Traffic Engineering 2: Layer 2 Prioritisation - CoS (Class of Service)

Local area networks of any significant size, which almost certainly encompasses all those at educational organisations, are complex and unpredictable systems. The traffic flows produced within these networks, and the interactions between different flows within network components such as switches, are highly complex. Classifying, policing and priority queuing allow the network administrator some control over how these flows transit the network, and – crucially for voice and video traffic – allow time-critical traffic to have priority over other, less time-sensitive traffic.

In cases where physical or logical traffic segregation is not possible, or not adequate in the case of VLANs alone, the traffic prioritisation (or QoS) features available on most Layer 2 and Layer 3 network equipment can be used to provide reasonably robust traffic flows through a normally congested network. It should be noted, however, that this will help only where traffic is oversubscribing an output interface on a piece of equipment. If the traffic level is causing high CPU or memory usage, applying QoS may not help at all.

The most common bandwidth problems in the LAN arise where either high speed core links, frequently now at gigabit speeds, break out to lower speed links to departments or when multiple links into a switch feed into a single uplink towards the core. It is at these points, where the link speed falls or where many links are aggregated into one, that prioritisation will ensure that the time-critical traffic will receive priority treatment and not have to sit in queues. The effect of traffic overload in these scenarios is either an increase in latency, or, if the latency becomes large enough, the packets risk being dropped off the tail of the queue, resulting in packet loss.

In order to configure QoS, there are a number of elements or processes that need to be considered – Classification and Marking, Policing and Queuing.

Classification and Marking

In order to treat some traffic in the switch differently to the rest of the traffic flowing through it, it is first necessary to divide the traffic up into different ‘Classes’. Depending on the equipment in your network, it may be possible to select eight or more different classes of traffic and treat them all differently. However, in practice, due to the limitations of most switches currently in circulation, it is normal to only consider two or four classes. This also makes configuration, monitoring and control far simpler.

Once the traffic has been classified it must be marked, via one of the standard methods, in order to indicate to later processes within the switch and to next-hop switches that this is the traffic that should receive some specific treatment; in this case prioritisation over non-H.323 traffic.

Traffic is normally classified by explicitly configuring the switch to select certain traffic flowing through it. It may be that all ingress traffic on a certain switchport is selected, or all traffic to and/or from a certain MAC address. Increasingly, Layer 2 switches are also adding Layer 3 functionality into their code, so in many cases with newer switches such as the 3Com® 4400s and Cisco® 3550s it is possible to select traffic by Layer 3 identifiers such as source and/or destination IP address.

Example.

To classify traffic on a Cisco® 2950 switch the following commands could be used:

```
Class-map match-all VIDEO
 Match access-group 190
!
!
access-list 190 permit ip host 10.1.1.1 any
!
```

[1]

In the above case a class-map called VIDEO is created into which is placed all traffic that matches the constraints of the access-list – i.e. all IP traffic from host 10.1.1.1.

Once the traffic has been put into classes, it is marked to enable treatment appropriate to its classification. Traffic is usually classified and marked on ingress to a port, but prioritisation, i.e. queuing, is only applied on egress from a port. Marking traffic at ingress allows the egress port to identify and prioritise traffic. In our Cisco® 2950, marking is achieved through the use

Example: Marking Traffic

```
Class map match all VIDEO
 Match access-group 190
!
policy-map INCOMING
 class VIDEO
 set ip dscp 46
!
access-list 190 permit ip host 10.1.1.1 any
!
```

[2]

In this case, adding to the above class-map example, a policy-map called INCOMING is created in which the class VIDEO has its DSCP value set to 46.

So in Cisco® terms – and this applies to Cisco® Layer 3 devices as well as switches – the normal course of action is to create one or more CLASS-MAPs which classify the traffic and then apply a POLICY-MAP to attach a defined policy to the classes created in the CLASSMAPs.

There are a number of ways that traffic can be marked, as defined by standards that have now become well established. The two standards most frequently used are the 802.1p CoS (Class of Service) and DSCP (Differentiated Services Code Point). At Layer 2, strictly speaking, only the 802.1p values are used; however, as was mentioned above, many new

Layer 2 switches are able, at least partly, to parse Layer 3 headers in order to extract or match certain information.

At the moment all that has been done is to specify a policy to mark the traffic with DSCP value 46. DSCP 46 is defined as being for EF (Expedited Forwarding) traffic flows and is the value usually assigned to all interactive voice and video traffic.

