Real-World PHP
Computer Science 50, Fall 2008

Keito Uchiyama
<keito at cs.harvard.edu>
The Real World

• What's so special about The Real World?
  – Lots more users (hopefully)
  – Frequent requests
  – Evil users
Topics

• Performance
• Security
• Templating
Performance

• How can we improve PHP performance?
  – Remove unnecessary code
  – Remove bottlenecks
  – Throw software at the problem
Remove bottlenecks

- Common bottlenecks
  - Disk
  - The PHP interpreter
  - Network
  - MySQL

- How can we remove these bottlenecks?
  - Optimize code
  - Cache requests to outside servers
  - Write better SQL
Bottleneck 1: Disk

- An online spellchecker
- A zip code table in a CSV file (Problem Set 8)
Bottleneck 2: The PHP Interpreter

- PHP is interpreted. Code has to be "compiled" on the fly to bytecode every time
- But code rarely changes!
- Solutions
  - Templating
  - PHP accelerators
Bottleneck 3: Network

• Networks are slow. Especially the interwebs!

```php
function goGetLolCat();
{
    $xml = new SimpleXMLElement(file_get_contents("http://feedproxy.google.com/ICanHasCheezburger?format=xml"));
    $item = $xml->channel->item[0];
    $title = htmlspecialchars(substr($item->description, strpos($item->description, "-") + 2), ENT_QUOTES);
    $link = $item->link;

    return "<a class='img' href='{$link}'><img alt='{$title}' border='0' src='" . $attributes['url'] . "' /></a>";
}
```
Solution: Caching

- Prevent your code from fetching the same data across the network every time

```php
define('CACHEFILE', 'lolcat.html);
define('MAXCACHETIME', 24 * 60 * 60); //Cache for 24 hours

if (filemtime(CACHEFILE) < time() - MAXCACHETIME) {
    //Cache is stale
    $loldata = goGetLolCat();
    file_put_contents(CACHEFILE, $loldata); //Cache loldata for later
} else {
    //Cache is still less than 24 hours old
    $loldata = file_get_contents(CACHEFILE); //Read in loldata from cache
}

echo $loldata;
```
Bottleneck 4: MySQL

• Don't fetch too much data (too many rows or columns)
  – Too many columns:
    SELECT * FROM students;
  – Too many rows:
    SELECT * FROM students WHERE student_username = 'jharvard';
Bottleneck 4: MySQL

• Use indexes!
• Types of indexes:
  – PRIMARY KEY
  – UNIQUE
  – INDEX
  – FULLTEXT
• What indexes to create
• Make MySQL EXPLAIN itself
Security

• Potential exploits
  – SQL injection
  – Cross-site scripting (XSS)
  – Invalid values
$username = $_GET['username'];
$sql = sprintf("SELECT * FROM Students WHERE username = '%s'", $username);

What if the user types in:
Robert
SELECT * FROM Students WHERE username = 'Robert';

But! What if the user types in:
Robert'; DROP TABLE Students; --
SELECT * FROM Students WHERE username = 'Robert'; DROP TABLE Students; --';
Security: Example 1

Hi, this is your son's school. We're having some computer trouble.

Oh, dear - did he break something? In a way -

DID YOU REALLY NAME YOUR SON Robert?); DROP TABLE Students;--?

Oh, yes. Little bobby tables, we call him.

Well, we've lost this year's student records. I hope you're happy.

And I hope you've learned to sanitize your database inputs.

http://xkcd.com/327/
Security: Example 2

//In register.php...
$username = $_GET['username'];
Store_Username_in_Database($username);

//In listusers.php...
foreach ($users as $user)
    echo $user->name;

What if user types in:
<script>alert('Pwned!');SendUserCookiesToBadSite();</script>
Security: Example 2

- Google search
- Gmail
- Yahoo! Mail
- Facebook
- MySpace
- Orkut
- TJMaxx

- Paypal
- Nokia
- eBay
- Microsoft Xbox
- Wikipedia
- Netscape
- Sourceforge

http://en.wikipedia.org/wiki/Cross-site_scripting
$numshares = mysql_real_escape_string($_GET['shares']);
$sql = "INSERT INTO Shares (name, shares) VALUES ('GOOG', $numshares);"

How can we do better?

- Regular expressions
- is_numeric()
- Check bounds
- filter_var()
Templating

• Lets you keep your logic and presentation separate

• Write cleaner code. No more ugly <?php ?> blocks in your presentation code.

• Bonus: Caches constant data for you so your code ends up running faster!
Templating: Smarty

Basic steps:

• Initialize Smarty engine
• Initialize variables
• Display template (a .tpl file)
// Get menu from database and print in table
$sql = 'SELECT dish_id, dish_name, dish_rating
       FROM dishes...';
if ($result = mysql_query($sql)) {
    $dishes = array();
    while ($row = mysql_fetch_assoc($result)) {
        $dishes[] = $row;
    }
    $smarty->assign('dishes', $dishes);
}
{if isset($dishes) }
<table class="menutable">
{foreach from=$dishes item=dish}
<tr>
<td><a href="/dish/{$dish.dish_id}">{$dish.dish_name|htmlspecialchars}</a></td>
<td>{stars rating=$dish.dish_rating}</td>
</tr>
{/foreach}
</table>
{else}
<p>No food is available at HUDS at this time!</p>
{/if}
Templating: Smarty

- http://www.smarty.net/
- Quick Install Guide
The Real World

• Never assume.
• Always doubt.
• Have fun*.
Real-World PHP
Computer Science 50, Fall 2008

Keito Uchiyama
<keito at cs.harvard.edu>