

# Evaluation Method

- Interim and Final Report
- Attendance is not Checked, but, ...
- Questions or Comments are Mandated
  - In the quater, questions or comments with technical content must be made at least twice during lecture (may be in Japanese)
  - Good questions and comments will be awarded with points
  - Declare your name and student ID after each lecture, if you make questions or comments

Advanced Lecture on Internet Applications

# 7. Text based Communication: Web, HTTP, HTML, JAVA Script

Masataka Ohta

[mohita@necom830.hpcl.titech.ac.jp](mailto:mohita@necom830.hpcl.titech.ac.jp)

<ftp://ftp.hpcl.titech.ac.jp/appli7e.ppt>

# BTW, What is the Internet?

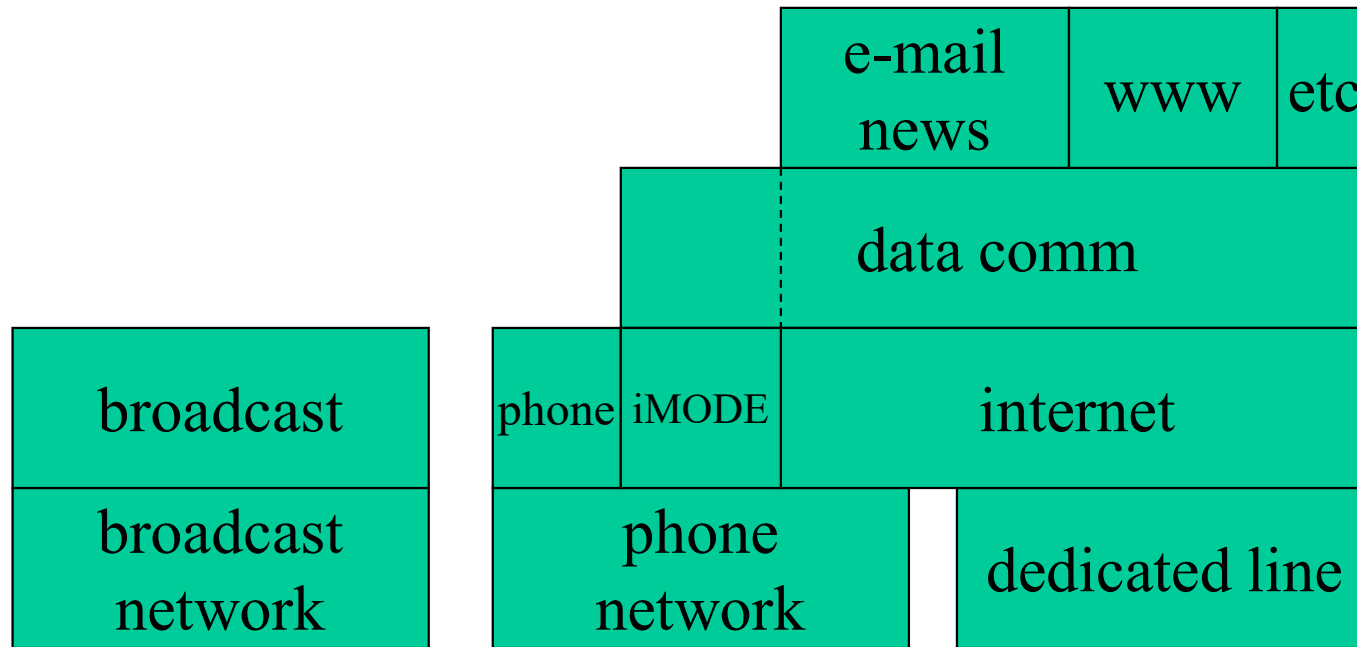
- Not e-mail
  - seriously thought so 20 years ago
- Not web, either
  - many still misunderstand so
- Is not applications
- The Internet is a network directly connecting terminals based on the principle of the Internet using IP (Internet Protocol)

