MCSE Enterprise

Capacity Planning

About Chapter 24

C hapter 24 explores the fundamentals of capacity planning, a process that includes determining current usage of a server's or network's resources, as well as tracking their utilization over time to predict future usage and any additional equipment that will be required to meet projected needs.

This chapter begins by defining the basic steps involved in capacity planning: establishing a baseline, gathering data over time, and using data to predict future utilization. Then it explains how to perform each of these tasks.

Next, this chapter explains how to capacity plan your server. This section suggests specific tools and techniques you can use to determine if your server's performance is adequate for now, and, if it is, to determine when the server will need upgrading.

Next, the chapter explains how to capacity plan your network, either for overall network utilization, or for a specific type of network traffic. Specific Performance Monitor counters and Network Monitor statistics are suggested.

Chapter 24 contains some interesting and worthwhile information, but can be considered optional reading unless you're preparing for the Enterprise exam. This chapter maps to the "Establish a baseline for measuring system performance" objective in the Monitoring and Optimization section for the Enterprise exam.

Fundamentals of Capacity Planning

Capacity planning is the process of determining current usage of server and/or network resources, and tracking utilization over time to predict future usage and the additional hardware that will be required to meet projected levels of utilization. Capacity planning can be performed on a single computer, such as a network server, or it can be performed on an entire network. Capacity planning enables you to plan for your organization's future needs and growth.

Two methods are commonly used to perform capacity planning. One method involves determining current server/network utilization to predict future utilization. The second method involves simulating the expected future utilization using various tools. These tools, which are sometimes called *stress test tools*, can simulate specified levels of server utilization and/or network traffic to determine how current hardware responds to expected future levels of use.

Because the second method requires the use of special tools that *don't* ship with Windows NT, and, therefore, aren't part of the Windows NT 4.0 Certified Professional exam objectives, this chapter focuses on the first method of capacity planning.

The two primary Windows NT administrative tools that can be used for capacity planning are Performance Monitor and Network Monitor. Additionally, Windows NT Diagnostics provides some statistics that can be useful for capacity planning.

The basic steps involved in capacity planning are:

- Establishing a baseline
- Gathering data over time
- Using this data to predict future utilization and future hardware requirements

The following sections explain each of these basic capacity planning steps. Additionally, this chapter recommends specific statistics to use when capacity planning a server and when capacity planning a network.

Establishing a Baseline

Establishing a baseline is the equivalent of determining the status quo — you gather data to determine current server and/or network utilization and performance. Determining the current situation is the first step in predicting what the future environment might be like.

You can use Performance Monitor and Network Monitor to gather data about your server's memory, processor, and disk. You can also obtain statistics about the network in general and about how a particular server utilizes the network specifically. The data you gather in this step will form an initial baseline of data about your server and/or network.

You should consider gathering data at various times to obtain an overall picture of utilization. For example, consider collecting data during business hours, when network usage is at its peak and production is heavy. Also consider collecting data during nonbusiness hours when network utilization levels should be lower. You might also consider gathering data over several days in a week, or perhaps over several days in a month to take into account normal cyclical business highs and lows. Taking several samples over a limited period of time normally provides the most accurate initial picture of your server and/or network environment.

Gathering Data Over Time

Once you've established a baseline of data, the next step in capacity planning is to continue gathering data over time to obtain additional statistics that can later be used for trend analysis.

A reasonable period of time for data gathering depends on many factors, including:

- The amount of time you can afford to allocate to capacity planning tasks
- How fast your organization is growing
- External requirements, such as equipment budget deadlines
- The changing skill levels of users in your organization (because more skilled users tend to create more data)
- The nature of your organization, particularly if your company experiences seasonal peaks and lows during the course of its operating cycle

Perhaps the most important consideration is that data should be gathered frequently enough so that you can perform reliable trend analysis to predict future utilization.

Using Data to Predict Future Utilization

Once you've established a baseline and gathered data over time, you're ready to analyze that data to predict future utilization, and the hardware that will be required to meet those utilization levels.

