ClamAV-Open-Source-Malware-Detektion

20 Jahre ClamAV: Open-Source-Malware-Detektion

20 Jahre nach der ersten Veröffentlichung des wohl bekanntesten quelloffenen Malware-Scanners legt Cisco zum Jubiläum die Version 1.0 auf. Das Antivirenprogramm ist beliebt, geht einige Wege jedoch anders als die Konkurrenz.

Von Ralf Spenneberg

-tract

- ClamAV ist einer der beliebtesten Open-Source-Malware-Scanner. Als Framework bildet er die Grundlage vieler freier und kommerzieller Sicherheitslösungen, denn Engine und Regeln sind unter der GPL verfügbar.
- Im Unterschied zu anderen Antivirenprogrammen nutzt ClamAV keine Verhaltensüberwachung, sondern legt den Fokus auf die Prüfung von E-Mails – dafür hält es Signaturdatenbanken bereit.
- Mit eigenen Signaturen und YARA-Regeln können Administratoren die Erkennung von ClamAV weiter ausbauen.

Es war ein langer Weg zur großen Eins, die Cisco dem nun 20 Jahre alten ClamAV Ende des Jahres 2022 verlieh. Dabei ist die Zahl mehr ein Symbol als ein Indikator für maßgebliche Neuerungen. Die hat das Programm auch gar nicht nötig, denn die Aufgabe, für die es entwickelt wurde, erledigt es gut – und bildet dabei gleichzeitig Grundlage und Pflichtkomponente

für viele modernere Securitylösungen.

Ursprünglich wurde der Malware-Scanner ClamAV für die Analyse von E-Mails auf E-Mail-Gateways entwickelt. Heute ist es ein Framework, das in beliebige Applikationen integriert werden kann. Im Unterschied zu vielen kommerziellen Produkten ist es aber nicht zur Endpoint Security, also der Absicherung eines Arbeitsplatzrechners, gedacht, denn ClamAV nutzt keine Verhaltensüberwachung. Es beobachtet laufende Prozesse und deren Ausführung also nicht. Deshalb ist ClamAV in den Augen vieler Anwender nicht mit modernen Virenscannern unter Windows vergleichbar. Für E-Mails oder das Erkennen von Phishing-URLs ist so ein Werkzeug jedoch immer noch sehr sinnvoll.

ClamAV agiert als ein signaturbasierter Scanner. Die Signaturen können unterschiedlichster Natur sein, sie können aus Hashes, URLs oder dem E-Mail-Inhalt bestehen, aber auch aus YARA-Regeln, Bytecode, Metadaten von Containern oder PE-Zertifikaten. Hash- und inhaltsbasierte Signaturen können Admins dabei auch selbst erzeugen, zum Beispiel mit dem Werkzeug Sigtool, das die einfache Generierung eigener Signaturen aus Schaddateien erlaubt. Für komplexe Prüfungen, die nicht mit einfachen Signaturen lösbar sind, eignen sich Entwickler-Bytecode-Signaturen. Sie bestehen aus ausführbarem Code, der fast beliebige Eigenschaften in einer Datei prüfen kann. Die von VirusTotal entwickelten YARA-Regeln stellen zusätzlich eine generische Möglichkeit zur Beschreibung von Malware dar.

Wie ClamAV zu Cisco stieß

In den ersten Jahren noch ohne kommerziellen Hintergrund entwickelt, ging das Projekt ClamAV 2007 in den Besitz der Firma Sourcefire über, die die Entwickler übernahm. Sourcefire wurde von Martin Roesch, dem Erfinder des Intrusion-Detection-Systems Snort, gegründet. Sourcefire hat ClamAV anschließend in seine kommerzielle IPS-Lösung integriert (AMP – Advanced Malware Protection).

Die ClamAV-Entwickler gingen im Vulnerability Research Team von Sourcefire auf, das sich auch um die von Snort bereitgestellten Regeln kümmerte. 2013 wurde Sourcefire dann von Cisco gekauft. Das ClamAV-Projekt blieb allerdings weiter quelloffen. Mittlerweile hat Cisco das Vulnerability Research Team in sein Talos-Team aufgenommen, das sowohl für die Betreuung von ClamAV als auch für Snort zuständig ist. ClamAV ist weiterhin eine der Komponenten in den kommerziellen IT-Security-Produkten von Cisco, beispielsweise Firepower.

