DIY Drupal 7 Performance

Mike Carper omcarper (mikeytown2)



Pacific Northwest Drupal Summit October 2015

Follow along https://goo.gl/L4P3MA



https://goo.gl/L4P3MA

About Mike

- November 2008 - Started using Drupal

- January 2009 - Discovered Boost module

- September 2009 - Boost 6.x-1.0 released

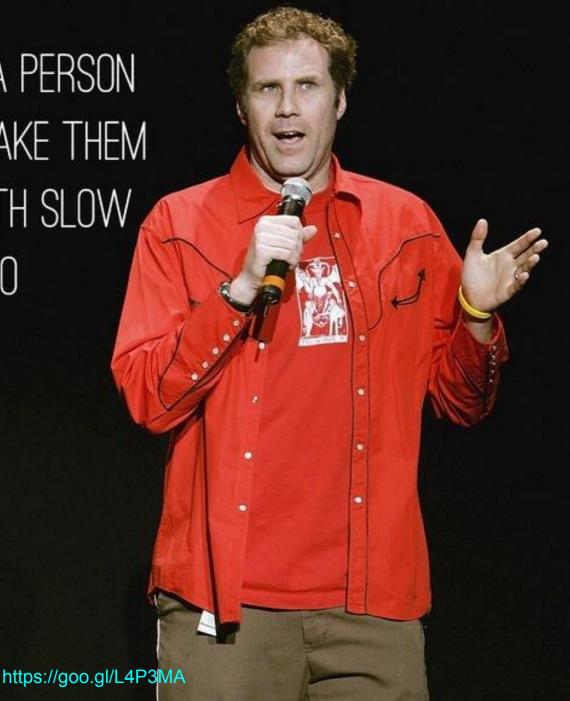
- November 2009 - Started @ Datasphere

About Mike (@ Datasphere)

- Our team scaled a single Drupal 6 domain access multisite up to 1,300 different domains
- Our Drupal 7 site is 99% authenticated with a nothing shared between users, making caching very hard to do
- Created a lot of performance related modules that helped us out a lot

"BEFORE YOU MARRY A PERSON
YOU SHOULD FIRST MAKE THEM
USE A COMPUTER WITH SLOW
INTERNET TO SEE WHO
THEY REALLY ARE."

WILL FERRELL





What can we do to make sure users enjoy using your site?

Make sure it's FAST



For logged in & logged out users

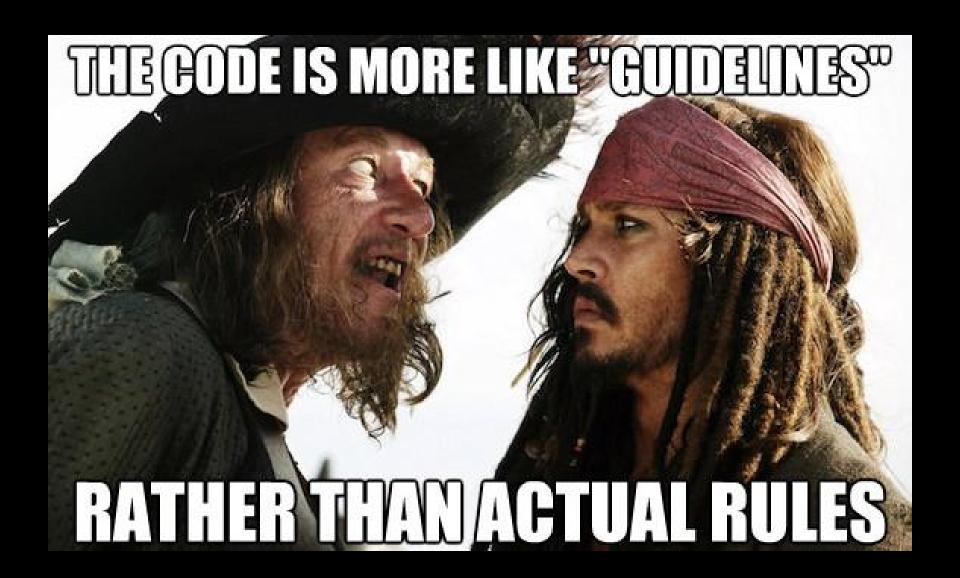
How?

Throw Money At



Or you can DIY!

Some rules to follow



General Guidelines

Anonymous visitor caching is easy.

Big performance gains come asynchronous/parallel and deferred code execution. Why do it now when you can do it later.

Less is better.

General Guidelines (cont'd)

Big improvements can be found on the frontend.

If you have to pick scalability over performance 9 times out of 10 pick scalability; you'll thank yourself later on when you have more than 20 users logged in at once.

Measure changes - Was that good or bad?

Page Caching

mainly useful for anonymous users

Caching to the rescue!



Popular Page Cache Options

Drupal Core

Boost

Memcache/Redis & Varnish

Drupal Core

Advantages:

Comes with core

Better than nothing

Disadvantages:

Relies on the database

Slow as PHP is needed



https://goo.gl/L4P3MA



Boost

Advantages:

 Works on almost all shared hosting

 Fast as no PHP is executed

Disadvantages:

 Need to copy in htaccess rules

 Cache is not easily shared with multiple servers



Memcache/Redis & Varnish

Advantages:

Disadvantages:

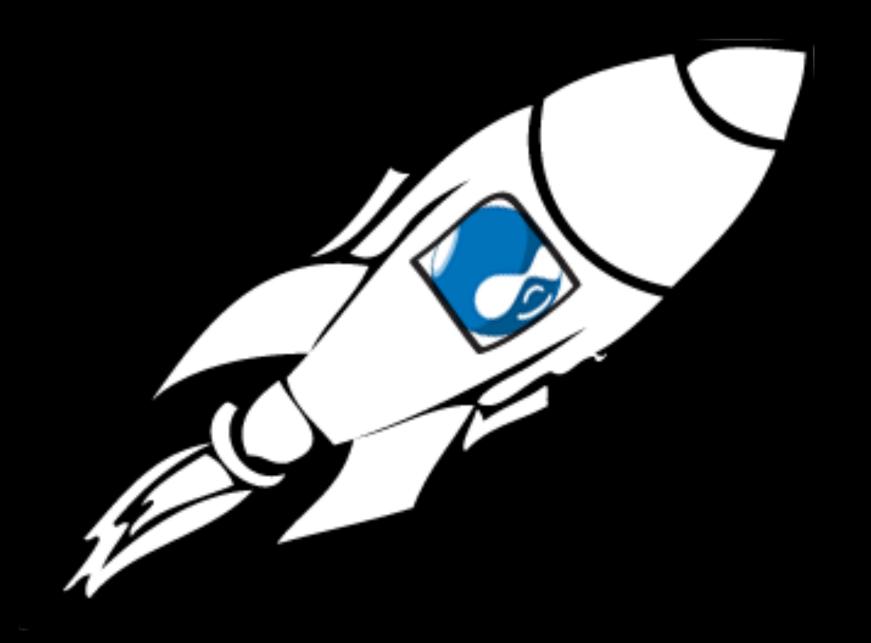
Fast

Need shell access

Can help logged in users as well

Not simple to setup

 High availability is hard



Optimization and Tuning of Drupal 7

Optimization Overview

Core page and block caching Core CSS & JS aggregation Panels & views caching Disable unnecessary modules Modules that usually help Poormanscron in core Enable fast 404 Core patches

Optimization Overview (cont'd)

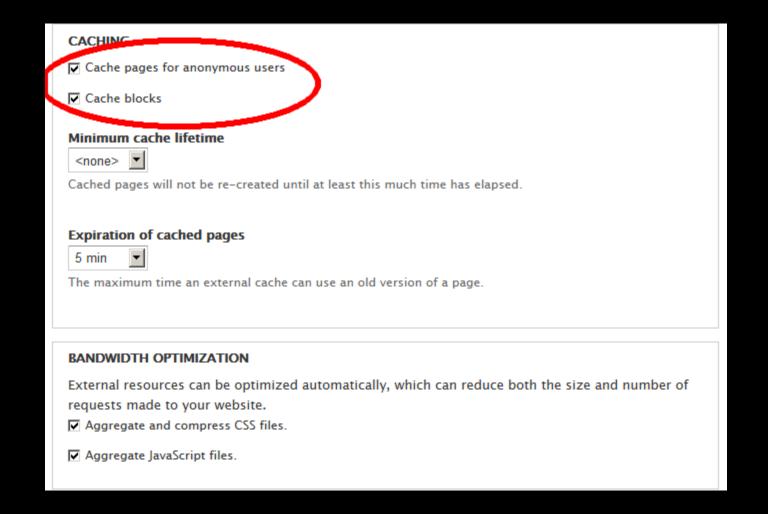
Database Connection Op Code php.ini PHP Notices/Warnings Basic database tuning **Deadlock Detection** Diagnosing Stuck Queries Slow Query Analysis Cachegrind

