OPUS one

Six Strategies for Defense-in-Depth

Securing the Network from the Inside Out

Joel Snyder

INTRODUCTION

The idea of perimeter defense when referring to a corporate network ignores common knowledge: that most successful and significant security breaches don't come from the outside. Serious issues often originate inside the network: everything from worms, viruses, and Trojan horses to unsecured wireless networks, peer-to-peer mobile communications and guest users can compromise the security of corporate networks.

To address these threats, the corporate network should no longer be a single homogeneous zone in which users connect from anywhere in the network and receive the same levels of access. Instead, the network requires internal perimeterization and defenses. Regulatory requirements also demand stringent controls on data flow within the corporate network. Logging and auditing requirements put pressure at one end of the spectrum, while rules regarding disclosure and information sharing are pushing against the other side.

In addition, the notion of a perimeter in a corporate network is fast disappearing. While site-to-site and remote access VPNs are extending the perimeter, employees themselves are eroding the perimeter and making it weaker—often without being aware of the impact they are having on network security.

For example, a mobile employee who connects a laptop to the Internet from a mobile hotspot and is exposed to a worms or viruses can infect the corporate network when the employee returns to the office. The firewall that stopped the worm at the perimeter is unable to stop this internal attack because it came from a trusted source. Similarly, an unsecured wireless access point (AP) in the corporate network can singularly jeopardize the security provided by the perimeter firewall.

Finally, mobility itself brings chaos to any network manager's attempt to segregate and segment traffic. Contractors and visitors require access to the Internet, while employees themselves move about within the campus connecting at different locations. Segmenting traffic based on source IP address is simply not enough in this environment, as a malicious client can easily assume another identity by changing its own IP address.

The response to address the new security environment of corporate networks is often referred to as defense-in-depth. The idea is to add protection at multiple layers rather than relying only on a perimeter firewall. Networks can no longer be partitioned into "inside" and "outside."

Defense-in-depth requires that relationships between network resources and network users be a controlled, scaleable and granular system of permissions and access controls that goes beyond simply dropping firewalls between network segments. The defense-in-depth banner has been handy for all sorts of other security products, from IDS to virus scanners---certainly useful additions to a corporate network security plan. But few security architects have taken the idea of defense-in-depth to its logical conclusion: turn the network inside out.

MAKING A NETWORK SECURE: DEFENSE-IN-DEPTH

Defense-in-depth is a dramatic departure from the transparent data corridor of the LAN. By pushing security into the network itself, the LAN changes from a public-access highway to a high-security network of roads, serving gated communities. Adding security into the LAN requires considering and implementing three key attributes of secure networking:

Access control - knowing who is on the network (authentication), what resources they are authorized to use, and applying these access controls to their traffic

Integrity - guaranteeing that the network itself is available as a business critical resource and that threats can be identified and mitigated.

Privacy - ensuring that traffic on the network is not accessible to unauthorized users.

Defense-in-depth is not a product, like a perimeter firewall. Instead, it is a security architecture that calls for the network to be aware and self-protective. In studying the problem of adding defense-in-depth, we've identified six key strategies that security architects can use to change significantly the security posture of enterprise wired and wireless LANs (WLANs):

Strategy 1: Authenticate and authorize all network users

Strategy 2: Deploy VLANs for traffic separation and coarse-grained security

Strategy 3: Use stateful firewall technology at the port level for fine-grained security

Strategy 4: Place encryption throughout the network to ensure privacy

Strategy 5: Detect threats to the integrity of the network and remediate them

Strategy 6: Include end-point security in policy-based enforcement

Problem	**Challenges**	**Solution**
We don't know who is on our network	Maintaining authentication databases for all types of users and systems; equipment that doesn't support authentication protocols	Authenticate users (and perhaps devices) within the network, leveraging tools like 802.1X, RADIUS and LDAP to provide both authentication and authorization information

STRATEGY 1:
Authenticate and authorize all network users

The starting point for any deployment of defense-in-depth is authentication. Authentication should be handled at the earliest point of connection of the system to the network: at the port level, even before the client is assigned a network address.

