QoS Deployment Experiences

SANOG VII Mumbai, India 16-24Jan’06

Rajesh Shinde
rajesh.shinde@relianceinfo.com
Agenda

❖ What is QoS?
❖ Why QoS?
❖ Benefits of QoS For Service Provider
❖ QoS Functional Components & Models
❖ QoS in MPLS Network
❖ QoS Deployment Experiences
What is QoS?

“Collection of technologies which allows applications/users to request and receive predictable service levels in terms of data throughput capacity (bandwidth), latency variations (jitter) and delay"
What is QoS? (cont.....)

- Measure of transmission quality and service availability of a network
- Transmission quality of the network is determined by the following factors: Latency, Jitter, and Loss.
- QoS from User Perspective: The network capability to provide the desired application performance for Voice, Video, and Data!
- QoS from Service Provider Perspective: Methods to utilize existing network capacity efficiently and meet performance requirements and achieve the maximum traffic throughput. “Managed unfairness”
What is QoS? (cont.....)

- **Objectives of QoS**
  - Supporting dedicated bandwidth
  - Improving loss characteristics
  - Avoiding and managing n/w congestion
  - To give priority to certain mission critical applications in the n/w
  - To maximize the use of the current network investment in infrastructure
  - Better performance for delay sensitive applications such as Voice and Video
  - To respond to changes in n/w traffic flows
Agenda

- What is QoS?
- Why QoS?
- Benefits of QoS For Service Provider
- QoS Functional Components & Models
- QoS in MPLS Network
- QoS Deployment Experiences
Why QoS?

**Technological Reasons**

- Real time applications are sensitive to delay, jitter, and packet loss
- Voice, Video, and Data application traffic demand varying service requirements
- Over subscription of available bandwidth for multiple applications
- Resultant congestion and ensuring of SLAs for priority traffic in networks
- Optimization of bandwidth utilization
Agenda

- What is QoS?
- Why QoS?
- Benefits of QoS For Service Provider
- QoS Functional Components & Models
- QoS in MPLS Network
- QoS Deployment Experiences
Benefits of QoS For Service Provider

- B/W Management ➔ by congestion management and avoidance mechanism
- Better handling of N/W capacity for new applications / services
- Increase in revenue by selling multiple service classes over shared infrastructure like MPLS backbone
- Predictive behavior guarantee on the data IP / MPLS network (which are usually best effort).
Agenda

- What is QoS?
- Why QoS?
- Benefits of QoS For Service Provider
- QoS Functional Components & Models
- QoS in MPLS Network
- QoS Deployment Experiences
QoS Functional Components & Models

- **QoS Functional Components**
  - Classification of packets on the basis of
    - Layer 2 parameters (802.1Q CoS bits)
    - Layer 3 parameters (IP Precedence, DSCP, source MPLS EXP bits or destination IP)
    - Source port, destination port, or stateful inspection
  - Policing / Shaping
    - Both identify the traffic rate violations similarly
    - Policing-Dropping violating traffic or marking it with higher drop probability
    - Shaping-Delays the excess traffic by using buffer & shapes the flow to configured rate
QoS Functional Components & Models (cont....)

- Marking / Rewriting
  - Eases the classification in Service Provider’s Core n/w
  - Carries packet’s service class & drop probability information
  - Can be tied to policing
  - Layer 2 and layer 3 specific (e.g. FR DE, ATM CLP, 802.1p/q, IP DSCP, IP Precedence, MPLS Experimental field)

- Scheduling (Congestion Management/Avoidance)
  - Congestion Management (Queuing- CBWFQ, LLQ, MDRR)
    - Creation of queues corresponding traffic classes
    - Managing the front end of queues by transmitting the packets from queues on the basis of priority & b/w share
  - Congestion Avoidance (Dropping- With RED)
    - Monitoring network traffic loads to anticipate and avoid congestion at common network bottlenecks
    - Achieved by selective dropping of packets on the basis of drop probabilities
QoS Functional Components & Models (cont.....)

**QoS Models**

- **Best Effort**
  - Traditional IP service with no state (no traffic classification)

- **InteServ (Integrated Service) – RFC 2210,2211,2212,2215**
  - First effort towards IP QoS
  - Signaled QoS with per flow state
  - Resource/policy admission control
  - Not scalable over internet

- **DiffServ (Differentiated Service) – RFC 2274,2275**
  - Provisioned QoS
  - Per flow aggregate QoS
  - No scalability issue—Better choice for Service Provider
  - No resource/policy admission control
QoS Functional Components & Models (cont.....)

