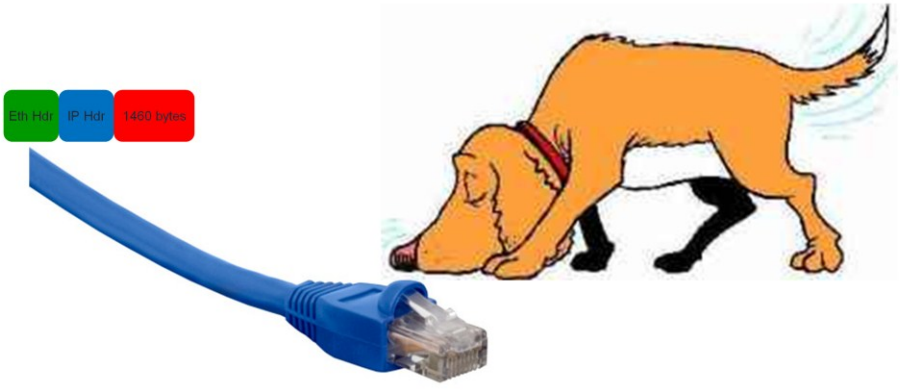
Jason Daniel
jdaniel@us.ibm.com

Nathan Flowers III
nflowers@us.ibm.com

Network Engineers

February 2013

Network Sniffer



2

There are many tools that provide the capability to capture and analyze network traffic. There are many names that are used as well such as network sniffer, trace tool and protocol analyzer.

Network Traffic Analysis

…110110110110 11…

3

Capturing a trace of network traffic is only the beginning. Making an accurate analysis of the captured data is the primary purpose in capturing network traffic.

Wireshark

http://www.wireshark.org/

4

Wireshark is an open source tools that provide both the ability to capture network traffic as well as analyze the captured traffic.

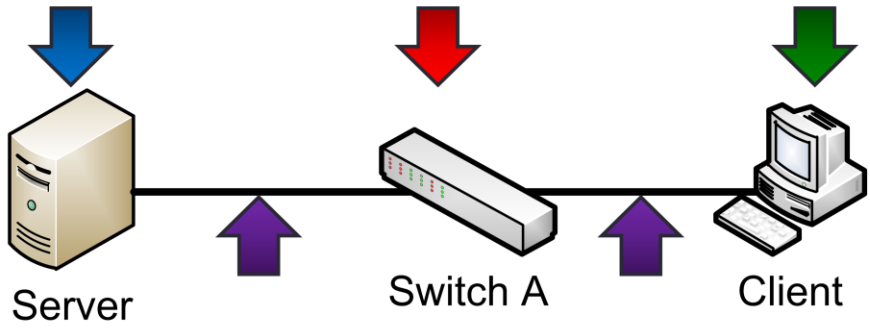
Common Uses

- Network Troubleshooting
- Development Tool
- Traffic Analysis
  - Intrusion Detection
- Protocol Analysis
- Performance Tuning

5

Tools such as Wireshark can be helpful in many areas. One area of concern to many system administrators is network performance tuning of servers and clients.

