

Oracle Golden Gate Lets Get Started



Michael Messina
Senior Managing Consultant
AdvizeX, A Rolta Company



Speaker Introduction

- Michael Messina
- Senior Managing Consultant AdvizeX, A Rolta Company
- Working with Oracle for over 20 years
- Background includes Performance Tuning, High Availability and Disaster Recovery
- Oracle Database OCP
- Oracle RAC Certified Expert
- Oracle Exadata Implementation Specialist
- Oracle ACE
- Mike.Messina@rolta.com
- www.rolta.com / www.advizex.com



Agenda

- What is Golden Gate
- Why Use Golden Gate
- Golden Gate Architecture
- Golden Gate Requirements/Limitations
- Golden Gate one-way Replication
 - Replace Streams
 - Near 0 Downtime Platform Migration
 - Near 0 Downtime Upgrade
 - Near 0 Downtime Maintenance
- Golden Gate Multi-master Replication



Agenda

- Golden Gate Extract, Transform, Load (ETL)
 - Load Your Datawarehouse
- Some Best Practices for Manager Process
- Some Best Practices for Extract Processes
- Some Best Practices for DataPump Processes
- Some Best Practices for Replicat Processes
- Golden Gate Basic Troubleshooting Tips
- Golden Gate Basic Tuning Tips
- Monitoring Golden Gate
- Questions/Discussion



What is Golden Gate

- Powerful data integration and replication solution
- One Way Replication data between one or more databases
- Multi-Master Replication
- Allows data transformation to be used to change/manipulate data through the replication
- Multi-Database Vendor Capable replication so we can replicate data between the databases in your environment.
 - MS SQL Server
 - DB2
 - TeraData
 - Sybase
 - MySQL
 - PostgreSQL



Why Use Golden Gate

- No More Streams in Oracle Database 12c
- No More Advanced Replication (Multi-Master Replication) in Oracle Database 12c
- No More Change Data Capture in Oracle Database 12c
- A Great Addition to a High Availability Environment
- Execute Near 0 downtime upgrades
- Execute Near 0 downtime maintenance capabilities
- Execute Near 0 downtime platform migrations
- Improve Live primary database performance by create custom reporting environments.
- Capable of Replicating Large volumes of data from highly active databases.



Golden Gate Architecture

- Source Database
 - Database where transactions are executed that need to be replicated to a destination.
- Destination/Target Database
 - Database where transactions from the source must be executed.
- Manager
 - Controls the starting and stopping of Golden Gate Processes
 - Trail File Management and Reporting
- Extract
 - Golden Gate Process on the source that captures/extracts executed transactions that are required to be executed on the target/destination.
- Local Trail
 - Local Trail is operating system files that are written by Golden Gate that contains the transactions extracted/captured from the Source database by the extract process.