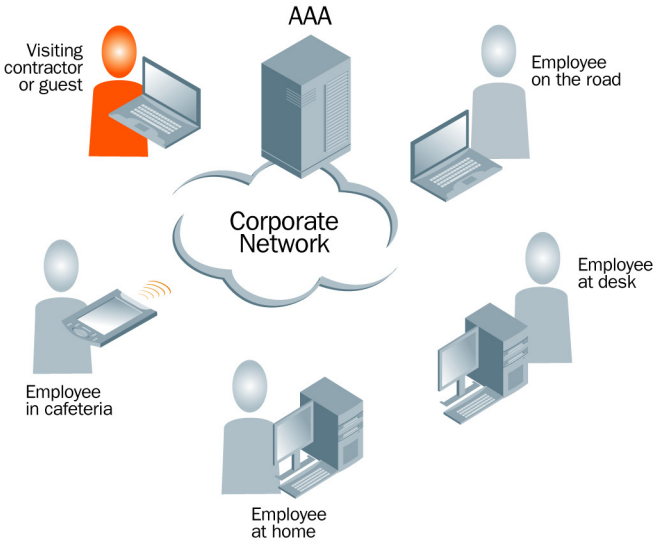
Associated with every positive authentication must also be authorization: now that we know who this person is, what does it really mean? What can they do? Where can they go? Unless every user in the authentication database has the same privileges and accesses, authentication must be tightly linked to authorization. The combination of positive authentication and user-based authorization information should form the basis for policy enforcement.

Challenges in Authentication

There are two key challenges in implementing network user authentication: the lack of a centralized authentication database, and the inability of some legacy systems to support modern protocols.

The clear choice for network authentication is IEEE 802.1X, the IEEE standard for network authentication. As an open

Figure
New approaches to security require authentication for all users prior to being granted network access. Centralized policy management drives this new security architecture. Sophisticated new systems that centralize security can now enforce user access based on location, device type and a myriad of other parameters.



standard with support for multiple authentication protocols, 802.1X is flexible enough to support everything from digital certificates to username/password authentication, and platforms from low-end PDA devices and mobile phones up to desktop and server operating systems.

802.1X has become a strong force and has already seen widespread adoption across network equipment manufacturers and operating system vendors

Strategies for Deploying Authentication in Networks

The obvious place to start deploying network-based authentication using 802.1X is in the wireless network. As a replacement for simple WEP authentication, 802.1X can be used by itself or in conjunction with WPA or 802.11i security. Since wireless is becoming an obligatory technology for most buildings, adding 802.1X both resolves the demand for wireless and offers the opportunity to get acquainted with the technology and the protocol.

Because 802.1X supplicant software is built into recent versions of both Windows and Macintosh operating systems, testing supplicants (clients) is rarely a difficult process. However, other platforms, such as PDAs and particularly embedded wireless devices (such as wireless print servers), may present a challenge.

Once there is experience with wireless deployments, it is time to move to wired device authentication. Although a full roll-out will probably require some replacement of equipment, it is likely that there is 802.1X-compatible hardware somewhere in the enterprise that can be used to begin wired testing and start deployment.

Defense-in-depth is successful only if authorization is implemented successfully following positive authentication. It is critical that a user's privileges on the network vary based not just based on their identity but also based on other intelligence about the user such as:

(1) machine identity
(2) security level of the machine
(3) location of the user
(4) time of day and
(5) authentication method

For example a user accessing email from a personal computer at home on the weekend may be given access to email only if the home PC is running appropriate version of the corporate-approved firewall. In the event of non-compliance the user may be directed to a download site to download such software. An interesting use of location-based authorization is enabled by intelligent WLAN systems that can pinpoint the location of the user. In such a scheme, a use can be prevented from accessing sensitive applications when sitting in the corporate cafeteria.

