Global Information Assurance Certification Paper

Secure Internet Gateways:

Backing Down from a Fight

GIAC (GCIA) Gold Certification

Author: Seth Polley, readingroom@nightvisionsecurity.org

Advisor: Chris Walker

Accepted: October, 2018

Abstract

When does a security agent become a double agent? On-premise corporate devices are protected by a stack of security products, whereas remote clients have traditionally relied on DNS, Proxy, and/or VPN solutions to obtain the same levels of protections. These remote clients typically utilize lightweight agents that run on the devices with the intention of enforcing security and/or policy-based protections - no matter which offsite networks the corporate device connects to. Frequently though, there are provisions to deactivate the agent when the computer connects to the local network. As the security professional charged with protecting your company's remote assets - do you know the extent of these 'back-off' scenarios? This Gold Paper will discuss some of the commonly known scenarios, but will delve further into the unknowns which may surprise you.

Introduction

The Internet has become a crucial cornerstone to the way organizations operate today. According to a 2013 Gartner report, "By 2018, 25% of corporate data traffic will bypass perimeter security and flow directly from mobile devices to the cloud." (Wagner, et al., 2013). An updated 2016 Gartner report and 2017 Gartner Security Summit figures indicated that "By 2021, 27% of corporate data traffic will bypass perimeter security, up from 10% today." (Moore, 2016). As corporate data becomes more heavily integrated in the cloud, so too does this become an increasing target for cyber criminals. On-premise users are protected by layered stacks of security products, while remote workers have traditionally used VPN solutions to get the same level of protections when off of the organization's network. As an organization's network begins to extend beyond the perimeter, so too must the organization's security.

1.1. Secure Web Gateways (SWGs)

As Internet access became more prevalent and bandwidth was very limited yet, web proxy solutions rose to control the web traffic and content available to end-users. The Secure Web Gateway (SWG), often integrated into a proxy architecture and alternatively known as a web proxy solution, helped screen unsecured user-initiated Web traffic through URL filtering. Reacting to the growing concerns over unwanted software, malicious websites, and hosted malware code, these gateways began to incorporate detections aimed at filtering out this additional undesirable content, further enforcing corporate and regulatory compliance. Preventing the filtered traffic from coming into the

Seth Polley,
readingroom@nightvisionsecurity.org

internal networks, organizations had a method to aid in maintaining productivity while protecting their users from infection.

A Secure Web Gateway is typically implemented to secure an organization against threats originating from the Internet, sourced from websites and other Web 2.0 products or services. These solutions were often designed as a hardware or software gateway device implemented at the outer boundaries of a network. With the rise of a mobile workforce, personnel frequently connecting to the Internet off-premise and away from the corporate network's protections, these archaic implementations were no longer sufficient.

1.2. Secure Internet Gateways (SIGs)

Software as a Service (SaaS) applications, sometimes called web-based or hosted software, are rising in popularity to fill these gaps. These applications run on the SaaS providers' servers where organizations access them via the Internet, rather than requiring the traditional methods of installation and maintenance on-premise. As networks grow decentralized, the user base shifting to a mobile workforce, the security solutions have also begun the same shift into the cloud. The SWG market has been moving strongly toward service-based SaaS delivery models, growing at more than 30% for the past five years (Gartner, Inc., 2017). An evolution of the Secure Web Gateway, the Next Generation Secure Web Gateway or Secure Internet Gateway (SIG), is emerging to address these demands and changes. Without an 'always on' VPN solution, off-network

Seth Polley,
readingroom@nightvisionsecurity.org

and off-VPN protection is required for the security of the mobile workforce. The SIGs are cloud security platforms which are meant to secure and protect an organization's user base in the manners they work today (and will increasingly work in the future). The SIGs function off-premise from the cloud, largely moving away from the traditional VPN protections which tunnel users' traffic back on-premise through inefficient backhaul routing, and scales more efficiently to secure users on and off network. By delivering these elements of the security stack as a cloud service, an organization not only eliminates the cost and complexity of traditional secure web gateway appliances, but it also provides a faster experience for the user through the globally distributed cloud infrastructure (Zscaler, Inc., 2017).