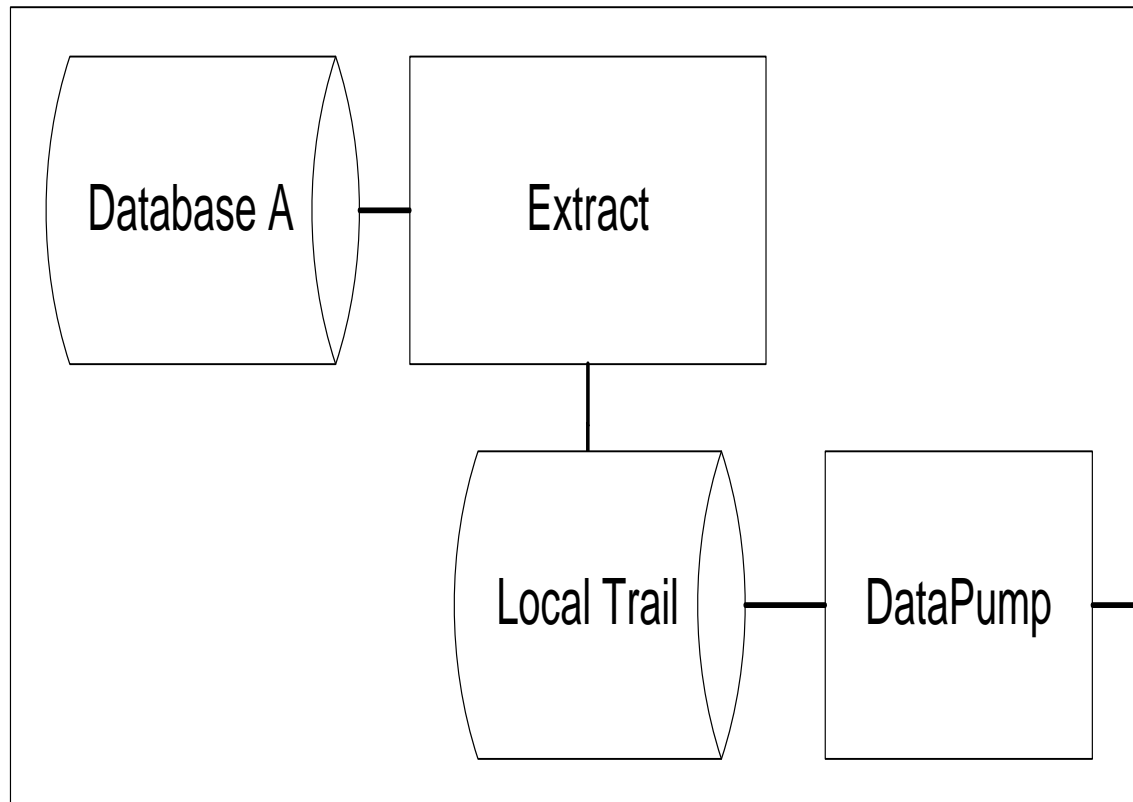


Golden Gate Architecture

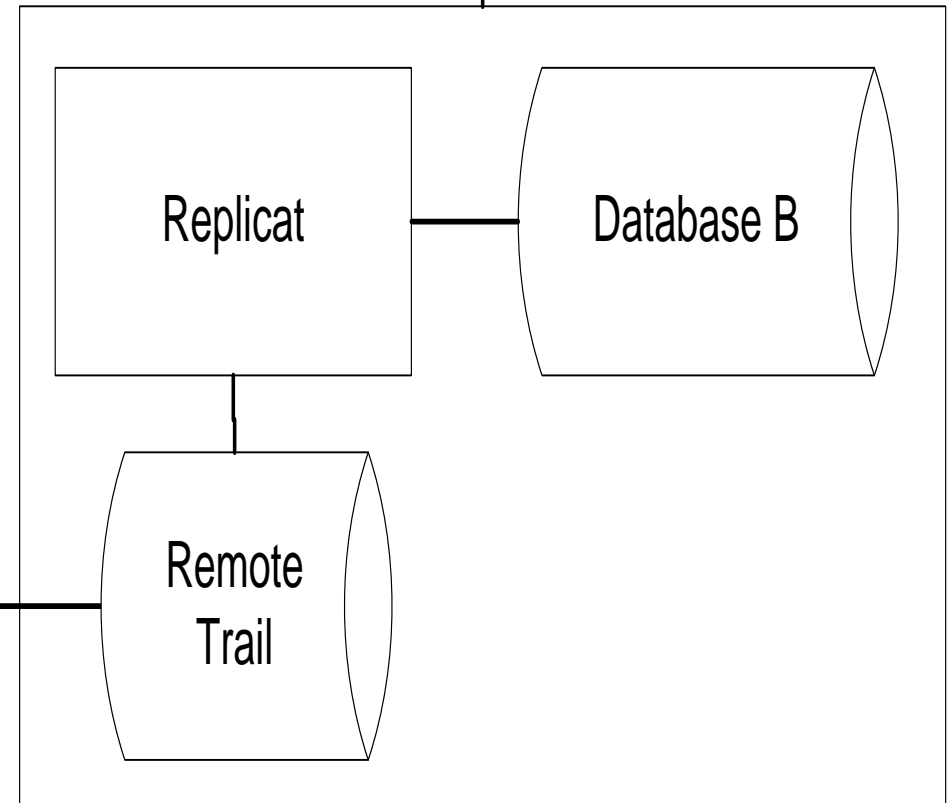
- DataPump
 - No Not Oracle Database DataPump which exports database objects to a file
 - Golden Gate Process that transfers transactions in the local trail to a remote trail on another server to be used to execute the transactions.
 - Optional really but a recommended best practice to use it.
 - Can take a single extract trail and send to multiple places by using multiple datapumps using same extract trail.
- Remote Trail
 - Golden Gate Operating System files that contain the transactions extracted/captured from the source database.
 - Written by the DataPump process for the replicat to read and execute the transactions.
- Replicat
 - Golden Gate Process on the target/destination that reads the transactions from the remote trail written by the DataPump and executes those transactions on the destination/target database.

