Firewalls — Are they enough protection for current networks?

J. Stuart Broderick

Strategic Consulting, Symantec Corporation, 911 Central Parkway North, Suite 300, San Antonio, TX 78232, USA

Abstract

Firewalls are widely deployed in most organizations connected to the Internet to implement their information security policy and to protect information from unauthorized access. Yet intruders still manage to attack systems and gain unauthorized access to information that is supposedly protected by these firewalls. This article examines some of the history and background of firewalls and what has led up to the current situation. Additionally this article will examine how any organization, even your home network, can be made a far safer place using readily available existing firewall technologies.

Introduction

Using a firewall to protect your computer(s) has become an automatic, no-brainer decision for most organizations; firewalls are even available as free or low-cost downloads when users sign up for Internet services. This was not the case even a few years ago. In the late 1980s/early 1990s firewalls were relatively new, as was the mass appeal of the Internet. Because of this, the technology solutions available were expensive, often difficult to configure and provided only a basic level of protection for the networks they protected. After the introduction of the World Wide Web and the Internet browser in the early 1990s they were many new firewall technology entrants on the market. Many of these technology companies were short-lived because although there was promise of a big market, selling the concept of security protection to a marketplace that didn’t want to or even need to think about security was very difficult at best.

So what changed between the mid-1990s and today? Legislation. But legislation wasn’t just created for the sake of it. Legislation was being driven by numerous factors including improper financial and other business processes, and for the purposes of this article the rampant increase in the number and sophistication of attacks on computers by people both inside and outside of the organization(s) affected. But shouldn’t the firewall have prevented this? The answer is yes and no. Both answers are correct to a certain extent — the question is why. Unfortunately the answer is
rooted in history, the history of how and why the Internet uses the TCP/IP protocol suite as its core engine for transporting and accessing information.

TCP/IP – a very brief history in time

When the TCP/IP protocol suite was invented in the 1970s, it was done so to enable a very small group of people to share information ‘relatively’ quickly over a telephone line. The concept of a local area network was still a decade away! At the time, 110 baud modems were the highest speed available; but that was quick compared to sending floppy disks or tapes in the mail. So the original developers of the earliest protocols (ftp and telnet) had to create a very lightweight transmission protocol with very few features above the essential features to make it work. One of the features that made the cut-list very early on was security. The developers all knew each other, why did they need security, they were sharing information willingly. Little did they know that 25+ years later over 800 million users (2005) would be using the concept and technology they developed to save time compared to sending information via regular postal service and fax.

Over the following 10–15 years from the initial development of the ftp and telnet protocols a number of other TCP/IP protocols and applications were developed such as WAIS, gopher, ntp. Even in 1992, most of the world that was connected together via the Internet accessed data as text. Graphical browsers, http and web servers hadn’t been invented yet. During this period, Internet attacks were known about, but were not widespread. Perhaps the most famous being the Morris worm in 1998. This worm only failed to bring the complete Internet to a crashing halt because of a coding error. This together with the infamous exploits of Kevin Mitnick slowly brought the network security world to its senses.

TCP/IP security awakening

The network attacks of the late 1980s encouraged firewall developers to move beyond the simple router approach that had previously dominated. They began to realize the security shortcomings associated with many of the TCP/IP protocols. In a few short years preceding 1995, a number of breakthrough technologies evolved, including proxy firewalls, SOCKS, and stateful firewalls. Each firewall technology had its good points and weak points, its protagonists and its enemies. All tried to achieve the same thing – to secure the network; to provide perimeter protection. None succeeded 100% of the time – why? The problem lay, and still lays, inherently in the design of the protocols designed 20+ years earlier. Put simply, the protocols as originally defined and as used by millions of systems across the world could not be secured without a significant redesign of TCP/IP itself. These facts, combined with the apparently never ending demands for Internet addresses led to the development of the TCP/IP V6 protocol suite which was finalized in the mid 1990s. Several RFCs (RFC 1883, 1995; RFC 2460, 1998) were published detailing this standard. This protocol suite promised to solve many if not all of the problems facing the Internet at that time. Today, in 2005, TCP/IP V6 is still not universally deployed due to a number of reasons:

1. Organizations were concerned about interoperability of their existing networks and TCP/IP V6 networks.
2. Many of the core applications on which TCP/IP V6 was dependent were late or very late to the marketplace. One of the key ones being the Domain Name Service (DNS).
3. Organizations were concerned about the anticipated increase in computing power required to run TCP/IP V6.
4. Core Internet providers were focused on signing up as many clients as possible, not on securing them.

