

Core - Bug # 66503

| | | | |
|--|---|---------------------|-------------------|
| Status: | Resolved | Priority: | Must have |
| Author: | Mateusz Wojtuła | Category: | Frontend |
| Created: | 2015-04-20 | Assigned To: | Markus Klein |
| Updated: | 2015-04-27 | Due date: | |
| TYPO3 Version: | 7 | | |
| PHP Version: | 5.5 | | |
| Complexity: | medium | | |
| Is Regression: | No | | |
| Sprint Focus: | Stabilization Sprint | | |
| Subject: | Core: Error handler (FE): PHP Warning: sem_get(): failed for key 0xbaa3533: No space left on device | | |
| Description | | | |
| <p>Hi,</p> <p>I got the following error (this information is from tab Log):</p> <p>Core: Error handler (FE): PHP Warning: sem_get(): failed for key 0xbaa3533: No space left on device in /var/www/project/typo3_src/typo3/sysext/core/Classes/Locking/SemaphoreLockStrategy.php line 100</p> <p>The problem is on page with single news. I also use fluidpages, fluidcontents, tq_seo. I added information about environment as print screen.</p> | | | |
| Related issues: | | | |
| related to Core - Feature # 47712: Refactor Locking | | Resolved | 2013-02-08 |

Associated revisions

Revision a1ed7cef - 2015-04-27 19:10 - Markus Klein

[BUGFIX] Multiple fixes for Locking API and TSFE locking

- Retrieve correct LockingStrategy for requested capabilities
- Prefix lock filenames to make them better visible in the folder
- Make all LockStrategies destroyable
- Semaphore locking now uses ftok() to generate a unique id
- Make the Mbox lock independent of the target file
- Introduce an access lock for each of the TSFE cache locks

We decrease the priority of Semaphore locking since this can be pretty dangerous for the average user. If something goes really wrong in the webserver (which is out of our control), we might leave behind stale semaphores, which might cause a permanent deadlock for an instance, which can only be resolved by a server admin.

We might raise the priority again at a later point in time, when we can provide better means of cleanup.

The new access locks protects the access to the cache locks in TSFE now, which allows us to safely remove those cache locks after using them. This way we don't spam the system with loads of locks.

Releases: master
Resolves: #66503
Change-Id: Ia19e6e7d47d7941e01785f5a6b67746a6c0fa368
Reviewed-on: <http://review.typo3.org/38840>
Reviewed-by: Andreas Allacher <andreas.allacher@gmx.at>
Tested-by: Andreas Allacher <andreas.allacher@gmx.at>
Reviewed-by: Christian Kuhn <lolli@schwarzbu.ch>
Tested-by: Christian Kuhn <lolli@schwarzbu.ch>
Reviewed-by: Markus Klein <klein.t3@reelworx.at>
Tested-by: Markus Klein <klein.t3@reelworx.at>

History

#1 - 2015-04-20 20:15 - Markus Klein

Please check your system status with

```
ipcs -s
```

```
sysctl -a |grep kernel.sem
```

#2 - 2015-04-20 20:16 - Markus Klein

The default on Ubuntu 14.04 is:

```
kernel.sem = 32000 1024000000 500 32000
```

#3 - 2015-04-20 20:24 - Mateusz Wojtuła

I get this:

```
root@matw:~# ipcs -s
```

```
----- Semaphore Arrays -----
```

| key | semid | owner | perms | nsems |
|------------|--------|----------|-------|-------|
| 0x0baa3533 | 0 | www-data | 666 | 3 |
| 0x7c9e0990 | 32769 | www-data | 666 | 3 |
| 0xa8e44899 | 65538 | www-data | 666 | 3 |
| 0x1e536c5e | 98307 | www-data | 666 | 3 |
| 0xb4375310 | 131076 | www-data | 666 | 3 |
| 0xea7be8dc | 163845 | www-data | 666 | 3 |
| 0x456c2c52 | 196614 | www-data | 666 | 3 |
| 0x02867477 | 229383 | www-data | 666 | 3 |
| 0xf0e544b5 | 262152 | www-data | 666 | 3 |
| 0x339a6321 | 294921 | www-data | 666 | 3 |
| 0x3ca1cf4d | 327690 | www-data | 666 | 3 |
| 0x4d9d2850 | 360459 | www-data | 666 | 3 |

