

Visa Scheme for Inter-Organization Network Security

Deborah Estrin and Gene Tsudik
Computer Science Department
University of Southern California
Los Angeles, California 90089-0782

Abstract

In this paper we describe a *visa* scheme for implementing access control in Inter-Organization Network (ION) gateways. The purpose of the scheme is to allow an organization to modify and trust only those internal systems that require ION access; all other internal systems can not communicate with the outside. Control is distributed among the ION participants so that each may make its own design tradeoffs between performance and trust.

It is desirable to implement controls at the *network*, i.e., packet, level because of the relative performance, flexibility, and ubiquity of network-level gateways. However, a new mechanism was called for because the only information available to existing network-level gateways is the network-level address in the packet header and such network-level addresses do not carry the higher-level, logical information (e.g., organization affiliation) needed to make access control decisions. To overcome these problems, a visa ION gateway works in concert with an Access Control Server (ACS). The ACS carries out high-level evaluation of communication requests and the gateway enforces the ACS's decision using the visa scheme. In order for a node to send a packet through a visa gateway, the node must obtain a key (visa) from the ACS of the visa-controlled networks that it wishes to leave and enter. If the node passes an ACS's policy filter, the ACS gives its local gateway the source and destination nodes' network IDs and a visa with which to authenticate packets coming from or to the source node as they pass through the gateway. The same visa is given to the source node to stamp all outgoing packets for the duration of the session. To prevent or inhibit the acquisition of visas through interception of packets the stamp included in each packet is a function of the visa and the packet checksum.

1 Motivation

This paper describes an access control protocol, called a *visa* scheme for use in Inter-Organization Networks (IONs). IONs are becoming widespread in the academic community as well as in the private sector. While necessary and convenient, inter-organization connections present a number of problems, most crucial of which is access control [1,2].

Ordinarily, gateways forward packets between networks indiscriminantly, i.e., based on routing information only. If such a gateway is used for an inter-organization connection, all internal resources are potentially accessible to all external machines, and all internal machines

can potentially gain access to external resources. Some organizations address the need for control of such connections by implementing high-level gateways with access control functions; for example, an electronic mail relay that forwards mail to and from registered users only. While suitable for some IONs, high-level gateways suffer from performance overhead of the gateway's high-level processing, and reduced generality and flexibility, since special high-level gateway software must be constructed for each high-level protocol supported. The purpose of our visa scheme is to implement access control in ION gateways without incurring the costs inherent to high-level gateways.[3]

One simple way of implementing access control is to place a source-destination filter in the packet-level gateway, i.e., to maintain an access control list based on internet addresses. However, this approach works only if the access control list is static or if the source and destination IDs carry sufficient information to inform access decisions. If there is a well-defined set of resources that are to be accessible by a well-defined set of entities, then the access control list could be managed manually. Alternatively, if internet addresses are structured in such a way that the gateway can classify a node according to the range into which its internet address falls, the gateway could maintain an access control list by node classes (internet address regions), and thereby achieve greater flexibility.

In this paper we are interested in the more general case of a dynamic environment where network addresses by themselves do not provide sufficient information for the gateway to make a policy decision about whether or not to permit access; the DARPA Internet is one such environment. As described in [3], internet numbers are assigned to carry topological, not logical information, while policy decisions are generally based on the latter. Because internet numbers do not carry enough information to assist access control decision making, our first proposal is that before an ION packet-level gateway starts passing packets between internal and external machines, it should require both internal and external participants to carry out a high-level conversation with an Access

Control Server (ACS). The ACS would decide whether or not the connection is authorized based on the high-level information provided. After authorization the ACS could inform the gateway that the connection between that source and destination was approved and the gateway could then check all address fields of arriving packets and reject packets whose source-destination pair was not registered.

Unfortunately, there remains a nagging problem that led us to develop a more sophisticated mechanism, referred to here as a visa scheme and described in the following pages. Namely, if the gateway relies solely on a list of approved address pairs provided by the ACS, the gateway, as well as the ACS and authorized internal nodes, must trust all internal nodes to not masquerade as other nodes, i.e., not to fake their internet addresses. In a decentralized environment with many personal computers and workstations it is not hard to modify one's internet address. As a result, this simple scheme does not provide internal nodes and gateways with enough of a mechanism to protect themselves from malicious or fraudulent traffic. Without additional control it would be unwise for an organization to accept liability for outgoing ION traffic, or for a particular internal node to accept responsibility for its own outgoing ION traffic. In summary, the visa scheme is developed to address two limitations of relying on internet addresses alone for access control: 1) internet addresses are bound to topological information, and 2) machines on a local network can claim a false address rather easily.

