Allocating Network Bandwidth to Match Business Priorities

Speaker
Peter Sichel
Chief Engineer
Sustainable Softworks
psichel@sustworks.com

MacWorld San Francisco 2006
Session M225
12-Jan-2006 10:30 AM - 12:00 PM

Introduction

Speaker Background

Macintosh developer specializing in kernel level TCP/IP networking for the past 10 years.

- IPNetMonitorX - Macintosh internet tools
- IPNetTunerX - network optimization tuner
- IPNetRouterX - native IP connectivity suite
- 14 years prior experience as software architect at Digital Equipment Corporation. Emphasis on user interface design and communications.

Overview

What is this session about:

- Improving network throughput and availability
- What happens when you push TCP/IP to the extreme?
- Techniques for addressing difficult performance problems

TCP tuning

- Why is it necessary to tune TCP/IP?
- Optimization 101 - measure, adjust, repeat

Beyond tuning

- When is tuning not enough?
- Packet Shaping and QoS (Quality-of-Service)
- Bandwidth Allocation Examples
- Transparent Proxy

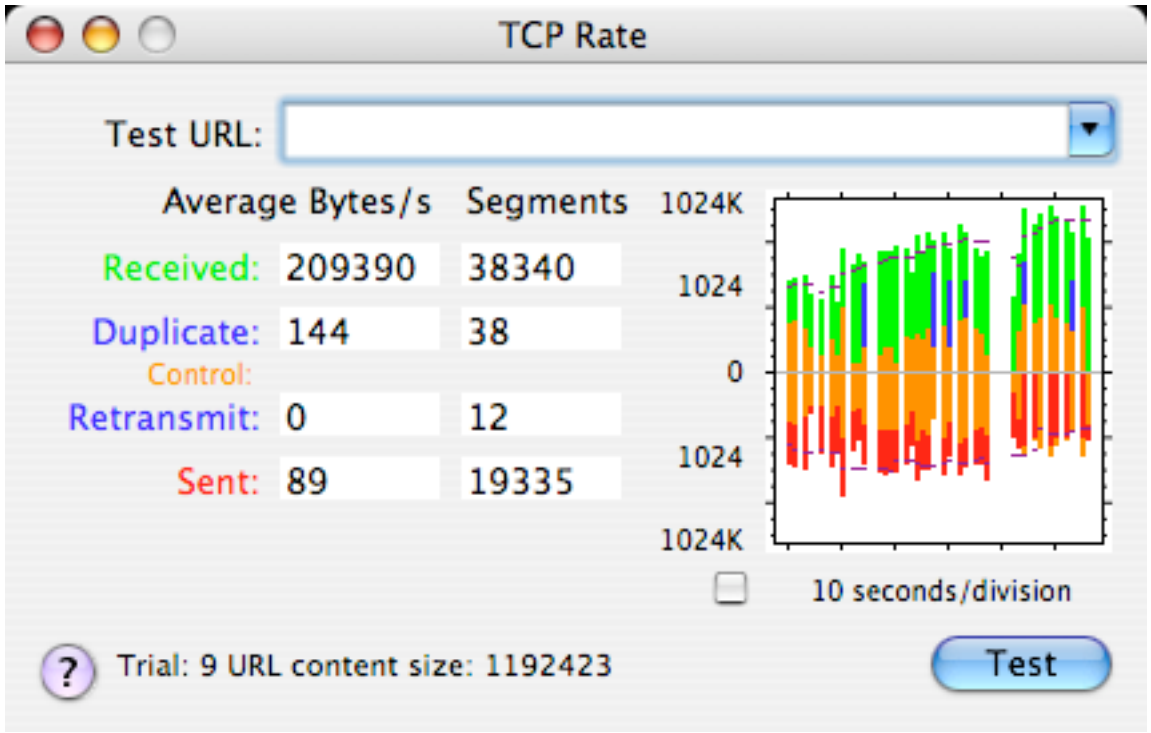
Why is it necessary to tune TCP/IP?

- TCP/IP is not controlled by any manufacturer and is designed to work on almost any kind of underlying network with widely varying characteristics.