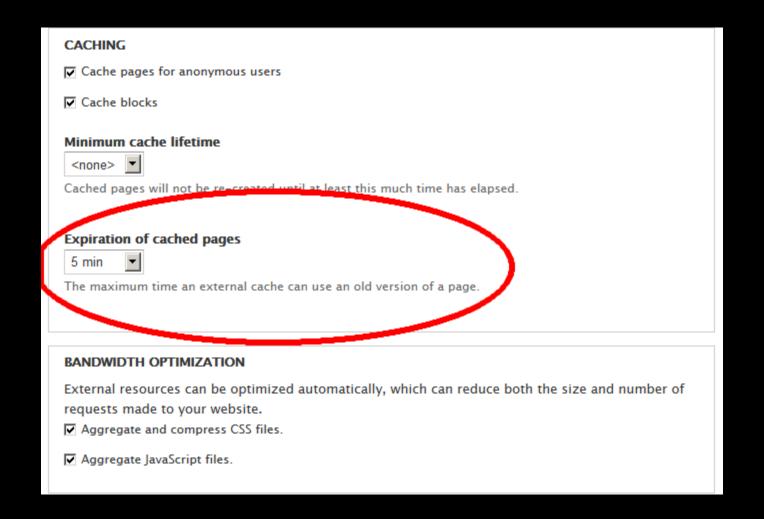
TLDR



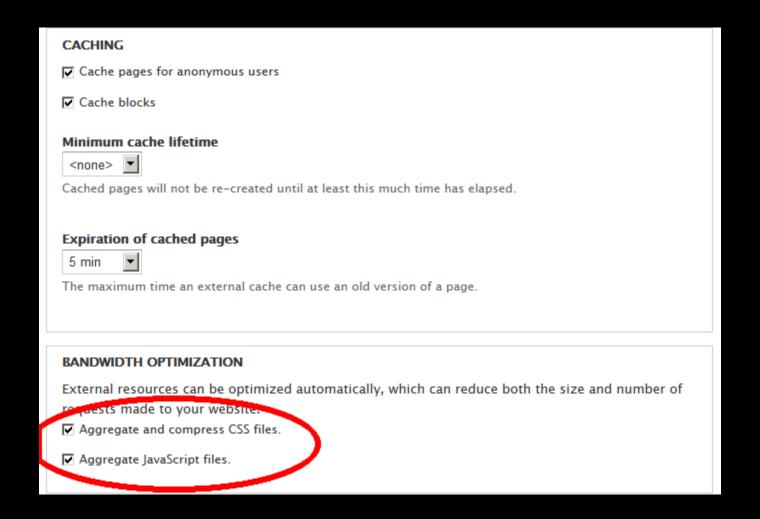
https://goo.gl/L4P3MA

Core and Block Caching

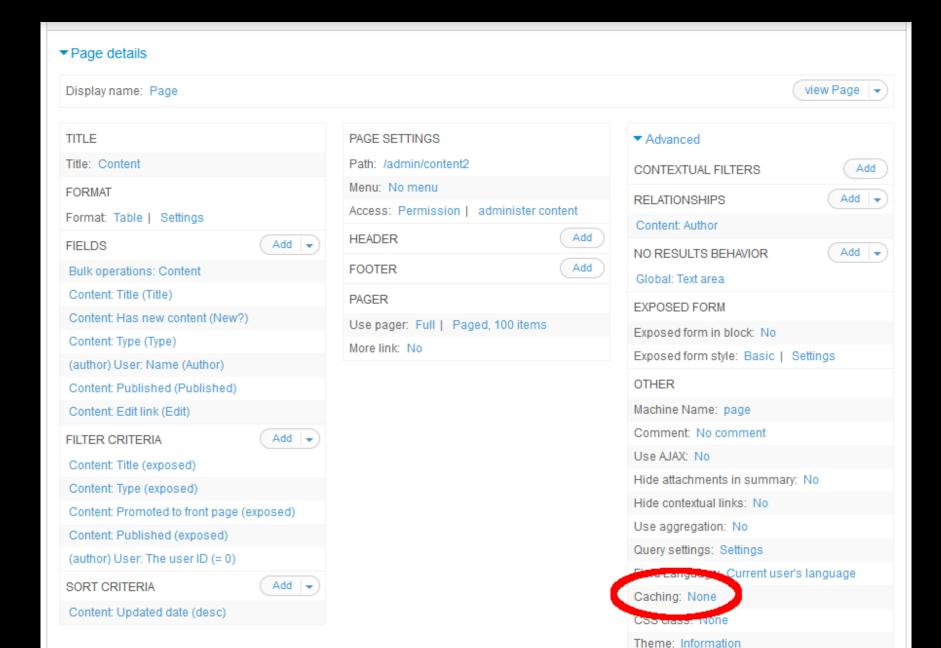


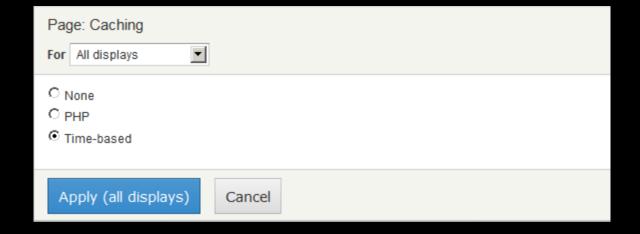


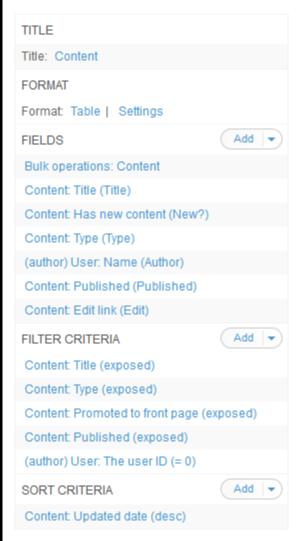
Enable CSS & JS Aggregation

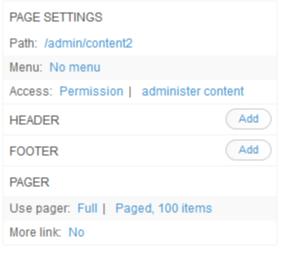


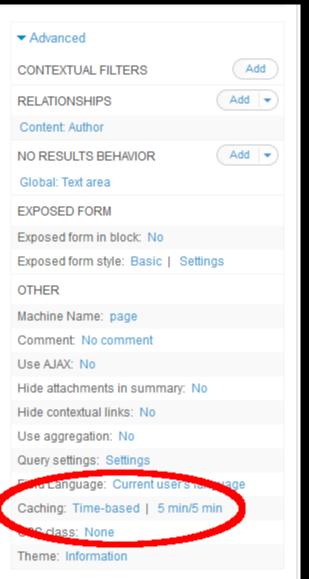
Views Caching



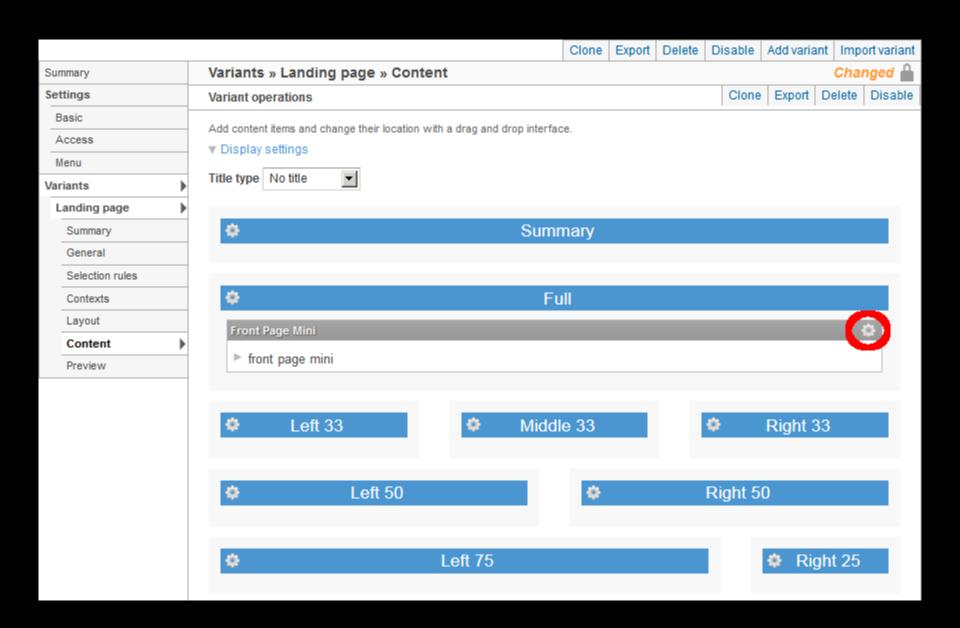


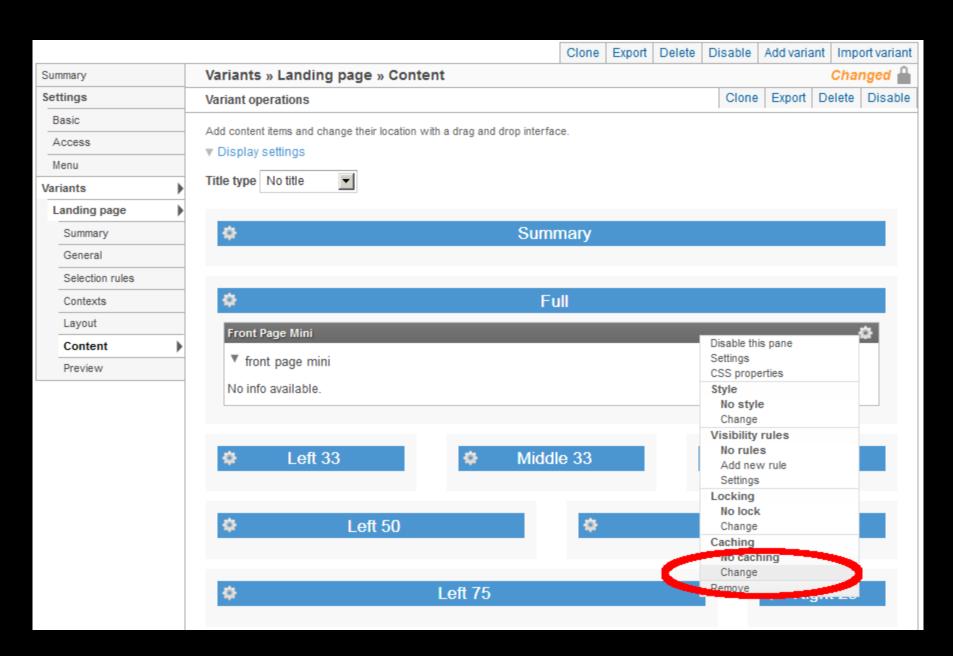






Panels Caching

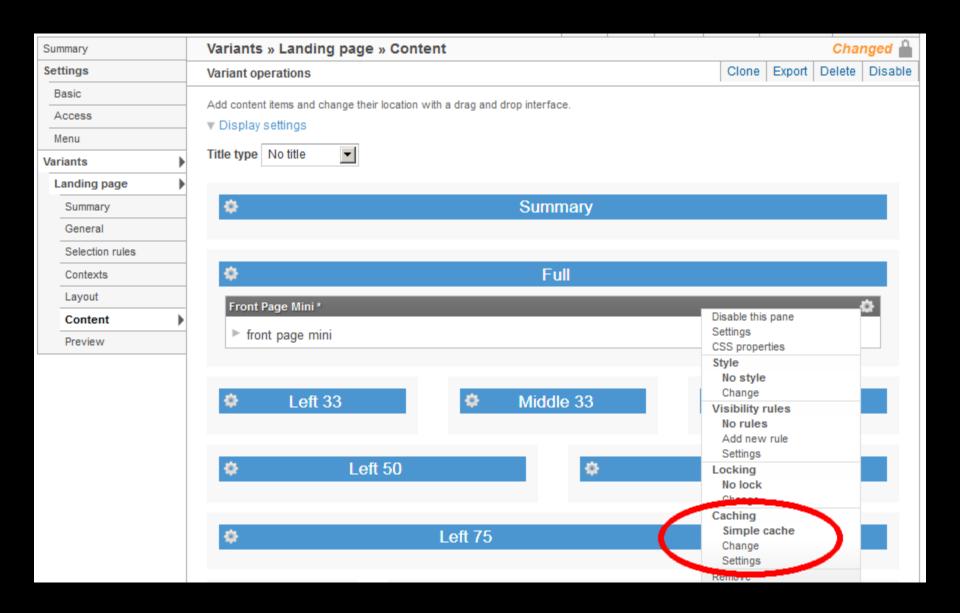




Method

- C No caching
- Simple cache

Next



Disable unnecessary modules







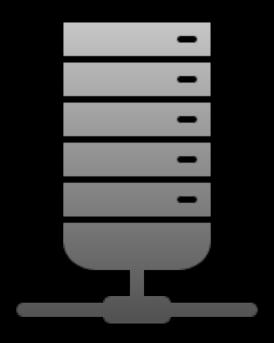


Most modules behave well; some do not. The poor performing modules are usually the ones without a lot of users.

The next release of D7 (7.40) contains a big fix that should speed up most drupal installs by 5-10%; the module_implements() function has been improved in core.

Modules that help with optimizing Drupal

Backend



https://goo.gl/L4P3MA

Entity Cache Module

A lot more useful if using memcache https://drupal.org/project/entitycache

Render Cache

Caches rendered entities currently https://www.drupal.org/project/render_cache

Cache Expiration

Works with the page cache https://www.drupal.org/project/expire

Imageinfo Cache

Generate image styles on file upload https://www.drupal.org/project/imageinfo cache

Asynchronous Prefetch Database Query Cache

