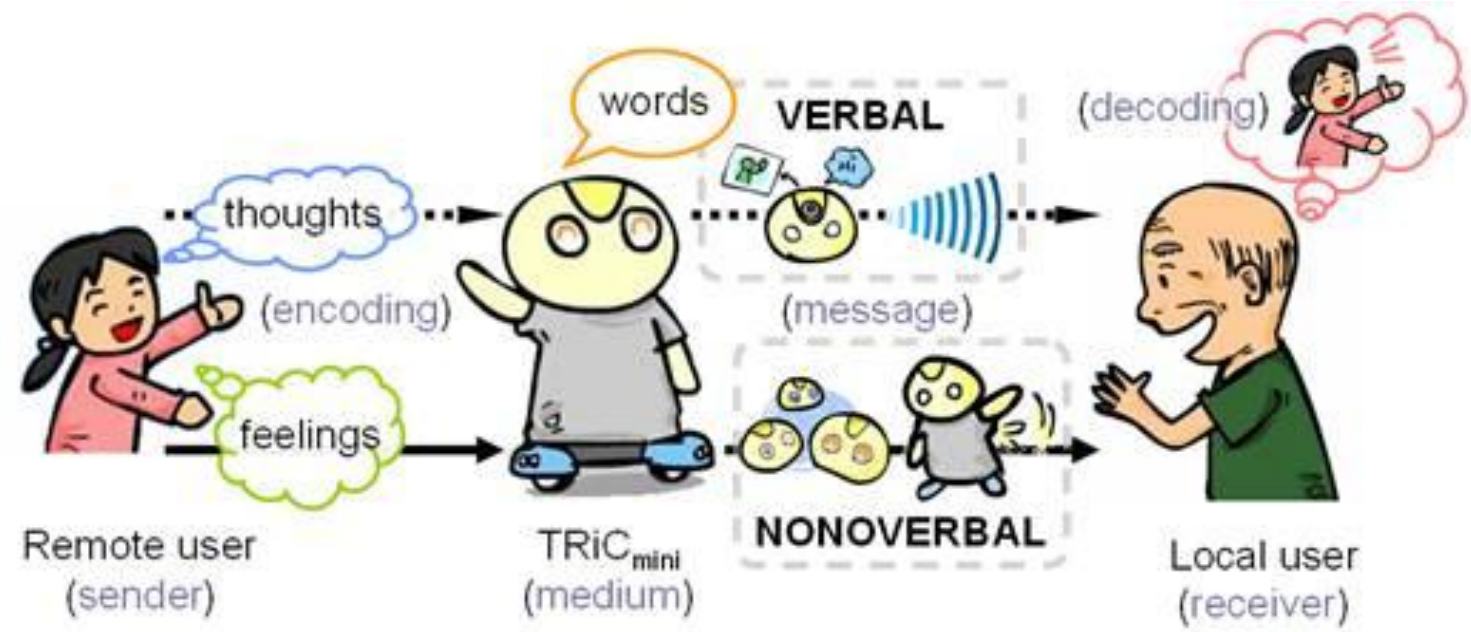
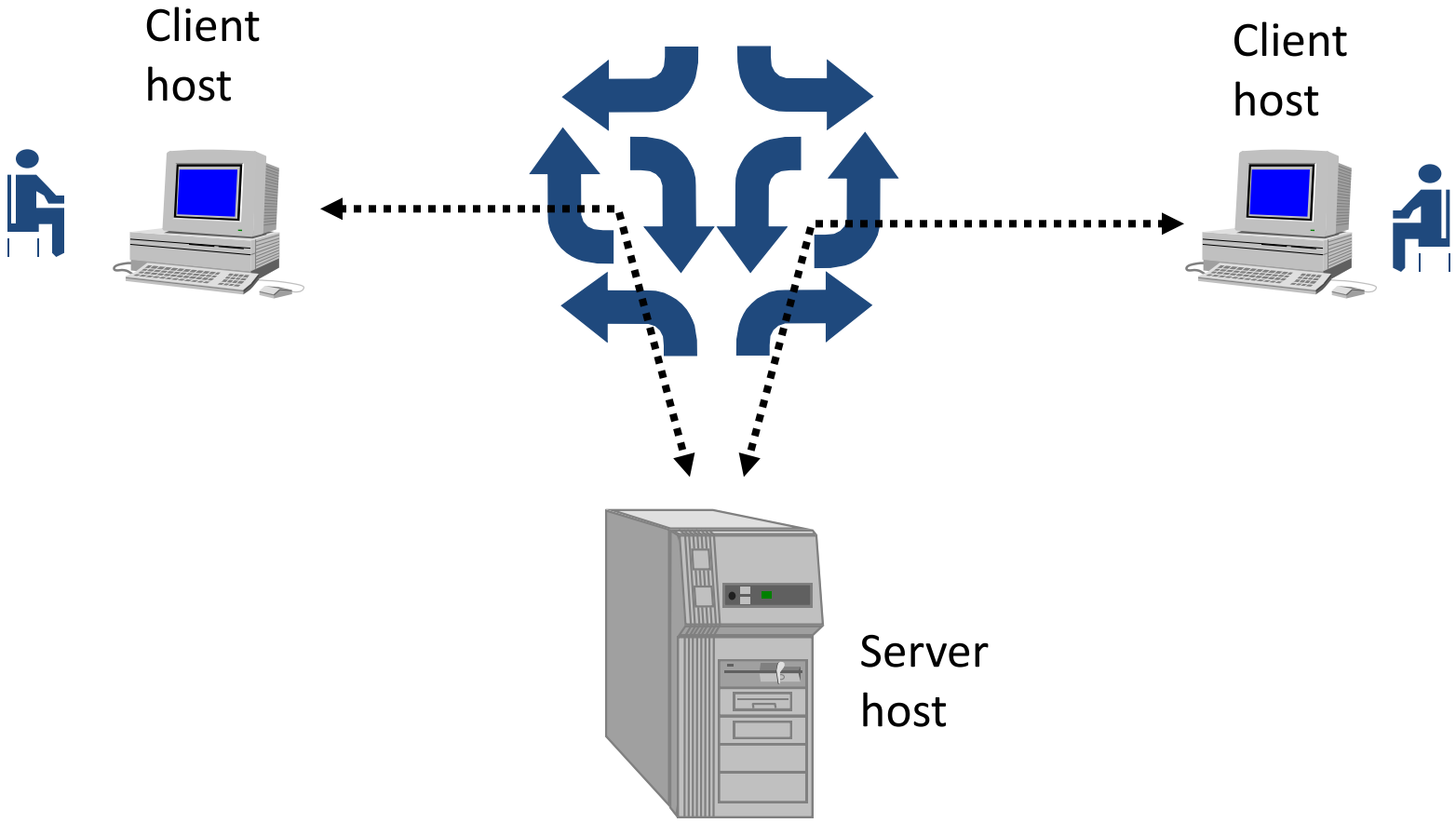
