CS118 Discussion 1A, Week 3

Yunqi Guo

Boelter Hall 5422, Friday 10:00—11:50 p.m.

Q&A

- **Q:** What's the difference between Expires and Cache-Control headers?

- **A:**

  - Cache-Control was introduced in HTTP/1.1 and offers more options than Expires. They can be used to accomplish the same thing but the data value for Expires is an HTTP date whereas Cache-Control max-age lets you specify a relative amount of time so you could specify "X hours after the page was requested".

  - Expires is recommended for static resources like images and Cache-Control when you need more control over how caching is done.

Cache-control in an HTTP response.

- Cache-Control: must-revalidate
- Cache-Control: no-cache
- Cache-Control: no-store
- Cache-Control: no-transform
- Cache-Control: public
- Cache-Control: private
- Cache-Control: proxy-revalidate
- Cache-Control: max-age=<seconds>
- Cache-Control: s-maxage=<seconds>

Q&A

User-Agent?

GET /118/index.html HTTP/1.1<cr><lf>Host: gaia.cs.umass.edu<cr><lf>User-Agent: Mozilla/5.0 (Windows;U; Windows NT 5.1; en-US; rv:1.7.2) Gecko/20040804 Netscape/7.2 (ax) <cr><lf>Accept:ext/xml, application/xml, application/xhtml+xml, text/html;q=0.9, text/plain;q=0.8,image/png,*/*;q=0.5<cr><lf>Accept-Language: en-us,en;q=0.5<cr><lf>AcceptEncoding: zip,deflate<cr><lf>Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7<cr><lf>Keep-Alive: 300<cr><lf>Connection:keep-alive<cr><lf><cr><lf>

Q&A: User-Agent:

Mozilla/5.0 (platform; rv:geckoversion) Gecko/geckotrail Firefox/firefoxversion

- Mozilla/5.0 is the general token that says the browser is Mozilla compatible, and is common to almost every browser today.
- platform describes the native platform the browser is running on (e.g. Windows, Mac, Linux or Android), and whether or not it's a mobile phone. Firefox OS phones simply say "Mobile"; the web is the platform. Note that platform can consist of multiple "; "-separated tokens. See below for further details and examples.
- rv:geckoversion indicates the release version of Gecko (such as "17.0"). In recent browsers, geckoversion is the same as firefoxversion.
- Gecko/geckotrail indicates that the browser is based on Gecko.
- On Desktop, geckotrail is the fixed string "20100101"
- Firefox/firefoxversion indicates the browser is Firefox, and provides the version (such as "17.0").

Outline

- Application Layer Protocol: DNS, CDN, P2P
- Transport Layer Protocol: UDP
- HW1, HW2 clarification

Application Layer: protocols

- DNS:
  - What is the transport layer protocol?
  - How the scalability is achieved?
  - Who will use iterative/recursive query?
  - Why is DNS resolver needed?