Fokus auf die E-Mails

Um E-Mails und deren Anhänge effizient analysieren zu können, kann ClamAV eine Vielzahl von Archiven und Dateiformaten automatisch extrahieren und scannen. Dazu gehören Formate von ZIP, RAR, 7zip, arj und tar über DMG, IMG, ISO9660, PKG, HFS+ und GPT bis hin zu OLE2, OOXML, CAB und CHM. Eine vollständige Liste findet sich in der Dokumentation. Noch 2022 hat Splunk die Effizienz von ClamAV in der Erkennung von Malware in unterschiedlichen Dateitypen getestet, wobei speziell die Erkennung in Word-Dokumenten und Windows-Bibliotheken sehr gut abschnitt. Insgesamt hat ClamAV etwa 60 Prozent von 400 000 Virussamples von MalwareBazaar erkannt. Vergleichende Ergebnisse kommerzieller Antiviruslösungen gibt es von Splunk allerdings nicht.

Für die Integration von ClamAV in einen Mailserver nutzte man in der Vergangenheit häufig Werkzeuge wie Amavis oder Amavisd-new. Diese haben die E-Mails zunächst zerlegt, Archive extrahiert und dann die einzelnen Dateien gescannt. Das war nötig, weil viele Virenscanner die Anhänge ansonsten nicht verarbeiten konnten. Da ClamAV das mittlerweile aber auch selbst kann, ist eine solche ressourcenaufwendige Zerlegung heute nicht mehr notwendig. Mit ClamAV-Milter existiert daher auch eine direkte Integration in MTAs über die Milter-Schnittstelle.

Viele Mailserver integrieren ClamAV direkt, beispielsweise

über genau diese Milter-Schnittstelle. Der Filter Amavisd-new hat seit 2018 keine Weiterentwicklung mehr erfahren. Er war in der Vergangenheit in Kombination mit SpamAssassin der bevorzugte Filter in Open-Source-Mail-Gateways. Als moderne Alternative nutzen Administratoren heute häufig den Dienst Rspamd, der neben einer gegenüber SpamAssassin verbesserten Spamerkennung auch ClamAV nutzen kann. Zusätzlich bietet er Module für DKIM-Signaturen, DMARC und weitere Authentifizierungsmöglichkeiten.

Clamd für weniger Ballast

ClamAV wird über die Kommandozeile mit dem Befehl clamscan aufgerufen und lädt dann die sehr umfangreiche Signaturdatenbanken des Scanners. Insgesamt handelt es sich um fast 1 GByte an Signaturen, die bei jedem Start von clamscan geladen und geparst werden müssen, was viel Arbeitsspeicher und CPU-Ressourcen benötigt. Damit das nicht bei jedem Aufruf erforderlich ist, bietet ClamAV eine daemonisierte Version, den Clamd. Der wird als Dienst gestartet und über einen Socket angesprochen. Das kann ein lokaler UNIX-Socket oder ein TCP-Socket sein. Der Befehl clamdscan scannt nun eine beliebige Datei und kommuniziert mit dem Dienst Clamd über einen Socket.

Unter Linux kann ClamAV auch als On-Access-Scanner eingesetzt werden. Hier werden die Dateien bei jedem Zugriff geprüft. Das ist auf Fileservern wie Samba oder Nextcloud interessant. Während der Clamd diese Aufgabe früher selbst übernommen hat, setzen moderne Implementierungen ab Version 0.102.0 dafür einen zusätzlichen Dienst namens Clamonacc ein. Daneben wird ein einigermaßen moderner Linux-Kernel ab Version 3.8 benötigt, der die FANOTIFY-API bereitstellt, über die der Zugriff auf schädliche Dateien verhindert wird. Ohne die API kann auch der On-Access-Scan genutzt werden. Dann kommt es allerdings nur zu einer Protokollierung der gefundenen Dateien und nicht zur Blockade der Zugriffe.

