

13 Πώς η κίνηση διασχίζει το Διαδίκτυο;

Έχουμε αναφέρει το Διαδίκτυο πολλές φορές ως τώρα, και όλα τα προηγούμενα κεφάλαια βασίζονται στην ύπαρξή του. Είναι ώρα να μπούμε στην αρχιτεκτονική του Διαδικτύου, ξεκινώντας σε αυτά τα δύο κεφάλαια με τη θεμελίωση του Διαδικτύου πάνω στο TCP/IP.

13.1 Μία σύντομη απάντηση

Θα πλοηγηθούμε σε πολλές βασικές έννοιες γύρω από την εξέλιξη του Διαδικτύου, παρέχοντας τα θεμέλια για τα τέσσερα επόμενα κεφάλαια. Έτσι η ενότητα «μία σύντομη απάντηση» θα είναι μεγαλύτερη από την ενότητα «μία εκτενής απάντηση» σε αυτό το κεφάλαιο.

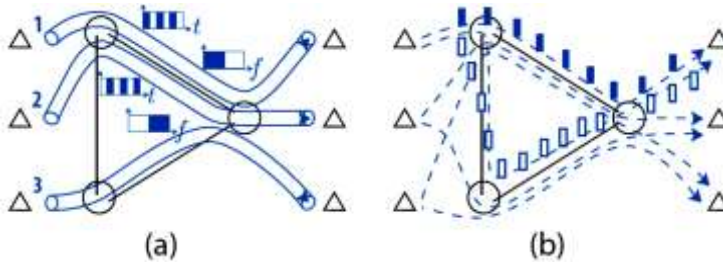
Είναι παρακινδυνευμένο να συζητηθεί η ιστορική εξέλιξη των τεχνολογιών, όπως το Διαδίκτυο. Μερικά στοιχεία τα οποία θα θέλαμε να πιστέψουμε ότι είναι αναπόφευκτα αποτελέσματα ενός προσεκτικού σχεδιασμού, στην πραγματικότητα είναι είτε ιστορική κληρονομιά κάποιων ατυχημάτων ή των αδόμετων απαιτήσεων για οπισθόδρομη συμβατότητα, της ανάγκης για βαθμιαία ανάπτυξη και οικονομικών κινήτρων. Επομένως, δεν είναι εύκολο να συμφωνήσουμε για το τι θα είχε συμβεί, ποια θα ήταν τα εναλλακτικά μονοπάτια στην εξέλιξη και ποιοι συμβιβασμοί θα μπορούσαν να υπήρχαν σε άλλη περίπτωση.

13.1.1 Μεταγωγή πακέτων

Η απάντηση στην ερώτηση αυτού του κεφαλαίου ξεκινά με τη βασική ιδέα της σχεδίασης ενός δικτύου: όταν οι τυπικοί σας χρήστες δεν απαιτούν πραγματικά κάποιο «αφοσιωμένο πόρο», πρέπει να επιτρέπετε στους χρήστες το διαμοιρασμό πόρων. Η λέξη «χρήστης» εδώ χρησιμοποιείται εναλλάξ με τη «σύνοδο» ή τη «συνεδρία». Η λογική μονάδα είναι μία σύνοδος επιπέδου εφαρμογής αντί ένας φυσικός χρήστης ή μία συσκευή. Για την ώρα, υποθέστε ότι η σύνοδος έχει μόνο μία πηγή και έναν προορισμό, δηλαδή είναι μία σύνοδος μονο-εκπομπής (unicast).

Στην περίπτωση της δρομολόγησης, ο πόρος βρίσκεται κατά μήκος ολόκληρης της διαδρομής από το ένα άκρο μίας συνόδου επικοινωνίας, στον αποστολέα, στο άλλο

άκρο, στο δέκτη. Μπορούμε να αφιερώσουμε είτε ένα σταθερό ποσοστό των πόρων κατά μήκος της διαδρομής σε κάθε σύνοδο ή μπορούμε να αναμείξουμε και να ταιριάξουμε πακέτα από διαφορετικές συνόδους και επίσης να μοιράσουμε όλες τις διαδρομές. Αυτή είναι η διαφορά μεταξύ δικτύων μεταγωγής κυκλώματος και μεταγωγής πακέτων.



Σχήμα 13.1 Ένα απλό δίκτυο με τρεις διασυνδεδεμένους δρομολογητές και τρεις συνόδους. (a) Μεταγωγή κυκλώματος: κάθε σύνοδος καταλαμβάνει ένα αφοσιωμένο κύκλωμα, είτε μία μερίδα από κάθε χρονοθυρίδα t ή ένα μέρος της ζώνης συχνότητας f , ακόμα και όταν δεν χρησιμοποιείται. (b) Μεταγωγή πακέτου: κάθε σύνοδος στέλνει πακέτα κατά μήκος μίας ή περισσότερων διαδρομών (όταν υπάρχουν πακέτα προς αποστολή) και όλες οι διαδρομές μοιράζονται μεταξύ χρονοθυρίδων και ζωνών συχνότητων.

Πριν τη δεκαετία του 1960, δικτύωση ήταν η σύνδεση τηλεφωνικών κλήσεων στα τύπου μεταγωγής κυκλώματος Δημόσια Τηλεφωνικά Δίκτυα Μεταγωγής (Public Switched Telephone Networks – PSTN). Στον τομέα αυτό, συνέχισε να υπάρχει ενεργή έρευνα μέχρι τις αρχές της δεκαετίας του 2000, συμπεριλαμβανόμενης της έρευνας γύρω από τη δυναμική δρομολόγηση, όπως θα δείτε σε μία άσκηση.

Η επανάσταση, η οποία έγινε γνωστή ως το Διαδίκτυο, ξεκίνησε κατά τη διάρκεια των δεκαετιών 1960 και 1970, με μία μετατόπιση στη μεταγωγή πακέτου ως το θεμελιώδες πρότυπο της δικτύωσης. Στις αρχές της δεκαετίας του 1960, οι ερευνητές ανέπτυξαν επίσημα την ιδέα του τεμαχισμού των μηνυμάτων μίας συνόδου σε μικρά πακέτα και της αποστολής τους με διαφορετικές πιθανές διαδρομές, με κάθε διαδρομή να μοιράζεται με άλλες συνόδους. Η εικόνα 13.1 αντιπαραβάλλει τη μεταγωγή κυκλώματος με τη μεταγωγή πακέτων. Κάθε κύκλωμα στη μεταγωγή κυκλώματος μπορεί να καταλάβει είτε μία συγκεκριμένη ζώνη συχνότητας είτε μία αφοσιωμένη μερίδα από χρονοθυρίδες. Σε αντίθεση, στη μεταγωγή πακέτων, δεν υπάρχει κανένα αφοσιωμένο κύκλωμα για κάθε σύνοδο. Όλες οι σύνοδοι επιτρέπουν στα πακέτα τους να μοιράζονται τις διαδρομές.

Το 1969, υποστηριζόμενα από την αμερικάνικη Υπηρεσία Προηγμένων Αμυντικών Ερευνών (Advanced Research Project Agency – ARPA), το UCLA και τρία άλλα ιδρύματα έφτιαξαν το πρώτο πρωτότυπο ενός δικτύου μεταγωγής πακέτων, το οποίο έγινε γνωστό ως ARPANET. Το ARPANET ξεκίνησε να αναπτύσσεται. Το 1974,

οι Cerf και Kahn ανέπτυξαν ένα πρωτόκολλο, δηλαδή ένα σύνολο κανόνων για την επικοινωνία μεταξύ των συσκευών, για τα δίκτυα μεταγωγής πακέτων, αποκαλούμενο Πρωτόκολλο Ελέγχου Μεταφοράς/Πρωτόκολλο Διαδικτύου (Transmission Control Protocol/Internet Protocol – TCP/IP). Αυτό το πρωτόκολλο επέτρεψε την κλιμακούμενη διασυνδεσιμότητα στο ARPANET. Από το 1985 ως το 1995 το αμερικανικό Εθνικό Ίδρυμα Επιστήμης (National Science Foundation – NSF) ανέλαβε την επόμενη φάση της ανάπτυξης, υποστηρίζοντας τη δημιουργία και λειτουργία ενός συνεχώς αυξανόμενου δικτύου των δικτύων, αποκαλούμενο NSFNET. Ξεκινώντας στις αρχές της δεκαετίας του '90, εμπορικά συμφέροντα και επιχειρηματικές δραστηριότητες επέκτειναν αυτό το διασυνδεδεμένο δίκτυο των δικτύων. Πράγματι, ως το 1994 ο Παγκόσμιος Ιστός και το περιβάλλον χρήσης των περιηγητών του ιστού είχαν ωριμάσει και ο κόσμος κινήθηκε γρήγορα στις επαγγελματικές εφαρμογές που κτίστηκαν πάνω σε αυτό το δίκτυο, γνωστό μέχρι τότε ως Διαδίκτυο. Σήμερα, το Διαδίκτυο έχει ανθίσει σε ένα βασικό μέρος του πώς οι άνθρωποι ζουν, παίζουν, μιλούν και σκέφτονται. Υπάρχουν τώρα περισσότερες συσκευές συνδεδεμένες στο Διαδίκτυο από τους ανθρώπους στον κόσμο και προβλέπεται ότι έως το 2020 θα υπάρχουν έξι φορές περισσότερες συνδεδεμένες συσκευές από ότι άνθρωποι. Είναι πραγματικά εκπληκτικό για μία τεχνολογική ανάπτυξη πέντε δεκαετιών.

Η συζήτηση μεταξύ της κατανομής αφοσιωμένων πόρων και της κατανομής διαμοιραζόμενων πόρων είναι μακρά και σε βάθος. Εκτός από τη σύγκριση μεταξύ μεταγωγής κυκλώματος και της μεταγωγής πακέτων εδώ, και της ορθογωνίας με τη μη-ορθογώνια κατανομή πόρων στο Κεφάλαιο 1, θα δούμε επίσης τρεις ακόμα ειδικές περιπτώσεις αυτής της επιλογής σχεδιασμού: πελάτη-εξυπηρετητή σε σχέση μεταξύ ομοτίμων (peer-to-peer), τοπική αποθήκευση σε σχέση με υπηρεσίες νέφους και χρονοδρομολόγηση χωρίς ανταγωνισμό (contention-free scheduling) σε σχέση με τυχαία προσπέλαση στα Κεφάλαια 15, 16 και 18 αντίστοιχα.

Σε γενικές γραμμές, υπάρχει ένα μεγάλο πλεονέκτημα της μεταγωγής πακέτου, ή της κατανομής αφοσιωμένων πόρων: η εγγύηση της ποιότητας. Δεδομένου ότι σε κάθε σύνοδο παρέχεται ένα κύκλωμα αφοσιωμένο σε αυτή, η ικανότητα διαβίβασης και η απόδοση καθυστέρησης είναι εγγυημένη και υπάρχει πολύ μικρό τρεμόπαιγμα (jitter) (οι διαφορές της καθυστέρησης). Σε αντίθεση, στη μεταγωγή πακέτων, η κυκλοφορία μίας συνόδου (ενδεχομένως) επιμερίζεται σε διαφορετικές διαδρομές, κάθε μία από τις οποίες μοιράζεται με άλλες συνόδους. Τα πακέτα φτάνουν εκτός σειράς και πρέπει να ανασυνταχθούν στον παραλήπτη. Οι ζεύξεις μπορεί να κορεστούν. Η ικανότητα διαβίβασης δεδομένων και οι επιδόσεις καθυστέρησης γίνονται αβέβαιες. Οι ερευνητές του διαδικτύου αυτό το ονομάζουν ως την υπηρεσία καλύτερης προσπάθειας (best-effort) που προσφέρει το Διαδίκτυο, το οποίο περιγράφεται ακριβέστερα ως υπηρεσία *καμιάς προσπάθειας* για να υπάρχει εγγυημένη επίδοση.

Από την άλλη πλευρά, όμως, υπάρχουν δύο μεγάλα πλεονεκτήματα της μεταγωγής πακέτων: (1) η *ευκολία της συνδεσιμότητας* και (2) η *κλιμάκωση λόγω της αποδοτικότητας*.

(1) Η ευκολία της συνδεσιμότητας είναι προφανής: δεν υπάρχει καμία ανάγκη να αναζητηθεί, καθιερωθεί, διατηρηθεί και τελικά να καταστραφεί ένα από άκρο-σε-άκρο κύκλωμα για κάθε σύνοδο.

(2) Η κλιμάκωση εδώ αναφέρεται στην ικανότητα υποστήριξης όσο το δυνατόν περισσότερων διαφορετικών τύπων συνεδριών, μερικές από τις οποίες είναι μακράς διάρκειας, ενώ άλλες είναι βραχείας. Υπάρχουν δύο υποκείμενοι λόγοι για την αποδοτικότητα της μεταγωγής πακέτων, οι οποίοι με τη σειρά τους οδηγούν στην υψηλή κλιμάκωση. Αυτοί οι δύο λόγοι αντιστοιχούν στο χαρακτηριστικό της μεταγωγής πακέτων «πολλές σύνοδοι μοιράζονται μία διαδρομή» και στο χαρακτηριστικό «κάθε σύνοδος μπορεί να χρησιμοποιήσει πολλαπλές διαδρομές». Ονομάζουμε αυτά τα δύο χαρακτηριστικά στατιστική πολυπλεξία και συγκέντρωση των πόρων.

- **Στατιστική πολυπλεξία:** η μεταγωγή πακέτου μπορεί ελαστικά να αντιστοιχίσει την απαίτηση της δυναμικότητας σε προσφορά χωρητικότητας. Αυτό ταιριάζει στα δυναμικά, κατόπιν παραγγελίας, σενάρια με εκρηκτική κυκλοφορία. Ειδικότερα, όταν μία πηγή είναι ανενεργή και δεν στέλνει καθόλου κυκλοφορία στο δίκτυο, δεν καταλαμβάνει καθόλου πόρους.
- **Συγκέντρωση των πόρων:** αυτό χρειάζεται λίγα μαθηματικά για να το επιδείξουμε, όπως θα δούμε σε μία άσκηση. Αλλά η βασική ιδέα είναι απλή: αντί να έχουμε δύο σύνολα από πόρους (π.χ., τις χωρητικότητες δύο ζεύξεων) σε απομόνωση, βάζοντάς τες σε μία ενιαία μειώνεται η πιθανότητα κάποια απαίτηση να πρέπει απορριφθεί, διότι ένα σύνολο πόρων χρησιμοποιείται πλήρως.

Τελικά, οι δυνατότητες να παρασχεθεί εύκολα συνδεσιμότητα και να κλιμακώνεται με πολλούς διαφορετικούς χρήστες επικράτησαν, αν και αυτό δεν ήταν ξεκάθαρο μέχρι τις αρχές της δεκαετίας του 2000. Σε αντίθεση με την εγγύηση ποιότητας, που είναι βέβαια καλό να έχουμε, αυτές οι δυνατότητες είναι ουσιαστικό να υπάρχουν για ένα δυναμικό και μεγάλο δίκτυο, όπως το Διαδίκτυο. Μόλις διευρυνθεί το δίκτυο με έναν απλό και κλιμακούμενο τρόπο, μπορούμε να ψάξουμε για άλλες λύσεις, ώστε να φροντίσουμε την ποιοτική παραλλαγή. Αλλά πρέπει πρώτα να διευρύνουμε το δίκτυο, από την άποψη του αριθμού των χρηστών και των τύπων των εφαρμογών. Αυτός είναι ο βασικός λόγος, γιατί το IP κατέκτησε τη βιομηχανία δικτύωσης και η μεταγωγή πακέτων επικράτησε, παρά τα εναλλακτικά σχέδια στο χώρο των πρωτοκόλλων (που δεν θα καλύψουμε εδώ), όπως τα X.25, ATM, αναμετάδοσης πλαισίου (frame relay), ISDN, κτλ.



