

Ποιότητα υπηρεσίας

Μιλτιάδης Αναγνώστου

27 Μαΐου 2022

Απαιτήσεις QoS εφαρμογής

Υπερβολική διάθεση πόρων

Ολοκληρωμένες υπηρεσίες

RSVP - Resource reSerVation Protocol

Διαφοροποιημένες υπηρεσίες

Ποιότητα υπηρεσίας - ποιότητα εμπειρίας εφαρμογής

- ▶ Υπάρχουν υπηρεσίες και εφαρμογές που πρέπει να φτάνουν στον πελάτη με συγκεκριμένες προδιαγραφές ποιότητας, π.χ. ήχος με συγκεκριμένο εύρος ζώνης συχνοτήτων και χωρίς διακοπές, video με συγκεκριμένο ρυθμό πλαισίων και ανάλυση, εντολές προς καθοδηγούμενα οχήματα με όριο στην καθυστέρηση άφιξης της εντολής κ.λπ.
- ▶ Η ποιότητα μιας υπηρεσίας ή ενός προϊόντος δεν ποσοτικοποιείται πάντοτε (π.χ. ενός έργου τέχνης). Όταν ποσοτικοποιείται, αυτό γίνεται με παραμέτρους που προσιδιάζουν στη συγκεκριμένη υπηρεσία ή το προϊόν. Συνήθως οι παράμετροι πρέπει να μένουν μέσα σε καθορισμένα αποδεκτά διαστήματα τιμών.
- ▶ Όταν μια υπηρεσία ή ένα σύστημα λειτουργεί σε δικτυακό περιβάλλον ο σχεδιασμός και η υλοποίηση πρέπει να γίνουν έτσι ώστε να ικανοποιούνται οι απαιτήσεις ποιότητας. Τίθενται τότε ερωτήματα όπως:
 - ▶ Μπορεί το δίκτυο να τις υποστηρίξει;
 - ▶ Πώς αυτές οι απαιτήσεις ποιότητας μεταφράζονται σε μετρήσιμα μεγέθη της λειτουργίας του δικτύου;

Παράδειγμα: Υπηρεσία video streaming

Μια υπηρεσία video streaming έχει τις εξής προδιαγραφές ποιότητας:

- ▶ Να μη γίνονται διακοπές μεγαλύτερες των 3 sec. Να γίνονται με συχνότητα το πολύ μια φορά στις 100 ώρες προβολής video.
- ▶ Η ποιότητα εικόνας να είναι ανάλυσης HD1080 ή ανώτερη με ρυθμό ως και 60 πλαίσια ανά sec. Υποβάθμιση στο αμέσως κατώτερο επίπεδο ανάλυσης ή σε κατώτερο ρυθμό είναι ανεκτή μόνο στο 1/200 του χρόνου.
- ▶ Η προβολή ενός video να αρχίζει το πολύ 20 sec μετά τη σχετική εντολή του πελάτη.

Έστω ότι η υπηρεσία αυτή προσφέρεται πάνω από ένα δίκτυο με εγγυημένες τις εξής παραμέτρους επίδοσης δικτύου: Ελάχιστος μέσος ρυθμός μετάδοσης, μέγιστη διασπορά ρυθμού μετάδοσης, μέγιστη μέση καθυστέρηση, μέγιστη διασπορά της καθυστέρησης. Οι εγγυήσεις αυτές προσφέρονται σε διαφορετικά επίπεδα επίδοσης με διαφορετικό τιμολόγιο.