Umwege zur Signatur

Für die Erkennung neuer Viren benötigt ClamAV, wie alle anderen Virenscanner auch, regelmäßige Updates seiner Signaturdatenbank. Die Datenbank wird in drei Dateien vorgehalten: bytecode.cvd, daily.cvd und main.cvd. Bereits bei seiner ursprünglichen Veröffentlichung 2002 führten die Popularität von ClamAV und die daraus resultierenden Downloads zu einem DoS auf den anbietenden Servern. Die Popularität hat in den letzten Jahren nicht nachgelassen, sondern eher zugenommen. Das führte dann im März 2021 zu einer Änderung der Download-Politik im Content Delivery Network von Cisco. Der direkte und vollständige Download der Signaturdatenbanken ist nun nicht mehr erlaubt. Stattdessen muss man Werkzeuge wie Freshclam oder cvdupdate nutzen. Während Freshclam die Datenbank auf dem lokalen System aktualisiert, ist cvdupdate für die Aktualisierung der Datenbank auf einem Spiegelserver verantwortlich. Da cvdupdate im Gegensatz zu Freshclam mehr Daten aus dem Internet laden muss, eignet sich Freshclam besonders für kleine Umgebungen. Beide Werkzeuge prüfen mithilfe von DNS TXT Records zunächst, ob ein Update bereitsteht. Nur wenn die DNS TXT Records ein Update andeuten (siehe Listing), laden sie differenzielle Updates der Datenbank.

Listing: DNS TXT Records weisen auf ein notwendiges Update hin

```
host -t txt current.cvd.clamav.net.
current.cvd.clamav.net descriptive text "0.103.7:62:26768:1672646400:1:90:49192:333"
```

Die einzelnen Werte werden durch Doppelpunkte getrennt und kodieren die folgenden Informationen:
Letzte unterstützte ClamAV Version: 0.103.7
Version der neuesten main.cvd: 62
Version der neuesten daily.cvd: 26768
Zeitstempel der letzten Aktualisierung: 1672646400
Soll eine Versionswarnung erfolgen: 1

Function-Level (f-level): 90
Version der neuesten safebrowsing.cvd: 49192
Version der neuesten bytecode.cvd: 333

Die Virusdatenbank wird täglich aktualisiert, üblicherweise einmal, manchmal zweimal. Die Werkzeuge cvdupdate und Freshclam können aber stündlich oder häufiger aufgerufen werden, da sie die DNS-Einträge prüfen. Wenn viele Clients aktualisiert werden müssen, kann Freshclam die Updates über einen cachenden Proxy laden oder – speziell in einem vom Internet getrennten Netz – auch über einen lokalen Spiegelserver bereitstellen. Die Homepage von ClamAV beschreibt die Einrichtung eines Spiegelservers mit cvdupdate (siehe ix.de/zn4q). Wenn man die Datenbanken via Freshclam aktualisiert, können statt CVD- auch CLD-Dateien lokal gespeichert sein. CVD-Dateien sind dabei schlicht komprimierte Versionen der CLD-Dateien. Um die täglichen Aktualisierungen zu integrieren, verwendet Freshclam die CLD-Dateien und löscht dann die daily.cvd.