Σχήμα 13.2 Μία αρθρωτή προσέγγιση στη δικτύωση: Ένα τυπικό μοντέλο μίας στοίβας διαστρωματωμένων πρωτοκόλλων. Κάθε επίπεδο (στρώμα) είναι υπεύθυνο για ένα συγκεκριμένο σύνολο εργασιών, χρησιμοποιώντας την υπηρεσία που παρέχεται από το επίπεδο που βρίσκεται από κάτω και με τη σειρά του παρέχοντας μία υπηρεσία στο επίπεδο που βρίσκεται από πάνω. Οι οριζόντιες γραμμές που χωρίζουν τα επίπεδα αντιπροσωπεύουν κάποιου είδους περιορισμό από το τι το κάθε επίπεδο μπορεί να δει και να κάνει. Με τα χρόνια, οι εφαρμογές έχουν εξελιχθεί από τη μεταφορά αρχείων βασισμένη σε εντολές γραμμής εντολών σε όλες αυτές τις εφαρμογές που ζούμε σήμερα. Οι τεχνολογίες του φυσικού επιπέδου και του επιπέδου συνδέσεων έχουν εξελιχθεί από το μόντεμ των 32 kbps στις οπτικές ίνες των 10 Gbps και το ασύρματο δίκτυο των 100 Mbps. Τα δύο μεσαία επίπεδα, ωστόσο, κυριαρχημένα από το TCP/IP, παρέμειναν κατά ένα μεγάλο μέρος αμετάβλητα με τα χρόνια. Τα δύο ενδιαμέσια επίπεδα είναι η «λεπτή μέση» του μοντέλου της στοίβας πρωτοκόλλων του τύπου «κλεψύδρα».

13.1.2 Διαστρωμένη αρχιτεκτονική

Η διαχείριση ενός δικτύου μεταγωγής πακέτου είναι περίπλοκη. Υπάρχουν πολλές εμπλεκόμενες εργασίες και ο έλεγχος κάθε εργασίας απαιτεί μία ακολουθία επικοινωνίας και υπολογισμού, που καλείται πρωτόκολλο για τον έλεγχο αυτό. Είναι μία φυσική πρακτική όταν σχεδιάζουμε ένα τόσο σύνθετο σύστημα να το σπάμε σε μικρότερα κομμάτια. Αυτή η διαδικασία της αρθρωτής προσέγγισης δημιουργείται από τη στοίβα διαστρωματωμένων πρωτοκόλλων για το Διαδίκτυο. Η ιδέα της αρθρωτής προσέγγισης του σχεδιασμού δεν παρακινείται από την αποδοτικότητα της

κατανομής των πόρων, αλλά από την οικονομική βιωσιμότητα μέσω των επιχειρησιακών μοντέλων διαφορετικών εταιριών, που ειδικεύονται σε διαφορετικά επίπεδα, και από την ανθεκτικότητα στις απρόβλεπτες καινοτομίες, που μπορεί να προκύψουν στο Διαδίκτυο. Αυτή η εξελισιμότητα ενισχύεται περαιτέρω από τα δίκτυα επικάλυψης, που μπορούν να δημιουργήσουν νέες τοπολογίες δικτύων και νέες λειτουργίες πάνω από τη συνδεσιμότητα του Διαδικτύου, όπως θα δούμε στο Κεφάλαιο 15.

Μία τυπική στοίβα επιπέδων πρωτοκόλλων φαίνεται στην Εικόνα 13.2. Το TCP/IP τοποθετείται ακριβώς στη μέση της, στα επίπεδα μεταφοράς και δικτύου. Πέρα από τη σύντομη διάρκεια της εξέλιξης του Διαδικτύου, η ταχύτητα μετάδοσης των φυσικών μέσων αυξήθηκε περισσότερες από 30.000 φορές και οι εφαρμογές έχουν εξελιχθεί από τη μεταφορά αρχείου βασισμένη σε εντολές γραμμών στο Netflix και στο Twitter. Ωστόσο, το ίδιο το Διαδίκτυο συνέχισε να λειτουργεί, κυρίως χάρη στη «λεπτή μέση» του TCP/IP, η οποία παρέμεινε σχεδόν η ίδια, καθώς οι εφαρμογές και τα μέσα επικοινωνίας συνέχιζαν να αλλάζουν.

Κάθε επίπεδο παρέχει μία υπηρεσία στο παραπάνω επίπεδο και χρησιμοποιεί μία υπηρεσία από το παρακάτω επίπεδο. Για παράδειγμα, το επίπεδο μεταφοράς παρέχει μία από-άκρο-προς-άκρο (end-to-end) σύνδεση, που υλοποιεί τις υπηρεσίες της δημιουργίας μίας συνόδου, της αναδιάταξης πακέτων και του ελέγχου συμφόρησης, στο επίπεδο εφαρμογών από πάνω του, που υλοποιεί εφαρμογές, όπως ο ιστός, το ηλεκτρονικό ταχυδρομείο και η κοινή χρήση περιεχομένου. Με τη σειρά του, το επίπεδο μεταφοράς παίρνει την υπηρεσία από το επίπεδο δικτύου από κάτω του, συμπεριλαμβανομένων των συνδεσιμοτήτων που δημιουργήθηκαν μέσω δρομολόγησης. Το επίπεδο ζεύξης είναι υπεύθυνο για τον έλεγχο της πρόσβασης στο μέσο επικοινωνίας και το φυσικό επίπεδο ελέγχει την πραγματική μετάδοση πληροφοριών στο φυσικό μέσο.

Υπάρχουν λειτουργικές επικαλύψεις στα επίπεδα. Για παράδειγμα, η λειτουργικότητα του ελέγχου σφαλμάτων κατανέμεται σε αρκετά επίπεδα: υπάρχει κωδικοποίηση ελέγχου σφαλμάτων στο φυσικό επίπεδο, αναμετάδοση κατά άλματα (hop-by-hop) στο επίπεδο σύνδεσης, δρομολόγηση πολλαπλών διαδρομών για την αξιοπιστία στο επίπεδο δικτύου και έλεγχος σφαλμάτων από άκρη σε άκρη στο επίπεδο μεταφοράς. Ο λειτουργικός πλεονασμός δεν είναι ατέλεια, υπάρχει εκεί από σχεδιασμό, πληρώνοντας το τίμημα της μείωσης της αποδοτικότητας για την ανθεκτικότητα και τα σαφή όρια των επιπέδων.

Πώς θα έπρεπε να διαθέσουμε τις λειτουργίες μεταξύ των επιπέδων και να τις βάλουμε μαζί στη σωστή διεπαφή και στο σωστό χρονοδιάγραμμα; Αυτή είναι η ερώτηση της αρχιτεκτονικής δικτύων που θα συνεχίσουμε να εξερευνούμε σε επόμενα κεφάλαια. Για παράδειγμα, οι οριζόντιες γραμμές στην Εικόνα 13.2, που δείχνουν τα όρια μεταξύ των επιπέδων των πρωτοκόλλων, είναι στην πραγματικότητα πολύ περίπλοκα αντικείμενα. Αντιπροσωπεύουν περιορισμούς ως προς τι μπορεί να κάνει

και να δει το κάθε επίπεδο. Στο επόμενο κεφάλαιο, θα πάρουμε μία γεύση από κάποιες μεθοδολογίες, ώστε να καταλάβουμε αυτή την αρχιτεκτονική απόφαση του «ποιος κάνει τι» και «πώς να κολλήσουμε τις ενότητες μαζί».

Μεταξύ των επιπέδων μεταφοράς και δικτύου, έχουν ήδη ληφθεί κάποιες αρχιτεκτονικές αποφάσεις στο TCP/IP, το οποίο αποτελεί την κυρίαρχη ειδική περίπτωση των πρωτοκόλλων των επιπέδων 4/3. Κατ' αρχάς, το επίπεδο μεταφοράς, υπεύθυνο για την από άκρο-προς-άκρο διαχείριση, είναι προσανατολισμένο προς τη σύνδεση (connection-oriented) στο TCP, ενώ το επίπεδο δικτύου, υπεύθυνο για τη διαχείριση της από άκρο-προς-άκρο συνδεσιμότητας, τρέχει κατά άλματα χωρίς σύνδεση (connectionless) δρομολόγηση στο IP. Σαν αναλογία, η τηλεφωνική κλήση σε κάποιον απαιτεί πρώτα να πραγματοποιηθεί μία προσανατολισμένη σε σύνδεση σύνοδος ανάμεσα στον τηλεφωνητή και τον καλούμενο. Σε αντίθεση, η αποστολή ενός μηνύματος ηλεκτρονικού ταχυδρομείου σε κάποιον χρειάζεται μόνο μία χωρίς σύνδεση σύνοδος, καθώς ο παραλήπτης δεν χρειάζεται να γνωρίζει ότι εισέρχεται μία σύνοδος. Η επιλογή της σχεδίασης του προσανατολισμένου σε σύνδεση TCP και του χωρίς σύνδεση IP, ακολουθεί μία «άκρο-προς-άκρο» αρχή, όπου οι τελικοί κόμβοι (end-host) είναι ευφρείς και το δίκτυο είναι «χαζό». Η πραγματοποίηση της συνδεσιμότητας θα πρέπει να υλοποιείται εξ' ολοκλήρου με μεταγωγή πακέτων στο επίπεδο δικτύου και η άκρο-προς-άκρο ανατροφοδότηση να τρέχει από το πάνω από αυτό επίπεδο. Αλλά αυτή η επιλογή της σχεδίασης δεν ήταν η μοναδική που δοκίμασε το Διαδίκτυο αυτές τις δεκαετίες της εξέλιξής του, π.χ. ένα χωρίς σύνδεση επίπεδο μεταφοράς πάνω από ένα προσανατολισμένο σε σύνδεση επίπεδο δικτύου είναι επίσης δυνατό και, πράγματι, χρησιμοποιήθηκε κάποτε.

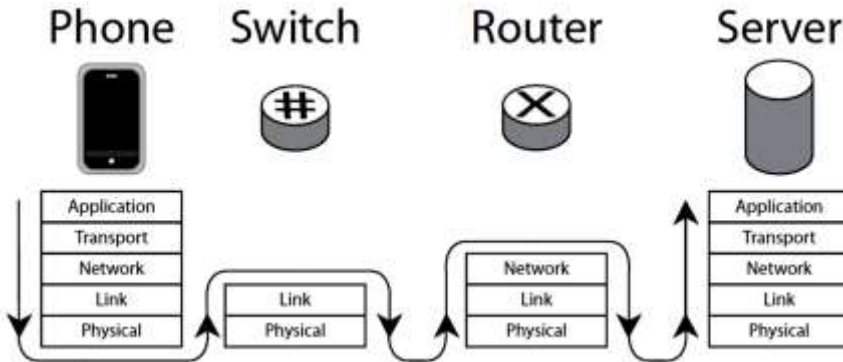
Δεύτερον, η δρομολόγηση στο IP είναι ανεξάρτητη από τις συνθήκες φόρτου στις συνδέσεις, λαμβάνοντας υπόψη ότι ο έλεγχος συμφόρησης στο TCP φέρει εις πέρας τη διαχείριση της ζήτησης στους τελικούς κόμβους, ανταποκρινόμενος στα φορτία των συνδέσεων. Επιπρόσθετα, στην αρχή του από-άκρο-προς-άκρο, η στρατηγική αυτή υποθέτει ότι είναι εύκολο να σταθεροποιηθεί η προσαρμογή του ρυθμού στους τελικούς κόμβους, από ότι η προσαρμογή της διαδρομής μέσα στο δίκτυο.

Όπως θα δούμε σε μία άσκηση, υπάρχει επίσης μία ενδιαφέρουσα αρχιτεκτονική καταμερισμού εργασίας (division-of-labor) ανάμεσα στο επίπεδο δικτύου και το επίπεδο σύνδεσης το οποίο βρίσκεται από κάτω του.

Επιστρέφοντας στη λίστα πρωτοκόλλων ξανά, το επίπεδο εφαρμογής τρέχει εφαρμογές που παράγουν μία ακολουθία μηνυμάτων. Κάθε ένα από αυτά διαιρείται σε τμήματα στο επίπεδο μεταφοράς, με μία επικεφαλίδα επιπέδου 4 να προστίθεται μπροστά από το πραγματικό περιεχόμενο, το αποκαλούμενο ωφέλιμο φορτίο (payload). Κατόπιν περνάει στο επίπεδο δικτύου, το οποίο διαιρεί και ενώνει τα τμήματα ως δεδομενογραφήματα ή πακέτα (packets), με μία επικεφαλίδα επιπέδου 3 μπροστά. Κάθε πακέτο μεταφέρεται περαιτέρω στο επίπεδο σύνδεσης, το οποίο προσθέτει ακόμα μία επικεφαλίδα επιπέδου 2, για να διαμορφώσει ένα πλαίσιο. Αυτό

τελικά περνάει στο φυσικό επίπεδο για μεταφορά. Αυτές οι επικεφαλίδες αποτελούν επιβαρύνσεις (overheads), αλλά περιέχουν χρήσιμες, μερικές φορές ουσιαστικές πληροφορίες, για ταυτοποίηση και έλεγχο. Για παράδειγμα, η επικεφαλίδα επιπέδου 3 περιέχει τη διεύθυνση του πηγαιού κόμβου και τη διεύθυνση του κόμβου προορισμού, που είναι χωρίς αμφιβολία χρήσιμα κατά τη δρομολόγηση. Θα εξετάσουμε τις επιπτώσεις αυτών των σημασιολογικών επιπρόσθετων στην απόδοση στο Κεφάλαιο 19.

Κάθε στοιχείο δικτύου, για παράδειγμα η οικιακή πύλη σας, ο εξοπλισμός WiFi της επιχείρησής σας, ο εξοπλισμός των κεντρικών γραφείων σας κοντά στο κέντρο της πόλης, ο μεγάλος δρομολογητής μέσα στο «νέφος», τρέχει ένα υποσύνολο της λίστας επιπέδων πρωτοκόλλων. Το κάθε ένα θα αποκωδικοποιήσει και θα διαβάσει μόνο τις πληροφορίες της επικεφαλίδας που συνδέονται με το υποσύνολό του. Αυτό αποδίδεται στο Σχήμα 13.3.



Σχήμα 13.3 Διαφορετικά στοιχεία δικτύου επεξεργάζονται τα διαφορετικά στρώματα στη λίστα πρωτοκόλλων. Οι τελικοί κόμβοι (end-hosts) επεξεργάζονται όλα προς τα πάνω μέχρι το επίπεδο εφαρμογής. Οι μεταγωγείς (switches), που προωθούν τα πλαίσια, επεξεργάζονται μέχρι το επίπεδο σύνδεσης. Οι δρομολογητές (routers), που μετακινούν τα πακέτα στο δίκτυο, επεξεργάζονται μέχρι το επίπεδο δικτύου. Σε μία πραγματική από άκρο-προς-άκρο σύνδεση, υπάρχουν συνήθως πολλά περισσότερα άλματα.

13.1.3 Κατανεμημένη Ιεραρχία

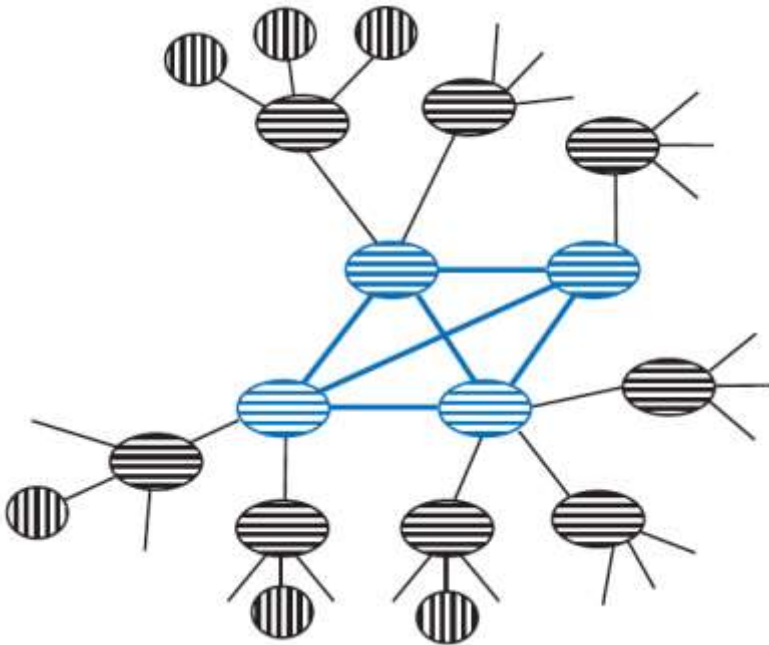
Το Διαδίκτυο δεν είναι σύνθετο μόνο από την άποψη του αριθμού των εργασιών που πρέπει να διαχειριστεί, αλλά και από την άποψη του αριθμού των χρηστών. Η ύπαρξη μίας ιεραρχίας είναι απαραίτητη. Μία από-άκρο-προς-άκρο σύνδεση, για παράδειγμα μία σύνδεση συνεχούς ροής στο YouTube από τους εξυπηρετητές του Google στο iPhone σας, μπορεί να διασχίσει μία ασύρματη εναέρια διασύνδεση, μερικές συνδέσεις στο δίκτυο κινητής τηλεφωνίας και μετά μία ακολουθία από ακόμα περισσότερες συνδέσεις σε πιθανόν πολλαπλούς Παρόχους Υπηρεσιών Διαδικτύου (ISPs) στο δημόσιο Διαδίκτυο.

