Livre blanc

Une brève histoire de Zero Trust : les grandes étapes de la refonte de la sécurité des entreprises

Pourquoi raconter l'histoire de Zero Trust ?

Nombreux sont ceux qui, dans le domaine de la sécurité informatique, pensent que Zero Trust constitue un véritable bouleversement, une refonte fondamentale de la sécurité des entreprises et de la protection des réseaux et des ressources qui abritent nos meilleures idées, connectent nos talents les plus brillants et donnent accès à des outils de productivité transformateurs.

Mais pour comprendre à quel point le modèle Zero Trust est révolutionnaire dans le domaine de la cybersécurité, il est nécessaire de connaître les faiblesses de l'approche traditionnelle de la sécurité du réseau et de comprendre comment l'idée d'une architecture Zero Trust a évolué pour devenir un modèle qui bouleverse fondamentalement un mode de pensée vieux de plusieurs décennies.

Les réseaux 2D et la sécurité cloisonnée

L'architecture réseau en étoile et la sécurité cloisonnée (« caste-and-moat » en anglais, pour château et douves) sont les deux principales métaphores utilisées pour décrire respectivement l'architecture réseau et la sécurité réseau traditionnelles. L'imagerie utilisée dans les deux cas ne date pas d'hier.

L'architecture de réseau en étoile fait référence à des réseaux satellites disposés autour d'un hub central. Ce modèle implique le routage du trafic interne et externe à travers une pile de sécurité dans un data center primaire avant qu'il ne soit envoyé à sa destination. Si cette approche a fonctionné pendant un certain temps, elle s'est avérée plus compliquée et plus coûteuse avec l'adoption du cloud, la décentralisation du personnel et l'importance croissante de la mobilité dans les entreprises.

La sécurité de type « château et douves » (sécurité cloisonnée), quant à elle, fait référence à des réseaux autonomes conçus pour admettre le trafic légitime tout en maintenant les ennemis à l'extérieur de leurs murs. Comme un gardien à la porte, les dispositifs de sécurité internes sont censés laisser entrer les personnes animées de bonnes intentions tout en repoussant les brigands. La transition massive des applications vers le cloud, associée à la migration des travailleurs en dehors du périmètre de l'entreprise, a rendu cette approche obsolète plus rapidement que les boulets de canon ne l'ont fait pour les véritables châteaux.

Les VPN et le Wi-Fi ont encore davantage compliqué le problème. L'ancienne architecture cloisonnée ne donnait aux administrateurs aucun moyen de connecter des hôtes à un réseau sans leur donner carte blanche. En fin de compte, il était impossible de connecter des terminaux à des réseaux sans une certaine forme de segmentation destinée à assurer la sécurité de ces derniers.

Il nous fallait quelque chose de mieux.

802.1X et les problèmes liés au NAC

En 2001, l'IEEE Standards Association a publié sa norme de protocole 802.1X pour le contrôle d'accès au réseau (NAC).

> "Un moyen d'authentifier et d'autoriser les appareils connectés à un port LAN ayant des caractéristiques de connexion point à point, et d'empêcher l'accès à ce port lorsque le processus d'authentification et d'autorisation échoue."
>
> IEEE sur 802.1X →

Peu après, les dispositifs sans fil ont commencé à inclure un demandeur, ou client, 802.1X, qui permettait aux réseaux d'authentifier le terminal avant d'autoriser une connexion. Cette avancée était destinée à offrir la possibilité de verrouiller les réseaux câblés et sans fil, afin que seuls les appareils gérés et les utilisateurs autorisés puissent se connecter. Imaginez le demandeur fournissant une pièce d'identité au videur à la porte du réseau, qui décide qui peut entrer et qui reste à l'extérieur.