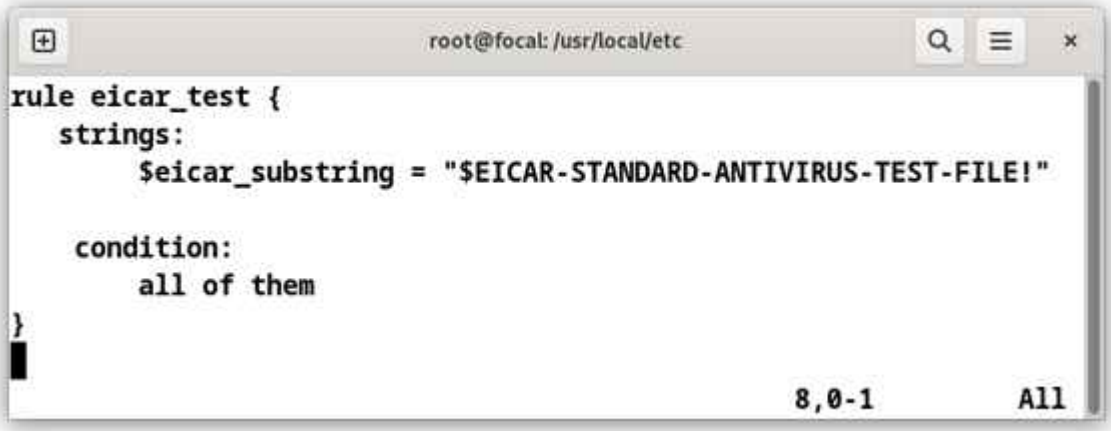
```
root@focal: /usr/local/etc
root@focal:/usr/local/etc# freshclam  -F
ClamAV update process started at Tue Jan  3 11:25:33 2023
daily database available for download (remote version: 26770)
Time:    1.0s, ETA:    0.0s [=========================>]   57.64MiB/57.64MiB
Testing database: '/usr/local/share/clamav/tmp.965f1b7007/clamav-989d89d46fe2d25
b3c559c2701a67b5e.tmp-daily.cvd' ...
Database test passed.
daily.cvd updated (version: 26770, sigs: 2015022, f-level: 90, builder: raynman)
main database available for download (remote version: 62)
Time:    0.1s, ETA:    0.0s [=========================>]        16B/16B
^Can't download main.cvd from https://database.clamav.net/main.cvd
WARNING: FreshClam received error code 429 from the ClamAV Content Delivery Netw
ork (CDN).
This means that you have been rate limited by the CDN.
 1. Run FreshClam no more than once an hour to check for updates.
    FreshClam should check DNS first to see if an update is needed.
 2. If you have more than 10 hosts on your network attempting to download,
    it is recommended that you set up a private mirror on your network using
    cvdupdate (https://pypi.org/project/cvdupdate/) to save bandwidth on the
    CDN and your own network.
 3. Please do not open a ticket asking for an exemption from the rate limit,
    it will not be granted.
WARNING: You are on cool-down until after: 2023-01-04 11:25:41
bytecode database available for download (remote version: 333)
```

Beim ersten Aufruf lädt Freshclam die vollständigen

Datenbanken. Unternimmt Freshclam zu viele Versuche, wird der Client gesperrt (Abb. 1).
Die Datenbank safebrowsing.cvd wird seit 2019 allerdings nicht mehr von ClamAV unterstützt und auch nicht zum Download angeboten. Zugriffe auf diese Datenbank beantwortet das Content Delivery Network mit dem HTTP-Code 403. Grund dafür ist eine Lizenz- und API-Änderung durch Google. Damit Freshclam die Bandbreitenlimitierungen respektiert und die Datenbank nicht mehr lädt, sollte mindestens die Version 0.103.2 eingesetzt werden.

YARA-Anbindung

ClamAV kann mit gewissen Einschränkungen auch YARA-Regeln nutzen. YARA ist ein Projekt von VirusTotal. Es handelt sich um einen eigenen Scanner, der in einer einfachen Sprache geschriebene Regeln unterstützt. Für den berühmten EICAR-Test-Virus könnte die Regel in Abbildung 2 benutzt werden.



Mit YARA-Regeln kann ein Administrator einfach eigene Signaturen erzeugen (Abb. 2).
Viele kommerzielle Malware-Scanner werten YARA-Regeln aus. Ein Administrator kann damit einfach vorhandene Scanner um eigene Regeln ergänzen. In der Community stellen viele Sicherheitsforscher ihre eigenen YARA-Regeln bereit und auch ClamAV-Regeln werden in der Community angeboten. Unter ix.de/zn4q findet man ein Skript, das das Laden dieser inoffiziellen Communityregeln vereinfacht.

Da viele kommerzielle Hersteller ClamAV in ihren Produkten einsetzen, wird das Produkt mit Long-Term Support angeboten. Die LTS-Versionen werden für wenigstens drei Jahre ab ihrer ursprünglichen Veröffentlichung betreut. Aktuell gelten zwei Versionen als LTS: Version 0.103 und die Jubiläumsausgabe 1.0. Weitere Versionen gelten als Feature Release, die für mindestens 4 Monate gepflegt werden – dazu gehören die Versionen 0.104 und 0.105. (kki@ix.de)

1. Quellen
2. Die Studie von Splunk und weitere Ressourcen zu ClamAV finden sich unter ix.de/zn4q.