Deploying QPM in an Enterprise

Session 2606
Agenda

- QoS Background
- QPM Product Update
- Tuning for Oracle and PeopleSoft
- Implementation Methodology
- Customer Case Study
- Additional Resources

Lots of Applications

<table>
<thead>
<tr>
<th>VoIP</th>
<th>ERP</th>
<th>Multimedia</th>
<th>Web/URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Latency</td>
<td>Latency Tolerant</td>
<td>Low Bandwidth</td>
<td>Bursty Bandwidth</td>
</tr>
</tbody>
</table>

Different Demands on the Network
QPM Architecture

- Policy and configuration management via CLI and COPS
- DiffServ and RSVP QoS standards
- Directory-enabled
  - User-based policies
  - Export policies
  - DEN/CIM compliant
- CiscoWorks 2000 device import

QoS Policy Manager 1.1

- Policy-based QoS management
  - Data, voice and video applications
  - Layer 7 application and user-based policies
  - Policy wizards/templates with rules-based validation
- Comprehensive device support
  - Cisco IOS routers: 2500 thru 7500
  - Catalyst switches: 5500, 6500, 8500
  - Extensive QoS features: classification, queuing, shaping, congestion avoidance, VOIP, fragmentation, NBAR, RSVP, ...
- Network-wide QoS management
  - Automated distribution to devices
- Reliable deployment
  - CLI preview, job control and audit,...
QPM Functionality

**WAN-Edge QoS**
- Introducing QPM 1.0
- Broad Device/IOS Support
- Extensive QoS Features
- Application Profiles

**Nov 99**
- QoS for LAN-WAN
  - QPM 1.1
  - Catalyst QoS for LAN
  - New WAN QoS Features
  - Voice QoS
  - L7 Application Recognition

**NEW!**
- QoS Policy Mgmt Solution
- Provisioned QoS
- New COPS QoS
  - Standards-Based—COPS
  - User Policies
  - Directory Integration
  - Web Browser-Based

**April 99**
- Catalyst QoS for LAN
- New WAN QoS Features
- Voice QoS
- L7 Application Recognition

End-to-End Policy-Based QoS

- Traffic classification
- Consistent differentiated service
- QoS enforcement mechanisms
- Central policy administration
### Reasons for Tuning

- Different mission-critical apps behave differently
- Varied business objectives
- Cisco IOS, Cat OS and QPM are rich network-focused products, unlike h/w based competition
- QoS mechanisms and setting
- Cisco IOS revs and devices
- Varied impact on background traffic

### Proven QoS Enforcement Mechanisms

- **CAR**
  - Committed Access Rate
- **WFQ/CBWFQ**
  - Weighted Fair Queuing
  - dWFQ for VIP
- **IP RTP**
  - Custom/Priority Queuing for FR
- **WRR**
  - Weighted Round Robin
- **Congestion Control**
  - Congestion Avoidance
- **Admission Control**
- **Policing/Metering**
- **Classification/Markdown**
- **NBAR**
- **RSVP Configuration**
- **Traffic Shaping**
- **GTS**
  - FRTS Traffic Shaping
- **WRED**
  - dWRED for VIP
- **Weighted Random Early Detection**
- **FRF12**
  - Fragmentation
- **cRTP**
- **Traffic Shaping**
Extensive Device and Software Support

- **Campus devices**
  - Catalyst 8510, 5000, 6000, 8540, LocalDirector, RSM

- **WAN devices**
  - Cisco 2500, 2600, 3600, 4X00, 7100, 7200, 7500, ATM interfaces, distributed VIP

- **Cisco software releases**
  - Cisco IOS: 11.1, 11.1cc, 11.2, 11.3, 12.0, 12.0(5)T, 12.0(5)XE, 12.(0)7T
  - Cat OS: 5.1, 5.3, 5.4

- **New device and software support can be downloaded**
Oracle Testing

- Showed that QoS policies can effectively protect Oracle traffic over competing traffic
- Design guidelines and test reports
Configurations Tested

- **Oracle application**
  - Oracle financials and manufacturing release 11.0
    - 3-tier, browser downloaded
      - Java applet
    - Oracle 8.05 database
- **Cisco IOS 11.2**
  - PQ, WFQ, GTS (limiting), CQ
Test Description

Replicated Oracle Application Performance over Low-Speed WAN

DCE Clockrate Changed on this Port to Vary Bandwidth from 64 KBPS to 256 KPS

Simulated Frame Relay

Policies Applied on 7200 Egress Interface

2 Mbps

Cisco 7200

2 Baselines Run

“Remote Office” LAN

“Headquarters LAN”

