How-To Library

Informatica

Tuning the Performance of the Monitoring Model Repository

Abstract

The monitoring Model repository is a relational database instance. The monitoring Model Repository Service monitors the Data Integration Service jobs, and stores the statistics in the monitoring Model repository. This article discusses the methods that you can use to improve monitoring Model repository performance.

Supported Versions

- Big Data Quality 10.2.1
- Big Data Management 10.2.1

Table of Contents

Overview

The monitoring Model Repository Service manages the monitoring Model repository. The service monitors statistics for ad hoc jobs, applications, logical data objects, SQL data services, web services, and workflows. The service then saves the statistics in the monitoring Model repository.

When you install Informatica products, you can configure a Model repository and a monitoring Model repository. When you configure the Model repositories, use separate database user accounts for monitoring Model repository and Model repository. You can also configure these repositories after the installation.

In Informatica Administrator, configure a monitoring Model repository at the domain level. After you configure the Monitoring Configuration parameters, you can view the job status in Informatica Administrator, Informatica Developer, and Informatica Analyst. In the Administrator tool, the Monitor tab displays the statistics and reports about the objects in the domain. You can use the Administrator tool or the infacmd command line program to administer the monitoring Model Repository Service.

Based on your requirement, you can configure one or more Model repositories to store metadata created by Informatica products. Informatica Developer, Informatica Analyst, Data Integration Service, and the Administrator tool store metadata in the Model repository. The Model Repository Service manages the Model repository.

Configuring Monitoring Model Repository Service

Specify a Model repository when you configure the monitoring settings to store run-time statistics about objects that the Data Integration Services run.

1. Log in to Informatica Administrator.
2. Navigate to the **Manage** > **Services and Nodes** view.
3. In the **Domain Navigator**, select the domain.
4. In the **Domain** view, click the **Monitoring Configuration** view.

   The default monitoring configuration parameters appear.
5. Click the **Edit** icon.

   The **Monitoring Configuration** dialog box appears.
6. You can edit the following options depending on your requirements:

Option	Description
Model Repository Service	Name of the monitoring Model repository that stores the historical information. The monitoring Model repository must not be integrated with a version control system.
Username	User name to access the monitoring Model Repository Service. Does not appear in domains that use Kerberos authentication.
Password	Password to access the monitoring Model Repository Service. Does not appear in domains that use Kerberos authentication.
Modify Password	Modify the monitoring Model Repository Service password.
Security Domain	Name of the security domain that the monitoring Model repository user belongs to.
Preserve Summary Historical Data	Number of days that the monitoring Model repository saves averaged data. If purging is disabled, then the monitoring Model repository saves the data indefinitely. Default is 180. Minimum is 0. Maximum is 366.
Preserve Detailed Historical Data	Number of days that the monitoring Model repository saves per-minute data. If purging is disabled, then the monitoring Model repository saves the data indefinitely. Default is 14. Minimum is 1. Maximum is 14.
Purge Statistics Every	Interval, in days, at which the monitoring Model Repository Service purges data that is older than the values configured in the Preserve Historical Data option. Default is 1 day.
Days At	Time of day when the monitoring Model Repository Service purges statistics. Default is 1:00 a.m.
Maximum Number of Sortable Records	Maximum number of records that can be sorted in the Monitor tab. If the number of records on the Monitor tab is greater than this value, then you can only sort by Start Time and End Time. Default is 3,000.

Option	Description
Maximum Delay for Update Notifications	Maximum time, in seconds, that the Data Integration Service buffers statistics before it stores them in the monitoring Model repository and displays them in the Monitor tab. If the Data Integration Service shuts down unexpectedly before it stores the statistics in the monitoring Model repository, then the statistics are lost. Default is 10.
Date Time Field	Include milliseconds for date and time fields in the **Monitor** tab.

7. Click **OK**.

To apply the settings, you must restart all of the Data Integration Services.

Recommendations

Consider the following guidelines when you configure or manage a monitoring Model Repository Service:

Database

- Use separate database user accounts for the Model repository and monitoring Model repository.
- As a best practice, configure the monitoring Model repository database instance on a remote server with reliable network connectivity between the database and the Informatica Server. The required minimum network bandwidth is 1 Gbps, and the recommended network bandwidth is 10 Gbps.
- As a general rule, consider that 1,000 jobs in Informatica Big Data Management might persist 400 MB of data in the database. Therefore, you might configure the monitoring Model repository database to accommodate 40 GB of data for around 100,000 jobs.
- If you modify the configuration of the monitoring Model Repository Service, you must restart the Data Integration Service for the changes to take effect.

