Queueing

Área de Ingeniería Telemática
http://www.tlm.unavarra.es

Máster en Comunicaciones

Objetivos

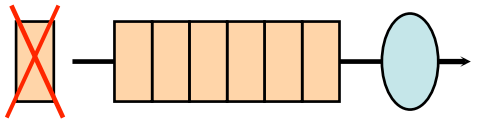
- Conocer los mecanismos de gestión activa de cola

Congestion Avoidance

- También llamado *Active Queue Management*
- Su objetivo es evitar que interfaces o colas se congestionen
- Diseñado para TCP pues es un protocolo que reacciona ante congestión reduciendo la tasa de envío

Passive Queue Management

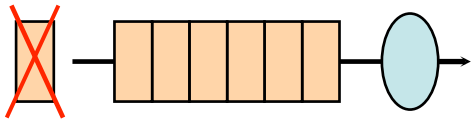
Drop-Tail (la más habitual)



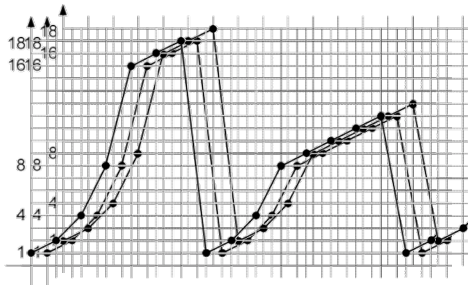
- Simple
- Descarta un paquete independientemente de su importancia
- Controla la congestión pero no la evita
- (...)

Passive Queue Management

Drop-Tail (la más habitual)

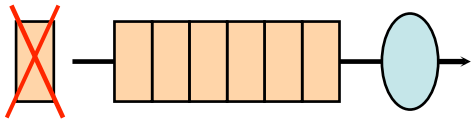


- Simple
- Descarta un paquete independientemente de su importancia
- Controla la congestión pero no la evita
- TCP puede enviar ráfagas (limitadas por el tamaño de la ventana)
- Si llega una ráfaga a una cola casi llena se perderán varios paquetes
- Ante pérdidas TCP reduce ventana de congestión y así velocidad de envío
- Varias conexiones pueden entrar simultaneamente en este proceso de control de la congestión y reducirse en gran medida el throughput global
- Introduce **sincronización global** con varias conexiones TCP en el enlace
- (...)

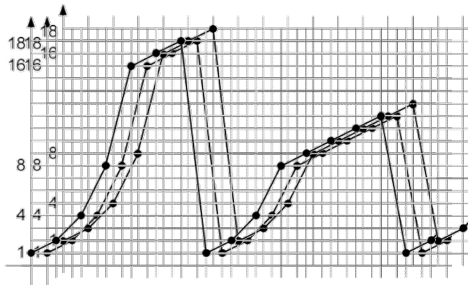


Passive Queue Management

Drop-Tail (la más habitual)



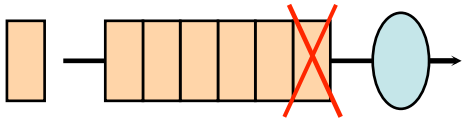
- Simple
- Descarta un paquete independientemente de su importancia
- Controla la congestión pero no la evita
- TCP puede enviar ráfagas (limitadas por el tamaño de la ventana)
- Si llega una ráfaga a una cola casi llena se perderán varios paquetes
- Ante pérdidas TCP reduce ventana de congestión y así velocidad de envío
- Varias conexiones pueden entrar simultaneamente en este proceso de control de la congestión y reducirse en gran medida el throughput global
- Introduce **sincronización global** con varias conexiones TCP en el enlace
- Unos pocos flujos pueden monopolizar el recurso
- La cola se mantiene llena por largos periodos
- Pero la cola es para absorber ráfagas, luego no se podrán absorber
- Colas llenas no lleva a mayor throughput sino a menor throughput
- **Colas poco ocupadas llevan a mayor throughput y menor retardo**



Passive Queue Management

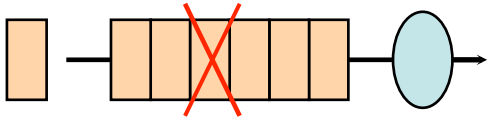
Head-drop

- Tira los paquetes que más tiempo llevan en el buffer
- Probablemente ya han sido retransmitidos (TCP)
- Probablemente ya llegan tarde (UDP/RTP)
- Controla la congestión pero no la evita, posible synch



