New issue

Jump to bottom

Bug in bin2llvmir Decoder #637

Closed **seviezhou** opened this issue on Sep 4, 2019 · 2 comments

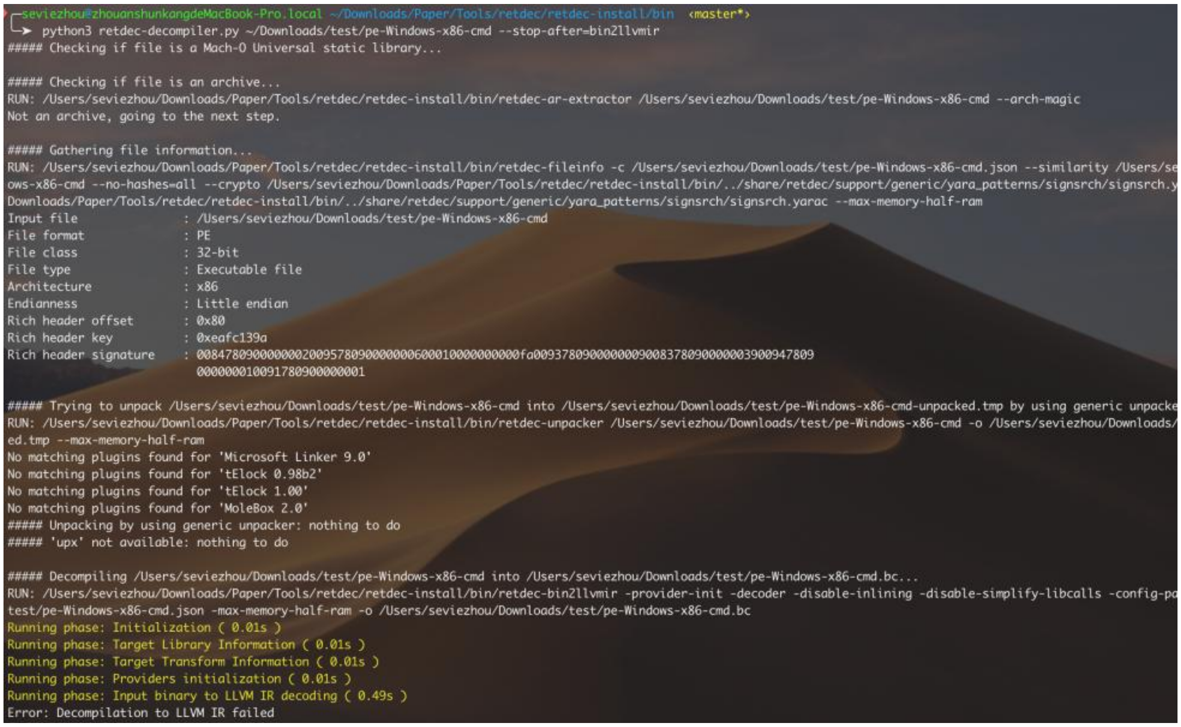
Labels bug C-bin2llvmir

seviezhou commented on Sep 4, 2019 Contributor

I try to translate the following PE file:

pe-Windows-x86-cmd.zip

But in the decoder phase, the retdec just gets an error and exits:



```
┌─seviezhou@zhouanshunkangdeMacBook-Pro.local ~/Downloads/Paper/Tools/retdec/retdec-install/bin  ‹master*›
└─➤ python3 retdec-decompiler.py ~/Downloads/test/pe-Windows-x86-cmd --stop-after=bin2llvmir
##### Checking if file is a Mach-O Universal static library...

##### Checking if file is an archive...
RUN: /Users/seviezhou/Downloads/Paper/Tools/retdec/retdec-install/bin/retdec-ar-extractor /Users/seviezhou/Downloads/test/pe-Windows-x86-cmd --arch-magic
Not an archive, going to the next step.

##### Gathering file information...
RUN: /Users/seviezhou/Downloads/Paper/Tools/retdec/retdec-install/bin/retdec-fileinfo -c /Users/seviezhou/Downloads/test/pe-Windows-x86-cmd.json --similarity /Users/se
ows-x86-cmd --no-hashes=all --crypto /Users/seviezhou/Downloads/Paper/Tools/retdec/retdec-install/bin/../share/retdec/support/generic/yara_patterns/signsrch/signsrch.y
Downloads/Paper/Tools/retdec/retdec-install/bin/../share/retdec/support/generic/yara_patterns/signsrch/signsrch.yarac --max-memory-half-ram
Input file                : /Users/seviezhou/Downloads/test/pe-Windows-x86-cmd
File format               : PE
File class                : 32-bit
File type                 : Executable file
Architecture              : x86
Endianness                : Little endian
Rich header offset        : 0x80
Rich header key           : 0xeafc139a
Rich header signature     : 0084780900000002009578090000000600010000000000fa00937809000000090083780900000039009478 09
                            000000010091780900000001

##### Trying to unpack /Users/seviezhou/Downloads/test/pe-Windows-x86-cmd into /Users/seviezhou/Downloads/test/pe-Windows-x86-cmd-unpacked.tmp by using generic unpacke
RUN: /Users/seviezhou/Downloads/Paper/Tools/retdec/retdec-install/bin/retdec-unpacker /Users/seviezhou/Downloads/test/pe-Windows-x86-cmd -o /Users/seviezhou/Downloads/
ed.tmp --max-memory-half-ram
No matching plugins found for 'Microsoft Linker 9.0'
No matching plugins found for 'tElock 0.98b2'
No matching plugins found for 'tElock 1.00'
No matching plugins found for 'MoleBox 2.0'
##### Unpacking by using generic unpacker: nothing to do
##### 'upx' not available: nothing to do

##### Decompiling /Users/seviezhou/Downloads/test/pe-Windows-x86-cmd into /Users/seviezhou/Downloads/test/pe-Windows-x86-cmd.bc...
RUN: /Users/seviezhou/Downloads/Paper/Tools/retdec/retdec-install/bin/retdec-bin2llvmir -provider-init -decoder -disable-inlining -disable-simplify-libcalls -config-pa
test/pe-Windows-x86-cmd.json -max-memory-half-ram -o /Users/seviezhou/Downloads/test/pe-Windows-x86-cmd.bc
Running phase: Initialization ( 0.01s )
Running phase: Target Library Information ( 0.01s )
Running phase: Target Transform Information ( 0.01s )
Running phase: Providers initialization ( 0.01s )
Running phase: Input binary to LLVM IR decoding ( 0.49s )
Error: Decompilation to LLVM IR failed
```