| | | | |
|--------------------|----------|-----|---|
| 0xb532d7b8 393228 | www-data | 666 | 3 |
| 0x5d85f034 425997 | www-data | 666 | 3 |
| 0xd2358e97 458766 | www-data | 666 | 3 |
| 0xaf36479e 491535 | www-data | 666 | 3 |
| 0xad937213 524304 | www-data | 666 | 3 |
| 0x2347b1c0 557073 | www-data | 666 | 3 |
| 0x14362d4d 589842 | www-data | 666 | 3 |
| 0xb1063a0a 622611 | www-data | 666 | 3 |
| 0xa0784163 655380 | www-data | 666 | 3 |
| 0xa1cc990c 688149 | www-data | 666 | 3 |
| 0x87e611ee 720918 | www-data | 666 | 3 |
| 0xc2b3fab1 753687 | www-data | 666 | 3 |
| 0x33e81829 786456 | www-data | 666 | 3 |
| 0xd8c99d51 819225 | www-data | 666 | 3 |
| 0x8783b42a 851994 | www-data | 666 | 3 |
| 0x60e5873a 884763 | www-data | 666 | 3 |
| 0xc314ead5 917532 | www-data | 666 | 3 |
| 0xa9442db3 950301 | www-data | 666 | 3 |
| 0xab52b406 983070 | www-data | 666 | 3 |
| 0x26551e77 1015839 | www-data | 666 | 3 |
| 0xd3fc66ce 1048608 | www-data | 666 | 3 |
| 0xf16f6100 1081377 | www-data | 666 | 3 |
| 0x25b03396 1114146 | www-data | 666 | 3 |
| 0x2ba70226 1146915 | www-data | 666 | 3 |
| 0xf10f5538 1179684 | www-data | 666 | 3 |
| 0x365b4f31 1212453 | www-data | 666 | 3 |
| 0x6a8d5d1a 1245222 | www-data | 666 | 3 |
| 0x6e2a442c 1277991 | www-data | 666 | 3 |
| 0x8264f29c 1310760 | www-data | 666 | 3 |
| 0x0973ce0a 1343529 | www-data | 666 | 3 |
| 0x043b027a 1376298 | www-data | 666 | 3 |
| 0xa250f4df 1409067 | www-data | 666 | 3 |
| 0x7c984307 1441836 | www-data | 666 | 3 |
| 0xebf7e556 1474605 | www-data | 666 | 3 |
| 0xf02169f2 1507374 | www-data | 666 | 3 |
| 0xc27603cd 1540143 | www-data | 666 | 3 |
| 0xd3c3c612 1572912 | www-data | 666 | 3 |
| 0x9cc25ec5 1605681 | www-data | 666 | 3 |
| 0x04629384 1638450 | www-data | 666 | 3 |
| 0xd5e85eaf 1671219 | www-data | 666 | 3 |
| 0x22687898 1703988 | www-data | 666 | 3 |
| 0xb40f30b2 1736757 | www-data | 666 | 3 |
| 0xe31f7de7 1769526 | www-data | 666 | 3 |
| 0x07aecb8a 1802295 | www-data | 666 | 3 |
| 0xcf62bda7 1835064 | www-data | 666 | 3 |
| 0x2c784309 1867833 | www-data | 666 | 3 |
| 0x0e8d31f2 1900602 | www-data | 666 | 3 |
| 0x4fd013e4 1933371 | www-data | 666 | 3 |
| 0xb44b91b2 1966140 | www-data | 666 | 3 |
| 0x9b85d8b5 1998909 | www-data | 666 | 3 |
| 0xe8994ff1 2031678 | www-data | 666 | 3 |
| 0xdeb8161d 2064447 | www-data | 666 | 3 |
| 0xa3c4fb43 2097216 | www-data | 666 | 3 |

| | | | | |
|------------|---------|----------|-----|---|
| 0x2ba5a06e | 2129985 | www-data | 666 | 3 |
| 0xd33d736a | 2162754 | www-data | 666 | 3 |
| 0x113edfef | 2195523 | www-data | 666 | 3 |
| 0x36a1dc1a | 2228292 | www-data | 666 | 3 |
| 0xf12befbb | 2261061 | www-data | 666 | 3 |
| 0x69911364 | 2293830 | www-data | 666 | 3 |
| 0xa9140a5f | 2326599 | www-data | 666 | 3 |
| 0x03c74337 | 2359368 | www-data | 666 | 3 |
| 0x6a014e24 | 2392137 | www-data | 666 | 3 |
| 0xbc24ca33 | 2424906 | www-data | 666 | 3 |
| 0xf114402a | 2457675 | www-data | 666 | 3 |
| 0x2a36aca6 | 2490444 | www-data | 666 | 3 |
| 0x2e918a24 | 2523213 | www-data | 666 | 3 |
| 0xe4a3c031 | 2555982 | www-data | 666 | 3 |
| 0x0c841d24 | 2588751 | www-data | 666 | 3 |
| 0xb380209d | 2621520 | www-data | 666 | 3 |
| 0x6149871e | 2654289 | www-data | 666 | 3 |

```

root@matw:~# sysctl -a |grep kernel\sem
kernel.sem = 250 32000 32 128
kernel.sem_next_id = -1

```

So this is my system error? Not TYPO3?