Ποιό επίπεδο επίδοσης θα συστήσει ο σχεδιαστής της υπηρεσίας στον πάροχο της υπηρεσίας; Ποια θα είναι η αναμενόμενη ποιότητα υπηρεσίας; Μπορούν να ικανοποιηθούν όλες οι απαιτήσεις; Πώς θα γίνει ο σχεδιασμός της υπηρεσίας ώστε να γίνεται πιο πιθανή η ικανοποίηση των προδιαγραφών;

Σημείωση ορολογίας

- ▶ Δυστυχώς ο όρος *ποιότητα υπηρεσίας* (quality of service - QoS) χρησιμοποιείται με δυο διαφορετικές σημασίες, ανάλογα με τα συμφραζόμενα (context):
 - ▶ Για να περιγράψει την ποιότητα της τελικής υπηρεσίας ή εφαρμογής που παρέχεται στον πελάτη.
 - ▶ Για να περιγράψει την ποιότητα της δικτυακής «υπηρεσίας» που υποστηρίζει την παραπάνω. Η δικτυακή υπηρεσία αποτελεί ένα επίπεδο επιδόσεων και περιγράφεται με ένα σύνολο παραμέτρων που περιγράφουν αυτές τις επιδόσεις του δικτύου. Την «υπηρεσία» προσφέρει και εγγυάται ο πάροχος δικτύου.
- ▶ Ο όρος *ποιότητα εμπειρίας* (quality of experience - QoE) υποδηλώνει τον βαθμό ικανοποίησης του πελάτη.

Ενδεικτικές απαιτήσεις εφαρμογών

Εφαρμογή	Εύρος ζώνης	Καθυστέρηση	Διασπορά καθ.	Πιθ. απώλειας
email	χαμηλό	αδιάφορη	αδιάφορη	χαμηλή
Αποθήκευση αρχείων	ποικίλο	χαμηλή	χαμηλή	πολύ χαμηλή
Audio streaming	μεσαίο	χαμηλή	χαμηλή	αρκετά χαμηλή
Video streaming	υψηλό	χαμηλή	χαμηλή	χαμηλή
VoIP	χαμηλό	χαμηλή	χαμηλή	χαμηλή

Υπερβολική διάθεση πόρων (overprovisioning)

- ▶ Η διαφοροποίηση της ποιότητας εξυπηρέτησης μεταξύ εφαρμογών με διαφορετικές απαιτήσεις μπορεί να επιτευχθεί με κατάλληλους μηχανισμούς ελέγχου της κίνησης στο δίκτυο.
- ▶ Εναλλακτικά μπορεί το δίκτυο να είναι «υπερπαραμετροποιημένο», δηλαδή να υιοθετηθούν υπερβολικές τιμές των παραμέτρων που χαρακτηρίζουν τους πόρους του, ώστε τελικά όλες οι εφαρμογές να εξυπηρετούνται «πολυτελώς».
- ▶ Για παράδειγμα, μπορεί να τοποθετηθούν ζεύξεις με χωρητικότητα πολλαπλάσια της μέγιστης αναμενόμενης κίνησης.
- ▶ Η παροχή ποιότητας μέσω υπερπαραμετροποίησης είναι
(α) δαπανηρή,
(β) παύει να ισχύει όταν αυξηθεί αρκετά η κίνηση στο δίκτυο και
(γ) δεν παρέχει εγγυήσεις ποιότητας σε όλα τα σενάρια.

Πρόβλεψη και σχεδιασμός για την ποιοτική παροχή υπηρεσιών από ένα δίκτυο

Το πρώτο βήμα για την παροχή ποιοτικών υπηρεσιών είναι ο κατάλληλος σχεδιασμός του δικτύου.

- ▶ Εφόσον υπάρχει η δυνατότητα να γίνει πρόβλεψη των εφαρμογών που θα εξυπηρετηθούν από ένα δίκτυο πριν να στηθεί το δίκτυο, πρέπει αυτό να σχεδιαστεί κατάλληλα.
- ▶ Οι εφαρμογές χαρακτηρίζονται (α) από την κίνηση που δημιουργούν ανά χρήστη, (β) από την ένταση της ζήτησης από σύνολα χρηστών, (γ) από τη γεωγραφική διασπορά των χρηστών.
- ▶ Στη βάση αυτών των δεδομένων ένα δίκτυο μπορεί να σχεδιαστεί ώστε να μπορεί να εξυπηρετεί συγχρόνως ένα σύνολο εφαρμογών.
- ▶ Ωστόσο κατά κανόνα τα δίκτυα είναι εξελισσόμενοι οργανισμοί και οι (συχνά απρόβλεπτες) εφαρμογές προστίθενται με το χρόνο.

