STUDY TO DETERMINE THE NEED FOR AND FEASIBILITY OF IMPLEMENTING A NATIONAL IP-BASED PUBLIC SAFETY INTERCONNECTIVITY AUTHENTICATION PROCESS

Prepared by the
National Law Enforcement and Corrections Technology Center
- Rocky Mountain Region
NPSTC Support Office

May, 2004

Points of view are those of the authors and do not necessarily represent the official position of the U.S. Department of Justice. This document is not intended to create, does not create, and may not be relied upon to create any rights, substantive or procedural, enforceable by any party in any matter civil or criminal.

The National Law Enforcement and Corrections Technology Center is supported by Cooperative Agreement #2001–RD–CX–K001 awarded by the U.S. Department of Justice, National Institute of Justice. Analyses of test results do not represent product approval or endorsement by the National Institute of Justice, U.S. Department of Justice; the National Institute of Standards and Technology, U.S. Department of Commerce; or Aspen Systems Corporation.

The National Institute of Justice is a component of the Office of Justice Programs, which also includes the Bureau of Justice Assistance, Bureau of Justice Statistics, Office of Juvenile Justice and Delinquency Prevention, and Office for Victims of Crime.

Author Contact:

Kathy J. Imel
President
La Loba International, Inc.
Phone: 303–438–9565
Fax: 303–438–1244
E-mail: kjimel@aol.com

NPSTC National Support Office Contact:
Laura Lippman
NPSTC NSO Office Administrator
Toll Free: 866-807-4755
Direct: 303-649-1843
Fax: 303-649-1844
Email: llippman@highlands-group.com

TABLE OF CONTENTS

STUDY TO DETERMINE THE NEED FOR AND FEASIBILITY OF IMPLEMENTING A NATIONAL IP-BASED PUBLIC SAFETY INTERCONNECTIVITY AUTHENTICATION PROCESS

EXECUTIVE SUMMARY

The Rocky Mountain Regional office of the NLECTC was asked by the National Public Safety Telecommunications Council to examine the feasibility of establishing a national database of public safety IP addresses for mobile devices. A list of functional requirements of public safety mobile interconnectivity were drafted and used as the basis for evaluating the NPSTC request. The study reviewed the current literature in the areas of Internet protocol, network heterogeneity, device variability, the number of public safety devices to be controlled, and security.

The study found that the public safety requirements for mobile interconnectivity cut across multiple technological and logistical areas, and no single technology change or action would be sufficient to accomplish the entire list of complex requirements of public safety. Furthermore, the majority of the NPSTC requirements were being addressed by existing research, government, or commercial efforts.

The study makes three recommendations:

1. Recommends not to implement a national IP-based authentication process, for a variety of reasons as identified in the study.
2. Recommends that a project to develop single sign-on for public safety applications be considered.
3. Recommends that next generation network development efforts that address the needs of public safety, as currently being carried out by Project MESA, be supported and monitored.

1. INTRODUCTION

The proliferation of Internet Protocol (IP)-based mobile communications networks is having a direct impact on the ability for public safety agencies to communicate, both internally and externally. Public and private infrastructures are both being used, and seamless communications is not always achieved. To properly manage the incident command activities associated with a large multi-jurisdictional public safety response, multiple agencies must be able to intercommunicate, data from a host of disparate sources must be accumulated, and secure integration, authentication, and administration of all agency resources must be accomplished quickly and easily.

1.1. BACKGROUND

In 2002, the National Public Safety Telecommunications Council (NPSTC) received a request from the Steering Committee of the FCC's National Coordination Committee (NCC) to examine the feasibility of establishing a national database of public safety IP addresses for mobile devices. Subsequently, in April, 2003, the Chair of the NCC's Interoperability Subcommittee prepared and submitted to NPSTC a White Paper discussing the concept in detail. Most recently, NPSTC requested acquisition of a secure domain name to be used when public safety units interfaced with the Internet.

As the NPSTC support office, the Rocky Mountain Regional office of the NLECTC (NLECTC-RM) was asked by NPSTC to examine these requests and to determine the level of effort which would be required to implement such a system. Under funding from the NIJ AGILE program, NLECTC-RM established a team which, as a first step, examined the goals of the project, reviewed other ongoing activities of a similar or related nature, and drafted an initial list of functional requirements for national IP-based interconnectivity as summarized below:

1) Ability of users to seamlessly roam between public and private data systems carrying with them their specific user privileges.
2) Establishment of standards for the certificate and physical communications layer protocols for wireless devices, consistent with pre-existing public safety standards.
3) Provide the technology and applications platforms necessary to access new telecommunications and automation tools and/or expand access to existing public safety information systems.
4) Emergency systems must be able to interoperate and integrate with other agencies systems.
5) Ability to establish ad-hoc mobile networks.
6) Ability to provide secure (authenticated, authorized, and/or encrypted) intra- and interagency communications.
7) Ability to provide multiple levels and geographical and service-specific system availability.
8) System capable of rapid transmission of data, through harsh operating environments, with minimum of data or transmission errors.

At this point, it was unclear whether development and operation of a new national-level IP-based authentication process was necessary, or whether current and rapidly advancing new technologies would be sufficient to meet the expressed needs of the public safety user community. What was also unclear was the impact on and requirements for integrating with other already existing databases and networks. Finally, if the project were feasible, the level of effort and cost to develop and administer such a system needed to be identified.

NLECTC-RM determined that the appropriate next action was to conduct a study to evaluate the need for and feasibility of implementing a national IP-based public safety interconnectivity authentication process.

1.2. GOALS OF THE STUDY

The intent of the study was to examine the following:

1. Determine whether a new national-level authentication process was necessary
2. Examine in more detail the current projects which are using or developing similar technology/processes
3. Identify the integration issues involved in developing such a system at the national level
4. Identify the impact of creating a new national-level process on the affected state and local public safety related entities
5. Identify the technologies that would be required to implement the requested system
6. Evaluate the impact and costs to implement the requested system, if judged feasible
7. Recommend a course of action

1.3. STUDY APPROACH

Initially, the study design called for using several groups (a Project Team, an Advisory Committee, a Technology Subcommittee, and a User Subcommittee) to participate in and provide input to the study. As the project proceeded, it became apparent that the wide variety of technologies and issues which surrounded what had seemed to be a simple goal made it impractical to try to provide a single clear mission to a group of, primarily, volunteers. As a result, the study approach was modified into a more classic literature review with subsequent peer review scenario. In this way, all of the major challenges surrounding the issue could be incorporated into the study by the Project Team as they were encountered. In addition, the results of the literature review would serve as a means for ensuring that the peer reviewers understood all of the issues that were encountered.

2. PUBLIC SAFETY MOBILITY FUNCTIONAL REQUIREMENTS

The initial *Operational Requirements for National IP-Based Public Safety Interconnectivity: Statement of Functional Requirements*[1] developed by the Project Team on behalf of NPSTC included a number of scenarios which corresponded directly with the public safety-related

scenarios developed by the Project MESA group.[2] In summary, the overall goals articulated in the *Functional Requirements* document were as follows:

1. Provide the ability for users to securely and seamlessly roam between public and private data systems carrying with them specific user privileges
2. Create a set of uniform specifications and/or standards for the associated technologies
3. Support multiple levels of security and encryption
4. System and network switching technology built from these specifications and/or standards to be capable of switching and transferring multiple applications, from multiple user devices through multiple system components, to one or more authorized hosts or network servers and/or to other user terminals
5. Results to satisfy the current and identified long-term needs and requirements of national and local public safety, public service, and public security communities from throughout the world

2.1. ASSUMPTIONS

The *Functional Requirements* document makes a number of basic assumptions, including:

1. Individuals will carry multiple devices (e.g., pager, PDA, portable radio, cell phone, etc.)
2. Same device could be used by different people
3. An agency may carry in IPv6-based mobile wireless broadband network to an incident scene to be used to connect to outside world
4. Units entering the scene would automatically register with the network and announce their availability and capability to the network area controller

2.1. DATA INVOLVED

At the scene of an incident information would need to be both accessed and transmitted. The data that would need to be accessed includes (but is not limited to):

Local law enforcement records systems
State law enforcement records and criminal history systems
State DMV files
National law enforcement records and criminal history systems
International law enforcement records and criminal history systems
Fingerprint files - local, state, national, international
Photo/mug shot files - local, state, national, international
Iris scan files - local, state, national, international
Hazardous materials files
Computer aided dispatch (CAD) files - local, state, national (including: unit capabilities)
EMS files, including: Patient information, Medical information

Triage information
Geographical information systems - local, state, national, international
Structure data, including building floor plans, fire control equipment locations, pre-fire plans, and occupant info
Hydrant and other water equipment related information
Environmental information (wind speed and direction, temperature, etc.)
Traffic/transportation management systems
Equipment management systems
Robotics transmissions
Streaming video transmissions (private and public security cameras, traffic monitoring cameras, etc.)

Among the various types of data that would need to be transmitted from the scene of the incident are included the following:

Unit status updates
Bio-medical information from electrical measuring devices
Patient information
Personal Alert Safety Systems information
AVL information
Personnel location information
Personnel emergency notifications
Streaming video
Equipment monitor status information
Equipment control information (traffic systems, public/private utilities)
Robotics control

3. PUBLIC SAFETY MOBILITY ISSUES AND CHALLENGES

The ability to allow a variety of different users and user devices to securely and seamlessly roam over a national (and even international) landscape between a number of heterogenous public and private data networks while carrying with them specific user and/or device privileges presents a multitude of challenges. This resolves into two primary areas: the network (maneuvering within it, as well as its characteristics) and security.

Some of the challenge includes establishing an understanding of all of the terminology used within mobile and security technologies, as well as the current state of technology development in a number of areas. To ensure that the issues discussed below are consistently understood by all readers, a certain amount of space has been dedicated to providing critical definitions and status of development for each of the challenges identified.

It should be noted that none of these issues exist in a vacuum. They are all interrelated to some degree, and therefore, must all be addressed in order to achieve the goals identified by NPSTC. Addressing any single one alone will not be sufficient.

In addition, a key point to note is that the technologies and standards involved in mobile communications and security are constantly evolving and changing. A multitude of standards bodies, international interest groups, and commercial entities are actively working in all of these areas (with new entities cropping up each year). As a result, any summary, such as this one, can only be a snapshot in time and is almost obsolete before it is written.

3.1 INTERNET PROTOCOL

The Internet Protocol (IP) is the basic building block upon which all Internet protocols are built. Applications, such as Web browsers and email, use TCP and UDP which are built on top of IP. In simple terms, "IP is a network protocol that routes data across a set of networks connected by routers (an internet)."[3] IP addresses (in IP version 4 [IPv4], these are 32-bit numbers written in the form of four decimal numbers, each between 0 and 255, separated by periods) are used to identify the locations within the networks. Because numeric IP addresses are difficult for humans to remember, IP addresses are often given meaningful, hierarchical names in the Domain Naming Service (DNS), names such as www.myinternetaddress.com. Each domain and device connected to the Internet network is assigned an IP address, either as a fixed address or as a dynamically assigned address.

Most researchers postulate that post-3rd generation mobile systems will include a variety of wireless access networks connected to an all-IP based core broadband network, enabling global roaming using the radio technique that is best suited to support the requested service, and by terminals that may consist of several different components, many of which will be assembled into personal-area networks (PANs). As time progresses, most, if not all, devices will be built with the ability to be accessed via their IP address. The movement towards universal IP addressing facilitates the convergence of fixed and mobile networks with the ultimate realization of a global common standard worldwide. However, at the moment, many of the devices used by public safety, chief among them being handheld portable radios, are not IP addressable.

Mobile IP is the current standard for supporting mobility in IP networks. Solutions for mobility management include Mobile IP (v4 and v6), Hierarchical Mobile IP, Regional Registration and various other micro-mobility protocols, each with its own security solution.

Aghvami and Jafarian maintain that "mobile IP is suitable for handling macromobility and slow-moving mobile hosts, i.e., when there is infrequent mobile host migration. However, for frequent mobile host migration (the wireless environment), the use of Mobile IP leads to an increase in handover latency and load on global networks."[4] Hence the use of wireless local area networks (WLANs) and even personal area networks (PANs, also called piconets). Separating the local mobility from the wide-area mobility can improve overall performance.

