White Paper

Una breve historia acerca de Zero Trust: los principales hitos del replanteo de la seguridad empresarial

¿Por qué contar la historia de Zero Trust?

Muchos en el sector de seguridad de TI creen que Zero Trust es revolucionario, un replanteo fundamental acerca de la seguridad empresarial y que protege las redes y los recursos que albergan nuestras mejores ideas, conectan a nuestros mejores talentos y otorgan acceso a herramientas de productividad transformadoras.

Pero para entender cuán revolucionario es el modelo Zero Trust en ciberseguridad, es necesario comprender las debilidades del enfoque de seguridad de red heredada y cómo la idea de la arquitectura Zero Trust ha evolucionado para reformar fundamentalmente ideas que tienen décadas de antigüedad.

Redes 2D y seguridad castle-and-moat (castillo y foso)

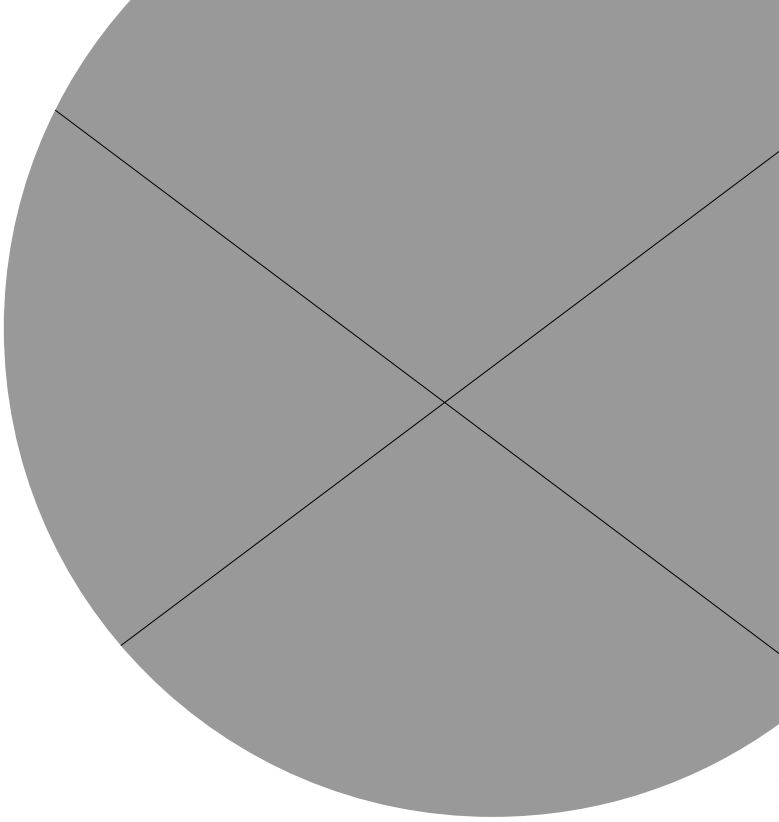
Hub-and-spoke (radial) y castle-and-moat (castillo y foso) son las dos metáforas principales utilizadas para describir la arquitectura de red heredada yla seguridad de la red, respectivamente. En ambos casos, esta manera de describirlas ha existido ya durante un tiempo.

La arquitectura de red radial se refiere a las redes satélite dispuestas alrededor de un concentrador central. Este modelo consiste en enrutar el tráfico interno y externo a través de una pila de seguridad en un centro de datos primario antes de que siga hacia su destino. Aunque este método ha funcionado durante un tiempo, se ha vuelto más complicado y costoso debido a la adopción de la nube, la distribución de las fuerzas de trabajo y la creciente importancia de la movilidad en la empresa.

Por otro lado, la seguridad de tipo castle-and-moat se refiere a redes autónomas diseñadas para admitir un tráfico confiable y, al mismo tiempo, mantener a los enemigos efectivamente fuera de sus muros. Del mismo modo que los guardias en la puerta, los dispositivos de seguridad internos están pensados para permitir el ingreso de las personas adecuadas y rechazar a los bandidos. La transición masiva de dispositivos a la nube, junto con la migración de los trabajadores fuera de los perímetros corporativos, hizo que este método quedara obsoleto más rápidamente que las balas de cañón para los castillos reales.

