IBM Global Services

Network performance and capacity planning: Techniques for an e-business world

e-business is about transforming key business processes with Internet technologies. In an e-business world, networks are the heart of your business.

Rapid response times. High availability. Adequate bandwidth. These are the demands placed on your business networks. Fueled by an explosion of e-business and remote access requirements, the performance of your network today directly affects your customer's first impression of you, and therefore your business success, more than ever before. Effective network baseline analysis gives you the ability to improve and maintain performance of your existing networks as well as make informed design and purchase decisions as you grow.

Baseline analysis requirements

Effective network capacity planning doesn't always translate to, "Buy more bandwidth." Network requirements should be based on analytical insights and business goals. Network baseline analysis provides essential information using industry-standard SNMP MIB data collection techniques and proven analysis methods. Through data gathering and analysis of network performance trends, an enterprise view of LAN/WAN availability, performance and capacity can be obtained. You can then effectively answer the following performance and capacity questions:

- What points in the network are directly impacting availability and performance? Where is the congestion?
- Are there times of the day when the business is vulnerable due to high use?
- Is existing capacity being used effectively? Are current resources allocated to maximize performance and minimize costs and problems?
- Is the correct mix of technology employed to meet business demands?
- Are network applications and protocols utilizing the network bandwidth effectively and efficiently?

e-business

Getting the full picture

To answer these questions requires an in-depth look at the network from a number of perspectives. Let's look at the key elements that need to be analyzed:

- **Traffic Characteristics**
  - Traffic volumes and rates
  - Prime versus non-prime traffic rates
  - Traffic volumes by technology
- **Load Distribution/Load Balancing**
  - Device throughput
- **Operational Capacity**
  - WAN percent capacity used
  - LAN percent capacity used
- **Resource Interface Utilization**
  - Top utilized interfaces
  - LAN interface (router stats) and segment (RMON stats)
  - WAN interface utilization
- **Evidence of Congestion**
  - Packet discards
  - Top error interfaces
- **Traffic Overhead**

Network traffic

Network traffic characteristics give you a view of network usage, helping you begin to develop a comprehensive picture.

If you don't know where the traffic is coming from on your network, then how do you accurately plan for it? Like cars on a highway, you need to know how many are on the road, what type they are, where they are coming from, and where they are going. What happens at "rush hour", when your traffic is the highest? If a road or route becomes unusable, can data still reach its destination?

Traffic analysis demands a comprehensive look at your network's behavior. To anticipate potential problems and prevent them, establish a networkwide view of how much data is being transferred across the network. Determine traffic patterns by various business periods for a given day or for the entire

collection period. Information about traffic characteristics includes views of significant network devices, reports on traffic volumes and rates that change per unit of time for major devices. Differences in network traffic between different network technologies can be valuable in determining network usage practices and guide network planning activities when the information is used effectively.

As an example, here are a few sample traffic characteristics taken from a distributed enterprise network analysis done over a period of 6 days. Chart 1 - Network Traffic, is an example of a network profile.

