CSCI-351
Data communication and Networks

Lecture 14: Content Delivery Networks (Over 1 billion served … each day)

The slide is built with the help of Prof. Alan Mislove, Christo Wilson, and David Choffnes's class

Outline

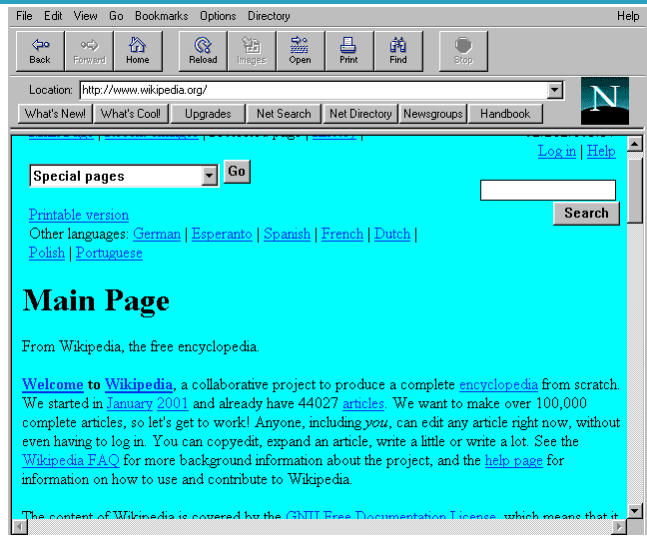
- **Motivation**
- CDN basics
- Prominent example: Akamai

Content in today's Internet

- Most flows are HTTP….
  - Web is at least 52% of traffic (as of early 2000), however..
- HTTP uses TCP, so it will
  - Be ACK clocked
  - For Web, likely never leave slow start
    - QUIC?
- Is the Internet designed for this common case?
  - Why?

Evolution of Serving Web Content

- In the beginning…
  - …there was a single server
  - Probably located in a closet
  - And it probably served blinking text



- Issues with this model
  - Site reliability
    - Unplugging cable, hardware failure, natural disaster
  - Scalability
    - Flash crowds (aka Slashdotting)

Replicated Web service

- Use multiple servers

- Advantages
  - Better scalability
  - Better reliability

- Disadvantages
  - How do you decide which server to use?
  - How to do synchronize state among servers?

Load Balancers

- Device that multiplexes requests across a collection of servers
  - All servers share one public IP
  - Balancer transparently directs requests to different servers
- How should the balancer assign clients to se
  - Random / round-robin
    - When is this a good idea?
  - Load-based
    - When might this fail?
- Challenges
  - Scalability (must support traffic for n hosts)
  - State (must keep track of previous decisions)

Load balancing: Are we done?

- Advantages
  - Allows scaling of hardware independent of IPs
  - Relatively easy to maintain
- Disadvantages
  - Expensive
  - Still a single point of failure
  - Location!

Where do we place the load balancer for Wikipedia?

