

Trans-Oceanic Backbones Basic Techniques

- Six Techniques used by the Internet Community to scale bandwidth:
 - ✓ Bigger Circuits
 - ✓ Inverse Multiplexing
 - ✓ Clear Channel E3 or DS3
 - ✓ PPP over SDH
 - ✓ Asymmetrical Satellite Systems
 - ✓ Hybrid Systems

© 1998, Cisco Systems, Inc.

www.cisco.com

Trans-Oceanic Backbones Basic Techniques

Easiest way to scale bandwidth is to upgrade the circuit. There are not many OC-12s (633Mbps) circuits across the Atlantic.

Problem: Many Telcos fail in their engineering and capacity planning for transmission capacity - leading to limits on the upgrade capability (i.e. locked to nxE1 upgrades).

© 1998, Cisco Systems, Inc.

www.cisco.com

Trans-Oceanic Backbones Basic Techniques

“ Pearl: It helps to work directly with the people who do the international transmission capacity planning and purchasing. They get to see your projections, you get on time upgrades. ”

© 1998, Cisco Systems, Inc.

www.cisco.com

Trans-Oceanic Backbones Basic Techniques

- Inverse Multiplexing (iMux) takes several circuits and *bundles* them into one or more logical circuits.
- Two major techniques:
 - ✓ Use protocol/forwarding features in the router
 - ✓ Use an external inverse multiplexer

Router Multiple Circuits Router

© 1998, Cisco Systems, Inc.

www.cisco.com

Trans-Oceanic Backbones Basic Techniques

- Router protocol/forwarding features as iMux
 - Parallel Links Across the Ocean. $n \times$ E1 circuits between the routers using the routing protocols to perform the load balancing and bundling of the parallel circuits - works up to 4 to 8 E1s.



Presentation_ID

© 1998, Cisco Systems, Inc.

7

Trans-Oceanic Backbones Basic Techniques

- Several Techniques:
 - Static Route - Per Packet
 - OSPF
 - eBGP Multihop - Per Flow (w/ Netflow & CEF) or Per Packet (w/ CEF)
 - Multi-Link PPP (MLPPP) - Tighter Bundling Options (up to 8)
 - CEF Load Balancing - Per packet or per flow
 - BGP Maximum Paths (up to 6 - different routers)

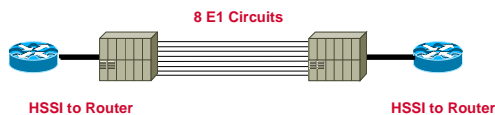
Presentation_ID

© 1998, Cisco Systems, Inc.

8

Trans-Oceanic Backbones Basic Techniques

- Many Telcos have not provisioned facilities to cater to oceanic circuits above E1.
- Many E1s grouped together into a larger pipe via iMux technology.
- Defacto Industry practice is to use Cisco Routers with HSSI ports connected to Larscom's Orion 4000 iMUXes.



Presentation_ID

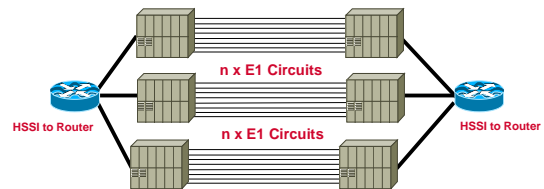
© 1998, Cisco Systems, Inc.

www.cisco.com

9

Trans-Oceanic Backbones Basic Techniques

- Several iMUX bundles can be grouped together on the same router to build 34M and 45M equivalent circuits
- eBGP Multihop is the preferred load balancing technique.
- Telstra Internet is now over 100M of iMUX backbone bandwidth!



Presentation_ID

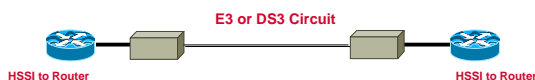
© 1998, Cisco Systems, Inc.

www.cisco.com

10

Trans-Oceanic Backbones Basic Techniques

- Clear Channel E3 (34M) or DS3 (45M)
 - Preferred method for high speed backbone links is a clear channel circuit.
 - Configuration is simple - connect the CSU/DSU to a HSSI/T1/E1 port on a Cisco router.



Presentation_ID

© 1998, Cisco Systems, Inc.

www.cisco.com

11

Trans-Oceanic Internet Backbones

- What do you do after DS3 (45M)?
 - Multiple DS3?
 - ATM at OC-3 (155M) or OC-12 (622M)?
 - PPP over SDH at OC-3 (155M) or OC-12 (622M)?
- ATM vs PPP over SDH (POSIP)
 - ATM is not the best choice when all you are doing is Internet traffic across the ocean.

Presentation_ID

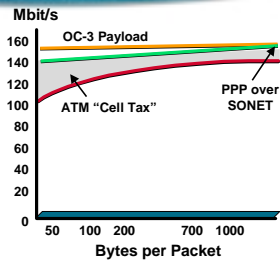
© 1998, Cisco Systems, Inc.

www.cisco.com

12

Trans-Oceanic Internet Backbones

- **Line utilization**
 - ✓ Packet over SONET provides 98% utilization
 - ✓ 20% - 30% overhead over an ocean is a lot of money!
- **Goodput verses throughput**
 - ✓ Sufficient buffering for large TCP flows
 - ✓ Congestion avoidance
 - ✓ ATM and TCP/IP Headers are part of "throughput," not "goodput"



Presentation_ID © 1998, Cisco Systems, Inc. www.cisco.com

13

Trans-Oceanic Backbones Basic Techniques

POSIP Overview

- ✓ Packet-oriented serial interface, OC3/STM-1, OC-12, OC-48
- ✓ Supports either SONET or SDH interface
- ✓ PPP packets are encapsulated in SONET STS-3c
 - POS available information bandwidth: 149.76 Mbps
 - ATM available information bandwidth: 128.36 Mbps
 - Cell tax ranges from 14-50% depends on the packet distribution
- ✓ Encapsulations
 - RFC 1619 point-to-point protocol over SONET/SDH
 - RFC 1662 point-to-point protocol in HDLC-like framing
- ✓ IP protocol only

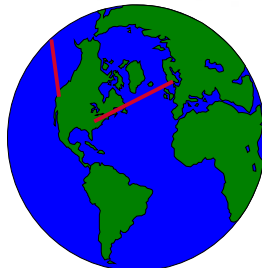
Presentation_ID © 1998, Cisco Systems, Inc.

www.cisco.com

14

Trans-Oceanic Internet Backbones

- Pensacken NY (Sprint NAP) to Stockholm Sweden (D-GIX).
- First 155 M Internet link across an ocean!
- First operational 155M POSIP connection!
- Second is from Japan to US



Presentation_ID © 1998, Cisco Systems, Inc. www.cisco.com

15

Trans-Oceanic Backbones Basic Techniques

- If an ISP's trans-oceanic traffic pattern is always asymmetrical, then why pay for that idle bandwidth going out?
 - ✓ Asymmetrical traffic pattern is when the ISPs is pulling down more information than sending out. Today a typical ISP is somewhere between a 80:20 - 60:40 traffic ratio to the US.
- Satellite Services allow ISPs to buy a circuit with different speeds in the two different directions.
 - ✓ For example - 2 Mbps in / 256 Kbps out
- Hence the ISPs only pays for what they need - no idle bandwidth giving the other side a free ride.

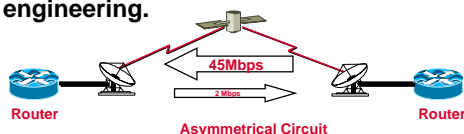
Presentation_ID © 1998, Cisco Systems, Inc.

www.cisco.com

16

Trans-Oceanic Backbones Basic Techniques

- Trans-Oceanic ISP links are a reality. They work and are used to backup terrestrial trans-oceanic links.
- Latency issues (around 700 ms RTT) can be minimized through creative engineering.



