

NAVAL POSTGRADUATE SCHOOL Monterey, California



THESIS

**TOWARD AN INTERNET SERVICE PROVIDER (ISP)
CENTRIC SECURITY APPROACH**

by

Patrick D. Price

March 2002

Thesis Advisor:

Timothy Levin

Thesis Co-Advisor:

Cynthia Irvine

**This thesis was completed in cooperation with the Institute for Information
Superiority and Innovation.**

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.			
1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE March 2002	3. REPORT TYPE AND DATES COVERED Master's Thesis	
4. TITLE AND SUBTITLE: Title (Mix case letters) Toward an Internet Service Provider (ISP) Centric Security Approach			5. FUNDING NUMBERS
6. AUTHOR(S)			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			10. SPONSORING/MONITORING AGENCY REPORT NUMBER
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited			12b. DISTRIBUTION CODE
13. ABSTRACT (maximum 200 words) Individual users, businesses, and governments have become functionally dependent on the Internet's connectivity to interact at the most basic levels of social and economic intercourse. Yet self-propagating worms and distributed denial of service attacks have demonstrated that disruption of the Internet infrastructure can be quickly achieved despite the vast knowledge of vulnerabilities and readily available subscriber-based countermeasures. In part, this condition is made possible because networks continue to operate under an obsolete subscriber-centric security paradigm that is based on all end users being trusted to act appropriately. This thesis develops the idea of an Internet Service Provider (ISP)- centric security approach by examining the types, roles, security mechanisms, and operational precepts of ISP's to illustrate their functional control within the infrastructure. Denial of service and worm attacks are detailed to provide the context for an emerging set of conditions that forms the basis of the requirement for the ISP approach. This paper concludes by examining four enabling technologies currently available that, used uniformly, provide ISPs with the framework to implement Internet based security that can serve to enhance the layered defense model and invoke the tenants of best practices.			
14. SUBJECT TERMS Internet Security, Internet Service Provider, Denial of Service Mitigation			15. NUMBER OF PAGES 91
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited
This thesis was completed in cooperation with the Institute for Information
Superiority and Innovation.

TOWARD AN INTERNET SERVICE PROVIDER (ISP) CENTRIC SECURITY
APPROACH

Patrick D. Price
Commander, United States Navy
B.A., The Citadel, 1986

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN INFORMATION SYSTEMS AND OPERATIONS

from the

NAVAL POSTGRADUATE SCHOOL
March 2002

Author:

Patrick D. Price

Approved by:

Timothy Levin, Thesis Advisor

Cynthia Irvine, Co-Advisor

Dan Boger, Chairman
Department of Information Sciences

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Individual users, businesses, and governments have become functionally dependent on the Internet's connectivity to interact at the most basic levels of social and economic intercourse. Yet self-propagating worms and distributed denial of service attacks have demonstrated that disruption of the Internet infrastructure can be quickly achieved despite the vast knowledge of vulnerabilities and readily available subscriber-based countermeasures. In part, this condition is made possible because networks continue to operate under an obsolete subscriber-centric security paradigm that is based on all end users being trusted to act appropriately. This thesis develops the idea of an Internet Service Provider (ISP)-centric security approach by examining the types, roles, security mechanisms, and operational precepts of ISP's to illustrate their functional control within the infrastructure. Denial of service and worm attacks are detailed to provide the context for an emerging set of conditions that forms the basis of the requirement for the ISP approach. This paper concludes by examining four enabling technologies currently available that, used uniformly, provide ISPs with the framework to implement Internet based security that can serve to enhance the layered defense model and invoke the tenants of best practices.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

I.	INTRODUCTION.....	1
A.	BACKGROUND.....	1
B.	OBJECTIVES.....	3
C.	ORGANIZATION.....	3
II.	EVOLUTION ON THE INFRASTRUCTURE.....	5
A.	INTRODUCTION.....	5
B.	DESCRIPTION OF CURRENT INTERNET ENVIRONMENT	6
1.	Types of Internet Service Providers	7
2.	Achieving Connectivity Between Networks	11
3.	Fundamental Elements Within An ISP Network.....	14
4.	Internet Routing and Associated Protocols	15
C.	STATE OF INTERNET SECURITY.....	16
1.	First Generation Techniques.....	17
2.	Second Generation Techniques.....	21
3.	Organizational Approach To Implement Distributed Security....	26
D.	INTERNET REQUIREMENTS TRANSFORMED.....	29
1.	Erosion of Trust.....	30
2.	Unsophisticated User Base.....	30
3.	Sophisticated and Ubiquitous Attack Techniques	31
4.	Evolving Influence of Legal Liability	32
E.	SUMMARY.....	38
III.	DEFINING THE THREAT.....	41
A.	INTRODUCTION.....	41
B.	DENIAL OF SERVICE / DISTRIBUTED DENIAL OF SERVICE	43
C.	DENIAL OF SERVICE IMPLEMENTATION.....	44
1.	SYN-ACK/TCP-SYN Flooding	44
2.	SMURF.....	45
3.	FRAGGLE / UDP.....	45
4.	Teardrop	46
5.	PING of Death/ Oversized Packet.....	46
D.	DISTRIBUTED DENIAL OF SERVICE IMPLEMENTATION	46
E.	WORMS AND ASSOCIATED HYBRIDS	48
F.	WORM IMPLEMENTATION.....	50
1.	Self-launch Method	50
2.	User-launch Method.....	50
3.	Hybrid Method.....	50
G.	SUMMARY.....	53
IV.	ISP CENTRIC APPROACH.....	55
A.	INTRODUCTION.....	55
B.	THE SECURITY POLICY ENFORCEMENT POINT	56

1.	Firewall	59
2.	Anti-spoofing (Ingress/Egress Filtering)	59
3.	Network Address Translation	60
4.	Quality of Service (QoS)	60
5.	Web Steering	62
C.	MITIGATING E-MAIL AS A PROPAGATION MECHANISM	62
1.	Early Warning	66
2.	Economy of Scale	67
D.	DENIAL OF SERVICE APPLIANCE	67
E.	DISTRIBUTED FIREWALL	68
F.	SUMMARY	69
IV.	CONCLUSIONS AND RECOMMENDATIONS	71
A.	CONCLUSIONS	71
B.	FUTURE RESEARCH	74
	LIST OF REFERENCES	75
	INITIAL DISTRIBUTION LIST	77

LIST OF FIGURES

Figure 1.	A Representative Tier-1 Internet Service Provider Backbone	9
Figure 2.	Internet Service Provider Tiered Architecture	11
Figure 3.	ISP Internal Network Architecture.....	15
Figure 4.	Port and Aggregate Rate-Limiting	25
Figure 5.	Flow Rate-Limiting	26
Figure 6.	Basic Topology and Communication Path of DDoS [From Ref 12]	48
Figure 7.	Network Attack Model.....	52
Figure 8.	Nortel Networks Shasta 5000 BSN Deployment Architecture	58
Figure 9.	Firewall Deployment within a Demilitarized Zone (DMZ)	64
Figure 10.	Representative Intelligent Scanning Architecture.....	65

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

I would like to express my gratitude to my advisors, Tim Levin and Cynthia Irvine, for their supervision and support during this project. I would also like to thank the efforts of Linda Zupan from the Nortel Corporation who provided me access to an abundance of product documentation as well as the human resources needed to understand it. Without all their help, this thesis would not have been possible.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. BACKGROUND

By any account, the technological revolution that now characterizes the Information Age has firmly engulfed the world and created a global society ever dependent on computers and computer systems. Fueled by the concepts of data sharing and distributed collaboration, the diverse community of isolated computers and networks has evolved into a meshed architecture of interconnected networks known more familiarly as the Internet. The progeny of ARPANET project designers, today's Internet has become a complex, nodal-structured, multi-tiered network of interconnected systems that serves the core functionality of moving packetized data from source to destination quickly and flexibly.

As dependence on the Internet has increased, the requirement to protect the systems, pathways, and data contained within this architecture has emerged as a major priority of users, commercial industry, and governments. Articulated at the highest levels of U.S. government, Presidential Decision Directive 63 (PPD 63)-October 1997, Protecting America's Critical Infrastructures, states that, "certain national infrastructures are so vital that their incapacity or destruction would have debilitating impact on the defense or economic security of the United States". Moreover, the significant inference of PPD 63 is that a national effort that transcends individual or community self-interest is needed to assure security of the increasingly vulnerable information content and infrastructure.

Critical to understanding the security of the Internet is the recognition of the conditions and requirements that govern the use of the system, and the identification of the appropriate stakeholders responsible for translating security concepts into action across the distributed infrastructure. Today's Internet stakeholders, composed of service providers, private network operators, academic and research institutions, individual users, and product vendors, generally cooperate informally to keep the overall system running and "protected". In this traditional context, participating users have been singly responsible for their operations and defensive preparations within their own domain as

dictated by their own unique, often conflicting, set of requirements. Conversely, the higher level Internet infrastructure and diverse market of service providers, collectively known as Internet Service Provider's, have emerged as the predominate influence over the Internet that operates under a commercial business model geared to provide ubiquitous access and maximum performance at the lowest cost possible. Accordingly, the Internet security environment has evolved under the same influences and can be aptly characterized as an end system or subscriber-centric approach to security and infrastructure protection.

From its beginnings, use of the Internet had always been a voluntary endeavor, governed by some commonly held principles and rules of etiquette that served to define the boundaries of acceptable behavior. However, participation in the networked system today is no longer considered voluntary for commercial enterprises to compete, for individuals to interact within society, or for governments to govern, defend, and provide for its citizens. Coincident with exponential growth and increased complexity, the system has given rise to new levels of vulnerability, in terms of degrading availability and reliability, made easily more exploitable by individuals, nation states, or sub-national threats bent on malicious intent. Despite the vast knowledge base of system vulnerabilities and countermeasures readily available to individual subscriber and network administrators, malicious hackers consistently demonstrate their ability to destroy and disrupt the telecommunications infrastructure and infect thousands of hosts in minimal time. Evidence of this pervasive trend has been most clear when one considers the disruption of Web based services for millions of end users that resulted from a rash of coordinated denial of service attacks that targeted some of the largest e-commerce sites during February 2000 or the litany of computer worm attacks such as Melissa, ILOVEYOU, Code Red, and Nimda that successively infected larger portions of the Internet domain space between 1998 and 2001.

Given the growing inefficiency of the subscriber-centric security model that characterizes today's environment to shield itself against attack, coupled with the emergence of the Internet Service provider as the critical pathway and keepers of the infrastructure, it is hypothesized that the most damaging attacks, specifically distributed denial of service and Worm propagation, can be better mitigated with an Internet Service

Provider's (ISP) -centric security approach to enhance the existing layered defense methodology.

B. OBJECTIVES

Through a process of examining available security policy, mechanisms and architectures, the primary objective of this thesis is to formulate a basic understanding of the role that Internet Service Providers, to include military and commercial organizations, play with respect to Internet security. As a secondary objective, this thesis will suggest an alternative organizational approach necessary to enhance and extend the concepts of layered defense as it relates to Internet security. By matching the emerging set of conditions or requirements that govern use of today's Internet within the context of currently available technology, this thesis will suggest potential modifications to the enforcement mechanisms and architectural implementations employed by ISPs that will further mitigate the effects of distributed denial of service (DDOS) and code propagation (Worm) attacks.

C. ORGANIZATION

This thesis is divided into five chapters. Chapter II examines the current types, roles, and operational precepts of Internet Service Providers to illustrate the aspects of their functional and positional control within the infrastructure. From an ISP perspective, this chapter will detail the dominant mechanisms and one organizational context that constitutes the state of Internet security intended to foster availability and reliability. It will conclude by identifying an emerging set of conditions affecting Internet security and ISP operations to form the basis of the requirement for an ISP-centric security approach. Considering some of the more critical threats facing future Internet security, Chapter III provides a detailed explanation on how the Internet infrastructure is targeted by both internal and external sources using the specific attack techniques associated with DDoS and Worm propagation. Within this context, Chapter IV identifies four enabling technologies that are available, but not widely deployed, that can serve as the ISP-centric framework to better mitigate these specific threats. The conclusion and recommendations are presented in Chapter V.

THIS PAGE INTENTIONALLY LEFT BLANK

II. EVOLUTION ON THE INFRASTRUCTURE

A. INTRODUCTION

Over a period of years, computers have evolved from the simple single-user computing environments that emphasized a stand-alone functionality to that which can now be characterized as largely distributed. Realizing tremendous benefits, advantages, and efficiencies gained from resource and information sharing, individual users, businesses, and governments have become functionally dependent on a network of ubiquitous connectivity and near absolute availability to operate at some of the most basic levels of social, political, and economic intercourse. Along with evolutionary changes in the computer processing capabilities and functionality, there has been an associated evolution in the computing environment that has shifted from a subscriber-centric emphasis to a network-centric approach that puts a premium on effective utilization of the information that resides on or passes between systems. Best articulated in the expression “*the right information, at the right time, to the right person*”, information is now described as the most valuable asset of organizations today.

Today, valuable information is largely exchanged across an infrastructure that fosters redundancy to eliminate single points of failure and to ensure the utmost reliability in transmission. In terms of access to the infrastructure, Internet Service Providers have emerged as the dominant providers of critical gateways or paths to the infrastructure. While both the computer and computing environments have evolved with advances in technology and organizational concepts, the Internet security environment has largely stagnated within the subscriber-centric security framework to provide a point solution approach that has arguably failed to keep pace with the evolving network connectivity and associated threats represented by distributed denial of service and malicious code attacks against key components on the infrastructure.

To gain an understanding of the role that ISPs play with respect to Internet security and how the Internet infrastructure contributes its own vulnerability in terms of DDoS and Worm propagation attacks, it is necessary to examine the logical and physical connectivity of the Internet itself. The objective of this section is to provide a general

overview of the network infrastructure to underscore the scope of the ISP influence and points of control while identifying an emerging set of conditions that govern the operation of the Internet. As an extension to the emerging conditions, this chapter will also examine the larger context of liability that has the potential to influence an ISP-centric approach above and beyond the technical issues of combating the denial of service and worm threats.

B. DESCRIPTION OF CURRENT INTERNET ENVIRONMENT

In the logical sense, the Internet has been commonly referred to as a network of networks. During each stage of its development, Internet design has been shaped by a succession of influential stakeholders who sought to impart its particular values in terms of functionality and security [Ref 1]. In the beginning, the academic and research community garnered the attention of users as they concentrated on the technology of internet service that fostered intellectual collaboration and scientific achievement. Then, the telecommunications industry regained position in an attempt to shape the data service itself, as evidenced in the formulation of the feature-oriented services such as the X.25 protocol or frame relay. Today, Internet access and availability is provided by a vast, sometimes convoluted, array of providers and telephone companies. While the organizations that make up the telecommunications industry and academia have been established for many years, Internet Service Providers (ISP) have only recently emerged as the dominant influence to provide and manage the extensive physical and logical implementations of the Internet.

On the immediate surface, the community of ISPs provide the switches, routers, and access services that link participating subscribers, both individual consumer and business enterprises, to the Internet. Somewhat hidden, but nonetheless just as vital, are the hosts of carriers (Interexchange Carriers (IEC) and Local Exchange Carriers (LEC)) that operate the public telephone network cabling and equipment of the telecommunication infrastructure that ISPs leverage against to provide national or even global interconnectivity. IECs are the largest of the telecommunication carriers that traditionally have been responsible for long-haul transport services of data and voice communication. Since few companies provide just long distance telephone services

anymore, IECs like AT&T and WorldCom have diversified into other areas to include becoming ISP's themselves¹. LECs are the telecom companies that own and operate most of the actual access lines into customer homes and businesses to provide "local" level transport services. Many LECs have expanded capabilities beyond their historical local context to achieve regional and national presence. As in the case of IECs, LECs have also introduced Internet services as a major component of their function. For example, Verizon Communications began operations as a leading local phone company and has since expanded to become a national multi-service provider that offers long distance voice, data, and wireless Internet/telephone services.

Together, these organizations constitute the underlying network architecture in today's environment since the Internet's expansion outside of the university and research settings. However, it must be noted that clear distinctions between division of labor, titles, and responsibility are constantly changing in response to the rapid growth of an industry dominated by market-based companies seeking to balance the myriad of economic forces and cost-revenue calculations.

1. Types of Internet Service Providers

According to a 2001 Internet survey conducted by Nua², it was estimated that approximately 6000 ISP's operate in the U.S alone. When combined with all other countries, they serve to provide Internet access and assorted services to more than 500 million hosts around the world. Self-organized into a three-level hierarchical structure as Tier 1, Tier 2, and Tier 3 providers, ISP's have been generically categorized within the tiered structure according to their network infrastructure and in terms of their supported customer base. An ISP's network classification has come to generally refer to its ownership of particular elements of the Internet infrastructure (routers, cable/fiber backbone, etc.) or leasing capacity of its network resources (bandwidth). Customer classification refers to the scale of the customer base under contract, in terms of national or regional coverage. Alternatively, ISPs that do not own or lease their own backbone

¹ The Telecommunications Act of 1996 lifted barriers to allow telephone companies to compete in both local and long distance markets as well as to provide services outside their traditional context to include Internet access and multimedia applications.

² Nua Internet Surveys located at <http://www.nua.ie>

resources but still contribute to the Internet connectivity model are commonly referred to as resellers. The three basic types of the ISP market include:

a. National Service Providers (Tier 1 NSP)

Tier 1 organizations are commonly known as national backbone providers. They are the ISPs that fund, install, operate, and lease capacity along the very high speed fiber optic cabling that spans the entire United States as illustrated in Figure 1. In this regard, the Tier 1 ISP is usually a subsidiary of a large IEC that is providing long distance connection between local or regional networks as part of its telecom business. Possessing a national presence level, Tier 1 providers are generally connected to all major interexchange points (IXP), administered by the exchange carriers, that constitute the focal meeting points of the Internet infrastructure and serve as the basis for contractual peering agreements between ISPs. Large business enterprises, Tier 2 ISPs, and various reseller providers usually characterize the Tier 1 customer base. Common Tier 1 providers of today include companies like MCI WorldCom, AT&T, UUNet, and Cable and Wireless.

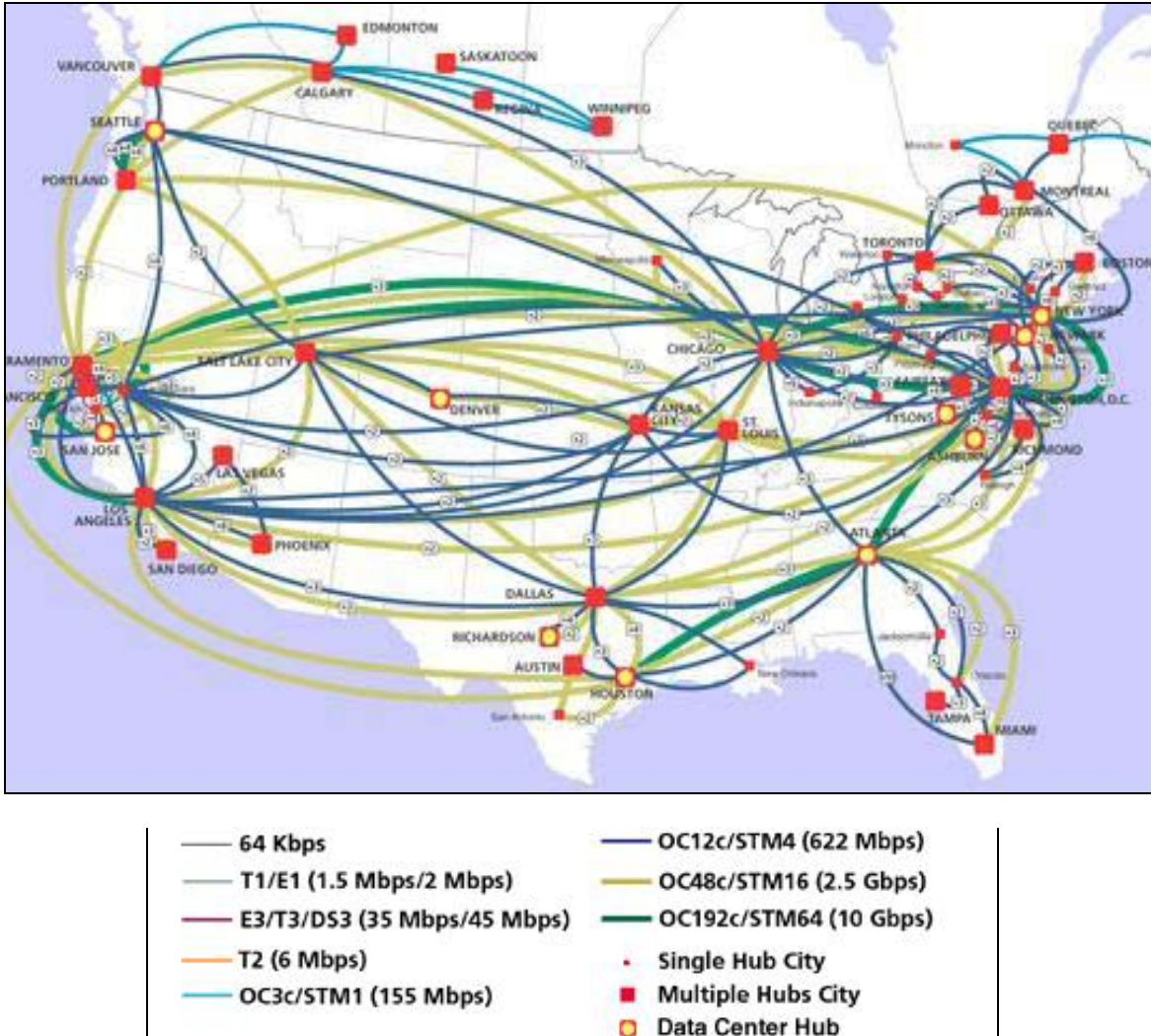


Figure 1. A Representative Tier-1 Internet Service Provider Backbone³

b. Regional Service Providers (Tier 2 RSP)

Tier 2 providers are generally considered the regional and even national level providers that connect to the Internet backbone via Tier 1 providers. Functionally, they are most similar to Tier 1 providers in that they also lease their backbone network to subscribers, but are much smaller in scale and are not considered full equals to the Tier 1

³ Source: WorldCom at <http://www1.worldcom.com>

community. Tier 2 ISPs will generally establish peering agreements with the national service providers to transfer data traffic. In this regard, they are typically configured to connect at a carrier IXP or couple directly with a national provider backbone. Tier 2 networks generally encompass a single geographic region of the country while servicing a more confined customer base. Accordingly, Tier 2 providers are uniquely positioned to offer more direct involvement with end subscribers in terms of security, customer training, help desk related functions, and configuration services.

c. Resellers (Tier 3)

The third level in the ISP hierarchy includes the Tier 3 organizations that are typically considered to be the “local” service providers. Tier 3 providers usually purchase services and capacity from larger national or regional providers and, in turn, resell this service to small businesses or individual residential consumers. Because of their relative smaller size, reseller ISP’s typically are single-sited businesses that operate fairly limited access modem banks and connect to the higher level tier structure using the lower end connection capability relative to the entire infrastructure. Tier 3 providers can be differentiated on the basis that they do not own significant portions of the network infrastructure themselves or lease its capacity. As such, they are also the most sensitive to operating costs that could greatly influence attempts to incorporate security infrastructure add-ons. Technically, lower Tier providers (Tier 4 and 5) do exist in today’s internet environment but given that their operations are considered extremely limited, often retaining six or fewer subscribers, they are considered to be the smallest subset of the same reseller market and not uniquely separated for this study. Figure 2 illustrates the levels of interconnectivity between the various types of service providers that represents the tiered architecture.

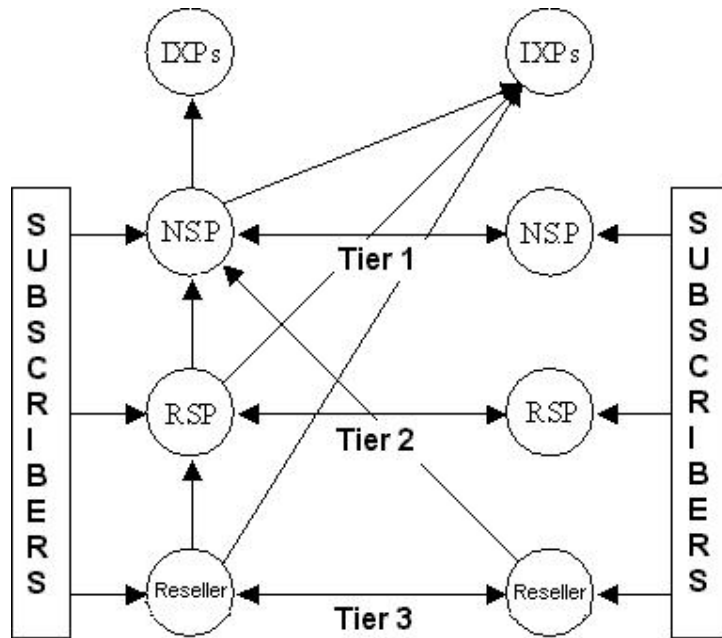


Figure 2. Internet Service Provider Tiered Architecture⁴

2. Achieving Connectivity Between Networks

In order to provide connectivity of national scale, ISP's interconnect with one another to exchange data traffic destined for their respective end-user/client base. Specifically, regional and local networks aggregate their dial-up, broadband, and wireless data traffic and systematically hand it off to upstream or backbone networks to which they are connected. Data packets that originate with a customer on one backbone segment destined for a customer on another backbone are transferred via inter-exchange points (IXP) located nationwide, of which there are approximately fifty in the U.S.⁵

IXP's are facility-based infrastructure components of the Internet architecture that provide ISPs with a common connection point to exchange traffic at both layer-2 (ATM, Ethernet) and layer-3 mechanisms (IP based routers). IXP's, also referred to as Network Access Points (NAP), Metropolitan Access Exchanges (MAE), and Commercial Information Exchanges (CIX) or otherwise, are privately owned and operated by the various Interexchange Carriers (IEC), Incumbent Local Exchange Carriers (ILEC),

⁴ Source: <http://www.ncs.gov>

⁵ North America (US only) figures compiled from listings at <http://www.ep.net> - visited January 2002

Competitive Local Exchange Carriers (CLEC) or the larger Tier 1 ISP's themselves. Various organizations include companies such as Bell Atlantic, Pac Bell, and MCI. Together, these IXP facilities consist of the very high-speed local area network architectures (LAN) or metropolitan area networks (MAN) that interconnect the various ISP's wide area network (WAN) technologies. An individual ISP is then connected to the IXP LAN by either high-speed routers or switches. All traffic routing and address information is, generally, provided by a unique IXP Route Server to the respective ISP routers based on one of two predefined policy agreements, either peering or transit.

Routers located at the IXP can be configured for access in generally two ways. An ISP can provide and manage their own routers at the IXP using dedicated lines. Alternatively, they can lease routers from the exchange carrier along with any necessary connections. However, the responsible exchange carrier will always retain management of the IXP local network while offering maintenance and repair services for collocated ISP equipment.

The interconnection agreements, or policies, that govern bi-directional data exchange at the IXP between parties are based on commercial negotiations and contracts to provide either peering services or transit services. These policies will determine how traffic is carried, transferred, monitored, or even manipulated. Depending on the nature of the agreement, these policies can serve as the basis to dictate priority or out-of-band traffic profiles to resolve traffic congestion issues. Exchange policies that govern peer or transit services are typically of a bilateral, multilateral, or multi-party bilateral nature. Bilateral agreements are those that specifically address the relationship of two ISPs at any one exchange point. Multilateral agreements are those that govern the behavior of multiple ISPs at one exchange point. Lastly, multi-party bilateral agreements are specialized hybrids that govern the behavior between small ISPs and larger Tier providers to carry traffic to other smaller ISPs, predominantly, not connected to the same local exchange point. In terms of Internet access, the more IXPs that an ISP is able to connect with only increases that network's chance of maintaining reliable service in the face of heavy traffic load, attack, or periodic outage that may result from equipment failure or maintenance action.

Unlike other national infrastructure initiatives like highway or rail, these Internet connections are not negotiated within a context of any specific industry regulation but instead are entered into on the basis of simple cost-benefit analysis at each connection. The ISP business model is driven by three factors: to generate revenue; to maximize performance; and to lower costs while treating embedded security as an ancillary function. Peering is considered worthwhile when the traffic exchanged between ISP networks is roughly equal or when there is a mutual technological or commercial benefit such as in the case of achieving a higher quality of service or faster delivery of data. Conversely, peering would not be considered beneficial if there existed a large traffic imbalance between providers, it was technically cumbersome to achieve interconnectivity, or if one provider was intent on preserving a level of service differentiation. While multiple ISPs may be connected to any one IXP, this does not guarantee or imply that an ISP can exchange traffic or have knowledge of other ISPs attached to the same exchange point. In fact, upstream providers may make it a business practice to contract peering arrangements within the umbrella of legal non-disclosure.

In the purest of peering arrangements, ISPs agree to exchange traffic with one another at essentially no cost. More specifically, however, this type of agreement is defined as the advertising of routes or Internet pathways, via specific routing protocols, for the associated parties and their respective customers. In peering relationships, ISPs obligate themselves to broadcast all their subscriber's routes to other parties and vice versa. In transit arrangements, one ISP pays for access to the upper Tier's Internet routing table, the connection point, and allocated bandwidth- in essence, becoming a subscriber itself. In receipt, the transit provider supplies a connection to all its associated end users [Ref 3].

It has been noted that, in the aggregate, the ISP industry is moving toward more peering agreements to lower costs and improve overall network performance. This shift has only recently been possible given the expanding IXP market and maturation of high performance technologies such as gigabit Ethernet and high-end optical technologies [Ref 2]. With increased peering comes closer cooperation between parties, closer cooperation can then foster higher levels of coordination, especially in terms of local or infrastructure level incident response. Coupled with the fact that the top ten competitors

generate just over 65% of all access revenues as of 2001⁶, the implications are that market consolidation trends will only bolster the potential of an ISP-centric security approach to take hold.

3. Fundamental Elements Within An ISP Network

The final aspect in understanding the critical position that ISPs retain within the Internet is understanding the architectural elements behind an ISP's internal network that serve the purpose of achieving access to the infrastructure as well as authenticating and distribution of data for the subscriber. ISP's network infrastructure can be separated into three distinct areas. Figure 3 provided as an overview of an ISP's internal network.

a. Access Network

The access network is that portion of the ISP's infrastructure that comprises the various access services and equipment used to connect subscribers to the Internet and ISP services from the point data leaves the local loop telephone system. Functionally, the access network sits between the subscriber edge networks and provides the ISP a point of aggregation for incoming traffic. For example, the access network will contain the remote access servers (RAS) that terminate the dial-up, DSL, or cable modem connections.

b. Distribution Network

The distribution network is that portion of the ISP's network that connects the access network to its backbone services. It will contain the Remote Access Dial In User Service (RADIUS) servers that contain username and password information to authenticate subscribers and end users. Given proper authentication by RADIUS, the RAS will be authorized to issue an appropriate IP address and finish the connection that allows the user continued routing along the internet backbone directly or via an upstream provider. Other services within the distribution network include the Domain Name Server (DNS) that can provide primary and secondary domain name resolution, E-mail via POP3 and SMTP services, and the World Wide Web. In an effort to generate even greater profitability and market diversity that extend beyond basic connectivity, a growing population of ISPs have begun offering other revenue-generating services such as Web

⁶ Market alert summary from a Cahners In-Stat Group report entitled "2001 Business ISPs-Service, Size, and Share"

Hosting, Virtual Private Networking, Voice over IP, managed security services, and remote data storage (archive) that can be serviced within the distribution network.

c. Core Network

The core network is that portion that provides the general data transfer service and connects the ISP to the wide area network (WAN) or the Internet, writ large, via an array of high-speed routers. Functionally, the core network is used by all the different applications that run over it and connects the local ISP to other ISPs at the central IXPs described previously.

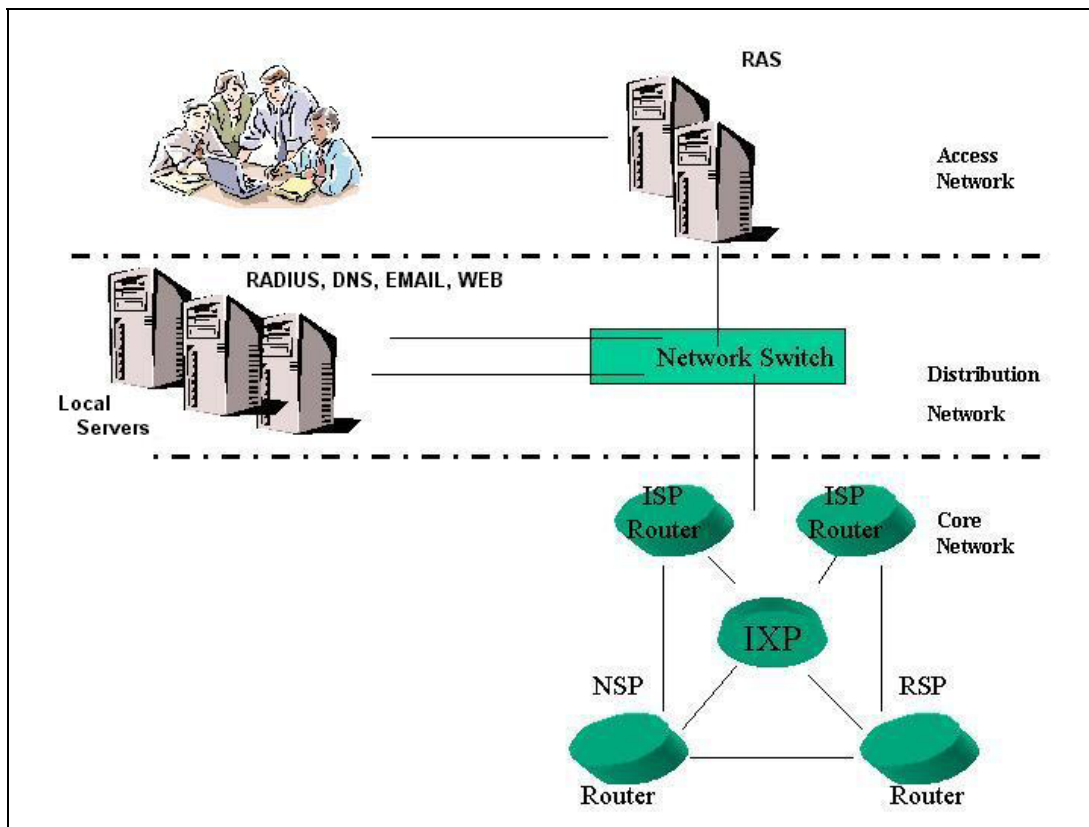


Figure 3. ISP Internal Network Architecture

4. Internet Routing and Associated Protocols

The “network of networks” that populates the Internet is connected by an array of high performance routers owned by the community of ISPs previously described. The process and flow of data, in the form of IP datagrams, requires that the connected router decide the optimal path to the next router along any path or, if at the intended network border, find the destination host. This series of routing across the entire infrastructure is

facilitated by, in fact depends on, accurate dynamically updated databases or routing tables resident within each router. These routers are used to move traffic within particular networks, known as autonomous systems (AS), or between them. Functionally, the router will appear as a connected host and will operate intelligently based on one of three gateway protocols: Routing Information Protocol (RIP); Open Shortest Path First (OSPF); and Border Gateway Protocol (BGP). Regardless of type, all three use distance-vector algorithms to compute routing tables locally and are considered dynamic in that they interact with adjacent routers to “learn” which networks or system each router believes is connected; and then, adjusts its table as conditions of the network change. When a BGP router advertises a particular route to another, several attributes are automatically associated with that route. These attributes are used to select or deny specific routes that reflect the ISP interconnection and peering policies. In terms of infrastructure security, the router network and associated protocols represent one of the major critical vulnerabilities that must be considered on balance with providing transparent and high performance connectivity. As we shall see, security solutions must foster higher levels of resistance against fraudulent updates in a manner that internal security procedures can use to both track incidents and apply upstream/downstream countermeasures.

C. STATE OF INTERNET SECURITY

As organizations began to rely heavily upon the networked computer environment (Internet) to enhance the use of information, increase creativity, and foster communication of ideas, the importance of information and its supporting systems mandate that effective measures to protect that information from unauthorized access and destruction be employed. Accordingly, arrays of control were developed to form the basis of protection and security that co-evolved along with the implied requirements of the computing environment. The earliest days of computing, pre-1980, were characterized mainly by the implementation of physical controls that prevented intruders from gaining access to the facilities that housed the computer resources. From 1980 through the mid-1990’s, the environment dramatically changed with the introduction of the desktop PC, client/server models, and a variety of operating systems like UNIX and Windows that

necessitated more technical measures to ensure protection. New forms of security controls that addressed system access, integrity, user authentication, and confidentiality from both local and remote entry points were introduced.

As the computer user base and importance of rapid information exchange expanded, new generations of operational controls in the form of policies and procedures were required to manage the risk of greater exposure. From an ISP perspective, the development of industry-specific best practices evolved to provide a framework for managing the Internet infrastructure in ways that made it “reasonably resistant to known security vulnerabilities” and “not easily hijacked by attackers for use in subsequent attacks” [Ref 4]. It is at this point that the modern mechanisms to enforce Internet security can be identified to capture the state of the "defense in depth" concept used by ISPs to defend the infrastructure. The deployment of these mechanisms can be divided into two distinct technology-focused phases and one, newly emerging, organizational context.

1. First Generation Techniques

Within the subscribers network and ISP internal configurations, point solutions such as firewalls, considered both a device and filtering mechanism, became the first generation of network security measures deployed in a perimeter defense to address the obvious vulnerabilities associated with access. Because of the expanding connectivity of computer networks to include remote users, firewalls were designed to mitigate the clear risks of unauthorized intrusion from the internet. The basic premise behind firewall technology emerged to compartmentalize the network into protected segments by establishing guarded gateways that serve the purpose of keeping users and information on the inside safe from access by non-authorized users on the outside. This strategy was based on the assumption that everyone inside the perimeter was trusted and everyone outside the gateway was not. In the ISP context, this task is considerably more complex given the connectivity model is based on any internet user being afforded access to business-related services inside of the network, a web server for instance. To implement reasonable identification of source traffic to counter the threat of denial of service,

various techniques of filtering were developed for implementation at the border routers of a subscriber's network and as applicable within an ISP's core network [Ref 5].

a. Ingress / Egress / Directed Broadcast Filtering

From an ISP's perspective, ingress filtering on source address is defined as the direction of filtering from the subscriber's network to the Internet. Given that denial of service attack model, outlined in Chapter II, is constructed to preserve the anonymity of the perpetrator, the attacker will attempt to obscure the origin of the attack by forging the source address of the attacking host machine in a way that exploits the trusted relationship between host machines on a network. Procedurally, an ISP will filter all traffic coming from the subscriber that has a source address other than an IP address assigned to that network within the access control list (ACL) for that router. In cases where ingress filtering is not possible for any particular ISP, the ISP is to "encourage" its use within its subscriber networks as close to the individual host machines as possible.

Egress filtering on source address, uniquely identified to acknowledge the perspective of bi-directional traffic, is defined as the direction of filtering from the Internet (ISP's network) to the subscriber's edge network. To reduce the vulnerability of an attack going into the customer's network from the Internet, ISPs are encouraged to filter traffic going into the border router that has a source address of any IP address(s) that have been already assigned to that customer. Further, specific policy filters on IP broadcasts to known broadcast addresses within internal subnets are recommended to limit multiplication type attacks associated with denial of service profiles. Again, if not possible, the ISP should encourage its subscribers to implement the technique.

Despite the extensive documentation associated with the benefits of employing these particular techniques, it is suggested elsewhere that they are not widely incorporated across the Internet for various reasons that include general risk assessment, inadequate training, or improper configuration practices. A primary resistance to its use can be put in terms of work load for an ISP with a large presence. For example, an ISP handling thousands of small business accounts would need to incorporate and actively manage an equivalent number of entries within the router ACL. More broadly, the direct beneficiary of the concerted effort is not the diligent ISP that acts as the transit agent for the malicious traffic; yet, that ISP bears all the financial costs in manpower or potential

network performance degradation. Direct benefit is gained by the destination ISP network that achieves a security service for no added cost in performance or labor. In the case of universities functionally performing as ISPs for their community of users, to include: researchers; administration; students; home networking; and distance learning enclaves, mechanisms like firewalls and filters are often considered hindrances to the pursuit of research and academic freedom or convenience. For others, the scale of their Internet presence in terms of hosts and services prove challenging to field adequate IT departments to operate and manage, or perhaps fund. Particularly relevant to DDoS, the lack of security associated with large universities has been so pervasive that many institutions received unwanted national recognition for hosting a majority of the attack zombies used to assault against Yahoo, eBay and other high profile e-commerce sites in February 2000⁷. Regardless, as currently implemented in the border router paradigm, this filtering mechanism remains susceptible to subversion if an attacker forges the source address of a host within the permitted IP filter block necessitating additional efforts.

b. Route Filtering

As previously described, the Internet connectivity is based on a scalable architecture of high end routers that continually exchange routing information (tables) between peers to advertise possible destinations and distribution paths between/within autonomous systems. These same routers may become overloaded in situations where excessive numbers of routing updates are exchanged in the normal course of operation. Beyond the context of normal failures or inherent router behavior traits, commonly referred to as oscillations, there exists the concern and possibility that the infrastructure is susceptible to subversion by hackers that could leverage this vulnerability on which to execute a denial of service⁸. As such, ISPs are tasked with implementing strong authentication processes [Ref 6] and damping mechanisms, known as BGP Route Flap Damping [Ref 7], to limit the risk of excessive loading between routers as well as any

⁷ In an interview conducted with ComputerWorld magazine (February 2001), Jeffrey Hunker, former Clinton administration's director of critical infrastructure outreach program, was quoted as saying "Universities were a major contributor to the DDoS attack". Many of the offending universities cited in the interview as facilitating the attack included prominent institutions such as Stanford University, Oregon State University, University of Washington, James Madison University, and The University of California at Santa Barbara.

⁸ InternetWeek, December 2001, interview with Carlos Recalde, Director of Telecommunications at KPMG conducted by Rutrell Yasin.

downstream router that could be affected from a cascade of peer failures. RFC 2439 makes note that this technique is being implemented within all BGP supported commercial products. However, continued effort to enhance the overall security of the routing infrastructure is being conducted in the form of revised standards project known as Secure BGP as a more complete solution.⁹

c. Authentication/Encryption

One of the primary mechanisms to protect the protocol updates from attack is through the use of router authentication. Authentication methods include both plain text and Message Digest Algorithm Version 5 (MD5). MD5 differs slightly from plain text in that no authentication key is exchanged, instead transmitting only a hash of the message. Still in use today, MD5 is compatible with the family of gateway protocols (BGP, OSPF, RIP, etc.) and remains the primary tool to validate the authentication of routing updates. Without it, malicious parties can divert or analyze the exchanged traffic. For example, a fictitious route injected into a router could divert legitimate traffic to an incorrect or bogus destination.

Encryption prevents unauthorized users or “third parties” from being able to read transmitted information even if they eventually gained access to it. While employed extensively by subscriber networks to protect data, use of encryption in the ISP context is largely restricted to local administrative processes such as terminal communications with network infrastructure devices during configuration management practices.

d. Black Hole Routing

A slight variation on the aforementioned packet filtering techniques is to create specific listings of static routes, the traffic intended to be removed, and forward the associated IP traffic into a pseudo-interface identified as Null0. This has the unique feature that when implemented, the Null0 interface does not forward or receive traffic and drops the packets in a manner that creates no associated processor overhead. Leveraging the inherent strength of routers to forward vice filter, this mechanism helps to avoid performance degradation associated with large ACLs. This is considered a drastic

⁹ Secure Border Gateway Protocol Project, <http://www.net-tech.bbn.com/sbgp/sbgp-index.html>

measure by ISPs in that all services associated with a particular site are then restricted vice just the offending packets.

2. Second Generation Techniques

As networks grew more complex in design and more varied by their supported platforms and operating systems, the security market responded by developing a sophisticated set of mechanisms to help proactively identify vulnerabilities before an intrusion occurred as well as a system to recognize when an attack was underway. These techniques are broadly categorized as vulnerability scanners and intrusion detection systems. Commensurate with the complexity and exponential growth of the Internet, two additional mechanisms, referred to as load-balancing and rate-limiting, have been employed to help mitigate the effects of denial of service.

a. Vulnerability Scanner and Probing Programs

Vulnerability scanners are both vendor specific and open source products designed to discover network vulnerabilities to known intrusion-based exploits. While each provides slightly different features, their primary focus has remained in identification of previously recorded vulnerabilities and of coarse-level risk analysis through the use of threat condition ranking (low, medium, high risk) profiles. Once known, administrators could then apply corrective software patches to eliminate the vulnerability or minimize its impact on the system by configuration changes. Product names such as TripWire, LanGuard, SATAN, and Nessus fall within this category and are largely deployed by subscriber networks, and some smaller service providers, to assess the security posture of their internal networks. However, these and other programs are not widely used by the ISP community to sample the state of subscriber defenses unless conducted within the context of their managed security service offerings to be detailed later.

b. Intrusion Detection System

Intrusion Detection Systems (IDS) based on threat profiles and sophisticated algorithm techniques, both rule and anomaly based, emerged to proactively identify and track patterns of activity that could signal potential intrusion attempts or misuse of information resources at near real-time levels. As in the subscriber network

architectures, ISPs are employing the IDS networks to monitor their access and distribution networks at both the host-based and network-based levels to detect and respond to inbound distributed attacks or malicious code. Accordingly, the IDS was designed to monitor for known attack signatures and sniff out suspicious network behavior. When it finds unusual activity, the IDS will send an alert to designated operations (IT) staff in the form of pages and automated e-mails, while logging and reporting the intrusions progress. Subsequent to the incident, audit logs are then analyzed to compile forensic data for evidentiary purposes and investigation, to feedback into the attack pattern library, or terminate the offending connection. The bulk of the effort to screen audit logs is still conducted manually by many organizations. Thus, the major failings with the IDS implementations, as currently available, remain that (1) the technique does little to deal with the immediate problem or stop the attack in progress, and (2) dealing with the large volumes of log data that is generated from the army of collectors/sensors deployed in any particular scheme. These deficiencies are seen as a major contributor behind why participating networks do not use, or stop using, IDSs.

Advances in the technology have begun to allow deployment of an IDS concept known as an Intrusion Prevent System (IPS), which looks out for the same signatures and anomalies, but focus equally on monitoring the behavior of network devices for indications of suspicious behavior. If an infected server or network device tries to execute a behavior out of the norm, the IPS will automatically neutralize it with a specific countermeasure without consideration to the type attack. The difficulty with the use of these technologies is in the area of accurately base-lining proper behavior of network machines to maintain a balance between responding to legitimate attacks and of acting on false positives at real time speeds. To address these issues and others, industry efforts remained focused on improving overall IDS functionality. Specifically, vendors are introducing more sophisticated correlation engines to better link the state of the network as a whole. Advanced visualization processes are being added at the monitor consoles to gain greater insight into how the deployed sensors are performing and to increase operator usability. Other groups are emphasizing automated policy management routines to help keep pace with network capacity. As the cost of such efforts continue to climb beyond the reach of many, subscriber organizations have turned to a third

alternative that emphasizes outsourcing of the same tasks. However, a fundamental limitation of an IDS approach is its attempt to provide an omniscient classifier of all traffic. Attempts to construct such mechanisms run afoul of the fundamental theories that underlie computational systems. Therefore, IDS will always be a mechanism that the sophisticated attacker can circumvent.

c. Load-Balancers

In response to the growth of World Wide Web, companies began to increasingly outsource their Web based presence, in the form of massive server farms, to ISPs under the scope of their managed services. Commensurate with these large architectures, load-balancers have been developed to manage the million connections per second requirements placed on Web servers or other resources. In essence, these load balancers are provided to dynamically distribute incoming requests across a group of servers running a common application or set of applications in a manner that the grouping appears to the client base as one server. Upon receipt of a connection request, the load balancer will pass the request to one of the servers based on specific criteria of availability, server health, and load handling capacity. In terms of security, the load-balancing technology has an additive quality in that it maintains state or knowledge of individual sessions by source origin, connection start, connection end, and real time status of individual resource loading. Additionally, inherent in the design is the fact that the load-balancer presents a virtual IP address to the requesting Internet client, thus hiding the presence of all attached resources akin to the Network Address Translation (NAT) technology used to address the decreasing IP address pool under IPv4. Retaining knowledge of state parameters, the load-balancer can dynamically shift to a low-load server, firewall, or IDS component as warranted to mitigate availability-related attacks against the internal infrastructure.

d. Rate-Limiting

As in the case of load balancing technology, rate-limiting mechanisms are often considered within the rubric of dual-use security. Initially conceived to optimize network connectivity, in terms of meeting quality of service levels within a service level agreement, rate-limiting techniques were developed to control the allocation of bandwidth and traffic delivery rates. Otherwise referred to as traffic shaping and traffic

policing, development of rate-limiting within routers provides ISPs with an ability to restrict traffic flow outright (drop) or to prioritize it, via a queuing technique, for preferred handling along the core router network infrastructure. Rate-limiting using the traffic shaping technique controls both rate and volume by dropping packets. Closely aligned, the traffic policing technique makes use of packet marking so as to defer the decision to drop a packet to the core router network when congestion actually develops. The objective of traffic policing, however, is to avoid dropping packets. In order to provide the most flexibility, rate-limiting can be employed by routers in one of three methods: port rate limiting, aggregate rate-limiting, or flow basis.

As depicted within Interface B of Figure 4, service providers desiring to restrict bandwidth allocation on either an inbound or outbound physical port, regardless of data traffic type or protocol, utilize port rate-limiting. In this manner, differential service, as potentially outlined in a service level agreement, can be achieved at a specific port to deliver varying levels of uplink/downlink capacity. Traffic that exceeds the bandwidth threshold can then be dropped outright or prioritized by overwriting the "type of service" field within the IP protocol header. Figure 4 reflects a 600 megabit per second (Mbps) rate limit at the point of entry for interface B, regardless of source/destination address or traffic type.

Alternately, aggregate rate-limiting can be utilized to restrict bandwidth consumption in terms of a specific protocol or traffic pattern based on a pre-established policy statement. By using traffic policing rules, limits can be applied across a multitude of applications based on source, destination, port number or protocol type fields within the IP header. Application traffic exceeding a predefined threshold can be dropped outright or prioritized using the "type of service" field in the IP header just as in the case of port rate-limiting previously described. Figure 4 reflects that given a similar 600 Mbps (A) interface, HTTP protocol traffic is restricted to 300 Mbps while others are subdivided into increments of 200 Mbps or less.

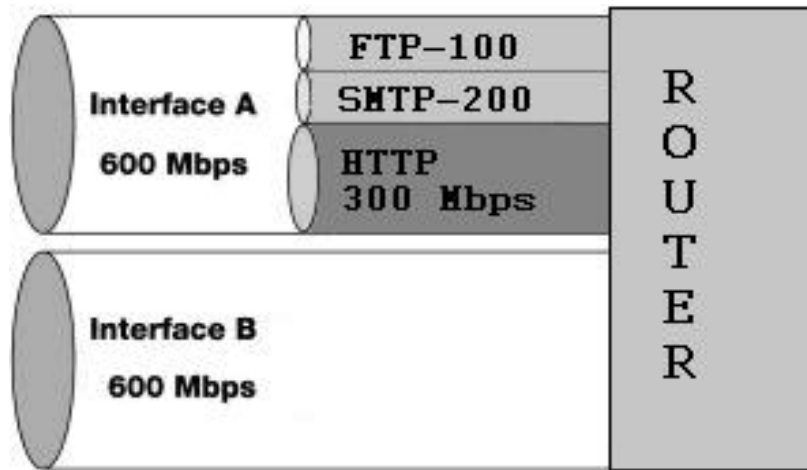


Figure 4. Port and Aggregate Rate-Limiting

Somewhat similar to aggregate rate-limiting, flow limiting adaptively restricts bandwidth or prioritizes traffic on the basis of a connection. Traffic policy is constructed to utilize source and destination addresses, port numbers, and protocol types of all the related packets in a stream and thereby applies an upper limit to overall bandwidth consumption during a particular initiated connection. This is a technique used often within switched routers that precede large server farms. Figure 5 depicts the total bandwidth of 600 Mbps available with interface B being evenly distributed between two TCP/IP sessions.

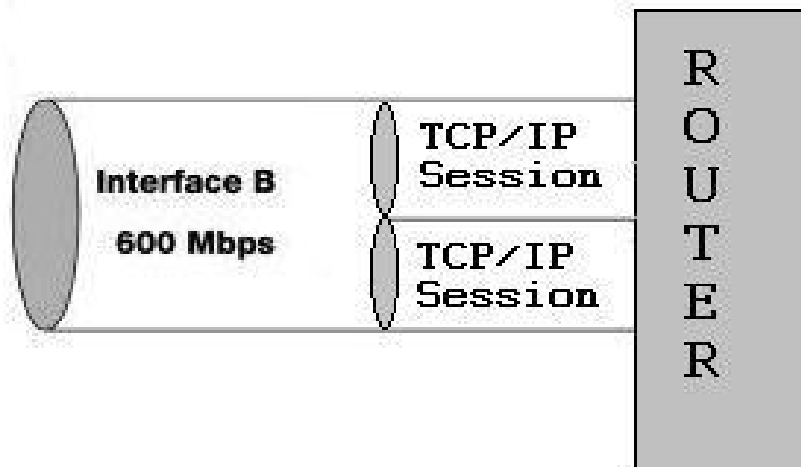


Figure 5. Flow Rate-Limiting

3. Organizational Approach To Implement Distributed Security

Many security programs are never adequately developed or correctly implemented. This can occur because a large base of network operators and end users often lack the requisite level of expertise to analyze security risks or perhaps the time to keep pace with rapid information technology change. Yet, it is fully appreciated that an effective security program requires a concerted effort to establish and is, by its very nature, difficult to maintain. Recently, organizations looking for help with their overall security programs have turned to outsourcing solutions that are centered on the concept of a managed security service provider (MSSP). This emerging trend to realign security responsibility, albeit in the form of a paid service, demonstrates the potential advantage gained from an organizational change approach to security rather than one based solely on technology

The MSSP community is generally comprised of independent and spin-off security focused companies that have emerged largely from established security consulting/integration firms. TruSecure Corp., The Salinas Group, and Counterpane Systems are but a few of the current providers. Similarly, the service may be provided by the traditional ISPs that have added point solution security devices and monitoring services to their list of their network management services. In both cases, MSSPs can be

hired to monitor and manage a variety of network components and services to include the range of anti-virus software/malicious code protection, firewall management, intrusion detection systems, perimeter scanning (penetration testing), and protected Web servers. With a current market share of \$630 million, estimated to go to \$2 billion by 2005¹⁰, the managed security concept also demonstrates an economic model that lends itself to the ISP-centric approach.

a. Advantages of the Managed Security Approach

When integrated within the ISP context, this approach can directly link the ISP to evolving security requirements of its networks. The ISP can often standardize, through restricted support packages, the types of applications, operating systems, and security hardware deployed over the Internet. For example, an ISP may support only NT and Solaris operating systems. For software/hardware combinations, they may choose a Sun server with Oracle and MS Access for database platforms. This approach can often limit the proliferation of untested or poorly researched equipment while bringing in economies of scale. Other MSSPs have demonstrated a greater degree customization by handling a variety of security products that can be selected from and are capable of being incorporated with subscriber specific security policies. In either case, economies of scale allow providers to set prices within the grasp of even the smallest businesses and end subscribers. An additional small network or home machine can typically be added to an already x-thousand node network with little to no upgrade in ISP resources. In addition to efficiencies associated with scale, standardization facilitates predictable risk and performance. Tracking and performing patch fixes across the homogeneous network remains a simpler task for administrators.

The dual-hatted ISPs are better able to reserve contingency bandwidth across multiple backbone providers to assure level-of-delivery standards required by its customers. In a better position to employ load-balancing technologies, large and medium sized ISPs are uniquely suited to employ dispersed server/service strategies to ensure availability in face of malicious attacks. ISPs acting in this capacity are also uniquely suited to respond across the entire infrastructure. In observing an attack profile, or trend, against one customer, they can move to protect other clients before the same happens

¹⁰ Source: The Yankee Group

again. No single network or end subscriber can develop this level of internet perspective to counter large-scale malicious activity.

Lastly, an ISP offering the value added MSSP service concept allows large multinational organizations, who typically have major network resources widely dispersed around the globe, and smaller users to augment their in-house IT security staff. MSSPs can be better positioned to be central warehouses of extensive security related experience, as opposed to any particular network subscriber that may have never experienced an attack or becomes complacent due to the infrequency of service disruption. As a centralized resource to collect valuable forensic information, the MSSP is ideally placed to contribute to the overall trend analysis of newer availability attacks or forward alerts to law enforcement agencies.

The cost to achieve well-managed security mechanisms is high and includes investments in hardware, software, and personnel. Given that it is difficult for any organization to quantify the return on investment for any major security installations, internet users and businesses are able to look to managed security as an economical way to address security for their operations.

b. Disadvantages

Recent economic conditions have demonstrated that outsourcing security has its conditional pitfalls. By the end of 2001, it was estimated that only 50 managed security providers remained in the market, down from an estimated 80 companies tracked by the Yankee Group¹¹. It had often been the case that security focused companies went out of business with no contingency plan to move customers to alternate providers or to transition the same customer to self-sufficiency. Subscribers were subsequently forced to pick up the pieces following a collapse or rapidly find an alternate service to run existing, but often times unsupported, equipment and configuration schema. In this regard, security analysts have been advising organizational strategies that spread out managed security services among multiple firms.

Beyond economic factors, other issues have cast an unfavorable light on the outsourcing approach. Most problems have been associated with the rush to market

¹¹ Edmund Dejesus, "Information Security Magazine: Managing Managed Security", January 2001.

by companies unprepared for the task or the problem itself. For example, initial offerings by some of the largest carriers centered on firewall management services. Yet, surveys conducted by analyst suggested that 30% of respondents felt their own IT staff members were more qualified than their contracted carrier representatives. In areas of Service Level Agreements (SLA), contracts that govern the roles and relations between provider and customer, the task of determining achievable levels of performance proved difficult for the earliest pioneers. Given that none existed previously, MSSPs and edge networks are still in the process of identifying the criteria and guarantees for levels of service and security needed to conduct operations, especially business, over the Internet. Obvious questions will arise in determining how much proprietary information or access to sensitive information is needed by the provider to ensure adequate protection for a customer. Further efforts will be required to determine the level of specificity required in a SLA to define the boundaries of incident response, routine testing, audits, and penetration exercises. As the managed services multiply or become more diverse, the SLA will also become necessarily complex and will have to incorporate provisions for variances based on location, penalties or refunds for violations, and incentives for increased performance measures.

D. INTERNET REQUIREMENTS TRANSFORMED

The Internet architecture, in general terms, has evolved in a manner that reflects the relative influence of all its principle stakeholders over the years. Thus far, all of the stakeholders can be categorized in one of five basic sectors that comprise the Internet community: universities and research institutions; the competitive marketplace of ISPs; commercial enterprises and small businesses; the telecommunications industry; and lastly, the public domain. The Internet has flourished because each of these stakeholders has been guided by one overarching goal to maximize connectivity, a single-minded purpose that has provided its own reward of global interconnectivity. However, as is the case of any system lifecycle, the conditions that influence its use change over time. Today, the goals of reliability and assured availability have risen to the forefront because of the relative position that the Internet has garnered within the fabric of critical infrastructure not to mention the financial aspects of commerce. Accordingly, new sets of

conditions have emerged to characterize the Internet in a manner that exceeds the capacity of subscriber-centric approach to effectively deal with its most disruptive threats. These conditions are divided into the following categories [Ref 8].

1. Erosion of Trust

The quality of openness of the Internet remains one its most powerful and useful features. However, one need only take account of the patchwork of point solutions to limit or monitor the conduct of its users to recognize that the computing environment has long since transcended its founding principles of trustworthiness and cooperative behavior. Regardless of how well one subscriber node postures itself in terms of security, the susceptibility to attack often depends on the state of security at another node. And yet, they have no influence on the security condition of that node. Increasingly, complex software is introduced with minimal consideration to security resulting in a cycle of vulnerabilities and exploits. Moreover, attack trend and analysis efforts suggest that the time between vulnerability discovery and exploitation is rapidly decreasing [Ref 9]. As the basis of trust erodes between parties that still have vested interests in communicating, a need for intermediaries to intercede becomes required.

2. Unsophisticated User Base

Many network computer systems remain increasingly vulnerable to attack in part because various elements of the subscriber base do not implement security measures that are already available. This deficiency extends from individual users to some of the largest university and enterprise networks. This has been attributed to several reasons such as lack of training and awareness, high costs of implementation, and complacency. For example, the migration from dial-up access to broadband services like DSL and cable has been ongoing for some years. However, review of open source press accountings suggest that many users are unaware of the vulnerability associated with the “always on” connection despite the availability of host based personal firewalls, procurable commercially or from open source forums. Worse yet, a Security Focus article described an event in which a national broadband service provider discouraged the use of firewalls on their connection because of configuration incompatibilities¹². Others suggest that larger organizations operating a diverse array of resources as part of their heterogeneous

¹² <http://www.SecurityFocus.com/news/287>, “Broadband ISPs Shouldn’t Knock Down Firewalls”, November 20, 2001.

networks simply can not keep pace with the practical process of tracking and installation of patches across the various operating systems and applications made available by product vendors. As Internet ready technology (PDAs, cell phones, game boxes, house appliances, etc.) become more pervasive, configuration issues will inherently challenge the overall security environment.

3. Sophisticated and Ubiquitous Attack Techniques

Attack technology is constantly evolving. In today's environment, attack and intrusion techniques continue to surface that demonstrate an increasing level of sophistication. For example, long past is the situation in which delivery of viruses and Trojan horse programs occurred from exchange of floppy disks between two users. Leveraging the growing standardization among PC users, malicious code developers can efficiently reach a larger base using a disproportionately small number of tools. With some 90 percent of the world PC market using Microsoft Windows products, malicious code writers can introduce a single virus particular to the MS operating system to achieve a more global impact. While efficient from a management or training perspective, the state of standardization effectively removes a natural barrier to propagation that may have been present given a more diverse computing environment. In other areas, the addition of newer polymorphic worm engines provide hackers with an automated ability to subvert intrusion detection systems by creating functionally equivalent attacks with different signatures as they spread between hosts. While the expertise required to develop sophisticated programs may extend to a relatively small group of hackers, the vast number of openly accessible "hacker" web sites provide access to powerful tools and exploits that lowers the overall barrier for entry and make it possible for the unsophisticated to achieve commensurate disruption.

While much of the focus of computer security deals with the external penetration by malicious users, sophisticated security attacks can be launched through a process defined as subversion. Subversion is uniquely identified because it involves the covert and more methodical compromise of the external and internal controls over the lifecycle of a computer system to allow unauthorized and undetectable access to system resources and information [Ref 10]. Subversion can involve the implanting of "artifices", Trojan Horse or backdoor programs, at any one of the design, implementation, distribution,

installation, or production phases of a computer system or software program. The artifices can be constructed to be completely undetectable by automated scanning techniques or manual inspection. One need only consider the more innocuous type of code known as the "easter egg" that commercial software developers embed in many popular software programs to appreciate the opportunity for subversion. A widely recognized example was the Flight Simulation program hidden within Microsoft Excel.

4. Evolving Influence of Legal Liability

As a major component of the critical national infrastructure, both within and outside the United States, the Internet must continually deliver on its fundamental promise of availability and reliability. In this vein, its future potential and contribution is governed by the level of confidence felt by participating stakeholders as they conduct business, of any kind, across the infrastructure. A commonly held, but unspoken, confidence in the Internet has seemingly been present since its inception. Arguably, it has been a major condition in its exponential growth. However, the 2001 request for information (RFI) submitted by the Director of White House Office of Cyberspace Security, Mr. Richard Clarke, to consider the feasibility and economic impact of a separate GOVNET began the serious debate regarding major Internet players disconnecting from the public infrastructure.

As previously described, the direct economic costs associated with a DDoS or worm attack can be sufficiently large enough to cause a major financial crisis for any organization. But it is the loss in public trust, customer goodwill, organizational reputations, and spillover effects that comprise the indirect costs sufficient to undermine market/sector stability or, much broader, the national economic condition. Although no such galvanizing attack has yet to occur in cyberspace, one need only consider the financial impact on the airline industry, leisure/travel industry, and U.S economy which resulted from the September 11, 2001 suicide terrorist attacks on the World Trade Center and Pentagon to gain an appreciation for the interdependence of institutional systems and public confidence.

The Internet's vulnerability and susceptibility to attack by malicious parties, especially in the form of DDoS and Worms, has begun to erode the requisite levels of individual and institutional confidence to a degree that an emerging condition of legal

liability, particularly for the commercial stakeholders, is making the existing security paradigm untenable. For example, it was reported that Washington's state attorney general demanded Qwest Communications (an ISP) to refund its business and individual digital subscriber line customers for service disruptions following the Code Red worm attack¹³. Citing the thousands of dollars in economic losses and poor response efforts from the company's technical staff, this incident illustrates an emerging consensus that ISPs must shoulder more of the direct burden to eliminate such attacks as well as bear more of the financial costs for failures in contracted availability. Should future judicial proceedings directly address the issues of negligence and security on the Internet, their rulings may influence a redesign of the Internet that makes an ISP centric approach sensible from a long-term financial point of view. To understand the changing requirements regarding an ISP's responsibility to provide for higher levels of security, the foundations of tort liability are briefly explored to establish the conditions that necessitate the shift in the security paradigm. Similarly, the state of the legislative environment will be described to contrast its influence over the Internet stakeholders to deter and defend against denial of service and worm attacks.

a. Tort Liability

In criminal law, DDoS and Worm attacks are already provisioned for within the Computer Fraud and Abuse Act, as are most forms of malicious computer activity. Thus, in context of causing harm, judgments regarding liability of the perpetrator have been established. Tort law, however, is more broadly understood to encompass the legal mechanism intended to deter undesirable behavior and to compensate those damaged by the action. Thus, tort law encompasses protocols that are based on harm caused through negligence and harm caused without fault. Causing harm without fault, also known as "strict liability", is tightly associated with claims over defective products that demonstrate a danger to public use and would not generally be considered in the context of Internet security. Assuming no other malfeasances, an automobile safety airbag that fails to deploy in an accident would fall into this category.

¹³ <http://www.infosecuritymag.com>, Information Security Magazine, Security Wire Digest, August 27, 2001

However, taking insufficient precautions against known risks has long been an established tenant underlying judicial findings of negligence because of three intertwined evaluation criteria: least-cost avoider; state of the art; and degree of control [Ref 11.]. In terms of security, failing to adequately defend computer networks may carry with it an increased level of legal liability that may compel ISPs to shoulder a larger responsibility for the overall Internet security posture regardless of having complete solutions available.

1. Least-Cost Avoider Principle. Within the general bounds of liability, the least-cost avoider principle has been adopted as an evaluation benchmark that seeks to assign legal liability to a party who is in the best position to know of an associated risk and to take the appropriate action to minimize that risk. In situations where no clear party can be determined, the least-cost avoider principle allows for the assignment of liability to a party that is simply in the position that can most cheaply discover who is the least-cost avoider. For example, an ISP's Term of Service agreement that states that the ISP is not responsible for the content within an e-mail is an attempt to contractually transfer liability to the least-cost avoider, as the subscribers are in the better position to protect or filter their own electronic correspondence and can provide that filtering more cheaply. In the context of DDoS and Worm attacks, the "hacker" behind the event is already determined to be criminally and tort liable for perpetrating the action. Unfortunately, most individuals that can be held responsible would likely possess insufficient finances to make a victim financially whole again. However, the victim sites themselves and the ISPs that provide the connectivity infrastructure may eventually be held liable for incurred losses since they may be in the best position to payout claims under the normal umbrella of business insurance or large financial resources, as in the case of the federal government.

2. State of the Art. Tort Liability is also deeply intertwined within the concept described as "state of the art" [Ref 11]. This principle is another evaluation benchmark used in an attempt to incorporate a "level of sensitivity" in regards to the availability of cost-effective solutions or precautions. Measures of this nature generally fall within the scope of current technology and the realm of best practices. While the current state of security is characterized by point solutions to combat malicious

attack, other technologies and procedures are just beginning to mature that could enable infrastructure-wide protection schemes. For example, the evolution of filter technologies within the Internet core routers, the wide spread deployment of network-based intrusion detection systems, or scope of reasonable incident response may all serve as minimal standard of practices. Once fully matured to be an effective counter to DDoS or even implemented by the majority of organizations as incomplete solution sets, the least-cost avoidance and state of the art evaluations may allow for the assignment of liability to the ISP if not adequately provided for during an attack. Still, organizations can be found liable for failing to exercise reasonable duty of care, or reasonable prudence, despite their compliance with standard industry practices. Considered experts in a field of connectivity, ISPs would be expected to stay informed of related advances and responsible to implement changes as warranted, even if not widely used.

3. Degree of Information and Control. “Degree of Information and Control” is described as the third evaluation benchmark that courts consider in the process of specifically identifying the least-cost avoider and assigning liability [Ref. 11]. This principle basically analyzes who is in the best position to exercise the necessary level of prudence and foresee potential harm. The intended targets of DDoS and worm attacks are often the commercial or military web hosting services and associated servers. Similarly, a growing number of these integrated services (information) are being housed, monitored, and managed (control) from within the physical and virtual facilities of ISPs or entirely by an ISP’s hosting division staff. Combined, these situations place the provider in the best position and with the most ability to prevent attacks. Accordingly, the ISP are conceivably more exposed under tort law if their service performance lags the standard of care afforded under the aforementioned state of the art analysis. Given the difficulty in tracing the creator of a worm, the most practical target of litigation is likely to be against those that failed to patch the systems.

b. Immunity

Under the current security paradigm, the providers have conducted themselves in a manner characteristic of agents for connectivity or as an underlying backbone that facilitates communication and data exchange between parties. Thus, ISPs have argued that they should be considered beyond the reach of liability for conduct of its

users. In fact, the safe harbor provisions extended to ISPs within the Communications Decency Act of 1996 related to indecent or defamatory content and the Digital Millennium Copyright Act of 2000 related to copyright infringement would seem to bolster their position. It must be noted, however, that both of these legislative measures give specific statutory exemption to a registered ISP in an attempt to foster dissemination of content. The motive behind these exemptions is explicitly linked to the particular sensitivity that the U.S has in protecting freedom of speech. Additionally, both statutes afford the safe harbor provision only under certain circumstances or conditions and are not considered absolute. For example, ISPs must have public policies against copyright infringement and remove/block instances in an expeditious manner when discovered or if appropriately notified. In terms of indecent content, ISPs must inform network users of screening software and report known violations, as defined by the Child Online Protection Act, or face financial penalty. Since activity associated with DDoS and worm attack would conceivably fall outside the context of free speech, it is unlikely that ISPs will be provided similar safe harbor.

Contractual shifting risk can provide an ISP with some degree of immunity against liability. As a strategy, contracts can be used to shift the security burden onto subscribers by compelling customers to take certain precautions or shift responsibility by simply disclaiming it as a term of service condition. Contracts that compel customer behavior to take precautions are common within the same Service Level Agreements that governed ISP peering arrangements or Term of Service Agreements used by end subscribers. Disclaiming liability, however, is considered to be non-binding on third parties participants, parties that do not directly sign the contract. So in the scenario of a DDoS, it would be unlikely that an ISP would be able to shift liability to a third party internet user (non subscriber of the hosting ISP) infected with zombie or handler programs to recover damages sought by the victim web site. Furthermore, disclaimers are only enforceable between signatories if they remain consistent with currently held public policy, appropriate jurisdiction, or considered non-coercive. Given an ISP's national or global presence, contractual immunity may be inconsistently applied or costly to defend and illustrates that further action is warranted.

c. Legislative Environment

In concert with the expanded use of the computers, Congress has enacted several laws that recognize the detrimental effect of malicious activity over the Internet. However, legislation enacted to date has emphasized the criminal law aspects of regulation by aiming its efforts at individuals or groups who abuse the networks for illegal purposes and not at other third parties who facilitate, knowingly or not, in the attack such as in a worm or DDoS scenario. For example, under the Computer Fraud and Abuse Act, persons who alter, damage, or gain unauthorized access to content on the network computers of the government, banking, and credit card companies are subject to fines, jail, or both. Further, it imposes penalties for unauthorized access to other computers with the intent to extort or defraud. It does not address the third party complicity of the computer owner that has failed to adequately protect his machine that was used as a surrogate in the attack. The most recent initiatives such as The Digital Millennium Copyright Protection Act (DMCA), Computer Decency Act, and Electronic Communications Protection Act only address the narrow concepts of copyright law, free speech, and privacy within the Internet context.

Three newer legislative initiatives that have a dramatic impact on ISPs ability or desire to stay as a neutral connectivity partner have recently emerged to fill part of the void. Specifically, state legislatures in California and Virginia have taken up the issue of anti-spam and trespass law in ways that have analogous implications for DDoS and worm code. Spam, the Internet version of junk mail, can functionally achieve the same result of a denial of service against a “relay” mail server as it attempts to mass produce a single inbound message into thousands of copies for hosted accounts. Sufficient performance degradation or system crash can occur as a result of a condition known as bounce back, in which large numbers of undeliverable addresses, very typical of spam listings, cause large quantities of mail to come back to the relay server. As spammers attempt to disguise themselves or utilize stealth mailing techniques, administrative and technical costs soar at the ISP while trying to track, contain, or filter the offender.

The Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism Act (USA-PATRIOT) of 2001 has greatly

expanded police powers to obtain customer account information and record transaction history from any agency providing connectivity to the Internet. As ISPs seek to comply with their accountability and data archival specifications, they will inherently gain broader monitoring capabilities necessary to “see” more of what is sent over their respective network infrastructure. For example, the prevailing Cable Act of 1984 specifically prevented law enforcement wiretaps across cable communication lines or the ability to obtain customer account information. The PATRIOT Act has superseded this restriction and now governs disclosure of both voice and data services. And most broadly, PATRIOT provisions now allow law enforcement to obtain a single search warrant capable of crossing multi-state jurisdictions to collect and/or “trap and trace” data across the Internet infrastructure.

E. SUMMARY

Because of the exponential growth of the Internet, the resulting network computing landscape shows little resemblance to the functionality and design architecture originally envisioned by its founders. Today, the Internet is heavily influenced, if not dominated, by a diverse community of service providers that cooperatively engage in mutually beneficial peering arrangements to form the critical gateways to the infrastructure from which users and network systems interface. In addition to providing the access control and linkage to other networks that facilitates global interconnectivity, these same service providers are increasingly operating and managing a complete set of services such as web servers, mail, and data storage, that have redefined their role beyond the transparent, best-effort delivery of packets.

While the Internet has flourished based upon a requirement to maximize connectivity, the evolving usage of the Internet, coupled with an inherently insecure technical foundation, has given rise to a fundamental erosion of trust and decreased technical sophistication of users ill suited to counter the growing sophistication of threats.

Despite the growing importance of the infrastructure to the national aims, responsibility for the security of the Internet has largely remained within a paradigm based on trusted users to act appropriately to preserve its availability and reliability under

the aegis of security related to "best practices" and layered defense. Largely due to scale, complexity, and inherently insecure computing designs, the subscriber-centric approach has failed to keep pace with an emerging threat of denial of service and malicious code that can now leverage the distributed nature of the infrastructure itself to severely disrupt critical functions and services required by society, commerce, and government users. In contrast, the ISP community has demonstrated some of the fundamental capability to employ established techniques and organization concepts to prevent, detect, and contain the vast majority of attacks, or at least mitigate their effects until more focused countermeasures can be employed.

THIS PAGE INTENTIONALLY LEFT BLANK

III. DEFINING THE THREAT

A. INTRODUCTION

Interconnectivity, distributed computing, and near instantaneous speed of transmission are among the most powerful features of the Internet. Ironically, these attributes are also the qualities that jeopardize its long-term security the greatest. Hackers increasingly use these specific characteristics to exploit the networks of hosts that exhibit little to no security. Two of the more serious security threats emerging in today's networking environment are the distributed denial of service (DDoS) attack and the self-propagating Worm.

In February 2000, eBay, Yahoo.com, Amazon .com, CNN.com, and several other top e-commerce sites fell victim to a DDoS which rendered their Web sites inaccessible to legitimate users for six consecutive days and cost the targeted companies a combined estimate of \$1.2 billion dollars in lost revenue, additional security measures, and damage recovery measures. It was later discovered that a lone juvenile from Montreal Canada perpetrated the attack. In terms of larger market impact, the Federal Trade Commission estimated that as much as \$15 billion dollars worth of e-commerce transactions go unrealized because of the lack of trust in the current business-to-business and business-to-customer electronic relations that are susceptible to similar kinds of Internet based attacks. Similarly, a Gartner Group survey suggested that as much as 86 percent of American adults refrain from conducting personal business over the Internet because of security concerns. While most of what the public is made aware of comes from a few high profile incidents, it is believed that the majority of attacks are still not well publicized¹⁴.

Since recent examples of prosecuted hackers have been juveniles, there remains a prevailing attitude that the reasons behind the trend are themselves juvenile or simplistic in nature and lack serious merit. DDoS continues to be a pervasive threat and one conducted with serious underpinnings as demonstrated in both March and December 2001 when the widely recognized Internet security watchdog organization, the Computer

¹⁴ Computer Security Institute, "2001 CSI/FBI Computer Crime and Security Survey", Volume 12., No.1, Spring 2001.

Emergency Response Team (CERT), became a victim of a targeted DDoS attacks. Later the same year, selected Department of Defense and defense industry sites were taken down by Chinese hackers in response to the in-flight collision between the U.S Navy's EP-3 surveillance aircraft and a Chinese fighter patrol aircraft. By attacking a recognized "security conscious" poster organization, the former attack served to send the larger message of ubiquitous vulnerability - no matter what you do, where you are, who you are, you can be reached. This is a message and context not unlike that which sponsored terrorism tries to achieve. The latter case demonstrates that such attacks are emerging as new forms of political and social protest. Both instances suggest an evolving threat that goes beyond juvenile bragging rights. Still more ominous, experts continue to warn the nation's public and private enterprises of future attacks by ever more sophisticated cyber-terrorist acting either independently or on behalf of adversarial nation states as a means to conduct an emerging form of cyberwar. In 1998, hackers traced to the Persian Gulf region tapped into the NASA network to gain access to data sensitive to the national air traffic control system. During the Kosovo air campaign of 1999, Serbian hackers conducted coordinated attacks against NATO servers and U.S governmental Web sites.

As to the proliferation of Worm attacks, high profile incidents during 2001 demonstrated that the development and use of such attack programs are evolving faster than previously recorded, about one every three months¹⁵, and that each new generation is being engineered to improve its repertoire of exploited security vulnerabilities, self-autonomy, distribution/replication methodology, and destructive payloads. As to its potential impact on the larger Internet infrastructure, Code Red, for example, managed to infect some 359,000 hosts worldwide within a 14-hour period and was reported to have caused an estimated \$2 billion dollars in damage, making it the second costliest single outbreak in Internet history¹⁶. Gone unchecked, experts widely believed that it could have set off an unprecedented Internet-based attack given that a principle component was to install a legion of sleeper attack daemons (programs) needed to conduct a distributed denial of service.

¹⁵ Edward Skoudis, Information Security Magazine, "The Year of the Worm", September 2001.

¹⁶ Anna Saita, Information Security Magazine, "Right Back Atcha", September 2001.

B. DENIAL OF SERVICE / DISTRIBUTED DENIAL OF SERVICE

The primary objective of any Denial of Service (DoS) attack is to render a target computer, web-based service, or device useless by dominating its system resources. The direct impact will be to crash computer resources, block legitimate subscriber use of the Internet, and consume valuable bandwidth capacity to degrade overall Internet performance¹⁷. In the grand scheme, any device with an IP address or connection to a network infrastructure, to include wireless and mobile phone platforms, becomes vulnerable to attack. Particularly menacing is the conceivable threat that an attack aimed at the major IP based routers and IXPs, described earlier, could render large segments of the Internet inaccessible.

The distinguishing characteristic of a DoS attack is that it originates from a single machine. This type of attack often holds an asymmetric advantage in that even relatively older PC and modem technology can still be used to disable faster and more sophisticated machines, networks, or devices. Conversely, the Distributed Denial of Service (DDoS) attack leverages the Internet's distributed architecture and some segment of the millions of interconnected computers that reside on it in the same basic techniques. In the case of the February 2000 attack already mentioned, some of the effected sites were receiving levels of traffic measured in the gigabits per second range.

While the ultimate effect on the target system is the same under DoS and DDoS, a DDoS is considered significantly more serious in that the malicious traffic originates from multiple machines over the wire versus originating from a single machine. Consequently, DDoS attacks are much faster at achieving desired levels of impact against the entire infrastructure as well as being much more difficult to defend against or isolate. Coupled with the accelerated deployment of "always-on", high-bandwidth access technologies such as DSL and cable modem, the frequency of DoS style attacks have already increased as much as 60 percent over the past three years¹⁸.

¹⁷ The world wide web security frequently asked questions, <http://www3.org/security/faq>.

¹⁸ Source: Computer Security Institute, "2001 CSI/FBI Computer Crime and Security Survey", Vol. VII, Spring 2001.

C. DENIAL OF SERVICE IMPLEMENTATION

Malicious hackers generally employ DoS using one of two techniques to assault the underlying TCP/IP or UDP/IP protocol standards that provide the basis for data exchange [Ref. 12]. The first, labeled flooding attacks (ICMP, SYN, SMURF, FRAGGLE (UDP)), seek to quantitatively flood the target system with spurious Internet Protocol (IP) packet traffic with the intent to overload targeted network servers. Legitimate data traffic in route to the target system is not directly prevented. Instead, legitimate packets become the smallest fraction of the total traffic processed. The processing of legitimate traffic becomes relegated to the background and thus service is denied. The second technique, generally referred to as a logic-based attack (PING of Death, Chargen, Teardrop), exploits known software bugs resident on the operating system (OS) of the target system to take it offline, crash, or reboot. Despite the varying methodologies available using these types of techniques, in the larger sense, the primary emphasis to be appreciated is that any communication infrastructure attached to the Internet could be disrupted, even crippled, if not adequately defended. The specific methodologies are summarized to further develop an understanding of the characteristics underlying the attack.

1. SYN-ACK/TCP-SYN Flooding

Briefly, the Transmission Control Protocol (TCP) is the primary transport mechanism used to provide the guarantee of a reliable connection between two hosts. To do so, TCP makes use of a three-way handshake to establish or terminate an active session. To initiate the process, a client machine will transmit a synchronization packet (SYN) to the desired server. In turn, the server will acknowledge that request using a SYN-ACK packet. Then, in an attempt to ensure unambiguous agreement, the client will acknowledge the server's acknowledgement with its own ACK packet. The TCP-SYN flood attack techniques are designed to leverage this handshake process to effect a denial of service condition. By initiating only the initial handshake portion of the session and not responding to the acknowledgment sequence, the SYN-ACK technique forces numerous half-open TCP connections that results in the victim server storing larger numbers of acknowledgement packets in its queue. Eventually, the queue will reach a state of overflow such that its capacity to issue any new, legitimate, acknowledgements

will become disabled. The typical targets of this technique are web servers and traffic load-balancers deployed in concert with firewalls, servers, or IDSs.

2. SMURF

A SMURF attack is a technique that leverages the availability of amplifier machines to multiply the total amount of traffic a target site will receive. Specifically, the attack will take advantage of the direct broadcast addressing mechanism of a network by spoofing the target machine's IP address and broadcasting Internet Control Message Protocol (ICMP) ping requests (ICMP_ECHO_REQUEST) packets across multiple subnets, if configured. Upon receipt, all machines that are connected to the subnet will reply to these requests and forward them to the intended target of the attack. This process will serve to congest or clog the victim's network resources with bogus packets making it unavailable to legitimate traffic. The intelligence in this particular attack profile is that all intermediate systems will contribute to overall congestion, thus magnifying a DoS, as each is drawn into the echo-response cycle. It is believed that as much as half of all DoS attacks flood a victim's site in this manner. SMURF attacks are possible against the full range of infrastructure components to include web servers, individual host machines, and routers.

3. FRAGGLE / UDP

Considered a derivative of the SMURF attack, FRAGGLE is a form of attack that exploits the connectionless delivery characteristics of the User Datagram Protocol (UDP) in that it has no inherent properties of flow control, ordered delivery of packets, or acknowledgements. As opposed to the SMURF attack, the FRAGGLE method will send UDP packets, vice ICMP packets, to the broadcast address using the forged source address of the intended target. A stereotypical attack might involve sending a series of UDP packets to the character generation (chargen) port on one host with the packet's source port set to echo on the same system or on another. After a connection is established between two UDP services on a network, excessive numbers of forged UDP broadcast packets between the two services will serve to consume all available bandwidth at the expense of further legitimate users. Chargen and echo are just two of the available, but considered non-essential, UDP-based diagnostic ports that can be used for debugging or maintenance purposes.

4. Teardrop

IP packet data is normally disassembled or fragmented to gain greater efficiencies in the transmission process across the Internet. Each packet fragment is constructed to be identical to its original IP packet with the exception of an offset value that indicates which “bytes” of the original (whole) packet are included. The Teardrop technique exploits the IP vulnerability as it relates to the reassembly of packets and purposely injects packet fragments with overlapping offset values to impede the sequencing, or reassembly, process in the victim’s machine.

5. PING of Death/ Oversized Packet

This type of attack exploits known bugs in the TCP/IP implementations of some operating systems to cause complete crash, reboots or system hangs. Specifically, this technique uses the ping utility to send packets that exceed the maximum 65,536 bytes of data allowed under the IP specification. While still a capable form of attack against older operating systems (OS), this particular exploit has largely been rendered ineffective through the normal course of OS replacement or software upgrades.

D. DISTRIBUTED DENIAL OF SERVICE IMPLEMENTATION

The fundamental attack techniques underlying a DDoS are the same as described above for a DoS. The use of tools to effect a distributed denial of service, generally, paralleled the transition to the larger bandwidth capability within the Internet itself. Preference for its use can also be viewed as part of the larger escalation response cycle to directly counter the commercial deployments of firewalls, load balancers, and large server farms by subscriber networks. The primary difference between DoS and DDoS, however, hinges on the degree of sophistication as well as the notions of scale, speed, and complicity in the attack model.

A DDoS is generally implemented by a malicious hacker by first identifying a network or number of non-secure host computers by the process of scanning large ranges of network IP blocks looking for specific, but commonly known vulnerabilities such as wu-ftp, RPC services, and others. Once a vulnerable system is identified, each is implanted with either a zombie program or a specific handler (client) program. Both programs are used to facilitate the hacker’s takeover or control of impregnated “third

party” systems that will be used to coordinate attacks against future targets, the intended victim of a DDoS.

Zombie machines, also known as z-bots or slaves, are utilized to directly attack the target machines while the handler/client system is used principally as the intermediate control station. As Figure 5 depicts, the handler machines are used to activate subordinate zombie machines in a decentralized command and control fashion using techniques ranging from simple UDP packet based commands to the more sophisticated TCP (telnet, ssh) and ICMP (ECHO_REPLY) based commands. In most instances, the process to establish and execute the army of z-bot attackers is fully automated such that much of routine work of scanning for vulnerable hosts, installing the daemon (attack program), concealing the intrusion, establishing peering relations are being accomplished through scripts, Trojan Horse, or self replicating worm applications.

Future trends expected in DDoS tools suggest that besides improving the graphical user interfaces, to address ease of use issues for hackers, more subtle approaches are being pursued to create a slower degradation, vice outright denial, of service in an effort to complicate the discovery of attacks by intrusion defense systems at the victim site [Ref 9]. The logic being that since detection technologies generally operate using pre-established thresholds, forming attacks that arrive below such levels could delay or outright subvert automated response mechanisms. Other subtle modifications to current implementations suggest that greater emphasis is also being placed on the stealth qualities of handlers and zombie programs to make them more invisible to signature based scanners. This is evidenced by wide spread use/installation of rootkits as part of the payload that can be used to hide the presence of programs, files, or connections.

In terms of command and control, hackers have also begun to deploy tools that have, in addition to the aforementioned adjustments, placed a primary emphasis on encryption to mask the communication channels between master-handler and handler-zombie programs. Subsequent versions of attack tools are envisioned to “pulse” on a periodic basis in the form of short encrypted bursts to further complicate the monitor and detection efforts.

Because much of the DoS coding is still based on cross pollination between existing publicly-accessible programs or as a result of incremental adjustments to the available coding, the profiles of known hacker toolkits such as Trin00, Tribal Flood Network (TFN), TFN2K and Stacheldraht are detectable by open source scanning tools.

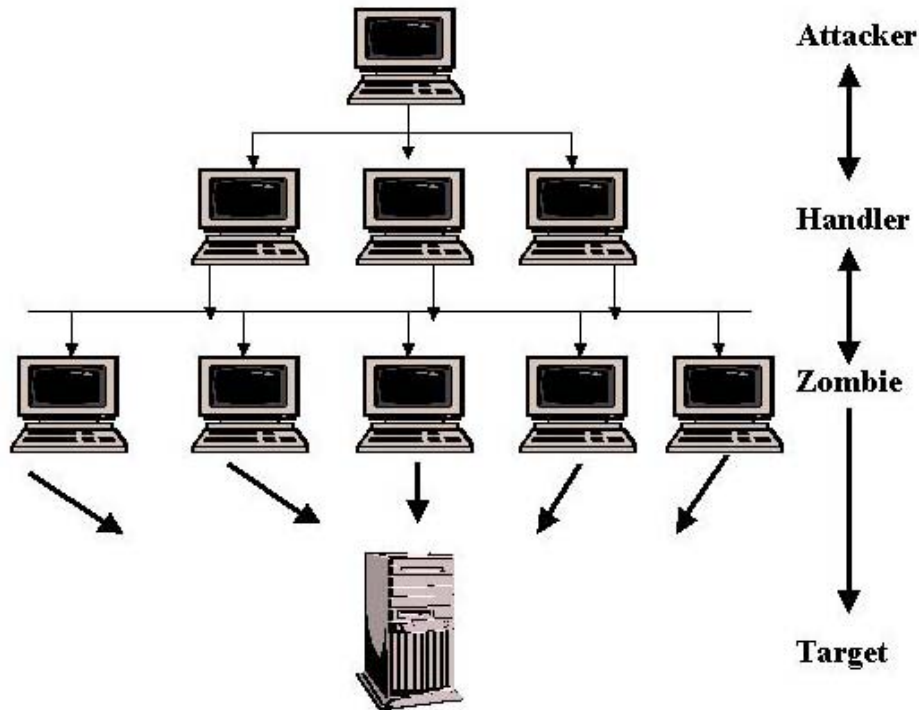


Figure 6. Basic Topology and Communication Path of DDoS [From Ref 12]

E. WORMS AND ASSOCIATED HYBRIDS

Computer systems have long been vulnerable to the threat posed by a family of malicious code known formally as a virus –described here simply as a program designed to spread from one file or computer instruction set to another on the same machine in order to modify or corrupt data. Although capable of wide spread damage in their own right, viruses have an inherent self-limitation in that they require some level of user interactivity to activate and sustain its lifecycle. Alternatively, a distinct subset of malicious code known as a worm poses an even greater threat to a network of computer systems.

A worm's distinct characteristic is that it has the ability to propagate without the help of any user interaction. Designed with a functional emphasis placed on self-replication, worms are engineered simply for the purpose of spreading themselves as widely as possible across a network of machines. While knowledge of their existence or use is not new¹⁹, their true impact has only recently been realized because of the size and ubiquitous interconnectivity of the Internet. In terms of threat to Internet security today, worm developers have enhanced the malignant behavior of their creations by bundling sophisticated distributed attack and remote control tools as part of their payload. While in the process of compromising thousands of vulnerable hosts, attackers are now exploiting the inherent self-replicating properties of worms to spread their tools and then use their victims in amplified attacks in such schemas as distributed password cracking, system vulnerability scanning, and the aforementioned denial of service attacks. This unique combination of replication and payload is what has lead some analyst within the anti-virus vendor community to the conclusion that over the next decade, worms will surpass the virus as the major malicious threat to any computer system or device attached to the Internet.

If not troubling enough for end users, the underpinnings of internet transmission and core routing will be increasingly jeopardized. As in the case of the Code Red incident, transmission capacity became the indirect victim of the worm's rapid propagation in that large allocations of bandwidth were consumed simply due to heavy vulnerability scanning and mass e-mailings. In a preliminary report sponsored by the Renesys Corporation, analysts concluded that there existed a "strong correlation between the propagation of Code Red and Nimda with a period of core internet router instability" most likely resulting from a combination of router CPU overload and the numerous network administrator's incident response activities enacted to adjust to the traffic surge²⁰.

¹⁹ Robert T. Morris is credited with authoring/releasing the first Internet worm as early as 1988 that infected an estimated 6000 computers.

²⁰ J. Crowe, A. Ogielski, B. Premore, Y. Yuan, "Global Routing Instability During Code Red II and Nimda Worm Propagation: Preliminary Report", Renesys Corporation, September 2001.

F. WORM IMPLEMENTATION

By definition, a worm is a form of malicious code that has the ability to propagate without any assistance from the computer user. Within the network environment, worm code is able to sustain propagation by either leveraging the transport capabilities of a communications service like e-mail in which it can attach itself to self-generated outgoing traffic or by making use of some arbitrary protocol such as HTTP, FTP, IRC or TCP/IP to probe and spread to other connected systems. Once received, execution is generally achieved through one of three methodologies engineered within the design of the code. The basic launch methods include: self-launch, user-launched, and hybrid.

1. Self-launch Method

A worm that propagates using the self-launch method exploits some particular vulnerability of the host machine operating system or installed applications to enable the worm to execute automatically once introduced to a new platform. The unique characteristic of this type is that a human user does not need to execute anything. For example, the worm may append itself to an available executable within the operating system that is used during the boot process.

2. User-launch Method

A worm that employs a user-launch method requires a user's intervention or action to trigger execution of a supplanted payload. This particular method usually takes advantage of our susceptibility to social engineering techniques that exploit human curiosity. For example, a user is convinced to open an infected e-mail attachment in hopes of receiving a reward or prize. Similarly, the worm may append its code to all HTML files within a web server directory in an effort to infect visiting web browsers.

3. Hybrid Method.

Hybrid worms use design features of both self and user-launch types. First executed by some user action on one host, the worm then automatically spreads to other systems via a given exploit and becomes activated through a variety of means like an anniversary date or a system reboot.

G. WORM ATTACK MODEL

In the general sense, the propagation mechanisms simply automate the network attack model, reflected in Figure 7, which has been used by human intruders for years to exploit computer technology. For example, the post incident analysis of the Nimda Worm, authored by the SANS Institute, illustrated four distinct modes of propagation to infect hosts running any version of the Windows operating system [Ref 12].

First, the Nimda Worm conducted automated scans of the Internet IP domain space to locate vulnerable web servers. Then, employing any of approximately one hundred known system exploits, such as an IIS transversal technique or backdoor program left resident from a previous worm infection, Nimda remotely downloaded a copy of its program from a previously compromised host. Second, having the capability to extract listings from a user's address book, the worm was able to self-generate mass e-mailings with an infected attachment in a manner that harnessed the power of social engineering to takeover where technical means fell short. Third, if successfully installed on a web server, the worm leveraged weaknesses within the HTTP service protocol by implanting visiting web-browsers with an infected executable program that ran automatically within certain versions of the popular Internet Explorer web browser²¹. As in the case of e-mail exploits, hijacking the legitimate protocols like HTTP provides the worm an additional route through firewall perimeters. Fourth, the worm would search out and append itself to any executable file within an accessible network share drive on the compromised host to lay in wait to be accessed by another host at some future date.

²¹ Security Focus virus news bulletin, "Toward More CyberSecurity in 2002" by Alex Salkever, Jan 2002, reported that in addition to 90% of the world's PC running Microsoft Windows more than 80% used Internet Explorer.

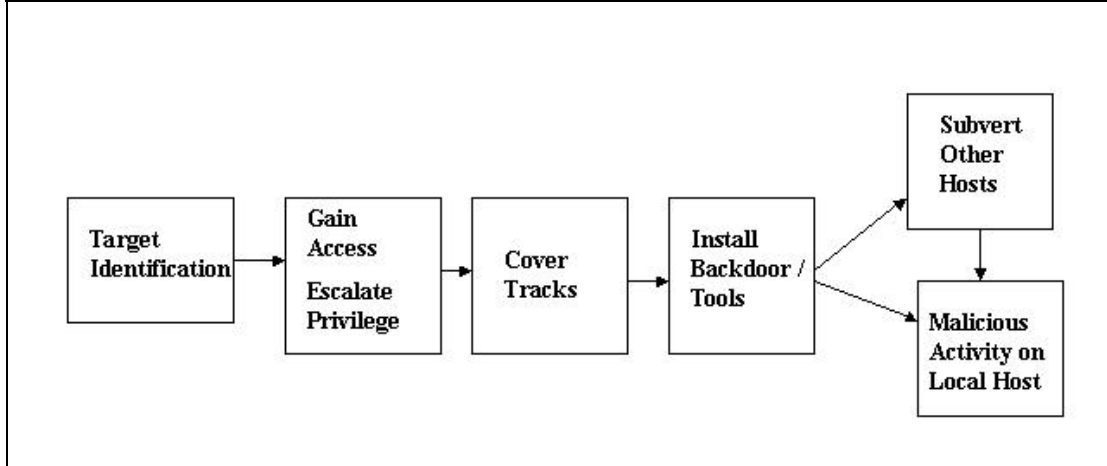


Figure 7. Network Attack Model

While Nimda is but one of many recently identified Internet-based worms, detailed here to provide granularity, the various propagation techniques (or combination thereof) consistently appear within the current generation of worm releases. The distinguishing characteristic of a particular worm, however, is a function of its author's choice to exploit one vulnerability over another, previously identified or otherwise.

In the larger context of Internet security, worm technology exposes several contributing factors that have exacerbated the weakness of relying solely on the subscriber-centric security approach. First, for a myriad of reasons that range from apathy to overload, administrators of both large and small networks have failed to keep pace with system updates despite the availability of software patches to eliminate known vulnerabilities and adequate foreknowledge of existing threats²². This was particularly evident in the mass infection rates experienced in the aftermath of Code Red II, approximately twenty-nine days after patches were made publicly available following the Code Red I worm release. Second, the ubiquitous nature of e-mail as a propagation mechanism has dramatically increased the rate of infection. Levels of exposure that previously took months or weeks to propagate are now possible in minutes and seconds. Once an e-mail account is exposed, individuals continually demonstrate their

²² ComputerWorld Magazine, "Study: Constant Security Fixes Overwhelming IT Managers" by Dan Verton, November 30, 2001.

susceptibility to even the most basic social engineering techniques that serve to perpetuate the infection cycle. Third, worm technology, coupled with an abundance of freely available attack code, has dramatically reduced the period in which the anti-virus signature update process can effectively respond. This is exacerbated by the prevailing desktop-centric approach.

G. SUMMARY

Threats to computer systems can come from many sources. While much of the focus of security efforts has been on addressing issues of insider abuse, there presently exist the means to leverage the interconnectivity and transmission capacity against other networks as well as the Internet infrastructure itself. Two of the more serious threats of distributed denial of service and self-propagating worm were presented with some detail to provide the reader with an understanding of the attack methodologies and to illustrate the increasing ineffectiveness of the subscriber-centric security approach as the primary means to achieve defense in depth.

While the techniques were discussed in the context of current capabilities (known attacks) that remain effective today, the identification of trends and projected capabilities highlight that effectively countering future attacks will become increasingly more difficult should the onus remain on the edge networks and individual subscribers.

Due to the highly interactive nature of the various systems and networks, it is recognized that no single system can be made adequately secure unless all the interconnected systems are made secure. Given that a substantial number of edge systems remain vulnerable for a variety of reasons, it is necessary for defensive mechanisms, such as those identified in Chapter II, to be implemented deeper within the infrastructure to isolate compromised systems attempting to generate malicious traffic. If released into the Internet, the ability of even the more security hardened systems to counter a combined attack by multiple end systems is greatly reduced or negated thoroughly. Within the context of the denial of service and worm capabilities identified here, Chapter IV will describe some technologies and tools that are currently available, but not widely used within the ISP community, that can strengthen the existing layered defense security model.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. ISP CENTRIC APPROACH

A. INTRODUCTION

As mentioned, one of the more critical aspects of assuring the requisite security for the Internet is a need to identify the appropriate entities that are in the best position to translate security concepts into action across the distributed infrastructure. Historically, the onus for that security has fallen singly on the shoulders of the participating subscribers or edge networks in a manner that best meets their own operating requirements. Given that initial use of the media was voluntary in nature, the notion of acceptable behavior or a self-imposed etiquette by participants was sufficient to mitigate most malicious activity.

Today, participation within the Internet is no longer considered voluntary and in many cases deemed mission critical to the needs of commerce, society, and governments at large. At the same time, as Chapter III indicated, new levels of ubiquitous interconnectivity, coupled with an abundance of system vulnerabilities inherent in computing products, now makes it possible for the growing population of malicious users to leverage millions of connected hosts against any element in a manner that jeopardizes the Internet's overall reliability and continued availability. As previous high profile incidents have demonstrated, the subscriber-centric security paradigm has been unable to effectively respond to the distributed nature of the DoS and Worm threats despite our vast knowledge base of vulnerabilities and customer premise countermeasures. The limiting factor to the impact such threats have had to date has been more a result of poor execution on the part of less capable adversaries rather than a triumph of vigilant network administrators.

Commensurate with the Internet's evolution, however, Internet Service Providers (ISP) have emerged as a dominant influence in expanding the infrastructure as well as services available to its end users. Acting as the central gateway for all but the smallest fraction of internet users, the Internet Service Providers command a functional advantage, by way of their access, core distribution networks, and peering/service level agreements, to exert a level of measured restraint over its subscribers in an effort to reestablish the

boundaries of acceptable use, most of which are already identified as best practices. More importantly, ISPs retain a unique positional advantage to globally enforce the mechanics of appropriate use that no single participating network can exact on any other network segment.

Given the fact that there presently exists no single technology to guarantee absolute security for every computing situation, the widely accepted and necessary concept of a layered defense has evolved within the network security environment to help mitigate the effects of malicious activity. Most of these mechanisms, summarized in Chapter II, that comprise the layered defense are inconsistently applied across networks or often misconfigured and/or unpatched to some degree. While most of the technologies to be suggested within the ISP centric security approach implement the same principles and security mechanism in use today (firewalls, anti-virus, intrusion detection), the fundamental difference is that the advances in technology have now made it possible to employ the techniques at the internet-work level sufficient to match the available rates of transmission. Coupled with a unifying organizational context emerging from the managed service provider business model, an ISP-centric security approach offers the benefit of mitigating the threat's advantage of scale once released into the infrastructure. This section will discuss four enabling technologies that provide ISPs with the internet-work level tools to better mitigate the effects of denial of service and geometric growth attacks, most of which have recently been made available.

B. THE SECURITY POLICY ENFORCEMENT POINT

Although the layered defense approach is universally accepted as the means to achieve the highest degree of protection in today's environment, all subscribers do not consistently apply the necessary components to achieve the requisite levels. In essence, all organizations implement a security posture based on their unique risk management analysis that often, in terms of Internet security, fails to strengthen the "system" as a whole. This has been most apparent in the case of individual subscribers who have begun to migrate from traditional dial-up service to the higher capacity broadband service such as DSL and cable. With traditional dial-up access, subscribers generally exhibit a low profile in terms of their Internet presence. When initiating a connection, users are

assigned a random IP address from a large address pool provided by their ISP and their connections are typically short in duration. Somewhat analogous to a small fish in a big pond, dial-up users could assume they were less at risk and avoid using a common defense mechanism such as a desktop firewall. Alternatively, broadband access is viewed as a higher profile in that, generally, users are assigned a fixed IP address and operate using a high bandwidth, always on connection. Since the vast majority of ISPs do not require subscribers to deploy any of the perimeter defenses as a term of service, hackers can easily detect their presence and have a much longer opportunity to exploit potential vulnerabilities.

As part of the movement to generate business revenue, ISPs have been leveraging newer technologies to deliver a greater number of valued added services to subscribers. Virtual Private Network (VPN) service is one such example that provides a secure communication path between two end points over the existing public (non-secure) Internet infrastructure. One specific class of technology, however, that has emerged is a service provisioning platform or security policy enforcement point (SPEP) designed with the intention to provide a more diverse security feature set and IP services from the point of network access. Two such platforms are the Nortel Networks Shasta 5000 Broadband Service Node (BSN) and the Lucent Technologies SpringTide 7000 that can be deployed across the range of service provider networks, commercial enterprise networks, and university settings. While these two platforms are vendor specific products, both demonstrate similar offerings that will be singularly highlighted within a context of mitigating DoS and worm propagation to illustrate the application of the ISP-centric security approach. As shown in Figure 8, the SPEP is deployed at the service providers network edge. It was designed with a functional emphasis to provide a centralized access point that can aggregate all the various access technologies typically found within the ISPs access network; DSL, cable, dial-up, ATM, Frame Relay, and wireless. The principle relevance in the ISP-centric security environment is that the SPEP enables the ISP to apply a range of service (security enhancing) policies to a large number of individual subscribers through a single interface. In terms of scaling that supports the tiered infrastructure of the ISP community, the SPEP is also designed to allow a multiple of other ISPs to use the same node yet retain separate address space to distribute the

service policies to their client base. This is made possible by the creation of what Nortel calls a “context or virtual router” scheme. Figure 8 illustrates the concept best. ISP A is considered the owner of the SPEP and owns the underlying backbone infrastructure leading to the Internet. ISP B, potentially a tier 3 “reseller”, is provided a distinct context from which security services can be distributed to its respective clients. In terms of security, the specific feature sets of interest that will be outlined include: firewall; anti-spoof filtering; Network Address Translation; Quality of Service (QoS) as a traffic management tool; and, to a limited extent, web steering [Ref 14].

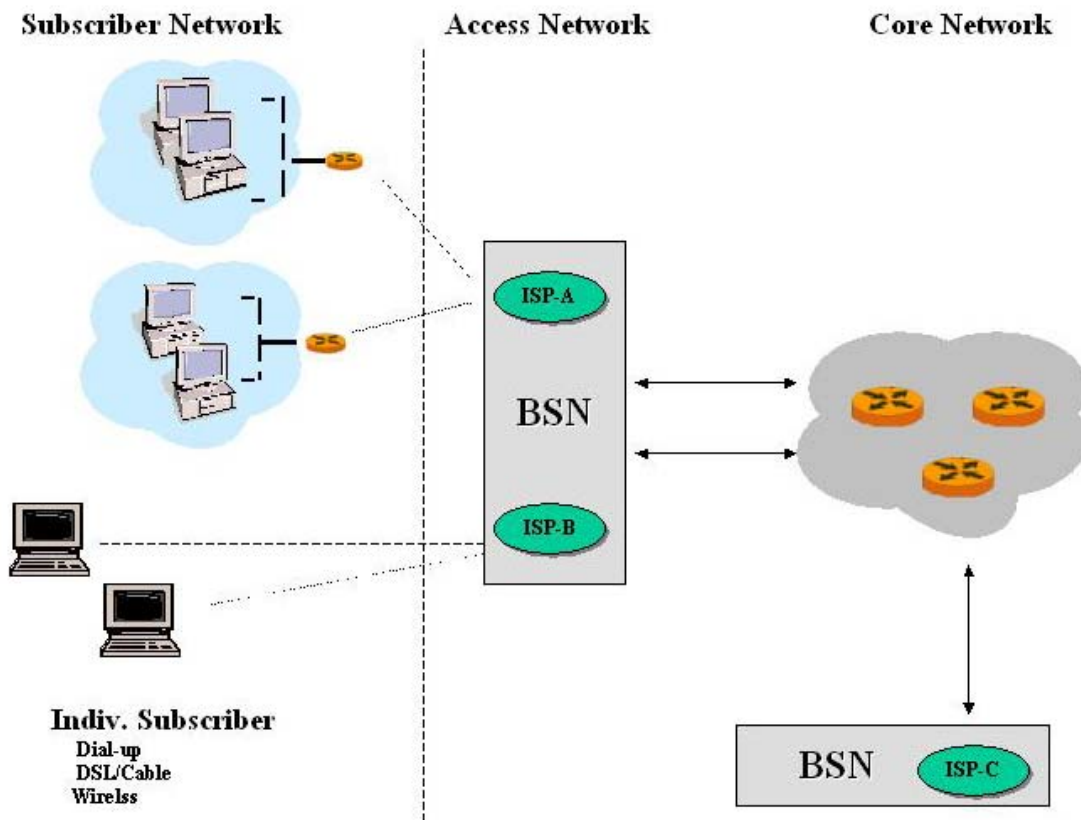


Figure 8. Nortel Networks Shasta 5000 BSN Deployment Architecture

1. Firewall

With a level of specificity that accommodates tens of thousands of individual subscribers per node, the SPEP provides a state-aware firewall technology at the point of access into the ISP backbone network. Deployed as a network-based firewall, vice a customer premise device, the ISP can employ all the commonly used principles of packet filtering and session level management described in the Chapter II to enforce security policy and block malicious traffic before it can be released into either the ISP core network or outbound into the subscriber network. In terms of utility and management, this capability allows an ISP to create policy statements customized to individual subscribers or groupings of subscribers using preformatted templates. Each firewall instance (per subscriber) is able to track the bi-directional application flows that use both static TCP and UDP ports as in the case of HTTP or protocols that use dynamic ports such as streaming media and FTP. This would allow for the filtering of unsolicited packets into the subscriber's network/host. More broadly, the advantage of this approach is that it allows ISP to enforce aspects of "best practices" that may have otherwise been inconsistently applied within the subscriber-centric model. As opposed to the situation in which the subscriber base deploys a variety of vendor specific firewalls, the SPEP allows ISPs to provide higher levels of customer support using the managed security service paradigm.

2. Anti-spoofing (Ingress/Egress Filtering)

Despite the presence of firewalls, hackers have always managed to exploit some known or unknown system vulnerabilities to gain access to computer resources within networks. Once penetrated, spoofing attacks are commonly used to generate legitimate or illegitimate source-addressed traffic that can bypass the perimeter firewall packet-checking processes. In this regard, the SPEP incorporates features to create subscriber specific anti-spoof policy profiles that can filter both inbound and outbound traffic to block source address spoof attacks. While the techniques of egress and ingress filtering are common defense mechanisms, described more completely in Chapter II, they have traditionally been deployed at the subscriber's edge router only. If a number of hosts within a particular network have been infected with a zombie program, the edge router represents a single point of vulnerability that must deal with the aggregated flood traffic.

By pushing the anti-spoofing and filtering policy closer to the individual subscriber level, the SPEP approach prevents an individual host from releasing forged source traffic. Having greater top-sight of the traffic that does enter the infrastructure, the SPEP concept contributes favorably to the ability to trace and isolate the source of malicious activity.

3. Network Address Translation

In general, Network Address Translation (NAT) is a mechanism to translate internal network IP addresses into a single unique IP address. It was designed with the intent to more efficiently manage the dwindling IP address space pool inherent under the IPv4 specification. However, NAT also has an additive value in terms of security in that it is an effective means to hide the internal (private) IP addresses from the public internet field of view. Its principle security strength is that no communication from the outside is allowed to reach into the network unless it is first initiated by the internal address. To capture this unique benefit, the BSN and Springtime 7000 are designed to internally support NAT down to the individual subscriber IP address. Migrating this functionality to the SPEP, from the subscriber's network, allows the level of "reachability" necessary for subscriber specific security profiles to be effective.

4. Quality of Service (QoS)

The Internet works within a basic operating concept described as "best-effort" delivery to reliably distribute packetized data between two end points. During this process, however, no commitment is made on the part of the infrastructure to guarantee appropriate levels of bandwidth capacity or timeliness of delivery (latency) for the transmission. In order to better support real-time network services, QoS emerged as mechanism to manage the prioritization and efficient scheduling of traffic to control congestion and delay. For example, streaming/multi-media applications such as live audio intuitively requires an orderly sequence of arrival and timeliness in delivery to provide the user with an intelligible message that mirrors the originators oration. Should portions of the traffic arrive out of order or excessively late, the message would sound fragmented and incoherent by the receiver.

As a by-product of this scheduling and prioritization function, QoS mechanisms can have a security enhancing character especially in terms of mitigating denial of service and worm propagation. For instance, a common attribute of DoS is one of bandwidth

consumption resulting from packet floods initiated by the zombie network. QoS can be used as a traffic management process to restrict the amount of traffic (shaping) that any one zombie can broadcast should some predefined threshold be exceeded. Similarly, the traffic can be tagged as low priority (traffic policing) to be dropped further downstream in favor of preferred or legitimate data packets.

Accordingly, the SPEP concept incorporates the mechanism to employ and enforce QoS policies down to the individual subscriber. Using class-based traffic policing to optimize the backbone network, ISPs can define the types, rate, and volume of type traffic an individual subscriber can introduce into the core infrastructure that has been pre-established by way of a contractual Service Level Agreement. In particular, the BSN supports policing categories based on committed rate, committed burst size, peak rate, and peak rate burst size thresholds within the four Differential Service Assured Forwarding (DiffServ AF) classes, used as a means to allocate forwarding resources such as buffer size and bandwidth of a DiffServ domain node. [Ref 15] For example, a subscriber's HTTP server traffic may be assigned as an AF-4 class and a committed rate of 10 kbps. This characterization could indicate the HTTP traffic is afforded a guaranteed 20 percent of available bandwidth for the contracted committed rate. Exceeding any predefined threshold, the traffic can be dropped outright, queued using selectable weight values, or marked for prioritization. Using a worm infection as an example, HTTP traffic associated with its propagation phase would likely exceed the committed rate and be marked for a drop action before release into Internet, thus minimizing the impact upon bandwidth availability until further incident response procedures could be taken.

In terms of traffic shaping, the SPEP can be used to similarly enforce a customized policy profile that manages traffic sent into the subscriber's network/host to provide rate "guarantees" or traffic priority specific to individual application flows. Consider the case of the worm using e-mail as the propagation mechanism, using traffic shaping to lower the priority of less mission critical applications like SMTP could be helpful in dampening the indirect resource consumption that results from the mass mailing until further action could be taken.

5. Web Steering

Web steering is a functional mechanism that allows an ISP or enterprise network to dynamically redirect a client's HTTP request for the purposes of funneling users to specific content. Instead of connecting to an intended web page, the session request is diverted to a preferred proxy cache web server. This process can be used for a variety of reasons, but some of the more common examples include the network authority steering employees to the host company's homepage upon initial login, an ISP directing customers to a service selection page that can be used to request a new value-added service, or a web host organization directing clients to a periodic customer service survey. Nortel's BSN also provides for this capability, defined as their Personal Content Portal. While this feature has some explicit business related applications, the ISP community can leverage this same functionality to reduce the spread of malicious code that is often embedded within subverted web pages. For example, part of the Nimda propagation profile was to infect the HTML files on a compromised web server. Similarly, hackers routinely embed Trojan Horse programs within subverted web pages. In most cases, incident response processes are slow in making repairs to compromised servers before large numbers of unsuspecting users visit their sites. As national and private organizations begin to track and maintain databases of infections close to real-time, it is conceivable in an ISP-centric model to push compiled listings of compromised servers to service providers in a manner that they can rapidly invoke HTTP request filters and automatically steer subscribers away from dangerous areas. Web steering could be used somewhat analogous to the emergency broadcast system over television and radio, where programming across all channels is interrupted and viewers are alerted or directed to safe harbor during times of crisis.

C. MITIGATING E-MAIL AS A PROPAGATION MECHANISM

Electronic mail (e-mail) has effectively become the indispensable communication mechanism in use over the Internet. E-mail is often cited as the number one reason subscribers elect to join the internet community. In an effort to subvert many of the static perimeter defense mechanisms used throughout networks, however, worm/virus

developers have increasingly leveraged e-mail and their attachments as one of the more effective propagation mechanisms to distribute malicious code.

While many ISPs host mail services for their subscribers, the bulk of the infrastructure still resides within the customer networks. This architectural approach was reasonable during the Internet's beginning, given that most networks developed internally before attaching to the global community. Further, this approach also lends itself to providing edge networks with the highest degree of autonomy to configure services commensurate with their own requirements. To guard their systems from the outside, most universities, small business, and enterprise networks deploy firewalls and e-mail servers within a "Demilitarized Zone" similar to that shown in Figure 9, to protect their internal networks, yet permit internal users to send and receive e-mail. Virus scanning can be accomplished at either the Internet facing firewall or within the mail server depending on a given architecture. The advantage of these techniques is that it provides scalability for the network in that multiple firewalls, servers, and scanning mechanisms can be added to manage network traffic loads for improved performance.

The disadvantages to this implementation, in terms of global security are three fold. From the anti-virus management perspective, the sheer volume of networks that are supported by virus product vendors creates the situation in which response time to deliver frequent signature updates provides sufficient windows of opportunity in which new viruses can be introduced into the system or existing viruses can achieve large infection rates. This situation is further exacerbated in the home subscriber market as anti-virus vendors attempt to keep pace at the individual desktop level. While vendors have responded by increasing the level of automation to push/pull updates, as in the case of Symantec Corporation's LiveUpdate feature, subscribers retain the discretion of enabling this functionality or establishing the periodicity of updates. If inconsistently applied across even a small percentage of attached networks/desktops, e-mail propagation retains the advantage of exploiting large portions of the Internet as a whole.

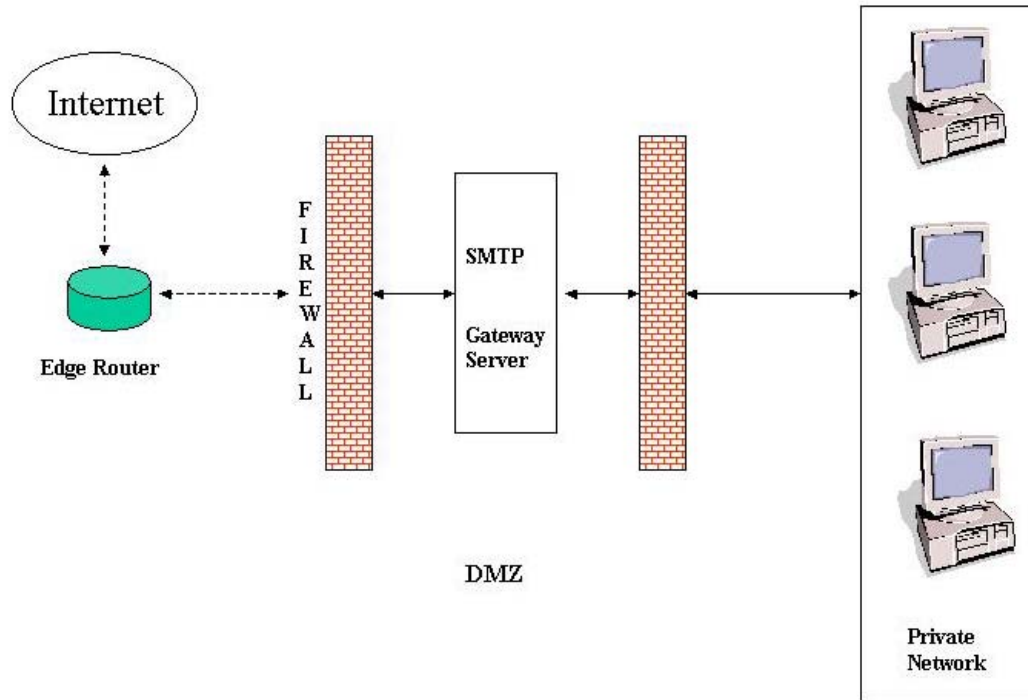


Figure 9. Firewall Deployment within a Demilitarized Zone (DMZ)

From the network-level perspective, anti-virus scanning technology can be employed at the firewall, the server proxy/gateway, or both. For networks that employ scanning at the firewall, all inbound and outbound data are intercepted and scanned, or blocked outright. The advantage of this scheme is that no additional server assets are required to implement protection and responsibility to manage each network client is alleviated. The disadvantage is that the firewall can quickly become a single point of vulnerability and bottleneck in terms of network performance. To reduce loading, the common practice becomes to relax policy constraints regarding outbound traffic that is assumed to be trusted while concentrating resources against untrustworthy inbound traffic. Again, given that exposure to viruses/worms can be achieved by a multitude of other methods, the Internet becomes vulnerable to large outbreaks of malicious code. In the case where Intelligent Scanning Architecture (ISA)²³ is employed (see Figure 10) within medium and larger sized networks to mitigate the performance penalties of the firewall implementation, common practice is also to relax policy regarding outbound

²³ A variation of scanning that incorporates a process that allows the firewall to offload suspicious traffic to a separate a proxy virus scanning sever while letting “legitimate” traffic to proceed.

traffic flow. Further, the deployment of additional components within networks again places increased burden of scale as anti-virus vendors attempt to distribute updates.

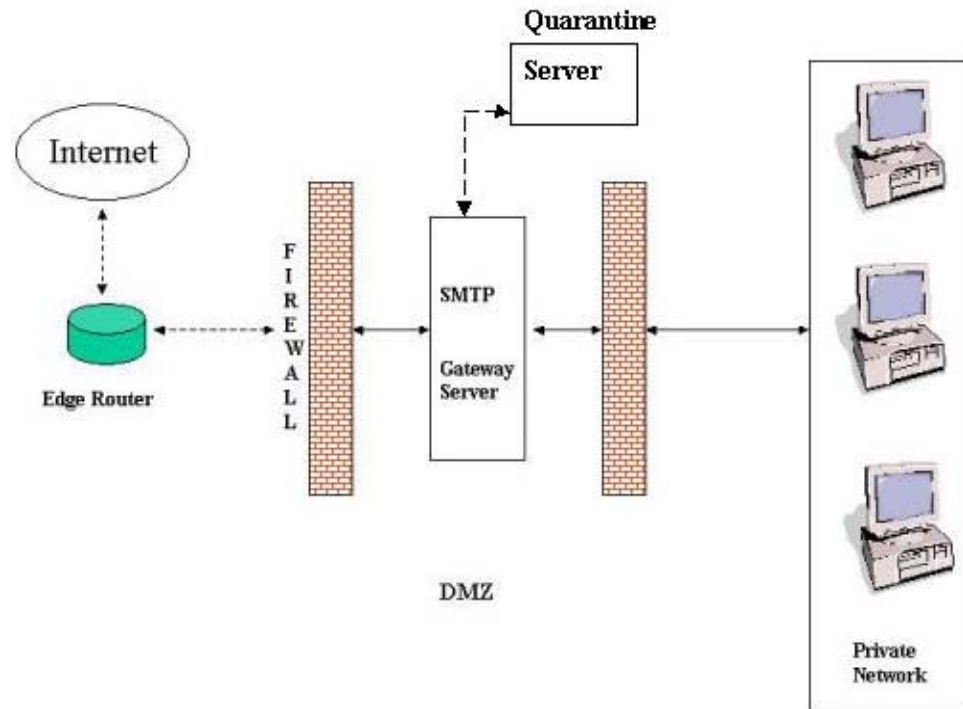


Figure 10. Representative Intelligent Scanning Architecture

Alternatively, maturation of virus scanning technology now makes it possible for an ISP to deploy sufficient processing capability to handle the volume of e-mail traffic being generated before viruses can reach subscriber systems [Ref 16]. Two examples include EarthLink and AT&T WorldNet. In the latter case, AT&T has recently deployed the BrightMail Anti-Virus/Spam solution suite that is designed to protect the integrity and security of e-mail systems and individual accounts deeper within the Internet's distribution network²⁴. Fundamentally, this technology works similar to other anti-virus scanning products. It employs both signature-based filters and heuristic-scanning techniques against all transmitted SMTP traffic, attachments, and various encoding schemes (MIME, UUENCODE, BinHex) for possible viruses. If detected, the entire message can be temporarily stored and disinfected before forwarding to its original

²⁴ The BrightMail solution suite is also marketed for enterprise networks. According to BrightMail's press release dated August 7, 2001, AT&T is cited as one of the first Tier-1 ISPs to deploy this concept.

destination. Should other forms of malicious code, like a known Trojan Horse program, be detected, the attachment is deleted and the intended recipient notified. Although such systems provide a measure of protection against known "signatures", they cannot offer a simple and complete solution to the malicious code problem. The only known solutions for which there is scientific basis are systems that provide confidence of non-subversion and enforcement of protection domains. Lack of these systems force networks to be based upon the weaker security technologies discussed here. However, if e-mail scanning is deployed within the ISP distribution network, this approach can bolster overall internet security and offers the following advantages over subscribe network deployments.

1. Early Warning

As in the BrightMail example, an array of dedicated e-mail accounts disguised as regular accounts are placed across all the participating networks to act as early warning devices. The logic being similar to a deployed sensor network, these "probe accounts" are established to detect the early beginnings of a virus' propagation to allow backend support to analyze and distribute appropriate countermeasures before mass infection can occur. BrightMail is jointly supported by Symantec to provide the virus definitions and scanning engine updates. Combined, the two organizations cooperate to provide the ISP with round the clock analysis of suspicious traffic and to deliver customized rule sets for disabling the attack.

The integration of a backend support infrastructure becomes beneficial to the network as a whole in addressing newer generations of malicious code, typified by the emergence of polymorphic techniques. A virus spawned by a polymorphic worm may appear as an isolated instance to any one ISP edge-network administrator. The cooperative pairings between an ISP and an anti-virus vendor allows the ISP to take advantage of the expertise of a specialized organization that potentially has a view of activities across ISP boundaries. Relieved of the burden to coordinate updates to the hundred of thousands of end systems, the vendor response is reduced to a more manageable level. As the backend support is also a repository for infected traffic, the code can be made available for broader forensic evaluation to aide in post-incident investigation and legal prosecution. The importance of this capability cannot be understated given the expanded scope of recent cyber-related legislation (PATRIOT Act,

2001) that has created a closer binding of computer crime and national defense. In contrast, anti-virus processes implemented strictly within subscriber networks tend to neglect this component of security and are generally excluded from the backend support infrastructure.

2. Economy of Scale

Deploying e-mail scanning technology along with the organizational focus within the ISP network serves to reduce the overall complexity of subscriber networks that must increasingly add more devices to match similar capabilities. Further, this approach lends itself to greater scalability in terms of security. For small additions to the server infrastructure, ISPs can distribute protection for a larger number of networks or individual subscribers. Obvious issues related to "single point of failure" can be addressed through availability and reliability mechanisms previously mentioned (load-balancing, advanced switching, etc.). As virus scanning mechanisms become more centralized, the responsiveness of supporting vendors to expedite timely upgrades is improved while mitigating vulnerabilities resulting from improper configuration or less than perfect interoperability between devices. Recognizing that increased centralization will potentially bring with it an increased focus by hackers, continued efforts to deploy intrinsically secure "mail systems" would be even more critical.

D. DENIAL OF SERVICE APPLIANCE

The internet environment is too complex to assume that all ISPs would uniformly opt for the solutions offered thus far, nor would it be consistent with the principle of layered defense to ensure adequate redundancy. Accordingly, the emergence of functional specific DoS appliances provides the ISP centric approach the ability to extend security deeper into the core infrastructure or the subscribers network. The DoS appliance is a unique technology that provides the monitoring, analysis, and response mechanism to enforce security within the network of core routers. From the ISP perspective, the availability of infrastructure add-on devices like that of Arbor Networks PeakFlow and Asta Networks' Vantage System provide a means to conduct an additional level of network-wide analysis and DoS mitigation.

Both systems are comprised of a collection of sensors that are deployable alongside key routers to monitor the flow of possible DoS signature traffic. Using anomaly and signature-based scanning algorithms, these particular DoS appliances sample the traffic passing through a selected router rather than sit "in-line" between two routers. Once suspicious traffic is detected, the analysis is conducted to determine source and type of attack. Passed relevant parameters, the enforcement mechanism within the appliance is capable of implementing a suitable policy filter recommendation to its associated router to block the DoS traffic, given the consent of the network administrator. Constraining the ability of the DoS appliance to deploy filters without an administrator's consent, however, is more a management issue rather than a technological barrier. The added value of these devices is that they provide redundant and distributed support to filter traffic as close to the flooding source as possible.

E. DISTRIBUTED FIREWALL

The SPEP concept does provide a means to better compartmentalize the internet user space through the implementation of network-based security profiles. However, the advantages provided by the BSN or Springtime 7000 can be eroded given unique subscriber network topologies. NAT, deployed externally to the BSN, was one such example that could hamper the "reachability" of the BSN to deliver subscriber level protection given that its function is to hide internal IP addresses behind a single interface. To address this potential shortfall in protecting hidden hosts deeper within the defense perimeter, it is prudent for the ISP-centric solution to incorporate the concept of distributed firewall technologies for at least the medium to large size networks [Ref 17].

One approach to deliver on the concept is the 3Com Embedded Firewall Architecture, provided by 3Com Corporation [Ref 18]. 3Com's solution is comprised of a firewall functionality embedded within a network interface card (NIC) that acts as the security policy enforcement mechanism. The individual NIC's then interact with the Firewall Policy Server that centrally manages and distributes security policy within the architecture.

The functional emphasis of the distributed firewall concept is that all security policy defined by an off board, central location is enforced locally at the individual hosts.

In the 3Com product, the enforcement mechanism is embedded within the NIC (hardware) to raise the level of tamper resistance and to ensure that the security mechanism remains “non-bypassable”. Citing the inherent vulnerability of software based solutions that result from a codependent relationship between host operating system and security application, the 3Com design stresses an operating system independent approach. Further, by embedding the firewall on the NIC, the 3Com approach serves to ensure that the enforcement mechanism always remains available between the host and network. In terms of scale, a central policy server simplifies the practical burdens of policy management of the distributed architecture to push out both rule-based and role-based security policy. Unlike rule-based policies, role-based security policy is based on the rights and duties of a particular person/position within an organization.

The 3Com Embedded Firewall Architecture is marketed as a subscriber-centric network security mechanism. However, its inherent centralized management capability does afford its application within the managed security service model. Because the BSN solution is outwardly focused to the subscriber networks, a similar implementation could be pursued by ISPs to protect their own internal networks to limit the potential of their hosts within the distribution network from being subverted and used to coordinate attacks against customer network resources. .

F. SUMMARY

In spite of the preparation and protective measures taken now at the subscriber's edge of the network, the problem of mitigating distributed denial of service and worm propagation requires a long-term centralized approach to ensure a higher level of uniform application. The Internet by design has created a level of ubiquitous interconnectivity and interdependency among computers that has exceeded the limits of a decentralized subscriber-centric security approach. By virtue of the Internet's maturation as a system, security of each node attached depends on the security of all other nodes. The distributed nature of the threat clearly demonstrates the responsibility of computer users. The onus for individual subscribers and networks to remain engaged in terms of best practices is prudent. But more importantly, the responsibility of the Internet Service Provider to

uniformly implement security deeper within the core infrastructure is demonstrated to be technologically viable. Possessing both a positional and functional advantage as the central gateway to the Internet, ISPs have the ability to deploy and manage effective countermeasures for the escalating DoS and worm threats. Three infrastructure level technologies were introduced to illustrate the concept of an ISP-centric security approach that at some levels is already underway: the security policy enforcement point or service node concept to deliver customized IP services to edge networks; a centralized e-mail scanning architecture to address malicious code propagation before subscriber infection; and the deployment of DoS appliance technology within router networks to create communities of "good neighbors".

The distributed nature of the attack technology, coupled with the high speed of propagation, demonstrates the requirement for distributed solutions that can be implemented at the internet level. Until security becomes intrinsic to the design, development, and deployment of all network-computing processes, equipment, etc., intermediate systems must be introduced to ensure the requisite levels of availability and reliability. All of the technologies discussed are currently available to ISPs today.

IV. CONCLUSIONS AND RECOMMENDATIONS

A. CONCLUSIONS

One of the more critical issues facing the Internet is the increasing ease with which malicious users can disrupt the network's availability and overall reliability. Distributed denial of service attacks and self-propagating worms were examined in this study to illustrate their distributed approach to rapidly deliver attacks from multiple points to overwhelm conventional countermeasures. This condition occurs in large part because most, if not all, detection and mitigation processes take place at the subscriber's edge of the network instead of farther upstream. Even if the targeted network applies the correct filters at its perimeters, the Internet pipeline is full of illegitimate traffic, which prevents legitimate traffic from accessing resources. Many of the techniques that prove effective in combating availability attacks are inconsistently applied across all subscriber networks. In many respects, this inconsistency can be attributed to an overall change in the sophistication of the user base unable, or ill-equipped, to deal with the maintenance of insecure computing designs.

In response to the increasingly disruptive activity, ISPs have been called upon to voluntarily interpose additional controls within their networks to help mitigate the impact of availability attacks. Various types of filtering, intrusion scanning, load distribution, and quality of service mechanisms were identified to characterize the current mechanisms deployed by ISP's. While effective, these measures have also been inconsistently applied for a variety of reasons to include training, performance degradation, and a general reluctance to expend resources beyond the traditional context of interconnectivity brokers. As more mission-critical networks become the targets of sustained attack, it is generally assumed that ISPs will eventually face growing levels of exposure under liability and legislative pressure to embrace a centralized security approach. Given these conditions, this thesis discusses the concept of ISP-centric security to enhance the notion of security "defense in depth".

The ISP community was shown to be essentially a three-tiered hierarchical system that continually seeks greater degrees of interconnectivity with other ISPs to leverage

access between providers. Beyond the technical implementations, interconnectivity is achieved through the mutually beneficial peering agreements that serve the purpose of characterizing the nature of the data exchange between providers. ISPs and other independent provider organizations have also expanded their scope of services to emphasize managed security (MSSP). While still in the initial stages of development, the MSSP concept demonstrates some of the functional advantages of outsourcing levels of security responsibility to the ISP as well as provides insight to the economic feasibility of a centralized approach.

Four enabling technologies were discussed that can serve as the framework to better mitigate the threat by providing the ISP with the capability to uniformly enforce the recognized tenants of best practices and traffic management.

The security policy enforcement point (SPEP) concept was identified as the centerpiece of the ISP-centric approach that acts as an enforcement device capable of imparting specific firewall, anti-spoofing, and network usage restrictions to individual subscribers before their traffic enters the Internet. The IP service node distributes the mitigation processes to all key access points under the governance of the ISP to achieve levels of security support unachievable today. At a minimum it greatly simplifies the tracking and tracing functions required to identify compromised hosts. The service node web-steering functionality was also illustrated to suggest a mechanism capable of minimizing the impact of compromised web-hosts to unsuspecting browsers.

The uniform application of centralized e-mail scanning was proposed to help counter the trend toward e-mail capable virus attacks made increasingly more disruptive since being incorporated as a principle worm propagation mechanism. By absorbing e-mail scanning at the ISP level, one of the more prevalent propagation mechanisms can be thwarted outright or at a minimum slowed to allow more stringent measures to be implemented by the subscriber networks. In cooperation with the backend support infrastructure, the ISP approach can help to increase response efforts of the anti-virus community to distribute relevant updates currently burdened by large numbers of subscriber networks and individual hosts.

The Denial of Service Appliance was highlighted as a means to mitigate levels of attack traffic deeper within the service provider's core network. Consistent with the layered defense paradigm, the addition of DoS appliances that can be tuned for specific attack signatures and profiles allows ISP's to monitor and respond to traffic that leaks through other perimeter mechanisms.

The use of distributed firewalls was proposed for two purposes. The service node and DoS appliance are employed as outwardly looking mechanisms to better control the subscriber's traffic. It is prudent to deploy the distributed firewall within the extensive service provider networks to ensure their systems are equally constrained from participating in denial of service attacks should they become subverted. Additionally, the ability of the service node to reach all hosts within a network may be undermined given unique subscriber topologies. The distributed firewall technology can be utilized as a means to augment an IP service node's capability for larger networks that are unable to adjust their configuration.

An ISP-centric approach is suggested as an attempt to consolidate some of the more basic but effective mechanisms available today to combat the challenges posed by a distributed threat. A redesign of the fundamental computing infrastructure with intrinsically secure systems is not immediately available due to many, largely non-technical, factors. The mechanisms discussed here are available today and would require smaller levels of investment.

B. FUTURE RESEARCH

This research illustrates the organizational and technical framework that can be used as the basis for an ISP centric approach. Further experimentation is needed to determine the robustness and efficiency of the architecture. In this regard, specific recommendations include:

The technologies outlined in the ISP centric framework, such as the BSN, currently do not support mechanisms that allow for automated and coordinated response across ISP nodes. An examination of the technical issues surrounding the networking of the service node concept to apply adaptive policy filters across nodes in response to denial of service attacks is warranted given migration to even greater transmission speeds.

Information was not available regarding the extent that military facilities functioning as ISPs are deploying these technologies, combined or in part. The design and implementation of the ISP-centric architecture, as described, within the NPS domain could provide the test bed for broader implementation across all military facilities.

One of the central components of a distributed firewall architecture is the centralized policy server. As employed in the 3Com solution, the central policy server is a standalone device that coordinates and manages policy distribution to the individual network interface cards (NIC). An examination of the technical issues surrounding the integration of the SpringTime 7000 or Shasta BSN to assume the role of policy server for the NIC-based firewall could be pursued in order to reduce the level of network complexity and infrastructure needed to support both.

LIST OF REFERENCES

1. R. Buddenberg, "Internet History", IS4502 Telecommunication Networks, Course Lecture, Winter Quarter 2001.
2. Chris Metz, "Interconnecting ISP Networks", IEEE Internet Computing March/April 2001.
3. Chicago NAP Peering Agreement version 3.2, May 1995 available at <http://nap.aads.net/MPLA.html>
4. T. Killalea, "Recommended Internet Service Provider Security Services and Procedures, RFC 3013, November 2000.
5. P. Ferguson, D. Senie, "Network Ingress Filtering", RFC 2827, May 2000.
6. A. Heffernan, "Protection of BGP Sessions via the TCP MD5 Signature Option", RFC 2385, August 1998.
7. C. Villamizar, R. Chandra, R. Gorvindan, "BGP Route Damping", RFC 2439, November 1998.
8. Marjory Blumenthal, David Clark, "Rethinking the Design of the Internet: The End-to-End Argument vs. the Brave New World", ACM Transactions on Internet Technology, Vol. 1, No. 1, August 2001, pp.70-109.
9. K. Houle, G. Weaver, N. Long, R. Thomas, "Trends in Denial of Service Attack Technology", Version 1.0, CERT Coordination Center, Carnegie Mellon University, October 2001.
10. P. Myers, "Subversion: The Neglected Aspect of Computer Security", Master's Thesis, Naval Postgraduate School, Monterey CA, June 1980.
11. M Radin, W. Scott, L. Scott, Whitepaper: "Distributed Denial Of Service Attacks: Who Pays?", Mazu Networks, Inc. 2001.
12. SANS Institute, NIMDA Worm/Virus Final Report, October 3, 2001.
13. J. Scambray, S. McClure, G. Kurtz, "Hacking Exposed: Network Security Secrets & Solutions Second Edition", McGraw-Hill, 2001.
14. Nortel Networks, Shasta 5000 BroadBand Service Node Service Creation System User's Guide, Version 2.1(2), Santa Clara, CA, 2000.
15. J. Heinanen, T. Finland, F. Baker, W. Weiss, J. Wroclawski, "Assured Forwarding PHB Group", RFC 2597, June 1999.

16. Brightmail Incorporated, "E-mail for the Twenty First Century: The Mailwall™ Solution", Whitepaper, version 1.0, 2001. Available at <http://www.brightlight.com>.
17. S. Bellovin, S. Keromytis, J. Smith, S. Ioannidis, "Implementing a Distributed Firewall", In Proceedings of Computer and Communications Security (CSS), pgs 190-199, November 2000. Available at <http://citeseer.nj.nec.com/ioannidis00implementing.html>.
18. 3Com Corporation, "3Com Embedded Firewall Architecture for E-Business: Stronger Security for Open Networks", Technical Papers, Santa Clara, CA, April 2001. Available at <http://www.3com.com>.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
Ft. Belvoir, VA
2. Dudley Knox Library
Naval Postgraduate School
Monterey, CA
3. Department of the Navy
Attn: CAPT Robert A. Zellman
Office of the Chief of Naval Operations (N6)
Washington, DC
4. Commander, Naval Security Group Command
Naval Security Group Headquarters
Fort Meade, MD
5. Fleet Information Warfare Center
NAB Little Creek
Norfolk, VA
6. Ms. Deborah M. Cooper
Deborah M. Cooper Company
Arlington, VA
7. Department of the Navy
Attn: Ms. Louise Davidson
Office of the Chief of Naval Operations (N643)
Arlington, VA
8. HQMC, C4
Branch Head, Information Assurance
Attn: Ms. Elaine Cassara
Washington, DC
9. Community CIO Office
Attn: Mr. William Dawson
Washington, DC
10. Community CIO Office
Attn: Ms. Deborah Phillips
Community Management Staff
Washington, DC

11. Office of Naval Research
Attn: Dr. Ralph Wachter
Arlington, VA
12. Defense Information Systems Agency
Attn: Mr. Richard Hale
Falls Church, VA
13. James P. Anderson Company
Attn: Mr. James Anderson
Ambler, PA
14. Superintendent
Attn: Dr. Cynthia E. Irvine
Naval Postgraduate School
Code CS/IC
Monterey, CA
15. Superintendent
Attn: Tim Levin
Naval Postgraduate School
Code CS/IC
Monterey, CA
16. Superintendent
Attn: LCDR Steven J. Iatrou
Naval Postgraduate School
Code IW/IS
Monterey, CA
17. Superintendent
Attn: Dan C. Boger
Naval Postgraduate School
Code IS
Monterey, CA
18. Linda Zupan
RTP, NC
19. HQMC, C4IA
Attn: Major Dan Morris
Washington, DC