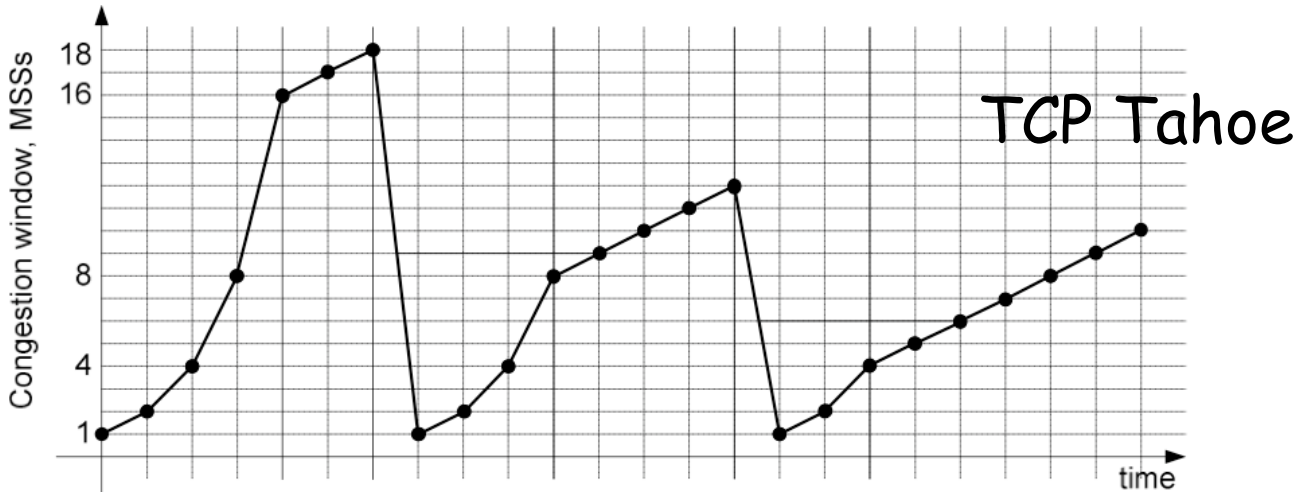
***Random-Drop* (ante cola llena)**

- Se puede reducir la sincronización global pero no controlar UDP
- Controla la congestión pero no la evita



Active Queue Management

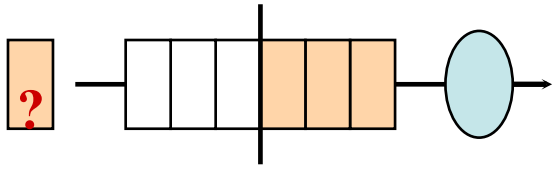
- Pensando en TCP, no controla UDP igual de bien
- Evita sincronizaciones, menores retardos y fluctuaciones
- TCP regula su tasa al detectar pérdidas (*Congestion avoidance*)



Active Queue Management

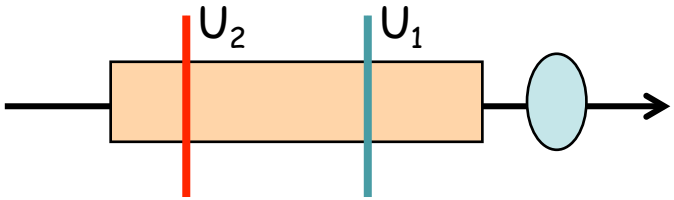
Early-Random-Drop (cola no llena)

- Si la cola excede un nivel se tira cada paquete que llega con una probabilidad fija



Weighted Tail Drop

- Se asignan umbrales en la cola a diferentes clases de tráfico
- Cuando se alcanza el umbral U_1 se descartan los paquetes que lleguen de la clase 1
- Cuando se alcanza el umbral U_2 se descartan también los paquetes que lleguen de la clase 2
- De la tercera clase se descartarán solo cuando se llene la cola

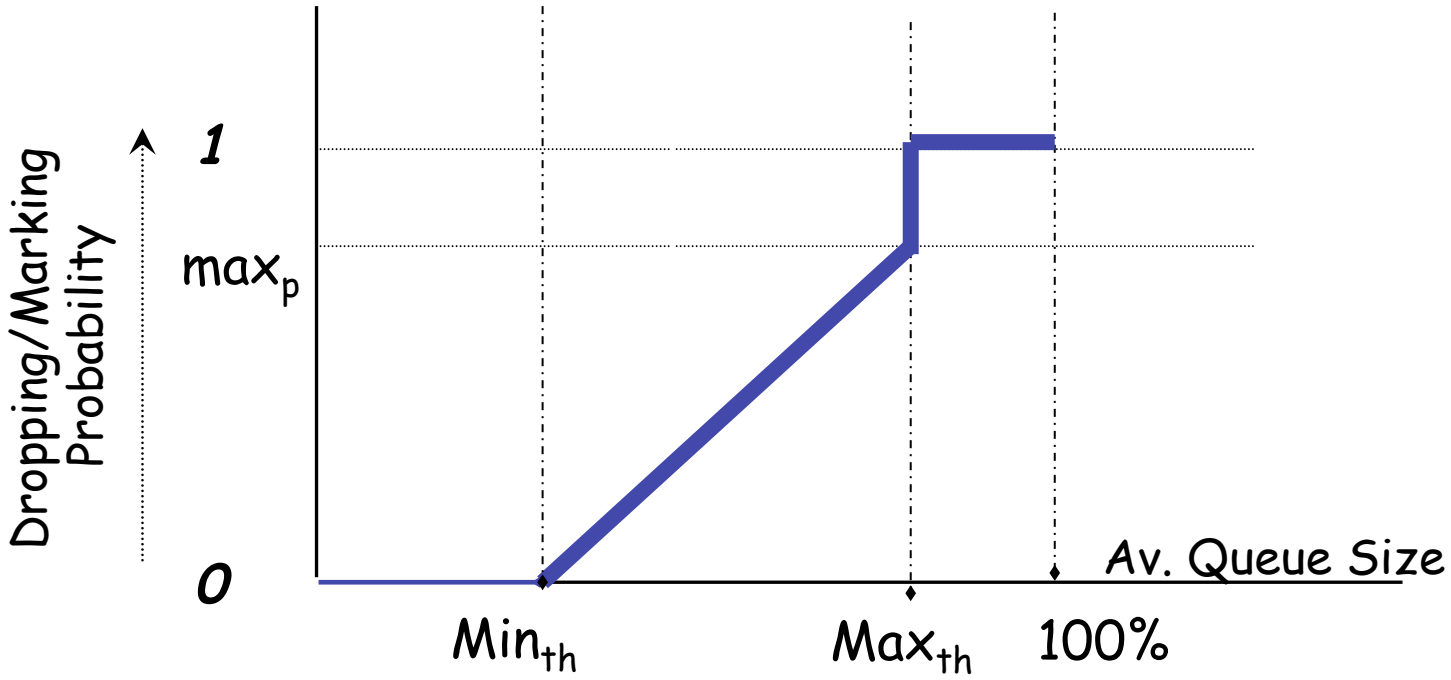


Active Queue Management

RED (Random Early Detection)