Golden Gate Architecture

A - Source



B - Target





Golden Gate Architecture

- Classic Capture
 - Original method for extracting database changes
 - Read directly from online redo logs and archive redo logs
 - Will read directly from database when needed (SECUREFILE LOBs stored in-row, LOB with cache, lob data missing from redo log)
- Integrated Capture
 - Introduced with Golden Gate version 11.2.1.0.1
 - Uses Logmining server, either locally or in a downstream capture
 - Main difference is logmining server does the extract work and not reading directly from the online redo logs and archive redo logs.



Golden Gate Architecture

• Directory Structure

- The directory structure of Golden Gate it is recommended to follow the install much like other Oracle installed products by installing Golden Gate in the ORACLE_BASE location.

- ORACLE_BASE /u01/app/oracle
- Golden Gate Install location base /u01/app/oracle/gg
- Golden Gate Trail files /u01/app/oracle/gg/dirdat
- Golden Gate Report Files /u01/app/oracle/gg/dirrpt
- Golden Gate Process Parameter Files /u01/app/oracle/gg/dirprm



Golden Gate Architecture

• RAC

- Oracle RAC is a high availability solution for the Oracle database that introduces an additional component to the Golden Gate Architecture. A shared file system for the Golden Gate installation binaries as well as for all the trail files. Typically this is done with a reliable NFS mounted file system.
- /u01/app/oracle/gg Shared File System for Golden Gate Binaries
- /u01/app/oracle/gg/dirdat Shared File System for Golden Gate Trail files



Golden Gate Requirements/Limitations

• Disk Space

■ Installation

- The Golden Gate installation requires between 50M and 150M depending on platform
- recommended going with a 1G file system
 - > Allows for keeping report files
 - > patch staging as well as ability to backup the existing Golden Gate Installation for patch or upgrade

■ Trail Files

- calculated for an Oracle database from the amount of archive redo generated by the database over a period of time.
- Golden Gate only utilizes about 30% of the redo logs conservative 35%
- recommend that enough space be available to handle 3 days



Golden Gate Requirements/Limitations

- Replication Requirements
 - Supplemental Logging
 - Supplemental Logging for database
 - ALTER DATABASE ADD SUPPLEMENTAL LOG DATA ;
 - Primary/Unique Keys
 - Golden Gate expects that all tables have a primary or unique key define or at minimum a unique index.
 - Lack of primary key, unique key or unique index will cause Golden Gate to use all columns as key, key can be defined manually, but should have index to support key columns if defined manually. Suggest not letting Golden Gate Chose all columns in most circumstances.



