



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



Image from techui.com.

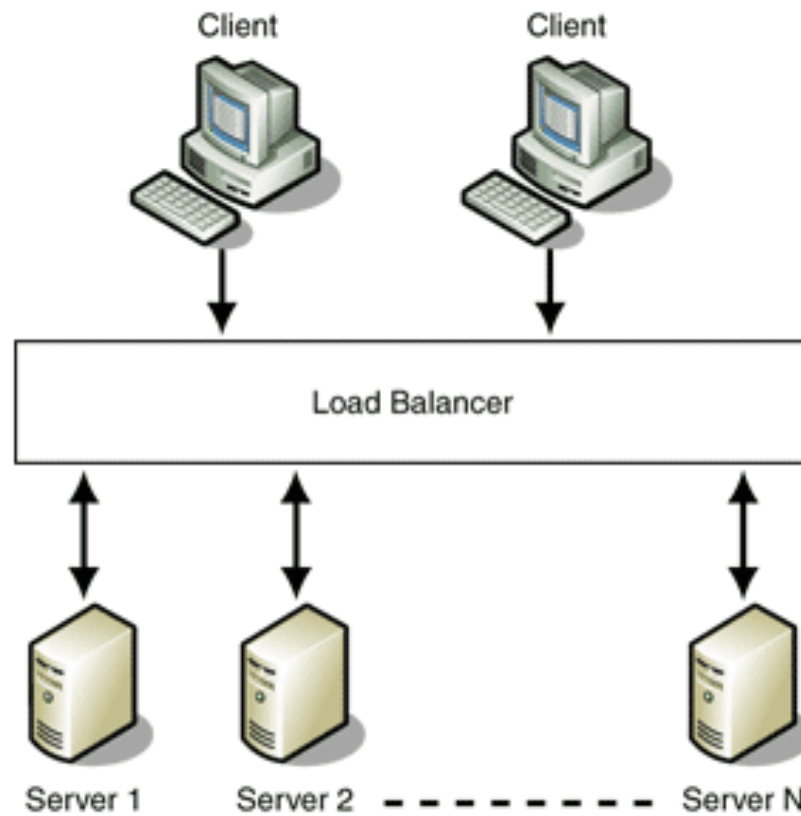
PHP Acceleration

- 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
<http://www.zend.com/en/products/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
- . . .

.html

The screenshot shows a Mozilla Firefox browser window with the address bar displaying `http://boston.craigslist.org/gbs/web/640115588.html`. The page content includes a breadcrumb trail: `boston craigslist > boston/camb/brook > web design jobs`. A yellow warning banner reads: **Avoid scams and fraud by dealing locally!** Beware any deal involving Western Union, Moneygram, wire transfer, cashier check, money order, shipping, escrow, or any promise of transaction protection/certification/guarantee. [More info](#). Below this is the job title: **PHP / MySQL Short Term Project (Cambridge)**. The posting details are: Reply to: job-640115588@craigslist.org, Date: 2008-04-12, 5:55PM EDT. The description states: 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. A bulleted list of details follows: Location: Cambridge; Compensation: Competitive with industry standards. Negotiable; This is a part-time job; This is a contract job; Principals only. Recruiters, please don't contact this job poster; Please, no phone calls about this job!; Please do not contact job poster about other services, products or commercial interests. The PostingID is 640115588. At the bottom, there is a copyright notice: Copyright © 2008 craigslist, inc. with links to [terms of use](#), [privacy policy](#), and [feedback forum](#). A right-hand sidebar contains a 'please flag with care:' section with links for [miscategorized](#), [prohibited](#), [spam/overpost](#), and [best of craigslist](#).

