

Trường Đại Học Bách Khoa Tp.HCM
Hệ Đào Tạo Từ Xa
Khoa Khoa Học và Kỹ Thuật Máy Tính

Mạng máy tính căn bản

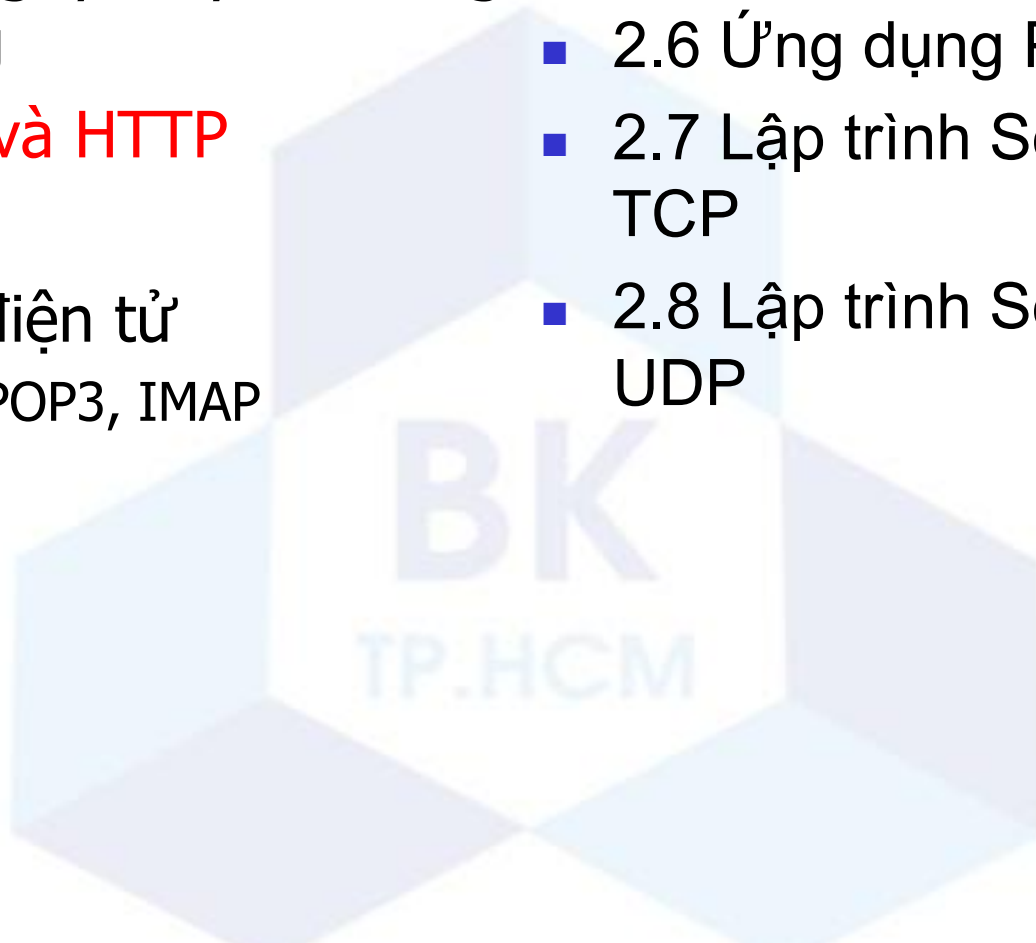
Bài giảng 4: Tầng ứng dụng (tt)

Tham khảo:

Chương 2: “Computer Networking – A top-down approach”
Kurose & Ross, 5th ed., Addison Wesley, 2010.

Chương 2: Tầng ứng dụng

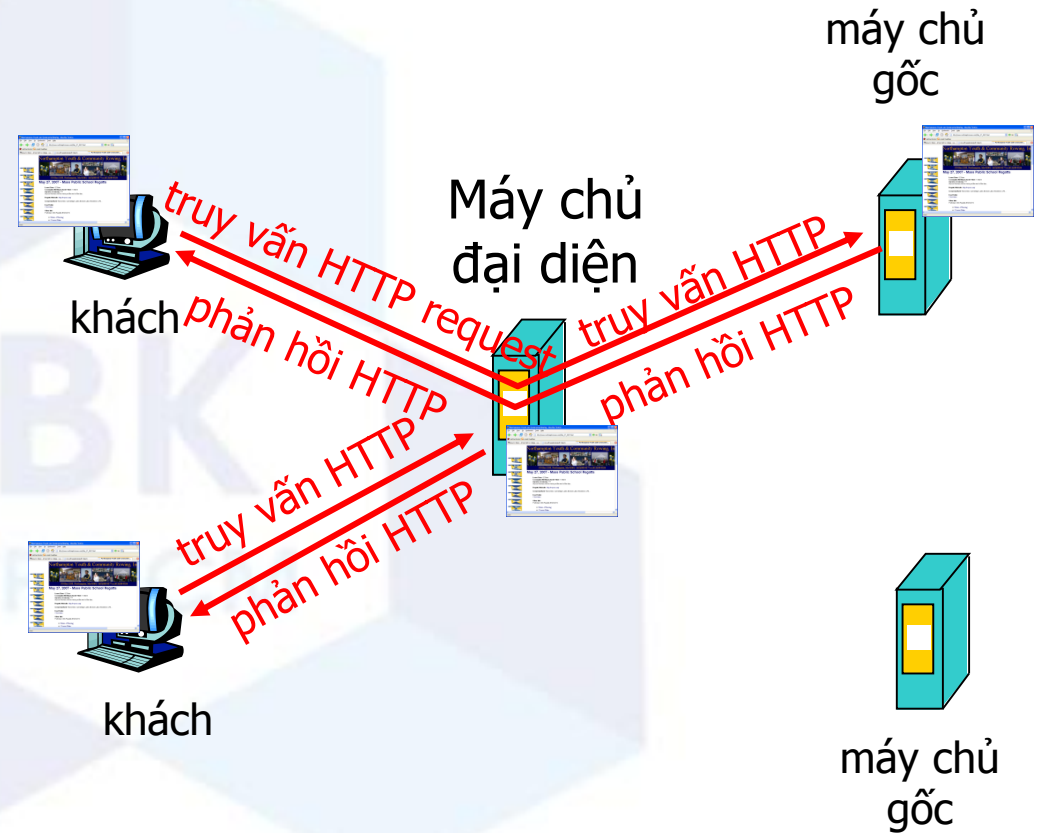
- 2.1 Các nguyên lý của tầng ứng dụng
- 2.2 Web và HTTP
- 2.3 FTP
- 2.4 Thư điện tử
 - SMTP, POP3, IMAP
- 2.5 DNS
- 2.6 Ứng dụng P2P
- 2.7 Lập trình Socket với TCP
- 2.8 Lập trình Socket với UDP



Bộ nhớ đệm Web (Web cache) và máy chủ đại diện (proxy server)

Mục đích: đáp ứng yêu cầu của n/d mà không cần máy chủ gốc

- thiết lập trình duyệt: truy cập web qua bộ nhớ đệm
- trình duyệt gửi toàn bộ truy vấn HTTP tới máy chủ đại diện
 - đối tượng có trong bnd: bnd gửi trả đối tượng
 - ngoài ra, bnd sẽ truy vấn đối tượng từ máy chủ gốc, sau đó gửi lại cho người dùng, đồng thời lưu lại trong bnd



Bộ nhớ đệm Web (tt)

- BND làm việc vừa như khách vừa như chủ
- thường thì BND thường được cài đặt bởi nccdv địa phương hoặc cục bộ (trường đại học, cơ quan)

Tại sao cần có BND Web?

- giảm thời gian phản hồi cho truy vấn n/dùng
- giảm lưu lượng trong một đường kết nối của công sở.
- cho phép các nccdv tăng thêm nội dung cung cấp cho người dùng

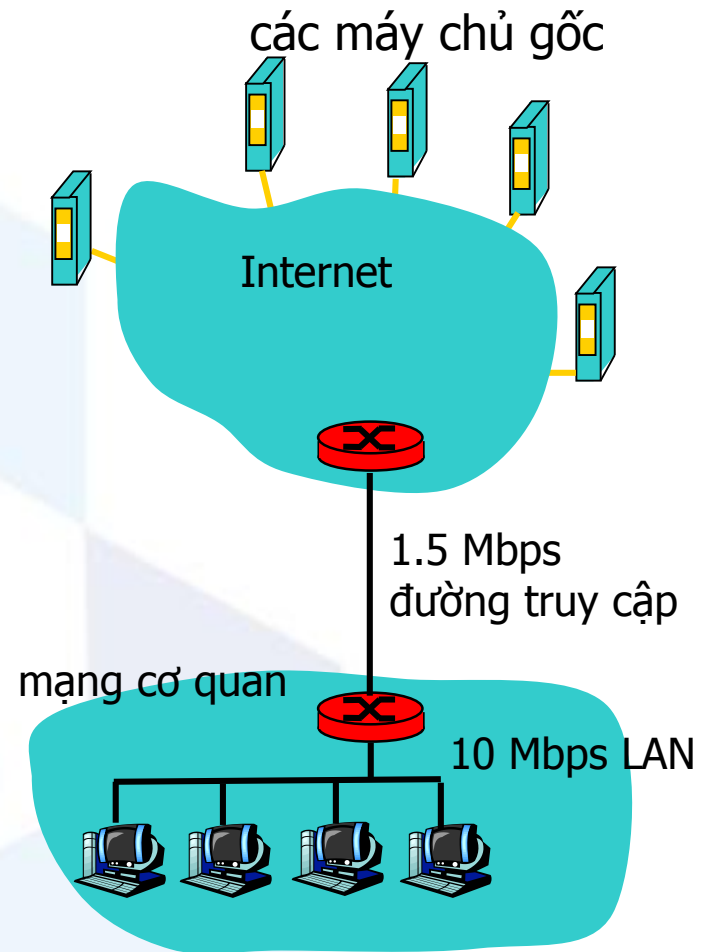
Ví dụ hiện thực BNĐ

Giả thiết

- kích thước trung bình của đối tượng = 100,000 bit
- tần số t/bình của truy vấn từ trình duyệt của cơ quan tới máy chủ gốc = 15 lần/s
- độ trễ xoay vòng từ bất cơ quan tới bất kì máy chủ gốc nào = 2 s

