



What to Monitor in IP on z/OS

By Tony Amies

Previously, we explored techniques for collecting IP monitoring data. In this article we examine what IP data to monitor in the system.

PREVIOUSLY in “Choosing an IP Monitor for z/OS” (*Technical Support* magazine, April 2003), we looked at the techniques available for collecting IP monitoring data, namely NetStat, SNMP, SMF exits and the packet trace interface. In this article, we will look at what should be monitored in the system with reference to the various techniques used.

WHY MONITOR IP PERFORMANCE?

A common myth is that there is often sufficient bandwidth in the network to negate the need for monitoring IP performance. Although this may reduce the need for some types of monitoring, other types of performance monitoring are not related to network bandwidth and are still required to ensure that the IP network and its related services are available and perform within acceptable limits.

Monitoring IP performance is critical to maintaining service levels. Monitoring IP performance in real-time can confirm that the network is operating correctly and highlight small variations which, if left unchecked, could lead to more serious problems at a later time. Without real-time performance monitoring, some of these variations can be missed and the first sign of the problem may be when an outage occurs. Longer-term analysis of performance data is equally important to look for changes in traffic volumes or connection

rates. These can be used for capacity planning purposes to ensure that sufficient bandwidth and system resources are readily available to support the network services.

WHAT SHOULD BE MONITORED?

There are two distinct categories of IP performance monitoring: *Service Delivery* and *System Resources*.

Service Delivery monitoring typically measures end-user response times, network transit times or the ability of the network to support a specific number of end-user connections. In many cases, service delivery is the key indicator for acceptable performance as this is what the end-users of the network are experiencing.

System Resource monitoring ensures that the various components on which the service is dependent are operating within acceptable limits. This can be measured in terms of utilization and availability of resources.

Although service delivery monitoring is, perhaps, the more visible of the two approaches, in practice both types of monitoring need to be implemented. Monitoring response times and throughput highlights the health of the network but does not aid in identifying or resolving problems detected. System resource monitoring is equally important to help diagnose the cause of any service delivery problems and, in many

cases, can be used to identify problems before end-user services are impacted.

RESPONSE TIME MONITORING

Historically, response time measurement has been the traditional method for monitoring performance. In IP networks, this still applies, but only for specific network services such as Telnet. Unlike SNA, IP is a stream-based, peer-to-peer protocol without the concept of a block of data forming a request and a block of data forming a response. Therefore, it tends to be the IP services that emulate former SNA 3270 services that qualify for response time measurement. For services where this is not the case, network transit time monitoring is more applicable where the time for the application data to be delivered to the partner system is measured rather than a round trip response time.

Most IP monitoring products on the market claim to measure network response time. However, many of these products use the PING command to obtain this measurement. PING offers some insight into the network performance but does not necessarily measure what the end-user is experiencing. PING sends out a block of data using an ICMP echo request. The response time is measured as the time it takes for the partner system to return the same block of data. This is not particularly reliable since:

- ▼ ICMP packets can be handled at different priorities by intermediate routers

- ▼ ICMP packets are, in many cases, blocked by firewalls
- ▼ Real end-users rarely use ICMP protocols and do not generally send and receive data of equal length

A more accurate form of TN3270 response time monitoring is documented in RFC 2562. The techniques described in the RFC can be adopted by IP monitoring tools using TCP level acknowledgements and the associated sequence numbers, rather than SNA level acknowledgements. This allows the methodology described in the RFC to be used to provide response time measurement for the TN3270E and TN3270 protocols. See FIGURE 1.

In FIGURE 1, the TN3270 response time is calculated by comparing the timestamps. The difference between timestamps one and two is the application response time. The difference between timestamps three and four is the network delay. The difference between timestamps one and four is the total response time. In this scenario, it is important for the monitoring tool to determine when all data sent to the TN3270 client has been acknowledged before taking timestamp four. This is achieved by comparing the TCP sequence numbers on the incoming acknowledgement packets. When the sequence numbers indicate that the total bytes sent have been acknowledged, the monitor can deem the transaction to be complete. In order for a monitoring tool to implement this form of TN3270 response time monitoring, the tool must either be operating inside the TN3270 server or have access to the packet data flow.