Popping up: HTTP performance

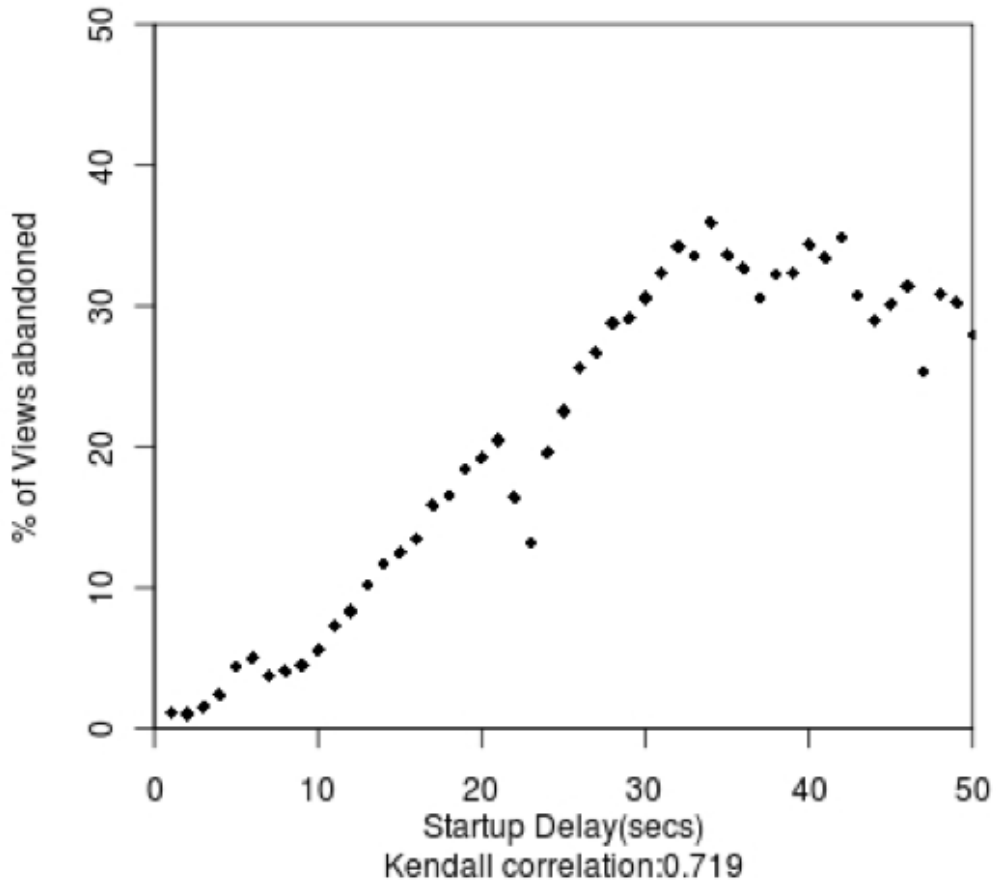
- For Web pages
  - RTT matters most
  - Where should the server go?
- For video
  - Available bandwidth matters most
  - Where should the server go?

- Is there one location that is best for everyone?

Server placement

Why speed matters

- Impact on user experience
  - Users navigating away from pages
  - Video startup delay



Why speed matters

- Impact on user experience
  - Users navigating away from pages
  - Video startup delay
- Impact on revenue
  - Amazon: increased revenue 1% for every 100ms reduction in PLT*
- Ping from ROC to LAX: ~100ms

Strawman solution: Web caches

- ISP uses a middlebox that caches Web content
  - Better performance – content is closer to users
  - Lower cost – content traverses network boundary once
  - Does this solve the problem?

- No!
  - Size of all Web content is too large
    - Zipf distribution limits cache hit rate
  - Web content is **dynamic** and **customized**
    - Can't cache banking content
    - What does it mean to cache search results?

Outline

- Motivation
- **CDN basics**
- Prominent example: Akamai

What is a CDN?

