

Enterprise Network Monitoring and Analysis in a Mission-Critical Environment.

NETWORKS AND HOW PEOPLE USE NETWORKS HAVE CHANGE considerably over the years. Network environments have also evolved and matured to the point where the focus has shifted from technology adoption and implementation issues to management activities needed to support business end users in a mission-critical environment. As a result, the strategic role that information technology (IT) plays in delivering business continuity across the company has been elevated from an afterthought to the group that adds a competitive edge to the corporation. The role of the chief information officer (CIO) (and therefore all IT personnel) has reflected this change. CIOs used to report to the chief financial officer (CFO) and were thought of as a facilities-type of people, whereas now CIOs now usually report directly to the chief executive officer (CEO) and can bring about changes to the entire corporation.

Business end users are also starting to expect error-free network connectivity with guaranteed uptime and response time. These users also expect network services to be delivered regardless of the underlying technology. This increase in reliance on client/server applications as a fundamental part of conducting business means that end users need to rely on the network as much as the phone. The combination of the maturing of the underlying technology and the end users' expectations means that the monitoring and analyzing of the network is a critical function of any IT group. This chapter discusses the role the network monitoring and analysis takes in administering networks. It starts by explaining the concept of service-level management (SLM) and where network monitoring fits into that "framework." We then show the range of functionality that network monitoring brings to the IT manager's arsenal.

SERVICE-LEVEL MANAGEMENT

A common misconception of IT managers is that it is the technology that provides the correct level of service to the business, but actually it is the correct implementation of a technology. IT managers (and engineers) must make the transition from the latest technology provider to providing the best service for the business. In many cases, applying the latest technology will not improve the service that is being provide. What is required is a strategy and a methodology for providing the correct service to the business.

IT service management is based on defining, achieving, and maintaining required levels of IT service to the business user population with the company. At this point in time, few client server IT organizations have realized that IT service management is the only strategy that allows them to meet the business end users' needs. Another way of looking at IT service management is that it is the "glue" that allows IT to align itself with the businesses that it supports. The language that IT and business groups communicate through is service-level agreements.

By implementing service-level agreements between IT and the business groups, one can increase the efficiency of IT staff and further automate and control the process of meeting service-level agreements. Because service-level management helps align multiple IT management processes to meet service objectives, it becomes increasingly important to coordinate the work of the IT staff toward that end.

IT organizations are typically very tactical. This stems primarily from the reactive “firefighting” mentality that historically IT has needed when implementing new technologies. A reactive approach to IT management is very tactical, not strategic. IT should aim to move to a proactive approach, and that requires planning. Planning requires setting objectives based on a clear vision. Objectives stem from a clearly defined set of strategies that can be expanded into implementation tactics to support the objectives defined in the planning process. Much can be learned from this methodology when tackling the problem of managing the distributed IT environment.

SERVICE-LEVEL AGREEMENTS

The “language” that can bind IT organization to the business is the service-level agreement (SLA). SLAs are documents that define the various levels of service that IT must deliver to end users. SLAs are written for individual applications from end to end. A business can have multiple SLAs for each application (i.e., web, E-commerce, order entry). One of the challenges IT organizations face is proper decompositions of the problem into manageable parts. This can be especially difficult in today’s client-server environment. Although there are many components to the SLA, not all are required. Many elements an SLA may contain include the following:

- Service volume. IT needs to quantify the volume of the service to be provided, such as average and peak rates and the time of day the demand is expected to occur. The business may also be provided with the incentive to receive better service, or a reduced cost for service, by avoiding peak resource usage periods (similar to discount phone rates at night). Being able to determine the volume of service allows IT to determine the infrastructure needed to support that service (i.e., do we need a two-lane road or a six-lane expressway).
- Service timeliness. IT needs a qualitative measure for most applications to be able to measure performance of the application. For some applications (i.e., an automated teller machine) it should be “90 percent of transactions processed within two seconds.”
- Service availability. When will the service be available to the user? IT must be able to account for both planned and unplanned downtime. The business must be able to specify when they expect the system to be available to achieve their specified levels of work.
- Service limitations. It is too costly to normally supply a service without some form of limitation. The limitations to the service are documented here.
- Service compensation. This is the penalty clause area. If the service is not provided, then some form of remuneration should be documented. Also, there is a

cost to providing the service. This cost can be recovered with a chargeback system, example.

