The Impact of Packet Loss on TCP Performance
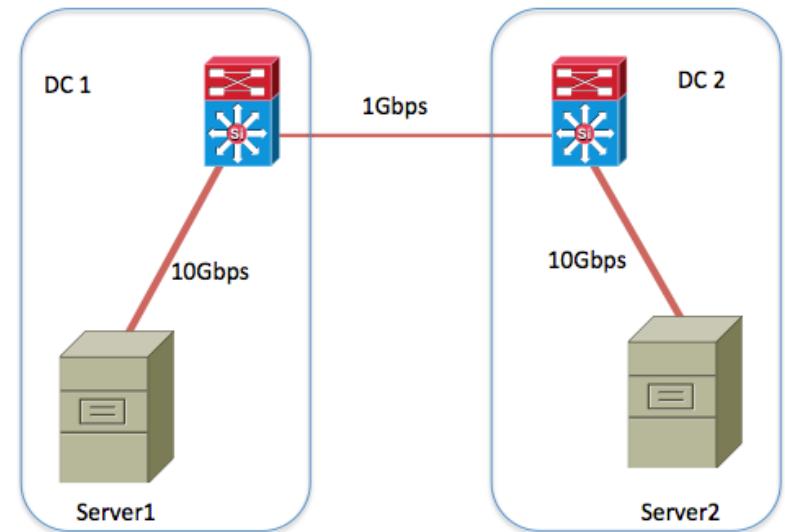
Terry Slattery
Chesapeake Netcraftsmen
Principal Consultant
CCIE #1026

Agenda

- **Application Slowness**

- The Impact of Packet Loss

- Causes of Packet Loss

- Detecting Packet Loss

- Fixing Packet Loss

- Background Reading

Application Slowness

- **Packet Loss (Prime Cause)**
- **Inadequate Server Performance**
  - **I/O bound process**
  - **DB locks**
- **Poor Application Architecture**
  - **Application ping-pong over high latency path**
  - **Choke points**
- **Buffer Bloat**
- **How Much Loss Is Acceptable?**
  - **A: 1%**
  - **B: .01%**
  - **C: .001%**
  - **D: .0001%**

