

ΚΕΦΑΛΑΙΟ 1

Τα κίνητρα για μια νέα αρχιτεκτονική δικτύων

Μια φορά κι έναν καιρό, υπήρχε αυτό που υπήρχε, και αν τίποτα δεν είχε συμβεί, δεν θα είχαμε τίποτα να πούμε.

—Charles de Lint

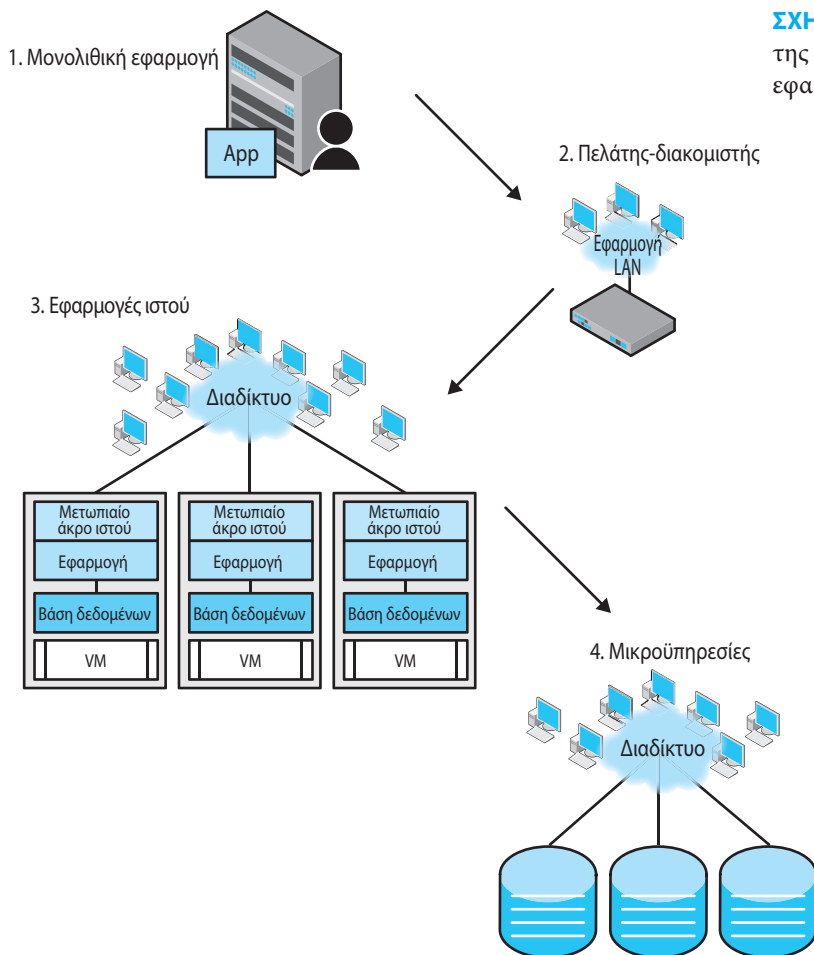
Εάν οι εφαρμογές δεν άλλαζαν ποτέ, θα υπήρχε κάποια άλλη ιστορία να πούμε από αυτήν που κρατάτε τώρα στα χέρια σας. Μια κατανεμημένη εφαρμογή συνθέτει μια χορογραφία με το δίκτυο, και αυτή που οδηγεί τα βήματα είναι η εφαρμογή. Η ιστορία του σύγχρονου δικτύου κέντρου δεδομένων ξεκίνησε όταν η εφαρμογή άρχισε να χορεύει σε διαφορετικό ρυθμό και το δίκτυο δεν μπόρεσε να ακολουθήσει. Η κατανόηση αυτής της μετάβασης βοηθά με πολλούς τρόπους, και όχι μόνο για να μάθουμε γιατί είναι απαραίτητη μια αλλαγή. Συχνά συναντώ πελάτες που θέλουν να δημιουργήσουν ένα σύγχρονο δίκτυο αλλά μένουν προσκολλημένοι σε παλιούς τρόπους σκέψης. Οι προγραμματιστές εφαρμογών που προέρχονται από την πλευρά της επιχείρησης τείνουν επίσης να σκέφτονται με τρόπους που δεν συνάδουν με το σύγχρονο κέντρο δεδομένων.

