

Προγραμματισμός Ιστού

Αρχιτεκτονική, Πρωτόκολλα και
Πρότυπα Βασικών Υπηρεσιών

Ευάγγελος Σακκόπουλος
Επικ. Καθηγητής

Αντικείμενο Μαθήματος...

- **Αρχιτεκτονική, Βασικά Πρωτόκολλα, εφαρμογές και υπηρεσίες Διαδικτύου (HTTP, DNS), Ο ρόλος του Εξυπηρετητή web.**
- **Η Γλώσσα HTML, CSS και εισαγωγή στη D-HTML.**
- **Client-Side Scripting: Προχωρημένα θέματα D-HTML & Javascript**
- **Προγραμματισμός εξυπηρετητή: PHP, ASP.NET και επικοινωνία με βάσεις δεδομένων.**

Αντικείμενο Μαθήματος

- **Ανάλυση εννοιών XML, XSLT.**
- **Προχωρημένα θέματα XML και Web Services.**
- **AJAX (Asynchronous JavaScript And XML)**
- **Σχεδιασμός και ανάπτυξη εφαρμογών**
- **Ασφάλεια στο Internet: Secure HTTP, PKI**
- **Τεχνολογίες Σημασιολογικού Διαδικτύου (Semantic Web)**

Διάρθρωση Μαθήματος...

- Διαλέξεις Μαθήματος Τετάρτη 6-10μμ
- Απαλλακτική Εργασία και Παρουσίαση σε αντικείμενα για Ανάπτυξη με Τεχνολογίες Εφαρμογών Διαδικτύου (ομάδες των 2 ή 3 ατόμων)
- Θα ενημερώνεστε από το gunet2

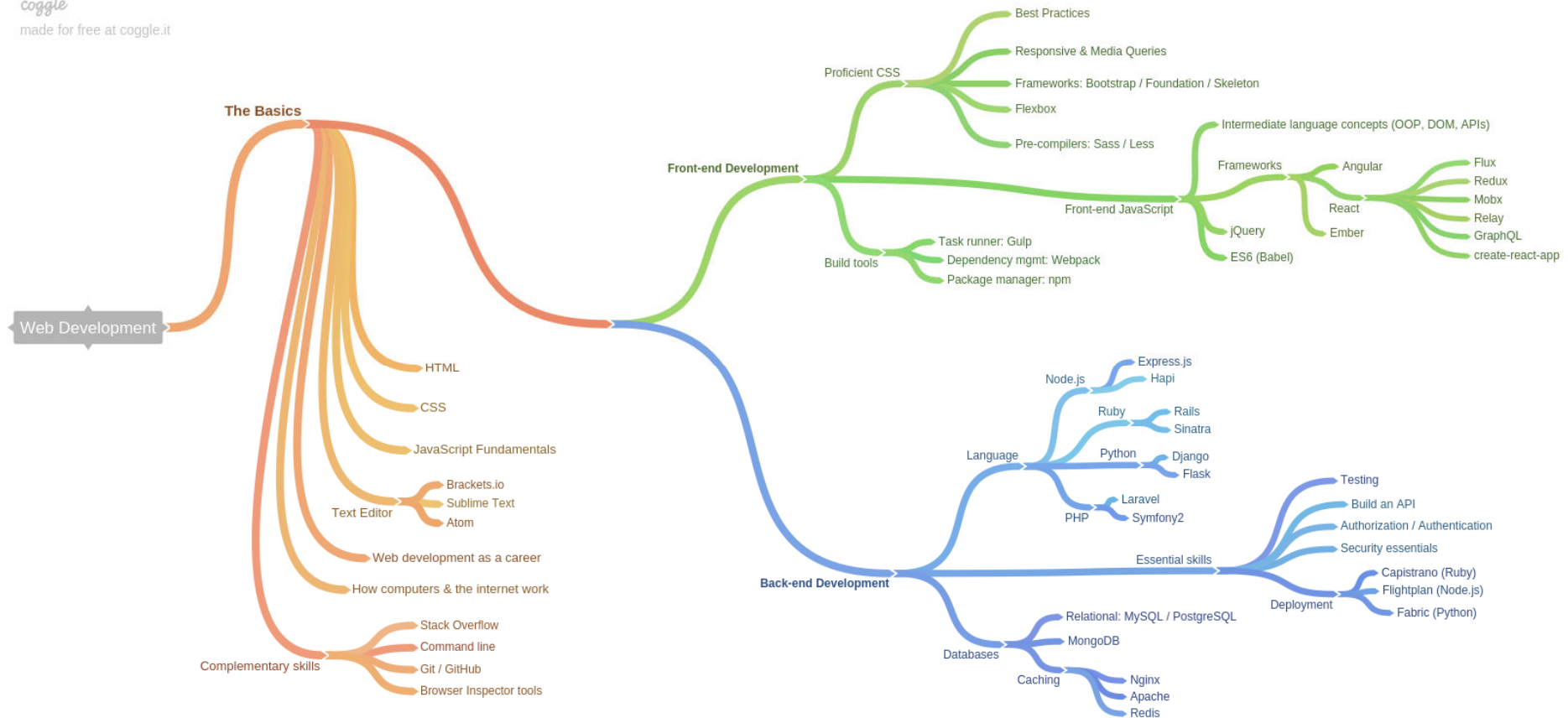
<http://gunet2.cs.unipi.gr/courses/TMH109/index.php>