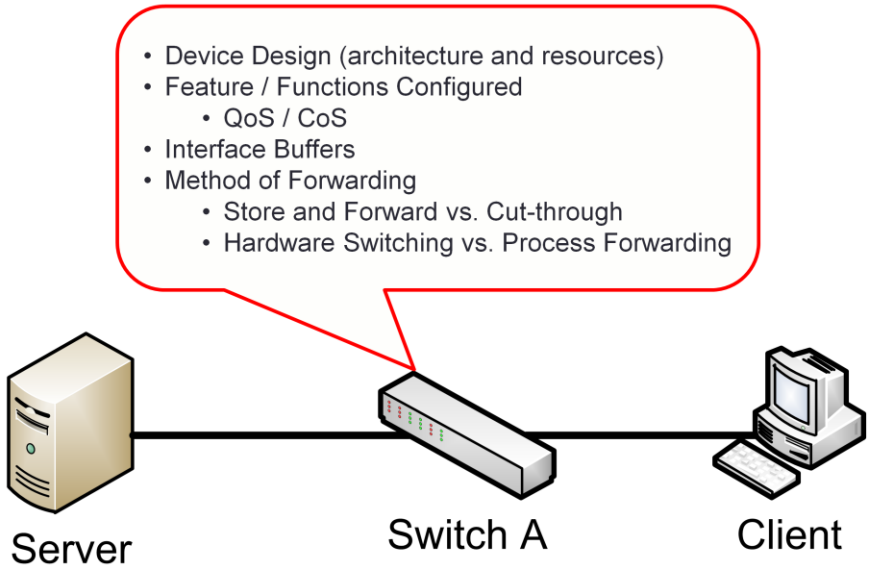
Location of Network Performance Impacts



6

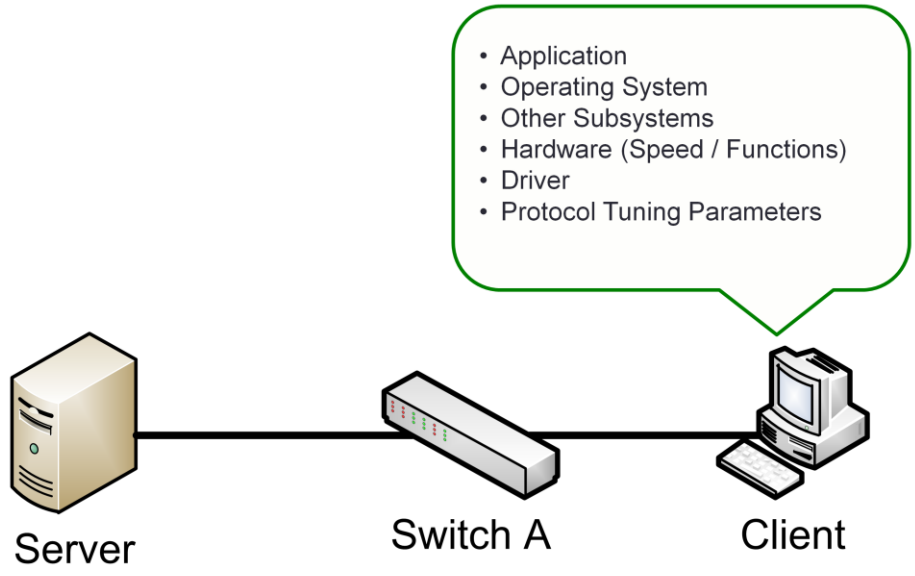
There are several points in a network path at which events can occur that can impact network performance.