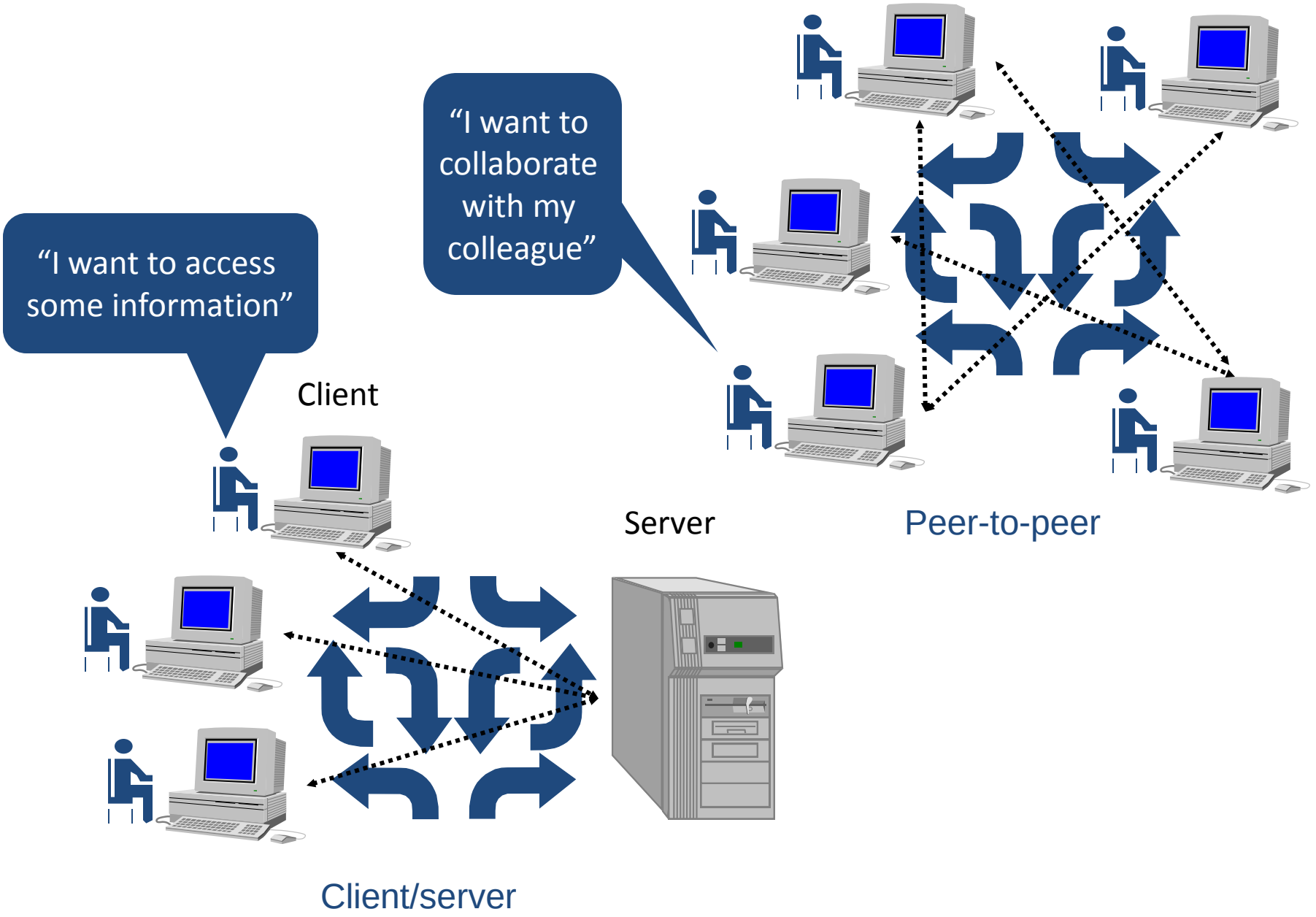
Communication models



