Optimizing Session Caches in PowerCenter
Abstract

The PowerCenter Integration Service allocates cache memory for XML targets and Aggregator, Joiner, Lookup, Rank, and Sorter transformations in a mapping. For optimal session performance, configure the cache sizes so that the PowerCenter Integration Service can run the complete transformation in memory. This article describes how the default auto cache mode works, how to use the cache calculator to estimate the cache sizes, and how to analyze the session log to determine the optimal cache sizes.

Supported Versions

- PowerCenter 9.1.0 - 9.6.1

Table of Contents

Overview .......................................................... 2
Cache Size .......................................................... 3
  Auto Cache Size ............................................. 3
  Calculate the Cache Size ................................... 4
  Numeric Cache Size .......................................... 7
Cache Size Increased by the PowerCenter Integration Service ......... 7
Cache Size for Partitioned Caches ................................ 8
  Configure the Cache Size for Partitioned Caches .......... 8
Cache Size Optimization ......................................... 8
  Step 1. Set the Tracing Level to Verbose Initialization .... 9
  Step 2. Run the Session ....................................... 9
  Step 3. Analyze Caching Performance ....................... 9
  Step 4. Configure Specific Cache Sizes .................... 9

Overview

When you run a session that uses an XML target or an Aggregator, Joiner, Lookup, Rank, or Sorter transformation, the PowerCenter Integration Service creates caches in memory to run the transformation. You can configure the cache sizes for these transformations. The cache size determines how much memory the PowerCenter Integration Service allocates for each transformation cache at the start of a session run.

If the cache size is larger than the available memory on the machine, the PowerCenter Integration Service cannot allocate enough memory and fails the session run.

If the cache size is smaller than the amount of memory required to run the transformation, the PowerCenter Integration Service processes some of the transformation in memory and stores overflow values in cache files to process the rest of the transformation. When the service pages cache files to the disk, processing time increases. For optimal performance, configure the cache size so that the PowerCenter Integration Service can process the complete transformation in memory.

By default, the PowerCenter Integration Service automatically configures the memory requirements at run time, based on the maximum amount of memory that the service can allocate to transformation caches for the session. You can use the cache calculator to estimate the total amount of memory required to run the transformation. You provide inputs to calculate the cache size for each transformation.
After you run a session using auto cache mode or using calculated cache sizes, you can tune the cache sizes for the transformations. You analyze the transformation statistics in the session log to determine the cache sizes required for optimal performance, and then configure numeric values for the cache sizes.

**Cache Size**

Cache size determines how much memory the PowerCenter Integration Service allocates for each transformation cache at the start of a session run. Configure the cache sizes in the session properties. The cache sizes specified in the session properties override the values set in the transformation properties.

XML targets and Aggregator, Joiner, Lookup, and Rank transformations require an index cache and a data cache. The PowerCenter Integration Service stores key values in the index cache and output values in the data cache. You configure both the index and data cache sizes for these transformations. Sorter transformations require a single cache. The PowerCenter Integration Service stores sort keys and the data to be sorted in the Sorter cache.

If the session is reusable, all instances of the session use the cache sizes configured in the reusable session properties. You cannot override the cache sizes in the session instance.

Use one of the following methods to configure a cache size:

**Auto cache mode**

Use auto memory to specify a maximum limit on the cache size that is allocated for processing the transformation. Use this method if the machine on which the PowerCenter Integration Service process runs has limited cache memory.

**Cache calculator**

Use the cache calculator to estimate the total amount of memory required to process the transformation based on your input.

**Numeric value**

Configure a specific value for the cache size. Configure a specific value when you tune the cache size.

**Auto Cache Size**

By default, cache size is set to Auto. The PowerCenter Integration Service automatically configures the cache memory requirements at run time. You define the maximum amount of memory that the service can allocate for all transformations that use auto cache mode in a single session.

To set the maximum cache memory for transformations in auto cache mode, configure the following session properties on the Config Object tab:

**Maximum Memory Allowed for Auto Memory Attributes**

Maximum amount of memory to allocate for the session cache. The PowerCenter Integration Service allocates memory from the session cache to all transformations with cache size set to Auto. The default unit is bytes. Append KB, MB, or GB to the value to specify other units. For example, 1048576 or 1024 KB or 1 MB.

**Maximum Percentage of Total Memory Allowed for Auto Memory Attributes**

Percentage of machine memory to allocate for the session cache. The PowerCenter Integration Service allocates memory from the session cache to all transformations with cache size set to Auto.
The following image shows the session properties that define the maximum cache memory:

When you run a session in auto cache mode, the PowerCenter Integration Service calculates the maximum percentage of memory and compares that against the maximum amount of memory that you specify. Then it allocates the lower amount of memory to transformations in auto cache mode. If multiple transformations are in auto cache mode, the PowerCenter Integration Service allocates the memory to all transformations in auto cache mode.