So, whereas previously it would have been necessary to mark all traffic in the LAN with 802.1p priority values from 0 to 7, and then remark traffic at the Layer 2 to Layer 3 boundary to DSCP values, it is now possible to use DSCP values across the LAN as well as the WAN, depending on the equipment in the network path.

At Layer 3 the IP Precedence value was previously used; DSCP is designed to replace it. However, some equipment at Layer 3 will only mark with IP Precedence values which, like 802.1p values, range from 0 through 7.

The 802.1p standard defines seven levels of CoS from 0 through to 7 (highest priority). 802.1p is a sub-set of the 802.1q standard which added additional fields into the header of a standard Ethernet frame allowing it to contain VLAN identifiers as well as the priority values.

DSCP defines 64 values (0-63) which can be used to treat traffic in a very granular manner, assuming the equipment has enough queues to support queuing at large numbers of different rates. However, most switches support only 2, 4 or 8 queues per output port so DSCP values in certain ranges, as well as multiple 802.1p priority values, will tend to map to a smaller range of queues, as shown in table 2 (below).

Table 2: 3Com® 4400 CoS to queue map

Priority	**Queue Index**	**802.1p Name**
0	1 (lowest)	Best Effort
1	1 (lowest)	Background
2	1 (lowest)	Reserved
3	2 Excellent	Effort
4	2 Controlled	Load
5	3	Video

6	4 (highest)	Voice
7	4 (highest)	Network

It is very important, having decided to run QoS at Layer 2 or Layer 3, that a policy is designed to specify which traffic flows will be marked, what value and scheme they will be marked with, and what sort of treatment the flow requires. The earlier in the process this ‘design’ is formulated, the simpler later configuration becomes.

It is also clear from the above that it is possible to configure all switches to classify and mark frames/packets on ingress. However, most large networks (at least in some areas) consist of edge switches which connect hosts, and core switches which only connect switches. It is critical that all the edges of the network are treated the same – which will consist of marking traffic in a consistent manner at all ingress points. If this is done at all the edges, it is no longer necessary to re-classify/police that traffic on ingress to next-hop switches as it moves closer to the core of the network. This can simplify configuration in core switches as they will only be required to queue traffic previously marked by other switches.

Again, as with designing a marking policy, designing a trust policy early on can simplify QoS management and administration later. It is possible to run QoS on one ‘corner’ of the network and gradually expand it outwards, but the ingress point to this ‘corner’ will need to have classification applied to prevent spuriously marked traffic from interfering with deliberately marked traffic.

It has been found that traffic in a number of networks is already being marked. Applications running on all current versions of desktop and server operating systems have the ability to mark traffic leaving their host, and some do. Most commonly these are voice- or video-based applications, but just as easily it could be game or denial-of-service applications that use it. It is highly advisable, therefore, not only that organisations classify and mark traffic that is explicitly selected for better treatment, but also that they ensure that other traffic entering the switch has not already been marked by an external host or application. It is possible to re-mark all traffic, except selected traffic, to CoS/DSCP/IP Precedence 0; or at least to re-mark to 0 traffic that is trying to use other CoS values. So for example, either all traffic entering the switch is re-marked to CoS 0 (except for H.323 traffic from certain IP/MAC addresses or from a certain switchport); or all traffic marked as CoS 5 is re-marked to CoS 0, except for relevant H.323 traffic. This way an organisation can be sure that the only traffic entering their Premium flow is traffic that they specifically want to allow.

A Note on Marking

At Layer 2 it is frequently the case that it is only necessary, or possible, to configure ingress marking to the switch and the priority treatment is pre-defined: in some cases behaviour can be changed, in others it is fixed. As an example the Cisco® 3524 switch allows no configuration at all; incoming packets with CoS value 4-7 will be prioritised at egress above those marked with CoS 0-3.

Other switches may have four queues and define the prioritisation treatment differently, e.g.

CoS 5 in one switch may be the equivalent of CoS 6 in another. This can become more complex if DSCP and even IP Precedence values are all mixed in, and even different switch models from the same manufacturer may map 802.1p priority values into different output queues. Fortunately many switches also now allow these mappings to be changed in order to suit your specific needs.

As mentioned above, defining your classification scheme early on will simplify matters. At the simplest level this would just be to define which values should be applied to which types of traffic, as shown in table 3 (below).