Γενικό Πλάνο

Ο προγραμματισμός Ιστού ΣΥΝΟΛΙΚΑ

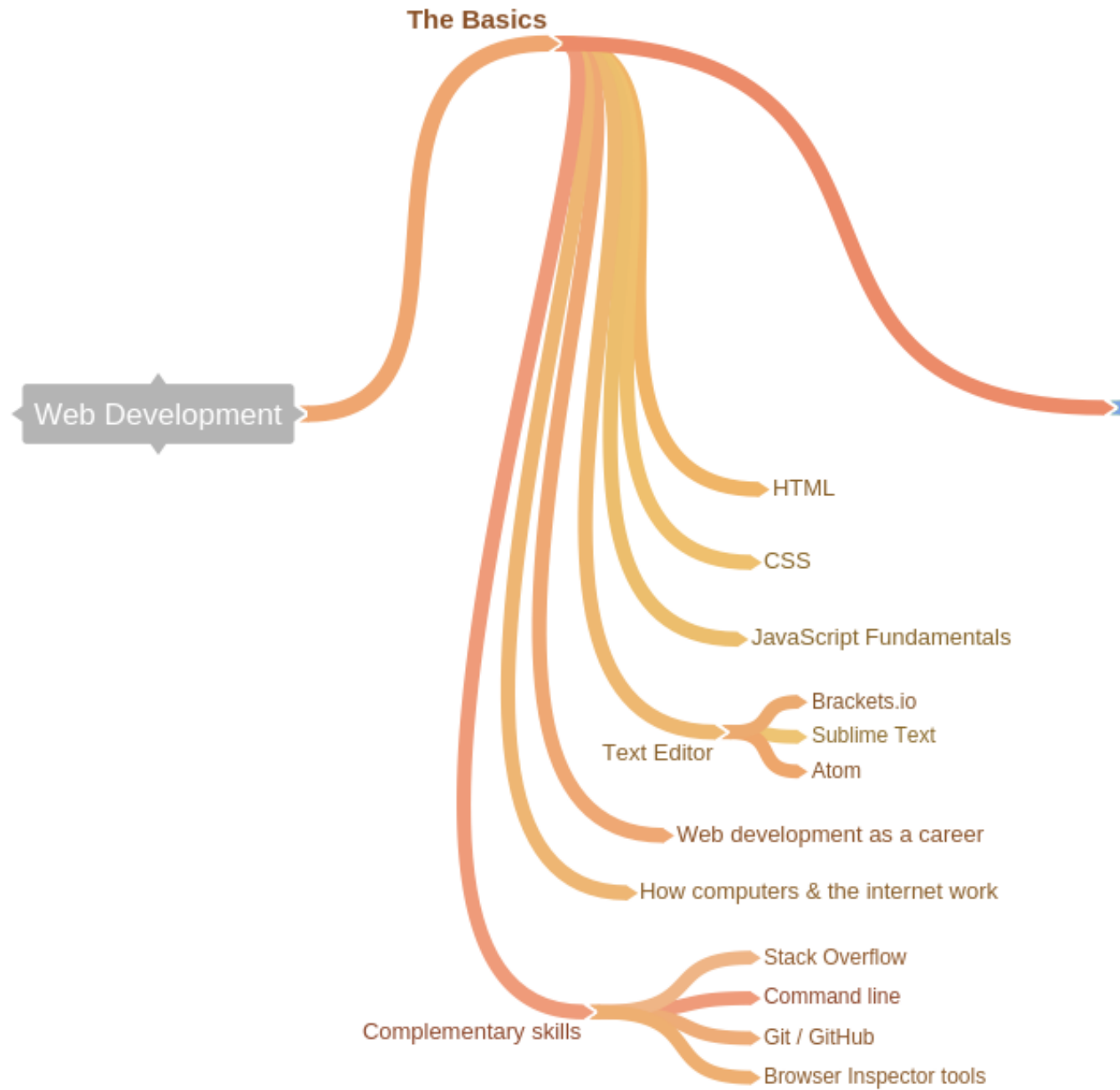
Ο προγραμματισμός Ιστού

coggle
made for free at coggle.it

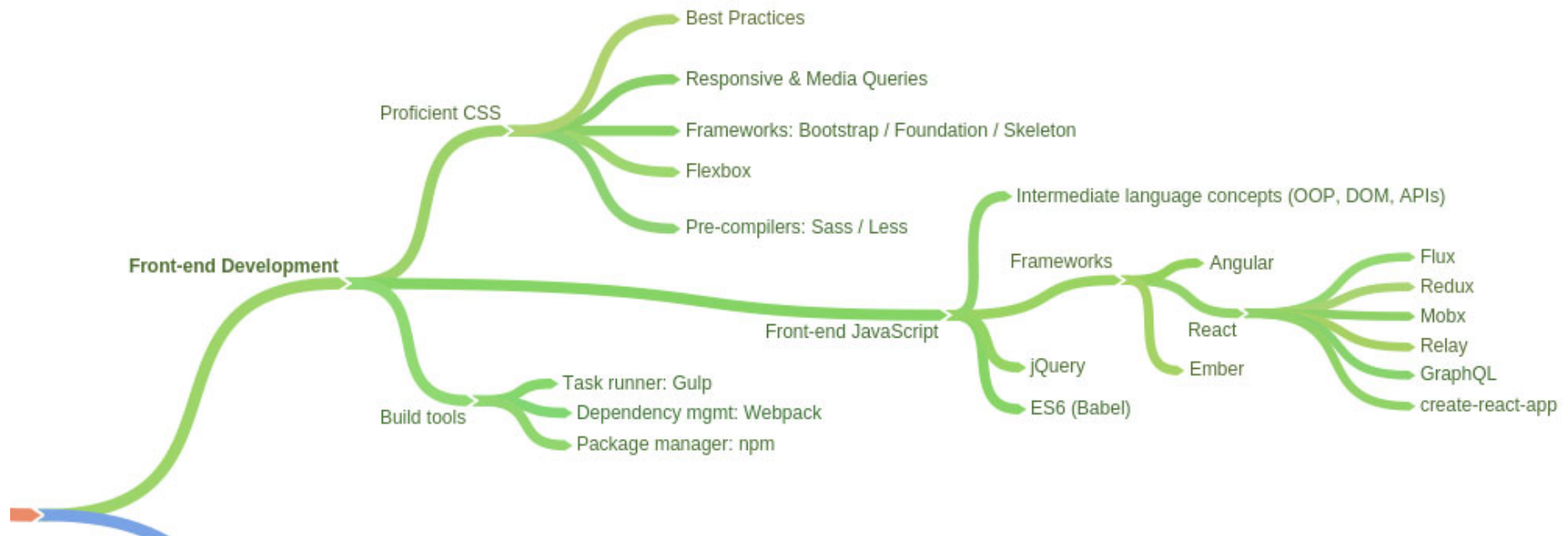


web programming flow

B
α
σ
ι
K
α

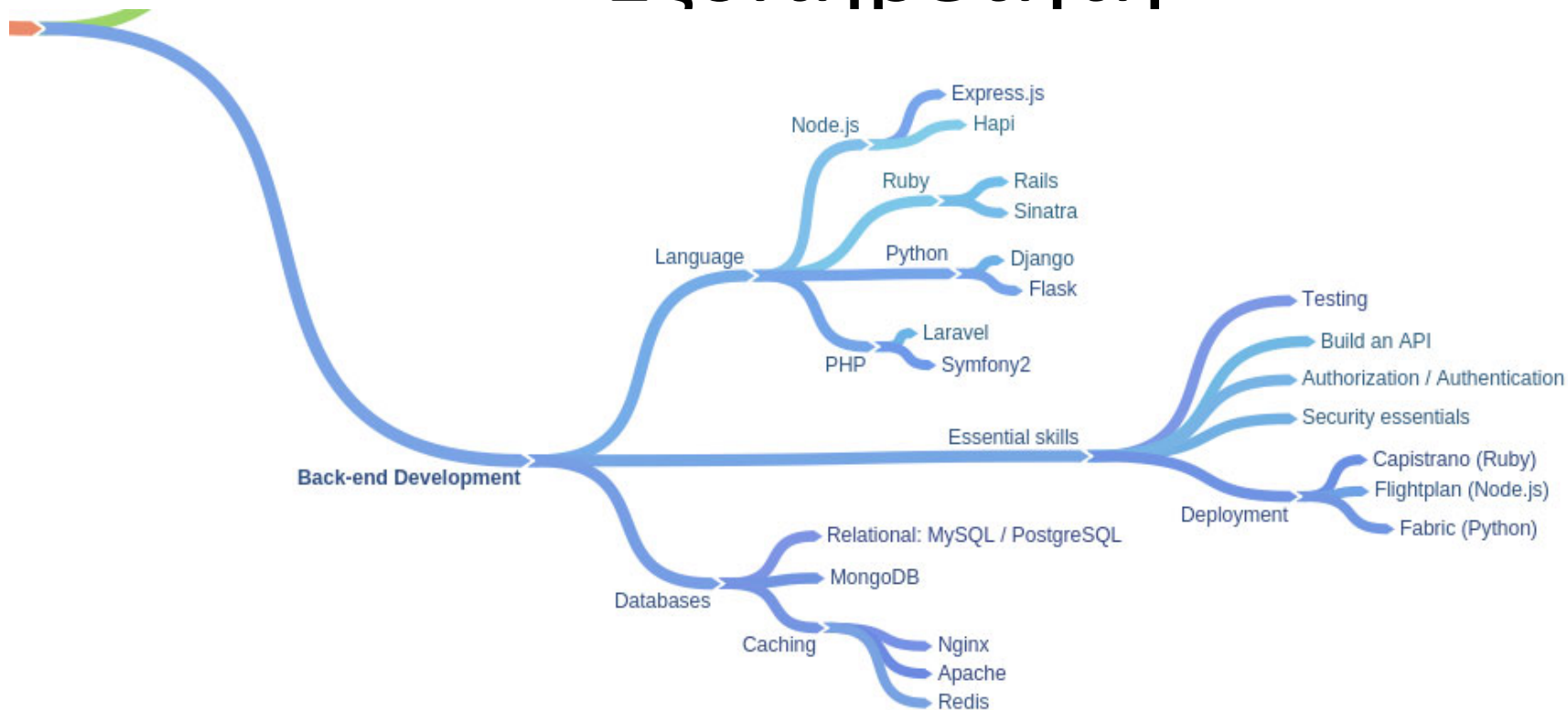


Προγραμματισμός στον Πελάτη



web programming flow

Προγραμματισμός στον Εξυπηρετητή



web programming flow

Αναδρομή

Internet και World Wide Web

- Ποιο ήρθε πρώτα το Internet ή το WWW?

The Internet

- Internet είναι ένα δίκτυο διασυνδεδεμένων υπολογιστών που πλέον έγινε παγκόσμιο
- Internet 1969 - ARPANET
- 1969 ARPANET συνέδεε το UCLA, Stanford, UCSB, Univ. Utah

Ιστορικά Στοιχεία

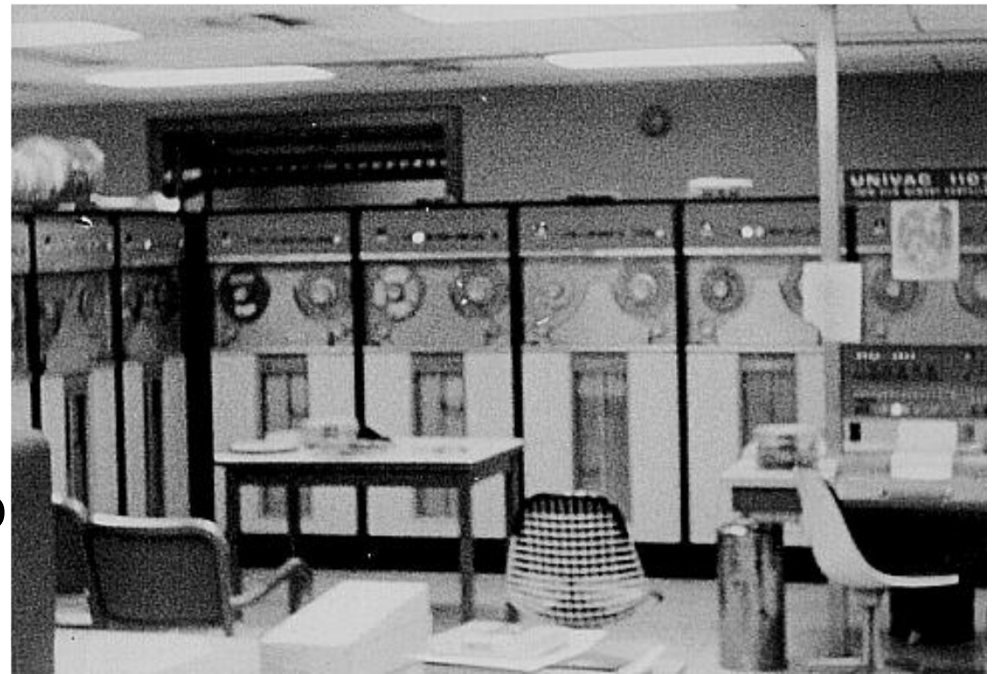
- ARPANET
 - Ο «παππούς» των δικτύων
 - Ανατύχθηκε από την DARPA (Advanced Research Projects Agency) στα τέλη του '60
 - Το '83 διασπάται σε MILNET και ARPANET
- NSFNET
 - Δίκτυο τριών επιπέδων (Backbone, μεσαίου επιπέδου, Campus)
- Σημαντικές Ημερομηνίες
 - Πρώτη Έκδοση HTTP → 1991
 - MOSAIC (browser) → 1993

Τεχνολογία Υπολογιστών;

- Ποια ήταν η τεχνολογία υπολογιστών στα τέλη του 1960 και στα 1970;