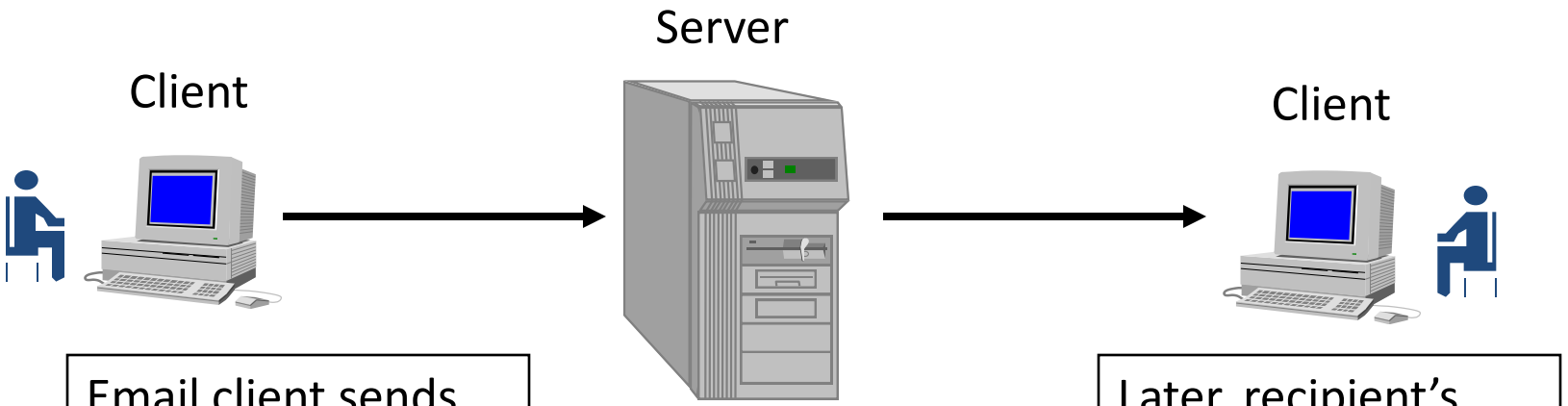

Client
host
Client
host
Server
host


"I want to collaborate with my colleague"
"I want to access some information"
Client
Server
Peer-to-peer
Client/server

