» Issue34656



This issue tracker **has been migrated to GitHub**, and is currently **read-only**.
For more information, **see the GitHub FAQs in the Python's Developer Guide.**



This issue has been migrated to GitHub:
https://github.com/python/cpython/issues/78837

classification

Title: [CVE-2018-20406] memory exhaustion in Modules/_pickle.c:1393	
Type: security	**Stage**: resolved
Components: Library (Lib)	**Versions**: Python 3.8, Python 3.7, Python 3.6, Python 3.4, Python 3.5

process

Status: closed	**Resolution**: fixed
Dependencies:	**Superseder**:
Assigned To:	**Nosy List**: dfmz77669, larry, ned.deily, vstinner
Priority: normal	**Keywords**: patch

Created on **2018-09-13 04:38** by **shuoz**, last changed **2022-04-11 14:59** by **admin**. This issue is now **closed**.

Files

File name	Uploaded	Description	Edit
poc	shuoz, 2018-09-13 04:38		
pk.py	shuoz, 2018-09-13 04:53		
CVE-2018-20406-pickle_LONG_BINPUT.patch	mcepl, 2019-01-23 17:24	Patch updated for Python 3.4.*	

Pull Requests

URL	Status	Linked	Edit
PR 9261	merged	benjamin.peterson, 2018-09-14 00:55	
PR 9465	merged	miss-islington, 2018-09-21 01:36	
PR 9466	merged	miss-islington, 2018-09-21 01:37	
PR 11869	merged	vstinner, 2019-02-15 11:44	
PR 11870	merged	vstinner, 2019-02-15 11:46	

Messages (18)

msg325230 - (view) **Author: shuoz (shuoz)** **Date: 2018-09-13 04:38**

```
python version:
   Python 3.8.0a0 (heads/master:4ae8ece, Sep 13 2018, 09:48:16)
   [GCC 5.4.0 20160609] on linux


I found a bug in python pickle.load func. Can cause memory exhaustion DDOS.

./python pk.py poc


cat ./pk.py
import pickle
import sys
filename = sys.argv[1]
with open(filename, 'rb') as f:
    aa = pickle.load(f)
    print(aa)
```

msg325231 - (view) **Author: shuoz (shuoz)** **Date: 2018-09-13 04:55**

```
[----------------------------------registers-----------------------------------]
RAX: 0x7ff9d401e010 --> 0x0
RBX: 0x7ffff7f48d00 --> 0x1
RCX: 0x7ff8ab58c800 --> 0x7ffff7ea5d80 --> 0x2
RDX: 0x7ffff3ac47d8 --> 0x1
RSI: 0x25152303
RDI: 0xfff3a803c00 --> 0x0
RBP: 0x7473078c
RSP: 0x7fffffffcf20 --> 0x7ffff3ac47d8 --> 0x1
RIP: 0x7ffff28a8a64 (<_Unpickler_MemoPut+1668>: add    r11,0x20)
R8 : 0xfff3a803bff --> 0x0
R9 : 0xfff3a803c01 --> 0x0
R10: 0xffffefe91a3 --> 0x0
R11: 0x128a917f8 --> 0x0
R12: 0xfff156b1922 --> 0x0
R13: 0xe8e60f18 --> 0x0
R14: 0x7ffff7f48d18 --> 0x7ff8ab58c800 --> 0x7ffff7ea5d80 --> 0x2
R15: 0xfff3a803c02 --> 0x0
EFLAGS: 0x216 (carry PARITY ADJUST zero sign trap INTERRUPT direction overflow)
[-------------------------------------code-------------------------------------]
   0x7ffff28a8a52 <_Unpickler_MemoPut+1650>:    cmp    BYTE PTR
[r15+0x7fff8000],0x0
   0x7ffff28a8a5a <_Unpickler_MemoPut+1658>:    jne    0x7ffff28a8ae1
<_Unpickler_MemoPut+1793>
   0x7ffff28a8a60 <_Unpickler_MemoPut+1664>:    add    rsi,0x4
=> 0x7ffff28a8a64 <_Unpickler_MemoPut+1668>:    add    r11,0x20
   0x7ffff28a8a68 <_Unpickler_MemoPut+1672>:    cmp    BYTE PTR
[r10+0x7fff8000],0x0
   0x7ffff28a8a70 <_Unpickler_MemoPut+1680>:    mov    QWORD PTR [rax],0x0
   0x7ffff28a8a77 <_Unpickler_MemoPut+1687>:    je     0x7ffff28a896d
<_Unpickler_MemoPut+1421>
   0x7ffff28a8a7d <_Unpickler_MemoPut+1693>:    nop    DWORD PTR [rax]
[-------------------------------------stack-------------------------------------]
0000| 0x7fffffffcf20 --> 0x7ffff3ac47d8 --> 0x1
0008| 0x7fffffffcf28 --> 0xffffefe91a3 --> 0x0
```

```
0016| 0x7fffffffcf30 --> 0x7ffff7f48da8 --> 0x20 (' ')
0024| 0x7fffffffcf38 --> 0x7ffff7f48d00 --> 0x1
0032| 0x7fffffffcf40 --> 0xffffffffa00 --> 0x0
0040| 0x7fffffffcf48 --> 0x0
0048| 0x7fffffffcf50 --> 0x7ffff7f48da0 --> 0x28 ('(')
0056| 0x7fffffffcf58 --> 0x7ffff7f48da8 --> 0x20 (' ')
[-------------------------------------------------------------------------------]
Legend: code, data, rodata, value
0x00007ffff28a8a64      1392        for (i = self->memo_size; i < new_size; i++)
gdb-peda$ p new_size
$5 = 0xe8e60f18
gdb-peda$ p self->memo_size
$6 = 0x20
gdb-peda$ p i


.....
for (i = self->memo_size; i < new_size; i++)
        self->memo[i] = NULL;
.....
```

