Content Distribution

Kai Shen
Dept. of Computer Science, University of Rochester

Content Distribution:
- Popular web sites like Yahoo;
- Breaking news from CNN;
- Microsoft is distributing a new version of Windows online.

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.

Means of Content Distribution

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

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

HTTP request

DNS query

Yahoo authoritative DNS server

IP address of S. America site

a client at S. America

Yahoo at North America

Yahoo at South America

Yahoo at Asia

Yahoo at Europe
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?

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 (N₁, ..., N₁₀₀₀₀).
  - Too costly to measure the distance from X to 10,000 nodes.

- Use a landmark L:
  - Measure the distance from N₁, ..., N₁₀₀₀₀ to L offline.
  - Given node X, we measure the distance from X to L.
  - For node N, if distance(X,L) differs significantly from distance(N,L), X and N are likely far apart.
  - So we prefer nodes whose distances to L are similar to distance(X,L).

- Use multiple landmarks

Another Related Question

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

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

- [Diagram of client, Web cache, and response flows]
A Quantitative Study: the Base Case

Assumptions
- 1.5Mbps wide-area access link, 10Mbps local-area network
- wide-area delay 2 seconds, local-area delay 10 milliseconds

Performance
- effective bandwidth to wide-area content = 1.5Mbps
- average access delay = Internet delay + LAN delay = 2.01 seconds

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

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%
  = 2.02 * 50% + 0.01 * 50%
  = 1.015 seconds

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

More on Web Caching

- Content providers lose direct control of content
- Is the cached page up-to-date?
  - using If-modified-since HTTP header.
  - removing pages that are too old.
- Cooperative web caching
  - A number of nearby caches (each from a UR department) cooperate with each other.
  - Key: knowing each other’s cache content.

Content Distribution Network

Motivation: limited scale for replication by content providers; lose control of content by caching
- CDN is a service: content providers are the CDN customers.
- CDN company installs hundreds of CDN servers throughout Internet
- CDN replicates its customers’ content in CDN servers. When provider updates content, CDN updates servers
### CDN in Action

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

2. **DNS query for** www.cdn.com

3. **CDNs authoritative DNS server**

4. **HTTP request for** www.cdn.com/yahoo/sports.gif

**Content provider**
- www.yahoo.com

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

### 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