DRAFT

More Bandwidth Doesn't Mean What You Think It Means: Why the Air Force Cannot Utilize the Full Potential of Its Enterprise Information Technology Systems

Sean M. Patrick

July 2003

Program on Information Resources Policy

Center for Information Policy Research

Harvard University

The Program on Information Resources Policy is jointly sponsored by Harvard University and the Center for Information Policy Research.

July 2003

PROGRAM ON INFORMATION RESOURCES POLICY

Harvard University **Center for Information Policy Research**

Affiliates

AT&T Corp.
Australian Telecommunications Users Group
BellSouth Corp.
The Boeing Company
Booz Allen Hamilton
Center for Excellence in Education
Commission of the European Communities
Critical Path
CyraCom International
Ellacoya Networks, Inc.
Hanaro Telecom Corp. (Korea)
Hearst Newspapers
Hitachi Research Institute (Japan)
IBM Corp.
Korea Telecom
Lee Enterprises, Inc.
Lexis–Nexis
John and Mary R. Markle Foundation
Microsoft Corp.
MITRE Corp.
Motorola, Inc.
National Security Research, Inc.
NEC Corp. (Japan)
NEST–Boston
Nippon Telegraph & Telephone Corp (Japan)
PDS Consulting
PetaData Holdings, Ltd.
Samara Associates
Skadden, Arps, Slate, Meagher & Flom LLP
Strategy Assistance Services
TOR LLC
United States Government:
Department of Commerce
National Telecommunications and Information Administration
Department of Defense
National Defense University
Department of Health and Human Services
National Library of Medicine
Department of the Treasury
Office of the Comptroller of the Currency
Federal Communications Commission
National Security Agency
United States Postal Service
Upoc
Verizon

Disclaimer

The views, opinions, and conclusions expressed in this paper are those of the author and should not be construed as an official position of the Department of Defense, the United States Air Force, or any other governmental agency or department.

EXECUTIVE SUMMARY

This paper presents a technical argument for changing the architecture of the current Air Force (AF) unclassified computer network. The AF is fielding enterprise level applications, such as the AF portal, that require users all over the world to connect to a single server or small number of servers. All AF users must establish a network connection through their local base network, out from the base network security perimeter, through the Defense Information Systems Agency (DISA) operated Non-secure Internet Protocol Router Network (NIPRNET), through another network security perimeter, and finally through the destination base network to the enterprise application server. This paper explains how the process described above prohibits the AF from maximizing the potential of its enterprise applications.

To maximize the potential of its enterprise applications the AF should be able to prioritize network applications and users in order to guarantee a high level of service to important network traffic. The current AF unclassified network architecture prohibits this "Quality of Service" (QoS) guarantee because the NIPRNET is not configured to implement the QoS features developed by the Internet community to support delay-sensitive applications.

Internet Protocol (IP) networks such as the Internet and NIPRNET have a limited amount of bandwidth at any given moment. There is no guarantee that the amount of bandwidth available is sufficient to transport all the data offered to the network at a particular moment. IP networks respond to congestion by dropping IP packets. Most network applications respond to dropped IP packets by resending the lost data and slowing the data transmission rate for that network connection. Some applications, such as multimedia (audio and video), are delay-sensitive. They simply ignore packets lost due to network congestion and do not adjust their transmission rate. This can result in a poor networking experience for application users on congested networks. A small amount of congestion is expected for IP networks because most applications will increase their data transmission rates until packet losses cause them to slow down. This paper contains an appendix that explains how the Internet community developed QoS techniques to control which IP packets the network drops and which ones get preferential treatment. All of these QoS techniques require that network routers between users and the application server have the QoS features enabled. Therefore, it is impossible to implement QoS for AF enterprise IT systems because NIPRNET, the long-haul path for connectivity of AF enterprise applications, does not support QoS.

Equally important as providing QoS for enterprise applications is protecting the AF network enterprise from hostile network activity. The current AF architecture leaves the AF vulnerable to distributed denial of service (DDoS) attacks. The current architecture also makes every AF location connected to the NIPRNET its own gateway to the outside world. This means that expensive network security equipment must be installed and maintained at over 100 locations. By changing the AF unclassified architecture and building an AF "Community of Interest Network (COIN)," the AF could substantially reduce the number of network ingress points to defend and protect itself from DDoS attacks.

From July 2000 until January 2002 the Air Staff conducted an effort to build an AF COIN within the Defense Information Systems Network (DISN). DISA fully supported this initiative. The difference between the proposed COIN architecture and the existing AF network architecture was that in the COIN, DISA would sell dedicated bandwidth between AF locations instead of connectivity to the NIPRNET at each AF location. This way, the AF would control all of the routers between AF users and enterprise network applications. The AF could then implement the QoS capability in its existing routers to ensure that network traffic from important applications and users received a high level of service. Unfortunately, the Air Staff dropped this proposal. Perhaps this paper will provide a catalyst to revive that effort.

Contents

Page

Illustrations

Page

Tables

Page

Chapter One

"Just Do It"

Many corporations in America today have revolutionized the way they conduct business by using information technology (IT) to reduce operating costs. Multinational corporations have standardized internal business applications such as electronic messaging, accounting, personnel and other applications so that all their employees worldwide use their corporate IT infrastructure and Web browsers to connect to corporate enterprise application servers. Currently, the U.S. Air Force (USAF) cannot implement this business practice effectively because it does not control the end-to-end quality of service (QoS) of its enterprise business applications. The reason is that USAF base networks (what would be called "campus" networks in industry) are not connected together by dedicated data links. Instead, a USAF computer user at one base who wishes to connect to a USAF server located at another base must use the Defense Information Systems Agency (DISA)-provided Non-secure Internet Protocol Router Network (NIPRNET) for its wide-area connectivity. From July 2000 until January 2002, the Air Staff undertook an effort to change the USAF unclassified network architecture to improve enterprise network performance and security. Unfortunately, this initiative was killed. Perhaps this paper might provide a catalyst to revive that effort.

In early July 2000, the then Secretary of the Air Force (SECAF), F. Whitten Peters, and then Chief of Staff of the Air Force (CSAF), General Mike Ryan, hosted a summit of Air Force four-star generals in Washington, D.C. The purpose of this Information Technology Summit was to describe to the senior Air Force leadership several examples of how U.S. industry was leveraging the power of IT to drastically reduce operating expenses and increase employee productivity. The generals were told that one multi-national company had cut annual expenses by $1 billion by consolidating hundreds of disparate company IT systems located in offices throughout the world into a single corporate Web-based system.[1] The SECAF and CSAF told the assembled group that they envisioned implementing industry's IT practices within the U.S. Air Force. They concluded the meeting by using a well-known slogan from the Nike Corporation: "Just Do It."

The task of implementing the new IT vision fell to the Deputy Chief of Staff for Communications and Information, Lieutenant General John L. (Jack) Woodward. General Woodward's staff organized several focus groups that met for two weeks to develop a strategy. IT experts from across the USAF assembled in Washington, D.C., and wrote plans for a USAF Enterprise Concept of Operations (CONOPS), communications and computing transport layer, server consolidation, a USAF portal, information assurance, funding, directory services,

---

[1]Oracle Corporation, "Making the Complex Simple," Oracle White Paper (Redwood Shores, Calif.: Oracle Corporation, April 2001).

acquisition agility, and marketing.[2] One common shortfall identified by several of the focus groups was a lack of available bandwidth to support consolidation of electronic mail (e-mail) servers over long distances. The USAF's existing unclassified network architecture simply could not support the new vision.

The computing transport and information assurance focus groups began working with DISA on plans to build a USAF "Community of Interest Network (COIN)."[3]The difference between the COIN architecture and the existing USAF network architecture was that the COIN would provide dedicated bandwidth between USAF locations, while the existing architecture connected all USAF locations to each other through the NIPRNET, which is shared by the entire Department of Defense (DoD). A private network provides some very important network security and performance benefits. It is important to note that the focus groups working on the COIN design did not believe that a private network would save money or manpower within the USAF. Instead, the COIN was intended to allow base-level network administrators and security personnel to focus their efforts on hardening the internal security and performance of their base network. An organization external to the base would be responsible for defending the USAF unclassified network from threats external to the USAF, but each base would still be responsible for defending itself against threats internal to the USAF.

The purpose of this paper is to examine the technical case for building a USAF intranet, or COIN. It is intended for the reader with a general understanding of computer networks. This paper describes why and how a USAF intranet or COIN could provide better overall network security and guaranteed performance for USAF enterprise applications not currently available in the USAF unclassified[4] network architecture. **Chapter Two** describes the current unclassified USAF network architecture and performance issues related to that architecture. **Chapter Three** examines how the current architecture leads to some security vulnerabilities. **Chapter Four** provides a general description of how a USAF Intranet or COIN could be constructed and how the USAF could use that COIN to enhance network enterprise security and provide guaranteed level of service for USAF enterprise applications.

The **Appendix** describes how IP router networks such as the Internet and NIPRNET work. It also explains why the NIPRNET architecture prohibits the use of QoS features. The appendix is

---

[2]LtGen John L. (Jack) Woodward, e-mail to Major Command Chief Information Officers, subject: IT Summit Call to Action, July 20, 2000.

[3]The term "COIN" was used instead of "Air Force Intranet" to avoid the appearance that the USAF endorsed the Navy/Marine Corps Intranet (NMCI) concept. While there are security and performance advantages to an intranet, the USAF does not support the total outsourcing of fixed-base IT support, whereas outsourcing is a key ingredient of NMCI.

[4]This paper is limited to the unclassified USAF network architecture. The technical limitations of an IP router network also apply to the DISA-managed Secret Internet Protocol Router Network (SIPRNET). USAF enterprise IT applications using the SIPRNET for wide-area connectivity are vulnerable to the same QoS limitations described for the NIPRNET in the appendix.

intended to impart a rudimentary understanding of the Open Systems Interconnect (OSI) reference model and how information flows in a packet-switched network. Readers that wish to refresh their understanding of Transmission Control Protocol (TCP)/Internet Protocol (IP) networks, the “bursty” nature of IP traffic, network congestion, and the causes of latency (delay) in IP networks are advised to read the appendix prior to reading the rest of this paper.

Chapter Two

The Air Force "Enterprise": This Spaceship Isn't Built for Light Speed

This chapter will describe the USAF unclassified network architecture and performance issues that affect the USAF network enterprise. For the purposes of this paper, the term "Air Force enterprise" is defined as:

> The aggregate of people, systems, resources, and processes that provide the information availability and assurance that enable Air Force core missions. It is not constrained by time or operating location, and assures access to the information required for decision-making and mission accomplishment.[1]

2.1 Every Base Is a "Gateway"

Figure 2-1 shows how USAF bases are connected to each other. The diagram has been simplified for clarity. Although it shows only 12 Air Force bases connected to the NIPRNET, in reality there are over 100 USAF connections to the NIPRNET.

For a network user on base A to connect to a server located on base B, he/she must pass through the local base A network, exit the base network perimeter, traverse the NIPRNET, pass through base B's perimeter, and then traverse the base B network to connect to the desired server. This figure shows that each USAF location connected to a NIPRNET router functions as its own network gateway to the outside world. This means that each location must protect itself from hostile network activity that originates off base. Whenever fast-spreading computer worms or viruses threaten the Internet (and therefore the NIPRNET), as the Code Red virus did in 2001, the USAF must implement defensive measures at over 100 locations.

---

[1]Concept of Operations for the USAF Information Enterprise, Draft Version 15, 18 June 2001, HQ USAF computer file folder: il_public/af ilc (dir of comm ops)/focus groups/af it enterprise

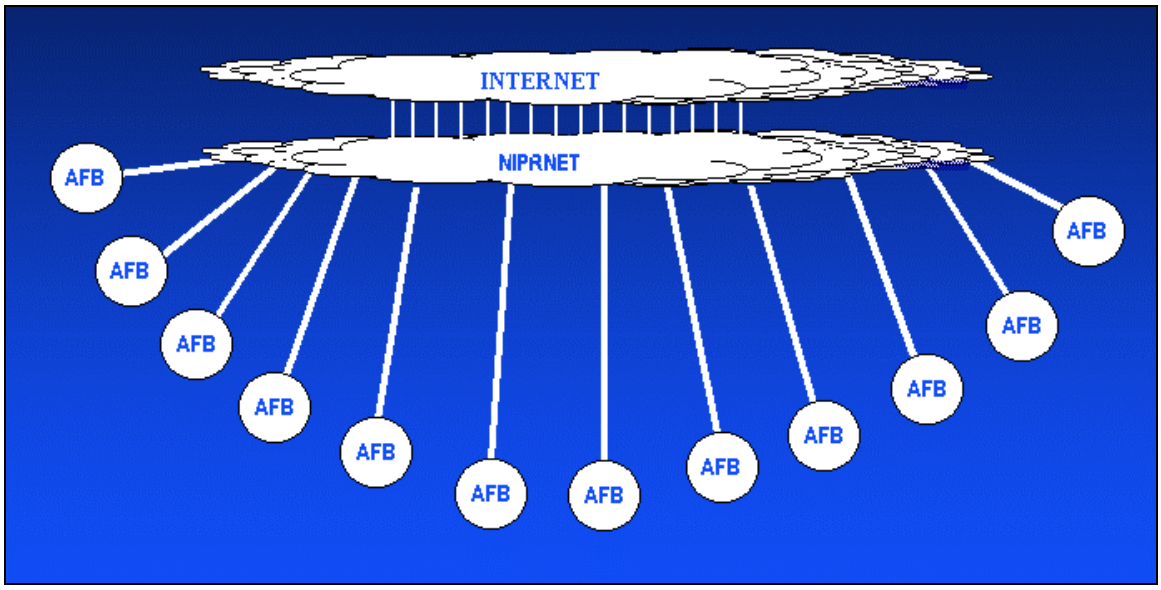


Figure 2-1
Air Force Unclassified Enterprise Architecture

Figure 2-2 is a diagram of the equipment used for the "standard" USAF network perimeter. The equipment is installed by the USAF Combat Information Transport System (CITS) Program Office, Electronic Systems Center (ESC), at Hanscom Air Force Base (AFB), Massachusetts. The drawing shows that there are seven single –points –of failure at each USAF base gateway: service delivery point (SDP) router, intrusion detection system (IDS) switch, external router, external switch, firewall, internal switch, and internal router. If any one of these devices fails, base network users cannot access enterprise applications located off base. The CITS program is currently changing this architecture to put the Web proxy in parallel with the base firewalls, but this process is not yet complete.

