Lecture 30

CS625: Advanced Computer Networks
Fall 2003

Wednesday, 05 November 2003

Bhaskaran Raman
CSE, IIT-Kanpur

http://www.cse.iitk.ac.in/users/braman/courses/cs625-fall2003/outline.html

Outline for today

- Web server redirection mechanisms
- Scribe for today?

Web Server Load Balancing

- Setting: server replicas, for fault-tolerance and load-balancing
- Goal: serve client request from “appropriate” server replica
- Where to redirect?
- How to redirect?

Client-Side Approaches

- Web clients: e.g., Netscape’s website
- Smart clients: using Java applets
- Client-side proxies
- Disadvantages:
  - Not much server information
  - Redirection “hand-crafted” for each website => will not work for all servers
DNS-based Approaches
• Have an address for all server replicas
• DNS is used for redirection
• Disadvantages:
  – Limited control
  – DNS caching problems
• Constant-TTL versus Adaptive-TTL algorithms

Constant-TTL DNS Redirection
• Stateless algorithm: Round-Robin (RR)
  – Ignores server capacity or availability
• Server-state-based algorithm
  – Redirection to least-loaded server
• Client-state-based algorithm
  – Client-state: Request-rate, Distance
  – Multi-tier Round-Robin possible
• Server- and client-state-based approach possible

Adaptive TTL DNS Redirection
• Two-step process
  – Choose server
  – Choose TTL
• Usually with server- and client-state information

Dispatcher-based Approaches
• Centralized dispatcher:
  – Has single virtual IP address
  – Identifies individual servers through some other address
• Packet single rewriting
• Packet double rewriting
• Packet forwarding
  – MAC address based
  – ONE-IP approach
• HTTP redirection: server-state, location-based
Server-based Approaches

- Redirection by server itself
- Can be used as a second-level redirection
- HTTP redirection, or packet rewriting

Some Remarks

- Comparing the approaches...
- Server load more important than network bandwidth
- Throughput versus latency