Web Caching

- Cache-Related Protocols
- Cache Software and Hardware
- Impediments to Caching
- Replication
- Content Distribution
- Content Adaptation

Sashidhar Reddy
Cache-related Protocols

1. Internet Cache Protocol (ICP):
   - Query Protocol.
   - Works well with hierarchies.
     Sets of caches connected together as peers and linked to a common parent. Child can query siblings, Child can query parent but parent can’t query child.
   - UDP is used to contact peers.
   - Limit to the depth of hierarchies?
2. Cache Array Resolution Protocol (CARP):

- Set of caching proxies can effectively function as a single logical cache.
- Responses are collectively cached between the group of proxies.
- A hash function is used to partition the URLs across the caches.
- Deterministic.
- To locate a resource among caches hash function constructs a *resolution path*.
- Fewer duplications.
- Load balancing and reconfiguration of caches is difficult.
3. Cache Digest Protocol

- Digest is an indication of collection of objects in a cache.
- Checks the digest of each cache and contacts the cache that has the required resource.
- UDP is used to exchange digests between caches.
- Staleness of the digest?
- Size of the digest?
4. Web Cache Coordination Protocol (WCCP)

- Intercepts HTTP request and reroutes to a cache engine.
- Coordination is required for load balancing across multiple caches.
- Periodically checks if all the caches are alive.
Cache Software

The Squid Cache

- Most widely used.
- Can handle many protocols (HTTP, FTP, SSL etc.).
- Configurable Cache.
- Must be configured at client’s side.
- Uses ICP to talk to peers. (Recent version can also use cache digests).
- Can work in clustering environment.
Squid uses Access Control Lists (ACLs) to decide which clients can access it and disable access to certain Web servers.

```
  acl dumb dstdomain abc.com, def.com
  http_access deny dumb
```

- Can function as Web server accelerator.
- It can prevent attacks on original server by making just itself visible to the external network

```
  acl safe dstdomain safe1.com safe2.com
  http_access deny !safe
```
Problems with Proxy Caches:

1. Browser must be configured to contact a particular caching proxy.
2. Inter-cache communication is required to prevent duplication across the caches.
3. If the proxy is unavailable, the browser would have to be reconfigured.
   (Recent browsers can also be automatically reconfigured
   http://home.netscape.com/eng/mozilla/2.0/relnotes/demo/proxy-live.html)

Solution: Deploy specific pieces of hardware that rely on networking infrastructure.
Cache hardware

Interceptors and Redirectors:
- Intercepts HTTP requests and reroute them to web cache servers or cache clusters.
- Transparent to end user.
- Violates the end-to-end argument by not maintaining constant end points of the connection.
- Can be done using routers or switches.
Routers use policy-based rerouting (e.g., 80 for HTTP) method.
Switch is network device that has software for intercepting Web traffic and redirecting it to proxies. Switches (L5) can examine the content of the request. This can add to latency.
Appliances

✦ Needs less administration on customer’s part.
✦ Some require a separate redirection device and some redirect internally.
✦ Not very flexible.
✦ Customer might not gather enough information about their cache traffic.
Adaptive Web Caching

- Targets *hot spot* phenomenon.
- Multiple, distributed caches dynamically join and leave cache groups (cache meshes) based on content demand.
- Cache Group Management Protocol (CGMP) specifies how meshes are formed and how individual caches join and leave those meshes.
- Content Routing Protocol (CRP) is used to locate cached content from within the existing meshes.

*World Wide Web Caching: Trends and Techniques*

*Greg Barish and Katia Obraczka*
Impediments to Caching

Cache Busting:

- Some servers might not be interested in having cached content delivered by proxy.
- Cache Busting is a technique to prevent responses from being cached.

E.g.: Set **Expires** header to a value in the past.

 Include **Cache-Control: no-cache** headers.
How to prevent cache busting:

- **Hit-metering**
  Include a new HTTP header `meter`, which would be used by a cache to inform the server about the number of cache hits for that particular resource.

- **Ad-insertion**
  Proxy adds advertisements to a page.

- **Clear GIFs or Web bugs.**
  These are small images invisible to users, which are placed in on a page to serve as a counter.
Privacy Issues

- Neither client nor the origin server can be sure if the response is cached or not.
- Response could be cached at intermediaries along the response path against the wishes of the origin server.
- Caches can aid privacy!!!!!
  - Cache can shield users from origin servers.
**Replication**

- Content of the origin server is copied to mirror sites.

**Advantages:**
- Updating resources is easier.
- Possible to multicast resources as set of replicas are known before hand.
- Novel techniques are employed to send only differences to replicas.
Content Distribution

- Instead of duplicating all data of the origin server, only selected content is copied.

Content Adaptation

- Includes converting resources to different formats.
- Removes heavier computation tasks from the origin server