Αυτό το κεφάλαιο θα σας βοηθήσει να δώσετε απαντήσεις σε ερωτήσεις όπως οι παρακάτω:

- Ποια είναι τα χαρακτηριστικά των νέων εφαρμογών;
- Τι είναι το δίκτυο πρόσβασης-συνάθροισης-κορμού;
- Γιατί δεν κατάφεραν να καλύψουν αυτές τις εφαρμογές τα δίκτυα πρόσβασης-συνάθροισης-κορμού;

Το ανακάτεμα εφαρμογής-δικτύου

Το Σχήμα 1.1 παρουσιάζει την εξέλιξη των εφαρμογών από μεμονωμένους μονόλιθους (monolith) στο εξαιρετικά κατανεμημένο μοντέλο των μικροϋπηρεσιών (micro-service). Με κάθε διαδοχικό βήμα στην εξέλιξη, εξελίσσονται επίσης οι απαιτήσεις από τα δίκτυα.



ΣΧΗΜΑ 1.1 Η εξέλιξη της αρχιτεκτονικής εφαρμογών

Στην εποχή των μεμονωμένων μονολιθικών εφαρμογών, οι οποίες συνήθως εκτελούνταν σε έναν κεντρικό υπολογιστή (mainframe), οι απαιτήσεις δικτύωσης ήταν ελάχιστες. Οι διασυνδέσεις ήταν ιδιόκτητες, όπως και τα πρωτόκολλα (π.χ., SNA ή DECnet). Συμβαδίζοντας με το εκκολαπτόμενο στάδιο της δικτύωσης και των κατανεμημένων εφαρμογών, οι απαιτήσεις για εύρος ζώνης από το δίκτυο ήταν αμελητέες βάσει των σημερινών δεδομένων.

Η επόμενη γενιά εφαρμογών προέκυψε με την εξάπλωση των σταθμών εργασίας και των προσωπικών υπολογιστών (Personal Computer – PC). Αυτό το κύμα χαρακτηρίστηκε από την ανάπτυξη της αρχιτεκτονικής πελάτη-διακομιστή (client-server architecture). Τα δίκτυα άρχισαν να γίνονται πιο εξεζητημένα, καθώς χαρακτηρίζονταν από την άνοδο του τοπικού δικτύου (LAN, local area network). Η απαίτηση δικτύωσης αυξήθηκε, αν και εξακολουθούσε να είναι ασήμαντη βάσει των σημερινών προτύπων. Με την πάροδο του χρόνου, τα δεδομένα που ανταλλάσσονταν μεταξύ

εφαρμογών εξελίχθηκαν από μόνο κείμενο και εικόνες σε ήχο και βίντεο. Όπως αρμόζει σε μια αναπτυσσόμενη τεχνολογία, οι διασυνδέσεις και τα πρωτόκολλα αφθονούσαν και ήταν κυρίως ιδιόκτητα. Το Ethernet, το δακτυλιοειδές δίκτυο αδειοδοτικού (Token Ring) και η διεπαφή δεδομένων κατανεμημένων μέσω οπτικών ινών (FDDI, Fiber Distributed Data Interface) ήταν οι πιο δημοφιλείς διασυνδέσεις. Οι ταχύτητες διασυνδέσεων έφταναν τα 100 Mbps. Πρωτόκολλα όπως το IPX της Novell και το VINES της Banyan Systems ανταγωνίζονταν το IP για την κυριαρχία στην κατηγορία πρωτοκόλλων δικτύου ανώτερου επιπέδου. Οι επιχειρήσεις ήταν επιφυλακτικές για τη στοίβα πρωτοκόλλων TCP/IP που προσφερόταν δωρεάν με ανοιχτό κώδικα. Πολλοί στοιχημάτιζαν υπέρ των αντίστοιχων εμπορικών προϊόντων. Οι εφαρμογές ήταν περιτειχισμένες στις επιχειρήσεις, χωρίς δυνατότητα πρόσβασης από έξω, διότι τα περισσότερα πρωτόκολλα ανώτερου επιπέδου δεν ήταν σχεδιασμένα για χρήση σε μεγάλη κλίμακα ή κατά μήκος επιχειρήσεων. Σήμερα, όλα αυτά τα εμπορικά πρωτόκολλα ανώτερου επιπέδου αποτελούν πλέον υποσημειώσεις στην ιστορία της δικτύωσης. Νικητές ανακηρύχθηκαν το Ethernet και το TCP/IP. Τα ανοιχτά πρότυπα και ο ανοιχτός κώδικας σημείωσαν την πρώτη τους νίκη έναντι των ιδιόκτητων αντίστοιχων προϊόντων.