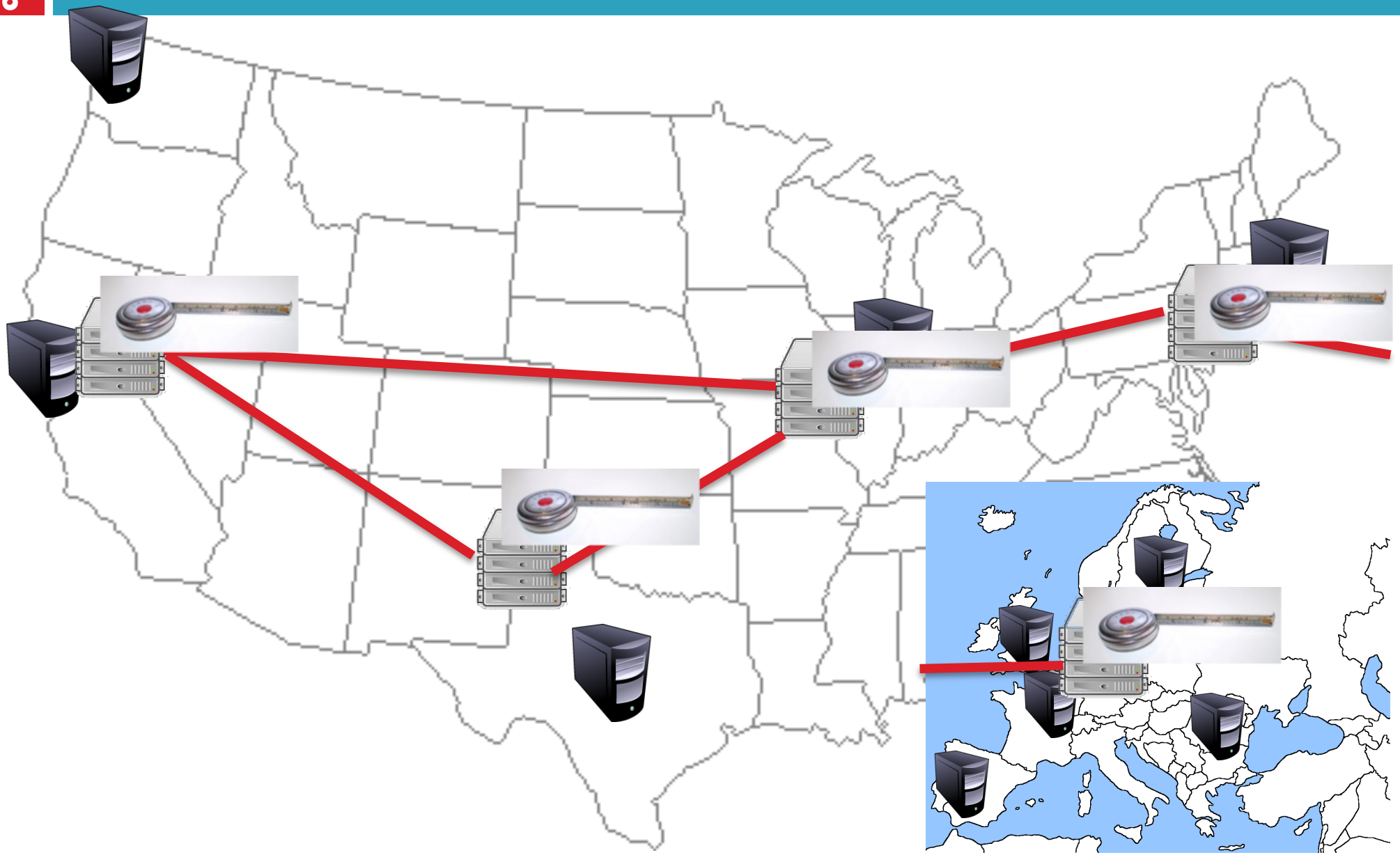
- Content Delivery Network
  - Also sometimes called Content Distribution Network
  - At least half of the world's bits are delivered by a CDN
    - Probably closer to 80/90%

- Primary Goals
  - Create replicas of content throughout the Internet
  - Ensure that replicas are always available
  - Directly clients to replicas that will give good performance

Key Components of a CDN

- Distributed servers
  - Usually located inside of other ISPs
    - Why?
- High-speed network connecting them
- Clients
  - Can be located anywhere in the world
  - They want fast Web performance
- Glue
  - Something that binds clients to “nearby” replica servers

Key CDN Components



Examples of CDNs

- Akamai
  - 147K+ servers, 1200+ networks, 92 countries (in early 2010)
  - 240K+ servers, 1700+ networks, 130 countries (now)
- Cloudflare, Limelight, Edgecast, and others web service providers (e.g., Google, Facebook)
  - Advice…?

Inside a CDN

- Servers are deployed in clusters for reliability
  - Some may be offline
    - Could be due to failure
    - Also could be "suspended" (e.g., to save power or for upgrade)
- Could be multiple clusters per location (e.g., in multiple racks)
- Server locations
  - Well-connected points of presence (PoPs)
  - Inside of ISPs

Mapping clients to servers (1)

- CDNs need a way to send clients to the "best" server
  - The best server can change over time
  - And this depends on client location, network conditions, server load, …
  - What existing technology can we use for this?

- URL Rewriting
  - Modifies the URL of specific content
  - netflix.com/movie1 to a17.akamai.com/movie1
  - Requires content modification in the origin websites
    - But it allows fine-grained controls

Mapping clients to servers (2)

- DNS-based redirection (most widely used)
  - DNS Server approaches
    - Clients request www.foo.com
    - DNS server directs client to one or more IPs based on request IP
    - Use short TTL to limit the effect of caching
  - CNAME approaches
    - Clients request www.foo.com
    - Returned record is foo.com CNAME a18.akamai.com