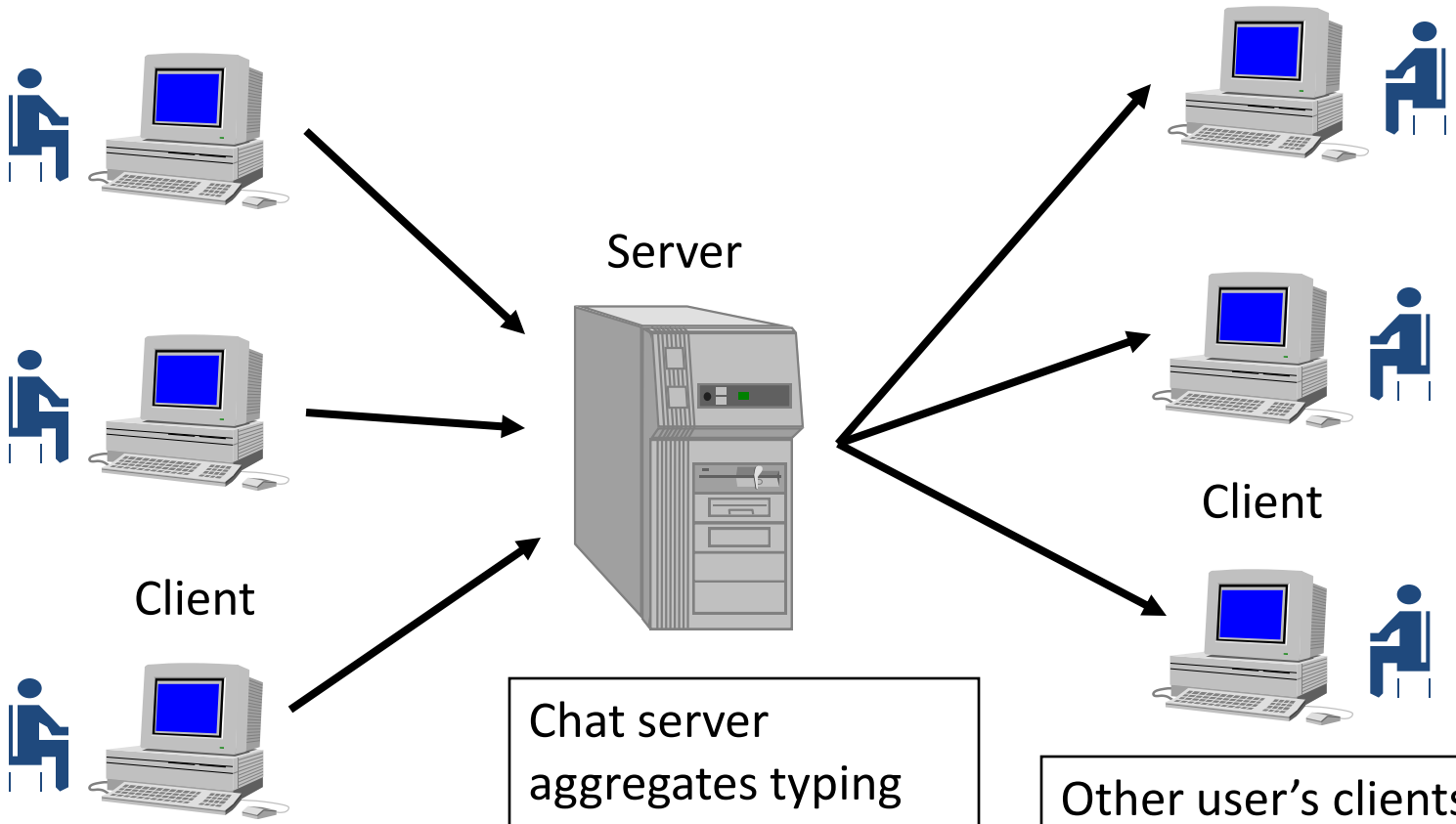
Distinctions

- Client-server
  - Asymmetric relationship
  - Client predominately makes requests, server makes replies
- Peer-to-peer
  - Symmetric relationship
  - **Peer-to-peer** (**P2P**) computing or networking is a distributed application architecture that partitions tasks or workloads between peers. Peers are equally privileged, equipotent participants in the application.