- Measurement of service. This section describes the monitoring process by which service levels will be compared against the agreed-upon service levels. IT must define how the service levels will be monitored and the frequency with which normal monitoring will take place.

SERVICE-LEVEL OBJECTIVES

For IT managers, being able to define service-level objectives (SLOs) from the SLA is the critical aspect of IT service management. SLOs are solely the domain of the IT manager and his or her staff. They are derived from what the stated business user service levels need to be based upon the development of the SLA. From the SLOs, IT can define the metrics (system, network, database, and application) that it needs to collect, monitor, measure, and report on. Some examples of metrics in a SLO include the following:

- Application availability: This metric reflects the application availability from the end users' standpoint. This metric can be different depending on the implementation of the application (i.e., client/server versus stand-alone).
- Application performance: Monitoring and measuring the applications to determine if user response time meets the service level specified in the SLA against specific business transactions. A volume metric can also be defined if defined in the SLA.
- Application security: Mechanisms must be put in place to insure secure access to certain applications due to company confidentiality or competitive threats. Measures to the effectiveness of this system must be put in place and monitored against service objectives (e.g., no unauthorized access, different levels of access to certain parts of an application, etc.).
- Application reliability: Measurements need to be established and collected to determine the accuracy of an application. This may require periodic sampling of the work accomplished by the application or may require the application to be instrumented to provide those measures in real time.

SLOs are defined from SLAs. An important aspect is defining the proper metrics to measure service objective compliance. This is especially true in the distributed environment where the components of the infrastructure that need to be measured are many (network, system, application, and database) and geographically dispersed. Using this method to refine the metrics needed means increase focus can be given to the quality of the metrics rather than the quantity. It also allows for a phased implementation of service-level monitoring and what it means.

The bottom line is this: focus on getting the right metric for the management task at hand. This implies the task is understood and the desired solution has been well-

defined. Ignore the urge to implement all management tools and collect every piece of data all the time with the thought that someday it may be needed.

One resource for finding more detail on service management processes and procedures is the IT Information Library. This is the ultimate guideline on IT service management, including recommendations for managing people, processes, and tools. The Information Technology Service Management Forum (IT SMF) is a global consortium of more than 400 international corporations responsible for advancing the IT Information Library.

Service management is important to the network manager, as it allows the correct resources to be allocated to the correct business. Service management allows the IT manager to decide where to allocate resources, what level of support is needed.

NETWORK MONITORING AND ANALYSIS DEFINED

Distributed network monitoring is the ability to view a remote network and perform monitoring and analysis on that remote network as if it were local. In the past, portable devices were carried to remote sites and placed on to the network when problems occurred on that segment. Having a network monitoring device on a segment only when there are problems means that the segment is not monitored 99 percent of the time. Monitoring devices permanently placed on mission-critical segments can constantly monitor traffic. That means analysis can be performed over and above fault management.

The Exhibit 28-1 shows an example of remote monitoring agents installed on a large enterprise network with a variety of media types such as WAN's, Switches and Media types such as FDDI and Ethernet.

The agents, or "probes," reside on the remote segments and collect information on the traffic that it sees. The segments can be of any media type from various local-area network (LAN) media types, such as Ethernet, FDDI, Token Ring, or some WAN protocol such as frame relay. The segments can be geographically dispersed, but in general must be interconnected. The network management console contains a suite of applications that collect the network information from these remote agents and interprets then using power graphical user interfaces. Interestingly, the network management console communicates with agents using the same network that the agents are monitoring. (Out-of-band communication between the manager and the agents is also possible.)

