

Checkpoint/Restore in Fuzzing

Malte Klaassen

17.11.2018

Zusammenfassung

TODO Fuzzing von komplexen Programmen als ganzes ist meist langsam, Fuzzing komplexer Programme in vielen kleinen Einzelblöcken ist aufwändig für den Tester da für sie jeweils die Fuzz-Targets geschrieben werden müssen. In diesem Stück Papier untersuchen wir die Nutzbarkeit von Checkpoint/Restore-Mechanismen zur Lösung solcher Probleme im Fuzzing, insb. mit Blick auf Netzwerkprotokolle mit einer wohldefinierten State Machine, präsentieren eine mögliche Architektur zum Fuzzern von Implementierungen ebensolcher Protokolle und untersuchen verschiedene Fuzzer und Checkpoint/Restore Tools auf ihre Anwendbarkeit.

Inhaltsverzeichnis

1	Einleitung	2
1.1	Fuzzing	3
1.2	Klassifizierung von Fuzzern	3
1.3	Historischer Überblick	3
1.4	Beschränkungen im Fuzzing	4
2	Checkpoint/Restore	4
2.1	C/R Typen und Unterscheidungen	4
2.1.1	Nach Typ	4
2.1.2	Nach Capabilities	6
3	Fuzzing mit C/R	6
3.1	C/R-Fuzzer	7
3.2	Non-C/R-Fuzzer mit C/R Ansätze	9
3.2.1	Naiv	9
3.2.2	Exploration des Zustandsgraphen in Client/Server Systemen	11
4	CRIU	14
4.1	Nutzung	14
4.2	Funktionsweise	15
4.3	Capabilities	17
4.4	Probleme	17
4.5	Performance	18

5	TODO: CRIU + Fuzzing Engines	21
5.1	Fuzzing-Engines	21
5.1.1	Funktionsweise: libfuzzer	21
5.1.2	afl	21
5.2	Fuzzing-Engines und C/R-Fuzzing	21
5.2.1	libfuzzer	21
5.2.2	afl	22
5.3	Fazit	22
6	Ausblick	23
7	Anhang	24

1 Einleitung

Mit zunehmender Prävalenz und Prominenz von IT-Systemen ist das Thema der Sicherheit eben dieser Systeme zunehmend in den Fokus gerückt, sowohl in Fällen häufig fehlerbehafteter Nutzung von Software in denen Nutzer selbst, direkt betroffen sind wie bei Hacks oder Leaks bei Online-Plattformen wie Facebook oder Malware-Kampagnen wie WannaCry, häufig ausgelöst durch technische Sicherheitslücken und verzögerte Patches, als auch vereinzelt in Fällen von rein technischen Sicherheitslücken wie beispielsweise Heartbleed.

Proaktiv vs Reaktiv

Über die Jahrzehnte haben sich dabei verschiedene Ansätze und Methoden zur proaktiven Erkennung, Behebung und Verhinderung von ebensolchen Sicherheitsproblemen entwickelt, beispielsweise:

- Formulierung von Sicherheitsanforderungen bereits während und vor der Designphase
- Formale Beweise beispielsweise bei kryptografischen oder Netzwerk-Protokollen oder sogar von Programmen, insbesondere in der Funktionalen Programmierung
- Compilerchecks und -safeguards sowie statische Codeanalyse
- (Externe) Reviews der Software und Penetrationstesting
- Klassisches Sicherheitstesting

Leider sind nicht immer alle diese Methoden anwendbar und haben jeweils ihre eigenen Stärken und Schwächen. So sind Formale Beweise der Sicherheit von Programmen außerhalb von einigen modernen Hochsprachen häufig nicht möglich, klassisches Sicherheitstesting ist zumeist beschränkt auf die Vorstellungskraft des Testers beziehungsweise auf sein Verständnis der zu testenden Software, Grenzfällen oder komplett unerwartete aber mögliche Eingaben sind mit klassischem Testing schwer umfassend abzudecken.

1.1 Fuzzing

TODO: Corpus, Seeds, Dictionaries

Fuzzing ist eine Test-Methode die teil-automatisiert versucht genau diese Art von mit klassischen Test-Methoden schwer abzudeckenden Fällen umfassend zu erkennen indem, statt wie im klassischen Testing nur in von Menschen mehr oder weniger präzise beschriebenen Testfällen getestet wird, ein Codestück, das so genannte Fuzz-Target, wiederholt mit computergenerierten zufälligen Inputs getestet wird, dabei wird die Ausführung genau von Fuzzer überwacht um beispielsweise Programmabstürze, das Lesen aus initialisiertem Speicher oder auch undefiniertes Verhalten zu Erkennen und zu der problematischen Code-stelle zurück zu verfolgen.

1.2 Klassifizierung von Fuzzern

Fuzzer werden anhand mehrerer Eigenschaften unterschieden, insbesondere ob sie sich beim Testen der Struktur des Programmes und der Struktur der erwarteten Eingaben bewusst sind. Hierbei

- Wenn der Fuzzer sich der Struktur des Programmes bewusst ist und dieses Wissen im Testing nutzt um bessere Code-Coverage zu erreichen spricht man, je nach Technologie, von einem White-Box- oder Gray-Box-Fuzzer. Besitzt der Fuzzer keine Kenntnisse über die Programmstruktur und generiert programmunabhängige Zufallsinputs spricht man von einem Black-Box-Fuzzer.
- Wenn der Fuzzer sich der erwarteten Struktur des Inputs, beispielsweise bei einem bestimmten Dateiformat oder Netzwerkprotokoll, bewusst ist spricht man von einem Smart Fuzzer, ansonsten von einem Dumb Fuzzer.

Die meisten modernen Fuzzer können sowohl als Black- als auch als Gray- beziehungsweise White-Box-Fuzzer und sowohl als Smart als auch als Dumb Fuzzer agieren, diese Terme werden zumeist auch zur Beschreibung von Operationsmodi ebensolcher moderner Fuzzer verwendet.

1.3 Historischer Überblick

Testing mit randomisierten Inputs ist kein sonderlicher neuer Ansatz, jedoch ist die Effizienz von naivem Random Input Test (also Black-Box-Fuzzing) offensichtlich beschränkt und der größte Teil der Fehler die auf diese Art und Weise gefunden werden können relativ simpel und selten in den Tiefen des Programmes versteckt, das heißt es handelt sich zumeist um Fehler die vergleichsweise simpel auch durch Code-Reviews oder klassische Tests gefunden werden könnten, entsprechenden hatte Random Input Testing für recht lange Zeit nur eine relativ geringe Relevanz, sowohl in der Qualitätssicherung als auch in der Sicherheitsforschung.

Mit dem Aufkommen von effizienteren Gray- und White-Box-Fuzzern und zunehmender Relevanz von IT-Sicherheit in den späten 2000ern entwickelte sich ein breiteres Interesse an Fuzzing, inklusive einiger Open Source Fuzzern und Fuzzing Frameworks, hin zu Fuzzing als Kernbestandteil des Sicherheits-Testings vieler relevanter Open Source Software Projekte, vorrangetrieben insbesondere

durch Projekte wie Googles OSS-Fuzz, ein Service zum kontinuierlichen Testen verschiedenster Open Source Software durch mehrere verschiedene Fuzzer.

1.4 Beschränkungen im Fuzzing

Leider ist natürlich auch Fuzzing nicht ohne Probleme und Beschränkungen. Da Fuzzing zu großen Teilen darauf beruht eine große Menge an Eingaben zu Testen sinkt die Effizienz des Fuzzings deutlich wenn die Ausführung des Fuzz-Targets zu lange dauert. Dies führt dazu, dass das Fuzzten komplizierter, insbesondere interaktiver Systeme nicht ohne weiteres effizient möglich ist. Dieses Problem wird zumeist gelöst indem man komplexe Systeme in kleinere, simple Subsysteme aufteilt und diese unabhängig voneinander fuzzt und so beispielsweise unnötigen Overhead im Setup des Fuzz-Targets vermeidet. Aber auch dies hat seine Nachteile: Der Tester muss hier nun mit der Struktur des zu testenden Systems bekannt sein, muss das komplexe System in kleinere Systeme selbst zerlegen und dann für jedes der Subsysteme einzelne Fuzz-Targets schreiben - dies ist zeit- und arbeitsaufwändig und, ähnlich wie beim klassischen Testing, anfällig für Fehler des Testers.

2 Checkpoint/Restore

Mit Checkpoint/Restore, kurz C/R, manchmal auch Checkpoint/Restart, bezeichnet man das Speichern eines Prozesses oder einer Anwendung in einer Art und Weise die eine spätere Wiederherstellung und weitere Ausführung der Anwendung erlaubt. C/R-Systeme finden Anwendung in verschiedenen Bereichen, beispielsweise im Debugging und Fault Recovery durch das Zurücksetzen eines Programmes auf einen vorherigen Zustand, in der Beschleunigung oder Optimierung von Prozessen durch das Fortsetzen von Anwendungen von Checkpoints anstatt rechenzeitaufwändigen Komplettausführungen und dem Aussetzen der Ausführung von Programmen wenn diese gerade nicht benötigt werden oder, je nach Fähigkeiten des zugrundeliegenden C/R-Systems, sogar die Migration von laufenden Anwendungen zwischen Maschinen.

