QoS for VOIP at the edge

LINX 65
London
18th May 2009

Richard Halfpenny
Network Engineering / Operations
Exa Networks Ltd (AS30740)
richard.halfpenny@exa-networks.co.uk

exa networks
let's keep things simple

Presentation Overview

- Introduction
- VOIP overview
- Quality of Service – for/against
- How VOIP breaks on packet switched networks
- Improving VOIP end-user experience
- QoS fundamentals
- QoS in service provider network
- Configuration examples

exa networks
let's keep things simple

Introduction

Introduction

- Exa Networks – AS30740
- Specialist northern based ISP
- Business and education markets
- HQ based near Bradford, West Yorkshire
- Present at both LINX and AMS-IX
- Focus on service quality and customer support

exa networks
let's keep things simple

Introduction

- Scope of presentation is overview of implementing QoS at network edge (PE->CE)
- We typically implement on leased private circuits – Ethernet and Serial
- DSL a different beast (and talk!). More layers to consider – PPPoATM sessions delivered over L2TP over IP over ATM (BT 20CN).

exa networks
let's keep things simple

Introduction

- Exa implementing QoS on Cisco equipment at the edge, mostly on 7301 and 7204.
- ISR's (1841/2811) for CPE
- Juniper M series in core for L3, Foundry at L2.
- QoS features we use not supported on M series without intelligent queuing PICs.
  - e.g. rate limiter/shaper on vlan with queues within rate limiter
- Intend to migrate fully to Juniper in future (both L2/L3

exa networks
let's keep things simple

VOIP
Overview

VOIP overview

- "Talking to people over t'interwebs"
- Signalling
  - Call setup and tear-down
- Media
  - Audio stream from caller A to caller B
  - Audio stream from caller B to caller A

exa networks
let's keep things simple

VOIP overview

- Signalling for call setup/teardown
  - Typically SIP or Cisco Skinny (SCCP)
  - Phone to Phone
  - Phone to PSTN Gateway or Session Border Controller
  - Call routing
  - Co-ordinates codecs used for media stream

VOIP overview

- Media (sending process)
  - Typically RTP/udp
  - Audio (speech) waveform, amplitude of which is sampled at constant rate X number of times per second. 8000 / 8Khz at 8 bit resolution but can vary.
  - Blocks of raw samples encoded in specific format (codec). Typically every 20ms or 160 samples (assuming 8Khz sample rate) but can vary depending on codec used. Codec may compress samples to save bandwidth.
  - Encoded blocks of samples encapsulated in RTP/udp packet (packetised). Sent at regular intervals, usually one encoded block per IP packet, to receiving device. Typical packet size of 200 bytes (IP+UDP+RTP headers 40 bytes, 160 bytes payload).

exa networks
let's keep things simple

VOIP overview

- Media (receiving process)
  - Reversal of encoding but with some additional safeguards to help maintain uninterrupted playback
  - RTP/udp packet is buffered (de-jitter buffer)
  - Encoded block is retrieved from packet
  - Block decoded to raw audio samples
  - Samples played back

exa networks
let's keep things simple

Quality of Service For/Against?

QoS – reasons to implement?

- Customers request it
  - Trend is to migrate towards converged data/voice networks.
- Improved perception of service provider “quality”
- Competitive advantage
- Can make VOIP just as reliable as TDM delivery if implemented correctly combined with resilient network design
- Turning on QoS inside a router usually doesn’t cost any money but upgrading a circuit does
- Upgrading circuit line rate won’t help.. applications can (and will) consume any sized link to the detriment of VOIP.

exa networks
let's keep things simple

QoS – reasons to NOT implement?