Email application



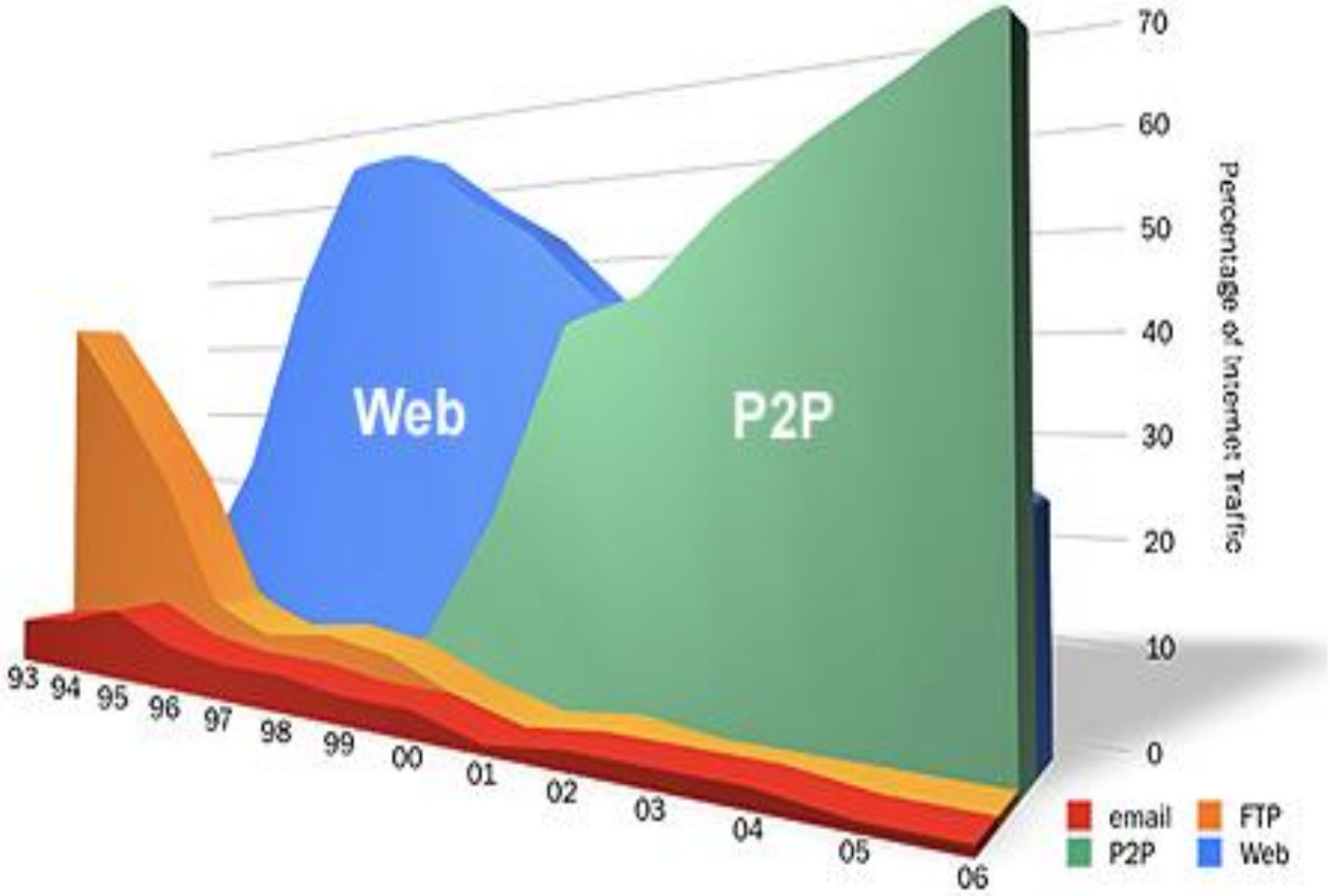
Chat application



Chat clients send user's typing to server

Chat server aggregates typing from all users and sends to all clients

Other user's clients display aggregated typing from chat server


Web
P2P
70
60
50
40
30
20
10
0
Percentage of Internet Traffic
93
94
95
96
97
98
99
00
01
02
03
04
05
06
email
FTP
P2P
Web

A Peer

- Peers are both suppliers and consumers while in the traditional client-server model, the server supplies while the client only consumes.

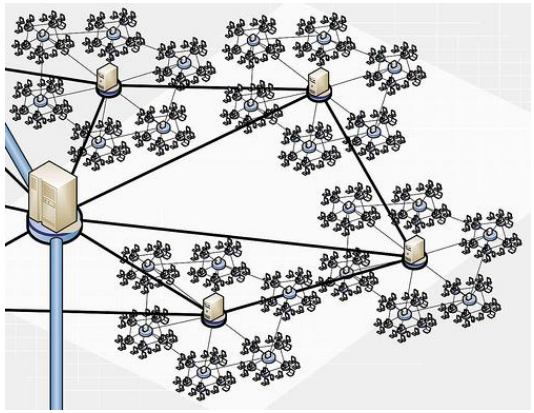
Peer-To-Peer VS Client-Server

A **server** based network	A **peer-to-peer** based network.
	

Network Architecture

- Typically they are ad-hoc networks, where the addition and the removal of nodes have no significant impact on the network.
- This distributed architecture allows peer-to-peer systems to provide enhanced scalability and service robustness.
- Often, implemented as an application layer overlay network that is placed on top of native or physical network. These are used for peer discovery and indexing.

Application Overlay Network



P2P applications : Bittorrent, **Content delivery, File-sharing networks.**

Create these overlay networks over the existing internet in order to perform indexing and peer collection functions.

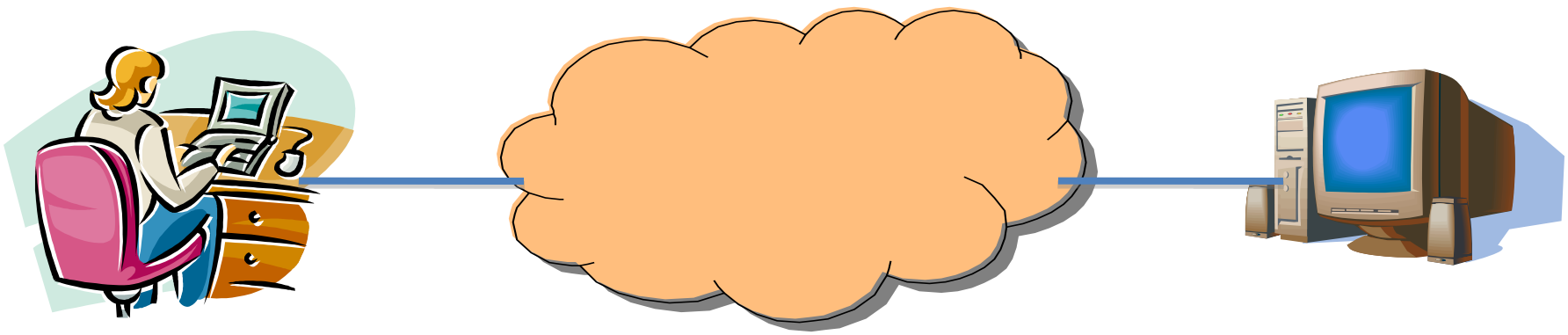
Copyright infringements
Peer-to-peer networking involves data transfer from one user to another without using an intermediate server. Companies developing P2P applications have been involved in numerous legal cases, primarily in the United States, over conflicts with copyright law.

