Commits ReMon

hash	author	date	message
00a3cee	lennertfranssens	Tue, 31 May 2022 19:01:26 +0200	Update list of allowed syscalls for nginx benchmark
d115647	lennertfranssens	Tue, 31 May 2022 15:56:13 +0200	Remove useless comment
e5677a6	lennertfranssens	Tue, 31 May 2022 15:54:45 +0200	Add old IP-MON process to script
57275d4	lennertfranssens	Tue, 31 May 2022 13:59:25 +0200	Refactor seccomp bpf benchmark artifact folder and contents …
e26c302	lennertfranssens	Tue, 31 May 2022 11:13:19 +0200	MVEE_RUNS_UNDER_MVEE_CONTROL is working again
dc11a34	lennertfranssens	Tue, 31 May 2022 10:59:11 +0200	Make nginx work for now with 1 process and no shared memory
d026cad	lennertfranssens	Mon, 30 May 2022 17:25:34 +0200	Only use non-zero bits in secret
5413b8d	lennertfranssens	Mon, 30 May 2022 15:54:57 +0200	Implement fake syscall to let the monitor know if the filter …
e38a40f	lennertfranssens	Mon, 30 May 2022 12:03:39 +0200	Random mapping for IP-MON per variant
c988b11	lennertfranssens	Mon, 30 May 2022 11:37:39 +0200	Invoke key exchange key is now random
0e8b9f3	lennertfranssens	Sun, 29 May 2022 20:38:49 +0200	Add full 64-bit key support in seccomp-BPF IP-MON …
08ce7f6	lennertfranssens	Fri, 27 May 2022 00:02:33 +0200	Add nginx to seccomp bpf benchmarks artifact bootstrap
7565005	lennertfranssens	Thu, 26 May 2022 23:35:21 +0200	Remove auto-generated file
c53e3df	lennertfranssens	Thu, 26 May 2022 23:16:41 +0200	Remove port mapping for docker script
28426d7	lennertfranssens	Thu, 26 May 2022 23:12:15 +0200	Map IP-MON on same address, even when on cloned variant
d4990f3	lennertfranssens	Wed, 25 May 2022 09:04:40 +0200	Copy IP-MON state of parent monitor in new monitors after …
f79b8a6	lennertfranssens	Wed, 4 May 2022 21:54:20 +0200	Update getpid benchmark program
2e2d471	lennertfranssens	Wed, 4 May 2022 16:35:27 +0200	Use make benchmark in bootstrap
13f5980	lennertfranssens	Wed, 4 May 2022 15:28:22 +0200	Update for benchmark tests
f4a7409	lennertfranssens	Wed, 4 May 2022 14:03:19 +0200	Update bootstrap
c50a7cf	lennertfranssens	Wed, 4 May 2022 14:01:55 +0200	Refactor microbenchmark
6e685c3	lennertfranssens	Wed, 4 May 2022 13:59:19 +0200	Register ipmon thread if RB is 0
9b3b476	lennertfranssens	Tue, 3 May 2022 22:30:41 +0200	Push with changes
cccdbce	lennertfranssens	Tue, 3 May 2022 21:56:07 +0200	Update benchmarks scripts
7723b3e	lennertfranssens	Tue, 3 May 2022 19:38:39 +0200	Do not apply patches to IP-MON
d409179	lennertfranssens	Tue, 3 May 2022 19:35:15 +0200	Reuse original parts of the bootstrap script
74a6cf2	lennertfranssens	Tue, 3 May 2022 19:23:11 +0200	Remove unused symbol
5fabfdb	lennertfranssens	Tue, 3 May 2022 19:13:18 +0200	Update folder names
ebeccca	lennertfranssens	Tue, 3 May 2022 19:00:33 +0200	Update scripts to benchmark ipmon
5332d4c	lennertfranssens	Tue, 3 May 2022 18:01:09 +0200	Copy eurosys2022-artifact folder
e898292	lennertfranssens	Mon, 2 May 2022 23:45:31 +0200	Use regex to find out if file is libipmon.so or has libipmon in its name
e6f0a06	lennertfranssens	Mon, 2 May 2022 17:18:14 +0200	Do not print out region span

c17d1e4	Lennert Franssens	Mon, 2 May 2022 16:56:55 +0200	Add dependencies for new IP-MON
3a49013	lennertfranssens	Fri, 29 Apr 2022 10:10:36 +0200	Update syscalls in bpf filter and add old whitelist to readme file
22b9e0a	lennertfranssens	Fri, 29 Apr 2022 09:52:21 +0200	Upload seccomp-bpf policy with original ipmon policy
78068a1	lennertfranssens	Thu, 28 Apr 2022 16:26:49 +0200	Remove unneeded files
30e21c3	lennertfranssens	Thu, 28 Apr 2022 15:29:07 +0200	Add instructions on how to generate the seccomp-bpf filter to …
ab20694	lennertfranssens	Thu, 28 Apr 2022 15:15:49 +0200	Update script to generate seccomp-bpf filter from json file
9bf94ff	lennertfranssens	Thu, 28 Apr 2022 15:14:59 +0200	Install perl and modules for perl to compile IP-MON
d392ead	lennertfranssens	Thu, 28 Apr 2022 12:13:21 +0200	Clean code
1a8a1bb	lennertfranssens	Thu, 28 Apr 2022 12:12:50 +0200	Start making a generated bpf filter
eca5df8	lennertfranssens	Thu, 28 Apr 2022 10:59:55 +0200	Do not use seccomp filters from docker
068a224	lennertfranssens	Thu, 28 Apr 2022 08:56:45 +0200	Switch back to PTRACE_CONT when using ipmon with seccomp-bpf
719aa9a	lennertfranssens	Wed, 27 Apr 2022 10:03:53 +0200	ipmon is almost working now
575114e	lennertfranssens	Wed, 27 Apr 2022 09:13:19 +0200	Add some debugging code in the monitor
04b2666	lennertfranssens	Wed, 27 Apr 2022 09:12:23 +0200	Unfold fake syscall
25a1db1	lennertfranssens	Tue, 26 Apr 2022 09:00:52 +0200	Map ipmon on same addresses after an execve
fe97ba6	lennertfranssens	Tue, 26 Apr 2022 09:00:03 +0200	Unfold checked syscall and do not return errno for any syscalls for now
bbdfef5	lennertfranssens	Mon, 25 Apr 2022 09:04:59 +0200	Set ipmon region in monitor for current variant only
176b27d	lennertfranssens	Mon, 25 Apr 2022 08:23:37 +0200	Try to fix segfault when using ipmon
9d0c160	lennertfranssens	Sun, 24 Apr 2022 19:50:55 +0200	Update some logic to let ipmon almost work
90115ac	lennertfranssens	Sun, 24 Apr 2022 19:50:14 +0200	Add ipmon executed unchecked syscalls to the filter to allow them …
ff3671a	lennertfranssens	Sun, 24 Apr 2022 11:03:44 +0200	RDTSC is working again, but MVEE_RUNS_UNDER_MVEE_CONTROL …
b380da1	lennertfranssens	Sun, 24 Apr 2022 11:04:21 +0200	Revert Add some logging and disable ipmon for now""
2f70efe	lennertfranssens	Sun, 24 Apr 2022 10:22:46 +0200	Add some logging and disable ipmon for now
5cc4ca0	Lennert Franssens	Fri, 22 Apr 2022 08:55:58 +0200	Update README.md
96287bf	Lennert Franssens	Fri, 22 Apr 2022 08:47:04 +0200	Update README.md
2a6a453	lennertfranssens	Thu, 21 Apr 2022 17:15:21 +0200	Remove part of MVEE_REGISTER_IPMON that is not used with bpf ...
c6e0de1	lennertfranssens	Thu, 21 Apr 2022 17:13:28 +0200	Update enclave functions to let them work with seccomp bpf filters
c9c4cda	lennertfranssens	Wed, 20 Apr 2022 13:25:21 +0200	Map ipmon on fixed addres (for now)
00a0f50	lennertfranssens	Wed, 20 Apr 2022 13:24:12 +0200	Update bpf filter and add labels to know some addresses
93debee	lennertfranssens	Mon, 18 Apr 2022 22:45:20 +0200	Copy ipmon registration to fake syscall handler…
a0ae225	lennertfranssens	Mon, 18 Apr 2022 22:43:54 +0200	Update bpf filter and add label to instruction
613391f	lennertfranssens	Sun, 17 Apr 2022 02:00:53 +0200	Let ipmon work with allow bpf filter
5cd456d	lennertfranssens	Sun, 17 Apr 2022 02:00:11 +0200	Let ipmon work with allow bpf filter
eaf1fa2	lennertfranssens	Sat, 16 Apr 2022 22:15:50 +0200	Install clang to docker container for ipmon installation

fd987e3	lennertfranssens	Sat, 16 Apr 2022 22:14:51 +0200	Handle ipmon mmap on a specific way
abf401a	lennertfranssens	Sat, 16 Apr 2022 22:13:53 +0200	Don't check kernel patch for ipmon and fix some errors
ae56c0c	lennertfranssens	Sat, 16 Apr 2022 20:19:35 +0200	Fix error with SHM TAG
81ba875	lennertfranssens	Wed, 13 Apr 2022 17:36:45 +0200	Prepare for new ipmon
581bbf8	lennertfranssens	Wed, 13 Apr 2022 17:35:38 +0200	Prepare for new ipmon
fbf0605	lennertfranssens	Wed, 23 Mar 2022 18:41:49 +0100	WIP let monitor know the address of syscall instruction of glibc
89834ea	lennertfranssens	Tue, 22 Mar 2022 17:42:29 +0100	Update bpf filter to work with custom_syscall_function
065e35c	lennertfranssens	Tue, 15 Mar 2022 16:19:36 +0100	Let BPF filters work
91de1cc	lennertfranssens	Wed, 9 Mar 2022 11:45:27 +0100	Add some comments, debug info and let the seccomp syscall work
b795e48	lennertfranssens	Thu, 3 Mar 2022 09:27:41 +0100	Start implementing seccomp-BPF filter
fe070ab	lennertfranssens	Wed, 16 Feb 2022 12:03:43 +0100	Add instructions on how to build the container
158e07a	lennertfranssens	Tue, 15 Feb 2022 23:30:47 +0100	Add script to create build directory
ba23256	lennertfranssens	Tue, 15 Feb 2022 23:21:03 +0100	Enable network access in the container
0fae601	lennertfranssens	Tue, 15 Feb 2022 23:20:48 +0100	Install python3 in the container

Commits ReMon-glibc

hash	author	date	message
b73adc634	lennertfran	Mon, 30 M	Only use non-zero bits in secret
6b4f1f767f	lennertfran	Sun, 29 Ma	Add full 64-bit key support in custom syscall function
e68ea7720	lennertfran	Thu, 26 Ma	Address of IP-MON lies in between [2048,4096[+ 12 '0' bits
3bc90cea1!	lennertfran	Mon, 2 Ma	Fix strict-prototypes error
ec0d0fbabe	lennertfran	Wed, 20 Ap	Implement custom syscall to work with new ipmon
1742cd478	lennertfran	Wed, 13 Ap	Prepare for new ipmon
100c9dcf0(	lennertfran	Wed, 23 M	WIP fake syscall to notify the address of syscall instruction of glibc
455c2055e	lennertfran	Tue, 22 Ma	Update custom_syscall_function
6f9ae8be0:	lennertfran	Wed, 23 Fe	Add some comments
b897b055d	lennertfran	Wed, 23 Fe	Add function (to) call when intercepting a syscall
b0f942b2el	lennertfran	Wed, 23 Fe	Multi-line macro for syscall instruction replacement
6ea88c8de	lennertfran	Tue, 22 Feb	Disable werror in configure scripts
b0236dcf6:	lennertfran	Tue, 22 Feb	Make syscall a custom function in x86_64
bbcb4bc1e	lennertfran	Tue, 22 Feb	Ignore R_X86_64_NONE on lazy relocation

Messages with prof. dr. ir. Bjorn De Sutter, prof. dr. Bart Coppens and dr. ir. Bert Abrath

---

bcoppens
4:33 PM
Ziehier het paper over een In-Process Monitor voor MVEEs
PDF

atc16-paper-volckaert.pdf
PDF

4:34
De code daarvan vind je hier: https://github.com/stijn-volckaert/ReMon
GitHubGitHub
GitHub - stijn-volckaert/ReMon
Contribute to stijn-volckaert/ReMon development by creating an account on GitHub.
(115 kB)
https://github.com/stijn-volckaert/ReMon

4:36
Die eBPF interfaces en seccomp interfaces kan je alvast hier wat meer info over vinden:
Filter and Modify System Calls with seccomp and ptrace
https://www.alfonsobeato.net/c/filter-and-modify-system-calls-with-seccomp-and-ptrace/
Syscall User Dispatch
https://www.kernel.org/doc/html/latest/admin-guide/syscall-user-dispatch.html
Seccomp user-space notification and signals https://lwn.net/Articles/851813/ (toen ik er laatst naar keek was dat nog niet in mainline, ik weet niet of dat intussen al aangepast is)
eBPF wordt dus ook gebruikt voor performance monitoring enzo, en daarvoor kan je bv kijken naar deze voorbeeldjes: http://www.brendangregg.com/perf.html
lwn.netlwn.net
Seccomp user-space notification and signals
The seccomp() mechanism allows the imposition of a filter program (expressed in "classic" BPF) that makes policy decisions on whether to allow each system call invoked by the target process. The user-space notification feature further allows those decisions to be deferred to another process. As this recent patch set from Sargun Dhillon shows, though, user-space notification still has some rough edges, especially when it comes to signals. This patch makes a simple change to try to address a rather complex problem brought to the fore by changes in the Go language's preemption model.
brendangregg.combrendangregg.com
Linux perf Examples
Examples of using the Linux perf command, aka perf_events, for performance analysis and debugging. perf is a profiler and tracer.

Lennert Franssens
  10:47 AM
Ik zou tegen 22 oktober een werkplan moeten indienen (http://masterproef.tiwi.ugent.be/Werkplan/). Moet ik dat zelf maken of kunnen jullie me daarbij helpen?

5 replies
Last reply 9 months agoView thread

Lennert Franssens
  11:28 AM
Ik heb ook de ReMon repo gedownload. Ik heb de stappen uit de README uitgevoerd (na een kleine aanpassing https://github.ugent.be/lefranss/ReMon/commit/14e5f920b806f04dd84bc217eda8f30ea80cc45b) maar blijf hangen bij het maken van IP-MON. In het bestand ReMon/MVEE/Inc/MVEE_fake_syscall.h staat MVEE_FAKE_SYSCALL_BASE niet gedefinieerd. Welke waarde moet die hebben? Of moet die waarde van ergens anders komen? (edited)

3 replies
Last reply 9 months agoView thread

babrath
  6:00 PM
replied to a thread:
Ik zou tegen 22 oktober een werkplan moeten indienen (http://masterproef.tiwi.ugent.be/Werkplan/). Moet ik dat zelf maken of kunnen jullie me daarbij helpen?
We kunnen je daar bij helpen, misschien dinsdag ergens een meeting?
View newer replies

babrath
  6:01 PM
replied to a thread:
Ik heb ook de ReMon repo gedownload. Ik heb de stappen uit de README uitgevoerd (na een kleine aanpassing https://github.ugent.be/lefranss/ReMon/commit/14e5f920b806f04dd84bc217eda8f30ea80cc45b) maar blijf hangen bij het maken van IP-MON. In het bestand ReMon/MVEE/Inc/MVEE_fake_syscall.h staat MVEE_FAKE_SYSCALL_BASE niet gedefinieerd. Welke waarde moet die hebben? Of moet die waarde van ergens anders komen?
Hmm ik weet het niet direct, maar dat is zoiezo een verouderde repository. Er is een nieuwere versie beschikbaar, maar nog niet publiek, denk ik?
View newer replies

babrath
  7:16 PM
replied to a thread:
Ik heb ook de ReMon repo gedownload. Ik heb de stappen uit de README uitgevoerd (na een kleine aanpassing https://github.ugent.be/lefranss/ReMon/commit/14e5f920b806f04dd84bc217eda8f30ea80cc45b) maar blijf hangen bij het maken van IP-MON. In het bestand ReMon/MVEE/Inc/MVEE_fake_syscall.h staat MVEE_FAKE_SYSCALL_BASE niet gedefinieerd. Welke waarde moet die hebben? Of moet die waarde van ergens anders komen?
Ja, we zullen in tussentijd kijken voor toegang tot de nieuwere repo

babrath
  7:17 PM
replied to a thread:
Ik zou tegen 22 oktober een werkplan moeten indienen (http://masterproef.tiwi.ugent.be/Werkplan/). Moet ik dat zelf maken of kunnen jullie me daarbij helpen?
Best online, ik ben heel dinsdagnamiddag vrij.
@bcoppens
?
View newer replies

bcoppens
  8:12 PM
replied to a thread:
Ik zou tegen 22 oktober een werkplan moeten indienen (http://masterproef.tiwi.ugent.be/Werkplan/). Moet ik dat zelf maken of kunnen jullie me daarbij helpen?
Dinsdag in 14u is goed!
View newer replies

Lennert Franssens
  8:16 PM
Dit is mijn werkplan gebaseerd op de milestones die we samen overlopen hebben :slightly_smiling_face:
PDF

lennert_franssens-werkplan-masterproef_UGent.pdf
PDF

8:17
Ik ben nog niet zo heel zeker over de 'Bestaande situatie en probleemstelling' en 'Doelstelling van het project'

bcoppens

1:35 PM
Ik besef net dat we het eigenlijk zonet even konden gehad hebben over je vraag over doelstelling :sweat_smile:
:smile:
1

Lennert Franssens
2:06 PM
Ja, ik ben gewoon niet zeker of dat juist geformuleerd is op die manier...

bcoppens
4:32 PM
Enkele opmerkingen :slightly_smiling_face:
Bij de probleemstelling: ik denk dat alles er wel in staat wat er in moet staan, maar dat de structuur niet zo ideaal is. Ik zou iets hebben in de aard van (maar dan verder uitgeschreven he ;)) : MVEEs veilige manier [etc]. MVEEs werken zo. De traditionele manier is met syscalls onderscheppen met de debug-API in een ander proces. Dat zorgt voor overhead. ReMon biedt de optie om een aantal 'ongevaarlijke' system calls niet meer via die trage omweg te laten gebeuren, maar door een aparte component te injecteren in de adresruimte van die applicatie, die die system calls veel sneller kan afhandelen. Dit vereist echter een kernel patch, en de meeste mensen zijn geen fan van custom kernel pathces :cry:
Doelstelling: 'Een nieuw ontwerp van de bestaande applicatie [...]' -> In deze thesis gaan we ReMon aanpassen, zodanig dat de functionaliteit van de kernelpatch (en de interactie die de monitor [uitleggen dus al in de probleemstelling wat een monitor is in 1 zin] daarmee heeft) wordt vervangen door functionaliteit die gebruik maakt van die nieuwe technologieen
'Woord vooraf schrijven' zou ik echt pas helemaal op het einde doen, want dat is best iets dat je pas schrijft als je zeker weet waarover je thesis gaat / wat je het afgelopen jaar allemaal hebt meegemaakt (afhankelijk van hoe je dat invult ;))
Versnelde innovatie door middel van eBPF -> geen idee wat je daarmee bedoelt :slightly_smiling_face:
Literatuurstudie beschrijven -> ik zou hier expliciet meerdere weken voor uittrekken (maar dan niet apart 'bronvermelding van literatuurstudie' hebben)
Ik zou op 20 feb dan expliciet dat eigen ontwerp ook als 'deadline' opgeven?
Die 'Waarover vertel ik' bij de 2e tussentijdse presentatie zou ik weglaten. Ik zou gewoon vermelden dat je er aan begint te werken oid
globalgoals is 2 woorden denkik?
Je hebt nu normaalgezien ook access tot deze code: https://github.ugent.be/bcoppens/ReMon/tree/shm-dev :smile: (da's direct de juiste branch)

Lennert Franssens
5:23 PM
Oké, ik wijzig het werkplan dan aan de hand van de opmerkingen :slightly_smiling_face:

En klopt, ik kan de repo zien dus dat is ook in orde :smiley:
:tada:
1

Lennert Franssens
  8:07 AM
Ik ben even aan het proberen om het docker_MVEE.sh script uit te voeren in download mode. Het script blijf hangen op git submodule update --init --recursive omdat ik geen toegang heb tot de llvm repo (https://github.com/csl-ugent/ReMon-llvm-project.git) in de deps folder. Misschien is de beste oplossing om die repo ook op het account van @bcoppens te zetten? Dan kan ik de submodule url van llvm vervangen naar die van de gekloonde repo.

babrath
  8:30 AM
Ahja
8:30
Uiteindelijk nog niet belangrijk, denk ik
8:30
Dus die kan je skippen

Lennert Franssens
  8:31 AM
Oké, dan haal ik die er even uit :slightly_smiling_face:

babrath
  8:32 AM
Die llvm kun je gebruiken om programma's te compileren op een manier die beter geschikt is voor Remon
8:32
Maar dat kan ook zonder :-)
:+1:
1

Lennert Franssens
  8:33 AM
En staat er ergens een beschrijving van alle componenten van ReMon?

babrath
  9:02 AM

Hmm, in de source code, wil je zeggen? Niet direct

Lennert Franssens
6:05 PM
Ja dat bedoelde ik. Het zal zichzelf wel uitwijzen als ik de code eens goed bekijk :slightly_smiling_face:

brdsutte
11:56 AM
Ik heb nieuws van de voorzitter van de OC: onze studenten ind. inf. moeten het werkplan pas indienen op het einde van week 6.
:+1:
2

Lennert Franssens
3:04 PM
Heeft er iemand tijd om me ReMon te helpen installeren? Ik heb al enkele pogingen ondernomen maar ik blijf steeds op verschillende plaatsen vastzitten...

11 replies
Last reply 9 months agoView thread

Lennert Franssens
4:54 PM
Dit is het aangepaste werkplan. Dan hebben jullie het al voor maandag :slightly_smiling_face:
2 files

lennert_franssens-werkplan-masterproef_UGent.docx
Word Document

lennert_franssens-werkplan-masterproef_UGent.pdf
PDF

1 reply
9 months agoView thread

babrath
6:25 PM

replied to a thread:
Heeft er iemand tijd om me ReMon te helpen installeren? Ik heb al enkele pogingen ondernomen maar ik blijf steeds op verschillende plaatsen vastzitten...
Zeker, ik heb desnoods morgen wel tijd om te bellen, en anders maandag fysiek? Waar zit je nu vast?
View newer replies

babrath
6:36 PM
replied to a thread:
Dit is het aangepaste werkplan. Dan hebben jullie het al voor maandag
:slightly_smiling_face:
Dat ziet er mij in het algemeen al redelijk goed qua detail en beschrijving uit. Ik vraag me wel af wat 'IP-MON uit het systeem halen' (28 feb) wil zeggen?

Lennert Franssens
6:44 PM
Ja dat was om het wat verschillend te maken maar ik ga dat veranderen naar dezelfde omschrijving als de week erna.

babrath
7:12 PM
OK, 2 files: Een patch die je eerst moet applyen op je git repo (via git am) en dan het script dat alle dependencies installeert, de build directory opzet (met nieuw aangemaakte makefiles daarin), en vervolgens de MVEE (de CP-MON in ReMon) build.
2 files

0001-Small-patch-to-show-that-all-deps-are-installed.patch
Diff

install.sh
Shell

:heart:
1

7:12
Eens je dit script hebt uitgevoerd, moet je enkel nog maar 'make' doen om de MVEE dan opnieuw te builden na enige aanpassingen.

Lennert Franssens
7:14 PM
Heel erg bedankt!

babrath
  7:15 PM
Geen probleem :slightly_smiling_face:

bcoppens
  4:58 PM
```
sudo apt-get update
sudo apt-get install linux-source-5.3.0
tar jxf /usr/src/linux-source-5.3.0/linux-source-5.3.0.tar.bz2
cd linux-source-5.3.0
patch -p1 < path_to_patches/linux-5.3.0-full-cerberus.patch
make menuconfig
scripts/config --disable DEBUG_INFO
# scripts/config --disable CONFIG_SYSTEM_TRUSTED_KEYS
make -j `getconf _NPROCESSORS_ONLN` deb-pkg LOCALVERSION=-cerberus
sudo dpkg -i ../linux-headers*.deb ../linux-image*.deb ../linux-libc-dev*.deb
```
4:59
Dat is om een aparte package te maken van de gepatchte kernel :stuck_out_tongue:
4:59
de CONFIG_SYSTEM_TRUSTED_KEYS is voor wanneer hij zaagt over iets van keys, security feature van Linux, maar maakt geen hol uit voor ons

Lennert Franssens
  4:57 PM
Dat is gelukt :slightly_smiling_face:
4:59
Is het normaal dat hij op dit punt blijft vastzitten?
Screenshot from 2021-10-14 09-08-43.png
Screenshot from 2021-10-14 09-08-43.png

5:00
De bovenstaande afbeelding is zonder patch op mijn host-os. De onderstaande screenshot is in de virtuele machine. Daar heb ik net dezelfde installatie gedaan. Daar krijg ik dan volgende fout (met en zonder kernelpatch).
5:02
Screenshot from 2021-10-14 17-02-31.png
Screenshot from 2021-10-14 17-02-31.png

5:04
In het logbestand staat daar ook dat fd-linux.so niet gevonden kan worden in /patched_binaries/libc/amd/.

bcoppens

5:07 PM
Dat eerste is niet normaal :disappointed:
5:08
Voor dat tweede, kan je eens proberen met VMWare?
5:08
Ikzelf heb dat probleem nog niet gehad, maar ik heb dat wel al bij anderen gezien
5:08
en ik dacht dat dat ofwel vanzelf werkte bij VMWare Player, of dat daar een optie voor was om dat te disablen oid
5:08
van die cpuid

Lennert Franssens
5:13 PM
Oké, dan ga ik sowieso eerst eens in VMWare proberen om al tot dezelfde fout te komen als op mijn host-os.
Er staan wel allemaal andere bestanden in die patched_binaries map, dus ik vermoed dat het ergens verdwenen is in de loop der tijd?
5:14
Ik ben nu net op de repo van meneer Volckaert gaan kijken en daar staat het wel nog (als symbolische link naar een ander bestand in die map)
https://github.com/stijn-volckaert/ReMon/tree/master/patched_binaries/libc/amd64
(edited)

GitHubGitHub
ReMon/patched_binaries/libc/amd64 at master · stijn-volckaert/ReMon
Contribute to stijn-volckaert/ReMon development by creating an account on GitHub.
(115 kB)
https://github.com/stijn-volckaert/ReMon/tree/master/patched_binaries/libc/amd64

5:16
Ik veronderstel dat ik zelf een symbolische link mag maken dan naar ld-2.31.so (of ld-2.27.9000.so)?

bcoppens
5:17 PM
Ah, hmmmmm, wacht, gebruik je dat docker script?
5:17
want ik denk dat die die symlinks goed zou moeten zetten iirc
5:18
'./scripts/switch_patched_binaries.sh ubuntu20' die lijn in de docker_MVEE.sh zou dat nornaal moeten doen
:heart:
1

Lennert Franssens

5:18 PM
Kan ik dat zelf ook op een virtuele machine doen? Het is maar om de kernel patch te kunnen blijven gebruiken (voor nu toch) waarvoor ik ubuntu18 nodig heb.

bcoppens
5:19 PM
dat heeft ook een mogelijk ubuntu18 argument :wink:
5:19
maar ja, je kan dat ook in een VM doen

Lennert Franssens
5:21 PM
Oké, dan kan ik weer verder :slightly_smiling_face: bedankt!
:tada:
1

Lennert Franssens
1:06 PM
Hebben jullie vandaag ook al vroeger tijd voor de meeting? Ik heb geen les meer vandaag dus zou vroeger kunnen komen :slightly_smiling_face:

babrath
1:18 PM
Voor mij op zich wel, misschien 15h?
@bcoppens

brdsutte
1:34 PM
Bart en ik hebben een andere call om 15u

Lennert Franssens
1:36 PM
Oké, dan kom ik om 16u.

bcoppens
4:11 PM
PDF

jenny.pdf
PDF

bcoppens
  11:19 PM
btw, Bert zei onlangs dat je voor BPF admin-capabilities moest hebben, maar 't is beperkter sinds 5.8:

```
       CAP_BPF (since Linux 5.8)
              Employ privileged BPF operations; see bpf(2) and
              bpf-helpers(7).

              This capability was added in Linux 5.8 to separate out BPF
              functionality from the overloaded CAP_SYS_ADMIN
              capability.
```

bcoppens
  11:25 PM
ah of hij had het mss op
SECCOMP_SET_MODE_FILTER
want daar staat natuurlijk niks over CAP_BPF maar wel CAP_SYS_ADMIN... Hoewel je daar ook de optie hebt voor PR_SET_NO_NEW_PRIVS... iiiiinteresting
11:27
ok ja, makes sense om daar ofwel voor CAP_SYS_ADMIN ofwel voor PR_SET_NO_NEW_PRIVS te gaan, nice

Lennert Franssens
  11:41 AM
Ja ik gebruik de PR_SET_NO_NEW_PRIVS om CAP_SYS_ADMIN niet te moeten 'zetten'.
11:44
Dat is nodig om prctl(PR_SET_SECCOMP, SECCOMP_MODE_FILTER, &prog) - met prog een struct met de filter en zijn lengte - te kunnen zetten.

bcoppens
  12:34 PM
Inderdaad, en dus heb je (daarvoor toch) geen admin privileges nodig, wat wel goed is :smile: Maar dus goed dat je het al op de 'niet-root' manier doet dan :smile:

Lennert Franssens
  10:58 AM
Is het eigenlijk wel mogelijk om de return value van een syscall terug te krijgen vanuit een signal handler?

babrath
  10:59 AM
ik zou niet zien waarom niet? Hoe roep je hem op?

bcoppens
  11:00 AM
Die gaat na de syscall toch in rax zitten vermoed ik?

Lennert Franssens
  11:00 AM
Ja, het programma werkt ondertussen al. Maar ik zou graag open(...) uitvoeren. Die wordt dan in de signal handler opgeroepen maar mijn file descriptor krijg ik nooit terug op de plaats waar ik open(...) origineel uitvoerde.
11:01
In de signal handler krijg ik die wel, maar ik vind niet hoe ik die return waarde vanuit de signal handler terug krijg naar die originele plaats dus.

bcoppens
  11:02 AM
Ja, maar dus dan moet de register context aangepast worden met de terugkeerwaarde
11:03
en dat deden ze gelijk maar raar / niet in die 'rechtse, oude' voorbeeldcode (edited)

Lennert Franssens
  11:04 AM
Ahja oke, dan ga ik dat eens proberen :slightly_smiling_face:
11:04
Screenshot from 2021-11-16 11-02-41.png
Screenshot from 2021-11-16 11-02-41.png

11:04
Buiten dat werkt het wel al

bcoppens
  11:04 AM
ah nice :smile:
11:04
dus in principe kan je die waarden in die context gewoon aanpassen

Lennert Franssens
  11:06 AM
Ik ga het direct proberen :smiley:

Lennert Franssens
  11:34 AM

Screenshot from 2021-11-16 11-33-57.png
Screenshot from 2021-11-16 11-33-57.png

11:34
Dat werkt inderdaad :stuck_out_tongue_closed_eyes:

bcoppens
  11:34 AM
:tada: :partying_face:

babrath
  11:36 AM
mooi :slightly_smiling_face:

Lennert Franssens
  8:56 AM
Is de meeting straks online of on campus? Ik heb vandaag geen les dus als het voor jullie lukt om ze online te doen, zou dat wel goed uitkomen :slightly_smiling_face:

bcoppens
  9:07 AM
wij zitten zelf on campus, maar online kan zeker ook :slightly_smiling_face:

Lennert Franssens
  9:09 AM
Oke, dan zie ik jullie straks online. Tot dan! :smiley:

bcoppens
  9:11 AM
Tot straks :smile:

Lennert Franssens
  9:30 PM
Ik zit vast op een syscall die wordt uitgevoerd in een signal handler en terug zou moeten gaan naar de plaats waar de initiele syscall vandaan komt. Kunnen we daar even voor bellen op dinsdag 23/11? Ik heb heel de dag tijd om online een meeting te doen :slightly_smiling_face:

bcoppens
  9:33 PM
Voor mij lukt dat dus, 10u is goed, Bert? :smile:

babrath
  10:16 PM
ja ze

Lennert Franssens
  10:56 PM
Super! Tot dan! :grinning:

bcoppens
  5:19 PM
BTW, 't zal dus morgen een thesiscall zijn he :wink:

Lennert Franssens
  5:25 PM
Ja, ik dacht het al :sweat_smile: rest gaat ook al online door, dus het komt beter uit zo :slightly_smiling_face:

bcoppens
  5:25 PM
:smile:

Lennert Franssens
  2:50 PM
Hebben wij nu een meeting of heb ik het juiste uur verkeerd genoteerd? :sweat_smile:

bcoppens
  2:51 PM
Ahja juist, tijd uit het oog verloren!
2:51
Geef me secondje

Lennert Franssens
  2:52 PM
Ja, geen probleem hoor :wink:

bcoppens
  3:08 PM
vim configure-libc-partial-order-debug.sh # (met --disable-werror)
vim configure-libc-woc.sh # (met --disable-werror)
mkdir build_debug

```
cd build_debug
# ../configure-libc-partial-order-debug.sh # beware: builddir removen als je tussen
beide switcht! (make clean duurt véél te lang)
../configure-libc-woc.sh
# potentially also edit out MVEE_LOG_EIPS in totalpartial-agent.c
time make -j3
make install # -> ~/glibc-build/ als prefix
3:09
debuggen:
GLRO(dl_debug_printf) ("mvee_shm_op %x - %lx - %lx\n", (unsigned)id, size, (unsigned
long)address);
(alle varianten horen hetzelfde te zijn)
cp /projects/ReMon-glibc/install/lib/libc-2.31.so
/opt/repo/patched_binaries/libc/amd64/
of
cp ~/glibc-build/lib/libc-2.31.so /opt/repo/patched_binaries/libc/amd64/libc-2.31.so


bcoppens
  2:02 PM
Call? :smile:


Lennert Franssens
  5:20 PM
Lukt het voor jullie om morgen tussen 13u en 14u eens te bellen om het verdere
verloop van mijn masterproef te bespreken?


bcoppens
  5:21 PM
zeker!
5:21
kijkt naar Bert


babrath
  5:23 PM
yep :slightly_smiling_face:


Lennert Franssens
  5:23 PM
Super, tot morgen! :slightly_smiling_face:


Lennert Franssens
  8:46 PM
Dag meneer Coppens en meneer Abrath, via dit berichtje stuur ik jullie het hoofdstuk
over de technologieverkenning door.
```

PDF

hoofdstuk1.pdf
PDF

:+1:
2

Lennert Franssens
  9:53 AM
Ik heb die nieuwe versie geïnstalleerd in de container en krijg deze error (die ik voordien niet kreeg):
lennertfranssens@18606d95c5a9:/projects/ReMon/MVEE/bin/Debug$ ./mvee -- test
Opening MVEE Monitor Log @ ./Logs/MVEE.log
Opening log for non-instrumented instructions @ ./Logs/non-instrumented.csv
MONITOR[0] - WARNING: Executing non-diversified variants: /usr/bin/bash -- bash -c test
MONITOR[0] - WARNING: caught fatal monitor exception: Variant:1 [PID:00575] - Reading/Writing memory failed - when: write runs_under_mvee_control shared memory tag - errno: Input/output error
MONITOR[0] - WARNING: Backtrace requested. current monitor state: STATE_NORMAL
MONITOR[-1] - WARNING: signalling monitor 0 for shutdown - monitor state is: STATE_NORMAL
all monitors terminated
Program terminated after: 0.298766 seconds

babrath
  9:56 AM
hoe heb je dit geinstalleerd?

Lennert Franssens
  9:59 AM
mkdir build/
cd build/
cmake ..
make emulate-shm
make -j$(nproc)
En dan in MVEE/bin/Release ./mvee -- test in mijn vorig bericht staat in de Debug directory maar dat was om eens te kijken of ik daar een andere output kreeg, maar dat geeft me net dezelfde error. (edited)
10:03
Ahaa ik heb ./bootstrap.sh niet gedaan. Dat zal het zijn

babrath

10:06 AM
als je in docker werkt zou bootstrap.sh eigenlijk niet nodig moeten zijn, alle dependencies moeten in de container zelf zitten

Lennert Franssens
10:07 AM
Zijn jullie nu aanwezig in het iGent gebouw?

babrath
10:07 AM
nee

Lennert Franssens
10:09 AM
Anders kijken we er morgen naar tijdens de meeting. En dan doe ik ondertussen verder met glibc :slightly_smiling_face:

babrath
10:10 AM
welke branch zit je?

Lennert Franssens
10:10 AM
De master branch van de ReMon-MVEE organisatie.
10:11
Buiten de laatste commit zie ik nu juist

babrath
10:11 AM
hm, vreemd, ik heb het probleem niet
10:12
ik build wel op /opt/repo/build, ipv /projects/ReMon/build
:heart:
1

10:12
heb je 2 checkouts?
10:12
twee keer de ReMon repo gecloned, wil ik zeggen

Lennert Franssens
10:13 AM
Ja, klopt

babrath
  10:13 AM
dat lijkt me geen goed idee :stuck_out_tongue:

Lennert Franssens
  10:14 AM
Dan probeer ik ook eens in /opt/repo/build :smile:

babrath
  12:12 PM
is het gelukt?