Λειτουργία του δικτύου για την παροχή υπηρεσιών

Όταν παγιωθεί η κατασκευή ενός δικτύου το δίκτυο σε λειτουργία μπορεί να κάνει τα εξής:

- ▶ Να αποφασίζει αν μπορεί να δεχτεί ένα ρεύμα κίνησης.
- ▶ Να παραχωρεί πόρους για την εξυπηρέτηση κίνησης.
- ▶ Να ρυθμίζει την κίνηση μέσα στο δίκτυο, μέσω δρομολόγησης κ.α. μεθόδων. Με την τεχνολογία των software defined networks και του υπολογιστικού νέφους επεκτείνονται οι δυνατότητες προσαρμογής των πόρων του δικτύου στην προσφερόμενη κίνηση.

Αλγόριθμοι χρονοπρογραμματισμού πακέτων

Εάν το δίκτυο έχει τη δυνατότητα να υποχρεώσει τα πακέτα ενός εισερχόμενου ρεύματος κίνησης να ακολουθήσουν την ίδια διαδρομή, μπορεί να κάνει κατάλληλη διαχείριση των εξής πόρων:

Εύρος ζώνης πρέπει τουλάχιστον κατά μέσον όρο η χωρητικότητα κάθε ζεύξης να υπερκαλύπτει το άθροισμα των μέσων ρυθμών των ρευμάτων που την χρησιμοποιούν.

Καταχωρητές που θα απορροφήσουν τις ανομοιομορφίες στην κίνηση. Η προστασία συγκεκριμένων ρευμάτων μπορεί να γίνει με διάθεση καταχωρητών στα ρεύματα αυτά.

Κύκλοι CPU δηλ. υπολογιστικοί πόροι για την επεξεργασία των πακέτων κάθε ρεύματος.

Σε software defined networks οι πόροι μπορούν να διατίθενται με μεγαλύτερη ευελιξία.

Διαχείριση προτεραιοτήτων σε κόμβο

- ▶ Ένας κόμβος μπορεί να επιλέγει τη σειρά με την οποία θα προωθήσει στην κατάλληλη εξερχόμενη ζεύξη (δεδομένης μιας δρομολόγησης) κάθε εισερχόμενο πακέτο.
- ▶ Η προφανής σειρά είναι να προωθείται πρώτο το πακέτο που έχει φτάσει πρώτο (First-In-First-Out, First-Come-First-Served).
- ▶ Ωστόσο η λογική FIFO «αδικεί» τις λιγότερο πυκνές ροές, διότι
 1. αυξάνει την πιθανότητα απόρριψης των πακέτων τους λόγω πλήρωσης των καταχωρητών και
 2. καθυστερεί υπερβολικά όσα δεν απορρίπτονται.

Δίκαιη αναμονή (Fair queueing)

- ▶ Κάθε ρεύμα έχει χωριστή ουρά προς την μια έξοδο.
- ▶ Επιλέγεται ένα πακέτο από κάθε ρεύμα κυκλικά (Nagle, 1987).
- ▶ Κατά συνέπεια όλα τα ρεύματα αναχωρούν από την έξοδο με τον ίδιο ρυθμό (εφόσον έχουν πακέτα προς αποστολή).



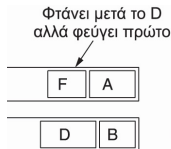
- ▶ Μειονέκτημα του αλγόριθμου είναι ότι οι ροές με μεγάλα πακέτα κερδίζουν.