Fixes everything that is wrong with the core database cache; MySQL only https://www.drupal.org/project/apdqc

HTTP Parallel Request & Threading Library

Doesn't do much out of the box but has some awesome tools when creating custom code https://www.drupal.org/project/httprl

Frontend



Frontend

I always test changes made with Google's PageSpeed Insights and webpagetest.org

I set the connection speed to 3G or slower when testing on webpagetest.org

ImageAPI Optimize

Optimize the image when it is saved https://www.drupal.org/project/imageapi optimize

Defer Image

Prevent images from slowing down the initial load and render of a page https://www.drupal.org/project/defer_image

Font Awesome SVG

Use font awesome? Use this to reduce the bytes downloaded.

https://www.drupal.org/project/fontawesome_svg

AdvAgg

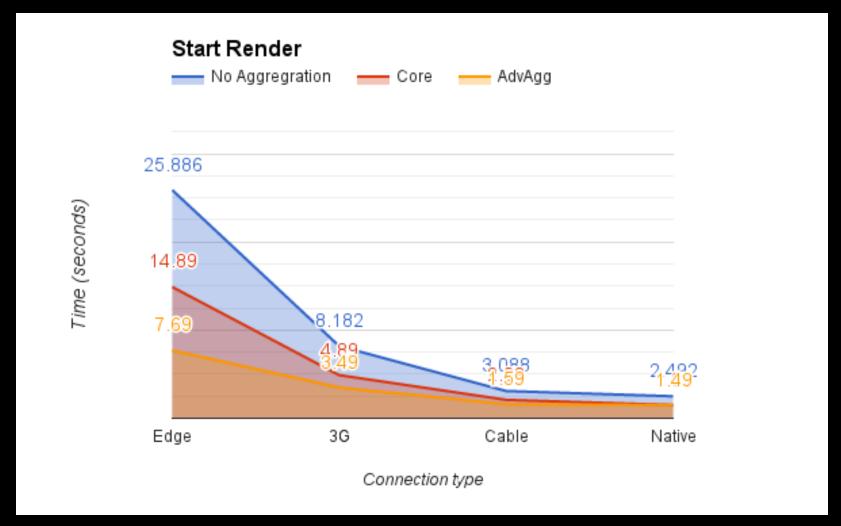
Faster CSS/JS aggregation https://www.drupal.org/project/advagg/

More info on AdvAgg

Drupal.org started using it this year. According to Google Analytics the average UC Browser page load time went from 13.3 seconds down to 5.4 seconds.

UC Browser is a mobile browser with a huge user base in China, strong adoption in India and continued growth in emerging regional markets.

Stats from https://www.drupal.org/project/drupal The slower the connection the bigger an effect AdvAgg has on improving the front end responsiveness.



AdvAgg D.O Settings.