Often the raw data, in and of itself, isn't extremely useful for trend analysis. It is usually beneficial to export the gathered data in a format acceptable to standard analysis tools (such as spreadsheets and databases). For example, Performance Monitor enables you to export data directly to a comma-separated value or tab-separated value file. These types of files can be imported directly into most spreadsheets and databases.

Once you have imported the data into a spreadsheet or database, you can chart it, graph it, or otherwise manipulate it to obtain meaningful results.

In most instances, the data gathered over the last several months can be used to graph growth trends in server/network utilization. This growth trend graph can be extended to predict future server/network utilization. When the extended graph indicates that future utilization of a specific resource will approach or exceed maximum capacity, additional hardware may be required to meet anticipated needs.

Up to this point, gathering data for capacity planning has been discussed in general terms. Now let's look at the specific types of data and statistics that are most useful for capacity planning your server, and for capacity planning your network.

Capacity Planning Your Server

Capacity planning your server involves using Performance Monitor and Network Monitor to gather server and network statistics over time to establish a baseline and to provide additional data for trend analysis. You can analyze these statistics to determine if the server's performance is adequate now, and, if it is, to determine approximately when the server will need to be upgraded, and what hardware will be required to maintain its performance.

> **tip** When planning for future utilization, plan for more disk space, more memory, more network bandwidth, and more processor power than is indicated by your statistical analysis alone. Trend analysis is helpful, but may not take into account that as time passes, operating systems and applications become more complex and require more system resources.

The two primary methods you can use to gather statistics using Performance Monitor are running Performance Monitor individually on each server to be monitored, and running Performance Monitor on a single, centralized Windows NT computer to monitor several servers simultaneously.

You can run Performance Monitor individually on each server to be monitored on the network. Be aware that when Performance Monitor is run on the computer being monitored, the memory, disk, and processor statistics obtained may be above normal levels because of the running of Performance Monitor itself.

You can also run Performance Monitor on a single, centralized Windows NT Server or Workstation computer to monitor several servers on the network simultaneously. When Performance Monitor is run on a different computer than the computer being monitored, the memory, disk, and processor statistics are accurate for the computer being monitored, but the network utilization statistics may be above normal levels. This is due to the sending of Performance Monitor statistics across the network to the centralized Windows NT computer from the servers being monitored.

When using Network Monitor, keep in mind that the version of Network Monitor that ships with Windows NT Server 4.0 can only be run on the computer being monitored—you *can't* monitor a remote computer by using this version of Network Monitor. If you want to monitor computers remotely, consider using the version of Network Monitor that ships with Microsoft Systems Management Server.

The four components that are of critical importance when capacity planning your server are memory, processor, disk, and network. Here's a closer look at each of these critical server components, and the specific statistics that you can use to monitor each.

Memory

Gathering statistics about a server's memory utilization is an important step when capacity planning your server. I recommend using the following counter to obtain this data:

○ **Memory Pages/sec:** This Performance Monitor counter is the primary counter that is used to monitor memory usage for capacity planning or troubleshooting purposes. It measures the amount of 4KB memory pages that are read from or written to the paging file during a one-second time period.

This counter is selected for monitoring in the Add to Chart, Add to Alert, Add to Report, and Add To Log dialog boxes in Performance Monitor. The Add to Chart dialog box is shown in 24-1. Notice that Memory is selected in the Object drop-down list box, and that Pages/sec is selected in the Counter list box.
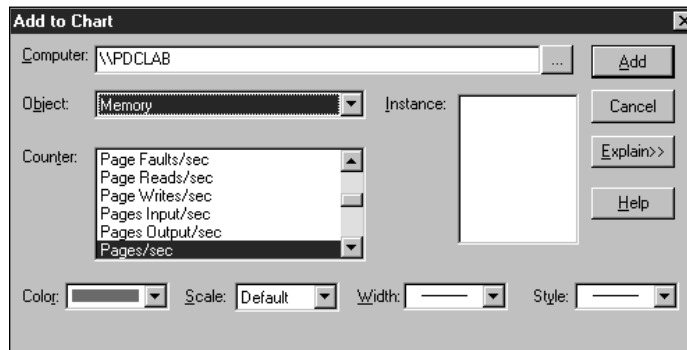

