Data and Computer Communications

Chapter 23 – Internet Applications Internet Directory Service and the World Wide Web

Eighth Edition

by William Stallings

Lecture slides by Lawrie Brown

Internet Applications
Internet Directory Service and the World Wide Web

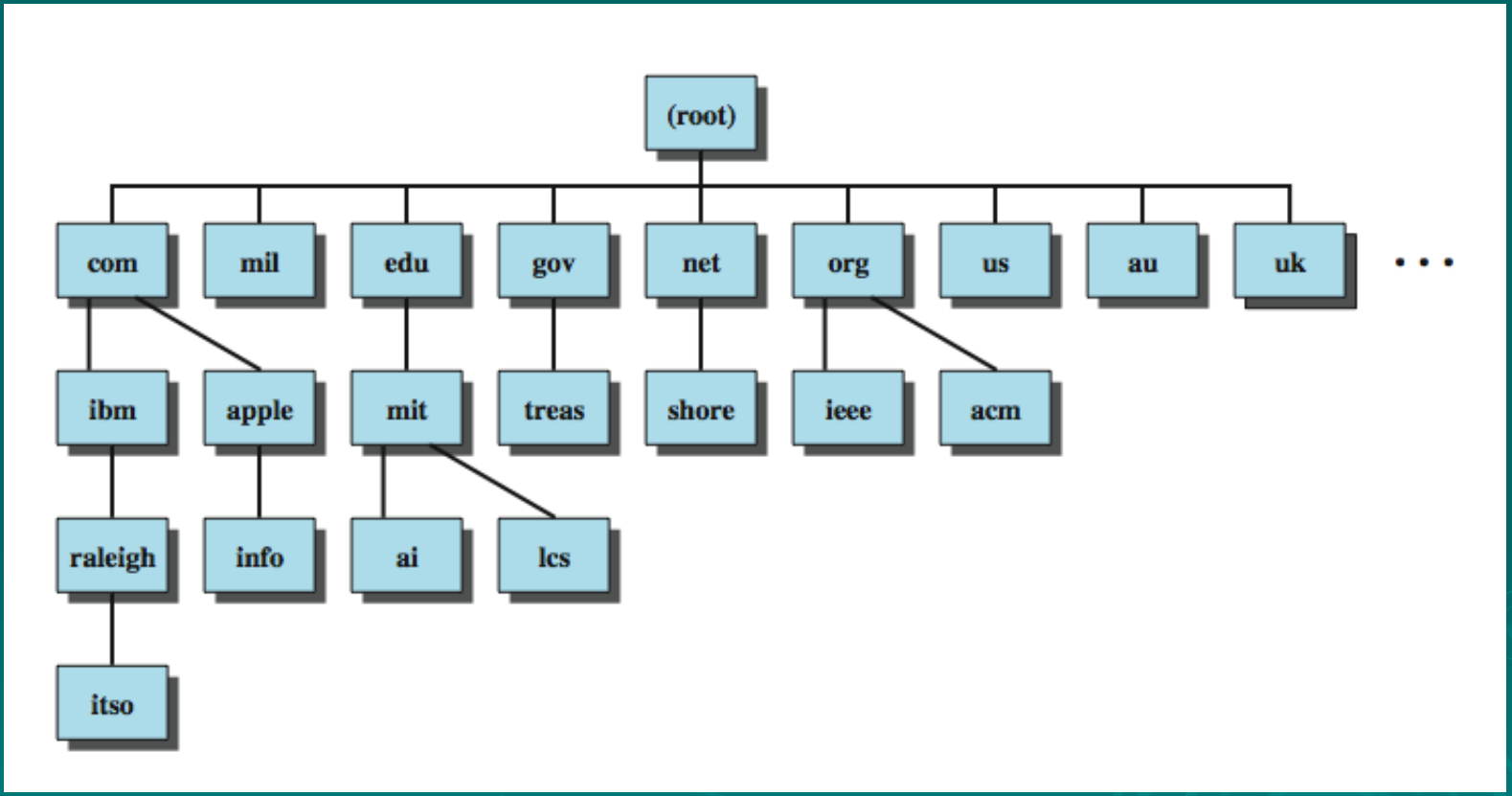
Life in the modern world is coming to depend more and more upon technical means of communication. Without such technical aids the modern city-state could not exist, for it is only by means of them that trade and business can proceed; that goods and services can be distributed where needed; that railways can run on schedule; that law and order are maintained; that education is possible. Communication renders true social life practicable, for communication means organization.