- RFC 2309
- Descartar paquetes probabilísticamente antes de la congestión
- Evalúa la ocupación media del buffer
- Cálculo mediante *exponential moving average*
- Para w bajo la media sigue los cambios rápidos del valor instanténeo

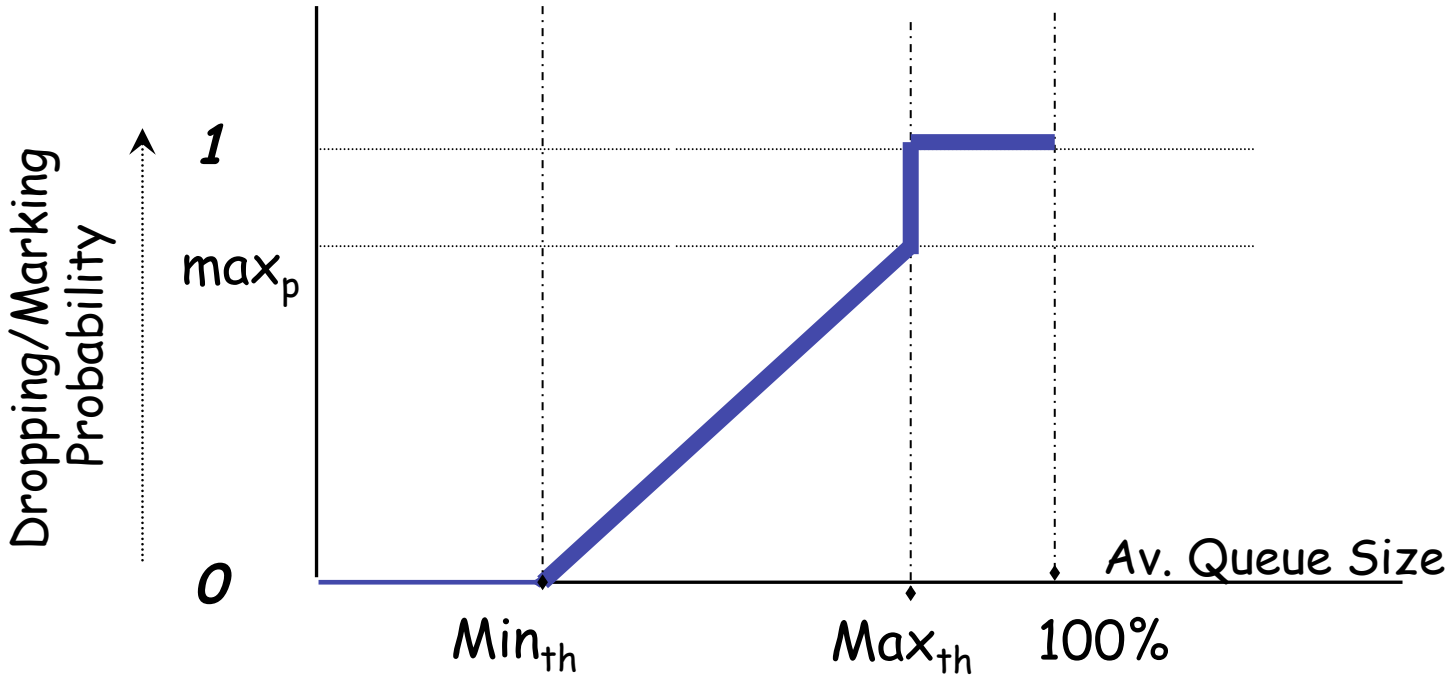
$$q_{avg} = q_{avg\_old}\left(1-\frac{1}{2^w}\right)+\left(q_{current}\frac{1}{2^w}\right)$$



Active Queue Management

RED (Random Early Detection)

- Difícil medir sus beneficios
- Con mala configuración podría comportarse peor que *drop-tail*
- Se han propuesto bastantes más algoritmos de AQM
- RED es el más extendido en implementaciones

