Compatibility with IP Precedence

- The set of codepoints 'xxx000' can be used, irrespective of bits 3-5 of the DSCP field, to yield a network that is compatible with historical IP Precedence use.
- Thus, for example, codepoint '011010' would map to the same PHB as codepoint '011000', that is, when used as an old IP precedence behavior, only three bits of the DSCP are relevant.
- Anyway, a Class Selector PHB does not necessarily require a static priority scheduling discipline.
- Other schedulers might be used to implement a CS PHB.
- This is a choice of the network administrator.
Back to relevant applications

- Applications can be divided into five broad categories, as far as their transport is concerned:
  - Application control
  - Media-oriented
  - Circuit emulation
  - Data
  - Best-Effort
Application control category

- This class of applications is referred to also as *Signaling service class*
- It is intended to be used to control applications or user endpoints
- Examples of protocols that would use this service class are SIP or H.248 for IP telephone service and SIP or Internet Group Management Protocol (IGMP) for control of broadcast TV service to subscribers
- Although user signaling flows have similar performance requirements as Low-Latency Data, they need to be distinguished and marked with a different DSCP
- The essential distinction is something like "administrative control and management" of the traffic affected as the protocols in this class tend to be tied to the media stream/session they signal and control
Media-oriented category

- Five of these classes are defined service
  - **Telephony service class** (VoIP service)
  - **Real-Time Interactive service class** is intended for inelastic video flows from applications such as SIP-based desktop video conferencing applications and for interactive gaming
  - **Multimedia Conferencing service class** is for video conferencing solutions that can adapt their transmission rate to network conditions (rate adaptive). The Real-Time Interactive service class should be used for inelastic video flows and the Multimedia Conferencing service class for rate-adaptive video flows.
  - **Broadcast Video service class** is to be used for inelastic traffic flows, which are intended for broadcast TV service
  - **Multimedia Streaming service class** is to be used for elastic multimedia traffic flows. This multimedia content is typically stored before being transmitted. It is also buffered at the receiving end before being played out.
Circuit-emulation

• The circuit emulation service (pseudowire) has the purpose of creating virtual constant-bit-rate circuits in the IP network
• The pseudowire service will be addressed in detail in the following
Data category

- The data category is divided into three service classes:
  - **Low-Latency Data** for applications/services that require low delay or latency for bursty but short-lived flows
  - **High-Throughput Data** for applications/services that require good throughput for long-lived bursty flows. High Throughput and Multimedia Steaming are close in their traffic flow characteristics with High Throughput being a bit more bursty and not as long-lived as Multimedia Streaming.
  - **Low-Priority Data** for applications or services that can tolerate short or long interruptions of packet flows. The Low-Priority Data service class can be viewed as "don't care" to some degree.
Best-Effort category

• All traffic that is not differentiated in the network falls into this category and is mapped into the default service class

• If a packet is marked with a DSCP value that is not supported in the network, it should be forwarded using the default service class
ITU-T categories of QoS as a function of delay

- The ITU-T G.1010 recommendation specifies a set of QoS categories, as a function of delay and application type.
- This categorization is also reported in some IETF RFCs (for example, RFC 4594).
- The four delay categories are:
  - Interactive (delay <<1 s)
  - Responsive (delay ~2 s)
  - Timely (delay ~10 s)
  - Non-critical (delay >>10 s)

<table>
<thead>
<tr>
<th>Interactive</th>
<th>Responsive</th>
<th>Timely</th>
<th>Non-critical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error tolerant</td>
<td>Error intolerant</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conversational voice and video</td>
<td>Voice/video messaging</td>
<td>Streaming audio and video</td>
<td>Fax</td>
</tr>
<tr>
<td>(e.g. Telnet, interactive games)</td>
<td>(e.g. E-commerce, WWW browsing, Email access)</td>
<td>(e.g. FTP, still image)</td>
<td>Background (e.g. Usenet)</td>
</tr>
</tbody>
</table>