- Use Aggressive Render Cache
- Move all inline scripts to the bottom of the execution order & Move all external scripts to the top of the execution order
- Use DNS Prefetch for external CSS/JS
- Combine CSS files by using media queries
- Move JS to the footer All but what is in the \$all_in_footer_list
- Deferred JavaScript Execution: Add The defer Tag To All Script Tags - All but external scripts

https://goo.gl/L4P3MA

AdvAgg D.O Settings. (cont'd)

- Deferred inline JavaScript Execution: Put a wrapper around inline JS so it runs from a setTimeout call
- Remove ajaxPageState CSS and JS data if ajax.js is not used on this page
- Change CSS bundles to 2
- Change JS bundles to 5
- Move Google Analytics analytics.js code from inline to be a file
- Prefetch stats.g.doubleclick.net/robots.txt

Other AdvAgg settings

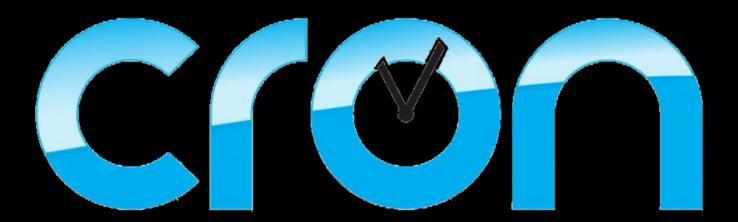
- Use JSMin
- Use Async Font loading

Future AdvAgg settings

- Inline critical css

https://www.drupal.org/node/2493801

Poormans cron in core



https://goo.gl/L4P3MA

admin/config/system/cron

Cron takes care of running periodic tasks like checking for updates and indexing content for search.

Run cron

Last run: 1 hour 8 min ago.

To run cron from outside the site, go to http://127.0.0.1/drupal7/cron.php?cron_key=3RW5NgKZ2obnWOYjloc93i2LRikY_Mzksjbpdle5Y4Q

Run cron every

3 hours

_

Save configuration

Follow a guide on how to setup cron on your server. Almost all shared hosting allows for cron jobs to be scheduled.

Even bad hosts have instructions



> Help > Web Hosting > Databases > Accessing Your Cron Job Manager with Shared Hosting

Accessing Your Cron Job Manager with Shared Hosting

Cron is a standard Linux feature that lets you schedule tasks, called "Cron Jobs," to run unattended at a specified frequency.

- To Create a Cron Job
 - 1. Log in to your Account Manager.
 - Click Web Hosting.

Click one of the following, based on the type of hosting you have. You can tell this by the second word in your hosting's description beneath your hosting account's domain name, e.g. Deluxe Web correlates to Web/Classic (Hosting Control Panel).

Web/Classic (Hosting Control Panel)

- 1. Next to the hosting account you want to use, click Launch.
- In the Tools section of the Hosting Control Panel, click the Cron Job Manager icon. This action displays the Cron Job Manager screen.
- Click edit next to the Email Address heading in the left navigation area.
- 4. Enter the email address to which you would like to have all information sent regarding this Cron Job.
- Click Save.
- Click Create Cron Job.
- 7. Select and enter the options you want to use, or click Custom if you want more granular settings.
- Click Save.

admin/config/system/cron

Cron takes care of running periodic tasks like checking for updates and indexing content for search.

Run cron

Last run: 1 Lour 8 min ago.

To run cron from outside the site, go to http://127.0.0.1/drupal7/cron.php?cron_key=3RW5NgKZ2obnWOYjloc93i2LRikY_Mzksjbpdle5Y4Q

Run cron every

3 hours

Save configuration

Once you have your host running the drupal cron job you can disable the one that comes with core by default

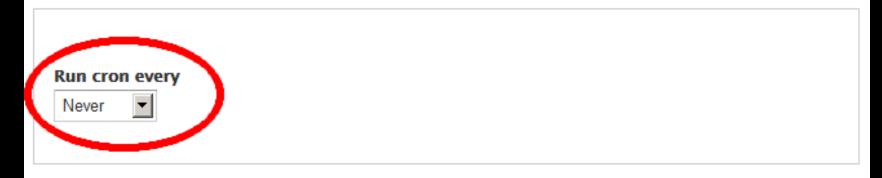
admin/config/system/cron

Cron takes care of running periodic tasks like checking for updates and indexing content for search.

Run cron

Last run: 1 hour 8 min ago.

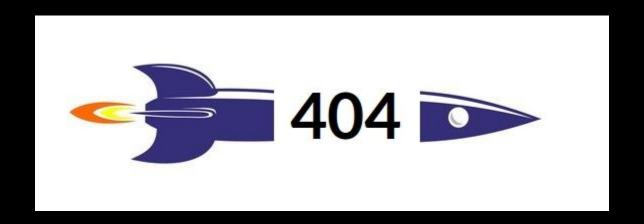
To run cron from outside the site, go to http://127.0.0.1/drupal7/cron.php?cron_key=3RW5NgKZ2obnWOYjloc93i2LRikY_Mzksjbpdle5Y4Q



Save configuration

Enable fast 404

Unless you're using the private file system



Go to the sites/default directory
Open up to the setting.php file
Around line 500 uncomment this
drupal_fast_404();
by removing #

Also use fast 404 to block urls commonly used by bots

Requests like /postnuke/article.php should be 404-ed ASAP. Do not waste server resources on serving a nice 404 to bots.

While inside the settings.php file above where we enabled fast 404 are the settings that control it. add

```
^(postnuke|wp-content|mailman|phpBB)|
to the start of the $conf['404_fast_paths']
variable, right after the first /
```

This means that if any path starts with any of these strings we'll skip booting drupal and 404 right here taking very little server resources.

In your settings.php Replace this:

```
$conf['404_fast_paths'] = '/\.(?:
txt|png|gif|jpe?
g|css|js|ico|swf|flv|cgi|bat|pl|dll|exe|asp)
$/i';
```

With this

```
$conf['404_fast_paths'] = '/^(postnuke|wp-
content|mailman|phpBB)|\.(?:txt|png|gif|jpe?
g|css|js|ico|swf|flv|cgi|bat|pl|dll|exe|asp)
$/i';
```

You can of course add more paths to this as you discover the other paths used by bots.

Lets say you also want to fast 404 any path that starts with browserconfig.xml. This

(postnuke|wp-content|mailman|phpBB)

Now should be this

(postnuke|wp-content|mailman|phpBB|browserconfig\.xml)

Be careful when doing this; you can 404 a lot of paths.

