

Oracle® Fusion Middleware

Release Notes for Oracle GoldenGate for Big Data

12c (12.2.0.1.1)

E67176-03

June 2016

Oracle GoldenGate for Big Data 12c streams transactional data into big data systems in real time, raising the quality and timeliness of business insights. This document contains the release notes for the 12c (12.2.0.1.1) release of Oracle GoldenGate for Big Data.

This document contains the following sections:

- [New and Changed Features](#)
- [New or Changed Properties](#)
- [Upgrade](#)
- [Deprecated Items](#)
- [Corrected Problems](#)
- [Known Issues](#)
- [Documentation Accessibility](#)
- [Related Documents](#)
- [Getting Help with My Oracle Support](#)

New and Changed Features

This section outlines the new features that are being introduced in the release, features that have been enhanced, and changes to existing features.

Patch Release 12.2.0.1.1 — May 2016

- New supported versions are:
 - HDFS Handler
 - Hortonworks HDP 2.4 (HDFS 2.7.1)
 - Cloudera CDH 5.7.x (HDFS 2.6.0)
 - Cloudera CDH 5.6.x (HDFS 2.6.0)
 - Cloudera CDH 5.5.x (HDFS 2.6.0)
 - HBase Handler
 - Apache HBase 0.98.x
 - Apache HBase 0.96.x
 - Hortonworks HDP 2.4 (HBase 1.1.2)
 - Hortonworks HDP 2.2 (HBase 0.98.4)

Cloudera CDH 5.7.x (HBase 1.2.0)
Cloudera CDH 5.6.x (HBase 1.0.0)
Cloudera CDH 5.5.x (HBase 1.0.0)
Flume Handler
Hortonworks HDP 2.4 (Flume 1.5.2)
Cloudera CDH 5.7.x (Flume 1.6.0)
Cloudera CDH 5.6.x (Flume 1.6.0)
Cloudera CDH 5.5.x (Flume 1.6.0)
Kafka Handler
Apache Kafka 0.9.0.x
Hortonworks HDP 2.4 (Kafka 0.9.0.0)
Cloudera CDH does not include Kafka but releases Kafka separately
Cloudera Distribution of Apache Kafka 2.0.x (Kafka 0.9.0.0)
Cloudera Distribution of Apache Kafka 1.x (Kafka 0.8.2.0)

- Kafka 0.9.0 added support for securing client connections to Kafka using SSL/TLS or Kerberos. The Kafka client libraries provide a complete abstraction for security functionality so no changes were required to the Kafka Handler. Enabling security does require additional configuration in the Kafka producer properties file. For more information about enabling Kafka security, see <http://kafka.apache.org/documentation.html#security>
- The HBase client interface has been evolving in the last several HBase releases. In the previous Oracle GoldenGate for Big Data release, only HBase 1.0.0 and higher was supported. Support for HBase 0.98.x and 0.96.x in the HBase Handler was added in this release using the new configuration property, `gg.handler.name.hBase98Compatible`. Valid values are `true` and `false`; the default is `false`. This property can also be used for Cloudera CDH 5.7.x - 5.4.x. The 5.7.x - 5.4.x CDH versions include HBase 1.x+ but binary compatibility with Apache HBase 1.0.0 was broken to retain backward compatibility.
- The new `gg.schemareplaceregex` and `gg.schemareplacestring` parameters allow you to configure search criteria of catalog names, schema names, table names, and column names using regular expressions then replace occurrences of matches with configured replacement values.
- The new `gg.contentreplaceregex` and `gg.contentreplacestring` parameters allow you to configure search criteria of column values using regular expressions then replace occurrences of matches with configured replacement values.
- Oracle Wallet integration was implemented to address security concerns. You can now create a credential store with Oracle Wallet and store your credentials in it using an alias and domain name. Properties like `username` and `password` do not need to be specified in clear text anymore.
- HDFS Handler support to write corresponding generated schemas from formatters to HDFS. Currently Avro and JSON formatters generate schemas.
- Support for writing files to HDFS in Avro Object Container File (OCF) format and optionally to create tables in Apache Hive automatically. Avro OCF integrates well with Apache Hive.
- Support for writing to HDFS in sequence file format.