Ενώ η αρθρωτή προσέγγιση βοηθά στην επίλυση του προβλήματος της πολυπλοκότητας με το «διαίρει και βασίλευε» από την άποψη των λειτουργιών, η ιεραρχία βοηθά στην επίλυση του προβλήματος του μεγάλου μεγέθους με το «διαίρει και βασίλευε» από την άποψη της φυσικής έκτασης. Αυτό είναι ένα επαναλαμβανόμενο θέμα σε πολλά κεφάλαια. Στο τρέχον, μπορούμε να δούμε ότι το Διαδίκτυο, αυτό το δίκτυο των δικτύων με περισσότερα από 30.000 Αυτόνομα Συστήματα (ΑΣ), έχει αρκετά κύρια ιεραρχικά επίπεδα, όπως φαίνεται στην Εικόνα 13.4.

- Λίγοι πολύ μεγάλοι πάροχοι Διαδικτύου με παγκόσμια παρουσία αποκαλούνται πάροχοι βαθμίδας-1 και διαμορφώνουν μία πλήρως συνεργαζόμενη ομότιμη σχέση μεταξύ τους: κάθε πάροχος βαθμίδας-1 έχει κάποια σύνδεση με κάθε έναν από τους άλλους παρόχους βαθμίδας-1. Αυτό το πλήρως συνεργαζόμενο δίκτυο μερικές φορές αποκαλείται η ραχοκοκαλιά του Διαδικτύου. Παραδείγματα παρόχων βαθμίδας-1 είναι οι AT&T, BT, Level 3, Sprint κτλ.
- Υπάρχουν πολλοί περισσότεροι πάροχοι βαθμίδας-2 με τοπική παρουσία. Κάθε πάροχος βαθμίδας-1 είναι συνδεδεμένος με μερικούς παρόχους βαθμίδας-2, διαμορφώνοντας μία σχέση πελάτη-προμηθευτή. Καθένας από αυτούς τους παρόχους βαθμίδας-2 παρέχει συνδεσιμότητα σε πολλούς παρόχους βαθμίδας-3 και αυτή η ιεραρχία συνεχίζει. Το σημείο στο οποίο οποιοδήποτε δύο πάροχοι συνδέονται μεταξύ τους ονομάζεται Σημείο Παρουσίας (Point of Presence – PoP).
- Ένας πάροχος οποιασδήποτε βαθμίδας μπορεί να παρέχει συνδεσιμότητα στο Διαδίκτυο απευθείας στους καταναλωτές. Αυτοί οι πάροχοι, που λαμβάνουν κυκλοφορία μόνο προς ή από τους δικούς τους καταναλωτές, αλλά δεν μεταφέρουν κυκλοφορία από άλλους παρόχους, ονομάζονται πάροχοι στελέχη (stub ISPs). Χαρακτηριστικά, οι πανεπιστημιακοί, οι εταιρικοί και οι αγροτικοί αστικοί πάροχοι ανήκουν σε αυτή την κατηγορία.

Μία άλλη χρήσιμη έννοια στην κατανεμημένη ιεραρχία είναι αυτή της περιοχής (domain). Κάθε επιχειρησιακή οντότητα σχηματίζει μία περιοχή, που ονομάζεται Αυτόνομο Σύστημα (ΑΣ). Συχνά υπάρχει ένας κεντρικός ελεγκτής μέσα σε κάθε ΑΣ. Όπως θα δούμε αργότερα σε αυτό το κεφάλαιο, η δρομολόγηση μέσα σε ένα ΑΣ και η δρομολόγηση ανάμεσα στα ΑΣ ακολουθούν πολύ διαφορετικές προσεγγίσεις.

Αργότερα, στο Κεφάλαιο 15, θα δούμε επίσης πώς οι λειτουργικές και οι χωρικές ιεραρχίες συνδυάζονται στο κτίσιμο δικτύων επικάλυψης.



Σχήμα 13.4 Χωρική ιεραρχία στη δικτύωση: Πολλαπλά επίπεδα παρόχων και οι σχέσεις τους. Κάθε κόμβος σε αυτό το γράφο είναι ένας πάροχος και κάθε σύνδεση αντιπροσωπεύει μία επιχειρησιακή σχέση και τις φυσικές συνδέσεις μεταξύ δύο παρόχων. Οι τέσσερις πάροχοι στο κέντρο είναι πάροχοι βαθμίδας-1, με ομότιμες συνδέσεις μεταξύ τους. Καθένας από αυτούς παρέχει συνδεσιμότητα σε αρκετούς πελάτες παρόχους. Οι πάροχοι στελέχη στα άκρα δεν παρέχουν υπηρεσία μεταφοράς σε κανέναν άλλον πάροχο. Ένας πάροχος οποιασδήποτε βαθμίδας μπορεί επίσης να παρέχει συνδέσεις σε τελικούς χρήστες, οι συνδέσεις αυτές δεν εμφανίζονται εδώ.

13.1.4 Δρομολόγηση IP

Η μεταγωγή πακέτου, η πολυεπίπεδη αρχιτεκτονική και η κατανεμημένη ιεραρχία είναι τρεις βασικές έννοιες του Διαδικτύου. Αφού έχουμε συζητήσει αυτά τα θέματα, μπορούμε να προχωρήσουμε στη δρομολόγηση στο Διαδίκτυο.

Τα δίκτυα μεταφοράς συχνά προσφέρουν ενδιαφέρουσες αναλογίες για την επικοινωνία και τα κοινωνικά δίκτυα. Σε αυτή την περίπτωση, μπορούμε να χρησιμοποιήσουμε μία χρήσιμη αναλογία από την ταχυδρομική υπηρεσία. Για να δρομολογήσουμε ένα γράμμα από τον αποστολέα στον παραλήπτη χρειαζόμαστε τρεις βασικές λειτουργίες.

- Ανάθεση διευθύνσεων. Κατ' αρχάς πρέπει να συνδέσουμε μία μοναδική ετικέτα σε κάθε κόμβο στο δίκτυο, αλλιώς δεν θα μπορούμε να προσδιορίσουμε ούτε τις πηγές και ούτε τους προορισμούς. Στο ταχυδρομικό σύστημα η ετικέτα είναι η ταχυδρομική διεύθυνση, όπως η διεύθυνση με χρήση οδών ή ο αριθμός ταχυδρομικής θυρίδας. Οι ταχυδρομικοί κώδικες μπορούν γρήγορα να σας εστιάσουν σε ένα υποδίκτυο της χώρας. Στο Διαδίκτυο χρησιμοποιούμε τη διεύθυνση IP, που είναι ένας αριθμός 32-bit που συχνά απεικονίζεται ως τέσσερα δεκαδικά νούμερα, χωρισμένα με τελείες. Καθένα από αυτά τα τέσσερα νούμερα εκτείνεται από το 0 ως το 255, δεδομένου ότι προσδιορίζεται από $32/4=8$ bit, για παράδειγμα το 127.12.5.88. Οι «ταχυδρομικοί κώδικες» εδώ ονομάζονται μάσκες υποδικτύου, π.χ. το 127.12.5.0/24 σημαίνει ότι τα πρώτα 24 bits δίνουν το πρόθεμα από όλες τις διευθύνσεις IP του υποδικτύου: κάθε διεύθυνση IP σε αυτό το υποδίκτυο πρέπει να ξεκινά με 127.12.5 και να τελειώνει σε οποιαδήποτε 8 bit. Ωστόσο, στο ταχυδρομικό σύστημα μία διεύθυνση και η ταυτότητα ενός ατόμου είναι ξεχωριστή. Στο Διαδίκτυο μία διεύθυνση IP είναι και η διεύθυνση για την πραγματοποίηση της σύνδεσης και το αναγνωριστικό μίας συσκευής. Αυτό το διπλό φορτίο της λειτουργικότητας σε διευθύνσεις IP προκάλεσε διάφορα προβλήματα ελέγχου στο Διαδίκτυο.
- Δρομολόγηση. Μετά πρέπει να αποφασίσετε τις διαδρομές: είτε μία διαδρομή για κάθε σύνοδο (δρομολόγηση μονής διαδρομής) ή πολλαπλές διαδρομές για κάθε σύνοδο (δρομολόγηση πολλαπλών διαδρομών). Το ταχυδρομείο χρησιμοποιεί δρομολόγηση μονής διαδρομής και η δρομολόγηση αποφασίζεται από πριν, σε ποιες ενδιάμεσες πόλεις περνάει το ταχυδρομείο, ώστε να φτάσει ας πούμε το Princeton, NJ από το Stanford, CA. Υπάρχουν δύο μεγάλοι εύρους κατηγορίες μεθόδων δρομολόγησης: δρομολόγηση βασισμένη σε μετρικές και δρομολόγηση βασισμένη σε πολιτικές. Μέσα σε ένα ΑΣ η δρομολόγηση βασίζεται σε κάποιου είδους μετρική, είτε διαλέγοντας τη συντομότερη διαδρομή ανάμεσα στη δεδομένη πηγή και τον προορισμό, είτε κατανέμοντας την κυκλοφορία σε διαδρομές, ώστε καμία διαδρομή να μην είναι πολύ φορτωμένη. Ωστόσο, ανάμεσα στα ΑΣ, η δρομολόγηση βασίζεται σε πολιτικές. Για παράδειγμα, το ΑΣ1 μπορεί να υποπτευτείται ότι υπάρχουν εισβολείς συνδεδεμένοι στο ΑΣ2, και γι' αυτό αποφεύγει να δρομολογεί πακέτα κατά μήκος οποιασδήποτε διαδρομής που διασχίζει το ΑΣ2.
- Προώθηση. Η προώθηση υλοποιεί την πολιτική της δρομολόγησης. Η πραγματική ενέργεια της προώθησης συμβαίνει κάθε φορά που ένα πακέτο λαμβάνεται σε έναν δρομολογητή ή κάποια επιστολή λαμβάνεται σε ένα ενδιάμεσο ταχυδρομικό γραφείο. Μερικοί μηχανισμοί προώθησης κοιτούν μόνο τη διεύθυνση προορισμού και αποφασίζουν το επόμενο άλμα, ενώ άλλοι, όπως το Πολυπρωτόκολλο Μεταγωγής Ετικέτας (MultiProtocol Label

Switching (MPLS)), διαβάζουν ορισμένες ετικέτες επισυναπτόμενες στο πακέτο, που δείχνουν ρητά το επόμενο άλμα. Σε κάθε περίπτωση, λαμβάνεται μία απόφαση προώθησης και επιλέγεται μία από τις συνδέσεις εξόδου για να σταλεί το πακέτο.

Ας δούμε τώρα καθένα από αυτά με λίγο περισσότερες λεπτομέρειες, πριν εστιάσουμε στο υπόλοιπο κεφάλαιο μόνο στο κομμάτι της δρομολόγησης.

Υπάρχουν δύο εκδόσεις του IP: η έκδοση 4 και η έκδοση 6. Το IPv4 χρησιμοποιεί 32 bit για διευθύνσεις, οι οποίες εξαντλήθηκαν στις αρχές του 2011. Το IPv6 χρησιμοποιεί 4 φορές περισσότερα bit, 128 bit, τα οποία μεταφράζονται σε 2^{128} , γύρω στις 10^{39} διαθέσιμες διευθύνσεις. Αυτό μπορεί να φαίνεται πολύ, αλλά με τη γρήγορη εξάπλωση των συσκευών που είναι συνδεδεμένες με το Διαδίκτυο, είμαστε στο δρόμο να χρησιμοποιήσουμε πολλές από αυτές τις διευθύνσεις. Ένας τρόπος να αναβαθμιστεί ένα δίκτυο IPv4 σε IPv6 είναι να δημιουργήσουμε μία «σήραγγα» μεταξύ δύο στοιχείων δικτύων IPv4, όπου τα πακέτα IPv6 ενσωματώνονται σε επικεφαλίδες IPv4.

Πώς διατίθενται αυτές οι διευθύνσεις IP; Δίνονταν σε μπλοκ, με διαφορετικά μεγέθη μπλοκ που καθορίζονταν από την «κλάση». Για παράδειγμα, κάθε μπλοκ διεύθυνσης κλάσης A έχει ένα σταθερό πρόθεμα 8-bit, έτσι $2^{32-8}=2^{24}$ διευθύνσεις στο μπλοκ κλάσης A. Αυτό συνήθως δίνεται σε έναν εθνικό πάροχο Διαδικτύου ή έναν μεγάλο προμηθευτή εξοπλισμού. Οι χαμηλότερες κλάσεις έχουν λιγότερες διευθύνσεις ανά μπλοκ. Αλλά το μπλοκ αυτό των 8-bit έφερε μεγάλη σπατάλη σε διαθέσιμες, αλλά αχρησιμοποίητες διευθύνσεις IP. Έτσι η κοινότητα του Διαδικτύου μετατοπίστηκε στην Αταξική Δρομολόγηση Δικτυακών Περιοχών (Classless Inter Domain Routing – CIDR), όπου το μπλοκ δεν χρειάζεται να είναι σε πολλαπλάσια των 8 bit.

Ως συσκευή, μπορείτε να έχετε είτε μία σταθερή στατική διεύθυνση IP ορισμένη σε εσάς είτε πρέπει να πάρετε μία δυναμική, που θα σας την αναθέσει ένας ελεγκτής (controller) ο οποίος βρίσκεται στο τοπικό δίκτυο. Αυτός ο ελεγκτής ονομάζεται εξυπηρετητής Πρωτοκόλλου Δυναμικής Διάρθρωσης Υποδοχέων (Dynamic Host Configuration Protocol (DHCP)). Μία συσκευή έρχεται σε επαφή με τον εξυπηρετητή DHCP, λαμβάνει μία τρέχουσα αχρησιμοποίητη διεύθυνση IP και την επιστρέφει στη δεξαμενή των διευθύνσεων IP, όταν δεν τη χρειάζεται πλέον. Θα αναρωτιέστε πώς μπορεί μία συσκευή να επικοινωνήσει με έναν εξυπηρετητή DHCP κατά πρώτο λόγο. Θα εξετάσουμε τα πρωτόκολλα που εμπλέκονται στο Κεφάλαιο 19. Μερικές φορές η διεύθυνση που δίνεται σε μία συσκευή σε ένα τοπικό δίκτυο, π.χ. είναι εταιρικό ενδοδίκτυο, είναι διαφορετική από αυτή που φαίνεται στον εξωτερικό κόσμο και ένας δρομολογητής Μετάφρασης Διευθύνσεων Δικτύου (Network Address Translation (NAT)) κάνει αυτή τη μετάφραση και στις δύο κατευθύνσεις.

Όπως αναφέρθηκε, η δρομολόγηση μεταξύ αυτονόμων συστημάτων (ΑΣ) είναι πολύ διαφορετική από την ενδο-ΑΣ δρομολόγηση. Το Πρωτόκολλο Συνοριακής Πύλης

(Border Gateway Protocol (BGP)) είναι το κυρίαρχο πρωτόκολλο για την ανακάλυψη διευθύνσεων και την επίτευξη συνδεσιμότητας στην ενδο-ΑΣ δρομολόγηση. «Κολλάει» μαζί το Διαδίκτυο. Ωστόσο, ως πρωτόκολλο δρομολόγησης βασισμένο σε πολιτικές, είναι ένα περίπλοκο, ακατάστατο πρωτόκολλο, με πολλές γκρίζες ζώνες. Θα το περιγράψουμε εν συντομία στο Προχωρημένο Υλικό.