Node

- Configure the monitoring Model Repository Service on the node where you configure the domain so that the monitoring Model Repository Service is on the same network as that of the Informatica product installation. This arrangement eliminates network latency that might occur when statistics are persisted in the monitoring Model repository.
- It is recommended that you configure the Informatica domain, Data Integration Service, Model Repository Service, and monitoring Model Repository Service on the same node because if you configure the domain and application services on different nodes, it might impact performance.

Purge monitoring data

- You can configure the **Purge Statistics Every** property in the Monitoring Configuration to purge statistics at regular intervals, or you can use the infacmd mrs DeleteContents command to delete the content in the monitoring Model repository.
- It is preferred to retain the content for 24 hours before you purge the content, therefore configure the **Purge Statistics Every** property to 1 day.

Monitoring Model Repository Uses Microsoft SQL Server Database

- To configure the isolation level for the database, run the `ALTER DATABASE DatabaseName SET ALLOW_SNAPSHOT_ISOLATION ON` and `ALTER DATABASE DatabaseName SET READ_COMMITTED_SNAPSHOT ON` commands.
  To verify whether the allow snapshot isolation and read committed isolation level are set, run the `select name, snapshot_isolation_state_desc, is_read_committed_snapshot_on from sys.databases` query.

  For more information about isolation level, see *Informatica Big Data Suite 10.2.1 Installation and Configuration Guide*.
- To update the statistics, run the infacmd mrs updateStatistics command.
  For more information about the infacmd command, see *Informatica 10.2.1 Command Reference*.

Other recommendations

- If you have configured high availability, perform regular backup of monitoring Model repository.

Guidelines to Improve the Performance

The monitoring Model repository is a relational database instance managed by the monitoring Model Repository Service. Over time, the monitoring Model repository performance might slow down for many reasons.

To improve the performance of the monitoring Model repository, you can perform the following tasks to improve the performance of the monitoring Model repository:

- Manage and purge repository statistics.
- Manage and purge log entries.
- Configure connections and memory settings to tune the Model repository.
- Configure the settings in Informatica clients and databases.
- Increase the **Maximum Heap Size** property to 2 GB.

Manage and Purge Statistics

To improve monitoring Model repository performance, you can control the amount of statistics that the monitoring Model repository retains and the amount of time that it retains them.

When you configure monitoring Model Repository Service in the domain, the Data Integration Services store statistics and reports in the monitoring Model repository. As mappings, workflows, and other objects run, the Data Integration Service Manager persists, updates, retrieves, and publishes run-time statistics in the monitoring Model repository. The statistics include historical information about objects that the Data Integration Services run. You can view the statistics on the **Monitor** tab in the Administrator tool.

To preserve optimal performance, the most important property to set is the **Purge Statistics Every** property. Set this property to purge the statistics at regular intervals. For faster performance, you can also limit the number of days that the monitoring Model repository preserves summary and detailed historical data and limit the number of sortable historical records.

The following table lists the Monitoring options and recommended settings that you can set that can improve performance:

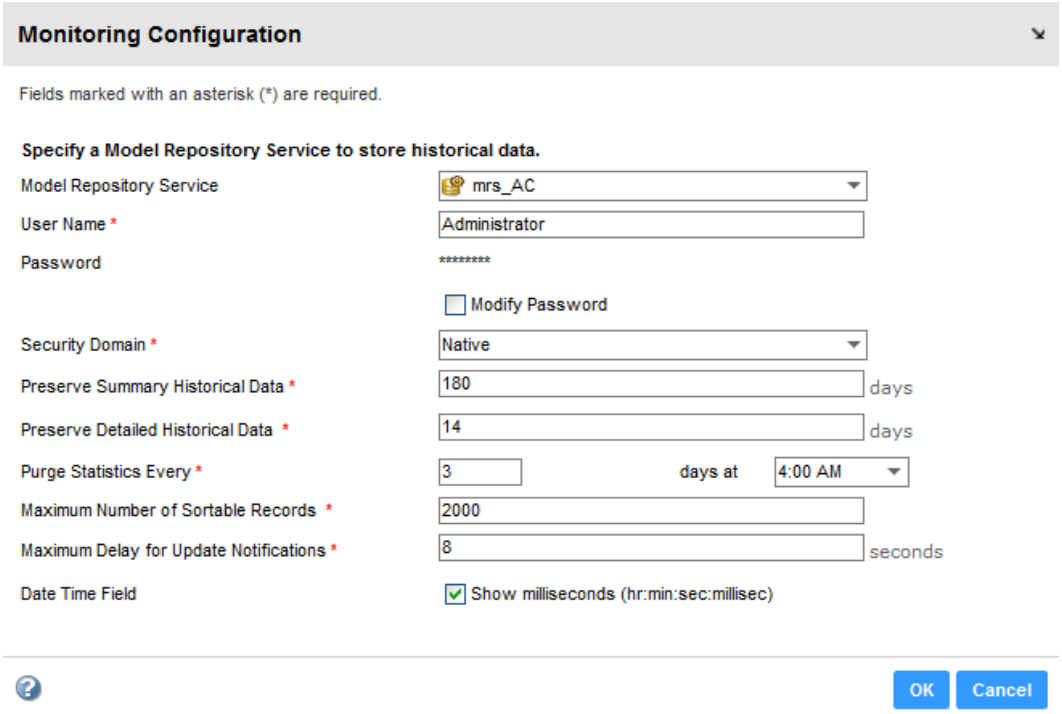
Option	Recommended Action
Preserve Summary Historical Data	Based on your requirement, you can reduce the setting.
Preserve Detailed Historical Data	Based on your requirement, you can reduce the setting.
Purge Statistics Every	Do not increase the interval for optimal performance. The default value is 1 day.
Days At	Choose a time of day when demands on the system are lowest for optimal performance.
Maximum Number of Sortable Records	Based on your requirement, you can reduce the setting.
Show Milliseconds	Clear the option for faster performance.

Managing and Purging Statistics

The monitoring Model repository saves historical data and sortable records for a period of time that you set and purges them at intervals that you set.

1. In Informatica Administrator, click **Monitor** > **Services and Nodes** > **Domain Navigator** view.
2. Click the **Monitoring Configuration** tab.

   The current monitoring configuration appears.
3. Click the **Edit** icon to change the monitoring configuration.



The **Monitoring Configuration** window opens.

4. Edit the **Purge Statistics Every** property to purge statistics at regular intervals.
5. Edit the other monitoring properties to preserve optimal performance, click **OK**, and then click **Save**.

To apply the settings, you must recycle all the Data Integration Services.

Manage and Purge Log Entries

To improve the monitoring Model repository performance, you can control the number of log entries that the monitoring Model repository retains and the amount of time that it retains them.

The Data Integration Service Manager accumulates log events for the domain, application services, and users. The log events contain operational and error messages for a domain. The Service Manager and the application services send log events to the Log Manager. When the Log Manager receives log events, it generates log event files. You can view service log events in the Administrator tool.

To maintain optimal performance, you can set the monitoring Model repository to control the automatic purge of log events. You can also purge log events manually.

The following table describes the log management options that you can set and that affect performance:

Option	Recommended Action
Preserve logs for number of days	Based on your requirements, you can reduce the setting.
Maximum size for logs in MB	Based on your requirements, you can reduce the amount of storage space.