Hélas, le modèle NAC n'était pas la panacée, et les problèmes ont commencé avec ce N (pour Network, réseau). Les réseaux internes ont été conçus avec une base de confiance implicite, et essayer de greffer l'authentification/autorisation après coup représentait un travail considérable. Pour que le NAC soit pleinement efficace, tous les ports accessibles doivent être verrouillés, mais tous les appareils ne sont pas compatibles avec la norme 802.1X. L'adoption croissante d'imprimantes connectées à Internet, de lecteurs de badges et d'autres appareils compatibles avec le réseau a généré une faille de sécurité manifeste. Imaginez maintenant que notre videur continue à surveiller cette seule porte du réseau alors que plusieurs (voire des dizaines) d'autres entrées sont disponibles.

Faire tomber les murs de Jéricho et repenser le rôle du périmètre dans la sécurité

En 2003, il était clair que l'utilisation des appareils personnels allait continuer à proliférer et que les entreprises devaient commencer à réfléchir à la manière de protéger les machines qui n'étaient pas enfermées derrière les murs du château. De plus, l'utilisation croissante du chiffrement diminuait l'efficacité des pare-feu de périmètre, ce qui obligeait à choisir entre évoluer pour répondre aux problèmes de capacité imposés par le déchiffrement et l'inspection, ou laisser passer le trafic chiffré sans le remettre en question.

Cette année-là, un groupe multinational de leaders technologiques européens s'est réuni pour aborder des sujets tels que l'authentification des utilisateurs, le chiffrement, la gestion des identités et l'application des politiques. Après s'être officiellement constitué en 2004, le forum de Jéricho a présenté au monde la notion de « dé-périmétrisation ».

Avec un nom rappelant l'histoire biblique des Israélites abattant les murs de l'ancienne ville de Jéricho, le forum a entrepris de résoudre le problème consistant à « faciliter des flux d'informations sécurisés et sans frontières entre les sociétés ».

Outre cette métaphore pertinente, le groupe a publié les Commandements du forum de Jéricho, ce qui se rapproche le plus des vérités venues d'en haut sur la gestion des réseaux sans périmètre. Malheureusement, l'ensemble des contrôles et des mesures d'atténuation prescrits dépassait les capacités de déploiement ou d'administration de la plupart des entreprises à ce stade.

La notion de « Zero Trust» fait son entrée dans le lexique informatique

En 2010, John Kindervag, analyste chez Forrester, a publié un article intitulé « No More Chewy Centers:

Introducing The Zero Trust Model Of Information Security » et, illico presto, nous avions un nouveau mot tendance représentant une nouvelle façon de penser la sécurité des réseaux. L'une des principales affirmations avancées dans l'article est que la simple présence sur un réseau n'est pas suffisante pour justifier la confiance.

C'est là que nous avons commencé à entendre des énoncés comme « l'identité est le nouveau périmètre », explique Lisa Lorenzin, directeur technique de Zscaler Field et experte du Zero Trust. « Nous authentifiions un utilisateur et utilisions cette identité pour déterminer ce qu'il pouvait faire. Avec un peu de chance, nous pouvions rassembler un certain contexte, comme savoir si nous étions en présence d'un appareil géré ou non, et prendre des décisions concernant l'accès sur la base de cette compréhension rudimentaire. »

Cela a représenté un certain progrès. Mais la sécurité de l'entreprise restait cantonnée à la sécurisation des réseaux eux-mêmes. Elle n'était pas encore prête à les abandonner complètement. Nous étions encore loin d'une approche transformationnelle, et l'adoption de ces principes a de nouveau échoué. Il faut dire qu'elle s'appuyait toujours sur le même ensemble d'outils axés sur le réseau : 802.1X et RADIUS au niveau 2, des pare-feu sensibles à l'identité au niveau 3, etc.

La nouvelle méthode n'était qu'un simple NAC avec un nom accrocheur.

