

ΤΕΙ ΠΕΙΡΑΙΑ

ΤΜΗΜΑ ΑΥΤΟΜΑΤΙΣΜΟΥ

ΑΙΓΑΛΕΩ

**ΣΗΜΕΙΩΣΕΙΣ
ΘΕΩΡΙΑΣ ΔΙΑΔΙΚΤΥΑΚΩΝ ΠΡΩΤΟΚΟΛΛΩΝ**

ΑΘΗΝΑ 2010

Ι. Δ. Αγγελόπουλος, καθηγητής

Τα πνευματικά δικαιώματα αυτού του βιβλίου ανήκουν στον Ι. Δ. Αγγελόπουλο. Απαγορεύεται αυστηρά η ανατύπωση ή φωτοτύπηση για άλλο σκοπό πλην της χρήσης από σπουδαστές του τμήματος Αυτοματισμού ΤΕΙ Πειραιά.

Βιβλιογραφία:

- Σημειώσεις Διαδικτυακών Πρωτοκόλλων, Ι.Δ. Αγγελόπουλος. Διαθέσιμες σε μορφή pdf στην ιστοσελίδα: <http://:auto.teipir.gr> (ιδέ σχετικό Μάθημα, αριστερή στήλη)

A. Στα Αγγλικά:

1. Tanenbaum, Andrew S., "Computer Networks", Third Edition, Prentice Hall, 1996, ISBN 0133499456.
2. James Kurose, Keith Ross, "Computer Networking, 3rd Edition, Int. Edition, Addison Wesley, ISBN:0-321-26976-4
3. Computer Networking, 5th edition, James F. Kurose, Keith W. Ross, Addison-Wesley, ISBN: 0-13-607967-9
4. TCP/IP Illustrated, W. Richard Stevens, Addison-Wesley
5. Interworking with TCP/IP vol. I, Principles, protocols and Architecture, 2nd ed. Douglas Comer, Prentice-Hall Int.
6. Jean Walrand, "Communication Networks", 2nd ed., McGraw Hill, 1998, ISBN 0256205671.
7. Data Networks των D.Bertsekas-Galagher.Μπορείτε να το κατεβάσετε από την ιστοσελίδα: <http://www.pdf-top.com/ebook/data%20networks%20bertsekas/>
8. Handbook of Computer Communications Standards του W.Stallings, Volume 2 - Edition H.W.Sams.

B. Στα Ελληνικά:

9. Αρης Αλεξόπουλος, Γιώργος Λαγογιάννης, «Τηλεπικοινωνίες και Δίκτυα Υπολογιστών, Πέμπτη έκδοση, 1999, ISBN: 960-220-086-3
10. Α. Tanenbaum, «Δίκτυα Υπολογιστών, 3^η έκδοση, Παπασωτηρίου, Αθήνα 2000 (Ελληνική έκδοση του 1)
11. Larry Peterson, Bruce Davie, «Δίκτυα Υπολογιστών, μια προσέγγιση από τη σκοπιά των συστημάτων», 4^η Αμερικανική Έκδοση, εκδόσεις Κλειδάριθμος

Κεφάλαιο 1.

Τα πρωτόκολλα TCP/IP και το Internet

Εισαγωγή στο Διαδίκτυο και τα πρωτόκολλά του.

1.1 Τι είναι το Διαδίκτυο (Internet)

Με τον όρο Διαδίκτυο (Internet) υποδηλώνεται το σύνολο των δικτύων της ανθρωπότητας τα οποία είναι διασυνδεδεμένα και προσπελάσιμα μέσω των προγραμμάτων πλοήγησης (explorers). Αποτελεί απώτερη εξέλιξη της διάδοσης των υπολογιστών και των μεταξύ τους ανταλλαγών πληροφοριών που οδήγησαν στην δημιουργία δικτύων δεδομένων. Στα τέλη της δεκαετίας του 60 άρχισε να γίνεται φανερό ότι η αξιοποίηση των υπολογιστικών αλλά και των ανθρώπινων πόρων (δηλ. αποδοτικότερη χρήση του πολύτιμου χρόνου των ερευνητών) των Πανεπιστημίων και των ερευνητικών κέντρων των ΗΠΑ δεν μπορούσε να βελτιστοποιηθεί αν έπρεπε να υποχρεώνει τους ερευνητές να πηγαίνουν κοντά στους υπολογιστές απομακρυνόμενοι από τους εργαστηριακούς χώρους όπου τηρούσαν τα στοιχεία της εργασίας τους. Χρειαζόταν ένα ποιο ανθρωποκεντρικό μοντέλο εργασίας όπου τα εργαλεία και η πληροφορία πάνε εκεί που βολεύει τον άνθρωπο και όχι ο άνθρωπος εκεί που βολεύει τα μηχανήματα.

Ξεκίνησε λοιπόν το 1968 από υπηρεσίες του Αμερικανικού Πρακτορείου Προχωρημένων Ερευνών (ARPA) η προδιαγραφή ενός πειραματικού δικτύου με 4 κόμβους που θα συνέδεε αρχικά τα Πανεπιστήμια της Καλιφόρνια USLA και UCSB, το Πανεπιστήμιο της Utah και το Ίδρυμα Ερευνών Stanford. Την υλοποίηση ανέλαβε η εταιρεία Bolt Beranek & Newman και το δίκτυο που ονομάστηκε ARPANET άρχισε να λειτουργεί το επόμενο έτος (1969). Το δίκτυο αυτό έμελλε να αποτελέσει το σπέρμα δημιουργίας ενός από τα πιο πολύπλοκα ανθρώπινα δημιουργήματα που εκτείνεται σε όλη την υδρόγειο και σε κάθε γωνιά όπου υπάρχει ανθρώπινη δραστηριότητα. Σε συνέργεια με τις κερηκτικές εξελίξεις που ακολούθησαν στον τομέα των υπολογιστών και μικρουπολογιστών, των οπτικών και ασύρματων επικοινωνιών αλλά και των πολυμέσων (ψηφιακή εικόνα και ήχος) που οδήγησαν σε ραγδαία επέκταση των δικτύων δεδομένων, γιγάντωσαν αυτό το δίκτυο ώστε να λάβει τη μορφή ενός ενιαίου παγκοσμιοποιημένου δικτύου υπό καταναεμημένη διαχείριση χιλιάδων οργανισμών και εταιρειών. Αυτό το δίκτυο για το οποίο γύρω στο 1980 θα αρχίσει να χρησιμοποιείται ο όρος Internet, δεν αποτελεί απλά ένα μεγάλο τεχνολογικό βήμα αλλά και όχημα για την ραγδαία εξέλιξη και όλων των άλλων επιστημών, εμπορικών, κοινωνικών, πολιτιστικών αλλά και και παντοίων άλλων (όχι πάντοτε θετικών) ανθρωπίνων δραστηριοτήτων επιφέροντας πρωτοφανείς βελτιώσεις στην οργάνωση, παραγωγή, διοίκηση και πολλούς άλλους τομείς της κοινωνικής ζωής.

Αλλά δεν είναι μόνο οι νεώτερες επινοήσεις που ευνοήθηκαν και ευνόησαν το Διαδίκτυο. Η εμφάνισή του έδωσε νέα αξία και σε παρελθόντα προϊόντα της ανθρώπινης δημιουργικότητας αφού αύξησε την προσιτότητά τους. Σαν παράδειγμα αξίζει να αναφερθεί το Project Gutenberg (<http://gutenberg.net>) το οποίο έκανε διαθέσιμα στον καθένα σχεδόν όλα τα λογοτεχνήματα της ανθρωπότητας, από τον Ηρόδοτο και τον Σοφοκλή μέχρι τον Δάντη και τον Σαίξπηρ εκτός από αυτά τα οποία είναι ακόμη σε καθεστώς προστασίας συγγραφικών δικαιωμάτων (copyright). Δυστυχώς τα αρχεία είναι μόνο στην Αγγλική γλώσσα αφού η προσπάθεια είναι ιδέα ενός Αμερικανού του Michael Hart που αξιοποίησε μια χορηγία του Πανεπιστημίου του Illinois και ο οποίος προσέλκυσε με τα χρόνια εκατοντάδες άλλους Αγγλόφωνους εθελοντές που ήταν διατεθειμένοι να δακτυλογραφήσουν τα έργα ώστε να αποκτήσουν ηλεκτρονική μορφή (Με την επινόηση της οπτικής αναγνώρισης χαρακτήρων τα τελευταία χρόνια, η δουλειά των εθελοντών απλοποιήθηκε σημαντικά, αλλά ταυτόχρονα εξαντλήθηκαν και τα αξιόλογα έργα με αποτέλεσμα να προστίθενται πλέον και έργα δευτερευούσης ή και αμφίβολης αξίας - αλλά αυτό συμβαίνει πάντα σε όλες τις εκφάνσεις της ανθρώπινης δραστηριότητας, αφού τα ποιοτικά κριτήρια είναι πάντα υποκειμενικά).

Ο όρος Διαδίκτυο επελέγη για να τονίσει τη δυνατότητά του να επιτρέπει επικοινωνία δια μέσου ανομοιογενών φυσικών δικτύων ανατρέποντας τη μέχρι τότε κατάσταση, όπου, ενώ τα τερματικά επικοινωνούσαν απρόσκοπτα εντός κάθε τύπου δικτύου, δύο τερματικά σε δίκτυα διαφορετικής τεχνικής δεν είχαν συνήθως δυνατότητες επικοινωνίας. Την κατάσταση αυτή ανέτρεψε σταδιακά η εισαγωγή στα τερματικά λογισμικού που υλοποιούσε πρόδρομες μορφές των πρωτοκόλλων του μετέπειτα Διαδικτύου που έδιναν τη δυνατότητα ανταλλαγής δεδομενογραμμάτων (datagrams). Οι υπολογιστές χρησιμοποιούσαν το πρωτόκολλο 1822 (από τον αριθμό του εγγράφου που το προδιέγραφε) για να επικοινωνούν με τους κόμβους (οι οποίοι τότε ονομάζοντο IMPs (Internet Message Processors)). Οι πρώτες εφαρμογές που ήταν remote login, ftp και αργότερα electronic mail, ήταν ενσωματωμένες στο πρωτόκολλο Network Control Program (NCP) που μετεξελίχθηκε στο σημερινό TCP/IP (Internet Protocol δηλ. Πρωτόκολλο Διαδικτύου που δίνει το όνομα και στο όλο

δίκτυο και στο TCP (Transmission Control Protocol). Το ζευγάρι αυτό των πρωτοκόλλων αποτελεί την καρδιά του Διαδικτύου και το κύριο αντικείμενο της μελέτης μας, αφού η κατανόησή τους ισοδυναμεί με γνώση της λειτουργίας του Διαδικτύου.

Από τους βασικούς στόχους της δημιουργίας αυτού του δικτύου ήταν η ανθεκτικότητα σε πυρηνικές προσβολές γεγονός που αντανάκλαται στην φιλοσοφία σχεδίασης και στον αυτοτελή χαρακτήρα των μονάδων πληροφορίας (που είναι πακέτα με πλήρη διεύθυνση προέλευσης, προορισμού και χαρακτηριστικά διαχείρισης που αποκλήθηκαν δεδομενογράμματα κατ' αναλογία προς τα γράμματα του κοινού ταχυδρομείου που επίσης έχουν πλήρεις διευθύνσεις αποστολέα παραλήπτη).

Πρέπει να τονισθεί ότι τουλάχιστον στα αρχικά στάδια οι κόμβοι (και φυσικά τα τερματικά) του Διαδικτύου δεν βασιζόντουσαν σε ειδικά μηχανήματα αλλά κοινούς υπολογιστές με μόνο νέο στοιχείο το λογισμικό. Δεν υπήρχε ειδικό υλικό για το Διαδίκτυο. Αυτό που απαιτείτο ήταν η απλή προσθήκη του λογισμικού TCP/IP στα τερματικά καθώς και η τοποθέτηση ενός-δύο υπολογιστών σε θέση δρομολογητή ώστε να διασυνδέουν το συγκρότημα (Πανεπιστήμιο ή εταιρεία ή ερευνητικό οργανισμό) προς άλλα φυσικά δίκτυα μέσω μισθωμένων γραμμών που περαιτέρω διασυνδέονταν με άλλα δημιουργώντας ένα ενιαίο πια διαδίκτυο. Σημειωτέον ότι ο υπολογιστής που έμπαινε σε θέση δρομολογητή (π.χ. ένας mini computer PDP-11 της εταιρείας DEC) δεν έτρεχε κάποιο διαφορετικό λογισμικό αλλά το ίδιο πρόγραμμα το οποίο εμπεριείχε και την λειτουργία της δρομολόγησης η οποία ήταν ενσωματωμένη στο σετ πρωτοκόλλων TCP/IP. Απλά εκτός από την κάρτα του τοπικού δικτύου που είχαν αναγκαστικά όλοι οι υπολογιστές, ο δρομολογητής είχε και ένα modem ή κάρτα X.25 ή διεπαφή PCM που του επέτρεπε μέσω κάποιου τηλεφωνικού ή μισθωμένου κυκλώματος να συνδέεται προς την αντίστοιχη ομοειδή συσκευή κάποιου άλλου οργανισμού. Αυτός με τη σειρά του συνδεόταν και κάπου αλλού εις τρόπον ώστε να μπορεί ένα δεδομένογραμμα ξεκινώντας από οποιοδήποτε τέτοιο φυσικό υποδίκτυο να καταλήξει σε οποιοδήποτε άλλο. Έτσι το Διαδίκτυο συνέδεε υπάρχοντα φυσικά δίκτυα πάσης τεχνολογίας με την υλοποίηση των πρωτοκόλλων TCP/IP πάνω από ποικιλία υπολογιστών και φυσικών δικτύων. Τα δίκτυα αυτά, (π.χ. SNA, Ethernet, X.25, Token ring, κτλ), με την προσθήκη του λογισμικού που υλοποιούσε το TCP/IP αποκτούσαν κοινή γλώσσα. Η προσφυγή στην υποδομή μετάδοσης του τηλεφωνικού δίκτυο χρησιμοποιήθηκε πολλές φορές επικουρικά για διασύνδεση μέσω modem λόγω της μεγάλης διαδοχής του.

Η εκρηκτική εξέλιξη του Διαδικτύου οδήγησε σε πολλές μεταβολές και σήμερα η υλοποίηση του πρωτοκόλλου IP στους κόμβους (δρομολογητές) δεν βασίζεται πάντα σε λογισμικό. Παρότι μερικοί μικροί δρομολογητές σε μικρές εταιρείες εξακολουθούν να είναι κοινοί υπολογιστές με ειδικό λογισμικό, η διεκπεραίωση εκατομμυρίων δεδομενογραμμάτων το δευτερόλεπτο που απαιτούν οι σημερινές συνθήκες άλλαξε τη σχεδίαση των μεγάλων δρομολογητών οι οποίοι έγιναν ειδικά μηχανήματα με ειδικό υλικό και λογισμικό. Επίσης οι ζεύξεις ανάμεσα στους δρομολογητές βασίζονται σε ειδικές τεχνικές (IPoverSDH, IPoverWDM) και μόνο σε μικρές εταιρείες ή σπίτια βασίζονται σε modem ή PCM. Επίσης δημιουργήθηκαν εταιρείες παροχής υπηρεσιών Διαδικτύου (ISP = Internet Service Providers) που παρέχουν διασύνδεση σε χρήστες (σπίτια και εταιρείες) και άλλες που διασυνδέουν τις πρώτες και ονομάζονται NSP (Network Service Providers). Δηλαδή οι δεύτερες δεν διασυνδέουν τερματικά αλλά δρομολογητές των πρώτων δημιουργώντας ένα δεύτερο ιεραρχικό επίπεδο.

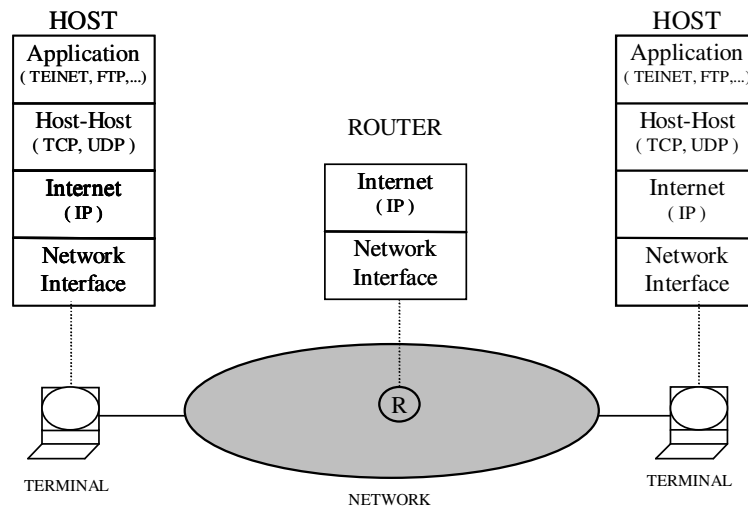
Φυσικά στα τερματικά η υλοποίηση με λογισμικό παρέμεινε καθώς και η συνήθεια να αποκαλούνται οι τερματικοί υπολογιστές που τρέχουν τα πρωτόκολλα του δικτύου (και κυρίως το TCP/IP) ως host computers δηλ. οικοδεσπότες, φιλοξενούντες το λογισμικό του δικτύου (σαν υπολογιστές γενικής χρήσεως που είναι τρέχουν και άλλο λογισμικό). Στην απόδοση στα ελληνικά προτιμήσαμε το σύντομο ξενιστής κατά μεταφορά του αντίστοιχου όρου της βιολογίας αντί του ομόρριζου περιφραστικού «φιλοξενών υπολογιστής» αλλά πολλές φορές θα λέμε υπολογιστής, ή τερματικό, ή σταθμός (station). Ο τελευταίος όρος είναι συνηθισμένος στα τοπικά δίκτυα.

Όπως ελέγχθη, η καρδιά του Διαδικτύου είναι το Internet Protocol (IP) που καλύπτει το 3ο στρώμα (δικτύου), και το Transmission Control Protocol (TCP) που υλοποιεί το 4ο στρώμα (μεταφοράς). Τα δύο αυτά πρωτόκολλα πρωτοεμφανίσθηκαν το 1974 και έχουν έκτοτε υποστεί πολλές βελτιώσεις. Τα πακέτα (δεδομενογράμματα) TCP/IP πρέπει να ενθυλακωθούν μέσα στα πλαίσια του εκάστοτε φυσικού δικτύου (π.χ. LAN, X.25, SNA, PCM ή απλά RS-232 σε modem πάνω από μισθωμένες ή και επιλεγόμενες τηλεφωνικές γραμμές κτλ.) για να διαβιβαστούν στο άλλο άκρο του αφού αλλιώς το φυσικό δίκτυο δεν θα ήξερε πώς να τα χειριστεί. Ωστόσο το πακέτο εκθυλακώνεται προτού παραδοθεί στο στρώμα IP στον ξενιστή ή στον δρομολογητή.

Το Διαδίκτυο δεν ακολουθούσε φυσικά την αρχιτεκτονική OSI αφού προηγήθηκε αυτής αλλά μια δική του που φαίνεται στο σχήμα 1.1.

Στο ανώτερο επίπεδο βρίσκονται όπως πάντα οι εφαρμογές π.χ FTP, Telnet, SMTP κτλ. (Αντιστοιχεί χονδρικά στα 3 ανώτερα στρώματα του OSI). Ακολουθεί το επίπεδο πρωτοκόλλων μεταξύ ξενιστών (host-to-host) δηλαδή απ' άκρου εις άκρο στο οποίο εδρεύουν οι λειτουργίες ελέγχου ροής και λαθών δηλαδή αυτό που αντιστοιχεί στις λειτουργίες του 4^{ου} στρώματος OSI. Εδώ χρησιμοποιείται το TCP όταν οι απαιτήσεις αξιοπιστίας είναι υψηλές και το UDP όταν ένα πιο ελαφρύ πρωτόκολλο χωρίς επιβεβαίωση είναι πιο κατάλληλο. Ακολουθεί το στρώμα δικτύου που εδώ είναι το στρώμα IP και εκτελεί τη παράδοση δεδομενογραμμάτων άνευ εγγυήσεων.

Στο κατώτερο επίπεδο (αντίστοιχο του φυσικού στρώματος και του στρώματος ζεύξης) βρίσκεται η πρόσβαση στο δίκτυο στην οποία ανήκουν η κάρτα δικτύου και το λογισμικό οδήγησης (driver) που εξαρτώνται από το εκάστοτε φυσικό δίκτυο που μπορεί να είναι όπως αναφέρθη μία απλή σειριακή πόρτα που οδηγεί σε ένα modem ή κάρτα τοπικού δικτύου ή X.25 κτλ. Το Διαδίκτυο δεν πρόσθεσε κάτι ιδιαίτερο σε αυτό το επίπεδο, απλά χρησιμοποίησε τα υπάρχοντα στοιχεία πατώντας σε αυτά. Η ικανότητα του Διαδικτύου να συνεργάζεται με κάθε φυσικό δίκτυο είναι και το μυστικό της τεράστιας επιτυχίας του. Είναι στην ουσία μία τεχνολογία διασύνδεσης δικτύων που δεν φιλοδοξεί να προδιαγράψει τα κατώτερα στρώματα και τα φυσικά μέσα αφήνοντάς τα να εξελίσσονται απρόσκοπτα. Η υλοποίηση μέσω λογισμικού παρέχει πλήρη ελευθερία στην επιλογή φόρμας πακέτων, διευθύνσεων, τεχνικών μετάδοσης κτλ. καθώς και τη δυνατότητα εύκολης τροποποίησης που οδηγεί σε νέες βερσιόν, δίνοντας και τη χαρακτηριστική ευελιξία που διακρίνει το Διαδίκτυο.



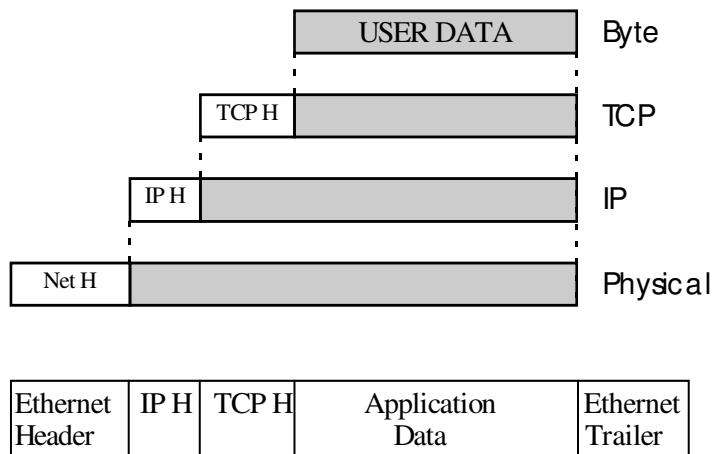
Σχήμα 1.1 Τα στρώματα στο Διαδίκτυο

Καθώς το Διαδίκτυο δεν μπορεί να νοηθεί χωρίς τα υποκείμενα φυσικά δίκτυα, είναι κεντρικής σημασίας ο μηχανισμός ενθυλάκωσης μέσα στις φόρμες των μονάδων των φυσικών δικτύων. Στο σχήμα 1.2 φαίνονται αφ' ενός οι σχέσεις των μονάδων πρωτοκόλλων του Διαδικτύου μεταξύ τους και ο τρόπος ενθυλάκωσης εκάστης σε αυτήν του υποκείμενου στρώματος καθώς και την τελική ενθυλάκωση στο φυσικό δίκτυο. Σαν παράδειγμα δίνεται η τελική μονάδα που προκύπτει με την ενθυλάκωση σε PDU του Ethernet (μια πολύ συνηθισμένη περίπτωση).

Προτού κλείσουμε αυτή την εισαγωγή αξίζει να δώσουμε μερικές παρατηρήσεις αναφορικά με την ορολογία. Ο όρος Διαδίκτυο σήμερα περιλαμβάνει όλα τα στοιχεία των δικτύων του κόσμου που είναι διασυνδεδεμένα μέσω του ομώνυμου πρωτοκόλλου (Internet Protocol). Αυτό το κολοσσιαίο παγκόσμιο υπερ-σύστημα (ίσως το πιο πολύπλοκο του πλανήτη) διασυνδέοντας τη συντριπτική πλειοψηφία των υπολογιστών του κόσμου μας καθιστά δυνατή την προσπέλαση σε πληροφορίες που τηρούνται σε οποιοδήποτε σημείο του πλανήτη από οποιοδήποτε άλλο σημείο. Επεκράτησε λοιπόν κατά συνεκδοχή, με τον ίδιο όρο να αναφερόμαστε και στους πληροφοριακούς πόρους που είναι ενταμιευμένοι σε αυτό. Έτσι ακούγεται πολλές φορές η φράση “Τις πληροφορίες αυτές τις βρήκα στο Internet”.

Για τους κόμβους ο όρος gateway αποδόθηκε με τον όρο πύλη διόδευσης ενώ ο όρος router με τον όρο δρομολογητής. (η διαφορά των δύο είναι δυσδιάκριτη, στην άκρη ενός διοικητικά ανεξάρτητου υποδικτύου χρησιμοποιείται όρος πύλη ενώ στη σε ένα δίκτυο κορμού οι κόμβοι αποκαλούνται δρομολογητές ωστόσο η λειτουργικότητα είναι ουσιαστικά ίδια). Ο ξενιστής αποκαλείται πολλές φορές αδοκίμωσ και τερματικό (απομεινάρι της προ του Διαδικτύου εποχής που η προσπέλαση των

υπολογιστών γίνονταν από απλή οθόνη / πληκτρολόγιο που χειριζότανε μόνο χαρακτήρες ASCII χωρίς υπολογιστική ικανότητα. Φυσικά ο ξενιστής σήμερα είναι συνήθως προσωπικός Η/Υ που εκτελεί και άλλες τοπικές υπολογιστικές εργασίες πλην της πρόσβασης στο Διαδίκτυο. Ωστόσο είναι τόσο διαδεδομένη η πρόσβαση που από τα WINDOWS 95 και μετά το TCP/IP έχει ενσωματωθεί στο λειτουργικό των PC. Ακόμη πολλές φορές χρησιμοποιείται για τον ξενιστή και ο όρος σταθμός καθώς σε εταιρικό ή πανεπιστημιακό περιβάλλον αποτελεί ταυτόχρονα σταθμό του τοπικού δικτύου της επιχείρησης. Στην Ελληνική βιβλιογραφία ο όρος ενίοτε αποδίδεται και σαν υπολογιστής υπηρεσίας, διότι είναι ένας υπολογιστής μέσω του οποίου δεχόμαστε τις υπηρεσίες του δικτύου, ήτοι φυλλομέτρηση ιστοχώρων (browsing), υπηρεσίες φωνής (π.χ. skype), βίντεο κλπ.



