

IBM Integration Bus V9 Good operating practices  
Version 9 Release 0

*Configuration recommendations to reduce  
memory usage*



**Note**

Before using this information and the product it supports, read the information in "Notices" on page 9.

---

# Contents

<b>Chapter 1. Introduction . . . . .</b>	<b>1</b>	<b>Notices. . . . .</b>	<b>8</b>
<b>Chapter 2. Deploying message flows to integration servers. . . . .</b>	<b>3</b>	Programming interface information . . . . .	10
<b>Chapter 3. Configuring JVM settings for the integration server . . . . .</b>	<b>5</b>	Trademarks . . . . .	10
<b>Chapter 4. Balancing the number of integration servers. . . . .</b>	<b>6</b>	<b>Sending your comments to IBM . . . . .</b>	<b>12</b>
Strategy. . . . .	6		

---

## Chapter 1. Introduction

The major influence on the amount of memory that is used in processing data is the combination of the messages that are being processed, and how the flow is designed. However, the configuration of the integration broker runtime also has a significant effect on the memory usage, and it cannot be ignored.

The message flow logic and product code of an integration server all use virtual memory. The message tree, variables, and output messages are all held at locations within virtual memory as they are processed, therefore the entire process size of the broker does not need to be held in physical memory at the same time. Integration servers that are configured differently and have different integration flows deployed to them, use different amounts of both virtual and real memory; there is not an assigned default amount of memory for each integration server.

When processing begins, a subset of the virtual memory that is allocated to an integration server must be in real memory when that data is processed (read or updated) with ESQL or Java™ for example. This set of essential memory is what forms the real memory requirement, the working set, of the integration server. The contents of this working set change over the lifetime of the integration server.

The whole working set must be in real memory for message processing to continue smoothly. If required pages are moved into real memory during processing, then this movement causes delays in processing. In a system under heavy load, the operating system must manage the movements of pages quickly. This behavior can slow down processing for one or more components because the demand for real memory is higher than the memory that is physically available. In extreme circumstances, the system might thrash and complete no actual useful work.

The working set requirements of all the integration servers that are running, plus the memory that is required by the `bipservice`, `bipbroker`, and `bipHTTPListener` processes form the total real memory requirements of the integration node. As integration flows and integration servers are started and stopped, the real memory requirement changes. The overall real memory requirement for a fixed configuration and workload usually remains within a narrow band of usage, with only slight fluctuations. If the real memory requirement continues to grow, then there is probably a memory leak. This memory leak might be in the integration logic, or in some rare cases, in the IBM® Integration Bus product.

The memory requirement for an integration node is entirely dependent on the processing that must be carried out. If all integration servers were stopped, the real memory requirements would be smaller. The integration servers would still have their initial memory footprint, as well as the memory needed for tasks that do not involve processing messages, such as processing a deploy. So, if 50 integration servers were active and processing with inefficient integration flows to process large messages, then real memory requirements might rise to 10's of gigabytes in size. As with integration servers, there is no fixed memory requirement for an integration node. The memory requirement depends on which integration flows are running, and the configurations of the integration node.

All these factors affect memory usage, which is why the configuration recommendations in the following topics are so important to the smooth processing of your messages.

---

## Chapter 2. Deploying message flows to integration servers

Every message flow has a memory requirement that depends on the routing or transformation logic that is run and the messages to be processed.

If you use an extra instance of a message flow, the memory usage obviously rises. But the memory usage for the instance is lower than if you deploy another identical flow to a different integration server. If you deploy integration flows that are different from each other to an integration server, then more virtual memory and real memory is used by that integration server.

There is an initial requirement of memory for deployment, and a subsequent higher requirement after messages are processed. Different integration flows use different amounts of extra memory. It is not possible to accurately predict the amount of virtual memory or real memory that is needed for an integration flow or message to be processed. Therefore, when you plan your flow, it is recommended to run and measure the requirement after the integration flow processes some messages. If you only process 1 or 2 messages a second for a few minutes, the memory usage is unlikely to reach a steady state. Now is the time make the flow work hard: Make all the additional instances start by processing 1000 messages per second for example. After 2 or 3 minutes of processing messages, the memory usage stabilizes if the mix of the messages remains more or less constant (for example, if the messages are not continually increasing in size).




If multiple messages are deployed to the integration server, then ensure that all messages are processed before you observe the peak memory usage of the integration server.

When you look at the memory usage of an integration server, you must focus on the real memory usage:

- ▶ **Linux** **UNIX** On Linux and UNIX systems, you are looking for the RSS value.
- ▶ **Windows** On Windows systems, you are looking for the Working Set.

This statistic is the key to understanding how much memory a process is using at a specific point in time. Other pages might be sitting in swap-space, which were used at some point in the processing, (for example: At the start of processing, or when a different part of the message flow was run) and those pages might not be in real memory anymore as the demand for memory increases.

To understand how much memory is used by processes on the system, use the following commands:

-  For AIX® systems: Use this command to display both virtual memory and real memory usage, and the command for the process:  
`ps -e -o vsz=,rss=,comm=`
-  For Linux systems: Use this command to display both virtual memory and real memory usage, and the command for the process:  
`ps -e -o vsz,rss,cmd`
-  For Windows systems: Start the task manager and select **View > Select Columns > Memory (Private Working Set)**. You can then display both virtual memory and real memory usage, and the command for the process:  
`ps -e -o vsz,rss,cmd`

---

## Chapter 3. Configuring JVM settings for the integration server

The default settings of the Java Virtual Machine (JVM) are sufficient for most situations. However, as your requirements on message flow logic or nodes grows, the settings must be adjusted to ensure continued levels of service.