# Application?

- face skin (important to be a beauty)
  - total revenue of official i-mode sites is about 4% of total revenue of Docomo
- “though US may invent web” (speech by a Swiss politician in INET @ Geneva)
- we need some application, but any application is OK

broadcast	phone	data comm
broadcast network	phone network	dedicated line

networks before the Internet



networks with the Internet

broadcast	phone	e-mail news	www	etc
streaming		data comm (batch)		
internet				
dedicated line (including wireless)				

network in the future

# History of Web

- a teletext system
  - NAPLPS (North America) and CAPTAIN (Japan) failed
  - MINITEL (around France) was successful
    - easy to provide servers, free terminals
- web is Internet version of MINITEL
- so successful because of the Internet and Moore's law
  - easy to provide servers, inexpensive terminals



# HTML (HyperText Markup Language, rfc1866)

- hypertext
  - text with hyperlinks (URI) embedded
  - very old concept with more than 50 years of history
- mark up
  - formatting directives
    - such as indentations by string surrounded by “<z” and “>”
  - originally European centric, though improving
    - vertical writing etc.

# URI (Uniform Resource Identifier, rfc2396)

- URN (Uniform Resource Name)
  - identify resource without locating
  - a method to reference something
- URL (Uniform Resource Locator)
  - identify resource with location
    - if in the Internet, address of servers and port#
- can click and surf
  - because URLs are instantly used free of charge

# Example of URLs

`ftp://ftp.is.co.za/rfc/rfc1808.txt`

-- ftp scheme for File Transfer Protocol services

`gopher://spinaltap.micro.umn.edu/00/Weather/California/Los%20Angeles`

-- gopher scheme for Gopher and Gopher+ Protocol services

`http://www.math.uio.no/faq/compression-faq/part1.html`

-- http scheme for Hypertext Transfer Protocol services

`mailto:mduerst@ifi.unizh.ch`

-- mailto scheme for electronic mail addresses

`news:comp.infosystems.www.servers.unix`

-- news scheme for USENET news groups and articles

`telnet://melvyl.ucop.edu/`

-- telnet scheme for interactive services via the TELNET Protocol

# Format of URIs

- most generally
  - `<scheme>:<scheme-specific-part>`
- usually
  - `<scheme>://<authority><path>?<query>`
- useful `<authority>` is `<server>`
  - `server = [ [ userinfo "@" ] hostport ]`
    - `hostport = host [ ":" port ]`
    - `host = hostname | IPv4address`

# Object Oriented Processing over the Internet

- resource pointed by URL = object
- “scheme” (and prefix of “path”) identify class
- path identify object within “server”
- “query” identify method
  - default method is “access”
- parameters may be specified in “query”

# TCP and Command

- commands and replies represented in ASCII strings are exchanged over TCP
  - reply often begins with 3 digits followed by a space and text explaining reply **in English**
- line is terminated by CR and LF
- data may be sent over the same TCP connection (SMTP) or other TCP connection (FTP)
  - separator for data is necessary for sending over the same TCP

# FTP (rfc959)

- File Transfer Protocol
- protocol to exchange files over the Internet
- port# 21 is used for command and response
- various file formats are supported

# HTTP (HyperText Transfer Protocol, rfc2616)

- like FTP and SMTP, ASCII based exchange of commands and responses
- OPTIONS
  - ask what are the available options
- GET
  - retrieve content of resource including header
- HEAD
  - retrieve header of resource



# HTTP (2)

- POST
  - send data to resource
- PUT
  - replace resource by a sent data
- DELETE
  - delete resource

# HTTP (3)

- TRARC
  - get list of proxy chain
    - against E2E principle?
- CONNECT
  - reserved key word
    - to update proxy by tunneling

# Meaning of HTTP Replies (1)

- 1xx: Informational
  - Request received, continuing process
- 2xx: Success
  - The action was successfully received, understood, and accepted
- 3xx: Redirection
  - Further action must be taken in order to complete the request

