Informatica PC Training

Day-8

Agenda:

- Performance Tuning
- Troubleshooting

Debadatta Mohanty

Housekeeping Tips

- Please mute your phone during the presentation.
- If there is too much noise, participants will be put on auto-mute.
- We shall open up the table for Q&A at the end of the session.
- Please feel free to post your questions over Chat as well.
- This session will be recorded and an email will be sent with links to the recordings after the session.
- At the end of the course, TEX will request you to provide feedback on the training.

Performance Tuning in Informatica

Performance Tuning

- First step in performance tuning is to identify the performance bottleneck in the following order:
  - Target
  - Source
  - Mapping
  - Session
  - System
- The most common performance bottleneck occurs when the Integration Service writes to a target database.

Target Bottleneck

- Identifying
  - A target bottleneck can be identified by configuring the session to write to a flat file target.
  - Read the thread statistics in the session log. When the Integration Service spends more time on the writer thread than the transformation or reader threads, you have a target bottleneck.
- Optimizing
  - Dropping Indexes and Key Constraints before loading.
  - Increasing commit intervals.
  - Increase the database network packet size.
  - Use of Bulk Loading / External Loading.

Source Bottleneck

- Identifying
  - Add a filter condition after Source qualifier to false so that no data is processed past the filter transformation. If the time it takes to run the new session remains about the same, then there is a source bottleneck.
  - In a test mapping remove all the transformations and if the performance is similar, then there is a source bottleneck.
  - Database query

- Optimizing
  - Optimizing the Query by using hints.
  - Use Informatica Conditional Filters if the source system lacks indexes.
  - Increase the database network packet size.
  - Create Index in Source System if it doesn't have indexes

Mapping Bottleneck

- Identifying
  - If there is no source bottleneck, add a Filter transformation in the mapping before each target definition. Set the filter condition to false so that no data is loaded into the target tables. If the time it takes to run the new session is the same as the original session, there is a mapping bottleneck.
  - Analyze performance counters. High errorrows and rows inlookupcache counters indicate a mapping bottleneck.
  - Read the thread statistics and work time statistics in the session log. When the Integration Service spends more time on the transformation thread than the writer or reader threads, you have a transformation bottleneck.
- Optimizing
  - Configure for Single-Pass reading
  - Avoid unnecessary data type conversions.
  - Avoid database reject errors.
  - Use Shared Cache / Persistent Cache
  - Delete unused ports from the pipeline or Transformations

Session Bottleneck

- Identifying
  - If there is no source, Target or Mapping bottleneck, then there may be a session bottleneck.
  - Use Collect Performance Details. Any value other than zero in the readfromdisk and writetodisk counters for Aggregator, Joiner, or Rank transformations indicate a session bottleneck. Low (0-20%) BufferInput_efficiency and BufferOutput_efficiency counter values also indicate a session bottleneck.

- Optimizing
  - Increase the number of partitions.
  - Use incremental Aggregation if possible.

Incremental Aggregation

- First Run creates idx and dat files.
- Second Run performs the following actions:
- For each i/p record, the Server checks historical information in the index file for a corresponding group, then:
  - If it finds a corresponding group, it performs the aggregate operation incrementally, using the aggregate data for that group, and saves the incremental change
  - If it does not find a corresponding group, it creates a new group and saves the record data.
- When writing to the target Integration Service
  - Updates modified aggregate groups in the target
  - Inserts new aggregate data
  - Deletes removed aggregate data
  - Ignores unchanged aggregate data
  - Saves modified aggregate data in the index and data files