Core Patches

Hacking core for fun and profit https://groups.drupal.org/node/210683



Add static cache to module_load_include()

https://www.drupal.org/node/1443308#comment-10329229

Usually shaves 90ms off of a big site

If item is hidden in _menu_tree_check_access() skip it right away

https://www.drupal.org/node/1710656#comment-6304412

Usually shaves 50ms off of a big site

Avoid re-scanning module directory when a filename or a module is missing

https://www.drupal.org/node/1081266? page=1#comment-10434917

If you have one missing module then this can shave up to a second off of every page load

Improve theme registry build performance

https://www.drupal.org/node/2339447#comment-10354975

Usually shaves 5-10 seconds off of a cache clear

Speed up drupal_parse_info_format()

https://www.drupal.org/node/2146643#comment-9592921

Reduces memory usage and shave a second off of cache clears

inline file_uri_scheme() in file_stream_wrapper_uri_normalize()

https://www.drupal.org/node/1443342#comment-5613306

Shaves 150ms on a cache clear

Fix Notice: Trying to get property of non-object in image_style_deliver()

https://www.drupal.org/node/1762772#comment-7387274

Needed if the files directory is mounted on a NFS drive.

Database Connection

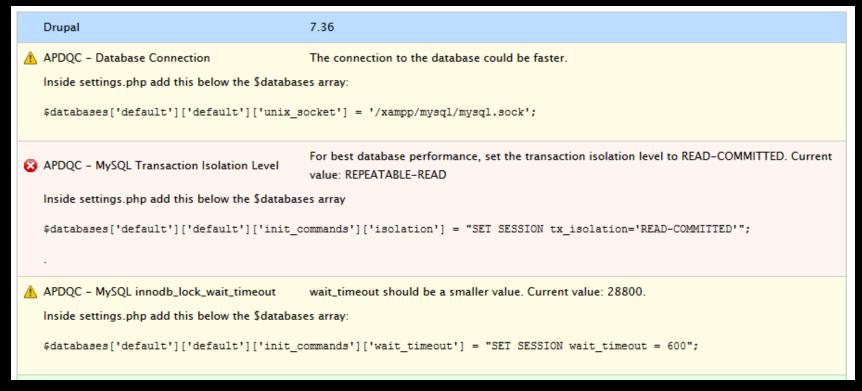
Use unix sockets if possible



https://goo.gl/L4P3MA

Use the APDQC Module

The status report has a section on setting up unix sockets or an IP instead of hostname.



Op Code Cache

Use the built in OPcache in php 5.5+



OPcache settings

```
zend_extension=opcache.so
opcache.enable=1
opcache.memory_consumption=256
opcache.validate_timestamps=0
opcache.max_accelerated_files=100000
opcache.fast_shutdown=1
opcache.interned_strings_buffer=16
```

php.ini



For all

```
realpath_cache_size = 1M
realpath_cache_ttl = 3600
```

Only if memcache is being used

memcache.hash strategy="consistent"

PHP Notices/Warnings

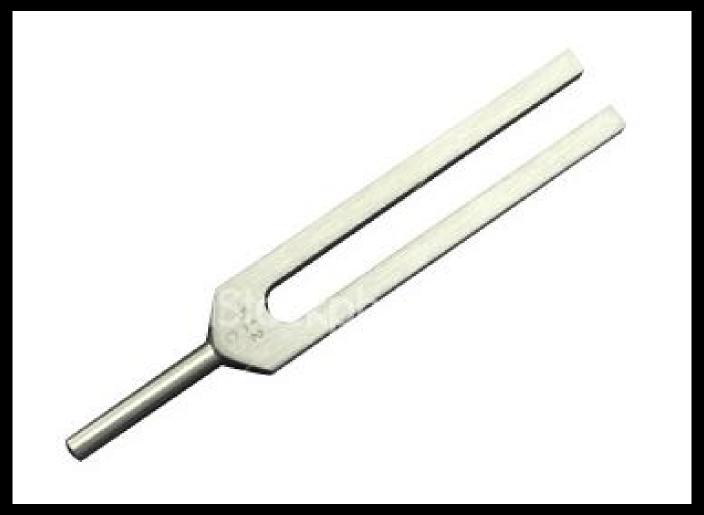
Fix them as these slow down your site



Put this in your settings.php file

```
<?php
if (!empty($ GET['errors'])) {
  // Show Errors in output
  ini set('display errors', '1');
  // Report all php errors.
  error reporting (-1);
  // Display errors using dsm().
  $conf['error level'] = 2;
```

Basic database tuning



Use the APDQC module

Status report is your friend

What it will check for

- That you're using the mysqlnd db driver
- using a unix socket if at all possible
- Verify tx_isolation is READ-COMMITTED
- max_allowed_packet is at least 32MB
- innodb_buffer_pool_size is big enough for your database to fit into memory
- Verify innodb_flush_log_at_trx_commit is 2
- Query cache is disabled
- innodb_lock_wait_timeout and wait_timeout are both sane values

Note about the MySQL query cache

This should be disabled.

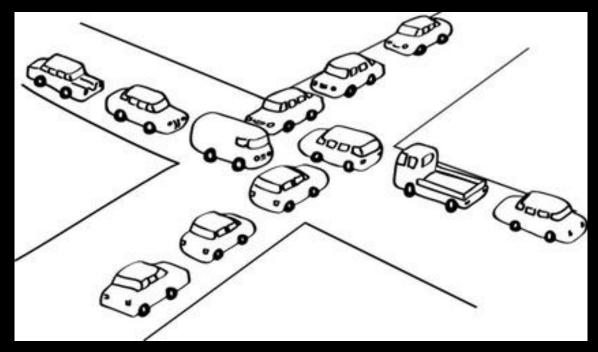
The built in MySQL query cache gets invalidated at the table level, any writes to that table will wipe out the query cache; it also has some nasty global locks on reads and writes.

More info about MySQL & InnoDB

http://www.mysqlperformanceblog.com/2013/09/20/innodb-performance-optimization-basics-updated/

Deadlock Detection

"deadlock found when trying to get lock try restarting transaction"



https://goo.gl/L4P3MA

SHOW ENGINE innodb STATUS

This command will output a bunch of information; the more interesting bit is under the "LATEST DETECTED DEADLOCK" section.