Δίκαιη αναμονή (Fair Queueing) κατά Demers

- ▶ Υπολογίζεται πότε θα τελειώνει κάθε πακέτο αν μεταδιδόταν ένα bit σε κάθε κύκλο.
- ▶ Επιλέγεται να μεταδοθεί πιο πριν το πακέτο που θα τελειώνει πιο πριν.
- ▶ Δηλαδή για κάθε χωριστό ρεύμα k υπάρχει χωριστή ουρά και ο χρόνος περάτωσης F_i^k του i -οστού πακέτου του ρεύματος k υπολογίζεται με τον τύπο

$$F_i^k = \max\{A_i^k, F_{i-1}^k\} + L_i^k$$

όπου A_i^k είναι ο χρόνος άφιξης του εν λόγω πακέτου και L_i^k είναι η διάρκεια μετάδοσής του στο εικονικό σχήμα.



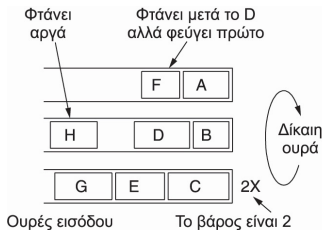
- ▶ Κατά συνέπεια τα πακέτα ενώ έχουν φτάσει με τη σειρά A, B, D, F

Σταθμισμένη δίκαιη αναμονή (Weighted Fair Queueing - WFQ)

- Είναι παραλλαγή του προηγ. σχήματος με τη δυνατότητα να δίνεται προτεραιότητα σε ορισμένα ρεύματα κατά παράγοντα W^k που πολλαπλασιάζει την ταχύτητα στο ρεύμα k :

$$F_i^k = \max\{A_i^k, F_{i-1}^k\} + L_i^k / W^k$$

- Στο πιο κάτω παράδειγμα $W^1 = W^2 = 1, W^3 = 2$.



Πακέτο	Χρόνος άφιξης	Μήκος	Χρόνος ολοκλήρωσης	Σειρά εξόδου
A	0	8	8	1
B	5	6	11	3
C	5	10	10	2
D	8	9	20	7
E	8	8	14	4
F	10	6	16	5
G	11	10	19	6
H	20	8	28	8

Χρονοπρογραμματισμός με προτεραιότητες ανά ρεύμα

- ▶ Κάθε πακέτο φέρει μια ένδειξη προτεραιότητας.
- ▶ Μέσα στο ίδιο ρεύμα τα πακέτα προωθούνται με FIFO.
- ▶ Μεταξύ πακέτων διαφορετικών ρευμάτων προηγείται το ανώτερης προτεραιότητας ανεξάρτητα από το χρόνο άφιξης.
- ▶ Τα ρεύματα υψηλής προτεραιότητας μπορούν να αποκλείσουν παντελώς τα ρεύματα χαμηλής προτεραιότητας αν εξαντλούν τη χωρητικότητα της εξόδου για ένα διάστημα.

Χρονοπρογραμματισμός βάσει χρονοσφραγίδων

- ▶ Κάθε πακέτο φέρει μια χρονοσφραγίδα και προωθούνται με τη σειρά που δείχνουν οι χρονοσφραγίδες.
- ▶ Κατά συνέπεια σε ένα κόμβο αποκτούν προτεραιότητα τα πακέτα που είναι πιο καθυστερημένα.
- ▶ Το αποτέλεσμα είναι ότι όλα τα πακέτα τείνουν προς την ίδια μέση καθυστέρηση.

