Evaluating the Impact of DNS and HTTP Session Characteristics on Consumer ISP Web Traffic

Jason But, Urs Keller and Grenville Armitage

Centre for Advanced Internet Architectures
Swinburne University of Technology

jbut@swin.edu.au

November 2005

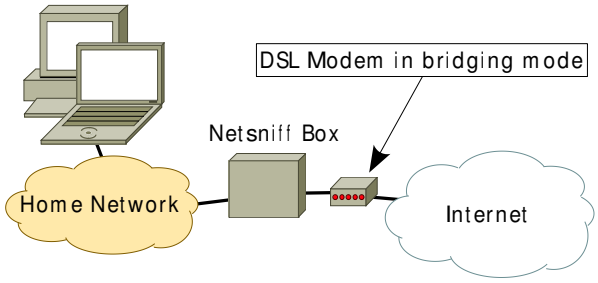
Outline

- Motivation
- Experimental Environment
- What makes up a HTTP Session
- Results
  - Gathered DNS Lookup Statistics
  - Gathered HTTP Transaction Statistics
- Caching Content
  - Cacheability of Content
  - Potential to Improve the Online User Experience

Motivation

- Web caches considered good
- Evaluate their effectiveness in the consumer ISP context
- Aim to meaure the entire user web experience
  - Correlate HTTP transactions with corresponding DNS queries
  - Measure the impact of client-server RTT with respect to HTTP transaction time
  - Estimate potential improvements to user web access times using different caching models

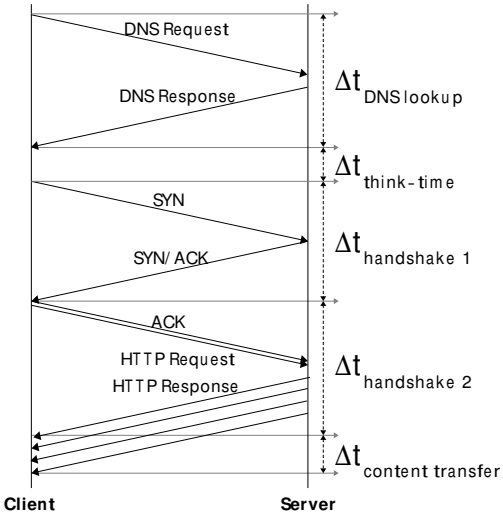
Experimental Environment



- NetSniff[1] Box running FreeBSD 5.3
  - Manages ISP Connection
  - Provides DHCP and DNS forwarding services within the home LAN
  - Provides a NAT service to any computer on the home LAN
  - Restarts data gathering upon failure
  - Logs rolled over daily and transfered to CAIA at 3AM

[1] Multi protocol layer network traffic analysis tool developed at CAIA - http://caia.swin.edu.au/ice/tools/netsniff

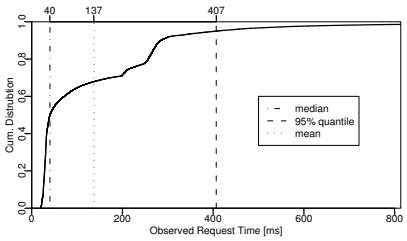
What Contributes to a HTTP Session



- DNS Query
  - If result is not locally cached
- HTTP Transaction
  - TCP Handshake
  - Content Transfer Time

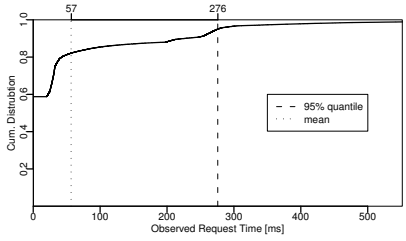
DNS Lookup Statistics

DNS Lookup Times for HTTP Transactions having a lookup



- About 73,500 DNS Requests
- Requests within 40ms handled by ISP DNS Server
- Responses > 250ms reflect access to an overseas DNS

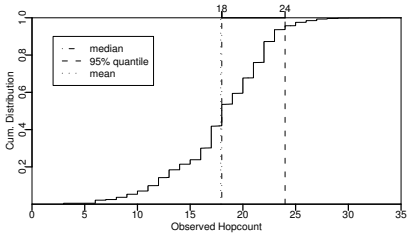
DNS Lookup Times for HTTP Transactions including browser cached entries



- Total 179,000 HTTP Transactions
- Other transactions incur no DNS lookups
  - DNS resolution cached at client
  - Equivalent to a DNS resolution of 0ms

