

# LAMP: Linux, Apache, MySQL, Perl/PHP/Python

Khalid Baheyeldin

Ontario Linuxfest  
October 24, 2009

<http://2bits.com>

The logo for '2bits' features the number '2' in a bright green color and the word 'bits' in a dark grey color. The letters are stylized with a 3D effect, including a drop shadow and a slight reflection below them. The '2' is positioned to the left of the 'bits'.



# Agenda



- Introduction
- What is LAMP?
- Linux, Apache, MySQL, PHP/Perl/Python
- Installation, Configuration
- Performance
- Alternatives
- Security
- Discussion





# About Khalid



- 25 years in software development and consulting
- Sinclair ZX Spectrum, mainframe, then UNIX since 1987
- Linux discovered 1990, using it regularly since 1995, “LAMP” since 1999
- Open source developer, contributor since 2003
- Full time open source consulting





# About 2bits.com



- Founded in 1999
- Drupal CMS/CMF since 2003
- Full time consulting
- Services
  - Drupal development
  - LAMP performance optimization and tuning
  - Server provisioning for performance and uptime
  - Manage huge sites for various clients
- <http://2bits.com>





# What is LAMP?

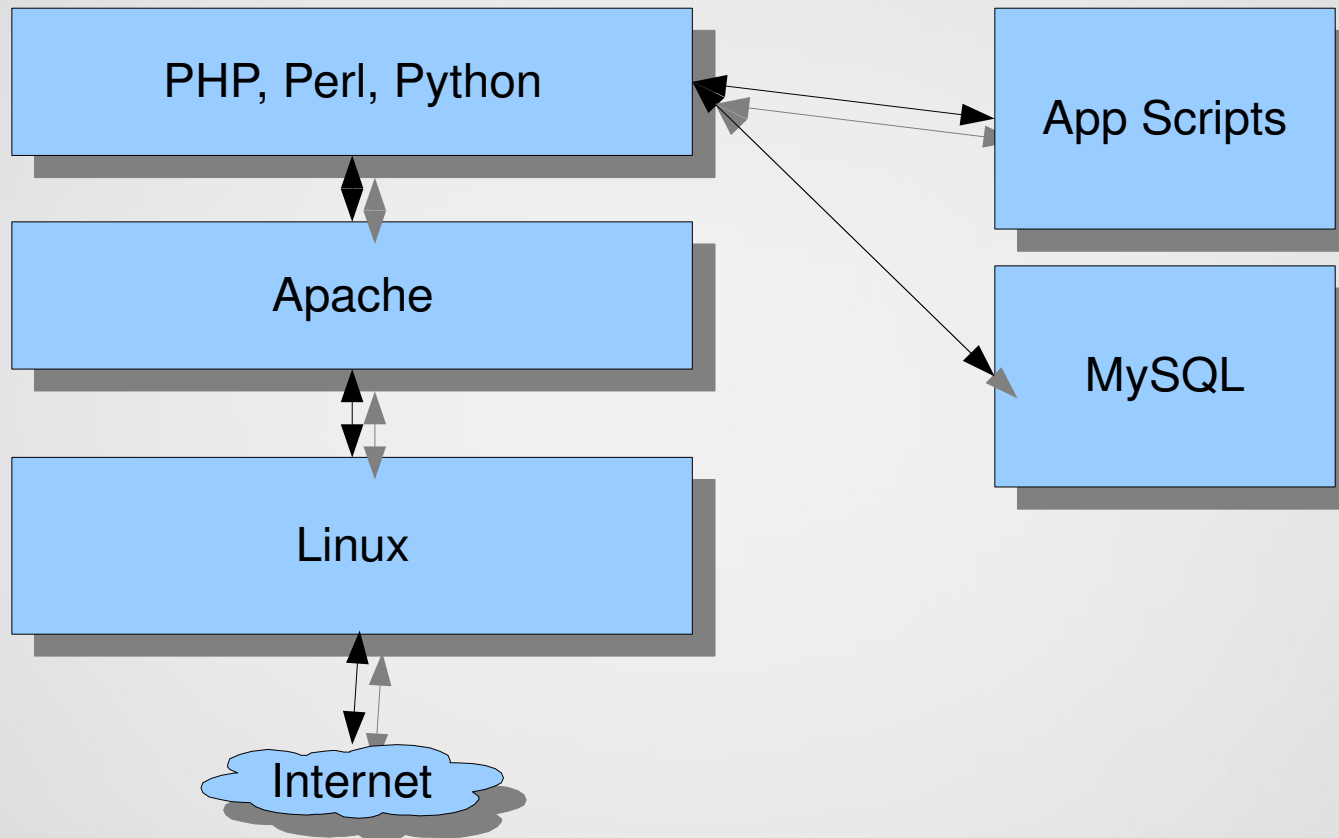


- Application platform: operating system, web server, database and scripting language
- LAMP stands for: Linux, Apache, MySQL and Perl/PHP/Python
- 2000: Term coined in Germany
- 2001: Term popularized by O'Reilly ([onlamp.com](http://onlamp.com))
- Technology stack itself was in use before that (e.g. Slashdot)





# LAMP Overview







# Benefits



- Perhaps the most widely used open source stack
- Poster child for the FOSS movement
- Ubiquitous, yet not “visible”, unlike the desktop or cell phones
- Business cost savings (no licensing needed)
- Low entry barrier for developers





# Who uses LAMP?



- Huge web sites
- Most of the top 20 sites, excluding Microsoft, Google and Chinese sites