FIGURE 24-1   Selecting the Memory Pages/sec counter in Performance Monitor

○ **Interpreting the results of this counter:** A consistently low number for this counter (less than 2) indicates that the computer most likely has sufficient memory. A consistently high number for this counter (more than 2-3) may indicate that the computer does *not* have sufficient RAM.

in the
real world   **Memory is probably the most important component of a server. Insufficient RAM can cause excessive disk and processor utilization. When it appears that the computer's disk, processor, and memory are all functioning at or near maximum capacity, try adding additional RAM *first.* This usually alleviates some of the disk and processor load because paging is significantly reduced when RAM is added.**

When in doubt, add RAM before proceeding to the more expensive processor and disk upgrades.

Processor

Gathering information about your server's processor performance is also important. I recommend using the following counters:

- **Processor % Processor Time:** This performance counter is the primary counter used to monitor processor utilization on a single-processor computer for capacity planning or troubleshooting purposes. It measures the percent of the time that the processor is actively used by processes other than the Idle process.

  This counter is selected for monitoring in the Add to Chart, Add to Alert, Add to Report, and Add To Log dialog boxes in Performance Monitor. The Add to Chart dialog box is shown in Figure 24-2. Notice that Processor is selected in the Object drop-down list box, and that % Processor Time is selected in the Counter list box.
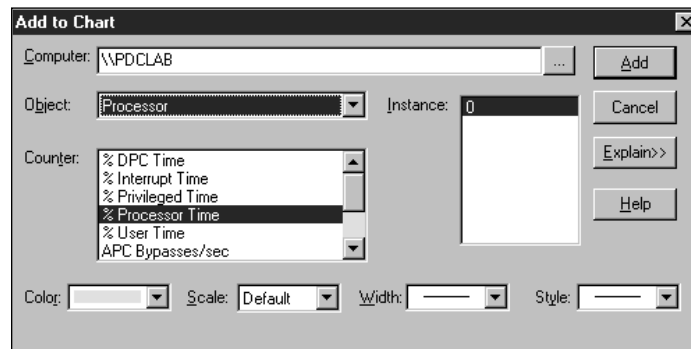

FIGURE 24-2   Selecting the Processor % Processor Time counter in Performance Monitor

- **System % Total Processor Time:** This is an additional Performance Monitor counter you can use to monitor overall processor utilization on a multiprocessor computer. It measures the percent of the time that all processors in the computer are actively used by processes other than the Idle process. Figure 24-3 shows this counter being selected in the Add to Chart dialog box.
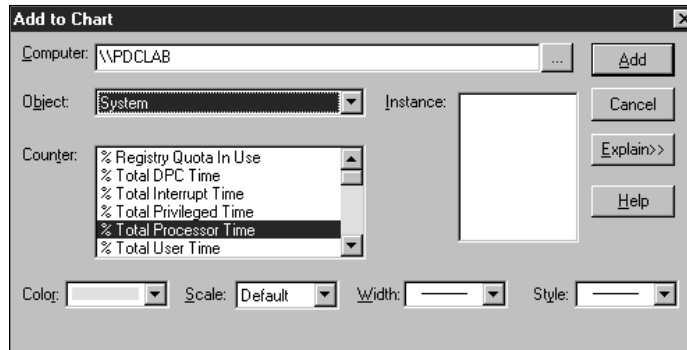

FIGURE 24-3   Selecting the System % Total Processor Time
counter in Performance Monitor

- **Interpreting the results of these counters:** A consistently low number for either of these counters (less than 60%) indicates that the computer most likely has sufficient processor power. A consistently high number for either of these counters (approaching 100%) may indicate that the computer does *not* have sufficient processor power. Before upgrading your computer's processor(s), check memory utilization first. Upgrade RAM, if necessary, *before* upgrading your processor(s).

Disk

There are three primary counters that can be used to monitor disk utilization on a server.

- **PhysicalDisk Avg. Disk Queue Length:** This Performance Monitor counter measures the average number of disk reads and writes waiting to be performed.