Overall, effective network security begins with authentication at the earliest possible stage and with intelligent authorization. This combination of authentication and authorization should form the basis of security policy in corporate networks today.

BOTTOM BAR
Virtual LANs extend the Ethernet standard by letting two different networks share the same wire. To keep the traffic separated, each frame from each network is tagged with a VLAN number. At either end of a physical link, devices such as switches or routers know how to interpret the VLAN tags and break the traffic apart. End systems only see the traffic from the LAN they "belong." In effect, what used to require two sets of equipment and two physical wires can now be done with a single set of VLAN-capable switches and routers.

Problem	Challenges	Solution
Need to separate network-connected entities into different security and service profiles without rewiring and reengineering the network	VLANs can be used for security isolation, but there are dangers in packet leakage and misconfiguration; switches now become firewalls. VLANs as generic security barriers do not scale to large networks, especially multi-site ones	Use dynamic assignment to VLANs for devices and users as a way to provide coarse-grained control of security at the building level

STRATEGY 2

Use VLANs for traffic separation and coarse-grained security

VLANs are, by their nature, unrouted chunks of network traffic. In most modern building networks, a fair amount of layer 3 IP routing takes place between wiring closets and the computer rooms. In a campus environment, routing is even more common. This makes pushing large numbers of VLANs around the infrastructure a fairly difficult-to-manage process.

Although most networks are heavily over-engineered with Gigabit (or 10 Gigabit) trunks, carrying a large number of VLANs around the network to represent different security profiles can stress not only the infrastructure, but also the management of the network itself. This difficulty is compounded as WLANs are added to the network. To maintain simplicity, enable inter-SSID mobility and preserve the current IP addressing scheme, it is essential that the WLAN architecture of choice have the ability to enable multiple VLANs across a single SSID. This is typically true of new generation of centralized WLAN solutions.

Strategies for Security VLANs

The key to successful use of security VLANs is dynamic assignment. While some ports in the network can be 'hard wired' to a particular VLAN (for example, in the server room or in the reception area of the company), assigning traffic to a VLAN should be done dynamically based on the authentication provided by the user (see Strategy 1, authenticate and authorize network users). Dynamic assignment is a critical requirement in building manageable networks. Static definition of security tends to cause long-term maintenance problems and impedes mobility of end users. By tying security to authentication information retrieved at the point of network access, secure networks can support quickly changing and moving user populations with minimum staffing costs.

There are multiple ways to assign devices to VLANs dynamically, including:

- based on 802.1X authentication information
- based on Web-based authentication information
- according to an SSID selected by the user in a wireless network
- based on detection of some other attribute, such as the MAC address of the device or the location of the user

Bringing dynamic assignment into the network requires a mechanism for providing authorization information at authentication time. In certain environments this can be maintained manually or using an out-of-band mechanism such as a user list. In the case of SSID selection in a wireless network, the user is asking for permission to connect to a particular network and then authenticates (or proves knowledge of the SSID or WEP key, depending on how secure of an environment is needed) to finalize access.

With 802.1X authentication, there is no way for a user to request a particular VLAN, which means that users must have VLAN information stored in the 802.1X authentication (RADIUS) server. Fortunately, an IETF-standardized mechanism exists to let a single RADIUS server send this information down to different devices. WLAN assignment may also be modified based on other information, such as the location of the user or the results of an end-point client security scan. The first step, then, should be to use VLANs and 802.1X/RADIUS authentication for assignment as this is most likely to be supported across multiple devices.

Problem	Challenges	Solution
Enforcing fine-grained security policy within the network based on who a user is	Policy management is difficult. Firewalling has become inexpensive, but still represents a considerable premium over simple switching	Build stateful security policies based on group information, applying policy at the port level

STRATEGY 3

Use firewall technology for fine-grained security

While using VLANs is sufficient for a coarse classification of some network users, the real solution to securing such a valuable resource is a fine-grained, user-based set of security policies enforced by the network.

