Domain Based Security: enabling security at the level of applications and business processes

K J Hughes

The Problem

Many recent articles in the computer press testify to increasing concerns about potential damage from ‘application level’ attacks. Modern firewalls have improved to such an extent that they can now be configured accurately and reliably to counter network level attacks while permitting commonly required services such as email and web traffic. In the past, would-be attackers have been able to use malformed packets and similar tricks to exploit flaws in the implementation of common protocols, but many of these loopholes have now been closed. Not only have technical solutions to these issues been devised, but the task of firewall management (traditionally difficult and error-prone) has also become much easier. This is because network-level firewalls can now be purchased as ‘appliances’, configured as a single product to supply commonly required protection needs.

However, no matter how well the firewall is configured, the information exchange that a firewall is required to facilitate carries an inherent security risk. Information received through the firewall may carry data that can be interpreted as instructions or code by the applications, which process it. Any instructions executed in this way enjoy the permissions and privileges of the user on whose behalf the applications run. In other words, the sender of information that becomes executable can do whatever the user receiving it may do: access and modify the user’s files, send messages in the user’s name, and use any encryption mechanisms and digital signatures that applications may invoke in order to hide, reveal or authenticate information. Unless the user’s workstation is ‘clamped down’, much more extensive damage may be caused to the wider network and its information and services by exploiting weaknesses in the local network operating system.

Despite the fact that these dangers have been widely and repeatedly discussed in recent months and years, the remedies offered have been consistently inadequate or ineffective. Concerned parties are exhorted to apply rigorous virus checking regimes, to make maximum use of Public Key Infrastructure (PKI) technology and to employ the latest Intrusion Detection Systems (IDSs) – in other words to buy into all available security gadgets in the hope that they may be of some help, whilst maintaining constant vigilance in the hope that an attack can be detected before significant damage is done. In reality, a professional attack launched at the application level would not be resisted by any of these approaches, although any of them may play a key role within an overall security architecture.

Anti-virus software only defends against code that is known to be harmful. It is not intended to prevent the introduction of all mobile code, since this is incompatible with the effective use of the Internet and with many modern applications used for sharing information. Even if the intention is to ban all executable code, this is extremely difficult to achieve in practice. Although it is a simple matter to remove all files with a ‘.exe’ or similar extension, there are many other ways of introducing executables. Many applications support the use of macros in various guises. New versions of applications may introduce such features in obscure or

Copyright © 2002 QinetiQ Ltd. All Rights Reserved.

unadvertised ways. Even worse, there are increasing reports of buffer overflow attacks on applications caused by processing of malformed documents. Given the complexity of modern applications, such attacks are effectively impossible to counter.

Intrusion detection examines information as it traverses a network. A wide range of checks is possible, revealing (for example) where back door modems have been connected, or code fragments in requests that may cause a buffer over-run on a server. However the network activity of an intruder working at the application level is indistinguishable from the activity of a legitimate user. As long as the attack makes use of the user's privileges and does not attempt to 'break out' into the network it will pass undetected, although it can do as much damage as the user in whose name it operates. An IDS does not therefore offer any effective defence against this sort of attack.

Encryption of any sort, including PKI, is of itself inherently incapable of resolving the problem. This issue is discussed at length by Schneier[1], but can be understood at an intuitive level: the applications in use must be able to handle the keys used to sign or encrypt user data, otherwise the user would either be unable to view and process the data or be unable to hide or sign it. Since these keys are available to any code acting in the user's name, no effective defence can be offered.

'Vigilance' is somewhat effective when an attacker wants to be noticed, which is something that most hackers appear to want. However, a person launching a professional, targeted attack would choose to hide their activity and modern systems offer many opportunities for concealment. General vigilance is therefore of limited use against a serious, targeted attack.