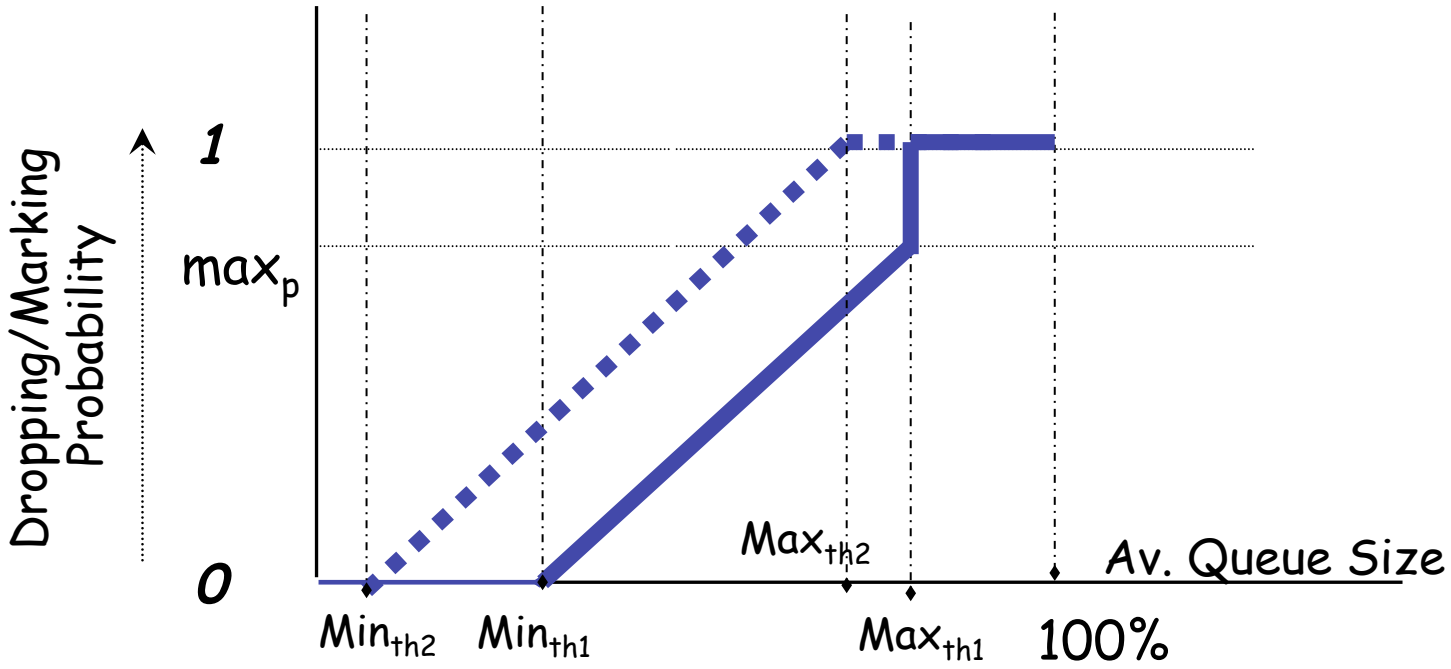


Active Queue Management

WRED (Weighted RED)

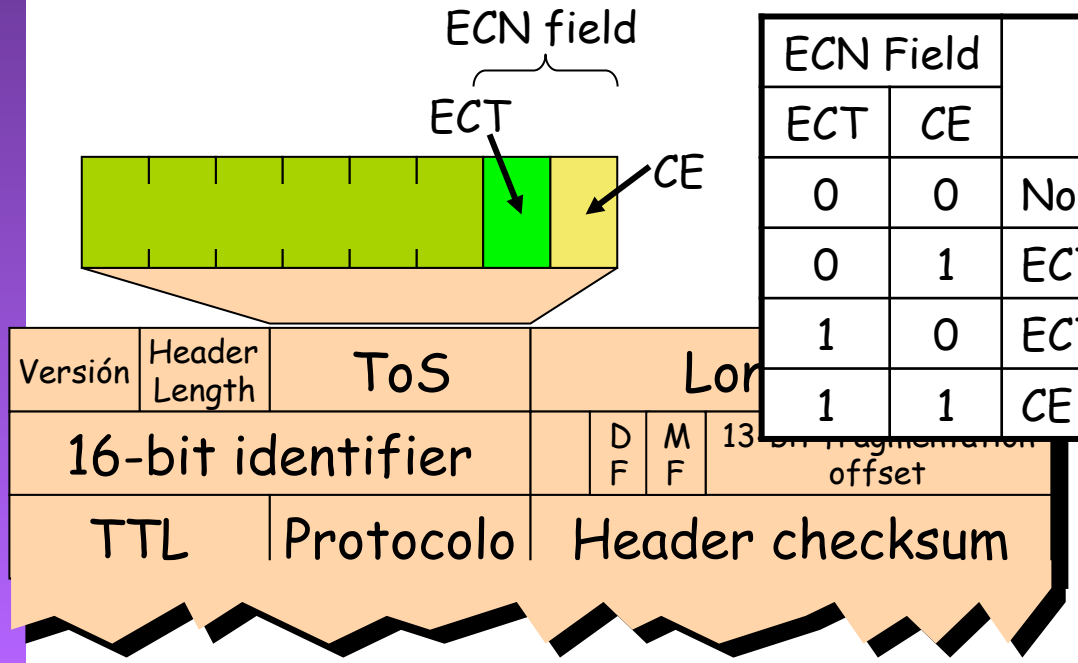
- Emplea un Min_{th} y Max_{th} diferente para diferentes clases de tráfico
- Mayor cuanto mayor es el valor de precedencia

Otras propuestas: Adaptive RED (ARED), RED In & Out (RIO), Flow weighted RED (FRED)...