CDN redirection example


```
tijay$ dig www.fox.com


;; ANSWER SECTION:
www.fox.com.                    510     IN      CNAME   www.fox-rma.com.edgesuite.net.
www.fox-rma.com.edgesuite.net. 5139 IN  CNAME   a2047.w7.akamai.net.
a2047.w7.akamai.net.            4       IN      A       23.62.96.128
a2047.w7.akamai.net.            4       IN      A       23.62.96.144
a2047.w7.akamai.net.            4       IN      A       23.62.96.193
a2047.w7.akamai.net.            4       IN      A       23.62.96.162
a2047.w7.akamai.net.            4       IN      A       23.62.96.185
a2047.w7.akamai.net.            4       IN      A       23.62.96.154
a2047.w7.akamai.net.            4       IN      A       23.62.96.169
a2047.w7.akamai.net.            4       IN      A       23.62.96.152
a2047.w7.akamai.net.            4       IN      A       23.62.96.186
```

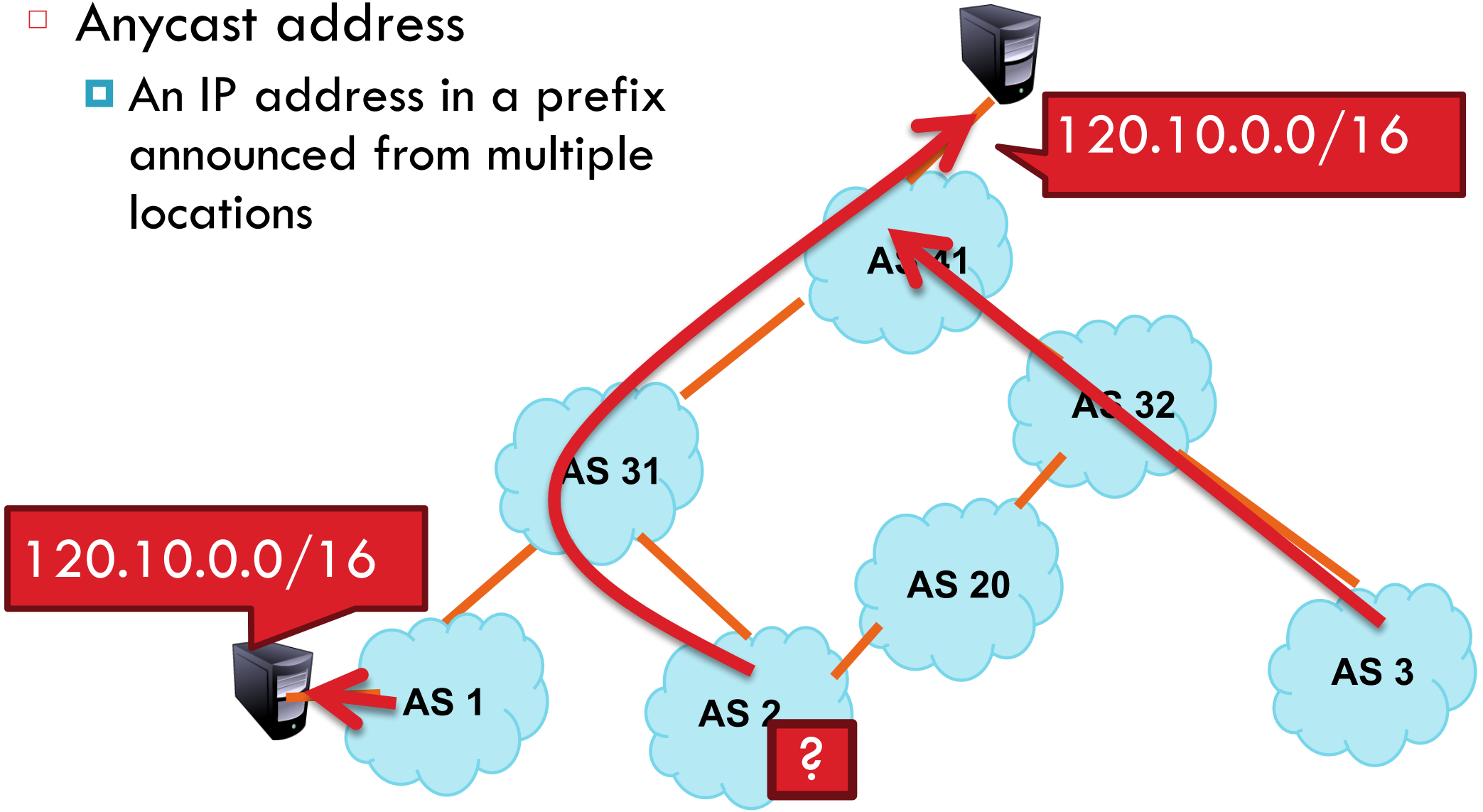
DNS Redirection Considerations

- Advantages
  - Uses existing, scalable DNS infrastructure
  - URLs can stay essentially the same
  - TTLs can control "freshness"

