

10101010100100110101  
101111010100100101  
1011010100101010  
1010101010101  
1010101010  
1010100  
101

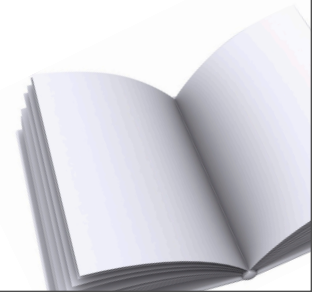
# When Your Pet Project Grows Up

Zack Huston & Robert Swarthout  
March 2007



Ummm....

We are just normal Joes sharing our experience. This is by no means the only way or even the correct way to do things. It works for us. Put these ideas into production at your own risk.



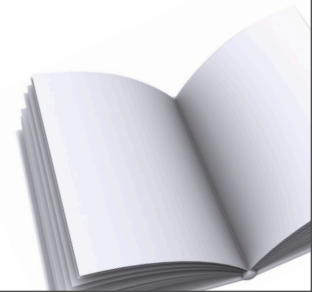
# Recap – Last Month

- IDEs Used
- Hardware
- Server Software
  - Apache
  - MySQL
  - PHP
- Safety and Security
  - Backups
  - Firewall
- Ease Your Pain
  - Server Administration
  - Utilities



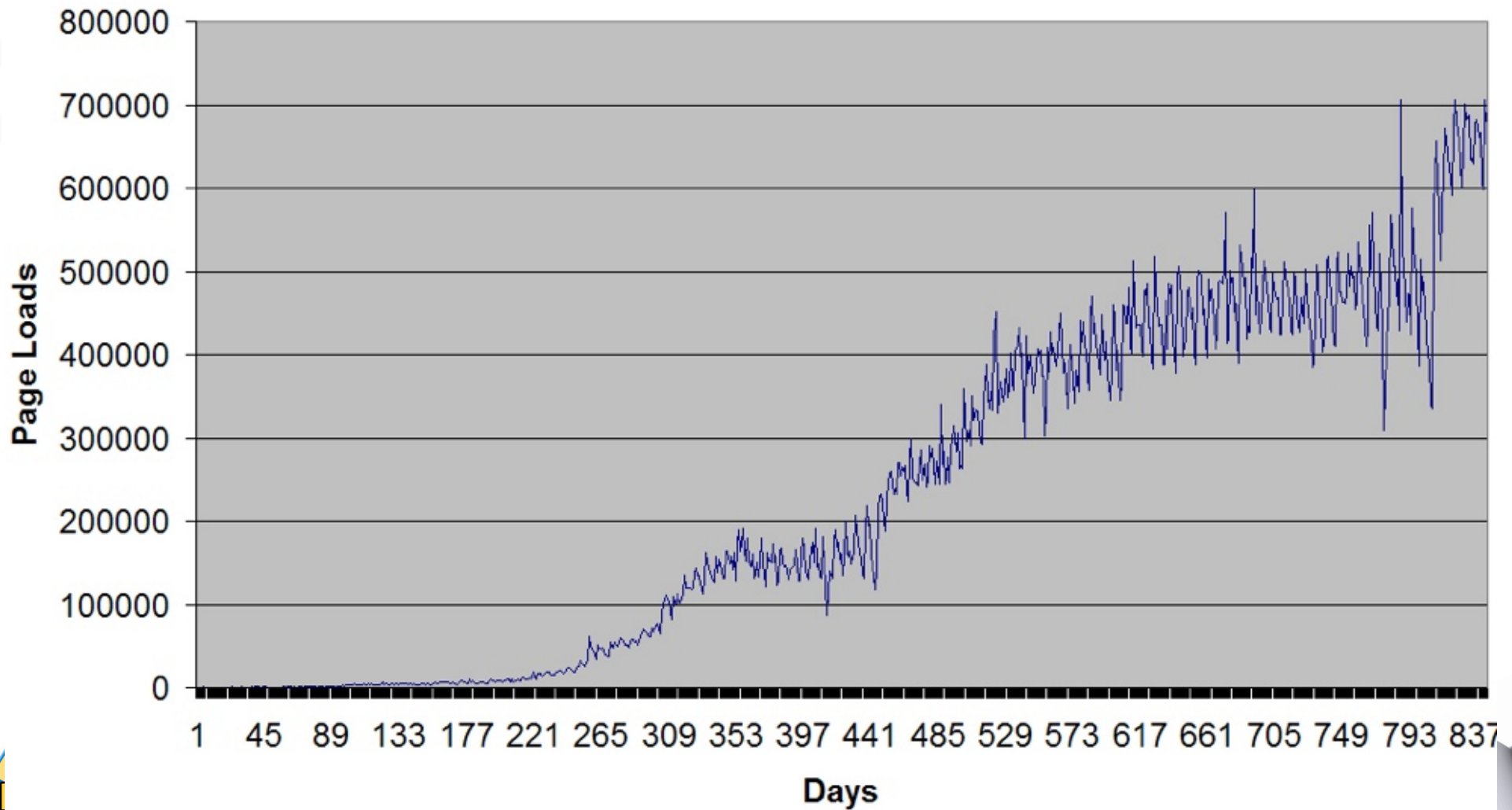
# It's Been A Busy Month

- Completed move to new servers
- Magazine publicity
  - Real Simple Magazine – March issue
  - Wired Magazine – March issue
- 1 million books received
- 1 million books posted