TODO: Was alles muss gespeichert werden?

2.1 C/R Typen und Unterscheidungen

Für das Checkpointing/Restoring von Anwendungen gibt es nicht einen einzelnen Standard sondern eine ganze Reihe von Modellen die wir grob nach zwei Faktoren klassifizieren: Typ bzw. Implementierungsebene und nach den Fähigkeiten des Checkpoint/Restore-Systems, also welche Aspekte des Programms genau wiederhergestellt werden können und welche nicht.

2.1.1 Nach Typ

Bei der Klassifizierung nach Typ oder Implementierungsebene unterscheiden wir grob nach drei Kategorien: Das Checkpointen im Userspace, aus dem zu checkpointenen Programm selbst, das Checkpointen auf Kernelebene und das Checkpointen von Virtuellen Maschinen beziehungsweise Containern.

Im klassischen Userspacecheckpointing findet das Checkpointing größtenteils

transparent gegenüber dem Kernel statt, das heißt das C/R-Tool nutzt aus dem Userspace zugängliche Ressourcen wie `/proc/` oder Intercepts von Library- oder Syscalls zum sammeln der notwendigen Checkpoint-Informationen und ebenfalls nur ohne besondere Privilegien zugängliche Ressourcen zur Wiederherstellung. Dieses Vorgehen führt in reinen Userspace-C/R-Systemen wie DMTCP (Distributed MultiThreaded CheckPointing) unter Umständen zu einigen Fallstricken oder Problemen, so können zum Beispiel nicht alle Eigenschaften eines Programmes, beispielsweise die PID, nicht ohne weiteres korrekt wiederhergestellt werden, die Intercepts benötigen alternative Libraries die spätestens zur Startzeit mit der Anwendung gelinkt werden müssen, häufig Beschränkungen darauf haben mit welchen Programmen (bspw. Sprachen) sie genutzt werden können und unter Umständen die Funktionsweise des Programmes ändern könnten sowie mögliche Performanceeinbußen durch das duplizieren von Kernelstrukturen im Userspace.

Dem gegenüber steht das Checkpointing auf Kernelebene, beispielsweise BLCR (Berkeley Lab's Linux Checkpoint/Restart) oder Linux-C/R. C/R-Systeme nutzen den Kernel direkt um alle nötigen Informationen während des Checkpointings zu sammeln, benötigen also keine Libraries zum Sammeln eben dieser Informationen durch das Abfangen von Syscalls o.ä., und haben alle nötigen Privilegien um, theoretisch, falls implementiert, alle Eigenschaften der Anwendung auch korrekt wiederherzustellen. Allerdings muss der Kernel eben diese Funktionalitäten anbieten damit Kernelebene-C/R-Tools eingesetzt werden können und obwohl es wiederholt versuche der Integration von C/R-Tools in den Mainline Linux Kernel gab ist dies bisher nicht geschehen, das heißt es wird zur Nutzung ein gepatchter Kernel oder ein Kernel mit zusätzlichen Modulen benötigt. Das Checkpointing mit Kernelebene-C/R-Tools folgt zumeist einem Schema ähnlich diesem:

1. Einfrieren und Synchronisieren der Prozesse und Threads
2. Erfassen von globalen Informationen (d.h. Informationen die nicht direkt Teil der zu checkpointenden Anwendung sind wie Namespaces oder Container)
3. Erfassen der Prozess- bzw. Thread-Hierarchien und -Informationen wie Parent/Child-Verhältnisse, PIDs, ...
4. Erfassen des Statuses der einzelnen Prozesse und Threads, bspw. Signals, geöffnete Files und FDs, ...
5. Beenden oder Weiterausführung der Anwendung, Schreiben all der erfassten Informationen

Dabei werden all diese Schritte natürlich nicht von der zu checkpointenden Anwendung selbst vorgenommen sondern von einem dafür gespawnten externen Prozess. Das Wiederherstellen funktioniert generell in ähnlicher Weise:

1. Wiederherstellen der globalen Information, beispielsweise Container bei Multiprocess-Anwendungen oder dem Prozess bei Multithreaded aber Single-process Anwendungen, wenn nötig

2. Erstellen einer Prozess- bzw. Threadhierarchie entsprechend der erfassten Hierarchie (in eingefrorenem Zustand)
3. Wiederherstellung des Statuses der einzelnen Prozesse und Threads
4. "Fortsetzen" beziehungsweise Auftauen der Prozesse/Anwendung

CRIU (Checkpoint/Restore in Userspace) ist ein Hybrid aus Userspace- und Kernelebene-Systemen, das heißt es agiert größtenteils in Userspace, nutzt jedoch Calls auf privilegierte Operationen wie das Forken mit bestimmten PIDs. All diese von CRIU benötigten Kernel-Capabilities sind seit einiger Zeit im Linux Mainline-Kernel vorhanden. Das Checkpointing und Restoring in CRIU ähnelt stark dem der Kernelebene-C/R-Systeme, benötigt insbesondere keine Intercept-Libraries wie normale Userspace-C/R-Tools, jedoch können nicht all diese Informationen von externen Programmen erfasst oder hergestellt werden. Dies umgeht CRIU indem es mittels ptrace sogenannten parasite code in die zu checkpointenen Prozesse einfügt (beziehungsweise den restorer blob bei der Wiederherstellung) die eben diese Informationen erfassen.

Das Checkpointen von Virtuellen Maschinen oder Containern nutzt zumeist einen der obigen Ansätze um die laufenden Virtuelle Maschine (als Anwendung auf dem Host-System) als ganzes zu Checkpointen. So nutzt Docker beispielsweise CRIU, verwaltet durch entsprechenden Aufrufe des docker Kontrollprogramms `docker checkpoint create [...]` und `docker start --checkpoint [...]`.

2.1.2 Nach Capabilities

Zusätzlich unterscheiden sich verschiedene C/R-Tools auch noch danach was genau sie Checkpointen und Restoren können. Einige Tools wie frühere BLCR-Versionen unterstützen nur einzelne multi- oder singlethread Prozesse, Linux-CR unterstützt das Checkpointen beliebiger Container oder Subtrees von Prozessen. Tools die innerhalb von multithreaded Prozessen arbeiten, beispielsweise einzelne Threads checkpointen und wiederherstellen können existieren anscheinend nicht, zumindest fanden wir keine. Dies ist auch plausibel, da solche Threads sich eben Ressourcen teilen und daher ein teilweises Wiederherstellen nur einiger Threads zu undefiniertem Verhalten führen würde.

Weitere Unterschiede gibt es in der Frage welche Eigenschaften von laufenden Programmen wiederhergestellt werden können, beispielsweise Netzwerkverbindungen wie TCP- oder UDP-Sockets, Dateien oder Namespaces, sowie bezüglich der Frage ob beliebige Programme gecheckpointed werden können oder ob bereits beim Start der Anwendung bekannt sein muss, dass diese gecheckpointed werden soll und entsprechende Vorbereitungen wie das Preloading von Libraries oder das Informieren des Kernels getroffen werden müssen.

3 Fuzzing mit C/R

Wie oben erwähnt gibt es im klassischen Fuzzing einige Probleme: Komplexe Systeme oder große Overhead im Set-Up des Fuzztargets führen zur Reduktion des Durchsatzes und damit Einschränkung der Effizienz, mehrstufige interaktive Programme wie beispielsweise Netzwerkprotokolle sind als ganzes nur mit

großem Aufwand oder sogar gar nicht vernünftig zu fuzzen. In dieser Arbeit wollen wir untersuchen, ob sich einige oder sogar alle dieser Probleme durch die Nutzung von Checkpoint/Restore-Mechanismen beheben oder zumindest in ihren Auswirkungen reduzieren lassen. Dazu betrachten wir zunächst einmal generell mögliche Ansätze und ihre theoretischen Vor- und Nachteile bevor wir ihre Implementierbarkeit mit kontemporären Fuzzern und Checkpoint/Restore-Systemen betrachten.

Generell sehen wir zwei Ansätze zur Integration von Checkpoint/Restore-Tools in Fuzzern:

1. Integration des C/R-Tools in den Fuzzer selbst, das heißt der Fuzzer selbst übernimmt das Checkpointing und Restoring
2. Implementierung des Checkpointing und Restoring im Fuzztarget, transparent zum Fuzzer

Die erste Alternative, ein C/R-Fuzzer, wäre aber natürlich eine sauberere Lösung als die ad hoc Lösung der Implementierung im Fuzztarget und bietet die selben oder sogar mehr Möglichkeiten wie die Fuzztarget-Lösung, benötigt aber natürlich eben einen C/R-Fuzzer, etwas das unseren Nachforschung nach bisher nicht einmal in Ansätzen existiert.

