

Content Distribution

Kai Shen

11/9/2009 CSC 257/457 - Fall 2009 1

Files/News/Software Distribution on Demand

- Content distribution:
 - Popular web directory sites like Yahoo;
 - Breaking news from CNN;
 - Online software downloads from Linux kernel repositories
- All downloading from a central content site creates performance problems:
 - Limited bandwidth at the central site;
 - Many users are far away from the central site.

11/9/2009 CSC 257/457 - Fall 2009 2

Means of Content Distribution

- By content providers, e.g., www.yahoo.com
 - replicated, distributed Internet sites
- By content consumers (or clients)
 - reduce bandwidth consumption of an organization through caching - Web caching
- By a third-party
 - content distribution network

11/9/2009 CSC 257/457 - Fall 2009 3

Replicated Internet Sites

- Content providers build a set of distributed sites on the Internet.
- DNS answers the IP address of the site close to the client

Is DNS caching a problem?

limited scale!

11/9/2009 CSC 257/457 - Fall 2009 4

Determining Nearby Content Server

- We create a "map", indicating distances between machines (or local networks) and content servers
- When a DNS query arrives at authoritative DNS server:
 - server determines network from which query originates
 - uses "map" to determine a nearby content server
- How to build the map?
 - offline pings from content servers to potential sources
 - just-in-time pings from content servers to the source
- What if the source does not respond to pings?

11/9/2009

CSC 257/457 - Fall 2009

5

A Related Question

- Scalable estimation of node-to-node relative proximity
 - A specific problem: given any node X, find a few nearby nodes in a group of 10,000 nodes (N_1, \dots, N_{10000}).
 - Too costly to measure the distance from X to 10,000 nodes.
- Use a landmark L:
 - Measure the distances from N_1, \dots, N_{10000} to L offline.
 - Given node X, we measure the distance from X to L.
 - For node N, if $\text{distance}(X,L)$ differs significantly from $\text{distance}(N,L)$, X and N are likely far apart.
 - So we prefer nodes whose distances to L are similar to $\text{distance}(X,L)$.
- Use multiple landmarks

11/9/2009

CSC 257/457 - Fall 2009

6

Another Related Question

- How to build a map indicating the geographical location of machines or local networks?

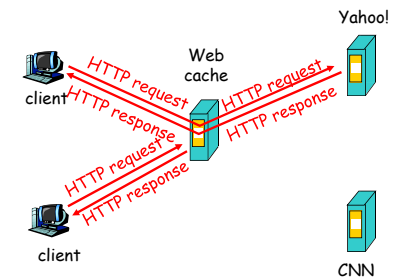
11/9/2009

CSC 257/457 - Fall 2009

7

Web caches (proxy server)

- Cache is installed and shared by users (university, company, residential ISP)
- **Goal:** satisfy client requests without involving the original server.
- user sets browser: accesses via Web cache
- browser sends all HTTP requests to Web cache
 - object in cache: cache returns object
 - otherwise cache requests object from the original server, then returns object to client



11/9/2009

CSC 257/457 - Fall 2009

8

A Quantitative Study: the Base Case

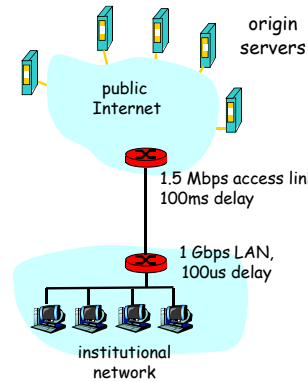
Assumptions

- 1.5Mbps wide-area access link, 1Gbps local-area network
- wide-area delay 100 milliseconds, local-area delay 100 microseconds

Performance

- effective bandwidth to wide-area content - 1.5Mbps
- average access delay = Internet delay + LAN delay = 100.1 milliseconds

Upgrade the access link is costly, doesn't help the access delay.