With this configuration, network administrators can use monitoring tools to manage the whole network. Some functions a network administrator can perform are as follows:

- Network performance management: The ability to continuously monitor certain network statistics to ensure adherence to the SLA. Setting network thresholds to identify anomalies and creating baselines to aid in determining “normal” network performance.
- Network security monitoring: ensuring that only authorized users access the network. This includes monitoring the effectiveness of fire walls as well as internal security monitoring.
- Fault management and availability: Being able to troubleshoot network problems in timely fashion and monitor the availability of servers from end users’ perspective.
- Network service simulation: Traffic profile modeling allows a network manager to do a quick “what-if” analysis before reconfiguring network resources. Having the appropriate data of past network trends determines what changes need to be made to handle the ever-growing network growth.
- Policy-based management: Being able to control who gets the limited amount of network bandwidth has always been a goal of network administrators, but until now some of the recent standards have been almost impossible to implement. By implementing some of the recent standards, true policy-base management is at the network manager’s fingertips.

NETWORK MONITORING AND ANALYSIS IN THE IT ENVIROMENT

The IT management environment covers the whole range of devices that reside on the network as well as the network that enable business end users to function. We can break this down into four components:

- Systems management: This is concerned with the performance of the computers on the network and usually deals with issues such as data-base performance and disk use on file servers.
- Element management: this is concerned with managing the various networking devices, such as bridges, routers, and hubs. Typical management issues deal with configuration tables, throughput, link states, and port partitioning. A device management application usually shows a picture of the device on your screen, complete with installed cards and indicator lights.
- Desktop management: This is concerned with the end-user workstations and PCs. The management issues are PC configuration files, disk use, application support, etc.
- Network monitoring and analysis: This is primarily concerned with the activity on the wire. It looks at the flow of data across the network in an effort to understand network performance and capacity and to resolve problems related to networking protocols.

Network monitoring and analysis allows the IT department to manage one part of the end-to-end management picture. System, database, and application management issues are not discussed in this chapter.

STANDARDS OVERVIEW

Network monitoring has benefited from several standards. The main standard in use for network monitoring is the remote monitoring (RMON) standard, which defines a method of monitoring traffic up to the DataLink layer (Layer 2) in the Open Systems Internet (OSI) stack, the RMON2 standard, which has not yet been ratified by the Internet Engineering Task Force (IETF) defines how to monitor traffic at the network layer (OSI Layer 3) and some portions of the application layer (Layer 7).

- Simple Network Management Protocol (SNMP)
- Simple Network Management Protocol version 2 (SNMPv2)
- Remote Monitoring (RMON) standard
- Remote Monitoring version 2 (RMON2)

Why Do Network Monitoring?

As part of an IT departments SLA with its business end users, IT must maintain a certain level of network service. To be able to do this, the network must be monitored to ensure error-free connectivity, responsiveness, and level of throughput. If the network is not monitored, it would be impossible for the IT department to guarantee any level of service.

In today's competitive environment, new client/server applications are quickly appearing in business environments; some examples are the World Wide Web and Lotus Development Corp.'s Notes. If the network is not being monitored, then the effect of adding one of this network-intensive applications is unknown and eventually one will bring the network to its knees. If the environment is being monitored, network bandwidth will always exceed future growth.

The ability to monitor trends changes IT from being reactive-waiting until something breaks before resolving the problem- to being proactive- resolving potential issues before they break. The IT department should now blend into the background, allowing business end users to focus on their functions.

Who Does Network Monitoring?

Because there are many parts to network monitoring, many people are involved. Here are some generic descriptions:

- **Network Manager:** Responsible for long-term strategic decisions regarding the network. Involved in looking at new technologies, such as 100Base-X or

asynchronous transfer mode (ATM), deciding where and when to modify bandwidth. This person tends to look at network trends, performing forecasting and capacity planning.

- Network engineer: Responsible for day-to-day operations of the network. Upgrades network devices, adds capacity. Also acts as a second line of support for problems that the operations center engineer cannot resolve.
- Operations center engineer: Most large corporations have a centralized monitoring center that is staffed with “level 1” engineers that attempt basic troubleshooting on problems. These engineers monitor for events that are triggered by servers, workstations, or network devices that can alert the operations center on potential problems. These engineers are the first line of support and are constantly in reactive mode.