Σχήμα 1.2. Ενθυλάκωση μονάδων πρωτοκόλλων στο Διαδίκτυο (και παράδειγμα με Ethernet)

Με την μεγάλη επιτυχία των πρωτοκόλλων TCP/IP άρχισε να γίνεται χρήση τους και για την προσπέλαση πληροφοριών και σε εσωτερικά υποδίκτυα μιας επιχείρησης ή ενός οργανισμού, οπότε εισήχθη για αυτή την περίπτωση ο όρος Intranet. Όταν από την άλλη μεριά μέσω λογικών λειτουργιών (π.χ. κρυπτογράφησης ή αποκλεισμού διευθύνσεων σε κλάδους των δρομολογητών) δημιουργούμε ένα υποσύνολο του Διαδικτύου για εταιρική χρήση αποκλείοντας προσπέλαση σε μη εγκεκριμένους χρήστες (εικονικά δίκτυα) παρότι χρησιμοποιούμε την υποδομή του δημοσίου Διαδικτύου και πέραν του εταιρικού υποδικτύου, τότε χρησιμοποιείται ο όρος Extranet.

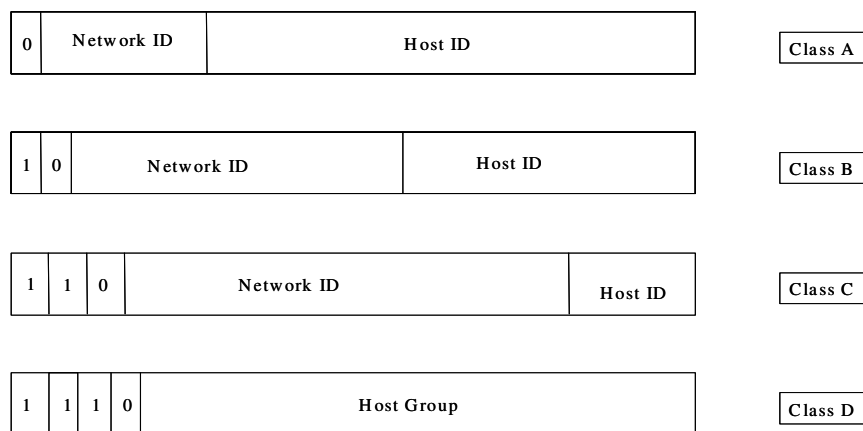
Και μία σύντομη ιστορική αναδρομή στην εξέλιξη του Διαδικτύου:

- 1969: Ξεκινά το ARPANET με 4 κόμβους
- 1973: Οι Vint Cerf και Bob Kahn δημοσιεύουν το TCP/IP Πρώτη εκτός ΗΠΑ σύνδεση (UCL).
- 1974 Πρώτες υλοποιήσεις του TCP/IP
- 1975: Προδιαγραφή ηλεκτρονικού ταχυδρομείου (RFC 733)
- 1983: Το TCP/IP γίνεται το αποκλειστικό στάνταρντ στο ARPANET. Πρώτη σύνδεση Ελλάδα στον κόμβο του Αμστερνταμ.
- 1984: Οι ξενιστές ξεπερνούν τους 1000. Πρώτη εισαγωγή του DNS (Domain Name Server).
- 1985: Δημιουργείται το NSFNET (National Science Foundation) με ταχύτητα κορμού 56kbps
- 1988: Ο κορμός του NSFNET φθάνει τα 1.544Mbps
- 1990: Ο κορμός του NSFNET εκτοπίζει το ARPANET. Μπαίνει και η Ελλάδα στο NSFNET.
- 1997: Το Internet φθάνει τα 16 εκατομμύρια ξενιστές και 60 εκατ. χρήστες! (Δεν πρέπει να λησμονάται ωστόσο ότι το τηλεφωνικό δίκτυο, ίσως το πιο πολύπλοκο ανθρώπινο κατασκεύασμα, έχει την ίδια εποχή 740 εκατ. συνδρομητές!)
- 2002. Οι χρήστες μετριούνται πλέον σε πολλές εκατοντάδες εκατομμύρια και είναι ζήτημα λίγων χρόνων η κατάκτηση του δισεκατομμυρίου!

1.2 Οι διευθύνσεις του Διαδικτύου (IP addresses)

Στην επιλογή διευθύνσεων ενός δικτύου επιδιώκονται πολλαπλές συχνά αντιφατικές σκοπιμότητες: Πρέπει να είναι συμπαγείς και μικρού μήκους για να μην καταλαμβάνουν πολύ χώρο στο πακέτο. Στο στόχο αυτό αντιτίθεται η ανάγκη να ευκολύνεται τη δρομολόγηση με την ομαδοποίηση όσων υπολογιστών βρίσκονται γεωγραφικά μαζί πράγμα που σημαίνει ότι οι προβλέψεις για μελλοντικές διευθύνσεις θα παραμένουν σαν κενά αυξάνοντας το συνολικό μέγεθος. Επίσης, επειδή πρέπει να χρησιμοποιηθούν και από ανθρώπους είναι καλό να είναι εύκολες στη προφορά και την απομνημόνευση πράγμα που απαιτεί πλεονασματικότητα και άρα σπατάλη.

Οι διευθύνσεις του Διαδικτύου (ή διευθύνσεις IP όπως αποκαλούνται εναλλακτικά), αποτελούνται από 32 δυαδικά ψηφία οργανωμένα σε δύο πεδία υιοθετώντας μια ιεραρχία δύο επιπέδων, ένα με τις διευθύνσεις δικτύων και ένα με διευθύνσεις υπολογιστών μέσα στο κάθε δίκτυο για την υποβοήθηση της δρομολόγησης. Έτσι η διεύθυνση χωρίζεται σε δύο πεδία τα οποία μαζί προσδιορίζουν μοναδικά στον κόσμο κάθε ξενιστή του Διαδικτύου. Τα αριστερά (πιο σημαντικά) ψηφία της διεύθυνσης όλων των ξενιστών που ανήκουν στο ίδιο δίκτυο είναι κοινά και προσδιορίζουν την ταυτότητα (Network Identity) του υποδικτύου στο οποίο είναι προσκολλημένος ο ξενιστής υπολογιστής και το υπόλοιπο μέρος δίνει την ταυτότητα του ξενιστή (host Identity) σε διάκριση από τους άλλους ξενιστές του δικτύου.



Σχήμα 1.3. Οι τύποι διευθύνσεων του Διαδικτύου

Για να αποφεύγεται η σπατάλη διευθύνσεων δεν ορίστηκε σταθερό μήκος για τα δύο πεδία Net Id και host Id, αλλά τρία μεγέθη που αντίστοιχα καθορίζουν τρεις κλάσεις διευθύνσεων (κλάσεις A, B or C) με φόρμες όπως φαίνεται στο σχήμα, ενώ υπάρχει και μία φόρμα (κλάση D) για πολλαπλή διεύθυνση (multicast) και μία τελευταία για μελλοντικές χρήσεις. Τα πρώτα ψηφία υψηλής τάξεως προσδιορίζουν την κλάση της διεύθυνσης. Η πρώτη κλάση με μέγεθος πεδίου net id ενός byte καλύπτει λίγα πολύ μεγάλα δίκτυα με πολλά εκατομμύρια υπολογιστές το καθένα π.χ. στο ARPANET οι διευθύνσεις αρχίζουν με 00001010. Η δεύτερη, δίκτυα που μπορεί να έχουν μέχρι 65536 ξενιστές, ενώ η τρίτη μικρά δίκτυα π.χ. τοπικά. Πρέπει να σημειώσουμε ότι η σύμβαση του Διαδικτύου είναι ότι πάντα προηγείται το πιο σημαντικό οκτέτο και αυτό ισχύει και για τα πεδία των διευθύνσεων όπως και για όλα τα πακέτα του Διαδικτύου. Έτσι οι υπολογιστές που εσωτερικά χρησιμοποιούν άλλη σύμβαση πρέπει προτού αποστείλουν τα πακέτα να τοποθετήσουν τα οκτέτα με το σημαντικής τάξεως μπροστά (big-endian).

Σημαντικό στην δομή της διεύθυνσης IP είναι ότι το λογισμικό μπορεί να βρει γρήγορα την διεύθυνση δικτύου και την διεύθυνση του ξενιστή αλλά και να βοηθηθεί η δρομολόγηση λόγω της ιεραρχικής δομής. Συγκεκριμένα οι ενδιάμεσες πύλες διόδευσης (gateways) εξετάζουν μόνο την ταυτότητα υποδικτύου (Net id) και προωθούν τα πακέτα προς το δίκτυο προορισμού όπου όταν φθάσουν θα εξετασθεί η ταυτότητα (Host id) του ξενιστή υπολογιστή για την ανεύρεση του. Έτσι οι πίνακες δρομολόγησης είναι πολύ πιο μικροί απ' ό,τι αν δεν είχαμε ιεραρχικές διευθύνσεις. Λόγω της σχεδόν εξάντλησης των διευθύνσεων από την εκρηκτική αύξηση ξενιστών, η διάκριση σε κλάσεις εγκαταλείφθηκε αργότερα. Αυτή την εξέλιξη θα την παρουσιάσουμε σε επόμενο κεφάλαιο.

Αξίζει να παρατηρήσουμε ότι στην πραγματικότητα οι διευθύνσεις προσδιορίζουν την θύρα επικοινωνίας σε ένα υπολογιστή και όχι την ίδια την μηχανή. Αυτό γίνεται καλύτερα αντιληπτό στην

περίπτωση των πυλών διόδευσης που χαρακτηρίζονται από πολλές θύρες αφού διασυνδέουν πολλά δίκτυα και έτσι πρέπει να έχουν τόσες διευθύνσεις IP όσα και τα υποδίκτυα που διασυνδέουν.

Οι διευθύνσεις δικτύου μπορεί να αποδοθούν και σε υποδίκτυα γενικώς. Στην περίπτωση αυτή το τμήμα Host id της διεύθυνσης είναι 0, η διεύθυνση π.χ. του ARPANET είναι 10.0.0.0. Αντίθετα, η διεύθυνση που σχηματίζεται όταν όλα τα bit της διεύθυνσης ξενίου είναι 1 χρησιμοποιείται για κοινοποίηση (broadcasting) στο υποδίκτυο. Φυσικά η δυνατότητα κοινοποίησης δεν εξασφαλίζεται απλά με την ύπαρξη της κατάλληλης διεύθυνσης, πρέπει και να την υποστηρίξει το υποδίκτυο. Αυτό μπορεί να είναι προφανές σε ένα υποδίκτυο Ethernet αλλά να απαιτεί ειδικό λογισμικό σε δίκτυα άλλης φυσικής τοπολογίας.

Η διεύθυνση [127.0.0.0] είναι μία ειδική διεύθυνση ανακλάσεως (loopback) που χρησιμοποιείται για ελέγχους μεταξύ του προγράμματος εφαρμογής και του προγράμματος επικοινωνίας μιάς μηχανής. Όταν το δεύτερο λάβει ένα πακέτο με αυτή την διεύθυνση, το επιστρέφει χωρίς να στείλει τίποτα στο δίκτυο. Δηλαδή δεν θα δούμε ποτέ ένα τέτοιο πακέτο να κυκλοφορεί στο δίκτυο παρά μόνο εσωτερικά μέσα στις μηχανές.

Για ευκολία απομνημόνευσης των δυαδικών διευθύνσεων και χειρισμό τους από διαχειριστές δικτύων, έχει επικρατήσει ένας εναλλακτικός συμβολισμός των διευθύνσεων ο εστιγμένος δεκαδικός κατά τον οποίο κάθε ένα από τα τέσσερα οκτέτα της διεύθυνσης δίδεται με τη δεκαδική τιμή του και χρησιμοποιούνται στιγμές για να τα χωρίζουν, εξ ου και το όνομα εστιγμένος δεκαδικός συμβολισμός (dotted decimal notation). Π.χ. η διεύθυνση 10010011 01100110 00000001 00000011 που ανήκει στο Ίκαρο (υπολογιστή του ΕΜΠ) γράφεται όταν ενημερώνουμε το λογισμικό κατά την εγκατάσταση της στοιβας TCP/IP και ως: 147.102.1.3. Επειδή οι δυαδικές διευθύνσεις δεν είναι εύκολο να τις χειρίζονται οι χρήστες, καθιερώθηκε και η παράλληλη χρήση ανθρωπομορφικών διευθύνσεων καθώς και μέθοδος αντιστοίχισης των δύο μέσω πινάκων που τηρούνται σε εξυπηρετητές ονομάτων ανά επικράτεια (DNS, Domain Name Server). Το σύστημα DNS εξετάζεται σε επόμενο κεφάλαιο. Για παράδειγμα η αντίστοιχη ανθρωπομορφική διεύθυνσή του Ίκαρου είναι: "icaros.softlab.ntua.gr". Διακρίνεται η ιεραρχική δομή και η αυξανόμενη εστίαση προς τον πλήρη προσδιορισμό από τα δεξιά προς τα αριστερά. Οι επικράτειες διαχωρίζονται μέσω των στιγμών.

Τις διευθύνσεις του Διαδικτύου διαχειρίζεται μία κεντρική υπηρεσία η ICANN (Internet Corporation for Assigned Names and Numbers) η οποία ορίζει τον αριθμό ταυτότητας υποδικτύου, ενώ τους αριθμούς ταυτότητας των ξενιστών υπολογιστών του υποδικτύου ορίζει ο διαχειριστής του κάθε υποδικτύου. Η ίδια υπηρεσία διαχειρίζεται και τους κεντρικούς εξυπηρετητές του DNS.

Ένα σημαντικό πρόβλημα με διευθύνσεις που ανήκουν σε θύρες (δηλ. σε πρίζες επικοινωνίες) όπως οι τηλεφωνικές ή και οι διευθύνσεις του Διαδικτύου που εξετάζουμε, είναι ότι πρέπει να αλλάζουν όταν μετακινούμε τον ξενιστή υπολογιστή σε άλλο δίκτυο. Για να λυθεί το πρόβλημα αυτό και να μπορεί κάποιος να συνδέεται στον πάροχό του και από ξένα δίκτυα δηλ. κάτι σαν την περιαγωγή (roaming) της κινητής τηλεφωνίας έχουν προταθεί λύσεις για κινητές διευθύνσεις IP (mobile IP) αλλά δεν θα εξετασθούν εδώ διότι δεν έχουν ακόμα παγιωθεί σαν λύσεις.

Το πλήθος των διευθύνσεων έχει αρχίσει να εξαντλείται γι' αυτό με το RFC 1519 οι ανωτέρω 4 κλάσεις διευθύνσεων καταργήθηκαν και αντ' αυτών εισήχθη η αταξική δρομολόγηση (CIDR-Classless InterDomain Routing) η οποία εξετάζεται σε επόμενο κεφάλαιο. Επίσης στην νέα έκδοση του IP που ονομάζεται IPv6, το πεδίο διευθύνσεων έχει αυξηθεί σε 128 bits. Το IPv6 παρουσιάζεται συνοπτικά σε επόμενο κεφάλαιο.

1.3 Το πρόβλημα συσχέτισης διευθύνσεων δικτύου με τις φυσικές διευθύνσεις και Το πρωτόκολλο Επίλυσης Διευθύνσεων ARP (Address Resolution Protocol)

Για να μπορούν δύο μηχανές να επικοινωνήσουν πρέπει να ανήκουν σε ενιαίο φυσικό δίκτυο (δηλ να υπάρχει φυσική διόδευση μεταξύ τους) και να είναι γνωστές οι μεταξύ τους διευθύνσεις υλικού (hardware) δηλ. οι διευθύνσεις φυσικού δικτύου. Όμως θα θέλαμε για λόγους ευχρηστίας να μην χρειάζεται να ξέρουμε τις φυσικές διευθύνσεις αλλά να μπορούμε να κάνουμε την επικοινωνία με μόνη την γνώση των διευθύνσεων διαδικτύου (δηλ διευθύνσεων IP).

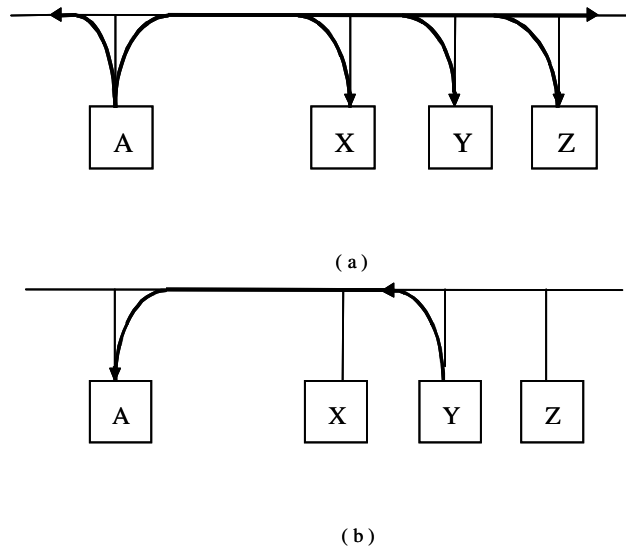
Οι διευθύνσεις φυσικού δικτύου εξαρτώνται από το υλικό (hardware) π.χ. οι διευθύνσεις των καρτών δικτύου Ethernet έχουν εγγραφεί μονίμως από τον κατασκευαστή της κάρτας και είναι παγκοσμίως μοναδικές διότι ένα μόνο γραφείο εκχωρεί ομάδες διευθύνσεων στους κατασκευαστές. Παρ'ότι κάθε κάρτα έχει μία μοναδική διεύθυνση, αυτή δεν έχει καμμία σχέση με την τοποθεσία που θα βρεθεί η μηχανή στην οποία έχει ενσωματωθεί η κάρτα. Αυτή η φυσική διεύθυνση (ή διεύθυνση MAC στην

περίπτωση των τοπικών δικτύων) τοποθετείται αυτομάτως σε κάθε πλαίσιο Ethernet που στέλνει η κάρτα. Από την άλλη μεριά σε κάθε μηχανή που χρησιμοποιεί το πρωτόκολλο TCP/IP εκχωρείται από μία (ή και περισσότερες από μία) διεύθυνση IP των 32-bits, που μπαίνει στο λογισμικό του IP (γιαυτό αναφέρεται και ως διεύθυνση υψηλού επιπέδου με την έννοια του υψηλότερου επιπέδου της στρωμάτωσης) και είναι ανεξάρτητη από την φυσική διεύθυνση της μηχανής και επίσης παγκοσμίως μοναδική.

Κατά την διασύνδεση των φυσικών δικτύων προς το Διαδίκτυο εμφανίζεται το πρόβλημα της αντιστοίχισης διευθύνσεων υψηλού επιπέδου προς φυσικές διευθύνσεις, ή άλλως το πρόβλημα επίλυσης διευθύνσεων (address resolution). Το πρόβλημα έχει αντιμετωπισθεί με διάφορους τρόπους. Κάποια πρωτόκολλα κρατάνε πίνακες, σε κάθε μηχανή, που περιέχουν ζευγάρια από διευθύνσεις υψηλού επιπέδου και φυσικές αλλά το πρόβλημα είναι το μεγάλο μέγεθος των πινάκων σε μεγάλα δίκτυα. Αυτή η μέθοδος ακολουθείται σε δίκτυα X.25 όπου οι πύλες διόδευσης αποθηκεύουν ζευγάρια από IP διευθύνσεις και φυσικές διευθύνσεις X.25, σε ένα πίνακα και ψάχνουν τον πίνακα, όταν επιλύουν μία διεύθυνση IP.

Άλλα συστήματα λύνουν το πρόβλημα ενσωματώνοντας διευθύνσεις hardware (φυσικές) σε υψηλού-επιπέδου διευθύνσεις (host id) πράγμα που είναι εφικτό μόνο όταν οι διευθύνσεις του φυσικού δικτύου μπορούν να επιλέγονται από τον διαχειριστή όπως γινότανε σε παλιά δίκτυα. Ωστόσο σε κάθε περίπτωση δεν είναι καλή ιδέα διότι συνιστά μεγάλη θυσία ευελιξίας.

Η πολύ κοινή περίπτωση του Ethernet παρουσιάζει ιδιαίτερο ενδιαφέρον. Οι φυσικές διευθύνσεις του Ethernet μήκους 48-bit είναι μοναδικές και έχουν ενσωματωθεί κατά οριστικό τρόπο από τον κατασκευαστή της κάρτας δικτύου. Κατά συνέπεια, όταν λόγω βλάβης και απαιτείται αλλαγή της κάρτας, τότε η φυσική διεύθυνση της μηχανής αλλάζει. Η βασική απαίτηση είναι να αποφευχθεί η εμπλοκή των χρηστών, έτσι επινοήθηκε η λύση να κάνει την αντιστοίχιση στις αναπόφευκτες φυσικές διευθύνσεις ένα λογισμικό χαμηλού επιπέδου που κρύβει από τα προγράμματα υψηλού επιπέδου τις φυσικές διευθύνσεις ώστε να μπορούν πλέον να δουλέψουν μόνο με διευθύνσεις Internet. Το πρωτόκολλο που χρησιμοποιείται λέγεται ARP (Address Resolution Protocol - Πρωτόκολλο επίλυσης διευθύνσεων) και έχει δυναμικό χαρακτήρα. Η λύση αυτή επιτρέπει να προστεθούν νέες μηχανές στο δίκτυο, χωρίς επανασύνταξη κώδικα και δεν απαιτεί υποστήριξη από κεντρική βάση δεδομένων και διατήρηση πίνακα αντιστοίχισης διευθύνσεων.



Σχήμα 1.4: Το πρωτόκολλο ARP.

Όπως δείχνει το σχήμα 1.4, η ιδέα, πίσω από τη δυναμική επίλυση, με το ARP είναι απλή: Όταν π.χ. μια πύλη A έχει δεχθεί ένα δεδομένογραμμα με διεύθυνση IP του υπολογιστή Y και πρέπει να το προωθήσει στον Y μέσω του τοπικού δικτύου πρέπει να το ενθυλακώσει σε πλαίσιο Ethernet στο οποίο πρέπει να τοποθετήσει τη φυσική διεύθυνση MAC. Προτού κάνει την ενθυλάκωση κοινοποιεί (broadcasts) ένα ειδικό πακέτο το οποίο ζητά από τον ξενιστή υπολογιστή Y να απαντήσει με τη φυσική του διεύθυνση. Όλοι οι ξενιστές υπολογιστές, συμπεριλαμβανομένου και του Y λαμβάνουν το αίτημα, αλλά μόνο ο ξενιστής υπολογιστής Y αναγνωρίζει την δική του διεύθυνση IP και στέλνει μία

απάντηση που περιέχει τη φυσική διεύθυνση του. Όταν ο Α λαμβάνει την απάντηση, χρησιμοποιεί τη φυσική διεύθυνση για να στείλει το πακέτο κατευθείαν στον Υ.

Η μέθοδος ARP συνδυάζεται με χρήση κρύπτης (cache) όπου φυλάσσονται οι πρόσφατες αντιστοιχίσεις IP - φυσικών διευθύνσεων, ώστε να μη χρειάζεται να χρησιμοποιείται το ARP συνεχώς. Οσάκις ένας ξενιστής υπολογιστής λαμβάνει μία απάντηση ARP, σώζει την διεύθυνση δικτύου της μηχανής και την αντίστοιχη φυσική διεύθυνση του hardware στην κρύπτη του για πιθανές μελλοντικές χρήσεις. Αν ο ξενιστής υπολογιστής βρει την επιθυμητή σύνδεση στην κρύπτη του, δε χρειάζεται εκπομπή προς το δίκτυο. Αυτό δεν συμβαίνει σπάνια διότι συνήθως μία επικοινωνιακή δραστηριότητα δεν εξαντλείται σε ένα πακέτο, έτσι η λύση της κρύπτης (cache) αξίζει τον κόπο.

Μπορεί να αναρωτηθεί κανείς γιατί να μην χρησιμοποιείται η λύση της κοινοποίησης για κάθε πακέτο (broadcasting) αντί να προηγείται η επίλυση μέσω του ARP. Στο κάτω-κάτω αν μπορούμε να βρούμε τον υπολογιστή την πρώτη φορά μπορούμε με την ίδια μέθοδο να τον βρούμε πάντα και έτσι να μην χρειάζομαστε ποτέ την φυσική του διεύθυνση όπως δεν την χρειάστηκε το πακέτο ARP. Ο λόγος είναι ωστόσο ότι η κοινοποίηση είναι πάρα πολύ σπάταλη, για να τύχει γενικευμένης χρήσης, γιατί απαιτεί κάθε μηχανή στο δίκτυο να επεξεργάζεται το πακέτο κοινοποίησης (broadcast) και ο φόρτος από τα κοινοποιούμενα πακέτα θα προκαλούσε συμφόρηση.

Συνοψίζοντας: Το πρωτόκολλο επίλυσης διευθύνσεως "Address Resolution Protocol" (ARP) αντιμετωπίζει το πρόβλημα της εύρεσης της φυσικής διεύθυνσης ενός υπολογιστή - στόχου, στο ίδιο φυσικό δίκτυο, δηλ. αφού το πακέτο χρησιμοποιώντας την διεύθυνση δικτύου (Netid) έχει φθάσει στην πύλη του υποδικτύου προορισμού.

1.4 Το πρωτόκολλο RARP

Συνήθως, η διεύθυνση IP μιας μηχανής, φυλάσσεται στη βοηθητική της μνήμη, όπου το λειτουργικό σύστημα τη βρίσκει στο ξεκίνημα. Η ερώτηση που προκύπτει είναι η εξής: «Πώς μια μηχανή χωρίς δίσκο, κάποια που δεν έχει πρόσβαση σε βοηθητική μνήμη, προσδιορίζει την IP διεύθυνσή της;». Το πρόβλημα είναι κρίσιμο για σταθμούς εργασίας χωρίς δίσκους, οι οποίοι χρησιμοποιούν IP διευθύνσεις για να επικοινωνήσουν με κάποιον εξυπηρετητή (server) αρχείων. Επιπλέον, επειδή πολλές μηχανές χωρίς δίσκο χρησιμοποιούν το τυποποιημένο πρωτόκολλο μεταφοράς αρχείων (FTP) για να αποκτήσουν το λογισμικό εκκίνησης (boot S/W), πρέπει να αποκτήσουν μια διεύθυνση IP προτού τρέξει το λειτουργικό σύστημα. Επειδή θέλουμε το λογισμικό κάθε μηχανής να είναι το ίδιο και να μην περιέχει την συγκεκριμένη IP διεύθυνση της μηχανής ενσωματωμένη μέσα στο λογισμικό εκκίνησης, πρέπει να υπάρχει ένας εξυπηρετητής στο δίκτυο στον οποίο να απευθυνθεί η μηχανή κατά την εκκίνηση για να μάθει την διεύθυνσή της. Για την πρώτη αυτή επικοινωνία η μηχανή χρησιμοποιεί τη φυσική της διεύθυνση σε ένα απλό δίκτυο. Έτσι, η μηχανή μαθαίνει την αντίστοιχη IP διεύθυνσή της και μπορεί πλέον να επικοινωνήσει δια μέσου του Internet.

