

TYPO3 Core - Bug #53862

Epic # 55070 (Closed): Workpackages

Epic # 55065 (Closed): WP: Overall System Performance (Backend and Frontend)

Bug # 52949 (Closed): Speed decrease since 4.5

isValidUrl idna converts whole URI instead of domain only

2013-11-22 12:05 - Michiel Roos

Status: Closed	Start date: 2013-11-22
Priority: Should have	Due date:
Assignee:	% Done: 100%
Category: Performance	Estimated time: 0.00 hour
Target version: next-patchlevel	Complexity: easy
TYPO3 Version: 6.2	Is Regression: No
PHP Version: 5.3	Sprint Focus:
Tags:	
Description	
<p>The expensive idna_convert is called from isValidUrl. Instead of feeding it just the domain part, it converts the whole URI.</p> <p>When feeding it just the domain part, a great speed gain can be had.</p> <p>This patch breaks apart the URI using parse_url(), converts the domain part, then re-assembles the uri using HttpUtility::buildUrl().</p> <p>Please find profiling runs attached for:</p> <ul style="list-style-type: none">• /typo3/ (login screen)• /typo3/backend.php• /typo3/mod.php?M=web_list	
Related issues:	
Related to TYPO3 Core - Bug #55475: Error when saving Domain system record	Closed 2014-01-30
Related to TYPO3 Core - Bug #55713: GeneralUtility' not found	Closed 2014-02-06

Associated revisions

Revision f8fdcea7 - 2013-12-21 15:50 - Michiel Roos

[BUGFIX] isValidUrl() idna converts whole URI

GeneralUtility::isValidUrl() idna converts whole URI instead of domain only.

The expensive idna_convert() is called from isValidUrl(). Instead of feeding it just the domain part, the whole URI is converted.

When supplying just the domain part, a great speed gain can be seen.

On seriously malformed URLs, parse_url() may return FALSE and emits an E_WARNING. So we check for that first.

PHP no longer supports the flags FILTER_FLAG_HOST_REQUIRED and FILTER_FLAG_SCHEME_REQUIRED. A scheme is now required by default. [1]
Return FALSE if the URL does not start with a scheme.

A public GeneralUtility::idnaEncode() method uses a static idna_convert instance and fetches converted strings from a string cache array to avoid multiple checks on the same domain.

All manual idna_convert instances are replaced with GeneralUtility::idnaEncode() calls.

Special characters are not allowed in the URL except in the domain part [2]. So the test with special characters in the path was removed from the GeneralUtilityTest class.

[1] <http://www.php.net/manual/en/filter.filters.flags.php#107382>

[2] <http://tools.ietf.org/html/rfc3986#appendix-A>

Change-Id: I7a0ab0a399d9d6cf68c824f413be6b6d621947c1
Resolves: #53862
Releases: 6.2, 6.1, 6.0
Reviewed-on: <https://review.typo3.org/25636>
Reviewed-by: Markus Klein
Reviewed-by: Wouter Wolters
Tested-by: Wouter Wolters
Reviewed-by: Michiel Roos
Tested-by: Michiel Roos
Tested-by: Markus Klein
Reviewed-by: Andreas Wolf
Reviewed-by: Jo Hasenau
Reviewed-by: Xavier Perseguers
Tested-by: Xavier Perseguers

Revision c99a07a9 - 2014-01-09 14:48 - Michiel Roos

[BUGFIX] isValidUrl() idna converts whole URI

GeneralUtility::isValidUrl() idna converts whole URI instead of domain only.

The expensive idna_convert() is called from isValidUrl(). Instead of feeding it just the domain part, the whole URI is converted.

When supplying just the domain part, a great speed gain can be seen.

On seriously malformed URLs, parse_url() may return FALSE and emits an E_WARNING. So we check for that first.

PHP no longer supports the flags FILTER_FLAG_HOST_REQUIRED and FILTER_FLAG_SCHEME_REQUIRED. A scheme is now required by default. [1]
Return FALSE if the URL does not start with a scheme.

A public GeneralUtility::idnaEncode() method uses a static idna_convert instance and fetches converted strings from a string cache array to avoid multiple checks on the same domain.

All manual idna_convert instances are replaced with GeneralUtility::idnaEncode() calls.

Special characters are not allowed in the URL except in the domain part [2]. So the test with special characters in the path was removed from the GeneralUtilityTest class.

[1] <http://www.php.net/manual/en/filter.filters.flags.php#107382>

[2] <http://tools.ietf.org/html/rfc3986#appendix-A>

Change-Id: I7a0ab0a399d9d6cf68c824f413be6b6d621947c1
Resolves: #53862
Releases: 6.2, 6.1, 6.0
Reviewed-on: <https://review.typo3.org/26501>
Reviewed-by: Xavier Perseguers
Tested-by: Xavier Perseguers
Reviewed-by: Markus Klein
Tested-by: Markus Klein

Revision d0c42764 - 2014-01-09 14:58 - Michiel Roos

[BUGFIX] isValidUrl() idna converts whole URI

GeneralUtility::isValidUrl() idna converts whole URI instead of domain only.

