



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
Εθνικόν και Καποδιστριακόν
Πανεπιστήμιον Αθηνών

ΤΜΗΜΑ
ΠΛΗΡΟΦΟΡΙΚΗΣ &
ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ



Προηγμένες Αρχιτεκτονικές Δικτύων (M132)

Δίκτυα Κέντρων Δεδομένων

Διδάσκων:

Κωνσταντίνος Χριστοδουλόπουλος

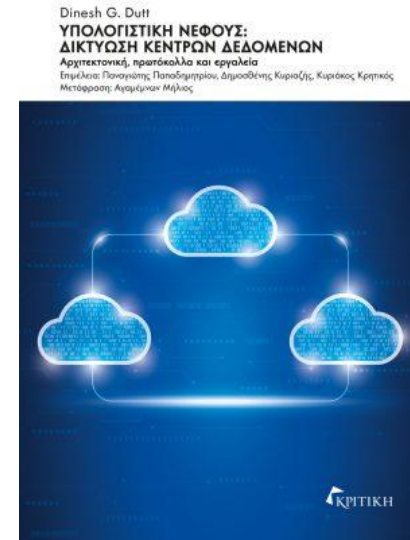
kchristodou@di.uoa.gr

<https://eclass.uoa.gr/courses/DI469/>

Προηγμένες Αρχιτεκτονικές Δικτύων

Θεματικές Ενότητες (ΘΕ) μαθήματος

- Εισαγωγή
- Δίκτυα Κορμού και Μητροπολιτικά Δίκτυα
(Οπτικά Δίκτυα Μεταγωγής Μήκους Κύματος)
- Δίκτυα πρόσβασης – Ενσύρματα
(Δίκτυα καλωδίου, DSL, παθητικά οπτικά δίκτυα)
- Δίκτυα πρόσβασης – Ασύρματα/Κινητά Δίκτυα
(κινητά δίκτυα 4G και 5G)
- **Δίκτυα κέντρων δεδομένων**
(Datacenter networks)



Σχετικά βιβλία:

1) [Dinesh G. Dutt](#) “Υπολογιστική νέφους: Δικτύωση κέντρων δεδομένων Αρχιτεκτονική, πρωτόκολλα και εργαλεία” εκδόσεις κριτική

2) Papers

<https://eclass.uoa.gr/modules/document/index.php?course=DI469&openDir=/64479193ZI65/69fb0101x0v4>

Κέντρα Δεδομένων (Data Centers)

■ Wikipedia

“Ένα κέντρο δεδομένων είναι μια εγκατάσταση που χρησιμοποιείται για τη στέγαση συστημάτων υπολογιστών και συναφών στοιχείων, όπως τηλεπικοινωνιακών συστημάτων και συστημάτων αποθήκευσης”

■ Υπάρχουν 3 είδη κέντρων δεδομένων

- Ιδιωτικά/εταιρικά (private)
- Κέντρα δεδομένων συν-τοποθεσίας (co-location)
- Κέντρα δεδομένων νέφους (cloud)



America's data center spending spree

Construction spending on data centers, seasonally adjusted (billions)

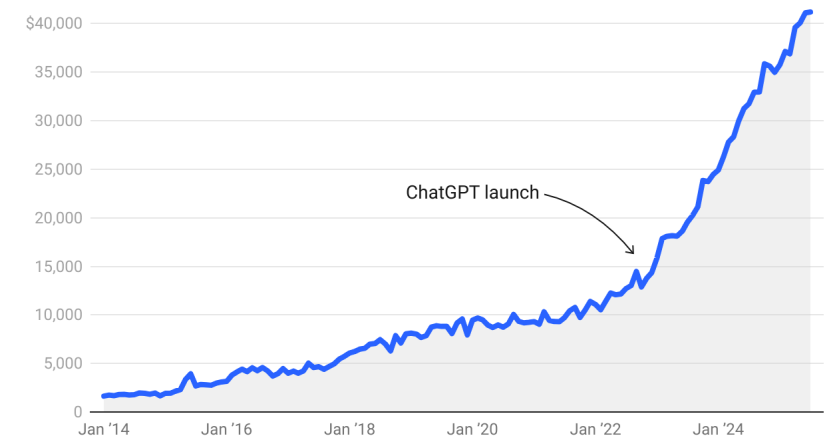


Chart: Michael Thomas • Source: Census • Created with Datawrapper

Ιδιωτικά κέντρα δεδομένων (Private)

- Ένα ιδιωτικό κέντρο δεδομένων είναι μια εγκατάσταση που ανήκει και λειτουργεί από την εταιρεία που το χρησιμοποιεί και είναι συχνά χτισμένο στην τοποθεσία της εταιρίας
- Σε ορισμένες περιπτώσεις μπορεί να βρίσκεται σε άλλη τοποθεσία (π.χ. όταν η εταιρία έχει πολλαπλές τοποθεσίες ή σε μια τοποθεσία χωρίς γραφεία της εταιρίας)
- Μπορεί να έχουν οριστεί τμήματα του κέντρου δεδομένων για την εξυπηρέτηση διαφορετικών τμημάτων/εφαρμογών της εταιρίας

Κέντρα δεδομένων συν-τοποθεσίας (Co-location)

- Η εταιρεία ενοικιάζει κενά racks στο κέντρο δεδομένων συν-τοποθεσίας – τοποθετεί διακομιστές
- Το υλικό (και λογισμικό) ανήκει στην εταιρεία και απλώς στεγάζεται στο κέντρο δεδομένων, το κέντρο δεδομένων παρέχει όλα τα 'γύρο' (ασφάλεια πρόσβασης, ηλεκτρικό + backup, δικτυακή σύνδεση με τον έξω κόσμο, κλπ)
- Δημοφιλής υπηρεσία για εταιρίες που δεν διαθέτουν χώρο και πόρους για το δικό τους κέντρο δεδομένων
- Οφέλη: χαμηλότερο κόστος, λιγότερο τεχνικό προσωπικό, αξιοπιστία, γεωγραφική τοποθεσία (καλή διασύνδεση με άλλα δίκτυα), προβλέψιμα έξοδα, επεκτασιμότητα



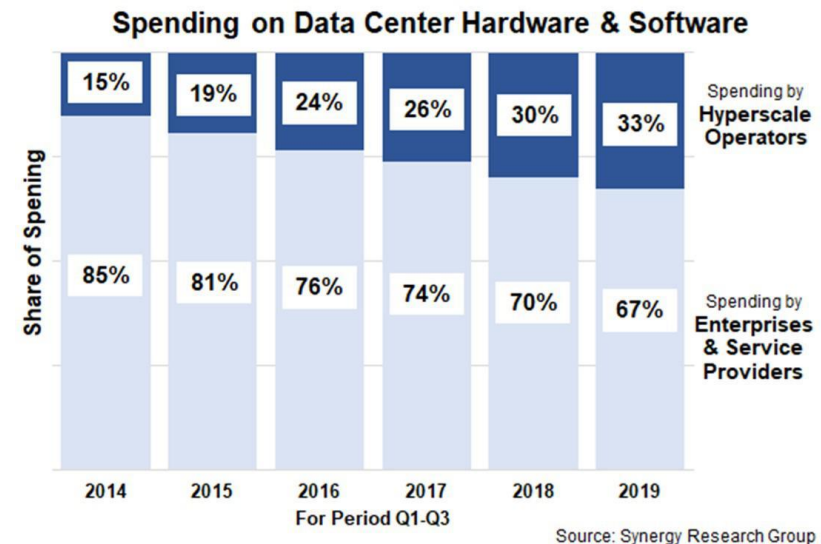
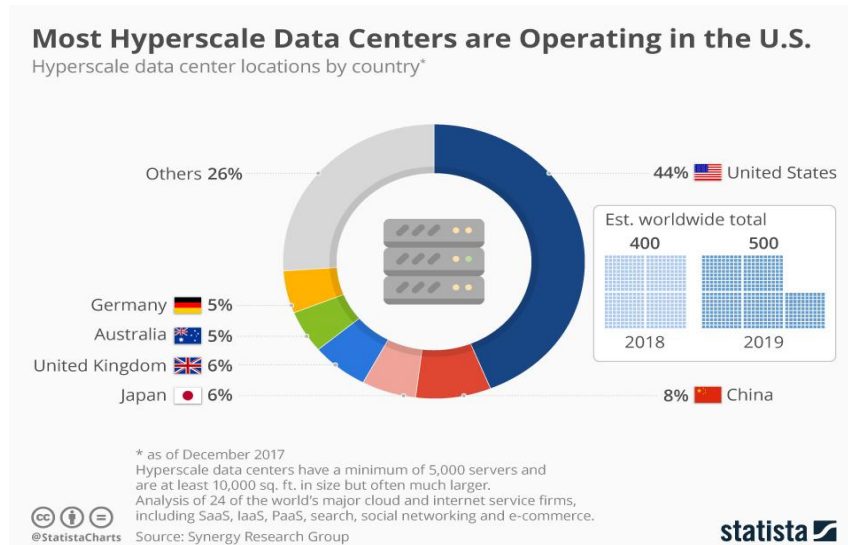
1U Colocation	2U Colocation	10U Colocation
 115€/month	 145€/month	 799€/month
Buy now	Buy now	Buy now
1U Colocation Space	2U Colocation Space	10U Colocation Space
1 Ampere Power Consumption	1 Ampere Power Consumption	6 Ampere Power Consumption
2 x C13 Power Sockets	2 x C13 Power Sockets	12 x C13 Power Socket
A + B Power Paths	A + B Power Paths	A + B Power Paths
1x100Mbps Switch Uplink	1x100Mbps Switch Uplink	2x100Mbps Switch Uplink
2 TB Traffic	2 TB Traffic	10 TB Traffic
1 x 1Gbps Switchport	1 x 1Gbps Switchport	6 x 1Gbps Switchports
1 x IPv4 Address	1 x IPv4 Address	16 x IPv4 Addresses
Greek Data Center	Greek Data Center	Greek Data Center
APC NetShelter Rack	APC NetShelter Rack	APC NetShelter Rack
Tier III Construction Certified	Tier III Construction Certified	Tier III Construction Certified
Cold Aisle Containment	Cold Aisle Containment	Cold Aisle Containment
Leed v4 Gold Energy Efficiency	Leed v4 Gold Energy Efficiency	Leed v4 Gold Energy Efficiency
ISO27001:2013 Certified	ISO27001:2013 Certified	ISO27001:2013 Certified

Κέντρα δεδομένων νέφους (Cloud)

- Κέντρα δεδομένων που προσφέρουν υπηρεσίες υπολογιστικού νέφους
 - Συνήθως μεγάλου μεγέθους (υπάρχουν και μεσαίου)
 - Ανήκουν σε έναν οργανισμό (Amazon, Google, Microsoft, IBM, etc)
 - Χρησιμοποιούν μια σχετικά ομοιογενή πλατφόρμα λογισμικού και υλικού
 - Κοινό σύστημα διαχείρισης και ευρεία χρήση εικονικοποίησης
 - Σχεδιασμένο να αντέχει βλάβες με μικρή ή καθόλου επίδραση στην απόδοση και τη διαθεσιμότητα των υπηρεσιών
- Εκτελούν εφαρμογές της εταιρίας (Google Search, Mail, Maps, youtube, etc – βλέπε ιδιωτικά κέντρα δεδομένων) ή/και προσφέρουν υπηρεσίες σε πελάτες
 - IaaS, or infrastructure as a service: on-demand access to cloud-hosted physical and virtual servers, storage and networking - the backend IT infrastructure for running applications and workloads in the cloud
 - PaaS, or platform as a service: on-demand access to a complete, ready-to-use, cloud-hosted platform for developing, running, maintaining and managing applications
 - SaaS, or software as a service: on-demand access to ready-to-use, cloud-hosted application software

Cloud Datacenters

- Κέντρα δεδομένων νέφους
 - Μεγάλα κέντρα δεδομένων (hyperscale): 'υπολογιστικά εργοστάσια' καταλαμβάνοντας μεγάλες εκτάσεις (>1000 m²) με δεκάδες χιλιάδες διακομιστές (>10000)
 - Μεσαία και μικρά κέντρα δεδομένων με χιλιάδες διακομιστές
- Περισσότερα από 500 hyperscale datacenters παγκοσμίως
- Ανήκουν σε μεγάλες εταιρίες πληροφορικής (cloud και content providers)
- Hyperscale αντιπροσώπευαν το 33% όλων των δαπανών για υλικό και λογισμικό για τα κέντρα δεδομένων το 2019



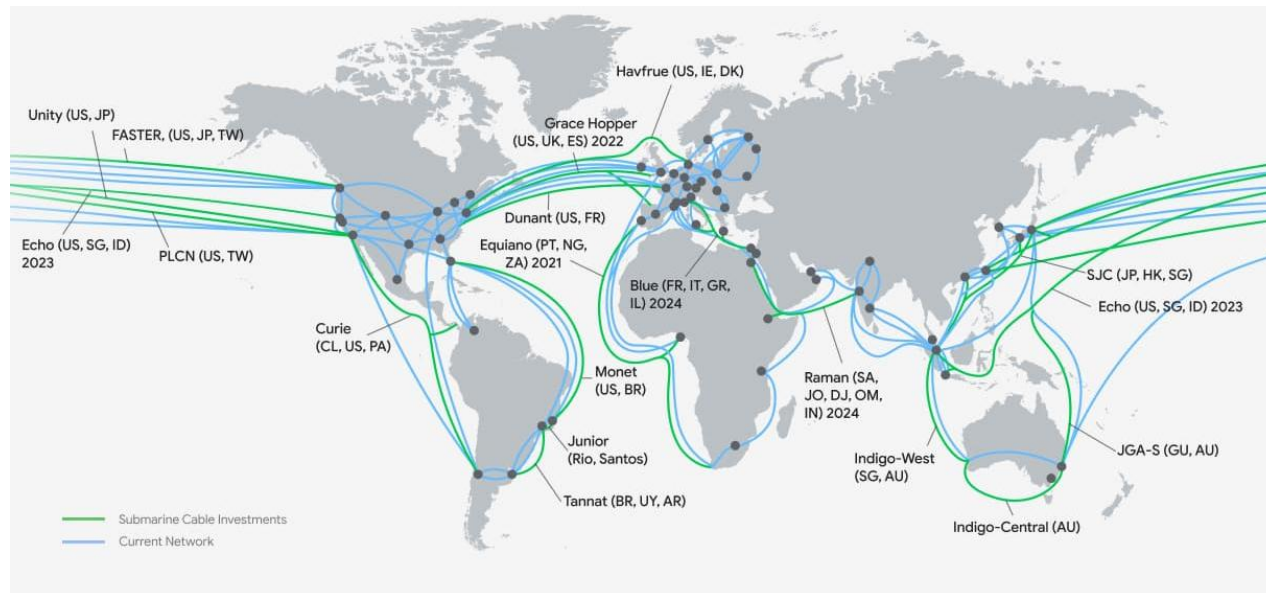
Χαρακτηριστικά ενός hyperscale datacenter

- Η ετήσια κατανάλωση ενέργειας ενός μεγάλου datacenter μπορεί να ξεπεράσει τα 30 GWh
 - Στις ΗΠΑ, η τυπική κατανάλωση ηλεκτρικού ρεύματος ενός νοικοκυριού είναι περίπου 11.700 kWh
 - Δλδ ίσο με ~2500 σπίτια
- 50 000 - 200 000 διακομιστές
- Το μεγαλύτερο κέντρο δεδομένων είναι πάνω από 100 000 m²
 - Περίπου 14 γήπεδα ποδοσφαίρου (γήπεδο ποδοσφαίρου ≈ 7 000 m²)
- Ο μέσος αριθμός των θέσεων εργασίας σε ένα κέντρο δεδομένων συνήθως περιορίζεται σε 30 εργαζομένους διαφόρων ειδικοτήτων (IT και μη)
- Το κόστος κατασκευής ενός κέντρου δεδομένων είναι μερικά εκατομμύρια δολάρια/ευρώ
 - Αναφέρεται συχνά σε τιμή ανά τετραγωνικό πόδι (0.30² m²): τουλάχιστον 200\$ ανά τετραγωνικό πόδι, αλλά μπορεί να κυμαίνεται τόσο υψηλά όσο 1.000\$ ανά τετραγωνικό πόδι