Η μεθοδολογία είναι απλή: η μηχανή χωρίς δίσκο στέλνει μία έκκληση σε κάποια άλλη μηχανή, τον server, και περιμένει μέχρι ο server να στείλει μια απάντηση. Πώς όμως, η μηχανή χωρίς δίσκο, γνωρίζει τη φυσική διεύθυνση του server; Συνήθως δεν την γνωρίζει - απλά κοινοποιεί (broadcasts) την έκκληση σε όλες τις μηχανές του τοπικού δικτύου. Ένας ή περισσότεροι servers απαντούν. Ο server έχει δίσκο όπου φυλάει μια βάση δεδομένων με τις διευθύνσεις Internet που αντιστοιχούν στις φυσικές διευθύνσεις. Στην έκκληση, η μηχανή η οποία αναζητά την IP διεύθυνσή της, πρέπει να προσδιορίσει μοναδικά την ταυτότητά της, ώστε ο server να μπορεί να κατατάξει τη σωστή Internet διεύθυνση και να ξέρει πού να στείλει την απάντηση. Για το σκοπό αυτό χρησιμοποιείται η διεύθυνση φυσικού δικτύου.

Το πρόβλημα αυτό είναι το ανάστροφο από το πρόβλημα επίλυσης διεύθυνσης, όπου είναι δεδομένη η διεύθυνση φυσικού δικτύου και αναζητείται μια διεύθυνση Internet. Το πρωτόκολλο που σκιαγραφήσαμε ονομάζεται RARP (Reverse Address Resolution Protocol = Ανάστροφο Πρωτόκολλο Επίλυσης Διεύθυνσης).

Όπως ένα μήνυμα ARP, έτσι και το RARP στέλνεται από μια μηχανή σε μια άλλη ενθυλακωμένο εντός του τμήματος δεδομένων ενός πλαισίου Ethernet. Ένα τέτοιο πλαίσιο, το οποίο μεταφέρει μια έκκληση RARP, έχει το συνηθισμένο προοίμιο, διευθύνσεις προέλευσης και προορισμού τύπου Ethernet, και πεδίο τύπου πακέτου στην αρχή του πλαισίου. Ο τύπος πλαισίου περιέχει την τιμή 8035₁₆ που προσδιορίζει ότι τα περιεχόμενα του πλαισίου είναι ένα μήνυμα RARP. Το πεδίο δεδομένων περιέχει τα 28 οκτέτα του μηνύματος RARP.

Κεφάλαιο 2

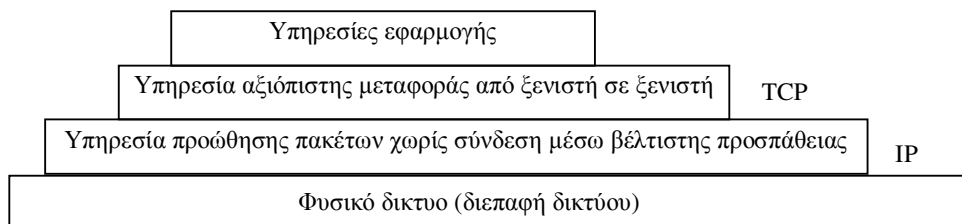
Το πρωτόκολλο Διαδικτύου IP (Internet Protocol)

2.1 Εισαγωγή

Το πρωτόκολλο IP που θα εξετάσουμε σ' αυτό το κεφάλαιο είναι πλέον το πιο διαδεδομένο πρωτόκολλο του στρώματος δικτύου και είναι συνώνυμο με το Διαδίκτυο. Μαζί με το πρωτόκολλο TCP του στρώματος μεταφοράς δίνουν το χαρακτήρα του Διαδικτύου, δεδομένου ότι στα κατώτερα στρώματα το Διαδίκτυο χαρακτηρίζεται από πανσπερμία αφού δεν πρόκειται για δίκτυο με κάποιο ενιαίο χαρακτήρα αλλά για πολλαπλά φυσικά υποδίκτυα διασυνδεδεμένα μέσω πυλών διόδευσης τύπου IP (IP gateways) που το καθένα έχει τα δικά του χαρακτηριστικά. Όσον αφορά τα ανώτερα στρώματα, αυτά ποτέ δεν μπορούν να χαρακτηρίσουν ένα δίκτυο αφού δεν εμφανίζονται παρά μόνο στις τερματικές συσκευές.

Αυτό το τόσο επιτυχημένο πρωτόκολλο χαρακτηρίζεται από το μοντέλο επικοινωνίας χωρίς σύνδεση (ConnectionLess) και τον μη εγγυημένο χαρακτήρα προώθησης που υιοθετεί και που έχει γίνει γνωστός σαν «προώθηση βέλτιστης προσπάθειας» ή «προώθηση κατά δύναμιν» (best effort forwarding). Η βασική υπόθεση εκκίνησης είναι ότι το IP εφαρμόζεται πάνω από διάφορα υποδίκτυα με εξ αντικειμένου μη εγγυημένη πάντα αξιοπιστία και επομένως δεν έχει νόημα να επιχειρηθεί κάποια αξιόπιστη επικοινωνία σε αυτό το επίπεδο αφού η συνολική αξιοπιστία θα χαρακτηρίζεται από αυτήν του λιγότερο αξιόπιστου ενδιαμέσου υποδικτύου. Υιοθετήθηκε λοιπόν μία λύση άνευ συνδέσεων βασισμένη σε δεδομενογράμματα (datagrams) μεγέθους έως 64K οκτέτων τα οποία προωθούνται από δρομολογητή σε δρομολογητή χωρίς παρακολούθηση της πορείας τους αλλά ούτε και της τύχης τους μέσα στο δίκτυο και τα οποία στην πρώτη δυσκολία απορρίπτονται χωρίς πολλές διαδικασίες.

Θα αναρωτηθεί κανείς πώς αφού το Διαδίκτυο μεταφέρει αρχεία δεδομένων μεγέθους εκατοντάδων megabytes χωρίς ούτε ένα λανθασμένο ψηφίο μπορεί να βασισθεί σε ένα αναξιόπιστο πρωτόκολλο δικτύου. Η απάντηση βρίσκεται στο άλλο πρωτόκολλο του αλληλοσυμπληρούμενου σετ δηλ. το TCP το οποίο προβαίνει από τα άκρα της σύνδεσης σε σχολαστική παρακολούθηση της αποστολής ώστε να μη χαθεί ούτε ένα bit χωρία να αναμεταδοθεί.



Σχήμα 2.1 Τα στρώματα στο Διαδίκτυο

Το μοντέλο στρωμάτων του Διαδικτύου φαίνεται στο σχήμα 2.1. Δεν ακολουθεί το (άλλωστε μεταγενέστερο) μοντέλο OSI αλλά συγχωνεύει τις λειτουργίες σε 4 στρώματα. Πάνω από το φυσικό που ενσωματώνει και τις λειτουργίες της ζεύξης υπάρχουν τρία στρώματα. Στο χαμηλότερο επίπεδο είναι η υπηρεσία παράδοσης πακέτων χωρίς σύνδεση μέσω του IP. Στο επόμενο επίπεδο είναι η υπηρεσία αξιόπιστης μετάδοσης και στο υψηλότερο επίπεδο είναι οι υπηρεσίες εφαρμογής.

Επειδή χρησιμοποιούνται πολύ ανομοιογενή φυσικά δίκτυα, έχει γίνει η προσπάθεια να δημιουργηθεί μία αφαίρεση που να δίνει στον χρήστη την εντύπωση ενός ενιαίου δικτύου κρύβοντας την ανομοιομορφία πίσω από την βασική λειτουργία κατά την οποία το δίκτυο παίρνει πακέτα και τα παραδίδει στον προορισμό τους επιτρέποντας στις εφαρμογές να τρέχουν από μακριά. Για τον χρήστη το διαδίκτυο είναι ένα απλό ενιαίο δίκτυο που διασυνδέει όλους τους ξενιστές υπολογιστές (hosts). Στην πραγματικότητα όταν στέλνουμε πακέτα από τον υπολογιστή μας στον οποίο έχουμε φορτώσει την στοίβα TCP/IP, πιθανότατα ξεκινάμε από ένα τοπικό δίκτυο του πανεπιστημίου μας ή της επιχείρησής μας για το οποίο τα πακέτα IP είναι πολύ μεγάλα και αφού σπάσουν σε πακέτα LLC, θα διεκδικήσουν το κοινό μέσο υπακούοντας στο πρωτόκολλο MAC του τοπικού δικτύου, στη συνέχεια περνώντας πιθανόν από κάποια γέφυρα διασυνδεδεμένων τοπικών δικτύων, θα φθάσουν

στην πύλη διόδευσης IP (IP gateway) όπου καταλήγουν όλα τα πακέτα που ανήκουν στο IP και έχουν προορισμό τον έξω κόσμο.

Εκεί θα ξανασυναρμολογηθούν και θα σταλούν μέσω κάποιας ζεύξης που διαθέτει ο οργανισμός του παραδείγματός μας προς κάποιον δρομολογητή IP (IP router) κάποιου άλλου οργανισμού που μας παρέχει υπηρεσίες Διαδικτύου. Από εκεί αρχίζει μία σειρά παρόμοιων δρομολογήσεων από πύλη σε πύλη μέχρι τον τελικό προορισμό που πιθανόν να είναι κάποια βάση δεδομένων όπου θέλουμε να αναζητήσουμε δωρεάν ή επί πληρωμή από αυτόν που προσφέρει αυτήν την υπηρεσία διάφορες πληροφορίες (π.χ. ένας δικηγόρος να αναζητήσει το αρχείο νομολογιών από τον Γερμανικό δικηγορικό σύλλογο αναφορικά με κάποια δικαστική εμπλοκή πελάτη του με Γερμανικά δικαστήρια κατόπιν αυτοκινητιστικού ατυχήματος στην Γερμανία). Για τον χρήστη η διαβίβαση των πακέτων μέσω των τοπικών δικτύων των οποίων η λειτουργικότητα είναι εντελώς διαφορετική από εκείνη της ζεύξης μεταξύ των δρομολογητών είναι διαφανής δηλ. τα ενδιάμεσα υποδίκτυα έχουν αποκρυβεί από τον χρήστη. Ο χρήστης κυρίως βλέπει την πλούσια λειτουργικότητα του λογισμικού υψηλού επιπέδου με το οποίο είναι εφοδιασμένος ο τερματικός υπολογιστής του.

Εν κατακλείδι η βασική φιλοσοφία του πρωτοκόλλου IP χαρακτηρίζεται από στιβαρότητα, αντοχή και προσαρμοστικότητα αλλά λίγες εγγυήσεις.

2.2. Επικοινωνία χωρίς συνδέσεις

Η θεμελιώδης υπηρεσία του πρωτοκόλλου διαδικτύου IP είναι λοιπόν μία υπηρεσία μεταφοράς πακέτων που χαρακτηρίζεται σαν μη αξιόπιστη ελλείψει των κατάλληλων μηχανισμών που θα εξασφάλιζαν εγγυημένη παράδοση σε αυτό το επίπεδο. Αυτό πρακτικά σημαίνει ότι το πακέτο μπορεί να καθυστερήσει ή να παραδοθεί εκτός της σειράς του, αλλά και να απορριφθεί, να διπλασιαστεί, ή να βρεθεί σε λάθος προορισμό. Η υπηρεσία IP ούτε ανιχνεύει τέτοιες καταστάσεις αλλά πολλάκις ούτε ειδοποιεί τον αποστολέα ή τον λήπτη. Η πιο συνήθης αστοχία (και μάλιστα με μεγάλη διαφορά) είναι η απώριψη πακέτου λόγω υπερχειλίσης ταμειυτήρα δρομολογητή.

Η υπηρεσία επίσης χαρακτηρίζεται χωρίς σύνδεση (connectionless) γιατί το κάθε πακέτο είναι ανεξάρτητο από τα άλλα. Μία ακολουθία από πακέτα που στέλνονται από ένα υπολογιστή σε άλλο, μπορεί να ακολουθήσουν διαφορετικά δρομολόγια και ορισμένα από αυτά να χαθούν. Η υπηρεσία που παρέχει το IP χαρακτηρίζεται ως της καλύτερης δυνατής προσπάθειας (best-effort delivery) γιατί το λογισμικό του Διαδικτύου καταβάλει προσπάθεια να παραδώσει τα πακέτα στον προορισμό τους χωρίς όμως και να έχει εξασφαλίσει τις προϋποθέσεις εκείνες που θα καθιστούσαν αυτή την προσπάθεια αρκετή για εγγυημένο αποτέλεσμα. (Αυτό δεν συμβαίνει με άλλα πρωτόκολλα δικτύου, και κυρίως αυτά που βασίζονται σε συνδέσεις και μηχανισμούς αποδοχής ώστε να εγγυώνται την επάρκεια των πόρων δικτύου για το ζητούμενο ποιοτικό αποτέλεσμα υπηρεσίας, όπως π.χ. το X.25). Η έκφραση “ειλικρινή” προσπάθεια αναφέρεται στο γεγονός ότι δεν γίνεται απώλεια πακέτων από μη αντικειμενικά αίτια όπως εξάντληση των πόρων του δικτύου (π.χ. υπερχειλίση ταμειυτήρων) ή βλάβη φυσικών υποδικτύων (π.χ. κομμένες ζεύξεις).

Το πρωτόκολλο IP ορίζει τρεις σημαντικές παραμέτρους της επικοινωνίας. Πρώτα το IP πρωτόκολλο ορίζει τη βασική μονάδα της μετάδοσης δεδομένων (πακέτο) που χρησιμοποιείται για οργάνωση των δεδομένων. Δεύτερο το λογισμικό IP εκτελεί την διαδικασία της δρομολόγησης, διαλέγοντας μια πόρτα εξόδου για να στείλει τα δεδομένα. Τρίτον, το IP παρέχει ένα σύνολο κανόνων που υλοποιούν την ιδέα της μη αξιόπιστης παράδοσης πακέτων. Οι κανόνες χαρακτηρίζουν τον τρόπο που οι ξενιστές υπολογιστές και πύλες διόδευσης θα επεξεργάζονται τα πακέτα, πώς και πότε θα παραχθούν τα μηνύματα λάθους και τις συνθήκες κάτω από τις οποίες κάποια πακέτα θα απορριφθούν.

2.3 Η φόρμα του δεδομενογράμματος IP

Η βασική μονάδα μετάδοσης του διαδικτύου καλείται δεδομένογραμμα Διαδικτύου (IP datagram). Όπως ένα τυπικό πλαίσιο κάθε φυσικού δικτύου το δεδομένογραμμα αποτελείται από την επικεφαλίδα και το πεδίο της πληροφορίας. Η επικεφαλίδα του περιέχει τις διευθύνσεις του προορισμού και της πηγής και το πεδίο της πληροφορίας περιέχει τα δεδομένα. Η διαφορά είναι ότι η επικεφαλίδα του δεδομενογράμματος περιέχει διευθύνσεις IP ενώ τα φυσικά πλαίσια περιέχουν φυσικές διευθύνσεις. Δεν πρέπει να λησμονείται ότι το Διαδίκτυο είναι ένα νοητό δίκτυο που έχει υπερτεθεί σε πολλά και ποικίλα φυσικά (δηλ. πραγματικά όπως τα διάφορα LAN, δορυφορικές ζεύξεις, ζεύξεις PCM, SONET κλπ.) δίκτυα.

Η φόρμα του δεδομενογράμματος είναι:

0	4	8	16	24	31
VERS	HLEN	TYPE OF SERVICE	TOTAL LENGTH		
IDENTIFICATION			D F	M F	FRAGMENT OFFSET
TIME TO LIVE	PROTOCOL	HEADER CHECKSUM			
SOURCE IP ADDRESS					
DESTINATION IP ADDRESS					
IP OPTIONS (IF ANY)				PADDING	
DATA					
...					

Σχήμα 2.2

Οι πρώτες 6 γραμμές του σχήματος ανήκουν στην κεφαλίδα. Στην τυπική της μορφή η κεφαλίδα δεν έχει επιλογές (options) και έρμα (padding) και έτσι έχει ένα μέγεθος 20 οκτέτων. Ακολουθούν τα δεδομένα το μήκος των οποίων προκύπτει από το πεδίο συνολικό μήκος (total length) με αφαίρεση του μήκους κεφαλίδας.

Επειδή η επεξεργασία των δεδομενογραμμάτων γίνεται από λογισμικό, τα περιεχόμενα και η φόρμα μπορούν να εξελίσσονται και δεν είναι υποχρεωτικά αυστηρά αμετάβλητα όπως π.χ. στα τοπικά δίκτυα. Για να διευκολύνεται η εξέλιξη, το πεδίο των πρώτων 4-bit (VERS) περιέχει την έκδοση του IP πρωτοκόλλου που χρησιμοποιήθηκε για να δημιουργηθεί το δεδομένογραμμα. Συνηθίζεται να επαληθεύεται ότι ο αποστολέας, ο λήπτης και κάθε δρομολογητής συμφωνούν μεταξύ τους στη μορφή των δεδομενογραμμάτων. Το λογισμικό ελέγχει το πεδίο έκδοσης πριν από την επεξεργασία καθενός για να εξασφαλίσει ότι η μορφή ταιριάζει με την αναμενόμενη μορφή του λογισμικού. Φυσικά νεώτερες εκδόσεις μπορούν να υποστηριχθούν από ξενιστές και δρομολογητές στους οποίους έχει φορτωθεί το νέο λογισμικό. Είναι αυτονόητο ότι, οι υπολογιστές θα απορρίψουν τα δεδομενογράμματα (datagrams) με εκδόσεις πρωτοκόλλων που δεν υποστηρίζουν, προλαμβάνοντας λανθασμένη μετάφραση των περιεχομένων τους. Η τρέχουσα έκδοση του IP είναι η IPv4. Έχει κυκλοφορήσει και η νέα έκδοση IPv6.

Το πεδίο μήκους επικεφαλίδας Header Length (HLEN) αποτελείται από 4bits και προσδιορίζει το μήκος της επικεφαλίδας του δεδομένογραμμα σε λέξεις των 32-bit. Για να μπορεί το λογισμικό να βρísκει και να ερμηνεύει κάθε πεδίο όλα τα πεδία της κεφαλίδας έχουν σταθερό μήκος που είναι πολλαπλάσιο της λέξης (32 bits). Επειδή το μήκος του πεδίου επιλογών (IP OPTIONS) είναι το μόνο μεταβλητό υπάρχει και το αντίστοιχο έρμα (PADDING) που ο ρόλος του είναι να συμπληρώσει το πεδίο με τα bits που λείπουν για να συμπληρωθεί η τελευταία λέξη.

Το πεδίο “συνολικό μήκος” (TOTAL LENGTH) δίνει το μήκος του IP δεδομένογράφματος μετρημένο σε οκτάδες, περιλαμβάνοντας την επικεφαλίδα και το πεδίο πληροφορίας. Τα 16 bits αυτού του πεδίου καθορίζουν και το μέγιστο πιθανό μέγεθος ενός IP δεδομένογράφματος ήτοι 2^{16} ή 65535 οκτάδες και αποτελεί καλό συμβιβασμό μεταξύ αφ'ενός της μονοπώλησης των ζεύξεων και χρήση μεγάλων ταμιευτήρων όπως συμβαίνει με μεγάλα πακέτα, και αφ'ετέρου της σπατάλης μεγάλου ποσοστού πληροφορίας σε κεφαλίδες όπως συμβαίνει με μικρά πακέτα.

Το πεδίο τύπος υπηρεσίας (TYPE OF SERVICE-TOS) προέβλεπε 3 bits για προτεραιότητες (PRECEDENCE) και άλλα 3 bits D,T και R που φιλοδοξούσαν να καθορίζουν κάποια ιδιαίτερα χαρακτηριστικά ποιότητας υπηρεσίας κατά την μετάδοση (Με το D bit ζητείται χαμηλή καθυστέρηση, με το T bit υψηλή παροχέτευση (throughput) ενώ το R bit υψηλή αξιοπιστία. Στην πράξη δεν χρησιμοποιήθηκαν και περιέπεσαν σε αχρηστία.

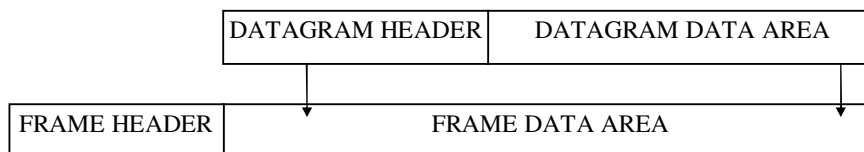
Ωστόσο στα τέλη της δεκαετίας του '90 με το RFC 2474 και πολλά άλλα σχετικά, έγινε αλλαγή της ονομασίας του πεδίου αυτού σε DS (Differentiated Services-Διαφοροποιημένες Υπηρεσίες). Και πάλι η φιλοδοξία ήταν να ορισθούν μηχανισμοί προώθησης με προτεραιότητες ανάλογα με τις ανάγκες της κάθε ροής, δηλ. οι δρομολογητές να προωθούν διαφορετικά (κατά προτεραιότητα) πακέτα που είναι σημειωμένα με κατάλληλους κωδικούς σε αυτό το πεδίο χρησιμοποιώντας τα 6 πρώτα bits αυτού του πεδίου. Η μέχρι τώρα χρήση είναι μάλλον περιορισμένη αλλά πιθανόν στο μέλλον να αυξηθεί η χρήση σε νέες υλοποιήσεις των δρομολογητών διότι οι ανάγκες για υποστήριξη ζωντανής φωνής και βίντεο αυξάνουν. Δεν θα παρουσιασθούν εδώ περισσότερες λεπτομέρειες που μπορούν να αναζητηθούν στο RFC 2474, 2475, 3260, 4594 κτλ.

Το 2001 με το RFC 3168 ορίσθηκαν τα δυο απομένοντα bits για να εισαχθεί ένας νέος τρόπος ελέγχου συμφόρησης με άμεση ειδοποίηση της συμφόρησης (ECN-Explicit Congestion Notification), η οποία απαιτεί συνεργασία του IP και του TCP και η οποία επίσης δεν έχει γνωρίσει εκτεταμένες υλοποιήσεις και παραμένει μια ενδιαφέρουσα πρόταση. Θα εξετασθεί συνοπτικά στο κεφάλαιο μερί ελέγχου συμφόρησης του TCP.

Τα επόμενα πεδία χειρίζονται το μήκος και το ζήτημα του τεμαχισμού των δεδομενογραμμάτων (datagram). Θα εξετασθούν στη συνέχεια από κοινού με το σχετικό ζήτημα της ενθυλάκωσης.

2.4 Ενθυλάκωση δεδομενογράμματος (datagram)

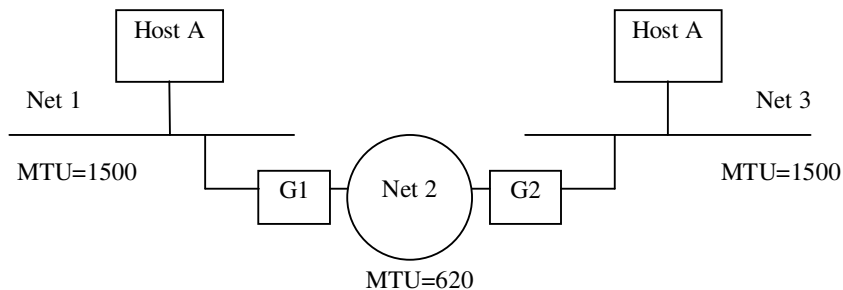
Αφού τα δεδομενογράμματα μεταδίδονται από φυσικά δίκτυα πρέπει να ταξιδεύουν μέσα στα πλαίσια αυτών των δικτύων. Η ιδέα της μεταφοράς ενός δεδομενογράμματος στο πεδίο δεδομένων ενός πλαισίου φυσικού δικτύου καλείται ενθυλάκωση (encapsulation). Έτσι όλα τα πεδία του δεδομενογράμματος μεταφέρονται χωρίς το υλικό (hardware) του φυσικού δικτύου να αναγνωρίζει την μορφή του δεδομενογράμματος ούτε π.χ. να καταλαβαίνει την διεύθυνση του IP αφού βρίσκεται μέσα στο πεδίο δεδομένων του πλαισίου (ιδέ σχήμα) το οποίο δεν προστελάζεται από το φυσικό δίκτυο, δηλ το φυσικό δίκτυο είναι διάφανο (transparent) όσον αφορά την μεταφορά των δεδομενογραμμάτων του Διαδικτύου.



Για να κάνουμε την μετάδοση πιο αποδοτική θα πρέπει κάθε δεδομένογραμμα να αντιστοιχεί ή δυνατόν σε ένα πραγματικό φυσικό πλαίσιο. Μόνο στην ιδανική περίπτωση, που είναι σπάνια σύμπτωση, το ολικό IP δεδομένογραμμα ταιριάζει ακριβώς στο φυσικό πλαίσιο. Αυτό όμως δεν είναι δυνατόν πάντα διότι τα διάφορα φυσικά δίκτυα που διασχίζει το κάθε δεδομένογραμμα έχουν διαφορετικά άνω όρια πλαισίου. Για παράδειγμα το Ethernet περιορίζει την μετάδοση στις 1500 οκτάδες ενώ το proNET επιτρέπει 2044 οκτάδες ανά πλαίσιο. Αυτά τα όρια συνιστούν την μέγιστη μονάδα μετάδοσης (Maximum Transfer Unit ή MTU). Πολλές φορές το μέγεθος της MTU μπορεί να είναι αρκετά μικρό: κάποιες τεχνολογίες υλικού (hardware) περιορίζουν την μετάδοση στις 131 οκτάδες (όπως π.χ. το X.25) ή και λιγότερο. Περιορίζοντας εν γένει τη φόρμα του δεδομενογράμματος ώστε να ταιριάζει στο μικρότερο δυνατό MTU που μπορεί να συναντήσει το δεδομένογραμμα δεν είναι ικανοποιητική λύση διότι η μετάδοση γίνεται λιγότερο αποδοτική στις άλλες περιπτώσεις όταν αυτά μεταδίδονται διαμέσου ενός δικτύου που έχει τη δυνατότητα να μεταδώσει μεγαλύτερα πλαίσια. Η λύση η οποία κρύβει τις υποκείμενες τεχνολογίες δικτύων (διαφάνεια) αλλά κρατά υψηλή αποδοτικότητα είναι να επιλέγεται ένα κατάλληλο αρχικό μέγεθος αλλά να δίδεται η δυνατότητα τεμαχισμού των δεδομενογραμμάτων στο κατάλληλο μικρότερο μέγεθος μόλις βρεθούν μπροστά από ένα φυσικό δίκτυο με περιορισμό μεγέθους. Τα μικρά κομμάτια στα οποία διαιρείται το δεδομένογραμμα λέγονται τεμάχια (fragments) και η διαδικασία λέγεται τεμαχισμός (fragmentation). Τα τεμάχια έχουν την ίδια φόρμα με το αρχικό δεδομένογραμμα.

2.5 Τεμαχισμός και επανασυναρμολόγηση των τεμαχίων

Η ανάγκη τεμαχισμού προκύπτει από το γεγονός ότι θα πρέπει να μεταφέρονται από ποικιλία φυσικών δικτύων που το καθένα υποστηρίζει διαφορετικό μέγιστο μήκος πακέτου. Τίθεται λοιπόν το ερώτημα: πόσο μεγάλο μπορεί να είναι ένα δεδομένογράμμα. Τα δεδομένογράμματα του Διαδικτύου πρέπει να χαρακτηρίζονται από ευελιξία σε αυτό το ζήτημα διότι τα πλαίσια των φυσικών δικτύων έχουν συνήθως ορισθεί από το υλικό (hardware) και δεν μπορούν να προσαρμόζονται στις ανάγκες του Διαδικτύου. Αντιθέτως η διαχείριση του μήκους των δεδομένογραμμάτων του Διαδικτύου γίνεται όπως ξέρουμε από λογισμικό και έτσι είναι εύκολο να έχουν οποιοδήποτε μήκος απαιτείται. Με πεδίο μήκους των 16bit το μήκος των δεδομένογραμμάτων μπορεί να φθάσει τις 65,536 οκτάδες, αλλά σε νεότερες εκδόσεις αυτό το όριο μπορεί να αλλάξει.



