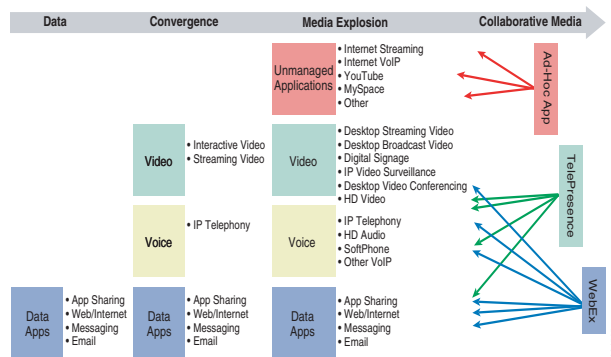


The Quality of Service Challenge for Medianets

Today there is a virtual explosion of media applications on the IP network with many different types of voice, video, and data applications. For example, voice streams can be standard IP Telephony, high-definition audio, Internet VoIP, etc. Similarly, there are many flavors of video, including on-demand or broadcast desktop video, low-definition interactive video (such as webcams), high-definition interactive video (such as Cisco TelePresence), IP video surveillance, digital signage, and entertainment-oriented video applications. In turn, there is a virtually limitless number of data applications. Additionally, beyond the current media explosion, many new applications are integrating numerous types of media streams into end-user applications. The evolution and trends in media applications are illustrated in Figure 1.

Figure 1 Media Application Convergence Evolution



The explosion of content and media types, both managed and un-managed, as well as highly-integrated collaboration applications, requires network architects to take a new look at their Quality of Service (QoS) strategies. Without a clear strategy, the volume and complexity of today's media applications on the IP network could very well exceed the ability of the network administrator to provision and manage.

Cisco Medianet QoS Design Strategy

Each group of media applications that has unique traffic patterns and service level requirements requires a dedicated QoS class in order to provision and guarantee service level requirements. There's simply no other way to

make service level guarantees. This fundamental QoS requirement leads one to ask how many classes of media should be provisioned and how should these individual classes be marked and treated?

To this end, Cisco advocates following relevant industry standards and guidelines whenever possible, as this extends the effectiveness of your QoS policies beyond your direct administrative control.

That being said, it may be helpful to overview a relevant RFC for QoS marking and provisioning, namely RFC 4594, "Configuration Guidelines for DiffServ Service Classes."

These guidelines are to be viewed as industry best-practice recommendations. As such, enterprises and service providers are encouraged to adopt these marking and provisioning recommendations, with the aim of improving QoS consistency, compatibility, and interoperability. However, since these guidelines are not standards, modifications can be made to these recommendations as specific needs or constraints require. To this end, to meet specific business requirements, Cisco has made a minor modification to its adoption of RFC 4594, namely the switching of Call-Signaling and Broadcast Video markings (to CS3 and CS5, respectively). A summary of Cisco's implementation of RFC 4594 is presented in Figure 2.

Figure 2 Cisco Differentiated Services (DiffServ) QoS Recommendations for Medianets

Application Class	Per-Hop Behavior	Admission Control	Queuing and Dropping
VoIP Telephony	EF	Required	Priority Queue (PQ)
Broadcast Video	CS5	Required	(Optional) PQ
Real-Time Interactive	CS4	Required	(Optional) PQ
Multimedia Conferencing	AF4	Required	BW Queue + DSCP WRED
Multimedia Streaming	AF3	Recommended	BW Queue + DSCP WRED
Network Control	CS6		BW Queue
Signaling	CS3		BW Queue
Ops/Admin/Mgmt (OAM)	CS2		BW Queue
Transactional Data	AF2		BW Queue + DSCP WRED
Bulk Data	AF1		BW Queue + DSCP WRED
Best Effort	DF		Default Queue + RED
Scavenger	CS1		Min BW Queue

RFC 4594 outlines twelve classes of media applications that have unique service level requirements:

- VoIP Telephony**—This service class is intended for VoIP telephony (bearer-only) traffic (VoIP signaling traffic is assigned to the "Call-Signaling" class). Traffic assigned to this class should be marked EF. This class is provisioned with an Expedited Forwarding (EF) Per-Hop Behavior (PHB). The EF PHB defined in RFC 3246 is a strict-priority queuing service and, as such, admission to this class should be controlled (admission control is discussed in the following section). Example traffic includes G.711 and G.729a.
- Broadcast Video**—This service class is intended for broadcast TV, live events, video surveillance flows, and similar "inelastic" streaming video flows ("inelastic" refers to flows that are highly drop sensitive and have no retransmission and/or flow control capabilities). Traffic in this class should be marked Class Selector 5 (CS5) and may be provisioned with an EF PHB; as such, admission to this class should be controlled. Example traffic includes live Cisco Digital Media System (DMS) streams to desktops or to Cisco Digital Media Players (DMPs), live Cisco Enterprise TV (ETV) streams, and Cisco IP Video Surveillance.
- Real-Time Interactive**—This service class is intended for (inelastic) room-based, high-definition interactive video applications and is intended primarily for voice and video components of these applications. Whenever technically possible and administratively feasible, data sub-components of this class can be separated out and assigned to the "Transactional Data" traffic class. Traffic in this class should be marked CS4 and may be provisioned with an EF PHB; as such, admission to this class should be controlled. An example application is Cisco TelePresence.
- Multimedia Conferencing**—This service class is intended for desktop software multimedia collaboration applications and is intended primarily for voice and video components of these applications. Whenever technically possible and administratively feasible, data sub-components of this class can be separated out and assigned to the "Transactional Data" traffic class. Traffic in this class should be marked Assured Forwarding (AF) Class 4 (AF41) and should be provisioned with a guaranteed

bandwidth queue with DSCP-based Weighted-Random Early Detect (DSCP-WRED) enabled. Admission to this class should be controlled; additionally, traffic in this class may be subject to policing and re-marking. Example applications include Cisco Unified Personal Communicator, Cisco Unified Video Advantage, and the Cisco Unified IP Phone 7985G.

