

Network Working Group
Request for Comments: 3457
Category: Informational

S. Kelly
Airespace
S. Ramamoorthi
Juniper Networks
January 2003

Requirements for IPsec Remote Access Scenarios

Status of this Memo

This memo provides information for the Internet community. It does not specify an Internet standard of any kind. Distribution of this memo is unlimited.

Copyright Notice

Copyright (C) The Internet Society (2003). All Rights Reserved.

Abstract

IPsec offers much promise as a secure remote access mechanism. However, there are a number of differing remote access scenarios, each having some shared and some unique requirements. A thorough understanding of these requirements is necessary in order to effectively evaluate the suitability of a specific set of mechanisms for any particular remote access scenario. This document enumerates the requirements for a number of common remote access scenarios.

Table of Contents

1. Introduction	2
1.1 Requirements Terminology	3
1.2 Reader Prerequisites	3
1.3 General Terminology	4
1.4 Document Content and Organization	4
2. Overview	5
2.1 Endpoint Authentication	6
2.1.1 Machine-Level Authentication	7
2.1.2 User-Level Authentication	7
2.1.3 Combined User/Machine Authentication	8
2.1.4 Remote Access Authentication	8
2.1.5 Compatibility With Legacy Remote Access Mechanisms	9
2.2 Remote Host Configuration	10
2.3 Security Policy Configuration	11
2.4 Auditing	12
2.5 Intermediary Traversal	13

3. Scenarios	13
3.1 Telecommuters (Dialup/DSL/Cablemodem)	14
3.1.1 Endpoint Authentication Requirements	15
3.1.2 Device Configuration Requirements	16
3.1.3 Policy Configuration Requirements	17
3.1.4 Auditing Requirements	18
3.1.5 Intermediary Traversal Requirements	18
3.2 Corporate to Remote Extranet	19
3.2.1 Authentication Requirements	19
3.2.2 Device Configuration Requirements	20
3.2.3 Policy Configuration Requirements	21
3.2.4 Auditing Requirements	21
3.2.5 Intermediary Traversal Requirements	21
3.3 Extranet Laptop to Home Corporate Net	22
3.3.1 Authentication Requirements	22
3.3.2 Device Configuration Requirements	23
3.3.3 Policy Configuration Requirements	23
3.3.4 Auditing Requirements	24
3.3.5 Intermediary Traversal Requirements	24
3.4 Extranet Desktop to Home Corporate Net	25
3.4.1 Authentication Requirements	25
3.4.2 Device Configuration Requirements	26
3.4.3 Policy Configuration Requirements	26
3.4.4 Auditing Requirements	26
3.4.5 Intermediary Traversal Requirements	26
3.5 Public System to Target Network	27
3.5.1 Authentication Requirements	27
3.5.2 Device Configuration Requirements	28
3.5.3 Policy Configuration Requirements	28
3.5.4 Auditing Requirements	29
3.5.5 Intermediary Traversal Requirements	29
4. Scenario Commonalities	29
5. Security Considerations	30
6. References	30
7. Acknowledgements	30
8. Editors' Addresses.	30
9. Full Copyright Statement	31

1. Introduction

Until recently, remote access has typically been characterized by dial-up users accessing the target network via the Public Switched Telephone Network (PSTN), with the dial-up connection terminating at a Network Access Server (NAS) within the target domain. The protocols facilitating this have usually been PPP-based, and access control, authorization, and accounting functions have typically been provided using one or more of a number of available mechanisms, including RADIUS [RADIUS].

Note that for such access, it has often been assumed that the communications infrastructure supporting the ISP connection (the PSTN) is relatively secure, and poses no significant threats to communications integrity or confidentiality. Based on this assumption, connection security has been limited to access control at the NAS based on username/passphrase pairs. In reality, PSTN dialup connections have never been impervious to a determined adversary.

The availability of widespread broadband access, in concert with the desire to reduce the cost of PSTN toll access, have driven the development of Internet-based remote access mechanisms. In some cases, PPP-based tunneling mechanisms have been used to provide remote access by allowing the dial user to first access a local ISP account, and then tunnel an additional PPP connection over the Internet into the target network. In the case of broadband users, such connections are tunneled directly over the Internet. While these mechanisms have been lacking in terms of security features, the increasing availability of IPsec renders it possible to provide more secure remote access to the remote resources via the Internet.

Remote access via the Internet has numerous benefits, financial and otherwise. However, security is paramount, and this presents strong incentives for migration from the old dial-up model to a more secure IPsec-based remote access model. Meeting the security requirements of various classes of remote access users presents a number of challenges. It is the aim of this document to explore and enumerate the requirements of various IPsec remote access scenarios, without suggesting particular solutions for them.

1.1 Requirements Terminology

The keywords MUST, MUST NOT, REQUIRED, SHALL, SHALL NOT, SHOULD, SHOULD NOT, RECOMMENDED, MAY, and OPTIONAL, when they appear in this document, are to be interpreted as described in [3].

1.2 Reader Prerequisites

Reader familiarity with RFCs 2401-2412 is a minimum prerequisite to understanding the concepts discussed here. Familiarity with concepts relating to Public Key Infrastructures (PKIs) is also necessary. Familiarity with RADIUS, PPP, PPTP, L2F, L2TP, and other remote access support protocols will also be helpful, though not strictly necessary.

1.3 General Terminology

- o Remote Access - this term is used to refer to the case in which the remote user does not necessarily reside at a fixed location, i.e., in which the user's IP address is not fixed, and therefore usually not known prior to connection establishment.
- o Secure Remote Access - this term refers to remote access which is secured using elements of the IPsec protocol suite.
- o IPsec Remote Access Client (IRAC)- this term is used to refer to the remote access user's system.
- o IPsec Remote Access Server (IRAS) - this term refers to the device providing access to the target network. An alternative term is "Security Gateway".
- o Security GateWay (SGW) - this refers to the device providing access to the target network. An alternative term is IRAS.
- o Virtual IP address (VIP) - this term describes an address from a subnet local to the target network which is assigned to a remote client, giving the appearance that the remote client actually resides on the target network.
- o Machine-Level Authentication - this term describes the case where the identity of a machine is verified by virtue of the machine's possession and application of some combination of authenticators. For a more complete definition, see section 2.
- o User-Level Authentication - this term describes the case where the identity of a user (as opposed to that of a machine) is verified by virtue of the user's possession and application of some combination of authenticators. For a more complete definition, see section 2.
- o NAT - Network Address/Port Translation