# Meaning of HTTP Replies (2)

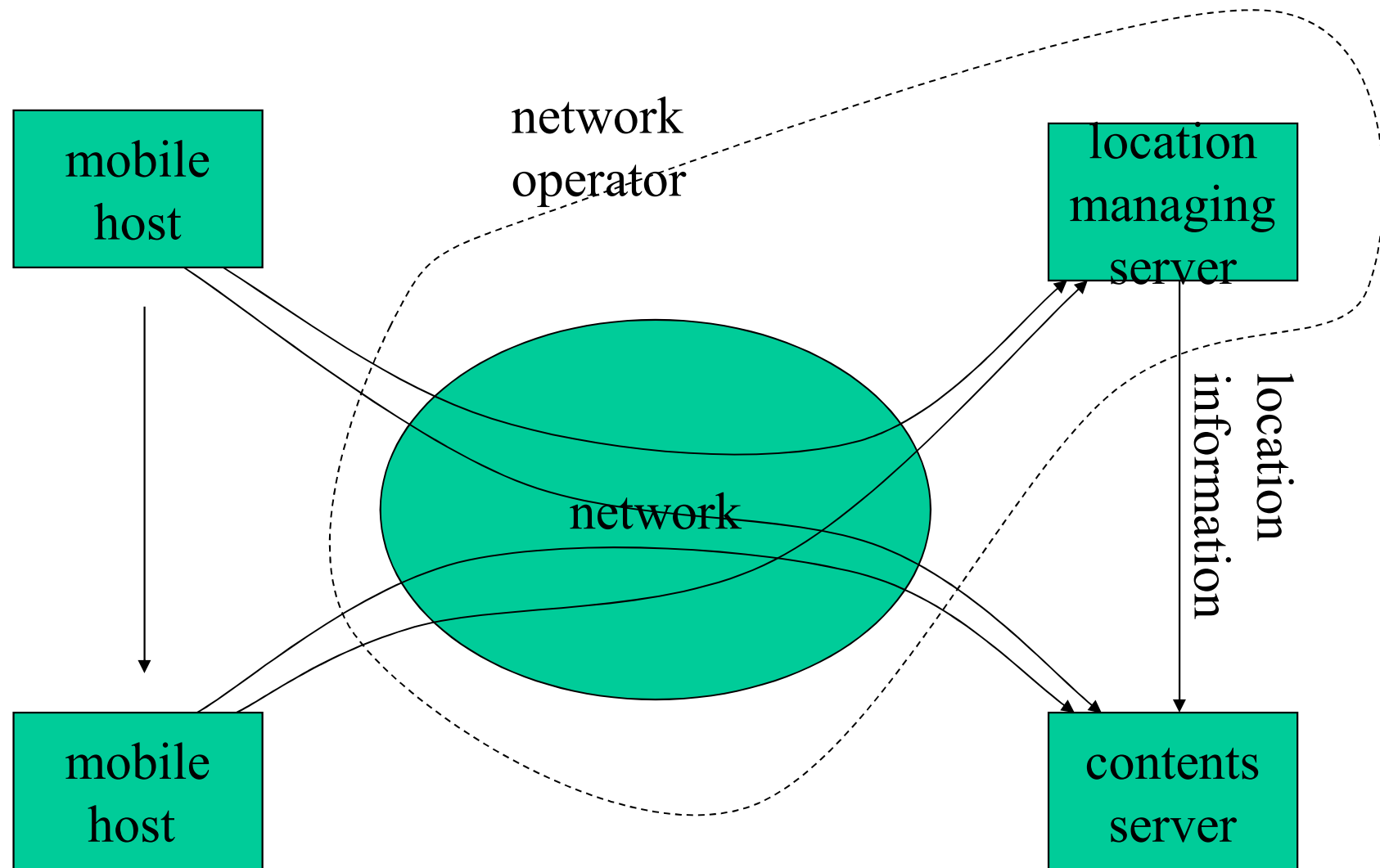
- 4xx: Client Error
  - The request contains bad syntax or cannot be fulfilled
- 5xx: Server Error
  - The server failed to fulfill an apparently valid request