What Data Is Provided?

Monitoring the network means that information on every packet on every segment can be gathered. Network monitoring really means deciding which data is important and should be gathered and which data is redundant. Corporations with many segments need to decide on only a few critical pieces of information, otherwise they are inundated with data. The cost of analyzing the network would exceed the actual cost of the network. Some of the most critical measurements that should be gathered are as follows:

- Utilization: Segment utilization information should be gathered to generate trends for capacity planning purposes, baselining purposes, performance information.
- Error rates: Total error rate information can give performance indicators; baselining the error rate of the network, correlated with utilization, can give indicators of physical layer network problems.
- Protocol distribution: This can generate trends for changing application mixes; monitoring the usage of new applications and the effect new applications on the network.
- Top talkers: These can also give indications on the performance of the network, performance of machines, load of application, and services on the network. Top talkers can also indicate potential new applications that are unknown to the network department (new Internet applications such as PointCast have been discovered using this method).
- Latency measurements (echo test): These lead to trends in performance.

How Does Network Monitoring Work?

Network monitoring is a large subject, and there are many proprietary protocols involved. This chapter covers only standards based protocols, plus the most widespread proprietary protocols.

The Simple Network Management Protocol (SNMP)

The Simple Network Management Protocol (SNMP) was a draft standard in 1988 and was finally ratified in April 1989. SNMP is described by Request For Comments (RFC) 1098. SNMP has three basic components:

- Agent: A software program that resides in managed element of the network such as a hub, router, or specialized device.
- Manager: This communicates with the agent using the SNMP commands.
- Management information base (MIB): A database that resides with the agent and holds relevant management information.

Exhibit 28-2 shows the relationship between those three components.

There are five types of SNMP commands, which are called protocol data units (PDUs):

1. Get request: A manager requests (from the agent) the value of a variable stored in the MIB.
2. Get-Next request: The manager uses this to request information in multiple variables. Used to reduce network traffic. If one variable is not available, no values are returned. It is also used to retrieve unknown rows if available.
3. Set request: The manager instructs the agent to set MIB variable to the desired value.
4. Get-response: This is send by the agent as a response to a secure electronic transaction (SET) or Get-Next command as either an error or identical to the Set to show it was accepted, or to a Get-Next with the value portions of the request filled in. the manager checks its list of previously send requests to locate the one that matches this response is discarded, otherwise it is handled.
5. Trap: One of two unsolicited messages send from the agent to the manager, often used for event notification.

THE MANAGEMENT INFORMATION BASE TREE

MIBs are hierarchical in nature (see Exhibit 28-3). This allows unique identifiers for each MIB variable (or Object). Some MIBs of interest are as follows:

- RFC1213 – MIB11 – basic system information and basic level statistics
- RFC1757 – RMON (Remote Monitoring)
- RFC1513 – RMON (Remote Monitoring) extension for Token Ring.

There are several advantages that network management applications have with SNMP:

- The protocol is easy to implement.
- The protocol requires few resources to operate.
- The protocol is mature, stable, and well-understood.

- The protocol is widely available (on most computers), and most network devices have some form of agent/MIB embedded within them.

However, as networks have grown and the need for network management has become more imperative several disadvantages with SNMP have become apparent. Some of these disadvantages include the following:

- limited security
- lack of block transfer
- polling-base protocol
- trap limitations

SNMPv2

SNMPv2 is a proposed standard that attempts to address these issues. Some of the proposed extensions to the standard include the following:

- Manager-to-manager communication capability
- Additional SNMP commands (PDUs):
 - Get BulkRequest – for getting whole tables
 - InformRequest – A manager-to-manager PDU
- Reliable traps

The last area of contention with SNMPv2 is security. To proposed drafts attempt to address the security issue.