Lennert Franssens
  12:35 PM
Ja dat werkte zoals het moest wanneer ik het op die plaats build. Maar ik heb ondertussen een veel groter probleem. Ik krijg mijn computer niet meer aan. De voeding geeft wel nog de juiste spanningen en ook de power-pins werken nog. Dus ik denk dat mijn moederbord niet meer werkt aangezien er niks gebeurt wanneer ik de power-pins verbind op mijn leeg moederbord. Ik heb vanmorgen alles op mijn laptop geïnstalleerd en heb tussen de lessen al wat met glibc geëxperimenteerd. Daar ben ik nu de syscall wrapper aan het aanpassen en een custom syscall functie aan het schrijven.

babrath
  12:40 PM
oei, das problematisch :slightly_smiling_face:

Lennert Franssens
  12:45 PM
Dat vind ik ook :sweat_smile:

bcoppens
  3:08 PM
@Lennert Franssens
 call? :slightly_smiling_face:

Lennert Franssens
  3:17 PM
Dat stond voor morgen in mijn agenda. Ben ik mis?

bcoppens
  3:17 PM
uuuuuuuuuuuuh
3:17
Morgen om dit uur past alvast niet, maar het uur daarna wel :sweat_smile:

Lennert Franssens
  3:18 PM
Wel er staat tussen 13u en 14u. Maar het kan dus ook aan mij liggen :sweat_smile:

bcoppens
  3:19 PM
euh
3:19
dat lijkt me onwaarschijnlijk :sweat_smile:

Lennert Franssens
  3:20 PM
Dan ligt het dus aan mij. Mijn excuses. Doen we het dan morgen wanneer het voor jullie past?

bcoppens
  3:20 PM
14u15-15u zou eigk beter passen dan :sweat_smile: (edited)
3:20
of of 16u
3:20
kan ook zijn dat wij onduidelijk waren!
3:20
of fout :sweat_smile:
3:21
14u15 dan? :slightly_smiling_face:

Lennert Franssens
  3:24 PM
Misschien 16u dan best? Dan is de les van compilers gedaan. Nu ik daaraan denk zal het toch wel mijn fout geweest zijn want anders had onze meeting overlapt met die les.

babrath
  3:24 PM
ok

bcoppens
  3:25 PM
ok! :slightly_smiling_face:

Lennert Franssens
  3:57 PM
Goed, tot morgen! :slightly_smiling_face:
:+1:
2

Lennert Franssens
  11:05 AM
Ik denk dat ik iets gevonden heb:
https://public-inbox.org/libc-alpha/e68e016fc1573fa57a14dbe419641fa7c1b22f9c.1568219400.git.isaku.yamahata@gmail.com/ (origineel:
https://sourceware.org/pipermail/libc-alpha/2019-September/106501.html) (edited)

babrath
  2:48 PM
dat kan zeker gebruikt worden om alle plaatsen waar je de syscall instructie moet vervangen te vinden

babrath
  2:49 PM
maar het build/linker-probleem dat we vorige week hadden gaat wel blijven

4 replies
Last reply 4 months agoView thread

babrath
  2:49 PM
ik heb al wel wat code om dat te op te lossen (voor 2 van die plaatsen waar je syscall moet vervangen), maar ik was nog niet 100% tevreden over het resultaat

Lennert Franssens
  3:02 PM
Ahaa, super om te horen! :slightly_smiling_face:

babrath
  9:47 AM
replied to a thread:

maar het build/linker-probleem dat we vorige week hadden gaat wel blijven
en werkt dat?
View newer replies

Lennert Franssens
10:18 AM
replied to a thread:
maar het build/linker-probleem dat we vorige week hadden gaat wel blijven
Ja, al een paar keer getest en het werkt zoals het zou moeten. Ik heb nu een kleine macro geschreven om meerdere instructies te kunnen uitvoeren en dat werkt ook nog altijd zonder problemen. Dus ik denk dat de "constructie" om de systeemaanroepen in glibc al te onderscheppen nu wel goed zit en verder uitgebreid kan worden.

```
#ifdef __ASSEMBLER__
.macro custom_syscall_function
        syscall
        nop
        nop
.endm
#else
# define custom_syscall_function        \
        "syscall\n\t"                   \
        "nop\n\t"                       \
        "nop\n\t"
#endif
```

View newer replies

babrath
2:57 PM
git patch (git applyen via 'git am FILENAME'), om de syscall instructie te vervangen met een call naar een functie die ook wel degelijk gelinked en al wordt
:slightly_smiling_face:
0001-work-in-progress.patch

```
From 8884a4c7f81887e2a8cfc2e179399a3ba207c319 Mon Sep 17 00:00:00 2001
From: babrath <bert.abrath@UGent.be>
Date: Wed, 23 Feb 2022 14:53:05 +0100
Subject: [PATCH] work in progress
```

Click to expand inline (66 lines)

2:58
die custom_syscall kan je natuurlijk hernoemen, zolang dat maar consistent overal gebeurt
2:58
dit is nog niet gecombineerd met die commits die je zelf gevonden had, dat moet je zelf nog doen
3:01
de custom_syscall functie is ook nog basically leeg, ik heb het niet zelf gestest dus ik weet niet 100% zeker of deze skeleton code eigenlijk zuccesvol een syscal

doet :wink:
3:01
ik zou denken van wel om dat alle argumenten via registers worden gepassed, maar ik weet het niet zeker
3:02
deze versie is ook nog niet ideaal, de custom_syscall functie wordt opgeroepen via de PLT, terwijl dat niet moet
3:02
dat is een linker/symbol-visibility issue, ik zal daar wel naar kijken, ik denk/hoop dat dat nog geen impact heeft op jou

Lennert Franssens
5:29 PM
Het is gelukt om het eenvoudig mee te implementeren (veel werk was het natuurlijk niet meer voor mij :sweat_smile:) met hetgeen ik hal had. Ik heb al eens snel getest en alles lijkt nog te werken op deze manier. Ik heb eens met ud2 i.p.v. syscall geprobeerd en dat geeft mij ook de verwachtte output (geen dus). - Link naar de commit:
https://github.com/lennertfranssens/ReMon-glibc/commit/b897b055d7e529c047c71f2371ee56527bfc0147

Lennert Franssens
5:38 PM
Dat lukt blijkbaar nog door gewoon de syscall-instructie op te roepen in de __custom_syscall functie :slightly_smiling_face:

babrath
ik zou denken van wel om dat alle argumenten via registers worden gepassed, maar ik weet het niet zeker
Direct Message | Feb 23rd | View conversation

Lennert Franssens
11:18 AM
Ik zit nog met een klein probleempje i.v.m. ptrace. Als ik mijn seccomp filter voor de varianten instel, loopt monitor 1 vast in de staat STATE_WAITING_ATTACH. Ik heb al geprobeerd om de ptrace-opties aan te passen in de attach-functie en waar de andere opties worden ingesteld maar dat lijkt niks te veranderen.

bcoppens
11:19 AM
op welke manier heb je de opties aangepast?

babrath
11:21 AM

voor zowel monitor als variant

Lennert Franssens
  11:23 AM
Hier in MVEE/Inc/backends/ptrace/MVEE_ptrace.h
Screenshot from 2022-03-02 11-22-11.png
Screenshot from 2022-03-02 11-22-11.png

11:23
PTRACE_O_TRACESECCOMP heb ik toegevoegd

babrath
  11:23 AM
en dat is het enige verschil?

Lennert Franssens
  11:23 AM
Ja

babrath
  11:23 AM
en als je dat weer weghaalt is het probleem opgelost?

Lennert Franssens
  11:24 AM
Neen

babrath
  11:25 AM
ah?

Lennert Franssens
  11:25 AM
Dat geeft net dezelfde fout

babrath
  11:26 AM
dan heeft het niets met seccomp te maken?

Lennert Franssens

11:26 AM
Screenshot from 2022-03-02 11-25-51.png
Screenshot from 2022-03-02 11-25-51.png

11:26
Ik denk van wel want het is begonnen nadat ik dit heb toegevoegd
11:26
Screenshot from 2022-03-02 11-26-31.png
Screenshot from 2022-03-02 11-26-31.png

11:27
```
// Set seccomp filter for this variant
                if (prctl(PR_SET_SECCOMP, SECCOMP_MODE_FILTER, &prog) == -1)
                {
            warnf("Failed to set seccomp filter\n");
            exit(-1);
            return;
                }
```
11:28
(in MVEE.cpp)
11:30
Het programma blijft in die while loop hangen
11:30
Dus de variant kan niet attachen aan de monitor (zoals de fout in de terminal ook zegt)

Lennert Franssens
11:31 AM
Nu heb ik al wat opgezocht de afgelopen dagen maar er staat nergens iets over die attach bij het gebruik van seccomp met ptrace

1 reply
4 months agoView thread

babrath
11:31 AM
wat is de bedoeling van die prctl(PR_SET_SECCOMP, SECCOMP_MODE_FILTER, &prog) ?

Lennert Franssens
11:36 AM
Dat stelt de bpf filter in voor dat proces (variant)
11:36
&prog is de struct die de filter bijhoudt
11:37
Ahja en het probleem is er niet als ik alle syscalls doorlaat (ALLOW) in de filter

11:38
Wel als ik ze naar de tracer stuur

babrath
11:47 AM
wat mij lijkt dat er gebeurt..
11:47
is dat je voor deze variant een filter opzet voor elke system call

Lennert Franssens
11:48 AM
Ja klopt eigenlijk wel
11:48
buiten de subset system calls die naar ipmon mogen

babrath
11:49 AM
en dan een system call doet als deel van het hele proces van tracing-setup (eerst is de main monitor thread de tracer, maar dit moet eigenlijk een aparte, dedicated monitor thread worden)
11:50
maar dat deze niet voorbij de filter geraakt? en je variant dus stilhangt
11:50
je kan naar alle processen/threads kijken in top (of htop, nog beter)
11:50
en zien welke state deze processen hebben
11:51
R is running, S is sleep, T is stopped, t is trace-stopped (i.e., wachtende op toestemming van de tracer om verder te gaan)

Lennert Franssens
11:56 AM
Ik zal dat eens bekijken

bcoppens
12:05 PM
replied to a thread:
Nu heb ik al wat opgezocht de afgelopen dagen maar er staat nergens iets over die attach bij het gebruik van seccomp met ptrace
Maar had je dat niet al gedaan met een kleiner testprogrammaatje?

Lennert Franssens
12:06 PM
Ja en daar werkt het wel, maar daar zit je natuurlijk niet met die attach van een

monitor. Daar is het gewoon een ouder- en kindproces.

babrath
  12:15 PM
je kan altijd proberen met die prctl later te plaatsen, na het overdragen van de ptrace-relatie, maar net voor dat de eigenljike variant-code begint

Lennert Franssens
  12:16 PM
Na de while-loop dan
12:16
Ik ga dat straks ook eens proberen dan
12:16
Lijkt me wel een goed idee :wink:

Lennert Franssens
  10:06 AM
Wacht ik even in de vergaderruimte?

bcoppens
  10:07 AM
Ja, wij zijn nog bezig :smile:

Lennert Franssens
  10:07 AM
Ja, ik zag het :smile:

babrath
  1:05 PM
Je kan ook eens naar alle vermelde gevallen van ENOSYS in https://man7.org/linux/man-pages/man2/seccomp.2.html kijken, of deze de oorsprong zijn

babrath
  1:45 PM
Ik denk dat ik weet at het probleem is: een leuke en zeer subtiele interactie tussen seccomp en ptrace. Ik ga ze later in detail uitschrijven, maar ik denk dat je al wel tekenen gaat opmerken als je voor elke syscall volgende acties logt: elke bpf-filter invocatie en result, en elke uitvoering van monitor::handle_event met status.reason
.
:heart:
1

5 replies
Last reply 4 months agoView thread

Lennert Franssens
2:15 PM
Ik zal dat straks eens doen :slightly_smiling_face:

Lennert Franssens
11:07 AM
Een kleine update al. Als write getracet wordt (en ook alle andere), wordt die syscall 2 keer uitgevoerd. Ik ben nog aan het zoeken hoe de logs door SECCOMP_FILTER_FLAG_LOG in het audit.log bestand komen.
2 files

allow.log
Plain Text

trace_write.log
Plain Text

11:10
En in static bool resume_until_syscall (pid_t variantpid, int pending_signal_to_be_delivered=0) staat ptrace(PTRACE_SYSCALL, variantpid, 0, (void*) (long) pending_signal_to_be_delivered) waardoor alle system calls, ook die waarvoor de bpf-filter een ALLOW-actie stuurt, getracet worden.

babrath
11:12 AM
idd
11:12
dat leek mij het probleem

Lennert Franssens
11:12 AM
In mijn eigen programma heb ik dat kunnen "oplossen" door ptrace(PTRACE_CONT,...) te gebruiken. Maar in de MVEE zorgt dat voor problemen met de synchronisatie.

babrath

11:13 AM
de monitor gebruikt altijd ptrace(PTRACE_SYSCALL) om de variant te laten continuen
11:13
op die manier zal de variant stoppen bij elke syscall entry, als syscall exit (i.e., net na het uitvoeren van de syscall in de kernel)
11:13
de monitor wordt dus 2x genotified per syscall
11:14
met het toevoegen van de bpf-filter wordt de monitor echter NOG een keer genotified bij de syscall entry
11:14
en kan de monitor dus 2x genotified worden bij entry, en 1x bij exit

Lennert Franssens
11:15 AM
Dus moet er een manier gevonden worden om die PTRACE_SYSCALL te veranderen naar PTRACE_CONT?

babrath
11:16 AM
ja, maar enkel voor exit events
11:16
bij entry events (veroorzaakt door de bpf-filter) wil je PTRACE_SYSCALL nog steeds gebruiken, anders zou je niet van de exit genotified worden
11:17
merk ook op de ik daarnet zij dat de monitor 2x KAN genotified worden bij entry
11:17
soms is het 1x
11:17
volgende excerpt uit de ptrace manpage is echt leuk:

> Syscall-enter-stop and syscall-exit-stop are indistinguishable from each other by the tracer. The tracer needs to keep track of the sequence of ptrace-stops in order to not misinterpret syscall-enter-stop as syscall-exit-stop or vice versa. In general, a syscall-enter-stop is always followed by syscall-exit-stop, PTRACE_EVENT stop, or the tracee's death; no other kinds of ptrace-stop can occur in between. However, note that seccomp stops (see below) can cause syscall-exit-stops, without preceding syscall-entry-stops. If seccomp is in use, care needs to be taken not to misinterpret such stops as syscall-entry-stops.

11:19
het is dit misinterpreteren van een entry event als een exit event dat leidt tot de -38 return code, dit is niet echt de return van de uitgevoerde system call. De system call is nog niet uitgevoerd, en deze -38 (aka -ENOSYS) wijst voor x86 eigenlijk op het feit dat dit een syscall entry is.
11:19
Zie volgende paragraaf uit dezelfde manpage:

> Some architectures allow the cases to be distinguished by

examining registers. For example, on x86, rax == -ENOSYS in
syscall-enter-stop. Since SIGTRAP (like any other signal) always
happens after syscall-exit-stop, and at this point rax almost
never contains -ENOSYS, the SIGTRAP looks like "syscall-stop
which is not syscall-enter-stop"; in other words, it looks like a
"stray syscall-exit-stop" and can be detected this way. But such
detection is fragile and is best avoided.

bcoppens
  11:20 AM
'almost never' :wink:
11:21
That said, mss is dit ook interessant?
PTRACE_GET_SYSCALL_INFO (since Linux 5.3)
11:21
The op field has one of the following values (defined in
<linux/ptrace.h>) indicating what type of stop occurred
and which part of the union is filled:

PTRACE_SYSCALL_INFO_ENTRY
The entry component of the union contains
information relating to a system call entry stop.

PTRACE_SYSCALL_INFO_EXIT
The exit component of the union contains
information relating to a system call exit stop.

PTRACE_SYSCALL_INFO_SECCOMP
The seccomp component of the union contains
information relating to a PTRACE_EVENT_SECCOMP
stop.

PTRACE_SYSCALL_INFO_NONE
No component of the union contains relevant
information.
11:22
(De vraag is wat dit allemaal met de overhead gaat doen, maar daarvoor hebben we deze thesis, right? :smile:)

babrath
  11:22 AM
idd :slightly_smiling_face:

Lennert Franssens
  11:23 AM
Oké dat zijn dingen waar ik nu wel even verder met kan :smile:

babrath
  11:23 AM
perfect :slightly_smiling_face:

bcoppens
  11:24 AM
Yay :smile:

babrath
  11:24 AM
aarzel niet om hulp te vragen, dit is duidelijk al wat complexer aan het worden :wink:

bcoppens
  11:24 AM
idd :stuck_out_tongue:
11:25
Maar ik vermoed dat het met die PTRACE_GET_SYSCALL_INFO wel moet lukken, en dus een geluk dat je kernelversie recent genoeg daarvoor is :wink:
:+1:
1

Lennert Franssens
  11:27 AM
Welja ik vroeg me nog af of het de bedoeling is dat ik dit allemaal zelf weet/vind of dat het ook nog steeds een soort leerproces is waarbij hulp wel toegelaten is om meer inzichten te verkrijgen? :sweat_smile: Beetje bij beetje wordt alles wel duidelijk maar het duurt soms even

babrath
  11:28 AM
je mag altijd om hulp vragen
11:29
we proberen dan te helpen op een manier die ook duidelijk maakt hoe wij aan onze info zijn geraakt, zodat je het hopelijk in het vervolg meer op eigen houtje tot de juiste inzichten kunt komen
:heavy_check_mark:
1

11:29
maar zoals gezegd, dit was echt een complexe en subtiele interactie, dat is niet vreemd dat je dat niet snel gevonden zou hebben
:heavy_check_mark:

1

11:30
en geen van ons zou er echt veel mee gewonnen zijn als je hier weken op had vast gezeten

Lennert Franssens
11:31 AM
Oké, super! :slightly_smiling_face:
11:32
Laten jullie ook weten wanneer jullie een meeting kunnen houden volgende week? Maandag-, dinsdag en woensdagnamiddag ben ik volledig vrij. Donderdag valt iets moeilijker volgende week. Maar vrijdagmiddag tussen 12u45 en 14u30 heb ik ook nog tijd :slightly_smiling_face:

babrath
11:34 AM
dinsdagnamiddag fysiek?

bcoppens
11:34 AM
Woensdag & vrijdag gaat, maar dan is het voor mij een call :wink:
11:34
Dinsdag kan fysiek idd :slightly_smiling_face:

Lennert Franssens
11:35 AM
Dinsdagnammidag is goed :slightly_smiling_face:

babrath
11:35 AM
hangt af van wanneer je in iGent per toeval bent ook he, remote mag ook zeker

Lennert Franssens
11:36 AM
Ik ben dinsdag toch in Gent dus dat is niet zoveel moeite om even naar daar te komen :wink:

bcoppens
11:37 AM
15:00? :slightly_smiling_face:

Lennert Franssens
  11:37 AM
Dat is goed :slightly_smiling_face:

bcoppens
  11:37 AM
Staat in onze agenda!
:+1:
2

babrath
  3:12 PM
PDF

Lennert_technologieverkenning.pdf
PDF

Lennert Franssens
  5:20 PM
Dit was het originele idee:
5:21
Screenshot_20220322-172030_Slack.png
Screenshot_20220322-172030_Slack.png

bcoppens
  7:59 PM
Euh ja, WOEPSIE FLOEPSIE
:laughing:
1

8:00
Dat lijkt me dus complexer dan nodig om retrospect, denk ik (edited)
8:00
Als in: het reproduceert de originele functionaliteit wel zo dicht als mogelijk
8:01
Maar dat biedt volgens mij in deze niet veel nuttigs

babrath
  8:18 PM
Aha!

8:18
Welja, doe het eerst simpel :slightly_smiling_face: en dan doen we een security evaluation om te zien of het complexer moet

Lennert Franssens
8:23 PM
Oké, dat komt in orde :smiley:

Lennert Franssens
10:40 AM
Is het voldoende om in glibc ergens bij het opstarten een syscall(...) uit te voeren met een fake syscall nummer en als argument het adres van die ene syscall-instructie? Die vang ik dan op in de monitor en in de handler zet ik dan een flag dat het adres gekend is en het adres natuurlijk ook.

babrath
10:41 AM
dat lijkt mij wel? :slightly_smiling_face:

Lennert Franssens
10:42 AM
En die syscall zal altijd uitgevoerd worden nadat de monitor aan de variant gekoppeld is?

babrath
10:46 AM
hmm
10:46
ik zou echt zoeken naar waar andere fake MVEE syscalls staan
10:47
die worden gegarandeerd op de juiste moment uitgevoerd
10:47
misschien kan je die zelfs uitbreiden

Lennert Franssens
10:47 AM
Ahja dat ik gewoon een extra argument meegeef
10:47
Dat is een goed idee

babrath
10:47 AM
idd

Lennert Franssens
  11:10 AM
Ja, die gebruikt al 6 argumenten... Dan maak ik een tweede maar doe hem op dezelfde plaats als MVEE_RUNS_UNDER_MVEE_CONTROL, die gebruikt wordt om het adres van die infinite loop enzo in te stellen
:+1:
1

Lennert Franssens
  12:43 PM
Nieuwe werking:
De initiele BPF-filter is een die op alles SECCOMP_RET_ALLOW teruggeeft. Dit zorgt ervoor dat er geen seccomp_stop events gebeuren en PTRACE_SYSCALL gewoon kan uitgevoerd worden zoals het vroeger werkte. Bij het binnenkomen van de fake syscall (die werkt ondertussen) verandert de state van de variant zodat die weet waar de syscall instructie in glibc zit en wordt een filter toegevoegd. Door het veranderen van die state wordt op alle plaatsen waar PTRACE_SYSCALL wordt uitgevoerd bij een syscall_exit-event nu een PTRACE_CONT gedaan zodat het seccomp_stop-event als entry gezien wordt.
12:44
Het enige waar ik nog niet helemaal uit ben is het type van het adres van de syscall-instructie...
12:48
Ik zou denken een unsigned long maar dat zal duidelijk moeten worden bij het testen

babrath
  12:49 PM
unsigned long, lijkt mij, maar kijk naar gelijkaardige gevallen
:slightly_smiling_face:
12:50
waarom zou je nog een initiële bpf-filter hebben, als deze toch altijd ALLOW'ed?

Lennert Franssens
  12:55 PM
Ja dat is iets dat ik eens kan testen of het ook nog werkt zonder initiele filter

babrath
  12:58 PM
dat lijkt mij op eerste zicht functioneel hetzelfde, dus ik zou verwachten van wel
:slightly_smiling_face:

Lennert Franssens

2:03 PM
Is het wel mogelijk om vanuit de tracee nog eens seccomp() uit te voeren na execve()? Want om een filter toe te voegen moet de tracee zelf die filter instellen, als ik het goed begrijp
2:04
En dat moet dan net op het moment zijn dat de monitor het adres krijgt van de syscall-instructie

babrath
2:10 PM
ok, ja, eigenlijk, dat is waar :sweat_smile:
2:10
maar dat maakt het op zich net gemakkelijker?

babrath
2:11 PM
als het de tracee zelf is die de bpf filter moet instellen, wel, die kent het adres van de syscall instructie zeker :slightly_smiling_face:

1 reply
3 months agoView thread

babrath
2:11 PM
een fake syscall naar de MVEE om de vlag te zetten en naar seccomp-modus over te schakelen is wel nog steeds nodig

babrath
2:12 PM
hoewel je de seccomp syscall zelf als dusdanig zou kunnen beschouwen?

1 reply
3 months agoView thread

Lennert Franssens
2:18 PM
replied to a thread:
als het de tracee zelf is die de bpf filter moet instellen, wel, die kent het adres van de syscall instructie zeker :slightly_smiling_face:
Ook al net na zijn fork()?

babrath
2:19 PM
ja, maar dit adres verandert wel na een execve

Lennert Franssens
2:20 PM
Dan kan ik de code voor de nieuwe filter toch niet meer toevoegen na execve? Want tijdens het toevoegen moet het adres al gekend zijn

babrath
2:28 PM
hoe wil je exact zeggen?

Lennert Franssens
2:29 PM
Het probleem is dus dat ik in de programmacode die door execve wordt uitgevoerd (i.e. het programma dat gemonitord moet worden) de systeemaanroepen om de filter toe te voegen moet uitvoeren. Maar aangezien dat het te monitoren programma is, kan ik die code niet aanpassen.

babrath
2:31 PM
na een execve wordt glibc opnieuw geladen en geïnitialiseerd
2:34
de bpf filters worden wel behouden overheen een execve, dus als die filters enkel syscalls vanop een specifiek adres doorlaten, gaat na een execve de syscall-instructie weer op dat exacte adres moeten staan :sweat_smile:

Lennert Franssens
2:35 PM
Ja, als een execve in het programma wordt uitgevoerd?
2:37
Maar ik zie dus nog niet in hoe ik die bpf-filter kan toevoegen na een execve? (Hoe ik de code daarvoor moet schrijven, vooral waar)

babrath
2:39 PM
op zich gewoon op dezelfde plaats als die fake andere syscalls staan?

1 reply
3 months agoView thread

Lennert Franssens
2:39 PM
Want nu deed ik dat in MVEE.cpp in de start_monitored() functie. Maar na die execve zal ik dus opnieuw

```
// Define empty BPF-filter while starting variant
                struct sock_filter filter[] = {
                    /* [0] Load the return address from 'seccomp_data' buffer into
accumulator */
                    BPF_STMT(BPF_LD | BPF_W | BPF_ABS, (offsetof(struct
seccomp_data, instruction_pointer))),
                    /* [1][A] Jump forward B-A-1 instructions if return address does
not match the syscall address. */
                    BPF_JUMP(BPF_JMP | BPF_JEQ | BPF_K, (unsigned
int)variants[variantnum].syscall_address_ptr, 0, (unsigned char)B_A_1),
                    /* [2] Load system call number from 'seccomp_data' buffer into
accumulator. */
                    BPF_STMT(BPF_LD | BPF_W | BPF_ABS, (offsetof(struct
seccomp_data, nr))),
                    /* [3-9] Jump forward 0 instructions if system call number does
not match '__NR_XXX'. */
                    BPF_JUMP(BPF_JMP | BPF_JEQ | BPF_K, __NR_getpid, 7, 0),
                    BPF_JUMP(BPF_JMP | BPF_JEQ | BPF_K, __NR_getegid, 6, 0),
                    BPF_JUMP(BPF_JMP | BPF_JEQ | BPF_K, __NR_geteuid, 5, 0),
                    BPF_JUMP(BPF_JMP | BPF_JEQ | BPF_K, __NR_getgid, 4, 0),
                    BPF_JUMP(BPF_JMP | BPF_JEQ | BPF_K, __NR_getpgrp, 3, 0),
                    BPF_JUMP(BPF_JMP | BPF_JEQ | BPF_K, __NR_getppid, 2, 0),
                    BPF_JUMP(BPF_JMP | BPF_JEQ | BPF_K, __NR_write, 1, 0), // TODO:
change to gettid to test
                    /* [10] Jump forward 1 instructions if system call number does
not match '__NR_XXX'. */
                    BPF_JUMP(BPF_JMP | BPF_JEQ | BPF_K, __NR_getuid, 0, 1),
                    /* [11] Execute the system call */
                    BPF_STMT(BPF_RET | BPF_K, SECCOMP_RET_ALLOW),
                    /* [12][B] Execute the system call in tracer */
                    BPF_STMT(BPF_RET | BPF_K, SECCOMP_RET_TRACE),


                    // Check if return address is from syscall instruction in glibc
                    // TRUE:
                    //  Execute filter. Can we allow the syscall?
                    //      True:
                    //          Allow
                    //      FALSE:
                    //          Return trace
                    // FALSE:
                    //      Return trace
                };

                // Set BPF-filter
                struct sock_fprog prog = {
                    (unsigned short)(sizeof(filter) / sizeof(filter[0])),
                    filter,
                };
```

```
                // Avoid the need for CAP_SYS_ADMIN
                if (prctl(PR_SET_NO_NEW_PRIVS, 1, 0, 0, 0) == -1)
                    warnf("Couldn't avoid the need for CAP_SYS_ADMIN\n");

                // Enable seccomp BPF-filtering
                //if (prctl(PR_SET_SECCOMP, SECCOMP_MODE_FILTER, &prog) == -1)
                if (syscall(__NR_seccomp, SECCOMP_SET_MODE_FILTER, 0, &prog) == -1)
                {
                    perror("ERROR");
                    warnf("Couldn't enable seccomp BPF-filtering\n");
                }
```
moeten uitvoeren. En ik vraag me dus af waar ergens ik dat nog kan doen aangezien ik de code van de tracee na de execve in MVEE.cpp eigenlijk kwijt ben

Lennert Franssens
  2:40 PM
replied to a thread:
op zich gewoon op dezelfde plaats als die fake andere syscalls staan?
Aah dat is deel van het process van de tracee ook?

babrath
  2:41 PM
ja
2:42
execve behoudt het proces (en bepaalde process-related dingen, zoals bpf filters) maar wist de address space en start een nieuw uitvoerbaar bestand
2:42
fork dupliceert de address space (en andere process-related dingen, zoals file handles), en creëert dus een nieuw proces met een nieuwe PID

Lennert Franssens
  2:45 PM
Oke, dat verklaart veel :sweat_smile:
2:45
Nu nog een kleine bedenking
2:46
If multiple filters exist, they are all executed, in reverse
       order of their addition to the filter tree—that is, the most
       recently installed filter is executed first. (Note that all
       filters will be called even if one of the earlier filters returns
       SECCOMP_RET_KILL. This is done to simplify the kernel code and
       to provide a tiny speed-up in the execution of sets of filters by
       avoiding a check for this uncommon case.) The return value for
       the evaluation of a given system call is the first-seen action
       value of highest precedence (along with its accompanying data)
       returned by execution of all of the filters.
2:47

Dat wil zeggen dat na een execve, en het instellen van een ALLOW als de syscall van het juiste nieuwe adres van de syscall-instructie komt, de vorige filters opnieuw worden uitgevoerd
2:47
Als de laatste filter TRACE teruggeeft is er geen probleem
2:48
Maar als die ALLOW teruggeeft, zal hij de vorige filter uitvoeren en dan komt het adres niet meer overeen met dat van de huidig uitgevoerde syscall-instructie
2:49
Die zal dan TRACE teruggeven en omdat die hogere prioriteit heeft dan ALLOW zal eigenlijk elke syscall na een execve getracet worden
2:49
En dat is wat we net niet willen (edited)

Lennert Franssens
2:57 PM
Om dat op te lossen kunnen we terug een key gebruiken die constant is/blijft in glibc tijdens de uitvoering van een process, over alle execve's heen. Het probleem is dan weer dat we niet kunnen verifieren of de syscall-instructie vanop de juiste plaats komt
2:58
Gebeurt dat nu ergens in ReMon, die controle van de syscall-instructie?

babrath
3:22 PM
nee
3:25
je kan geen bpf filters verwijderen, dus eens een filter geïnstalleerd wordt, zal deze aanwezig blijven, ongeacht hoeveel execve's er volgen, en voor ook voor alle child processes die via fork geproduceerd worden
3:27
als je dus een bpf-filter 'specialiseert' en verwacht dat alle syscall's van op 1 specifiek adres komen, moet dit zo blijven, na alle execve's, en voor alle child processes
3:27
dat is inderdaad een beperkende factor
3:30
langs de andere kant wil dit ook zeggen dat we 'gewoon' de beslissingslogica kunnen omdraaien: de MVEE stelt de bpf filter in met een bepaald adres, VOOR de execve, en forceert vervolgens dat glibc (of IP-MON, of misschien een andere shared library component die enkel die syscall instruction sequence bevat) altijd op dezelfde locatie geladen wordt
3:31
ik denk ook dat "de syscall moet gebeuren vanop deze instructie" ook enkel een vereiste is om de syscall zonder CP-MON monitoring te kunnen laten gebeuren? i.e., mocht er toch een syscall vanop een ander adres komen, kan deze nog steeds via TRACE naar CP-MON doorgestuurd worden?

bcoppens
  3:32 PM
Dat lijkt me idd de meest logische manier eigk, dat we die locaties gewoon a priori vastleggen

babrath
  3:32 PM
zolang de adressen tussen varianten verschillen, lijkt mij dit nog steeds dezelfde security guarantee te bieden?

bcoppens
  3:33 PM
idd
3:34
zelfs meer
3:34
omdat je zo meer garanties hebt :stuck_out_tongue:
3:35
(over welke code daar staat/dat mag uitvoeren)

babrath
  3:38 PM
nu vraag ik mij af, hoe werkt IP-MON nu weer exact? worden alle syscall oproepen vanuit glibc gehooked en naar IP-MON gestuurd (zodat ook weer enkel van daar syscall instructies gebeuren), of is dit iets dat de in-kernel broker doet?

bcoppens
  3:38 PM
Ik dacht dat de IK-B dat deed?
3:39
Niet 100% zeker

babrath
  3:39 PM
want dat heeft wel implicaties voor hoe dit systeem eigenlijk moet werken, en waar de beslissing tot het aanzetten van seccomp eigenlijk genomen moet worden :stuck_out_tongue:

bcoppens
  3:41 PM
Ja maar seccomp kan dat iig niet doen :stuck_out_tongue:

babrath
  3:43 PM
nee het hooken en doorsturen naar IP-MON gaat glibc zelf moeten doen, en het is dan enkel de syscall instructie uit IP-MON die mag doorgelaten worden door de filter?

bcoppens
  3:45 PM
ja, dat was iig mijn bedoeling. Je kan daar dan nog wel extra foefelen om dat extra 'te verbergen' I guess, omdat je niet wil dat die syscall-instructie misbruikt kan worden. Vandaar ook die redirect in IKB leek me
3:45
Maar ik zou eerst dit laten werken en dan pas extra foefelare :p
3:46
Voor de extra foefelare is mss mogelijk dat we toch een 2-traps syscall approach gaan moeten hebben, bedenk ik me :sweat_smile:

babrath
  3:47 PM
welja, het lijkt mij dan, dat we uiteindelijk willen dat de seccomp filter geïnstalleerd wordt vanuit IP-MON, en dat enkel de syscall-instructie vanuit IP-MON door de filter mag, en dat enkel IP-MON overal op hetzelfde adres geladen gaat moeten zijn

bcoppens
  3:48 PM
HMMMMMM
3:48
ik ga hier nog eens over moeten nadenken

babrath
  3:48 PM
de filter kan dan de entrypoint tot IP-MON returnen (analoog aan de key in het oorspronkelijke design), en glibc kan hier dan naartoe springen

bcoppens
  3:49 PM
ja exact
3:49
dat is exact wat ik aan het denken was met foefelare voor een 2-traps syscall approach :stuck_out_tongue:

babrath
  3:49 PM
ja ik was niet zeker wat ge met 2-traps wou zeggen :stuck_out_tongue:

bcoppens
  3:49 PM
2x syscall :wink:
3:49
technisch gezien kan je dan zelfs dat entrypoint over tijd laten varieren
:stuck_out_tongue:

babrath
  3:49 PM
niet 2x trappen? :stuck_out_tongue:

bcoppens
  3:50 PM
:smile:

babrath
  3:50 PM
wel, maar de filter kan niet aangepast worden

bcoppens
  3:50 PM
want je kan extra filters toevoegen, en de return staat op dezelfde prioriteit
:stuck_out_tongue:
3:50
en dan gaat hij de meest recente nemen :wink:

babrath
  3:50 PM
eww

bcoppens
  3:50 PM
I KNOW RIGHT :smile:
3:50
(prolly wel een performance issue :stuck_out_tongue_winking_eye:)

babrath
  3:50 PM
das wel vertraging elke keer ge zo'n filter toevoegt
3:50
idd

3:50
@Lennert Franssens
, volg je? :smile:

Lennert Franssens
  3:52 PM
Ik ben even niet thuis en heb het wel proberen volgen maar ben nog niet helemaal mee :sweat_smile: ik ga het sowieso straks nog eens goed moeten lezen :smile:

babrath
  3:52 PM
das ok :smile:
3:52
morgenochtend in iGent per toeval?

Lennert Franssens
  3:52 PM
Ja ik kan tot 12u45 naar daar komen :wink:

babrath
  3:53 PM
tussen 10u-12u heb ik alvast (en Bart possibly ook?) tijd om het eens bij te lichten en uit te tekenen op een bord :p

Lennert Franssens
  3:55 PM
Oh dat is een goed idee :smiley:
3:55
Tot morgen dan! :wink:

bcoppens
  3:56 PM
Wel ik zou aan een projectvoorstel moeten schrijven enzo :sweat_smile:
3:56
Maar dit is betere procrastinatie :sweat_smile:

Lennert Franssens
  11:05 AM
https://www.alfonsobeato.net/c/filter-and-modify-system-calls-with-seccomp-and-ptrace/

Lennert Franssens

3:52 PM
Het is de calculate_data_mapping_base functie in MVEE_mman.cpp die ik kan hergebruiken denk ik om het adres in te stellen zodat enkel bits 12->28 niet 0 kunnen zijn

babrath
3:54 PM
klinkt plausibel :smile:

Lennert Franssens
3:56 PM
Ik ben er eigenlijk wel redelijk zeker van aangezien in de syscall handler van mmap in de monitor de returnwaarde van die functie gebruikt wordt om arg1 (de adrespointer) van mmap te overschrijven :slightly_smiling_face:

bcoppens
3:57 PM
:smile:

babrath
4:09 PM
goed :smile:

bcoppens
3:49 PM
https://man7.org/linux/man-pages/man2/seccomp_unotify.2.html waarom heb ik die pagina nog nooit gelezen? :sweat_smile: alle 't is niet relevant voor deze thesis eh, maar wel cool stuff ook

babrath
3:52 PM
ja ma wel ingewikkeld weer :stuck_out_tongue:

bcoppens
3:52 PM
jaja, niet voor deze thesis :stuck_out_tongue_winking_eye:
3:55
> Note well: this mechanism must not be used to make security policy decisions about the system call, which would be inherently race-prone for reasons described next.

Sad :disappointed:

babrath
  3:56 PM
hooray!

bcoppens
  3:58 PM
:smile:

Lennert Franssens
  4:21 PM
Inderdaad, heb er in het eerste semester ook wat met geëxperimenteerd maar door die notes over de security policy decisions heb ik daar niet verder met gewerkt. Maar wel leuke dingen voor de mensen :smile:

bcoppens
  4:22 PM
idd :smile:

Lennert Franssens
  2:03 PM
PDF

Boosting_MVX_Systems_Through_Modern_OS_Extensions.pdf
PDF

babrath
  11:57 AM
feedback is aanwezig in comments
PDF

Boosting_MVX_Systems_Through_Modern_OS_Extensions_feedback.pdf
PDF

11:57
vraag maar als iets niet duidelijk is
11:57
ik denk dat we bij de volgende meeting ook nog eens wat uitleg moeten geven over top-down presentatie van design/implementatie enzo :stuck_out_tongue:

Lennert Franssens
  11:58 AM

Oké, goed. Ik zal het tegen dan al eens overlezen ook :smiley:

Lennert Franssens
  6:16 PM
Oké, goed nieuws... Ik heb een voorbeeldje gevonden waarin ze ook die seccomp_ret_trace gebruiken en de output daarvan is (als ik 226 meegeef als errno return waarde):
write(1, "something's gonna happen!!\n", 27) = -1 (errno 226)
write(1, "it will not definitely print thi"..., 39) = -1 (errno 226)
6:17
Dus 226 zit weldegelijk in errno en rax zal -1 bevatten veronderstel ik (edited)

babrath
  6:26 PM
Waar komt die logging vandaan? Strace?

Lennert Franssens
  6:26 PM
ja strace

babrath
  6:28 PM
Hmm
6:29
Ik ben paranoïde
6:29
Doe eens met gdb? :p

Lennert Franssens
  7:04 PM
Ik ben net thuis. Ik heb ondertussen die inline asm nog even gemaakt. Ik stuur de output en het bronbestand door. Ik test het straks ook nog eens met gdb :slightly_smiling_face: maar ik denk dus dat het ergens anders aan ligt aangezien met die inline asm wel de juiste waarde wordt verkregen...
2 files

test.c
C

test.log
Plain Text

babrath
  9:21 AM
het is niet dat je die waarde niet via errno kan krijgen
9:22
op het moment dat het in die test.c gebeurd
9:22
het is dat die waarde nog niet in errno zit op het moment dat jij het in glibc wil doen (edited)

Lennert Franssens
  9:22 AM
Klopt, het lijkt toch alsof er in rax iets anders zit dan -1 maar ben het nu aan het testen

babrath
  9:22 AM
syscalll return value == RAX == -ERRNO
9:22
in dit geval:
9:23
Screenshot from 2022-04-20 09-23-15.png
Screenshot from 2022-04-20 09-23-15.png

9:23
wat -226 is
9:24
glibc neemt de waarde in RAX, en verplaatst die naar errno
9:24
dat verplaatsen gebeurt echter NA de locatie in de code waar jij je syscall fragmentje hebt staan

Lennert Franssens
  9:25 AM
Aaah dat verklaart veel, ik was ook al zoiets aan het denken dat het pas daarna in errno wordt geplaatst en door dat te testen zag ik net dat rax niet -1 is
9:26
dan nemen we gewoon het inverse van rax als die groter is dan 0x085
9:26
En ik heb voor de zekerheid ook nog eens getest en de max waarde is inderdaad 4095

Lennert Franssens

9:37 AM
Ik denk dat ik dan toch goed nieuws heb nu. Het werkt om naar de ipmon syscall functie te springen :slightly_smiling_face:

babrath
9:39 AM
Hoera :smile:

bcoppens
9:43 AM
:partying_face:

Lennert Franssens
9:43 AM
Dat was even lastig :sweat_smile: nu ga ik de ipmon syscall functie nog even debuggen en hoop dat ik dan kan nagaan of de instruction pointer overeenkomt met het label dat ik in de ipmon syscall functie gezet heb. Als dat dan werkt kan ik het probleem met rdtsc oplossen en normaal zou het dan moeten werken. En als het dan werkt, maak ik een programma om filters met op te stellen (bv adhv json? waarin dan een soort van default trace staat buiten die dat in dat bestand beschreven staan). En dan kan ik hopelijk wel aan de benchmarks beginnen :slightly_smiling_face:

babrath
9:47 AM
ok :slightly_smiling_face:
9:47
klinkt goed

Lennert Franssens
11:47 AM
Wat moet ik eigenlijk van de ipmon_enclave functie overhouden?

babrath
11:50 AM
alles wat niet deel is van het vervangen van de in-kernel broker? :stuck_out_tongue:

Lennert Franssens
11:53 AM
Ahja oké :smile:

Lennert Franssens
1:34 PM

Die ipmon_enclave functie zorgt wel voor heel wat problemen...

babrath
  1:34 PM
hoezo?

Lennert Franssens
  1:35 PM
Als ik alles laat doorgaan als unchecked syscall dan werkt alles zoals het zou moeten maar vanaf er gewisseld kan worden tussen unchecked en checked syscalls of er met de RB gewerkt wordt, krijg ik MAPERR's vanuit de cross-process monitor

babrath
  1:36 PM
checked syscalls zijn de syscalls die vanuit de CP-MON afgehandeld worden?

Lennert Franssens
  1:37 PM
Ja, klopt. Maar als ik alles op checked syscalls zet dan werkt het ook
1:37
Maar dan gaan ze natuurlijk allemaal naar de cross process monitor

babrath
  1:37 PM
huh
1:37
en aan de hand van wat beslis je tussen checked en unchecked?

Lennert Franssens
  1:39 PM
Logica die al in die ipmon_enclave functie zit.
Bv.:
if (RB->have_pending_signals & 2) -> checked
if (ipmon_syscall_is_unsynced(args, syscall_no) -> unchecked
1:40
Nog een "mooi" voorbeeld van iets dat niet werkt is de ipmon_prepare_syscall functie. Als ik die oproep krijg ik ook een MAPERR fout :confused:

babrath
  1:41 PM
en het is CP-MON die MAPERR geeft? niet de IP-MON?

