IT Administrator
Das Magazin für professionelle System- und Netzwerkadministration

Support und Administration

Vergleichstest Remoting-Tools

Fernzugriffe sicher einrichten

Wichtige Daten schützen
Maßnahmen gegen Ransomware

Optimierte Virtualisierung
Hardware für vSphere konfigurieren

Fehlerfrei automatisieren
Best Practices für Bash

Mobile Höchstleitung

Lenovo stellt eine neue mobile Workstation vor – das ThinkPad P16. Der Rechner ist in einem neuen, dünneren und leichteren Gehäuse untergebracht und soll die höchste Leistung aller jemals gebauten mobilen Workstations der ThinkPad-P-Serie bieten. Es ist mit den neuen Intel-12th-Gen-HX-CPUs ausgestattet und kommt mit einer NVIDIA-RTX-A5500-Grafikkarte mit 16 GByte dediziertem Speicher daher. Das ThinkPad P16 ist außerdem mit der neuesten DDR5-Speichertechnologie versehen, die eine Kapazität von bis zu 128 GByte unterstützt sowie mit einer maximalen Speicherkapazität von 8 TByte über Performance-SSDs für höhere Lese-/Schreib- und Datenübertragungsgeschwindigkeiten sorgen will. Das neue ThinkPad P16 wird voraussichtlich im August 2022 ab 1979 Euro verfügbar sein. *(jp)*

Lenovo: www.lenovo.de

Tunnelbauer

Der NCP Secure Windows-Client unterstützt ab der Version 13.0 den WPA3-Standard zur Absicherung von WLAN-Verbindungen. WPA3 hebt das Sicherheitslevel im WLAN durch stärkere Verschlüsselungsalgorithmen und Maßnahmen gegen Brute-Force-Angriffe weiter an. Bei der Nutzung eines öffentlichen Hotspots müssen sich Anwender derweil häufig über eine Website des Hotspot-Betreibers, ein sogenanntes Captive Portal, anmelden und dafür am VPN-Tunnel vorbei kommunizieren.

Die "Secure Hotspot"-Anmeldung soll diesen Vorgang absichern und basiert ab Version 13.0 auf Microsoft Edge. Der Chrome-basierte Webbrowser wird dabei ausschließlich für den Zweck der Anmeldung an einem Hotspot verwendet und nur ihm ist für diesen Augenblick die Kommunikation am Tunnel vorbei erlaubt. Durch die Umstellung auf den Edge-Browser bei der Hotspot-Anmeldung soll die Kompatibilität beispielsweise bei der WLAN-Nutzung in Fernzügen der Deutschen Bahn erhöht werden. *(dr)*

NCP: www.ncp-e.com

Wasserdicht

Zyxel Networks stellt den wetterfesten Outdoor-Access-Point NWA55AXE vor. Das Gerät ist laut Anbieter speziell für raue Umgebungen im Außenbereich konzipiert und eignet sich darüber hinaus zum Erweitern bestehender WLAN-Netzwerke ins Freie. Der NWA55AXE ist abwärtskompatibel zu den älteren Standards IEEE802.11 ac/n/g/b/a und bietet den neuesten WiFi-6-Standard IEEE802.11 ax mit bis zu 1800 MBit/s in den Frequenzbändern 2,4GHz und 5GHz. Zyxel gibt an, dass das Gerät über ein Dual-Band-WiFi-6-Funkmodul verfüge, das bis zu 25 Prozent schnellere Geschwindigkeiten biete als bisherige Wireless-Standards. *(jp)*

Zyxel: www.zyxel.com/de

Windows-Lücke im Visier

Die Ende Mai entdeckte Schwachstelle im Windows-Diagnosetool namens "Follina" schlug im Juni weiter Wellen. So hatten es es Angreifer auf europäische Regierungen und US-amerikanische Verwaltungen abgesehen. Hierfür versendeten sie Phishing-Mails mit manipulierten RTF-Dokumenten, bei denen der Exploit bereits ohne das Öffnen des Word-Dokuments abläuft. Die kritische Sicherheitslücke im Diagnosetool ermöglicht es unter anderem, über ein präpariertes Word-Dokument Schadcode aus dem Internet nachzuladen und auf dem Rechner auszuführen. Im erwähnten Fall der RTF-Dateien genügt es, wenn die Opfer das File im Windows Explorer als Vorschau betrachten. Anfang Juni stürzten sich dann die Akteure des Qbot-Netzwerks auf die Lücke und verteilten darüber ihren Schadcode. Auf Seite 94 stellen wir ein Video des SANS-Institutes vor, in dem die Schwachstelle sowie Schutzmaßnahmen ausführlich beleuchtet werden. *(dr)*

Vergleichstest Remote-Zugriff und -Support

Fernweh stillen

von Thomas Bär und Frank-Michael Schlede

Nicht erst die lang andauernde Pandemie hat es gezeigt: Die Zeit der Turnschuhadministration, in der der IT-Verantwortliche höchstpersönlich zum betreuten System eilte, sind endgültig Vergangenheit. Fernzugriff und Fernwartung sind inzwischen gut etablierte Lösungen. Wir haben uns mit AnyDesk, LogMeIn Pro, TeamViewer und VNC Connect vier bekannte Vertreter dieser Gattung näher angeschaut.


Quelle: nprause – 123RF

Für IT-Profis ist der Fernzugriff eine bewährte Sache: So konnten altgediente Systembetreuer mit dem Einsatz von Telnet glänzen oder einfach mal eben eine RDP-Sitzung auf Windows-Systeme aufsetzen. Doch die Ansprüche an die Disziplin des Fernzugriffs sind gestiegen, sodass inzwischen selbst an der Kommandozeile geschulte Administratoren oft lieber zu einer der vielen Remote-Desktop-Lösungen greifen, die sowohl als freie Software als auch als kommerzielle Lösungen bereitstehen.

Vier Programme im Testfeld

Natürlich können wir mit diesem Test keinen allumfassenden Marktüberblick geben, sondern haben uns vier Produkte aus diesem Softwaresegment herausgesucht. Dazu gehört mit dem Produkt TeamViewer des gleichnamigen Unternehmens aus dem schwäbischen Göppingen ein Werkzeug, das gerade im deutschsprachigen Raum schon fast so etwas wie ein Synonym für Fernwartung geworden ist. Aber auch der ebenfalls schwäbische Konkurrent AnyDesk aus Stuttgart ist sicher nicht nur den rein professionellen Anwendern ein Begriff, wenn es um die Wartung und Betreuung aus der Ferne geht.

Das US-amerikanische Unternehmen LogMeIn aus Boston hat sich nach eigenen Aussagen ebenfalls auf Software aus dem Umfeld der Fernwartung spezialisiert und ist mit LogMeIn Pro in unserem Vergleichstest vertreten. Das vierte Produkt hat zum Teil einen Open-Source-Hintergrund: VNC Connect. Die Software wird vom englischen Unternehmen RealVNC, das aus einem AT&T-Forschungslabor in Cambridge entstand, weiterentwickelt, betreut und vertrieben.

Wir haben alle vier Programme auf Rechnern unter Windows 10 Enterprise und Professional installiert und getestet. Ein Windows Server 2019 kam ebenfalls zum Einsatz. Auch eine Installation auf der Vorabversion von Windows 11 funktionierte bei allen Probanden problemlos. Die Microsoft-Aussage, dass sämtliche Programme, die auf Windows 10 laufen, dies ebenso auf Windows 11 tun werden, trifft also zumindest für die Vorabversion mit der Build-Nummer 22000.51 und die von uns untersuchten Fernwartungswerkzeuge zu.

Für die Überprüfung der Verbindung zu Linux-Systemen kamen Ubuntu in der Version 20.04.2 LTS und Fedora in der Version 34 in virtuellen Maschinen zum Einsatz. Neben der Oberfläche und Bedienbarkeit haben wir uns auch die Möglichkeiten bezüglich Dateitransfer und Kommunikation bei den einzelnen Programmen angeschaut. Die Reihenfolge, in der wir hier im Artikel die einzelnen Programme vorstellen, stellt keine Wertung dar – wir sind einfach nach dem Alphabet vorgegangen.

AnyDesk

Die deutsche Software AnyDesk steht auf der Webseite des Unternehmens in den Versionen Essentials, Performance und Enterprise zur Verfügung. Nutzer können die Performance-Version für einen kostenlosen 14-Tage-Test herunterladen. Dies ist möglich, ohne dass Interessierte dafür Kreditkarteninformationen oder andere Daten zur Zahlung angeben müssen. Eine freie Version der Software für Privatnutzer steht ebenfalls bereit. So ist es problemlos möglich, die Software zunächst einmal unverbindlich im eigenen Netzwerk auszuprobieren.

AnyDesk 6.3.2

Hersteller
AnyDesk Software
https://anydesk.com/de

Preis
AnyDesk steht in den Versionen Essentials, Performance und Enterprise zur Verfügung. Die Essentials-Version kostet 9,90 Euro pro Monat für einen Nutzer, der dann ein Gerät verwalten kann. Die Professional-Version schlägt pro Nutzer und Monat mit 19,90 Euro zu Buche, wobei jeder dieser Nutzer bis zu drei Geräte administrieren kann. Die Enterprise-Version berechnet der Hersteller jeweils als individuelle Lösung.

Systemvoraussetzungen
Windows (Workstation ab XP SP2, Server ab 2003 SP2), macOS ab 10.11 (El Capitan), iOS ab Version 11, Android ab Version 4.4, ChromeOS, FreeBSD (ab Version 10), diverse Linux-Versionen und Raspberry Pi 2.

Technische Daten
www.it-administrator.de/downloads/datenblaetter

Für den Vergleichstest hat uns das Unternehmen eine Performance-Lizenz für einen etwas längeren Zeitraum zur Verfügung gestellt. Uns standen für Windows die Version 6.3.2 und für Linux die Version 6.1.1-1 zur Verfügung. Die Software berechtigt jeden Nutzer, bis zu drei Geräte zu verwalten. Als Nutzer bezeichnet der Hersteller immer den Initiator einer Sitzung. Er muss – ähnlich wie es TeamViewer anbietet – die Software nicht auf seinem System installieren, sondern kann die ausführbare Datei direkt starten und Verbindung zu einem anderen System unter AnyDesk aufnehmen. Dort muss die Software ebenfalls installiert sein und der AnyDesk-Systemdienst muss laufen, damit eine Verbindung funktioniert. Für Admins, die eben schnell mal Hilfe leisten müssen, ist das eine sehr praxisnahe Vorgehensweise.

Alleinstellungsmerkmal Kacheln

Die Oberfläche von AnyDesk unterscheidet sich recht deutlich von den anderen getesteten Programmen. Die Stuttgarter Entwickler setzen komplett auf ein Kachel-Design, wie es Microsoft bei Windows 8 versucht hat. Das tut der Funktionalität der AnyDesk-Software aber keinen Abbruch. Ganz im Gegenteil – wir haben die Oberfläche der Software als übersichtlich, gut strukturiert und leicht zu bedienen empfunden.

Mit der aktuellen Version 6 bietet AnyDesk nun unter anderem eine Zwei-Faktor-Authentifizierung, Unterstützung von Wake-On-LAN und Gruppenrichtlinien. Mit Letzteren ist es möglich, alle Windows-Arbeitsplätze von einem zentralen Ort aus einzurichten und zu konfigurieren.

Während wir bei einer früheren Version (5.2) die AnyDesk-Software auf einem Fedora-System im Windows-Stil installieren konnten, indem wir nach dem Download des Installationspakets einen Doppelklick darauf ausführten, führte das bei der aktuellen Version 6.1.1-1 von AnyDesk zu einer Fehlermeldung. Es galt zunächst, ein entsprechendes Repository einzurichten, um dann die Software samt benötigten Bibliotheken auf das System zu bringen. Danach war es erst einmal nicht möglich, von unserem Windows-10-Rechner auf Fedora-34 zuzugreifen, da dort standardmäßig der Display-Server Wayland aktiv ist, der von AnyDesk nicht unterstützt wird. Durch eine Umstellung auf den X11-Display-Server ließ sich dieses Problem aber lösen.

AnyDesk 6.3.2 So urteilt IT-Administrator

Umsetzung Benutzerkonto	8
Oberfläche und Bedienbarkeit	8
Features für den Dateitransfer	8
Plattformunterstützung	8
Preis/Leistungsverhältnis	7

Die Details unserer Testmethodik finden Sie unter www.it-administrator.de/testmethodik

Sowohl bei Verbindungen zwischen Windows-Systemen als auch zwischen Windows und Linux konnte AnyDesk uns während der Testphase überzeugen. Die Art und Weise, wie die Dateiübertragung als ein wichtiger Faktor bei der Fernwartung gelöst wird, ist bei drei der vier Testkandidaten weitgehend identisch: Der Nutzer erhält zwei Fenster mit dem jeweiligen Quell- und Zielverzeichnis in der Art und Form, wie er es von Anwendungen wie FileZilla oder vielen Klon-Versionen des Windows-Explorers her kennt. Das funktioniert unter AnyDesk auf Anhieb schnell und reibungslos. Die aktuelle Version 6 erwies sich während der Testphase als schnell und zuverlässig.

LogMeIn Pro

Im Gegensatz zu AnyDesk oder TeamViewer ist es bei LogMeIn nicht möglich, die Software einfach direkt und ohne Installation zu nutzen. Wer das Werkzeug einset-

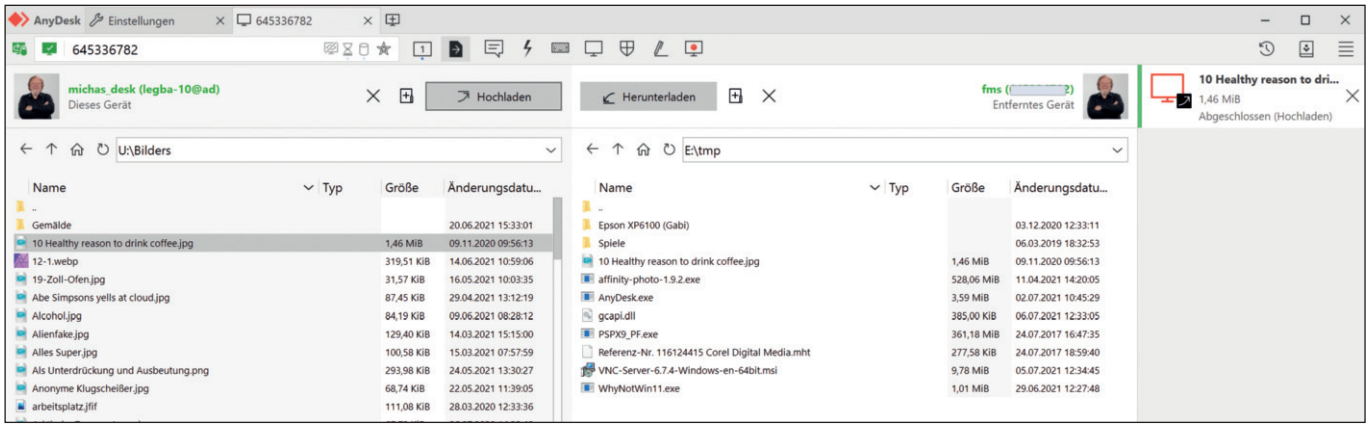

Bild 1: Der Dateitransfer von AnyDesk, hier zwischen zwei Windows-Systemen, ähnelt bekannten Programmen und Explorer-Alternativen.

zen will, muss sich zunächst auf der LogMeIn-Webseite anmelden und ein Konto anlegen. Danach können Nutzer aus dieser Oberfläche heraus Computer hinzufügen, indem sie das Installationsprogramm für den Host herunterladen oder dafür einen Link erstellen, den sie dann an einen Empfänger weiterleiten. Mit der Installation gelangt dann auch der LogMeIn-Client auf das System, über den der Nutzer auf seine mit der Software ausgestatteten Rechner zugreifen kann. Hat er sich auf LogMeIn Central mit seinem Konto angemeldet, findet er dort dann im Menü unter "Netzwerke" den Eintrag "Softwareverteilung". Dort kann er ein eigenes Installationspaket zur Ferninstallation erstellen und einen entsprechenden Link anlegen. Dazu findet standardmäßig die VPN-Software Hamachi von LogMeIn Verwendung.

Die Lösung unterstützt als Hostcomputer – so bezeichnet das Unternehmen die Rechner, auf die der Nutzer aus der Ferne zugreifen kann – nur Windows- und macOS-Geräte. Gerade die fehlende Unterstützung von Linux-Systemen haben wir als echten Nachteil empfunden, da mittlerweile doch recht viele Rechner mit diesen Betriebssystemen zum Einsatz kommen. Eine Beta-Version für Linux steht allerdings auf LogMeIn Central bereit. Dort sowie in den entsprechenden Stores können Nutzer auch Viewer-Apps für Android und iOS finden.

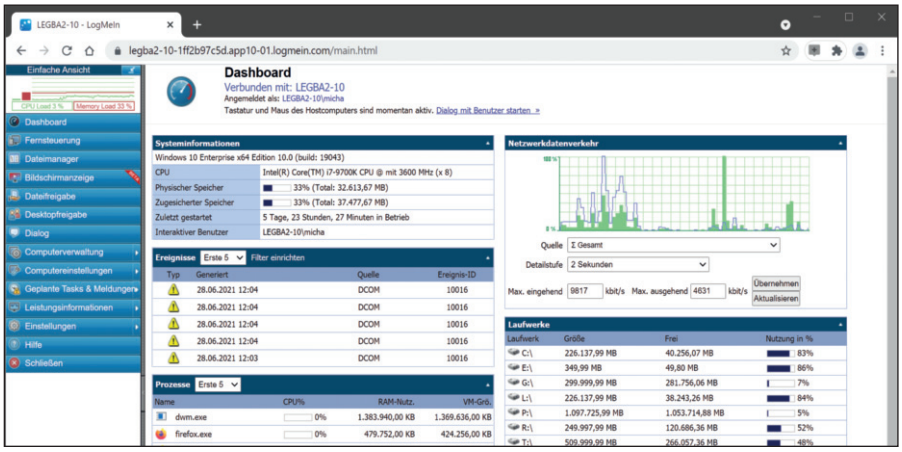

Bild 2: Das Dashboard von LogMeIn Pro bietet umfangreiche Informationen für IT-Profis, die weit über die Auflistung der Zugriffe hinausgehen.

LogMeIn Pro überraschte uns zudem bei der Installation auf einem unserer Windows-10-Testsysteme mit der Meldung, dass wir mit "Eingeschränkter Funktionalität" zu rechnen hätten, wenn wir die Software nicht, wie standardmäßig vorgesehen, im Ordner "C:\ProgrammFiles(x86)\" installieren. Abgesehen davon, dass uns die Software und die Webseite einer Erklärung schuldig blieben, warum das denn so sein sollte, ist das für den Einsatz im professionellen Umfeld nicht gerade ideal, wollen Systemadministratoren die Software von Drittanbietern unter Umständen einheitlich in bestimmten, von ihnen festgelegten Verzeichnissen installieren. Wir haben diese Warnung ignoriert und konnten auf unserem Testsystem keine eingeschränkte Funktionalität feststellen.

Forscher Online-Speicher

Die Oberfläche der Software ist gut gelungen, auch wenn es uns etwas verwunderte, dass die Bildschirmanzeige dort als "Beta" markiert war. Insgesamt bietet das Dashboard für das jeweilige System sehr viel Informationen, die gerade für die Fernwartung sehr interessant sind: So kann der Administrator dort unter "Computerverwaltung" nicht nur Informationen zu Tasks und Diensten auf dem entfernten System aufrufen, sondern er hat unter anderem direkten Zugriff auf den Registry-Editor. Die Übertragung der Bildschirminhalte funktionierte während der Testphase schnell und problemlos.

Die Möglichkeit, Dateien zwischen den Systemen mittels zweier Explorer-Fenster zu übertragen, funktionierte auch hier gut. Allerdings bietet LogMeIn Pro noch eine weitere Möglichkeit zum Teilen von Dateien an. Diese ist im Client unter "LogMeIn Files" zu finden. Wählt der Nutzer diesen Punkt aus, legt die Software auf dem lokalen System ein Laufwerk an, das pro Subskription mit 1 TByte Online-Speicherplatz verbunden ist. Ganz wie es die Nutzer von Diensten wie OneDrive kennen, lassen sich auf diese Weise Dateien für andere Nutzer freigeben. Die Software agierte dabei aber für unseren Geschmack etwas zu forsch, indem sie sich auf dem lokalen Windows-System einfach einen Laufwerksbuchstaben griff, der allerdings

LogMeIn Pro 4.1.1

Hersteller
LogMeIn
www.logmein.com/de/pro

Preis
LogMeIn bieten den Nutzern drei sogenannte Pro-Abos an: Privatanwender zahlen 30 Euro im Monat. Damit kann der Nutzer auf bis zu zwei Systeme zugreifen. Die Version für "sehr aktive Nutzer" schlägt monatlich mit 70 Euro zu Buche und erlaubt den Zugriff auf bis zu fünf Computer. Kleinunternehmen können für 127 Euro im Monat auf bis zu zehn Computer zugreifen.

Systemvoraussetzungen
Auf den von LogMeIn als "Hostcomputer" bezeichneten Systemen (das sind die Rechner, auf die der Nutzer aus der Ferne zugreifen möchte) werden Windows 7 oder neuer, Windows Server ab 2008 R2, Intel-basierende Apple-Systeme unter macOS 10.12 oder neuer unterstützt. Linux und Chromebooks sowie Windows RT finden keine Unterstützung, eine Beta für Linux ist allerdings erhältlich. Zum Zugriff auf die Hostsysteme steht auch eine Android- und eine iOS-App bereit.

Technische Daten
www.it-administrator.de/downloads/datenblaetter

LogMeIn Pro 4.1.1 So urteilt IT-Administrator

Umsetzung Benutzerkonto	4
Oberfläche und Bedienbarkeit	7
Features für den Dateitransfer	6
Plattformunterstützung	5
Preis/Leistungsverhältnis	6

Die Details unserer Testmethodik finden Sie unter www.it-administrator.de/testmethodik

TeamViewer 15.19.5

Hersteller
TeamViewer
www.teamviewer.com/de/

Preis
TeamViewer bietet die Lizenzen "Single User" (Business), "Multi User" (Premium) und "Für Teams" (Corporate) an. Bei "Single User" kann ein Nutzer eine Verbindung mit einem Gerät zu einem monatlichen Preis von 29,90 Euro aufbauen. "Multi User" beinhaltet für 59,90 Euro 15 lizenzierte Nutzer, von denen aber nur jeweils ein Nutzer gleichzeitig eine Remote-Sitzung abhalten kann. Die Corporate Lizenz umfasst 30 lizenzierte Nutzer, von denen drei gleichzeitig eine Remote-Sitzung abhalten können. Sie kostet 129 Euro pro Monat. Zudem stehen diverse Zusatzpakete wie Monitoring & Asset Management, Endpoint Protection, Backup und Web Monitoring kostenpflichtig zur Verfügung.

Systemvoraussetzungen
Die aktuelle Version 15 unterstützt Windows-Systeme ab Windows 7 und Windows Server ab der Version 2008 R2. Bei den Apple-Systemen funktioniert die aktuelle TeamViewer-Version mit macOS 10.11 (El Capitan) und den nachfolgenden Versionen. Die Firma hat nach eigenen Aussagen die Software bereits an die "Apple Silicon Macs" angepasst und kündigt vollen Support dieser Hardware für TeamViewer in Version 15.12 und neuer an. Die Software steht nicht zuletzt für die meisten der aktuellen Linux-Versionen zur Verfügung. Für Mobilgeräte gibt es eine Host-Version von TeamViewer ab Android 4.4 und iOS wird aktuell ab iOS 12 sowie iPadOS 13 unterstützt.

Technische Daten
www.it-administrator.de/downloads/datenblaetter

schon von einer lokalen externen Festplatte belegt war. Auf diese konnten wir dann erst nach dem Ausschalten des LogMeIn-Clients und einem Neustart der Festplatte wieder unter diesem Eintrag zugreifen. Dabei gingen zwar keine Daten verloren, aber es wäre doch zu wünschen, dass die Entwickler von LogMeIn hier etwas mehr Sorgfalt walten lassen.

Insgesamt war die LogMeIn Pro-Software gut und problemlos für die Fernwartung und -betreuung einzusetzen. Wer sie nutzt, muss sich jedoch bewusst sein, dass seine Verbindung über die Cloud und einen Server in den Vereinigten Staaten läuft. Allerdings versichert der Anbieter, dass dabei jede Kommunikation grundsätzlich immer verschlüsselt abläuft.

TeamViewer

Schon die Versionsnummer (zum Testzeitpunkt 15.19.5) von TeamViewer zeigt, dass es sich hier um ein Produkt handelt, das sich schon eine geraume Zeit im Markt behauptet. So ist dann wohl gerade in der Home-Office-Zeit innerhalb der letzten anderthalb Jahre fast jeder Nutzer schon mal von einem Helpdesk-Mitarbeiter aufgefordert worden, "schnell mal den TeamViewer" zu installieren, damit sich der IT-Spezialist das Problem selbst anschauen kann. Ebenso wie AnyDesk können Nutzer die Software einfach herunterladen und sie dann auch ohne Installation direkt zum Verbindungsaufbau nutzen. Das und die Tatsache, dass TeamViewer sein Werkzeug für den privaten Einsatz kostenlos zur Verfügung stellt, dürfte nicht unerheblich zu dessen Verbreitung und Popularität beigetragen haben.