- Engineering time and/or expertise costs money (especially if lacking the skill-set in-house)
- Router hardware/OS might not support
  - More of a problem on ASIC forwarding platforms as hardware could lack some required features
- Provider might be against VOIP adoption
  - Want to dissuade customers from using it over their own PSTN/mobile offerings or they might provide their own competing VOIP service.
- Network neutrality?

exa networks
let's keep things simple

What can break VOIP on a packet switched network?

exa networks
let's keep things simple

VOIP - what can break on packet switched networks

- Numerous factors affect the experience of VOIP for an end-user on a packet switched network!
  - Jitter
  - Packet Loss
  - Out of order packets
  - Lack of bandwidth

exa networks
let's keep things simple

VOIP - what can break on packet switched networks

- Jitter
  - Any variation in the delay of media packets sent from caller A reaching caller B.
  - Receiving end must see packets at regular intervals.
  - De-jitter buffer is used at receiving end – packets clock in to buffer at irregular intervals and clock out uniformly for decoding. Typical length 30-50ms.
  - Too much jitter and the buffer becomes exhausted. Length can be increased but affects end-to-end delay. ITU G.114 recommendation for end-to-end is 150ms max – at 400ms two way conversation becomes very difficult. Codec delays eat into this figure.

exa networks
let's keep things simple

VOIP - what can break on packet switched networks

```
[root@lts ~]#
[root@lts ~]# ping -s 1472 -M do -i 0.05 82.219.4.24
PING 82.219.4.24 (82.219.4.24) 1472(1500) bytes of data.
1480 bytes from 82.219.4.24: icmp_seq=1 ttl=54 time=18.7 ms
1480 bytes from 82.219.4.24: icmp_seq=2 ttl=54 time=18.9 ms
1480 bytes from 82.219.4.24: icmp_seq=3 ttl=54 time=19.6 ms
1480 bytes from 82.219.4.24: icmp_seq=4 ttl=54 time=19.4 ms
1480 bytes from 82.219.4.24: icmp_seq=5 ttl=54 time=18.3 ms
1480 bytes from 82.219.4.24: icmp_seq=6 ttl=54 time=19.4 ms
1480 bytes from 82.219.4.24: icmp_seq=7 ttl=54 time=18.7 ms
1480 bytes from 82.219.4.24: icmp_seq=8 ttl=54 time=19.5 ms
1480 bytes from 82.219.4.24: icmp_seq=9 ttl=54 time=18.8 ms
1480 bytes from 82.219.4.24: icmp_seq=10 ttl=54 time=18.7 ms
1480 bytes from 82.219.4.24: icmp_seq=11 ttl=54 time=19.3 ms
1480 bytes from 82.219.4.24: icmp_seq=12 ttl=54 time=18.8 ms
1480 bytes from 82.219.4.24: icmp_seq=13 ttl=54 time=18.8 ms
1480 bytes from 82.219.4.24: icmp_seq=14 ttl=54 time=19.5 ms
1480 bytes from 82.219.4.24: icmp_seq=15 ttl=54 time=21.1 ms
1480 bytes from 82.219.4.24: icmp_seq=16 ttl=54 time=19.2 ms
1480 bytes from 82.219.4.24: icmp_seq=17 ttl=54 time=18.0 ms
1480 bytes from 82.219.4.24: icmp_seq=18 ttl=54 time=18.8 ms
```

exa networks
let's keep things simple

VOIP - what can break on packet switched networks

```
[root@lts ~]# ping -s 1472 -M do -i 0.05 82.219.194.12
PING 82.219.194.12 (82.219.194.12) 1472(1500) bytes of data.
1480 bytes from 82.219.194.12: icmp_seq=1 ttl=243 time=90.3 ms
1480 bytes from 82.219.194.12: icmp_seq=2 ttl=243 time=92.1 ms
1480 bytes from 82.219.194.12: icmp_seq=3 ttl=243 time=115 ms
1480 bytes from 82.219.194.12: icmp_seq=4 ttl=243 time=91.4 ms
1480 bytes from 82.219.194.12: icmp_seq=5 ttl=243 time=124 ms
1480 bytes from 82.219.194.12: icmp_seq=6 ttl=243 time=122 ms
1480 bytes from 82.219.194.12: icmp_seq=7 ttl=243 time=89.9 ms
1480 bytes from 82.219.194.12: icmp_seq=8 ttl=243 time=89.8 ms
1480 bytes from 82.219.194.12: icmp_seq=9 ttl=243 time=89.6 ms
1480 bytes from 82.219.194.12: icmp_seq=10 ttl=243 time=90.8 ms
1480 bytes from 82.219.194.12: icmp_seq=11 ttl=243 time=91.2 ms
1480 bytes from 82.219.194.12: icmp_seq=12 ttl=243 time=91.3 ms
1480 bytes from 82.219.194.12: icmp_seq=13 ttl=243 time=90.9 ms
1480 bytes from 82.219.194.12: icmp_seq=14 ttl=243 time=90.1 ms
1480 bytes from 82.219.194.12: icmp_seq=15 ttl=243 time=91.0 ms
```