The expensive idna_convert() is called from isValidUrl(). Instead of feeding it just the domain part, the whole URI is converted.

When supplying just the domain part, a great speed gain can be seen.

On seriously malformed URLs, parse_url() may return FALSE and emits an E_WARNING. So we check for that first.

PHP no longer supports the flags `FILTER_FLAG_HOST_REQUIRED` and `FILTER_FLAG_SCHEME_REQUIRED`. A scheme is now required by default. [1]
Return `FALSE` if the URL does not start with a scheme.

A public `GeneralUtility::idnaEncode()` method uses a static `idna_convert` instance and fetches converted strings from a string cache array to avoid multiple checks on the same domain.

All manual `idna_convert` instances are replaced with `GeneralUtility::idnaEncode()` calls.

Special characters are not allowed in the URL except in the domain part [2]. So the test with special characters in the path was removed from the `GeneralUtilityTest` class.

[1] <http://www.php.net/manual/en/filter.filters.flags.php#107382>

[2] <http://tools.ietf.org/html/rfc3986#appendix-A>

Change-Id: I7a0ab0a399d9d6cf68c824f413be6b6d621947c1

Resolves: #53862

Releases: 6.2, 6.1, 6.0

Reviewed-on: <https://review.typo3.org/26531>

Reviewed-by: Xavier Perseguers

Tested-by: Xavier Perseguers

Reviewed-by: Markus Klein

Tested-by: Markus Klein

Revision 3dcc61df - 2014-01-30 23:46 - Wouter Wolters

[BUGFIX] Regression in DataHandler

The fix for issue #53862 calls `GeneralUtility::idnaEncode` without fully qualified class namespace.

Follow-up to: c99a07a9

Resolves: #55475

Releases: 6.1, 6.0

Change-Id: I8ba161ee73e7456da53d2182b4a22d87dad9d53c

Reviewed-on: <https://review.typo3.org/27173>

Reviewed-by: Steffen Müller

Tested-by: Steffen Müller

Reviewed-by: Christian Kuhn

Tested-by: Christian Kuhn

Revision dc271e4f - 2014-01-30 23:57 - Wouter Wolters

[BUGFIX] Regression in DataHandler

The fix for issue #53862 calls `GeneralUtility::idnaEncode` without fully qualified class namespace.

Follow-up to: c99a07a9

Resolves: #55475

Releases: 6.1, 6.0

Change-Id: I8ba161ee73e7456da53d2182b4a22d87dad9d53c

Reviewed-on: <https://review.typo3.org/27174>

Reviewed-by: Christian Kuhn

Tested-by: Christian Kuhn

Reviewed-by: Wouter Wolters

Tested-by: Wouter Wolters

History

#1 - 2013-11-22 12:14 - 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 <https://review.typo3.org/25636>

#2 - 2013-11-22 12:21 - 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 <https://review.typo3.org/25636>

#3 - 2013-11-22 13:01 - Michiel Roos

- File *fu_idna_4.7.patch* added

#4 - 2013-11-22 14:57 - 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 <https://review.typo3.org/25636>

#5 - 2013-11-22 14:58 - 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 <https://review.typo3.org/25636>

#6 - 2013-11-22 15:00 - 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 <https://review.typo3.org/25636>

#7 - 2013-11-22 15:39 - 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 <https://review.typo3.org/25636>

#8 - 2013-11-22 15:52 - Michiel Roos

- File *patch_set_6_backend_php.png* added

#9 - 2013-11-22 18:26 - 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 <https://review.typo3.org/25636>

#10 - 2013-11-22 18:46 - 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 <https://review.typo3.org/25636>

#11 - 2013-11-22 23:10 - 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 <https://review.typo3.org/25636>

#12 - 2013-11-22 23:12 - 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 <https://review.typo3.org/25636>

#13 - 2013-11-23 00:05 - 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 <https://review.typo3.org/25636>

#14 - 2013-11-23 00:58 - Gerrit Code Review

Patch set 13 for branch **master** of project **Packages/TYPO3.CMS** has been pushed to the review server.
It is available at <https://review.typo3.org/25636>

#15 - 2013-11-23 01:06 - Michiel Roos

- File *invalidFiles_login_page.txt* added

These files are tested but come back invalid.

We may gain some more speed by finding out who is testing them and why.

#16 - 2013-11-23 07:53 - Gerrit Code Review

Patch set 14 for branch **master** of project **Packages/TYPO3.CMS** has been pushed to the review server.
It is available at <https://review.typo3.org/25636>

#17 - 2013-11-23 12:10 - Michiel Roos

- File Screen_Shot_2013-11-23_at_12.08.54.png added

#18 - 2013-11-23 15:50 - Gerrit Code Review

Patch set 15 for branch **master** of project **Packages/TYPO3.CMS** has been pushed to the review server.
It is available at <https://review.typo3.org/25636>

#19 - 2013-11-23 21:07 - Gerrit Code Review

Patch set 16 for branch **master** of project **Packages/TYPO3.CMS** has been pushed to the review server.
It is available at <https://review.typo3.org/25636>