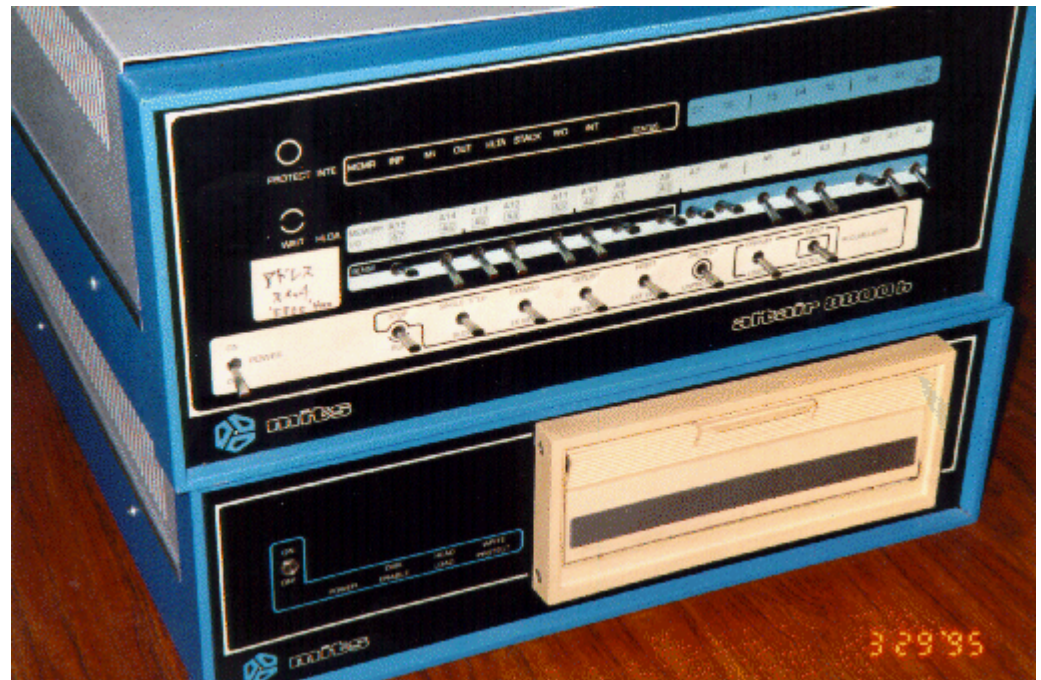
Οι υπολογιστές το 60 & 70

- Δεν υπήρχαν PC αλλά mainframe computers 60
- Στα μέσα 1970 – υπήξαν κάποιοι personal computers
 - Altair: Box with blinking lights
- Τέλη 1970 – Apple 2, το πρώτο ουσιαστικά PC



Personal Computing?

- Απλά μια συσκευή για απλούς υπολογισμούς - Just a box with blinking lights
- Χωρίς δίκτυο και διασύνδεση



Internet - 1970

- 1972 – Αναπτύχθηκε το Telnet ως τρόπος σύνδεση απομακρυσμένων υπολογιστών
- 1972 – Παρουσιάστηκε το Email
 - 1977 - U. Wisconsin το πρώτο «Μεγάλο» σύστημα email με 100 χρήστες
- 1973 - ARPANET γίνεται διεθνές
- 1973 - File Transfer Protocol (FTP)

Τεχνολογία Υπολογιστών; (2)

- Ποια ήταν η κατάσταση στα 1980;

Computers 1980

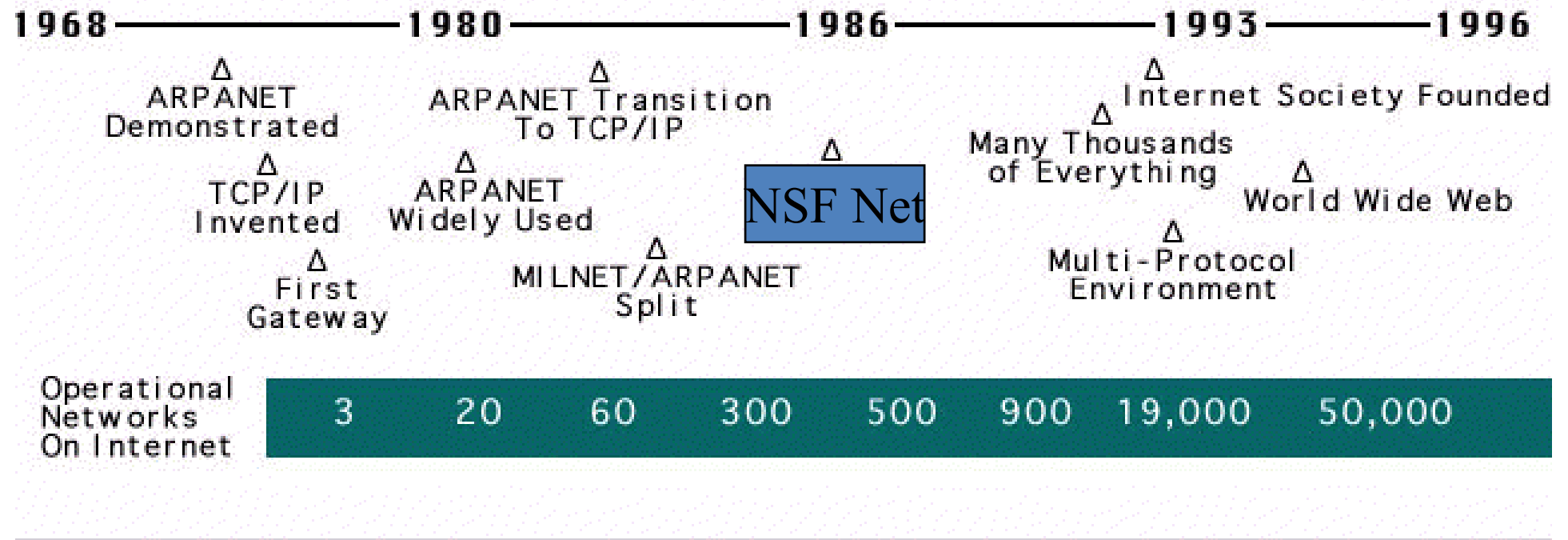
- 1981 – IBM PC
- 1984 – Apple Macintosh
- 1986 – Modem ως περιφερειακό του PC

Internet - 1980

- 1984 - Domain Name Server
 - Επιτρέπει την ονομασία (όχι πια μόνο αριθμοί)
- 1986 - NSFNET
 - Στα 1990, γίνεται το backbone του σύγχρονου Internet
 - Ιδιωτικό μετά το 1995

Barry M. Leiner, Vinton G. Cerf, David D. Clark, Robert E. Kahn, Leonard Kleinrock, Daniel C. Lynch, Jon Postel, Larry G. Roberts, Stephen Wolff. A Brief History of the Internet. Internet Society.
<http://www.isoc.org/internet/history/brief.shtml>

Internet Timeline



Internet 1990

- 1991 - Tim Berners-Lee παρουσιάζει το **World Wide Web!**
 - TBL είναι προγραμματιστής στο CERN, ένα εργαστήριο φυσικών επιστημών στην Ελβετία
- 1993 - Mosaic (becomes Netscape) designed by graduate students at University of Illinois
 - first point-and-click browser
 - later developed into Netscape Navigator
- These are the two most significant events in the formation of the WWW

Internet 1990s

- 1991 - Tim Berners-Lee releases **World Wide Web!**
 - TBL is computer programmer at CERN, a physics lab in Europe (book *Weaving the Web* by TBL)
- 1993 - Mosaic (που γίνεται στη συνέχεια Netscape) από φοιτητές του University of Illinois
 - Ο πρώτο διαφυλλιστής - browser (point-and-click)
- These are the two most significant events in the formation of the WWW

World Wide Web

- Μέσω του Internet, οι υπολογιστές μπορούν να επικοινωνούν μεταξύ τους
- Τα δημοσιευμένα αρχεία μπορούν να τα διαβάσουν απομακρυσμένοι χρήστες
 - HyperText Markup Language (.html)
- URL - Universal Resource Locator – το όνομα συνήθως ενός υπολογιστή κάπου μακριά
 - <http://www.msu.edu/~urquhar5/tour/active.html>

HTTP

- World Wide Web χρησιμοποιεί HTTP Servers, γνωστούς και ως web server
- Λαμβάνει HTTP type request και στέλνει τα αρχεία σε πακέτα

Web Browsers

- **Mosaic (1993)** ο πρώτος **first point-and-click browser**
- **Web browsers** είναι το λογισμικό για να προσπελαύνουμε ιστοσελίδες
- Netscape Navigator, Firefox και ο Internet Explorer είναι οι πλέον διάσημοι

Τεχνολογία υπολογιστών; (3)

- Στα μέσα του 1990 τι τεχνολογίες έχουμε;

Ιστορία στα 1990

- Windows 95 GUI έκανα δημοφιλή τη χρήση PC στο ευρύ κοινό.
- Windows 95 + Internet (AOL, others) → Τεράστια αύξηση του αριθμού των οικιακών PC
- Επιπλέον ένας υπολογιστής βρέθηκε σε κάθε γραφείο (χώρων εργασίας)

Universal Resource Locator

<http://www.msu.edu/~urquhar5/tour/active.html>

http://

Δηλώνει τύπο
μεταφοράς

/~urquhar5/tour/active.html

Θέση αρχείου στον απομακρυσμένο
υπολογιστή

www.msu.edu

Domain Name -

Το όνομα του απομακρυσμένου υπολογιστή

21^{ος} Αιώνας – Διαμοίραση αρχείων - File Sharing

- Το Internet επέτρεψε τη διαμοίραση απλής πληροφορίας
- Το FTP υπήρξε ένα αρχικό σύστημα διαμοίρασης αλλά είναι κάπως δύσκολο στη χρήση
- Το WWW επέτρεψε τη πρόσβαση σε προηγμένες πληροφορίες αλλά δεν ήταν σχεδιασμένο για file-sharing
- Napster, KaZaA, Morpheus, e-mule and LimeWire είναι file-sharing.