As with response time methods in SNA, the time taken for the transmission of the acknowledgement packets is substituted for that of the initial end user request. There can be a delay in the remote stack when sending the acknowledgement packet, but, as it is only a few milliseconds, it can be ignored as it is orders of magnitude smaller than time taken for the packet to traverse the network.

For other TCP services that do not have a simple request/response structure like TELNET, a subset of the RFC 2562 technique can be used to measure network transit times or throughput for the service. By using the TCP sequence numbers, time stamping and measuring data length, a monitor can simply measure the time it takes for each outbound packet to be acknowledged.

FIGURE 2 shows a typical full-duplex data flow for which standard response time measurement would not apply.

In this example the remote TCP/IP application is sending data at the same time as the local TCP/IP application on the same connection. The concept of response time does not exist because the data transmitted by the local application is not necessarily sent as a result of the incoming data. It is, therefore, more practical to look at the network transit time for the application data being sent. At point one, the monitoring tool takes a timestamp and records the data length being sent. At point two, additional data being sent is counted. The network transit time for the data can be calculated when the TCP acknowl-

edgement is returned with the TCP sequence numbers indicating all bytes have been acknowledged. Depending on the data flow, this may be indicated by the sequence numbers in multiple acknowledgement packets, or another inbound data packet.

This technique for network transit time monitoring makes similar assumptions to TN3270 monitoring and it, too, will be accurate to a few milliseconds. The ultimate aim is to provide consistent and acceptable response times to end-users and to highlight when variations occur. Calculating response and transit times using these techniques

IP and SNA Network Management tools for z/OS and OS/390

WDS offers you a wide selection of Network Management tools specifically designed for z/OS and OS/390.

implex

- The only real-time IP monitor that can tell you what is NOT working as well as what IS working in your IP stack.
- VIPA Activity
- TN3270 Response time
- NetView interface
- Packet Fragmentation monitor
- Connection Failure and Route Change indicator plus much more.

routeview

- Automated VTAM subarea path table design and generation.
- Import your existing path tables.
- Automatic health check
- Point and click to design your new network.
- Automatically generate your new path tables.

exigence

- Comprehensive diagnostics for IP and SNA Without the use of IPCS, GTF or ACF/TAP.
- All major protocols supported including APPN, OSPF, LPD, GDS variables, X25, Enterprise Extender and many more.

Download a free trial copy of any of these products from our website, or call us and we will be happy to ship it to you on tape or CD, free of charge.

ftpalert

- Give FTP the security and management it really needs.
- Integrate FTP with your automated operator.
- Automate FTP failure recovery procedures.
- Make FTP a separate RACF resource.
- Full audit trail of FTP activity.



*William Data Systems, LLC
99 Canal Center Plaza, Suite G-10
Alexandria, VA 22314
Toll free 877-723-0008,
Tel 703-299-0008,
Fax 703-299-9776*

For network tools that really make a difference.

www.willdata.com

enables inconsistencies or gradual degradation to be easily detected and acted upon.

MONITORING RESOURCES

Monitoring the performance of key resources can help with the early detection of performance problems that could cause degraded services or even system outages. For performance monitoring, the key metrics are the amount of work being done by a resource in relation to the total it can practically handle.

For a z/OS TCP/IP implementation, an overall health check of the work being done by the stack can be a useful spot check on the performance of the system. The health check would typically indicate the number of active connections, together with the current packet and byte rates. For more detailed monitoring, similar data is required for individual resources such as interfaces, services (applications) and gateways. Ultimately, for any resource being monitored, the same data needs to be collected:

- ▼ Current number of active connections (for connection-based protocols)
- ▼ Current packet and byte rates (packets and bytes per second)
- ▼ Error conditions

Obtaining the number of active connections for connection-oriented services can be achieved in several ways, such as SNMP, NetStat, SMF Exits or the packet trace interface, with the key consideration being the system overhead associated with collecting the monitoring data.

Monitoring packet and byte rates becomes more complex because in order to determine the rate, the monitoring tool needs to be able to calculate the number of bytes and packets over a specific period of time. All the main data collection techniques can be used to collect byte and packet counts, but the overhead consideration becomes more important when there are a high number of resources to monitor. Polling techniques such as NetStat and SNMP may need to be throttled down to reduce the polling interval for collecting the data. This reduces the system overhead, but dilutes the real-time value of the data collected. For an in-depth discussion of data collection techniques see “Choosing an IP Monitor for z/OS” (*Technical Support* magazine, April 2003).