#20 - 2013-11-25 10:49 - Gerrit Code Review

Patch set 17 for branch **master** of project **Packages/TYPO3.CMS** has been pushed to the review server.
It is available at <https://review.typo3.org/25636>

#21 - 2013-11-26 10:31 - Gerrit Code Review

Patch set 18 for branch **master** of project **Packages/TYPO3.CMS** has been pushed to the review server.
It is available at <https://review.typo3.org/25636>

#22 - 2013-11-26 19:51 - Michiel Roos

- File TYPO3_CMS_Login__6.2.master.workbench.local___Workbench-7.jpg added

#23 - 2013-11-26 20:05 - Michiel Roos

- File Screen_Shot_2013-11-26_at_20.03.58.png added

Basic test against:

/typo3/backend.php from a 'just cleared the cache' state.

#24 - 2013-11-28 15:17 - Gerrit Code Review

Patch set 19 for branch **master** of project **Packages/TYPO3.CMS** has been pushed to the review server.
It is available at <https://review.typo3.org/25636>

#25 - 2013-11-28 15:26 - Gerrit Code Review

Patch set 20 for branch **master** of project **Packages/TYPO3.CMS** has been pushed to the review server.
It is available at <https://review.typo3.org/25636>

#26 - 2013-11-30 01:37 - Gerrit Code Review

Patch set 21 for branch **master** of project **Packages/TYPO3.CMS** has been pushed to the review server.
It is available at <https://review.typo3.org/25636>

#27 - 2013-11-30 01:53 - Gerrit Code Review

Patch set 22 for branch **master** of project **Packages/TYPO3.CMS** has been pushed to the review server.
It is available at <https://review.typo3.org/25636>

#28 - 2013-12-02 16:01 - Gerrit Code Review

Patch set 23 for branch **master** of project **Packages/TYPO3.CMS** has been pushed to the review server.
It is available at <https://review.typo3.org/25636>

#29 - 2013-12-05 11:35 - Gerrit Code Review

Patch set 24 for branch **master** of project **Packages/TYPO3.CMS** has been pushed to the review server.
It is available at <https://review.typo3.org/25636>

#30 - 2013-12-05 12:28 - Gerrit Code Review

Patch set 25 for branch **master** of project **Packages/TYPO3.CMS** has been pushed to the review server.
It is available at <https://review.typo3.org/25636>

#31 - 2013-12-19 20:13 - Gerrit Code Review

Patch set 1 for branch **TYPO3_6-1** of project **Packages/TYPO3.CMS** has been pushed to the review server.
It is available at <https://review.typo3.org/26501>

#32 - 2013-12-20 11:26 - Michiel Roos

- File Patch_25_backendPhp_unpopulated_Cache_dir.png added
- File Patch_25_dbList_siteRoot_populated_Cache_dir.jpg added
- File Patch_25_login_unpopulated_Cache_dir.jpg added

The numbers for Helmut.

Helmut and others . . .

Please consider verifying the numbers using:

- <https://github.com/Tuurlijk/Typo3Workbench>
- <https://github.com/Tuurlijk/xhprof>

;-)

#33 - 2013-12-21 15:52 - Gerrit Code Review

Patch set 1 for branch **TYPO3_6-0** of project **Packages/TYPO3.CMS** has been pushed to the review server. It is available at <https://review.typo3.org/26531>

#34 - 2013-12-21 16:30 - Michiel Roos

- Status changed from Under Review to Resolved
- % Done changed from 0 to 100

Applied in changeset [f8fdcea7949a5c71f6078f7dda76f08881f56c01](https://review.typo3.org/26531).

#35 - 2014-01-09 16:39 - Markus Klein

- Parent task set to #52949

#36 - 2018-10-02 12:09 - Benni Mack

- Status changed from Resolved to Closed

Files

backend_php.png	34.2 KB	2013-11-22	Michiel Roos
login_screen.png	34.4 KB	2013-11-22	Michiel Roos
web_list.png	34.7 KB	2013-11-22	Michiel Roos
fu_idna.v3.diff	1.01 KB	2013-11-22	Michiel Roos
fu_idna_4.7.patch	1.07 KB	2013-11-22	Michiel Roos
patch_set_6_backend_php.png	36.8 KB	2013-11-22	Michiel Roos
invalidFiles_login_page.txt	6.48 KB	2013-11-23	Michiel Roos
Screen_Shot_2013-11-23_at_12.08.54.png	38.9 KB	2013-11-23	Michiel Roos
TYPO3_CMS_Login__6.2.master.workbench.local___Workbench-7.jpg	39 KB	2013-11-26	Michiel Roos
Screen_Shot_2013-11-26_at_20.03.58.png	102 KB	2013-11-26	Michiel Roos
Patch_25_backendPhp_unpopulated_Cache_dir.png	35.4 KB	2013-12-20	Michiel Roos
Patch_25_dbList_siteRoot_populated_Cache_dir.jpg	45.4 KB	2013-12-20	Michiel Roos
Patch_25_login_unpopulated_Cache_dir.jpg	43.6 KB	2013-12-20	Michiel Roos