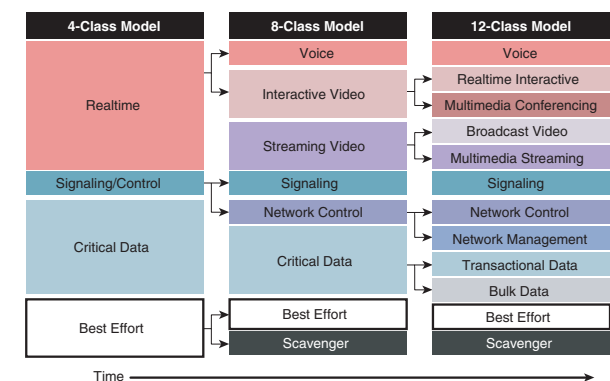
- **Multimedia Streaming**—This service class is intended for Video-on-Demand (VoD) streaming video flows, which, in general, are more elastic than broadcast/live streaming flows. Traffic in this class should be marked AF Class 3 (AF31) and should be provisioned with a guaranteed bandwidth queue with DSCP-based WRED enabled. Admission control is recommended on this traffic class (though not strictly required) and this class may be subject to policing and re-marking. Example applications include Cisco Digital Media System Video-on-Demand (VoD) streams.
- **Network Control**—This service class is intended for network control plane traffic, which is required for reliable operation of the enterprise network. Traffic in this class should be marked CS6 and provisioned with a (moderate, but dedicated) guaranteed bandwidth queue. WRED should not be enabled on this class, as network control traffic should not be dropped (if this class is experiencing drops, then the bandwidth allocated to it should be re-provisioned). Example traffic includes EIGRP, OSPF, BGP, HSRP, IKE, etc.
- **Signaling**—This service class is intended for signaling traffic that supports IP voice and video telephony. Traffic in this class should be marked CS3 and provisioned with a (moderate, but dedicated) guaranteed bandwidth queue. WRED should not be enabled on this class, as signaling traffic should not be dropped (if this class is experiencing drops, then the bandwidth allocated to it should be re-provisioned). Example traffic includes SCCP, SIP, H.323, etc.
- **Operations/Administration/Management (OAM)**—As the name implies, this service class is intended for network operations, administration, and management traffic. This class is critical to the ongoing maintenance and support of the network. Traffic in this class should be marked CS2 and provisioned with a (moderate, but dedicated) guaranteed bandwidth queue. WRED should not be enabled on this class, as OAM traffic should not be dropped (if this class is experiencing drops, then the bandwidth allocated to it should be re-provisioned). Example traffic includes SSH, SNMP, Syslog, etc.
- **Transactional Data (or Low-Latency Data)**—This service class is intended for interactive, “foreground” data applications (“foreground” refers to applications from which users are expecting a response—via the network—in order to continue with their tasks; excessive latency directly impacts user productivity). Traffic in this class should be marked AF Class 2 (AF21) and should be provisioned with a dedicated bandwidth queue with DSCP-WRED enabled. This traffic class may be subject to policing and re-marking. Example applications include data components of multimedia collaboration applications, Enterprise Resource Planning (ERP) applications, Customer Relationship Management (CRM) applications, database applications, etc.
- **Bulk Data (or High-Throughput Data)**—This service class is intended for non-interactive “background” data applications (“background” refers to applications from which users are not awaiting a response—via the network—in order to continue with their tasks; excessive latency in response times of background applications does not directly impact user productivity). Traffic in this class should be marked AF Class 1 (AF11) and should be provisioned with a dedicated bandwidth queue with DSCP-WRED enabled. This traffic class may be subject to policing and re-marking. Example applications include: E-mail, backup operations, FTP/SFTP transfers, video and content distribution, etc.
- **Best Effort (or Default Class)**—This service class is the default class. The vast majority of applications will continue to default to this Best-Effort service class; as such, this default class should be adequately provisioned. Traffic in this class is marked Default Forwarding (DF or DSCP 0) and should be provisioned with a dedicated queue. WRED is recommended to be enabled on this class.
- **Scavenger (or Low-Priority Data)**—This service class is intended for non-business related traffic flows, such as data or video applications that are entertainment and/or gaming-oriented. The approach of a “less-than Best-Effort” service class for non-business applications (as opposed to shutting these down entirely) has proven to be a popular, political compromise. These applications are

permitted on enterprise networks, as long as resources are always available for business-critical voice, video, and data applications. However, as soon as the network experiences congestion, this class is the first to be penalized and aggressively dropped. Traffic in this class should be marked CS1 and should be provisioned with a minimal bandwidth queue that is the first to starve should network congestion occur. Example traffic includes YouTube, Xbox Live/360 Movies, iTunes, BitTorrent, etc.

A Phased Approach

While RFC 4594 outlines a 12-class model, Cisco recognizes that not all enterprises are ready to implement such a complex QoS design. Therefore, Cisco recommends a phased approach to media application class expansion, as illustrated in Figure 3.

Figure 3 Medianet Application Class Expansion



Administrators can incrementally deploy QoS policies across their medianets in a progressive manner, utilizing such a phased approach to application class expansion to keep in step with their evolving business needs.

Summary

The explosion of new media applications presents complicated service level requirements of enterprise networks. However Cisco's approach—based on RFC 4594—does much to simplify QoS designs to meet the evolving service level needs of medianets. For more information, see:

http://www.cisco.com/en/US/docs/solutions/Enterprise/WAN_and_MAN/QoS_SRND_40/QoSIntro_40.html.