Lennert Franssens
  1:41 PM
De CP-MON

babrath
  1:41 PM
kun je de log eens tonen? :slightly_smiling_face:

Lennert Franssens
  1:43 PM
Ik zal een doorsturen waarin ik die ipmon_prepare_syscall oproep
:slightly_smiling_face:
1:43
Binary

MVEE.log
Binary

babrath
  1:56 PM
het is wel degelijk IP-MON die de MAPERR heeft (edited)
1:57
0.305908 - MONITOR[0] - Variant:0 [PID:10554] - variant crashed - trapping ins:
0000000000022d1b: ipmon_pos_to_pointer(ipmon_buffer*) at ??:?
(/opt/repo/IP-MON/libipmon.so)
0.305915 - MONITOR[0] - Variant:0 [PID:10554] - Received signal SIGSEGV (11)
0.305919 - MONITOR[0] - Variant:0 [PID:10554] - signal arrived while variant was
executing ins: 0000000000022d1b: ipmon_pos_to_pointer(ipmon_buffer*) at ??:?
(/opt/repo/IP-MON/libipmon.so)
0.305921 - MONITOR[0] - Variant:0 [PID:10554] - ret is currently: 0
0.305931 - MONITOR[0] - WARNING: Warning: SIGSEGV in variant 0 (PID: 10554)
0.305937 - MONITOR[0] - WARNING: IP: 0000000000222d1b, Address: 0000000000000246,
Code: SEGV_MAPERR (1), Errno: 0
0.305942 - MONITOR[0] - Variant:0 [PID:10554] - ==================================
0.305944 - MONITOR[0] - Variant:0 [PID:10554] - generating local backtrace for
variant
0.305945 - MONITOR[0] - Variant:0 [PID:10554] - > variant is currently suspended
0.305948 - MONITOR[0] - pid: 10554 - 000: 0000000000022d1b:
ipmon_pos_to_pointer(ipmon_buffer*) at ??:? (/opt/repo/IP-MON/libipmon.so)
0.306029 - MONITOR[0] - pid: 10554 - 001: 0000000000022d6e:
ipmon_prepare_syscall(ipmon_buffer*, ipmon_syscall_args&, unsigned long) at ??:?
(/opt/repo/IP-MON/libipmon.so)
0.306062 - MONITOR[0] - pid: 10554 - 002: 00000000000231d9: ipmon_enclave at ??:?
(/opt/repo/IP-MON/libipmon.so)
0.306082 - MONITOR[0] - pid: 10554 - 003: 0000000000026038:

ipmon_enclave_entrypoint_alternative at ??:? (/opt/repo/IP-MON/libipmon.so)
0.306264 - MONITOR[0] - pid: 10554 - 004: 00000000001050c5:
glibc_custom_syscall_exit at custom_syscall.c:?
(/opt/repo/patched_binaries/libc/amd64/libc-2.31.so)
1:57
dus de crash gebeurt in variant 0, specifiek de instructie op adres 0x22d1b

Lennert Franssens
1:58 PM
hier dus: 22d1b: 8b 07 mov (%rdi),%eax

babrath
1:59 PM
het is een SIGSEGV (segmentation fault), specifiek MAPERR (niet gemapped geheugen):
zie https://man7.org/linux/man-pages/man2/sigaction.2.html
2:00
de stack trace geeft aan dat het dus in de functie ipmon_pos_to_pointer gebeurt. De specifieke source line staat er niet bij, maar die zou je normaal wel krijgen als je ipmon met debug-versie build

Lennert Franssens
2:01 PM
Dat is maar een kleine functie

babrath
2:01 PM
wel, dan kan je wss op zicht zeggen welke dereference er fout loopt
:stuck_out_tongue:

Lennert Franssens
2:02 PM
```
return (void*)((unsigned long)RB +
                                    64 * (RB->numvariants + 1) +
                                    RB->variant_info[ipmon_variant_num].pos);
```
2:02
RB zelf dan toch?

babrath
2:02 PM
dat lijkt me waarschijnlijk
2:04
gezien het adres dat de variant probeert te accessen 0x246 (== NULL page) is, zou ik vermoeden dat RB hier simpelweg NULL is
2:04

maar dat kan je gemakkelijk testen door het eens uit te printen

Lennert Franssens
  2:07 PM
Ja, dat is dus NULL

babrath
  2:09 PM
volgende vraag: op welke plaats zou dat eigenlijk ingesteld moeten worden op iets anders dan NULL, :smile:

Lennert Franssens
  2:13 PM
De een paar lijnen onder waar ik het nu zet :sweat_smile: of het zou meegegeven kunnen worden als argument vanop de plaats waar ipmon_enclave opgeroepen wordt

babrath
  2:14 PM
hoe was het vantevoren, en hoe heb je het aangepast?

Lennert Franssens
  2:45 PM
Ik ben eens teruggegaan naar het begin en ik heb dus de registratie van ipmon ipv een prctl via een fake syscall gedaan. Wanneer ik dat nu terug omdraai blijft hij hangen net na die prctl. Ik ga het even aan de kant leggen en er vanavond nog eens naar kijken

Lennert Franssens
  9:46 AM
Kunnen we nog eens kort bellen voor wat meer info over de RB in ipmon?

babrath
  10:02 AM
OK
10:02
nu?

Lennert Franssens
  10:04 AM
Ja, binnen 5 minuutjes ben ik er :slightly_smiling_face:

babrath
  10:09 AM
teams doet irritant, sec

Lennert Franssens
  10:09 AM
Oké :slightly_smiling_face:

Lennert Franssens
  2:38 PM
Als ik de args variabele meegeef in ipmon_enclave naar de functie ipmon_syscall_is_unsynced krijg ik een segmentation fault. Als ik die niet meegeef, dan krijg ik die fout niet.
2:38
```
/*-----------------------------------------------------------------------------
    ipmon_syscall_is_unsynced
-----------------------------------------------------------------------------*/
unsigned char ipmon_syscall_is_unsynced(struct ipmon_syscall_args& args, unsigned long syscall_no)
{
        switch(syscall_no)
        {
#include "MVEE_ipmon_is_unsynced.h"
        }
        return 0;
}
```
2:39
Args wordt niet gebruikt in die functie dus ik kan die weglaten
2:40
Maar verderop in ipmon_enlcave heb ik die args toch nodig om door te geven met ipmon_prepare_syscall bijvoorbeeld.

babrath
  2:41 PM
ben je zeker dat args niet gebruikt wordt in die ge-include header?

Lennert Franssens
  2:43 PM
Ja, daar staat enkel dit in:
```
case __NR_munmap: return ipmon_handle_munmap_is_unsynced();
case __NR_uname: return ipmon_handle_uname_is_unsynced();
case __NR_sched_yield: return ipmon_handle_sched_yield_is_unsynced();
case __NR_madvise: return ipmon_handle_madvise_is_unsynced();
```

babrath

2:43 PM
als args niet gebruikt wordt, hoe leidt het dan tot een segfault?

Lennert Franssens
2:44 PM
En ik heb ook geprobeerd om return 1 boven de switch te zetten. En met de args geeft het een fout en zonder args niet.
2:46

```
0.318786 - MONITOR[0] - Variant:0 [PID:15296] - SYS_FSTAT64 return
0.318788 - MONITOR[0] - File type:                character device
0.318790 - MONITOR[0] - I-node number:            3
0.318792 - MONITOR[0] - Mode:                     20620 (octal)
0.318794 - MONITOR[0] - Link count:               1
0.318796 - MONITOR[0] - Ownership:                UID=1000   GID=5
0.318798 - MONITOR[0] - Preferred I/O block size: 1024 bytes
0.318800 - MONITOR[0] - File size:                0 bytes
0.318802 - MONITOR[0] - Blocks allocated:         0
0.318813 - MONITOR[0] - Last status change:       Tue Apr 19 17:59:22 2022
0.318816 - MONITOR[0] - Last file access:         Thu Apr 21 14:45:52 2022
0.318819 - MONITOR[0] - Last file modification:   Thu Apr 21 14:45:52 2022
0.318875 - MONITOR[0] - INFO: syscall_info = UNKNOWN
0.322476 - MONITOR[0] - Variant:0 [PID:15296] - variant crashed - trapping ins:
00000000000231c5: ipmon_enclave at ??:? (/opt/repo/IP-MON/libipmon.so)
0.322490 - MONITOR[0] - Variant:0 [PID:15296] - Received signal SIGSEGV (11)
0.322498 - MONITOR[0] - Variant:0 [PID:15296] - signal arrived while variant was
executing ins: 00000000000231c5: ipmon_enclave at ??:?
(/opt/repo/IP-MON/libipmon.so)
0.322504 - MONITOR[0] - Variant:0 [PID:15296] - ret is currently: 120864061795256
0.322529 - MONITOR[0] - WARNING: Warning: SIGSEGV in variant 0 (PID: 15296)
0.322537 - MONITOR[0] - WARNING: IP: 00000000002231c5, Address: 0000000000000000,
Code: SI_KERNEL (128), Errno: 0
```

babrath
2:47 PM
klinkt vreemd
2:48
kan je de huidige code eens doorsturen?

Lennert Franssens
2:48 PM

```
       if (ipmon_syscall_is_unsynced(args, syscall_no)) {
               long ret = ipmon_unchecked_syscall(syscall_no, args.arg1, args.arg2,
args.arg3, args.arg4, args.arg5, args.arg6);
#ifdef MVEE_IP_PKU_ENABLED
               erim_switch_to_untrusted;
#endif
               return ret;
```

```
        }
```

babrath
  2:48 PM
ook verder terug, vanaf het begin van de functie aub :stuck_out_tongue:

2 replies
Last reply 2 months agoView thread

Lennert Franssens
  2:48 PM
Wanneer ik de code hieronder uitvoer, krijg ik geen fout...

```
        ipmon_syscall_args a;
        if (ipmon_syscall_is_unsynced(a, syscall_no)) {
                long ret = ipmon_unchecked_syscall(syscall_no, args.arg1, args.arg2,
args.arg3, args.arg4, args.arg5, args.arg6);
#ifdef MVEE_IP_PKU_ENABLED
                erim_switch_to_untrusted;
#endif
                return ret;
        }
```

2:50
Het verschil met de code erboven is dat ik een "lege" variabele van het type ipmon_syscall_args aanmaak

babrath
  2:52 PM
heb je de assembly code die ipmon_enclave oproept ook aangepast?

Lennert Franssens
  2:53 PM
Ja, een klein beetje

babrath
  2:53 PM
hoe? want aangezien je de functie-signatuur hebt aangepast maakt dit wel uit

Lennert Franssens
  2:54 PM
Screenshot from 2022-04-21 14-54-08.png
Screenshot from 2022-04-21 14-54-08.png

2:55

Die mov r12,r9 is de boosdoener zeker?

babrath
  2:57 PM
wel, je overschrijft daar r12, ik weet niet of dat al dan niet mag op die moment
2:57
met de mov r9, r8 overschrijf je echter r9, en dat mag zeker niet
2:57
die werd in de oorspronkelijk code nog ge pushed en later gepopped

Lennert Franssens
  2:58 PM
Oké, dan ga ik die code eens beter proberen "begrijpen" en schrijven
:slightly_smiling_face:

babrath
  2:58 PM
je moet je aan de calling convention houden hier: sommige registers zijn caller-saved, andere zijn callee-saved, en op het moment van de call moet de stack aligned zijn op 16 bytes
2:58
amd64 calling convention, meer specifiek
:+1:
1

2:59
System V AMD64 ABI, nog specifieker :stuck_out_tongue:
:heart:
1

Lennert Franssens
  4:53 PM
Ik denk dat het helemaal werkt. Ik ga sowieso nog wat testen vanavond zodat ik niks over het hoofd zie, maar op het eerste zicht lijkt het dus te werken
:stuck_out_tongue:

babrath
  4:55 PM
hola? dat klinkt goed :smile:

Lennert Franssens
  5:01 PM
Ja hé :stuck_out_tongue_closed_eyes: dus dan blijft de planning zoals die dat ik een

paar dagen geleden hier had doorgestuurd. Met dat verschil dat de instruction pointers ook al juist gecheckt worden in de bpf filter. Een die uit de functie unchecked komt, krijgt SECCOMP_RET_ALLOW en een die uit checked komt wordt naar de CP-monitor gestuurd. Iets dat ik wel nog moet doen is het "random" bepalen van het adres waarop ipmon gemapt wordt maar dat lijkt mij niet zo heel veel werk meer :slightly_smiling_face:

babrath
5:03 PM
ok :slightly_smiling_face:

Lennert Franssens
5:05 PM
Ahja mag ik ervan uitgaan dat het enkel op x86_64 moet werken?

babrath
5:09 PM
ja
5:10
maar wees daar wel expliciet in als je het schrijft, en gebruik architectuur-specifieke code enkel in de juiste plaatsen, niet in het midden van generieke code

Lennert Franssens
5:11 PM
Oké :slightly_smiling_face:

brdsutte
9:05 AM
@Lennert Franssens
zou een proefverdediging kunnen op 4, 5, of 6 mei?

Lennert Franssens
9:11 AM
Ja, dat zou lukken. Enkel op 5 mei ben ik in de voormiddag niet beschikbaar.

Lennert Franssens
8:35 AM
Goeiemorgen, past het vandaag om even te bellen? Ik zit vast op een paar foutjes m.b.t. het mappen van ipmon in het geheugen met meerdere varianten. En specifieker, wanneer een execve wordt uitgevoerd in de variant. (edited)

1 reply

2 months agoView thread

babrath
10:15 AM
ja hoor
10:16
dus het werkt met 1 variant, en nu ben je aan het proberen met 2, en daar loopt het fout?
10:16
ik heb tijd tot 11u15 ongeveer
10:17
en anders na 15u

Lennert Franssens
10:18 AM
Ik heb nu ook een uurtje tijd :slightly_smiling_face:

babrath
10:19 AM
ok, teams?

Lennert Franssens
10:20 AM
Ja :slightly_smiling_face:

bcoppens
2:14 PM
Ter info: de tussentijdse thesisverdediging gaat door op 6 mei 13:00-14:30, meeting room 1.2 in iGent

Lennert Franssens
5:13 PM
Oke, geen idee wat ik de vorige keer verkeerd heb gedaan. Maar nu werkt IP-MON op de oude manier wel op Ubuntu 20.04 met een 5.4.0 kernel met de patch.

bcoppens
5:15 PM
:partying_face:

babrath
6:29 PM
Vreemd

6:29
Maar goed om horen :D

Lennert Franssens
6:32 PM
Is er nog iemand kort beschikbaar? :slightly_smiling_face:

babrath
6:39 PM
Hmm
6:39
Straks, rond 20u ook goed?

Lennert Franssens
6:39 PM
Ja dat is goed, als het past voor u he :wink:

bcoppens
8:00 PM
Ik ben op restaurant :sweat_smile:

babrath
8:00 PM
ik ben op teams

Lennert Franssens
8:17 PM
Smakelijk :wink:

bcoppens
Ik ben op restaurant :sweat_smile:
Direct Message | May 2nd | View conversation

bcoppens
11:07 PM
@Lennert Franssens
wegens corona gaan de thesispresentaties online door op vrijdag ipv fysiek
:grimacing:

Lennert Franssens
11:13 PM

Dat is geen probleem :slightly_smiling_face:

Lennert Franssens
5:07 PM
Screenshot from 2022-05-04 16-58-28.png
Screenshot from 2022-05-04 16-58-28.png

:tada:
1

5:07
Dit is het resultaat van de benchmarks

babrath
5:08 PM
waar staat de 0-1-2-3-4 egk voor?

2 replies
Last reply 2 months agoView thread

Lennert Franssens
5:09 PM
Links is de nieuwe versie van ipmon, rechts de oude. Exact dezelfde benchmarks gedaan. Bij ipmon met seccomp-bpf zet ik alles eens op trace, wat in de oude versie overeenkomt op USELESS_POLICY. De run daarna zet ik alles op trace buiten getpid en bij de oude implementatie zet ik daar FULL_SYSCALLS aan.

babrath
5:10 PM
en je moet dus 2x langs de bpf-filter gaan op hier, right?

Lennert Franssens
5:11 PM
Ja, twee keer. Een keer om errno terug te sturen. De tweede keer om RET_ALLOW te doen.

babrath
5:12 PM
ja, ok
5:13
dus de bpf-filter is sneller dan de in-kernel broker, wat mij wel een beetje verbaast

Lennert Franssens
  5:13 PM
Ja dat had ik ook helemaal niet verwacht
5:14
Ik zou wel nog een policy kunnen maken voor de oude versie van ipmon die enkel getpid doorlaat ipv alle systemcalls die hij ondersteunt

babrath
  5:14 PM
er is wel geen feature parity, de in-kernel broker doet nog wel een aantal checks die misschien vertraging geven?

Lennert Franssens
  5:14 PM
Ja de key uit register 12 is volledig weg
5:15
en ook de RB die vanuit de kernel telkens moest komen is nu ook gewoon thread_local in ipmon
5:15
Dat zijn twee zaken die anders door de kernel ook gedaan werden

babrath
  5:15 PM
er is natuurlijk ook het feit dat de IK broker 2 64-bits secrets teruggeeft ipv de 12-bit errno hier, maar zo'n groot verschil zou ik daar niet van verwachten (wel, ik zou geen groot verschil verwachten in het ideale geval dat de bpf-filter een 128-bits secret zou kunnen teruggeven...)
5:18
ja.. misschien dat het daar van komt, hoewel een key in een register zetten (en er later weer uit lezen en comparen) me niet zo traag had geleken dat ik het als oorzaak van de overhead zou verwachten? Langs de andere kant, dat is wss een random waarde die de broker moet genereren?

Lennert Franssens
  5:18 PM
Ja, dat zal het zijn
5:19
Ik kijk snel even in de kernelpatch maar het zal die generate van die key zijn die het verschil geeft

babrath
  5:20 PM
wel, voor meer uitsluitsel, op een later moment: nginx evalueren, en proberen een

meer secure versie met meerdere bpf-filter-calls die meerdere 12-bits secrets doorgeven

1 reply
2 months agoView thread

Lennert Franssens
  5:21 PM
Ja, dat lijkt me ook een goed idee :slightly_smiling_face:
5:22
Maar ookal ligt het niet volledig in de lijn van onze verwachtingen, toch is het beter zo dan dat we nu al met veel vertraging zouden zitten :slightly_smiling_face:

babrath
  5:22 PM
true

bcoppens
  5:42 PM
Oh nice!
5:45
Dus zowel de originele ipmon als de bpfmon hebben dezelfde policy geïmplementeerd?

Lennert Franssens
  5:46 PM
De originele heeft een iets uitgebreidere policy maar de timing wordt gedaan op 500000 getpid's (edited)
5:46
Ik ga straks nog een extra policy toevoegen aan de originele om enkel getpid toe te laten. Dan hebben ze exact dezelfde policy.

babrath
  5:47 PM
je kan de originele ook eens aanpassen om geen get_random_int meer te doen, en gewoon secret key 0 te gebruiken ofzo
5:47
puur per wijze van experiment om de bron van extra overhead te zoeken

bcoppens
  5:47 PM
Right, idd
5:47
die randomness zou't wel eens kunnen zijn :stuck_out_tongue:

babrath
  5:48 PM
't zou het easy kind of randomness moeten zijn, die niet gaat stallen (als ik het juist lees), maar dan nog, extra calls en locks dieper in de kernel enzo :slightly_smiling_face:

bcoppens
  5:48 PM
idd
5:48
'iets uitgebreidere policy' kan ook 'iets meer code uitgevoerd' betekenen :stuck_out_tongue:
:sweat_smile:
1

5:49
maar 't zal allicht een combinatie van beide zijn
5:49
en 't zou wel goed zijn als je dan die aanpassingen eens probeert zoals bert zei, om de bron van overhead te zoeken (evt meerdere bronnen dus!)

Lennert Franssens
  5:51 PM
Dat ga ik zeker nog proberen doen voor vrijdag. Ik ga vanavond eerst de slides nog wat afwerken :slightly_smiling_face:
:+1:
2

Lennert Franssens
  9:29 PM
IP-MON kernel met ipmon_key = 0 en IP-MON policy waarbij enkel getpid in IP-MON uitgevoerd wordt bij het testgeval 'ipmon enabled mvee with getpid allowed'.
Screenshot from 2022-05-04 21-26-53.png
Screenshot from 2022-05-04 21-26-53.png

9:31
(het is wel de omgekeerde volgorde van schermen, sorry :sweat_smile:)
9:32
De nieuwe versie van IP-MON is een klein beetje sneller in vergelijking met de vorige benchmarks omdat ik geen onnodige achtergrondprocessen meer heb draaien. De originele versie is een heel pak sneller in vergelijking met de vorige benchmarks (inderdaad, door het uitschakelen van de random key). Ik zie wel nog steeds een klein verschil in snelheid in het voordeel van de nieuwe implementatie. (edited)

babrath
  9:36 PM
Dus allebei de aanpassingen verlagen de overhead. Is de bpf filter nog steeds sneller?

Lennert Franssens
  9:37 PM
Ja de overhead wordt inderdaad kleiner in de originele versie van IP-MON door die twee aanpassingen te doen (1. ipmon_key is altijd 0 en 2. enkel getpid wordt in IP-MON uitgevoerd). Maar met seccomp-bpf blijft het sneller.
9:40
Ik heb daarstraks voor de zekerheid de testen ook nog eens manueel gedaan zonder benchmark mode en er treden ook geen fouten op (waardoor de uitvoersnelheid ook sneller zou zijn door het vroeg optreden van een fout bijvoorbeeld, maar dat is dus niet het geval :slightly_smiling_face:). (edited)

bcoppens
  9:49 PM
Ah nice :smile:

babrath
  9:51 PM
:+1:

Lennert Franssens
  12:04 PM
Ik heb ondertussen ook al eens kort nagedacht over die secret die groter moet worden dan de 12 bits die we nu gebruiken. Aangezien de seccomp-bpf op voorhand gedefinieerd wordt, zal de secret constant blijven doorheen de levensloop van de seccomp-bpf filter. Dat wil niet zeggen dat we op die 12 bits moeten blijven. In plaats van in de eerste errno-waarde direct het adres mee te geven, kunnen we een waarde meegeven waarmee we de start van de uitwisseling van de secret aangeven (gewoon groter dan 0x85, de hoogst gebruikte standaard errno-waarde). Glibc doet dan een syscall met als syscall-nummer de errno-waarde. In de bpf-filter wordt deze waarde opgemerkt en weten we dat we de eerste 12 bits van de secret kunnen teruggeven als nieuwe errno-waarde. Dit komt terug aan in glibc, waar we die waarde bijhouden in een register en daarna doen we een OR operatie van 0x800 op de errno-waarde om het nummer van de volgende syscall te bepalen. Dit doen we om zeker te zijn dat in onze seccomp-bpf filter de waarde niet overlapt met een 'echte' syscall. In de seccomp-bpf filter gaan we weten dat wanneer dat nummer voorbijkomt, we de volgende 12 bits moeten doorsturen. Dit doen we 5 keer om een secret van 60 bits door te sturen. De laagste 4 bits zijn van weinig belang, omdat we de enclave zullen alignen op 4 bits (16 decimaal).
In glibc weten we dan op die manier de secret van 64 bits, het adres van de enclave, waarna we een call naar dat adres kunnen doen. Wanneer we terugkeren clearen we de

registers waarin we de (deel)secret(s) hebben opgeslagen. (edited)
12:06
Enkel de secret voor de RB moet dan nog eens bekeken worden, maar gaat een pak lastiger zijn aangezien we dat niet zomaar in een register kunnen steken zonder het altijd te weten in IP-MON... Of we moeten het in de seccomp-bpf filter steken en garanderen dat we die op een of andere manier kunnen meenemen over execve's heen. (edited)

bcoppens
1:15 PM
Dat klinkt iig redelijk

babrath
1:16 PM
klinkt goed idd
1:17
secret in RB is idd gelijkaardig

bcoppens
1:17 PM
Nu, dat wil wel zeggen dat de uiteindelijke totale filter over alle processen heen groter en groter gaat worden, lijkt me? en dan ga je ook moeten de invloed op uitvoeringstijd daarvan moeten weten/meten :slightly_smiling_face:
1:17
idd

Lennert Franssens
1:18 PM
Ja, en dat zal dan de vertraging introduceren die we eigenlijk nu al verwacht hadden

babrath
1:18 PM
maar daar kun je de vraag stellen hoe belangrijk die secret daar bijhouden is
1:18
eens je alle filters hebt om 60 bits door te sturen voor het IP-MON address, hebben we alvast een beter zicht op welke overhead het secret voor RB zou geven
1:19
also, de hoogste 12-bits de IP-MON enclave kun je eigk ook negeren, die gaan altijd 0 zijn

Lennert Franssens
1:19 PM
Dat is ook al waar

bcoppens
  1:19 PM
idd

Lennert Franssens
  4:56 PM
Welke grafiek(en) neem ik best in de presentatie?
https://ugentbe-my.sharepoint.com/:x:/g/personal/lennert_franssens_ugent_be/EbhmdMBR
m4FIq92eLTOQ3GgBzrIysn52wcKl_b1JE13OTw?e=4riqQZ
4:56
Of moeten er nog andere zijn?

bcoppens
  5:02 PM
De meest linkse, maar dan met een log-schaal voor de Y-as?

Lennert Franssens
  5:03 PM
Dat is inderdaad een heel goed idee die log :wink:

bcoppens
  5:04 PM
:smile:

Lennert Franssens
  5:19 PM
Enkel mijn resultaten moet ik er nog inzetten
PDF

proefverdediging.pdf
PDF

babrath
  9:39 PM
op slide 12 zou ik vermelden dat het adres van RB ook secret is
9:40
op slide 18 ben ik niet zeker waar ineens de term 'interceptor' vandaan komt? die is
niet eerder geïntroduceerd, en stond op vorige slides ook niet bij IK-broker
9:41
ah, die interceptor en verifier zijn de afzonderlijke delen/rollen van de
IK-broker..

Lennert Franssens
  9:42 PM
Ja, maar ik moet die inderdaad misschien gewoon op de vorige slides ook al gebruiken

babrath
  9:42 PM
wel, of dat niveau van detail tijdens deze presentatie iets bijdraagt is mij niet duidelijk, je kan gwn 2x zeggen dat het de IK-broker is

Lennert Franssens
  9:42 PM
voor de duidelijkheid
9:42
Kan ook ja

babrath
  9:44 PM
je nieuw design ziet er wel goed uit, maar de kernel is ineens verdwenen, en in plaats daarvan zit glibc (wat natuurlijk wel een rol speelt en die je wel hebt aangepast) ineens op die plaats
9:44
dat is wat verwarrend
9:44
allee, glibc zit eigk in de variant
9:44
en de bpf-filter zitin de kernel
9:45
wel, in de user space van de variant, net als IP-MON

Lennert Franssens
  9:46 PM
Aaah ik weet al waarom ik dat daar gezet heb. Die staat wel op dezelfde plaats omdat die een deel van de werking van de interceptor overneemt. Maar ik ga dan mijn blok dat de kernel moet voorstellen kleiner maken en het blok van de user space uitbreiden naar onder toe

babrath
  9:46 PM
de slides lijken me wel een duidelijk verhaal te vertellen, maar het is wss te veel

Lennert Franssens
  9:46 PM

Ik kom op 19 minuten uit :sweat_smile:

babrath
  9:46 PM
ik denk dat de slides over glibc (50*) verborgen mogen
9:46
ahh?
9:46
hmm
9:47
zonder gigantisch snel te babbelen? :stuck_out_tongue:

Lennert Franssens
  9:47 PM
Ja op een normaal tempo
9:47
die slides van glibc ga ik eigenlijk snel over

babrath
  9:47 PM
hmm ok dan

Lennert Franssens
  9:47 PM
dat is op minder dan een minuut uitgelegd
9:48
Maar ik ga het nog een paar keer oefenen morgenvroeg en als ik zie dat ik toch te lang bezig ben (door het echt te vertellen) dan weet ik dat ik die slides kan wegdoen :slightly_smiling_face:

babrath
  9:49 PM
ja idd
9:49
goed :slightly_smiling_face:

Lennert Franssens
  12:47 PM
Is er een link naar de meeting?

bcoppens
  12:51 PM
Ik hoop dat die nog gaat komen :stuck_out_tongue:

12:51
Normaal zal't de zoom meeting room zijn van prof De Sutter :slightly_smiling_face:

Lennert Franssens
12:51 PM
Oke :slightly_smiling_face:

bcoppens
12:52 PM
Ik ga er van uit dat hij die zoomlink hier wel zal posten, maar een reminder kan geen kwaad
@brdsutte
^ :smile:

brdsutte
12:58 PM
https://ugent-be.zoom.us/j/3335288368
Call

Zoom meeting
Ended at 9:32 PM - Lasted 1 day
Meeting ID: 333-528-8368
0 people joined
Added by Zoom

brdsutte
2:38 PM
Uit mijn notities:
- CP-MON niet in originele figuren
- complexe figuren tegelijkertijd op slides en daarna is het niet precies duidelijk over welk deel je aan het praten bent of zelfs wat de onderdeeltjes van je figuur voorstellen.
- waarom slide 12 niet vervangen door een figuur, of extra zaken toevoegen aan bestaande figuren?
- slides 14 ev: niet telkens tekst in een nieuwe titel zetten (stel je voor dat per seconde beeld in een doc een nieuwe titel verschijnt …)
- slide 21: GHUMVEE op slide, CP-MON in wat verteld wordt
- slide 26: twee stijlen van visualisatie (niveaus van abstractie) door elkaar gebruik
- slide 52 ev: veel te grote fonts

bcoppens
2:39 PM
en uit mijn notities:
Slide 9 had eigk mss geintegreerd kunnen worden in slide 8, had je dan kunnen duiden

op figuur
11 en 12 zijn mss saai: veel tekst (volzinnen), maar mss adhv een figuur was dat leuker geweest?
Rond slide 30'ish, bij de downsides van seccomp-bpf ook vermelden dat je enkel kan toevoegen en dat enkel de zwaarste actie doorkomt. Zeker dat eerste is belangrijk, want dat heeft implicaties op de secrets…
Bij slide 40 zei je iets ala ‘als de system call veranderd is [in de 2e syscall tov de 1e] dan blokkeren we dat ook’, maar seccomp kan dat toch niet weten?
Slide 57 mss ook wat te lange zinnen? (En mismatched fontsizes)
hoe ziet je microbenchmark er uit? wat wil je daarmee meten? Ok, dat was nu de 2e slide, ik zou die omdraaien
De labels op de as links (1 10 100 …) slide 64 en verder groter aub! Die titel ‘Time to execute 1 getpid syscall (ns)’ had je beter links 90 graden naar links als aslabel gezet. aantal decimalen is beetje overkill
Dat van die ‘secrets uitgezet voor een gelijke benchmark te bekomen’ was niet duidelijk, was zeker al niet duidelijk dan dat al je resultaten eigk op een gepatchte ‘originele’ ipmon waren. Was ook niet altijd duidelijk wat belangrijk was om naar te kijken in de grafieken / of er iets belangrijk was

Lennert Franssens
2:40 PM
Bedankt!

Lennert Franssens
9:30 AM
Ik heb de opmaak nog wat aangepast want die was nog zoals voordien. Dit weekend vul ik de uitleg over de microbenchmarks aan. Volgende week focus ik op het oplossen van de fout bij een fork in de nieuwe implementatie waarna ik die nginx benchmarks kan doen en daarover kan schrijven, samen met een conclusie. Dan moet ik enkel nog een abstract schrijven tegen 9 juni.
PDF

Boosting_MVX_Systems_Through_Modern_OS_Extensions.pdf
PDF

Lennert Franssens
1:02 PM
Zijn jullie morgen in het iGent gebouw?

babrath
1:05 PM
Nee

Lennert Franssens
  1:05 PM
Oké

babrath
  1:05 PM
Woensdag ben ik (en miss Bart) er wel

Lennert Franssens
  1:07 PM
Ik werk vanavond de tekst over de microbenchmarks af (want heb nu nog een inhaalles). Morgen ga ik proberen het probleem met fork op te lossen. Afhankelijk van hoe goed/snel dat gaat zal ik morgenavond laten weten of ik dan woensdag langskom :slightly_smiling_face:

babrath
  1:42 PM
Ok
1:43
We zullen vandaag of morgen eens lezen

bcoppens
  2:03 PM
Woensdag ben ik er zeker

Lennert Franssens
  4:20 PM
Oké, goed :smiley:
4:21
Heel veel is er niet gewijzigd. Gewoon een iets betere top-down uitleg. En ik heb nog wat kleine, essentiële dingen toegevoegd hier en daar om dat deel (H1, H2 & H3) helemaal "klaar" te maken.

babrath
  10:21 AM
ik ben het nu aan het lezen, en op eerste zicht is het wel nog redelijk kort
10:22
de evaluatie moet inderdaad nog komen, maar vergeet ook niet nog een inleiding te schrijven?
10:22
ahh ja, daar
10:22
die heeft gwn geen nummer

Lennert Franssens
  10:22 AM
Ja die indruk had ik ook al

babrath
  10:22 AM
verwarrend :stuck_out_tongue:

Lennert Franssens
  10:23 AM
Inleiding was hoofdstuk 0, maar die nummering daarbij moest ik verwijderen :smile:
10:24
Ik weet ook niet echt hoe het nog veel langer geschreven moet worden?

babrath
  10:24 AM
dat mag ook gwn hoofdstuk 1 zijn?

Lennert Franssens
  10:24 AM
Ahja dan ga ik dat zo aanpassen

babrath
  11:09 AM
alvast wat feedback, met ook hier en daar suggesties om dingen uitgebreider uit te leggen, of figuren toe te voegen
PDF

feedback_lennert.pdf
PDF

:heart:
1

Lennert Franssens
  8:03 PM
Er zijn al vorderingen met het probleem bij een fork :slightly_smiling_face:
8:04
Maar ik kom zeker morgen naar het iGent gebouw om het nog eens met jullie te bekijken want ergens verderop (ik denk bij het teruggaan naar de parent als ik het juist interpreteer) is er een synchronisatiefout

babrath
  8:05 PM
Ok!

Lennert Franssens
  8:18 PM
Het probleem is trouwens enkel nog als er meerdere varianten zijn. Met 1 variant is het fork-probleem al opgelost :smiley:
:tada:
1

babrath
  8:19 PM
ah nice :slightly_smiling_face:

Lennert Franssens
  7:40 AM
Bij het terugsturen van een waarde als antwoord op een signaal, zou de monitor dat op een exit-event moeten doen maar hij springt naar een entry-event (en skipt het exit-event dus). Dus ik vermoed dat ik ergens nog een aanpassing moet maken, of een aanpassing teveel heb gemaakt, van PTRACE_SYSCALL naar PTRACE_CONT.

Lennert Franssens
  8:05 AM
En dat is dan weer bij een STOP_SIGNAL event, waarna een STOP_SYSCALL komt wat eigenlijk een STOP_SECCOMP zou moeten zijn denk ik

Lennert Franssens
  4:50 PM
De resultaten van een aangepaste versie waarbij de filter meerdere keren doorlopen worden. Er wordt een secret van 64 bits doorgegeven aan de hand van een secret die 12 bits groot is. Die eerste secret wordt nu gebruikt om de volgende segmenten te berekenen, maar daarvoor zou in praktijk ook een afgeleide waarde van het "deelgeheim" gebruikt kunnen worden.
microbenchmarks_chart.png
microbenchmarks_chart.png

4:51
Het is niet zo slecht als ik eerst dacht, maar toch ook nie zo goed als een secret van 12 bits :wink:
4:51

De achtste en tiende bar zijn hier het nuttigst om te bekijken

babrath
  4:52 PM
ahja
4:52
76, omdat 64 + 12

Lennert Franssens
  4:52 PM
Ja

babrath
  4:52 PM
kan je ook eens 48 proberen? of 47 = 11 + 3*12?

Lennert Franssens
  4:52 PM
Ja

babrath
  4:53 PM
en de logaritmische bars vind ik persoonlijk nogal verwarrend

Lennert Franssens
  4:54 PM
Mag er dan een beginsecret blijven?
4:54
Om aan te geven dat de secret doorgegeven zal worden
4:55
Of moet dat ook al een deel van de 47/48 bits zijn?
4:56
Aangezien het natuurlijk een soort afgesproken waarde is, is het natuurlijk geen secret maar het telt wel mee in de overhead die gecreëerd wordt

Lennert Franssens
  5:23 PM
Ik maak het nu zo:
1x 12 bits om aan te geven dat het adres zal doorgegeven worden.
3x 12 bits om bits 12->47 van het adres door te geven. Aangezien we toch alignen op 4096 moeten we de eerste 12 bits toch niet doorgeven.

Lennert Franssens
  5:31 PM
Ik ben de testomgeving aan het builden voor die nieuwe implementatie en laat iets weten wanneer ik de resultaten heb :slightly_smiling_face:
5:31
Dit gaat minder lang duren dan de vorige aangezien ik nu maar 1 versie moet testen :smile:

Lennert Franssens
  9:11 PM
Hier de resultaten met een secret van 48 bits t.o.v. alle andere implementaties die getpid doorlaten
microbenchmarks_chart_v2.png
microbenchmarks_chart_v2.png

babrath
  11:07 PM
Ok, dus dat valt beter mee dan verwacht qua overhead. Nu zou ik het wel eens voor nginx willen zien :slightly_smiling_face:

brdsutte
  12:50 PM
ziet er goed uit!

Lennert Franssens
  10:03 AM
Dit zijn de resultaten van de nginx benchmark. We zien dat de originele implementatie van IP-MON ongeveer een gelijke snelheid heeft als een native uitvoering. De implementatie met een seccomp-BPF filter is gevoelig trager dan de originele implementatie, maar wel sneller dan een default MVEE. Het verschil tussen de originele implementatie en de nieuwe implementatie met seccomp-BPF is wel te verklaren. De nieuwe implementatie heeft nog enkele bugs voor enkele kritische systeemaanroepen die het geheel nog meer kunnen versnellen. Die systeemaanroepen zijn:
__NR_fstat
__NR_read
__NR_readv
__NR_pread64
__NR_preadv
__NR_pwrite64
__NR_mprotect
__NR_brk
__NR_open
__NR_openat

__NR_close
Wanneer we de originele implementatie van IP-MON zo configureren dat het deze systeemaanroepen ook naar CP-MON zal doorsturen, zien we dat die versie gelijkaardige snelheden behaalt als onze nieuwe implementatie. Dat wijst erop dat wanneer we de hierboven opgesomde lijst compatibel kunnen maken met de nieuwe versie van IP-MON, we normaal ook snelheden kunnen behalen in dezelfde grootteorde als die van de originele versie van IP-MON met zijn volledige set van compatibele systeemaanroepen. (edited)
2 files
nginx_latency.png

nginx_througput.png

:heavy_check_mark:
1

bcoppens
  11:01 AM
Ziet er goed uit, maar wat bedoel je met 'De nieuwe implementatie heeft nog enkele bugs voor enkele kritische systeemaanroepen die het geheel nog meer kunnen versnellen'? :sweat_smile:

Lennert Franssens
  11:06 AM
Wel de lijst van systeemaanroepen hierboven zijn er die nog niet werken in combinatie met de implementatie met seccomp-BPF. En dat zijn net systeemaanroepen die nginx sneller laten werken als ze naar IP-MON gestuurd kunnen worden. En ik denk wel dat daar een oplossing voor bestaat, bijvoorbeeld meer aanpassingen in de handlers in IP-MON daarvoor, maar die aanpassingen heb ik niet kunnen maken.
:+1:
1

bcoppens
  9:32 AM
image.png
image.png

9:32
de ugent-spamfilter wordt gelijk almaar erger...
9:34
(zat dus in mijn junk mail)
:hushed:

1

brdsutte
  10:59 AM
Zelfde bij mij...

Lennert Franssens
  1:50 PM
Ik stuur mijn voorlopige versie van mijn scriptie door. Enkel de samenvatting van 10 lijnen en het abstract moeten nog gemaakt worden.
PDF

MVX_systemen_versnellen_met_moderne_besturingssysteemextensies.pdf
PDF

:heavy_check_mark:
2

bcoppens
  10:29 AM
@brdsutte

@babrath

@Lennert Franssens
 Thesisverdediging 1 juli om 13:00 ergens in iGent (lokaal nog te bepalen :slightly_smiling_face:)

Lennert Franssens
  10:30 AM
Goed, bedankt! :slightly_smiling_face:

brdsutte
  10:41 AM
@bcoppens
 Leg je de plaats snel vast en geef je alles in in Plato? Als we lang wachten gaan we daarover weer spam krijgen :disappointed:

bcoppens
  10:45 AM
ik ga dat straks na de usenix deadline doen :wink:

babrath
  1:11 PM
@Lennert Franssens
 die tabel/lijst met syscalls die bij seccomp niet meer supported zijn via IP-MON en dus nu via CP-MON, staat die ook in jou thesis?