# Internet and Contents Regulation

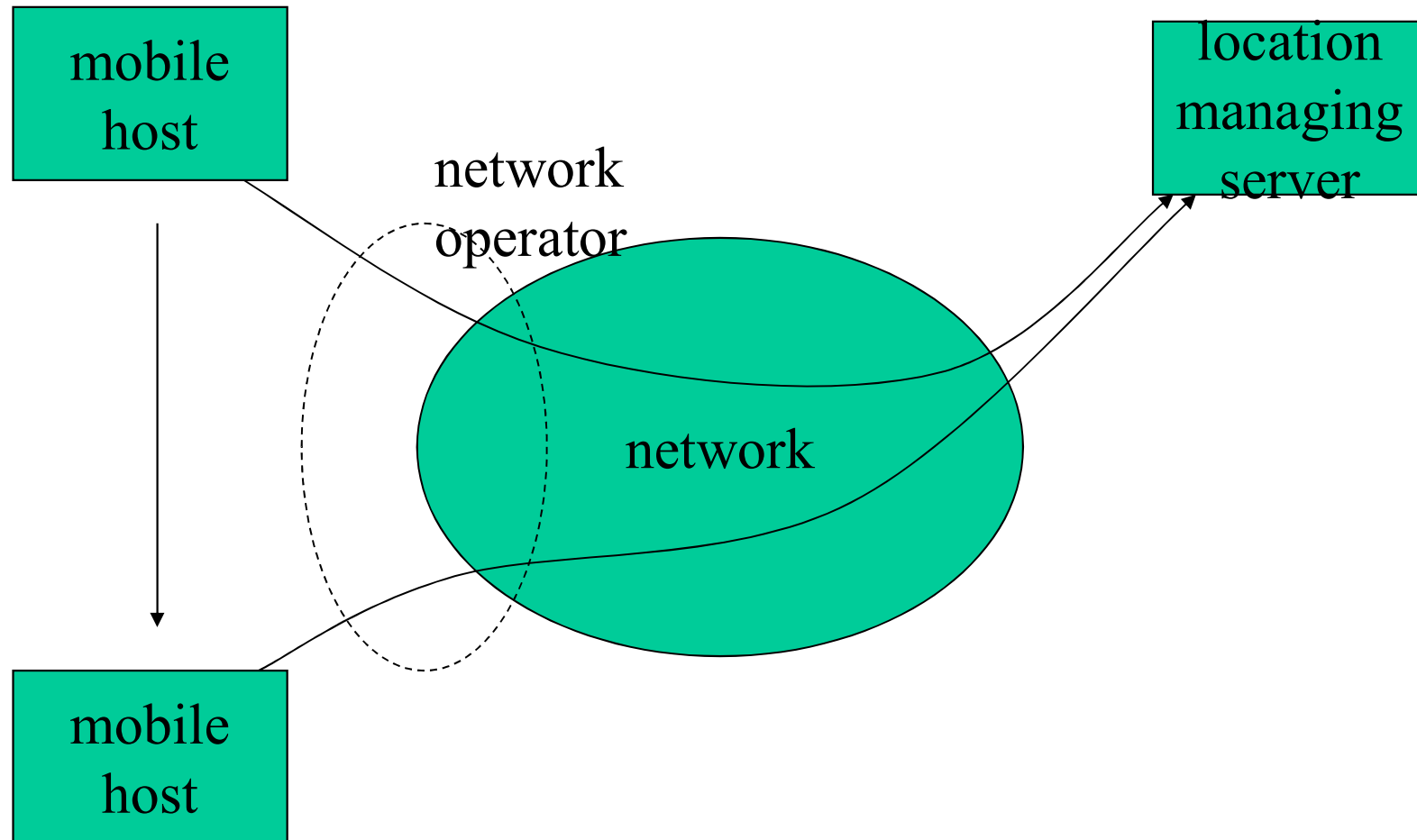
- with dial-up Internet access
  - server is centrally maintained by ISPs
  - (commercial) ISPs are easy to be regulated
- with persistent connectivity
  - anyone can operate servers
    - peer to peer model
  - must regulate all the end users
    - hard, though, not impossible

# Location Dependent Service

- contents change according to location
- impossible at IP layer
  - no location information in IP header
  - may be possible at application layer
- how can location information obtained automatically?