Μέσα σε ένα ΑΣ, υπάρχουν δύο κύριοι τύποι πρωτόκολλων δρομολόγησης τα οποία είναι βασισμένα σε μετρικές: το Πρωτόκολλο Πληροφοριών Δρομολόγησης (Routing Information Protocol (RIP)) χρησιμοποιεί τη μέθοδο διανύσματος απόστασης, όπου ο κάθε κόμβος συλλέγει πληροφορίες σχετικά με τις αποστάσεις μεταξύ του ίδιου και άλλων κόμβων, και το πρωτόκολλο του Πρώτου Συντομότερου Ανοικτού Μονοπατιού (Open Shortest Path First (OSPF)) χρησιμοποιεί τη μέθοδο συνδεδεμένης κατάστασης, όπου κάθε κόμβος προσπαθεί να κατασκευάσει μία σφαιρική άποψη της τοπολογίας ολόκληρου του δικτύου. Θα εστιάσουμε στο πιο απλό RIP στην επόμενη ενότητα, αφήνοντας το OSPF για το Προχωρημένο Υλικό.

Όταν ένα πακέτο φτάνει σε μία διεπαφή δρομολόγησης, αυτή η διεπαφή ενεργεί ως πύλη εισόδου για το πακέτο, ενώ μία άλλη διεπαφή ενεργεί ως πύλη εξόδου. Στο ενδιάμεσο είναι ο ιστός μεταγωγής, που μετακινεί το πακέτο από μία πύλη εισόδου σε μία πύλη εξόδου φυσικά. Αν τα πακέτα φτάνουν πολύ γρήγορα, εμφανίζεται συμφόρηση μέσα στο δρομολογητή. Μερικές φορές αυτό συμβαίνει, επειδή οι πύλες εξόδου που έχουν επιλεγεί είναι κατειλημμένες, μερικές φορές επειδή το υλικό μεταγωγής είναι απασχολημένο και μερικές φορές επειδή ένα πακέτο περιμένει να έρθει η σειρά του στην ουρά αναμονής της πύλης εισόδου, εμποδίζοντας κατά συνέπεια όλα τα πακέτα πίσω του στην ίδια ουρά αναμονής εισόδου.

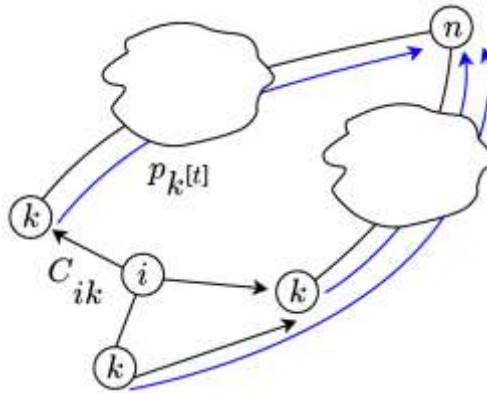
Ποια πύλη εξόδου είναι η «σωστή»; Αυτό αποφασίζεται κοιτώντας τον πίνακα προώθησης, ο οποίος είναι ή αποθηκευμένος κεντρικά στο δρομολογητή ή αντιγραμμένος με ένα αντίγραφο σε κάθε πύλη εισόδου. Ο πίνακας προώθησης συνδέει τις αποφάσεις δρομολόγησης με πραγματικές δράσεις προώθησης. Ένας κοινός τύπος πίνακα προώθησης απεικονίζει όλες τις διευθύνσεις προορισμού IP στο Διαδίκτυο και προσδιορίζει σε ποια πύλη εξόδου, κατά συνέπεια τον επόμενο δρομολογητή στο μονοπάτι, πρέπει να μεταβεί ένα πακέτο, με βάση τη διεύθυνση προορισμού που είναι γραμμένη στην επικεφαλίδα. Υπάρχουν πάρα πολλές διευθύνσεις IP εκεί έξω, και έτσι ο πίνακας δρομολόγησης συχνά ομαδοποιεί κάποιες διευθύνσεις σε μία ισοδύναμη κλάση εισαγωγής.

Τώρα πρόκειται να μελετήσουμε ένα μέλος της οικογένειας της ενδο-ΑΣ δρομολόγησης και μετά πώς κατασκευάζονται οι πίνακες προώθησης από καταναμημένα μηνύματα που περνούν ανάμεσα στους δρομολογητές.

13.2 Μία Εκτενής Απάντηση

Θεωρήστε ένα κατευθυνόμενο γράφημα $G = (V, E)$, που αναπαριστά την τοπολογία μέσα σε ένα ΑΣ, όπου κάθε κόμβος στο σύνολο των κόμβων V είναι ένας δρομολογητής και κάθε σύνδεση στο σύνολο συνδέσεων E είναι μία φυσική σύνδεση από έναν δρομολογητή i σε έναν άλλο δρομολογητή j .

Κάθε σύνδεση έχει ένα κόστος c_{ij} . Είναι συχνά ένας αριθμός περίπου ανάλογος με το μήκος της σύνδεσης. Αν είναι 1 για όλες τις συνδέσεις, τότε ελαχιστοποιώντας το κόστος κατά μήκος μίας διαδρομής είναι το ίδιο με το να ελαχιστοποιήσουμε τον αριθμό των αλμάτων. Αν απεικονίζει δυναμικά τη συνθήκη συμφόρησης σε αυτή τη σύνδεση, θα οδηγούσε σε δυναμική, ευαίσθητη στο φορτίο δρομολόγηση. Αλλά το IP δεν ασκεί τη δυναμική δρομολόγηση, αφήνοντας την ευαισθησία στο φορτίο στον έλεγχο συμφόρησης του TCP.



Σχήμα 13.5 Η εξίσωση του Bellman για τη δρομολόγηση ελάχιστου κόστους. Το ελάχιστο κόστος από έναν κόμβο i σε έναν προορισμό n είναι το μικρότερο, ανάμεσα σε όλους τους γείτονές του, άθροισμα του κόστους από τον i σε ένα γείτονα και του κόστους από αυτόν το γείτονα στο n . Ο κόμβος i δεν χρειάζεται να γνωρίζει πώς οι γείτονές του πηγαίνουν στο n , μόνο το κόστος της προσπέλασης στο n , μέσω κάθε γείτονα.

Το πρόβλημα της συντομότερης διαδρομής είναι μία σημαντική ειδική περίπτωση του προβλήματος ροής δικτύου, που είναι με τη σειρά του μία σημαντική ειδική περίπτωση του γραμμικού προγραμματισμού. Ένα πιο ακριβές όνομα είναι το πρόβλημα διαδρομής ελάχιστου κόστους, καθώς στην πραγματικότητα βρίσκουμε τη διαδρομή ελάχιστου κόστους για κάθε ζεύγος κόμβων σε ένα δεδομένο γράφο. Το «ελάχιστο κόστος» θα ήταν ισοδύναμο με το «συντομότερο», μόνο όταν τα κόστη σύνδεσης είναι οι φυσικές αποστάσεις. Αλλά και πάλι, το «πρόβλημα συντομότερης διαδρομής» ταιριάζει και για το λόγο αυτό θα χρησιμοποιούμε αυτόν τον όρο.

Το πρόβλημα συντομότερης διαδρομής έχει μελετηθεί εκτενώς από τις αρχές της δεκαετίας του 1950 και υπάρχουν αρκετές γνωστές προσεγγίσεις: ο αλγόριθμος Bellman-Ford, ο αλγόριθμος Dijkstra, η δυαδικότητα Lagrange κτλ. Θα εστιάσουμε στον αλγόριθμο Bellman-Ford, γιατί είναι απλός, κατανεμημένος και απεικονίζει τις βασικές αρχές του δυναμικού προγραμματισμού. Επίσης οδηγεί σε μία πλήρως κατανεμημένη και ασύγχρονη εφαρμογή που χρησιμοποιείται στο RIP, εν μέρει στο BGP και στη μέθοδο δρομολόγησης στο αρχικό ARPANET.

Κατ' αρχάς, μερική σημειογραφία. Για την ώρα, καθορίστε έναν κόμβο προορισμού n : μπορούμε να γενικεύσουμε σε πολλαπλούς προορισμούς εύκολα. Θέτουμε $p_i[t]$ το μήκος της συντομότερης διαδρομής από τον κόμβο i στον προορισμό n χρησιμοποιώντας το πολύ t συνδέσεις. Δεν είναι σύμπτωση, ότι χρησιμοποιούμε το χρονικό σύμβολο t για να συλλάβουμε αυτόν το χωρικό ορισμό. Θα δούμε σύντομα ότι το t πράγματι δεικτοδοτεί και τις επαναλήψεις.

Αν ο κόμβος i δεν μπορεί να φτάσει στον προορισμό n σε t βήματα, λέμε ότι $p_i[t] = \infty$. Προφανώς, κατά την έναρξη $p_i[0] = \infty$ για όλους τους κόμβους i , εκτός κι αν είναι ο ίδιος ο προορισμός n .

Εδώ έρχεται η βασική ιδέα του αλγορίθμου Bellman-Ford. Προφανώς, ο i πρέπει να φτάσει στο n μέσω κάποιων γειτονικών κόμβων. Αντιλαμβανόμαστε ότι το $p_i[t+1]$ δεν μπορεί να διασπαστεί σε δύο μέρη, όπως φαίνεται στο Σχήμα 13.5:

- Το κόστος c_{ik} για να πάει από τον κόμβο i σε έναν από τους εξερχόμενους γείτονες του k . Ένας εξερχόμενος γείτονας είναι ένας κόμβος, όπου υπάρχει μία σύνδεση που δείχνει από τον i σε αυτόν και συμβολίζουμε το σύνολο αυτών των γειτόνων για τον κόμβο i ως $O(i)$.
- Το ελάχιστο κόστος για να πάμε από αυτόν το γείτονα k στον προορισμό n , χρησιμοποιώντας το πολύ t βήματα, καθώς έχουμε ήδη χρησιμοποιήσει 1 βήμα (από τα $t+1$ βήματα) για να πάμε στον κόμβο k .

Η διαδρομή ελάχιστου κόστους επιλέγει το γείτονα που ελαχιστοποιεί το άθροισμα των δύο παραπάνω κοστών:

$$p_i[t+1] = \min_{k \in O(i)} \{c_{ik} + p_k[t]\}. \quad (13.1)$$

Ας υποθέσουμε ότι κάθε κόμβος γνωρίζει το κόστος για να φτάσει στους εξερχόμενους γείτονές του. Τότε, με επαναληπτικές ανανεώσεις του $p_i[t]$ και μεταδίδοντας ένα διάνυσμα που περιγράφει αυτούς τους ανανεωμένους αριθμούς στους γείτονές του, μπορούμε να εκτελέσουμε τη (13.1) με έναν κατανεμημένο τρόπο. Δεν υπάρχει

αμφιβολία ότι ονομάζεται αλγόριθμος δρομολόγησης διανύσματος απόστασης. Η μετάδοση του διανύσματος απόστασης τριγύρω και η αρχικοποίηση των τιμών του c_{ik} απαιτούν ένα μικρό πρωτόκολλο, όπως θα δούμε σύντομα με το RIP.

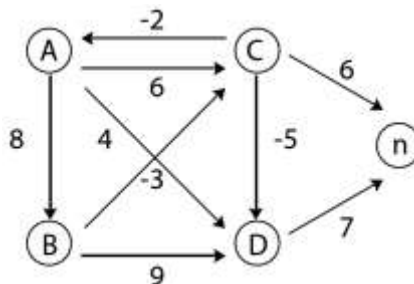
Μία γρήγορη παράκαμψη: στην πραγματικότητα υπάρχει μία πολύ ευρεία ιδέα πίσω από την επαναληπτική εξίσωση Bellman (13.1). Η βελτιστοποίηση μέσα από μία ακολουθία χρονοθυρίδων ή χωρικών σημείων ανήκει στον ερευνητικό τομέα του δυναμικού προγραμματισμού. Για πολλά μοντέλα συστημάτων όπου το κόστος που θέλουμε να ελαχιστοποιήσουμε είναι προσθετικό με το χρόνο ή το χώρο, και η εξάρτηση μεταξύ των σταδίων του προβλήματος είναι χωρίς μνήμη (το επόμενο στάδιο εξαρτάται μόνο από το τρέχον), γνωρίζουμε ότι (αυτό είναι αργκό) η «ουρά» της βέλτιστης λύσης στην ελαχιστοποίηση του κόστους είναι η βέλτιστη λύση στην «ουρά» της ελαχιστοποίησης του κόστους. Ο αλγόριθμος Bellman-Ford είναι μία ειδική περίπτωση αυτής της γενικής αρχής.

13.3 Παραδείγματα

13.3.1 Κεντρικοποιημένος υπολογισμός Bellman-Ford

Κατ' αρχάς, ένα παράδειγμα του αλγόριθμου Bellman-Ford. Υποθέστε ότι έχουμε μία τοπολογία δικτύου όπως φαίνεται στο Σχήμα 13.6. Τα αρνητικά βάρη των συνδέσεων είναι εκεί μόνο για να δείξουν ότι ο αλγόριθμος Bellman-Ford μπορεί να τα προσαρμόσει, εφόσον δεν υπάρχουν κύκλοι με αρνητικά βάρη, καθώς αυτοί μπορεί να μειώσουν το κόστος κάποιας διαδρομής στο αρνητικό άπειρο.

Σε αυτό το μικρό παράδειγμα με τέσσερις κόμβους (χωρίς να υπολογίζουμε τον προορισμό), γνωρίζουμε ότι μπορούμε να σταματήσουμε μετά από 4 επαναλήψεις, καθώς κάθε διαδρομή που περνάει πέντε κόμβους ή παραπάνω θα πρέπει να κάνει έναν κύκλο και αυτό μπορεί μόνο να προσθέσει κόστος στη διαδρομή, συνεπώς δεν θα είναι ποτέ βέλτιστη. Αλλά σε μία πραγματική, κατανεμημένη εφαρμογή ενός πρωτόκολλου διανύσματος απόστασης δεν γνωρίζουμε πόσοι κόμβοι υπάρχουν, κι έτσι θα πρέπει να βασιστούμε στην έλλειψη νέων μηνυμάτων για να καθορίσουμε πότε είναι ασφαλές να τερματίσουμε τον αλγόριθμο.



Σχήμα 13.6 Ένα παράδειγμα για να απεικονίσουμε τον αλγόριθμο Bellman-Ford. Εδώ θέλουμε να βρούμε τη συντομότερη διαδρομή από τους κόμβους A, B, C και D σε έναν κοινό κόμβο προορισμού n. Υπάρχουν συνδέσεις με αρνητικά βάρη μόνο για να απεικονίσουμε ότι ο αλγόριθμος μπορεί να τις χειριστεί. Αλλά δεν υπάρχουν κύκλοι με αρνητικά βάρη, γιατί θα καθιστούσαν το πρόβλημα συντομότερης διαδρομής εσφαλμένα καθορισμένο.

Προσπαθούμε να βρούμε τις ελάχιστες διαδρομές από τους κόμβους A, B, C και D στον προορισμό n. Αρχικοποιούμε τις αποστάσεις

$p_A[0] = p_B[0] = p_C[0] = p_D[0] = \infty$. Και φυσικά δεν υπάρχει κόστος για να φτάσει τον εαυτό του: $p_n[t] = 0$ για όλα τα t.

Για $t = 1$, με την εξίσωση Bellman, έχουμε

$$p_A[1] = \min\{c_{AB} + p_B[0], c_{AC} + p_C[0], c_{AD} + p_D[0]\} = \min\{8 + \infty, 6 + \infty, 4 + \infty\} = \infty,$$

$$p_B[1] = \min\{c_{BC} + p_C[0], c_{BD} + p_D[0]\} = \min\{-3 + \infty, 9 + \infty\} = \infty,$$

$$p_C[1] = \min\{c_{Cn} + p_n[0], c_{CD} + p_D[0], c_{CA} + p_A[0]\} = \min\{6 + 0, -5 + \infty, -2 + \infty\} = 6,$$

$$p_D[1] = \min\{c_{Dn} + p_n[0]\} = 7 + 0 = 7.$$

Προσέξτε ότι ο κόμβος D έχει μόνο μία εξερχόμενη σύνδεση, έτσι $p_D = 7$ και δεν χρειάζεται να συνεχίσουμε να το υπολογίζουμε.