1.4 Document Content and Organization

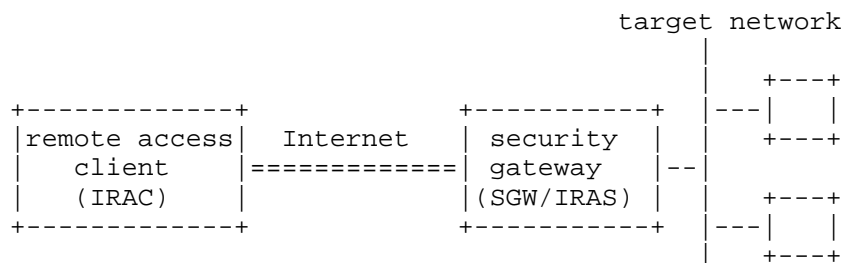
This document, while initially intended to simply outline requirements for various remote access scenarios, has come to include somewhat more than this. This document has evolved from discussion within the IPsec Remote Access (IPSRA) working group. As a result, it in some respects documents the evolution of this thought process. While this represents a departure from the typical form of a

requirements document, the associated historical context should prove useful in interpreting the conclusions reached in the IPSRA working group.

The balance of this document is organized as follows: First, there is a general overview of the basic requirements categories, including definitions relevant to these categories. Following this is a section devoted to each remote access scenario. Within each of these sections there are subsections detailing requirements specific to that scenario in each of the following areas: endpoint authentication, remote host configuration, policy configuration, auditing, and intermediary traversal.

2. Overview

In a very general sense, all secure remote access scenarios have a similar high-level appearance:



In all cases, a remote client wishes to securely access resources either behind a SGW or on an IPsec-protected host, and/or wishes to provide other (specific) systems with secure access to the client's own resources. There are numerous details which may differ, depending on the particular scenario. For example, the IRAC may be within another corporate network, or connected to an ISP via dialup, DSL, or CATV media. There may be additional intermediaries between the remote client and the security gateway, but ultimately, all of these configurations may be viewed somewhat equivalently from a high level.

In general, there are several basic categories of requirements relevant to secure remote access scenarios, including endpoint authentication, remote host configuration, security policy configuration, auditing, and intermediary traversal. Endpoint authentication refers to verification of the identities of the communication partners (e.g., the IRAC and the IRAS). Remote host configuration refers to the device configuration parameters of the IRAC system. Security policy configuration refers to IPsec policy configuration of both the security gateway and the remote host, and

might also be termed "access control and authorization configuration". Auditing refers to the generation and collection of connection status information which is required for the purpose of maintaining the overall security and integrity of the connected networks. Intermediary traversal refers to the ability to pass secured traffic across intermediaries, some of which may modify the packets in some manner. Such intermediaries include NAT and firewall devices. These various categories are treated in more detail below.

2.1 Endpoint Authentication

Before discussing endpoint authentication with respect to remote access, it is important to distinguish between data source authentication and end user authentication. Data source authentication in the IPsec context consists in providing assurance that a network packet originates from a specific endpoint, typically a user, host, or application. IPsec offers mechanisms for this via AH or ESP. End user authentication within the IPsec context consists in providing assurance that the endpoint is what or who it claims to be. IPsec currently offers mechanisms for this as part of IKE [IKE].

While the two types of authentication differ, they are not unrelated. In fact, data source authentication relies upon endpoint authentication, because it is possible to inject packets with a particular IP address into the Internet from many arbitrary locations. In many instances, we cannot be certain that a packet actually originates from a particular host, or even from the network upon which that host resides. To resolve this, one must first authenticate the particular endpoint somehow, and then bind the addressing information (e.g., IP address, protocol, port) of this endpoint into the trust relationship established by the authentication process.

In the context of secure remote access, the authenticated entity may be a machine, a user (application), or both. The authentication methods currently supported by IPsec range from preshared secrets to various signature and encryption schemes employing private keys and their corresponding public key certificates. These mechanisms may be used to authenticate the end user alone, the device alone, or both the end user and the device. These are each discussed in more detail below.

2.1.1 Machine-Level Authentication

In the case where no user input is required in order for an authentication credential to be used, the entity authenticated will primarily be the device in which the credential is stored and the level of derived assurance regarding this authentication is directly related to how securely the machine's credential is maintained during both storage and use. That is, a shared secret or a private key corresponding to a public key certificate may be either stored within the device or contained in another device which is securely accessible by the device (e.g., a smartcard). If the knowledge required for the use of such authentication credentials is entirely contained within the subject device (i.e., no user input is required), then it is problematic to state that such credential usage authenticates anything other than the subject device.

In some cases, a user may be required to satisfy certain criteria prior to being given access to stored credentials. In such cases, the level of user authentication provided by the use of such credentials is somewhat difficult to derive. If sufficiently strong access controls exist for the system housing the credential, then there may be a strong binding between the authorized system user and the credential. However, at the time the credential is presented, the IRAS itself has no such assurance. That is, the IRAS in isolation may have some level of assurance that a particular device (the one in which the credential resides) is the one from which access is being attempted, but there is no explicit assurance regarding the identity of the user of the system. In order for the IRAS to derive additional assurance regarding the user identity, an additional user credential of some sort would be required. This is discussed further below.

2.1.2 User-Level Authentication

In some cases, the user may possess an authentication token (preshared key, private key, passphrase, etc.), and may provide this or some derivative of this whenever authentication is required. If this token or derivative is delivered directly to the other endpoint without modification by the IRAC system, and if the IRAC system provides no further credentials of its own, then it is the user alone which has been authenticated. That is, while there may be some assurance as to the network address from which the user is originating packets, there is no assurance as to the particular machine from which the user is attempting access.

2.1.3 Combined User/Machine Authentication

To authenticate both the user and the system, user input of some sort is required in addition to a credential which is securely stored upon the device. In some cases, such user input may be used in order to "complete" the credential stored on the device (e.g., a private key is password-encrypted), while in others the user's input is supplied independently of the stored credential. In the case where the passphrase is applied to the credential prior to use, the level of assurance derived from successful application of the credential varies according to your viewpoint.

From the perspective of a system consisting of user, IRAC, IRAS, and a collection of system protections and security procedures, it may be said that the user has been authenticated to an extent which depends upon the strength of the security procedures and system protections which are in place. However, from the perspective of the IRAS alone, there is little assurance with respect to user identity. That is, schemes requiring that stored credentials be modified by user input prior to use may only be said to provide user-level authentication within the context of the larger system, and then, the level of assurance derived is directly proportional to the weakest security attribute of the entire system.