In most cases, there is an eighth single –point –of failure, because most bases have a single data link from the SDP router to the DISA NIPRNET edge router (DISA uses the term "hub" for its edge routers). If the DISA edge router fails, the base is also cut off from unclassified data traffic external to the base. Although the following discussion of the availability of enterprise applications does not include this router, it is important to consider it for end-to-end QoS for these applications.

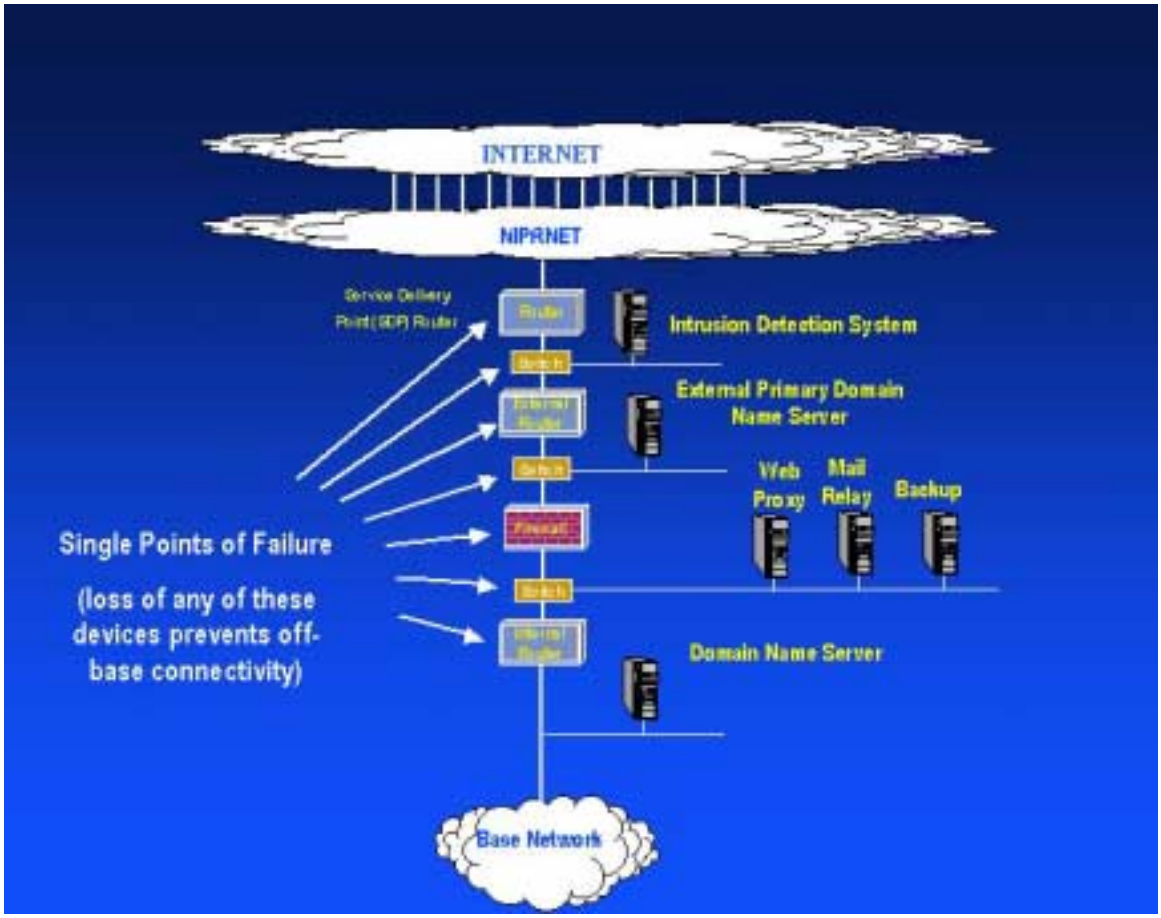


Figure 2-2
USAF Base Perimeter

2.2 Availability of Enterprise Applications

Many factors affect the availability of enterprise applications. Since this paper is intended to show how a COIN would enhance security and performance, it confines the discussion of the difference in availability to the factors that differ between the current architecture and a potential COIN architecture. In a COIN architecture, not every base would be a gateway to the Internet (explained in more detail in **Chapter Three**). Instead, bases would be connected to at least two other USAF locations by Defense Information Switched Network (DISN) data links (layer 2), not through the NIPRNET (layer 3). Therefore, the end-to-end path for a user accessing a USAF enterprise system would be different with regard to the base perimeter and long-haul path.

For the purpose of analysis, let us assume that the availability of the long-haul path is 100 percent for both NIPRNET (layer 3) and DISN (layer 2). Actually the availability of the DISN should always be as high as or higher than that of the NIPRNET, since the NIPRNET (layer 3)

uses the DISN (layer 2) for connectivity. If the DISN fails, the NIPRNET fails also. The reverse is not true.

We can now analyze the difference in expected availability for enterprise applications between the current architecture and COIN architecture by comparing the expected availability using the current base perimeter and a proposed COIN base perimeter that would have no single points of failure for off-base connectivity.

The military *Electronic Reliability Design Handbook* gives equations for the overall availability of electronic systems in series (A_S) and electronic systems in parallel (A_P), as shown in **Figure 2-3**. In these equations, if any system in a series fails, the whole system fails (single – point –of failure). For parallel systems, only one path is needed for communication. If one system fails, other paths are available.

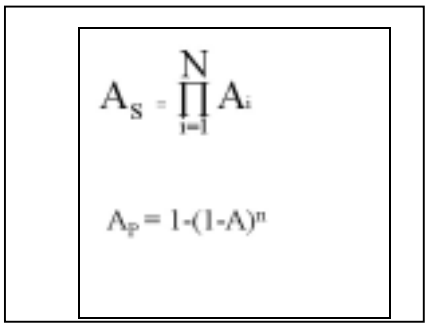


Figure 2-3
System Availability Equations

N = The number of systems in series with each individual availability A_i

n = The number of systems in parallel with the same availability A.

Individual availability = MTBF/(MTBF+MTTR)

MTTR = Mean Time to Repair = Maintenance time + administrative and logistics delay

MTBF = Mean Time Between Failures

The first equation means that the overall system availability is equal to the product of the individual system availabilities ($A_S = A_1 * A_2 ***A_N$). The second equation is straightforward, but it is important to note the exponential nature of this equation.

Table 2-1 indicates the expected availability of perimeter networking equipment using MTBF information provided by the CITS program office for each device. The availability values assume a one-hour repair time per outage and a three-hour administrative and logistics delay

(time required to detect the failure, order, ship, and receive a replacement part). The numbers do *not* reflect outages in a perimeter device caused by software or human errors. Only hardware failures are considered.

Table 2-1

USAF Perimeter Equipment Availability

Perimeter Device	**Availability**
SDP Router	.99980
IDS Switch	.99996
External Router	.99980
External Switch	.99996
Firewall	.99980
Internal Switch	.99996
Internal Router	.99980

Using this table, A_S = .99908. That is, the overall expected availability of a base perimeter when considering only hardware failure of the components is 99.908 percent. Stated another way, the expected amount of downtime per year is (1 – .99908) * 24 * 365 = 8.06 hours per year. Since an enterprise network user must pass two boundaries to reach a server located on another USAF base, the expected availability for any given user on one base to cross both boundaries is .99908 * .99908 = .9982. This equates to an expected downtime of **16.1 hours per year**.

Now let us consider an enterprise user in a COIN architecture. The minimum requirement for the base perimeter in a COIN is two SDP routers connected to different locations in the base network backbone (in parallel to the outside world from the user's point of view). The expected availability, A_P when n = 2 is .99999996. This equates to an expected downtime of **1.26 seconds per year**. But since the server is also located on a base with two access points with an A_P of .99999996, the end-to-end expected availability is $(.99999996)^2$ (because the two base perimeters are in series with each other), or .99999992. This equates to an expected downtime of **2.52 seconds per year**.

To put this in context, the expected availability of USAF enterprise applications in a COIN architecture is **23,000** times the expected availability of those applications in the current USAF unclassified network architecture if one considers only hardware failures in the base network perimeter equipment.

The exponential nature of the availability equation means that when devices are implemented in parallel and a third SDP is included in the COIN architecture, the expected availability is **114 million** times the current expected availability when considering only network perimeter device failure. For all practical purposes, the expected downtime due to a perimeter device failure in this case is **zero** ($1.4 * 10^{-7}$ hr/year).

Some USAF enterprise systems, such as the USAF portal, are hosted by DISA. The network perimeter architecture for those systems is different than the standard USAF base perimeter architecture and probably does not have as many single points of failure in series. The expected availability for these systems is higher than for those enterprise systems hosted on USAF bases, but it is still not nearly as high as the expected availability of USAF enterprise systems hosted in a COIN.

2.3 Quality of Service (QoS) for USAF Enterprise Applications

The **Appendix** describes the store-and-forward nature of network routers and the ability to provide QoS guarantees to different types of IP packets. Unfortunately, the current USAF unclassified network architecture (Fig. 2-1) prohibits the USAF from using this capability because USAF-to-USAF packets must pass through the NIPRNET routers.

2.3.1 Quality of Service (QoS) in the NIPRNET

Routers can be used to give special handling to high-value packets in order to provide a better level of service for those packets. None of the methods described to provide QoS for IP traffic is used in the NIPRNET today. Since the QoS methods described in the Appendix require that *every* router between an enterprise application user desktop and the enterprise application server *that may experience any congestion at all* must have the same QoS capability enabled to take advantage of this technology, the current USAF network architecture cannot support QoS for USAF enterprise applications.

2.3.2 Congestion in the NIPRNET

Congestion in IP networks is expected and results in packet loss. When packets are dropped, the application expecting the packet either ignores the missing packet or retransmits the missing packet and slows down its transmission rate depending on which layer 4 protocol is being used (see the Appendix for explanation). DISA monitors some congestion statistics for the NIPRNET. In data obtained from November 2002, for 248 tests between NIPRNET routers, 116 (47%) indicated some level of packet loss. Most of the loss was minimal (less than 1%), which is to be expected. Because of the “bursty” nature of Internet traffic, it is difficult to determine if these tests give a fair indication of the level of congestion in the NIPRNET. It is possible for a test to indicate no packet loss during one run and then show high packet loss on another run just seconds later. Also, the test results indicate the packet loss and latency measurements were done on

backbone trunks of the NIPRNET. The results did not show congestion data between the NIPRNET edge routers (DISA calls them hub routers) that support all USAF-to-USAF traffic.

Since network congestion is usually more likely to occur at the edge of a network than in the core of a network, these test do not provide an accurate indication of how congestion in the NIPRNET affects USAF-to-USAF traffic. Since most ABs have a single connection to a NIRNET hub router, if there is any congestion on the upstream links of that router as a result of other DoD traffic, there is a likelihood that USAF packets will be dropped and USAF applications will slow down as a result. Further study is needed to determine the true impact of NIPRNET congestion on USAF-to-USAF traffic sent through the NIPRNET.

2.4 USAF Enterprise Application Performance Issues

The previous discussion showed why the USAF cannot guarantee end-to-end QoS for enterprise applications. Simply stated, the USAF cannot control what packets get dropped in the NIPRNET.

2.4.1 Impact of NIPRNET Congestion on USAF Enterprise Applications

Even if the USAF were to establish policies requiring that it be a good steward of wide area network (WAN) bandwidth, there is no assurance that other DoD agencies will have those same policies. If the NIPRNET links in the DISA edge router providing service to a given USAF base were to become congested, packets would be dropped according to DISA's congestion management policy. Some of those packets would very probably be USAF TCP segments, which would cause the USAF applications to retransmit the segments and slow the data flow rate. If the DoD agencies connected to the other interfaces of that edge router had poor traffic policies—for example, allowing streaming audio from Internet radio broadcasts (User Datagram Protocol packets) during periods of congestion—USAF TCP applications would "throttle back" and give up wide area bandwidth for the other DoD traffic. Thus, the current architecture does not permit the USAF to control the end-to-end QoS for its enterprise IT applications.