#4 - 2015-04-20 21:19 - Markus Klein

What server do you use?
Is this a shared hosting or so?

You have a 250 limit for the number of semaphores.

TYPO3 currently does not remove them, because we have no "controlled" environment - in the sense that a request does not know about other requests - and we don't want to have race conditions because of removing a semaphore too early.

Thanks for you report, btw, because this is very valuable feedback as the Locking API was changed not long ago and we really need field experience.

Currently the semaphore locking is the preferred method, if available. But if it turns out that it causes too much trouble, we might change that again.

#5 - 2015-04-20 21:37 - Markus Klein

- Category set to Frontend
- Status changed from New to Accepted
- Assigned To set to Markus Klein
- Priority changed from Should have to Must have
- Target version set to 7.2 (Frontend)
- Complexity set to medium
- Sprint Focus set to Stabilization Sprint

After checking the code, I see that we potentially use quite a lot of semaphores.
I'll try to come up with a patch to limit the number.

#6 - 2015-04-20 21:50 - Mateusz Wojtula

Thanks, for this information. So what can I do when this error appears again?
This site is on [DigitalOcean](#) with the smallest virtual server (512MB ram, 1 core, 20 GB SSD).

If you want to check it you can use this link to register and get 10\$ for free <https://www.digitalocean.com/?refcode=96665686914b>

#7 - 2015-04-20 22:31 - Markus Klein

More infos about the numbers.

```
kernel.sem = 32000    1024000000    500    32000
kernel.sem = 250      32000          32     128
```

The order is: SEMMSL, SEMMNS, SEMOPM, and SEMMNI

SEMMSL: maximum number of semaphores per semaphore set
SEMMNS: total number of semaphores (not semaphore sets) for the entire Linux system
SEMOPM: maximum number of semaphore operations that can be performed per semop(2) system call
SEMMNI: maximum number of semaphore sets for the entire Linux system

Description of the settings eg. <http://www.puschitz.com/TuningLinuxForOracle.shtml#SettingSemaphores>

It can clearly be seen that the second set of numbers limits the total number of semaphores on the whole system to 128(!)

#8 - 2015-04-20 23:15 - Markus Klein

| So what can I do when this error appears again?

Search for how you can increase the maximum number of semaphores or how you can release existing ones on the web.

#9 - 2015-04-20 23:37 - Markus Klein

After digging around in the code I realize:

We use a lot of information to generate a unique key for the lock. So we actually have a key per variation of a page. (md5 of serialize of id, type, gr_list, MP, cHash, startPage)
\\TYPO3\CMS\Frontend\Controller\TypoScriptFrontendController::createHashBase(TRUE)

This means we spam the system with a semaphore per variation, but OS might limit us to 128 semaphores!
This approach is brutally wrong, even if we don't use semaphores but file locks, which need a file per variation then.

We need to define a better strategy.

Idea: Use a single lock to access some shared resource, which keeps track of the process currently rendering a page.

#10 - 2015-04-21 00:27 - Markus Klein

Current procedure:

- check cache, if empty proceed with
- generate hash
- get lock for this hash
- generate page
- write page to cache
- release lock

This creates a lock per hash, which might be manifold.

I propose the following procedure:

- check cache, if empty proceed with
- generate hash
- get lock for cache access (1 unique lock per instance)
- write to cache that we're working on the content (set lock for this hash)
- release the lock
- generate page
- get lock again
- remove lock for the hash
- write page to cache
- release lock

This is some sort of simple reader/writer pattern, where we allow inconsistent reads, while generating the page.

Regarding the Semaphore locking in general:

We need to change the semaphore keys. Important to say that the keys must be **unique per instance**, otherwise two T3 instances on the same server might block each other (even if only for a very short time).

The best way to generate a key (that also does not collide with OS keys) is to use `ftok(filename, projId)`, which needs a filename.

We could use a fixed filename, but that would limit the possible range of `projId` to 256 (8bit), which is a no-go, since we can't map arbitrary \$subject to 256 `projIds` in a sane way.

Therefore we propose to create a file in `typo3temp/locks/sem_<md5 of $subject>` used for `ftok`.

This way one also has a reference how many semaphores have been created in the system by the instance, without looking into command line tools.

#11 - 2015-04-21 02:33 - Gerrit Code Review

- *Status changed from Accepted to Under Review*

Patch set 1 for branch **master** of project **Packages/TYPO3.CMS** has been pushed to the review server.

It is available at <http://review.typo3.org/38840>

#12 - 2015-04-21 02:33 - Gerrit Code Review

Patch set 2 for branch **master** of project **Packages/TYPO3.CMS** has been pushed to the review server.