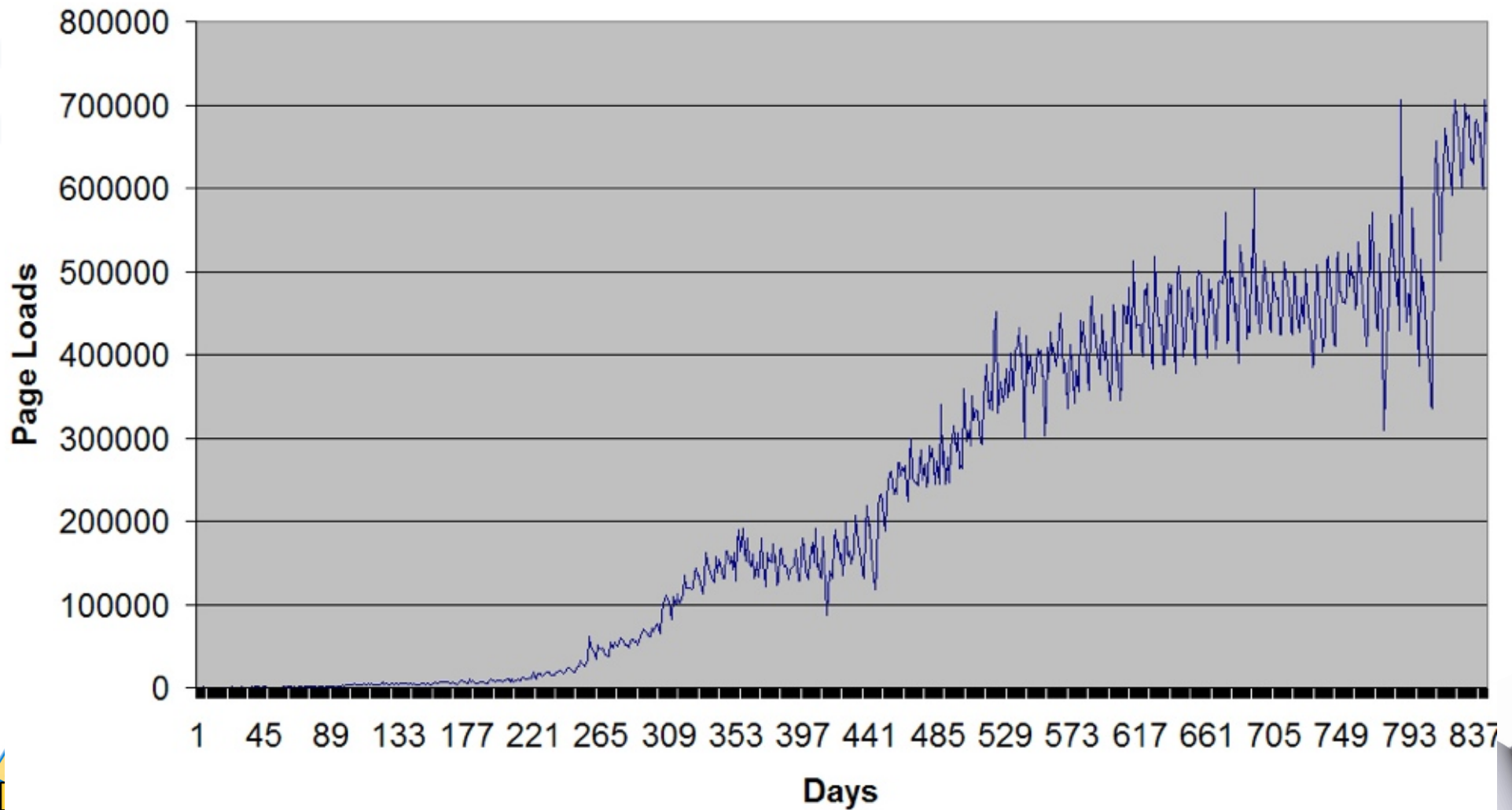
# PaperBackSwap Last Month

## Page Loads per Day



# PaperBackSwap This Month

## Page Loads per Day



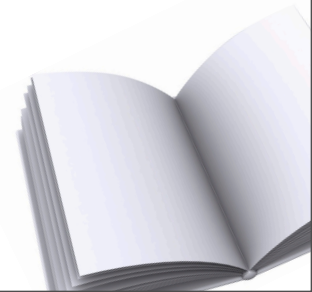
# Part 2

- Coding Practices
- Database Sessions
- Memcached
- MySQL Replication
- Subversion
- APC
- LVS (Linux Virtual Server)
- NFS (Network File System)
- Sendmail
- Redundancy



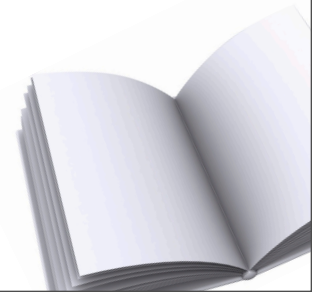
# Coding Practices

- PHP is our web and server scripts language, one common language
- Make things portable but don't get carried away
- Our code is designed for some reuse but it also designed specifically for our implementation
- Make functions to prevent code copying and pasting
- One way in, one way out
- Keeping code similar so that it can be used across PaperBackSwap, SwapaCD and any new sites



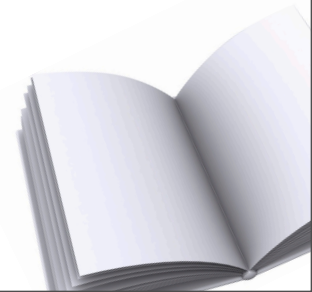
# Database Sessions

- Solves the problem of sharing session information across multiple web servers
- Easy to setup
- Uses PHP's custom session handler functionality