Moreover, if the USAF were to purchase additional bandwidth from a base to an already congested NIPRNET, there is no guarantee that the USAF would see the expected benefit from the additional bandwidth. Dr. David Clark of the Massachusetts Institute of Technology (MIT),[2] one of the leading pioneers of the Internet,[3] agrees with the author's analogy that increasing bandwidth to an IP "cloud" is like widening an on-ramp to a Los Angeles freeway. If there is no congestion on the highway, more traffic will flow into the network. However, when (not if) the highway is congested, the increased bandwidth will fail to provide the expected benefit. The

---

[2]David Clark, personal interview with author, 200 Technology Square, Cambridge MA, 15 Nov. 2002.

[3]The Internet Society, "A Brief History of the Internet" (Reston, Va.: The Internet Society), [On-line]. URL: http://www.isoc.org/internet/history/brief.shtml (Accessed on 11 Nov. 2002.)

increased traffic in the “on-ramp” must compete with the other network traffic for the available wide area bandwidth. Without some type of end-to-end QoS capability, the effective throughput through the IP cloud (like the Internet or NIPRNET) is unpredictable.

In a COIN architecture, the USAF would “own the highway” and could establish policies to give preferential treatment to important enterprise applications or users. For example, the USAF could decide that during times of network congestion, e-mail and file transfer packets would be dropped before Web connections to enterprise servers or interactive video sessions. It does not matter if an e-mail or file transfer takes several seconds longer to reach a recipient (the TCP protocol will retransmit any lost packets and slow the transfer), but any additional delay in an interactive network session can become annoying to the user.

Chapter Three

Enterprise Security—Our Shields Are Fragile

No computer network is totally secure. As long as humans write computer code and operate computers, there will be network security incidents. They are an unavoidable risk of doing business, much like the unavoidable risks of aircraft accidents in aviation. In response to unfavorable publicity surrounding several fatal Army plane crashes, Brigadier General Oscar Westover, the Assistant Chief of the Army Air Corps and commander of the Army Air Corps Mail Operation (AACMO), the military department tasked to fly the air mail in 1934 when the commercial air mail contracts were canceled, sent a message to his air mail zone commanders. It stated: "There will be no more accidents." One of his zone commanders, B.Q. Jones, replied: "There will be no more flying."[1] Likewise, the appropriate response to any prospective order "There will be no more network security incidents" would be "There will be no more networking." Security problems go with the territory. However, just as there are ways to reduce the probability of aircraft accidents, there are ways to reduce the probability of network security incidents.

3.1 Current USAF Network Security Architecture—Defending Everywhere

Figure 2-1 in the previous chapter shows that each USAF base must defend itself from malicious outside network traffic. Thus, there are over 100 locations where USAF personnel must employ a combination of hardware, software, and procedures to defend the USAF enterprise against hostile network activity.

3.1.1 Firewall Types

There are three basic types of firewalls: application proxy, stateful inspection, and packet filter. All three types of firewalls have at least two network interface cards that split the network into a trusted (internal) zone and an untrusted (external) zone.

Application proxy firewalls actually make connections on behalf of the user. If a user on the untrusted side of the network attempts to connect to a server on the trusted side, the firewall will intercept the connection request. If the connection is allowed based on pre-existing rules within the firewall, the firewall will make the connection to the internal server.

Application proxy firewalls can look into the highest layer of the protocol stack to enforce rules on what types of connections are allowed. For example, the File Transfer Protocol (FTP) has both "put" and "get" commands. The "put" command transfers a file from the connection

---

[1]DeWitt S. Copp, *A Few Great Captains: The Men and Events That Shaped the Development of U.S. Air Power* (McLean, Va.: EPM Publications, 1980), 196.

originator's computer to the target computer, while the "get" command downloads a file from the target computer to the connection originator's computer. An application proxy firewall can be configured to allow outbound FTP connection "puts" but deny "gets" across the security boundary to prevent downloading of files that may contain viruses. While application proxy firewalls are considered by some to be the most secure type of firewall, they suffer from performance limitations because they must establish and maintain two connections (one external, one internal) across the security perimeter.[2]

Stateful inspection firewalls keep track of connections between the trusted and untrusted network zones (and vice versa). When a user on the untrusted side of the network wants to connect to a server inside the trusted network, the firewall will either permit or deny the connection based on pre-set rules in the firewall. If the connection is allowed, the firewall will maintain a state table for each connection across the security perimeter. When packets arrive on the external interface of the firewall, the firewall first checks to see if they are part of an existing authorized connection and, if so, forwards them inside the network. If not, the firewall must check its rule set to see if the packet is allowed to pass through the security perimeter. For example, a stateful inspection firewall could be configured to allow FTP connections only in one direction (outbound or inbound). Since these types of firewalls do not look into the application layer of the stack they could not be used to enforce rules on FTP "put" or "get" direction (assuming the connection is allowed in the first place). Stateful inspection firewalls have a much greater traffic capacity than application proxy firewalls.[3]

Packet filter firewalls are just that: filters. When packets arrive on the external or internal interface of the firewall, they are checked against a rule set within the firewall (usually a list of prohibited IP address and port numbers). If the packet matches any prohibited criteria, it is dropped. If not, it passes across the security perimeter of the network. A packet filter firewall does not require a separate hardware device because the rule sets used for packet filtering can be implemented in network routers. For this reason, some in industry do not consider packet filters to be "firewalls."[4]

3.1.2 Air Force Firewalls

The primary firewall used by the USAF is an application proxy firewall. Although Figure 2-2 shows one firewall "box" at each base perimeter, in reality many bases have more than one firewall because of throughput limitations. The CITS program office recently conducted a live

---

[2]The MITRE Corporation, *Firewall Concepts and Selected Products Pertinent to the Air Force Combat Information Transport System (CITS)*, MP 01B0000057 (Bedford, Mass.: The MITRE Corporation, November 2001).

[3]Ibid.

[4]SecurityDogs.com, Firewall Security Overview, [On-line]. URL: http://www.securitydogs.com/firewall_overview.html (Accessed on 23 Feb. 2003.)

network evaluation of the capabilities of the USAF firewall. The results showed that the maximum throughput without encountering network problems was approximately 25 MBps.[5]

3.1.3 Intrusion Detection

The Air Force uses a network-based intrusion detection system, located at each USAF base network perimeter to monitor traffic traversing the base security perimeter. If it detects any connections meeting a predefined malicious signature, an alert is sent to the USAF Computer Emergency Response Team (AFCERT) and the Major Command Network Operations and Security Center (NOSC). If the AFCERT identifies a hostile source IP address, it contacts the USAF Network Operations Center (AFNOC), which then updates the Access Control List (ACL) of each USAF SDP router to deny packets from the hostile IP address.

3.1.4 Web Proxy

The standard AFB network perimeter also includes a proxy for Web traffic. This device is used to cache Web traffic and filter malicious Web connections. When a network user wishes to connect to a Web server, his/her browser connects to the Web proxy, which establishes the Web connection on his/her behalf. When the Web server returns the Web page content, it is stored (cached) on the Web proxy server for a period of time in case there is another request for the same Web page. If so, the proxy server sends the stored content directly to the Web browser instead of forwarding the request outside the base. This improves network efficiency. Since these devices operate at the application layer of the protocol stack, they can be used to enforce policy to prevent Web connections to prohibited sites.

3.1.5 E-mail Relay

Internet viruses are often carried as attachments to e-mail. The standard AFB perimeter includes a mail relay that can be used to filter e-mail messages known to contain malicious code.

3.1.6 Network Security Software

Each USAF location has a suite of network security software to protect the internal network. This includes anti-virus software for desktops and servers, scanning software to determine if software patches have been installed on network equipment, and other software to keep track of compliance with network vulnerability alerts.

---

[5]CITS Program Office, “Combat Information Transport System (CITS) Evaluation of Secure Computing Corporation’s Sidewinder 5.2, 08 January–01 February 2002,” 20 Feb. 2002 test report (Hanscom Air Force Base, Mass.: CITS Program Office, 2002).

3.2 Impacts of Current Security Architecture

3.2.1 Multiple Proxies for Enterprise Applications

As shown above, the USAF security architecture is based on the idea that maximum security is obtained if there are no direct connections between nodes on the two different sides of the network security perimeter. Connections are proxied by the firewall or by the Web proxy server if the latter is in parallel with the firewall (boundary configuration implemented at some USAF bases). For a user on one base (base A) to connect to a server at another base (base B), the firewall (or Web server if it is in parallel with the firewall) at base A must proxy the connection off base A. Since there is another proxy at the perimeter of base B, the connection is "proxied between the proxies" before the connection finally reaches the intended server. That is, the user at base A establishes a connection with the firewall (or Web proxy) at base A, which then establishes a connection with the firewall (or Web proxy) at base B, which then establishes a connection with the target server (if the rules allow). The many handshakes and flow control messages needed to establish and maintain multiple proxy connections (see Appendix) add considerable latency to the end-to-end connection. As mentioned in Chapter Two, the CITS program is changing the standard architecture and placing the Web proxy in parallel with the firewall. Those locations that still have the Web proxy behind the firewall have even more proxy connections. The Web proxy must proxy the connection with the firewall before it is sent again by the firewall to the distant Web server. If the Web server is behind another proxy firewall, it is proxied again before reaching the server.

3.2.2 Vulnerability to Distributed Denial of Service (DDoS) attacks

Figures 2-1 and 2-2 show that the USAF is vulnerable to DDoS type attacks. If a DDoS attack were to target any given base, it could consume all the available bandwidth between the SDP router and the NIPRNET edge router. It is not likely that USAF network security personnel would detect the attack. The USAF would probably learn of the attack from disgruntled network users who complained of lost or severely degraded off-base connectivity. The USAF would then have to work with DISA to identify the offending traffic (a non-trivial event) and have it blocked at the gateways from the Internet to the NIPRNET.

3.2.3 DISN Expansion Program

DISA is pursuing a program to greatly increase the amount of available bandwidth within the DISN. This program, if funded, will be capable of providing multiple, very high bandwidth rates (155 MBps and higher) to each USAF base. The current USAF network security architecture cannot support these high data rates because of the previously mentioned limited capability of the application proxy firewalls. It is unclear if the current network intrusion detection system can handle such high data rates.

A COIN could take advantage of the DISN Expansion Program by removing the bottlenecks at each base to allow multiple high-speed connections between USAF locations. USAF-to-NIPRNET gateways would require high-speed, survivable boundary protection devices.

3.2.4 Damage Mitigation and Reconstitution During Hostile Internet Activity

A recent article in the *NATO Review* describes four ways of reducing the vulnerability of computer networks during any future cyber war: anticipation and assessment, preventive or deterrent measures, defensive measures, and damage mitigation and reconstitution. The current USAF unclassified network security architecture incorporates the first three methods but does not support the fourth. The authors of the article, all affiliated with the Computer Emergency Response Team (CERT) Analysis Center at Carnegie Mellon University, argue that networks should be able to perform critical missions without relying on outside connectivity that could be lost during hostile network activity. They write: "Insulated intranets that can operate efficiently and safely without wider connections offer considerable promise in this respect."[6] Since the USAF does not have an intranet or COIN, it cannot maintain connectivity for mission-critical USAF enterprise applications to mitigate the impact of hostile cyber events that flow from the Internet into the NIPRNET.

[6]Timothy Shimeall, Phil Williams, and Casey Dunlevy, "Countering Cyber War," *NATO Review*, Winter 2001/2002, [On-line]. URL: http://www.cert.org/archive/pdf/counter_cyberwar.pdf (Accessed on 24 Jan. 2003.)

Chapter Four

What a COIN Can Do for the Air Force: Recommendations for Adding Warp Speed to the Enterprise

Figure 4-1 shows a high-level diagram of what a USAF COIN could look like. Each base is connected to two other USAF locations using DISA-provided data links (layer 2) to provide multiple paths for base users to connect off base. The USAF would reduce the external gateways to defend from malicious outside activity. USAF enterprise application servers (such as the USAF portal) would be moved into the USAF network. Since the USAF would "own the highway" for USAF-to-USAF network connections, it could establish policies to provide a higher degree of service to important enterprise applications and users.