Ομοίως, για $t = 2$ έχουμε

$$p_A[2] = \min\{c_{AB} + p_B[1], c_{AC} + p_C[1], c_{AD} + p_D[1]\} = \min\{8 + \infty, 6 + 6, 4 + 7\} = 11,$$

$$p_B[2] = \min\{c_{BC} + p_C[1], c_{BD} + p_D[1]\} = \min\{-3 + 6, 9 + 7\} = 3,$$

$$p_C[2] = \min\{c_{Cn} + p_n[1], c_{CD} + p_D[1], c_{CA} + p_A[1]\} = \min\{6 + 0, -5 + 7, -2 + 8\} = 2.$$

Για $t = 3$,

$$p_A[3] = \min\{c_{AB} + p_B[2], c_{AC} + p_C[2], c_{AD} + p_D[2]\} = \min\{8 + 3, 6 + 2, 4 + 7\} = 8,$$

$$p_B[3] = \min\{c_{BC} + p_C[2], c_{BD} + p_D[2]\} = \min\{-3 + 2, 9 + 7\} = -1,$$

$$p_C[3] = \min\{c_{Cn} + p_n[2], c_{CD} + p_D[2], c_{CA} + p_A[2]\} = \min\{6 + 0, -5 + 7, -2 + 11\} = 2.$$

Για $t = 4$,

$$p_A[4] = \min\{c_{AB} + p_B[3], c_{AC} + p_C[3], c_{AD} + p_D[3]\} = \min\{8 - 1, 6 + 2, 4 + 7\} = 7,$$

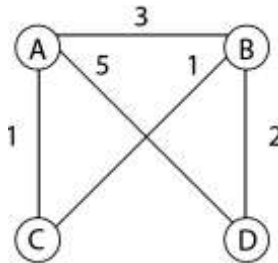
$$p_B[4] = \min\{c_{BC} + p_C[3], c_{BD} + p_D[3]\} = \min\{-3 + 2, 9 + 7\} = -1,$$

$$\begin{aligned}
 p_C[4] &= \min\{c_{Cn} + p_n[3], c_{CD} + p_D[3], c_{CA} + p_A[3]\} \\
 &= \min\{6 + 0, -5 + 7, -2 + 8\} = 2.
 \end{aligned}$$

Μπορούμε επίσης να εύκολα να παρακολουθήσουμε τις πορείες που πάρθηκαν από κάθε κόμβο για να φτάσει το n : το D πηγαίνει απευθείας στο n , το C πηγαίνει μέσω του D στο n , το B πηγαίνει μέσω του C και το A πηγαίνει μέσω του B.

13.3.2 Κατανεμημένο Πρωτόκολλο Πληροφοριών Δρομολόγησης (RIP)

Μέχρι τώρα έχουμε υποθέσει ότι έχουμε κεντρικοποιημένο υπολογισμό. Αλλά φανταστείτε ότι είστε ένας από τους κόμβους που προσπαθείτε να βρείτε πώς θα φτάσετε σε άλλους κόμβους στο δίκτυο. Πώς γνωρίζετε το κόστος για να φτάσετε σε άλλους κόμβους και, επίσης, πώς ξεκινάτε;



Σχήμα 13.7 Ένα παράδειγμα για να απεικονίσουμε το κατανεμημένο μήνυμα που περνά το RIP, όπου κάθε κόμβος θέλει να βρει μία από τις συντομότερες διαδρομές προς κάθε κόμβο, μέσω της διαβίβασης μηνυμάτων. Πάλι, επιλέγουμε ένα εξαιρετικά μικρό δίκτυο, ώστε να περάσουμε λεπτομερώς από τα αριθμητικά βήματα. Μία σημαντική πρόκληση της δρομολόγησης στο Διαδίκτυο είναι στην πραγματικότητα η κλίμακα του δικτύου, ως εκ τούτου προκύπτει και η επιθυμία για τις κατανεμημένες λύσεις κατά πρώτο λόγο.

Τώρα θα περιγράψουμε το πρωτόκολλο μεταφοράς μηνύματος στο RIP, το οποίο επιτρέπει την ανακάλυψη και την ανανέωση των c_{ik} και $p_i[t]$ κατά μήκος των κόμβων. Για απλότητα της παρουσίασης, υποθέστε ότι όλες οι συνδέσεις είναι αμφίδρομες: αν ο κόμβος i μπορεί να στείλει μηνύματα στον κόμβο j , έτσι μπορεί και ο j στον i .

Το μήνυμα που μεταβιβάζεται στα πρωτόκολλα δρομολόγησης διανύσματος απόστασης έχει την ακόλουθη μορφή:

[NodeID, DestinationID, CostofMinCostPath]. .

Στην αρχή, στο 0, κάθε κόμβος γνωρίζει μόνο τη δική του ύπαρξη, έτσι κάθε κόμβος i μπορεί να μεταβιβάσει μόνο σε αυτό το διάνυσμα $[node_i, node_i, 0]$. Αλλά μόλις

κάθε κόμβος λάβει μηνύματα από τους γείτονές του, μπορεί να ανανεώσει τη λίστα με τα διανύσματα.

Υπάρχουν αρκετά κύρια χαρακτηριστικά στη διαβίβαση μηνυμάτων.

- Μικρά μηνύματα: Όλες οι λεπτομερείς πληροφορίες της τοπολογίας (για το ποιος συνδέεται με ποιον και ποια είναι τα κόστη των συνδέσεων) μπορούν να συνοψιστούν σε αυτές τις λίστες των διανυσμάτων απόστασης.
- Τοπική αλληλεπίδραση: Η μεταφορά μηνυμάτων από γείτονα σε γείτονα είναι αρκετή για να διαδώσει αυτή τη σύνοψη, έτσι ώστε να ανακαλυφθούν συντομότερες διαδρομές από το ένα άκρο στο άλλο. Δεν είναι αναγκαία η μετάδοση όλων των συνόψεων σε όλους τους κόμβους.
- Τοπική άποψη: Ακόμα και όταν λαμβάνονται αυτές οι βέλτιστες διαδρομές, κάθε κόμβος έχει τοπική άποψη: γνωρίζει μόνο ποιος γείτονας θα στείλει το πακέτο με μία δεδομένη διεύθυνση προορισμού, αλλά δεν έχει ιδέα πώς θα είναι η πραγματική από άκρο σε άκρο διαδρομή. Όσο για την αποστολή των πακέτων με χρήση της γνώσης της διεύθυνσης προορισμού IP, δεν χρειάζεται να τη γνωρίζει, γιατί οι δρομολογητές θα προωθήσουν τα πακέτα βήμα-βήμα.

Εδώ παρουσιάζεται ένα παράδειγμα για να απεικονίσει αυτή την επαναληπτική και κατανεμημένη μέθοδο δρομολόγησης, για το μικρό δίκτυο που φαίνεται στο Σχήμα 13.7. Ψάχνουμε να βρούμε τις ελάχιστες διαδρομές ανάμεσα σε όλους τους κόμβους. Μπορούμε να συλλέξουμε όλα τα διανύσματα απόστασης

$[NodeID, DestinationID, CostofMinCostPath]$, μαζί με την απόφαση για το επόμενο βήμα, σε έναν πίνακα δρομολόγησης για κάθε κόμβο. Κάθε κόμβος αποθηκεύει μόνο τις πληροφορίες για το δικό του NodeID.

Για $t = 0$ έχουμε τους παρακάτω, σχεδόν κενούς, πίνακες, έναν για κάθε κόμβο:

NodeID	DestinationID	Cost of MinCost Path	Next Node
A	A	0	A

NodeID	DestinationID	Cost of MinCost Path	Next Node
B	B	0	B

NodeID	DestinationID	Cost of MinCost Path	Next Node
C	C	0	C

NodeID	DestinationID	Cost of MinCost Path	Next Node
D	D	0	D

Σε κάθε επανάληψη, κάθε κόμβος στέλνει τα διανύσματα απόστασης (τον παραπάνω πίνακα δρομολόγησης, εκτός από τις πληροφορίες του επόμενου βήματος) στους γείτονές του. Για παράδειγμα, στο $t = 1$, ο κόμβος B λαμβάνει μηνύματα από τους κόμβους A, C, και D. Ο κόμβος B λαμβάνει $[A, A, 0]$ από τον κόμβο A, $[C, C, 0]$ από τον κόμβο C και $[D, D, 0]$ από τον κόμβο D.

Με όρους της εξίσωσης Bellman (13.1), αυτό λέει στον κόμβο B ότι $p_A[0] = 0$ για τον προορισμό A, $p_C[0] = 0$ για τον προορισμό C και $p_D[0] = 0$ για τον προορισμό D. Όλες οι άλλες αποστάσεις είναι άπειρες. Ο κόμβος B τότε χρησιμοποιεί την (13.1) για να υπολογίσει τις νέες αποστάσεις στον πίνακα δρομολόγησης.

Ας εργαστούμε στους υπολογισμούς για τον προορισμό A στο $t = 1$:

$$p_A[1] = 0,$$

$$p_B[1] = \min\{c_{BA} + p_A[0], c_{BC} + p_C[0], c_{BD} + p_D[0]\} = \min\{3 + 0, 1 + \infty, 2 + \infty\} = 3,$$

$$p_C[1] = \min\{c_{CA} + p_A[0], c_{CB} + p_B[0]\} = \min\{1 + 0, 1 + \infty\} = 1,$$

$$p_D[1] = \min\{c_{DA} + p_A[0], c_{DB} + p_B[0]\} = \min\{5 + 0, 2 + \infty\} = 5.$$

Για τον προορισμό B στο $t = 1$, έχουμε

$$p_A[1] = \min\{c_{AB} + p_B[0], c_{AC} + p_C[0], c_{AD} + p_D[0]\} = \min\{3 + 0, 1 + \infty, 5 + \infty\} = 3,$$

$$p_B[1] = 0,$$

$$p_C[1] = \min\{c_{CA} + p_A[0], c_{CB} + p_B[0]\} = \min\{1 + \infty, 1 + 0\} = 1,$$

$$p_D[1] = \min\{c_{DA} + p_A[0], c_{DB} + p_B[0]\} = \min\{5 + \infty, 2 + 0\} = 2.$$

Για τον προορισμό C στο $t = 1$, δεν υπάρχει $p_D[1]$ καθώς ο κόμβος C δεν έχει αντιληφθεί ακόμα την ύπαρξή του:

$$p_A[1] = \min\{c_{AB} + p_B[0], c_{AC} + p_C[0], c_{AD} + p_D[0]\} = \min\{3 + \infty, 1 + 0, 5 + \infty\} = 1,$$

$$p_B[1] = \min\{c_{BA} + p_A[0], c_{BC} + p_C[0], c_{BD} + p_D[0]\} = \min\{3 + \infty, 1 + 0, 2 + \infty\} = 1,$$

$$p_C[1] = 0.$$

Για τον προορισμό D στο $t=1$, δεν υπάρχει $p_C[1]$, καθώς ο κόμβος C δεν έχει αντιληφθεί ακόμα την ύπαρξή του:

$$p_A[1] = \min\{c_{AB} + p_B[0], c_{AC} + p_C[0], c_{AD} + p_D[0]\} = \min\{3 + \infty, 1 + \infty, 5 + 0\} = 5,$$

$$p_B[1] = \min\{c_{BA} + p_A[0], c_{BC} + p_C[0], c_{BD} + p_D[0]\} = \min\{3 + \infty, 1 + \infty, 2 + 0\} = 2,$$

$$p_D[1] = 0.$$

Αν και κάθε σύνολο από τους υπολογισμούς, που γίνονται παραπάνω είναι για την αποστολή πολλών κόμβων σε ένα μοναδικό προορισμό, κατά την εκτέλεση του RIP γίνεται το αντίθετο: κάθε κόμβος εκτελεί υπολογισμούς για διάφορους προορισμούς. Στο $t=1$ κάθε κόμβος αποθηκεύει τον δικό του πίνακα, που φαίνεται παρακάτω. Ο C και ο D δεν μπορούν να φτάσουν ο ένας τον άλλον, γιατί μέχρι τώρα το μήνυμα που περνάει αποκαλύπτει μόνο μηδενικού ή ενός βήματος διαδρομές και αυτοί οι δύο κόμβοι χρειάζονται τουλάχιστον δύο βήματα για να συνδεθούν.

NodeID	DestinationID	Cost of MinCost Path	Next Node
A	A	0	A
A	B	3	B
A	C	1	C
A	D	5	D

NodeID	DestinationID	Cost of MinCost Path	Next Node
B	A	3	A
B	B	0	B
B	C	1	C
B	D	2	D

NodeID	DestinationID	Cost of MinCost Path	Next Node
C	A	1	A
C	B	1	B

NodeID	DestinationID	Cost of MinCost Path	Next Node
D	A	5	A
D	B	2	B
D	D	0	D

Τώρα, στο $t = 2$, όλοι οι κόμβοι στέλνουν τα ανανεωμένα διανύσματα απόστασης στους γείτονές τους.

- Ο κόμβος B λαμβάνει $[A, B, 3], [A, C, 1], [A, D, 5]$ από τον κόμβο A. Αυτό λέει στον κόμβο B ότι $p_A[1] = 3$ για τον προορισμό B, $p_A[1] = 1$ για τον προορισμό C και $p_A[1] = 5$ για τον προορισμό D.
- Ο κόμβος B λαμβάνει $[C, A, 5], [C, B, 1]$ από τον κόμβο C. Αυτό λέει στον κόμβο B ότι $p_C[1] = 5$ για τον προορισμό A, $p_C[1] = 1$ για τον προορισμό B και $p_C[1] = \infty$ για τον προορισμό D.
- Ο κόμβος B λαμβάνει $[D, A, 5], [D, B, 2]$ από τον κόμβο D. Αυτό λέει στον κόμβο B ότι $p_D[1] = 5$ για τον προορισμό A, $p_D[1] = 2$ για τον προορισμό B και $p_D[1] = \infty$ για τον προορισμό C.

Ανανεώνουμε όλους τους πίνακες δρομολόγησης στο $t = 2$. Εστιάζουμε στον πίνακα του κόμβου A για να απεικονίσουμε την πηγή.

Για τον προορισμό B στο $t = 2$,

$$p_B[2] = \min\{c_{AB} + p_B[1], c_{AC} + p_C[1], c_{AD} + p_D[1]\} = \min\{3 + 0, 1 + 1, 5 + 2\} = 2.$$

Για τον προορισμό C στο $t = 2$ έχουμε

$$p_C[2] = \min\{c_{AB} + p_B[1], c_{AC} + p_C[1], c_{AD} + p_D[1]\} = \min\{3 + 1, 1 + 0, 5 + \infty\} = 1.$$

Για τον προορισμό D στο $t = 2$, έχουμε το ίδιο κόστος είτε μέσω του κόμβου B, είτε απευθείας στον D:

$$p_D[2] = \min\{c_{AB} + p_B[1], c_{AC} + p_C[1], c_{AD} + p_D[1]\} = \min\{3 + 2, 1 + \infty, 5 + 0\} = 5.$$

Κάθε κόμβος αποθηκεύει τις δικές του πληροφορίες δρομολόγησης σε έναν πίνακα στο $t = 2$. Τα παραπάνω επεξηγούν τις καταχωρίσεις στον πρώτο πίνακα παρακάτω και θα δούμε επίσης ότι οι κόμβοι C και D τώρα γνωρίζουν για την ύπαρξη ο ένας του άλλου:

NodeID	DestinationID	Cost of MinCost Path	Next Node
A	A	0	A
A	B	2	C
A	C	1	C
A	D	5	B

NodeID	DestinationID	Cost of MinCost Path	Next Node
B	A	2	C
B	B	0	B
B	C	1	C
B	D	2	D

NodeID	DestinationID	Cost of MinCost Path	Next Node
C	A	1	A
C	B	1	B
C	C	0	C
C	D	2	D

NodeID	DestinationID	Cost of MinCost Path	Next Node
D	A	5	A
D	B	2	B
D	C	3	B
D	D	0	D

Ομοίως, κάθε κόμβος αποθηκεύει τις πληροφορίες δρομολόγησής του σε έναν πίνακα στο $t = 3$:

NodeID	DestinationID	Cost of MinCost Path	Next Node
A	A	0	A
A	B	2	C
A	C	1	C
A	D	4	C

NodeID	DestinationID	Cost of MinCost Path	Next Node
B	A	2	C
B	B	0	B
B	C	1	C
B	D	2	D

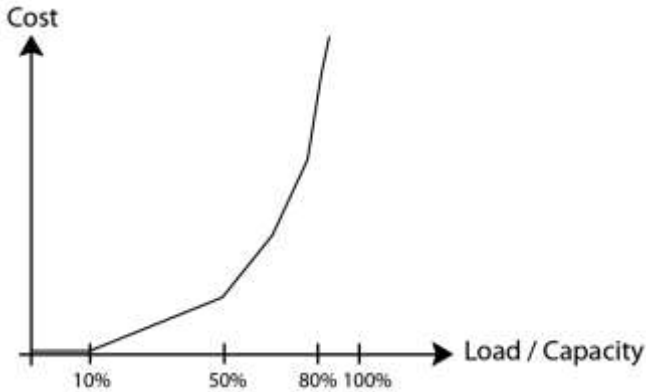
NodeID	DestinationID	Cost of MinCost Path	Next Node
C	A	1	A
C	B	1	B
C	C	0	C
C	D	3	B

NodeID	DestinationID	Cost of MinCost Path	Next Node
D	A	4	B
D	B	2	B
D	C	3	B

D D 0 D

Περαιτέρω επαναλήψεις δεν επιφέρουν άλλες αλλαγές. Οι πίνακες δρομολόγησης έχουν συγκλίνει προς τη σωστή λύση, μέσω των κατανεμημένων μηνυμάτων που περνάει το RIP.