Figure 2/G.1010 – Model for user-centric QoS categories
<table>
<thead>
<tr>
<th>Application category</th>
<th>Service class</th>
<th>Signaled</th>
<th>Flow behavior</th>
<th>ITU-T G.1010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Media oriented</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Application control</td>
<td>Signaling</td>
<td>N.A.</td>
<td>Inelastic</td>
<td>Responsive</td>
</tr>
<tr>
<td></td>
<td>Telephony</td>
<td>YES</td>
<td>Inelastic</td>
<td>Interactive</td>
</tr>
<tr>
<td></td>
<td>Real-time interactive</td>
<td>YES</td>
<td>Inelastic</td>
<td>Interactive</td>
</tr>
<tr>
<td></td>
<td>Multimedia conferencing</td>
<td>YES</td>
<td>Rate adaptive</td>
<td>Interactive</td>
</tr>
<tr>
<td></td>
<td>Broadcast video</td>
<td>YES</td>
<td>Inelastic</td>
<td>Responsive</td>
</tr>
<tr>
<td></td>
<td>Multimedia streaming</td>
<td>YES</td>
<td>Elastic</td>
<td>Timely</td>
</tr>
<tr>
<td></td>
<td>Low-latency data</td>
<td>NO</td>
<td>Elastic</td>
<td>Responsive</td>
</tr>
<tr>
<td></td>
<td>High-throughput data</td>
<td>NO</td>
<td>Elastic</td>
<td>Timely</td>
</tr>
<tr>
<td></td>
<td>Low-priority data</td>
<td>NO</td>
<td>Elastic</td>
<td>Non-critical</td>
</tr>
<tr>
<td></td>
<td>Standard</td>
<td>Not specified</td>
<td>Not specified</td>
<td>Non-critical</td>
</tr>
</tbody>
</table>
## Characteristics of service classes (I)

<table>
<thead>
<tr>
<th>Service Class</th>
<th>Traffic Characteristic</th>
<th>Tolerance To</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Loss</td>
</tr>
<tr>
<td>1. Network Control</td>
<td>Variable size packets, mostly inelastic short messages, but traffic can also burst (BGP)</td>
<td>Low</td>
</tr>
<tr>
<td>2. OAM</td>
<td>Variable size packets, Elastic &amp; inelastic flows</td>
<td>Low</td>
</tr>
<tr>
<td>3. Telephony</td>
<td>Variable size packets, constant emission rate inelastic and low-rate flows</td>
<td>Very</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Low</td>
</tr>
<tr>
<td>4. Signaling</td>
<td>Variable size packets, somewhat bursty Short-lived flows</td>
<td>Low</td>
</tr>
<tr>
<td>5. Multimedia</td>
<td>Variable size packets, Costant transmit interval Rate adaptive, reacts to loss</td>
<td>Low</td>
</tr>
<tr>
<td>Conferencing</td>
<td></td>
<td>Medium</td>
</tr>
<tr>
<td>6. Real-Time</td>
<td>RTP/UDP streams, inelastic Mostly variable rate</td>
<td>Low</td>
</tr>
<tr>
<td>Interactive</td>
<td></td>
<td>Low</td>
</tr>
</tbody>
</table>
## Characteristics of service classes (II)

<table>
<thead>
<tr>
<th>Service Class</th>
<th>Traffic Characteristic</th>
<th>Tollerance To</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Loss</td>
</tr>
<tr>
<td>7. Multimedia Streaming</td>
<td>Variable size packets, Elastic with variable rate</td>
<td>Low</td>
</tr>
<tr>
<td>8. Broadcast Video</td>
<td>Constant and variable rate Inelastic, non-bursty flows</td>
<td>Very</td>
</tr>
<tr>
<td>9. Low-Latency Data</td>
<td>Variable rate, bursty Short-lived elastic flows</td>
<td>Low</td>
</tr>
<tr>
<td>10. High-Throughput Data</td>
<td>Variable rate, bursty Long-lived elastic flows</td>
<td>Low</td>
</tr>
<tr>
<td>11. Standard</td>
<td>A bit of everything</td>
<td></td>
</tr>
<tr>
<td>12. Low-Priority Data</td>
<td>Non-real-time and elastic</td>
<td>High</td>
</tr>
</tbody>
</table>
# Possible selection of PHBs (IETF RFC 4594)