  This counter is selected for monitoring in the Add to Chart, Add to Alert, Add to Report, and Add To Log dialog boxes in Performance Monitor. The Add to Chart dialog box is shown in Figure 24-4. Notice that PhysicalDisk

is selected in the Object drop-down list box, that Avg. Disk Queue Length is selected in the Counter list box, and that Total is selected in the Instance list box.
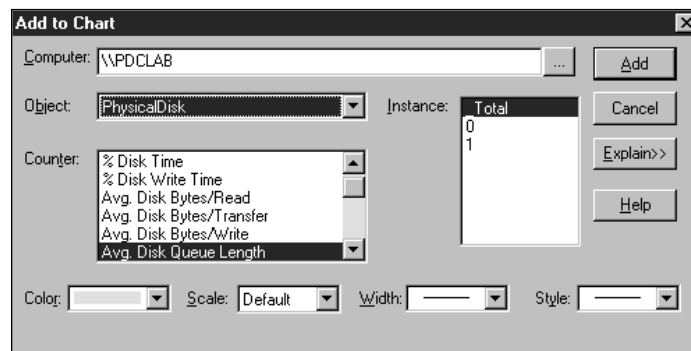

FIGURE 24-4   Selecting the PhysicalDisk Avg. Disk Queue Length counter in Performance Monitor

- **Interpreting the results of this counter:** A consistently high number for this counter (greater than 2–3) may indicate that a faster hard disk and/or hard disk controller is required for adequate system performance.

- **PhysicalDisk % Disk Time:** This Performance Monitor counter measures the percentage of time that the disk performs reads and writes. Figure 24-5 shows this counter being selected in the Add to Chart dialog box.
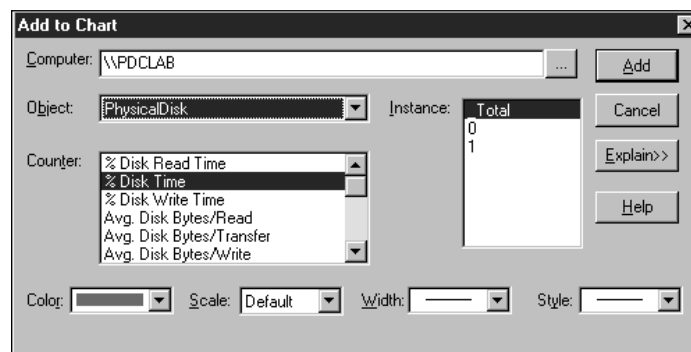

FIGURE 24-5   Selecting the PhysicalDisk % Disk Time counter in Performance Monitor

- **Interpreting the results of this counter:** A consistently high number for this counter (between 70% and 100%) may indicate that a faster hard disk and/or hard disk controller is required for adequate system performance.

- **LogicalDisk % Free Space:** This Performance Monitor counter measures the percentage of unused disk space. Figure 24-6 shows this counter being selected in the Add to Chart dialog box. Notice that LogicalDisk is selected in the Object drop-down list box, that % Free Space is selected in the Counter list box, and that Total ==> Total is selected in the Instance list box.
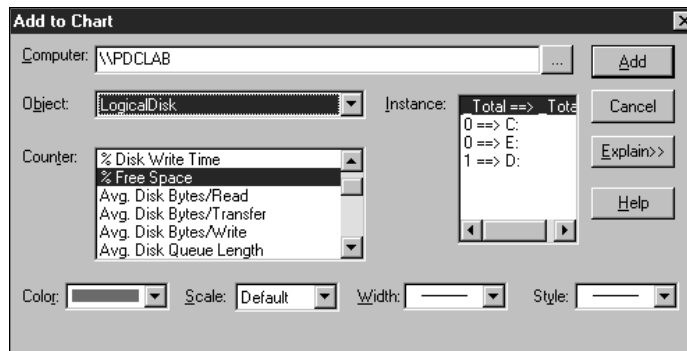

FIGURE 24-6   Selecting the LogicalDisk % Free Space counter in Performance Monitor

- **Interpreting the results of this counter:** A consistently high or gradually increasing number for this counter (approaching 100%) indicates that the server does *not* have sufficient disk space available. An additional disk, or a replacement disk that has more capacity, may be required.