3.1.1. SEAMLESS ROAMING

Internet protocol was designed for networks with fixed location users. It is a 'best effort' protocol designed for a pure data network. It does not support real-time services, something required to support the multimedia applications envisioned in public safety's future. For quality of service in an IP-based network, other protocols (such as TCP and UDP) are necessary. However, according to a researcher at Stanford,

> "Reliable transport protocols such as TCP that have been traditionally used for wired networks do not perform as well on wireless networks. This is because TCP assumes that packet loss and unusual delays are mainly caused by congestion. TCP is thus tuned to adapt to such congestion losses by slowing down the amount of data it transmits. It drops its transmission window size and backs off its retransmission time, thus reducing the load and congestion on the network. In wireless networks however, packet loss is often caused due to other factors besides congestion. Wireless channels often suffer from high bit error rates (BER) and intermittent connectivity due to handoffs. Moreover, the BER may vary continuously during a session. TCP unfortunately assumes that these losses are due to congestion and invokes its congestion control measure. This results in an unnecessary reduction in end-to-end throughput and thus sub-optimal performance."[5]

As a result, research into other technologies that would operate better in a mobile environment continues. Other technologies which may be employed in the future to ensure quality of service include MPSL (Multi Protocol Label Switching) and RSVP (Resource reSerVation Protocol).

Aghvami and Jafarian state that "mobile IP can support portability but not mobility. In other words, the computer can be operated at any of a set of points of attachment, but not during the time that the computer is changing its point of attachment."[6] This means that, within a single wide area network (the core network), devices can move around and remain connected, and thus, Mobile IP will be sufficient, but not when roaming between cellular (or other wireless) networks or within WLANs and PANs (micro-mobility) where handoff becomes an issue.

IP version 4 (IPv4) has a number of difficulties managing mobile devices for several reasons, including:

> 1) Mobile computers need to use a forwarding address at each new point of attachment to the network. With IPv4, getting this address may be difficult.
>
> 2) Good authentication facilities are required to inform any agent in the routing infrastructure about the new location of the mobile node. These facilities are not commonly deployed in IPv4 nodes.
>
> 3) In IPv4, it can be difficult for mobile nodes to determine whether or not they are attached to the same network.

4) In IPv4, mobile nodes usually cannot inform their communication partners about a change in location.[7]

A common standard for micro-mobility management protocols has yet to be determined. Several have been proposed. The European Information Society Technologies' (IST) "Moby Dick" project has been working to address present and future mobile communications services by focusing on a paradigm based upon IP version 6 (IPv6). They recently demonstrated the ability to transmit uninterrupted streaming data to a terminal, despite changing the point of attachment to the network over three access technologies – WLAN, TD-CDMA, and Ethernet. They are currently conducting similar tests using PDAs.[8]

Although the final technologies to be used in next generation networks are still being developed, once established, it should be possible to have a transparent core and access network.

3.1.2. CURRENT SITUATION - IP Version 4 (IPv4)

Most researchers in the field seem to agree that the next generation networks will rely on the use of the newest standard for Internet Protocol, IP version 6. However, in order to understand the historical context from which the request for this study partially derived, a summary of the issues facing the previous standard, IPv4, as well as the ramifications of the new standard, are included below.

GPRS (General Packet Radio Service) is the network protocol to which commercial mobile data network providers in the U.S. are moving or have already moved to. Most GPRS networks currently use DHCP (Dynamic Host Configuration Protocol) to dynamically assign IP addresses to mobile devices. DHCP's purpose is to enable individual computers on an IP network to extract their configurations from a server (the 'DHCP server') or servers, in particular, servers that have no exact information about the individual computers until they request the information. The overall purpose of this is to reduce the work necessary to administer a large IP network. The most significant piece of information distributed in this manner is the IP address. (Although most commonly used for dynamic IP address assignment, there is nothing in the DHCP protocol that precludes the support of statically defined addresses.)[9]

Addresses can change frequently during the day – whenever the user roams out of coverage, is inactive for a while, logs out of the application, or restarts their device. In a wired environment, a typical client-server application does not have to worry about continuously and automatically re-establishing connections during the day. User sessions are normally established once (when the user starts the application), with the connection persisting for as long as the user is working with the application. In the wireless world, the user session may disconnect five times in an hour, with the device assigned a different IP address each time. Wireless applications with time-critical dispatch requirements need a mechanism in place to keep user sessions active, regardless of frequent disconnects and changing IP addresses.

In addition, GPRS operators often implement a timeout for inactive sessions (this can be as short as 5 minutes for most U.S. GPRS operators[10]). If there is no data exchanged between the client and server for this period, the connection is dropped and the IP address goes back into the pool to be reassigned to another device. If this happens, the server application cannot reestablish the connection after a timeout, and therefore cannot push data to the device until the device reconnects with the server.

For a variety of reasons, the handset model in commercial services has been implemented with a dynamic IP address. This means that each time a wireless packet client registers with its serving network it will be assigned a new private IP address, known only to the serving network. For browsers, email clients, and derivatives (such as calendar programs), dynamic addressing is not a problem, as these applications access well-known server addresses that are almost always located on the 'tethered' Internet. But, if you want to connect to a remote wireless video server, a high-speed wireless modem, or any packet-side machine-to-machine (M2M) interface, you have to know the IP address. If the address changes every time the unit registers, the address is, essentially, unknowable.

End users originate browser and mobile office transactions, allowing for an automatic and transparent re-registration. Mobile-to-mobile (M2M) endpoints must be awakened, and the resulting temporary IP address must be communicated to the connecting client process. On current wireless networks, the process of waking up a packet side endpoint involves either sending it a message via the Short Message Service (SMS) or dialing its corresponding circuit side phone number.[11]

A large number of solutions to the dynamic IP address issue have been proposed. The details of the implementation vary, but when simplified there are essentially two models using IPv4: a two-party model and a three-party model. As Ctek explains "In the two party model, the client wishing to access a packet-side endpoint has all the knowledge and connectivity required to make the wake up circuit call, or SMS transmission, and receive back the answer in the form of a dynamically assigned IP address. The problem with this is that each end has to maintain the knowledge of how to get to the other, hardly a scalable solution. Variants of this approach would have the callback information passed along in-band as a part of the wake up message. This is slight improvement but still has the one-to-one baggage of circuits and/or logic required to send the wake up. The two party models require dedicated resources and a level of shared information that makes it cost prohibitive for many applications."

"In the three party model we add a mediator, a shared resource with a well-known address that maintains a list of permanently assigned names and resolves them to temporary IP addresses... [using DNS]. With DNS we are free to move our web pages from one hosting firm to another, poke a change at DNS, and expect that the pages will load with the same URL that was used at the old hosting company. A vanilla DNS server properly loaded with the names and temporary addresses will meet most of the requirements of disambiguating a dynamically assigned IP address using a well-known name. By adding end point wakeup and enhanced DNS update functionality to the public domain DNS application, a workable wireless name server application

can be built."[12] Name servers applications such as this have already been made commercially available for existing networks.

The issue for public safety as it relates to dynamic IP addresses is also one of security. Many of the legacy applications for accessing State and local public safety databases rely on associating a terminal with a static IP address in order to ensure that device is authorized to receive confidential data. A number of these applications have been modified recently to modify and/or eliminate the requirement for all mobile devices to have a static IP address. But end-to-end security with IPv4 cannot be guaranteed.

3.1.3. FUTURE SITUATION - IP Version 6 (IPv6)

IP version 6 was designed to replace IP version 4. IPv4 has not been substantially changed since the initial protocol was published in 1981. The protocol has proven to be robust, stable, easily implemented, interoperable, and has supported tremendous expansion to the size of the Internet today. About 10 years ago, work was started on a successor protocol, originally called IPng (for "next generation").

One of the primary reasons for the creation of a new protocol was the anticipated exhaustion of available IP address space. A 32-bit address space (as used in IPv4) would theoretically support approximately 4 billion IP devices. However, in reality, all of the potential addresses are not available democratically because a large percentage of them are pre-allocated (the government, universities, and organizations within the United States alone hold nearly 74 percent of the total IPv4 address space[13]). For some entities outside of the United States, particularly in Asia, IPv4 addresses have already stopped being available. However, even for Europe and the United States, estimates predict that the official IPv4 address space will be exhausted by 2005 or 2006.[14]

The current IP address space will not satisfy the huge increase in the number of users or the geographical needs of the continuing Internet expansion, let alone the requirements for emerging applications that need even more permanent IP addresses (e.g., cars, TVs, household appliances, etc.). While technologies such as Network Address Translation (NAT), Classless Interdomain Routing (CIDR), and temporary-use allocations (DHCP and RADIUS/PPP) have been developed as short-term fixes, they are not the best long-term solution. They don't allow for end-to-end security; they break the peer-to-peer model; and they create bottlenecks in the network.

IPv6 was the draft standard that emerged from the IPng project. (Draft Standard status is granted to protocols for which multiple interoperable implementations have been demonstrated and for which the technology is believed to be mature and stable.) IPv6 uses 128-bit addresses, which should be sufficient to provide enough globally unique IP addresses for the foreseeable future (128-bits provides roughly 340 billion billion billion billion unique addresses). Put another way, this means that every citizen on earth could have over 1 million unique IP addressable devices.[15] However, the larger address space required a change to the basic format of the IP packets.

Because such format changes are not backwards-compatible, the IPng project also considered other changes that would correct known deficiencies in IPv4. These included the need for simpler autoconfiguration and renumbering, requirements for security at the IP level, and the need for better support for real-time deliver of data (also called quality of service, or QoS), among others.

Some of the improvements designed into IPv6 include:

1) IPv6 has built-in security support (IPsec). Thus, all IPv6 nodes support cryptographic-based authentication and encryption and so can provide end-to-end security services.

2) IPv6 facilitates efficient renumbering of sites by explicitly supporting multiple addresses on an interface.

3) IPv6 eliminates the need for network address translation and application's layered gateway since every IP device will be able to have its own globally unique IP address (thus eliminating the need to translate hundreds of internal IP addresses into a few global IP addresses).

4) IPv6 supports an increased number of multicast addresses. Multicast allows IP packets (such as streaming video) to be sent to multiple destinations at the same time, saving network bandwidth.

5) IPv6 supports autoconfiguration and "plug and play" operation. A new machine can generate its own address automatically and can immediately communicate with other nodes. The autoconfiguration can either be performed in a *stateless* manner (in which a node forms the addresses itself) or in a *stateful* manner (in which the node uses DHCPv6 to obtain addresses and other configuration information).[16] "Autoconfiguration eases the management burden and administrative cost to the operator by allowing end stations to auto-configure their IPv6 address based on the subnetwork prefix of a default gateway."[17]

6) IPv6 provides enhanced support for Mobile IP and mobile computing devices. In Mobile IPv6, macro-mobility is built in and any IPv6 node can use mobility as needed. IPv6 packets addressed to the home address of a mobile node are transparently routed to its 'care-of' address (without having to pass through the home network). Home agents proxy on behalf of nodes that consider a particular link their "home," but are currently located elsewhere. Home agents intercept packets sent to a mobile node's home address and resend them to the mobile node's care-of address (i.e., the mobile node's current location). When a mobile node is visiting a link, it obtains a temporary care-of address, using the IPv6 autoconfiguration mechanisms.[18] At the same time, the globally unique "home address" is retained for authentication, authorization, and accounting (AAA) purposes and to ensure that applications that are dependent upon transport and higher level layer connections are not broken.[19]

While Mobile IPv6 and some related routing solutions have been developed to better support mobile networks, the focus has been more on terminal relocation rather than the handover process. Thus Mobile IPv6 as it currently stands supports macro-mobility (inter-domain handoffs); while micro-mobility (intra-domain handoffs) is being addressed through a number of proposed extensions to Mobile IP which have not been finalized in the IETF framework.[20] Candidates for implementation include fast handover (FMIPv6) and hierarchical schemes (e.g., HMIPv6).

A number of transition mechanisms have been developed to allow IPv6 to be used with IPv4 backbones (both network and Internet). As a result, new network segments can be implemented now with IPv6 without waiting for the backbone to be upgraded. However, until mobile terminals can have unique permanent IPv6 addresses, NAT must continue to be used to map private/public IPv4 addresses as well as to map roaming bearer traffic within application layer gateways (within the GPRS Tunneling Protocol [GTP]).[21] Also, during the transition period (and possibly forever), many implementations will decide to take a dual-stack approach, whereby the applications will choose which IP version to access based on the format of the IP address returned by the DNS server when sending.