Die Oberfläche der Software hat über die Jahre ein eher geringes Facelifting erfahren. Aber das tut der Funktionalität keinen Abbruch und bietet zudem den Vorteil, dass Nutzer sich aufgrund jahrelanger Erfahrung mit der Software schnell zurechtfinden. Die Entwickler aus Göppingen haben ihrem Produkt im Verlauf der letzten Jahre eine reiche Palette an Erweiterungen hinzugefügt. So präsentiert die Anwendung nach dem Start im Menü auf der linken Seite auch gleich verschiedene Möglichkeiten, auf diese Erweiterungen zuzugreifen: Wählt der Nutzer beispielsweise den Eintrag "Remote Management" aus, so bekommt er die drei Möglichkeiten Monitoring & Asset Management, Endpoint Protection und Backup angezeigt. Ein Klick auf das jeweilige Pluszeichen und der Nutzer wechselt in das Dashboard seines TeamViewer-Kontos, wo er die entsprechenden Einstellungen vornehmen beziehungsweise die Daten auslesen kann. Bei diesen drei Features handelt es sich um Zusatzprodukte, die einzeln lizenziert und pro Monat beziehungsweise pro Monat und Endpoint oder beim Backup nach der Menge des Speichervolumens zu bezahlen sind. Als weiteres Tool steht dann noch eine IoT-Variante der TeamViewer-Software zur Verfügung.

Extras kosten

Zu den Standardfeatures zählt neben der Chat-Funktionalität die Möglichkeit, mittels TeamViewer Meetings abzuhalten. Das funktionierte in einem kurzen Test sehr gut. Die Linux-Version der TeamViewer-Software erfordert etwas Einarbeitung ins Linux-Umfeld bei der Installation, bietet dann aber das exakt gleiche Erscheinungsbild mit den gleichen Funktionen, wie sie unter Windows zu finden sind. Das gilt ebenso für die macOS-Version des Software. Für die Möglichkeit,

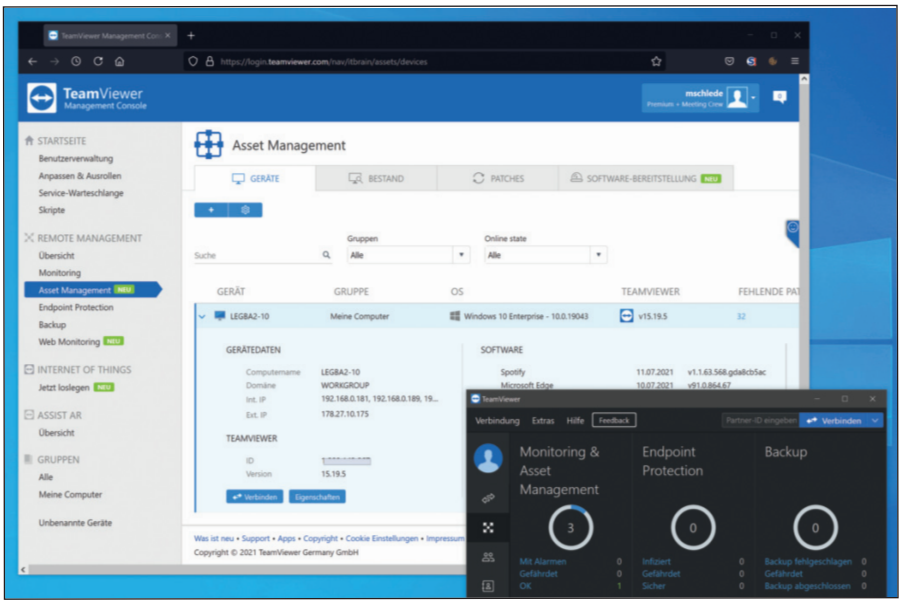

Bild 3: Der Blick auf die Managementkonsole von TeamViewer. Die Software lässt sich um eine ganze Reihe weiterer Fähigkeiten erweitern – wie hier das Asset Management.

TeamViewer 15.19.5
So urteilt IT-Administrator

Umsetzung Benutzerkonto	6
Oberfläche und Bedienbarkeit	7
Features für den Dateitransfer	8
Plattformunterstützung	8
Preis/Leistungsverhältnis	6

Die Details unserer Testmethodik finden Sie unter www.it-administrator.de/testmethodik

Dateien zu übertragen, bietet auch TeamViewer eine "Zwei-Fenster"-Version an, die aber von allen vier Produkten dem Original Windows-Explorer am nächsten kommt, da die Nutzer dort wirklich alle Elemente des Explorers vorfinden.

Was Schnelligkeit und Zuverlässigkeit angeht, zeigte TeamViewer während der Testphase keine Probleme. Verbindungen werden schnell und zuverlässig aufgebaut und den Anwendern stehen neben Chat und Dateiübertragung auch Möglichkeiten wie ein Whiteboard zur Verfügung. Die Einstellungen für die TeamViewer-App findet der Administrator in einem sehr übersichtlichen Menü unter dem Menüpunkt "Extras". Hier kann er sogar eine Vorlage für die Einladungsnachricht entsprechend einrichten.

Dieser Test bestätigte, dass es sich bei TeamViewer um so etwas wie den Rolls Royce der Fernwartungssoftware handelt: Sehr viele Extras und Möglichkeiten, die selbstverständlich ihren Preis haben. Wer viele Features mitsamt Erweiterungen für unter anderem Endpoint-Sicherheit, Backup und Asset-Management benötigt, wird bei diesem Produkt auf jeden Fall fündig.

VNC Connect

Wer mit Linux- oder gar Unix-Veteranen spricht, dürfte häufig den Namen VNC gehört haben: Für diese Technik stehen Remote-Desktop-Clients in den unterschiedlichsten freien und kommerziellen Ausprägungen zur Verfügung. Für unseren Bericht haben wir die VNC-Connect-Software von der Webseite des Unternehmens RealVNC heruntergeladen. Dort steht sie ohne weitere Zugriffsbarrieren als 30-Tage-Testversion bereit.

Wer VNC Connect nutzen will, muss die VNC Server App auf dem entfernten Computer installieren, den er kontrollieren will. Auf diesen kann er dann mit Hilfe des VNC Viewers zugreifen. Die Serverkomponente ist auf jedem der Systeme zu lizenzieren, das kontrolliert werden soll. Will eine IT-Abteilung die Professional- oder Enterprise-Edition von VNC Connect mit der Sofortunterstützung betreiben, müssen Teammitglieder als "Technician" eingetragen sein. Diese Menschen können dann VNC Viewer dazu benutzen, sich in das VNC-Connect-Konto einzuloggen und den Knopf für "Sofortunterstützung" in der oberen Leiste der App auszuwählen. Dann muss der Endnutzer, der die Hilfe benötigt, eine entsprechende App herunterladen und einen neunstelligen Sitzungscode eingeben, den der Administrator im Viewer angezeigt bekommt und an den Nutzer weitergibt. Danach steht die Verbindung und der Techniker kann entsprechend helfen.

Das entspricht durchaus dem Vorgehen, wie es auch bei AnyDesk und TeamViewer möglich ist, nur dass hier eine separate App heruntergeladen werden muss. Das funktionierte im Test gut, da der Besitzer eines Testkontos bei VNC Connect automatisch als "Technician" eingetragen ist.

Um die Ecke gedacht

VNC Connect unterscheidet bei der Verbindung zwischen einer Cloud- und einer direkten Verbindung. Bei der Cloudvariante wird die Verbindung über die Cloudressourcen von RealVNC aufgebaut. Danach kommt dann eine Peer-to-Peer-Verbindung zwischen Server und Viewer zum Einsatz. Ein Vorteil der Cloudverbindung besteht darin, dass der VNC Viewer automatisch die Verbindungen findet. Mit der Enterprise Edition sind dann auch direkte Verbindungen möglich, bei denen keine Verbindung zum Clouddienst von VNC nötig ist. Die Verbindung läuft unmittelbar über die IP-Adresse des Zielsystems. Natürlich muss der Administrator dann dafür Sorge tragen, dass Firewall- und Netzwerkeinstellungen den Zugriff zulassen. Diese Methode ist zum Beispiel für den reinen Einsatz in einem abgeschotteten Firmennetzwerk sinnvoll.

Als einziges der vier Programme nutzt VNC Connect nicht die Darstellung in zwei Fenstern, um Dateien zwischen den Systemen zu übertragen, sondern arbeitet mit den Standarddialogen, wie sie auch im lokalen Explorer bereitstehen. Der Transfer funktioniert zwar auf diese Art, aber wir empfinden die Darstellung mit den zwei Fenstern – gerade für wenig er-

VNC Connect 6.21 (Viewer), 6.74 (Server)

Hersteller
RealVNC
www.realvnc.com/de/connect/

Preis
VNC Connect steht in den Ausprägungen Professional und Enterprise bereit. Einen Unterschied stellen die Verbindungsmöglichkeiten dar: Während die Professional-Version nur Cloudverbindungen ermöglicht, erlaubt die Enterprise-Variante zudem Direktverbindungen. Bei beiden Version gibt es dann noch einmal eine Unterscheidung zwischen Gerätezugriff und "Technikerunterstützung". Fordert der Gerätezugriff eine Installation auf dem Remote-Gerät, funktioniert die Sofortunterstützung auch ohne. Die Version Gerätezugriff wird nach der Zahl der Remote-Computer abgerechnet: 2,79 Euro bei der Professional- und 3,99 Euro bei der Enterprise-Version pro Computer und Monat. Die Variante Sofortunterstützung kostet 13,99 Euro pro Techniker und Monat bei drei gleichzeitigen Sitzungen in der Professional- und 27,99 Euro bei zehn gleichzeitigen Sitzungen.

Systemvoraussetzungen
Der VNC Server unterstützt Windows 7 bis 10 und Windows-Server ab 2008 R2. Apple-Systeme lassen sich ab macOS 10.10 (Yosemite) bis macOS 11 (Big Sur) fernwarten. Im Linux-Bereich läuft die Servervariante auf Ubuntu (ab 16.04 LTS), RHEL 6 bis 8, CentOS 6 bis 8 sowie Suse Enterprise 12. Auch auf Raspberry Pi OS sowie Debian 9 und 10 auf dieser Hardware kann VNC Server zum Einsatz kommen. VNC Viewer findet auf den gleichen Systemen Unterstützung und steht zudem für Android 6 bis 10, iOS 9 bis 14 und iPadOS 13 und neuer bereit.

Technische Daten
www.it-administrator.de/downloads/datenblaetter

fahrene Nutzer – doch als übersichtlicher und besser zu bedienen. Zumal der Nutzer beim Transfer um die Ecke denken muss: Will er von dem zu wartenden Rechner eine Datei empfangen, muss er dort über die Taskleiste auf "Dateiübertragung / Datei senden" klicken. Das ist zwar logisch, aber nicht unbedingt intuitiv und schnell.

Im Vergleich zur "üppigen" Ausstattung der anderen Programme kommt VNC Connect erst einmal sehr viel schlichter daher, ist aber gut für die Aufgaben der Fernwartung und Betreuung gerüstet. Das Prinzip mit den unterschiedlichen Apps halten wir insgesamt für nicht so praktisch, es ist aber im alltäglichen Einsatz sicher kein Hinderungsgrund. Auf der Webseite von Real VNC steht eine sehr ausführliche Dokumentation zu VNC Connect bereit. Aber während VNC Viewer sogar auf den Linux-Systemen mit einer deutschen Oberfläche und einigen deutschsprachigen Hilfstexten auf die Rechner kommt, steht diese Dokumentation nur in englischer Sprache bereit.

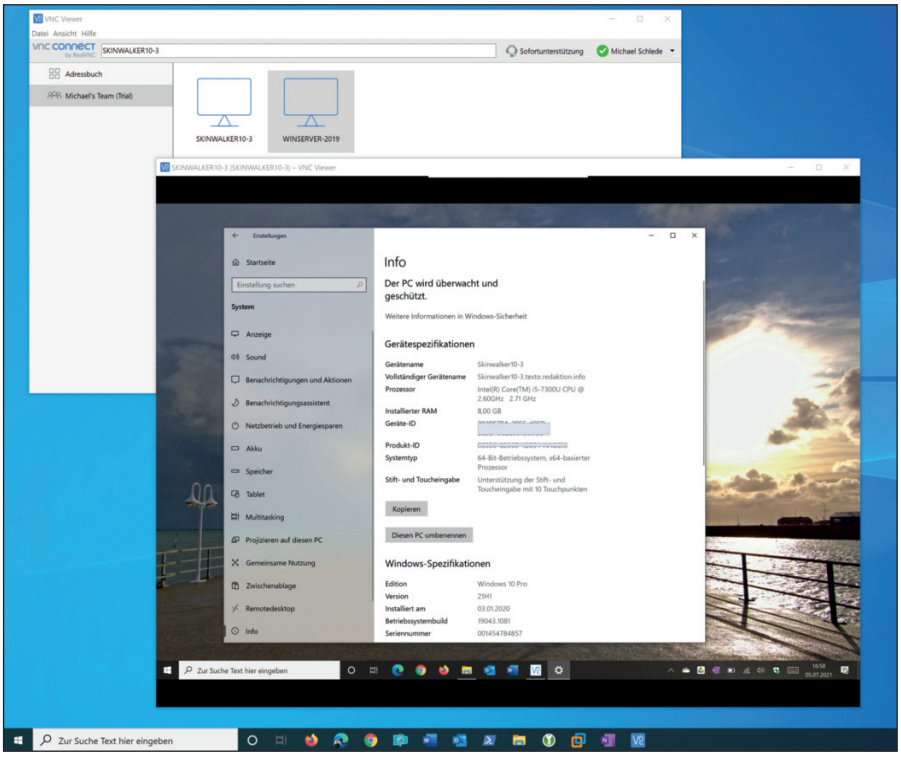

Bild 4: Gerade im LAN kann VNC Connect durch seine Schnelligkeit überzeugen. Der Administrator arbeitet sehr flüssig und direkt auf dem entfernten System.

Fazit

Wir haben diesen Vergleich nicht im Sinne eines "Shootouts" durchgeführt, bei dem es um Biegen und Brechen darum ging, dass eine Software die höchste Punktzahl erhält. Das wäre schon deshalb wenig sinnvoll, weil die vier von uns getesteten Programme alle Standardaufgaben ohne Schwierigkeiten zu unserer Zufriedenheit ausgeführt haben. Negative Ausreißer konnten wir bei keinem der vier Programme ausmachen. Auch wenn dem Testteam die Oberfläche von AnyDesk schon aufgrund der Aufgeräumtheit am besten gefallen hat, bietet doch das gesamte Testfeld dem Nutzer auf den unterschiedlichen Plattformen eine gut zu bedienende GUI.

Die Linux-Kandidaten kamen in der Regel weniger gut mit unserem UHD-Bildschirm und dessen hoher Auflösung zurecht. Das galt bei VNC Connect auch für die Windows-Version. Alle Werkzeuge bieten zumindest einen Viewer-Client für die mobilen Plattformen Android und iOS an. Da es nachweislich keine Freude ist, beispielsweise einen Windows-Rechner auf dem Bildschirm eines Smartphones zu bedienen oder gar zu konfigurieren, dürfte dieser im professionellen Umfeld aber nur in Notfällen zum Einsatz kommen.

TeamViewer bietet ohne Frage die vielfältigsten Möglichkeiten an und lässt sich mit Erweiterungen sehr gut in Richtung umfassendes Systemmanagement ergänzen. AnyDesk ist TeamViewer in vielen Beziehungen aber inzwischen dicht auf den Fersen. LogMeIn Pro hat nach unserer Einschätzung seit den letzten Versionen deutliche Fortschritte gemacht, was sich auch in der Oberfläche der Software deutlich zeigt. Allerdings existieren hier noch einige "Baustellen", so beispielsweise bei der Linux-Unterstützung.

VNC Connect ist ob der strengen Trennung von Server und Viewer etwas gewöhnungsbedürftig, arbeitete im Testzeitraum aber sehr zuverlässig. Im LAN ist die Geschwindigkeit dieser Software unübertroffen. Leider bleibt es den Administratoren auch bei diesen aktuellen Versionen der Fernwartungswerkzeuge, die Linux unterstützen, nicht erspart, dass sie auf diesen Systemen bei der Installation etwas basteln müssen. Das ist aber weniger den Tools als der Diversität der unterschiedlichen Linux-Derivate bei den Installationspaketen geschuldet.

Vor einer Entscheidung sollte die IT-Mannschaft und hier vor allem die Kollegen vom Helpdesk austesten, welche der Werkzeuge ihren Anforderungen am besten entsprechen und welche Anwendung den Endnutzern in der täglichen Praxis am wenigsten Probleme bereitet. Schließlich ist das Verhältnis Kosten und Nutzen ein gewichtiger Faktor: Wer eine Software einsetzen will (oder muss), die für alle Eventualitäten gewappnet ist und die sich später in Richtung Asset- und Endpoint-Management/-Schutz sowie Backup ausbauen lässt, muss selbstverständlich mit einem höheren Preis pro Nutzer und System rechnen. *(ln)*

VNC Connect 6.21 / 6.74
So urteilt IT-Administrator

Umsetzung Benutzerkonto	4
Oberfläche und Bedienbarkeit	6
Features für den Dateitransfer	5
Plattformunterstützung	7
Preis/Leistungsverhältnis	8

Die Details unserer Testmethodik finden Sie unter www.it-administrator.de/testmethodik

Delinea Secret Server 11 Professional Edition

Vertrauliches teilen

von Dr. Christian Knermann

Quelle: zerbor – 123RF

Der Delinea Secret Server hilft, die Vielzahl an Zugangsdaten in den Griff zu bekommen, die nicht nur den Alltag von IT-Administratoren begleiten. Die Software verwaltet lokale Accounts, Domänenkonten, SSH-Schlüssel und weitere schützenswerte Informationen. Ein umfangreiches Rollenkonzept sorgt dafür, dass diese nicht in die falschen Hände geraten. Sämtliche hilfreichen Funktionen bietet jedoch nur die Platinum-Edition.

Der Anbieter Delinea von Produkten rund um das Privileged Account Management (PAM) firmiert unter diesem Namen erst seit Beginn dieses Jahres und doch handelt es sich um ein Unternehmen mit langjähriger Erfahrung. Delinea ist entstanden aus dem Zusammenschluss der Hersteller Thycotic und Centrify. Erstgenannter hat sein Kernprodukt, das zum Zeitpunkt unseres Tests noch den Namen Thycotic Secret Server trug, bereits vor über zehn Jahren erstmals vorgestellt und seitdem eine Familie von dazu passenden Softwareprodukten entwickelt.

Neben dem Secret Server kümmert sich der Privilege Manager darum, mit möglichst wenig lokalen Admin-Rechten auf Endpunkten auszukommen. Beim separat erhältlichen Connection Manager handelt es sich um eine Clientapplikation für Windows und macOS, die nahtlos in den Secret Server integriert ist und Sitzungen verwaltet. Weitere Produkte rund um den Lebenszyklus von Accounts und sicheren Zugriff runden das Angebot ab.

Im Fokus unseres Interesses stand mit dem Secret Server aber das Flaggschiff der Produktpalette, das Delinea sowohl als Cloudangebot wie auch zur Installation on-premises anbietet. In letzterem Fall handelt es sich um eine Webapplikation auf Basis von Microsoft Internet Information Server (IIS) und einer auf Microsofts SQL-Server basierenden Datenbank.

Zugangsdaten sicher im Team verwaltet

Möchte ein Team von Administratoren gemeinsam genutzte Zugangsdaten sicher verwahren, ist der naheliegende Schluss oftmals der Einsatz eines dateibasierten Passwort-Safes. Eine solche Lösung funktioniert für kleine Teams hervorragend, doch ihre Grenzen treten schnell zutage. So haben alle Teammitglieder mit Zugang zur Passwortdatei uneingeschränkt lesenden wie auch schreibenden Zugriff auf alle enthaltenen Informationen. Eine solche Passwortdatei unterstützt weder granulare Berechtigungen noch Freigabeprozesse oder gar Protokollierung.

An diesem Punkt setzt der Secret Server an. Die Software ist auf Basis jährlicher Mietlizenzen pro namentlich benanntem Benutzer erhältlich. Dabei unterscheidet Delinea die Anwender nach IT-Administratoren und Business-Usern. Erstere bilden als Mitarbeiter der IT-Abteilung eines Unternehmens die Kernzielgruppe, die intensiv mit dem Secret Server arbeitet. Bei einem Business-Nutzer handelt es sich um einen User außerhalb der IT-Abteilung – etwa im Einkauf –, der die Daten zur Firmenkreditkarte hinterlegen möchte. Oder in der Marketingabteilung, die sämtliche Zugangsdaten zu den Social-Media-Konten des Unternehmens sicher verwahren möchte.

Erweiterte Funktionen nur mit Platinum

Delinea bietet den Secret Server in zwei Ausbaustufen als Professional- und als Platinum-Edition an. Beide verwalten lokal bis zu 10.000 geheime Schlüssel, nur in der Cloud sind es unbegrenzt viele. Unter einem geheimen Schlüssel versteht Delinea nahezu beliebige Datensätze schützenswerter Informationen und gibt dem Secret Server ab Werk über 40 sogenannte Schablonen für geheime Schlüssel mit, die die Struktur der Datensätze beschreiben. So besteht etwa die Schablone für ein lokales Windows-Konto aus Textfeldern für Maschinen- und Benutzernamen, einem Feld für das Passwort und einem Hinweisfeld für Notizen. Die Schablone für ein Active-Directory-Konto ersetzt das Textfeld für den Namen der Maschine durch eines für die Domäne.

Neben diversen Spielarten von SSH- und Telnet-Zugängen kennt der Secret Server etwa auch SSH-Schlüssel, Anmeldungen an Datenbanken, Mainframes und Firewalls verschiedener Hersteller sowie an den Clouds der drei

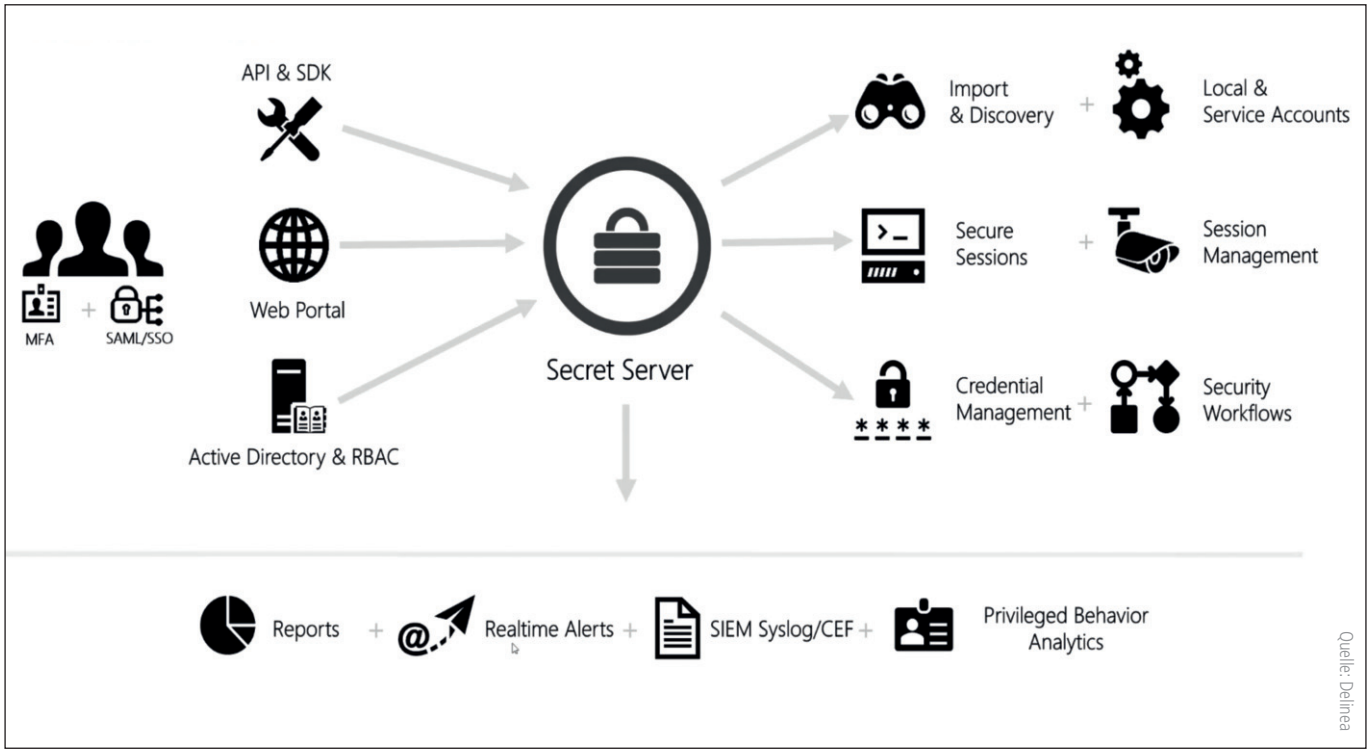

Bild 1: Der Secret Server positioniert sich als umfassendes Werkzeug zum Umgang mit geheimen Schlüsseln.

großen Hyperscaler Amazon, Google und Microsoft. Wer nicht fündig wird, darf mit dem Designer für Schablonen per Webfrontend selbst Vorlagen erstellen oder per Import/Export im XML-Format mit anderen Nutzern des Secret Servers austauschen.