Σχήμα 2.3

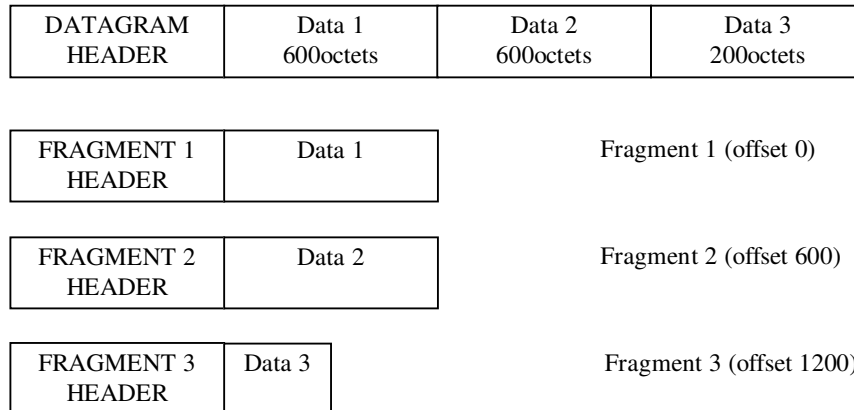
Ο τεμαχισμός συμβαίνει συνήθως στην πύλη διόδου η οποία λαμβάνει το δεδομένογράμμα από το δίκτυο με ένα μεγάλο MTU και πρέπει να το δρομολογήσει διαμέσου ενός δικτύου του οποίου το MTU είναι μικρότερο από το μέγεθος του δεδομένογράφματος. Στο σχήμα 2.3 και οι δύο ξενιστές υπολογιστές που συνδέονται κατευθείαν σε Ethernet έχουν MTU των 1500 οκτάδων. Έτσι και οι δυο σταθμοί μπορούν να παράγουν και να στέλνουν δεδομένογράμματα μέχρι 1500 οκτάδες. Το μονοπάτι μεταξύ αυτών περιλαμβάνει δίκτυο με MTU 620. Αν ο ξενιστής A στείλει ένα δεδομένογράμμα, στο B μεγαλύτερο από 620 οκτάδες η πύλη G1 θα το τεμαχίσει. Ομοίως εάν το B στείλει ένα μεγαλύτερο δεδομένογράμμα τον τεμαχισμό θα τον αναλάβει η G2.

Κατά τον τεμαχισμό πρέπει να ενσωματωθούν στα τεμάχια αρκετές πληροφορίες ώστε να είναι δυνατή η συναρμολόγηση στον προορισμό. Για τον σκοπό αυτό χρησιμοποιούνται τα πεδία της τρίτης τετράδας δηλ. IDENTIFICATION, DF MF ΚΑΙ FRAGMENT OFFSET. Σε όλα τα τεμάχια μπαίνει ο ίδιος αριθμός των 16 bit στο πεδίο IDENTIFICATION ώστε να διακρίνονται ότι ανήκουν στο ίδιο σύνολο σε συνδυασμό με τις ίδιες διευθύνσεις προέλευσης και προορισμού πράγμα που ελαχιστοποιεί τον κίνδυνο σύγχυσης τεμαχίων διαφορετικών πακέτων από σύμπτωση. Για να είναι αντιληπτή η θέση του κάθε τεμαχίου στο αρχικό δεδομένογράμμα, μπαίνει το FRAGMENT OFFSET που είναι ένας αριθμός 8 bit που δείχνει την απόσταση της αρχής του τεμαχίου από την αρχή του αρχικού δεδομένογράφματος σε bytes. Έτσι το πρώτο τεμάχιο έχει OFFSET μηδέν, το δεύτερο ίσο με το μήκος δεδομένων του πρώτου, κτλ. Για να ξέρει ο δέκτης ότι έχει λάβει και το τελευταίο και μπορεί να κάνει τη συναρμολόγηση, όλα τα τεμάχια εκτός του τελευταίου έχουν στη σημαία MF (More Fragments) τιμή 1.

Το μέγεθος του τεμαχισμού επιλέγεται έτσι ώστε κάθε τεμάχιο να χωρά σε ένα πλαίσιο του ενδιαμέσου δικτύου. Επιπλέον επειδή το IP περιγράφει την διάσταση των τεμαχίων με το πεδίο FRAGMENT OFFSET σε πολλαπλάσια των οκτώ οκτάδων, το μέγεθος τεμαχισμού επιλέγεται να είναι το μεγαλύτερο πολλαπλάσιο του 8 που χωρά στο πεδίο δεδομένων του φυσικού πλαισίου. Φυσικά το τελευταίο κομμάτι είναι συνήθως μικρότερο από τα προηγούμενα εκτός αν συμπτωματικά το αρχικό μέγεθος ήταν ακριβές πολλαπλάσιο του νέου. Τα τεμάχια συναρμολογούνται στον προορισμό και παράγουν ένα ολοκληρωμένο αντίγραφο του αρχικού δεδομένογράφματος πριν τον τεμαχισμό. Το πρωτόκολλο IP δεν εγγυάται λοιπόν ότι δεν θα φθάσουν τεμαχισμένα μεγάλα δεδομένογράμματα.

Η προδιαγραφή του IP απαιτεί οι ξενιστές (συναρμολόγηση) και οι πύλες (τεμαχισμός) να μπορούν να χειριστούν δεδομένογράμματα μεγέθους τουλάχιστον μέχρι 576 οκτάδες.

Ένα παράδειγμα τεμαχισμού φαίνεται στο σχήμα. Οι κεφαλίδες και των τριών τεμαχίων είναι ίδιες με του αρχικού δεδομενογράμματος εκτός των bit στο τριψήφιο πεδίο FLAGS που βρίσκεται πριν το FRAGMENT OFFSET. Τα δύο πρώτα τεμάχια έχουν το τρίτο bit MF που ονομάζεται more fragments (κι' άλλα τεμάχια), σε τιμή 1. Το τρίτο έχει αυτό το πεδίο μηδέν αφού είναι το τελευταίο τεμάχιο και δεν ακολουθείται από άλλα τεμάχια.



Σχήμα 2.4

Απαξ και τεμαχισθεί ένα δεδομένογραμμα σε μικρά κομμάτια τα οποία μεταδίδονται ξεχωριστά προκειμένου να περάσουν από ένα υποδίκτυο που δεν υποστηρίζει το μέγεθός του, δεν ξανασυναρμολογείται παρά στον τελικό προορισμό. Αυτό προτιμήθηκε από τις διαδοχικές συναρμολογήσεις και τεμαχισμούς ανάλογα με το μέγεθος της εκάστοτε MTU. Φυσικά υπάρχει το μειονέκτημα ότι μετά τον πρώτο τεμαχισμό εάν συναντήσουμε υποδίκτυο με μεγαλύτερο μέγεθος πλαισίου έχουμε κάποια σπατάλη με το να παραμένουμε στο μικρότερο μέγεθος. Επίσης επιμένοντας στον τεμαχισμό αυξάνουμε κάπως τις πιθανότητες απώλειας του δεδομενογράμματος αφού αρκεί να χαθεί ένα μόνο τεμάχιο για να κηρυχθεί χαμένο όλο το αρχικό δεδομένογραμμα και να αναμεταδοθεί από το TCP. Ο σταθμός προορισμού έχει ένας χρονιστής (timer) και εάν σε ένα καθορισμένο χρονικό διάστημα ο προορισμός δεν λάβει όλα τα τεμάχια τότε η μηχανή συναρμολόγησης που είχε ενεργοποιηθεί στον παραλήπτη απορρίπτει τα συγκεκριμένα δεδομένογραμματα. Αναμεταδίδεται φυσικά ολόκληρο αφού ο αποστολέας δεν ξέρει τίποτα για τον τεμαχισμό ο οποίος φυσικά θα ξαναλάβει χώρα στην πύλη μπροστά από το υποδίκτυο με την μικρότερη MTU όπως και πριν. Ο λόγος της αποφυγής της επανασυναρμολόγησης και στις ενδιάμεσες πύλες (δρομολογητές) παρά αυτά τα δύο μικρά μειονεκτήματα ήταν για να κρατηθούν απλούστερες οι ενδιάμεσες πύλες αποφεύγοντας την αποθήκευση των τεμαχίων μέχρι την άφιξη και των υπολοίπων, τους χρονιστές κτλ. και κρατώντας ανεξάρτητη την δρομολόγηση των τεμαχίων.

Τρία πεδία στην επικεφαλίδα του δεδομενογράμματος τα IDENTIFICATION, FLAGS και FRAGMENT OFFSET ελέγχουν την κατάτμηση και την συναρμολόγηση των δεδομενογραμμάτων. Το πεδίο IDENTIFICATION περιέχει ένα μοναδικό αριθμό που αποτελεί την ταυτότητα του δεδομενογράμματος. Έτσι όταν κόβεται ένα δεδομένογραμμα σε μικρά τεμάχια σε κάθε τεμάχιο υπάρχει το ίδιο πεδίο IDENTIFICATION για να μπορεί ο λήπτης να ξέρει σε ποια μηχανή συναρμολόγησης να οδηγήσει το εν λόγω κομμάτι, δεδομένου ότι τα τεμάχια δεν φθάνουν κατ' ανάγκη με σειρά.

Το πεδίο FRAGMENT OFFSET καθορίζει την θέση του τεμαχίου στο αρχικό δεδομένογραμμα με βάση την απόσταση του πρώτου οκτέτου από την αρχή του δεδομενογράμματος πριν την κατάτμηση μετρημένη σε μονάδες των 8 οκτέτων για οικονομία στο μέγεθος του πεδίου FRAGMENT OFFSET. Η συναρμολόγηση αρχίζει με το τεμάχιο που έχει OFFSET μηδέν και συνεχίζει με τα επόμενα. Το τελευταίο είναι σημειωμένο με την σημαία (FLAG) MF (more fragments).

Το πεδίο FLAGS αποτελείται από 3 bits αλλά το πρώτο δεν χρησιμοποιείται. Τα 2 χαμηλότερης τάξης bits ελέγχουν την κατάτμηση. Το πρώτο bit χρησιμοποιείται για να απαγορεύσει την κατάτμηση για λόγους ελέγχου και ονομάζεται DF δηλ. "do not fragment". Όταν έχει τιμή 1 δηλώνει ότι το δεδομένογραμμα δεν πρέπει να κατατμηθεί. Συνήθως το λογισμικό εφαρμογών που χρησιμοποιεί

TCP/IP δεν ενδιαφέρεται για την κατάσταση και τη συναρμολόγηση που είναι αυτόματες διαδικασίες που συμβαίνουν σε χαμηλό επίπεδο στο λειτουργικό σύστημα και είναι διαφανείς για την εφαρμογή. Ωστόσο ενίοτε είναι σημαντικό να ελεγχθεί εάν έχει επισυμβεί κατάσταση. Π.χ. εάν αποστέλλεται το πρόγραμμα εκκίνησης ενός μικροϋπολογιστή που μόλις ξεκινά από την ROM του και χρειάζεται για συμπλήρωση της εκκίνησης κάποιο κομμάτι κώδικα μέσω του Διαδικτύου, αυτό είναι άχρηστο εάν κατατμηθεί διότι ο υπολογιστής δεν είναι σε θέση να το συναρμολογήσει αφού δεν έχει ακόμη εκκινήσει. Εάν μία πύλη λάβει ένα δεδομένογραμμα μεγέθους μεγαλύτερου από αυτό που μπορεί να διαβιβάσει στο υποδίκτυο με το bit do not fragment ίσον με 1, τότε το απορρίπτει και στέλνει μήνυμα λάθους πίσω στην πηγή

Το τελευταίο bit MF του πεδίου FLAGS δηλώνει αν το τεμάχιο περιλαμβάνει δεδομένα από τη μέση ή από το τέλος του αρχικού δεδομένογραμματος. καλείται "more fragments" bit. Το πεδίο TOTAL LENGTH αναφέρει το μέγεθος του τεμαχίου και όχι του αρχικού δεδομένογραμματος και έτσι ο προορισμός δεν χρησιμοποιεί αυτό το πεδίο για να καταλάβει αν έχει συλλέξει όλα τα τεμάχια. Το more fragments bit λύνει το πρόβλημα. Όταν ο προορισμός λάβει ένα τεμάχιο με το bit MF ίσο με μηδέν, γνωρίζει ότι αυτό το είναι το τελευταίο του αρχικού δεδομένογραμματος. Προσθέτοντας το πεδίο TOTAL LENGTH στο πεδίο FRAGMENT OFFSET ο παραλήπτης βρίσκει ποιο πρέπει να είναι το FRAGMENT OFFSET του επόμενου τεμαχίου και έτσι να βρει τη σειρά που θα πρέπει να μπουν τα τεμάχια και να καταλάβει εάν λείπει κάποιο οπότε πρέπει να περιμένει μέχρι ή να τα έχει όλα ή να λήξει ο χρονιστής οπότε όλα τα τεμάχια απορρίπτονται.

Το πεδίο TIME TO LIVE (χρόνος ζωής) καθορίζει πόσο χρόνο σε δευτερόλεπτα το δεδομένογραμμα επιτρέπεται να παραμείνει στο Διαδίκτυο. Όταν ένας υπολογιστής ορίζει σε αυτό το πεδίο ένα μέγιστο χρόνο ζωής του δεδομένογραμματος, οι πύλες και ξενιστές που επεξεργάζονται το δεδομένογραμμα ελαττώνουν αυτό το χρόνο και εάν φθάσει η τιμή αυτού του πεδίου στο 0, το δεδομένογραμμα απορρίπτεται. Υπάρχει φυσικά εδώ η δυσκολία εκτίμησης του χρόνου διάδοσης σε κάθε φυσική ζεύξη αλλά κυρίως του χρόνου παραμονής στη μνήμη του κάθε δρομολογητή. Στην πράξη αποδείχθηκε ότι αρκούσε α μέτρηση του αριθμού δρομολογήσεων αντί του χρόνου και έτσι απλά κάθε δρομολογητής μειώνει αυτό το πεδίο κατά ένα. Σε περίπτωση που το πεδίο TIME TO LIVE μηδενιστεί, ο δρομολογητής απορρίπτει το δεδομένογραμμα και στέλνει ένα μήνυμα λάθους πίσω στην πηγή μέσω του πρωτοκόλλου ICMP το οποίο εξετάζεται σε επόμενο κεφάλαιο.

Το πεδίο PROTOCOL είναι ανάλογο του πεδίου "τύπος πλαισίου" του δικτύου Ethernet. Η τιμή του καθορίζει ποιο υψηλού επιπέδου πρωτόκολλο χρησιμοποιήθηκε για να δημιουργήσει το μήνυμα που μεταφέρεται στο πεδίο δεδομένων του δεδομένογραμματος. Στην ουσία η τιμή του πεδίου PROTOCOL περιγράφει τη φόρμα των δεδομένων.

Το πεδίο HEADER CHECKSUM διασφαλίζει την ακεραιότητα των τιμών της επικεφαλίδας. Το CHECKSUM του IP προκύπτει από την επεξεργασία της επικεφαλίδας σαν ακολουθία ακεραίων μήκους 16-bits οι οποίοι προστίθενται χρησιμοποιώντας αριθμητική συμπληρώματος του ένα και παίρνοντας το συμπλήρωμα του ένα του προκύπτοντος αθροίσματος. Το ίδιο το πεδίο HEADER CHECKSUM κατά την άθροιση λογαριάζεται με περιεχόμενα 0. Το CHECKSUM καλύπτει μόνο την κεφαλίδα και όχι τα δεδομένα. Τα δεδομένα αφήνονται στη φροντίδα ελέγχου λαθών των πρωτοκόλλων των ανωτέρων στρωμάτων (TCP και εφαρμογής).

Τα πεδία SOURCE IP ADDRESS και DESTINATION IP ADDRESS περιέχουν τις IP διευθύνσεις των 32-bit του αποστολέα και του λήπτη των δεδομένογραμμάτων δηλαδή είναι τα κύρια πεδία βάσει των οποίων γίνεται η δρομολόγηση και η φόρμα τους εξετάστηκε σε προηγούμενο κεφάλαιο.

Το πεδίο IP OPTIONS (επιλογές) δεν υπάρχει πάντα, ωστόσο η δυνατότητα επεξεργασίας των επιλογών πρέπει να υπάρχει πάντα στο λογισμικό που υλοποιεί το IP. Χρησιμοποιείται κυρίως για ελέγχους και εκσφαλμάτωση. Είναι μεταβλητού μήκους ανάλογα των χρησιμοποιούμενων επιλογών.

Κάποιες επιλογές έχουν μήκος μίας οκτάδας, και αποτελούνται από ένα option code. Άλλες επιλογές είναι ποικίλου μήκους. Όταν οι επιλογές παρουσιάζονται στο δεδομένογραμμα, παρουσιάζονται συνεχόμενες χωρίς διαχωριστές ανάμεσα τους. Κάθε επιλογή περιλαμβάνει μία option code απλής οκτάδας, η οποία μπορεί να ακολουθείται από μία οκτάδα με το μήκος και ένα σετ από οκτάδες δεδομένων για την συγκεκριμένη επιλογή. Το option code octet διαιρείται σε τρία πεδία όπως φαίνεται στο σχήμα:

0	1	2	3	4	5	6	7
COPY	OPTION CLASS	OPTION NUMBER					

Τα πεδία περιλαμβάνουν ένα 1-bit COPY flag, 2-bits OPTION CLASS και 5-bits OPTION NUMBER. Το COPY flag ελέγχει το πώς οι πύλες (gateways) χειρίζονται τις επιλογές κατά την κατάρτηση. Όταν είναι 1 τότε καθορίζει ότι η επιλογή θα αντιγραφεί σε όλα τα τεμάχια. Όταν είναι 0 σημαίνει ότι θα αντιγραφεί μόνο στο πρώτο τεμάχιο.

Τα OPTION CLASS και OPTION NUMBER bits καθορίζουν την γενική κλάση της επιλογής και δίνουν μία συγκεκριμένη επιλογή σε αυτή την κλάση. Ο επόμενος πίνακας δείχνει τις κλάσεις:

Κλάση επιλογής (Option Class)	Σημασία
0	Δεδομένογραμμα ή έλεγχος δικτύου
1	Για μελλοντικές χρήσεις
2	Εκσφαλμάτωση (Debugging) και μετρήσεις
3	Για μελλοντικές χρήσεις

Το πεδίο PADDING (έρμα) εξαρτάται από τις επιλογές που έχουν γίνει. Απαρτίζεται από τόσα bits που περιέχουν 0 όσα χρειάζονται για να φέρουν το μήκος της επικεφαλίδας του δεδομένογραμματος σε ακριβές πολλαπλάσιο των 32 bits αφού το μήκος της επικεφαλίδας προδιαγράφεται σε μονάδες λέξεων μήκους 32 bits.

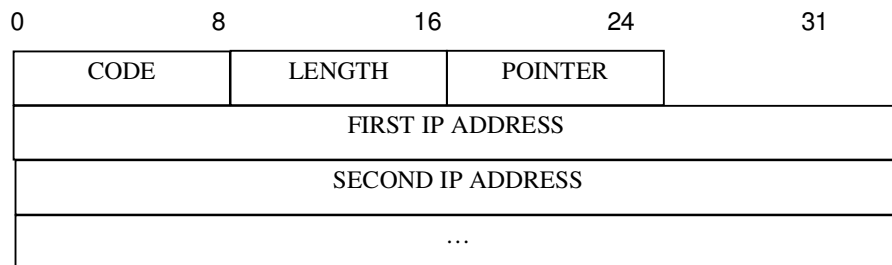
Τέλος το πεδίο DATA δείχνει την αρχή του πεδίου των δεδομένων του δεδομένογραμματος.

Ας δούμε τώρα μερικές σημαντικές επιλογές

Καταγραφή διαδρομής (Record Route Option)

Οι επιλογές της δρομολόγησης και του χρονοσφραγίδας (timestamp) είναι οι περισσότερο ενδιαφέρουσες γιατί παρέχουν ένα τρόπο παρακολούθησης ή ελέγχου της διαδρομής που οι πύλες επιλέγουν για τα δεδομένογραμματα. Η επιλογή record route επιτρέπει στην πηγή να δημιουργήσει μία άδεια λίστα IP διευθύνσεων στην οποία κάθε gateway που διαχειρίζεται το δεδομένογραμμα θα προσθέσει την IP διεύθυνσή του. Στο σχήμα δείχνεται η μορφή του record route.

Το πεδίο CODE περιέχει τον αριθμό και την κλάση επιλογής (7 για το record route). Το πεδίο LENGTH καθορίζει το συνολικό μήκος της επιλογής όπως εμφανίζεται στο IP datagram περιλαμβάνοντας τις τρεις πρώτες οκτάδες. Τα πεδία αρχίζουν με ένα FIRST IP ADDRESS περιέχοντας την δεσμευμένη (reserved) περιοχή για καταγραφή των διευθύνσεων. Το πεδίο POINTER υποδεικνύει την διεύθυνση της πρώτης ελεύθερης θέσης όπου μπορεί να γραφεί η επόμενη διεύθυνση.



Όταν ένας υπολογιστής διαχειρίζεται ένα δεδομένογραμμα που έχει ενεργοποιημένη την επιλογή καταγραφής δρομολόγησης (route option set) τότε θα καταχωρίσει τις διευθύνσεις του στην λίστα των καταγραφών. Ο υπολογιστής πρώτα συγκρίνει τον δείκτη και τα μήκη των πεδίων. Αν ο δείκτης είναι μεγαλύτερος από το μήκος, η λίστα είναι γεμάτη, και έτσι ο υπολογιστής προωθεί το δεδομένογραμμα χωρίς ενημέρωση. Αν η λίστα δεν είναι γεμάτη, τότε εισάγει την διεύθυνσή του μήκους 4-οκτάδων στη θέση που καθορίζεται από το POINTER, και τον αυξάνει κατά τέσσερα.

Επιλογή δρομολόγησης πηγής (Source Route Option)

Δρομολόγηση πηγής είναι μία μέθοδος δρομολόγησης κατά την οποία μπορεί ο αποστολέας να απαγορεύσει ένα μονοπάτι διαμέσου του δικτύου. Η δυνατότητα να γίνονται έλεγχοι, όπως της παροχέτευσης (throughput) ενός συγκεκριμένου μονοπατιού, απαιτεί ο διαχειριστής του δικτύου να μπορεί να δρομολογεί τα δεδομενογράμματα των χρηστών από διαδρομές που μπορεί να μην επελέγοντο από τους δρομολογητές. Δεν είναι χρήσιμο για τον μέσο χρήστη που μπορεί να μη ξέρει την τοπολογία του δικτύου

Επιλογή χρονοσφραγίδας (Timestamp Option)

Το timestamp option λειτουργεί όπως το record route option δηλ. περιέχει μία αρχικά άδεια λίστα στην οποία κάθε πύλη που βρίσκεται κατά μήκος του δρομολογίου από την πηγή στον προορισμό συμπληρώνει μία θέση της λίστας. Κάθε εγγραφή στη λίστα περιέχει δύο μέρη των 32bit: την IP διεύθυνση της πύλης (gateway) που κάνει την εγγραφή και μία χρονοσφραγίδα των 32-bit. Στο σχήμα φαίνεται η μορφή:

0	8	16	24	31
CODE(68)	LENGTH	POINTER	OFLOW	FLAGS
FIRST IP ADDRESS				
FIRST TIMESTAMP				
...				

Στο σχήμα τα πεδία LENGTH και POINTER χρησιμοποιούνται για να καθορίσουν το μήκος του χώρου που κρατείται για την επιλογή και την τοποθεσία των επομένων καταχωρίσεων. Το πεδίο OFLOW (4bit) περιέχει ένα ακέραιο αριθμό που εκφράζει το πλήθος των πυλών που δεν μπορούν να καταχωρίσουν χρονοσφραγίδα γιατί η επιλογή ήταν πολύ μικρή.

Οι χρονοσφραγίδες δίνουν τον χρόνο και την ημερομηνία επεξεργασίας του δεδομενογράμματος από την πύλη διόδευσης (gateway) εκφράζεται σε milliseconds (Universal Time). Βέβαια αν τα timestamps εκδίδονται από διαφορετικούς υπολογιστές μπορεί να διαφέρουν στην τοπική ώρα και γιαυτό πρέπει να θεωρούνται απλά χονδρικές εκτιμήσεις.

Τέλος ας εξηγήσουμε τη χρήση του COPY bit στο πεδίου CODE των επιλογών. Όταν τεμαχίζεται ένα δεδομένογράμμα, η πύλη αντιγράφει μερικές επιλογές σε όλα τα τεμάχια, ενώ άλλες τις γράφει μόνο σ' ένα τεμάχιο. Για παράδειγμα, η επιλογή καταγραφής της δρομολόγησης ενός δεδομενογράμματος δεν μπορεί να περνιέται σε όλα τα τεμάχια διότι αυτά δεν ακολουθούν πάντα τον ίδιο δρόμο και στο τέλος δεν θα μπορεί να συνταχθεί μία συνεπής λίστα δρομολογίου. Έτσι περιοριζόμαστε στην καταγραφή μόνο της διαδρομής του πρώτου τεμαχίου. Αντιθέτως, η επιλογή δρομολόγησης πηγής αντιγράφεται σε όλα τα τεμάχια, ώστε αυτά να εξαναγκασθούν όλα να πάνε από το ίδιο δρομολόγιο.

Κεφάλαιο. 3

ΔΡΟΜΟΛΟΓΗΣΗ ΠΑΚΕΤΩΝ ΣΤΟ INTERNET

3.1 Τι είναι δρομολόγηση

Η δρομολόγηση (routing) δεδομενογραμμμάτων στο Διαδίκτυο είναι η λειτουργία επιλογής του δρομολογίου που θα ακολουθήσει το κάθε πακέτο. Ο όρος δρομολόγηση έχει επικρατήσει για την περιγραφή της λειτουργίας της μεταγωγής των δεδομενογραμμμάτων. Εκτελείται κατά κατανεμημένο τρόπο από κάθε κόμβο και περιλαμβάνει δύο λειτουργίες και δύο αντίστοιχες οικογένειες αλγορίθμων. Η πρώτη είναι η προώθηση (forwarding) κάθε πακέτου που φθάνει σε ένα κόμβο με την επιλογή της θύρας εξόδου και κατά συνέπεια του επόμενου βήματος (δηλαδή της απόφασης σε ποιο γειτονικό δρομολογητή να προωθηθεί το πακέτο). Αν και κάθε δρομολογητής αποφασίζει αυτόνομα στη βάση τοπικά διαθέσιμης πληροφορίας (που φυλάσσεται στους πίνακες δρομολόγησης), αυτό δεν μπορεί να γίνει χωρίς καμία γνώση της τοπολογίας του δικτύου, γνώση η οποία εμμέσως πρέπει να έχει εκ των προτέρων ενσωματωθεί στους πίνακες δρομολόγησης. Η δεύτερη λοιπόν λειτουργία της δρομολόγησης είναι η εύρεση της τοπολογίας (topology discovery) που οδηγεί στην δημιουργία και ενημέρωση των πινάκων στους οποίους βασίζεται η προώθηση. Ο όρος δρομολόγηση (routing) χρησιμοποιείται πολλές φορές μόνο για τη πρώτη ή τη δεύτερη λειτουργία ή και για τις δύο μαζί όταν αναφερόμαστε στη συνολική λειτουργία του δρομολογητή. Θα ασχοληθούμε πρώτα με την προώθηση που αφορά πιο άμεσα τους χρήστες γιατί διαχειρίζεται τα πακέτα με τις πληροφορίες τους ενώ η διαχείριση των πινάκων είναι μια πιο προχωρημένη λειτουργία και θα εξετασθεί πιο κάτω.