—*On Human Communication*, Colin Cherry

DNS
The Internet Directory Service

- the Domain Name Service (DNS) provides mapping between host name & IP address
- defined in RFCs 1034 / 1035
- key elements
  - domain name space
  - DNS database
  - name servers
  - name resolvers

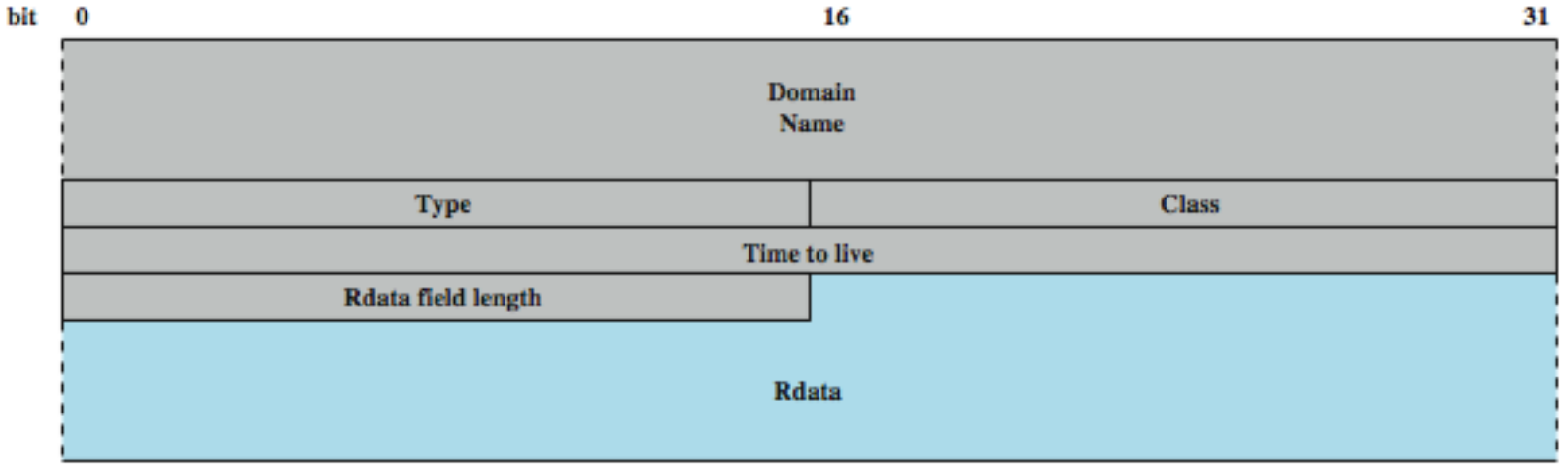
Domain Names