With the continuing focus on expansion of Internet usage, cheaper computers, provision of broadband communications to small offices and home, the surge of security interest on firewalls waned somewhat. After all what do firewalls do? They simply act as a guard to the network and make decisions on whether a particular packet or packet stream can pass through them based on a set of simple rules. Something now seen as ‘cool’ to the legions of new software developers entering the software world in the last 10 years. Instead many of these developers wanted to build a company to develop new ‘cool’ applications using ‘cool’ development techniques such as eXtreme Programming (XP) very quickly so that they could take their company to IPO status and make huge amounts of money. Security and secure programming was generally low on their list of priorities – but don’t worry, the firewall will protect us …

TCP/IP firewall: universal panacea for security

What do firewalls do? They simply act as a guard to the network and make decisions on whether
a particular packet or packet stream can pass through them based on a set of simple rules. In other words, they enforce your security policy.

To understand what a firewall does and how it works requires a basic understanding of how information is packaged up to permit it to be transported on an IP network. Simply put, each information package consists of three basic parts: header, payload (original data to be sent) and trailer. In reality, there are likely to be multiple parts in the header and trailer that contain information collected in the layers of the ISO communications stack pertinent to a particular communication. This is shown graphically in Fig. 1 for a data packet being encapsulated from ISO layer 3.

The header and trailer defines:

- where it is going to,
- where it is coming from,
- what TCP protocols the packet expects to be processed with,
- how large the original data segment is,
- the sequence number that identifies each packet in a stream of packets making up a given communication stream.

The payload contains protocol specific information and the user information that is being transmitted.

*It is vital to note that NONE of this information is protected against modification.* This means that anything that is sent across a TCP/IP network (including the Internet itself) is open to view by any third party. This was one of the compromises made in the original lightweight design of the TCP/IP protocol suite!

So how does a firewall determine whether a packet has been altered — it can’t, unless the protocol rules in the payload have been broken and the firewall checks the payload content. Hmmm … doesn’t sound too good does it? It isn’t, but the chances of this type of attack occurring, unless the target recipient is a high profile recipient, is quite low — the attack is simply too resource hungry for the attacker.

If an organization or individual wants to protect their information in-transit — they have only one option. They must encrypt the information in some way and manage the consequences of that encryption. Consequences include management of the encryption keys and out of band communication of the key to the intended recipient of the information, and the additional computing resources needed to encrypt and decrypt the information.

**CIAA**

When considering implementing a firewall solution, the organization must consider what it is attempting to do and what level of security and reporting it requires. In most cases security managers are concerned about protecting the confidentiality of information. Business/application owners may assume that confidentiality is the concern of the security manager and their concern is availability and integrity of the information i.e. the content has not been unexpectedly altered, and the information is there whenever they need it. The Security Auditor is concerned with all of these factors and additionally with being able to show that the firewall is working as required. This will be determined through analysis of the log files. Additionally, the log files provide accountability from what system, if not what user, communicated with certain systems and applications. If a proxy firewall is used, the content of what anomalies or unusual activity the communications contained can be known.

**Internal address protection**

Many organizations use so-called 'private' TCP/IP addresses that anyone can use, providing they...
don’t announce to the Internet that they are using them. Due to the relative shortage of public IP addresses, thousands, if not millions of users are using the same IP address on their systems. Routing of packets on the Internet would fail if these addresses were public. As a result a firewall is frequently used to provide Network Address Translation (NAT) services for the internal network.

