

Network Performance Engineering Issues and Strategies

by Steve Waldbusser, Chief Strategist, VitalSoft
November 1996

Today's network managers know that in order to take back control of their network, they must make decisions based on real data about their current and historical network performance. However, that network performance data is often not available and network managers are forced to budget, plan, and operate with less than the full facts at their disposal. Their dilemma is that their organization is already working at full capacity responding to network failures and add/move/change requests. They simply don't have time to perform the repetitive tasks involved in gathering and reporting on network performance. Performance management software offers a way out of this dilemma. Which software solution is right for your organization depends on the technical issues raised by your evolving network.
This paper describes the many ways an organization can benefit from basing its decision-making process on historical performance information and reviews the technical issues that must be considered when choosing a solution.

Benefits of Managing with Network Performance Information

Many people in your organization can benefit day-to-day by having network performance information available to them. In fact, there are probably people outside your organization such as local LAN administrators, financial planners, or even users, who can similarly benefit.

- Planning
- Budget Support
- Confirmation of Network Upgrades
- Monitoring Service Level Agreements
- Communications with Internal Clients
- Quality Control
- Baselining Your Network

Planning

Only with real facts on past and current network utilization can one effectively plan for the future. This information can give valuable insights as to where the coming priorities will be in planning for expenditures, short- and long-term staff objectives, and technology upgrades. These are all areas that require significant lead-time in order for the organization to work in a planned, coordinated manner.
In order to use data effectively for planning, at least 3 months of historical data must be accumulated so that long-term trends can be tracked. The best way to follow these trends is to plot the average and peak utilization values for each day's prime time (i.e., 9AM-5PM) during the time period. The next step is to draw best-fit lines for each data series that follow the trends. The more historical data is available, the more evident the trends will be. These best-fit lines can then be plotted out into the future, predicting what the average and peak values will be. When plotted out into the future, these trend lines will also show when a network component is going to reach critical thresholds that cause performance to degrade.
The first way to use this information is to determine when the peak utilization for a resource will reach 80% and then to schedule an upgrade or optimization to begin 5 months before that time. It is necessary

© 2001 Lucent Technologies



to plan for a five-month lead time to account for delays in receiving equipment or provisioning services, delays in scheduling downtime, and time for adequate design, implementation, and testing. This lead time will allow these processes to occur in a coordinated manner, rather than as a crisis. Further, when preparing plans and budgets, any resources whose lead time begins during the planning period should have an upgrade or optimization scheduled during that period (e.g., the next fiscal year). Also, note that the 80% threshold is appropriate for all network resources, WAN, LAN, and router CPU, except for Ethernet, where a 40% utilization threshold should be used.

Finally, while it is useful to track errors historically, they should not be used for trending. The first reason for this is that errors usually do not follow trends. The second is that while utilization can be viewed as a leading indicator of network performance, errors are not. In other words, utilization can be tracked to predict beforehand when users will experience degraded performance, but if errors exist, users are already experiencing degraded performance. So monitoring errors may be useful for quality control, but it is more reactive than proactive.

Budget Support

Network performance information plays a valuable role in justifying budget requests. It is clear that budget requests that are backed up with real data in an easy to understand graphical format receive more attention and better consideration. It is also clear that networking organizations will be under more and more pressure to justify their financial planning in this way.

When preparing network performance management reports to support budget requests, it is important to understand that the audience of the report may include non-technical people. The report should be designed so that it makes its point without the need for a detailed technical description. One of the most important ways to achieve this is to present resource utilization figures as a percentage of capacity so that it is clear where the trend lies with respect to 100% of maximum capacity.

Another crucial point to remember is that most networks are very busy during part of the day and idle during other parts. Averaging the entire day would result in a report that vastly understates the need for additional funding. When preparing such reports, it is therefore important to only compute daily averages based on the prime-time hours appropriate for your particular business. This would be 9AM-5PM for many businesses, while a brokerage network might concentrate on 9:30AM-4PM and an Internet Service Provider might concentrate on home users connecting between 6PM and midnight.

Confirmation of Network Upgrades

After a budget request has been approved and the work completed, another opportunity develops - to confirm that the predicted benefit has been achieved. From the historical database, "before" and "after" reports show the benefits gained in decreased user response time, decreased error rates, and decreased network utilization. When these values meet or exceed the predicted benefits, this is an important report to share with management and end users so that they can see the benefits of the investment.