Application Layer: protocols

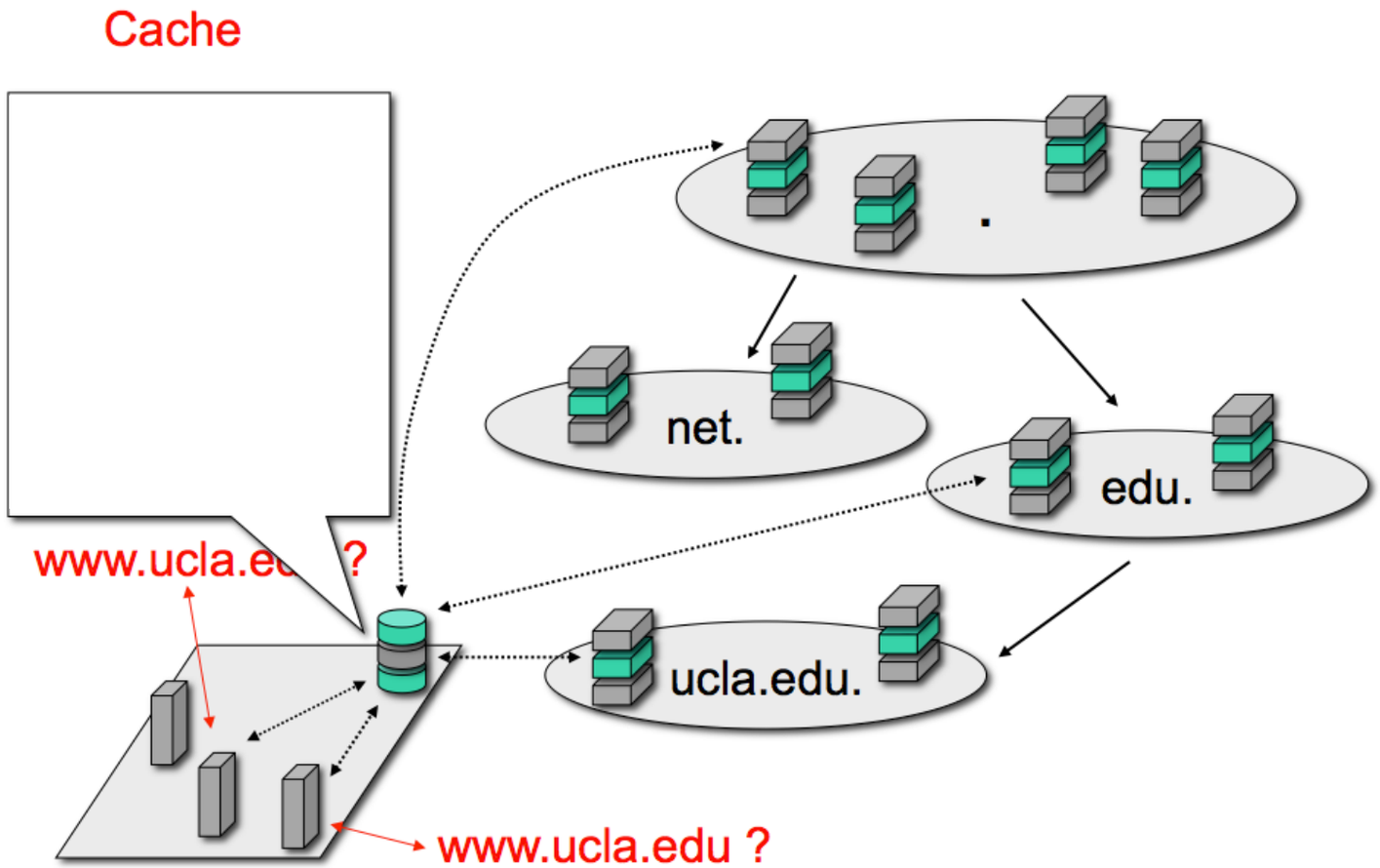
- DNS: convert hostname to IP address (and more)
- A distributed and hierarchical database
    - Root DNS servers
    - Top-level domain (TLD) servers
    - Authoritative DNS servers
    - local DNS server (caching resolver, stub resolver)

DNS protocol: exercise

- Assume the caching resolver's cache is empty initially
- Host A queries www.ucla.edu, how many queries should the caching resolver issue?
- After A's DNS query, host B queries www.mit.edu, how many queries should the caching resolver issue?