A number of ISPs in Asia have already begun offering IPv6 service on a commercial level. The 6bone Network is a global IPv6 test network that was started in 1996 and which now includes more than 1,000 participating hosts. It is projected that by 2006, at least 50 percent of ISPs will be offering IPv6 services.[22]

In January, 2003, members of the European Union-based IPv6 Wireless Internet Initiative (6WINIT) conducted what they claimed to be the first IPv6 over 3G UMTX/WCDMA network demonstration using a medical application. The medical emergency system, called Guardian Angel, was able to move seamlessly between different access networks. From the hospital setting, doctors were able to observe the patient in the ambulance, check the heart rate and blood pressure (using GMS/GPRS or UMTS/WCDMA). Once the ambulance reached the hospital, the system automatically switched over to the hospital's indoor WLAN hotspot. In addition, data flows were able to operate in parallel; if the WLAN did not have sufficient reliability, the data transmissions could simultaneously use the GSM/GPRS or UMTS/WCDMA channel.[23]

3.2. NETWORK HETEROGENEITY

To accomplish all of the functional requirements identified by NPSTC, a number of heterogeneous (dissimilar) networks will be used, some types that are in existence today and some that are either in development or yet to be developed. Unfortunately, each of these different network types uses one, or in some cases multiple competing, protocols operationally because of the various technologies utilized. While the individual protocols may be quite robust for communications within the network, they do not necessarily address the requirements for communications between networks. In addition, the various networks operate over different frequency bands. As a result interoperability between the various networks has been difficult to non-existent.

Two different models for wireless communications, and standards development, exist today. One model is derived from the radio and telephone industry. The second model is based upon computer network technology. Public safety related wireless networks include private radio networks (such as those used in the majority of public safety agencies today and which can be analog or digital), commercial cellular telephone networks, and wireless computer networks (i.e., wireless local area networks [WLANs], personal area networks [PANs], broadband wireless networks, and mobile ad hoc networks [MANETs]). The latter two of these are defined and described below, particularly in regards to the standards under which they operate.

3.2.1. CELLULAR TELEPHONE NETWORK MODEL

For many lay people, wireless communications means 'cellular telephones.' First generation (1G) wireless cellular networks were analog and have largely been replaced by second generation (2G) digital networks (although many 1G networks are still in operation). Most of the 2G networks are circuit-switched networks which support analog fax, circuit-switched voice and data, and cellular digital packet data (CDPD) with typical speeds limited to 9.6 kbps to 19.2 kbps. The primary technologies used were code division multiple access (CDMA), time division multiple access (TDMA) and Global System for Mobile Communications (GSM). These 2G mobile radio systems have been deployed successfully worldwide.

To meet the needs of today's subscribers, many wireless service providers are in the process of or have completed upgrading their 2G networks to 2.5G networks, which will continue to use the 2G architecture to deliver voice and circuit-switched data applications while adding a packet data overlay (i.e., General Packet Radio Systems, or GPRS) to support additional packet data services. Upgrading a 2G wireless infrastructure to support 2.5G enables subscribers to obtain data rates up to 170 kbps. In the near future, third generation (3G) networks will have voice and data converging over one common, shared packet (i.e., IP or ATM) backbone using technologies such as CDMA 2000, Enhanced data rates for Global Evolution (EDGE), and Universal Mobile Telecommunications System (UMTS). These next-generation networks will provide even higher data rates of up to 2 Mbps to support streaming music, multimedia, and voice and data services. While already deployed in some locations (primarily in Europe and Asia), it is predicted that 3G systems will be in wide use no earlier than 2005.[24]

Third generation wireless systems have been standardized via the International Mobile Telecommunications-2000 (IMT-2000) standards developed by the International Telecommunications Union - Radiocommunications Sector (ITU-R). The ITU is the traditional body for the publication of international standards in telecommunications. It is a United Nations agency with national governments as member organizations as well as private sector participation.

The open question is how the 3G systems will evolve. The literature seems to agree that the fourth generation (4G) network (also referred to as the Next Generation Network in Europe) is not likely to be a single standardized air interface and networking infrastructure like 3G. Rather, 4G will include several different networking technologies, and this heterogeneous architecture

will interoperate through IP and the wireless application layer (WAL) in order to provide the best possible networking services wherever the user is located. The future 4G will likely include a large variety of different access networks. The terminals and base stations will be software radio-based. Because new radio access methods and software radio breakthroughs are anticipated, new transmission methods will probably be utilized. Finally, all seem to agree that 4G networking means providing an all-IP architecture and connectivity to anywhere at any time. Since almost every decade has brought forth a new mobile system, it is anticipated that this next generation system to provide completely new services may occur as early as 2010.

How quickly 4G is achieved will depend to a large extent on the amount of effort and money spent on developing such technologies. In 2000, only 8 organizations in the world were involved in 4G research and development. By 2003, the 4G community had expanded to over 2500 bodies worldwide. By 2008, it is estimated that over $400 billion will be invested in 4G mobile research and development, of which only $48 billion will be government funds.[25]

The scenarios described in the NPSTC functional requirements document identified above rely to a large extent on the availability of various broadband 4G networks, networks that are obviously not yet available and for which the standards have not yet been finalized. Many of these standards are in the process of being discussed and developed, with a number of organizations (including Project MESA) making contributions in this area. It is important to note that in a narrowband network (as we primarily have today), the scarce resource that needs to be allocated is bandwidth (computing resources for header compression are cheaper than the cost of the bandwidth). Conversely, in a broadband network, the scarce resource will revolve around routing (cost of bandwidth will be cheaper than the high-performance routing engine).[26]

3.2.2. COMPUTER TECHNOLOGY BASED WIRELESS MODEL

IMT-2000 systems are extremely important, but are not the only wireless data technologies available. Other important wireless Internet technologies have grown from the computer networking industry.

3.2.2.1. WIRELESS LOCAL AREA NETWORK (WLAN)

A wireless local area network (WLAN) is a network of computers or terminals connected by radio frequencies. Wireless LANs were conceived to complement fixed wired networks. Wireless access points (similar to traditional Ethernet hubs) provide access to devices that have wireless network interface cards. The access points connect to an Ethernet switch and are often configured on their own virtual private network (VPN). In general, wireless LANs cannot yet achieve speeds as fast as wired networks. The size of the antenna and the transmission power setting of the interface card determine the distance, data speed, and area between the transmitter and receiver.

The Institute of Electrical and Electronic Engineers (IEEE), a professional association of engineers which is headquartered in the U.S., is the most influential standards-making body in

the world of LANs. The IEEE 802 Standards Committee is responsible for LAN and MAN (metropolitan area network) standards. The Committee comprises several Working Groups, each focusing on particular areas. The 802.11 Working Group looks at WLAN standards. 802.11 is also called WiFi (short for wireless fidelity). Within the 802.11 Working Group are various Task Groups, some of which concentrate on air interface standards, some on security, and some on roaming and quality of service standards, among other things.

The European Telecommunications Standards Institute (ETSI) , a trade association, performs a similar role within the European Union.

WLANs generally operate in either the 2.4 GHz frequency band or the 5 GHz band. There are different (and sometimes multiple) air interface standards for each. In the 2.4 GHz band, the two primary WLAN air interface standards are 802.11b and 802.11g. In the 5 GHz band, the primary standard (in the U.S., but not approved in Europe) is 802.11a, with a new standard, 802.11h, on the horizon.

3.2.2.1.1. The 802.11a Air Interface Standard

The 802.11a standard is actually newer than the 802.11b standard, contrary to what its name implies. 802.11a operates in the license-exempt 5 GHz band, at speeds up to 54 Mpbs, using a modulation technique called Orthogonal Frequency Division Multiplexing (OFDM). OFDM uses multiple sub-carriers operating at slightly different frequencies. As there is more license-exempt spectrum available in the 5 GHz band than at 2.4 Ghz, it is possible to have more non-overlapping channels to increase coverage and density. In addition, there are currently fewer existing sources of interference. However, 802.11a systems generally have a shorter range than 802.11b systems because of power limits. In addition, because of certain limitations, this standard is not usable in Europe.

3.2.2.1.2. The 802.11b Air Interface Standard

802.11b is today's most widely-used wireless LAN technology. 802.11b equipment operates in the 2.4 GHz band using a modulation system known as Direct Sequence Spread Spectrum (DSSS), which is similar to that used by CDMA cellular systems. There are three non-overlapping channels available for 802.11b networks in the license-exempt 2.4 GHz band, which enables users to run up to three access points in the same physical area.

The theoretical maximum speed for 802.11b systems under perfect conditions is 11 Megabits per second (Mbps). The distance between the user and access point and/or the amount of background noise or interference can reduce this theoretical maximum in stages to as low as 1 Mbps. Also, because of system overheads, the actual data transfer speed will typically run at about 5-7 Mbps. Typical operational distances range from 50 to 300 feet indoors to well over 1,000 feet line-of-sight outdoors.

3.2.2.1.3. The 802.11g Air Interface Standard

The 802.11g standard, approved in 2003, allows operation at up to 54 Mbps in the 2.4 GHz band. The standard states that 802.11g devices must support existing 802.11b modulation techniques to ensure that an 802.11b client device can talk to an 802.11g access point and vice versa. For higher data rates, the standard specifies OFDM as the mandatory modulation technique. Thus, 802.11g is envisioned as the favored replacement for 802.11b systems, providing 802.11a speeds at 802.11b distances, while being backwards compatible.

3.2.2.1.4. The 802.11h Air Interface Standard

The 802.11h Task Group is working on a modified version of the 802.11a standard that would support Transmit Power Control (the ability to turn the power 'volume' up and down automatically) and Dynamic Frequency Selection (the ability to move to a different part of the frequency band to avoid interference). This is in response to European frequency regulations which require both in WLAN systems operating in the 5 GHz band. The European HiperLAN2 standard supported both of these features. With the development of 802.11h, vendor support for HiperLAN2 is shifting to the newer 802.11h.

3.2.2.1.5. WLAN Security Standards

802.11 Wired Equivalent Privacy (WEP) was envisioned as a combined access control, link privacy, and message integrity system for WLANs. However, during the development of WEP, certain compromises were made which have resulted in it being less secure than intended in all three areas. As a result, the 802.11i Task Group and the 802.1x Task Group have worked on improving both network security and authentication, respectively (the 802.1aa Task Group is currently working on revisions to 802.1x). 802.1x provides for port-based user authentication utilizing Remote Authentication Dial-In User Service (RADIUS) servers, and Extensible Authentication Protocol (EAP) as its authentication framework.

The 802.11i proposed standard includes an Enhance Security Network (ESN) that will use 802.1x to deliver its authentication and key management services. All stations and access points in an ESN will be required to contain an 802.1x port entity and an 802.11i authentication agent. In addition, the standard will require an authentication server that participates in the authentication of all mobile devices and access points. It may authenticate these devices itself, or it may provide information that the devices can use to authentication each other.

3.2.2.1.6. WLAN Roaming and Quality of Service Standards

For a mobile device to move seamlessly from one WLAN access point to another, the access points need to communicate with each other using an inter-access point protocol (IAPP). Because the original 802.11 standard did not specify this IAPP, most access point vendors implemented their own proprietary IAPPs with the result that seamless roaming between access

points form different vendors was generally not possible. 802.11f remedies this situation by specifying a standard IAPP for 802.11 networks.

802.11f does not address the issue of roaming between access points on different IP sub-nets, nor does it address the issue of roaming between networks belonging to different companies.

The 802.11e standard is being designed to make improvements to 802.11 to make it better able to handle voice traffic, MPEG video at up to 3 Mbps, and data streams of up to 10 Mbps. The goal is to minimize jitter and delay variations and to maximize access point throughput. This standard has not yet been approved.

3.2.2.2. PERSONAL AREA NETWORK (PAN)

As defined by Gehrmann et. al., a personal area network is "a collection of fixed, portable, or moving components within or entering a Personal Area, which form a Network through local interfaces. A Personal Area is a sphere around a person (stationary or in motion) with a typical radius of about 10 meters. The definition includes components that are carried, worn, or located near the body, e.g., personal digital assistants (PDAs)/handheld personal computers (HPCs), printers, microphones, speakers, headsets, bar code readers, sensors, displays, pagers, mobile phones, and smart cards."[27] The network interconnection is through local communications links such as short-range wireless connections (e.g., Bluetooth). PANs are also referred to as piconets, and sometimes as 'distributed dynamically configurable terminals.' All of the components in the PAN need not have communications capability to the global network but may be restricted to intra-PAN communications. Thus, the security model for a PAN has both a local and a global component for its requirements.