The NATing firewall manages potentially thousands of simultaneous connections from computers connected to the internal network through a small number (often one) of externally facing Internet addresses. Now the firewall is acting as a guard and an accountant to check for authorized packets and to ensure they get routed to the appropriate internal computer.

**Distributed network protection**

So far, network protection has been addressed using a single firewall that connects the entire corporate network to the Internet at a single point. For many organizations this is not acceptable for business continuity reasons and for throughput reasons.

**Business continuity factors**

If the business of a given organization is entirely dependent on being connected to the Internet, a single firewall solution is unacceptable. Computers crash or fail sometimes in their operational lifetime — its a fact. System failure may be due to software, hardware, or facilities (e.g. Electricity, HVAC) failures. Regardless of the cause of the failure, the organization is off-line. Prudent organizations that require 100% online uptime opt to distribute their online connectivity across multiple firewalls, often geographically distributed to minimize the impact of upstream failures of the Internet or ISP. There are no issues with doing this. Its simply an additional cost in terms of hardware, software, line rental, administration and management.

Similar solutions can be adopted by organizations that require very high data volumes. Such organizations may have to distribute their load via multiple ISPs.

Organizations that require their Internet connectivity to be always on may opt for high-availability firewalls. In most cases, high availability does not equate to non-stop. Typically, high-availability firewalls are a pair of firewalls acting in Master/slave or Primary/Secondary operation where the secondary or slave firewall simply monitors whether the primary or master firewall is running. Once it detects failure, the secondary or slave firewall promotes itself to be the Master or Primary firewall until it is switched back to its default mode — either manually or automatically. Regardless of the detailed operations, the communications passing through the original firewall at the time of failure are lost and the session is terminated. This means the session has to be re-established. This is why it is high availability and not non-stop. It simply means that firewall functionality is always available except for the small window or a few seconds or less when the transition of the standby firewall to its elevated status is completed.

**Firewall location matters**

A firewall is a network perimeter protection device, therefore it is always located at the perimeter of what is being protected. So the real question is ‘Where is the perimeter?’

So far we’ve considered the perimeter of the network to be the point or points at which the organization connects to the Internet via the ISP. Increasingly, organizations are starting to use firewalls to segment their network in the same way that they use routers and switches. Typically, these inner perimeters surround mission critical, or legislative-critical portions of their networks such as finance, human resources, sales, and executive file-servers.

Deploying firewalls to protect sensitive parts of the networks makes sense because the firewall rule sets can often be made increasingly restrictive as they get closer to the critical parts of the organizations computing infrastructure. Perhaps the ultimate firewall would only permit one protocol to pass during certain periods of the day/month. In such cases those system(s) would be all but impenetrable from systems external to their network for most of their operational lifetime. The systems external to their protected subnet could be subject to whatever the next layer of less restrictive firewall permitted passage to.

**Internal network protection**

Most business managers believe that unauthorized intrusions and access attempts against their sensitive data are performed by external agents. While the most damaging attempts may be from external sources, it has been repeatedly proven over many years that most attacks against data and systems are from internal sources. Statistical evidence shows that 60–80% of all attacks are internal attacks by the people the organization trusts.

Many organizations still operated a ‘flat-security’ network model. This means that all systems and all data in the organization are protected to the same level and that level is defined by the
configuration of one or more network perimeter firewalls that provide Internet connectivity. While the 'flat-security' model provides some deterrent to external attackers, it provides no protection against internal attacks. Moreover, would-be internal attackers are trusted and often have a complete knowledge of how the network is constructed.

Managers of 'flat-security' networks focus most if not all of their security attention on the firewall(s) and can't figure out why they still get compromised. They get compromised from the direction in which they are not looking! To remedy this situation, the management of the organization must accept three things:

1. not all attacks are from external sources,
2. successful defense against attacks requires a defense in depth approach,
3. some information may be intentionally or accidentally disclosed using approved communication mechanisms e.g. email.

The first two items can be addressed using a technology based solution, the third cannot (in most cases). All three items can be addressed by increasing security awareness within the organization as a whole.