Napster

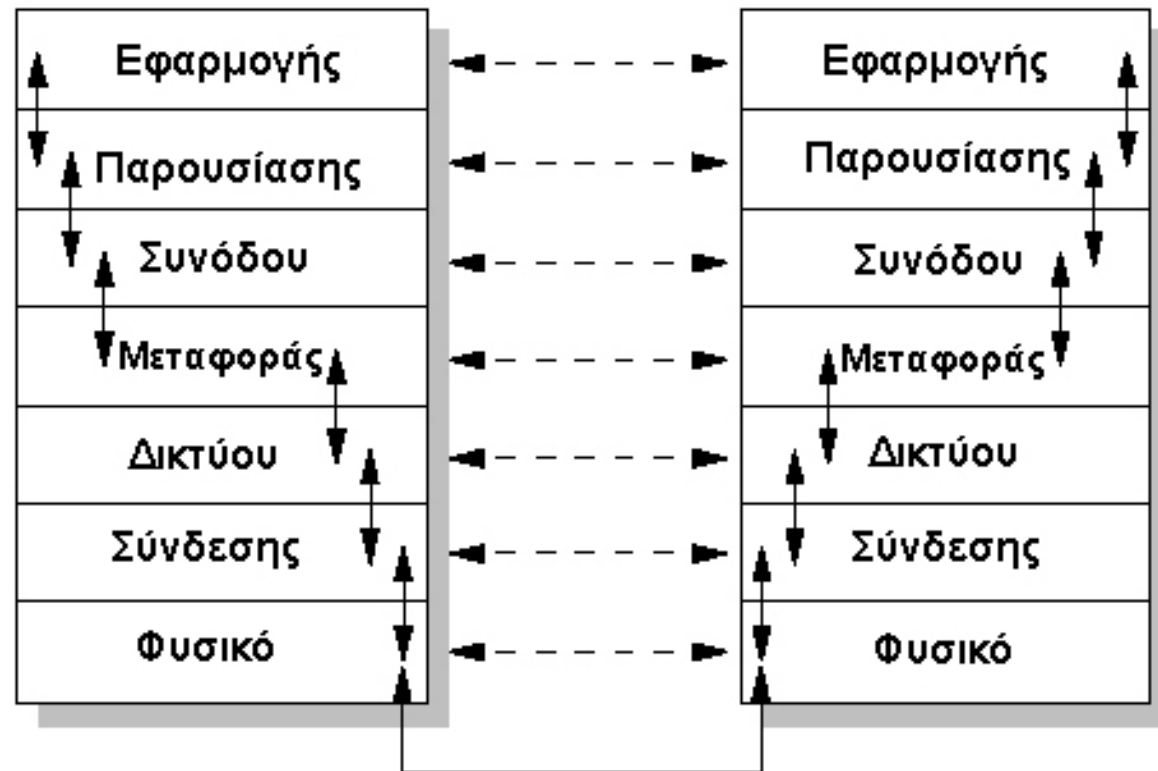
- Napster ήταν μια κοινότητα διαμοίρασης μουσικής
- Χρησιμοποίησε ένα κεντρικό server όπου καταγραφόταν ποιος είχε τι και που
- Παραβίαζε τα copyrights
- Το Napster πλέον ελέγχει τι μεταφέρεται για copyrights

Κατάρρευση των Οικονομίας της Πληροφορίας

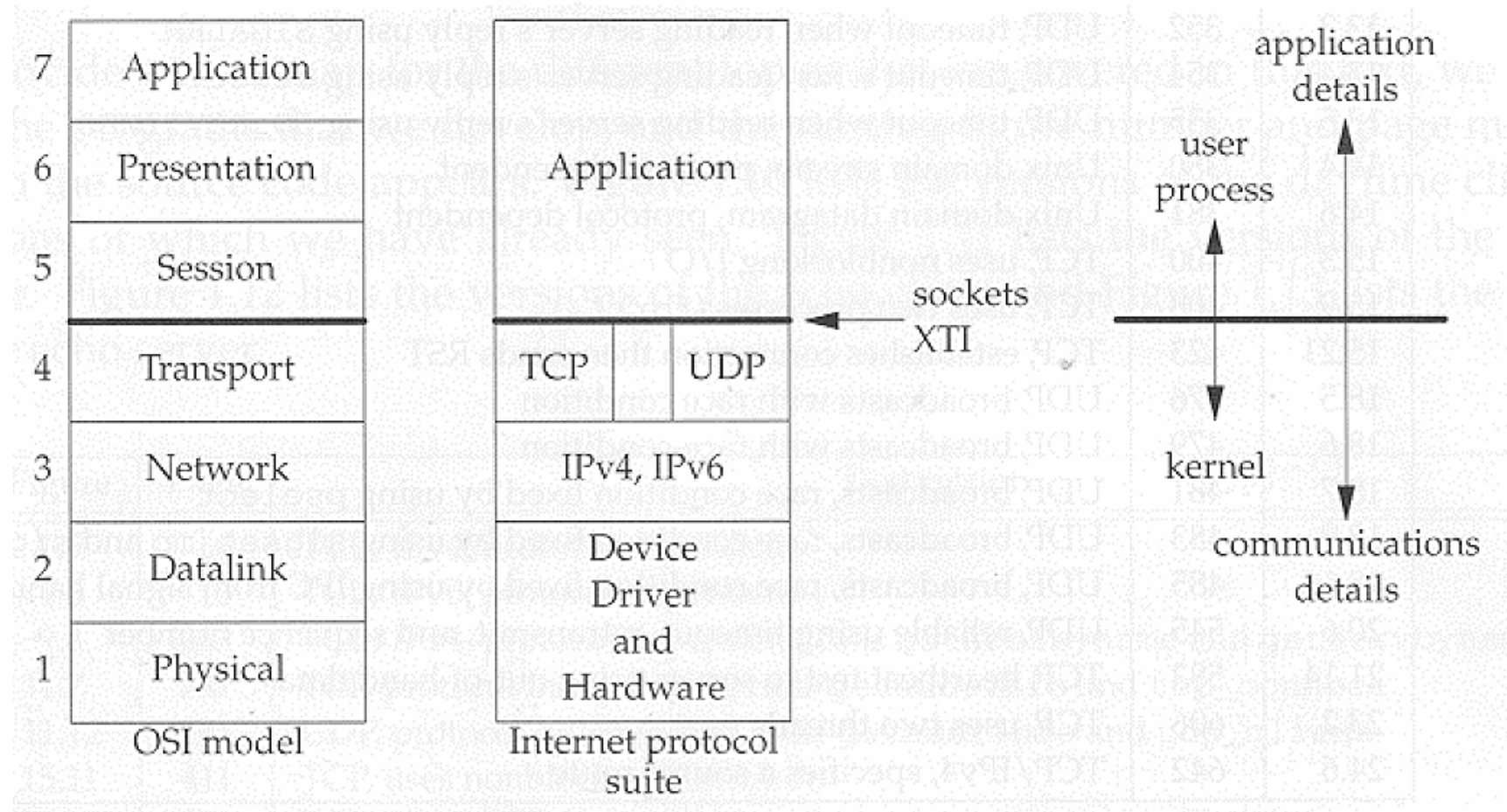
- Η τεράστια οικονομική ανάπτυξη στα τέλη του 1990 έγινε λόγω της υποσχόμενης ανάπτυξης..των Internet εταιρειών
- Οι περισσότερες δεν ήταν ποτέ κερδοφόρες
 - Amazon.com δήλωσε τα πρώτα του κέρδη το 2003 από όταν ξεκίνησε το 1997!