Γενικό σχήμα αποδοχής μιας ροής

1. Ο χρήστης προσφέρει μια ροή που συνοδεύεται από απαιτήσεις ποιότητας υπηρεσίας (QoS).
 2. Το δίκτυο αποφασίζει αν μπορεί να δεχτεί τη ροή δεδομένων των ήδη εγκατεστημένων ροών και των χωρητικοτήτων.
 3. Εάν η ροή γίνει δεκτή, γίνονται οι σχετικές κρατήσεις πόρων στους δρομολογητές κατά μήκος της διαδρομής ή ενός συνδυασμού διαδρομών.
- ▶ Οι σχετικοί πόροι είναι εύρος ζώνης σε ζεύξεις, χώρος σε καταχωρητές και κύκλοι CPU.
 - ▶ Ωστόσο οι πόροι και οι παράμετροι ποιότητας συνδέονται συνήθως μόνο στατιστικά, όχι με απόλυτες εγγυήσεις.
 - ▶ Η διαδρομή μπορεί να μην είναι η συντομότερη ή πιο οικονομική, αλλά αυτή που εξασφαλίζει την ποιότητα (μία ή συνδυασμός) - *δρομολόγηση QoS*.

Στατιστικές εγγυήσεις και ανοχή

- ▶ Για πολλές από τις απαιτήσεις το δίκτυο μπορεί να δώσει στατιστικές εγγυήσεις, π.χ. ότι η καθυστέρηση μεταξύ των δύο άκρων δεν θα υπερβεί τα 200 msec στο 99% των πακέτων.
- ▶ Η παροχή απόλυτων εγγυήσεων (στο 100% των πακέτων) θα οδηγούσε στην αποδοχή λιγότερων ροών - ή μπορεί να είναι και αδύνατη.
- ▶ Ορισμένες εφαρμογές είναι ανεκτικές σε ελεγχόμενη υποβάθμιση της ποιότητας, π.χ. το video να πέσει από τα 30 πλαίσια ανά sec στα 25.

Προδιαγραφή ροής

Ένα παράδειγμα προδιαγραφής ροής δίνεται από τις RFC 2210/2211 για Ολοκληρωμένες Υπηρεσίες (Integrated Services) με πέντε παραμέτρους:

Παράμετρος	Μονάδα
Ρυθμός κουβά κουπονιών	Bytes/sec
Μέγεθος κουβά κουπονιών	Bytes
Μέγιστος ρυθμός μετάδοσης	Bytes/sec
Ελάχιστο μήκος πακέτου	Bytes
Μέγιστο μήκος πακέτου	Bytes

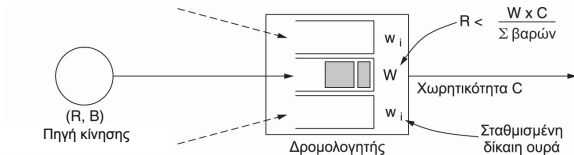
Μοντέλο Parekh-Gallagher

- ▶ Κάθε ρεύμα εισόδου περιορίζεται από ένα κουβά (R, B) .
- ▶ Σε κάθε νέα ροή διατίθεται ένα βάρος W σε σχήμα WFQ. Το βάρος πρέπει να ικανοποιεί την

$$R < \frac{W \times C}{\sum \text{βαρών}}$$

διαφορετικά η ροή δεν γίνεται δεκτή.

- ▶ Αν η κίνηση φτάνει σε ριπές, η μέγιστη καθυστέρηση είναι αυτή της απορρόφησης μιας ριπής, ήτοι B/R . Οι ριπές στη συνέχεια εξομαλύνονται και δεν επιφέρουν περαιτέρω καθυστερήσεις στη διαδρομή.