mobile hosts and location dependent contents in mobile phone network

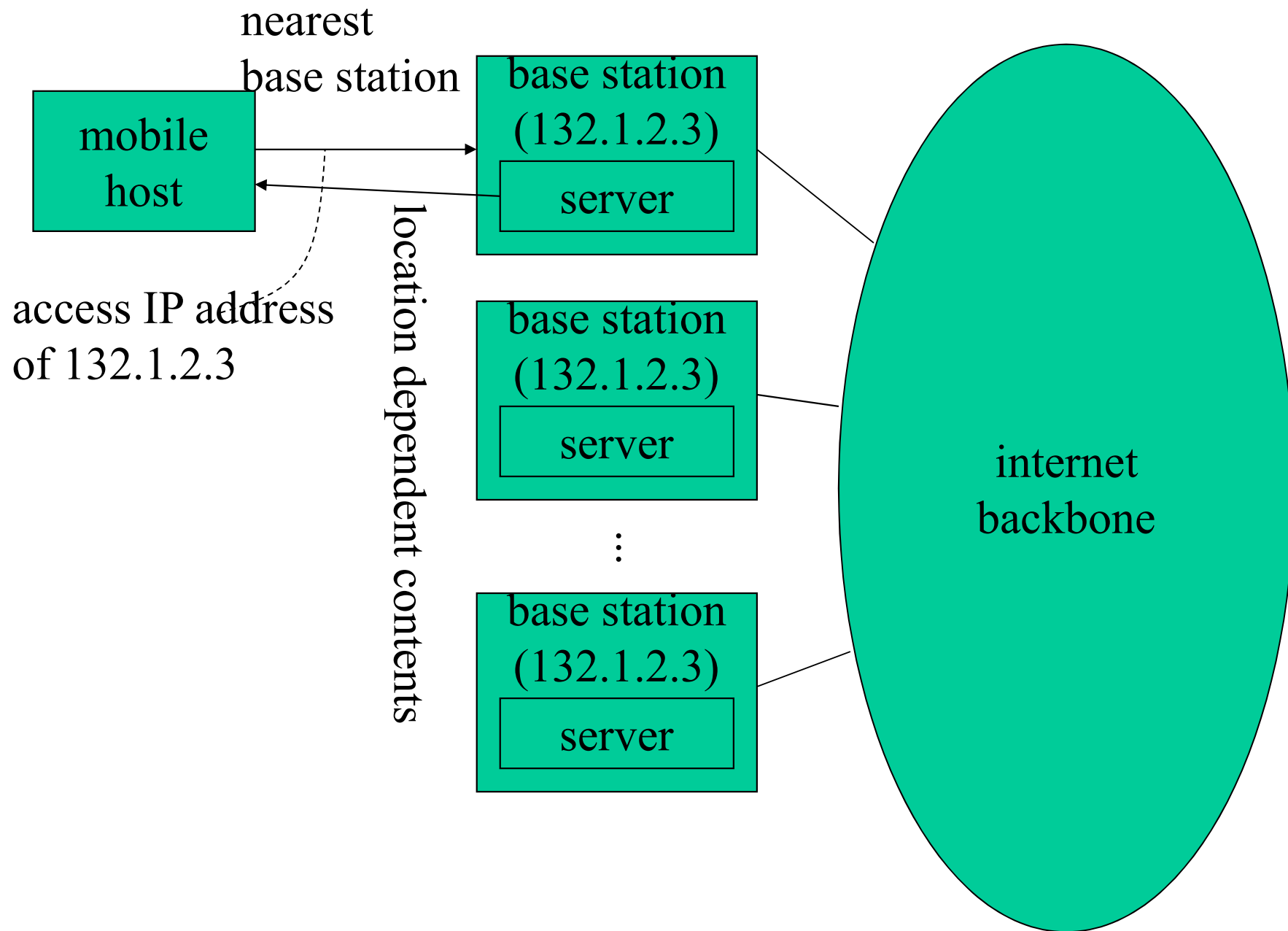


mobile hosts in the internet

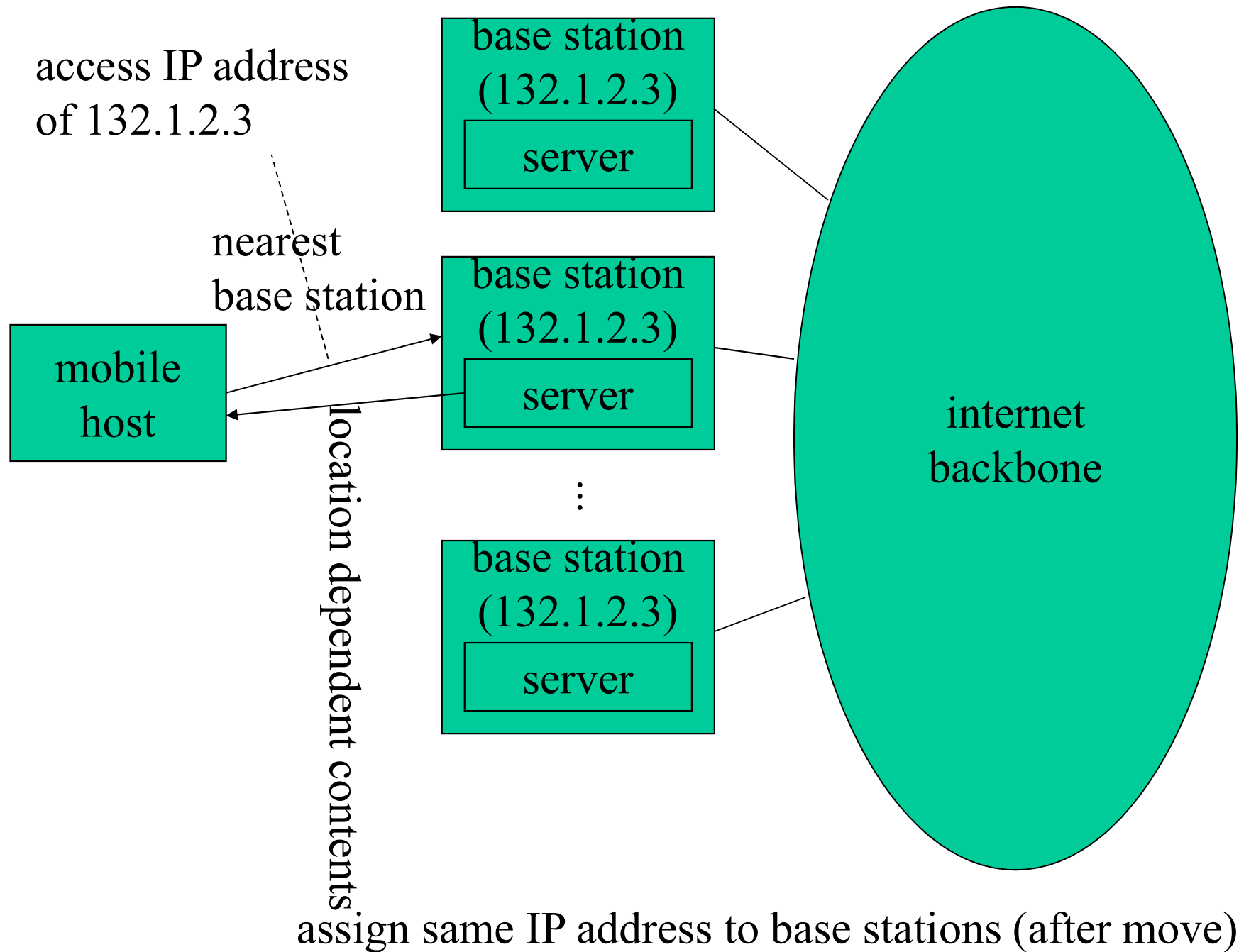


# Distributed Location Dependent Service by Anycast

- assign an (anycast) IP address to multiple base stations at different locations
- mobile hosts access nearest base station by the address (specified by URL)
- base station offers location dependent contents
  - or, redirect to other URL with location information added



assign same IP address to base stations



1) access URL:

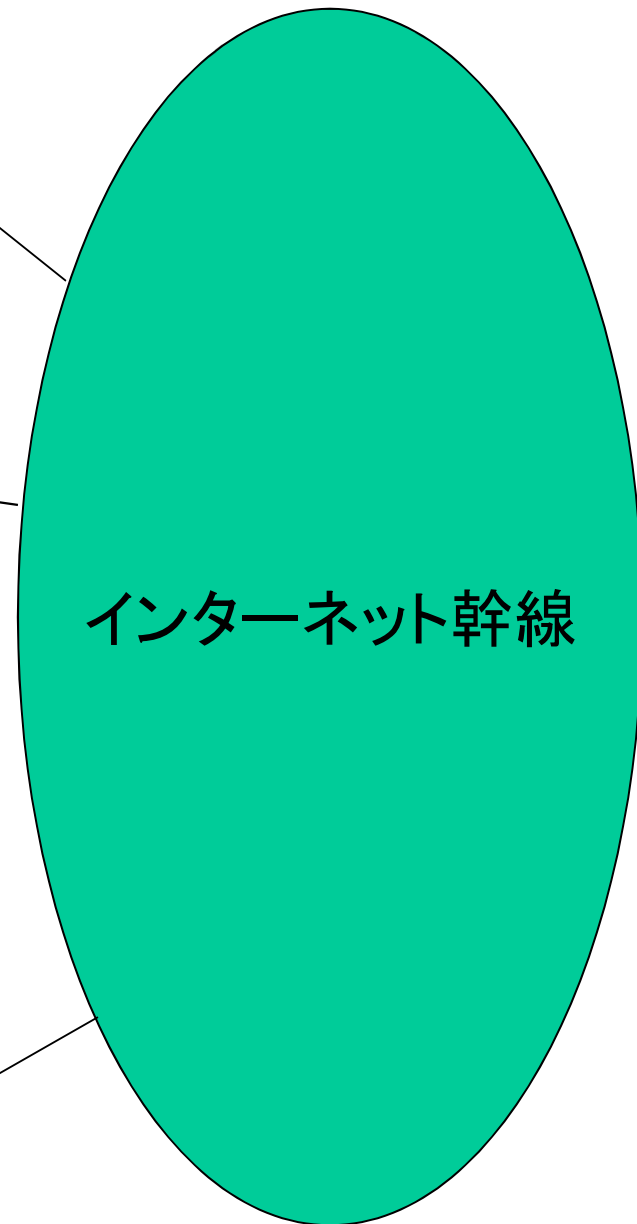
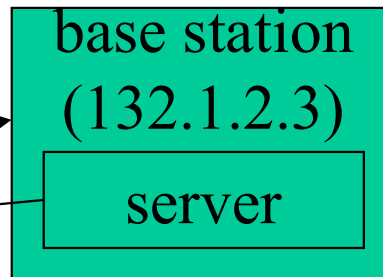
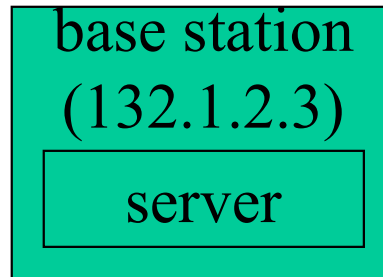
<http://map.here/>

(IP address of  
map.here is

132.1.2.3)

