Quality of Service

on Arista, Cisco and Juniper

Lutz Donnerhacke

IKS Service

Available devices

- Existing
  - Cisco 4500(X)
  - Cisco ASR 1001-X
- Replacement
  - Arista 7020 SR
  - Arista 7280 SR3
  - Juniper MX 204
- Long term coexistence
  - Major expanding of the network
  - Smooth transition
  - Missing feature parity

QoS – Problem space

- Reasons
  - Bandwidth mismatch: in > out.
  - Interleaving: multiple inputs at the same time
- Decisions
  - Another packet is still on wire – queue or drop?
  - Other packets in still waiting – bypass, queue, or drop?
  - Queue is full – Which packet to drop?
- Also known as
  - Packet loss – oops
  - Jitter – delivery between 8am and 6pm

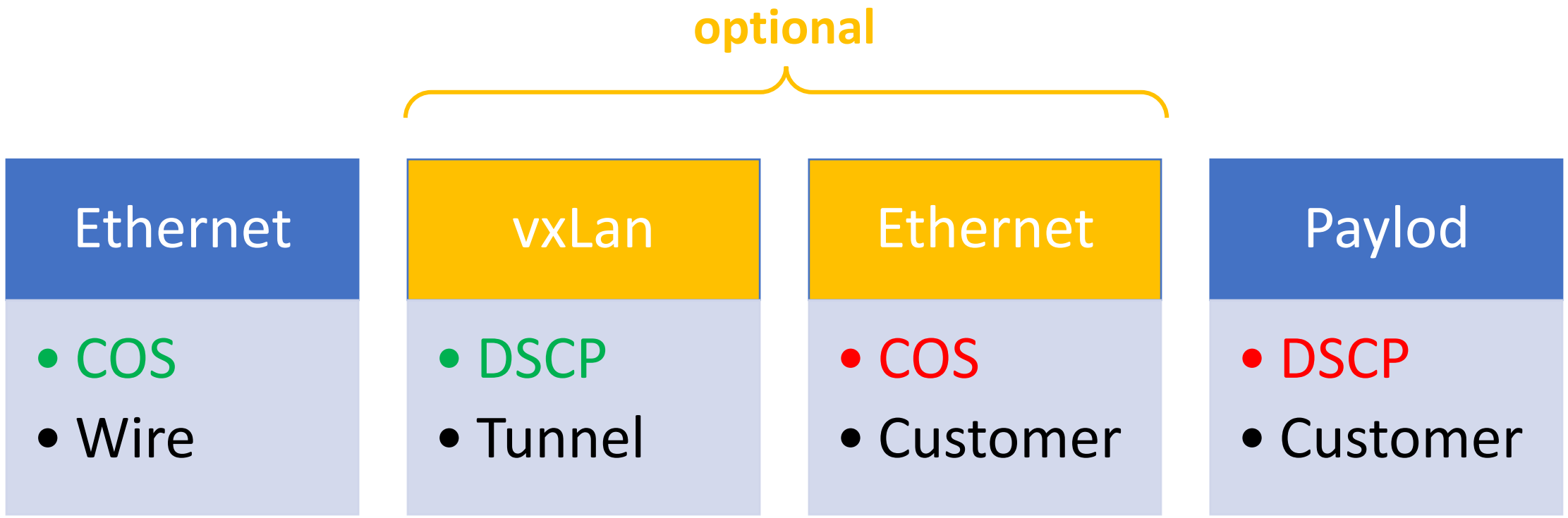
QoS – Solution space

- Avoid the problems
  - Telco approach – reserve a certain slot through the whole network
  - Virtual end2end wire with guaranteed latency and bandwidth
  - Insane amount of unused bandwidth
  - Plan and buy in advance
- Codify the decisions (made anyway)
  - Classify traffic
  - Manage queues
  - Schedule to wire
  - Mark for others

QoS – Marking Data

- Move complex classification to the edge
  - Much easier to classify (i.e. by incoming interface / vlan)
  - Trust classification inside the network
- Mark packets in DSCP
  - Available in IPv6 and IPv4 (other protocols?)
  - 6 bits – 64 possible classifications
  - Persistent mark (hard to remove, impossible to recover)
- Mark frames in COS
  - Available in 802.1p (requires tagged interconnects)
  - 3 bits – 8 possible classifications
  - Hop-by-hop mark (needs to be copied)
- Other markings (i.e. MPLS-EXP)