Αρχιτεκτονική Δικτύων

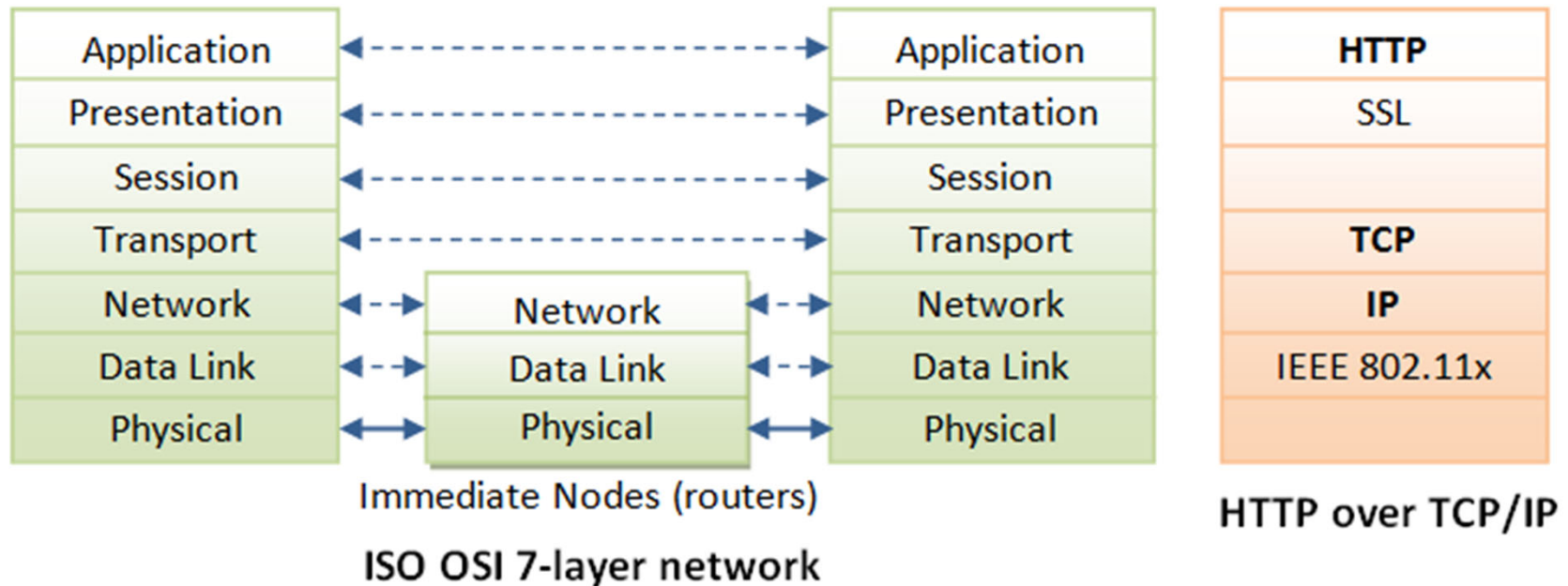
- Μοντέλο Αναφοράς OSI



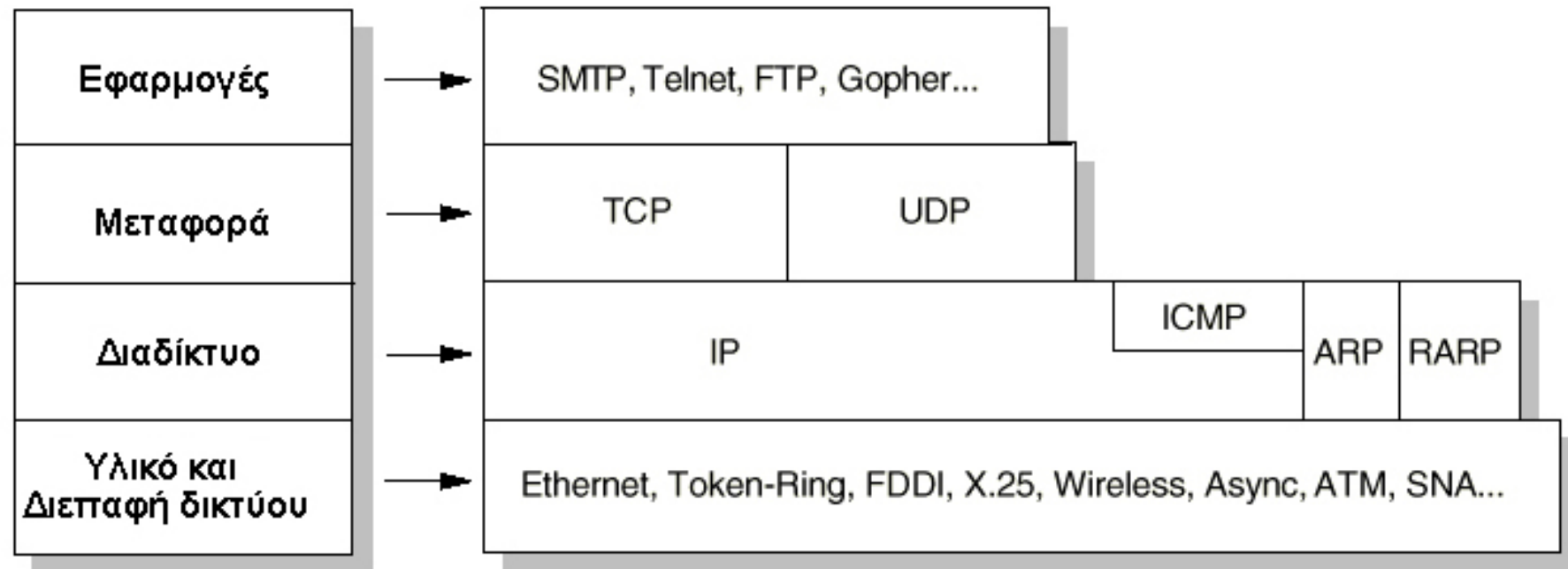
Οικογένεια Πρωτοκόλλων TCP/IP



Οικογένεια Πρωτοκόλλων TCP/IP



Οικογένεια Πρωτοκόλλων TCP/IP



Εφαρμογές: Echo, Ping, Traceroute, Whois, FTP, E-mail (SMTP, POP3, IMAP), TELNET, HTTP

Οικογένεια Πρωτοκόλλων TCP/IP

- UDP
 - Απλό πρωτόκολλο του επιπέδου μεταφοράς
 - Υπηρεσία Χωρίς Σύνδεση
 - Χωρίς Έλεγχο Λαθών
- TCP
 - Υπηρεσία Με Σύνδεση
 - Υπάρχει Έλεγχος Λαθών και Επαναμετάδοση
 - Αξιοπιστία
 - Full Duplex επικοινωνία

Οικογένεια Πρωτοκόλλων TCP/IP

- IP
 - Η καρδιά της οικογένειας πρωτοκόλλων
 - Μεταφορά Δεδομενων σε hosts διαφορετικών δικτύων
 - Προσφέρει Δρομολόγηση και Έλεγχο Λαθών
 - Τρόπος Λειτουργίας
 - Σπάσιμο σε αυτοδύναμα πακέτα
 - Μετάδοσή τους στο δίκτυο
 - Επανασυναρμολόγηση στον προορισμό

IP Διευθύνσεις

- Μοναδικοί αριθμοί 32-bit
- Δίνονται από το InterNIC
- Γενική Μορφή: x.x.x.x (x: 0-255)
- Τα πρώτα bits της διεύθυνσης δηλώνουν την κατηγορία δικτύου
- Υπάρχουν 5 κατηγορίες δικτύων
 - Class A-E

IP Διευθύνσεις

		7 bits	24 bits
Class A	0	Netid	Hostid

		14 bits	16 bits
Class B	1 0	Netid	Hostid

		21 bits	8 bits
Class C	1 1 0	Netid	Hostid

		28 bits
Class D	1 1 1 0	Multicast Group id

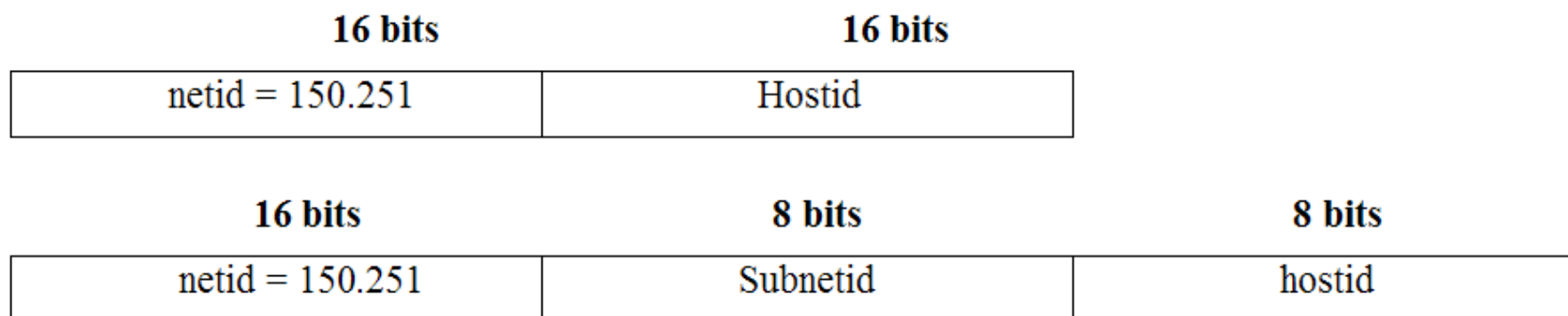
		27 bits
Class E	1 1 1 1 0	(Δεσμευμένα για μελλοντική χρήση)

IP Διευθύνσεις

Κλάση	Εύρος διευθύνσεων
A	0.0.0.0 ως 127.255.255.255
B	128.0.0.0 ως 191.255.255.255
C	192.0.0.0 ως 223.255.255.255
D	224.0.0.0 ως 239.255.255.255
E	240.0.0.0 ως 247.255.255.255

IP Διευθύνσεις

- Διαχείριση Hostid



Subnet Mask

- 32-bit αριθμός
- Ψηφίο "1" για το network ID και το subnet ID
- Ψηφίο "0" για το host ID

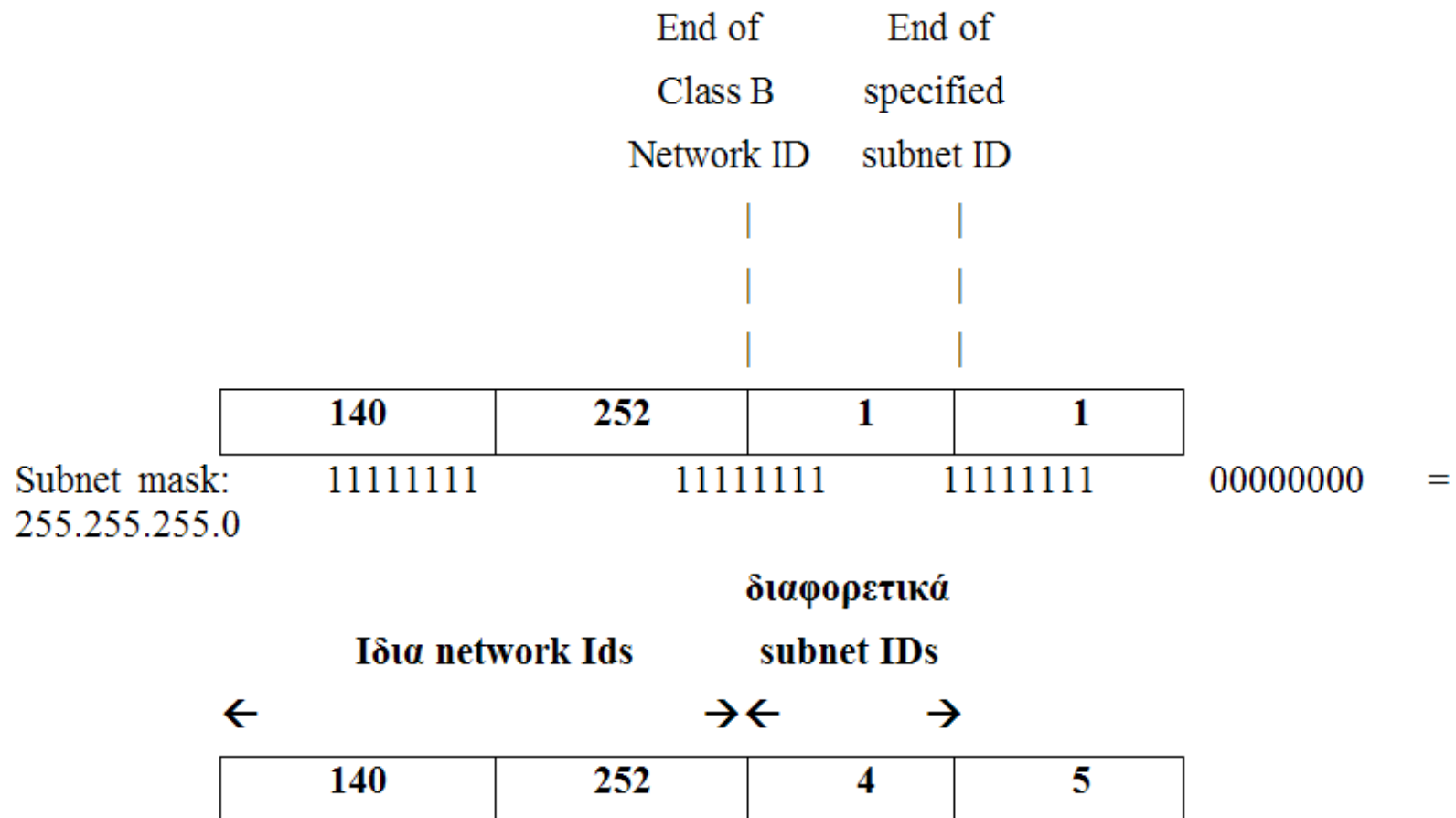
16 bits	8 bits	8 bits
Netid	subnetid	hostid
Subnet Mask: 1111111111111111	11111111	00000000
= 0xffff00 = 255.255.255.0		