1.3.1 Gartner Magic Quadrant Leaders

Gartner is considered by many to be the leading research and advisory company for insight into the product and service offerings of the top technology providers. One popular research publication is the Gartner Magic Quadrant. The "Gartner Magic Quadrant research methodology provides a graphical competitive positioning of four types of technology providers in fast-growing markets: Leaders, Visionaries, Niche Players and Challengers." (Gartner, Inc., 2018). The three Magic Quadrant Leaders in the Secure Internet Gateway space are Symantec, Zscaler, and Cisco (Orans & Firstbrook, 2017):

- Symantec Secure Web Gateway
- Zscaler Internet Access

Seth Polley,
readingroom@nightvisionsecurity.org

- Cisco Umbrella

2. Configurations

There are four main types of configurations commonly seen within the major SIGs and any mix of the four may be implemented within an organization - network-based, site-based, client-based agent, or client browser-based (proxy) configurations. Policies dictate the security protection, category settings, and individual destination lists you can apply to some or all of your identities. The logs are often viewable (to some degree) through the web portal and can be configured for direct ingest into a Security Information and Event Management (SIEM) system or for storage in a cloud bucket/folder.

2.1. Network-based Configurations

A network-based configuration is defined by the public IP space of the organization itself. The Wide Area Network (WAN) traffic originating within the defined IP space(s) egresses an edge device, often times a router or firewall, and is tunneled or forwarded to the Secure Internet Gateway's services. These configurations allow an organization to extend its protections to any device that is connected to the Internet through the defined networks' IP space. A network configuration may be anywhere from a single public IP address (static or dynamic) to a range of public IP addresses defined through Classless Inter-Domain Routing (CIDR) notation. CIDR

Seth Polley,
readingroom@nightvisionsecurity.org

notation is a method that allows for more flexible allocation of IP addresses through a compact representation indicating the size of these networks. Once one or more networks have been added, creating a global identity scope for an organization, the policies can be defined and enforced.

2.2. Site-based Configurations

Site-based configurations allow one or more fixed locations to establish a secure connection to the SIG. The fixed locations tend to follow data center and/or geographic office location. Through the use of GRE or IPSec VPN tunnels, a secure connection can be made from an internal router behind the firewall to an inline SIG that inspect all Internet traffic bi-directionally. By configuring a primary tunnel at one data center and a secondary tunnel at a secondary, these afford an organization failover protections or granular configuration of security policies.

2.3. Client-based Configurations

Client-based configurations, as the name may suggest, focus the application of configurations around an agent installed on the mobile device. Using traditional VPN configurations, the clients may be installed and active, but remain in a disconnected state until the user initiates the secure connection. The SIG clients change this paradigm, removing the dependence on user action to enforce protects, and automatically actions a connected or disconnected state based on pre-defined modes. A few examples of modes a SIG client may operate under are:

- Always On: The client is configured to be 'Always On', regardless of network it is utilized on. In this mode, the policies and configurations of the client are always enforced, whether on the trusted organization's network or that of a third-party entity.
- Selectively On: Similar to the Always On configuration, the client is configured to be active for all non-corporate services and provides selective VPN tunneling connections for chosen corporate resources.
- Deferred: The client is active and communicating with the SIG service, but the network- and/or site-based configurations are utilized for policy enforcement and reporting purposes. Here, the network policies are given a higher precedency than those defined by the local client's configurations, when evaluated before the policy defined for the roaming clients.
- Disabled: The client is automatically disabled, or 'backs off', when it is connected to a protected network that meets pre-configured criteria. As with the Deferred mode, the network- and/or site-based configurations are utilized.

Alternatively, client configurations may specify the use of Proxy Auto-Config (PAC) files within the browser for a proxy-based approach that can only analyze HTTP, HTTPS, FTP and SOCKS traffic. The PAC file directs the browser to forward its traffic to the enforcement proxy for inspect and actioning. A proxy-based approach may introduce greater latency by forwarding through the middleman service and attempts at caching may affect downstream authentication, policy enforcement, and/or reporting. As with the

Seth Polley,
readingroom@nightvisionsecurity.org

other configurations though, this PAC file and proxy approach may be paired with network-based configurations or a client application.

3. Client Behavior Analysis

While security may never stop 100% of the threats, it must work 100% of the time (Cisco, 2017). Client-based protections may claim to be enforced the entire time the mobile device is off the corporate or trusted network, no matter of destination, but this claim is not always true. Unlike VPN clients, where the user initiates the connection before working or to access internal/sensitive resources, the client-based protections are designed to always protect the device no matter what foreign network it may connect to. Compared to antivirus engines on the local device which are 'always on' (unless there is user and/or administrator intervention to manually deactivate the protections), these claims may seem reasonable. To understand the erratic agent behaviors though, one must first accept that the agent may not be 'always on', and that third-party actors or networks can disable the client-based protections without the awareness of the organization and/or user. Agents, such as the SIG or VPN clients, commonly have 'back-off' scenarios which lead the client to disable protections in favor of the native protections offered on the corporate networks. These bolstered native protections may include firewalls, proxies, Intrusion Detection Systems (IDS) or Intrusion Prevention Systems (IPS), antivirus, DLP, SIEM alerting, and more. The known scenarios which make the agent stand down may application-based, domain name-based, or even location-based. The following scenarios will discuss the lesser known or unknown cases one may not immediately consider being a risk factor.

