

DiffServ

Area de Ingeniería Telemática
<http://www.tlm.unavarra.es>

Grado en Ingeniería en Tecnologías de
Telecomunicación, 3º

Origen de los Servicios Diferenciados

Propuestas del IETF

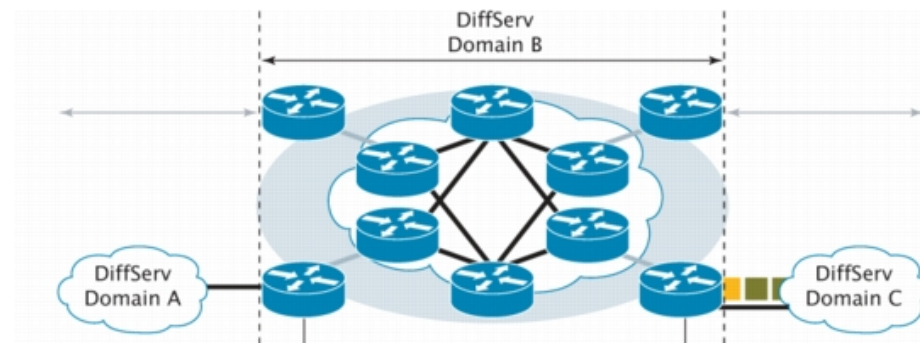
- **IntServ** (Integrated Services)
 - Filosofía: reserva de recursos
 - Cada router del trayecto ha de tomar nota y efectuar la reserva solicitada
- **DiffServ** (Differentiated Services)
 - (...)

Propuestas del IETF

- **IntServ** (Integrated Services)
 - Filosofía: reserva de recursos
 - Cada router del trayecto ha de tomar nota y efectuar la reserva solicitada
- **DiffServ** (Differentiated Services)
 - Filosofía: priorización de tráfico
 - El usuario o un primer equipo de red marca los paquetes con un determinado nivel de prioridad
 - Los routers van agregando las demandas de los usuarios y propagándolas por el trayecto
 - A cada clase de paquetes se le da un trato diferenciado
 - Esto le da al usuario una confianza razonable de conseguir la QoS solicitada
- Pueden coexistir

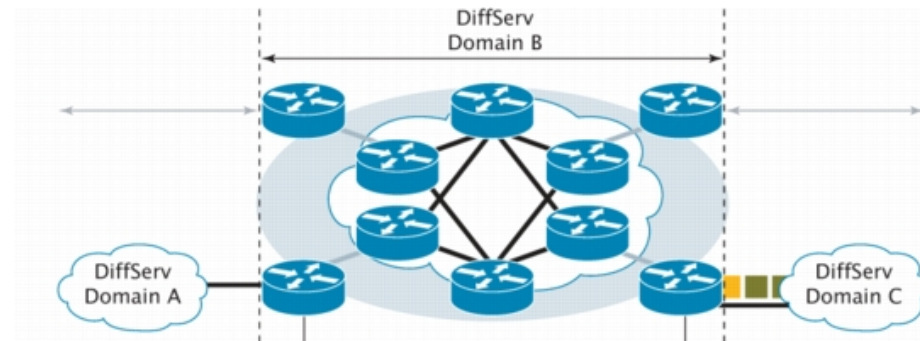
DiffServ

- IntServ se suele considerar que no escala bien
- RFC 2475 : “An Architecture for Differentiated Services”
- DiffServ no es sensible a los requisitos de un flujo individual y gracias a eso logra escalabilidad
- Clasificar el tráfico en pocas clases
- Clasifican los *ingress routers* (complejidad en la frontera)
- Asignan un *codepoint* en la cabecera IP
- DiffServ mapea en cada nodo el *codepoint* del paquete a un comportamiento en concreto (scheduling, gestión de cola...)