Table 3: Sample classification scheme

Traffic Type	**Description**	**CoS value**	**DSCP Value**	**IP**
Best Effort	All web, e-mail etc	0	0	0
Premium IP	H.323 and VoIP	5	46	5

Policing

If you imagine the case where priority queuing is applied without any restriction, then a host, or hosts, which managed to inject the full link bandwidth, suitably marked as Premium traffic, could completely starve the Best Effort flow of any bandwidth. This is obviously not desirable, and policing allows the network administrator to limit the amount of traffic entering the Premium flow at a given point in the network. Policing could also re-mark traffic, either from unauthorised hosts or from a host that was trying to inject too much traffic into the Premium class. This must of course be applied carefully, as policing a flow and not allowing adequate bandwidth for all possible Premium flows that may be set up will severely impact all those flows – causing packet loss and hence degradation in video or voice applications. Policing, unfortunately, is not available on all switches.

In many circumstances, traffic shaping is often applied in conjunction with policing. For H.323, traffic-shaping is not recommended as, given the nature of UDP traffic, there is no mechanism, and no time, for a retransmission of traffic should it be dropped by a shaper. Shaping should only be applied to UDP voice and video streams where an adequate buffer size can be allocated to the shaper to prevent packet loss from occurring due to the shaping.

Queuing

Once the traffic has been selected, marked and policed, it must be treated differently in some manner in the switch. Usually this is effected by putting different classes of traffic into different egress queues at the output port. The queues are then emptied by the port onto the link in deference to their priority, so high priority queues will be served by the port more frequently than low priority ones. The frequency and nature of the difference between queues can be fixed in a switch, or on larger switches may be fully configurable.

Some options that may be encountered are:

- Strict Priority Queuing. Some traffic is selected and put into a strict priority queue. This queue is always serviced by the switch if there is any traffic in it, irrespective of any other traffic in any other queues. This can mean that if there is enough traffic in the strict priority queue, other traffic can be completely starved of bandwidth.
- Priority Queuing (with bandwidth limit). The use of strict priority queuing (above) is far more useful if it provides a mechanism to police the bandwidth allowed into the priority queue. Voice and video traffic will normally use a priority queue with a limit to prevent starvation of other flows, whilst protecting adequate bandwidth for the calls.

Note that while the 802.1p standard specifies eight different priority levels, many switches, including many high-end ones, simplify the choices available, sometimes giving only a high and low priority queue.

There are also often many options available for drop-precedence in queues, allowing traffic in certain classes to be dropped before others. In most cases, as the aim is to provide the best possible performance for the H.323 traffic, these will not apply. In general the default behaviour for the Best Effort class should be satisfactory.

On most edge switches, such as the Cisco® and 3Com® switches, there is a default setting for queues, and frames/packets marked with a certain CoS or DSCP value will be placed into the appropriate queue and shipped out onto the network.

As has been mentioned above, the hard part is not getting a particular piece of network equipment to prioritise traffic, but rather designing a scheme that will suit the network, and which will cover all the capabilities of the various switches in that network. Simplifying provisioning to four, or fewer, classes makes this process easier and is highly recommended in the first instance.

Table 4 (below) shows detail from a 3Com® 4400 switch which has four independent queues, and the default mapping between CoS value and queue:

Table 4: Current CoS to queue mappings from switch command line (from: trafficManagement|qos|trafficQueue|Summary)

Priority CoS Value	**Queue Index**
0	1 (lowest)

1	1 (lowest)
2	1 (lowest)
3	2
4	2
5	3
6	4 (highest)
7	4 (highest)

Configuration Samples

The following sections look at specific configurations on Cisco® 3524, Cisco® 2950 and 3Com® 4400 switches. It is only possible here to cover a fraction of the options available. Please see the Further Reading section for links to more comprehensive documentation.

Applying QoS on a Cisco® 3524 Switch

The standard Cisco® 3524 switch has two available queues – high and low. By default, incoming traffic to the switch which is tagged with a CoS of 0-3 will be placed in the low-priority queue. Traffic with CoS 4-7 will be placed in the high-priority queue. This behaviour is not configurable and is enabled by default. Traffic entering the switch which is untagged can be marked with a CoS value.

The issue with this, of course, is that any host attached to a Cisco® 3524 switch can put priority marked traffic into that switch and it will be honoured, above other traffic. This could be used as a denial-of-service mechanism.

Note that the Cisco® 3524-PWR-XL switch with Software version 12.0(5)XU or higher does support CoS remarking (Flannagan et al, 2003).