**Defense in depth**

The concept of defense in depth is one that security practitioners have touted for many years. Although use of multiple firewalls arranged in series to provide defense in depth is used by many government agencies across the world, its use in non-governmental areas has been somewhat limited — perhaps due to the perceived costs involved. A few years ago, the cost argument was valid. However, with the introduction of smaller and faster firewall devices and security appliances, perhaps it is time for organizations to re-evaluate their prior decisions.

For best protection of an organization's information, their networks should be designed (or redesigned) from a security standpoint (King et al.) and not necessarily from a business unit or geographic standpoint. Designing a network from a security standpoint would make unauthorized access to information from internal or external sources very difficult. An example of partitioned or segmented network is shown in the following diagram.
In this diagram, transition to each successive and increasingly sensitive layer as users move towards the center of the diagram requires passage through an increasingly restrictive firewall. This diagram is a simple illustration of a secure segmented network, it is not provided as a one-size fits all solution. Each organization can, and probably will justify modifications to this model based on their risk tolerance, the value of information they are protecting and their relationships with third parties that may or may not require access to certain information.

In real-life implementations of this diagram, only highly trusted administrators and security personnel are ever permitted to directly access the mission critical systems. They are mission critical for a reason; the operation of the business depends on them, so unauthorized access is not permissible. For trusted administrators and other specifically authorized personnel, usage of strong authentication and VPN technology will enable them to reach to whatever depth they wish to get to in the computing environment.

This type of network segmentation is not new. Governments across the globe have used similar solutions for years. But it is only relatively recently that the cost of firewall solutions has reached a level that it is very practical for virtually any organization to implement. In large organizations network partitioning using tens or hundreds of firewalls currently segment their complex computer networks in an effort to severely limit any damage a would-be intruder (internal or external) could inflict on their systems before being stopped at a firewall.

This type of network solution will only be effective if it is supported by strong IT Security Policy, Standards and Procedures. Ensure these are established first prior to investing in this type of solution or the benefit the organization will accrue will be relatively minor. This type of solution, with thought and attention to detail could be employed by any organization, or even your home network.

Unless the information that your organization holds and processes is extremely sensitive and/or valuable, the level of protection that it needs is simply one or two steps higher than that of your competitor organizations. Unless an intruder has a specific motive to access your information, they will most often access that information which is easiest to get at. If you implement a security solution that presents a higher frustration factor than your competitors, all but the most dedicated intruders will move on to an easier target.

**Device hosted firewall/personal firewall**

The ultimate network perimeter is the point where the computer connects to the network. There is an increasing trend to install firewalls at this point — sometimes called personal firewalls. Some solutions are effective and secure and their security settings can be managed centrally, others are dependent on the skills and the knowledge of the user. Typically laptop and desktop users are not security experts, so to expect them to be able to manage their own personal firewall is wishful thinking at best.

The centrally managed personal firewall solution can offer many advantages providing the managers are knowledgeable about security. However, for the power-user, these firewalls sometimes get in the way of them performing their job, so they request (or hack) special holes in the standard firewall configuration. These holes may or may not be a source of intrusion and should be carefully evaluated to minimize any impact on the organization’s risk posture.

**Application security**

So far we’ve only discussed firewalls, network protocols and their usage. At the end of the line, it is the application that protects the data itself. So even if all the firewall infrastructure is perfectly secure, an application that doesn’t respond as expected to either a bogus request, or invalid data, or an unexpected combination of valid requests could result in unauthorized access being granted. It has become very apparent over the last few years that the quality of application code with respect to secure programming principles has dropped significantly. Part of the reason for this may be the demand for features over security, time to market, or simply that secure programming skills are not being taught (perhaps they are not deemed as ‘cool’ skills).

Security related software bugs are not restricted to the end user application alone. They can exist anywhere in the communication chain between the communicating parties. Areas where security bugs may be present include:

- TCP/IP protocol,
- implementation of the TCP/IP protocol containing the bug,
- operating system (with patches) in any of the systems through which the data stream passes,
- firewall software and/or configuration files controlling the operations of that software,
- cryptography software (where used).

