

Application Control Defined – The Top 7 Capabilities Required to Restore Firewall Effectiveness

Author: Mark Bouchard

AimPoint Group
keeping IT on target

Executive Summary

For long the backbone of most enterprise's network security infrastructure, traditional stateful packet inspection firewalls are essentially blind to modern applications and threats. This condition has put many IT/Security managers in a precarious position. Should they attempt to utilize coarse control capabilities currently at their disposal, blocking ports or entire protocols just to eliminate access to a handful of known bad applications? For that matter, will such an approach even work? Or do they plot a more permissive course, allowing access to every application the business deems necessary or useful, simultaneously exposing the organization to greater risk and a myriad of unwanted applications that sap user productivity and consume valuable computing resources?

The answer is "none of the above." What organizations need instead is a next-generation firewall featuring application control. Among its many benefits, application control promises to restore firewall effectiveness by enabling IT to implement and enforce granular policies governing application access while helping prevent application-layer threats and unwanted exposure of sensitive data. Realizing this promise, however, is by no means guaranteed. The approaches being taken by both incumbent and emerging firewall vendors are many and varied, and so too are the results. Accordingly, this paper provides detailed explanations for the top criteria and capabilities that must be fulfilled to ensure enterprise expectations and objectives are fully met with regard to this crucial facet of next-generation firewalls.

Why Application Control is Necessary

The unfortunate reality is that traditional stateful inspection firewalls have failed to keep pace with ongoing changes to the application and threat landscapes. In particular:

- **Many applications now look the same (at least on the surface)** – Widespread webification of client-server applications has been taking place for some time. Add to this a rapid rise in the adoption rate for cloud-based web services, and it's no surprise that HTTP and HTTPS are estimated to account for approximately two thirds of all network traffic in today's enterprises.
- **Many applications are now taking advantage of evasive measures** – Techniques such as port hopping, the use of non-standard ports, and protocol tunneling have been common "features" of P2P file sharing practically since its inception. They are also standard fare for the majority of social networking and personal productivity applications that have since emerged. Moreover, many business applications are now being designed to incorporate these same capabilities to help enable operation in the broadest set of scenarios by minimizing the need for changes to intervening network and security infrastructure.

The Top Criteria and Capabilities that Define Application Control for Next-Generation Firewalls

To ensure a superior level of effectiveness and otherwise maximize an organization's return on investment, application control feature sets incorporated as part of next-generation firewalls should match up well with the following key characteristics and capabilities:

1. **Extensive Application Intelligence**
2. **Essential Contextual Information**
3. **Flexible, Granular Control**
4. **Application-layer Threat Prevention**
5. **Extensive Physical Coverage**
6. **Effective and Scalable Management**
7. **The Power to Perform**

- **Many applications can be classified as both “good” and “bad”** – Particularly with the emergence of Enterprise 2.0 – that is, the use of Web 2.0 technologies and applications to serve legitimate business purposes – it is no longer possible to classify many applications definitively as good or bad. Although some applications clearly reside at one end of the spectrum or the other, many fall somewhere in between. Not only that, but their classification is likely to vary from organization to organization, and even within a given organization from use case to use case.

Of course the bad guys are well aware of these changes too, which is why they have been focused for some time now on developing threats that not only target applications in general but also attempt to take advantage of their reputation and/or evasive capabilities.

For firewalls that rely primarily on IP addresses, ports, and protocols for classification purposes, the result is the inability to reliably distinguish network traffic associated with applications being used for legitimate business activities from that associated with applications being used for other reasons. Another significant issue is a higher potential for successful threat penetrations. In particular, for those cases where all allowed traffic is not automatically inspected for threats by default, the inability to classify applications accurately will carry over into the process for deciding which sessions get passed to the firewall's deep packet inspection engine for further analysis.

The impact for IT/Security is that they have essentially lost control. Assuming they continue to rely on traditional firewalls, none of the options available to them is particularly attractive. They can *attempt* to block every application of a questionable nature and in doing so risk impeding the business. Alternately, by adopting a more permissive policy, they can practically guarantee accessibility of all applications the business desires to use, but must then deal with a myriad of unwanted ones as well – not to mention the threats that hitch a ride with them.

The solution to this apparent conundrum is application control. A feature of next-generation firewalls, application control is *intended* to correct the far-sightedness of traditional firewalls by providing the means to resolve not only the overlying application responsible for a given traffic stream, but other important details as well. Examples include who is behind a given activity, whether its purpose is good or bad, and whether it involves the transmission of sensitive data.