msg325430 - (view) **Author: Serhiy Storchaka (serhiy.storchaka) *** **Date: 2018-09-15 11:02**

```
>>> import pickletools
>>>
pickletools.dis(b'\x80\x04\x95\x1d\x00\x00\x00\x00\x00\x00\x00}\x94(\x8c\x03age\x94
K\x17\x8c\x03jobr\x8c\x07student\x94u.')
    0: \x80 PROTO      4
    2: \x95 FRAME      29
   11: }    EMPTY_DICT
   12: \x94 MEMOIZE    (as 0)
   13: (    MARK
   14: \x8c     SHORT_BINUNICODE 'age'
   19: \x94     MEMOIZE    (as 1)
   20: K        BININT1    23
   22: \x8c     SHORT_BINUNICODE 'job'
   27: r        LONG_BINPUT 1953695628
   32: u        SETITEMS   (MARK at 13)
   33: d    DICT       no MARK exists on stack
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
  File "/home/serhiy/py/cpython/Lib/pickletools.py", line 2457, in dis
    raise ValueError(errormsg)
ValueError: no MARK exists on stack
```

Ignore the error of unbalanced MARK. The problem code is LONG_BINPUT with the excessive large argument **1953695628**. The C implementation of pickle tries to resize the the memo list to the size twice larger than this index. And here an integer overflow occurred.

This unlikely occurred in real world. The pickle needs to have more than 2**30-1 ≈ 10**9 memoized items for encountering this bug. It means that its size on disk and in memory should be tens or hundreds of gigabytes. Pickle is not the best format for serializing such amount of data.

msg325937 - (view) **Author: Benjamin Peterson (benjamin.peterson) *** **Date: 2018-09-21 01:36**

New changeset **a4ae828ee416a66d8c7bf5ee71d653c2cc6a26dd** by Benjamin Peterson in branch 'master':
closes ~~**bpo-34656**~~: Avoid relying on signed overflow in _pickle memos. (~~**GH-9261**~~)
https://github.com/python/cpython/commit/a4ae828ee416a66d8c7bf5ee71d653c2cc6a26dd

msg325938 - (view) **Author: miss-islington (miss-islington)** **Date: 2018-09-21 01:52**

New changeset **ef4306b24c9034d6b37bb034e2ebe82e745d4b77** by Miss Islington (bot) in branch '3.7':
closes ~~**bpo-34656**~~: Avoid relying on signed overflow in _pickle memos. (~~**GH-9261**~~)
https://github.com/python/cpython/commit/ef4306b24c9034d6b37bb034e2ebe82e745d4b77

msg325939 - (view) **Author: miss-islington (miss-islington)** **Date: 2018-09-21 02:00**