Kết quả

- hiệu suất sử dụng LAN = 15%
- hiệu suất sử dụng trên đường truy cập = 100%
- độ trễ tổng = độ trễ Internet + độ trễ truy cập + độ trễ LAN
= 2s + vài phút + vài ms



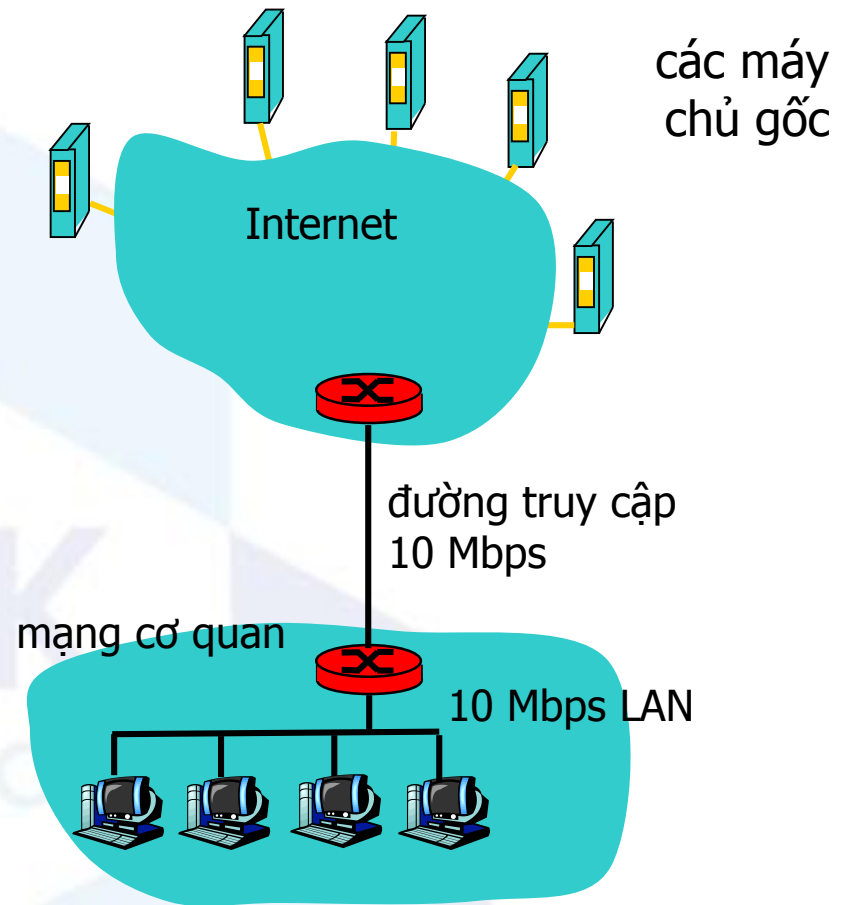
Ví dụ hiện thực BNĐ (tt)

những giải pháp khả dĩ

- tăng băng thông của đường truy cập lên khoảng 10 Mbps

kết quả

- hiệu suất sử dụng LAN = 15%
- hiệu suất sử dụng đường truy cập = 15%
- độ trễ tổng = độ trễ Internet + độ trễ truy cập + độ trễ LAN
- = 2s + ms + ms
- rất tốn kém khi nâng cấp



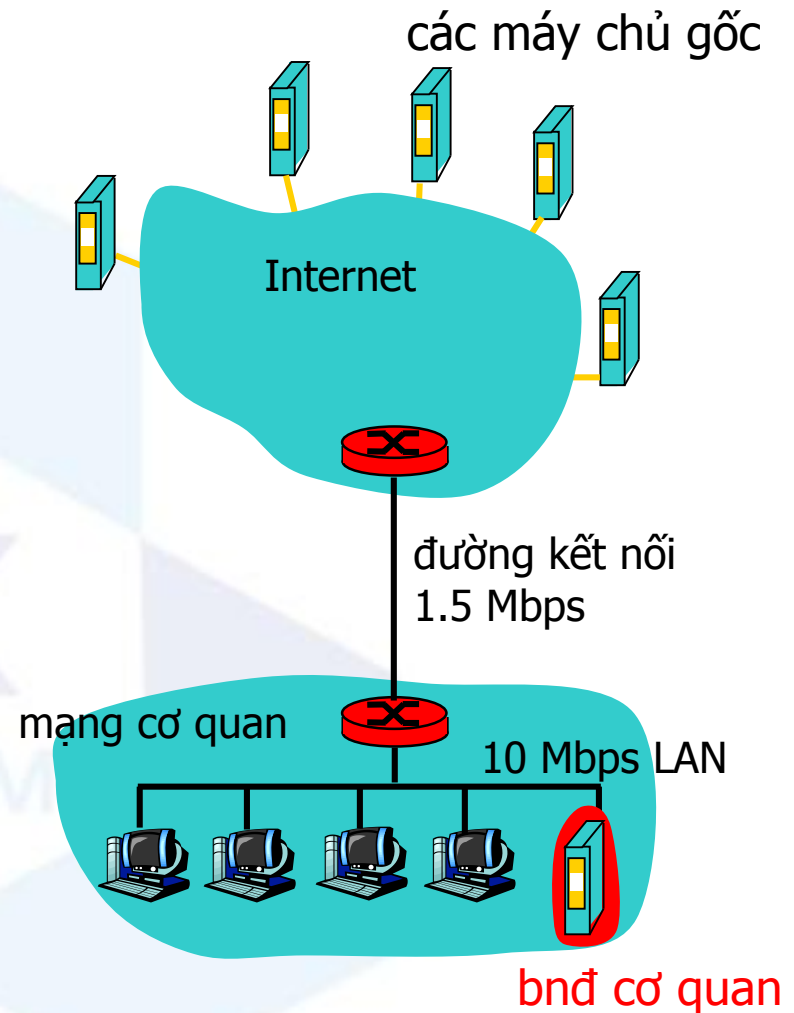
Ví dụ hiện thực BNĐ (tt)

giải pháp: cài đặt BNĐ

- giả sử tần số truy cập BNĐ là 0.4

kết quả:

- 40% truy vấn sẽ được đáp ứng ngay lập tức bởi BNĐ
- 60% truy vấn sẽ được đáp ứng bởi máy chủ gốc
- hiệu suất sử dụng đường kết nối giảm xuống còn 60%, kéo theo độ trễ không đáng kể (khoảng 10 msec)
- tổng độ trễ tối đa = độ trễ Internet + độ trễ truy cập + độ trễ LAN = $.6 * (2.01) \text{ secs} + \text{msec} < 1.4 \text{ secs}$



GET có điều kiện

Mục đích: không gửi đối tượng nếu như BND đã có bản cập nhật mới nhất của nó

- BND: chỉ rõ thời gian của bản lưu BND trong truy vấn HTTP

If-modified-since:
<date>

- máy chủ: sẽ không gửi lại đối tượng được yêu cầu nếu bản lưu là bản cập nhật mới nhất:

HTTP/1.0 304 Not
Modified

máy chủ

HTTP request msg
If-modified-since:
<date>

đối tượng
không
thay đổi

HTTP response
HTTP/1.0
304 Not Modified

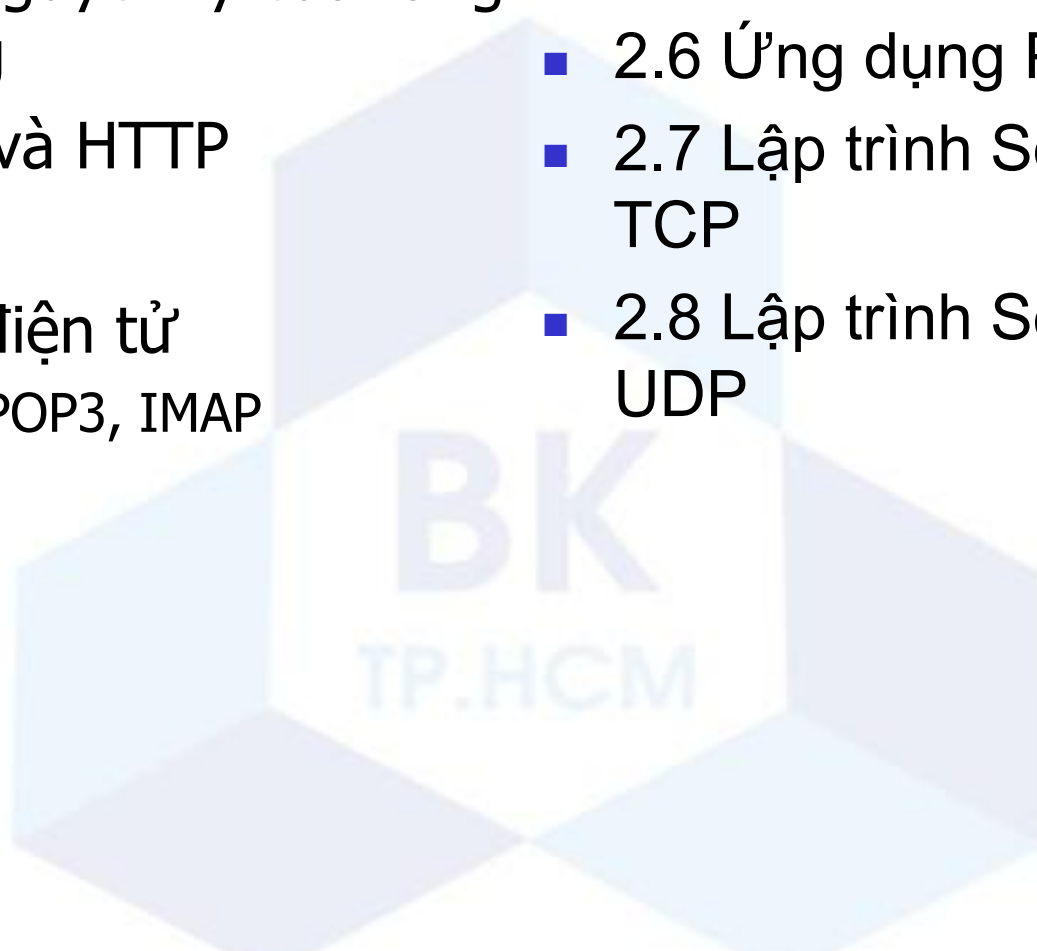
HTTP request msg
If-modified-since:
<date>

