HUGE!

381 modules and 174GB database

Khalid Baheyeldin

http://2bits.com

Drupal Camp Toronto 2012

About Khalid

- 27 years in software development and consulting
- First computer: Sinclair ZX Spectrum
- Experience: Mainframe, UNIX
- Open Source: Linux, LAMP
- Full time open source developer, contributor and consultant

About Khalid (cont'd)

- Drupal since 2003
- Core features as well as 37+ contrib modules
- Member of the Advisory Board of the Drupal Association
- Co-Founder of the Waterloo Region Drupal Users Group
- Drupal talks at DrupalCons, DrupalCamps, DUGs
- Also Ontario Linux Fest, KW LUG
- Google Summer of Code (2005 to 2010)

About 2bits.com

- Founded in 1999 as a partnership, incorporated since 2007
- Drupal since 2003
- Services
  - Specialize in Drupal scalability and performance
    - Performance Assessment (various flavours)
    - Hosting provisioning, tuning and management
    - Drupal custom module development
- International clientele
- Extensive in depth articles and testimonials at http://2bits.com

How HUGE?

- 381 modules (more than we've ever seen or heard of)
- 178,446 MB database (174GB)
- 707 million rows total
  - 49.65 million nodes
  - 215 million rows (content_field_biblio_mesh_nids)
  - 117 million rows (content_field_mesh_qualifier_names)
  - 71.4 million biblio_contributor

Importing the database

2bits

Data format

- 18GB database dump (using mysqldump) compressed tar archive
- On a 3TB HPFS+ (Mac/OSX filesystem)
- Linux cannot mount HPFS+ file systems greater than 2TB!
- Find a Mac laptop ...
- Copy over the network to Linux
- Fast using Ethernet ...

Importing the database

Traditional Method

- Usual way, using the command:
  - gunzip -c db.sql.gz | mysql dbname
- 5 days for 84,449 MB to be imported
- Extrapolated, it would be 10+ days to complete
- Not linear, gets slower as it progresses

Importing the database

2bits

Maatkit mk-parallel-dump

- Perl script
- Deprecated, many warnings against using it
- Not in Maatkit's successor, Percona Toolkit

Importing the database

2bits

Mydumper / Myloader

- Written in C, by current and ex-MySQL employees
- Configurable number of parallel threads
- Recommended : one thread per physical CPU
- Uses standard mysqldump format, compressed:
  - One db.table-schema.sql.gz
  - Multiple db.table.0000n.sql.gz
- Full article on http://2bits.com

Importing the database

2bits

Mydumper / Myloader (cont'd)

- Version in Ubuntu 12.04 LTS repository does not work (segmentation fault)
- We compiled it from source
- Also, needs patching (details in article)
- Load time: 3 days, 16 hours, 15 minutes

Importing the database

2bits

Drop Indexes Before Loading

- Create tables as usual using db.tbl-schema files.
- Wrote a script to massage db.table-schema files, and remove the secondary indexes, using ALTER TABLE DROP INDEX ...
- Load time: 8 hours 15 minutes!
- Adding the indexes: 6 hours 32 minutes
- Total: 14 hours 47 minutes!

Importing the database

2bits

Barracuda File Format

- Default file format is “antelope”
- Changed by: innodb_file_format=barracuda
- Load time: 10 hours 30 minutes
- Adding the indexes: 6 hours 52 minutes
- Total: 17 hours and 22 minutes
- Database size is 165,955 MB

Importing the database

2bits

Compressed Row Format

- Must have: innodb_file_format=barracuda
- For each table: ALTER TABLE ... ROW_FORMAT=COMPRESSED KEY_BLOCK_SIZE=8
- Load time: 20 hours 8 minutes
- Adding the indexes: 9 hours 2 minutes
- Database size is only 87,000 MB!
- Think Solid State Disks (SSDs)

Solid State Disks

2bits

- IOPS (I/O operation per second)
  - 7200 RPM SATA disks: 100 IOPS
  - 15K RPM SAS: 210 IOPS
  - Desktop SSD: 50,000 to 80,000 IOPS
  - Enterprise SSD (PCI-e): 300,000 to 1,200,000 IOPS
- Marketing figures (maximums, idealized tests)
- Watch for limitations elsewhere (SATA chipset)
- Finite number of read/write cycles, wear levelling
- Expensive

Side Benefit: Backup

2bits

- Using Mydumper, you can cut down backup time drastically
- Because of parallel processing, and chunking large tables
- On another site, down from 8 minutes to 35 seconds!