When considering remote access from a general perspective, assumptions regarding the overall system are liable to prove incorrect. This is because the IRAS and the IRAC may not be within the same domain of control; extranet scenarios are a good example of this. Hence, the most desirable joint user/machine authentication mechanisms in this context are those which provide a high level of assurance to both the IRAS and the IRAC, independently of the larger system of which the user, IRAS, and IRAC are a part.

2.1.4 Remote Access Authentication

In the general case for remote access, authentication requirements are typically asymmetric. From the IRAC's perspective, it is important to ensure that the IRAS at the other end of the connection is indeed what it seems to be, and not some rogue system masquerading as the SGW. That is, the IRAC requires machine-level authentication for the IRAS. This is fairly straightforward, given the authentication mechanisms supported by IKE and IPsec. Further, this sort of authentication tends to persist through time, although the extent of this persistence depends upon the mechanism chosen.

While machine-level authentication for the IRAS is sufficient, this is not the case for the IRAC. Here, it is often important to know that the entity at the other end of the connection is one who is

authorized to access local resources rather than someone who happened upon an unoccupied but otherwise authorized system, or a malicious Trojan horse application on that user's system, or some other unauthorized entity. Authenticating the user presents different requirements than authenticating the user's machine; this requires some form of user input, and often the authentication must be periodically renewed.

In situations where a high level of physical security does not exist, it is common to require a user-input secret as part of the authentication process, and then to periodically renew the authentication. Furthermore, since such circumstances may include the possibility of the presence of a Trojan horse application on the IRAC system, one-time passphrase mechanisms are often advisable. Choosing passphrase mechanisms and renewal intervals which provide an acceptable level of risk, but which do not annoy the user too much, may be challenging. It should be obvious that even this approach offers limited assurance in many cases.

Clearly, there are variable assurance levels which are attainable with the various endpoint authentication techniques, and none of the techniques discussed offer absolute assurance. Also, there are variations in the authentication requirements among different remote access scenarios. This means there is no "cookie cutter" solution for this problem, and that individual scenarios must be carefully examined in order to derive specific requirements for each. These are examined on a case by case basis below in the detailed scenario descriptions.

2.1.5 Compatibility With Legacy Remote Access Mechanisms

There are a number of currently deployed remote access mechanisms which were installed prior to the deployment of IPsec. Typically, these are dialup systems which rely upon RADIUS for user authentication and accounting, but there are other mechanisms as well. An ideal IPsec remote access solution might utilize the components of the underlying framework without modification. Inasmuch as this is possible, this should be a goal. However, there may be cases where this simply cannot be accomplished, due to security and/or other considerations. In such cases, the IPsec remote access framework should be designed to accommodate migration from these mechanisms as painlessly as is possible.

In general, proposed IPsec remote access mechanisms should meet the following goals:

- o should provide direct support for legacy user authentication and accounting systems such as RADIUS

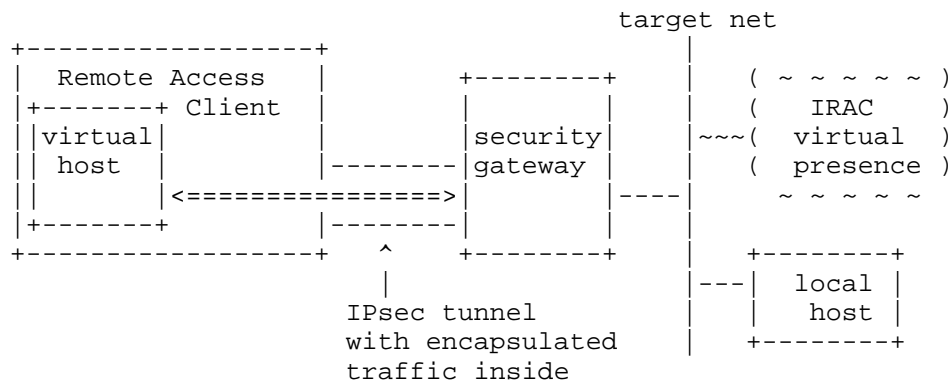
- o should encourage migration from existing low-entropy password-based systems to more secure authentication systems
- o if legacy user authentication support cannot be provided without some sort of migration, the impact of such migration should be minimized
- o user authentication information must be protected against eavesdropping and replay (including the user identity)
- o single sign-on capability should be provided in configurations employing load-balancing and/or redundancy
- o n-factor authentication mechanisms should be supported

2.2 Remote Host Configuration

Remote host configuration refers to the network-related device configuration of the client system. This configuration may be fixed or dynamic. It may be completely provided by the administrator of the network upon which the remote user currently resides (e.g., the ISP), or it may be partially provided by that administrator, with the balance provided by an entity on the remote corporate network which the client is accessing. In general, this configuration may include the following:

- o IP address(es)
- o Subnet mask(s)
- o Broadcast address(es)
- o Host name
- o Domain name
- o Time offset
- o Servers (e.g., SMTP, POP, WWW, DNS/NIS, LPR, syslog, WINS, NTP, etc.)
- o Router(s)
- o Router discovery options
- o Static routes
- o MTU
- o Default TTL
- o Source routing options
- o IP Forwarding enable/disable
- o PMTU options
- o ARP cache timeout
- o X Windows options
- o NIS options
- o NetBIOS options
- o Vendor-specific options
- o (other options)

Cases where such configuration is fixed are uninteresting; it is the cases where specific IRAC configuration occurs as a result of remote access with which we are concerned. For example, in some cases the IRAC may be assigned a "virtual address", giving the appearance that it resides on the target network:



In this case, the IRAC system begins with an externally routable address. An additional target network address is assigned to the IRAC, and packets containing this assigned address are encapsulated, with the outer headers containing the IRAC's routable address, and forwarded to the IRAS through the tunnel. This provides the IRAC with a virtual presence on the target network via an IPsec tunnel. Note that the IRAC now has two active addresses: the ISP-assigned address, and the VIP.

Having obtained this virtual presence on the corporate network, the IRAC may now require other sorts of topology-related configuration, e.g., default routers, DNS server(s), etc., just as a dynamically configured host which physically resides upon the target network would. It is this sort of configuration with which this requirements category is concerned.