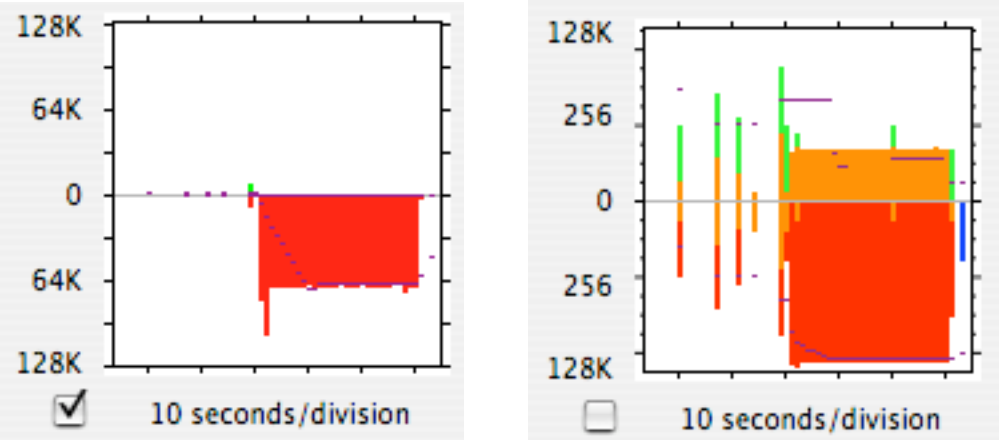
- To meet this challenge, the protocol designers made TCP adaptive. TCP is self monitoring and optimizes its own behavior to match the network environment. Adaptation takes time and the default settings cannot optimize for every possible kind of network.

- There are many independent implementations, the protocols continue to evolve in response to practical experience, some implementations have peculiar compatibility constraints.

Optimization 101 - measure, adjust, repeat



Problematic transfer with duplicate/retransmit data



Efficient transfer (Linear and Log scale)

When is tuning not enough?

When data links are saturated.

When you want to allocate bandwidth based on business priorities.

When you want redundancy, automatic fail over, or load balancing to increase available bandwidth.

Packet Shaping and QoS (1)

Each TCP connection tries to use as much bandwidth as it can get, and backs off when packets are lost or delayed.

When a data link becomes heavily saturated by multiple competing applications:

- ACKs can be delayed limiting throughput
- Efficiency drops as data is retransmitted
- Connections back off to avoid congestion

The purpose of Packet Shaping and QoS is to adapt network behavior to increase TCP/IP efficiency over the “slow link” or “last mile”.

Packet Shaping and QoS (2)

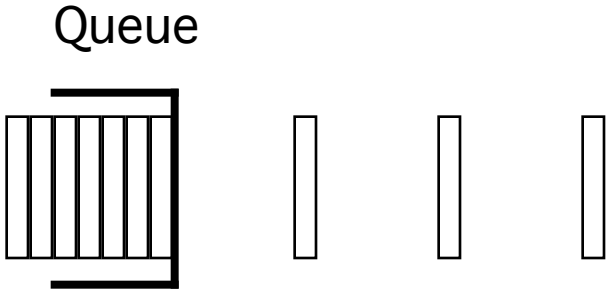
Three basic techniques:

- Queuing (buffering and waiting) to limit send rate. Since TCP is adaptive, it will self adjust to the impaired data link.
  `Examples: throttled, LinkSys WRT-54, Xincom`

- Expedite higher priority traffic by sending out of order or reserving dedicated bandwidth. May support TOS field in IP header.
  `Example: throttled`

- Use TCP's built-in flow control mechanism to modulate the send rate for maximum efficiency.
  `Examples: Packeteer, IPNetRouterX, IPNetSentryX`

