15-441 15-641 Computer Networking

Lecture 17: Delivering Content
Content Delivery Networks
Peter Steenkiste

Fall 2016
www.cs.cmu.edu/~prs/15-441-F16

Overview

- Web
- Consistent hashing
- Peer-to-peer
- CDN
  - Motivation
  - Edge servers
  - Content delivery
  - Mapping
  - Impact on Internet
- Video

Mapping: Server Selection

- Which server?
  - Lowest load → to balance load on servers
  - Best performance → to improve client performance
    - Based on Geography? RTT? Throughput? Load?
  - Any alive node → to provide fault tolerance
- How to direct clients to a particular server?
  - As part of routing → anycast
    - Not covered ☹
  - In a data center → cluster load balancing
  - As part of application → HTTP redirect
  - As part of naming → DNS

Application Based

- HTTP supports simple way to indicate that Web page has moved (30X responses)
- Server receives Get request from client
  - Decides which server is best suited for particular client and object
  - Returns HTTP redirect to that server
- Can make informed application specific decision
- May introduce additional overhead → multiple connection setup, name lookups, etc.
- While good solution in general, but…
  - HTTP Redirect has some design flaws

Naming Based

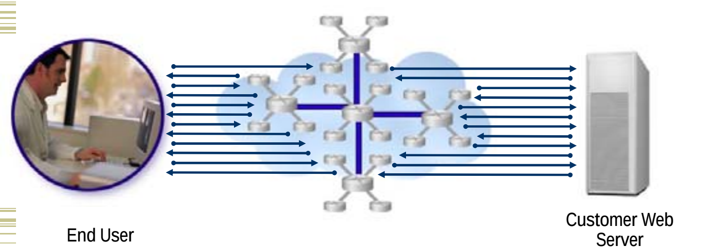
- Client does name lookup for service
- Name server chooses appropriate server address
  - A-record returned is “best” one for the client
- What information can name server base decision on?
  - Server load/location → must be collected
  - Information in the name lookup request
    - Name service client → typically the local name server for client

5

Mapping Algorithms

- Three main components to finding “closest” edge cache to end user from a Network point of view:
  - Packet Loss + Throughput + Latency
  - Listed in order of importance (roughly)
- Mapping also takes into account edge cache performance
  - Does a server have an object on its hard drive?
  - Uses consistent hashing algorithm
  - Does the edge cache have CPU, RAM, bandwidth, etc. available to serve end-user?

Access to Web Site without CDN



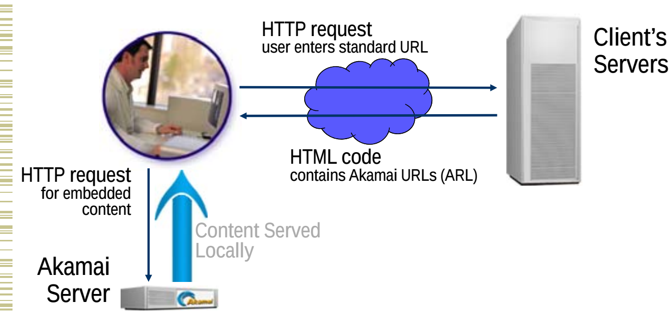
1 User enters standard URL
www.customer.com

2 Customer’s Web Server returns HTML with embedded URLs
<img src=/images/logo.gif>

3 User’s browser requests embedded objects from customer Web server

4 Objects served with round trips across the Internet

FreeFlow – Akamai’s Object Delivery Service



Example ARL: img src= a1000.g.akamai.net/…/www.customer.com/images/logo.gif

Steps in Content Retrieval

- Clients fetch html document from primary server
  - E.g. fetch index.html from cnn.com
- URLs for replicated content are replaced in html
  - E.g. <img src=“http://cnn.com/af/x.gif”> replaced with <img src=“http://a73.g.akamaitech.net/7/23/cnn.com/af/x.gif”>
  - Note that modified name includes original file name
  - Drawback: must modify the content
- Client is forced to resolve aXYZ.g.akamaitech.net hostname
- Focus is on static content
  - Can be freely cached