Golden Gate Requirements/Limitations

● Requirements

- Network Connectivity
 - Adequate Network connectivity should exist between the source host and target host in a Golden Gate configuration
- Disable integrity constraints and triggers on target/destination database.
- Recommended to use checkpoint table
 - Replicat maintains checkpoints which are positions in the trail file which tells Replicat where to start after an expected or unexpected shutdown.
 - add checkpointtable
 - EDIT PARAMS ./GLOBALS
 - CHECKPOINTTABLE <ggschema>.CHKPT



Golden Gate Requirements/Limitations

• Limitations

- Unsupported Data Types - Oracle Golden Gate has several Data Types that are not supported by Golden Gate:
 - Abstract data types with scalar, LOBs, VARRAYS, nested tables and/or REFs
 - ORDDICOM
 - ANYDATA
 - ANYDATASET
 - ANYTYPE
 - BFILE
 - MLSLABEL
 - TIMEZONE_ABBR
 - TIMEZONE_REGION
 - URITYPE
 - UROWID



Golden Gate Requirements/Limitations

• Single Column Tables Limitations

- Supported except where the column is one of the following data types:
 - LOB
 - LONG
 - Nested Table
 - User defined data type
 - VARRAY
 - XML



Golden Gate Requirements/Limitations

- Limitations
 - Classic Capture
 - Only option when database version is less than 11.2.0.3
 - No Compressed tables for Classic Capture, however Extract can deliver to Compressed tables as of Golden Gate 11.1.1.1
 - Does not handle distributed XA transactions
 - How to Stop Creating Table DBMS_TABCOMP_TEMP_UNCMP (Doc ID 1326118.1) ABEND issue, auto restart can work around in most cases.
 - Integrated Capture (Golden Gate Version 11.2.1)
 - Database version 11.2.0.3 or above
 - Can handle HCC and OLTP Compressed tables (Doc ID 1266389.1)
 - Can Handle OLTP Compressed
 - Can Handle distributed XA transactions
 - Can be used for lower database versions using downstream capture using 11.2.0.3



Golden Gate one-way Replication

- Replace Streams
 - Streams as of Database 12c is a deprecated feature and therefore will need a replacement.
- Near 0 Downtime Platform Migration
 - Move to new Operating System
 - Move to new servers and/or new location
 - Move to Oracle from another database vendor
- Near 0 Downtime Upgrade
 - Skip database version
- Near 0 Downtime Maintenance
 - Create a second environment that can support normal load
 - is kept in sync, allow maintenance such as reorgs, index rebuilds, table partitioning, etc.
 - Switch to the copy that has had maintenance applied and run there.



Golden Gate Multi-master Replication

- Replace Advanced Replication
 - Also known as Multi-master or bi-directional replication
 - Data updates happening in multiple locations and synced between them
 - Advanced Replication as of Database 12c is a deprecated feature and therefore needs a replacement.
 - Must handle collisions even at a minimal level.
 - Recommended to divide data so that data being manipulated is more localized to minimize potential collisions.
 - Can be used for a 0 downtime migration where application servers are migrated to new environment over time and both are active at the same time therefore from an application perspective there was no downtime.

Golden Gate Extract, Transform, Load (ETL)



- Golden Gate allows the following
 - Mapping of tables/schemas
 - Mapping of columns
 - Filter of data based on conditions
 - Consolidation of table data from multiple sources to single target
- All this potentially allows you to load your data warehouse incrementally using Golden Gate.
- Can help build real-time data warehouse with streaming transactions from OLTP real-time as they happen, transform those transactions and put into the warehouse or feed to staging area.
- Can work as an addition to other ETL tools and processes.