3.1 C/R-Fuzzer

Klassische moderne Fuzzer wie libfuzzer oder afl folgen während des Fuzzings im Allgemeinen diesem Aufbau:

Listing 1: Struktur Klassischer Non-C/R-Fuzzer

1	Bis ein Problem auftritt:
2	Generiere einen neuen Input (aus dem Seed, Corpus, Dictionary Regeln)
3	Führe das Fuzz-Target mit diesem Input aus
4	Analysiere das Fuzz-Target während der Ausführung und update den Corpus wenn nötig

Das heißt das Fuzz-Target wird so oft mit verschiedenen Inputs ausgeführt bis ein Problem auftritt. Dabei hat das Fuzz-Target in Normalfall eine Form wie diese (oder kann in diese Form gebracht werden):

Listing 2: Struktur Klassisches Fuzztarget

1	[Inputunabhängiges Set-Up]
2	[Inputabhängiges Set-Up]
3	[Zu testender Code oder Funktion]

Hierbei ist das Ergebnis von 1. im Allgemeinen entweder konstant oder zumindest austauschbar wenn es beispielsweise externe (pseudo-)zufällige Einflüsse wie Nutzung der Systemzeit oder Aufrufe von Zufallszahlgeneratoren gibt. Dies heißt aber nun, dass wir das Fuzz-Target aufteilen können wie folgt:

1. Inputunabhängiges Set-Up
2. Inputabhängiges Set-Up und zu testende Funktionalität

Wenn wir dies machen können wir nun als Alternative zum klassischen Fuzzer einen C/R-Fuzzer nutzen, der folgendem Aufbau folgt:

Listing 3: Struktur einfacher C/R-Fuzzer

1	Führe den Inputunabhängigen Teil des Fuzz-Targets aus und checkpointe ihn
2	Bis ein Problem auftritt:
3	Generiere einen neuen Input (aus dem Seed, Corpus, Dictionary Regeln)
4	Restore das Fuzz-Target, übergebe diesen Input und führe den Inputabhängigen Teil des Fuzz-Targets aus
5	Analysiere den Input-Abhängigen Teil des Fuzz-Targets während der Ausführung und update den Corpus wenn nötig

Hierbei müssen natürlich ein paar Dinge beachtet werden: Die Übergabe des Inputs muss während der Laufzeit beispielsweise über stdin erfolgen, eine Übergabe des Inputs durch Parameter zum Zeitpunkt des Aufrufes wie beispielsweise bei libfuzzer ist nicht ohne weiteres möglich. Das Aufteilen des Fuzz-Targets heißt nicht das Aufteilen in mehrere unabhängige Code-Segmente - damit Checkpoint/Restore ohne Modifikationen funktionieren kann muss der inputabhängige Teil bereits von Anfang an im Fuzz-Target vorhanden sein. Ein solches Fuzz-Target könnte dann wie folgt aussehen:

Listing 4: Struktur C/R-Fuzzer-Fuzztarget

1	[Inputunabhängiges Setup]
2	Breakpoint
3	[Inputabhängiges Setup]
4	[Zu testender Code oder Funktion]

Hierbei hält der Breakpoint die Ausführung des Fuzztargets an und informiert den Fuzzer darüber, dass er an dieser Stelle das Fuzz-Target checkpointen sollte. Ein solcher C/R-Fuzzer kann unter Umständen das Problem großen Overheads im Setup lösen oder zumindest reduzieren. Mit T_{Setup} die Laufzeit des inputunabhängigen Setups, $T_{Execution}$ die (mittlere) Laufzeit des inputabhängigen Setups und des zu testenden Codes, $T_{Checkpoint}$ und $T_{Restore}$ die benötigte Zeit für eine Checkpointing- beziehungsweise Restoreoperation. Dann ergibt sich für die Laufzeit eines normalen, nicht-C/R Fuzzers bei n benötigten Operationen

$$T_{Non-C/R}(n) = n(T_{Setup} + T_{Execution}) \quad (1)$$

und für einen C/R-Fuzzer wie oben beschrieben

$$T_{C/R}(n) = T_{Setup} + T_{Checkpoint} + n(T_{Restore} + T_{Execute}) \quad (2)$$

Für große n und $T_{Restore} < T_{Setup}$ ergibt sich hier nun also eine Zeitersparnis pro Iteration:

$$\begin{aligned} \frac{T_{Non-C/R}(n) - T_{C/R}(n)}{n} &= \frac{(n-1)T_{Setup} - T_{Checkpoint} - nT_{Restore}}{n} \\ &=_{n \rightarrow \infty} T_{Setup} - T_{Restore} \end{aligned} \quad (3)$$

T_{Setup} ist dabei natürlich vom jeweilig zu testenden Programm abhängig, $T_{Restore}$ von mehreren Faktoren, unter anderem dem Programm selbst, insbesondere die Größe im Speicher und die zu checkpointenen Features wie Netzwerksockets, und natürlich das genutzte Checkpoint/Restore-Tool, eine genauere Betrachtung davon findet in Kapitel CRIU.Performance statt.

3.2 Non-C/R-Fuzzer mit C/R Ansätze

Ein anderer Ansatz zur Kombination von Fuzzing mit Checkpoint/Restore-Mechanismen wäre, statt des impliziten Checkpointing durch den Fuzzer, explizites Checkpointing und Restoring im Fuzztarget. Dies hätte den Vorteil einer größeren Flexibilität im Einsatz und würde im Idealfall keinerlei Änderungen am Fuzzer benötigen, könnte also unter Umständen bereits mit vorhandenen Fuzzern genutzt werden - darauf welche Umstände dies sind und wie es sich ob der Anwendbarkeit bei modernen Fuzzern verhält gehen wir in Kapitel CRIU+Fuzzing Engines ein. Wir betrachten hier zwei Varianten dieses Ansatzes: Zuerst eine recht naive Nutzung von Checkpoint/Restore im Fuzztarget eines Non-C/R-Fuzzers mit ähnlichem Ziel und einem ähnlichen Ansatz zu dem oben beschriebenen C/R-Fuzzer, danach eine Architektur die durch Anwendung von Checkpoint/Restore versucht das effiziente Fuzzern komplexer mehrstufiger Programme, beispielsweise Netzwerkprotokolle, durch Exploration des Zustandsgraphen als ganzes zu ermöglichen.

3.2.1 Naiv

Eine konzeptuell recht simple Variante der Kombination von klassischen, Non-C/R-Fuzzern mit C/R-Implementierung im Fuzztarget wäre naives Restoring nach dem inputunabhängigen Setup, ähnlich wie bei dem oben beschriebenen C/R-Fuzzer. Anders als beim obigen C/R-Fuzzer muss hier jedoch im Fuzztarget entschieden beziehungsweise bestimmt werden, ob wir uns in der ersten Iteration befinden, das Setup also ausgeführt werden und danach gecheckpointet werden muss, oder ob wir uns in einer späteren Iteration befinden, wir also das Setup überspringen und direkt, mittels Restoring, zur Ausführung des zu testenden Codes springen. Dies kann beispielsweise durch das Setzen von Systemvariablen oder das Erstellen von Tokens im Dateisystem geschehen - ist die Variable bei einer Ausführung also gesetzt oder existiert die Datei so befinden wir uns in einer späteren Ausführung und sollten das Setup durch einen Restore überspringen. Konkret sollte ein solches Fuzztarget dann hierfür in etwa so aussehen:

Listing 5: Struktur C/R-Fuzztarget für Non-C/R-Fuzzer

```
1 Falls der Token existiert :
2     Restore
3 Sonst :
4     Führe Setup aus
5     Setze Token
6     Checkpointe das Programm (ohne Dump)
7     Hole dir den Input
8     Führe den Test mit diesem Input aus
```

Wenn in Zeile 2 `restore` wird so wird die alte, gecheckpointete Instanz wieder hergestellt und setzt ihre Ausführung in Zeile 7 fort, holt sich also neuen Input und testet mit diesem.

Hierbei können jedoch einige Probleme auftreten die zu beachten sind:

- Wie beim C/R-Fuzzer muss natürlich das Setup inputunabhängig sein beziehungsweise in einen inputunabhängigen Teil, welcher vor dem Checkpoint ausgeführt wird, und einen inputabhängigen Teil, welcher mit dem Test selbst ausgeführt wird, aufgebrochen werden.

- Der Fuzzer muss mit Multiprocess-Programmen umgehen können und über das Vorhandensein des restorten Prozesses informiert sein, insbesondere muss er Exceptions, undefiniertes Verhalten und anderes unerwünschtes Verhalten im restorten Prozess erkennen können. Dies könnte beispielsweise durch eine restore-as-child Funktionalität wie beispielsweise von CRIU angeboten realisiert werden, benötigte jedoch trotzdem noch einen robusten Multiprocessfähigen Fuzzer
- Der restorte Prozess muss in irgendeiner Form an den Input kommen - dies könnte mit kontemporären Fuzzern unter Umständen ein Problem darstellen, da diese zumeist den Input entweder beim Aufruf des Fuzztargets als Argument oder durch einen FD wie `stdin` übergeben werden - beides Ressourcen auf die der wiederhergestellte Prozess im Allgemeinen keinen Zugriff hat. Dieses Problem könnte gelöst werden indem der restorte Prozess den Input direkt von Fuzzer erhält, beispielsweise indem der Fuzzer die Inputs über einen FIFO oder eine ähnliche Ressource im Filesystem, außerhalb des Prozesses selbst, übergibt oder indem der restorende Prozess vor dem Restore den Input in einer Art und Weise abspeichert, so dass der restorte Prozess diesen Input auslesen kann, beispielsweise wieder durch eine Zwischenspeicherung im Filesystem (siehe Listing WSAUCHIMMER).
- Die Messung der Codecoverage bei White- oder Gray-Box-Fuzzern dürfte wenig zuverlässig sein wenn im Fuzztarget selbst restoret wird - einerseits durch die Einführung von neuen Codeelementen durch das C/R-Tool, wobei dies natürlich vom C/R-Ansatz und dem genutzten C/R-Tool abhängt, saubere Kernelebene-C/R-Tools dürften hier beispielsweise wenig Einfluss haben, andererseits durch die massive und durch den eigentlich Programmcode nicht beschriebene Veränderung des Prozesses durch das Restore. Dies dürfte dazu führen, dass ein solcher Non-C/R-Fuzzer wenig besser als ein Black-Box-Fuzzer funktionieren dürfte.

