Defining the Next-Generation Firewall

Gartner RAS Core Research Note G00171540, John Pescatore, Greg Young, 12 October 2009, R3210 04102010

Firewalls need to evolve to be more proactive in blocking new threats, such as botnets and targeted attacks. Enterprises need to update their network firewall and intrusion prevention capabilities to protect business systems as attacks get more sophisticated.

Key Findings

- The stateful protocol filtering and limited application awareness offered by first-generation firewalls are not effective in dealing with current and emerging threats.
- Using separate firewalls and intrusion prevention appliances results in higher operational costs and no increase in security over an optimized combined platform.
- Next-generation firewalls (NGFWs) are emerging that can detect application-specific attacks and enforce application-specific granular security policy, both inbound and outbound.
- NGFWs will be most effective when working in conjunction with other layers of security controls.

Recommendations

- If you have not yet deployed network intrusion prevention, require NGFW capabilities of all vendors at your next firewall refresh point.
- If you have deployed both network firewalls and network intrusion prevention, synchronize the refresh cycle for both technologies and migrate to NGFW capabilities.
- If you use managed perimeter security services, look to move up to managed NGFW services at the next contract renewal.

WHAT YOU NEED TO KNOW

An NGFW is a wire-speed integrated network platform that performs deep inspection of traffic and blocking of attacks. There are products today with NGFW characteristics, but these must not be confused with well-marketed first-generation firewalls or products more appropriate for small businesses (see Note 1).

ANALYSIS

Changing business processes, the technology that enterprises deploy, and threats are driving new requirements for network security. Increasing bandwidth demands and new application architectures (such as Web 2.0) are changing how protocols are used and how data is transferred. Threats are focusing on getting vulnerable users to install targeted malicious executables that attempt to avoid detection. Simply enforcing proper protocol use on standard ports and stopping attacks looking for unpatched servers are no longer of sufficient value in this environment. To meet these challenges, firewalls need to evolve into what Gartner has been calling "next-generation firewalls." If firewall vendors do not make these changes, enterprises will demand price concessions to reduce first-generation firewall costs substantially and look at other security solutions to deal with the new threat environment.

What Is a Next-Generation Firewall?

To meet the current and coming generation of network security threats, Gartner believes firewalls need to evolve yet again to what we have been calling "next-generation firewalls". For example, threats using botnet delivery methods have largely been invisible to first-generation firewalls. As service-oriented architectures and Web 2.0 grow in use, more communication is going through fewer ports (such as HTTP and HTTPS) and via fewer protocols, meaning port/protocol-based policy has become less relevant and less effective. Deep packet inspection intrusion prevention systems (IPSs) do inspect for known attack methods against operating systems and software that are missing patches, but cannot effectively identify and block the misuse of applications, let alone specific features within applications. Gartner has long used the term "next-generation firewall" to describe the next stage of evolution to deal with these issues.

Gartner defines a network firewall as an in-line security control that implements network security policy between networks of different trust levels in real time. Gartner uses the term "next-generation firewall" to indicate the necessary evolution of a firewall to deal with changes in both the way business processes use IT and the ways attacks try to compromise business systems. As a minimum, an NGFW will have the following attributes:

- Support in-line bump-in-the-wire configuration without disrupting network operations.

- Act as a platform for network traffic inspection and network security policy enforcement, with the following minimum features:

- Standard first-generation firewall capabilities: Use packet filtering, network-address translation (NAT), stateful protocol inspection, VPN capabilities and so on.

- Integrated rather than merely colocated network intrusion prevention: Support vulnerability-facing signatures and threat-facing signatures. The IPS interaction with the firewall should be greater than the sum of the parts, such as providing a suggested firewall rule to block an address that is continually loading the IPS with bad traffic. This exemplifies that, in the NGFW, it is the firewall correlates rather than the operator having to derive and implement solutions across consoles. Having high quality in the integrated IPS engine and signatures is a primary characteristic. Integration can include features such as providing suggested blocking at the firewall based on IPS inspection of sites only providing malware.

- Application awareness and full stack visibility: Identify applications and enforce network security policy at the application layer independent of port and protocol versus only ports, protocols and services. Examples include the ability to allow Skype use but disable file sharing within Skype or to always block GoToMyPC.

- Extrafirewall intelligence: Bring information from sources outside the firewall to make improved blocking decisions, or have an optimized blocking rule base. Examples include using directory integration to tie blocking to user identity, or having blacklists and whitelists of addresses.

- Support upgrade paths for integration of new information feeds and new techniques to address future threats.