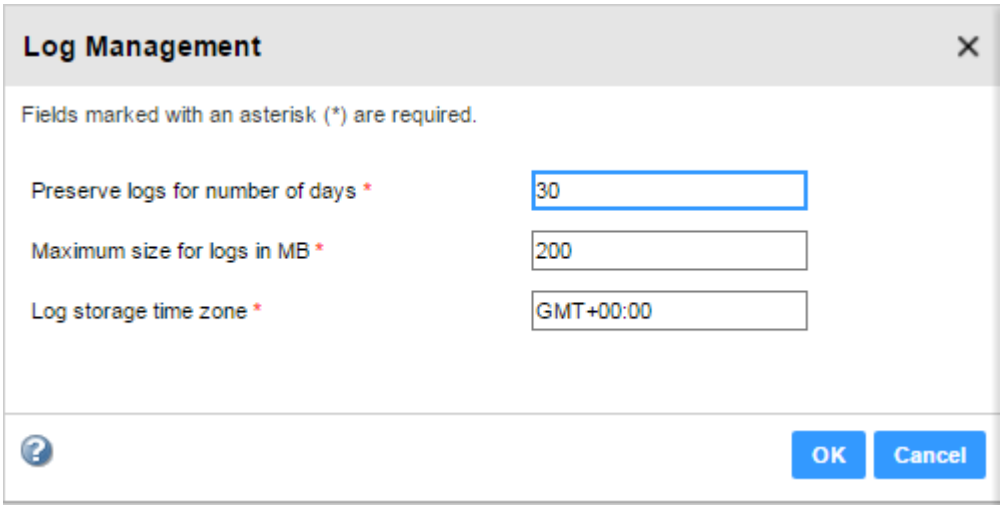
Managing and Purging Log Entries

The monitoring Model repository accumulates log events for the domain, application services, and users. The log events contain operational and error messages for a domain. To improve performance, you can set the monitoring Model repository to control the automatic purge of log events.

To maintain optimal performance, limit the size of logs and the number of days to retain them.

1. In the Administrator tool, click the **Logs** tab.
2. Select **Actions** > **Log Management**.

   The **Log Management** dialog box appears.



3. To configure the monitoring Model repository to automatically purge log entries more frequently, set one or both of the following properties:
    - Reduce the value of the **Preserve logs for number of days** property.
    - Reduce the value of the **Maximum size for logs in MB** property.
4. Click **OK**.

Recycle the monitoring Model Repository Service for the changes to take effect.

Tune the Monitoring Model Repository

Heap size and connection pool properties can affect monitoring Model repository performance.

The effect on performance when you tune the monitoring Model repository parameters depends on usage and the configuration settings. When a particular setting is causing a bottleneck, then increasing the value of that property might fix the issue. For example, you can increase the Java heap size to increase the performance. However, the Java heap size of the monitoring Model repository is limited by the amount of physical memory on a machine, and setting the parameter to exceed the physical memory available could cause a problem. Consult with system administrators when you plan to tune these parameters.

You can tune the following settings to improve the monitoring Model repository performance:

- Maximum Heap Size property
- Java Stack Size property
- Memory settings
- Hibernate Connection Pool Size property

Setting the Maximum Heap Size Property

You can increase the Maximum Heap Size property to increase monitoring Model repository performance.

Maximum heap size is the amount of RAM allocated to the Java Virtual Machine (JVM) that runs the monitoring Model Repository Service. The default value is 1 GB.

You can increase this property to increase the monitoring Model repository performance. Generally, Informatica recommends that you set this property at approximately two times the size of the monitoring Model repository. For example, if the Model repository consumes 1 GB of disk space, set this property to 2 GB.

You can also increase this property when you have a large pool of concurrent users. Monitoring Model repository users include users of the Developer tool and the Analyst tool. Because mappings launch the Data Integration Service at run time, and the Data Integration Service connects to the monitoring Model repository, you can also include in the user pool mappings that run automatically.

Note: A value that exceeds the physical memory on the machine that hosts the monitoring Model repository might cause an issue. Do not exceed the available physical memory.

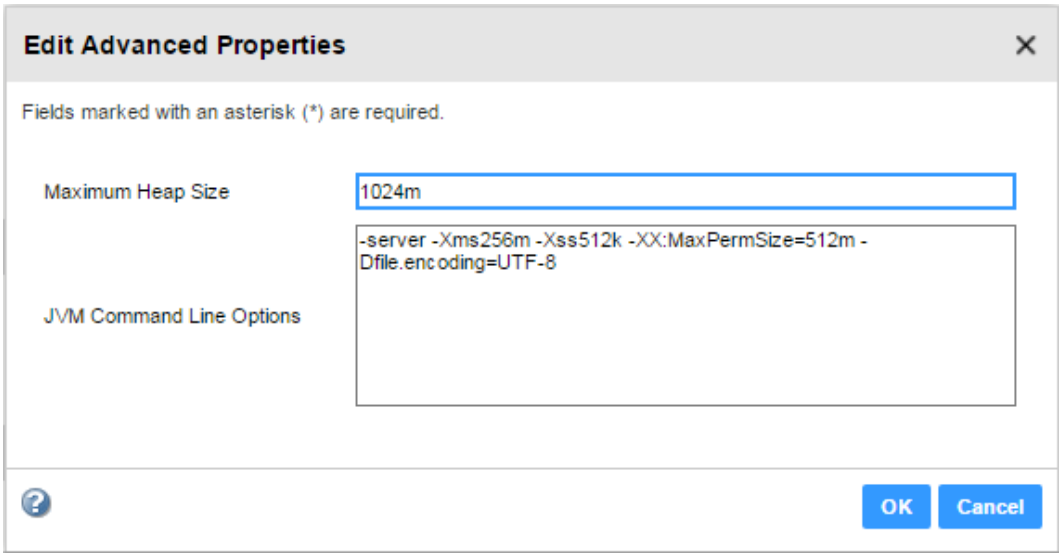
The following table lists guidelines for this property based on the number of concurrent users:

Number of Concurrent Users	Max Heap Size Value
Single user	1 GB
< 10	2 GB

Number of Concurrent Users	Max Heap Size Value
10-50	4 GB
> 50	8 GB

1. In the Administrator tool, click the **Domain** tab.
2. In the **Domain Navigator**, select the monitoring Model Repository Service.
3. Click the **Edit** icon in the Advanced Properties section.

   The **Edit Advanced Properties** dialog box appears.



4. Increase the **Maximum Heap Size** parameter.
5. Click **OK**.

Recycle the monitoring Model Repository Service for the changes to take effect.

Setting the Java Stack Size Property

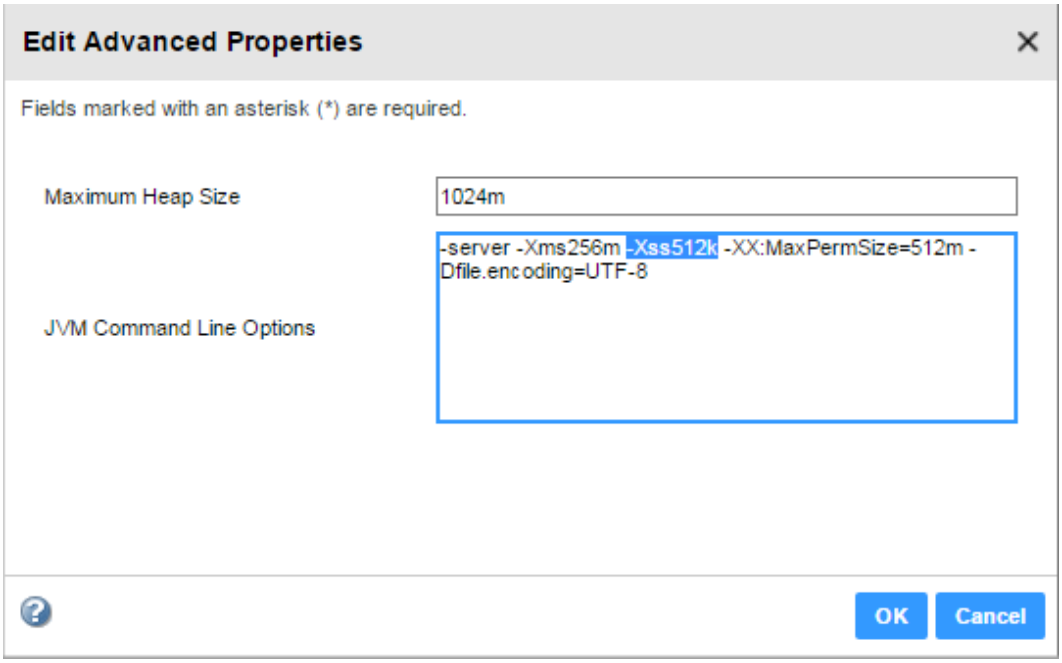
You can increase the Java Stack Size property to increase monitoring Model repository performance. The Java stack size is the size limit of each Java thread in the Java Virtual Machine (JVM) that runs the monitoring Model Repository Service. The default value is 512K. You can increase this property to make more memory available to monitoring Model Repository Service processes.

Note: As with other tunable parameters, you should consult system administrators before changing this setting. The amount of available memory is limited by the physical memory of the machine.

1. In the Administrator tool, click the **Domain** tab.
2. In the **Domain Navigator**, select the monitoring Model Repository Service.
3. Click the **Edit** icon in the Advanced Properties section.

   The **Edit Advanced Properties** dialog box appears.

