



Deployment Guide

Sizing and Scalability Guidelines

6.2



COPYRIGHT AND TRADEMARK NOTICE

Copyright © 1997-2021 GraphOn Corporation. All Rights Reserved.

This document, as well as the software described in it, is a proprietary product of GraphOn, protected by the copyright laws of the United States and international copyright treaties. Any reproduction of this publication in whole or in part is strictly prohibited without the written consent of GraphOn. Except as otherwise expressly provided, GraphOn grants no express or implied right under any GraphOn patents, copyrights, trademarks or other intellectual property rights. Information in this document is subject to change without notice.

GraphOn, the GraphOn logo, GO-Global, and AppController are trademarks or registered trademarks of GraphOn Corporation in the US and other countries. Microsoft, Windows, Windows NT, Internet Explorer, and Remote Desktop Services are trademarks of Microsoft Corporation in the United States and/or other countries. Linux is a registered trademark of Linus Torvalds. Red Hat is a trademark or registered trademark of Red Hat, Inc. in the United States and other countries. Adobe, Acrobat, and Reader are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States and/or other countries. iPhone, iPad, iPod, Mac, and OS X are registered trademarks of Apple Inc.

Portions copyright © 1998-2017 The OpenSSL Project. All rights reserved. This product includes software developed by the OpenSSL Project for use in the OpenSSL Toolkit (www.openssl.org). Portions copyright © 1995-1998 Eric Young (eay@cryptsoft.com). All rights reserved. This product includes software written by Eric Young (eay@cryptsoft.com). This product includes software written by Tim Hudson (tjh@cryptsoft.com).

Portions of this software are licensed from United Mindworks LLC.

All other brand and product names are trademarks of their respective companies or organizations.



CONTACT INFORMATION

GraphOn Corporation
6 Loudon Road, Suite 200
Concord, NH 03301 USA
sales@graphon.com
www.graphon.com



CONTENTS

Chapter I — Introduction

GO-Global	p. 1
-----------	------

Chapter II — Application Design Considerations

Considerations	p. 2
Significant Performance Factors	p. 4

Chapter III — Host Requirements

Host Requirements	p. 5
Disk and Memory Requirements	p. 5
Networks and Modems	p. 6
Processor, Bus, and Memory	p. 6
Hard Disks	p. 6
Network Interfaces	p. 6
Using Performance Monitoring Tools	p. 6
Admin Console	p. 8

Chapter IV — Capacity Planning

Overview	p. 8
Memory Calculation Example	p. 9
Basic Allocation Example	p. 9
Sizing a GO-Global Host	p. 10
Troubleshooting	p. 17
Relay Load Balancer Considerations	p. 18

Chapter V — Cloud Deployment

Configuring a Cloud Deployment	p. 19
Monitoring Usage	p. 21
Taking a Host Offline	p. 22

Appendix

GO-Global Host Deployment Checklist	p. 25
-------------------------------------	-------

GO-Global

GO-Global allows Windows Server operating systems to serve Windows-based applications to desktops, terminals, and Web-browsers running on PC and non-PC devices. Because GO-Global clients are available on several platforms, users can gain access to Windows applications from a variety of client desktops. With its thin-client, server based architecture, GO-Global is optimized for reliable, secure, and scalable application delivery.

Unlike the traditional client/server environment, an application runs only on the server in the GO-Global Host environment. The GO-Global client performs virtually no local processing of applications. The GO-Global Host transmits only the application presentation—the Graphical User Interface (GUI) —down to the client. Each user logs on and perceives only his or her own *session*, which is managed by the server operating system and is independent from any other client session.

From an application development perspective, one of the biggest benefits of GO-Global is that the majority of existing application programs run "as is"—no programming changes are required to run them on a GO-Global Host. This does not mean that all existing applications run equally well under GO-Global. As with the addition of any new operating system service, understanding how to design applications that take advantage of these new capabilities is critical for a successful deployment.

Considerations

An important difference between the standard Windows server environment and the GO-Global Host environment is the design of applications that run in these environments. In a centralized host environment, applications are designed specifically to run in that environment. Utilizing GO-Global, applications designed for a Windows-based PC work without having to be explicitly developed for the GO-Global Host environment. Applications that run on Windows servers today will run on the GO-Global Host without modification.

This is important when you consider the implications of multiple users sharing a common Windows server system concurrently. Instead of having different users running applications using their own local hardware resources (such as CPU, memory, and disk), and local software resources (such as registry settings, preference files, and dynamic link libraries), users in the GO-Global Host environment share the same hardware and software resources. For example, if two users run the same application in a GO-Global Host environment, two copies of that application are started in the same system, each operating under a different user context. All of this is managed behind-the-scenes by the operating system.

Having multiple users access the same set of applications in a common system can create a number of areas of contention, however. These areas include:

- **Users contend for CPU time.** In the GO-Global Host environment, each user has his or her own application environment and can run whatever applications are available to that user. However, all applications run by all users are contending for the central CPU resources available on the server machine. If one user runs a poorly written, CPU-intensive application, other users on that server could experience considerable performance loss.