- Best to use InnoDB table engine
- Does not scale well
- Every page load has at least a one database read and write to your sessions table



# Memcached

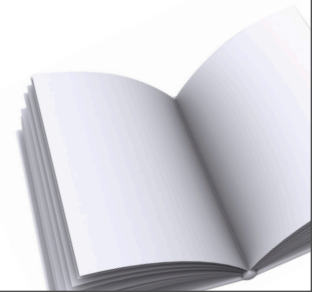
- *Memcached* is a high-performance, distributed memory object caching system, generic in nature, but intended for use in speeding up dynamic web applications by alleviating database load.
- Replacement for database sessions
- How does it work?
- Very fast, easy to use
- Only for use on internal networks (no user authentication)
- Developed by Danga.com (LiveJournal.com parent company), also used by Facebook.com
- Open source
- PHP PECL Extension, Written in C, \*unix only



# MySQL Replication

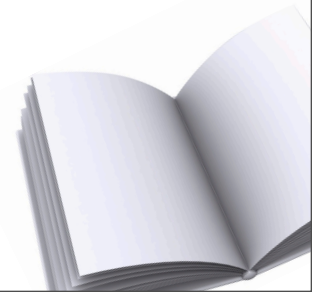
## Master > Slave

- Numerous possibilities for replication server topologies
- The master records all queries that modify the database to binary logs
- The slave processes the binary logs to stay in sync
- Great solution to assist in backing up a database
- Provides a hot spare if the master server fails