Listing 6: Struktur C/R-Fuzztarget für Non-C/R-Fuzzer mit Inputdump

```

1 Falls der Token existiert:
2     Hole den Input
3     Schreibe den Input in eine Datei
4     Restore
5 Sonst:
6     Führe Setup aus
7     Setze Token
8     Checkpointe das Programm (ohne Dump)
9     Hole den Input aus der Datei
10    Führe den Test mit diesem Input aus

```

Ebenso wie der oben beschriebene C/R-Fuzzer versucht dieser Ansatz nur das Setup-Overhead-Problem im Fuzzing zu lösen - mit ähnlichen Laufzeitverbesserungen. Wie auch beim C/R-Fuzzer gilt hier

$$T_{Non-C/R} = n(T_{Setup} + T_{Execution}) \quad (4)$$

$T_{C/R}$ ist etwas komplizierter, da das Fuzztarget hier einige der C/R-Fuzzer-Funktionalitäten ad-hoc implementieren muss, es ergibt sich zunächst:

$$\begin{aligned}
T_{C/R}(n) &= 1(T_{Tokencheck} + T_{Setup} + T_{Tokenset} + T_{Checkpoint} + T_{Input} + T_{Execute}) \\
&\quad + (n-1)(T_{Tokencheck} + T_{Restore} + T_{Input} + T_{Execute}) \\
&\stackrel{(a)}{\simeq} T_{Setup} + T_{Checkpoint} + T_{Execute} + (n-1)(T_{Restore} + T_{Execute}) \\
&= T_{Setup} + T_{Checkpoint} + (n-1)T_{Restore} + nT_{Execute}
\end{aligned} \tag{5}$$

Die Vereinfachung in (a) kommt dabei aus der Annahme, dass die diversen Token- und Inputoperationen verglichen mit den recht langsamen Restore- und Setup-Operationen für die letztendliche Laufzeit nicht ins Gewicht fallen beziehungsweise im Falle der Inputoperationen auch im Non-C/R-Fall ausgeführt werden müssen, dies dort jedoch versteckt geschieht.

Das $T_{C/R}$ in diesem Fall unterscheidet sich nicht relevant vom $T_{C/R}$ in Gleichung (2), der Unterschied um $1T_{Restore}$ ergibt sich daraus, dass wir im C/R-Fuzzer nach dem ersten Setup nicht direkt weiter ausführen sondern den Prozess beenden und dann später für die erste Ausführung wieder Restoren. Natürlich ließe sich im C/R-Fuzzer-Fall auch diese Optimierung vornehmen.

Auch hier ergibt sich, analog zu Gleichung (3) die Laufzeitverbesserung je Iteration als

$$\begin{aligned}
\frac{T_{Non-C/R}(n) - T_{C/R}(n)}{n} &= \frac{(n-1)T_{Setup} - T_{Checkpoint} - (n-1)T_{Restore}}{n} \\
&\stackrel{n \rightarrow \infty}{=} T_{Setup} - T_{Restore}
\end{aligned} \tag{6}$$

3.2.2 Exploration des Zustandsgraphen in Client/Server Systemen

Die bisher beschriebenen Ansätze bieten nur sehr naive Ansätze zur Lösung nur eines der beschriebenen Fuzzingprobleme - Laufzeitoverheadreduktion in einstufigen, simplen Anwendungen. In diesem Kapitel beschreiben wir einen Ansatz zum Testen einer Seite komplexerer, insbesondere mehrstufiger, interaktiver Client-Server-Anwendungen. Hierbei wird durch die Fuzzing-Inputs Schritt für Schritt ein Zustandsgraph beispielsweise des Servers aufgebaut, parallel dazu werden bereits diese Zustände jeweils mit den Inputs auf Exceptions getestet. Zur besseren Lesbarkeit werden wir in diesem Kapitel davon ausgehen, dass ein Server gefuzzt werden soll und das Fuzztarget die Client-Aufgaben übernimmt, dies lässt sich natürlich auch umkehren.

Das Fuzztarget hat dabei in diesem Ansatz mehrere Aufgaben:

- **Client-Server-Interaktion:** Das Fuzztarget implementiert nicht einen wirklichen Client der das erwünschte Protokoll spricht sondern simuliert nur einen solchen indem es die von der Fuzzing-Engine generierten Inputs als Client-Input an den Server übergibt, beispielsweise über eine etablierte TCP-Verbindung zwischen dem Server und Client.
- **Verwaltung des Servers:** Das Fuzztarget hält in persistenter Form, beispielsweise durch Schreiben ins Dateisystem, den bisher gefundenen Zustandsgraphen des Servers, also die gefundenen Zustände und die Inputs die zu Übergängen führen, vor. Zusätzlich wird jeder Zustand wenn er das

erste mal erreicht wird gecheckpointet und das entsprechende Image mit dem Zustand verknüpft, das heißt das Fuzztarget muss auch bestimmen wann neue Zustände erreicht werden.

Hierbei werden dann neu generierte Inputs aber nicht nur auf einen bestimmten Zustand angewendet sondern entweder hintereinander auf alle bisher bekannten Zustände oder auf zufällig gewählte Zustände, so dass der Zustandsgraph weiter ausgebaut wird.

Ein solches Fuzztarget könnte in Pseudocode wie folgt aussehen:

Listing 7: Fuzztarget zur Exploration des Zustandsgraphen

```

1  if token not set:
2      server.start()
3      connection <- connectTo(server)
4      c <- checkpoint(connection)
5      state_id <- server.state
6      s <- cr.checkpoint(server, dump=true)
7      transition = [] // [(input, new_state)]
8      states <- {state_id = {s = s, c = c, t = transition, id =
          state_id}}
9  for state in states:
10     server <- cr.restore(state[s])
11     connection <- restore(state[c])
12
13     input <- fuzzer.input
14     connection.send(input)
15     if state.id ≠ server.state and server.state not in states.keys:
16         new_c <- checkpoint(connection)
17         new_state_id <- server.state
18         new_s <- cr.checkpoint(server, dump=true)
19         new_t = []
20         states[state][t].append((input, new_state_id))
21         states[new_state_id] = {s = new_s, c = new_c, t = new_t, id
            = new_state_id}
22     else if state.id ≠ server.state:
23         states[state][t].append(input, server.state)

```

Hierbei kann es sich bei **connection** um eine beliebige Verbindungsart zwischen dem Client und dem Server handeln, beispielsweise eine TCP-Verbindung oder UNIX-Sockets, solange das Wiederherstellen dieser Verbindung vom C/R-Tool unterstützt wird. Das clientseitige Checkpointing (beispielsweise Zeile 4) wird dabei nicht durch das normale C/R-Tool durchgeführt, da nicht der ganze Client gecheckpointet wird, sondern durch andere, der Art von **connection** angemessene Ansätze. Für eine TCP-Verbindung kann dies beispielsweise durch einen Userspace TCP/IP-Stack geschehen oder durch das Nutzen der **TCP_REPAIR** Option in modernen Linux-Kerneln (seit 3.5).

Zusätzlich muss sich der Server darüber bewusst sein, in welchem Zustand er sich gerade befindet und dies dem Fuzztarget/Client mitteilen können (siehe: **state_id <- server.state** etc.), beispielsweise durch Einfügen entsprechender Funktionalität in den Servercode und die Übergabe durch einen entsprechenden IPC-Ansatz oder direkt durch das Auslesen des States aus den C/R-Dumps.

Für einen Server mit bereits bekannten Zuständen s_0, s_1, s_2, s_3 mit den Zustandsübergängen wie gezeigt und entsprechenden Inputs für die Übergänge i_1, \dots, i_4 würde nun also für einen neuen Input i^* wie in Abbildung 1 gezeigt

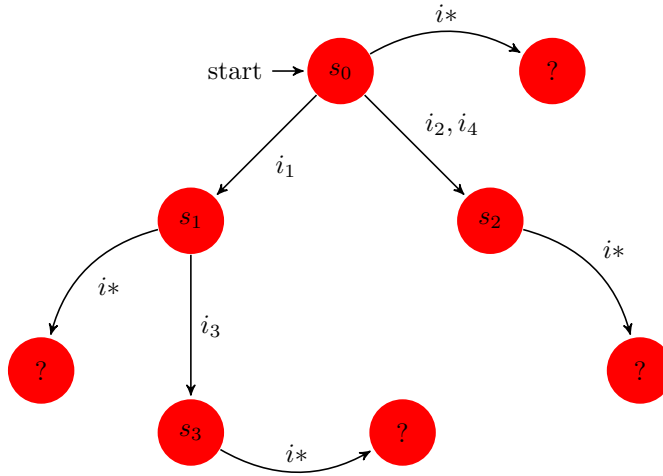


Abbildung 1: Zustandsübergangsexploration mit Input i^*

jeder der Zustände s_0, \dots, s_3 einmal wiederhergestellt, mit i^* getestet und falls Zustand s_i mit Input i^* einen neuen Zustand erzeugt dieser gecheckpointed und als Knoten s_4 mit der Kante (s_i, s_4, i^*) dem Graphen hinzugefügt.