Although performance monitoring of a resource typically means looking at high-water marks, looking at low-water marks can be of equal importance. For example, monitoring excessive values for packets per second on an OSA adaptor or a high number of concurrent connections to a Telnet server is a good way of detecting potential performance degradation. However, an exceptionally low value is an indicator of existing poor performance, not necessarily of the resource itself, but perhaps of another component in the network.

Detecting error conditions for specific resources or connections requires the analysis of several different data types and sources. Monitoring ICMP activity can be used to detect possible performance problems in real-time. This typically requires the monitor to see the ICMP packet data looking for specific ICMP message types, specifically:

Source Quench: Packets are being transmitted too fast for the receiver to handle.

Router Redirect: Packets are being returned by a router for re-direction.

Fragmentation Required: Packets are too large for an intermediate router to handle.

FIGURE 1: A TYPICAL TN3270 TRANSACTION

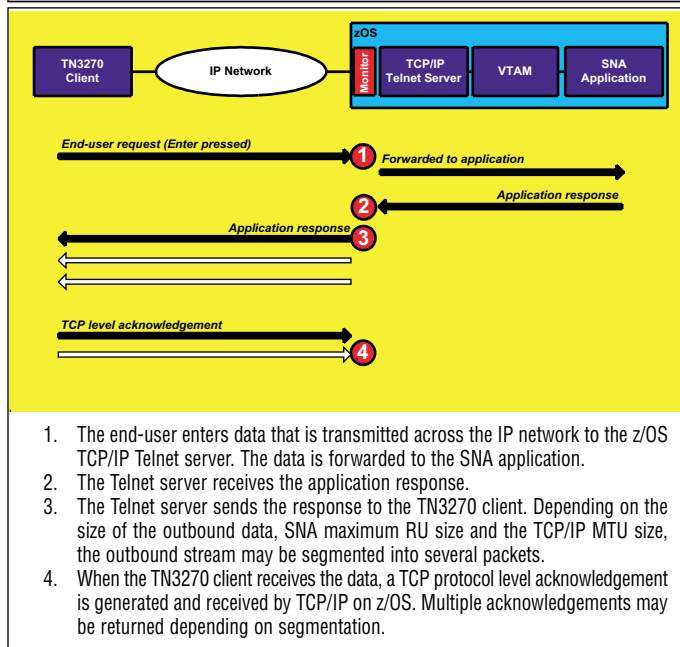
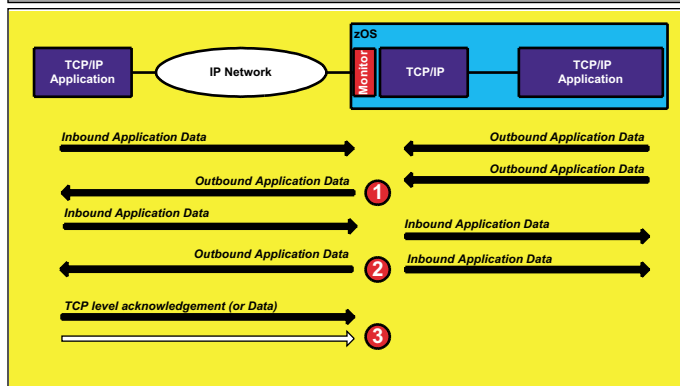


FIGURE 2: A TYPICAL FULL-DUPLEX DATA FLOW



Fragmentation occurs when the sender of the data sends a packet larger than one or more of the intermediate routers in the network can handle. A ‘fragmentation required’ ICMP packet is returned and the data is fragmented into multiple smaller packets. Fragmentation can cause performance problems and, although the ICMP ‘fragmentation required’ packet might be seen by z/OS, detecting fragmentation of inbound packets in real-time can only be detected by looking at the packet headers.

Retransmissions are another indicator of potential performance problems and can only practically be detected in real-time by looking at the sequence numbers in the packet headers. Monitoring tools capable of doing this can check for a sequence number smaller than the previous packet’s sequence number. If this event occurs, retransmission has taken place and the difference between the current sequence number and the previous is the number of bytes being retransmitted. As the underlying IP protocol is connectionless and with no guaranteed delivery, some retransmissions over a multiple hop connection may be acceptable. However, a large number of retransmitted packets or bytes will lead to throughput degradation for the connection(s) experiencing the problems and lead to increased

network traffic, perhaps fuelling the root cause of the retransmission.