Location of Network Performance Impacts



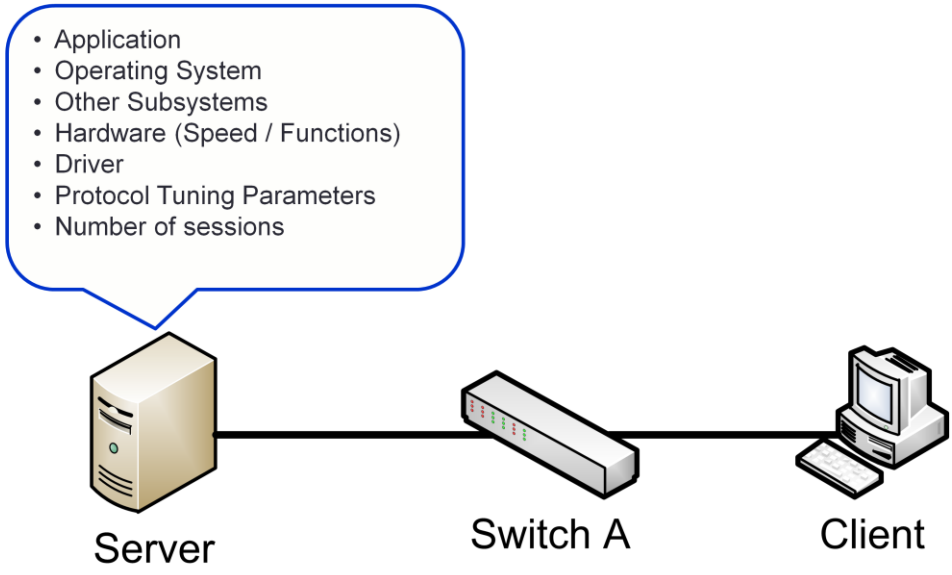
7

Location of Network Performance Impacts



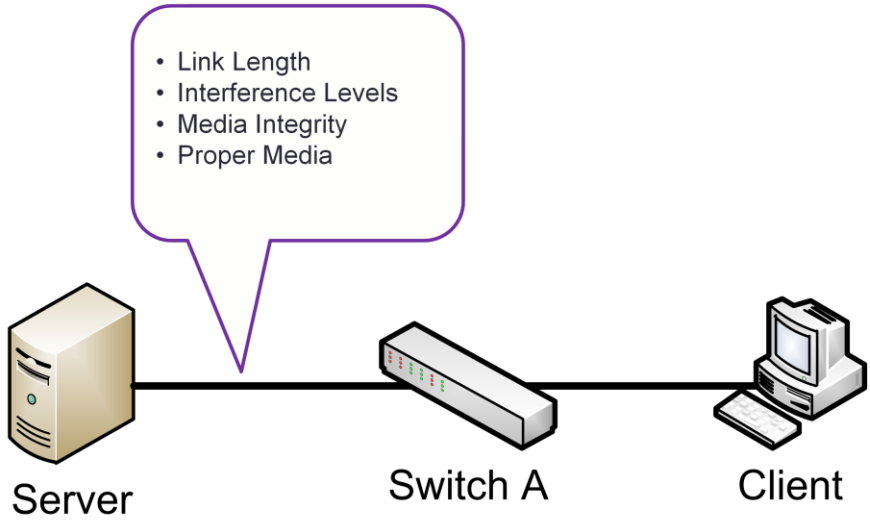
8

Location of Network Performance Impacts



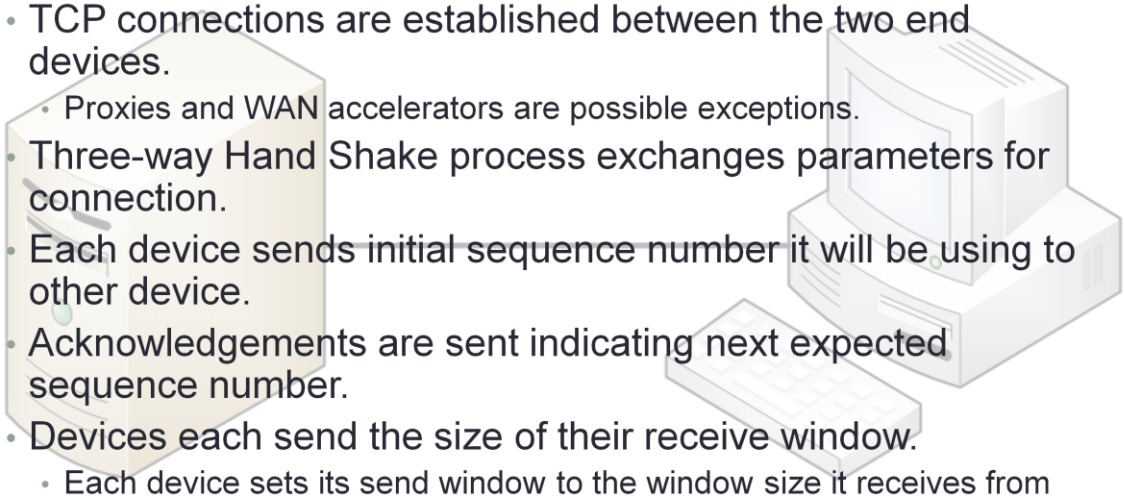
9

Location of Network Performance Impacts

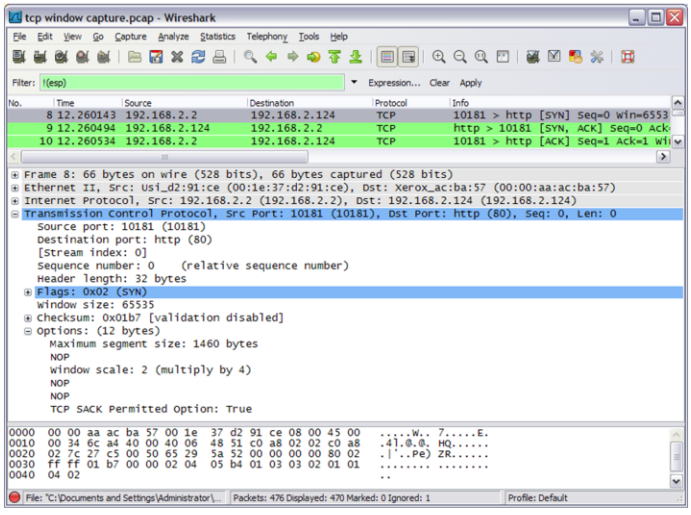


TCP Connections

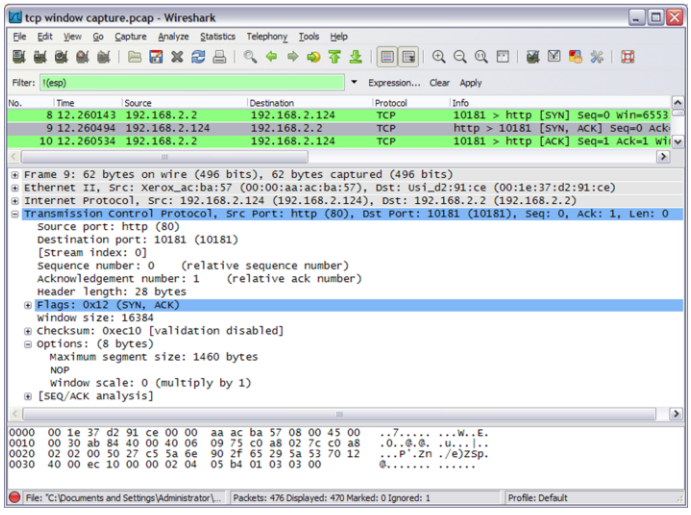
- TCP connections are established between the two end devices.
  - Proxies and WAN accelerators are possible exceptions.
- Three-way Hand Shake process exchanges parameters for connection.
- Each device sends initial sequence number it will be using to other device.
- Acknowledgements are sent indicating next expected sequence number.
- Devices each send the size of their receive window.
  - Each device sets its send window to the window size it receives from other device.
- Optional TCP parameters are also exchanged.