QoS – Marking Data



Modelling – defining global classes

Class	Purpose	COS	Queuing	Bandwidth	Effective BW
7	Device Management	**7**	Priority	10%	-
6	Network Control	**6**	Priority	10%	-
5	Real Time	5	Priority	10%	-
4	3th Party – Real Time	4	Priority	10%	-
3	3th Party – Important	3	Round-Robin	10%	16%
2	3th Party – Normal Internet	2	Round-Robin	10%	16%
1	Normal Internet	**0**	Round-Robin	40%	66%
0	Don’t Care	1	Round-Robin	Remaining	2%

Codify definitions

- Only on Juniper

  [class-of-service code-point-aliases ieee-802.1]

  unwichtig 001
  normal 000
  tp-normal 010
  tp-wichtig 011
  tp-echtzeit 100
  echtzeit 101
  netzwerk 110
  kritisch 111

- Others only “numbers”

Codify definitions

- Only on Juniper

  [class-of-service forwarding-classes]

  class kritisch queue-num 7 priority high policing-priority premium
  class netzwerk queue-num 6 priority high policing-priority premium
  class echtzeit queue-num 5 priority low policing-priority premium
  class tp-echtzeit queue-num 4 priority low policing-priority premium
  class tp-wichtig queue-num 3 priority low
  class tp-normal queue-num 2 priority low
  class normal queue-num 1 priority low
  class unwichtig queue-num 0 priority low

- Others only “numbers”

Classifying host-generated traffic

- Juniper

  [class-of-service host-outbound-traffic]

  forwarding-class netzwerk
  ieee-802.1 default netzwerk

  [interfaces lo0 unit 0]

  family inet filter output classify-loopback4
  family inet6 filter output classify-loopback6

  [firewall family inet filter classify-loopback4]

  term icmp
  from protocol icmp
  then loss-priority high; forwarding-class normal; accept

  term mgmt
  from address <mgmt-net>
  then loss-priority low; forwarding-class kritisch; accept

  term finally then accept

- Others define per service, routing daemon, etc. (if possible)

Classifying incoming traffic – Juniper

- Per interface

```
[class-of-service interfaces xe-0/1/2]
        unit <vlan> forwarding-class echtzeit
```

- Trust interconnection

```
[class-of-service classifiers ieee-802.1 intern]
        forwarding-class unwichtig loss-priority high code-points unwichtig
        forwarding-class normal loss-priority medium-high code-points normal
        forwarding-class tp-normal loss-priority medium-high code-points tp-normal
        forwarding-class tp-wichtig loss-priority medium-low code-points tp-wichtig
        forwarding-class tp-echtzeit loss-priority low code-points tp-echtzeit
        forwarding-class echtzeit loss-priority low code-points echtzeit
        forwarding-class netzwerk loss-priority low code-points netzwerk
        forwarding-class kritisch loss-priority low code-points kritisch

[class-of-service interfaces et-0/0/0]
        unit * classifiers ieee-802.1 intern
```

Classifying incoming traffic – Arista

- Per interface (matching requires “tcam feature qos …”)

```
class-map type qos match-any class-mgmt
        match vlan 12,246-248
policy-map type quality-of-service classify
        class class-mgmt / set traffic-class 7
        …
qos profile extern / no qos trust / service-policy type qos input classify
interface Ethernet1
        service-profile extern
```

- Trust interconnection

```
 qos profile intern
        qos trust cos
interface Ethernet32
        service-profile intern
```

Classifying incoming traffic – Arista

- show qos maps

```
Cos-tc map:
     cos:  0  1  2  3  4  5  6  7
     ---------------------------
     tc:   1  0  2  3  4  5  6  7
```

- Using own mappings (per interface)

  qos map cos to traffic-class name dtag
  cos 0-3 to traffic-class 0
  cos 4 to traffic-class 2
  cos 5 to traffic-class 3
  cos 6,7 to traffic-class 4

