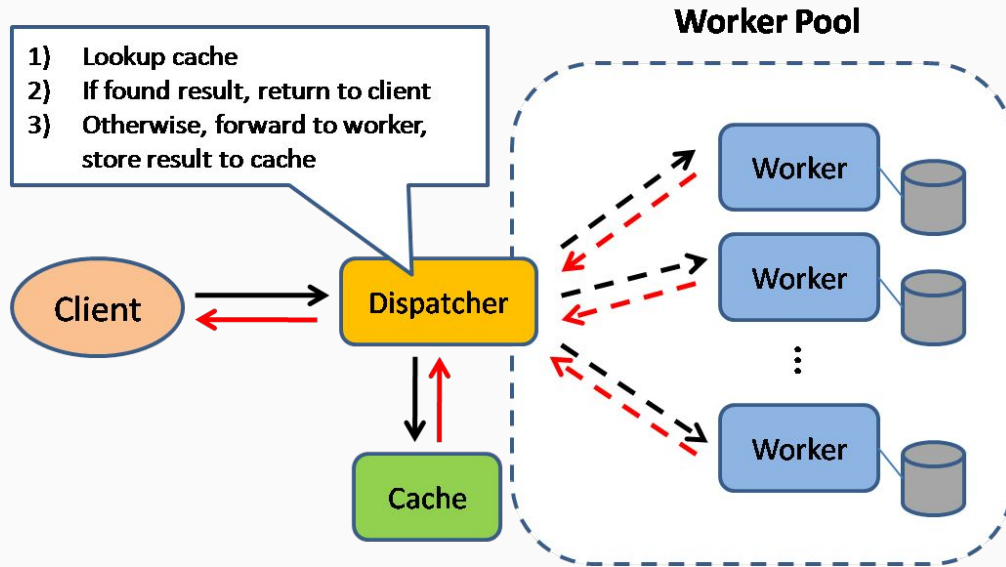


Caching in large scale distributed systems

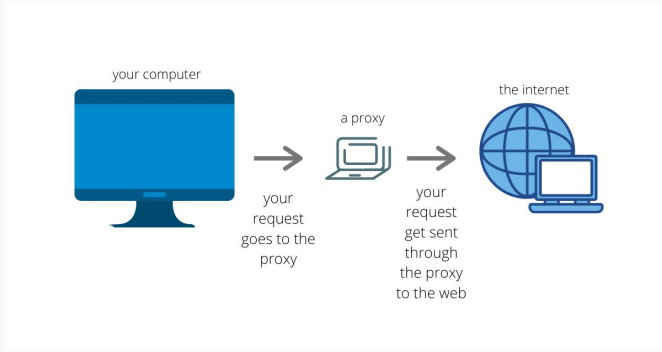




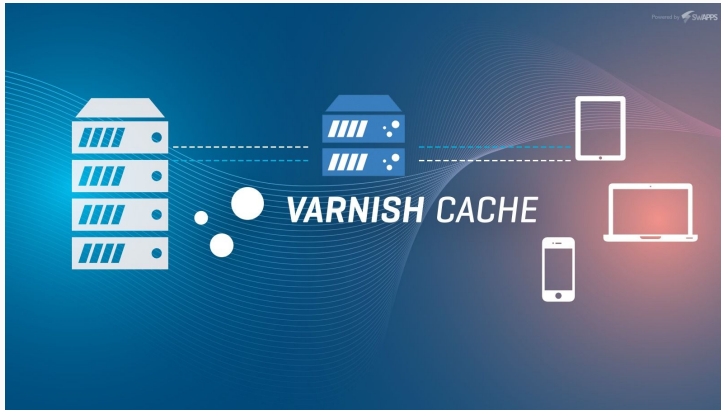
- Improves page load
- Reduce load on db servers
- Cache in front of database



- Serve content from locations closer to the user
- Request fulfilled by CDN does not go to servers
- Push/Pull CDNs

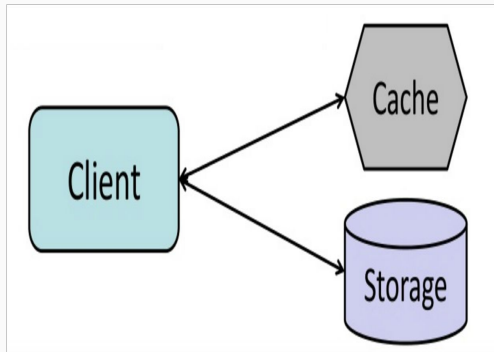


- Proxy server with static content.
Return response for cached requests
- Varnish for servers with http protocols
- Default caching for database server

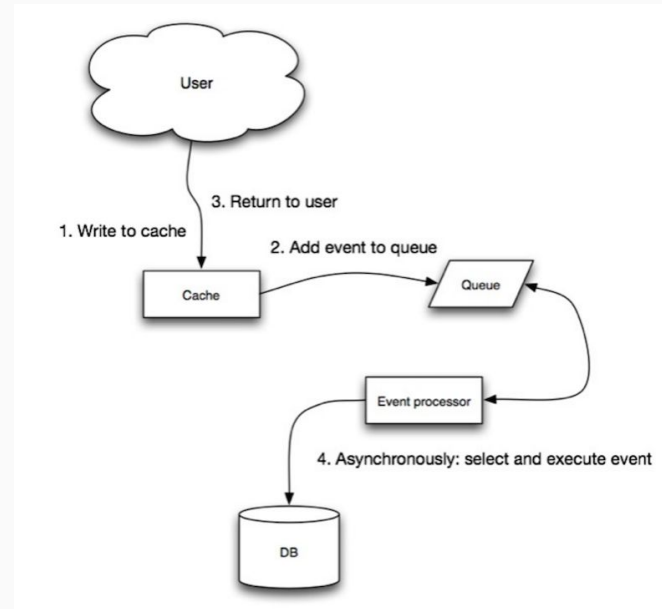
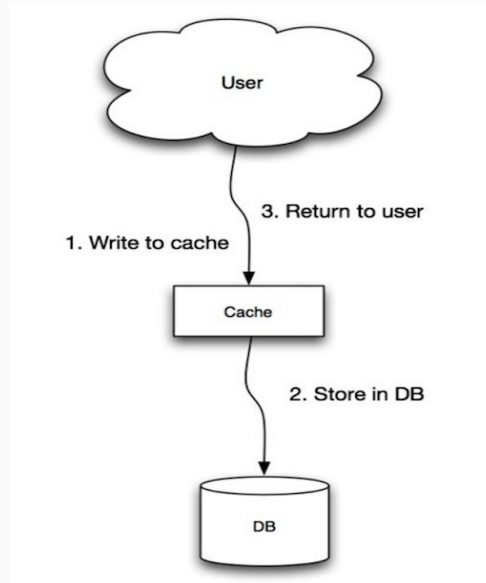


Caching types : Web server / database caching

- In-memory caches Key-value stores
- Data stored in RAM
- Row level
- Query-level
- Fully-formed serializable objects
- Fully-rendered HTML



- Cache side
- Write through
- Write behind



Thanks!

Resources :

- <https://github.com/donne-martin/system-design-primer#cache>