2.3 Security Policy Configuration

Security policy configuration refers to IPsec access policies for both the remote access client and the security gateway. It may be desirable to configure access policies on connecting IRAC systems which will protect the target network. For example, since a client has access to the Internet (via its routable address), other systems on the Internet also have some level of reciprocal access to the client. In some cases, it may be desirable to block this Internet

access (or force it to pass through the tunnel) while the client has a tunneled connection to the target network. This is a matter of client security policy configuration.

For the security gateway, it may also be desirable to dynamically adjust policies based upon the user with which a connection has been established. For example, say there are two remote users, named Alice and Bob. We wish to provide Alice with unrestricted access to the target network, while we wish to restrict Bob's access to specific segments. One way to accomplish this would be to statically assign internal "virtual" addresses to each user in a one-to-one mapping, so that each user always has the same address. Then, a particular user's access could be controlled via policies based upon the particular address. However, this does not scale well.

A more scalable solution for remote client access control would be to dynamically assign IP addresses from a specific pool based upon the authenticated endpoint identity, with access to specific resources controlled by address-based policies in the SGW. This is very similar to the static mapping described above, except that a given group of users (those with identical access controls) would share a given pool of IP addresses (those which are granted the required access), rather than a given user always mapping to a given address. However, this also has scaling issues, though not as pronounced as for the static mapping.

Alternatively, an arbitrary address could be assigned to a user, with the security gateway's policy being dynamically updated based upon the identity of the remote client (and its assigned virtual address) to permit access to particular resources. In these cases, the relevant security policy configuration is specific to the IRAS, rather than to the IRAC. Both IRAS and IRAC security policy configuration are encompassed by this requirements category.

2.4 Auditing

Auditing is used here to refer to the collection and reporting of connection status information by the IRAS, for the purpose of maintaining the security and integrity of the IRAS protected network. For remote access, the following auditing information is useful from a security perspective:

- o connection start time
- o connection end time

Note that the requirement for a connection-end-time attribute implies the need for a connection heartbeat mechanism of some sort so that the IRAS can accurately determine this quantity in cases where the

IRAC does not explicitly terminate the connection. Also note that the heartbeat mechanism in this case is always directed from the IRAC to the IRAS.

In some cases, use of a heartbeat may negatively influence a connection. For example, if the heartbeat interval is very short, and the connection is reset after loss of very few heartbeat packets, there is a possibility that network congestion could lead to unnecessary connection resets. The heartbeat interval and reset threshold should be chosen with this in mind, and it should be possible to adjust these quantities either through configuration or negotiation.

2.5 Intermediary Traversal

Intermediary traversal is used here to refer to passing a secured data stream through an intermediary such as a firewall or NAT device. In the case of firewalls, numerous deployed products do not recognize the IPsec protocol suite, making it difficult (sometimes impossible) to configure them to pass it through. In such cases, a mechanism is required for making the data stream appear to be of a type which the firewall is capable of managing.

In the case of NAT devices, there are a number of issues with attempting to pass an encrypted or authenticated data stream. For example, NAT devices typically modify the source IP address and UDP/TCP port of outgoing packets, and the destination IP address and UDP/TCP port of incoming packets, and in some cases, they modify additional fields in the data portion of the packet. Such modifications render the use of the AH protocol impossible. In the case of ESP, the UDP/TCP port fields are sometimes unreadable and always unmodifiable, making meaningful translation by the NAT device impossible. There are numerous other protocol-field combinations which suffer similarly. This requirements category is concerned with these issues.

3. Scenarios

There are numerous remote access scenarios possible using IPsec. This section contains a brief summary enumeration of these, followed by a subsection devoted to each which explores the various requirements in terms of the categories defined above.

The following scenarios are discussed:

- o dialup/dsl/cablemodem telecommuters using their systems to access remote resources

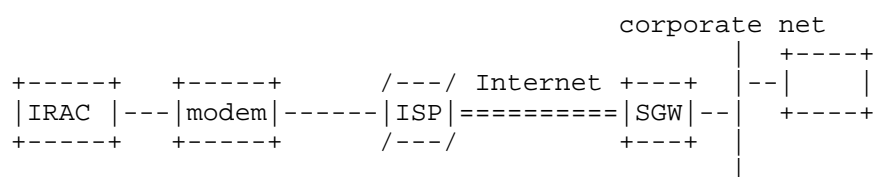
- o extranet users using local corporate systems to access the remote company network of a business partner
- o extranet users using their own system within another company's network to access their home corporate network
- o extranet users using a business partner's system (located on that partner's network) to access their home corporate network
- o remote users using a borrowed system (e.g., an airport kiosk) to access target network resources

3.1 Telecommuters (Dialup/DSL/Cablemodem)

The telecommuter scenario is one of the more common remote access scenarios. The convenience and wide availability of Internet access makes this an attractive option under many circumstances. Users may access the Internet from the comfort of their homes or hotel rooms, and using this Internet connection, access the resources of a target network. In some cases, dialup accounts are used to provide the initial Internet access, while in others some type of "always-on" connection such as a DSL or CATV modem is used.

The dialup and always-on cases are very similar, with two significant differences: address assignment mechanism and connection duration. In most dialup cases, the IRAC's IP address is dynamically assigned as part of connection setup, and with fairly high likelihood, it is different each time the IRAC connects. DSL/CATV users, on the other hand, often have static IP addresses assigned to them, although dynamic assignment is on the increase. As for connection duration, dialup remote access connections are typically short-lived, while always-on connections may maintain remote access connections for significantly longer periods of time.

The general configuration in either case looks like this:



An alternative to this configuration entails placing a security gateway between the user's system and the modem, in which case this added SGW becomes the IRAC. This is currently most common in cases where DSL/CATV connections are used.

3.1.1 Endpoint Authentication Requirements

The authentication requirements of this scenario depend in part upon the general security requirements of the network to which access is to be provided. Assuming that the corporate SGW is physically secure, machine authentication for the SGW is sufficient. If this assumption regarding physical security is incorrect, it is not clear that stronger authentication for the SGW could be guaranteed, and derivation of an effective mechanism for that case is beyond the scope of this document.

For the IRAC, there are numerous threats to the integrity of the user authentication process. Due to the open nature of common consumer operating systems, some of these threats are quite difficult to protect against. For example, it is very difficult to assert, with any level of certainty, that a single user system which permits the downloading and running of arbitrary applications from the Internet has not been compromised, and that a covert application is not monitoring and interacting with the user's data at any point in time.