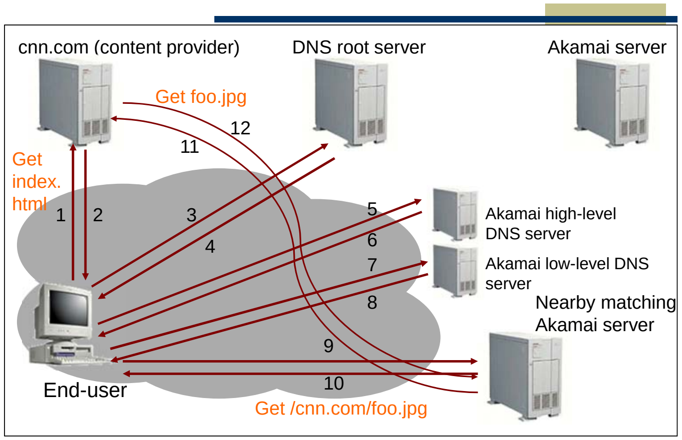
9

Resolution of Modified Name

- Root server gives NS record for akamai.net
- Akamai.net name server returns NS record for g.akamaitech.net
  - Name server chosen to be in region of client’s name server
  - TTL is large
- G.akamaitech.net nameserver chooses server in region
  - Should try to chose server that has file in cache - How to choose?
  - Uses aXYZ name and hash
  - TTL is small → why?

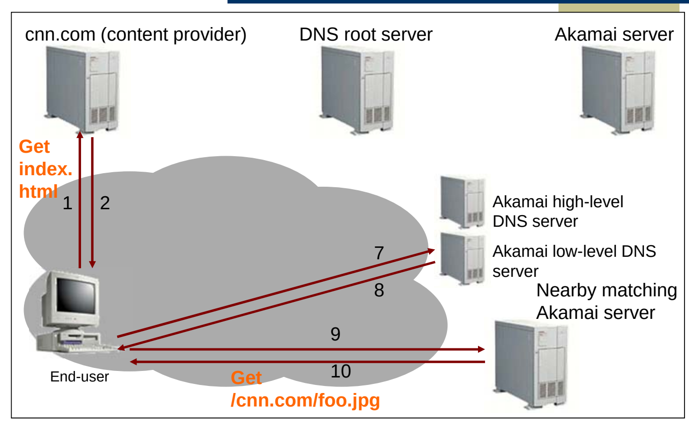
10

How Akamai Works



11

Akamai – Subsequent Requests



12

Redirecting to CDN for First Page

Customer CNAME's (aliases)
www.customer.com

- Anyone looking up www.customer.com will be redirected to an Akamai hostname - "customer.d4p.net"
- customer.d4p.net is CNAME'd to aXXX.g.akamai.net
- Standard Akamai mapping magic sends returns the closest edge server for aXXX.g.akamai.net

Impact of Redirection

End user never communicates with origin server
- High reliability: thousands of servers backing each other up
- If one geographic area is disabled, no other area is affected
- Mitigates some DoS attacks

Uncacheable content is tunneled back to origin
- Can reduce need for tunneling with server-side scripts running on edge caches – now very common

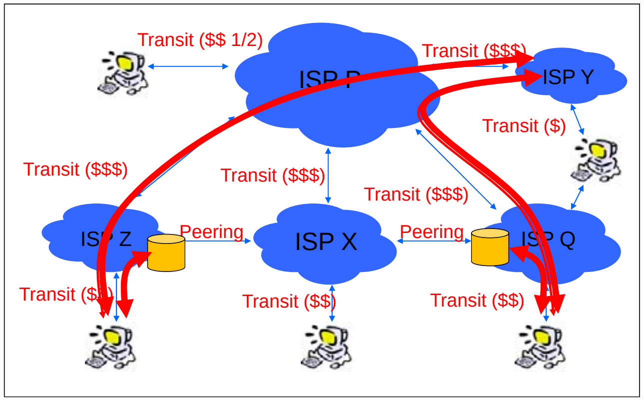
Persistent TCP connections increase performance
- Helps with downloading of objects to end caches

Can be used to outsource entire service to CDN
- Customer's code runs on CDN servers
- Must coordinate with origin server, e.g., consistent state

Overview

- Web
- Consistent hashing
- Peer-to-peer
- CDN
  - Motivation
  - Edge servers
  - Content delivery
  - Mapping
  - Impact on Internet
- Video

Change in Traffic with CDN Edge Caches



Impact of CDN Growth

- Flattening of the Internet
  - More content is served from the edge of the network
  - Also see an increase in peering
  - More traffic remains at edge – reduced load on core
- Changes in the economic relationships
  - Caches benefit users: better performance, reliability
  - Happy customers benefit the CDN
  - ISP benefits since more content is served locally
    - Reduces traffic from provider – direct economic benefit
  - CDNs sometimes place caches in “eye ball” ISPs for free, but economic models change all the time

Some Recent Trends

- If CDN’s can deploy caches, why can’t I?
- Content providers have started to deploy CDNs
  - Reduce cost, assuming you are large enough
  - Optimize caching to their specific requiremens
  - Can still use CDNs, e.g., in certain regions, ….
- Internet Service Providers also try deploy CDNs
  - Sometimes difficult to build the business relationships with content owners – too many ISPs!
  - How about the know-how?
  - Hybrid solutions are emerging, e.g., ISPs install hardware and license software from CDNs

Summary

- Caching improves web performance
- Caching only at client is only partial solution
  - Not enough locality
- Content Delivery Networks move data closer to user, balance load, increase availability, …
  - Is having impact on structure of the Internet
  - No longer just a solution for static content

Overview

- Web
- Consistent hashing
- Peer-to-peer
- CDN
- Video
  - Context: network performance
  - Video technologies

Yesterday's Applications

- FTP: transfer files to a host
  - No distributed file systems!
  - Mostly replaced by "the web" – http
- Telnet: use a computer remotely
  - Similar to ssh today (minus the security)
- Mail: exchange electronic e-mail
  - Similar today (kind of)
  - Initially host-to-host: name@my.computer.edu
- Very exciting (30 years ago)!
- Requirements?

21

Today's Applications

- Amazon, Facebook, etc.
  - What matters most?
  - 2009 quote: "*Amazon* found every 100ms of *latency* cost them 1% in sales"
- Video streaming
  - Accounts for very high percentage of bandwidth
  - Interactive versus broadcast versus playback
  - What matters most?
- Skype audio and video conferencing
  - Traditional telephone app
  - What matters most?

22

Requirements

- Network reliability
  - Network service must always be available
- Scalability.
  - Scale to large numbers of users, traffic flows, …
- Performance: latency and throughput
- Security: for users and the network itself
  - Privacy, authentication, deal with various attacks, …
  - Attacks on the network versus enabled by the network
- Manageability: monitoring, enforcing policies, billing, …

23

What Service Does an Application Need?

Data loss

Timing

Bandwidth

24

Basic Network Requirements

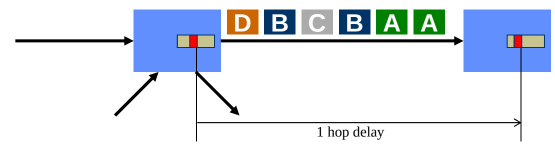
Application	Data loss	Bandwidth	Time Sensitive
file transfer	no loss	elastic	no
e-mail	no loss	elastic	no
web documents	no loss	elastic	no
real-time audio/ video	loss-tolerant	audio: 5Kb-1Mb video:10Kb-5Mb	yes, 100's msec
stored audio/video	loss-tolerant	same as above	yes, few secs
interactive games	loss-tolerant	few Kbps	yes, 100's msec
financial apps	no loss	elastic	yes and no

Somewhat dated – How?

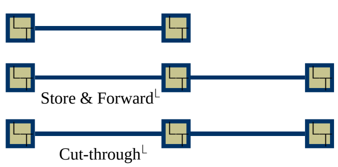
Packet Delay Components for One Link

Sum of a number of different delay components:

- Queuing delay on each router: contention with other traffic.
  - Depends on the traffic load and queue size
- Transmission delay on each link: time to put bits on wire.
  - Proportional to the packet size and 1/link speed
- Propagation delay on each link: ~speed of light for 1 bit.
  - Proportional to the length of the link
- Processing delay on each router: mostly for header.
  - Depends on the speed of the router



Packet Delay in a Network



Propagation + Transmit + other

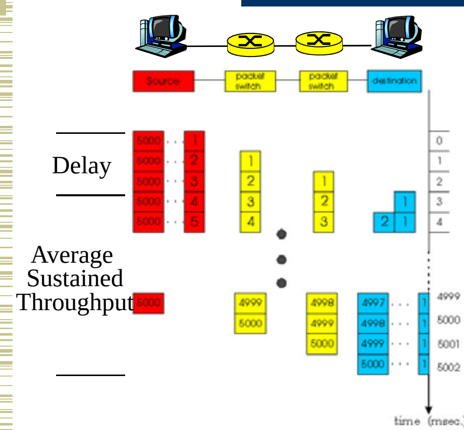
2*(Prop + xmit+ other)

2*(prop + other) + xmit

When does cut-through matter? *

- Only possible if there is no queue

Application-level Delay



$$\text{Delay}^* + \frac{\text{Size}}{\text{Throughput}}$$

Units: seconds + bits/(bits/seconds)

* For minimum sized packet

Some Examples

- How long does it take to send a 100 Kbit file?
  - Assume a perfect world
- Is the transfer latency or throughput limited?
- What about a 10 Kbit file?

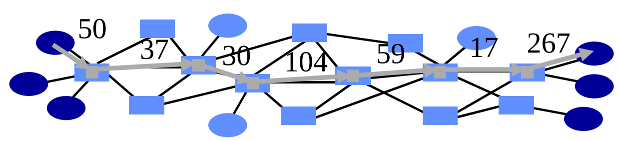
Throughput / Latency	100 Kbit/s	1 Mbit/s	100 Mbit/s
500 µsec			
10 msec			
100 msec			

A Word about Units

- What do “Kilo” and “Mega” mean?
  - Depends on context
- Storage works in powers of two.
  - 1 Byte = 8 bits
  - 1 KByte = 1024 Bytes
  - 1 MByte = 1024 Kbytes
- Networks work in decimal units.
  - Network hardware sends bits, not Bytes
  - 1 Kbps = 1000 bits per second
  - To avoid confusion, use 1 Kbit/second
- Why? Historical: CS versus ECE.

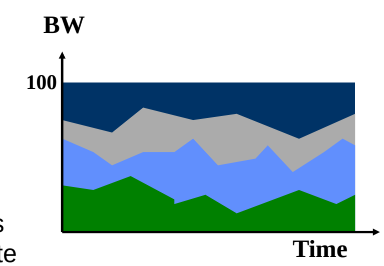
A Closer Look at Throughput

- When streaming packets, the network works like a pipeline
  - All links forward different packets in parallel
- Throughput is determined by the slowest stage
  - Called the bottleneck link
- Does not matter why the link is slow!
  - Low link bandwidth
  - Many users sharing the link bandwidth



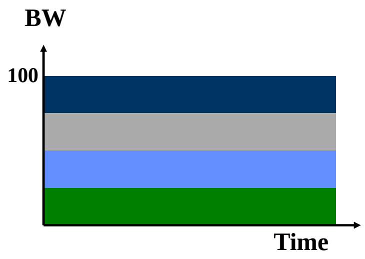
Bandwidth Sharing

- Bottleneck link bandwidth determines end-to-end throughput
- Router before the bottleneck link decides how much bandwidth each user gets
- User bandwidth can fluctuate quickly as flows join, leave, or change rate
- Queues are critical in “keeping the pipe full”



Fair Sharing of Bandwidth

- All else being equal, fair means that users get equal treatment
  - Sounds fair
- When things are not equal, we need a policy that determines who gets how much bandwidth
  - Users who pay more get more bandwidth
  - Users with a higher “rank” get more bandwidth
  - Certain classes of applications get priority



33