Durch diese schrittweise Erforschung des Zustandsgraphen kann der Fuzzer nun an den verschiedenen Zuständen ansetzen und sollte so durch Formulierung nur eines Fuzztargets in der Lage sein die ganze Funktionalität dieses Servers zu testen - anstatt wie im klassischen Fuzzing einzelne Fuzztargets für jeden Pfad in diesem Zustandsgraphen zu schreiben oder sogar nur einzelne Funktionen in Isolation zu testen.

Hierbei gilt natürlich zu beachten, dass auch dieser Ansatz einige Probleme mit sich bringen dürfte:

- White-Box-Fuzzing beziehungsweise klassisches Gray-Box-Fuzzing durch Code-Coverage werden, wie schon bei den naiven Ansätzen, durch die permanente Variation des tatsächlich ausgeführten Codes erschwert beziehungsweise könnten in bisherigen Fuzzern tatsächlich gar nicht nutzbar sein, da der eigentlich ausgeführte Code, der Client, gar nicht der Code ist den wir Fuzzern wollen - dies wäre der Server. Stattdessen stellt diese Art des Fuzzings jedoch eine eigene Art des Gray-Box-Fuzzings dar, da das Fuzztargets nach und nach die Struktur des Servers, wenn auch auf einem höheren Level als klassische Code-Coverage, herausarbeitet und darüber versucht das Fuzzing zu optimieren.
- Selbst wenn Code-Coverage-Messungen möglich sind so ist ihre Effektivität doch dadurch abgeschwächt, dass diese Inputs auf alle Zustände angewendet werden, nicht nur auf diejenigen Zustände bei denen sie ob der Code-Coverage ausgewählt wurden.
- Ebenso wie bei der naiven C/R-Implementierung im Fuzztarget muss der Fuzzer auch hier mit Multiprocess-Anwendung umgehen können, insbesondere den Server - als vom Fuzztarget gestartetes, eigenständiges Programm - in die Suche von Exceptions einbeziehen.

- Das Formulieren von Regeln, Dictionaries, ... für den Inputs eines Smart-Fuzzers wird verkompliziert, da diese Regeln nun nicht nur für einen Pfad oder eine bestimmte Funktion gelten sollten sondern für das ganze zu testende Programm. Dies kann unter Umständen zu einer Vergrößerung des Suchraums oder sogar der nicht-Nutzbarkeit von Smart-Fuzzern führen was wiederum zu Performanceeinbußen führen kann
- In Fällen in denen es nur wenige gültige Übergänge auf einen Zustand mit hoher Distanz vom Ausgangsknoten gibt könnte es passieren, dass die Inputs bereits untersucht wurden oder als wenig vielversprechend von der Fuzzing-Engine verworfen wurden. Um dies zu vermeiden sollte die Optimierung der Inputs, wenn überhaupt möglich, nicht zu zielgerichtet erfolgen.

4 CRIU

Für den Rest der Arbeit werden wir CRIU - Checkpoint/Restore in Userspace - näher betrachten und CRIUs Nutzbarkeit für Fuzzing, insbesondere in Kombination mit kontemporären Fuzzing-Engines, untersuchen. Die Wahl von CRIU als C/R-Tool für diese Untersuchung ist dabei auf mehrere Gründe zurückzuführen:

- Verfügbarkeit & Aktualität: CRIU ist ein vergleichsweise modernes, gut dokumentiertes Open Source C/R-Tool das geringe Anforderungen - es benötigt keine Kernelpatches oder -module außerhalb des Linux Mainline Kernels - an das System stellt und ist daher auf einer Vielzahl von Systemen verfügbar. Zusätzlich wird CRIU von den Entwicklern noch unterstützt und weiterentwickelt, im Gegensatz zu einigen anderen vielversprechenden C/R-Tools wie Linux-C/R.
- Unveränderte Binary: Anders als einige andere Userspace-C/R-Tools nimmt CRIU keine Veränderungen an der Binary vor oder benötigt bestimmte Schritte bei der Kompilierung - CRIU benötigt nicht einmal besondere Schritte beim Anwendungsstart, erst zum Zeitpunkt des Checkpointens wird CRIU aktiv.
- Capabilities: CRIU bietet eine breite Palette an Capabilities die mit CRIU wiederhergestellt werden können, insbesondere TCP- und andere Netzwerkverbindungen, mehr dazu in Kapitel CRIU.Capabilities.

CRIU findet unter anderem Anwendung in einigen prominenten Container-Anwendungen wie Docker, LXC/LXD und OpenVZ, anstatt des Legacy OpenVZ C/R-Systems, zur Implementierung des Container-Checkpointings.

4.1 Nutzung

Es gibt zwei Kategorien wie CRIU genutzt werden kann:
TODO RPC details

- Die Nutzung über das Command Line Interface ohne Remote Procedure Call beziehungsweise ohne CRIU Service

- Die Nutzung des RPCs entweder durch das CLI oder eine Programmiersprachen API, insbesondere die C-API

Unabhängig davon welcher der beiden Optionen man nutzt folgt das Checkpointing und Restoring in etwa den gleichen Schritten:

- Falls RPC genutzt werden soll muss der Service laufen (beispielsweise zu starten durch `criu service`), falls die C-API genutzt wird muss diese mittels `criu_init_opts()` initialisiert werden
- Man bestimmt welchen laufenden Prozess man checkpointen möchte, meist über die PID durch Nutzung der `--tree $PID` Option beziehungsweise `criu_set_pid(pid)`, falls sich ein Prozess über die C-API selbst checkpointen möchte muss `criu_set_pid` nicht explizit aufgerufen werden. Zusätzlich muss angegeben werden, ob der Prozess nach dem Checkpointing fortgesetzt werden oder beendet werden soll.
- Man gibt einen Pfad oder einen FD für das Image-Directory an, entweder über die `--images-dir $IMGDIR` Option oder `criu_set_images_dir_fd(fd)`, bei der Nutzung der C-API muss `fd` ein bereits geöffneter FD auf das Verzeichnis sein in das das Image geschrieben werden soll
- Spezifizierung der nötigen Zusatzinformationen, insbesondere `--shell-job` beziehungsweise `criu_set_shell_job(true)`, `--tcp-established` beziehungsweise `criu_set_tcp_established(true)` und andere externe Ressourcen falls nötig
- Modifikation der Socket-Location des Services, des Log-Levels, Log-Files, ... falls nötig oder erwünscht

Diese Optionen werden dann mit der entsprechenden Operation, beispielsweise Checkpointing (`criu dump`, `criu_dump()`) oder Restoring (`criu restore`, `criu_restore()`), kombiniert, diese wird dann von CRIU ausgeführt. Wird ein Shell-Job restored so wird dieser direkt an das laufende Terminal attached.

4.2 Funktionsweise

Wie bereits in der Übersicht verschiedener C/R-Systeme angeschnitten ist CRIU ein Userspace-C/R-Tool welches jedoch dabei Kernel-Features wie beispielsweise `prctl` aus dem `CONFIG_CHECKPOINT_RESTORE` Kernel-Feature oder die `TCP_REPAIR`-Flag verwendet um Anwendungen möglichst vollständig wiederherstellen zu können ohne dabei eine komplette Kopie aller Kernelinformationen im Userspace aufbauen und permanent erhalten zu müssen wie dies beispielsweise bei DMTCP der Fall ist. Dies bedeutet insbesondere auch, dass die zu checkpointene Anwendung weder zur Kompilierzeit noch zum Zeitpunkt des Anwendungsstartes sich der zukünftigen Checkpointoperation bewusst sein muss und keine besonderen Vorbereitungen für das Checkpointen einer Anwendung getroffen werden müssen. Die von CRIU genutzten Kernel-Features sind im Mainline-Linux-Kernel enthalten und es sind sogar alle nötigen Features aktiviert, insbesondere `CONFIG_CHECKPOINT_RESTORE`, ein Feature welches eine Reihe von für C/R nützlichen Tools und Informationen enabled, beispielsweise `prctl` zur Manipulation von Prozessen oder eine Erweiterung der in `/proc/`

verfügbaren Informationen beispielsweise um die `/proc/<pid>/map_files` erweitert, `CONFIG_NAMESPACES` sowie Features für diverse spezielle Namespaces, namentlich UTS, IPC, PID und Network Namespaces, sowie diverse Features zum Monitoring verschiedener Socket-Arten.