Las redes VPN y Wi-Fi complican aún más el problema. La antigua arquitectura castle-and-moat no permitía a los administradores conectar a los invitados a una red sin darles libertad total mientras estuvieran allí. En última instancia, no existía una buena manera de conectar los puntos finales a las redes sin cierta segmentación para mantener la seguridad de las redes.

Necesitábamos algo mejor.



802.1X y los problemas con NAC

En 2001, la Asociación de Estándares aprobados por el Instituto de Ingenieros Eléctricos y Electrónicos (IEEE, por su sigla en inglés) publicó su estándar del protocolo 802.1X para el control de acceso a la red (NAC).

> "Un medio para autenticar y autorizar dispositivos conectados a un puerto LAN que tiene características de conexión punto a punto, y para impedir el acceso a ese puerto cuando el proceso de autenticación y autorización falle."
>
> IEEE en 802.1X flecha-derecha

Poco después, los dispositivos inalámbricos empezaron a incluir un solicitante o cliente 802.1X que permitía a las redes autenticar el punto final antes de permitir una conexión. Este avance tenía por objeto ofrecer la capacidad de bloquear las redes cableadas e inalámbricas, de modo que solo los dispositivos administrados y los usuarios autorizados pudieran conectarse. Imagine que el solicitante proporciona una identificación al portero que protege la puerta de la red y decide quién puede entrar y a quién se deja afuera.

Alas, el modelo de NAC no era la solución universal, y los problemas comenzaron con esa red. Las redes internas se diseñaron con una confianza implícita y tratar de agregar la autenticación/autorización después de ello era un gran esfuerzo. Para que el NAC sea totalmente eficaz, era necesario bloquear todos los puertos accesibles, pero no todos los dispositivos eran compatibles con 802.1X. La creciente adopción de impresoras, lectores de insignias y otros dispositivos conectados a Internet y habilitados para la red era un problema de seguridad evidente. Ahora bien, imaginemos que nuestro inspector seguía cuidando esa única puerta de la red cuando existían varias (o incluso docenas) de entradas alternativas disponibles.

Derribando las paredes de Jericó y replanteando el rol del perímetro en la seguridad

En 2003, era claro que el uso de dispositivos personales seguiría proliferando, y las organizaciones debían empezar a plantearse cómo proteger las máquinas que no estaban encerradas tras los muros del castillo. Además, el aumento del uso del cifrado estaba reduciendo la eficacia de los firewalls perimetrales, lo cual forzaba la opción entre escalar para resolver los problemas de capacidad que imponía el descifrado e inspección o permitir que el tráfico cifrado pasara sin controles.

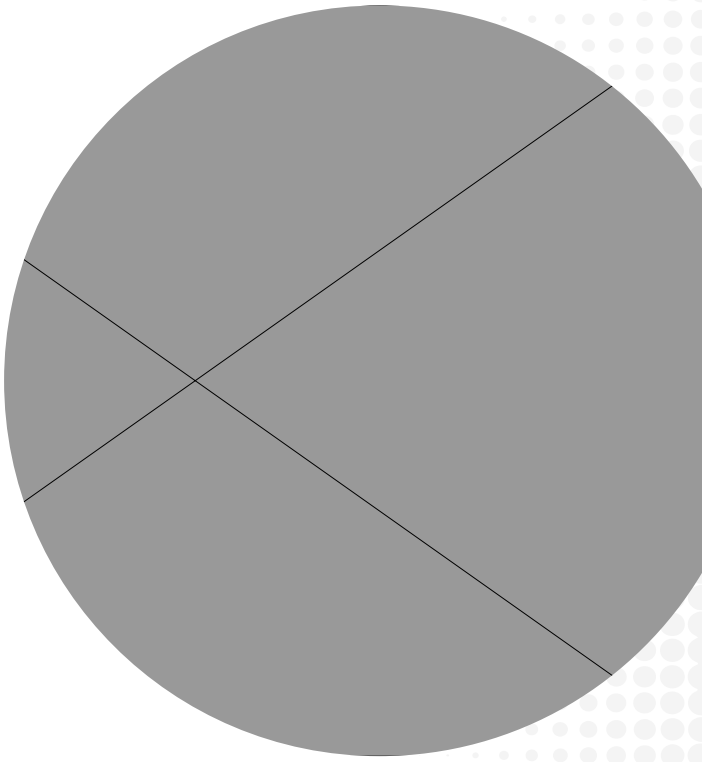
Ese año, un grupo multinacional de líderes tecnológicos europeos se reunió para abordar temas como la autenticación de usuarios, el cifrado, la administración de identidades y el cumplimiento de las políticas. Después de establecerse formalmente en 2004, el Foro Jericho le presentó al mundo la noción de "desperimetrización".

