

Ποιότητα Υπηρεσίας
στο Διαδίκτυο
Quality of Service (QoS)

Το Διαδίκτυο : δίκτυο βέλτιστης προσπάθειας

➤ Συνταγή επιτυχίας : Απλό δίκτυο & έξυπνες εφαρμογές

- **QoS** : μερικές εφαρμογές λειτουργούν πολύ καλύτερα όταν οι καθυστερήσεις είναι περιορισμένες και η ρυθμαπόδοση επαρκής

- Απαιτήσεις διαφορετικών εφαρμογών

R_{min} : Ελάχιστη ρυθμαπόδοση D_{max} : Μέγιστη καθυστέρηση

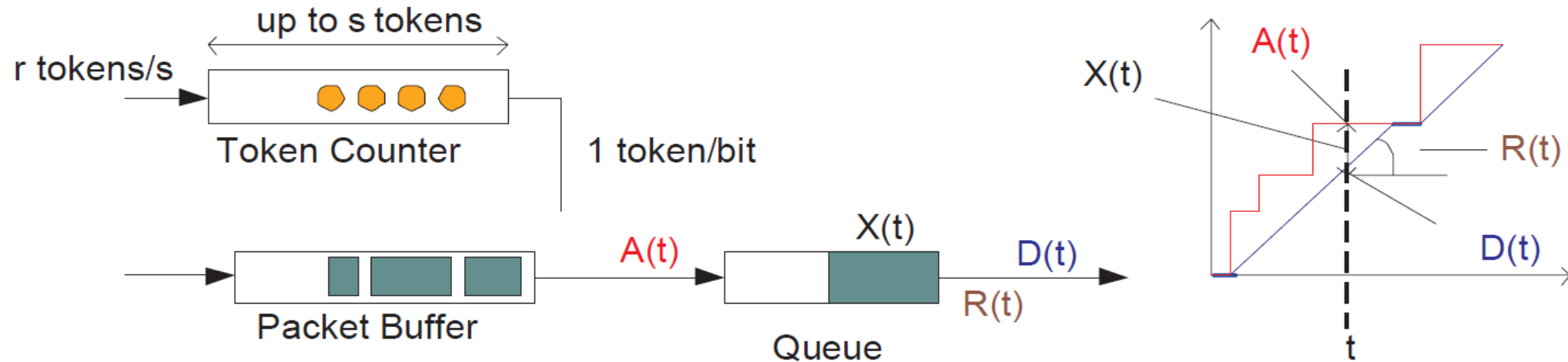
Application	R_{min} (Kbps)	D_{max} (s)
Video Streaming	60	10
VoIP	10	0.2
Download	30	200
Web Browsing	50	2
Video Conference	80	0.2

Πρόκληση: QoS με απλές τροποποιήσεις πρωτοκόλλων ???

- Δυστοκία υλοποιήσεων : επιπλοκές στη διαχείριση & στη συμφωνία μεταξύ ISP
- **IntServ** (for integrated services): users request a minimum bandwidth along a path (reservation protocol - **RsVP**) and promise to regulate their flows ---like Virtual Circuits: too complicated does not scale
- **DiffServ** (for differentiated services) : mark packets and use priority queuing at routers

Ρύθμιση Ροών
&
το Σύστημα Αναμονής WFQ

Μηχανισμός Κουπονιών : Ρύθμιση Αφίξεων στην Ουρά



$R(t)$: service rate (bps)

$D(t)$: number of bits that have departed from the queue in $[0, t]$

$X(t)$: number of bits in the queue at time t

$A(t)$: number of bits that enter the queue in $[0, t]$

Μηχανισμός Κουπονιών : Ρύθμιση Αφίξεων στην Ουρά

$$A(t) - A(u) \leq s + r(t - u) \quad \text{for all } 0 \leq u < t$$

Απόδειξη: ο μετρητής κουπονιών έχει το πολύ s κουπόνια την χρονική στιγμή u και θα συλλέξει $r(t - u)$ περισσότερα κουπόνια κατά τη διάρκεια $[u, t]$. Δεδομένου ότι κάποιος χρειάζεται ένα κουπόνι ανά bit που στέλνει κάποιος στην ουρά, ο **μέγιστος αριθμός bits** που μπορούν να μετακινηθούν από το buffer πακέτων στην ουρά κατά τη διάρκεια του $[u, t]$ ισούται με $s + r(t - u)$.

$$\text{Burst} \leq s$$

$$\text{Arrival rate} \leq r$$

Fact 1 Assume that $R(t) \geq r_1 \geq r$ for all $t \geq 0$ and $X(0) = 0$. Then $X(t) \leq s$ for all $t \geq 0$. Moreover, the queuing delay of every packet in the queue is at most equal to s/r_1 .