PANs using the Bluetooth standard operate in the 2.4 Ghz band. In this band, Bluetooth transmits voice and data at rates less than 1 Mbit per second. Bluetooth devices can function in either (or both) circuit switched mode, or packet switched mode (the mode for Internet data, as well as for higher bandwidth mobile communications systems, such as GPRS. A Bluetooth network (piconet) can allow the interconnection of eight devices in a radious of ten meters. As many as ten piconets can overlap to form a Scatternet, linking up to 80 Bluetooth devices.

The latest Class 1 Bluetooth standard will extend the range of Bluetooth devices to up to 100 meters, which opens it up to the same issues of security that have been raised regarding 802.11b.[28] According to Microsoft, Bluetooth has the following potential security risks:

- Bluetooth is designed to run over a short-range wireless peer-to-peer network. If one or more devices are used as gateways to other networks, and if the security of Bluetooth is compromised, it could expose the device or its attached networks.

- Bluetooth supports third party extensions. If these extensions do not use proper security and authentication procedures, they could compromise the security of a device or local network.[29]

Besides the above security issues, the Bluetooth version 1.1 standard does not support native IP, so makes support of TCP/IP and WLAN applications difficult.

Standards for Bluetooth are set by the Bluetooth Special Interest Group (SIG), an industry consortium. In addition to the Bluetooth SIG, IEEE has formed a group to set standards for wireless PANs (WPAN), the 802.15 working group, which has four Task Groups (described below). According to the IEEE web site: "The 802.15 WPAN™ effort focuses on the development of consensus standards for Personal Area Networks or short distance wireless networks. These WPANs address wireless networking of portable and mobile computing devices such as PCs, Personal Digital Assistants (PDAs), peripherals, cell phones, pagers, and consumer electronics; allowing these devices to communicate and interoperate with one another. The goal is to publish standards, recommended practices, or guides that have broad market applicability and deal effectively with the issues of coexistence and interoperability with other wired and wireless networking solutions."[30]

The IEEE Std 802.15.1™-2002 was approved as a new standard by the IEEE-SA Standards Board in April, 2002. The original version of 802.15.1 was based upon portions of the Bluetooth v1.1 Specification. The Bluetooth SIG is in the process of finalizing v1.2 of its specification. The IEEE 801.15.1a Task Group (Task Group 1, or TG1) will review the SIG changes and update 802.15.1. But, as the Task Group states, "802.15.1a will correspond to the changes made in the new Bluetooth Spec, but are to retain the IEEE-specific information included in the original Standard."[31]

The IEEE 802.15 Coexistence Task Group 2 (TG2) for Wireless Personal Area Networks is developing Recommended Practices to facilitate coexistence of Wireless Personal Area Networks™ (802.15) and Wireless Local Area Networks (802.11). The Task Group is developing a Coexistence Model to quantify the mutual interference of a WLAN and a WPAN™. The Task Group is also developing a set of Coexistence Mechanisms to facilitate coexistence of WLAN and WPAN™ devices.

The IEEE P802.15.3 High Rate (HR) Task Group (TG3) for Wireless Personal Area Networks (WPANs) is chartered to draft and publish a new standard for high-rate (20Mbit/s or greater) WPANs. Besides a high data rate, the new standard will provide for low power, low cost solutions addressing the needs of portable consumer digital imaging and multimedia applications. The approved draft standard will support data rates up to 55 Mbps.

The IEEE 802.15 TG4 is chartered to investigate a low data rate solution with multi-month to multi-year battery life and very low complexity. It is intended to operate in an unlicensed, international frequency band. Potential applications are sensors, interactive toys, smart badges, remote controls, and home automation. The draft standard 802.15.4 was approved in May, 2003. It supports data rates of 250 kbps, 40 kbps, and 20 kbps. It is designed to operate in 16 channels in the 2.4 GHz ISM band, 10 channels in the 915 MHz I band , and one channel in the 868 MHz band.

3.2.2.3. SUMMARY OF WLAN AND PAN STANDARDS

Wireless LAN Standards

The following chart was excerpted from webopedia.com.[32]

Standard	Data Rate	Modulation Scheme	Security	Pros/Cons
IEEE 802.11	Up to 2Mbps in the 2.4GHz band	FHSS or DSSS	WEP & WPA	This specification has been extended into 802.11b.
IEEE 802.11a (Wi-Fi)	Up to 54Mbps in the 5GHz band	OFDM	WEP & WPA	Products that adhere to this standard are considered "Wi-Fi Certified." Eight available channels. Less potential for RF interference than 802.11b and 802.11g. Better than 802.11b at supporting multimedia voice, video and large-image applications in densely populated user environments. Relatively shorter range than 802.11b. Not interoperable with 802.11b.
IEEE 802.11b (Wi-Fi)	Up to 11Mbps in the 2.4GHz band	DSSS with CCK	WEP & WPA	Products that adhere to this standard are considered "Wi-Fi Certified." Not interoperable with 802.11a. Requires fewer access points than 802.11a for coverage of large areas. Offers high-speed access to data at up to 300 feet from base station. 14 channels available in the 2.4GHz band (only 11 of which can be used in the U.S. due to FCC regulations) with only three non-overlapping channels.
IEEE 802.11g (Wi-Fi)	Up to 54Mbps in the 2.4GHz band	OFDM above 20Mbps, DSSS with CCK below 20Mbps	WEP & WPA	Products that adhere to this standard are considered "Wi-Fi Certified." May replace 802.11b. Improved security enhancements over 802.11. Compatible with 802.11b. 14 channels available in the 2.4GHz band (only 11 of which can be used in the U.S. due to FCC regulations) with only three non-overlapping channels.
Bluetooth	Up to 2Mbps in the 2.45GHz band	FHSS	PPTP, SSL or VPN	No native support for IP, so it does not support TCP/IP and wireless LAN applications well. Not originally created to support wireless LANs. Best suited for connecting PDAs, cell phones and PCs in short intervals.

HomeRF	Up to 10Mbps in the 2.4GHZ band	FHSS	Independent network IP addresses for each network. Data is sent with a 56-bit encryption algorithm.	**Note:** HomeRF is no longer being supported by any vendors or working groups. Intended for use in homes, not enterprises. Range is only 150 feet from base station. Relatively inexpensive to set up and maintain. Voice quality is always good because it continuously reserves a chunk of bandwidth for voice services. Responds well to interference because of frequency-hopping modulation.
HiperLAN/1 (Europe)	Up to 20Mbps in the 5GHz band	CSMA/CA	Per-session encryption and individual authentication.	Only in Europe. HiperLAN is totally ad-hoc, requiring no configuration and no central controller. Doesn't provide real isochronous services. Relatively expensive to operate and maintain. No guarantee of bandwidth.
HiperLAN/2 (Europe)	Up to 54Mbps in the 5GHz band	OFDM	Strong security features with support for individual authentication and per-session encryption keys.	Only in Europe. Designed to carry ATM cells, IP packets, Firewire packets (IEEE 1394) and digital voice (from cellular phones). Better quality of service than HiperLAN/1 and guarantees bandwidth.

Source: Webopedia.com. Last updated: June 26, 2003

3.2.2.4. BROADBAND WIRELESS ACCESS

The IEEE 802.16 Working Group on Broadband Wireless Access was formed to develop the WirelessMAN™ air interface standards for broadband wireless metropolitan area networks, as well as related compliance and coexistence standards. It has completed core specifications for fixed broadband wireless access from 2 to 66 GHz and is enhancing that work to support mobile and fixed terminals from the same base station.

WiMax is the new shorthand name for 802.16 (short for Worldwide Interoperability for Microwave Access). Unlike, 802.11a, WiMax was designed from the beginning to be compatible with European standards. The initial version of the 802.16 standard, approved in 2002, operates in the 10-to-66 GHz band and requires line-of-sight towers. The 801.16a extension, ratified in March, 2003, doesn't require line-of-sight transmission and allows the use of lower frequencies (2 to 11 GHz), many of which are unregulated. It claims a 31-mile range and 70 Mbps data transfer rates that can support thousands of users.

Additional 802.16 extensions are being worked, and will cover:

802.16b – Quality of Service
802.16c – Interoperability

802.16d – Fixing additional items not covered by 802.11c (the standard for developing access points)
802.116e – Support for mobile wireless technology (slow-speed, lightly mobile user). First draft targeted for July, 2004.

In December, 2002, the IEEE Standards Board approved the establishment of IEEE 802.20, the Mobile Broadband Wireless Access (MBWA) Working Group. The mission of IEEE 802.20 is to develop the specification for an efficient packet based air interface for Wide Area Networks (WAN) that is optimized for the transport of IP based services. The goal is to enable worldwide deployment of affordable, ubiquitous, always-on and interoperable multi-vendor mobile broadband wireless access networks that meet the needs of business and residential end user markets.

The scope of the MBWA is to specify the physical and medium access control layers of an air interface for interoperable mobile broadband wireless access systems, operating in licensed bands below 3.5 GHz, optimized for IP-data transport, with peak data rates per user in excess of 1 Mbps. It will support various vehicular mobility speeds up to 150 miles per hour (250 Km/h) in a MAN environment and will target spectral efficiencies, sustained user data rates and numbers of active users that are all significantly higher than achieved by existing mobile cellular systems. The ability to support handoffs between 802.20 networks and 802.11-based WLANs will likely be part of the specifications. The target date for approval of the new standard by the IEEE-SA Standards Board is 2005.

At the moment, it is unclear how the 4G wireless cellular efforts will coexist and/or interoperate with the MBWA standards under development.

3.2.2.5. MOBILE AD HOC NETWORK (MANET)

A Mobile Ad-hoc NETwork (MANET) is a collection of mobile nodes which communicate over radio and do not need any pre installed communication infrastructure. Communication can be performed if two nodes are close enough to exchange packets.

Specifically, MANET is a wireless network routing technology that enables users to maintain robust IP (Internet Protocol) communication links while moving around arbitrarily. Mobile ad hoc networks are envisioned to be self-forming, self-maintained, self-healing and will not require any existing infrastructure. While MANETs can be completely self contained, they can also be tied to an IP-based global or local network (e.g. Internet or private networks). These are referred to as Hybrid MANETs. There are numerous MANET protocols currently in existence and being developed, including Ad Hoc On-demand Distance Vector Routing (AODV), Temporally Ordered Routing Algorithm (TORA), DSR, and OLSR. Each protocol will evolve over time to better suit the dynamic nature of mobile ad hoc networks.

Nodes communicate with each other without the intervention of centralized access points or base stations. In such a network, each node acts both as a router and as a host. Due to the limited

transmission range of wireless network interfaces, multiple hops may be needed to exchange data between nodes in the network, which is why the literature sometimes uses the term multi-hop network for a MANET.

Some of the main restrictions of a MANET, to date, include:

Dynamic topology -Nodes are mobile and can be connected dynamically in an arbitrary manner. Links of the network vary over time and are based on the proximity of one node to another node. They are also subject to frequent disconnection during node's mobility.

Bandwidth constrained - Wireless links have significantly lower capacity than the wired links; they are affected by several error sources that result in degradation of the received signal and high bit error rate.

Energy constrained - Mobile nodes rely on battery power, which is a scarce resource; the most important system design criteria for optimization may be energy conservation.

Limited physical security - Mobility implies higher security risks than static operation because portable devices may be stolen or their traffic may cross insecurely wireless links. Eavesdropping, spoofing and denial-of-service attacks should be considered.

Autonomous - No centralized administration entity is required to manage the operation of the different mobile nodes.

Infrastructure-less and self operated

Much research is currently being done on the practical application and deployment of MANETs.

3.3. DEVICE VARIABILITY

True data interoperability will require not only interoperability across heterogeneous networks, but also across heterogeneous devices. In mobile commerce, a terminal or mobile device is used to communicate over the mobile telecommunications network. A myriad of such devices exist, including mobile phones, PDAs, smart phones (a combination of mobile phone and PDA), laptop computers, ear pieces (as part of a personal area network), and others. Each has its own specific characteristics and features which directly impact its usability. These include:

* Size and color of the display
* Input device (i.e., availability of keyboard and/or mouse)
* Memory
* CPU power
* Battery life
* Network connectivity/bandwidth capacity
* Supported operating system

* Availability of internal smart card reader

Depending upon the combination of factors present, the services that the end user can receive will differ considerably. For example, the ability of mobile phones, with their limited memory and CPU power, to process complex security algorithms (such as PKI - see below) is much more restricted than that of a laptop computer or even a PDA.[33] As a result, mobile phone manufacturers have utilized devices such as internal smartcards (e.g., Subscriber Identity Module [SIM] in mobile phones) to handle device authentication (see more on smartcards below).