- **Users contend for disk access.** This is analogous to how users contend for disk access using the traditional client/server network connections. In the GO-Global Host environment, input/output demands are more intense because users contend not only for access to applications and related application files, but for server operating system disk access. For example, multiple users may be making different dynamic link library (DLL) calls at the same time or swapping between real and virtual memory areas.
- **Users contend for Random Access Memory (RAM).** Each user has an independent session, which he or she can attempt to fill with all the memory-intensive applications at his or her disposal. Some users may try to open as many applications as they can, while others take a more conservative approach and run only those applications that they need. Nonetheless, the needs of all the users are satisfied from the same core set of server memory resources.
- **Users contend for network access.** As in any distributed processing environment, the network provides the pipeline for communication between the desktop and the servers. In the GO-Global Host environment, the need for network access is more critical than in the traditional distributed client/server architecture because all desktop activity—graphical output and mouse/keyboard input—flows over the network links between the desktop and the server. Without a functioning network connection to the server, a GO-Global client cannot function at all.
- **Users contend for server-side hardware components.** Hardware components such as CD-ROMs, disk drives, serial ports, and parallel ports are typically server-based, not client-based. Therefore, when a user goes to access server-side hardware components, he or she may be competing with other users making similar requests. Sharing these traditionally non-shared components creates new considerations for both the users and for the applications that access these components.
- **Users contend for access to global Windows objects and resources.** In the GO-Global Host environment, users do not run individual copies of the Windows server. Some of the core components are cloned, but the remaining components are shared among GO-Global users. As such, users compete for access to the registry, the paging file, system services, and other global objects and resources.

Many of these points of contention can be mitigated by sizing the GO-Global Host system with sufficient CPU, memory, and disk resources to handle the client demand. For example, a multiple processor configuration can maximize CPU availability. Memory availability can be maximized by installing extra physical memory. Finally, disk access performance can be maximized by configuring multiple SCSI channels and distributing your operating system and application loads across different physical drives. Properly configuring a GO-Global Host is a critical element to yielding optimal application performance.

Significant Performance Factors

Know Your Application

Knowing your application may require asking the following questions: How graphics-intensive is the application? How much memory does it require? How much of the application's memory can be shared among users? How does the application refresh the screen? Does it require a lot of user input? What is considered acceptable performance?

Know Your Users

Knowing your user may require asking the following questions: How fast do your users type? Do your users leave the applications open? When do users log on? Do the users stay logged on throughout the day? Are logons evenly distributed throughout the day or are most logons at a specific time of the day?

Know Your Network

Because GO-Global provides application sessions to the user, network communications are very important. Even infrequent network slowdowns can provide unacceptable performance to your users. Knowing your network may include the following questions: How fast do your users type? Are the applications graphics-intensive? What is the typical display resolution of your client workstations? What is the network bandwidth required for a user running the application?

GO-Global Host Requirements

The GO-Global Host requires one of the following 64-bit Windows operating systems:

- Windows Server 2019 — Standard and Datacenter
- Windows Server 2016 — Standard and Datacenter
- Windows Server 2012 R2 — Standard and Datacenter
- Windows 10 (latest SAC release) — Professional and Enterprise
- Windows 10 (latest LTSC release) — Enterprise

The GO-Global Host is supported on computers that have the latest Windows Updates installed.

Disk and Memory Requirements

In addition to the Windows operating system requirements for your server, the following are required for GO-Global:

- 70 MB disk space for standard GO-Global Host installation, including the Admin Console
- 85 MB for Web files
- 50 MB disk space for GO-Global client files

Networks and Modems

In addition to connections over LAN or WAN based networks, GO-Global supports dial-up networks utilizing modem as well as VPN connections. TCP/IP connectivity is a minimum requirement. For best results, GraphOn recommends a dual processor system, 1 GB of RAM, and a hard disk with at least 256 MB of free space.

Processor, Bus, and Memory

The processor and bus architecture are fundamental to GO-Global Host performance. The ISA (AT bus) architecture is low-bandwidth and is not recommended for GO-Global Hosts. Use a higher-performance bus, such as EISA or PCI, for best performance. These buses support the sustained data transfer rates of a GO-Global Host.

The memory (RAM) requirement for GO-Global is 16 MB plus 4 MB for each typical user or 8 MB for each power user. In many cases, adding RAM has a greater effect on system performance than upgrading to a faster processor. In general, processor and RAM requirements for GO-Global scale linearly. You can roughly double the number of users supported on a multi-processor capable system by doubling the number of processors and doubling the amount of RAM. Purchasing multi-processor capable systems (even if you initially purchase only one processor), provides for convenient system scaling as your requirements grow.