Examples of enforcement by an NGFW include blocking or alerting on fine-grained network security policy violations, such as the use of Web mail, anonymizers, peer-to-peer or PC remote control. Simply blocking access to known sources of these services by destination IP addresses is not enough. Policy granularity requires the blocking of only some types of application communication to an otherwise permissible destination, and redirectors make a definitive blacklist impossible to achieve. This means that there are many undesirable applications that an NGFW can identify and block even when they are designed to be evasive or are encrypted with SSL. An additional benefit of application identification can be bandwidth control, since removing, for example, undesired peer-to-peer traffic can greatly reduce the bandwidth usage.

What Is an NGFW Not?

There are network-based security product spaces that are adjacent to NGFW but not equivalent:

- **Small or midsize business (SMB) multifunction firewalls or unified threat management (UTM) devices:** These are single appliances that host multiple security functions. While they invariably include first-generation firewall and IPS functions, they do not provide the application awareness functions and are not generally integrated, single-engine products. They are appropriate for cost saving in branch offices and for use by smaller companies, but they do not meet the needs of larger enterprises. This category of exclusion includes first-generation firewalls paired with low-quality IPS, and/or having deep inspection and application control features merely colocated in the appliance rather than a tight integration, which is greater than the sum of the parts.

- **Network-based data loss prevention (DLP) appliances:** These perform deep packet inspection of network traffic, but focus on detecting if previously identified types of data are transiting the inspection point. They implement data security policy with no real-time requirement, not wire-speed network security policy.
- **Secure Web gateways (SWGs):** These focus on enforcing outbound user access control and inbound malware prevention during HTTP browsing over the Internet, through integrated URL filtering and through Web antivirus. They implement more user-centric Web security policy, not network security policy, on an "any source to any destination using any protocol" basis.
- **Messaging security gateways:** These focus on latency-tolerant outbound content policy enforcement and inbound mail anti-spam and anti-malware enforcement. They do not implement wire-speed network security policy.

While these products may be network-based and use similar technology, they implement security policies that are the responsibility and authority of different operational groups within most businesses. Gartner believes these areas will not converge before IT and security organizational responsibilities have radically changed.

An NGFW is also not an "identity firewall" or an identity-based access control mechanism. In most environments, the network security organization has neither the responsibility nor the authority for enforcing user-based access control policies at the application level. Gartner believes that NGFWs will be able to incorporate user identity information at the group level (that is, shadowing Active Directory) to make better network security decisions, but they will not be routinely used for enforcing granular user-level enforcement decisions.

NGFW Adoption

Large enterprises will replace existing firewalls with NGFWs as natural firewall and IPS refresh cycles occur or as increased bandwidth demands or successful attacks drive upgrades to firewalls. Today, there are a few firewall and IPS vendors that have advanced their products to provide application awareness and some NGFW features, and there are some startup companies that are focused on NGFW capabilities. Gartner believes that changing threat conditions and changing business and IT processes will drive network security managers to look for NGFW capabilities at their next firewall/IPS refresh cycle. The key to successful market penetration by NGFW vendors will be to demonstrate first-generation firewall and IPS features that match current first-generation capabilities while including NGFW capabilities at the same or only slightly higher price points.

Gartner believes that less than 1% of Internet connections today are secured using NGFWs. We believe that by year-end 2014 this will rise to 35% of the installed base, with 60% of new purchases being NGFWs.

Note 1

First-Generation Firewalls

First-generation firewalls came about when connecting trusted internal systems to the Internet resulted in the rapid and disastrous compromise of vulnerable internal systems, as evidenced by the impact of the Morris worm in 1988. Their use evolved to include implementing security separation of internal network segments at different trust levels as well, such as DMZ layers in an extranet or in data center zones. A network firewall can be implemented in a wide range of form factors, but it must always operate at network speeds and, at a minimum, cause no disruption to normal operation of the network.

Standard network security policy consists of two parts:

- **Block all that is not explicitly allowed:** Early firewalls blocked connections at the source/destination IP address level and then evolved to do so at the port and protocol level. As firewalls matured, this enforcement of proper protocol state became mainstream. More recently, advanced firewalls have developed the capability to recognize and block connections:
- At the application level
- Based on characteristics of the source address associated through external information sources (such as geolocation, known sources for malware, or which user is connecting)
- **Inspect what is allowed to detect and block attacks and misuse:** In the early years of firewalls, proxy-based firewalls performed more detailed inspection of the traffic allowed to pass through the firewall and attempted to detect and block malicious actions. However, early proxy firewalls were software-based and did not have the horsepower to keep up with the increasing speed of networks or the increasing complexity of applications and attacks, and the increase in new applications outstripped the ability to create new application-specific proxies. IPSs based on purpose-built appliances, to perform deep packet inspection, have evolved as the primary network security control implementing this function.