Using Palo Alto Networks to Protect the Datacenter

July 2009
# Table of Contents

**Introduction** 3

**Granular Policy Control** 3
- Isolating the DC With Network Segmentation 3
- Controlling Application Access 4
- User-based Access Control With Active Directory 5
- Monitoring and Inspecting the Content 5

**Datacenter Protection Without Performance Degradation** 6
- Single Pass Software 6
- Parallel Processing Hardware 7

**Granular Control + Performance = Datacenter Protection** 8
INTRODUCTION

The Palo Alto Networks next-generation firewall restores visibility and control over applications, users and content with three unique identification technologies; App-ID, User-ID and Content-ID. Using a Palo Alto Networks next-generation firewall enables IT to first determine which of the more than 850 applications are traversing the network (App-ID) and then tie security policies to specific users and groups in Active Directory (User-ID). For those applications that are allowed, traffic can be scanned for threats (Content-ID).

When deployed at the Internet gateway, the types of applications that are commonly found include P2P, external proxies, webmail, instant messaging, social networking and media. Mixed with this wide range of end-user applications are some business applications such as Exchange, Lotus Notes, and SharePoint. Clearly when deployed at the gateway, the return on investment comes from the ability to see the wide range of end-user applications and either positively enable and inspect them or block them, depending on the business requirement. As the deployment location moves away from the gateway towards the datacenter, the value proposition changes from breadth of control to granularity of control, threat prevention and low latency, high performance throughput.

- Granular control means locking down the datacenter first by isolating it physically or logically into a security zone and then defining specific business applications and specific sets of users that can access the datacenter.

- The datacenter represents the heart and soul of the enterprise, acting as the location that touches all business applications and transactions, as such, it is targeted by attackers and must be protected from vulnerabilities (inbound) and unauthorized data transfer (outbound).

- Performance takes into account the business focus of the deployment location where any delay in traffic means potential lost revenue.

Palo Alto Networks’ next-generation firewall can isolate datacenter environments through security policies that are based on a combination of application and user or group identity from within Active Directory. The user-based policy can also include detection and prevention of threats and unauthorized data transfer. This level of granular control is unmatched by any firewall solution on the market.

GRANULAR POLICY CONTROL

There are several aspects to applying more granular policy controls to the datacenter (DC). The first is the process of isolating the datacenter using segmentation in order to limit the avenues of access. The next aspect is to apply policies that control traffic and user access while inspecting traffic bi-directionally—but without hindering the business.

ISOLATING THE DC WITH NETWORK SEGMENTATION

While there are many different ways to segment a network, Palo Alto Networks next-generation firewalls bring a unique combination of hardware and software related segmentation capabilities that enable customers to isolate key sections of their network such as a datacenter. Every Palo Alto Networks firewall supports security zones, which, for purposes of datacenter isolation are equivalent to network segments. A security zone is a logical container for physical interface(s), VLANs, a range of IP addresses or a combination thereof. Interfaces that are added to each security zone can be configured in layer 2, layer 3 or mixed mode thereby enabling deployment in a wide range of network environments without requiring network topology modifications.
Security zones can first be applied to isolate the DC as a means of protecting the data. Once the network has been divided into distinct zones, positive control model security policies can be applied that control, at a very granular level, which applications, users and content are allowed in and out of the DC security zone. The unique capability that Palo Alto Networks provides is the ability to allow or deny access to a particular zone based on the specific application and the specific user or group from within Active Directory (as opposed to solely by IP address).

From a hardware platform and performance perspective, the combination of 10 Gbps firewall performance and high interface density [up to (24) 1 Gbps interfaces] means that a single firewall can be used to physically separate the network into distinct zones and secure them without creating a performance bottleneck.

CONTROLLING APPLICATION ACCESS

Palo Alto Networks is the only firewall on the market that uses a patent-pending technology called App-ID™ to identify and control more than 850 applications, irrespective of port, protocol, SSL encryption or evasive tactic employed.

The traffic classification by App-ID is done inline (not proxied) using four different techniques (decoders, decryption, signatures and heuristics) to determine the application identity which is then used as the basis for all policy decisions including appropriate usage, content inspection, logging and reporting. From a DC security perspective, using the exact identity of the application (e.g., Oracle, Sybase, SAP, MS SQL) as the basis for the security policy means IT has far greater control over traffic than broad-based terms such as IP address range, port and protocol might provide.

The list of applications that App-ID detects is growing steadily with 3-5 new applications added weekly based on input from customers, partners and market trends. In the event that there are custom applications within the datacenter, customers can apply an application override to effectively rename the application for visibility and control purposes. For additional flexibility in identifying HTTP applications, customers can create custom signatures to identify an application.