Many enterprise network managers have reached the same conclusion and have begun embedding perimeter-style firewalls throughout the network interior in an attempt to apply security policy at points other than at the Internet access gateway. There

are two major problems with using interior firewalls to enforce policy. First, packets do not come with authentication information stapled to them. This means that when a firewall deep in the network has to make an access control decision, either it has to depend on highly unreliable information (the IP address that the packet is coming from) or it has to put up a new roadblock and insist on user authentication to that firewall. While a variety of proprietary approaches to this issue have been offered, typically based on some VPN-like authentication and encryption scheme, all are attempting to solve the problem the long way around.

The second problem with using interior firewalls is that there are never enough of them. Systems connect to the network at the edge, not at the core, and the traffic from those systems needs to be controlled at the point of entry to the network. Catching it hops down the line is too late: the control needs to be tightly bound to the point of entry. All of this points to a simple sounding strategy: firewall at the port level.

Until recently, firewalling at the port level was impractical at best. Now that we have technologies such as 802.1X authentication (don't forget Strategy 1, authenticate and authorize users) and firewall systems with very high port densities, enforcement of stateful security policy at the port level is a reasonable and economical goal.

Challenges to Enforcement of Fine-Grained Access Controls

There are two main challenges to enforcing fine-grained access controls. The first is management: how to define and create the stateful security policies and then how to bind those to an authenticating user. The second is economics. If a high-end managed LAN switch costs $50/port, the firewall vendor accustomed to getting $500 to $1000/port for appliance-style firewalls is going to have an uphill battle selling something ten times the price just to add security.

Overcoming the management issue is more significant, because no amount of budget flexibility can solve an unmanageable problem. The key to breaking through the problem of managing per-user network security policies is to move to a role-based management model. Although everyone is different, people are not really that different---and it is simpler to define access controls and security policy based on roles that the user has within the organization than try and answer the question "what should this specific person have access to?"

Starting with role-based management, users are assigned to groups that represent the roles that they play. A key requirement is the ability to assign a user to multiple groups. Because users do have multiple roles, they must be able to take on those roles simultaneously.

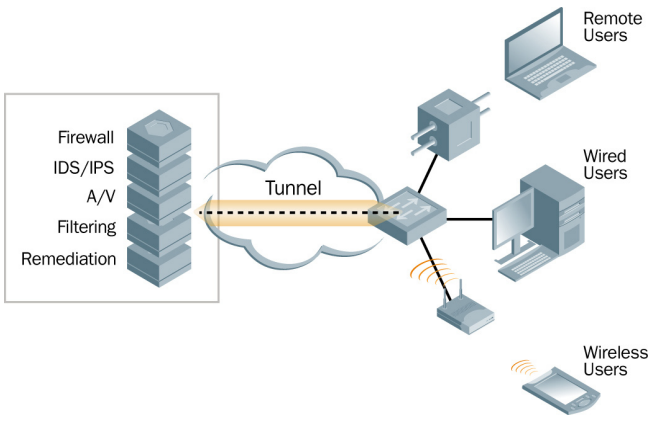
Strategies for Applying Fine-Grained Access Controls

Any application of security policy has to start with definition of the policy. Although technology solves many problems, the difficulty of defining policy has never changed and must be tackled first. If a security policy for inside the network cannot be defined and agreed upon, there's no point in going any further.

Security policy should be role- and resource-based, defining who has access to what resource, how the resource is accessed (Read? Write? Put? Get?) and any other modifiers, such as time of day or user location.

One winning approach to pushing out security policy enforcement to the port level is to start with the wireless network because wireless security is predicated on user-identity, given that users are no longer associated with physical ports, intelligent wireless network products integrate policy-enforcement directly into the system.

Figure
Enterprises should consider ways to virtualize security services for all users and all ports without having to deploy security appliances in every wiring closet.