Über die manuelle Eingabe geheimer Schlüssel und den Import aus anderen Systemen hinaus erkennen Professional und Platinum auf Wunsch lokale Konten und solche im Active Directory automatisch. Sie können gleichermaßen geheime Schlüssel zeit- oder ereignisgesteuert ändern und über deren Verwendung mittels erweiterter Audit- und Report-Funktionen Auskunft geben. Beide Editionen arbeiten mit diversen Produkten und Standards für Mehrfaktorauthentifizierung sowie Single Sign-on zusammen. Und beide verschlüsseln die ihnen anvertrauten Informationen nach dem Standard AES-256 auf Applikationsebene, auf Wunsch in Verbindung mit einem Hardware Security Module (HSM).

Mit einigen erweiterten Funktionen setzt sich die Platinum-Edition ab. So kann sie auch Service-Accounts, wie sie unter Windows für Dienste und geplante Tasks zum Einsatz kommen, mit ihren Abhängigkeiten verwalten. Sie bietet weiterhin integrierte Freigabe-Workflows, sodass Nutzer beim Admin den Zugriff auf geheime Schlüssel anfordern können. Auf Wunsch ändert der Secret Server den geheimen Schlüssel nach der Verwendung umgehend.

Erweiterter Schutz von Zugangsdaten für Unix- und Linux-Systeme mit automatischer Rotation von SSH-Schlüsseln, fortgeschrittenes Scripting sowie Mechanismen für hohe Verfügbarkeit gehören ebenfalls ab Werk zum Umfang der Platinum-Edition. All diese Funktionen sind zwar auch für die Professional-Variante verfügbar, jedoch nur in Form kostenpflichtiger Add-Ons.

Windows Server, SQL Server und IIS als Basis

Der Secret Server unterstützt die 64-Bit-Varianten von Windows Server 2012 bis 2019, sofern es sich mindestens um die Standard-Edition handelt. Microsofts Client-Betriebssysteme, der frühere Small Business Server sowie die Essentials-Varianten neuerer Server eignen sich nicht. Weiterhin darf der Server weder als Domaincontroller noch als SharePoint-Server fungieren.

Delinea Secret Server 11 Professional Edition

Produkt
Software für das Privileged Access Management (PAM).

Hersteller
Delinea
https://delinea.com/products/secret-server/

Preis
Die Lizenzierung erfolgt auf Basis einer jährlicher Mietlizenz pro namentlich zu benennendem Benutzer und Funktionsumfang. Die Kosten liegen bei einem typischen KMU-Betrieb mit bis zu 250 Nutzern in der Größenordnung von 3000 und 8000 Euro pro Jahr.

Systemvoraussetzungen
Software: Microsoft Windows Server 2012 bis 2019, Microsoft Internet Information Server (IIS) 7 oder neuer, Microsoft .NET 4.8 oder neuer, Microsoft SQL Server 2012 bis 2019 mit Sortierung auf Serverebene (Collation) "SQL_Latin1_General_CP1_CI_AS".

Minimale Hardware-Anforderungen: Zwei CPU-Kerne und 4 GByte RAM; 25 GByte Festplattenplatz für den Webserver und 50 GByte Festplattenplatz für den Datenbankserver.

Clients: Alle aktuellen Browser, optional Protokoll-Handler für Windows und macOS.

Technische Daten
www.it-administrator.de/downloads/datenblaetter

Bild 2: Geheime Schlüssel verwaltet der Secret Server in Ordnern mit fein abgestuften Berechtigungen.

Generell empfiehlt der Hersteller, zwei separate Systeme zu verwenden, eines für den Webserver und ein weiteres für die Datenbank. Den SQL-Server unterstützt Delinea sowohl in der Standard- als auch der Enterprise-Edition mitsamt sämtlichen Hochverfügbarkeitskonzepten, die Microsoft anbietet. Die Express-Edition akzeptiert Delinea grundsätzlich auch, weist aber darauf hin, dass diese Ausgabe des Datenbankservers sich eher für kleinere Testinstallationen eignet und im praktischen Betrieb eine reduzierte Benutzererfahrung zur Folge haben kann.

Im Rahmen unseres Tests wagten wir dennoch die Installation aller Komponenten auf nur einem virtuellen Server, was in unserer sehr kleinen Umgebung keine Probleme bereitete. Dazu hatten wir eine frische Installation vom Windows Server 2019 vorbereitet und dieses System als Mitglied in unsere Domäne aufgenommen.

Von Delinea hatten wir eine für 30 Tage gültige Trial-Lizenz für den Funktionsumfang der Platinum Edition zur lokalen Installation erhalten. Nachdem wir die DSGVO- und Lizenzbedingungen akzeptiert hatten, luden wir die Setuproutine für den Server sowie die Connection-Manager-Clients für Windows und macOS herunter.

Die Setuproutine enthält sowohl den Secret Server als auch den Privilege Manager. Wir entschieden uns im ersten Dialogschritt, nur Ersteren zu installieren. Daraufhin durften wir wählen, SQL Server Express einzurichten oder die Verbindung zu einem bestehenden SQL-Server aufzunehmen. Auch hier war die erste Option für unser Vorhaben die richtige. Der Assistent überprüfte daraufhin sämtliche Voraussetzungen und scheiterte zunächst am fehlenden .NET Framework sowie dem IIS mitsamt seinen Komponenten.

Installation der Voraussetzungen nur halbautomatisch

Mit Hilfe der Schaltfläche "Fix Issues" veranlassten wir die Installation und Konfiguration der fehlenden Elemente, mussten aber nach der Installation von .NET zunächst manuell einen Neustart initiieren. Anschließend scheiterte die Überprüfung erneut aufgrund fehlender Komponenten des IIS und auch der Button "Fix Issues" half uns nicht weiter. Immerhin teilte uns der Assistent im Detail mit, welche Features des IIS er vermisste. So mussten wir über den Servermanager die IIS-Verwaltungskonsole sowie acht weitere Features manuell hinzufügen. Anschließend konfigurierte die Setuproutine selbsttätig ein selbstsigniertes Zertifikat für das Webfrontend des Secret Servers, verbunden mit der Warnung und Empfehlung, dass ein Zertifikat einer vertrauenswürdigen Zertifizierungsstelle zum Einsatz kommen möge.

Nun blieb uns nur noch, Namen und Passwort für den initialen Admin-Benutzer des Secret Servers festzulegen und eine SMTP-Konfiguration für den Versand von Benachrichtigungen zu hinterlegen. Eine funktionierende SMTP-Konfiguration ist Voraussetzung, damit lokale Nutzer später ihr Passwort zur Anmeldung am Webfrontend zurücksetzen können, falls sie es vergessen sollten.

Der Assistent präsentierte uns abschließend eine Zusammenfassung der Optionen, die wir per Klick auf "Install" quittierten. Den Rest erledigte das Setup dann komplett eigenständig, lud die aktuelle Version von SQL Server Express herunter und leitete uns schließlich zur Anmeldung am Webfrontend des Secret Servers.

Basiskonfiguration schnell erledigt

Dort konnten wir uns mit unserem initialen Admin-Konto anmelden, einem lokalen Benutzer, der nur innerhalb der Datenbank existiert. Das übersichtliche Dashboard gibt mit seiner vertikalen Menüstruktur auf der linken Seite keine Rätsel auf. Bevor wir uns an die Verwaltung von geheimen Schlüsseln begeben konnten, folgten wir den

Handreichungen in der Onlinedokumentation der Software und erledigten einige Voreinstellungen.

Zunächst trugen wir unsere Lizenzen im Bereich "Verwaltung / Lizenzen" ein und mussten diese anschließend online aktivieren. Für Server ohne Internetkontakt bietet Delinea hier alternativ ein Offlineverfahren an. Der Hersteller empfiehlt, auch bei der Verknüpfung mit einem AD den initialen Benutzer sicher aufzubewahren, sodass immer ein Konto zur Anmeldung am Secret Server zur Hand ist. So aktivierten wir im Bereich "Verwaltung / Konfiguration" auf der Registerkarte "Kennwörter lokaler Benutzer" das Zurücksetzen vergessener Kennwörter, um für den Ernstfall gerüstet zu sein.

Weiterhin aktivierten wir auf der Registerkarte "Anmelden" die Option mit dem sperrigen Namen "AD-Berechtigungsnachweise im Cache zur Verwendung bei Offline-Engines speichern", was nichts anderes besagt, als dass der Secret Server die Anmeldung mit einem AD-Account für 30 Tage im Cache behält, falls das AD nicht erreichbar sein sollte.

Einfache Verzahnung mit dem Active Directory

Im Bereich "Verwaltung / Verzeichnisservices" konfigurierten wir dann die Verbindung zu unserer lokalen AD-Domäne. Alternativ dazu spricht der Secret Server auch mit Azure-AD- und OpenLDAP-Domänen. Für unser lokales AD konnten wir optional die Verbindung per LDAPS aktivieren und mussten einen ersten geheimen Schlüssel zur Anmeldung an der Domäne hinterlegen. Ebenfalls optional unterstützt Delinea für dieses Konto eine Mehrfaktorauthentifizierung per FIDO2-Token, TOTP-Authenticator oder Bestätigung per E-Mail.

In den Eigenschaften der Verbindung fügten wir dann auf der Registerkarte "Gruppen" eine oder mehrere Gruppen aus unserer Domäne hinzu, deren Mitglieder der Secret Server anschließend synchronisierte. Diese Synchronisation ist auf der Registerkarte "Verwaltung / Verzeichnisservices / Konfiguration" standardmäßig mit einem Intervall von einer Stunde aktiv. Secret Server legt automatisch für alle Mitglieder der konfigurierten Gruppen aktive Anmeldungen am Webfrontend an. Das funktioniert natürlich nur, solange Lizenzen vorhanden sind.

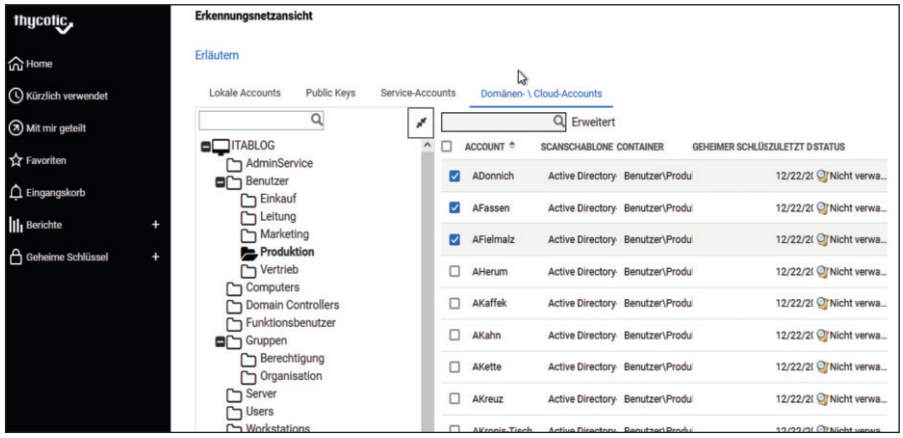

Bild 3: Die automatische Erkennung ermittelt zu verwaltende Konten auf lokalen Systemen, im AD und in Cloudumgebungen.

Umfangreiches Sicherheitskonzept

Unsere aus dem AD synchronisierten Benutzer fanden wir im Dashboard unter "Verwaltung / Benutzerverwaltung" wieder, wo wir sie einzeln oder anhand ihrer AD-Gruppe mit Rollen versehen konnten. Der Secret Server bringt dazu ein weitläufiges Konzept für Role Based Access Control (RBAC) mit und hatte alle unsere Benutzer zunächst mit der Rolle "User" ausgestattet. Alternativ dazu durften wir aus den vorgefertigten Rollen "Administrator", "Basic User" und "Read Only User" wählen oder aber benutzerdefinierte Rollen modellieren. Dazu stellt der Secret Server insgesamt weit über 100 einzelne Berechtigungen bereit, die alle Aspekte der Arbeit mit dem System granular regeln. Für den Einstieg reichte uns aber zunächst die Unterscheidung in Administratoren und Benutzer. Letztere dürfen geheime Schlüssel hinzufügen und bearbeiten, jedoch nicht die Konfiguration des Servers verändern.

Zusätzlich zu den Rollen regeln Ordnerberechtigungen detailliert, wer welche geheimen Schlüssel sehen und nutzen darf. Um uns davon zu überzeugen, legten wir im Bereich "Geheime Schlüssel" des Hautmenüs manuell Unterordner und darin einige geheime Schlüssel, Zugangsdaten zu Windows- und Linux-Systemen sowie unserer Virtualisierungsinfrastruktur an (Bild 2). Der Bereich "Persönliche Ordner" mit allen Ordnern und geheimen Schlüsseln darunter dient, wie sein Name vermuten lässt, der Aufbewahrung individueller Zugangsdaten, die nur der jeweils angemeldete Benutzer sieht.

Alle übrigen Ordner außerhalb dieser Struktur konnten wir mit Berechtigungen versehen und so für andere Benutzer zugänglich machen. Der Secret Server unterscheidet dabei Berechtigungen für Ordner und für einzelne geheime Schlüssel, die sich ähnlich den Berechtigungen in einem Dateisystem auf untergeordnete Objekte vererben. Eigentümer haben Vollzugriff. Alternativ konnten wir die Berechtigungen auf Bearbeiten, Anzeigen oder lediglich Auflisten beschränken. Indem wir uns mit verschiedenen Benutzern am Dashboard anmeldeten, überzeugten wir uns davon, dass das RBAC-Konzept tadellos funktionierte und die einzelnen Nutzer nur auf die geheimen Schlüssel zugreifen konnten, für die sie tatsächlich berechtigt waren.

Konten schnell importiert

Über die feinteiligen Berechtigungen hinaus wusste der Secret Server beim Importieren und automatischen Erkennen von Zugangsdaten zu überzeugen. So fanden wir in der Online-Dokumentation den Download des "Migration Tools", eines Hilfsprogramms, das wir ohne Installation unter Windows ausführen konnten. Das Werkzeug führte uns durch den Export aus den Passwort-Safes KeePass in den Versionen 1.x und 2.x, Password Cor-

ral, Password Safe sowie Passwords Max bis zum XML-Import über das Dashboard des Secret Servers.

Auch ohne Import aus einer bestehenden Lösung mussten wir nicht sämtliche geheimen Schlüssel in unserer Umgebung von Hand eintragen. Als mächtiges und komfortables Werkzeug erwies sich die automatische Erkennung, die wir unter "Verwaltung / Erkennung" konfigurieren konnten. Eine sogenannte Erkennungsquelle legt das Ziel eines Scans fest. Der Secret Server untersucht wahlweise eine gesamte AD-Umgebung mit allen darin enthaltenen Computern auf Domänen-Benutzer und lokale Accounts ab. Letzteres setzt voraus, dass der jeweilige Zielcomputer zum Zeitpunkt des Scans am Netz ist. Gleiches gilt für einen Scan vom Typ "UNIX", der Konten aller in einem definierten IP-Adressbereich auffindbaren UNIX- und Linux-Maschinen erkennt.

Laut Empfehlung von Delinea sollte eine Erkennung kein einmaliges Ereignis sein, sondern regelmäßig laufen, sodass sich Änderungen in der Umgebung erfassen lassen und keine Accounts außerhalb der Verwaltungshoheit des Secret Servers unentdeckt bleiben. Standardmäßig verwendet der Secret Server ein tägliches Erkennungsintervall. Als weitere Erkennungsquellen dürfen VMware-ESXi-Infrastrukturen sowie Tenants in den Clouds von Amazon und Google dienen.

Exemplarisch definierten wir in unserer Testumgebung Erkennungsquellen der Typen "Active Directory" und "UNIX". Für den AD-Zugriff verwendeten wir den Account eines Domänen-Admins, für die Linux-Maschinen jeweils ein lokales Konto. Sobald wir die Scans ausgeführt hatten, präsentierte uns das Dashboard in der "Erkennungsnetzansicht" das Ergebnis in Form einer Baumansicht, in der wir für unsere AD-Umgebung und die lokalen Linux-Maschinen die erkannten lokalen und Domänen-Accounts auflisten und in den Secret Server importieren konnten (Bild 3).

Automatische Kennwortänderungen

Beim Import eines oder mehrerer Accounts konnten wir den Typ des geheimen Schlüssels, also die gewünschte Schablone, den Zielordner im Secret Server und den Namen des geheimen Schlüssels auswählen. Beim Import mehrerer Konten arbeitet der Secret Server mit Variablen der Form "$DOMAIN\$USERNAME", sodass jeder geheime Schlüssel mit einem individuellen Namen in der Datenbank erscheint.

Beim Import ließ sich pro Account das zugehörige Passwort, falls bekannt, von Hand eintragen. Alternativ wiesen wir den Secret Server an, das Passwort des jeweiligen Kontos zu ändern, entweder auf einen manuell festgelegten Wert oder aber auf ein automatisch generiertes Passwort.

Für alle verwalteten geheimen Schlüssel konnten wir auch später noch ihr jeweiliges Passwort ändern und das auf Wunsch regelmäßig und automatisch nach einem definierten Zeitplan. Dazu fanden wir in den Eigenschaften eines jeden Datensatzes die Registerkarte "Ferne Kennwortänderung". Im Fall von Domänen-Accounts durften wir hier zwischen zwei Methoden wählen. Entweder der Secret Server verwendet bei der Kennwortänderung das Kennwort des Benutzers selbst oder ein privilegierter Account, also der geheime Schlüssel eines Domänen-Administrators oder Konten-Operators, ändert das Kennwort für den betroffenen Benutzer.

Remote-Sitzungen solide gelöst

Um ein Kennwort zu verwenden, konnten wir es – Leseberechtigung vorausgesetzt – aus dem Webfrontend in die Zwischenablage befördern. RDP-Sitzungen zu Windows-Systemen oder PuTTY-Sitzungen zu UNIX und Linux initiiert der Secret Server alternativ direkt aus dem Dashboard heraus. Dazu fanden wir in den Eigenschaften von Windows- und Linux-Konten jeweils Links zum "RDP-Startprogramm" und dem "PuTTY-Startprogramm".

Dahinter verbirgt sich der Protokoll-Handler, den der Secret Server bei erster Verwendung zur Installation unter Windows und macOS anbietet. Nach seiner Installation nimmt dieses Werkzeug Links aus dem Dashboard entgegen und startet die passende Remote-Sitzung. Mehr Komfort beim Aufbau von Sitzungen verspricht der separat zu lizenzierende Connection Manager. Darüber hinaus bietet Delinea auch die Integration in zahlreiche Produkte von Drittanbietern, wie etwa Devolutions Remote Desktop Manager, an.

Fazit

Der Secret Server positioniert sich als umfassende Software für das Privileged Access Management überall dort, wo dateibasierte Ansätze nicht mehr ausreichen. Besonders gut gefallen hat uns das RBAC-Konzept. Die Berechtigungen sorgen dafür, dass Nutzer nur Zugriff auf die geheimen Schlüssel erhalten, die sie auch tatsächlich benötigen.

Nicht vergessen werden darf, dass der Secret Server bei konsequentem Einsatz eine betriebskritische Komponente darstellt. High Availability und Disaster Recovery sollten folglich implementiert und getestet sein. Entsprechende Features offeriert der Hersteller als Add-on für die Professional Edition oder im Umfang der Platinum Edition. *(ln)*

So urteilt IT-Administrator

Schablonen für geheime Schlüssel	8
AD-Integration	7
Rollen und Berechtigungen	7
Automatische Erkennung	8
Protokoll-Handler	6

Die Details unserer Testmethodik finden Sie unter www.it-administrator.de/testmethodik

Dieses Produkt eignet sich

optimal für mittlere und große Unternehmen, die die Zugangsdaten ihrer Nutzer sicher verwalten möchten.

bedingt für sehr große Organisationen mit vielen Teams. Je nach Anforderungen ist hier die Platinum Edition eher geeignet.

nicht für kleine Teams, denen ein dateibasierter Passwort-Safe ausreicht.

Neugierde geweckt?

Profitieren Sie
vom Plus
an Wissen

Im Schnupperabo mit **sechs**
Ausgaben zum Preis von **drei**

Heinemann Verlag
Im Dialog mit Spezialisten.

Abo- und Leserservice
IT-Administrator
vertriebsunion meynen
Herr Stephan Orgel
D-65341 Eltville
Tel: 06123/9238-251
Fax: 06123/9238-252
leserservice@it-administrator.de

it-a.eu/abo63

shop.heinemann-verlag.de

Schutz vor Ransomware

Der Feind in meinem Netz

von Dr. Matthias Wübbeling

Ransomware ist seit einigen Jahren auf dem Vormarsch. Unternehmen, die bisher nicht selbst betroffen waren, wissen dabei, dass es nur eine Frage der Zeit ist, bis ein Angriff gegen sie gerichtet wird. Doch ob dieser erfolgreich ist und wie weit ein Angreifer tatsächlich kommt, hängt von vielen unterschiedlichen Faktoren ab. Daher gilt es, sich auf Attacken vorzubereiten und die Abwehrmechanismen in Stellung zu bringen.

Die Anzahl der Cyberangriffe mit Ransomware steigt seit einigen Jahren kontinuierlich. Neben Emotet hat die Ransomware WannaCry im Jahr 2017 für großes Aufsehen gesorgt. Über eine Sicherheitslücke in Microsofts SMB-Protokoll wurden hunderttausende Windows-Systeme infiziert und die darauf vorhandenen Daten wurden verschlüsselt. Die Malware nutzte für die Verbreitung den von einer Hackergruppe veröffentlichten NSA-Exploit namens EternalBlue. Zwar gab es bereits rechtzeitig vor dem WannaCry-Ausbruch einen Patch von Microsoft, der diese Lücke schloss, allerdings waren viele Systeme noch nicht aktualisiert und damit weiterhin verwundbar.

Eher zufällig fand der britische Sicherheitsforscher Marcus Hutchins eine Möglichkeit, WannaCry zu deaktivieren. Der Schädling prüft vor der Verschlüsselung von Dateien die Existenz einer speziellen Domain. Ist diese nicht erreichbar, beginnt WannaCry mit der Verschlüsselung. Nach der Registrierung dieser Domain im weltweiten DNS-System war damit nach nur vier Tagen die weitere Verbreitung zunächst gestoppt. Bis dahin wurden bereits Bitcoin im Wert von mehreren Hunderttausend Euro auf das Wallet der Angreifer überwiesen.

Große Aufmerksamkeit in den deutschen Medien erreichte etwa die Verschlüsselung von mehr als 30 Servern im Computernetzwerk des Universitätsklinikums Düsseldorf im Herbst 2020 durch eine modifizierte WannaCry-Variante. Infolge der Nichtverfügbarkeit der IT in der Uniklinik verstarb sehr wahrscheinlich eine Patientin, da sie nicht aufgenommen werden konnte und in ein weiter entferntes Krankenhaus transportiert werden musste.

Einfallstore für Ransomware

Das Ziel von Ransomware ist in den meisten Fällen die Erpressung eines Lösegelds. Emotet und EternalBlue sind nur zwei von vielen Möglichkeiten, die Kriminelle haben, um die Kontrolle über Computer zu übernehmen und vorhandene Dateien zu verschlüsseln. Das Emotet-Botnet wurde Anfang 2021 von Ermittlern der Europol zerschlagen und die Infrastruktur abgeschaltet. Dieser große Erfolg europäischer Behörden konnte jedoch nur für einen kurzen Moment die Bedrohungslage entspannen. Bereits im November waren die ersten modifizierten Emotet-Varianten im Umlauf.

Neben E-Mails und Sicherheitslücken sind auch geleakte Logindaten von Mitarbeitern regelmäßig ein Einfallstor in Unternehmensnetzwerke. Das zeigt der Fall des US-amerikanischen Pipelinebetreibers Colonial Pipeline [1]. Hier haben die Angreifer offenbar im Darknet gültige Kontodaten eines Mitarbeiters gefunden und konnten diese verwenden, um per VPN auf die Computersysteme des Unternehmens zuzugreifen und Ransomware zu installieren. Als Folge der verschlüsselten Dateien und der Lösegeldforderung hat Colonial das gesamte Pipelinesystem für mehr als fünf Tage heruntergefahren. Die so eingetretene Versorgungslücke mit Treibstoff führte zu einem sprunghaften Anstieg der Kraftstoffpreise in Teilen der USA.

Colonial bezahlte umgerechnet fast vier Millionen Euro Lösegeld an die Erpresser. Diese haben sich jedoch nicht darauf verlassen, die Dateien von Colonial einfach nur zu verschlüsseln. Vielmehr kopierten sie zuvor fast 100 GByte an Files von den Colonial-Systemen. Die Erpresser drohten zusätzlich damit, diese Informationen zu veröffentlichen, sollte Colonial das Lösegeld nicht bezahlen. Dieses Vorgehen, Dateien vor der Verschlüsselung zu kopieren, eröffnet einen weiteren Angriffsvektor. Betroffene Unternehmen bezahlen dann auch, wenn die Daten mit einem Backup einfach wiederherstellbar sind, um eine Veröffentlichung interner Daten

und möglicher Unternehmensgeheimnisse zu verhindern.