Lennert Franssens
  1:26 PM
Ahaa neen nog niet. Die zet ik er ook in :wink:

Lennert Franssens
  9:09 PM
Ik heb mijn presentatie verder aangevuld met nog enkele nieuwe zaken en de feedback van de proefverdediging ook verwerkt in mijn nieuwe presentatie. Ziet dat er goed uit voor jullie of zijn er nog zaken die nog moeten veranderen tegen volgende week? De dianummers moet ik nog updaten, tussen de dia's staan nog enkele oude dia's die ik nog niet verwijderd heb maar wel al onzichtbaar staan :slightly_smiling_face: (edited)
PDF

masterproefverdediging.pdf
PDF

babrath
  3:54 PM
slide 24: vergeet niet dat IP-MON wel degelijk ook nog aan checking van argumenten doet
op slide 29/30 zou ik duidelijk maken dat "ptrace" eigenlijk een tracer proces is, dat dus de beslissing maakt over het al dan niet doorlaten van de system call
slide 31 kwam nogal abrupt? is het de samenvatting van de voorgaande slides? met dat allow en trace visueel getoond worden, had ik voor errno iets gelijkaardigs verwacht
slide 33: "combinere"
slide 34: ik zou het niet "omzeilen" noemen, eerder vervangen of iets in die aard
op slide 43 en voorgaande heb je het over "CP-MON", maar op slide 44 is het "MVEE"
op slide 47/48 zou ik duidelijk aanduiden wat de eerste en tweede trap zijn
slide 49: is de secret in de oorspronkelijke IP-MON ook uniek per variant, of niet?
in slides 49-53 ga je er duidelijk van uit dat de secret eigenlijk het adres van IP-MON is, maar dit staat toch nergens expliciet op de slides
slide 61: en wat met binaries die hun eigen syscall instructies uitvoeren ipv glibc aan te roepen hiervoor?
3:55
voor de rest lijkt mij dat wel OK, de visuele voorstelling van hoe alle componenten met elkaar interageren is zeker goed

:+1:
1

Lennert Franssens
  4:55 PM
slide 49: is de secret in de oorspronkelijke IP-MON ook uniek per variant, of niet? Hangt er een beetje vanaf wat we als secrets zien... De secret in de originele implementatie/design is de random waarde die per systeemaanroep wordt gegenereerd. Dat geheim wordt gebruikt in IP-MON en de broker verifier. Maar in de nieuwe implementatie hebben we zo geen 'geheim'. Het geheim dat we daar hebben is direct ook ons adres waar naartoe gesprongen moet worden. Dat is tevens ook een geheim in de originele implementatie, maar wordt niet in beschouwing genomen daar. Moet ik dan de vergelijking doen met de originele implementatie met als geheim het adres van IP-MON, en de gegenereerde secret per systeemaanroep in de originele implementatie als "onmogelijk over te nemen in de nieuwe implementatie" beschrijven en na een korte vermelding niet meer meenemen in de vergelijking. Of houd ik het zoals het er nu staat?

Lennert Franssens
  5:25 PM
Ik heb het aangepast naar 3 secrets die vergeleken worden. 1) De unieke secret per systeemaanroep in de originele, die niet in de nieuwe te implementeren was. 2) De unieke secret per variant, het adres van IP-MON. 3) De unieke secret per variant, het adres van de RB.
PDF

veiligheidsaspect_vergelijking_aangepast.pdf
PDF

babrath
  9:57 PM
dat was ongeveer wat ik ging voorstellen, ja

Lennert Franssens
  10:04 PM
Zo goed dan, of moest er nog iets bij?

babrath
  10:15 PM
Lijkt mij wel goed
10:17
Als je het visueler uitgelegd krijgt is dat beter dan een hoop tekst, maar aangezien

dit een vergelijking van 2 systemen is met redelijk wat kleine details zou ik het zelf ook niet direct beter weten :sweat_smile:

Lennert Franssens
  10:20 PM
Ja, ik zit daar zelf ook wat met gewrongen dat er bij sommige stukken niet zoveel visuele voorstellingen zijn. Maar vaak is het niet evident om het visueel voor te stellen :sweat_smile:
10:20
waar ik nog iets vind om het via een figuur duidelijker te maken, zal ik het nog aanpassen

Messages with prof. dr. Bart Coppens and dr. ir. Bert Abrath
---

bcoppens
6:01 PM
Dag Lennert,
Zoals beloofd heb ik de Raspberry Pi-opgave wat herwerkt. Hier in bijlage is de nieuwe versie. De repo om van te starten bevindt zich alhier: https://github.ugent.be/bcoppens/besturingssystemen-raspberrypi De geannoteerde PDF-manuals waar in de opgave naar verwezen wordt, zijn dezelfde gebleven. Dus aangezien je al op het vak van afgelopen academiejaar ingeschreven bent, kan je die manuals kan je dan hier vinden: https://ufora.ugent.be/d2l/le/content/228922/Home (Voor je thesis zelf zijn die Linux-practica wel nuttiger, daar kan je uiteraard nu ook al aan via Ufora :slightly_smiling_face:)
PDF

opgave.pdf
PDF

Lennert Franssens
9:34 AM
Dag meneer Coppens,
Bedankt voor het berichtje. Ik ga de komende dagen al eens proberen om deze opgave te maken. Op 30 augustus heb ik mijn laatste herexamen. Dan kan ik ook de beginnen Linux-practica maken.
Als er problemen zijn of ik zit ergens vast laat ik het hier weten.
Ik heb er heel veel zin in en kijk ernaar uit om aan de thesis te werken.
Tot binnenkort!
Lennert

bcoppens
9:37 AM
Dat hoor ik graag! Alvast veel plezier en ook al veel succes met de herexamens :-)
(Als je onduidelijkheden of schrijffouten zou vinden hoor ik dat ook graag ;-))
Tot binnenkort!

Lennert Franssens
2:48 PM
Dag meneer Coppens, Is het goed dat er Raspberry Pi OS op het SD-kaartje, dat in de Raspberry Pi zit, staat? Er staat ook een klein foutje op pagina 5. Er staat dat de witte kabel op pin 8 (GPIO 12/TXD) aangesloten wordt. Dit zou GPIO 14/TXD volgens de bovenstaande figuur moeten zijn.

bcoppens
2:53 PM

Dag Lennert, Ja, da's geen probleem, je moet gewoon wat files op de boot-partitie aanpassen dan zodanig dat onze bootloader draait :slightly_smiling_face: (Ik stuur het zometeen door). Dat is inderdaad een foutje, bedankt! Ik fix het in mijn lokale versie :slightly_smiling_face:

bcoppens
  3:03 PM
Backup je oude config.txt, ze daar deze config.txt in de plaats, en zet in dezelfde map (dus de rootfolder van de bootpartitie) de bootloader.bin die je krijgt als je make typt
config.txt
#BART:
enable_uart=1
arm_64bit=1
disable_commandline_tags=1
Click to expand inline (14 lines)

Lennert Franssens
  3:05 PM
Oke, bedankt! :slightly_smiling_face:

Lennert Franssens
  5:53 PM
Is het normaal dat ik een error krijg bij het uitvoeren van make voor de eerste keer? Ik heb even gezocht en vind niet direct zelf een oplossing die werkt om de error op te lossen. Ik heb het ook getest met de VM, maar die geeft me dezelfde error.
Screenshot from 2021-07-30 17-50-28.png
Screenshot from 2021-07-30 17-50-28.png

bcoppens
  5:54 PM
Neen :stuck_out_tongue: Het is vreemd dat je de error ook in de VM krijgt, though. Ben je in de VM van een cleane checkout gestart, of heb je verdergebouwd op de build van buiten je VM?
5:55
Als het enkel buiten de VM zou gebeurd zijn zou ik denken dat je een te recente g++-versie hebt die wat extra features gebruikt waar ik niet op voorzien ben
5:55
Dus ik hoop een beetje dat je nu gewoon diezelfde files in je VM gebruikt, dat zou dat verklaren, ze zijn dan immers erbuiten gemaakt met een andere g++ :wink:

Lennert Franssens
5:56 PM
Ik dan eens een cleane checkout proberen. Want ik ben in dezelfde map gaan werken in de VM als die erbuiten (waar de error al was) :slightly_smiling_face:

bcoppens
6:00 PM
ja dan gaat het allicht daardoor zijn

Lennert Franssens
6:03 PM
Ja inderdaad, nu werkt het wel :slightly_smiling_face:

bcoppens
6:03 PM
:tada:
6:03
for reference, welke g++-versie gebruikte je buiten de VM? :wink:

Lennert Franssens
6:04 PM
Ahaa 9.3, die moest dus 8.3 zijn :sweat_smile:

bcoppens
6:05 PM
:smile:
6:06
nuja, als we ooit zouden switchen naar 9.3 zou dat wel makkelijk te fixen moeten zijn :slightly_smiling_face: maargoed, zorgen voor later :wink:
:laughing:
1

Lennert Franssens
8:30 PM
Is het de bedoeling dat ik iets zie als ik de Raspberry Pi aan een scherm hang? Hij blijft op het 'gekleurde scherm' staan. Ik heb het met RPi OS 32 bit en 64 bit getest.
8:33
Ik had wel al gevonden dat bij een gekleurd scherm: "The RPi isn’t able to load the kernel Image, which can occur if something happened during the write or as a result of a failing SD card."
8:34

De kernel (bootloader.bin) heb ik in de boot partitie gestoken en ook het aangepaste config.txt bestand zit daarin. Als ik de originele config.txt terugzet dan start hij wel normaal op.

bcoppens
  8:36 PM
dat klinkt wel vreemd... :confused: Ik zal je straks / morgen eens wat meer bestanden van mij sturen als ik terug aan mijn laptop zit :stuck_out_tongue:

Lennert Franssens
  8:37 PM
Oké, perfect! Er zit geen haast achter hoor :wink:

bcoppens
  8:37 PM
:smile:

bcoppens
  4:51 PM
Maar dus met die bootloader.bin, heb je daar dan een seriele kabel aangehangen?
4:52
Want initieel bij het eerste practicum, en het merendeel van het 2e practicum, ga je enkel IO doen op je seriele kabel
4:52
En gaat die communiceren met die bootloader.py

Lennert Franssens
  1:13 PM
Ja en als ik dan een screen-sessie probeer te openen over de seriële kabel lukt dat ook niet. De kernel (bootloader.bin) wordt gewoon niet geladen denk ik. Ik heb wel de 8GB versie maar denk niet dat dat het probleem is.
1:15
En als ik het script uitvoer, blijft dat op de bootloader wachten. Heb ik dan een verkeerde versie van Raspberry Pi OS?

bcoppens
  1:16 PM
Ik heb ook de 8GB versie :wink:
1:16
Het idee is in principe dat we ons eigen OS hebben dus dat dat niet zou uit mogen maken
1:16
Maar mss is er iets veranderd in de firmware-files die erbijgelever zijn
1:18

Backup je originele files, en zet deze eens in de plek?
9 files

start4x.elf
Binary

start4db.elf
Binary

start4cd.elf
Binary

start4.elf
Binary

bcm2711-rpi-4-b.dtb
Binary

fixup4x.dat
Binary

fixup4db.dat
Binary

fixup4cd.dat
Binary

fixup4.dat
Binary

Lennert Franssens

1:35 PM
Ja, ik ga dat doen. Moet ik gebruik maken van 64 of 32 bit? Of maakt dat ook niet uit dan?

bcoppens
1:48 PM
64-bit
1:48
alle, je boot-partitie moet de 64-bit files bevatten

Lennert Franssens
1:55 PM
Oké :slightly_smiling_face:

Lennert Franssens
2:30 PM
Een screen-sessie openen lukt nu wel - met veel "ERROR" geprint. Het script bootloader.py blijft wel nog hangen
Screenshot from 2021-08-01 14-26-23.png
Screenshot from 2021-08-01 14-26-23.png

2:30
Screenshot from 2021-08-01 14-28-31.png
Screenshot from 2021-08-01 14-28-31.png

Lennert Franssens
2:36 PM
Als ik nu het oude config.txt bestand terugzet start de Pi ook niet meer op.
2:39
Dit is het gekleurde scherm dat ik te zien krijg :sweat_smile:
image.png
image.png

Lennert Franssens
3:19 PM
Ik heb dit op het SD-kaartje gezet: https://downloads.raspberrypi.org/raspios_arm64/images/raspios_arm64-2021-05-28/. Via deze link stuur ik u mijn boot-partitie met de 'originele' .elf, .dtb en .dat bestanden, de aangepaste config.txt en de bootloader.bin die gemaakt is via de VM :slightly_smiling_face: Is het andere studenten wel gelukt om via de seriële

verbinding en de aangepaste bootloader hun kernel op de Pi te zetten?

bcoppens
3:25 PM
Ok, dussss:
Die ErrorErrorError is eigenlijk een goed teken! Want dat wil zeggen dat mijn bootloader geladen wordt en de seriele verbinding correct opzet. Want mijn bootloader-code bevat deze code: if (!read_and_compare_line("kernel")) { while (true) printstring_uart("Error"); } Dus de eerste lijn die je moet sturen over de seriele connectie moet "kernel" zijn, anders faalt het :wink: (Da's ook wat bootloader.py doet: uart_connection.send_line("kernel"))
De kans bestaat idd dat je bootloader.py idd wijst naar het foute device? Is het device dat in bootloader.py gebruikt wordt hetzelfde als je voor screen gebruikt?
Dat kleurenscherm komt door de bootcode van de Pi, dat geeft enkel aan dat de GPU zelf de monitor correct kan aanspreken :wink:
Als dat met de bootloader.py niet zou werken kan ik je idd eens mijn bootpartitie doorsturen :stuck_out_tongue: Hoewel ik niet zeker ben dat het daar aan ligt, nu blijft dat de bootloader wel ErrorErrorError... print

Lennert Franssens
3:35 PM
Oh dat eerste maakt me blij! :slightly_smiling_face: Ik heb ondertussen nog eens geprobeerd (met de nieuwe bestanden die u me doorstuurde) en alles uitgevoerd zoals in de opgave beschreven en het is gelukt om de kernel door te sturen :smiley:
3:35
Dan kan ik beginnen aan de eigenlijke opgave :slightly_smiling_face:

bcoppens
3:37 PM
:tada:

Lennert Franssens
3:28 PM
Dag meneer Coppens, ik ben de eerste opgave aan het maken. Er wordt gevraagd om block en lower_attributes te maken. Ik heb in de armv8 manual op D5-2573 gevonden welke bits gebruikt worden voor de lower en upper attributes in een block descriptor. Met de uitleg die daar stond heb ik volgende code geschreven voor de functies block en lower_attributes:

```
uint64_t block(uint64_t output_address /* full 64 bit start address of the block */, uint64_t upper_attr, uint64_t lower_attr) {
  uint64_t addr { 0 };

  addr |= lower_attr << 2;
  addr |= output_address << 17;
  addr |= upper_attr << 51;
```

```
  return addr;
}
uint64_t lower_attributes(int mair_index, int ap, int shareability) {
  uint64_t attr { 0 };

  // (NS) is always 0 - here bit 3
  attr |= mair_index << 0; // Indx bits - 2:0
  attr |= ap << 4; // AP bits - 5:4
  attr |= shareability << 6; // SH bits - 7:6
  attr |= 1 << 8; // AF bit - 8

  return attr;
}
```

De functie upper_attributes is op voorhand al geschreven en daar staat volgende code:

```
uint64_t upper_attributes(bool pxn, bool uxn) {
  uint64_t attr { 0 };
  if (pxn)
    attr |= 1 << 3;
  if (uxn)
    attr |= 1 << 4;
  return attr;
}
```

Aangezien upper attributes begint op bit 51 en PXN op bit 53 staat weet ik nu niet goed of mijn code of de op voorhand gemaakte code van upper_attributes nu een fout heeft. Ik zou veronderstellen dat upper_attributes voor PXN 2 keer een left shift doet van 0b1 aangezien dat ook 2 vanaf bit 51 ligt - wat ook het begin van de upper attributes is. (edited)

bcoppens
  9:58 PM
Daar ga ik morgen eens naar kijken anders :sweat_smile:
9:59
(Ben op vakantie in Zwitserland en ben pas nu terug op de kamer, da's wat uitzonderlijk laat tbh :sweat_smile: Normaal ga ik morgen op een normaler uur terug zijn, dan kan ik er eens op mijn gemak naar kijken :stuck_out_tongue:)

Lennert Franssens
  10:19 PM
Oké geen probleem hoor! Ik ben toch aan het studeren voor de herexamens dus veel haast zit er niet achter :wink: Veel plezier nog op vakantie en tot binnenkort dan wel :slightly_smiling_face:

bcoppens
  10:20 PM
ok :smile: bedankt :smile:
10:21

en jij dan ook veel succes intussen met het studeren :smile:
:heart:
1

bcoppens
  5:25 PM
HMMMMMM, ik ben nu naar die code aan het kijken, en daar is inderdaad iets verwarrend aan... :sweat_smile: Ik denk dat je gelijk hebt: in mijn eigen oplossing staat daar bij block: upper_attr << 52 (en dan kloppen die shifts in upper_attributes ook terug). Maar ik heb geen idee waarom ik daar aan die 52 kom (of aan die 3 en 4), want dat komt idd niet overeen met die figuren die ik zelf gehighlight heb in de opgave... Dus ik denk dat dat idd gewoon fout is in mijn opgave. Misschien dat de andere 2 studenten die die opgave al geprobeerd hebben daar niet op gelet hebben / het mij niet is opgevallen toen ik naar hun code keek, met dat het in dit geval (nog) niet uitmaakt omdat die upper attributes in de huidige code toch op 0 staan...
5:26
Ahja
5:26
URGH
5:26
Ik zie het al :stuck_out_tongue:
5:27
Er is een verschil tussen de layout van een block descriptor (figuur D5-15, p D2566) en een page descriptor, (figuur D5-17, p D2569) :stuck_out_tongue:
5:28
Bij die block is er dus effectief een shift van 52
5:32
En dus er is een subtiel verschil in de page en block upper attributes in die zin dat er 1 bit meer (lager) zit bij de page descriptor tov de block upper descriptor
5:32
Maar mss dat ik dat wel wat moet verduidelijken somehow :sweat_smile:
5:33
Bedankt om me er op te wijzen en me ook even te verwarren :smile:

Lennert Franssens
  10:02 AM
Oké, dat is een duidelijk antwoord :slightly_smiling_face:

bcoppens
  2:17 PM
We lopen een beetje uit
2:17
Ten laatste 15:00 zijn we klaar
2:17
Sorry :disappointed:

Lennert Franssens
  2:18 PM
Geen probleem hoor :slightly_smiling_face:

Lennert Franssens
  5:37 PM
Ik zal ook nog wat verderzoeken. Bedankt voor de hulp al en tot volgende week! :slightly_smiling_face:

bcoppens
  10:42 PM
FYI, 't zal ofwel morgen een call zijn, ofwel doe je het live met enkel Bert :sweat_smile:
10:42
mijn mama heeft net positief getest :disappointed:
10:42
Dus ik ga vooral van thuis werken voorlopig

Lennert Franssens
  8:33 AM
Ja ik zit ook met de ondervraging van compilers om 13u35 dus een call gaat wel maar ik zit ook op jullie bureau dan. Dus het gaat allebei. Laat maar weten wat u het liefst heeft. En hopelijk is ze er niet te ziek van :crossed_fingers:

Lennert Franssens
  8:40 AM
Ahja en buiten de errno die maximaal 16 bit groot is, is ook de seccomp_data.IP (instruction pointer) maar 32 bit groot bij het oproepen. Dat is een kleiner probleem aangezien we daar geen rekening moeten houden met een bepaald interval dat ook op andere gegevens kan wijzen. Maar dus het hele adres vergelijken gaat ook niet in de bpf filter... (edited)

bcoppens
  1:16 PM
Ik ben thuis :wink:

babrath
  1:17 PM
wilt ge joinen of niet? :slightly_smiling_face:

bcoppens

1:18 PM
Dat mag :stuck_out_tongue:

babrath
1:19 PM
teams? :slightly_smiling_face:

bcoppens
1:19 PM
ja :stuck_out_tongue:

Messages with prof. dr. Bart Coppens

Lennert Franssens
9:01 AM
joined Slack

Lennert Franssens
3:49 PM
Dag meneer Coppens, ik ben terug begonnen aan het labo van de raspberry pi. Het is toch moeilijker dan gedacht. Ik weet niet zo goed wat ik moet doen bij de eerste opgave om de paginatabel aan te maken. En hoe een block daarbij gebruikt kan worden. Zou u me kunnen helpen of meer info geven?

bcoppens
4:13 PM
Zeker. Een block wordt gebruikt om meteen een ganse 'blok' (van in dit geval 1GiB) aan consecutive geheugen te vertalen, zonder dat we daarvoor alle individuele page table entries moeten aanmaken. Dus als je wil zeggen dat adres 0 -> 0, adres 4K -> 4K, 8K -> 8K, ... dan is dat veel nodeloze data / entries, en die block zorgt gewoon voor een shortcut om direct te zeggen 0-1GiB mapt op 0-1GiB
4:13
En dus we willen in het allereerste deel gewoon een heel simpele identische mapping opzetten
4:14
die zegt dat alle virtuele adressen tussen 0 en 1 GiB en tussen 3 en 4 GiB mappen op fysieke adressen met exact dezelfde waarde
4:14
dus je kan (en mag) dat doen met gewone page tables
4:14
(wat we in het tweede deel ook gewoon zo gaan doen voor 0..1GiB)
4:14
maar je kan daar ook gewone page tables aanmaken die die identische mapping opzetten
4:15
(ik ga er van uit dat je intussen de lessen van prof De Bosschere bekeken hebt?)

Lennert Franssens
4:16 PM
Ja, inderdaad. Dus in de eerste opgave moet er enkel een block aangemaakt worden als ik het goed begrijp?
4:17
En via welke input weet die dan hoe deze moet mappen (op welke pagina)?
4:21
Ah dat staat eigenlijk al in uw eerste bericht als ik het nu nog eens grondig nalees.
4:21
Ik dacht altijd dat een block gebruikt werd om per pagina te gaan mappen. Maar dat

hoeft dus niet persé?

bcoppens
  4:25 PM
In de eerste opgave map je 2 blocks :slightly_smiling_face: En inderdaad, de mapping staat op voorhand (door de opgave) al vast, en hangt dus niet aan inputs vast :slightly_smiling_face:
4:26
Neen, pagina per pagina mappen gebeurt aan de hand van die page table entries in een table in te vullen. Die block 'omzeilt' dat zogezegd :slightly_smiling_face:

Lennert Franssens
  4:27 PM
Oké, ik denk dat ik het doorheb :slightly_smiling_face:

bcoppens
  4:30 PM
:tada:
4:30
En stel gerust nog vragen he :smile:

Lennert Franssens
  4:31 PM
Bedankt! Dat zal ik zeker doen als er me iets niet duidelijk is :stuck_out_tongue:

Lennert Franssens
  11:40 AM
Dag meneer Coppens, ik ben er nog steeds niet uit hoe ik het moet implementeren... Is het mogelijk om daarover eens te bellen (op msteams)?

bcoppens
  11:41 AM
ja, ik ben nu even bezig en ga daarna eten, past 13u30 voor je?

Lennert Franssens
  11:41 AM
Ja, dat past. Tot dan!

Lennert Franssens
  7:17 PM
Dag meneer Coppens, zou u me in de loop van de week nog eens wat extra uitleg willen geven over de functie table() van het raspberry-pi labo?

bcoppens
  8:08 PM
Sure, woensdag in de loop van de namiddag ergens? :slightly_smiling_face:
8:08
(En je niet te veel vaststaren op die raspberry pi practica :wink:)

Lennert Franssens
  8:13 PM
Lukt het ook eerder op het einde van de week? Woensdagnamiddag ben ik niet thuis. En nene :wink: Ik heb wel een beetje het gevoel dat het wat boven mijn niveau ligt :sweat_smile: Of dat het toch op zijn minst iets is waar ik niet mee gewend ben. Maar het is wel interessant, dus het is leuk om naar de oplossing te zoeken :slightly_smiling_face:

bcoppens
  8:13 PM
Sure, donderdagnamiddag en vrijdagnamiddag zijn ook goed voor mij :slightly_smiling_face:
8:14
En nja, dit is ook meer low-level, dus trek het je niet aan :smile: En als je het interessant vindt, dan is dat een goed teken :stuck_out_tongue:
:heart:
1

Lennert Franssens
  8:19 PM
Ja, donderdagnamiddag past :slightly_smiling_face:

bcoppens
  8:21 PM
Donderdagmiddag 14u? :slightly_smiling_face:

Lennert Franssens
  8:28 PM
Dat is goed. Tot dan! :slightly_smiling_face:

bcoppens
  8:28 PM
Tot dan! :slightly_smiling_face:

Lennert Franssens
  2:46 PM
Dag meneer Coppens, ik heb ondertussen volgende code https://github.ugent.be/lefranss/besturingssystemen-raspberrypi-oplossing/blob/master/mm.cc die nog niet werkt. Het probleem zit nog in de table functie.

bcoppens
  2:48 PM
ik kijk eens :slightly_smiling_face:

bcoppens
  2:56 PM
De page functie lijkt niet echt correct: net zoals je bij table aangeeft dat zo een entry een page entry is (|=0b11), moet je aangeven dat een page entry een valid page entry is. Zoals jij het nu schrijft, gaan de laagste-orde bits 0 zijn, wat aangeeft dat die entry invalid is :wink:
Ik vermoed eigenlijk dat dat het enige probleem is?
2:56
Kan je daar eens naar kijken?

Lennert Franssens
  3:01 PM
Ja, ik heb het veranderd. Ik heb ook in die functie nu het output_address op bit 0 laten starten aangezien ik in de init functie het adres volledig doorgeef. Ik krijg nu enkel nog een Synchrounous Exception.
Screenshot from 2021-09-17 15-01-19.png
Screenshot from 2021-09-17 15-01-19.png

bcoppens
  3:02 PM
En wat denk je dat die exception wil zeggen :smile:
3:02
of wacht :p

Lennert Franssens
  3:06 PM
Het is bij het uitvoeren van read32(0x0), die naar de invalid entry zou moeten wijzen
3:07
Maar ik weet nu niet of het de bedoeling is om die exception dan te krijgen :sweat_smile:

bcoppens
  3:07 PM
Wel :stuck_out_tongue:
3:07
Check eens hoe je code gestructureerd is :stuck_out_tongue:
3:08
x = read(15GiB)
y = read(0)
print(x)
print(y)
3:08
waar gaat het crashen :stuck_out_tongue:

Lennert Franssens
  3:08 PM
Bij y = read(0) waardoor print(x) niet meer uitgevoerd wordt
:sweat_smile::sweat_smile:

bcoppens
  3:08 PM
Indeed :smile:

Lennert Franssens
  3:09 PM
Dan is het eigenlijk juist hoe het nu in elkaar zit? :slightly_smiling_face:

bcoppens
  3:09 PM
Ja :stuck_out_tongue:
3:09
en probeer nu eens de eerste printk ervoor te moven
3:09
zodat je effectief ziet of de 15G read de juiste waarde leest :stuck_out_tongue:
3:10
En wees gerust, ik zou dit soort fout ook maken :sweat_smile:
3:10
alles mooi netjes groeperen :stuck_out_tongue:
:laughing:
1

Lennert Franssens
  3:11 PM
Ja, hij geeft dezelfde waarde terug

bcoppens
  3:12 PM
:tada:
3:13
Mss wel je repo private maken :stuck_out_tongue:

Lennert Franssens
  3:16 PM
Eindelijk :joy: ik ga eerst de labo's van Linux maken en dan kan ik nog eens kijken naar opgave 2 van dit labo nadien :slightly_smiling_face: Achteraf bekeken is het eigenlijk niet zo supermoeilijk maar het is een ander soort van programmeren dan het procedureprogrammeren dat we gewoon zijn in andere vakken. Hier is het meer configuratie. Zoals alles is het begin moeilijk maar ik denk wel dat dit een tof project kan zijn. En zoals u gisteren zei ook om op verder te kunnen werken. Als je zo iets groot kan maken blijf je wel gestimuleerd doorheen het semester om elke week je best te doen om alles tijdig te maken.
3:17
Ja, die zet ik terug op privé. Ik weet ook dat er vaak naar oplossingen wordt gezocht via de ugent github :smiley:

bcoppens
  3:18 PM
Yeah, inderdaad :sweat_smile: Da's helaas vaak je lot als je zo lowest-level dingen wil doen: 50% van de tijd manuals begrijpen, 10% code schrijven, 40% uitzoeken waarom je code niet werkt omwille van een voetnoot die je vergeten lezen was in de manual :sweat_smile:
3:19
Maar jouw thesis zelf bevindt zich op een iets hoger niveau :p
3:19
Dus dat zou wel (famous last words :sweat_smile:) ok moeten zijn :wink:
3:19
En ja, 't is dat :stuck_out_tongue: In een toekomstige wereld waar je dit dan gebruikt als basis voor projectjes is dat natuurlijk minder erg :stuck_out_tongue:
3:20
Maar voor een eerste iteratie van het practicum zou ik toch willen proberen dat ze dan niet al iets ter beschikking hebben... :sweat_smile:

Lennert Franssens
  3:23 PM
Gelukkig haha, dat is toch iets waar ik wat meer met vertrouwd ben :smile: en dat snap ik wel, het zou een beetje zonde zijn van de tijd en het werk dat jullie er al in hebben gestoken om het te maken :slightly_smiling_face:
3:25
Bedankt om me zo snel te helpen en waarschijnlijk is het nu wel dan tot maandag in jullie kantoor! Nog een fijn weekend en tot dan! :slightly_smiling_face:

bcoppens
  3:26 PM
Jij ook nog een fijne vakantie en tot dan! :slightly_smiling_face:

Lennert Franssens
  9:07 PM
Heeft u nog even tijd vandaag?

bcoppens
  9:09 PM
Uuuuuh, depends hoe lang dat moet duren? :sweat_smile:

Lennert Franssens
  9:10 PM
Niet zo heel lang denk ik, het is een klein vraagje... :smile:

bcoppens
  9:10 PM
secondje, ik pak mijn gerief :stuck_out_tongue:

Lennert Franssens
  9:11 PM
Op het eerste zicht lijkt het niet direct te werken om SUD en seccomp te combineren. Ik kon me herinneren dat ik iets over user notifier en trapping had gelezen bij seccomp. En dat is eigenlijk bijna exact wat ik nu met SUD doe. Met behulp van SECCOMP_RET_USER_NOTIF en SECCOMP_RET_TRAP zou ik dus ook in staat moeten zijn om alle syscalls die niet door ptrace bekeken worden, terug te sturen naar userspace. Daar kan ik ze dan, net als ik nu bij SUD doe, opnieuw uitvoeren nadat ik alle registers bekeken heb.
9:11
En oke :slightly_smiling_face:

Lennert Franssens
  3:34 PM
Nog een korte oplijsting van wat ik (denk ik) goed begrepen heb:
glibc gebruiken om syscalls 'te onderscheppen' uit het te monitoren programma en van daar uit te voeren (zelfde syscall, zelfde argumenten)
dat adres uit glibc weten we
we kunnen in de seccomp filter via de instructiepointer te weten komen of de syscall uit onze glibc adresruimte komt
indien deze daar niet uit komt wordt hij rechtstreeks naar de cross monitor gestuurd
als deze daar wel uitkomt, wordt de syscall door de filter gestuurd
De functie van de tweede 'plaats' om de syscall uit te voeren in onze aangepaste

glibc was me nog niet helemaal duidelijk. Er missen misschien ook nog stukjes in de eerste oplijsting. Anders ga ik me met de volgende puntjes al eens bezighouden en dan bellen we nog eens als ik daar mee klaar ben:
glibc gebruiken om syscalls 'te onderscheppen' uit het te monitoren programma en van daar uit te voeren (zelfde syscall, zelfde argumenten)
dat adres uit glibc weten we
we kunnen in de seccomp filter via de instructiepointer te weten komen of de syscall uit onze glibc adresruimte komt

bcoppens
3:38 PM
glibc aanpassen zodat alle syscalls die daarin gebeuren eigenlijk via een apart stukje code doorgestuurd worden
dat stukje code bevat twee syscalls:
3:39
oei te rap op enter geduwd
3:43
dat stukje code bevat dus twee syscalls:
de eerste die de originele syscall volledig doorstuurt naar de kernel, waar die door de seccomp filter passeert. Die seccomp filter checkt: komt het terugkeeradres overeen met die 1e syscall? Zo ja: dan kijken we of het een syscall is die in-process mag afgehandeld worden. Indien ja, laten we seccomp een 'errorcode' teruggeven die overeenkomt met een secret 'key'
Die code in glibc na de syscall vangt die errorcode op.
De code erna doet dan de check op argumenten / buffers. Indien ze gelijk zijn, doet deze code een tweede syscall, die als extra argument die 'errorcode' meegeeft (wat dus enkel gaat werken met syscalls met max 5 argumenten...). De seccomp filter kijkt dan: is het terugkeeradres van deze syscall corresponderend met deze tweede syscall-locatie? Ja? Dan check ik dat extra argument, en check ik terug of het een allowed syscall is. Zo ja, laat ik die gewoon door
In alle andere gevallen (ander terugkeeradres, unallowed syscall, incorrecte sleutel) gaat de syscall volledig naar de CP-MON (die dan het terugkeeradres kan aanpassen naar een locatie in die glibc-code die dan correct kan terugkeren naar de gewone code)
Dunno of dat duidelijker is zo? :smile:
:heart:
1

Lennert Franssens
3:50 PM
Ja, zo is het helemaal duidelijk :smiley: ik heb het allemaal wel gevolgd maar het was wat te veel om alles te onthouden :sweat_smile: zou het erg zijn als ik de volgende calls opneem? Dan kan ik zo'n dingen op mijn gemak nog eens opnieuw bekijken/beluisteren :slightly_smiling_face:

bcoppens

3:50 PM
Voor mij is dat goed :stuck_out_tongue:
3:50
Best ook eens met Bert checken :wink:

Lennert Franssens
3:53 PM
Zal ik doen, of het gewoon dan vragen voor ik het doe :slightly_smiling_face:
3:56
En bedankt om het nog eens uit te schrijven :grinning:

bcoppens
3:57 PM
Geen probleem :smile:
3:57
Graag gedaan :slightly_smiling_face:

Lennert Franssens
9:14 PM
Goeieavond, sorry voor het storen. Kent u iets van kubernetes en hoe je daarop kan deployen?

bcoppens
9:17 PM
Helaas niet :sweat_smile:

Lennert Franssens
9:18 PM
Oké, geen probleem :slightly_smiling_face:

Lennert Franssens
1:13 PM
Heeft u straks even tijd (5 minuutjes) om even te bellen?

bcoppens
1:33 PM
Afhankelijk van hoe snel deze meeting gedaan is :sweat_smile:
1:33
Ik laat je nog iets weten

Lennert Franssens
1:43 PM

Oke, dat is goed :slightly_smiling_face:

bcoppens
  2:17 PM
Ik ben beschikbaar :slightly_smiling_face:

Lennert Franssens
  3:11 PM
Klein vraagje... Moet ik iets zeggen over de replication buffers? Ik zou het zelf niet doen aangezien die werking niks met mijn onderwerp te maken heeft en dat eigenlijk iets is dat inwendig in ReMon wordt gedaan.

bcoppens
  3:12 PM
Dan mag je dat wel weglaten :slightly_smiling_face:

Lennert Franssens
  3:13 PM
Oké :slightly_smiling_face:

Lennert Franssens
  5:40 PM
Dit is de link naar mijn presentatie. In de notes staat ongeveer wat ik ga zeggen. Ik heb na sommige delen toch een slide met wat tekst gezet om de belangrijkste punten in mijn 'conclusie' op te lijsten.
https://ugentbe-my.sharepoint.com/:p:/g/personal/lennert_franssens_ugent_be/EZeS6xHKGsFKlh57OSyOpwcBC00DjI2jeT2jIXvs50SL6w

bcoppens
  5:46 PM
Aha!
5:46
Ik ga nu niet al je notes doornemen wel :wink:
5:47
maar de figuren zien er mooi clean uit alvast, animaties ook
5:47
Ik zou wel proberen 'master/slave' te veranderen in 'leader/follower'
5:47
Ik weet dat we die termen vroeger gebruikten, maar, laten we zeggen, er was voortschrijdend inzicht ;)

Lennert Franssens
  5:48 PM

Oke, dat is iets dat ik direct al zal aanpassen :slightly_smiling_face:

bcoppens
  5:48 PM
ik zag ook toevallig in je slides 'een nieuw design uitgevonden' staan, mss is 'bedacht' beter :slightly_smiling_face:
:+1:
1

5:51
Ik zie hier ook toevallig staan 'Alle functionaliteit die ik ga vertellen zit in de user space.', 't is eerder dat ze volledig aangestuurd wordt vanuit de userspace vermoed ik; de checks van BPF zelf gaan nog in de kernel uitgevoerd worden qua functionaliteit? Mss best voorzichtig daar zijn in het verwoorden?

Lennert Franssens
  5:52 PM
Ahja dat is een goede tip, ik zal het nog eens herbekijken hoe ik het uitleg :slightly_smiling_face:

bcoppens
  5:56 PM
Maar dus wel cleane figuren zo :smile:

Lennert Franssens
  5:56 PM
Dan is het goed zo? :smiley:

bcoppens
  5:58 PM
Wat mij betreft wel :smile:

Lennert Franssens
  5:59 PM
Super! Bedankt om nog eens na te kijken en tot morgen! :slightly_smiling_face:

bcoppens
  6:03 PM
Tot morgen! :slightly_smiling_face:

Lennert Franssens
  5:00 PM

Dag meneer Coppens, ik ga vanavond het hoofdstuk over de technologieverkenning nog eens nalezen op schrijffouten en stuur het morgen door.

bcoppens
5:01 PM
Ok!

Lennert Franssens
5:02 PM
En is er een mogelijkheid voor persoonlijke feedback voor het vak Software Hacking and Protection? Het resultaat was net wat lager dan verwacht...

bcoppens
5:10 PM
Jazeker :slightly_smiling_face: Dan stuur je best wel een mail naar prof De Sutter, want dat examen was opgedeeld in verschillende delen dus dat is ook door verschillende mensen verbeterd die dan uitleg kunnen geven :wink:

Lennert Franssens
5:11 PM
Oké, dat zal ik doen. Bedankt!
5:16
Werken jullie morgen nog of bel ik gewoon maandag eens om het verloop van het volgend semester nog eens te bespreken? :slightly_smiling_face:

bcoppens
5:17 PM
Wij werken morgen nog :stuck_out_tongue: So far heb ik, behalve tussen 15u-16u niets ingepland, en ik denk Bert ook niet, dus zet het gewoon in de groepschat wanneer je wil afspreken :slightly_smiling_face:
:heart:
1

Lennert Franssens
11:17 AM
Ik zit nog met een klein probleempje i.v.m. ptrace. Als ik mijn seccomp filter voor de varianten instel, loopt monitor 1 vast in de staat STATE_WAITING_ATTACH. Ik heb al geprobeerd om de ptrace-opties aan te passen in de attach-functie en waar de andere opties worden ingesteld maar dat lijkt niks te veranderen.
11:18
Oeps, dat was voor het gesprek bedoeld. Ik stuur het daar even opnieuw :)

bcoppens
  11:19 AM
:smile: ok :slightly_smiling_face:

Lennert Franssens
  10:06 AM
Dat laatste examen valt heel slecht... Kan ik daar iets aan doen?
Screenshot from 2022-03-17 10-05-38.png
Screenshot from 2022-03-17 10-05-38.png

bcoppens
  10:12 AM
Ouch
10:12
Ja, ik denk niet dat daar veel aan te doen gaat zijn wel...