Con un nombre que recuerda la historia bíblica de los israelitas derribando las murallas de la antigua ciudad de Jericó, el foro se propuso resolver el problema de cómo "permitir flujos de información seguros y sin límites entre empresas".

Además de la acertada metáfora, el grupo dejó Los mandamientos del Foro Jericó, lo más cercano hasta el momento a las verdades desde las alturas sobre el control de las redes sin perímetro. Lamentablemente, el conjunto de controles y mitigaciones indicado estaba más allá de la capacidad de implementación o administración de la mayoría de las empresas en ese momento.

"Zero Trust" aparece en el léxico de TI por primera vez

En 2010, el analista de Forrester John Kindervag publicó un artículo titulado "No More Chewy Centers:



Introducing The Zero Trust Model Of Information Security" e inmediatamente surgió una nueva palabra de moda que representaba una nueva manera de pensar en la seguridad de la red. Una afirmación clave del artículo era que la mera presencia en una red no era suficiente para otorgar confianza.

"Aquí es donde empezamos a escuchar cosas como 'la identidad es el nuevo perímetro'", dice la directora de tecnología de campo de Zscaler y veterana de Zero Trust, Lisa Lorenzin. "Autenticamos a un usuario y utilizamos esa identidad para determinar lo que puede hacer. Tal vez, si teníamos suerte, podíamos reunir algo de contexto, como, por ejemplo, si teníamos un dispositivo administrado o no administrado, y tomar decisiones sobre el acceso basadas en esa comprensión básica."

Progreso. Pero esto dejó a la seguridad empresarial estancada en proteger las propias redes. Aún no estaba lista para abandonarlas por completo. Todavía no llegábamos a un método transformador, por lo que la adopción de estos principios volvió a fracasar. Por un lado, seguía basándose en el mismo conjunto de herramientas centradas en la red: 802.1X y RADIUS en la capa 2, firewalls con reconocimiento de identidad en la capa 3, etc.

La nueva forma era simplemente NAC con un nombre pegadizo.

Beyond(the perimeter)Corp

Mientras tanto, los hackers vinculados al Ejército de Liberación Popular (PLA) de China estaban haciendo que los mejores y más brillantes de la industria tecnológica reconsideraran por completo el problema de la confianza. En 2010, Google reveló una operación de 2009 que la había tenido por objetivo así como a otras empresas tecnológicas de alto perfil, como Akamai, Adobe y Juniper Networks. La campaña fue llamada por los investigadores de seguridad de McAfee "Operación Aurora".

Al patear el avispero de los talentos de los mejores en TI, los hackers chinos, sin saberlo, aceleraron las tareas para la arquitectura Zero Trust en los principales laboratorios tecnológicos del país. Google desarrolló BeyondCorp en respuesta a la Operación Aurora, que se centró en **"trasladar los controles de acceso del perímetro de la red a los usuarios individuales... [permitiendo] trabajar de manera segura desde prácticamente cualquier sitio sin necesidad de una VPN tradicional".**

Pero "Google es una empresa dirigida por ingenieros, para ingenieros, con un presupuesto efectivamente infinito y una infraestructura heredada pequeña en comparación con muchas empresas", dice Lorenzin. "Y aún así, les llevó siete años y seis documentos técnicos de diseño e implementación."

Incluso con el ejemplo bien documentado de Google, la verdadera arquitectura Zero Trust seguía estando fuera del alcance para la mayoría de las empresas. A pesar de que intenta "allanar el camino para que otras organizaciones realicen su propia implementación de una red Zero Trust", el futuro que Google imaginaba estaba todavía muy lejos.

Mientras tanto, para los usuarios, la popularidad de la nube y el continuo énfasis en la movilidad significaron que había más datos disponibles accesibles desde fuera del perímetro de la red que desde dentro de ella. La necesidad de un método generalizado de confianza era mayor que nunca.

Guía de mercado para el acceso a la red de Zero Trust de Gartner®

La empresa de investigación tecnológica Gartner® fue responsable de los siguientes avances significativos en Zero Trust como un marco ampliamente adaptable. Aunque aún estaba presente, el término "Zero Trust" no era la prioridad en 2010 cuando la empresa publicó su Evaluación de Riesgos y Confianza Adaptativa Continua (CARTA).