SYSTEM RESOURCES

Monitoring some system resources is another key part of overall IP performance monitoring, specifically TCP/IP CPU utilization, storage use of CSM and ECSA and some specific Unix System Services resources without which TCP/IP will not operate efficiently and, in some cases, not at all.

Several tools and techniques are available to measure CPU utilization of the TCP/IP address space and any other address spaces that may impact IP performance. CSM and ECSA storage utilization can be obtained from VTAM commands via the VTAM SPO or via analysis of system control blocks. Without sufficient CSM or ECSA storage, TCP/IP will not function and performance degradation will quickly be followed by outage of services.

Unix System Services provides an integral part of TCP/IP services on z/OS and shortage of USS resources can lead to poor TCP/IP performance and system outage. USS data is available from callable services (Assembler and C), and these can be used by performance monitoring tools to ensure that there are sufficient resources available to provide the services.

There are finite limits in many USS components and although not an exhaustive list, the following metrics need to be carefully monitored:

- ▼ Active processes
- ▼ Active userids
- ▼ Shared memory

A common problem is when an application that is experiencing problems leaves active child processes in a *zombie* state rather than terminating the process. If this is left unchecked, the active processes limit can be reached and many TCP/IP services will be disabled.

AUTOMATED PERFORMANCE MONITORING

Many of the techniques discussed, especially response time measurement and real-time detection of fragmentation and retransmissions, require access to the detailed packet data flowing through IP. Much of the required information is not available through standard facilities such as NetStat or SNMP. Viewing trace data or even parsing collected trace data with batch utilities may be acceptable for diagnosing specific problems, but is not practical for monitoring the performance of even small networks. The optimum method for real-time performance monitoring is to analyze the packet data in real-time, which enables many performance indicators to be quickly and efficiently detected, such as:

- ▼ Packet rates
- ▼ Byte rates
- ▼ Fragmentation
- ▼ Retransmissions
- ▼ Network transit time
- ▼ End-user response times (TN3270)

Tools that monitor IP performance without real-time access to the packet data must compromise by frequently polling SNMP MIBs or constantly using NetStat commands. These techniques allow packet and bytes rates to be calculated over the polling period, but do not detect retransmissions, fragmentation or enable end-user response times to be determined.

This article has concentrated on those metrics that will give a good overall picture of your system and the level of service being given to your users. The next question to ask is: "What levels should we set the alert systems at?"

It is almost impossible to suggest levels without being site-specific. However, devices such as adaptors will rarely be able to achieve the maximum capacity and still function, so setting their threshold at, say, 75Mb for a 100Mb adapter might be prudent. Response

time thresholds should initially be set at what is considered reasonable for the user, based upon accepted service levels within the organization. Any system should be monitored for a period of time to establish what the normal levels of service delivery and resource usage are prior to the setting of alert thresholds.

CONCLUSION

In summary, now that IP is becoming the standard network connection for a mainframe, not to use an IP monitor would be to ignore the impact of IP on your network. The IP metrics that you decide are important to your installation will often be a major influence in which IP monitor you decide to use. However, if you already have an IP monitor, that will decide which IP metrics are available for you to choose from. The careful and considered use of an IP monitor means that you will see problems before the users. Just how long before depends on how well you monitor your system. 🚀

Questions or comments? Please email editor@naspa.com.



Tony Amies is a product architect with William Data Systems. He has been working with IP on mainframes for over 10 years, specializing in TCP/IP management, monitoring and integration. Tony has worked in IT since 1978. He can be reached at t.amies@willdata.com

*©2003 Technical Enterprises, Inc. Reprinted with permission from **Technical Support** magazine. For subscription information go to www.naspa.com, email mbrship@naspa.com or call 414-768-8000, Ext. 115.*

William Data Systems, LLC
99 Canal Center Plaza, Suite G-10
Alexandria, VA 22314
(703) 299-0008
[Toll Free] (877) 299-0008
www.willdata.com
info@willdata.com