- Support to encode binary column value data as hex encoding. Default encoding for binary data is Base64.
- JSON formatter adds the ability to optionally include an array of the primary key column names in the generated JSON output.
- Support for writing operation data to Kafka using multiple topics. Topic name selected corresponds to the fully qualified table name of the operation.
- In the previous release, a field for tokens was not suppressible in the delimited text formatter. Setting the configuration parameter `gg.handler.name.includeTokens=false` would cause the tokens field to be empty though the field was still present. In the current release, setting `gg.handler.name.includeTokens=false` suppresses the tokens field content in the delimited text output and no field for tokens will be included in the output.

Initial Release — December 2015

- The Flume, Kafka, HBase, HDFS, Pluggable Formatters, and Metadata Provider handlers are introduced.
All of these handlers support the Delimited Text, JSON, Avro, and XML formatters, which are also introduced.
All handlers support truncate operations.
- You can retrieve object metadata from the Java metadata provider since the target metadata needs to be retrieved from the actual target.
- Replicat for Java was enhanced as follows:
 - To include commit or checkpoint events to Java per the Replicat for Java-based targets.
 - DDL event propagation to Java module is introduced.
 - Retrieval of entity metadata from the target metadata provider is introduced.
 - The checkpoint flow was enhanced to support non-RDBMS Java targets.
 - Support for using the existing callback API being called from Java is introduced.
 - Support for propagated LCR events to Java is introduced.
 - Support for using multi-byte object names (columns, tables, schema).
 - The source metadata is used if no mapping is specified.
- The flushing of the outstanding operations and checkpoint within the group that has not met the `minGroupSize` rule is introduced. The flushing occurs based on the byte size related rule.
- Only one handler can be used at a time to be consistent with the behavior that Replicat expects.

- The checkpoint behavior was modified to support the Oracle GoldenGate Replicat mode with respect to `maxGroupSize` (for operation-based grouping) and `maxByteGroupSize` (for size-based grouping) behavior.
- The HBase Handler has moved to the new recommended 1.0.0 HBase client interface. As a result, the HBase Handler is not compliant with releases of HBase older than 1.0.0. If an integration is required with 0.99.x or older release of HBase, this can be accomplished using the 12.1.2.1.x HBase Handler. Contact Oracle Support to obtain a ZIP file of the 12.1.2.1.x HBase Handler.

The HBase Handler includes a configuration parameter for encoding. In the previous release, native system encoding was assumed. Explicit configuration of the encoding gives you better control of the encoding of data streamed to HBase.

The HBase Handler has improved support for primary key updates. The HBase Handler includes a configuration parameter to perform one of the following three actions when encountering a primary key update, `abend`, treat the action as a normal update, treat the action as a delete, and subsequent insert. Primary key updates are somewhat problematic for HBase because data is only appended. The queried contents of an HBase row are simply the consolidation of all the individual HBase operations. The configuration choice that most accurately models what is occurring is to perform a delete in HBase using the before change primary key values and a subsequent insert using the after change primary key values. If this behavior is configured, you must have full supplemental logging enabled (full before and after images of all columns) so that the data in the HBase insert operation is complete.

The HBase Handler provides a new configuration value for NULL values. In the previous release NULL values were inserted into HBase as NULL. This is the default though you have the ability to override the default.

- The HDFS Handler source code was included in the 12.1.2.1.x release. The HDFS Handler source code is not included in the 12.2.0.1 release. The 12.1.2.1.x release included limited functionality, therefore the source code was provided to allow you to extend the functionality to their specific needs. The functionality in the 12.2.0.1 release is more robust so eliminates or greatly reduces your need modify it.