Έχουμε δει ότι η βασική υπηρεσία που παρέχεται από το IP συνίσταται σε μία μη αξιόπιστη, άνευ σύνδεσης παράδοση πακέτων “κατά δύναμιν” (best-effort). Η υποστήριξη της επικοινωνίας χωρίς σύνδεση μέσω διασυνδεδεμένων δικτύων όπως γίνεται στο Διαδίκτυο απαιτεί τη δρομολόγηση των πακέτων IP (IP routing) μέσω ειδικών κόμβων που καλούνται δρομολογητές (routers) ή πύλες διόδευσης (gateways) μέχρι την παράδοσή τους στον τελικό τους προορισμό. Η δρομολόγηση είναι η κυριότερη λειτουργία του στρώματος δικτύου και στην προκειμένη περίπτωση του πρωτοκόλλου IP σε βαθμό που η χρησιμοποιούμενη τεχνική να χαρακτηρίζει το όλο δίκτυο σαν δίκτυο μεταγωγής δεδομενογραμμμάτων. Επειδή από την άλλη τα δεδομενογράμματα σε αντίθεση με τα προγενέστερα δίκτυα πρέπει να αποθηκευθούν στη μνήμη μέχρι να αποφασισθεί σε ποια διεπαφή εξόδου θα προωθηθεί, το Διαδίκτυο χαρακτηρίζεται και σαν δίκτυο τεχνικής αποθήκευσης-προώθησης (store-and-forward).

Το IP ακολουθεί σαν βασική φιλοσοφία την δρομολόγηση βήμα-πρός-βήμα (hop-by-hop routing) δηλ. μία κατανεμημένη διαδικασία χωρίς κεντρικό έλεγχο. Ο κάθε κόμβος έχει μόνο μία επιλογή να κάνει ήτοι να διαλέξει την πόρτα εξόδου και άρα την επόμενη ζεύξη μέσα από την οποία θα μεταδοθεί το δεδομένογράμματα (datagram) μέχρι τον επόμενο κόμβο, αφήνοντας σε αυτόν την ευθύνη για την συνέχιση της διαδικασίας μέχρι τον τελικό ξενιστή. Αυτό δεν σημαίνει ωστόσο ότι δεν λαμβάνει χώρα καμία συνεννόηση για τη βελτιστοποίηση της διαδρομής κάτι που όπως θα δούμε γίνεται μέσω ανταλλαγών σχετικών πληροφοριών μεταξύ των κόμβων το προϊόν των οποίων διοχετεύεται στην ενημέρωση των πινάκων δρομολόγησης.

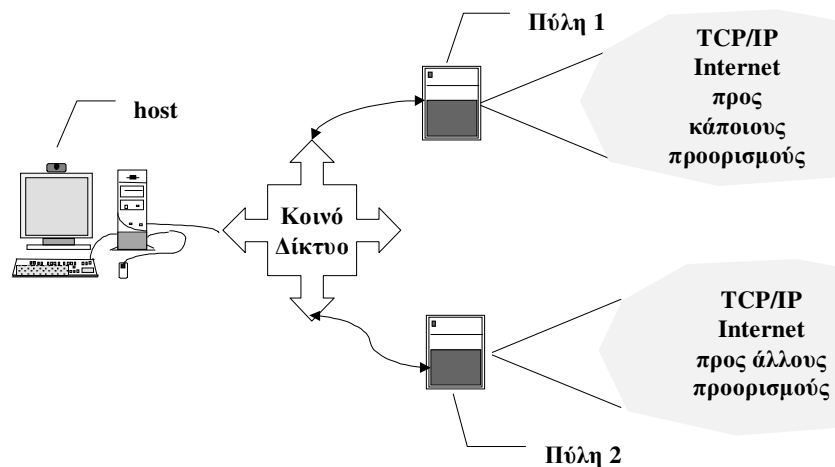
Την λειτουργία της δρομολόγησης μπορούν να εκτελέσουν και οι ξενιστές αλλά με κάποιες διαφορές από τους δρομολογητές. Για τον ξενιστή η δρομολόγηση είναι κατά βάση μια απλή λειτουργία: εάν ο προορισμός συνδέεται κατ' ευθείαν με τον ξενιστή (π.χ. μέσω τοπικού δικτύου, ή ζεύξης T1, ή ζεύξης X.25 κτλ.) τότε το δεδομένογράμματα αποστέλλεται κατευθείαν στην αντίστοιχη διεπαφή (Interface) δικτύου, αλλιώς αποστέλλεται σε μία προεπιλεγμένη πύλη (default gateway) με την οποία ο ξενιστής έχει άμεση φυσική σύνδεση, και η οποία αναλαμβάνει τα περαιτέρω. Η τοποθέτηση μίας τέτοιας πύλης είναι η πρώτη ενέργεια όταν ένα απομονωμένο δίκτυο συνδέεται στο Διαδίκτυο, εάν δεν υπάρχει, τότε δεν υπάρχει και διασύνδεση προς το Διαδίκτυο. Το στρώμα IP ενός ξενιστή δέχεται κατά κανόνα δεδομενογράμματα δημιουργούμενα εσωτερικά από ανώτερα στρώματα (π.χ. TCP, UDP) που είναι υλοποιημένα στο λογισμικό του για να τα στείλει στις διεπαφές δικτύου και αντιστρόφως. Αντίθετα, ο δρομολογητής τα προωθεί κατά κανόνα από διεπαφή εισόδου σε διεπαφή εξόδου (ή σπανίως τα απορρίπτει). Τα δεδομενογράμματα πηγάζουν ή καταλήγουν σε ξενιστές ενώ απλά διέρχονται από τους δρομολογητές. Είναι φυσικά ο κανόνας να είναι ένας δρομολογητής άμεσα συνδεδεμένος σε πολλά φυσικά δίκτυα με αντίστοιχες διεπαφές (τουλάχιστον δύο).

Τα ανωτέρω δεν αποκλείουν και τις λιγότερο συχνές περιπτώσεις όπου ένας δρομολογητής είναι προορισμός δεδομενογραμμμάτων - συνήθως από πρωτόκολλα ελέγχου π.χ. ICMP, EGP, IGP (ιδέ

3.7) - τα οποία πρέπει τότε να διαβιβασθούν σε ειδικές εφαρμογές του δρομολογητή. Αντίστοιχα υπάρχει η δυνατότητα και ένας ξενιστής να διαμορφωθεί σε ρόλο δρομολογητή οπότε και προωθεί δεδομενογράμματα από διεπαφή σε διεπαφή, αλλά στην πιο συνήθη περίπτωση που είναι απλά ξενιστής, τότε όσα δεδομενογράμματα φθάνουν και δεν έχουν διεύθυνση προορισμού την δική του, τα απορρίπτει (και δεν τα επιστρέφει από την κατά κανόνα μία και μοναδική διεπαφή φυσικού δικτύου που διαθέτει). Η αιτιολόγηση αυτής της συμπεριφοράς θα δοθεί πιο κάτω στο 3.4). Αυτό δεν σημαίνει ότι δεν υπάρχουν περιπτώσεις που και ένας απλός ξενιστής πρέπει να επιλέξει προορισμό επομένου βήματος παρότι έχει μία διεπαφή δικτύου όπως στο παράδειγμα του σχήματος 3.1. Αυτή η πρακτική πηγάζει από το γεγονός ότι αρχικά οι διαφορές στο λογισμικό IP μεταξύ ξενιστή και δρομολογητή ήταν μηδαμινές, σήμερα ωστόσο οι δρομολογητές έχουν εξελιχθεί σε ειδικά μηχανήματα όπου πολλές λειτουργίες γίνονται πολύ διαφορετικά και μάλιστα όχι πια από λογισμικό αλλά από ειδικό υλικό (κυκλώματα) και αρχιτεκτονικές που αναμιγνύουν και άλλες τεχνικές μεταγωγής στα χαμηλά επίπεδα (π.χ. MPLS multi-protocol label switching).

Στη γενική περίπτωση η δρομολόγηση στο Διαδίκτυο μπορεί να είναι μια περίπλοκη λειτουργία, ειδικά ανάμεσα σε πύλες με πολλαπλές συνδέσεις φυσικών δικτύων. Ιδανικά το λογισμικό δρομολόγησης θα πρέπει να εξετάζει όταν επιλέγει το καλύτερο μονοπάτι θέματα όπως:

- * το φορτίο του δικτύου,
- * την ταχύτητα των ζεύξεων
- * τις αποστάσεις προς τους γειτονικούς κόμβους
- * τον τύπο υπηρεσίας που καθορίζεται στην επικεφαλίδα.



Σχήμα 3.1. Host με μία μοναδική σύνδεση σε δίκτυο που πρέπει να δρομολογήσει τα IP πακέτα προς μετάδοση είτε προς άμεση παράδοση, είτε σε μία από της δύο πύλες/δρομολογητές ανάλογα με τον προορισμό τους.

3.2 Άμεση και Έμμεση Δρομολόγηση (Πρώθηση)

Το πρώτο πράγμα που πρέπει να κάνει η μηχανή που θέλει να δρομολογήσει ένα δεδομένογράμμο είναι πρώτον να βεβαιωθεί ότι δεν είναι η ίδια προορισμός και ακολούθως να εξετάσει εάν ο προορισμός βρίσκεται στο ίδιο υποκείμενο φυσικό σύστημα μετάδοσης δηλαδή διαθέτει μία διεπαφή προς το φυσικό δίκτυο του υπολογιστή προορισμού (π.χ. βρίσκονται στο ίδιο τοπικό δίκτυο Ethernet, ή στα δύο άκρα μιας ζεύξης T1). Με βάση αυτό το κριτήριο έχουμε δύο εναλλακτικές μεθόδους δρομολόγησης: την άμεση και την έμμεση. Στην περίπτωση που υπάρχει φυσική σύνδεση λαμβάνει χώρα άμεση δρομολόγηση, ενώ στην αντίθετη περίπτωση απαιτείται έμμεση δρομολόγηση οπότε τα πακέτα πρέπει να περάσουν από τουλάχιστον άλλη μία πύλη διόδου (αλλά ίσως και περισσότερες) για πρώθηση προς το τελικό φυσικό δίκτυο προορισμού.

Η μετάδοση ενός δεδομενογράμματος IP ανάμεσα σε δύο μηχανές στο ίδιο φυσικό δίκτυο (άμεση δρομολόγηση) γίνεται ως εξής. Ο αποστολέας ενθυλακώνει το δεδομένογραμμα στο φυσικό πλαίσιο του κοινού υποκείμενου φυσικού δικτύου, συσχετίζει την IP διεύθυνση προορισμού με την αντίστοιχη φυσική διεύθυνση υλικού (hardware), - την οποία προσδιορίζει με μία από τις μεθόδους που εξετάσαμε, π.χ. από πίνακες ή κρύπτη (cache) που διαθέτει, ή μέσω πρωτοκόλλου ARP, - και ακολούθως στέλνει το πλαίσιο κατευθείαν στον προορισμό (εξ ου και ο όρος άμεση δρομολόγηση). Ο προορισμός θα αναγνωρίσει και θα διαβάσει το δεδομένογραμμα από την φυσική διεύθυνση που χαρακτηρίζει την διεπαφή του με το φυσικό δίκτυο. Ο προσδιορισμός του κατά πόσον ο προορισμός βρίσκεται στο ίδιο φυσικό δίκτυο γίνεται με σύγκριση της ταυτότητας δικτύου (net id) του προορισμού προς αυτήν της διεπαφής της πύλης. Αν είναι ίδιες, τότε μπορεί να χρησιμοποιηθεί άμεση δρομολόγηση.

Εάν η διεύθυνση προορισμού έχει διαφορετική ταυτότητα δικτύου, τότε ο αποστολέας πρέπει να προσδιορίσει σε ποια επόμενη πύλη διόδευσης (gateway) πρέπει να προωθήσει το δεδομένογραμμα για να πλησιάσει κατά ένα βήμα το δίκτυο προορισμού (έμμεση δρομολόγηση). Το ζήτημα που εγείρεται σ' αυτό το σημείο είναι πώς μπορεί να ξέρει μία πύλη σε διεπαφή (πόρτα εξόδου) να το προωθήσει ούτως ώστε να το σπρώξει πλησιέστερα προς τον τελικό προορισμό και μάλιστα κατά τον καλύτερο και "οικονομικότερο" τρόπο. Τίθενται δηλαδή - όταν έχουμε πολλές εναλλακτικές διαδρομές, - και ζητήματα βέλτιστης επιλογής αποφυγής κάποιων δρομολογίων που έχουν προσωρινά προβλήματα, αποφυγής αέναης κυκλικής δρομολόγησης κτλ. Τα ζητήματα αυτά θα εξετάσουμε στην συνέχεια.

3.3 Ο ρόλος των πινάκων δρομολόγησης

Κάθε μηχανή στο Διαδίκτυο διαθέτει ένα πίνακα δρομολόγησης Διαδικτύου (ή αλλιώς πίνακα δρομολόγησης IP) όπου αποθηκεύει πληροφορίες σχετικά με πιθανούς προορισμούς. Ο πίνακας κατά βάσιν αντιστοιχίζει ομάδες διευθύνσεων προορισμού σε κάθε μια από τις διεπαφές εξόδου που διαθέτει και οι οποίες αντιστοιχούν στην διεύθυνση του κόμβου του επόμενου βήματος. Οι διευθύνσεις προορισμού μπορεί να είναι διευθύνσεις ξενιστών ή δικτύων (αυτό καθορίζεται από μία σημαία).

Είναι προφανές ότι δεν είναι πρακτικά δυνατόν κάθε πίνακας δρομολόγησης να περιέχει πληροφορία για κάθε δυνατό προορισμό του κόσμου. Αρκεί να σκεφθεί κανείς το ρυθμό αύξησης των υπολογιστών του Διαδικτύου για να αντιληφθεί ότι δεν είναι εφικτή μία κεντρική διαχείριση. Η καταναμενημένη προσέγγιση σε συνδυασμό με την αρχή της απόκρυψης πληροφορίας (information hiding)* είναι η μόνη ρεαλιστική προσέγγιση. Δηλαδή η κάθε μηχανή να μην κάνει παρά ένα μέρος της δουλειάς και να χρησιμοποιεί μόνο την τοπικά αναγκαία πληροφορία. Σ' αυτό βοηθά και η ιεραρχική οργάνωση των διευθύνσεων του Διαδικτύου και είδαμε ήδη πως χρησιμοποιείται αυτό για να προσδιορισθεί εάν μπορεί να χρησιμοποιηθεί άμεση δρομολόγηση. Ομοίως η ιεραρχική δομή των διευθύνσεων επιτρέπει να χρησιμοποιούνται στους πίνακες δρομολόγησης μόνο το πρόθεμα (δηλ το net id) στους πίνακες δρομολόγησης περιορίζοντας δραστικά το μέγεθός τους. Δηλαδή αποφεύγεται η ανάμιξη σε λεπτομέρειες για τους ξενιστές του κάθε υποδικτύου εκτός από την περίπτωση των ξενιστών που ανήκουν στο ίδιο φυσικό δίκτυο με τη μηχανή που εκτελεί τη δρομολόγηση. Έτσι οι πύλες διόδευσης χρειάζεται να τηρούν μόνο καταχωρίσεις στους πίνακες με τα ζεύγη G, N όπου για κάθε πύλη G άμεσα προσπελάσιμη από την παρούσα πύλη (δηλ. που ενώνονται με ζεύξη) δίδονται όλες οι διευθύνσεις δικτύων N που "κρύβονται" πίσω της δηλ. που εξυπηρετούνται μέσω αυτής της πύλης. (Όπου είναι δυνατόν Καταβάλλεται προσπάθεια αυτές οι διευθύνσεις να μπορούν να συνοψισθούν σε μια καταχώριση με ίδια πιο σημαντικά ψηφία). Έτσι κάθε πύλη κατέχει πληροφορία μόνο για το επόμενο βήμα δρομολόγησης αδιαφορώντας για το ποια άλλα βήματα θα χρειασθούν για να φθάσει το δεδομένογραμμα στον προορισμό του.

Το μέγεθος του πίνακα δρομολόγησης εξαρτάται από το πλήθος των διασυνδεδεμένων δικτύων. Υπάρχει επίσης και η επιλογή της *μερικής ενημέρωσης* του πίνακα δρομολόγησης. Σε αυτή την περίπτωση ο κόμβος αναγνωρίζει ένα συγκεκριμένο πλήθος δικτύων και τους επόμενους κόμβους στους οποίους πρέπει να μεταβιβάσει το πακέτο, ενώ για διευθύνσεις που δεν περιέχονται στον πίνακα δρομολογεί το πακέτο σε κάποια άλλη προκαθορισμένη πύλη (δρομολογητή) η οποία θα

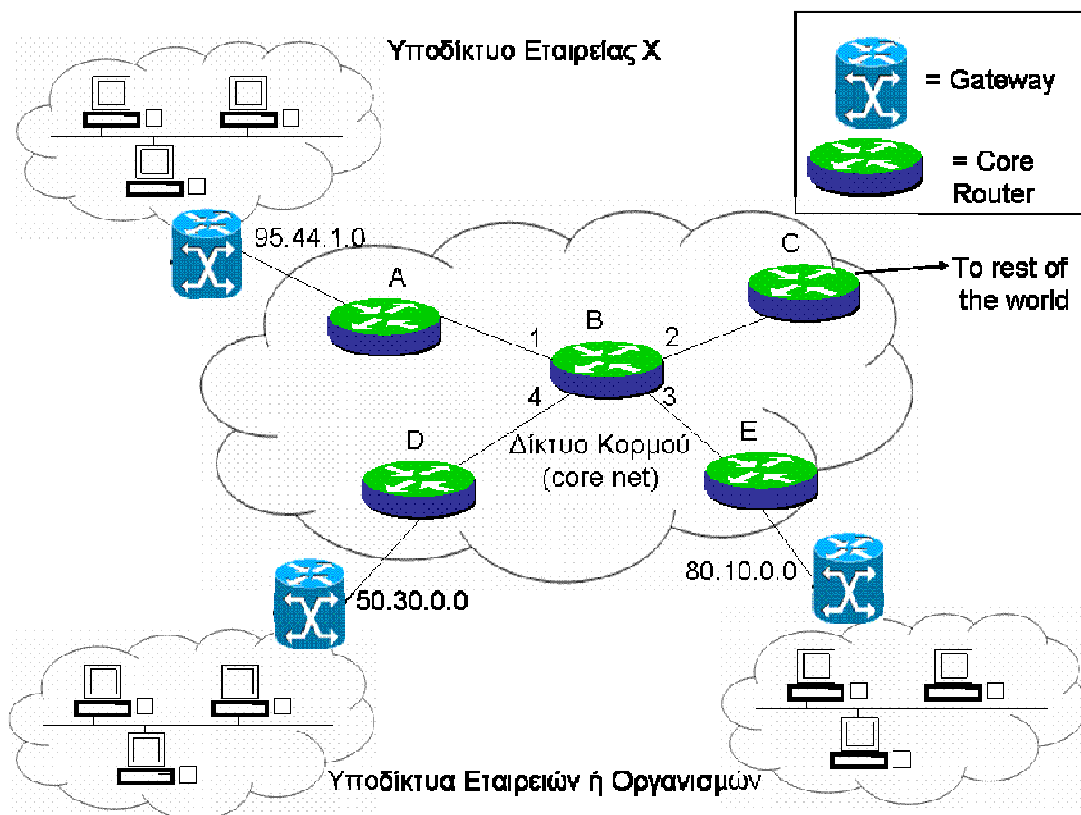
*Η αρχή αυτή ευρίσκει ευρεία εφαρμογή κατά τον προγραμματισμό όπου το κυρίως πρόγραμμα καλεί υπορουτίνες που λύνουν επιμέρους προβλήματα απαλλασσόμενο της ανάγκης να ξέρει πως λύνονται αυτά και αντιστρόφως οι συντάκτες των υπορουτινών δεν χρειάζεται να ασχοληθούν με το τι κάνουν οι άλλες υπορουτίνες ούτε το κυρίως πρόγραμμα.

προσπαθήσει με τον ίδιο τρόπο να δρομολογήσει το πακέτο. Προφανώς η επιλογή της πλήρους ή μερικής ενημέρωσης του πίνακα δρομολόγησης και ο τρόπος ενημέρωσής του εξαρτάται από την αρχιτεκτονική του διασυνδεδεμένου δικτύου και τις προτεραιότητες που τίθενται από τους διαχειριστές του για τη λειτουργία του. Η πολυπλοκότητα του προβλήματος της δρομολόγησης θα φανεί παρακάτω όπου θα εξετάσουμε αναλυτικότερα την επίδραση της αρχιτεκτονικής και την ανάπτυξη του ήδη εγκατεστημένου Internet καθώς και τις τεχνικές που αναπτύχθηκαν για την αντιμετώπιση των επιμέρους προβλημάτων σε κάθε περίπτωση.

Άλλη μία τεχνική για απόκρυψη πληροφορίας και μικρούς πίνακες δρομολόγησης είναι η προαποφασισμένη δρομολόγηση (default route) κατά την οποία αν δεν υπάρχει διαδρομή στον πίνακα, οι ρουτίνες δρομολόγησης στέλνουν το datagram σε μία προαποφασισμένη πύλη (default gateway). Αυτή η τεχνική είναι αρκετά χρήσιμη π.χ. όταν μία εγκατάσταση έχει ένα μικρό αριθμό από τοπικές διευθύνσεις και μόνο μία σύνδεση με τον έξω κόσμο. Κάθε φορά που η διεύθυνση δεν ανήκει σε ένα από τα τοπικά δίκτυα, αποστέλλεται σε αυτή την προαποφασισμένη πύλη όποιος και αν είναι ο προορισμός.

Υπάρχει τέλος και η τεχνική της δρομολόγησης πηγής η οποία είναι μια από τις επιλογές (options) που χρησιμοποιεί το πεδίο επιλογών της επικεφαλίδας του IP. Κατ' αυτήν ο αποστολέας προσδιορίζει πλήρως όλες τις ενδιαμέσες πύλες διόδευσης. Αυτή η δρομολόγηση ωστόσο έχει μόνο ειδική χρήση π.χ. σε διαγνωστικές εφαρμογές από διαχειριστές δικτύων, διότι απαιτεί γνώση της τοπολογίας που δεν είναι προσιτή στον χρήστη.

Στο σχήμα 3.2 περιγράφεται σε μία απλή περίπτωση διασύνδεσης δικτύων και στη συνέχεια δίδεται ο πίνακας δρομολόγησης του κόμβου B. Οι δρομολογητές A, B, C, D, E κάνουν έμμεση δρομολόγηση, ενώ οι πύλες (gateway) κάνουν άμεση δρομολόγηση μέσα στο φυσικό δίκτυο χρησιμοποιώντας τις φυσικές διευθύνσεις που βρίσκουν μέσω του πρωτοκόλλου ARP.



Σχήμα 3.2. Παράδειγμα διασύνδεσης 4 δικτύων μέσω τριών πυλών/δρομολογητών και ο πίνακας δρομολόγησης της πύλης 2 για όλες τις πιθανές διευθύνσεις προορισμού.

Πίνακας δρομολόγησης κόμβου Β	
Πιο σημαντικά ψηφία διεύθυνσης IP του πακέτου	Πόρτα εξόδου πακέτουIP
50.30.X.X	4
80.10.X.X	3
95.44.X.X	1
Λοιπές Διευθύνσεις	2

3.4 Επεξεργασία των πακέτων IP που φθάνουν σ' ένα κόμβο

Πριν προχωρήσουμε στην παρουσίαση των επί μέρους προβλημάτων και στον τρόπο σύνταξης και ενημέρωσης των πινάκων δρομολόγησης ας δούμε πώς ενεργεί ένας κόμβος όταν δέχεται ένα εισερχόμενο IP πακέτο από μία διεπαφή δικτύου. Θα διακρίνουμε δύο περιπτώσεις: ενός απλού ξενιστή και μιας πύλης δρομολόγησης δεδομένου ότι έχουμε διαφορετική συμπεριφορά.

Όπως είδαμε στην ιεράρχηση των διαφόρων επιπέδων ενός τηλεπικοινωνιακού δικτύου και την διαστρωμάτωση των πρωτοκόλλων το επίπεδο δικτύου στο οποίο αντιστοιχεί το πρωτόκολλο IP παραλαμβάνει πακέτα από το υποκείμενο του φυσικό στρώμα μέσω του στρώματος ζεύξης. Στην περίπτωση ενός απλού ξενιστή, εάν η διεύθυνση IP του πακέτου ταυτίζεται με αυτή του ξενιστή, το πακέτο θα απογυμνωθεί από την επικεφαλίδα που προσθέτει το στρώμα IP και το περιεχόμενό του θα προωθηθεί στο ανώτερο στρώμα για επεξεργασία. Φυσικά πρέπει να ελεγχθούν όλες οι πιθανές διευθύνσεις που μπορεί να έχει ένας ξενιστής, ώστε να αποσαφηνιστεί το αν το πακέτο έχει φτάσει τον τελικό του προορισμό δεδομένου ότι το συγκεκριμένο μηχανήμα μπορεί να έχει περισσότερες της μιας IP διευθύνσεις με τις οποίες πρέπει να παραβάλει την διεύθυνση του παραληφθέντος πακέτου στην περίπτωση που βρίσκεται συνδεδεμένο σε περισσότερα του ενός φυσικά δίκτυα,. Ομοίως πρέπει να γίνει έλεγχος για τυχόν ομαδικές διευθύνσεις στις οποίες πιθανόν να συμμετέχει ο ξενιστής (λειτουργία multicast). Εάν η διεύθυνση IP του πακέτου δεν ταυτίζεται με αυτή του ξενιστή, το πακέτο θα πρέπει υποχρεωτικά να απορριφθεί και να μην επαναδρομολογηθεί προς την κατεύθυνση της τελικής διεύθυνσης που περιλαμβάνει το πακέτο (πράγμα που θα έκανε μία πύλη).