ECN

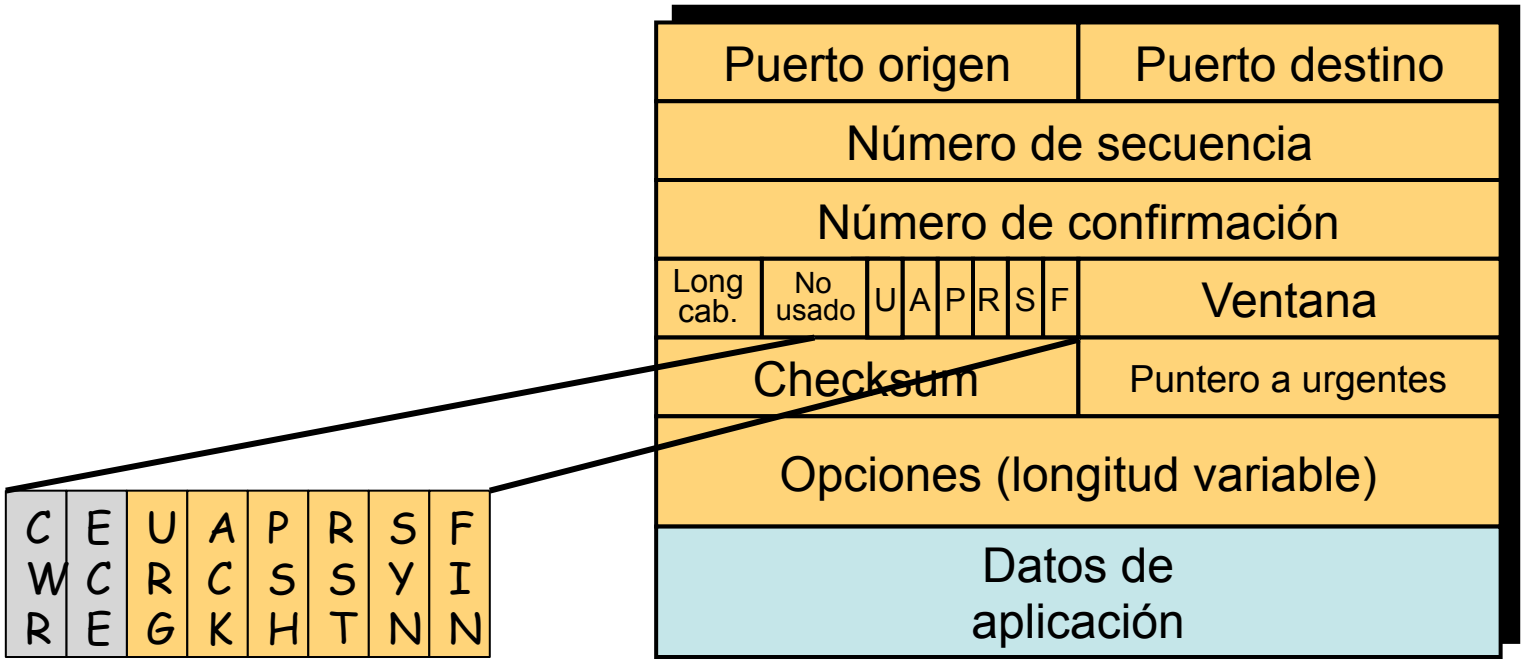
- *Explicit Congestion Notification*
- RFC 3168
- Extensión a RED: marcar en vez de descartar (salvo cuando se alcanza ocupación máxima que sí se descarta)
- Bit ECT = *ECN-Capable Transport*
- Bit CE = *Congestion Experienced*
- Requiere extender el control de congestión de TCP



ECN Field			
ECT	CE		
0	0	Not-ECT	Paquete no emplea ECN
0	1	ECT(1)	Extremos "ECN-capable"
1	0	ECT(0)	idem (recomendado)
1	1	CE	Congestion experienced

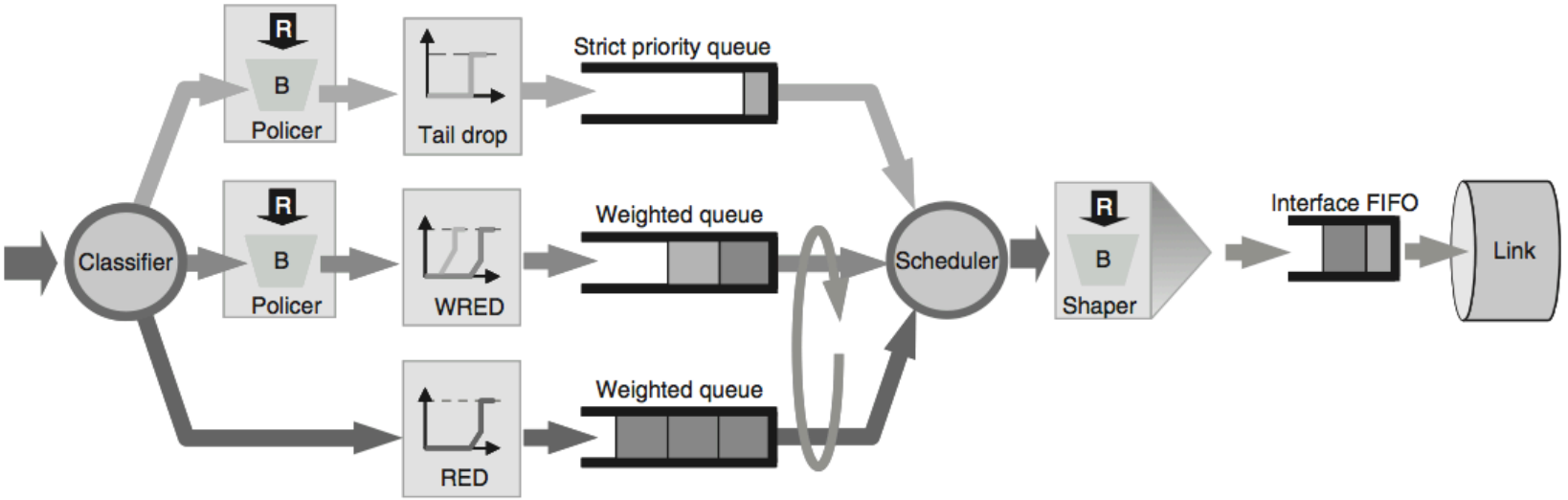
ECN y TCP

- Emplea dos flags nuevos en la cabecera
- ECN-Echo (ECE): para que el receptor del paquete con CE activo devuelva esta indicación a emisor
- Emisor reacciona como si hubiera detectado una pérdida
- CWR flag: emisor notifica a receptor en el siguiente paquete de que ha recibido el ECE
- Poco extendida la implementación de ECN en routers y hosts

