Content Distribution Networks

Lecture 16

COS 461: Computer Networks

Kyle Jamieson
Content Distribution Network (CDN)

• **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**
  - Reactive by TTL or 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
Server Selection Mechanism

- **Application**
  - HTTP redirection

- **Advantages**
  - Fine-grain control
  - Selection based on client IP address

- **Disadvantages**
  - Extra round-trips for TCP connection to server
  - Overhead on the server
Server Selection Mechanism

• Routing
  – Anycast routing

• Advantages
  – No extra round trips
  – Route to nearby server

• Disadvantages
  – Does not consider network or server load
  – Different packets may go to different servers
  – Used only for simple request-response apps
Server Selection Mechanism

- **Naming**
  - DNS-based server selection

- **Advantages**
  - Avoid TCP set-up delay
  - DNS caching reduces overhead
  - Relatively fine control

- **Disadvantage**
  - Based on IP address of local DNS server
  - “Hidden load” effect
  - DNS TTL limits adaptation
How Akamai Works
Akamai Statistics

• **Distributed servers**
  – Servers: ~275,000
  – Networks: 1,500
  – Countries: 136

• **Many customers**
  – 50% of Fortune Global 500

• **Network**
  – Up to 50 Tbps daily
  – 2019 Cricket World Cup: 25.3M concurrent viewers
  – 85% of Internet is one network hop from Akamai servers
How Akamai Uses DNS

cnn.com (content provider) → DNS root server

GET
index.html

HTTP

cache.cnn.com/foo.jpg

1

2

Akamai global DNS server

Akamai regional DNS server

End user

Nearby Akamai cluster

Akamai cluster
How Akamai Uses DNS

cnn.com (content provider)

DNS TLD server

DNS lookup
cache.cnn.com

End user

1

2

3

4

ALIAS:
g.akamai.net

Nearby
Akamai
cluster

Akamai
global
DNS server

Akamai
regional
DNS server

Akamai
cluster
How Akamai Uses DNS

cnn.com (content provider)  DNS TLD server

1. DNS lookup
2. g.akamai.net
3. Akamai global DNS server
4. Akamai regional DNS server
5. Nearby Akamai cluster
6. End user

ALIAS
a73.g.akamai.net

Akamai cluster
How Akamai Uses DNS

cnn.com (content provider) → DNS TLD server

1. cnn.com queries DNS TLD server for a73.g.akamai.net.
2. DNS TLD server returns the IP address 1.2.3.4.
3. End user queries DNS a73.g.akamai.net.
4. DNS a73.g.akamai.net returns the IP address 1.2.3.4.
5. End user connects to 1.2.3.4.
How Akamai Uses DNS

- cnn.com (content provider)
- DNS TLD server
- Akamai global DNS server
- Akamai regional DNS server
- Nearby Akamai cluster

Flow:
1. End user requests GET /foo.jpg
2. Host: cache.cnn.com
How Akamai Uses DNS

cnn.com (content provider)  DNS TLD server

GET foo.jpg

End user

GET /foo.jpg
Host: cache.cnn.com
How Akamai Uses DNS

cnn.com (content provider) → DNS TLD server

1. cnn.com sends a DNS query to the top-level domain (TLD) server.

2. The TLD server forwards the query to an Akamai global DNS server.

3. The global DNS server routes the query to an Akamai regional DNS server.

4. The regional DNS server resolves the domain to an IP address.

5. The IP address is returned to the TLD server.

6. The TLD server forwards the IP address to Akamai global DNS server.

7. The global DNS server routes the IP address to the nearest Akamai cluster.

8. The nearest Akamai cluster returns the content to the end user.

End user → CNN

Nearby Akamai cluster
How Akamai Works: Cache Hit

cnn.com (content provider)  
DNS TLD server

Akamai global DNS server

Akamai regional DNS server

Nearby Akamai cluster

End user

1 2 3 4 5 6
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, and connectivity
Routing Client Requests within 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
Selecting server inside cluster

- **“Consistent hashing”**
  - content_key = hash(URL) mod N
  - node_key = hash(server ID) mod N
  - Content belongs to server’s node_key is “closest” to URL’s content_key
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
- Failed cluster or network path
  - High-level map updated quickly
- Failed path to customer’s origin server
  - Route packets through an intermediate node
Akamai Transport Optimizations

• Bad Internet routes
  – Overlay routing through an intermediate server

• Packet loss
  – Sending redundant data over multiple paths

• TCP connection set-up/teardown
  – Pools of persistent connections

• TCP congestion window and round-trip time
  – Estimates based on network latency measurements
Akamai Application Optimizations

• Slow download of embedded objects
  – Prefetch when HTML page is requested

• Large objects
  – Content compression

• Slow applications
  – Moving applications to edge servers
  – E.g., content aggregation and transformation
  – E.g., static databases (e.g., product catalogs)
Conclusion

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

• **Moving content towards client is key**
  – Reduces latency, improves throughput, reliability

• **Contribution distribution solutions evolved**
  – Reactive caching, load balancing, to
  – Proactive content distribution networks