For IBM Integration Bus, the default memory settings of the JVM are as follows:

- Minimum: 32 MB
- Maximum: 256 MB

The amount of Java heap that is required depends on the message flows, and in particular the amount of Java that they contain. Remember, some nodes such as the FileInput node, or the SAPInput node are written in Java. Therefore, different integration servers might require different Java heap settings, so the same values cannot be reused for every integration server.

A larger Java heap might be needed when there is a heavy requirement from the message flow logic, or from the nodes that are used within it. The best way to determine whether there is sufficient Java heap is to look at the resource statistics for the integration server to check the level of garbage collection.

- For batch processing, low garbage collection resource statistics is the target. A garbage collection resource statistic of 1% is ideal.
- For real-time processing, low pause times are the target. A low pause time of under 1 second is ideal.

For guidance on resource statistics, see Starting resource statistics collection in the online product documentation.

The IBM Integration Bus JVM heap settings can be changed with the **mqsichangeproperties** command. The following example increases the maximum JVM heap to 512 MB for the integration server *IN\_OUT*, in the integration node *PERFBRKR*:

```
mqsichangeproperties PERFBRKR -o ComIbmJVManager -e IN_OUT -n jvmMaxHeapSize -v 536870912
```



---

## Chapter 4. Balancing the number of integration servers

An IBM Integration Bus integration node supports any number of integration servers, there is no finite limit. Practical limitations such as the amount of real memory or swap-space, do force a maximum limit to ensure continued levels of service.

Large systems typically have under 100 integration servers. A broker with over 100 integration servers is considered 'very large', and it is recommended to review the strategy that is used to create integration servers and assign message flows to them. The number of integration servers that are used does not affect the amount of either virtual memory or real memory that is used by an individual integration server. However, the number of integration servers does have a strong effect on the amount of real memory that is required by the integration node. It is therefore necessary balance business requirements against performance requirements when you calculate how many integration servers are required.

Just because an integration server can host hundreds of message flows, it is not always a good idea for it to do so. For example, if the integration server fails, many applications are lost until the integration server is restarted. And that restart time increases significantly the more message flows that the integration server hosts.

---

### Strategy

It is good operating practice to have a strategy in place to control the assignment of message flows to integration servers. A number of schemes are commonly used, such as:

- Limit message flows for one business application to one integration server: Large business applications might require multiple integration servers depending on the number of message flows for the business area.
- Assign flows together when they have the same hours of availability: There is little point in assigning flows for the online day and the batch activities at night into the same integration server; it is not possible to shut down the integration server without interrupting the service of one set of flows.
- Assign 'high profile' flows to their own integration server: When you use a high profile or high performance application for example, assign their message flows to their own integration server. Specific tuning and configuration can then be applied to just those high profile components, and the loss or maintenance of lower priority flows does not affect these ones.

- Use general utility integration servers to host many less important flows.
- In a deployment with many integration nodes, avoid deploying all message flows to all integration nodes. Instead, use some integration nodes for processing specific applications only. By deploying everything everywhere, the memory usage rises considerably, so this pattern reduces that usage. When you have message flows that you know use large amounts of memory, try to assign them to the lowest number of integration servers as possible. This assignment reduces the number of integration servers that become large and use more memory, and therefore reduces the amount of real memory that is required.

Whichever strategy you use, plan for future expansion and ensure that the topography still works when hundreds more services are added. Decisions that are made in 'the early days' often do not scale when hundreds more services are added, and the demand for more real memory becomes a problem.

---

## Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing  
IBM Corporation  
North Castle Drive  
Armonk, NY 10504-1785  
U.S.A.

For license inquiries regarding double-byte (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

IBM World Trade Asia Corporation  
Licensing 2-31 Roppongi 3-chome, Minato-ku  
Tokyo 106-0032, Japan

**The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law:**  
INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Corporation  
Software Interoperability Coordinator, Department 49XA  
3605 Highway 52 N  
Rochester, MN 55901  
U.S.A.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this information and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement, or any equivalent agreement between us.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM

products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

All statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

#### COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

---

## Programming interface information

Programming interface information, if provided, is intended to help you create application software for use with this program.

However, this information may also contain diagnosis, modification, and tuning information. Diagnosis, modification and tuning information is provided to help you debug your application software.

**Important:** Do not use this diagnosis, modification, and tuning information as a programming interface because it is subject to change.

---

## Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by

IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at Copyright and trademark information ([www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml)).



---

## Sending your comments to IBM

If you especially like or dislike anything about this book, please use one of the methods listed below to send your comments to IBM.

Feel free to comment on what you regard as specific errors or omissions, and on the accuracy, organization, subject matter, or completeness of this book.

Please limit your comments to the information in this book and the way in which the information is presented.

**To make comments about the functions of IBM products or systems, talk to your IBM representative or to your IBM authorized remarketer.**

When you send comments to IBM, you grant IBM a nonexclusive right to use or distribute your comments in any way it believes appropriate, without incurring any obligation to you.

You can send your comments to IBM in any of the following ways:

- By mail, to this address:

User Technologies Department (MP095)  
IBM United Kingdom Laboratories  
Hursley Park  
WINCHESTER,  
Hampshire  
SO21 2JN  
United Kingdom

- Electronically, use the appropriate network ID:
  - IBM Mail Exchange: GBIBM2Q9 at IBMMAIL
  - IBMLink: HURSLEY(IDRCF)
  - Email: idrcf@hursley.ibm.com

Whichever method you use, ensure that you include:

- The publication title and order number
- The topic to which your comment applies
- Your name and address/telephone number/fax number/network ID.



Printed in USA