Presentation_ID © 1998, Cisco Systems, Inc. www.cisco.com

17

Trans-Oceanic Backbones Basic Techniques

- **Minimizing Latency on Asymmetrical Satellite Links:**
 - ✓ Good Traffic Engineering (i.e. avoid congestion)
 - ✓ Hybrid Asymmetrical Links
 - ✓ Hybrid Simplex Links
 - ✓ WWW Caching
 - ✓ Content Routing

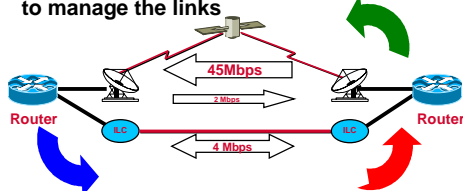
Presentation_ID © 1998, Cisco Systems, Inc.

www.cisco.com

18

Trans-Oceanic Backbones Basic Techniques

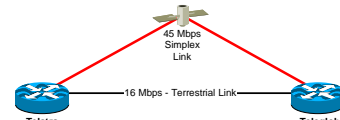
- Hybrid Asymmetrical Satellite links combine terrestrial and satellite together.
 - ✓ Reduces the latency by 1/3 to 1/2.
 - ✓ Static, BGP filtering, or eBGP multihop is used to manage the links



Presentation_ID © 1998, Cisco Systems, Inc. www.cisco.com 19

Trans-Oceanic Backbones Basic Techniques

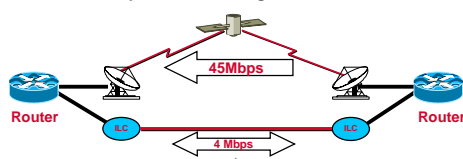
- Telstra and Teleglobe were the first two ISPs who pioneered this technique.
 - ✓ Several other ISPs in Asia and Europe are using this technique.
 - ✓ Very few Tier 1 & 2 NSPs in the US will terminate these trans-oceanic systems - hence the growth of co-lo business (i.e. AboveNet)



Presentation_ID © 1998, Cisco Systems, Inc. www.cisco.com 20

Trans-Oceanic Backbones Basic Techniques

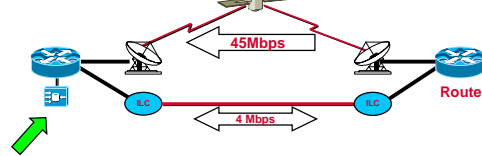
- Hybrid Simplex Satellite links combine terrestrial and satellite's capability to for a circuit with only direction..
 - ✓ Takes advantage of uni-directional nature of satellite circuits
 - ✓ Reduces the latency by 1/3 to 1/2.
 - ✓ Static, BGP filtering, or eBGP multihop is used to manage the links with new protocols coming.



Presentation_ID © 1998, Cisco Systems, Inc. www.cisco.com 21

Trans-Oceanic Backbones Basic Techniques

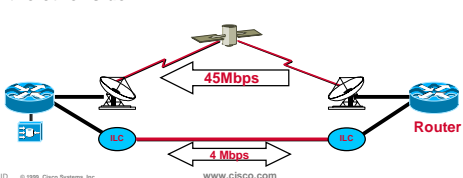
- WWW Caching adds an additional "buffer" to the higher latency.
 - ✓ Caches content as it comes over the link.
 - ✓ Minimizes the download of the same content over and over again.
 - ✓ WWW Caching is a main stream Internet technology - all issues with it's use have been resolved via technical means via IETF and W3O.



Presentation_ID © 1998, Cisco Systems, Inc. www.cisco.com 22

Trans-Oceanic Backbones Basic Techniques

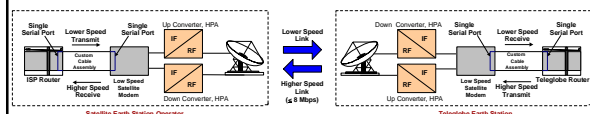
- Since asymmetrical systems will limit any "free rides" from people who should be paying for the traffic from the ISP to the world.
 - ✓ The system is designed around the ISP's traffic profile.
 - ✓ Choke point in the ISP's space limits excessive pull from the other side.



Presentation_ID © 1998, Cisco Systems, Inc. www.cisco.com 23

Trans-Oceanic Internet Backbones

Asymmetric Satellite Service - Scenario 1A
ISP Router in Earth Station
Receive and Transmit Links (less than or equal to 8 Mbps)



Single low-speed satellite modem supports up-link and down-link speeds up to 8 Mbps

Slides courtesy of Teleglobe www.cisco.com 24

Trans-Oceanic Internet Backbones

Asymmetric Satellite Service - Scenario 1B
 ISP Router in Earth Station
 Receive Link to ISP (greater than 8 Mbps)
 Transmit Link from ISP (greater than 8 Mbps)

Single high-speed satellite modem supports up-link and down-link speeds up to 45 Mbps

Slides courtesy of Telelobe www.cisco.com

Trans-Oceanic Internet Backbones

Telelobe & Intelsat Asymmetric E1/512 Kbps Satellite Link
 for Africa Telecom '98 Internet Cafe

Presentation_ID © 1998, Cisco Systems, Inc. www.cisco.com

Trans-Oceanic Internet Backbones

- **Equipment Issues**
 - ✓ Need a interface card in a router with lots of buffering. VIP2-50 with max memory connected to a HSSI PA or POS PA (DS-3) are known to work.
 - ✓ Need Random Early Detection (RED). Needed to insure effective utilization of the link.

Presentation_ID © 1998, Cisco Systems, Inc. www.cisco.com

New Trends

What others are doing with their Trans-Oceanic Links

Presentation_ID © 1998, Cisco Systems, Inc. www.cisco.com

New Trends

- Co-Location or Lease of Router in the US.
- Dual Sided Content Routing.
- Heavy Localization of traffic (IXPs)

Presentation_ID © 1998, Cisco Systems, Inc. www.cisco.com

New Trends

- ISP's Trans-Oceanic Backbones are migrating into *systems* designed to get maximized efficiency from the link.
- High Cost of trans-oceanic bandwidth, Exponential Growth, and new demand for Value Added Services (QoS, Content Routing, and VoIP) are all driving factors.

Presentation_ID © 1998, Cisco Systems, Inc. www.cisco.com

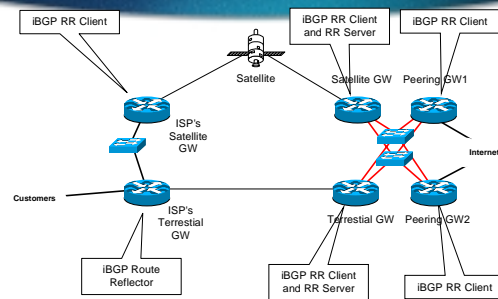
New Trends

- These Trans-Oceanic Systems will consist of:
 - ✓ Mix of Satellite and Terrestrial Circuits
 - ✓ WWW Caching and Performance Enhancing Proxies
 - ✓ QoS Services (Premium and Basic)
 - ✓ Application Redirection (Voice and Multicast)

Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com

31

New Trends

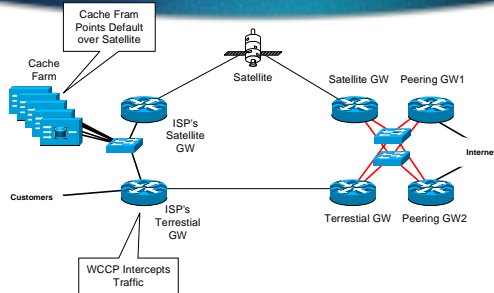


Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com

www.cisco.com

32

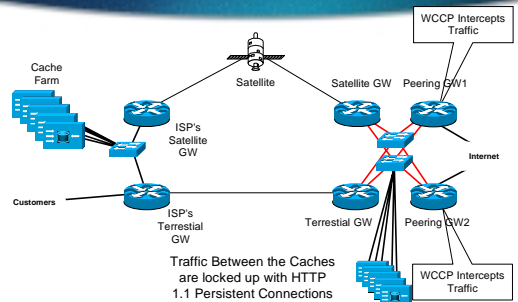
New Trends



Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com

33

New Trends



Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com

www.cisco.com

34

New Trends

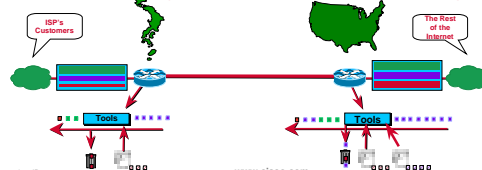
- HTTP 1.1 Persistent connections between the two caches move the average packet size from ~40 bytes to 512 - 1500 bytes.
 - ✓ More Goodput vs Overhead
 - ✓ Satellite Modifications to TCP increase efficiencies (RFC 2488 or equivalent)
- Technique is also called *Performance Enhancing Proxies (PEP)*
 - ✓ draft-ietf-pilc-pep-00.txt

Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com

35

New Trends

- Any QoS, CoS, or DiffServ tools need to be applied on the *upstream router's* interface.
 - ✓ Applying the tools on the *downstream* side would force the ISP to pay for the packets before they are dropped.
 - ✓ Most US NSPs are reluctant to apply any special configurations on the US side.
 - ✓ That means the ISP outside of the US needs to co-locate their own router in the US (or lease one from a NSP in their facilities).



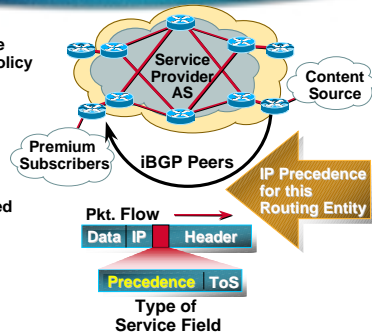
Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com

www.cisco.com

36

New Trends

- Use BGP to propagate precedence-setting policy as needs change
- BGP community label used to bind precedence policies to destination prefixes
- Return traffic classified with appropriate QoS

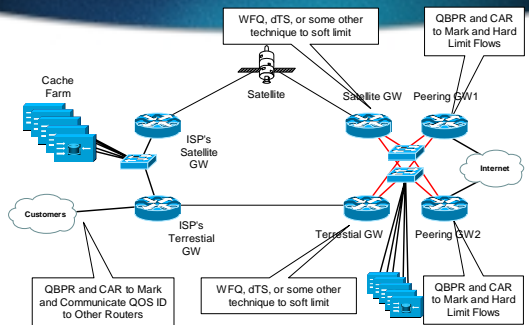


Presentation_ID © 1999, Cisco Systems, Inc.

www.cisco.com

37

New Trends



Presentation_ID © 1999, Cisco Systems, Inc.

www.cisco.com

38

New Trends

What's Next?

- ✓ VoIP and IP Telephony. Using routing tricks or NPR (Netflow Policy Based Routing) to keep voice traffic on the terrestrial circuits.
- ✓ Content Replication. Pushing content from the international side to the US side of the link.

Presentation_ID © 1999, Cisco Systems, Inc.

www.cisco.com

39

Collecting and Reporting Capacity Information

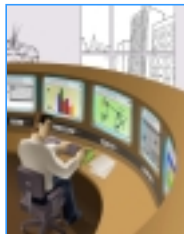
Presentation_ID © 1999, Cisco Systems, Inc.

www.cisco.com

40

Internet Traffic Measurement

- Aggressive collections and analysis of network data is critical to any ISPs who wishes to tackle the problems of CoS, QoS, and I3F
- Ironically, most ISPs do not collect this information, even when most of the tools are public domain on the Internet.
- The concern is that so many people are talking about buzzword and not enough about the fundamentals of what is actually happening on the Internet.



Presentation_ID © 1999, Cisco Systems, Inc.

www.cisco.com

41

Performance Management and Capacity Planning Definitions

- Capacity planning
 - ✓ The process of determining the likely future network resource requirements to prevent a performance impact on business critical applications
- Performance management
 - ✓ The practice of managing network service response time, consistency and quality for individual services and services overall

Presentation_ID © 1999, Cisco Systems, Inc.

www.cisco.com

42

Increasing Importance of Capacity Planning

- Frequent application deployment failure
- Increased reliance on network services for business applications
- Exponential growth in business and nonbusiness related traffic
- Network Failure is typically capacity related

Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com

43

Capacity Related Risks

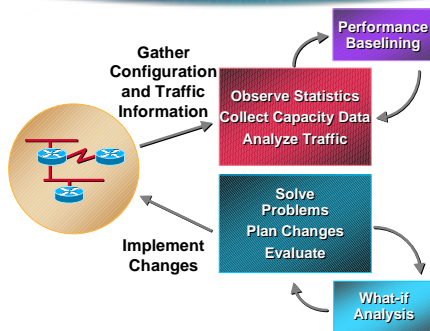
- Network degradation and failure
- Application timeouts and failure
- Application performance degradation

Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com

www.cisco.com

44

Effective Capacity Management



Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com

45

Resource Constraints or Bottlenecks

- CPU
- Memory
- Buffering, queuing and latency
- Interface and pipe sizes
- Speed and distance
- Application characteristics

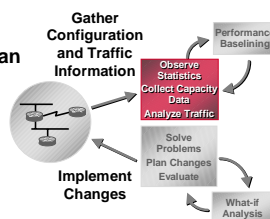
Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com

www.cisco.com

46

Collecting and Reporting Capacity Information

- Development of information collection plan
- Tools for collecting capacity information
- Defining capacity areas
- Reporting and interpreting results



Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com

47

Information to Collect

- Link utilization
- CPU
- Memory
- Performance (ping response time)
- Queue/buffer drops
- Broadcast volume
- Frame Relay DE, FECN, BECN, traffic-shaping parameters
- NetFlow statistics
- RMON

Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com

www.cisco.com

48

Link Utilization

Resource	Address	Segment	Avg. Util (%)	Peak Util (%)
JTKR01S2	10.2.6.1	128 Kbps	66.3	97.6
JYKR01S0	10.2.6.2	128 Kbps	66.3	97.8
FMC18S4/4	10.2.5.1	384 Kbps	51.3	109.7
PACR01S3/1	10.2.5.2	384 Kbps	51.1	98.4

Presentation_ID © 1998, Cisco Systems, Inc.

www.cisco.com

49

CPU Utilization

Resource	Polling Address	Avg. Util (%)	Peak Util (%)
FSTR01	10.28.142.1	60.4	80
NERT06	10.170.2.1	47	86
NORR01	10.73.200.1	47	99
RTCR01	10.49.136.1	42	98

Presentation_ID © 1998, Cisco Systems, Inc.

www.cisco.com

50

Performance (Ping Response Time)

Resource	Address	AvRes T (mS) 09-09-98	AvRes T (mS) 09-09-98	AvRes T (mS) 09-24-98	AvRes T (mS) 10-01-98
AADR01	10.190.56.1	469.1	852.4	461.1	873.2
ABNR01	10.190.52.1	486.1	869.2	489.5	880.2
APRR01	10.190.54.1	490.7	883.4	485.2	892.5
ASAR01	10.196.170.1	619.6	912.3	613.5	902.2
ASRR01	10.196.178.1	667.7	976.4	655.5	948.6
ASYR01S					503.4
AZWR01	10.177.32.1	460.1		444.7	
BEJR01	10.195.18.1	1023.7	1064.6	1184	1021.9