Diagnosing Stuck Queries

What if SHOW PROCESSLIST shows a lot of simple queries that are taking too long

Processes	ID	User	Host	Database	Command	Time	Status	SQL query ← T→
Kill	236	root	localhost:51083	d7	Sleep	3		
Kill	237	root	localhost:51084	d7	Query	3	update	<pre>INSERT INTO cache (cid, serialized, created, expire, data) VALUES ('node_types:en', '1', '1415747825</pre>
Kill	238	root	localhost:51085	d7	Query	3	update	<pre>INSERT INTO cache (cid, serialized, created, expire, data) VALUES ('node_types:en', '1', '1415747825</pre>
Kill	239	root	localhost:51086	d7	Query	3	update	INSERT INTO cache (cid, serialized, created, expire, data) VALUES ('node_types:en', '1', '1415747825
Kill	240	root	localhost:51087	d7	Query	3	update	<pre>INSERT INTO cache (cid, serialized, created, expire, data) VALUES ('node_types:en', '1', '1415747825</pre>
Kill	241	root	localhost:51088	d7	Query	3	update	<pre>INSERT INTO cache (cid, serialized, created, expire, data) VALUES ('node_types:en', '1', '1415747825</pre>
Kill	243	root	localhost:51092	mysql	Query	0	init	SHOW PROCESSLIST

95% of the time database transactions are to blame

What to do?

https://www.drupal.org/project/apdqc Is the answer. Module if used correctly will fix the deadlock and metadata locking issues with core's db cache.

I want to monitor this

You can run this next query and you can then kill that mysql process id if you desire to. This only works in real time; every time you get a locked database, run the query and manually kill that process id.

ID stuck thread that holds the lock

```
SELECT trx.trx id AS trx id, trx.trx mysql thread id AS thread id, trx.
trx query AS query, trx.trx tables locked AS tables locked, trx.
trx rows locked AS rows locked, trx.trx state AS trx state, p.DB AS db, p.
STATE AS process state, trx.trx operation state AS trx op state, p.COMMAND
AS command, p.TIME AS query age, DATEDIFF( trx.trx started, NOW()) AS
trx age, DATEDIFF( trx.trx wait started, NOW()) AS wait age, wr.
requested lock id AS requested lock id, wb.requested lock id AS
blocking lock id, trx.trx isolation level AS iso level
FROM information schema.innodb trx AS trx
LEFT JOIN information schema.processlist p ON trx.trx mysql thread id = p.
ΙD
LEFT JOIN information schema.innodb lock waits w ON trx.
trx mysql thread id = p.ID
LEFT JOIN information schema.innodb lock waits wb ON trx.trx id = wb.
blocking trx id
LEFT JOIN information schema.innodb lock waits wr ON trx.trx id = wr.
requesting trx id
WHERE trx tables locked >0 OR trx rows locked >0
GROUP BY trx.trx mysql thread id
```

trx_id	thread_id	query	tables_locked	rows_locked	trx_state	db	process_state	trx_op_state	command	query_age	trx_age	wait_age	requested_lock_id	blocking_lock_id	iso_level
3805696	236	NULL	0	1	RUNNING	d7		NULL	Sleep	7	0	NULL	NULL	3805701:60646:24:3	REPEATABLE READ
3805697		INSERT INTO cache (cid, serialized, created, expir	1	1	LOCK WAIT	d7	update	inserting	Query	7	0	0	3805697:60646:24:3	3805701:60646:24:3	READ UNCOMMITTED
3805698		INSERT INTO cache (cid, serialized, created, expir	1	1	LOCK WAIT	d7	update	inserting	Query	7	0	0	3805698:60646:24:3	3805701:60646:24:3	READ UNCOMMITTED
3805699	239	INSERT	1	1	LOCK	d7	update	inserting	Query	7	0	0	3805699:60646:24:3	3805701:60646:24:3	READ

Query

Query

inserting

inserting

UNCOMMITTED

UNCOMMITTED

UNCOMMITTED

NULL READ

0 3805700:60646:24:3 3805701:60646:24:3 READ

0 3805701:60646:24:3

https://goo.gl/L4P3MA

WAIT

1 LOCK

1 LOCK

WAIT

WAIT

d7 update

d7 update

INTO

cache (cid, serialized, created, expir...

240 INSERT

241 INSERT

INTO

cache (cid,

INTO

cache (cid, serialized, created, expir...

3805700

3805701

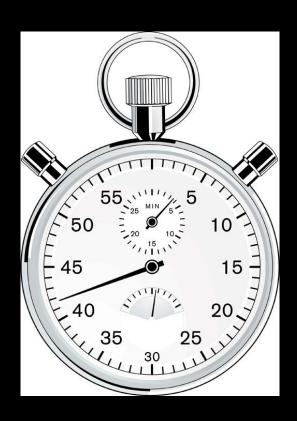
trx_id	thread_id	query	tables_locked	rows_locked	trx_state	db	process_state	trx_op_state	command	query_age	trx_age	wait_ag	requested_lock_id	blocking_lock_id	ko_level
3805696	236	NULL	0	1	RUNNING	d7		NULL	Sleep	7	0	NUL	NULL	3805701:60646:24:3	READ
3805697		INSERT INTO cache (cid, serialized, created, expir	1	1	LOCK WAIT	d7	update	inserting	Query	7	0	0	3805697:60646:24:3	3805701:60646:24:3	READ UNCOMMITTED
3805698		INSERT INTO cache (cid, serialized, created, expir	1	1	LOCK WAIT	d7	update	inserting	Query	7	0	0	3805698:60646:24:3	3805701:60646:24:3	READ UNCOMMITTED
3805699		INSERT INTO cache (cid, serialized, created, expir	1	1	LOCK WAIT	d7	update	inserting	Query	7	0	0	3805699:60646:24:3	3805701:60646:24:3	READ UNCOMMITTED
3805700		INSERT INTO cache (cid, serialized, created, expir	1	1	LOCK WAIT	d7	update	inserting	Query	7	0	0	3805700:60646:24:3	3805701:60646:24:3	READ UNCOMMITTED
3805701		INSERT INTO cache (cid,	1	1	LOCK WAIT	d7	update	inserting	Query	7	0	0	3805701:60646:24:3	NULL	READ UNCOMMITTED

Slow Query Analysis

Percona's toolkit has a nice perl script called pt-query-digest. Recommend using that.



Basic query performance measurements with the devel module



https://drupal.org/project/devel

This module can tell you what queries are slow and how many ms is spent in the database and in php.

admin/config/development/devel

QUERY LOG

✓ Display query log

Display a log of the database queries needed to generate the current page, and the execution time for each. Also, queries which are repeated during a single page view are summed in the # column, and printed in red since they are candidates for caching.