The **Skype protocol** is a proprietary Internet telephony network based on peer-to-peer architecture, used by Skype.

Content Distribution Networks (CDNs)

A **content delivery network** or **content distribution network** (**CDN**) is a globally distributed network of proxy servers deployed in multiple data centers. The goal of a CDN is to serve content to end-users with high availability and high performance.

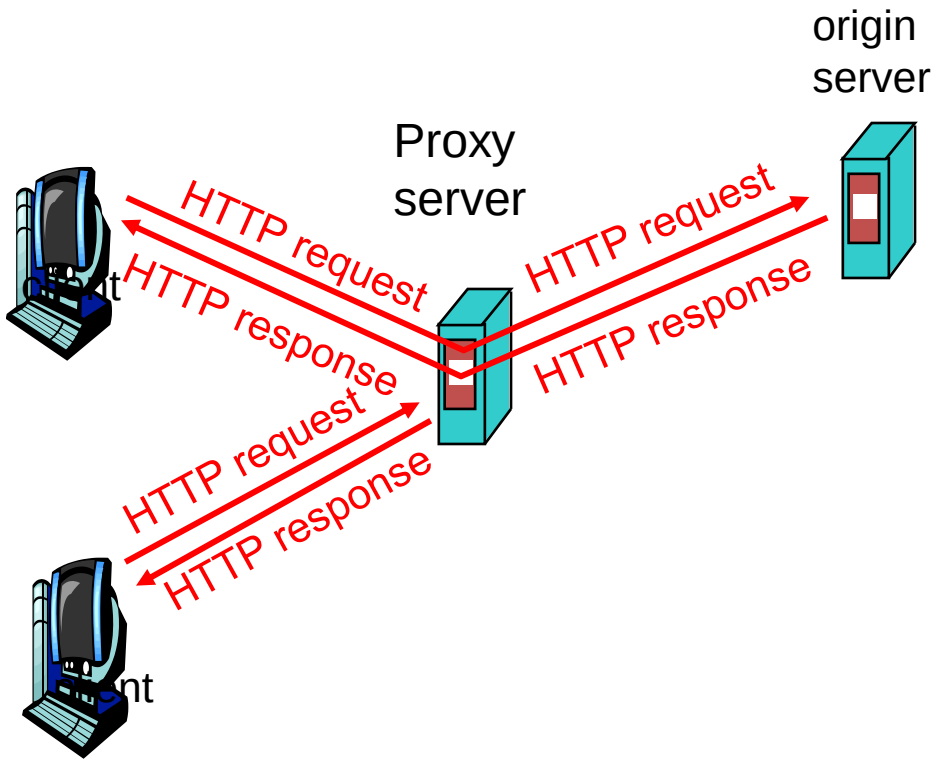
Single Server, Poor Performance



- Single server
  - Single point of failure
  - Easily overloaded
  - Far from most clients
- Popular content
  - Popular site
  - “Flash crowd” (aka “Slashdot effect”)
  - Denial of Service attack

Web Caching

Proxy Caches

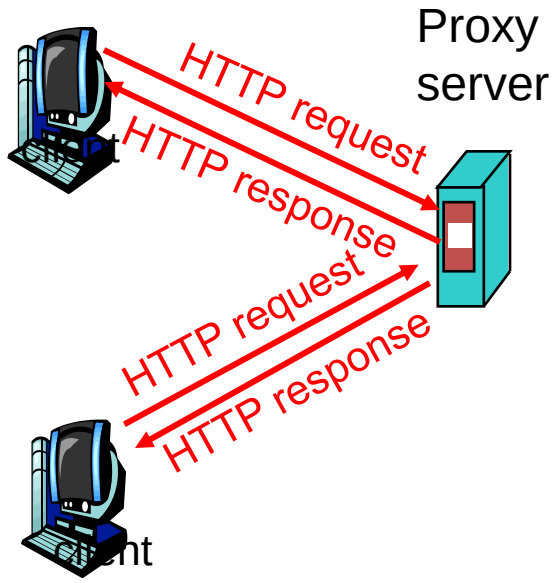


Proxy Operation

- Explicit proxy
  - Requires configuring browser
- Implicit proxy
  - Service provider deploys an “on path” proxy
  - … that intercepts and handles Web requests