Some Best Practices for Manager Process

- The manager process is the main controller for Golden Gate and therefore the best location for certain higher level parameters. It is a process there by default
- Parameters
 - PORT – Specifies the Port the manager process uses
 - > Port 7840
 - PURGEOLDEXTRACTS – Specifies the removal of trail files already processed so that old trail files do not consume space for new trail files.
 - > PurgeOldExtracts dirdat/*, UseCheckpoints, MinKeepDays 2
 - AUTOSTART - This will cause the processes to autostart and not require a manual start.
 - > For all processes AutoStart ER *
 - > For specific process AutoStart EXTRACT <extractname>

Some Best Practices for Manager Process



- Parameters
 - AUTORESTART - The auto restart will have the GoldenGate processes restart automatically if found to be stopped or abended. Autorestart tries to restart the processes two times after waiting two minutes inbetween. Every two hours, GoldenGate tries restart again. You can adjust these defaults if needed.
 - > AUTORESTART EXTRACT eproc

Some Best Practices for Extract Processes



- Reads from online redo logs, whenever possible do not let extract fall far enough behind to have to read archive logs, this can cause catch up to take longer if extract falls too far behind.
- Keep number of extracts from a single source to the minimum number required
 - As the number of extracts increases if Extract falls behind and has to read from archive logs then contention can increase and slow processes down the more extract processes there are
- Extract read all records and determines which ones are written in the trail via the extract parameters therefore use the capture method required to avoid issues.

Some Best Practices for Extract Processes



- Parameters
 - NLS_LANG - The NLS_LANG setting is an environment setting set using the SETENV of the parameter file, this is highly recommended to be specified to match the database NLS setting and not allowed to default in anyway.
 - SETENV (NLS_LANG=AMERICAN_AMERICA.WE8ISO8859P1)
 - USERID - In the extract process this would be the golden gate user login to the source database.
 - USERID ggadmin@destdb, PASSWORD passwd

Some Best Practices for Extract Processes

- Parameters
 - TRANLOGOPTIONS - When ASM is utilized in the extract database the GoldenGate local extract must log into the ASM instance to read the redo logs.
 - TranLogOptions ASMUser “sys@asm”, asmpassword “<encrypted password>”, ENCRYPTKEY default
 - ExtTrail - This specifies the location of the extract trail files and the leading characters of the trail files.
 - EXTTRAIL dirdat/ea
 - EXTTRAIL dirdat/processname/ea
 - Table - The table parameter defines the schema.table or schema.* for all tables in a schema to be replicated so that the extract process extracts the transactions for the associated table(s).
 - Table schema.table_name ;
 - Table schema.* ;

Some Best Practices for Extract Processes



- Parameters

- TableExclude - put in a table exclude for each of the tables that to be excluded from replicating.
 - TableExclude schema.<table_name> ;
- Sequences - The sequences parameter is utilized in extract processes to handle sequences. This parameter like the Table parameter can accept a wild card character (*) to indicate all sequences.
 - sequence <schema>.* ;
 - sequence <schama>.<sequencename> ;

Some Best Practices for Extract Processes



● Parameters

- REPORT - Utilized to insert runtime statistics into the report and report at specified time
 - Report at 02:30

- REPORTCOUNT - force GoldenGate to execute a report with runtime statistics every X minutes.
 - ReportCount Every 120 minutes, Rate

- REPORTROLLOVER - Replicat and Extract processes run continuously and over time report file can grow quite large. To ensure that the report does not get too large use the REPORTROLLOVER parameter to start a new report at a specific time.
 - ReportRollover at 02:00



Add an Extract Process

- Adding an extract process is done via the ggsci command line utility
 - Run the ggsci utility
 - ./ggsci
 - Add the extract process
 - add extract eprocc, tranlog, THREADS 2, begin now
 - add exttrail dirdat/eprocc/ec, extract eprocc
 - Edit the parameters to complete the process configuration
 - edit param eprocc

Some Best Practices for DataPump Processes

- Match at least one datapump per extract process.
- Use multiple datapumps when sending transactions to multiple target destinations.
- Use Datapump as a best practice when sending trail to remote host destination.

