Differentiated Services

Moving towards Quality of Service on the Ethernet

Executive Summary

Throw out the 80/20 rule when it comes to network traffic in today's enterprises. The majority of traffic no longer comes from within the workgroup. IT managers are seeing as much traffic coming from outside the workgroup and over the Internet, because:

- Companies are moving mission-critical applications from local systems to enterprise-wide server farms.
- Inter-network-based global communications are increasing as corporate businesses expand offshore and technology partnerships grow.
- Users are increasingly accessing off campus databases using web-based technologies for research, product search and evaluations, and collaborative development.

In addition to traffic pattern changes like these, both volume and type of data have expanded with:

- Higher speed PCs increasing productivity and generating traffic more quickly than before
- Fast Ethernet at the desktop, Gigabit Ethernet along the backbone, and Layer 3 switching technology, moving more traffic more quickly
- Increasing use of more demanding applications, such as multimedia and voice-over IP (VOIP), requiring faster than "best-effort" service

The increased demands for inter-network access, the advancements in PC and network technologies, and convergence of different, more demanding applications stress the network's boundary routers. These devices become a severe bottleneck. They inhibit higher service quality for the more demanding applications, resulting in observable degradations at the destination, including delay of mission-critical data, jittery video, and choppy VOIP communications.

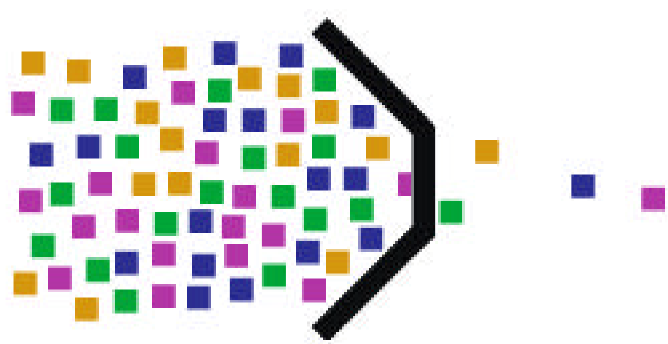


Figure 1. Increased traffic and more bandwidth-demanding applications converge at boundry routers, resulting in service degradation, caused by the router bottleneck.

On campus, IT Managers respond to service degradation with yet bigger pipes and faster equipment, often a necessary but costly alternative. Outside the campus, faster technology is not within a manager's control and comes at a significant price for access. However, prioritizing network resources to traffic requirements is within the IT managers domain. Allocating traffic resources to requirements takes managers beyond the brute force approach of adding bandwidth through more advanced technologies. It adds the realm of policy-based management and mechanisms to implement this kind of management.

Quality of Service (QoS) and Class of Service (CoS) mechanisms, such as Differentiated Services (DS), provide building blocks for policy-based network management and prioritizing resources to requirements. This article briefly describes QoS and CoS and introduces the Differentiated Services mechanism.

QoS and CoS: Service Quality Systems

Today, getting the highest network service quality is as important as best of class networking products. Network managers proactively look for better ways to service their customers. IT managers, service providers, backbone operators, and standards organizations explore and recommend service quality systems, such as QoS and CoS, for networks of today and tomorrow, while trying to maintain a migration path to newer technology that is both practical and economically feasible.

QoS

Quality of Service (QoS) mechanisms provide the necessary level of service (bandwidth and delay) to an application in order to maintain an expected quality level. To a mission-critical application, QoS means guaranteed bandwidth with zero frame loss. For a telephony application, such as voice over IP (VOIP), QoS means guaranteed frame latency. Fine-grain control provided by QoS places a significant burden on the network infrastructure. Each device must keep an entry in its forwarding table for each flow. In a large corporate network, devices can become overwhelmed with the millions of flows, especially at the boundaries.

ATM, Frame Relay, and Multi-Protocol Label Switching (MPLS) are examples of protocols that deliver level of service by application flow. Resource Reservation Protocol (RSVP) is another protocol that has gained acceptance as a valid QoS mechanism. RSVP delivers end-to-end service by reserving bandwidth and availability of resources along a particular path. RSVP is implemented in some routers, Layer 3 switches, and in Microsoft's NT 5.0 Server. However, complete end-to-end QoS implementations are difficult to achieve in today's corporate LAN environment and come at a high price.