<table>
<thead>
<tr>
<th>Service Class</th>
<th>DSCP NAME</th>
<th>DSCP VALUE</th>
<th>APPLICATION EXAMPLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Network Control</td>
<td>CS6</td>
<td>100000</td>
<td>Network Routing</td>
</tr>
<tr>
<td>2. OAM</td>
<td>CS2</td>
<td>010000</td>
<td>OAM&amp;P</td>
</tr>
<tr>
<td>3. Telephony</td>
<td>EF</td>
<td>101110</td>
<td>IP Telephony Bearer</td>
</tr>
<tr>
<td>4. Signaling</td>
<td>CS5</td>
<td>101000</td>
<td>IP Telephony Signaling</td>
</tr>
<tr>
<td>5. Multimedia Conferencing</td>
<td>AF41</td>
<td>100010</td>
<td>H.323/V2 video</td>
</tr>
<tr>
<td></td>
<td>AF42</td>
<td>100100</td>
<td>Conferencing (adaptive)</td>
</tr>
<tr>
<td></td>
<td>AF43</td>
<td>100110</td>
<td></td>
</tr>
<tr>
<td>6. Real-Time Interactive</td>
<td>CS4</td>
<td>100000</td>
<td>Video Conferencing and Interactive gaming</td>
</tr>
</tbody>
</table>
## Possible selection of PHBs (IETF RFC 4594)

<table>
<thead>
<tr>
<th>Service Class</th>
<th>DSCP NAME</th>
<th>DSCP VALUE</th>
<th>APPLICATION EXAMPLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>7. Multimedia Streaming</td>
<td>AF31</td>
<td>011010</td>
<td>Streaming video and audio</td>
</tr>
<tr>
<td></td>
<td>AF32</td>
<td>011100</td>
<td>On demand</td>
</tr>
<tr>
<td></td>
<td>AF33</td>
<td>011110</td>
<td></td>
</tr>
<tr>
<td>8. Broadcast Video</td>
<td>CS3</td>
<td>011000</td>
<td>Broadcast TV and Live events</td>
</tr>
<tr>
<td>9. Low-Latency Data</td>
<td>AF21</td>
<td>010010</td>
<td>Client/server transaction</td>
</tr>
<tr>
<td></td>
<td>AF22</td>
<td>010100</td>
<td>Web-based ordering</td>
</tr>
<tr>
<td></td>
<td>AF23</td>
<td>010110</td>
<td></td>
</tr>
<tr>
<td>10. High-Troughput Data</td>
<td>AF11</td>
<td>001010</td>
<td>Store and forward</td>
</tr>
<tr>
<td></td>
<td>AF12</td>
<td>001100</td>
<td>Applications</td>
</tr>
<tr>
<td></td>
<td>AF13</td>
<td>001110</td>
<td></td>
</tr>
<tr>
<td>11. Standard</td>
<td>DF (CS)</td>
<td>000000</td>
<td>Undifferentiated applications</td>
</tr>
<tr>
<td>12. Low-Priority Data</td>
<td>CS1</td>
<td>001000</td>
<td>Any flow that has no BW assurance</td>
</tr>
</tbody>
</table>
## Possible conditioning and scheduling schemes

<table>
<thead>
<tr>
<th>Service Class</th>
<th>DSCP NAME</th>
<th>CONDITIONING AT DS EDGE</th>
<th>QUEUING</th>
<th>AQM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Network Control</td>
<td>CS6</td>
<td>Police using sr+bs</td>
<td>Rate</td>
<td>YES</td>
</tr>
<tr>
<td>2. OAM</td>
<td>CS2</td>
<td>Police using sr+bs</td>
<td>Rate</td>
<td>YES</td>
</tr>
<tr>
<td>3. Telephony</td>
<td>EF</td>
<td>Police using sr+bs</td>
<td>Priority</td>
<td>NO</td>
</tr>
<tr>
<td>4. Signaling</td>
<td>CS5</td>
<td>Police using sr+bs</td>
<td>Rate</td>
<td>NO</td>
</tr>
<tr>
<td>5. Multimedia Conferencing</td>
<td>AF41</td>
<td>Using two-rate</td>
<td>Rate</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>AF42</td>
<td>Three-color marker (rfc 2698)</td>
<td>Rate</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>AF43</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Real-Time Interactive</td>
<td>CS4</td>
<td>Police using sr+bs</td>
<td>Rate</td>
<td>NO</td>
</tr>
</tbody>
</table>
## Possible conditioning and scheduling schemes