A coding error in any of these areas could manifest itself as a security vulnerability that
could be exploited. This is not a simple problem and doesn’t have a simple solution.

As discussed earlier, IP V6 offered a solution that would address many of the security issues associated with sending data via TCP/IP. What it couldn’t and didn’t address was that of poor application coding from a security standpoint. Until the security quality of applications is raised dramatically, the application will frequently be the weakest security link.

Wireless networks and wireless network integration

The introduction of wireless communication technologies for computers has introduced significant issues when trying to protect the network from unauthorized access. This is primarily because wireless networks effectively dissolve the normal perimeter unless the network is properly constructed and managed. Two wireless culprits have been identified as the primary sources for the erosion of the network perimeter. They are:

1. laptops or other computers connected to the corporate wired network with the wireless NIC enabled,
2. inexpensive and unauthorized rogue access points or routers connected to the wired network for the convenience of the user.

In the majority of cases, the insecurity introduced is not intentional, simply a lack of education on the part of the user. Typically the insecurity is introduced and justified as a convenience feature.

The potentially more sinister source of insecurity is the technology itself. Wireless devices (WiFi) are sold with promises of limited range and string encryption to protect the information in-transit. The truth is somewhat more telling. WiFi transmissions can be picked up using relatively simple technologies not 150 feet away (as often seen in WiFi device specifications) but several hundred yards or indeed over a mile away in certain conditions. Non-WiFi specialists think that the transmission of data is point to point — just like a wired network. They forget that it is a radio transmission and that transmits in 360° and three dimensions — now where is the network perimeter. The WiFi vendors and security committees thought of that — they included data encryption. Unfortunately, the specification of that encryption is poor and it can easily be broken. Worse still, many WiFi users don’t even enable it! Now the network perimeter is not only porous, it has been eliminated.

Firewall technologies

Firewall solutions (Mandy Andress; Zwicky et al.) have progressed in many ways over the last 25 or so years from requiring sizable hardware, costly software and services to at the other extreme becoming consumer devices that can fit in the palm of your hand and cost less than $100. Despite this apparent progress, they still depend on essentially the same things: an insecure protocol suite and software (whether loaded on a hard drive or downloaded into firmware).

There are 3 basic types of firewalls:

- **Proxy firewalls**
  - Use rules that not only check the packet headers, but then check the content of the payload for issues if the packet header has been authorized.
- **Packet filter firewalls**
  - Essentially two types: regular packet filters check the header of every packet, stateful packet filtering checks the initial packet header and if permitted simply monitors the other packet headers for packet sequence number consistency.
- **Circuit level firewalls** — of historical interest only, so have been omitted.

An almost religious fanatical marketing battle between proxy based firewalls and packet filter based firewalls has raged for years. In reality, many commercial firewalls are a hybrid of the two. Its just a matter of the proportions of proxy and packet filtering components. Most hybrid firewalls tend to favor one side over the other — it depends on the security beliefs of the development team. The organization buying and using the firewall doesn’t really care, they simply want a firewall that works and protects their data.

Whichever solution is chosen, the firewall software provides a reasonable (but often imperfect) level of security only if they are

- developed with a focus on security of the application itself in addition to providing perimeter protection for the device(s) the firewall protects,
- properly installed and configured according to the organization’s security requirements,
- only installed on adequately secured foundation hardware,
- located appropriately to protect what data requires protection,
- configured to provide the minimum number of services and options to support the business requirement,
actively managed to maintain the security of the firewall itself and its firewall operational configuration.

The protection cannot be perfect since we are trying to absolutely protect something that contains fundamental design flaws — it cannot be done.

These requirements listed above apply both to appliance based firewalls or to firewalls hosted on generally available operating systems and hardware.

When the marketing, promotional and technical niceties are stripped away, a firewall is simply a software application that is loaded on an operating system that permits TCP/IP traffic to be routed from one network interface to another subject to the rules imposed by the firewall software. As such a firewall is subject to all of the software-related issues identified previously. The difference between firewall software and other software is that the firewall developers should have paid much more attention to security than perhaps other software developers do.