Network

There are two primary Performance Monitor counters for monitoring network utilization.

- **Network Segment % Network utilization:** This Performance Monitor counter measures the total network utilization on a given network segment as a percentage of the maximum amount of network traffic possible on that segment. If your server is connected to more than one network segment, each segment should be monitored.

This counter is selected for monitoring in the Add to Chart, Add to Alert, Add to Report, and Add To Log dialog boxes in Performance Monitor. The Add to Chart dialog box is shown in Figure 24-7. Notice that Network Segment is selected in the Object drop-down list box, that % Network utilization is selected in the Counter list box, and that the first network adapter on the server is selected in the Instance list box.
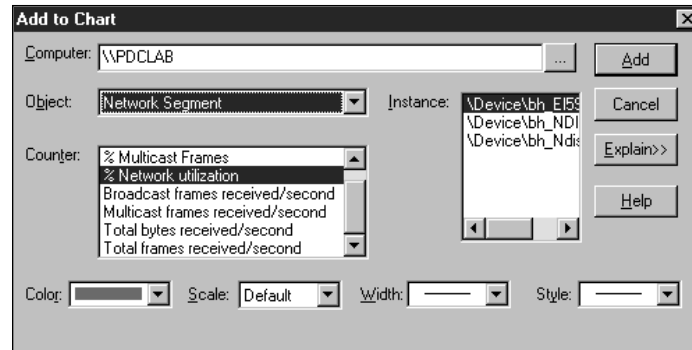

FIGURE 24-7   Selecting the Network Segment % Network Utilization counter in Performance Monitor

- **Interpreting the results of this counter:** A consistently high number for this counter (between 60% and 100%) may indicate that there are too many computers (or too much network traffic) on the network segment, and that an additional network adapter may need to be installed in the server, or a router may need to be installed on the network to further segment the network.

- **Server Bytes Total/sec:** This Performance Monitor counter can be used in place of the Network Segment % Network Utilization counter if the Network Monitor Agent is not installed on the server. This counter measures the total amount of bytes sent to and received by the server. If this server is the only server on the network segment (and the server is only connected to one network segment), the measurement obtained from this counter approximates the total network utilization on that network segment. (The counter is said to approximate total network utilization because it is assumed that most network traffic is to and from the server.)

Figure 24-8 shows this counter being selected in the Add to Chart dialog box.
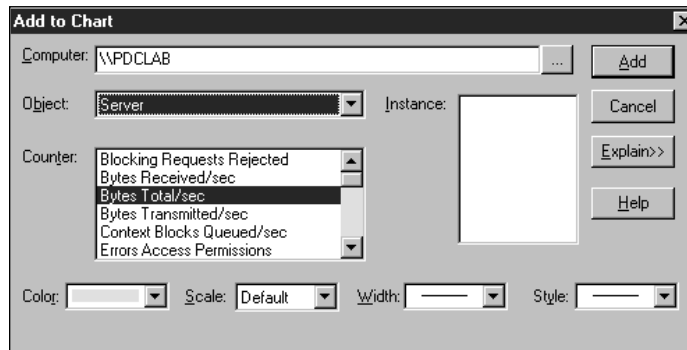
FIGURE 24-8 Selecting the Server Bytes Total/sec counter in
Performance Monitor

○ **Interpreting the results of this counter:** A consistently high number for
this counter (approaching the maximum number of bytes that can be sent
on the network segment in one second) may indicate that there are too
many computers (or too much network traffic) on the network segment,
and that an additional network adapter may need to be installed in the
server, or a router may need to be installed on the network to further
segment the network.

Network Monitor can also be used to monitor network usage. You can view
the Network Statistics section in the Total Stats pane after a capture is completed.
Using the information presented in this section, you can determine the average
number of bytes transmitted per second by dividing the number of bytes shown by
the number of seconds shown in the Time Elapsed statistic at the top of the Total
Stats pane.