CoS

Class of Service (CoS) mechanisms reduce flow complexity by mapping multiple flows into a few service levels. Network resources are then allocated based on these service levels, and flows can be aggregated and forwarded according to the service class of the packet. Instead of the fine grain control of QoS, not easily achievable with current technology, CoS applies bandwidth and delay to different classes of network services. CoS easily scales with network expansion. As the network grows, traffic continues to be managed based on a few service levels, keeping infrastructure burdens to a minimum.

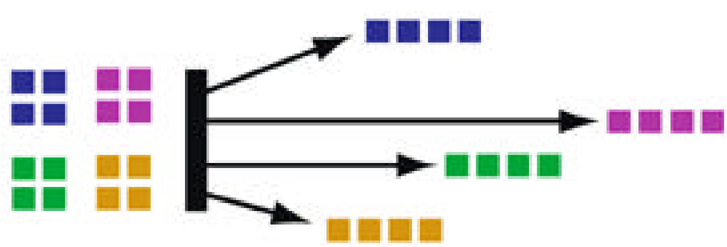


Figure 2. CoS maps multiple flows to a few service levels, reducing network burdens while increasing service quality. CoS easily scales with network expansion.

Class of Service and Differentiated Services

Two common CoS mechanisms are IEEE 802.1p tagging (Layer 2) and Type of Service (ToS) prioritization (Layer 3). Both provide prioritization but have their limitations. Differentiated Services (DS) is an advanced architecture of ToS; it is currently in draft stages by the IEEE's Internet Engineering Task Force (IETF).

802.1p
- Priority Based
- Adds 16 bits to the Layer 2 header
- Specifies 6 different priorities
- Costly upgrades
- Instability in older networks
- Layer 2 only

ToS
- Priority Based
- Layer 3 header
- 3 Precedence bits
- Specifies 7 different priorities

Differentiated Services
- Policy/Rule Based
- 6 bits wide
- (64 forwarding behaviors)
- Based on PHB's (Per Hop Behaviors)
- Backward compatible with ToS

Figure 3. Class of Service mechanisms provide service level classifications by marking packets and frames.

802.1p
IEEE 802.1p adds 16 bits to the Layer 2 header, including three bits that can be used to classify priority (the tag). Frames with 802.1p implementation are called "tagged frames". The standard specifies six different priorities, which do not offer extensive policy-based service levels. In addition, implementing 802.1p in a network with non-802.1p switches could lead to instability, because older switches would misinterpret the unexpected 16 bits specified by the standard. Implementing 802.1p in older networks could require a costly upgrade of all switches. Most importantly, since the Layer 2 header is only read at the switch level, the boundary routers, where the bottlenecks occur, cannot take advantage of prioritization based on 802.1p unless it is mapped to a Layer 3 prioritization scheme. While prioritization is achieved within the switched network, it is lost at the LAN/WAN boundary routers.

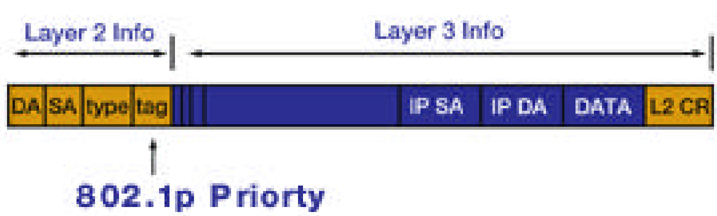


Figure 4. IEEE 802.1p adds three priority bits (tag) to the Layer 2 header.

Type of Service (ToS)
Although ToS was defined in the early 80s, it was largely unused until recent IP traffic bottlenecks at the boundary routers required prioritization for better service levels.

The Type of Service octet in the Ipv4 header includes three precedence bits defining seven different priority levels ranging from highest priority for network control packets to lowest priority for routine traffic. The remaining five bits are not implemented. Many boundary routers and ToS-enabled Layer 3 switches read the precedence bits and map them to forwarding and drop behaviors.