HTTP Used as Background

Test Scripts Supplied by Oracle

Replicated Oracle Application Performance over Low-Speed WAN

Policies Applied on 7200 Egress Interface

2 Mbps

Cisco 7200

2 Baselines Run

“Remote Office” LAN

“Headquarters LAN”

HTTP Used as Background

Test Scripts Supplied by Oracle

Packet Distribution Histogram

- High occurrence of 1500 byte packets for Oracle and Web
- Higher number of Web to Oracle packets indicates need to have strong policy on Oracle
- Can use this information to
  Calculate maximum delay behind low priority packet, and determine need for fragmentation
  Forward from priority queue in whole multiples of largest packet size
What We Found

- Best improvements were made with CQ
- Adjusting CQ highly effective for improving response time
- CQ support available starting Cisco IOS 11.1

![Graph showing response times for different configurations](image)

Custom Queuing

- Specify ratio of bytes (percentage) to forward from a queue
- Specify maximum number of packets in each queue, default is 20 packets
Assess Improvement in Oracle Response Time

Oracle
Finer Tuned Policies Yield Better Response Times for Oracle

<table>
<thead>
<tr>
<th>CQ95%</th>
<th>CQ99%</th>
<th>CQ</th>
<th>CQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>127.17</td>
<td>101.09</td>
<td>83.58</td>
<td>70.81</td>
</tr>
</tbody>
</table>

Resp Time (sec)

Background
Finer Tuned Policies May Impact Packet Loss for Background Traffic

<table>
<thead>
<tr>
<th>CQ95%</th>
<th>CQ95%</th>
<th>CQ Byte30K-6</th>
<th>CQ Byte30K-400</th>
<th>CQ Byte30K-400 Pkt40-1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Packet Loss

Interaction of Oracle vs Web

- High CQ policy settings still lets web traffic through

Custom Queue 30000-400 Byte/40-4 Limit

Oracle

Web
PeopleSoft Testing

• Showed that QoS policies can effectively protect PeopleSoft traffic over competing traffic
• Recommended practices, test reports, and templates

Packet Distribution

<table>
<thead>
<tr>
<th>Range</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-64</td>
<td>10%</td>
</tr>
<tr>
<td>65-127</td>
<td>15%</td>
</tr>
<tr>
<td>128-252</td>
<td>5%</td>
</tr>
<tr>
<td>253-511</td>
<td>55%</td>
</tr>
<tr>
<td>512-1023</td>
<td>10%</td>
</tr>
<tr>
<td>1024-1518</td>
<td>5%</td>
</tr>
</tbody>
</table>
Interaction of PeopleSoft vs Background

CBWFQ with 50% Min BW for PeopleSoft

What We Found

• Best improvements were made with CBWFQ
Bandwidth Utilization

- BL, w cong
- WFQ
- CBWFQ, 25%
- CQ, 80%
- PQ
- CBWFQ, 50%
- BL, w/o cong

CBWFQ

- Supports user-defined traffic classes
- Queue reserved for each class

Set BW to 50% for PeopleSoft
### Recommended Implementation Scheduling

- Business requirements document
- Lab testing
  - Characterize mission critical and background app
  - Determine QoS mechanisms and settings
  - Determine impact on background traffic
- Phase 1—Proof of concept
  1–2 links, 1–2 apps
- Phase 2—Larger scale
  Select division, more apps, more links

### Implementation Methodology

```
Monitor Impact on-going
Implement Policy
Test/Validate Policies
Define Policy Settings
Application Profiling
Baselining
Business Req.
```

- Measure
- Implement
- Design
- Plan
Business Requirements

- Project scope
- Define business objectives
- Define mission critical apps and background
- Define classes
- Survey LAN, WAN, mgmt apps, apps, lab set-up
- Have sign off

Bus Req Interview Questions

- What are the mission critical apps?
- What is the background traffic?
- What needs to be prioritized?
- What are the business goals?
- Where are the perceived choke points?
- How are the apps deployed? On which servers?
- What does the network look like?
- What version of Cisco IOS, Cat OS is running?
- What is the time frame for implementation?
Baselining

• Collect topology information
• Collect traffic information on utilization and traffic flows
• Analyze collected information for the flow of current traffic and volumes

Application Profiling

• Drill down
• Profile business critical applications: packet distribution, bandwidth utilization, and burstyness
• Document each transaction of interest
• Verify that the application can be classified by either IP address, host name, protocol, port number, URL, or NBAR support
  - IP address
    Common to see critical apps on dedicated servers
  - Port number
    Additional granularity for multi-purpose servers
  - NBAR
  - URL
Enforcement Guidelines

- CBWFQ has desirable results
- Start with conservative setting
- IP RTP header compression recommended on links < 1.544 Mbps for real-time applications