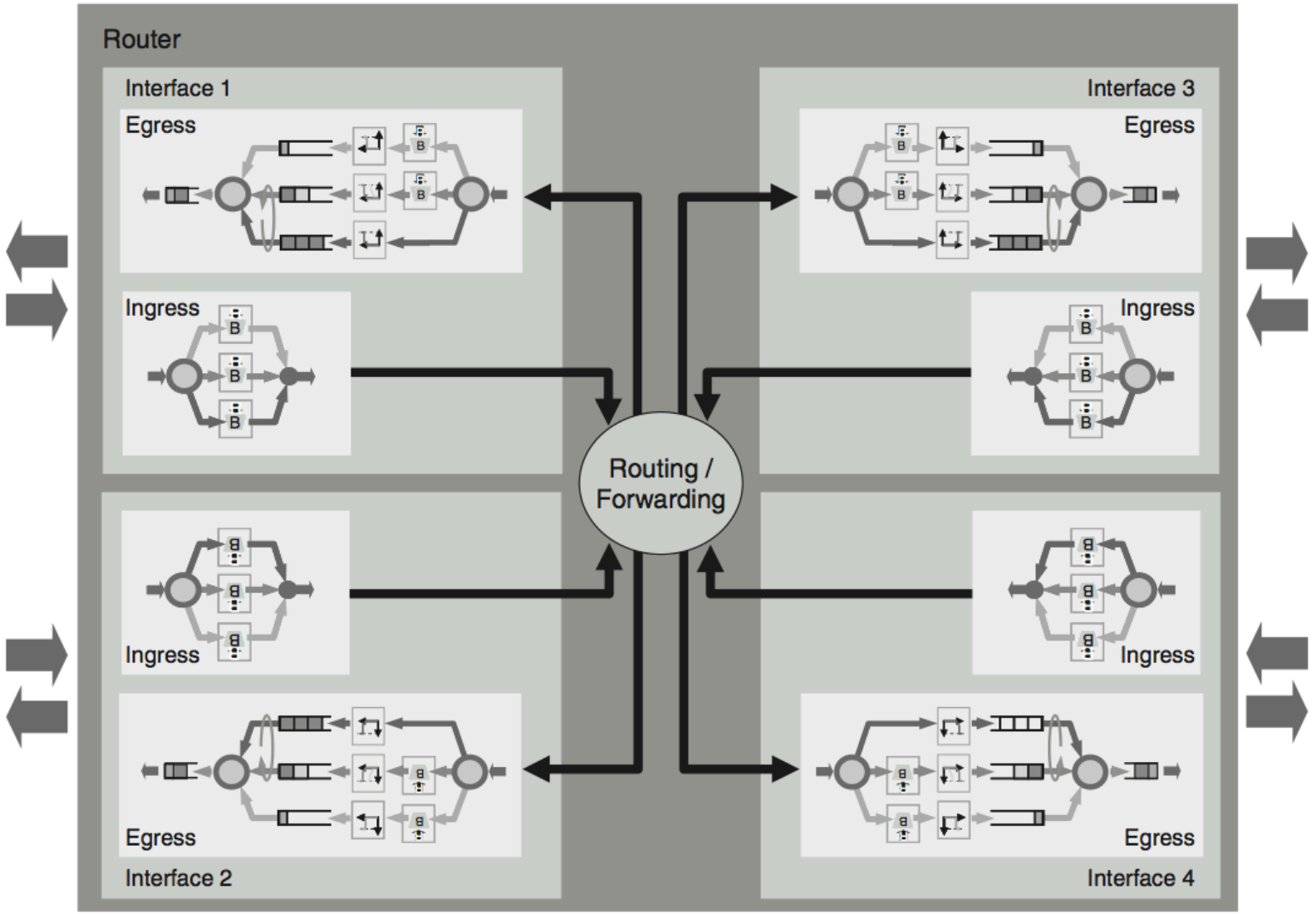


Implementación

Implementación



Implementación



Equipos

D-Link DES-1210

VoIP Deployment

- Highest priority for VoIP services
- Auto Voice VLAN

QoS

- Ensure time-sensitive data gets delivered efficiently
- Supports IEEE 802.1p QoS up to four 802.1p Priority Queues
- DSCP QoS

QoS, Bandwidth Control

The switches are perfect for deployment in a VoIP environment, as they support Auto Voice VLAN and Differentiated Services Code Point (DSCP) QoS for VoIP application. Auto Voice VLAN will automatically place voice traffic from an IP phone to an assigned VLAN and by doing so enhance the VoIP service.

With a higher priority and individual VLAN, this feature guarantees the quality and security of VoIP traffic. DSCP marks parts of an IP packet, enabling different levels of service to be assigned for network traffic. With Bandwidth Control, the network administrator can reserve bandwidth for important functions that require a larger bandwidth or might have high priority.

Auto Voice VLAN	• Max. 10 user defined OUI, Max. 8 default OUI
QoS (Quality of Service)	
802.1p Quality of Service	• 4 Queues per port
Queue Handling	• Strict, Weighted Round Robin (WRR)
CoS based on	• 802.1p Priority Queues, DSCP
Bandwidth Control	• Port-based (Ingress/Egress, min. granularity 64 Kb/s)

D-Link DES 3828P

Priority Queues	8 Queues - Strict Priority or Weight Round Robin Queue Scheduling
Classification ACLs	MAC Address, IPv4/v6 Address, TCP/UDP Port Number & Payload, Physical Port, DSCP, 802.1p,VLAN, Protocol Type, IPv6 Traffic Class & Flow Label
Classification QoS	MAC Address, IP Address, TCP/UDP Port Number, Physical Port, DSCP, ToS, IPv6

Cisco Catalyst 2950 (SI)