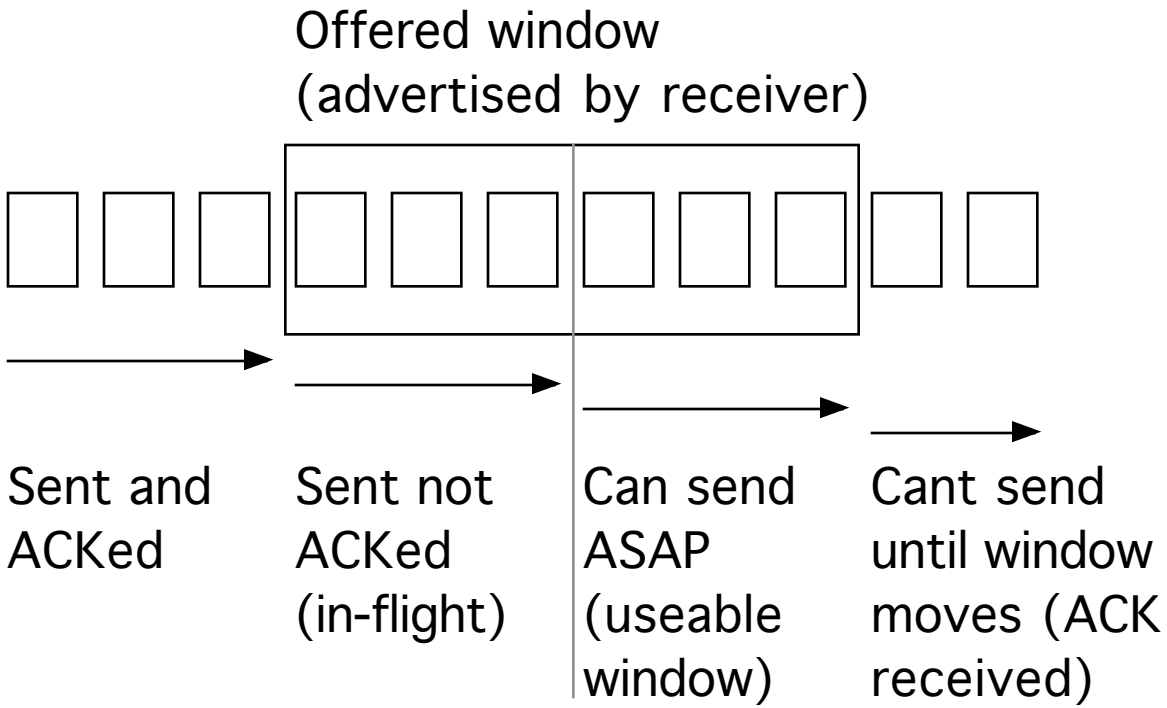
Packet Shaping Techniques (1)



Queuing (buffering and waiting)

- Easiest to implement, can be applied to any protocol
- Can only rate limit send traffic
- Limited number of queues
- “Shaping” amounts to re-ordering and dropping packets
- Can be effective, but sacrifices some efficiency

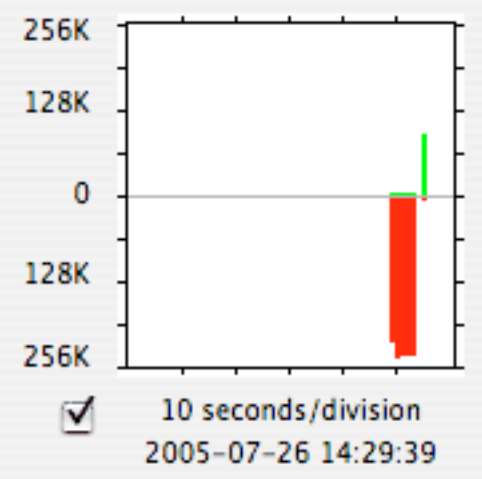
Packet Shaping Techniques (2)



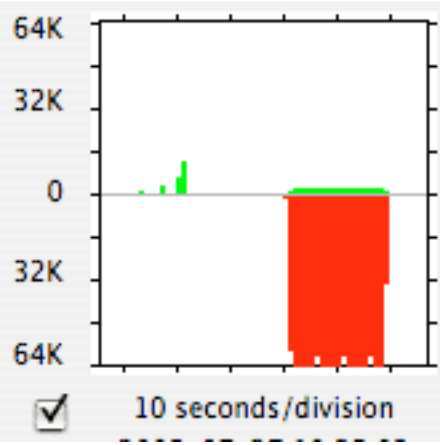
TCP Rate Limiting

- Adjust receive window on the fly to control when and how much the window moves.
- Can rate limit send and receive traffic
- Avoids congestion for best use of available bandwidth.