11/9/2009

CSC 257/457 - Fall 2009

9

A Quantitative Study: Web Caching

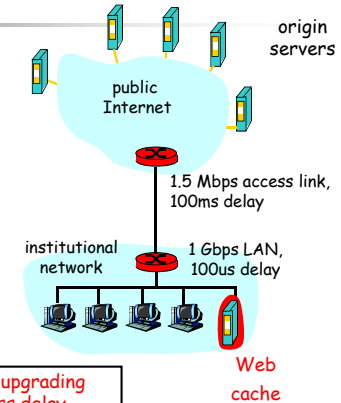
Install a Web cache

- suppose hit rate is 50%

Performance

- effective bandwidth to wide-area content = 3.0Mbps
- average access delay = cache miss delay * 50% + cache hit delay * 50% = $100.2 * 50\% + 0.1 * 50\%$ = 50.15 milliseconds

Not only higher bandwidth without upgrading access link, but also shorter access delay.



11/9/2009

CSC 257/457 - Fall 2009

10

Cooperating Web Caching

- Caching cooperation
 - Several caches (often nearby) cooperate with each other
 - Benefits: more cache hits; less load on original servers
- Probing all caches at each access is too costly
- Challenge:** knowing (roughly) each other's cache content
 - Each cache periodically broadcast its content to other caches (what data structure?)

11/9/2009

CSC 257/457 - Fall 2009

11

Cache Content Staleness

- Is the cached page up-to-date?
 - using If-modified-since HTTP header.
 - removing pages that are too old.
- Content providers lose direct control of content

11/9/2009

CSC 257/457 - Fall 2009

12

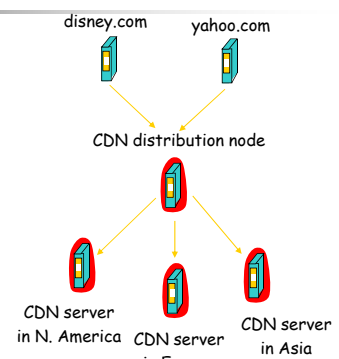
Mining of Cache Logs

- Web cache logs contain a wealth of information
 - List of who accessed what at when
- User privacy
- Aggregate statistics
 - Most popular web objects, distribution of web object popularity
 - General user access models (think time between accesses, pattern of page browsing sequence, ...)
 - ...

11/9/2009 CSC 257/457 - Fall 2009 13

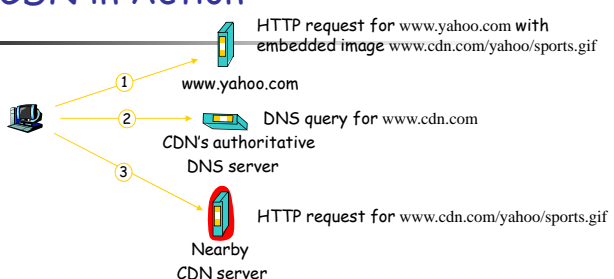
Content Distribution Network

- Motivation:
 - limited scale for replication by content providers
 - lose control of content by caching
- Content distribution network:
 - hundreds of CDN servers throughout Internet
 - content providers' content replicated on CDN servers
- Goals:
 - transparent to clients
 - content providers retain primary control



11/9/2009 CSC 257/457 - Fall 2009 14

CDN in Action



HTTP request for www.yahoo.com with embedded image www.cdn.com/yahoo/sports.gif

- 1 www.yahoo.com
- 2 DNS query for www.cdn.com
CDN's authoritative DNS server
- 3 HTTP request for www.cdn.com/yahoo/sports.gif
Nearby CDN server

Content provider

- www.yahoo.com
- Replaces www.yahoo.com/sports.gif with www.cdn.com/yahoo/sports.gif

CDN company


- replicates content at CDN servers
- uses its authoritative DNS server to redirect requests to a nearby CDN server

11/9/2009 CSC 257/457 - Fall 2009 15

Summary

- Three means:
 - By content providers - replicated, distributed Internet sites
 - By content consumers (or clients) - Web caching
 - By a third-party - content distribution network
- Differentiate them on the following:
 - Scalability
 - Content staleness
 - Transparency to clients
 - Reliance on the Domain Name System

11/9/2009 CSC 257/457 - Fall 2009 16



BitTorrent

- Peer-to-peer distribution of large files
- A large file is divided into chunks
- Chunks are distributed in peer-to-peer fashion
 - High scalability
- Rarest-first download
 - High data availability

11/9/2009 CSC 257/457 - Fall 2009 17