Deploying Services over an Optical Internetworking Infrastructure

Session 3005
Agenda

- Introduction
- Target Architecture
- Deploying IP Services Over Optical Infrastructure
- The need for QoS in IP Networks
- MPLS in IP Networks
Service Provider’s Paradox: A New Business Model

• Most inter-city links are 2 x 2.5 Gbps today with low average utilization
• Backbone bandwidth doubling, for most providers, every 6 to 9 months

"By 2010, 80% of service provider profits will come from data services." Tom Nolle, BCR 7/99

Data

Voice

Content Increasing in Value

Source: Datamonitor, 1998

Optical Market Projections

Global Optical Network Markets, 1999-2004

Source: Pioneer 1999
IP Core Devices Market Forecast

$ Billions

1999 2000 2001 2002 2003

Aggressive  Conservative

Target Architecture

The Challenges...
What’s the Problem?

So the problem we must solve is: API to API service creation wherever that API resides, i.e. Laptop, PC, mobile, palmtop, phone, etc. The network must support IP applications with the Full IP network service portfolio on an end-to-end (API-to-API) basis.

Complexity’s an Issue

Many Flows
Simplifying Technology

Scalability

Services

Network Services
- IP QoS
- IP Multicasting
- IP Any-2-Any
- IP VPN
- IP Security
- IP Traffic
- IP Engineering

MPLS

Transport Services

Time to Revenue

OSS and BSS Costs

The Optical Internetworking

Traditional Model

IP
ATM/FR
SONET/SDH
Optical

Optical Internetworking

IP
Optical

- Lower equipment cost
- Lower operational cost
- Simplified architecture
- Scalable capacity
Target Architecture

Realization

Target Content Architecture

Content

Audio and Video

Application SP

Content Aware Hosting

Network Services

Core Operators

Single Fabric IP Service Core

Access Operators

Mobile Voice

Fixed Voice

ISDN

DSL

Cable

ATM

F/RSDH

ISP

Fixed Radio

Consumers and Local Delivery Mechanism

UM
Target Network Architecture

Connectivity

Hosted Services

Product Services

Network Services

Internet Backbone

DSL

HFC

Centrex

ISDN

PSTN

Voice Signalling VPN

TCP/UDP/ATM

Data VPN

IP VPN (ATM/FR/PDH/SDH)

MPLS QoS Core

MPLS QoS

Text

Audio

Video

IN

SCP

UM

TV

Radio

Content and Applications

Trusted Backbone

Other Licensed Operators

LMDS/MMDS

WLL (LMDS, MMDS)

Direct Access

Leased Lines (SDH)

X.25

Frame Relay

ATM

Managed CPE

Broadband Access

Persistent Connection

a DSL (Cell or Frame)

Cable

WLL (LMDS, MMDS)

Dialup

PSTN

ISDN

GSM (V110, GPRS)

Core Functionality

MPLS “Edge”

Intra-PoP Connectivity

Backbone

(STM1, STM4, STM16...)

IP and MPLS

POS or DWDM

Provider “Edge”

Content and Caches

UCP

DNS, AAA

Mgmt

MPLS “Edge”

Provider “Edge”
**Target Backbone Architecture**

**PoS Packet-over-SONET/SDH**
- Runs over dark fiber, SONET, or WDM
- Enables transport “mix and match”
- Provides efficient evolution path for incumbents
- Provides optimized transport for greenfield builds
- Standards based

**IP Services: Essentials**

- Open standards
- Scalability
- Reliability
- Manageability
- Low latency
- Network services
  - DNS, DHCP, hosting, caching, AAA, number translation, SCP, firewall, NAT, etc...

- IP QoS
- IP traffic Eng
- IP VPN
- IP any-2-any
- IP multicast
- IP security
Deploying IP Services over Optical Infrastructure

Services: The Challenge

- The optical IP network needs to meet the requirements of the applications
- Different services have different requirements:
  - Voice, videoconferencing, distance learning-low latency, low jitter
  - Bulk data transfer-high bandwidth
  - Email-no demanding requirements
To ensure service quality, each application or flow, needs to be differentiated from one another.