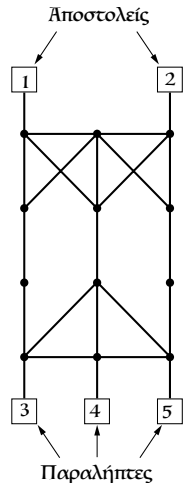


Ολοκληρωμένες υπηρεσίες

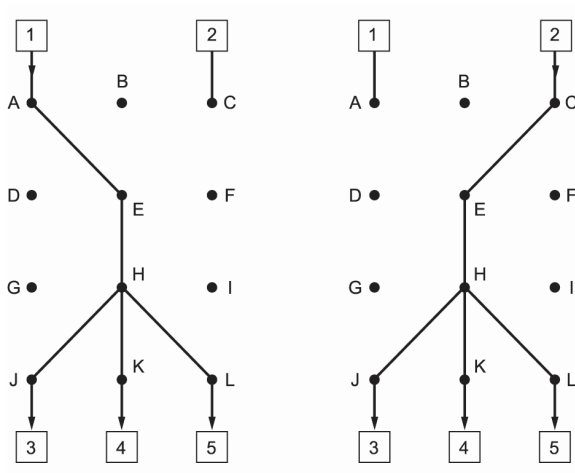
- ▶ Μεταξύ 1995 και 1997 η IETF (Internet Engineering Task Force) παρήγαγε τις RFC 2205 ως 2212 με αντικείμενο τη διευκόλυνση εφαρμογών για streaming πολυμέσων, με δυνατότητα πολυεκπομπής, π.χ. μετάδοσης ενός τηλεοπτικού προγράμματος προς πολλούς αποδέκτες.
- ▶ Μια από τις απαιτήσεις ήταν η ευελιξία στις ομάδες αποδεκτών σε πραγματικό χρόνο.

RSVP - Resource reSerVation Protocol

- ▶ Το πρωτόκολλο κράτησης πόρων RSVP επιτρέπει σε μια ομάδα αποστολέων να μεταδίδει ρεύματα προς ομάδες αποδεκτών και ενώ οι τελευταίοι μπορούν αλλάζουν το επιθυμητό τους ρεύμα.
- ▶ Η μετάδοση γίνεται έτσι ώστε να βελτιστοποιείται η χρήση του εύρους ζώνης και να μη συμβαίνει συμφόρηση.
- ▶ Για την μετάδοση αξιοποιείται κάποια προσέγγιση του δέντρου Steiner (spanning tree προς επιλεγμένους κόμβους).
- ▶ Ο αποστολέας βάζει στα πακέτα τη διεύθυνση μιας ομάδας των αποδεκτών. Στη συνέχεια ο αλγόριθμος πολυεκπομπής βρίσκει το κατάλληλο δέντρο.



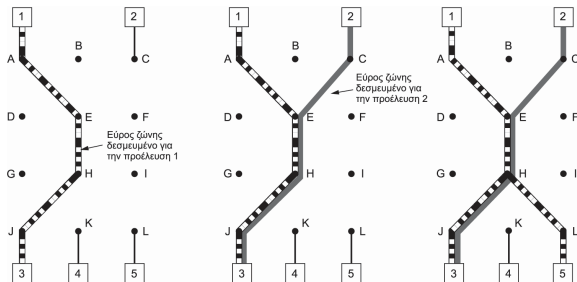
Δέντρα πολυεκπομπής για αποστολές 1 και 2



Μηνύματα κράτησης

- ▶ Καθένας από τους παραλήπτες μπορεί να ζητήσει να γίνουν κρατήσεις ως τον αντίστοιχο αποστολέα.
- ▶ Το μήνυμα χρησιμοποιεί τον αλγόριθμο προώθησης αντίστροφου μονοπατιού.
- ▶ Σε κάθε βήμα ο δρομολογητής κάνει τις απαραίτητες κρατήσεις, π.χ. με WFQ, αλλιώς απαντάει ότι απέτυχε.
- ▶ Εφόσον το μήνυμα φτάσει ως τον αποστολέα, θα έχουν γίνει όλες οι απαραίτητες κρατήσεις.