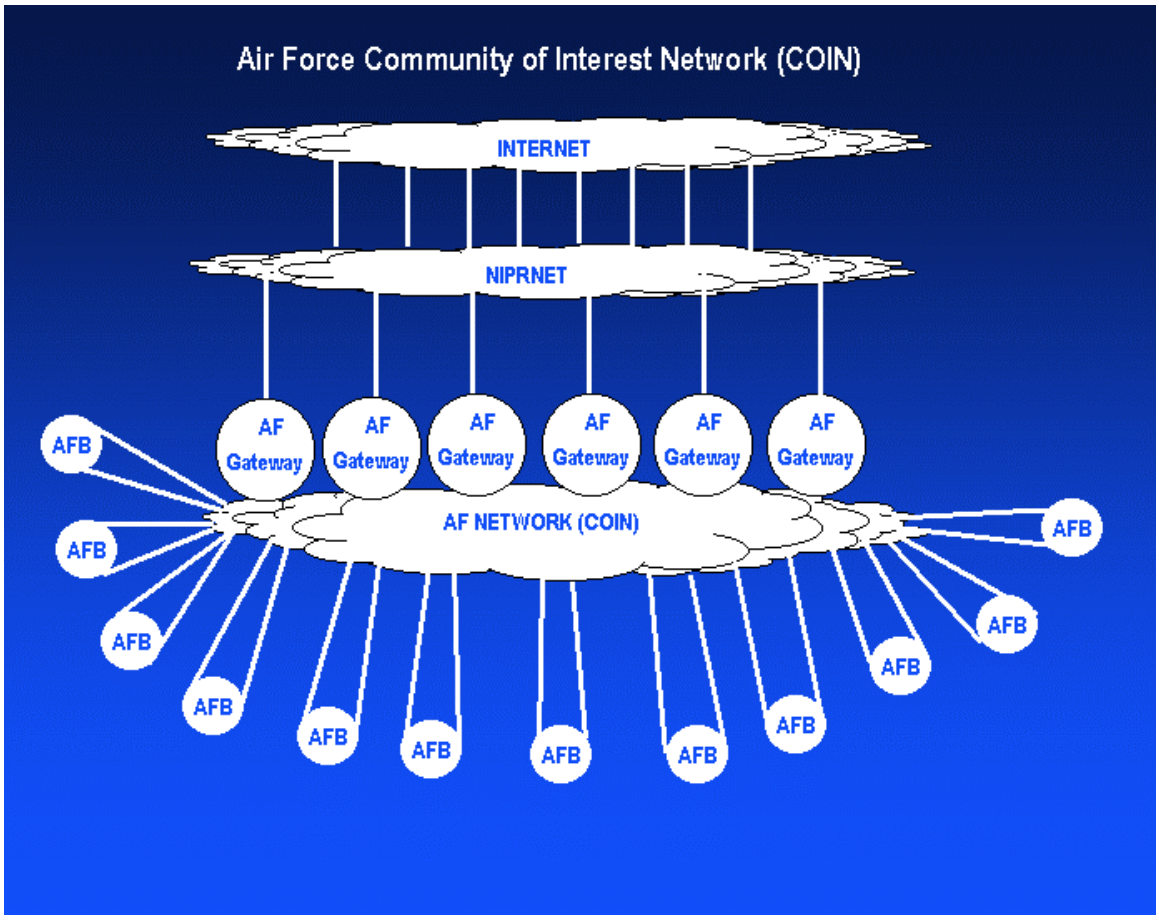


Figure 4-1
Hypothetical Air Force Unclassified Community of Interest Network

4.1 Improving Enterprise-Level Network Performance

4.1.1 Provide QoS Capability for Important Applications and Users

The USAF should decide which enterprise applications and users deserve QoS guarantees and then implement the QoS capabilities already present in USAF routers[1] to support these data flows. For example, the USAF may decide that video packets should have special handling. The USAF could implement a QoS scheme in its routers to make sure video packets are not dropped and receive precedence in router queues. The USAF could also decide which user traffic should have special handling in the network. For example, video packets could receive one level of service but video flows between Major Command Commanders (four-star generals) and their Wing Commanders (mostly one-star generals) could be given an even higher level of service.

If the Air Force portal and Global Combat Support System Integration Framework (GCSS–IF) were brought into the USAF COIN, traffic flows to these critical IT assets could be given an increased level of service in order to provide QoS to users of the portal and GCSS-IF. The USAF portal is currently hosted in a DISA facility on Gunter Annex of Maxwell AFB; the AFNOC is located on a different floor of the same facility. From a technical standpoint, it would not be terribly difficult to move the USAF portal inside any future USAF COIN.

4.1.2 Eliminate Multiple Proxies Between Users and Enterprise Applications

A USAF COIN architecture should allow network users to connect to enterprise application servers through the minimum number of proxies needed to provide efficient connectivity and security. This would eliminate the latency caused by the multiple proxies currently in use between USAF users and USAF enterprise application servers.

4.2 Enhancing Enterprise-Level Network Security

4.2.1 Reduce the Number of USAF External Gateways

A USAF COIN architecture could drastically reduce the number of enterprise boundaries. This would eliminate the need for costly and bandwidth-limiting application proxy firewalls at over 100 locations. Enterprise-level boundary defense should include Application Specific Integrated Circuit (ASIC)-enabled firewalls and intrusion detection systems to allow high-speed throughput without sacrificing security at the enterprise gateways. Internet-bound USAF traffic should pass through the USAF gateways into the NIPRNET, and out of the NIPRNET gateways to the Internet.

---

[1]Cisco offers an Enterprise QoS Policy Manager for its routers at a one-time cost of around $15,000 (unlimited licenses).

In a COIN, bases would be “shielded” behind USAF-level gateways. In the event of a DDoS attack against the USAF, the USAF gateways could block the external malicious traffic and still permit network communication between bases inside the COIN.

In a COIN architecture, base-level network security personnel would no longer be responsible for protecting the local network from threats external to the USAF. Local security experts would then be able to dedicate their time to protecting the USAF from insider threats.

4.2.2 Harden the Internal Networks

The USAF is pursuing a server consolidation initiative. In some cases, many servers which were previously scattered across the USAF base are being located in the base Network Control Center (NCC). The USAF may wish to consider placing network security devices (firewalls, intrusion detection, et cetera) at each of these locations to protect these assets from internal threats. The USAF could also explore using desktop firewalls and intrusion detection systems to protect each machine on the base network. While this sounds like a daunting task for bases that have thousands of users, computer security companies have developed management software that makes this task feasible.

Local base security personnel could also establish a practice of scanning the local infrastructure for unauthorized paths into the network, such as wireless access points, remote access servers (modem banks), and commercial network connections. For those USAF offices that must protect sensitive data such as personnel, financial, legal, security, or medical information, the local base could implement an encrypted file system (EFS).

4.2.3 Develop Continuity of Operations Plans (COOPs)

A COIN architecture would require Continuity of Operations Plans (COOPs) to support mission-critical applications in the event an enterprise gateway is lost due to natural disaster or hostile action. This COOP needs to include relocation of critical support functions to alternate locations within the USAF enterprise. It could also include procedures for “network minimize” to limit unnecessary traffic during times of reduced capability. COOPs need to be exercised on a regular basis.

4.2.4 Establish an Air Force Network Operations and Security Center (AFNOSC)

Currently, the AFNOC is responsible for USAF enterprise network operations, while the AFCERT is responsible for enterprise network security. When network events occur, it is often difficult to determine if the incident was caused by malicious activity, hardware or software failure, or operator error. Although the AFNOC and AFCERT work well together, the USAF might be better served by a single organizational structure responsible for both enterprise network

security and network operations. USAF enterprise gateways could be managed by Major Command NOSCs or by the AFNOSC if/when it is established.

4.3 Drawbacks to a COIN

Despite its overall advantages, a COIN would have some drawbacks. The most important ones are described below. However, these disadvantages pale in comparison to the advantages.

4.3.1 Enterprise Network Management Complexity

A USAF COIN would require additional enterprise management functions not required by the current USAF network architecture. The AFNOC would be required to implement interior and exterior routing protocols on USAF routers. For those bases inside the COIN, the AFNOC would need to implement an interior routing protocol such as Open Shortest Path First (OSPF) to direct AFB-to-AFB traffic. For the enterprise gateways, the AFNOC would be required to implement a Border Gateway Protocol (BGP) to advertise routes to USAF bases to the outside world.

4.3.2 DISA NIPRNET QoS Efforts

DISA may implement one of the QoS techniques described in the Appendix within the NIPRNET. If that were to happen, the USAF would need to determine if there were a sufficient guarantee that USAF mission-critical traffic would receive priority treatment within the NIPRNET. If DISA were to implement QoS in the NIPRNET, each DoD agency would be required to mark its traffic according to agreed-upon levels of service. However, there is no assurance that all DoD agencies would adhere to the agreed-upon level of service markings; in other words, there would be no way to ensure that a DoD agency did not mark traffic for a higher level of service than it should receive. USAF mission-critical applications would compete for NIPRNET resources (bandwidth, router processing, et cetera) with all other DoD traffic marked with the same level of service.

The Differentiated Services (DS) Architecture RFC[2] recommends that Internet service providers (ISPs) examine packets entering edge routers to confirm that customers are not marking packets for preferential treatment that they are not entitled to mark under the service-level

[2]IP specifications are contained in documents called “Requests for Comments” (RFCs). Since the Internet was to be based on open standards, the original intent was that researchers and software developers would propose new protocols and post them for public comments. After a sufficient period for editing, the RFCs were to be published as Internet Engineering Notes (IENs). Since some RFCs were good enough as written and others were completely rewritten and republished as new RFCs, the term IEN was never used and the accepted protocol standards are published as RFCs. (Source: Douglas Comer, *The Internet Book* (Upper Saddle River, N.J.: Prentice Hall, 1997), 57. RFCs can be found at URL: http://www.ietf.org/rfc.html (Accessed on 22 Feb. 2003.)

agreement (SLA).[3] DISA routers would not be able to examine USAF packets for compliance with an SLA because the USAF virtual private network (VPN) encrypts USAF-to-USAF packets.

Even if DISA were to implement a QoS technique within the NIPRNET, the USAF would still be susceptible to DDoS attacks as described earlier. A NIPRNET QoS scheme would not prevent the additional latency caused by multiple proxying of USAF applications at base network boundaries.

[3]Network Working Group, RFC 2475, "An Architecture for Differentiated Services," Dec. 1998, 12, [On-line]. URL: http://www.ietf.org/rfc/rfc2475.txt (Accessed on 1 March 2003.)

Chapter Five

Wrap-up: Why the Air Force Should Just "Do It"

The current USAF unclassified network architecture does not support end-to-end QoS of enterprise applications. As shown in this paper, network users at one USAF base must connect through a maze of network equipment and processes to obtain information from a server located at another USAF base or DISA location. Each step along the way adds a delay to the connection time (latency) and is subject to packet loss.

Network traffic is "bursty" and results in congestion in the NIPRNET. This congestion results in packet loss and subsequent slowing of TCP traffic. The USAF has no way to influence what traffic is slowed in the NIPRNET due to congestion. There are techniques to provide end-to-end QoS for network traffic. However, since the USAF does not control all of the devices that implement this end-to-end QoS capability, the USAF cannot utilize the full potential of its enterprise applications. In a COIN architecture, the USAF could implement QoS techniques for its mission-critical enterprise applications because the USAF would control all the routers between USAF users and the enterprise application servers.

The current USAF unclassified network architecture leaves the USAF susceptible to DDoS attacks. A COIN architecture would allow internal USAF enterprise applications to communicate unimpaired by a DDoS attack or other hostile event such as a computer worm because these types of events could be detected and blocked at USAF enterprise gateways.

At some point in the future, DISA may implement a QoS technique in the NIPRNET. The USAF should support such an effort, because a USAF COIN architecture would still use the NIPRNET to provide Internet access for USAF users. A NIPRNET with QoS-enabled routers could guarantee that USAF mission-critical applications received priority treatment from a USAF SDP router at one base to an SDP router at another base. However, unless the USAF base perimeter architecture is changed, this mission-critical traffic would still be subject to the delays caused by proxy devices and to potential loss of availability caused by multiple singe points of failure in series at each boundary.

Applications have been developed that are sensitive to IP network latency, jitter, and packet loss. The NIPRNET does not support QoS for these applications. DISA is willing to sell survivable layer 2 bandwidth to the USAF, which would allow the USAF to control the end-to-end QoS of its enterprise applications. The USAF needs to seize this opportunity.

Appendix

Everything You Always Wanted to Know About How the Internet Works but Were Afraid to Ask

The Internet is a fascinating thing. Originally designed under contract to the Advanced Research Projects Agency (ARPA)—now called the Defense Advanced Research Projects Agency (DARPA)—as an experiment to see if computers could "talk" to each other in a survivable network, it has grown to phenomenal significance in the way Americans live their lives. The September 2002 Netcraft Survey found that there are over 35 billion Web sites currently on line.[1] According to an e-commerce projection by International Data Corporation (IDC), the value of business-to-business transactions over the Internet is expected to reach $4.3 trillion by 2005.[2]

But what is the Internet and how does it work? In October 1995 the Federal Networking Council passed a resolution defining the Internet. It stated:

> The Federal Networking Council (FNC) agrees that the following language reflects our definition of the term "Internet." "Internet" refers to the global information system that – (i) is logically linked together by a globally unique address space based on the Internet Protocol (IP) or its subsequent extensions/follow-ons; (ii) is able to support communications using the Transmission Control Protocol/Internet Protocol (TCP/IP) suite or its subsequent extensions/follow-ons, and/or other IP-compatible protocols; and (iii) provides, uses or makes accessible, either publicly or privately, high level services layered on the communications and related infrastructure described herein.[3]

The keys to the Internet are packet switching and a layered set of protocols with standards for moving up or down the layers. To understand the Internet it is helpful to understand the Open Systems Interconnect (OSI) reference model of the seven different "layers" used to describe how computer to computer communications are broken up into independent processes. The following sections give a rudimentary explanation of the OSI Reference Model and a "physical world" analogy to try to explain how information is sent in the "cyber world." Following the example are discussions of the "bursty" nature of Internet traffic and of methods to provide QoS in the Internet.

---

[1]Netcraft Web Server Survey, October 2002, [On-line]. URL: http://www.serverwatch.com/news/article.php/1475371 (Accessed on 29 Oct. 2002.)