16 bits	10 bits	6 bits
Netid	subnetid	hostid
Subnet Mask: 1111111111111111	1111111111	000000
= 0xfffffc0 = 255.255.255.192		

Subnet Mask

- Από την IP και το Subnet Mask, ένας υπολογιστής καθορίζει αν η πληροφορία προορίζεται για υπολογιστή στο:
 - Υποδίκτυό του
 - Διαφορετικό Υποδίκτυο, αλλά ίδιο Δίκτυο
 - Διαφορετικό Δίκτυο
- Ουσιαστικά το χρησιμοποιεί για να αναγνωρίσει το όριο μεταξύ του SubnetID και του HostID

Subnet Mask

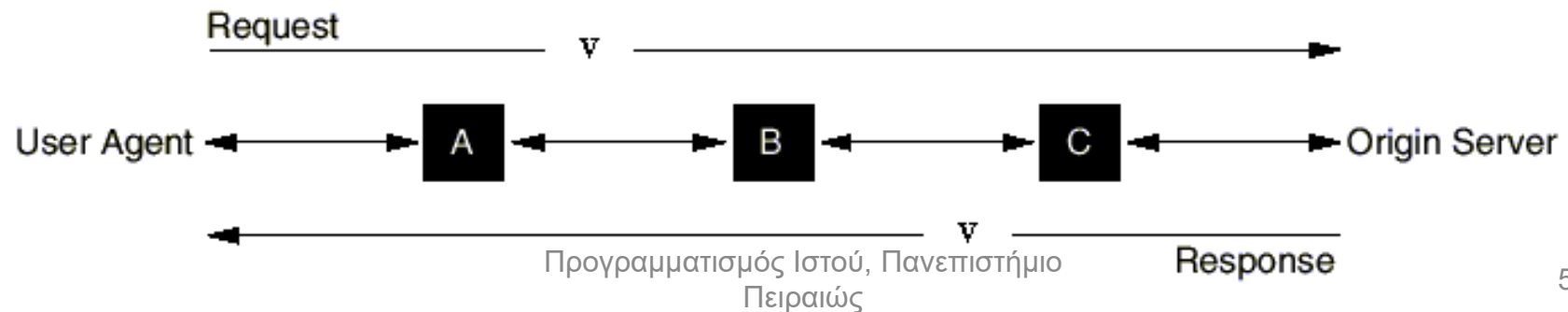
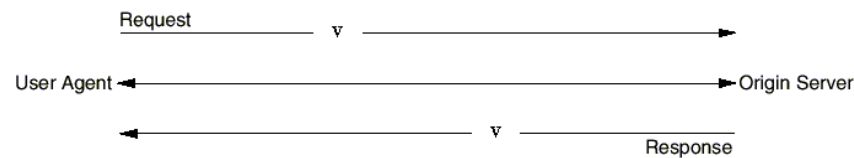


Domain Name Service

- Αντιστοίχιση IP σε ονόματα
- Π.χ.
195.251.225.81 → saas11.cs.unipi.gr
- Μοναδικά, Ιεραρχία
- Διαχείριση Ονομάτων
 - InterNic
 - CENTR
 - GR - HOSTMASTER

HTTP

- Μοντέλο Αίτησης / Απόκρισης
 - Αιτήσεις: Ξεκινούν από τον πελάτη
 - Μπορεί να υπάρχουν ενδιάμεσοι κόμβοι



HTTP

- Uniform Resource Identifier
 - Είναι απλά δομημένες σειρές χαρακτήρων που ορίζουν — είτε μέσω ονόματος, είτε μέσω διεύθυνσης, είτε μέσω άλλων χαρακτηριστικών — ένα πόρο
- HTTP URL
 - Χρησιμοποιείται για τον εντοπισμό πόρων στο http

`http_URL = "http:" "//" host [":" port] [abs_path ["?" query]]`

- Π.χ

`http://gigas.ceid.upatras.gr/activities/index.html`

HTTP

- Τύποι Μηνυμάτων

HTTP-message = Request | Response ; HTTP/1.1 message

- Και οι δύο τύποι μηνύματος αποτελούνται από:
 - Μια αρχική γραμμή
 - Μηδέν οι περισσότερες γραμμές επικεφαλίδων (header lines)
 - Μια κενή γραμμή που αποτελεί την διαχωριστική γραμμή (CRLF)
 - Μια προαιρετική γραμμή που αποτελεί το σώμα του μηνύματος και μπορεί να περιέχει ένα αρχείο, δεδομένα αναζήτησης κ.α. Αυτά είναι και τα πραγματικά δεδομένα.

HTTP

- Μέθοδοι Request

- GET

- GET `http://www.w3.org/pub/www/theproject.html` HTTP/1.1

- HEAD

- POST

- Response

- Η πρώτη γραμμή: Γραμμή κατάστασης → 3 πεδία

- Protocol version , Status-Code, Reason-Phrase

- HTTP/1.1 200 OK

HTTP – Media Types

- Ορίζονται από το RFC1590
- Πρόκειται για επέκταση των MIME Types
- Τα mime types χρησιμοποιούνται από το SMTP, για τον καθορισμό του τύπου δεδομένων που μεταφέρονται
- Υπάρχουν 6 τύποι δεδομένων

- Text

- Audio

- Structured

- Image

- Application

- Message

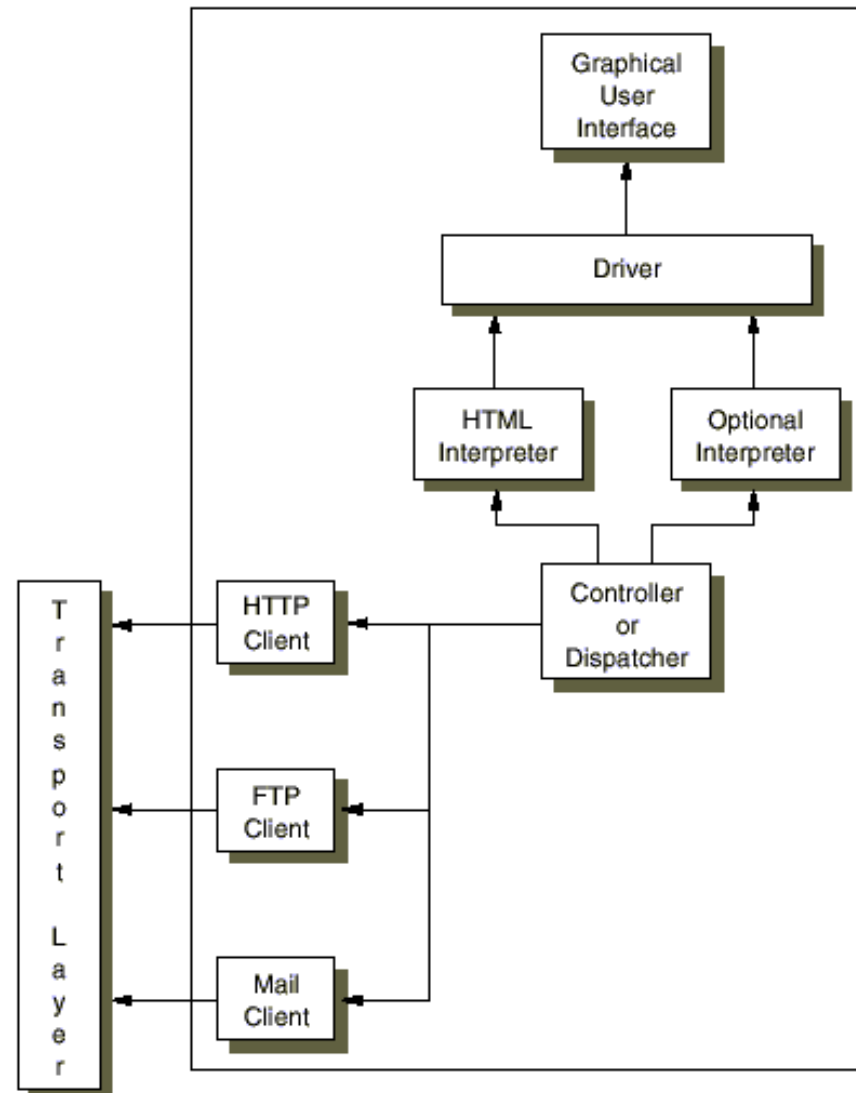
IPv6

- Το υπάρχον IPv4 αποδεικνύεται ανεπαρκές
- Νέο πρωτόκολλο: IPv6
- Μέγεθος διευθύνσεων 128 bit
- Μορφή διευθύνσεων:
x : x : x : x : x : x : x : x (όπου x hex number)
Π.χ.: FEDC:BA98:4798:8798:FEDC:BA98:78D9:3425
- Κατηγορίες Διευθύνσεων
 - Unicast
 - Anycast
 - Multicast