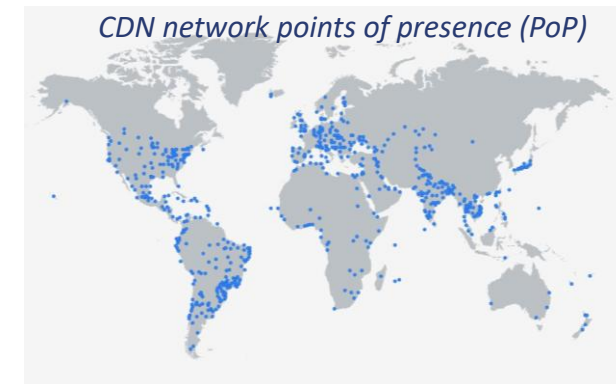
Το παράδειγμα του Google Cloud

Η υποδομή δικτύου της Google έχει τέσσερα διαφορετικά στοιχεία

- Data Centers
- Edge Nodes (Google Global Cache, or GGC)
- Points of Presence (POPs)
- Ιδιωτικό (οπτικό) δίκτυο διασύνδεσής τον παραπάνω



Google Global Network



39
REGIONS

118
ZONES

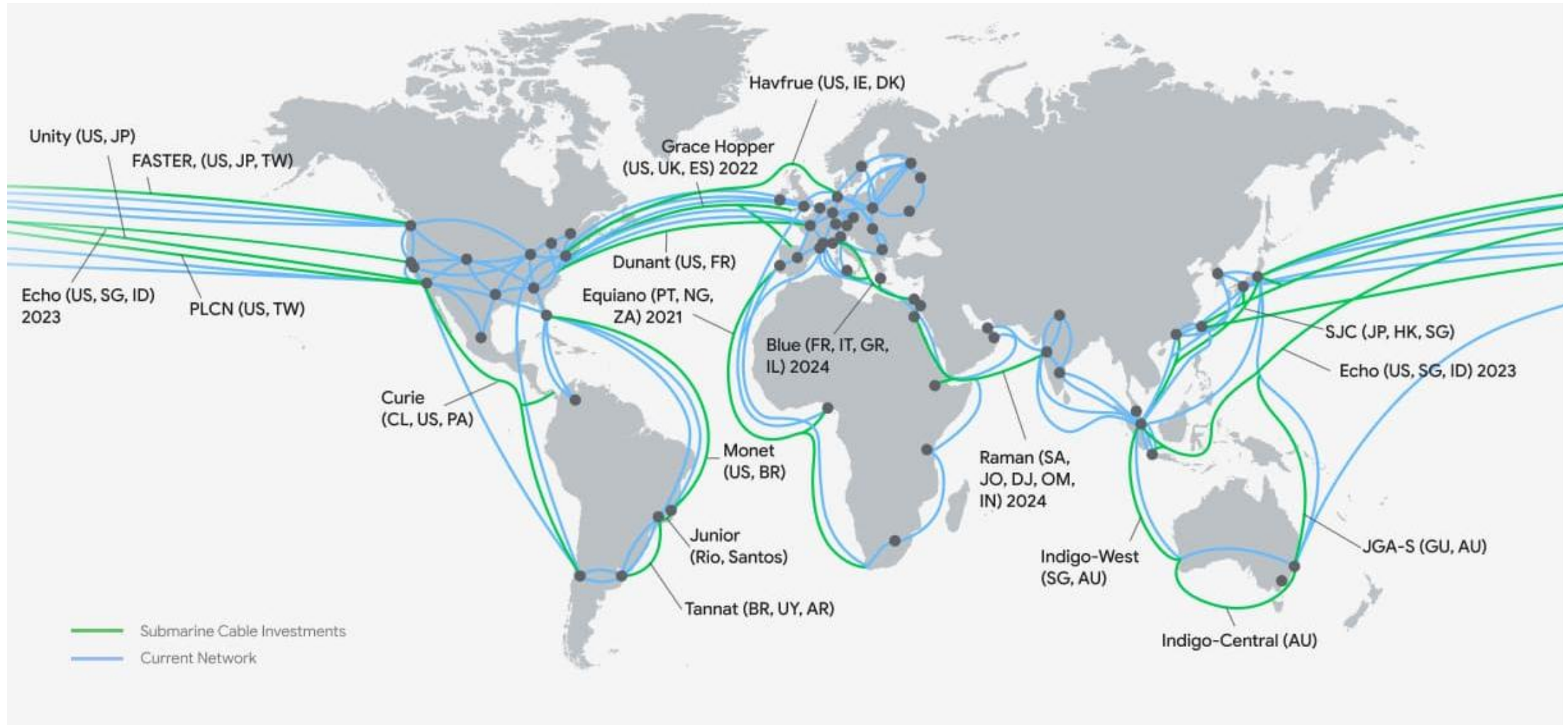
187
NETWORK EDGE LOCATIONS

AVAILABLE IN
200+
COUNTRIES AND TERRITORIES

COMING SOON! Google Cloud will continue expanding into the following regions: Querétaro (Mexico), Malaysia, Thailand, New Zealand, Greece, Norway, South Africa, Austria and Sweden.

<https://cloud.google.com/about/locations#lightbox-cloud-media-map>

Google Datacenters



Google Datacenters

Products available by location

Deploy resources in specific zones, regions and multi-regions.

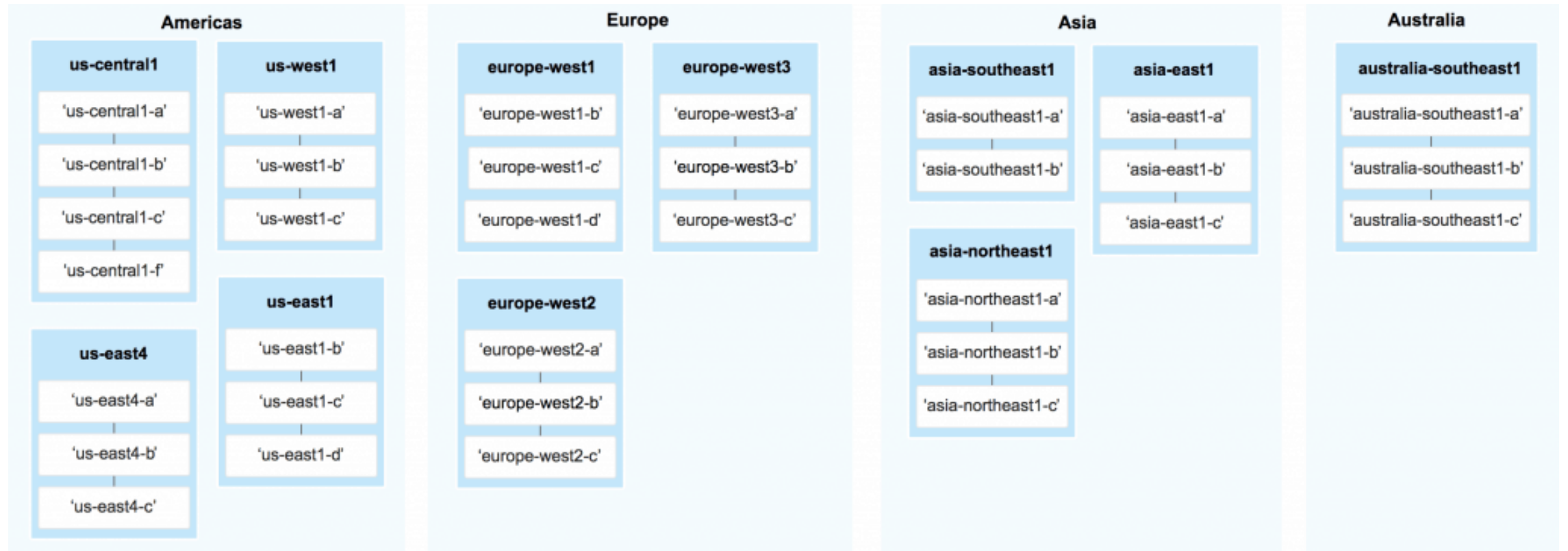
	AMERICAS		EUROPE		ASIA PACIFIC		MULTI-REGION		
PRODUCTS	OREGON (us-west1)	LOS ANGELES (us-west2)	SALT LAKE CITY (us-west3)	LAS VEGAS (us-west4)	IOWA (us-central1)	SOUTH CAROLINA (us-east1)	N. VIRGINIA (us-east4)	MONTRÉAL (northamerica-northeast1)	SÃO PAULO (southamerica-east1)
COMPUTE									
Compute Engine ^{5,6}	●	●	●	●	●	●	●	●	●
App Engine		●	●	●	●	●	●	●	●
Google Kubernetes Engine ^{5,6}	●	●	●	●	●	●	●		
Cloud Functions		●	●	●	●	●	●		

Οι **περιοχές** είναι ανεξάρτητες γεωγραφικές περιοχές που αποτελούνται από ζώνες. Οι ζώνες εντός μίας περιοχής έχουν καθυστέρηση δικτύου μετ' επιστροφής κάτω του <1 ms.

 39	 118	 187	 200+
REGIONS	ZONES	NETWORK EDGE LOCATIONS	COUNTRIES AND TERRITORIES

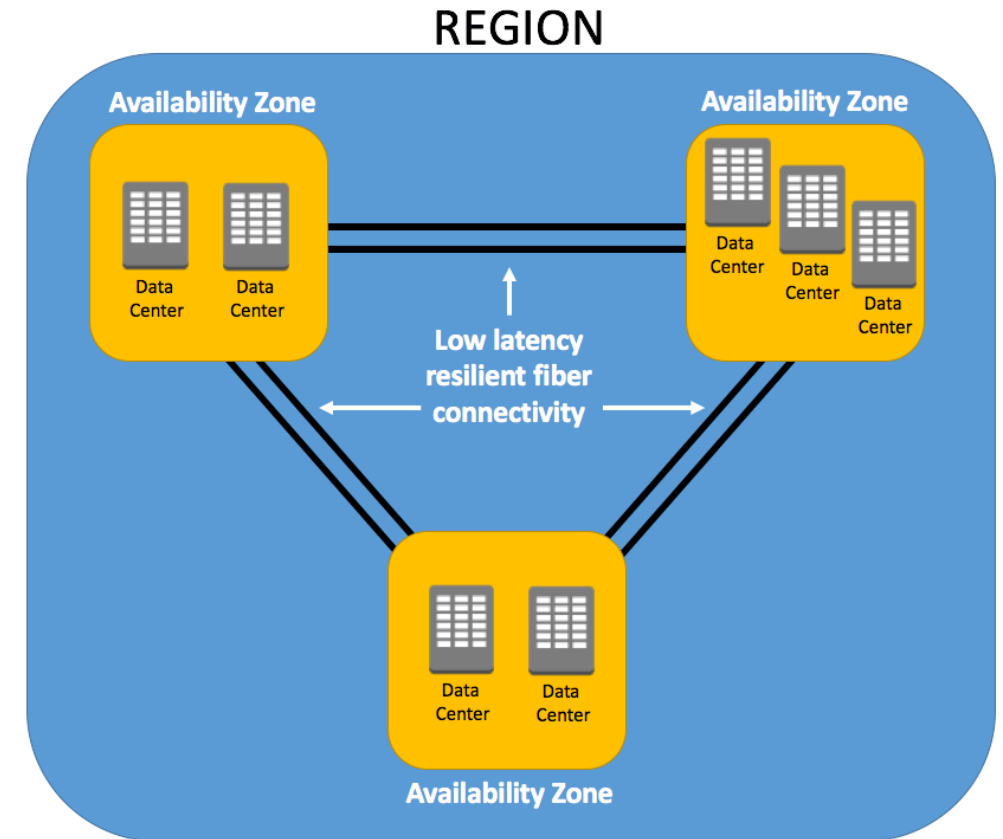
COMING SOON! Google Cloud will continue expanding into the following regions: Querétaro (Mexico), Malaysia, Thailand, New Zealand, Greece, Norway, South Africa, Austria and Sweden.

Google Datacenters



Google Datacenters

- Η τοποθέτηση εφαρμογών σε διαφορετικές ζώνες σε μια περιοχή παρέχει απομόνωση από τους περισσότερους τύπους αποτυχιών φυσικών υποδομών και υπηρεσιών λογισμικού υποδομής (π.χ. λειτουργικό)
- Η τοποθέτηση πόρων σε διαφορετικές περιοχές παρέχει ακόμη μεγαλύτερο βαθμό ανεξαρτησίας από αποτυχίας
 - π.χ. σε περίπτωση απώλειας μιας ολόκληρης περιοχής λόγω φυσικής καταστροφής



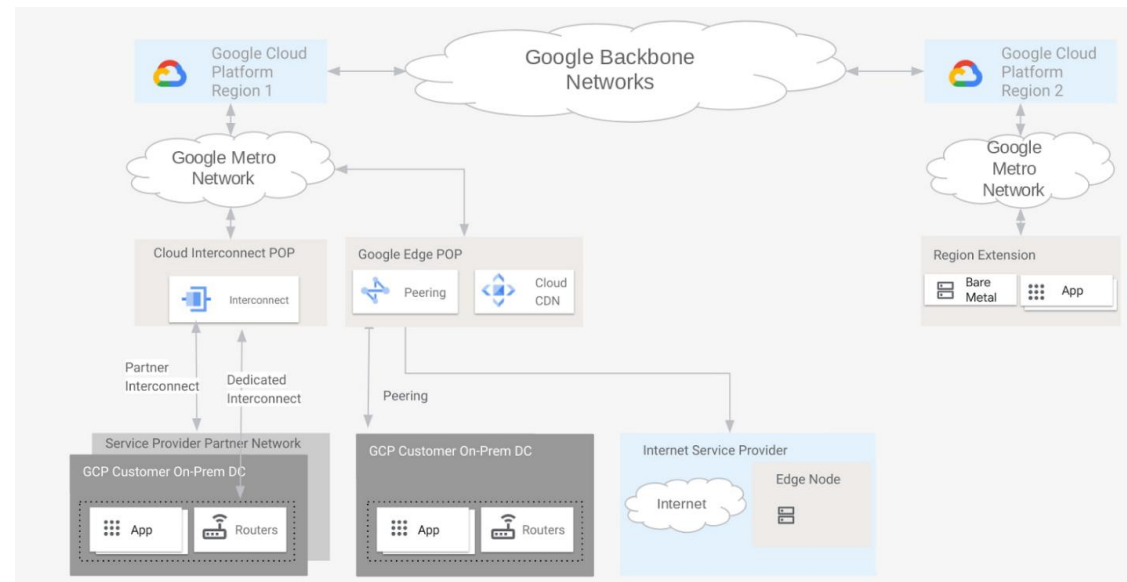
Points of Presence (POPs)

■ Google Edge POP

- Μέσω των Edge POPs πελάτες ή ISP μπορούν να συνδεθούν στο δίκτυο της Google μέσω peering
- Οποιοδήποτε δίκτυο φτάνει σε κάποιο Google POP με δικά του έξοδα
- Υπάρχουν τερματικοί δρομολογητές στα POPs και οι πελάτες μπορούν να διασυνδεθούν με αυτούς
- Πέρα από τη διεπαφή με το δίκτυο της Google, παρέχονται και υπηρεσίες Content Delivery Network (CDN)