- Examples:
  - Digg (Apache, PHP, MySQL)
  - Wikipedia (Apache, PHP, MySQL)
  - Yahoo (Apache, PHP, MySQL)
  - WordPress.com (PHP, MySQL)
  - Youtube (Apache)







# Who else?



- Many other large sites
  - Facebook (Apache, PHP, other stuff)
  - Craigslist (Apache)
- Geeky large web sites
  - Slashdot (Apache, MySQL, perl)
  - Linux.com (Apache, PHP)
  - Drupal.org (Apache, PHP, MySQL)





# Tip: HTTP Headers



- Curious about a site?
- Here is how to glean some info:
  - `time wget -S -O /dev/null http://digg.com`
- Shows HTTP Headers
  - Language used
  - Web server





# Applications



- A very large number of applications are build on the LAMP stack
- Content Management Systems (CMS)
  - Drupal, Joomla, Wordpress
  - MediaWiki (Wikipedia)
- Frameworks
  - Symfony, Django





# Installation



- Distro binaries
  - Far easier to install and maintain
  - Easier to apply security updates
  - Good for most cases
- Compiling from Source
  - Can customize further
    - e.g. Remove certain modules to save memory
    - Optimize for a given architecture
  - You take ownership of security patches and bug fixes





# Distro install



- Debian/Ubuntu

- Minimal install:

- aptitude install apache2 mysql-server php5 php5-mysql

- Details on 2bits.com at <http://bit.ly/j27Kc>

- CentOS

- Yum





# Other options



- Ubuntu Server CD has an “install LAMP server” option.
- Appliances
  - Can run in a virtual machine
  - <http://www.turnkeylinux.org/>
  - LAMP, MediaWiki, Drupal, Java, and much more







L

abits

Linux





# Linux



- Needs no introduction (this is Linux event!)
- Kernel + tools making up a distro
- Focus on server distros
  - Debian stable (Lenny)
  - Ubuntu Server edition LTS (8.04.2 Hardy Heron)
  - CentOS, free version of RedHat
- Seen LAMP running on Android phone, with Drupal on it!