Das Checkpointing folgt im Großen und Ganzen grob dem oben beschriebenen Muster:

1. Einfrieren des Prozessbaumes. Dies geschieht auf eine von zwei Arten: Entweder, falls möglich, durch die Nutzung einer freezer cgroup, oder ansonsten durch das Traversieren des PID-Baumes in `/proc/` und anhalten der Prozesse durch die Nutzung der `PTRACE_SEIZE` Anweisung.
2. Erfassen von Informationen wie den Virtuellen Speicherbereichen, Registern, ... durch das Lesen entsprechender `/proc/` Einträge und `ptrace`.
3. Injektion eines Parasiten-Code-Blobs in die Anwendung mittels `ptrace`. Dieser Parasit kann nun die fehlenden Prozessinformationen extrahieren, hierfür wird die Anwendung kurzfristig wieder aufgetaut.
4. Entfernung des Parasiten mittels `ptrace`, Schreiben des Images und Fortsetzung der Ausführung der Anwendung.

Der Parasit ist hierbei eine PIE welche mittels `ptrace` in die Anwendung injiziert wird. Dies geschieht in zwei Schritten: Zuerst wird ein Stück Code mittels `ptrace` so in die Anwendung injiziert, dass dieses nach dem Auftauen der Anwendung ausgeführt wird und somit Shared Memory für den eigentlichen Parasiten zur Verfügung stellen kann. Dann wird die Anwendung kurzfristig aufgetaut, das Shared Memory zur Verfügung gestellt und die Ausführung des Parasiten angestoßen. Der Parasit wartet auf Anweisung des CRIU-Prozesses, zumeist ob des Erfassens von Informationen wie Memory-Inhalten und Credentials, und führt diese aus. Zuletzt sendet der CRIU-Prozess noch eine `PARASITE_CMD_FINI` Anweisung an den Parasiten, dieser terminiert sich selbst und wird anschließend vom CRIU-Prozess durch das unmappen des Shared Memories entfernt.

Das Restoring weicht etwas weiter vom oben beschriebenen Muster ab:

1. Lesen des Images und Zuweisung von Shared Ressourcen zu Prozessen so das diese nur einmal wiederhergestellt werden.
2. Forken der CRIU-Anwendung um einen Prozessbaum entsprechend dem der Anwendung zu erzeugen.
3. Wiederherstellen eines Großteils der Prozessinformationen, insbesondere Speicherinhalte (jedoch nicht die exakten Mappings), Sockets, Namespaces, ...
4. Ersetzen der Forks durch einen weiteren Code-Blob, den Restorer Context, Wiederherstellung aller verbleibender Informationen, insbesondere exakte Memory Mappings, Credentials, Operationen die von Credentials abhängen wie `fork-with-pid` und Threads durch den Restorer-Context, un-mapping des Restorer-Context und Fortsetzen der planmäßigen Ausführung der Anwendung.

Der Restorer-Blob ist, ebenso wie der Parasit, ein PIE-kompiliertes Stück Code welches kurzfristig in Shared Memory im Anwendungsprozess lebt um dort Aufgaben der CRIU-Anwendung zu übernehmen. Dabei agiert der Restorer-Context als Verbindungsstück zwischen dem CRIU-Fork und dem eigentlich Anwendungsprozess und reorganisiert beispielsweise die bereits in Schritt 3 gefüllten VMAs so dass die Mappings mit der Anwendung übereinstimmen oder stellt die korrekten PIDs wieder her.

Die entstehenden CRIU Images bestehen aus mehreren Dateien organisiert in dem beim Checkpointing spezifizierten Ordner, diese fallen in drei Kategorien: Dateien im Protobuf-Format welche einen Großteil der eigentlichen Checkpoint-Informationen enthalten wie zum Beispiel generelle Checkpointing Informationen in `inventory.img` oder Informationen über die Credentials wie uid in `creds.img`, CRIU-spezifische Binary-Images welche beispielsweise die Memory-Inhalte und Pagemaps enthalten und Dateien die durch Drittparteianwendungen generiert wurden uns beispielsweise Routing- und Netzwerkinformationen enthalten.

4.3 Capabilities

TODO

What can we restore with CRIU? Pretty much any program, independent of language etc, does not require linking with some interception library, marking at start, ... CRIU also can restore pretty much any feature we could find, especially established TCP connection.

4.4 Probleme

Da CRIU nicht für eine Anwendung in Fuzzing- oder ähnlichen Hoch-Frequenz-Restoring-Kontexten entwickelt wurde entstehen bei der Anwendung zunächst ein paar Probleme:

1. Das mehrfache Wiederherstellen eines Anwendungspaares mit etablierter TCP-Verbindung scheitert da die Verbindung bei späteren Restores resettet wird. Dies wird normalerweise, bei Checkpointing/Dumping und einfachem Restoring, durch entsprechende Firewallregeln, gesetzt direkt nach dem Dumping, verhindert, diese Regeln werden während des Restoring wieder entfernt. Bei weiteren Restores sind diese Regeln also nichtmehr vorhanden.
2. CRIU stellt Anwendungen unter der gleichen PID wie zum Zeitpunkt des Checkpointings wieder her - wurde diese PID aber nach dem Checkpointing und vor dem Restoring wieder vergeben so scheitert das Restoring.
3. Bei existierenden Netzwerkverbindungen ist zu beachten, dass diese zumeist eine beschränkte Lebenszeit haben, eine Zeit nach der die Verbindung geschlossen wird wenn keine Kommunikation statt findet. Diese Zeit wird beispielsweise für TCP-Verbindung unter Linux nicht in tatsächlicher Prozesslaufzeit gemessen sondern als Differenz zwischen gespeicherter Zeit für die letzte Kommunikation und momentaner Systemzeit berechnet - wenn nun also ein Prozess nach diesem Zeitfenster, der Linux-Default sind

7200 Sekunden, wiederhergestellt wird so wird die Verbindung direkt ob des Überschreitens der TCP Keepalive Time geschlossen.

Glücklicherweise gibt es für diese Probleme Lösungen die eine Nutzung in einem Fuzzing-Umfeld ermöglichen sollten.

Das TCP-RST-Problem lässt sich durch manuelles Setzen eben dieser Regeln direkt nach der Beendigung der Ausführung der Anwendungen lösen, beispielsweise wie in Listing [TODO] zu sehen.

Listing 8: Setzen der Firewallregeln

```
1 // [...]
2
3 // create the netfilter rules & iptables calls
4 char iptables1[200], iptables2[200], iptables3[200], iptables4
  [200];
5 sprintf(iptables1, "iptables -A OUTPUT -s 127.0.0.1/32 -d
  127.0.0.1/32 -p tcp -m mark ! --mark 0xc114 -m tcp --sport
  %i --dport %i -j DROP", clientport, serverport);
6 sprintf(iptables2, "iptables -A OUTPUT -s 127.0.0.1/32 -d
  127.0.0.1/32 -p tcp -m mark ! --mark 0xc114 -m tcp --sport
  %i --dport %i -j DROP", serverport, clientport);
7 sprintf(iptables3, "iptables -A INPUT -s 127.0.0.1/32 -d
  127.0.0.1/32 -p tcp -m mark ! --mark 0xc114 -m tcp --sport
  %i --dport %i -j DROP", clientport, serverport);
8 sprintf(iptables4, "iptables -A INPUT -s 127.0.0.1/32 -d
  127.0.0.1/32 -p tcp -m mark ! --mark 0xc114 -m tcp --sport
  %i --dport %i -j DROP", serverport, clientport);
9
10 // [...]
11
12 // wait for restored processes to terminate
13 waitpid(serverpid, NULL, 0);
14 waitpid(clientpid, NULL, 0);
15
16 // apply the netfilter rules
17 system(iptables1);
18 system(iptables2);
19 system(iptables3);
20 system(iptables4);
```

Die Problematik belegter PIDs lässt sich durch die Isolation der Anwendung durch beispielsweise eine virtuelle Maschine beziehungsweise einen Container oder die Nutzung von PID Namespaces lösen. Unter Umständen ließe sich das PID-Problem auch durch die Deaktivierung der `--shell-job` Option lösen, dies würde jedoch das Ausführen der Anwendungen in einer eigenen Session erfordern - etwas das Fuzzing-Engines im Allgemeinen nicht unterstützen.

Das TCP Keepalive Timeout und ähnliche Probleme ließen sich auf mehrere Arten und Weise lösen sollten sie ein Problem sein - entweder könnte die Zeit bis zum Timeout verlängert werden durch beispielsweise `echo $new_time_in_s > /proc/sys/net/ipv4/tcp_keepalive_time` auf einen genügend großen Wert oder durch Manipulation des Zeitpunktes der letzten Nachricht im Image.