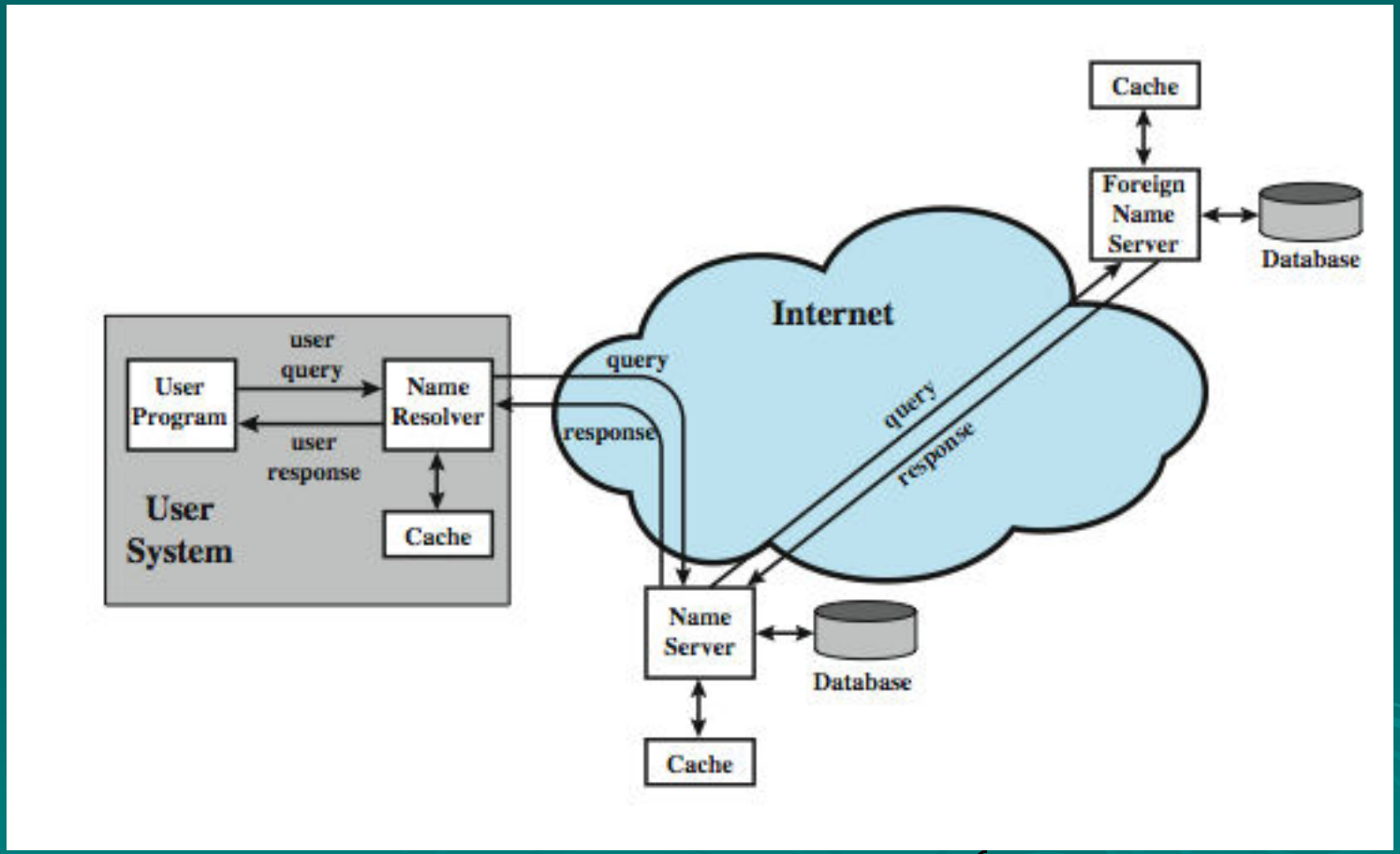
DNS Database

- hierarchical database
- containing resource records (RRs)
- features
  - variable-depth hierarchy for names
  - distributed database
  - distribution controlled by database
- provides name-to-address directory service for network applications

Resource Records (RRs)