Presentation_ID © 1998, Cisco Systems, Inc.

www.cisco.com

51

Capacity and Performance Tools

Presentation_ID © 1998, Cisco Systems, Inc.

www.cisco.com

52

Three Essential Tools

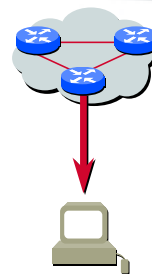
- **S**imple **N**etwork **M**anagement **P**rotocol (SNMP)
- **R**emote **M**ONitoring Protocol (RMON)
- **N**etFlow - Flow Based TCP/IP Analysis

Presentation_ID © 1998, Cisco Systems, Inc.

www.cisco.com

53

Traffic Management Elements



Data Collection

Data Export

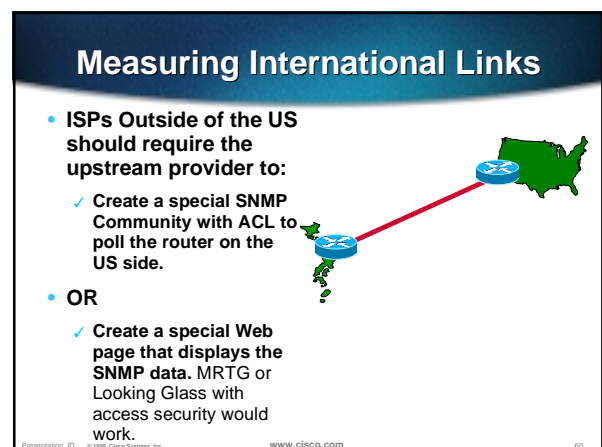
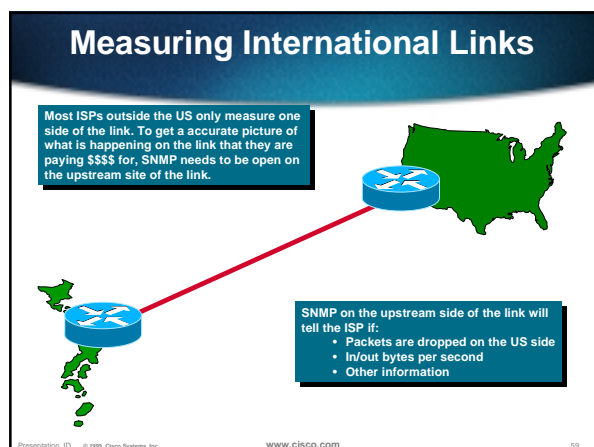
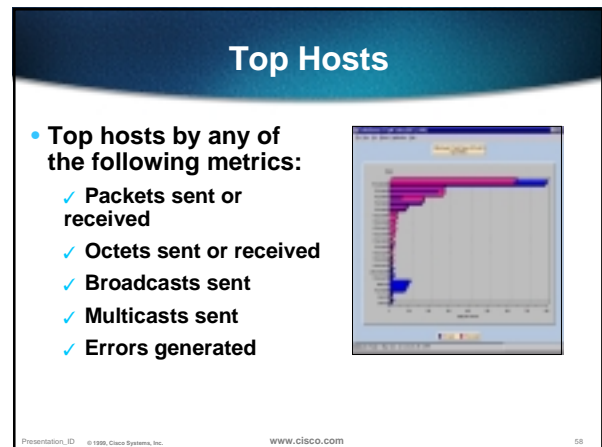
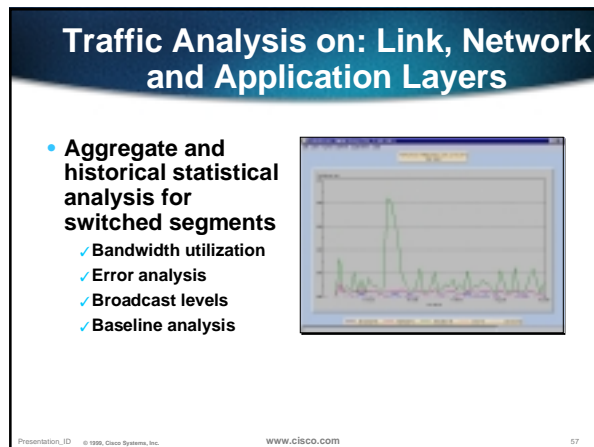
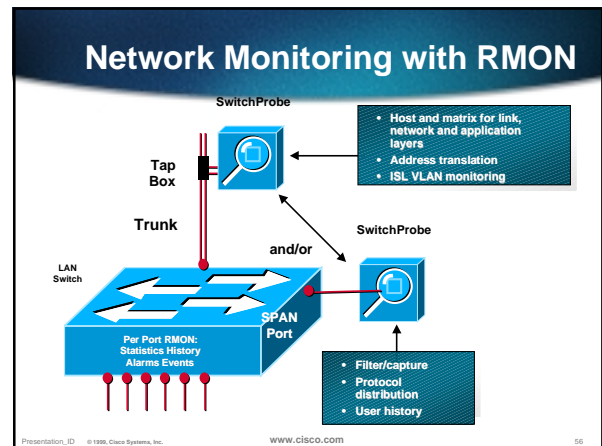
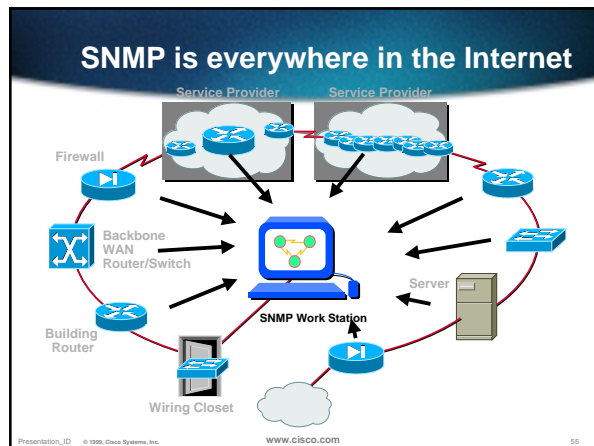
Data Analysis
and
Visualization

- Data collection mechanisms on network equipment
- Data export mechanisms to applications
- Data analysis and visualization

Presentation_ID © 1998, Cisco Systems, Inc.

www.cisco.com

54



Measuring International Links



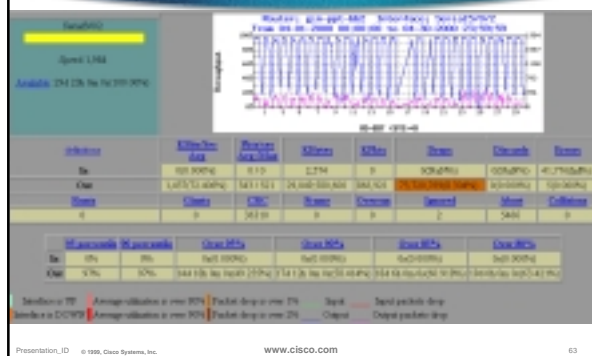
Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com 61

Measuring International Links



Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com 62

Measuring International Links



Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com 63

Measuring International Links



Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com 64

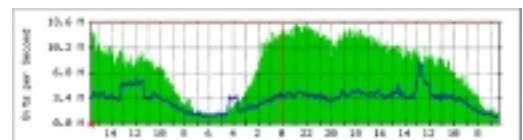
Fundamental Tools are Cheap!

- The tools to create a simple network manage system that will give an ISP the basics comprise of the following:
 - ✓ PC with LINUX (free UNIX)
 - ✓ CMU SNMP (free SNMP)
 - ✓ PERL5 (free UNIX script language)
 - ✓ GNU Plot (free graphic plot tool)
 - ✓ Printer

Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com 65

Fundamental Tools are Cheap!

