

python / cpython Public

<> Code Issues 5k+ Pull requests 2.1k Actions Projects Security and q

gh-135462: Fix quadratic complexity in processing special input in HTMLParser #135464

Merged

serhiy-storchaka merged 3 commits into python:main from serhiy-storchaka:htmlparser-quadr... on Jun 13, 2025

Conversation 19 Commits 3 Checks 53 Files changed 3



serhiy-storchaka commented on Jun 13, 2025 • edited by bedevere-app Bot Member

- Issue: [Worst case quadratic complexity in HTMLParser #135462](#)

serhiy-storchaka requested review from ezio-melotti and sethmlarson 11 months ago

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 needs backport to 3.13 needs backport to 3.14 labels on Jun 13, 2025



bedevere-app Bot added the awaiting core review label on Jun 13, 2025

bedevere-app Bot mentioned this pull request on Jun 13, 2025

Worst case quadratic complexity in HTMLParser #135462

Closed

pythongh-135462: Fix quadratic complexity in processing special input... ✓ c87cb49

  **serhiy-storchaka** [force-pushed](#) the `htmlparser-quadratic-complexity` branch from `77d5125` to `c87cb49` [11 months ago](#) [Compare](#)

serhiy-storchaka commented [on Jun 13, 2025](#)

Member

Author

The solution has been written in a way that simplifies backporting. There are other issues, and the code will be refactored in new versions after fixing them.

 **ezio-melotti** reviewed [on Jun 13, 2025](#)

[View reviewed changes](#)

Lib/test/test_htmlparser.py

Comment on lines [+724](#) to [+727](#)

```
724 + @support.requires_resource('cpu')
725 + def test_eof_no_quadratic_complexity(self):
726 +     # Each of these examples used to take about an hour.
727 +     # Now they take a fraction of a second.
```



ezio-melotti [on Jun 13, 2025](#)

Member

If they now take a fraction of a second, is there a reason to require the `cpu` resource?

My understanding is that:

- with the `requires_resource('cpu')` decorator:
 - this test would normally be skipped
 - in case of regression, we won't notice unless the `cpu` is enabled
- without the decorator:
 - this test is always run and completes quickly
 - in case of regression, the test will timeout/fail and expose the problem



serhiy-storchaka [on Jun 13, 2025](#)

Member

Author

They totally take 1.3 seconds on my computer. All other tests take 0.1-0.2 seconds. It is a waste of time to run it several times for every update of any PR. Some buildbots are slower than my computer.

I think that it is enough to run this test only on the fastest buildbots. We already used `requires_resource('cpu')` in similar tests.



Lib/test/test_htmlparser.py

```

433 | - | ('data', '\n.

I have update the NEWS entry.



> Lib/html/parser.py Outdated

Show resolved

> Lib/html/parser.py Outdated

Show resolved



**serhiy-storchaka** and others added 2 commits 11 months ago



**serhiy-storchaka** Apply suggestions from code review ...

✗ [fe9730d](#)



**serhiy-storchaka** Update a NEWS entry.

✓ [5fed8ef](#)

✓ **ezio-melotti** approved these changes [on Jun 13, 2025](#)

[View reviewed changes](#)

**bedevere-app** (Bot) added **awaiting merge** and removed **awaiting core review** labels [on Jun 13, 2025](#)

✓ **sethmlarson** approved these changes [on Jun 13, 2025](#)

[View reviewed changes](#)

**sethmlarson** left a comment Contributor

Approach seems sensible to me.

**serhiy-storchaka** merged commit **6eb6c5d** into `python:main` [on Jun 13, 2025](#) [View details](#)

43 checks passed

**miss-islington-app** (Bot) commented [on Jun 13, 2025](#)

Thanks [@serhiy-storchaka](#) for the PR 🙌🎉... I'm working now to backport this PR to: 3.9, 3.10, 3.11, 3.12, 3.13, 3.14.



**bedevere-app** (Bot) removed the **awaiting merge** label [on Jun 13, 2025](#)

**serhiy-storchaka** deleted the `htmlparser-quadratic-complexity` branch [11 months ago](#)

**miss-islington** pushed a commit to `miss-islington/cpython` that referenced this pull request [on Jun 13, 2025](#)

**pythongh-135462**: [Fix quadratic complexity in processing special input...](#) ... ✓ [ebd88a8](#)

**miss-islington** pushed a commit to `miss-islington/cpython` that referenced this pull request [on Jun 13, 2025](#)

**pythongh-135462**: [Fix quadratic complexity in processing special input...](#) ... ✓ [9043eda](#)

**bedevere-app** Bot commented on Jun 13, 2025

[GH-135481](#) is a backport of this pull request to the [3.14 branch](#).

 **bedevere-app** Bot removed the **needs backport to 3.14** label on Jun 13, 2025

**miss-islington-app** Bot commented on Jun 13, 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 6eb6c5dbfb528bd07d77b60fd71fd05d81d45c41 3.12
```