El documento describía la necesidad de comprender quién solicita acceso y conceder ese acceso en función de una evaluación dinámica del entorno, el contexto disponible y las responsabilidades de un usuario garantizadas.

Lorenzin describe CARTA como "un gran modelo que nunca tuvo la tracción que merecía".

En Gartner®, CARTA terminó por convertirse en el acceso a la red de Zero Trust (Zero Trust Network Access, ZTNA) después de que el marco original no lograra convencer a los profesionales de la tecnología (¡obsérvese que las redes siguen siendo el centro de interés como objetivo del acceso!) Pero, fundamentalmente, CARTA sigue siendo importante para la historia de Zero Trust porque los principios que estableció siguen vigentes en ZTNA.

La siguiente contribución significativa de Gartner® a este debate fue el reconocimiento de la convergencia de los campos de las redes y la seguridad. En 2019, expresó este matrimonio con la introducción del perímetro de servicio de acceso seguro (Secure Access Service Edge, SASE). Sin embargo, fue una unión de corta duración y, para 2021, una vez más estaba dividiendo las categorías con la presentación de la categoría de mercado del perímetro de servicio seguro (Secure Service Edge, SSE): SASE sin WAN.

Independientemente del nombre, Gartner® se había establecido para entonces como un árbitro importante de lo que hacía o no hacía para Zero Trust. Los proveedores ya estaban luchando por encajar en una de sus nuevas categorías de mercado.

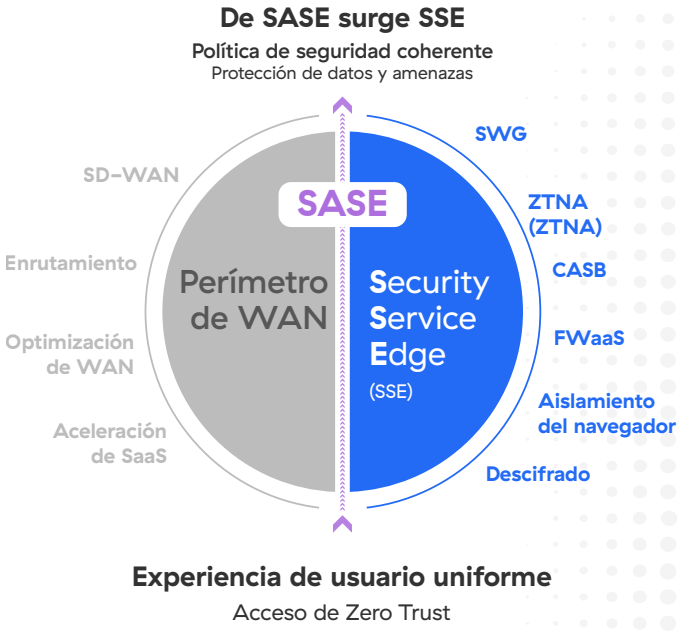
"El Hombre" ingresa al escenario: el respaldo del NIST, la OMB y el gobierno a la ZTA.

En 2020, el Instituto Nacional de Estándares y Tecnología (NIST) replanteó el diálogo con su norma NIST 800-207 para la arquitectura Zero Trust. Este nuevo paradigma de ciberseguridad se centraba en la protección de los recursos y en la premisa de que nunca debe concederse confianza de manera implícita, sino que debe evaluarse continuamente.

Con este artículo, finalmente se liberaron el perímetro y la red privada virtual. El centro de interés pasó de proteger la red a proteger a los usuarios, los datos y las aplicaciones que interactúan a través de la red. Zero Trust significa ahora simplemente un acceso basado en el contexto y con privilegios mínimos, que se utiliza con una variedad mucho más amplia de casos de uso y flujos de tráfico.

El estándar 800-207 estipula principios clave y suposiciones para Zero Trust. Los siguientes son tres de los puntos más críticos (de una lista mucho más larga):

1. Ningún recurso es inherentemente confiable.
2. Todas las comunicaciones están protegidas independientemente de la ubicación de la red. Terminando e inspeccionando la solicitud; mirando todo el contexto disponible asociado con el usuario y la solicitud.
3. Todas las autenticaciones y autorizaciones de recursos son dinámicas y se cumplen estrictamente antes de permitir el acceso.