<table>
<thead>
<tr>
<th>Service Class</th>
<th>DSCP NAME</th>
<th>CONDITIONING AT DS EDGE</th>
<th>QUEUING</th>
<th>AQM</th>
</tr>
</thead>
<tbody>
<tr>
<td>7. Multimedia Streaming</td>
<td>AF31</td>
<td>Using two-rate</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AF32</td>
<td>Three-color marker</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AF33</td>
<td>(rfc 2698)</td>
<td>Rate</td>
<td>YES</td>
</tr>
<tr>
<td>8. Broadcast Video</td>
<td>CS3</td>
<td>Police using sr+bs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Low-Latency Data</td>
<td>AF21</td>
<td>Using single-rate</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AF22</td>
<td>Three-color marker</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AF33</td>
<td>(rfc 2697)</td>
<td>Rate</td>
<td>YES</td>
</tr>
<tr>
<td>10. High-Troughput Data</td>
<td>AF11</td>
<td>Using two-rate</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AF12</td>
<td>Three-color marker</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AF13</td>
<td>(rfc 2698)</td>
<td>Rate</td>
<td>YES</td>
</tr>
<tr>
<td>11. Standard</td>
<td>DF</td>
<td>Not applicable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Low-Priority Data</td>
<td>CS1</td>
<td>Not applicable</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Notes on the assignment of PHBs, conditioning schemes and scheduling policies

- Conditioning:
  - "sr+bs" represents a policing mechanism that provides single rate with burst size control
  - The single-rate, three-color marker (srTCM) behavior is specified in RFC 2697, and the two-rate, three-color marker (trTCM) behavior is specified in RFC 2698
  - The PHB for Real-Time Interactive service class should be configured to provide high bandwidth assurance
  - The PHB for Broadcast Video service class should be configured to provide high bandwidth assurance
Example of a deployment scenario

- A network administrator determines that he needs to provide different performance levels, in particular:
  - Reliable VoIP (telephony) service, equivalent to Public Switched Telephone Network (PSTN)
  - A low-delay assured bandwidth data service
  - Support for current Internet services

- The network administrator's deploys the following six service classes:
  - Network Control service class for routing and control traffic that is needed for reliable operation of the provider's network
  - Standard service class for all traffic that will receive normal (undifferentiated) forwarding treatment through the network for support of current Internet service
  - Telephony service class for VoIP (telephony) bearer traffic
  - Signaling service class for Telephony signaling to control the VoIP service
  - Low-Latency Data service class for the low-delay assured bandwidth differentiated data service
  - OAM service class for operation and management of the network
### Example of deployment

<table>
<thead>
<tr>
<th>Service class</th>
<th>DSCP</th>
<th>Conditioning</th>
<th>PHB</th>
<th>Queueing</th>
<th>AQM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network control</td>
<td>CS6</td>
<td>***</td>
<td>RFC 2474</td>
<td>Rate</td>
<td>YES</td>
</tr>
<tr>
<td>Telephony</td>
<td>EF</td>
<td>sr+bs</td>
<td>RFC 3246</td>
<td>Priority</td>
<td>NO</td>
</tr>
<tr>
<td>Signaling</td>
<td>CS5</td>
<td>sr+bs</td>
<td>RFC 2474</td>
<td>Rate</td>
<td>NO</td>
</tr>
<tr>
<td>Low-latency data</td>
<td>AF21</td>
<td>Single-rate three-color marker</td>
<td>RFC2597</td>
<td>Rate</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>AF22</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AF23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>OAM</td>
<td>CS2</td>
<td>RFC 2474</td>
<td>Rate</td>
<td>YES</td>
</tr>
<tr>
<td>Standard</td>
<td>DEFAULT</td>
<td>N.A.</td>
<td>RFC 2474</td>
<td>Rate</td>
<td>YES</td>
</tr>
</tbody>
</table>
Notes on network control traffic (I)

- The Network Control service class is used for transmitting packets between network devices (routers) that require control (routing) information to be exchanged between nodes within the administrative domain as well as across a peering point between different administrative domains.
- Traffic transmitted in this service class is very important as it keeps the network operational, and it needs to be forwarded in a timely manner.
- The Network Control service class should be configured using the DiffServ Class Selector (CS) PHB, defined in [RFC2474]
- This service class should be configured so that the traffic receives a minimum bandwidth guarantee, to ensure that the packets always receive timely service (rate-based scheduling)
Notes on network control traffic (II)