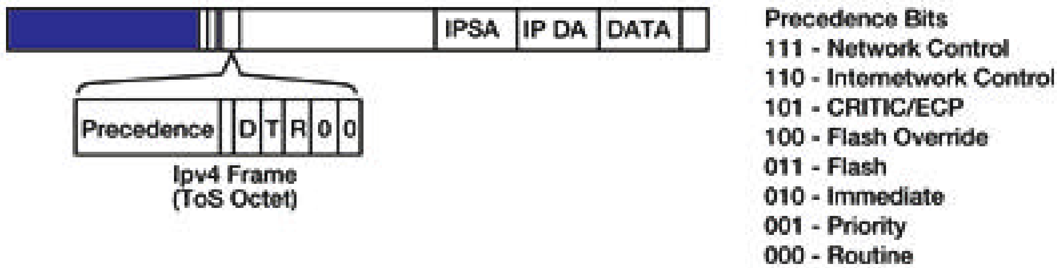


Figure 5. ToS provides 3 prioritization bits in the Layer 3 header.

Differentiated Services (DS)

In response to demand for a robust, common system for service classification, an IETF Working Group (diff-serv) is drafting a framework and definitions for a Differentiated Services (DS) mechanism. DS will not be based on priority, application, or flow, but on the possible forwarding behaviors of packets, called Per Hop Behaviors (PHBs). (Refer to Appendix A – DSCP and PHBs)

DS is rule based. Therefore, it is a unique mechanism for policy-based network management. Instead of applying faster, more expensive, advanced technology, networks can be managed by appropriate network policies, applying current network technology and resources while considering in-house traffic and upstream and downstream networks, whether they are the corporate LAN backbone or external WANs.

DS is analogous to consumer-based differentiated service industries, such as travel services. A person can travel by bus, train, or airplane; 1st class, business class, coach, or standby. Each class of service can be characterized by how fast you reach your destination, how many stops you make along the way, and what kind of service amenities you receive enroute, if any. Some services may have limitations, such as when you can travel, and others, such as standby, include risk of not reaching your destination in the time frame expected. In all cases, you pay more for higher quality services.

The Differentiated Services framework offers the same kind of classification system. Based on network policies, different kinds of traffic can be marked for different kinds of forwarding. Resources can then be allocated according to the marking and the policies.

Just as a travel ticket "encodes" your travel service level, in the DS architecture,

- the IP header includes a Differentiated Services Code Point (DSCP), indicating the level of service desired;
- the DSCP maps the packet to a particular forwarding behavior (PHB) for processing by a DS-compliant router;
- the PHB provides a particular service level (bandwidth, queuing, and dropping decisions) in accordance with network policy.

For example, mission-critical packets could be encoded with a DSCP that indicated a high bandwidth, 0-frame-loss routing path. Interactive video conferencing data and all data from the CEO's computer may carry the same requirement and be aggregated with mission-critical packets. E-mail and web browsing data could be coded with a DSCP indicating routine traffic handling with minimal packet drops. The DS-compliant boundary router would then make route selections and forward the packets accordingly as defined by network policy and the PHBs the network supports. The highest class traffic would get preferential treatment in queuing and bandwidth while the lower class packets would be relegated to slower service.

The DSCP is six bits wide, allowing coding for up to 64 different forwarding behaviors. The DSCP replaces the ToS bits, and it retains backward compatibility with the three precedence bits so that non-DS compliant, ToS-enabled devices will not conflict with the DSCP mapping.

Differentiated Services Implementation

DS requires investment from network managers. Whether IT managers in the enterprise, ISP administrators, or backbone operators, each manager will:

- define supported service levels and forwarding rules for the network in network policies, inter-domain Service Level Agreements (SLAs), and Traffic Conditioning Agreements (TCAs)
- implement DS mechanisms, including network devices, such as DS-compliant routers, protocols, and rule management entities
- monitor and maintain service quality both across their respective domains and at the interfaces between upstream and downstream domains

A set of contiguous DS-compliant networks containing DS-compliant nodes is a DS domain. Service agreements between networks define service level commitments, traffic flow constraints, and inter-network DSCP marking rules. At the different network boundaries in the domain, ingress and egress routers assure inbound and outbound packets are accordingly marked and that traffic complies with the Service Level Agreements and Traffic Conditioning Agreements.