97 hidden items

[Load more...](#)

 This was referenced on Aug 14, 2025

**Change: move asset host iterator to asset files** [greenbone/gvmd#2514](#)

 Merged

**Change: move asset OS iterator to asset files** [greenbone/gvmd#2515](#)

 Merged

**Remove delete\_agent\_installer GMP command** [greenbone/gvmd#2516](#)

 Merged

**Deps: Bump the python-packages group with 3 updates** [greenbone/greenbone-feed-sync#292](#)

 Merged

**Use different commands for adding user and group to container** [greenbone/gvm-tools#1243](#)

 Merged

**Deps: Bump actions/checkout from 4 to 5 in the github-actions group**[greenbone/greenbone-feed-sync#293](#) Merged**Bump actions/checkout from 4 to 5 greenbone/gvmd#2517** Merged**Add: Add support for creating tasks with OCI image targets. greenbone/gvmd#2497** Merged**Fix: Use "filter\_string" param in scripts greenbone/gvm-tools#1244** Merged**Change: move remaining iterators to the asset files greenbone/gvmd#2518** Merged**Deps: Bump the github-actions group with 2 updates greenbone/gvm-tools#1245** Merged

**Agent-Hellboy** pushed a commit to Agent-Hellboy/cpython that referenced this pull request on Aug 19, 2025



[pythongh-135462](#): [Fix quadratic complexity in processing special input...](#) ...

[4bce1bc](#)



This was referenced on Aug 19, 2025

**Change: move delete and modify functions to asset files greenbone/gvmd#2519** Merged**Change: move report\_host\_\* to the asset files greenbone/gvmd#2520** Merged**Fix CI greenbone/ospd-openvas#1054** Merged**Fix: Add missing and amend G\_LOG\_DOMAIN #defines greenbone/gvmd#2521** Merged**Add scan agent config support greenbone/gvmd#2522** Merged**Change: move rest of Assets section out of manage\_sql.c greenbone/gvmd#2523**

 Merged

**Fix result and report id fetching in monthly-report script** [greenbone/gvm-tools#1246](#)

 Merged

**Change: move host\_notice into asset files** [greenbone/gvmd#2524](#)

 Merged

**Deps: Bump the python-packages group with 5 updates** [greenbone/greenbone-feed-sync#294](#)

 Merged

**Add: Ability to get, modify and verify scanners of type 'container-image'** [greenbone/gvmd#2525](#)

 Merged

**Fix: Behavior with ENABLE\_AGENTS feature flag** [greenbone/gvmd#2526](#)

 Merged

This was referenced on Aug 27, 2025

**Remove separate default openvasd scanner** [greenbone/gvmd#2527](#)

 Merged

**Fix: handle 0 ports case in get\_port\_lists filter** [greenbone/gvmd#2528](#)

 Merged

**Fix: Address warnings when building with the gcc version in Debian testing** [greenbone/gvmd#2529](#)

 Merged

**Add: Filter settings for compliance types** [greenbone/gvmd#2530](#)

 Merged

**Fix: Add check if UUID exists when creating config** [greenbone/gvmd#2531](#)

 Merged

**Change: Update openvasd to use the generic http scanner library** [greenbone/gvmd#2532](#)

 Merged

 waylan mentioned this pull request on Sep 24, 2025

## Malformed HTML comment closer -- > causes Markdown parser to hang / OOM

[Python-Markdown/markdown#1554](#)

 Closed

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

### Reviewers

-  **sethmlarson** 
-  **ezio-melotti** 

### Assignees

-  **serhiy-storchaka**

### Labels

- 

### Projects

None yet

### Milestone

No milestone

### Development

Successfully merging this pull request may close these issues.

None yet

### 3 participants