Lennert Franssens
  10:14 AM
Oké :sweat_smile::sweat_smile:

bcoppens
  10:19 AM
Je kan het altijd eens aan de FSA vragen, maar ik vrees er voor

Lennert Franssens
  10:21 AM
Welja, ik zal dat eens proberen. Ik heb nog nooit zo'n goed verspreid rooster gehad. En nu komt dat natuurlijk slecht uit dat het zo goed verspreid is... Het is Compilers trouwens dat op 1 juli valt :grimacing:

bcoppens
  10:24 AM
ja ik had het intussen al opgezocht :sweat_smile:
10:24
maar daarom ook dat ik vrees dat er weinig aan te doen valt, veel studenten volgen dat vak, en 't is een plichtvak in de master informatica
:cry:
1

Lennert Franssens

7:07 PM
Is het mogelijk om een adres te reserveren dat niet meer overschreven kan worden tot het terug vrijgegeven wordt (in asm)?

bcoppens
7:09 PM
euh, tenzij je direct begint met ganse pagina's read-only te mappen: niet direct? :sweat_smile:

Lennert Franssens
7:15 PM
Hmn spijtig
7:15
Want nu blijkt dat we enkel 16-bit waarden kunnen doorsturen via errno codes.

bcoppens
7:18 PM
..... aaaaaaaah
7:18
juistja
7:18
alleja
7:19
WEETE

Lennert Franssens
7:19 PM
En we moeten ook zeker zijn dat die waarde groter is dan 0x0083 want alle waarden even groot of kleinder daaraan kunnen ook errno-codes uit het programma zijn. Dus ik zou errno gebruiken om aan te geven dat het om een errno uit de bpf-filter gaat. Maar het adres om van glibc naar ipmon te springen zou ik dan in een gereserveerd register steken :wink:

bcoppens
7:19 PM
laten we beginnen met gewoon 64/16 waarden door te sturen :stuck_out_tongue:
7:19
ahja
7:20
maar ik zou voorlopig gewoon dan meerdere syscalls doen en de errnos daarvan combineren
7:20
dat is slecht voor snelheid
7:20
maar is wel de beste approximatie van het originele idee :stuck_out_tongue:

7:20
dan kunnen we dat later verbeteren :wink:

Lennert Franssens
7:22 PM
Het adres van dat gereserveerd register kan ik via een fake syscall vanuit glibc doorsturen naar de monitor. En op het moment dat via een fake syscall ipmon geregistreerd wordt, kan ik op dat gereserveerd adres de entry naar de ipmon syacall instellen :slightly_smiling_face:

bcoppens
7:23 PM
ja, maar dan riskeert dat te lekken?
7:23
dat kan eh :stuck_out_tongue:

Lennert Franssens
7:23 PM
Ja dat is ook al waar :sweat_smile:

bcoppens
7:23 PM
en hoe garandeer je dat dat register niet door online assembly overschreven wordt :sweat_smile:

Lennert Franssens
7:24 PM
Grote problemen, grote problemen :sweat_smile::sweat_smile:

bcoppens
7:24 PM
:smile:

Lennert Franssens
7:25 PM
Want via de errno code lukt het dus eigenlijk ook niet om de entry mee te geven omdat die waarde ook een niet-entry-adres kan zijn :confused:
7:25
Anders hebben we het morgen eens over tijdens de meeting :slightly_smiling_face:

bcoppens
7:26 PM

da's misschien het best idd :slightly_smiling_face:

Lennert Franssens
  1:28 PM
Goeiemiddag meneer Coppens, heeft u in de late namiddag nog even tijd om iets kleins uit te leggen over de mvee? Ik ben beschikbaar vanaf 15u30. Als het niet past, kan ik morgen ook bellen :slightly_smiling_face:

bcoppens
  1:30 PM
Het kan zijn dat ik straks ga wandelen met mijn ouders, dus ik weet het nog niet :sweat_smile:

Lennert Franssens
  1:31 PM
Helemaal begrijpelijk met dit mooie weer :wink: Ik hoor het dan straks wel :smile:

bcoppens
  1:34 PM
ok :smile:

bcoppens
  2:07 PM
Yep we gaan wandelen :sweat_smile:
2:07
't zal voor morgen zijn vrees ik :sweat_smile:

Lennert Franssens
  2:09 PM
Geen probleem! Veel plezier en tot morgen :wink:
:+1:
1

Lennert Franssens
  5:41 PM
Ik heb zeer raar gedrag geïntroduceerd. Soms werkt ipmon als er meerdere varianten gebruikt worden. Maar soms ook niet :face_with_raised_eyebrow:
Goed nieuws is dat het probleem met RDTSC is opgelost en dat het nu soms met meerdere varianten werkt in plaats van nooit.

bcoppens

5:45 PM
dat klinkt een beetje als een race? :disappointed:

Lennert Franssens
5:45 PM
Ja, of een verkeerde mapping?
5:46
Want het gebeurt bij het registreren van ipmon

bcoppens
6:00 PM
ah, dat kan ook
6:01
is het nog steeds als je de mapping deterministisch dezelfde kiest elke keer? (edited)

Lennert Franssens
6:02 PM
Ik doe elke eerste mapping op base 0x200000 en de volgenden varianten op base+=0x100000
6:02
Dus ipmon komt elke keer op hetzelfde adres bij elke nieuwe uitvoering van de mvee

bcoppens
6:04 PM
da's dan wel raar als het aan de mapping zou liggen. Wat als je alle andere dingen ook hetzelfde mapt?

Lennert Franssens
6:05 PM
Alle andere dingen weet ik natuurlijk niet want die gebeuren niet via die disjoint functie

bcoppens
6:06 PM
right :disappointed:

Lennert Franssens
6:11 PM
Ik ga er nog eens goed over nadenken hoe ik dat kan oplossen. En als dat opgelost is, moet ik ook nog een fout oplossen die optreedt wanneer ik "ls" uitvoer.
6:12
Dan treedt een segfault op

6:12
Het heeft ook iets met die mapping te maken veronderstel ik

bcoppens
6:15 PM
vreemd :disappointed:
6:16
we kunnen er morgen naar kijken met Bert :slightly_smiling_face:

Lennert Franssens
7:47 PM
Ja, dat is goed. Ik zal morgenvroeg in de chat een berichtje sturen om te vragen wanneer het past :slightly_smiling_face:
7:48
Wat ik wel al weet ondertussen, is dat het hierop vastloopt (bij de tweede uitvoering van MVEE_REGISTER_IPMON):

```
if (!interaction::fetch_ip(variants[i].variantpid, ip))
    throw RwRegsFailure(i, "fetch IP-MON registration site");
```

bcoppens
9:32 PM
dat is toch ook maar vreemd, dat klinkt alsof het eigk stiekem eerder al is fout gegaan :disappointed:

Lennert Franssens
12:06 PM
Dag meneer Coppens, weet u hoe ik deze error kan oplossen?

```
/usr/bin/ld:
/home/lennertfranssens/Documents/ReMon-seccomp-bpf/ReMon/ext/ReMon-glibc/build_debug
/libc_pic.os: relocation R_X86_64_PC32 against symbol `__custom_syscall' can not be
used when making a shared object; recompile with -fPIC
/usr/bin/ld: final link failed: Bad value
collect2: error: ld returned 1 exit status
../Makerules:699: recipe for target
'/home/lennertfranssens/Documents/ReMon-seccomp-bpf/ReMon/ext/ReMon-glibc/build_debu
g/libc.so' failed
make[2]: ***
[/home/lennertfranssens/Documents/ReMon-seccomp-bpf/ReMon/ext/ReMon-glibc/build_debu
g/libc.so] Error 1
make[2]: *** Waiting for unfinished jobs....
mv -f
/home/lennertfranssens/Documents/ReMon-seccomp-bpf/ReMon/ext/ReMon-glibc/build_debug
/elf/ld-linux-x86-64.so.2.new
/home/lennertfranssens/Documents/ReMon-seccomp-bpf/ReMon/ext/ReMon-glibc/build_debug
/elf/ld-linux-x86-64.so.2
make[2]: Leaving directory
```

'/home/lennertfranssens/Documents/ReMon-seccomp-bpf/ReMon/ext/ReMon-glibc/elf'
Makefile:470: recipe for target 'elf/subdir_lib' failed
make[1]: *** [elf/subdir_lib] Error 2
make[1]: Leaving directory
'/home/lennertfranssens/Documents/ReMon-seccomp-bpf/ReMon/ext/ReMon-glibc'
Makefile:9: recipe for target 'all' failed
make: *** [all] Error 2
12:07
Ik heb al -fPIC toegevoegd aan CXXFLAGS maar dat heeft niet geholpen.
12:08
Het dit probleem treedt op wanneer ik ReMon-glibc build op een 'native' Ubuntu 18.04. In de docker container werkt het wel. Maar ik wil de benchmarks op een 'native' Ubuntu 18.04 doen dus het zou handig zijn als ik die error kan oplossen :slightly_smiling_face:

bcoppens
1:15 PM
Da's wel vreemd
1:15
Ik heb niet meteen een idee
1:15
Is de docker 18.04 ook, of 20.04?

Lennert Franssens
1:15 PM
Ik heb Bert daarnet ook een berichtje gestuurd en hij zei me dat ik het gewoon op Ubuntu 20.04 moest proberen want nu doe ik het op Ubuntu 18.04.
1:16
En de docker is 20.04

bcoppens
1:20 PM
Yeah dan zou ik de 20.04 proberen native :smile:

Lennert Franssens
1:21 PM
Daar werkte de kernelpatch dan weer niet op :smile: maar ik heb het nog geen tweede keer geprobeerd, dus ga dat nu eens doen :slightly_smiling_face:

bcoppens
1:22 PM
welke kernel is 20.04?

Lennert Franssens

1:22 PM
5.4.0

bcoppens
1:22 PM
https://github.com/ReMon-MVEE/ReMon/blob/master/patches/linux-5.4.0-full-ipmon.patch
die werkt niet?
GitHubGitHub
ReMon/linux-5.4.0-full-ipmon.patch at master · ReMon-MVEE/ReMon
Contribute to ReMon-MVEE/ReMon development by creating an account on GitHub. (36 kB)
https://github.com/ReMon-MVEE/ReMon/blob/master/patches/linux-5.4.0-full-ipmon.patch

Lennert Franssens
1:23 PM
nee, de eerste keer toen ik probeerde toch niet

bcoppens
1:23 PM
Vreemd
1:23
Dan zal je eens moeten zeggen wat er faalt :stuck_out_tongue:

Lennert Franssens
1:27 PM
In IP-MON (de oude versie) wordt die is_ipmon_kernel_compatible() uitgevoerd, en die geeft 0 terug. Maar ik ga het nog eens opnieuw proberen.

bcoppens
1:27 PM
zeker dat je de juiste kernel geboot hebt? :sweat_smile:

Lennert Franssens
1:27 PM
ja als ik cat /proc/version deed, kwam hij mij de juiste versie zeggen (met -ipmon achter)
1:28
Maar om alles uit te sluiten zal ik hem gewoon nog eens opnieuw builden :slightly_smiling_face:

bcoppens
1:29 PM
vreemd :disappointed:

1:29
je kan altijd proberen de nieuwere kernel te booten met de oude ubuntu? :grimacing:

Lennert Franssens
1:33 PM
Kan ik doen. Ik laat iets weten. En als het niet werkt, dan gebruik ik Ubuntu 18.04 met kernel 5.3.0 met daarop de Ubuntu 20.04 docker container.

bcoppens
1:38 PM
:+1:

Lennert Franssens
9:48 PM
Gaat professor De Sutter daar nu vrijdag een vraag over stellen waarom dat met seccomp-bpf sneller is dan met de kernelpatch? Want ik heb daar eerlijk gezegd eigenlijk nog geen antwoord op momenteel & heb nog veel werk voor andere vakken waardoor ik het onmogelijk tegen vrijdag kan uitzoeken :sweat_smile:

bcoppens
9:48 PM
Wel, als hij dat doet, kan je altijd zeggen dat je daar nog aan bezig bent en hoopt dat nog kan verwerken in je finale thesis? :smile:

Lennert Franssens
9:49 PM
Ahja oke haha, dat is kleine geruststelling toch :relieved:

bcoppens
9:50 PM
:partying_face:

Lennert Franssens
10:02 AM
Ik zit nu even te denken over glibc. Stel dat een applicatie inline asm heeft met de syscall-instructie. Gaat die dan ook nog via glibc?

bcoppens
10:03 AM
nope, die heeft dan pech :stuck_out_tongue:
10:03
dat zal dan via ptrace gaan

Lennert Franssens
  10:04 AM
Eigenlijk zou ik in mijn bpf filter dan ook nog moeten checken dat de syscall-instructie dan via onze glibc binnenkomt om die daarna pas doorheen de rest van de filter te sturen en anders inderdaag gewoon te tracen met CP-MON.

bcoppens
  10:04 AM
Inderdaad!
10:04
Dat is zelfs vrij essentieel :stuck_out_tongue:

Lennert Franssens
  10:05 AM
Oke hahha dat heb ik dus helemaal over het hoofd gezien :sweat_smile:
10:05
Allez, nu denk ik er wel aan

bcoppens
  10:05 AM
Want anders kan het zelfs een syscall-gadget zijn die niet eens in het originele programma zo bedoeld was :wink:
10:05
Gelukkig denk je er dan nu aan en niet pas in juli ofzo :stuck_out_tongue:

Lennert Franssens
  10:05 AM
Klopt :joy:
10:06
Ik ga dat wel vermelden in mijn presentatie dat dat iets is dat ik nog voor de veiligheid moet implementeren :wink:

bcoppens
  10:06 AM
Is goed :smile:

Lennert Franssens
  10:06 AM
Want prof. De Sutter zou dat anders wel eens kunnen vragen :joy:

bcoppens

10:07 AM
idd :stuck_out_tongue:

Lennert Franssens
10:07 AM
Gaat hij eigenlijk aanwezig zijn tijdens de presentatie?

bcoppens
10:07 AM
Ik denk het wel

Lennert Franssens
10:07 AM
Oke :slightly_smiling_face:

Lennert Franssens
7:40 PM
Weet u welke instructie ik kan gebruiken in assembler om een kill te veroorzaken?
7:41
Is dat int3?

bcoppens
7:45 PM
ud2 geeft een sigill
7:45
Dat kan ook handig zijn
7:45
En idd, int3

Lennert Franssens
7:45 PM
ud2 zorgt dat de mvee een backtrace genereert?

bcoppens
7:54 PM
euh, ah, daar ben ik niet zeker van, maar dat kan je snel checken (zelfde met int3) op een klein testprogrammaatje?

Lennert Franssens
7:54 PM
Ja, ben het nu even aan het testen :slightly_smiling_face:
7:55

Ben aan de laatste benchmarks bezig om de overhead van een grote key te checken. Maar daarvoor moest ik dat nog implementeren natuurlijk :sweat_smile:

bcoppens
7:56 PM
:partying_face:

Lennert Franssens
8:35 PM
De implementatie is al gelukt. Vannacht laat ik de benchmarks lopen. Dan laat ik morgen iets weten over de resultaten :slightly_smiling_face: Ik weet nog niet of het veel zin heeft om de impact van de groeiende filter te testen, aangezien voor clone's en execve's IP-MON op hetzelfde adres gemapt blijft. Daardoor blijft ook de filter statisch en moet gewoon een manier bedacht worden om te checken dat wij onze filter hebben toegevoegd (via een dubbel fake syscall systeem kan dat gedaan worden. 1.) Laten weten dat de variant de filter gezet heeft na het uitvoeren van de seccomp-functie in IP-MON. 2.) Net voor het instellen van de filter checken of 1. in een ouder-proces van de huidige potentiele clone of execve al uitgevoerd is.)

bcoppens
9:02 PM
Ik ben benieuwd naar de resultaten alvast :smile:

Lennert Franssens
8:43 AM
Ze zitten al over de helft :smile:
:sweat_smile:
1

Lennert Franssens
11:07 AM
Nog 2u nu, maar ik heb al wat resultaten zien voorbij komen en het is goed en slecht
11:08
Het is 10x trager ongeveer dan wanneer we het met 1x de errno-waarde terugsturen, maar wel nog steeds 10x sneller dan de traditionele CP-MON :slightly_smiling_face:
11:08
Straks zal ik het in de groep zetten als ik er een mooi grafiekje van heb :slightly_smiling_face:

Lennert Franssens
11:56 AM
En ik heb ook de implementatie nog wat aangepast zodat we nu ook een random adres voor IP-MON hebben, dat wel constant blijft voor elke variant over execve's en

clone's heen
11:57
Het enige wat ik nu nog moet doen is dat de filter geen 2x geïnstalleerd wordt

bcoppens
  11:59 AM
Dat is wel een beetje jammer van die 10x trager, maar 10x sneller is dan weer wel goed :wink: :smile:

Lennert Franssens
  12:10 PM
Klein probleempje... Er is een figuur te klein voor een niet-digitale lezer. Kan dat kwaad?

bcoppens
  12:12 PM
Hoe bedoel je?

Lennert Franssens
  12:12 PM
image.png
image.png

bcoppens
  12:13 PM
Vergroot daar gewoon de fonts in de boxes :stuck_out_tongue:
12:13
Want voor een digitale lezer is dat ook wel vervelend :sweat_smile:

Lennert Franssens
  12:13 PM
Oke :smile:
12:15
Ik blijf wel met het probleem zitten dat het in totaal 41 bladzijden corpus is :confused: maar ik ben zelf wel tevreden met de inhoud en denk dat alles duidelijk is voor de lezer en alles wat er kan instaan, er ook instaat. Dus het is aan de korte kant, maar ik heb ook gewoon veel tijd moeten steken aan het vinden van oplossingen die opzich niet heel moeilijlk zijn, maar wel moeilijk te vinden waren in het geheel :sweat_smile:

bcoppens

12:16 PM
Sure :slightly_smiling_face:
12:16
Maar is dat 41 bladzijnden waarbij p1 de 1e pagina van de introductie is?
12:16
of 41 bladzijden in de PDF
12:16
met inhoudsopgave enzo er allemaal bij

Lennert Franssens
12:16 PM
1e pagina is de introductie

bcoppens
12:17 PM
ok sure

Lennert Franssens
12:17 PM
Met voorblad, oplijsting van figuren enzo, inhoudsopgave en bijlagen is het rond de 60 denk ik

bcoppens
12:17 PM
dat is idd niet echt lang, maar ook niet zo dramatisch kort :slightly_smiling_face:

Lennert Franssens
12:17 PM
Oke, dat is toch al een beetje een geruststelling dan :smile: ik ben net die laatste figuren nog wat groter aan het maken en dan stuur ik het door

bcoppens
12:18 PM
ok! :smile:

Lennert Franssens
12:18 PM
Ik moet dan juist mijn abstract en samenvatting van 10 lijnen nog schrijven, maar dat doe ik dinsdag en woensdag (want ik heb dinsdagvoormiddag mijn 1e examen :wink:)

bcoppens

12:19 PM
ahja :sweat_smile:
12:19
Veel succes!

Lennert Franssens
12:20 PM
Maar opzich zie ik het wel goedkomen tegen 9 juni hoor :smile: en bedankt!
:tada:
1

Messages with dr. ir. Bert Abrath

Lennert Franssens
  1:00 PM
Ik heb het probleem gevonden. execvp overschrijft natuurlijk de programmacode waardoor de daaropvolgende code niet meer wordt uitgevoerd. De lus met PTRACE_CONT werkt dus wel :slightly_smiling_face:

babrath
  1:01 PM
ahhhhhh
1:01
ja exec is natuurlijk een speciale syscall :smile: (edited)

Lennert Franssens
  12:27 PM
Nog een klein vraagje i.v.m. ptrace en signal handlers. Het is het laatste dat ik wil proberen in die combinatie. Dus in de tracee zeg ik dat er een signal handler is voor SIGSYS signals. Maar de tracer vangt deze signalen op. De signal handler wordt dus nooit uitgevoerd. Kan daar iets tegen gedaan worden of is dat gewoon hoe ptrace werkt? De code van het programma staat hier: https://github.ugent.be/lefranss/thesis/blob/master/broker_demo/main.c#L86. Op lijn 86 zet ik de signal handler voor het SIGSYS signaal; lijn 117 zet ik dat dit kind de tracee is; lijn 134 stuurt een SIGSTOP zodat de tracer vanaf daar alles kan onderscheppen; lijn 152 zegt de tracer dat de tracee gestopt moet worden als er een SECCOMP_RET_TRACE waarde naar de tracer gestuurd wordt.
In de man-page van seccomp staat wel:
Note that a tracer process will not be notified if another
filter returns an action value with a precedence greater
than SECCOMP_RET_TRACE.
en
In decreasing order of precedence, the action values that may be
returned by a seccomp filter are:
SECCOMP_RET_KILL_PROCESS
...
SECCOMP_RET_TRAP
...
SECCOMP_RET_TRACE
(edited)

babrath
  12:35 PM
als er een signal wordt verstuurd naar de tracee zal dit signal altijd eerst passeren bij de tracer
12:36
de tracer heeft wel de keuze om dit signal door te laten of niet (voor de meeste

signals, SIGSYS included)
12:37
als je wil dat het signal doorgelaten wordt, geef het mee als argument aan de PTRACE_CONT die je doet om de tracee te laten continuen

Lennert Franssens
12:41 PM
Maar eigenlijk zorgt dat dan niet voor een snellere uitvoering? Aangezien dat signaal dan toch altijd eerst de tracer moet "passeren"?

babrath
12:42 PM
nee idd, eigk niet :open_mouth:
12:42
voor welk mechanisme is dat?

Lennert Franssens
12:44 PM
SECCOMP_RET_TRAP, en dat had ik dus ook willen gebruiken in plaats van SUD. Of toch op zijn minst proberen wat de mogelijkheden zijn
12:45
Maar ja aangezien het signaal dan toch al aan de tracer zit, kan ik beter de argumenten enzo ineens daar bekijken i.p.v. nog eens door te sturen.

babrath
12:55 PM
hmm
12:55
en RET_ERRNO?

Lennert Franssens
12:59 PM
Eens even proberen. Want met ptrace(PTRACE_CONT, pid, 0, SIGSYS); wordt er nu soms wel naar de signal handler gegaan, maar niet bij de juiste system calls.

babrath
1:01 PM
hm
1:02
dat hangt van u seccomp filter af, zeker? :stuck_out_tongue:

Lennert Franssens
1:03 PM

Normaal wel, maar nu blijkbaar niet want read wordt uitgevoerd in de signal handler hoewel die gewoon uitgevoerd zou moeten worden (SECCOMP_RET_ALLOW in de filter) :sweat_smile:

babrath
  1:08 PM
zeker dat het de signal handler is?
1:08
en hij stopt niet in de tracer omdat het een syscall is?

Lennert Franssens
  1:09 PM
Screenshot from 2021-11-23 13-09-14.png
Screenshot from 2021-11-23 13-09-14.png

1:10
Ja, ik print enkel dat SYSCALL NO ... deel uit vanuit de signal handler. En ik heb enkel bij de close syscall een SECCOMP_RET_TRAP gezet.
1:11
En zoals te zien is op de screenshot doet hij ook nog steeds de close syscall in ptrace en even later in de signal handler.

Lennert Franssens
  1:17 PM
En de tracer vangt ook de syscalls uit de signal handler op.

babrath
  1:32 PM
hm, ik weet aan die output wel niet direct wat door wat uitgeprint wordt (tracer, signal handler, seccomp?)

Lennert Franssens
  1:34 PM
Het is van dit programma: https://github.ugent.be/lefranss/thesis/blob/master/broker_demo/main.c. Het enige probleem dat er dus nog is, is dat syscalls vanuit de signal handler (write syscalls die ik daar uitvoer) ook door de tracer worden opgevangen en dat de sigsys nog steeds eerst in de tracer komt.
1:35
[waitpid status: 0x0007057f]
received a SIGTRAP
syscall number: 1
first argument: 2

```
SYSCALL NO 39
        ARG1 22
        ARG2 2
        ARG3 140728286860784
        ARG4 140447464361630
        ARG5 0
        ARG6 0

--------------------------------
[waitpid status: 0x00001f7f]
received a SIGSYS, but why am I not in the signal_handler?
syscall number: 39
--------------------------------
[waitpid status: 0x0007057f]
received a SIGTRAP
syscall number: 1
first argument: 2
------------------------------------------------
SYSCALL NO 39
        ARG1 140728286860784
        ARG2 -128
        ARG3 140447464331421
        ARG4 0
        ARG5 140728286869896
        ARG6 140728286869896

--------------------------------
[waitpid status: 0x0007057f]
received a SIGTRAP
syscall number: 1
first argument: 1
Hello, World!
--------------------------------
[waitpid status: 0x0007057f]
received a SIGTRAP
syscall number: 0
first argument: 3
--------------------------------
[waitpid status: 0x0007057f]
received a SIGTRAP
syscall number: 1
first argument: 1
Hello, World!
--------------------------------
[waitpid status: 0x0007057f]
received a SIGTRAP
syscall number: 0
first argument: 3
--------------------------------
[waitpid status: 0x00001f7f]
```

```
received a SIGSYS, but why am I not in the signal_handler?
syscall number: 3
--------------------------------
[waitpid status: 0x0007057f]
received a SIGTRAP
syscall number: 1
first argument: 2
------------------------------------------------
SYSCALL NO 3
        ARG1 3
        ARG2 140728286861392
        ARG3 8192
        ARG4 0
        ARG5 0
        ARG6 140728286860320

--------------------------------
[waitpid status: 0x00000000]
CHILD is exiting
1:35
En dit is de output


babrath
  1:36 PM
ja
1:36
dus wat is het resultaat voor write? ALLOW of TRACE?


Lennert Franssens
  1:36 PM
Write heeft TRACE


babrath
  1:37 PM
dat sigsys dan eerst nog in tracer komt lijkt mij jammer genoeg niet zo vreemd
1:37
dan is dat toch ook juist? :stuck_out_tongue:


Lennert Franssens
  1:38 PM
Ja, zeker :smile: maar bestaat er geen manier om syscalls uit de signal handler niet
te tracen? Of is dat nu gewoonweg deel van de tracee?


babrath
  1:39 PM
```

ahh
1:39
wel
1:40
die komen ook bij de seccomp filter terecht
1:40
en daar kan jehopelijk kijken naar een eigenschap om ze te onderscheiden en ze gewoon te ALLOWen ipv te TRACEn ?
1:40
een eigenschap zoals het adres van waar de syscall gebeurt?

Lennert Franssens
1:41 PM
Ahaa dat is inderdaad een goed idee :wink:
1:43
Maar ik denk dat we dan wel tot de conclusie komen dat het gewoon met SECCOMP_RET_TRACE zal werken. En de syscalls die dan SECCOMP_RET_ALLOW hebben, moeten gebruik maken van de ipmon die er nu al is in ReMon.

babrath
1:54 PM
ja, zo klinkt het wel :slightly_smiling_face:

Lennert Franssens
9:42 AM
Goeiemorgen meneer Abrath, ik heb zonet te horen gekregen dat ik een hoogrisicocontact heb gehad en moet nu in quarantaine. Daardoor zal ik niet fysiek aanwezig kunnen zijn in de les die deze namiddag doorgaat. Ik zal de les dan uiteraard wel online volgen.

babrath
10:06 AM
cava
10:06
het labo achteraf gaat ook online gaan

Lennert Franssens
10:30 AM
Oke, dat is goed :slightly_smiling_face:

babrath
10:35 AM
Nog wat verduidelijking voor het labo:
- De communicatie gaat dan via Teams (ik zal online zijn). Zowel tekstueel (simpele

vragen of aankondigingen) als via een call (voor ingewikkeldere vragen en je scherm te delen).
- Qua tijdstip: De starttijd van de labo's is afhankelijk van wanneer de Q&A sessie eindigt. We zullen je dit wel laten weten, via Teams.
- Mochten er verduidelijkingen en algemene aankondigingen zijn tijdens het labo, laten we dit ook via Teams weten.
:+1:
1

Lennert Franssens
10:40 AM
Wordt de Q&A ook gestreamd?

babrath
10:41 AM
Normaal wel
10:42
daar is precies wel nog niet over gecommuniceerd
10:42
als ik prof. De Sutter zie, zal ik het hem eens vragen

Lennert Franssens
10:43 AM
Super! Bedankt en tot vanmiddag :smiley:

babrath
11:19 AM
Het is via Zoom, ik heb er sessies voor toegevoegd op Ufora. Is dat duidelijk te vinden? :sweat_smile:

Lennert Franssens
11:20 AM
Ja, het staat er duidelijk op :slightly_smiling_face:
:+1:
1

babrath
11:21 AM
ok

Lennert Franssens

3:05 PM
Dag meneer Abrath, kan ik het labo van Software Hacking and Protection morgen opnieuw van thuis uit maken?

babrath
3:06 PM
Natuurlijk :slightly_smiling_face:
3:07
het labo zal op dezelfde manier als dat van vorige week georganiseerd worden

Lennert Franssens
3:08 PM
Oke, super! :slightly_smiling_face:

Lennert Franssens
2:45 PM
Zou ik u al even mogen bellen op teams?

babrath
3:53 PM
ah oei, gemist
3:53
probleem opgelost? :stuck_out_tongue:

Lennert Franssens
3:54 PM
Ja hoor, dat was voor die preprocessor macro :wink:

babrath
4:00 PM
ok!

Lennert Franssens
2:23 PM
Hoe kan ik het return address van een instructie in glibc vinden?

Lennert Franssens
12:00 PM
Het enige dat ik nu nog aan het aanpassen ben is IPMON, de rest staat klaar in glibc en remon. Maar nu vraag ik me juist af hoe ik in de code in MVEE_ipmon.cpp de code van de syscall-instructie kan veranderen. Het is dus de bedoeling om vanuit ipmon naar return-adres uit glibc te springen. Dat adres heb ik in MVEE_ipmon.cpp. Maar de

code van de ipmon syscall staat in MVEE_ipmon_syscall.S en ik vind niet hoe ik de code die daar staat kan aanpassen zodat ik het return-adres naar glibc daarin krijg.
12:04
Het zou dus eigenlijk een stuk inline asm code moeten zijn. Maar gaan de FUNCTION_START_EXPORTED en FUNCTION_END macro's uit MVEE_ipmon_syscall.S dan nog werken?

Lennert Franssens
12:31 PM
Of het is ook nog altijd mogelijk om dat adres even op de stack te plaatsen bij de overgang van glibc naar ipmon. Of het adres in een register plaatsen gaat ook zeker werken.

babrath
12:35 PM
Op de stack is misschien het beste, ik zal enr straks eens naar kijken
12:35
Of een korte call?

Lennert Franssens
12:35 PM
Ja gaat ook. Wanneer juist dan?

babrath
12:50 PM
14u30 kort?

Lennert Franssens
12:51 PM
Dat valt nogal moeilijk. Om 16u30 zou beter passen of voor 13u30. Alst dat voor u lukt tenminste :slightly_smiling_face:

babrath
12:55 PM
16u30 dan :slightly_smiling_face:

Lennert Franssens
12:58 PM
Oké, goed. Tot dan! :slightly_smiling_face:

Lennert Franssens
12:24 PM

Een klein vraagje... Ik heb ondertussen onderstaande code voor de glibc custom_syscall. Maar om alles te laten werken (opvangen van geldige errno codes) moet ik checken of de errno die gestuurd wordt een geldige errno is (eigenlijk kijk ik dus of die kleiner of gelijk is aan de maximale waarde van een errno). Als die errno code een geldige is, gaat de call naar ipmon niet door want de waarde van errno is niet het adres van de ipmon syscall. Maar mag ik dat zomaar doen? Ben ik er zeker van dat de ipmon syscall nooit op het adres 0x0083 of lager zal zitten?
12:25
Screenshot from 2022-03-29 12-22-56.png
Screenshot from 2022-03-29 12-22-56.png

12:26
mvee_glibc_syscall_ret_ptr moet ik nog een betere naam geven want dat is nu een gewoon label waar ik het adres nooit meer van moet weten achteraf
12:27
En nu hoop ik ook dat bij het returnen van een errno code vanuit de bpf filter, ook de carry bit wordt gezet zoals bij een normale errno gebeurt
12:30
En nog bijkomend goed nieuws, alles werkt nog als IPMON niet aanstaat. Dat is zeker de bedoeling maar door alle veranderingen had ik daar toch wat schrik voor dat het niet van de eerste keer meer ging werken :smile:

babrath
2:00 PM
ipmon zal nooit op dat adres zitten
2:00
maar ik zal eens zien of het nog properder zou kunnen
2:00
haha goed :stuck_out_tongue:
2:00
gohja, dat label mag gewoon weg denk ik?

babrath
2:12 PM
Ik denk niet dat er een betere manier is.. Maar SECCOMP_RET_DATA wel maar 16 bits kan zijn? Dat is wel teleurstellend :disappointed:

Lennert Franssens
3:24 PM
Dat wil zeggen dat het adres niet doorgestuurd kan worden als errno?

babrath
3:26 PM
niet in zijn volledigheid
3:27

die 16 bits kunnen een deel zijn (met de rest dan fixed), of een key om het adres te decoderen
3:27
redelijk lastig,iig

Lennert Franssens
3:41 PM
Dat is minder goed nieuws... :confused:
3:42
Ik zie net dat de max value van een errno ook maar 16 bits kan zijn

Lennert Franssens
3:56 PM
Ik ga nog eens nadenken of we op een andere manier de entry naar ipmon kunnen doorgeven

babrath
4:28 PM
ahja, dat zal daar aan liggen
4:28
ik vrees van niet

Lennert Franssens
6:31 PM
Nieuw probleem... De code die ik afgelopen namiddag heb doorgestuurd zal dan ook niet meer werken aangezien we niet zeker weten dat de laatste 16 bits van het adres van de ipmon syscall groter is dan 0x0083. Dit is net de laatste stap die nodig is om het te laten werken :sweat_smile:

Lennert Franssens
6:56 PM
Als ik nu een adres "reserveer" in het geheugen vanuit glibc (variabel zodat het per variant verschilt), dat adres doorstuur naar de monitor via een fake syscall en op dat adres het adres van de ipmon syscall (dat ik ook via een fake syscall later in de monitor krijg) bijhoud. Alleen weet ik niet of zo iets kan?

Lennert Franssens
7:01 PM
En dan kan ik nog steeds die errno waarde vergelijken met 0x0083. Vanuit de bpf filter stuur ik een errno waarde uit die groter is dan 0x0083. Enkel de bpf filter zal dus een waarde groter dan 0x0083 genereren, waardoor ik weet dat ipmon actief is (en ingeladen is in het geheugen). Dat wil zeggen dat ik dan in de custom_syscall in glibc enkel naar de "call (RESERVED_ADDRESS)" instructie kan gaan op het moment dat op het adres RESERVED_ADDRESS het adres van de ipmon syscall staat.

7:02
Maar zoals ik zei moet ik dan wel zeker zijn dat ik een adres kan reserveren dat niet tijdens de uitvoering van een variant mag overschreven worden.

babrath
8:09 PM
Hmm
8:10
Dat kan je in principe wel vanuit de mvee forc3ren, waar ipmon wordt geladen

Lennert Franssens
8:13 PM
Ah dat zou goed zijn :slightly_smiling_face: ik moet er alleen nog eens over nadenken hoe ik dat adres (dat in glibc dan wel gekend is) dan in die call instructie krijg

babrath
8:15 PM
Maar wat wil je zeggen met gekend in glibc exact? :p

Lennert Franssens
8:25 PM
Tijdens het starten van glibc kan ik dat adres random kiezen. Dat adres wordt dan via een fake syscall naar de monitor gestuurd en op het moment van de activatie van ipmon ingevuld. Daarna maak ik dat adres read-only vanuit de monitor. Nu, in het begin zei ik al dat ik dat adres random kan kiezen en dat gebeurt in csu/libc-start.c. En eigenlijk zou ik dat adres dat ik daar bijhoud, graag gebruiken in misc/custom_syscall.c in de inline asm...

babrath
8:26 PM
Hmm, het adres zou nog altijd geheim moeten zijn
:confused:
1

Lennert Franssens
4:27 PM
Als ik andere waarden als errno terugstuur (11, 226, 55...) krijg ik ook 26 ik de glibc custom syscall functie. Dus het lijkt alsof die errno-waarde die ik in glibc opvraag (héél toevallig?) hetzelfde is als het adres van de ipmon syscall entry

babrath

4:28 PM
hm
4:28
da zou vreemd zijn?
4:28
doe eens call %rax om de return value van de syscall te krijgen?
4:29
ahh, de entry in de filter? kun je die eens aanpassen?

Lennert Franssens
4:29 PM
dat doe ik nu al
4:29
heb ik ook al gedaan
4:29
en hij blijft 26 terugsturen

babrath
4:29 PM
ah
4:30
doe eens
4:30
een compare van 0x85 op %rax (zonder errno op te halen)
4:30
en dan call %rax als dit groter is dan 0x85

Lennert Franssens
4:31 PM
Ik heb net een align gedaan op 8192 om het adres te verplaatsen en het zit nu in ipmon op 228000, en toch krijg ik in de filter ook nog altijd 26
4:31
Dat zal ik eens doen

babrath
een compare van 0x85 op %rax (zonder errno op te halen)
Direct Message | Apr 19th | View conversation

babrath
4:31 PM
uhu
4:32
ik denk eigk dat de ERRNO_DATA niet terecht gaat komen in glibc's errno, maar in de syscall return value gaat zitten?

Lennert Franssens
  4:35 PM
Dat werkt niet, geeft een segfault
4:36
maar die -1 zit wel in rax want alls ik bij write een errno terugstuur, dan springt hij naar dat ene stukje na die compare in glibc, anders niet
4:39
En het zou toch in errno moeten zitten volgens de beschrijving op kernel.org:
SECCOMP_RET_ERRNO:
Results in the lower 16-bits of the return value being passed to userland as the errno without executing the system call.

Lennert Franssens
  4:39 PM
Het heeft niet geholpen :disappointed: het loop op hetzelfde punt vast, ookal zit ipmon op hetzelfde adres gemapt (edited)

1 reply
2 months agoView thread

babrath
  4:47 PM
wel het is niet hetzelfde punt, de backtrace is anders
4:48
en het is een 'illegal instruction' ipv een segmentation fault
4:52
ik zou debuggen door uit te printen vanuit het enclave entrypoint wanneer je daar komt?

Lennert Franssens
  4:54 PM
Oke :slightly_smiling_face:

Lennert Franssens
  9:06 PM
De shmat die voor problemen zorgt, wordt niet vanuit ipmon_register_thread verstuurd. En er staat altijd waarde -38 als eerste arguments wanneer het foutloopt, wat verdacht veel lijkt op errno code 38 "function not implemented" :face_with_raised_eyebrow:

babrath
  9:31 PM
Ahh? En van waar komt die dan wel?
9:31

Maar ja das waar

Lennert Franssens
9:33 PM
Dat ben ik nog aan het uitzoeken. Ik heb de code in ipmon aangepast zodat ik de MVEE_GET_SHARED_BUFFER apart oproep. Bij de eerste execve staat die voor de shmat, en dus bij de tweede execve niet, wat mij doet vermoeden dat het dus niet van die specifieke shmat komt

babrath
9:37 PM
Je kan de backtrace van het proces uitprinten vanuit de mvee
9:37
Specifiek voor shmat bvb
9:39
Er is daar een functie voor, ik ben de naam kwijt, maar het is iets met log_backtrace ofzo denk ik. Die gaat zeker ook gebruikt worden als de varianten crashen, om dan de backtrace uit te printen

Lennert Franssens
9:40 PM
Oke, dan ga ik daar eens op zoeken. Bedankt! :slightly_smiling_face:

Lennert Franssens
9:21 AM
Oké, turns out dat het toch vanuit ipmon op die ene plaats komt. Ik heb geen idee waarom de monitor die fake syscall voor MVEE_GET_SHARED_BUFFER niet logt. Maar aangezien die een errno -38 teruggeeft, neemt shmat dat als eerste argument, wat niet werkt natuurlijk. Dus ik ga nu bekijken hoe het komt dat ik een errno -38 krijg voor fake syscalls op het moment na een execve

babrath
9:25 AM
Hmm ja
9:25
Kijk eens naar de handler voor fake syscalls
9:26
Misschien is er een speciale regeling voor ipmon, en wordt dat per ongeluk niet gelogd?

Lennert Franssens
9:26 AM
Goed idee :slightly_smiling_face:
9:31

Hij komt na die execve niet meer in call_call_dispatch dus het probleem zit nog voor de fake syscall handler. Ik ga nog wat "terug" in de code

babrath
10:15 AM
Ok

Lennert Franssens
10:32 AM
Nieuwe vaststelling: na het opzetten van de bpf filter in ipmon (dus ook na de switch van syscall_stop voor alles naar afwisselend seccomp_stop/syscall_stop) werken de fake syscalls niet meer. Dus nu even uitzoeken op welk punt ze niet meer herkend worden in de code.

babrath
11:21 AM
dus voor de execve al?
11:21
ahja
11:21
wel, dat zijn stiekem allemaal getpid's
11:21
probeer die eens by default door te sturen naar de tracer?

Lennert Franssens
11:22 AM
alles wordt momenteel naar de tracer doorgestuurd

babrath
11:22 AM
je kan specifieke beslissen adhv het argument dat aan getpid meegestuurd wordt, dat zal in IP-MON ook al wel zo gebeuren denk ik
11:22
ahh?
11:22
en het werkt nog steeds niet?