[2]CyberAtlas, "B2B E-Commerce Headed for Trillions," [On-line]. URL: http://cyberatlas.internet.com/markets/b2b/article/0,,10091_986661,00.html (Accessed on 29 Oct. 2002.)

[3]The Internet Society, "A Brief History of the Internet" (Reston, Va.: The Internet Society), [On-line]. URL: http://www.isoc.org/internet/history/brief.shtml (Accessed on 6 Nov. 2002.)

A.1 The OSI Reference Model

The OSI Reference Model was developed to describe how an application on one networked computer communicates with the application on another networked computer. The theoretical model identifies seven layers: Physical, Data Link, Network, Transport, Session, Presentation, and Application. The OSI layers that must be understood in order to grasp the importance of using a private network to ensure QoS of network traffic and enhancing enterprise security are Data Link (layer 2), Network (layer 3) and Transport (layer 4). (NOTE: The hard-core computer scientist reader will be quick to point out that the Internet is actually built around the TCP/IP protocol suite which is a four layer model. The OSI reference model is used in this paper because the author believes it provides a better educational tool for understanding how computers "talk" to each other).

The important thing to remember is that this model represents layers of protocols. Each layer builds on the one below it and passes information to the one above it and vice versa. This way, if there are improvements or other changes in one layer, the other layers do not have to change as long as the standards for passing information from one layer to the next remain the same.

Figure A-1 is a diagram of the OSI model. The model is often referred to as the "protocol stack." As information passes from a sender on one node of a network it moves "down the stack" and then back "up the stack" at the receiving node of the network. At each layer different protocol headers and trailers are added to the application data as it moves down the stack and then stripped off the data as it moves up the stack toward the application at the other end.[4]

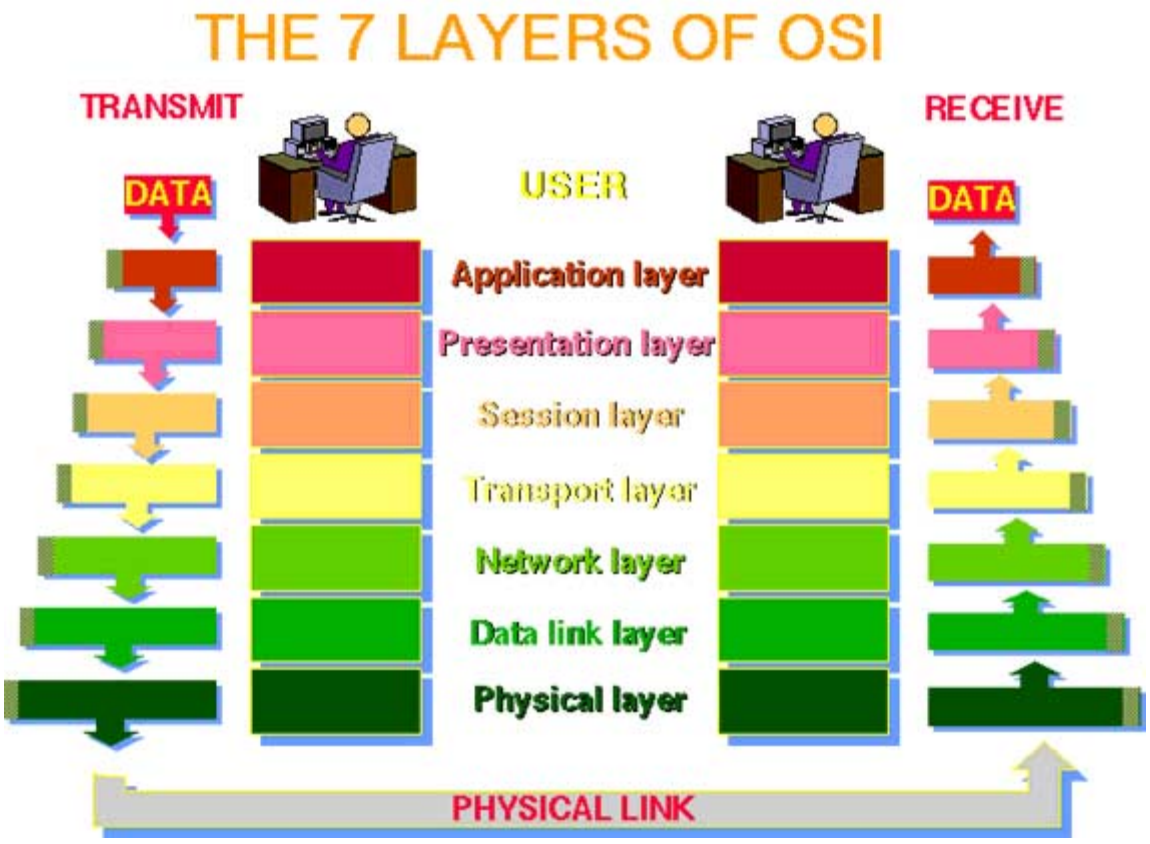
A.1.1 Layer 1: The Physical Layer

The physical layer refers to media used to transfer data over some distance. Category 5 (Cat-5) copper wire is often used to connect computers to an Ethernet local area network (LAN) using RJ-45 connectors, which look like standard phone jacks but are slightly wider.[5] The computer's network interface card (NIC) applies a voltage change to the LAN wire to indicate a one or zero. Other types of media used to support the physical movement of data in networks are fiber optic cable, coaxial cable, and the radio waves used for wireless LANs.

---

[4]Neil Briscoe, "Understanding the OSI 7-Layer Model," *PC Network Advisor*, No. 120, July 2000, 13–14, [On-line]. URL: http://www.itp-journals.com/nasample/t04124.pdf (Accessed on 22 Feb. 2003.)

[5]Webopedia [On-line]. URL: www.webopedia.com/TERM/C/Cat_5.htm



Source: http://webopedia.internet.com/quick_ref/OSI_layers.asp (Original taken from The Abdus Salam International Centre for Theoretical Physics)

Figure A-1

OSI Reference Model

A.1.2 Layer 2: The Data Link Layer

The data link layer determines how two machines on a network "talk" to one another. The most common data link in use is Ethernet (IEEE 802.3).[6] Standard Ethernet speeds are 10 million bits per second (10 Mbps), 100 million bits per second (100 Mbps, also known as "fast Ethernet") and 1 gigabit per second (1 Gbps, also known as "Gig-E"). The Ethernet standard uses carrier-sense multiple-access with collision detection (CSMA/CD). This means that all machines on an Ethernet segment share a common medium. Each machine "listens" to the network segment for any Ethernet message (called a "frame") that is addressed to its physical machine address—also called its Media Access Control (MAC) address. Each network interface card has a unique 48-bit MAC address programmed into it by the manufacturer. If a machine detects a frame that has its

[6] Julian Moss, "Understanding TCP/IP," *PC Network Advisor*, No. 87, September 1997, 3 [On-line]. URL: http://www.itp-journals.com/nasample/c04100.pdf (Accessed on 22 Feb. 2003.)

MAC address, it will extract the data from that frame and pass it up to the next protocol layer. If an Ethernet frame is transmitted that has the broadcast MAC address (all "1"s or, written in hexadecimal, FF:FF:FF:FF:FF:FF), all the nodes on the segment will process the data payload. If a node needs to transmit data, it will listen to the network to determine when no other node is transmitting and then send the frame out on the LAN segment. If two nodes on the same network segment transmit at the same time, a collision occurs and each node will detect the collision, wait a random time (very short), and then retransmit.[7]

In the past, it was common for several machines to share a single Ethernet segment, which meant that all the machines on that segment shared the bandwidth of that segment. Advances in computer technology have led to full-duplex layer 2 network switches that allow each machine to connect directly to the network switch. In this configuration, each Ethernet segment contains only one node. Each node gets full use of that segment bandwidth and collisions do not occur.

Ethernet frames consist of data and headers. Ethernet frames can range in size from 64 to 1518 bytes. The maximum data payload of an Ethernet frame is 1492 bytes. This is called the maximum transmission unit (MTU) for Ethernet. Higher layer protocols must fragment any data payload exceeding this length.[8]

A.1.3 Layer 2 Devices—Hubs, Bridges, and Switches

A hub is a device used to connect network devices to a LAN segment. If an Ethernet frame enters one port on a hub it is simultaneously transmitted to all other ports on the hub.[9] A bridge is a device used to connect different LANs or different segments from the same LAN. A layer 2 network switch is used to connect many LAN segments together. Ethernet switches read the destination MAC address from the Ethernet frame header. If the MAC address is for a device on the same network segment of the switch port that read the header, the frame is ignored because the destination device will be "listening" on that segment for frames containing its MAC address in the header. If the MAC address belongs to a device on a different network segment, the switch will relay the entire frame to the switch port connected to the network segment where the destination device is located.

A.1.4 Layer 3: The Network Layer

The Network layer is the heart of the Internet. This layer deals with Internet Protocol (IP) datagrams. An IP version 4 (IPv4) datagram consists of a block of data and a header, as shown in **Figure A-2**. (IP version 6 is discussed later in this section.)

---

[7]Martin C. Libicki, *Information Technology Standards* (Boston, Mass.: Digital Press, 1995), 79.

[8]*PC Network Advisor,* Issue 87, 3.

[9]Webopedia [On-line]. URL: http://www.webopedia.com/TERM/h/hub.html

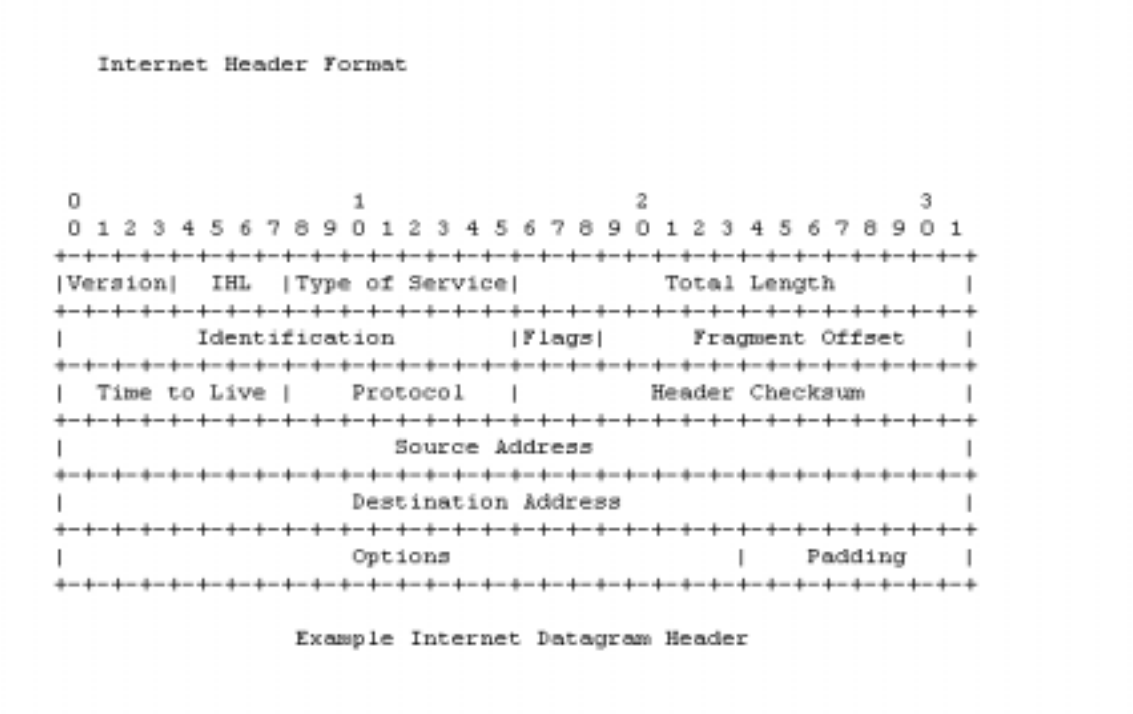


Figure A-2

IP V4 Header

The IP Specification, RFC 791, describes the purpose of the IP layer (equivalent to the Network layer in the OSI model). IP provides addressing and fragmentation of blocks of data called datagrams. IP accepts data from a layer 4 protocol and sends it out onto a network (layer 2) in the form of packets. If the datagram is too large to send directly down to the layer 2 protocol, IP fragments the datagram into two or more packets.[10] Therefore, in a "packet switched network" the data portion of packets moving from node to node can contain an entire IP datagram or fragments of a datagram. The network treats each packet as an independent entity.

The IP header contains some important fields, including the 32-bit source and destination IP addresses, an 8-bit Time to Live (TTL) field, and an 8-bit Type of Service (ToS) field. Since the original publication of the IP specification, the ToS field has been re-designated as the "Differentiated Services" (DS, or Diff-Serv) field, and the first six bits of this field are called the Differentiated Services Code Point (DSCP). The remaining two bits of the original ToS field are currently unused.