Non time critical applications, e.g. www, could impact time sensitive applications, such as voice or video.

IP will be the predominant traffic on wide area networks, optimized over an optical infrastructure.

Today’s networks are built around the requirements from legacy voice transport.

Tomorrows networks use IP as transport; voice, will be a relatively small amount of the overall traffic.
Mapping Services to Features

<table>
<thead>
<tr>
<th>Services</th>
<th>Enablers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Virtual Private Networks</td>
<td>MPLS VPNs, BW Guarantees</td>
</tr>
<tr>
<td>Bandwidth on Demand</td>
<td>Optical Integration with WLR</td>
</tr>
<tr>
<td>Content and Video Distribution</td>
<td>IP Multicast and CoS</td>
</tr>
<tr>
<td>Packet Circuit Emulation</td>
<td>MPLS Traffic Engineering</td>
</tr>
<tr>
<td>Real Time Services: Voice</td>
<td>Fast Restoration and CoS</td>
</tr>
<tr>
<td>Migration of Legacy Services</td>
<td>FR/AAL5 over MPLS</td>
</tr>
<tr>
<td>Tiered Service Offerings</td>
<td>CoS, MPLS Traffic Eng...</td>
</tr>
</tbody>
</table>

The Need for QoS in IP Networks

Consider Voice...
Traditional Voice Networks

SONET/SDH Rings and or ATM Network Provides Low Delay, Low Jitter, and Protection, (BLSR/MSP Ring)

Voice Quality: Compression Standards

Unacceptable

Utility

Business

Toll

Bandwidth (kbps)

Quality

PCM

ADPCM 32 (G.723)

LDCELP 16 (G.728)

CS-ACELP 8 (G.729)

ADPCM 16 (G.726)

LPC 4.8 (G.726)

LD CELP 16 (G.728)

MPMLQ (G.723.1)
Why Is it So Important to Minimise Delay?

Usability of Voice Circuit as a Function of End-to-End Delay

To Hit the Target:
- Eliminate PC delay
- Lower network latency
- Tighten network jitter

180 ms End to End

Delay Considerations in IP Networks

- CODEC
- Packetization
- Output queuing
- Access (up) link transmission
- Backbone network transmission
- Access (down) link transmission
- Input queuing
- Jitter buffer
- CODEC
Calculating a Delay Budget (G.729)

<table>
<thead>
<tr>
<th>Component</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Encoder/decoder delay (algorithmic plus processing and VAD)</td>
<td>30 ms</td>
</tr>
<tr>
<td>Waiting and framing (10 ms frames)</td>
<td>10 ms</td>
</tr>
<tr>
<td>Move to output queue</td>
<td>negligible</td>
</tr>
<tr>
<td>Wait in queue (depends on queuing and congestion)</td>
<td>0–8 ms</td>
</tr>
<tr>
<td>Access up link - backbone - down link</td>
<td>Variable</td>
</tr>
<tr>
<td>(1 ms per 100 Miles)</td>
<td></td>
</tr>
<tr>
<td>Move from input queue to DSP</td>
<td>Negligible</td>
</tr>
<tr>
<td>Jitter buffer</td>
<td>4 ms–40 ms</td>
</tr>
<tr>
<td>Coder process delay</td>
<td></td>
</tr>
</tbody>
</table>

*Total: (excluding up link, backbone, and down link)* 45 ms

What Is Referred to as QoS?

- The following parameters are usually equated with QoS:
  - Bandwidth guarantees
  - Admission control
  - Delay
  - Jitter
QoS with IP Class of Service

- Implemented at network Layer 3
- Media independent
- Efficient use of available bandwidth through statistical multiplex
- Bandwidth guarantees, delay, and jitter limits via advanced queuing mechanisms
- Admission control only for applications that need certain guarantees (gatekeeper)
- Connectionless—no TDM

Changing QoS Requirements

- Changing applications change requirements
- Classic voice requires dedicated 64k channel in TDM network
- Voice over IP designed to work on IP transport networks, can tolerate delay and jitter within limits
- IP transport is inherently statistically multiplexed, but can limit delay and jitter to suit VoIP applications
VoIP in the Intranet