DNS Operation



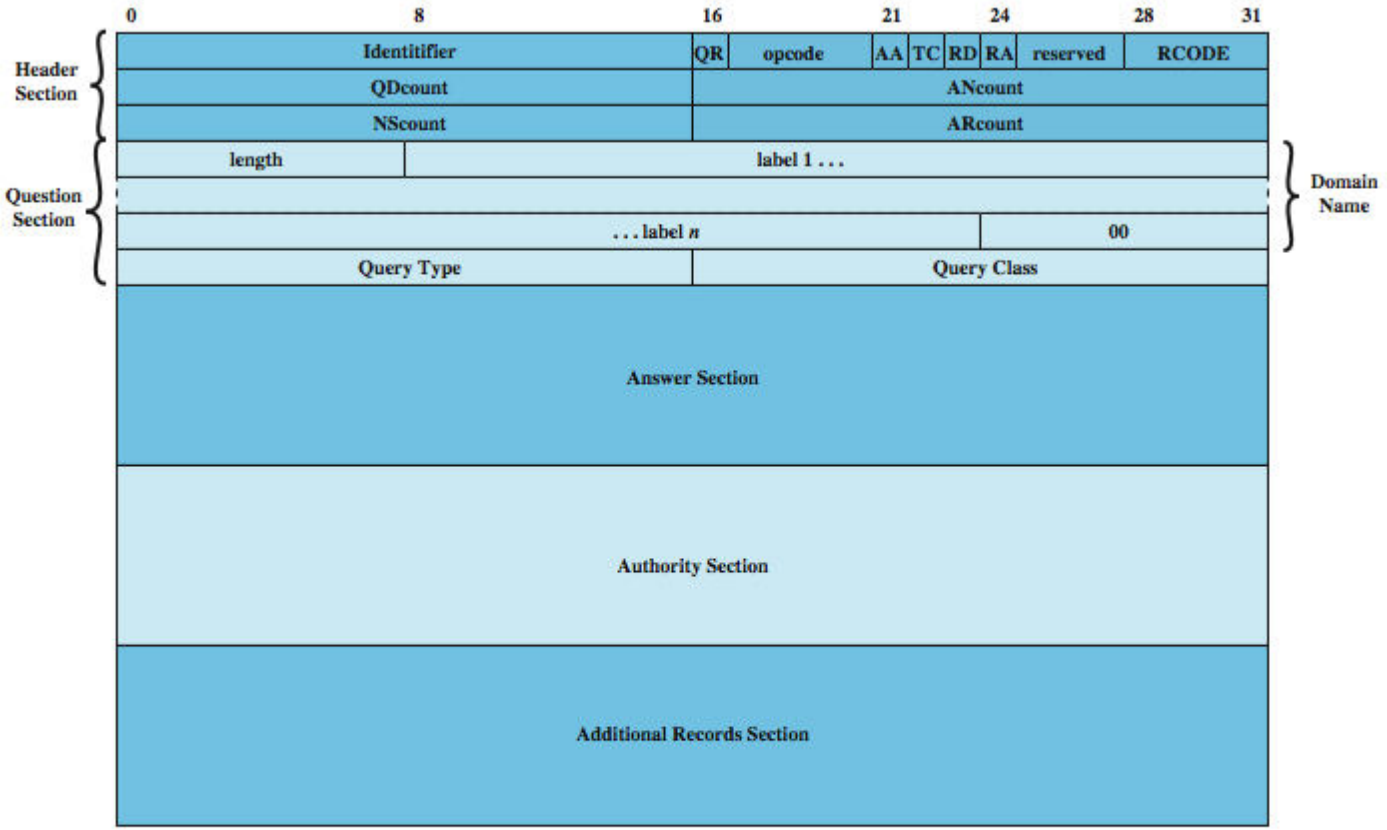
DNS Server Hierarchy

- DNS database is distributed hierarchically
  - may extend as deep as needed
- any organization owning a domain can run name servers
- each server manages authoritative name data for a zone
- 13 **root name servers** at top of hierarchy share responsibility for top level zones

Name Resolution

- query begins with name resolver on host
- knows name/address of local DNS server
- given a name request, the resolver can:
  - return name from cache if already known
  - send DNS query to local server which may return answer, or query other servers
- recursive technique - server queries other servers for resolver
- iterative technique - resolver queries servers in turn as needed

DNS Messages



QR = query/response bit
AA = authoritative answer
TC = truncated
RD = recursion desired
RA = recursion available

RCODE = response code
QDcount = number of entries in question section
ANcount = number of resource records in answer section
NScount = number of name server resource records in authority section
ARcount = number of resource records in additional records section

Hypertext Transfer Protocol HTTP

- base protocol for World Wide Web
- for any hypertext client/server application
- is a protocol for efficiently transmitting information to make hypertext jumps
  - can transfer plain text, hypertext, audio, images, and Internet accessible information
- versions 0.9, 1.0, & now 1.1 (RFC2616)