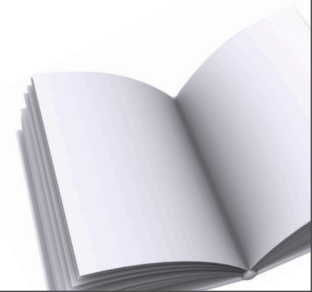
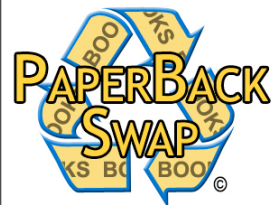
# MySQL Replication Dual Master

- Provides 2 servers that can share the load of updates and selects
- Code can be written to instantly fail over if one server goes down
- Can be very tricky to setup correctly
- Results from our efforts to use this topology
  - Servers do not stay in perfect sync
  - Temporarily not in use, pending further study



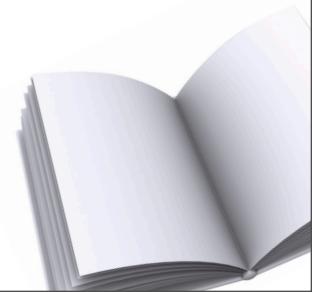
# Subversion

- Used to store revision history of code
- Highly recommend to use from the beginning of any project
- TortoiseSVN – windows client/server
  - <http://tortoisesvn.tigris.org>
- Now we know who to “blame” for bugs or alterations
- Each developer now has their own development environment
- Staging environment
- PHP front-end to subversion client? - server side
- Winmerge – helpful utility for comparing revisions
- <http://subversion.tigris.org/> and <http://svnbook.red-bean.com/>