Täter und Opfer

Die Täter großer Ransomwarevorfälle sind zumeist gut organisierte Gruppen. Im Rahmen von Ermittlungen und Veröffentlichungen im Anschluss an die Vorfälle werden die Hacks immer wieder zu Hackergruppen in Osteuropa oder Russland zugeordnet. Diese Attribuierung findet vor allem im Rahmen der Analyse der Schadsoftware und der Kommunikation der Täter mit ihren Opfern statt. Aber auch kleine Tätergruppen nutzen Ransomware, wenngleich sie diese nicht selbst entwickeln. In einschlägigen Foren gibt es Baukästen für Ransomware bereits für geringe Beträge.

Quelle: jaretera – 123RF

Game over: Sehen Sie eine Meldung wie diese auf dem Bildschirm, hat die Ransomware – in dem Fall WannaCry – zugeschlagen.

Betrachten wir die öffentliche Berichterstattung über Ransomware-Vorfälle, erhalten wir das Gefühl, dass vor allem große Institutionen und öffentliche Einrichtungen im Visier stehen. Dieses Bild täuscht jedoch. Kleine und mittelständige Unternehmen sind ebenso Opfer von Ransomware wie Privatpersonen. Vor allem KMU erleiden mitunter schwere Schäden nach Cyberangriffen. Offizielle Zahlen, in wie vielen Fällen das Lösegeld tatsächlich floss, gibt es nicht. Die weiterhin hohe Anzahl an Angriffen ist aber ein Hinweis darauf, dass die Methode erfolgreich ist.

Bei der Wahl ihrer Opfer sind die Tätergruppen häufig zwar nicht wählerisch, gehen nach einem erfolgreichen Einbruch aber durchaus gezielt vor. Sie analysieren die Netzwerkinfrastruktur und übernehmen möglichst viele Systeme darin. Mögliche Backupsysteme werden identifiziert, um den Wiederherstellungsprozess zu erschweren. Erst wenn die Angreifer dann auch der sensiblen Daten habhaft sind, beginnt die Verschlüsselung und die Lösegeldforderung folgt auf dem Fuße. Während private Computer auf den ersten Blick nicht besonders lohnenswert erscheinen, können diese eine gewisse Brisanz entwickeln – selbst dann, wenn der Besitzer kein Lösegeld zahlt. Die Angreifer sammeln nämlich auch Passwörter aus den üblichen Passwortsafes, aus dem Browser oder dem E-Mail-Programm. Diese sind zwar mitunter durch ein Masterpasswort gesichert, allerdings lässt sich die Eingabe dieses Masterpasswords mitlesen. Oft sind darin auch Zugänge zu weiteren Computersystemen erhalten. So finden Kriminelle auch immer wieder Logindaten zu Fernwartungszugängen. Mit diesen gelangen die Täter dann in Unternehmensnetze, womit sie ein weiteres Opfer für ihre Schadsoftware gefunden haben.

Technische und menschliche Schutzmaßnahmen

Um sich erfolgreich vor Ransomware und den möglichen Folgen einer Infektion zu schützen, sind unterschiedliche Maßnahmen erforderlich. Technische Lücken, wie bei WannaCry und EternalBlue, lassen sich durch regelmäßige Updates aller Systeme im Unternehmen schließen. Dabei müssen Sie sich im Hinblick auf automatisierte Updates für jedes System die Frage stellen, ob ein potenzieller Ausfall durch einen Fehler beim Update schwerer wiegt als eine existierende Sicherheitslücke. Sie sollten also wo möglich automatische Updates aktivieren, auch auf die Gefahr eines Ausfalls durch eine fehlgeschlagene Aktualisierung hin. Je länger die Prüfung und Freigabe eines Updates dauert, desto mehr Zeit haben Angreifer, um auf die Unternehmenssysteme zuzugreifen.

Nutzen Sie zentrale Antivirus-Programme und Application-Layer-Gateways, um Anhänge in E-Mails zu untersuchen, bevor diese an Mitarbeiterkonten zugestellt werden. Zumindest die von Baukästen erzeugten Varianten der Schadsoftware lassen sich damit häufig ausfindig machen, wenngleich Sie gegen die individuellen Varianten der größeren Gruppen häufig nicht funktionieren.

Um das Nachladen von Schadsoftware nach einer erfolgten Infektion zu unterbinden, können Sie Ihre Mitarbeiter über Webproxys umleiten, binäre Downloads verbieten beziehungsweise separat freischalten und alle übrigen Anfragen am Proxy vorbei zunächst im Paketfilter blockieren. Anbieter von Sicherheitsprodukten bieten hierfür Listen von IP-Adressen und Domains an, die Sie filtern sollten. Diese Maßnahmen sind oft nicht ohne Einschränkung der Mitarbeiter umsetzbar und führen mitunter zu Konflikten bei der alltäglichen Arbeit. Daher verzichten viele Unternehmen darauf. Allerdings sollten Sie versuchen, gemeinsam mit Ihren Kollegen auszuprobieren, welche Maßnahmen Sie umsetzen können. Berücksichtigen Sie dabei auch Zeiten, in denen üblicherweise niemand arbeitet. Wenn Ihr Büro geschlossen ist, können Sie deutlich strengere Regeln umsetzen und überwachen, und diese zu den normalen Öffnungszeiten wieder lockern.

Auch wenn die Benutzer Ihres Unternehmens immer wieder als eigentliche Schwachstelle dargestellt werden, sind diese doch vielmehr die letzte Verteidigungslinie, die im Gegensatz zu den industriellen Sicherheitsprodukten wie Antivirus-Programmen, Application-Layer-Gateways und speziellen Firewalls die tatsächliche Infektion noch verhindern können.

Angreifer nutzen Spearphishing-Maßnahmen und wie im Fall von Emotet existierende Kommunikation, um Benutzer zum Ausführen der Schadsoftware zu bringen. Diese perfiden Techniken zu durchschauen, ist selbst für gut ausgebildete und trainierte Mitarbeiter schwierig. Die Ausbildung Ihrer Benutzer sollte also ein integraler Bestandteil der gesamten IT-Sicherheitsstrategie sein. Gezielte Schulungen sensibilisieren Ihre Benutzer, sodass diese nicht unfreiwillig zu Helfern der Angreifer werden. Dazu gehören auch eine aktive Fehlerkultur und die Möglichkeit für Ihre Mitarbeiter, auffällige Aktivitäten zu melden.

Zusätzlich können Sie Ihre Mitarbeiter technisch unterstützen. Etablieren Sie die Signatur von E-Mails in Ihrem Unternehmen. Das erschwert zumindest ein Stück weit die Erstellung glaubwürdiger E-Mails. Wenn alle E-Mails in Ihrem Unternehmen signiert sind, fallen die auf, die es nicht sind. Je weniger Ausnahmen es gibt, umso zuverlässiger können Ihre Mitarbeiter gefälschte E-Mails erkennen. Sollte dennoch ein Benutzer einen schadhaften Anhang öffnen, schützen Sie das System durch aktive Gruppenrichtlinien, die die Ausführung von Makros in diesen Dateien verbieten. Benötigen Benutzer Makros für ihre tägliche Arbeit, sollten Sie zumindest signierte Makros verwenden und die Ausführung unsignierter Makros unterbinden.

Cyberversicherungen

Wenn Sie sich gegen die Folgen von Cyberangriffen schützen möchten, können Sie zum klassischen Mittel einer Versicherung greifen und eine sogenannte Cyberversicherung abschließen. Je nachdem, in welchen Branchen Sie aktiv sind, können Cyberversicherungen eine günstige Ergänzung in Ihrem Risiko- und Notfallplan sein. Versicherungen können Ihnen auch bei ersten Schritten nach einem erfolgten Angriff helfen und erprobte Partner bei der Vorfallsanalyse vermitteln. Ob beziehungsweise in welchem Umfang die Folgen des Angriffs getragen werden und ob vielleicht sogar ein Lösegeld von der Versicherung übernommen wird, hängt von Ihren individuellen Verträgen ab. Einige große Versicherungen haben im letzten Jahr jedoch angekündigt, kein Lösegeld mehr an Kriminelle zu zahlen.

Zugriffe einschränken

Gültige Logindaten in den Händen Krimineller sind neben den technischen Schwachstellen ein großes Problem für die Sicherheit Ihrer Computer und Dienste. Im Darknet erhalten Hacker umfangreiche Datensammlungen von Identitätsdaten und Logindaten. Da Benutzer dazu neigen, dieselben Passwörter für unterschiedliche Dienste zu verwenden, sind Sie möglicherweise auch gefährdet, wenn andere Dienste Opfer von Hackerangriffen werden. Tatsächlich benutzen mehr als zwei Drittel der Benutzer bereits geleakte Logindaten noch länger als ein Jahr weiter. Das amerikanische NIST, aber auch der IT-Grundschutz weisen auf diese Gefahren hin und empfehlen die regelmäßige Überprüfung der Benutzerkonten und Passwörter.

Es gibt unterschiedliche Dienstleister, die sogenannte Identity Leak Checker anbieten. Der kostenfreie amerikanische Dienst "Have I Been Pwned" (HIBP) [2] ist der wohl bekannteste Anbieter von Leak-Informationen. Durch die Eingabe einer E-Mail-Adresse erhalten Sie Informationen darüber, ob diese Teil eines Datenlecks war. Auch wenn es sich um die betriebliche E-Mail-Adresse handelt, ist die Nutzung von HIBP aus Datenschutzgründen fraglich und sollte mit dem Personalrat oder Betriebsrat abgesprochen werden. Zudem erhalten Sie bei HIBP keinen direkten Zugriff auf das betroffene Passwort zu diesem Account, sodass Sie es nicht direkt gegen Ihre Systeme prüfen können. Dafür gibt es aber spezialisierte Dienstleister am Markt, die auch DSGVO-konforme Prüfungen von Log-indaten umsetzen.

Bei der Vergabe von Benutzerrechten sollten Sie die Möglichkeit gestohlener Log-indaten berücksichtigen. Geben Sie Mitarbeitern nur so viele Zugriffsrechte, wie diese unbedingt für ihren normalen Berufsalltag benötigen. Das gilt vor allem beim Zugriff auf Server und gemeinsame Dateien. Räumen Sie etwa für existierende Dateien auf einem Server für Benutzer ausschließlich Leserechte ein, müssen die User zwar für jede Änderung eine neue Datei hochladen, die Dateien können dafür von diesem Benutzer dann nicht gelöscht oder verschlüsselt werden.

Das Prinzip der geringsten Berechtigungen hat vor allem zur Folge, dass Sie Prozesse etablieren müssen, die regelmäßig existierende Berechtigungen überprüfen. Das kommt insbesondere dann zum Tragen, wenn Mitarbeiter die Abteilungen wechseln oder für Projekte abteilungsübergreifend zusammenarbeiten. Das Phänomen ist etwa bei Praktikanten üblich, die im Laufe ihres Praktikums unterschiedliche Abteilungen durchlaufen. Einmal erteilte Privilegien werden häufig nicht wieder entzogen, neue aber regelmäßig hinzugefügt. Zum Abschluss eines Praktikums hat der Praktikant dann Zugriff auf ein Benutzerkonto mit vielen sicherheitsrelevanten Zugriffsmöglichkeiten.

Ausbreitung verhindern

Wenn Angreifer trotz aller Schutzmaßnahmen doch einmal Zugriff auf Computer in Ihrem Unternehmensnetzwerk haben, muss das nicht zwangsweise bedeuten, dass diese am Ende auch erfolgreich mit ihrem Angriff sind. Sie sollten versuchen, den Schaden in solchen Fällen möglichst klein zu halten. Um eine Ausbreitung zu verhindern, trennen Sie unterschiedliche Abteilungen und unterschiedliche Teams derselben Abteilung bestenfalls netzwerktechnisch voneinander und bringen diese in eigenen Subnetzen unter. Zwischen diesen sollten Sie eine Firewall haben, die übergreifenden Netzwerkverkehr reguliert und auf das Nötigste beschränkt.

Je schneller Sie reagieren, desto größer ist Ihre Chance, einen großen Schaden abzuwenden. Dafür etablieren Sie ein umfangreiches Monitoring Ihrer Ressourcen, um betroffene Systeme rasch zu erkennen und zu isolieren. Eventuell sollten Sie das

gesamte Team beziehungsweise die gesamte Abteilung gemeinsam abschotten. So bleiben die Arbeitsfähigkeit und der Schutz der anderen Einheiten zunächst erhalten. Diesen Vorgang müssen Sie natürlich auch regelmäßig durchspielen. Oft sind dafür nur wenige Firewallregeln nötig. Je nach Aufbau Ihrer Infrastruktur können Sie betroffene Computer auch automatisch in einem separaten VLAN isolieren. Ein Angreifer hat dann zwar noch Zugang zu einem System, kann von dort aber keine weiteren Rechner infizieren. Loggen Sie interne Netzwerkverbindungen auf den Routern zwischen Ihren Abteilungen, können Sie sogar im Nachgang feststellen, ob bereits eine Ausbreitung – das sogenannte Lateral Movement – stattgefunden hat.

Selbst wenn der Angreifer über eine Schwachstelle in Ihr Unternehmensnetzwerk eingedrungen ist, heißt das nicht, dass er dieselbe Schwachstelle auch auf anderen Systemen vorfinden muss. Angreifer benutzen daher unterschiedliche Tools für ihre Bewegungen. Tatsächlich wird dabei auch das von Microsoft mitgelieferte Remote-Desktop-Protokoll verwendet. Insbesondere in Zeiten von Home Office und VPN-Verbindungen von zu Hause in das Unternehmensnetz hinein erfreuen sich Remotedesktops großer Beliebtheit. Die Zugriffsmöglichkeit ist in den meisten Fällen über das Active Directory schnell eingeräumt. Überwachen Sie auch hier an zentralen Stellen die Verbindungen, die mit dem Verzeichnisdienst aufgebaut werden. Reagieren Sie möglichst automatisiert auf unvorhergesehene Verbindungsversuche.

Backups schützen

Das Anlegen von Backups gehört zu den Standardaufgaben eines Administrators. Sie sollten aber nicht nur das Anlegen übernehmen und überwachen, sondern auch die Absicherung und den Zugriff auf existierende Backups. Bestenfalls gibt es ohne weiteres keinen Zugriff auf das Backupsystem. Vielmehr sollte das Backupsystem Zugriff auf die einzelnen Dienste haben, die es sichern soll. Können die Benutzer Ihrer Systeme selbst nicht auf das Backup zugreifen, kann dieses auch nicht von einer Ransomware verschlüsselt werden, die unter einem normalen Benutzerkonto gestartet wurde.

Um dennoch eine einfache Wiederherstellung von Dateien für den normalen Benutzungsfall zu erlauben, haben Sie bestenfalls unterschiedliche Backupsysteme etabliert. Eines, das Ihre Benutzer selbst verwalten können, und eines, auf das nur im äußersten Notfall und nur von wenigen Administratoren zugegriffen werden kann. Das schützt Sie zwar nicht vor Lösegeldforderungen, um kopierte Geschäftsgeheimnisse zu schützen. Sie können aber nach einem Ransomware-Vorfall schnell wieder den Betrieb aufnehmen.

Abläufe regeln

Ist das Kind in den Brunnen gefallen und Ihre Systeme sind unkontrollierbar von einem Ransomwarevorfall betroffen, müssen Sie adäquat reagieren. Bestenfalls haben Sie dafür im Vorfeld Risiko- und Notfallpläne erarbeitet und Verantwortlichkeiten festgelegt. Die Pläne beinhalten Informationen über die Kritikalität einzelner Systeme und legen auch fest, wie weit Sie andere Systeme abschalten müssen. Ein gezielter und geplanter Shutdown kann Ihr Unternehmen vor existentiellen Schäden schützen, selbst wenn dabei andere Schäden in Kauf genommen werden müssen. Die Betreiber der Colonial-Pipeline haben hier vorbildlich reagiert und das System gezielt vom Netz genommen.

Informieren Sie zeitnah die Ansprechpartner in den jeweiligen Abteilungen und bringen Sie Backupsysteme entsprechend der festgelegten Kritikalität in Betrieb. Wenn Sie rechtlich zur Meldung von Cybervorfällen verpflichtet sind, sollten Sie entsprechende Formulare vorgefertigt haben und zeitnah eine Meldung machen. So verhindern Sie spätere Strafen aufgrund von Unterlassung. Die betroffenen Systeme halten Sie für weitere forensische Analysen vor. Diese können Sie selbst durchführen, wenn Ihr Unternehmen groß genug ist und Sie entsprechende Fähigkeiten in Ihrer IT-Abteilung haben. Ansonsten beauftragen Sie einen externen Dienstleister damit. Das Ziel der Analyse sollte vor allem die Identifikation der Schwachstelle sein – Ihre Daten haben Sie hoffentlich aus dem Backup wieder einspielen können. Bauen Sie nach und nach Ihre Infrastruktur wieder auf, sobald Sie die Schwachstelle beseitigt haben. Vergessen Sie dabei nicht, neue Backupsysteme für einen erneuten Angriff einzurichten.

Für jede Abteilung haben Sie im Optimalfall noch einen internen Notfallplan. Dieser beinhaltet dann auch die Information von Zulieferern, Partnern oder Endkunden im jeweiligen Bereich. Wenn Sie selbst Zulieferer sind, müssen Sie abhängige Unternehmen in der Lieferkette zeitnah informieren. Auch Ihre eigenen Zulieferer sollten Sie in Zeiten von lagerlosen Lieferketten und Just-in-Time-Produktionen informieren.

Wenn Sie von dem Angriff überrascht wurden, während Sie noch in den Ausarbeitungen zu Ihren Risiko- und Notfallplänen sind, versuchen Sie wenigstens zu retten, was noch zu retten ist. Das kann auch beinhalten, über die Zahlung eines Lösegelds nachzudenken. Das sollten Sie jedoch gemeinsam mit den Behörden absprechen, die Sie nach dem Vorfall informiert haben. Richten Sie mit allen Personen, die Sie zeitnah als relevant identifizieren können, einen Krisenstab ein und besprechen dort die notwendigen Maßnahmen.

Fazit

Ransomware ist die große Bedrohung für Unternehmen, öffentliche Einrichtungen und Privatpersonen. In den letzten Jahren haben die Folgen von Ransomware-Angriffen immer größere Ausmaße angenommen. Der Artikel zeigte Ihnen unterschiedliche Angriffsvektoren und Ausprägungen von Ransomware anhand von tatsächlichen Vorfällen. Wir haben das existierende Risiko diskutiert und wie Notfallpläne Ihnen bei der Reaktion auf Angriffe helfen können, den Betrieb wiederherzustellen. *(dr)*

Link-Codes

[1] Ransomware-Angriff auf Colonial-Pipeline
m3z91

[2] Have I Been Pwned
h0z52

Redundanz für das Rechenzentrum

In Sicherheit

von Evgenij Smirnov

Quelle: pboonrit – 123RF

Auf den ersten Blick erscheint das Konzept schlüssig: Die Verfügbarkeit des eigenen Rechenzentrums soll durch das Verteilen der Infrastruktur auf zwei Serverräume erhöht werden. Doch nicht jedes Design geht im Fehlerfall gleichermaßen gut auf und bringt das gewünschte Maß an Ausfallsicherheit. Lesen Sie in diesem Beitrag, worauf Sie beim Design Ihres räumlich aufgeteilten Rechenzentrums achten müssen.

Wer heutzutage ein Rechenzentrum plant, und sei es nur ein Serverraum in einem kleinen Unternehmen, versucht meistens, die dort untergebrachte Technik auf mehr als einen Raum zu verteilen. Vorzugsweise befinden sich diese Räumlichkeiten dann in verschiedenen Brandabschnitten oder sogar Gebäuden zum Beispiel auf einem Werksgelände. Die Idee dahinter ist, bei einem geplanten oder ungeplanten Ausfall eines Rechenzentrum-Raums die dort bereitgestellten IT-Dienste von dem anderen Raum aus aufrechterhalten zu können. Virtualisierungs- und Storage-Plattformen, Datenbanken und Anwendungen bieten dafür äußerst umfangreiche Hilfsmittel.

Doch stellen IT-Verantwortliche oft fest, dass es mit zwei RZ-Räumen, redundanter Stromversorgung, Klima und Vernetzung allein nicht getan ist. Damit sich die Schwachstellen der geplanten RZ-Redundanz nicht erst im Failover-Fall offenbaren, ist ein gutes Verständnis der eingesetzten Technologien und vor allem eine klare Definition der angestrebten Redundanz vonnöten. Die wichtigsten Merkmale, die Sie dabei berücksichtigen müssen, sind (in dieser Reihenfolge):

- Single Points of Failure,
- Failover-Verhalten der einzelnen Systeme und Dienste,
- Performance-Einschränkungen durch Wegfall von Systemen.

Das Ziel klar definieren

Bevor Sie weitreichende und im Nachhinein schwer zu korrigierende Entscheidungen für Ihr RZ-Design treffen, müssen Sie festlegen, welcher Ausfall mit welchen Einschränkungen toleriert werden soll. Die oft formulierte Idealvorstellung lautet: "Wenn eines der Gebäude komplett abbrennt, muss für den Rest der Firma alles weiterlaufen, höchstens ein wenig langsamer". Mit einer solchen Leistungszusage ist es in der Regel auch am einfachsten, ein Budget für Hochverfügbarkeit genehmigt zu bekommen. Doch was bedeutet eine derart allgemein gefasste Aussage in der praktischen Durchführung?

In der modernen IT spielt die Anbindung an das Internet und damit an die Außenwelt eine große Rolle. Für viele Dienste und Anwendungen ist der Internetzugang oft wichtiger als die Anbindung an die internen Netze und Systeme. Die Vorgabe, einen RZ-Raum ohne funktionale Einschränkungen abschalten oder verlieren zu können, bedeutet daher, dass die Internetzugänge inklusive Zuführung seitens der Provider ebenfalls auf beide Räume verteilt werden müssen. Das bringt einige Herausforderungen mit sich. Eine entsprechende Konfiguration der externen Router, die ebenfalls auf beide Räume zu verteilen sind und über entsprechende Failover-Mechanismen verfügen müssen, ist in der Regel herstellerseitig vorgesehen und schnell erledigt.

Die Leitungszuführung hingegen kann problematisch werden, insbesondere wenn die beiden RZ-Standorte in einem Gebäude dicht beieinander liegen und nur durch Brandabschnitte getrennt sind. Da kann es schon passieren, dass die Leitungsführung zum Raum A durch den Brandabschnitt mit dem Raum B verläuft und bei Problemen in diesem Bereich in Mitleidenschaft gezogen wird. Bei Serverräumen, die an unterschiedlichen Seiten eines weitläufigen Fabrikgeländes platziert sind, müssen Leitungsanbieter manchmal überzeugt werden, einen Leitungsstrang rund ums Gelände zu führen.

Viel spannender wird die Herausforderung, den Internet-Uplink auf zwei Räume zu verteilen, wenn aus dem Rechenzentrum heraus Dienste ins Internet veröffentlicht werden. Es muss also auch bei Ausfall eines der Räume sichergestellt sein, dass externe Kommunikationspartner den entsprechenden Endpunkt finden und erreichen können. Dazu gibt es zwei technische Lösungsansätze:

- Sie lassen die externen IP-Adressen Ihrer DMZ stets in einem der Räume terminieren, beim Ausfall des dortigen Routers übernimmt der Router des verbleibenden Raumes und teilt mittels Border Gateway Protocol (BGP) dem Rest der Welt mit, dass er von nun an für diese Subnetze verantwortlich ist. Diese Technik muss Ihr ISP explizit unterstützen, Sie müssen das Grenznetz

also zwingend mit ihm zusammen designen und aufbauen. Wie gravierend sich Fehler im Umgang mit BGP auswirken können, zeigt der jüngste Ausfall sämtlicher Dienste des Facebook-Konzerns [1].
- Sie richten in jedem Raum eine eigene DMZ ein und routen Ihre extern veröffentlichten Dienste innerhalb Ihres eigenen Netzwerks. Um sicherzustellen, dass externe Kommunikationspartner die angebotenen Dienste auch bei Ausfall eines der Räume finden, müssen Sie entweder zu GeoDNS (zum Beispiel ClouDNS mit der "Failover and Monitoring"-Option [2]) oder zum Geo-Loadbalancing greifen. Ein Beispiel hierfür ist "Elastic Load Balancing" von AWS [3]. In beiden Fällen nutzen Sie einen Dienst aus der Public Cloud, womit Sie genau genommen nicht mehr zwei, sondern drei "Serverräume" haben.