DiffServ – QoS Model

Traffic Classification and Conditioning

Classification/Marking/Policing

Per-Hop Behavior

Queuing/Dropping

Ingress Node

Interior Node

Egress Node

TCB

PHB

TCB

PHB
QoS Functional Components & Models (cont.....)

DiffServ – QoS Model

➢ Per-Hop Behaviors (PHB)
  – Expedited Forwarding (EF): RFC2598
    • Dedicated low delay queue
    • Comparable to Guaranteed B/W in IntServ
  – Assured Forwarding (AF): RFC2597
    • 4 queues × 3 drop preferences
    • Comparable to Controlled Load in IntServ
  – Class Selector: Compatible with IP Precedence
  – Default (best effort)
QoS Functional Components & Models (cont.....)

**DiffServ – QoS Model (AF PHB)**

- 4 independently-forwarded AF classes
- Within each AF class, 3 levels of drop priority! This is very useful to protect conforming to a purchased, guarantee rate, while increasing chances of packets exceeding contracted rate being dropped if congestion is experienced in the core.

Eg. AF12 = Class 1, Drop 2, thus “001100”

![Diagram showing AF Class 1 to 4 with drop preference]

- AF Class 1: 001dd0
- AF Class 2: 010dd0
- AF Class 3: 011dd0
- AF Class 4: 100dd0

dd = drop preference
Agenda

- What is QoS?
- Why QoS?
- Benefits of QoS For Service Provider
- QoS Functional Components & Models
- QoS in MPLS Network
- QoS Deployment Experiences
QoS in MPLS Network

- **MPLS support for QoS**
  - MPLS can support both IntServ and DiffServ
  - MPLS support for DiffServ being standardized (draft-ietf-mpls-diff-ext)
Co-existence of MPLS & DiffServ is scalable

Co-existence of MPLS & Diff-Serv possible because of same scalability goals, both models do:
- aggregation of traffic on Edge & processing of Aggregate only in Core
QoS in MPLS Network (cont...)

MPLS & DiffServ
Label Header for Packet Media

Label 20 bits
EXP Experimental Field, 3 bits
S Bottom of Stack, 1 Bit
TTL Time to Live, 8 Bits
  - Can be used over other layer-2 technologies
  - Contains all information needed at forwarding time
  - One 32-bit word per label
QoS in MPLS Network (cont...)

- **DSCP & MPLS EXP Bits**
  
  - DSCP field is not directly visible to MPLS Label Switch Routers (except edge LSR)
  
  - Information on DiffServ must be made visible to LSR in MPLS Header (using EXP field / Label)
Agenda

- What is QoS?
- Why QoS?
- Benefits of QoS For Service Provider
- QoS Functional Components & Models
- QoS in MPLS Network
- QoS Deployment Experiences
QoS Deployment Experiences

QoS deployment in MPLS network

• **Step 1:** Identifying application requirements
• **Step 2:** Defining Policies
• **Step 3:** Testing policies
• **Step 4:** Implementing policies
• **Step 5:** Monitoring & adjusting
Step 1: Identifying Application Requirements

- Importance of an application to the customer
  - What applications are considered “mission critical”?
- Derived from application properties
  - Application performance/quality requirements
  - Properties of the underlying transport protocol stack
- Applications with different properties/requirements should be queued separately
- Interaction of SP’s (Service Provider) business team with various enterprise customers (for understanding various service classes profiles in finalizing QoS products)
## QoS Deployment Experiences (Cont....)

### Step 1: Identify Application Requirements (cont..) - QoS Requirements of Applications example:

<table>
<thead>
<tr>
<th></th>
<th>Bandwidth</th>
<th>Delay</th>
<th>Jitter</th>
<th>Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voice Payload</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Video Payload -</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interactive</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>(2-Way)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Video Payload -</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Streaming (1-Way)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Video Signaling</td>
<td>Low</td>
<td>Low</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>Voice Signaling</td>
<td>Low</td>
<td>Low</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>Data:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interactive,</td>
<td>Variable, typical</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>Mission Critical</td>
<td>medium</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data: Not</td>
<td>Variable, typically</td>
<td>High</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>Interactive,</td>
<td>high</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mission Critical</td>
<td>high</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data: Interactive, Not Critical</td>
<td>Variable, typical</td>
<td>High</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>Mission Critical</td>
<td>medium</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data: Not</td>
<td>Variable, typically</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Interactive,</td>
<td>high</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not Critical</td>
<td>high</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data: Not</td>
<td>Variable, typically</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Interactive,</td>
<td>high</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not Critical</td>
<td>high</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Step 2: Defining Policies