For example, the machine that hosts the PowerCenter Integration Service has 1 GB of memory. You set the Maximum Memory Allowed for Auto Memory Attributes property to 800 MB. You also set the Maximum Percentage of Total Memory Allowed for Auto Memory Attributes property to 10%. The PowerCenter Integration Service allocates 102.4 MB of memory to the session cache and divides the cache memory among all transformations in auto cache mode.

The maximum session cache size that you set affects only transformations with cache mode set to Auto. The PowerCenter Integration Service allocates memory separately to transformations for which you configure a specific cache size. If a session has multiple transformations that require caching, you can set the cache mode for some transformations to Auto and specify a cache size for other transformations. The PowerCenter Integration Service allocates the memory specified for transformations in auto cache mode in addition to the memory it allocates to transformations configured with numeric cache sizes.

For example, a session has three transformations that require caching. You set two transformations to auto cache mode and specify a maximum memory cache size of 800 MB for the session. You also specify a cache size of 500 MB for the third transformation. The Integration Service allocates a total of 1,300 MB of memory.

### Calculate the Cache Size

Use the cache calculator to estimate the cache size based on your input. The PowerCenter Integration Service allocates the calculated amount of memory to the transformation cache at the start of the session run.

The cache calculator requires different inputs for each transformation. You must select the applicable cache type to apply the calculated cache size. For example, to apply the calculated cache size for the data cache and not the index cache, select only the Data Cache Size option.

**Note:** You cannot use the cache calculator to estimate the cache size for an XML target.
Calculating the Cache Size for a Transformation

Use the cache calculator to estimate the total amount of memory required to process an Aggregator, Joiner, Lookup, Rank, or Sorter transformation.

1. In the Workflow Manager, open the session.
2. Click the **Mapping** tab.
3. Select the transformation in the left pane. The right pane of the Mapping tab shows the transformation properties where you can configure the cache size.

4. For the data or index cache size property, click the **Open** button to open the cache calculator. The **Cache Calculator** dialog box appears.

5. Select the **Calculate** mode to calculate the total memory requirement for the transformation.
6. Provide the input based on the transformation type.
If the input value is too large and you cannot enter the value in the cache calculator, use auto cache mode.

The following table describes the input that you provide for each transformation type:

<table>
<thead>
<tr>
<th>Transformation</th>
<th>Option Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>Data Movement Mode</td>
<td>The data movement mode of the PowerCenter Integration Service. The cache requirement varies based on the data movement mode. Each ASCII character uses one byte. Each Unicode character uses two bytes.</td>
</tr>
<tr>
<td>Aggregator</td>
<td>Number of Groups</td>
<td>Number of groups. The Aggregator transformation aggregates data by group. Calculate the number of groups using the group by ports. For example, if you group by Store ID and Item ID, you have 5 stores and 25 items, and each store contains all 25 items, then calculate the number of groups as: 5 * 25 = 125 groups.</td>
</tr>
<tr>
<td>Joiner</td>
<td>Number of Master Rows</td>
<td>Number of rows in the master source. Applies to a Joiner transformation with unsorted input. The number of master rows does not affect the cache size for a sorted Joiner transformation. <strong>Note:</strong> If rows in the master source share unique keys, the cache calculator overestimates the index cache size.</td>
</tr>
<tr>
<td>Lookup</td>
<td>Number of Rows with Unique Lookup Keys</td>
<td>Number of rows in the lookup source with unique lookup keys.</td>
</tr>
<tr>
<td>Rank</td>
<td>Number of Groups</td>
<td>Number of groups. The Rank transformation ranks data by group. Determine the number of groups using the group by ports. For example, if you group by Store ID and Item ID, have 5 stores and 25 items, and each store has all 25 items, then calculate the number of groups as: 5 * 25 = 125 groups.</td>
</tr>
<tr>
<td>Rank</td>
<td>Number of Ranks</td>
<td>Number items in the ranking. For example, if you want to rank the top 10 sales, you have 10 ranks. The cache calculator populates this value based on the value set in the Rank transformation.</td>
</tr>
<tr>
<td>Sorter</td>
<td>Number of Rows</td>
<td>Number of rows.</td>
</tr>
</tbody>
</table>

7. Click **Calculate**.

   The cache calculator calculates the cache sizes in kilobytes.