A

abits

Apache





# Apache



- Patches to NCSA web server
- “A patchy server” -> “Apache”
- Free software
- Current stable version 2.2.x





# Apache marketshare



- Most widely used web server, Sept 2009 per Netcraft
- Active sites:
  - 54.32% of active sites
  - Next most used server is Microsoft IIS (20.05%)
- Top sites (busiest 1 million sites):
  - 66.9% Apache vs. 18.14% for IIS
- Used to be higher, Microsoft “bought” market share in 2007





# Apache modes



- MPM Pre Fork
- MPM Worker







# Apache pre-fork



- Each request is handled by a separate process
- Apache pre-forks a configurable number of processes, leaving spares around
- Incoming requests always have a handler for them, as long as the maximum is not reached
- Most common for PHP on VPS or dedicated, shared hosting more often CGI
- Configuration is important
  - Too low and you can get users complaining!
  - Too high and you get the server swapping!





# Pre-fork configuration



- Parameters for many things

<code>StartServers</code>	5
<code>MinSpareServers</code>	5
<code>MaxSpareServers</code>	10
<code>MaxClients</code>	150
<code>MaxRequestsPerChild</code>	0





# Apache MPM Worker



- Threaded mode
- Each request is handled by a thread
- Much less memory usage
- Does not work with certain setups (e.g. PHP with mod\_php)





# Worker configuration



- Parameters for:

<code>StartServers</code>	2
<code>MaxClients</code>	150
<code>MinSpareThreads</code>	25
<code>MaxSpareThreads</code>	75
<code>ThreadsPerChild</code>	25
<code>MaxRequestsPerChild</code>	0





# Tip: Worker + FastCGI

- Using Worker (threaded server) and FastCGI has the potential for memory savings





# Modules



- Modular design
- Modules do various specific things
- Can be static (compiled in) or shared (pluggable, implemented as a dynamic library)
- Can be disabled/enabled, reduce size







# Modules



- mod\_rewrite
  - manipulate URLs in various ways
  - e.g. Prettier URLs `example.com/index.php?q=something` becomes just `example.com/something`
  - Can be used it for mapping old content to new
- mod\_deflate
  - compresses web pages, css, js to save bandwidth
- mod\_status
  - statistics, number of requests, number of bytes, number of processes





M

abits

MySQL





# MySQL



- Relational Database
- History
  - Initially there was mSQL
  - MySQL emulated MySQL
  - Was used for decision support (read heavy)
  - Offered it for free to hosting companies
  - More adoption, more growth
  - “Previous version GPL'd”, then full GPL





# MySQL



- Pluggable “Engines”
- Commercial support
- Current stable version is 5.0.51





# MyISAM Engine



- Non transactional engine
- Lightweight
- Fast!
- Table level locking
  - Bad for high traffic sites
- Indexes and data in separate files





# InnoDB Engine



- Transactional engine
- Developed by InnoDB
- Row level locking
- Tablespaces (like Oracle), or one file per table
- Index and data stored in the same file
- Each table can be MyISAM or InnoDB
- More resource intensive







# Tip: Convert to InnoDB

- InnoDB is better for not locking the whole table
- Better concurrency
- Simple to convert tables from MyISAM to InnoDB
  - `ALTER TABLE table1 Engine=InnoDB;`
- Be careful with large tables though
  - Took 6 hours to complete for one client!





# MySQL Advantages



- Easy to use and administer
- Supported by all languages and frameworks
- Small enough
- Powerful enough
- Upgrades are easy (data format remains the same)





# Disadvantages



- Ignores statements that the engine does not support:
  - e.g. Non transactional engine (MyISAM) but using BEGIN TRANSACTION, ...etc.
- No sub second measurement
  - SHOW PROCESSLIST
  - 3<sup>rd</sup> Party patches have it (Percona)





# MySQL future



- Oracle bought InnoDB, makes of InnoDB a few years ago
  - Perhaps to meddle/control MySQL?
- Sun bought MySQL in 2008 for \$1B
- Sun is bought by Oracle in 2009
- MySQL founders have projects
  - Maria, Drizzle, OurDelta, ...etc.
- We can speculate all we want!
- MySQL is GPL!