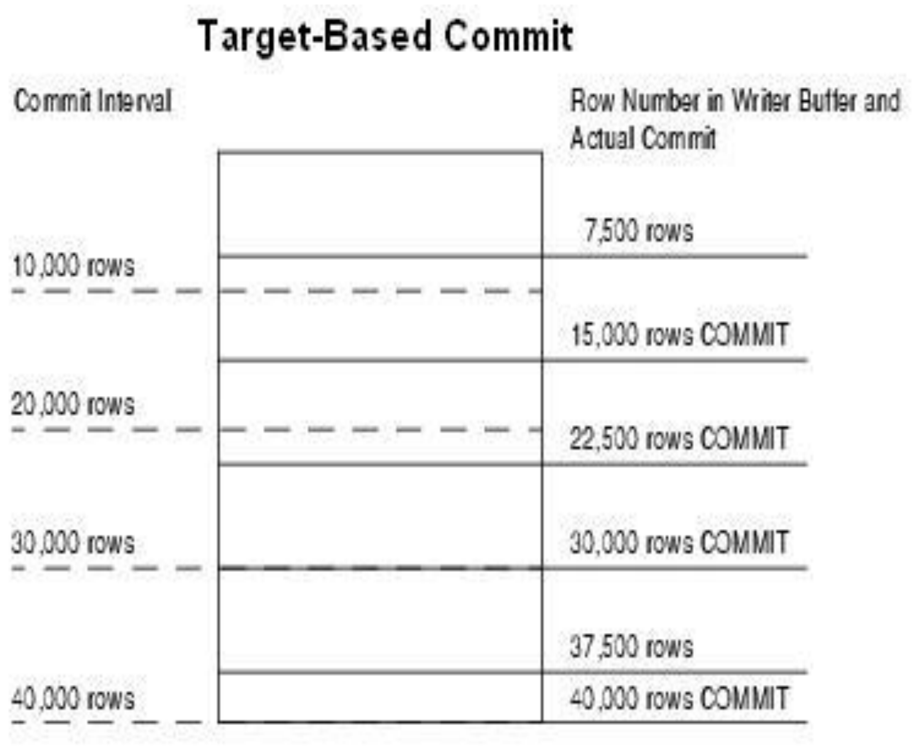
System Bottlenecks

- Identifying
  - If there is no source, Target, Mapping or Session bottleneck, then there may be a system bottleneck.
  - Use system tools to monitor CPU usage, memory usage, and paging.
  - On Windows :- Task Manager
  - On Unix – Systems toots like sar, iostat. For E.g. sar –u (%usage on user, idle time, i/o waiting time)
- Optimizing
  - Improve network speed.
  - Improve CPU performance
  - Check hard disks on related machines
  - Reduce Paging
  - Use multiple CPUs

Commit Points

- A commit interval is the interval at which the Integration Service commits data to relational targets during a session
- The commit point can be a factor of the commit interval, the commit interval type, and the size of the buffer blocks
- The commit interval is the number of rows you want to use as a basis for the commit point
- The commit interval type is the type of rows that you want to use as a basis for the commit point
- Can choose between the following types of commit interval
  - Target-based commit
  - Source-based commit
  - User defined Commit
- During a source-based commit session, the Integration Service commits data to the target based on the number of rows from an active source in a single pipeline

Commit Points



During a target-based commit session, the Informatica Service continues to fill the writer buffer after it reaches the commit interval

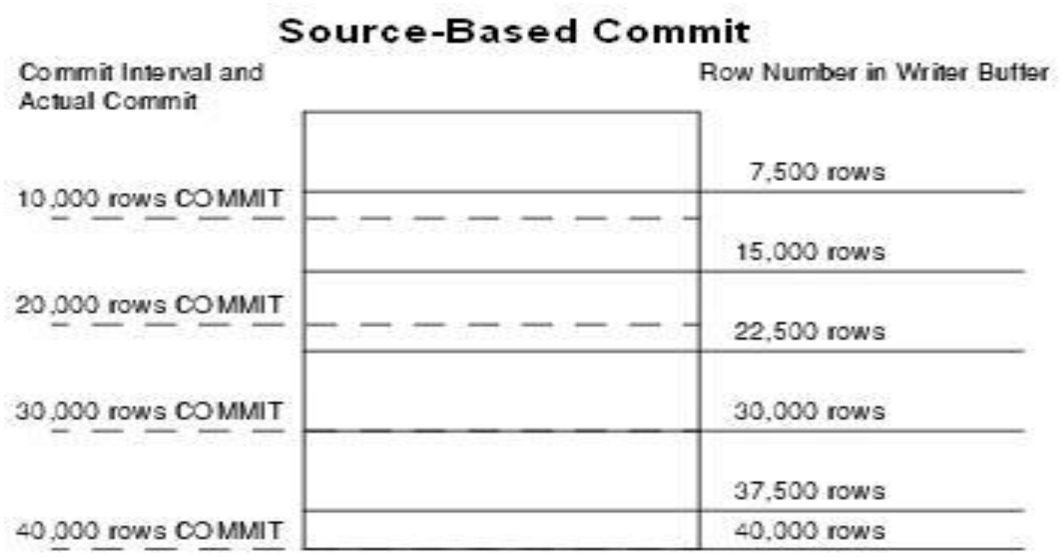
When the buffer block is filled, the Integration Service issues a commit command