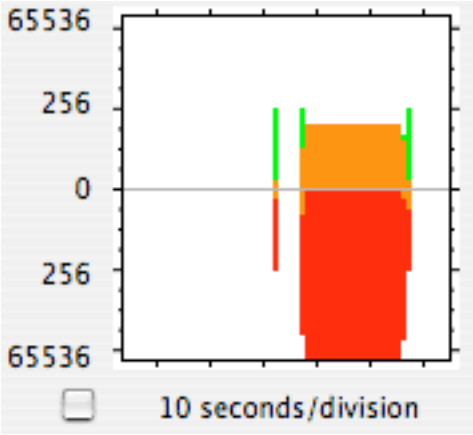
Packet Shaping Example

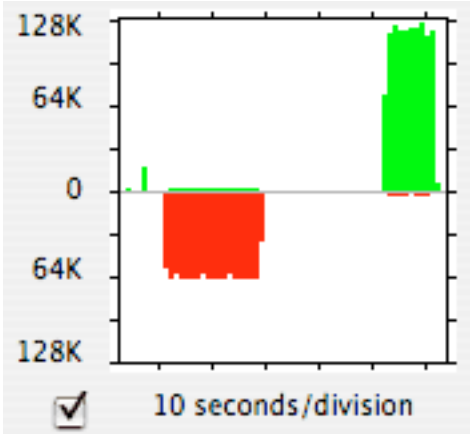


Send 1.1 MB file at 2 Mpbs



512 Kbps send limit



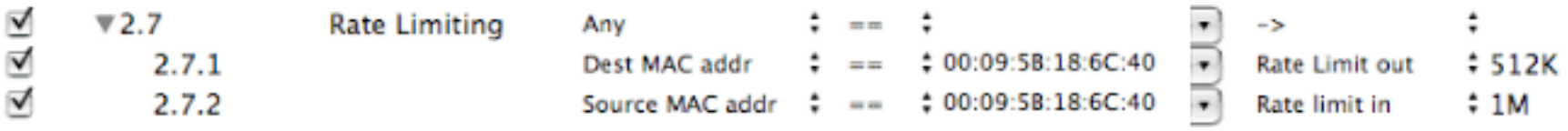


1 Mbps receive limit

Bandwidth Allocation (1)

Classify and monitor traffic to identify data flows of interest.

- Devices (Internet router, server, host,...)
- Services (Web, Email, File Transfer, VOIP)



Bandwidth Allocation (2)

Maintain network responsiveness during heavy Email or file transfers.

Improve network performance during simultaneous upload and download.

Improve server responsiveness when frequently accessed.

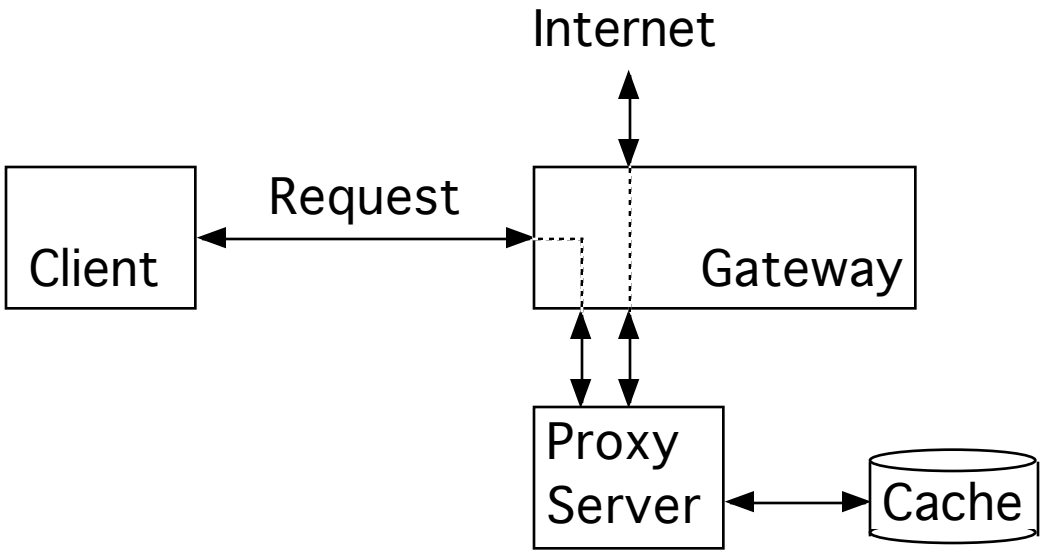
Reserve bandwidth for VOIP.

Transparent Proxy

Redirect client requests without any special configuration or knowledge at the client.

Cache frequently requested pages or files to reduce WAN traffic.

Block even the most aggressive peer-to-peer protocols.



Demo / Q & A

More Information:
http://www.sustworks.com
psichel@sustworks.com
(send Email for updated slides)