Because the capabilities are so varied, to date this has required end user services to be customized for each type of device. And, as mentioned above, depending on the network technology used for transmission, the bandwidth capacity varies and will influence the kind of services that the end user is able to receive.

According to Schwiderski-Grosche and Knospe, "the heterogeneity of devices, operating systems, and network technologies is a challenge for a uniform end user platform. For this reason, standardisation bodies consisting of telecommunication companies, device manufacturers, and value-added service providers integrate their work."[34]

3.3.1. SMART CARDS

Smartcards may be a data storage device only or may contain a processor to execute some functions, even programs. Storage cards can't be written again after manufacture, so once their validity expires, they must be replaced with a new card, and inconvenient and costly proposition.

Processor cards allow the electronic storage of data, plus it functions as a mobile PC, powered and handled by another terminal. These functions can include security features to prevent unauthorized access (and therefore can be personalized). Processor cards generally contain RAM, ROM, EEPROM, and a CPU. This allows them to provide cryptographic functions, user authentication functions, and device authentication functions. The newer processor cards allow downloadable program code into the EEPROM (e.g., a Java applet).

Smartcards can serve as personal mobile storage devices, avoiding the necessity of storing identification data (e.g., PIN, biometric data) in a central database. Through on-card matching, the card itself can recognize its legal user because the verification process is incorporated into the card itself (e.g., electronic signatures).[35]

Smartcards can either be contact type cards or "contactless" cards. Contact cards, such as those used in mobile phones and ATMs, require the card to be put in direct contact with a card reader in order for the stored information to be exchanged. On the other hand, a contactless card can be accessed without direct contact (but the proximity between the card and the reader currently cannot exceed one meter). These are used where certain actions need to be performed quickly while passing by (e.g., access control systems, fare payment systems). Researchers expect the

future to have dual-contact cards that will provide both contact and contactless communications capabilities.[36]

As the technology evolves, it is expected that smartcards will be used in tandem. For example, an individual could carry a smartcard with them that contains all of their personal medical information. A doctor or other healthcare provider (such as an emergency medical technician) could use a smartcard with the appropriate authorization which would then allow the patient data to be accessed.

The Federal Identity and Credentialing Committee, chartered by OMB, recently released guidelines for developing interoperable federal identification systems based on smartcards. The policy would establish a common Federal ID card that could be used for both physical and logical assess control.[37]

3.4. NUMBER OF PUBLIC SAFETY DEVICES TO BE CONTROLLED

A national authentication database of fixed IP addresses would theoretically contain an IP address for every device used in the national public safety community. No census of the entire public safety device population exists, so in order to estimate the size of such a database (and the associated administrative activities that would be associated with its management, best estimates were made based upon available data.

According to statistics from BJS,[38] FEMA,[39] and PSWN,[40] there are approximately 2.3 million uniformed State and local law/fire/EMS personnel (this does not include all of the Federal first responder personnel who would also need to be included for a comprehensive estimate). Currently, let's assume that each of them averages 3 devices each (pager, cell phone, portable radio); some will have more, some less. At the current rate, that would mean 6.9 million devices (IP addresses) just to handle today's load. Conservatively, if that number were only to double in the next five years, we would be facing the administration of at least 13.8 million IP addresses.

Since these devices change and the personnel who use them change constantly, the administrative overhead for updating the profiles associated with these IP addresses in a centralized manner would be quite extensive. Regarding static IP addressing under IPv4, Ctek, Inc. concluded, "as a practical matter provisioning static IP addresses on wireless networks is a service order nightmare, compounded by roaming and network resale issues. Static addressing is completely orthogonal to the notion of 'push button' activations and will require additional layers of customer service if it is deployed."[41]

3.5. SECURITY (AUTHENTICATION, AUTHORIZATION, AND ENCRYPTION)

One of the primary concerns of the NPSTC requirements is that of security. The topic of security encompasses a huge number of topics, definitions, and issues. In this review, the focus

has been on the broad topics of authentication, authorization, and encryption. Each of these is first defined below and then further described individually.

3.5.1. DEFINITIONS

Authentication. Authentication refers to the ability to know the identity of an entity with some degree of assurance (i.e., *who* the entity is). The entity being authenticated may be a person, a network, a device, or all of the above.

Authorization. Authorization refers to determining what an entity is allowed (or authorized) to do, usually within an application. The specific *what* that an entity is allowed to do is often called the permissions or access privileges of the entity. Authorization usually occurs *after* the entity's identity has been authenticated.

Encryption. Encryption is a means of cryptographically changing (coding) plain text information such that it is unintelligible until it is cryptographically returned to its original state (decoding). All or only part of a message may be encrypted depending upon the sensitivity of the information being transmitted.

3.5.2. AUTHENTICATION

Authentication is a topic of immense concern to the Federal government. The Office of Management and Budget recently published a document, E-Authentication Guidance for Federal Agencies, that defines four levels of authentication in terms of the consequences of authentication errors and misuse of credentials.[42] Agencies are directed to complete a risk assessment and to map the identified risks to the required assurance level, at which time the agency can then select the appropriate technology to meet the minimum technical requirements for that level of assurance.

As a follow-up to that guidance, the National Institute of Standards and Technology (NIST) was directed to develop technical guidance for agencies in the implementation of electronic authentication (e-authentication). By NIST's definition, e-authentication is the remote authentication of individual people over a network, for the purpose of electronic government and commerce. The draft guidance has recently been published and is open for comment until March 15, 2004.[43]

The NIST draft technical guidance nicely summarizes the terminologies and issues involved in e-authentication. Since any authentication solution which the NPSTC Support Office might provide to NPSTC would need to adhere to this guidance, we are using the NIST terminology. The following sections about the E-authentication Model have been excerpted in their entirety from portions of sections 5 and 6 of that document.

E-AUTHENTICATION MODEL**.** In the guidance, the individual claiming an identity is called a *claimant* and the party verifying that identity is called a *verifier*. E-authentication begins with *registration*. Before an individual can claim an identity, he or she must demonstrate that the identity is a real identity, and that he or she is the person who is entitled to use that identity. This process is called *identity proofing*, and, in the lexicon of the guidance, identity proofing is performed by a *Registration Authority (RA)*, a trusted entity that registers individual *subscribers* with a *Credentials Service Provider (CSP)*. The CSP registers or gives the subscriber a *token* to be used in an authentication protocol and issues *credentials* as needed to bind that token to the identity, or to bind the identity to some other useful attribute. When a *claimant* successfully demonstrates possession and control of a token in an on-line authentication to a *verifier* through an *authentication protocol*, the verifier can establish the identity of the subscriber. A verifier can pass along an assertion about the identity or provide an attribute of the claimant to a *relying party*. The relying party can used the authenticated identity and other factors to make access control or authorization decisions.

Subscriber, RA and CSP. In the conceptual e-authentication model, a claimant in an authentication protocol must be a subscriber to some CSP. At some point, the subscriber registers with an RA, which verifies the identity of the subscriber, typically through the presentation of paper credentials and by records in databases. This process is called identity proofing. The RA, in turn, vouches for the identity of the subscriber to a CSP. The CSP registers or gives the subscriber a token to be used in an authentication protocol and issues credentials as needed to bind that token to the identity, or to bind the identity to some other useful attribute. The subscriber may be given electronic credentials to go with the token at the time of registration, or credentials may be generated later as needed.

There is always a relationship between the RA and CSP. In the simplest and perhaps the commonest case, the RA/CSP are separate functions of the same entity. However, an RA might be part of a company or organization that registers subscribers with an independent CSP, or several different CSPs. Therefore a CSP may have an integral RA, or it may have relationships with multiple independent RAs, and an RA may have relationships with different CSPs as well. (Another section of the draft guidance provides recommendations for the identity proofing and registration process.)

Tokens**.** Tokens are something that the claimant possesses and controls that may be used to authenticate the claimant's identity. In e-authentication, the claimant authenticates to a system or application over a network. Therefore, a token used for e-authentication shall include some secret information and it is important to provide security for the token. In fact, the three factors often considered as the cornerstone of authentication:

- Something you know (for example, a password)
- Something you have (for example, a cryptographic key or smart card)
- Something you are (for example, a voice print or other biometric)

influence the security provided by tokens. Tokens that incorporate all three factors are stronger than tokens that only incorporate one or two of the factors.

The secrets are often based on either *public key pairs* or *shared secrets*. *Public keys* are one half of a public key pair (also known as an asymmetric key), and the other half, a related *private key*, is used as a token. A verifier, knowing a claimant's public key through some credential (typically a *public key certificate*), can use an authentication protocol to verify the claimant's identity, by proving that the claimant has control of the associated private key token (*proof of possession*).

Shared secrets are either *symmetric keys* or passwords. In a protocol sense, all shared secrets are similar, and can be used in similar authentication protocols; however, passwords, since they are often committed to memory, are something the claimant knows, rather than something he has. Passwords, because they are committed to memory, usually do not have as many possible values as cryptographic keys, and, in many protocols, are vulnerable to network attacks that are impractical for keys. Moreover the entry of passwords into systems (usually through a keyboard) presents the opportunity for very simple keyboard logging or "shoulder surfing" attacks. Therefore keys and passwords demonstrate somewhat separate authentication properties (something you know rather than something you have) and passwords often have lesser resistance to network attacks. However, when using either public key pairs or shared secrets, the subscriber has a duty to maintain exclusive control of his token, since possession or control of the token is used to authenticate the subscriber's identity.

Biometrics are unique personal attributes that can be used to identify a person. They include facial pictures, fingerprints, DNA, iris and retina scans, voiceprints and many other things. The view of this guidance is that biometrics values should not be considered secrets in authentication processes, since the biometrics can often be observed, and since they are innate to the person and cannot be changed. Since they are not secrets, biometrics cannot serve as tokens for e-authentication. Therefore biometrics by themselves are of limited value in the remote e-authentication processes that are the subject of the guidance. Biometrics are valuable where the claimant is physically present at a reader controlled by the verifier, in registration processes to be able to later prove who actually registered and received credentials, and, in some cases, to unlock the keys of hardware tokens.

The guidance recognizes four kinds of claimant tokens: hard tokens, soft tokens, one-time password device tokens and password tokens.

Electronic Credentials. Paper credentials are documents that attest to the identity or other attributes of an individual or entity called the subject of the credentials. Some common paper credentials include passports, birth certificates, driver's licenses, and employee identity cards. The credentials themselves are authenticated in a variety of ways: traditionally perhaps by a signature or a seal, special papers and inks, high quality engraving, and today by more complex mechanisms, such as holograms, that make the credentials recognizable and difficult to copy or forge. In some cases, simple possession of the credentials is sufficient to establish that the physical holder of the credential is indeed the subject of the credentials. More commonly, the credentials contain biometric information such as the subject's description, a picture of the subject or the handwritten signature of the subject that can be used to authenticate that the holder of the credentials is indeed the subject of the credentials. When these paper credentials are presented in-person, authentication biometrics contained in those credentials can be checked to confirm that the physical holder of the credential is the subject.

Electronic identity credentials bind an identity and perhaps other attributes to a token. The recommendation does not prescribe particular kinds of electronic credentials. There are a variety of electronic credential types in use today, and new types of credentials are constantly being created. Electronic credentials may be general purpose credentials or targeted to a particular verifier. Some common types of credentials are:

- X.509 public key identity certificates that bind an identity to a public key;
- X.509 attribute certificates that bind an identity or a public key with some attribute;
- Kerberos tickets that are encrypted messages binding the holder with some attribute or privilege.

Electronic credentials may be stored as directory objects. These credentials may be signed objects (e.g., X.509 certificates), in which case they are self-authenticating. In this case, the directory may be an untrusted entity. Alternatively, the directory may be a trusted entity. When the directory is trusted, credentials may simply be stored as a directory entry.