Beyond Corp (au-delà du périmètre)

Pendant ce temps, des hackers affiliés à l'Armée populaire de libération (APL) de la Chine ont amené les plus grands noms de l'industrie technologique à reconsidérer la question de la confiance. En 2010, Google a révélé une opération menée en 2009 qui l'avait visé, ainsi que plusieurs autres sociétés technologiques de premier plan, dont Akamai, Adobe et Juniper Networks. Cette campagne a été baptisée « Opération Aurora » par les chercheurs en sécurité de McAfee.

En s'attaquant à l'élite des ingénieurs informatiques, les pirates chinois ont involontairement accéléré le développement de l'architecture Zero Trust dans les meilleurs laboratoires technologiques du pays. Google a développé BeyondCorp en réponse à l'opération Aurora, qui visait à **« déplacer les contrôles d'accès du périmètre du réseau vers les utilisateurs individuels... [pour permettre] de travailler en toute sécurité depuis pratiquement n'importe quel endroit sans devoir recourir à un VPN traditionnel ».**

Mais « Google est une société dirigée par des ingénieurs, pour des ingénieurs, avec un budget pratiquement infini et une infrastructure existante relativement réduite par rapport à de nombreuses entreprises », déclare Lisa Lorenzin. « Et cela leur a quand même pris sept ans et six livres blancs de conception et de mise en œuvre. »

Même avec l'exemple bien documenté de Google, une véritable architecture Zero Trust était encore inabordable pour la plupart des sociétés. Malgré la tentative de « préparer le terrain pour que d'autres entreprises puissent réaliser leur propre mise en œuvre d'un réseau Zero Trust », on était encore bien loin du futur imaginé par Google.

Parallèlement, pour les utilisateurs, la popularité du cloud et l'importance persistante de la mobilité signifiaient que davantage de données étaient disponibles et accessibles depuis l'extérieur du périmètre du réseau que depuis l'intérieur. Une approche généralisée de la confiance était plus urgente que jamais.

Gartner et l'arrivée imminente de l'accès réseau Zero Trust

La société d'études technologiques Gartner est à l'origine des principales avancées de Zero Trust en tant que cadre adaptable à grande échelle. Bien qu'il existait toujours, le terme « Zero Trust » n'était pas encore à l'ordre du jour en 2010 lorsque la société a publié son rapport intitulé « Continuous Adaptive Risk and Trust Assessment (CARTA) ».

Ce document décrivait la nécessité de comprendre qui demande un accès, et d'accorder cet accès

sur la base d'une évaluation dynamique de l'environnement, du contexte disponible et des responsabilités de l'utilisateur.

Lisa Lorenzin décrit CARTA comme « un excellent modèle qui n'a jamais obtenu la portée qu'il méritait ».

Chez Gartner, CARTA s'est finalement transformé en « Zero Trust Network Access » (ZTNA) lorsque le cadre initial n'a pas réussi à s'imposer auprès des professionnels de la technologie (notez l'accent mis sur les réseaux en tant que point central de l'accès !) Mais, fondamentalement, CARTA conserve son importance dans l'histoire de Zero Trust parce que les principes qu'il a définis perdurent sous la forme de ZTNA.

L'importante contribution suivante de Gartner à cette réflexion est venue de la reconnaissance de la convergence des domaines de la mise en réseau et de la sécurité. En 2019, il a concrétisé ce rapprochement en présentant le Secure Access Service Edge (SASE). Cette union a été de courte durée, cependant, et en 2021, Gartner a une nouvelle fois divisé les catégories en introduisant la catégorie de marché Secure Service Edge (SSE) : le SASE sans WAN.

Quel que soit le nom retenu, Gartner s'était déjà imposé comme un arbitre important de ce qui constituait ou non le Zero Trust. Les fournisseurs se bousculaient désormais pour se ranger dans l'une de ses nouvelles catégories de marché.