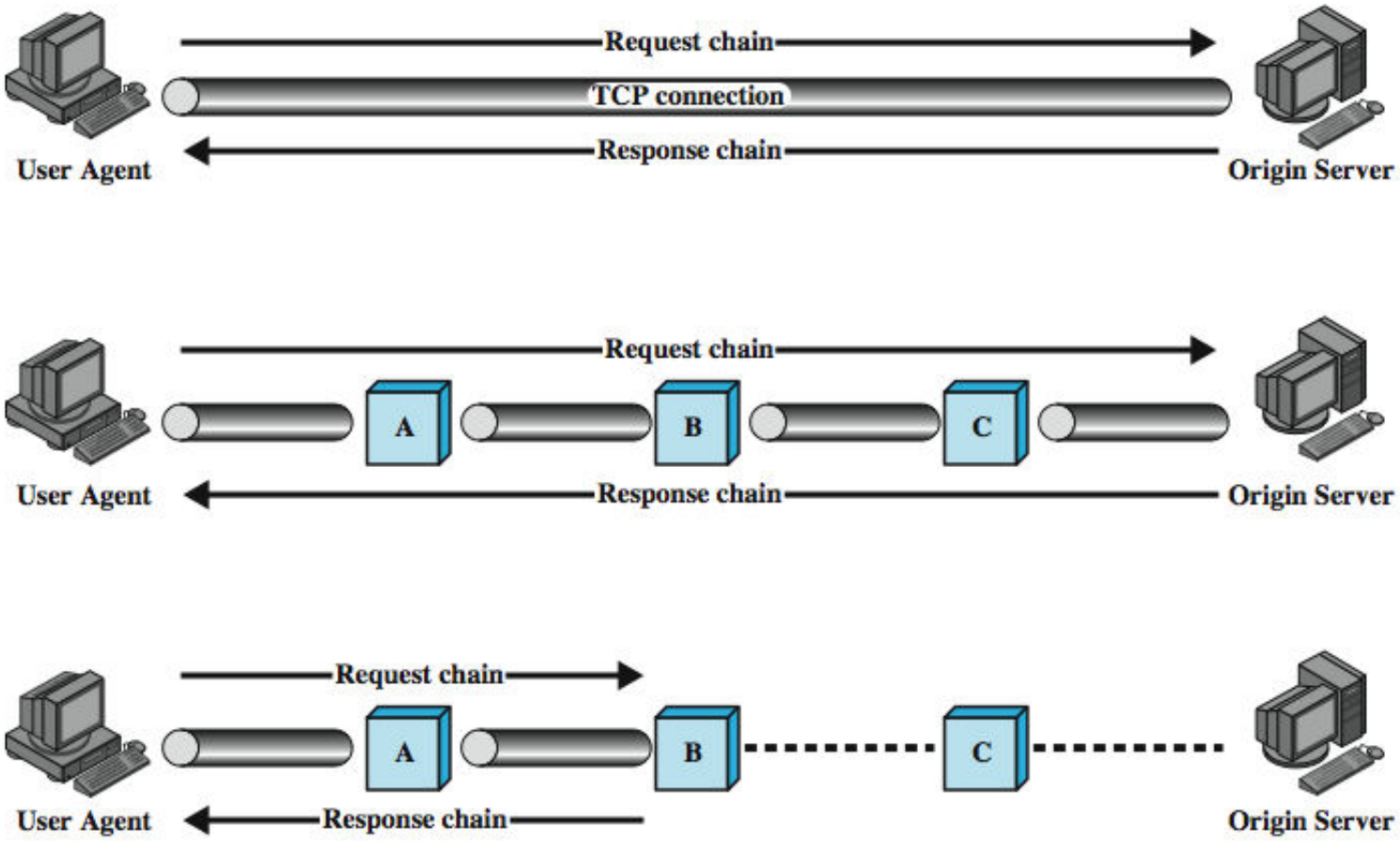
HTTP Overview

- transaction oriented client/server protocol
- between Web browser (client) and Web server
- uses TCP connections
- stateless
  - each transaction treated independently
  - each new TCP connection for each transaction
  - terminate connection when transaction complete
- flexible format handling
  - client may specify supported formats

Key Terms

- cache
- client
- connection
- entity
- gateway
- message
- origin server
- proxy
- resource
- server
- tunnel
- user agent

