Content Distribution Networks

COS 518: Advanced Computer Systems
Lecture 18
Mike Freedman

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, ...

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

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 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

A DNS lookup traverses DNS hierarchy

DNS caching

- **Performing all these queries takes time**
  - And all this before actual communication takes place

- **Caching can greatly reduce overhead**
  - Top-level servers very rarely change, popular sites visited often
  - Local DNS server often has information cached

- **How DNS caching works**
  - All DNS servers cache responses to queries
  - Responses include a time-to-live (TTL) field, akin to cache expiry
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 Uses DNS

1. CNN.com (content provider) sends a GET request for `index.html` to the DNS root server.
2. The root server forwards the request to the TLD server for `cnn.com`.
4. The request is sent to the nearest Akamai cluster.
5. The request is served from the cache at `cache.cnn.com`.
6. The end user receives the requested content.
How Akamai Uses DNS

cnn.com (content provider)

DNS TLD server

1
2
3
4
5
6

DNS lookup

g.akamai.net

a73.g.akamai.net

Akamai global
DNS server

Akamai regional
DNS server

Nearby
Akamai
cluster

end user

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

How Akamai Uses DNS

cnn.com (content provider)

DNS TLD server

1
2
3
4
5
6

DNS a73.g.akamai.net

Address
1.2.3.4

Akamai global
DNS server

Akamai regional
DNS server

Nearby
Akamai
cluster

end user

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

Mapping System

- Map each IP class to a preferred server cluster
  - Based on performance, cluster health, etc.
  - Updated roughly every minute

- 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

• Failed cluster
  — High-level map updated quickly

• Failed path to customer’s origin server
  — Route packets through an intermediate node

Conclusion

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

• Contribution distribution solutions
  — Reactive caching
  — Proactive content distribution networks