Lennert Franssens
11:22 AM
nee
11:23
dus het loopt vast op het feit dat de fake syscalls niet meer herkend, en dus ook niet meer uitgevoerd worden

babrath
  11:27 AM
en komen ze nog in de handler? allee, ik zou wat meer debug statements toevoegen :smile:

Lennert Franssens
  11:28 AM
Heb ik al gedaan :smile: maar ze komen daar ook niet meer in
11:28
ze komen zelfs niet meer in de handle_event

babrath
  11:29 AM
hm
11:30
en ook niet in de bpf filter, right?

Lennert Franssens
  11:31 AM
dat weet ik eigenlijk niet

babrath
  11:31 AM
dat zou ik ook eens proberen loggen :slightly_smiling_face:

Lennert Franssens
  12:10 PM
Er komen geen logs in het audit bestand

Lennert Franssens
  12:22 PM
Maar ik ben er wel zeker van dat het door de bpf filter geraakt want ik heb een klein testprogramma geschreven met zulke grote getallen voor een syscall, en die komen door de bpf filter naar een tracer
12:26
langs de andere kant zie ik niet van waar het probleem anders zou moeten komen aangezien er ook geen aanwijzigingen zijn dat er een synchronisatieprobleem is m.b.t. het "dubbel" afhandelen van een systeemaanroep (seccomp_stop/syscall_stop)

babrath
  1:30 PM
er komen geen logs van getpid in het audit bestand? of van geen enkele syscall?

1:30
zie je iets in de log over "vdso"?

babrath
  1:56 PM
hmmm wait a second
1:57
je krijgt elke keer die warning
It seems that you are running a 64-bit kernel with vsyscall set to emulate.
GHUMVEE requires that vsyscall be set to native, so it can intercept calls to
timing related functions such as sys_gettimeofday.
1:57
probeer eerst eens dat te fixen (via de uitleg die er bij staat), misschien lost het dat al op :sweat_smile:

Lennert Franssens
  3:43 PM
Dat heeft het niet verholpen
3:44
En gewoon geen logs in het algemeen over seccomp in het audit bestand, ookal heb ik de flag aangezet, staat audit aan en heb ik aangegeven alle acties van seccomp te willen loggen

babrath
  3:44 PM
dat is vreemd
3:44
ik zou dat eerst proberen oplossen, dat is wel nuttige debug-info :open_mouth:

Lennert Franssens
  4:28 PM
Ja, ik krijg het niet aan

Lennert Franssens
  4:35 PM
Seccomp audit messages van slack komen met overvloed binnen maar van mijn eigen programma's niet

babrath
  4:35 PM
hmm
4:35
dat is bizar

Lennert Franssens
  4:36 PM
Ja, want ik heb het een paar maanden geleden wel aan kunnen zetten

babrath
  4:36 PM
misschien maakt het een verschil dat het vanuit een docker is? not sure
4:36
kan je in de docker zelf naar de audit logs kijken?

Lennert Franssens
  4:36 PM
Ik heb ook al een klein testprogramma op mijn host gedraait en dat komt ook niet in de logs

babrath
  4:37 PM
hm ok

Lennert Franssens
  4:37 PM
en vanuit docker geraak ik niet in dmesg of audit.log
4:37
De moed zakt mij nu toch wat in de schoenen

babrath
  4:38 PM
mja, als je het op de host ook niet krijgt, zal het vanuit de docker wel geen verschil maken

Lennert Franssens
  4:39 PM
nee dat denk ik ook niet
4:39
dat het dan aan docker zou liggen

babrath
  4:40 PM
ok
4:40
heb je morgen tijd om eens te bellen?

Lennert Franssens
  4:40 PM
ja, dat moet lukken
4:41
maakt niet echt uit wanneer, ik ben heel de dag thuis

babrath
  4:41 PM
neem anders even pauze, of schrijf wat, of lees papers intussentijd :wink:
4:41
Call om 11u?

Lennert Franssens
  4:41 PM
Ik was net al eens gaan fietsen om mijn gedachten eens te verzetten :smile:

babrath
neem anders even pauze, of schrijf wat, of lees papers intussentijd :wink:
Direct Message | Apr 26th | View conversation
4:42
Ja, dat is goed :slightly_smiling_face:

babrath
Call om 11u?
Direct Message | Apr 26th | View conversation

babrath
  4:42 PM
Ok, perfect :slightly_smiling_face:
4:42
ja, als je met iets vast zit kan het helpen om gewoon even iets anders te doen
:sweat_smile:

Lennert Franssens
  4:43 PM
Ja, dat is waar :smile:

Lennert Franssens
  8:40 AM
Ik denk dat ik wel al wat verder zit in de "zoektocht" naar waar het misloopt.
Wanneer ik (met seccomp-bpf enabled) niet meer ptrace_cont doe, maar alles via
ptrace_syscall zie ik een fake syscall terug verschijnen na het instellen van de

filter. Het geheel werkt niet meer, maar dat komt omdat we dan niet meer in sync zitten omdat er per syscall 3 keer naar de monitor gegaan wordt en de monitor dat eigenlijk voorziet voor 2 per syscall.

Lennert Franssens
  8:54 AM
Oké, nog verder. Een normale syscall (één die een "echte" handler heeft), gaat met ptrace_syscall en een filter geïnstalleerd door deze drie stappen: entry, seccomp_stop, exit. Maar als ik een fake syscall doe, met nog steeds de filter geïnstalleerd gaat die enkel door deze stappen: entry, exit. Dus het heeft met ptrace_cont te maken.

Lennert Franssens
  9:17 AM
Dus de monitor skipt de fake syscall volledig omdat bij ptrace_cont naar seccomp_stop geluisterd wordt. Dat treedt niet op in de fake syscall en de monitor zal dus bij de volgende syscall terug inpikken (op het seccomp_stop signaal).
9:18
Een eerste "idee" is om 3 states in de monitor te maken. Entry, seccomp_stop en exit. En bij een fake syscall zetten we zelf de state op seccomp_stop om vervolgens naar de exit state te gaan.

Lennert Franssens
  9:40 AM
Fixed
9:41
Ik heb gewoon op het moment van een seccomp_stop nogmaals een resume_all() uitgevoerd zodat er naar exit gesprongen wordt. Heel de implementatie van ipmon_active in de monitor mag verwijderd worden.

Lennert Franssens
  9:46 AM
Maar misschien best wel nog eens een call. Want nu werkt de execve volledig en de fake syscalls. Maar nu werkt het weer niet meer met meerdere varianten :sweat:

babrath
  10:24 AM
ik ga het eens volledig lezen en nadenken he :stuck_out_tongue:
10:29
ik ben al volledig verbaasd dat fake syscalls (i.e., getpid's) niet langs de seccomp filter komen?

Lennert Franssens
  10:31 AM

Ja dat vind ik ook wel raar want in een klein testprogramma werkt het wel met een soort van fake syscall (waar het nummer veel groter is dan syscall_max)

babrath
  10:32 AM
hmm
10:32
wacht, hoe werken die fake syscalls weer?
10:33
ik dacht dat dat dezelfde syscall nr als getpid was, maar het zijn eigk andere nummers?

Lennert Franssens
  10:34 AM
Ja het zijn andere nummers met als base 0x69999999 dacht ik

babrath
  10:34 AM
uh oh
10:34
ja, dat zal het wel verklaren
10:34
de kernel zegt "die ken ik niet", en geeft het niet eens door aan seccomp, presumably
10:35
vervroegde call in 5m anders? :stuck_out_tongue:

Lennert Franssens
  10:42 AM
Ja, ik ben er :slightly_smiling_face:

babrath
  10:43 AM
wacht ik ben even in de kernel src aan het zien
10:43
ik bel wel als ik uitluitsel heb :stuck_out_tongue:

Lennert Franssens
  10:43 AM
oké :smile:
10:43
Ik moet wel zeggen dat dus mijn voorbeeldprogramma wel werkt met seccomp-bpf en ptrace_cont
10:44

dus ik denk niet dat het aan de kernel ligt die zulke grote getallen als syscall number niet zou doorgeven

Lennert Franssens
  12:07 PM
Het geeft een fout op die syscall 158 (normaal aangezien ik enkel syscall(158) doe), maar daarvoor zou er al iets moeten zijn over 0x699...4 en dat staat er niet. Ik stuur het programma mee en de uitvoer van ./mvee -- "./program" (edited)
2 files

program.c
C

MVEE.log
Binary

12:10
IP-MON is niet gebruikt voor de uitvoering hiervan, het programma zet zelf de filter.
12:12
Ik ben voor de zekerheid ook nog eens naar de code van MVEE.cpp gaan kijken en daar wordt geen seccomp-bpf filter meer geïnstalleerd

babrath
  4:23 PM
het programma gaat verkeerd als je het met de MVEE uitvoert?
4:23
of hoe moet ik je begrijpen :sweat_smile:

Lennert Franssens
  4:30 PM
Ja, het geeft hetzelfde gedrag als met ipmon. Alle syscalls worden getracet buiten die met een syacall number dat hoger is dan 400 ofzo

babrath
  4:31 PM
en hetzelfde programma buiten MVEE lukt wel?

Lennert Franssens
  4:41 PM
Ja

Lennert Franssens
  9:38 AM
Starting with Linux 4.8, the PTRACE_EVENT_SECCOMP stop was
       reordered to occur between syscall-entry-stop and syscall-exit-
       stop. Note that seccomp no longer runs (and no
       PTRACE_EVENT_SECCOMP will be reported) if the system call is
       skipped due to PTRACE_SYSEMU.
Dat is het enige dat ik nog vind dat een mogelijke oorzaak kan zijn. Maar ik zie nergens in de monitor dat PTRACE_SYSEMU gebruikt wordt dus normaal zou het daar niet aan kunnen liggen

babrath
  9:41 AM
hmm, lijkt mij onwaarschijnlijk
9:45
is het gemakkelijk voor mij om jou versie van CP-MON met het testprogramma'tje eens te runnen?
9:46
en welke versie van de Linux kernel heb je?

Lennert Franssens
  9:47 AM
Ja en versie 5.13
9:47
Moet ik even toegang geven tot mijn computer? Dat is misschien sneller

babrath
  9:48 AM
nee ik ga het wel op mijn systeem doen :sweat_smile:

Lennert Franssens
  9:48 AM
Oké :smile:
9:48
Ik zal het even terug in elkaar zetten en pushen
9:48
op een aparte branch

babrath
  9:48 AM
ok

Lennert Franssens
  9:50 AM
Ah het zou moeten lukken met de versie die nu op de master branch staat
9:50
https://github.com/lennertfranssens/ReMon

Lennert Franssens
  9:51 AM
program.c
#include <stdio.h>
#include <string.h>
#include <stdlib.h>
#include <errno.h>
#include <stddef.h>
Click to expand inline (50 lines)

1 reply
2 months agoView thread

Lennert Franssens
  9:51 AM
USE_IPMON moet aanstaan voor het maken van de MVEE zelf
9:52
in MVEE.ini moet "use_ipmon": false zijn

babrath
  9:53 AM
welke branch?

Lennert Franssens
  9:53 AM
Het schakelen van PTRACE_SYSCALL voor alles naar een wisselende versie met PTRACE_SYSCALL en PTRACE_CONT, gebeurt automatisch in de syscall handler POSTCALL(seccomp) waar ik de flag ipmon_active op true zet.
9:54
En gewoon master

babrath
  9:54 AM
ok
9:54
en het maakt niet uit welke versie van glibc ik gebruik, neem ik aan?

Lennert Franssens
  9:54 AM
Nee, klopt
9:55
https://github.com/lennertfranssens/ReMon-glibc/tree/wip is degene die ik gebruik maar dat zou normaal niet mogen uitmaken

babrath
  9:59 AM
en waar vind ik het programma'tje dat werkt buiten, maar niet binnen de MVEE?

Lennert Franssens
  10:01 AM
Ik heb uw naam als comment bij het programma geschreven, hier een paar berichten boven

babrath
  10:01 AM
ok

Lennert Franssens
  10:02 AM
Het gaat crashen op die syscall 158, maar daarvoor zou het die 0x699...4 moeten tracen en dat wordt ook niet gedaan
10:02
Bij mij toch niet :smile:

babrath
  10:04 AM
en hoe compile je dat?

Lennert Franssens
  10:04 AM
gewoon gcc program.c

babrath
  10:04 AM
ok
10:08
hmm ik krijg bij het testprogramma'tje gewoon al "when setting seccomp filter: Permission denied"

10:08
enig idee?

Lennert Franssens
  10:09 AM
En het is zeker mijn versie van ReMon?

babrath
  10:09 AM
zonder remon

Lennert Franssens
  10:09 AM
Want dat doet mij denken aan prctl(PR_SET_NO_NEW_PRIVS, 1, 0, 0, 0)

babrath
  10:09 AM
maar het is fixed
10:09
ja
10:10
ik had dat eerst zonder die ,0,0,0 gedaan
10:10
maar blijkbaar moeten die extra argumenten daar bij :sweat_smile:

Lennert Franssens
  10:10 AM
Aah oké :smile:
10:11
2 files

main.c
C

program.c
C

10:11
Dit is mijn testprogramma met met main de tracer (en ook de tracee)
10:11

En dat van daarstraks is het testprogramma om in remon te runnen

babrath
  10:13 AM
ok
10:14
ik ga een tijdje debuggen, ik laat iets weten

Lennert Franssens
  10:14 AM
Oké :slightly_smiling_face:

babrath
  10:16 AM
ik krijg alvast seccomp logging :stuck_out_tongue:

Lennert Franssens
  10:16 AM
In het audit bestand? :open_mouth:

babrath
  10:20 AM
geen idee, ik kijk in dmesg
10:20
sudo dmesg | less
10:20
en dan vanonder kijken

Lennert Franssens
  10:20 AM
heb ik ook gedaan en daar stond niks bij mij

babrath
  10:54 AM
het is wel degelijk een extra BPF filter die roet in het eten gooit
10:54
maar de filter zit niet waar we gedacht hadden! :smile:
10:54
het is in docker :slightly_smiling_face:
:smile:
1

10:55

we voeren de MVEE (en de testprogramma's) uit in docker om het environment beter te controleren, maar voor veiligheidsredenen gebruikt docker zijn eigen seccomp filters
10:56
we kunnen deze uitzetten in het docker_MVEE.sh script
10:56
via een --security-opt
10:56
mijn x11docker command is nu als volgt:
10:56
```
x11docker --hostdisplay --hostipc --gpu --pulseaudio --interactive --home --sudouser
--clipboard --cap-default -- \
        --security-opt seccomp=unconfined --cap-add SYS_PTRACE -p 8088:8080 -ti
$VOLUMES -- \
        $IMAGE bash
```
10:56
ik denk dat dit zou moeten werken
:heart:
1

10:57
kudos aan Bart trouwens, die opmerkte dat docker zijn eigen filters installeerde

Lennert Franssens
  10:58 AM
Dan kan ik eindelijk weer verder, nu hopen dat er niet te veel andere problemen meer opduiken :smile:
10:58
Heel erg bedankt!!!

babrath
  10:59 AM
Geen probleem, hopelijk werkt het meteen :slightly_smiling_face:

Lennert Franssens
  11:03 AM
Ja, het werkt. Zelfs "ls" werkt :stuck_out_tongue_closed_eyes:
11:03
En zelfs met meerdere varianten

babrath
  11:03 AM
nice!
11:04
wat is het volgende op de agenda dan? :slightly_smiling_face:

Lennert Franssens
  11:06 AM
De mapping van ipmon iets meer random maken. En dan nog een scriptje maken dat een json bestand krijgt om de seccomp bpf filter policy voor ipmon in te stellen. Ik denk dat ik dat tegen deze namiddag wel af kan werken. En dan ga ik morgen proberen testen? :slightly_smiling_face:

babrath
  11:07 AM
ok!

Lennert Franssens
  11:07 AM
Zijn jullie morgen beschikbaar om de benchmarks wat uit te leggen?

babrath
  11:09 AM
ja
11:09
10u30?

Lennert Franssens
  11:09 AM
Oké, super! :slightly_smiling_face: Ik laat iets weten als alles klaar is. Misschien dat ik de code dan ook nog wat opkuis.
11:10
Ja, dat is goed. Mag dat online?

babrath
  11:10 AM
in de namiddag hebben we geen tijd, en in de voormiddag kan Bart niet
11:10
ja hoor
11:10
maar ik kan dus vanaf 10u30

Lennert Franssens
  11:10 AM
Ik zet het al in mijn agenda :slightly_smiling_face:

babrath
  11:19 AM
ik denk eigenlijk dat de benchmarks al in de repo zitten, in de map

eurosys2022-artifact
11:19
zie README https://github.com/ReMon-MVEE/ReMon/tree/master/eurosys2022-artifact
GitHubGitHub
ReMon/eurosys2022-artifact at master · ReMon-MVEE/ReMon
Contribute to ReMon-MVEE/ReMon development by creating an account on GitHub. (36 kB)
https://github.com/ReMon-MVEE/ReMon/tree/master/eurosys2022-artifact

11:19
en dan meteen step 2, run experiments for nginx

Lennert Franssens
  11:59 AM
Ah oke dat is goed :slightly_smiling_face:

Lennert Franssens
  6:48 PM
Ik heb de kernelpatch gedaan op een fresh install van ubuntu 20.04, met linux-source-5.4.0. Maar hij blijft dus hangen tijdens het booten. Ik heb wel CONFIG_EFI_STUB moeten aanzetten omdat grub de kernel niet liet booten.
IMG_20220429_184552.jpg
IMG_20220429_184552.jpg

Lennert Franssens
  1:05 PM
Oké, na een heel weekend builden van die kernel met verschillende configs is het eindelijk gelukt om hem op te starten :smile:

babrath
  3:30 PM
Ah oaepi?
3:30
Oei
3:30
Wat heeft het verschil gemaakt?

Lennert Franssens
  3:35 PM
In menuconfig de volgende zaken uitschakelen: Cryptographic API > Certificates for signature checking > Provide system-wide ring of blacklisted keys & Reserve area for inserting a certificate without recompiling & Provide a keyring to which extra trustable keys may be added & Additional X.509 keys for default system keyring = ""

babrath
  3:42 PM
Ahh

Lennert Franssens
  3:48 PM
Ik had dat nog niet gedaan met een "nieuwe" kernel waardoor ik dat dus niet wist :sweat_smile:

babrath
  6:40 PM
Ja ok :sweat_smile:

Lennert Franssens
  10:54 AM
Mag ik ook met de ipmon kernel de seccomp bpf versie benchmarken? (edited)

babrath
  10:56 AM
hmm
10:56
in principe wel
10:56
maar ik zou het toch met de unpatched kernel doen? :stuck_out_tongue:

Lennert Franssens
  10:58 AM
Ja ik was wat aan het twijfelen, waarschijnlijk denken we hetzelfde. Voor de zekerheid toch zonder die patch. Ik ga wel dezelfde versie nog eens builden, zonder patch dan. :wink:

babrath
  11:03 AM
idd
11:03
goh
11:05
welja, eigenlijk op dezelfde manier builden is eigk wel de beste vergelijking, goed punt

Lennert Franssens
  12:02 PM

Als ik ReMon-glibc nu build op een 'native' ubuntu 18.04, krijg ik volgende error:
/usr/bin/ld:
/home/lennertfranssens/Documents/ReMon-seccomp-bpf/ReMon/ext/ReMon-glibc/build_debug
/libc_pic.os: relocation R_X86_64_PC32 against symbol `__custom_syscall' can not be
used when making a shared object; recompile with -fPIC
/usr/bin/ld: final link failed: Bad value
collect2: error: ld returned 1 exit status
../Makerules:699: recipe for target
'/home/lennertfranssens/Documents/ReMon-seccomp-bpf/ReMon/ext/ReMon-glibc/build_debu
g/libc.so' failed
make[2]: ***
[/home/lennertfranssens/Documents/ReMon-seccomp-bpf/ReMon/ext/ReMon-glibc/build_debu
g/libc.so] Error 1
make[2]: *** Waiting for unfinished jobs....
mv -f
/home/lennertfranssens/Documents/ReMon-seccomp-bpf/ReMon/ext/ReMon-glibc/build_debug
/elf/ld-linux-x86-64.so.2.new
/home/lennertfranssens/Documents/ReMon-seccomp-bpf/ReMon/ext/ReMon-glibc/build_debug
/elf/ld-linux-x86-64.so.2
make[2]: Leaving directory
'/home/lennertfranssens/Documents/ReMon-seccomp-bpf/ReMon/ext/ReMon-glibc/elf'
Makefile:470: recipe for target 'elf/subdir_lib' failed
make[1]: *** [elf/subdir_lib] Error 2
make[1]: Leaving directory
'/home/lennertfranssens/Documents/ReMon-seccomp-bpf/ReMon/ext/ReMon-glibc'
Makefile:9: recipe for target 'all' failed
make: *** [all] Error 2
Ik heb -fPIC al toegevoegd aan CXXFLAGS, maar dat helpt niet.

babrath
  12:12 PM
Ahh
12:13
Waarom niet 20.04? (edited)
12:13
Dat lukt toch wel?

Lennert Franssens
  12:13 PM
Omdat IP-MON daar de kernelpatch niet herkende (op een 5.4.0 kernel) en op Ubuntu
18.04 met kernel 5.3.0 wel.

babrath
  12:14 PM
Neem gewoon 20.04 met een oudere kernel? :p

Lennert Franssens
  12:15 PM
De laagste kernel source die ik kon downloaden voor 20.04 was 5.4.0 :smile:
12:15
Of mag ik het ook in een container doen op Ubuntu 18.04?

babrath
  12:16 PM
Sure?

Lennert Franssens
  12:16 PM
Ja

babrath
  12:16 PM
Allee zolang het werkt
12:17
Ik weet niet wat er veranderd is dat glibc build doet galen?

Lennert Franssens
  12:18 PM
Ik ook niet. Op Ubuntu 20.04 werkte het gewoon en nu op 18.04 niet. Ik zal het nog een laatste keer proberen op 20.04 en als het niet lukt, dan doe ik het in een container op 18.04.

Lennert Franssens
  12:25 PM
Een klein vraagje over llvm. Hoe maak je met een IRBuilder een CreateCall waarin een call gedaan wordt naar een functie uit de system v abi?

babrath
  12:33 PM
Uhhh, hoe wil je zeggen? Een system call?

Lennert Franssens
  12:34 PM
Ja ik zou voor het vak compilers de __assert functie moeten oproepen maar ik vind niet hoe ik dat juist moet doen (edited)
12:34
CreateCall heeft als parameter een pointer naar de functie nodig maar eigenlijk wil ik gewoon een string meegeven

babrath
  12:36 PM
Er zijn andere functies om een pointer naar een functie te krijgen
12:36
Ik denk getorinsertfunction of zoiets
12:37
Dooorzoek gwn de llvm codebase naar voorbeelden van het gebruik van createcall

Lennert Franssens
  12:37 PM
Oke, bedankt! :slightly_smiling_face:

Lennert Franssens
  9:28 AM
Grote problemen, grote problemen... Ik heb alle benchmarkresultaten kunnen doen buiten die van ipmon met seccomp-bpf waarbij getpid in ipmon wordt uitgevoerd. Ik heb de indruk dat ook die syscall niet juist (meer) wordt afgehandeld door ipmon. Ik weet wel dat write een is die wel zeker werkt, maar ja...
9:29
Doe ik ze dan gewoon met write of probeer ik getpid werkend te krijgen?

babrath
  9:31 AM
Wat is het probleem met getpid?
9:31
Het is een kleine benchmark, dus als het niet werkt, zou je het deze keer wel sneller moeten kunnen vinden, lijkt mij

Lennert Franssens
  9:32 AM
Ik krijg geen logs meer van getpid in het log-bestand maar als output op mijn scherm krijg ik dit:
Opening MVEE Monitor Log @ ./Logs/MVEE.log
Opening log for non-instrumented instructions @ ./Logs/non-instrumented.csv
MONITOR[0] - WARNING: Executing non-diversified variants: /usr/bin/bash -- bash -c ../../../seccomp-bpf-benchmarks-artifact/benchmarks/microbenchmark/getpid
MONITOR[0] - WARNING: The variants are loading a binary with non-instrumented atomic operations.
MONITOR[0] - WARNING: Binary name: /home/lennertfranssens/Documents/ReMon-new/ReMon/IP-MON/libipmon.so
MONITOR[0] - WARNING: If this is a multi-threaded program, you will probably see divergences because of this.
MONITOR[0] - WARNING: Please refer to our EuroSys 2017 paper for more details:
MONITOR[0] - WARNING: Taming Parallelism in a Multi-Variant Execution Environment
MONITOR[0] - WARNING: Stijn Volckaert, Bart Coppens, Bjorn De Sutter, Koen De

Bosschere, Per Larsen, and Michael Franz.
MONITOR[0] - WARNING: In 12th European Conference on Computer Systems
(EuroSys'17). ACM, 2017.
MONITOR[0] - WARNING:
MONITOR[0] - WARNING: Executing non-diversified variants:
/home/lennertfranssens/Documents/ReMon-new/ReMon/seccomp-bpf-benchmarks-artifact/ben
chmarks/microbenchmark/getpid --
../../../seccomp-bpf-benchmarks-artifact/benchmarks/microbenchmark/getpid
MONITOR[0] - WARNING: The variants are loading a binary with non-instrumented atomic
operations.
MONITOR[0] - WARNING: Binary name:
/home/lennertfranssens/Documents/ReMon-new/ReMon/IP-MON/libipmon.so
MONITOR[0] - WARNING: If this is a multi-threaded program, you will probably see
divergences because of this.
MONITOR[0] - WARNING: Please refer to our EuroSys 2017 paper for more details:
MONITOR[0] - WARNING: Taming Parallelism in a Multi-Variant Execution Environment
MONITOR[0] - WARNING: Stijn Volckaert, Bart Coppens, Bjorn De Sutter, Koen De
Bosschere, Per Larsen, and Michael Franz.
MONITOR[0] - WARNING: In 12th European Conference on Computer Systems
(EuroSys'17). ACM, 2017.
MONITOR[0] - WARNING:
MONITOR[0] - WARNING: Warning: SIGSEGV in variant 1 (PID: 1349145)
MONITOR[0] - WARNING: IP: 00000000002431f0, Address: 0000000000000008, Code:
SEGV_MAPERR (1), Errno: 0
MONITOR[0] - WARNING: DWARF: couldn't map EIP 00000000000212af to a known region for
variant: 1 (pid: 1349145)
MONITOR[0] - WARNING: DWARF: couldn't map EIP 00000000000212af to a known region for
variant: 1 (pid: 1349145)
MONITOR[0] - WARNING: Warning: SIGSEGV in variant 0 (PID: 1349144)
MONITOR[0] - WARNING: IP: 00000000002231f0, Address: 0000000000000008, Code:
SEGV_MAPERR (1), Errno: 0
MONITOR[0] - WARNING: DWARF: couldn't map EIP 00000000000212af to a known region for
variant: 0 (pid: 1349144)
MONITOR[0] - WARNING: DWARF: couldn't map EIP 00000000000212af to a known region for
variant: 0 (pid: 1349144)
MONITOR[-1] - WARNING: signalling monitor 0 for shutdown - monitor state is:
STATE_NORMAL
all monitors terminated
Program terminated after: 0.437940 seconds
9:33
Het komt wel uit ipmon die fout (want de IP staat in het bereik waar ipmon gemapt
zit)

babrath
  9:33 AM
is dat met benchmark mode aan?

Lennert Franssens

9:34 AM
Met benchmark mode uit omdat de benchmarks voor dat geval veel trager waren dan alle andere gevallen

babrath
9:34 AM
..?
9:34
de benchmarks waren trager met benchmark mode uit?

Lennert Franssens
9:35 AM
En omdat mij dat vreemd leek heb ik met debug mode gebuild en dan ook dat programma uitgevoerd
9:35
Nene
9:35
Dus ik heb benchmarks gedaan in 4 gevallen
9:36
En die waren allemaal 8000/5 ns ongeveer (buiten de originele ipmon), maar in het geval van de nieuwe ipmon met getpid allowed was het veel trager dan die 8000/5 ns. Dus dat leek mij raar en ik ben gaan kijken of getpid dan wel werkt door benchmark mode uit te zetten, en blijkbaar werkt het dus niet.

babrath
9:38 AM
en wat voor stack trace krijg je in de logs?

Lennert Franssens
9:39 AM
Binary

MVEE.log
Binary

9:43
Dus ik denk dat het aan de ipmon_enclave ligt.

babrath
9:45 AM
welja, ergens vanuit de ipmon_enclave probeer je adres 0x8 te accessen
9:45
en aangezien dat een adres op de NULL page is, lijkt mij dat weer een NULL pointer dereference

9:46
(wel een pointer naar een struct die NULL Is, en je probeert en veld binnen die struct te lezen)
9:46
ik kzou print statements binnenin de enclave toevoegen? van bepaalde pointers die vermoedelijk het probleem kunnen zijn?

Lennert Franssens
  9:46 AM
Oke :slightly_smiling_face:

babrath
  9:52 AM
je kan ook altijd proberen ipmon met debuginfo te builden, dan gaat de stacktrace ook meer info geven (sourc eline numbers)

Lennert Franssens
  9:53 AM
Dat gaat inderdaad veel sneller gaan :smile:
9:56
if (RB->have_pending_signals & 2) {
9:56
Weer de RB

babrath
  9:58 AM
aha
9:58
en waarom deze keer? :stuck_out_tongue:
9:59
is het ook kapot met 1 variant?

Lennert Franssens
  10:00 AM
Ik ga snel eens proberen
10:05
Dan is het ook kapot

babrath
  10:11 AM
ok
10:12
ahja ok
10:12

maar debug maar verder met 1 variant :stuck_out_tongue:

Lennert Franssens
  10:13 AM
Dat ga ik doen :slightly_smiling_face:

Lennert Franssens
  10:55 AM
Het is heel moeilijk te debuggen want op print statements blijft hij "hangen". Elke keer dat ik printf uitvoer, wordt de enclave terug opgeroepen (voert printf achterliggend ook een getpid uit?)

babrath
  10:56 AM
hoezo hangen? wat zegt de log?
10:56
printf doet geen getpid, enkel write normaal gezien

Lennert Franssens
  10:56 AM
die stopt op een fout met malloc
10:56
Ah maar printf doet achterliggend wel een write
10:58
Ah nee
10:58
toch niet

babrath
  11:01 AM
wat niet? :sweat_smile:

Lennert Franssens
  11:01 AM
Dat printf een write doet

babrath
  11:01 AM
doet wel of niet een write?

Lennert Franssens
  11:02 AM

Ik ben het even aan het testen
11:03
Wel een write

babrath
11:03 AM
wel, normaal gezien, is printf een buffered write
11:04
je print naar het scherm (via een write syscall), maar dit wordt gebuffered (de writes worden opgespaard en gebundeld) in een buffer aangemaakt via malloc, TENZIJ er een newline is, dan wordt de write niet gebufferd

Lennert Franssens
11:04 AM
printf zorgt voor een write in strace. Dus aangezien write een getpid uitvoert blijft het dus loopen
11:04
En ja ik doe een newline dus het is een write

babrath
11:04 AM
aangezien write een getpid uitvoert?

Lennert Franssens
11:05 AM
Aaah zo
11:05
Wacht ja ik heb het helemaal verkeerd begrepen

babrath
11:08 AM
maar nu is het duidelijk? geeft dat andere inzichten? :sweat_smile:

Lennert Franssens
11:09 AM
Een beetje al :smile:

Lennert Franssens
11:16 AM
Opening MVEE Monitor Log @ ./Logs/MVEE.log
Opening log for non-instrumented instructions @ ./Logs/non-instrumented.csv
MONITOR[0] - WARNING: Executing non-diversified variants: /usr/bin/bash -- bash -c ../../../seccomp-bpf-benchmarks-artifact/benchmarks/microbenchmark/getpid

```
MONITOR[0] - WARNING: The variants are loading a binary with non-instrumented atomic
operations.
MONITOR[0] - WARNING: Binary name:
/home/lennertfranssens/Documents/ReMon-new/ReMon/IP-MON/libipmon.so
MONITOR[0] - WARNING: If this is a multi-threaded program, you will probably see
divergences because of this.
MONITOR[0] - WARNING: Please refer to our EuroSys 2017 paper for more details:
MONITOR[0] - WARNING:   Taming Parallelism in a Multi-Variant Execution Environment
MONITOR[0] - WARNING:   Stijn Volckaert, Bart Coppens, Bjorn De Sutter, Koen De
Bosschere, Per Larsen, and Michael Franz.
MONITOR[0] - WARNING:   In 12th European Conference on Computer Systems
(EuroSys'17). ACM, 2017.
MONITOR[0] - WARNING:
INFO: Init IP-MON
INFO: Registering IP-MON
MONITOR[0] - WARNING: Executing non-diversified variants:
/home/lennertfranssens/Documents/ReMon-new/ReMon/seccomp-bpf-benchmarks-artifact/ben
chmarks/microbenchmark/getpid --
../../../seccomp-bpf-benchmarks-artifact/benchmarks/microbenchmark/getpid
MONITOR[0] - WARNING: The variants are loading a binary with non-instrumented atomic
operations.
MONITOR[0] - WARNING: Binary name:
/home/lennertfranssens/Documents/ReMon-new/ReMon/IP-MON/libipmon.so
MONITOR[0] - WARNING: If this is a multi-threaded program, you will probably see
divergences because of this.
MONITOR[0] - WARNING: Please refer to our EuroSys 2017 paper for more details:
MONITOR[0] - WARNING:   Taming Parallelism in a Multi-Variant Execution Environment
MONITOR[0] - WARNING:   Stijn Volckaert, Bart Coppens, Bjorn De Sutter, Koen De
Bosschere, Per Larsen, and Michael Franz.
MONITOR[0] - WARNING:   In 12th European Conference on Computer Systems
(EuroSys'17). ACM, 2017.
MONITOR[0] - WARNING:
MONITOR[0] - WARNING: Warning: SIGSEGV in variant 0 (PID: 1370615)
MONITOR[0] - WARNING: IP: 00007f413239c5e6, Address: 00007ffdf07feff8, Code:
SEGV_MAPERR (1), Errno: 0
MONITOR[0] - WARNING: DWARF: couldn't map EIP 00000000000212af to a known region for
variant: 0 (pid: 1370615)
MONITOR[0] - WARNING: DWARF: couldn't map EIP 00000000000212af to a known region for
variant: 0 (pid: 1370615)
all monitors terminated
Program terminated after: 0.705886 seconds
11:17
Hij komt geen tweede keer in init() of ipmon_register_thread()


babrath
  11:17 AM
terwijl er wel een execve is gebeurd?
```

Lennert Franssens
  11:19 AM
Die staat niet in de logs

babrath
  11:19 AM
ah?

Lennert Franssens
  11:19 AM
Maar waarom komt hij dan 2x die warnings zeggen?
11:19
Hij laadt wel 2x libipmon in dan toch?

babrath
  11:20 AM
wel, het lijkt dat er weleenexecve is gebeurd, want de getpid-binary wordt uitgevoerd
11:20
idd
11:20
dus libipmon wordt ingeladen, de 2e keer, maar ie komt nog niet tot het registreren?

Lennert Franssens
  11:20 AM
neen, inderdaad
11:20
Dus ik vermoed dat voor init() ergens een getpid wordt opgeroepen

babrath
  11:21 AM
aha

Lennert Franssens
  11:21 AM
En die springt dan naar zijn eigen enclave al (door het overerven van de vorige bpf filter) maar de thread_local RB variabele is nog niet ingeladen

babrath
  11:21 AM
wel, das problematisch

Lennert Franssens
  11:22 AM
:neutral_face:

babrath
  11:22 AM
second hoor
11:22
dus dat gaat elke keer gebeuren dat je getpid via IP-MON wil laten afhandelen, right?

Lennert Franssens
  11:22 AM
ja
11:23
buiten het geval dat er geen execve's meer komen na het instellen van de eerste filter

babrath
  11:28 AM
en van waar komt de getpid in IP-MON?
11:29
wel, dat van die execve is vreemd, maar misschien een logging artifact?
11:29
de getpid binary begint uit te voeren lijkt mij, dus dan moet de execve wel degelijk gebeurd zijn

Lennert Franssens
  11:31 AM
Ik zal de log nog eens doorsturen voor de zekerheid
11:32
Binary

MVEE.log
Binary

babrath
  11:34 AM
ik zie hem toch gebeuren op lijn 1217?

Lennert Franssens
  11:36 AM

Dan ligt het niet aan die execve?

babrath
  11:37 AM
nee hoor, er is niets mis met execve, er gebeuren er genoeg en op de juiste manier
11:37
libipmon wordt ook mooi twee keer ingeladen
11:37
gewoon niet meer geinitialiseerd nadat ie de 2e keer wordtd ingeladen

Lennert Franssens
  11:38 AM
En is dat normaal of abnormaal gedrag?

babrath
  11:38 AM
wel, als ipmon niet geïnitaliseerd wordt blijft RB NULL
11:38
dus dat zou ik abnormaal noemen?
11:38
de vraag lijkt mij waarom initialisatie niet gebeurt
11:39
en daarvoor zou ik naar de code kijken: waar gebeurt dit normaal, wat zou er anders kunnen zijn nu..

Lennert Franssens
  11:39 AM
Het is dus enkel als ik getpid allow in de bpf filter, met write bijvoorbeeld heb ik dat probleem niet

babrath
  11:39 AM
en wat zegt je logging in IP-MON?
11:40
kom je bij getpid in IP-MON op een moment dat je nog niet mag? terwijl dat bij write niet zo is?

Lennert Franssens
  11:57 AM
Ja ik zie het niet meer eigenlijk... Dit is een log van een filter met write allowed
Binary

MVEE.log
Binary

babrath
  12:05 PM
allow of trace?

Lennert Franssens
  12:06 PM
Write is allow in die log

babrath
  12:06 PM
ah, dus dan wordt die niet gemonitored?

Lennert Franssens
  12:06 PM
Klopt
12:06
Moet ik ook nog een log sturen met alles op trace?

babrath
  12:06 PM
nee nee, wacht, en waarom doe je dan allow?

Lennert Franssens
  12:07 PM
Om in de enclave te kunnen komen

babrath
  12:07 PM
ik dacht dat dat via errno was?

Lennert Franssens
  12:07 PM
Ja
12:07
En die gaat naar de enclave_entry en die roept ipmon_enclave op

babrath
  12:07 PM

dus is het allow, of ret_errno?

Lennert Franssens
12:08 PM
Eerst ret_errno en dan allow
12:08
2-traps

babrath
12:08 PM
right, en dat werkt?

Lennert Franssens
12:08 PM
Ja

babrath
12:08 PM
maar getpid met ret_errno niet?

Lennert Franssens
12:09 PM
En als ik het verander naar getpid om te ret_errno en allow'en, dan werkt het dus niet

babrath
12:09 PM
omdat dat gebeurt VOOR de initialisatie van libipmon, right?

Lennert Franssens
12:09 PM
Ja

babrath
12:11 PM
en wanneer gebeurt die initialisatie?
12:14
en als je de MVEE log voor getpid vergelijkt met die van write
12:14
hoe verschillen ze?

Lennert Franssens
  12:16 PM
Ik ga nog eens goed kijken

babrath
  12:17 PM
je kan ze ook diffen met vimdiff, of andere textuele diff tools

Lennert Franssens
  12:17 PM
(Klein vraagje tussendoor: weet u het antwoord al? :smile:)

babrath
  12:17 PM
nee
12:17
het lijkt mij dat getpid gebeurt (om een of andere reden) vanuit IP-MON, voor dat IP-MON geïnitaliseerd is
12:18
misshien moet er elke keer bij het binnenkomen van IP-MON enclave ge-checked worden of er al geïnitialiseerd is, en indien niet, dat nogmaals doen
12:18
langs de andere kant, waarom zou dit probleem nu dan pas voorkomen
12:19
maar door meer logging in IP-MON toe te voegen moet je wel snel kunnen zien welke code er wanneer uitgevoerd wordt

Lennert Franssens
  12:20 PM
Het probleem is een beetje dat ik geen printf's kan toevoegen in de enclave wanneer ik write allow omdat hij dan in een lus terechtkomt waarbij de enclave zichzelf blijft oproepen
12:21
Dus dan ga ik het doen met getpid op allow

babrath
  12:22 PM
en write trace'n?
12:22
gaat dat niet?

Lennert Franssens
  12:22 PM
ja dat gaat inderdaad wel

Lennert Franssens
  12:30 PM
Hij doet geen SYS_GETPID net voor MVEE_REGISTER_IPMON
Binary

MVEE.log
Binary

12:30
toch zeker niet de tweede keer

babrath
  12:33 PM
vreemd
12:35
dus hier wordt IPMON wel twee keer geïnitialiseerd?

Lennert Franssens
  12:35 PM
Ja
12:35
MVEE_REGISTER_IPMON geeft dat aan

babrath
  12:35 PM
maar dit is niet met getpid

Lennert Franssens
  1:04 PM
Neen
1:04
Dus ik heb geen idee meer waarom het misloopt op dat moment

babrath
  1:05 PM
wat loopt er mis? ik volg je uitleg even niet :sweat_smile:
1:05
allee, het lijkt mij dat die write werkt?