Hard Disks

The hard disk subsystem in a server is an important factor in system throughput. Small Computer System Interface (SCSI) disk drives and adapters, especially Fast Narrow SCSI (SCSI-2), Fast Wide SCSI, Wide Ultra SCSI, and Wide Ultra2 SCSI devices, have significantly better throughput than ST-506, Integrated Device Electronics (IDE), or Enhanced Small Device Interface (ESDI) disk drives and adapters.

For the highest disk performance, consider using a SCSI-based Redundant Array of Independent Disks (RAID) controller. RAID controllers automatically place data on multiple disk drives and can increase disk performance and improve data reliability. Use NTFS for all disk partitions on your GO-Global Hosts. NTFS allows security configuration, better performance, and more fault tolerance.

Network Interfaces

The RXP protocol is highly compressed and causes negligible loading on a network, but because the GO-Global Host handles all network requests, a high performance network interface card (NIC) is recommended.

Using Performance Monitoring Tools

GraphOn recommends using performance monitoring tools to get accurate accounts of system performance and the effects of configuration changes on system throughput. The most important measurements for performance monitoring are the percentage of total processor time, memory pages per second, percentage of network utilization, and hard disk I/O rates. An effective way to estimate how many users a server can support is to measure system performance with two to five users on the system and then scale the results. This method has been found to yield reliable results.

Admin Console

The Admin Console is a centralized management and monitoring utility. It is used by administrators to monitor or manage user sessions, server processes, and to register and manage applications on the GO-Global Host. The Admin Console allows for remote management of all GO-Global Hosts from a central location.

Overview

GO-Global supports multiple users on Windows servers. A multi-user system requires more system resources than a single-user system. This section provides some system sizing guidelines that can help you decide on a hardware configuration that will support your users with optimal performance. Most companies find that their users can be categorized as either typical users or power users.

- **Typical user**

A typical user generally uses one or two applications — normally only one at a time. Little actual program data is transferred between the client and server, and the users rarely use Object Linking and Embedding (OLE).

- **Power user**

A power user is a more sophisticated user who uses three or more applications, often with several active at the same time. Data is often cut and pasted between local and remote applications, and OLE is used heavily. Power users consume more resources than typical users. A good rule of thumb is that one power user is equivalent to two typical users in processor utilization and RAM requirements.

Memory Calculation Example

A basic GO-Global Host allocates memory as follows:

Service	Memory allocation	Program
Application Publishing Service (one instance per server)	15 MB	aps.exe
License Service (two instances per server)	2 MB	lmgrd.exe
License Interface (one instance per user)	2 MB	blm.exe
Update Client Service (one instance per server)	2 MB	uc.exe
Program Window (one instance per user, unless using a desktop shortcut)	3 MB	pw.exe
Logon Graphical Authentication (one instance per user)	2 MB	logon.exe
Remote Clipboard (one instance per user)	2 MB	lremoteClipboard.exe

Note: It is important to realize that any subsequent application memory is in addition to the basic GO-Global Host configuration.

Basic Allocation Example

In the following example we calculate the total GO-Global Host consumption plus any subsequent user allocation based upon such things as logon services and licensing. A sample application such as Microsoft Word is factored in to the final equation. If we utilize Word 2003, the additional allocated RAM for an average document will be ~25 MB per user.

Program	MB		Total
aps.exe	15	→	
lmgrd.exe	2		
uc.exe	2		
Total server consumption	19		19 MB
lremoteClipboard.exe	2		
pw.exe	3		
blm.exe	2		
logon.exe	2		
winword.exe	25		
Total user consumption	34		
		* Number of users (15)	510 MB
		Total RAM requirement	529 MB

The example above shows that a total of 529MB of RAM would be required to support 15 users utilizing Microsoft Word on the server. These numbers are approximate and are subject to change depending on specific system and hardware configurations. Note that these numbers reflect only the GO-Global system services and do not take into account other applications or system services, such as a Web server. Standard system overhead is in addition to any GO-Global services.

Sizing a GO-Global Host

The following section describes a more advanced method for determining the maximum number of users that can be supported on a GO-Global Host.

The number of users that can be supported from a GO-Global Host is influenced by many factors, including:

- The speed and number of processors
- The amount of physical memory
- Memory requirements of the application(s)
- Processor utilization requirements of the application(s)
- Kernel memory requirements of hardware devices (for example, RAID hard disk controllers, video cards, etc.)

A combination of these factors will generally determine the number of users that can effectively run from a GO-Global Host. This section describes how to evaluate the scalability impact of these factors and estimate the maximum number of users that can be supported from a GO-Global Host.

The number of users that can be supported on a given GO-Global Host is generally limited by one or more of the following system resources:

- **Processor (CPU) Utilization:** The percentage of time that the server's processors are executing non-idle threads.
- **Physical Memory:** The amount of physical memory used by processes running on the computer.
- **Committed Memory:** The amount of memory for which space has been reserved in the disk paging file.
- **Paged Pool Memory:** The amount of memory allocated within the Paged Pool, an area of kernel-mode memory used to store data for the operating system and other kernel-mode modules.
- **System Page Table Entries (PTEs):** The amount of space consumed within the System Page Table Entry (PTE) area, an area of kernel-mode memory that is used by the operating system to store kernel stacks, and by device drivers (e.g., RAID hard disk controllers, video cards, etc.) to map views of device-specific data into the kernel address space.