Agenda

- **Application Slowness**

- **The Impact of Packet Loss**

- **Causes of Packet Loss**

- **Detecting Packet Loss**

- **Fixing Packet Loss**
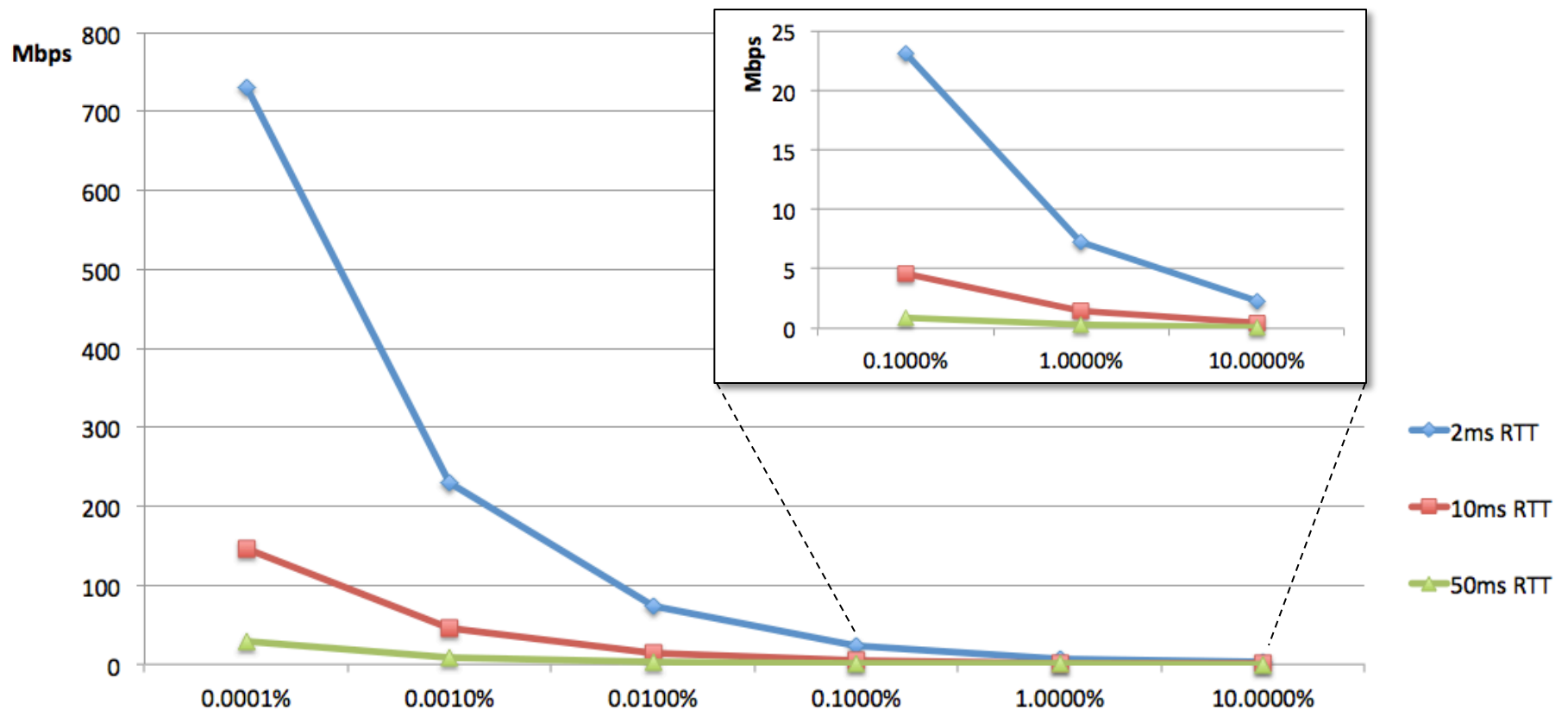
- **Background Reading**

The Mathis Equation

- ## TCP Throughput is based on Bandwidth-Delay product, assuming variable window size
   Rate = bandwidth x round-trip-delay

- ## TCP Throughput Experiments and Simulations
   – Paper: *The Macroscopic Behavior of the TCP Congestion Avoidance Algorithm*

$$rate \leqslant \frac{mss}{rtt} \times \frac{1}{\sqrt{p}}$$

- ## Typical loss rates:
   – Fiber: p=10E-13 (1 bad bit in 10E13 bits)

Huh? Please Explain.

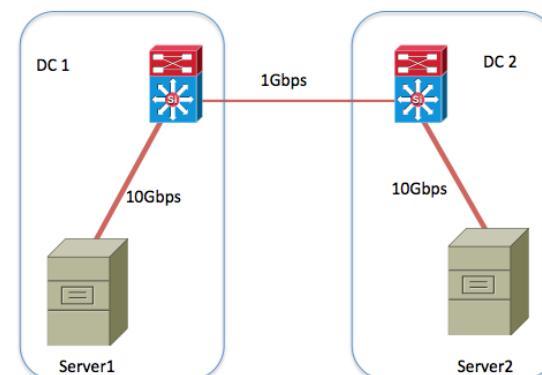- **Mathis Equation Describes TCP "Goodput"**
- **1Gbps Link**