The 12.1.2.1.x HDFS Handler only supports output in delimited text format. The 12.2.0.1 HDFS Handler supports output in delimited text, JSON, XML, and two different Avro formats.

The 12.1.2.1.x HDFS Handler did not support metadata change events. The 12.2.0.1 HDFS Handler can support metadata change events (assuming that upstream replication supports metadata in trail and DDL propagation).

The truncate table operation is supported in the 12.1.2.1.x HDFS Handler.

- The Flume Handler source code was included in the 12.1.2.1.x release; The Flume Handler source code is *not* included in the 12.2.0.1 release. The 12.1.2.1.x release included limited functionality and therefore the source code was provided to allow you to extend the functionality for your needs. The functionality in the 12.2.0.1 release is more robust thus eliminating or at least greatly reducing the your need to modify it.

The 12.1.2.1.x Flume Handler can work with the 12.2.0.1 release with some modifications. The most significant modification being that the Oracle Java package

names have changed from `com.goldengate` to `oracle.goldengate` to be compliant with official Oracle Java package naming conventions. If you want to use the 12.1.2.1.x Flume Handler in the 12.2.0.1 release, you must change the package names, import statements, and then recompile your project. Alternatively, Oracle Support can provide you with a release of the 12.1.2.1.x Flume Handler with the properly changed package names.

The 12.2.0.1 Flume Handler supports pluggable formatters. Delimited text is supported and it is simply one option. Oracle Golden Gate 12.2.0.1 ships with five pluggable formatters, delimited text, XML, JSON, Avro row format, and Avro operation format. You can also write custom formatters.

Metadata change events (DDL) are supported by the Flume Handler in release 12.2.0.1 release. However, this functionality depends on the source replicated database and the upstream Oracle GoldenGate Capture process to capture and replicate DDL events. This feature is not immediately available for all database implementations in Oracle GoldenGate 12.2. See the Oracle GoldenGate documentation for your database implementation for information about DDL replication.

The 12.1.2.1.x Flume Handler did not support operation mode. The 12.2.0.1.0 Flume handler can push operations as flume events one at a time as they are received.

The 12.1.2.1.x Flume Handler did not require or support the Flume Agent configuration file in the configured classpath. The 12.2.0.1 Flume Handler uses the Flume Agent configuration file to resolve the host and port of the Flume source. Additionally, the 12.2.0.1 Flume Handler can support advanced Flume features available in Flume 1.6.0 including load balancing, fail over, and Kerberos authentication.

The Flume 12.1.2.1.x Handler provided in-built headers that could be only be modified by changing the source code. The Flume Handler 12.2.0.1 provides an interface that can be implemented and used for custom headers, in addition to pre-existing default headers that can be over-ridden. Any dynamic transaction or operation data can be used to populate the event message headers.

New or Changed Properties

This section lists the properties that have been added, changed, or documented for the first time for the release.

Patch Release — May 2016

- Traceability of schema changes was added so that you can add versions to schemas using the `gg.handler.name.format.versionSchemas` property. Schema versioning is optional. Schema versioning causes a schema with a timestamp to be created in the schema directory on the local file system every time a new schema is created. You can set schema versioning as follows:

```
gg.handler.name.format.versionSchemas
```

For the schemas created by the JSON and Avro formatters. Valid values are `true` and `false`; the default is `false`.

- The `gg.handler.name.schemaFilePath` property was added to the HDFS Handler to persist created schemas to HDFS. You set this new configuration

property, `gg.handler.name.schemaFilePath=`, to a file directory inside HDFS to the location that you want to write the schemas. A metadata change event causes the schema to be overwritten when the next operation for the associated table is encountered. Schemas follow the same naming convention as for schemas written to the local file system.