Στη συνέχεια ήρθαν η πανταχού παρουσία του διαδικτύου και η στέψη της στοίβας πρωτοκόλλων TCP/IP. Οι εφαρμογές έσπασαν τα δεσμά των επιχειρήσεων και άρχισαν να σχεδιάζονται για πρόσβαση από όλο τον κόσμο. Με τη δυνατότητα των πελατών να συνδέονται από οπουδήποτε, η κλίμακα που έπρεπε να ικανοποιήσουν οι διακομιστές των εφαρμογών αυξήθηκε σημαντικά. Στις περισσότερες επιχειρήσεις δεν ήταν δυνατόν ένα στιγμιότυπο διακομιστή να αντεπεξέλθει στον φόρτο που δημιουργούνταν από όλους τους πελάτες. Επίσης, αυτά τα μεμονωμένα στιγμιότυπα διακομιστή έγιναν ένα μοναδικό σημείο αστοχίας. Ως εκ τούτου, έγινε δημοφιλής η λειτουργία πολλαπλών στιγμιότυπων ενός διακομιστή, τα οποία τοποθετούνταν πίσω από έναν μηχανισμό εξισορρόπησης φόρτου. Οι ίδιοι οι διακομιστές χωρίστηκαν σε πολλαπλές μονάδες, καλύπτοντας συνήθως το μετωπιαίο άκρο ιστού, την εφαρμογή και τη βάση δεδομένων ή τον χώρο αποθήκευσης.

Το Ethernet έγινε η de facto διασύνδεση εντός της επιχείρησης. Μαζί με το TCP/IP, το Ethernet διαμόρφωσε το έδαφος για τα δίκτυα επιχειρήσεων και τις εφαρμογές τους. Ακόμα θυμάμαι την ημέρα στα μέσα ή τα τέλη του 1998 που η διοίκηση της εταιρείας όπου εργαζόμουν ανακοίνωσε το τέλος της γραμμής ανάπτυξης για το Token Ring και τη FDDI στην οικογένεια μεταγωγέων Catalyst 5000 και 6500. Η επιτυχία του Ethernet στις επιχειρήσεις άρχισε να διαβρώνει τη δημοτικότητα των υπόλοιπων διασυνδέσεων, όπως του σύγχρονου οπτικού δικτύου (SONET, Synchronous Optical Network), στα δίκτυα των εταιρειών παροχής υπηρεσιών. Στο τέλος αυτού του κύματος, το Gigabit Ethernet είχε γίνει η πιο δημοφιλής διασύνδεση για δίκτυα επιχειρήσεων.

Καθώς η υπολογιστική ισχύς συνέχισε να αυξάνεται, ήταν αδύνατο για τις περισσότερες εφαρμογές να αξιοποιήσουν το σύνολο της επεξεργαστικής ισχύος της κεντρικής μονάδας επεξεργασίας (CPU, Central Processing Unit). Αυτή η αφθονία υπο-

λογιστικής ισχύος προκάλεσε κορεσμό όταν οι βελτιώσεις στην κατασκευή των επεξεργαστών πρόσφεραν τη δυνατότητα να υπάρχουν περισσότεροι από ένας πυρήνας CPU σε ένα κύκλωμα επεξεργαστή.