Διαδοχικές κρατήσεις πόρων



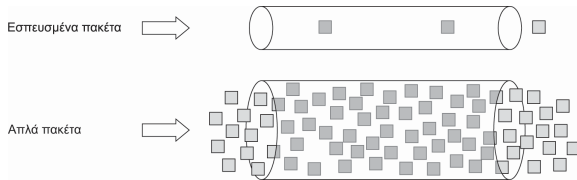
- ▶ Στο σχήμα αρχικά ο παραλήπτης 3 κάνει μια κράτηση ως τον αποστολέα 1, μετά πάλι ο 3 ως τον 2 για άλλο ρεύμα και τέλος ο 5 προς τον 1 για το πρώτο ρεύμα.
- ▶ Αν η τρίτη κράτηση δεν χρειάζεται πιο ψηλή ποιότητα από την πρώτη, οι κρατήσεις από τον H και πάνω δεν μεταβάλλονται.

Διαφοροποιημένες υπηρεσίες

- ▶ Με τις *διαφοροποιημένες υπηρεσίες* (differentiated services, RFC 2474, 2475 κ.α.) μειώνεται σε σύγκριση με το RSVP η πολυπλοκότητα των κρατήσεων και των λειτουργιών που απαιτούνται στους κόμβους.
- ▶ Κάθε πακέτο που κάνει χρήση των διαφοροποιημένων υπηρεσιών φέρει μια ένδειξη της κλάσης στην οποία ανήκει στο σχετικό πεδίο.
- ▶ Οι δρομολογητές ενός διαχειριστικού πεδίου (π.χ. ενός ISP) χειρίζονται τα πακέτα με ενδείξεις διαφοροποιημένων υπηρεσιών ανάλογα με την κλάση τους.
- ▶ Το πλεονέκτημα αυτού του πρωτοκόλλου είναι ότι δεν χρειάζεται αρχική εγκατάσταση διαδρομών με καθορισμένους πόρους. Εγγυάται όμως την ποιότητα μόνο στατιστικά.
- ▶ Τα πακέτα με διαφ. υπηρεσίες μπορεί να είναι υποχρεωμένα στην πηγή να τηρούν κάποιους περιορισμούς, π.χ. leaky bucket με συγκεκριμένες παραμέτρους.
- ▶ Αντίστοιχα ο ISP μπορεί να έχει μια πολιτική χρέωσης ή μια πολιτική άνω ορίου για τέτοιες υπηρεσίες.

Επισπευδόμενη προώθηση (Expedited forwarding)

- ▶ Τα πακέτα χωρίζονται σε δύο κλάσεις, κανονικά και επισπευδόμενα (RFC 3246).
- ▶ Τα επισπευδόμενα πακέτα προωθούνται κατά προτεραιότητα. Π.χ. μπορεί να μπαίνουν σε διαφορετικές ουρές από τα κανονικά και να δίνεται πιο συχνή εξυπηρέτηση στις ουρές των επειγόντων.
- ▶ Ο χαρακτηρισμός γίνεται είτε από την πηγή είτε από τον πρώτο δρομολογητή στη διαδρομή ενός ρεύματος.



Εξασφαλισμένη προώθηση (Assured forwarding)

- ▶ Τα πακέτα χωρίζονται σε 4 κατηγορίες προτεραιότητας (RFC 2597).
- ▶ Ορίζονται επίσης τρεις κατηγορίες διαγραφής όταν διαπιστώνεται συμφόρηση. Κατά συνέπεια τα πακέτα διαφοροποιούνται σε 12 κατηγορίες.
- ▶ Ο χαρακτηρισμός γίνεται στην πηγή ή στον πρώτο δρομολογητή.
- ▶ Στη συνέχεια τα πακέτα περνούν από έλεγχο που τα βάζει σε τόσο πιο ψηλή κατηγορία διαγραφής όσο ανήκουν σε ρεύμα που αποκλίνει από το επιθυμητό σχήμα.
- ▶ Ένας δρομολογητής μπορεί να προωθεί τα πακέτα ανάλογα με την κατηγορία τους (από τις 4), π.χ. με WFQ. Όταν αυξάνεται η συμφόρηση μπορεί να τα διαγράψει ανάλογα με την κατηγορία διαγραφής.