The following image shows the `-Xss` property highlighted in the **Edit Advanced Properties** dialog box:



4. In the **JVM Command Line Options** window, increase the value of the `-Xss` parameter to set the Java stack size property. The `-Xss` option expresses the Java stack size.
5. Click **OK**.

Recycle the monitoring Model Repository Service for the changes to take effect.

Configuring Cache Memory Settings

You can configure the monitoring Model repository to use cache memory. When you configure the monitoring Model repository to use cache memory, the monitoring Model Repository Service stores objects in memory. Access to objects in memory is faster than access to objects in the repository database.

To configure the amount of memory allocated to cache, you can update the **Cache JVM Options** parameter. The **Cache JVM Options** parameter uses the -Xmx option to set the maximum heap size for the monitoring Model Repository Service. When the number of objects in the monitoring Model repository is very large, you can increase the maximum heap size to increase performance. The default property value is `-Xmx128m`. Change the integers to 256 or another multiple of 128 to increase the JVM cache.

Note: The Model Repository Service cache process runs as a separate process. The Java Virtual Manager (JVM) that runs the Model Repository Service is not affected by the JVM options you configure for the Model Repository Service cache.

1. In the Administrator tool, navigate to the **Manage** > **Services and Nodes** > **Domain Navigator** section.
2. Select the Model Repository Service.
3. In the **Cache Properties** section, click the **Edit** icon.

The following image shows the **Edit Cache Properties** dialog box:

4. In the **Edit Cache Properties** dialog box, specify the amount of memory allocated to cache in the **Cache JVM Options** field. Click **OK.**

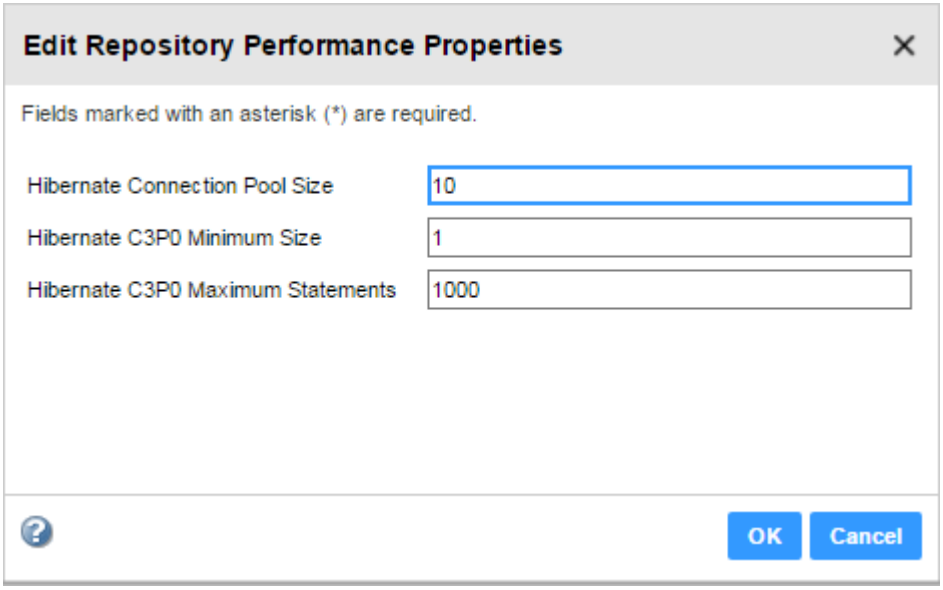
Recycle the Model Repository Service for the changes to take effect.

Setting the Performance Properties for the Model Repository Service Process

In deployments with many concurrent connections, you can configure performance tuning properties for storage of data objects in the monitoring Model Repository Service. The Model Repository Service uses an open source object-relational mapping tool called Hibernate to map and store data objects and metadata to the Model repository database. For each service process, you can set Hibernate options to configure connection and statement pooling for the Model repository.

1. In the Administrator tool, click the **Domain** tab.
2. In the **Domain Navigator**, select the monitoring Model Repository Service.
3. Click the **Processes** view.
4. Click **Edit** in the **Repository Performance Properties** section.

The **Edit Repository Performance Properties** dialog box appears.



5. Change the properties as required.
6. Click **OK**.

Restart the monitoring Model Repository Service for the changes to take effect.