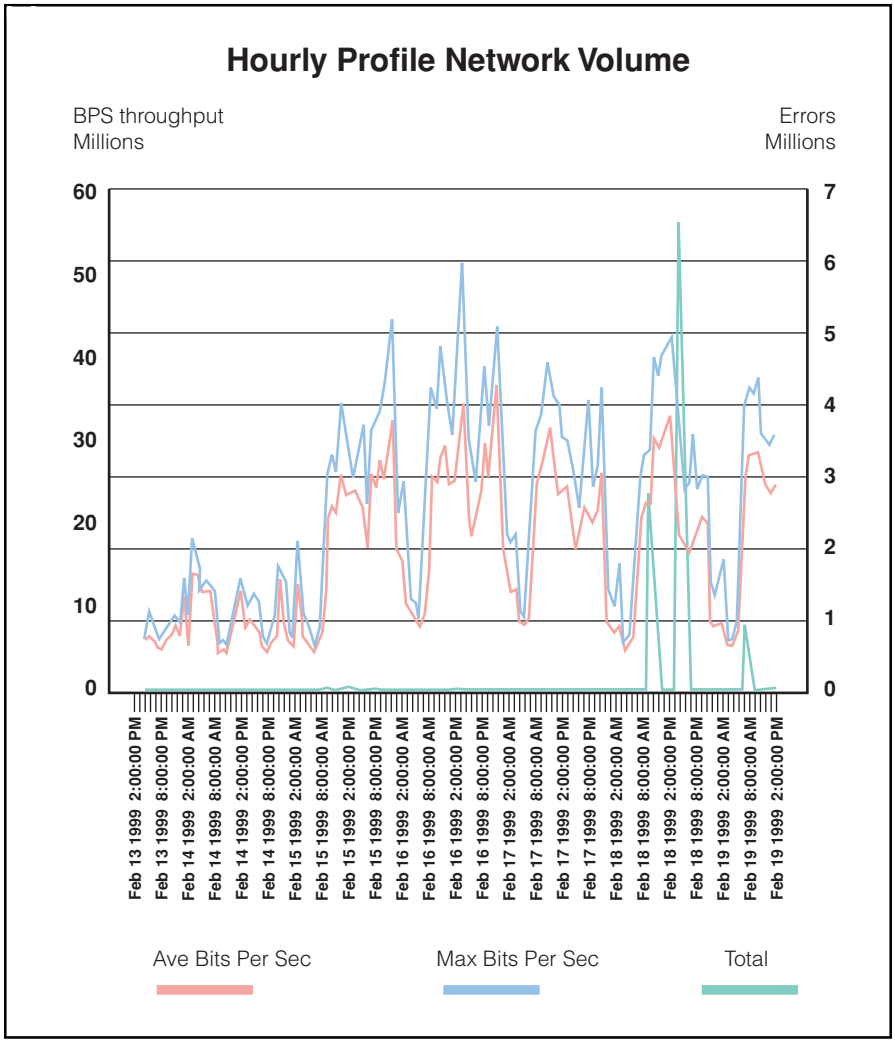

Chart 1: Hourly Profile Network Volume

- 66 percent of the traffic seen during the collection period was Ethernet, followed by 28 percent WAN and 5 percent ISDN.

- The hourly network traffic profile identifies two distinct periods of network activity: 8 a.m. - 5 p.m. Monday to Friday during the prime shift and a daily 5 - 12 p.m. batch window.

- The hourly network traffic profile identifies peaks of congestion activity between 8 a.m. and 5:00 p.m. Monday to Friday, and peak collision activity on the RMON-monitored Ethernet segments during the 5 - 12 p.m. batch window.

Network load balancing and distribution

Identifying areas to implement or modifying load balancing for key network components can improve availability and performance.

To keep traffic flowing on your network highway, you will need to constantly evaluate the load on your significant network components, especially the routers providing Internet and intranet connections for your internal and external customers. This will help you identify areas of vulnerability as well as opportunities to better utilize existing resources. Using this information will help you more effectively implement new and redundant paths throughout your network to lower the risk of a single point of failure. Performing this analysis will also uncover when components are either not operating, or not configured as you had expected or planned during your design. Our experience is that actual component utilization is often quite different from expectations. Without this information, you may incorrectly focus your problem determination activities on symptoms, rather than causes, of network performance concerns.

Examples of network loading statistics from our enterprise analysis include:

- Of the 31 devices monitored, 4 devices contribute 60 percent of the total network volume (Bytes) during the collection period.

- DWN-Chan1 router contributes over 40 percent of the overall daily network average of 92 Mbps, highlighting this resource as critical to the core network infrastructure and operation of the network.

Some network device vendors provide the capability to monitor the device CPU, memory and buffer resources. If available, this information may also be looked at during the baseline analysis process. Incorrect or inadequate resourcing of network devices can be a significant contributor to network errors and device bottlenecks which constrain network traffic.

e-business networks require an increased focus on WAN capacity and performance to support critical business applications.