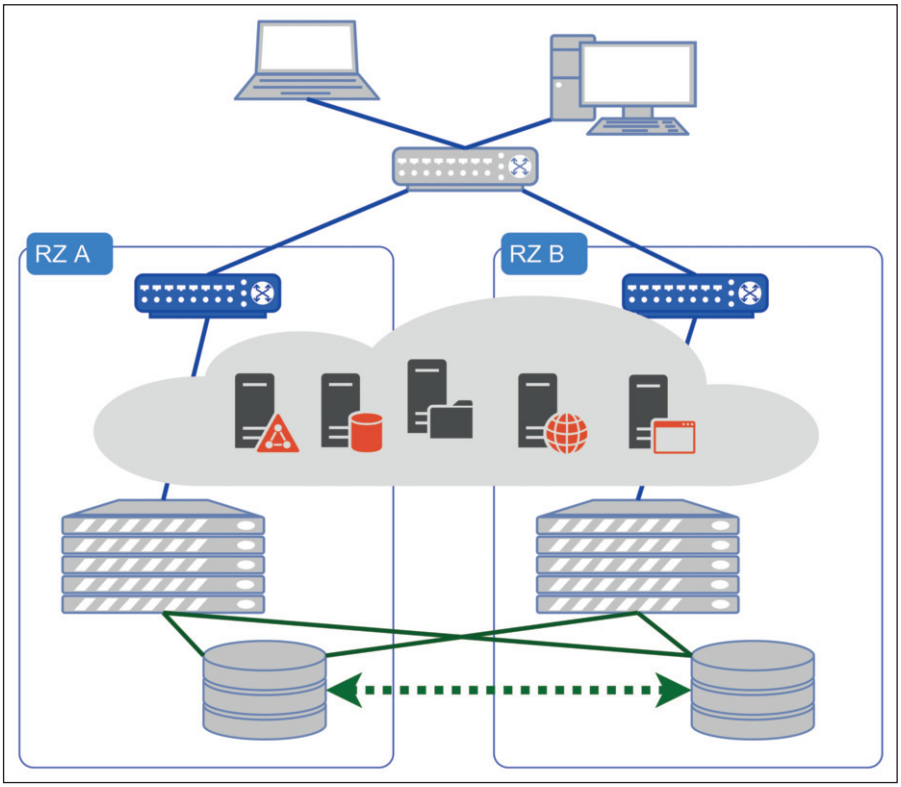

Bild 1: Die scheinbare Einfachheit der Plattform-HA birgt viele Herausforderungen.

Doch auch für Ihre internen Applikationen und Dienste müssen Sie sich auf eine jeweils passende Hochverfügbarkeits-Methodik festlegen, um keine Überraschungen im Falle eines Ausfalls oder einer geplanten Wartung zu erleben. Dabei ist es wichtig, dass Sie die folgenden Szenarien in Bezug auf Ihre konkrete Infrastruktur genau betrachten und gegeneinander gewichten (nach Wahrscheinlichkeit des Eintretens und Dauer der Wiederherstellung):
- Unkontrollierter Ausfall des RZ A oder RZ B
- Kontrollierte Abschaltung des RZ A oder RZ B
- Verlust der externen Konnektivität im RZ A oder RZ B
- Verlust der Verbindung zwischen RZ A und RZ B

Besonders das letzte Ausfallszenario birgt sowohl im Design als auch später im Betrieb viele Tücken. Es wird bei der Planung eines hochverfügbaren internen RZ jedoch oft übersehen oder nicht ausreichend beachtet.

Schließlich müssen Sie möglichst frühzeitig ein in der Verfügbarkeitsplanung häufiges Missverständnis auflösen – spätestens dann, wenn Sie Ihre RZ-Planung dem Management vorstellen. Oft ist in technischen Designs und Konzepten von "Hochverfügbarkeit" (High Availability oder HA) die Rede. Dies bedeutet jedoch nicht mehr, als dass bei Ausfall eines Systemteils das Gesamtsystem automatisch und innerhalb einer definierten Zeit die Bereitstellung des jeweiligen Diensts wiederaufnimmt.

HA sorgt für keine Zusicherung des unterbrechungsfreien Betriebs oder eine 100-prozentige Datenkonsistenz nach Wiederherstellung. Genau dies ist jedoch die Erwartungshaltung der Nichttechniker (und leider auch vieler Techniker) in Bezug auf den Begriff "HA". Die korrekte Bezeichnung für diese deutlich höhere Servicequalität ist "Fehlertoleranz" (Fault Tolerance oder FT). Fehlertolerante Designs sind in vielen Fällen möglich, jedoch meist mit deutlich höheren Kosten verbunden als die "einfache" HA.

Die beiden Designextreme

Bei der Projektierung eines hochverfügbaren Rechenzentrums mit zwei oder mehreren Abschnitten stehen Ihnen zahlreiche Designmöglichkeiten zur Verfügung. Sie alle haben gemeinsam, dass jeder der Räume eine aktuelle Kopie aller Anwendungsdaten vorhalten muss, Sie zu jeder Zeit eine Netzwerkverbindung zu den Anwendungs- und Infrastrukturservern ermöglichen müssen und außerdem eine Datendivergenz zwischen den Räumen ("split brain") in allen vorgesehenen Szenarien zu vermeiden ist. Die beiden Designansätze, mit denen Sie Ihre Planung beginnen müssen, sind "Plattform-Hochverfügbarkeit" und "Anwendungs-Hochverfügbarkeit".

Die Plattform-Hochverfügbarkeit, die erst mit dem Fortschreiten der Servervirtualisierung möglich geworden ist, abstrahiert die Compute-, Storage- und Netzwerkleistung von ihrem jeweiligen Standort. Alle Datenspeicher, in denen sich die virtuellen Festplatten befinden, sind synchron gespiegelt – entweder mit der klassischen SAN-Spiegelung oder mit "Software-defined-Storage"-Technologien wie VMware vSAN oder Microsoft Storage Spaces Direct. Sowohl Daten- als auch Storage-Netzwerke sind zwischen den RZ-Räumen voll vermascht, sodass die eingesetzten Switching-Protokolle schnell auf Veränderungen der Netzwerktopologie reagieren und die Pakete in den "richtigen" Raum zustellen können. Sämtliche Anwendungen und Infrastrukturdienste, auf die Clients zugreifen sollen, werden auf virtuellen Maschinen oder in

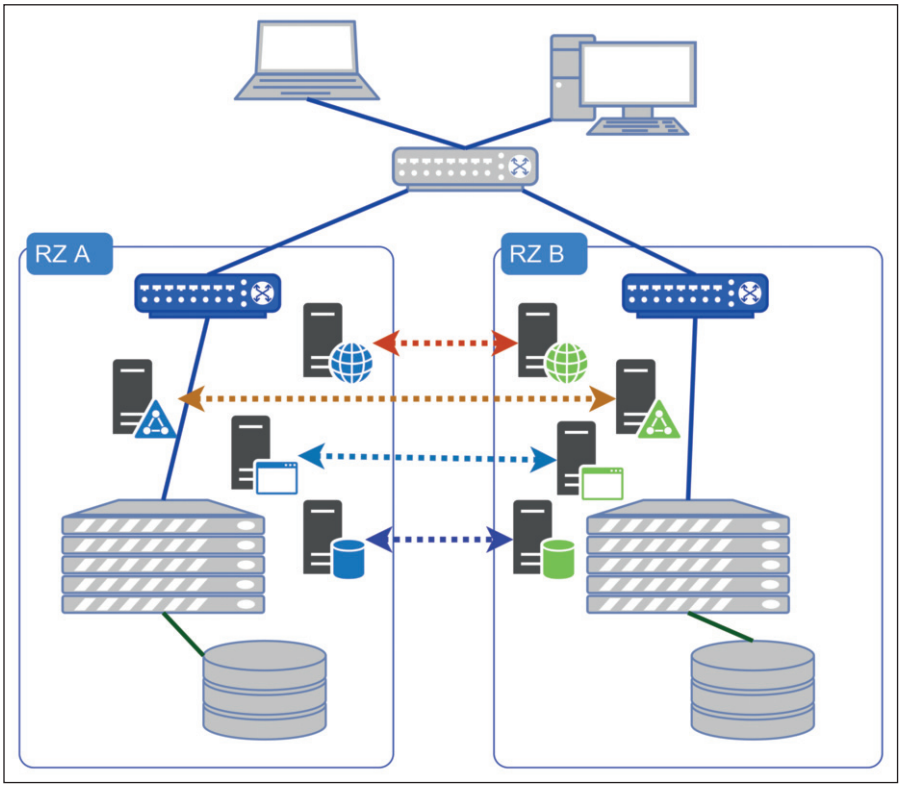

Bild 2: Bei der Anwendungs-HA muss jeder Dienst separat berücksichtigt werden.

Containern ausgeführt. Der Hochverfügbarkeitsansatz besteht darin, dass jede virtuelle Maschine stets den aktuellen Datenunterbau hat und netzwerktechnisch unter ihrer IP-Adresse erreichbar ist – egal in welchem RZ-Abschnitt sie gerade ausgeführt wird.

Anwendungs-Hochverfügbarkeit verfolgt den gegenteiligen Ansatz: Storage-Bestände und IP-Adressen, somit auch einzelne VMs, sind dauerhaft an einen der Räume gebunden, und jeder Dienst beziehungsweise jede Anwendung nutzt die dort eingebauten Replikations- und Failover-Mechanismen. Active Directory, DNS, DHCP, doch auch Exchange, SQL, ADFS oder andere Webanwendungen verfügen über zuverlässige und gut skalierbare HA-Fähigkeiten.

Zu diesem Ansatz gibt es keine wirkliche Alternative, wenn Sie gezwungen sind, eine Anwendung auf physischen anstatt virtuellen Servern bereitzustellen. Doch auch im virtuellen Umfeld ist die anwendungseigene Hochverfügbarkeit oft die einzige, die vom jeweiligen Hersteller unterstützt wird – so ist dies beispielsweise bei Microsoft Exchange oder Skype for Business der Fall.

Beide Architekturen haben ihre Stärken und Schwächen. Sie können diese Stärken ausnutzen, wenn Sie die beiden Designansätze geschickt miteinander kombinieren. Dafür ist es wichtig, die Auswirkungen der verschiedenen HA-Architekturen genau zu kennen. Denken Sie daran, dass keines der HA-Konzepte per se den Anspruch der Integrität und Vollständigkeit der Daten erhebt und alle Aussagen sich lediglich auf die Verfügbarkeit der Dienste beziehen.

Failover vs. Switchover

Bei der Planung hochverfügbarer Infrastrukturen wird oft der Extremfall ausführlich diskutiert: Ein ganzer RZ-Standort geht unerwartet und unkontrolliert vom Netz. Das kann durch ein katastrophales Ereignis oder menschliches Versagen erfolgen, aber auch durch einen Cyberangriff etwa auf die Stromversorgung. In diesem Fall ist das erwartete Verhalten Ihrer Server- und Netzwerklandschaft ein sogenanntes "Failover". Ein viel häufiger auftretendes Ereignis ist jedoch ein "Switchover", das heißt, ein kontrolliertes Abschalten eines RZ-Abschnitts für Baumaßnahmen oder im Rahmen behördlich vorgeschriebener Untersuchungen, falls Ihr Hochverfügbarkeitskonzept einer Auditpflicht unterliegt.

Nicht jede HA-Architektur ist für Failover und Switchover gleichermaßen geeignet. Bei der Anwendungs-HA werden Sie als Administrator versucht sein, für die meisten Anwendungen den Switchover-Vorgang administrativ "anzukündigen". Das beinhaltet das Verschieben aktiver Cluster-Instanzen, das Versetzen von einzelnen Servern in einen Wartungsmodus oder deren kontrolliertes Herunterfahren.

Obwohl dieses behutsame Vorgehen eine bessere Verfügbarkeit aller Anwendungen während und nach dem Switchover verspricht, ist diese Herangehensweise aufwendig und bei vielen Anwendungen nur denkbar, wenn Sie sie möglichst vollständig automatisieren. Im Fall der Plattform-HA müssen Sie "nur" alle virtuellen Maschinen aus dem abzuschaltenden RZ-Abschnitt verschieben und die dortigen Virtualisierungshosts in den Wartungsmodus versetzen, damit die VMs nicht automatisch zurückmigriert werden.

Bei einem ungeplanten Failover fallen alle Dienste, die im havarierten RZ-Abschnitt gelaufen sind, erst einmal aus. Im Fall der Plattform-HA werden sie durch die Hochverfügbarkeitsfunktion der Virtualisierung neu gestartet. Je nach Ressourcen kann es einige Minuten dauern, bis alle VMs wieder laufen. Die Betriebssystem- und Datenfestplatten sind in diesem Fall "doppelt-Crash-konsistent", zum einen durch das abrupte Ausschalten der VM selbst und zum anderen durch den Abbruch der Replikation der Storage-Blöcke zwischen den RZ-Abschnitten. Moderne Betriebssysteme haben inzwischen eine sehr hohe Resilienz gegenüber solchen Vorgängen entwickelt.

Datenbanken und Anwendungen können hingegen einen Schaden davontragen, wenn Sie für ihre interne Datenhaltung keine transaktionelle Integrität erzwingen. Die Hochverfügbarkeit auf Anwendungsebene leistet hier, sofern sie sauber konfiguriert ist, deutlich bessere Dienste. Für viele Anwendungen können Sie hier mit einer ausfallfreien Fortsetzung des Betriebs rechnen. Ist neben Hochverfügbarkeit auch Lastverteilung Teil Ihrer Bereitstellungsstrategie, ist vom Ausfall eines RZ-Abschnitts ohnehin nur die Hälfte der Verbindungen betroffen.

12 Monatsausgaben im Print- & E-Paper-Format

+

Zwei Sonderhefte
im Print- & E-Paper-Format

+

Jahres Archiv-CD
mit allen Monatsausgaben im PDF-Format

Das Abo All-Inclusive
shop.heinemann-verlag.de

Abo- und Leserservice IT-Administrator

vertriebsunion meynen
Herr Stephan Orgel
D-65341 Eltville
Tel: 06123/9238-251
Fax: 06123/9238-252
leserservice@it-administrator.de

IT Administrator

Physische Standorte planen

Die Anwendungs-Hochverfügbarkeit beinhaltet meist Konzepte, die die Bindung einer Serverinstanz an einen bestimmten RZ-Raum durch standortspezifische IP-Subnetze und davon abgeleitete Active-Directory-Standorte realisieren. So findet beispielsweise die AD-Replikation innerhalb eines Standorts stets sofort nach einer Änderung statt, zwischen den Standorten unterliegt sie einem Turnus und berücksichtigt die Qualität der Standortverbindungen. Weitere Beispiele für standortabhängiges Verhalten sind DFS (Priorisierung von Ordnerzielen) oder Exchange (Autodiscover, Mailrouting, Safety Net, Datacenter Activation Coordination).

Doch auch Loadbalancer können davon profitieren, zwischen unterschiedlichen Subnetzen für ihre "realen Server" unterscheiden zu können. Bei einer sauber konfigurierten Anwendungs-HA können die Applikationen also erkennen, dass nicht einfach nur eine beliebige Hälfte der Anwendungsserver vom Netz gegangen ist, sondern ein ganzer Standort. Auch die Authentifizierungsanforderungen an das Active Directory gehen nicht ins Leere, wenn ein anderer Standort als der eigene ausfällt, was für den flüssigeren Betrieb der AD-abhängigen Anwendungen sorgt.

Bei der Plattform-HA gibt es aus Sicht der virtualisierten Anwendungsserver nur einen Standort. In der Konfiguration der Virtualisierung und des Storage müssen die einzelnen Knoten zwar bestimmten Standorten zugewiesen werden, um die Datenreplikation und das Failover-Verhalten entsprechend der physischen Servertopologie zu lenken.

Dies geschieht meist durch Definition sogenannter Fault Domains (siehe [4] für VMware und [5] für Microsoft), die Zuweisung der Knoten zu Fault Domains ist jedoch den Administratoren vorbehalten. Falls Sie innerhalb einer hochverfügbaren Plattform Anwendungen unter Benutzung ihrer eigenen HA-Mechanismen bereitstellen, müssen Sie Vorkehrungen treffen, damit die anwendungseigene HA adäquat auf den Ausfall eines RZ-Standorts reagiert.

Der dritte Zwilling

Die meisten HA-Konzepte, ganz gleich auf welcher Ebene sie wirken, setzen auf eine Clustering-Logik, um auf Ausfälle einzelner Server zu reagieren. Die Algorithmen, nach denen jeder Server bestimmt, ob er seinen Dienst weiter verrichten oder einstellen soll, unterscheiden sich von Produkt zu Produkt. Alle Dienste, die zur Ausfallerkennung auf Microsoft-Cluster (Windows Server Failover Clustering, WSFC) setzen, bedienen sich eines auf Stimmenmehrheit basierenden Entscheidungsverfahrens.

Dieses ist in den modernen Windows-Server-Versionen mehrfach verfeinert worden, sodass Sie bei einem kontrollierten Herunterfahren alle Clusterknoten bis auf einen (und den Quorum-Zeugen) abschalten können, ohne dass das Cluster seinen Dienst einstellt. VMware hat eine andere Logik implementiert, nach der jeder Host versucht zu ermitteln, ob er andere Hosts, die Datenspeicher und bestimmte Punkte im Netzwerk erreicht. Mit vSphere 7.0U1 ist dieser Mechanismus um Cluster-Services-VMs erweitert worden, sodass jeder Cluster auch dann seine Vorgänge koordinieren kann, wenn das vCenter gerade ausgefallen oder nicht erreichbar ist.

Jeder dieser Mechanismen stellt Sie als RZ-Planer vor eigene Herausforderungen. Die auf Stimmenmehrheiten basierende Microsoft-Logik funktioniert sehr gut bei einem Ausfall einzelner Knoten. Geht jedoch ein ganzer RZ-Abschnitt gleichzeitig vom Netz, fehlt bei einer symmetrischen Verteilung der Knoten die Hälfte der Stimmen und das Überleben des Clusters hängt davon ab, ob der Quorumzeuge noch erreichbar ist.

Diese Logik hat VMware auch für vSAN Stretched Cluster übernommen – inklusive der dynamischen Stimmenzählung (geplant ab 7.0U3), sodass bei einem konsekutiven Ausfall eines RZ-Standortes und des Quorums der verbleibende Standort weiterhin den vSAN-Datastore zur Verfügung stellt. Die ausschließlich auf Erreichbarkeit beruhende Logik eines vSphere-Clusters scheitert dann, wenn die Netzwerkadressen, welche als "Beacon" eingetragen sind, vom ausgefallenen Rechenzentrum in Mitleidenschaft gezogen wurden.

Die größten Entscheidungsprobleme haben Cluster jeder Art jedoch zu bewältigen, wenn nicht einer der RZ-Standorte ausfällt, sondern lediglich die Verbindung zwischen den Standorten. Hier muss die Failover-Logik wirkungsvoll verhindern, dass ein Split Brain-Szenario entsteht, bei dem jeder Knoten der Meinung ist, dass er der legitime Überlebende eines Ausfalls ist und die Dienste weiter ausführt. In der Regel wird hier als Lösungsansatz ein dritter Standort empfohlen, der unabhängig vom übrigen Netzwerk an die produktiven RZ-Standorte angebunden ist und lediglich den Quorum-Zeugen beinhaltet.

Bei einer symmetrischen Knotenverteilung stellt Sie dieser Ansatz zunächst einmal vor ein Dilemma: Ist der Anschluss an das Quorum bei der Trennung der Querverbindung intakt, sehen beide Seiten das Quorum und sprechen sich damit die Stimmenmehrheit zu, sodass es zum Split-Brain-Phänomen kommt. Geht die Quorum-Verbindung mit der Heartbeat-Verbindung des Clusters zusammen offline, hat keine der Seiten eine Stimmenmehrheit und der gesamte Cluster geht offline.

Ein Ansatz, der hier mehr Erfolg verspricht, besteht darin, das Backend-Netzwerk des Clusters (Heartbeat und gegebenenfalls Datenreplikation) mit dessen Frontend- und Quorum-Netzwerk zusammenzulegen. Die Separation und Priorisierung des Traffics kann in diesem Fall durch VLANs und QoS erfolgen, aber der Ausfall einer Netzwerkverbindung wirkt sich auf Client- und Quorumzugriff gleichermaßen aus, sodass weder Split-Brain noch eine unerwartete Gesamtabschaltung des Clusters auftreten können.

Doch auch der strikt auf Erreichbarkeit basierende HA-Ansatz eines vSphere-Clusters bedarf beim Einsatz über RZ-Abschnitte hinweg einer sorgfältigen Behandlung. Hier wird über die Host-Isolation anhand der Zugriffe auf Datenspeicher entschieden. Bei einem

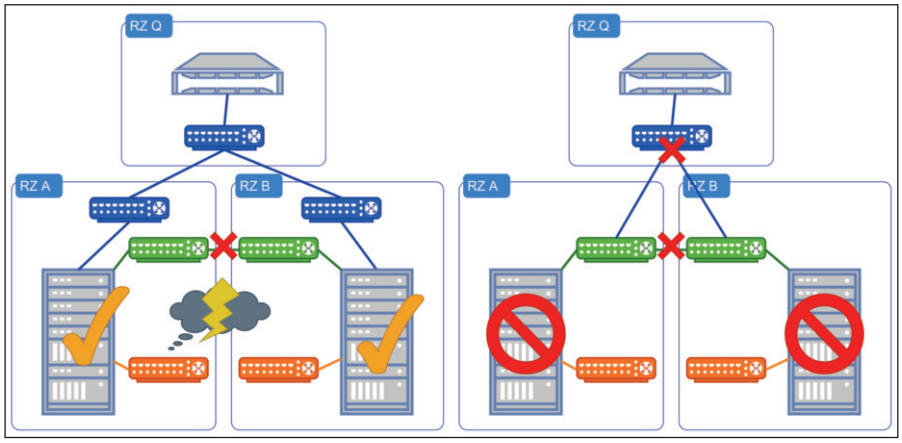

Bild 3: Eine ungünstige Platzierung der Quorum-Leitung führt zu anderem Failover-Verhalten als beabsichtigt.

klassischen, gespiegelten SAN kann es passieren, dass der Ausfall der Spiegelverbindung von den Hosts gar nicht erkannt wird. Folgt dann ein Ausfall der Querverbindung für das Managementnetzwerk, spielen die sogenannten Isolationsadressen die Rolle des Quorums – ist eine von ihnen erreichbar, gilt der Host als nicht isoliert und setzt seine Arbeit fort [6].

Auch hier gilt: Das Zusammenlegen (auf Hardware-Ebene) des Management- und des Frontend-Netzwerks kann für eine präzisere Erkennung von Ausfällen sorgen als die generelle Best-Practices-Empfehlung, diese Netzwerke voneinander zu trennen. Stellen Sie nach Möglichkeit pro RZ-Standort zwei Datenspeicher für Ihre vSphere-Cluster bereit, die nur in diesem Standort vorhanden sind und fixieren Sie den Hartbeat-Datastore auf diese. Dann wissen Ihre ESXi-Hosts genau, wenn sie isoliert sind, und können den Ausfall eines einzelnen Hosts dennoch korrekt behandeln.

Das Beste aus beiden Welten

Beim Design Ihres RZ sollten Sie sich nicht zu früh darauf versteifen, einen der oben beschriebenen HA-Konzepte für alle Dienste und Anwendungen durchzusetzen. Einige Dienste (Lizenzserver, beispielsweise für RDS, oder Datenbanken, für die HA-Funktionen nicht lizenziert wurden) verfügen gar nicht über native HA-Mechanismen und würden von der Plattform-Hochverfügbarkeit immens profitieren.

Andere Anwendungen müssen vielleicht aus Gründen, die Sie nicht beeinflussen können, auf physischen Servern bereitgestellt werden, sodass Sie hier auf die Anwendungs-HA angewiesen sind. Bei allen anderen müssen Sie eine Entscheidung treffen, wobei der Hersteller-Support und das in Ihrem IT-Team vorhandene Know-how eine Rolle spielen. Diese Entscheidungen lassen sich sehr gut an Microsoft-Produkten illustrieren, deren native Hochverfügbarkeit auf "Availability Groups" basiert: Exchange und SQL.

Wenn Sie mehr als nur eine Handvoll Exchange-Postfächer hosten, ist eine Database Availability Group (DAG) praktisch Pflicht. Damit verdoppelt sich der Speicherplatz, den Sie für die Exchange-Daten bereitstellen müssen, schon auf der Anwendungsebene mindestens. Sie werden kaum davon profitieren, dass Sie ihn noch einmal verdoppeln, indem Sie die Postfachserver auf gespiegelten Datenspeichern laufen lassen.

Es kann also sinnvoll sein, die Postfachserver fest in den RZ-Abschnitten anzusiedeln und die Hochverfügbarkeit ausschließlich auf Anwendungsebene zu betreiben. Ein Hybrid-Exchange-Server für Ihre Office-365-Bereitstellung hingegen benötigt keine derartige Sonderbehandlung und lässt sich durchaus "schwebend" innerhalb einer hochverfügbaren Virtualisierungsplattform betreiben.

Eine häufige Designfrage bei der Kombination beider HA-Ansätze innerhalb einer Umgebung betrifft Domaincontroller für das Active Directory. Diese nehmen in der Regel wenig Ressourcen in Anspruch, können aber nicht unbegrenzt lange ausfallen. Hier erreichen Sie eine optimale Bereitstellung, indem Sie bestimmte IP-Subnetze zwar organisatorisch einem RZ-Standort widmen, sie jedoch technisch über die gesamte Virtualisierungsplattform spannen.

So können Domaincontroller und Systeme, die die Standort-Topologie im AD ignorieren, bei Ausfall oder Wartung eines RZ-Standorts umziehen. Systeme, die physisch an einen RZ-Standort gebunden sind, werden ihre Domaincontroller finden, auch wenn sie gerade im anderen Serverraum laufen. Sie müssen sich bei geplanter Wartung oder längerem Ausfall eines Standortes zudem nicht damit befassen, ob der aktuelle PDC Emulator oder RID Master gerade online ist.

