Core - Bug # 63615

Bug # 63692 (Accepted): Memory consumption while bulk inserting

Status:	Resolved	**Priority:**	Must have
Author:	Stephan Großberndt	**Category:**	Performance
Created:	2014-12-05	**Assigned To:**	
Updated:	2014-12-10	**Due date:**	

TYPO3 Version:	6.2
PHP Version:	
Complexity:	easy
Is Regression:	No
Sprint Focus:	

Subject:	Huge memory consumption in DataHandler->processClearCacheQueue while bulk inserting

Description

I found a memory consumption problem in   TYPO3\CMS\Core\DataHandling\DataHandler::processClearCacheQueue()
 due to this dubious construct, which is called for each and every record added:

```
foreach (static::$recordsToClearCacheFor as $table => $uids) {
  foreach (array_unique($uids) as $uid) {

    ...

    foreach ($pageIdsThatNeedCacheFlush as $pageId) {
      if ($pageId >= 0) {
        $tagsToClear[] = 'pageId_' . $pageId;
      }
    }
    // Queue delete cache for current table and record
    $tagsToClear[] = $table;
    $tagsToClear[] = $table . '_' . $uid;
    $tagsToClear = array_unique($tagsToClear);
  }
}
```

I am adding 6500 records, so **FOR EACH AND EVERY** record 'pageId_' . $pageId, $table and $table . '_' . $uid are added to
$tagsToClear and afterwards 'pageId_' . $pageId and $table are removed again for 6499 times. There is no reason to do it like this.
array_unique() is meant here to reduce memory consumption in $tagsToClear but for big arrays with array_unique() this is terrible -
xhprof showed 2,020,470,496 bytes (~2 gigabyte !!!) of used memory for array_unique() in processClearCacheQueue()

It is possible to set   clearCache_disable=1
 to disable clearing of the cache but there is also a simple code solution to solve this since entries in this array are meant to be unique
anyway:

```
foreach (static::$recordsToClearCacheFor as $table => $uids) {
  foreach (array_unique($uids) as $uid) {

    ...

    foreach ($pageIdsThatNeedCacheFlush as $pageId) {
      if ($pageId >= 0) {
        $tagsToClear['pageId_' . $pageId] = 'pageId_' . $pageId;
      }
```

```
        }
        // Queue delete cache for current table and record
        $tagsToClear[$table] = $table;
        $tagsToClear[$table . '_' . $uid] = $table . '_' . $uid;
      }
    }
```

array_unique memory consumption drops from

```
    IMemUse%: 13064.1%
    Incl. MemUse (bytes): 2,020,470,496 bytes (~2 gigabyte !!!)
```

to

```
    IMemUse%: 4.1%
    Incl. MemUse (bytes): 629,464 bytes
```

There may be optimizations possible like

```
    foreach (static::$recordsToClearCacheFor as $table => $uids) {
      foreach (array_unique($uids) as $uid) {
        ...

        foreach ($pageIdsThatNeedCacheFlush as $pageId) {
          $pageKey = 'pageId_' . $pageId;
          if ($pageId >= 0 && !isset($tagsToClear[$pageKey])) {
            $tagsToClear[$pageKey] = $pageKey;
          }
        }
        // Queue delete cache for current table and record
        if (!isset($tagsToClear[$table])) {
          $tagsToClear[$table] = $table;
        }
        $tagsToClear[$table . '_' . $uid] = $table . '_' . $uid;
      }
    }
```

but even without those its **SOOOO** much better...

Regards,
Stephan

Associated revisions

Revision 0f2f447e - 2014-12-10 01:10 - Stephan Großberndt

[BUGFIX] Memory consumption in DataHandler->processClearCacheQueue

Reduce memory consumption in DataHandler->processClearCacheQueue() for
bulk inserts by filling $tagsToClear as an associative array instead of
a normal array with array_unique()-call after each added record.