đối tượng
đã
thay đổi

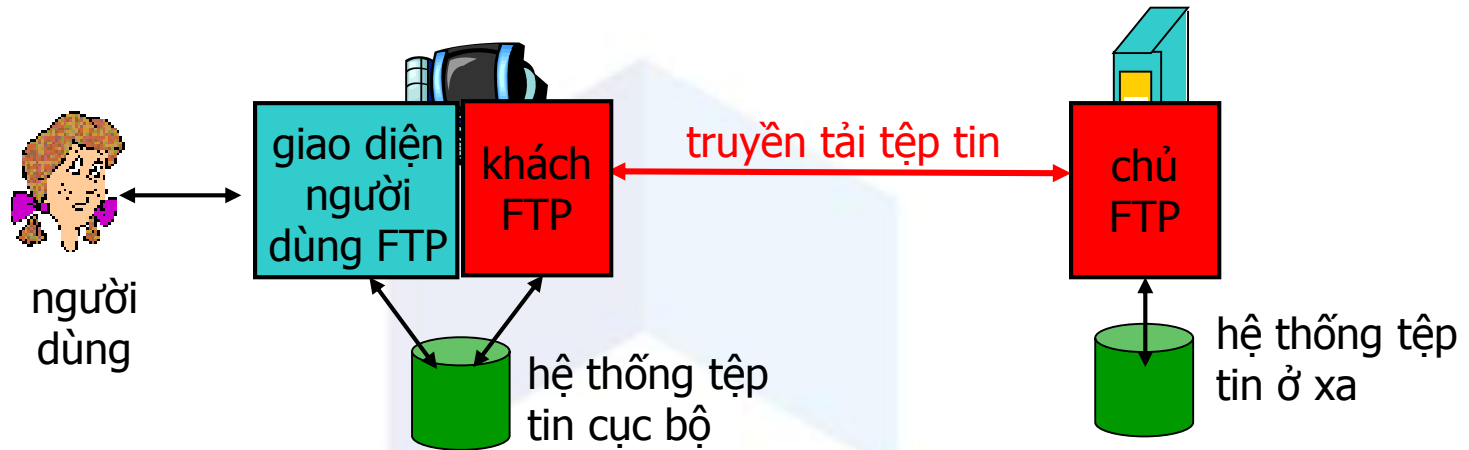
HTTP response
HTTP/1.0 200 OK
<data>

Chương 2: Tầng ứng dụng

- 2.1 Các nguyên lý của tầng ứng dụng
- 2.2 Web và HTTP
- 2.3 FTP
- 2.4 Thư điện tử
 - SMTP, POP3, IMAP
- 2.5 DNS
- 2.6 Ứng dụng P2P
- 2.7 Lập trình Socket với TCP
- 2.8 Lập trình Socket với UDP



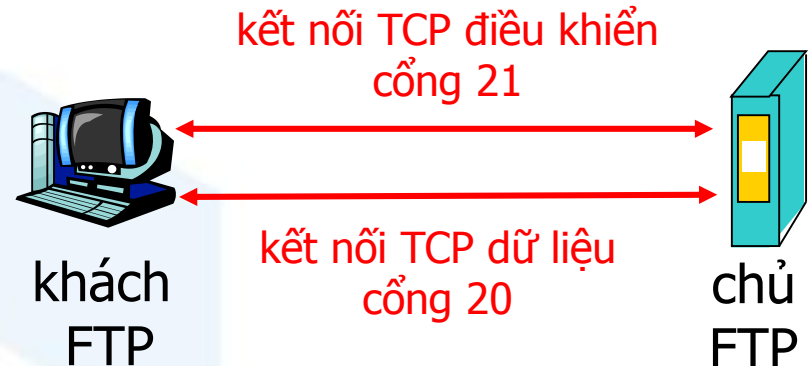
FTP: Giao thức truyền tải tệp tin



- truyền tải tệp tin đến/từ máy ở xa
- mô hình khách/chủ
 - **khách**: là phía bắt đầu quá trình truyền tải
 - **chủ**: máy ở xa lắng nghe kết nối
- ftp: RFC 959
- máy chủ ftp: cổng 21

FTP: điều khiển riêng biệt, kết nối dữ liệu

- Khách FTP kết nối tới máy chủ FTP tại cổng 21, giao thức truyền tải là TCP
- khách được kiểm tra danh tính thông qua *kết nối điều khiển*
- khách duyệt các thư mục trên máy từ xa bằng cách gửi các câu lệnh thông qua *kết nối điều khiển*.
- khi chủ nhận được câu lệnh truyền tải tệp, chủ mở *kết nối TCP thứ 2* tới khách
- sau khi truyền tải xong 1 tệp, chủ đóng *kết nối dữ liệu*.
- chủ mở một kết nối TCP dữ liệu khác để truyền tải một tệp khác.
- điều khiển kết nối: “ngoại tuyến” (out of band)
- máy chủ FTP lưu lại “trạng thái”: thư mục hiện tại, thông tin về danh tính



Các câu lệnh và phản hồi FTP

Ví dụ câu lệnh:

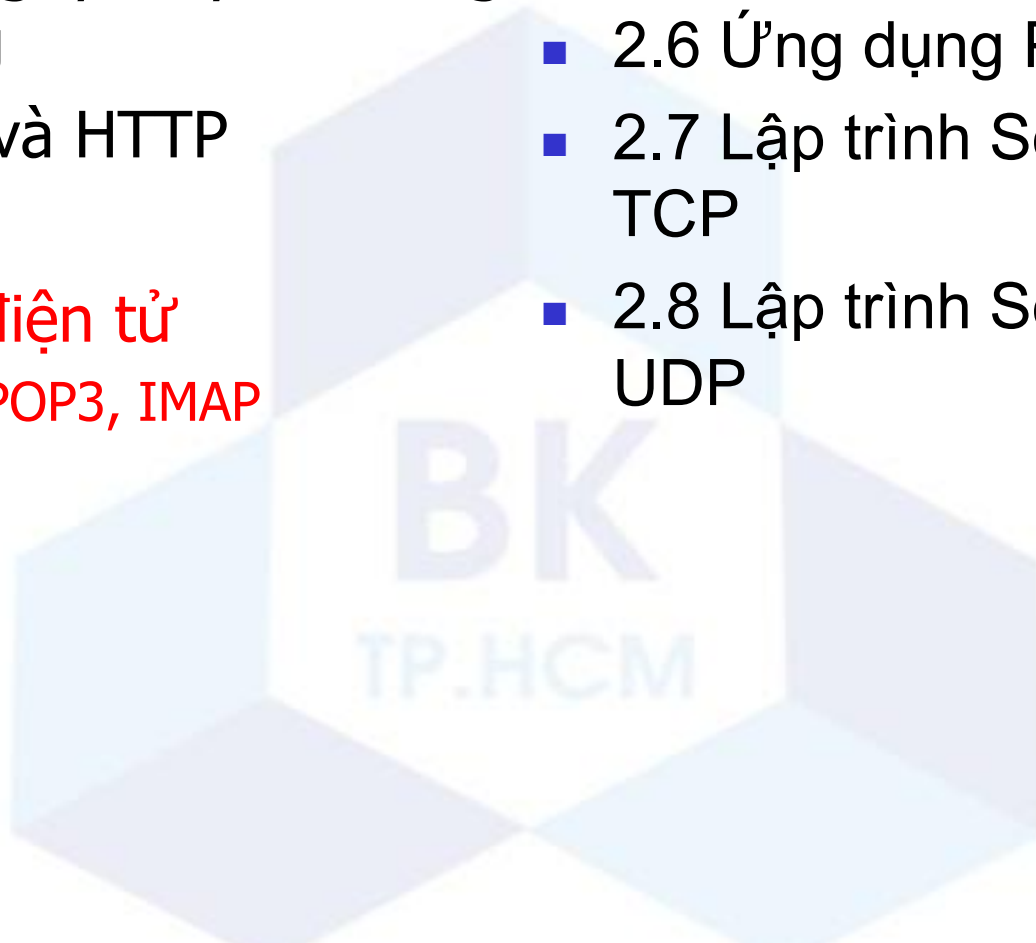
- gửi đi dưới dạng văn bản ASCII qua kết nối điều khiển
- `USER tên_người_dùng`
- `PASS mật_khẩu`
- `LIST` liệt kê danh sách các tệp trong thư mục hiện tại
- `RETR tên_tệp` tải tệp tin về
- `STOR tên_tệp` tải tệp tin lên máy từ xa

Ví dụ mã trả về

- mã trạng thái và mô tả (như trong HTTP)
- `331 Username OK, password required`
- `125 data connection already open; transfer starting`
- `425 Can't open data connection`
- `452 Error writing file`

Chương 2: Tầng ứng dụng

- 2.1 Các nguyên lý của tầng ứng dụng
- 2.2 Web và HTTP
- 2.3 FTP
- 2.4 Thư điện tử
 - SMTP, POP3, IMAP
- 2.5 DNS
- 2.6 Ứng dụng P2P
- 2.7 Lập trình Socket với TCP
- 2.8 Lập trình Socket với UDP