Security to meet Business Needs

All of the above approaches are seeking a solution that is independent of enterprise objectives and business processes. If there is no requirement to do business with people except those who can be completely trusted, then the conduct of this business can be isolated from the rest of the world using a combination of physical separation and cryptography. Such techniques are entirely generic. However, there is an increasing need to use IT solutions to conduct business with parties that can only be partly trusted. Under these circumstances, it is necessary to share information with people while defending against their actions. Such people may work within the organisation or outside it. Effective security in these circumstances requires a combination of separation and controlled sharing.

While information separation can be achieved with the general-purpose methods of physical separation and cryptography, the need for controlled sharing is not a generic problem and there is no panacea. The only available course of action is to focus on protecting what is most important for the organisation. Each organisation must define its own priorities, identifying:

- What information must be shared, in what ways and with whom?
- What information and services need to be protected, from whom, against what and to what degree?

On this basis, the organisation must develop specific security requirements. These requirements define constraints and controls over the ways that information may be shared. They must permit those interactions required to conduct the operations of the organisation and defend against those eventualities that are considered most damaging. They must be defined in terms of the business processes of the organisation, its staff, the information they work with and the people they interact with. Using this approach, it becomes possible to position security controls where they support rather than interfere with the business processes. Crucially, controls placed in this way should allow responsible staff members within the organisation to control what happens to critical information and services. They should also enable these staff members to be held accountable for their actions.

[1] Bruce Schneier, 'Secrets and Lies: digital security in a networked world', Wiley 2000.

For some organisations confidentiality is a key concern, so the actions that must be controlled are those that permit the selection of information for release to ensure it does not contain sensitive information. If the integrity of particular transactions is paramount, then actions that commit those key transactions must be controlled. If message authentication is important, the 'From' field of the message, as displayed to the recipient, must match the actual identity of the user sending it, so that it cannot be spoofed; the release of the message and confirmation of its contents must be under the direct control of the sender.

Users cannot be held to account for their actions if the software that carries them out is unpredictable or may be modified in unexpected ways. The actions that an organisation identifies as critical must therefore be carried out only by 'well behaved' software that is under the direct control of the person responsible for them. The organisation must be selective, because to be successful such measures require the operators to check their actions carefully before confirming their intention. Therefore, the provision of such measures is not only costly, but may be disruptive to business processes unless they are deployed appropriately. People will become careless if they are called upon to make unnecessary checks and the security measures become intrusive.

This is the theory that underpins Domain Based Security. To be successful in practice, the approach requires:

- Methods to help the organisation develop and specify the constraints and controls required to support its particular protection needs, and to ensure that they are applied across the enterprise as a whole;
- Cost effective mechanisms that provide targeted protection for security-critical actions initiated by users. These mechanisms must work at the application level, because only the applications can display the information in ways the users understand. They cannot be part of the applications because they must behave in known and trusted ways. They must be sufficiently generic to work even when the business use of an application varies.

QinetiQ Trusted Information Management has addressed both of these needs, but this paper concentrates on the former.

Policy and Architecture

It is generally accepted that effective security requires an enterprise-wide approach. The key drivers for security are defined by the organisation's Security Policy, which is determined through an assessment of the values of certain assets or services, the potential cost to the organisation should they be compromised in specific ways, and the likelihood that such compromise might be instigated by specific groups of people. These considerations determine the degree of protection to be afforded to prevent or mitigate the different kinds of compromise identified.

Ideally, policy should be high level and enduring. It should be concerned with the ends, not the means, so that it remains sound in the face of changes in technical solutions and business processes. At the same time, it must be sufficiently meaningful to be interpreted in consistent ways. Categories of assets must be defined in terms of the value systems that are appropriate to them. Levels of protection must be defined in terms of the kinds of mistake that might lead to compromise, as well as the motivations and skills of possible groups of attackers.