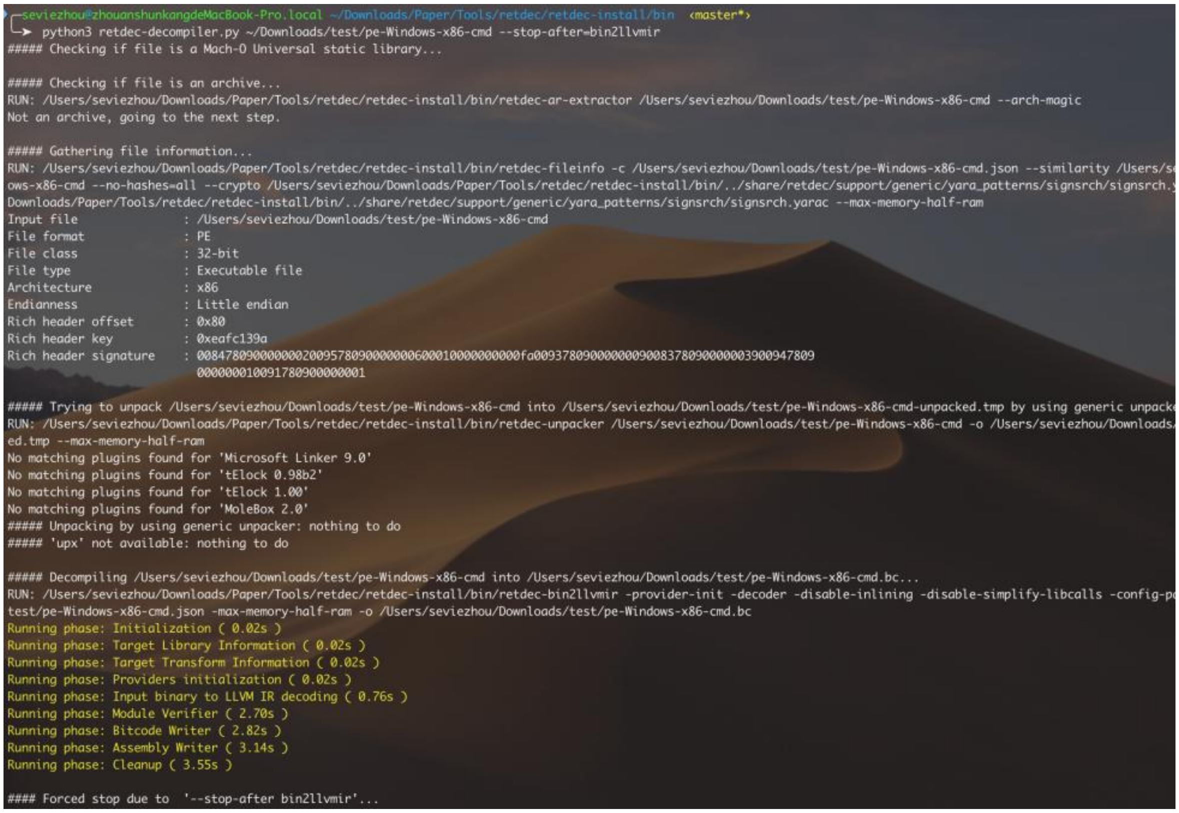
The problem is in file `src/bin2llvmir/optimizations/decoder/ir_modifications.cpp` , function `Decoder::canSplitFunctionOn` , line 426 and line 445, two portions of code:

```
...
                        auto up = fncStarts.upper_bound(bAddr);
                        --up;
                        Address bFnc = *up;

...
                                auto up = fncStarts.upper_bound(pAddr);
                                --up;
                                Address pFnc = *up;

...
```

The problem here is, if `up` equals to `fncStarts.begin()` , the `--up` will crash, the possible fix is:

```
            auto up = fncStarts.upper_bound(bAddr);
            if (up == fncStarts.begin()) {
                return false;
            }
            --up;
            uint64_t bFnc = *up;
```