Η *εικονικοποίηση διακομιστή* (server virtualization) ήταν μια έννοια που επινοήθηκε με σκοπό τη μεγιστοποίηση της αποδοτικής χρήσης των CPU. Η εικονικοποίηση διακομιστή ικανοποίησε δύο ανεξάρτητες ανάγκες. Πρώτον, η εκτέλεση πολλών διακομιστών χωρίς να απαιτείται η χρήση πολλαπλών φυσικών υπολογιστικών κόμβων χάρισε στο μοντέλο το πλεονέκτημα της οικονομικής αποδοτικότητας. Δεύτερον, αντί να απαιτείται ο επαναπρογραμματισμός των εφαρμογών έτσι ώστε αυτές να γίνουν πολυνηματικές (multi-threaded) και παράλληλα να εκμεταλλεύονται τους πολλαπλούς πυρήνες, η εικονικοποίηση διακομιστή επέτρεψε στις επιχειρήσεις να επαναχρησιμοποιούν τις εφαρμογές τους, εκτελώντας πολλές από αυτές σε απομονωμένες μονάδες που ονομάστηκαν *εικονικές μηχανές* (VMs, virtual machines).

Το κυρίαρχο λειτουργικό σύστημα (OS, operating system) για την ανάπτυξη εφαρμογών στις επιχειρήσεις ήταν τα Windows. Τα λειτουργικά συστήματα που βασίζονταν στο Unix, όπως το Solaris, δεν είχαν την ίδια δημοτικότητα ως πλατφόρμες εφαρμογών (application platforms). Αυτό οφειλόταν εν μέρει στο κόστος των σταθμών εργασίας που ήταν απαραίτητοι για τη λειτουργία τέτοιων συστημάτων Unix ή τύπου Unix. Η κατακερματισμένη αγορά αυτών των λειτουργικών συστημάτων ήταν ακόμα μια αιτία για την έλλειψη εφαρμογών. Δεν υπήρξε ένα σύστημα που να ξεχώριζε στην κατηγορία λειτουργικών συστημάτων τύπου Unix. Όλα φαίνονταν ίδια σε υψηλό επίπεδο, αλλά διέφεραν σε αμέτρητες λεπτομέρειες, δυσκολεύοντας το έργο των προγραμματιστών εφαρμογών. Από την άλλη μεριά, το Linux σταθεροποιούνταν και αύξανε τη δημοτικότητά του, αλλά δεν είχε γεφυρώσει ακόμα το χάσμα μεταξύ των πρώτων που το υιοθέτησαν και της ευρείας ανάπτυξης.

Η επιτυχία του διαδικτύου, ιδίως του ιστού, σήμαινε ότι οι στάλες των online πληροφοριών έγιναν ορμητικό ποτάμι. Η εύρεση σχετικών πληροφοριών απέκτησε κρίσιμη σημασία. Οι μηχανές αναζήτησης άρχισαν έναν πόλεμο επικράτησης. Η Google, αν και άργησε να εισέλθει σε αυτό το μέτωπο, τελικά ξεπέρασε όλες τις άλλες μηχανές αναζήτησης. Κατάφερε μάλιστα να μπει στα λεξικά ως ρήμα, *γκουγκλάρω*, που σημαίνει *αναζητώ πληροφορίες στον ιστό*, περίπου όπως η Xerox ήταν κάποτε (και συνεχίζει σε κάποια μέρη του πλανήτη) το συνώνυμο της φωτοτυπίας. Τη δεδομένη στιγμή, η μηχανή αναζήτησης Google χειρίζεται 3,5 δισεκατομμύρια αναζητήσεις ανά ημέρα, με μέσο όρο 40.000 το δευτερόλεπτο. Για να αντεπεξέλθουν σε αυτή την κλίμακα, οι διακομιστές έπρεπε να κατανεμηθούν ακόμη περισσότερο και νέες τεχνικές, ιδίως αρχιτεκτονικές εφαρμογών που βασίζονταν σε συστάδες (cluster), όπως η MapReduce, απέκτησαν κυρίαρχη θέση. Όλα αυτά συνέβαλαν σε μια ιστορική μετατόπιση από την κυριαρχία της επικοινωνίας μεταξύ πελατών και διακομιστών στην επικράτηση της κυκλοφορίας μεταξύ διακομιστών.