Fazit

Bei der Planung Ihres hochverfügbaren, auf mehrere Abschnitte verteilten Rechenzentrums müssen Sie eine Balance zwischen Ausfallsicherheit, Komplexität und Betriebsfähigkeit finden. Setzen Sie sich bei der Planung mit den Eigenarten der technischen HA-Ansätze auseinander und designen Sie Ihr Rechenzentrum so, dass Sie von den technologischen Vorteilen der Plattform- und Anwendungs-Hochverfügbarkeit größtmöglich profitieren. *(dr)* IT

Link-Codes

[1] **Pressemitteilung von Facebook**
m1z81

[2] **ClouDNS Failover and Monitoring**
m1z82

[3] **AWS Elastic Load Balancing**
m1z83

[4] **Fault Domains in VMware vSAN**
l6z44

[5] **Fault Domains im Windows Server Failover Clustering**
m1z85

[6] **vSphere Isolation Addresses**
m1z86

Automatisierte Datensicherung

Viel hilft nicht viel

von Martin Loschwitz

Ganze Systeme in ein Backup zu gießen, war in der Vergangenheit eine valide Strategie, heute wirkt der Ansatz jedoch antiquiert. Nicht nur belegen Vollbackups unnötigen Speicherplatz, auch der Restore-Prozess beansprucht viel Zeit. Eine gute Automatisierung ermöglicht kleinere Datensicherungen und kurze Wiederanlaufzeiten. Der Beitrag zeigt den Weg zu einer automatisierten Backupstrategie.

Quelle: Wichan Kijchanpaiboon – 123RF

Vor allem in produktiven Umgebungen kommt der Administrator nicht mehr ohne eine valide Backupstrategie aus. In Zeiten, in denen viele Firmen ohne digitale Dienste schlicht nicht überleben können, kann der Verlust wichtiger Daten dem Unternehmen schnell den Garaus machen.

Sonderlich gerne befassen sich Admins mit der Thematik allerdings bis heute nicht. Das macht sich vorrangig durch Backupkonzepte bemerkbar, die zwar auch heute vielerorts noch in Nutzung sind, jedoch eher den technischen Stand von vor zehn Jahren repräsentieren. Das ist schade: Im Jahre 2021 ist es nämlich längst nicht mehr nötig, komplette Systeme als wiederherstellbares Backup auf ein Bandlaufwerk zu schreiben, um vor Ausfällen geschützt zu sein. Richtig eingesetzt greift Automatisierung dem Admin an vielen Stellen unter die Arme, wo früher noch eine typische Backupsoftware zum Einsatz gekommen wäre.

Nur Datenbankinhalte sichern

Was damit gemeint ist, zeigt das Beispiel einer Datenbank deutlich. Fast jedes Setup der modernen IT beinhaltet eine Datenbank. Oft ist das MySQL oder dessen Konkurrent MariaDB. Und weil Daten das Blut in den Adern moderner Anwendungen sind, bilden Datenbanken stets einen neuralgischen Punkt in der Infrastruktur. Gehen Daten verloren, hat das meist katastrophale Auswirkungen. Bei einem Backup geht es allerdings selten um das komplette Datenbanksystem: Meistens zählt nur der Inhalt selbst. Oftmals sind das nur wenige GByte. Dementgegen steht die Backuprealität in vielen Unternehmen: Für Datenbanksysteme werden in vielen Fällen Vollbackups erstellt. Das System lässt sich aus solchen zwar komplett wiederherstellen, doch fressen sie viel Platz auf den Sicherungsmedien. Und das ist nicht der einzige technische Nachteil.

Ist ein Rollback auf einen früheren Datenstand nötig, würde es viel zu lange dauern, die Datenbank selbst aus dem Vollbackup zu extrahieren. Zu allem Überfluss sichern Unternehmen heute deshalb häufig zusätzlich zum ganzen Datenbanksystem auch noch die Inhalte der Datenbank separat. Was umso schlimmer ist eingedenk der Tatsache, dass sich praktisch der gesamte Datenbankserver – mit Ausnahme der tatsächlichen Nutzdaten – ohnehin automatisiert wiederherstellen lässt. Denn praktisch ist ein Datenbankserver heute ja nicht viel mehr als eine Linux-Installation mit Datenbankpaketen.

Das beschriebene Beispiel legt den Wahnsinn vieler Backupstrategien offen. Unternehmen sichern wie wild ganze Systeme, obwohl sie diese in der Praxis nur sehr selten wiederherstellen müssen. Die eigentlich regelmäßig benötigten Backups kommen separat hinzu. In Summe verstopft der Sammelwahn die Backup-Medien und vergeudet so unnötig Platz. Die Lösung für das Problem liegt indes auf der Hand: Wer seine Automatisierung im Griff hat, kann sich das Sichern ganzer Systeme sparen – und gewinnt sogar beim Recovery der Datenbankdaten selbst viel Zeit. Die folgenden Schritte zeigen, wie eine sinnvolle Backup-Stragie auf Basis von Automation aussehen kann.

Workload erfassen

Wer sich schon mal mit den Themen Automation und Backups beschäftigt hat, dem ist vielleicht auch der Begriff "Immutable Environment" begegnet. Die Idee dahinter ist, Backups und Automatisierung so zu kombinieren, dass sich jedweder Workload schnell und ohne viel Bastelei sichern und wiederherstellen lässt. Der Weg zu einem funktionalen Immutable-Environment-Design führt stets über eine Bestandsaufnahme im ersten Schritt. Das heißt konkret: Bevor Sie automatisieren und Backups zusammen-

streichen, erfassen Sie die Stellen, an denen Automation sinnvoll ist, und jene, wo Nutzdaten lokal zu sichern sind.

Die zu automatisierenden Systeme unterscheiden sich dabei je nach Umgebung erheblich. Haben Sie das Glück, eine eher homogene Serverumgebung zu betreuen, ist diese Aufgabe nicht sehr komplex. Sehen Sie sich einer heterogenen Umgebung aus Clients und Servern gegenüber, wird das Gespann aus Automatisierung und Backups jedenfalls komplexer in der Implementation – aber nicht unmöglich. Wir beziehen uns in diesem Beitrag bevorzugt auf Linux-Server. Am Ende folgt ein kurzer Exkurs für Clientsysteme, der auch das Thema Windows einschließt.

Ganz gleich, ob Workloads im produktiven Einsatz in Container verpackt sind oder direkt auf dem Blech laufen: Eine basale Linux-Distribution ist stets eine Grundvoraussetzung für die Nutzung eines Servers. Das gilt auch für virtuelle Instanzen, die ebenfalls Linux als Grundgerüst benötigen. Setzt eine Umgebung verstärkt auf Container, ist das Grundsystem vermutlich schlank, weil eine Container-Distribution zum Einsatz kommt. Und selbst wenn Sie hier auf Systeme wie Red Hat Enterprise Linux oder SUSE Linux Enterprise Server setzen, wird ein nur für den Betrieb von Containern vorbereiteter Host kein umfassend großes System werden.

Gerade weil ein funktionierendes Grundsystem die oberste Prämisse für die Nutzung jedes Servers ist, liegt bei der Automation-Strategie ein Hauptaugenmerk auf dem Prozess, der aus einem normalen Server von der Stange ein System der eigenen Umgebung macht. Folgerichtig ist der erste Schritt, die automatische Installation von Linux zu ermöglichen.

Konfiguration automatisieren

Was viele IT-Verantwortliche leider aus den Augen verlieren, ist, dass auch die Konfiguration der jeweiligen Systeme einen wichtigen Faktor darstellt. Ein installiertes, sehr basales RHEL oder SLES bietet Vorteile, ist aber ohne Anpassung an die Voraussetzungen vor Ort eher nutzlos. Klar: Über die automatische Installation per Kickstart, AutoYaST oder Preseeding lassen sich viele Parameter des Systems bestimmen. Ein Ersatz für echte Automatisierer wie Ansible ist das aber nicht.

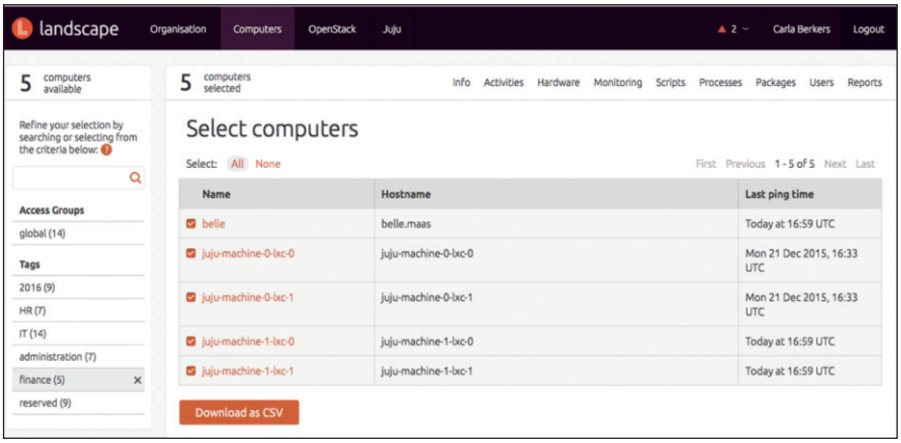

Bild 1: Beim Lifecycle-Management hat der Admin die Wahl zwischen Tools der großen Hersteller wie Landscape von Ubuntu oder einem beliebigen anderen Werkzeug.

Mancher Admin bearbeitet seine Systeme etwa noch während des Deployment-Vorgangs mit Skripten, die der jeweilige Installer während oder unmittelbar nach der Installation ausführt. Doch diese Mechanismen sind eigentlich nicht für die umfassende Systemkonfiguration gedacht, und mit einem echten Automatisierer wären die meisten Admis viel besser bedient. Denn nur so entsteht ein durchgehendes Lifecycle-Management: Im ersten Schritt installiert der Admin den Server per Netzwerkinstallation neu. Dann folgt die basale Konfiguration per Automatisierer. Abschließend wird der Datensatz der tatsächlichen Nutzdaten eingespielt.

Backup einspielen

Wer seine Systeme automatisiert unter Kontrolle hat, muss im letzten Arbeitsschritt Datensätze für spezifische Zielprogramme aus dem Backup zurückspielen. Diese Aufgabe ist kaum sinnvoll automatisierbar – es liegt auch in der Natur der Sache, dass das Einspielen von Backups üblicherweise eine Aufgabe ist, die der Admin sorgfältig von Hand durchführen will. Umso wichtiger ist es, die automatisierbaren Schritte der Arbeit tatsächlich zu automatisieren. Ihr Weg dorthin ist steinig oder leicht, je nach vorhandenem Setup.

Automation heißt auch Lifecycle-Management

Viele Admins haben das Thema Automatisierung auf die lange Bank geschoben. Das betrifft vielerorts vor allem die Ebene des Betriebssystems. Dabei beginnt hier die Erfolgsgeschichte der Automation einer Umgebung. Denn es ergibt auch in kleineren Setups aus Sicht des Admins kaum Sinn, sich mit der Installation eines Betriebssystems regelmäßig herumzuschlagen. Zumal diese Aufgabe wirklich mittlerweile gut automatisierbar ist. Die Protokolle für alle Teile des Setups – PXE, DHCP, TFTP & Co. – stehen seit Jahrzehnten bereit und funktionieren hervorragend.

Aus heutiger Sicht sieht das Soll-Szenario ungefähr so aus: Bekommt der Admin einen Server geliefert, packt er diesen aus, montiert den Server im Rack und verkabelt ihn passend. Idealerweise ist der Server ab Werk darauf konfiguriert, per PXE einen Netzwerkstart durchzuführen. Die meisten Anbieter konfigurieren ihre Systeme entweder gleich so vor oder bieten zumindest die Möglichkeit, den gewünschten Startmodus bei der Bestellung anzugeben. Per Netz startet der Server dann in ein Inventarisierungssystem, wo die vor Ort eingesetzte Lifecycle-Management-Software erstmals Notiz von ihm nimmt.

Danach weist der Admin dem System eine Rolle zu, aus der sich im Idealfall die Konfiguration des Servers automatisch ergibt. Der Rest ist dann reine Lappalie: Der Server bootet in den Installationsmechanismus für das Betriebssystem, bekommt dabei auch gleich seine Netzwerkkonfiguration und wird danach per Automatisierer auf den richtigen Stand gebracht. Wenn alle Räder dieses Mechanismus richtig ineinandergegriffen haben, läuft auf dem Zielsys-

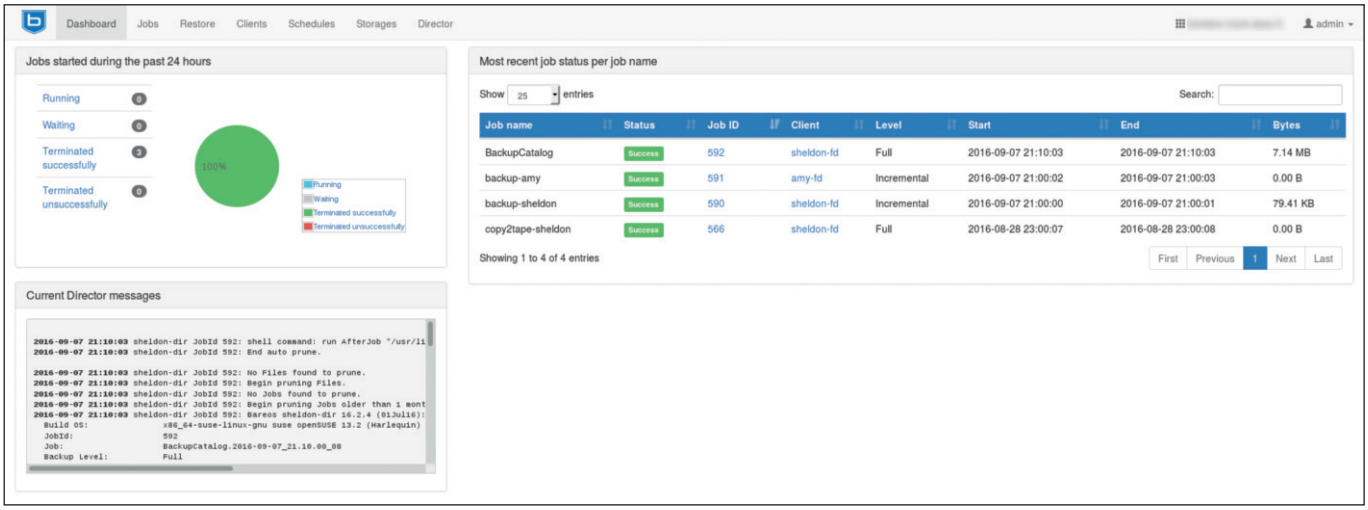

Bild 2: Datenbankbackups sind ein gutes Beispiel dafür, dass heutzutage oft nur der Inhalt einer Anwendung wiederhergestellt wird.

tem jetzt auch schon die benötigte Software – etwa MariaDB, falls es sich um einen Datenbankserver handelt.

Ihr primäres Ziel haben Sie als Administrator damit bereits erreicht – Sie sind in der Lage, in kürzester Zeit eine neue Datenbankinstanz aufs Blech oder virtuell aus dem Ärmel zu zaubern. Davon profitieren Sie, wenn Sie die bestehenden Datenbanken der Installation erweitern möchten, aber eben auch, wenn ein ausgefallenes System durch ein neues zu ersetzen ist. Das Backup des gesamten Datenbanksystems können Sie an dieser Stelle daher schon abschalten. Denn mit dem beschriebenen Automatismus stellen Sie ein System her, das einem zerstörten Vorgänger gleicht wie ein Ei dem anderen.

In aller Regel ist der Vorgang der automatisierten Neuinstallation zudem viel schneller als die komplette Wiederherstellung aus dem Backup. Denn das Lifecycle-Management, das auch Details zur Konfiguration der Systeme enthält, läuft im Normalfall permanent – es genügt für einen Server also, in die PXE-basierte Netzwerkinstallation zu booten, um sie zu verwenden.

Tools in der Praxis

Welches der auf dem Markt angebotenen Lifecycle-Management-Systeme sich für Ihr spezifisches Setup am besten eignet, hängt von verschiedenen Faktoren ab. Zwei Ansätze stehen sich diametral entgegen: Die eine Option besteht darin, die Werkzeuge des Herstellers zu verwenden, der auch ansonsten das Setup dominiert. Das funktioniert in homogenen Umgebungen sehr gut. Wer etwa ausschließlich auf Red Hat Enterprise Linux oder CentOS setzt, wird mit Red Hat Satellite [1] zufrieden sein. Wer ein typischer Ubuntu-Shop ist, kommt mit Landscape [2] (Bild 1) oder MAAS [3] gut zurecht, weil Ubuntu ideal an diese angepasst ist. Und für SUSE-Enthusiasten steht der SUSE Manager [4] zur Verfügung.

Haben Sie es mit einer heterogenen Umgebung zu tun, zahlt es sich vielleicht aus, das Lifecycle-System selbst zu bauen und nicht auf ein fertiges Produkt eines Anbieters zu setzen. Keine Sorge, auch dabei müssten Sie nicht bei Null anfangen: Foreman [5] etwa gilt als äußerst erprobtes, versatiles System für das Lifecycle-Management. Zusätzlich haben viele Automatisierer mittlerweile auch Lifecycle-Manager im Gepäck. Erheben Sie wie beschrieben im ersten Schritt Ihren Workload und lassen Sie sich in der weiteren Folge am besten die Angebote der verschiedenen Anbieter genau vorstellen – bedenken Sie dabei allerdings, dass das Lifecycle-Management Sie im Normalfall eine ganze Weile begleiten wird. Es schadet deshalb nicht, sich hier am Anfang eher kompromisslos zu geben.

Restore mithilfe von Snapshots

Bis hierhin haben wir uns vorrangig mit der Notwendigkeit beschäftigt, Betriebssysteme nicht aus dem Backup zu ziehen, sondern automatisiert wiederherzustellen. Automation und Orchestrierung greifen dem Admin allerdings auch beim Restore von Datenbanken und anderen Diensten unter die Arme. Beispielsweise wenn es um die Rettung von Daten geht, die ein Nutzer versehentlich gelöscht hat oder die anderweitig korrumpiert wurden. Mit dem richtigen Konzept und der passenden Backupsoftware lässt sich nämlich auch der Wiederherstellungsprozess durch Automation und richtiges Anwendungsdesign deutlich schneller gestalten als in konventionellen Umgebungen.

Gegeben sei also erneut ein Server mit einer MariaDB-Datenbank. Ob das System auf Blech läuft oder virtualisiert, ist zunächst zweitrangig. Wichtig ist aber, dass der Admin ein sinnvolles Konzept für Backups des Datensatzes seiner Datenbank hat. Hier ergeben sich regelmäßig zwei Optionen: Die Sicherung des Datenträgers, auf dem die Datenbank beheimatet ist, oder das Erstellen eines Backups über die Datenbank selbst.

Beide Ansätze haben Vor- und Nachteile, mit denen der Administrator vertraut sein sollte. Und beide Methoden stehen nicht in jedem Szenario zur Verfügung. Backups über die Datenbank selbst bieten den Vorteil, mit hoher Wahrscheinlichkeit keine korrupten Backupdateien zu produzieren, weil die Datenbank aktiv am Prozess beteiligt ist – sie weiß also, dass sie gerade gesichert wird. Und das ist wichtig: Laufende Transaktionen kann die Datenbank beenden und das Starten neuer Transaktionen verzögern. Der Nachteil: Damit das funktioniert, muss die Datenbank selbst laufen.

Stürzt der Datenbankprozess reproduzierbar ab, wird es auf diesem Weg nicht mehr möglich sein, an Backups zu kommen oder einen Restore durchzuführen.

Die Sicherung des Datenträgers mit den Datenbankdaten hat hingegen den Nachteil, dass ein solcher Prozess stets dafür sorgen muss, dass ein Write in die Datenbank nicht genau im Moment des Snapshot stattfindet. Ansonsten ist der Snapshot potenziell korrupt und dadurch natürlich unbrauchbar. Agenten, die die Datenbank während des Ziehens des Snapshot kurz anhalten, sollen bei vielen Backuplösungen dieses Problem umgehen (Bild 2).

Backupsoftware ist umgebungsbedingt

Wie bereits erwähnt, gibt es für das Sichern von Workloads verschiedene Optionen. Welche Software und welchen Weg Sie für Ihr eigenes Setup nutzen, hängt stark von den jeweiligen Begebenheiten ab. Am Markt existieren durchaus potente Programme und All-in-one-Lösungen für das Sichern von Datenbanken wie MariaDB oder PostgreSQL.

Andererseits findet sich im Werkzeugkasten der Open-Source-Gemeinde für nahezu jede Anwendung auch Software, die spezifisch auf den Einsatzzweck hin optimiert ist. Von einer Empfehlung sehen wir an dieser Stelle daher ab – wir gehen jedoch im weiteren Verlauf davon aus, dass Sie eine intakte Sicherung der Daten besitzen, die Sie für den Restore-Prozess benötigen.

Rollback in Windeseile

In einem Szenario, in dem Sie die Datenbank zurücksetzen müssen, ist ein entscheidendes Kriterium der Arbeitsaufwand und die Downtime der Umgebung. Denn der Rollback soll in aller Regel so schnell wie möglich geschehen. Wohl dem Admin, der seine Automation im Griff hat: In kürzester Zeit installiert dieser nämlich komplett automatisch eine zweite, parallel zur ersten Datenbank laufende Instanz. In diese spielt er anschließend den letzten funktionierenden Datensatz ein und schwenkt dann die IP-Adresse von der Instanz mit der alten Datenbank auf die Instanz mit der neuen – und fertig ist die Laube.

In virtualisierten Umgebungen spielt dieser Ansatz seine Vorteile freilich deutlicher aus als auf echtem Blech. Denn hier lässt sich sogar ad hoc ein neues Volume mit den Inhalten des Snapshot anlegen, das der Admin der neuen Datenbankinstanz von Anfang an mit auf den Weg gibt – oder im Rahmen einer kurzen Downtime der alten Datenbankinstanz unterjubelt. Aber auch auf physischen Servern ist der gesamte Vorgang mit deutlicher Zeitersparnis verbunden.

Lifecycle-Management für Windows-Clients

Der Beitrag befasst sich in erster Linie mit Linux-Servern im RZ – doch weiß jeder Administrator, dass die Wiederherstellung eines Windows-Systems aus einem Backup ebenfalls ein lästiger Vorgang ist. Die Idee, herkömmliche Backups durch Automation zu ersetzen, greift indes auch bei klassischen Clients nicht vollständig ins Leere. Hier kommt es freilich darauf an, wie das jeweilige System verwaltet wird.

Gerade in Firmenumgebungen finden sich Netzwerklaufwerke als Speicherort für Nutzerdaten. Diese sind mit dem Benutzer zusammen mobil. Meldet sich also der User an einem anderen Gerät als üblich an, findet er dort wie gewohnt sein Netzlaufwerk und seine Daten vor. Für den Client selbst heißt das, dass er austauschbar ist – und genau hier liegt das mögliche Fundament für eine umfassende Automationsstrategie.

Desktopbetriebssysteme wie Windows oder Linux lassen sich nämlich ebenso gut in ein Lifecycle-Management einbinden wie die Geräte aus dem Serverraum. Lediglich macOS ist hier eine Ausnahme, denn eine "unattended" Installation bietet Apple nicht an. Andere Anbieter einschlägiger Hardware stellen hier gar eigene Tools zur Verfügung, die aber freilich nur den Bestand der eigenen Marke abdecken.

Wer eine heterogene Clientlandschaft etwa aus HP-, Dell- und Microsoft-Geräten hat, wird mit diesen spezifischen Lösungen also nicht glücklich. Auch ohne kann es aber funktionieren. Windows etwa bietet das Preinstallation Environment (PE), aus dem heraus Windows sich automatisiert installieren lässt. Und der Start des WinPE aus einer PXE-Umgebung heraus ist seitens Microsoft ausdrücklich unterstützt. Ferner gibt es auch für Windows natürlich Automatisierer – zum Teil funktionieren gar die, die Admins schon von Linux her kennen. Wer also einmal die Zeit investiert, Lifecycle-Management für Windows-Clients zu bauen, automatisiert deren Installation bei Bedarf ebenso und schafft in Kombination mit Netzwerklaufwerken eine Art Immutable Environment, wie es auch im Serverraum möglich ist. Gerade für IT-Abteilungen, die regelmäßig große Mengen an Systemen neu ausrollen, kann das eine erhebliche Zeitersparnis sein.

Fazit

Wer schlau automatisiert, verkürzt vor allem die Wiederanlaufzeit nach Ausfällen oder dem versehentlichen Löschen von Daten dramatisch. So ungern sich mancher Admin auch mit dem Backupthema befassen mag: Vollbackups kompletter Systeme sind im Jahr 2021 weder ökonomisch sinnvoll noch in irgendeiner Art und Weise State of the Art. Es ist schlicht nicht sinnvoll, sich seine Tapes mit Inhalten zuzukleistern, die bei Bedarf jederzeit wieder aus dem Internet zu beziehen sind.

Damit Automation und Restore Hand in Hand gehen, hat der IT-Verantwortliche allerdings ein paar Hausaufgaben zu erledigen. Doch keine Sorge: Diese lohnen sich nicht nur im Backupkontext, sondern erlauben dem Admin gleichzeitig sinnvolles Lifecycle-Management seiner Instanzen. *(jm)* IT

Link-Codes

[1] Red Hat Satellite
l4z21

[2] Landscape
f3p81

[3] MAAS
l4z23

[4] SUSE Manager
l4z24

[5] Lifecycle-Management mit Foreman
l4z25

Hardwarekonfiguration für vSphere

Hart am Blech

von Evgenij Smirnov

Obwohl VMware in Sachen Virtualisierung meist nur noch vom Software-defined Datacenter spricht, liegt stets Hardware mit zahlreichen Konfigurationsmöglichkeiten zugrunde. Unser Workshop zeigt, welche BIOS-Einstellungen für neue und bestehende Server wichtig sind, wenn sie als ESXi-Hosts zum Einsatz kommen sollen. Dabei kümmern wir uns um die optimale Leistung, den Energieverbrauch und die Sicherheit.