- The following are examples of protocols and applications that should use the Network Control service class:
  - Routing packet flows: OSPF, BGP, ISIS, RIP
  - Control information exchange within and between different
  - LSP setup using CR-LDP and RSVP-TE

- User traffic should not use the Network Control service class

- The following are traffic characteristics of packet flows in the Network Control service class:
  - Most messages are sent between routers and network servers

- The recommended DSCP marking is CS6 (Class Selector 6)
Notes on network control traffic (III)

- The recommended Network Edge Conditioning policy is
  - At peering points (between two DiffServ networks) where SLAs are in place, CS6 marked packets should be policed, e.g., using a single rate with burst size (sr+bs)
  - CS6 marked packet flows from untrusted sources (for example, end user devices) should be dropped DiffServ network
  - Packets from users/subscribers should be dropped

- The fundamental service offered to the Network Control service class is enhanced best-effort service with high bandwidth assurance

- Since this service class is used to forward both elastic and inelastic flows, the service should be engineered so that the Active Queue Management (AQM) is applied to CS6 marked packets
Notes on OAM service class (I)

- The OAM (Operations, Administration, and Management) service class is recommended for protocols such as Simple Network Management Protocol (SNMP), Trivial File Transfer Protocol (TFTP), FTP, Telnet, and Common Open Policy Service (COPS).
- Applications using this service class require a low packet loss but are relatively not sensitive to delay.
- The OAM service class should use the CS PHB, configured to provide a minimum bandwidth assurance for CS2 marked packets.
- The OAM service class should be configured to use a Rate Queuing system.
- All flows in this service class are marked with CS2 (Class Selector 2).
Notes on OAM service class (II)

• The following applications should use the OAM service class:
  - Provisioning and configuration of network elements
  - Performance monitoring of network elements
  - Any network operational alarms

• The following are traffic characteristics:
  - Variable size packets
  - Intermittent traffic flows
  - Traffic may burst at times
  - Both elastic and inelastic flows
  - Traffic not sensitive to delays
Notes on OAM service class (III)

- The fundamental service offered to "OAM" traffic is enhanced best-effort service with controlled rate
- The service should be engineered so that CS2 marked packet flows have sufficient bandwidth in the network to provide high assurance of delivery
- Since this service class is used to forward both elastic and inelastic flows, the service should be engineered so that Active Queue Management is applied to CS2 marked packets
Notes on the telephony service class (I)

• The Telephony service class is recommended for applications that require real-time, very low delay, very low jitter, and very low packet loss for relatively constant-rate traffic sources (inelastic traffic sources)
• This service class should be used for IP telephony service
• The EF PHB has the required features, but also a well-engineered CS or AF PHB could be used for telephony
• The call admission procedure should verify that the newly admitted flow will be within the capacity of the Telephony service class forwarding capability in the network
• For VoIP (telephony) service, call admission control is usually performed by a telephony call server/ gatekeeper using signaling (SIP, H.323, H.248, MEGACO, etc.) on access points to the network
• The bandwidth in the core network and the number of simultaneous VoIP sessions that can be supported needs to be engineered and controlled so that there is no congestion for this service
Notes on the telephony service class (II)

- The Telephony service class should use Expedited Forwarding (EF) PHB, but also other PHBs can be used if well engineered, for some applications like VoIP, but not for pseudowire.
- The Telephony service class should be configured to use a Priority Queuing system.
- The following applications should use the Telephony service class:
  - VoIP (G.711, G.729 and other codecs).
  - Voice-band data over IP (modem, fax).
  - T.38 fax over IP.
  - Circuit emulation over IP, virtual wire, etc.
- The following are traffic characteristics:
  - Mostly fixed-size packets for VoIP (60, 70, 120 or 200 bytes in size).
  - Packets emitted at constant time intervals.
  - Admission control of new flows is provided by telephony call server, media gateway, gatekeeper, edge router, end terminal, or access node that provides flow admission control function.
Notes on the signaling service class (I)