If the average number of bytes per second approaches the segment's maxi-
mum capacity (an Ethernet 10BaseT segment, for example, has a maximum capac-
ity of 10Mbps or 1,250,000 bytes per second), this may indicate that there are too
many computers or too much network traffic on the segment being analyzed.

Capacity Planning Your Network

Capacity planning your network involves using Performance Monitor and Network
Monitor to gather network statistics over time to establish a baseline and to pro-
vide additional data for trend analysis. You can analyze these statistics to deter-

mine if the network's performance is adequate now, and if it is, to determine approximately when the network will need to be upgraded.

Capacity planning a network typically falls into one of two categories: capacity planning *overall* network utilization, or capacity planning for a *specific type* of network traffic.

Capacity Planning Overall Network Utilization

There are two primary Performance Monitor counters for monitoring overall network utilization: Network Segment % Network utilization, and Server Bytes Total/sec.

concept link   **Detailed information about these counters is presented in the "Network" section earlier in this chapter.**

Network Monitor can also be used to monitor network usage. The Network Statistics section in the Total Stats pane provides various network utilization statistics.

You can use these Performance Monitor and Network Monitor statistics to determine your network's overall baseline performance, and also for trend analysis that will indicate when the network should be upgraded to meet projected overall network utilization needs.

Capacity Planning for a Specific Type of Network Traffic

You might occasionally want to perform capacity planning for a specific type of network traffic, such as browsing, name resolution, or broadcasts. For example, if you are planning a new WAN link, you might be concerned about the amount of browsing traffic that will exist on the WAN link once it is installed.

To begin the process of capacity planning your new WAN link for browsing traffic, configure Network Monitor to capture and analyze all browsing traffic on an existing network. Once the browsing packets are captured, Network Monitor is used to determine how large each of these packets is, how many browsing packets are transmitted per workstation and per server on the network, and how frequently browsing packets are transmitted.

Once you have determined the amount of browsing traffic that each workstation and server generates over a given time period, you can use this information, in conjunction with your planned network configuration (including the number of workstations and servers that will be used on each side of the WAN link) to project how much browsing traffic will be sent over the new WAN link.

Network Monitor is the primary tool for capacity planning specific types of network traffic, because you can configure it to capture and display only the types of network traffic you want to analyze. Our example features browsing, but you can use Network Monitor to analyze virtually any type of network traffic, including: WINS and DNS name resolution traffic, broadcast traffic, NetBIOS multicast traffic, IP address assignment traffic (DHCP), Directory Services synchronization traffic, and so on.
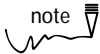
note ▽ **To use Network Monitor for this type of capacity planning, you may need to use the version of Network Monitor that ships with Systems Management Server. (The version that ships with Windows NT Server 4.0 is designed to only capture packets sent to and from the Windows NT Server computer that is running Network Monitor.)**

concept link **For more information on how to use Network Monitor, see Chapter 23.**

Key Point Summary

This chapter introduced the basics of capacity planning. The following points highlight the major issues.

- *Capacity planning* is the process of determining current usage of server and/or network resources, as well as tracking utilization over time to predict future usage and the additional hardware that will be required to meet projected needs. Capacity planning can be performed on a single computer, such as a network server, or it can be performed on an entire network.

- Performance Monitor and Network Monitor are the two primary Windows NT tools that can be used for capacity planning.

- The basic steps involved in capacity planning are establishing a baseline, gathering data over time, and using this data to predict future utilization and future hardware requirements.

  - Establishing a baseline determines the status quo. Consider gathering several samples of data during both business and nonbusiness hours when establishing an initial baseline. Once you've established a baseline, continue gathering data over time to obtain additional statistics that can later be used for trend analysis.

  - Finally, use the data you've gathered to predict future utilization. It is usually beneficial to export this data to a comma-separated value or tab-separated value file, which can then be imported directly into most spreadsheets and databases for trend analysis.

- Capacity planning your server involves using Performance Monitor and Network Monitor to gather server and network statistics over time. These statistics, when analyzed, help you to determine if the server's performance is adequate now, and, if it is, to determine approximately when the server will need to be upgraded.

