Internet business security — prudence or paranoia?

Tim Roper

Introduction

The Personal Computer (PC) revolution resulted in a proliferation of hardware and software, much of it more affordable and powerful than any of the information technology (IT) that preceded it. A PC on every desktop and personal productivity software priced in the hundreds of dollars gave IT to the masses. But this empowerment brought with it new problems: "computer virus" is now a household term. To assume that we have survived its onslaught is to ignore the price already paid as well as to ignore the new strains and species that have not even been invented.

The Internet offers us a revolution in distributing IT. It promises to be even bigger than the PC in its impact on business and society. Similarly, it promises to threaten our security in even bigger ways.

Comparing the revolutions

The PC and modem revolutions were characterised by at least the following:

- affordable access to technology, both hardware and software
- affordable access to information, both local and remote
- the evolution of networks, both local and wide area.

The rate of growth in these areas was impressive, especially when compared to progress during the previous eras of stand-alone mainframes in closed IT departments.

On the other hand, the following undesirable aspects became apparent:

- the unprotected nature of the PC architecture and its consequent susceptibility to attack
- the inventiveness of hackers (and indeed a change in meaning for that word)
- the epidemiological manner in which computer viruses spread, sometimes in spite of protective measures.

As with most popular new technologies, increasing volumes and decreasing prices drove each other to turn the PC and modem into commodity products.

If there is one difference between the PC and Internet revolutions of particular significance to security, it would appear to be the rate at which the number of systems connected to the Internet is growing compared to the previously impressive growth in the PC market.

At the risk of being simplistic, the ability of viruses to spread can be attributed to the existence of a large number of susceptible (compatible) bodies and a series of unprotected contacts between pairs of them. How does the Internet compare to this? Very closely it seems. The large numbers of bodies are already there, thanks to the PC and modem revolutions. Many of them are susceptible to the same threats by virtue of their compatibility (eg. PC with PC, and UNIX system with UNIX system). For those that are not compatible, the requirement that they support a common protocol suite (TCP/IP) for connection to the Internet added compatibility where there was none. As for the series of contacts, that is what networks are all about. We are left to consider whether these contacts can be protected and at what cost.

The PC revolution experienced a previously unknown level of compatibility. Technical information as well as hardware was widely available to the masses. What wasn't available was reverse engineered and made available. That this happened was due in a large way to the sheer number of inventive people with sufficient interest and free time (or other motivation).

One consequence was the "success" of the computer virus as a technical and social phenomenon. Given the even larger growth rate, compatibility, and number of educated inventive users that the Internet enjoys, we can reasonably expect security problems of a magnitude only hinted at by the computer virus.

Technical background

This section attempts to provide minimal technical background, sufficient to understand the following sections fully.

Definitions

By an **Internet** we mean a collection of often disparate physical networks connected by gateways over which common applications can run by virtue of a common foundation protocol. This is known as the **Catenet** model (Cerf, 1978) and the common protocol is the Internet Protocol, or just "IP" (Postel, 1981a). Hence internet refers to a class of network.

On the other hand, by **Internet** we mean a specific instance of an internet, namely the catenet consisting of NSFnet, AARnet, and other national and international networks plus the networks of Internet Service Providers (ISPs) and their customers world-wide.

The Internet Protocol

The version of IP currently deployed is version 4. Together with the higher level protocols, such as TCP (Postel, 1981b), it combines a high degree of functionality together with ubiquitous acceptance by computer manufacturers. It has survived and prospered during its 14 year life and it is a credit to its designers that it is now providing the basis for the exploding international network we know as the Internet.

The salient technical feature of IP that we need to observe here is that it provides a connectionless service. The data that it carries for us is broken up into packets, or *datagrams*, and which are carried over whatever currently appears to be the most suitable route from sender to receiver. Datagrams are autonomous chunks of data in that:

- they each carry their sender's (*source*) address and receiver's (*destination*) address
- the datagrams making up a logical piece of data may, in general, be routed differently through an internet on their way from sender to receiver.

Figure 1 illustrates a simple network in which there are two possible routes from sender A to receiver B, via C and via D.