Resolves: #63615
Releases: master, 6.2
Change-Id: I1f557f435088cab3aa928ec3f44b0ded9265d7a7
Reviewed-on: http://review.typo3.org/35085
Reviewed-by: Lukas Krieger <lukas.krieger@me.com>
Tested-by: Lukas Krieger <lukas.krieger@me.com>
Reviewed-by: Stefan Neufeind <typo3.neufeind@speedpartner.de>
Reviewed-by: Alexander Opitz <opitz.alexander@googlemail.com>
Tested-by: Alexander Opitz <opitz.alexander@googlemail.com>
Reviewed-by: Markus Klein <klein.t3@reelworx.at>
Tested-by: Markus Klein <klein.t3@reelworx.at>

Revision 31518c96 - 2014-12-10 01:11 - Stephan Großberndt

[BUGFIX] Memory consumption in DataHandler->processClearCacheQueue

Reduce memory consumption in DataHandler->processClearCacheQueue() for
bulk inserts by filling $tagsToClear as an associative array instead of
a normal array with array_unique()-call after each added record.

Resolves: #63615
Releases: master, 6.2
Change-Id: I1f557f435088cab3aa928ec3f44b0ded9265d7a7
Reviewed-on: http://review.typo3.org/35211
Reviewed-by: Markus Klein <klein.t3@reelworx.at>
Tested-by: Markus Klein <klein.t3@reelworx.at>

Revision 6da66858 - 2014-12-10 14:34 - Stephan Großberndt

[BUGFIX] Memory consumption in DataHandler->processClearCacheQueue

Reduce memory consumption in DataHandler->processClearCacheQueue() for
bulk inserts by filling $tagsToClear as an associative array instead of
a normal array with array_unique()-call after each added record.

Resolves: #63615
Releases: master, 6.2
Change-Id: I74bef31fcf1f08665e51cee7efe847e32f501ef2
Reviewed-on: http://review.typo3.org/35232
Reviewed-by: Oliver Hader <oliver.hader@typo3.org>
Tested-by: Oliver Hader <oliver.hader@typo3.org>
Reviewed-by: Stephan Großberndt <stephan@grossberndt.de>
Tested-by: Stephan Großberndt <stephan@grossberndt.de>
Reviewed-by: Wouter Wolters <typo3@wouterwolters.nl>
Tested-by: Wouter Wolters <typo3@wouterwolters.nl>

History

#1 - 2014-12-05 19:28 - Stephan Großberndt

This may be related: http://forum.typo3.org/index.php?t=msg&th=206887

#2 - 2014-12-05 20:59 - Gerrit Code Review

- Status changed from New 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/35085

#3 - 2014-12-07 11:57 - 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/35085

#4 - 2014-12-07 11:59 - 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/35085

#5 - 2014-12-07 17:51 - 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/35085

#6 - 2014-12-09 07:03 - Philipp Gampe

- Parent task set to #63692

#7 - 2014-12-09 12:18 - 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/35085

#8 - 2014-12-09 12:30 - Gerrit Code Review

false commit removed

#9 - 2014-12-09 12:53 - Gerrit Code Review

false commit removed

#10 - 2014-12-09 15:12 - Lukas Krieger

the gerrit issue-ID is wrong

Gerrit Code Review wrote:

> *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/35185*

#11 - 2014-12-09 15:28 - Stephan Großberndt

This was an accidental commit to this issue. It should have been #63666 in the first place. It was fixed in the commit message and is shown in the correct issue now.

Stefan Neufeind corrected it here (strike-out and "wrong issue-ID"), but Benjamin Mack did another commit based on the old version. You can safely ignore it here. I requested another delete / strikeout in the typo3.teams.core mailinglist.

#12 - 2014-12-09 17:13 - Alexander Opitz

removed false commits here.

Last state is:

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/35085

#13 - 2014-12-10 01:11 - Gerrit Code Review

Patch set 1 for branch **TYPO3_6-2** of project **Packages/TYPO3.CMS** has been pushed to the review server.
It is available at http://review.typo3.org/35211

#14 - 2014-12-10 01:30 - Stephan Großberndt
- Status changed from Under Review to Resolved
- % Done changed from 0 to 100

Applied in changeset commit:0f2f447e2ec834b69436bd5cca766353bf6d7024.

#15 - 2014-12-10 11:26 - Gerrit Code Review
- Status changed from Resolved to Under Review

Patch set 1 for branch **TYPO3_6-2** of project **Packages/TYPO3.CMS** has been pushed to the review server.
It is available at http://review.typo3.org/35232

#16 - 2014-12-10 15:00 - Stephan Großberndt
- Status changed from Under Review to Resolved

Applied in changeset commit:6da6685887adf0386345d8c13ea092e70c3ace1b.