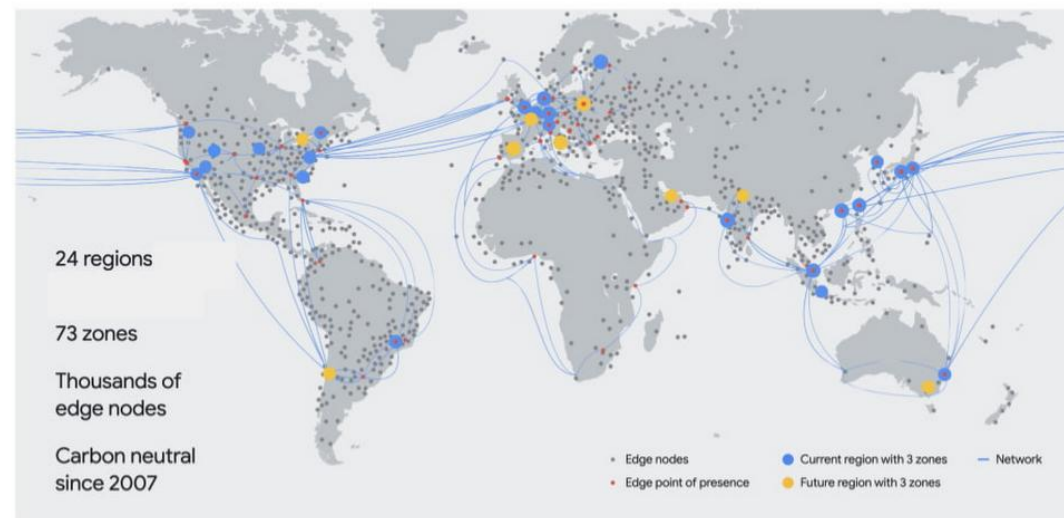
■ Cloud Interconnect POP

- Προσφέρει άμεσες, αποκλειστικές συνδέσεις με υπηρεσίες Google Cloud Platform (GCP)
- Για διασύνδεση on-premises υποδομής με GCP (hybrid cloud) ή ενός συνεργάτη παρόχου υπηρεσιών (Service provider Partner Network)

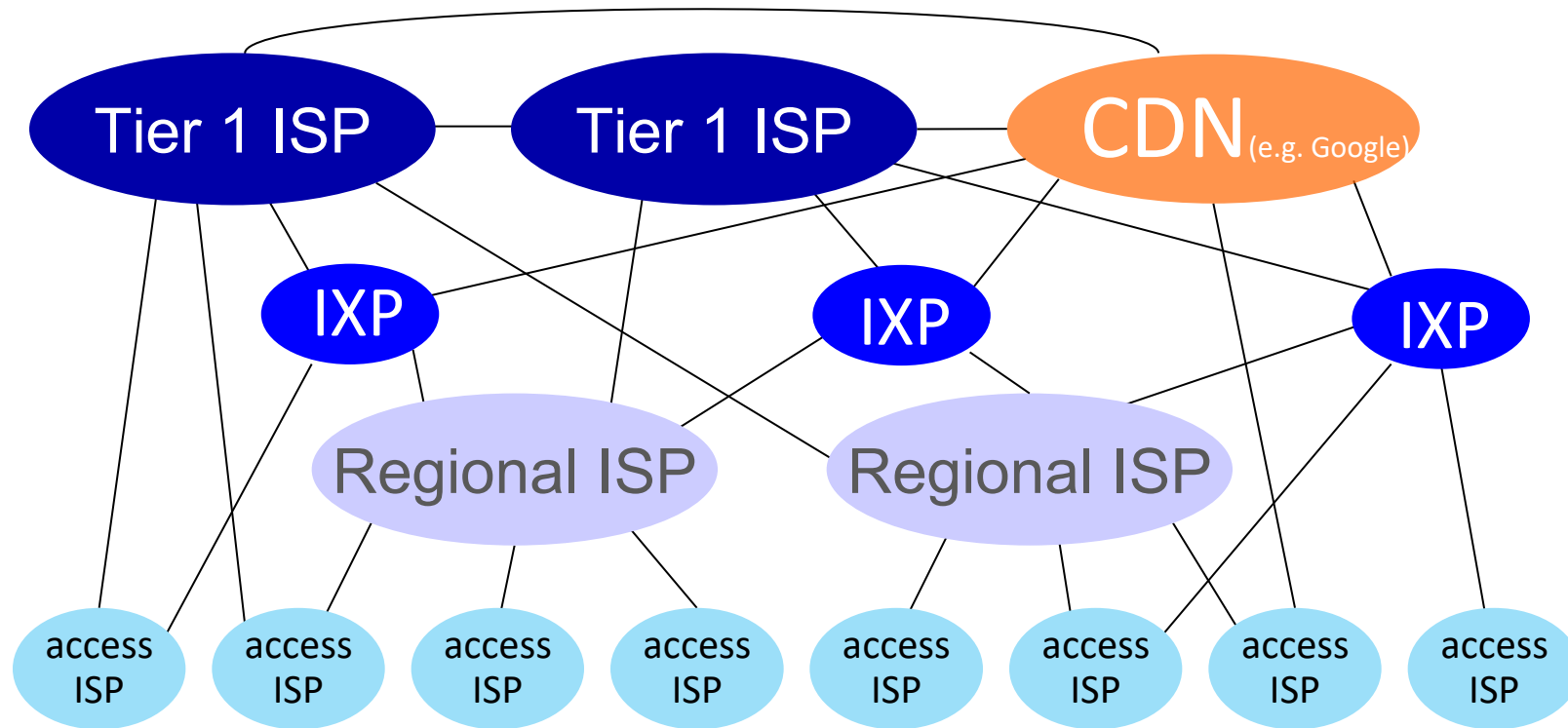


Edge Points of Presence (POPs)

- Η Google είναι παρούσα σε πάνω από 90 internet exchange points (IXP) και σε περισσότερες από 100 ιδιωτικές εγκαταστάσεις διασύνδεσης σε όλο τον κόσμο
- Λειτουργεί ένα μεγάλο, παγκόσμιο δίκτυο που συνδέει τα PoPs με τα κέντρα δεδομένων της
- Με τον τρόπο αυτό η κίνηση της Google έρχεται πιο κοντά στους χρήστες, παρέχοντας καλύτερης ποιότητας υπηρεσίες
- Το Peering μειώνει το συνολικό κόστος των εταιρειών για τη μεταφορά δεδομένων της Google σύγκριση με την χρήση transit παρόχων (tier 1 ISPs)



Δομή του Διαδικτύου: δίκτυο δικτύων

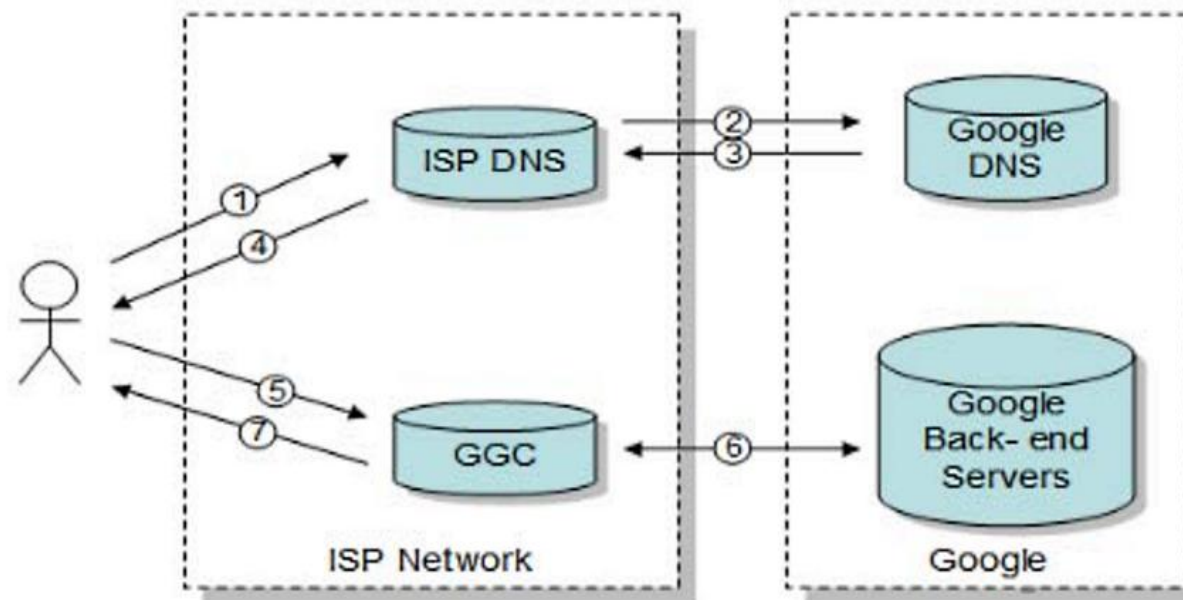


Στο κέντρο του Διαδικτύου: μικρός αριθμός διασυνδεδεμένων μεγάλων δικτύων

- **“tier-1” ISPs** (e.g., AT&T, Level 3, Sprint, NTT), εθνική και διεθνής κάλυψη
- **δίκτυα παρόχων περιεχομένου – CDN** (πχ, Google, Facebook): ιδιωτικά δίκτυα που συνδέουν τα κέντρα δεδομένων τους στο Διαδίκτυο, συχνά παρακάμπτοντας τους tier-1 και κάποιους περιφερειακούς ISPs

Google Global Cache (GGC)

- Το Google Global Cache (GGC) επιτρέπει στους ISP να εξυπηρετούν συγκεκριμένο περιεχόμενο της Google μέσα από το δικό της δίκτυο
- Αυτό μειώνει τη συμφόρηση και βελτιώνει την εμπειρία των χρηστών



Google Data Center - Council Bluffs, Iowa, USA



Google datacenter pictures: Council Bluffs



Κέντρα δεδομένων στην Ελλάδα



Κέντρα δεδομένων στην Ελλάδα



Microsoft: Πότε ξεκινά και πώς θα είναι το τριπλό data center campus στην ανατολική Αττική

Δημήτρης Μαλλός - CNN Greece
Σάββατο, 21 Φεβρουαρίου 2026 07:22



ΕΔΥΤΕ - Εθνικό Δίκτυο
Υποδομών
Τεχνολογίας και
Έρευνας - GRNET



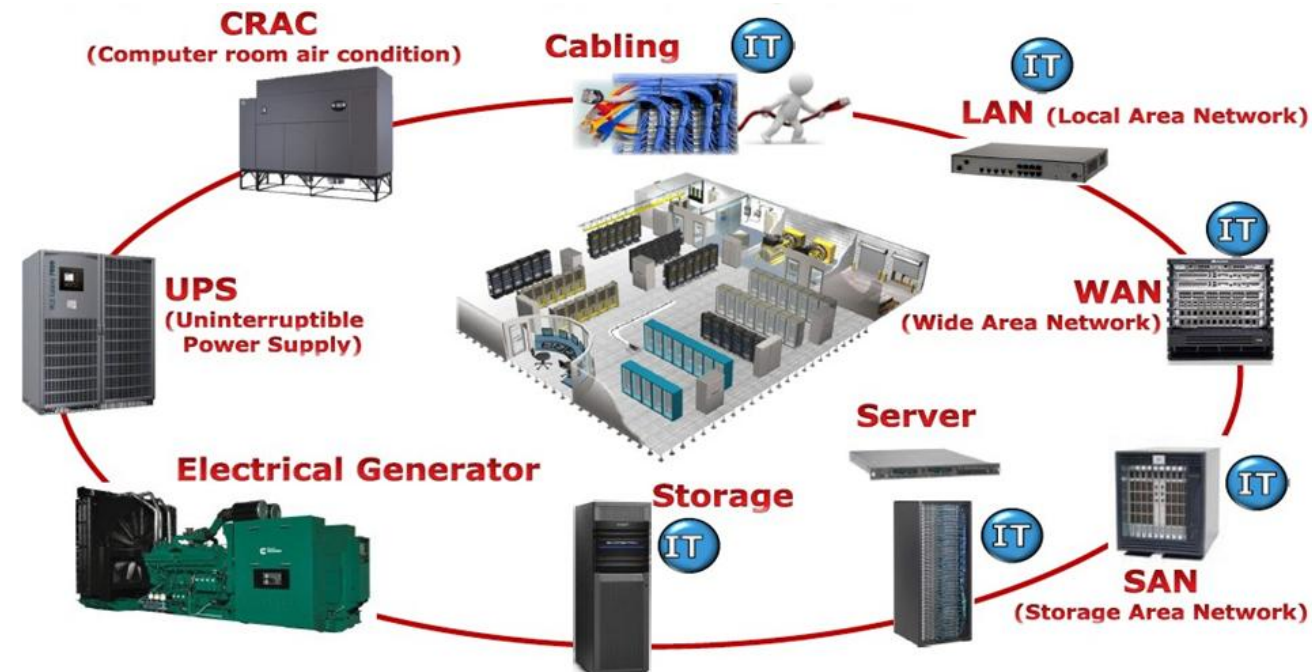
- «Πράσινο» κέντρο δεδομένων στον Λούρο
- Κέντρο δεδομένων Πανεπιστημίου Κρήτης (Ηράκλειο) για Υπηρεσίες Υγείας (31 διασυνδεδεμένα νοσοκομεία σε 8 περιφέρειες της χώρας)
- 3 κέντρα δεδομένων στο Υπουργείο Παιδείας, Έρευνας και Θρησκευμάτων
- Κέντρο δεδομένων στο Εθνικό Κέντρο Τεκμηρίωσης
- Κέντρο δεδομένων (μικρό) Κωλέττη, Internet Exchange point (IXP)- διασύνδεση Αθήνας με άλλες πόλεις