Consideration must be given to the cost of protection as well as to the cost of compromise, but simple allocation of funds for security in proportion to the importance of each kind of compromise is not the answer. The benefits of security provision are seldom in simple proportion to their cost: rather they accrue in a step-wise fashion. Significant benefits are achieved through clear assessment of need followed by proper implementation of those measures identified as effective for the required purposes. Little or no benefit accrues from half-hearted, unfocused efforts, although significant costs may be incurred and much damage may be done to operational effectiveness. A holistic approach is therefore required, to achieve cost-effective and coherent security solution to meet enterprise objectives.

Copyright © 2002 QinetiQ Ltd. All Rights Reserved.

For a large organisation, IT is acquired in a piecemeal fashion. It is impractical to build, manage and replace the entire infrastructure and its applications as a single entity, although it is likely that most of it will be interconnected in some way. Changes are likely to be implemented as individual projects, often developing in parallel with each other. However the security implications of a project may extend far beyond the project boundary. Indeed, the provision of any additional functionality is prone to allow existing security controls to be bypassed and this may have far-reaching consequences. Hence an IT project to provide support for one part of the enterprise might easily undermine the security required to protect the assets of another part, which the project members are entirely unaware of. Typically this sort of problem only comes to light when a project is at or near completion and it is too late or too expensive to repair the damage in a satisfactory way.

To overcome this problem, an enterprise-wide security architecture is required. The architecture must be defined using a language that is security-relevant. It should be concerned with requirements for separation that must not be violated by added functionality, as well as with controls over the exchange of information. The architecture must also be defined at the application and business level. There are two clear reasons for this:

- Constraints defined below the application and business level can always be violated by new applications that exploit the permitted functionality in different ways. Constraints defined at the application level, however, impose specific requirements on all lower levels to support only the functionality specified at the highest level and prevent all other interaction.
- Constraints defined in terms of technical solutions become outdated when new technology becomes available.

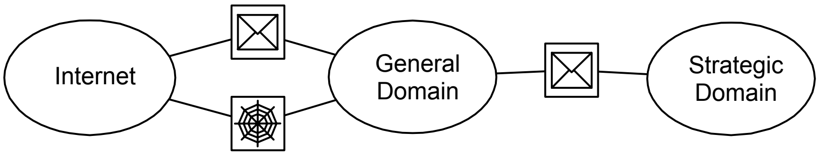
Infosec Architecture Models

Domain Based Security uses a graphical modelling technique to define the security architecture for an organisation. The same technique can be used by projects to define their specific security requirements and show that they comply with the architecture.

The technique uses two different but related views, which when combined provide an Infosec Architecture Model. The first is a business level view, which represents requirements to control the way information is exchanged between different groups of people. This is referred to as an Infosec Business Model.

In this model, domains represent groups of people and the information they work with using a computer system. Connections between domains are defined in the model where there is a need for the people working in the domains to interact and share information. Superficially therefore, it looks like any other business interaction model, except for a simple but crucial difference. Whereas conventional interaction models show minimum required connectivity, Infosec models show the maximum connectivity that is consistent with security requirements.

For example, in the model shown below, an organisation has defined two distinct domains in which its employees may work. The Strategic domain is used to handle highly sensitive information requiring special protection. In this model, people working in the General domain may use Internet email and browse the World-Wide Web, as well as exchanging messages with people working in the Strategic domain. Importantly, however, this model requires that there shall be no interaction between users working in the Strategic domain and the Internet, as this is considered too risky. It also stipulates that interaction with users in the General domain must be via the permitted message connection, for which special controls may be defined.

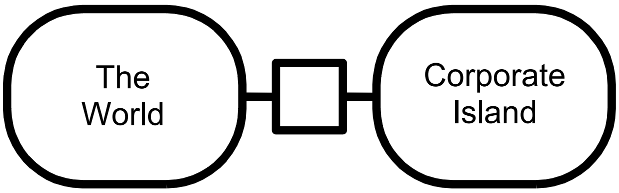


Copyright © 2002 QinetiQ Ltd. All Rights Reserved.

