Dirk Fox

Zero Trust

Hintergrund

Das Konzept einer Zero-Trust-Achitektur (ZTA) geht zurück auf Diskussionen des 2004 gegründeten (und 2014 in The Open Group aufgegangenen) Jericho Forums [1]. Jenes verfolgte das Ziel einer „De-Perimetrisierung" der Netzwerk-Sicherheit: Gesucht wurde ein Sicherheitsverständnis, das nicht die Vertrauenswürdigkeit des internen Netzwerks zur Voraussetzung hat. Den Begriff „Zero Trust" prägte 2010 der Forrester-Analyst John Kindervag in einer Veröffentlichung [2].

Eine ZTA ist der Versuch einer Antwort auf das Erfordernis einer IT-Sicherheitsarchitektur, die auch private Geräte und (unternehmens-) fremde Nutzer im internen Netz oder in Cloud-Diensten zulässt und durch eine Kombination unterschiedlicher Schutzmechanismen verschiedene Sicherheitslevel in einer komplexen vernetzten Infrastruktur unterscheidet. Zugleich soll sie vor Angriffen schützen, die über mit Schadsoftware infizierte Geräte aus dem internen Netz erfolgen oder nach einem erfolgreichen Eindringen ins interne Netz von dort ungehindert fortgesetzt werden.

Der Ansatz wurde in den vergangenen Jahren weiterentwickelt, unter anderem von Google („Beyond Corp" [3]), und mündete in der Veröffentlichung einer Special Publication des NIST im Jahr 2020 [5]. Inzwischen bieten viele große IT-Anbieter ihre Produkte in einem Zero-Trust-Kontext an, darunter Citrix, Oracle und Microsoft [6]. Mit der wachsenden Nutzung von Cloud-Diensten und der Verbreitung von Home-Offices hat das Konzept in den vergangenen beiden Jahren erneut an Bedeutung gewonnen.

Grundgedanke

Die Bezeichnung „Zero Trust" ist, wörtlich verstanden, irreführend: Tatsächlich geht es im Kern darum, in immer komplexeren Netzwerkinfrastrukturen das Vertrauen in Endgeräte und Nutzer aus unterschiedlichen Schutzmechanismen und Echtzeit-Analysen zu gewinnen, es ständig zu überprüfen und abhängig vom Ergebnis der Prüfung den Zugang zu bestimmten Anwendungen und Daten zu gewähren. Der Bereich, in dem einem Nutzer oder Endgerät aufgrund einer Überprüfung vertraut wird, soll dabei möglichst klein gehalten werden.

Eine ZTA umfasst zunächst überwiegend bekannte und verbreitete Schutzmechanismen wie die (möglichst starke) Authentifikation von Nutzern (*Multi-Faktor Authentication*, MFA) und Geräten (*Network Access Control*, NAC), die Analyse des Netzwerkverkehrs (*Intrusion Detection System*, IDS sowie *Security Information and Event Management*, SIEM) und den Zugriffsschutz auf Anwendungen und Daten (*Single Sign On*, SSO).

Neu ist, dass für die Bewertung der Vertrauenswürdigkeit von Geräten und Nutzern neben der Authentifikation auch „Indizien" aus Analysen der Geräte und Benutzer herangezogen werden, wie beispielsweise die Aktualität des Betriebssystems, die Zeiten des Zugriffs, der Einsatz eines Virenscanners oder die genutzten Dienste und Protokolle. Dazu soll die ZTA eine permanente Risikobewertung durchführen und Zugang zu Daten und Anwendungen jeweils abhängig von deren Ergebnis gewähren.

Einige Anbieter zählen auch Filtertechnologien wie beispielsweise eine Keyword-basierte Entscheidung über die Zulässigkeit der Versendung von E-Mails bestimmten Inhalts über weniger vertrauenswürdige Netzwerkverbindungen zu den Zero-Trust-Techniken.