Το Linux επικράτησε ως προς την ανάπτυξη αυτής της νέας γενιάς εφαρμογών. Η φιλοσοφία της ανάπτυξης εφαρμογών στο Unix (έχει ενσωματωθεί πλέον στο Linux) ήταν ο σχεδιασμός προγραμμάτων που έκαναν αποκλειστικά ένα πράγμα, το έκαναν

καλά και μπορούσαν να συνεργαστούν με άλλα παρόμοια προγράμματα προκειμένου να παρέχονται διάφορα είδη υπηρεσιών. Κατά συνέπεια, οι διακομιστές χωρίστηκαν ακόμα περισσότερο, δημιουργώντας ένα κίνημα που έχει γίνει πλέον γνωστό ως «μικροϋπηρεσίες» (microservice). Αυτό οδήγησε στην ανάπτυξη και στην άνοδο των περιεκτών (container), οι οποίοι ήταν ελαφρύτεροι από τις VM. Εάν τα πάντα εκτελούνταν στο Linux, η εκκίνηση μιας μηχανής ήταν μάλλον ένα μεγάλο αντίτιμο για την εκτέλεση ενός μικρού μόνο τμήματος μιας μεγαλύτερης υπηρεσίας.

Η άνοδος του Linux και της οικονομίας μεγάλης κλίμακας οδήγησε στην άνοδο του υπολογιστικού νέφους (cloud), μιας υπηρεσίας που περιόρισε στις επιχειρήσεις τον πονοκέφαλο που προκαλεί η λειτουργία ενός οργανισμού ΤΠ (Τεχνολογίας Πληροφοριών). Εξάλειψε την υποχρέωση να πρέπει να γνωρίζουμε ποιο δίκτυο και ποια υπολογιστική υποδομή θα αγοράσουμε, πόσο κάνουν και πώς θα τις κλιμακώσουμε, αναβαθμίσουμε και επιδιορθώσουμε. Μάλιστα το έπραξε αυτό σε ένα κλάσμα του κόστους κατασκευής και διαχείρισης μιας ιδιωτικής υποδομής. Η ιστορία του υπολογιστικού νέφους κρύβει βεβαίως περισσότερες λεπτομέρειες, αλλά περισσότερα για αυτό θα ειπωθούν αργότερα.

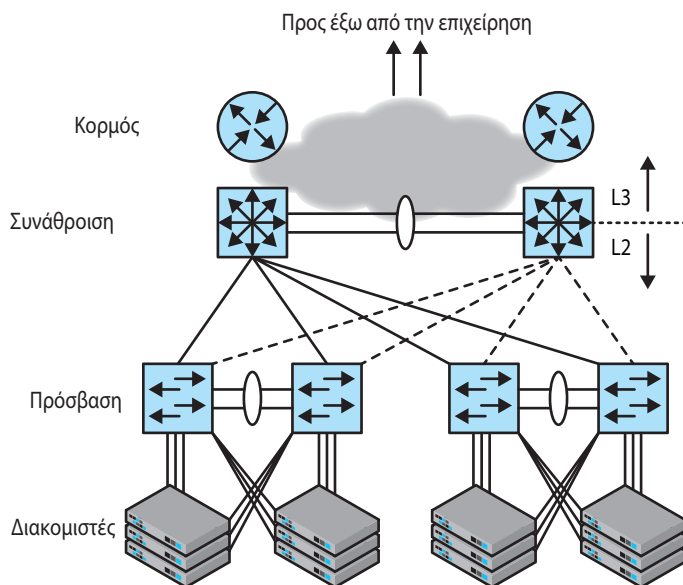
Η κλίμακα των πελατών που υποστηρίχθηκε σε μια υποδομή υπολογιστικού νέφους υπερέβαινε κατά πολύ τις τεχνολογίες που αναπτύσσονταν σε επίπεδο επιχείρησης. Οι περιέκτες και η δικτύωση του υπολογιστικού νέφους επέτρεψαν επίσης μεγαλύτερο όγκο επικοινωνίας μεταξύ των διακομιστών, όπως συνέβη επίσης στην περίπτωση της νέας γενιάς εφαρμογών όπως η αναζήτηση στον ιστό. Όλες αυτές οι νέες εφαρμογές απαιτούσαν επίσης υψηλότερο εύρος ζώνης από το δίκτυο, οπότε το 10 Gigabit Ethernet έγινε το κοινό πρότυπο διασύνδεσης. Οι ταχύτητες διασύνδεσης εξακολουθούν να αυξάνονται, με τους τελευταία ανακοινωμένους μεταγωγείς να υποστηρίζουν 400 Gigabit Ethernet.