Problem	**Challenges**	**Solution**
The network must protect the data	The best privacy would be encryption at the application layer. Applications don't generally do that.	Build encryption into the network where possible and where risk is high

STRATEGY 4
Place encryption throughout network to ensure privacy

Privacy of data throughout the enterprise is becoming a significant issue. Because the network itself carries very sensitive data, there is a strong need to protect that data from accidental or in-

tentional disclosure. The obvious case is in WLANs: no network manager would consider deploying a wireless network solution that does not enforce strong encryption.

In the wired environment, encryption can also be appropriate. The wake-up call for most network managers has come in the form of regulatory requirements. For any health care provider touched by the Health Insurance Portability and Accountability Act (HIPAA) requirements, wide-spread encryption of data even when inside of the corporate network may be required by law. Regulations such as California's SB1386 (on publication of information when private information is exposed) are also pushing companies to encrypt more data to reduce the risk of disclosure of protected information.

As every student of network security knows, encryption to assist in privacy can be done at any layer, from the physical link all the way to the application. There are tradeoffs with each alternative, generally revolving around coverage and generality. Encrypt lower for a general-purpose solution that covers all applications. Encrypt higher and protect the data end-to-end, eliminating any potential for exposure. The lower in the stack the encryption, the more traffic that can be encrypted and the lower the likelihood of network eavesdropping.

However, encrypting "low" in the stack means that the data are in the clear as they move from the application to the first encrypted link. In that sense, the likelihood of exposure grows larger, especially at network control points handling unencrypted traffic. From a host security point of view, the ominous presence of malware also poses a risk to having unencrypted data flowing through the guts of an end-system. The alternative, then, is application-layer encryption. However, this approach means that each application server and client must be modified to support encrypted data---a huge task.

Challenges when Adding Privacy to Networks

When approaching privacy from a defense-in-depth point-of-view, the natural inclination is to build encryption into the network itself. While IEEE 802.11i is helpful when discussing privacy in WLANs, for the wireline side of the network, the alternatives are less clear-cut. Unfortunately, 802.11i doesn't apply in a wired environment. Network managers are left in a standards void, then, with no obvious analog to 802.11i for the wired LAN. Additional alternatives for wireline encryption are proprietary link-encryption systems that offer security at the data link layer, but no higher. Higher-layer encryption that can traverse data links but stops short of the end systems is also an option. For example, cooperating network equipment at the network jack and the data center could encrypt data across multiple links and switching/routing points.

Strategies for Adding Privacy to Networks

For wireless networks, adding privacy to networks is easy. The IEEE has published IEEE 802.11i, a specification of wireless security that describes exactly how to provide privacy and integrity on a WLAN using state-of-the-art encryption algorithms and state-of-the-art authentication based on 802.1X (See Strategy 1). Any network manager considering privacy and encryption within the network should be looking at 802.11i for standardized and widely interoperable solutions for wireless.

Alternatives for wireline encryption include proprietary link-encryption systems that offer security at the data link layer, but no higher. Higher-layer encryption that can traverse data links but stops short of the end systems is also an option. For example, cooperating network equipment at the network jack and the data center could encrypt data across multiple links and switching/routing points.The table below provides some guidance on best practices in incorporating encryption and message authentication into data networks.

Table

ENVIRONMENT	COMMON SOLUTIONS
All wireless	802.11i combined with 802.1X using either TKIP or AES encryption
Server-to-server wired	IPsec in transport or tunnel mode between servers or server farm subnets
Client-to-server wired	Ideally application layer encryption. Common alternative is typically link layer encryptin between wiring closet and data center. New alternative is IPsec encryption from each network jack to data center
Client-to-server remote access	VPN protocol such as IPsec or SSL corporate VPN gateway

Wireless networks of all kinds require strong encryption at the link layer. Although link-layer encryption is desirable in the enterprise network, for all but the most sensitive of applications it is unlikely to be a requirement. As a stop-gap, application-layer encryption adds tremendous privacy, while proprietary link-layer wired LAN solutions extend the security perimeter at reasonable cost and without disrupting existing systems or applications.