Another aspect of protection that App-ID and the Palo Alto Networks firewall can provide is visibility into poorly configured or unauthorized applications that may be target the datacenter. Here too, the Palo Alto Networks next-generation firewall can be configured to monitor and log all attempts at DC access.

The key difference between Palo Alto Networks and other solutions that claim to perform application control is the use of a positive control model (allow only that which is expressly defined). By defining a short list of applications (using App-ID) that are allowed into the datacenter, using a positive control model, all other applications are implicitly denied. So in a datacenter deployment, all other applications that might try to gain access to the DC are blocked from accessing the zone and that activity is logged for forensics and auditing purposes.

POSITIVE VERSUS NEGATIVE CONTROL

Negative-model countermeasures operate on the basis of enumerating all communications and content that is known to be bad by virtue of its potential to cause damage. Antivirus software and intrusion detection systems are classic examples of this type of tool. The challenge is that new applications or threats cannot be stopped until they are identified and the tools are updated with the specific means to detect them (e.g., a signature).

In contrast, positive-model countermeasures operate on the basis of allowing all communications that are known either to be appropriate or necessary in a given situation, and then excluding everything else. The advantage is that such communications can be defined in advance, thereby enabling associated tools, such as firewalls, to automatically block a wide range of both known and unknown applications (or threats).
USER-BASED ACCESS CONTROL WITH ACTIVE DIRECTORY

The next step in isolating the DC is to associate the application identity with specific user and group information from Active Directory. Palo Alto Networks delivers this capability with User-ID, a technology that seamlessly integrates with Active Directory, enabling user- and group-based policy control, without the burdensome requirements of being in line with the authentication process or installing/maintaining an agent on every desktop.

Visibility into, and control over the application activity at a user level, not just an IP address level, is a required step in protecting the datacenter. User-ID is seamless to the end-user, once they have logged onto the network using their Active Directory credentials, the security policy controls DC access. There is no secondary web authentication login requirement (unless one is warranted), nor is there any duplication of users and groups, as is the case in some offerings.

User-ID enables policies that map the user and group identity (e.g., finance users, sales, marketing) stored within Active Directory to the application (e.g., Oracle). The policy can be created to allow only inbound traffic based on a specific application from a specific set of users and in so doing, limit the DC exposure.

In addition to allowing only Oracle into the DC, a policy can be established that dictates which IT tools are allowed into the DC for maintenance. Palo Alto Networks can establish a policy that assigns a set of IT tools (telnet, SSH, RDP) to the IT group, and forces those tools over a specific port. This allows a customer to shut out anyone other than IT from accessing the DC with these tools—on all ports, for all users. This level of visibility and control over both applications and users is unmatched in the firewall market. Every other firewall uses port and protocol to classify traffic along with IP addresses to control users, and as such, they are unable to provide this level of security and control.

MONITORING AND INSPECTING THE CONTENT

Controlling the applications and users that can access the DC solves only part of the visibility and control challenge that IT departments face when trying to secure the DC. With the understanding that the DC represents a significant corporate asset, the process of monitoring and inspecting the application traffic traversing each zone becomes the next significant challenge and one that is addressed by the real-time content inspection engine within Content-ID.

The threat prevention engine takes full advantage of the traffic classification that App-ID delivers and in so doing, enables IT to build an inspection policy that detects and blocks a specific set of threats flowing inbound towards the DC over a particular application (e.g., Oracle DB, MS SQL). Palo Alto Networks’ threat prevention engine combines several innovative features to scan all the traffic for viruses, spyware and vulnerability exploits in a single pass:

- **Uniform signature format:** Rather than use a separate set of scanning engines and signatures for each type of threat, Palo Alto Networks uses a uniform threat engine and signature format to detect and block a wide range of malware while dramatically reducing latency.

- **Stream-based threat scanning:** Virus, spyware, and vulnerability prevention is performed through the use of stream-based scanning, a technique that begins scanning as soon as the first packets of the file are received, as opposed to waiting until the entire file is loaded into memory to begin scanning. This eliminates the performance and latency issues with the traditional proxy approach by receiving, scanning, and sending traffic to its intended destination immediately without having to first buffer and then scan the file.

In contrast, traditional file-based AV engines start scanning only after a complete file is received and buffered into memory, where files are typically held until the scanning is finished and
deemed to be clean. Traditional file-based threat prevention introduces significant performance delays and latency, and can often result in a timed-out connection while waiting for the file to fully transfer and be scanned. In datacenter deployments, Palo Alto Networks’ high speed, stream-based scanning can scan traffic (including SMB/CIFS) at a multi-gbps throughput speeds which can eliminate software-based virus prevention offerings along with the related performance degradation.