# Interpreter Modes



- CGI
- Inline
- FastCGI





# Common Gateway Interface

- Oldest method of running dynamic stuff inside a web site (early to mid 90s?)
- Workflow:
  - Request comes in e.g. `/cgi-bin/something.cgi?param=value`
  - Web server forks a new process (`fork()`) then `exec()`
  - New process executes an interpreter (e.g. Perl)
  - A script is executed by the interpreter, with parameters passed
  - Passing back the result
  - Terminate the process







# CGI (cont'd)



- Advantages
  - Provided a way to create dynamic web sites
  - Security (if setup correctly, interpreter process can run other than the web server user, with limited privileges)
- Drawbacks
  - Slow (because of the fork system call)
  - Wasteful on resources (fork is expensive)
  - Not suitable for large web sites (does not scale)





# Inline (modules)



- Interpreter inside the Apache process
- PHP (mod\_php)
- Perl (mod\_perl)
- Python (mod\_python)





# Inline



Apache

Script interpreter (e.g. PHP)





# FastCGI

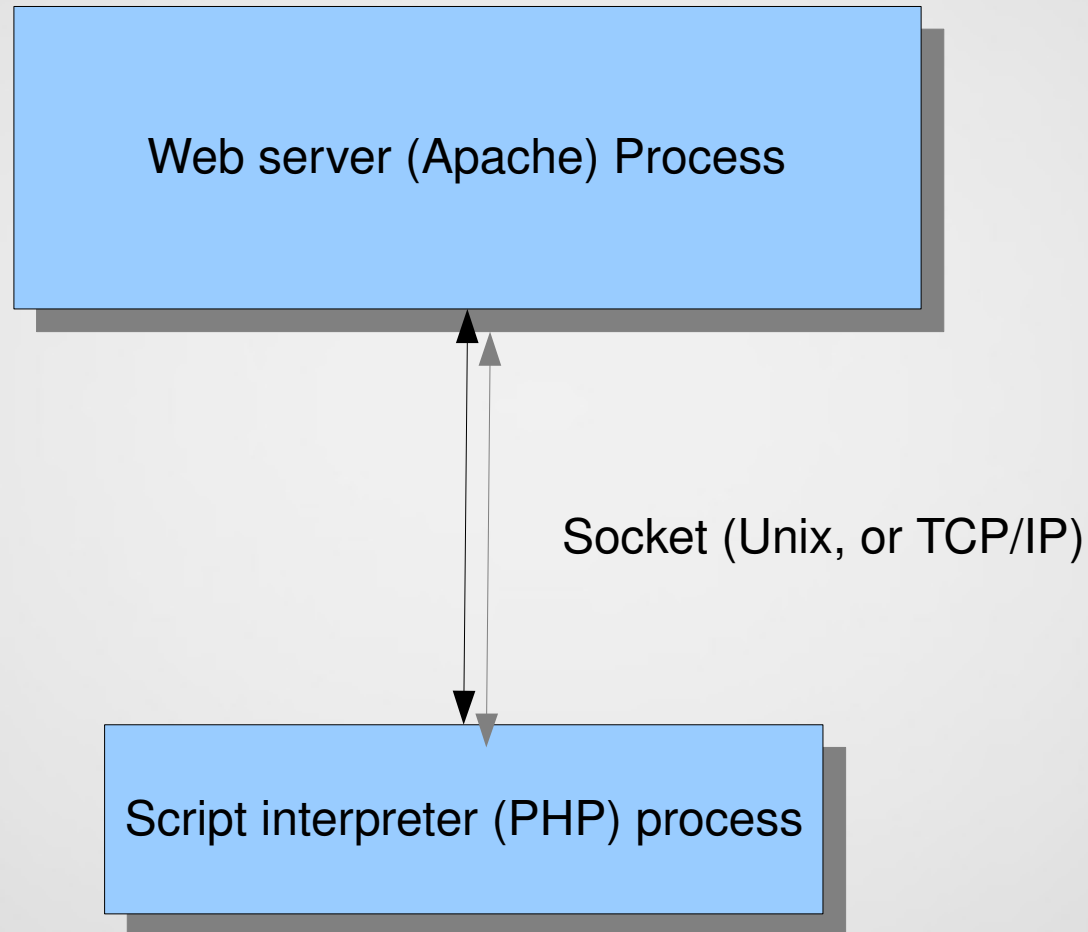


- Similar to CGI
- Does not fork per request
- Instead, uses a socket to communicate with an existing process that holds the interpreter
- Instead of Apache being bloated with 100s of processes, we have fewer dynamic processes, a small pool to handle dynamic requests
- Apache has mod-fastcgi (questionable stability) and mod-fcgid (much more stable)
- Best of both worlds!





# FastCGI





# FastCGI (fcgid)



- Advantages
  - Significant memory savings!
  - Speed is very close to mod\_php
  - Can be used with threaded server (MPM Worker)
  - Less connections to the database
- Drawbacks
  - Watch the timeouts (e.g. Large site with Drupal and cron)







P

abits

PHP, Perl, Python





# Scripting languages



- Interpreted, not compiled
  - Easy of develop
  - Easy to deploy
- Dynamic typing (integer vs string)
- Rich libraries
  - HTTP, file I/O, databases, string, math, arrays, XML, ...etc.)





# Why not compiled?



- C, C++?
- Can be used, but more cumbersome
- Need to be compiled and linked to the web server
- Harder to develop, harder to deploy
- Some large sites do that
  - Amazon, eBay
  - Probably Google (very secretive)





# PHP, Perl, Python



- PHP is popular and easy language, specifically for web development