Each network's marking rules would determine the DSCP to be applied to each packet. Packet marking might be done by DS-compliant NICs, by a traffic classifier entity in the network, or at the first hop router in the network. However, defining marking rules and implementing them at the NIC will reduce latency at routers and increase overall traffic flow through bottlenecks. A default DSCP, mapping to today's "best effort" forwarding, will be applied if rules are not defined or the NICs are not set up to mark the packet entering and processed by the network

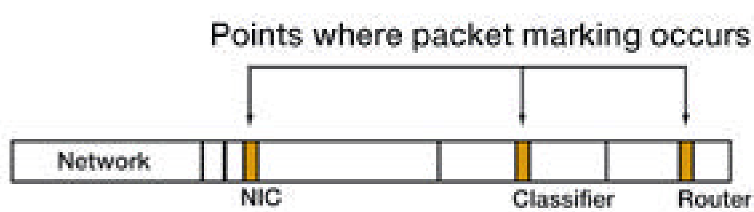


Figure 6. Setting the initial DSCP may occur at different nodes in the network.

In the DS-compliant network and at boundary routers, forwarding would follow the DSCP prioritization rules. Routers and DS-compliant Layer 3 switches may make delay/drop decisions on lower class packets in order to assure delivery of higher-class packets and, if available, may select faster, higher quality routes for the highest classifications. Marking entities can change the DSCP as packets enter or leave the network based on the upstream and downstream marking rules defined in the SLA and TCA (refer to Appendix B – The Traffic Conditioner), ensuring packet marking integrity within the network. Differentiated Services may trigger accounting mechanisms at network boundaries to track each service usage for quality level monitoring and billing purposes.

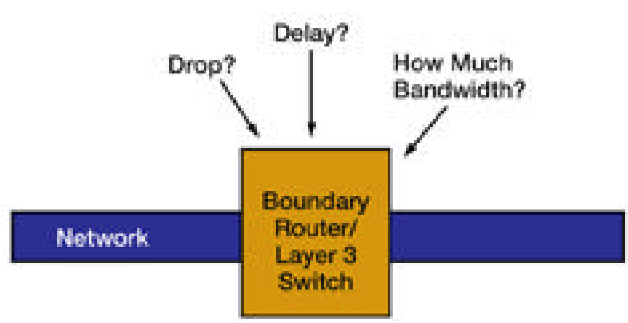


Figure 7. Routing decisions include delay/drop behaviors and bandwidth selection.

Conclusion: Differentiated Services, Moving Toward QoS

Increasing demand for inter-network access, technology advancements at both the desktop and in the network, and increasing use of multi-media and other interactive technologies are burdening the core network devices and LAN/WAN boundary routers in busy corporate networks. Service quality will continue to suffer without QoS mechanisms or CoS traffic classification schemes, such as Differentiated Services.

In today's corporate LAN environment, QoS is a costly, difficult implementation. Differentiated Services can alleviate bottlenecks by more efficient management of current corporate network resources.

DS is a policy-based management tool for networks. By mapping multiple flows to a few service levels, Differentiated Services reduces the burden on network devices and easily scales with network growth. DS is backward compatible with the existing Layer 3 ToS prioritization scheme, allowing administrators to mix DS-compliant devices with existing ToS-enabled equipment. Differentiated Services expands the prioritization scheme of ToS and allows forwarding decisions to be made based on network policies that are other than priority of traffic.

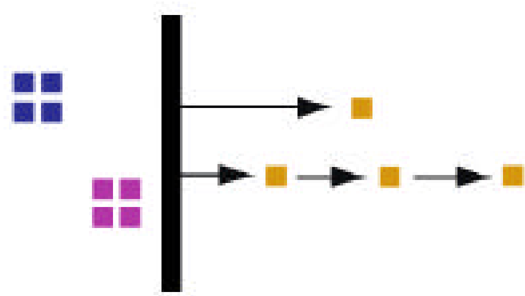


Figure 8. Differentiated Services can alleviate core network and boundry router bottlenecks by better managing high priority traffic.

Using Differentiated Services, an IT manager can move network service quality closer to QoS without adding unmanageable network burden and costly system upgrades.