New changeset **71a9c65e74a70b6ed39adc4ba81d311ac1aa2acc** by Miss Islington (bot) in branch '3.6':
closes ~~**bpo-34656**~~: Avoid relying on signed overflow in _pickle memos. (~~**GH-9261**~~)
https://github.com/python/cpython/commit/71a9c65e74a70b6ed39adc4ba81d311ac1aa2acc

msg333292 - (view) **Author: Miro Hrončok (hroncok) *** **Date: 2019-01-09 08:08**

Should this go to 3.4 and 3.5 as well, since it is a security thing?

http://people.canonical.com/~ubuntu-security/cve/2018/CVE-2018-20406.html

msg333294 - (view) **Author: Serhiy Storchaka (serhiy.storchaka) *** **Date: 2019-01-09 09:33**

I am not sure this issue should be classified as a security issue. It can cause DDOS, because pickle should not be used with untrusted data. If it is used, the program has more severe security issues than just DDOS.

The crash could be triggered by accident, but this is very unlikely. I doubts that this happened even once in real world. Libraries used for handling a large amount of data (like NumPy) use more efficient pickle representation, and can provide even more efficient alternate serialization methods. Note that integers and floats are not memoized, this increases the complexity and size of data that could be affected by this bug.

But I think that this fix needs a news entry. Do you mind to add it Benjamin?

msg334108 - (view) **Author: Matej Cepl (mcepl) *** **Date: 2019-01-20 23:18**

Does it even make sense to make a security patch for 2.7 for this one?

msg334179 - (view) **Author: STINNER Victor (vstinner) *** **Date: 2019-01-21 20:58**

Python 2.7 is not affected:

* Python 2.7 has no C accelerator _pickle (**Modules/_pickle.c**)
* Python 2.7 doesn't support protocol 4 (attached proof of concept)

I reopen the issue because the issue should be fixed in 3.4 and 3.5 as well, since it has been marked as a vulnerability (it got a CVE number).

msg334183 - (view) **Author: Matej Cepl (mcepl) *** **Date: 2019-01-21 21:47**

> * Python 2.7 has no C accelerator _pickle (**Modules/_pickle.c**)

And **Modules/cPickle.c** is that drastically different?

msg334208 - (view) **Author: STINNER Victor (vstinner) *** **Date: 2019-01-22 12:24**

> And **Modules/cPickle.c** is that drastically different?

Stupid me. I was surprised that Python 2.7 had no C accelerator. I was looking for Modules/*pickle*.c on my case sensitive Linux filesystem...

msg334209 - (view) **Author: STINNER Victor (vstinner) *** **Date: 2019-01-22 12:34**

New changeset **a4ae828ee416a66d8c7bf5ee71d653c2cc6a26dd** by Benjamin Peterson in branch 'master':
closes ~~**bpo-34656**~~: Avoid relying on signed overflow in _pickle memos. (~~**GH-9261**~~)
https://github.com/python/cpython/commit/a4ae828ee416a66d8c7bf5ee71d653c2cc6a26dd

It seems like this patch changes the implementation of the internal "memo" object which is a custom C type in Python 3.

In Python 2 cPickle, the memo is a regular dictionary and so I'm not sure that Python 2 is affected by this vulnerability.

Can someone please confirm?

msg334267 - (view) **Author: Matej Cepl (mcepl) *** **Date: 2019-01-23 17:24**

Python 3.4 doesn't allow C99 constructs, so I had to update the patch to reorder iterator declarations. Just if any future colleague Python Linux distro maintainer needs it.

msg336381 - (view) **Author: dfmz77669 (dfmz77669)** **Date: 2019-02-23 14:14**

In python2, Picklertype donot has tp init which has bug in python3 Pickler_Type.
I think it not effect python2.
Can you arch more infor?
thanks

msg336508 - (view) **Author: STINNER Victor (vstinner) *** **Date: 2019-02-25 10:12**

As I wrote in my previous comment, I don't think that Python 2.7 is affected by this issue.

msg336568 - (view) **Author: Larry Hastings (larry) *** **Date: 2019-02-25 22:44**

New changeset **4b42d575bf0fb01192b3ec54b7e224b238691527** by larryhastings (Victor Stinner) in branch '3.4':
[3.4] ~~**bpo-34656**~~: Avoid relying on signed overflow in _pickle memos (~~**GH-9261**~~) (~~**#11870**~~)
https://github.com/python/cpython/commit/4b42d575bf0fb01192b3ec54b7e224b238691527

msg336589 - (view) **Author: Larry Hastings (larry) *** **Date: 2019-02-26 00:41**