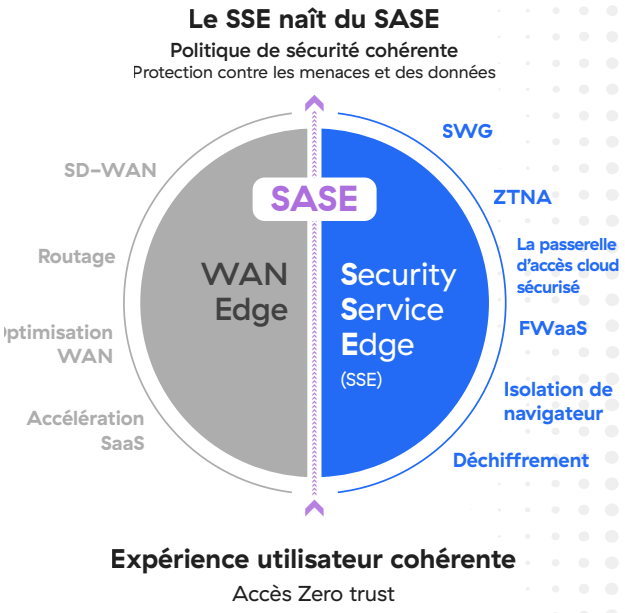
« The Man » (le gouvernement) entre dans la danse : approbation de ZTA par le NIST, l'OMB et le gouvernement

En 2020, le National Institute for Standards and Technology (NIST) a recadré la conversation avec sa norme NIST 800-207 pour l'architecture Zero Trust. Ce nouveau paradigme de cybersécurité était axé sur la protection des ressources et sur le principe selon lequel la confiance ne doit jamais être accordée implicitement, mais doit être constamment évaluée.

Avec ce document, les carcans du périmètre et du VPN ont finalement été éliminés. La priorité est passée de la protection du réseau à la protection des utilisateurs, des données et des applications qui interagissent via le réseau. Zero Trust est désormais synonyme d'accès contextuel et de moindre privilège, applicable à un éventail beaucoup plus large de cas d'utilisation et de flux de trafic.

La norme 800-207 stipule les principes et hypothèses clés de Zero Trust. Voici trois des points les plus critiques (parmi une liste beaucoup plus longue) :

1. Aucune ressource n'est intrinsèquement fiable.
2. Toute communication est sécurisée, quel que soit l'emplacement du réseau. Interrompre et inspecter la demande ; examiner tout le contexte disponible associé à l'utilisateur et à la demande.
3. L'authentification et l'autorisation de toutes les ressources sont dynamiques et rigoureusement appliquées avant que l'accès ne soit autorisé.



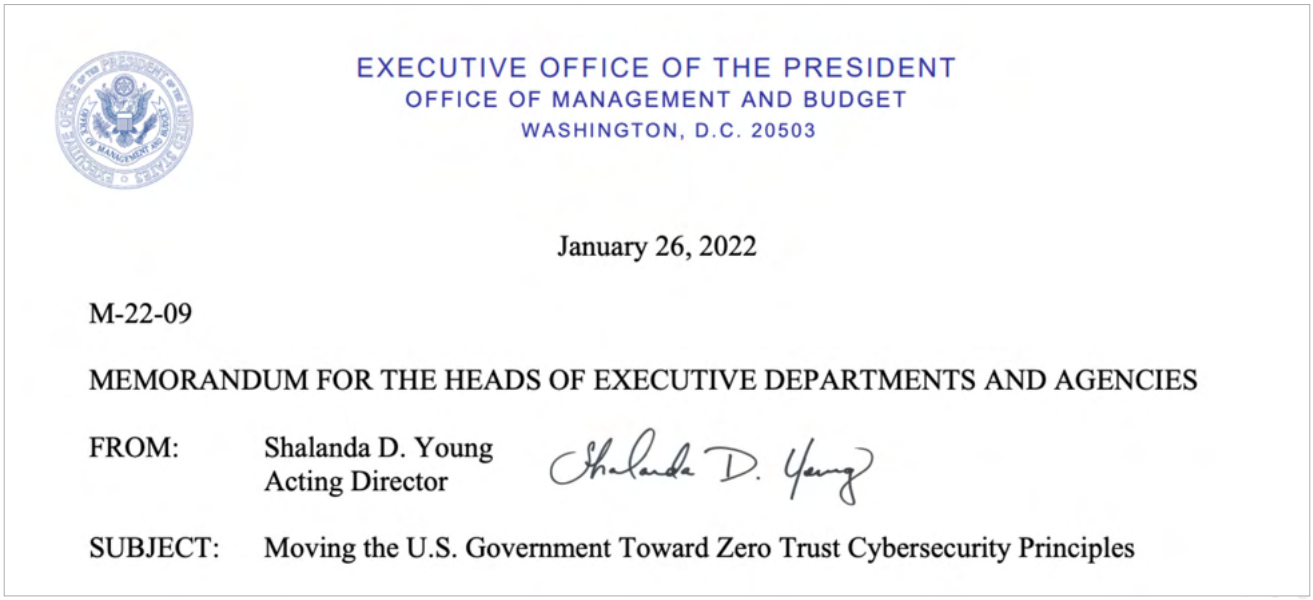
Mais le véritable point de non-retour pour la promotion des principes de Zero Trust est venu du plus haut niveau, du moins aux États-Unis. L'Office of Management and Budget des États-Unis, chargé de mettre en œuvre les politiques présidentielles, a publié sa directive M-22-09 en 2022, stipulant que tous les bureaux du gouvernement fédéral doivent adopter les principes de l'architecture Zero Trust d'ici 2024 et définissant des étapes et des dates cibles claires à cet égard.