Appendix A – DSCPs and PHBs

The Differentiated Services Code Point (DSCP) octet, in the IP header, classifies the packet service level. The DSCP maps to a particular observable forwarding behavior called a Per Hop Behavior (PHB).

The DSCP replaces the ToS Octet in the Ipv4 header and the Class Octet in the Ipv6 header. Currently, only the first six bits are used. Two bits of the DSCP are reserved for future definitions. This allows up to 64 different classifications for service levels.

The DSCP is unstructured, but it does reserve some values to maintain limited backward compatibility with the precedence bits in the ToS octet, because some systems do use these bits for controlling traffic. The DTR bits will not be supported. The DSCP will support standardized meanings, mapping to standardized PHBs, and also allow values to support local policies.

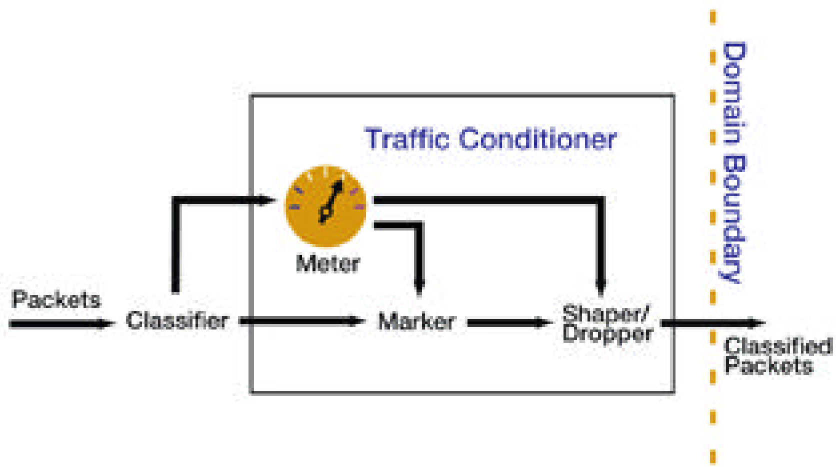
Within the DS domain and at the egress and ingress boundaries classifier entities check packet markings and change them if necessary to indicate the service level the packet should receive according to the current network service policy (inbound traffic) or for the downstream DS compliant network (outbound

Appendix B – The Traffic Conditioner

Implementing DS for network traffic management may require agreements, called Traffic Conditioning Agreements (TCAs), between the enterprise and its network service providers. The TCA

- describes traffic characteristics, called traffic profiles, for the traffic flow being sent into the WAN and coming from the WAN
- defines policies governing the amount and kind of traffic that can be sent into and received from the WAN connection
- lists the Differentiated Services (DS) rules to follow in order to adhere to the TCA, such as re-marking the DSCP or dropping packets

A Traffic Conditioner mechanism, usually at the network boundary, ensures the traffic flow follows the TCA rules.



Using meters, markers and shapers/droppers, as necessary, the Traffic Conditioner ensures the TCA is adheres to.

A classifier reads the DSCP or a combination of DSCP and other fields (source IP, destination IP, etc.) and selects and routes packets to a Traffic Conditioner.

Like a police officer, a Traffic Conditioner may be needed to monitor and control the traffic flow into the downstream DS domain by:

- Monitoring the temporal traffic flow of each packet stream (meter) to see if it is within the required profile and triggering re-marking, dropping, or shaping functions, if it is out of profile
- Changing the DSCP, if necessary, to change the forwarding behavior (marker)
- Changing the queuing management (shaper)
- Or dropping the packet, if allowed (dropper)

For example, the classifier may read a packet stream indicating "Gold" service and route the packets to a Traffic Conditioner. If the current Gold traffic falls within the traffic profile characteristics, the packet may be aggregated with other packets on a high priority route and sent immediately to the next hop.

However, if the packet stream is out of profile, the packet DSCP may be changed to map the packet to a "Silver" service level, or the packet may be re-queued or dropped altogether, according to the rules of the TCA. Alternatively, "lower class" packets may by re-marked or dropped, according to the policy.

A Traffic Conditioner may not include all these functions. Only the functions needed to adhere to the TCA policies would be implemented.