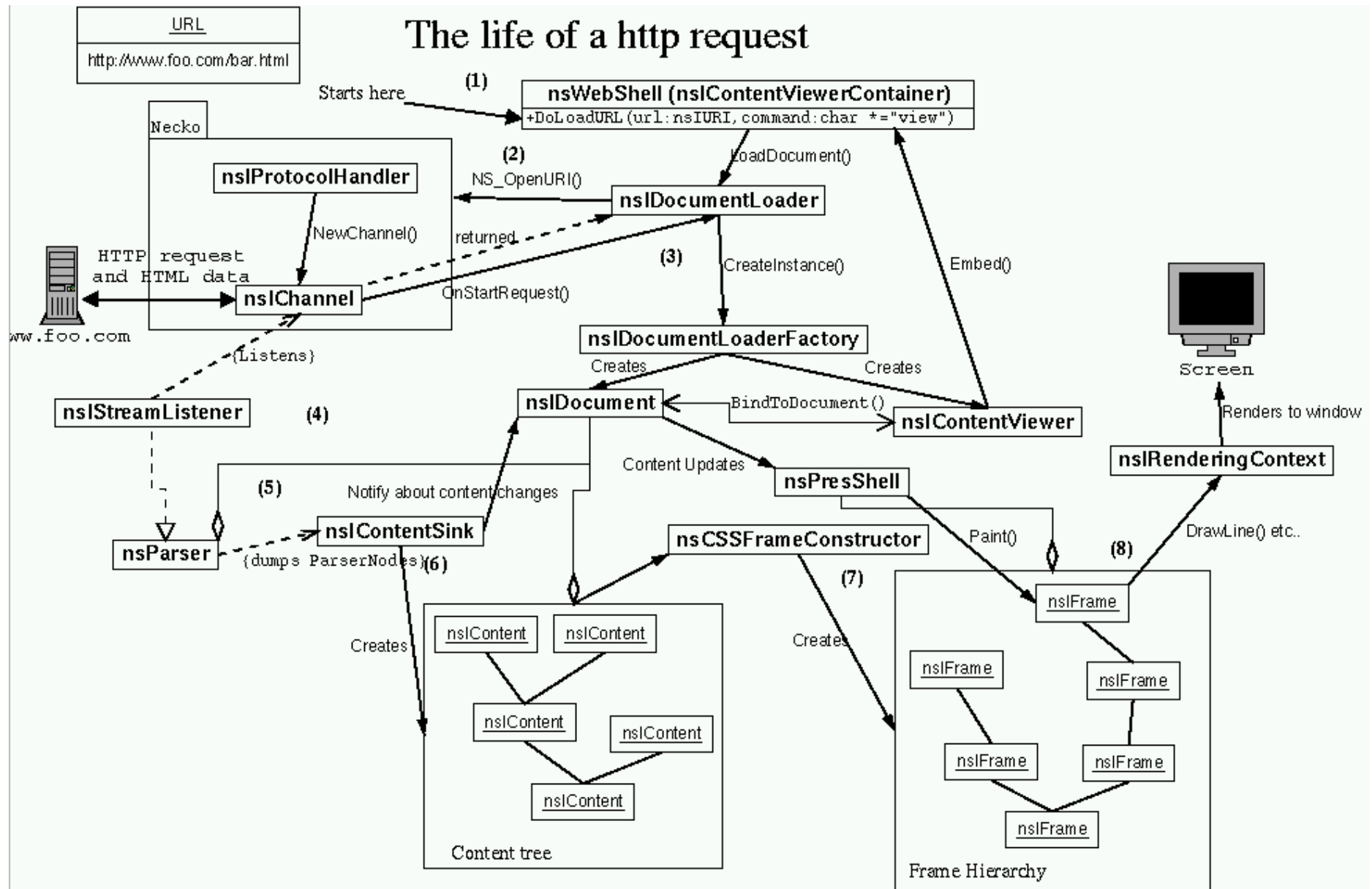
Client – Server Μοντέλο

- Βασικός Τύπος Δικτυακών Εφαρμογών
- Η σχέση Client και Server μπορεί να είναι many-to-many
 - Ένας server εξυπηρετεί ταυτόχρονα πολλούς clients
 - Ένας client μπορεί να συνδεθεί ταυτόχρονα σε πολλούς server
- Web Clients → Web Browsers
 - MS IExplorer, Mozilla, Netscape, Opera
- Web Servers
 - MS IIS, Apache

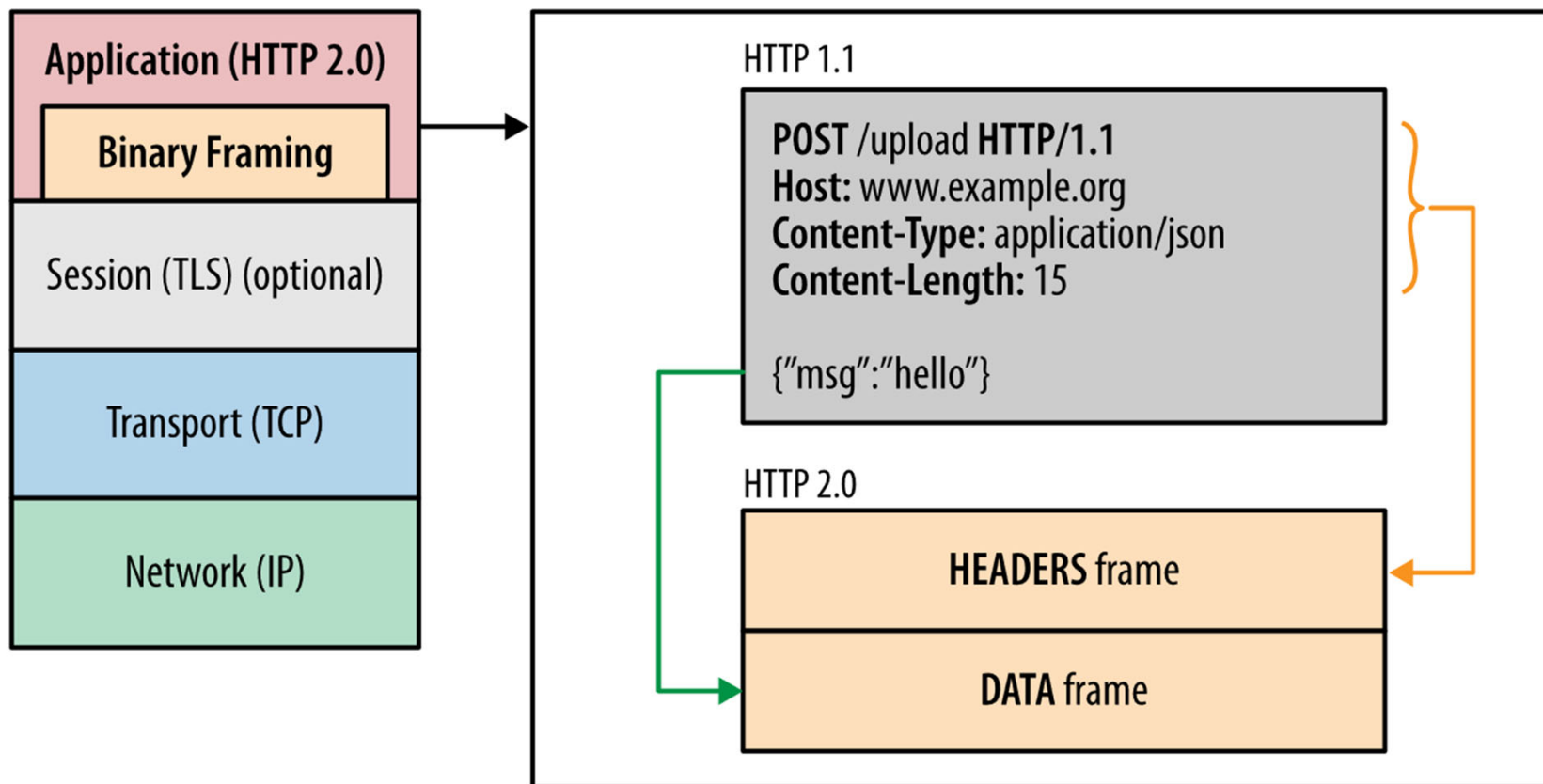
Δομή WebBrowser



The life of a http request



HTTP 2.0



HTTP 2.0

HTTP/2 ορολογία:

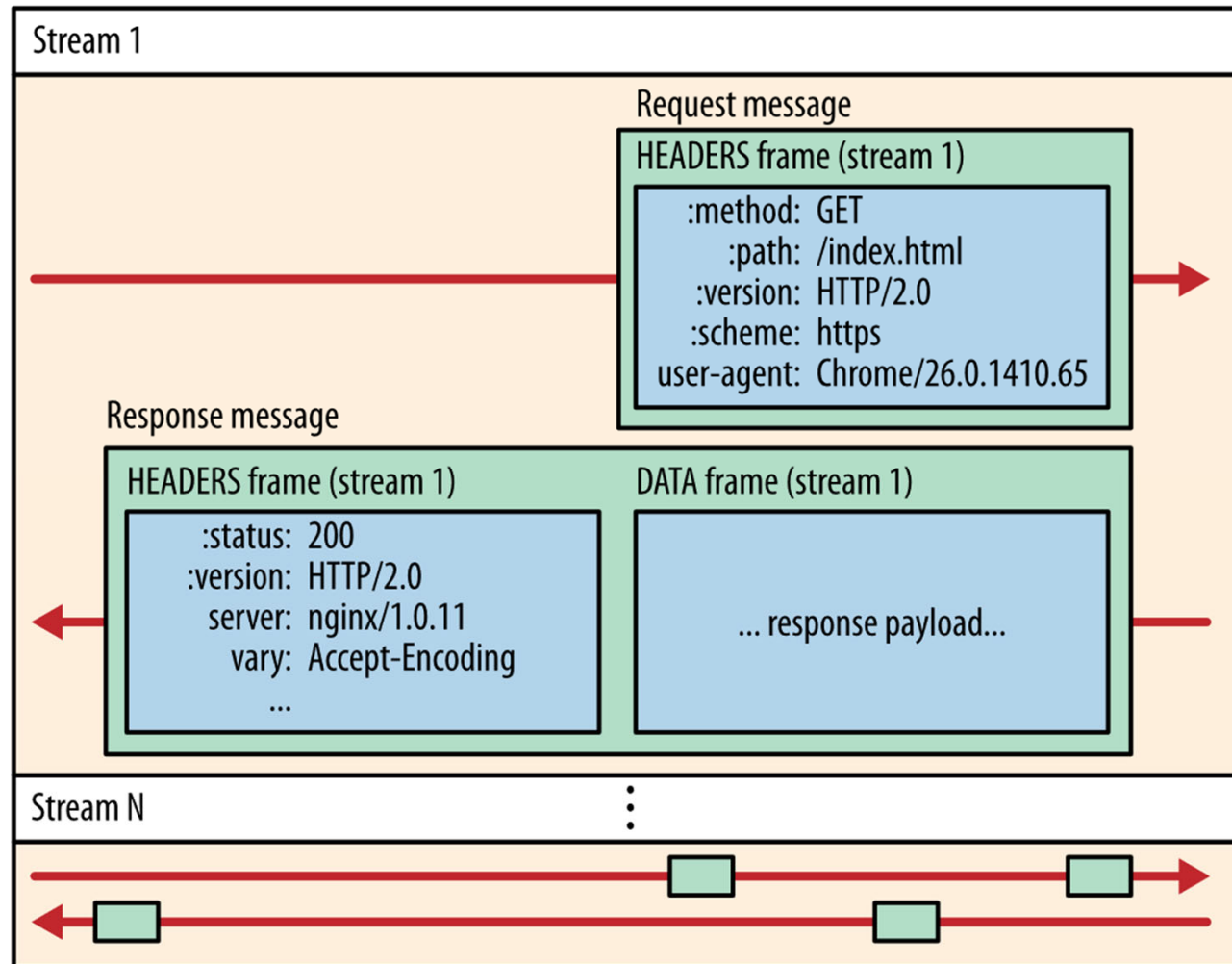
- *Stream*: ροή byte δύο κατευθύνσεων εντός μίας σύνδεσης που μπορεί να μεταφέρει ένα ή περισσότερα messages.
- *Message*: μία αλληλουχία από frames που αντιστοιχεί σε ένα αίτημα ή σε μία απόκριση.
- *Frame*: το πιο μικρό στοιχείο κατά την επικοινωνία του HTTP/2, το οποίο περιλαμβάνει έναν frame header, που κατ' ελάχιστο αντιστοιχεί το stream στο οποίο ανήκει το frame.