Test/Validate Policies

- Set up lab with like routers and switches
- Like app environment
- Adtech delay simulator
- Capture/playback tool like Mercury WinRunner
- Load generator like LoadRunner or Chariot
First Case Study

- Customer
- Vendor
- Consultant

Netigy

- QPM implementation partner
- Experts in both network and application consulting
- Have gone through several implementations
- Authored QPM/PeopleSoft test report
- Annette Clewett, QPM Practice Director, Author of “Network Resource Planning for SAP R/3, BAAN IV, and PeopleSoft”
- 500 professionals trained in Cisco technology and supporting network competencies
- Contact: Tina Burke, tina.burke@netigy.com
Conexant Systems

- $2B independent company focused exclusively on providing semiconductor solutions for communications electronics
- Spin off of Rockwell in 1999
- Have bought QPM for strategic purpose
- SAP, Voice as key apps
- HQ in Newport Beach, California

LAN Environment

- Newport Beach HQ largest LAN with 4000 nodes
- Core switches Cabletron 9000, to be replaced by Catalyst 6500
- Engineering LAN converted to Catalyst 6500 in core and Catalyst 5000 in wiring closet
- Central Newport Beach data center home to all servers, except Lotus Notes
WAN Environment

- Cisco 7500, 3600, 2500
- Cisco IOS 11.2 to 12.0(7)T
- Domestic circuits Frame Relay PVC speeds 256 Kbps to T1, 0 CIR
- High-speed circuits exist between Newport Beach, San Diego (DS3), Newbury Park (DS3) and Mexicali (2-T1s)
- International circuits are also Frame Relay 64 Kbps to E1, PVC with non-zero CIR
- Dual Frame Relay PVCs are used for domestic sites with VoIP traffic

Application Environment

- SAP recently upgraded to 4.6B
- Deployed logical 3-tier, physical 2 tier running on Compaq Alpha server
- 700 total users, no more than 200 concurrent users
- Remote sites 15–25 users
- Concerned about performance of SAP across international low speed Frame Relay with non-zero CIR
- Treat SAP as one service level for all SAP modules
Application Environment

- Voice over IP (VoIP) is being deployed for toll-bypass
  - Between Newport Beach and select remote sites
- QoS goal is to support VoIP and data on a single PVC for use on international circuits
  - Dual PVCs are not cost effective
- The current VoIP infrastructure
  - Cisco 3640 router at Newport Beach
  - Cisco 3640 remote site routers running Cisco IOS 12.0(7)T

Business Requirements

- Business objectives
  - Improve response time for critical applications
  - Problem with background traffic impacting remote sites
  - Uses dual PVC domestic, single PVC international
  - Converge voice and data on single PVC for international
  - Also would like to converge for domestic links
  - 3 classes of service
- Apps
  - SAP, Voice, Notes, Promis, Resumix
- Network
  - Cisco 7500, Catalyst 6000, Cisco 3640, Cisco AS5300, Cisco IGX, L2 switches, Legacy PBX
- Duration: Two weeks, conducted with phone interview
First Site Visit

- Reconfirm business requirements doc
- Install QPM
- Capture traffic for playback in lab

Phase 1 Plan

- Proof of concept
  Prioritize SAP and voice
  Santa Clara and Tokyo links
- Deliverables:
  Business requirements document
  Install QPM
  Replicate test environment
  Determine policy settings
  Validate response time improvement
Policy

- 3 classes
  - Gold—Strict priority for voice
  - Silver—Mission critical traffic (SAP)
  - Bronze—Best effort (other)
- Classification
  - IP address for SAP
  - RTP port range for voice: 16384 to 32767
- Enforcement
  - IP RTP priority (strict priority queuing) for voice
  - Frame Relay traffic shaping
  - Frame Relay Fragmentation (FRF.12)
  - cRTP header compression
  - Color SAP
  - WFQ for SAP

Traffic Classification and Coloring

Newport Beach

SAP

Cisco 3640

Cisco 3640

256K BIR
32K CIR

Rolm PBX

Tokyo/Santa Clara

SAP Client
Role of QPM

Conexant Requirements

- Cisco 7500, 3600, 2500
- IOS 11.2 to 12.0(7)T
- Multiple policies
- Voice, SAP
- Ease of use
- Network focus

QPM

- CLI commands generated for multiple devices
- Multiple IOS support
- Central policy database support
- Multiple application support
- GUI-based, rules-based,
- Network focus

Fragmentation Guidelines

- Use FRF.12 on slow-speed FR links
- FRF.12 recommended on links whose endpoints are both Frame Relay
- Recommend fragment sizes

