

python / cpython Public

<> Code Issues 5k+ Pull requests 2.2k Actions Projects Security and q

gh-133767: Fix use-after-free in the unicode-escape decoder with an error handler #129648

Merged

serhiy-storchaka merged 3 commits into python:main from

serhiy-storchaka:unicode-escape-d... on May 12, 2025

Conversation 14 Commits 3 Checks 73 Files changed 8



serhiy-storchaka commented on Feb 4, 2025 • edited by bedevere-app Bot

Member

If the error handler is used, a new bytes object is created to set as the object attribute of UnicodeDecodeError, and that bytes object then replaces the original data. A pointer to the decoded data will become invalid after destroying that temporary bytes object. So we need other way to return the first invalid escape from `_PyUnicode_DecodeUnicodeEscapeInternal()`.

`_PyBytes_DecodeEscape()` does not have such issue, because it does not use the error handlers registry, but it should be changed for compatibility with `_PyUnicode_DecodeUnicodeEscapeInternal()`.

- Issue: [Use-after-free in unicode_escape decoder with error handler #133767](#)



Fix use-after-free in the unicode-escape decoder with error handler

3a939ff



serhiy-storchaka requested review from Yhg1s, ericvsmith and sethmlarson last year

Member

gpshead commented on Feb 4, 2025

Nice! This is similar enough, but clearly far more polished, than what I quickly whipped up while trying to understand the problem and linked to on the PSRT mailing list... that I won't bother posting my own draft PR.

I don't have a good feel for if we need to retain the older internal-use-only C APIs or not, but doing this change via ones with a suffix as you seem to be proposing and leaving the old, though now unused by our own internals, ones in place in case something else references them makes sense to me.

serhiy-storchaka commented on [Feb 4, 2025](#)

Member

Author

I experimented with several different solutions. One of them was similar to yours, except that I copied all three bytes. It was also necessary to distinguish "no invalid escape" from "escaped null byte". In the end, the currently proposed solution is the simplest.

This PR does not leave the old C API. I do not think that it is needed. The functions are renamed because an error at link time is more preferable than undefined behavior at run time.



1

 **gpshead** reviewed on [Feb 4, 2025](#)

[View reviewed changes](#)

> `Parser/string_parser.c` Outdated

 [Show resolved](#)



 **serhiy-storchaka** added 2 commits [11 months ago](#)

  [Merge branch 'main' into unicode-escape-decode-errors](#)

 [be5d80c](#)

  [Add a NEWS entry.](#)

 [7194b4d](#)

  **serhiy-storchaka** changed the title [Fix use-after-free in the unicode-escape decoder with error handler gh-133767: Fix use-after-free in the unicode-escape decoder with an error handler](#) on [May 9, 2025](#)

  **bedevere-app** Bot mentioned this pull request on [May 9, 2025](#)

[Use-after-free in unicode_escape decoder with error handler #133767](#)

 Closed

  **serhiy-storchaka** marked this pull request as ready for review [11 months ago](#)

  **serhiy-storchaka** requested review from [lysnikolaou](#) and [pablogsal](#) as [code owners](#) [11 months ago](#)

 **bedevere-app** (Bot) added the **awaiting core review** label on May 9, 2025

serhiy-storchaka commented on May 9, 2025

Member

Author


After adding a NEWS entry it is ready for review.

The code is now more complex, decoding functions now return both the invalid char and its positions. This is because the new code in the Python parser needs the position. It can be returned if there was no decoding errors handled by the error handler. The Python parser does not use the error handler.



 **gpshead** approved these changes on May 10, 2025

[View reviewed changes](#)

 **bedevere-app** (Bot) added **awaiting merge** and removed **awaiting core review** labels on May 10, 2025

 **gpshead** added **needs backport to 3.13** **needs backport to 3.14** **test-with-buildbots** labels on May 10, 2025

bedevere-bot commented on May 10, 2025

 New build scheduled with the buildbot fleet by [@gpshead](#) for commit [7194b4d](#) 

Results will be shown at:

<https://buildbot.python.org/all/#/grid?branch=refs%2Fpull%2F129648%2Fmerge>

If you want to schedule another build, you need to add the **test-with-buildbots** label again.

 **bedevere-bot** removed the **test-with-buildbots** label on May 10, 2025

 **serhiy-storchaka** added **type-security** **needs backport to 3.9** **needs backport to 3.10** **needs backport to 3.11** **needs backport to 3.12** labels on May 10, 2025

 **serhiy-storchaka** merged commit **9f69a58** into **python:main** on May 12, 2025

140 of 142 checks passed

[View details](#)

543 hidden items

[Load more...](#)

miss-islington-app Bot commented [on May 12, 2025](#)

Sorry, [@serhiy-storchaka](#), I could not cleanly backport this to `3.12` due to a conflict.
Please backport using [cherry_picker](#) on command line.

```
cherry_picker 9f69a58623bd01349a18ba0c7a9cb1dad6a51e8e 3.12
```



miss-islington-app Bot commented [on May 12, 2025](#)

Sorry, [@serhiy-storchaka](#), I could not cleanly backport this to `3.11` due to a conflict.
Please backport using [cherry_picker](#) on command line.

```
cherry_picker 9f69a58623bd01349a18ba0c7a9cb1dad6a51e8e 3.11
```



miss-islington-app Bot commented [on May 12, 2025](#)