Die meisten Veröffentlichungen fordern übereinstimmend die Umsetzung der folgenden fünf elementaren Schutzmechanismen, die in Unternehmensinfrastrukturen auch unabhängig von einer ZTA heute bereits üblich sein sollten:

- **Minimale Berechtigungen** (Least Privileges): Jedem Endgerät und jedem Benutzer dürfen nur die Berechtigungen auf Daten und Peripheriegeräte eingeräumt werden, die zur Erfüllung einer Aufgabe erforderlich sind. Dazu zählt nicht nur eine rollenbasierte Rechtevergabe für Anwendungen, File Shares und Daten, sondern insbesondere auch, dass Benutzern keine lokalen Administrationsrechte eingeräumt werden – eine hoch riskante Unsitte, die noch immer in vielen Unternehmen vorherrscht.
- **Mehr-Faktor-Authentifizierung** (MFA): Für die Authentifikation der Benutzer sind MFA-Verfahren wie Smartcards Voraussetzung. Werden Cloud-Dienste genutzt, müssen dabei ggf. unterschiedliche Verfahren unterstützt werden, wie Token oder TANs.
- **Netzwerksegmentierung**: Netzwerke sollten in unterschiedliche virtuelle LANs (VLANs) unterteilt werden, die jeweils verschiedenen Sicherheitsniveaus entsprechen und je nach Vertrauen in den Benutzer oder das Endgerät automatisch zugewiesen werden. Dabei können z. B. ein Gäste-Netz (nur mit Internet-Zugang), ein Mitarbeiter-Netz, ein Produktions-Netz, ein Peripherie-Netz und ein Server-Netz unterschieden werden.
- **Verschlüsselung**: Ausnahmslos alle Kommunikationsverbindungen – auch im internen Netz – verwenden eine starke, Zertifikats-basierte Authentifikation und werden verschlüsselt.
- **Geräteauthentifikation**: Alle (eigenen) Endgeräte und Peripheriegeräte eines Unternehmens sollten nach dem IEEE-Standard 802.1x nur nach gegenseitiger Authentifikation unter Verwendung von Maschinenzertifikaten und einem RADIUS-Server (Remote Authentication Dial-in User Service) in interne Netzsegmente hineingelassen werden.

Bewertung

Die für eine ZTA geforderte, über die o.g. technischen Schutzmechanismen hinaus gehende Analyse von Geräteeigenschaften (Betriebssystemversion, Update-Status, Vorhandensein eines Virenscanners, Sicherheitseinstellungen, Aktualität der verwendeten Anwendungsprogramme) und des Benutzerverhaltens (Zugriffszeiten, verwendete Protokolle, genutzte Bandbreite, Kommunikationspartner, Verhaltenshistorie) können Hinweise auf

eine mögliche Gefährdung geben. Sie sind allerdings in mehrerlei Hinsicht problematisch:

- Erfahrungen mit IDS- und SIEM-Systemen zeigen, dass automatisierte Analysen eine sehr große Zahl an Fehlalarmen („False Positives“) erzeugen, die behandelt werden müssen und dazu führen können, dass echte Alarme übersehen oder erst mit Verzögerung erkannt werden. Beispielsweise kann ein Zugriff außerhalb üblicher Arbeitszeiten durch eine Dienstreise in einer anderen Zeitzone begründet sein, oder eine spezielle Recherche einen ungewöhnlich großen Bandbreitebedarf verursachen und einen Alarm auslösen.
- Analysen des Netzwerkverkehrs sind bei verschlüsselter Kommunikation nahezu unmöglich; sie erfordern daher meist, dass eine sichere Verbindung „aufgebrochen“ wird (bspw. durch ein Umverschlüsseln von TLS-Verbindungen an der Firewall). Das schafft einen kritischen Angriffspunkt und kann auch zu Verstößen gegen das grundrechtlich geschützte Fernmeldegeheimnis führen.
- Aussagekräftige Analysen des Geräts und des Verhaltens eines Benutzers erfordern eine umfangreiche Protokollierung von Daten, die wiederum eine Verhaltens- und Leistungskontrolle ermöglichen. Ihr Einsatz erfordert eine Betriebsvereinbarung, enge Zugriffsberechtigungen sowie kurze Löschfristen – was wiederum mit dem Ziel konfligiert, möglichst viel Information über Benutzer und Geräte in die Analyse einfließen zu lassen.
- Schließlich kann ein Angreifer, der über eine Schadsoftware ein Gerät im Netz kontrolliert, die Analyse der ZTA täuschen, indem er modifizierte Geräteeigenschaften (bpsw. das Vorhandensein des von ihm deaktivierten Antivirenscanners, falsche Sicherheitseinstellungen oder Betriebssystemversionen) meldet.