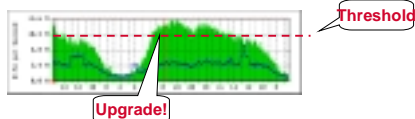
- Example of what can be done with another Shareware tool - MRTG



Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com 66

Fundamental Tools allow for the Baseline!

- Baseline Quality Levels are critical for any ISP Server.
 - Average Utilization and Packet Loss need to be monitored on the entire network.
 - QoS Threshold need to be set and acted on to maintain any sort of foundation to build advanced IP services. This is *ISP 101* which most new ISPs forget!
 - All you need is SNMP! It's not rocket science.



Presentation_ID © 1998, Cisco Systems, Inc. www.cisco.com 67

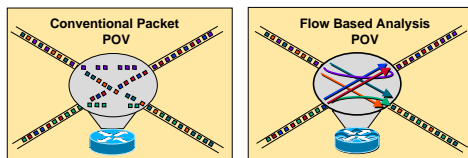
Limitations of SNMP and RMON

- SNMP and RMON will tell you what is happening on the network (i.e. load, PPS, packet drops).**
- SNMP and RMON will not tell you who is doing what to where and when.**
- For that sort of details, *TCP/IP Flow Based Analysis* is needed.**

Presentation_ID © 1998, Cisco Systems, Inc. www.cisco.com 68

Flow Based Analysis

**Connectionless IP But...
The Network Is Full of Flows (Conversations)**



- Flows are Unidirectional
- Flows are Granular
 - IP address and app. port# pairs
 - (TOS/Protocol/Interface)

Presentation_ID © 1998, Cisco Systems, Inc. www.cisco.com 69

Flow Based Analysis

- Key IETF work:**
 - Real Time Traffic Flow (RTFM) working group
 - <http://www.auckland.ac.nz/net/Internet/rtfm/>
 - IP Provider Metrics (IPPM)
- Public Domain and Commercial Tools now available.**

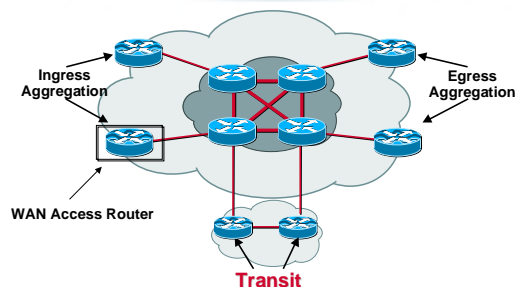
Presentation_ID © 1998, Cisco Systems, Inc. www.cisco.com 70

Flow Based Analysis

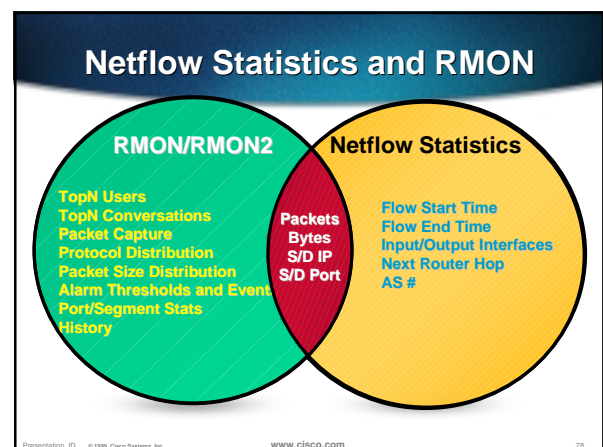
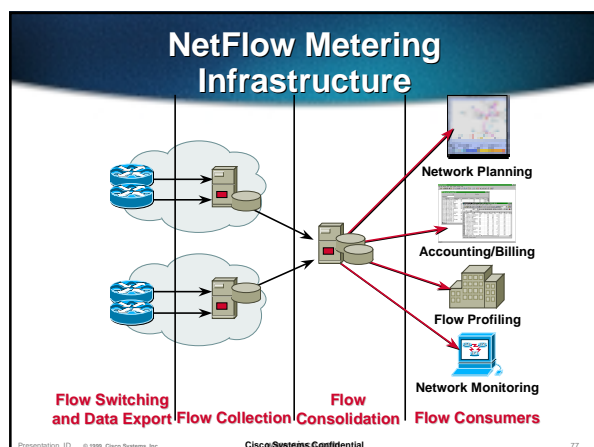
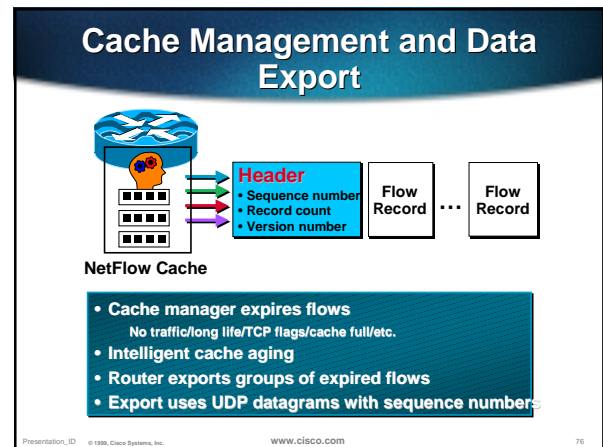
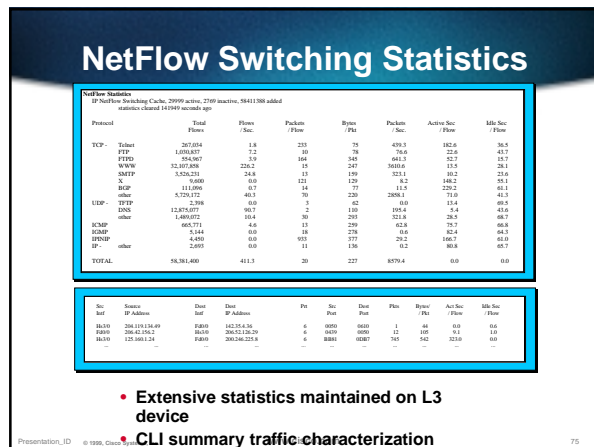
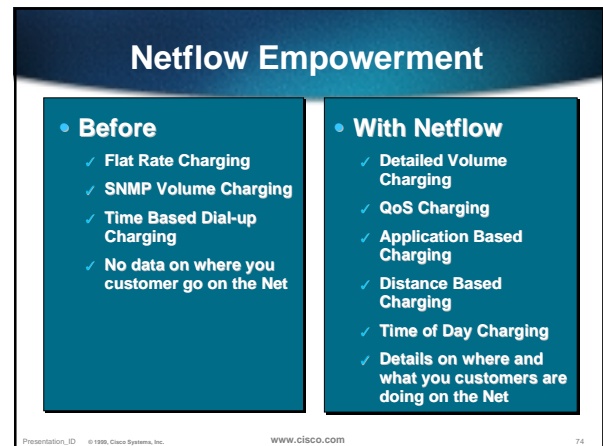
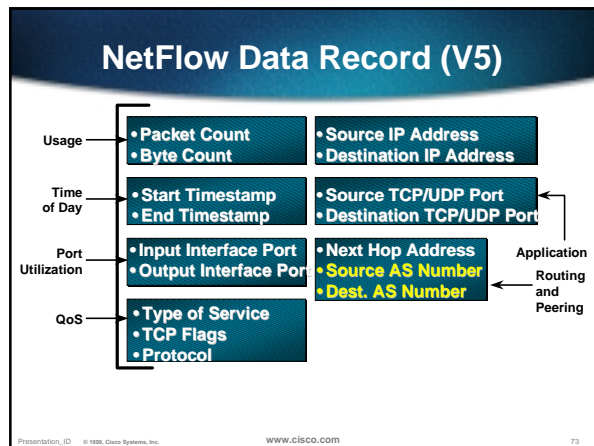
- Key Tools Used Today**
 - NetScarf** - Global and Regional Internet Analysis project (<http://www.merit.edu/~netscarf>)
 - Traffic Flow Analysis** - NLANR (<http://www.nlanr.net>)
 - NetFlow** - Analysis and IP switching technology build into Cisco's IOS.
 - NeTraMet** - Free Flow analysis software. (<http://www.auckland.ac.nz/net/Accounting/ntm.Release.note.html>)
 - Many new Netflow based commercial tools