Τι γίνεται με τις αποτυχίες των συνδέσεων; Είναι η δρομολόγηση διανύσματος απόστασης εύρωστη απέναντι σε γεγονότα όπως μία διακοπή σύνδεσης; Θα το ανακαλύψουμε σε μία άσκηση.



Σχήμα 13.8 Μία τυπική συνάρτηση κόστους $C(f)$: αύξουσα, κυρτή και τμηματικά γραμμική. Αυξάνεται απότομα όταν το ποσό του φορτίου (load) πλησιάζει περίπου το 50% της χωρητικότητας (capacity) της σύνδεσης και φτάνει σε πολύ υψηλά μεγέθη, όταν αυτό το φορτίο είναι 80%.

13.4 Προχωρημένο Υλικό

Σε αυτή την ενότητα, θα προχωρήσουμε σε περισσότερες λεπτομέρειες σχετικά με τη δρομολόγηση μέσα σε ένα ΑΣ και ανάμεσα στα ΑΣ. Κατόπιν, σε μία άσκηση, θα μεταβούμε επίσης σε ουσιαστικές ιδέες της μεταγωγής μέσα σε ένα μικρό, τοπικό δίκτυο και ένα κατανεμημένο πρωτόκολλο, που υπολογίζει ένα δένδρο συνδέσεων, για να συνδέσει ένα δεδομένο σύνολο από κόμβους.

13.4.1 Δρομολόγηση κατάστασης συνδέσεων: *Open Shortest Path First (OSPF)*

Η επιλογή της συντομότερης διαδρομής είναι μόνο μία μετρική από πολλές άλλες λογικές. Ένα ακόμα δημοφιλές μέτρο είναι η ελαχιστοποίηση των φορτίων των συνδέσεων, όπου το φορτίο μίας σύνδεσης ορίζεται ως το ποσοστό της χωρητικότητας που χρησιμοποιείται. Αυτό εξάγεται από μία διαδικασία που ονομάζεται κυκλοφοριακή μηχανική (traffic engineering) από τους παρόχους Διαδικτύου. Ο στόχος είναι

να εξισορροπήσει το φορτίο της κυκλοφορίας μεταξύ των διαδρομών, έτσι ώστε κανένα φορτίο σύνδεσης να μην είναι πολύ υψηλό και να δημιουργήσει συμφόρηση.

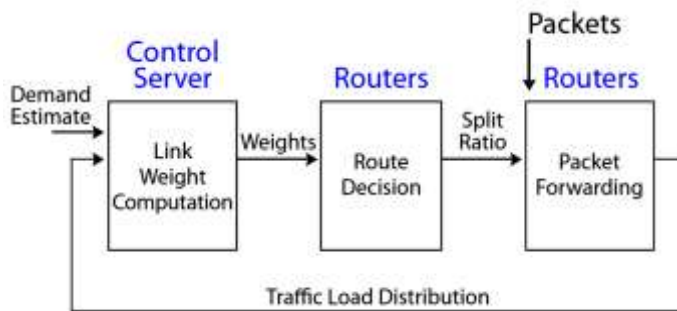
Η μέτρηση της απόδοσης της κυκλοφοριακής μηχανικής ορίζεται από το πρόβλημα της πολυπροϊοντικής ροής (multi-commodity flow). Είναι ένα βασικό πρόβλημα βελτιστοποίησης, που αντιμετωπίζεται σε πολλά δίκτυα, για να σχεδιαστεί μία αντιστοίχιση των ροών σε διαφορετικές διαδρομές με μία δεδομένη τοπολογία $G = (V, E)$.

Θεωρούμε ότι κάθε προορισμός n έχει εισερχόμενη ροή από διάφορους άλλους κόμβους στο δίκτυο. Ας πούμε ότι το $f_{i,j}^n$ υποδηλώνει το ποσό της ροής σε μία σύνδεση (i, j) , που προορίζεται να φτάσει τον κόμβο n και f_{ij} το άθροισμα του φορτίου σε μία σύνδεση (i, j) προς όλους τους προορισμούς. Η αντικειμενική συνάρτηση είναι να ελαχιστοποιηθεί κάποια συνάρτηση κόστους: όσο υψηλότερο είναι το φορτίο, τόσο υψηλότερο είναι το κόστος. Μία τυπική συνάρτηση κόστους σύνδεσης είναι μία τμηματικά γραμμική, αύξουσα και κυρτή συνάρτηση, όπως φαίνεται στο Σχήμα 13.8. Όσο το φορτίο πλησιάζει στη χωρητικότητα της σύνδεσης, το κόστος αυξάνεται ραγδαία. Ο περιορισμός είναι απλά μία εξίσωση συντήρησης ροής: η εισερχόμενη κυκλοφορία σε έναν κόμβο v , συν τη $D(v, n)$, την κυκλοφορία που δημιουργείται στο v και προορίζεται να φτάσει στο n , πρέπει να είναι ίση με την εξερχόμενη κυκλοφορία από τον κόμβο v :

Ελαχιστοποιείστε $\sum_{i,j} C(f_{ij})$

υπό τους περιορισμούς $\sum_{j:(v,j) \in E} f_{v,j}^n = \sum_{i:(i,v) \in E} f_{i,v}^n + D(v, n)$ για όλα τα $v \neq n$ (13.2)

$$f_{ij} = \sum_t f_{ij}^t, \text{ για κάθε } (i, j) \in E.$$



Σχήμα 13.9 Τρεις κύριες μονάδες στα πρωτόκολλα δρομολόγησης κατάστασης συνδέσεων, όπως το OSPF. Ο υπολογισμός των βαρών των συνδέσεων γίνεται σε έναν συγκεντρωτικό εξυπηρετητή διαχείρισης, βασισμένος σε μία εκτίμηση της απαίτησης κυκλοφορίας. Αυτά τα βάρη των συνδέσεων δίνονται κατόπιν στους δρομολογητές,

όπου ο καθένας αποφασίζει αυτόνομα για την αναλογία της διάσπασης για την ισορροπία των φορτίων, με βάση αυτά τα βάρη. Όταν ένα πακέτο φτάσει πραγματικά σε έναν δρομολογητή, η προώθησή του σε ένα επόμενο βήμα γίνεται διαβάζοντας τη διεύθυνση IP του προορισμού του.

Χρησιμοποιώντας την ορολογία από το Κεφάλαιο 4, γνωρίζουμε ότι αυτό το πρόβλημα (13.2) είναι ένα γραμμικά περιορισμένο, κυρτό πρόβλημα βελτιστοποίησης. Είναι πολύ εύκολο να λυθεί, τουλάχιστον μέσω ενός κεντρικοποιημένου αλγόριθμου. Και υπάρχουν πολλοί ειδικευμένοι αλγόριθμοι για την περαιτέρω επιτάχυνση του υπολογισμού.

Αλλά το IP δεν επιτρέπει την από άκρο-προς-άκρο διάνοιξη, κι έτσι δεν μπορεί να παρακολουθήσει το $f_{i,j}^n$. Αυτό είναι μία επιλογή σχεδιασμού που οδηγείται από την απλότητα της διαχείρισης του δικτύου. Ούτε επιτρέπει το IP τη δυναμική προσαρμοσμένη στα φορτία των συνδέσεων δρομολόγηση· αυτό είναι δουλειά του TCP. Αυτό που πραγματικά συμβαίνει είναι ότι το δημοφιλέστερο ενδο-ΑΣ πρωτόκολλο δρομολόγησης, το OSPF, λύνει το παραπάνω πρόβλημα έμμεσα, μέσω της βελτιστοποίησης των βαρών των συνδέσεων. Αυτό φαίνεται στο Σχήμα 13.9 και συνοψίζεται παρακάτω.

- Υπολογισμός βαρών σύνδεσης. Ένας συγκεντρωτικός εξυπηρετητής διαχείρισης συλλέγει ή εκτιμά την κυκλοφορία πηγής-προορισμού, π.χ. όλα τα $D(v, n)$ στην (13.2), κάθε 12 ώρες ή κάτι τέτοιο. Αυτή η χρονική κλίμακα δεν είναι όσο δυναμική όσο η κλίμακα διακύμανσης της κυκλοφορίας. Τότε ο εξυπηρετητής υπολογίζει ένα σύνολο βαρών συνδέσεων, ένα για κάθε σύνδεση στο δίκτυο, και φορτώνει αυτές τις πληροφορίες σε όλους τους δρομολογητές. Κάθε δρομολογητής έχει μία σφαιρική οπτική της τοπολογίας και όχι απλά μία οπτική της γειτονικής τοπολογίας, όπως στη δρομολόγηση διανύσματος αποστάσεων.
- Χρήση βαρών συνδέσεων για να διαχωριστεί η κυκλοφορία. Με δεδομένα τα βάρη που έχουν υπολογιστεί από το συγκεντρωτικό εξυπηρετητή διαχείρισης, κάθε δρομολογητής κατασκευάζει πολλές διαδρομές για κάθε πιθανό προορισμό. Στο OSPF, κάθε δρομολογητής κατασκευάζει μόνο τις συντομότερες διαδρομές με τα δεδομένα βάρη συνδέσεων και διαχωρίζει την εισερχόμενη κυκλοφορία σε ίσα μέρη ανάμεσα σε όλες τις συντομότερες διαδρομές.
- Προώθηση των πακέτων. Καθώς κάθε πακέτο φτάνει στο δρομολογητή, το επόμενο βήμα αποφασίζεται μόνο από τη διεύθυνση IP του προορισμού και η κυκλοφορία που διαχωρίζεται αποφασίζεται στο παραπάνω βήμα. Δεν έχει σημασία ποια είναι η πηγή και ποιους δρομολογητές έχει περάσει ως τώρα το πακέτο. Αυτό ονομάζεται προώθηση βάσει του προορισμού και μετάδοση κατά άλματα.

Μπορούμε να συγκρίνουμε τη δρομολόγηση κατάστασης συνδέσεων, όπως το OSPF, με τη δρομολόγηση διανύσματος απόστασης, όπως το RIP. Η δρομολόγηση κατάστασης συνδέσεων περνάει λεπτομερή μηνύματα για κάθε τοπική τοπολογία, ενώ η δρομολόγηση διανύσματος απόστασης περνάει πιο χονδροειδή μηνύματα για την καθολική τοπολογία. Υπάρχουν επίσης διαφορετικά σημεία ανταλλαγών μεταξύ της επιβάρυνσης στην επικοινωνία και στον τοπικό υπολογισμό μεταξύ των πρωτόκολλων κατάστασης συνδέσεων και διανύσματος απόστασης.

Είναι εύκολο όμως να γυρίσουμε ένα κουμπί βαρών συνδέσεων και να ελπίζουμε ότι τα σωστά βάρη θα προκαλέσουν έμμεσα μία κατανομή της κυκλοφορίας $\{f_{ij}^t\}$ και να λύσουν την (13.2); Η απάντηση είναι όχι. Η επιλογή των βαρών των συνδέσεων για το OSPF, για να επιφέρει μία λύση στην (13.2) αποτελεί ένα υπολογιστικά δυσεπίλυτο πρόβλημα.

Αλλά δεν χρειάζεται να χρησιμοποιήσετε το OSPF. Υπάρχουν και άλλα μέλη αυτής της οικογένειας των πρωτόκολλων δρομολόγησης κατάστασης συνδέσεων, π.χ. το PEF, όπου τα βάρη των συνδέσεων χρησιμοποιούνται για να ορίσουν τα βάρη για όλες τις διαδρομές που θα χρησιμοποιηθούν, όχι μόνο τις συντομότερες διαδρομές, αλλά και τις μεγαλύτερες διαδρομές. Ωστόσο, το κομμάτι της κυκλοφορίας που τοποθετείται σε κάθε διαδρομή μειώνεται εκθετικά, ως συνάρτηση του μήκους της πορείας. Αν τα βάρη των συνδέσεων χρησιμοποιούνται κατ' αυτό τον τρόπο, συμπεραίνουμε ότι ο υπολογισμός των σωστών βαρών των συνδέσεων γίνεται ένα υπολογιστικά επιλύσιμο πρόβλημα. Όπως και στο μηχανισμό σχεδιασμών παιχνιδιών, αυτή είναι μία περίπτωση της αρχής «σχεδιασμός για δυνατότητα βελτιστοποίησης»: ο σχεδιασμός του πρωτόκολλου ενός δικτύου γίνεται ώστε η λειτουργία του να μπορεί εύκολα να βελτιστοποιηθεί.

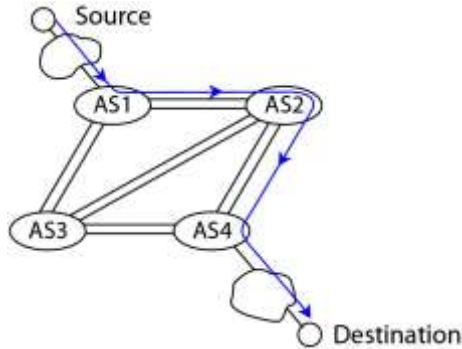
13.4.2 Ενδο-ΑΣ Δρομολόγηση: Το πρωτόκολο συνοριακών πυλών (Border Gateway Protocol (BGP))

Αφού το Διαδίκτυο είναι το δίκτυο των δικτύων, χρειάζεται όλα τα ΑΣ να συνεργάζονται μεταξύ τους, έτσι ώστε οι πελάτες ενός ΑΣ να φτάνουν τους πελάτες ενός άλλου ΑΣ. Η δια-ΑΣ δρομολόγηση συγκολλά το Διαδίκτυο, αλλά δεν είναι καλά οργανωμένη, διότι η καλύτερη διαδρομή (ορισμένη από μέτρηση απόστασης ή βάρους) ανάμεσα στα ΑΣ συχνά δεν είναι αυτή που έχει επιλεγεί λόγω πολιτικών και οικονομικών συμφερόντων.

Θεωρήστε το γράφο μικρής κλίμακας συνδεσιμότητας ΑΣ του Σχήματος 13.10. Κάθε ΑΣ έχει έναν αριθμό, που ονομάζεται ASN, όπως ακριβώς κάθε τελικός κόμβος και κάθε διεπαφή ενός δρομολογητή έχει μία διεύθυνση IP.

Το BGP ελέγχει τη δρομολόγηση ανάμεσα στα ΑΣ (ονομάζεται μία σύνοδος eBGP) και μετακινεί ενδο-ΑΣ πακέτα μέσα σε ένα ΑΣ (ονομάζεται μία σύνοδος iBGP). Η επιλογή του ΑΣ του επόμενου βήματος στο BGP γίνεται περίπου όπως η επιλογή

του δρομολογητή επόμενου βήματος στο RIP, καθώς κάθε ΑΣ μεταδίδει μία λίστα με χαρακτηριστικά BGP στα γειτονικά ΑΣ. Μέρος αυτών των χαρακτηριστικών ονομάζεται ΔΙΑΔΡΟΜΗ-ΑΣ (AS-PATH), που καταγράφει τα ΑΣ που αυτό το ΑΣ πρέπει να διασχίσει, για να φτάσει ένα συγκεκριμένο εύρος διευθύνσεων IP.



Σχήμα 13.10 Ένα παράδειγμα, όπου το BGP δρομολογεί μία σύνοδο σε πολλαπλά ΑΣ στην ίδια βαθμίδα. Κάθε κόμβος είναι ένα ΑΣ και κάθε σύνδεση είναι μία ομότιμη σχέση που εκδηλώνεται μέσω κάποιας φυσικής συνδεσιμότητας ανάμεσα στους συνοριακούς δρομολογητές. Το ποιο ΑΣ θα επιλεγεί ως το επόμενο βήμα δεν εξαρτάται μόνο από τις μετρικές επίδοσης, αλλά επίσης από πολιτικές σε σχέση με οικονομικά συμφέροντα και θέματα ασφάλειας. Το ποιος δρομολογητής θα επιλεγεί για την προώθηση του πακέτου έξω από ένα ΑΣ βασίζεται στη δρομολόγηση καυτής πατάτας (hot-potato).