Operational capacity

On your highways, how many lanes do you have available for traffic, what are the speed limits, and how full are they? An understanding of your true network speed and capacity can uncover a number of potential problem areas and opportunities to utilize or eliminate unused resources. Historically, the bulk of network traffic was local with workstations accessing local file or application servers. WAN traffic has traditionally been for communications between company sites. The network of today is WAN intensive, a result of the higher Internet/intranet traffic capacity requirements needed to support e-business ventures and initiatives.

Adding WAN capacity represents a real cost because of the relatively high expense of WANs compared to LANs. New WAN usage should be evaluated in terms of business application requirements and how they contribute to the business. WAN capacity that has already been added to meet anticipated needs should be reevaluated periodically. This ensures that available capacity is being used to the planned-for degree. Chart 2 shows how capacity for a WAN connection tracks over a six-day period.

Operational WAN Capacity BPS		Average percent In Capacity	Potential Max Percent In Capacity	Average percent Out Capacity	Potential Max Percent In Capacity
13 Feb.	46731000	2.4		3.58	8.41
14 Feb.	46731000	2.2	7.19	3.17	8.19
15 Feb.	46731000	6.33	29.8	7.5	31.05
16 Feb.	46731000	6.38	27.3	7.62	37.4
17 Feb.	46731000	6.69	27.12	7.61	31.16
18 Feb.	46731000	6.1	25.06	7.7	31.69

Chart 2: Average WAN Capacity

Effective network baseline analysis provides an accurate picture of the relationship between available versus utilized capacity for the network. This evaluation includes links to network vendors where the consumption of available capacity should be evaluated 24 hours a day for 7 days to evaluate trends. Significant cost savings can be uncovered in purchased capacity that is either unused or may be reallocated to a link with greater business requirements.

Analysis of network capacity should include a percent of network capacity being used, as well as a breakdown of the capacity being used on both the LAN and WAN. In addition, it is useful to look at WAN capacity use across diverse WAN technologies, such as frame relay or ISDN, where the unique characteristics of the technology circuit should be reviewed by an experienced consultant.

Our enterprise network analysis example shows the following operational capacity:

- WAN operational capacity is 46.7 Mbps for the monitored interfaces. Maximum average inbound utilization is 6.69 percent, with a potential maximum inbound utilization of 29.8 percent on Feb. 15. Maximum average outbound utilization is 7.70 percent, with a potential maximum outbound utilization of 37.4 percent on Feb. 16.

- Ethernet operational capacity is 5,050 Mbps for the monitored interfaces. Maximum average utilization is .28 percent, with a potential maximum utilization of 1.23 percent on Feb. 16.

- Of the monitored devices, router resource DWN-Chan7, with operation capacity of 20,256 Mbps, had the highest average capacity. Capacity used was at 13.7 percent on Feb. 16, with a maximum capacity of 72.5 percent on Feb. 16.

- Overall, operational network capacity is plentiful.

Network interface utilization

Utilization analysis can give you the performance improvements you need to meet business demands without increasing your infrastructure expense.

How busy are your highway interchanges, points where traffic leaves or enters a network segment? Network utilization for any segment or link in the network is determined by calculating the relationship between the volume of traffic and the speed of the segment link. Baseline network analysis requires characterizing the utilization of network device interfaces. Knowing which interfaces have the highest utilization, when high utilization occurs, and if the utilization pattern regularly exceed some identified threshold may identify a need for more (or less) capacity. Highly utilized interfaces can also represent potential bottlenecks. Underutilized network segments may represent potential cost savings. Balancing traffic between over- and underutilized segments can dramatically improve performance for your customers.

Network utilization analysis should include information identifying the most highly used segments in the enterprise network. It is also important to characterize the direction of the incoming versus outgoing traffic. This identifies which traffic is affecting performance and helps to identify the source so it can be addressed. Analysis should include the average utilization, top interfaces by maximum utilization and time periods when send/receive interfaces exceeded a threshold. Frame-relay circuit information must be analyzed to identify top frame-relay interfaces by average percent sent/received utilization and frame-relay circuits with average utilization over committed information rate (CIR).