HTTP 2

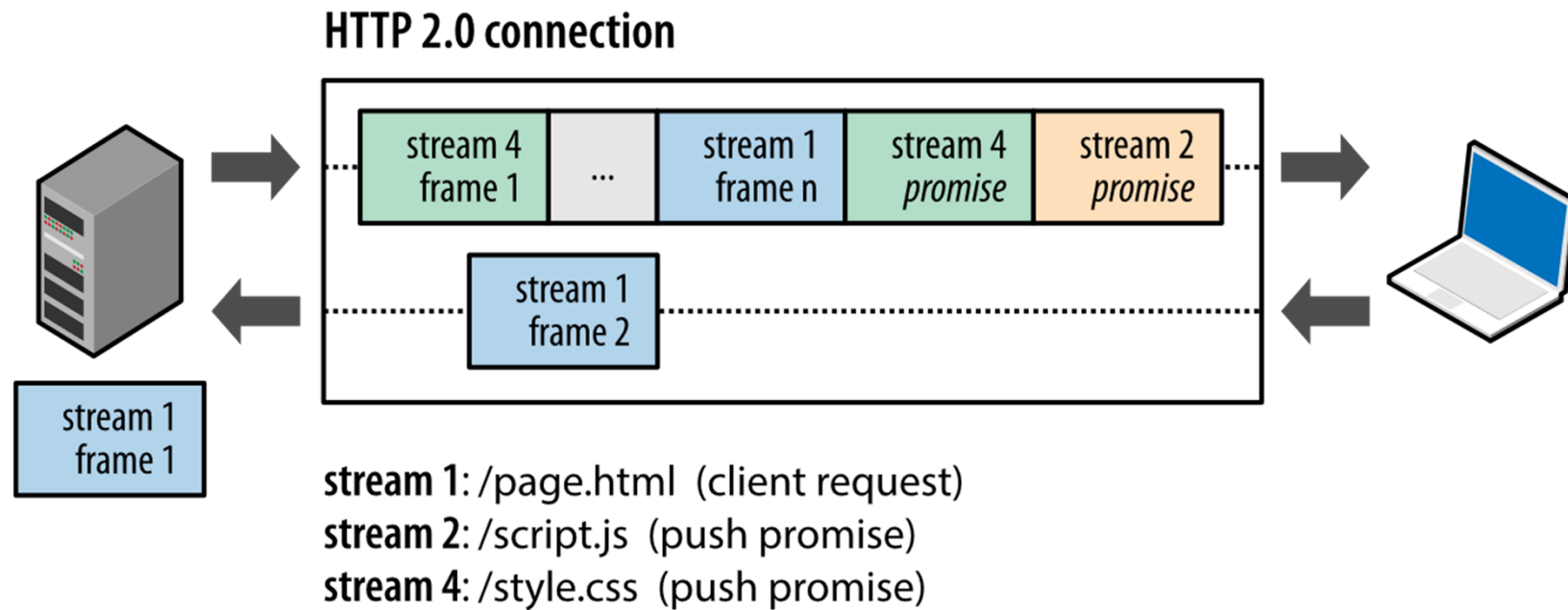
- The relation of these terms can be summarized as follows:
- All communication is performed over a single TCP connection that can carry any number of bidirectional streams.
- Each stream has a unique identifier and optional priority information that is used to carry bidirectional messages.
- Each message is a logical HTTP message, such as a request, or response, which consists of one or more frames.
- The frame is the smallest unit of communication that carries a specific type of data—e.g., HTTP headers, message payload, and so on. Frames from different streams may be interleaved and then reassembled via the embedded stream identifier in the header of each frame.

HTTP 2

Connection



HTTP 2



HTTP 2 PUSH_PROMISE 101

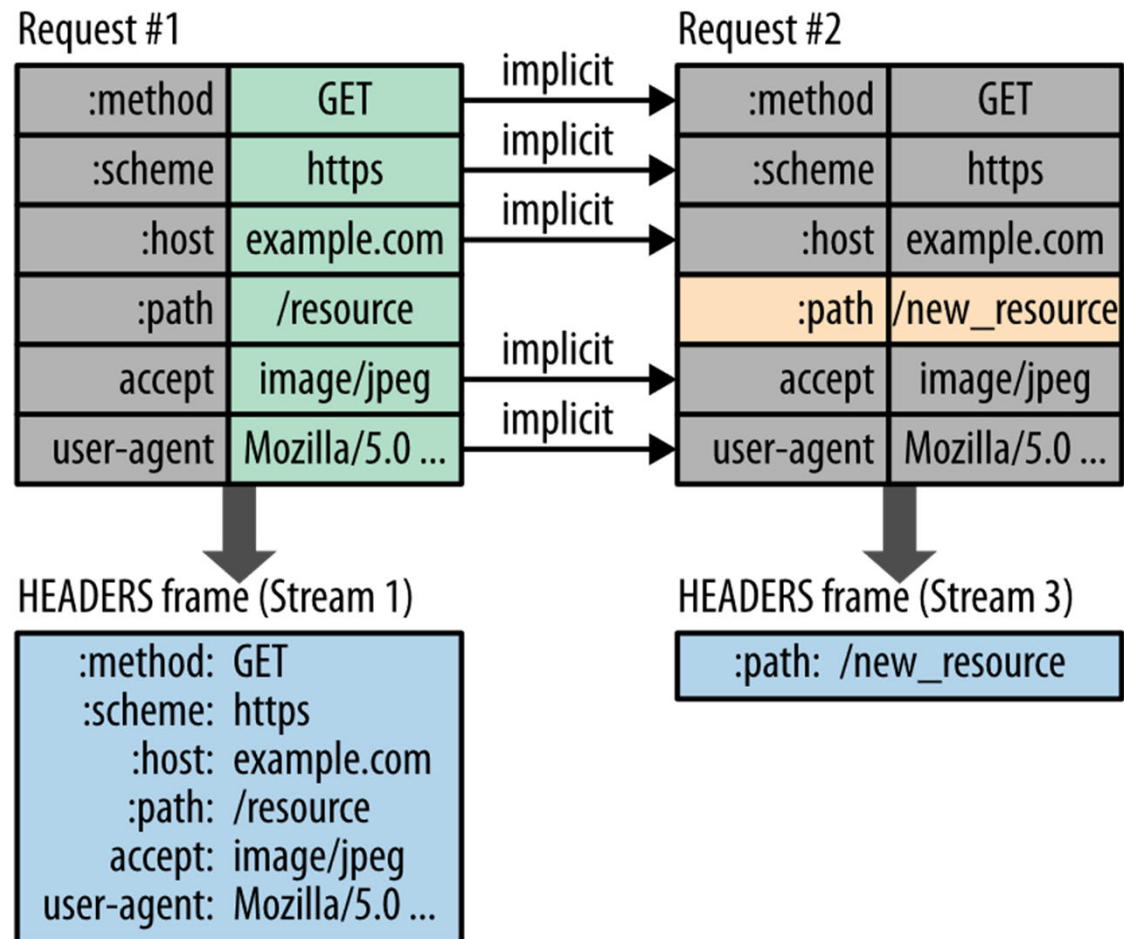
- All server push streams are initiated via PUSH_PROMISE frames
- signal the server's intent to push the described resources to be delivered ahead of the response data that requests the pushed resources.
- critical: the client needs to know which resources the server intends to push to avoid creating duplicate requests for these resources.
- The simplest strategy to satisfy this requirement is to send all PUSH_PROMISE frames, which contain just the *HTTP headers* of the promised resource, ahead of the parent's response (in other words, DATA frames).
- option to decline the stream (via a RST_STREAM frame) if it wants to. (e.g. resource is already in cache.)
- This is an important improvement over HTTP/1.x. By contrast, the use of resource inlining, which is a popular "optimization" for HTTP/1.x, is equivalent to a "forced push": the client cannot opt-out, cancel it, or process the inlined resource individually.

HTTP 2 PUSH_PROMISE 101

- With HTTP/2 the client remains in full control of how server push is used.
- The client can limit the number of concurrently pushed streams
- These preferences are communicated via the SETTINGS frames at the beginning of the HTTP/2 connection
- Each pushed resource is a stream that, unlike an inlined resource, allows it to be individually multiplexed, prioritized, and processed by the client.
- The only security restriction, as enforced by the browser, is that pushed resources must obey the same-origin policy

HTTP 2 Header compression

- In HTTP/1.x, this metadata is always sent as plain text and adds anywhere from 500–800 bytes of overhead per transfer, and sometimes kilobytes more if HTTP cookies are being used.
- HTTP/2 compresses request and response header metadata using the HPACK
- Early versions of HTTP/2 and SPDY used zlib, with a custom dictionary, to compress all HTTP headers. This delivered an 85% to 88% reduction in the size of the transferred header data, and a significant improvement in page load time latency



HTTP 2 Apache

- <https://httpd.apache.org/docs/2.4/howto/http2.html>