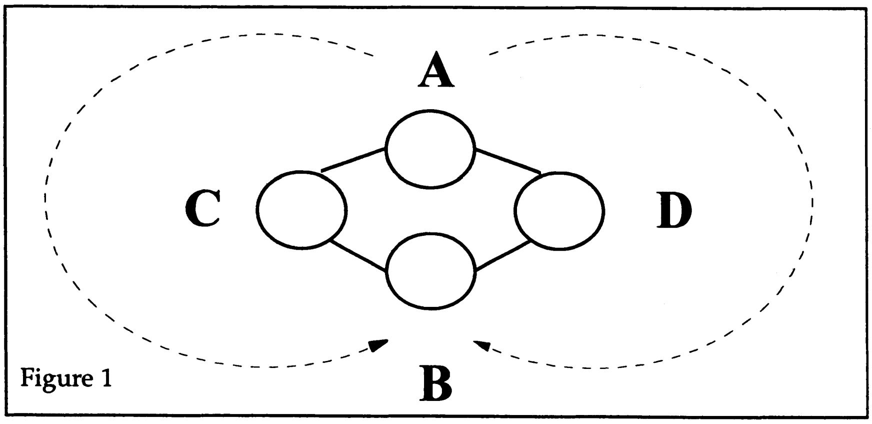


Figure 1

It is not surprising, given its age and the extent to which it has been scaled, that IP version 4 has some deficiencies. The best known of these is its 32 bit address size limit, and its partitioning of those bits, which we now see as sub-optimal. However, it is another deficiency with which we are concerned here, namely authentication. There is nothing in an IP diagram that provides any assurance that the source address contained in a received datagram is really the address of the sender. There is also nothing that reliably tells the receiver by which of the possible routes the datagram followed in its trip from the alleged sender.

So, given a datagram received from an apparently legitimate source, we face at least the following questions.

- From where did the datagram really come?
- What happened to it on the way?

In the general case, we have no answer. Note that data encryption at the application level does not protect the privacy or security of the datagram addressing and other control information.

Types of attack

We distinguish three ways in which an Internet user's property (eg data) may be compromised by an outsider via the Internet:

- file oriented
- remote access
- interception.

I now describe these mechanisms further. Note that this is by no means a taxonomy of the known methods of attack on a computer system, or the data stored therein. Rather, it is simply a distinction that is useful later in this paper.

File oriented attacks

By a *file oriented attack*, we mean the class of attacks that can be launched by the once-off transfer of a file. Such files may be transferred by diskette, down-loaded from a bulletin board, or attached to electronic mail (*e-mail*). The best known of these mechanisms is the virus, a term used popularly to describe any undesirable action taken by a computer system against itself resulting from the introduction of a file from outside. More correctly, virus refers to such a mechanism that both causes damage to the computer system it infects and which is capable of reproduction.

This form of attack is facilitated by sharing of software and data via diskettes and bulletin boards, or simply by the re-use of infected diskettes or hard drives. It pre-dates the Internet revolution, being founded in the PC and modem revolutions. However, global, shared networks such as the Internet provide a quantum leap over diskettes and bulletin boards in their provision of simple methods for distributing files.

Historically PC viruses have been transmitted via disk control information (boot sectors) and executable files (.EXE, .COM, .BAT, etc.). More recently it has been demonstrated that the powerful macro facilities provided by modern word processors are in fact general purpose programming language with more ability to affect the computer system than the application itself. The market for virus detection utilities does not appear to be drying up.

Remote access attacks

Whereas by*file oriented attack* we refer to a once-off opportunity to transfer virulent material onto a target computer system, by *remote access* we mean that the attacker achieves some form of "on-line" access to the target system and uses that to exploit weaknesses in the system otherwise only available to in-house personnel or through breaches of physical security.

Interception

With *file oriented* and *remote access* attacks we were concerned with intrusion into our computer system. In the case of *interception* we are concerned with what happens to our data when it is not stored in our computer systems but rather when it is in transit between our system and that of a service provider or consumer.

A simple example of this is the privacy breach that occurs when a third party intercepts e-mail as it is routed from sender to receiver. Of course it may become more than a privacy issue, for example where the e-mail discloses trade secrets or credit card details.

Internet access options

There are four models for accessing the Internet applicable to businesses:

- the shell account
- the proxy server
- the end-node
- the inter-connection

The Shell Account

The shell account is the simplest and cheapest form of access to the Internet. The business is allocated an account on a computer system operated by a service provider, and connects to that system by some means *other* than TCP/IP. The means of connection is typically via PC with terminal emulator and modem, but includes dropping by an Internet café.

Because the IP datagrams stop at the provider's computer system, no remote access attack is possible on the PC. However it is common under this model to down-load files from the provider's system to one's own, using serial line protocols such as *kermit, X-Modem*, etc., so that *file transfer* attack is open to be exploited.

The Proxy Server

The proxy server method refers to one party providing an Internet-based service on behalf of another. Typically the first party is an ISP, providing World Wide Web (WWW) and File Transfer Protocol (FTP) servers on behalf of the second. This model locates the technology and skills with the ISP. The second party provides the first with the material to be published, for example Hyper Text Mark-up Language (HTML) files in the case of a WWW server or arbitrary file types for an FTP server. The proxy server model is illustrated in Figure 3.

The means by which the second party transfers the material to the first may be by disk, tape, modem or other media. In some cases it may be in hard copy or conceptual form only and the ISP provides or sub-contracts a WWW authoring service.

The End-Node

Whereas the Shell Account terminates the IP datagrams at the service provider's system, this access model terminates them at a computer at the business's premises. This is often a PC running Microsoft Windows and Internet access software. Again, the connection to the ISP is by dial-up modem, however in this case IP datagrams are encapsulated and transferred via the modem in place of the simple character stream in the case of the Shell Account.

This model gives the business's PC full access to the Internet by running clients such as WWW browsers and FTP clients. By running suitable server software, it can also make information available to other Internet users, such as customers of the business. (Doing this would benefit from the use of a suitable multi-tasking operating system on the PC, such as Windows NT or UNIX, rather than a single-tasking operating system such as MS-DOS or Windows). The End-Node model is illustrated in Figure 4.

Figure 2

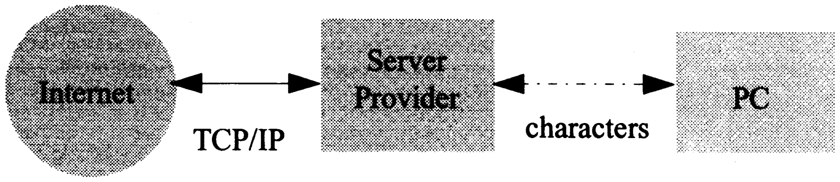


Figure 3

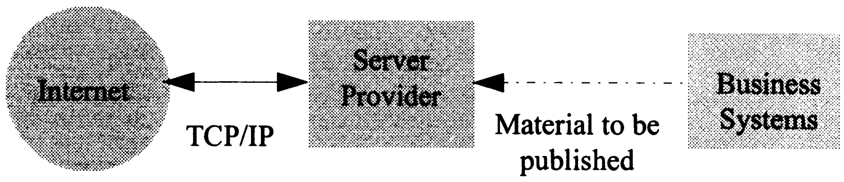
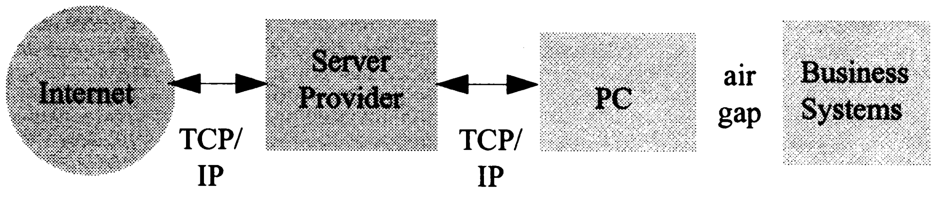


Figure 4



It is fundamental that this model does not provide any other computer systems at the business with direct access to, or from, the Internet. The "air gap" shown in Figure 4 is a critical component in securing the business data from attack by remote access.

The business data is still susceptible to file transfer attacks since it is expected that selected software or data will be transferred to or from the Internet by diskette or other forms of "sneaker net".

It is important to be, and to remain, aware that although the purchase price of the air gap is small, its cost of maintenance can be very high. It is extremely easy and tempting to connect the end-node PC in Figure 4 to the corporate Local Area Network (LAN) in order to simplify the problem of moving data and software files between the end-node PC and the remainder of the business systems. Network administrators charged with maintaining the air gap will find themselves frequently challenged by staff who accuse them of wasting their (the staff's) time.

The Inter-Connection

The fourth, and most powerful and risk prone, model for connection to the Internet is the Inter-Connection. In this case the business's LANs and attached computer systems are directly connected to the Internet via a router (aka gateway). This has the advantage of allowing Internet applications such as Web browsers to run on any desk-top computer in the business, and of allowing Web servers, for example, to run on any suitable system. The disadvantages should now be clear: access from the Internet to any of the business systems is now possible. Whilst this is actually an advantage for staff travelling or working from home (telecommuting), it is also an advantage for unauthorised users of those business systems.

This type of connection to the Internet is recommended only for businesses with a compelling need and the in-house technical know-how to set up and, as importantly, to manage and maintain it. Clearly, this makes it the most expensive. Even then, judicious use of the air gap is recommended, for example to separate corporate financial, personnel, marketing and research data from the Internet connected systems. We note that if connecting one corporate LAN to the Internet exposes this kind of data, questions should be asked about internal security before that happens.

Strategies

Just as people who can see the world from their window can probably be seen by the world[1], does access *to* global information imply access *from* throughout the globe? Does one's interface to the Internet pass everything through just as a glass window passes all light through? What is the equivalent to one-way mirrors in this brave new world?

The discussion is concerned with choosing which of the models of Internet connection is suitable in a given situation and some guidelines to be followed when doing so.

Requirements Analysis

As with any IT project, requirements analysis is critical. If the needs of the business with respect to the Internet are not understood, the solution implemented is likely to:

- fail to meet expectations
- be rejected by the user community
- overrun cost and time budgets
- create more problems than it solves.

Similarly, it is often better to start with reduced expectations.

The following is a list of some of the functions a business may be trying to achieve by connecting to the Internet.

- access services provided by other business, e.g. product catalogues
- provide services to other businesses, e.g. by advertising services offered on a Web page
- communicate with other business, e.g. by e-mail
- provide remote access to telecommuting staff.

Some of the critical decisions to make relate to what nature of services are required to be accessed or provided, and what level of convenience is required.

The following services may be considered for access and for provision:

- e-mail
- World Wide Web (WWW)
- File Transfer Protocol (FTP).

Note that whereas e-mail is a symmetric (peer to peer) system, WWW and FTP are examples of client/server systems and exhibit different characteristics depending on whether the client or server side is being considered.

Clients

The following possibilities for where the selected services are to be accessed include:

- from every desk
- from a dedicated computer system, physically separate from the business systems.

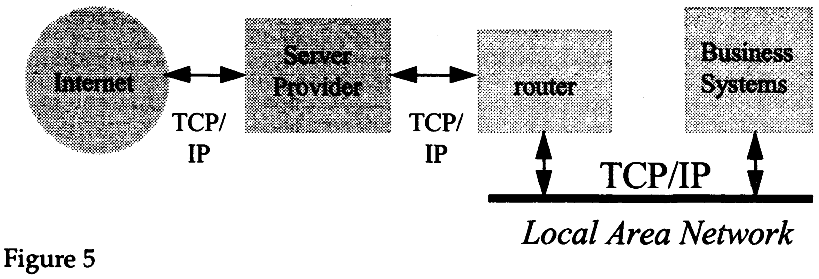


Figure 5

Is the first essential or would the second be acceptable?

Possibilities from where services may be provided include:

- the service provider's computer system
- a single, separate computer system at the business's premises
- on arbitrary computer systems at the business's premises.

The first case is simplest and cheapest, the third is most risk prone and the second may be a reasonable compromise.

Servers

Some attributes of a server to be provided by a business, directly or through a proxy, that should be considered are:

- the name under which the service is offered to the Internet, e.g. the ISP's name versus your name
- the convenience with, and cost at which the service can be updated.

Design Decisions

Given a realistic analysis of the business requirements for Internet access, the next step is to choose a model of access that meets these requirements without creating unnecessary security risks and without costing more than necessary.

The characteristics of the access models described above in *Internet Access Options* are summarised by Table 1 and compared by the following sections.

The Shell Account

The Shell Account allowed access by businesses not directly connected to the Internet to services provided by other Internet users. It is the cheapest way to access the services provided by others.

However, the Shell Account provides inconvenient and limited access to services provided by other businesses. It does not allow the use of client software that relies on a Graphical User Interface (GUI) for its operation, such as the popular WWW browser *Netscape*. On its own, it offers no means to publish one's own Web page.

This model exposes the business systems to file-oriented attacks if files are down-loaded from the Shell Account. Conversely it may permit interception attacks if used to transmit sensitive information to service providers or other correspondents.

Questions for the ISP include the following:

- What do you do to ensure that my account is not used by someone else?
- What measures to you take to prevent attacks on your systems and hence my account?
- Can I have an e-mail address for my business name,

  e.g. `my.name@my.business?`

The Proxy Server

As with the Shell Account, the Proxy Server provides access for the business without direct connection to the Internet at a relatively low cost but with some inconvenience. Whereas the Shell Account allows access by the business to services on the Internet, the Proxy Server allows the provision of services by the business to other users of the Internet.

Proxy WWW, FTP, etc. servers are fully functional. They may lack convenience for businesses who prefer to have a greater degree of timely control over the content of their servers. On the other hand, for businesses without technical expertise or time, they offer a managed service without the need to support hardware and software of one's own.

If considering this type of service, be sure to specify your naming requirements to your prospective ISP. Just as businesses prefer to advertise

	Shell Account	Proxy Server	End Node	Inter-Connection
e-mail	✓	✗	✓	✓
WWW client	✗	✗	✓	✓
WWW server	✗	✓	✓	✓
FTP client	✓	✗	✓	✓
FTP server	✗	✓	✓	✓
cost	low	low	medium	high
convenience	low	low	medium	high
risk	low	low	medium	high

Table 1: Comparison of Services, Cost and Access Models

under their own name rather than that of their advertising agency and many prefer to create the perception that they are large enough to have an in-house advertising department, a Universal Resource Locator (URL) of

http://www.my.name

is typically preferred to

http://www.isp.name/~myname

for example.

With this model of access, remote access attacks are restricted to the ISP's systems. File oriented attacks against the business are possible, for example if an FTP server allows up-loading or a WWW server has a forms or similar interface permitting the user to submit data. Interception attacks are possible if a WWW server dispenses private information, presumably with the intention of only doing so to authorised users.

Some questions to ask the ISP including the following:

- What steps will you take to protect the integrity of the material you are publishing for me?
- Can you make my WWW page (or FTP server) appear under my business's name?
- Can you support forms input and how can you deliver that input to me?

The End-Node

The End-Node brings the IP datagrams onto the business's premises. It therefore offers convenience and functionality but at a higher cost and security risk than the Shell Account or Proxy Server.

This model directly exposes the connected PC to file oriented and remote access attacks, and indirectly exposes other business systems to file oriented attacks. Interception attacks are also a risk. Provided that the air gap is maintained and files being transferred to and from business systems are vetted in the way they should be if being transferred from or to the outside, it should not introduce file oriented attacks that are not already possible, however it may increase their likelihood by encouraging more frequent transfers.

Questions to ask the ISP include the following:

- What do you do to prevent forged or otherwise undesirable datagrams from reaching my site?
- Do your other customers have greater ability to access my PC compared with the rest of the Internet?
- How would you respond if I reported suspicious datagrams apparently originating at another of your customers?
- Can you make my PC appear in a domain named after my business?
- Can you provide my PC with access to an e-mail mail-box?

The Inter-Connection

This model extends the Internet into the business systems, providing maximum flexibility and convenience at greatest cost and risk. Direct remote access attacks against business systems are possible, as well as the file oriented and interception attacks of the other models.

References to firewalls occur frequently in discussions and the literature of Internet security (Cheswick 1994). Here I simply summarise their function without attempting a tutorial. The remainder of this section may be safely skipped by readers not concerned with the Inter-Connection model of Internet access.

A firewall aims to preserve the convenience and functionality but mitigate the risks of the Inter-Connection model discussed in the sections on interconnection above.

All firewalls take the place of the router in Figure 5, taking advantage of the fact that this unit partitions the Internet into two and forces all datagrams sent to and from the business system to pass through it. This simplifies the problem of source address authentication illustrated in Figure 1 since it is now possible to apply some validation to source addresses. For example, one does not expect to receive from the Internet a datagraph containing a source address belonging to oneself.

One result of using a firewall to reduce risk is some reduction in fucntionality and convenience. Some applications and protocols are simply not tractable with a functioning firewall between client and server.

The purchase price of firewall hardware and software can range from $5,000 to $30,000 and higher for military-grade systems. The investment in training and recruitment and the recruitment cost of support and maintenance is commensurately high.

The risks inherent in the Inter-connection model remain even with a firewall: they cannot be completely removed. This only services to increase the cost of supporting such a connection as skill and vigilance is required to track and prevent the latest methods of attack.

Questions for the ISP follow

- Do your other customers have greater ability to access my network compared with the rest of the Internet?
- How would you respond if I reported suspsicious datagrams originating at another of your customers?
- Can you provide security consulting services, inlcuding the provision and maintenance of a firewall?

Conclusion

We live with flammable gas piped under our streets, high voltage power cables above our streets, and motor vehicles driviing on our streets. We study them, regulate them and attempt to contract technical barriers between them and those places where we do not want them. Of course, we fail to control them completely, but we constrain them sufficiently to accpet them and to enjoy benefits. But every so often someone gets badly hurt.