Presentation_ID © 1998, Cisco Systems, Inc. www.cisco.com 71

Flow Activation Locations



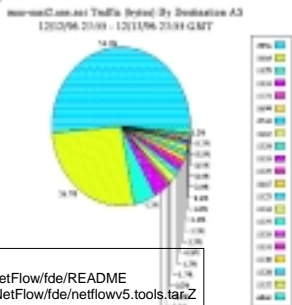
Presentation_ID © 1998, Cisco Systems, Inc. www.cisco.com 72



NetFlow Provides Open Interfaces

Cflowd
by ANS & BBN
and maintained by
CAIDA

[Http://www.caida.org](http://www.caida.org)
<ftp://ftp-eng.cisco.com/ftp/NetFlow/de/README>
<ftp://ftp-eng.cisco.com/ftp/NetFlow/de/netflowv5.tools.tar.Z>

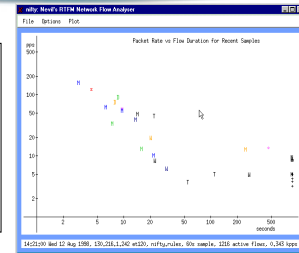


Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com 79

NetFlow Provides Open Interfaces

NetFlowMet
by Nevil Brownlee

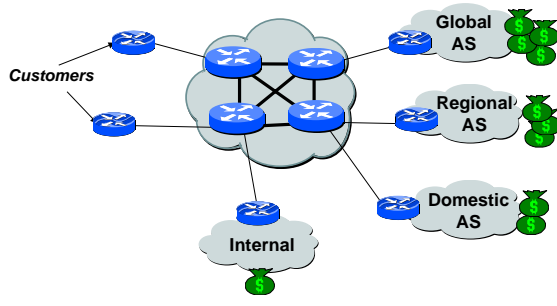
Uses the work from
the IETF's Realtime
Traffic Flow
Measurement
(RTFM) WG



<http://www.auckland.ac.nz/net/Accounting/>
<http://www.auckland.ac.nz/net/Internet/rtfm/>

Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com 80

NetFlow Distance-Based Accounting



Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com 81

NetFlow Distance-Based Accounting

- General Information page for Cisco Netflow services:
<http://www.cisco.com/warp/public/732/netflow/>
- Cisco's NetFlow FlowCollector v2.0 and NetFlow FlowAnalyzer v2.0:
http://www.cisco.com/warp/public/732/netflow/netan_o_v.htm
- Case Study - How to implement Netflow in a network. *Traffic Accounting Using Netflow and Cflowd* by Roberto Sabatino (DANTE/TEN-34)
<http://www.dante.net/pubs/dip/32/32.html>

Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com 82

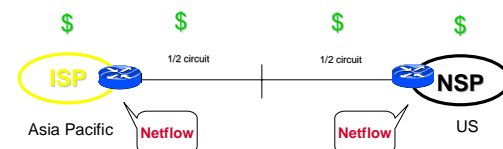
NetFlow Distance-Based Accounting

- 3rd Party Solutions:
 - Belle Systems <http://www.belle.dk>
 - Solect <http://www.solect.com>
 - XACCT Technologies <http://www.xacct.com>
 - Apogee Networks, Inc. <http://www.Apogeenet.com>
 - RODOPI <http://www.rodopi.com>

Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com 83

Netflow as a tool

Netflow statistics empowers all
ISPs with the ability to know the
who, what, where, and how
much.



Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com 84

Conclusions

- Aggressive measurement and analysis is critical to an ISPs and the Internet's survival.
- Not enough measurement and actual data analysis is taking place on the Internet. Too many people are speculating with weak data to back up their claims.

Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com

85

Capacity and Performance Best Practices

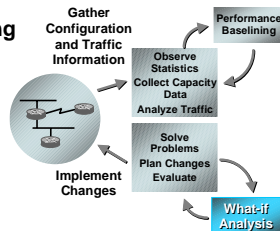
Presentation_ID © 1999, Cisco Systems, Inc.

www.cisco.com

86

What-If Analysis

- Simulation applications
- Lab application modeling
 - ✓ Protocol analyzer, WAN emulator, packet generator, NETSYS performance analyzer
- Lab network modeling
 - ✓ NVS/NVT, lab network modeling



Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com

87

Service Level Management

- Define performance requirements
- Define Upgrade criteria by capacity area
- Measure capacity and performance
- Review thresholds and baseline
- Take action!

Presentation_ID © 1999, Cisco Systems, Inc.

www.cisco.com

88

Service Level Management

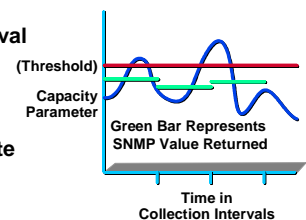
Threshold	WAN	LAN
CPU	75-90%	75-90%
Link	80-90%	40-90%
Memory	50%	50%
Output Queue	200	25
Buffer Misses	Any	Any
Broadcast Vol	10/Sec	300/Sec
FECN/BECN	10/Sec	N/A

Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com

89

Peak and Average Utilization

- Solution to narrow collection interval
- Low collection interval = high overhead
- Recommend ≥ 5 minutes
- Peak values not quite what they seem
- Close to threshold indicates likely exceed condition



Presentation_ID © 1999, Cisco Systems, Inc.

www.cisco.com

90

Capacity Exception Management

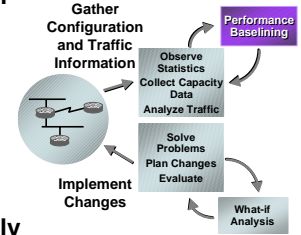
- Alarm critical capacity thresholds (CPU, critical link)
- Develop notification, escalation and action plan for threshold violations
- Take action!

Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com

91

Performance and Capacity Baselineing

- Interface utilization
- Device CPU, memory, buffer, I/O utilization
- Network overhead
- Raw performance characteristics
- Monthly or quarterly baseline report



Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com

92

Upgrade Planning

- Understand lead times for circuits, equipment, planning and design
- upgrade criteria based on service level management

Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com

93

QoS Management

- Prioritize applications by business impact
- Understand networked application behavior (packet size, timeouts, flows, bandwidth requirements)
- Develop QoS management plan

Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com

94

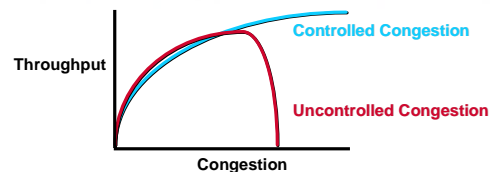
Managing TCP Traffic

Moving Mountains of Data
Without Incurring the World
Wide Wait

319 1096, 05/19_12 © 1999, Cisco Systems, Inc. www.cisco.com

95

Problems with Congestion



- Uncontrolled congestion will seriously degrade system performance
 - ✓ The system buffers fill up
 - ✓ Packets are dropped, resulting in retransmissions
 - ✓ This causes more packet loss and increased latency
 - ✓ The problem builds on itself until the system collapses

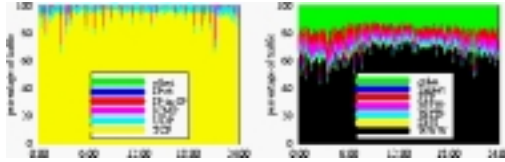
Presentation_ID © 1999, Cisco Systems, Inc. www.cisco.com