It is available at <http://review.typo3.org/38840>

#13 - 2015-04-21 02:38 - Gerrit Code Review

Patch set 3 for branch **master** of project **Packages/TYPO3.CMS** has been pushed to the review server.
It is available at <http://review.typo3.org/38840>

#14 - 2015-04-21 03:20 - Gerrit Code Review

Patch set 4 for branch **master** of project **Packages/TYPO3.CMS** has been pushed to the review server.
It is available at <http://review.typo3.org/38840>

#15 - 2015-04-21 04:06 - Gerrit Code Review

Patch set 5 for branch **master** of project **Packages/TYPO3.CMS** has been pushed to the review server.
It is available at <http://review.typo3.org/38840>

#16 - 2015-04-21 04:30 - Gerrit Code Review

Patch set 6 for branch **master** of project **Packages/TYPO3.CMS** has been pushed to the review server.
It is available at <http://review.typo3.org/38840>

#17 - 2015-04-21 04:35 - Gerrit Code Review

Patch set 7 for branch **master** of project **Packages/TYPO3.CMS** has been pushed to the review server.
It is available at <http://review.typo3.org/38840>

#18 - 2015-04-21 04:53 - Markus Klein

Testing instructions:

Insert a sleep(10) in the tsfe::generate_pre...() function, right after the release of the lock. This allows you to see the "Page is generated" message" if you visit the page with two browsers at the same time. Clear the cache first and logout from BE.
At the same time you can watch the lock files come and go in typo3temp/locks.

#19 - 2015-04-22 19:51 - Gerrit Code Review

Patch set 8 for branch **master** of project **Packages/TYPO3.CMS** has been pushed to the review server.
It is available at <http://review.typo3.org/38840>

#20 - 2015-04-24 14:33 - Gerrit Code Review

Patch set 9 for branch **master** of project **Packages/TYPO3.CMS** has been pushed to the review server.
It is available at <http://review.typo3.org/38840>

#21 - 2015-04-24 14:38 - Gerrit Code Review

Patch set 10 for branch **master** of project **Packages/TYPO3.CMS** has been pushed to the review server.

It is available at <http://review.typo3.org/38840>

#22 - 2015-04-24 18:49 - Andreas Allacher

Test-script to produce some locking states, this would result in deadlock with patchset 8 because of missing no block option of patchset 9

```
#!/bin/sh
./typo3cms cache:flushgroups --groups=pages
time wget --output-document=index_temp.1.html --content-on-error "http://localhost/" &
time wget --output-document=index_temp.2.html --content-on-error "http://localhost/" &
time wget --output-document=index_temp.3.html --content-on-error "http://localhost/" &
time wget --output-document=index_temp.4.html --content-on-error "http://localhost/" &
time wget --output-document=index_temp.5.html --content-on-error "http://localhost/" &
time wget --output-document=index_temp.6.html --content-on-error "http://localhost/" &
time wget --output-document=index_temp.7.html --content-on-error "http://localhost/" &
time wget --output-document=index_temp.8.html --content-on-error "http://localhost/" &
time wget --output-document=index_temp.9.html --content-on-error "http://localhost/" &
time wget --output-document=index_temp.10.html --content-on-error "http://localhost/" &
time wget --output-document=index_temp.11.html --content-on-error "http://localhost/" &
time wget --output-document=index_temp.12.html --content-on-error "http://localhost/" &
```

Regarding usleep in locking part it according to times it seems to be better to increase it (at least for me). But I think it is also depends on how much memory/resources are available

With 5us I am about 1second slower per request than with e.g. 5000 but I guess that really depends on the system usage.

But I think 5us might really be to little 5000 would at least be 5 milliseconds.

Event better (but not much normally than 5000) seem to be 50000 or 100000 but might also be system related.

Maybe others can do some tests too?

#23 - 2015-04-24 18:50 - Andreas Allacher

- File *lock_test.sh* added

#24 - 2015-04-25 20:37 - Gerrit Code Review

Patch set 11 for branch **master** of project **Packages/TYPO3.CMS** has been pushed to the review server.

It is available at <http://review.typo3.org/38840>

#25 - 2015-04-27 18:19 - Gerrit Code Review

Patch set 12 for branch **master** of project **Packages/TYPO3.CMS** has been pushed to the review server.

It is available at <http://review.typo3.org/38840>

#26 - 2015-04-27 19:30 - Markus Klein

- Status changed from *Under Review* to *Resolved*

- % Done changed from 0 to 100

Files

| | | | |
|-----------------------------|---------|------------|------------------|
| application_information.png | 20.6 kB | 2015-04-20 | Mateusz Wojtuła |
| lock_test.sh | 1.1 kB | 2015-04-24 | Andreas Allacher |