THE REMOTE MONITORING PROTOCOL (RMON)

The RMON standard is specific standard for performing remote monitoring on networks. The RMON standard is defined by two standards RFC 1757 and RFC 1513. The standard defines a MIB that is broken down into ten groups, the first nine define monitoring of Ethernet networks and the tenth defines extensions for Token Ring there are no standards for monitoring FDDI, 100Base-X, or WAN networks. RMON vendors have added their own proprietary extensions for these additional media types. RMON is limited, as it gives visibility only up to the data layer (Layer 2) in the OSI stack.

- Statistics group: This group contains many segment statistics in 32-bit counters, such as packets, dropped packets, broadcasts, and multicasts. These are just counters, not studies.
- History group: This group contains segment history statistics for various counters such as broadcast, errors, multicasts, packets, and octets. These numbers are for

certain time periods. RMON defines two default time periods – 5 seconds and 1800seconds.

- Alarms group: this covers threshold monitoring and trap generation when that threshold has been reached. It allows alarms to be set of various counters and patch match. Traps can start and stop packet capture.
- Host group: This contains host table and traffic statistic counters, plus a time table of discovery.
- Host top N: This contains studies for X time and X hosts, listing top talker for the study group.
- Traffic matrix group: This group contains matrix of Medium Access Control (MAC) layer (Layer 2) conversations. Information such as error, packets, and octets sorted by MAC address.
- Packet capture/filter group: these two groups are use together. Packet capture group contains the packets that have been captured. Multiple instances can be created.
- Token Ring group: Contains specific information about Token Ring, such as ring order, ring station table, and packet size distribution for history studies.

Remote Monitoring version 2 (RMON2) Protocol

The RMON standard brought many benefits to the network monitoring community, but it also left out many features. The RMON2 standard tries to address this (see Exhibit 28-4) by allowing the monitoring of layer 3 (Network Layer) information as well as protocol distribution up to Layer 7 (Application Layer).

NETWORK PERFORMANCE MANAGEMENT

Performance management means being able to monitor segment activity as well as intrasegment traffic analysis. Network managers must be able to examine traffic patterns by source, destination, conversations, protocol/application type, and segment statistics such as utilization and error rates. Network managers must define the performance goals and how notification of performance problems should happen and with what tolerances. Some objectives that network managers are faced with are as follows:

- Baselineing and network trending: How to determine the true operating envelope for the network by defining certain measurements such as segment utilization, error rate, and network latency, to check SLOs and irregular conditions that, if gone unchecked, may have drastic consequences on network business users productivity.
- Application usage and analysis: This helps managers answer questions such as “What is the overall load of your WWW traffic?” and “What times of the day do certain applications load the network?” This allows network managers to discover

- important performance information (either real-time or historic) that will help define performance SLOs for applications in the client/server environment.
- Internetwork perspective: Is traffic between remote sites and interconnect devices critical to the business? With internetwork perspective capabilities one can discover traffic rates between subnets and find out which nodes use WAN link to communicate. It can also help one define “typical” rates between interconnect devices. Internetwork perspective can show how certain applications use the critical interconnect paths and define “normal” WAN use for applications.
 - Data correlation: This allows one to select peak network usage points throughout the day and discover which nodes contributed to the network load at that peak point in time, which nodes they are sending traffic to, and which applications were running between them.

Exhibit 28-5 shows an example of traffic flow between several segments. The thickness of the line indicates the volume of traffic. With this information it is easy to identify potential WAN bottlenecks.

Exhibit 28-6 shows clients and servers correlated with a time graph. Being able to determine how much a particular server affects the network can help in the positioning of that server and again improve performance.

Network Security Monitoring

Security management encompasses a broad set of access control policies that span network hosts, network elements, and network access points (firewalls). Consistent policies are the key here; the objective is to support access and connectivity that is appropriate to the business need while restricting clearly inappropriate network-base access. As in other activities, constant monitoring for specific violations is critical, as is a notification mechanism. For certain conditions, immediate, automatic action may be required (i.e., “Shut down this connection” or “Shut down the firewall”). Monitoring should include both passive and active monitoring (probing).