As a result, the amount of data committed at the commit point generally exceeds the commit interval

Commit Points

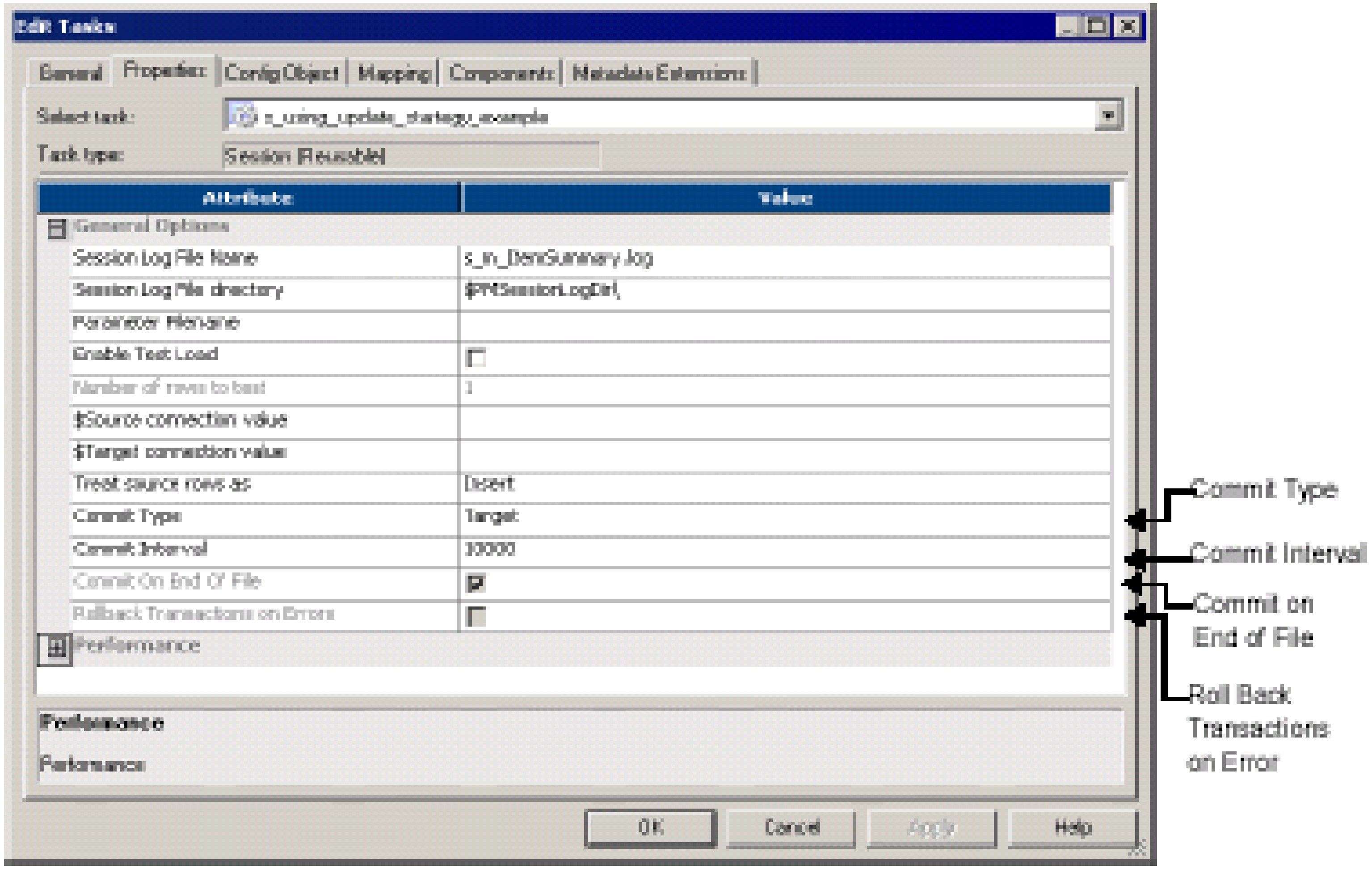
- During a source-based commit session, the Informatica Service commits data to the target based on the number of rows from an active source in a single pipeline
- These rows are referred to as source rows
- A pipeline consists of a source qualifier and all the transformations and targets that receive data from the source qualifier
- An active source can be any of the following active transformations:
  - Advanced External Procedure
  - Source Qualifier
  - Normalizer
  - Aggregator
  - Joiner
  - Rank
  - Mapplet, if it contains one of the above transformations.