<https://grnet.gr/infrastructure/datacenters-grnet/>

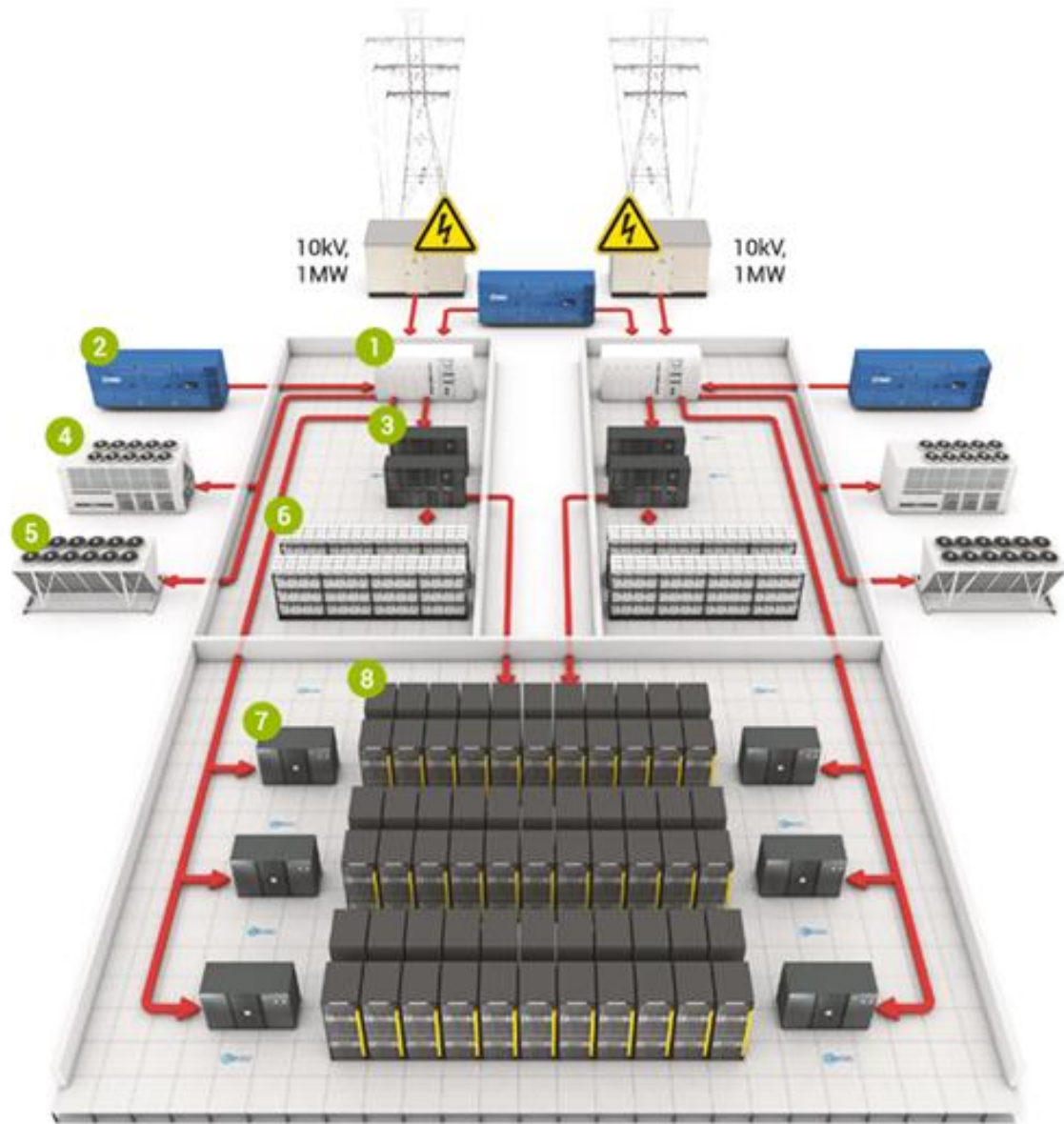


Τμήματα ενός τυπικού datacenter

- Υπολογιστική υποδομή
- Αποθηκευτική υποδομή
- Δικτυακή υποδομή (LAN, WAN, SAN)
- Καλωδίωση (cabling)
- Uninterrupted power source systems (UPS)
- Συστήματα παραγωγής ενέργειας (electrical generator)
- Συστήματα μετάδοσης ενέργειας (PDU)
- Συστήματα ελέγχου περιβάλλοντος (CRAC)
- Συστήματα πυρασφάλειας
- Συστήματα φυσικής ασφάλειας



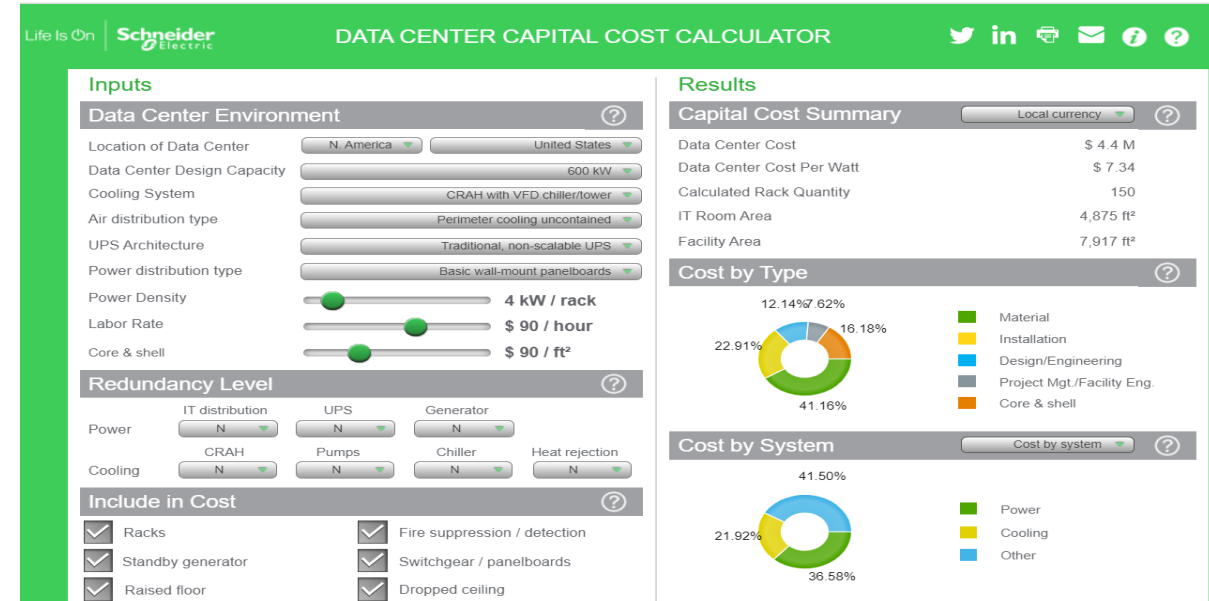
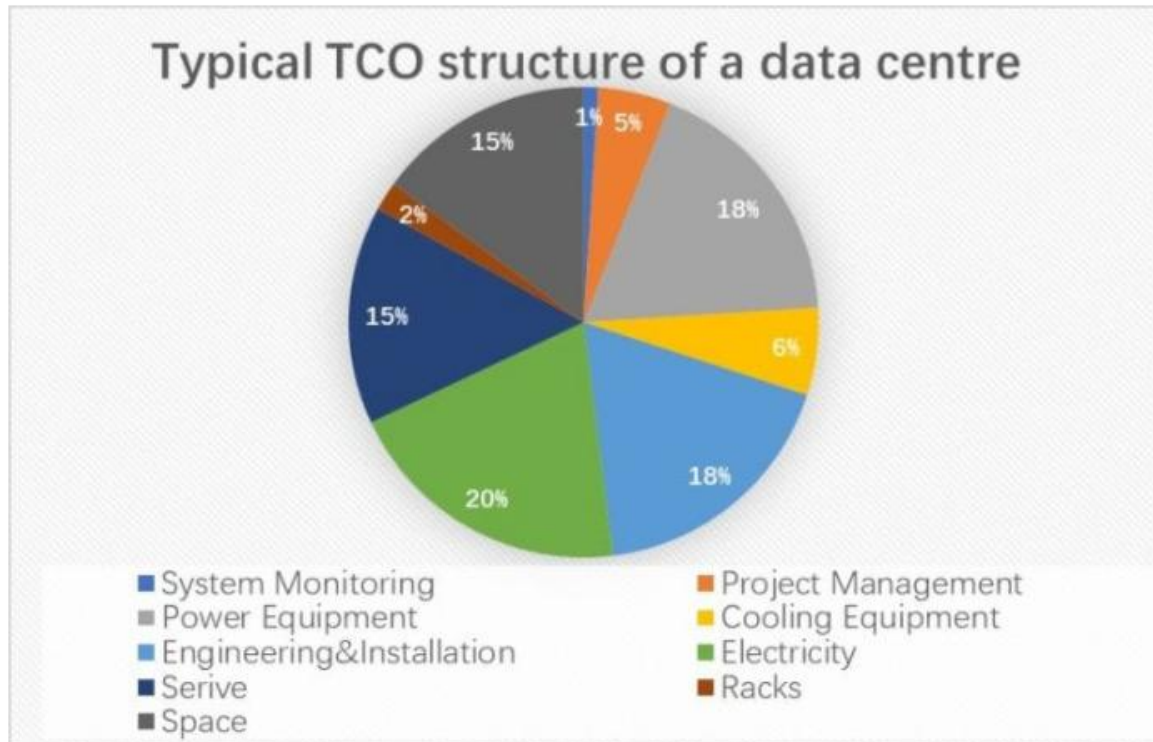
Φυσική διαμόρφωση ενός τυπικού datacenter



- 1 Main power distribution with Automatic Transfer Switch (ATS)
- 2 SDMO Diesel generator
- 3 APC Symmetra PX500 (Uninterruptible power supply)
- 4 Emerson Network Power chiller
- 5 DryCooler LU-VE
- 6 UPS batteries
- 7 Emerson Network Power climate control
- 8 Servers, storage, networking equipment

Καταμερισμός κόστους ενός datacenter

- Το total cost of ownership (TCO) περιέχει τρία μέρη, το κόστος επένδυσης, το λειτουργικό κόστος και το κόστος ενέργειας

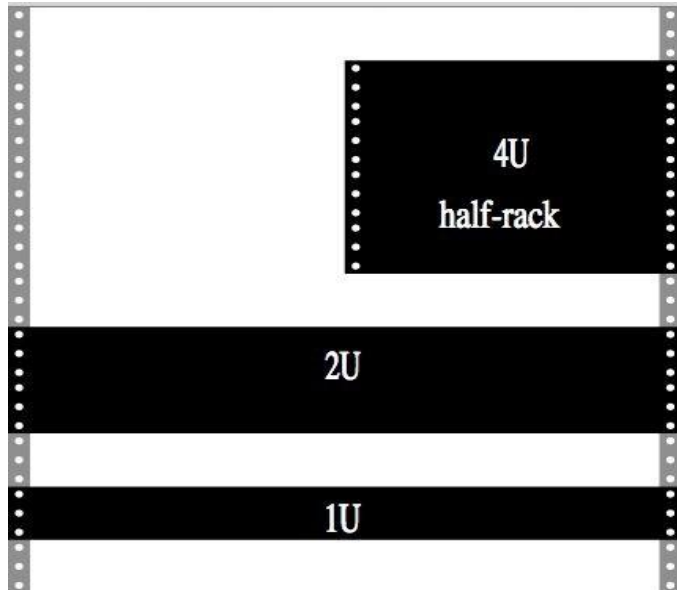


<https://www.se.com/uk/en/work/solutions/for-business/data-centers-and-networks/trade-off-tools/>

Rack

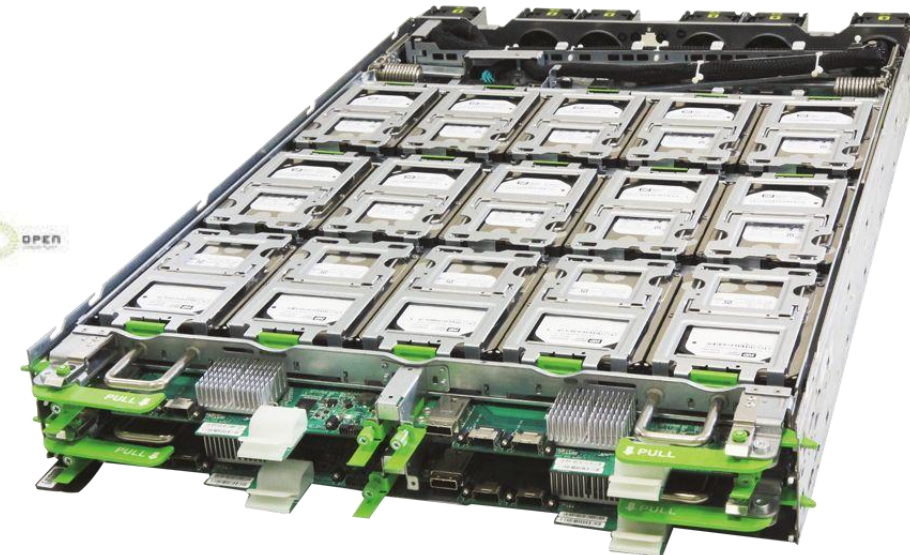
- Typically is 48cm wide, 91 or 102 cm depth, 180 cm height
- Rack unit U: 4.45 cm height
- Typically, Rack = 42 U

- Slots:



Rack Slots

- Slots hold power distribution, servers, storage, networking equipment
- Typical server: 2U
 - 8-128 cores
 - DRAM: 32-512 GB
- Typical storage: 2U
 - 30 drives
- Typical Network: 1U
 - 64 x 10GB



Διακομιστές ενός datacenter

- Δύο είδη διακομιστών χρησιμοποιούνται πιο συχνά στα κέντρα δεδομένων: διακομιστές rack και blade
- Ο διακομιστής rack
 - Υπολογιστής γενικής χρήσης (συνήθως high-end) που μπορεί να διαμορφωθεί ανάλογα με τις απαιτήσεις
 - Έχει κατασκευαστεί ειδικά για να μπαίνει σε racks (χρησιμοποιώντας βίδες στήριξης ή ράγες)
 - Το ύψος, ή το μέγεθος μπορεί να διαφέρει (1U, 2U, 4U)
 - Τοποθετούνται ο ένας πάνω από τον άλλο μέσα σε ένα ράφι, για να ελαχιστοποιείτε ο χώρος
- Ο διακομιστής Blade
 - Είναι μικρός και μερικές φορές περιέχει μόνο CPU, μνήμη ή αποθηκευτικό χώρο
 - Μπαίνει σε ένα περίβλημα blade, το οποίο χωράει πολλούς διακομιστές blade, παρέχοντας τροφοδοσία, ψύξη, δικτύωση, διάφορες διασυνδέσεις και διαχείριση
 - Επιλογές σχετικά με το τι θα είναι το κάθε blade και πως να τα συνδυάσεις σε ένα περίβλημα



Row/Cluster/POD

- 30+ racks

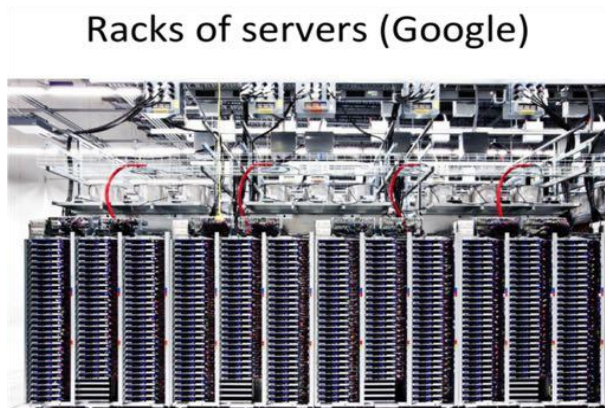


Row/Cluster/POD



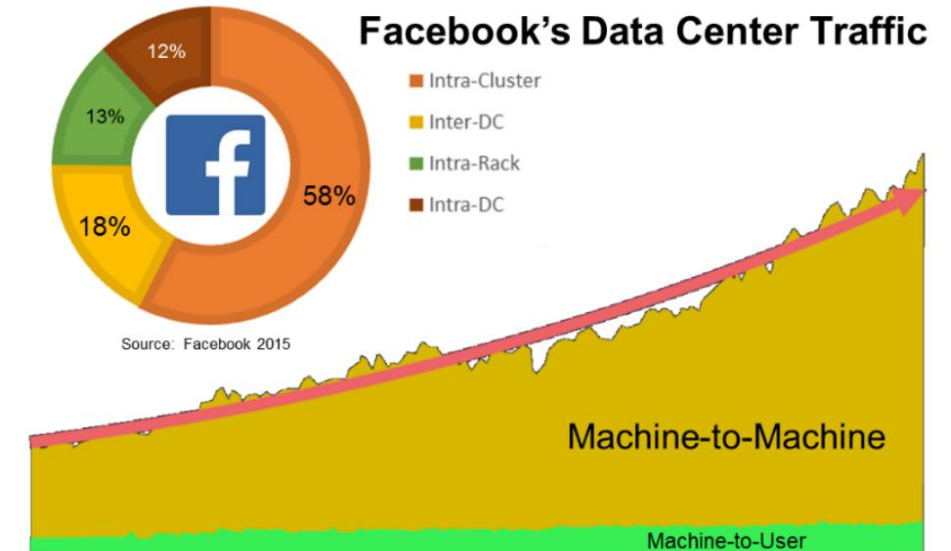
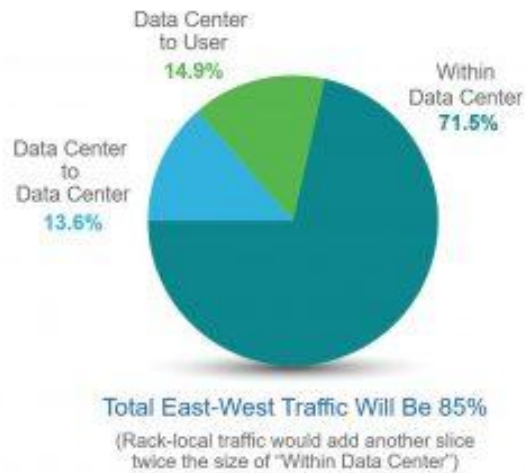
Δικτύωση κέντρων δεδομένων