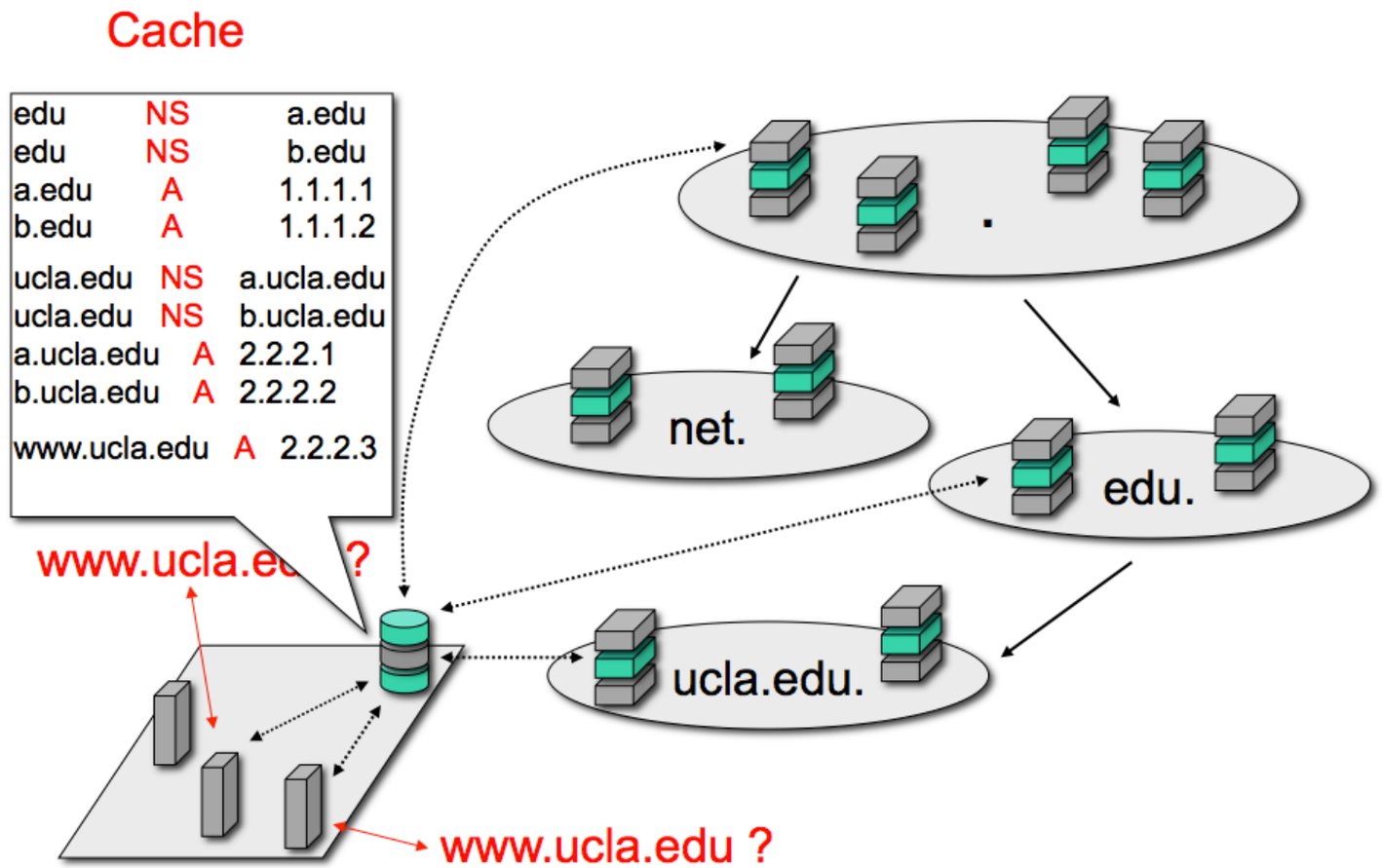
DNS protocol: exercise

- Assume the caching resolver's cache is empty initially
- Host A queries www.uc
  the caching resolver iss
- After A's DNS query, hc
  many queries should th



DNS protocol: exercise

- Assume the caching resolver's cache is empty initially
- Host A queries www.uc
  the caching resolver iss
- After A's DNS query, hc
  many queries should th



A fun experiment: DNS query

```
$ dig google.com
; <<>> DiG 9.8.3-P1 <<>> google.com
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 44777
;; flags: qr rd ra; QUERY: 1, ANSWER: 1, AUTHORITY: 4, ADDITIONAL: 4

;; QUESTION SECTION:
;google.com.            IN  A

;; ANSWER SECTION:
google.com.     76  IN  A   172.217.2.14

;; AUTHORITY SECTION:
google.com.     85950   IN  NS  ns3.google.com.
google.com.     85950   IN  NS  ns4.google.com.
google.com.     85950   IN  NS  ns1.google.com.
google.com.     85950   IN  NS  ns2.google.com.

;; ADDITIONAL SECTION:
ns1.google.com.     59591   IN  A   216.239.32.10
ns2.google.com.     50756   IN  A   216.239.34.10
ns3.google.com.     40354   IN  A   216.239.36.10
ns4.google.com.     36005   IN  A   216.239.38.10

;; Query time: 84 msec
;; SERVER: 158.69.209.100#53(158.69.209.100)
;; WHEN: Thu Jan 19 20:37:48 2017
;; MSG SIZE  rcvd: 180
$ dig any mit.edu
$ dig 206.5.217.172.in-addr.arpa
```