Η κλίμακα, η διαφορά στο στιλ επικοινωνίας εφαρμογών και η ανάπτυξη ακόμα πιο καταναεμημένων εφαρμογών έθεσαν νέες απαιτήσεις στο δίκτυο, προκειμένου να μπορεί να ανταποκρίνεται σε αυτά τα νέα απαιτούμενα. Παρακάτω θα ασχοληθούμε με το δεύτερο μέλος του ζευγαριού του χορού, το δίκτυο, για να μάθουμε τι υπήρχε τότε και γιατί δεν μπορούσε πλέον να ακολουθήσει τον νέο ρυθμό.

Ο σχεδιασμός δικτύων μετά την αλλαγή του αιώνα

Το Σχήμα 1.2 παρουσιάζει ένα σχέδιο δικτύου που κυριαρχούσε στο τέλος του περασμένου αιώνα. Πρόκειται για το δίκτυο με το οποίο επιχειρήσαν να «χορέψουν» οι εφαρμογές των σύγχρονων κέντρων δεδομένων. Αυτό το δίκτυο ονομάζεται *πρόσβασης-συνάθροισης-κορμού* (access-aggregation-core) ή, για συντομία, access-agg-core ή απλώς access-agg.

Τα τελικά σημεία, ή υπολογιστικοί κόμβοι, συνδέονται σε μεταγωγείς πρόσβασης, οι οποίοι βρίσκονται στο χαμηλότερο επίπεδο στο Σχήμα 1.2. Οι μεταγωγείς συνάθροισης, ή αλλιώς *μεταγωγείς διανομής* (distribution switch), συνδέονται με τη σειρά



ΣΧΗΜΑ 1.2 Αρχιτεκτονική δικτύων πρόσβασης-συνάθροισης-κορμού

τους με το δίκτυο κορμού, με αποτέλεσμα να διασυνδέεται το δίκτυο πρόσβασης με τον υπόλοιπο κόσμο. Δεν εμφανίζονται ρητές ζεύξεις (link) μεταξύ των μεταγωγέων συνάθροισης και των μεταγωγέων κορμού επειδή δεν είναι ίδιες σε όλα τα δίκτυα και κυρίως δεν αφορούν τη συζήτηση που ακολουθεί.

Ο λόγος για την ύπαρξη δύο μεταγωγέων συνάθροισης ήταν βεβαίως η αποφυγή της απομόνωσης του δικτύου όταν παρουσίαζε βλάβη κάποιος από τους μεταγωγείς συνάθροισης. Εκείνη την εποχή, δύο μεταγωγείς θεωρούνταν επαρκείς για την επίτευξη κατάλληλου ρυθμού διεκπεραίωσης εργασιών (throughput) και τον πλεονασμό. Όμως, εν τέλει, αποδείχθηκαν ανεπαρκείς.

Η κυκλοφορία μεταξύ των μεταγωγέων πρόσβασης και συνάθροισης προωθείται μέσω γεφύρωσης (bridging). Βόρεια των διατάξεων συνάθροισης τα πακέτα προωθούνται μέσω δρομολόγησης. Επομένως, οι διατάξεις συνάθροισης ήταν κουτιά Janus, με τη νότια πλευρά να παρέχει γεφύρωση και τη βόρεια δρομολόγηση. Αυτό υποδεικνύεται στο Σχήμα 1.2 με την παρουσία των L2 και L3, τα οποία αναφέρονται στη δικτύωση επιπέδου 2 (γεφύρωση) και επιπέδου 3 (δρομολόγηση), αντίστοιχα.



