

Building caches for web applications

By
Vivek Khurana

Presented at Foss.Nitc
NIT Calicut, 4th march 2007

What is a web application

Data + Delivery mechanism

Redundant data

Web server transmits same page multiple times

HTTP Protocol

Keep alive
Cache expires
Document Expiration
Expiration Dates and Ages
If-Modified-Since
etc.

<http://www.w3.org/Protocols/rfc2616/rfc2616.txt>

Fine tune the DB

Indexing
Heap

Steps for caching

- 1) Receiving
- 2) Parsing
- 3) Lookup
- 4) Freshness check
- 5) Response creation
- 6) Sending
- 7) Logging

Receiving

- Receive data from multiple clients simultaneously
- Begin processing before complete message arrives

Parsing

- Normalize the headers
- Store headers in proper data structures

Lookup

- Cache stores Header and Body of the response
- On receiving a request cache will check local memory.
- In case of cache hit, cache will serve cache object

Freshness check

- cache checks for the freshness of the cache object
- HTTP protocol defines a method for checking freshness of cache
- HTTP rules are complex

Sending response

- Cache checks the original response header and adapts it according to the client.
- Cache-Control, Age, and Expires headers etc. (freshness check in short)

Logging

Log the response if you need to

Age calculation algorithm*

```
$apparent_age = max(0, $time_got_response - $Date_header_value);  
$corrected_apparent_age = max($apparent_age, $Age_header_value);  
  
$response_delay_estimate = ($time_got_response - $time_issued_request);  
  
    $age_when_document_arrived_at_our_cache =  
  
        $corrected_apparent_age + $response_delay_estimate;  
  
$how_long_copy_has_been_in_our_cache = $current_time - $time_got_response;  
  
  
$age = $age_when_document_arrived_at_our_cache +  
  
    $how_long_copy_has_been_in_our_cache;
```

*Taken from [http protocol RFC](http://http://tools.ietf.org/html/rfc2616#section-14.17).

Cache implementation

- As Objects
- As files
- As database files

3rd party caching libraries

- Jcache, Jcachestore
- php APC
- perl-cache

Thank you
hiddenharmony@gmail.com