Upstream pcre developers are working on pcre2 and have been for a while. I would like to know the impact on Suricata of supporting this.

- can they coexist in a single program (we use pcre for match and parsing)
- what does distro support look like for 'tier 1' in Support Status
- are there features we use that are missing from pcre2?
- estimate of effort for just supporting it for matching
- estimate of effort for completely replacing it

Subtasks:
Task # 4446: pcre2: document changes vs prce1 for rule writers

History

#1 - 08/07/2020 02:01 PM - Victor Julien
- Status changed from New to Assigned
- Assignee changed from OISF Dev to Jeff Lucovsky
- Target version changed from 70 to 7.0rc1

#2 - 10/30/2020 03:42 PM - Philippe Antoine
Is the goal of this ticket to replace pcre by pcre2?
If so, it will be useful for fuzzing, as it finds bugs in pcre which will likely not be fixed because pcre is at end of life and all the development is now in pcre2
of https://bugs.exim.org/show_bug.cgi?id=2661

#3 - 10/31/2020 06:08 AM - Victor Julien
Ultimately yes. I would like us to get away from using pcre for rule/config parsing, so that we only use it for the "pcre" rule keyword.

#4 - 11/01/2020 07:57 PM - Philippe Antoine
Ok, so this is about rustifying all input handling, right?

From the fuzzing point of view, it is useful to replace pcre by pcre2 even as long as we keep it for rule/config parsing

#5 - 12/06/2020 09:55 AM - Philippe Antoine
For reference:
https://bugs.chromium.org/p/oss-fuzz/issues/detail?id=26226&q=label%3AProj-suricata
https://bugs.chromium.org/p/oss-fuzz/issues/detail?id=26498&q=label%3AProj-suricata

#6 - 02/09/2021 10:33 AM - Philippe Antoine

- can they coexist in a single program (we use pcre for match and parsing)

I do not see why not
I do not see the point though, but I neither see the difference between supporting it for matching and completely replacing it
Is it just using pcre2 in detect-pcre*.c?
what does distro support look like for 'tier 1' in Support Status

I guess it should be pretty good.
PCRE1 (the 8.xx series) is at end of life.
Should CI check this?

estimate of effort for just supporting it for matching
estimate of effort for completely replacing it

I do not see a difference between these two

Used methods seem
- pcre_free_substring replaced straight by pcre2_substring_free
- pcre_get_substring seems replaced by pcre2_substring_get_bynumber which does not have the same arguments...
- pcre_exec replaced by pcre2_match
- pcre_copy_substring replaced by pcre2_substring_copy_bynumber
- pcre_free
- pcre_compile replaced by pcre2_compile
- pcre_study obsolete in pcre2: The new API is more extensible, and it was simplified by abolishing the separate "study" optimizing function
- pare jit stuff present in pcre2

To investigate: the suricata.yaml pcre.match-limit and match-limit-recursion
Seems like it is supported with PCRE2_CONFIG_DEPTHLIMIT

Looks like a few days work to get something testable
And then the work needed to test it...

are there features we use that are missing from pcre2?

It looks complete, needs to be tested...
pcre_study is now obsolete but the features must be there

#7 - 02/09/2021 10:34 AM - Philippe Antoine
For reference: https://github.com/catenacyber/suricata/tree/pcre2-3194-v1

Incomplete work: needs pcre_get_substring and pcre_copy_substring to be translated...

#8 - 02/12/2021 05:29 PM - Victor Julien
CentOS7 has pcre2 support in base, version is 10.23-2.el7. We should make sure that works with our code.

#9 - 02/12/2021 05:29 PM - Victor Julien
- Subject changed from investigate pcre2 support to pcre2 support
- Status changed from Assigned to In Progress
- Assignee changed from Jeff Lucovsky to Philippe Antoine

#10 - 03/16/2021 07:49 PM - Philippe Antoine
Draft https://github.com/OISF/suricata/pull/5981

#11 - 04/13/2021 12:40 PM - Philippe Antoine
- Status changed from In Progress to In Review

https://github.com/OISF/suricata/pull/6054

#12 - 09/29/2021 11:30 AM - Philippe Antoine
- Status changed from In Review to Closed

https://github.com/OISF/suricata/pull/6414

#13 - 11/18/2021 09:23 PM - Sascha Steinbiss
FYI https://bugs.debian.org/cgi-bin/bugreport.cgi?bug=999952 <- Moving to PCRE2 over the deprecated old PCRE has also been suggested by Debian
Thanks Sasha for answering there.

Is a backport of this work desired? It's somewhat intrusive, but not impossible of course.

Well, if you think it's going to be a while until Suri 7 is released, then a backport would be appreciated in the meantime. If you think it will be released in the coming few months, then I don't think so. Obviously, the sooner we can switch the dependency, the better. But AFAICS the goal is to retire the old PCRE library in Debian before the Bookworm freeze, which is not even planned yet.