The purpose of IP is to move packets from the source IP address to the destination IP address, but it is important to remember that there are certain functions it does not perform. IP

[10]Network Working Group, RFC 1122, "Requirements for Internet Hosts—Communications Layers," October 1989 [On-line]. URL:. http://www.ietf.org/rfc/rfc1122.txt (Accessed on 1 March 2003.)

does not provide any means to guarantee delivery. However, the Diff-Serv field can be used to change the probability that a packet will be dropped, as discussed later. Packets are not acknowledged either end-to-end or hop-by-hop. There is no error control in the IP header for the data payload, but there is a checksum in the header to determine if the header has been damaged. There is no flow control in IP—the Network layer—of the OSI model. Internet Control Message Protocol (ICMP) is used at this layer to report errors.[11]

IP addressing is used to identify network nodes or hosts. The IP address is the logical address of each device on a network. It is assigned to each device by the network system administrator from IP addresses that are globally managed by the Internet Corporation for Assigned Names and Numbers (ICANN). IP addressing is complex, and a detailed discussion is beyond the scope of this paper. The important thing to remember is that IP addresses are used to identify network hosts within the Internet. Therefore, for security reasons, the IP address is used to identify the originating organization of network traffic (since IP addresses are assigned to organizations).

The IP address is an important field in the IP header that is used to filter out or drop unwanted packets. It is possible for a network user to "spoof" or replace a source IP address in a packet to make it appear that it comes from a particular location when, in fact, the packet originated elsewhere. Many companies and ISPs filter IP packets at their network boundaries to check if they contain invalid IP addresses in the packet header. They drop inbound packets with source IP addresses from the range of addresses assigned to their internal network and they drop outbound packets that do not have a source IP address from their assigned block of addresses.

The TTL field is used to prevent an IP packet from getting caught in a loop and traveling around the Internet forever. Each time a packet makes a hop (moves from one layer 3 device to another), the TTL field is decremented by one. When the TTL field reaches zero, the packet is discarded and an ICMP error message is sent to the source IP address. This field is also used in an interesting manner in the "traceroute" utility to map the route between two network devices.[12]

IP version 6 (IPv6) was developed because the explosive growth of the Internet will eventually result in a lack of available IPv4 addresses. IPv6 uses a 128-bit IP address instead of a 32-bit address. IPv4 has over 4.2 billion possible addresses (2^{32}); however, because large blocks of numbers are assigned to organizations that are not using them, there are far fewer actual numbers available for new network node addresses. IPv6 has over $3 * 10^{38}$ possible addresses

---

[11]Network Working Group, RFC 791, "Internet Protocol Specification," September 1981, 2, [On-line]. URL:. http://www.ietf.org/rfc/rfc791.txt (Accessed on 1 March 2003.)

[12]In traceroute, the utility starts by sending a packet addressed to a nonexistent port at the destination IP address with the TTL field set to 1. When the first router gets the packet, it discards it and sends a "time-to-live equals 0" error message back to the sending IP address. Then traceroute sends a packet with a TTL set to 2 and the second router sends the TTL error message. This continues until the packet reaches the target host and returns an ICMP message with a "destination port unreachable" error message. See *PC Network Advisor* Issue 87, "Understanding TCP/IP," 5.

(2^{128}). Put another way, IPv6 provides over $6 * 10^{23}$ possible IP addresses for every square meter on the face of the earth.[13] In practice, IPv6 is being deployed to recognize 2^{64} different networks.

The IPv6 header is different from the IPv4 header in that multiple headers can be used. The first header contains the minimum required fields, including the version, priority, flow label, payload length, and hop limit as well as the source and destination IP addresses. It also includes a field to indicate if another header is included in the packet.[14]

A.1.5 Layer 3 Devices—Routers

Diagrams often represent the Internet as a cloud. This cloud represents many interconnected network routers. Routers are specialized network devices that are used to move packets along from their source to their destination IP address. A router has three functions. First, the router needs to determine what the next hop should be for the best route to send packets so that they reach the destination IP addresses. Second, the router has to forward packets from an input network interface to an output network interface. Third, a router needs to be able to buffer or temporarily store packets and possibly queue packets using a priority-weighted scheme.[15] Thus, routers are store-and-forward devices. They receive packets, briefly store them, then forward them on to the next network hop or final destination.

Routers use several types of protocols to communicate with neighboring routers in order to develop a routing database, called a routing table, which is used to send packets across optimum paths in a network. Some common protocols include Border Gateway Protocol (BGP), Open Shortest Path First (OSPF), Intermediate System-to-Intermediate System (IS-IS), and Routing Information Protocol (RIP).[16]

Routers use a forwarding engine to move packets from one network interface of the router to another. When a packet arrives, the forwarding engine consults the routing table and sends the packet to the appropriate output interface.[17]

Routers must also buffer packets. If many packets simultaneously arrive at several router interfaces and the routing table indicates they should all go to the same output interface, they

---

[13]David Morton, "Understanding IPV6," May 1997, *PC Network Advisor*, Issue 83, 18, [On-line]. URL: http://www.itp-journals.com/nasample/c0655.pdf (Accessed on 1 March 2003.)

[14]Ibid., 20.

[15]Cisco Systems, "The Evolution of High-End Router Architectures," White Paper (San Jose, Calif.: Cisco Systems, no date, downloaded on 18 November 2002, from: http://www.cisco.com/en/US/products/hw/routers/ps167/products_white_paper09186a0080091fdf.shtml), 2.

[16]Ibid.

[17]Ibid, 3.

must be queued until they can be transmitted by the output network interface.[18] If the number of arriving packets exceeds the size of the router buffer needed to temporarily store the packets, the router will discard some packets using pre-set router configuration settings to decide which packets will be dropped.

There are several ways routers can control queues to manage packet loss.[19] The simplest congestion management technique is First In First Out (FIFO) queuing. This is also known as first-come, first-serve queuing because there is no analysis of the packet types. In FIFO queuing under default configurations, all packets are treated the same way and are buffered and forwarded in order of arrival. If the buffer fills, all arriving packets are dropped until there is sufficient room in the buffer to accept more traffic.[20] Techniques used by routers to selectively drop lower priority packets or give precedence to higher priority packets are discussed later.

A.1.6 Layer 4: The Transport Layer

Two Transport layer protocols ride on top of IP to send information on the Internet: the User Datagram Protocol (UDP) and Transmission Control Protocol (TCP). Like IP, the Transport layer protocols consist of data and headers. In fact, the Transport layer header and data are carried across the network in the data portion of the IP layer packet. This is a key aspect of the Internet. The data portion of each layer of the OSI stack contains the header and the data (or a fragment of the data) from the layer above it.

An important element of the Transport layer protocol is the port address field. This 16-bit field or port number is not related at all to the physical switch or hub ports discussed in section **A.1.3**. They are not physical connections to a network. Layer 4 port numbers are logical software addresses used to pass data to and from the higher layer protocols. The lower port numbers (0–1023) are called “well known ports” and are reserved for common Internet protocols.[21] For example, port 25 is used for Simple Mail Transfer Protocol (SMTP) and port 80 is used for Hyper-Text Transfer Protocol (HTTP—the protocol used in the World Wide Web [WWW]). These common-use numbers are assigned by the Internet Assigned Numbers Authority (IANA). Developers can write software so that their applications “listen” to the transport layer for data arriving on any layer 4 port.

UDP is a connectionless protocol that, like IP, does not guarantee delivery of data. UDP packets (datagrams) consist of application layer data and a header comprising four 16-bit fields: a

---

[18]Ibid.

[19]Cisco Systems, “Congestion Management Overview,” White Paper, no date, downloaded on 18 November 2002 from: http://www.cisco.com/univercd/cc/td/doc/product/software/ios122/122cgcr/fqos_c/fqcprt2/qcfconmg.htm

[20]Ibid.

[21]RFC 1122, 81.

length field, checksum field, and source and destination port fields. UDP is a good choice of Transport layer protocol for those applications that have no need to establish a connection to ensure complete receipt and proper ordering of the application data. Multimedia (audio and video) data fits this category. UDP is used in real-time multimedia applications where it is better to ignore late-arriving and missing packets than to attempt to insert the packets in the proper sequence.[22]

TCP is the protocol that provides a reliable data stream for the Internet.[23] TCP is connection oriented. That means that if a node on a network wishes to exchange information with another node using TCP, the two machines must first establish a connection (explained later in this section). TCP guarantees that all data has been exchanged without errors and that it is in the proper sequence. It also implements a flow control process to avoid overwhelming a receiving node and to react to periods of network congestion by reducing the amount of data sent out on the network. Most Internet applications, such as e-mail, file transfers, and Web servers, use TCP to exchange data.[24] Since TCP is the most common way of communicating on the Internet, it is explained in more detail below.

Figure A-3 is a diagram of the TCP header. The 16-bit source and destination port numbers are used to pass data to and from higher layer protocols in the same way as in UDP. The combination of the port number and a node IP address from the IP header uniquely identifies a service running on that node and defines a connection socket.[25] The sequence number and acknowledgment number fields are used to keep track of application data that is sent and received. They represent blocks of data to be exchanged and are expressed as numbers of octets (8 bits) of data. The data offset field is used to indicate where the data begins. The reserved field is not used and must be set to 0. The six control flag bits (URG, ACK, PSH, RST, SYN, FIN) are used as shown in **Table A-1**. The window field is used to control the flow of data during the connection. The checksum field is used to verify that no errors occurred during transmission.[26]

For TCP communication to occur, one host (the server) must be “listening” for incoming connections on a given port number. The client host sends a TCP SYN segment to the server with its destination port number the same as the one that the server is “listening” to. The client source port number is set by a higher layer protocol. The SYN flag is set and the client’s initial sequence number is given.

---

[22] *PC Network Advisor*, Issue 88, 13.

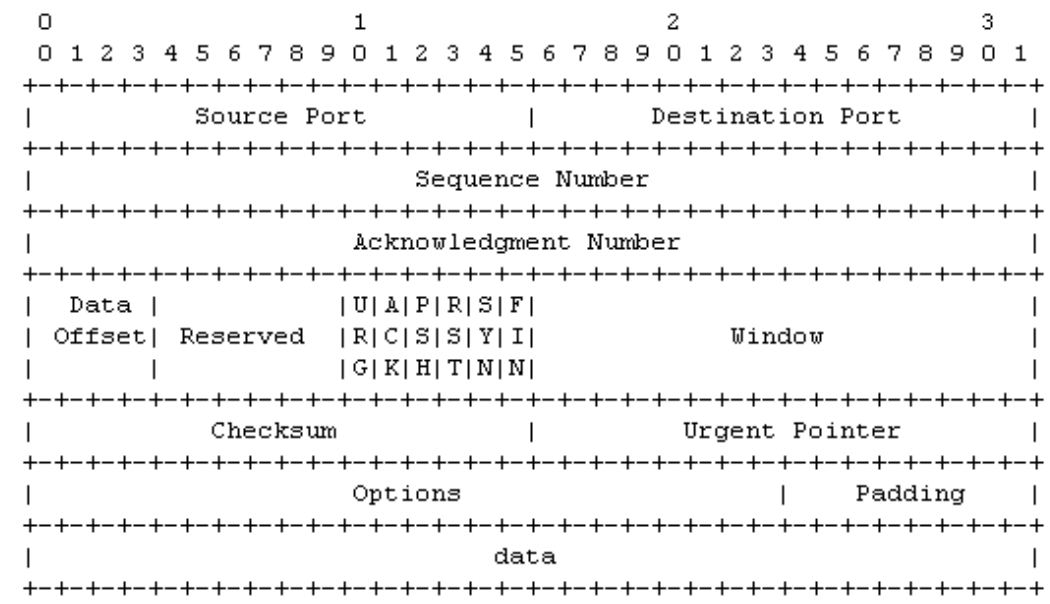
[23] RFC 1122, 81.

[24] Julian Moss, “Understanding TCP/IP,” *PC Network Advisor*, Issue 88, October 1997, 14 [On-line]. URL: http://www.itp-journals.com/nasample/c04100-2.pdf (Accessed on 1 March 2003.)

[25] Ibid.

[26] Network Working Group, RFC 793, “Transmission Control Protocol Specification,” Sept. 1981, 14—15 [On-line]. URL: http://www.ietf.org/rfc/rfc793.txt (Accessed on 1 March 2003.)

TCP Header Format



TCP Header Format

Note that one tick mark represents one bit position.

Figure A-3

TCP Header

Table A-1

TCP Header Flags

URG	URGENT pointer field significant
ACK	Acknowledgement field significant
PSH	Push Function
RST	Reset the connection
SYN	Synchronize sequence numbers
FIN	No more data from sender

Source: RFC 793

When the server receives the SYN segment it returns a segment with its own initial sequence number for its return data, destination port number set to the client's source port number, both SYN and ACK flags set, a window size set to the amount of data (octets) it is willing to receive (usually based on the available buffer size at the server), and an acknowledgment number set to the client's initial sequence number plus one, indicating which segment it is ready to receive. The client confirms the connection by sending a segment with the ACK flag set and the acknowledgment field containing the server's sequence number plus one, indicating it is ready to receive the next segment of data from the server.