Problem	Challenges	Solution
Identifying and remediating threats to network integrity in a cost-effective way	Balancing threat identification with resource cost	Analyze requirements for intrusion detection and remediation and find a solution which fits the network's real requirements

STRATEGY 5

Detect threats to the integrity of the network and remediate them

The challenge for implementing internal IPS/AV schemes is that boxes have to be located in every closet and even then they cannot prevent a PC from potentially affecting its peer on the same network. A better way to address this problem is to encrypt traffic from each network jack and bring it into back to a central location where all the policies are applied. This method is non-disruptive to addressing schemes and is far better than distributing multiple firewalls and IPS/AV systems in each wiring closet.

If there is a trinity of security concerns in access control, privacy, and integrity , the third of these gets the least interest. The main reason for this is simple: detecting threats to the network can be very difficult. While some threats to network and data integrity are easy to identify and remediate, others can be extremely hard to detect---and even more difficult to protect against. While many companies focus on 'towards the firewall' threat management, the threats can come from anywhere: worms and viruses, wireless, guests, and careless or malicious insiders. It is worth while to identify as many of these threats as possible and either notify or attempt remediation.

The security community's first attempt at threat identification came in the form of IDS, intrusion detection systems. While IDS have proven their worth as a tool in the arsenal of the security analyst, most enterprises have discovered that the information they get from their IDS is not primarily useful in detection and remediation of immediate threats. An IDS is like a protocol analyzer: it's a tool for the security analyst to use in diagnosing and identifying problems, more than a first-line-of-defense against network integrity threats.

To support the continuing need for threat detection and management, security vendors have flooded the market with products ranging from in-line intrusion prevention systems (IPS) based on the same core technology as IDS, to application-layer firewalls and highly specific tools designed to catch a particular type of threat, such as a network worm. Even more mundane areas such as anti-virus scanners have moved into network-based devices, hoping to catch viruses on the fly in the network, no matter what protocol they use to propagate.

Threat management tools also extend into more analytical areas, such as vulnerability analyzers and security information management (SIM) tools that collect data from multiple network and security devices and attempt to identify threats by correlating log information.

Challenges in Ensuring Network Integrity

The greatest challenge in managing network threats is defining the appropriate risk/reward balances. We've already discussed, briefly, the difficulty of determining ROI of security products in general. With threat management and network integrity assurance, the ROI calculation is as hard as it gets. Obviously, you are hoping to protect against total network failure, but adding integrity checking tools to the network doesn't give a good metric of how much less frequently the network is unavailable or degraded for security reasons---or whether the tool will necessarily catch the problems that beset the network.

A commonly encountered challenge with deployment of network integrity products, such as intrusion detection systems, is the highly distributed nature of most networks. In a highly-switched network, monitoring the integrity of the network becomes a very difficult task. If you can't see the traffic, you can't detect threats and anomalous behavior.

When looking for areas to deploy network integrity tools, you may be stymied with another metric of difficulty: measuring risk itself. Successful integration of these tools requires understanding what the threats are you care about and what you need to do to detect them. While risk itself is generally unmeasurable, the threats to your network are not difficult to enumerate. Simply listing the threats, the consequences, and your remediation strategy will go a long way towards identifying the right strategy for ensuring network integrity.

Strategies for Ensuring Network Integrity

The most successful strategies will identify the areas of greatest risk and concentrate on those first. That's half of the best path forward. The other half is to examine the technologies that have the lowest cost, both in terms of capital and continuing operations and support.

A good example of the former is Trojan horses, viruses, and malware. These threats have the ability to degrade not only network and system performance, but they can also expose and disclose sensitive information or cause a complete denial of service. More importantly, it's very easy to become infected with various kinds of malware. The risk of infection is high, and the risk to the network is high in the case of an infection. This is why enterprise system managers universally have virus identification

and mitigation strategies already in place. Those who have not added malware/spyware to their anti-virus tools will be doing so shortly---the risk is too great to ignore.