96

Backbone Traffic Mix

Transport Breakout

TCP Applications



Source: MC/NSF OC-3MON via <http://www.nlanr.net>, 1998

Presentation_ID © 1998, Cisco Systems, Inc.

www.cisco.com

97

TCP Technology Issues

- Single drops communicate from network to sending host
✓ “You need to slow down”
- Multiple drops in round trip trigger time-outs
✓ “Something bad happened out here”

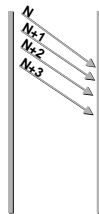
Presentation_ID © 1998, Cisco Systems, Inc.

www.cisco.com

98

Behavior of a TCP Sender

- Sends as much as credit allows
- Starts credit small
✓ Avoid overloading network queues
- Increases credit exponentially
✓ To gauge network capability



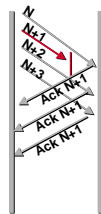
Presentation_ID © 1998, Cisco Systems, Inc.

www.cisco.com

99

Behavior of a TCP Receiver

- When in receipt of “next message,” schedules an ACK
- When in receipt of something else, acknowledges all it can immediately



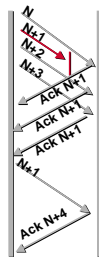
Presentation_ID © 1998, Cisco Systems, Inc.

www.cisco.com

100

Sender Response to ACK

- If ACK acknowledges something
✓ Update credit and send
- If not, presume it indicates a lost packet
✓ Send first unacknowledged message right away
✓ Halve current credit
✓ Increase linearly to gauge network throughput



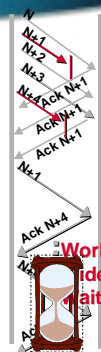
Presentation_ID © 1998, Cisco Systems, Inc.

www.cisco.com

101

Multiple Drops in TCP

- In the event of multiple drops within the same session:
✓ Current TCPs wait for time-out
✓ Selective acknowledge may work around (but see INFOCOM '98)
✓ New Reno “fast retransmit phase” takes several RTTs to recover



Presentation_ID © 1998, Cisco Systems, Inc.

www.cisco.com

102

How Can We Make TCP in a Network Act Predictably?

- Predictable amount of traffic in the network:
 - ✓ Well-written TCP implementations manage their rates to the available bandwidth
- Router needs to
 - ✓ Provide predictable treatment of packets
 - ✓ Queue delay and drop characteristics

Presentation_ID © 1998, Cisco Systems, Inc.

www.cisco.com

103

Fundamental FIFO Queue Management Technologies

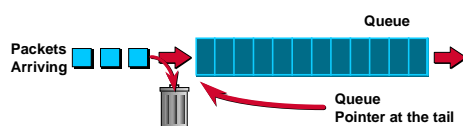
- Tail drop
 - ✓ Network standard behavior
 - ✓ Causes session synchronization when waves of traffic experience correlated drops
- Random Early Detection (RED)
 - ✓ Random drops used to desynchronize TCP sessions and control rates

Presentation_ID © 1998, Cisco Systems, Inc.

www.cisco.com

104

Tail Drop



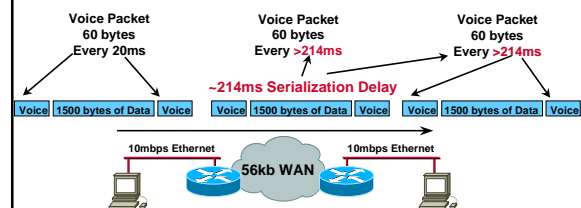
- Without RED, when the queue fills up all packets that arrive are dropped—**Tail drop**

Presentation_ID © 1998, Cisco Systems, Inc.

www.cisco.com

105

Large Packets “Freeze Out” Voice



- Large packets can cause playback buffer underrun, tail drops
- Jitter or playback buffer can accommodate some delay/delay variation

Presentation_ID © 1998, Cisco Systems, Inc.

www.cisco.com

106

Session Synchronization

- Session synchronization results from synchronized losses
- Tail drop from waves of traffic synchronizes losses

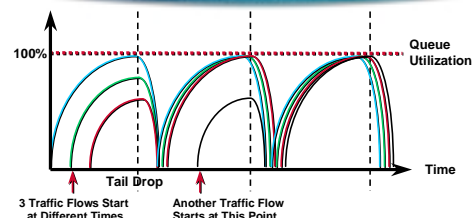


Presentation_ID © 1998, Cisco Systems, Inc.

www.cisco.com

107

Session Synchronization



- TCP is a sliding window protocol that uses **self-clocking** to adjust its use of the network to match available bandwidth using slow-start and congestion avoidance algorithm
- Session synchronization is when many TCP connections go through TCP Slow-Start mode at the same time

Presentation_ID © 1998, Cisco Systems, Inc.

www.cisco.com

108

Random Early Detection (RED)



- With RED, as oppose to doing a tail drop, the router monitors the **average queue size** and using randomization it chooses connections to notify that a congestion is impending

Note: **Avg. queue size** is not an

109

Effect of Random Early Detection



Courtesy of Sean Doran, Ebone

RED Enabled

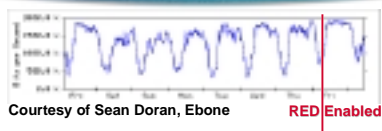
- One day, below 100% throughput
 - ✓ Simple FIFO with tail drop
- Starting 10:00 second day, 100% throughput
 - ✓ Random Early Detection enabled

Presentation_ID © 1999, Cisco Systems, Inc.

www.cisco.com

110

Was that a Fluke?



Courtesy of Sean Doran, Ebone

RED Enabled

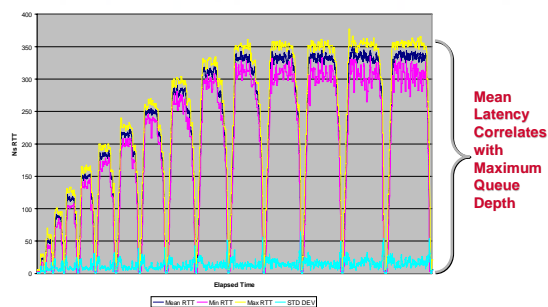
- No, here's what happened that week...
- Session synchronization reduced throughput until RED enabled

Presentation_ID © 1999, Cisco Systems, Inc.

www.cisco.com

111

FIFO Traffic Timings

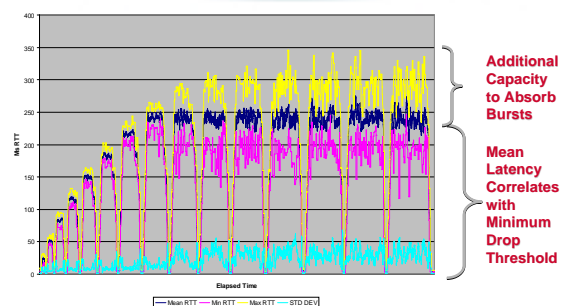


Presentation_ID © 1999, Cisco Systems, Inc.

www.cisco.com

112

RED Traffic Timings

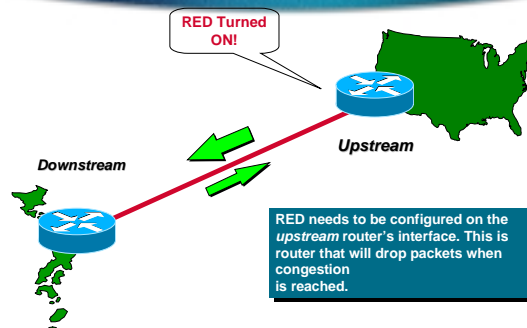


Presentation_ID © 1999, Cisco Systems, Inc.

www.cisco.com

113

Where to apply RED?