exa networks
let's keep things simple

VOIP - what can break on packet switched networks

- Jitter
  – No media re-transmissions.. if playback "timeslot" for an encoded sample block is missed or de-jitter buffer is exhausted an audio drop-out occurs.
  – Small RTP packet (200 bytes) stuck behind a number of larger data packets (1500 bytes) within interface output queue.
  – Lower circuit line rate means higher serialisation delay and packets take longer to clock onto wire.

exa networks
let's keep things simple

VOIP - what can break on packet switched networks

- Jitter
  - Larger non-VOIP data packets queued in front of RTP will have a bigger affect on jitter for RTP on circuits with lower line rates.
  - 1500 byte packet at 100Mbps takes 0.12ms to serialise onto wire.
  - 1500 bytes at 2Mbps = 6ms
  - 1500 bytes at 250kbps = **48ms**

exa networks
let's keep things simple

VOIP - what can break on packet switched networks

- Jitter
  - At 250kbps (e.g. typical DSL upload) it takes 48ms to clear a 1500 byte packet that has just started to be serialised onto wire.
  - If a small RTP packet was waiting behind this packet being transmitted then a 40ms de-jitter buffer would be emptied before RTP reaches the receiving end.
  - RTP packet queued behind two 1500 byte packets in a simple FIFO would experience jitter of 96ms.

exa networks
let's keep things simple

VOIP - what can break on packet switched networks

- Packet loss
  - Interface CRC errors.
  - Interface output buffer exhaustion leading to tail drop. Problematic on equipment with small buffers – microbursts not absorbed.
  - No re-transmission mechanism in media stream.
  - Leads to break-up of audio.
  - Control channel sensitive to packet loss but not jitter.

exa networks
let's keep things simple

VOIP - what can break on packet switched networks

- Out of order packets
  - Multiple paths to a destination over links of unequal path length the most likely cause.
  - Per-packet load balancing at L2 or L3 is very bad idea for VOIP transmission.
  - Causes discards in decoder leading to loss of sections of audio.

VOIP - what can break on packet switched networks

- Lack of bandwidth
  - Even the best QoS won't shift 4Mbps of RTP down a 1Mbps circuit!
  - Unless some form of CAC (Call Admission Control) is used, a single additional 80kbps call on a 1M circuit carrying 960kbps of RTP (12 calls) will cause audio break-up for all users.

exa networks
let's keep things simple

Improving end-user experience

Improving end-user experience

- Ensure low jitter for media (RTP) stream
- Minimise packet loss
- Fix poorly implemented routing
- Increase circuit bandwidth

exa networks
let's keep things simple

Improving end-user experience

- Ensure low jitter for media (RTP) stream
  - Prioritise RTP over other data when queuing for output/transmission on router interfaces.
  - Prefer to deploy higher line rates if possible (reduces serialisation delay).
  - If using 3rd party packet switched circuits (e.g. EoMPLS) ensure SLA's are agreed.

Improving end-user experience

- Minimise packet loss
  - Increase output buffers if possible but this can have a serious negative effect on jitter if priority queuing is not implemented.
  - Ensure network monitoring systems are watching for interface errors, CRC’s etc.

exa networks
let's keep things simple

Improving end-user experience

- Out of order packets
  - For multiple links of equal cost between routers it's generally best practice to use load sharing that chooses egress interface towards next hop based on a hash of destination IP address.
  - 802.3ad Ethernet link aggregation.
  - MLPPP on lower speed bundles such as multiple E1's can work effectively and preserves packet ordering over links of unequal end-to-end latency.

exa networks
let's keep things simple

Improving end-user experience

- Increase circuit bandwidth
  - Reduction in serialisation delay from higher line rate can improve end-to-end delay. No need for as large de-jitter buffer.
  - Customers need to be aware of limitations on how many concurrent calls can be passed. At some point an upgrade will be required.
  - Priority queued RTP will starve bitrate available for other data, including control channel. Link fully utilised with RTP may be useless for any best efforts traffic (e.g. http/smtp).

exa networks
let's keep things simple

QoS Fundamentals

QoS Fundamentals

- Classification
  - Identify interesting traffic
- Marking
  - Mark packet with classification
- Queuing
  - Transmit traffic based on classification / marking

exa networks
let's keep things simple

QoS Fundamentals

- Classification
  - Identify traffic based on layer 3 packet headers.
  - ACL's to match IP protocol, port range, src/dst IP address, etc.
  - Deep packet inspection devices.
  - Cisco NBAR (Network Based Application Recognition)

exa networks
let's keep things simple

QoS Fundamentals

- Classification
  - Agree with customer exactly what traffic to match.
  - Some customer configurations may run on non standard port numbers.

QoS Fundamentals

- Marking
  - Mark packets with information representing classification for use later when queuing.
  - Layer 2 and Layer 3 handled separately.
  - Ethernet has CoS field but only inside .1q frame. Marking lost if frame becomes untagged.
  - IP header has 1 byte TOS field for IP Precedence or Differentiated Services.

exa networks
let's keep things simple

QoS Fundamentals

- Marking
  - CoS and IP Precedence have 3 bits usable giving maximum of 8 possible values.
  - DSCP (DiffServ CodePoint) has 6 bits usable for 64 possible values.

  - DSCP see RFC2474

QoS Fundamentals

- Marking
  - Try to do only once, as packet enters SP network.
  - Use markings elsewhere in network rather than doing a classify/mark at each router.
  - Overwrite markings set by customer for traffic that obviously isn't priority. Attacker could set their own markings on junk traffic and flood.

QoS Fundamentals

- Queuing
  - Queue based on value contained in markings.
  - Priority queuing on output of interface only.
  - Multiple queues pointless on input as there's nothing you can do about packet priority at receiving end.
  - Know what your hardware is capable of!
  - Linecard limitations etc.

QoS Fundamentals

- Queuing
  - Software and hardware queues on most non ASIC accelerated platforms.
  - Software queue by default is a single FIFO. Can be split to form multiple queues, each being serviced with different priority and/or policy.
  - Hardware queue (aka TX Queue or TX Ring) is a small FIFO designed to ensure there is data to serialise onto wire as soon as current frame being transmitted fully clears interface.

exa networks
let's keep things simple

QoS Fundamentals

- Queuing
  - IOS usually automatically reduces the TX ring to a small value when a queuing mechanism is enabled on an interface.
  - Too small and link becomes under-utilised due to empty clock cycles between L2 frames.
  - Too big and increases jitter. TX ring is a final FIFO *after* priority/best effort/etc queues.

exa networks
let's keep things simple

QoS Fundamentals

- Queuing
  - Queuing only important when there is output congestion on an interface.
  - Congestion forms whenever there is more than one packet waiting to be transmitted in the output buffer/queue.
  - If an interface never experiences congestion then only a simple FIFO would be needed. This is very unlikely however in practice due to micro traffic bursts.

exa networks
let's keep things simple

QoS in Service Provider network

QoS in service provider network

- Congestion point usually at edge between PE-CE routers
  - Lower bitrate links mean higher serialisation delay and bigger effect on jitter even when priority queuing is in place.
  - Faster links can easily be consumed by applications (e.g. Bittorrent) – bandwidth not the answer so QoS needed.

exa networks
let's keep things simple

QoS in service provider network

- Needed in core for end-to-end
  - QoS in core mitigates problems caused by very large micro bursts introducing significant jitter and packet loss.
  - DoS conditions – mark bad traffic at edges and drop / police everywhere based on marking rather than use IP ACL’s.
  - Serialisation delays generally lower due to faster line rates. Faster interfaces usually have larger buffers but just as much probability of jitter being introduced.
  - RTP packet stuck behind 1000 x 1500 byte packets within interface FIFO on 1Gbps circuit adds 12ms jitter. This however won’t exhaust most de-jitter buffers.