MySQL Query Cache

```
query_cache_type = 1
```

<http://dev.mysql.com/doc/refman/5.1/en/query-cache.html>

memcached

```
$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);  
}
```

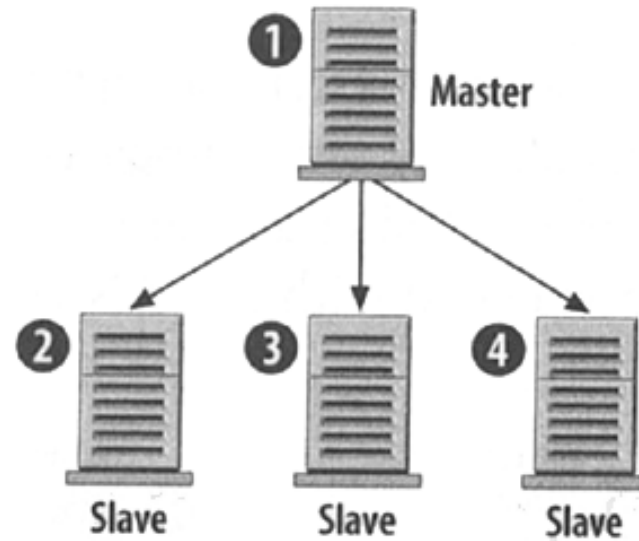
<http://memcached.org/>

<http://www.php.net/memcache>

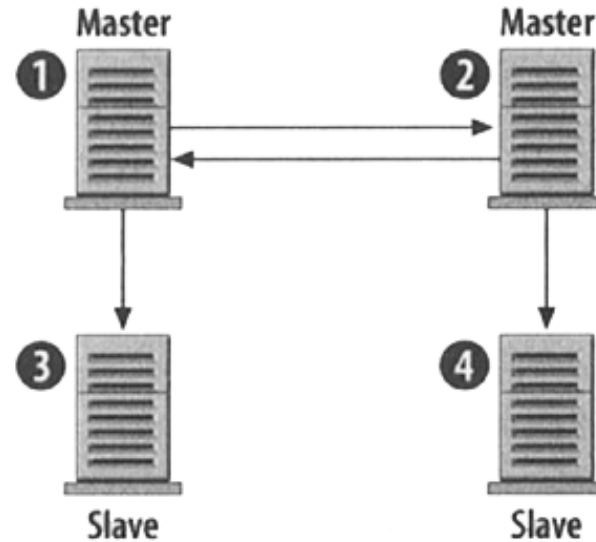
MySQL

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