- The Signaling service class is recommended for delay-sensitive client-server (VoIP) and peer-to-peer application signaling.
- Telephony signaling includes signaling between IP phone and soft-switch, soft-client and soft-switch, and media gateway and soft-switch as well as peer-to-peer using various protocols.
- This service class is intended to be used for control of sessions and applications.
- The Signaling service class should be configured so that the probability of packet drop or significant queuing delay under peak load is very low.
- The Signaling service class should use the CS5 PHB.
- This service class should be configured to provide a minimum bandwidth assurance for CS5 marked packets.
- The Signaling service class should be configured to use a Rate Queuing system.
Notes on the signaling service class (II)

- The following applications should use the Signaling service class:
  - Peer-to-peer IP telephony signaling (e.g., using SIP, H.323)
  - Peer-to-peer signaling for multimedia applications (e.g., using SIP, H.323)
  - Peer-to-peer real-time control function
  - Client-server IP telephony signaling using H.248, MEGACO, MGCP, IP encapsulated ISDN, or other proprietary protocols
  - Signaling to control IPTV applications using protocols such as IGMP

- The following are traffic characteristics:
  - Variable size packets, normally one packet at a time
  - Intermittent traffic flows
  - Traffic may burst at times
  - Delay-sensitive control messages sent between two end points
Notes on the multimedia conferencing service class (I)

- The Multimedia Conferencing service class is recommended for applications that require real-time service for rate-adaptive traffic, for example video conferencing equipment (e.g. H.323) with dynamic bandwidth adjustment.

- The traffic sources in this service class have the ability to dynamically change their transmission rate based on feedback from the receiver:
  - When the receiver detects a pre-configured level of packet loss, it signals to the transmitter the indication of possible on-path congestion.
  - When available, the transmitter then selects a lower rate encoding codec.

- Typical video conferencing configurations negotiate the setup of multimedia session using protocols such as H.323.

- The bandwidth in the core network and the number of simultaneous video conferencing sessions that can be supported should be engineered to control traffic load for this service.

- The Multimedia Conferencing service class should use the Assured Forwarding (AF) PHB, configured to provide a bandwidth assurance for AF41, AF42, and AF43 marked packets.

- Multimedia Conferencing service class SHOULD be configured to use a Rate Queuing system.
Notes on the multimedia conferencing service class (II)

- The following applications should use the Multimedia Conferencing service class:
  - H.323/V2 and later versions of video conferencing applications (interactive video)
  - Video conferencing applications with rate control or traffic content importance marking
  - IP VPN service that specifies two rates and mean network delay that is slightly longer than network propagation delay
  - Interactive, time-critical, and mission-critical applications

- The following are traffic characteristics:
  - Variable size packets
  - The higher the rate, the higher the density of large packets
  - Constant packet emission time interval
  - Variable rate
  - Source is capable of reducing its transmission rate based on detection of packet loss at the receiver
Notes on the multimedia conferencing service class (III)

• General recommended DSCP marking (“where "A" < "B"”):
  - AF41 = up to specified rate "A"
  - AF42 = in excess of specified rate "A" but below specified rate "B"
  - AF43 = in excess of specified rate "B"
  - "A" approximates the sum of the mean rates and "B" approximates the sum of the peak rates

• Recommended DSCP marking when performed by H.323/V2 video conferencing equipment:
  - AF41 = H.323 video conferencing audio stream RTP/UDP
  - AF41 = H.323 video conferencing video control RTCP/TCP
  - AF41 = H.323 video conferencing video stream up to specified rate "A"
  - AF42 = H.323 video conferencing video stream in excess of specified rate "A" but below specified rate "B"
  - AF43 = H.323 video conferencing video stream in excess of specified rate "B"

• Recommended conditioning: The two-rate, three-color marker
Notes on the multimedia conferencing service class (IV)

- The fundamental service offered to "Multimedia Conferencing" traffic is enhanced best-effort service with controlled rate and delay.
- For video conferencing service, typically a 1% packet loss detected at the receiver triggers an encoding rate change, dropping to the next lower provisioned video encoding rate.
- As such, AQM should be used primarily to switch the video encoding rate under congestion, changing from high rate to lower rate, i.e., 1472 kbps to 768 kbps.
- The probability of loss of AF41 traffic must not exceed the probability of loss of AF42 traffic, which in turn must not exceed the probability of loss of AF43 traffic.
- Setting of AQM thresholds:
  - min-threshold AF43 < max-threshold AF43
  - max-threshold AF43 <= min-threshold AF42
  - min-threshold AF42 < max-threshold AF42
  - max-threshold AF42 <= min-threshold AF41
  - min-threshold AF41 < max-threshold AF41
  - max-threshold AF41 <= memory assigned to the queue.
- This configuration tends to drop AF43 traffic before AF42 and AF42 before AF41.
Notes on the real time interactive service class (I)