However, there are 2 general threats we might realistically hope to somehow mitigate with appropriate authentication mechanisms if we can assume that the system has not been compromised in this manner. First, there is the possibility that a secure connection is established for a particular user, but that someone other than the intended user is currently using that connection. Second, there is the possibility that the user's credential (password, hardware token, etc.) has been somehow compromised, and is being used by someone other than the authorized user to gain access.

Mitigation of the first threat, the possibility that someone other than the authorized user is currently using the connection, requires periodic renewal of user authentication. It should be clear that machine authentication will not suffice in this case, and that requiring periodic re-entry of an unchanging user password (which may be written on a post-it note which is stuck to the user's monitor) will have limited effectiveness. Convincing verification of the continued presence of the authorized user will, in many cases, require periodic application of a time-variant credential.

Mitigation of the second threat, credential compromise, is difficult, and depends upon a number of factors. If the IRAC system is running a highly secure operating system, then a time-variant credential may again offer some value. A static password is clearly deficient in this scenario, since it may be subject to either online or offline guessing, and eventually compromised - which is the threat we are attempting to mitigate. However, if the IRAC operating system is not

hardened, the use of a time-variant credential is only effective if simultaneous access from more than one location is forbidden, and if the credential generation mechanism is not easily compromised.

A second approach to the credential compromise problem entails using a PKI-based credential which is stored within a secure container of some sort, and which requires some user interaction prior to operation (e.g., a smartcard). If such a credential requires periodic user interaction to continue operating (e.g., pin re-entry), this may help to limit the access of an unauthorized user who happens upon a connected but unattended systems. However, choosing an acceptable refresh interval is a difficult problem, and if the pin is not time-variant, this provides limited additional assurance.

To summarize, the following are the authentication requirements for the IRAS and IRAC:

IRAS

- o machine authentication MUST be provided.

IRAC

- o support for user authentication SHOULD be provided
- o support for either user or machine authentication MUST be provided
- o support for user authentication MUST be provided if protection from unauthorized connection use is desired.
- o if user authentication is provided for short-lived dialup connections, periodic renewal MAY occur
- o if user authentication is provided for always-on connections, periodic renewal SHOULD occur

3.1.2 Device Configuration Requirements

There are 2 possibilities for device configuration in the telecommuter scenario: either access to the target network is permitted for the native ISP-assigned address of the telecommuter's system, or the telecommuter's system is assigned a virtual address from within the target address space. In the first case, there are no device configuration requirements which are not already satisfied by the ISP. However, this case is the exception, rather than the rule.

The second case is far more common, due to the numerous benefits derived by providing the IRAC with a virtual presence on the target

network. For example, the virtual presence allows the client to receive subnet broadcasts, which permits it to use WINS on the target network. In addition, if the IRAC tunnels all traffic to the target network, then the target policy can be applied to Internet traffic to/from the IRAC.

In this case, the IRAC requires, at minimum, assignment of an IP address from the target network. Typically, the IRAC requires anywhere from several more to many more elements of configuration information, depending upon the corporate network's level of topological complexity. For a fairly complete list, see section 2.2.

To summarize, the following are the device configuration requirements for the IRAC:

- o support for a virtual IP (VIP) address MAY be provided
- o if VIP support is provided, support for all device-related parameters listed in section 2.2 above SHOULD be provided
- o support for address assignment based upon authenticated identity MAY be provided
- o if authenticated address assignment is not supported, an identity-based dynamic policy update mechanism such as is described in [ARCH] MUST be supported.

3.1.3 Policy Configuration Requirements

In terms of IRAC policy configuration, the most important issue pertains to whether the IRAC has direct Internet access enabled (for browsing, etc.) while a connection to the target network exists. This is important since the fact that the IRAC has access to sites on the Internet implies that those sites have some level of reciprocal access to the IRAC. It may be desirable to completely eliminate this type of access while a tunnel is active.

Alternatively, the risks may be mitigated somewhat by forcing all Internet-bound packets leaving the IRAC to first traverse the tunnel to the target network, where they may be subjected to target network policy. A second approach which carries a bit less overhead entails modifying the IRAC's policy configuration to reflect that of the target network during the time the IRAC is connected. In this case, traffic is not forced to loop through the target site prior to exiting or entering the IRAC. This requires some sort of policy download (or modification) capability as part of the SA establishment process. A third approach is to provide a configuration variable for the IRAC which permits specification of "tunnel-all", or "block all traffic not destined for the target network while the SA is up".

In terms of IRAS configuration, it may be necessary to dynamically update the security policy database (SPD) when the remote user connects. This is because transit selectors must be based upon network address parameters, but these cannot be known a priori in the remote access case. As is noted above, this may be avoided by provision of a mechanism which permits address assignment based upon authenticated identity.

To summarize, the following are the policy configuration requirements for the IRAS and IRAC:

IRAS

- o dynamic policy update mechanism based upon identity and assigned address MAY be supported.
- o if address assignment-based policy update mechanism is not supported, address assignment based upon authenticated identity SHOULD be supported.

IRAC

- o IRAC SHOULD provide ability to configure for "tunnel-all" and/or "block-all" for traffic not destined for the remote network to which IPsec remote access is being provided.
- o support for dynamic IRAS update of IRAC policy MAY be provided.

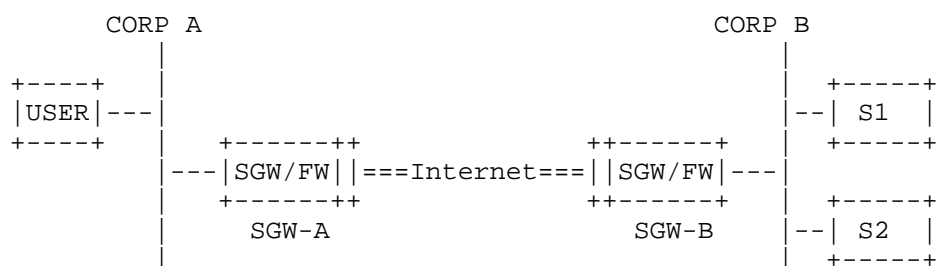
3.1.4 Auditing Requirements

For telecommuter sessions, session start/end times must be collected. Reliable derivation of session end time requires that the IRAC somehow periodically signify that the connection remains active. This is implied if the IRAS receives data from the IRAC over the connection, but in cases where no data is sent for some period of time, a signaling mechanism is required by which the IRAC indicates that the connection remains in use.