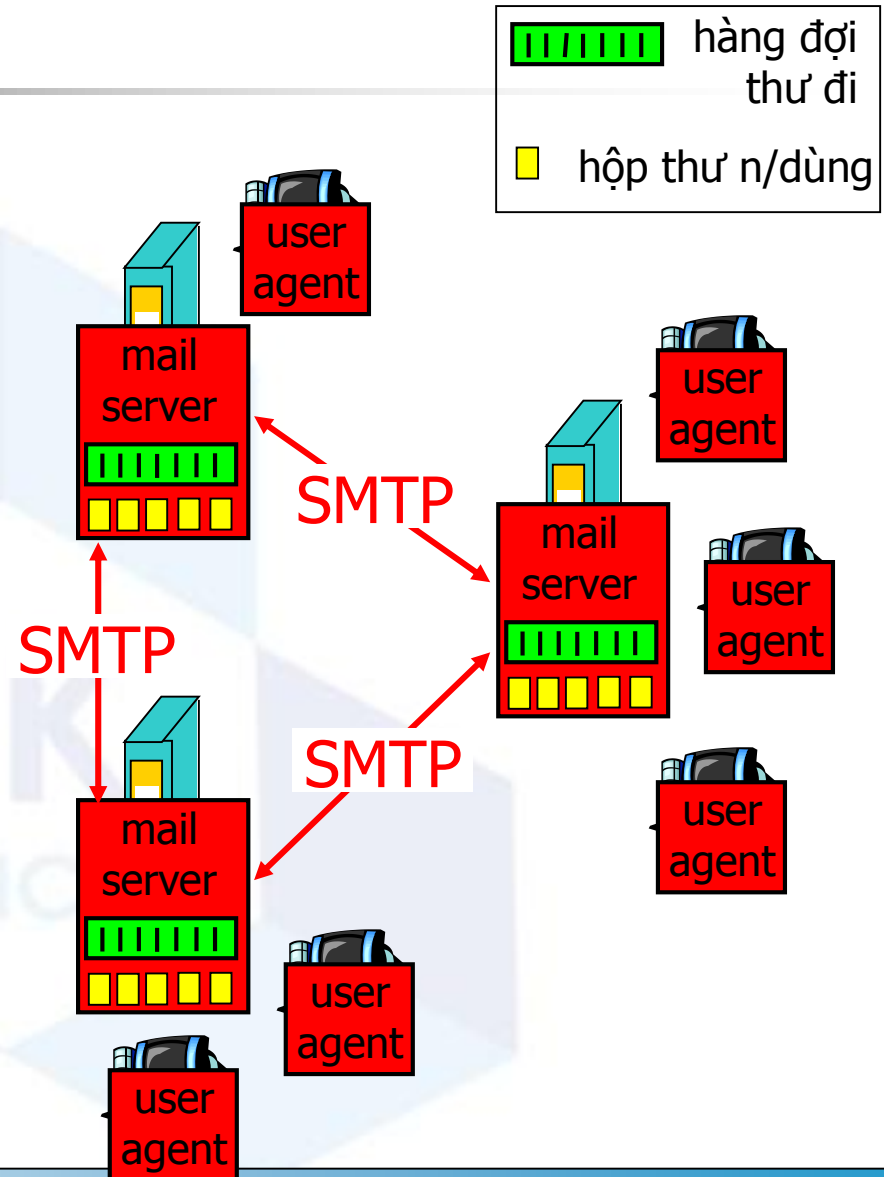
Thư điện tử (e-mail)

Ba thành phần chính:

- công cụ quản lý thư người dùng (user agent - UA)
- máy chủ thư
- giao thức truyền tải thư: SMTP

Công cụ quản lý thư của người dùng

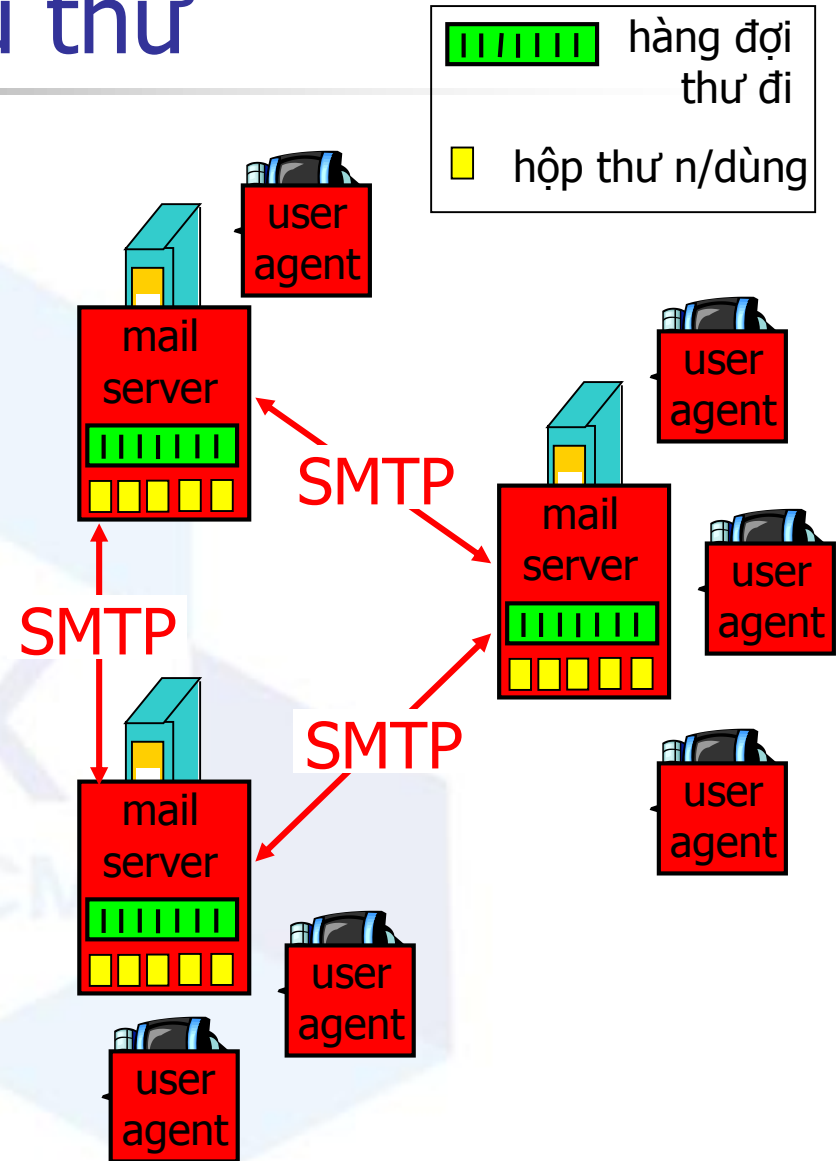
- soạn thảo, chỉnh sửa, lưu, v.v...
- vd: Outlook, Mozilla Thunderbird
- thư đi và thư đến được lưu trên máy chủ



Thư điện tử: máy chủ thư

Máy chủ thư

- **hộp thư** chứa thư đến của người dùng
- **hàng đợi thông điệp** của thư đi
- **giao thức SMTP** giữa các máy chủ trao đổi thư cho nhau
 - khách: là máy chủ gửi thư
 - “chủ”: là máy chủ nhận thư

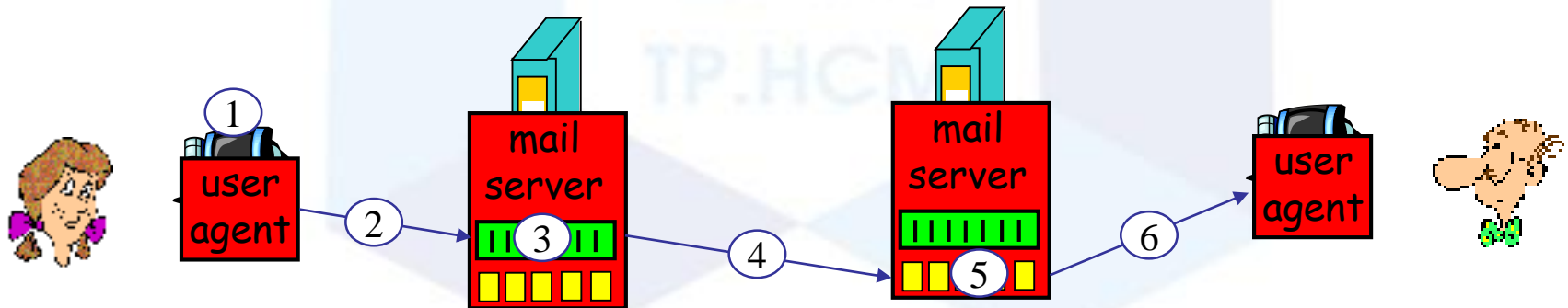


Thư điện tử: SMTP [RFC 2821]

- sử dụng TCP để gửi thư giữa khách và chủ, cổng 25
- vận chuyển trực tiếp: từ máy chủ gửi và máy chủ nhận
- quá trình vận chuyển gồm 3 pha:
 - bắt tay (tạo kết nối)
 - trao đổi thông điệp
 - kết thúc
- tương tác câu lệnh/phản hồi
 - **câu lệnh:** văn bản ASCII
 - **phản hồi:** mã trạng thái và mô tả
- thông điệp phải có định dạng 7-bit ASCII

Kịch bản: Alice gửi thông điệp Bob

- 1) Alice sử dụng UA để soạn thư và gửi tới `bob@someschool.edu`
- 2) UA của Alice gửi thông điệp tới máy chủ thư của cô ta; thư được đặt trong hàng đợi
- 3) phía khách của SMTP mở kết nối TCP tới máy chủ thư của Bob
- 4) khách SMTP gửi thư của Alice qua kết nối TCP
- 5) máy chủ thư của Bob đặt thư vào hộp thư của Bob
- 6) Bob bật UA của mình lên để đọc thư