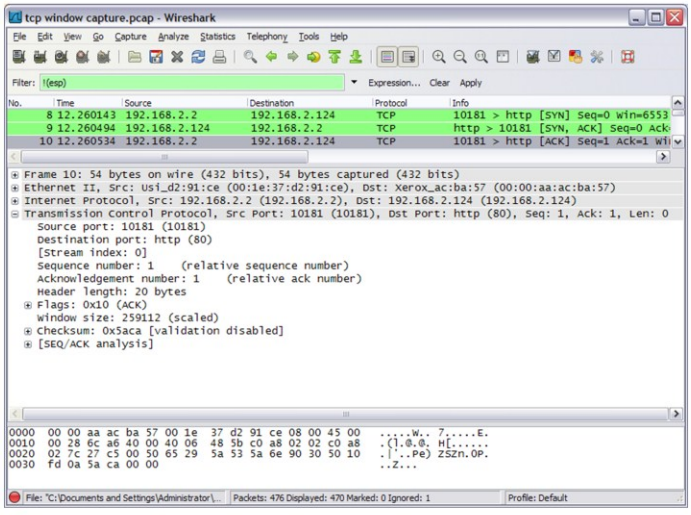
TCP “Three-Way Handshake” SYN



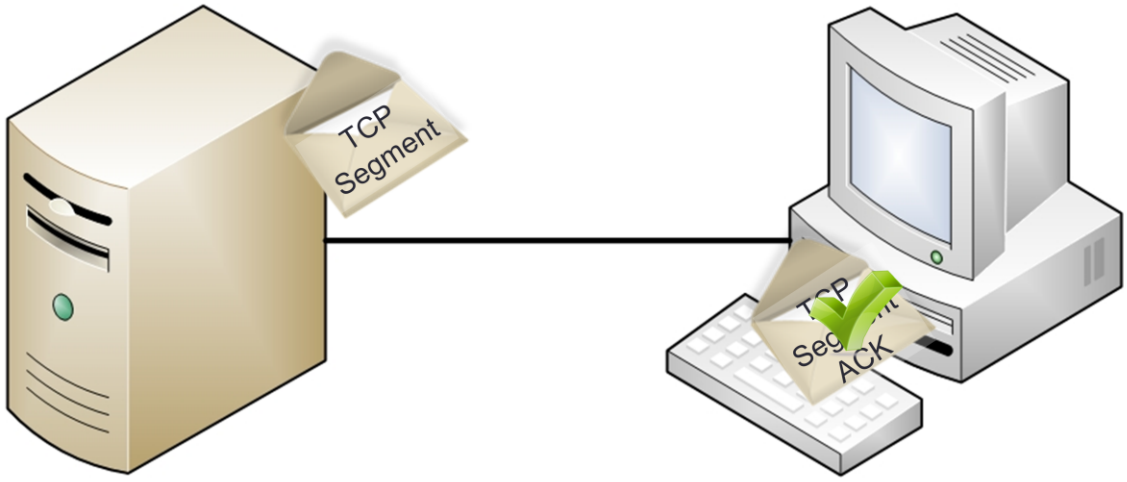
TCP “Three-Way Handshake” SYN - ACK



TCP “Three-Way Handshake” ACK



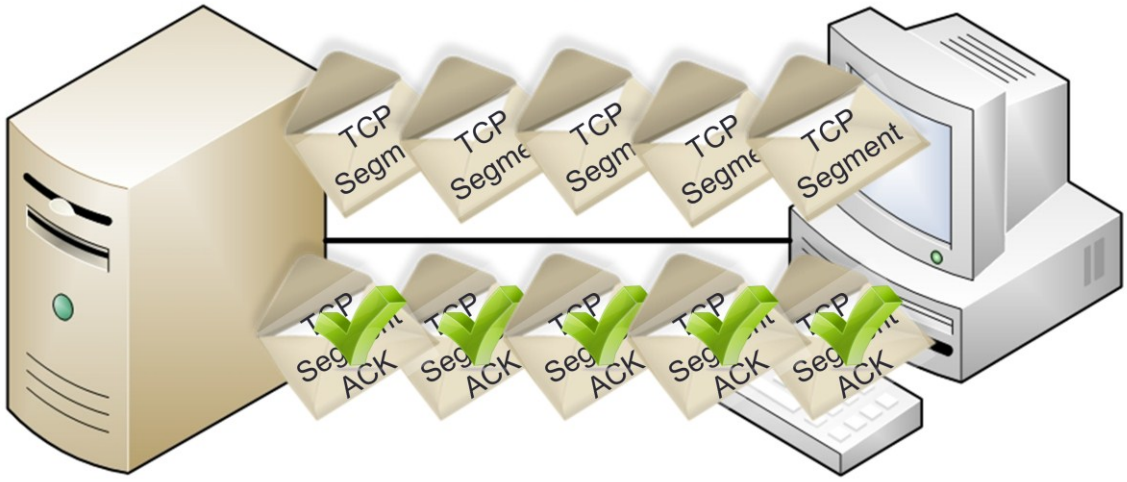
TCP Window Size (non-optimal)



15

Default or mistuned protocol parameters can result in inefficient transmission of data across the network.

TCP Window Size (optimal)



16

Properly tuned protocols provide for more efficient transmission of traffic. Tuning is a process that tailors the protocol parameters to the specific network environment that is in use. If the networking environment changes then tuning for those changes will be necessary to insure optimal configuration of the protocols on the servers and clients.

TCP Tuning Basics

$TP \leq RWIN / RTT$

- TP = Throughput
- RWIN = Receive Window
- RTT = Round-Trip Time

What is the max TP of a session with a .002 sec RTT and RWIN of 65535 bytes?