- Κάθε κέντρο δεδομένων έχει το δικό του δίκτυο
 - διασυνδέει τους υπολογιστές του μεταξύ τους
 - διασυνδέει το κέντρο δεδομένων με το διαδίκτυο
- 15% το κόστους ενός μεγάλου κέντρου δεδομένων σχετίζονται με τη δικτύωση
 - δικτυακός εξοπλισμός (μεταγωγείς, δρομολογητές, ισοσταθμιστές φορτίου)
 - κόστος εξωτερικών ζεύξεων
- Ενώ η δικτύωση δεν είναι το μεγαλύτερο κόστος, είναι πολύ σημαντική στη λειτουργικότητα, στη μεγιστοποίηση της απόδοσης και τη μείωση όλου του κόστους



Δικτύωση κέντρων δεδομένων

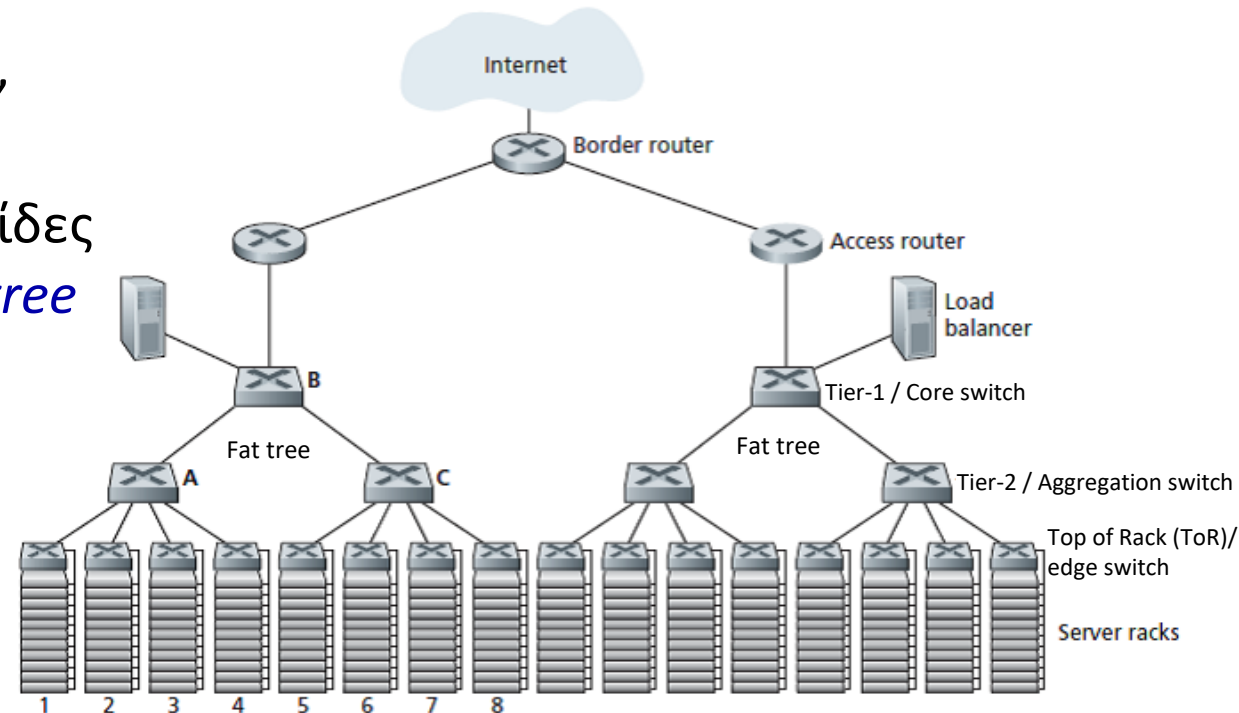
- Είδη κίνησης:
 - Μέσα στο κέντρο δεδομένων (“machine to machine”, “east-west”)
 - Μεταξύ κέντρου δεδομένων και χρηστών (“north-south”)
 - Μεταξύ κέντρων δεδομένων (datacenters interconnect DCI – μέρος του “north-south”)
- Το μεγαλύτερο ποσοστό της κίνησης είναι μέσα στο κέντρο δεδομένων !



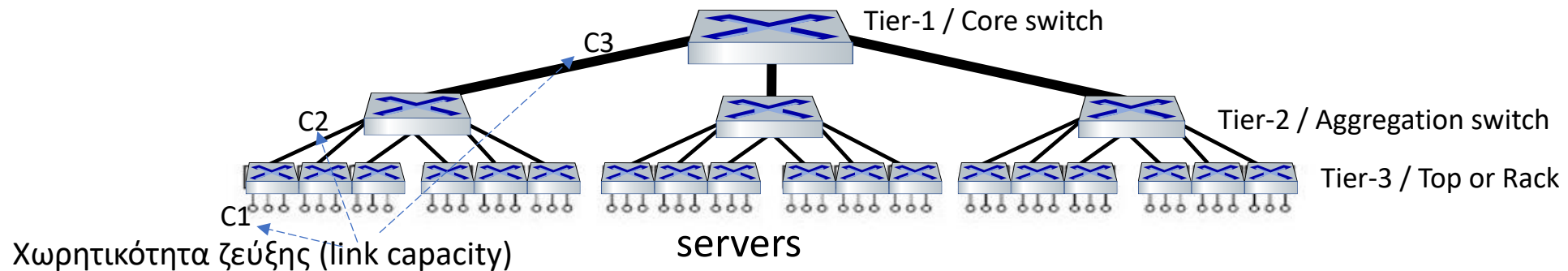
<https://blogs.cisco.com/sp/five-things-that-are-bigger-than-the-internet-findings-from-this-years-global-cloud-index>

Δίκτυο κέντρου δεδομένων: αρχιτεκτονική

- Ιεραρχική τοπολογία: κίνηση από εξωτερικούς χρήστες προς πολλαπλούς servers (north-south NS) αλλά και κίνηση μεταξύ των servers (east-west EW)
- border router(s) για σύνδεση με το Διαδίκτυο, access routers για χωρισμό σε τμήματα
- Κάτω από έναν access router υπάρχουν βαθμίδες μεταγωγέων (switches) που σχηματίζουν *fat tree* (επόμενες διαφάνειες)
- Εξισορροπητές (υπολογιστικού) φορτίου, λαμβάνουν την εξωτερική κίνηση και την κατανομίζουν στους servers
- Πολλοί servers μπορεί να τρέχουν την ίδια υπηρεσία (π.χ. κάτω από τον ίδιο access router)

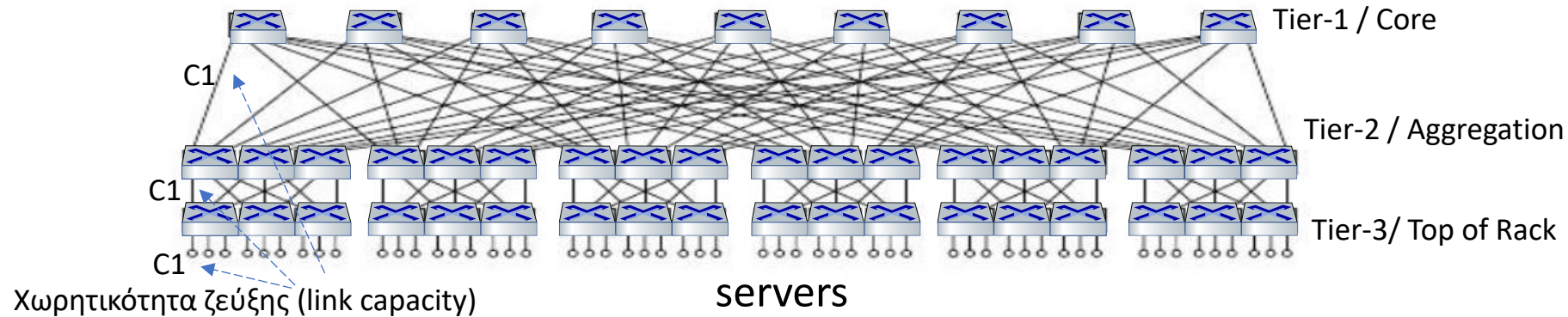


Δίκτυο κέντρου δεδομένων - Fat tree

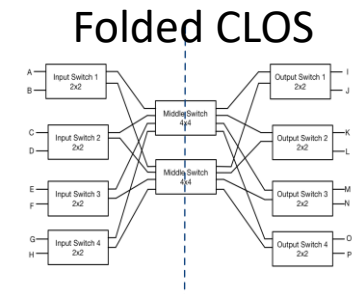


- 3 επίπεδα μεταγωγών
 - Tier 1 (**core**) switch, Tier 2 (**aggregation**) switch, Tier 3 or **Top of Rack** (ToR) switch
- Στόχος: πλήρης ρυθμός προς/από το Διαδίκτυο στους servers (NS), και μεταξύ των servers (EW)
 - Το άθροισμα της χωρητικότητας των ζεύξεων να είναι το ίδιο σε κάθε επίπεδο, π.χ. $18 \cdot C1 = 6 \cdot C2 = C3$
 - Οι ζεύξεις προς τη ρίζα του δέντρου είναι «χοντρές» (**fatter**), σε σύγκριση με τα φύλλα / servers
 - Οι μεταγωγείς προς τη ρίζα είναι μεγάλοι (π.χ. μεταγωγέας στη ρίζα=όλοι οι μεταγωγείς φύλλα)
- Προβλήματα
 - Single point(s) of failure
 - «Χοντρές» ζεύξεις και μεγάλοι μεταγωγείς στα πάνω επίπεδα

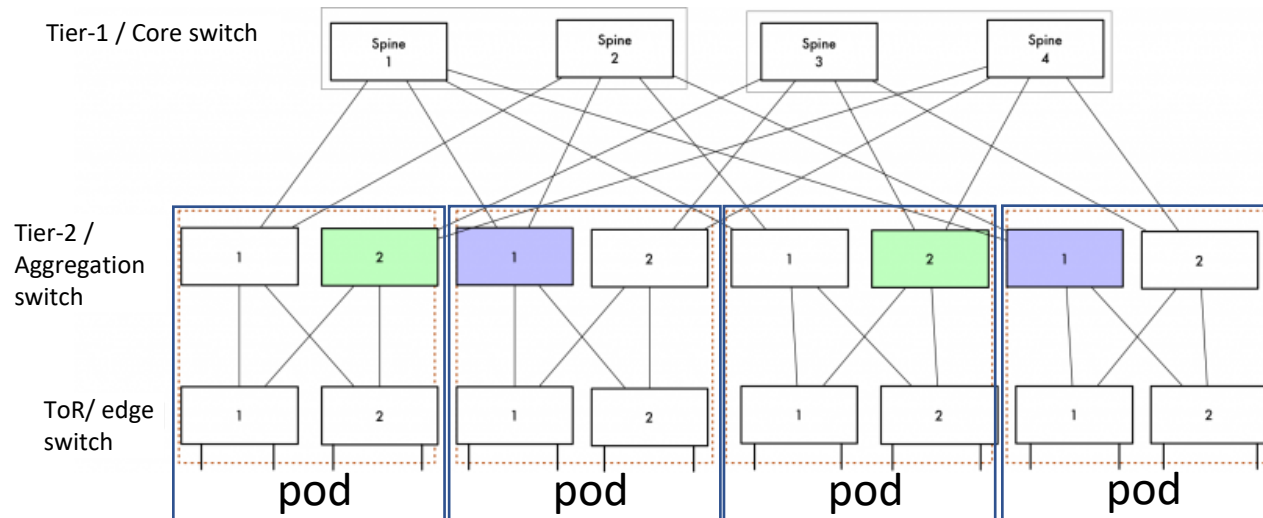
Δίκτυο κέντρου δεδομένων - Fat tree/folded Clos



- Ισοδύναμη υλοποίηση με μικρούς μεταγωγείς (συνήθως όλοι ίδιοι)
- Πολλαπλές ζεύξεις από έναν μεταγωγέα σε μεταγωγείς του πάνω και κάτω επιπέδου
- Δίκτυο Folded Clos (Charles Clos 1953)
 - Γνωστή θεωρία, πολλές εφαρμογές (τηλεφωνικά κέντρα, router internals)
- Διατηρεί τις ιδιότητες (ίδια χωρητικότητα σε κάθε επίπεδο) και αντιμετωπίζει τα προβλήματα
 - Single point of failure: Πολλαπλά μονοπάτια - αυξημένη αξιοπιστία
 - Χοντρές ζεύξεις, μεγάλοι μεταγ.: μικροί μεταγωγείς (και ίδιοι -οικονομία κλίμακας), χαμηλοί ρυθμοί

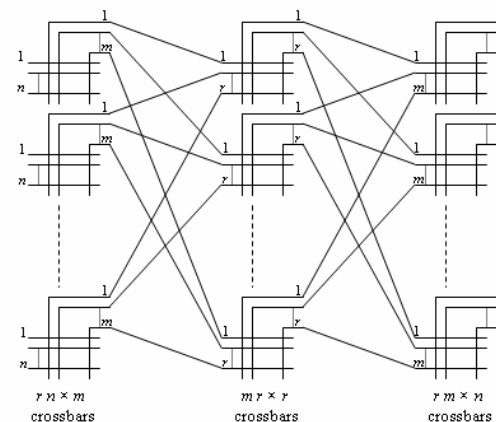


Fat tree – folded Clos - theory



Assuming a switch has “ k ” ports and the number of Levels is “ L ”

	Fat-tree with L levels	Two level Fat Tree $L=2$	Three Level Fat Tree $L=3$	Four Level Fat Tree $L=4$
Number of Core switches	$\left(\frac{k}{2}\right)^{L-1}$	$\frac{k}{2}$	$\left(\frac{k}{2}\right)^2$	$\left(\frac{k}{2}\right)^3$
Number of Hosts supported	$2\left(\frac{k}{2}\right)^L$	$2\left(\frac{k}{2}\right)^2$	$2\left(\frac{k}{2}\right)^3$	$2\left(\frac{k}{2}\right)^4$
Total Switches	$(2L-1)\left(\frac{k}{2}\right)^{L-1}$	$3\left(\frac{k}{2}\right)$	$5\left(\frac{k}{2}\right)^2$	$7\left(\frac{k}{2}\right)^3$
Number of Edge Switches	$2\left(\frac{k}{2}\right)^{L-1}$	k	$2\left(\frac{k}{2}\right)^2$	$2\left(\frac{k}{2}\right)^3$
Number of Pods	$2\left(\frac{k}{2}\right)^{L-2}$	NA	k	$k^2/2$