After this fix, the Decoder works well:



```
┌─seviezhou@zhouanshunkangdeMacBook-Pro.local ~/Downloads/Paper/Tools/retdec/retdec-install/bin  <master*>
└─➤ python3 retdec-decompiler.py ~/Downloads/test/pe-Windows-x86-cmd --stop-after=bin2llvmir
##### Checking if file is a Mach-O Universal static library...

##### Checking if file is an archive...
RUN: /Users/seviezhou/Downloads/Paper/Tools/retdec/retdec-install/bin/retdec-ar-extractor /Users/seviezhou/Downloads/test/pe-Windows-x86-cmd --arch-magic
Not an archive, going to the next step.

##### Gathering file information...
RUN: /Users/seviezhou/Downloads/Paper/Tools/retdec/retdec-install/bin/retdec-fileinfo -c /Users/seviezhou/Downloads/test/pe-Windows-x86-cmd.json --similarity /Users/se
ows-x86-cmd --no-hashes=all --crypto /Users/seviezhou/Downloads/Paper/Tools/retdec/retdec-install/bin/../share/retdec/support/generic/yara_patterns/signsrch/signsrch.y
Downloads/Paper/Tools/retdec/retdec-install/bin/../share/retdec/support/generic/yara_patterns/signsrch/signsrch.yarac --max-memory-half-ram
Input file                : /Users/seviezhou/Downloads/test/pe-Windows-x86-cmd
File format               : PE
File class                : 32-bit
File type                 : Executable file
Architecture              : x86
Endianness                : Little endian
Rich header offset        : 0x80
Rich header key           : 0xeafc139a
Rich header signature     : 008478090000000200957809000000006000100000000000fa00937809000000009008378090000003900947809
                            00000001009178090000000 1

##### Trying to unpack /Users/seviezhou/Downloads/test/pe-Windows-x86-cmd into /Users/seviezhou/Downloads/test/pe-Windows-x86-cmd-unpacked.tmp by using generic unpacke
RUN: /Users/seviezhou/Downloads/Paper/Tools/retdec/retdec-install/bin/retdec-unpacker /Users/seviezhou/Downloads/test/pe-Windows-x86-cmd -o /Users/seviezhou/Downloads/
ed.tmp --max-memory-half-ram
No matching plugins found for 'Microsoft Linker 9.0'
No matching plugins found for 'tElock 0.98b2'
No matching plugins found for 'tElock 1.00'
No matching plugins found for 'MoleBox 2.0'
##### Unpacking by using generic unpacker: nothing to do
##### 'upx' not available: nothing to do

##### Decompiling /Users/seviezhou/Downloads/test/pe-Windows-x86-cmd into /Users/seviezhou/Downloads/test/pe-Windows-x86-cmd.bc...
RUN: /Users/seviezhou/Downloads/Paper/Tools/retdec/retdec-install/bin/retdec-bin2llvmir -provider-init -decoder -disable-inlining -disable-simplify-libcalls -config-pa
test/pe-Windows-x86-cmd.json -max-memory-half-ram -o /Users/seviezhou/Downloads/test/pe-Windows-x86-cmd.bc
Running phase: Initialization ( 0.02s )
Running phase: Target Library Information ( 0.02s )
Running phase: Target Transform Information ( 0.02s )
Running phase: Providers initialization ( 0.02s )
Running phase: Input binary to LLVM IR decoding ( 0.76s )
Running phase: Module Verifier ( 2.70s )
Running phase: Bitcode Writer ( 2.82s )
Running phase: Assembly Writer ( 3.14s )
Running phase: Cleanup ( 3.55s )

#### Forced stop due to  '--stop-after bin2llvmir'...
```

xkubov added C-bin2llvmir bug labels on Sep 9, 2019

xkubov commented on Sep 9, 2019 • edited Member

Hi, thank you for the report. Indeed it looks like an issue as the boundaries are not checked in those parts of the code. If you want, you can open a pull request to fix this issue.

seviezhou added a commit to seviezhou/retdec that referenced this issue on Sep 9, 2019

`Try to fix issue avast#637` … c517d07

seviezhou mentioned this issue on Sep 9, 2019

Try to fix issue #637 #641

Merged

PeterMatula pushed a commit that referenced this issue on Sep 10, 2019

`Try to fix issue #637` … 517298b

PeterMatula added a commit that referenced this issue on Sep 10, 2019

`CHANGELOG.md: add entries for #637 and #641.` a894749

PeterMatula commented on Sep 10, 2019 Collaborator

- Fixed by Try to fix issue #637 #641.
- Added `CHANGELOG.md` entry in `a894749`.
- Since this was not a decompilation quality bug, but a crash bug, I added the binary to the internal nightly tests suite that checks that RetDec does not crash.

PeterMatula closed this as completed on Sep 10, 2019

Assignees

No one assigned

Labels

bug C-bin2llvmir

Projects

None yet

Milestone

No milestone

Development

No branches or pull requests

3 participants