

Protocol Interactions and Their Effects on Internet-Based E-Commerce

Hans Kruse

J. Warren McClure School of Communication Systems Management
Ohio University
hkruse1@ohiou.edu

Abstract

Internet protocols are based on the concept of protocol layers. This layered design is intended to provide needed information to each type of network device independent of information required for other devices. However, a number of special circumstances have led to the creation of devices that rely on information from protocol layers that they would not normally access. These so-called "layer violations" have been possible because security services had been rarely used in the early internet. Electronic commerce applications on the other hand rely heavily on security, both to authenticate information flows and to keep them confidential. Devices attempting to access information from higher protocol layers may fail depending on the type of security that is in effect, and depending on the placement of the device within the network. In this paper we develop a classification scheme, or model, for the types of devices that interact with information outside their normal protocol layer. We also develop a model for the administrative network structure that defines where security systems may be deployed. The combination of these two classification schemes leads to a set of deployment rules for points where layer interaction can be successfully implemented.

1. Introduction

This paper addresses compatibility issues which may arise when the Internet is used to widely deploy electronic commerce applications. The Internet protocols are designed in a layered fashion that is supposed to separate various types of network functions and limit the possibility of harmful interaction between these functions. However, there are several examples in the Internet today where "violations" of the layered model have been introduced for good reasons, and in

full knowledge that compatibility problems may arise. For example, Network Address Translators [3] [8] [1] (NATs) were introduced to delay the exhaustion of the IP version 4 address space, and to ease a user's transition from one network provider to another. At the same time, NATs do cause compatibility problems and prevent the deployment of some end-to-end security models. As the Internet expands it increasingly encompasses diverse networks, including terrestrial wireless systems such as packet radio, cell phones, and PCS, space-based systems including geostationary and low-earth-orbit satellites, and a variety of symmetric and asymmetric access networks at the customer premise. Some of these network types rely on transport layer gateways to improve performance for typical Internet users. These gateways violate the separation of layers in ways that are similar to NATs, except that the gateways may be deployed deeper inside the network. The use of LAN emulation over Frame Relay and ATM as well as the support for mobile TCP/IP [7] [9] users further complicates the seemingly simple layer structure and end-to-end semantic of TCP/IP. Electronic commerce applications commonly require the implementation of procedures to insure confidentiality (usually via encryption), authentication and non-repudiation, and availability (especially the prevention of denial of service attacks). In many cases these systems are based on end-to-end semantics that rely on the transport layer information remaining invisible and unchanged in the network. It is well known that many of the "layer violating" systems outlined above either hinder the deployment of a security infrastructure, or they are rendered useless by such a deployment. We do not imply that the interactions outlined above are unknown, nor that they have been ignored. In fact, each of the protocols and procedures mentioned above is described in either an internet standards document, or an informational RFC; each of these documents contain extensive descriptions of possible protocol interactions. The purpose of this paper is not to propose protocol changes, or present major new protocol interaction problems. Rather, it is our aim to develop a classification scheme and reference model that allows for an organized planning process prior to the deployment of security and devices requiring layer interactions.

2. Layer Interactions in Network Devices

A network device outside a host should normally operate at or below the network layer of the protocol stack, e.g. the IP layer in TCP/IP. For the purpose of this paper we use the term "layer interaction" for any device function that requires access to information above the network layer. We organize devices



that introduce layer interactions within the network into a grid with six entries as shown in table 1. These devices make use of information housed in headers that belong to protocol layers above the network layer.

	Optional Access to Transport Layer Information	Mandatory Access to Transport Layer Information	Access to Application Layer Information
Read/Modify	Type I/mod	Type II/mod	Type III/mod
Read Only	Type I/read	Type II/read	Type III/read

Table 1

Classification of Layer Interactions in Network Devices

In our classification scheme, we distinguish between devices that act on information read from the headers of higher layers, and devices that potentially modify information in higher layer headers. In this context we therefore include in our definition of Layer Interactions the operation of devices that seek to change fields in the network layer information that are normally considered fixed, such as the origin and destination addresses.