4.5 Performance

Bei der Betrachtung der Performance von C/R-Systemen gibt es zumeist zwei Komponenten zu betrachten:

Iterationen	100 ⁰ byte	100 ¹ byte	100 ² byte	100 ³ byte	100 ⁴ byte	TCP
10	0.292s	0.290s	0.292s	0.301s	0.627s	0.982s
50	1.269s	1.269s	1.256s	1.298s	2.332s	4.615s
100	2.459s	2.494s	2.477s	2.558s	4.482s	9.170s
500	12.205s	12.228s	12.250s	12.648s	21.564s	45.541s
$T_{Restore}$	0.024s	0.024s	0.024s	0.025s	0.043s	0.091s

Abbildung 2: Performancemessung

- Die Performanceeinflüsse zur Laufzeit des Programms, ausgelöst beispielsweise durch das Sammeln von Informationen die für das Checkpointing relevant sind
- Die benötigte Zeit für das Durchführen der tatsächlichen Checkpointing- und Restore-Operationen

Ersteres, Performanceeinflüsse des laufenden Programms, sind bei C/R-Systemen wie CRIU jedoch nicht-existent, da vor dem Checkpointing weder dem Kernel noch der Anwendung selbst bewusst ist, dass sie gecheckpointed werden wird, und nach dem Restoring zwar Reste des Restorer Context Blobs noch in Memory vorhanden sein können, diese sich jedoch nicht mit dem Anwendungscode überschneiden und entsprechenden im Normalfall keinen Einfluss auf die weitere Performance haben werden.

Die Geschwindigkeit mit der Checkpointing und Restore Operationen durchgeführt werden können spielt für klassische Zwecke wie Debugging oder Fault-Recovery zumeist eine untergeordnete Rolle, da wir hier zumeist die Anwendung nur eine handvoll Male checkpointen und restoren, die Grenzen der Akzeptabilität der Wartezeit hier zumeist durch menschliche Interaktion oder die normal erwartete Startup-Zeit bestimmt werden. Bei Nutzung von C/R-Systemen für Fuzzing sieht dies natürlich anders aus - die Geschwindigkeit mit der wir Anwendungen wiederherstellen können $T_{Restore}$ ist wie wir in Kapitel 3 von hoher Relevanz für das C/R-fuzzzen von Anwendungen, da wir für jeden Testfall mindestens einmal restoren müssen.

Die benötigten Zeiten für C/R-Operationen hängen dabei von einer Reihe von verschiedenen Umständen ab, unter anderem:

- Allokierter/Genutzter Speicher der Anwendung
- Größe des Prozessbaumes
- Auswahl zu restorender Features wie TCP-Sockets
- TODO: Mehr?

Da natürlich auch externe Einflüsse wie Hardware, Auslastung, ... einen Einfluss auf die Performance haben und diese Zeiten zumeist, wie oben beschrieben, nicht von hoher Relevanz in der Nutzung sind ist der Vergleich der Performance aus der Literatur heraus zumeist nicht möglich.

Abbildung [TODO: Performancemessung] zeigt die Ergebnisse von Performancemessung in Hinsicht auf Laufzeit eines Launchers der Anwendungen startet, checkpointed und dann diese wiederholt wiederherstellt sobald die Anwendung terminiert. Diese Messung fand auf einem Lenovo Thinkpad X230 statt, die Laufzeiten sind das arithmetische Mittel jeweils mehrerer Messungen, für die gesamten Ergebnisse siehe Anhang [TODO]. Die Anwendungen sind dabei

1. Eine einzelne Anwendung die eine bestimmte Anzahl bytes zufällige Daten ließt, sich selbst checkpointed und dann eine sehr simple Operation auf den Daten ausführt und das Ergebnis auf `stdout` ausgibt (TODO: Anhang)
2. Ein Client-Server-Paar das eine TCP-Verbindung aufbaut, sich selbst checkpointen und dann sich gegenseitig jeweils ein kurzes Datenpaket zusenden, dabei Informationen über den Zustand auf `stdout` ausgeben (TODO: Code im Anhang)

Dabei setzen sich die Zeitmessungen über n Iterationen wie in Gleichung [TODO] zusammen. Durch die Kombination von Messungen mit verschiedenen Iterationen, siehe Gleichung [TODO], lässt sich dies dann als lineares Gleichungssystem formulieren und sowohl T_1 als auch T_2 abschätzen. Dabei fällt auf, dass $T_{Execute}$ mit individuell gemessenen Zeiten von relevant unter $1ms$ in T_2 vernachlässigt werden kann, also der so erhaltene Wert von T_2 bereits eine sehr gute Approximation für $T_{Restore}$ ist.

$$T(n) = T_1 + nT := (T_{Setup} + T_{Checkpoint}) + n(T_{Restore} + T_{Execute}) \quad (7)$$

$$\begin{aligned} T(n_1) &= T_1 + n_1 T_2 \\ T(n_2) &= T_1 + n_2 T_2 \end{aligned} \quad (8)$$

Damit ergeben sich nun also nun mittlere Wiederherstellungszeiten, je Iteration, von ca. $24ms$ für die einfach Wiederherstellung sowie zusätzliche Zeit je nach Menge von genutztem Speicher und Zeiten von ca. $90ms$ für die Wiederherstellung von einem Client-Server-Paar mit bestehender TCP-Verbindung - dies sollte auch schon bereits die Dauer der Wiederherstellung einer einzelnen solchen TCP-Anwendung sein, da der Restore in mehreren Prozessen zugleich geschieht.

Damit liegen zumindest die Checkpoint- und Restore-Zeiten für die Nicht-TCP-Anwendungen, für TCP-Verbindungen waren wir nicht in der Lage vergleichbare Zahlen zu finden, relevant über denen von Linux-CR - dort wurden auf älterer Hardware Restore-Zeiten von $\leq 1ms$ erreicht. Wir vermuten, dass dies auf die verschiedenen Ansätze zum Restoring zurückzuführen ist, insbesondere benötigt Linux-CR nicht die Injektion der Blobs durch `ptrace` und den damit verbundenen Overhead da Linux-CR direkt mit Kernel-Ressourcen arbeiten kann. Insbesondere heißt dies aber, dass, obwohl kontemporäre starke C/R-Tools recht lange Checkpoint- und Restore-Zeiten haben, diese nicht optimal sind und es noch Optimierungspotential in dieser Richtung gibt, wir können bei einer generellen Bewertung der Möglichkeiten von C/R-Fuzzing also noch mit einer Verbesserung dieser Werte rechnen.

5 TODO: CRIU + Fuzzing Engines

Schlussendlich wollen wir in diesem Kapitel nun einen Blick auf kontemporäre Fuzzing-Engines werfen, insbesondere ob ihrer Nutzbarkeit für C/R-Fuzzing mit gegebenen C/R-Tools, speziell CRIU. Die Fuzzing Engines die wir dabei betrachten werden sind afl-fuzz (american fuzzy lop) und libfuzzer.

5.1 Fuzzing-Engines

Die hier betrachteten Fuzzing Engines libfuzzer und afl wurden von uns ausgewählt, da sie zwei der etablierteren Fuzzing Engines sind. Beide haben zur Auffindung einer ganzen Reihe von Bugs beigetragen und sind Teil von Googles OSS-Fuzz-Projekts - ein Projekt welches durch kontinuierliches Fuzzing durch Google versucht die Sicherheit und Stabilität von relevanter Open Source Software zu verbessern. Sowohl libfuzzer als auch afl folgen im allgemeinen der obigen Beschreibung von Fuzzern, unterscheiden sich jedoch in der exakten Umsetzungen in einigen Punkten.

5.1.1 Funktionsweise: libfuzzer

Die Nutzung libfuzzers folgt einem recht strikten aber simplen Schema dank der Integration in das llvm-Projekt und insbesondere in dessen Frontend, clang:

1. Schreibe ein Fuzz-Target, ein Stück Code welches eine zu fuzzende Funktion, in Normalfall `LLVMFuzzerTestOneInput`, zur Verfügung stellt welche als Argumente ein Array von `uint8_t` und die Länge dieses Arrays in `size_t` erwartet.
2. TODO

TODO

5.1.2 afl

TODO

5.2 Fuzzing-Engines und C/R-Fuzzing

5.2.1 libfuzzer

Um die Nutzbarkeit libfuzzers für C/R-Fuzzing zu testen implementierten wir ein simples Fuzz-Target wie in Kapitel 3.2.1, Listing 6, siehe Anhang [TODO]. Der Umweg des Inputdumpings ist zwingend notwendig, da libfuzzer den Input immer als Argument bei Aufruf übergibt. Bei der Ausführung ergaben sich jedoch bei uns einige Probleme:

- Ist AddressSanitizer aktiviert so scheitert CRIU bereits in der ersten Ausführung während des Checkpointings: Eine Page wird als $\tilde{300GB}$ groß erkannt, die Allokierung entsprechendes Speichers scheitert und damit auch das Page-mapping und das Pagedumping mittels des Parasitencodes (siehe Listing TODO). Dies liegt daran, dass AddressSanitizer 20TB virtuellen Speicher belegt. Restore-Versuche scheitern daran, dass das Image entsprechend nicht komplett ist, insbesondere `inventory.img`, die Top-Level-Beschreibung des Images, fehlt, siehe Listing TODO.

- Die Nutzung von CRIU in libfuzzer führt dazu, dass Wiederherstellung mit korrekten PIDs scheitert - und die Einbettung in einen Fuzzer macht auch die Nutzung von einem entsprechenden PID-Namespace ein Ding der Unmöglichkeit.