Schließlich darf der Aufwand für eine solche kontinuierlich an aktuelle Entwicklungen (bspw. neue Sicherheitslücken) anzupassende Echtzeit-Analyse nicht unterschätzt werden: Um messbare Verzögerungen bei der Zugriffsgewährung zu vermeiden, muss der Dienst mit ausreichenden Ressourcen ausgestattet werden – und stellt obendrein einen möglichen „Single Point of Failure“ dar: Fällt er aus, sind keine Zugriffe im Netz mehr möglich.

Fazit

Der Grundgedanke von Zero Trust, nicht das gesamte interne Netzwerk pauschal zum „vertrauenswürdigen Bereich“ zu erklären, sondern auch dort alle Geräte einer starken Authentifikation zu unterziehen, das Netzwerk zu segmentieren und den gesamten internen Datenverkehr zu verschlüsseln, ist uneingeschränkt richtig. Wer diese Mechanismen nach dem Stand der Technik noch nicht umgesetzt hat, erreicht mit ihrer Einführung eine erhebliche Risikoreduktion, da Angriffe aus dem internen Netz über eingebrachte Fremdgeräte oder mit „eingeschleppter“ Schadsoftware erheblich erschwert werden.

Auch bei der Umsetzung des (uralten) „Least Privilege“-Prinzips im Zusammenhang mit der Vergabe von Berechtigungen haben zweifellos sehr viele Unternehmen noch Handlungsbedarf. Mit der wachsenden Komplexität heutiger IT-Infrastrukturen wird das eine Daueraufgabe bleiben, deren Bewältigung die bei einem Angriff möglichen Schäden jedoch deutlich begrenzen kann.

Der mögliche Sicherheitsgewinn durch darüber hinaus gehende analytische Methoden zur Bewertung der Vertrauenswürdigkeit von Nutzern und Geräten muss allerdings gegen den damit verbunden Aufwand für Betrieb, Pflege und Wartung sowie die diesen inhärenten Risiken (Single Point of Failure, Persönlichkeitsrechte der Betroffenen) abgewogen werden.

Literatur

[1] Jericho Forum: *Commandments v1.2*. https://collaboration.opengroup.org/jericho/commandments_v1.2.pdf, 2007.

[2] John Kindervag (Forrester Research): *Build Security Into Your Network's DNA: The Zero Trust Network Architecture*. http://www.virtualstarmedia.com/downloads/Forrester_zero_trust_DNA.pdf, 2010.

[3] Rory Ward, Betsy Beyer: *BeyondCorp: A New Approach to Enterprise Security*, https://research.google/pubs/pub43231/, Vol. 39, No. 6, 2014, pp. 6-11.

[4] Identity Defined Security Alliance: *The Path to Zero Trust starts with Identity*. Whitepaper, https://www.idsalliance.org/wp-content/uploads/2019/07/IDSA_Zero-Trust_Whitepaper.pdf, 2019.

[5] Scott Rose, Oliver Borchert, Stu Mitchell, Sean Connelly: *Zero Trust Architecture*. NIST Special Publication 800-207, https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-207.pdf, August 2020.

[6] Ken Machtley: *Evolving Zero Trust*. Microsoft Position Paper, https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RWJJdT, Nov. 2021.