To determine the number of users that can run effectively from a given server

1. Determine how much of each of the above resources is required for an average user.

2. Determine how much of each resource is initially available on the server computer.
3. Calculate the user limit using the information from steps 1 and 2.
4. Verify the calculated result in a production environment.

The details of this process are described in the pages that follow.

1. Determining User Resource Requirements

User resource requirements are generally determined by the applications that users will run on the server. If an application supplier has tested an application with GO-Global, they may be able to provide per-user resource requirements for the application. Otherwise, the best way to determine user requirements is to run a multi-user test that simulates the operations that typical users will perform on the server.

To determine the per-user resource requirements for a given set of applications

1. Restart the server computer.
2. Begin logging system resources using the **Windows Performance Monitor**.
 - a. On a separate Windows computer, start the **Windows Performance Monitor**.
 - b. Right-click **\Performance Logs and Alerts\Counter Logs**.
 - c. Choose **New Log Settings**.
 - d. Type a name for the log.
 - e. Click **Add**.
 - f. Enter the name of the server computer in the **Select counters from computer** field.
 - g. Add the following counters:
 - i. \Processor\% Processor Time
 - ii. \Memory\Available Bytes
 - iii. \Memory\Committed Bytes
 - iv. \Memory\Pool Paged Bytes
 - v. \Memory\Free System Page Table Entries
 - h. Set the **Sample Interval** to 10 seconds.
 - i. Select the **Log Files** tab.
 - j. Change the **Log file type** to Text File - CSV.
 - k. Click **OK**.
 - l. Right-click the icon for the new log and choose **Start**.
3. Log 10 users on to the GO-Global Host. Each user should be running on a separate client computer.
4. In each session, start all the programs that you expect users to run during a typical session. In each session, simulate typical user interaction with the applications. This may be done with live users driving each session, or with a script that is recorded and played back on each client using a third-party automated testing tool.
5. Run the test for 15-30 minutes.
6. Stop the performance monitor log.
7. Right-click the icon for the log and choose **Stop**.
8. Open the log file in a spreadsheet program such as Microsoft Excel and chart the results.
9. Determine the per-user resource requirements from the data as follows:

a. The average processor utilization per user:

$$\text{AverageProcessorUtilizationPerUser} = (\text{AveragePercentProcessorTime} * \text{NumberOfProcessors}) / 10$$

where AveragePercentProcessorTime is the average value of the \Processor\% Processor Time_Total counter during the period when typical user interaction was simulated, and NumberOfProcessors is the number of processors installed in the computer.

b. The physical memory required per user:

$$\text{PhysicalMemoryPerUser} = (\text{InitialAvailableBytes} - \text{MinimumAvailableBytes}) / 10$$

where InitialAvailableBytes and MinimumAvailableBytes are respectively the initial and minimum values of the \Memory\Available Bytes counter.

c. The committed memory required per user:

$$\text{CommittedMemoryPerUser} = (\text{MaximumCommittedBytes} - \text{InitialCommittedBytes}) / 10$$

where MaximumCommittedBytes and InitialCommittedBytes are respectively the maximum and initial values of the \Memory\Committed Bytes counter.

d. The Paged Pool memory required per user:

$$\text{PagedPoolMemoryPerUser} = (\text{MaximumPagedPoolBytes} - \text{InitialPagedPoolBytes}) / 10$$

where MaximumPagedPoolBytes and InitialPagedPoolBytes are respectively the maximum and initial values of the \Memory\Pool Paged Bytes counter.

e. The number of System PTEs required per user:

$$\text{SystemPTEsPerUser} = (\text{InitialFreeSystemPTEs} - \text{MinimumFreeSystemPTEs}) / 10$$

where InitialFreeSystemPTEs and MinimumFreeSystemPTEs are respectively the initial and minimum values of the \Memory\Free System Page Table Entries counter.

2. Determining Available Resources

Once the user requirements for the desired set of applications is known, the next step is to measure the available resources of the computer(s) that will be used to run the applications. This may be done as follows:

1. If it is not already installed, install GO-Global on the server computer.
2. Restart the computer.
3. Obtain the initial values of the following performance counters using the **Windows Performance Monitor**:
 - a. \Memory\Available Bytes
 - b. \Memory\Committed Bytes
 - c. \Memory\Commit Limit
 - d. \Memory\Pool Paged Bytes
 - e. \Memory\Free System Page Table Entries
4. Use this data to calculate the following available resources for the computer:

a. The available processing capacity:

$$\text{AvailableProcessorUtilization} = (\text{NumberOfProcessors} * \text{ProcessorSpeed} * \text{MaximumPercentUtilizationAllowed}) / \text{BaseProcessorSpeed}$$

where NumberOfProcessors is the number of processors installed in the server, ProcessorSpeed is the speed in MHz of the server's processors, and BaseProcessorSpeed is the speed of the processor used to determine the AverageProcessorUtilizationPerUser in step 1.8.a above.