Lennert Franssens
  1:05 PM

Ja dat sowieso

babrath
1:05 PM
en getpid niet?

Lennert Franssens
1:06 PM
Ja dat bedoel ik dus

babrath
1:06 PM
kan je loggen elke keer je in ipmon_enclave komt, met onder meer welke syscall nr?

Lennert Franssens
1:09 PM
Ja

babrath
1:10 PM
welk, dat lijkt me alvast handig

Lennert Franssens
1:15 PM
Er komt geen printf uit die enclave functie, ookal staat die er wel. printf("INFO: In enclave, syscall_no = %lu\n", syscall_no);
Binary

MVEE.log
Binary

Lennert Franssens
1:20 PM
Hij crasht eigenlijk o phet moment dat hij die printf zou moeten uitvoeren

babrath
1:23 PM
en waar crasht ie? op welke system call of wat?

Lennert Franssens
  1:24 PM
dat kan ik niet zien aangezien hij printf in de enclave niet uitvoert

babrath
  1:24 PM
hm
1:24
er komt nooit een printf uit die enclave?
1:24
ook als write allow'ed is?

Lennert Franssens
  1:24 PM
neen

babrath
  1:24 PM
uhh
1:25
trace'd, wil ik zeggen

Lennert Franssens
  1:25 PM
inderdaad
1:25
write staat op trace
1:25
getpid op allow

babrath
  1:25 PM
en als je hier getpid naar trace doet?

Lennert Franssens
  1:25 PM
dan komt er niks meer inde enclave

babrath
  1:26 PM
hmm, ja
1:26
ok

1:26
en wat is de stack trace die uit de MVEE log komt?

Lennert Franssens
  1:26 PM
Wat ook wel raar is, is dat er superveel van deze in de log staan:
0.389787 - MONITOR[0] - Variant:0 [PID:1382665] - SYS_FSTAT(1, 0x00007ffe3cedef60)
0.389788 - MONITOR[0] - SYS_FSTAT - >>> Dispatch as SYNCED NORMAL
0.389795 - MONITOR[0] - Variant:0 [PID:1382665] - SYS_FSTAT64 return
0.389796 - MONITOR[0] - File type: character device
0.389796 - MONITOR[0] - I-node number: 4
0.389797 - MONITOR[0] - Mode: 20620 (octal)
0.389797 - MONITOR[0] - Link count: 1
0.389798 - MONITOR[0] - Ownership: UID=1000   GID=5
0.389799 - MONITOR[0] - Preferred I/O block size: 1024 bytes
0.389799 - MONITOR[0] - File size: 0 bytes
0.389800 - MONITOR[0] - Blocks allocated: 0
0.389801 - MONITOR[0] - Last status change: Tue May 3 23:10:27 2022
0.389802 - MONITOR[0] - Last file access: Wed May 4 13:13:04 2022
0.389803 - MONITOR[0] - Last file modification: Wed May 4 13:13:04 2022
0.389809 - MONITOR[0] - Variant:0 [PID:1382665] - SYS_BRK(0x0000000000000000)
0.389810 - MONITOR[0] - SYS_BRK - >>> Dispatch as SYNCED NORMAL
0.389817 - MONITOR[0] - Variant:0 [PID:1382665] - SYS_BRK return: -563 (Unknown
error 563)
0.389818 - MONITOR[0] - Variant:0 [PID:1382665] - SYS_BRK(0x0000000000000000) return
= 0xfffffffffffffdcd
1:27
0.414355 - MONITOR[0] - Variant:0 [PID:1382665] - variant crashed - trapping ins:
00000000000275e6: mvee_atomic_preop_internal at ??:?
(/home/lennertfranssens/Documents/ReMon-new/ReMon/patched_binaries/libc/amd64/libc-2
.31.so)
0.414357 - MONITOR[0] - Variant:0 [PID:1382665] - Received signal SIGSEGV (11)
0.414359 - MONITOR[0] - Variant:0 [PID:1382665] - signal arrived while variant was
executing ins: 00000000000275e6: mvee_atomic_preop_internal at ??:?
(/home/lennertfranssens/Documents/ReMon-new/ReMon/patched_binaries/libc/amd64/libc-2
.31.so)
0.414361 - MONITOR[0] - Variant:0 [PID:1382665] - ret is currently: 0
0.414369 - MONITOR[0] - WARNING: Warning: SIGSEGV in variant 0 (PID: 1382665)
0.414372 - MONITOR[0] - WARNING: IP: 00007f591d1a55e6, Address: 00007ffe3cd4eff8,
Code: SEGV_MAPERR (1), Errno: 0
0.414373 - MONITOR[0] - Variant:0 [PID:1382665] - ================================
0.414374 - MONITOR[0] - Variant:0 [PID:1382665] - generating local backtrace for
variant
0.414375 - MONITOR[0] - Variant:0 [PID:1382665] - > variant is currently suspended
0.414377 - MONITOR[0] - pid: 1382665 - 000: 00000000000275e6:
mvee_atomic_preop_internal at ??:?
(/home/lennertfranssens/Documents/ReMon-new/ReMon/patched_binaries/libc/amd64/libc-2
.31.so)
0.414491 - MONITOR[0] - pid: 1382665 - 001: 000000000009049a: _int_malloc at

malloc.c:?
(/home/lennertfranssens/Documents/ReMon-new/ReMon/patched_binaries/libc/amd64/libc-2
.31.so)
0.414566 - MONITOR[0] - pid: 1382665 - 002: 0000000000090a6a: tcache_init.part.0 at
malloc.c:?
(/home/lennertfranssens/Documents/ReMon-new/ReMon/patched_binaries/libc/amd64/libc-2
.31.so)
0.414669 - MONITOR[0] - pid: 1382665 - 003: 0000000000091b1e: __GI___libc_malloc at
:?
(/home/lennertfranssens/Documents/ReMon-new/ReMon/patched_binaries/libc/amd64/libc-2
.31.so)
0.414767 - MONITOR[0] - pid: 1382665 - 004: 000000000007a05d:
__GI__IO_file_doallocate at :?
(/home/lennertfranssens/Documents/ReMon-new/ReMon/patched_binaries/libc/amd64/libc-2
.31.so)
0.414867 - MONITOR[0] - pid: 1382665 - 005: 0000000000089480: __GI__IO_doallocbuf at
:?
(/home/lennertfranssens/Documents/ReMon-new/ReMon/patched_binaries/libc/amd64/libc-2
.31.so)
0.414929 - MONITOR[0] - pid: 1382665 - 006: 00000000000885b8: __GI__IO_file_overflow
at :?
(/home/lennertfranssens/Documents/ReMon-new/ReMon/patched_binaries/libc/amd64/libc-2
.31.so)
0.414986 - MONITOR[0] - pid: 1382665 - 007: 0000000000087686: _IO_new_file_xsputn at
:?
(/home/lennertfranssens/Documents/ReMon-new/ReMon/patched_binaries/libc/amd64/libc-2
.31.so)
0.415040 - MONITOR[0] - pid: 1382665 - 008: 000000000006f41a: __vfprintf_internal at
:?
(/home/lennertfranssens/Documents/ReMon-new/ReMon/patched_binaries/libc/amd64/libc-2
.31.so)
0.415093 - MONITOR[0] - pid: 1382665 - 009: 000000000005bdd2: __printf at :?
(/home/lennertfranssens/Documents/ReMon-new/ReMon/patched_binaries/libc/amd64/libc-2
.31.so)
0.417473 - MONITOR[0] - pid: 1382665 - 010: 000000000003013d: ipmon_enclave at
/home/lennertfranssens/Documents/ReMon-new/ReMon/IP-MON/MVEE_ipmon.cpp:3618
(/home/lennertfranssens/Documents/ReMon-new/ReMon/IP-MON/libipmon.so)
0.417502 - MONITOR[0] - pid: 1382665 - 011: 000000000003302e:
ipmon_enclave_entrypoint at ??:?
(/home/lennertfranssens/Documents/ReMon-new/ReMon/IP-MON/libipmon.so)
0.417567 - MONITOR[0] - pid: 1382665 - 012: 00000000001050c5:
glibc_custom_syscall_exit at custom_syscall.c:?
(/home/lennertfranssens/Documents/ReMon-new/ReMon/patched_binaries/libc/amd64/libc-2
.31.so)
0.417571 - MONITOR[0] - pid: 1382665 - 013: couldn't find code address:
00000000000212af
0.417584 - MONITOR[0] - WARNING: DWARF: couldn't map EIP 00000000000212af to a known
region for variant: 0 (pid: 1382665)
1:27
Dat eerste fragment is wat heel veel in de log staat en dat tweede fragment is de

stack trace

babrath
1:27 PM
hm
1:27
die brk, wordt die ge-traced?

Lennert Franssens
1:27 PM
ja
1:28
want anders zou hij niet in de log staan :slightly_smiling_face:

babrath
1:28 PM
ja idd

Lennert Franssens
1:28 PM
is dat een die eigenlijk vanuit ipmon komt en daarin moet afgehandeld worden?

babrath
1:28 PM
zeer vreemd dat die negatief is

Lennert Franssens
1:28 PM
ja

babrath
1:29 PM
misschien, dat weet ik niet zeker

Lennert Franssens
1:29 PM
en zo groot negatief
1:33
Die zou in de aangepaste glibc dan toch naar die errno-waarde moeten jumpen?
1:33
/* Custom syscall function. */
void __custom_syscall(void)

```
{
   asm volatile(
       "pushq %r14\n\t"
       "pushq %r15\n\t"
       "movq %rax,%r14\n\t"
       "syscall\n\t"
       "cmpq $-0x085,%rax\n\t"
       "jge glibc_custom_syscall_exit\n\t"
       "neg %rax\n\t"
       "shl $12,%rax\n\t"                        // we know that the 12 lsb's are 0
so we shift the errnocode 12 bits to the left
       "movq %rax,%r15\n\t"
       "movq %r14,%rax\n\t"
       "call %r15\n\t"
       "glibc_custom_syscall_exit:\n\t"
       "popq %r15\n\t"
       "popq %r14\n\t"
       "nop\n\t"
   );
}
```

babrath
  1:34 PM
dat lijkt mij dat die dat zou doen ja :open_mouth:
1:36
als je alles via trace laat gaan, komen er dan toch nog via een te grote errno terecht in de enclave?

Lennert Franssens
  1:37 PM
Nee, want dan print hij die printf(...) ook niet uit

babrath
  1:37 PM
en dan geeft brk dus wel normale waardes terug?

Lennert Franssens
  1:38 PM
Ja

babrath
  1:39 PM
dat is... vreemd
1:39
en als je alles trace'd, BUITEN brk?

Lennert Franssens
1:40 PM
Dat zal ik ook eens doen
1:40
Dit is met alles op trace
Binary

MVEE.log
Binary

1:40
Dit is met alles, buiten getpid, op trace
Binary

MVEE.log
Binary

1:40
Die fouten met SYS_BRK blijven maar komen
1:40
wat niet is bij het eerste bestand

babrath
1:41 PM
alles behalve getpid of alles behalve brk?

Lennert Franssens
1:41 PM
alles behalve getpid
1:41
die met 'behalve brk' ga ik nu uitvoeren
1:43
Dit is alles trace, behalve brk
Binary

MVEE.log
Binary

1:43
Geeft ook en fout op die printf
1:44
Nu ga ik een doen met alles trace, behalve getpid en brk

Lennert Franssens
  1:51 PM
Ik heb toch gewoon dit stukje code nu toegevoegd:
```
if (RB == 0) {
        ipmon_register_thread();
        RB = ipmon_RB;
}
```
1:51
Nu werkt het wel
1:52
Om dit van vroeger te vervangen:
```
// The kernel sets the highest bit of the RB pointer after every fork/clone
        // Check if the highest bit is set
        if ((unsigned long)RB & (1UL<<(sizeof(unsigned long)*8 - 1))) {
                // The way the variants attach to RB/regfile map has a big drawback
and we need to do some work
                // After fork the child has also mapped the parent's shared memory
segments in its memory
                //     1) these segments are protected with PKU (if it exists)
                //     2) If we call execve (which happens really often after fork)
everything is cleared and we are fine
                // We need to unset to highest bit of the address to take the actual
address of parent's RB (see the kernel patch for more details)
                ipmon_checked_syscall(__NR_shmdt, (void*)((unsigned long)RB & ~(1UL
<< (sizeof(unsigned long)*8 - 1)))); // detach from parent's RB
                ipmon_checked_syscall(__NR_shmdt, ipmon_reg_file_map); // detach
from parent's file map

                RB = (ipmon_buffer *) ipmon_register_thread();
        }
```
1:53
Hoewel ik niet weet hoe ik dit dan moet vervangen:
```
ipmon_checked_syscall(__NR_shmdt, (void*)((unsigned long)RB & ~(1UL <<
(sizeof(unsigned long)*8 - 1)))); // detach from parent's RB
ipmon_checked_syscall(__NR_shmdt, ipmon_reg_file_map); // detach from parent's file
map
```

babrath
  1:53 PM
dat is voor na een fork, dit is niet na een fork

Lennert Franssens
  1:53 PM
Ah

babrath

1:53 PM
maar het opnieuw registreren is goed
1:53
dat moet zoiezo gebeuren
1:53
ik weet niet waarom dat niet zo was?

Lennert Franssens
1:54 PM
ik weethet ook niet

babrath
1:54 PM
en doet brk nog altijd zo vreemd?

Lennert Franssens
1:54 PM
ik ben gaan vergelijken met de oude code
1:54
& neen

babrath
1:54 PM
ik heb de logs vergeleken, en voor de getpid-versie is brk broken vanaf de IP-MON registratie
1:55
en ik heb geen idee waarom :sweat_smile:
1:55
iets in de state lijkt mij corrupted, maar ik zou niet weten waar
1:55
het lijkt mij enkel maarin de MVEE zelf te kunnen zijn
1:55
maar als het nu ineens werkt? ook goed?

Lennert Franssens
1:56 PM
Ja, maar ik ga het wel opschrijven want dit is wel iets dat duidelijk nog op een andere manier kan en moet opgelost worden :smile:
1:57
Misschien is dat probleem met die && hierdoor ook wel opgelost

babrath
1:58 PM
ja

1:58
het lijkt mij dat er idd iets belangrijks ergens mis is

Lennert Franssens
  1:58 PM
Oke neen, dat is ook nog steeds stuk

babrath
  1:58 PM
maar zolang micro-benchmarks met getpid deftig werken is het voor momenteel wel goed

Lennert Franssens
  1:58 PM
Ja, ik ga ze snel eens laten lopen
1:59
En als ze werken kan ik helemaal focussen op de presentatie :slightly_smiling_face:

babrath
  1:59 PM
yup

Lennert Franssens
  2:05 PM
 > native run
   > 0: (207.600006+190.199997+192.100006+183.699997+74.900002)/5 ns
   > 1: (64.599998+56.700001+60.5+55.0+60.0)/5 ns
   > 2: (56.700001+58.700001+63.400002+54.099998+53.599998)/5 ns
   > 3: (56.900002+54.0+58.299999+58.599998+53.099998)/5 ns
   > 4: (56.299999+54.299999+57.0+53.0+52.900002)/5 ns
 > default mvee
   > 0: (16222.299805+14956.299805+13255.799805+13735.700195+13206.5)/5 ns
   > 1: (16651.300781+13203.0+13786.400391+13158.200195+13050.099609)/5 ns
   > 2: (12771.200195+13086.200195+12993.299805+13468.799805+12672.400391)/5 ns
   > 3: (14999.900391+13730.700195+13294.799805+13407.099609+14289.799805)/5 ns
   > 4: (15147.700195+13142.799805+13040.599609+13165.900391+13482.900391)/5 ns
 > ipmon enabled mvee with getpid traced
   > 0: (16187.299805+15055.299805+16162.599609+14724.299805+16208.900391)/5 ns
   > 1: (15288.799805+14887.599609+14330.0+14177.299805+15037.900391)/5 ns
   > 2: (14785.900391+14987.799805+14119.200195+14122.799805+15328.5)/5 ns
   > 3: (14712.799805+14796.700195+14070.400391+14984.099609+14639.200195)/5 ns
   > 4: (14808.400391+14806.099609+14007.900391+14189.5+14741.400391)/5 ns
 > ipmon enabled mvee with getpid allowed
   > 0: (2198.399902+2181.800049+2502.0+2237.399902+2441.199951)/5 ns
   > 1: (1744.800049+1730.599976+1741.400024+1763.800049+1743.0)/5 ns
   > 2: (1705.300049+1683.099976+1684.800049+1821.199951+1712.900024)/5 ns

> 3: (1702.400024+1683.800049+1692.400024+1714.199951+1709.300049)/5 ns
> 4: (2090.399902+2022.5+2755.100098+2075.100098+2055.300049)/5 ns
2:05
Dat ziet er goed uit he

babrath
2:06 PM
ok :smile:

Lennert Franssens
2:06 PM
Enkel ipmon op de oude manier moet ik nog eens benchmarken

babrath
2:06 PM
kun je de resultaten wel even duidelijk schrijven, al berekenen?

3 replies
Last reply 2 months agoView thread

babrath
2:06 PM
ok!

babrath
2:13 PM
replied to a thread:
kun je de resultaten wel even duidelijk schrijven, al berekenen?
gewoon, hoeveel ns per getpid syscall :-)
View newer replies
2:13
ahja, en het moet natuurlijk wel in benchmark mode :stuck_out_tongue:
:sweat_smile:
1

2:14
hoe ziet je microbenchmark er nu uit qua source code?

Lennert Franssens
2:15 PM
```
#include <chrono>
#include <stdio.h>
#include <unistd.h>
```

```
#include <string.h>
#include <stdlib.h>
#include <errno.h>

#define GETPID_TEST_COUNT    500000000
#define LOOP_TIMES            5

int main()
{
    // getpid benchmark setup
--------------------------------------------------------------------------------
------
    for (int size_i = 0; size_i < LOOP_TIMES; size_i++)
    {
        auto start = std::chrono::high_resolution_clock::now();
        for (int cnt_i = 0; cnt_i < GETPID_TEST_COUNT; cnt_i++)
            getpid();
        auto end = std::chrono::high_resolution_clock::now();
        float result = std::chrono::duration_cast<std::chrono::nanoseconds>(end -
start).count();
        printf("\t> %u: %f ns\n", size_i, result / GETPID_TEST_COUNT);
    }
    return 0;
}
```

babrath
  2:15 PM
ziet er wel goed uit
2:16
in benchmark mode ga je zo'n groot verschil tussen native run en ip-mon enabled with getpid allowed niet meer hebben

Lennert Franssens
  2:17 PM
Het staat nu in benchmark mode en ik had ook GETPID_TEST_COUNT ook aangepast van 10 naar 500000000 dus binnen een kwartiertje heb ik de resultaten wel denk ik

babrath
  2:17 PM
ok

Lennert Franssens
  2:17 PM
Of misschien dat ik het straks nog eens opnieuw laat lopen op een tty en gnome afgesloten zodat er geen onnodige processen draaien

babrath
  2:21 PM
ja dat kan ook
2:21
ma dit zijn al goede indicatieve eerste resultaten he :stuck_out_tongue:

Lennert Franssens
  2:22 PM
Awel ja misschien dat ik dat dan zal doen voor in het definitieve verslag

babrath
  2:23 PM
ja, lijkt mij wel goed

Lennert Franssens
  3:19 PM
 > native run
   > 0: (57.525532+60.262741+55.475948+58.16674+61.036579)/5 ns
   > 1: (57.03336+57.407681+56.233768+61.588428+56.724312)/5 ns
   > 2: (57.899479+56.596821+54.102482+56.864159+55.959438)/5 ns
   > 3: (60.287552+56.610142+56.306782+57.941761+56.89304)/5 ns
   > 4: (53.57206+57.820999+58.044231+59.20525+56.034191)/5 ns
 > default mvee
   > 0: (7981.168457+8014.01416+8152.656738+7919.495117+8043.988281)/5 ns
   > 1: (8031.134766+8315.574219+8018.098633+7989.904785+7948.680176)/5 ns
   > 2: (7870.553711+8228.581055+7893.627441+7995.883301+7907.382812)/5 ns
   > 3: (7926.350586+7910.257812+7911.106445+8004.90625+7998.393555)/5 ns
   > 4: (8089.540527+8109.269043+8052.602051+8081.146973+7951.020996)/5 ns
 > ipmon enabled mvee with getpid traced
   > 0: (8410.400391+8281.348633+8530.999023+8252.611328+8383.428711)/5 ns
   > 1: (8388.288086+8245.620117+8146.510254+8332.170898+8319.286133)/5 ns
   > 2: (8289.200195+8236.555664+8251.634766+8302.802734+8245.514648)/5 ns
   > 3: (8237.897461+8286.208008+8252.667969+8377.048828+8244.000977)/5 ns
   > 4: (8233.333008+8198.499023+8325.657227+8159.114746+8296.112305)/5 ns
 > ipmon enabled mvee with getpid allowed
   > 0: (251.836838+256.654968+256.475433+254.082764+258.075989)/5 ns
   > 1: (252.085007+256.669708+255.106567+253.833237+256.828552)/5 ns
   > 2: (251.73616+258.07962+255.912964+254.223755+257.189392)/5 ns
   > 3: (252.185364+256.888855+255.349518+253.776199+257.314209)/5 ns
   > 4: (252.268356+258.711609+255.270935+253.598557+257.009491)/5 ns
3:22
Hierop kunnen we zien dat de filter wel wat vertraging introduceert bij een TRACE van alle syscalls. Maar bij het returnen van een errno om vervolgens een ALLOW te kunnen doen, is het wel spectaculair sneller, toch?

babrath
  3:28 PM
hmm idd
3:28
en wat is het bij IPMON zonder bpf-filter, dus met in-kernel broker?
:slightly_smiling_face:

Lennert Franssens
  3:30 PM
Daar ben ik nu een benchmark test voor aan het maken :slightly_smiling_face:
:+1:
1

3:30
(een automatische)

Lennert Franssens
  4:41 PM
Wow
4:41
 > native run
   > 0: (58.445621+63.259781+58.284672+59.867001+58.976509)/5 ns
   > 1: (73.30024+73.392159+58.891891+60.09304+60.030369)/5 ns
   > 2: (59.00309+56.078739+58.753391+60.236191+58.806889)/5 ns
   > 3: (58.804852+57.02285+59.158901+59.332031+59.118832)/5 ns
   > 4: (52.79187+56.093658+61.240551+58.134941+59.83102)/5 ns
 > default mvee
   > 0: (7822.694336+7856.40332+7888.188965+7935.361328+7937.004883)/5 ns
   > 1: (7909.129395+7762.630859+7932.024902+7992.662598+7842.092285)/5 ns
   > 2: (7873.143066+8885.042969+7848.388672+7806.313477+7936.63916)/5 ns
   > 3: (7989.40332+7881.083496+8142.883301+7828.874023+7702.525391)/5 ns
   > 4: (7887.935547+7960.20459+8012.307617+7847.413086+7794.720703)/5 ns
 > ipmon enabled mvee with getpid traced
   > 0: (7953.077637+7837.025391+7780.78418+7935.316895+7882.363281)/5 ns
   > 1: (8038.478027+7805.273438+7912.146484+7833.367188+7821.367188)/5 ns
   > 2: (7978.974609+8009.303223+7853.672363+7797.820312+7907.481445)/5 ns
   > 3: (7865.509766+7793.019043+7832.789551+7776.965332+8054.753906)/5 ns
   > 4: (7907.522461+7875.815918+7939.48584+7824.627441+7910.172363)/5 ns
 > ipmon enabled mvee with getpid allowed
   > 0: (325.939728+318.341888+329.07666+323.765442+320.140991)/5 ns
   > 1: (325.734985+318.566589+328.808838+320.826385+320.906921)/5 ns
   > 2: (322.452789+319.366211+328.364105+321.134125+319.48053)/5 ns
   > 3: (321.165527+319.49231+330.650665+318.673828+319.114685)/5 ns
   > 4: (321.872101+318.056091+327.474213+320.079773+318.506592)/5 ns
4:42
Dat wil zeggen dat het met bpf het sneller is dan met de oude implementatie (edited)

Lennert Franssens
  4:48 PM
Kan dat wel eigenlijk?

Lennert Franssens
  4:58 PM
Dit kan toch niet
Screenshot from 2022-05-04 16-58-28.png
Screenshot from 2022-05-04 16-58-28.png

babrath
  5:05 PM
hmm
5:05
heb je nog een bpf filter aanstaan? het is wat verwarrend dat er nog steeds staat "getpid allowed"
5:05
of is dat IP-MON-speak?
5:05
also, post het eens in de gedeelde chat met @bcoppens
 erbij
5:06
dit is belangrijke info :stuck_out_tongue:

Lennert Franssens
  5:06 PM
Haha oke dat zal ik doen & neen geen bpf-filter meer, zou zelfs geen impact mogen hebben want ik doe die benchmark op jullie repo (zonder mijn aanpassingen)

babrath
  5:06 PM
ok

Lennert Franssens
  10:16 AM
Ik heb nog eens kort wat meer proberen te weten te komen hoe het komt dat ls && ls blijft hangen. En in de bijgevoegde log staat tijdens de postcall van wait4 een debug message van de vorm "INFO: In POSTCALL of wait4". Nu, die log is zonder het gebruik van IP-MON. Wanneer use_ipmon op true staat, blijft het hangen op die eerste SYS_WAIT4. En tussen de aanroep van die syscall en het uitvoeren van de postcall, zou een execve moeten uitvoeren die de eerste ls van "ls && ls" uitvoert. Dat

gebeurt dus niet... De log met het gebruik van IP-MON stuur ik ook door. Die uitvoering wordt na de eerste SYS_WAIT4 afgebroken door CTRL+C omdat het dus blijft "hangen". Ik heb het gevoel dat er een component niet wordt opgeroepen door de een of andere reden. Net zoals eigenlijk ook de init()-functie van IP-MON zelf niet altijd wordt aangeroepen bij elke execve, om een nog onbekende reden. Dat hebben we opgelost door die vanuit IP-MON zelf op te roepen als dat nog niet gebeurd was.
2 files

MVEE.log
Binary

MVEE_with_ipmon.log
Binary

babrath
  11:23 AM
uit de ipmon log: "0.332657 - MONITOR[0] - no suitable mapping found for pid -1 "
11:24
hmm, dat heeft de andere log ook
11:25
wel, wat je iig aan de ipmon log kan uitlezen is het volgende:
11:25
eerst is er maar 1 proces (met 2 varianten) dat gemonitored wordt door monitor 0 (MONITOR[0])
11:25
vervolgens wordt er een nieuw proces (ook met 2 varianten) gecreëerd door de fork: "0.332296 - MONITOR[1] - monitor running! - created by monitor: 0 "
11:26
dat wordt dan gemonitored door monitor 1 (MONITOR[1])
11:26
proces 0 gaat kort daarna wachten op iets (vermoedelijk proces 1), via SYS_WAIT4
11:27
als we kijken naar de laatste actie die proces 1 (waar proces 0 op aan het wachten is), dan is dit: "0.332423 - MONITOR[1] - SYS_GETPID - >>> Dispatch as SYNCED MASTERCALL "
11:27
proces 1 hangt dus vast in een getpid
11:27
de vraag is waarom

Lennert Franssens
  11:28 AM
Ja want ik heb een filter in ipmon die "leeg" is en op alles trace terugstuurt.
11:28

De werking zou dus hetzelfde moeten zijn als zonder ipmon.
11:30
Hoewel nu met die tweede monitor... Door de fork van dat eerste proces wordt wel de bpf filter meegenomen. Maar misschien stuurt die geen register_ipmon fake syscall uit zodat die met ptrace_syscall werkt ipv te switchen tussen ptrace_syscall en ptrace_cont. Daardoor zou hij kunnen blijven hangen omdat er nooit een geldig antwoord zal komen

babrath
  11:31 AM
wel er gaat zoiezo een ptrace_syscall gebeuren lijkt mij, anders zou getpid niet gelogd worden door CP-MON?
11:32
ah nee, ik ben in de war, ja
11:32
na een fork moet je terug die state switchen in de monitor wil je zeggen?

Lennert Franssens
  11:32 AM
Ja
11:33
Maar ik ben niet zeker of dat al niet automatisch gebeurt bij een fork dat ipmon zichzelf terug registreert
11:34
Het zal wel problematisch blijven aangezien de syscalls al in die nieuwe states moeten afgehandeld worden nog voor ipmon zichzelf opnieuw registreert
11:36
Dan is de oplossing dus om de state van de voorgaande monitor over te nemen bij een fork
11:36
Denk/hoop ik? :smile:

babrath
  11:45 AM
wel bij een fork wordt simpelweg de adresruimte gedupliceert
11:46
IP-MON zou nog aanwezig moeten zijn op dezelfde adressen

Lennert Franssens
  11:48 AM
En de monitor weet dus nog niet dat hij van state moet wisselen?
11:48
De nieuwe monitor*

babrath

12:07 PM
Dat weet ik niet? Ik zou eens naar de code kijken en logging toevoegen
12:07
Hoe zit het met het schrijven? :slightly_smiling_face:

Lennert Franssens
12:48 PM
Goed en niet goed. Ik had tegen zondag en gisteren nog 2 grote projecten. Dus ik heb enkel de wijzigingen op basis van de feedback nog maar kunnen doen. Maar vanmiddag, morgen en vrijdagochtend heb ik niks anders meer te doen. Dus dat kan ik helemaal spenderen aan het toevoegen van de nieuwe schema's en het nog iets uitgebreider brengen zodat alles duidelijk is in de eerste twee hoofdstukken van wat gebruikt wordt in de volgende :slightly_smiling_face:
12:48
Maar dus tegen vrijdagmiddag zou ik normaal wel een goede versie kunnen doorsturen.
12:49
Dan kan ik in het weekend schrijven over de microbenchmarks
12:49
En volgende week het probleem met die fork proberen aanpassen
12:49
Om dan nog meer benchmarks te doen en daarover te schrijven.
12:50
En dan zou het af moeten zijn :slightly_smiling_face:

babrath
12:54 PM
ok :slightly_smiling_face:

Lennert Franssens
11:03 AM
Ik heb nog een hoodfstuk gevonden
11:03
Ik kan ook nog iets zeggen over het genereren van de filter
11:03
en dan door het oplossen van het probleem met fork kan ik ook nog een paar extra alinea's schrijven. (zowel in het hoofdstuk van het nieuw ontwerp en de implementatie ervan, en in het hoofdstuk over de technologieverkenning) (edited)
11:04
En bij de benchmarks heb ik nu ook nog 3 extra pagina's, wat ook nog eens extra kan uitgebreid worden door de benchmarks met nginx.
11:05
Dus aan 40 pagina's geraak ik wel denk ik. Maar ik vind het moeilijk om dingen te schrijven over hetgeen wat er nu al staat, aangezien ik dan eigenlijk alles begin te herhalen :confused:
11:06
En is het dan de kwaliteit of de kwantiteit die telt?

babrath
  11:09 AM
ik weet niet direct wat je wil zegen over het genereren van de filter?
11:10
fork kan je idd meer over zeggen, zeker over implementatie

Lennert Franssens
  11:10 AM
Het scriptje dat aan de hand van een policy de filter in ipmon zet

babrath
  11:10 AM
ahh
11:10
hmm, ik zou vermelden dat je daar een script voor hebt, maar voor de rest is dat geen interessant implementatiewerk vrees ik :confused:

Lennert Franssens
  11:11 AM
En ik mag ook mijn benchmarks om de overhead van een groeiende filter te meten niet vergeten en de overhead van een grotere sleutel is ook nog cruciaal

babrath
  11:11 AM
meer benchmarking evaluatie is goed :stuck_out_tongue:
11:11
idd
11:12
en mogelijke suggesties tot future work, bespreking van mogelijke alternatieven, zijn altijd welkom

Lennert Franssens
  11:12 AM
Ahja ik zal dat titeltje al toevoegen in de conclusie, dan vergeet ik dat ook niet

babrath
  11:13 AM
:+1:

Lennert Franssens
  11:13 AM
Buiten mijn examens is dit nu nog het enige (gisteren had ik nog een deadline en een

presentatie van een project), dus het komt wel in orde denk ik :slightly_smiling_face:

babrath
11:13 AM
Ok :slightly_smiling_face:

Lennert Franssens
11:16 AM
En hoe groot moet de interlinie juist zijn? Ik heb de masterproef van meneer Coppens eens bekeken en hij gebruikte een dubbele lijn per nieuwe lijn. Moet die ook zo groot zijn bij mij?

babrath
11:21 AM
huh
11:21
goede vraag, zijn daar geen meer algemene guidelines voor van de faculteit ofzo?

Lennert Franssens
11:23 AM
Ja, ze zeggen 16 punt maar ik weet niet op wat dat juist neerkomt? :smile:

babrath
11:24 AM
16-punt interlinie? of typesize?

Lennert Franssens
11:24 AM
Regelafstand 16 punt staat er
11:24
https://www.ugent.be/ea/nl/faculteit/studentenadministratie/masterproef/vorm1920.pdf
11:25
Ongeveer anderhalf dan waarschijnlijk?

babrath
11:26 AM
afhankelijk van u fontsize he? :stuck_out_tongue:

Lennert Franssens
11:26 AM
Euhm ja? :smile:

babrath
  11:27 AM
welja, als u interline 16punt is en u fontsize ook 16punt is, dan is het maar 1?

Lennert Franssens
  11:31 AM
Ik ben eens in het boek van mevrouw Pollefliet aan het kijken en zij zegt ook dat 16 punt goed is, wat neerkomt op een regelafstand van 1,3.

babrath
  11:32 AM
ok

Lennert Franssens
  10:06 AM
Het loop toch weer vast. Daarnet heb ik met de patched glibc-2.31.so die al in de repo zit geprobeerd. Door op alles een RET_TRACE te sturen was het niet nodig om de door mij aangepaste versie van glibc te gebruiken (met de custom_syscall functie). Wanneer ik nu toch mijn aangepaste versie van glibc gebruik, loop het vast, ookal stuur ik op alles een RET_TRACE... :confused:

babrath
  10:09 AM
en welke stack trace?

Lennert Franssens
  10:12 AM
2 files

MVEE_custom_glibc.log
Binary

MVEE_orig_glibc.log
Binary

Lennert Franssens
  10:26 AM
Het is dus weldegelijk op de syscall instructie zelf

image.png
image.png

Lennert Franssens
  10:36 AM
Kan ik komen?

Lennert Franssens
  12:02 PM
0.569636 - MONITOR[0] - INFO: calling resume for syscall and ipmon_active=1
0.569638 - MONITOR[0] - WARNING: postcall handler handled resume. not resuming...
0.569643 - MONITOR[0] - INFO: Syscall status = STOP_SIGNAL
0.569646 - MONITOR[0] - INFO: Handling signal event...
0.569724 - MONITOR[0] - Variant:0 [PID:18734] - variant crashed - trapping ins:
00000000001050ab: __custom_syscall at ??:?
(/opt/repo/patched_binaries/libc/amd64/libc-2.31.so)
0.569726 - MONITOR[0] - Variant:0 [PID:18734] - Received signal SIGSEGV (11)
0.569729 - MONITOR[0] - Variant:0 [PID:18734] - signal arrived while variant was
executing ins: 00000000001050ab: __custom_syscall at ??:?
(/opt/repo/patched_binaries/libc/amd64/libc-2.31.so)
0.569731 - MONITOR[0] - Variant:0 [PID:18734] - ret is currently: 13
0.569737 - MONITOR[0] - WARNING: Warning: SIGSEGV in variant 0 (PID: 18734)
0.569741 - MONITOR[0] - WARNING: IP: 00007f44e02570ab, Address: 0000000000000000,
Code: SI_KERNEL (128), Errno: 0
(edited)

babrath
  12:02 PM
is dat goed of niet? :stuck_out_tongue:

Lennert Franssens
  12:03 PM
Neen haha :sweat_smile:
12:03
Het zou zo moeten gaan:
0.555526 - MONITOR[0] - INFO: calling resume for syscall and ipmon_active=1
0.555527 - MONITOR[0] - WARNING: postcall handler handled resume. not resuming...
0.555537 - MONITOR[0] - INFO: Syscall status = STOP_SECCOMP
0.555547 - MONITOR[0] - Variant:0 [PID:18487] - SYS_RT_SIGACTION(2 - SIGINT -
SIG_DFL)
0.555552 - MONITOR[0] - SYS_RT_SIGACTION - >>> Dispatch as SYNCED NORMAL
0.555561 - MONITOR[0] - INFO: Syscall status = STOP_SYSCALL
0.555565 - MONITOR[0] - Variant:0 [PID:18487] - SYS_RT_SIGACTION return: 0
0.555566 - MONITOR[0] - In call_postcall_return

```
0.555568 - MONITOR[0] - sighandler changed for sig: SIGINT
0.555569 - MONITOR[0] - > SIGACTION sa_handler   : 0x0000000000000000 (= SIG_DFL)
0.555570 - MONITOR[0] - > SIGACTION sa_sigaction : 0x0000000000000000
0.555571 - MONITOR[0] - > SIGACTION sa_restorer  : 0x00007fb529b1e230
0.555572 - MONITOR[0] - > SIGACTION sa_flags     : 0x04000000 (= )
0.555574 - MONITOR[0] - > SIGACTION sa_mask      :
```

12:03
Dus met de aangepaste glibc krijg ik een STOP_SIGNAL in plaats van een STOP_SECCOMP
12:08
Dat signaal is door de crash, dus het komt niet door de MVEE zelf denk ik

babrath
  12:38 PM
welk signaal komt door de craash? STOP? lijkt mij niet een signal voor een crash?

Lennert Franssens
  12:39 PM
Is een SIGSEGV geen signal?

babrath
  12:40 PM
ik zie geen SIGSEGV? (edited)

Lennert Franssens
  12:40 PM
Ik pas het even aan
12:41
Dus de voorlaatste log bevat nu alle info
12:41
En daar staat die SIGSEGV

Lennert Franssens
  3:49 PM
Ik krijg geen logs meer uit IP-MON, die geven ook segmentation faults

Lennert Franssens
  3:56 PM
RB en entry zijn niet 0 :confused:

babrath
  3:57 PM
en in de handler zelf?
3:58

welke variabelen zijn er < 0x20?

Lennert Franssens
4:15 PM
accept4 - (int fd, struct sockaddr* peer_sockaddr, int* peer_addrlen, int flags)
4:15
ARG2 en ARG3 zijn 0
4:15
en ARG4 ook maar dat kan geen kwaad

Lennert Franssens
6:24 PM
Het lijkt dus alsof die brk-aanroep op een of andere manier (misschien vanuit de filter?) altijd een adres in IP-MON teruggeeft op dezelfde manier als wij nu doen vanuit de filter.
6:28
Binary

MVEE.log
Binary

6:30
Ik heb ipmon nu gemapt op 0x800000 (zodat het adres boven 2048 + 12 lsb's op 0, decimaal ligt). Ook glibc is aangepast. Maar dusja de brk-aanroep weet ip-mon altijd "te vinden" als errno-waarde. En dat is dus wat nu exact gebeurt in de filter, maar niet voor brk...
6:30
2 files

MVEE_ipmon_seccomp_bpf_policy.h
C

seccomp_bpf_policy.json
JavaScript/JSON

babrath
10:24 PM
dat is bizar
10:24
hb je al eens de manpage voor brk bekeken?

(https://man7.org/linux/man-pages/man2/sbrk.2.html)
10:25
brk kan idd een adres teruggeven, maar dan gaat dit echt een adres zijn, niet het adres omgezet naar die errno-waarde :stuck_out_tongue:
10:27
dus de filter kan de juiste errno-waarde teruggeven, maar brk (die dan vermoedelijk een CP-MON syscall is) ook?

Lennert Franssens
  10:21 AM
Blijkbaar
10:25
Het is wel zo dat met een "lege" filter, die op alles RET_TRACE stuurt, het wel werkt momenteel. Dus de fout heeft dan toch misschien iets te maken met de werking van de filter in het geheel. Maar waar weet ik dus niet direct :sweat_smile:

Lennert Franssens
  10:51 AM
En met bijvoorbeeld __NR_write in de filter om in IP-MON uit te voeren, lukt "echo test && echo test", "ls", "echo test && ls" maar niet "ls && echo test" of "ls && ls"...

Lennert Franssens
  11:03 AM
Ik denk dat ik het probleem weet
11:04
Volgens mij is het de groeiende filter
11:06
Ik ga nu proberen uitzoeken of ik kan nagaan of in het proces al een filter zit. Bij een fork moet die dan niet meer opnieuw ingesteld worden, want het is exact dezelfde die nu opnieuw wordt ingesteld, maar volgens mij iets speciaal doet om de vorige aan de nieuwe te koppelen, wat voor problemen zorgt in onze implementatie?