Listing 9: Pagemapping/dumping Errors mit ASan

```

1 (00.010591) Error (criu/pagemap-cache.c:54): pagemap-cache:
   pmc_init: Can't allocate 30064246792 bytes
2 (00.010597) Error (criu/pagemap-cache.c:85): pagemap-cache: Failed
   to init pagemap for 28814
3 (00.010600) Error (criu/mem.c:516): Can't dump page with parasite

```

Listing 10: Restoring Error mit ASan

```

1 (00.000108) No inventory.img image
2 (00.000133) Error (criu/util.c:693): Can't read link of fd -404: No
   such file or directory
3 (00.000142) Error (criu/protobuf.c:75): Unexpected EOF on (null)

```

Listing 11: Restoring Error durch PID mismatch

```

1 (00.007567) Error (criu/cr-restore.c:1668): Pid 9177 do not match
   expected 8154
2 (00.007683) Error (criu/cr-restore.c:2309): Restoring FAILED.

```

Selbst wenn diese Probleme gelöst werden sollten wäre eine Nutzbarkeit libfuzzers für C/R-Fuzzing, sowohl mit CRIU als auch anderen kontemporären C/R-Systemen, vorerst unklar, da libfuzzer zwar mit Multi-Threaded, aber nicht notwendigerweise korrekt mit Multi-Prozess-Anwendungen umgehen kann, eine allgemeine Nutzbarkeit, unabhängig von den zu testenden Programmen, unwahrscheinlich.

5.2.2 afl

Ebenso wie bei der Untersuchung libfuzzers implementierten wir bei afl ein simples C/R-Fuzz-Target wie in Kapitel 3.2.1, Listing 6, siehe Anhang [TODO]. Das Inputdumping ist an dieser Stelle nicht notwendig, da afl die Übergabe des Inputs über einen FD erlaubt.

Allerdings war es uns nicht möglich in afl auch nur den Checkpointing-Vorgang zu starten [TODO: Genau Analyse was da bei AFL schiefgeht, da arbeite ich grad noch dran... ich mag afl echt nicht].

5.3 Fazit

Zusammenfassend lässt sich nun also feststellen, dass C/R-Fuzzing in der in dieser Arbeit beschriebenen Form noch vor einigen Problemen steht die einer Nutzung im Wege stehen. Hierbei begegnen wir zwei Kategorien von Problemen: Erstens Probleme die aus der Wahl des Fuzzers und C/R-Systems resultieren beziehungsweise aus der Implementierung eben dieser resultieren, beispielsweise die Inkompatibilität von CRIU mit ASan ob des von ASan belegten Speichers oder die Belegung von benötigten PIDs zur Wiederherstellung durch libfuzzer. Zweitens Probleme die aus der Architektur existierender Fuzzer und

C/R-Systeme entstehen, insbesondere die Unmöglichkeit mit den betrachteten Fuzzern an wiederhergestellte Fuzz-Targets anzuknüpfen und diese, wie für das Fuzzing nötig, zu überwachen. Erstere Klasse von Probleme ließen sich unter Umständen durch Workarounds, Anpassungen der Implementierung oder Wahl eines anderen Systems umgehen, Zweitere würde, da diese Architekturentscheidungen von den meisten wenn nicht allen modernen und nutzbaren entsprechenden Systemen so getroffen wurden, die Nutzung oder sogar die Entwicklung eines Systems erfordern die eben nicht dieser Architektur folgt, beispielsweise ein C/R-System welches innerhalb eines Prozesses auf einzelnen Threads arbeitet oder ein Fuzzer der in der Lage ist an neue Kinder des Fuzz-Targets anzuknüpfen. Thread-C/R ist dabei nach unserer Einschätzung weniger vielversprechend, da, anders als Prozesse, Threads eben keine eigenständige Einheiten, sich Ressourcen mit anderen Threads teilen - dies würde ein C/R-System vor einige Probleme stellen wenn nur einige der Threads gecheckpointed werden und später wiederhergestellt werden, insbesondere die Frage welchen der Zustände Variablen etc annehmen sollen, den zum Zeitpunkt des Checkpointings oder den zum Zeitpunkt vor dem Restoring, und wie damit umzugehen ist wenn vom restoreten Thread genutzter Speicher wieder vergeben wurde, dies würde zuerst eine Anpassung aller entsprechender Speicheradressen im Image erfordern - hierbei sollte bedacht werden, dass eine Differenzierung zwischen Speicheradresse und einfachem gespeicherten Wert in einer Momentaufnahme des Zustandes eines Threads nicht im Allgemeinen möglich ist.

6 Ausblick

Auch wenn das Fazit zunächst etwas vernichtend klingen mag sollte die Idee des C/R-Fuzzing nicht ohne weitere, tiefergehende Betrachtung verworfen werden, da wir uns in dieser Arbeit, bedingt durch die Architekturen und Designentscheidungen populärer Fuzzer und C/R-Systeme, nur mit einer von vielen möglichen Facetten des C/R-Fuzzings in größerer Tiefe auseinandergesetzt haben: Mit der Kombination von Single-Process-Fuzzern mit Process-Level-C/R-Systemen und einer Implementierung des C/R-Fuzzings im Fuzz-Target. Zwei Ansätze wären bei weiterer Betrachtung anderer dieser Facetten aus unserer Sicht besonders vielversprechend:

- Die Nutzung oder Implementierung eines Multi-Process-Fuzzers, also Fuzzer die die Arbeit auf Anwendungen mit mehreren Prozessen erlauben
- Die Implementierung eines dedizierten C/R-Fuzzers

Beide Ansätze würden die grundlegende Problematik der Inkompatibilität der von uns betrachteten Systeme beheben.

Dies sollte jedoch, mit Blick auf die Performance im C/R-Fuzzing mit effizienten C/R-Systemen mit einer breiten Palette an Capabilities kombiniert werden - das ein solches C/R-System möglich ist hat Linux-CR gezeigt, für eine Weiterverfolgung des C/R-Fuzzings eine Wiederbelebung von Linux-CR oder eine Implementierung eines ähnlichen C/R-Systems also von hohem Interesse.

7 Anhang

TODO

Literatur

- [1] Heather Kelly. *Facebook hack exposed 50 million users' info – and accounts on other sites*. CNN Business, 2018. <https://money.cnn.com/2018/09/28/technology/facebook-breach-50-million/index.html>
- [2] *Cyberangriff "WannaCry Hochfahren, einloggen, hoffen* Tagesschau, 2017. <https://www.tagesschau.de/ausland/wannacry-microsoft-103.html>.
- [3] Ole Reißmann. *OpenSSL-Sicherheitslücke – Warum "Heartbleed" Millionen Web-Nutzer gefährdet*. Spiegel Online, 2014. <http://www.spiegel.de/netzwelt/web/heartbleed-openssl-fehler-verraet-passwoerter-a-963381.html>
- [4] B. Potter, G. McGraw. *Software Security Testing*, IEEE Security & Privacy, vol. 2, no. 5, pp. 81-85, 2004. <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=1341418>
- [5] Joe W. Duran, Simeon Ntafos. *A report on random testing*, ICSE '81 Proceedings of the 5th international conference on Software, pp. 179-183, 1981. <https://dl.acm.org/citation.cfm?id=802530>
- [6] Ari Takanen. *Fuzzing: the Past, the Present and the Future*. Actes du 7ème symposium sur la sécurité des technologies de l'information et des communications (SSTIC), pp. 202-212 2009. [https://codeengn.com/archive/Reverse%20Engineering/Fuzzing/Fuzzing%20the%20Past,%20the%20Present%20and%20the%20Future%20\[Ari%20Takanen\].pdf](https://codeengn.com/archive/Reverse%20Engineering/Fuzzing/Fuzzing%20the%20Past,%20the%20Present%20and%20the%20Future%20[Ari%20Takanen].pdf)
- [7] Jason Ansel, Kapil Arya, Gene Cooperman. *DMTCP: Transparent Checkpointing for Cluster Computations and the Desktop*. 23rd IEEE international parallel and distributed processing symposium, Rome, Italy, pp. 1–12, 2009. <http://dmtcp.sourceforge.net/papers/dmtcp.pdf>
- [8] Oren Laadan, Serge E. Hallyn. *Linux-CR: Transparent Application Checkpoint-Restart in Linux* Proceedings of the Linux Symposium, Ottawa, Canada, pp. 159-172, 2010. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.366.6842&rep=rep1&type=pdf#page=159>
- [9] Jason Duell. *The design and implementation of Berkeley Lab's linux checkpoint/restart*. Lawrence Berkeley National Laboratory, 2005. <https://escholarship.org/uc/item/40v987j0>
- [10] Jonathan Corbet. *TCP connection repair*. lwn.net, 2012. <https://lwn.net/Articles/495304/>
- [11] *Man page: pid namespaces*. Linux Man Pages. http://man7.org/linux/man-pages/man7/pid_namespaces.7.html

- [12] Konstantin Serebryany, Derek Bruening, Alexander Potapenko, Dmitry Vyukov. *AddressSanitizer: A Fast Address Sanity Checker*. Usenix Annual Technical Conference, 2012. <https://storage.googleapis.com/pub-tools-public-publication-data/pdf/37752.pdf>