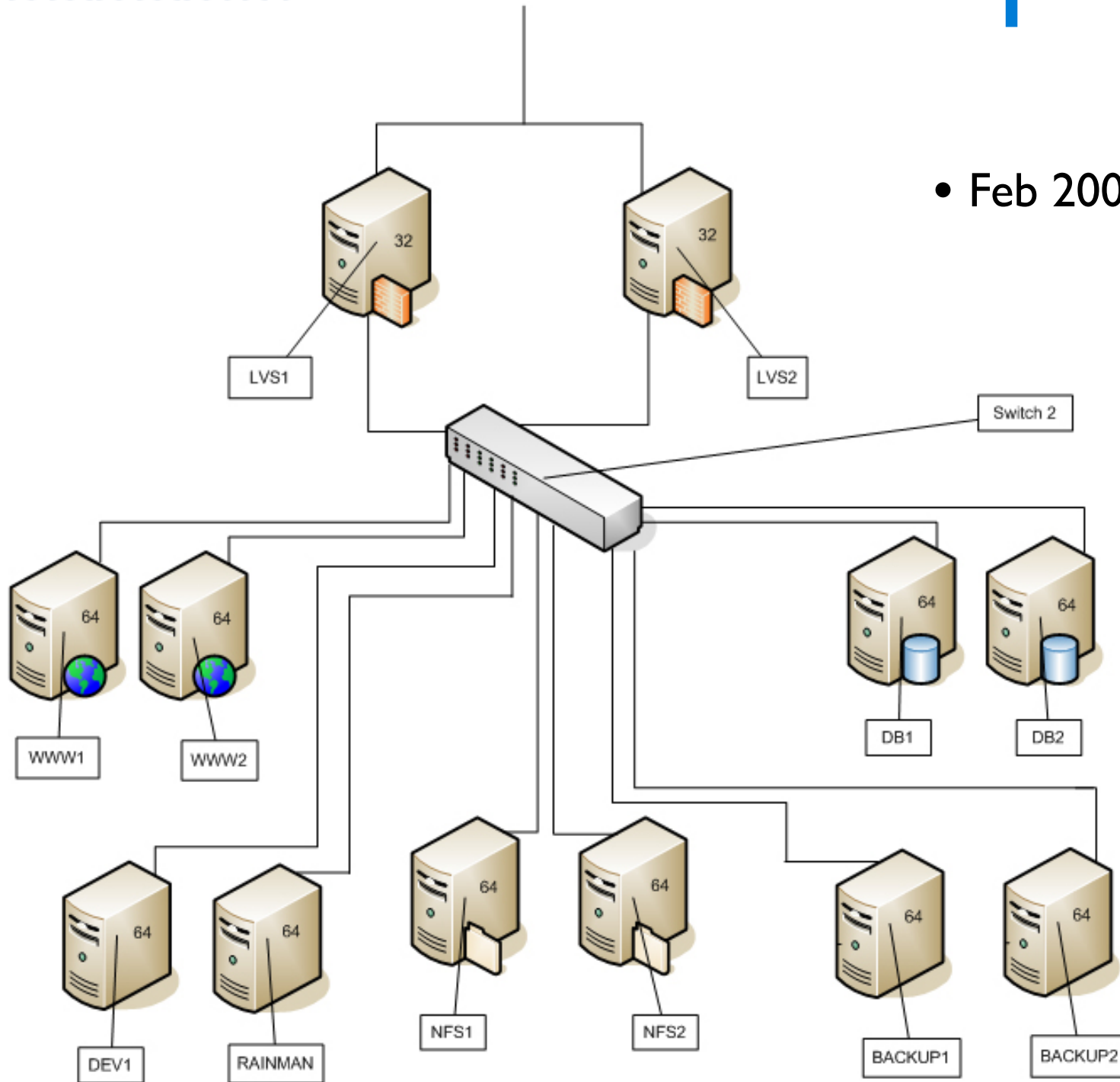


# APC

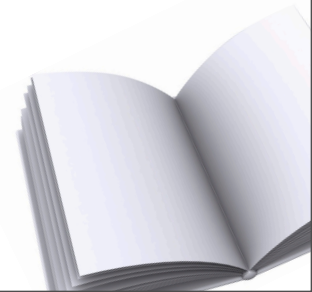
- Alternative PHP Cache - opcode cache for PHP
- Code caching typically increases server capacity 50%
- Used across all web nodes to cache compiled copies of code (in new setup)
- PHP PECL extension
- FREE!
- <http://www.php.net/apc> and <http://pecl.php.net/apc>



# Hardware – Setup

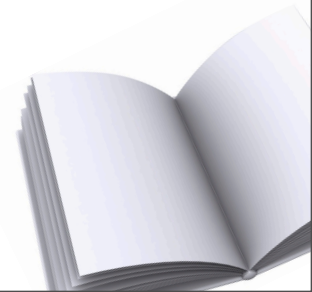


• Feb 2007 - ?



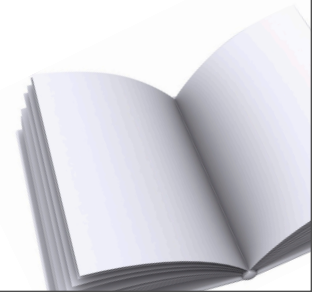
# LVS

- LVS stands for Linux Virtual Server
- Can load balance all network traffic to server cluster
- Serves as a partial firewall
- Heartbeat provides for redundant LVS servers
- Ldirectord monitors the health of internal servers
- Very flexible, scalable
- Used by hundreds of sites including Wikipedia.org, Real.com and Google.com



# NFS

- NFS stands for Network File System
- Samba/Windows Share with a twist
- Only available for the \*unix platforms
- Why do we use NFS?
- DRBD (Drive Block Device) - enables redundancy via network RAID
- Heartbeat – Allows for a hot spare
  - Starts needed daemons
- Multiple variations of NFS, some used by Flickr.com and YouTube.com



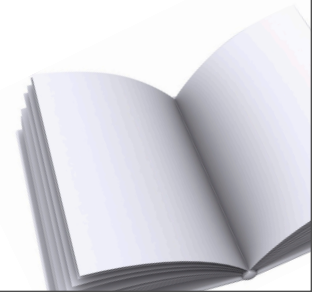
# Sendmail

- Configured on multiple servers to accept and send mail
- Incoming mail is load balanced to multiple servers by using the LVS
- Email is stored centrally on the NFS server
- DNAT (Destination Network Address Translation)
  - Allows all outgoing email to go out from one IP
  - Allows for the setup of a single reverse DNS entry



# Redundancy is key

- Downtime almost always cost you money in some way
  - Directly
  - Indirectly
- Allows for those times when traffic spikes exceed the capacity of a single server
- Provides the ability to fix or modify a server without taking the site down
- <http://www.linux-ha.org/> - High Availability Linux



# And the Fat Lady Sings...

Thank you  
for your  
time and attention!

Zack Huston and Robert Swarthout

[www.PaperBackSwap.com](http://www.PaperBackSwap.com)

[www.SwapaCD.com](http://www.SwapaCD.com)