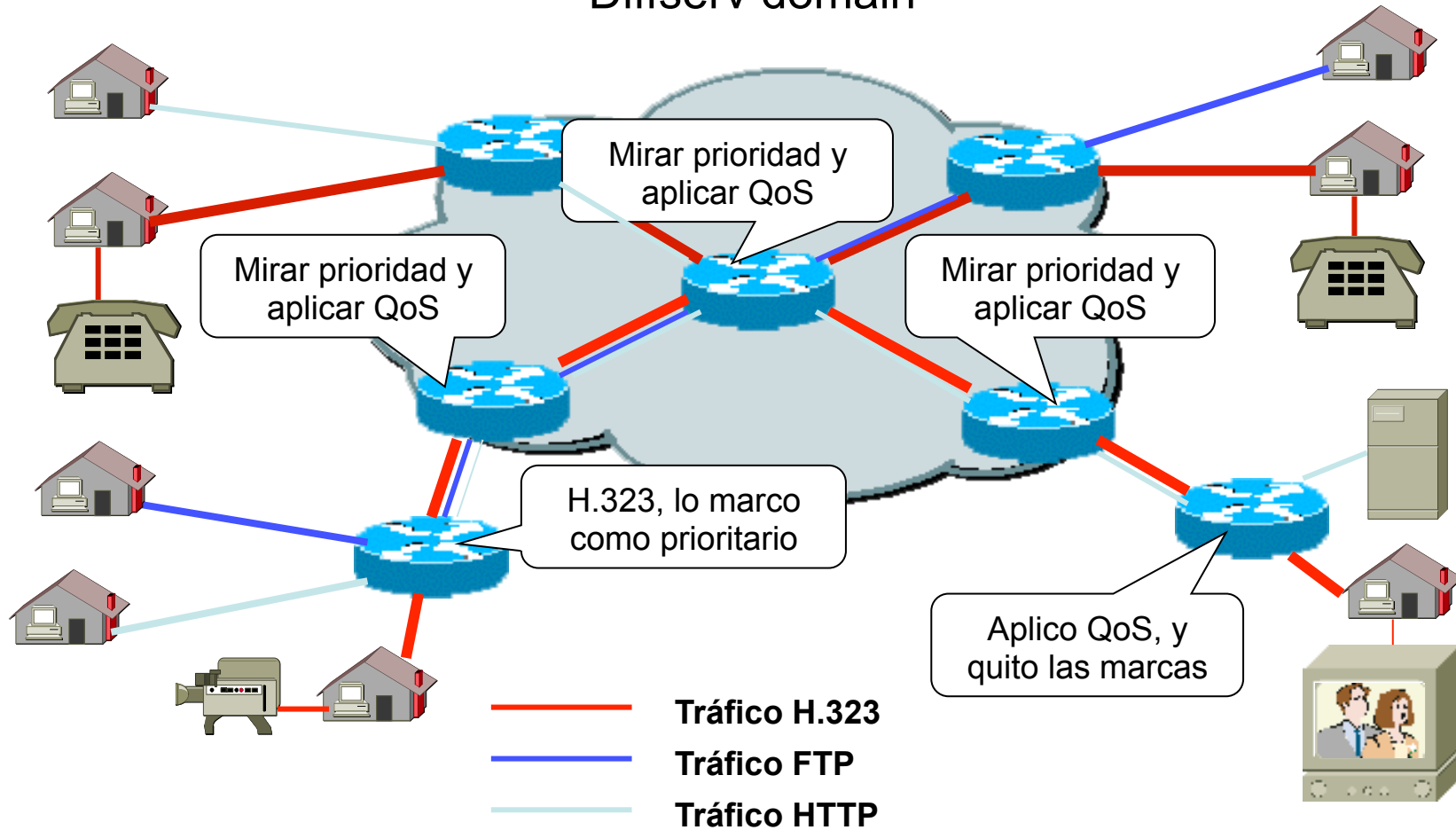
DiffServ

- Configuración manual o mediante NMS (Network Management System), no automática
- Se puede aplicar a IPv4, a IPv6 y a MPLS



Ejemplo

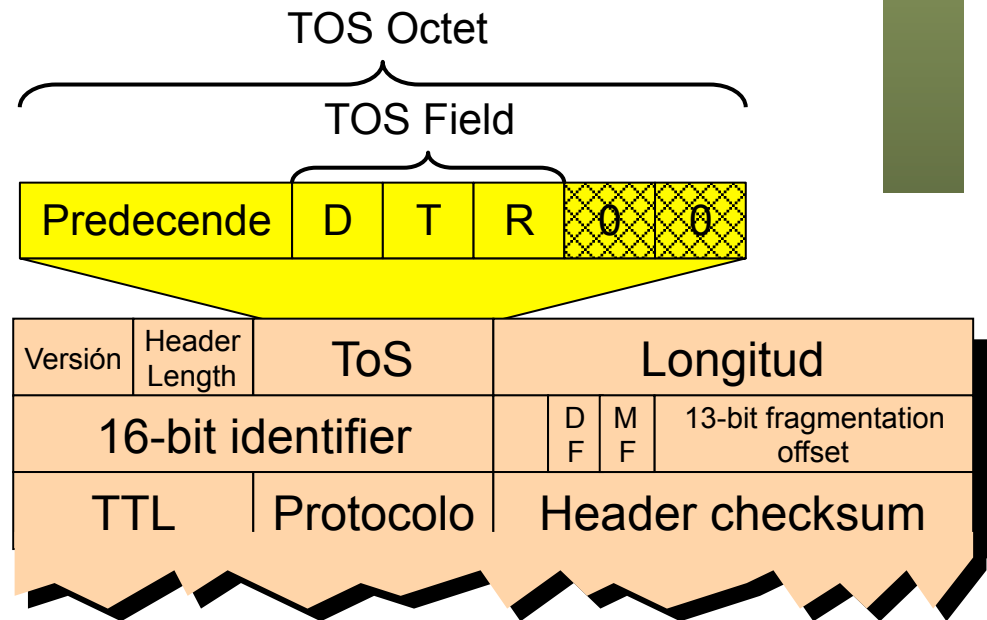
Diffserv domain



TOS: RFC 791

- RFC 791, RFC original de IP: Type Of Service 8 bits
- Bits 0-2 : Precedence (fácil mapeo a bits de 802.1p)
- Bit 3: 0 = Normal delay, 1 = Low Delay (subjetivos)
- Bit 4: 0 = Normal Throughput, 1 = High Throughput
- Bit 5: 0 = Normal Reliability, 1 = High Reliability
- Bits 6-7: Reservados para uso futuro
- Los bits de precedencia se han usado, el resto no

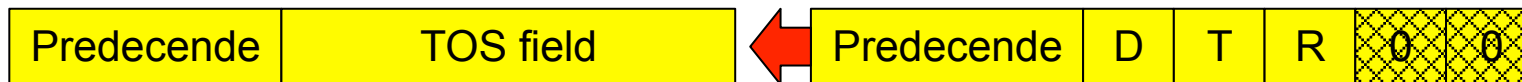
Precedence	Significado
111	Network Control
110	Internetwork Control
101	CRITIC/ECP
100	Flash Override
011	Flash
010	Immediate
001	Priority
000	Routine



Precedence = "An independent measure of the importance of this datagram"

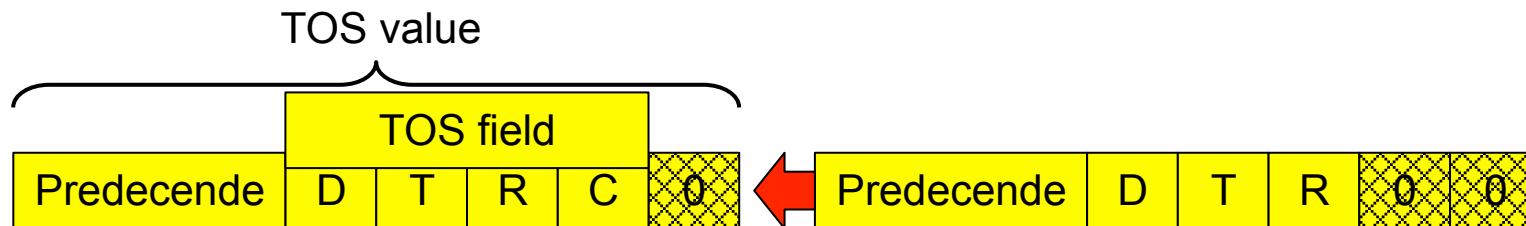
TOS: RFC 1122

- RFC “Requirements for Internet Hosts – Communication Layers”
- El segundo campo engloba los últimos 5 bits aunque no se define el uso de los dos últimos



TOS: RFC 1349

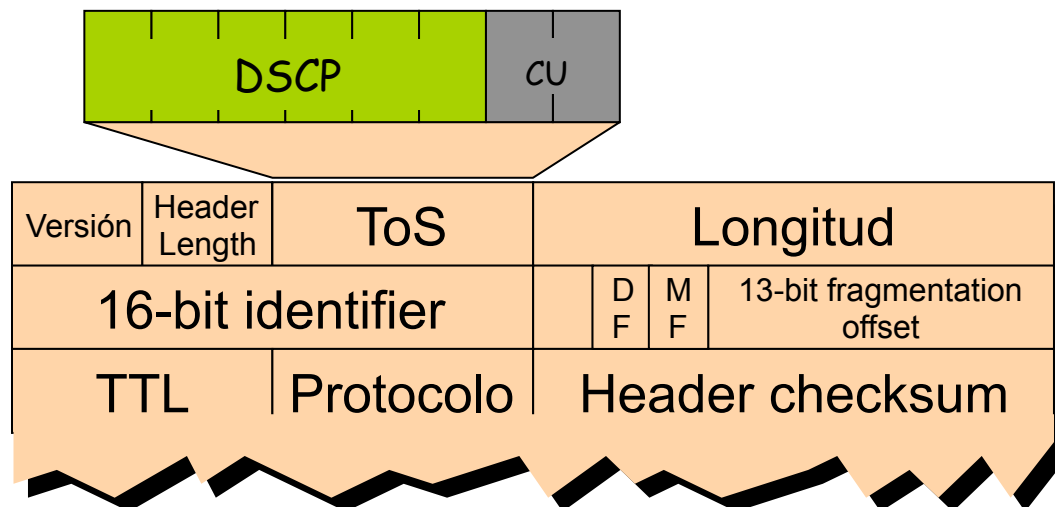
- El TOS field indica cómo debe hacer la red los compromisos respecto a este paquete
- Se mantienen los bits para Delay, Throughput y Reliability y se añade un cuarto para “Coste” (monetario)
- El último bit se llama “MBZ” (*Must Be Zero*)
- Ya no se interpretan como bits independientes sino que solo hay estos valores válidos:
 - 1000 : minimizar delay (pero puede no ser suficiente para el usuario)



- 0100 : maximizar throughput
- 0010 : maximizar reliability
- 0001 : minimizar coste económico
- 0000 : servicio normal
- Precedence para determinar la cola
- TOS para determinar el camino en la red (usándolo en el IGP)

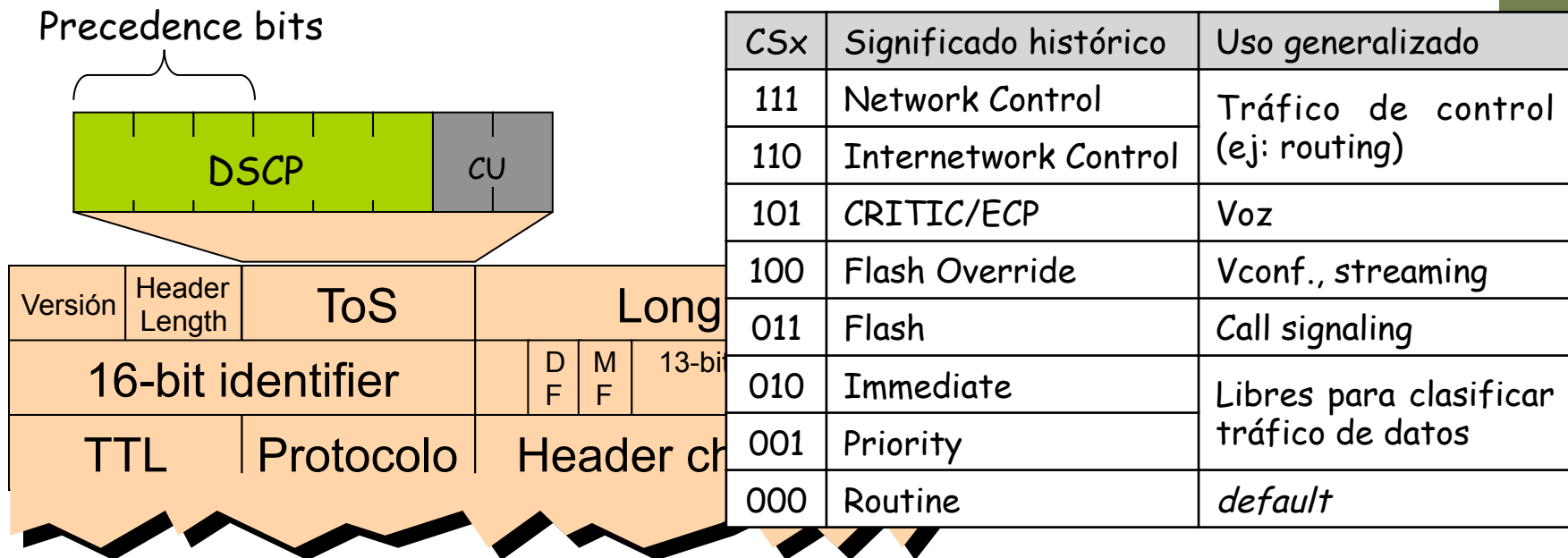
DiffServ: DSCP

- RFC 2474 : “Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPV6 Headers”
- RFC 2475 : “An Architecture for Differentiated Services”
- Las definiciones anteriores quedan obsoletas
- ToS ahora se llama DS (*Differentiated Services*) field
- 2 bits sin usar (“CU” = “Currently Unused”) (hoy ya se usan en ECN)
- 6 de sus bits son el DSCP (*Differentiated Services CodePoint*)



DiffServ: Precedence

- Se mantiene en los *Class Selector Codepoints*: CSx = XXX000
- Cada *Class Selector Codepoint* corresponde a una probabilidad de un reenvío “a tiempo” no inferior al que se obtendría con un CS inferior
- Un paquete descartado se considera un caso de entrega “fuera de tiempo”



Per-Hop Behaviors

PHBs

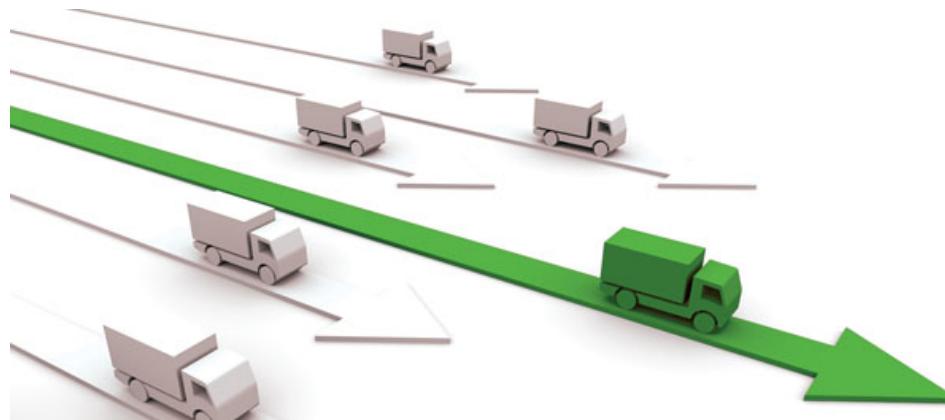
- PHB = *Per Hop Behavior*
- El tratamiento que se le da al paquete en cuestión de scheduling y gestión de cola en ese nodo
- El mapeo *codepoint* \leftrightarrow *PHB* debe ser configurable
- No se definen planificadores o controles de cola sino una visión externa de “caja negra” del comportamiento en un salto

PHBs

- *Best-Effort* (BE, codepoint 0) o *Default PHB*
- *Assured Forwarding* (AFxy)
- *Expedited Forwarding* (EF)

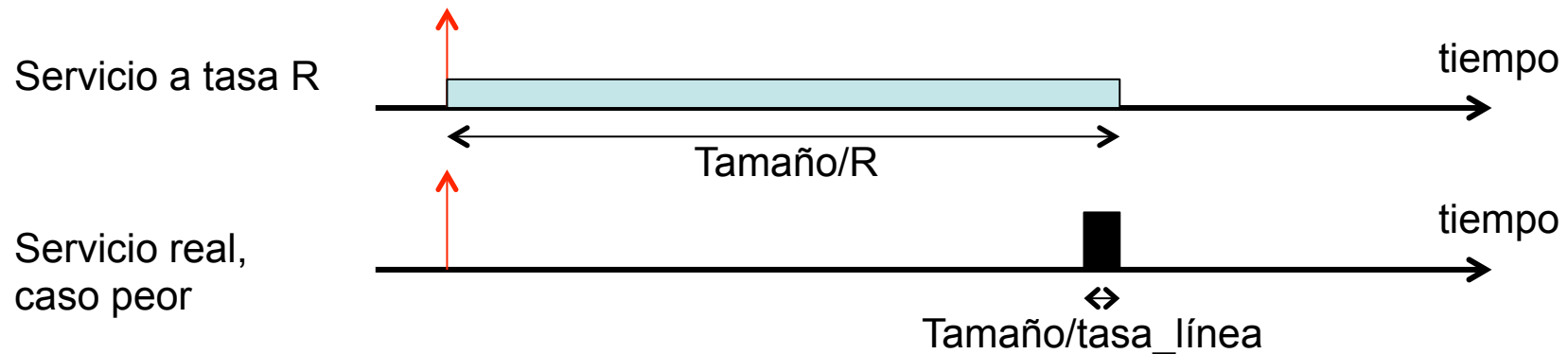
PHB: Expedited Forwarding (EF)

- RFC 3246 : “An Expedited Forwarding PHB (Per-Hop Behavior)”
- Alta prioridad: pocas pérdidas, baja latencia, bajo jitter, bw garantizado
- DSCP recomendado 101110 (46)
- Debe ser servido al menos a la tasa configurada (R , medida en un intervalo), independiente del tráfico no-EF



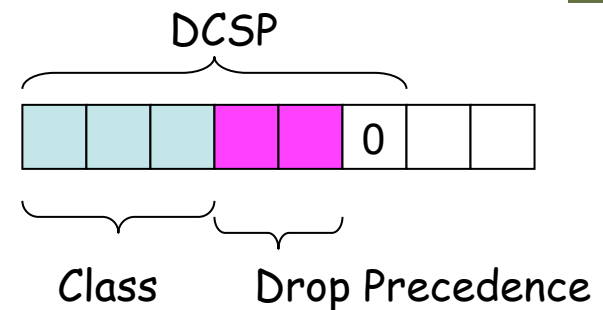
PHB: Expedited Forwarding (EF)

- El instante de salida de un paquete se toma como el de terminar de transmitir el último bit
- Si llega un paquete EF y no hay ninguno anterior
 - Empezaría su servicio cuando llega
 - Se le sirve a tasa R y con eso y su tamaño se determina su instante de salida
 - Por supuesto luego se le envía a tasa de línea, no a R



PHB: Assured Forwarding (AF)

- RFC 2597 : “Assured Forwarding PHB Group”
- Clasificar y tratar tráfico que exceda el contrato para ser descartado con mayor probabilidad
- En realidad son 4 grupos de PHBs (AF1x, AF2x, AF3x y AF4x) independientes
- Cada uno tiene una reserva en cada nodo (BW, buffer)
- Cada uno con 3 probabilidades de descarte (*drop*) (en total 12)
- DSCP `xxxyy0` : `xxx` la clase, `yy` la *drop precedence*
- (...)

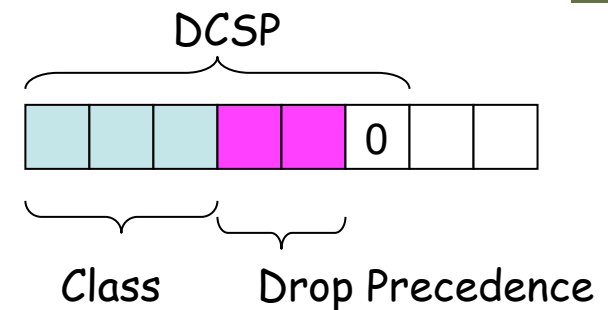


$$p_L \leq p_M \leq p_H$$

PHB: Assured Forwarding (AF)

- DSCP $xxxyy0$: xxx la clase, yy la *drop precedence*
- Drop precedence alta implica mayor probabilidad de descarte
- Para cada drop precedence, empezando desde la más baja, la probabilidad de descarte debe mayor que para la anterior
- No hay relación entre probabilidades de descarte de clases diferentes

Drop	Class 1	Class 2	Class 3	Class 4	Drop prob.
Low	001010 AF11	010010 AF21	011010 AF31	100010 AF41	p_L
Medium	001100 AF12	010100 AF 22	011100 AF32	100100 AF42	p_M
High	001110 AF13	010110 AF23	011110 AF33	100110 AF43	p_H

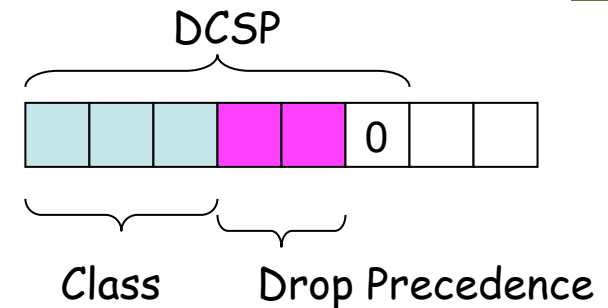


$$p_L \leq p_M \leq p_H$$

PHB: Assured Forwarding (AF)

- Busca que ante congestión se descarte con mayor probabilidad de los de drop precedence alta
- Debe emplear AQM (RED, WRED, ...) con diferentes parámetros según drop precedence
- Se suelen usar solo AFx1 y AFx2 pues es difícil definir qué entendemos por una tercera subclase de prioridad
- Policers para la clasificación

Drop	Class 1	Class 2	Class 3	Class 4	Drop prob.
Low	001010 AF11	010010 AF21	011010 AF31	100010 AF41	p_L
Medium	001100 AF12	010100 AF 22	011100 AF32	100100 AF42	p_M
High	001110 AF13	010110 AF23	011110 AF33	100110 AF43	p_H



$$p_L \leq p_M \leq p_H$$

PHB: Class Selector (CS)

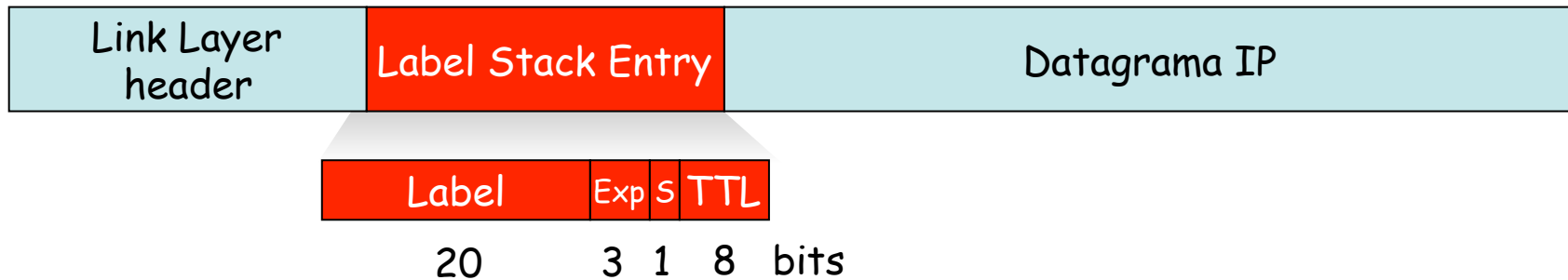
- DSCP xxx000
- Mantienen el significado de los 3 bits de “precedence”
- A mayor valor mayor probabilidad de entrega a tiempo

Nombre	DSCP
CS0	000000
CS1	001000
CS2	010000
CS3	011000
CS4	100000
CS5	101000
CS6	110000
CS7	111000

DiffServ y las tecnologías

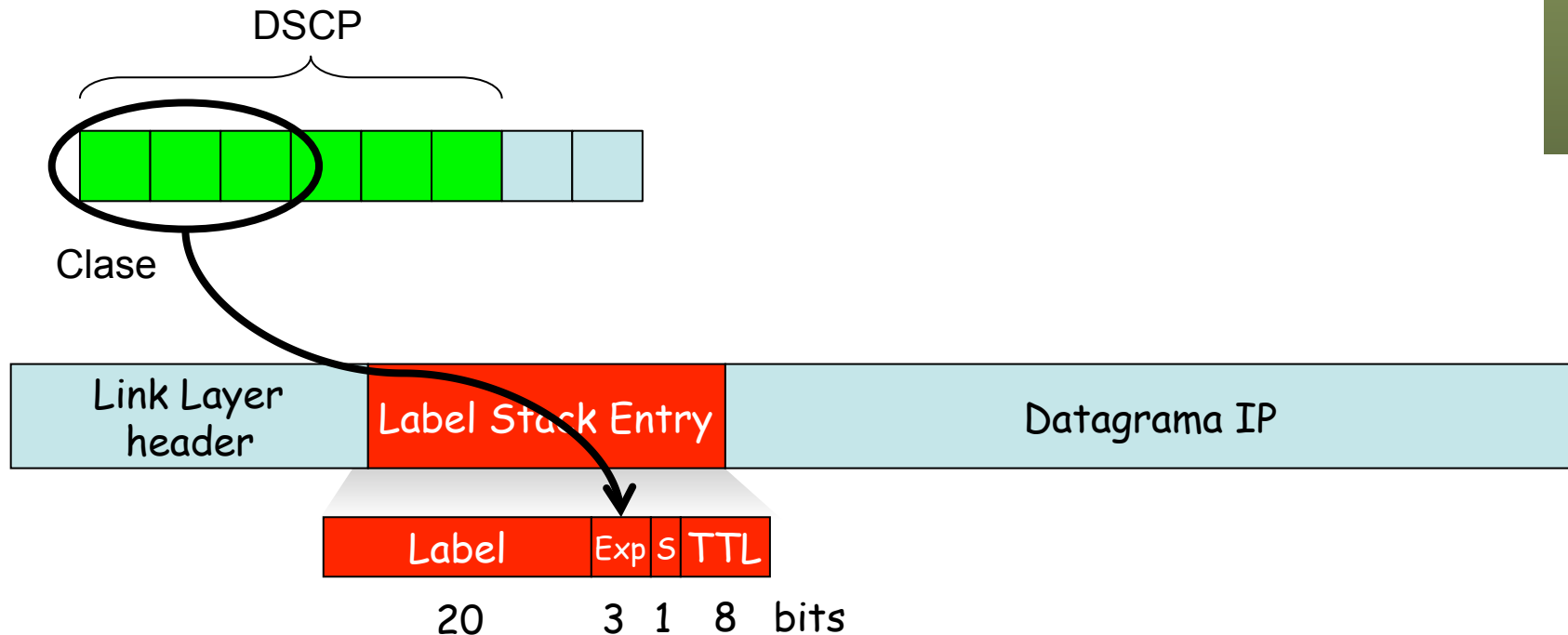
MPLS : Exp field

- 3 bits en una entrada de etiqueta
- Definido inicialmente (RFC 3032) para uso experimental
- RFC 5462 lo renombra a "*Traffic Class field*"



