 上に提供されるデータウェアハウス・インフラストラクチャです。Oracle GoldenGate は Hive データ・ストアにリアルタイム・データを配信できるため、スケーラブルで俊敏性に優れ、コスト効率の高いリアルタイム・データウェアハウスを Hadoop 上に構築できます。

HBase 用 Oracle GoldenGate ハンドラ

Apache HBase は、Hadoop 分散ファイル・システム（HDFS）上で動作する非リレーショナル（NoSQL）データベースです。列形式であり、フォルト・トレラントなストレージと大量のスパース・データへの高速アクセスを提供します。また、Hadoop にトランザクション機能を付加し、ユーザーが更新、挿入および削除を実行できるようにします。Oracle GoldenGate for Big Data を使用すると、挿入、更新および削除をリアルタイムに HBase に適用し、大量のデータセットから価値を導き出すことができます。

Apache Flume 用 Oracle GoldenGate ハンドラ

Apache Flume は、大容量のストリーミング・データ・フローを効率的に転送する分散型サービスです。データ・フローをリアルタイムに分析し、ペースの速い現代のデジタル・ビジネス環境でのタイムリーな洞察と対処をサポートします。Oracle GoldenGate for Big Data を使用してデータを Flume にリアルタイムに配信することにより、リアルタイム分析アプリケーションを構築できます。

Kafka 用 Oracle GoldenGate ハンドラ

Apache Kafka は、高スループットのパブリッシュ・サブスクライブ型メッセージ・システムです。分散型で、パーティション化、レプリケートされ、フォルト・トレラントであると同時に、従来のメッセージ・システムよりも強力なメッセージ・シーケンスの順序付け機能を備えています。Kafka は、Hadoop を超える汎用エンタープライズ・メッセージ・システムでもあります。GoldenGate は Kafka Pub/Sub API、Kafka Connect API、Kafka REST Proxy API を使用して Confluent Platform と統合できます。Oracle GoldenGate for Big Data を使用してリアルタイム・データを Kafka にストリーミングすることにより、フォルト・トレラントで信頼性が高く、拡張可能なリアルタイム分析アプリケーションを構築できます。

MongoDB 用 Oracle GoldenGate ハンドラ

MongoDB は、高いパフォーマンスおよび可用性と自動スケーリング機能を備えたオープンソースのドキュメント・データベースです。高速でのデータの読み取りと消費ユーザーへの公開が可能のため、ペースの速い現代のデジタル・ビジネス環境でタイムリーに洞察を引き出し、状況に対処できます。Oracle GoldenGate for Big Data を使用してデータを MongoDB にリアルタイムに配信することにより、スケーラブルで可用性の高いリアルタイム分析アプリケーションを構築できます。

Cassandra 用 Oracle GoldenGate ハンドラ

Apache Cassandra は、複数のマシンに渡る水平方向のスケーリングにより大量のデータを格納するように設計された NoSQL データベース管理システムです。デバイスや高速コンピュータでデータが生成される、ペースの速い現代の大規模なデジタル・ビジネス環境において、非常に高速にデータを書き込むことのできるデータベースの 1 つです。Oracle GoldenGate for Big Data を使用してリアルタイムにデータを取得して Cassandra にリアルタイムに配信することによ

り、スケーラビリティに優れた高速で動的なリアルタイム・ベースの分析システムを構築できます。

JDBC 用 Oracle GoldenGate ハンドラ

JDBC (Java Database Connectivity) は、もっとも汎用性の高いテクノロジー・インタフェースの 1 つで、オンプレミスおよびクラウド上で動作するスプレッドシート、データベースおよびアプリケーション・プロバイダに広く使用されています。Oracle GoldenGate for Big Data を使用してデータを JDBC インタフェースにリアルタイムに配信することにより、Amazon Redshift、IBM Netezza、Greenplum などの大規模なロングテール・ターゲット・データベースにトランザクション・データをレプリケートし、リアルタイム分析レポート・システムを構築できます。

Amazon Kinesis 用 Oracle GoldenGate ハンドラ

Amazon Kinesis はストリーミング・データをリアルタイムに取得、バッファ、処理できる Amazon Web Services のフルマネージド型アプリケーションで、数秒または数分でインサイトを得ることができます。Amazon Kinesis を使用すると、データの取得と同時にデータを処理、分析して即座に回答することができ、すべてのデータが収集されてからでなければ処理を開始できないということはありません。Kinesis のストリームは、Amazon S3 や Amazon Redshift といった他の Amazon Cloud アプリケーションにデータをストリーミングする場合にも使用できます。Oracle GoldenGate for Big Data を使用すると、どのトランザクション・システムやメッセージ・バスからでも Amazon Kinesis にリアルタイムにデータをストリーミングして、貴重なインサイトを数秒で取得することができます。

Elasticsearch 用 Oracle GoldenGate ハンドラ

Elasticsearch は拡張性の高いオープンソースのフルテキスト検索および分析エンジンです。Elasticsearch を使用すると、大量のデータをすばやく、ほぼリアルタイムで保存、検索および分析できます。通常、複雑な検索機能を備えたアプリケーションを駆動する基盤となるエンジンやテクノロジーとして使用されます。Oracle GoldenGate for Big Data を使用してリアルタイムにデータを Elasticsearch に配信することにより、スケーラブルで可用性が高いリアルタイムの複雑な検索機能を分析アプリケーションに組み込むことができます。