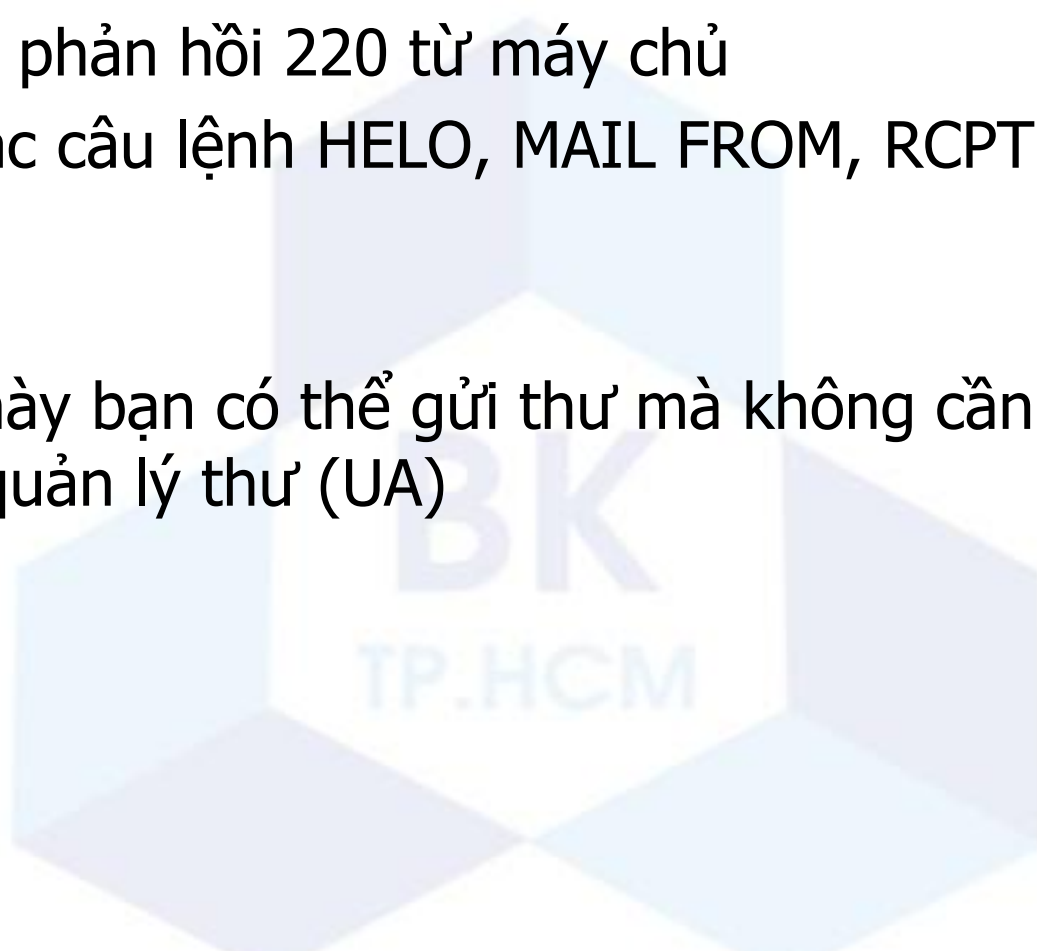
Ví dụ tương tác SMTP

```
S: 220 hamburger.edu
C: HELO crepes.fr
S: 250 Hello crepes.fr, pleased to meet you
C: MAIL FROM: <alice@crepes.fr>
S: 250 alice@crepes.fr... Sender ok
C: RCPT TO: <bob@hamburger.edu>
S: 250 bob@hamburger.edu ... Recipient ok
C: DATA
S: 354 Enter mail, end with "." on a line by itself
C: Do you like ketchup?
C: How about pickles?
C: .
S: 250 Message accepted for delivery
C: QUIT
S: 221 hamburger.edu closing connection
```

Thử nghiệm tương tác SMTP:

- `telnet tên_máy_chủ_thư 25`
- nhìn thấy phản hồi 220 từ máy chủ
- gõ vào các câu lệnh HELO, MAIL FROM, RCPT TO, DATA, QUIT

bằng cách này bạn có thể gửi thư mà không cần sử dụng công cụ quản lý thư (UA)



SMTP: tổng kết

- SMTP sử dụng kết nối ổn định
- SMTP đòi hỏi thông điệp (mào đầu và thân) phải dùng kí tự ASCII 7-bit
- máy chủ SMTP sử dụng `CRLF.CRLF` để xác định điểm kết thúc của thông điệp

So sánh với HTTP:

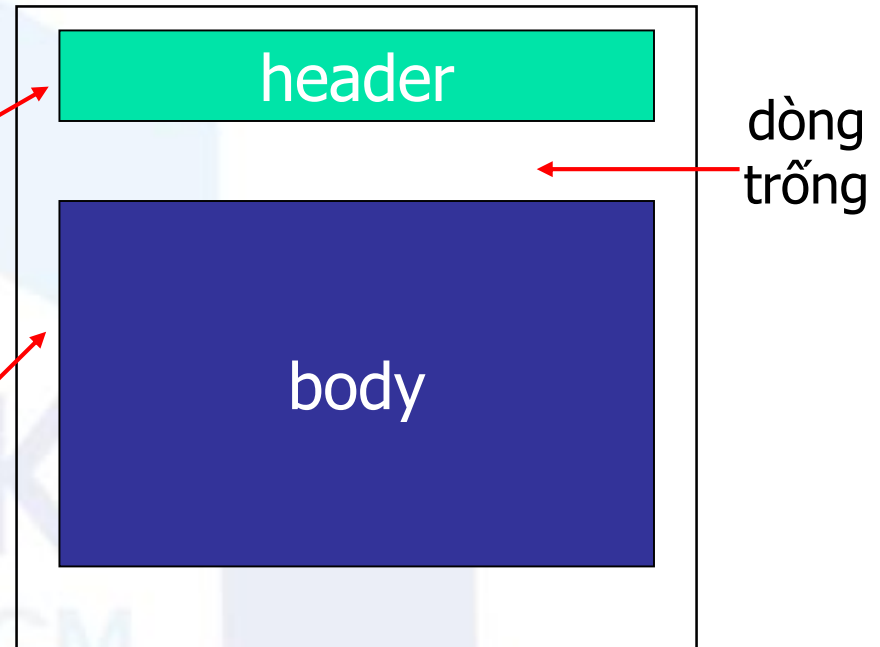
- HTTP: tải về
- SMTP: đẩy đi
- cả hai đều có tương tác câu lệnh/phản hồi dưới dạng ASCII, các mã trạng thái
- HTTP: mỗi đối tượng được đóng gói trong thông điệp phản hồi của riêng nó
- SMTP: nhiều đối tượng được gửi chung trong nhiều phần của một thông điệp

Định dạng thông điệp thư

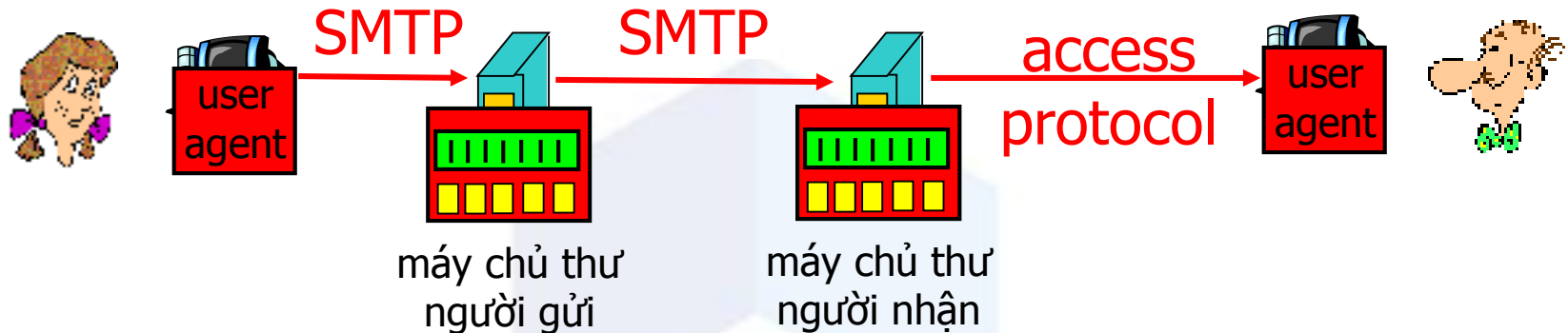
SMTP: giao thức để trao đổi thư điện tử

RFC 822: chuẩn dùng cho định dạng thông điệp văn bản:

- các dòng mào đầu, vd:
 - To:
 - From:
 - Subject:*khác nhau phụ thuộc vào câu lệnh SMTP!*
- thân
 - nội dung thư, chỉ chứa kí tự ASCII



Giao thức truy cập thư



- SMTP: giao/nhận thư đến/từ máy chủ khác
- Giao thức truy cập thư: lấy thư từ máy chủ
 - POP: Post Office Protocol [RFC 1939]
 - kiểm tra danh tính (UA<-->máy chủ) và tải về
 - IMAP: Internet Mail Access Protocol [RFC 1730]
 - nhiều chức năng hơn (phức tạp hơn)
 - quản lý thư lưu trên máy chủ
 - HTTP: gmail, Hotmail, Yahoo! Mail, v.v..