New changeset **ef33dd6036aafbd3f06c1d56e2b1a81dae3da63c** by larryhastings (Victor Stinner) in branch '3.5':
closes ~~**bpo-34656**~~: Avoid relying on signed overflow in _pickle memos. (~~**GH-9261**~~) (~~**#11869**~~)
https://github.com/python/cpython/commit/ef33dd6036aafbd3f06c1d56e2b1a81dae3da63c

History

Date	User	Action	Args
2022-04-11 14:59:05	admin	set	github: 78837
2019-05-10 18:14:59	ned.deily	set	messages: - **msg342084**
2019-05-10 17:36:37	ned.deily	set	nosy: + **ned.deily** messages: + **msg342084**
2019-02-26 00:41:36	larry	set	status: open -> closed resolution: fixed messages: + **msg336589** stage: patch review -> resolved
2019-02-25 22:44:16	larry	set	nosy: + **larry** messages: + **msg336568**
2019-02-25 10:12:46	vstinner	set	messages: + **msg336508** components: + Library (Lib), - ctypes versions: + Python 3.4, Python 3.5, Python 3.6, Python 3.7, Python 3.8, - Python 2.7
2019-02-23 14:14:10	dfmz77669	set	versions: - Python 3.4, Python 3.5, Python 3.6, Python 3.7, Python 3.8 nosy: + **dfmz77669** messages: + **msg336381** components: + ctypes, - FreeBSD
2019-02-23 13:42:06	dfmz77669	set	nosy: - **alexandre.vassalotti**, **benjamin.peterson**, **mcepl**, **serhiy.storchaka**, **hroncok**, **miss-islington**, **xtreak**, **shuoz**
2019-02-15 11:46:03	vstinner	set	pull_requests: + **pull_request11903**
2019-02-15 11:44:19	vstinner	set	stage: resolved -> patch review pull_requests: + **pull_request11902**
2019-01-23 17:24:30	mcepl	set	files: + **CVE-2018-20406-pickle_LONG_BINPUT.patch** messages: + **msg334267**
2019-01-22 12:34:09	vstinner	set	messages: + **msg334209**
2019-01-22 12:24:56	vstinner	set	messages: + **msg334208** versions: + Python 2.7
2019-01-21 21:47:40	mcepl	set	messages: + **msg334183**
2019-01-21 20:58:29	vstinner	set	status: closed -> open versions: + Python 3.4, Python 3.5, Python 3.6, Python 3.7 nosy: + **vstinner** messages: + **msg334179** resolution: fixed -> (no value)

2019-01-21 20:41:36	vstinner	set	title: memory exhaustion in Modules/_pickle.c:1393 -> [CVE-2018-20406] memory exhaustion in Modules/_pickle.c:1393
2019-01-20 23:18:13	mcepl	set	nosy: + **mcepl** messages: + **msg334108**
2019-01-09 09:33:25	serhiy.storchaka	set	messages: + **msg333294**
2019-01-09 08:08:59	hroncok	set	nosy: + **hroncok** messages: + **msg333292**
2018-09-21 02:00:41	miss-islington	set	messages: + **msg325939**
2018-09-21 01:52:40	miss-islington	set	nosy: + **miss-islington** messages: + **msg325938**
2018-09-21 01:37:06	miss-islington	set	pull_requests: + **pull_request8879**
2018-09-21 01:36:57	miss-islington	set	pull_requests: + **pull_request8878**
2018-09-21 01:36:50	benjamin.peterson	set	status: open -> closed nosy: + **benjamin.peterson** messages: + **msg325937** resolution: fixed stage: patch review -> resolved
2018-09-15 11:02:11	serhiy.storchaka	set	messages: + **msg325430**
2018-09-14 04:04:40	koobs	set	nosy: - **koobs**
2018-09-14 00:55:24	benjamin.peterson	set	keywords: + **patch** stage: patch review pull_requests: + **pull_request8718**
2018-09-13 09:47:19	xtreak	set	nosy: + **xtreak**
2018-09-13 05:46:03	serhiy.storchaka	set	nosy: + **alexandre.vassalotti**, **serhiy.storchaka**
2018-09-13 04:55:08	shuoz	set	messages: + **msg325231**
2018-09-13 04:53:32	shuoz	set	files: + **pk.py**
2018-09-13 04:38:47	shuoz	create	