- Python is a “better” language, but more generic.
- Perl a “worse” language, harder maintain, also not web specific





# PHP



- Most widely used language on the web
- Current stable version is 5.2.4
- Can be run in many modes
  - CGI
    - Drawbacks
  - mod\_php
  - FastCGI





# mod\_php



- Least problematic way of running PHP
- PHP is inside each and every Apache process
- Can be a tad faster than FastCGI
- Can also consume much more memory, since each Apache process has it, whether it is serving dynamic or static content







# Apache's config



- MPM-Worker for a large site

```
<IfModule mpm_worker_module>
    ServerLimit          500
    StartServers         10
    ThreadsPerChild     10
    MaxClients           600
    MinSpareThreads     30
    MaxSpareThreads     50
    MaxRequestsPerChild 3000
</IfModule>
```





# PHP and fcgid



- First install fcgid
- aptitude install libapache2-mod-fcgid
- Configure as follows:

```
<IfModule mod_fcgid.c>
    AddHandler fcgid-script .fcgi .php
    DefaultInitEnv PHPRC "/etc/php5/cgi"
    FCGIWrapper /usr/bin/php-cgi .php
    MaxRequestsPerProcess 1500
    MaxProcessCount      50
    IPCCommTimeout       240
    IdleTimeout           240
</IfModule>
```





# PHP op-code cache



- Op-code caches / Accelerators
- Reads, parses, and tokenizes scripts
- Stores the result in memory (op-code cache)
- Executes the code from the cache
- Significant resource savings for large sites





# PHP APC



- APC
  - The “official” code cache from the PHP development team
  - Constantly tracks the latest PHP versions
  - Installed via PEAR/PECL repositories





# eAccelerator



- eAccelerator
  - Used to be the fastest and uses least memory
  - No longer maintained
  - Instabilities (segfaults, WSODs)





# XCache



- Spin off from lighttpd web server
- Independent effort
- Maintained
- Debian/Ubuntu packages available
- Instabilities (segfaults, WSODs)







# Python



- mod\_python
- FastCGI as well





# Python Frameworks



- Zope, a framework
  - Grand daddy of all CMS
  - NATO uses it
- Plone CMS built on top of Zope
- Django framework
  - WashingtonPost
  - TorontoLife.com
- Pylons





# Python Frameworks



- CherryPy framework
  - [cuil.com](http://cuil.com)
  - Turbogears framework based on it