Proof:

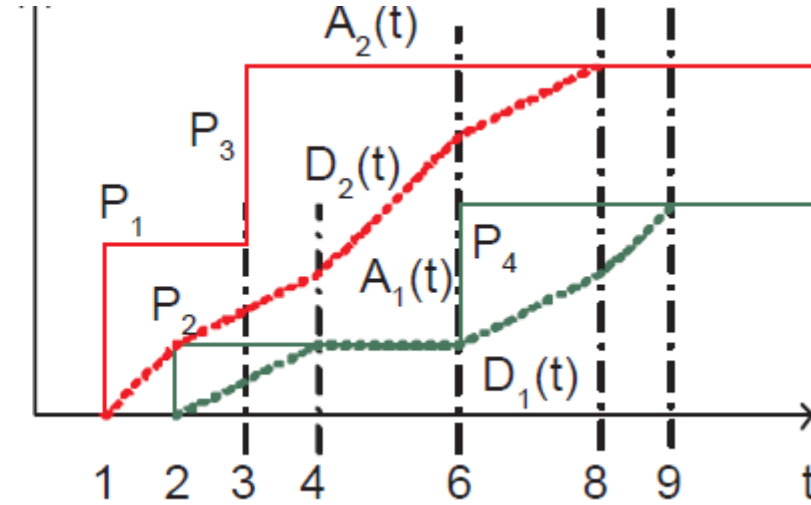
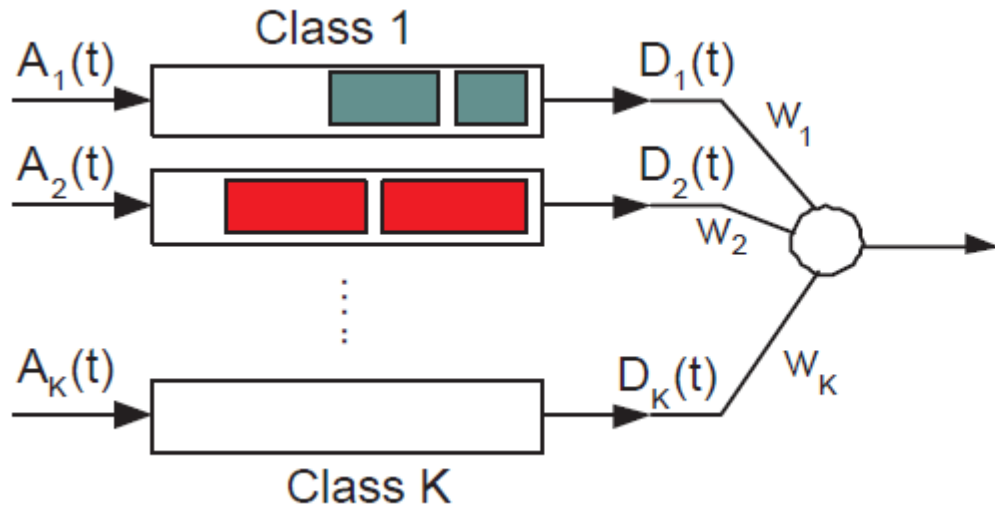
Fix some time $t > 0$ such that $X(t) > 0$. Let u be the last time before time t that $X(u) = 0$. That is, $X(u) = 0$ and $X(v) > 0$ for $v \in (u, t)$. Let $D(t)$ be the number of bits that leave the queue in $[0, t]$, for $t \geq 0$. Note that the number of bits $D(t) - D(u)$ that left the queue during $[u, t]$ is at least equal to $r_1(t - u)$ since the service rate is at least equal to r_1 whenever the queue is nonempty. Also, the number of bits $A(t) - A(u)$ that entered the queue in $[u, t]$ is at most $s + r(t - u)$. Since $X(u) = 0$, we see that

$$X(t) = A(t) - A(u) - [D(t) - D(u)] \leq s + r(t - u) - r_1(t - u) \leq s$$

since $r_1 \geq r$.

The queuing delay of a packet that enters the queue at time t , when the backlog is $X(t)$, is less than $X(t)/r_1$ since the queue clears that backlog at least at rate r_1 . We have seen that $X(t) \leq s$. Consequently, the queuing delay is always bounded by s/r_1 . □

GPS : Generalized Processor Sharing



➤ Στιγμιαίος ρυθμός εξυπηρέτησης της κλάσης k

$$w_k C/W$$

C : ταχύτητα της γραμμής εξόδου του δρομολογητή

W : άθροισμα των βαρών w_i των ουρών που δεν είναι άδειες σε χρόνο t

Fact 2 Assume that the traffic of class k is regulated with parameters (s_k, r_k) such that $\rho_k := w_k C / (\sum_j w_j) \geq r_k$. Then the backlog of class k never exceeds s_k and its queuing delay never exceeds s_k / ρ_k .

WFQ : Weighted Fair Queueing

GPS is not implementable. Weighted Fair Queueing approximates GPS. WFQ is defined as follows. The packets are classified and queued as in GPS. The scheduler transmits one packet at a time, at the line rate. Whenever it completes a packet transmission, the scheduler starts transmitting the packet that GPS would complete transmitting first among the remaining packets.