- Limitations
  - DNS servers see only the DNS server IP
    - Assumes that client and DNS server are close. Is this accurate?
  - Content owner must give up control
  - Unicast addresses can limit reliability; the client will connect to "one" IP

CDN Using Anycast

- Anycast address
  - An IP address in a prefix announced from multiple locations



Anycasting Considerations

- Why do anycast?
  - Simplifies network management
    - Replica servers can be in the same network domain
  - Uses best BGP path

- Disadvantages
  - BGP path may not be optimal
  - Stateful services can be complicated

Optimizing Performance

Key goal

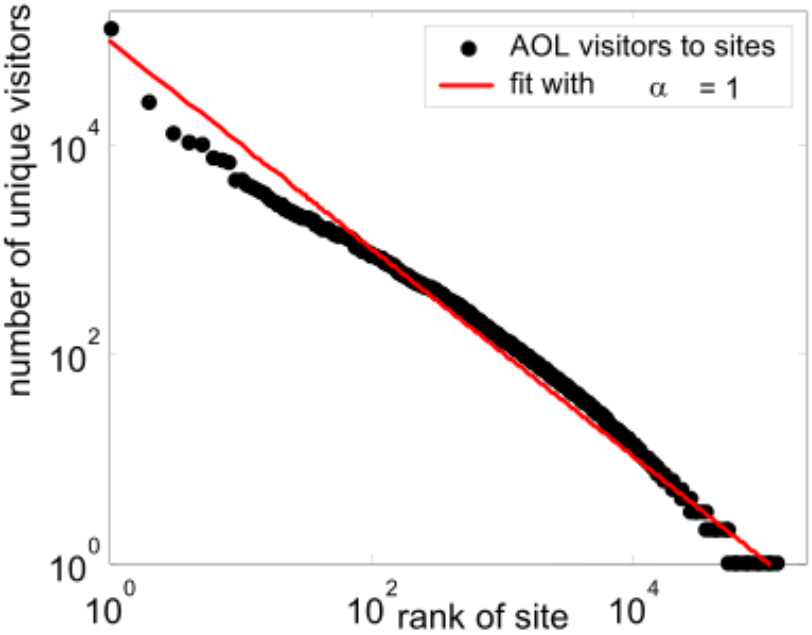
Send clients to server with best end-to-end performance

- Performance depends on
  - Server load
  - Content at that server
  - Network conditions

- Optimizing for server load
  - Load balancing, monitoring at servers
  - Generally solved

Optimizing performance: caching

- Where to cache content?
  - Popularity of Web objects is Zipf-like
    - Also called heavy-tailed and power law
  - $N_r \sim r^{-1}$
  - Small number of sites cover large fraction of requests
  - Different popularity depending on the location
    - Temporal and spatial popularity
  - Do you think it is easy to predict?



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

Optimizing performance: Network

- Example approximation strategies
  - Geographic mapping
    - Internet paths do not take shortest distance
  - Active measurement
    - Ping from all replicas to all routable prefixes
    - 56B * 100 servers * 500k prefixes = 500+MB of traffic per round
  - Passive measurement
    - Send fraction of clients to different servers, observe performance
    - Downside: Some clients get bad performance

THE ACCIDENTAL LEAK —

Google goes down after major BGP mishap routes traffic through China

Google says it doesn't believe leak was malicious despite suspicious appearances.