TP ≤ 65535 bytes / .002 sec
TP ≤ 32,767,500 bytes / sec or 262,140,000 bits / sec (262.14 Mb/s)

TCP Tuning Basics

RWIN ≥ TP * RTT

- TP = Throughput
- RWIN = Receive Window
- RTT = Round-Trip Time

What is the minimum RWIN required to achieve max TP with a single TCP session on a 1 Gb/s link with .002 sec RTT?

RWIN ≥ 1,000,000,000 b/s *.002 sec
RWIN ≥ 2,000,000 bits (divide by 8 to get bytes which is typical RWIN units)
RWIN ≥ 250,000 bytes

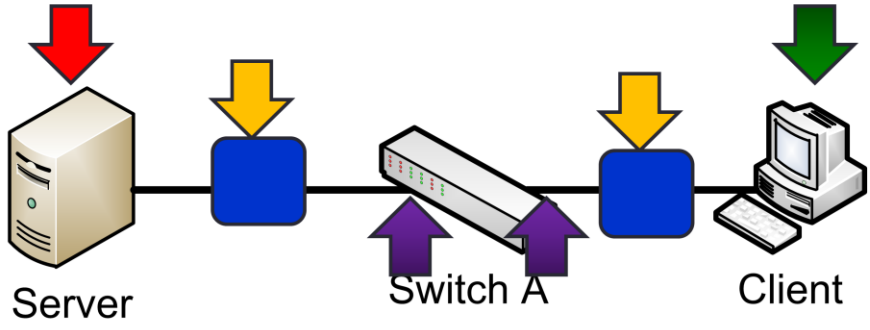
TCP Tuning Notes

- Default TCP parameters for devices typically do not provide optimal performance in every case; therefore, calculations need to be made to determine appropriate values.
- The RWIN value configured on the system should be a multiple of the TCP Maximum Segment Size (MSS). For Ethernet using standard 1500 as the MTU size the MSS would be 1460 bytes. In the prior example, 251120 would be the RWIN configured.
- Tuning for the highest possible throughput rate for a single session is not always desirable for overall system performance. The characteristics of the traffic and the number of concurrent sessions must be taken into consideration for optimizing traffic rates and system resource utilization. Increasing the TCP RWIN size allocates more memory resources for TCP use. This extra memory is allocated on a per session basis. Therefore, a server with a large number of concurrent sessions with the RWIN size configured very large would require a large amount of memory for TCP use.

TCP Tuning More Notes

- Some operating systems such as Windows Server 2008, Windows Vista, Windows 7 and Linux kernel versions 2.6.17 (and later) implement auto-tuning of the TCP parameters. Operating systems such as Windows XP, Windows Server 2003 and Linux kernel versions prior to 2.6.17 require manual tuning of the TCP parameters.

Where to Capture?



21

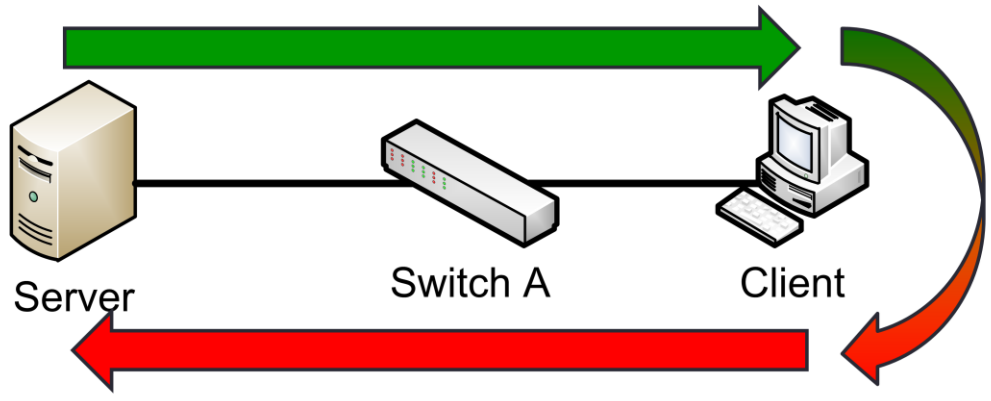
Captures of network traffic can be taken at various locations in the network. Depending on the environment, the access that is available and the purpose of the capture one position may be more desirable than others.