In our scheme, Type III devices reach all the way into the application layer information. These devices are typically not transparent to the user, and their use is in most cases required for the correct operation of the application. While we allow for a device of type III/read (a device that only reads application layer information but passes it on unchanged), we believe that this type is probably quite rare.

Type II devices only need access to the transport protocol header, such as UDP or TCP headers. However, a type II device is required for the correct operation of some or all of the network.

Type I devices are inserted into the network to improve performance. However, their use is optional and bypassing a type I device does not render the network or the application unusable. Bypassing the device may, however, carry a performance penalty.

We explain examples of some of these devices below:

- Type III: Web proxies and proxies found in many firewalls fall into this category. More important to our discussion are the Application Level Gateways (ALGs) used in conjunction with Network Address Translators

(NATs), for example for the translation of embedded IP addresses in FTP commands.

- Type II: The most prominent device in this area is the Network Address Translator itself. Most NATs perform one-to-many translations in which the a local host address and a host's TCP port are translated into a single global IP address and a globally unique TCP port number. This type of NAT requires read and modify access to the port numbers in the TCP header. Note, however, that even in the case where the NAT translates one-to-one between local and global addresses, the modification of the IP address in the packet constitutes a layer interaction in the context of our definition of the term. A packet-filter based firewall is an example of a type II/read device in that it cannot perform its function without access to the transport layer. Its use is mandatory because it will not allow unrecognized packets to pass. The firewall will, however, not modify packets it allows to pass.
- Type I: Many proposed and implemented Performance Enhancing Gateways (PEPs) for satellite and wireless links are in this category. This includes gateways that split TCP connections (if they can be bypassed, otherwise these are type II devices), and devices that aim to prevent acknowledgement congestion on asymmetric links. Many routers are capable of priority queuing procedures based in transport layer information (usually port numbers), which makes them type I/read devices.

3. Protocol Structure in the Presence of Security Systems

Security systems in the Internet fall into two categories. Applications can choose to implement security above the transport layer, either through a standard protocol like Transport Layer Security (TLS) [2] , or in a proprietary way. In this case, the transport layer protocol header is neither authenticated nor confidential. In the terminology of section 2, only type III devices are effected by TLS or similar security systems.

Considerable effort has been directed to the deployment of a more comprehensive, and less application-specific security infrastructure. This combination of security implementation guidelines, key distribution protocols, and additional network layer protocol headers is typically referred to as "Secure

IP” or IPsec [4] [5] [6] . The concept of VPNs¹ is largely based on deploying IPsec. IPsec can be deployed using authentication, which is designed to allow a recipient to detect unauthorized modifications of the packet, and/or confidentiality (i.e., encryption).

IPsec may be deployed in Transport Mode, meaning that two hosts establish secure communications directly with each other, or in Tunnel Mode. In Tunnel Mode, at least one of the endpoints of the secure communications path is a gateway that adds security to traffic passing through it on behalf of several hosts. For simplicity we will assume that both ends of a tunnel mode association are gateways in the descriptions below.

Table 2 summarizes the access that is possible at the network and the transport protocol layer for each of the deployment options.

Note that authentication and encryption can be combined with the effect of removing the ability to modify indicated in the encryption entries in table 2. In principle it is also possible to apply IPsec more than once, creating more than two sets of network layer protocol headers. If authentication is used, all headers are protected against modification (but readable). If encryption is used, only the outermost header is readable.

Deployment Type	Access to the Network Layer Information	Access to the Transport Layer Information
Transport Mode Authentication	Read-only access is possible. Modifications ² will result in the packet being discarded.	Read-only access is possible. Modifications will result in the packet being discarded.
Tunnel Model Authentication	The network layer header is designed to deliver the packet to the destination gateway. Unless a device knows to look past the authentication header, it will not “see” the true network-layer information. Both the “outer” and the “inner” protocol header are read-only and cannot be modified.	Read-only access is possible. Modifications will result in the packet being discarded.
Transport Mode Encryption	Reading and modifications are possible	No access.
Tunnel Mode Encryption	Only the “outer” header is accessible (read and modify).	No access

Table 2

Summary of access to protocol headers when IPsec is in effect.