Εδώ επεισέρχεται το όχι τόσο καλά δομημένο, αλλά σημαντικό κομμάτι του eBGP: κάθε ΑΣ μπορεί να έχει μία ακολουθία από φίλτρα, που ορίζουν εξ αρχής τις προτεραιότητες στις επιλογές των γειτονικών ΑΣ, προκειμένου να φτάσουν σε ένα συγκεκριμένο εύρος διευθύνσεων IP. Στην κορυφή αυτής της λίστας φίλτρων είναι μία τακτική τοπικής προτίμησης: το ΑΣ1 μπορεί να αποφασίσει ότι απλά δεν θέλει να περάσει μέσω του ΑΣ3, έτσι κάθε ΔΙΑΔΡΟΜΗ-ΑΣ που περιλαμβάνει το ΑΣ3 βρίσκεται στον πυθμένα της λίστας προτεραιότητας. Αυτή η απόφαση μπορεί να οφείλεται σε κάποια υπόθεση ασφάλειας ή οικονομικού συμφέροντος: το ΑΣ3 μπορεί να γίνεται αντιληπτό από το ΑΣ1 ως μη ασφαλές ή το ΑΣ1 μπορεί να μη θέλει να στείλει υπερβολική κυκλοφορία στο ΑΣ3, σε περίπτωση που γέρνει την ισορροπία κυκλοφορίας-όγκου μεταξύ τους και φέρνει ως αποτέλεσμα ένα νέο συμβόλαιο πληρωμής ανάμεσα σε αυτά τα δύο ομότιμα ΑΣ.

Συνήθως υπάρχουν πολλαπλοί συνοριακοί δρομολογητές που συνδέουν δύο ΑΣ. Ποιον λοιπόν πρέπει να χρησιμοποιήσουμε; Το iBGP χρησιμοποιεί έναν απλό κανόνα: επέλεξε το συνοριακό δρομολογητή με το ελάχιστο κόστος για να τον φτάσουμε (από εκεί που τοποθετείται το πακέτο μόλις εισέλθει στο ΑΣ). Αυτό σημαίνει ότι κάθε ΑΣ θέλει να ξεφορτωθεί το πακέτο όσο το δυνατό πιο σύντομα. Αυτό ονομάζεται δρομολόγηση καυτής πατάτας (hot-potato).

Υπάρχουν πολλές λεπτομέρειες του BGP γύρω από την παραπάνω κεντρική ιδέα, συμπεριλαμβανομένου ενός ενδιαφέροντος μοντέλου που ονομάζεται πρόβλημα σταθερής διαδρομής το οποίο λαμβάνει υπόψη του όλα τα ζητήματα ευστάθειας του BGP.

Αλλά τώρα θα πρέπει να ολοκληρώσουμε αυτό το κεφάλαιο με τα θεμελιώδη του Διαδικτύου. Υπάρχουν στην πραγματικότητα πολλοί τύποι δρομολόγησης στα δίκτυα επικοινωνίας. Υπάρχουν εξειδικευμένα πρωτόκολλα δρομολόγησης στα ασύρματα δίκτυα πλέγματος και στα οπτικά δίκτυα. Έχει γίνει επίσης πολλή δουλειά για τη δρομολόγηση εγγυημένης ποιότητας των υπηρεσιών, όπως του χρόνου παράδοσης. Μερικά από αυτά τα πρωτόκολλα δρομολόγησης είναι κεντρικοποιημένα ενώ άλλα είναι καταναμημένα. Μερικά έχουν κεντρικοποιημένα επίπεδα για τη βελτιστοποίηση των παραμέτρων, αλλά καταναμημένα επίπεδα δεδομένων που προωθούν τα πραγματικά πακέτα. Μερικά από αυτά τα πρωτόκολλα είναι στατικά, ενώ άλλα είναι δυναμικά, ως μία συνάρτηση των φορτίων των συνδέσεων σε ένα γρήγορο χρονοδιάγραμμα.

Έχουν δε αφιερωθεί πολλά έτη εργασίας και στην ερευνητική κοινότητα και στη βιομηχανία για την εφαρμογή της δρομολόγησης πολλαπλής διανομής (multicast routing). Κοιτάξαμε μόνο τη δρομολόγηση μονής διαδρομής (unicast routing): από έναν αποστολέα σε έναν παραλήπτη. Καθώς η χρήση του Διαδικτύου αυξάνεται για τη διανομή περιεχομένου και την παράδοση ψυχαγωγικών βίντεο, συχνά υπάρχουν πολλοί παραλήπτες την ίδια στιγμή. Η δημιουργία μίας συνόδου δρομολόγησης μονής διαδρομής για κάθε έναν από τους παραλήπτες είναι ανέφικτη. Αλλά η πολλαπλή διανομή στο επίπεδο IP αποδεικνύεται δύσκολο να διαχειριστεί και συχνά αντικαθίσταται από μία εναλλακτική αρχιτεκτονική: την πολλαπλή διανομή στο επίπεδο εφαρμογής με ένα δίκτυο επικάλυψης. Θα αναλύσουμε το θέμα αυτό στο Κεφάλαιο 15, στη συζήτηση για το P2P.

Σύνοψη

Πλαίσιο 13: Αρχιτεκτονικές αρχές του Διαδικτύου

Η μεταγωγή πακέτου για κλιμακούμενη συνδεσιμότητα, τα αρθρωτά διαστρωματωμένα πρωτόκολλα και η καταναμημένη ιεραρχία με επικάλυψη είναι τρεις βασικές ιδέες του Διαδικτύου. Η δρομολόγηση με χρήση του IP στο επίπεδο δικτύου γίνεται διαφορετικά, βάσει του αν είναι μέσα σε ένα τομέα ή ανάμεσα σε τομείς. Η δρομολόγηση μέσα σε έναν τομέα έχει δύο κύριους τύπους: του διανύσματος απόστασης ή της κατάστασης σύνδεσης. Η δρομολόγηση διανύσματος απόστασης εφαρμόζει τον αλγόριθμο Bellman-Ford, μέσω της μετάδοσης καταναμημένων μηνυμάτων.

Περαιτέρω Μελέτη

1. Υπάρχει μία ολόκληρη βιβλιοθήκη με εγχειρίδια για τη δικτύωση υπολογιστών και το σχεδιασμό του Διαδικτύου και χιλιάδες εργασίες για όλα τα είδη των απόψεων για τη δρομολόγηση.
2. Το παρακάτω βιβλίο παρέχει μία κατανοητή και περιεκτική επισκόπηση των πρωτόκολλων δρομολόγησης του Διαδικτύου:
C. Huitema, *Routing in the Internet*, 2nd Ed., Prentice Hall, 1999
3. Η δρομολόγηση είναι ένα δημοφιλές θέμα στα μαθήματα αλγορίθμων στα τμήματα επιστήμης υπολογιστών. Το παρακάτω είναι ένα από τα πρότυπα εγχειρίδια στο θέμα:
T. H. Cormen, C. E. Leiserson, R. L. Rivest and C. Stein, *Introduction to Algorithms*, 3rd edn., MIT Press, 2009
4. Ένα από τα πολλά πρότυπα εγχειρίδια στη δικτύωση των υπολογιστών, που καλύπτει τη δρομολόγηση με περισσότερες λεπτομέρειες είναι:
L. L. Peterson and B. Davie, *Computer Networks: A Systems Approach*, 5th edn., Morgan Kauffman, 2012
5. Παρακάτω παρατίθεται ένα πρόσφατο ερευνητικό άρθρο, που απέδειξε τη δυναμικότητα επίτευξης του αποτελεσματικού υπολογισμού των βέλτιστων βαρών συνδέσεων στη δρομολόγηση κατάστασης συνδέσεων με προώθηση κατά άλματα:
D. Xu, M. Chiang, and J. Rexford, “Link state routing protocol can achieve optimal traffic engineering,” *IEEE/ACM Transactions on Networking*, vol. 19, no. 6, pp. 1717-1730, November 2011
6. Δεν είχαμε χρόνο να καλύψουμε το BGP λεπτομερώς. Η παρακάτω αναφορά είναι ένα σχετικά πρόσφατο ερευνητικό άρθρο για το BGP:
M. Caesar and J. Rexford, “BGP routing policies in ISP networks,” *IEEE Network*, vol. 19, no.6, pp. 5-11, November 2005

Ασκήσεις

13.1 Μεταγωγή πακέτου

Σε αυτό το πρόβλημα θα εξετάσουμε ποσοτικά τα δύο πλεονεκτήματα της μεταγωγής πακέτου.

α. Στατιστική πολύπλεξη:

Υποθέστε ότι έχουμε μία σύνδεση 10 Mbps, που διαμοιράζεται σε πολλούς χρήστες. Κάθε χρήστης της σύνδεσης παράγει 1 Mbps δεδομένων στο 10% του χρόνου και στο 90% του χρόνου είναι ανενεργός.

Αν χρησιμοποιήσουμε ένα δίκτυο μεταγωγής κυκλώματος και η κατανομή εύρους ζώνης είναι ίση για τους χρήστες, πόσους χρήστες μπορεί να υποστηρίξει η σύνδεση; Ας πούμε αυτόν τον αριθμό N . Τώρα θεωρήστε ένα δίκτυο μεταγωγής πακέτου. Ας πούμε ότι έχουμε M χρήστες συνολικά και θέλουμε την πιθανότητα να αρνηθεί υπηρεσία σε κάποιον χρήστη να είναι μικρότερη από 1%. Γράψτε την έκφραση που πρέπει να λυθεί με τη μορφή $f(M, N) < 0,01$. Λύστε την αριθμητικά για το M .

(Υπόδειξη: Χρησιμοποιήστε την αθροιστική συνάρτηση κατανομής της διωνυμικής κατανομής.

β. Συγκέντρωση πόρων (Resource pooling):

Θα θεωρήσουμε ότι μοντελοποιούμε έναν κοινό πόρο και θα δούμε τι συμβαίνει όταν τόσο η ζήτηση για το συγκεκριμένο πόρο όσο και η δυναμικότητα να εκπληρωθούν τα αιτήματα αυξάνονται. Όταν εισέρχεται μία αίτηση, ένας ανενεργός εξυπηρετητής απαντά στην αίτηση. Αν όλοι οι εξυπηρετητές είναι απασχολημένοι, η αίτηση απορρίπτεται. Ο παρακάτω μαθηματικός τύπος του Erlang δίνει την πιθανότητα να απορριφθεί μία αίτηση, με δεδομένους m εξυπηρετητές και E μονάδες κυκλοφορίας:

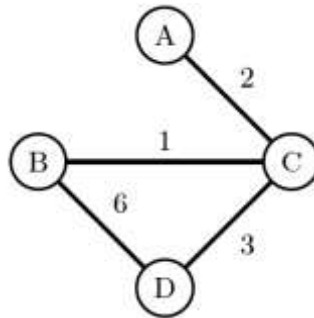
$$P(E, m) = \frac{\frac{E^m}{m!}}{\sum_{i=0}^m \frac{E^i}{i!}}.$$

Υπολογίστε το $P(3,2)$. Τώρα υπολογίστε το $P(6,4)$. Τι παρατηρείτε; Γενικά $P(wx, wy) < P(x, y)$, $\forall w > 1$. Αυτός είναι ένας από τους αρκετά συνήθεις τρόπους να εκτιμήσουμε την ιδέα των πλεονεκτημάτων της συγκέντρωσης πόρων.

13.2 Πρωτόκολλο Πληροφοριών Δρομολόγησης (Routing Information Protocol (RIP))

Θεωρήστε ένα δίκτυο με την τοπολογία, όπως φαίνεται στην Εικόνα 13.2.

α. Τρέξτε ένα παράδειγμα του RIP σε αυτό το δίκτυο για να βρείτε τις ελάχιστες διαδρομές ανάμεσα σε όλους τους κόμβους. Δείξτε τους πίνακες δρομολόγησης για κάθε βήμα.



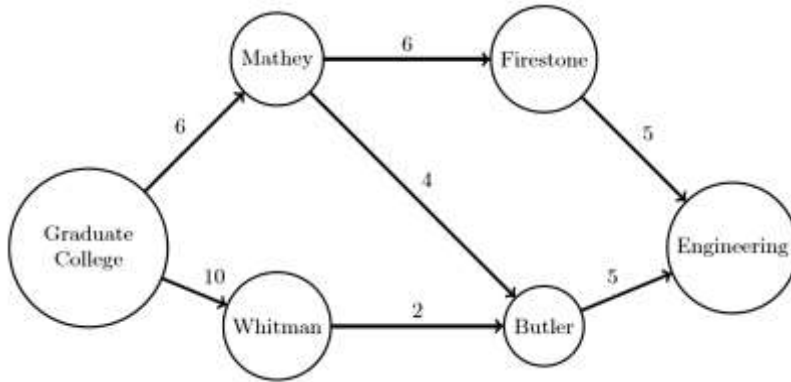
Σχήμα 13.11 Ένας μικρός γράφος για ένα παράδειγμα του RIP

- β. Τώρα η σύνδεση ανάμεσα στο A και το C καταρρέει, με αποτέλεσμα ένα κόστος ίσο με ∞ για τις δύο κατευθύνσεις της μεταφοράς. Το B και το C αντιλαμβάνονται αμέσως την κατάρρευση της σύνδεσης και ανανεώνουν τους δικούς τους πίνακες δρομολόγησης για τέσσερις επαναλήψεις μετά την αποτυχία της σύνδεσης. Πρέπει να δείξετε μόνο τους πίνακες δρομολόγησης που άλλαξαν. Τι συμβαίνει με τις διαδρομές στο A;
- γ. Προτείνετε μία λύση για το πρόβλημα που εμφανίστηκε στο (β).

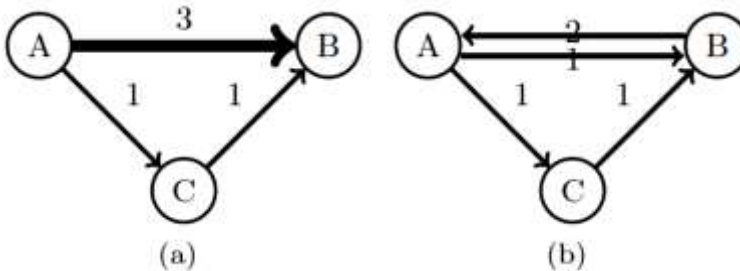
13.3 Ο αλγόριθμος Ford-Fulkerson και το πρόβλημα μέγιστης ροής

Είστε στη βιβλιοθήκη της Πολυτεχνικής Σχολής στο Princeton University και μελετάτε για τις τελικές εξετάσεις, που θα ξεκινήσουν σε 1 ώρα. Ξαφνικά αντιλαμβάνεστε ότι δεν παρευρεθήκατε σε μία σημαντική διάλεξη για τον έλεγχο ισχύος. Η ευγενική βοήθος του καθηγητή προσφέρεται να σας στείλει το βίντεο της διάλεξης. Δυστυχώς, μένει στο Graduate College, που είναι κάπου μακριά από την πανεπιστημιούπολη (δεν ξέρετε καν πού).

Καθώς θέλετε να πάρετε το βίντεο όσο το δυνατόν συντομότερα, αποφασίζετε να το χωρίσετε σε πολλά κομμάτια πριν το στείλετε στο δίκτυο του Princeton. Υποθέστε ότι το δίκτυο του Princeton έχει τις χωρητικότητες αγωγών που δίνονται στην Εικόνα 13.12. Πόσο όγκο δεδομένων θα πρέπει να στείλετε σε κάθε αγωγό, ώστε να μεγιστοποιήσετε το μέγιστο ρυθμό σας;



Σχήμα 13.12 Ένα απλοποιημένο δίκτυο της πανεπιστημιούπολης του Princeton, με έξι κόμβους με συντομεύσεις G, M, F, E, B και W.