IP QoS Technologies

• Two fundamental different approaches

• Device oriented
  Each device uses own criteria to prioritize traffic

• Network wide
  Packets carry CoS information in their header-IP precedence, diffserv
Device Oriented IP QoS

- Each device in the network has its own set of rules to classify and forward traffic
- Often flow-based, very granular
- Difficult to manage in large networks
- Most solutions add more devices to the network
- No Network wide intelligence

Network Wide IP QoS

- Packets are marked at the ingress or by the application and carry their classification throughout the network
- Core devices use this information to provide required services
- Easier to manage
IP Optical Internetworking
Congestion Issues

• Routers by design can have congested links when an interface is offered more load than it can support for an extended period of time

• This can cause three things:
  - Packet loss due to output queue overflow—some traffic might be sensitive to it
  - Added and varying delay—real-time applications impacted
  - Bandwidth for some important data is insufficient

Dealing with IP Congestion

• Impact can be minimized the following way:
  - Drop less important traffic before high priority traffic
  - Handle delay sensitive traffic different than bulk data
  - Ensure bandwidth for specific traffic is managed
Edge Functions

- Packet classification
  Precedence setting with CAR
- Bandwidth management
  Rate limiting with CAR
  Traffic shaping
- L3 metering
  NetFlow data export

Backbone Functions

- High-speed switching and transport
  Distributed switching (CEF)
- QoS enforcement
  Congestion avoidance (WRED)
  Congestion management (MDRR)
- QoS interworking
  IP-ATM QoS interworking (VC per CoS)
GSR Class of Service Features

- Network approach using IP precedence
- Committed Access Rate (CAR)
- Weighted Random Early Detection (WRED)
- Modified Deficit Round Robin (MDRR)

Committed Access Rate Rate Limiting and Classification

- ‘Rule Based’ engine
- CoS packet classification (Set-ToS) based on flexible rule
  - IP precedence/IP access list/incoming interface/MAC address
- Rate limiting functionality
- Generally deployed at the network edge
CAR Policy Examples

Drop

Re-Color

Per Application CAR

Multimedia

Mission-Critical

CAR Action Policies

• Configurable actions
  Transmit
  Drop
  Continue (go to the next rate-limit in the list)
  Set precedence and transmit (rewrite the IP precedence bits and transmit)
  Set precedence and continue (rewrite the IP precedence bits and go to the next rate-limit in the list)

• Cascadable rate-limit statements
Random Early Detection

Random Early Detection (RED)

Discard Test Based on: Average Queue Depth

Drop Probability

Min
Max

Average Queue Depth

1/mpd

1

0
Weighted RED (WRED)

- ToS: Incoming IP Packet
- IP Precedence: 7, 5, 0
- Queue
- Drop Probability
- Average Queue Depth (Red)

Modified Deficit Round Robin (MDRR)

- Alternate Priority
- LLQ—Low Latency Queue
- Strict Priority
MDRR Operation
Alternate Priority

Weight

1500  500

Deficit Counter (bytes)

1500 + 512x0 = 1500

1500 + 512*1 = 2012

1500 + 512*2 = 2524

1500 - 500 - 1500 = 500

MDRR Operation
Alternate Priority

Weight

1500  500

Deficit Counter (bytes)

1500 - 500 - 1500 = 500

2012

2524
MDRR Operation Alternate Priority

Weight | Deficit Counter
--- | ---
1 Q1(LLQ) | -500
2 Q2 | 2012-1500-1500= -988
3 Q3 | 2524

Weight | Deficit Counter
--- | ---
1 Q1(LLQ) | -500+1500-500= 500
2 Q2 | -988
3 Q3 | 2524
MDRR Operation
Alternate Priority

Weight 1500 500

1 Q1(LLQ) 0
2 Q2 -988
3 Q3 2524-500-1500-500= 24

L3 CAR
Apply Ingress Rate Thresholds
Determine Packet Class
Administer Packet Class
Apply Egress Rate Thresholds
Traffic Metering

Summary of Network IP QoS
Network Level Protection

- **Layer one**
  - APS/MSP

- **Layer two**
  - Intelligent Protection Switching, (IPS), using DPT rings

- **Layer three**
  - OSPF load balancing
  - MPLS, traffic engineering and fast reroute