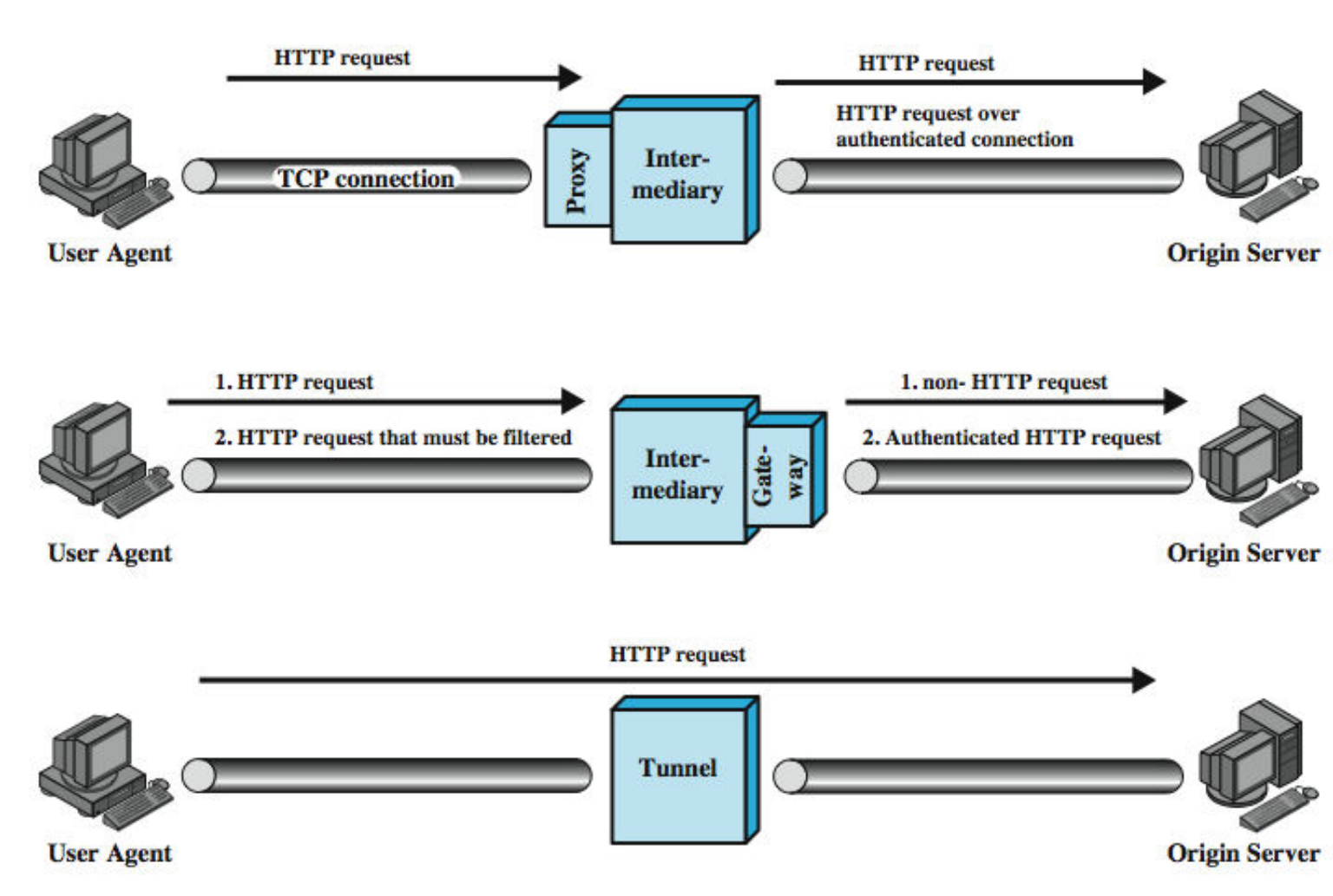
Examples of HTTP Operation



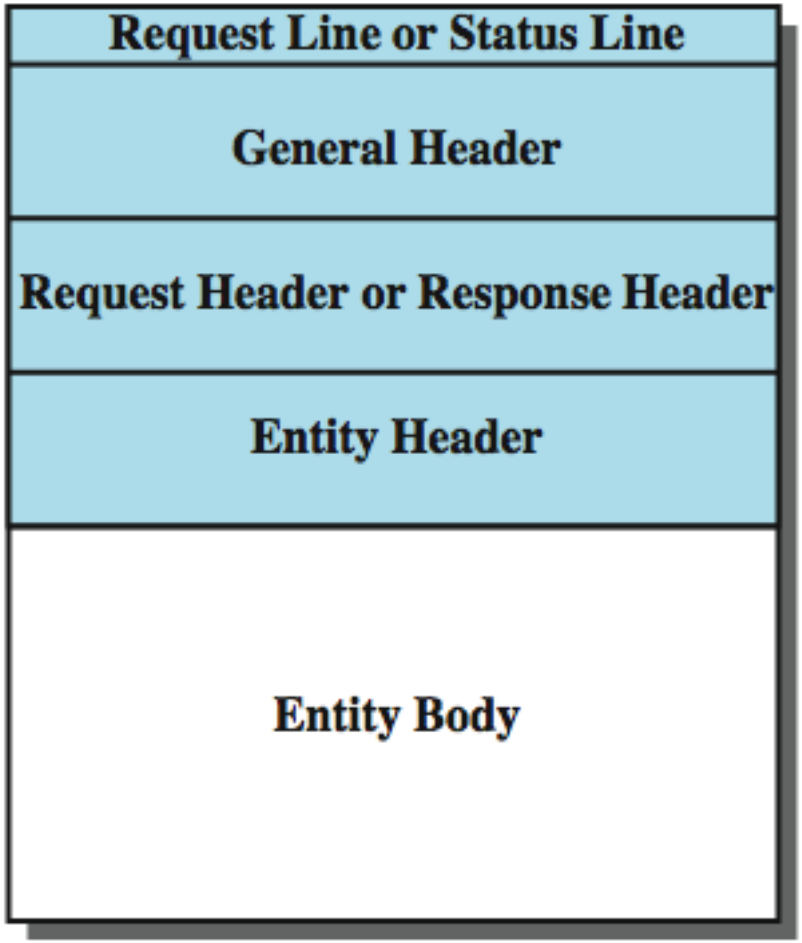
HTTP Operation - Caches

- often have a web cache
- stores previous requests/ responses
- may return stored response to subsequent requests
- may be a client, server or intermediary system
- not all requests can be cached

Intermediate HTTP Systems



HTTP Messages



HTTP Messages BNF Format

HTTP-Message = Simple-Request | Simple-Response |
Full-Request | Full-Response

Full-Request = Request-Line
*(General-Header | Request-Header | Entity-Header)
CRLF
[Entity-Body]

Full-Response = Status-Line
*(General-Header | Response-Header | Entity-Header)
CRLF
[Entity-Body]

Simple-Request = "GET" SP Request-URL CRLF

Simple-Response = [Entity-Body]

HTTP General Header Fields

- Cache-Control
- Connection
- Data
- Forwarded
- Keep-Alive
- Mime-Version
- Pragma
- Upgrade

Request Methods

- request-line has
  - method
  - Request URL
  - HTTP version
  - Request-Line = Method Request-URL HTTP-Version CRLF
- HTTP/1.1 methods:
  - OPTIONS, GET, HEAD, POST, PUT, PATCH, COPY, MOVE, DELETE, LINK, UNLINK, TRACE, WRAPPED, Extension-method

Request Header Fields

- Accept, Accept-Charset, Accept-Encoding, Accept-Language, Authorization, From, Host, If-Modified-Since, Proxy-Authentication, Range, Referrer, Unless, User-Agent

Response Messages