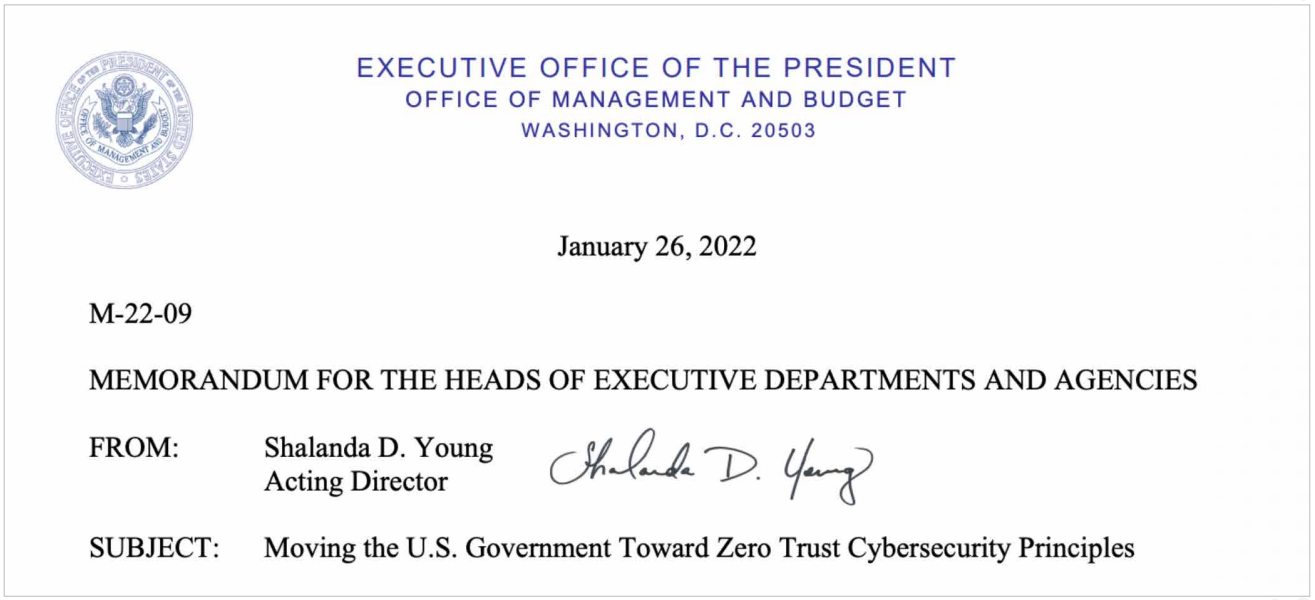
Pero el verdadero punto sin retorno para la promoción de los principios de Zero Trust vino de lo más alto, al menos en los Estados Unidos. La Oficina de Gestión y Presupuesto de los Estados Unidos, oficina responsable de la implementación de políticas presidenciales, emitió su directiva M-22-09 en 2022, en la que se afirma que todas las oficinas del gobierno federal deben adoptar los principios de la arquitectura Zero Trust para 2024 y se definen hitos y fechas objetivo claros a lo largo del camino.

"Hasta ahora, hemos tenido documentos orientativos. Hemos tenido modelos de administrador. Pero este es el primer punto en el que la teoría se pone en práctica, con la estrategia federal Zero Trust", según Lorenzin.

El ataque de la cadena de suministro contra la plataforma de gestión de TI Solar Winds, divulgado en 2021 y responsable de afectar al menos nueve organismos federales, entre ellos el Estado, el Tesoro, la Seguridad Nacional, el Comercio y la Energía, fue quizás el ataque más descarado y dañino patrocinado por un estado desde la Operación Aurora. En respuesta, el gobierno federal ha apoyado únicamente a Zero Trust, adoptando ese enfoque como su lema de ciberseguridad para los próximos años.

Implementación de Zero Trust

El método de Zscaler para la arquitectura Zero Trust concuerda perfectamente con el marco ZTA del NIST y la definición de Gartner® para SSE. Pero va más allá de cualquier norma debido a su compromiso con tres avances fundamentales en el criterio Zero Trust. Juntos, estos principios avanzados ayudan a dirigir la aplicación de Zero Trust hacia algunas conclusiones lógicas.