Improve Client Interactions

You can configure external settings in Informatica clients and databases to improve the speed at which they interact with the monitoring Model repository.

To improve the monitoring Model repository performance, you can perform the following tasks:

Reduce the number of open connections.

When developers have a large number of projects and folders open in the Developer tool, the connection time between the Developer tool and the monitoring Model repository might degrade.

To mitigate this issue, instruct developers to open a limited number of projects and folders. Ideally, each developer opens only the project and folder in which they are developing objects, and closes the folder when the development task ends or pauses.

Note: You cannot use this workaround if Informatica is installed in a Citrix virtual environment.

Tune Oracle database settings.

When you use Oracle as the monitoring Model repository database, you can tune the **open_cursors** parameter to increase performance. The **open_cursors** parameter specifies the number of open cursors for a session. The primary purpose of the parameter is to prevent an excessive number of open cursors. Cursor use increases when the Data Integration Service deploys applications. When you increase the number of open cursors, you can increase performance.

Allocate additional memory to the Developer tool.

When you deploy applications from the Developer tool, the Deployment Manager fetches all of the objects in the application to Developer tool memory for validation before it deploys the application. When you increase the memory available to this process, the process runs faster.

To allocate additional memory to this process, edit the **-Xmx** parameter in the Developer tool **DeveloperCore.ini** file.

Note: The maximum allowed setting is 1500 MB. If you enter a higher number, the Developer tool will not start.

Allocate additional memory to the Data Integration Service process.

The Data Integration Service deploys and persists the runtime statistics in the monitoring Model repository. When you allocate more memory to this process, the process runs faster.

To allocate additional memory to this process, edit the **Maximum Heap Size** property in the Advanced Properties for the Data Integration Service process.

Allocating Additional Memory to the Developer Tool

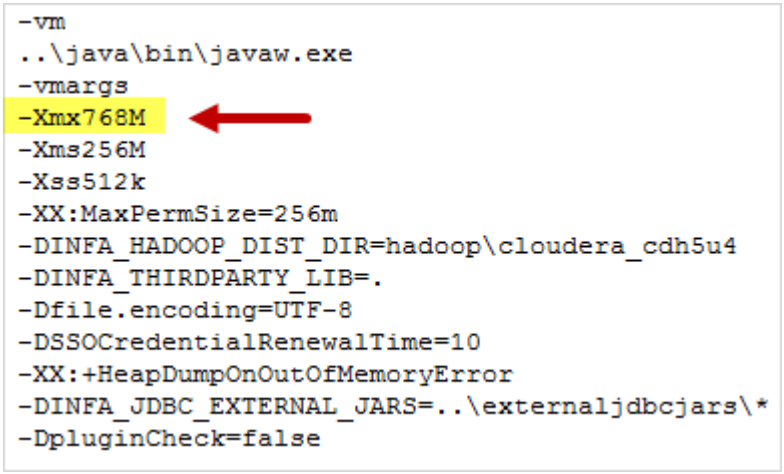
You can increase the maximum Java heap size to allot more memory to the Developer tool. When the Developer tool has more memory to use, it validates applications faster.

1. Browse to the following location:

   `<InformaticaInstallationDir>\clients\DeveloperClient`

2. Edit the DeveloperCore.ini file in a text editor.
3. Edit the line beginning with the characters `-Xmx` to increase the maximum Java heap size.

   For example, to set the Maximum Java heap to 2 GB, edit the parameter to be `-Xmx2G`.
   The following image shows sample text in a DeveloperCore.ini file. Adjust the highlighted memory parameter:



4. Save and close the DeveloperCore.ini file.

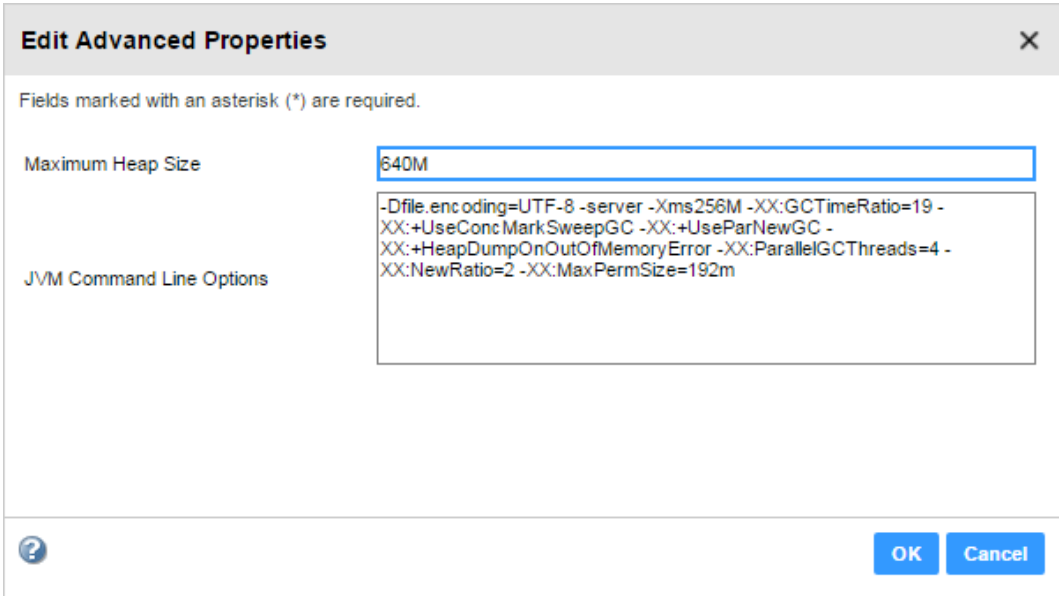
Restart the Developer tool for the changes to take effect.

Allocating Additional Memory to the Data Integration Service Process

You can make more memory available to the Data Integration Service process. When this process has more memory to use, it runs mappings, workflows, and applications faster.

1. In the Administrator tool, click the **Services and Nodes** tab.
2. In the **Domain Navigator**, select the Data Integration Service.
3. Click the **Processes** view.
4. Click **Edit** in the Advanced Properties section.

The **Edit Advanced Properties** dialog box appears.



5. Increase the value of the **Maximum Heap Size** parameter.
6. Click OK.

Restart the Data integration Service for the changes to take effect.

Frequently Asked Questions

What options can I use to configure the retention period of monitoring data in the monitoring Model repository?

You can configure the **Preserve Summary Historical Data** and **Preserve Detailed Historical Data** options to retain the monitoring data as necessary. You can also configure the **Purge Statistics Every** option to purge the data.

What is the criteria to configure the retention period of monitoring data?

You can configure the retention period based on your requirement. Performance improves when the retention period is low. Therefore, it is recommended that you configure the **Purge Statistics Every** option to 1 day or as required.

I am using the same Model Repository Service to handle metadata and run-time statistics. I see performance issues after I upgraded to version 10.2.1. What do I do to resolve this issue?

Some issues might occur when you use the same Model repository to store metadata and run-time statistics. In version 10.2.1, you need to use separate database user accounts for Model repository and monitoring Model repository. Configure a Model Repository Service to handle metadata, and a monitoring Model Repository Service to handle run-time statistics.

To separate the Model Repository Service, use one of the following methods:

- Purge existing statistics before upgrade. Run the command infacmd.sh isp purgemonitoringdata to purge the statistics. Open the domain logs to view the progress of the purging job. After the purge operation is complete, configure a monitoring Model Repository Service in the **Monitoring Configuration** view.
- Retain existing statistics and upgrade. After upgrade, you can configure a monitoring Model Repository Service in the **Monitoring Configuration** view.

Is there a upgrade impact if I purge statistics before an upgrade?

Upgrade is considerably faster if you purge the monitoring data before an upgrade. The upgrade process resets the Model Repository Service maximum heap size to 4 GB. After the upgrade, you can reset the maximum heap size property to the value to which it was set prior to the upgrade.

To reset the maximum heap size, select the Model Repository Service in the **Domain Navigator**, click the **Properties** view, and expand **Advanced Properties**. Set the **Maximum Heap Size** property to the value before upgrade. Set the **MaxMetaspaceSize** property to the minimum of 512 MB.

References

Based on your requirements, you can refer to the following guides to know more about monitoring Model Repository Service:

- Install monitoring Model Repository Service. See *Informatica Big Data Suite 10.2.1 Installation and Configuration Guide*.
- Configure monitoring Model Repository Service. See *Informatica 10.2.1 Application Service Guide* .
- View monitoring data. See *Informatica 10.2 .1 Administrator Guide* .

Author

Lavanya S