APS/MSP

- Available on all POS line cards
- Bellcore GR-253 compliant
- K1/K2 link-layer control
- Three levels of protection
  - Port failure
  - Line card failure
  - Router failure
Intelligent Protection Switching

- SRP IPS
  - Automatic or manual initiation
  - Wraparound event
  - Signaling via SRP control messages
  - Multilevel event hierarchy
  - Intermediate node transit
  - <50-ms goal for large ring

TE: Fast Re-Route

Uses MPLS TE Tunnels to Protect Failed Link
We Have Seen that Different Services Have Different Requirements; (Delay, Jitter Etc.) Using the Tools Available, CAR, WRED, MDRR, Different Traffic Flows Can be Effectively Marked, and Differentiated From Each Other, this Enables Service Providers to Offer Billable, Robust, Differentiated IP Services to Customers.
Traffic Engineering with MPLS

- Why traffic engineer?
  - Optimize link utilization
  - Specific paths by customer or class
  - Balance traffic load
- Traffic follows pre-specified path
- Path differs from normally routed path
- Controls packet flows across a L2 or L3 network infrastructure
MPLS: Forwarding

1a. Existing Routing Protocols (e.g. OSPF, IGRP) Establish Reachability to Destination Networks

1b. Label Distribution Protocol (LDP) Establishes Label to Destination Network Mappings

2. 'Labels' are VCS on Links: LDP Configures Label Swapping Operations at MPLS Switches by Setting Up VC Cross-connects

3. Ingress MPLS Edge Router Receives Packet, Performs Layer 3 Value-added Services, and “Label” Packets

4. MPLS Routers Switch Labeled Packets Using Label Swapping

5. MPLS Edge Router at Egress Removes Label and Delivers Packet

MPLS VPN Services

- MPLS allows carriers to offer VPNs based on the peer model which enables them to bundle value added services with the VPN

Customer Sites Connected to Network with Frame Relay, ATM, XDSL, Etc.

Customer Sites Have Ordinary IP Equipment, Not MPLS or Any Special VPN Equipment
Cisco MPLS VPNs: Overview

- Customer sites run ordinary IP, no special protocols
- VPN service is based on standard mechanisms:
  - BGP for distributing VPN routing information
  - MPLS for forwarding packets

VPN-IP Address Format

VPN-ID is a 64-bit customer identifier
Never carried on packets, only in label tables

Each customer network can use
Registered IP addresses
Illegal, unregistered addresses
Private addresses (RFC 1918), e.g. 10.x.x.x
Learning VPN Addresses

・MPLS edge routers and switches learn the addresses of attached sites
・Each site has a VPN identifier associated with it
・Ordinary IP routing protocols run to the customer sites

Distributing VPN Addresses

・Edge routers/switches communicate via BGP4 with multiprotocol extensions (RFC 2283)
・Edge routers switches get complete VPN routing information
・Distinct forwarding tables per VPN
Controlled Route Distribution

- Each customer site gets access to routing information for its VPN only
- Access is by RIP, BGP, maybe OSPF, or static configuration

Route Distribution: Key Points

- IP routing protocols are used to learn addresses from customer sites
- BGP, RIP, or static routing to customer sites is best
- A single core BGP session is used to support a large number of VPNs (1000s now, 100,000s or more later): very highly scalable
Forwarding: Summary

- Packets are forwarded using ordinary MPLS
- Each packet carries a label identifying the destination site: privacy equivalent to Frame Relay
- Network core knows nothing about VPNs: no need for per-VPN VCs

Core Functions in MPLS VPNs

- The core uses MPLS to forward packets across the network
- The network core knows nothing about VPNs: does not require separate VCs for separate VPNs
Traffic Separation in MPLS VPNs

Data Packet Two Labels

The Second-Level Label Identifies the Destination VPN and Customer Site

The Ingress MPLS Edge Switch or Router Puts Two Labels on Each Packet

- Each packet has a label identifying the destination VPN and customer site
- This provides the same level of privacy as Frame Relay

Summary

- Designing IP networks requires taking into account API to API
- An IP Optical Internetworking infrastructure allows:
  - Lower equipment cost
  - Lower operational cost
  - Simplified architecture
  - Design for low latency, low jitter
  - Differentiated Service delivery
- Features such as QoS, MPLS, and TE are service enablers
Questions?