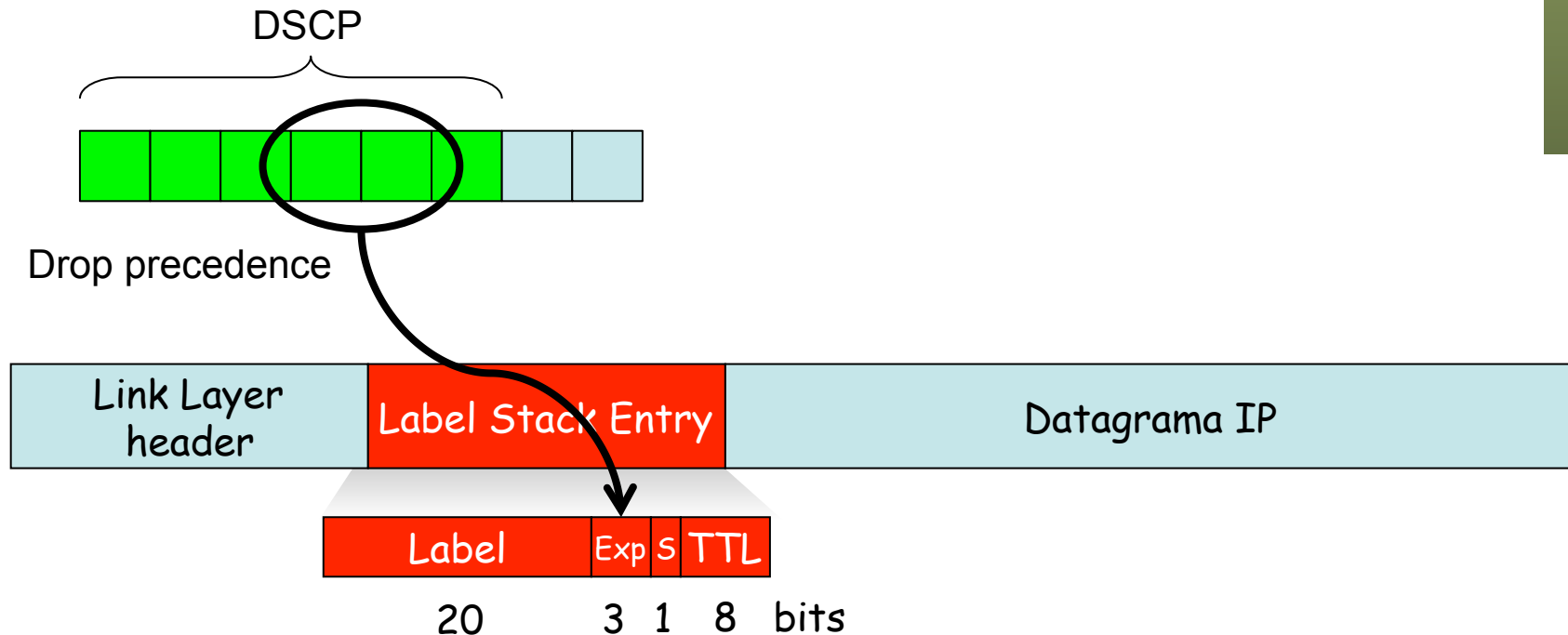
MPLS y DiffServ

- RFC 3270 : “Multi-Protocol Label Switching (MPLS) Support of Differentiated Services”
- Los LSRs no miran la cabecera IP y por lo tanto no ven el DSCP
- Hay 64 DSCPs y los bits EXP dan para 8 valores
- Mapear varios PHBs a un valor: PHB Scheduling Class (PSC)
- E-LSPs:
 - EXP-Inferred-PSC LSPs
 - Se emplea el campo EXP para determinar un PSC



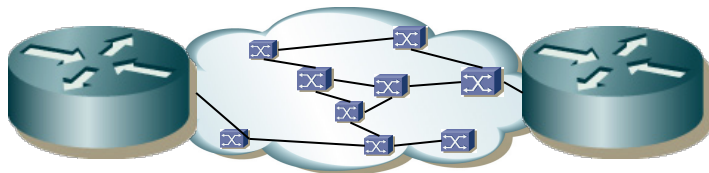
MPLS y DiffServ

- L-LSPs:
 - Label-Only-Inferred-PSC LSPs
 - El LSP lleva tráfico de solo un PSC
 - Se ponen de acuerdo los extremos en cuál en el establecimiento
 - Los bits EXP se emplean para representar la *drop precedence* si hay shim
 - Si no hay shim header (ATM LSR) se emplea campos específicos del nivel de enlace (en ATM el bit CLP)
 - No muy empleados



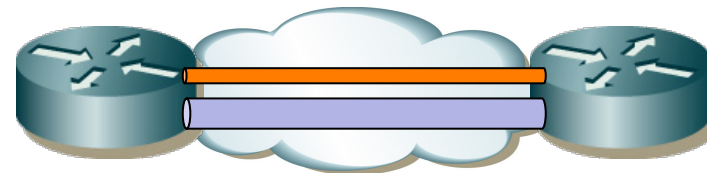
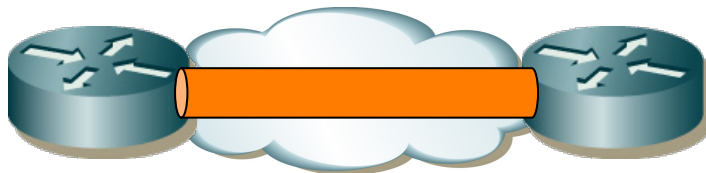
DiffServ y ATM

- ATM ofrece varias categorías de servicio: CBR, VBR-rt, VBR-nrt, UBR, ABR
- Ofrecen garantías respecto a delay (CTD) y jitter (CDV)
- Tráfico entre routers DiffServ puede atravesar una WAN ATM
- Hay que mapear las clases de tráfico a uno o varios VCs
- (...)