Fact 3 *Let F_k and G_k designate the completion times of packet P_k under WFQ and GPS, respectively, for $k \geq 1$. Assume that the transmission times of all the packets are at most equal to T . Then*

$$F_k \leq G_k + T, k \geq 1. \quad (1)$$

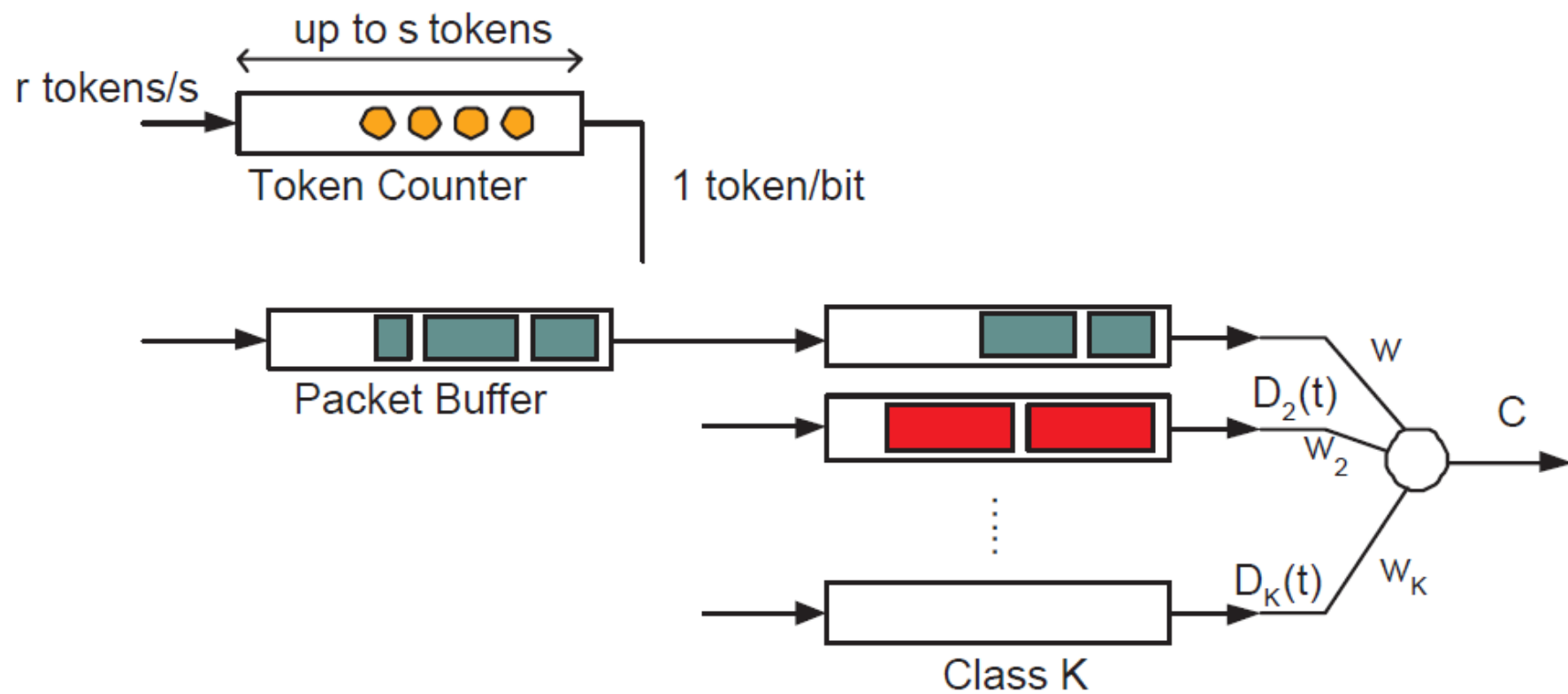


Figure 3: Regulated traffic and WFQ scheduler.

Fact 4 Assume that $r < \rho := wC/W$ where W is the sum of the scheduler weights. Then the maximum queuing delay per packet is

$$\frac{s}{\rho} + \frac{L}{C}$$

where L is the maximum number of bits in a packet/

Net Neutrality : «ίση μεταχείριση για όλα τα πακέτα»

- Εμποδίζει τους «μεγάλους» να ελέγξουν το Διαδίκτυο
- Υπερασπίζει την ελευθερία του λόγου και το δημόσιο αγαθό του Διαδικτύου για τους οικονομικά ασθενέστερους
- Απαγορεύει την επωφελή για τους χρήστες διαφοροποίηση υπηρεσιών
- Οι ISPs δεν μπορούν να χρεώσουν για καλύτερες υπηρεσίες: έλλειψη κινήτρου για επένδυση σε υποδομές
- Απαιτητικό περιεχόμενο δεν μπορεί να μεταδοθεί ποιοτικά
- Ρυθμιστικό πλαίσιο που επιτρέπει διαφοροποίηση υπηρεσιών και διατηρεί ένα ποσοστό της χωρητικότητας ουδέτερο και ανοιχτό ?