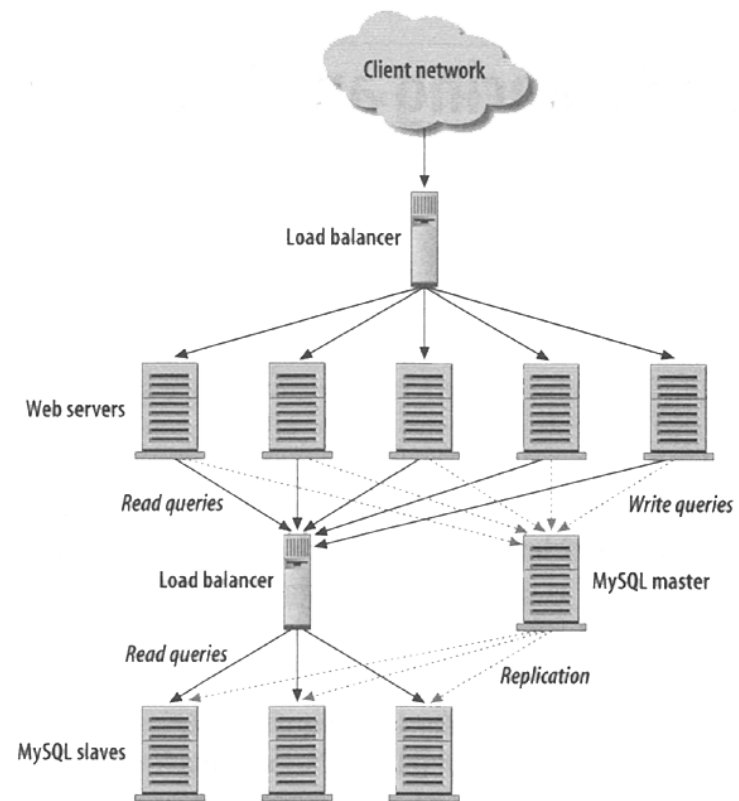
Replication: Master-Slave



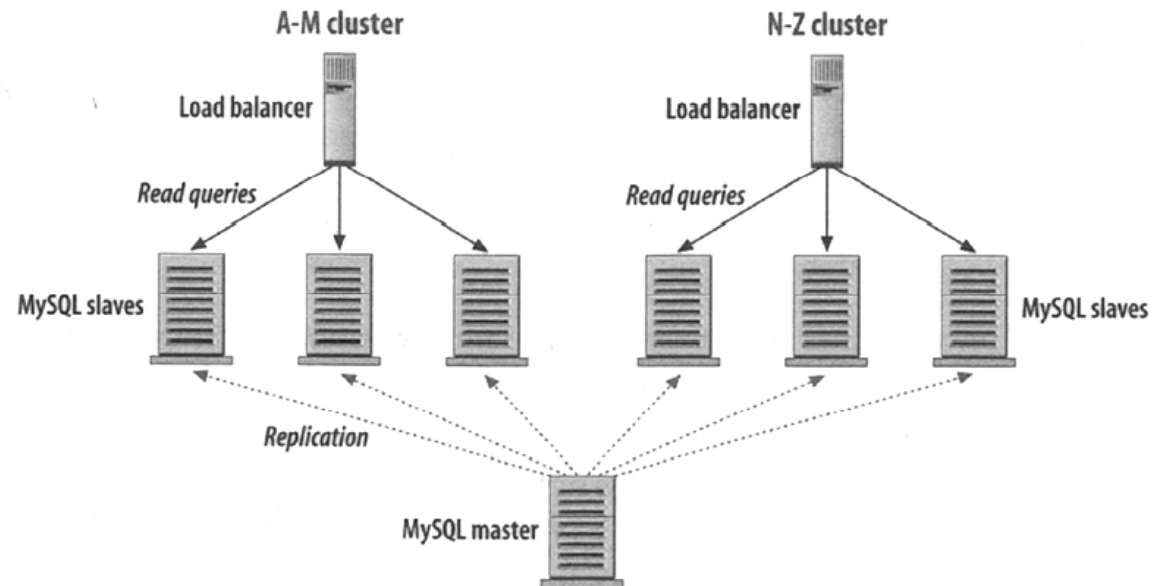
Replication: Master-Master



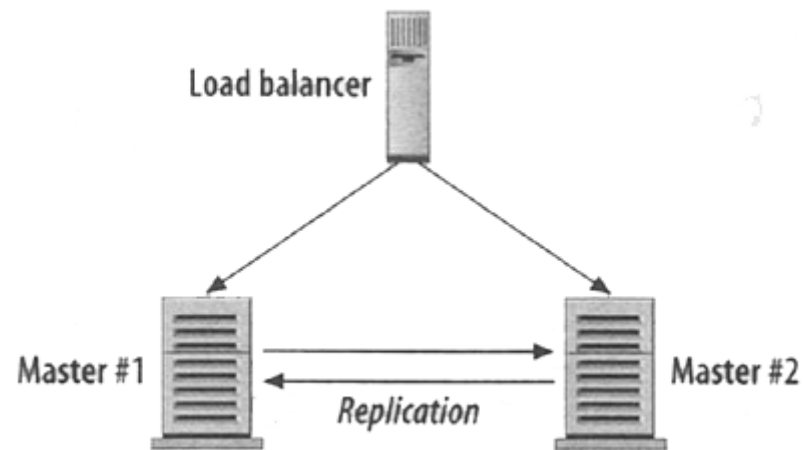
Load Balancing + Replication



... + Partitioning



High Availability





Computer Science S-75

Building Dynamic Websites

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

Lecture 9: Scalability

David J. Malan
malan@harvard.edu