Verifiers**.** In any authenticated on-line transaction, the verifier must verify that the claimant has possession and control of the token that verifies his or her identity. A claimant authenticates his or her identity to a verifier by the use of a token and an authentication protocol. This is called *Proof of Possession (PoP).* Many PoP protocols are designed so that a verifier, who has no knowledge of the

token before the authentication protocol is run, learns nothing about the token from the run. It is undesirable for verifiers to learn shared secrets unless they are also the CSP who registered the tokens. The verifier and CSP may be the same entity, the verifier and relying party may be the same entity or they may all three be separate entities. Where the verifier and the relying party are separate entities, the verifier must convey the result of the authentication protocol to the relying party. The object created by the verifier to convey this result is called an assertion.

Assertions**.** Assertions can be used to pass information about the claimant or the e-authentication process from the verifier to a relying party. Assertions support the claimant's identity but are not bound to the token possessed by the claimant. A relying party trusts an assertion based on the source, the time of creation, and attributes associated with the claimant.

Examples of assertions include:

- SAML assertions, specified using a mark up language intended for describing security assertions, can be used by a verifier to make a statement to a relying party about the identity of a claimant.
- Cookies, character strings placed in a web browser's memory, are available to websites with the same domain name as the entity that placed them in the web browser. Cookies may simply be tickets or may contain pointers to verified credentials.

Assertions may be stored as directory objects. Where assertions are signed objects (e.g., signed SAML assertions), they are self-authenticating. Alternatively, the directory may be a trusted entity. When the directory is trusted, unsigned assertions may be accepted based on the source.

The Relying Party**.** A relying party relies on results of an on-line authentication to establish the identity or attribute of a subscriber for the purpose of some transaction. The verifier and the relying party may be the same entity, or they may be separate entities. If they are separate entities, the relying party receives an assertion from the verifier. The relying party is responsible to validate that the received assertion came from a verifier trusted by the relying party. Where the assertions indicate time of creation or attributes associated with the claimant, the relying party is also responsible for verifying this information.

TOKENS**.** The guidance recognizes four kinds of claimant tokens for e-authentication. Each type of token incorporates one or more of the factors (something you know, something you have, and something you are.) Tokens that provide a higher level of assurance incorporate two or more of the factors. The four kinds of tokens include:

· Hard token – a hardware device that contains a protected cryptographic key. Authentication is accomplished by proving possession of the device and control of the key. Hard tokens shall:
    - o require the entry of a password or a biometric to activate the authentication key;
    - o not be able to export authentication keys;
    - o be FIPS 140-2 validated.

· Soft token – a cryptographic key that is typically stored on disk or some other media. Authentication is accomplished by proving possession and control of the key. The soft token shall be encrypted under a key derived from a password known only to the user, so knowledge of a password is required to activate the token. The cryptographic module used with the soft token shall be validated at FIPS 140-2 Level 1 or higher. Each authentication shall require entry of the password and the unencrypted copy of the authentication key shall be erased after each authentication.

· One-time password device token - a personal hardware device that generates "one time" passwords for use in authentication. The device may or may not have some kind of integral entry pad, an integral biometric (e.g., fingerprint) reader or a direct computer interface (e.g., USB port). The passwords shall be generated by using a FIPS approved block cipher or hash algorithm to combine a symmetric key stored on a personal hardware device with a nonce to generate a one-time password. The nonce may be a date and time, or a counter generated on the device, or a challenge sent from the verifier (if the device has an entry capability). The device shall be validated at FIPS 140-2 Level 1 or higher. The one-time password typically is displayed on the device and manually input (direct electronic input from the device to a computer is also allowed) to the verifier and as a password.

· Password token – a secret character string that a claimant memorizes and uses to authenticate his or her identity.

Impersonation of an identity using a hard or soft token requires that the impersonator obtain two separate things: either the key (token) and a password, or the token and the ability to enter a biometric into the token. Therefore both hard and soft tokens provide more assurance than passwords by themselves normally provide. Moreover, a hard token is a physical object and its theft is likely to be noticed by its owner, while a soft token can sometimes be copied without the owner being aware. Therefore a hard token offers more assurance than a soft token.

> Impersonation of an identity using a password token requires only that the impersonator obtain the password. The ability of humans to remember long, arbitrary passwords is limited, so password tokens are often vulnerable to a variety of attacks including guessing, dictionaries of commonly used passwords, and simple exhaustion of all possibilities. There are a wide variety of password authentication protocols that differ significantly in their vulnerabilities, and many password mechanisms are vulnerable to passive and active network attacks. While some cryptographic password protocols resist nearly all direct network attacks, these techniques are not at present widely used and all password authentication mechanisms are vulnerable to keyboard loggers and observation of the password when it is entered. Therefore password tokens generally provide less assurance than hard or soft tokens.[44]

In conjunction with the draft technical guidance, Federal agencies can evaluate their IT systems in terms of OMB's proscribed levels of assurance. To summarize, the four levels of assurance are:

> Level 1 – requires no identity proofing and allows a wide range of authentication technologies and tokens, including a simple personal ID number. There is no requirement for Federal Information Processing Standard (FIPS)-approved cryptography.
>
> Level 2 – requires some identity proofing and at least a password as a token. FIPS-approved cryptography is required to thwart eavesdropping or hacker attacks.
>
> Level 3 – requires a high level of identity proofing and FIPS-approved cryptography to protect the authentication token as well as prevent eavesdropping or attacks. Tokens can be either software or hardware.
>
> Level 4 – provides the highest practical remote network authentication assurance. It is similar to Level 3 but requires hardware tokens with cryptographic modules validated at FIPS 140-2 Level 2 or higher, providing robust two-factor remote authentication.

The General Services Administration has been in the process of developing an e-authentication infrastructure that will serve as a model and architecture for Federal agencies that need to authenticate users of e-government services. This project, the E-Authentication project, is part of the government's Quicksilver initiatives. Initially envisioned as a centralized e-authentication gateway through which all Federal agencies would connect,[45] the project has since been shifted to taking a federated approach to authenticating users.[46] In the federated model, agencies will depend on third-party credential providers to validate transactions between agencies and the public. This model is more in line with the way it is done in the commercial sector, as promoted by an industry consortium called the Liberty Alliance.[47, 48] As summarized by George Schu, vice president of VeriSign, Inc., "the idea of building one megagateway for all credentials would have been hard to pull off because technically it was too complicated."[49]

The draft technical guidance establishes requirements that Federal IT systems and service providers participating in authentication protocols be authenticated to subscribers. However, it does not address machine-to-machine authentication, nor provide requirements for issuing authentication credentials and tokens to machines and servers when they are used in e-authentication protocols with people. Neither does the guidance provide any technical recommendations for authorization.

Although not mentioned in the guidance, users can be authenticated via the host name or IP address of their machine. As one author described IP address-based authentication, "while this authentication mechanism could be deployed on a small scale in a closed system, it is just impractical in an open system."[50] From a security point of view, this technique of authentication is vulnerable to IP spoofing,[51] a technique used to gain unauthorized access to computers, whereby the intruder sends messages to a computer with an IP address indicating that the message is coming from a trusted host. To engage in IP spoofing, a hacker must first use a variety of techniques to find an IP address of a trusted host and then modify the packet headers so that it appears that the packets are coming from that host.

IP Spoofing is a problem without an easy solution, since it's inherent to the design of the TCP/IP suite. While some of types spoofing attacks are no longer often seen, such as session hijacking for host-based authentication services, IP spoofing is still prevalent in network scanning and probes, as well as denial of service attacks.

To defend against spoofing, precautions can either be taken through filtering at the network router (using access control lists), or by the more elaborate authentication mechanisms described in the guidance, in conjunction with encryption. Both authentication and encryption features are included in Ipv6, which should help eliminate current spoofing threats.

3.5.3. AUTHORIZATION

Authentication simply establishes identity, not what that identity is authorized to do or what access privileges he or she has; this is a separate decision. Relying parties, typically government agencies, who rely on the authenticated identity of a party, will use that authenticated identity and other factors to make access control or authorization decisions. In many cases, the authentication process and services will be shared by many applications and agencies, but the individual agency or application is the relying party that must make the decision to grant access or process a transaction based on the specific application requirements. The NIST guidance provides technical recommendations for the process of authentication, but not for authorization.

In the case of public safety, a myriad of applications and database exist, each with different authorization mechanisms and requirements. Each application has its own security rules and authorization process that must be complied with in order to gain access to its data. Each state provides access to its own state level information as well as to the Federal databases and other states (via the NLETS switch). However, at the local level, little has been done to consolidate access to local databases and applications. Within a single jurisdiction, it is possible to have

multiple authorization processes which must be traversed in order to access the needed information (e.g., one for CAD, one for Police Records Management, one for Fire Records Management, etc.).

3.5.3.1. SINGLE SIGN-ON (SSO)

As IT systems proliferate to support business and government processes, users and system administrators are faced with an increasingly complicated interface to accomplish their job functions. Users typically have to sign-on to multiple systems, necessitating an equivalent number of sign-on dialogues, each of which may involve different user names and authentication information. System administrators are faced with managing user accounts within each of the multiple systems to be accessed in a co-ordinated manner in order to maintain the integrity of security policy enforcement.

In a typical legacy system, an end-user has to identify and authenticate himself independently to each of the domains with which he wishes to interact. The end user interacts initially with a Primary Domain to establish a session with that primary domain, generally by supplying a set of user credentials applicable to the primary domain, for example a username and password. The primary domain session is typically represented by an operating system session shell executed on the end user's workstation. From this primary domain session shell the user is able to initiate the services of the other domains, such as platforms or applications.

To access the services of a secondary domain, an end user is required to perform a Secondary Domain Sign-on. This requires the end user to supply a further set of user credentials applicable to that secondary domain. An end user has to conduct a separate sign-on dialogue with each secondary domain that the end user wants to use. The secondary domain session is typically represented by an operating system shell or an application shell.

From the management perspective, the legacy approach requires independent management of each domain and the use of multiple user account management interfaces. Considerations of both usability and security have given rise to a need to co-ordinate and where possible integrate user sign-on functions and user account management functions for the multitude of different domains now found within an enterprise.

A service that provides such coordination and integration can provide real cost benefits to an enterprise through:

* reduction in the time taken by users to sign-on to individual domains, including reducing the possibility of such sign-on operations failing.
* improved security through the reduced need for a user to handle and remember multiple sets of authentication information.
* reduction in the time taken, and improved response, by system administrators in adding and removing users to the system or modifying their access rights.

* improved security through the enhanced ability of system administrators to maintain the integrity of user account configurations including the ability to inhibit or remove an individual user's access to all system resources in a co-ordinated and consistent manner.

Such a service has been termed Single Sign-On (SSO) based on the end user's perception of the impact of this service. However, both the end user and management aspects of the service are equally important. In the single sign-on approach, as part of the primary sign-on the system is required to collect from the user all the identification and user credential information necessary to support the authentication of the user to each of the secondary domains that the user may potentially require to interact with. The information supplied by the user is then used by Single Sign-On Services within the primary domain to support the authentication of the end user to each of the secondary domains with which the user actually requests to interact. From a management perspective the single sign-on model provides a single user account management interface through which all the component domains may be managed in a coordinated and synchronized manner.

In order to implement such a service, significant security issues of the Single Sign-On model must be resolved. Most importantly, the secondary domains have to trust the primary domain to:

1) correctly assert the identity and authentication credentials of the end user, and

2) protect the authentication credentials used to verify the end user identity to the secondary domain from unauthorized use. In other words, the authentication credentials have to be protected when transferred between the primary and secondary domains against threats arising from interception or eavesdropping leading to possible masquerade attacks.

The task of developing a standard for single sign-on applications has been undertaken by the Open Group, an international vendor and technology-neutral consortium. Their scope for the Single Sign-On Standard (code-named XSSO at the present) is to define services in support of:

* the development of applications to provide a common, single end-user sign-on interface for an enterprise, and

* the development of applications for the co-ordinated management of multiple user account management information bases maintained by an enterprise.[52]

However, support for single sign-on <u>across</u> enterprise system boundaries is not considered to be within the current scope of XSSO.

Several vendors now offer products aimed towards providing single sign-on capabilities within a single enterprise. However, to date, no universal national single sign-on (SSO) mechanism for public safety has been developed, nor is any group currently working on developing such a

scheme nor the data and security standards which would be necessary for it to function given the huge number of different government and commercial entities involved.