To increase the level of perimeter protected afforded by a firewall, it should be supplemented with additional security technologies such as intrusion detection/prevention technologies and malware technologies. This technology combination can be implemented and integrated separately or, as is now becoming more common, integrated into a single security appliance. The former approach can be customized for any requirement while the latter is dependent on the manufacturer’s belief of what should be done. Neither approach is right nor wrong — its simply a matter of what works for a given organization.

Firewalls cannot protect your organizations against:

- session hijacking,
- snooping of network data,
- modification of network data,
- re-routing of network data,
- spoofing of network messages,
- users giving information away,
- users attaching modems to systems protected by the firewall,
- social engineering attacks,
- downloading/accepting files, or messages containing viruses or other malware,
- internal attacks on your data.

The most secure firewall implementations occur when:

- develop a company wide IT security policy (Security policies, standards and procedures, 2002; Greenberg, 2003) and comply with it,
- think like an ‘attacker’ — not a ‘defender’,
- educate users and management,
- use ‘Proactive Security’ both on the firewall and for access control to critical systems.
  - harden the OS supporting the firewall software if possible,
  - disable ALL unnecessary services,
  - use strong authentication where possible,
  - check for ‘back-doors’,
  - check for unauthorized configuration modifications.

Aside from the firewall, a good information protection program (Greenberg, 2003) will also:

- implement and monitor an Intrusion Detection System so that you know when you are being attacked,
- implement incident handling and response process to handle the situation when someone tries or does get in.

At first sight, a firewall may be thought of as a simple plug and play device that you place on your network to protect it. In reality, to introduce a firewall solution and make it work effectively requires a significant amount of knowledge and security expertise.

Next steps

Since you have read this far, it is clear that you have concerns about your existing firewalled environment or an upcoming firewall related solution. The key to resolving concerns is to step back and attempt to take a very analytical view of what you want to do.

Depending on your comfort level at doing this yourself, or through your own organization, you may want to use an independent third party. Whichever is chosen, a very thorough and critical review of your current situation must be done. As with other security related issues, its all a matter of risk tolerance and how much your organization believes it can accept.

Summary

Properly designed, implemented and managed firewall technologies provide significant protection for current computer networks. They are not necessarily the universal security panacea as was sometimes touted in the past. They are simply part of an organization’s armor against unauthorized
intrusions into their computer systems. For now and for the foreseeable future, firewalls will remain a stalwart defense option for organizations. The perimeter on which firewalls can be installed has now reached the device level. It is surely only a matter of time before they are embedded within the devices themselves.

Despite the strengths of firewall technologies, they are being used to attempt to protect data arriving on insecure communications’ links that will be used or generated by applications that expend little or no effort validating that the information being sent or received is indeed valid information. This has resulted in firewall technologies continually facing an uphill battle against an increasingly sophisticated attacker, and an increasingly less security aware application and user. In reality a firewall is only part of the overall security solution; a much stronger protective blanket can be implemented to surround the organization’s data if the firewalls are supplemented with other technologies such as intrusion detection systems, malware protection systems and of course secured applications. The later issue is rapidly becoming a very serious consideration for many organizations and one that increasingly cannot be ignored as the application is the final step before reaching the data itself.

References

Mandy Andress. Surviving security: how to integrate people, process and technology [chapter 7; ISBN:0-672-32129-7].

Dr. Stuart Broderick has been designing, installing, teaching, lecturing and writing about network firewalls across the world for over 15 years. During the last 9 years he has been a guest lecturer at the University of London for the M.Sc. programme in Information Security and has lectured internationally on a variety of security topics in addition to firewall technologies.

From an initially highly focused approach to securing the network perimeter using a single firewall, Dr. Broderick designs and implements information security programs that cover a wide range of security requirements and business operations.

Dr. Broderick is also co-chairman of the US Chapter of the International Information Security Management Systems User Group.