■ Extract1 -> Datapump1a (Destination 1) -> Replicat1

Datapump1b (Destination 2) -> Replicat1

Some Best Practices for DataPump Processes

- Parameters
 - PASSTHRU - Tells GoldenGate that there will be no filtering or column mapping between source and targets structures.
 - PASSTHRU
 - RMTHOST - Parameter indicates the target server for the trail files as well as the Manager Port on the target Server.
 - RMTHOST myserver.mydomain.com, MGRPORT 7840
 - RMTTRAIL - Parameter indicates the location where the trail files should be located on the target server.
 - dirdat/rproc/ra



Add the Datapump Process

- Adding a datapump is done via the ggsci command line .
 - Run the ggsci utility
 - `cd /u01/app/oracle/gg`
 - `./ggsci`
 - Add the datapump process as well as add the remote trail for the process
 - `add extract pprocc, extrailsources dirdat/eprocc/ec`
 - `add rmttrail dirdat/rprocc/rc, extract pprocc`
 - Edit the parameters to complete the process configuration
 - `edit param pprocc`

Some Best Practices for Replicat Processes



- At least 1 replicat process for a extract and datapump flow.
- For Heavy Transactional Systems multiple replicat processes for a single extract and pump flow.

- Extract1 -> Datapump1a (Destination 1) -> Replicat1a
Replicat1b

Datapump1b (Destination 2) -> Replicat1a
Replicat1b

- Extract2 -> Datapump2 -> Replicat2

Some Best Practices for Replicat Processes



- Parameters
 - NLS_LANG - The NLS_LANG setting is an environment setting set using the SETENV of the parameter file, this is highly recommended to be specified to match the database NLS setting and not allowed to default in anyway.
 - SETENV (NLS_LANG=AMERICAN_AMERICA.WE8ISO8859P1)
 - USERID - In the replicat process this would be the golden gate user login to the target database.
 - USERID ggadmin@destdb, PASSWORD passwd
 - HandleCollisions – This is used to handle inserts that violate primary keys and deletes where the record is not found
 - HandleCollisions

Some Best Practices for Replicat Processes



● Parameters

- AssumeTargetDefs - This parameter is utilized to tell GoldenGate to assume that the target table definition is the same as the source table definitions. Not used in ETL Type Replicat processes.
 - AssumeTargetDefs
- Map - The Map command in the replicat configuration is where the map of a table in the source to the table in the target
 - Map schema.table_name, Target schema.table_name
 - Map schema.*, Target schema.*
- REPORT - Utilized to insert runtime statistics into the report and report at specified time
 - Report at 02:30

Some Best Practices for Replicat Processes



- Parameters

- REPORTCOUNT - force GoldenGate to execute a report with runtime statistics every X minutes.
 - ReportCount Every 120 minutes, Rate
- REPORTROLLOVER - Replicat and Extract processes run continuously and over time report file can grow quite large. To ensure that the report does not get too large use the REPORTROLLOVER parameter to start a new report at a specific time.
 - ReportRollover at 02:00

Some Best Practices for Replicat Processes



- Parameters

- DISCARDFILE - The replicat could have issues with records replicating them to the target for various reasons, it is best to make a record of these so that they can be handled and easily identified for resolution.
 - > DiscardFile <gginstall>/dirrpt/r1.dsc, Append
- DISCARDROLLOVER - To ensure the discard file that contains discard records from the replication does not grow to large.
 - > DiscardRollover at 02:30 ON SUNDAY



Add Replicat Process

- Adding a replicat is done via the ggsci command line utility which is located in the Golden Gate Binary installation location for example /u01/app/oracle/gg
 - Run the ggsci utility
 - cd /u01/app/oracle/gg
 - ./ggsci
 - Add the replicat process specifying the trail the replicat will read from
 - add replicat rprocc, EXTTRAIL dirdat/rprocc/r1
 - Edit the parameters to complete the process configuration
 - edit param rprocc



Golden Gate Basic Troubleshooting Tips

- Check Golden Gate process status
 - Info all
 - Info <golden gate process name>
- Golden Gate Process Status
 - STOPPED – Reported when process is not running and was stopped normally.
 - RUNNING – Reported when process is running normally.
 - ABENDED – Reported with process is not running and stopped due to error
- Golden Gate error log
 - Records issues, good first place to look when process abends
 - Located in the Golden Gate installation location
 - ggserr.log