Sort query log

- O by source
- by duration

The query table can be sorted in the order that the queries were executed or by descending duration.

Slow query highlighting

5

Enter an integer in milliseconds. Any query which takes longer than this many milliseconds will be highlighted in the query log. This indicates a possibly inefficient query, or a candidate for caching.

admin/config/development/devel

V	Display page timer
	Dispile, page execution time in the query log box.
ν	Display memory usage
	Display how much memory is used to generate the current page. This will show memory usage when devel_init() is called and when devel_exit() is called.
	Display redirection page
	When a module executes drupal_goto(), the query log and other developer information is lost. Enabling this setting presents an intermediate page to developers so that the log can be examined before continuing to the destination page.
	Display \$page array
	Display \$page array from hook_page_alter() in the messages area of each page.
	Display machine names of permissions and modules
	Display the language-independent machine names of the permissions in mouse-over hints on the Permissions page and the module base file names on the Permissions and Modules pages.

ms	#	where	ops	query	target
0.66	3	menu_get_item	P) E	SELECT * FROM menu_router WHERE path IN (:ancestors_0, :ancestors_1, :ancestors_2, :ancestors_3, :ancestors_4, :ancestors_5, :ancestors_6, :ancestors_7, :ancestors_9, :ancestors_10, :ancestors_11, :ancestors_12, :ancestors_13, :ancestors_14, :ancestors_15, :ancestors_16, :ancestors_17, :ancestors_18, :ancestors_19, :ancestors_20, :ancestors_21, :ancestors_22, :ancestors_23, :ancestors_21, :ancestors_22, :ancestors_23, :ancestors_23, :ancestors_24, :ancestors_25, :ancestors_26, :ancestors_27, :ancestors_27, :ancestors_28, :ancestors_28, :ancestors_29, :ancestors_29, :ancestors_21, :anc	default
0.6	3	menu_get_item	(A)E	SELECT * FROM menu router WHERE path IN ('admin/structure/block/manage/system/navigation', 'admin/structure/block/manage/system/%', 'admin/structure/block/manage/\$/navigation', 'admin/structure/block/manage/\$/\$', 'admin/structure/block/%/system/navigation', 'admin/structure/block/%/system/%', 'admin/structure/block/manage/system', 'admin/structure/block/manage/system', 'admin/structure/block/manage/system', 'admin/structure/block/%', 'admin/structure/block/manage/system', 'admin/structure/block/\$/\$', 'admin/structure/block/manage', 'admin/structure/block/system', 'admin/structure/system', 'adm	default %',
0.57	3	menu_get_item	PAE)		default

index

PRIMARY

Using where

Executed 34 queries in 6.86 ms. Queries exceeding 5 ms are highlighted. Page execution time was 123.86 ms. Memory used at: devel_boot()=3.08 MB, devel_shutdown()=15.14 MB(PHP peak=15.25 MB.)

menu_router

SIMPLE

Executed 34 queries in 6.86 ms. Operies exceeding 5 ms are highlighted. Page execution time was 123.86 ms. Memory used

ms	#	where	ops	query
0.66	3	menu_get_item	P) E	SELECT * FROM menu_router WHERE path IN :ancestors_6, :ancestors_7, :ancestors_8 :ancestors_15, :ancestors_16, :ancestors_ORDER BY fit DESC LIMIT 0, 1
0.6	3	menu_get_item	I A	SELECT * FROM menu_router WHERE path IN 'admin/structure/block/manage/%/navigati 'admin/structure/block/%/system/%', 'admin/structure/block/%/system', 'admin/block/%/system', 'admin/structure/block 'admin/structure/block', 'admin/structure
			\sim	

menu_get_item

0.57

id	select_type	table
1	SIMPLE	menu_router

```
evel_boot()=3.08 MB, devel_shutdown()=15.14 MB(PHP peak=15.25 MB.
```

target

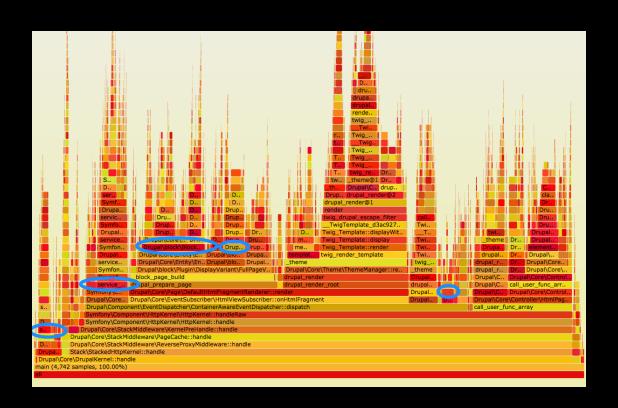
```
cestors_0, :ancestors_1, :ancestors_2, :ancestors_3, :ancestors_4, :ancestors_5, default ncestors_9, :ancestors_10, :ancestors_11, :ancestors_12, :ancestors_13, :ancestors_14, :ancestors_18, :ancestors_19, :ancestors_20, :ancestors_21, :ancestors_22, :ancestors_23)
```

```
min/structure/block/manage/system/navigation', 'admin/structure/block/manage/system/%', default
   'admin/structure/block/manage/%/%', 'admin/structure/block/%/system/navigation',
/block/manage/%/%', 'admin/structure/block/manage/system', 'admin/structure/block/manage/%',
ucture/block/%/%', 'admin/structure/%/manage/system', 'admin/structure/%/%', 'admin/%
age', 'admin/structure/block/%', 'admin/structure/%/manage', 'admin/%/block/manage',
, 'admin/%/block', 'admin/structure', 'admin/%', 'admin') ORDER BY fit DESC LIMIT 0, 1
```

default

type	possible_keys	key	key_len	ref	rows	Extra
index	PRIMARY	fit	4		15	Using where

Xdebug/Xhprof



Generating a cachegrind file

This allows you to pinpoint slow points. Once you know the bottlenecks, you can try to figure out what options are available to fix them.

xhprof is a good alternative to xdebug.

Another alternative to this is to use New Relic. It provides a lot of good information and it comes with a 14 day free trial.

Questions?

Twitter: omcarper

D7 performance wiki: <a href="https://groups.drupal.groups

Link to this presentation https://goo.gl/L4P3MA