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

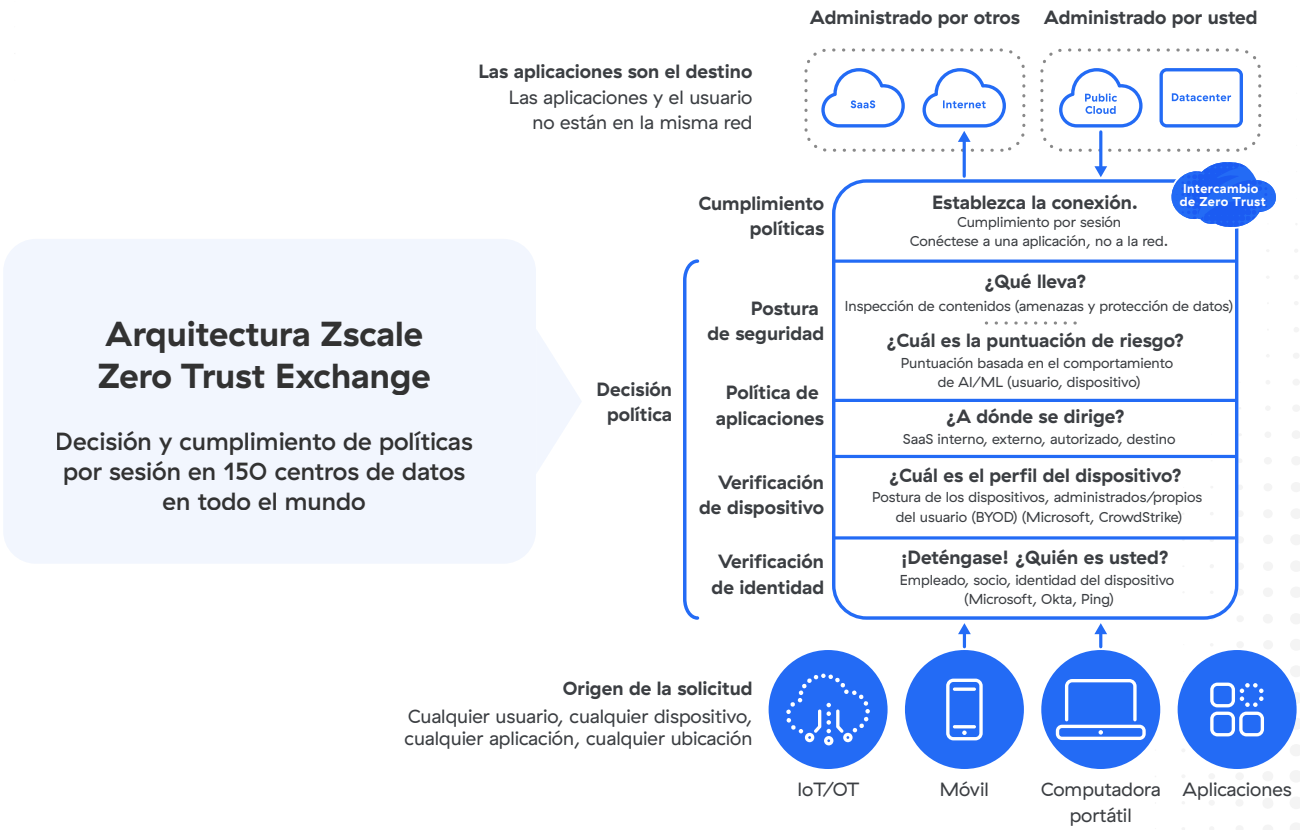
Todo el tráfico es tráfico Zero Trust

Zero Trust comenzó como una manera novedosa de proteger las redes. Eventualmente, se expandió más allá de las redes en las instalaciones, pero aún se centraba principalmente en el tráfico de aplicaciones privadas.

Por mucho tiempo se consideró el tráfico en base a su relación con una red, en lugar de desechar la red por completo.

Pero ahora sabemos que los principios de Zero Trust pueden utilizarse para proteger las aplicaciones SaaS, el tráfico hacia y desde las nubes públicas e incluso, a los usuarios cuando acceden a la Internet pública. Y los originadores de ese tráfico pueden ser tanto cargas de trabajo como usuarios. Se puede tener acceso independientemente del transporte, con tráfico que fluye a través de cualquier enrutador y que llega a través de cualquier red, cableada o inalámbrica, 4G o 5G, etc.