- The Real-Time Interactive service class is recommended for applications that require low loss and jitter and very low delay for variable rate inelastic traffic sources.
- Interactive gaming and video conferencing applications that do not have the ability to change encoding rates or to mark packets with different importance indications are such applications.
- Applications in this service class are configured to negotiate the setup of RTP/UDP control session.
- The bandwidth in the core network and the number of simultaneous Real-time Interactive sessions that can be supported should be engineered to control traffic load for this service.

- The Class Selector CS4 PHB should be used, configured to provide a high assurance for bandwidth for CS4 marked packets.
- Rate Queuing should be used.
Notes on the real time interactive service class (II)

• The following applications should use the Real-Time Interactive service class:
  - Interactive gaming and control
  - Video conferencing applications without rate control or traffic content importance marking
  - IP VPN service that specifies single rate
  - Inelastic, interactive, time-critical, and mission-critical applications requiring very low delay

• The following are traffic characteristics:
  - Variable size packets
  - Variable rate.
  - Lost packets are usually ignored by application.

• The fundamental service offered to "Real-Time Interactive" traffic is enhanced best-effort service with controlled rate and delay

• The service should be engineered so that CS4 marked packet flows have sufficient bandwidth in the network

• Normally, traffic in this service class does not respond dynamically to packet loss, thus, AQM should not be applied
Notes on the multimedia streaming service class (I)

• The Multimedia Streaming service class is recommended for applications that require near-real-time packet forwarding of variable rate elastic traffic sources (streaming audio and video, some video (movies) on-demand applications, and webcasts)

• Traffic is buffered at the source/destination; therefore, it is less sensitive to delay and jitter

• The Assured Forwarding (AF) PHB should be used, in particular, AF31, AF32, and AF33

• A Rate Queuing system should be used
Notes on the multimedia streaming service class (II)

- The following applications should use the Multimedia Streaming service class:
  - Buffered streaming audio (unicast)
  - Buffered streaming video (unicast)
  - Webcasts
  - IP VPN service that specifies two rates and is less sensitive to delay and jitter

- The following are traffic characteristics:
  - Variable size packets
  - The higher the rate, the higher the density of large packets
  - Variable rate
  - Elastic flows
  - Some bursting at start of flow from some applications

- Settings of AF thresholds
  - AF41 = up to specified rate "A"
  - AF42 = in excess of specified rate "A" but below specified rate "B"
  - AF43 = in excess of specified rate "B"
  - "A" approximates the sum of the mean rates and "B" approximates the sum of the peak rates
Notes on the multimedia streaming service class (III)

- The fundamental service offered to "Multimedia Streaming" traffic is enhanced best-effort service with controlled rate and delay
- Since the AF3x traffic is elastic and responds dynamically to packet loss, Active Queue Management should be used primarily to reduce forwarding rate to the minimum assured rate at congestion points
- The probability of loss of AF31 traffic must not exceed the probability of loss of AF32 traffic, which in turn must not exceed the probability of loss of AF33
- The following inequality should hold in queue configurations:
  - min-threshold AF33 < max-threshold AF33
  - max-threshold AF33 <= min-threshold AF32
  - min-threshold AF32 < max-threshold AF32
  - max-threshold AF32 <= min-threshold AF31
  - min-threshold AF31 < max-threshold AF31
  - max-threshold AF31 <= memory assigned to the queue
- This configuration tends to drop AF33 traffic before AF32 and AF32 before AF31
Notes on the broadcast video service class (I)

- The Broadcast Video service class is recommended for applications that require near-real-time packet forwarding with very low packet loss of constant rate and variable rate inelastic traffic sources that are not as delay sensitive as applications using the Real-Time Interactive service class.

- Such applications include broadcast TV, streaming of live audio and video events, some video-on-demand applications, and video surveillance.

- In general, the Broadcast Video service class assumes that the destination end point has a dejitter buffer, for video application usually a 2 - 8 video-frame buffer (66 to several hundred of milliseconds), and therefore that it is less sensitive to delay and jitter.