NETWORK CONTROL

Cisco Catalyst 2950SX-48, 2950T-48, 2950SX-24, 2950-24, and 2950-12 switches deliver LAN-edge QoS, **supporting two modes of reclassification**. **One mode-based on the IEEE 802.1p** standard-honors the class-of-service (CoS) value at the ingress point and assigns the packet to the appropriate queue. **In the second mode, packets can be reclassified based on a default CoS value assigned to the ingress port** by the network administrator. In the case of frames that arrive without a CoS value (such as untagged frames), these Cisco Catalyst 2950 Series switches support classification based on a default CoS value per port assigned by the network administrator. **After** the frames have been classified or reclassified using one of the above modes, **they are assigned to the appropriate queue at the egress**. Cisco Catalyst 2950 Series switches support **four egress queues**, which allow the network administrator to be more discriminating and granular in assigning priorities for the various applications on the LAN. **Strict Priority Scheduling** configuration ensures that time-sensitive applications, such as voice, always follow an expedited path through the switch fabric. **Weighted Round Robin (WRR) scheduling**, another significant enhancement, ensures that lower-priority traffic receives attention without comprising the priority settings administered by a network manager. These features allow network administrators to prioritize mission-critical, time-sensitive traffic, such as voice (IP telephony traffic), enterprise resource planning (Oracle, SAP, etc.), and computer-assisted design and manufacturing, over less time-sensitive applications such as FTP or e-mail (Simple Mail Transfer Protocol).

Layer 2 QoS	• Support for reclassifying frames is based either on 802.1p class-of-service (CoS) value or default CoS value per port assigned by network manager. • Four queues per egress port are supported in hardware. • The Weighted Round Robin (WRR) scheduling algorithm ensures that low-priority queues are not starved. • Strict priority queue configuration via Strict Priority Scheduling ensures that time-sensitive applications such as voice always follow an expedited path through the switch fabric.

Cisco Catalyst 2950EI

Network Control Through Advanced Quality of Service and Rate Limiting

The Cisco Catalyst 2950 Series offers superior, highly granular **QoS based on Layer 2 to 4 information**, helping to ensure that network traffic is classified and prioritized, and that congestion is avoided in the best possible manner. Configuration of QoS is greatly simplified through **automatic QoS (auto-QoS), a feature that detects Cisco IP phones and automatically configures the switch for the appropriate classification and egress queuing**. This optimizes traffic prioritization and network availability without the challenge of a complex configuration.

Cisco Catalyst 2950 Series switches can **classify, reclassify, police** (determine if the packet is in or out of predetermined profiles and affect actions on the packet), and **mark or drop** the incoming packets before the packet is placed in the shared buffer. Packet classification allows the network elements to discriminate between various traffic flows and to enforce **policies based on Layer 2 and Layer 3 QoS fields**.

To implement QoS, these switches first identify traffic flows, or packet groups, and **classify or reclassify these groups using the DiffServ Code Point (DSCP) field in the IP packet and/or the 802.1p class of service (CoS) field in the Ethernet packet**. **Classification and reclassification can also be based on criteria as specific as the source/destination IP address, source/destination MAC address, or the Layer 4 Transmission Control Protocol (TCP)/User Datagram Protocol (UDP) ports. At the ingress (incoming port) level, the Catalyst switches also perform policing and marking of the packet.**

Cisco Catalyst 2950EI

After the packet goes through classification, policing, and marking, it is assigned to the appropriate queue before exiting the switch. Cisco Catalyst 2950 Series switches support **four egress (outgoing port) queues per port**, which allows the network administrator to be more discriminating and specific in assigning priorities for the various applications on the LAN. At the egress level, the switch performs scheduling, which is an algorithm (process) that determines the order in which the queues are processed. The switches support **Weighted Round Robin (WRR) scheduling or strict priority scheduling**. The WRR scheduling algorithm ensures that lower-priority packets are not entirely starved for bandwidth and are serviced without compromising the priority settings administered by the network manager. Strict priority scheduling ensures that the highest priority packets will always get serviced ahead of all other traffic, and that the other three queues will be serviced using WRR best effort.

These features allow network administrators to prioritize mission-critical and/or bandwidth-intensive traffic, such as enterprise resource planning (ERP) (Oracle, SAP, and so on), voice (IP telephony traffic), and CAD/CAM over less time-sensitive applications such as FTP or e-mail (Simple Mail Transport Protocol [SMTP]). For example, it would be highly undesirable to have a large file download destined to one port on a wiring closet switch and have quality implications, such as increased latency in voice traffic, destined to another port on this switch. This condition is avoided by ensuring that voice traffic is properly classified and prioritized throughout the network. Other applications, such as Web browsing, can be treated as low-priority and handled on a best-effort basis.

Cisco Catalyst 2950EI

Cisco Catalyst 2950 Series switches are **capable of allocating bandwidth based on several criteria, including MAC source/destination address, IP source/ destination address, and TCP/UDP port number**. Bandwidth allocation is essential in network environments requiring service-level agreements, or when it is necessary for the network manager to control the bandwidth given to certain users. Cisco Catalyst 2950 Series switches support **up to six policers per Fast Ethernet port and up to 60 policers on a Gigabit Ethernet port**. This gives the network administrator granular control of LAN bandwidth.