3.1.5 Intermediary Traversal Requirements

If the address assigned by the ISP to the IRAC system is globally routable, and no intermediate devices between the IRAC and the IRAS perform NAT operations on the data stream, then there are no additional requirements. If NAT operations are performed on the data stream, some mechanism must be provided in order to render these modifications transparent to the IPsec implementation.

Extranets are becoming increasingly common, especially as IPsec becomes more widely deployed. In this scenario, a user from one corporation uses a local corporate system to access resources on another corporation's network. Typically, these corporations are cooperating on some level, but not to the degree that unbridled access between the two networks would be acceptable. Hence, this scenario is characterized by limited access. The general topological appearance is similar to this:



This is purposely simplified in order to illustrate some basic characteristics without getting bogged down in details. At the edge of each network is a combination security gateway and firewall device. These are labeled "SGW-A" and "SGW-B". In this diagram, corporation B wishes to provide a user from corporation A with access to servers S1 and/or S2. This may be accomplished in one of several different ways:

- 1) an end-to-end SA is formed from USER to S1 or S2
- 2) a tunnel-mode SA is formed between SGW-A and SGW-B which only permits traffic between S1/S2 and USER.
- 3) a tunnel-mode SA is formed between USER and SGW-B which only permits traffic between S1/S2 and USER.

These various cases are individually discussed with respect to each requirements category below.

For the corporate extranet scenario, the authentication requirements vary slightly depending upon the manner in which the connection is accomplished. If only a particular user is permitted to access S1/S2, then user-level authentication is required. If connection types (1) or (3) are used, this may be accomplished in the same manner as it would be for a telecommuter. If connection type (2) is

used, one of two things must occur: either SGW-A must provide some local mechanism for authenticating USER and SGW-B must trust this mechanism, or SGW-B must have some mechanism for authenticating USER independently of SGW-A.

If access is permitted for anyone within corporation A, then machine authentication will suffice. However, this is highly unlikely. A slightly more likely situation might be one in which access is permitted to anyone within a particular organizational unit in corporation A. This case is very similar the single user access case discussed above, and essentially has the same requirements in terms of the mechanism required for SGW-A, although machine authentication might suffice if the organizational unit which is permitted access has a sufficient level of physical security. Again, this requires that corporation B trust corporation A in this regard.

To summarize, the following are the authentication requirements for the IRAS and IRAC:

IRAS

- o machine authentication MUST be provided.

IRAC

- o support for either user or machine authentication MUST be provided
- o support for a combination of user and machine authentication SHOULD be provided
- o if user authentication is used, periodic renewal SHOULD occur

3.2.2 Device Configuration Requirements

It is possible that corporation B would want to assign a virtual address to USER for the duration of the connection. The only way this could be accomplished would be if USER were a tunnel endpoint (e.g., in cases (1) and (3)). It is not clear what benefits, if any, this would offer.

To summarize, the following are the device configuration requirements for the IRAC:

- o support for a virtual address MAY be provided
- o if VIP support is provided, support for all device-related parameters listed in section 2.2 above SHOULD be supported

- o support for address assignment based upon authenticated identity SHOULD be supported
- o if authenticated address assignment is not supported, an identity-based dynamic policy update mechanism such as is described in [ARCH] MUST be supported.

3.2.3 Policy Configuration Requirements

Any of the cases discussed above would present some static policy configuration requirements. Case (1) would require that SGW-A and SGW-B permit IPsec traffic to pass between USER and S1/S2. Case (3) would have similar requirements, except that the IPsec traffic would be between USER and SGW-B. Case (2) would require that the appropriate transit traffic be secured between USER and S1/S2.

None of these cases require dynamic policy configuration.

3.2.4 Auditing Requirements

For cases (1) and (3), session start/end times must be collected. Reliable derivation of session end time requires that the IRAC somehow periodically signify that the connection remains active. This is implied if the IRAS receives data from the IRAC over the connection, but in cases where no data is sent for some period of time, a signaling mechanism is required by which the IRAC indicates that the connection remains in use.

For case (2), the type(s) of required auditing data would depend upon whether traffic from multiple users were aggregated within a single tunnel or not. If so, the notion of individual connection start/stop times would be lost. If such measures are desired, this requires that per-user tunnels be set up between SGW-A and SGW-B, and that some sort of timeout interval be used to cause tunnel teardown when traffic does not flow for some interval of time.

3.2.5 Intermediary Traversal Requirements

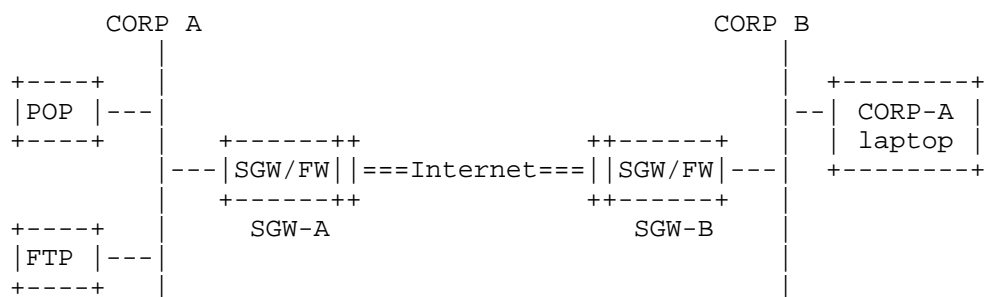
If the address assigned by the host network to the IRAC system is globally routable, and no intermediate devices between the IRAC and the IRAS perform NAT operations on the data stream, then there are no additional requirements in this regard. If NAT operations are performed on the data stream, some mechanism must be provided in order to render these modifications transparent to the IPsec implementation.

If a firewall situated at the edge of the host network cannot be configured to pass protocols in the IPsec suite, then some mechanism must be provided which converts the data stream to one which the

firewall may be configured to pass. If the firewall can be configured to pass IPsec protocols, then this must be accomplished prior to connection establishment.

3.3 Extranet Laptop to Home Corporate Net

The use of a laptop while visiting another corporation presents another increasingly common extranet scenario. In this case, a user works temporarily within another corporation, perhaps as part of a service agreement of some sort. The user brings along a CORP-A laptop which is assigned a CORP-B address either statically or dynamically, and the user wishes to securely access resources on CORP-A's network using this laptop. This scenario has the following appearance:



This is very similar to the telecommuter scenario, but it differs in several important ways. First, in this case there is often a SGW and/or firewall at the edge of CORP-B's site. Second, there may be a significantly increased risk that a long-lived connection could become accessible to someone other than the intended user.

3.3.1 Authentication Requirements

In most cases, the only acceptable connections from CORP-A's perspective are between the laptop and either SGW-A or the CORP-A servers the laptop wishes to access. Most of the considerations applied to the telecommuter also apply here, and user-level authentication is required to provide assurance that the user who initiated the connection is still the active user. As an added precaution, a combination of user-level and machine-level authentication may be warranted in some cases. Further, in either case this authentication should be renewed frequently.

To summarize, the following are the authentication requirements for the IRAS and IRAC:

IRAS

- o machine authentication MUST be provided.

IRAC

- o support for machine authentication SHOULD be provided
- o support for user authentication MUST be provided
- o support for a combination of user and machine authentication SHOULD be provided
- o periodic renewal of user authentication MUST occur

3.3.2 Device Configuration Requirements

The device configuration requirements in this scenario are the same as for the telecommuter, i.e., the laptop may be assigned a virtual presence on the corporate network, and if so, will require full infrastructure configuration.

To summarize, the following are the device configuration requirements for the IRAC:

- o support for a virtual address MAY be provided
- o if VIP support is provided, support for all device-related parameters listed in section 2.2 above SHOULD be supported
- o support for address assignment based upon authenticated identity SHOULD be supported
- o if authenticated address assignment is not supported, an identity-based dynamic policy update mechanism such as is described in [ARCH] MUST be supported.

3.3.3 Policy Configuration Requirements

The policy configuration requirements in this scenario differ from those of the telecommuter, in that the laptop cannot be assigned a policy which requires all traffic to be forwarded to CORP-A via the tunnel. This is due to the fact that the laptop has a CORP-B address, and as such, may have traffic destined to CORP-B. If this traffic were tunneled to CORP-A, there might be no return path to CORP-B except via the laptop. On the other hand, Internet-bound traffic could be subjected to this restriction if desired, and/or all traffic other than that between CORP-A and the laptop could be blocked for the duration of the connection.

IRAC

- o support for IRAS update of IRAC policy MAY be provided.
- o if IRAS update of IRAC policy is not supported, IRAC MAY support IRAS directives to "block-all" for non-tunneled traffic.
- o IRAC SHOULD provide ability to configure for "tunnel-all" and/or "block-all" for traffic not destined for the remote network to which IPsec remote access is being provided.

3.3.4 Auditing Requirements

The auditing requirements in this scenario are the same as for the telecommuter scenario. Session start/end times must be collected. Reliable derivation of session end time requires that the IRAC somehow periodically signify that the connection remains active. This is implied if the IRAS receives data from the IRAC over the connection, but in cases where no data is sent for some period of time, a signaling mechanism is required by which the IRAC indicates that the connection remains in use.

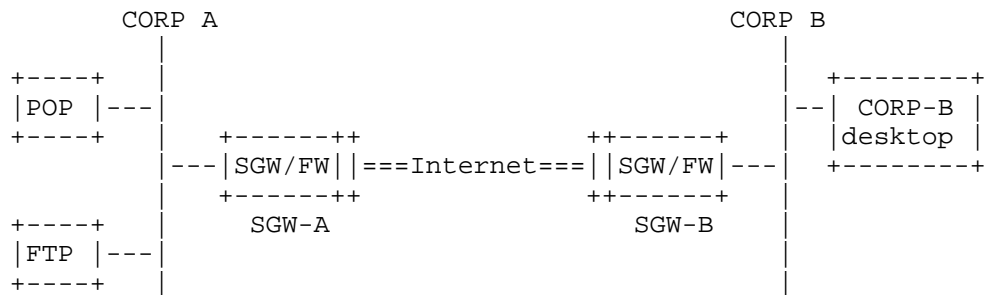
3.3.5 Intermediary Traversal Requirements

If the address assigned by the host network to the IRAC system is globally routable, and no intermediate devices between the IRAC and the IRAS perform NAT operations on the data stream, then there are no additional requirements in this regard. If NAT operations are performed on the data stream, some mechanism must be provided in order to render these modifications transparent to the IPsec implementation.

If a firewall situated at the edge of the host network cannot be configured to pass protocols in the IPsec suite, then some mechanism must be provided which converts the data stream to one which the firewall may be configured to pass. If the firewall can be configured to pass IPsec protocols, then this must be accomplished prior to connection establishment.

3.4 Extranet Desktop to Home Corporate Net

This is very similar to the extranet laptop scenario discussed above, except that a higher degree of trust for CORP-B is required by CORP-A. This scenario has the following appearance:



3.4.1 Authentication Requirements

The authentication requirements for the desktop extranet scenario are very similar to those of the extranet laptop scenario discussed above. The primary difference lies in the authentication type which may be used, i.e., in the laptop case, CORP-A can derive some assurance that the connection is coming from one of CORP-A's systems if a securely stored machine credential is stored on and used by on the laptop. In the desktop case this is not possible, since CORP-A does not own the IRAC system.

To summarize, the following are the authentication requirements for the IRAS and IRAC:

IRAS

- o machine authentication MUST be provided.

IRAC

- o support for machine authentication MAY be provided
- o support for user authentication MUST be provided
- o support for a combination of user and machine authentication MAY be provided
- o periodic renewal of user authentication MUST occur

3.4.2 Device Configuration Requirements

The device configuration requirements in this scenario are the same as for the laptop extranet scenario, i.e., the desktop system may be assigned a virtual presence on the corporate network, and if so, will require full infrastructure configuration. However, this seems less likely than in the laptop scenario, given CORP-A's lack of control over the software configuration of CORP-B's desktop system.

3.4.3 Policy Configuration Requirements

The policy configuration requirements are quite similar to those of the extranet laptop, except that in this scenario there is even less control over CORP-B's desktop than there would be over the laptop. This means it may not be possible to restrict traffic in any way at the desktop system.

3.4.4 Auditing Requirements

The auditing requirements in this scenario are the same as for the telecommuter scenario. Session start/end times must be collected. Reliable derivation of session end time requires that the IRAC somehow periodically signify that the connection remains active. This is implied if the IRAS receives data from the IRAC over the connection, but in cases where no data is sent for some period of time, a signaling mechanism is required by which the IRAC indicates that the connection remains in use.

3.4.5 Intermediary Traversal Requirements

If the address assigned by the host network to the IRAC system is globally routable, and no intermediate devices between the IRAC and the IRAS perform NAT operations on the data stream, then there are no additional requirements in this regard. If NAT operations are performed on the data stream, some mechanism must be provided in order to render these modifications transparent to the IPsec implementation.

If a firewall situated at the edge of the host network cannot be configured to pass protocols in the IPsec suite, then some mechanism must be provided which converts the data stream to one which the firewall may be configured to pass. If the firewall can be configured to pass IPsec protocols, then this must be accomplished prior to connection establishment.

3.5 Public System to Target Network

This scenario entails a traveling user connecting to the target network using a public system owned by someone else. A commonly cited example is an airport kiosk. This looks very similar to the extranet desktop scenario, except that in the extranet scenario, CORP-A might have a trust relationship with CORP-B, whereas in this scenario, CORP-A may not trust a publicly accessible system. Note that a trust relationship between CORP-A and the owner of the public system may exist, but in many cases will not.

3.5.1 Authentication Requirements

There are two variations to this scenario. In the first, no trust relationship exists between the target network and the borrowed system. In the second, some trust relationship does exist. In the case where no trust relationship exists, machine authentication is out of the question, as it is meaningless in this context. Further, since such a system could easily capture a passphrase, use of a static passphrase from such a system would seem to be ill-advised.

If a one-time passphrase were used, this would mitigate the risk of passphrase capture by the public system. On the other hand, if it is acknowledged that such capture is a real threat (i.e., the system itself is malicious), then it must also be recognized that any data transmitted and received via the resulting session would not be confidential or reliable with respect to this malicious system, and that the system could not be trusted to have actually disconnected when the user walks away. This suggests that accessing non-trivial information from such a system would be imprudent.

Another possible user authentication option would be a smartcard. However, many smartcards require a pin or passphrase to "unlock" them, which requires some level of trust in the kiosk to not record the pin. Hence, this approach suffers from drawbacks similar to those of the static passphrase in this regard. The primary difference would be that the pin/passphrase could not be used alone for access in the smartcard case.

In cases where a trust relationship with the owner of the public system exists, the trust level would modulate the risk levels discussed above. For example, if a sufficient level of trust for the system owner exists, use of a static passphrase might present no more risk than if this were permitted from a system owned by the accessed target. However, the primary benefit of such a trust relationship would be derived from the ability to authenticate the machine from

which the user is attempting access. For example, a security policy requiring that remote access only be permitted with combined user/machine authentication might be effected, with further control regarding which machines were allowed.

An additional issue to be dealt with in either case pertains to verification of the identity of the IRAS. If the IRAC were to be misdirected somehow, a man in the middle attack could be effected, with the obtained password being then used for malicious access to the true IRAS. Note that even a one-time password mechanism offers little protection in this case. In order to avert such an attack, the IRAC must possess some certifiable or secret knowledge of the IRAS prior to attempting to connect. Note that in the case where no trust relationship exists, this is not possible.

To summarize, the following are the authentication requirements for the IRAS and IRAC:

IRAS

- o machine authentication MUST be provided.

IRAC

- o in cases where no trust relationship exists between the accessed network and the system owner, sensitive data SHOULD NOT be transmitted in either direction.
- o in cases where a trust relationship exists between the accessed network and the system owner, machine authentication SHOULD be supported.
- o in cases where a trust relationship exists between the accessed network and the system owner, a static passphrase MAY be used in conjunction with machine-level authentication of the IRAC system.
- o frequent renewal of user authentication MUST occur

3.5.2 Device Configuration Requirements

None.

3.5.3 Policy Configuration Requirements

None.

3.5.4 Auditing Requirements

The auditing requirements in this scenario are the same as for the telecommuter scenario. Session start/end times must be collected. Reliable derivation of session end time requires that the IRAC somehow periodically signify that the connection remains active. This is implied if the IRAS receives data from the IRAC over the connection, but in cases where no data is sent for some period of time, a signaling mechanism is required by which the IRAC indicates that the connection remains in use.

3.5.5 Intermediary Traversal Requirements

If the address of the IRAC system is globally routable, and no intermediate devices between the IRAC and the IRAS perform NAPT operations on the data stream, then there are no additional requirements in this regard. If NAPT operations are performed on the data stream, some mechanism must be provided in order to render these modifications transparent to the IPsec implementation.

4. Scenario Commonalities

As we examine the various remote access scenarios, a general set of common requirements emerge. Following is a summary:

- o Support for user authentication is required in almost all scenarios
- o Machine authentication for the IRAS is required in all scenarios
- o A mechanism for providing device configuration for the IRAC is required in most scenarios. Such a mechanism must be extensible.
- o Machine authentication for IRAC is generally only useful when combined with user authentication. Combined user and machine authentication is useful in some scenarios.
- o Dynamic IRAC policy configuration is useful in several scenarios.
- o Most scenarios require auditing for session start/stop times.
- o An intermediary traversal mechanism may be required in any of the scenarios.

5. Security Considerations

The topic of this document is secure remote access. Security considerations are discussed throughout the document.

6. References

- [ARCH] Kent, S. and R. Atkinson, "Security Architecture for the Internet Protocol", RFC 2401, November 1998.
- [KEYWORDS] Bradner, S., "Key Words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RADIUS] Rigney, C., Rubens, A., Simpson, W. and S. Willens, "Remote Authentication Dial In User Service (RADIUS)", RFC 2865, June 2000.
- [IKE] Harkins, D. and D. Carrel, "The Internet Key Exchange (IKE)", RFC 2409, November 1998.

7. Acknowledgements

The editors would like to acknowledge the many helpful comments of Sara Bitan, Steve Kent, Mark Townsley, Bernard Aboba, Mike Horn, and other members of the ipsra working group who have made helpful comments on this work.

8. Editors' Addresses

Scott Kelly
Airespace
110 Nortech Pkwy
San Jose CA 95134 USA

Phone: +1 (408) 941-0500
EMail: scott@hyperthought.com

Sankar Ramamoorthi
Juniper Networks
1194 North Mathilda Ave
Sunnyvale CA 94089-1206 USA

Phone: +1 (408) 936-2630
EMail: sankarr@juniper.net

9. Full Copyright Statement

Copyright (C) The Internet Society (2003). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.