Access level monitoring ensures that the controls and security that are in place actually perform to expectations. Monitoring the traffic flow to a firewall for instance, ensures that no intruders access it internally. Access level monitoring polices the “police” and ensures that nothing has been overlooked by the security. (See Exhibit 28-7).

Management and Availability

Fault. Fault management is the continuous monitoring of the network and its elements and the detection of failures within the network environment. When a failure is detected, notification of the failure must occur in a timely fashion. The failure must be qualified with respect to other failures and prioritized.

Fault management systems include software to detect and notify a centralized system of these failures. The centralized system normally includes some form of discover and mapping software that allows the network manager to have a graphical view of the network. These notifications must be correlated so that event storms are eliminated. A trouble ticketing system can also be incorporated so that a document trail is kept to the problem and a mechanism can communicate the status of the problem to the end users.

Another aspect to fault management is availability. This is the monitoring of servers from business and users' perspectives to ensure that the machine is available to the end users. Tracking and notification of any interruption of client/ server access is a critical part of the IT department's function.

Capacity planning. Network demand is growing at unprecedented rates. New applications, such as SAP and the World Wide Web (WWW), are encouraging extensive use of the network. Graphics are now sent regularly over the network (either through a corporation's intranet or over the internet). As network managers increase bandwidth, new applications for the network (such as voice-over-IP or multimedia) become viable. This causes another spurt to demand for the network.

Capacity planning allows the network manager to look forward by looking at the past and helps the manager to forecast what the demand will be. This means the IT department can keep one step ahead of demand

Network Reporting. Part of the IT department's function is to demonstrate that members perform their functions to the prescribe level. Being able to document that the network runs at the level agreed to in the SLA is an important justification tool.

It is critical that reports are generated automatically, otherwise reports will not be generated or the effort to generate them will be to substantial to make it worthwhile.

NETWORK SERVICE SIMULATION

Being able to move from reactive troubleshooting to a proactive approach is critical to the success of IT. Part of this is being able to predict accurately the impact of network and application changes in a dynamic network environment. By simulating the effect of network and network centric application changes, IT can intelligently provide the quality of service for its end users. By being able to simulate the network, it will be able to do the following:

- deploy network-centric applications with confidence
- choose and implement new technologies with assurance
- prepare network for an increasing number of users
- create attainable network SLOs
- reduce the risk of service disruptions and SLA violations

Replication of the Network

One issue with simulation is that it must model the network. A bad model will lead to bad simulation. The best simulators will take network topology as well as network performance data gathered from your own network. Creating network topology and network traffic from scratch is a fruitless exercise. The closer the representation of your existing network, the more accurately you can simulate your network.

Test-Driving New Technologies

Most simulation tools have a library of modeled network devices and media based on real-world characteristics. When testing new devices (such as the latest switch) or new technologies (such as Gigabit Ethernet), one should be able to place these devices/technologies within your simulated environment to see the changes. Using a high-fidelity model library allows network managers to test-drive products and new technologies, such as moving from Ethernet to fast Ethernet, moving to ATM, adding a FDDI backbone, or creating completely new network segments. This reduces the risk in implementing a new technology and determines the best combination of technologies that will provide the service your business demands.

New Application Load Simulation

Simulation can reduce the impact of deploying network-centric applications by simulating the result of the load in your network. Being able to simulate new applications (i.e., SAPR/3) or adding demand to existing applications (i.e., the Web) allows the network manager to determine if the current network infrastructure will be sufficient for the new load. This is critical to ensuring that any changes will not cause problems to the existing SLAs the IT manager already has in place.

Simulating Organizational scaling

Organizational restructuring is common in today's business. These changes can significantly affect the performance of the network and business. Simulation allows one to test the impact of adding new users and their applications to the network before the network change is even made. If the simulation contains both accurate network topology and traffic data, the simulation can anticipate end-users response times under differing scenarios, allowing you to plan for smooth transition while maintaining acceptable network service level for all users.