Fügen Sie einen vSphere-Cluster zur bestehenden Virtualisierungsfarm hinzu, bauen ein ganz neues virtuelles Rechenzentrum auf VMware-Basis auf oder planen ein größeres Update Ihrer bestehenden vSphere-Umgebung, lohnt es sich, den physischen Unterbau der Virtualisierung genau anzuschauen. Zu den wichtigsten Punkten gehören dabei die Konfigurationseinstellungen der Hardware, oft salopp als "BIOS-Einstellungen" bezeichnet. Dabei booten heutzutage die meisten Server nicht mehr im BIOS-, sondern im UEFI-Modus.

Es gibt jedoch generell eine Fülle von Einstellungen, mit denen Sie das Verhalten der Hardware (CPU, RAM, Storage, Netzwerk) beeinflussen und die sich nicht ohne einen Neustart verändern lassen. Die verfügbaren Settings unterscheiden sich sehr stark je nach Serverhersteller, Modellreihe und Generation. Die möglichen Stellschrauben genau unter die Lupe zu nehmen, lohnt sich also selbst dann, wenn Sie bereits Erfahrung mit dem Servertyp in einer früheren Generation haben.

In größeren Umgebungen sind Server oft in ein zentrales Hardwaremanagement eingebunden (beispielsweise OneView bei HPE oder OpenManage bei DELL). Diese Werkzeuge ermöglichen eine zentrale Konfiguration der Hardware. So können IT-Verantwortliche neu hinzukommende Server mit dem gleichen Satz an Hardwarekonfigurationen versehen wie die Bestandsserver, ohne dass sie bei jeder Maschine einzeln Hand anlegen müssen. In einer solchen Umgebung ist es kontraproduktiv, Einstellungen an einzelnen Servern per Hand vorzunehmen, denn dadurch würde eine Diskrepanz zwischen zentral gespeicherten Profilen und tatsächlichen Konfigurationen entstehen.

Grundlegende BIOS-Einstellungen anpassen

Als erste Maßnahme sollten Sie stets dafür sorgen, dass das BIOS und auch die sonstige Firmware Ihrer Hosts auf dem aktuellen Stand sind. Dabei lohnt sich ein Blick in die VMware-Compatibility-Datenbank, denn die Support-Aussage von VMware für jeden Hardwaretyp bezieht sich auf bestimmte Firmware-Versionen.

Viele Servermodelle verfügen über "Workload-Profile", also eine Möglichkeit, schnell verschiedene Systemparameter auf Werte einzustellen, die einem bestimmten Einsatzzweck des Servers (Rechenleistung, Antwortzeit, Energieeffizienz et cetera) entsprechen. Gehört Ihr Servertyp dazu, sollten Sie das Workload-Profil, das Ihren Wünschen am ehesten entspricht, zuerst wählen und alle weiteren Einstellungen ausgehend von diesem Profil vornehmen.

Wir betrachten nun die wichtigsten Einstellungen, die sich auf das Verhalten Ihres Hardwareservers in seiner Rolle als ESXi-Host auswirken. Als Erstes sind die "Virtualization Features" zu nennen, die aktiviert sein müssen, damit der Typ 1-Hypervisor ESXi seinen Dienst verrichten kann. Auf Intel-Systemen ist das Feature "Virtualization Technology" (VT) für die Implementierung der "Secure Virtual Machine" zuständig, die das Ausführen von 64-Bit-Gastbetriebssystemen erlaubt. Auf den aktuellen Chipset- und CPU-Generationen heißt die grundlegende Funktion "VT-x", wobei das x für "x86" steht. Weitere im Artikel [1] beschriebene Virtualisierungs-

funktionen sind VT-D (exklusives Durchreichen von PCI-Geräten an virtuelle Maschinen) und VT-C, meist als SR-IOV in den BIOS-Menüs abgebildet (Virtualisierung von Gerätefunktionen). Eine ähnliche Auswahl von Virtualisierungsfunktionen, die ebenfalls über das BIOS steuerbar sind, bietet auch AMD [2] – interessanterweise ist VMware dort nicht als "Virtualisierungspartner" aufgeführt.

Der nächste wichtige Wert ist "Boot Time Optimizations": Obwohl ein Host mitunter Wochen oder sogar Monate ohne Neustart laufen kann, haben Admins ein großes Interesse daran, dass ein Reboot möglichst schnell abläuft. In diesem BIOS-Abschnitt haben Sie die Möglichkeit, einige Prüfungen und Optimierungen während der Bootsequenz abzuschalten. Allerdings kann sich ein zu rigoroses Beschleunigen des Bootvorgangs negativ auf die Stabilität des Servers auswirken, zum Beispiel wenn wichtige Hardwarechecks entfallen. Mit dem Feature "Boot Order" legen Sie – je nachdem, ob Sie einen lokalen Datenträger, eine LUN auf einem SAN als Bootdisk oder PXE für AutoDeploy zum Starten Ihres ESXi verwenden – das gewählte Bootmedium auf Platz 1 in der Bootreihenfolge.

Unter "Processor and Memory Options" nehmen Sie normalerweise für einen klassischen Virtualisierungshost keine besonderen Einstellungen vor. Doch für Workloads, die eine vollständige Konsistenz des Arbeitsspeichers erfordern, ist hier beispielsweise die Spiegelung der RAM-Bausteine einstellbar. Eine Funktion, von der Sie im ESXi-Betrieb definitiv die Finger lassen sollten, ist das bedarfsgesteuerte temporäre Abschalten nicht verwendeter CPU-Kerne, das einige Chipsätze anbieten. Und auch die permanente Deaktivierung von CPU-Kernen sollte nur in absoluten Ausnahmefällen zum Einsatz kommen. Meist sorgen Lizenzgründe dafür, dass Admins die Anzahl der Kerne pro Sockel künstlich reduzieren und hier Hand anlegen.

Komplexes Energiemanagement

Die aus ESXi-Sicht wichtigste Einstellung, die Sie im BIOS vornehmen können und in der Regel auch müssen, ist

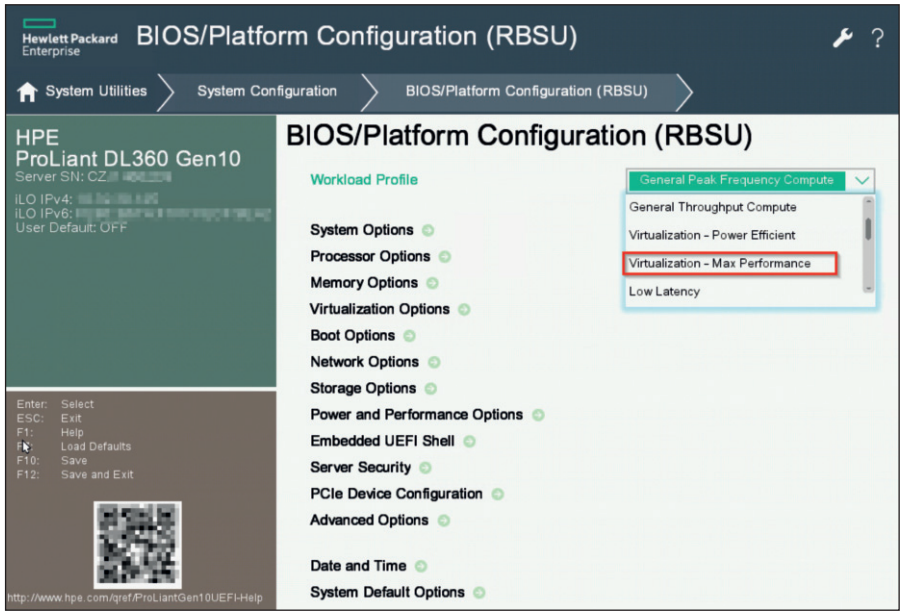

Bild 1: Die Nutzung des Workload-Profils legt bereits viele BIOS-Einstellungen fest.

das Energieprofil des Servers und vor allem der CPU. Die Zeiten des Turbo-Mode-Schalters am Computer sind schon lange vorbei. Prozessoren haben heute die Fähigkeit, ihre Taktfrequenz abhängig von der Rechenlast anzupassen und sich sogar selbst in den Schlaf zu versetzen, falls das ausgeführte Betriebssystem länger inaktiv ist. Und die Motherboards verfügen über Schnittstellen, die es dem Betriebssystem erlauben, die Energiefunktionen nach eigenen Algorithmen zu steuern.

Im Sinne der Green IT und der Klimaneutralität der Rechenzentren erscheint es IT-Verantwortlichen oft attraktiv, diese Energiesparfunktionen auch im Server- und speziell im Virtualisierungsbereich einzusetzen. Schließlich verbraucht eine aktuelle Intel-Xeon-Gold-CPU je nach Modell zwischen 140 und 235 Watt, was im Durchschnitt für einen Zwei-Wege-Server 9 kWh pro Tag und über 3 MWh pro Jahr allein für die CPUs ergibt.

Beim Betrieb virtualisierter Rechenzentren bleibt all das leider eine schöne Theorie: Das "ESXi Performance Best Practices"-Whitepaper [3] empfiehlt, die Steuerung der Energieoptionen dem ESXi-Hypervisor zu übertragen, und verweist weiter auf den "ESXi Resource Management Guide", um die richtige Richtlinie für Ihren speziellen Einsatz zu wählen. Diese Konstellation erwartet auch die "VMware Health Check Analyzer"-Appliance, falls Sie Ihre vSphere-Infrastruktur einem Hersteller-Healthcheck unterziehen.

Die Erfahrung aus dem Betrieb und dem Performance-Troubleshooting von virtualisierten Rechenzentren spricht jedoch eine andere Sprache. In Umgebungen mit hoher Dichte und Auslastung von virtuellen Maschinen muss das Power-Profil

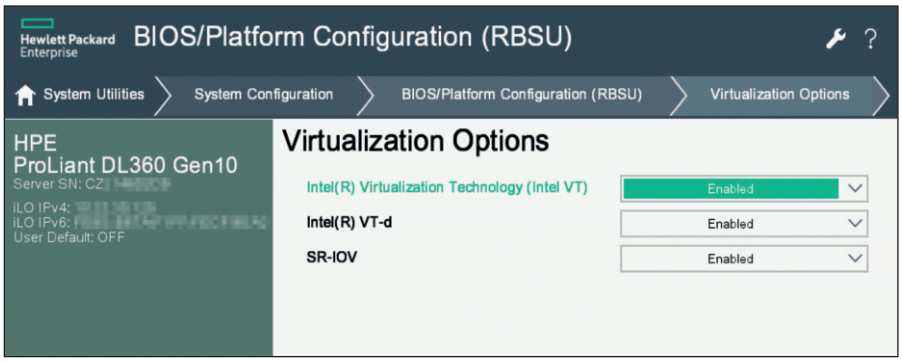

Bild 2: Das Workload-Profil aktiviert alle Virtualisierungsoptionen, was die Konfiguration jeder einzelnen Maschine unnötig macht.

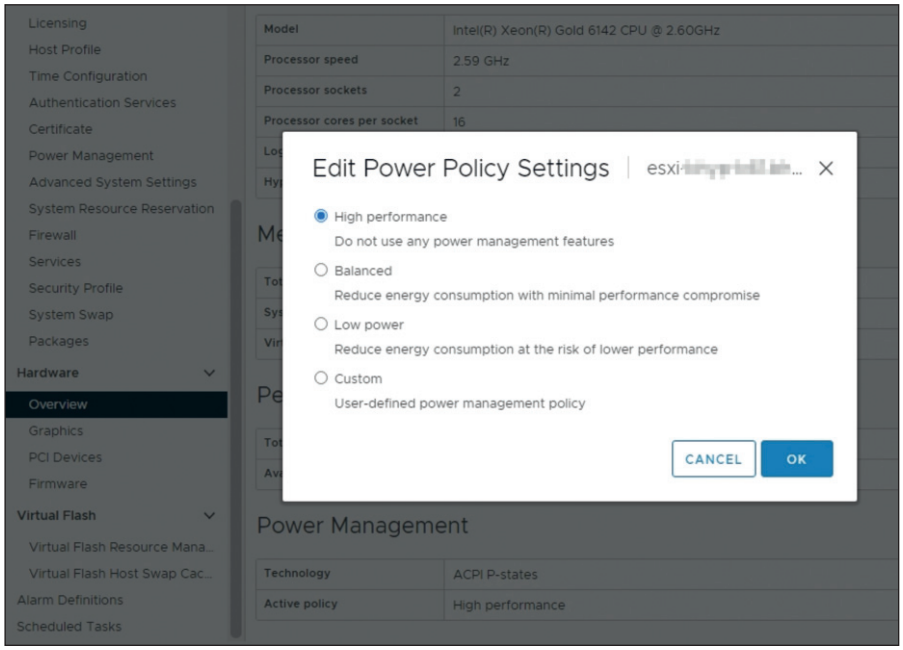

Bild 3: ESXi erlaubt eine granulare Konfiguration des Energiesparverhaltens.

der Hosts auf "High Performance" stehen, um optimale Ergebnisse zu erzielen. Bei aktuellen Servern führender Hersteller kann die Einstellung im BIOS auf "OS Controlled" lauten und in ESXi auf "High Performance". Auf der sicheren Seite sind Sie jedoch, wenn Sie das "High Performance"-Profil bereits im BIOS festlegen. In diesem Fall laufen Ihre CPUs mit der maximalen Taktfrequenz und liefern stetig die Rechenleistung aus, die für den Betrieb Ihrer VMs notwendig ist.

Sind Sie gehalten, Energie in Ihrem virtualisierten RZ zu sparen, können Sie dies mit dem Herunterfahren einzelner Hosts erreichen, wenn – beispielsweise nach Feierabend – der Anspruch and die Leistung sinkt. Besonders gravierend ist dieser Effekt in Clustern, die VDI-Maschinen oder Terminalserver beherbergen und deren Leistungsaufnahme daher linear mit der Anzahl der aktuell angemeldeten Benutzer skaliert. VMware bietet für diesen Anwendungsfall bereits seit vSphere 3 das "Distributed Power Management" (DPM), das bei sinkender Last die VMs mittels DRS auf wenige Hosts im Cluster konsolidiert und die überschüssigen Hosts herunterfährt.

Um die Hosts bei steigender Last wieder hochzufahren, war ursprünglich das standardmäßige Wake-On-LAN (WoL) vorgesehen – eine weitere Funktion, die Sie bei Bedarf im BIOS einstellen müssen. Seit vSphere 4 unterstützt DPM die weitaus komfortablere IPMI-Schnittstelle, mit der alle kompatiblen Server ausgestattet sind. Dabei kommuniziert Ihr vCenter direkt mit dem Hardwaremanagement des einzuschaltenden Servers (iLO, iDRAC, iRMC und so weiter). Hierfür ist ein entsprechend autorisiertes Benutzerkonto erforderlich. Dieses können Sie über das zentrale Servermanagement oder lokal auf jedem Host einzeln hinzufügen und berechtigen.

Die für das Einschalten eines Servers über IPMI notwendigen Berechtigungen sind von Hersteller zu Hersteller unterschiedlich. In manchen Konstellationen muss der entsprechende Benutzer volle Administratorrechte auf der Hardware-Verwaltungsschnittstelle bekommen.

Leider stößt DPM genau in dem Bereich auf seine Grenzen, wo die größten Energieeinsparungen zu erwarten sind: bei hyperkonvergenten Clustern mit VDI-Workloads. Durch die Notwendigkeit, Daten im vSAN in hinreichender Güte vorzuhalten, lassen sich nicht so viele Hosts gleichzeitig herunterfahren, wie für eine signifikante Energieeinsparung nötig wäre. Hier bleibt Ihnen nur die Möglichkeit, mittels Orchestrierung die Energierichtlinie der Hosts nach Feierabend anzupassen, sodass die Energiesparfunktionen der Hardware zum Einsatz kommen. Ein solches Konstrukt müssen Sie unbedingt ausgiebig testen und im Betrieb genau beobachten, damit das Nutzererlebnis der VDI-Nutzer nicht beeinträchtigt wird.

Hyperthreading und die Frage nach Spectre

Bereits seit vielen CPU-Generationen verfügen Intel-Prozessoren über das Hyperthreading-Feature (HT), womit dem Betriebssystem doppelt so viele CPU-Kerne angeboten werden wie in der CPU physisch vorhanden sind. Nach anfangs recht widersprüchlichen Empfehlungen ist VMware zwischenzeitlich zur Ansicht gelangt, dass Hyperthreading auch in der Virtualisierung ein wichtiges Feature ist, das die Flexibilität der CPU-Zuweisung erhöht und so die Gesamtperformance deutlich steigert.

Die im Januar 2018 veröffentlichten Sicherheitslücken "Spectre" und "Meltdown" sowie deren Nachfolger "Zombieload" und "L1TF-VMM", die von der Hyperthreading-Implementierung Gebrauch machen, befeuerten die Diskussion über Hyperthreading erneut. Die Behandlung dieser Sicherheitslücken erfolgt inzwischen auf drei Ebenen:

- In der CPU selbst: Aktuelle Intel-CPUs sind nicht mehr gegenüber bekannten Implementierungen von Spectre und Meltdown anfällig. Falls Sie noch ältere CPUs (Baujahr 2017 oder älter) im Einsatz haben, müssen Sie davon ausgehen, dass diese betroffen sind.
- In der BIOS-Firmware: Die meisten Mainboard-Hersteller haben inzwischen Microcode-Updates veröffentlicht, die zumindest die bekannten Sicherheitslücken neutralisieren.
- Im Hypervisor: VMware hat bereits für vSphere 5.5 einen alternativen Scheduler ("Side Channel aware Scheduler"; SCA) veröffentlicht, der die Auswirkungen von Spectre & Co. auf Kosten von einem Teil der Performance stark begrenzt.

Mit vSphere 6.7U2 stellte VMware eine neue Version (SCAv2) vor, deren Einfluss auf die Performance deutlich kleiner ist. Der Preis dafür ist der etwas schwächere

Schutz – der unerlaubte Zugriff auf Prozessdaten innerhalb einer VM wird vom SCAv2 nicht verhindert. Die Sicherheitslücke und ihre Behandlung durch Aktivieren von SCA sind im Knowledge-Base-Artikel [4] beschrieben.

Sollten Sie nach der Analyse Ihrer spezifischen Bedrohungslage beschließen, auf Nummer sicher zu gehen und Hyperthreading abzuschalten, können Sie dies nur im BIOS tun. Die Änderung des Hyperthreading-Status erfordert zwingend einen Neustart.

In die Karten geschaut

Bisher haben wir die BIOS-Einstellungen betrachtet, die das Verhalten der zentralen Einrichtungen des Servers wie Mainboard, CPU oder RAM beeinflussen. Nicht weniger wichtig sind Konfigurationen der im Server verbauten Peripherie. Viele dieser Einstellungen sind nur im Bootvorgang über die BIOS-Oberfläche möglich, für andere wiederum stellen die Hersteller ein ESXCLI-Plug-in oder ein separates Managementwerkzeug innerhalb des ESXi-Servers zur Verfügung.

Eine der ersten Konfigurationen, die Sie noch vor der Installation des Hypervisors vornehmen müssen, betreffen den Storage-Controller und die logischen Volumes. Für das Boot-Laufwerk ist eine RAID1-Redundanz üblich. Dieses Spiegel-Volume müssen Sie natürlich erzeugen, bevor Sie ESXi darauf installieren. Planen Sie eine vSAN-Konfiguration, müssen Sie sicherstellen, dass der Storage-Controller, an den die Festplatten angeschlossen sind, keinerlei RAID-ähnliche Intelligenz aktiviert hat und die Platten direkt an das Betriebssystem durchreicht.

Falls in Ihrer Umgebung alle Server und somit auch die ESXi-Hosts vom SAN booten, müssen Sie die erforderliche Konfiguration des Storage Adapters ebenfalls vor der Installation vornehmen. Hält sich die am einzelnen Server vorzunehmende Konfiguration eines Fibre-Channel-HBA normalerweise in Grenzen, so ist die Konfiguration eines iSCSI-HBA um einiges komplexer und schließt die IP-Konfiguration des Adapters, der Portaladresse und der CHAP-Authentifizierung sowie weitere iSCSI-spezifischen Einstellungen mit ein.

Ein weiterer Anwendungsfall, der einen direkten Eingriff in die Hardwarekonfiguration der Systemperipherie erfordert, betrifft das Link Layer Discovery Protocol (LLDP). ESXi unterstützt dieses zwar nativ, jedoch haben einige Netzwerkkarten (beispielsweise Intel X710-basierende Adapter von HPE) einen in der Firmware integrierten LLDP-Agenten, der die LLDP-Frames selbst auswertet und nicht an den Hypervisor weitergibt. Obwohl die von Intel veröffentlichte ESXCLI-Erweiterung für die Verwaltung von Netzwerkkarten [5] einen Befehl zur Deaktivierung von LLDP im Treiber unterstützt, ist es in den meisten Fällen dennoch erforderlich, LLDP auf der Karte selbst zu deaktivieren.

Fazit

Das BIOS eines Servers, der für den Einsatz als ESXi-Host vorgesehen ist, beinhaltet viele Einstellungen, die für die Performance und Betriebsstabilität der Umgebung entscheidend sind. Um diese Konfigurationen über viele Hosts hinweg zu vereinheitlichen, sollten Sie das BIOS stets aktuell halten, Workload-Profile einsetzen und Ihre Server in eine zentrale Hardwareverwaltung integrieren. Bei den Energiesparfunktionen müssen Sie sicherstellen, dass die CPUs Ihrer Hosts zumindest im Hochbetrieb auf maximale Performance eingestellt sind. Nur so steht die erwartete Leistung Ihrer Server- und Desktop-VMs auch tatsächlich zur Verfügung. *(jp)*

Link-Codes

[1] Intel-Virtualisierungstechnik
m3p31

[2] AMD-Virtualisierung
m3p32

[3] ESXi Performance Best Practices Whitepaper
ls1ca

[4] VMware KB: L1TF-VMM-Sicherheitslücke
m3p33

[5] Intel ESXCLI-Plug-in für Netzwerkkarten
m3p34

Sichere Shell-Skripte schreiben

Kleinvieh macht auch Mist

von Tam Hanna

Holistische Systemsicherheit bedeutet, auch dem kleinsten Teil Aufmerksamkeit zu widmen, auf dass es nicht als Einfallsloch in das System dient. Ein Beispiel für kleine, wenig beachtete, aber dennoch potenziell sehr gefährliche Teile der IT sind Shell-Skripte. Sie sorgen immer wieder für ernste Sicherheitsprobleme, weshalb wir Best Practices zur sicheren Skript-Programmierung präsentieren.

Einen eindrucksvollen Beleg für die mögliche Schädlichkeit von Shell-Skripten lieferte das amerikanische Softwareunternehmen Valve. Die Linux-basierte Version des Spielediensts Steam brachte ein Skript mit, das normalerweise nur für kleinere Einrichtungsaufgaben zuständig war. Verhängnisvollerweise [1] fand sich dort folgende Passage:

```
rm -rf "$STEAMROOT/"*
```

Dieses für das Löschen des Verzeichnisses "$STEAMROOT" verantwortliche Kommando bekommt Probleme, wenn die Umgebungsvariable nicht gesetzt ist. Die Bash-Shell löst dann keinen Fehler aus, sondern "zerlegt" die Umgebungsvariable einfach zu einem leeren String. Lohn der Mühen ist folgendes Kommando, das sich rekursiv durch das gesamte Dateisystem arbeitet und alle Informationen zerstört:

```
rm -rf "/"*.
```

Einige Benutzer entgingen einem Totalverlust dadurch, dass ihre Steam-Ausführungsumgebung unter SELinux-Jail lief. Andere waren nicht so glücklich, weshalb es an der Zeit ist, sich Maßnahmen zur defensiven Programmierung von Shell-Skripten näher anzusehen.

Shell-Variante festlegen

Unter unixoiden Betriebssystemen stehen Dutzende von Shells zur Verfügung, die sich nur durch die Unterstützung des POSIX-Standards ähneln und diverse proprietäre Funktionen mitbringen. Bei der Nutzung von Shell-spezifischem Code in anderen Shells tritt oft ein undefiniertes Verhalten auf. Dies mag in einer kontrollierten VM-Umgebung kein Problem sein, doch das Deployment in einem Docker- oder sonstigem Cluster ändert die Lage.

Das häufigste Problem ist die Verwendung der als Shebang bezeichneten Sequenz "#!/bin/sh", die gemäß POSIX-Standard die vom System vorausgewählte Shell betrifft. Ein Shebang ist unter Unix eine mit "#!" beginnende erste Zeile eines Skripts. Sie legt fest, welcher Interpreter für die Abarbeitung des Skripts zu verwenden ist. In der Literatur finden sich dafür auch die Begriffe shabang, hash-bang, pound-bang oder hash-pling.