Classifying incoming traffic – Cisco

- 4500
  - Traffic-classes are internal DSCP
  - no intermediate classification possible, classification on output
  - Defaults per Interface (and Vlan) to use it on output

```
policy-map mark-netzwerk
    class class-default
        set cos 6
interface Port-channel21
    switchport mode trunk
    switchport trunk allowed vlan 14,23,…
    vlan-range 14,23
        service-policy input mark-netzwerk
    vlan-range …
```

Classifying incoming traffic – Cisco

- ASR
  - 64 qos-groups

```
class-map match-any cos-kritisch
      match cos  7
policy-map cos-trust
      class cos-kritisch
                  set qos-group 7
      ...
interface TenGigabitEthernet0/0/1
       service-policy input cos-trust
```

Marking traffic – Juniper

- Rewrite COS bits on outgoing interface

```
[class-of-service rewrite-rules ieee-802.1 intern forwarding-class normal]
        loss-priority high code-point unwichtig
        loss-priority medium-high code-point normal
        loss-priority medium-low code-point normal
        loss-priority low code-point normal

[class-of-service interfaces et-0/0/0]
        unit * rewrite-rules ieee-802.1 intern
```

Marking traffic – Arista

- “qos rewrite cos” – command not found (?!)
  - You cannot activate a feature, which is always on
  - Other chipsets have other defaults
- Inverse of classify
  - Using a map: “show qos maps”
  - Map can be customized per interface

```
qos map traffic-class to cos name dtag
  traffic-class 0 drop-precedence 0 to cos 0
  traffic-class 0 drop-precedence 1 to cos 0
  traffic-class 0 drop-precedence 2 to cos 0
```

Marking traffic – Cisco

- 4500
  - Outgoing bits set on input policy (no intermediate class)
  - Use common marks on **all** interfaces

```
policy-map cos-trust
  class class-default
     set cos cos [optional-map]
interface TenGigabitEthernet0/0/0
  service-policy input cos-trust
```

- ASR
  - Part of the output policy

```
class-map match-any qos-kritisch
      match qos-group 7
policy-map qos-queuing
  class qos-kritisch
     set cos 7
interface TenGigabitEthernet0/0/0
  service-policy output qos-queuing
```

Queues: Policy, Scheduling, and Dropping

- Each interface has different queues
  - 7020: 10G with 2 HW-Queues, 100G with 8 HW-Queues
  - Others: 8 HW-Queues
- Policy
  - Restricting bandwidth
  - Enforce contractual limits
  - Rate limiter or shaper
  - Useful for stopping traffic to the control plane
- Scheduling
  - Distribute available bandwidth
- Drop profiles
  - Early dropping before queue is filled
  - May send ECN first
  - May select special TCP streams (WFQ)

Drop Profiles

- Default linear
  - 10% filled queue causes 10% packet loss
- Juniper

  [class-of-service drop-profiles]
  aggressiv interpolate fill-level [25,60,80] drop-probability [40,80,90]
  moderat interpolate fill-level [50,75,95] drop-probability [10,25,40]
  locker interpolate fill-level [80,100] drop-probability [0,100]
- Cisco
  - Output policy-map action “random-detect cos <x> percent <min> <max>”
- Arista
  - Output queue action “random-detect ecn …”
  - drop-precedence is for counters

Managing hardware queues – Juniper

- Queue is bound to forwarding-class

```
[class-of-service schedulers echtzeit]
        transmit-rate percent 10
        priority medium-high
        excess-priority low
        drop-profile-map loss-priority any protocol any drop-profile locker
[class-of-service schedules normal]
        transmit-rate percent 40
        priority medium-low
        excess-priority low
        drop-profile-map loss-priority high protocol any drop-profile aggressive
        drop-profile-map loss-priority any protocol any drop-profile moderat
        drop-profile-map loss-priority low protocol any drop-profile locker
[class-of-service scheduler-maps default]
        forwarding-class normal scheduler normal
        forwarding-class tp-echtzeit scheduler echtzeit
        forwarding-class echtzeit scheduler echtzeit
class-of-service interfaces et-0/0/0 scheduler-map default
```

Managing hardware queues – Arista

- Only config the non-standard values

  qos profile intern
  ```
      tx-queue 0     bandwidth percent 5
      tx-queue 1     bandwidth percent 50
      tx-queue 3     no priority
      tx-queue 4     bandwidth percent 10
      tx-queue 5     bandwidth percent 10
  ```