POLICY-BASED MANAGEMENT

Policy-based network management allows IT administrators to control who gets which network services when. This allows IT to manage a network to provide services according to business needs. For example, one user wants to surf the web, one wants to

process SAPR/3 transactions, and the last wants to submit an order. This means that three levels of service are needed to ensure that mission-critical traffic requirements are met while providing adequate services for less-critical traffic. Another aspect is that there are now more uses for the same network (voice, video, data). Voice-over-IP, videoconferencing, and E-commerce are now being merged onto the same network infrastructure. Policy-based network management allows network managers to set and configure policies to control their network environment to ensure the various SLAs are met. Many vendors (i.e., 3Com, Cisco Systems, etc.) now offer equipment (i.e., switches, routers, etc.) that understand policy-based management. Critical to the success of policy-based management is central management of these policies.

What is a Policy?

A policy is an association between a service and the rules that govern the use of or access to the service. Policies consist of two distinct elements:

1. The service, such as “priority bandwidth access,” “security actions,” or “access control.”
2. The rules describing the conditions under which the service is available, such as “application SAP has access to 10M bytes of bandwidth between 8 a.m. and 6 p.m.” A resource is the network element of software that provides the service.

Components of Policy-Based Network Management

Most policy-based network management implementations consists of a console, where the IT manager defines, edits, administers, and distributes the policies, and a set of agents that enforces the policies. In many cases the console is broken into two distinct components, a GUI and a policy server.

There are three generic types of agents:

1. Outsourcing agent. This runs on a network component (router, switch, etc.) and controls the resources in that device (i.e., bandwidth). For example, when an element receives a Resource Reservation Protocol (RSVP) request, that request is sent to the policy server using the Common Open Policy Service (COPS). The policy server decides how the request should be handled and sends that information back to the agent for enforcement.
2. Configuration agent. This normally runs on servers or machines with dedicated CPUs. This combines the policy server with the agent. An example is a policy-based software firewall that will make changes to its configuration as policies change.
3. Proxy agent. This agent takes the policy information and makes changes to the clients it manages (i.e., switches that are not “policy aware”) on behalf of the policy server. For example, using voice/IP, a port associated with a voice

application can be given a higher priority than a port associated with a data transfer, thereby avoiding the latency and jitter often associated with voice transmission over IP. Policy-based management is still in the early stages of its life cycle. Few corporations have widely implemented policy-based management strategies. Fast adoption will come as more corporations implement service level management and as voice, video, and data start to converge on the same network infrastructure over the coming months (and years). Many vendors claim policy-based management “aware” components, but in reality it is extremely difficult to implement policy-based management in the heterogeneous environment.

LIMITATIONS OF NETWORK MONITORING AND ANALYSIS

Monitoring the network with the RMON standard means the only data link layer (Layer 2) information is collected. This is not high enough in the OSI stack to gather information about traffic trends of client/server applications. RMON has reached a point where it is widely deployed within switches, but it is not at the level in the stack to add much value from an application monitoring view point.

The RMON2 standard defines a method of monitoring up to Layer 7 at certain times. RMON2 does not define continuous monitoring of all Layer 7 traffic, nor does RMON2 define any metrics for performance management.

As more types of traffic (i.e., voice, video) are mixed in to the network, current monitoring tools will lag behind in being able to diagnose problems with these types of traffic. This is where implementing SLAs and some form of policy-based management becomes critical to the success of the organization to ensure that mission-critical data is processed.

SUMMARY

Enterprise network management is a fast-changing environment. From the early days of monitoring the physical layer of the networks to the future of application-layer service-level management, the whole arena is helping IT management take control of the distributed environment that it spans.

Network management will always have several aspects that have been described in this chapter, and the tools for implementing SLAs between business end users and IT departments are quickly maturing.

However, network management is only part of the total end-to-end solution that must include the whole environment the business end users operate. This means systems, databases, and application monitoring tools must be deployed in conjunction with the network monitoring tools so that the whole environment can be viewed. Tools such as Hewlett-Packard’s PerfView are being released that for the first time and can integrate

seamlessly database, application, network, and system information on a single pane of glass for the end-to-end view that is necessary in this complex environment that IT must now work.