- The `gg.handler.name.format.schemaDirectory` property was added to the Avro row and operation formatters so that you can configure the location on your local file system where the created Avro schemas are written.
- The `gg.handler.name.format.iso8601Format` is a new property on the delimited text, XML, Avro, and JSON formatters. Set this property to `false` to replace the T between the date and time in the current timestamp with a space. Valid values are `true` and `false`; the default is `true`.
- The `gg.handler.name.rowkeyDelimiter` property was added to the HBase Handler. It configures the delimiter between primary key values from the source table when generating the HBase rowkey. The default delimiter is the pipe character `|`.
- The Kafka handler added the new property `gg.handler.name.topicPartitioning`. Set property value to `table` to write operation data from the source trail file to different Kafka topics based on table name. The default configuration value is `none` which uses the single configured Kafka topic for all the data.
- Support for hex encoding of binary data. The `gg.binaryencoding` property allows you to enable hex encoding. The default encoding of binary data is Base64.
- Support for writing to HDFS in sequence file format. The `gg.handler.hdfs.format=sequenceFile` property enables writing to HDFS in sequence file format.
- In the previous Oracle GoldenGate for Big Data release, only HBase 1.0.0 and higher was supported. Support for HBase 0.98.x and 0.96.x in the HBase Handler was added in this release using the new configuration property, `gg.handler.name.hBase98Compatible`. Valid values are `true` and `false`; the default is `false`. This property can also be used for Cloudera CDH 5.7.x - 5.4.x. The 5.7.x - 5.4.x CDH versions include HBase 1.x+ but binary compatibility with Apache HBase 1.0.0 was disabled to retain backward compatibility.
- The JSON formatter added the new property `gg.handler.name.format.includePrimaryKeys` which can be set to `true` or `false`. When set to `true` an array of the primary key column names is included in the JSON output. The default is `false`.

Initial Release — December 2015

- Encoding is supported on the HBase handler using the `gg.handler` property. Configuration for encoding is optional. You can set the encoding for HBase as in following:

```
gg.handler.handlername.encoding=UTF-8
```

For the HDFS, Flume, and Kafka Handlers configured encoding is supported using the delimited text, JSON, Avro, and XML formatters. Configuration for encoding is optional.

Upgrade

The upgrade from Oracle GoldenGate for Big Data release 12.2.0.1.0 to the 12.2.0.1.1 release is seamless and does not require any action on your part.

Deprecated Items

This section lists all items that were deprecated in each release.

Initial Release — December 2015

- The `goldengate.userexit.rawchars` and `goldengate.userexit.utf8mode` properties have been deprecated. This is applicable to both flat file writers and Java user exit.

Corrected Problems

This section describes the problems corrected in each release. The Bug *number* is the number of the BugDB ticket. For questions on specific tickets or issues, consult Oracle Support.

Patch Release — May 2016

Bug 22675718 - Java Adapter abends when table names are not mapped

Fixed an issue that was causing empty table names when handling DDL on table that was not part of the mapping.

Bug 23017779 - How to encrypt the password in `gg.mdp.password`

This new feature was added in this release, see [New and Changed Features](#).

Bug 22743724 - Global parameter that will filter out carriage return (hex od)

This new feature was added in this release, see [New and Changed Features](#).

Bug 22518330 - HBase Handler key column delimiter/separator is not configurable

Functionality was added to the HBase Handler to make the HBase rowkey delimiter configurable. The default value is still a pipe, |. You can configure the this delimiter using the following configuration:

```
gg.handler.name.rowkeyDelimiter=your_delimiter
```

Bug 22265877 - Adapter abends with: Unable to create column metadata: Unknown Data SubType: 7

An issue with Java Delivery encountering a trail file data sub-type that may not be internally defined in the Java namespace no longer causes an abend and processing now continues with relevant defaults values for the data type.

Bug 22951116 - Java Adapter 12.1.2.1.7 `dsOperation.getToken("TKN-TRANSACTION-CS")` is null

A new API was added so that all user and Oracle GoldenGate user tokens are accessed.

Initial Release — December 2015

Bug 21394677 - Flat File: escaped chars are replaced with bad characters

Fixed an issue where unexpected extra data after a NULL terminated the string in the flat file writer.