MaximumPercentUtilization is the maximum percentage of processor utilization that should occur during normal operation. The recommended value for MaximumPercentUtilization is 0.80 (80%).

b. The available physical memory:

$$\text{AvailablePhysicalMemory} = \text{InitialAvailableBytes} - \text{MinimumAvailableBytesAllowed}$$

where InitialAvailableBytes is the initial value of the \Memory\Available Bytes counter, and MinimumAvailableBytesAllowed is the minimum number of bytes of physical memory that should be available during normal operation. The recommended value for MinimumAvailableBytesAllowed is 41943040 bytes (40 MB).

c. The available committed memory (i.e., available space in the disk paging file):

$$\text{AvailableCommittedMemory} = \text{CommitLimit} - \text{InitialCommittedBytes} - \text{MinimumAvailableCommittedBytesAllowed}$$

where CommitLimit and InitialCommittedBytes are respectively the values of the \Memory\Commit Limit and \Memory\Committed Bytes counters, and MinimumAvailableCommittedBytesAllowed is the minimum number of bytes of committed memory that should be available during normal operation. The recommended value for MinimumAvailableCommittedBytesAllowed is 0.20 times the Commit Limit.

d. The available Paged Pool memory:

$$\text{AvailablePagedPoolMemory} = \text{SizeOfPagedPool} - \text{InitialPagedPoolBytes} - \text{MinimumAvailablePagedPoolBytesAllowed}$$

where SizeOfPagedPool is the size of the paged pool in bytes (generally 0x0A000000 (167772160) bytes on Windows 2000 Servers) and InitialPagedPoolBytes is the initial value of the \Memory\Pool Paged Bytes counter. MinimumAvailablePagedPoolBytesAllowed is the minimum number of paged pool bytes that should be available during normal operation. The recommended value for MinimumAvailablePagedPoolBytesAllowed is 1048576 bytes (10 MB).

Note: The actual size of the paged pool can be determined using a kernel debugger as follows:

1. Break into the debugger.
2. Type !vm. Then press ENTER.
3. Information about the computer's virtual memory will be displayed. The size of the Paged Pool will be listed as *PagedPool Maximum*.

For information on how to use a kernel debugger, refer to article 151981 in the Microsoft Knowledge Base.

e. The available System PTEs:

AvailableSystemPTEs = InitialFreeSystemPTEs - MinimumAvailableSystemPTEsAllowed

where InitialFreeSystemPTEs is the value of the \Memory\Free System Page Table Entries counter, and MinimumAvailableSystemPTEsAllowed is the minimum number of SystemPTEs that should be available during normal operation. The recommended value for MinimumAvailableSystemPTEsAllowed is 4000.

3. Calculating the User Limit

Using the values obtained in steps 1 and 2, a user limit for the computer may be estimated as follows:

1. Calculate the maximum number of users that can be supported by each of the computer's resources:

- a. Calculate the maximum number of users that can be supported by the computer's processors:
ProcessorUserLimit = AvailableProcessorUtilization / AverageProcessorUtilizationPerUser

- b. Calculate the maximum number of users that can be supported by the computer's available physical memory:

PhysicalMemoryUserLimit = AvailablePhysicalMemory / PhysicalMemoryPerUser

- c. Calculate the maximum number of users that can be supported by the computer's available committed memory:

CommittedMemoryUserLimit = AvailableCommittedMemory / CommittedMemoryPerUser

- d. Calculate the maximum number of users that can be supported by the computer's available paged pool memory:

PagePoolMemoryUserLimit = AvailablePagedPoolMemory / PagedPoolMemoryPerUser

- e. Calculate the maximum number of users that can be supported by the computer's available System PTEs:

SystemPTEUserLimit = AvailableSystemPTEs / SystemPTEsPerUser

2. Calculate the maximum number of users for the computer using the following formula:

**MaximumNumberOfUsers = Minimum (ProcessorUserLimit,
PhysicalMemoryUserLimit,
CommittedMemoryUserLimit,
PagePoolMemoryUserLimit,
SystemPTEUserLimit)**

3. If the maximum number of users is lower than required, attempt to modify the configuration of the computer to increase the availability of the limiting resource.

- a. If **Processor Utilization** is the limiting resource, increase the number of servers, or use servers with faster and/or a larger number of processors.
 - b. If **Physical Memory** is the limiting resource, install additional memory in the computer. The maximum amount of memory that may be installed in a Windows 2000 Server is 4 GB.
 - c. If **Committed Memory** is the limiting resource, increase the size of the disk-paging file:

- i. Right-click **My Computer** and select **Properties**.
- ii. Click **Advanced | Performance Options | Change**.
- iii. Increase the maximum size of the paging file.
- iv. Click **Set**.