Side Benefit: Backup

2bits

- Tables in chunks
  - 1.5M live.nodewords.sql.gz
  - 1.9M live.node_revisions.sql.gz
  - 2.4M live.search_index.sql.gz
  - 3.7M live.profile_values.00000.sql.gz
  - 3.7M live.profile_values.00001.sql.gz
  - 6.6M live.votingapi_vote.00000.sql.gz
  - ...
  - 8.5M live.votingapi_vote.00010.sql.gz
  - 12M live.sessions.sql.gz
  - 20M live.users.sql.gz

Importing the database

2bits

Helpful configuration params

- innodb-file-per-table
- innodb_flush_method = O_DSYNC
- innodb_flush_log_at_trx_commit = 2
- Up to 10X improvement for write performance
- Comment out these, since they increase writes to disk
  - log_slow_queries
  - log_bin

Importing the database

2bits

MySQL Parallel Script

- By Andrew Berry
  - https://github.com/deviantintegral/mysql-parallel
- Wrapper Script
- Requires GNU Parallel
  - Not in Ubuntu's repo 12.04 LTS
- Parallel Bzip2 (optional)
- Does not break down large tables into chunks

Modules

2bits

- 381 modules enabled
  - 18 core modules
  - 36 custom modules
  - 92 exported feature
  - 16 contrib modules, forked
  - Rest are contrib modules

Modules

- 760 files are being loaded for each request
  - Think require_once() from ctools, views, ...
  - APC helps with that, but still, lstat() calls are done to each file to check its timestamp
- 10-20% of page load time was just file loading!
- Eating CPU
  - 27,782 calls to module_implements() per page request!
  - 23,243 calls to module_hooks()

Modules

- Disable stuff like:
  - devel
  - table_wizard, schema
  - memcache, memcache_admin
  - views_ui
  - admin_menu
- Other sites (two recent sites, unused)
  - context, context_ui, context_layouts

Modules

2bits

- The good news:
  - Site was able to reduce the number of modules from 381 to 207 only!
  - Script to aggregate features in less files than what they are exported in.
  - Plan to contribute the script

Views

2bits

- Make sure views are cached, goes a long way ...
- Consider reducing number of items displayed in a view
  - Users are not going to read through a list of 50 nodes, 10 or 15 are sufficient
  - Reduces rendering time
  - Reduces memory usage per page

Features

- 92 exported modules
- Many of them have a .module that is empty, or only does a require_once() to a .inc.
- Could be combined
  - “mega feature” (cf. University of Waterloo)
  - Creative post-processing script

Memcache

2bits

- You don't need the “memcache” nor the “memcache_admin” modules enabled, just the include in settings.php
- Make sure you have PHP memcache 2.2.6 or later
- Add this to php.ini, if you use more than one memcached server
  - memcache.hash_strategy = "consistent"

Page Cache vs. Varnish

2bits

- Observed 3 modules that have variations of these:
  - `$conf['cache'] = FALSE;`
  - `$GLOBALS['conf']['cache'] = CACHE_DISABLED;`
- Attempt to disabling page cache: “flag”, “biblio”, “invite”
- Means some pages will not be cached for anonymous users

Logging

- As recommended for large sites, use syslog and disable dblog.
- However, a subsite shared the database, and had a prefixed *xyz_watchdog* table in the same database
- A special use script that ran daily did not hush PHP notices for each field
  - Undefined property: stdClass::$foo
- Millions of rows logged to *xyz_watchdog* in the same database!

Node Access

2bits

- No node access modules used on the site
- However, some were enabled in the past and disabled, causing over 6 million rows to still be present in the *node_access* table.
- The table needs to be reset back to one row, its pristine state

Elysia Cron

2bits

- Allows fine grained tuning of cron
- How frequent individual cron hooks run, and when
- For example: Can defer slow hooks to after midnight
- On Drupal 7, set the internal cron to “Never”, rather than the default, and run cron externally

PHP network calls

2bits

- Site uses Sphinx
- Multiple calls per page load
- Can add up to 300 milliseconds
- Frequent issue with sites we investigate

Fast 404

2bits

- Built in feature in Drupal 7
- Prevents Drupal from booting for static files (.jpg, gif, .png, .css, .js, ...)
- Uncomment it in your settings.php

Hosting

2bits

- Acquia, which is Amazon AWS based
- No access to the database server
- Usual “Cloud” pros and cons
  - Disks cannot be virtualised (bad neighbors? Can't be sure)
  - Variability in performance (good, good, sloooow, good, ...)
  - Gets expensive as you grow
  - Amazon AWS downtime (3 major incidents in 2012, so far!)
- Evaluating alternatives (e.g. dedicated hosting)

What is this site?

2bits

Anyone interested to know?