Lennert Franssens
  4:32 PM
Moet bij elke nieuwe monitor opnieuw ipmon ingeladen worden, inclusief de filter?

1 reply
1 month agoView thread

Lennert Franssens
  4:53 PM
Ja, ik heb het probleem nu echt gevonden
4:53
het is de groeiende filter

4:54
en ook gewoon het feit dat de filter blijft bestaan na een fork
4:55
dat zijn twee dingen waar de MVEE eigenlijk niet met overweg kan, of toch niet in zijn huidige implementatie

Lennert Franssens
5:02 PM
IMG_2022-05-26-17-02-13-976.jpg
IMG_2022-05-26-17-02-13-976.jpg

Lennert Franssens
5:39 PM
Dus IP-MON wordt ingeladen in Proces 1. Daarna gebeurt een fork want het commando "ls && echo test && ls && ls && ls" zal onderverdeeld worden in "ls && echo test", "ls", "ls" en "ls". De eerste fork is van Proces 1, met de geïnstalleerde filter, naar Proces 2. De filter uit Proces 1, zit sowieso in Proces 2, en werkt vanaf het begin, nog voor die zijn eigen IP-MON zal inladen. Monitor 1 (monitor 1 monitort proces 2, monitor 2 monitort proces 3...) is wel al op de hoogte van de geïnstalleerde (overgedragen) filter en registreert al STOP_SECCOMP events. Dat wil zeggen dat de filter die in Proces 1 werd geïnstalleerd, ook RET_TRACE terugstuurt voor systeemaanroepen uit Proces 2 nog voor die IP-MON heeft ingeladen. Op dat ogenblik zou er al een "allowed" systeemaanroep in Proces 2 kunnen gebeuren die naar IP-MON van Proces 1 springt vanuit Proces 2, wat niet de bedoeling is en segmentation faults oplevert. Maar stel nu dat we nog geen allowed systeemaanroepen hebben en IP-MON wordt ingeladen in Proces 2. Dan zal daarna wel een allowed systeemaanroep gebeuren (wat de bedoeling is van de filter, want het gaat stuk wanneer we allowed systeemaanroepen uitvoeren en niet bij TRACE-only :wink:). Op dat ogenblik is de nieuwe filter aan de oude toegevoegd. Maar omdat onze oude filter al zo strict is, en voor elke mogelijke flow een geldige uitkomst weet te genereren die een minstens even hoge prioriteit heeft als de nieuwe filter, zal die ook voor systeemaanroepen uit Proces 2 die vanuit IP-MON komen en RET_ALLOW in de nieuwe filter krijgen, ook nog in de "oude" filter gaan en een antwoord daaruit genereren als die een hogere prioriteit heeft, wat opnieuw op RET_ERRNO zal neerkomen en naar IP-MON van Proces 1 zal springen en dus een segfault zal genereren (ik stuur na dit bericht een foto met een duidelijk schema waarom dit zo is). Aangezien die dus een antwoord uit die oude filter zal terugsturen, wil dat zeggen RET_TRACE, RET_ALLOW of RET_ERRNO.
5:39
En hetzelfde gebeurt dus ook voor alle andere cloned processen
5:44
En hoe komt dat nu juist door de filter? Hieronder enkele fragmenten uit de manpage van seccomp:
If fork(2) or clone(2) is allowed by the filter, any child
processes will be constrained to the same system call
filters as the parent. If execve(2) is allowed, the

existing filters will be preserved across a call to
execve(2).
If multiple filters exist, they are all executed, in reverse
order of their addition to the filter tree—that is, the most
recently installed filter is executed first. (Note that all
filters will be called even if one of the earlier filters returns
SECCOMP_RET_KILL. This is done to simplify the kernel code and
to provide a tiny speed-up in the execution of sets of filters by
avoiding a check for this uncommon case.) The return value for
the evaluation of a given system call is the first-seen action
value of highest precedence (along with its accompanying data)
returned by execution of all of the filters.
In decreasing order of precedence, the action values that may be
returned by a seccomp filter are:

SECCOMP_RET_KILL_PROCESS

SECCOMP_RET_KILL_THREAD (or SECCOMP_RET_KILL)

SECCOMP_RET_TRAP

SECCOMP_RET_ERRNO

SECCOMP_RET_USER_NOTIF

SECCOMP_RET_TRACE

SECCOMP_RET_LOG

SECCOMP_RET_ALLOW
(https://man7.org/linux/man-pages/man2/seccomp.2.html)

Lennert Franssens
5:55 PM
IMG_2022-05-26-17-55-10-524.jpg
IMG_2022-05-26-17-55-10-524.jpg

5:57
Dus in het geval van een systeemaanroep die RET_TRACE zal krijgen is er geen probleem. Maar in het geval van een die in IP-MON geanalyseerd mag worden, en dus RET_ALLOW op een bepaald moment moet krijgen, en er een filter uit een "parent" reeds geïnstalleerd is, is er wel een probleem.

Lennert Franssens
6:05 PM
Stel dat we het nu over write hebben in Proces 2. Proces 2 zal filter 1 uitvoeren en in de eerste "run" van de filter naar RET_ERRNO evalueren. Daarna wordt filter 0

uitgevoerd. Die zal ook naar RET_ERRNO evalueren, maar omdat die dezelfde prioriteit heeft als de eerste RET_ERRNO, zal de eerste returnwaarde genomen worden (zie de man-page daarvoor). Dan zal glibc naar de enclave in IP-MON van Proces 2 springen. Daar zal de systeemaanroep opnieuw uitgevoerd worden, en dus ook de filter opnieuw uitgevoerd worden. Die zal terug beginnen met filter 1 uit te voeren (want de filters worden in omgekeerde volgorde van toevoegen uitgevoerd). Filter 1 evalueert naar een RET_ALLOW, waarna filter 0 uitgevoerd wordt. Filter 0 zal opnieuw naar RET_ERRNO evalueren. Aangezien RET_ERRNO een hogere prioriteit heeft dan RET_ALLOW, zal RET_ERRNO teruggestuurd worden, met als waarde het adres van de enclave van IP-MON van Proces 1. Aangezien Proces 2 dan naar dat adres vanuit de adresruimte van Proces 1 (waarin de enclave van IP-MON van Proces 1 zit) wil jumpen, krijgen we een fout.
6:06
En er is geen enkele manier hoe we dit kunnen oplossen aangezien RET_ERRNO altijd boven RET_ALLOW zal verkozen worden als er clones gebeurd zijn.
6:07
In de uitleg die ik gegeven heb, is de check of de systeemaanroep de eerste keer vanuit glibc komt achterwege gelaten, aangezien dat geen invloed heeft op de werking. RET_ERRNO zal dan boven RET_TRACE gekozen worden door een hogere prioriteit, en heeft dus geen invloed op hoe de systeemaanroepen in de aaneengeschakelde filter zal evalueren.
6:10
Ook het wijzigen van de manier waarop we de secret "doorsturen/bijhouden" wat we nu via RET_ERRNO doen, zal geen oplossing bieden. Want dan zal RET_ALLOW vervangen worden door een RET_TRACE, door hogere prioriteit, uit een hogere filter.
6:12
Het mappen van IP-MON op hetzelfde adres is dan misschien nog een mogelijkheid, gelijkaardig aan wat we nu doen bij een execve maar dan ook voor clones. Maar het biedt dan ook weer veel minder veiligheid...
6:13
Plus het gegeven dat veel programma's tegenwoordig met een seccomp-BPF filter werken, gooit roet in het eten voor onze filter. Er zullen beslissingen genomen worden die niet verwacht worden zowel in het programma zelf, als voor de MVEE.

babrath
8:08 PM
replied to a thread:
Moet bij elke nieuwe monitor opnieuw ipmon ingeladen worden, inclusief de filter? Elke variant moet ipmon hebben, maar zolang de filter al geïnstalleerd is (en ik denk dat je dit kan navragen) moet deze niet opnieuw geïnstalleerd worden

babrath
8:17 PM
Ik heb momenteel geen tijd om alles te lezen, maar er is geen reden om na elke fork de filter opnieuw te installeren. Zou dit het probleem oplossen?

Lennert Franssens

8:26 PM
Wel, de filter moet wel het adres van ipmon hebben uit de clone. Dus de filter moet wel opnieuw geïnstalleerd worden als we de locatie van ipmon via die errno willen doorgeven.

babrath
8:41 PM
Of de locatie constant houden over forks?

Lennert Franssens
8:41 PM
Van IP-MON?

babrath
8:43 PM
Ja

Lennert Franssens
8:43 PM
Kan dat?

babrath
8:44 PM
Ja

Lennert Franssens
8:44 PM
Dan ga ik dat nog eens proberen
8:46
Oké, dat heeft alles opgelost
8:47
nu werkt ls && ls && ls ... wel
8:48
dan zorg ik nu voor een "juiste" implementatie van het behouden van IP-MON op hetzelfde adres
8:48
En dan probeer ik nginx toch nog uit

babrath
10:56 PM
Ok :ok_hand: :smile: :tada:

Lennert Franssens
  12:14 AM
Het blijft maar niet ophouden... :face_with_rolling_eyes::joy: Nu is er nog een probleem met "ls && ls && ls" wanneer ik 2 of meer varianten uitvoer. En ook nginx krijg ik niet aan de praat, in de logs staat iets over futex :confused:

babrath
  10:08 AM
is het misschien iets gelijkaardig, met het mappen van IP-MON?
10:08
krijg je een segfault, of wat gebeurt er?

Lennert Franssens
  11:46 AM
Ik ga eens kijken :slightly_smiling_face:
11:49
Ahja nu ik het zie weet ik het weer
11:49
Het is soms dat het niet werkt
11:49
met 2 varianten of meer
11:50
en er is een fout door een mismatch van de argumenten van een systeemaanroep in beide varianten
11:50
ik stuur de log even door
11:51
Dus hier ziet u dat het wel werkt
image.png
image.png

11:51
En hier dat het niet werkt
11:51
image.png
image.png

11:52
(met deze log)
Binary

MVEE.log
Binary

babrath
  12:18 PM
En is dat met futex afgehandeld door cpmon of ipmon?

Lennert Franssens
  12:32 PM
Futex-aanroepen die vanuit IP-MON zelf komen (voor de werking van IP-MON), gebeuren door IP-MON. Andere futex aanroepen, uit het proces, worden naar CP-MON gestuurd.

babrath
  1:10 PM
Hmm
1:10
Dat klinkt alvast goed

Lennert Franssens
  2:41 PM
Als ik de filter leeglaat, of zelfs IP-MON niet gebruik (use_ipmon: false), krijg ik dezelfde fout

Lennert Franssens
  5:10 PM
Dat is allemaal op mijn laptop gedaan wat ik hiervoor stuurde. Nu probeer ik het op mijn desktop, en daar geeft het geen fouten. Vreemd, niet?

babrath
  7:41 PM
Ja
7:41
Kernel versies, meer of minder seccomp filters?
7:42
Maar werkt het dus volledig op je desktop?
7:42
Kan je dazr seccomp filter acties loggen?

Lennert Franssens
  9:47 PM
Ik heb voor de zekerheid een clean install op mijn laptop en op mijn desktop gedaan van Ubuntu 20.04. Daarop heb ik dan een native install van ReMon met IP-MON gedaan, met exact dezelfde filters. En op mijn desktop werkt het, en op mijn laptop niet (1 op ongeveer 5 keer wel)
9:48
Ik zal nog eens proberen om de acties te loggen

babrath
  10:40 PM
Wel, als het op u desktop werkt zou ik die even gebruiken om verder op te ontwikkelen :sweat_smile:
10:41
Ja, er zal nog wel ergens iets verkeerd zitten, mogelijk met het repliceren van synchronisatie, maar er zijn andere dingen die best eerst gedaan worden :slightly_smiling_face:

Lennert Franssens
  10:42 AM
Inderdaad :smile:
10:45
Het enige verschil is dat ik op mijn laptop dit heb uitgevoerd om mijn trackpad te laten werken:
sudo sed -i 's/GRUB_CMDLINE_LINUX_DEFAULT="[^"]*/& i8042.nopnp=1 pci=nocrs/' /etc/default/grub

babrath
  10:55 AM
Ah
10:56
Ja dat is verkeerd
10:56
Krijg je die warning bij het begin van het uit voeren van de mvee niet over grub cmdline en vdso?

Lennert Franssens
  11:33 AM
Neen, dat commando heb ik ook al uitgevoerd (edited)
11:34
En ik heb die twee boot parameters die ik had toegevoegd ook eens verwijderd, maar dat helpt ook niet :/ maar zoals u al zei ga ik dan best op mijn desktop verder
11:35
Zou dit iets kunnen zijn dat aan de cpu ligt? Of is het zeker een implementatie- of installatiefout?

babrath
  12:33 PM
Misschien cpu, niet zeker
12:33
Je laptop is ook zeker heropgestart en er zijn dus geen warnings meer?

Lennert Franssens
  12:33 PM
Ja
12:37
En 2 clean installs gedaan. En getest op native ubuntu 20.04 en in de container. Met alle juiste instellingen, glibc, IP-MON enabled... Het is zelfs zo dat zonder IP-MON het ook niet werkt op mijn laptop

babrath
  1:02 PM
Uhu
1:02
Ja das vreemd

Lennert Franssens
  3:13 PM
Ja, laat ons zeggen dat het aan de cpu zal liggen :smile: (edited)

babrath
  3:43 PM
het kan hoor, onwaarschijnlijk, maar ik kan mij minstens 1 scenario bedenken :stuck_out_tongue:

Lennert Franssens
  3:44 PM
En wat is dat dan juist? :slightly_smiling_face:

babrath
  3:44 PM
als u cpu Intel TSX heeft

Lennert Franssens
  9:47 AM
Met welke versie van nginx moet ik dat testen?

Lennert Franssens
  9:52 AM
Of gewoon de versie van de eurosys benchmarks, maar dan zonder ln -s "$__current_dir/../patches/nginx/b/src/os/unix/ngx_shmem.c" -f "$__current_dir/src/os/unix/ngx_shmem.c" voor "default"?

babrath

9:57 AM
uhh
9:57
die versie ja
9:57
wat is eht probleem met die ln?

Lennert Franssens
9:58 AM
Niks denk ik maar ik krijg een warning op mmap en dacht dat het daar aan lag. Maar het ligt aan iets anders

babrath
9:58 AM
ah, nee
9:58
die is normaal vrees ik :stuck_out_tongue:

Lennert Franssens
9:59 AM
Binary

MVEE.log
Binary

9:59
Normaal in welke zin? :smile:
10:00
Nu build ik de MVEE wel zonder MVEE_ALLOW_SHM

babrath
10:01 AM
hmm
10:01
wel
10:01
dat zou moeten werken
10:01
misschien zijn er nog SHM-related defines die uit moeten?
10:02
anders zou ik ook eens met proberen

Lennert Franssens
10:02 AM

Deze is met MVEE_ALLOW_SHM op ON:
Binary

MVEE.log
Binary

10:02
En use_ipmon is false
10:02
Dus het kan nog niet aan ipmon liggen :smile:

babrath
10:03 AM
hm
10:03
ik zou het proberen zonder ALLOW_SHM te doen
10:04
maar je kan de nginx config ook aanpassen om geen shared memory te proberen doen

Lennert Franssens
10:04 AM
Aah dat wist ik nog niet. Dan ga ik dat eens proberen :slightly_smiling_face:

babrath
10:05 AM
dat is in conf/nginx.conf
10:06
limit_req_zone en limit_conn_zone uit commentarieren?

Lennert Franssens
10:11 AM
Het geeft al een andere fout nu: nginx: [emerg] zero size shared memory zone "one"

babrath
10:14 AM
Lol

Lennert Franssens
10:15 AM
Ahnee ze stonden niet gecommentarieerd

babrath

10:15 AM
Ahja, hebt ge andere config opties die daar mee te maken hebdn ook aangepast?

Lennert Franssens
10:15 AM
en ik heb ze wel in commentaar gezet
10:15
daardoor komt het
10:16
Hij zoekt een lijn zoals die dat ik in commentaar heb gezet limit_req_zone $binary_remote_addr zone=one:10m rate=5000000r/s;

babrath
10:17 AM
hmm?

Lennert Franssens
10:18 AM
En zijn er nog manieren om die shared memory zone uit te zetten?
10:19
Bijvoorbeeld door geen meerdere processen in nginx te gebruiken?

babrath
10:22 AM
ja
10:22
uiteindelijk willen we liefst wel met meerdere processen meten, maar je kan met 1 beginnen

Lennert Franssens
10:29 AM
Dat lost het probleem ook niet op :confused:

babrath
10:29 AM
welke error krijg je?
10:29
stuur je config eens door?

Lennert Franssens
10:30 AM
nginx.conf
#user nobody;

worker_processes 2;
#error_log logs/error.log;
Click to expand inline (131 lines)

10:31
Daar had ik worker_processes al eens naar 1 gezet

babrath
  10:33 AM
en wat krijg je dan?

Lennert Franssens
  10:35 AM
Opnieuw dezelfde warning als in het begin:
WARNING: Trying to set up shared memory from a location other that mvee_shm_mmap
10:35
Binary

MVEE.log
Binary

babrath
  10:36 AM
en als je dan alles van limit_req en limit_conn uitcomment?

Lennert Franssens
  10:39 AM
Nu geeft het ook dezelfde error, daarstraks had ik de referenties daarnaar in de attribute location in de config file niet in commentaar gezet.

babrath
  10:40 AM
right
10:40
ok

Lennert Franssens
  10:41 AM
Ik weet niet of deze twee ook nog nuttig kunnen zijn?
Syntax: master_process on | off;

Default:
master_process on;
Context:	main
Determines whether worker processes are started. This directive is intended for nginx developers.

Syntax: multi_accept on | off;
Default:
multi_accept off;
Context:	events
If multi_accept is disabled, a worker process will accept one new connection at a time. Otherwise, a worker process will accept all new connections at a time.

babrath
  10:42 AM
je kan eens proberen met master_process off
10:43
en dan zou ik nog eens proberen met SHM niet te allowen

Lennert Franssens
  10:48 AM
Het werkt :partying_face::stuck_out_tongue_closed_eyes:

babrath
  10:48 AM
voor master_process off?

Lennert Franssens
  10:49 AM
Ja en SHM niet allowed. Ik moet het nu wel nog met ipmon testen... :sweat_smile:

babrath
  10:50 AM
ok
10:51
je kan daar alvast resultaten voor genereren
10:51
nginx met meerdere processen gaat ook wel interessant zijn, maar ik kan daar anders wel eens voor zien waarom dat niet direct werkt

Lennert Franssens
  10:52 AM
Met IP-MON werkt het ook :slightly_smiling_face:

babrath
  10:52 AM
Nice :slightly_smiling_face:

Lennert Franssens
  10:53 AM
Ik heb wel het gevoel door de aanpassingen die ik daar de laatste dagen nog in kunnen maken heb, dat het deel IP-MON wel op punt staat nu
10:53
Buiten het probleem met SHM, maar dat komt door een aanpassing in de MVEE
10:53
ik weet nog dat dat nog eens ergens stuk was en ik dat deel gewoon in commentaar heb gezet
10:53
Ik zal die commit nog eens zoeken
10:54
https://github.com/lennertfranssens/ReMon/commit/ff3671aff67a2a79eba000eb930c9163f42dfb8f#diff-718a4fba4a7760af48d6d1[…]18eb7fe15d34109ec66d3e5bR558
10:55
Dat zou wel eens het probleem met SHM kunnen zijn denk ik

babrath
  10:56 AM
welke lijn specifiek? :sweat_smile:

Lennert Franssens
  10:57 AM
558 in MVEE_syscalls.cpp
10:57
en 559 ook
11:01
Mag ik trouwens die 48-bits versie als "master" in mijn repo nemen? Aangezien die toch snel genoeg is, iets veiliger is, en meer random adressen heeft?
11:04
Ik ben nu even aan het testen met die twee lijnen terug uit commentaar en op het eerste zicht zijn er geen problemen
11:04
met echo test en ls
11:04
ik ga ook eens met nginx nu proberen
11:04
stel u voor dat die fout opgelost is door iets anders aan te passen in de MVEE de laatste maand :sweat_smile:
11:07
Het werkt nog altijd niet, maar die lijnen mogen terug uit commentaar want die veroorzaken geen fouten meer

babrath
  11:08 AM
48-bits als master lijkt mij goed
11:08
hm
11:08
welja
11:08
de instructie langswaar shared memory aangevraagd wordt is volgens mij ook geregistreerd in de MVEE
11:09
dus als die ergens anders ligt, omdat jij alle syscall instructies wrapped, dan is dat wss wat er kapot gaat
11:09
maar daar zou ik eens voor moeten kijken, dat is al even geleden

Lennert Franssens
  11:11 AM
Ja, nu opzich kan ik al verder met deze versie zonder shared memory. Het zal wel al een goed eerste beeld geven van de impact van de filter op zo'n programma

babrath
  11:11 AM
idd

Lennert Franssens
  11:28 AM
Waarom moeten we eigenlijk een ethernet verbinding hebben tussen client en server en mogen we niet op hetzelfde toestel werken? Ik ben er van overtuigd dat het met iets te maken heeft, maar weet gewoon niet wat :smile:

babrath
  11:46 AM
dat kan hoor, dat is gewoon geen realistische setup
11:46
als je op localhost werkt gaat je overhead ook veel groter zijn, omdat de network latency veel lager is
11:47
maar dat is niet realistisch :slightly_smiling_face:

Lennert Franssens
  11:47 AM
Ahja op die manier. Dan ga ik het ook via een gigabit verbinding doen

:slightly_smiling_face:

Lennert Franssens
  1:35 PM
Het was nog even zoeken en ik heb een paar syscalls naar cp-mon moeten sturen om het te laten werken, maar met de korte testjes die ik nu al heb gedaan (op localhost) zie ik dat de latency 0.8ms lager is dan zonder ip-mon. En dat op een latency van 3.7ms, is dat toch wel al een mooi verschil denk ik. Nu ga ik de testomgeving pushen en doe ik de testen tussen 2 computers.

babrath
  1:39 PM
ok
1:39
klinkt goed :slightly_smiling_face:

Lennert Franssens
  2:44 PM
Slecht nieuws
2:44
Alle redelijk slecht

babrath
  2:44 PM
Ja?

Lennert Franssens
  2:45 PM
Mijn netwerkkaart wordt niet ondersteund in kernel 5.4.0, wat ik nodig heb voor ipmon met de kernelpatch
2:45
of hebben jullie ook al een nieuwere patch?

babrath
  2:45 PM
hmm

Lennert Franssens
  2:46 PM
https://github.com/lennertfranssens/ReMon/blob/ipmon_48_bit_secret/patches/linux-5.10-full-ipmon-pku-assisted.patch
2:47
Ik zie die wel staan, maar kan het kwaad dat die pku-assisted daar ook staat?

babrath
  2:47 PM
ik denk van wel
2:47
maar eigk verschilt die niet zo hard van de 5.4-full-ipmon
2:47
wat mij doet denken dat het relevante deel van de linux src niet zo hard verandert is tussen 5.4 en 5.10
2:48
dus je kan proberen om de 5.4 patch te applyen op 5.10? Of whatever versie jouw netwerkkaart wel ondersteunt?

Lennert Franssens
  2:48 PM
Dus dan kan ik gewoon de 5.4 patch op een 5.10 kernel toepassen

babrath
  2:48 PM
idd (edited)

Lennert Franssens
  2:48 PM
Oke ik ga proberen :slightly_smiling_face:

babrath
  2:48 PM
dat kan wel tot conflicten leiden natuurlijk
2:48
maar die gaan wss gemakkelijk op te lossen zijn

Lennert Franssens
  2:48 PM
Dat is nu echt de laatste test die ik moest doen en dan ben ik klaar :smile:

babrath
  2:48 PM
is dat ook al voor multi-process ngixn met SHM? :stuck_out_tongue:

Lennert Franssens
  2:49 PM
Nog niet maar ik ga eerst mijn veslag afwerken de komende dagen en als er dan nog

tijd is, kan ik daar nog eens naar zoeken :slightly_smiling_face:

babrath
  2:50 PM
perfect
2:51
ik kan in tussentijd ook wel eens kijken daarvoor hoor, dat gaat wss sneller gaan

Lennert Franssens
  2:52 PM
Ja dat is ook een goed idee :smile: ik zal alles op punt zetten in de repo straks/vanavond en dan zou het probleem makkelijk repliceerbaar moeten zijn :slightly_smiling_face:
2:56
Ik heb een beter idee, de ethernet adapter (usb) die ik gebruik op mijn laptop werkt ook op mijn desktop met kernel 5.4.0. Ik weet dat mijn broer er ook een heeft, dus ik ga snel even naar hem thuis om die op te halen. Dan moeten we niet debuggen voor die kernelpatch :wink:

babrath
  2:58 PM
dat kan ook :slightly_smiling_face:

Lennert Franssens
  4:25 PM
Hoe moet ik die resultaten juist interpreteren? :slightly_smiling_face:
 > nginx native run, 1 worker
   > average latency: (1120.0+1120.0+1120.0+1130.0+1130.0)/5 us
   > average throughput: (8785.36+8797.89+8804.41+8762.56+8753.19)/5 requests/sec
 > nginx default ReMon run, 1 worker
   > average latency: (1620.0+1630.0+1630.0+1640.0+1620.0)/5 us
   > average throughput: (6164.95+6129.46+6104.13+6068.34+6161.14)/5 requests/sec
 > nginx ReMon with IP-MON with seccomp-BPF run, 1 worker
   > average latency: (1370.0+1360.0+1370.0+1370.0+1360.0)/5 us
   > average throughput: (7252.01+7292.58+7265.02+7267.9+7311.41)/5 requests/sec
 > nginx ReMon with IP-MON with kernel patch, 1 worker
   > average latency: (1120.0+1120.0+1120.0+1120.0+1120.0)/5 us
   > average throughput: (8787.86+8790.49+8787.78+8792.5+8798.03)/5 requests/sec
4:27
Mag ik gewoon een waarde per configuratie berekenen (het gemiddelde) om die dan in een grafiek te zetten?

babrath
  4:27 PM
ja

4:28
een gemiddelde met std dev mag ook, maar ze zijn precies redelijk gelijkaardig

Lennert Franssens
  4:28 PM
Dit resultaat is wel minder goed dan mijn microbenchmarks

babrath
  4:28 PM
ja idd
4:28
enig idee waarom?

Lennert Franssens
  4:29 PM
Omdat de set van doorgelaten systeemaanroepen kleiner is
4:29
Ik kan die nog verder finetunen
4:29
Maar er is een deel die niet werken met mijn implementatie en wel met de oude
4:31
Ik ga even uittesten welke nu juist wel en niet werken

babrath
  4:32 PM
ah
4:32
hm
4:32
ja
4:32
dat is idd wel belangrijk :slightly_smiling_face:

Lennert Franssens
  4:54 PM
Tiens, er zijn ondertussen meer systeemaanroepen in ipmon dan in de paper van meneer Volckaert staan. Waarschijnlijk is dat waarover ik gekeken heb :smile:

Lennert Franssens
  7:00 PM
Dit is de lijst van alle systeemaarnoepen die mogelijk zijn in de originele versie van ipmon. In commentaar staat soms TRACE om aan te geven dat die systeemaanroepen voor nginx sowieso naar CP-MON moeten, ook in de originele implementatie van IP-MON. En bij systeemaanroepen die ervoor zorgen dat nginx niet werkt, staat in commentaar

CAUSES ERROR.
7:00
"__NR_getegid",
"__NR_geteuid",
"__NR_getgid",
"__NR_getpgrp",
"__NR_getppid",
"__NR_gettid",
"__NR_getuid",
"__NR_getpid",
"__NR_gettimeofday",
"__NR_time",
"__NR_clock_gettime",
"__NR_sched_yield",
"__NR_getcwd",
"__NR_uname",
"__NR_getpriority",
"__NR_nanosleep",
"__NR_getrusage",
"__NR_sysinfo",
"__NR_times",
"__NR_capget",
"__NR_getitimer",
"__NR_access",
"__NR_faccessat",
"__NR_stat",
"__NR_lstat",
"__NR_fstat", // CAUSES ERROR
"__NR_newfstatat",
"__NR_getdents",
"__NR_readlink",
"__NR_readlinkat",
"__NR_getxattr",
"__NR_lgetxattr",
"__NR_fgetxattr",
"__NR_lseek",
"__NR_alarm",
"__NR_setitimer",
"__NR_timerfd_gettime",
"__NR_madvise",
"__NR_fadvise64",
"__NR_read", // CAUSES ERROR
"__NR_readv", // CAUSES ERROR
"__NR_pread64", // CAUSES ERROR
"__NR_preadv", // CAUSES ERROR
"__NR_select",
"__NR_poll",
"__NR_ioctl",
"__NR_futex", // TRACE
"__NR_timerfd_settime",

```
"__NR_sync",
"__NR_fsync",
"__NR_fdatasync",
"__NR_syncfs",
"__NR_write",
"__NR_writev",
"__NR_pwrite64", // CAUSES ERROR
"__NR_pwritev",
"__NR_epoll_wait",
"__NR_recvfrom",
"__NR_recvmsg", // TRACE
"__NR_recvmmsg",
"__NR_getsockname",
"__NR_getpeername",
"__NR_getsockopt",
"__NR_sendto",
"__NR_sendmsg", // TRACE
"__NR_sendmmsg",
"__NR_sendfile",
"__NR_shutdown",
"__NR_setsockopt",
"__NR_epoll_ctl",
"__NR_mmap", // TRACE
"__NR_munmap", // TRACE
"__NR_mremap",
"__NR_mprotect", // CAUSES ERROR
"__NR_brk", // CAUSES ERROR
"__NR_open", // CAUSES ERROR
"__NR_openat", // CAUSES ERROR
"__NR_close", // CAUSES ERROR
"__NR_fcntl",
"__NR_dup",
"__NR_dup2",
"__NR_dup3",
"__NR_pipe",
"__NR_pipe2",
"__NR_inotify_init",
"__NR_inotify_init1",
"__NR_chdir",
"__NR_fchdir",
"__NR_mkdir",
"__NR_socket",
"__NR_socketpair",
"__NR_bind",
"__NR_connect",
"__NR_listen",
"__NR_accept4",
"__NR_accept",
"__NR_epoll_create",
"__NR_epoll_create1"
```

babrath
8:32 PM
tiens
8:32
is die lijst niet ongeveer af te leiden uit de src van de in-kernel broker?

Lennert Franssens
8:33 PM
Dat is de lijst van de IK broker maar syscalls zoals open, brk... zorgen voor problemen

babrath
8:35 PM
huh
8:35
vreemd
8:35
en wat voor errors?
8:35
en is de performance op deze manier beter?

Lennert Franssens
8:35 PM
om te kunnen vergelijken ga ik nu de benchmark nog eens laten lopen op de versie met de kernel patch, met een aangepaste versie van ip-mon waarin dezelfde systeemaanroepen worden doorgelaten als in mijn filter

1 reply
29 days agoView thread

Lennert Franssens
8:36 PM
En de performance is een beetje beter, maar niet heel veel
8:37
En de errors zijn verschillend per syscall, maar vaak zijn het segmentation faults en synchronisatieproblemen.

babrath
8:40 PM
replied to a thread:
om te kunnen vergelijken ga ik nu de benchmark nog eens laten lopen op de versie met de kernel patch, met een aangepaste versie van ip-mon waarin dezelfde systeemaanroepen worden doorgelaten als in mijn filter

dat is wel enigszins interessant, maar ook geen eerlijke vergelijking.
8:40
Welja
8:41
heb je enig idee wat die segfaults zou kunnen veroorzaken? Het is jammer dat het een beetje een open vraag is nu of het merendeel van die syscalls alsnog gesupport zou kunnen worden of niet
8:41
maar dat is ook weer debugwerk, dus ik zou eerst verder werken aan je tekst enzo :slightly_smiling_face:

Lennert Franssens
8:42 PM
Oke, ik neem het nog mee in mijn todo om te doen na de tekst :slightly_smiling_face:

Lennert Franssens
9:54 PM
Ik heb alles nog eens opnieuw laten lopen en nu krijg ik deze resultaten:
9:54
> nginx native run, 1 worker
> average latency: (460.96+456.33+457.26+466.98+462.18)/5 us
> average throughput: (21237.95+21572.84+21524.6+21325.13+21276.2)/5 requests/sec
> nginx default ReMon run, 1 worker
> average latency: (1620.0+1610.0+1630.0+1640.0+1630.0)/5 us
> average throughput: (6143.54+6191.64+6127.39+6107.12+6119.49)/5 requests/sec
> nginx ReMon with IP-MON with seccomp-BPF implementation run, 1 worker
> average latency: (1150.0+1150.0+1140.0+1140.0+1140.0)/5 us
> average throughput: (8638.27+8641.35+8721.27+8736.54+8774.96)/5 requests/sec
> nginx ReMon with original IP-MON with kernel patch, 1 worker
> average latency: (464.59+459.83+462.13+461.25+464.29)/5 us
> average throughput: (21202.37+21403.59+21213.09+21276.54+21178.18)/5 requests/sec
> nginx ReMon with original IP-MON with kernel patch and equal set as seccomp-BPF implementation, 1 worker
> average latency: (1050.0+1080.0+1070.0+1050.0+1060.0)/5 us
> average throughput: (9503.93+9278.05+9301.38+9490.33+9358.96)/5 requests/sec (edited)

babrath
10:11 PM
Dan is het verschil dus vnl te verklaren door die (nog?) niet ondersteunde system calls

Lennert Franssens
10:11 PM
Ja, klopt

babrath
  10:11 PM
Tis wel een groot verschil :sweat_smile:

Lennert Franssens
  10:11 PM
Ja het zijn natuurlijk redelijk cruciale system calls die nog niet werken
10:12
"__NR_fstat", // CAUSES ERROR
"__NR_read", // CAUSES ERROR
"__NR_readv", // CAUSES ERROR
"__NR_pread64", // CAUSES ERROR
"__NR_preadv", // CAUSES ERROR
"__NR_pwrite64", // CAUSES ERROR
"__NR_mprotect", // CAUSES ERROR
"__NR_brk", // CAUSES ERROR
"__NR_open", // CAUSES ERROR
"__NR_openat", // CAUSES ERROR
"__NR_close", // CAUSES ERROR
10:16
Maar zoals bij de originele versie zijn dat wel zaken die nog geimplementeerd kunnen worden. Hoewel ik niet denk dat dat mij nog gaat lukken. En ookal hadden we dat een paar weken eerder geweten, dan denk ik ook dat dat iets was geweest dat mij niet echt gelukt was door de complexiteit in het algemeen. Ik wil maar zeggen, nu heb ik wel door hoe het een en ander werkt in de MVEE en hoe IP-MON met seccomp omgaat. Maar om die systeemaanroepen werkend te krijgen is er toch nog wat meer inzicht en ervaring nodig volgens mij. Langs de andere kant, we zien wel dat er potentieel gebruik in zit en dat met eenzelfde configuratie resultaten in dezelfde grootteorde behaald worden, wat wel perspectief biedt :slightly_smiling_face:
10:17
En het gaat ook slechts voor een beperkte set van applicaties zeker werken aangezien veel applicaties tegenwoordig zelf met seccomp-BPF filters werken. En vanaf ze gecombineerd worden, loopt het mis :confused: (edited)
10:19
Maarja, ik heb genoeg resultaten om nog dingen over te schrijven. En ik denk dat ik ook nog wel een hoofdstuk future work kan neerschrijven met wat we nu de afgelopen dagen nog te weten zijn gekomen :smile:

babrath
  8:10 AM
Dat zijn inderdaad allemaal dingen om over te schrijven :slightly_smiling_face:

Lennert Franssens
  1:51 PM
Kunnen we dinsdag nog eens bellen in de namiddag over wat juist in mijn abstract

moet? Ik ben de laatste hand aan het leggen aan mijn scriptie. Nog 2 figuren die ik moet maken, mijn conclusie en verder onderzoek. De rest is af, dus ik kan normaal tegen vanavond mijn "definitieve versie" doorsturen :slightly_smiling_face:

1 reply
23 days agoView thread

Lennert Franssens
  1:53 PM
Ik zit wel nog steeds met het probleem dat ik nog niet aan 50 bladzijden zit. Maar na het te herlezen vind ik het wel een coherent geheel waar alles duidelijk in staat. Dus ik vraag me af of die 50 bladzijden een must is, als het kwalitatief ook goed en duidelijk is met een paar bladzijden minder?

1 reply
23 days agoView thread

Lennert Franssens
  12:01 AM
Ik heb de figuren geschetst en moet ze nog in visio maken, maar lang gaat dat morgen niet meer duren :slightly_smiling_face:

Lennert Franssens
  2:31 PM
Ik ben thuis, heeft u nu tijd?

babrath
  2:40 PM
in 5m? :slightly_smiling_face:

Lennert Franssens
  2:40 PM
Ja, dat is goed :slightly_smiling_face:

Lennert Franssens
  12:11 AM
Het extended abstract is af maar voor ik het doorstuur ga ik het morgenvroeg nog eens nalezen op spelfouten :wink:
12:12
Mijn nederlandstalige samenvatting en mijn abstract (ik gebruik hetzelfde als dat wat ik in het begin van mijn extended abstract heb gezet) zijn ook af

Lennert Franssens

8:47 AM
PDF

abstract.pdf
PDF

Lennert Franssens
8:49 AM
Als abstract op plato zou ik hetzelfde nemen als het abstract in het begin van mijn extended abstract. En dit is mijn samenvatting in het Nederlands:
Multi Variant Execution (MVX) systemen bieden een techniek voor het detecteren van veiligheidsproblemen met betrekking tot geheugencorruptie in software. Traditionele MVX-systemen zijn zeer veilig, maar werken relatief langzaam. Relaxed Monitoring (ReMon) is een ontwerp dat de uitvoering van MVX-systemen effectief versnelt. Een kernel-patch werd gebruikt om dat ontwerp te implementeren. Veel mensen geven er de voorkeur aan om geen kernelpatch op hun systeem toe te passen. Dat vermindert het gebruik van ReMon.

In dit proefschrift onderzoeken we of het gebruik van een kernelpatch bij de implementatie van het ReMon-ontwerp kan worden vervangen door moderne OS-extensies. Om de kernelpatch te vervangen maken we gebruik van seccomp-BPF filters. We passen het ontwerp aan en implementeren het. Door snelheidsmetingen te doen zien we een duidelijk potentieel voor het gebruik van het nieuw ontwerp. Verder onderzoek kan gedaan worden naar het verder uitdiepen van de ondersteunde functies en de daarbij horende veiligheid. Er kan ook nog verder onderzoek gedaan worden naar het gecombineerd gebruik van de nieuwe implementatie van ReMon met door ReMon te analyseren software dat ook gebruik maakt van de nieuwe seccomp-BPF technologie.

2 replies
Last reply 20 days agoView thread

babrath
2:45 PM
feedback op je EA
PDF

abstract_feedback.pdf
PDF

:+1:
1

babrath

2:46 PM
replied to a thread:
Als abstract op plato zou ik hetzelfde nemen als het abstract in het begin van mijn extended abstract. En dit is mijn samenvatting in het Nederlands:…
mits wat extra conclusie (zie feedback) in dat abstract is dat voor mij OK
View newer replies

babrath
2:47 PM
replied to a thread:
Als abstract op plato zou ik hetzelfde nemen als het abstract in het begin van mijn extended abstract. En dit is mijn samenvatting in het Nederlands:…
samenvatting is goed voor mij

Lennert Franssens
4:46 PM
Moeten de hoofdstukken altijd op een rechterpagina beginnen?

babrath
4:47 PM
nee
4:47
maar het mag wel, dat kan mooier ogen
4:47
allee, het moet niet voor zover ik weet, tenzij het toch in de modaliteiten staat zonder dat ik het weet :stuck_out_tongue:

Lennert Franssens
4:48 PM
neen ik heb het ook nog nergens zien staan maar ik weet nog dat mijn broer dat ook had gedaan bij zijn masterproef dus ik vroeg het mij af of het nu wel of niet moest :sweat_smile: maar ik ga het wel doen want het oogt inderdaad wel mooier zo :slightly_smiling_face:
4:48
en moeten jullie een afgedrukte versie hebben?

babrath
4:55 PM
nee :slightly_smiling_face:

Lennert Franssens
4:59 PM
Oke :slightly_smiling_face:

Lennert Franssens
  9:28 PM
Moeten de repo's publiek gemaakt worden zodat bijvoorbeeld ook professor De Sutter de code kan bekijken?

babrath
  9:32 PM
Nee hoor, wij kijken :smile:

Lennert Franssens
  9:32 PM
Ah oké :smile: