Network Translation, Inc.

Private Internet Exchange (PIX)
Technical White Paper

May 16, 1995

Rev 2.0 - KWP

Network Address Translation, Internet Security, and the Private Internet Exchange

Internet Address Depletion

The combination of explosive growth in TCP/IP networking and the long-standing practice of assigning globally unique IP addresses to all hosts on TCP/IP networks has resulted in rapid depletion of the available IP address space. Since a unique address is required for each host connected to the global Internet, this presents a serious problem for new enterprise connections. The *IP: next generation* (IPng) area of the Internet Engineering Task Force (IETF) is currently considering proposals for a long-term solution. However, the process of selecting the new standard, managing the transition, and finally achieving ubiquitous implementation will take several years, if indeed it can be accomplished at all. "Given the potential proliferation of network address translation devices, it is not clear that IPng will secure sufficient following to attain market viability."[1] In the meantime, three primary strategies have emerged for maximizing the longevity of the current IP standard.

Address Allocation Guidelines and Private Internets

The traditional method of Internet address allocation split the address space into three classes of networks based on the number of hosts within them. For convenience, address classes were divided on 8-bit boundaries, allowing roughly 250 hosts on a Class C network, 64,000 on a Class B, and 16 million on a Class A. Unfortunately, this lack of granularity does not reflect the realities of enterprise networking. Many organizations have networks which fall somewhere between the Class C and B magnitudes. A network manager with 4,000 hosts, for example, faces the dilemma of using 16 Class C registrations or 1/16 of a Class B.

Because of the underutilization of address space within assigned Class B addresses, they are now nearly impossible to get. The Internet Registry will assign blocks of multiple Class C addresses to applicants who do not meet the IR requirements for a Class B allocation, which include a minimum of 4,096 hosts and the submission of a detailed network plan.

> The restrictions in allocation of Class B network numbers may cause some organizations to expend additional resources to utilize multiple Class C numbers. This is unfortunate, but inevitable if we implement strategies to control the assignment of Class B addresses. The intent of these guidelines is to balance these costs for the greater good of the Internet.[2]

Organizations have historically been assigned globally unique IP network addresses regardless of their intent to connect their private network to the Internet. Even those which do join the Internet usually only allow Internet access to a small percentage of the hosts on their leaf domain. The ratio of hosts with direct Internet access to hosts without such connectivity is typically between 1:1,000 and 1:10,000 in large corporate networks.[3] The use of registered, globally unique IP addresses for such large numbers of hosts which don't need them has further exacerbated the address depletion problem.

Recently, the Internet Assigned Numbers Authority reserved three blocks of the address space for use by private networks: one Class A, 16 Class B, and 255 Class C network numbers. These addresses may be used on the enterprise LAN for hosts that will never have direct IP connectivity with external hosts. But in real world enterprise networks, hosts fall naturally into three categories, not the "*always/never* connected to the Internet" dichotomy implied by RFC 1597.[4]

Categories of Connectivity	
Level	*Examples*
Always	Email, FTP, World-Wide Web servers
Sometimes	User's workstation or PC
Never	Secure hosts, corporate database servers

The private network addressing scenario laid out in RFC 1597 relegates the connectivity needs of this middle category of hosts to "application layer relays", otherwise known as proxy servers. Demand from end users for the direct Internet connectivity they need to run World-Wide Web browsers on their desktops is growing daily, but the deployment of proxy servers to meet this need will add yet another layer of complexity (along with potential maintenance and administrative headaches) to the network.

Classless Inter-Domain Routing (CIDR)

CIDR, as described in a series of Internet RFCs† , is primarily aimed at increasing the efficiency (and reducing the size) of the Internet routing tables. This is being accomplished by a policy of allocating IP addresses in a way which allows routing information for multiple networks to be aggregated into a single routing table entry. Internet service providers are now being assigned contiguous blocks of the Class C address space, which are in turn reallocated to their new customers. The incorporation of variable-length netmask information into the routing protocols makes it possible for these multiple Class C networks to be served by a single routing table entry on the Internet. The designation of this mechanism as classless comes from the fact that it enables routing at intermediate levels between the traditional 8-bit boundaries of IP network classes.

One unfortunate side-effect of Classless Inter-Domain Routing is that, in order to maximize its effectiveness, existing domains may need to be renumbered. This will incur a high administrative cost for the networks involved.

Network Address Translation (NAT)

The third (and most easily deployable) strategy for alleviating IP address depletion is Network Address Translation.[5] NAT is based on the concept of address reuse by private networks, and operates by mapping the reusable IP addresses of the leaf domain to the globally unique ones required for communication with hosts on other networks. It would be difficult indeed to take full advantage of reusable addresses on a private network without employing NAT functionality.

It is also unlikely that many network managers will voluntarily incur the expense of renumbering their networks, as will eventually be necessary for full deployment of CIDR. The insertion of a Network Address Translator at the Internet connection point makes this a one-step operation, eliminating the need to visit each host on the corporate LAN to change its IP address.

Network Address Translation dovetails with both reusable addressing and CIDR, simplifying or eliminating many of the obstacles associated with the deployment of these initiatives. But NAT also provides simple solutions for a number of other network management problems.

Private Internet Exchange (PIX)

Network Translation, Inc., introduced the Private Internet Exchange, the first commercially available implementation of Network Address Translation, in late 1994. Since that time, a number of Internet Firewall vendors have introduced so-called "address translation" features in their UNIX-based software products. To date, none of these do genuine RFC 1631 translation, relying instead on a scheme which maps connections to high-numbered ports on the firewall's single visible IP host address.

† RFC 1467, RFC 1481, RFC 1517, RFC 1518, RFC 1519, and RFC 1520

The Private Internet Exchange, which incorporates firewall and proxy server functionality along with its address translation mission, comes in a standard 19-inch rack-mountable package and is equipped with two Ethernet ports. In a typical installation, the inside (local) port is connected to the private network and the outside (global) port connects the PIX to the DMZ segment where the Internet router resides. Configuration is accomplished with the familiar **ifconfig** and **route** commands for each network interface. A **global** command specifies the virtual network number to which private host IP addresses will be mapped. The PIX broadcasts a default route to the inside network and provides proxy ARP within the DMZ segment for hosts on the inside network.

Dynamic Address Allocation

Mapping between local and global addresses is done dynamically. An Internet-bound packet sent by a host on the inside network follows default routes to the inside interface of the Private Internet Exchange. Upon receipt of the outbound packet, the source address is extracted and compared to an internal table of existing translations. If the inside host's address does not appear in the translation table, a new entry is created for that host, assigning a globally unique IP number from the pool of available addresses. The actual translation is accomplished by changing the source address of the packet to this "legal" address. Since the differences between the original and translated versions of the packet are known, the checksums are efficiently updated with a simple adjustment rather than complete recalculation. After a user-configurable timeout period during which there have been no translated packets for a particular address mapping, the entry is removed and the global address is freed for use by another inside host.

As mentioned above, the number of hosts needing Internet connectivity from inside the corporate firewall is generally a very small percentage of the domain. But even fewer of them will require access simultaneously. Dynamic address allocation as implemented in the PIX efficiently leverages a relatively small registered address space to serve the Internet connectivity needs of a much larger user population. In the course of expanding the private network, Internet access is available to the new hosts without reconfiguration.

Adaptive Application Security

Dynamic address translation is only enabled for connections initiated from the inside network, and is port-specific. The translation for an outbound HTTP connection from a World-Wide Web client, for example, would only forward packets from the the external Web server which were destined for port 80 of the client machine. In the case of FTP, which uses an ephemeral port for its data connection, the PIX takes note of the port number passively opened by the client's request and will only allow inbound FTP data for sessions which were initiated from inside the private network.

This level of selectivity is enabled by retaining state information for each TCP connection established through the Private Internet Exchange. A table containing the destination address, port numbers, sequencing information, byte counts, and internal flags for each TCP connection associated with a particular host address translation is maintained for the life of the translation entry. Inbound packets are compared against entries in the connection table and are permitted through the PIX only if an appropriate connection exists to validate their passage.

Thus, the Private Internet Exchange provides the functionality of a proxy server without the extra administrative overhead and without the need for special client software. Typical proxy servers run at the user level on a multi-user operating system and operate by copying data between separate TCP connections. The PIX operates on the packets directly, resulting in much higher performance.

Static Translations and Conduits

Unlike clients, internal hosts acting as Internet servers (email, anonymous FTP, World-Wide Web, etc.) require a predictable registered address and a different access policy. The Private Internet Exchange offers static translations which hard-wire an internal address to a specific global address, and do not time out. Static translations default to a "wide open" state, allowing any host on the Internet to connect to any port on the inside server, but security criteria for each conduit may be enforced via a mechanism similar to that of a packet-filtering router:

Protocol	TCP or UDP
IP address	Remote hosts/networks permitted access
Netmask	Applied to the above IP address
Port	IP port number for which access is allowed

Once conduits are created for a static translation, all connections not specifically allowed are denied. These disallowed packets are silently dropped (from the intruder's perspective) and logged for postmortem by the systems administrator.

Sequence Number Randomization

The technique of IP address spoofing has been well-known since it was first described by Robert T. Morris in 1985.[6] Recently a rash of such attacks on the Internet precipitated a Security Advisory from CERT (Computer Emergency Response Team).[7] Essentially, spoofing IP addresses requires the ability to guess the sequence numbers of TCP packets. Most TCP/IP implementations use a simple additive algorithm for incrementing sequence numbers, making it a trivial matter for an intruder to guess the next number in a connection (from even a single intercepted packet) and subsequently hijack that session. The Private Internet Exchange makes the process of guessing TCP sequence numbers extremely difficult, if not impossible, by using a randomizing algorithm for their generation.

Logging

Using the standard Berkeley syslog mechanism, The Private Internet Exchange is capable of logging extensive security and administrative information to a designated host:

Private Internet Exchange Syslog Messages	
Category	*Logged Events*
System	Console logins and logouts, PIX reboots
Resource	Exhaustion of connection and/or translation slots
Accounting	Bytes transferred, IP addresses and ports
Security	TCP connections rejected and UDP packets dropped

These syslog messages enable detailed monitoring of attempted security violations and network resource usage.

Summary

Going far beyond the address translation device originally described in RFC 1631, the Private Internet Exchange offers a number of unique advantages:

- Greater security and better performance than proxy servers, without requiring special "proxy-enabled" client software
- Transparent Internet access for end users
- Concealment of Internal network architecture from the outside world, except for those hosts explicitly allowed

- Fine-grained control of internal server access
- Firewall functionality without the administrative overhead and security risks associated with UNIX-based systems
- Rapid Internet connection of existing unregistered networks, without changing individual host IP addresses
- Unparalleled flexibility in network design: any combination of address class, subnetting, address reuse, etc.

The PIX is a versatile and powerful new tool for the network administrator's arsenal.

Contact Information

Network Translation, Inc.
1901 Embarcadero Road, Suite 108
Palo Alto, CA 94303
Phone: (415) 494-NETS (6387)
Fax: (415) 424-9110
Email: info@translation.com

References

1. J. Curran, "IPng White Paper on Market Viability," RFC 1669, BBN, August 1994.
2. E. Gerich, "Guidelines for Management of IP Address Space," RFC 1466, Merit, May 1993.
3. E. Fleischman, "A Large Corporate User's View of IPng," RFC 1687, Boeing Computer Services, August 1994.
4. Y. Reckhter, B. Moskowitz, D. Karrenberg, and G. de Groot, "Address Allocation for Private Internets," RFC 1597, T.J. Watson Research Center, IBM. Corp., Chrysler Corp., RIPE NCC, March 1994.
5. K. Egevang and P. Francis, "The IP Network Address Translator (NAT)," RFC 1631, Cray Communications, NTT, May 1994.
6. R. Morris, "A Weakness in the 4.2BSD Unix (tm) TCP/IP Software," Technical Report. AT&T Bell Laboratories, February 1985.
7. Computer Emergency Response Team, "IP Spoofing Attacks and Hijacked Terminal Connections," CA-95:01, Carnegie-Mellon University, January 1995.