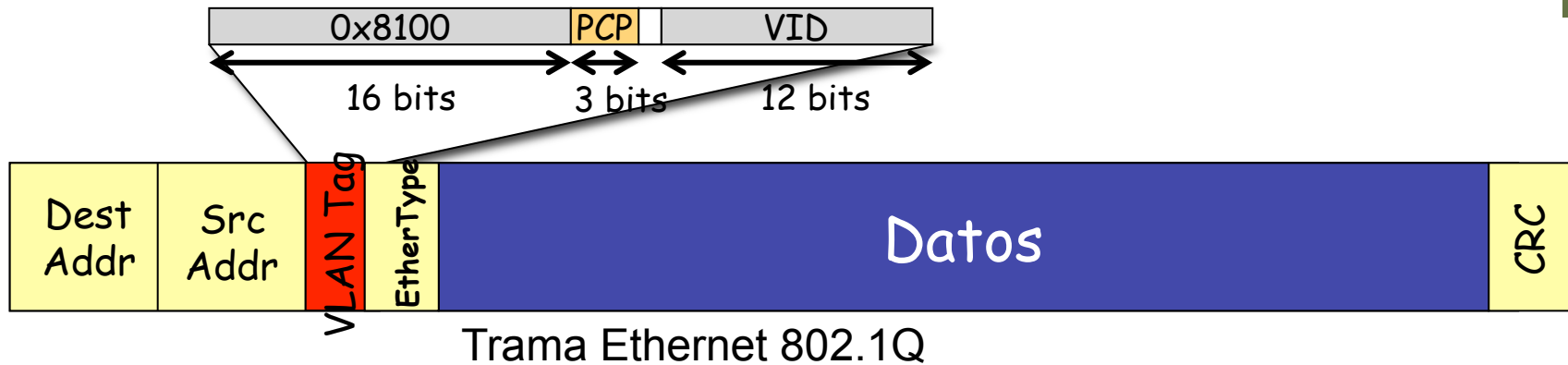
DiffServ y ATM

- Si se mapean todas a uno:
 - Si hay tráfico de tiempo real (voz) se necesita un VC para todos que ofrezca garantías de CTD y CDV
 - Circuitos *real time* serán generalmente más caros
 - Pero al menos se puede aprovechar mejor la multiplexación estadística
- Si se mapean a circuitos independientes:
 - Cada uno puede emplear la clase más adecuada a su tipología
 - Solo las clases *real time* necesitarán ese tipo de VCs
 - No se aprovecha multiplexación estadística entre ellos



DiffServ y Ethernet

- El campo de prioridad en 802.1Q es análogo al DSCP
- Solo 8 valores frente a los 64 DSCPs
- Los conmutadores suelen soportar una cola de prioridad (EF) y una serie de colas con pesos (AF)



Codepoints

- <http://www.iana.org/assignments/dscp-registry/dscp-registry.xml>
- Los estándar son xxxxx0
- Experimentales o para uso local: xxxx11
- Idem pero podrían estandarizarse: xxxx01

Nombre	DSCP
CS0	000000
CS1	001000
CS2	010000
CS3	011000
CS4	100000
CS5	101000
CS6	110000
CS7	111000

Nombre	DSCP
AF11	001010
AF12	001100
AF13	001110
AF21	010010
AF22	010100
AF23	010110
AF31	011010
AF32	011100
AF33	011110
AF41	100010
AF42	100100
AF43	100110
EF	101110
Voz CAC	101100

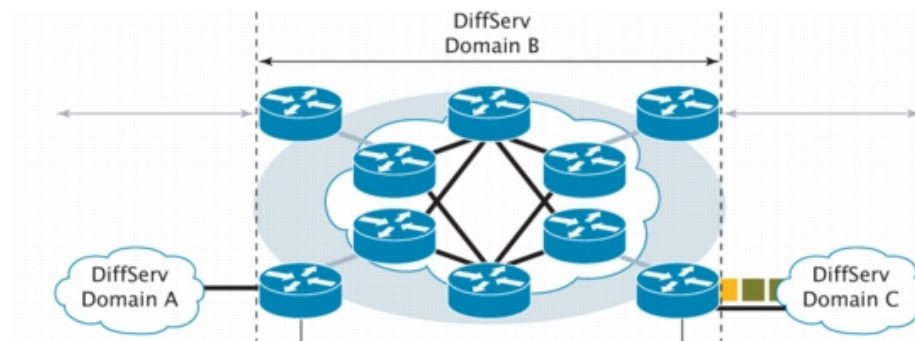
DiffServ SLAs

- Definidos en base a :
 - Traffic Conditioning Specification (TCS)
 - Define el tráfico al que se debe garantizar el SLA
 - Parámetros y sus valores que especifican las reglas de clasificador y el perfil de tráfico
 - Service Level Specification (SLS)
 - El conjunto de parámetros y sus valores que definen el servicio ofrecido a un flujo



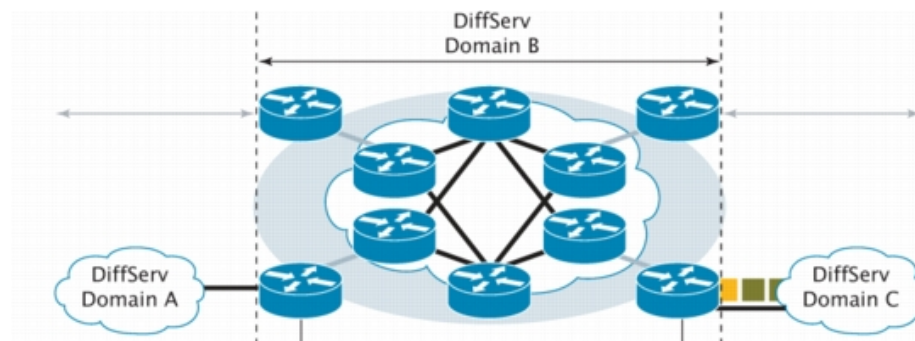
DiffServ SLAs : PDBs

- Per-Domain Behaviors (PDBs) (RFC 3086)
- Equivalente a un PHB pero en la escala de todo un dominio DiffServ
- Solo está definido el “lower-effort PDB” (LE) (RFC 3662)
- (...)



PDBs LE

- LE es menos prioritario que best effort y se descarta si hay cualquier otro tráfico
- Para BE al menos hay un “intento” de cursarlo lo mejor posible, para LE ni eso
- Se envían si el enlace está inactivo (por ejemplo WRR peso=0)
- No se especifica el codepoint a usar. Si es un CS se sugiere el CS1 pero puede ser un PHB AF o uno definido localmente
- Ejemplo: tráfico multimedia UDP no adaptativo, el exceso que no compita con best effort pues no compite justamente, hacerlo LE
- Ejemplo: para tráfico P2P, para tráfico de spiders web



Resumen

- Mercado el DSCP
- Clases
 - EF: tasa asegurada
 - AF: subclases que van degradando la calidad, emplean AQM
 - BE: Best Effort