<table>
<thead>
<tr>
<th>Link Speed</th>
<th>Fragment Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>56kbps</td>
<td>70 Bytes</td>
</tr>
<tr>
<td>64</td>
<td>80</td>
</tr>
<tr>
<td>128</td>
<td>160</td>
</tr>
<tr>
<td>256</td>
<td>320</td>
</tr>
<tr>
<td>512</td>
<td>640</td>
</tr>
<tr>
<td>768</td>
<td>1000</td>
</tr>
<tr>
<td>1536</td>
<td>Not Necessary</td>
</tr>
</tbody>
</table>
VoIP Is Classified and Placed into a Priority Queue which Controls Jitter and Delay

SAP Packet Distribution

0% 5% 10% 15% 20% 25% 30% 35% 40%

Interaction of SAP, VoIP and Background

What We Found

- After placing voice in “Gold”
- Impact on SAP in “Silver”

Baseline w Congestion (SAP,VoIP,BG)
With Policy Added (WFQ)
Baseline w/o Congestion (SAP only)
Second Case Study

- Customer
  [Image of First American Financial]

- Vendor
  [Image of Cisco Systems]

- Consultant
  [Image of Netigy]

First American Financial

- $2.8B leading provider of real-estate related services
- 1100 branches worldwide
- #2 in PC Week (1999) Top 100 Financial Services E-Business Innovators
- Running Cisco 7206, 7507, 3640, 3620, 2600, 2500, 1601, Catalyst 4000, 6509
- Various releases of Cisco IOS: 11.1 to 12.0(7)T
Business Requirements

• Business objectives
  Improve response time for “FAST” application
  Facilitate future deployment of voice and video applications
  Increase bandwidth efficiency
  Proactively prepare network for voice
  Improve policy administration

• Apps
  FAST

• Network
  Cisco 7206, 7507, Catalyst 6509, Catalyst 4000, Cisco 3640, 3620, 2600, 2500 and 2601

Classes

• Platinum
  Delay sensitive “real-time” applications, such as video and voice

• Gold
  Critical, non-delay sensitive applications such as ERP and e-commerce

• Silver
  Important, internal applications such as intranet websites

• Bronze
  For non-critical applications; bulk file transfers, casual Internet browsing
Application

- The FAST application
  Manages title and escrow documents
- Design policy with FAST in “Gold” class
- The application architecture is web-based with backend application and database servers
  All servers, including the first tier “web-server” are co-located at FAFC’s HQ
  Remote users are heavily dependent upon reliable and predictable network access to the HQ
- FAFC is planning to deploy Voice over IP (VoIP) for toll bypass to its remote sites

LAN

- Santa Ana, the headquarters (HQ)
  Contains the largest LAN and a centralized data center
- All application servers
  Front ended by a Catalyst 6509 server switch
- A Cisco 7206VXR router running 12.0(7)T
  Performs HQ routing functions
**WAN**

- Structured in a three-layer hierarchy
  - Core, distribution and access layer
- Core layer, or backbone,
  - 16-7206 series routers running Cisco IOS 12.0(3.4)T and
  - 2-7206VXR series routers running Cisco IOS 12.0(7)T
- Distribution layer provides connectivity to geographic regions
  - 1- 7507 router running IOS 11.1(28)CC,
  - 2- 7206 routers running IOS 12.0(3.4)T and
  - Numerous 3640 routers running IOS 11.3(3a)T
- Access layer is made up of a large quantity of branch offices
  - Catalyst 4000, Cisco 3640, 3620, 2600, 2500 and 2601 series routers running a variety of IOS ranging from 11.1.(12)AA to 11.3(10)T
Packet Distribution

Packet Size Distribution

Bytes

0-64 65-127 128-255 256-511 512-1023 1024-1518

% 45 40 35 30 25 20 15 10 5 0

Background

Fast

Policy Findings for Fast Application

Congested Baseline
CQ—Ratio Set to PS=50%, BG=50%
WFO with IP Precedence
CQ—Ratio Set to PS=80%, BG=20%
PQ—FAST = High Queue, BG = Normal
Non-Congested Baseline

Seconds

0 5 10 15 20 25 30 35 40 45
Dos and Don’t Summary

• Dos:
  - Keep your implementation simple
  - Bring in implementation experts
  - Generate business requirements plan
  - Address Layer 8 “organizational” issues
  - Use a small number of classes
  - Determine impact on background traffic
  - Phase implementation
  - Apply to small number of links and apps
  - Do SLA in later phase

• Don’ts
  - Add unnecessary complexity

Resources

• Web
  - Product overview, data sheets, test reports, ordering info, etc.


Deploying QPM in an Enterprise

Session 2606

Please Complete Your Evaluation Form

Session 2606