What Enterprises Should Look for in a Solution

Having a certain intention and delivering on it are completely separate things. Moreover, no two application control feature sets are created equal. Only by selecting a solution that fully meets the criteria and capabilities described in the following sections will organizations be assured of maximizing the gains from this all-important component of the next-generation firewall.

Extensive Application Intelligence

The foundation for any application control feature set is the ability to identify applications from the network traffic they generate. Clearly, the coverage provided should be broad in terms of the types of applications that can be detected, with support for everything from web, Web 2.0, and conventional client-server applications to business, non-business, and every class of application in between. However, it should also be deep in terms of being able to distinguish individual functions within many applications – for example, ordinary chat sessions versus file transfer traffic for common instant messaging applications.

The basis for these capabilities will typically be derived from the presence of an extensive application signature library, application knowledgebase, and URL database, as well as the corresponding infrastructure required to maintain them. The latter includes a worldwide sensor grid, a highly experienced application research team,

and geographically distributed update servers. Heuristic and behavioral algorithms may also be used in some cases to supplement application signatures, but these will generally exhibit lower detection accuracy and are by no means sufficient on their own.

Another major consideration is the specifics of how the various detection mechanisms are implemented. In particular, it *must* be done in a manner that renders them immune to common evasion techniques, ideally while also providing the means to handle encrypted sessions effectively.

Essential Contextual Information

The ability for many applications to be both good and bad means that simply identifying an application is not sufficient. Effectiveness also depends on gathering additional contextual information that can be used to establish further the relative appropriateness of an application session and the individual activities that comprise it.

The single most useful attribute in this regard is the identity of the user involved – and not just their username but also role and group affiliations. Ideally, the solution should seamlessly and transparently leverage existing enterprises sources of this information while also incorporating provisions to account for guest users. Other, attributes the solution should be capable of factoring into an eventual response include but are not limited to: the type, ownership, and security status of the endpoint device being used; the user's location; the type of network connection being used; the type and volume of content involved; the time of day, week, month, or quarter; and the reputation of the party or site at the other end.

Flexible, Granular Control

Another implication of the various shades of gray that modern applications can exhibit is that it is no longer enough to merely deny or allow access. Instead, a much broader and more granular set of actions should be supported to match the spectrum of situations that will inevitably arise.

Besides allowing certain functions within applications but not others, it should also be possible to invoke any of the contextual data covered in the previous section to specify the conditions that must be met for full or even partial access.

Application bandwidth management is another response mechanism that IT managers should be able to invoke on a similar set of bases. In other words, they should also be able to prioritize processing and/or allocate throughput levels for specific applications, users, groups, and so forth. This way bandwidth-hogging lifestyle applications that support streaming video – like YouTube – can be demoted in favor of business-critical applications, which, in turn, can be granted different levels of service.

Application-layer Threat Prevention

An historical shortcoming of traditional firewalls is that allowed communication sessions could very well include malware and other types of threats capable of infecting and damaging an enterprise's network. Accordingly, the primary objective of application-layer threat prevention is the closure of this loophole, particularly for those threats geared to exploit application-layer vulnerabilities (note: threats focused at the network layer are outside the domain of application control and should be covered by another part of the overall solution). This necessitates the inclusion of multiple, *real-time* threat/malware detection mechanisms common to today's standalone gateway anti-virus and network intrusion prevention systems, such as signature, protocol anomaly, and behavior anomaly-based analysis engines. Accounting for threat inspection in this manner also has the advantage of being highly efficient. It eliminates the need for a separate device or solution and reduces associated capacity requirements by limiting inspection to only that traffic that is otherwise allowed.

A second objective of this capability is the prevention of unwanted leaks of sensitive information. The idea in this case is to leverage application identification and in-depth inspection routines to examine traffic also for individual pieces of data, entire documents, excerpts of documents, or entire types of files that can then be blocked as appropriate. To maximize the effectiveness of this sub-capability, it should not be limited to a subset of protocols, such as SMTP, FTP, and HTTP – as it often is for many other data leak prevention (DLP) tools and feature sets.

Extensive Physical Coverage

The criteria and capabilities that have been covered thus far are essentially table stakes. Although there is room for minor bits of differentiation in these areas, the greatest opportunity for a solution to separate itself from the competition lies with providing extensive physical coverage and effective, scalable management.