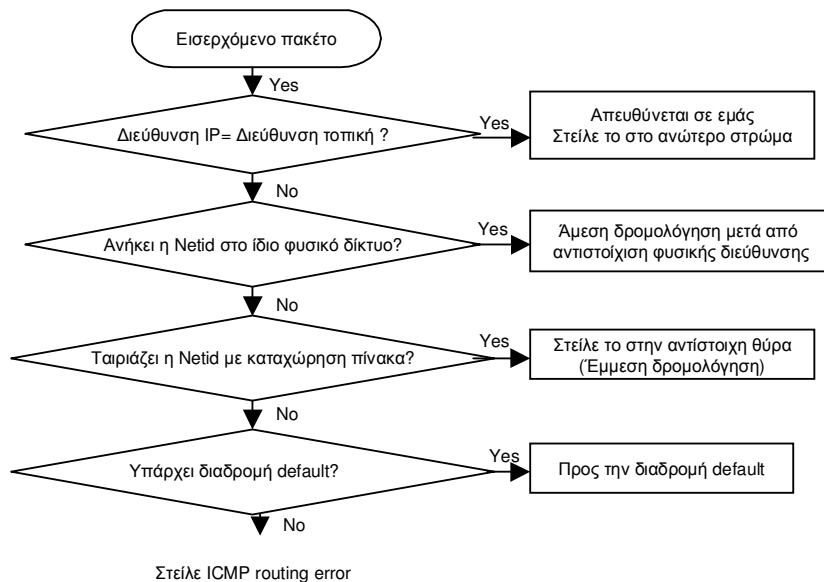
Αυτό πρέπει να γίνει για μία σειρά από λόγους. Κατ' αρχήν η λήψη ενός πακέτου που έχει άλλο προορισμό υποδηλώνει κάποιο λάθος στη διαδικασία της δρομολόγησης που δεν πρέπει να διαδοθεί. Επίσης με την επαναδρομολόγηση σπαταλώνται πόροι του δικτύου και του συγκεκριμένου ξενιστή που θα αποπειραθεί να δρομολογήσει το πακέτο. Υπάρχει επίσης η περίπτωση από απλά λάθη να προκύψει χάος. Σαν παράδειγμα μπορούμε να σκεφτούμε την εσφαλμένη εκπομπή προς όλους πακέτου που κατευθύνεται σε έναν συγκεκριμένο ξενιστή. Αν οι υπόλοιποι που θα λάβουν το πακέτο επιχειρήσουν να το δρομολογήσουν προς το σωστό παραλήπτη αυτός θα βρεθεί να «βομβαρδίζεται» από πολλαπλά αντίγραφα του ίδιου πακέτου. Τέλος όπως θα δούμε οι υπολογιστές που είναι εξουσιοδοτημένοι δρομολογητές κάνουν χρήση συγκεκριμένων πρωτοκόλλων για να ανταλλάσσουν μεταξύ τους πληροφορίες δρομολόγησης και να ενημερώνουν τους πίνακες δρομολόγησης και να ελέγχουν πιθανά σφάλματα (ιδέ 3.6). Οι απλοί ξενιστές λοιπόν θα πρέπει να βασίζονται σ' αυτούς και να μην κάνουν απόπειρες δρομολόγησης οι ίδιοι.

Ας δούμε τώρα πώς ενεργούν οι πύλες, οι οποίες όπως ελέχθη είναι κυρίως επιφορτισμένες με τη λειτουργία της δρομολόγησης με την λήψη ενός πακέτου IP. Κατ' αρχήν εξετάζουν και αυτές εάν η διεύθυνση του πακέτου είναι η δική τους τότε το πακέτο παραδίδεται όπως είδαμε στα ανώτερα στρώματα (πολύ μικρό ποσοστό πακέτων έχει για τελικό προορισμό πύλες, και προκύπτουν συνήθως κατά την εκτέλεση ειδικών πρωτοκόλλων μεταξύ πυλών, π.χ. IGP, EGP, or BGP που θα σκιαγραφηθούν στο τμήμα 3.6). Μετά εξετάζεται εάν η διεύθυνση υποδικτύου (net id) συμπίπτει με αυτή φυσικού δικτύου στο οποίο είναι προσκολλημένη η πύλη και εάν συμπίπτουν, το πακέτο προωθείται αφού ενθυλακωθεί στο φυσικό πλαίσιο και ευρεθεί η φυσική διεύθυνση του ξενιστή (άμεση δρομολόγηση). Εάν αυτό δεν συμβαίνει, εξετάζεται εάν υπάρχει καταχώριση στον πίνακα στην οποία να εμπίπτει η διεύθυνση υποδικτύου οπότε και προωθείται προς την διεπαφή που δείχνει η καταχώριση (έμμεση δρομολόγηση). Εάν δεν ταιριάζει καμία καταχώριση, το πακέτο προωθείται στην *προκαθορισμένη ή εξ ορισμού διαδρομή (default route)*. Εάν τέτοια δεν υπάρχει τότε αποστέλλεται σφάλμα δρομολόγησης μέσω ICMP (ιδέ κεφάλαιο 4). Τη σειρά των αποφάσεων δείχνει το επόμενο διάγραμμα.

Θα πρέπει εδώ να τονισθεί ότι σε κάθε βήμα μετάδοσης του πακέτου από πύλη σε πύλη, η IP διεύθυνση δεν αλλάζει και παραμένει πάντα αυτή του τελικού προορισμού. Αυτή είναι και η αρχή της υπηρεσίας χωρίς σύνδεση αντίθετα με δίκτυα που παρέχουν υπηρεσία με σύνδεση, (π.χ. X25, ATM) οπότε η διεύθυνση που φέρουν τα πακέτα υπό μορφή ετικέτας λογικού (νοητού) καναλιού αλλάζει σε κάθε κόμβο και αναγράφεται η ετικέτα του επόμενου ενδιαμέσου μονοπατιού στην προκαθορισμένη διαδρομή που θα ακολουθήσουν.

Η αρχικοποίηση γίνεται σε κάθε εκκίνηση μίας πύλης και χάρις στην ύπαρξη δικτύου η συνήθης λύση είναι να αποκτηθεί ένας αρχικός πίνακας δρομολόγησης είτε από γειτονικά μηχανήματα είτε από ένα συγκεκριμένο χώρο αποθήκευσης όπου αποθηκεύεται μόνιμα. Στη συνέχεια υπάρχει συνεχής ενημέρωση για την ανεύρεση περισσότερων διευθύνσεων καθώς και για οποιοδήποτε επιβεβλημένη αλλαγή (π.χ. λόγω αστοχίας κάποιας πύλης ή αχρήστευσης κάποιας σύνδεσης).

Η ενημέρωση αποτελεί διαδικασία η οποία στην απλούστερη περίπτωση όπως τα πρώτα χρόνια του ARPANET μπορεί να γίνεται και δια χειρός, δηλαδή οι διαχειριστές του δικτύου συμπληρώνουν τις θέσεις του πίνακα και ενημερώνουν με γνωστές διαδρομές σε κάθε περίπτωση. Αυτό βέβαια μπορεί να γίνει μόνο σε βασικές μορφές διασύνδεσης σε μικρή κλίμακα. Στη γενικότερη περίπτωση οι διαδικασίες αυτές γίνονται αυτόματα μέσω ειδικών πρωτοκόλλων που ανταλλάσσουν οι δρομολογητές μεταξύ τους. Τα πρωτόκολλα αυτά έχουν σαν κεντρικό σκοπό την ανακάλυψη της τοπολογίας του δικτύου την εξακρίβωση του «κόστους» κάθε διαδρομής ώστε να υπάρχουν προτιμώμενες διαδρομές και άλλες που χρησιμοποιούνται σαν εφεδρικές όταν η κύρια διαδρομή κόβεται ή υποφέρει από συμφόρηση. Με όλα αυτά τα στοιχεία ενημερώνονται οι καταχωρήσεις των πινάκων δρομολόγησης.



Διάγραμμα αποφάσεων δρομολογητή

Βασική αρχή στη συμπλήρωση των πινάκων δρομολόγησης είναι ότι δεν χρειάζεται όλοι οι διασυνδεδεμένοι υπολογιστές και όλες οι πύλες/δρομολογητές να περιλαμβάνουν καταχωρημένες όλες τις πιθανές διευθύνσεις προορισμού των πακέτων IP. Ένας πίνακας δρομολόγησης περιλαμβάνει επιλεκτικά κάποιες αντιστοιχίες προορισμών και πυλών στις οποίες πρέπει να δρομολογηθεί το πακέτο με διεύθυνση μία από τις γνωστές καταχωρημένες διευθύνσεις, για να προωθηθεί περαιτέρω στο επόμενο βήμα. Σημειώτεον ότι αναζητείται το τμήμα της διεύθυνσης που αντιστοιχεί στο net-id. Όταν δεν βρίσκεται ακριβές ταίριασμα με το πεδίο «net-id» αναζητείται η καταχώριση που εμφανίζει το μακρύτερο ταίριασμα, δηλ. ταιριάζουν τα περισσότερα bits ξεκινώντας από αριστερά (most significant bits). Η καταχώριση αυτή δίνει την πόρτα εξόδου στην οποία θα προωθηθεί το πακέτο (longest match). Η τεχνική αυτή παρακάμπτει τις κλάσεις διευθύνσεων και γι' αυτό λέγεται classless interdomain routing (CIDR). Αντί της χρήσεως της κλάσεως για τη εύρεση του ορίου μεταξύ net-id και host-id χρησιμοποιείται πλέον η μάσκα η οποία περιέχει 1 για τα Bits που ανήκουν στο net-id και 0 για αυτά που ανήκουν στο πεδίο host-id.

Στη συνέχεια θα εξετάσουμε πως λειτουργούν τα πρωτόκολλα ανακάλυψης τοπολογίας (topology discovery) πράγμα που απαιτεί μελέτη πρώτα της οργάνωσης του διαδικτύου σε αυτόνομες περιοχές.

3.6 Αλγόριθμοι εύρεσης τοπολογίας δικτύου (και διαχείριση των πινάκων)

Η δρομολόγηση δεδομενογραμμάτων στο Διαδίκτυο όπως ελέγχθη είναι ένας γενικός όρος που περιλαμβάνει την προώθηση (forwarding) την οποία έχουμε ήδη εξετάσει και την εύρεση της τοπολογίας (topology discovery) και την συνακόλουθη διαχείριση των πινάκων δρομολόγησης..

Ενώ η προώθηση λαμβάνει χώρα με κάθε εισερχόμενο πακέτο, η εκτέλεση των αλγορίθμων εύρεσης τοπολογίας και επακόλουθη ενημέρωση των πινάκων είναι μια λιγότερο συχνή διαδικασία που αποσκοπεί στην ανίχνευση μεταβολών στην τοπολογία (διακοπή ζεύξεων, αστοχία γειτονικών δρομολογητών, προσθήκη νέων κόμβων ή ζεύξεων κτλ.) ώστε η συνολική λειτουργική κατάσταση της δρομολόγησης στο δίκτυο να είναι βέλτιστη. Οι αλγόριθμοι που χρησιμοποιούνται είναι προσαρμοστικοί (adaptive) όπως απαιτείται σε ένα τόσο πολύπλοκο σύστημα) και εκτελούνται με λογισμικό διότι αφ' ενός δεν υπάρχει ζήτημα ταχύτητας εκτέλεσης και αφ' ετέρου είναι πολύ πολύπλοκοι για να μπορούν να εκτελεστούν από υλικό. Οι προσαρμοστικοί αλγόριθμοι δεν διατηρούν μόνιμες διαδρομές αλλά βασίζουν τις αποφάσεις δρομολόγησης στις συνθήκες που επικρατούν στο δίκτυο.

Ορισμένες ιδιότητες, οι οποίες είναι επιθυμητές σε έναν αλγόριθμο δρομολόγησης είναι: ορθότητα (προφανές), απλότητα (γρήγορη εκτέλεση χαμηλό κόστος), ανθεκτικότητα (να συνεχίζει να λειτουργεί κατά το δυνατόν εν μέσω αστοχιών υλικών και ζεύξεων), ευστάθεια (να μην οδηγεί μια στην μια επιλογή και μια στην άλλη υπό ίδιες συνθήκες αλλά να συγκλίνει γρήγορα στην βέλτιστη λύση), δικαιοσύνη (να μην ευνοεί ορισμένες ροές πακέτων έναντι άλλων) και βελτιστοποίηση. Η ανθεκτικότητα απαιτείται για την διατήρηση της λειτουργικότητας του παγκόσμιου δικτύου παρά τις απρόσμενες περιπτώσεις κατάρρευσης πυλών και απώλειας συνδέσεων που ενδέχεται να παρουσιαστούν κατά τη λειτουργία του. Η ευστάθεια ενός αλγορίθμου δεν είναι δεδομένη, καθώς υπάρχουν αλγόριθμοι που ενδέχεται να αργούν ή να μη συγκλίνουν ποτέ σε μία διαδρομή. Η αριστοποίηση τέλος εξαρτάται από τις προτεραιότητες που τίθενται και τις παραμέτρους που επιλέγονται για αριστοποίηση. Για παράδειγμα επιλέγοντας μέγιστο ρυθμό εξυπηρέτησης στο όριο της χωρητικότητας του δικτύου δεν μπορούμε να περιορίσουμε τη μέση καθυστέρηση των πακέτων λόγω των μεγάλων χρόνων αναμονής στις ουρές των δρομολογητών.

Παρακάμπτοντας τους αλγορίθμους κεντρικής δρομολόγησης που δεν είναι σύμφωνοι με τη φιλοσοφία του Διαδικτύου θα εξετάσουμε στη συνέχεια μερικούς τοπικής ή κατανεμημένης δρομολόγησης. Στην κατηγορία αυτή ανήκουν οι αλγόριθμοι της *καυτής πατάτας* (hot potato) και της *ανάστροφης ενημέρωσης* (backward learning) του Baran. Ο πρώτος προέβλεπε ότι κάθε δρομολογητής θα προσπαθούσε να «απαλλαγεί» από κάθε εισερχόμενο πακέτο τοποθετώντας το στη μικρότερη ουρά εξόδου χωρίς να ενδιαφέρεται που οδηγεί αυτή. Η ανάστροφη ενημέρωση προβλέπει την καταγραφή της διεύθυνσης κάθε δρομολογητή από τον οποίο διέρχεται ένα πακέτο και η αύξηση κατά ένα για κάθε βήμα ενός μετρητή. Έτσι με την παραλαβή κάθε πακέτου κάθε πύλη μπορεί να υπολογίσει την απόστασή της από τις προηγούμενες πύλες που έχει διέλθει το πακέτο και να ενημερώνει τον πίνακα δρομολόγησης υλοποιώντας τον τοπικό αλγόριθμο. Άλλοι αλγόριθμοι απομονωμένης δρομολόγησης είναι η πλημμυρίδα (flooding), η επιλεκτική πλημμυρίδα (selective flooding), ο τυχαίος περίπατος (random walk) κ.α.

Ωστόσο σήμερα στο Διαδίκτυο κυριαρχούν δύο δυναμικοί αλγόριθμοι κατανεμημένης δρομολόγησης: η δρομολόγηση διανύσματος αποστάσεων (distance vector routing) και η δρομολόγηση κατάστασης ζεύξεων (link state routing).

Ο αλγόριθμος διανύσματος αποστάσεων είναι επίσης γνωστός και ως κατανεμημένος αλγόριθμος Bellman- Ford από τους ερευνητές που τους ανέπτυξαν (Bellman 1957 και Ford and Fulkerson 1962). Ο αλγόριθμος καταστάσεως ζεύξης ονομάζεται και αλγόριθμος Dijkstra από το όνομα του ερευνητή που τον πρότεινε. Παράδειγμα πρακτικού πρωτοκόλλου που χρησιμοποιεί τον πρώτο είναι το πρωτόκολλο RIP (Routing Information Protocol) του δε δευτέρου το πρωτόκολλο OSPF (Open Shortest Path First).

Αξίζει να σημειωθεί ότι η ίδια διαδικασία εύρεσης τοπολογίας που χρησιμοποιούν τα δίκτυα δεδομενογραμμάτων τα οποία προσφέρουν υπηρεσία χωρίς σύνδεση (connectionless), ισχύει και για δίκτυα νοητών κυκλωμάτων που παρέχουν υπηρεσίες με σύνδεση (connection-oriented) όπως το X.25, μόνο που η επιλογή θύρας εξόδου γίνεται άπαξ κατά την εγκατάσταση μιας κλήσης. Δηλαδή το πακέτο εγκατάστασης κλήσης πρέπει να προωθηθεί με επίγνωση της τοπολογίας και των γειτονικών κόμβων για λόγους βελτιστοποίησης. Ωστόσο από κει και πέρα όλα τα άλλα πακέτα που ακολουθούν στην ίδια σύνδεση προωθούνται πάντα στην ίδια αρχικώς επιλεγείσα διαδρομή μέχρι την απόλυση

της σύνδεσης. Έτσι οι αλγόριθμοι που θα περιγράψουμε (διανύσματος αποστάσεων, και καταστάσεως ζεύξης) χρησιμοποιούνται και στα δίκτυα X.25, ATM, SNA, κτλ.

3.6.1 Δρομολόγηση Διανύσματος Αποστάσεων

Οι αλγόριθμοι δρομολόγησης διανύσματος αποστάσεων λειτουργούν με το να αναγκάζουν κάθε δρομολογητή να διατηρεί ένα σει τιμών (δηλ. ένα διάνυσμα) που δίνει τη βέλτιστη γνωστή απόσταση για κάθε προορισμό, καθώς και ποια γραμμή πρέπει να χρησιμοποιηθεί για να φτάσει κανείς εκεί. Αυτά τα διανύσματα ενημερώνονται με πληροφορίες που ανταλλάσσονται με τους γείτονες και σε αυτά βασίζεται η σύνταξη των πινάκων δρομολόγησης..

Ο αλγόριθμος διανύσματος αποστάσεων ήταν ο αρχικός αλγόριθμος δρομολόγησης του ARPANET και χρησιμοποιήθηκε επίσης και στο Internet με το όνομα RIP (routing Information Protocol) και σε πρώιμες εκδόσεις του DECnet και του IPX της Novell.

Στη δρομολόγηση διανύσματος αποστάσεων, κάθε δρομολογητής διατηρεί έναν πίνακα δρομολόγησης ο οποίος περιέχει μια καταχώρηση για κάθε δρομολογητή του υποδικτύου. Αυτή η καταχώρηση περιέχει δύο τμήματα: την προτεινόμενη εξερχόμενη γραμμή που πρέπει να χρησιμοποιηθεί για αυτόν τον προορισμό και μια εκτίμηση του χρόνου ή της απόστασης για τον προορισμό αυτόν. Το χρησιμοποιούμενο μέτρο μπορεί να είναι ο αριθμός βημάτων, ή και η καθυστέρηση σε msec, ο συνολικός αριθμός αποθηκευμένων πακέτων κατά μήκος της διαδρομής ή κάτι παρόμοιο.

Ο δρομολογητής μετρά την "απόσταση" για καθέναν από τους γείτονες του. Εάν η μονάδα μέτρησης είναι τα βήματα, οι γείτονες απέχουν ένα βήμα. Εάν το είναι το μήκος της ουράς, το μήκος κάθε ουράς θα μπει στον πίνακα. Εάν το μέτρο είναι η καθυστέρηση, ο δρομολογητής μπορεί να τη μετρήσει απ' ευθείας με ειδικά πακέτα ECHO, στα οποία ο δέκτης βάζει μια χρονική σφραγίδα και τα στέλνει αμέσως πίσω χωρίς χρονοτριβή. Ας δούμε την σειρά των ενεργειών για να γίνει κατανοητή η αρχή λειτουργίας του αλγορίθμου στην περίπτωση που η απόσταση μετρείται με την καθυστέρηση. Ο δρομολογητής μετρά πρώτα την καθυστέρηση για καθέναν από τους γείτονες του. Κάθε T_{msec} κάθε δρομολογητής στέλνει σε κάθε γείτονα του έναν κατάλογο με τις εκτιμημένες απ' αυτόν καθυστερήσεις για κάθε προορισμό. Επίσης λαμβάνει έναν παρόμοιο κατάλογο από κάθε γείτονα. Φανταστείτε ότι ένας από αυτούς τους πίνακες έχει μόλις φτάσει από τον γείτονα X, με το X_j να είναι η εκτίμηση του X για το πόσο χρόνο χρειάζεται για να φτάσει στον δρομολογητή i . Εάν ο δρομολογητής γνωρίζει ότι η καθυστέρηση μέχρι τον X είναι m msec, γνωρίζει επίσης ότι μπορεί να φτάσει στον i μέσω του X, σε χρόνο X_j+m msec. Με το να εκτελεί αυτόν τον υπολογισμό για κάθε γείτονα, ένας δρομολογητής μπορεί να βρει ποια εκτίμηση φαίνεται να είναι η καλύτερη και να χρησιμοποιήσει αυτήν την εκτίμηση και την αντίστοιχη γραμμή στον νέο του πίνακα δρομολόγησης. Σημειώστε ότι ο παλιός πίνακας δρομολόγησης δεν χρησιμοποιείται στον υπολογισμό.

3.6.2 Δρομολόγηση Κατάστασης Ζεύξεων (link state)

Η δρομολόγηση διανύσματος αποστάσεων χρησιμοποιήθηκε στο ARPANET μέχρι το 1979, όταν αντικαταστάθηκε από τον αλγόριθμο κατάστασης ζεύξεων. Δύο ήταν τα βασικά προβλήματα που οδήγησαν στην αντικατάστασή του. Πρώτον, επειδή το μέτρο καθυστέρησης ήταν το μήκος της ουράς, δεν έπαιρνε υπόψη του τη χωρητικότητα στην επιλογή των διαδρομών. Αρχικά, όλες οι γραμμές ήταν των 56 Kbps, έτσι η χωρητικότητα δεν αποτελούσε πρόβλημα, αλλά όταν μερικές γραμμές αναβαθμίστηκαν στα 230 kbps και άλλες στο 1,544 Mbps, το ότι δεν λαμβανόταν υπόψη η χωρητικότητα ήταν μεγάλο πρόβλημα. Βέβαια, ήταν δυνατή η αλλαγή του μέτρου καθυστέρησης έτσι ώστε να συμπεριληφθεί η χωρητικότητα της γραμμής. Όμως, το δεύτερο πρόβλημα ήδη υπήρχε, δηλαδή, μερικές φορές ο αλγόριθμος καθυστερούσε υπερβολικά να συγκλίνει. Για τους λόγους αυτούς, αντικαταστάθηκε από έναν εντελώς καινούργιο αλγόριθμο που τώρα ονομάζεται δρομολόγηση κατάστασης ζεύξεων. Παραλλαγές της δρομολόγησης κατάστασης ζεύξεων χρησιμοποιούνται ευρέως σήμερα (π.χ. το πρωτόκολλο OSPF που αντικατέστησε το RIP στο Διαδίκτυο).

Η ιδέα πίσω από τον αλγόριθμο κατάστασης ζεύξεων είναι απλή και μπορεί να συνοψισθεί σε πέντε βήματα. Ο κάθε δρομολογητής πρέπει:

1. Να ανακαλύπτει τους γείτονές του και να μαθαίνει τις διευθύνσεις δικτύου τους.
2. Να υπολογίζει την καθυστέρηση ή το κόστος για τον καθένα από τους γείτονές του.
3. Να κατασκευάζει ένα πακέτο που πληροφορεί όλα αυτά τα οποία μόλις έχει μάθει.
4. Να στέλνει αυτό το πακέτο σ' όλους τους άλλους δρομολογητές.
5. Να υπολογίζει τη συντομότερη διαδρομή προς κάθε άλλον δρομολογητή.

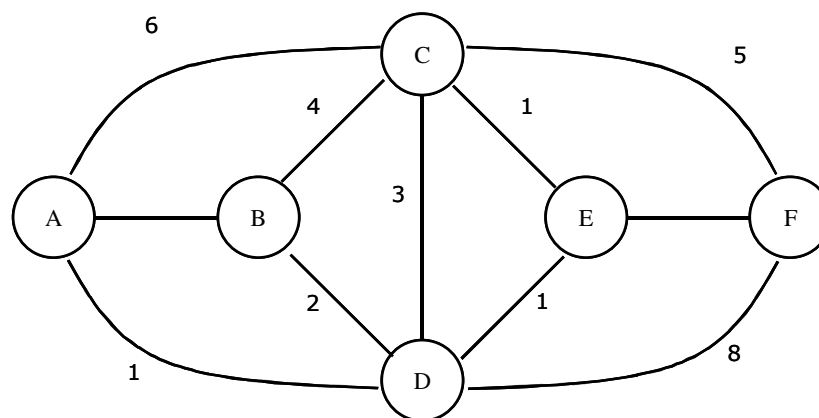
Στην ουσία, η πλήρης τοπολογία και όλες οι καθυστερήσεις μετριοούνται πειραματικά και διανέμονται σε κάθε δρομολογητή. Ο αλγόριθμος του Dijkstra μπορεί κατόπιν να χρησιμοποιηθεί για να βρει τη συντομότερη διαδρομή προς κάθε άλλο δρομολογητή.

Κατά την εκκίνηση ενός δρομολογητή, το πρώτο καθήκον του είναι να μάθει ποιοι είναι οι γείτονές του. Επιτυγχάνει αυτόν τον στόχο στέλνοντας ένα ειδικό πακέτο HELLO σε όλες τις διεπαφές του (πόρτες). Ο δρομολογητής που βρίσκεται στο άλλο άκρο της κάθε ζεύξης στέλνει μια απάντηση και τον πληροφορεί για το ποιος είναι. Τα ονόματα αυτά θα πρέπει να είναι μοναδικά σ' όλο το σύστημα, επειδή όταν ένας απομακρυσμένος δρομολογητής αργότερα μάθει ότι τρεις δρομολογητές είναι όλοι, τους συνδεδεμένοι στον δρομολογητή X είναι σημαντικό να μπορεί να καθορίσει το κατά πόσον και οι τρεις εννοούν τον ίδιο X ή όχι.

3.7 Οι Αλγόριθμοι Dijkstra και Bellman-Ford

Κλείνοντας την περιγραφή των μηχανισμών δρομολόγησης θα δώσουμε την αφηρημένη θεωρητική περιγραφή των δύο πιο ευρέως διαδεδομένων αλγόριθμων εύρεσης τοπολογίας που αναπτύχθηκαν από τον Dijkstra και τους Bellman-Ford για την εύρεση βέλτιστων διαδρομών. Έχουν βρει ωστόσο χρήση και σε άλλα πρακτικά προβλήματα βελτιστοποίησης διαδρομών. Η μαθηματική περιγραφή των αλγόριθμων αυτών βασίζεται στην μοντελοποίηση του δικτύου σε μορφή αφηρημένου γράφου, όπου οι ακμές (κόμβοι) αναπαριστούν τους δρομολογητές και οι γραμμές που τις συνδέουν τις ζεύξεις επικοινωνίας. Ως παράδειγμα θα χρησιμοποιήσουμε το γράφο του παρακάτω σχήματος 3.10 στον οποίο οι αριθμοί-ετικέτες δίπλα στις ζεύξεις παριστούν το κριτήριο κόστους (π.χ. απόσταση) και θα αναζητήσουμε τις βέλτιστες διαδρομές (μικρότερου συνολικού κόστους) από τον κόμβο εκκίνησης A προς κάθε άλλο κόμβο με χρήση των αλγορίθμων. Έτσι όταν θα αποφασίζουν την προώθηση κάθε πακέτου οι κόμβοι θα βελτιστοποιούν κάποιο κριτήριο.