Seth Polley,
readingroom@nightvisionsecurity.org

3.1. When agents are misdirected.

Considering again, the first option discussed for SIG configuration, network-based configurations can be utilized to define the public IP space of a corporate. When multiple organizations utilize network-based configurations, they will all forward the network traffic to one or more SIG service termination points (often geographically oriented). Now, consider a 'back-off' scenario where a client deactivates its protections when on a network configured to forward traffic to the SIG. Company #1's mobile device, a laptop equipped with a SIG client, backs off when connected to Company #1's network space which forwards traffic to the SIG. Knowingly or unknowingly, Company #1's laptop is placed on Company #2's configured network space. Company #2 also utilizes the same SIG service. The protections assumed to be 'always on', protecting Company #1's assets, now deactivate on Company #2's network. Network-based configurations on a foreign network have now misdirected the agent from pursuing its original mission.

The problems:

The SIG agent will disable itself, deferring to the network's local Domain Name System (DNS) server for DNS resolution. Thus, the corporately defined content and security blocks of Company #1 may no longer be enforced, as these decisions are now inherited based on Company #2's local network settings. All policy enforcement dictated for the mobile client will no longer apply.

Seth Polley,
readingroom@nightvisionsecurity.org

Furthermore, with the foreign network's local DNS server being utilized, the foreign entity may now gain at least elementary visibility into Company #1's corporate traffic. Unless there is a local device mechanism recording and saving or forwarding the DNS traffic too, which is not a standard dataset logged locally on most user devices, the original organization has now lost visibility into their user's DNS resolution (and therefore, web traffic which may be indicative of compromise from access of phishing or otherwise malicious domains).

The solution:

If vulnerable, the solution will vary by SIG service. Some SIGs provide documented or undocumented configurations which can be implemented to prevent the agent from disabling on any network configured to forward traffic to the its service. By ignoring this particular 'back-off' mechanism, the device agent will then remain active on both the original organization's network AND any foreign networks which may utilize similar SIG services.

3.2. When agents become double-agents.

When the SIG client operates within a Deferred state, the client remains active, but gives higher precedency to the configured network policies when they are enforced earlier than the policies of the roaming client. If the trusted network's connection is broken, then the configured roaming policy or default policy (if no roaming policy is configured) takes effect. A device configuration such as this is convenient for users who

Seth Polley,
readingroom@nightvisionsecurity.org

work out of the office most of the time, allowing for different on and off network policy configurations. However, if the organization's device is connected to a foreign network which utilizes the same SIG services, foreign network-based policies may take precedence over those defined for the local client if they have a higher enforcement priority.

The problems:

If the client's policies are deferred to those of a foreign network's which have taken higher precedence, the corporately defined content and security blocks may no longer be enforced when one of two scenarios play out. Assuming strong corporate client configurations, the door is left open for accidental or malicious weaker foreign network configurations to trump the desired, stricter policies. Conversely, supposing a weaker set of policies have been applied to the client, there would be a lowered risk through the application of foreign network configurations with stricter allowances – at the expense of an unexpected or worse user experience. As with the first scenario posed, the client's logging and reporting tasks are deferred to the foreign network too, leaving the original organization with visibility gaps and potentially exposing DNS queries to a foreign entity.

The solution:

Seth Polley,
readingroom@nightvisionsecurity.org

Ideally, the SIG service would have native logic being applied which allows the selection and enforcement of the most restrictive policy. There are two difficulties posed by this solution – how would the most restrictive policy be measured on the fly or how would the two be merged appropriately without a deadlock scenario? Without native support for this logic problem, an organization can configure the roaming client policy at the highest precedence level, preventing any network policy from dictating preferred configurations. By doing so, the choice of policy enforcement is returned to the original organization.

3.3. When agents are defeated.

One of the foundational steps to most web browsing is performing DNS lookups – or the resolution of human readable domain names (such as www.google.com) to the computer routable Internet Protocol (IP) addresses (such as 216.58.216.4) that allow browsers to load Internet resources. Think for a moment how this can be tampered with, as a malicious attacker might attempt on a foreign network he/she controls. The first thought may be to block the ports/services on 53 and 443 destined for the SIG service that performs the DNS lookups. When the client-based agents attempt to resolve a domain name but cannot communicate over the blocked port(s) and service(s) or to the destination(s) it expects, the agent may fail open to favor usability and the user experience.