Our enterprise resource interface utilization analysis shows:

- 10 WAN circuits report 60 percent or greater receive or sent utilization at minimum 10 percent of time during prime-shift hours.
- 4 ISDN circuits report 60 percent or greater receive or sent utilization at minimum 10 percent of time during prime-shift hours.
- 4 10Mbps Ethernet interfaces report peak 15-minute intervals of greater than 30 percent utilization during prime-shift hours.
- 4 10Mbps Ethernet segments report peak 15-minute intervals of greater than 30 percent utilization during prime-shift hours.

Network congestion

Analysis of congestion conditions allows you to improve response times and availability.

Is the data on your network highway being delayed or not even reaching its expected destination on the first trip? When data packets generated by an application do not reach their destination due to congestion, this information must be retransmitted. In addition, network error traffic may be generated to indicate the source of the delivery failure. The additional traffic associated with resending the information as well as the error traffic can worsen network congestion and may cause your customers to experience a delay or outage. If network congestion is a rare occurrence, it is insignificant. If it's a common occurrence, there may be a significant impact on business operations.

Analysis of network congestion includes error packet activity reports for the network as a whole and critical network devices specifically. The

top 25 with high error activity should be identified and examined in more detail to identify the cause. Any interface with errors greater than 10 packets per second should also be analyzed in greater detail. Reports on discarded packets for the network as a whole as well as critical devices can also reveal problem areas. Any discard is an undesirable event, but it is a normal adaptive mechanism during peak network traffic events. Occasional discard events are insignificant, but they must be analyzed in terms of network traffic characteristics in general. This is where the consultant experienced in many varieties of network implementations depends on professional skill to make the correct judgment. Chart 3 a high-level summary of traffic activity that shows which network technology is experiencing high rates of errors and discards.

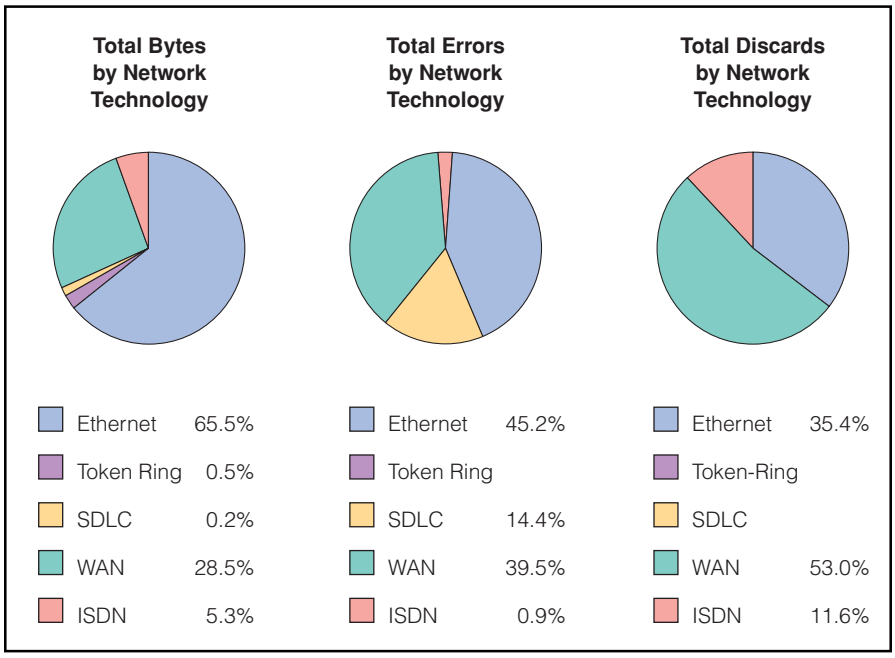

Chart 3: Network Technology Rates of Errors and Discards

The following is an example of an analysis showing congestion and error traffic activity for an enterprise.

- Three interfaces contribute over 70 percent of the total network discards. One of these interfaces contributes 56 percent of the 70 percent.

- Ethernet interface Ethernet3 on router resource CAS-Hosp1 reports congestion activity during prime-shift hours. The hourly network profile suggests that the Ethernet segment connected to this router interface is congested. This interface contributes the majority of discards.