Captures may be obtain at the server or client by running tools such as Wireshark or tcpdump. This is not always an option as some device may not provide the option for such tools. The capture applications running on the server/client does consume system resources and this may impact the flow of traffic in a sensitive environment.

Captures may also be made by using port mirror functions of the network switches. Not all switch provide mirror capabilities. Port mirroring function in the switch tend to be lower priority functions and may not mirror all traffic to the target port in high utilization situations.

Captures may also be made by the insertion of a in-line trace tool or a media tap. This require an outage be scheduled for insertion and removal.

Sending Device Capture - Full Path Latency

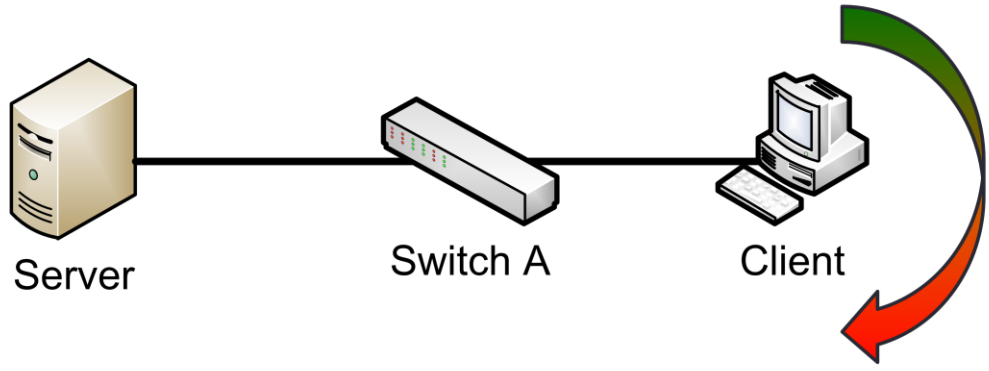


22

Network traffic captures at the sending device of a flow of traffic allows measurement of the latency through the network (round trip) including latency that occurs within the receiving system.

Capturing at the sender potentially provides an inflated measurement of real throughput of the network due to retransmission being captured.

Receiving Device to Capture – Internal Latency



23

Network traffic captures at the receiving device of a flow of traffic allows measurement latency that occurs within the receiving system only.

Capturing at the receiver provides for a more accurate measurement of real throughput of the network.

tcpdump vs Wireshark

- tcpdump is a command line tool for traffic capture or text display.
- tshark is similar tool provided with Wireshark distribution.
- Wireshark provides capture capability as well as detailed decode and analysis traffic.
  - WinPcap provides capture functions for Wireshark.
  - Wireshark can decode and perform basic trace analysis.
  - The ability to graph traffic details is often helpful.
- tcpdump capture files can be read by Wireshark.

Common tcpdump Commands

- tcpdump -w file.cap -i eth0
- tcpdump -w file.cap -i any
- tcpdump -w file.cap -i eth0 -s 0
- tcpdump -w file.cap -i eth0 not port 22
- tcpdump -w file.cap -i eth0 host 10.1.1.3 and host 10.1.1.5

25

tcpdump -w file.cap -i eth0

All traffic on eth0 interface is captured and written to file.cap The first 96 bytes of each frame are captured

tcpdump -w file.cap -i any

All traffic occurring on any network interface will be captured and written to file.cap

tcpdump -w file.cap -i eth0 -s 0

All traffic on eth0 interface is captured as a full frame capture and written to file.cap

tcpdump -w file.cap -i eth0 not port 22

All traffic other than port 22 traffic on eth0 interface is captured and written to file.cap The first 96 bytes of each frame are captured.

tcpdump -w file.cap -i eth0 host 10.1.1.3 and host 10.1.1.5

All traffic that contain both the addresses of 10.1.1.3 and 10.1.1.5 on eth0 interface is captured and written to file.cap The first 96 bytes of each frame are captured.

White Papers

Performance Issues with 10 Gigabit Ethernet
http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101954

Network Connectivity for Systems with Multiple Network Interfaces
http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP102133

Linux Bonding and VLANs with IBM BladeCenter
http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101773

Questions?

Thank You!