CONTENT-DISTRIBUTION NETWORKS

George Porter
March 6 and 8, 2019

ATTRIBUTION

• These slides are released under an Attribution-NonCommercial-ShareAlike 3.0 Unported (CC BY-NC-SA 3.0) Creative Commons license
• These slides incorporate material from:
  • Kyle Jamieson, Princeton University (also under a CC BY-NC-SA 3.0 Creative Commons license)
  • David Choffnes, Northeastern University
ANNOUNCEMENTS

OUTLINE

1. Web caching
2. Content-distribution networks
   • Featuring Akamai
   • Overlay networks
WEB CACHING

- Many clients transfer the same information
  - Generates redundant server and network load
  - Also, clients may experience high latency

WHY WEB CACHING?

- Motivation for placing content closer to client:
  - User gets better response time
    - Content providers get happier users
  - Network gets reduced load
- Why does caching work? Exploits locality of reference
- How well does caching work?
  - Very well, up to a limit
  - Large overlap in content
  - But many unique requests
CACHING WITH REVERSE PROXIES

- Cache data close to origin server → decrease server load
- Typically done by content providers
- Client thinks it is talking to the origin server (the server with content)
- Does not work for dynamic content

CACHING WITH FORWARD PROXIES

- Cache close to clients → less network traffic, less latency
- Typically done by ISPs or corporate LANs
- Client configured to send HTTP requests to forward proxy
- Reduces traffic on ISP-1’s access link, origin server, and backbone ISP
CACHING & LOAD-BALANCING: OUTSTANDING PROBLEMS

- Problem ca. 2002: How to reliably deliver large amounts of content to users worldwide?
- Popular event: “Flash crowds” overwhelm (replicated) web server, access link, or back-end database infrastructure
- More rich content: audio, video, photos

- Web caching: Diversity causes low cache hit rates (25–40%)

OUTLINE

1. Web caching
2. Content-distribution networks
   - Featuring Akamai
   - Overlay networks
CONTENT DISTRIBUTION NETWORKS

- **Proactive content replication**
  - Content provider (e.g. CNN) pushes content out from its own *origin server*
- **CDN replicates** the content
  - On many servers spread throughout the Internet
- **Updating the replicas**
  - Updates pushed to replicas when the content changes

REPLICA SELECTION: GOALS

- **Live server**
  - For availability
- **Lowest load**
  - To balance load across the servers
- **Closest**
  - Nearest geographically, or in round-trip time
- **Best performance**
  - Throughput, latency, reliability...
AKAMAI

- **Deployment**
  - 147K+ servers, 1200+ networks, 650+ cities, 92 countries
  - highly hierarchical, caching depends on popularity
  - 4 yr depreciation of servers
  - Many servers inside ISPs, who are thrilled to have them
  - Deployed inside 100 new networks in last few years

- **Customers**
  - 250K+ domains: all top 60 eCommerce sites, all top 30 M&E companies, 9 of 10 top banks, 13 of top 15 auto manufacturers

- **Overall stats**
  - 5+ terabits/second, 30+ million hits/second, 2+ trillion deliveries/day, 100+ PB/day, 10+ million concurrent streams
  - 15-30% of Web traffic

CIRCA 2007 OR SO

- 30000+ Servers
- 1450+ POPs
- 950+ Networks
- 67+ Countries

Current Installations
HOW AKAMAI USES DNS

1. **DNS TLD server**
2. **Akamai global DNS server**
3. **Akamai regional DNS server**
4. **Nearby Akamai cluster**
5. **End user**

1. **cnn.com (content provider)**
2. **GET index.html**
3. **cache.cnn.com/foo.jpg**

15

16
HOW AKAMAI USES DNS

HTTP

cnn.com (content provider)

DNS TLD server

DNS lookup
g.akamai.net

Akamai global
DNS server

Akamai regional
DNS server

Nearby
Akamai
cluster

ALIAS
a73.g.akamai.net

End user

Address
1.2.3.4

Nearby
Akamai
cluster
HOW AKAMAI USES DNS

End user GET /foo.jpg Host: cache.cnn.com

GET foo.jpg

10

19
HOW AKAMAI USES DNS

HTTP

cnn.com (content provider)

DNS TLD server

Akamai global DNS server

Akamai regional DNS server

Nearby Akamai cluster

End user

HOW AKAMAI WORKS: CACHE HIT

HTTP

cnn.com (content provider)

DNS TLD server

Akamai global DNS server

Akamai regional DNS server

Nearby Akamai cluster

End user
OPTIMIZING PERFORMANCE: NETWORK

- There are good solutions to server load and content
  - What about network performance?
- Key challenges for network performance
  - Measuring paths is hard
    - Traceroute gives us only the forward path
    - Shortest path != best path
  - RTT estimation is hard
    - Variable network conditions
    - May not represent end-to-end performance
  - No access to client-perceived performance

Example approximation strategies

- Geographic mapping
  - Hard to map IP to location
  - Internet paths do not take shortest distance
- Active measurement
  - Ping from all replicas to all routable prefixes
  - $56B \times 100 \text{ servers} \times 500k \text{ prefixes} = 500+\text{MB of traffic per round}$
- Passive measurement
  - Send fraction of clients to different servers, observe performance
  - Downside: Some clients get bad performance
MAPPING SYSTEM

• Equivalence classes of IP addresses
  • IP addresses experiencing similar performance
  • Quantify how well they connect to each other

• Collect and combine measurements
  • Ping, traceroute, BGP routes, server logs
    • e.g., over 100 TB of logs per days
  • Network latency, loss, throughput, and connectivity

ROUTING CLIENT REQUESTS WITH THE MAP

• Map each IP class to a preferred server cluster
  • Based on performance, cluster health, etc.
  • Updated roughly every minute
    • Short, 60-sec DNS TTLs in Akamai regional DNS accomplish this

• Map client request to a server in the cluster
  • Load balancer selects a specific server
  • e.g., to maximize the cache hit rate
ADAPTING TO FAILURES

• Failing hard drive on a server
  • Suspends after finishing “in progress” requests

• Failed server
  • Another server takes over for the IP address
  • Low-level map updated quickly (load balancer)

• Failed cluster, or network path
  • High-level map updated quickly (ping/traceroute)

TAKE-AWAY POINTS: CDNS

• Content distribution is hard
  • Many, diverse, changing objects
  • Clients distributed all over the world

• Moving content to the client is key
  • Reduces latency, improves throughput, reliability

• Content distribution solutions evolved:
  • Load balancing, reactive caching, to
  • Proactive content distribution networks
OUTLINE

1. Web caching
2. Content-distribution networks
   • Featuring Akamai
   • Overlay networks

ABSTRACTIONS AND OVERLAYS

Overlay Network

Physical Network
OVERLAYS FOR ROUTING: WHY?

• Triangle inequality doesn’t hold in networks!
OVERLAY NETWORKS FOR ROUTING

- Underlying network
  - Internet connectivity (IP Routing)

OVERLAY NETWORKS

- Potential overlay connectivity
  - SF as root
OVERLAY NETWORKS

• Determine edge weights
  • E.g., bandwidth, latency

OVERLAY NETWORKS

• Build overlay connectivity
  • An application-layer distribution tree