Sorry, [@serhiy-storchaka](#), I could not cleanly backport this to `3.10` due to a conflict.
Please backport using [cherry_picker](#) on command line.


```
cherry_picker 9f69a58623bd01349a18ba0c7a9cb1dad6a51e8e 3.10
```



miss-islington-app Bot commented [on May 12, 2025](#)

Sorry, [@serhiy-storchaka](#), I could not cleanly backport this to `3.9` due to a conflict.
Please backport using [cherry_picker](#) on command line.


```
cherry_picker 9f69a58623bd01349a18ba0c7a9cb1dad6a51e8e 3.9
```

 **serhiy-storchaka** added a commit to serhiy-storchaka/cpython that referenced this pull request on May 12, 2025

 [3.13] [pythongh-133767](#): [Fix use-after-free in the unicode-escape deco...](#) ... ✓ [0afcdf1](#)


bedevere-app Bot commented on May 12, 2025

[GH-133944](#) is a backport of this pull request to the [3.13 branch](#).

 **bedevere-app** Bot removed the **needs backport to 3.13** label on May 12, 2025

 **serhiy-storchaka** removed **needs backport to 3.9** **needs backport to 3.10** **needs backport to 3.11** **needs backport to 3.12** labels on May 12, 2025

 **serhiy-storchaka** removed their assignment on May 12, 2025

 **serhiy-storchaka** added a commit that referenced this pull request on May 13, 2025

 [3.14] [gh-133767](#): [Fix use-after-free in the unicode-escape decoder wi...](#) ... ✗ [69b4387](#)

 This was referenced on May 16, 2025

[Change: Replace global ignore_max_rows_per_page greenbone/gvmd#2447](#)

 Merged

[Throw an error when providing an invalid LDAP certificate. greenbone/gvmd#2442](#)

 Merged

[Deps: Bump ruff from 0.11.9 to 0.11.10 in the python-packages group greenbone/greenbone-feed-sync#278](#)

 Merged

 **jamie-albert** mentioned this pull request on May 19, 2025

[python-3.10/11/12/13/CVE-2025-4516 adv update wolfi-dev/advisories#18858](#)

Merged

mcepl mentioned this pull request on May 19, 2025

[3.12] Fix use-after-free in the unicode-escape decoder with error handler (GH-133767)

#134255

Closed

encukou pushed a commit that referenced this pull request on May 20, 2025

[3.13] gh-133767: Fix use-after-free in the unicode-escape decoder wi... 6279eb8

serhiy-storchaka added a commit to serhiy-storchaka/cpython that referenced this pull request on May 20, 2025

[3.12] pythongh-133767: Fix use-after-free in the unicode-escape deco... a75953b

serhiy-storchaka added a commit to serhiy-storchaka/cpython that referenced this pull request on May 20, 2025

[3.11] pythongh-133767: Fix use-after-free in the unicode-escape deco... 0c33e5b

serhiy-storchaka added a commit to serhiy-storchaka/cpython that referenced this pull request on May 20, 2025

[3.10] pythongh-133767: Fix use-after-free in the unicode-escape deco... 8b528ca

serhiy-storchaka added a commit to serhiy-storchaka/cpython that referenced this pull request on May 20, 2025

[3.9] pythongh-133767: Fix use-after-free in the unicode-escape decod... 0d5d68f

Yhg1s pushed a commit that referenced this pull request on May 26, 2025

[3.12] gh-133767: Fix use-after-free in the unicode-escape decoder wi... 4398b78


ambv pushed a commit that referenced this pull request on Jun 2, 2025

[3.11] gh-133767: Fix use-after-free in the unicode-escape decoder wi... 73b3040


 **ambv** pushed a commit that referenced this pull request [on Jun 2, 2025](#)

 [3.10] [gh-133767](#): [Fix use-after-free in the unicode-escape decoder wi...](#) ...  [ab9893c](#)

 **ambv** pushed a commit that referenced this pull request [on Jun 2, 2025](#)

 [3.9] [gh-133767](#): [Fix use-after-free in the unicode-escape decoder wit...](#) ...  [8d35fd1](#)

 **Pranjal095** pushed a commit to [Pranjal095/cpython](#) that referenced this pull request [on Jul 12, 2025](#)

 [pythongh-133767](#): [Fix use-after-free in the unicode-escape decoder wit...](#) ... [dfd3ae0](#)

 **gentoo-bot** pushed a commit to [gentoo/cpython](#) that referenced this pull request [on Jul 30, 2025](#)

 [pythongh-133767](#): [Fix use-after-free in the unicode-escape decoder wit...](#) ... [6e77767](#)

 **taegyunkim** pushed a commit to [taegyunkim/cpython](#) that referenced this pull request [on Aug 4, 2025](#)

 [pythongh-133767](#): [Fix use-after-free in the unicode-escape decoder wit...](#) ... [9d30d3b](#)

[Sign up for free](#) to join this conversation on **GitHub**. Already have an account? [Sign in to comment](#)

Reviewers

-  **gpshead** ✓
-  **Yhg1s** ●
-  **ericvsmith** ●
-  **sethmlarson** ●
-  **pablogsal** 🛡️ ●
-  **lysnikolaou** 🛡️ ●

Assignees

No one assigned

Labels

[type-security](#)

Projects

None yet

Milestone

No milestone

Development

Successfully merging this pull request may close these issues.

None yet

3 participants