« Jusqu'à présent, nous avons disposé de documents d'orientation. Nous avons eu des modèles d'administrateur. Mais c'est la première fois que les choses se mettent réellement en place, avec la stratégie fédérale de Zero Trust », selon Lisa Lorenzin.

L'attaque de la chaîne d'approvisionnement visant la plateforme de gestion informatique Solar Winds (révélée en 2021 et responsable de la compromission d'au moins neuf agences fédérales, dont l'État, le Trésor, la Sécurité intérieure, le Commerce et l'Énergie) était peut-être l'attaque visant l'État la plus cinglante et la plus préjudiciable depuis l'opération Aurora. En réponse, le gouvernement fédéral a tout misé sur Zero Trust, en adoptant cette approche comme référence en matière de cybersécurité pour les années à venir.

Mise en œuvre de Zero Trust

L'approche de Zscaler en matière d'architecture Zero Trust s'aligne étroitement sur le cadre ZTA du NIST et sur la définition de SSE proposée par Gartner. Mais elle va au-delà de ces normes, grâce à son engagement envers trois avancées fondamentales de la pensée Zero Trust. Ensemble, ces principes permettent de mener l'application du principe de Zero Trust à des conclusions logiques.



EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF MANAGEMENT AND BUDGET
WASHINGTON, D.C. 20503

January 26, 2022

M-22-09

MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

FROM: Shalanda D. Young
Acting Director

SUBJECT: Moving the U.S. Government Toward Zero Trust Cybersecurity Principles

Tout le trafic est du trafic Zero Trust

L'approche Zero Trust a démarré comme un nouveau moyen de protéger les réseaux. À terme, elle s'est étendue au-delà des réseaux sur site, mais elle était encore principalement axée sur le trafic d'applications privées.