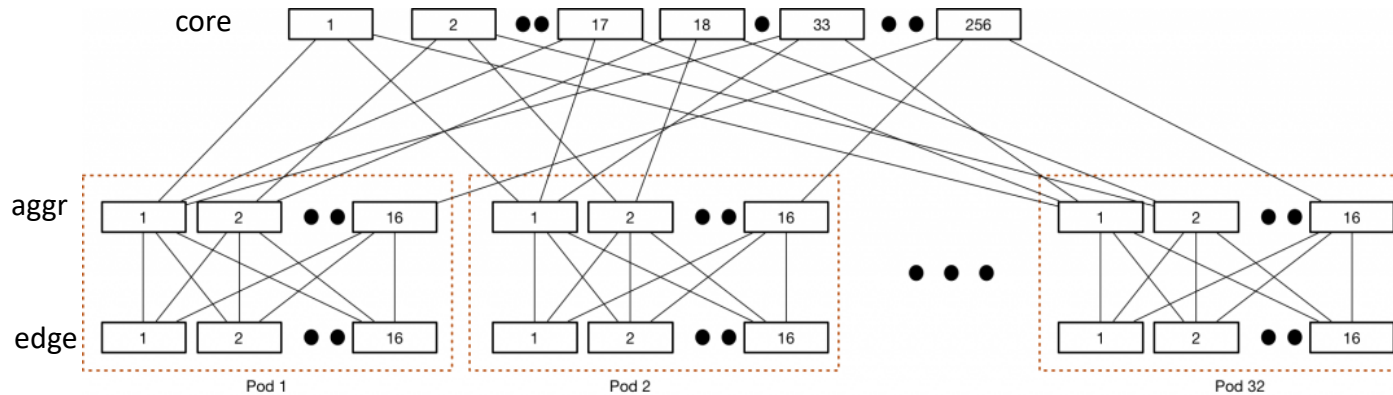
Patent US2244004A, by Edson Erwin, Bell Labs, 1938

Formalized by Charles Clos

[“A study of non-blocking switching networks”](#)

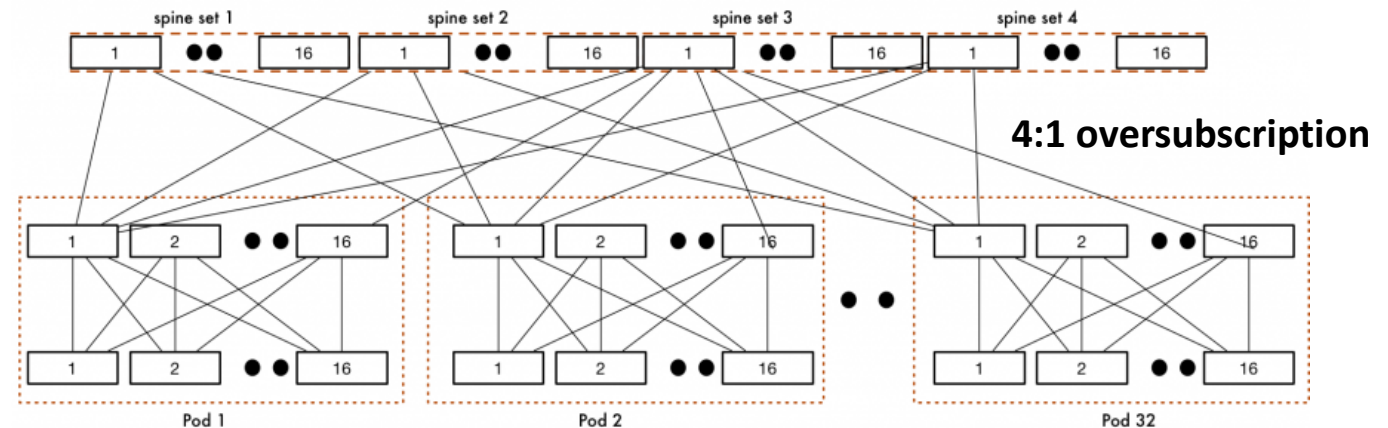
The Bell System Technical Journal (Volume: 32, Issue: 2, March 1953)

Fat – tree - oversubscription

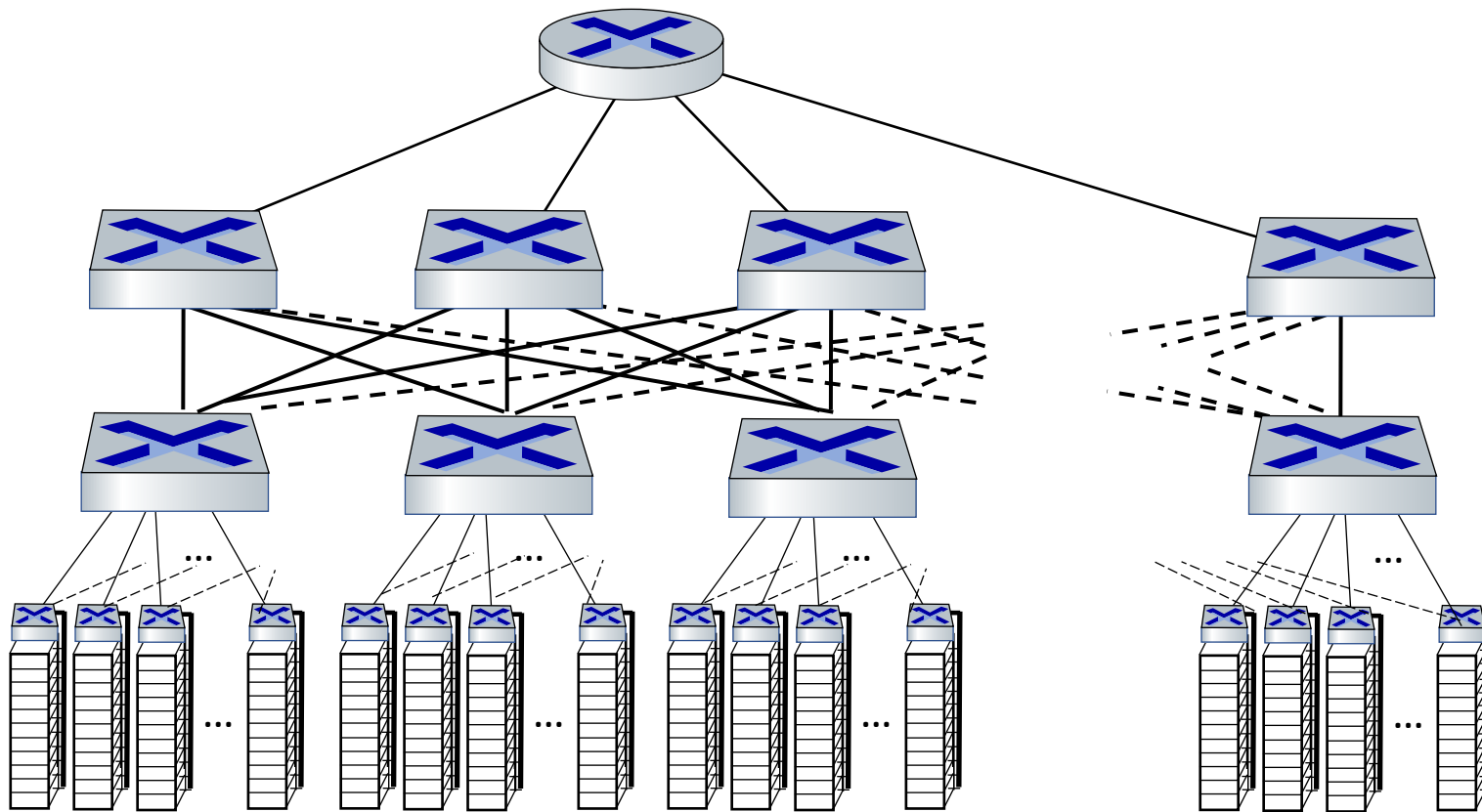


- K=32 (typical switch), L = 3
- Maximum of 256 core (spine) switches
- Maximum of 512 aggregation switches
- A total of 512 TOR (leaf) switches gives a total of 8192 (512 x 16) servers
- K=64 (also typical): 65K servers

only 64 core (spine) switches: 4 spine sets
(with 16 switches each)
Oversubscription of 4:1 at aggregation-
core and 1:1 at edge-aggregation levels



Κέντρο δεδομένων: αρχιτεκτονική



Δρομολογητής πύλης και (ανάλογα
το μέγεθος) δρομολογητές
πρόσβασης (Border, access router)

Μεταγωγείς βαθμίδας 1 (Tier-1
/spine/core)

- Π.χ συνδέονται με 32 Tier-2 switches

Μεταγωγείς βαθμίδας 2 (Tier-2,
aggregation/leaf)

- Π.χ συνδέονται με ~32 TORs switches

Top of Rack (TOR) /edge switch

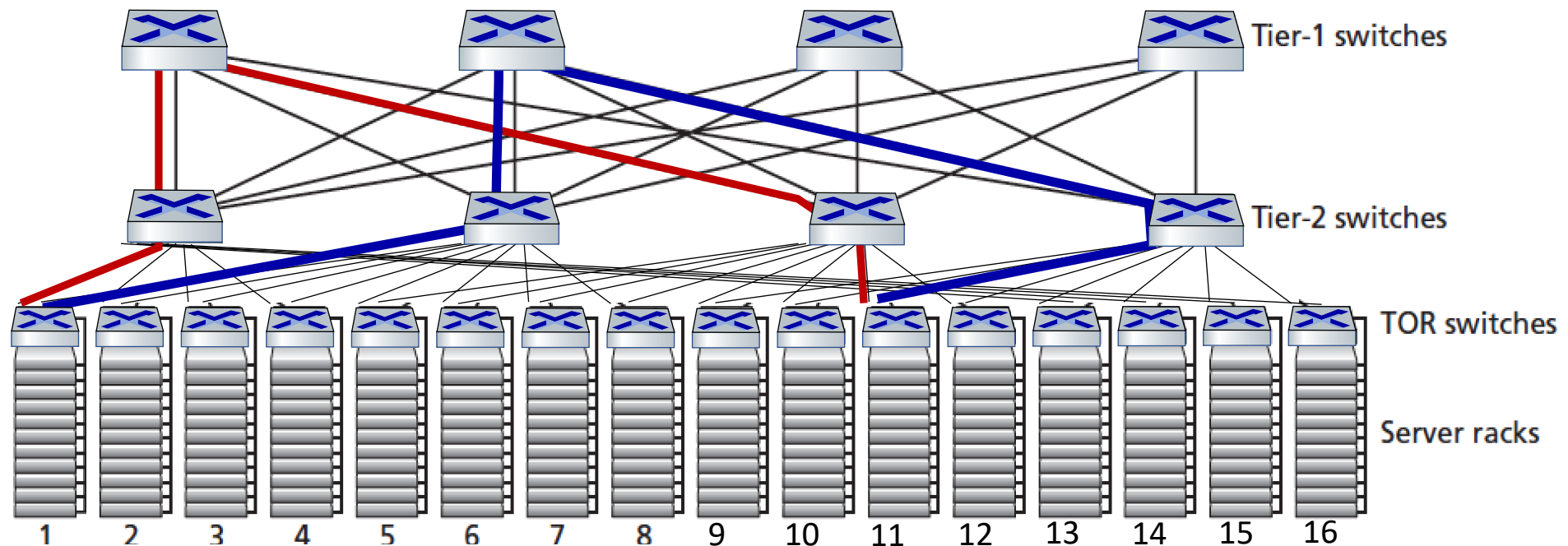
- ένας για κάθε ικρίωμα (rack)
- 25-40-100Gb/s Eth στις λεπίδες

Rack (Ικρίωμα/ράφι)

- 20- 40 λεπίδες (blades)

Πολλαπλά μονοπάτια

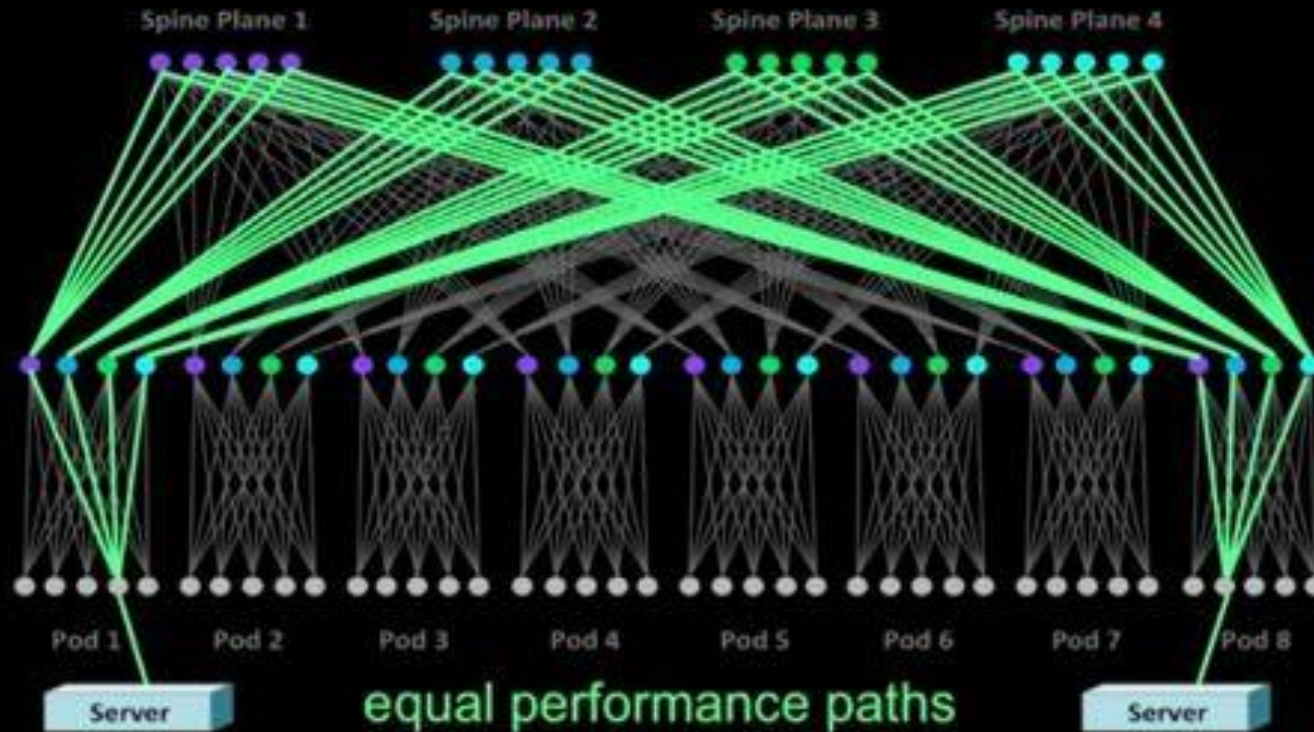
- Αυξημένη απόδοση μεταξύ racks (δυνατότητα πολλαπλών και ανεξάρτητων διαδρομών)
- Αυξημένη αξιοπιστία μέσω πλεονασμού
- Παραδοσιακά τα Ethernet - L2 δίκτυα τρέχουν spanning tree protocol (STP) και αυτό-εκμάθηση διαδρομών και δεν μπορούν να χρησιμοποιήσουν πολλαπλά μονοπάτια
- Το δίκτυο λειτουργεί με SDN / L2&L3 και χρησιμοποιεί παραλλαγές του Equal Cost Multiple Paths (ECMP) protocol (L3) που επιτρέπουν πολλαπλά μονοπάτια, και δεν επαφίονται στην αυτοεκμάθηση



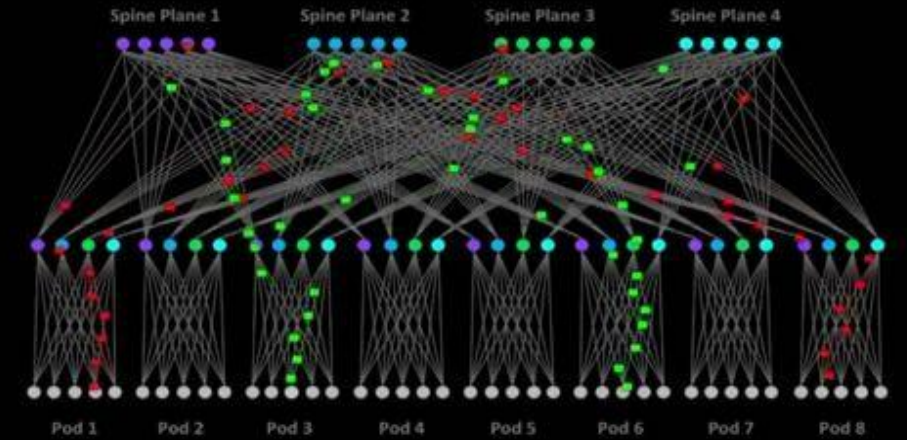
two **disjoint** paths highlighted between racks 1 and 11