Presentation_ID © 1999, Cisco Systems, Inc.

www.cisco.com

114

Applying RED/WRED

- Enabling WRED
 - ✓ [no] random-detect <weight-constant>
 - ✓ weight-constant = <1-16> is an integer used in weighted average to mean $2^{\text{weight-constant}}$. 10 is the default.
- Tuning weight constant affects loss rate
 - ✓ rule-of-thumb:
 - ✓ DS-3/OC-3 Links: Value of 10 might achieve $\sim 10^{-4}$ drop rate, recommended for DS-3/OC-3 link.
 - ✓ T1/E1 Links: Value of 7 might achieve a loss rate around 10^{-3} .
 - ✓ Actual recommended value should be determined in real operational network.

Presentation_ID © 1998, Cisco Systems, Inc. www.cisco.com

115

Therefore—TCP QoS Definition:

- Normally at most one drop per round trip
- Mean variation in latency bounded by predictable network

Presentation_ID © 1998, Cisco Systems, Inc.

www.cisco.com

116

TCP Flow Statistics

- >90% of sessions have ten packets each way or less
 - ✓ Transaction mode (mail, small web page)
- >80% of all TCP traffic results from <10% of the sessions, in high rate bursts
 - ✓ It is these that we worry about managing

Presentation_ID © 1998, Cisco Systems, Inc. www.cisco.com

117

An Interesting Common Fallacy about RED:

- “RED means you will have more drops”
 - ✓ Statement derives from observed statistics
- RED means that you will have
 - ✓ Closer to 100% utilization of your line
 - ✓ Less average delay per packet
- But queuing theory?
 - ✓ As a line approaches 100% utilization, drops will increase, **even though served load increases**

Presentation_ID © 1998, Cisco Systems, Inc.

www.cisco.com

118

TCP Traffic Management Issues

- Applications
 - ✓ Often have site-specific policy associated with them
 - ✓ Traffic often identifiable by port numbers
- Sites
 - ✓ Generally identifiable by address prefix or interface traffic is received on

Presentation_ID © 1998, Cisco Systems, Inc. www.cisco.com

119

TCP Bandwidth Policy Questions to Answer

- Particular site or application wants at **least** a certain bandwidth
- Particular site or application wants at **most** a certain bandwidth
- Particular site or application wants to **average** about a certain bandwidth

Presentation_ID © 1998, Cisco Systems, Inc.

www.cisco.com

120

What to look for in an Upstream Provider

600

1042_OSPF9_c2

© 1999, Cisco Systems, Inc.

www.cisco.com

121

Preparation

- List all potential providers
- Get Maps of oceanic cable systems
- Get Maps of satellite foot prints.

Presentation_ID

© 1999, Cisco Systems, Inc.

www.cisco.com

122

What to ask from the prospective providers?

- Network Maps with landing/termination points of your links.
- List of IXPs and Private Peers
- URLs of NOC Pages
- Do they lease routers and/or co-locations space?
- Do they have upstream caches?

Presentation_ID

© 1999, Cisco Systems, Inc.

www.cisco.com

123

Example - Network Map



What to require from the Upstream Provider

- Statistics Page and Weekly Reports
- 24x7 NOC Contacts
- RED or WRED on their router's interface
- CAR ICMP Rates Limits for DoS Protection
- Back-up contingencies in writing.

Presentation_ID

© 1999, Cisco Systems, Inc.

www.cisco.com

125

CDN Technology on Trans Oceanic Links

600

1042_OSPF9_c2

© 1999, Cisco Systems, Inc.

www.cisco.com

126

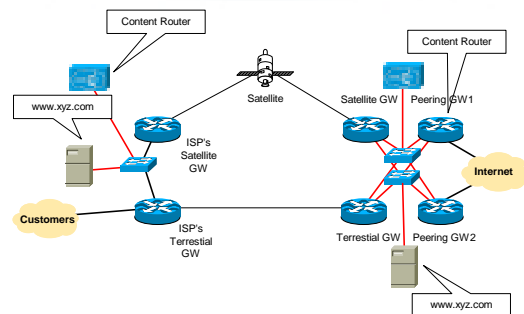
Content Routing

- Many International ISPs have Co-location Services in the US for their domestic customers.
 - Allows local customers to have their websites bypass congested links - allowing people outside the country to access the site.
 - The problem is that consumers in country then have to get through the congested link.

Presentation_ID © 1998, Cisco Systems, Inc. www.cisco.com

127

Content Routing over Trans-Oceanic Links

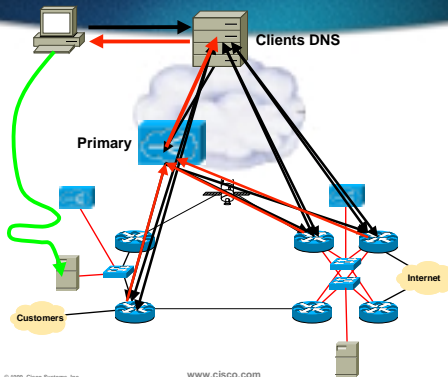


Presentation_ID © 1998, Cisco Systems, Inc. www.cisco.com

www.cisco.com

128

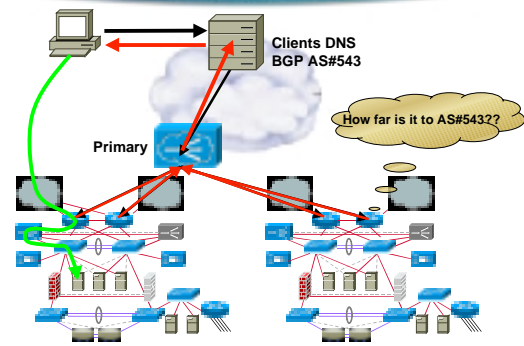
RTT Metric



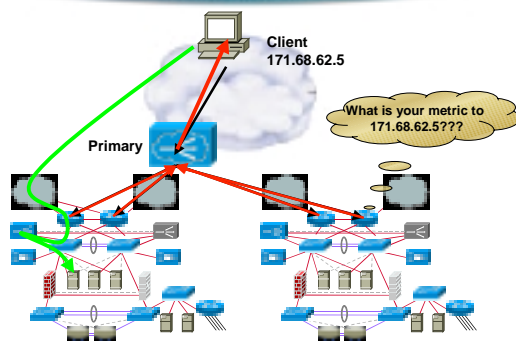
Presentation_ID © 1998, Cisco Systems, Inc. www.cisco.com

129

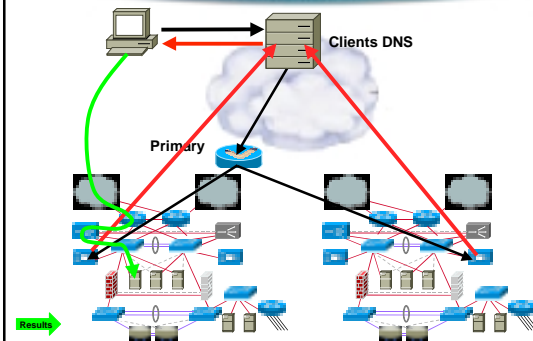
BGP-External Metric



IGP-Internal Metric



Race Condition (Boomerang)



Boomerang Technology Testing

VERIO™



- Cisco Needed to Prove the Boomerang Technology in a “real-world” environment
- Cisco chose Verio in six collocation facilities
- Solid network, geographic dispersion.
- Cisco Powered Network

Presentation_ID © 1998, Cisco Systems, Inc.

www.cisco.com

133

CISCO SYSTEMS



EMPOWERING THE
INTERNET GENERATIONSM

Presentation_ID © 1998, Cisco Systems, Inc.

www.cisco.com

134