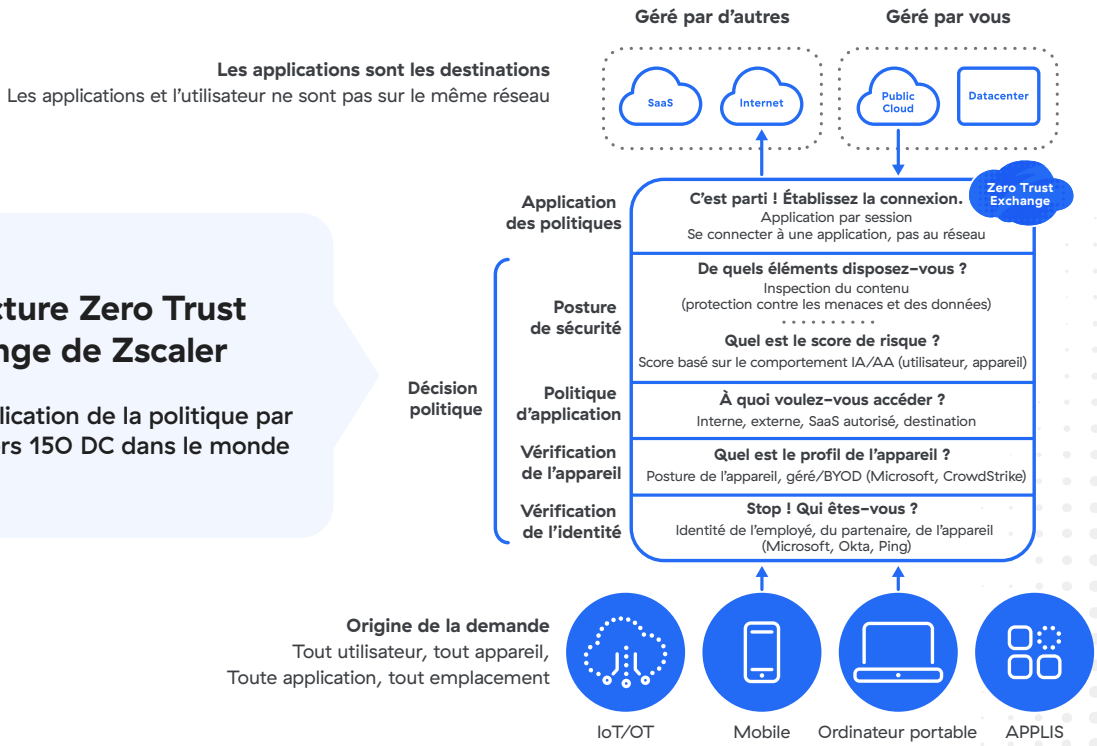
Pendant trop longtemps, le trafic a été considéré en fonction de sa relation avec un réseau, au lieu de supprimer complètement le réseau.

Mais nous savons maintenant que les principes de Zero Trust peuvent servir à protéger les applications SaaS, le trafic vers et depuis les clouds publics, et même les utilisateurs lorsqu'ils accèdent à l'internet public. Et les initiateurs de ce trafic peuvent être des charges de travail, mais également des utilisateurs. L'accès peut être indépendant du transport, le trafic passant par n'importe quel routeur et empruntant n'importe quel réseau, câblé ou sans fil, 4G ou 5G, etc.

Il est grand temps d'appliquer les principes de Zero Trust à tout le trafic, quelle que soit son origine, quelle que soit sa destination. Nous avons déjà supprimé les distinctions entre ce qui est fiable et ce qui ne l'est pas, sur le réseau ou en dehors de celui-ci. Il est maintenant temps d'arrêter de considérer quelle entité se connecte à quel réseau, et d'utiliser Zero Trust pour connecter toutes les entités directement en appliquant des politiques d'entreprise. Internet est le nouveau réseau de l'entreprise et tout le trafic est considéré comme loyal.

Architecture Zero Trust Exchange de Zscaler

Décision et application de la politique par session à travers 150 DC dans le monde



1 L'identité et le contexte passent toujours avant la connectivité

La vérification de l'identité est au cœur de la notion de Zero Trust. Mais dans le passé, nous avons confondu identité et connectivité, ce qui nous a menés à des modèles défaillants. Les adresses IP, les adresses MAC et les ports et protocoles ne sont pas des identités.

Les appareils OT peuvent se connecter aux réseaux depuis des usines. Les utilisateurs peuvent se connecter depuis des cafés. Mais cela ne signifie pas que nous savons quoi que ce soit à leur sujet. Nous devons donc commencer par l'identité et le contexte. Ce n'est qu'à partir de là que nous pouvons autoriser la connectivité.

