Oracle GoldenGate for Big Data

Oracle GoldenGate for Big Data 12c product streams transactional data into big data systems in real time, without impacting the performance of source systems. It streamlines real-time data delivery into most popular big data solutions, including Apache Hadoop, Apache HBase, Apache Hive, and Apache Flume, and facilitates improved insight and timely action.

Oracle GoldenGate for Big Data Product Overview

Big data analytics solutions play a pivotal role in improving business insight and enhancing customer experience. It is increasingly evident that organizations can realize the full potential value of their data assets by combining the structured transactional data with semi-structured and unstructured data. Businesses also notice that to be agile and react to situations in real time, access to transactional data with low latency is essential.

Low-latency transactional data brings additional value especially for dynamically changing operations that day-old data, structured or unstructured, cannot deliver. Oracle GoldenGate for Big Data product offers a comprehensive platform that streams transactional data into big data solutions in real time, without degrading the performance of the source production systems. It lays the foundation for more efficient operations and improved customer experience. With its easy-to-use solution, Oracle GoldenGate for Big Data enables IT organizations to quickly integrate into big data systems without extensive training and management resources.

Oracle GoldenGate is a time tested and proven product for real-time data integration and heterogeneous database replication. It supports organizations with real-time data integration for analytical systems, zero downtime migration and consolidation to cloud architectures, offloading reporting from transactional systems—on premises and cloud—for maximum performance, and active-active database replication for data distribution and continuous availability. In this new offering, Oracle GoldenGate for Big Data provides optimized and high performance delivery to Flume, HDFS, Hive and Hbase to support customers with their real-time big data analytics initiatives. Oracle GoldenGate for Big Data includes Oracle GoldenGate for Java, which enables customers to easily integrate to additional big data systems, such as Oracle NoSQL, Apache Kafka, Apache Storm, Apache Spark, and others.

Oracle GoldenGate Big Data’s real-time data streaming platform also allows customers to keep their big data reservoirs, or big data lakes, up to date with their production systems. With this solution, customers can perform ad-hoc discovery, organization, and enrichment of low-latency data before it traverses to more refined sets of analytics tools.
Oracle GoldenGate for Big Data helps organizations raise the quality and timeliness of their business insights by streaming real-time transactional data to big data environments. It lays the foundation for more efficient operations and improved customer experience.

**Related Products**

The following products support and complement Oracle GoldenGate for Big Data:

- Oracle GoldenGate
- Oracle GoldenGate Veridata
- Oracle Management Pack for Oracle GoldenGate
- Oracle GoldenGate Application Adapters

Oracle GoldenGate for Big Data is a key component of Oracle’s big data integration offering along with Oracle Data Integrator 12c. Oracle combines big data integration with Oracle’s Big Data Architecture to give customers the ability to capture the performance and real time advantages of Hadoop technologies.

Oracle GoldenGate for Big Data supports real-time data streaming to most popular big data targets.

Customers can combine Oracle GoldenGate for Big Data with Oracle GoldenGate product to have an end-to-end platform for real-time streaming from heterogeneous systems.

**Oracle GoldenGate Adapter for Apache Flume**

Apache Flume is a distributed service for efficiently transporting high-volume, streaming data flows. It performs real-time analytics on data flows to support timely insight and action in today’s fast-paced digital business environment. By streaming real-time data into Apache Flume systems via Oracle GoldenGate for Big Data, organizations can build fault-tolerant, highly reliable, and extensible real-time analytical applications.

**Oracle GoldenGate Adapter for HDFS**

Hadoop Distributed File System (HDFS) is a Java-based file system that provides scalable and reliable data storage that is designed to span large clusters of commodity servers. HDFS, MapReduce, and YARN form the core of Apache Hadoop. Using Oracle GoldenGate for Big Data to stream changed data to HDFS enables the downstream applications to further process data natively on Hadoop for better business insights and for lower total cost of ownership.

**Oracle GoldenGate Adapter for Hive**

Hive is a data warehousing infrastructure provided on top of Hadoop for doing summarization and analysis. Oracle GoldenGate can now provide real-time data to the Hive Data Store to enable scalable, agile, and cost effective real-time data warehousing on Hadoop.

**Oracle GoldenGate Adapter for HBase**

Apache HBase is a non-relational (NoSQL) database that runs on top of the Hadoop Distributed File System (HDFS). It is columnar and provides fault-tolerant storage and quick access to large quantities of sparse data. It also adds transactional capabilities to
Hadoop, allowing users to conduct updates, inserts and deletes. Now with Oracle GoldenGate for Big Data, inserts, updates and deletes can be applied to HBase in real-time for deriving value from large data sets.

**Key Technical Value Drivers**

Oracle GoldenGate offers a modular architecture designed for extreme performance, fault tolerance, and flexibility. Oracle GoldenGate for Big Data is built on the same architecture to enable extensible solutions for customers’ big data environments.

Oracle GoldenGate captures data from heterogeneous source systems including Java-based messaging systems, non-invasively and with negligible overhead. It stores database transactions in Trail Files and pumps it into the Java Adapter framework. Oracle GoldenGate for Big Data includes adapters for different big data technologies, such as HDFS, Hive, HBase or Flume. It also includes Oracle GoldenGate for Java, which enables customers to easily deliver to additional big data systems and support specific use cases that their environment demands.

With Oracle GoldenGate’s real-time streaming platform, customers can capture from the source system once and deliver all, or portion, of the changed data to multiple targets on premises and cloud, including databases, messaging systems, and now big data environments.

With constantly expanding heterogeneous support, Oracle GoldenGate enables organizations to standardize on a single real-time data streaming platform for their enterprise. Oracle GoldenGate supports log-based capture from, and delivery to, Oracle, DB2 for z/OS, i Series, & LUW (Linux, Unix, Windows), SQL Server, MySQL, Informix, Sybase ASE, SQL/MX, JMS messaging systems and more. Oracle GoldenGate’s delivery capabilities also include Oracle TimesTen In Memory Database and PostgreSQL, in addition to Hadoop-based big data systems.

**Summary**

Oracle GoldenGate for Big Data offers high-performance, fault-tolerant, easy-to-use, and flexible real-time data streaming platform for big data environments. It easily extends customers’ real-time data integration architectures to big data systems without impacting the performance of the source systems and enables timely business insight for better decision making.