CLOS Fat-tree – multiple paths – load balancing

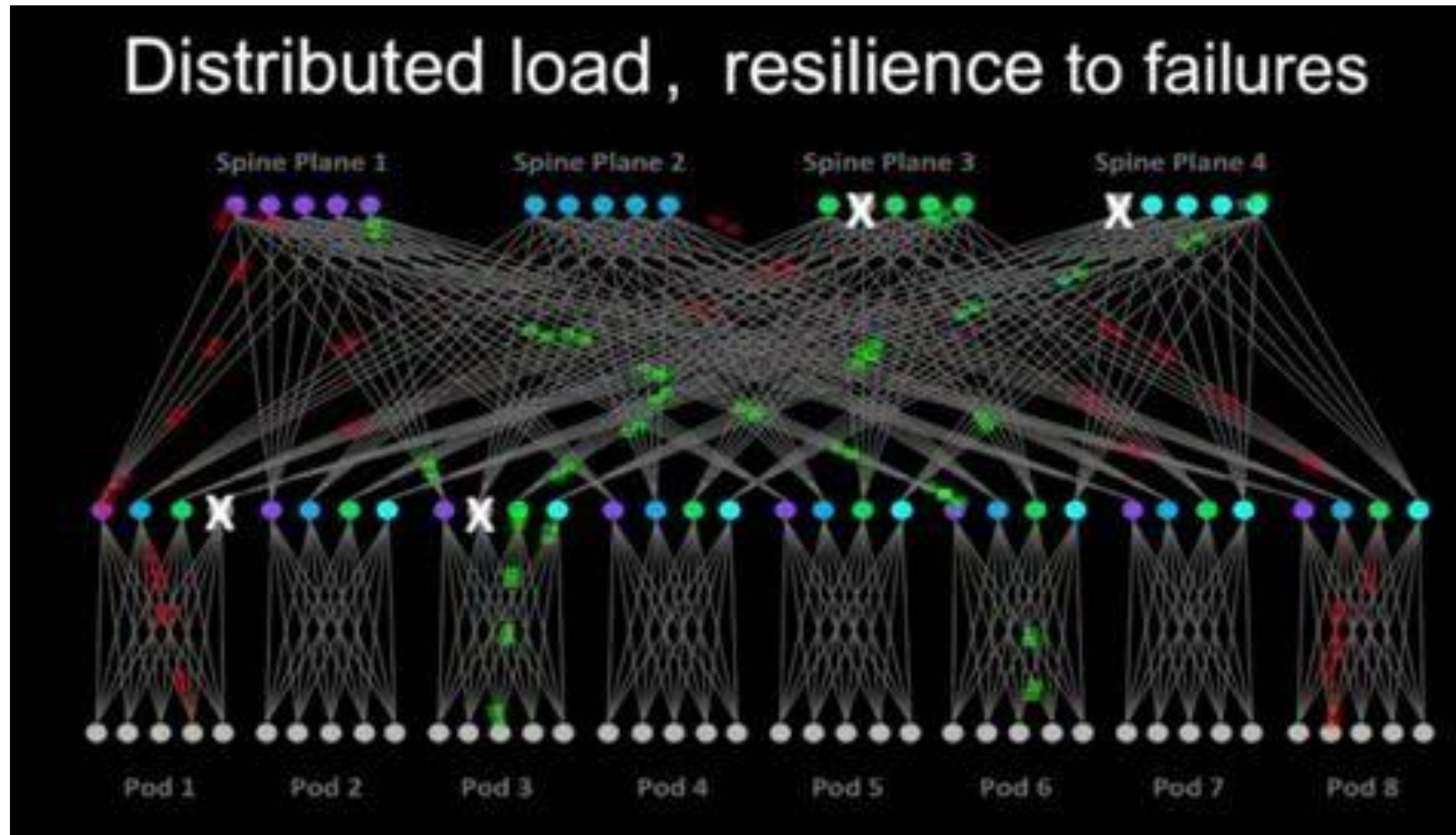
Many paths between servers



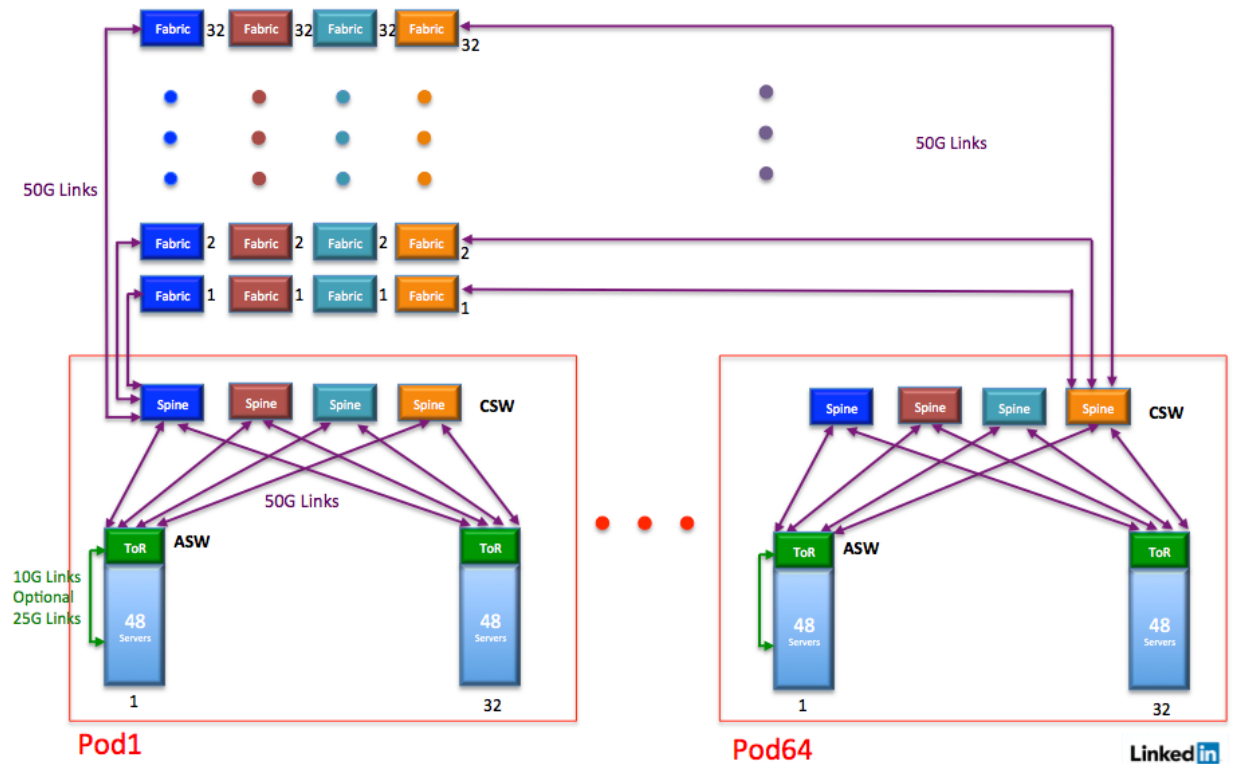
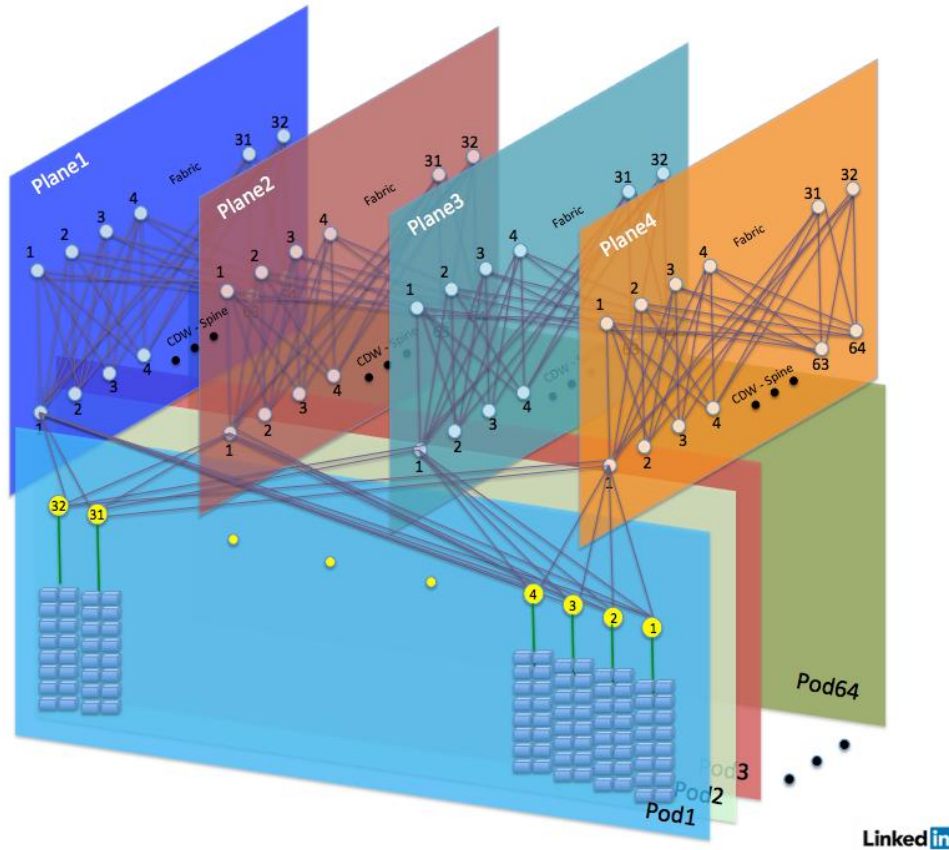
Distributed load



CLOS Fat-tree – resilience to failures

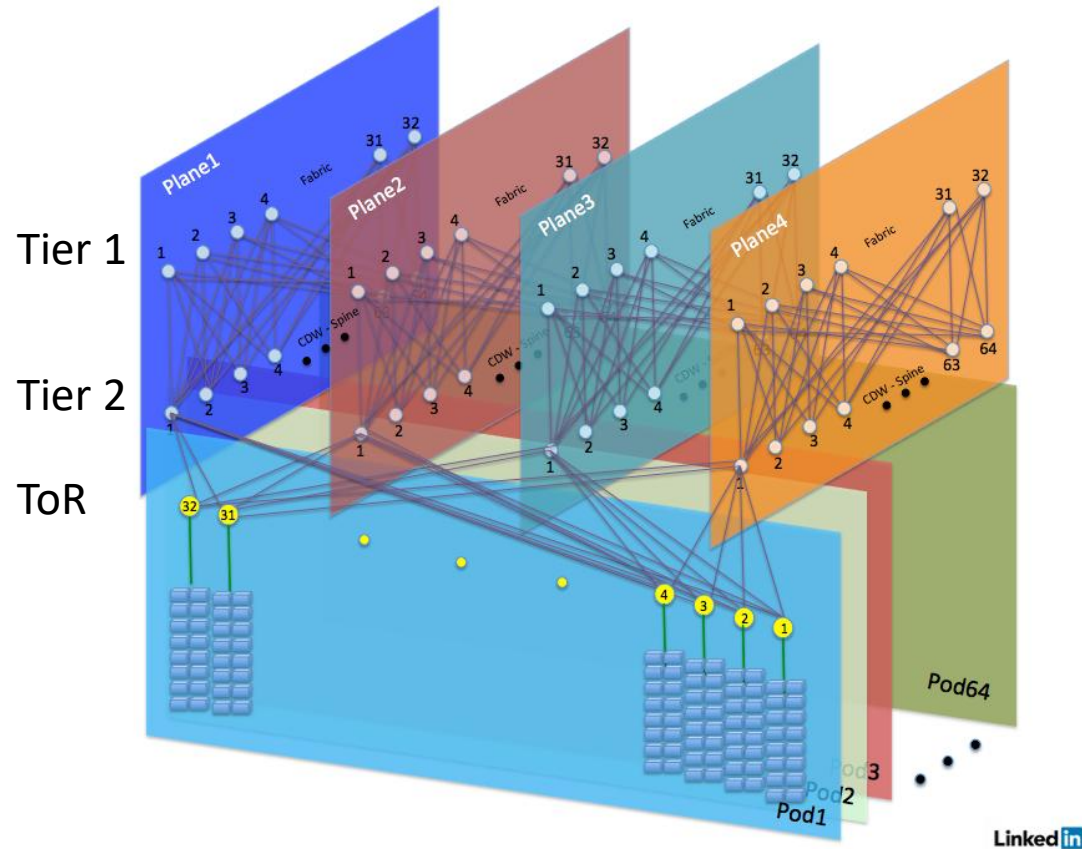


Linkedin Data Center network architecture



<https://engineering.linkedin.com/blog/2016/03/the-linkedin-data-center-100g-transformation>

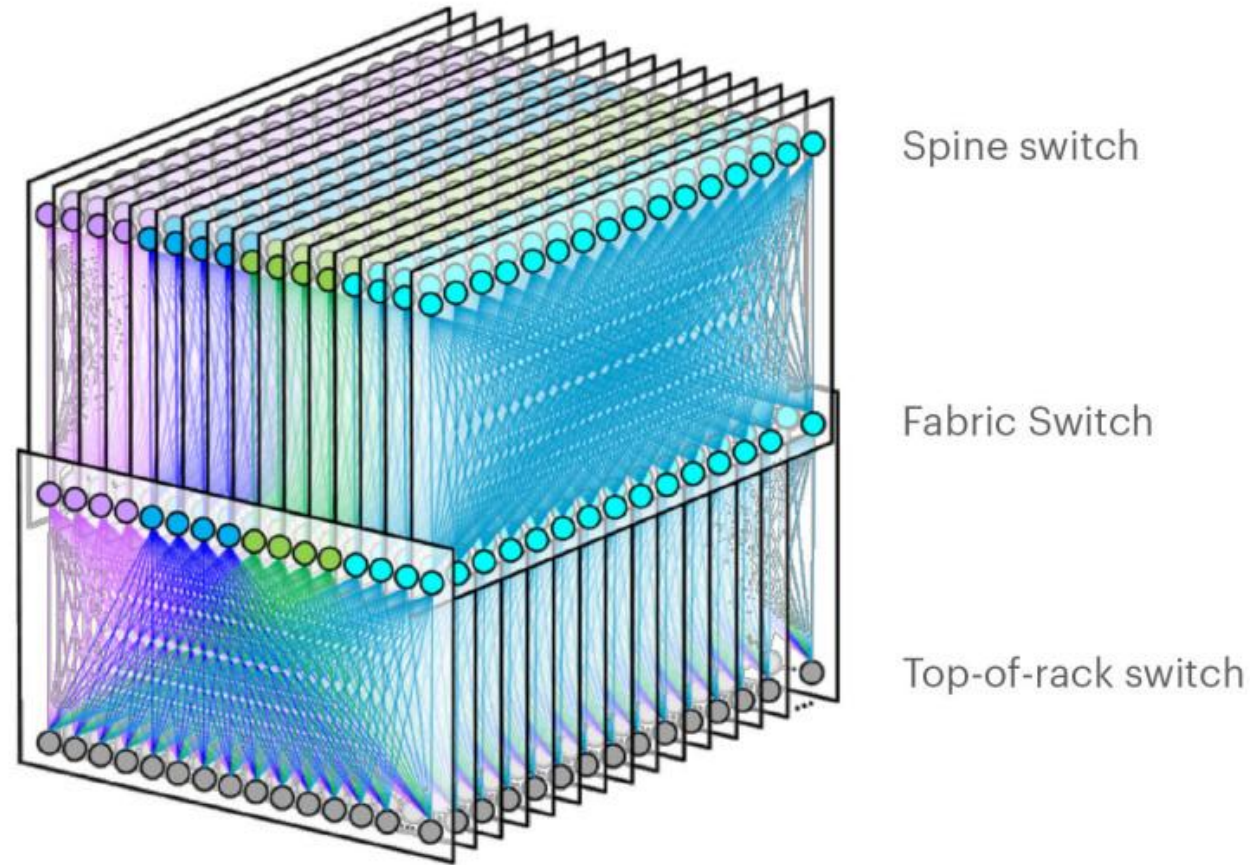
Linkedin Data Center network architecture