- Check results

  show qos interfaces …

Tx Queue	Bandwidth (percent)	Shape Rate (units)	Burst-Size (units)	Priority	ECN/WRED
7	- / -	- / -	(-)	- / -	SP / SP	D
6	- / -	- / -	(-)	- / -	SP / SP	D
5	- / 10	- / -	(-)	- / -	SP / SP	D
4	- / 10	- / -	(-)	- / -	SP / SP	D
3	23 / -	- / -	(-)	- / -	RR / RR	D
2	22 / -	- / -	(-)	- / -	RR / SP	D
1	50 / 50	- / -	(-)	- / -	RR / SP	D
0	5 / 5	- / -	(-)	- / -	RR / SP	D

Managing hardware queues – Arista

- Always check results (how much bandwidth is for normal Internet?)

  show qos interfaces …

```
Tx    Bandwidth         Shape Rate              Burst-Size          Priority   ECN/WRED
 Queue  (percent)          (units)                 (units)
 ------------------------------------------------------------------------------------
   1     91 / 50      - / -          ( - )           -  /  -           SP / SP       D
   0      9 /  5      - / -          ( - )           -  /  -           SP / SP       D

Tc - tx-queue map for interface with 2 tx-queues:
     tc:         0  1  2  3  4  5  6  7
     --------------------------------
     tx-queue:  0  0  0  0  1  1  1  1
```

- But there is more!

  show interface … counters queue rates

	Offered load (include drop)			Config		Drop rates		Output rates				
	bps	pps	% of link	% of total	bps	SP/WRR	bps	pps	bps	pps	% of link	% of total	WRR %
UC0	2.8M	2.2k	0.0%	30.6%	N/A	5.0	0.0	0.0	2.8M	2.2k	0.0%	30.6%	5.0%
UC1	6.5M	2.5k	0.0%	67.5%	N/A	50.0	0.0	0.0	6.5M	2.5k	0.0%	67.5%	50.0%
UC2	0.0	0.0	0.0%	0.0%	N/A	22.0	0.0	0.0	0.0	0.0	0.0%	0.0%	22.0%
UC3	847.2	1.0	0.0%	0.0%	N/A	23.0	0.0	0.0	847.2	1.0	0.0%	0.0%	23.0%
UC4	0.0	0.0	0.0%	0.0%	N/A	SP	0.0	0.0	0.0	0.0	0.0%	0.0%	N/A
UC5	0.0	0.0	0.0%	0.0%	N/A	SP	0.0	0.0	0.0	0.0	0.0%	0.0%	N/A
UC6	186.3k	37.1	0.0%	1.9%	N/A	SP	0.0	0.0	186.3k	37.1	0.0%	1.9%	N/A
UC7	355.8	0.6	0.0%	0.0%	N/A	SP	0.0	0.0	355.8	0.6	0.0%	0.0%	N/A

Managing hardware queues – Cisco

- Everything is the outgoing policy-map
  - Must not more class statements, than hardware queues
  - Policy on aggregated interfaces, queuing on the hardware
- 4500 is limited
  - ECN and more only available on class-default
  - Only one priority queue
  - Set only bandwidth (round-robin)
  - Control queue-length manually (otherwise massive drops)
- ASR
  - More elaborate
  - Two priority levels

Managing hardware queues – Cisco

```
asr# show run | sec qos-queuing
  policy-map qos-queuing
        class qos-kritisch
              priority level 1 percent 10
        class qos-netzwerk
              priority level 1 percent 10
        class qos-echtzeit
              priority level 2 percent 10
        class qos-tp-echtzeit
              priority level 2 percent 10
        class qos-tp-wichtig
              bandwidth percent 9
        class qos-tp-normal
              bandwidth percent 9
        class qos-normal
              bandwidth percent 40
        class class-default
```

QoS – Useful debugging commands

arista# show qos interfaces …
arista# show interfaces … counters queue [rates]
arista# show hardware counter drop *! not counted at the interface*

cisco# show policy-map interface …
cisco# show logging *! for config errors*

juniper# show class-of-service interface … comprehensive


QOS DOES NOT HARM
QUESTIONS?
imgflip.com