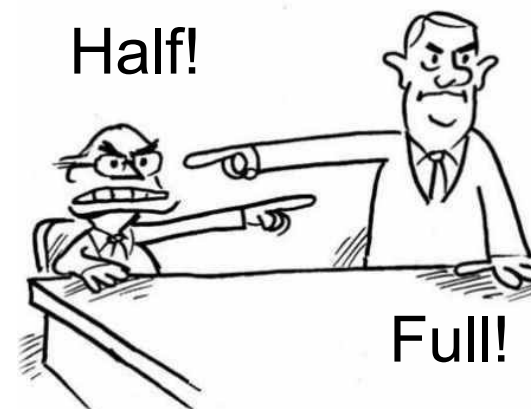# Agenda

- **Application Slowness**

- **The Impact of Packet Loss**

- **Causes of Packet Loss**

- **Detecting Packet Loss**

- **Fixing Packet Loss**

- **Background Reading**

Chesapeake
NETCRAFTSMEN

CISCO SYSTEMS | Premier Certified Partner

Causes of Packet Loss

- ## **Duplex Mismatch!**
  - **Either source or destination**

- ## **Bad Cabling (Occasional)**
  - **Pinched cable**
  - **Cat-3 cable**
  - **Bad connectors, poor crimps, corrosion**
  - **Scratched or dusty fiber connectors**

- ## **Egress Congestion**
  - **Too much data for the link capacity**

- ## **Ingress Overruns**
  - **Fast servers, old switch blade**

Buffer Bloat

- **Confuses TCP retransmission algorithm**
- **Retransmits at 2*RTT**
- **Excessive buffering holds more than 2*RTT of data**

- **Result: TCP retransmits after 2*RTT**
  - Some packets are delivered more than once
  - Wastes network bandwidth

Agenda

- **Application Slowness**

- **The Impact of Packet Loss**

- **Causes of Packet Loss**

- **Detecting Packet Loss**

- **Fixing Packet Loss**

- **Background Reading**

Detecting Packet Loss

- **Monitor Interface Stats**
  - **Duplex mismatch**
    - **FCS, CRC, Runts on full duplex interface**
    - **Late collisions on half duplex interface**
  - **Errors of any type**
  - **NMS or periodic 'show interface | include'**
- **NMS Caveat**
  - **A log of low volume interfaces obscure high volume interfaces**
  - **Two thresholds:**
    - **% errors**
    - **Total count of errors**
- **Interface Discards/Drops**
  - **Congestion indicator**

Increasing Loss on a Fiber Link

- ## 10G Fiber Backbone
  - ### Bad patch cable
  - ### All other similar links were clean

Detecting Packet Loss

- **Monitor Client or Server TCP Stats**
    - **Windows reports a rate via SNMP, not a count**
    - **TCP retransmission counts via CLI**

        ```
        netstat -sp tcp
        -s = display stats
        -p tcp = tcp protocol stats
        ```

- **Warning…**
    - **TCP flow control is based on packet loss, so…**
    - **Some retransmissions are normal**
    - **Look for** $\dfrac{\text{Retransmissions}}{\text{Transmitted Packets}}$ < 0.0001% or better

Agenda

- **Application Slowness**

- **The Impact of Packet Loss**

- **Causes of Packet Loss**

- **Detecting Packet Loss**

- **Fixing Packet Loss**

- **Background Reading**

Fixing Packet Loss

- ## Check Duplex Settings!
  - Memory of negotiation problems still exists
  - Auto-Auto works with newer equipment
  - Document devices that need fixed speed/duplex
- ## Replace Cables and Connectors
- ## Add QoS on Congested Interfaces
  - Heavily oversubscribed links need more BW
  - 95th percentile of 50% is a reasonable threshold
- ## Bad Switch Port
  - Look for ingress overruns on old blades
    - Several 1G servers on port group of 1G ASIC

Agenda

- **Application Slowness**

- **The Impact of Packet Loss**

- **Causes of Packet Loss**

- **Detecting Interface Errors**

- **Fixing Interface Errors**

- **Background Reading**

Background Reading

- **Blogs at Netcraftsmen.net**
  - **Search for TCP or Mathis**

- **TCP Performance Papers**
  http://www.linuxsa.org.au/meetings/2003-09/ tcpperformance.print.pdf

- **"Buffer Bloat"**

Summary

- **TCP is sensitive to small percentages of packet loss**

- **The causes are easy to determine and fix**

- **Few people understand the impact**

- **Fewer people work to clean up their network**

- **"What's In Your ~~Wallet~~ Network?"**

Questions?

Terry Slattery
Chesapeake Netcraftsmen