DAN GOODIN - 11/13/2018, 2:25 AM

https://www.flickr.com/photos/bjsnaaow/5668460525

Enlarge

How to understand CDNs

- Deploy multiple servers across the globe
  - How to make users connect the CDN server?
    - Changing all URLs of the content
    - DNS Authoritative server
  - How to find the best server?
    - IPs of DNS resolver
    - Anycast
    - EDNS (not covered from the class)

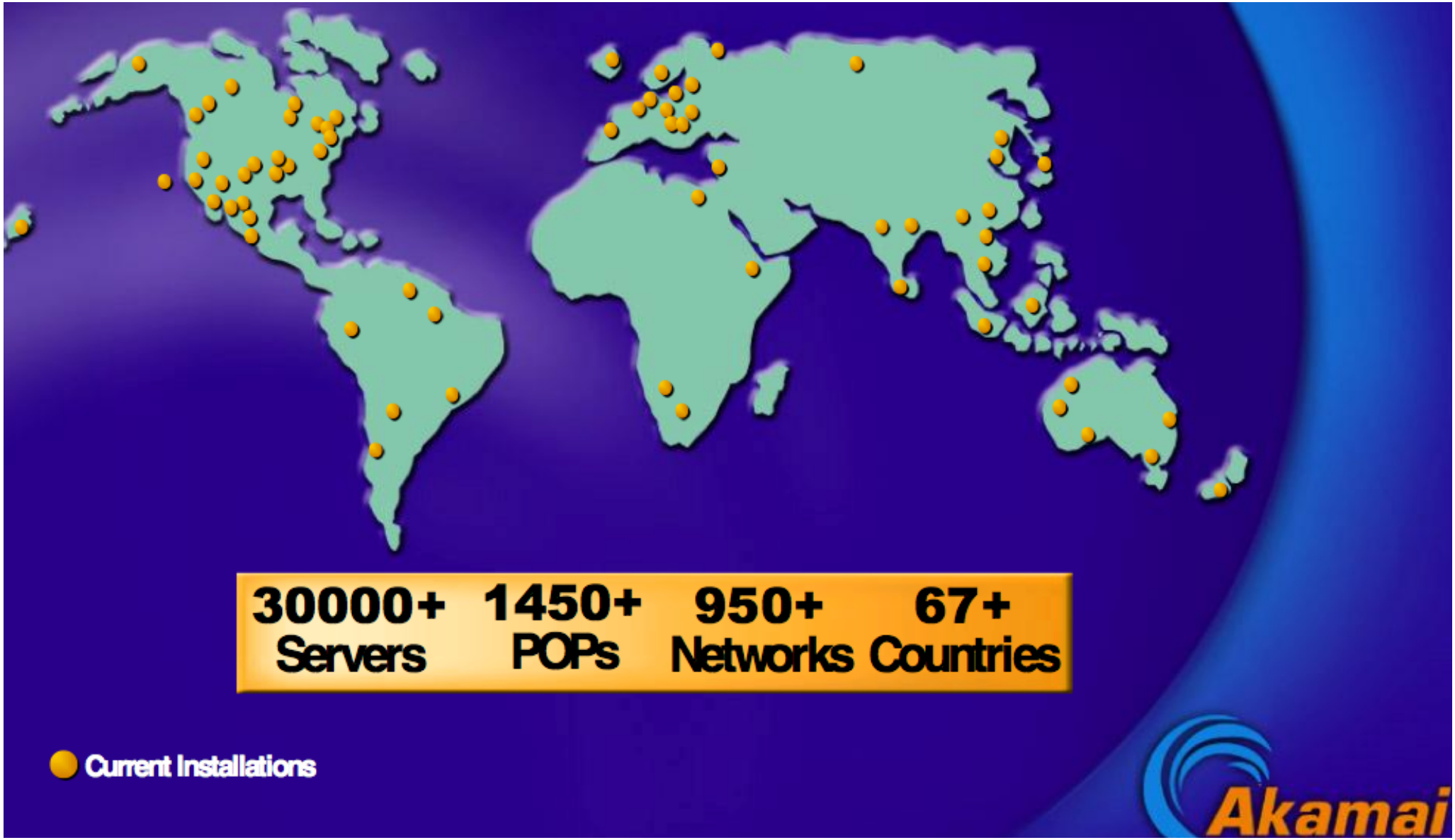
Outline

- Motivation
- CDN basics
- **Prominent example: Akamai**

Akamai case study

- 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 inside100 new networks in last few years
- Customers
  - 250K+ domains: all top 60 eCommerce sites, all top 30 M&E companies, 9 of 10 to 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

Somewhat old network map



DNS Redirection

- Web client's request redirected to 'close' by server
  - Client gets web site's DNS CNAME entry with domain name in CDN network
  - Hierarchy of CDN's DNS servers direct client to 2 nearby servers