¹ VPN, or Virtual Private Network, is a much overworked term. We refer to carrier offerings that allow corporate IP-based traffic to flow securely over public network facilities, including the Internet.

² IPsec will permit modification of header fields that are normally changed by network layer devices; we refer here to the modification of fields that are normally fixed.

4. A Model of the Physical Network Structure

The physical path taken by the information flow - as it leave the workstation of the end user – can be divided into 4 categories:

Network Attachment	This is the connection between the workstation and the network. This can be a network interface card, a modem, or a wireless link.
Local Area Network	This is the first shared network encountered by the information flow. It typically serves a small, homogeneous user community.
Local Area Backbone	This is the aggregation network which combines traffic from many LANs and from many different users.
Wide Area Network	The portion of the network that passes out of the immediate geographic area.

Note that for some users the two local area categories may not exist; for a dial-up user the Network Attachment leads directly into the Wide Area Network. This portion of the model helps to define the point at which the end user perceives security needs. Encryption needs to be deployed before information enters a portion of the network where eavesdropping can take place and would cause harm. This is typically either Wide Area Network (this is where firewalls are usually deployed), but it can be the Local Area Network or the Local Area Backbone if the information is sensitive enough. Finally, eavesdropping can take place in the Network Attachment if a wireless system is used.

Authentication must be deployed before the information flow reaches a point where forged packets can be inserted; for simplicity we will assume that this is the same as the onset of the need for encryption, but this does not need to be the case.

The physical network categories are used here mainly to determine the security mode that is likely to be deployed. Users whose security needs begin at the Wide Area Network are able to deploy Tunnel Mode security, since they can permit information to flow in the clear over the local portions of the physical network. Users who must secure information in one of the first three categories will usually deploy TLS or Transport Mode IPsec. The dial-up or wireless access user is an exception; these users often are placed at one end of a security tunnel that terminates at the interface between the Wide Area Network and the Local Area Backbone of the target site.

Section 5 below discusses the administrative structure that governs the deployment of security relationships, especially in cases where security tunnels are used.

5. A Model of the Administrative Network Structure

We divide the administration of the network into 4 levels, or areas. We envision these areas are embedded inside each other, i.e. in the most general case information will pass from one End User area through a Local network, a Regional network, a Global network, a second Regional network, a second Local network, and into the destination End User area.

End User	A network area which contains either a single workstation or a very small network of workstations. These workstations can act as clients or as servers. However, the end user has full control over all security policies applied to this area.
Local Network	This area serves a small number of end-user areas. Resources in a local area may include one or more of the following: non-authoritative name server, authentication server storage server small-scale caching server This area may administered by a separate entity which is however trusted by all End User areas contained in it. Local Network areas can have security trust relationships with other Local Network Areas which can be reached only by traversing Regional or Global areas.
Regional Network	This type of area contains the resources normally associated with a single organization in a multi-organization network, or an Internet Service Provider in the public Internet. In developing nations, most or all of a country may form a regional network. In the US a college campus or a large company would be examples of regional networks. Resources in this region are: domain name server,

authentication server,
large-scale caching server
intranet server (information of interest to this region only)

Administration of the Regional area is separate from the Local area, and there is in general not a (security) trust relationship between the two levels.

Global Network

The top level of this hierarchy represents the “network-at-large”; i.e. it is traversed by flows which cannot be served in the regional network. Resources placed here limited, although the root name servers may be considered to be in this area.

No security assumptions can be made for this network area. Any desired security must be applied to traffic before it flows into the Global network.

Application level security like TLS, and IPsec in Transport Mode are deployed between devices in the end user areas. Tunnel mode IPsec is deployed between devices which are placed at the boundaries between areas. Since the gateway creating the security tunnel has access to the information flow, it must be deployed at the edge of the local area or the end user area.