Application Layer: CDN

- CDN: Content Distribution Network
  - Globally distributed network of web servers
  - Stores and replicates images, videos and other files

- https://eclass.uoa.gr/modules/document/file.php/D245/2015/cdn.02f.ppt

Application Layer: protocols

- P2P: no always-on server, peers are intermittently connected

- Calculate content distribution time

Calculate content distribution time

14) Consider distributing a file of $F = 10\ Gbits$ to N peers. The server has an upload rate of $u_s = 20Mbps$, and each peer has a download rate of $d_i = 1\ Mbps$ and an upload rate of u. For $N=10,\ 100$ and 1000 and $u = 200Kbps,\ 600\ Kbps$ and $1\ Mbps$, prepare a chart giving the minimum distribution time for each of the combinations of N and u for both client-server distribution and P2P distribution.

Calculate content distribution time

14) Consider distributing a file of $F = 10\ Gbits$ to N peers. The server has an upload rate of $u_s = 20Mbps$, and each peer has a download rate of $d_i = 1\ Mbps$ and an upload rate of u. For $N=10,\ 100$ and 1000 and $u = 200Kbps,\ 600\ Kbps$ and $1\ Mbps$, prepare a chart giving the minimum distribution time for each of the combinations of N and u for both client-server distribution and P2P distribution.

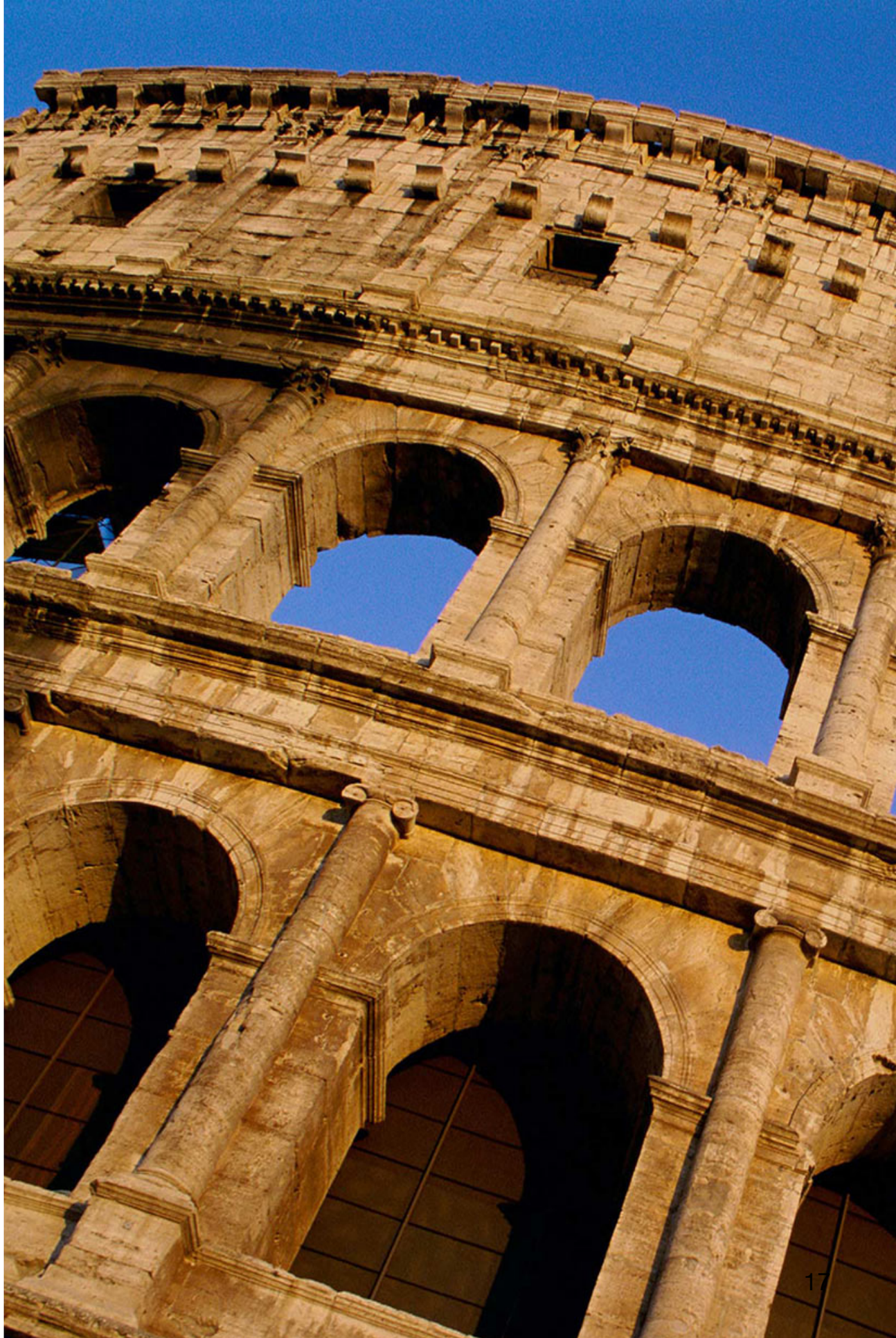
Answer: For calculating the minimum distribution time for client-server distribution, we use the following formula:

$$D_{cs} = max\ \{NF/u_s,\ F/d_{min}\}$$

Similarly, for calculating the minimum distribution time for P2P distribution, we use the following formula:

$$D_{P2P} = max\{F/u_s,\ F/d_{min},\ NF/(u_s + \sum_{i=1}^{N} u_i)\}$$

Transport Layer



Transport Layer V.S. Network Layer

- Network layer: logical communication between **hosts**
  - **IP address** is used for identifying a host
- Transport layer: logical communication between **processes**
  - **IP address and port number** are used for identifying a process

Multiplexing and De-multiplexing

- Multiplexing at send host: gather data from multiple sockets

- De-multiplexing at receiving host: deliver received segments to the right socket

- **Five tuples** (src_ip, src_port, dst_ip, dst_port, protocol) are used for multiplexing/demultiplexing

  - How to identify a TCP/UDP socket?

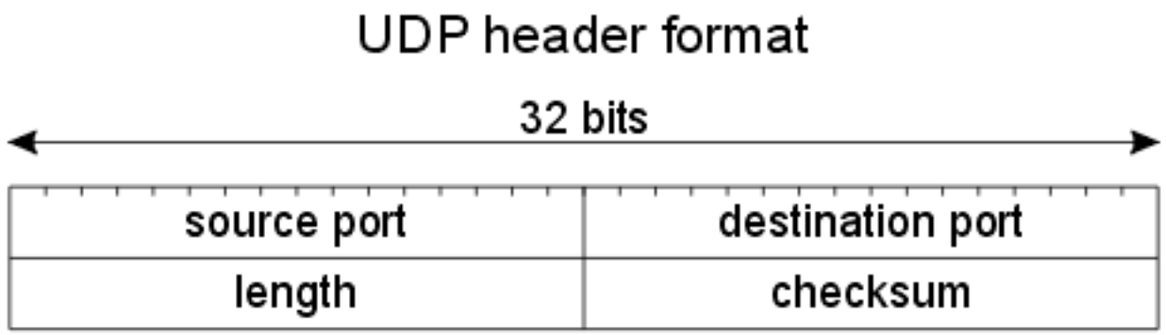
  - Can TCP and UDP share the same port numbers?

Multiplexing and De-multiplexing

- Multiplexing at send host: gather data from multiple sockets
- De-multiplexing at receiving host: deliver received segments to the right socket
- **Five tuples** (src_ip, src_port, dst_ip, dst_port, protocol) are used for multiplexing/demultiplexing
  - How to identify a TCP/UDP socket? **lsof -i**
  - Can TCP and UDP share the same port numbers? **Yes! e.g. DNS**

UDP

- No connection establishment
- No connection state
- Small packet overhead (8 byte,



- How to calculate checksum?
  - **Pseudo header** + **UDP header** + data
  - Why pseudo header?