WARNING:

The recommended size of the disk paging file will generally be less than or equal to twice the amount of physical memory installed in the machine. Use care when increasing the size of the paging file above the recommended size. If the amount of committed memory is allowed to rise far above the amount of physical memory, the system may begin performing an excessive number of disk I/O operations to retrieve data from paging file. This can seriously degrade system performance.

The following information involves opening and manipulating the Windows Registry. Carrying out operations other than those described here may cause configuration errors, possibly rendering your system unusable. Please use extreme caution any time you work in the Registry.

d. If **Paged Pool Memory** is the limiting resource, increase the size of the Paged Pool:

- i. Choose **Start | Run**.
- ii. Type **regedit**, **ENTER**.
- iii. Select the following registry value:
\\HKEY_LOCAL_MACHINE\\SYSTEM\\CurrentControlSet\\Control\\Session Manager\\Memory Management\\SystemPages.
- iv. Choose **Edit | Modify**.
- v. Type **0** in the **Value data field**. This will increase the size of the paged pool by decreasing the number of System PTEs to the system's default value.
- vi. Click **OK**.
- vii. Restart the computer.

Note:

Specify the minimum amount of paged pool memory that must be available for users to log on to the host, as follows:

1. Run the Registry Editor (regedit) on the server.
2. Select the following registry value:
\\HKEY_LOCAL_MACHINE\\Software\\GraphOn\\Bridges\\1.0.0\\AppServer\\SessionPagedPoolMemory.
3. Click **Edit | Modify**.
4. Select **Hexidecimal**.
5. In the **Value data** field, type the number of bytes that must be available in order to start a session (e.g., 10000).
6. Click **OK**.

e. If the number of **System PTEs** is the limiting resource, ensure that the system is configured to maximize the number of System PTEs:

- i. Double-click the following registry value:
\\HKEY_LOCAL_MACHINE\\SYSTEM\\CurrentControlSet\\Control\\Session Manager\\Memory Management\\SystemPages.
- ii. Select **Hexidecimal**.

- iii. Check the value. It should be FFFFFFFF. The GO-Global Setup Program sets this field to this value when the product is installed. If the value has been subsequently modified, reset it to FFFFFFFF.
- iv. Check the value of the **PagedPoolSize** value (also located under the Memory Management key). It should be zero (0).
- v. If the system is configured to maximize the number of System PTEs, review the requirements of hardware devices installed in the computer. Some devices, such as RAID disk controllers and video cards can consume large numbers of System PTEs. If such a device is installed in the server, ensure that it is required, and, if it is not, uninstall the device.
- vi. If any configuration modifications are made in step 3.3 above, recalculate the affected available resource, the resource user limit, and the maximum user limit for the server computer.

4. Verifying the Result

The result of the procedure outlined in steps 1-3 is an estimate of the maximum number of users that can be supported on a given GO-Global Host. Several assumptions are made to arrive at this result. For example, it is assumed that the operations performed during the multi-user resource requirements test are representative of the operations that real users will perform when using the product in a production environment. It is also assumed that the number of users will be limited by one of the system resources measured in the test.

To ensure the accuracy of these assumptions and the final result, the server configuration should be tested in a production environment, generally as part of a pilot deployment to the number of users calculated above. During this deployment, system resources should be monitored using the **Windows Performance Monitor**, and users should be asked to report the nature and time of any problems that they observe. In addition to the performance counters for the resources mentioned above, the following performance counters should also be monitored:

- **\System\Processor Queue Length:** The Processor Queue Length is the number of active threads that are available but unable to execute due to another active thread. One of the clearest indicators of a processor bottleneck is a sustained processor queue length greater than 1-3 threads per processor. High processor queue lengths generally occur when processor utilization is high, but they can develop when processor utilization is well below 90%.
- **\Memory\Pages/sec:** The Pages/sec performance counter measures the number of pages read from or written to disk to resolve hard page faults. Hard page faults occur when memory must be read from or written to the disk paging file. System-wide performance delays can occur when the number of hard page faults is high, even when CPU usage is low.