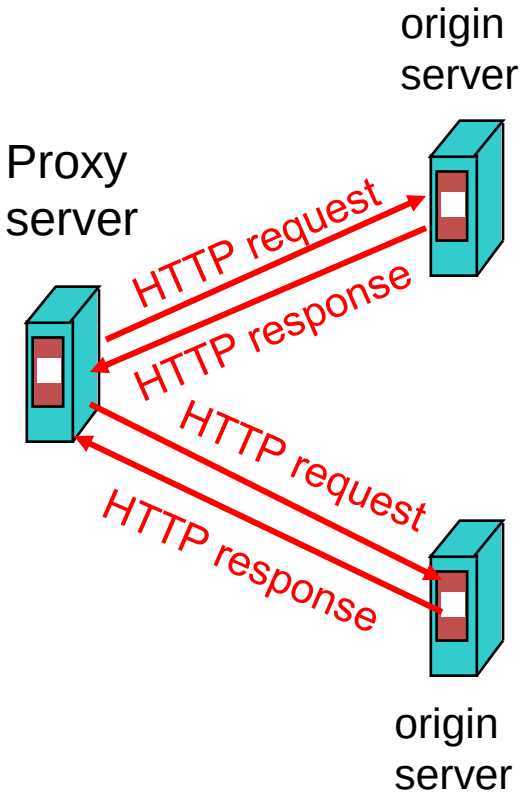
Forward Proxy

- Cache “close” to the client
  - Under administrative control of client-side AS

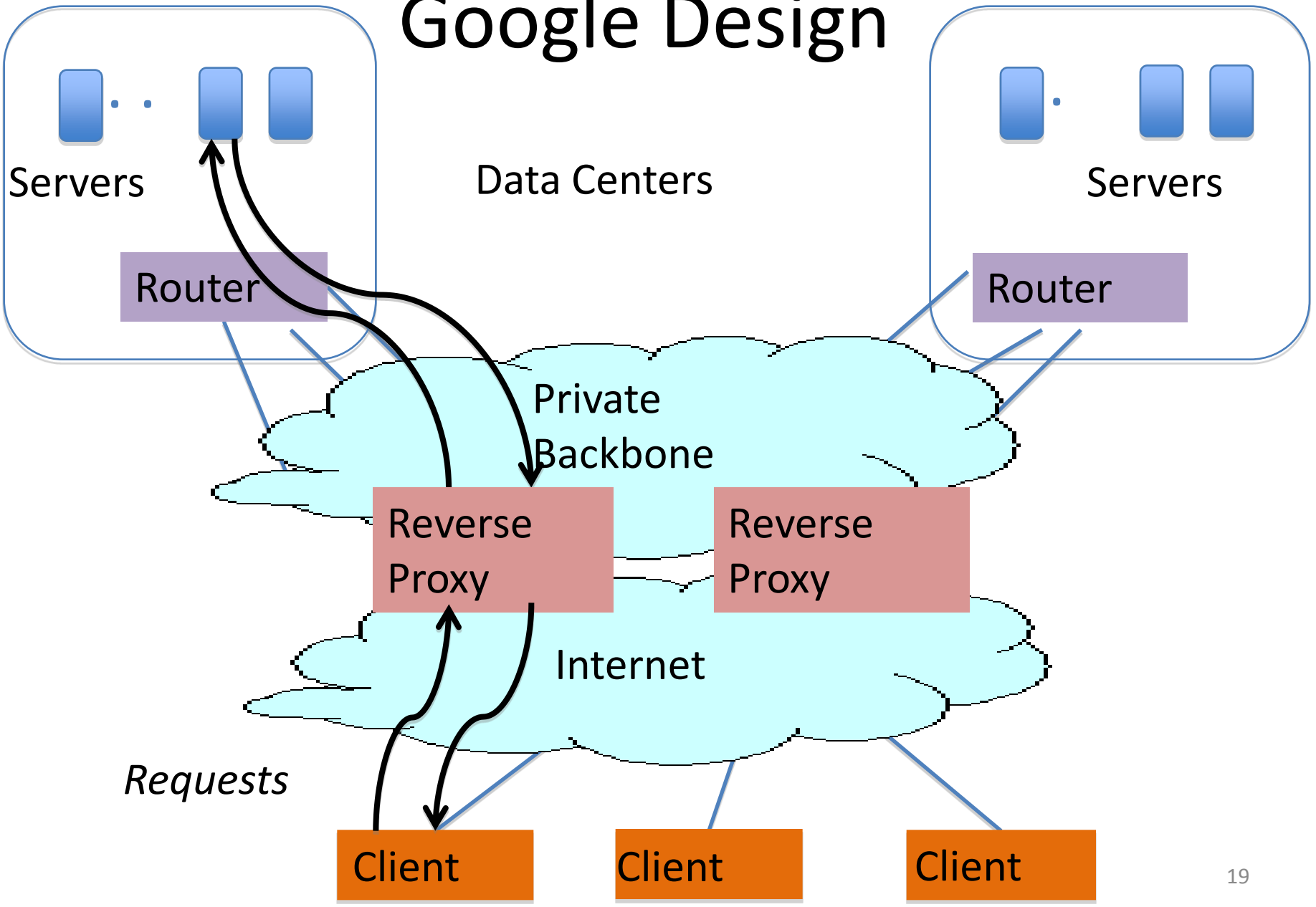


Reverse Proxy

- Cache “close” to server
  - Either by proxy run by server or in third-party content distribution network (CDN)
- Directing clients to the proxy
  - Map the site name to the IP address of the proxy