Bug 21437275 - REG: Flat file: VARCHAR column value is truncating

Fixed an issue with column values for the VARCHAR data type truncating.

Bug 18268283 - Replicat for Java: Support multi-byte object names

Support for multi-byte object names was added in Replicat for Java targets.

Bug 20816615 - Java Delivery: Limit max/min group size INFO messages to transactions

Fixed an issue with INFO messages being moved to DEBUG level to prevent the logs from getting too big or detailed at INFO level logging. Typically, this is the preferred logging level in production environments.

Bug 9660839 - Event action do not work with user exits

Support for user exits event actions was added for Replicat-based features.

Bug 20933435 - Using user exit checkpoint with custom handler

Ensure that you review and follow the instructions in the "Java Delivery Transaction Grouping" section of *Administering Oracle GoldenGate for Big Data*.

Bug 21180036 - Three Part Naming: Hive and HVBase handlers are missing catalog names in the output

The HDFS subdirectory name is the fully qualified table 3-part table name: table name + current timestamp + suffix. In this release the Hive handler was changed to the HDFS handler.

Bug 21942367 - Enable EVENTACTION functionality when you do not have a database

Support for using EVENTACTION without a database was added for Replicat so that you can invoke scripts at the shell level.

Bug 20426021 - HDFS Delivery: Warning messages due to extra whitespaces in the property values

Support for wrapping configuration property values with CDATA[] was added so that you can use whitespace and non-whitespace characters inside this wrapper in the HDFS handler. This allows you to preserve the values and does not create any whitespace warnings in the log file.

Bug 18238949 - Group records based on size

Support for message size grouping in custom handlers, as well as JMS and any XML formatter that needs to be used, was added.

Bug 13450906 - ER: javaue should support the truncate operation too

Support for the use of truncate operations with `javaue` was added.

Bug 20989753 - Declarative Configuration for Flume Delivery

Support for the Flume handler integrating with the Delimited Text, JSON, Avro, and XML formatters was added.

Known Issues

This section describes the known issues, with any available workarounds, identified in each release. The Bug *number* is the number of the BugDB ticket. For questions on specific tickets or issues, consult Oracle Support.

Patch Release — May 2016

BUG 23261763 - kafka replicat sends duplicate records in Async mode upon Stop/restart

Kafka replicat sends duplicate record to topics in async mode on graceful stop and restart. As a workaround, you can use `blockingSend=true`.

Initial Release — December 2015

BUG 22294888 - WIN: Reports are not displaying correct label version

Tokens within Extract do not work in a Windows environment. As a workaround, you can use Replicat instead of Extract.

Documentation Changes

This section discusses changes to the Oracle GoldenGate for Big Data documentation.

Initial Release — December 2015

A new Oracle GoldenGate for Big Data documentation set and library are introduced and found at

<http://docs.oracle.com/goldengate/bd1221/gg-bd/>

Related Documents

The Oracle GoldenGate for Big Data documentation set includes the following components:

- *Release Notes for Oracle GoldenGate for Big Data*
- *Integrating Oracle GoldenGate for Big Data*
- *Installing Oracle GoldenGate Big Data*
- *Administering Oracle GoldenGate for Big Data*

Getting Help with My Oracle Support

Use My Oracle Support to find knowledge solutions, workarounds, and other information that is reported by customers, partners, and Oracle employees. My Oracle Support also enables you to open a Service Request. If a patch is required to resolve a

service request, you will receive instructions on how to download it from My Oracle Support.

Note:

If you purchased Oracle GoldenGate and support through a distributor, contact your distributor instead of attempting to create a service request through My Oracle Support.

