

cpan-authors / XML-Parser Publicforked from [chorny/XML-Parser](#)[Code](#) [Issues](#) [Pull requests](#) 3 [Actions](#) [Projects](#) [Security and quality](#)

New issue



# XML::Parser::Expat crashes on utf8 stream [rt.cpan.org #19859] #64

Closed#109

Labels

Has Patch

todd opened on Sep 24, 2019

Member

Migrated from [rt.cpan.org#19859](#) (status was 'open')

Requestors:

- [rantwijk@science.uva.nl](mailto:rantwijk@science.uva.nl)

Attachments:

- [XML-Parser-2.34-unicodcrash.patch](#)

From on 2006-06-13 08:49:30

:

```
I encountered a UTF-8 related bug in the Expat library wrapper.  
The symptom of this bug is a Perl interpreter crash with the following  
error message:  
*** glibc detected *** double free or corruption (!prev): 0x081e2c00 ***
```



```
This error is caused by heap corruption from a buffer overflow in  
Expat.xs, line 388:  
Copy(tb, buffer, br, char)
```

```
This buffer overflow happens because the code assumes that the number of  
bytes copied (br) will never exceed the number of characters read from  
the input (buffsize). This assumption is invalid if the input stream is  
in utf8 mode.
```

```
The best solution is to have the Perl programmer set the stream to raw
```

mode, since this is also what libexpat expects. I think however, that the internal buffer overflow should be fixed anyway. The encoding issues could also be documented more clearly.

Sample program which triggers the bug on certain input files:

```
---
use strict;
use encoding 'utf8';
use XML::Parser;
# (if i uncomment this, the bug disappears) binmode(STDIN, ':bytes');
my $parser = XML::Parser->new( Style => 'Debug' );
$parser->parse(\*STDIN);
---
```

If the package maintainer agrees that this bug should be fixed, I am willing to provide a patch and do some testing. Just let me know if this is appreciated.

Package: XML-Parser-2.34  
Perl version: v5.8.5 built for i386-linux-thread-multi  
OS: Fedora Core release 3

Bye,  
Joris.

From [atourbin@cpan.org](mailto:atourbin@cpan.org) on 2006-08-11 00:46:06

:

> If the package maintainer agrees that this bug should be fixed, I am  
> willing to provide a patch and do some testing. Just let me know if  
this  
> is appreciated.



It looks like XML::Parser is unmaintained for quite some time, which is contrary to its wide-spread use. The patch for the problem you've reported can be appreciated not only by maintainer, but also by software vendors and ultimately by perl developers and users. So... If you have a patch, please post it here.

--  
Alexey Tourbin  
ALT Linux Team

PS: I filed a few bugs on XML::Parser too, e.g. #11917 and #13204.

From [rantwijk@science.uva.nl](mailto:rantwijk@science.uva.nl) on 2006-08-11 05:54:29

:

On Thu, 2006-08-10 at 20:47 -0400, via RT wrote:

> It looks like XML::Parser is unmaintained for quite some time, which  
> is contrary to its wide-spread use. The patch for the problem you've  
> reported can be appreciated not only by maintainer, but also by  
> software vendors and ultimately by perl developers and users. So...  
> If you have a patch, please post it here.

I have a patch that seems to work (attached to this message).  
This patch has also been posted to the Debian bug tracking system:  
<http://bugs.debian.org/cgi-bin/bugreport.cgi?bug=378411>



The patch only fixes the overflow condition. The issue remains that  
Expat expects its input to be raw (encoded) bytes, while the Perl  
programmer may accidentally pass it a decoded stream. This will lead to  
double decoding and incorrect XML parsing (except when the encoding was  
utf8, because Perl happens to use utf8 internally for unicode strings).

Perhaps we should say in the Expat.pm documentation that input streams  
must be set to ':raw' mode.

Joris.



  **toddrr** added **Has Patch** on Sep 24, 2019

  **toddrr-bot** added a commit that references this issue 3 weeks ago



Fix buffer overflow in parse\_stream when filehandle has :utf8 layer 

5361c2b

  **toddrr-bot** mentioned this 3 weeks ago

 [fix: buffer overflow in parse\\_stream on utf8 filehandle #109](#)


  **toddrr** closed this as completed in [#109](#) 3 weeks ago

 **hartwork** 2 weeks ago 

This seems to be [CVE-2006-10002](#)

 1

  **LEEKIYOON-SEC** mentioned this 2 weeks ago

 [\[Argus\] CVE-2006-10002: Perl의 XML::Parser 2.47 버전 이하에서 힙 손상 및 충돌 가능성](#)  
[LEEKIYOON-SEC/Argus-AI-Threat-Intelligence#461](#)

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

### Metadata

#### Assignees

No one assigned

#### Labels

**Has Patch**

#### Type

No type

#### Projects

No projects



#### Milestone


No milestone

#### Relationships

None yet

#### Development

 Code with agent mode 

 **fix: buffer overflow in parse\_stream on utf8 filehandle**

[cpan-authors/XML-Parser](#)

#### Participants