8. If the transformation has a data cache and index cache, select **Data Cache Size**, **Index Cache Size**, or both.

9. Click **OK** to apply the calculated values to the cache sizes that you selected.

**Calculating the Cache Size for an XML Target**

You cannot use the cache calculator to estimate the cache size for an XML target. You can use formulas to calculate an estimated cache size for an XML target.

To calculate the cache size, perform the following tasks:

1. Estimate the number of rows in each group.
2. Use the following formula to calculate the cache size for each group:

\[
\text{Group cache size} = \text{Data cache size} + \text{Primary key index cache size} + \text{Foreign key index cache size}
\]

The following equation shows how to calculate the size of the data cache for a group:

\[
\text{(Number of rows in a group)} \times \text{(Row size of the group)}
\]

The following equation shows how to calculate the size of the primary key index cache for a group:

\[
\text{(Number of rows in a group)} \times \text{(Primary key index cache size)}
\]

The following equation shows how to calculate the size of the foreign key index cache for a group:

\[
\text{Sum} \left( \text{(Number of rows in parent group)} \times \text{(Foreign key index cache size)} \right)
\]

3. Use the following formula to calculate the total cache size:

\[
\text{Total cache size} = \text{Sum(Cache size of all groups)}
\]

**Numeric Cache Size**

You can configure a numeric value for a cache size. The PowerCenter Integration Service allocates the specified amount of memory to the transformation cache at the start of the session run. Configure a numeric value when you tune the cache size.

The first time that you configure a cache size, use auto cache mode or use the cache calculator. After you run the session, analyze transformation statistics in the session log to determine the cache sizes required to process the transformations in memory.

When you configure the cache size to use the value specified in the session log, you can ensure that no allocated memory is wasted. However, the optimal cache size varies depending on the size of the source data. Review the session logs after subsequent session runs to monitor changes to the cache size. If you configure a numeric cache size for a reusable transformation, verify that the cache size is optimal for each use of the transformation in a session.

To define numeric values for cache sizes, configure the cache sizes on the Mapping tab in the session properties.

**Cache Size Increased by the PowerCenter Integration Service**

The PowerCenter Integration Service creates each memory cache based on the configured cache size. In some situations, the PowerCenter Integration Service might increase the configured cache size because it requires more cache memory.

The PowerCenter Integration Service might increase the configured cache size for one of the following reasons:

- **The configured cache size is less than the minimum cache size required to process the operation.** The PowerCenter Integration Service requires a minimum amount of memory to initialize each session. If the configured cache size is less than the minimum required cache size, then the PowerCenter Integration Service increases the configured cache size to meet the minimum requirement. If the PowerCenter Integration Service cannot allocate the minimum required memory, the session fails.

- **The configured cache size is not a multiple of the cache page size.** The PowerCenter Integration Service stores cached data in cache pages. The cached pages must fit evenly into the cache. Thus, if you configure 10 MB (1,048,576 bytes) for the cache size and the cache page size is 10,000 bytes, then the PowerCenter Integration Service increases the configured cache size to 1,050,000 bytes to make it a multiple of the 10,000-byte page size.

When the PowerCenter Integration Service increases the configured cache size, it continues to run the session and writes a message similar to the following message in the session log:

```
MAPPING> TE 7212 Increasing [Index Cache] size for transformation <transformation name> from <configured index cache size> to <new index cache size>.
```
Cache Size for Partitioned Caches

When you create a session with multiple partitions, the PowerCenter Integration Service might use cache partitioning for the Aggregator, Joiner, Lookup, Rank, and Sorter transformations.

When the PowerCenter Integration Service partitions a cache, it creates a separate cache for each partition and allocates the configured cache size to each partition. The PowerCenter Integration Service stores different data in each cache, where each cache contains only the rows needed by that partition. As a result, the PowerCenter Integration Service requires a portion of total cache memory for each partition.

When the PowerCenter Integration Service uses cache partitioning, it accesses the cache in parallel for each partition. If it does not use cache partitioning, it accesses the cache serially for each partition.

The following table describes the situations when the PowerCenter Integration Service uses cache partitioning for each applicable transformation:

<table>
<thead>
<tr>
<th>Transformation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregator</td>
<td>You create multiple partitions in a session with an Aggregator transformation. You do not have to set a partition point at the Aggregator transformation.</td>
</tr>
<tr>
<td>Joiner Transformation</td>
<td>You create a partition point at the Joiner transformation.</td>
</tr>
<tr>
<td>Lookup Transformation</td>
<td>You create a hash auto-keys partition point at the Lookup transformation.</td>
</tr>
<tr>
<td>Rank Transformation</td>
<td>You create multiple partitions in a session with a Rank transformation. You do not have to set a partition point at the Rank transformation.</td>
</tr>
<tr>
<td>Sorter Transformation</td>
<td>You create multiple partitions in a session with a Sorter transformation. You do not have to set a partition point at the Sorter transformation.</td>
</tr>
</tbody>
</table>

Configure the Cache Size for Partitioned Caches

You configure the memory requirements differently when the PowerCenter Integration Service uses cache partitioning. If the PowerCenter Integration Service uses cache partitioning, it allocates the configured cache size for each partition.