# Perl



- Used to be popular as a web site language in the CGI days
- Still in use by some sites (Slashdot)
- mod\_perl for embedding in Apache
- CMS based on Perl
  - Bricolage, popular for newspapers
  - WebGUI
  - Krang
  - Cyclone3





# Speed and scaling



- Speed: faster serving of pages
- Scalability: ability to handle more concurrent requests
- Use a PHP op-code cache/accelerators
- Page caching
  - Static HTML (Squid, Varnish)
- Object caching
  - Memcached
- [2bits.com](http://2bits.com), for more performance presentations





# Alternatives



Some components can be replaced for specific reasons







# Operating System



- WAMP
  - LAMP on Windows
  - Several ready to install stacks out there
  - One download
  - Good, because it encourages developers to cross into open source





# Operating Systems



- \*BSD
  - FreeBSD seems to be a common alternative
- Mac OS/X
  - MAMP stacks exist
  - Mainly for development





# Web servers



- Lighttpd
  - Nicknamed “lighty”
  - Fast and much less memory footprint
  - Memory leaks
- Ngnix
  - Same as lighty
  - Solves memory leaks
  - Still young, Russian documentation
  - Getting popular (shows on Netcraft)





# Web servers



- Good for serving static content
  - All images on a separate server that has the alternate light weight web server
  - Or use Apache in a reverse proxy, front ended by the other web server on different ports
- All run PHP as FastCGI
- Lack certain features (e.g. URL rewriting, .htaccess, ...etc.)





# Databases



- PostgreSQL
  - Very capable database
  - Community driven, not controlled by any company
  - Has transactional features by default
  - Not as easy to use as MySQL
  - Only recently incorporated replication
  - Upgrades mean export/import





# Languages



- One language that has been gaining traction in recent years is Ruby
- Ruby on Rails a much touted (hyped?) platform
- Twitter was on RoR, now uses other technologies for messaging
  - Scala: a functional language that uses the Java Virtual Machine







# Security



- Web applications *will* get remote probes *all the time*
  - Windows exploits
    - Worms targeting IIS and SQL Server (Code Red, SQL Slammer, Nimda, ...)
    - .DLL, vti, ...
    - Waste bandwidth and resources, but mostly harmless
  - Others are generic
  - Some are targeted to \*NIX systems
    - Attempted logins via SSH





# Types of exploits



- SQL Injection <http://xkcd.com/327/>
- Cross Site Scripting (XSS): code injection from other sites, mainly Javascript
- Cross Site Request Forgery (CSRF): tricking you into actions when you are logged in
- Remote File inclusion, can lead to Trojan
- Never trust user input, sanitize it
- [http://en.wikipedia.org/wiki/Category:Web\\_security\\_exploits](http://en.wikipedia.org/wiki/Category:Web_security_exploits)
- <http://code.google.com/edu/security/index.html>





# Best practice



- Run ssh on a non standard port (e.g. 4023)
- Watch your logs
  - The logwatch package emails you daily with a summary
- Check CPU usage, a spike can mean a trojan
- Install only what is needed, as little as possible
  - Your app is PHP? Don't install mod\_python
- Read a primer
  - e.g. <http://drupal.org/writing-secure-code>





# Best practice



- Consider Apache's `mod_security` (web application firewall)
- PHP Suhosin hardening patch (enabled in Ubuntu by default)
- Stay informed
  - Subscribe to the security mailing list of your distro
  - If on Debian/Ubuntu, install `apticron` and get an email next day updates are available





# Conclusion



- Success story for FOSS
- Powerful
- Capable
- Proven
- Low cost







# Resources



- 2bits.com
  - <http://2bits.com> Performance and Optimization
- Wikipedia
  - [http://en.wikipedia.org/wiki/LAMP\\_\(software\\_bundle\)](http://en.wikipedia.org/wiki/LAMP_(software_bundle))
- O'Reilly's ONLAMP <http://www.onlamp.com>
- Windows <http://www.wampserver.com>







# Discussion



Questions?

Comments?

