Improving Quality of Service

Using Dell PowerConnect 6024/6024F Switches

Quality of service (QoS) mechanisms classify and prioritize network traffic to improve throughput. This article explains the basic elements of QoS, focusing on how administrators can facilitate QoS with Dell™ PowerConnect™ 6024/6024F switches.

BY MARVELL SEMICONDUCTOR

Because network traffic can be so unpredictable, administrators often can provide only best-effort traffic delivery. Similar to the legal concept of a best-effort contract, best-effort delivery is considered to meet network service requirements—regardless of outcome—as long as the responsible party makes a genuine effort. However, network problems may cause loss, delay, or misdirection of traffic. Quality of service (QoS) mechanisms in network software help address this challenge by prioritizing packets to better meet required throughput levels.

Dell is helping to improve QoS on its computing platforms by adding flexibility and differentiation levels to its network switches. Dell™ OptiPlex™ and Dell Latitude™ PCs include integrated Gigabit Ethernet1 connections that increase traffic flowing through the core switches. The Dell PowerConnect™ 6024 and 6024F switches can help provide advanced QoS and traffic engineering capabilities—including eight output queues, advanced queue mapping, and easy traffic metering—to meet the growing need for more advanced QoS services. This article explores hardware and software components of QoS and their implementation in PowerConnect 6024/6024F switches.

Given today’s fast Ethernet switches operating with nonblocking multilayer switches, all at wire speed, prioritizing traffic may no longer seem necessary. However, many IT organizations are reluctant to invest in expensive new Ethernet equipment. Converged voice and data service—which requires guaranteed minimum delays and assured bandwidth—increases network congestion even further. By implementing QoS, administrators can manage network congestion more effectively.

A QoS mechanism comprises three main elements that work together:

- **Access Control Lists (ACLs):** Both network security and QoS mechanisms apply ACL rules to determine what traffic can enter a switch, under what circumstances traffic can enter a switch, and what traffic is dropped. Only traffic that meets ACL criteria is subject to QoS settings.

- **Hardware queues:** The QoS mechanism assigns each packet to a hardware queue based on the traffic class to which the packet belongs.

- **Traffic-class handling attributes:** The QoS mechanism manages different traffic classes, such as bandwidth management, shaping, and policing, based on QoS attributes for each traffic class.

1 This term does not connote an actual operating speed of 1 Gbps. For high-speed transmission, connection to a Gigabit Ethernet server and network infrastructure is required.
Policing packets controls bandwidth

Often, network traffic is asymmetrical. For example, Web users typically require ten times as much download bandwidth as upload bandwidth. Traffic policing (which is available only in advanced QoS mode) gives administrators the tools to control bandwidth precisely and deliver tiered services. Policing can be applied either to a specific traffic flow or to a group of traffic flows. In group assignment, or aggregate policing, administrators assign a traffic rate that governs all packets flowing in the designated group.

If a packet matches the predetermined profile, the policing function admits the packet to the network. If a packet does not fit the ACL profile, the policing function will either drop the packet immediately or admit the packet to the network according to the specified rules. If admitted, the packet is marked so that it will be handled differently at downstream switches; its differentiated services code point (DSCP) is modified to lower the preferred status. The only limitations on policing are those imposed by the ACLs.

Alternatively, in basic mode, the DSCP of a packet can be rewritten in accordance with a DSCP-to-DSCP mutation map—a useful capability for a switch that resides at a QoS domain boundary. For example, one domain could use a given DSCP to designate a specific service level, while the domain to which this switch belongs uses different nomenclature. Rewriting the DSCP can retain service-level definitions between QoS domains.

The DSCP-to-DSCP translation applies only to ingress ports because they are DSCP-trusted. Applying this map to a port causes IP packets to be rewritten with newly mapped DSCP values at the ingress ports.

Policing uses a simple token bucket algorithm. This algorithm adds a token to the bucket as the switch receives each packet. The rate at which the bucket fills is a function of two factors: the rate at which tokens are removed, or committed information rate (CIR), and the bucket depth, or committed burst size (CBS). Packets that meet the CIR and CBS, called conforming packets, continue through the QoS path.

Classifying packets determines service levels

Packets are first classified according to ACL rules specified by the network administrator. The classification process inspects packet fields, including IP address source and destination; IP subnet source and destination; TCP/UDP (Transmission Control Protocol/ User Datagram Protocol) port information; virtual LAN (VLAN) ID; and Media Access Control (MAC) address source and destination. Packet classification helps determine what service level will be applied to a frame as it traverses the switch, and is identical to the classification performed for network security. Thus, the packet classification process results in either dropping the packets from the switch or forwarding them according to the specified ACL rules.