To configure the memory requirements for a transformation with cache partitioning, calculate the total requirements for the transformation and divide by the number of partitions.

For example, you create four partitions in a session with an Aggregator transformation. You determine that an Aggregator transformation requires 400 MB of memory for the data cache. Configure 100 MB for the data cache size for the Aggregator transformation. When you run the session, the PowerCenter Integration Service allocates 100 MB for each partition, using a total of 400 MB for the Aggregator transformation.

Use the cache calculator to calculate the total estimated requirements for the transformation. If you use dynamic partitioning, you can determine the number of partitions based on the dynamic partitioning method. If you use dynamic partitioning based on the nodes in a grid, the PowerCenter Integration Service creates one partition for each node. If you use dynamic partitioning based on the source partitioning, use the number of partitions in the source database.

Cache Size Optimization

For optimal session performance, configure the cache sizes so that the PowerCenter Integration Service can run the complete transformation in memory.

To configure optimal cache sizes, perform the following tasks:

1. Set the tracing level to verbose initialization.
2. Run the session using auto cache mode or using calculated cache sizes.
3. Analyze caching performance in the session log.
4. Configure specific values for the cache sizes.

**Step 1. Set the Tracing Level to Verbose Initialization**

In the Workflow Manager, set the tracing level to verbose initialization to enable the PowerCenter Integration Service to write transformation statistics to the session log. The transformation statistics list the cache sizes required for optimal performance. By default, the tracing level is set to normal.

Set the tracing level on the Config Object tab in the session properties.

**Step 2. Run the Session**

The first time that you run the session, use auto cache mode or use calculated cache sizes.

You can run the entire workflow that contains the session task. Or, you can run just the session task.

**Step 3. Analyze Caching Performance**

After you run the session using auto cache mode or using calculated cache sizes, analyze the transformation statistics in the session log to determine the cache sizes required for optimal performance.

When an Aggregator, Joiner, Lookup, or Rank transformation pages to the disk, the session log specifies the index and data cache sizes required to run the transformation in memory. For example, you run an Aggregator transformation called AGG_TRANS. The session log contains the following text:

```
INFO: MAPPING, CMN_1791, The index cache size that would hold [1098] aggregate groups of input rows for [AGG_TRANS], in memory, is [286720] bytes
INFO: MAPPING, CMN_1790, The data cache size that would hold [1098] aggregate groups of input rows for [AGG_TRANS], in memory, is [1774368] bytes
```

The log shows that the index cache requires 286,720 bytes and the data cache requires 1,774,368 bytes to run the transformation in memory without paging to the disk.

When a Sorter transformation pages to the disk, the session log states that the PowerCenter Integration Service made multiple passes on the source data. The PowerCenter Integration Service makes multiple passes on the data when it has to page to the disk to complete the sort. The message specifies the number of bytes required for a single pass, which is when the PowerCenter Integration Service reads the data once and performs the sort in memory without paging to the disk.

For example, you run a Sorter transformation called SRT_TRANS. The session log contains the following text:

```
INFO: TRANSF_1_1_1, SORT_40427, Sorter Transformation [SRT_TRANS] required 2-pass sort (1-pass temp I/O: 13126221824 bytes). You might try to set the cache size to 14128 MB or higher for 1-pass in-memory sort.
```

The log shows that the Sorter cache requires 14,128 MB so that the PowerCenter Integration Service makes one pass on the data.

**Step 4. Configure Specific Cache Sizes**

For optimal performance, configure the cache sizes to use the values specified in the session log. Update the index and data cache size session properties.

1. In the Workflow Manager, open the session.
2. Click the Mapping tab.
3. Select the transformation in the left pane.
   The right pane of the **Mapping** tab shows the transformation properties where you can configure the cache size.

4. Enter the values that the session log recommended for the index and data cache sizes.
   When you enter a value, all values are in bytes by default. However, you can enter a value and specify one of the following units: KB, MB, or GB. If you enter the units, do not enter a space between the value and unit. For example, enter 350000KB, 200MB, or 1GB.
   The following image shows a session that has specific values configured for the index and data cache sizes for an Aggregator transformation:

5. Click **OK**.

**Author**

Alison Taylor  
Principal Technical Writer