Google Design



Proxy Caches

(A) Forward (B) Reverse (C) Both (D) Neither

- Reactively replicates popular content
- Reduces origin server costs
- Reduces client ISP costs
- Intelligent load balancing between origin servers
- Offload form submissions (POSTs) and user auth
- Content reassembly or transcoding on behalf of origin
- Smaller round-trip times to clients
- Maintain persistent connections to avoid TCP setup delay (handshake, slow start)

Proxy Caches

(A) Forward (B) Reverse (C) Both (D) Neither

- Reactively replicates popular content (C)
- Reduces origin server costs (C)
- Reduces client ISP costs (A)
- Intelligent load balancing between origin servers (B)
- Offload form submissions (POSTs) and user auth (D)
- Content reassembly, transcoding on behalf of origin (C)
- Smaller round-trip times to clients (C)
- Maintain persistent connections to avoid TCP setup delay (handshake, slow start) (C)

Limitations of Web Caching

- Much content is not cacheable
  - Dynamic data: stock prices, scores, web cams
  - CGI scripts: results depend on parameters
  - Cookies: results may depend on passed data
  - SSL: encrypted data is not cacheable
  - Analytics: owner wants to measure hits
- Stale data
  - Or, overhead of refreshing the cached data

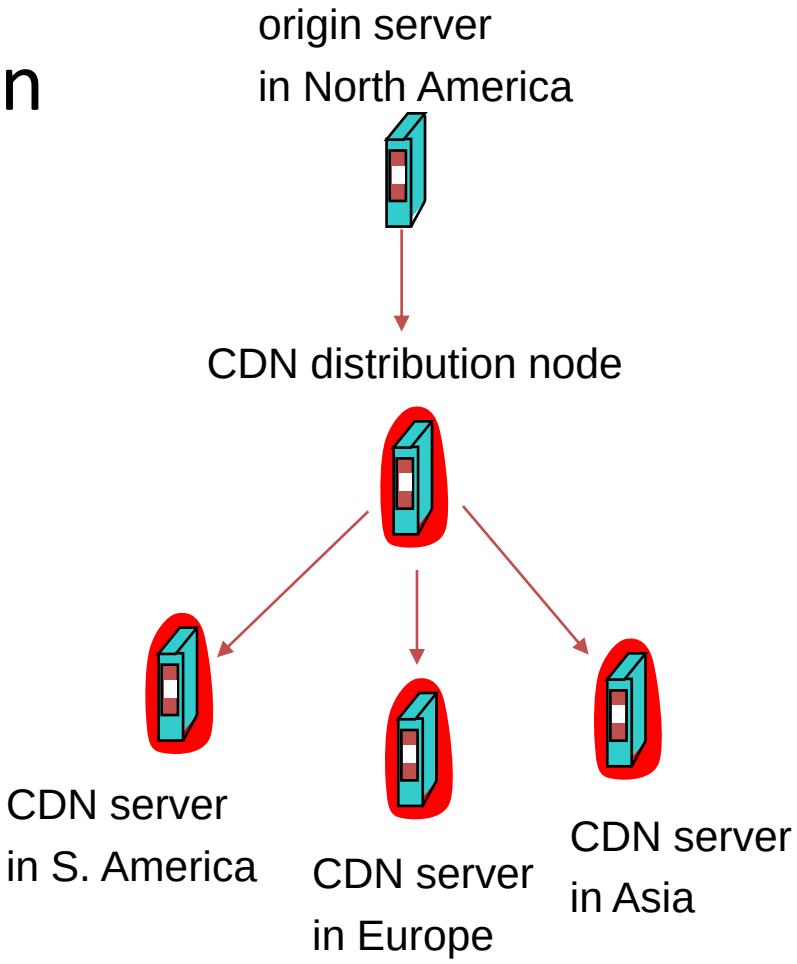
Modern HTTP Video-on-Demand

- Download “content manifest” from origin server
- List of video segments belonging to video
  - Each segment 1-2 seconds in length
  - Client can know time offset associated with each
  - Standard naming for different video resolutions and formats: e.g., 320dpi, 720dpi, 1040dpi, …
- Client downloads video segment (at certain resolution) using standard HTTP request.
  - HTTP request can be satisfied by cache: it’s a static object
- Client observes download time vs. segment duration, increases/decreases resolution if appropriate

Content Distribution Networks

Content Distribution Network

- Proactive content replication
  - Content provider (e.g., CNN) Contracts with a CDN
- CDN replicates the content
  - On many servers spread throughout the Internet
- Updating the replicas
  - Updates pushed to replicas when the content changes



Server Selection Policy

- Live server
  - For availability
- Lowest load
  - To balance load across the servers
- Closest
  - Nearest geographically, or in round-trip time
- Best performance
  - Throughput, latency, …
- Cheapest bandwidth, electricity, …

Requires continuous monitoring of liveness, load, and performance