Troubleshooting

If problems are reported:

1. Attempt to identify the resource that caused the problem by reviewing the data from the Windows Performance log at the time the problem occurred:
 - a. If the user reported sluggish performance, check:

- i. **% Processor Utilization:** If greater than 90%, the computer's processors may be the bottleneck.
 - ii. **Processor Queue Length:** If greater than 1-3 threads per processor for a sustained length of time, the computer's processors may be the bottleneck.
 - iii. **Available Bytes:** If this value is low (e.g., less than 10% of the installed physical memory), check the Pages/sec counter. If the Pages/sec value is significantly higher than normal, the system is frequently swapping memory to and from disk. This problem is generally caused by having insufficient physical memory.
- b. If the user reported an inability to connect to the server or reported errors while running an application, check:
- i. **Available Bytes:** GO-Global prevents new users from connecting to the server if Available Physical Memory is less than 20 MB.
 - ii. **Committed Bytes:** If the value of this counter is near the Commit Limit, random errors may occur in all sessions. Also, high values for Committed Memory often coincide with high disk paging (Pages/second) values and sluggish performance.
 - iii. **Paged Pool Bytes:** If near the maximum value (0x0A000000 (167772160) bytes on Windows 2000 Servers), the system may be running out of paged pool memory. If this is the case, random errors will occur in all running sessions.
 - iv. **Free System Page Table Entries:** GO-Global prevents new users from connecting to the server if the number of free System PTEs is less than 2000. If the number of free System PTEs drops below 1000, random errors can occur in all running sessions.

2. After identifying the resource that caused the problem:

- a. Attempt to increase the availability of the resource as described in step 3.3
—or—
- b. Reduce the maximum number of users for the specified resource to a safe value.

Relay Load Balancer Considerations

The only way to estimate per-user requirements of a given resource on a system hosting a given set of applications is to measure resource usage when the system is supporting x users, subtract the base value of the resource (the amount of the resource that is consumed when there are no users connected to the system), and then divide this number by the number of users (x).

For a relay load balancer, the key resources are CPU, memory and available bandwidth. Any one of these resources could be a bottleneck. To determine the per-user requirements of these resources for a given relay server that is providing access to a given set of applications, customers should monitor resource usage and use the method described above to estimate the per-user requirements of these resources.

There is no known hard limit on the number of concurrent users that a relay load balancer can support. In theory, a relay load balancer can support any number of users as long as there is sufficient CPU, memory and network bandwidth capacity.

Configuring a Cloud Deployment

In cloud deployments, host instances are often started and stopped as demand, respectively, rises and falls. To configure GO-Global for this type of environment:

I. Create the Application Host Manager

1. Install GO-Global on a computer that will be accessible to all the hosts.
2. Activate GO-Global:
 - a. If you are using a cloud license, run the Activation Wizard and select the cloud license you want the Application Host Manager to use.
 - b. If you are using an on-premises license, create three redundant license servers and configure the Application Host Manager to use them:
 - i. Install the GO-Global Licensing component on three AWS instances and configure the instances as described in the ***Using On-Premises Licenses in Cloud Environments*** section of the *Administrator Guide*.
 - ii. Install a three-server redundant license file on each of the license servers.
 - iii. Point the Application Host Manager to use the three license servers as described in the ***Multiple Host Environments*** section for on-premises licenses.
3. Run the Admin Console and click Tools | Host Options | Configuration.
4. Under Server Role, select **Application Host Manager**, then select either **Farm Manager** or **Relay Load Balancer**. Select **Farm Manager** if you will be using an AWS load balancer. Select **Relay Load Balancer** if you will be using GO-Global's built-in load balancing functionality.
5. Click **OK**.
6. Restart the **Application Publishing Service**.

II. For High Availability, Create a Backup Application Host Manager

1. Install GO-Global on another computer that will be accessible from all hosts and from the Application Host Manager. For availability across cloud availability zones, install GO-Global on a computer that is in a different cloud availability zone than the Application Host Manager.
2. Activate GO-Global as in section I, step 2 above.
3. Run the Admin Console and click Tools | Host Options | Configuration.
4. Under Server Role, select **Application Host Manager**, then select either **Farm Manager** or **Relay Load Balancer**, as in section I, step 4 above.
If you are using a cloud license, type the address of the *primary* Application Host Manager computer configured in section I above in the **Backup License Manager(s)** field.
5. Click **OK**.
6. Restart the **Application Publishing Service**.
7. If you are using a cloud license, run the Admin Console on the Application Host Manager computer configured in step 1 and specify its Backup License Manager:
 - a. Click Tools | Host Options | Configuration.
 - b. In the **Backup License Manager(s)** field, type the address of the Backup Application Host Manager computer configured in steps 1-6.
 - c. Click **OK**.

The screenshot shows the 'Host Options' dialog box with the 'Configuration' tab selected. The 'Server Role' section has the following options:

- ☐ Application Host
- ☐ Independent Host
- ☐ Dependent Host
 - Relay Load Balancer address:
- ☐ Farm Host
 - Farm Manager address:
- ☒ Application Host Manager
 - ☐ Relay Load Balancer
 - ☒ Farm Manager

At the bottom, there is a field for 'Backup License Manager(s):' with an empty text box.

Buttons at the bottom right: OK, Cancel.