Configuration Examples

Configuration Examples

- Basic examples from dev hardware
- Cisco IOS 12.4 SP Services
- CE router is 1841 with two FastEthernet interfaces
- PE is 2811 with two FastEthernet interfaces

exa networks
let's keep things simple

Configuration Examples

- Classification
  - class-map
  - Match even numbered RTP ports beginning at port 16384
  - Match packets already marked DSCP EF (expedited forwarding) elsewhere (e.g. other edge routers)
  - Match both unmarked RTP and marked

```
!
class-map match-any voip-rtp
 match ip rtp 16384 16383
!
class-map match-any dscp-ef
 match dscp ef
!
class-map match-any voip-rtp-and-dscp-ef
 match class-map voip-rtp
 match class-map dscp-ef
!
```

Configuration Examples

- Marking
  - policy-map (service policy on interface for input packets)
  - Mark packets classified by the voip-rtp class-map with DSCP EF (expedited forwarding)
  - Set everything else to DSCP 0 (default)

```
!
policy-map mark-traffic
  class voip-rtp
   set dscp ef
  class class-default
   set dscp 0
!
```

Configuration Examples

- Queuing
  - policy-map (service policy on interface for output packets)
  - Two queues – intended to run on a 10M interface
  - Low latency priority queue of upto 9M for RTP
  - All other traffic fair queued
  - 9M limit/policer on RTP to avoid starving default class of bandwidth

```
!
policy-map customerXX-qos
  class voip-rtp-and-dscp-ef
   priority 9000
  class class-default
   fair-queue
!
```

Configuration Examples

- Interface configuration on PE
  - 10M circuit (WES), customer commit 10M
  - Marking for input packets from CE
  - Queuing for output packets towards CE

```
!
interface FastEthernet0/0
 description Customer XX via cct ref XXXX (CIR 10M)
 ip address 82.219.120.1 255.255.255.252
 speed 10
 duplex full
 service-policy output customerXX-qos
 service-policy input mark-traffic
!
```

exa networks
let's keep things simple

Configuration Examples

- Interface configuration on CE
  - Marking for input packets from customer equipment
  - Queuing for output packets towards PE

```
!
interface FastEthernet0/0
 description ServiceProvider via cct ref XXXX (CIR 10M)
 ip address 82.219.120.2 255.255.255.252
 speed 10
 duplex full
 service-policy output customerXX-qos
!
interface FastEthernet0/1
 description Customer Equipment
 ip address 82.219.123.254 255.255.255.0
 speed 100
 duplex full
 service-policy input mark-traffic
!
```

exa networks
let's keep things simple

Configuration Examples

- Queuing
  - Also need to guarantee bandwidth for the signalling channel.
  - 200 kilobits/sec in this example
  - class-map “voip-signalling” classifies this traffic (defined based on customer signalling choice)

```
!
policy-map customerXX-qos
  class voip-rtp-and-dscp-ef
   priority 9000
  class voip-signalling
   bandwidth 200
  class class-default
   fair-queue
!
```

exa networks
let's keep things simple

Configuration Examples

- Hierarchical policy maps
  - Useful for when customer fixed CIR is less than line rate (e.g. 20M over 100M circuit)
  - Apply a shaper to cause congestion at CIR
  - Queue within shaper

Configuration Examples

```
!
policy-map customerXX-qos
  class voip-rtp-and-expedited
   priority 19000
  class voip-signalling
   bandwidth 200
  class class-default
   fair-queue
!
policy-map shape-customerXX
  class class-default
   shape average 20000000
   service-policy customerXX-qos
!
```

Configuration Examples

- Interface configuration on PE
  - 100M circuit (WES), customer commit to 20M
  - Output service policy changes to shape-customerXX

```
!
interface FastEthernet0/0
 description Customer XX via cct ref XXXX (CIR 20M)
 ip address 82.219.120.1 255.255.255.252
 speed 100
 duplex full
 service-policy output shape-customerXX
 service-policy input mark-traffic
!
```

exa networks
let's keep things simple

Thank you for your time

Questions?

exa networks
let's keep things simple