ファイル・ライター用 Oracle GoldenGate ハンドラ

Oracle GoldenGate フラット・ファイルを使用すると、ETL、独自アプリケーションまたは従来型アプリケーションなどのバッチ・ファイル入力を想定したツールで使用するディスクにマイクロ・バッチを書き込むことができます。データは、デリミタ区切り値、長さ区切り値、ORC、Parquet 形式など、ターゲット・アプリケーションの仕様に合わせて書式化されて HDFS や AWS S3 に書き込まれます。バッチ・ファイルのロールオーバーに充てる時間枠を数分もしくは数秒に短縮すると、ほぼリアルタイムでイベントをこれらのシステムにフィードできます。

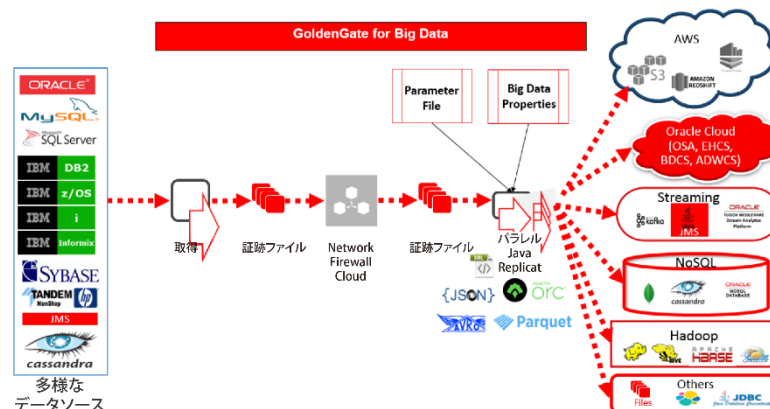
Oracle Stream Analytics

Oracle Stream Analytics では、高度な相関パターン、エンリッチメント、機械学習を利用して大量のリアルタイム情報を処理および分析できます。ビジネスに実際に役立つリアルタイムの情報をストリーミング・データから取得して、ビジネスを俊敏に推し進めるためのアクションを自動化します。Oracle GoldenGate for Big Data には Oracle Stream Analytics が組み込まれていて、Kafka 用 Oracle GoldenGate ハンドラから配信されるイベントを処理できるようになっています。

新たな価値を生み出す重要技術

Oracle GoldenGate が備えるモジュール型アーキテクチャは、きわめて優れたパフォーマンス、フォルト・トレランスおよび柔軟性を実現するように設計されています。Oracle GoldenGate for Big Data は、1 つのアーキテクチャを基盤に、お客様のビッグ・データ環境に合った拡張性の高いソリューションを構築します。

Oracle GoldenGate は、既存システムの機能を妨げることなくわずかなオーバーヘッドで、Java ベースのメッセージ・システムなどの異種ソース・システムからデータを取り込みます。また、データベース・トランザクションを証跡ファイルに格納するとともにパラレル Java Replicat フレームワークに投入します。Oracle GoldenGate for Big Data には、HDFS、Hive、HBase、Flume、Kafka、NoSQL、JDBC、Amazon Kinesis、Amazon S3 など、さまざまなビッグ・データ・テクノロジー用のハンドラが含まれています。データを XML、JSON、Avro、ORC、Parquet、デリミタ付きテキストなどの標準書式に変換するためのプラグラブルな書式設定テクノロジーも備えています。拡張性の高い Java アーキテクチャをベースとしているため、ビジネス・ニーズに応じて追加のビッグ・データ・システムを提供したり、特定のユースケースをサポートすることも可能です。



Oracle GoldenGate リアルタイム・ストリーミング・プラットフォームがあれば、ソース・システムから取り込んだ変更データのすべてまたは一部を、オンプレミス、クラウド、またはそのハイブリッドで運用されるデータベースやメッセージ・システム、さらにビッグ・データ環境を含む複数のターゲットに配信できます。

Oracle GoldenGate では、サポートされる異種システムが継続的に拡大されており、エンタープライズ全体を単一のデータ・ストリーミング・プラットフォームに標準化できます。現在、Oracle、DB2 for z/OS、iSeries、LUW (Linux、Unix、Windows)、SQL Server、MySQL、Informix、Sybase ASE、SQL/MX、SQL/MP、Enscribe、フラット・ファイルなどからのログベースのキャプチャおよびこれらのシステムへの配信をサポートしています。また、Hadoop ベースおよび非 Hadoop ベースのビッグ・データ・システムに加え、Oracle Stream Analytics、Oracle TimesTen In Memory Database、Teradata および PostgreSQL への配信機能も備えています。

まとめ

Oracle GoldenGate for Big Data は、パフォーマンスとフォルト・トレランスに優れた、使いやすく柔軟性の高い、ビッグ・データ環境のためのリアルタイム・データ・ストリーミング・プラットフォームです。ソース・システムのパフォーマンスに影響を与えることなく、既存のリアルタイム・データ統合アーキテクチャをビッグ・データ・システムへと容易に拡張し、ビジネスに対するタイムリーな洞察とよりの確な意思決定を可能にします。



お問い合わせ

Oracle GoldenGate for Big Data の詳細については、oracle.com を参照するか、+1.800.ORACLE1 でオラクルの担当者にお問い合わせください。

CONNECT WITH US



blogs.oracle.com/oracle



facebook.com/oracle



twitter.com/oracle



oracle.com

Hardware and Software, Engineered to Work Together

Copyright © 2018, 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。

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。0518



Oracle is committed to developing practices and products that help protect the environment