The problem:

Seth Polley,
readingroom@nightvisionsecurity.org

Unlike the first scenario in which the SIG agent completely disables itself, all or part of the agent's policy enforcement may be deferred to local configurations. When the ports or services performing the DNS lookups are prevented, the agent may defer to the local DNS servers, in order to avoid any loss of DNS resolution – sidestepping any potential negative impact to usability, therefore giving leeway to favor user experience over security concerns. Along with the aforementioned concerns where content and/or security blocks may no longer be enforced, and loss of logging incurred, there are further risks to an organization's security stance. By deference to the local DNS servers, interception becomes possible, and opens the door to falsification of the records - redirection or rewriting of the expected DNS requests to that of an attacker's choosing.

The solution:

When the SIG agents do not enforce policy, there is a resulting gap in protections. Holding a mindset that a security product must work 100% of the time may come down to fail open or fail closed decisions. While some SIG services provide configuration options allowing their customers to determine whether the agent fails open or closed, others are VERY averse to any fail closed scenario. The solutions available will be highly dependent on the chosen SIG service.

Seth Polley,
readingroom@nightvisionsecurity.org

4. Conclusion

When corporate data resided solely in on-premise data centers, the focal point for organizations was rightfully that of perimeter security. Shifting the balance as corporate data traffic bypasses the perimeter security, flowing directly to the cloud, has resulted in the traditional security models becoming broken. The investment in on-premise security appliances is being reconsidered by many organizations, as visibility into corporate data traffic is slowly lost and the prevalence of breaches continue to plague the information security realm.

Secure Internet Gateways are being considered as a replacement for some or all of the on-premise security architecture. Organizations compare and evaluate the products that will offer the greatest flexibility while maintaining a strong security stance. The said organizations may be unaware of SIG client behaviors that will undermine the reliability of these security technologies and their enforcement mechanisms. Though the configurations available may vary by SIG solution, the expectation is that the solution will work 100% of the time. The aforementioned scenarios are some examples of cases where this claim is not always true.

The bottom line is that Secure Internet Gateway solutions are not 'always-on' despite the provider's claims, an organization's attempts to follow industry best practices, or implementation of locked down policies. As the security professional tasked with handling the risk evaluation process or one charged with protecting the organization's

Seth Polley,
readingroom@nightvisionsecurity.org

remote assets, an awareness of potential 'back-off' scenarios can help more appropriately vet the available solutions and the organization's mobile workforce exposure.

Seth Polley,
readingroom@nightvisionsecurity.org

References

Cisco. (2017). Cisco Umbrella: Roaming Package. Retrieved August 25, 2018, from https://www.cisco.com/c/dam/en/us/products/collateral/security/firewalls/umbrella-roaming-customer-facing.pdf

Cisco. (2017). The Rise of the Secure Internet Gateway. Retrieved August 25, 2018.

Gartner, Inc. (2017, September 12). Competitive Landscape: Secure Web Gateways. Retrieved August 25, 2018.

Gartner, Inc. (2018). Gartner Magic Quadrant & Critical Capabilities. Retrieved August 25, 2018, from https://www.gartner.com/en/research/magic-quadrant

Moore, S. (2016, August 11). 5 Steps to Closing SaaS Security Gaps. Retrieved August 25, 2018, from https://www.gartner.com/smarterwithgartner/five-steps-to-closing-saas-security-gaps-2/

OpenDNS, Inc. (2014). Out of the Box and Into the Cloud: Comparing Network Security Delivery Platforms. Retrieved August 25, 2018, from http://info.umbrella.com/rs/opendns/images/WP-Appliances-vs-The-Cloud.pdf

Seth Polley,
readingroom@nightvisionsecurity.org

Orans, L., & Firstbrook, P. (2017, June 12). Magic Quadrant for Secure Web Gateways. Retrieved August 25, 2018.

Wagner, R., Kavanagh, K., Nicolett, M., Chuvakin, A., Walls, A., Feiman, J., . . . Keene, I. (2013, November 25). Predicts 2014: Infrastructure Protection. Retrieved August 25, 2018, from https://www.gartner.com/doc/2629230/predicts--infrastructure-protection

Zscaler, Inc. (2017). Zscaler Web Security. Retrieved August 25, 2018, from https://www.zscaler.com/resources/solution-briefs/swg-web-security.pdf

Seth Polley,
readingroom@nightvisionsecurity.org