Computer Science S-75
Building Dynamic Websites

Harvard Extension School
https://www.cs75.net/

Lecture 9: Scalability

David J. Malan
malan@harvard.edu
Web Hosts

- Bluehost
- DreamHost
- Go Daddy
- Host Gator
- pair Networks
- ...
VPSes

- DreamHost
- Go Daddy
- Host Gator
- Linode
- pair Networks
- Slicehost
- VPSLAND
- …
Vertical Scaling

- CPU
  - cores, L2 Cache, ...
- Disk
  - PATA, SATA, SAS, ...
  - RAID
- RAM
- ...
Horizontal Scaling
PHP Acceleration

- Code Optimization
- Opcode Caching
- . . .
PHP Accelerators

- Alternative PHP Cache (APC)
  http://pecl.php.net/package/APC
- eAccelerator
  http://eaccelerator.net/
- XCache
  http://xcache.lighttpd.net/
- Zend Platform
- ...
Load Balancing

Load Balancing with BIND

www  IN  A  64.131.79.131
www  IN  A  64.131.79.132
www  IN  A  64.131.79.133
www  IN  A  64.131.79.134
Sticky Sessions

- Shared Storage?
  FC, iSCSI, MySQL, NFS, *etc.*

- Cookies?
Load Balancers

- **Software**
  - ELB
  - HAProxy
  - LVS
  - ...

- **Hardware**
  - Barracuda
  - Cisco
  - Citrix
  - F5
  - ...

Caching

- .html
- MySQL Query Cache
- memcached
- ...
PHP / MySQL Short Term Project (Cambridge)

Reply to: job-640115588@craigslist.org
Date: 2008-04-12, 5:55PM EDT

We are students at Harvard Business School looking for someone to work with us in adding a subscription base to our website. Project would include adding a log in screen, payment setup, and helping to create the subscription pages which would need to draw information from a database (likely MySQL).

In talking with classmates, project appears to be around 50-100 hours, and would need to be completed in the next 2 months. Compensation is competitive, and we will supply graphics, data, and information.

Candidate should be strongly qualified in PHP, and MySQL.

- Location: Cambridge
- Compensation: Competitive with industry standards. Negotiable.
- This is a part-time job.
- This is a contract job.
- Extremely easy. Remember, please don't contact that job poster.
- Please, no phone calls about this job!
- Please do not contact job poster about other services, products or commercial interests.

PostingID: 640115588
MySQL Query Cache

`query_cache_type = 1`

$memcache = memcache_connect(HOST, PORT);
$user = memcache_get($memcache, $id);
if (is_null($user))
{
    $dbh = new PDO(DSN, USER, PASS);
    $result = $dbh->query("SELECT * FROM users WHERE id=$id");
    $user = $result->fetch(PDO::FETCH_ASSOC);
    memcache_set($memcache, $user['id'], $user);
}

http://memcached.org/
http://www.php.net/memcache
## MySQL

<table>
<thead>
<tr>
<th>Feature</th>
<th>MyISAM</th>
<th>Memory</th>
<th>InnoDB</th>
<th>Archive</th>
<th>NDB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Storage limits</td>
<td>256TB</td>
<td>RAM</td>
<td>64TB</td>
<td>None</td>
<td>384EB</td>
</tr>
<tr>
<td>Transactions</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Locking granularity</td>
<td>Table</td>
<td>Table</td>
<td>Row</td>
<td>Table</td>
<td>Row</td>
</tr>
<tr>
<td>MVCC</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Geospatial data type support</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Geospatial indexing support</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>B-tree indexes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Hash indexes</td>
<td>No</td>
<td>Yes</td>
<td>No [a]</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Full-text search indexes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Clustered indexes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Data caches</td>
<td>No</td>
<td>N/A</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Index caches</td>
<td>Yes</td>
<td>N/A</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Compressed data</td>
<td>Yes [b]</td>
<td>No</td>
<td>Yes [c]</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Encrypted data [d]</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Cluster database support</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Replication support [a]</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Foreign key support</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Backup / point-in-time recovery [a]</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Query cache support</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Update statistics for data dictionary</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

[a] Depends
[b] CSV, gzip, and zlib
[c] Some engines only
[d] Only with AES encryption

Replication: Master-Slave

Excerpted from *High Performance MySQL*.
Replication: Master-Master
Load Balancing + Replication
... + Partitioning
High Availability
Lecture 9: Scalability

David J. Malan
malan@harvard.edu