Viewing the Oracle GoldenGate Knowledge Base

To view the Oracle GoldenGate Knowledge Base, follow these steps:

1. Browse to [the My Oracle Support web site at http://support.oracle.com](http://support.oracle.com).
2. Select your language and then log in with your email and Oracle password.
3. Click the **Knowledge** tab.
4. In the **Select a product or product line** field, enter **GoldenGate**, and then select an Oracle GoldenGate product from the context-sensitive list to narrow your results. (You may have to wait a few seconds for this list to appear).
5. In the **Enter search terms** field, enter a search keyword or multiple keywords to focus the query.

Additional information about how to use this tab is obtained by clicking **Help**.

Creating an Oracle GoldenGate Support Request Ticket

If you cannot find an answer to your question or problem in the Knowledge Base, you can open a support request ticket with Oracle Support by following these steps:

1. Browse to [the My Oracle Support web site at http://support.oracle.com](http://support.oracle.com).
2. Select your language, and then sign in with your credentials. If you have not done so, you will have to register to use this web site.
3. Click the **Service Requests** tab.
4. Click **Create SR**.
5. Use the Create Service Request wizard to complete and submit your SR. Additional information about how to use this tab is obtained by clicking **Help**.

Documentation Accessibility

For information about Oracle's commitment to accessibility, visit the Oracle Accessibility Program website at <http://www.oracle.com/pls/topic/lookup?ctx=acc&id=docacc>.

Access to Oracle Support

Oracle customers that have purchased support have access to electronic support through My Oracle Support. For information, visit <http://www.oracle.com/pls/>

topic/lookup?ctx=acc&id=info or visit <http://www.oracle.com/pls/topic/lookup?ctx=acc&id=trs> if you are hearing impaired.

Oracle® Fusion Middleware Release Notes for Oracle GoldenGate for Big Data, 12c (12.2.0.1.1)
E67176-03

Copyright © 2015, 2016, Oracle and/or its affiliates. All rights reserved.

This software and related documentation are provided under a license agreement containing restrictions on use and disclosure and are protected by intellectual property laws. Except as expressly permitted in your license agreement or allowed by law, you may not use, copy, reproduce, translate, broadcast, modify, license, transmit, distribute, exhibit, perform, publish, or display any part, in any form, or by any means. Reverse engineering, disassembly, or decompilation of this software, unless required by law for interoperability, is prohibited.

The information contained herein is subject to change without notice and is not warranted to be error-free. If you find any errors, please report them to us in writing.

If this is software or related documentation that is delivered to the U.S. Government or anyone licensing it on behalf of the U.S. Government, then the following notice is applicable:

U.S. GOVERNMENT END USERS: Oracle programs, including any operating system, integrated software, any programs installed on the hardware, and/or documentation, delivered to U.S. Government end users are "commercial computer software" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, use, duplication, disclosure, modification, and adaptation of the programs, including any operating system, integrated software, any programs installed on the hardware, and/or documentation, shall be subject to license terms and license restrictions applicable to the programs. No other rights are granted to the U.S. Government.

This software or hardware is developed for general use in a variety of information management applications. It is not developed or intended for use in any inherently dangerous applications, including applications that may create a risk of personal injury. If you use this software or hardware in dangerous applications, then you shall be responsible to take all appropriate fail-safe, backup, redundancy, and other measures to ensure its safe use. Oracle Corporation and its affiliates disclaim any liability for any damages caused by use of this software or hardware in dangerous applications.

Oracle and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

Intel and Intel Xeon are trademarks or registered trademarks of Intel Corporation. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. AMD, Opteron, the AMD logo, and the AMD Opteron logo are trademarks or registered trademarks of Advanced Micro Devices. UNIX is a registered trademark of The Open Group.

This software or hardware and documentation may provide access to or information about content, products, and services from third parties. Oracle Corporation and its affiliates are not responsible for and expressly disclaim all warranties of any kind with respect to third-party content, products, and services unless otherwise set forth in an applicable agreement between you and Oracle. Oracle Corporation and its affiliates will not be responsible for any loss, costs, or damages incurred due to your access to or use of third-party content, products, or services, except as set forth in an applicable agreement between you and Oracle.