Lorsqu'un utilisateur demande l'accès à une ressource, nous devons d'abord examiner son identité, d'autres informations le concernant, comme son rôle ou son service, l'appareil qu'il utilise, et enfin les politiques de sécurité. Qu'est-ce que l'utilisateur essaie de faire ? À quoi veut-il accéder ? Qu'est-ce qui, dans l'environnement, pourrait influencer notre décision d'autoriser ou de refuser cette action ?

Le contexte va au-delà de l'identité et est évalué en permanence. D'autres facteurs peuvent être recoupés pour détecter des anomalies, notamment la géolocalisation, l'adresse IP, la posture de l'appareil et le moment de la journée. Enfin, une solution Zero Trust doit être capable de déchiffrer le trafic, de détecter les menaces et les risques d'exfiltration de données inline et à grande échelle.

Dans le cas de Zero Trust Exchange, nous corrélons également les renseignements sur les menaces pour déterminer les risques et prendre des décisions en matière de politique et d'accès, que ces menaces proviennent de notre cloud mondial ou de partenaires technologiques tiers tels que les fournisseurs de sécurité et de vérification d'identité.

2 Les applications et, au-delà, les environnements d'applications, doivent rester invisibles aux utilisateurs non autorisés

Maintenant que nous avons résolu le problème de savoir qui vous êtes avant de vous accorder l'accès, nous pouvons nous attaquer au défi suivant : comment vous connecter à vos ressources autorisées, tout en réduisant les risques et en minimisant le potentiel de compromission ? Une fois que le contexte entourant un utilisateur, un appareil, une politique et un environnement a été recueilli et analysé, nous pouvons franchir les étapes suivantes dans cette voie.

En éliminant l'auditeur entrant pour les connexions à distance, nous éliminons la surface d'attaque externe. Faute de quoi, les attaquants trouveraient trop facilement des passerelles VPN vulnérables ou des applications exposées afin de compromettre leurs cibles. Les VPN qui attendent des connexions entrantes sont des cibles faciles, et les acteurs malveillants le savent. Il s'agit d'un problème indépendant du fournisseur, qui ne peut être résolu qu'en modifiant le modèle architectural.

Zero Trust Exchange de Zscaler établit des connexions sortantes uniquement depuis l'utilisateur et l'environnement de l'application vers notre cloud sécurisé, des micro-tunnels chiffrés servant d'intermédiaire entre les demandes et leurs destinations.

Ce « tiers-lieu » en ligne constitue un tampon entre les utilisateurs vérifiés et toute ressource à laquelle ils sont autorisés à accéder. Une fois qu'un utilisateur est connecté à la ressource souhaitée, des politiques granulaires garantissent qu'il n'a pas la possibilité de s'aventurer au-delà. Les déplacements latéraux deviennent pratiquement impossibles.

3 Le dernier chapitre ?

Les principes exposés ci-dessus nous permettent de nous défaire véritablement et définitivement d'une conception traditionnelle des périmètres de réseau protégés par des pare-feu et des terminaux distants connectés via des réseaux privés virtuels. Ils ne se contentent pas de reproduire les contrôles de sécurité existants dans une instance virtuelle hébergée dans le cloud, ni de s'appuyer sur une compréhension artificielle de ce qui se trouve sur le réseau et de ce qui ne s'y trouve pas.

Une architecture complète conçue pour assurer une sécurité Zero Trust pour les utilisateurs, les charges de travail, les applications, les appareils OT et IoT, et au-delà, réduit les risques, améliore la protection, simplifie l'expérience de l'utilisateur et représente une amélioration fondamentale de la façon dont nous envisageons la sécurité de l'entreprise.

+1 408 533 0288 Zscaler, Inc. (HQ) • 120 Holger Way • San Jose, CA 95134 zscaler.fr