Οι αριθμοί του κόστους κάθε ζεύξης μπορεί να βασίζονται σε οποιαδήποτε μετρική σχέση έχει επιλεγεί από τους διαχειριστές του δικτύου. Έτσι μπορεί να εκφράζει το φυσικό μήκος (και άρα καθυστέρηση διάδοσης) της σύνδεσης, την καθυστέρηση αναμονής (το φορτίο) προς τη συγκεκριμένη διαδρομή ή το λογιστικό κόστος λειτουργίας, συντήρησης κτλ. Προφανώς το μήκος δεν μεταβάλλεται παρά σε περίπτωση νέας ζεύξης ή βλάβης υπάρχουσας, αλλά η καθυστέρηση αναμονής είναι κάτι που μεταβάλλεται συνεχώς άρα ο αλγόριθμος πρέπει να επαναλαμβάνεται συχνά δίνοντας την απαιτούμενη δυναμική και προσαρμοστικότητα. Από την άλλη μπορεί να επιθυμούμε την ελαχιστοποίηση του πλήθους των βημάτων δρομολόγησης οπότε το κόστος κάθε διαδρομής παίρνει τιμή 1 ανεξάρτητα του μήκους ζεύξης. Πολύ συνηθισμένη είναι και η χρήση τιμών αντιστρόφως ανάλογων προς το ρυθμό (bit rate) της κάθε ζεύξης, ή κάποιος συνδυασμός παραγόντων. Το κόστος των ζεύξεων μπορεί να είναι διαφορετικό σε κάθε κατεύθυνση, όπως π.χ. όταν χρησιμοποιείται η καθυστέρηση αναμονής στις ουρές εξόδου, ή ίδιο όπως στην περίπτωση που είναι το φυσικό μήκος της ζεύξης. Η τελευταία περίπτωση χρησιμοποιείται στον γράφο του παραδείγματος του σχήματος 3.10 που θα χρησιμοποιηθεί πιο κάτω. Σε κάθε περίπτωση οι αλγόριθμοι δεν εξαρτώνται από τον τρόπο επιλογής των ετικετών κόστους.



Σχήμα 3.10. Γράφος δικτύου και παράδειγμα εφαρμογής του αλγορίθμου του Dijkstra. Σε κάθε περίπτωση η διατύπωση του προβλήματος μπορεί να δοθεί ως εξής:

«Δοθέντος γράφου δικτύου με κόμβους συνδεδεμένους μέσω ζεύξεων με κόστος εκάστης που δίδεται για κάθε κατεύθυνση, να ευρεθεί για κάθε συνδυασμό δύο κόμβων η διαδρομή ελαχίστου συνολικού κόστους.»

Όλοι οι γνωστοί αλγόριθμοι βρίσκουν το κόστος από τον κόμβο αφετηρίας προς όλους τους άλλους αφού δεν υπάρχει γνωστός τρόπος να μπορεί να βρει αποδεδειγμένα το ελάχιστο κόστος προς ένα κόμβο χωρίς να βρει το ελάχιστο κόστος προς κάθε κόμβο και από κάθε άλλο κόμβο. Αυτό δεν είναι σοβαρό μειονέκτημα αφού ο κάθε κόμβος θέλει να ξέρει το κόστος προς όλους τους άλλους.

3.7.1 Ο αλγόριθμος Dijkstra

Θα εξετασθεί η λύση για ένα κόμβο εκκίνησης (τον A του σχήματος 3.10) προς όλους τους άλλους κόμβους. Ομοίως μπορεί να εκτελεσθεί ο αλγόριθμος για κάθε άλλο κόμβο εκκίνησης. Ωστόσο στην πράξη στο Διαδίκτυο όπου εκτελείται καταναμημένα ο κάθε κόμβος ενδιαφέρεται για το κόστος των διαδρομών που έχουν σαν εκκίνηση τον εαυτό του. Ο αλγόριθμος απαιτεί την ύπαρξη θετικού κόστους σε κάθε ακμή (όπως όντως συμβαίνει στην περίπτωση γράφων τηλεπικοινωνιακών δικτύων) και προχωρεί με διαδοχικά βήματα κατά τα οποία συμμετέχει ένα υποσύνολο M των κόμβων του δικτύου ενώ σε κάθε βήμα προστίθεται ένας ακόμη κόμβος.

Στην εκκίνηση μόνο ο κόμβος αφετηρίας A ανήκει στο υποσύνολο M και το κόστος προς όλους τους κόμβους που δεν είναι άμεσα συνδεδεμένοι προς τον κόμβο αυτόν θεωρείται άπειρο. Το κόστος K_j προς όλους τους άμεσα συνδεδεμένους με τον A κόμβους είναι απλά το κόστος της αντίστοιχης ζεύξης $k_{A,j}$. Ακολούθως αλγόριθμος εκτελεί τα εξής δύο βήματα προσθέτοντας ένα νέο κόμβο στο M κάθε φορά μέχρι εξάντλησης όλων των κόμβων:

Βήμα 1: Βρίσκει τον γειτονικό κόμβο που δεν ανήκει ήδη στο M και ο οποίος έχει το χαμηλότερο κόστος προς τον A και τον προσθέτει στο M.

Βήμα 2: Ενημερώνει το νέο προσωρινό κόστος και τις νέες ελάχιστες διαδρομές προς κάθε κόμβο από τον A μετά την προσθήκη του νέου κόμβου.

Τα ανωτέρω δύο βήματα επαναλαμβάνονται μέχρι να προστεθούν όλοι οι κόμβοι.

Έτσι, εάν με K_i συμβολίσουμε το προσωρινό κόστος διαδρομής του εξεταζόμενου κόμβου i και k_{iA} το κόστος της ζεύξης i-A, η διατύπωση του αλγορίθμου έχει ως εξής:

- Αρχικά $M = \{A\}$, $K_A = 0$ και $K_j = k_{A,j}$, για $j \neq A$
- **Βήμα 1.** (βρες επόμενο κοντινότερο κόμβο) Βρες $i \notin M$ τέτοιο ώστε:

$$K_i = \min_j K_j, \text{ για όλα τα } j \notin M$$

Θέσε $M = M \cup \{i\}$. Εάν το M περιέχει όλους τους κόμβους: τέλος, αλλιώς:

- **Βήμα 2.** (ενημέρωσε τις επικέτες) Για όλα $j \notin M$ θέσε:

$$K_j = \min[K_j, K_i + k_{ij}]$$

Επιστροφή στο βήμα 1.

Βλέπουμε στο σχήμα 3.11 τα σχετικά βήματα για το γράφο του παραδείγματος 3.10

Η όλη αξία του αλγορίθμου για την λύση σε εύλογο χρόνο αυτού του τυπικού NP-hard προβλήματος βελτιστοποίησης (δηλ προβλήματος που δεν λύνεται στη γενικότητά του σε πολυωνυμικό χρόνο) κρύβεται στην παρατήρηση ότι με την προσθήκη του νέου κόμβου η διαδρομή ελαχίστου κόστους περνά **μόνο** μέσω κόμβων που ήδη ανήκουν στο M ως εκ του τρόπου και των ελέγχων που γίνονται προκειμένου να επιλεγεί ο επόμενος κάθε φορά. Το δε ελάχιστο κόστος του νέου κόμβου j προκύπτει με πρόσθεση στο κόστος του πλησιέστερου γείτονα i (που ξέρουμε ότι είναι ελάχιστο), το άμεσο κόστος της επικέτας kij και αυτό ακριβώς εκφράζει η εξίσωση του βήματος 2.

3.7.2 Ο αλγόριθμος Bellman-Ford

Το ίδιο πρόβλημα λύνει με παρόμοια αλλά διαφορετική διαδικασία και ο αλγόριθμος των Bellman-Ford, ο οποίος μπορεί να χρησιμοποιήσει και αρνητικό κόστος διαδρομών. Και οι δύο βασίζονται στην μαθηματική επαγωγή και εκτελούν βήματα με προκαθορισμένες ιδιότητες και καταλήγουν στην ίδια λύση. Ο αλγόριθμος αυτός δεν περιορίζεται σε υποσύνολο των κόμβων κάθε φορά αλλά περιορίζεται κατά το πλήθος των βημάτων την κάθε φορά. Αρχίζει με την εύρεση πρώτα της διαδρομής μικρότερου κόστους για διαδρομές ενός βήματος, κατόπιν για διαδρομές δύο βημάτων κ.ο.κ. Φορμαλιστικά ο αλγόριθμος αυτός διατυπώνεται με τον παρακάτω τρόπο:

Ονομάζουμε $K_j^{(h)}$ το ελάχιστο κόστος της διαδρομής το πολύ h βημάτων από τον κόμβο A στον j και k_{ij} το κόστος της ζεύξης μεταξύ των γειτονικών κόμβων i και j .

Αρχικοποίηση: Δεχόμαστε $K_j^{(0)} = \infty$ για όλους τους κόμβους $j \neq A$ δηλαδή το αρχικό κόστος είναι άπειρο για μηδενικό βήμα Επίσης ότι για κάθε i : $K_A^{(h)} = 0$, δηλαδή το κόστος από τον κόμβο εκκίνησης A στον εαυτό του είναι μηδέν για κάθε πλήθος βημάτων.

Επαναλαμβανόμενο βήμα: Διαδοχικά για κάθε κόμβο i και κάθε $h \geq 0$: $K_i^{(h+1)} = \min [K_j^{(h)} + k_{ij}]$, όπου ο κόμβος j είναι γειτονικός του i άρα η διαδρομή από A προς τον i τερματίζει με την άμεση ζεύξη ij .

Δηλαδή εάν έχουμε βρει το ελάχιστο κόστος για την μετάβαση από τον A στον κόμβο i με h βήματα, τότε εξετάζουμε εάν δίνει μικρότερο κόστος η χρήση $h+1$ βημάτων δηλαδή εάν υπάρχει γείτονας j με ελάχιστο κόστος h βημάτων τέτοιο που η πρόσθεση του κόστους της ζεύξης ij να δίνει μικρότερη τιμή. Αλλιώς παραμένει σαν προσωρινό ελάχιστο κόστος για το κόμβο i το προηγούμενο.

Στον κατωτέρω πίνακα δίδονται τα αποτελέσματα των επαναλήψεων του αλγορίθμου Bellman-Ford για τον ίδιο γράφο του σχήματος 3-10, που καταλήγουν στην ίδια λύση με το αλγόριθμο Dijkstra.

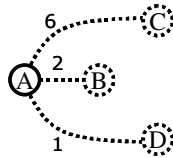
βήμα	Κόμβος B		Κόμβος C		Κόμβος D		Κόμβος E		Κόμβος F	
	Κόστ.	Διαδρομή	Κόστ.	Διαδρομή	Κόστ.	Διαδρομή	Κόστ.	Διαδρομή	Κόστ.	Διαδρομή
0	$K_B^0 = \infty$	--	$K_C^0 = \infty$	--	$K_D^0 = \infty$	--	$K_E^0 = \infty$	--	$K_F^0 = \infty$	--
1	$K_B^1 = 2$	A-B	$K_C^1 = 6$	A-C	$K_D^1 = 1$	A-D	$K_E^1 = \infty$		$K_F^1 = \infty$	
2	$K_B^2 = 2$	A-B	$K_C^2 = 4$	A-D-C	$K_D^2 = 1$	A-D	$K_E^2 = 2$	A-D-E	$K_F^2 = 9$	A-D-F
3	$K_B^3 = 2$	A-B	$K_C^3 = 3$	A-D-E-C	$K_D^3 = 1$	A-D	$K_E^3 = 2$	A-D-E	$K_F^3 = 4$	A-D-E-F
4	$K_B^4 = 2$	A-B	$K_C^4 = 3$	A-D-E-C	$K_D^4 = 1$	A-D	$K_E^4 = 2$	A-D-E	$K_F^4 = 4$	A-D-E-F

Θα κλείσουμε με ένα μικρό σχόλιο σχετικά με την πολυπλοκότητα των δύο ανωτέρω αλγορίθμων και την προκύπτουσα ταχύτητα εκτέλεσης. Και οι δύο αλγόριθμοι απαιτούν στη χειρότερη περίπτωση $N-1$ βήματα όπου N το πλήθος των κόμβων. Κάθε βήμα απαιτεί πράξεις ανάλογες προς το N για τον αλγόριθμο Dijkstra ενώ στη περίπτωση του Bellman-Ford πρέπει να γίνουν πράξεις ανάλογες του N με $N-1$ συνδυασμούς σε κάθε περίπτωση, άρα συνολικά οι πράξεις είναι ανάλογες του N^2 για τον Dijkstra και ανάλογες του N^3 για τον Bellman-Ford. Αυτό συμβολίζεται με $O(N^2)$ και $O(N^3)$ αντίστοιχα, που σημαίνει. Order of (N^2) , δηλαδή οι απαιτούμενες πράξεις είναι της τάξεως του (N^2) . Αυτό θα οδηγούσε στο συμπέρασμα ότι ο αλγόριθμος Dijkstra είναι αποδοτικότερος αλλά αυτές είναι οι εκτιμήσεις για τη χειρότερη περίπτωση. Ωστόσο αν δει κανείς πιο προσεκτικά τις απαιτούμενες πράξεις που απαιτούνται για τον Bellman-Ford είναι $O(mA)$ όπου A είναι το πλήθος των ζεύξεων και m το μέσο πλήθος των επαναλήψεων για κάθε κόμβο μέχρι τον τερματισμό του αλγορίθμου (που συμβαίνει όταν δεν έχουμε βελτίωση καμίας διαδρομής) και ταυτόχρονα είναι το μέγιστο πλήθος των ζεύξεων στη βέλτιστη διαδρομή). Σε πολλούς πρακτικούς γράφους όπου $N^2 \gg A$ και $m \ll N$ το $O(mA)$ είναι πολύ καλύτερο από το $O(N^2)$. Έτσι ο ένας αλγόριθμος μπορεί να υπερτερεί σε κάποια περίπτωση και σε κάποια άλλη ο άλλος. Ωστόσο, στην πράξη ο αλγόριθμος Bellman-Ford αργεί πολλές φορές να συγκλίνει και αυτό οδήγησε στην αντικατάσταση του RIP που τον χρησιμοποιούσε από το OSPF που χρησιμοποιεί τον Dijkstra.

ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ ΑΛΓΟΡΙΘΜΟΥ Dijkstra

Αρχικοποίηση (1η επανάληψη)

$M = \{A\}$



Αρχικό κόστος διαδρομών από A

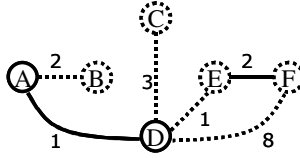
$K_B=2$ μέσω διαδρομής A-B
 $K_C=6$ μέσω διαδρομής A-C
 $K_D=1$ μέσω διαδρομής A-D
 $K_E=(\text{άπειρο})$
 $K_F=(\text{άπειρο})$

Βήμα 1: Βρες γείτονα με min κόστος (Απ. D)

Βήμα 2: Πρόσθεση του D στο σύνολο M και μετά επανάληψη ενημέρωσης πίνακα κόστους

2η επανάληψη

$M = \{A, D\}$



Νέο κόστος διαδρομών από A

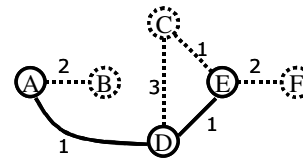
$K_B=2$ μέσω διαδρομής A-B
 $K_C=4$ μέσω διαδρομής A-D-C*
 $K_D=1$ μέσω διαδρομής A-D
 $K_E=2$ μέσω διαδρομής A-D-E*
 $K_F=9$ μέσω διαδρομής A-D-E*

Βήμα 1: Νέος γείτονας με min κόστος (Απ. B, E έστω E)

Βήμα 2: Πρόσθεση του E στο σύνολο M και 3η επανάληψη ενημέρωσης πίνακα κόστους

3η επανάληψη

$M = \{A, D, E\}$



Νέο κόστος διαδρομών από A

$K_B=2$ μέσω διαδρομής A-B
 $K_C=3$ μέσω διαδρομής A-D-E-C*
 $K_D=1$ μέσω διαδρομής A-D
 $K_E=2$ μέσω διαδρομής A-D-E
 $K_F=4$ μέσω διαδρομής A-D-E-F*

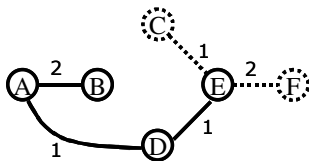
Βήμα 1: Νέος γείτονας με min κόστος (Απ. B)

Βήμα 2: Πρόσθεση του B στο σύνολο M και 4η επανάληψη ενημέρωσης πίνακα κόστους

*Ο αστερίσκος δείχνει βελτίωση κόστους μέσω του νέου κόμβου

4η επανάληψη

$M = \{A, B, D, E\}$



Νέο κόστος διαδρομών από A

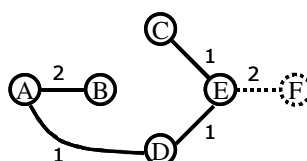
$K_B=2$ μέσω διαδρομής A-B
 $K_C=3$ μέσω διαδρομής A-D-E-C
 $K_D=1$ μέσω διαδρομής A-D
 $K_E=2$ μέσω διαδρομής A-D-E
 $K_F=4$ μέσω διαδρομής A-D-E-F

Βήμα 1: Νέος γείτονας με min κόστος (Απ. C)

Βήμα 2: Πρόσθεση του C στο σύνολο M και 5η επανάληψη ενημέρωσης πίνακα κόστους

5η επανάληψη

$M = \{A, B, C, D, E\}$



Νέο κόστος διαδρομών από A

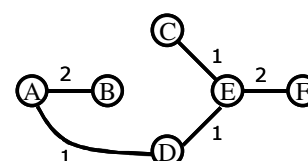
$K_B=2$ μέσω διαδρομής A-B
 $K_C=3$ μέσω διαδρομής A-D-E-C
 $K_D=1$ μέσω διαδρομής A-D
 $K_E=2$ μέσω διαδρομής A-D-E
 $K_F=4$ μέσω διαδρομής A-D-E-F

Βήμα 1: Νέος γείτονας (και τελευταίος) με min κόστος: E

Βήμα 2: Πρόσθεση του F στο σύνολο M και 6η επανάληψη ενημέρωσης πίνακα κόστους

6η επανάληψη

$M = \{A, B, C, D, E, F\}$



Νέο κόστος διαδρομών από A

$K_B=2$ μέσω διαδρομής A-B
 $K_C=3$ μέσω διαδρομής A-D-E-C
 $K_D=1$ μέσω διαδρομής A-D
 $K_E=2$ μέσω διαδρομής A-D-E
 $K_F=4$ μέσω διαδρομής A-D-E-F
 (Ουδεμία αλλαγή)

Έχουν ενταχθεί στο M όλοι οι κόμβοι και άρα αυτό είναι το τελικό κόστος

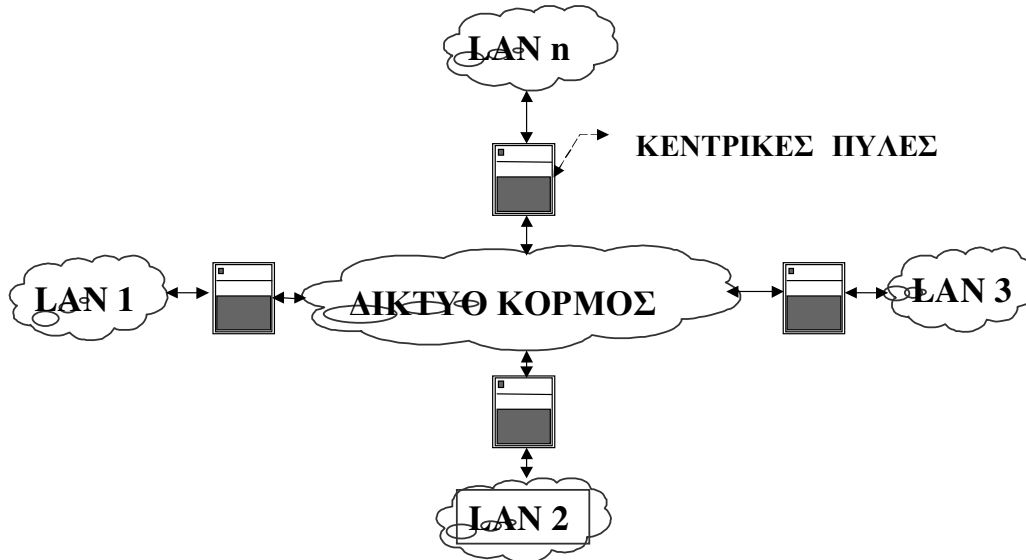
Θα υπήρχε τελική διαφορά εάν είχαμε επιλέξει τον B πριν τον E μετά την 2η επανάληψη;

Σχήμα 3.11. Γράφος δικτύου και αποτελέσματα των διαδοχικών επαναλήψεων του αλγορίθμου Dijkstra.

3.8 Επιπτώσεις της οργάνωσης του δικτύου στην δρομολόγηση – ιστορικά στοιχεία

Δρομολόγηση μέσω δικτύου κορμού

Το πρώτο παράδειγμα ευρείας διασύνδεσης δικτύων βασισμένης στα πρωτόκολλα TCP/IP και πρόγονο του σύγχρονου Internet αποτέλεσε το δίκτυο ARPANET το οποίο χρησιμοποιήθηκε ως *δίκτυο κορμού (backbone)* μέσω του οποίου επιτεύχθηκε επικοινωνία πολλών τοπικών δικτύων. Η δομή του απεικονίζεται στο παρακάτω σχήμα.



Σχήμα 3.3. Παράδειγμα διασύνδεσης δικτύων μέσω ενός δικτύου κορμού και πρόσβαση σε αυτό μέσω κεντρικών πυλών.

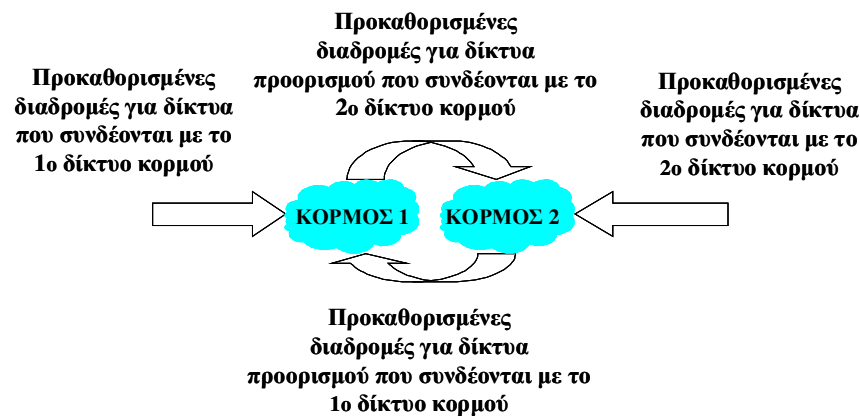
Η δρομολόγηση πακέτων από ένα τοπικό δίκτυο προς κάποιο απομακρυσμένο έπρεπε να γίνει μέσω κάποιας κεντρικής πύλης προς το δίκτυο κορμού, μέσω του οποίου θα μεταδιόταν στην πύλη η οποία τελικά ήταν συνδεδεμένη με το τοπικό δίκτυο προορισμού, ώστε να έχει πρόσβαση στον τελικό host προορισμού. Όλες οι διευθύνσεις λοιπόν που αντιστοιχούσαν σε απομακρυσμένα δίκτυα είχαν σαν προκαθορισμένη διαδρομή την κεντρική πύλη που ένωνε το τοπικό δίκτυο με «τον έξω κόσμο» (το δίκτυο κορμού). Οι κεντρικές πύλες, θα μπορούσαν να ακολουθήσουν την τακτική της μερικής ενημέρωσης και να χρησιμοποιούν προκαθορισμένες διαδρομές, ώστε να αποθηκεύουν μικρούς σε μέγεθος πίνακες δρομολόγησης. Αυτή η τακτική παρουσιάζει τα εξής προβλήματα που εκδηλώθηκαν όταν το δίκτυο μεγάλωσε:

- Για υπάρχει δυνατότητα πρόσβασης σε όλες τις πιθανές περιπτώσεις προορισμού όλες οι κεντρικές πύλες θα έπρεπε να σχηματίζουν μια γιγαντιαία κλειστή αλυσίδα, ώστε μέσω προκαθορισμένων διαδρομών να μπορεί ένα πακέτο να μεταδοθεί βήμα, βήμα προς όλα τα συνδεδεμένα δίκτυα.
- Απαιτείται συντονισμός όλων των κεντρικών πυλών ως προς τις προκαθορισμένες διαδρομές που θα ορίσει κάθε μία, ώστε να σχηματιστεί ένας κλειστός βρόχος.
- Πιθανά λάθη με ανύπαρκτες διευθύνσεις θα περνούσαν απαρατήρητα.

Για να αποφευχθούν τα παραπάνω επιλέχθηκε η πλήρης ενημέρωση όλων των κεντρικών πυλών με όλες τις πιθανές διευθύνσεις των διασυνδεδεμένων δικτύων, ώστε τα πακέτα να δρομολογούνται με ένα μόνο βήμα μέσα στο δίκτυο κορμού.

Δρομολόγηση μεταξύ ομοτίμων δικτύων κορμού

Παρ' όλο που το Internet «γεννήθηκε» βασισμένο στο ανωτέρω περιγραφέν κεντρικά ελεγχόμενο δίκτυο κορμού, αναπτύχθηκε σε μεγαλύτερο βαθμό από αυτόν που μπορούσε να ελέγχεται κεντρικά. Το μεγάλο πλήθος κεντρικών πυλών δυσχέρανε τον συντονισμό και την ενημέρωση των πινάκων δρομολόγησης, ενώ δεν ήταν δυνατή η πρόσβαση κάποιων τοπικών δικτύων μέσω μίας πύλης στο δίκτυο κορμού. Το επόμενο βήμα στη ανάπτυξη του Internet ήταν η καθιέρωση του δικτύου NSFNET ως *ομότιμο* δικτύου κορμού. Τα δύο κεντρικά δίκτυα συνδέθηκαν αρχικά μέσω μίας πύλης και οι εσωτερικές κεντρικές πύλες κάθε δικτύου περιελάμβαναν τις διευθύνσεις του δικού τους δικτύου και χρησιμοποιούσαν την πύλη αυτή ως προκαθορισμένη διαδρομή για πακέτα με διευθύνσεις που αντιστοιχούσαν στο άλλο δίκτυο. Η λύση αυτή εγκυμονεί τον κίνδυνο δημιουργίας ατέρμωνων βρόχων σε περίπτωση χρήσης λανθασμένης IP διεύθυνσης. Η λογική αυτής της διάκρισης απεικονίζεται στο σχήμα 3.4



Σχήμα 3.4. Διαίρεση της συγκεντρωτικής αρχιτεκτονικής σε δύο ομότιμα δίκτυα κορμού, που δημιουργεί δύο ομάδες πυλών με μερική πληροφόρηση των πινάκων τους και χρήση προκαθορισμένων διαδρομών.