In advanced QoS mode, only conforming packets go through the trusting phase; nonconforming packets undergo an administrator-defined action. In basic mode, the administrator specifies “trust” for particular network domains. Within the trusted domain, the PowerConnect 6024/6024F marks each packet based on administrator-specified fields that signal the type of service the packet should receive. These fields also are used to assign the packet to one of eight output queues. The administrator determines the trust behavior by designating the fields upon which output service assignment is performed: 802.1p tag–based2 fields, 802.1p port–based fields, Layer 3 predefined fields, and Layer 4 predefined fields.

Trusting in basic mode is a useful tool for switches at the edge of a QoS domain. In this case, the packets are classified at the edge of the domain and assigned to a queue. Sophisticated switches provide more flexible bandwidth management and control by supporting more queues. The PowerConnect 6024/6024F switch supports eight queues per port, allowing network administrators to configure eight differentiated levels of service for individuals and applications using the network.

In trusting (for either advanced or basic mode), the PowerConnect 6024/6024F switch assigns packets to a queue based on an administrator-configured map. The switch supports four different types of maps—class of service (CoS) to queue, class of service (CoS) to queue,

2 The terms class of service and 802.1p tag are used interchangeably.
By offering three quality of service modes, Dell PowerConnect 6024/6024F switches help provide network administrators with the versatility to meet various networking needs. Although best-effort and basic QoS modes each offer advantages for managing network traffic, advanced QoS mode enables administrators to use policing, trusting, marking, and queue scheduling techniques to further streamline the flow of information through a network.

The following applications and network services provide examples of the traffic needs that administrators can address by implementing advanced QoS mode rather than simply increasing bandwidth:

- **H.323 voice over IP (VoIP) and telephony**: VoIP requires limited bandwidth but is sensitive to latency and jitter. VoIP applications include both call setup and control protocols; these traffic types are transmitted over TCP using known ports. Voice streaming is transmitted over UDP using a known port.

- **H.323 video**: Video is less delay-sensitive than VoIP but requires more bandwidth. Although the call setup and control protocols are the same as for voice and video streaming, H.323 video traffic uses a different UDP port.

- **Data transfer**: For network backup, data transfer requires significant bandwidth and low packet loss but has no specific delay requirements.

Figure A shows a typical network configuration to meet H.323 VoIP and telephony, H.323 video, and data transfer requirements.

The following example shows how administrators can ensure adequate QoS for a network with several applications, each presenting its own specific requirements. The example is based on the following assumptions:

1. The H.323 protocol uses TCP port 1720 for call setup (port number assigned by the Internet Assigned Numbers Authority, or IANA, for H.323) and port 1736 for audio call protocol.
2. UDP port 2776 is used for voice streaming.
3. UDP port 5004 is used for video streaming.
4. Input is derived from a Layer 3 switch that supports differentiated services code point (DSCP) tagging.
5. The voice data stream is tagged with the DSCP value of 63. (Its default mapping is to the highest priority queue—queue 8—for any value between 56 and 63.)
6. The video data stream DSCP and the Virtual Path Terminator (VPT) tags are unknown because neither is supported by the specific originating switch or subnet. For that reason, VPT or DSCP cannot be used as the trust parameter.

**Configuring a PowerConnect 6024 switch**

The following steps describe the process to configure a PowerConnect 6024 switch for the example scenario:

1. In the graphical user interface (GUI) provided with Dell OpenManage Switch Administrator, set the system to advanced mode on the QoS Settings screen shown in Figure B. Implement the following requirements using the advanced QoS mode of the PowerConnect 6024 switch:
   - Reserve enough bandwidth by limiting all other traffic to 700 Mbps—this is enough bandwidth for a worst-case scenario, protecting the network against faulty conditions.
b. Prevent each real-time application from interfering with other real-time applications by limiting:
   • Voice data stream to 60 Mbps
   • Video data stream to 200 Mbps
   • H.323 control to 5 Mbps

The committed bandwidth is adequate for the application. However, because these applications are higher priority, administrators must protect the lower-priority applications from unexpected traffic load of the higher-priority applications.

2. On the TCP to Queue screen, map the TCP call setup and control of H.323 to queues 5 and 6. Map TCP port 1736 to queue 6; after clicking the Apply Changes button, map TCP port 1720 to queue 5. Queues 5 and 6 are lower priority than queue 8 (used for voice) and queue 7 (used for video). The mapping of voice and video follows the default DSCP mapping table.

3. On the QoS Aggregate Policer screen, create a policy aggregator to limit all non-real-time traffic to 700 Mbps. In this example, the Policer is named “agany” and assigned an ingress committed information rate (CIR) of 700,000 Kbps and an ingress committed burst size (CBS) of 1,400,000 bytes.