- Understanding network topology and traffic flow
- Assessing capacity of network devices (CPU, software, hardware etc.) and network links (speeds, overhead, congestion, etc.)
- Identifying bottleneck and non-bottleneck links
- Trusted and untrusted sources for QoS markings
- Layer 2 vs. Layer 3 service model
- Study & Research by SP’s business team to design QoS product profiles (Interactions with various enterprise customers to understand their business application's priorities)
- Co-ordination among SP’s business team & technical team till the completion of QoS deployment
QoS Deployment Experiences (Cont....)

Step 2: Defining Policies (cont....) –

• Defining SP’s aggregate service class model (Keep minimum number of classes)
• Mapping of customer’s service class model to SP’s one
• Internet traffic is to be classified as Best Effort
• Defining SLA for customer service classes
• QoS mechanisms to be implemented at network edge & core
• QoS Transparency with MPLS DiffServ Tunneling Modes (Handing of customer’s IPP/DSCP bits settings)
Step 2: Defining Policies (cont....) –  
*Example: QoS Standard Classification and Marking Rule Recommendations*

<table>
<thead>
<tr>
<th>Application</th>
<th>L3 Classification</th>
<th>L2 CoS/MPLS EXP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IPP</td>
<td>PHB</td>
</tr>
<tr>
<td>Routing</td>
<td>6</td>
<td>CS6</td>
</tr>
<tr>
<td>Voice</td>
<td>5</td>
<td>EF</td>
</tr>
<tr>
<td>Interactive-Video</td>
<td>4</td>
<td>AF41</td>
</tr>
<tr>
<td>Streaming Video</td>
<td>4</td>
<td>CS4</td>
</tr>
<tr>
<td>Mission-Critical Data</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Call Signaling</td>
<td>3</td>
<td>AF31/CS3</td>
</tr>
<tr>
<td>Transactional Data</td>
<td>2</td>
<td>AF21</td>
</tr>
<tr>
<td>Network Management</td>
<td>2</td>
<td>CS2</td>
</tr>
<tr>
<td>Bulk Data</td>
<td>1</td>
<td>AF11</td>
</tr>
<tr>
<td>Scavenger</td>
<td>1</td>
<td>CS1</td>
</tr>
<tr>
<td>Best Effort</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Step 3: Testing Policies

- QoS policies function of customer requirement (provided by SP’s business team) & SP’s n/w’s scalability, functionality

- Test QoS policies in the lab first
  - Testing feasibility of QoS product profiles provided by business team
  - Testing these profiles under simulation of customer’s application traffic entering the SP’s n/w stochastically (under L2/L3 MPLS VPN scenario)
Step 3: Testing Policies (cont....)

- **Test QoS policies in the lab first (cont....)**
  - Testing SP’s vendors QoS implementation for it’s routers’/switches’ hardware/software (currently functional in SP’s n/w) under different L2 encapsulations/markings & L3 protocol (IP)/markings
  - Testing SP’s vendor’s n/w OS for whether classification & marking rules are obeyed under default as well as user-defined conditions or not as well as scheduling tools are functioning as expected or not
QoS Deployment Experiences (Cont....)

Step 3: Testing Policies (cont....)

- *Test QoS policies in the lab first (cont....)*
  - ✔ Testing of QoS service profiles’ actual behavior vs. exhibited behavior (both should match)
  - ✔ Testing of QoS profiles for whether there is flexibility in allocation of n/w resources among different service classes
  - ✔ Testing of SLA probing feature available in routers’/switches’ software for SLA monitoring & Capacity Planning purpose
QoS Deployment Experiences (Cont....)

Step 3: Testing Policies (cont....)

- Test QoS policies in the lab first (cont....)
  - Testing integration of this SLA probing feature with SP’s current NMS/OSS
  - Test policy in a small portion of the production network
  - Run baseline tests with and without QoS under congestion conditions
QoS Deployment Experiences (Cont....)

Step 4: Implementing Policies

• Perform Classification on MPLS PE routers
• Perform policing and marking on MPLS PE routers
• Work toward core applying inbound/outbound policies
• Phased deployment—apply your policies incrementally
Step 5: Monitoring and Adjusting

- Measure delay and loss for different service classes
- Monitor application performance
- Adjust policies where necessary
Thank You