III. Create the Host Image

1. Install GO-Global on another computer that can communicate with the Application Host Manager and the Backup Application Host Manager.
2. Run the Admin Console and click Tools | Host Options | Configuration.
3. Under Server Role, select **Application Host**, then select either **Farm Host** or **Dependent Host**. Select **Farm Host** if the **Application Host Manager** is a **Farm Manager** and select **Dependent Host** if the **Application Host Manager** is a **Relay Load Balancer**.
4. In the corresponding **Farm Manager address** or **Relay Load Balancer address** field, type the address of the Application Host Manager, followed by a semicolon, followed by the address of the Backup Application Host Manager.
For example: application_host_manager;backup_application host_manager.
5. Click **OK**.
6. Restart the **Application Publishing Service**.
7. Publish applications and make other configuration changes that are desired.
8. If using an AWS load balancer, configure the instance to be part of the load balancer's target group.
9. Shut down the instance.
10. Create an image of the instance.

From here, host instances can be created as needed from the host image. When a host instance is started from the image, the host will automatically connect to the Application Host Manager.

When sessions start on the host, the host will check out seats from the Application Host Manager. Generally, one seat will be consumed for each concurrent session.

When a host stops, the host's connection to the Application Host Manager is broken. When a host's connection to the Application Host Manager is broken, the Application Host Manager automatically releases any seats that were in use on the host. As such, there is no risk that seats will be consumed by hosts that no longer exist. Nothing needs to be done to clear hosts that have been shut down; this is done automatically.

Monitoring Usage

Administrators can monitor load by monitoring the value of the GO-Global Host | Active Sessions Performance Counter. On the Application Host Manager, the value of this Performance Counter is equal to the total number of sessions running on all hosts. On each host, the value of this Performance Counter is equal to the number of sessions running on the host.

Taking Hosts Offline

When a host needs maintenance or the load drops to a level where a host is no longer needed, administrators can either manually or automatically take the host offline. This can generally be done without losing any user sessions by first preventing new sessions from being started on a host and then waiting for the sessions running on a host to close.

When using a Relay Load Balancer, administrators can use the following procedure to shut down a host without losing any user sessions.

1. Prevent new sessions from starting sessions on the host by setting **Maximum sessions on this host** to 0 using one of the following methods:
 - a. Set the value of the **maxUsersAllowed** property in the host's HostProperties.xml file to 0.
 - b. Set **Maximum sessions on this host** to zero via the Admin Console:
 - i. Run the Admin Console on the Application Host Manager and select the desired host from the list of All Hosts or run the Admin Console on the host itself.
 - ii. Click Tools | Host Options.
 - iii. Click the **Session Startup** tab.
 - iv. In the **Maximum sessions on this host** box, set the value to 0.
 - v. Click **OK**.
2. Monitor the number of sessions running on the host. When the number of sessions reaches zero, shut down the Dependent Host.

Taking a Farm Host Offline

When using third-party load balancers with a Farm Manager, administrators can use the following procedure to shut down a host without losing any user sessions.

1. If the host is configured to terminate sessions immediately when clients disconnect, enable reconnects by changing the value of the **SessionTimeoutBrokenConnection** property in the HostProperties.xml file to 1. With this setting, sessions will continue running for 1 minute after the session's client disconnects.
2. Remove the target host, i.e., the Farm Host you want to take offline, from the load balancer's target group. This will prevent new connections from being routed to the target host, and at some point, depending on the load balancer, it should close the open connections to the target host.

When connections to the target host are closed, GO-Global clients will automatically reconnect to their sessions running on the target host via other Farm Hosts in the cluster. Specifically, they will open a new connection to the load balancer, the load balancer will route the connection to one of the active Farm Hosts, and the active Farm Host will relay the connection to the host that has been removed from the load balancer's target group.

3. Monitor the number of sessions running on the target host. When the number of sessions reaches zero, shut down the Farm Host.

Note: On AWS Network Load Balancer, set **deregistration_delay.connection_termination.enabled** to true so connections will be closed when the target host transitions to the unused state.

APPENDIX

GO-Global Host DEPLOYMENT CHECKLIST

- ☐ Configure hardware
- ☐ Configure operating system
- ☐ Add international keyboard support (as needed)
- ☐ Attach server to the network
- ☐ Add DNS records for the host
- ☐ Optionally install web server
- ☐ Install application(s)
- ☐ Configure application(s) (as needed)
- ☐ Create/modify Windows user and server policies
- ☐ Ensure that users accounts have “local logon rights” in Local Security Policy, Domain Security Policy, and Domain Controller Security Policy
- ☐ Install GO-Global Host software
- ☐ Install GO-Global license
- ☐ Reboot the host
- ☐ Sign in with a user account that has administrative privileges
- ☐ Publish applications using the Admin Console
- ☐ Add server options to the Admin Console (as needed)
 - ☐ Data Encryption
 - ☐ Client Drives
 - ☐ Client Clipboard
 - ☐ Client Printers
 - ☐ Authentication
 - ☐ Connection transport mode (TLS or TCP)
- ☐ Secure the host and applications
- ☐ Apply corporate branding using the Branding dialog (optional)
- ☐ Deploy client software

[REVISION DATE: 2021.07.30]