Applying QoS on a Cisco® 2950 Switch

QoS is not enabled by default on the Cisco® 2950 switch.

On many Cisco® switches there is the option of using basic commands (e.g. mls qos trust

dscp) or using the Cisco® MQC (Modular QoS Command line). In general the MQC is preferable as it is useable not only across much of the Layer 2 range but also across the Layer 3 routers as well.

This code sample (below) uses Modular QoS Command-line commands to configure priority ...ited to) 20Mbit/s. See the previous ... explanation of the commands.

```
!
Class-map match-all VIDEO
 Match access-group 190
!
policy-map INCOMING
 class VIDEO
 set ip dscp 46
 police 20000000 4096 exceed action drop
!
!
!
interface FastEthernet 0/23
 service-policy input INCOMING
!
access list 190 permit ip host 10.1.1.1 any
!
```

[3]

Applying QoS on a 3Com® 4400 Switch

Firstly, create a Classifier for your chosen traffic. The code sample below will create a Classifier numbered 323 and named VIDEO-CLASSIFIER. It will class traffic from the stated IP address – here 10.1.1.1:

```
: Traffic|qos|classifier|create
Enter Classifier Number(): 323
Enter Classifier Name:VIDEO-CLASSIFIER
Enter Classifier Type: ipaddr
Enter IP address[0.0.0.0]: 10.1.1.1
Enter number of bits in subnet mask: 32
```

[4]

Next, a QoS Profile must be created. Drop down a menu level by typing q and then:

```
: profile | create
Enter profile number: 323
Enter profile name: VIDEO-PROFILE
```

[5]

At most points entering the command ‘summary’ will show the current status of the configuration of each element.

The profile just created will include the default classifier 1. This is not required as it is the default traffic running through the switch. Remove it as follows:

```
: Remove
Select profile number: 323
Select classifier number: 1
```
[6]

Add to this (323) profile the classifier (323) that has already been created (see above) and make it use service profile 5, as follows:

```
: AddClassifier
Select profile number: 323
Select classifier number: 323
Enter service level number: 5
```
[7]

(See table 5 for the service levels available)

Finally, this profile must be applied to all necessary ports on the switch:

```
: Assign
Select ports (unit:port…,?): 1:1-24
Enter profile number: 323
```
[8]

Table 5. Service levels available in the 3Com® 4400 switch (from: trafficManagement|qos|serviceLevel|summary)

		Conforming to		
Num	**Name**	**Priority**	**DSCP**	**QoS**
1	Drop		-	Non
2	Best Effort	0	0	Non
3	Business Critical	3	16	Non
4	Video Applications	5	24	Non
5	Voice Applications	6	46 (EF)	323

6	Network Control	7	48	Nor

Summary

Priority queuing can be applied to H.323 traffic in order to increase the reliability of delivery across a campus network. Queuing and VLANs can be configured independently or together. Providing a dedicated H.323 VLAN would seem to be the best currently available method of running voice and video services reliably and consistently.

In testing the configurations in this document, the team has been pleasantly surprised by the fact that the equipment behaved exactly as expected when QoS was applied, even with severe overloading of links. Only when the edge switches were pushed to their packet switching limits, with the injection of very large numbers of small packets, did the switches have problems maintaining forwarding rates.

The complexity of configuring QoS is frequently more in the planning and administration rather than in the actual command-line configuration on the switches. In networks with homogenous equipment, it is far simpler and quicker to provision QoS than in the case where a number of different manufacturers' equipment interleaves. As QoS is strictly provisioned on a per-hop basis, there are no issues with inter-working between equipment. However, ensuring that different configurations maintain the appropriate behaviours for specific classes of service, and maintaining CoS ingress protection at all the edges of the network, can become somewhat challenging in a large network – which is often a dynamic and fluid environment.

Even with the added complexity, QoS is one of those tools that will add value in the reliability to video or voice services – and its deployment should be seriously considered.

Links
[1] http://community.ja.net/system/files/h323guide-code07.jpg
[2] http://community.ja.net/system/files/h323guide-code08.jpg
[3] https://community-stg.jisc.ac.uk/system/files/h323guide-code09.jpg
[4] http://community.ja.net/system/files/h323guide-code10.jpg
[5] http://community.ja.net/system/files/h323guide-code11.jpg
[6] http://community.ja.net/system/files/h323guide-code12.jpg
[7] http://community.ja.net/system/files/h323guide-code13.jpg
[8] http://community.ja.net/system/files/h323guide-code14.jpg