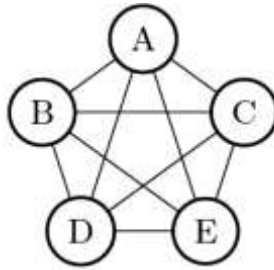
Σχήμα 13.13 Ένα παράδειγμα σχεδίασης ενός γράφου κατάλοιπων. Αποφασίζουμε να προωθήσουμε δύο μονάδες ροής από το A στο B στο γράφο (a), που μας δίνει το γράφο κατάλοιπων στο (b).

Θα περάσουμε μέσα από μία βήμα με βήμα προσέγγιση για τη λύση αυτού του προβλήματος μέγιστης ροής. Ένα χρήσιμο εγχείρημα θα είναι να παραγάγουμε ένα γράφο κατάλοιπων (residual graph). Δεδομένης μίας ροής, για κάθε σύνδεση κατά μήκος της διαδρομής της ροής, σχεδιάζουμε μία σύνδεση αντίθετης φοράς με το ποσό της ροής, αφήνοντας μία σύνδεση προς τα εμπρός με την υπολειπόμενη χωρητικότητα της σύνδεσης. Ένα παράδειγμα φαίνεται στο Σχήμα 13.13.

- Να κατανεύμετε 4 Mbps στη διαδρομή G-M-B-E. Σχεδιάστε το γράφο κατάλοιπων.
- Να κατανεύμετε 2 Mbps στη διαδρομή G-W-B-M-F-E στο γράφο κατάλοιπων από το (a). Σχεδιάστε το νέο υπολειπόμενο γράφο.
- Να κατανεύμετε 2 Mbps στη διαδρομή G-M-F-E στο γράφο κατάλοιπων από το (b). Σχεδιάστε το νέο υπολειπόμενο γράφο.

δ. Δεν υπάρχουν άλλες διαδρομές στο γράφο κατάλοιπων από το (c), κι έτσι ο αλγόριθμος τερματίζεται. Η κατανομή της χωρητικότητας δίνεται από τη χωρητικότητα του δικτύου από τα βήματα (a), (b) και (c). Σχεδιάστε το γράφο με την τελική κατανομή χωρητικότητας σε κάθε έναν από τους κόμβους.

Αυτός ο κλασικός αλγόριθμος ονομάζεται αλγόριθμος Ford-Fulkerson. Περισσότερα για αυτό στο Κεφάλαιο 16.



Σχήμα 13.14 Ένας μικρός γράφος πλήρους πλέγματος για την απεικόνιση του DAR

13.4 Δυναμική εναλλακτική δρομολόγηση

Δεν θα έχουμε την ευκαιρία να μιλήσουμε για τη δρομολόγηση σε δίκτυα μεταγωγής κυκλώματος. Μία σημαντική τεχνική εκεί είναι η Δυναμική Εναλλακτική Δρομολόγηση (Dynamic Alternative Routing – DAR), που υιοθετήθηκε από την British Telecom το 1996 για τα δίκτυά της.

Υποθέστε ότι, αντί να έχουμε σταθερές διαδρομές για να δρομολογήσουμε τα πακέτα από την πηγή στον προορισμό, θέλουμε οι δρομολογήσεις να αλλάζουν δυναμικά ως απάντηση στις συνθήκες του δικτύου. Θα υπολογίσουμε ένα τέτοιο πρωτόκολλο, το DAR, στην περίπτωση ενός πλήρως συνδεδεμένου γράφου.

Θέλουμε η δρομολόγηση να προσαρμόζεται δυναμικά στην υλοποίηση της σύνδεσης και να επιλέγει μία νέα διαδρομή, αν η τρέχουσα είναι πολύ απασχολημένη. Κάθε πιθανή σύννοδος (ζεύγος πηγής-προορισμού) έχει έναν συνδεδεμένο εφεδρικό κόμβο. Όταν αρχίζει μία σύννοδος, προσπαθεί στην αρχή να στείλει την κυκλοφορία της κατά μήκος μίας απευθείας σύνδεσης. Αν η απευθείας σύνδεση αποτύχει επειδή είναι πλήρης, η σύννοδος προσπαθεί να χρησιμοποιήσει τη διαδρομή δύο βημάτων με τον εφεδρικό κόμβο. Αν η εφεδρική πορεία αποτύχει και αυτή επειδή είναι απασχολημένη, η σύννοδος αποτυγχάνει και επιλέγει ένα νέο εφεδρικό κόμβο από το δίκτυο. Προφανώς ο εφεδρικός κόμβος δεν πρέπει να είναι ο ίδιος κόμβος με τον προορισμό της συνόδου.

Ένα πιθανό πρόβλημα με αυτό το σχήμα είναι ότι μπορεί πολλές σύννοδοι να τερματίσουν χρησιμοποιώντας διαδρομές δύο αλμάτων. Τότε δεν θα είμαστε και πολύ αποτελεσματικοί, καθώς χρησιμοποιούμε τη διπλάσια χωρητικότητα σε σχέση με τη

διαδρομή ενός άλματος. Γι' αυτό, το DAR κρατά ένα κομμάτι t_l από κάθε σύνδεση l για απευθείας συνόδους ενός άλματος (σύνοδοι που χρησιμοποιούν την απευθείας σύνδεση ανάμεσα στην πηγή και τον προορισμό, που ανατίθεται στην εφεδρική πορεία). Ονομάζουμε αυτήν την παράμετρο, $t_l \in (0,1)$, συντελεστή κράτησης διαύλων (trunk reservation coefficient). Κάθε σύνδεση l στο δίκτυο έχει μία συνολική δυναμικότητα c_l Mbps και είναι αμφίδρομη. Η μη-δεσμευμένη χωρητικότητα, $c_l - t_l c_l$, μπορεί να χρησιμοποιηθεί για απευθείας ή για έμμεσες συνόδους.

- α. Υποθέστε ότι έχουμε ένα πλήρες πλέγμα όπως φαίνεται στο Σχήμα 13.14. Ας πούμε ότι $c_l = 10$ Mbps, $t_l = 0,1 \forall l$. Οι εφεδρικοί κόμβοι αρχικοποιούνται στον Πίνακα 13.1.

Πίνακας 13.1 Εφεδρικοί κόμβοι για ένα παράδειγμα DAR

Σύνοδος	Εφεδρικός κόμβος
(B,C)	A
(C,B)	A
(A,C)	E

Πίνακας 13.2 Ο πίνακας των υλοποιήσεων των συνδέσεων για να συμπληρωθεί στο παράδειγμα DAR.

Σύνδεση	Μη κρατημένη δυναμικότητα που χρησιμοποιείται (σύνοδος)	Κρατημένη δυναμικότητα που χρησιμοποιείται (σύνοδος)
(A,B)		
(A,C)		
(A,E)		
(B,C)	9 Mbps [9 × (B, C)]	1 [1 × (B, C)]
(C,E)		

Τα παρακάτω γεγονότα συμβαίνουν σε σειρά.

1. Ξεκινούν δέκα παράλληλες σύνοδοι των (B,C).
2. Ξεκινούν δέκα παράλληλες σύνοδοι των (C,B).
3. Ξεκινούν δέκα παράλληλες σύνοδοι των (A,C).

Υποθέστε ότι οι σύνοδοι κρατούν πολλή ώρα και κάθε σύνοδος καταναλώνει 1 Mbps. Συμπληρώστε στον πίνακα 13.2 μετά την παραπάνω σειρά γεγονότων. Να θυμάστε ότι οι σύνοδοι και οι συνδέσεις είναι δύο διαφορετικές έννοιες. Μία σειρά έχει συμπληρωθεί για εσάς ως παράδειγμα.

β. Επαναλάβετε το (α) χωρίς το σχήμα κράτησης διαύλων.

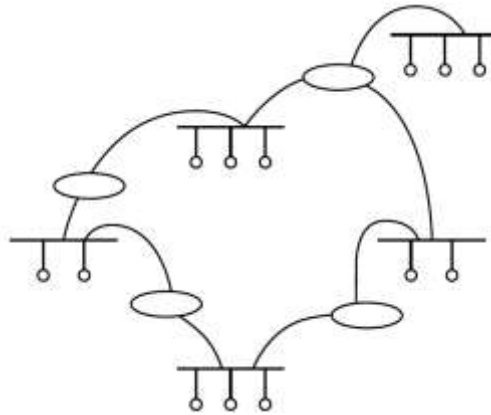
γ. Ποια είναι η αποδοτικότητα της χρησιμοποίησης των συνδέσεων, δηλαδή ο αριθμός των συνόδων διαιρεμένος με τη συνολική χωρίζονται του δικτύου που χρησιμοποιείται στο (α) και στο (β) αντίστοιχα;

(Για περισσότερες πληροφορίες ανατρέξτε στο R. J. Gibbens, F. P. Kelly, and P. B. Key, "Dynamic alternative routing," Routing in Communication Networks, M. Steenstrup Eds., Prentice Hall, 1995)

13.5 Επικαλύπτων Δένδρο

Η δρομολόγηση σε ένα διασυνδεδεμένο σύνολο από τοπικά δίκτυα (LANs), όπως αυτά σε μία επιχείρηση ή σε ένα πανεπιστήμιο, είναι πολύ ευκολότερη από ό,τι σε όλο το Διαδίκτυο και συνδέει συνολικά συσκευές με πολλές διαφορετικές τεχνολογίες επιπέδων 1 και 2. Οι συνδέσεις μεταξύ των LAN ονομάζονται γέφυρες ή μεταγωγείς (switch). Κάθε μία έχει πολλαπλές θύρες, με μία θύρα να συνδέεται σε ένα LAN. Θα μπορούσαμε να ρυθμίσουμε τα αναγνωριστικά του καθενός τοπικού δικτύου (LAN ID) σε κάθε θύρα της γέφυρας, αλλά ένας πιο εξελικτικός και αυτόματος τρόπος είναι για κάθε γέφυρα να «ακούσουμε» την άφιξη του κάθε πακέτου στη θύρα και να αντιγράψουμε τη διεύθυνση της πηγής στην επικεφαλίδα αυτού του πακέτου στη βάση δεδομένων των υπολογιστών (hosts), που είναι εφικτοί από αυτή τη θύρα.

Σε αυτή την άσκηση, ωστόσο, θα καταπιαστούμε με ένα απλούστερο αλλά σημαντικό πρόβλημα, να ανακαλύψουμε καταναμημένα ένα επικαλύπτων δένδρο (όχι απαραίτητα το ελάχιστο δένδρο συνδέσεων). Το πρωτόκολλο εφευρέθηκε το 1985 από τον Perlman. Όπως τα πρωτόκολλα δρομολόγησης κατάστασης συνδέσεων, θα δούμε παρακάτω ότι το πρωτόκολλο δένδρου συνδέσεων είναι ένα παράδειγμα του επιτεύγματος μίας συνολικής συνεπούς άποψης της τοπολογίας, μόνο μέσω τοπικών αλληλεπιδράσεων, που τελικά μεταδίδονται μέσα από το δίκτυο.



Σχήμα 13.15 Ένα παράδειγμα τοπικών δικτύων συνδεδεμένων με γέφυρες. Υπάρχουν δύο τύποι κόμβων σε αυτόν το γράφο: κάθε τοπικό δίκτυο (μία γραμμή με τους υπολογιστές υποδοχής της (*hosts*) να αναπαρίστανται με μικρούς κύκλους) είναι ένας κόμβος και κάθε γέφυρα (ένα οβάλ) είναι επίσης ένας κόμβος. Οι συνδέσεις σε αυτό το γράφο συνδέουν γέφυρες με τοπικά δίκτυα. Αν χρησιμοποιούνται όλες οι γέφυρες, ο γράφος γίνεται κυκλικός. Τα κατανεμημένα πρωτόκολλα δένδρου συνδέσεων ανακαλύπτουν ένα ακυκλικό υπογράφο, ο οποίος περιλαμβάνει όλα τα τοπικά δίκτυα, αλλά όχι όλες τις συνδέσεις.

Θεωρήστε ένα σύνολο από τμήματα LAN (οι οριζόντιες γραμμές, η κάθε μία με μερικές συσκευές συνδεδεμένες πάνω τους) και γέφυρες (οι οβάλ), που απεικονίζονται στο Σχήμα 13.15. Προφανώς, υπάρχουν κύκλοι στο γράφο: μπορούμε να πάμε από ένα τμήμα μέσω άλλων τμημάτων και πίσω στον εαυτό του, χωρίς να περάσουμε από την ίδια σύνδεση δύο φορές. Θέλουμε να καθορίσουμε έναν τρόπο, χωρίς κύκλους, για να παρέχουμε σύνδεση σε όλα τα τμήματα. Ένας τρόπος για να φτάσουμε σε ένα συνεπές δέντρο συνδέσεων είναι να βάλουμε τη γέφυρα με το μικρότερο αριθμό αναγνώρισης (ID) ως ρίζα του δένδρου και κάθε μία από τις άλλες γέφυρες να φτάνουν τη γέφυρα-ρίζα μέσω της διαδρομής μικρότερου αριθμού αλμάτων. Αυτό είναι εύκολο, τουλάχιστον για τόσο μικρά δίκτυα, αν έχουμε μία γενική άποψη. Αλλά πώς μπορούμε να το κάνουμε αυτό κατανεμημένα, μόνο με μετάδοση μηνυμάτων μεταξύ γειτόνων; Πώς μπορούν οι κόμβοι έστω να συμφωνήσουν ποια γέφυρα είναι η ρίζα του δένδρου;

Μία πιθανότητα είναι να ζητήσουμε από κάθε γέφυρα να ανακοινώσει το παρακάτω μήνυμα, που αποτελείται από τρία πεδία, κατά τη διάρκεια κάθε χρονοθυρίδας:

- το ID της γέφυρας που πιστεύει ότι είναι η ρίζα,
- τον αριθμό των βημάτων για να φτάσει τη γέφυρα-ρίζα από αυτή τη γέφυρα, και

- το ID αυτής της γέφυρας.

Στη συνέχεια, έρχονται οι ερωτήσεις.

- Αρχικά, κάθε γέφυρα έχει μόνο τοπικές πληροφορίες για τον εαυτό της. Ποια είναι τα μηνύματα από τις τέσσερις γέφυρες στο Σχήμα 13.15;
- Όταν λάβει ένα μήνυμα, κάθε γέφυρα επιλέγει τη γέφυρα-ρίζα και ανακαλύπτει τον τρόπο να την προσεγγίσει, εφαρμόζοντας την παρακάτω ταξινομημένη λίστα με κριτήρια.
 - Η γέφυρα με το χαμηλότερο αριθμό ID κερδίζει και γίνεται η νέα γέφυρα-ρίζα (όπως πιστεύεται από αυτή τη γέφυρα)
 - Αν υπάρχουν πολλαπλές διαδρομές για να προσεγγίσει την ίδια γέφυρα-ρίζα, η διαδρομή με το μικρότερο αριθμό βημάτων κερδίζει.
 - Αν υπάρχουν πολλαπλές διαδρομές ίσου αριθμού βημάτων για να προσεγγίσει την ίδια γέφυρα-ρίζα, η διαδρομή που ανακοινώνεται από τη γέφυρα με το μικρότερο αριθμό ID κερδίζει.

Κάθε γέφυρα τότε ανανεώνει το πεδίο γέφυρας-ρίζας του μηνύματος και αυξάνει τον αριθμό βημάτων κατά 1. Τότε στέλνει το νέο μήνυμα προς τους γείτονές της, εκτός φυσικά από αυτούς τους γείτονες που έχουν μικρότερη διαδρομή προς την ίδια γέφυρα-ρίζα.

Γράψτε την εξέλιξη των μηνυμάτων για τις γέφυρες του Σχήματος 13.15. Συγκλίνει προς ένα δένδρο συνδέσεων;

- Ακόμα και μετά τη σύγκλιση, η γέφυρα-ρίζα συνεχίζει να στέλνει το μήνυμα κατά τακτικά διαστήματα. Γιατί συμβαίνει αυτό; Υποθέστε τι συμβαίνει όταν η γέφυρα καταρρέει.

(Το παρακάτω βιβλίο παρέχει μία εκτενή έρευνα της μεταγωγής επιπέδου 2 και των πρωτόκολλων δρομολόγησης επιπέδου 3. Είναι επίσης ένα σπάνιο παράδειγμα βιβλίου που είναι προσηλωμένο στα πρωτόκολλα δικτύων και συνάμα διατηρεί την αίσθηση του χιούμορ: δείτε το Κεφάλαιο 8 για παράδειγμα: R. Perlman, *Interconnections: Bridges, Routers, Switches and Internetworking Protocols*, 2nd edn., Addison-Wesley, 1999).