Golden Gate Basic Troubleshooting Tips



• LAG

Program	Status	Group	Lag at Chkpt	Time Since Chkpt
MANAGER	RUNNING			
EXTRACT	RUNNING	DP3	00:00:03	00:00:00
EXTRACT	RUNNING	EXT3	00:00:00	00:00:00

- How far behind is Golden Gate from live source activity?
- Determined through info <processname> or info all
- Time Since Chkpt— How long has it been since process did a checkpoint
- Lag at Chkpt - How far behind was the process when last checkpoint was taken
- Total Lag (how far behind total) is Time Since Chkpt + Lag at Chkpt



Golden Gate Basic Tuning Tips

- Utilize as few Golden Gate Extracts as required for a database
 - Golden Gate Extracts when using several extracts adds overhead for reading transactions from redo/archive redo.
 - Too many extracts can cause extracts to contend with each other causing extracts to actually be slower than fewer extract processes. Esp. if Golden Gate get behind and has to read from Archive Logs.
- Increasing number of Replicats can help improve transaction apply on target database.
 - Multiple replicats can read from same remote trail.



Golden Gate Basic Tuning Tips

- Adjust TCP buffer
 - Default TCP socket buffer size is 30,000 bytes
 - TCPBUFSIZE can be increased to help traffic transfer
 - Rmthost <targethostname>, MgrPort 7809, TCPBUFSIZE 100000
- COMPRESS
 - Can compress traffic from Golden Gate to Remote host by using the COMPRESS in the
 - Compress ratios of 4:1 and potentially more
 - Adds CPU overhead to local and remote hosts
 - Rmthost <targethostname>, MgrPort 7809, TCPBUFSIZE 100000, COMPRESS



Golden Gate Basic Tuning Tips

- BATCHSQL
 - Can improve Replicate by as much as 300%
 - Only valid for Oracle, SQL Server, DB2 and Teradata databases
 - Groups smaller transactions together which allows processing at faster rate
 - Replicate parameter example: BATCHSQL
 - New section on report file, Check batches in error every error indicates a slowdown to the benefits of BATCHSQL
 - Check Golden Gate error log for OGG-01137 BATCHSQ suspending, continuing in normal mode, many of these indicate the loss of the benefits of BATCHSQL



Golden Gate Basic Tuning Tips

- GROUPTRANSOPS
 - Controls the grouping of smaller transactions from BATCHSQL
 - Replicat minimum of 1000 SQL operations in a transaction
 - Can increase this value to improve performance
 - Reduces number of checkpoints written to checkpoint table
 - Replicat parameter example:
 - GROUPTRANSOPS 2000

Monitoring Golden Gate



- Golden Gate Plug-in for Oracle Enterprise Manager 12c
 - The Oracle Enterprise Manager 12c Cloud Control is a comprehensive monitoring solution that can cover monitoring of Golden Gate configuration.
 - The Oracle Management Pack for Oracle Golden Gate license must be acquired.



Monitoring Golden Gate

- Oracle GoldenGate Monitor
 - Separate component that can be setup to monitor GoldenGate.
 - The GoldenGate Monitor consists of the GoldenGate Monitor Agent, GoldenGate Monitor Server and GoldenGate Monitor Repository therefore will require additional server resources to utilize.

*** Note: The Oracle GoldenGate Monitor can be downloaded from the Oracle Technology Network*



Monitoring Golden Gate

- Manual Automated Monitoring Processes
 - Manually monitor main processes of Golden Gate
 - extract
 - datapump
 - replicat
 - Done through scheduled scripted processes that check Golden Gate Status for each of the main processes.
 - Monitor how far behind (LAG) a monitored Golden Gate Process has.

Questions/Discussion

THANK YOU