Giao thức POP3

pha kiểm tra danh tính

- câu lệnh của khách:
 - **user**: khai báo tên
 - **pass**: mật-khẩu
- phản hồi của chủ
 - **+OK**
 - **-ERR**

pha giao dịch, khách:

- **list**: liệt kê số thứ tự thư
- **retr**: tải thư theo stt
- **dele**: xóa
- **quit**: kết thúc

```
S: +OK POP3 server ready
C: user bob
S: +OK
C: pass hungry
S: +OK đăng nhập thành công
```

```
C: list
S: 1 498
S: 2 912
S: .
C: retr 1
S: <message 1 contents>
S: .
C: dele 1
C: retr 2
S: <message 1 contents>
S: .
C: dele 2
C: quit
S: +OK máy chủ POP3 đồng ý ngắt
```


POP3 (tt) và IMAP

POP3

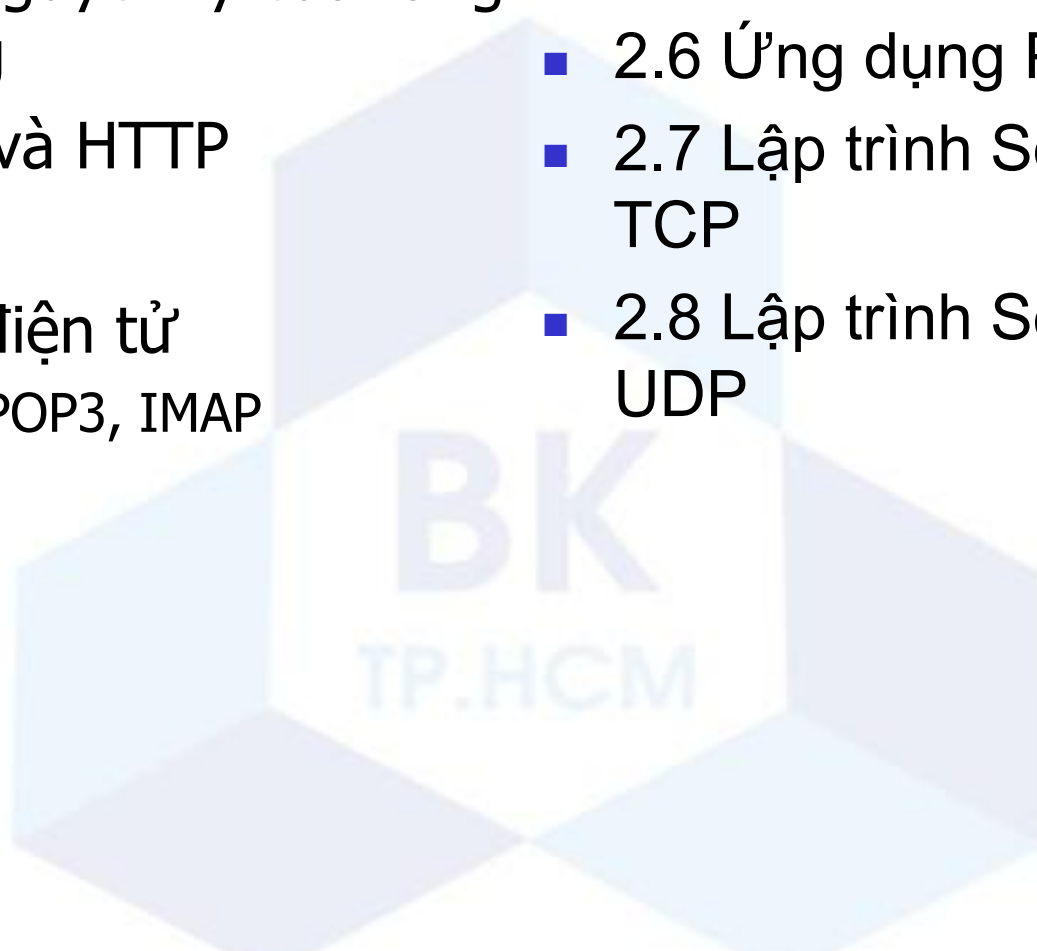
- ví dụ trước sử dụng cơ chế “tải-và-xóa”.
- Bob không thể đọc lại thư nếu như anh ta đổi UA
- “tải-và-giữ lại”: bản sao của thư sẽ được lưu trên nhiều UA
- POP3 là giao thức không trạng thái

IMAP

- Giữ tất cả thư trong một nơi: trên máy chủ
- Cho phép người dùng tổ chức thư theo thư mục
- IMAP giữ lại trạng thái người dùng qua các phiên làm việc:
 - tên của thư mục cũng như ánh xạ giữa ID của thông điệp và tên thư mục

Chương 2: Tầng ứng dụng

- 2.1 Các nguyên lý của tầng ứng dụng
- 2.2 Web và HTTP
- 2.3 FTP
- 2.4 Thư điện tử
 - SMTP, POP3, IMAP
- 2.5 DNS
- 2.6 Ứng dụng P2P
- 2.7 Lập trình Socket với TCP
- 2.8 Lập trình Socket với UDP



Hệ thống tên miền – DNS (Domain Name System)

Con người: nhiều kí hiệu nhận dạng:

- tên, số CMND, MSSV

Máy tính và thiết bị trên Internet:

- địa chỉ IP (32 bit) – được sử dụng để liên lạc giữa máy tính
- “tên”, vd: www.yahoo.com – được sử dụng bởi con người

Hỏi: ánh xạ giữa địa chỉ IP và tên?

Hệ thống tên miền:

- *có sở dữ liệu phân tán* được hiện thực trong một hệ thống phân cấp của nhiều *máy chủ tên miền*
- *giao thức tầng ứng dụng* cho phép máy tính, bđt và máy chủ tên miền liên lạc với nhau để *phân giải* tên miền (địa chỉ/tên)

DNS

Các dịch vụ DNS

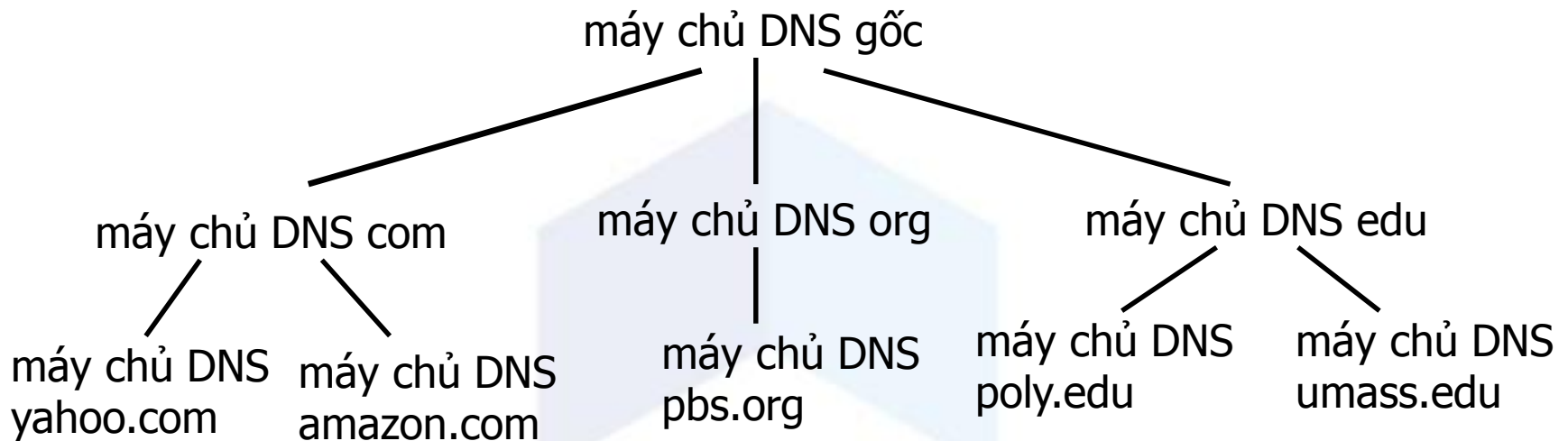
- phân giải tên miền sang địa chỉ IP
- đặt tên thay thế cho máy
 - tên chính thức, và tên thay thế
- đặt tên thay thế cho máy chủ email
- phân bổ tải
 - những trang web lớn thường có nhiều máy chủ với địa chỉ IP khác nhau đằng sau một tên miền chính thức

Tại sao không tập trung DNS?

- DNS bị lỗi -> cả mạng ngừng hoạt động
- quá tải lưu lượng
- độ trễ cao do đường truyền dài
- khó bảo trì

không mở rộng được!

Cơ sở dữ liệu phân tán và phân cấp

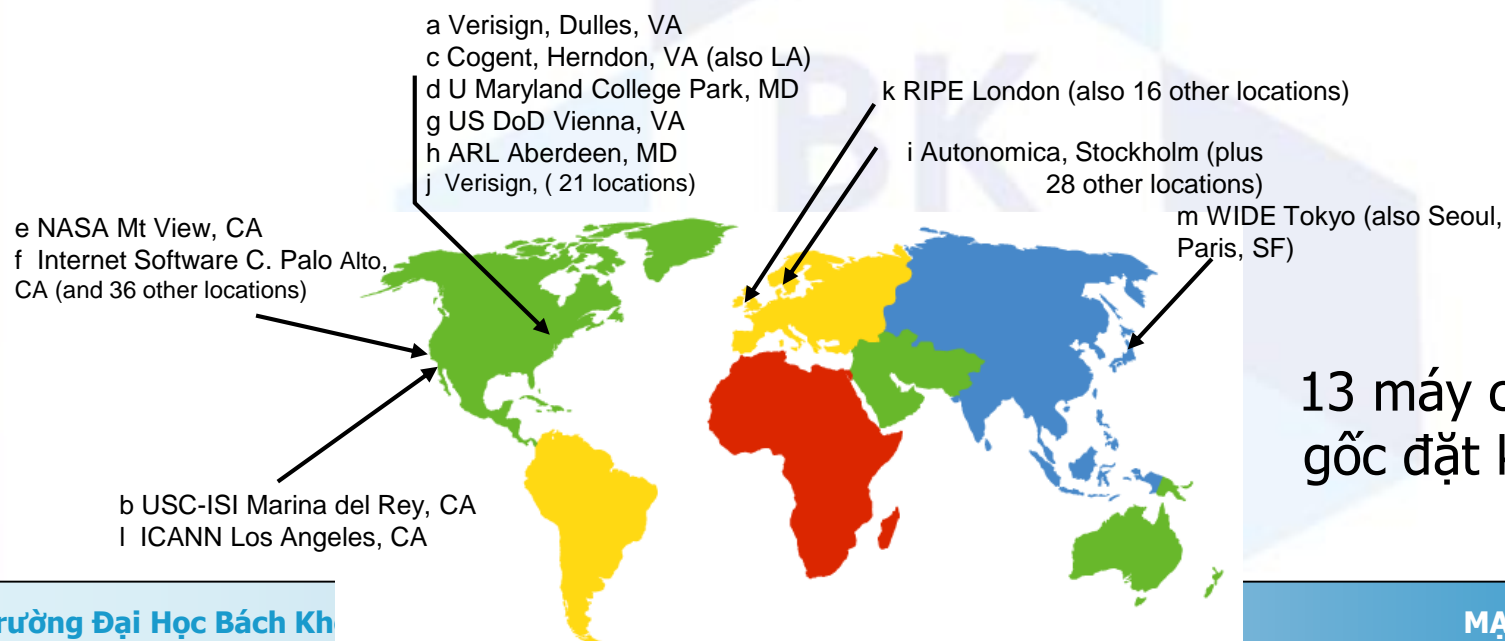


Khách muốn có IP của www.amazon.com:

- khách truy vấn một máy chủ DNS gốc để tìm ra máy chủ DNS ".com"
- khách truy vấn máy chủ DNS ".com" để lấy địa chỉ của máy chủ DNS amazon.com
- khách truy vấn máy chủ DNS "amazon.com" để lấy địa chỉ IP của www.amazon.com

DNS: Máy chủ gốc

- các máy chủ tên miền cục bộ sẽ liên hệ máy chủ gốc nếu chúng không thể tự phân giải tên miền
- máy chủ tên miền gốc:
 - liên hệ những máy chủ có thẩm quyền nếu không tìm ra ánh xạ
 - lấy thông tin ánh xạ
 - trả thông tin ánh xạ về cho các máy chủ tên miền cục bộ



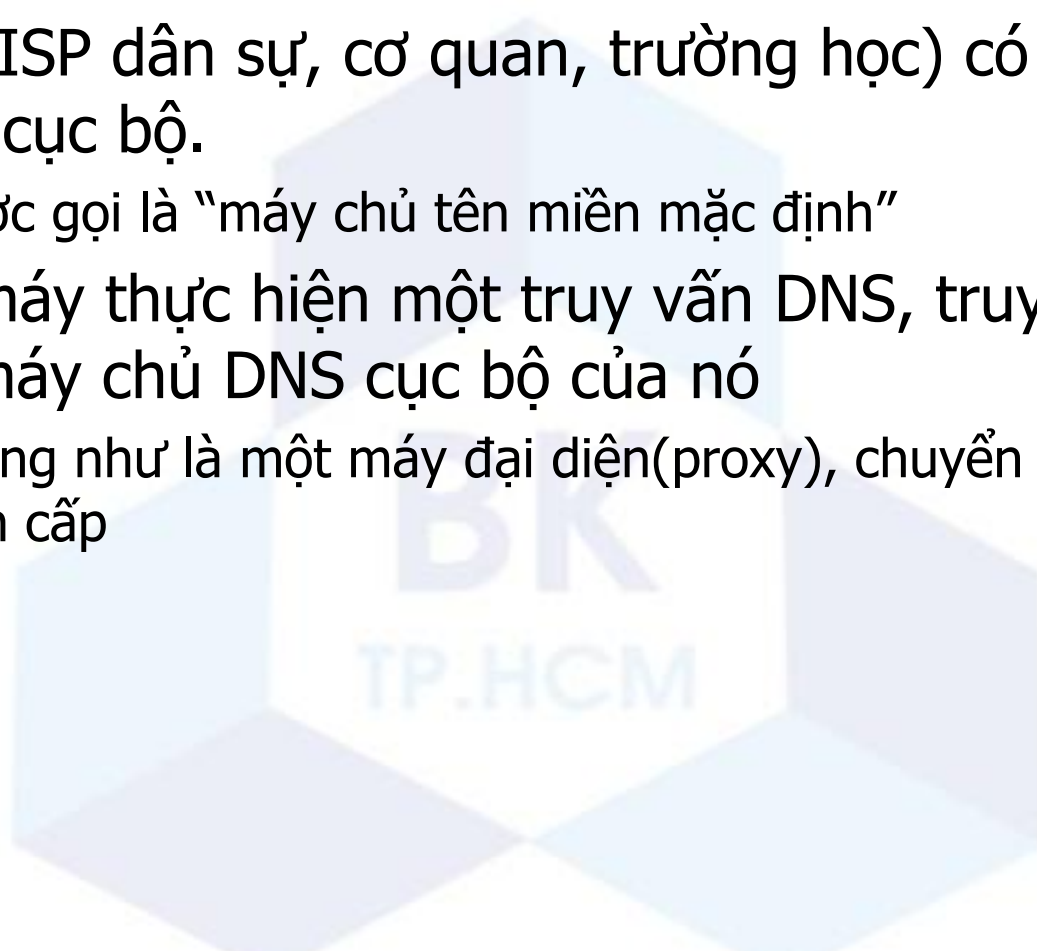
13 máy chủ tên miền gốc đặt khắp thế giới

TLD và Máy chủ thẩm quyền

- **Máy chủ tên miền cấp cao (Top-level domain - TLD):**
 - chịu trách nhiệm cho .com, .org, .net, .edu, .etc, và tất cả các tên miền quốc gia cấp cao .vn, .au, .ca, .jp
 - Network Solutions duy trì máy chủ TLD .com
 - Educause duy trì máy chủ TLD .edu
- **Máy chủ DNS thẩm quyền:**
 - máy chủ DNS cơ quan, cung cấp những ánh xạ tin cậy giữa tên miền và IP cho những máy chủ của chính cơ quan đó (vd: Web, mail).
 - có thể được vận hành bởi cơ quan hay nhà cung cấp dịch vụ

Máy chủ tên miền cục bộ

- không phụ thuộc một cách rõ ràng vào hệ thống phân cấp
- mỗi ISP (ISP dân sự, cơ quan, trường học) có một máy chủ tên miền cục bộ.
 - còn được gọi là “máy chủ tên miền mặc định”
- khi một máy thực hiện một truy vấn DNS, truy vấn sẽ được gửi cho máy chủ DNS cục bộ của nó
 - hoạt động như là một máy đại diện(proxy), chuyển tiếp truy vấn lên hệ phân cấp

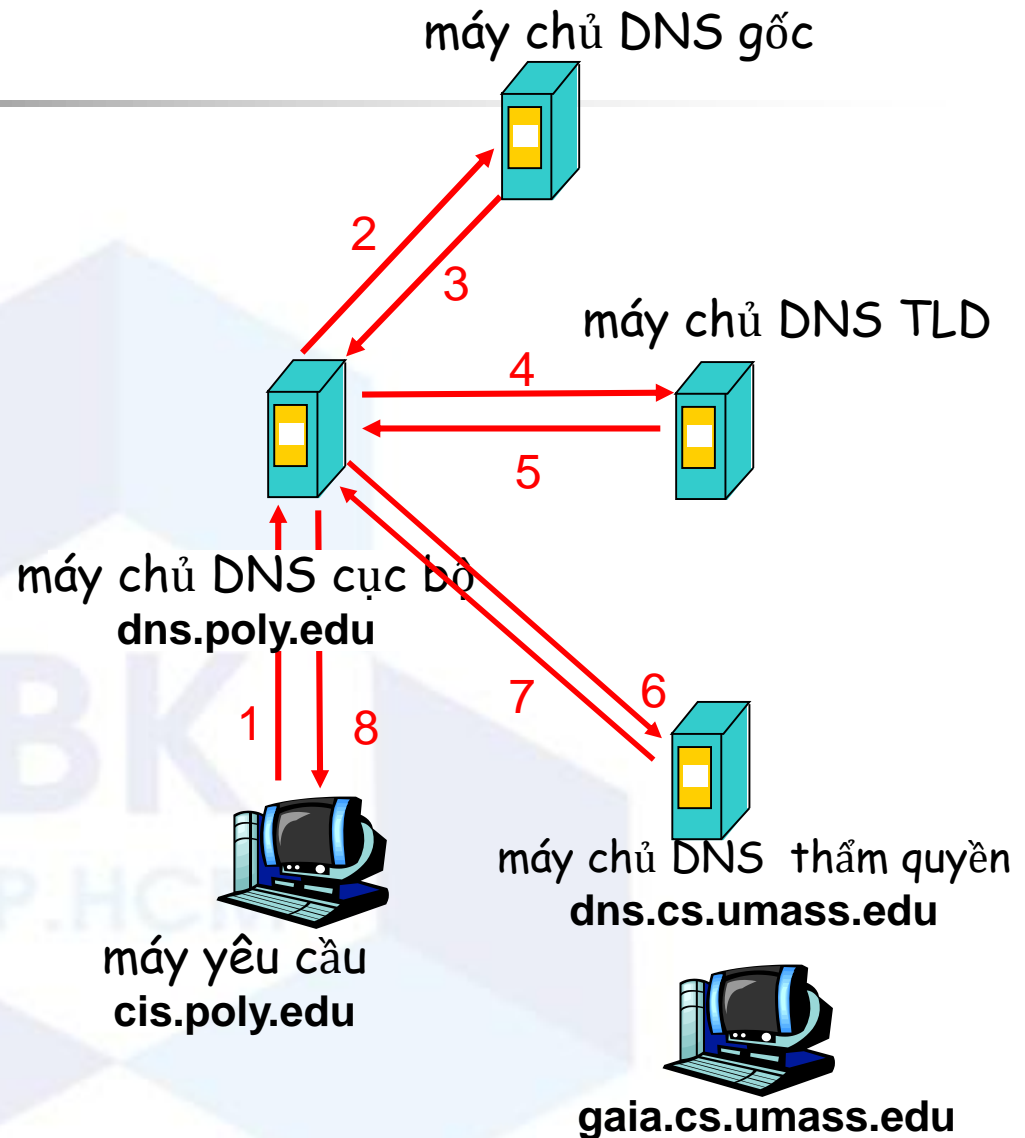


Ví dụ phân giải tên miền DNS

- Máy tại cis.poly.edu muốn có IP của gaia.cs.umass.edu

truy vấn lặp:

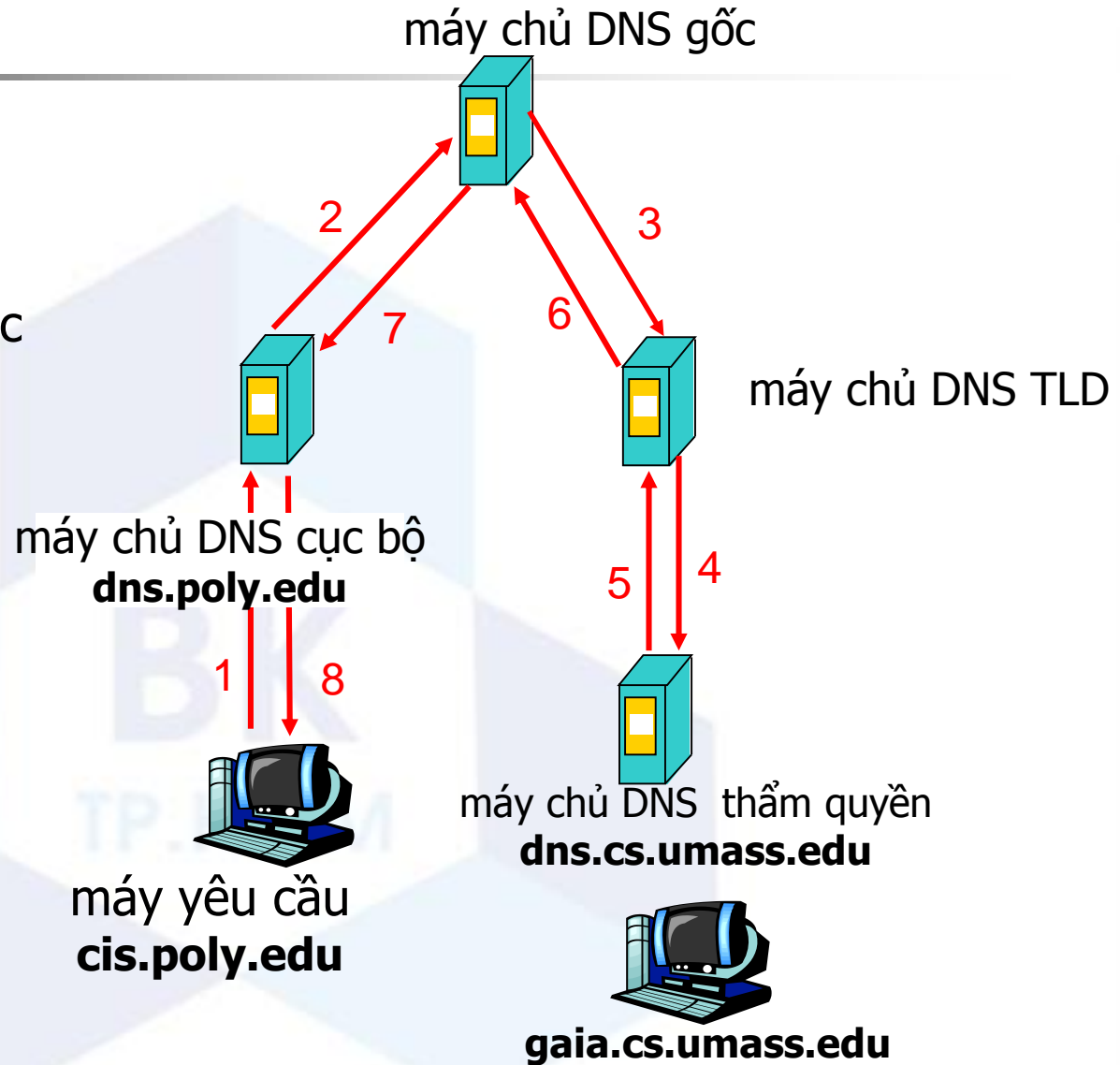
- máy chủ được liên hệ trả về tên của máy chủ khác để liên hệ tiếp
- "Tôi không biết tên miền này, nhưng hãy hỏi máy chủ này xem"



Ví dụ phân giải tên miền DNS

truy vấn đệ quy:

- ☐ giao toàn bộ công việc cho máy chủ cấp cao hơn
- ☐ tải cao?



DNS: nhớ đệm và cập nhật thẻ ghi

- khi máy chủ tên miền nhận được giá trị ánh xạ, nó sẽ *nhớ đệm* lại thông tin đó
 - các thông tin trong bộ nhớ đệm sẽ hết hạn và bị xóa sau một thời gian nhất định
 - tên các máy chủ TLD thường được nhớ đệm trong máy chủ cục bộ
 - vì vậy máy chủ DNS gốc thường ít khi được truy cập
- cơ chế cập nhật/thông báo được thiết kế bởi IETF
 - RFC 2136
 - <http://www.ietf.org/html.charters/dnsind-charter.html>

Thẻ ghi DNS

DNS: CSDL phân bố chứa thẻ tài nguyên (RR)

định dạng RR: (tên, giá trị, loại, TGS)

- Loại A (Type=A)
 - **tên** là tên máy
 - **giá trị** là địa chỉ IP
- Type=NS
 - **tên** là tên miền (vd: foo.com)
 - **giá trị** là tên máy của máy chủ DNS có thẩm quyền cho miền này
- Type=CNAME
 - **tên** là tên thay thế cho tên chính thống
www.ibm.com thực chất là servereast.backup2.ibm.com
 - **giá trị** là tên chính thống
- Type=MX
 - **giá trị** là tên của máy chủ thư liên hệ tới tên máy "tên"

Giao thức và thông điệp DNS

DNS protocol : *truy vấn (query)* và *phản hồi (reply)* có cùng *định dạng*

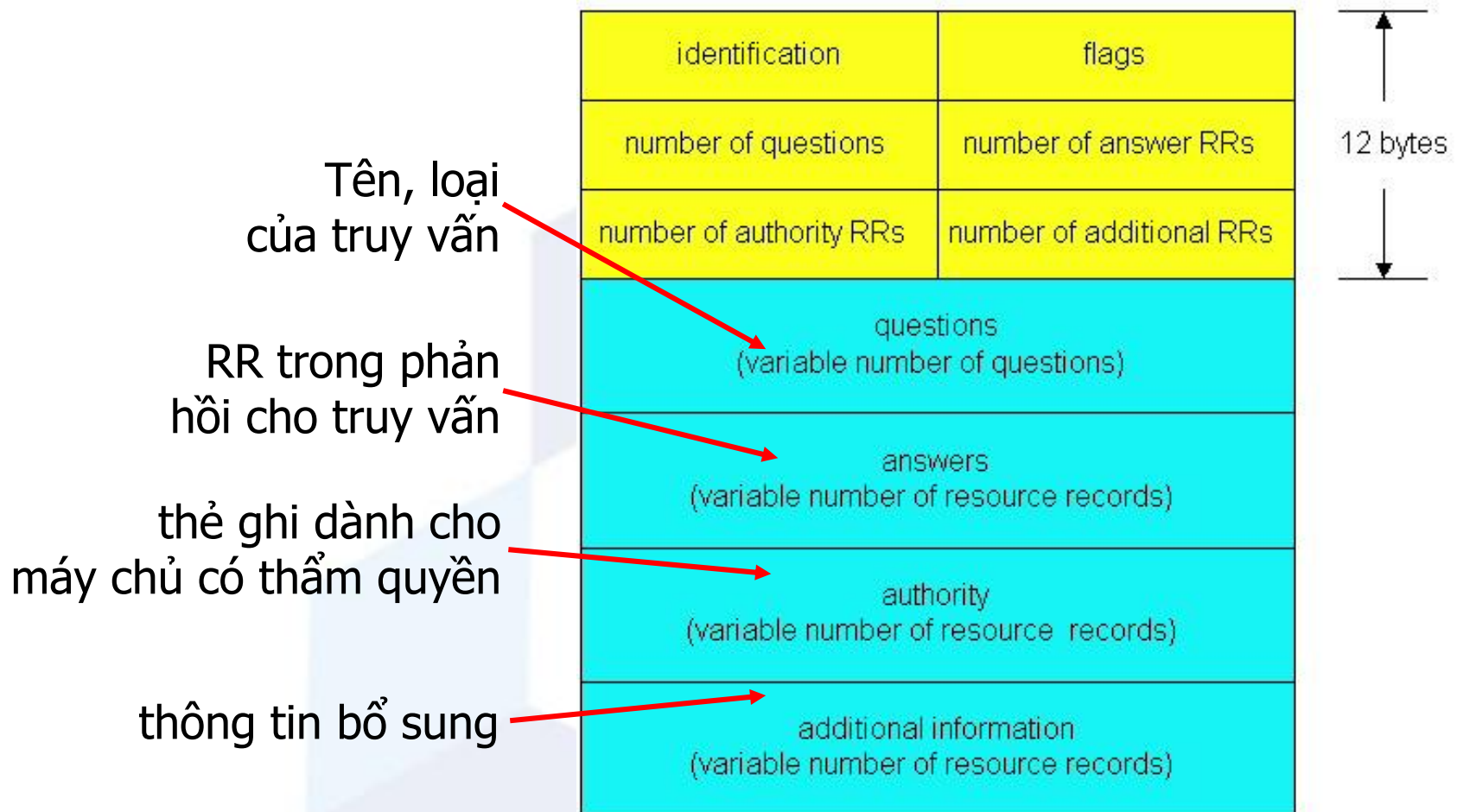
Phần mào đầu

- **số hiệu định danh**: là một số 16 bit trong thông điệp truy vấn, thông điệp phản hồi sử dụng chính số đó
- **cờ hiệu**:
 - truy vấn hay phản hồi
 - mong muốn đệ qui
 - đệ qui có sẵn
 - phản hồi là có thẩm quyền

identification	flags
number of questions	number of answer RRs
number of authority RRs	number of additional RRs
questions (variable number of questions)	
answers (variable number of resource records)	
authority (variable number of resource records)	
additional information (variable number of resource records)	

↑
12 bytes
↓

Giao thức và thông điệp DNS



Chèn thẻ ghi vào DNS

- ví dụ: một công ty mới thành lập “Network Utopia”
- đăng kí tên miền networkutopia.com tại *DNS nhà quản lý tên miền*
 - cung cấp tên, địa chỉ IP, địa chỉ IP của máy chủ DNS có thẩm quyền (sơ cấp và thứ cấp)
 - nhà quản lý chèn 2 thẻ RR vào máy chủ TLD “com” :

```
(networkutopia.com, dns1.networkutopia.com, NS)
(dns1.networkutopia.com, 212.212.212.1, A)
```

- trong máy chủ tên miền cục bộ tạo ra thẻ loại A cho `www.networkutopia.com`; thẻ loại MX cho `mail.networkutopia.com`
- *Làm sao mọi người lấy được địa chỉ IP của trang web của bạn?*