3.5.3.2. SAML STANDARD

For web-based single sign-on, an additional specification of importance is called SAML (short for Security Assertions Markup Language. Using SAML, the user's authentication, authorization, profiles, and preferences are transmitted from an original source service provider to subsequent destination service providers selected by the user during the session.

The SAML 1.0 specification was approved by the Organization for the Advancement of Structured Information Standards (OASIS) in March, 2002. A related Java-based application programming interface (API) standard for SAML is being reviewed by the Java Community Process program.

SAML is an XML (extensible markup language) framework for exchanging security information over the internet. It resides within a system's security mechanisms to enable exchange of identity and entitlement with other services. SAML does not define the mechanism by which authentication or authorization takes places. Rather, it defines the structure of the documents that transport security information between services.

SAML depends upon a concept called an assertion. An assertion is a declaration of a known fact about a subject (e.g., a user or a code). For example, this might be the fact that an application has been granted a certain class of access to a resource under certain specific conditions. The applications define and exchange assertions using a variety of request/response protocols.

In the current specification, there are four types of assertions:

* Attribute assertions
* Authentication assertions
* Authorization assertions
* Subject assertions

Assertions are only generated by a SAML "issuing authority" which can be either of two types: third-party security service providers, or individual businesses. The latter depends upon using Liberty Alliance technologies for establishing trusted private 'federations' of partners.[53]

3.5.4. ENCRYPTION

Data that travels across unsecured networks is potentially vulnerable to anyone who has the ability to intercept it. Encryption provides a means to scramble and protect data as it travels across otherwise unsecured networks. Different levels of encryption provide increasingly secure data protection.

Almost all methods of encryption rely on two basic constructs: codes and keys. A mathematical code, or algorithm, is structured so that only those with the correct keys to the equation can properly code and decode messages. The mathematics can be so complex that without knowing both the encryption code and the right key, it is essentially impossible to determine the original message.

Because they work in tandem, the second critical piece of encryption is the distribution and protection of keys. There are a number of different key exchange mechanisms. For example, in the Diffie-Hellman Key Exchange allows an authorized party to obtain sessions keys without any third party obtaining them, even when the keys are exchanged over unsecure links. The exchange is secure because the keys are never transmitted in clear text, and they are exceptionally difficult to decipher.

Encryption can be implemented at any of three OSI layers: the application layer, the network layer, or the data link layer. Application layer encryption requires each application to support encryption, and all hosts that communication with the application must speak the same encryption language. Implementing such a scheme could require replacing all the hosts in a network, but it does not necessarily require any network upgrades because network traffic is unaffected.

Implementing encryption at the network layer can be done anywhere in the network, and the hosts to not need to be upgraded. As a result, it provides a reasonable balance of security versus cost. Lastly, data link layer encryption is potentially the most secure since it encrypts everything (including the IP address). However, the downside is that each router must decrypt the traffic at every link in the network and then re-encrypt it once the appropriate path is determined, which can result in a significant slowdown in overall throughput.

The government, private industry, and other organizations contribute to the vast collection of standards on cryptography. A few of these are ISO, ANSI, IEEE, NIST, and IETF. The Data Encryption Standard (DES) was the first official U.S. government cipher intended for commercial use. DES is the most widely used cryptosystem in the world.

DES is defined in Federal Information Processing Standard (FIPS) 46-3, which describes the data encryption algorithm (DEA). The DEA is also defined in the ANSI standard X3.92. The National Institute of Standards and Technology (NIST) has re-certified DES every five years. FIPS 46-3 reaffirms DES usage as of October 1999, but single DES is permitted only for legacy systems. FIPS 46-3 includes a definition of triple-DES (3DES), which is now the FIPS preferred symmetric algorithm. Within a few years, DES and triple-DES will be replaced with the Advanced Encryption Standard (AES).

The Digital Signature Standard (DSS) is the digital signature algorithm (DSA) developed by the U.S. National Security Agency (NSA) to generate a digital signature for the authentication of electronic documents. DSS was put forth by the NIST in 1994, and has become the United States government standard for authentication of electronic documents. There are four key

properties of written signatures: authentication of the signer's identity, acceptance of the terms stated in the signed document, proof of the integrity of the document's contents, and nonrepudiation (in other words, the signer cannot deny he or she has signed). Digital signatures are electronic signatures that address all these properties.

DSS is specified in Federal Information Processing Standard (FIPS) 186. The DSA is a pair of large numbers that are computed according to the specified algorithm within parameters that enable the authentication of the signatory, and as a consequence, the integrity of the data attached. Digital signatures are generated through DSA, as well as verified. Signatures are generated in conjunction with the use of a private key; verification takes place through incorporation of a corresponding public key. Each signatory has their own paired public (assumed to be known to the general public) and private (known only to the user) keys. Because a signature can only be generated by an authorized person using their private key, the corresponding public key can be used by anyone to verify the signature.

4. SUMMARY OF FINDINGS

The NPSTC requirements identified in section 2 were broken down into their component parts and compared to the five issues above to determine which issues are addressing which requirements. As broken down, the requirements needed to satisfy future needs of local, national, and international public safety are as follows:

1. Support seamless roaming between public and private data systems
2. Support rapid transmission of data with minimum of data or transmission errors
3. Support heterogeneous networks
4. Support establishment of ad-hoc mobile networks
5. Utilize/develop standards for associated technologies
6. Solutions should be interoperable across local, national, and international boundaries
7. Maintain user privileges during roaming
8. Provide multiple levels of security and encryption
9. Support seamless interoperation and integration of multiple agencies' applications
10. Expand access to existing public safety information systems
11. Support heterogeneous devices

Each of these requirements have been grouped together below based upon common requirements. Each grouping includes a list of the main issues which need to be addressed in order to meet these requirements, a projected time frame within which it is expected that applicable solutions will be commercially available, and an evaluation of whether an government provided IP-based authentication process is required in order to ensure that this solution occurs.

4.1. NEXT GENERATION NETWORKS

The NPSTC public safety requirements which fall under this general category include:

1. Seamless roaming between public and private data systems
2. Support rapid transmission of data with minimum of data or transmission errors
3. Support heterogeneous networks
4. Support establishment of ad-hoc mobile networks
5. Utilize/develop standards for associated technologies
6. Solutions should be interoperable across local, national, and international boundaries

Achieving the vision of the next generation networks will require all of the following to some extent:

Migration to IPv6 (e.g., home/care-of addressing, autoconfiguration)
Further IPv6 extension to support better handover during roaming (i.e., micro-mobility)
Quality of Service technology enhancements (e.g., MPSL, RSVP, etc.)
Software defined radio based networks (e.g., supporting multiple interchangeable frequencies)
Convergence of cellular telephone network model and computer network model
Development of next generation network standards (e.g., 802.20, 4G)
Smart antenna enhancements
Enhancements to existing standards (e.g., various 802.11 task groups, 3GPP, IMT-2000, 802.15, 802.16)
Continued development of MANET specifications

Current Status:

IPv6 standard has been finalized. Migration has begun in Europe and Asia. In U.S., the Army is looking to begin migration in 2006.

IETF Working groups are addressing the necessary extensions to IPv6 for micro-mobility, the enhancements for Quality of Service, and 802.20, additional enhancements to 802.11, and MANET. ITU is working on areas for "beyond IMT-2000."

Manufacturers and researchers are working on development of software defined radio and smart antenna enhancements.

The U.S. and European standards bodies have been cooperating more closely than in the past. It appears that this trend will continue, particularly in light of the desire for global roaming in the next generation networks.

Estimated time frame for commercial availability:

No sooner than 2010 for the next generation backbone, although parts of the technologies will be deployed sooner.

Government-provided national IP-based authentication process needed to accomplish this:

No.

4.2. INTEGRATED SECURITY AND APPLICATIONS INTEROPERABILITY

The NPSTC public safety requirements which fall under this general category include:

7. Maintain user privileges during roaming
8. Provide multiple levels of security and encryption
9. Support seamless interoperation and integration of multiple agencies' applications
10. Expand access to existing public safety information systems

Achieving the NPSTC requirements for security and applications interoperability will require all of the following to some extent:

Migration to IPv6 (i.e., IPsec built-in and mandatory; globally unique home address, not just national; built-in authentication and encryption features)
Continued enhancements to security standards (e.g., IPsec; 802.11i/802.11aa/EAP)
Enhancements to XSSO standard, to include single sign-on across enterprise boundaries
Development of single sign-on protocols and data standards for public safety across all government levels
Develop (if needed) standardized public safety-specific security assertions

Current Status:

IPv6 standard has been finalized. Migration has begun in Europe and Asia. In U.S., the Army is looking to begin migration in 2006.

The Open Group is only working on within-enterprise SSO at the moment.

No group is addressing public safety single sign-on protocols and data standards to facilitate intergovernmental application interoperability.

No group is evaluating the need for standardized public safety-specific security assertions.

Estimated time frame for commercial availability:

Migration to IPv6 by commercial ISPs worldwide is projected to be 50% complete by 2006.
No time frame for SSO standards across enterprise boundaries
No time frame for development of public safety SSO protocols and data standards
No time frame for evaluation of need for or development of public safety security SAML-compliant assertions

Government-provided national IP-based authentication process needed to accomplish this:

No.

4.3. HETEROGENEOUS DEVICES

Solutions needed to accomplish:

Migration to IPv6 (i.e., autoconfiguration; built-in authentication and encryption features
Incorporation of IPv6 addresses into all devices
Battery life enhancements for portable units
Memory enhancements for portable units
ALU/CPU enhancements for small, portable units
Software radio based (e.g., supporting multiple interchangeable frequencies)
Support for multiple standards and network models (e.g., both cellular based and computer network based)
Internal authentication/authorization/cryptography mechanism enhancements

Current Status:

IPv6 standard has been finalized. Migration has begun in Europe and Asia. In U.S., the Army is looking to begin migration in 2006.

Development of new features and standards is vendor driven.

Ongoing efforts by researchers and manufacturers to increase capabilities of portable units

Manufacturers and researchers are working on development of software defined radio capability

Estimated time frame for commercial availability:

No specific time frame exists for all of these. Enhancements for most will likely be incremental in nature (with the exception of software based radio)

The SDR Forum expects that widespread adoption of software radio will take place by 2006 or 2007.

Government-provided national IP-based authentication process needed to accomplish this?

No.

5. RECOMMENDATIONS

As summarized in this extract from the Vision Statement for IST-FET's Global Computing Initiative, "the envisioned systems are highly dynamic: physical devices are mobile, connectivity and bandwidth are changing, computational processes and data can migrate, and applications come and go. The availability and responsiveness of the resources that are active in an application at any given point in time are unpredictable and difficult to control. The scale of the systems is expected to be extremely large, in number of components and, in some cases, in

geographic area spanned. The design issues involved in the construction of systems that can be configured in such environments pose enormous challenges to computer science, calling for focused research on the foundational problems that need to be solved."[54]

In summary, there is no easy solution to achieving the public safety goals as envisioned by NPSTC. The public safety requirements for mobile interconnectivity cut across multiple technological and logistical areas. As discussed above, no single technology change or action is sufficient to accomplish the complex requirements of public safety. Fortunately, the provision of mobile data, that for many years was only of interest to the public safety community, has now become viable for the general population. As a result, various commercial interests and researchers are actively pursuing solutions to many of the issues identified above.

The initial impetus for this study was to evaluate the feasibility of implementing a national IP-based interconnectivity authentication process. To function, such a system would require that all of the devices utilized be IP addressable. One of the first questions to be asked was whether such a process was practical, either now or in the future. If such a system were practical, would it be feasible, or were there other alternative options available? And ultimately, would it provide a solution to any of the requirements identified by NPSTC?

5.1. NATIONAL IP-BASED AUTHENTICATION PROCESS

The questions were looked at from the perspective of both the current version of IP (IPv4) and the successor version (IPv6), since the answers might be different in each case. In other words, might a national IP-based authentication process benefit public safety in the short run, even if it would not be necessary in the long run. Based upon the research conducted, development of an IPv4-based national authentication process is not recommended due to the following:

1) Subject matter experts, as well as the Federal e-authentication project, recommends against the concept of a large, centralized authentication process,
2) The number of devices that would need to be centrally managed is huge and will only get larger over time, with a significant administrative overhead.
3) A large number of public safety devices which are not yet IP addressable, and so would not be able to use the IP-based process.
4) IPv6 will be extensively available within a fairly short time frame.
5) None of the above NPSTC requirements depend upon the development of such a process in order to be implemented.