If the predicted benefits are not met, the "before" and "after" reports serve as objective data with which to challenge the responsible equipment suppliers and/or service providers. If the contract specifies acceptance criteria, performance management data can be used to prove that these measures have not been met.

Monitoring Service-Level Agreements

Service-level agreements (SLAs) are a fast-growing way of benefiting client and supplier alike. It benefits the client by providing an objective grading criteria and protection from poor service, and it benefits the supplier by providing a way of ensuring that expectations are set correctly and will be judged fairly. A Corporate IT department may find itself in both of these roles, as it may have SLAs with its clients, the end users; and at the same time have an SLA with a frame-relay service provider, for example.

It is important for Corporate IT departments to monitor both types of SLAs. In the case of the SLA with end users, IT must verify that it is in compliance with the service level and have some measure of early warning if there is risk of not meeting the service level so that corrective measures can be taken.

© 2001 Lucent Technologies



Communicating the actual service levels to internal clients can serve to advertise the value of the corporate IT department to the company. This information can also be used to pinpoint where a perceived service problem exists, potentially avoiding unnecessary work - in case the problem lies elsewhere or doesn't exist at all.
In the case of the SLA with the service provider, significant effort may have been expended negotiating mandatory minimum service levels and appropriate penalties for non-compliance. Without some means of monitoring these service levels, this effort has achieved little.

Communications with Internal Clients
With corporate IT organizations stretched to the limit, very little time and energy is available to spend communicating with internal clients. Communication is typically limited to calls to the trouble desk, not the most positive interaction with the client.
The advent of Web technology can change this. The Web can be used to distribute information throughout the organization to clients in multiple sites throughout the world, on any platform. If the web solution is designed correctly, it can make distribution of this information effortless and automatic.
In choosing a network performance monitoring solution with a web interface, it is important to make sure that

- it is scalable to hundreds of users
- it is interactive so that it can be used for real-time decision support
- it is customizable to suit the needs of the organization and its users
- it publishes information automatically so that no effort needs to be expended to keep it up to date

Quality Control
Many organizations have a network management platform running in their operations center. Among other things, this system checks availability of critical network links and components and quickly notifies personnel when failures occur. This does a fine job when network components fail completely but it doesn't provide any visibility when the quality declines due to excessive errors, delay, or congestion. These types of quality problems will affect users nearly as much as network failures because they can seriously degrade application performance. Diagnostic tools will often show error, delay, or congestion parameters but only on the node being diagnosed - it would be an extremely repetitive process to run such a tool on every network device even only once a day.
A network performance monitoring tool can be used to solve this problem. It can monitor utilization and errors on every WAN, Frame Relay, router, LAN, switch, and probe in your network. With the use of thresholds, you can be notified of those resources that are causing slow application performance. Quality control may focus on ensuring that no network resources exceed these thresholds. Because quality control problems could occur anywhere, it is important to monitor all resources on the network.

Baselining Your Network
It is difficult to improve or troubleshoot a network unless its normal operating characteristics are known. It is important to know the normal characteristics for all resources in a network. A history of the utilization and error characteristics of an entire network for at least a month is called a baseline.
The baseline can be used to support troubleshooting activity. When a network is experiencing problems, the current characteristics of the network can be compared with the baseline (normal) behavior for the network. This can aid the technician greatly in determining which parameter is abnormal and therefore the most likely cause of the problem. Because failures could occur anywhere, it is important to monitor all resources on the network.

© 2001 Lucent Technologies



Technical Challenges in Choosing a Network Performance Solution
Network Performance Management is not a new field and many lessons have been learned about what to do and what not to do in order to achieve the highest success. A few of those lessons are discussed here.

- Provide an Enterprise-Wide View
- Scaleability
- Performance
- Reliability

Provide an Enterprise-Wide View
It is important to provide an enterprise-wide view of network performance. There are two components to providing an enterprise-wide view. The first is to monitor the entire network and the second is to integrate that entire view into one reporting structure.
It is important to monitor the entire network because any portion of the network can encounter problems. If part of the network is left unmonitored, that part is likely to develop poor quality. User's perception of the service quality depends on the sum of the parts in the network. Problems may exist on local area networks scattered across the enterprise - these can cause as much difficulty as backbone problems.
A single integrated view of the entire network is necessary not only for convenience, but also to allow the entire network to be rolled up into summary and exception reports. Without a single consistent view, there is no way to sort network components to find the most heavily used; nor is there a way to provide a single exception report that contains all of the network components that have crossed critical thresholds.