6. Rules Derived from the Model

Sections 2 through 5 define various aspects and classifications of the reference model. We can now derive a number of rules that govern the deployment of network-based applications that use both security and layer interactions in the network devices. The rules are ordered from more general ones to more specific ones. We do not claim that this list is exhaustive.

1. Devices of type I/read, II/read, and III/read can be deployed anywhere if only authentication security services are in use.
2. TLS and TLS-like authentication security is compatible with all devices except Type III/mod, unless the type III/mod device is deployed in the same Local area as the end user it serves.
3. TLS and TLS-like encryption security is compatible with all devices except Type III (both /mod and /read), unless the type III device is deployed in the same Local area as the end user it serves.

4. Transport Mode Authentication prevents the use of Type II/mod and Type III/mod devices and prevents the information flow from using the services of a Type I/mod device.
5. Transport Mode encryption is compatible only with Type I devices, and renders their services unusable for the encrypted flow. In some cases a high percentage of encrypted flows (which have to bypass the type I device) may have a sufficiently negative impact on the network to cause the deployment of type II devices.
6. If Tunnel Mode encryption is in use, devices of Type II and III must be deployed in the same Local network area as the end user they serve. Type I devices need to be deployed in the same Local network area as the end user to be able to provide a useful service.

We note that adherence to these rules is easier in some cases than in others. For example, NATs are typically used at the edge of a Local network area anyway. As long as tunnel mode security is sufficient, and the tunnel is created outside the NAT, the network will function correctly. On the other hand, PEPs used for wireless links are currently most likely deployed by the carrier who provides the wireless link. The deployment of a PEP in the Regional or Global area defined in section 5 prevents the use of either transport or tunnel mode security, although TLS is possible.

7. Conclusions and Future Work

We have presented a network application reference model into which we can organize a variety of protocols and requirements. The model classifies technical interfaces within the network, which drive deployment decisions for secured applications. It also identifies the importance of boundaries of administrative control. We present a series of rules that define deployment strategies for network devices with layer interactions in the presence of security requirements. It is our hope that this model can help both network providers and application owners deploy internet-based systems without having to analyze a large number of protocol specifications in detail. The application owner will be able to determine which protocols and procedures can be used without reservation, and which ones carry performance and interoperability risks. The network provider can determine how to structure a network service offering with a basket of protocols and services that maximize user and network performance while minimizing interoperability problems.

This paper primarily defines the model and its components and derives a set of basic rules for deployment. To obtain the benefits indicated above, we need to apply the classification schemes defined here to a set of specific networks and devices. This work is currently in progress. We expect that in the course of this work more detailed deployment rules will emerge which can be added to the model.

Even the basic rules presented in this paper make it clear that the deployment of security and certain types of network devices are largely incompatible, leaving network managers and network providers a limited set of options. The rules we have derived make it clear that the deployment of certain types of devices (PEPs are the most prominent example) needs to be move towards the end user, if security is to be used effectively. Further work is needed to develop strategies to facilitate this type of deployment.

Bibliography

[1] RFC 2428 FTP Extensions for IPv6 and NATs. M. Allman, S. Ostermann, C. Metz. September 1998³

[2] RFC 2246 The TLS Protocol Version 1.0. Dierks, T. and C. Allen, January 1999.

[3] RFC 1631 The IP Network Address Translator (NAT). K. Egevang, P. Francis. May 1994.

[4] RFC 2401 Security Architecture for the Internet Protocol. S. Kent, R. Atkinson. November 1998.

[5] RFC 2402 IP Authentication Header. S. Kent, R. Atkinson. November 1998

[6] RFC 2406 IP Encapsulating Security Payload (ESP). S. Kent, R. Atkinson. November 1998

[7] RFC 1853 IP in IP Tunneling. W. Simpson. October 1995.

[8] RFC 2391 Load Sharing using IP Network Address Translation (LSNAT). P. Srisuresh, D. Gan. August 1998.

[9] RFC 2341 Cisco Layer Two Forwarding (Protocol) "L2F". A. Valencia, M. Littlewood, T. Kolar. May 1998.

³ All documents in the bibliography are available on-line at <http://www.rfc-editor.org>