- **Network and application vulnerability exploit prevention:** Incorporated into the uniform signature format is a set of intrusion prevention system (IPS) features that block known and unknown network and application-layer vulnerability exploits from compromising and damaging enterprise information resources. Vulnerability exploits, buffer overflows, DoS attacks and port scans are detected along side viruses and spyware by scanning the traffic only once, using proven threat detection and prevention (IPS) mechanisms.

- **File and data filtering:** Operating under the assumption that multiple layers of protection are key to protecting the datacenter, a policy can be put in place that scans outgoing traffic for the unauthorized transfer of files and data. Taking full advantage of the in-depth analysis performed by App-ID, the Content-ID engine enables administrators to implement data filtering policies to detect social security numbers, credit card numbers and custom data patterns. Files based on type (as opposed to looking only at the file extension) can also be detected. Response options for data filtering include blocking the transfer, logging it, sending an alert or a combination of all three.

**DATACENTER PROTECTION WITHOUT PERFORMANCE DEGRADATION**

One of the most critical elements to protecting the datacenter is to do so while keeping up with the high volume of traffic. Any injection of latency or slowdown in throughput will commonly result in the security being pulled out of the line of traffic. Palo Alto Networks next-generation firewalls SP3 use a single pass parallel processing (SP3) architecture to protect datacenter environments at speeds of up to 10 Gbps.

The two key elements that make up the SP3 architecture are the single pass software architecture and the custom-built hardware platform. The SP3 architecture is a unique approach to hardware and software integration that simplifies management, streamlines processing and maximizes performance.

**SINGLE PASS SOFTWARE**

Palo Alto Networks single pass software is designed to accomplish two key functions within the Palo Alto Networks next-generation firewall. First, the single pass software performs operations once per packet. As a packet is processed, networking functions, policy lookup, application identification and decoding, and signature matching for any and all threats and content are all performed just once. This significantly reduces the amount of processing overhead required to perform multiple functions in one security device.

Second, the content scanning step in Palo Alto Networks’ single pass software is stream-based, and uses uniform signature matching to detect and block threats. Instead of using separate engines and signature sets (requiring multi-pass scanning) and instead of using file proxies (requiring file download prior to scanning), the single pass software in our next-generation firewalls scans content once and in a stream-based fashion to avoid latency introduction.

This single pass traffic processing enables very high throughput and low latency – with all security functions active. It also offers the additional benefit of a single, fully integrated policy, enabling simplified management of enterprise network security.
PARALLEL PROCESSING HARDWARE

The other critical piece of Palo Alto Networks SP3 architecture is hardware. Palo Alto Networks next-generation firewalls use multiple banks of function specific processing operating in parallel to ensure that the single pass software operates as efficiently as possible.

- **Networking**: routing, flow lookup, stats counting, NAT, and similar functions are performed on network-specific processor.
- **Security**: User-ID, App-ID, and policy lookup are all performed on a multi-core security-specific processing engine with acceleration for encryption, decryption, and decompression.
- **Threat prevention**: Content-ID uses a dedicated content scanning processor to analyze content for all manner of malware.
- **Management**: A dedicated management processor drives the configuration management, logging, and reporting without touching data processing hardware.

The final element of the architecture revolves around built-in resiliency which is delivered by the physical separation of data and control planes. This separation means that heavy utilization of one won’t negatively impact the other – for example, an administrator could be running a very processor-intensive report, and yet the ability to process packets would be completely unhindered, due to the separation of data and control planes.

The single pass parallel processing (SP3) architecture

The single pass parallel processing architecture is completely unique in network security, and enables Palo Alto Networks next-generation firewalls to secure enterprise datacenter environments at very high levels of performance.
The datacenter has always been the heart and soul of the enterprise, acting as the location that touches all business applications and transactions. At one time, data center applications were restricted to internal networks where ample bandwidth was available, and delays or latency issues had little or no effect on business.

Times have changed in several respects. Datacenter applications must now deal with traffic from internal and external sources and the transaction size has shrunk – due to the HTTP-centric nature of most applications. Unlike ever before, the nature of the business today is online and real-time -- latency, performance degradation or unplanned outages are unacceptable. These traffic characteristics, combined with the need to protect applications against vulnerability exploits, place an unprecedented strain on existing security infrastructure in terms of performance.

Palo Alto Networks’ next-generation firewall addresses these issues with a combination of granular policy control, threat prevention and high performance, low latency throughput that today’s datacenters require.