Scalability
In order to scale up to medium to large networks, a performance management solution must be scaleable to thousands of network interfaces. This scaleability shouldn't be achieved by increasing SNMP polling intervals because that will eliminate the potential to capture peak values and because long SNMP polling intervals will cause inaccuracies on high-speed networks.
It is important that reports show peak values so that one can understand the effects of network bursts on end-user response time. Performance averages are interesting numbers as well, but they may hide periods when network performance has become intolerable. Peak values are measured over periods - an example might be that the 5-minute peak utilization between 1PM and 2PM was 67%. This is measured by collecting 12 separate 5-minute samples between 1PM and 2PM and at the end of the period choosing the highest value. The polling period and the peak values will both be affected by the peak period chosen. In order to provide 5-minute peaks, the network polling rate must be at least 1 poll per device every 5 minutes. As a rule, the peak values will be higher for shorter peak periods, because shorter but higher bursts of network load are captured. The desired polling period can vary for different networks or different applications but a period of 5 minutes is ideal and a period of longer than 5 minutes is never recommended.
SNMP counter values are 32 bits long and therefore hold a maximum value of roughly 4 billion. If the counter ever exceeds the maximum value before it is polled, it will wrap, and highly inaccurate data will be collected. These values wrap more quickly on faster network technologies and the shortest period in which they can wrap can be predicted based on the speed of the network. On Ethernet, these counters take longer than an hour to wrap, but on high-speed networks of 100Mbps or greater, SNMP counter values will wrap in less than 10 minutes. It is therefore important to have a polling period of 5 minutes or faster.
While the SNMP polling period should not exceed 5 minutes, it can be difficult to achieve this for networks with large numbers of interfaces to poll. Traditionally, SNMP polling engines have been non-multithreaded, which means that they exploit no parallelism when they poll the network and are subject to

© 2001 Lucent Technologies



a slow-down of the polling cycle with every network delay. A partially multithreaded polling engine allows limited parallelism and allows some requests to continue even when others are slowed down by a slow WAN link or a down device. Unfortunately, it is easy for a large network to overwhelm a partially multithreaded polling engine. On the other hand, a fully multithreaded polling engine allows full parallelism and will continue to poll no matter how many slow network links or devices are encountered. This fully multithreaded design allows extremely large networks to be polled from a single management station.

Performance

The performance of the database and report writers is another technical consideration. If the database and report writers take a long time to function, it can rule out any interactive use of the system to perform what-if analysis or to answer questions about how the network is currently performing. Often databases become even slower as more and more data is put into them, exacerbating the problem. The lack of interactive use can severely hamper the usability of the system and can result in a mainframe-oriented model where users make report customization requests to an administrator who spends a lot of repetitive time designing reports and waiting for them to be created. The end result is that fewer people actually use the reports to make decisions.

Reliability

The reliability of the network performance management system can have a major effect on how expensive it is to maintain and how much value it provides to the organization. An unreliable system will consume a lot of technically-skilled labor configuring and tuning the system. Traditionally, network performance management tools for medium to large networks require one or two people to configure, maintain, and tune the system. A lot of this repetitive work is due to changes in the network as well as changes in the reporting requirements of users. Further, if the system is not reliable, users will tend to not use the system and instead may simply make decisions without network data to support them. This is particularly true if the unreliability causes months of data to be lost. It is therefore important to make regular backups of the system so that when a failure occurs, there is an opportunity to restore the database and not lose months of trending information.

Conclusion

In order to take back control of your network, you must make decisions based on real data about the network's current and historical network performance. Because collecting that information has traditionally been so difficult, the information is often not available and network managers are forced to budget, plan, and operate with less than the full facts at their disposal. There are many benefits to be achieved with this information, and once the technical challenges are understood and solved, a convenient, workable solution can be achieved.

© 2001 Lucent Technologies