Content Distribution

- How to distribute content without requiring centralized, heavy-duty servers?
- Examples:
  - Bittorrent
    - Peer-to-peer content distribution
  - Akamai
    - Content distribution service

Bottlenecks for Web Downloads

- 3 potential bottleneck locations:
  - The first mile
    - User’s access connection may be limited
  - The last mile
    - Link connecting server to Internet can get overloaded by too many requests
  - Peering points
    - ISPs have little motivation to provide high-capacity connectivity to their peers

- Idea: replication can address the 2\textsuperscript{nd} and 3\textsuperscript{rd} issues
  - Only end-user can solve the 1\textsuperscript{st} one

- Content Distribution Networks
  - Systems that provide such replication
### CDN Infrastructure

- **aaa.com**
- **bbb.com**
- **ccc.com**

![CDN Infrastructure Diagram]

### Content Distribution vs Caching

- **Explicitly manage cache content**
  - Sell as service to web site owners for performance

- **“Push” content to caches at major Internet providers**
  - Make content appear “nearby” (low latency) no matter where the requester is located

- **Change cache content when necessary**
  - Operates as a proxy cache to refresh content
**Akamai Content Distribution Network**

http://www.akamai.com/html/about/facts_figures.html

- **56,000 servers in 70 countries within nearly 1000 networks**
  - 85% of the world's Internet users have a single "network hop" to an Akamai server

- **Delivers 15 - 20% of all Web traffic**
  - Web traffic at times reaching more than 2 Terabits/sec
  - Hundreds of billions of daily Internet interactions

---

**Akamai CDN Example**

[Map of Akamai CDN Example]
Akamai Content Types

- Static (HTML, images, PDF, etc.)
  - Expiration time (0 – infinity) assigned by customer
  - On-demand cache invalidation available to customer
  - Special features (authenticated access, transfer encodings, etc.)

- Dynamic
  - Assembles cacheable and non-cacheable elements of page at cache (contacts origin only for non-cacheable)

- Streaming
  - Uses redundant streams and jitter-control to ensure quality playback

Web Site Redirection to Akamai

- DNS CNAME aliases
  - e.g., images.pcworld.com
  - CNAME=images.pcworld.com.edgesuite.net

- Modified URLs (“Akamized”)
  - Prefix with domain name in Akamai
  - e.g., http://a1694.g.akamai.net/8675/images.pcworld.com..
“Akamai-zing” Web Pages

An “Akamaized” URL

Serial# maps content to a large virtual space — allows for optimal and load balanced mapping to physical servers

Fingerprint cryptographic hash of data used to ensure freshness of content

Akamai URL Example (WRAL-TV)

Image: http://a1844.g.akamai.net/7/1844/95/22/html.wral.com/images/structures/menus/nav_curve.gif
Image: http://a1844.g.akamai.net/7/1844/95/22/html.wral.com/images/structures/spacer.gif
Image: http://a1844.g.akamai.net/7/1844/95/22/html.wral.com/images/structures/misc/personalize_left.gif
Image: http://a1844.g.akamai.net/7/1844/95/22/html.wral.com/images/structures/spacer.gif
Image: http://a1844.g.akamai.net/7/1844/95/22/html.wral.com/images/structures/spacer.gif
Image: http://a1844.g.akamai.net/7/1844/95/22/html.wral.com/images/structures/spacer.gif
Image: http://a1844.g.akamai.net/7/1844/95/22/html.wral.com/images/structures/spacer.gif
Image: http://a1844.g.akamai.net/7/1844/95/22/html.wral.com/images/structures/spacer.gif
Image: http://a1844.g.akamai.net/7/1844/95/22/html.wral.com/images/structures/spacer.gif
Image: http://a1844.g.akamai.net/7/1844/95/22/html.wral.com/images/structures/spacer.gif
Image: http://a1844.g.akamai.net/7/1844/95/22/html.wral.com/images/structures/spacer.gif
Image: http://a1844.g.akamai.net/7/1844/95/10m/images.ibsys.com/2001/0830/940459.jpg
Image: http://a1844.g.akamai.net/7/1844/95/10m/images.ibsys.com/2001/0830/940452_120x90.jpg
Image: http://a1844.g.akamai.net/7/1844/95/22/html.wral.com/images/structures/spacer.gif
Akamai -- DNS “Request Routing” (finding the “closest” cache)

http://images.pcworld.com/shared/graphics/cms/BIZDEV_hotspotFinder_50.png

Akamai DNS Processing
- Monitors all proxy servers and end-to-end conditions on paths from 1000s of network locations to servers
- DNS servers updated with new loads and maps every few seconds

**Monitored conditions**
- Server load and operational status
- Routing topology
- Latency and packet losses per route
- Available bandwidth per route

---

**How Much Server Diversity Exists?**

A. Su, et. al., "Drafting Behind Akamai (Travelocity-Based Detouring)", SIGCOMM 2006

Yahoo!

![Graph showing server diversity](image-url)
**Server Diversity per Customer & User Location**

![Graph showing server diversity per customer & user location.](image)

**Redirection Dynamics**

![Graph showing redirection dynamics.](image)