4. On the Add ACE to IP-Based ACL screen, create five ACLs with the associated access control entries (ACEs). An ACL is created by using the Add button to add an ACL with the first associated ACE rule. A new ACE rule is added using the main screen. The five ACLs include:
   a. Web traffic ACL: First ACE permits any traffic with IP destination port 80 (the general assigned value used for HTTP by the IANA); second ACE permits traffic to TCP port 20 (the general assigned value used for file transfer by the IANA); and third ACE permits traffic to TCP port 21
   b. All non-real-time traffic ACL
   c. Voice data stream ACL: Permits traffic with IP destination port 2776
   d. Video data stream ACL: Permits traffic with IP destination port 5004

5. On the QoS Class Map screen shown in Figure C, bind the ACLs to five class maps. In advanced QoS mode, ACLs are not bound directly to interfaces and must be included in a policy, which includes a class map.

6. On the Policy screen shown in Figure D, create a new policy using the Add button. In this example scenario, the policy is named “polmap1” and includes the five class maps. Administrators define the class maps by completing the form on this screen for each field of information. Bind all the class maps to the policy and perform the following operations on the different classes:

   a. **Video class**
      a. Limit the traffic using the Policer to 200 Mbps, reserving adequate bandwidth for the video traffic.
      b. Change the DSCP frames tag value classified by the video class to 53 (marking assumption 6). The default queue for this DSCP value is queue 7.

   b. **Voice class**
      a. Limit the traffic using the Policer to 60 Mbps, reserving adequate bandwidth for the voice traffic.
      b. Use DSCP Trust mode to set DSCP queuing policy for all voice-classified frames. The default queue for DSCP 63 is queue 8.

   c. **Control class**
      a. Limit the traffic using the Policer to 5 Mbps, reserving adequate bandwidth for the H.323 control.
      b. Use the TCP/UDP command to set the queuing policy for classified TCP/UDP frames.

   d. **Web class**
      a. Direct all classified frames to queue 3 (lower-priority queue).
      b. Use the Policer to limit traffic according to the aggregated policy.

(continued on following page)
DSCP to queue, TCP to queue, and UDP to queue—allowing packets to be prioritized based on Layer 2, Layer 3, or Layer 4 behaviors.

Marking packets assigns administrator-defined actions
If a packet does not conform to policing rules, then it is marked for further action. DSCP values of nonconforming packets marked with the administrator-defined action policed-dscp-transmit will be rewritten according to the policed-DSCP map.

Egress actions: Controlling network congestion
Network congestion occurs when packets arrive at an output port faster than they can be transmitted. Switches use two general methods for controlling congestion in the network: queue scheduling and traffic shaping.

Queue scheduling prioritizes packets
Queue scheduling allows administrators to control service-class access to a limited network resource: link bandwidth. By managing the amount of bandwidth allocated to each service class on an output port, administrators can help control network congestion.

Strict priority queuing. This type of queuing is a simple method for supporting differentiated service classes. Strict priority queuing first classifies packets by the switch and then places them in different priority queues. Packets from the highest priority queues are scheduled first. Within each queue, packets are scheduled in first in, first out (FIFO) order.

Weighted round-robin (WRR) queuing. WRR queuing addresses the limitations of the strict priority model by ensuring that lower-priority queues are not denied access to output buffer space and bandwidth. In WRR queuing, packets are first categorized into different service classes—FTP, multimedia, and voice—and then assigned to a specific queue dedicated to that particular class. Each queue is serviced in round-robin order.

WRR queuing efficiently supports the differentiated service classes for a manageable number of highly aggregated traffic flows. Administrators can implement WRR queuing in hardware and apply it to high-speed interfaces in both the core and the edge of enterprise networks.

Traffic shaping regulates packet flow
Traffic shaping smoothes packet flow and regulates the rate and volume of traffic entering the network. Traffic-shaping tools set limits on the token generator, token bucket, and packet queue length. Traffic that adheres to the token bucket parameters can be transmitted on the link, but traffic that does not conform to the administrator-specified profile remains in the queue until it does conform.

For example, a network administrator can assign a per-queue shaper and a per-port shaper. If a given flow meets the shaping criteria for a specific queue and passes through it, the flow will then be subject to the aggregate shaper on the port. Thus, packets may pass the queue shaper, but be dropped on the port shaper.

Managing quality of service
A key feature of the QoS application in PowerConnect 6024/6024F switches is ease of management. Dell OpenManage™ Switch Administrator provides an intuitive graphical user interface that enables administrators to assign services and treatment to traffic flow. An industry-standard command-line interface also helps administrators manage the switch.

QoS mechanisms help administrators solve the problems inherent in unpredictable network traffic by classifying and prioritizing incoming and outgoing packets. Using hardware that facilitates QoS, such as Dell PowerConnect 6024/6024F switches, IT administrators can help maximize their QoS implementations, relieve network congestion, and better meet guaranteed minimum delays. Dell OpenManage Switch Administrator further simplifies configuring switches for QoS.

Marvell Semiconductor (http://www.marvell.com) is a leading global semiconductor provider of complete broadband communications and storage solutions. The company’s diverse portfolio includes switch controllers and processors, transceivers, communications controllers, wireless devices, and storage products that power the communications infrastructure.

For more information
Dell PowerConnect switches: http://www.dell.com/networking