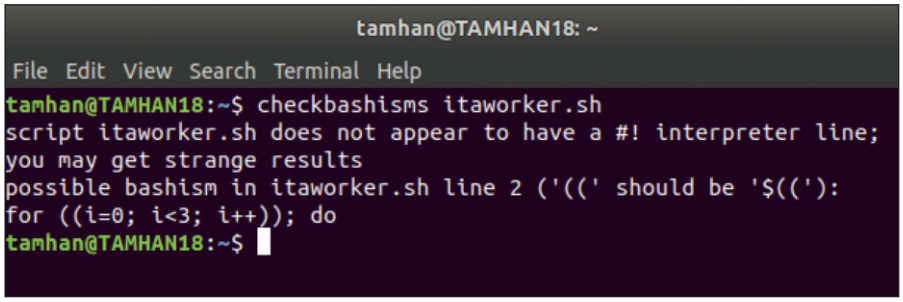

Bild 1: Mit dem checkbashisms-Befehl lässt sich Bash-spezifischer Code identifizieren.

Die unter Skriptprogrammierern weit verbreitete Bash ist nur selten die Standard-Shell. Auf unserer auf Ubuntu 18.04 basierten Workstation lässt sich dies durch Eingabe des which-Kommandos prüfen. So liefert *which sh* die Ausgabe "/bin/sh" zurück und der Befehl *which bash* ergibt "/bin/bash". Zur Umgehung des Problems stehen mehrere Methoden zur Verfügung. Am einfachsten ist es, im Shebang explizit die Verwendung der Bash-Shell über *#!/bin/bash* anzuweisen. Insbesondere in Cloudumgebungen gilt, dass es nicht vernünftig ist, das Vorhandensein der Bash-Shell anzunehmen. Wer sein Skript ohne spezifische Instruktionen aufbaut, spart sich beim Deployment Aufwand.

Der checkbashisms-Befehl hilft Administratoren dabei, Bash-spezifische Programmelemente zu finden und zu beseitigen. Zur Installation des Werkzeugs müssen Sie das Paket "devscripts" laden. Danach erzeugen wir ein kleines Shell-Skript, das eine Bash-spezifische For-Schleife enthält:

```
!#/bin/sh
for ((i=0; i<3; i++)); do
 echo "$i"
done
```

Wenn Sie dieses Skript mit *checkbashisms* analysieren, beschwert es sich über die Schleife (Bild 1). Wichtig ist, dass das Werkzeug Bash-spezifischen Code nur moniert, wenn der Shebang nicht auf die Bash verweist. Die folgende Version passiert die checkbashisms-Kontrolle ohne Probleme:

```
#!/bin/bash
for ((i=0; i<3; i++)); do
 echo "$i"
done
```

Passwörter sicher speichern

Administratoren verwenden Shell-Skripte gerne zur Automatisierung von Systemaufgaben, etwa zum Kopieren von Dateien oder dem Aktualisieren von auf Servern befindlichen Informationen. Dazu sind ein Passwort und/oder ein Benutzername erforderlich, die Sie nicht mit der Allgemeinheit teilen sollten.

Andererseits müssen Sie die Credentials zur Verfügung stellen, da Sie diese sonst bei jeder Ausführung des Skripts von Hand eingeben müssen. Angreifer erbeuten auf gekaperten Systemen gern alle Shell-Skripte, die sie bekommen. Finden sie dabei eine Gruppe von Passwörtern, potenziert sich der entstandene Schaden.

Der erste Weg zum Erfüllen dieser Bedingung besteht darin, die Credentials über Parameter an das Skript zu übergeben. Hierzu wäre folgende Vorgehensweise geeignet, wobei die Variablen "$1" und "$2" für den ersten und zweiten Parameter stehen:

```
#!/bin/bash
a=$1
b=$2
while [ TRUE ]; do
  sleep 1;
done
```

Die Ausführung erfolgt per *./itaworker.sh tam pass*. Dies ist aus sicherheitstechnischer Sicht kritisch. Das Linux-Kommandozeilenwerkzeug "ps" zur Auflistung laufender Prozesse gibt auf Wunsch die zum Aufruf übergebenen Parameter aus. Bild 2 zeigt, wie ein Angreifer in diesem Fall die Credentials frei Haus geliefert bekommt – die Eingabe von *ps* ist eine der ersten Aktionen eines Angreifers.

Sofern die dynamische Parametrisierung des Shell-Skripts unbedingt erforderlich ist, empfiehlt sich die Auslagerung der Credentials in Umgebungsvariablen. Diese lassen sich durch einen gewöhnlichen ps-Scan nicht finden, denn der Zugriff auf "/proc/self/environ" setzt bei korrekter Konfiguration fortgeschrittene Benutzerrechte voraus. Diese Vorgehensweise ist allerdings auch nicht universell akzeptiert – das Linux Documentation Project [2] empfiehlt, die Informationen über eine Pipe oder eine Redirection zu liefern.

In der Literatur herrscht durch die Bank Einigkeit, dass es auf jeden Fall besser ist, wenn sich das Passwort nicht im Klartext in der Skriptdatei findet. Ein Weg zum Erreichen dieses Ziels ist die Verwendung einer umkehrbaren Hash-Funktion, die ihre als Salt bezeichneten Ausgabewerte wieder in den Ursprungswert konvertiert.

Sofern Sie für die Salt-Rückumwandlung ein Programm verwenden, dass der Angreifer bei einem Scan nicht erwischt, erschweren sie ihm das Leben. Denn möchte er die Credentials im Klartext haben, muss er sich nochmals einloggen und das Kommando zur Ausführung bringen. Im

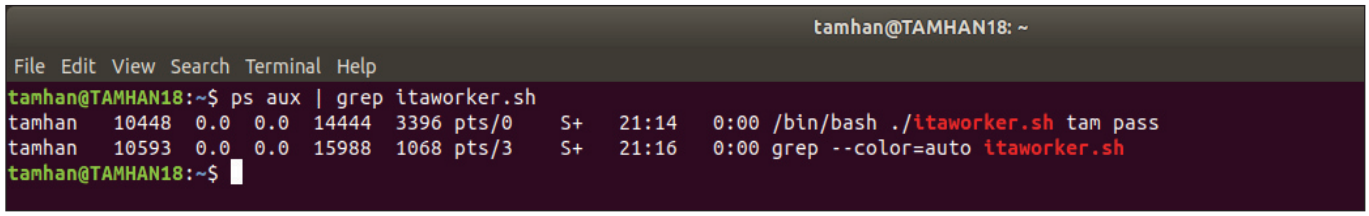

Bild 2: Die Verwendung von Parametern mit Credentials ist der Sicherheit nicht zuträglich.

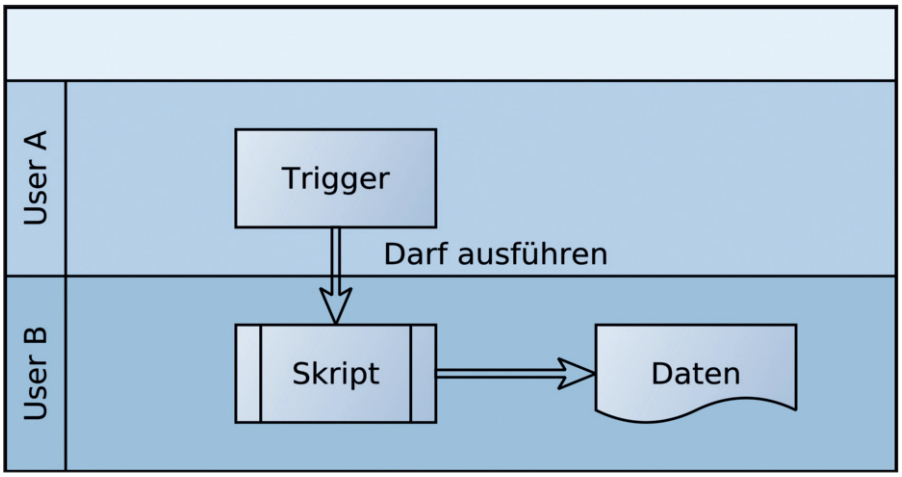

Bild 3: Das Unix-Benutzerrechtesystem schützt die Passwortkonfigurationsdatei vor unberechtigten Zugriffen.

Idealfall gehört die Binärdatei dann einem Benutzer, der ausschließlich für die Shell-Skript-Ausführung verantwortlich ist.

Möglichkeit Nummer zwei setzt (Bild 3) auf eine separate Konfigurationsdatei, die andere Leserechte aufweist. Sofern der für die Ausführung der Skripte verantwortliche User der Einzige ist, der diese Datei lesen darf, ist ein Angriff auf die anderen Benutzerkonten in Bezug auf Credentials unkritisch.

Um die Struktur des Skripts zu verstecken, bietet sich die von Francisco Rosales entwickelte Software SHC [3] an. Das Paket lässt sich unter Ubuntu über den Paketmanager installieren. Der Aufruf erwartet den Parameter "-f" zum Festlegen der Quelldatei:

```
shc -f itaworker.sh

ls itaworker.sh
```

Nach der Abarbeitung finden wir zwei zusätzliche Dateien: Neben der ausführbaren Datei mit der Endung ".X" finden wir auch ein File namens ".X.C", das den C-Code des Skripts bereitstellt.

SHC verpackt das Shell-Skript in einen C-Wrapper, den es danach über den C-Compiler in eine Binärdatei umwandelt. Löschen Sie die Dateien "itaworker.sh" und "itaworker.sh.x.c", ist es für den Angreifer danach schwieriger, den Code des Skripts zu erbeuten. Angemerkt sei, dass die große Verbreitung von SHC dazu geführt hat, dass es Dekompilationswerkzeuge gibt [4].

Temporäre Dateien als Sicherheitslücke

Ein Skript muss die Credentials anderen Applikationen zur Verfügung stellen. Hierzu kommen gern temporäre Dateien zum Einsatz. Legen Sie das File im Ordner "/tmp/" an, sollte es nach der Skript-Abarbeitung verschwinden. Ein klassisches Beispiel dafür ist folgendes Skript, das eine Datei mit einem Passwort anlegt:

```
#!/bin/sh
SECRETDATA="Der ITA sendet Gruesse"
echo > /tmp/mydata
chmod og-rwx /tmp/mydata
echo "$SECRETDATA" > /tmp/mydata
```

Probleme treten auf, wenn ein Angreifer die Temporärdatei über die sehr schnell reagierende Filesystem-Watcher-API überwacht. Im Fall des vorliegenden Skripts könnte er *fopen* aufrufen und die Datei als geöffnet markieren, um die abgelegten Informationen sofort oder später abzuernten.

Wie zuvor bei den Plaintext-Passwörtern gilt auch hier, dass der sicherste Weg der ist, die Credentials nicht in eine temporäre Datei zu schreiben. Ist dies zur Erfüllung der Funktion jedoch notwendig, bietet sich folgender Weg an:

```
#!/bin/sh
SECRETDATA="ITA sendet Gruesse"
umask 0177
FILENAME="$(mktemp
  /tmp/mytempfile.XXXXXX)"
echo "$SECRETDATA" > "$FILENAME"
```

Der umask-Befehl weist neu erzeugten Dateien die als Parameter übergebenen Zugriffsattribute zu. Zweitens erzeugen Sie durch das mktemp-Kommando einen jedes Mal zufälligen Dateinamen, was das Konfigurieren der Watcher-API erschwert. Zahlreiche Quellen sehen die Verwendung temporärer Dateien in Shell-Skripten allgemein kritisch. Unter [5] finden Sie eine Liste von Schwachstellen, die im Zusammenhang mit Tempfiles auftreten.

Benutzereingaben gegen Fehl-Parsings absichern

Häufigster Kritikpunkt gegen die Shell-Ausführungsumgebung ist die unzureichende Ausstattung mit Methoden zur String-Verarbeitung. Im Zusammenspiel mit der Möglichkeit, Variablen direkt an die Shell zu übergeben, entstehen gravierende Sicherheitslücken. Ein absoluter Klassiker ist das vorliegende Skript, das einen Parameter von STDIN einliest und diesen per eval-Befehl ausführt:

```
#!/bin/bash
read BAR
eval $BAR
```

Der eval-Befehl ist normalerweise zum Ausführen von Berechnungen und harmlosen Kommandos vorgesehen. Eine legitime Vorgehensweise wäre das Anliefern des folgenden Echo-Kommandos (bei Bedarf mit einem Pipe-Redirect):

```
./itaworker.sh
echo hello
```

Ein bösartiger User übergibt an dieser Stelle keinen Echo-Aufruf samt Redirection, sondern ein Kommando wie das weiter oben erwähnte "rm". Die Shell erkennt dies prinzipbedingt nicht, weshalb das Skript Daten zerstört.

Da manche eval-basierten Aufgabenstellungen ohne dieses Kommando überhaupt nicht machbar sind, monieren Shell-Überprüfungswerkzeuge wie "shellcheck" die Verwendung des Kommandos nicht – das vorliegende Skript würde die Fehlermeldung "Double quote to prevent

globbing and word splitting." auslösen, die sich folgendermaßen beheben lässt:

```
#!/bin/bash
read BAR
eval "$BAR"
```

Da diese Änderung aber keine Verifikation der in "BAR" enthaltenen Werte durchführt, ist die Sicherheitslücke nicht behoben. Aus sicherheitstechnischer Sicht ist der einzig richtige Weg zum Handhaben des eval-Kommandos dessen Nicht-Nutzung. Ist dies keine Option, achten Sie darauf, die eingegebenen Kommandos auf Strings wie "rm" und auf nicht alphanumerische Zeichen zu überprüfen.

Abfragen sind ein weiterer Weg, um Probleme zu provozieren. Im nächsten Beispiel wollen wir ein Skript ansehen, das überprüft, ob der Benutzer den String "foo" eingibt. Die Payload, die hier nur aus einem Aufruf von Echo besteht, soll nur bei Eingabe von "foo" zur Ausführung gelangen:

```
#!/bin/bash
read FOO
if [ x$FOO = xfoo ] ; then
    echo $FOO
fi
```

Bei einer oberflächlichen Überprüfung der Skriptverhaltens sehen wir, dass beispielsweise die Eingabe "itahallo" nicht zur Ausführung der Payload führt. Über-

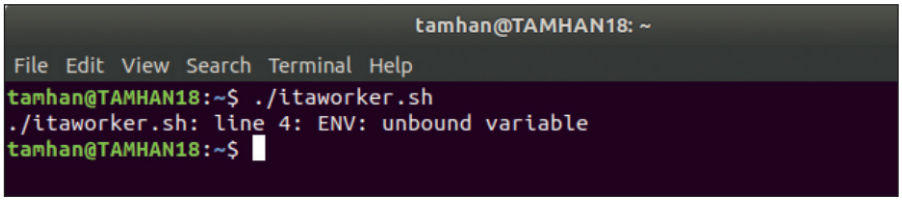

Bild 4: Die Option "nounset" beendet die Programmausführung beim Antreffen nicht existierender Variablen.

geben wir jedoch einen String nach dem Schema "foo = xfoo -<beliebig>", kommt die Payload trotzdem zur Abarbeitung:

```
./itaworker.sh
foo = xfoo -o sdjs
foo = xfoo -o sdjs
~$
```

Der Schadensfall tritt auf, wenn statt der Echo-Payload etwas Empfindlicheres wie ein eval-Aufruf abzusichern ist. Im schlimmsten Fall befindet sich in der Payload dann ein Kommando, das aufgrund der ungültigen Werte Schaden anrichtet.

Zur Umgehung des Problems bietet sich – wie auch in anderen Einsatzszenarien der Shell – die Verwendung von Quoting an:

```
if [ "$FOO" = "foo" ] ; then
    echo $FOO
fi
```

Die "neue" Version des Skripts evaluiert den Inhalt der Variable komplett, was die Abweisung von ungültigen Eingaben zur Folge hat:

```
./itaworker.sh
foo
foo
./itaworker.sh
foo = xfoo
```

Bei der Arbeit mit älteren Shells ist das Quoting noch wichtiger. Ein nach dem folgenden Schema aufgebautes Programm ließe sich durch Eingabe von einem per Semikolon aufgeteilten String zum Ausführen von beliebigem Code animieren:

```
#!/bin/sh
read VAL
echo $VAL
```

Ursache dieses Problems ist, dass ältere Shells keine Parsing-Läufe durchführen, bevor sie die Variable substituieren – das Semikolon würde das Kommando auftrennen und den darauffolgenden Teil zur Ausführung bringen.

Parsingfehler ausschließen

Die recht archaische Sprachsyntax sorgt mitunter für seltsames Verhalten wie

im folgenden Beispiel, das die Datei "/bin/werkzeug_VAR" aufrufen soll:

```
ENV_VAR="fehlwert"
ENV="werkzeug"
echo /bin/$ENV_VAR
```

Wenn Sie es in der Kommandozeile ausführen, sehen Sie, dass es den in der Variable "ENV_VAR" gespeicherten Wert verwendet. Die Shell trennt "echo /bin/$ENV_VAR" am Unterstrich ab. Zur Umgehung dieses Problems ist es empfehlenswert, prinzipiell alle Variablenzugriffe in geschwungene Klammern zu setzen. Eine korrigierte Version unseres Skripts sieht folgendermaßen aus:

```
ENV_VAR="fehlwert"
ENV="werkzeug"
echo /bin/${ENV}_VAR
```

Das Ausführen des korrigierten Skripts zeigt, dass der ausgegebene String nun richtig ist:

```
./itaworker.sh
/bin/werkzeug_VAR
```

Problem Nummer zwei betrifft die Analyse von Benutzerberechtigungen unter Verwendung der Shell. Mit "$UID" und "$USER" stehen zwei Variablen zur Verfügung, die sich zur Analyse der Berechtigungen des gerade aktiven Benutzers einspannen lassen, etwa:

```
if [ $UID = 100 -a $USER = "myuser-
  name" ] ; then
    cd $HOME
fi
```

Leider ist nicht sichergestellt, dass die beiden Variablen zur Laufzeit wirklich den Wert enthalten, der dem aktiven Benutzer entspricht. Während zahlreiche Shells mittlerweile die Variable "$UID" sichern, gilt dies meist nicht für "$USER" – in älteren Shells ist es sogar möglich, beide Variablen mit beliebigen Werten auszustatten. Zur Umgehung dieses Problems bietet es sich an, Benutzername und User-ID durch Aufrufen von Betriebssystemkommandos zu ermitteln. Die so zurückgegebenen Werte lassen sich nur durch Setzen der Path-Variable beeinflussen, was den Angriff verkompliziert.

In besonders kritischen Ausführungsumgebungen ist stets ratsam, die Utility-Pfade immer voll auszuschreiben. Rufen Sie "Echo" beispielsweise über seinen Pfad "/bin/echo" auf, ist das aktivierte Programm von der Path-Variablen unabhängig.

Abschließend sei noch auf den Umstand hingewiesen, dass die Bash-Shell eine "Spezialversion" des Shebangs kennt (*#!/bin/bash -p*). Diese Variante weist die Shell dazu an, die Dateien ".bashrc" und ".profile" nicht im Rahmen des Starts auszuführen. Die Variable "SHELLOPTS" und die in "ENV" und "BASH_ENV" angelegten Startskripte fallen in diesem Betriebsmodus ebenfalls unter den Tisch. Sinn dieser Vorgehensweise ist, dass vom Angreifer durchgeführte Manipulationen der Umgebung nicht durch zufällige Skript-Ausführungen aktivierbar sind.

Strengere Verifikation der Ausführung aktivieren

Shell-Skripte sind für Nicht-Programmierer vorgesehen, die damit häufige Aufgaben in ihrem täglichen Leben mit UNIX-Betriebssystemen beschleunigen. Als Entgegenkommen an diese Nutzerschicht sind so gut wie alle Shells "permissiv" eingestellt: Im Fall eines Fehlers versuchen diese, das Skript weiter auszuführen. Diese für einen weniger technisch versierten Nutzer durchaus freundliche Vorgehensweise ist kritisch, weil Sicherheitsprobleme verursachende Zweideutigkeiten die Abarbeitung des Skripts nicht beenden.

Das zuvor genannte Problem der Zerstörung von Dateien wäre bei folgender Einstellung der Shell nicht aufgetreten:

```
#!/bin/bash
set -o nounset
TAMS_VAR="fehlwert"
echo /bin/${ENV}_VAR
```

Der set-Befehl aktiviert bei Aufruf mit dem Parameter "-o" Ausführungsoptionen, die das Laufzeitverhalten des Interpreters beeinflussen. Wir verwenden hier "nounset", was nicht gesetzte Variablen als Fehler betrachtet. Die Ausführung des zuvor gezeigten Shell-Skripts scheitert nun mit dem in Bild 4 gezeigten Fehler.

Mit *set -o errexit* erfolgt eine Programmbeendigung, wenn ein vom Shell-Skript ausgerufener Befehl nicht den Rückgabewert null zurückgibt. Sein Einsatz lässt sich durch Kopieren einer nicht vorhandenen Datei überprüfen:

```
#!/bin/bash
set -o errexit
cp "existiertnicht" "baum.txt"
echo "Kopieren erfolgreich"
```

Das cp-Kommado liefert keine Null zurück, wenn es einen Parameter nicht findet, weshalb die an Echo übergebene Statusmeldung nicht in der Kommandozeile erscheint.

In Skripten gibt es immer wieder Situationen, in denen ein Teil des zu erledigenden Codes unkritisch ist und die per "-o" festgelegten Befehle sich durch Aufruf von "+o" deaktivieren lassen. Im folgenden Skript ist die errexit-Option bei der Abarbeitung des Kopierbefehls nicht mehr aktiv:

```
#!/bin/bash
set -o errexit
set +o errexit
```

Link-Codes

[1] Probleme mit Steam-Skript
l3z11

[2] Linux Documentation Project
l3z12

[3] SHC-Manpage
l3z13

[4] Grenzen von SHC
l3z14

[5] Sicherer Einsatz von Temp-Dateien
l3z15

[6] Shellcheck
l3z16

[7] Liste der Shellcheck-Checks
l3z17

[8] Shellcheck im Web
l3z18

```
cp "existiertnicht" "baum.txt"
echo "Kopieren erfolgreich"
```

Zu guter Letzt sei noch auf ein Problem bei der Verarbeitung von per Pipe verbundenen Kommandos hingewiesen. Normalerweise scheitert ein derartiger Befehl nur, wenn das in der Pipe-Folge letzte Kommando nicht den Wert null zurückliefert. Sollen Fehler irgendwo in der Hierarchie für einen Abbruch sorgen, aktivieren Sie die Option "pipefail" per *set -o pipefail*.

Funktionen verändern globale Variable

Ein weiterer Nebeneffekt der laxen Sprachsyntax ist, dass in Unterfunktionen definierte Variablen unerwartete Nebeneffekte entwickeln. Als Beispiel wollen wir uns ein Skript ansehen, das die Variable "o" sowohl innerhalb als auch außerhalb einer Funktion verwendet:

```
#!/bin/bash

hello_world () {
   o=2
   echo 'hello, world'
}

o=1
echo $o
hello_world
echo $o
```

Mit anderen Skriptsprachen aufgewachsene Entwickler erwarten, dass beide Aufrufe von "echo $o" den Wert "1" zurückgeben. Von Haus aus ist dem nicht so, die in der Funktion durchgeführten Änderungen bleiben auch nach ihrer Ausführung gültig . Zur Behebung des Problems reicht es aus, die Variablen in der Funktion mit dem local-Schlüsselwort auszustatten:

```
#!/bin/bash

hello_world () {
   local o=2
   echo 'hello, world'
}

o=1
echo $o
hello_world
echo $o
```

Obwohl längere Shell-Skripte aus Wartungsaspekten nicht unbedingt empfehlenswert sind, sollten Sie in Funktionen verwendete Variablen immer über das Schlüsselwort "local" absichern. Recycelt ein Kollege irgendwann doch Shell-Code, ist er vor unerwarteten Nebeneffekten besser geschützt.

Automatische Shell-Skript-Prüfung mit Shellcheck

C- und C++-Programmierer prüfen die Ergebnisse ihrer Arbeit seit langer Zeit durch statische Analyse. Ein statisches Analysewerkzeug hat eine Wissensbasis, in der es Programmierfehler vorhält und diese gegen den vorliegenden Code anwendet. Für Shell-Skripte steht mit Shellcheck [6] ein ähnliches Werkzeug zur Verfügung. Das unter der GPLv3 stehende Tool verhält sich von der Syntax her analog zu Lint – es erwartet als Parameter den Namen der Shell-Datei, den es auf unsaubere Elemente überprüft.

ShellCheck sucht nicht nur nach sicherheitsrelevanten Fehlern. In der unter [7] bereitstehenden Liste finden sich vier Dutzend Testkriterien, die auch allgemeine Programmierfehler erkennen und monieren. Zu guter Letzt gibt es unter [8] noch eine Webversion des Werkzeugs, die sich mit Skripten füttern lässt und diese ohne lokale Installation auf Korrektheit und Sicherheit überprüft.

Fazit

Die enge Integration zwischen Skripten und der für die Ausführung von Kommandos vorgesehenen Shell sorgt im Zusammenspiel mit einer laxen Syntax-Verifikation dafür, dass eine unsaubere Shell-Programmierung schnell beeindruckend gefährliche Sicherheitslücken produziert.

Die hier vorgestellten Kriterien umschiffen den Gutteil der Probleme. Anders als bei komplizierten Sicherheitslücken ist im Fall von Shell-Skripten die Ursache meist schlichtes Unverständnis dessen, dass eine bestimmte Konstruktion Nebeneffekte entfaltet. *(jp)*