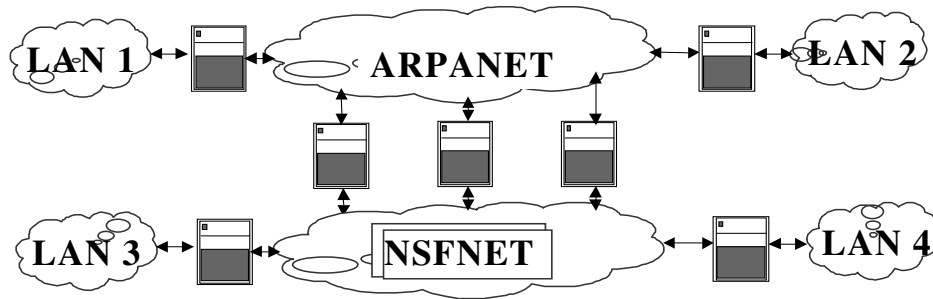
Η κατάσταση έγινε ακόμη πιο περίπλοκη με την πολλαπλή διασύνδεση του ARPANET με το NSFNET μέσω πολλών πυλών οπότε πολλαπλασιάστηκαν οι πιθανές διαδρομές για την προσέγγιση κάθε προορισμού. Αυτή η τοπολογία παρουσιάζεται στο σχήμα 3.5. Το βασικό συμπέρασμα ήταν ότι απαιτείται μία αυτόματη διαδικασία ενημέρωσης των πινάκων δρομολόγησης, ώστε να ενημερώνονται γρήγορα και αποτελεσματικά στην περίπτωση ενός ταχύτητα αναπτυσσόμενου και δυναμικά μεταβαλλόμενου Internet στη θέση της χειροκίνητης ενημέρωσης που χρησιμοποιείτο μέχρι τότε.

Το πρώτο πρωτόκολλο που αναπτύχθηκε για τη διακίνηση πληροφορίας ενημέρωσης των πινάκων δρομολόγησης στην περίπτωση ενός συμπλέγματος δρομολογητών με κεντρική διαχείριση είναι το πρωτόκολλο GGP (Gateway-to-Gateway Protocol) το οποίο βασιζόταν σε μία καταναμημένη μορφή του αλγορίθμου Bellman-Ford.

Το επόμενο πρόβλημα που έγινε αντιληπτό ήταν η δρομολόγηση του μεγαλύτερου μέρους της κίνησης μέσω περισσότερων πυλών απ' ότι ήταν αναγκαίο. Αυτό συνέβαινε, γιατί ένα τοπικό δίκτυο συνδεόταν μόνο μέσω μίας συγκεκριμένης κεντρικής πύλης προς το δίκτυο κορμό, έστω και αν ήταν συνδεδεμένο άμεσα με περισσότερες κεντρικές πύλες, οπότε απαιτούνταν η εκτροπή σε άλλες κεντρικές πύλες όσων πακέτων κατευθύνονταν σε τοπικά δίκτυα, τα οποία δεν ήταν συνδεδεμένα με την συγκεκριμένη κεντρική πύλη. Αυτό θα λύνονταν, εάν τα τοπικά δίκτυα είχαν πληροφορία δρομολόγησης από το κεντρικό δίκτυο και μπορούσαν να επιλέξουν εξ' αρχής τη βέλτιστη δρομολόγηση προς το δίκτυο κορμό. Συνεπώς προέκυψε η ανάγκη πρωτοκόλλου ανταλλαγής πληροφορίας δρομολόγησης μεταξύ κεντρικού και τοπικών δικτύων.

Η αντιμετώπιση αυτής της κατάστασης εξελίχθηκε σε όλο και πιο περίπλοκο πρόβλημα με την ανάπτυξη όλο και συνθετότερων δομών διασυνδεδεμένων δικτύων με το δίκτυο κορμό. Μία κεντρική πύλη μπορούσε να δώσει πρόσβαση σε ένα σύνολο διασυνδεδεμένων μεταξύ τους δικτύων μέσω εσωτερικών πυλών. Η πρόσβαση στα υποδίκτυα του κάθε επί μέρους φορέα έπρεπε να γίνει σε

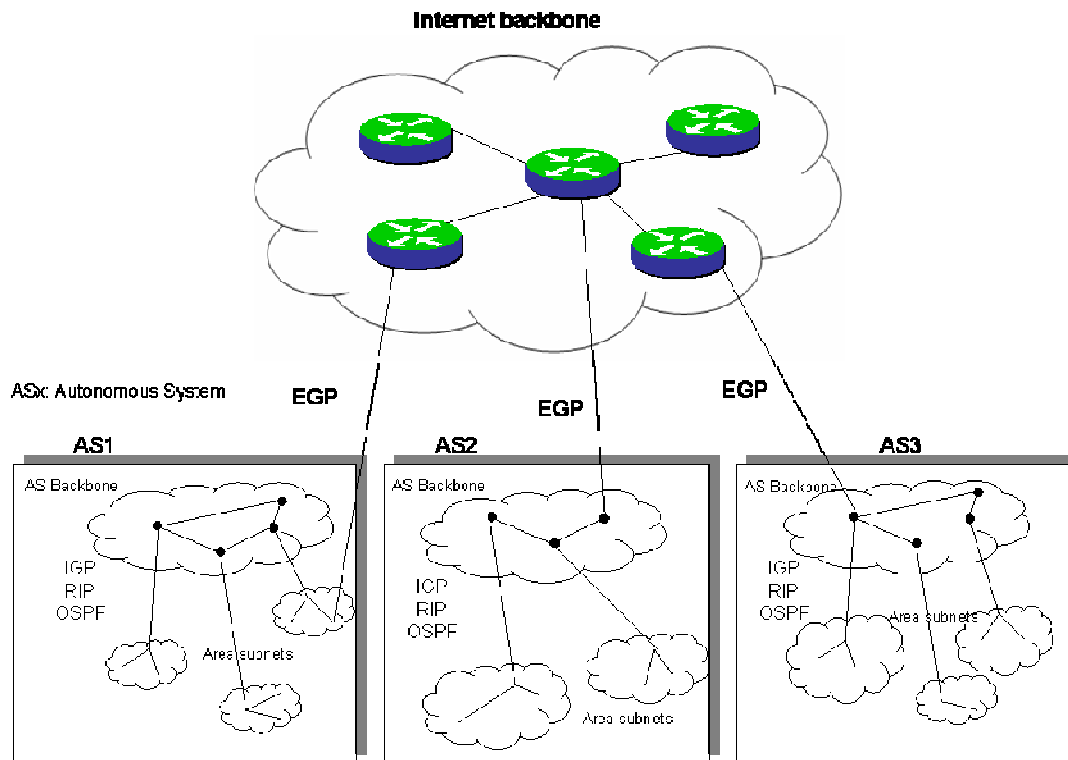
πολλές περιπτώσεις έμμεσα με δρομολόγηση μέσω των τοπικών πυλών και όχι άμεσα από την κεντρική πύλη του δικτύου κορμού.



Σχήμα 3.5. Διασύνδεση ομοτίμων δικτύων κορμού μέσω πολλαπλών πυλών.

Δρομολόγηση εντός αυτονόμων συστημάτων

Για να δοθεί λύση στα προβλήματα της δρομολόγησης που ανέκυψαν έγινε επιτακτική η διάκριση του κεντρικού δικτύου από τα επί μέρους δίκτυα που διασυνδέονταν μέσω αυτού. Η αντιμετώπιση των επί μέρους δικτύων βασίστηκε στην ιδέα των *αυτόνομων συστημάτων (autonomous systems)*, τα οποία κατ' αυτό τον τρόπο θα μπορούσαν να έχουν οποιαδήποτε πολύπλοκη μορφή και ανεξάρτητους εσωτερικούς αλγόριθμους δρομολόγησης, υπό την προϋπόθεση ότι θα υπήρχε μία πύλη, η οποία θα μετέδιδε πληροφορία δρομολόγησης από τα εσωτερικά επί μέρους δίκτυα στην κεντρική πύλη μέσω της οποίας θα γινόταν η πρόσβαση στο δίκτυο κορμού. Η πύλη του τοπικού δικτύου που αναλαμβάνει την επικοινωνία με την κεντρική πύλη αποκαλείται και *εξωτερική πύλη (exterior Gateway)*. Η εξωτερική πύλη πρέπει προφανώς να έχει απ' ευθείας σύνδεση με την κεντρική πύλη.

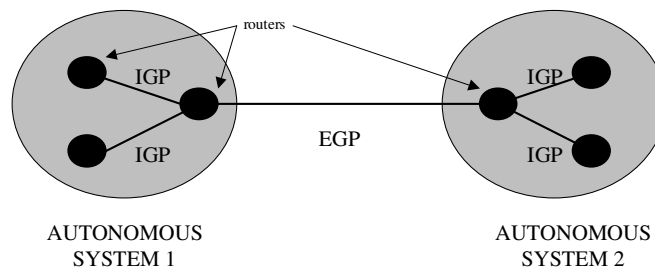


Σχήμα 3.6. Η αρχιτεκτονική των αυτόνομων συστημάτων με πρόσβαση μέσω κεντρικών πυλών σε δίκτυο κορμού και ανταλλαγή πληροφορίας δρομολόγησης μέσω εξωτερικών πυλών.

Η διάκριση αυτή ήταν επιβεβλημένη και από διοικητικούς λόγους καθώς κάθε αυτόνομο σύστημα βρίσκεται υπό τον έλεγχο μιας κοινής αρχής διαχείρισης (δηλ. μιας εταιρείας ή οργανισμού) που θέλει να ελέγχει τις επιλογές μέσα στο δίκτυο ιδιοκτησίας της. Η αρχιτεκτονική των αυτόνομων συστημάτων επιτρέπει να οριοθετούνται σαφώς και με ευελιξία τα όρια κάθε διαχειριστικής αρχής η οποία αποφασίζει τους αλγορίθμους δρομολόγησης, τα σημεία σύνδεσης με το δίκτυο κορμού κτλ. Παραδείγματος χάριν, το δίκτυο της ΟΤΕΝΕΤ και της HELLAS ON LINE αποτελούν αυτόνομα συστήματα. Η φιλοσοφία της οργάνωσης του Internet σε ένα κεντρικό δίκτυο κορμό και επί μέρους αυτόνομα συστήματα απεικονίζεται στο σχήμα 3.6.

Για την δρομολόγηση λοιπόν πακέτων προς υποδίκτυα στο εσωτερικό αυτόνομων συστημάτων η εξωτερική πύλη θα πρέπει να έχει ενημερώσει την κεντρική πύλη για την ύπαρξη αυτών των δικτύων και το δρόμο πρόσβασης μέσω αυτής. Το κεντρικό δίκτυο κορμού μπορεί κατ' αυτήν την έννοια να θεωρηθεί ένα μεγάλο αυτόνομο σύστημα. Επίσης η εξωτερική πύλη οποιουδήποτε αυτόνομου συστήματος μπορεί να ανταλλάσσει πληροφορία δρομολόγησης με την εξωτερική πύλη οποιουδήποτε άλλου αυτόνομου συστήματος με το οποίο βρίσκεται σε σύνδεση. Δύο τέτοια αυτόνομα συστήματα αποκαλούνται *εξωτερικοί γείτονες*. Το αντίστοιχο πρωτόκολλο που χρησιμοποιήθηκε για την ανταλλαγή πληροφορίας δρομολόγησης μεταξύ εξωτερικών πυλών είναι το αποκαλούμενο EGP (Exterior Gateway Protocol).

Πρέπει να τονίσουμε ότι το πρωτόκολλο αυτό δεν αποτελεί αλγόριθμο δρομολόγησης, καθώς πρόκειται πρωτόκολλο ανταλλαγής πληροφοριών ύπαρξης μονοπατιού για πρόσβαση σε κάποιο δίκτυο μέσω κάποιας εξωτερικής πύλης σε ένα αυτόνομο σύστημα. Το πρωτόκολλο δεν αναζητά την εύρεση ενός βέλτιστου μονοπατιού (όπως κάνουν τα RIP, OSPF κτλ.) καθώς δεν χρησιμοποιεί πληροφορία για τις διάφορες εναλλακτικές διαδρομές δρομολόγησης μέσα στο αυτόνομο σύστημα, το οποίο και χρησιμοποιεί εσωτερικά ανεξάρτητα πρωτόκολλα δρομολόγησης όπως το IGP-Interior Gateway Protocol (στη ίδια κατηγορία στην οποία ανήκουν τα RIP και OSPF). Η σχέση μεταξύ των πρωτοκόλλων IGP και EGP και οι επικράτειες εφαρμογής τους φαίνονται στο σχήμα 3.6 αλλά και πιο απλουστευμένα στο σχήμα 3.7.



Σχήμα 3.7. Επικράτειες χρήσεις των IGP και EGP

Εσωτερικά σε ένα αυτόνομο σύστημα όπως είπαμε μπορεί να χρησιμοποιηθούν κατά βούληση διάφοροι αλγόριθμοι δρομολόγησης, αρκεί να ενημερώνεται η κεντρική πύλη του δικτύου κορμού όπως περιγράψαμε. Όλες οι πύλες ενός αυτόνομου συστήματος αποκαλούνται *εσωτερικές μεταξύ τους*. Οι μηχανισμοί ενημέρωσης των πινάκων δρομολόγησης εντός του αυτόνομου συστήματος αφήνονται στο διαχειριστή του συστήματος να επιλέξει και ποικίλλουν ανάλογα με το μέγεθος, τις ανάγκες και τις συνεργασίες που υπάρχουν. Αν και κάποτε μπορούσε να γίνει στατικά και να ενημερώνονται με το χέρι για κάθε προσθήκη ή διαγραφή υπολογιστή στο σύστημα η επιλογή αυτή δεν είναι πια πρακτική. Μηχανισμοί αυτόματης ενημέρωσης των πινάκων δρομολόγησης οδήγησαν όπως ελέχθη στην ανάπτυξη των διαφόρων πρωτοκόλλων εσωτερικής δρομολόγησης σε αυτόνομα συστήματα όπως τα πιο γνωστά IGP (Interior Gateway Protocol), RIP (Routing Information Protocol) και HELLO. Μεταξύ κεντρικών πυλών χρησιμοποιήθηκε το πρωτόκολλο SPREAD και το πρωτόκολλο EGP.

Κεφάλαιο 4

Το Πρωτόκολλο Μηνυμάτων Ελέγχου του Internet (ICMP Internet Control Message Protocol)

Υπάρχουν σ' ένα δίκτυο μια σειρά από απρόβλεπτες καταστάσεις που δεν ανήκουν στην ομαλή εξέλιξη των λειτουργιών του δικτύου. Για τις πιο συνήθεις και σοβαρές απ' αυτές μια σωστή σχεδίαση των πρωτοκόλλων έχει προβλέψει στη μηχανή καταστάσεων τις κατάλληλες ενέργειες ανίχνευσης και επαναφοράς, ώστε να προσδώσει στιβαρότητα στο δίκτυο. Στην κατηγορία αυτή ανήκουν: η διόρθωση λαθών που προκλήθηκαν από παράσιτα, η επαναφορά και η επανεκκίνηση του X.25 κτλ. Πάντα ωστόσο παραμένουν καταστάσεις για τις οποίες δεν είναι ξεκάθαρο ποια πρέπει να είναι η σωστή ενέργεια μιας μηχανής ή το κόστος από την ενσωμάτωση λύσεων αυτόματης ανάνηψης να είναι πολύ μεγάλο. Το IP με την κατανεμημένη αρχιτεκτονική του και την άνευ συνδέσεων υπηρεσία που υποστηρίζει ακολουθεί τη φιλοσοφία ότι σε τέτοιες περιπτώσεις είναι καλύτερο η μηχανή που ανιχνεύει αυτήν την κατάσταση να ειδοποιεί το δίκτυο παρά να συνεχίζει όπως μπορεί. Παρότι η υπηρεσία δεν είναι εγγυημένη και απλώς "κατά δύναμιν" δεν μπορεί η απόρριψη ενός δεδομενογράμματος να γίνεται χωρίς να ειδοποιείται ο αποστολέας και παρά την αυτόνομη λειτουργία της κάθε πύλης στο Διαδίκτυο δεν πρέπει να ξεχνάμε ότι ο στόχος παραμένει να λειτουργεί σωστά το όλο δικτυακό σύστημα.

Οι ανώμαλες καταστάσεις περί των οποίων πρέπει να λαμβάνουν γνώση και οι άλλες μηχανές εκτός αυτής που πρώτη τις ανιχνεύει είναι πολλές. Πέραν των βλαβών των γραμμών επικοινωνίας και επεξεργαστών, το IP αποτυγχάνει να παραδώσει ένα δεδομένογραμμα όταν ο υπολογιστής προορισμού είναι αποκομμένος από το δίκτυο, όταν ο μετρητής χρόνου ζωής (time-to-live) έχει εκπνεύσει, οι πύλες έχουν υπερφορτωθεί με κίνηση ώστε δεν μπορούν να προωθήσουν την εισερχόμενη πληροφορία κτλ.

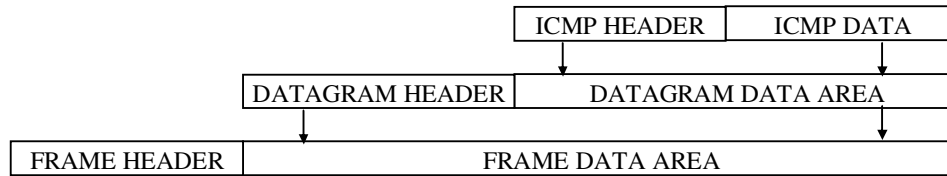
Για την αναφορά λαθών ή την παροχή πληροφορίας σχετικά με τέτοιες αναπάντεχες καταστάσεις, το IP διαθέτει το πρωτόκολλο Internet Control Message Protocol (ICMP) που επιτρέπει στις πύλες να στέλνουν μηνύματα λάθους ή ελέγχου σε άλλες πύλες ή ξενιστές. Ανάλογα με το μήνυμα προορισμός που θα αναλάβει να κάνει κάποια ενέργεια μπορεί να είναι το IP, ή το TCP ή ακόμη και το επίπεδο εφαρμογής, όπως επί παραδείγματι το μήνυμα destination port unreachable φθάνει στον φυλλομετρητή και προκαλεί το γνωστό μήνυμα που όλοι έχουμε ενίοτε συναντήσει ότι δεν μπορεί να βρεθεί η πηγή της πληροφορίας που ζητήσαμε. Το ICMP παρέχει επικοινωνία μεταξύ των λογισμικών πρωτοκόλλων δύο μηχανών. Βοηθά επίσης τους χειριστές στην εκσφαλμάτωση βλαβών δεδομένου ότι το ίδιο το IP δεν έχει τρόπο να καταλάβει εάν μία κακή λειτουργία οφείλεται σε τοπική ή μακρινή βλάβη υλικού αφού το ίδιο συνίσταται εξ' ολοκλήρου από λογισμικό.

Τα μηνύματα του ICMP χρησιμοποιούν και αυτά το τμήμα δεδομένων του δεδομενογράμματος. Η διαφορά είναι ότι ο προορισμός τους δεν είναι ένα πρόγραμμα εφαρμογής όπως στην κοινή περίπτωση αλλά το λογισμικό του IP της μηχανής σε μηχανή. Φυσικά εάν απαιτείται ενημερώνεται και η εφαρμογή δηλ. ο χρήστης. Έτσι, το ICMP λειτουργεί σαν μηχανισμός πληροφόρησης του αποστολέα περί σφαλμάτων που κατέστησαν αδύνατη την σωστή παράδοση των δεδομένων στον προορισμό. Τυπικά μηνύματα ICMP είναι: destination unreachable, unknown host, unknown network, change route, time-to-live of datagram was exceeded, fragmentation needed but DF set, κτλ.

Η εφαρμογή του αποστολέα πρέπει να ενεργήσει για τη διόρθωσή τους αφού πολλά λάθη προέρχονται από τον αποστολέα. Όμως δεν φταίει πάντα ο αποστολέας, μπορεί κάποια ενδιάμεση πύλη να έχει λανθασμένους πίνακες. Δυστυχώς το ICMP μπορεί να αναφέρει προβλήματα μόνο πίσω στην αρχική πηγή και δεν μπορεί να χρησιμοποιηθεί για να ενημερώσει άμεσα τις ενδιάμεσες πύλες σχετικά με το πρόβλημα Αυτό διότι τα δεδομενογράμματα έχουν μόνο πεδία διευθύνσεων αποστολής και προορισμού και όχι για τους ενδιάμεσες πύλες που δεν είναι πάντα οι ίδιες αφού δεν χρησιμοποιούνται σταθερά δρομολόγια. Αν κάποια πύλη ανιχνεύσει ένα πρόβλημα, δεν μπορεί να γνωρίζει το σύνολο των ενδιάμεσων μηχανών της διαδρομής και έτσι δεν μπορεί να τις ενημερώσει για το σφάλμα. Απλά χρησιμοποιεί το ICMP για να ενημερώσει την αρχική πηγή για το πρόβλημα.

Το ICMP απαιτεί δύο επίπεδα ενθουλάκωσης όπως φαίνεται στο παρακάτω σχήμα 4.1. Τα δεδομενογράμματα που μεταφέρουν τα μηνύματα ICMP, δρομολογούνται ακριβώς όπως τα δεδομενογράμματα που μεταφέρουν πληροφορία για τους χρήστες. Εν τούτοις το ICMP δεν θεωρείται πρωτόκολλο ανωτέρου επιπέδου αλλά μέρος του IP. Για να αναγνωρίζονται φέρουν στο

πεδίο πρωτοκόλλου του δεδομενογράμματος την τιμή 1. Τα μηνύματα ICMP μπορούν και αυτά να χαθούν ή και να απορριφθούν όπως κάθε δεδομένογραμμα IP. Επιπλέον σε ένα ήδη συμφορημένο δίκτυο το μήνυμα σφάλματος μπορεί να δημιουργήσει επιπλέον συμφόρηση με πληθώρα μηνυμάτων ICMP. Γι' αυτό γίνεται εξαίρεση στις διαδικασίες διαχείρισης δικτύου και δεν αποστέλλονται μηνύματα ICMP αναφορικά με λάθη που προέκυψαν από ένα IP δεδομένογραμμα που μεταφέρει ένα μήνυμα ICMP. Η εξαίρεση γίνεται για να αποφευχθεί ο φαύλος κύκλος μηνυμάτων σφάλματος.



Σχήμα 4.1

Αν και κάθε μήνυμα ICMP έχει τη δική του φόρμα, όλα αρχίζουν με τα παρακάτω τρία πεδία: Ένα πεδίο 8-bit με τον τύπο (TYPE) μηνύματος το οποίο είναι η ταυτότητα του μηνύματος, ένα πεδίο 8-bit CODE που παρέχει επιπλέον πληροφορία σχετικά με το τύπο του μηνύματος και ένα πεδίο 16-bit CHECKSUM. Επιπλέον τα μηνύματα ICMP που αναφέρουν τα σφάλματα που προκύπτουν, περιλαμβάνουν την επικεφαλίδα και τα πρώτα 64 bits των δεδομένων του δεδομενογράμματος που προκαλεί το πρόβλημα. Όλα αυτά δίδονται για να βοηθήσουν στον προσδιορισμό του σφάλματος.

Μερικοί κωδικοί μηνυμάτων ICMP είναι και οι εξής: 0 echo the reply
 3 destination unreachable
 4 source quench
 5 redirect (change route)
 11 time-to-live of datagram was exceeded
 κτλ.

Ένα από τα εργαλεία που προσφέρει το ICMP στους διαχειριστές δικτύου για εντοπισμό προβλημάτων είναι η αίτηση ηχούς (echo request) και απόκριση ηχούς (echo reply) σε ένα καθορισμένο προορισμό. Κάθε υπολογιστής που λαμβάνει μία αίτηση ηχώ ηχούς δημιουργεί μια απόκριση ηχούς και την επιστρέφει στον αρχικό αποστολέα. Η αίτηση περιέχει μία προαιρετική περιοχή δεδομένων. Η απόκριση περιέχει ένα αντίγραφο των δεδομένων που στέλνονται στην αίτηση. Η αίτηση ηχούς χρησιμοποιείται για να προσδιοριστεί η προσπελασιμότητα ενός προορισμού. Έτσι αφού η αίτηση και η απόκριση μεταφέρονται με δεδομένογραμμα IP, η επιτυχής παραλαβή της απόκρισης επιβεβαιώνει ότι τα κυριότερα μέρη του συστήματος μετάδοσης λειτουργούν. Αυτό εκμεταλλεύεται η απλή διαγνωστική εφαρμογή **ping**.

Η αποστολή του μηνύματος "time-to-live of datagram was exceeded" όταν μηδενιστεί το πεδίο TTL χρησιμοποιείται από την μικρή διαγνωστική εφαρμογή **trace route** η οποία στέλνει διαδοχικά πακέτα με TTL ίσο με 1, 2, 3, 4 κτλ. έτσι ώστε να αναγκάσει τον πρώτο, δεύτερο, τρίτο κτλ δρομολογητή να στείλουν πίσω τα αντίστοιχα μηνύματα αποκαλύπτοντας την ταυτότητά τους.

Ως γνωστόν, όταν μία πύλη δεν προλαβαίνει να επεξεργαστεί τα δεδομένογραμμα επειδή φθάνουν πολύ γρήγορα, εξαντλείται η μνήμη του ταμιευτήρα και απορρίπτονται. Το μήνυμα ICMP **source quench** (καταστολή πηγής) είχε σχεδιαστεί σαν μέτρο αντιμετώπισης συμφόρησης. Είναι μία αίτηση προς τον ξενιστή προέλευσης του πακέτου ώστε να μειώσει τον τρέχοντα ρυθμό μετάδοσης. Αυτό θα μεταφέρει το πρόβλημα στον προηγούμενο κόμβο που θα αναγκαστεί να στείλει παρόμοιο μήνυμα πτό πίσω και έτσι θα καταλήξουμε σε μείωση του ρυθμού σε όλο το μήκος της διαδρομής.

Ωστόσο στην πράξη αποδείχθηκε ότι ο μηχανισμός αυτός δεν ήταν και τόσο αποτελεσματικός. Καλύτερα είναι να ειδοποιείται απ' ευθείας η πηγή να μειώσει το ρυθμό της και έτσι μεταγενέστερα υλοποιήσεις μετά το 1993 κατήργησαν την απαίτηση να στέλνεται τέτοιο μήνυμα. Ο έλεγχος συμφόρησης γίνεται πολύ πιο αποτελεσματικά στο επίπεδο του στρώματος μεταφοράς μέσω μηχανισμού επιβεβαιώσεων δημιουργώντας κλειστό βρόχο ελέγχου και αυτό ακριβώς γίνεται στο TCP το οποίο και διαθέτει αποτελεσματικό μηχανισμό ελέγχου συμφόρησης που θα μελετηθεί σε επόμενο κεφάλαιο.