The point of providing extensive physical coverage is that organizations actually need application control to be implemented uniformly across their entire computing environment, from main and regional headquarters locations to smaller branch offices and facilities. This requires having access to a suite of application control enabled devices that span a wide range of price-performance points. Not only that, but coverage should be provided not just for fixed LAN users, but also for:

- Mobile users, such as guest and itinerant employees – for example, by having both an efficient way to accommodate unknown users and the ability for user-specific policies to effectively follow them from one location to the next
- WLAN users – for example, by having integral support for wireless controller functionality, or otherwise integrating with common, standalone controllers
- Remote users, such as telecommuters – for example, by having integral support for a full-featured SSL VPN

Robust, distributed and delegated administration functionality is also necessary to support this capability of course, and is covered in the following section.

Effective, Scalable Management

Achieving effective and scalable management depends on three primary sub-capabilities: visualization, scalable rules generation, and scalable distributed administration.

Visualization – Reports, data navigation utilities, and analytic tools are required to help understand what is actually happening on the network – which specific applications are being used, by which users, when, to what extent, and so forth. These insights are necessary for refining policies and establishing rules in the first place, as well as for illustrating the impact of rule enforcement and the need for changes over time. Minimum capabilities include real-time monitoring, intelligently organized activity summaries with drill downs that expose further details, and flexible access to at least a modest amount of historical data (e.g., 30 to 60 days).

Scalable rules generation – A significant and often overlooked consequence of application control is that it results in administrators having to deal with substantially more information and attributes when formulating access rules. Countering this expansion in scope requires a solution with a logically organized, object-based model that centers on people/groups, applications/categories, and actions as the main elements that comprise a rule. Powerful grouping features are also a must-have, as is an application library that supports categorization along numerous dimensions, such as type, risk level, productivity impact, resource utilization, etc.

Another essential element for scalable rules generation is a unified policy model. The goal in this case is to have a homogenous set of rules – as opposed to managing a set of traditional firewall rules in one part of the product and application control rules in another. Failure to deliver in this area is a sure sign of a product where application control has been bolted on, instead of being embedded as a foundational component of the firewall's architecture – a condition that is characterized by several significant shortcomings, including patchy inspection and control capabilities, inadequate performance, and a cumbersome, error-prone method for configuring rules.

Scalable distributed administration – This item ties back to the need to provide extensive physical coverage. The associated management system should be capable of centrally administering distributed instances of the host firewall individually and in groups. In addition, the ability to set global application control policies should be complemented by delegated administration that allows for localized customization, particularly with regard to application access and bandwidth management.

The Power to Perform

Application control involves a compute-intensive set of processes and capabilities. An adjunct requirement, therefore, is that the solution has the power to deliver all inspection and control capabilities at high throughput levels and with minimum introduced latency.

In this regard there are numerous telltale characteristics that can help identify a high-performance solution. Innovative, low-latency inspection techniques, a dedicated management plane, and specialized processors definitely have a role to play. What matters more, however, is (a) whether the solution was designed (or re-designed) from the outset to support application control – as opposed to it being bolted on at some later point in the product's evolution, and (b) real-world performance results, preferably from a pilot test conducted in your own organization's network.

What Enterprises Stand to Gain

The bottom line is that changes to the application and threat landscapes have eroded the effectiveness of traditional stateful packet inspection firewalls. In this regard, enterprises will be served best by purchasing and implementing a next-generation firewall featuring application control capabilities. This conclusion only holds true, however, to the extent the included application control feature set aligns with the criteria and capabilities outlined herein that define a robust, enterprise-class solution. For enterprises that embrace next-generation firewalls that do in fact meet these requirements, the result is the ability to efficiently and effectively:

- Enable all applications that are being used to support legitimate business activities, regardless of type or reputation, including difficult to pin-down Web 2.0 and social networking applications
- Block all “bad” applications, including “good” ones that are being used in a bad way
- Prevent malware, attacks, and data leaks via allowed application traffic, without having to deploy additional, standalone devices/solutions
- Prioritize allowed application traffic based on its relative importance to the business
- Demonstrate not just compliance but also proficiency when it comes to meeting regulatory requirements pertaining to firewall implementation, network segmentation, and/or user and application-oriented access control

About The Author

Mark Bouchard, CISSP, is the founder of AimPoint Group, an IT research and analysis company specializing in information security, compliance management, application delivery, and infrastructure optimization. A former META Group analyst, Mark has analyzed the business and technology trends pertaining to a wide range of information security and networking topics for more than 14 years. During this time, he has assisted hundreds of organizations worldwide with both strategic and tactical initiatives, from the development of multi-year strategies and high-level architectures to the justification, selection, and operation of security and networking solutions. A veteran of the U.S. Navy, Mark is passionate about helping enterprises address their IT challenges.