Will the undeniable securirty threats that the Internet poses both to istelf and our previously "safe" businesses prevent it from delivering its promised mass empowerment? Of course it will not, any more than computer viruses prevented the use of PCs and diskettes.

However, the sheer size of this technology and the speed with which it is being deployed must give us serious cause for concern and reason for action. The number of people able to walk down the street twisting door knobs to check for poor quality or absent locks is large as is the number that can be twisted per second. The "security through obscurity" behind many existing security measures simply does not survive in this neighbourhood.

Compared to the problem of virally-infected diskettes occasionally being injected into the periphery of our bodies corporate, the Internet offers an intravenous drip: often nutritious and even pleasurable, but unpredictably laced with toxic and even fatal substances. It is bad enough to find used hypodermic syringes on the street outside your foyer but what could be worse than finding hackers getting high on your data in your Data Warehouse? (Presumably, not finding them even though they are there.)

In many cases the risk mitigation is to be found in reducing expectations, which may have been unreasonably high anyway due to the hype surrounding the Internet, and employing a lower technology solution. For example, a Shell Account plus a Proxy Server can leave many, though not all, of the security problems with the service provider. Alternatively, an End-Node provides a compromise between convenience and security, but at the price of increased vigilance and expectations. The Inter-Connection model provides maximum functionality and convenience at increased risk, but for businesses prepared to commit to the capital and recurrent expense, firewall products and consulting expertise is available.

Glossary

AARnet	Australian Academic and Research Network, the provider of the Internet back-bone in Australia
catenet	Concatenated Network, a homogeneous network built from a set of heterogenous networks using a common protocol and a per network means of carrying that protocol over each of them, plus a gateway at each point of connection between those networks that understands the means used to carry the common protocol by both of the networks that it connects together
datagram	From data telegram, a self-contained chunk of data independently addressed and routed; a.k.a. packet
e-mail	Electronic Mail
firewall	A barrier between the Internet at large and one's business systems that aims to provide general access to the Internet from the business but restrict modes of access to the business from the Internet
FTP	File Transfer Protocol, the standard Internet protocol for the sequential transfer of files
GUI	Graphical User Interface
HTML	Hyper Text Markup Lauguage, a standard markup language suited to hyper-text
HTTP	Hyper Text Transfer Protocol, a protocol for transfer HTML and other file formats on the WWW
Hyper Text	Text in which cross-references, for example between a defined term and its definition, are encoded in the text and which may be followed automatically using suitable browsing software
Internet café	A shop front service providing casual access to the Internet
IP	Internet Protocol, a connectionless datagram oriented node-to-node protocol that must be supported by any node for it to be part of an internet
ISP	Internet Service Provider
IT	Information Technology
LAN	Local Area Network, a network that serves a localised user community and typically exhibits high bandwidth and short distances; the term is also (mis)used to refer to work-groups of PCs using a network operating systems for shared access to files, printers, etc.
NSFnet	National Science Foundation Network, the provider of the Internet back-bone in the USA
Sneaker Net	A communications mechanism in which the data to be transferred is encapsulated in a medium able to be carried by instances of the technology *homo sapiens*; named after the style of shock absorbing material normally fitted to the base of the supporting poles of these carriers.
TCP	Transmission Control Protocol, a connection-oriented host-host protocol that provides a reliable, byte-stream oriented services between clients and servers
URL	Universal Resource Locator, a syntactic construct identifying a network resource in a way that is independent of the site from where it is being referenced
WWW	World Wide Web, a network within the Internet in which the common protocol is HTTP and the client software is usually graphical and designed for browsing

References

Cerf, V., "The Catenet Model for Internetworking," IEN 48, Information Processing Techniques Office, Defense Advanced Research Projects Agency, July 1978.

Postel, J. (ed.), "Internet Protocol - DARPA Internet Program Protocol Specification," RFC 791, USC/Information Sciences Institute, September 1981.

Postel, J. (ed.), "Transmission control Protocol - DARPA Internet Program Protocol Specification," RFC 793, USC/Information Sciences Institute, September 1981.

Cheswick, W. and Bellovin, S., "Firewalls and Internet Security - Repelling the Wily Hacker," Addison-Wesley Publishing Company, 1994.

Tim Roper has been in the computing industry since 1979. Tim is currently employed by TechNIX Consulting Group International, where he manages a systems integration, software development and facilities management project, builds firewalls and consults on network design and security. Under pressure he confesses to authoring a home page and even to writing sendmail configurations.

[1] cf. "people who live in glass houses should get undressed in the dark".