- WAN circuits Serial5/4 and Serial4/1 on router resource DWN-Chan1 show congestion activity during prime-shift peaks in outbound utilization. The discards, as a percent of total traffic during the congestion periods, is less than 1 percent of the total traffic.

- 94 percent of network errors are reported by three router resources; DWN-Chan1, DWN-Chan11, and DWN-Chan12. DWN-Chan1 contributes 76 percent of the 94 percent of total network errors.

- 68 percent of the total errors are reported by WAN circuit Serial4/6 on router resource DWN-Chan1. The errors were reported during two distinct time periods: 9 - 10 a.m. and at 5 p.m. on Feb 18. This link reports light utilization throughout the collection period.

- WAN circuit Serial3 on router resource DWN-Chan11 consistently reports error activity throughout the collection period. Error activity as a percent of total packets is less than 1 percent of total.

- Overall, network error activity is minimal and indicates little if any detrimental impact to performance.

Worth noting is the fact that congestion may occur without high network traffic loads. Inadequately provisioned network devices or network devices that have been left at default values may create network congestion in an otherwise well-provisioned environment. Network devices may not be tuned to handle network traffic requirements. Monitoring network device statistics and analyzing trends may uncover issues in the network device buffer allocation. Device buffer pools may be inadequate or the device may be provisioned with inadequate memory resources. These resources may have

been appropriate when the device was installed, but are now inadequate because they were not reviewed through periodic network baselines analysis to keep them current. Reports on buffer allocation failures or buffer misses that include 24-hour periods over at least 7 days are required to effectively analyze this type of problem.

Controlling network and server management reporting traffic can increase available bandwidth

Ethernet media is also subject to collision errors on shared segments. While the switched environment has offered considerable relief from this type of error on segments dedicated to a single workstation, there continue to be shared Ethernet segments on some enterprise backbones.

Traffic overhead

Managing the traffic crossing your network highway for the purpose of delivering management information is important for effective capacity planning. Broadcast traffic often competes with business application traffic and consumes network bandwidth. But broadcast traffic may be useful and may significantly add to the apparent speed of operations by the LAN user. Broadcast traffic over a WAN link should be analyzed and filtered to manage this more costly and limited bandwidth.

Another part of overhead traffic is management traffic. SNMP generally adds less than two percent of traffic, but there is potential for it to grow uncontrollably if different groups within an enterprise implement management stations which poll for this information. Controlling the number of polling management stations is part of a larger management strategy and should be part of the overall network capacity and utilization plan.

Reports showing broadcast/multicast traffic at each interface allow the source of the traffic to be localized at the most appropriate resource. Monitoring for 24 hours a day for 7 days identifies patterns and cyclic broadcast activity that would not be detectable with a standard data-capturing tool. For accurate assessments, SNMP traffic should be reported as a percent of total traffic.

Summary

Your enterprise's ability to be competitive and run cost effectively depends on efficient networks that have the right capacity to run business applications. Collecting and analyzing SNMP MIB data can provide a quantitative analysis of network performance and capacity parameters in a router- or switch-based network. It can be done with equipment from any of the major network vendors. Effective analysis requires appropriate tools and effective analysis methods.

Most important, however, is experience in SNMP MIB analysis. SNMP data analysis is not a simple activity because of the flexibility in the definition of the standards and the desire of network vendors to add value. Analysis of MIB data for event-type activity differs from the analysis done when trending large volumes of data to understand the impact on network performance and capacity.

Investing in the right skills and performing periodic network baseline analysis will allow you to maximize the performance, capacity, and availability of your enterprise network.

For more information

For more information on IBM Performance Management and Capacity Planning Services, call 1-800-426-4682 (in the US), 919-301-4141 (from outside the US) or e-mail us at capacity@us.ibm.com.

William Nametka, the author of this paper, is a Senior Project Manager for IBM Performance Management and Capacity Planning Services. His areas of responsibility include network process, LAN design, technology migration and MIB analysis.

IBM Global Services
3200 Windy Hill Road
Mail Drop WG 15C/B27
Atlanta, GA 30339
U.S.A.

Produced in the United States of America
7-99

G563-0239-00