The client and server then begin sending and receiving segments containing data. They acknowledge data received by sending segments with the ACK flag set and containing the sender's sequence number plus one to indicate the sequence they are expecting next. If the receiver has application data to send to the sender, the acknowledgment of received data (ACK) and data to be sent in the other direction can be sent in the same TCP segment to improve efficiency. If the receiver has no application data to send, the ACK will be sent to the sender by itself. The two nodes continue sending, receiving, and acknowledging segments until all data is sent in both directions or an error occurs. If the connection closes normally, one side will send its last segment of data with the FIN flag set. The other side will then complete sending its data and send its final segment with the FIN and ACK flags set. The connection is then closed.[27]

Many things can go wrong in IP networks. Packets can and do get dropped (remember, IP does not guarantee delivery) or arrive out of order. Network hosts can crash, data can become corrupted in transit, or the path between the two nodes may become unusable for any number of reasons. The TCP specification allows for retransmitting segments if they are not acknowledged in a timely manner or are received with errors. TCP can also reset the connection by using the RST flag to attempt to recover from network errors. As a last resort, the TCP protocol will give up and close the connection on both ends via a time-out feature.

Perhaps the most significant attribute of the TCP protocol is the flow control provided by the TCP window. As previously mentioned, the TCP window value is used by the receiving node to tell the sending node how much data it is willing to accept. The sender uses this window value to adjust the amount of data sent in subsequent segments (and adjust the sending sequence number accordingly).

The TCP window is a sliding window in both directions. This means that the TCP window can grow (send more data) or shrink (send less data) based on changing network conditions. Both sides of a TCP connection must keep track of what segments have been sent, received, and acknowledged (ACK), what segments have been sent but not acknowledged, and what segments are waiting to be sent. Both sides use the sliding TCP window values and changing round-trip-times (RTT), determined by keeping track of sent and acknowledged segment numbers, to adjust

[27] *PC Network Advisor*, No. 88, 15.

their flow of data onto the network. Usually, applications will start off slowly (called “TCP Slow Start”) and increase their TCP window (increase data rate) in a linear fashion until either all data are transferred or an error occurs. If the error is a full buffer at the receiver a TCP window value of zero will be returned to tell the sender to wait before sending any more data. If the error is a failure to receive an ACK after an expected RTT, the sender will usually resend the segment and cut its data flow rate in half. This is done because the most probable cause of this error is a packet dropped due to network congestion.[28]

Because TCP can adjust its flow of data into a network, it automatically responds to changing network loading conditions. For example, during periods of congestion, TCP applications decrease network loading in reaction to lost packets. It is important to note that this is not the case for applications that use UDP. These applications continue to send data at a rate dictated by the application software even if the network is congested.[29]

A.1.7 Layers 5, 6 and 7: The Session, Presentation, and Application Layers

The Session layer uses start, hold, and stop messages to control connections. It creates message logs and ensures continued throughput of data. The Presentation layer is where data is made ready for end-user applications. It is used to pack and unpack data, convert data to a common language (such as ASCII [American Standard Code for Information Interchange]), and conduct encryption and decryption. The Application layer organizes information for specific applications such as e-mail, file transfer, et cetera.[30] Application developers must include the functions of the session and presentation layers in their application, so it is very common when discussing TCP/IP networks to refer to the top three layers of the stack simply as the “application layer.”

One way to look at the OSI layers is to view them as three super layers. The first two are concerned with the physical movement of bits from one machine to another. The middle two are concerned with moving information from one network to another. The upper three are concerned with exchanging information between and among applications.[31]

A.2 A Simple “Physical World” Example

Suppose an executive from Acme Corporation, Mr. Smith, in Los Angeles (LA) wants to send a 100-page document to another Acme executive, Ms. Jones, in New York (NY). Mr. Smith

---

[28]Ibid., 16.

[29]Mark A. Parris, “Class-Based Thresholds: Lightweight Active Router-Queue Management for Multimedia Networking,” doctoral dissertation, University of North Carolina (Chapel Hill, N.C.: University of North Carolina Press), 1.

[30]Libicki, 78.

[31]Ibid.

gives the document to his assistant, Sue, but she only has small envelopes that can hold 25 pages. To send the document via overnight courier, she breaks up the document and puts the pieces in four different envelopes. On the front of each envelope she writes whom the envelope is for (Ms. Jones), whom it is from (Mr. Smith), and which pages of the document are included in each envelope (pages 1–25, 26–50, et cetera). She calls Ms. Jones's assistant, Jim, in NY and tells him to expect the envelopes.

Sue then sends all four envelopes to the Acme mail room. The mail room worker takes each envelope and puts it in a separate overnight courier package because the mail room only has packages large enough to hold one envelope. On the front of each package, the worker writes where the package is from (Acme Corporation address, LA) and where it needs to go (Acme corporation address, NY).

It just so happens that the overnight courier truck arrives at the LA Acme office when only two of the four packages are ready. The courier driver picks up the two packages along with all the other overnight delivery packages and takes them to the LA airport. All the packages in the truck are dumped and sorted by destinations. The two packages for the Acme NY office are put on a plane to Memphis with all the other East Coast-bound packages, because the overnight company does not have a direct flight to New York. Once the two packages reach Memphis, they are sorted by destination and the two for the Acme NY office are put on a plane for New York with other New York-bound packages.

Meanwhile, another courier company driver is making his afternoon run, stops by the Acme mail room, and picks up the remaining two packages for the Acme NY office along with any other overnight packages. After the remaining packages reach the airport, they are put on a plane for St Louis, because the flight for Memphis has already gone and the courier company knows that the packages will be able to take a New York-bound flight out of St Louis later that evening.

Eventually, both planes land in New York and the packages are taken to the courier company's holding area at JFK airport. Early in the morning the packages are sorted and put in bins according to their destinations and loaded on a company truck. The bin destined for the Acme NY office is very full and the third package (containing pages 51–75 of the document) is on top of the stack. As the courier company driver heads toward the Acme office, honking the horn and slamming on the brakes, the third package slides off the bin and falls between the bin shelves and the wall of the truck, out of view.

The Acme NY mail room opens all the courier company packages, notices that three envelopes are supposed to go to Ms. Jones, and sends them up to her assistant. Jim looks at the envelopes, notices they are out of order, sorts them, and realizes he is missing the envelope with pages 51–75. He calls Sue in LA and leaves a voice mail (the work day has not started in LA) asking her to re-send the third envelope. Sue gets the voice mail and re-sends the third envelope, which goes through the same process as the original third envelope.

The next day, this envelope is delivered to Jim. He then opens the envelopes, puts the document back together, and gives it to Ms. Jones, who is happy to get the entire document, in order, even if it took longer than it should have. Jim then calls Sue and tells her everything is okay.

This is a very simple physical example of independent processes working together to move information from one location to another.

A.3 Tying It All Together

The previous sections described the OSI model and gave a physical example of a layered process to move information. The following analysis may help to clarify the importance of layered protocols.

A.3.1 Identifying the Layers

In the above example, the executives (Smith and Jones) represent the top three layers of the stack. They know what the information means and how to use it. Jim and Sue represent the Transport layer. They package the information and make sure all of it gets from one user to the other in the correct order. The mail rooms represent the Network layer nodes. They take information from the Transport layer package it and put their “node address” on the outside. The courier company airport hubs represent intermediate network nodes (routers). They don’t particularly care what order the packages are in or whom they are for; they just want to move packages from one courier company hub (IP address) to the other. The overnight courier trucks and planes represent the Data Link and Physical layers. Their job is to move packages from one physical location to the next. They are only concerned with the next hop in the chain, not the final destination.

A.3.2 Improvements and Disasters

Suppose the overnight courier company bought bigger, faster airplanes. This would be transparent to the Acme Company. The result would be that the packages *might* arrive sooner at Ms. Jones’ office. Of course, it would depend on the capacity of the courier’s JFK airport office and its schedule of truck deliveries to the Acme NY office. This is analogous to increasing bandwidth between routers in a large network. The end-to-end delay (network latency) may or may not decrease. It just depends on what is happening in the rest of the network. Perhaps the newer, faster planes are filled with packages from other companies and there is no room for the Acme packages. In this case, the total capacity of the transportation network increases, but there is no way to tell if the information will reach the other end any faster, unless Acme controls what packages get on what planes. The point is that good things can happen at the lower layers of the network stack (for instance, increasing bandwidth) that are transparent to the Application layer,

and the result may or may not be an increase in Application layer performance. It depends on network traffic conditions.

On the other hand, suppose something bad happens in the lower layers of the stack. Suppose the Memphis airport is closed due to weather. The Acme Company may see a delay in receiving its packages if the courier company cannot route all the packages through St Louis. This is analogous to a network router going down in a packet-switched network. The traffic gets re-routed and may or may not take longer to get to its destination. Again, it just depends on the network traffic.

Suppose something really drastic happens in the lower layers. Say that all the courier company planes are grounded by the Federal Aviation Administration. This would *not* be transparent to the Acme Company. It would take days for Ms. Jones to get her information, if she gets it at all.

Disasters that occur at the lower layers of the protocol stack can and do impact Application layer communications in the Internet. The Internet is a hostile place. For example, on 2 September 2000 a DDoS attack was launched against Yahoo! Inc. A DDoS attack occurs when hundreds of computer hosts (called "zombies") start sending many packets to a targeted host. The zombies are hosts that have been hacked by a "master" and await instructions to launch various types of attacks. The line in **Figure A-4** shows the Round Trip Time (RTT) and the bars show packet loss at the Yahoo! server during the attack. The figure shows that for a four-hour period the Yahoo server had close to 100 percent packet loss rate. Yahoo! was out of business during this period.

Source: www.visualware.com/visualpulse/yahoo.html

Figure A-4

DDoS Attack Against Yahoo!

As recently as 21 October 2002, another DDoS attack was launched against a core element of the Internet. Unknown attackers launched a "ping flood" attack against the 13 root-level Domain Name Servers that function as the authoritative directory assistance for the Internet. Luckily, the sheer size and redundancy of the Internet enabled it to absorb the flood of traffic until the packets could be blocked by cooperating ISPs.[32]

Another indication of the hostile nature of the Internet is the number of computer incidents reported to the CERT at Carnegie Mellon University. The CERT Coordination Center (CERT/CC) received 21,756 incident reports in 2000, and 52,658 reports in 2001. During the first three quarters of 2002 they received 73,359 incident reports.[33]

These figures emphasize the point that the Internet and all IP networks represent "best effort" networks. Every effort will be made to move packets from one location to another but no delivery guarantee is given—or possible. Packets can be lost due to hostile events, network congestion, human error, or other reasons. ISPs can make service-level agreements with their customers to provide a guaranteed level of service only within their own networks. Once they pass traffic outside of their network, there are no guarantees.

A.4 The Nature of Network Traffic

Network traffic is unpredictable. As explained earlier, all hosts on a 10 Mbps Ethernet LAN segment cannot send 10 Mbps worth of data at the same time, because they share the 10 Mbps capability of that network segment. The actual data transfer rate between any two hosts depends on how much the other hosts on the network need to "talk." If all the hosts need to talk a lot, the effective data throughput rate between any two hosts will be lower than if they were the only two hosts that needed to communicate.

A.4.1 Self-Similar or "Bursty"

Computers tend to communicate in bursts. Will E. Leland, in his paper *On the Self-Similar Nature of Ethernet Traffic*, analyzed the nature of computer network traffic and described the distribution of packets as a function of time as "self-similar." [34] This means that the distribution diagrams of packets per time unit appeared very similar, no matter what time scale he chose to plot. The plot of packets per minute looked similar to the plot of packets per hour which looked similar to the plot of packets per second which looked similar to the plot of packets per

---

[32] *Business Week* On-line, "The Day the Net Nearly Choked," [On-line]. URL: www.businessweek.com/technology/content/oct2002/tc20021030_3147.htm (Accessed on 30 Oct. 2002).

[33] Carnegie Mellon Software Engineering Institute, "CERT/CC Statistics 1988-2002" (Pittsburgh, Pa.: Software Engineering Institute, 2002), [On-line]. URL: www.cert.org/stats/cert_stats.html (Accessed on 6 Nov. 2002).

[34] Will E. Leland, "On the Self-Similar Nature of Ethernet Traffic (Extended Version)," *IEE/ACM Transactions on Networking* **2**, 1 (February 1994).

millisecond. All of the plots showed spikes and valleys in the number of packets per time unit. Leland noted that "at every time scale ranging from milliseconds to minutes and hours, bursts consist of bursty subperiods separated by less bursty subperiods."[35]

David Clark, William Lehr, and Ian Liu of MIT described the bursty nature of Internet traffic as follows:

> A typical traffic source on the Internet does not generate data at a constant rate, but is very bursty in nature. A person cruising the Web, for example, alternates between transferring a page and looking at it. User satisfaction is increased if each active transfer occurs as quickly as possible, so the resulting network load is a series of very high bandwidth bursts intermixed with relatively long periods of silence. Other activities generate bursty traffic patterns—retrieving a succession of mail messages, or interacting in a chat room.[36]

This discussion indicates that there is a high ratio between peak network load and average network load. In determining the amount of bandwidth to provide for network users, the bursty nature of network traffic needs to be considered. If the network links are sized to handle the possible aggregate peak load of all the users on the network, the resulting trunk utilization of that bandwidth will be unacceptably small.[37]

A.4.2 Congestion in the Internet

The MIT paper provides the following amplification regarding congestion in the Internet:

> The statistical nature of traffic aggregation means that there is no guarantee that there is enough capacity in the Internet to carry all the offered load at any instant. Some level of congestion is to be expected, especially during peak periods of usage. The design of the Internet deals with congestion in a straight-forward manner—when congestion is detected, the sources of traffic are expected to slow down, and when there is no congestion, they are permitted to speed up. (In fact, given this rule, some degree of congestion is the norm, since sources will speed up until congestion slows them down.) The implication of this approach to congestion is that the actual rate that any source can achieve at any instant is unpredictable.[38]

---

[35]Ibid., 4.