- The Class Selector CS3 PHB should be used.

- A Rate Queuing system should be adopted.
Notes on the broadcast video service class (II)

- The following applications should use the Broadcast Video service class:
  - Video surveillance and security
  - TV broadcast including HDTV
  - Video on demand (unicast) with control (virtual DVD)
  - Streaming of live audio events
  - Streaming of live video events

- The following are traffic characteristics:
  - Variable size packets
  - The higher the rate, the higher the density of large packets
  - Mixture of variable rate and constant rate flows
  - Fixed packet emission time intervals
  - Inelastic flows

- The fundamental service offered to "Broadcast Video" traffic is enhanced best-effort service with controlled rate and delay

- Normally, traffic in this service class does not respond dynamically to packet loss and AQM should not be applied
Notes on the low-latency data service class (I)

- The Low-Latency Data service class is recommended for elastic and responsive typically client-/server-based applications, requiring a relatively fast response and typically have asymmetrical bandwidth need, i.e., the client typically sends a short message to the server and the server responds with a much larger data flow back to the client.

- The most common example of this is when a user clicks a hyperlink (~ few dozen bytes) on a web page, resulting in a new web page to be loaded (Kbytes of data).

- This service class is configured to provide good response for TCP short-lived flows.
  - The Assured Forwarding (AF) PHB should be used (AF21, AF22, AF23).
  - A Rate Queuing system should be used.
Notes on the low-latency data service class (II)

• The following applications should use the Low-Latency Data service class:
  - Client/server applications
  - Web-based transactions (E-commerce)
  - Credit card transactions
  - Financial wire transfers
  - Enterprise Resource Planning (ERP) applications (e.g., SAP)
  - VPN service that supports Committed Information Rate (CIR) with up to two burst sizes

• The following are traffic characteristics:
  - Variable size packets
  - Variable packet emission rate with packet bursts of TCP window size
  - Short traffic bursts

• AQM and rate queueing should be used
Notes on the high-throughput data service class (I)

- The High-Throughput Data service class is recommended for elastic applications that require timely packet forwarding of variable rate traffic sources and, more specifically, is configured to provide good throughput for TCP longer-lived flows (the FTP protocol is a common example)
- The Assured Forwarding (AF) PHB should be used (AF11, AF12, and AF13)
- AQM and rate queuing should be used
Notes on the high-throughput data service class (II)

- The following applications should use the High-Throughput Data service class:
  - Store and forward applications
  - File transfer applications
  - Email
  - VPN service that supports two rates (committed information rate and excess or peak information rate)

- The following are traffic characteristics:
  - Variable size packets
  - Variable packet emission rate
  - Variable rate with packet bursts of TCP window size

- The recommended conditioning is the two-rate, three-color marker
Notes on the standard service class (I)

• The Standard service class is recommended for traffic that has not been classified into one of the other supported forwarding service classes in the DiffServ network domain

• This service class provides the Internet's "best-effort" forwarding behavior

• This service class may have a minimum bandwidth guarantee

• Rate Queuing should be used

• The following applications should use the Standard service class:
  - Network services, DNS, DHCP, BootP
  - Any undifferentiated application/packet flow transported through the DiffServ enabled network

• The following is a traffic characteristic:
  - Non-deterministic, mixture of everything

• The recommended DSCP marking is DF (Default Forwarding) '000000'.

• Network Edge Conditioning:
  - There is no requirement that conditioning of packet flows be performed for this service class

• The fundamental service offered to the Standard service class is best-effort service with active queue management to limit overall delay
Notes on the low-priority data service class (I)

- The Low-Priority Data service class serves applications that run over TCP and that the user is willing to accept service without guarantees.

- The following applications MAY use the Low-Priority Data service class:
  - Any TCP based-application/packet flow transported through the DiffServ enabled network that does not require any bandwidth assurances.

- The following is a traffic characteristic:
  - Non-real-time and elastic.

- Network Edge Conditioning:
  - There is no requirement that conditioning of packet flows be performed for this service class.

- The recommended DSCP marking is CS1 (Class Selector 1).

- The fundamental service offered to the Low-Priority Data service class is best-effort service with zero bandwidth assurance.

- By placing it into a separate queue or class, it may be treated in a manner consistent with a specific Service Level Agreement.