- status line plus one or more general, response, entity headers, then optional entity body
- status line contains
  - HTTP version
  - status code
  - reason phrase
  - Status-Line = HTTP-Version SP Status-Code SP Reason-Phrase CRLF

Status Codes

- informational - headers only
- successful - headers & body if relevant
- redirection - further action needed
- client error - has syntax or other error
- server error - failed to satisfy valid request

Response Header Fields

- Location
- Proxy-Authentication
- Public
- Retry-After
- Server
- WWW-Authenticate

Entity Header Fields

- Allow
- Content-Encoding
- Content-Language
- Content-Length
- Content-MD5
- Content-Range
- Content-Type
- Content-Version
- Derived-From
- Expires
- Last-Modified
- Link
- Title
- Transfer-Encoding
- URL-Header
- Extension-Header

Entity Body

- entity body is an arbitrary sequence of octets
- HTTP can transfer any type of data including:
  - text, binary data, audio, images, video
- data is content of resource identified by URL
- interpretation data determined by header fields:
  - Content-Type - defines data interpretation
  - Content-Encoding - applied to data
  - Transfer-Encoding - used to form entity body

Summary

- domain name service (DNS)
  - names, database, name resolution, messages
- HyperText Transfer Protocol (HTTP)
  - overview
  - request and response messages