Hop Counts to HTTP Servers

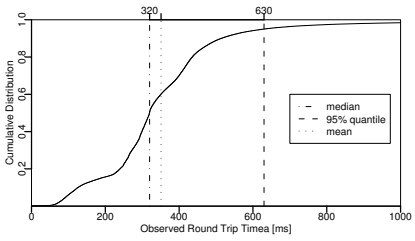
Path Hop Count Distribution for HTTP Transactions



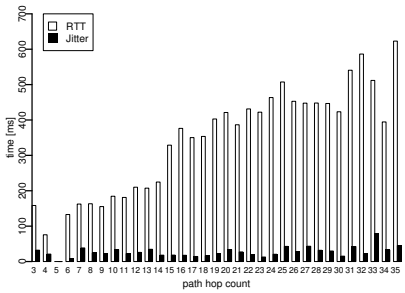
- Results indicate the absence of a Web Cache at the ISP
- Mean/median hop count to a web server is 18

RTT Distributions

RTT Distribution for HTTP Transactions



Mean RTT and Jitter vs. Path Hop Count for HTTP Transactions

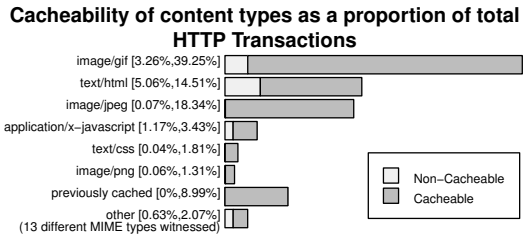


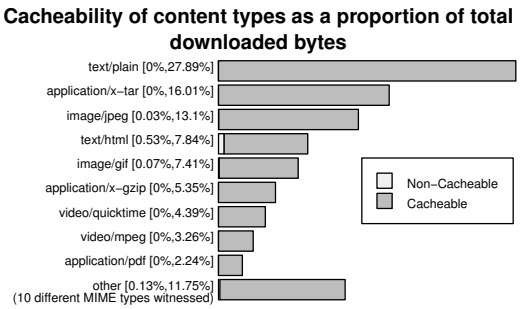
- Resonably linear spread of HTTP RTT distributions
  - Interrupted at 200ms (overseas servers)
- RTT and Jitter both related to number of hops in path

Would a Web Cache Help

- Decrease the Hop Count to the server
- Decrease the $2\ x\ RTT$ factor that makes up the overall content download time
  - Which will remain constant even as Internet access rates increase
- Minimise DNS query time
- Is enough content cacheable to justify?
- How much of an impact would it have?

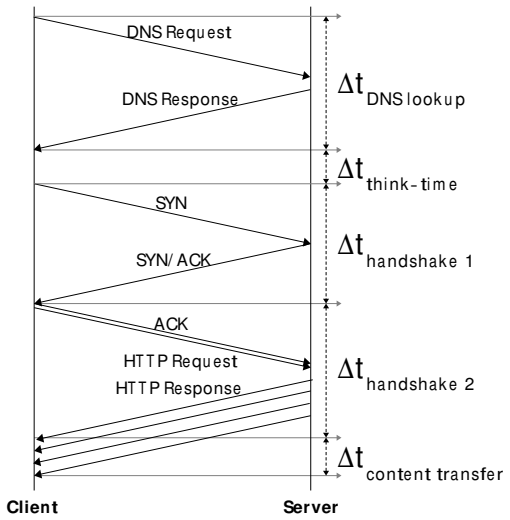
Cacheability of Content

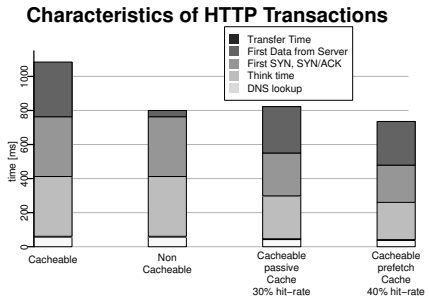




- Majority of downloaded content is cacheable
  - ~ 90% of individual HTTP transactions
  - > 99% of downloaded bytes
- Whether content is cached depends on the hit-rate of the cache
- Potential for large savings in access times

Potential Benefits





- Transfer time calculated as download at 512kbps
- For a cache hit
  - DNS lookup will be 0ms
  - HTTP RTT is equal to 20ms
- More potential for savings:
  - Up to 88% of content is cacheable

Conclusions

- ISP's we used did not cache web content locally
- RTT to the web server forms a significant proportion of the HTTP Transaction time
- Web Caches would reduce this RTT
  - Potential for 24-32% savings in average HTTP transaction time
- Reduce core traffic for other Internet applications
- Potential to become more important as access speeds increase
  - Data transfer time drops
  - RTT to server becomes a greater part of the HTTP transaction time