nearest

base station

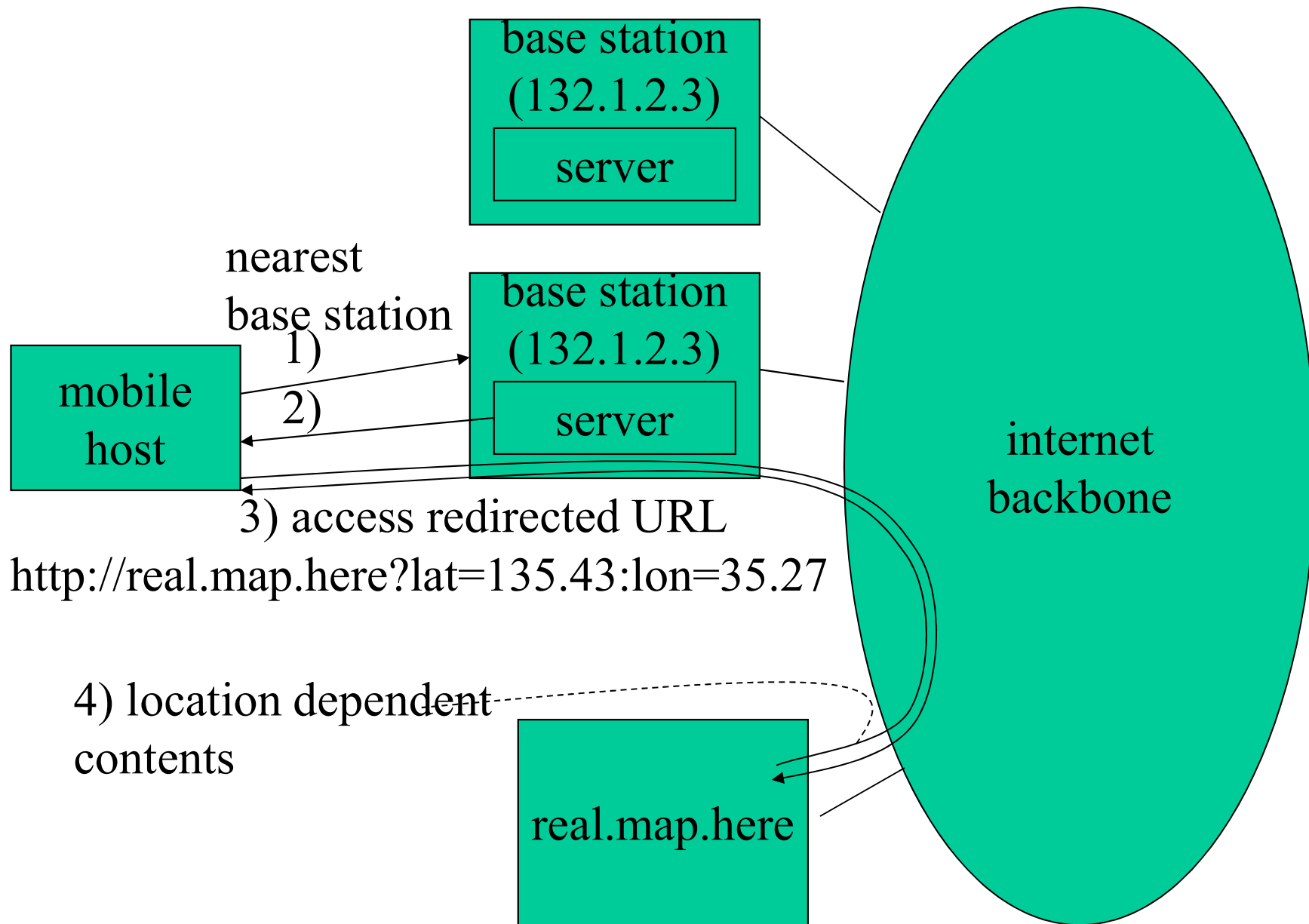


2) redirect to URL: :

<http://real.map.here?lat=135.43:lon=35.27>



example with URL and redirect



example with URL and redirect (cont'd)

# Merit of Anycast based Approach

- fully distributed
- no privacy concern
  - location information is first offered to the user

# JAVA (Script)

- embedded in HTML downloaded from server
- interpreted by client side
- client GUI can
  - initiate JAVA applet
    - may cause protocol actions
      - JAVA applet downloaded from server understand protocol
      - client do not have to understand protocol?

# JAVA Applet

- even if client does not know protocols recognized by server
  - JAVA applet interpreted by browsers
    - can communication with server
- because human beings are involved through GUI
  - human beings know protocols used for text/graphical representations in web pages
    - ■: stop, ►: play, ►: record etc.



# Standardization of JAVA API

- even if protocols are not standardized
  - if JAVA API is standardized
    - can control clients by JAVA code sent from server
- easier than standardizing protocols?
  - for people not familiar with networking, maybe
  - protocol standardization makes applets unnecessary
  - unconditional execution of programs from server may cause security problems
    - carefully designed protocols are better

# Wrap Up

- web is the most commonly used application
  - smoothly migrated from e-mail/bbs era
    - thanks to E2E principle
  - client server model
- HTML is transferred by HTTP
  - URL represents objects
- JAVA is for computer experts without much network expertise