Deploying Services Over an Optical Internetworking Infrastructure

Session 3005
Please Complete Your Evaluation Form

Session 3005
Backup Material

Configuring CAR

Rate-limit {Input | Output} [Access-group [Rate-limit] ACL-index | QoS-group Number] BPS Burst-normal Burst-max Conform-action Action Exceed-action Action

Example 1
Interface Serial1/0
  Rate-limit Input 20000000 24000 24000 Conform-action Transmit Exceed-action Drop
  Rate-limit Output 20000000 24000 24000 Conform-action Transmit Exceed-action Drop

Example 2
Interface Serial1/0
  Rate-limit Input Access-group 101 20000000 24000 32000 Conform-action Set-prec-transmit 5 Exceed-action Set-prec-transmit 0
  !
  Access-list 101 Permit Tcp Any Any Eq Www
Monitoring CoS Features CAR

GSR# show interfaces serial 1/0 rate-limit

Serial1/0
Input
matches: all traffic
params: 20000000 bps, 24000 limit, 24000 extended limit
conformed 8 packets, 428 bytes; action: transmit
exceeded 0 packets, 0 bytes; action: drop
last packet: 8680ms ago, current burst: 0 bytes
last cleared 00:03:59 ago, conformed 0 bps, exceeded 0 bps
Output
matches: all traffic
params: 20000000 bps, 24000 limit, 24000 extended limit
conformed 0 packets, 0 bytes; action: transmit
exceeded 0 packets, 0 bytes; action: drop
last packet: 8680ms ago, current burst: 0 bytes
last cleared 00:03:59 ago, conformed 0 bps, exceeded 0 bps

Configuring WRED

Label/Name for this Entire Group of WRED Parameters

Label for this set
Min Queue Threshold
Max Queue Threshold

Maps Precedence Values to Set of WRED Parameters to Use

Determines how closely Average follows instantaneous queue depth

(config)#cos-queue-group test-rx
(config-cos-que)#random-detect-label 0 25 100 1
(config-cos-que)#random-detect-label 1 50 125 1
(config-cos-que)#random-detect-label 2 75 150 1
(config-cos-que)#precedence 0 random-detect-label 0
(config-cos-que)#precedence 1 random-detect-label 1
(config-cos-que)#precedence 2 random-detect-label 2
(config-cos-que)#precedence 3 random-detect-label 2
(config-cos-que)#precedence 4 random-detect-label 2
(config-cos-que)#precedence 5 random-detect-label 2
(config-cos-que)#precedence 6 random-detect-label 2
(config-cos-que)#precedence 7 random-detect-label 2
(config-cos-que)#exponential-weighting-constant 2
Receive Side

Create global `slot-table-cos` to map cos-queue-groups to various destination slots as follows:

```
(config)# slot-table-cos test-table
(config-slot-cos)# destination-slot all test-rx
(config-slot-cos)# exit
```

Apply slot-table-cos to ingress slot as follows
```
(config)# rx-cos-slot 1 test-table
```

Transmit side
```
(config-if)# tx-cos test-rx
```

Configuring WRED

```
(config)#cos-queue-group test
(config-cos-que)# precedence 0 queue 0
(config-cos-que)# precedence 1 queue 1
(config-cos-que)# precedence 2 queue 2
(config-cos-que)# precedence 7 queue low-latency
(config-cos-que)# queue 0 50
(config-cos-que)# queue 1 100
(config-cos-que)# queue 2 150
(config-cos-que)# queue low-latency alternate-priority
(config-cos-que)#exit
```
Monitoring CoS Features

# Exec slot 1 show controller tofab queue
Displays breakdown of how buffer space allocated to queues

# Show interfaces pos 1/0 random-detect
Displays WRED drop statistics in the transmit direction

# Show cos
Displays WRED and MDRR parameters set for all configured cos-queue-group

# Show cos statistics
Displays WRED drop statistics