- The two primary methods you can use to capacity plan your server by using Performance Monitor are running Performance Monitor individually on each server to be monitored; and running Performance Monitor on a single, centralized Windows NT computer to monitor several servers simultaneously.

  - When Performance Monitor is run individually on each server to be monitored, the memory, disk, and processor statistics produced may be above normal levels due to the running of Performance Monitor itself. When Performance Monitor is run on a single, centralized computer, network utilization statistics may be above normal levels due to the sending of statistics across the network to the centralized computer running Performance Monitor.

- The four components that are of critical importance when capacity planning your server are memory, processor, disk, and network.

  - *Memory:* The primary Performance Monitor counter used to monitor memory utilization is Memory Pages/sec. A consistently high number for this counter (greater than 2-3) may indicate that the server doesn't have sufficient RAM.

- *Processor:* The primary Performance Monitor counter used to monitor processor utilization on a single-processor computer is Processor % Processor Time. The primary Performance Monitor counter used to monitor processor utilization on a multi-processor computer is System % Total Processor Time. A consistently high number for either of these counters (approaching 100%) may indicate that the computer doesn't have sufficient processor power.

- *Disk:* The three primary Performance Monitor counters used to monitor disk utilization on a server are: PhysicalDisk Avg. Disk Queue Length, PhysicalDisk % Disk Time, and LogicalDisk % Free Space. The first two counters indicate whether the speed of the disk is adequate, and the third counter indicates whether the disk has sufficient capacity.

- *Network:* The two primary Performance Monitor counters used to monitor network utilization are Network Segment % Network utilization and Server Bytes Total/sec. A consistently high number for either of these counters may indicate there are too many computers or too much network traffic on the network segment. Network Monitor can also be used to provide network usage statistics.

- Capacity planning your network involves using Performance Monitor and Network Monitor to gather network statistics over time. These statistics, when analyzed, help you determine if the network's performance is adequate now, and if it is, determine approximately when the network will need to be upgraded.

- Capacity planning a network typically falls into one of two categories: capacity planning *overall* network utilization, or capacity planning for a *specific type* of network traffic.

  - When capacity planning overall network utilization, two primary Performance Monitor counters are used. These counters are Network Segment % Network utilization and Server Bytes Total/sec. Network Monitor can also be used to monitor network usage.

  - When capacity planning for a specific type of network traffic, such as browsing, name resolution, or broadcasts, Network Monitor is often used. You can configure Network Monitor to capture and display only the specific type of network traffic you want to analyze.

Applying What You've Learned

Now it's time to regroup, review, and apply what you've learned in this chapter.

The following Instant Assessment questions will bring to mind key capacity planning facts and concepts that you've learned in this chapter.

Instant Assessment

1. What is capacity planning?

2. What are the three basic steps in capacity planning?

3. When establishing a baseline, when should you consider collecting data?

4. Performance Monitor enables you to export data directly to two types of files that can be imported directly into most spreadsheets and databases.

   What are these two types of files called?

5. What are the two primary methods that you can use to gather server and network statistics by using Performance Monitor?

6. What is the primary Windows NT administrative tool used when capacity planning for a specific type of network traffic?

7. List two categories of network capacity planning.

8. List the four components that are of critical importance when capacity planning a Windows NT server.

9. What is the primary Performance Monitor counter used to monitor memory usage for capacity planning purposes?

10. Which Performance Monitor counter is used to monitor processor utilization on a single-processor computer?

11. List three Performance Monitor counters that can be used to monitor disk utilization.

12. What is the primary Performance Monitor counter used to monitor network utilization?

T/F

13. Running Performance Monitor on a single, centralized
Windows NT computer to monitor several servers on
the network simultaneously can increase the network
utilization statistics above normal levels.   _____

14. Running Performance Monitor on each individual
server to be monitored can increase the memory, disk,
and processor statistics of these computers above normal
levels due to the running of Performance Monitor itself.   _____

concept link   **For answers to the Instant Assessment questions see Appendix D.**