Looking at the question from an IPv6 viewpoint, development of a national IPv6-based authentication process is also not recommended based upon the following:

1) Subject matter experts, as well as the Federal e-authentication project, recommends against the concept of a large, centralized authentication process,
2) The number of devices that would need to be centrally managed is huge and will only get larger over time, with a significant administrative overhead.

3) IPv6 has built-in features which address a number of the issues for which the national process was proposed, including built-in authentication and encryption features, incorporation of home/care-of addressing, plus the ability to support globally unique IP addressing.
4) None of the NPSTC requirements depend upon the development of such a process in order to be implemented.

Therefore, since it was determined that development of a national IP-based authentication process would not be practical nor feasible, no further study actions were taken (i.e., no budget estimate was calculated) in this area. However, the study did identify two areas of interest for which further study and/or action could be taken. These are identified below.

5.2. SINGLE SIGN-ON FOR PUBLIC SAFETY APPLICATIONS

While most of the component requirements are currently being addressed by the global research and commercial communities, two requirements in particular are not adequately addressed:

* Support for seamless interoperation and integration of multiple agencies' applications
* Expanding access to existing public safety information systems

As mentioned above, the necessary actions to address these areas are not being addressed by any group at the moment. Therefore, further research and/or action could be taken as follows:

1. Work with local, state, and Federal agencies to develop the application protocols and data standards which be necessary to support true applications interoperability for public safety across all government levels.

2. Work with local, state, and Federal agencies to develop the security protocols and standards which be necessary to support single sign-on capability for public safety across all government levels.

3. Determine if any standardized public safety-specific SAML-compliant credentials and/or assertions need to be developed for national authorization; if so, develop them.

5.3. NEXT GENERATION NETWORK DEVELOPMENT SUPPORT

Members of the public safety community from around the world have been actively involved in Project MESA, whose goal is to "develop advanced mobile broadband technical specifications that can be used to support the communications requirements of the PPDR [public protection and disaster relief] community."[55] For the most part, the requirements identified in section 4.1 as common to next generation networks are those that are also in common with Project MESA. Rather than develop an entirely new group and/or process for monitoring progress in these areas, it is recommended that NPSTC support and monitor Project MESA in these requirement areas.

REFERENCES

The following references are specifically cited in the text of this report. Other reference material in addition to that cited here was also reviewed. A full list of all 180 references reviewed is available by request.

1. Powell, J.: *Operational requirements for national IP-based public safety interconnectivity: Statement of functional requirements*. NLECTC internal document, 2003.

2. Project MESA web site: http://www.projectmesa.org/

3. Aghvami, H. and Jafarian, B.: *A vision of UMTS/IMT-2000 evolution*. Electronics & Communication Engineering Journal, June, 2000.

4. Aghvami and Jafarian, above.

5. Shah, A.: *TCP performance over wireless links*. Stanford University, November, 2001.

6. Aghvami and Jafarian, above.

7. Ladid, L.: *IPv6 – The next-generation Internet*. Ericsson Review, No. 1, 2000.

8. NEC: *NEC announces successful demonstration of beyond-3G IP-based converged mobile architecture*. Press Release, November 12, 2003, http://www.neceurope.com.

9. http://www.dhcp-handbook.com/dhcp_faq.html#widxx

10. Broadbeam Corporation: *Six Things An IT Manager Should Know About GPRS*. White paper, June, 2003, http://www.broadbeam.com

11. Ctek, Inc.: *Dynamic IP Addressing in M2M Applications*. White Paper, 2003, http://www.ctekproducts.com/

12. Ibid.

13. Hagen, S.: *IPv6: Revitalizing the Internet revolution*. The O'Reilly Network, September 25, 2002, http://www.oreillynet.com.

14. Ibid.

15. Nortel Networks: *IPv6: Building the foundation of the 3G wireless Internet*. White paper, May, 2001. http://www.nortelnetworks.com

16. Narten, T.: *Neighbor discovery and stateless autoconfiguration in IPv6*. IEEE Internet Computing, July/August, 1999.

17. Nortel Networks, above.

18. Narten, above.

19. Nortel Networks, above.

20. Einsiedler, H., Aguiar, R., Jahnert, J., Jonas, K., Liebsch, M., Schmitz, R., Pacyna, P., Gozdecki, J., Papir, Z., Moreno, J., and Soto, I.: *The Moby Dick project: A mobile heterogeneous all-IP architecture*. Advanced Technologies, Applications and Market Strategies for 3G ATAMS 2001, Kraków, Poland, June 17-20, 2001.

21. Nortel Networks, above.

22. Hagen, above.

23. Ericsson: *Ericsson conducts the world's first IPv6 over 3G UMTS/WCDMA network demonstration*. Press release, January 31, 2004.

24. Becchetti, L., Delli Priscoli, F., Inzerilli, T., Mahonen, P., and Munoz, L.: *Enhancing IP service provision over heterogeneous wireless networks: A path toward 4G*. IEEE Communications Magazine, August, 2001.

25. Fourth Generation Mobile Forum web site, http://www.4gmf.org.

26. Nortel Networks, above.

27. Gehrmann, C., Kuhn, T., Nyberg, K. and Windirsch, P.: *Trust model, communication and configuration security for Personal Area Networks*. Proceedings from SHAMAN Conference, 2000. http://www.ist-shaman.org

28. Jaques, R.: *Bluetooth security 'crisis' looming*. Vnunet.com, December 17, 2003, http://www.vnunet.com.

29. Microsoft web site, http://msdn.microsoft.com/library/default.asp?url=/library/en-us/wcebluet/html/ceconbluetoothsecurityissues.asp

30. IEEE web site, http://www.ieee802.org

31. Ibid.

32. http://www.webopedia.com/quick_ref/WLANStandards.asp

33. Dankers, et al, above.

34. Schwiderski, S. and Knospe, H.: *Secure mobile commerce*. Electronics & Communication Engineering Journal, October, 2002.

35. Scheuermann, D.: *The smartcard as a mobile security device*. Electronics & Communication Engineering Journal, October, 2002.

36. Ibid.

37. Jackson, W.: *Government releases guidelines for governmentwide smart cards*. Government Computer News, February 14, 2004.

38. Bureau of Justice Statistics web site: http://www.ojp.usdoj.gov/bjs/

39. Federal Emergency Management Agency (U.S. Fire Administration) web site: http://www.fema.gov/

40. Public Safety Wireless Network: http://www.pswn.gov.

41. Ctek, Inc., above.

42. Bolten, J.B.: *E-Authentication guidance for federal agencies*. Memorandum to the heads of all departments and agencies. Executive Office of the President, Office of Management and Budget. December 16, 2003, http://www.whitehouse.gov/omb/memoranda/fy04/m04-04.pdf

43. Burr, W., Polk, W., and Dodson, D.: *Draft Recommendation for Electronic Authentication.* National Institute of Standards and Technology, NIST Special Publication 800-63, January, 2004.

44. Ibid.

45. GSA: E-authentication Gateway. Various presentations, November, 2002. http://www.cio.gov/eauthentication

46. Miller, J.: *New authentication plan takes shape*. Government Computer News, November, 2003.

47. Liberty Alliance Project: *Introduction to the Liberty Alliance identity architecture*. White paper, March, 2003. http://www.projectliberty.org

48. Landau, S. and Hodges, J.: *A brief introduction to Liberty*. Liberty Alliance White paper, February, 2003. http://www.projectlibrary.org

49. Miller, above.

50. Claessens, J., Preneel, B., and Vandewalle, J.: *A tangled world wide web of security issues*. First Monday, March, 2002, http://firstmonday.org

51. Bellovin, S.: *Security problems in the TCP/IP protocol suite*. Computer Communication Review, April, 1989.

52. The Open Group web site: http://www.opengroup.org/security/sso/sso_scope.htm

53. Byous, J.: Single sign-on simplicity with SAML. Sun Microsystems white paper, May 9, 2002. http://java.sun.com/features/2002/single-signon.html.

54. Information Societies Technologies, Future and Emerging Technologies web site: http://www.cordis.lu/ist/fet/home.html

55. Project MESA: *Project MESA: An update, Making progress toward an international PPDR standard*. White paper, September, 2003. http://www.projectmesa.org

ACRONYMS AND ABBREVIATIONS

3GPP - Third Generation Partnership Project
6WINIT - IPV6 wireless internet initiative

AAA - Authentication, authorization, and accounting
ALU - Arithmetic and logic unit
ANSI - American National Standards Institute
AODV - Ad hoc on-demand distance vector routing
API - Application programming interface
ATM - Asynchronous transfer mode
ATM - Automatic teller machine
AVL - Automatic vehicle location

BER - Bit error rate
BJS - Bureau of Justice Statistics

CAD - Computer aided dispatch
CCK - Complementary code keying
CDMA - Code division multiple access
CDPD - Cellular digital packet data
CIDR - Classless interdomain routing
CPU - Central processing unit
CSMA/CA - Carrier sense multiple access/collision avoidance
CSP - Credentials service provider

DES - Data encryption standard
DEA - Data encryption algorithm
DHCP - Dynamic host configuration protocol
DMV - Department of Motor Vehicles
DNS - Domain naming service
DSR - Data signaling rate
DSS - Digital signature standard
DSSS - Direct sequence spread spectrum

EAP - Extensible authentication protocol
EDGE - Enhanced data rates for global evolution
EEPROM - Electrically erasable programmable read-only memory
EMS - Emergency medical service
ESN - Enhance security network
ETSI - European Telecommunications Standards Institute

FCC - Federal Communications Commission

FEMA - Federal Emergency Management Administration
FIPS - Federal information processing standard
FHSS - Frequency hopping spread spectrum
FMIPv6 - Fast mobile internet protocol version 6

GPRS - General packet radio service
GTP - GPRS tunneling protocol
GSM - Global system for mobile communications

HMIPv6 - Hierarchical mobile Internet protocol version 6
HPC - Handheld personal computer

IAPP - Inter-access point protocol
IETF - Internet Engineering Task Force
IMT-2000 - International Mobile Telecommunications-2000
IEEE - Institute of Electrical and Electronic Engineers
IP - Internet protocol
IPng - Internet protocol next generation
IPsec - Internet protocol security
IPv4 - Internet protocol version 4
IPv6 - Internet protocol version 6
ISO - International Standards Organization
ISP - Internet service provider
IST - Information Society Technologies
ITU - International Telecommunications Union

M2M - Machine-to-machine
M2M - Mobile-to-mobile
MAN - Metropolitan area network
MANET - Mobile ad hoc network
MBWA - Mobile broadband wireless access
MESA - Mobility for emergency and safety applications
MPEG - Moving pictures experts group
MPSL - Multi protocol label switching

NAT - Network address translation
NCC - National Coordination Committee
NIJ - National Institute of Justice
NIST - National Institute of Standards and Technology
NLECTC - National Law Enforcement and Corrections Technology Center
NLETS - National Law Enforcement Telecommunications System
NPSTC - National Public Safety Telecommunications Council
NSA - National Security Agency

OASIS - Organization for the Advancement of Structured Information Standards
OFDM - Orthogonal frequency division multiplexing
OLSR - Optimized link state routing
OMB - Office of Management and Budget
OSI - Open systems interconnection

PAN - Personal area network
PDA - Personal digital assistant
PIN - Personal identification number
PKI - Public key infrastructure
PoP - Proof of possession
PPDR - Public protection and disaster relief
PSWN - Public Safety Wireless Network

QoS - Quality of service

RA - Registration authority
RADIUS - Remote authentication dial-in user service
RAM - Random access memory
RF - Radio frequency
ROM - Read only memory
RSVP - Resource reservation protocol

SAML - Security assertions markup language
SDR - Software defined radio
SIG - Special interest group
SIM - Subscriber identity module
SMS - Short message service
SSL - Secure socket layer
SSO - Single sign-on

TCP - Terminal control protocol
TDMA - Time division multiple access
TORA - Temporally ordered routing algorithm

UDP - User datagram protocol
UMTS - Universal mobile telecommunications system
UMTX - Ultra miniature FM room transmitter

VPN - Virtual private network

WAL - Wireless application layer
WCDMA - Wireless code division multiple access

WEP - Wireless equivalent privacy
WiFi - Wireless fidelity
WiMax - Wireless interoperability for microwave access
WLAN - Wireless local area network
WPA - Wireless protected access
WPAN - Wireless personal area network

XML - Extensible markup language
XSSO - Single sign-on standard