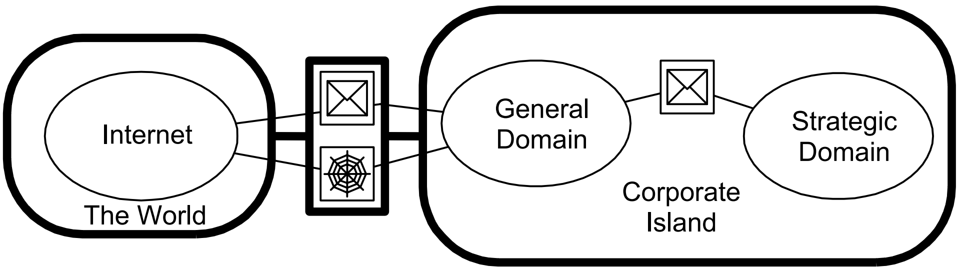
This model provides a useful medium for discussing and achieving an appropriate balance between two key sets of business imperatives: the need to exchange and share information and the need to protect information and services. Domains help to focus on the information and services that need protection, Connections must clearly be adequate to support business processes. However, connection itself represents an inherent source of risk, where users working with critical or sensitive information need to interact with users from whom this information is to be protected. It is clear that the exchange of information must be controlled on such a connection, to reduce the risk that legitimate interaction does not enable corruption or disruption of the system by outside influences or inappropriate release. Using the models, the positioning of security controls can be determined on a rational basis, giving confidence that they will be effective. The models also help to achieve user buy-in for security, by enabling users to understand why the controls are needed.

The second view models the infrastructure required to provide strong boundaries and support the implementation of non-bypassable controls. This is referred to as an Infosec Infrastructure Model. The need to provide clearly separate networks for different purposes can carry a substantial cost. It is important therefore that this cost is incurred only where required and that the investment is made to work to the greatest possible advantage.

The model distinctly shows the islands of infrastructure, which have clearly defined separation from all other islands. Each island may be geographically distributed, using cryptographic separation, but the cryptography, combined with physical separation, forms part of the island boundary as it does not allow any information to be shared. The model also shows the ‘causeways’ between islands: clearly identifiable and manageable points of connection that provide the sole means of exchanging information between islands. In the diagram below there are two islands, linked by a causeway. This design requires that all interaction between the ‘Corporate’ infrastructure and the other systems must be via this managed interface, so that the ad hoc connection of modems for outside connections is explicitly forbidden.



The Infosec Business Model can now be superimposed on the islands and causeways of the infrastructure model, to show which domains are supported by each island of infrastructure. Causeways implement the connections between domains that are hosted on different islands. Their role is to implement the controls that are required on the business connections and to ensure that they cannot be bypassed.



Copyright © 2002 QinetiQ Ltd. All Rights Reserved.

The result is an Infosec Architecture Model. It provides a clear mapping between the interactions to support business processes and the constraints and controls that provide protection for critical and sensitive information and services. It is concerned with both the business level and the network level, ensuring that strong controls imposed at one level cannot be undermined by weaknesses in the other.

Conclusion

Domain Based Security is not an ‘Off the Shelf’ solution to the dilemma of application level security. It does however provide the methods and tools for an organisation to create its own strategy to tackle the problem. It enables the organisation to define its priorities for security and position its defences such that they provide security where it is most needed, while offering minimal disruption of the conduct of necessary business activity. The security architecture must be tuned to specific security requirements, to achieve a balance between the need for information exchange and the protection of key assets, and to place controls where they work best and where users can operate them effectively. The approach also identifies clearly the need for an implementation that enables responsibility for the protection of information to be placed where it belongs: with the people who generate it and handle it.

Customer Contact Team
QinetiQ
Cody Technology Park
Ively Road Farnborough
Hampshire GU14 0LX
United Kingdom

Tel: +44 (0) 8700 100942
www.QinetiQ.com

Copyright © 2002 QinetiQ Ltd. All Rights Reserved.