While high-risk is easy to identify, low-cost is also a good way to find tools for your network integrity and protection arsenal. For example, most firewalls have some limited IPS capabilities built-in, such as denial of service (DoS) protection. Although these devices need a small amount of tuning, they can increase the level of network integrity without capital expense and with very low operational cost.

Although highly distributed networks are common deployment architectures, you may want to consider a more centralized strategy when it comes to monitoring network integrity and identifying threats. For example, pulling traffic back to a central data center and a small number of switching and routing points offers the opportunity to both monitor the network inexpensively and, where appropriate, install choke point technologies such as intrusion prevention systems.

Problem	**Challenges**	**Solution**
Even though a user on the network has been positively identified, they may be on a compromised platform and pose an unintentional threat to the network	Security posture is, at best, a coarse-grained barometer	At entry points and policy management points, incorporate end-point security determination; when defining policy, include end-point security as a factor

STRATEGY 6
Include End-Point Security in Policy Enforcement

User systems may range from tightly controlled laptops owned and managed by corporate IT to spyware-infected, keystroke-logging, Trojan-hosting systems at public Internet kiosks. A user who successfully identifies to the network should be given different privileges depending on the system they are using for access.

Most network managers are already aware of the problem of end-point security and have tools such as anti-virus, personal firewall, and patch management in place on many systems. The next step is verification: enforcement of policy regarding end-point security by varying access based on the security posture of the end system.

This technology and the thinking behind it is most evident in the world of SSL VPN where vendors are vying hard to differentiate themselves and incorporate end-point security posture detection and enforcement into their products. Remote access VPN tools, such as SSL VPN, have a particular vulnerability in this area because they are specifically designed to extend network access to users who may have a wide variety of untrusted computing platforms. However, this same thinking is beginning to move into the enterprise network.

The idea is simple: detect the security posture of the end system, and use that information to control access. Actually implementing end-point security policy enforcement is another matter entirely.

Challenges in Enforcing End-Point Security Policies

There are two significant challenges to deployment of policy enforcement based on end-point security posture. The primary difficulty comes from the wide variety of systems being used on corporate networks. Attempting to load and execute a security posture assessment tool on every system that connects to the network is an exercise fraught with danger and guaranteed only mixed success.

The second challenge is one of granularity. For example, if the end-point security assessment discovers that a personal firewall is loaded, but the firewall policy is out-dated, is the system in compliance or not? At the same time, security posture may mean different things depending on who the user is and what resources they are using. Thus, the compliance status may be difficult to determine without considering other factors. Building and, more importantly, maintaining the business logic to deal with different platforms and security postures will only get more complex, not less so.

Strategies for Enforcing End-Point Security Policies

Best practice solution strategies depend on a combination of forward thinking stateful security policy definitions and flexible remediation systems.

To use end-point security information within stateful security policies, the definition of every policy entry has to include the end-point security compliance level. A best practice is to reduce the analysis of end-point security to a zone definition, keeping the number of zones as small as practicable: three or four should be sufficient for most enterprises

A second best practice in this area is the provision of remediation resources. The idea is that if a user attempts to connect to the network but is considered out of compliance with security policy, simply blocking their access is not the right answer. Instead, the user should be connected to a section of the network that offers resources, such as the corporate personal firewall or anti-virus scanner or updated policies and virus definitions. Then the user has the opportunity to move into a compliant state and re-connect to the network to get full access.

CONCLUSION

Defense in depth is process not a product. It's a proactive approach to thinking about security from the inside out. Certain architectural approaches such as centralized security overlays lend themselves well to solve today interior security problems. Security continues to be an ongoing process and constant vigilance and user awareness play equally important roles in building the best security posture for enterprise networks.

END