Commit Points



- When the Integration Service runs a source-based commit session, it identifies the active source for each pipeline in the mapping
- The Integration Service generates a commit row from the active source at every commit interval
- When each target in the pipeline receives the commit row, the Integration Service performs the commit

Commit Points



Select task: s_using_update_strategy_example

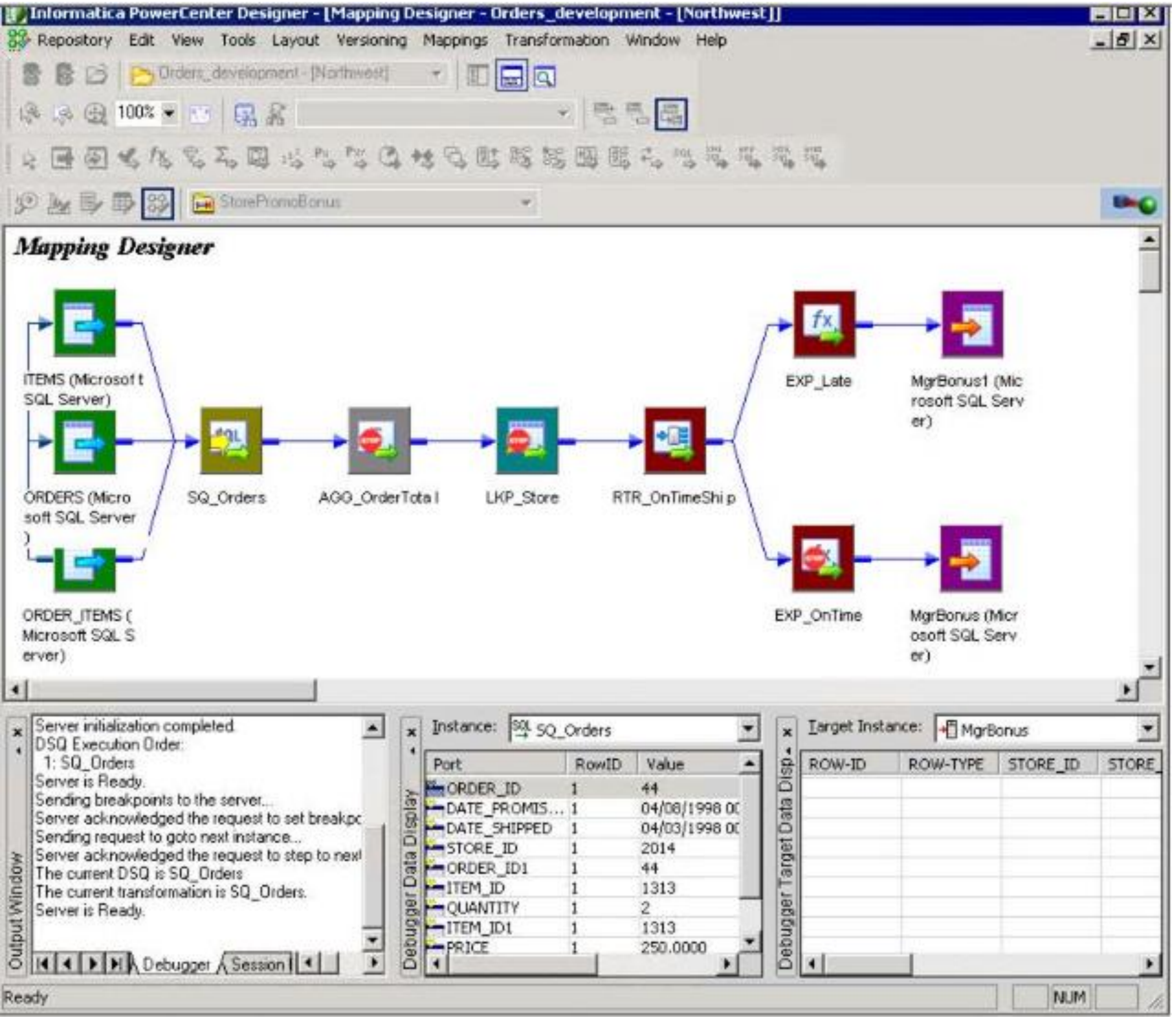
Task type: Session (Reusable)