Κατευθυντικότητα σε διαγράμματα δικτύων

Οι μηχανικοί δικτύων σχεδιάζουν το διάγραμμα δικτύου με τον δικτυακό εξοπλισμό στο επάνω μέρος και τους υπολογιστικούς κόμβους κάτω. Οι κατευθύνσεις με σημεία του ορίζοντα, όπως «βόρεια» και «νότια», έχουν ειδική σημασία σε τέτοια διαγράμματα. Η κυκλοφορία προς τον βορρά κινείται από τους υπολογιστικούς κόμβους στο δίκτυο και επιτρέπει την κίνηση από το εταιρικό δίκτυο προς το διαδίκτυο, ενώ η κυκλοφορία προς τον νότο κινείται από το διαδί-

κτυο πίσω στους τοπικούς υπολογιστικούς κόμβους. Με άλλα λόγια, πρόκειται για μια κλασική επικοινωνία πελάτη-διακομιστή. Η κυκλοφορία στη διεύθυνση ανατολής-δύσης ρέει μεταξύ διακομιστών μέσα στο ίδιο δίκτυο. Αυτή η κυκλοφορία ήταν μικρή την εποχή του μοντέλου πελάτη-διακομιστή, αλλά είναι το επικρατέστερο μοτίβο κυκλοφορίας για τις σύγχρονες εφαρμογές κέντρου δεδομένων, οι οποίες είναι κατανεμημένες με μεγάλο βαθμό συσταδοποίησης.

Η γοητεία της γεφύρωσης

Το δίκτυο πρόσβασης-συνάθροισης-κορμού εξαρτάται σε μεγάλο βαθμό από τη γεφύρωση, αν και γεννήθηκε μια εποχή όπου το διαδίκτυο εξελισσόταν ήδη με γοργά βήματα σε κάτι σημαντικό. Εάν η δρομολόγηση IP ήταν το στοιχείο που πυροδότησε το διαδίκτυο, γιατί η δικτύωση μέσα σε μια επιχείρηση στράφηκε στη γεφύρωση και όχι στη δρομολόγηση; Υπήρχαν τρεις κύριοι λόγοι: η άνοδος της προώθησης πακέτων με μεταγωγείς πυριτίου (silicon-switched packet forwarding), η άνοδος των ιδιόκτητων στοιβών λογισμικού δικτύωσης μέσα στην επιχείρηση και η υπόσχεση της μηδενικής διαμόρφωσης στη γεφύρωση.

Μεταγωγή πακέτων με υλικό

Τόσο στην πλευρά του παρόχου υπηρεσιών –που συνδέει τις επιχειρήσεις με το διαδίκτυο– όσο και στις εφαρμογές πελάτη-διακομιστή, άρχισαν να εμφανίζονται προμηθευτές που ειδικεύονται στην κατασκευή εξοπλισμού δικτύωσης. Ο εξοπλισμός δικτύωσης μεταπήδησε από τους σταθμούς εργασίας υψηλών επιδόσεων με πολλαπλές κάρτες διεπαφής δικτύου (NIC, network interface card) σε εξειδικευμένο υλικό που υποστήριζε μόνο την προώθηση πακέτων. Η προώθηση πακέτων βίωσε μια επανάσταση όταν αναπτύχθηκαν ειδικοί επεξεργαστές ολοκληρωμένων κυκλωμάτων εξειδικευμένων ως προς την εφαρμογή (ASIC, application-specific integrated circuit) για την προώθηση των πακέτων. Η έλευση διατάξεων πυριτίου για τη μεταγωγή πακέτων επέτρεψε τη σύνδεση πολλών περισσότερων διεπαφών σε ένα μόνο κουτί, καθώς και τη δυνατότητα προώθησης πακέτων με πολύ μικρότερη καθυστέρηση από όσο ήταν ποτέ δυνατόν στο παρελθόν. Ωστόσο, αυτή η τεχνολογία υλικού μεταγωγής υποστήριζε αρχικά μόνο τη γεφύρωση.

Δρομολογητές, γέφυρες και μεταγωγείς – επικαλυπτόμενοι όροι

Ένας δρομολογητής (router) είναι μεταγωγέας (switch); Ή ο μεταγωγέας είναι συνώνυμο της γεφύρωσης (bridging); Αυτό το ερώτημα παραμένει εδώ και καιρό στον νου των μηχανικών δικτύων. Ένας πελάτης μου, την εποχή του Cumulus, είχε την ιδέα ότι ο μεταγωγέας αφορούσε μόνο τη γεφύρωση και επομένως οι διαχωρισμένοι μεταγωγείς ήταν ικανοί μόνο για γεφύρωση – ως εκ τούτου,