[36]David Clark, William Lehr, and Ian Liu, "Provisioning for Bursty Internet Traffic: Implications for Industry and Internet Structure," MIT Workshop on Internet Quality of Service, November 1999, 2–3, [[On-line]. URL: http://www.ana.lcs.mit.edu/anaweb/PDF/ISQE_112399_web.pdf (Accessed on 1 March 2003.)

[37]Ibid., 3.

[38]Ibid.

Since the NIPRNET is an IP-routed network like the Internet, it is reasonable to extrapolate the above discussion to the NIPRNET. Congestion occurs in the NIPRNET, it results in packet loss, and this is to be expected. When provisioning bandwidth for an USAF enterprise network, the architecture of the enterprise and the methodology for dealing with network congestion are critical issues. **Chapter Three** discussed how these issues affect the current USAF unclassified network.

The Cooperative Association of Internet Data Analysis (CAIDA) conducted a study of Internet traffic at a large internet gateway from May 1999 to March 2000. The study found that approximately 85 percent of the traffic was TCP (mostly HTTP and FTP). The remaining traffic was mostly UDP, with a very small percentage of other traffic such as ICMP.[39]

A.5 Network Delay (Latency)

Network delay is the time it takes for information to travel from a sender to a receiver. In circuit-switched networks there is a short delay for call set-up and then the sender and receiver are connected (if network resources are available and the receiver is not “busy”). Most of the delay in the ensuing conversation is caused by propagation of the signal over a given distance. In IP based networks, there are many more causes of network delay. A few are discussed below.

A.5.1 Network Applications

Network applications determine how much information needs to be transferred between a sender and receiver. If the application requires a great deal of data transfer between sender and receiver, such as graphic objects contained in Web pages, the delay in communicating over the network will be higher than if the application is developed in a way that minimizes the amount of data that must be transferred.

A.5.2 Network Application Server Capacity

All computers, including network servers, have limited resources such as memory, CPU processor capacity, et cetera. Therefore every server can only support a certain number of simultaneous connections before all resources are consumed and the server capacity degrades. If a network server is expected to receive more connection requests than its internal resources can support, it is possible to balance the expected connection load across several servers that are “mirrored.”

---

[39]Sean McCreary and kc Claffy, *Trends in Wide Area IP Traffic Patterns: A View from Ames Internet Exchange*, CAIDA Technical Report (San Diego, Calif.: Cooperative Association of Internet Data Analysis, February 2000) [On-line]. URL: http://www.caida.org/outreach/papers/2000/AIX0005/AIX0005.pdf (Accessed on 1 March 2003.)

A.5.3 Client Computer Capacity

Client computers that have limited resources (memory, CPU capacity) will take longer to transfer data over a network than highly capable computers.

A.5.4 User Training

Computer users often contribute to delays in communicating over IP networks. If Windows users open too many applications on their machines they may consume all available random access memory (RAM) and their computers will begin swapping data between the RAM and hard drive swap space. This will delay the downloading of information when connecting over an IP network. Users should be trained to limit the number of open applications on their computers when they desire to surf the WWW.

A.5.5 Propagation Delay

There is a finite limit to how fast light can pass through fiber-optic cable and how fast radio waves can travel through the atmosphere or space. The further the sender is from the receiver, the longer it will take for information to travel across an IP network.

A.5.6 Router Capability

Every router must determine the appropriate output interface to send packets it receives. State-of-the-art high-end routers will transfer packets faster than older, less capable, routers.

A.5.6 Network Congestion

As previously described, when network congestion occurs, packets get dropped. If the packets are part of a TCP connection, the sending application will normally cut its data transfer rate in half. If a network user wishes to communicate with a server and there is congestion anywhere between the user and the server, there is a possibility that some of the packets will get dropped and the connection data rate will be lower than during times when there is no congestion.

A.5.7 Packet Fragmentation

If a packet exceeds the maximum transmission unit (MTU) size for Ethernet, it will be fragmented into two packets that do not exceed the MTU prior to being transmitted out over an Ethernet link. One common situation that leads to packet fragmentation is the use of virtual private networks (VPNs). VPNs use Internet Protocol Security (IPSec), which encrypts entire packets (data and header) and then adds an additional layer 3 header to the encrypted packet. If the original packet was as large as the MTU, the encrypted packet (with the additional header) will exceed the MTU and must be fragmented. This results in more packets being transmitted to complete a given network connection and adds delay to the connection time. When the USAF

deployed VPN cards in all of its SDP routers, the AFNOC had to modify the router software to disregard the setting of the "do not fragment" bit in the IP header of AF packets.

A.6 QoS in IP Networks

As described earlier, the Internet was designed to be a "best effort" network. No guarantee was made that any given packet would reach its intended destination. This worked well for applications such as e-mail and file transfers, where TCP retransmits any segments missing due to packet loss. However, newer applications that transmit voice and video by IP packets may become unusable if too many packets are lost or are subject to excessive jitter (variations in end-to-end latency between packets). To support these types of applications, the Internet community developed methods to provide QoS across IP networks.

QoS is all about resource allocation. Packets in an IP network compete with each other for link bandwidth and router buffer space. For an IP network to treat some network traffic better than other traffic, the network routers must be configured to recognize which packets deserve preferential treatment. The Internet community developed integrated and differentiated service models to deal with this issue in order to provide QoS in IP networks. Computer network traffic engineering using Multi-Protocol Label Switching (MPLS) can also be used to improve the level of service provided to some traffic in IP networks.

A.6.1 Integrated Services

The integrated services model is based on a per-flow resource reservation. For an application to be guaranteed some level of network resources (such as bandwidth), the application must first set up a resource reservation in all the routers between the sender and receiver. To do this, the Resource Reservation Setup Protocol (RSVP) is used. RSVP makes a reservation for network resources for traffic flow in only one direction. If the application is duplex (transmit and receive in both directions), two separate resource reservation setups are required.[40]

For example, suppose a user at one location wishes to videoconference with someone at another location. The videoconference application could use RSVP to connect through the IP network and establish a minimum bandwidth reservation for each router in the path. The receiver application would also have to use RSVP to reserve a minimum amount of bandwidth in the reverse direction. All of the routers at any possible congestion point between the two users must be RSVP capable and have sufficient bandwidth available to support the connection before any video packets can be sent.

---

[40]Zhen Wang, *Internet QoS; Architectures and Mechanisms for Quality of Service* (San Francisco, Calif.: Morgan Kaufman Publishers, 2001), 39.

The integrated service model works well for providing QoS for long-lasting IP network connections. However, deployment of this architecture in ISP networks has been slow for several reasons. This model was originally developed to support long-lasting and delay-sensitive applications (such as video teleconferencing) over the Internet. The WWW changed the nature of Internet traffic from a majority of long-lasting connections such as e-mail, remote access, and file transfer to short "bursty" connections where Internet browsers make many short connections to download Web page objects. The overhead required to establish resource reservations in every router between two network nodes is very inefficient for WWW-type traffic. Another problem with this model is the overhead associated with the accounting and billing management required to charge for the preferential treatment of RSVP traffic. Billing agreements must be made with different ISPs when connections traverse multiple ISP domains.[41]

A.6.2 Differentiated Services

The DS model uses a much different approach to allocate resources to preferential traffic. Instead of classifying traffic on a per-flow basis, the DS model marks packets in different forwarding classes. DS uses the first six bits of the original IPv4 8-bit ToS field in the IP packet header (Fig. A-2) to prioritize network traffic.[42] These six bits, called the Differentiated Services Code Point (DSCP), are used by DS-capable routers to give preferential treatment to higher priority traffic. The router treats all packets with the same DSCP the same way. Packets can be marked so that they are not likely to be dropped during congestion or marked so that they are given forwarding priority (minimize delay). The DSCP can be used to assure that levels of link bandwidth are provided to high-priority classes of traffic.[43] How the DS-capable router treats each forwarding class is called the per-hop behavior (PHB) of the router.

At the edge of the IP network, incoming traffic is marked (the DSCP bits are set) according to the management rules established by the network customers. For example, packets from or to key network nodes (such as the company portal or the chief executive officer's desktop) can be given preferential treatment. Packets from important company applications, such as network collaboration tools or financial transactions, can be given priority over routine applications such as e-mail and file transfers. In the core of the network the routers do not have to classify traffic so the preferential PHB can be implemented at very high data rates. It is important that all routers in the DS domain be DS capable. Otherwise, high-priority packets may be dropped and the QoS

---

[41]Ibid., 6.

[42]Network Working Group, RFC 2474, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers (Standards Track)," defines the DS field and class selector and default PHBs. [On-line]. URL: http://www.ietf.org/rfc/rfc2474.txt (Accessed on 1 March 2003.)

[43] Wang, *Internet QoS*, 104.

level cannot be guaranteed.[44] Too much high-priority traffic can still overwhelm a DS-capable network and result in packet loss.

A.6.3 Multi-Protocol Label Switching (MPLS)

MPLS is a technology developed to provide traffic engineering in an IP network. Traffic engineering means managing network bandwidth and traffic flows to selectively determine the route of traffic to prevent congestion. One motivation behind the development of MPLS is the difference in speed between switching and routing. In IP routing, the packet-forwarding engine in the router must examine the destination IP address of the packet, use the routing table to look up the appropriate egress interface of the router, and then send the packet on its way. Every router in the IP network must perform these functions. In MPLS, only the packet label (inserted into the packet by a network edge router called a Label Switch Router—LSR) is used to determine the egress interface on the router. Since there is no path lookup, switching is faster than routing.

The drawback to MPLS is that prior to sending traffic through the network, the routers must first establish Label Switch Paths (LSPs). The LSRs then swap labels on packets as the packets are switched from router to router according to the previously established LSP. This way, the IP header data is examined only at the ingress point to the MPLS network. When a packet enters the MPLS network, the LSP is chosen (or established if no existing LSP supports moving the packet to its destination IP address), the LSR adds the label appropriate for the chosen LSP, and the packet is switched through the network until the label is removed at the exit point of the network. MPLS networks can be set up with virtual circuits (LSPs) that take advantage of network provisioning and traffic engineering to provide the desired QoS level for IP traffic. Advances in silicon technology and ASIC that permit very high-speed IP routing have prompted some debate over the need and potential cost saving of packet switching verses packet routing.[45]

[44]Ibid., 105–106.

[45]Ibid., 141–142.

Acronyms and Abbreviations

ACK	acknowledgment field significant
ACL	access control list
AFB	Air Force base
AFCERT	Air Force computer emergency response team
AFNOC	Air Force network operations center
ASCII	American Standard Code for Information Interchange
ASIC	application-specific integrated circuit
BGP	Border Gateway Protocol
CERT	computer emergency response team
CERT/CC	CERT coordination center
CITS	Combat Information Transport System
COIN	community of interest network
CONOPS	concept of operations
COOP	continuity of operations plan
CPU	central processing unit
CSAF	Chief of Staff, Air Force
DDoS	distributed denial of service
DISA	Defense Information Systems Agency
DISN	Defense Information Services Network
DoD	Department of Defense
DS	differentiated services
DSCP	Differentiated Services Code Point
ESC	Electronic Systems Center (U.S. Air Force)
FIFO	first in first out
FIN	no more data from sender
FTP	File Transfer Protocol
Gbps	gigabits (billion bits) per second
GCSS-IF	Global Combat Support System Integration Framework
HTTP	Hypertext Transfer Protocol
IANA	Internet Assigned Numbers Authority
ICANN	Internet Corporation for Assigned Names and Numbers
ICMP	Internet Control Message Protocol
IEN	Internet Engineering Note

IP	Internet Protocol
IPv	Internet Protocol version
IS-IS	Intermediate System to Intermediate System
ISDN	Integrated Services Digital Network
ISP	Internet service provider
IT	information technology
LAN	local area network
LSP	label switch path
LSR	label switch router
MAC	Media Access Control
Mbps	megabits (million bits) per second
MTBF	mean time between failures
MTTR	mean time to repair
MIT	Massachusetts Institute of Technology
MPLS	Multi-Protocol Label Switching
MTU	maximum transmission unit
NCC	network control center
NIPRNET	Non-secure Internet Protocol Router Network
NMCI	Navy–Marine Corps Intranet
NOSC	Network Operations and Security Center
OSI	Open Systems Interconnect
OSPF	Open Shortest Path First
PHB	per-hop behavior
PRI	primary rate interface
PSH	push function
QoS	quality of service
RAM	random access memory
RFC	Request for Comment
RST	reset the connection
RSVP	Resource Reservation Setup Protocol
RTT	round trip time

SDP	service delivery point
SECAF	Secretary of the Air Force
SIPRNET	Secure Internet Protocol Router Network
SLA	service-level agreement
SMTP	Simple Mail Transfer Protocol
SYN	synchronize sequence numbers
TCP	Transmission Control Protocol
ToS	type of service
TTL	time to live
UDP	User Datagram Protocol
URG	URGENT point field significant
USAF	Air Force
VPN	virtual private network
WAN	wide area network
WWW	World Wide Web