Ya es hora de aplicar los principios de Zero Trust en todo el tráfico, independientemente del origen y del destino. Ya hemos hecho distinciones entre confiable y no confiable, dentro o fuera de la red. Ahora es el momento de dejar de pensar en qué entidad se está conectando a qué red y, en su lugar, utilizar Zero Trust para conectar a todas las entidades directamente utilizando políticas empresariales. Internet es la nueva red corporativa y todo el tráfico es válido.



1 La identidad y el contexto siempre tienen prioridad con respecto a la conectividad

La verificación de la identidad es la base de Zero Trust. Pero en el pasado hemos confundido la identidad con la conectividad, y eso nos ha llevado a modelos defectuosos. Las direcciones IP, las direcciones MAC y los puertos y protocolos no son identidad.

Los dispositivos OT pueden conectarse a las redes desde fábricas. Los usuarios pueden conectarse desde cafeterías. Pero eso no significa que sepamos nada de ellos. Por lo cual debemos empezar con la identidad y el contexto. Solo a partir de ahí podemos autorizar la conectividad.

Cuando un usuario solicita acceso a un recurso, primero debemos tener en cuenta quién es, otra información sobre él, tal como su función o departamento, el dispositivo que utiliza y, a continuación, las políticas de seguridad. ¿Qué intenta hacer el usuario? ¿A dónde se dirige? ¿Qué elementos del entorno pueden contribuir a nuestra decisión de permitir o denegar la acción?

El contexto va más allá de la identidad y se evalúa continuamente. Otros factores que se pueden evaluar para detectar anomalías son la geolocalización, la dirección IP, la postura del dispositivo y la hora del día. Y una solución Zero Trust debería poder descifrar el tráfico para inspeccionar las amenazas y los riesgos de exfiltración de datos en línea y a escala.

En el caso de Zero Trust Exchange, también correlacionamos la información sobre amenazas, de toda nuestra nube global, así como de socios tecnológicos externos como proveedores de seguridad y de verificación de identidad, para determinar el riesgo y tomar decisiones de política y acceso.

2 Las aplicaciones e incluso, los entornos de las aplicaciones, deben permanecer invisibles para los usuarios no autorizados

Ahora que hemos resuelto el problema de saber quién es antes de concederle acceso, podemos abordar el siguiente problema: ¿cómo conectar al usuario a los recursos autorizados, al tiempo que reducimos el riesgo y minimizamos la posibilidad de compromiso? Una vez que se ha reunido y analizado el contexto que rodea al usuario, el dispositivo, la política y el entorno, podemos dar los pasos siguientes en esa dirección.

Al eliminar a la unidad de escucha entrante que busca conexiones remotas, eliminamos la superficie de ataque externa. De lo contrario, es demasiado fácil para los atacantes ubicar puertas de enlace VPN vulnerables o aplicaciones expuestas para afectar los objetivos. Las VPN que están a la espera de conexiones entrantes son un blanco fácil y los atacantes se dan cuenta. Este es un problema independiente del proveedor que solo se puede resolver si se cambia el modelo arquitectónico.

Zscaler Zero Trust Exchange lo logra estableciendo conexiones hacia el exterior solamente, tanto del usuario como del entorno de la aplicación a nuestra nube de seguridad mediante microtúneles cifrados para establecer conexiones con agentes intermediarios entre las solicitudes y sus destinos.

Este "tercer lugar" en línea proporciona un amortiguador entre los usuarios verificados y cualquier recurso al que estén autorizados a acceder. Una vez que un usuario está conectado al activo solicitado, las políticas granulares garantizan que no tenga la opción de aventurarse más allá de él. El movimiento lateral se vuelve básicamente imposible.

3 ¿El capítulo final?

Los principios expuestos anteriormente nos permiten superar por fin la concepción heredada de perímetros de red protegidos por firewalls y puntos finales remotos conectados a través de redes privadas virtuales. No se limitan a replicar los controles de seguridad existentes en una instancia virtual alojada en la nube, ni se basan en una comprensión artificial de lo que está en la red y lo que no está.

Una arquitectura integral diseñada para brindar seguridad Zero Trust, para usuarios, cargas de trabajo, aplicaciones, dispositivos OT e IoT y mucho más, reduce el riesgo, mejora la protección, simplifica la experiencia del usuario y representa una mejora fundamental en nuestra manera de pensar acerca de la seguridad empresarial.

+1 408.533.0288 Zscaler, Inc. (HQ) • 120 Holger Way • San Jose, CA 95134 zscaler.es