Attribute	Value
General Options	
Session Log File Name	s_m_DeniSummary.log
Session Log File directory	$PMSessionLogDir\
Parameter Filename	
Enable Test Load	☐
Number of rows to test	1
$Source connection value	
$Target connection value	
Treat source rows as	Insert
Commit Type	Target
Commit Interval	10000
Commit On End Of File	☑
Rollback Transactions on Errors	☐
Performance	

Performance

Performance

Troubleshooting


Informatica PowerCenter Designer - [Mapping Designer - Orders_development - [Northwest]]
Repository Edit View Tools Layout Versioning Mappings Transformation Window Help
Orders_development - [Northwest]
100%
StorePromoBonus
Mapping Designer
ITEMS (Microsoft SQL Server)
ORDERS (Microsoft SQL Server)
ORDER_ITEMS (Microsoft SQL Server)
SQ_Orders
AGG_OrderTotal
LKP_Store
RTR_OnTimeShip
EXP_Late
EXP_OnTime
MgrBonus1 (Microsoft SQL Server)
MgrBonus (Microsoft SQL Server)
Output Window
Debugger Data Display
Debugger Target Data Display
Debugger
Session
Ready
NUM

```
Server initialization completed.
DSQ Execution Order:
  1: SQ_Orders
Server is Ready.
Sending breakpoints to the server...
Server acknowledged the request to set breakpc
Sending request to goto next instance...
Server acknowledged the request to step to next
The current DSQ is SQ_Orders.
The current transformation is SQ_Orders.
Server is Ready.
```

Instance: SQ_Orders

Port	RowID	Value
ORDER_ID	1	44
DATE_PROMIS...	1	04/08/1998 0C
DATE_SHIPPED	1	04/03/1998 0C
STORE_ID	1	2014
ORDER_ID1	1	44
ITEM_ID	1	1313
QUANTITY	1	2
ITEM_ID1	1	1313
PRICE	1	250.0000

Target Instance: MgrBonus

ROW-ID	ROW-TYPE	STORE_ID	STORE_

<u>Debugger Overview:</u> You can debug a valid mapping to gain troubleshooting information about data and error conditions. To debug a mapping, you configure and run the Debugger from within the Mapping Designer. The Debugger uses a session to run the mapping on the Integration Service. When you run the Debugger, it pauses at breakpoints and you can view and edit transformation output data.

Refer Manual for further Troubleshooting Tips

Feedback

If you have any suggestion regarding the presentation or training, want to suggest inclusion of any topic etc., please mail to debadatta.mohanty@hcl.com