Cisco Catalyst 2950EI

Overview	• Cisco Catalyst 2950 Series switches support the aggregate QoS model by enabling classification, policing/metering, and marking functions on a per-port basis at ingress and the queuing/scheduling function at egress. • The switches support configuring QoS ACPs on all ports to ensure proper policing and marking on a per-packet basis using ACPs. Up to four ACPs per switch are supported in configuring either QoS ACPs or security filters. • Auto-QoS greatly simplifies the configuration of QoS in VoIP networks by issuing interface and global switch commands that allow the detection of Cisco IP phones, the classification of traffic, and egress queue configuration.
QoS Classification Support at Ingress	• The switches support QoS classification of incoming packets for QoS flows based on Layer 2, Layer 3, and Layer 4 fields. • The following Layer 2 fields (or a combination) can be used for classifying incoming packets to define QoS flows: source/destination MAC address, or 16-bit Ethertype. • The switches support identification of traffic based on Layer 3 type of service (ToS) field DSCP values. • The following Layer 3 and Layer 4 fields (or a combination) can be used to classify incoming packets to define QoS flows: source/destination IP address, TCP source/destination port number, or UDP source/destination port number.
QoS Metering and Policing at Ingress	• Support for metering and policing of incoming packets restricts incoming traffic flows to a certain rate. • The switches support up to six policers per Fast Ethernet port, and 60 policers on a Gigabit Ethernet port. • The switches offer granularity of traffic flows at 1 Mbps on Fast Ethernet ports, and 8 Mbps on Gigabit Ethernet ports.

Cisco Catalyst 2950EI

QoS Marking at Ingress	• The switches support marking and remarking packets based on the state of policers and meters. • The switches support marking and remarking based on the following mappings: from DiffServ Code Point (DSCP) to 802.1p, and 802.1p to DSCP. • The switches support 14 well-known and widely used DSCP values. • The switches support classifying or reclassifying packets based on default DSCP per port. They also support classification based on DSCP values in the ACL. • The switches support classifying or reclassifying frames based on default 802.1p value per port. • The switches support 802.1p override at ingress.
QoS Scheduling Support at Egress	• Four queues per egress port are supported in hardware. • The WRR queuing algorithm ensures that low-priority queues are not starved. • Strict-priority queue configuration via Strict Priority Scheduling ensures that time-sensitive applications such as voice always follow an expedited path through the switch fabric.
Sophisticated Traffic Management	• The switch offers the ability to limit data flows based on MAC source/destination address, IP source/destination address, TCP/UDP port numbers, or any combination of these fields. • The switch offers the ability to manage data flows asynchronously upstream and downstream from the end station or on the uplink.

Cisco Catalyst 3560

Advanced QoS

The Cisco Catalyst 3560 offers superior multilayer, granular QoS features to help ensure that network traffic is classified and prioritized, and that congestion is avoided in the best possible manner. Configuration of QoS is greatly simplified through automatic QoS (Auto QoS), a feature that detects Cisco IP phones and automatically configures the switch for the appropriate classification and egress queuing. This optimizes traffic prioritization and network availability without the challenge of a complex configuration.

The Cisco Catalyst 3560 can classify, reclassify, police, mark, queue, and schedule incoming packets, and can queue and schedule packets at egress. **Packet classification allows the network elements to discriminate between various traffic flows and enforce policies based on Layer 2 and Layer 3 QoS fields.**

To implement QoS, the Cisco Catalyst 3560 Series Switch first identifies traffic flows or packet groups, and **classifies or reclassifies these groups using the Differentiated Services Code Point (DSCP) field or the 802.1p Class of Service (CoS) field. Classification and reclassification can be based on criteria as specific as the source or destination IP address, source or destination MAC address, or the Layer 4 TCP or UDP port. At the ingress, the Cisco Catalyst 3560 also polices to determine whether a packet is in or out of profile, marks to change the classification label, passes through or drops out of profile packets, and queues packets based on classification.** Control- and data-plane ACLs are supported on all ports to help ensure proper treatment on a per-packet basis.

Cisco Catalyst 3560

The Cisco Catalyst 3560 supports **four egress queues per port**, allowing the network administrator to be more discriminating and specific in assigning priorities for the various applications on the LAN. At egress, the switch performs scheduling and congestion control. Scheduling is an algorithm or process that determines the order in which the queues are processed. The Cisco Catalyst 3560 Series Switch **supports shaped round robin (SRR) and strict priority queuing**. The SRR algorithm helps ensure differential prioritization.

[...]

The Cisco Catalyst 3560 Series **can perform rate limiting through its support of the Cisco Committed Information Rate (CIR) function. Through CIR, bandwidth can be guaranteed in increments as low as 8 kbps. Bandwidth can be allocated based on several criteria, including MAC source address, MAC destination address, IP source address, IP destination address, and TCP or UDP port number**. Bandwidth allocation is essential when network environments require service-level agreements or when it is necessary for the network manager to control the bandwidth given to certain users.

Cisco Catalyst 3560

Advanced QoS	• Standard 802.1p CoS and DSCP field classification are provided, using marking and reclassification on a per-packet basis by source and destination IP address, source and destination MAC address, or Layer 4 TCP or UDP port number. • Cisco control- and data-plane QoS ACLs on all ports help ensure proper marking on a per-packet basis. • Four egress queues per port enable differentiated management of up to four traffic types. • SRR scheduling helps ensure differential prioritization of packet flows by intelligently servicing the ingress and egress queues. • Weighted tail drop (WTD) provides congestion avoidance at the ingress and egress queues before a disruption occurs. • Strict priority queuing guarantees that the highest-priority packets are serviced ahead of all other traffic. • There is no performance penalty for highly granular QoS functions.
Granular Rate Limiting	• The Cisco Committed Information Rate (CIR) function guarantees bandwidth in increments as low as 8 kbps. • Rate limiting is provided based on source and destination IP address, source and destination MAC address, Layer 4 TCP and UDP information, or any combination of these fields, using QoS ACLs (IP ACLs or MAC ACLs), class maps, and policy maps. • Asynchronous data flows upstream and downstream from the end station or on the uplink are easily managed using ingress policing and egress shaping. • Up to 64 aggregate or individual policers are available per Fast Ethernet or Gigabit Ethernet port.