- Switches with radix of 64x50G where ports can be combined/split so as to be 25/50/100G
- Fat tree ($k=64, l=3$)
- TOR switches: 32x50G to servers + 32x50G to Tier 2 ('spine') switches
- Tier 2 switches: 32x50G to TORs and 32x50G to Tier 1 ('fabric') switches
- Tier 1 switches: 64x50G to Tier 2 switches
- Actually, each TOR uses 28x50G ports, it is connected to 48 servers with 25G (24x50G ports to servers) servers and to 4 Tier 2 switches with 50G (4x50G ports to Tier 2)
 - Oversubscription of $24:4 = 6:1$ between ToR - Tier 2 levels

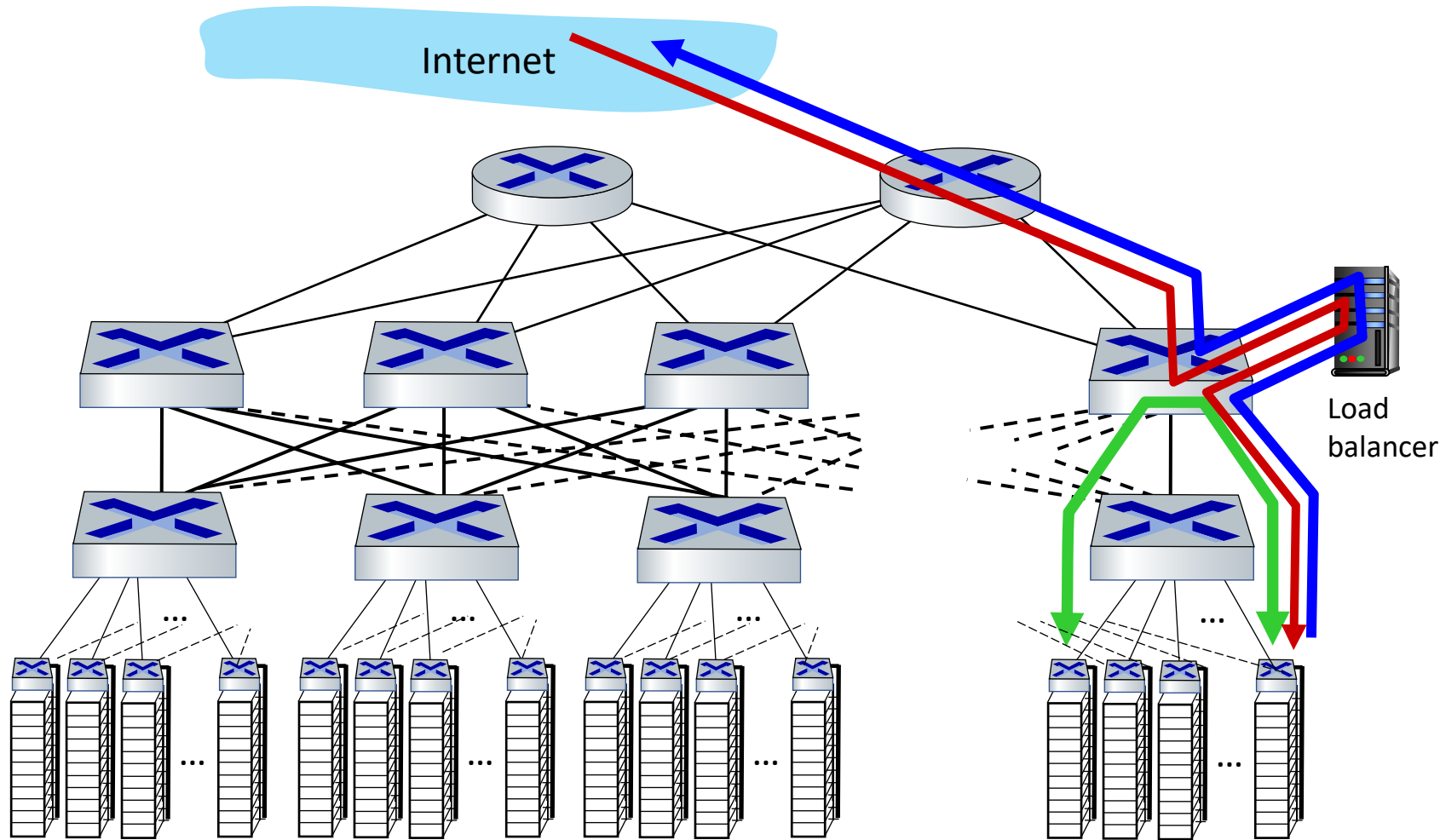
Αρχιτεκτονική Facebook F16

Equivalent implementation
with single type of switches /
single type of links



<https://engineering.fb.com/data-center-engineering/f16-minipack/> (posted 3/2019)

Δίκτυα κέντρων δεδομένων: δρομολόγηση επιπέδου εφαρμογής

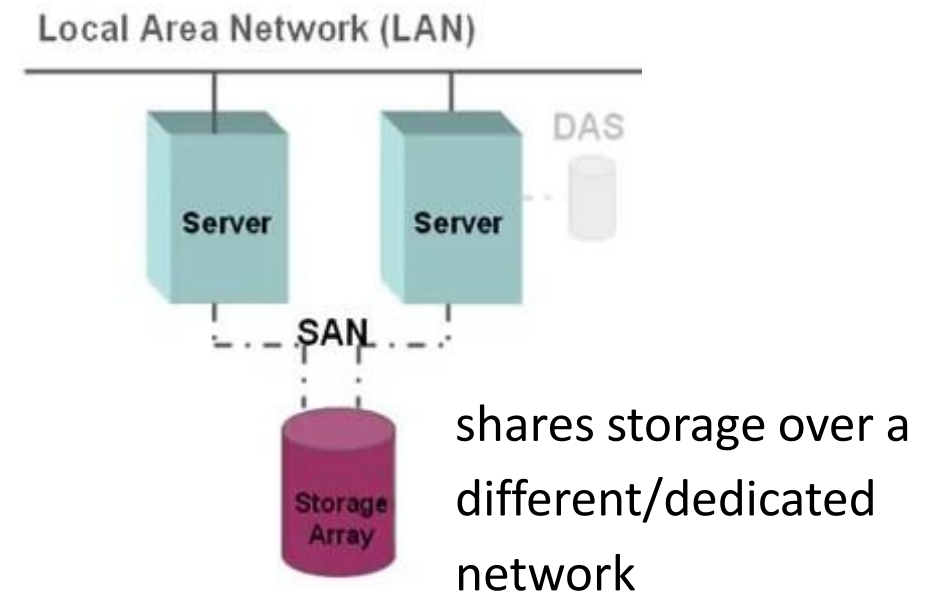
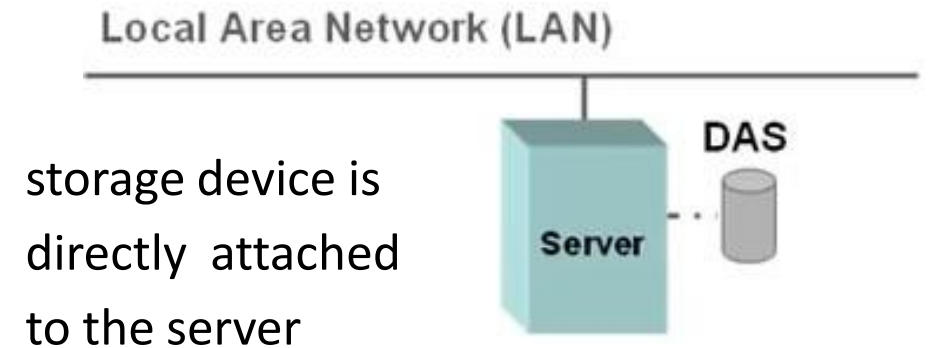
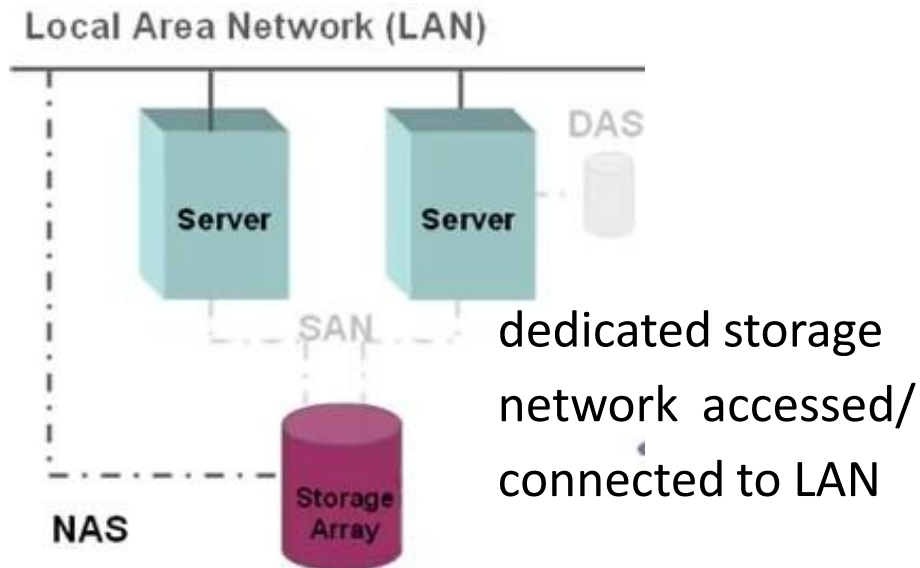


Εξισορροπιστής φορτίου: δρομολόγηση επιπέδου εφαρμογής

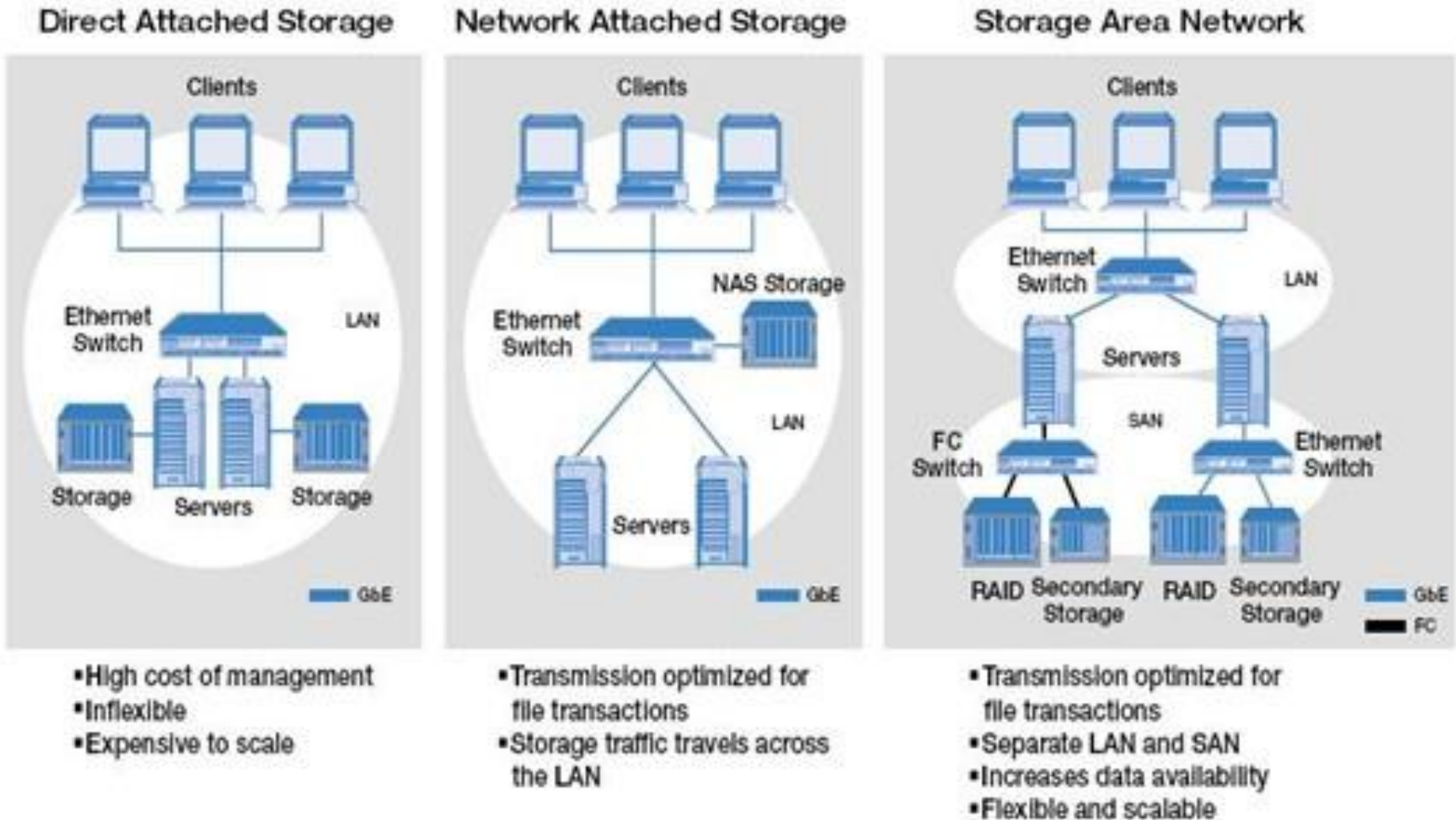
- λαμβάνει εξωτερικές αιτήσεις πελατών
- κατευθύνει/αναθέτει το φόρτο στους servers
- επιστρέφει τα αποτελέσματα στον εξωτερικό πελάτη (κρύβοντας τα εσωτερικά του datacenter από τον πελάτη)

Storage in Datacenters

- Direct Attached Storage (DAS)
- Network Attached Storage (NAS)
- Storage Area Network (SAN)



Storage Networking options



Δίκτυα κέντρων δεδομένων: καινοτομίες στα πρωτόκολλα

- Επίπεδο Ζεύξης (Link layer):
 - RoCE: remote direct memory access (RDMA) πάνω από Ethernet
- Δρομολόγηση (routing)
 - SDN χρησιμοποιείται κατά κόρον σε datacenters
 - Συνδυασμός L2 & L3 τεχνικών, VLAN, ECMP extensions/variations
- Επίπεδο μεταφοράς (Transport layer):
 - ECN (explicit congestion notification) χρησιμοποιείται στον έλεγχο συμφόρησης νέων γεύσεων (flavor) TCP (DCTCP, DCQCN)
 - Πειραματισμός με έλεγχο συμφόρησης σε κάθε άλμα (hop-by-hop - backpressure)
- Διαχείριση (management)
 - Common platform, virtualization, etc
 - Μηχανισμοί για την τοποθέτηση υπηρεσιών και των δεδομένων τους όσο το δυνατόν πιο κοντά (π.χ. στο ίδιο rack ή κοντινά) για ελαχιστοποίηση της επικοινωνίας