The visa scheme described below implements controls in a packet-forwarding gateway by working in concert with an ACS. The ACS carries out the high-level evaluation of communication requests and the gateway enforces the ACS's decision using the visa scheme. The visa scheme allows an organization to trust only those internal and external nodes that it explicitly provides with unique visas. If an authorized connection is abused, or a visa is passed from an authorized user to an unauthorized user, the responsibility can be isolated to a specific node and session. Without such a mechanism an organization, and the authorized machines within that organization, have inadequate means of protecting their liability for ION traffic.

In the following section we present the visa mechanism and several design goals. Section 3 describes the visa system components. Section 4 illustrates the use of the visa mechanism and Section 5 concludes with implementation issues.

2 Overview of Design Goals

A visa scheme was first suggested by D. Reed (M.I.T.) and documented by Mracek [5] and Estrin [3]. It is referred to as a visa scheme because gateways are analogous to border crossing stations, access control servers to embassies, and keys to visas.

In this scheme, in order for a host to send a packet via an ION gateway, it must obtain keys(visas) from the ACSs of the visa networks it wishes to exit and enter. If the host passes an ACS's policy filter, the ACS gives its local gateway the source and destination hosts' network IDs and a visa with which to authenticate packets coming from or to the source host as they pass through the gateway. The same visa is given to the source host to stamp all outgoing packets for the duration of the session. To prevent or inhibit (depending on the strength of the stamping function) the acquisition of visas through interception of packets the stamp included in each packet is a function of the visa and the packet checksum. Abuse of a visa is therefore possible only if (1) the source or gateway machine releases the visa value, or does not protect it adequately, or (2) the attacker is able to invert the function used to stamp packets.

2.1 Liability

The visa mechanism is designed to allow an organization to connect to the outside world without modifying *all* internal systems to defend themselves from external access, and without having to trust all internal systems to not abuse the external connection in the name of the organization. In other words, our goal is for an organization to *modify* and *trust* only those internal systems that explicitly request or require ION access. All other internal systems (the majority) would be unreachable by external packets and would not be able to export packets.¹

The requirement for control of incoming traffic (i.e., external access to internal information and resources) is rather straight forward, namely, controlled access to proprietary resources. In addition to incoming flows, we are also concerned with outgoing traffic because generally when an organization, A, connects to an external organization, B, A must agree to assume responsibility for the actions of persons and machines within its orga-

¹Many workstations and personal computers may be designated to receive electronic mail from external sources. However for such applications, these hosts need not be directly connected to the ION gateway; rather a mail server would be one of the ION accessible machines and it would in turn forward mail to individual hosts after applying appropriate controls.

nization boundaries (e.g., to stand by purchase orders or other contracts written by its employees). In particular, A must vouch for the authenticity of internal entities that are able to export packets to B. If A is not confident as to the identity of an internal entity, then A should not allow it to use the gateway. Alternatively, A should not agree to ION connections for which the liability exceeds the level of confidence that A has in its internal access control mechanism.

The visa mechanism allows an organization to isolate trust and identify fault but it *does not* in and of itself provide any particular level of security. The security of the mechanism depends upon each organizations internal security; in particular, the ability of the source and gateway machines to prevent access to their visa values, the protection of visas during distribution, and the strength of the stamping function. The value of the visa mechanism is that it allows an organization to exert control over ION connections in a way that is consistent with its security guidelines. Moreover, an organization doesn't have to trust all its internal entities. It only trusts those that it explicitly permits to use the connection to the outside (see section 5.2).

2.2 Flexibility

One of the main benefits of this scheme is its flexibility. Each organization employing the visa scheme should be able to tailor it to reflect that organization's policy regarding incoming and outgoing traffic and to make its own trade-offs in performance and security. The scheme is designed to support this diversity in addition to minimizing requirements for trust and a priori agreements across network boundaries. Where such requirements remain, the placement of trust is explicit and well-isolated.

2.3 Transparency

In addition to flexibility, transparency of the underlying mechanism is an important design goal. This scheme must allow an organization to connect some subset of its internal resources to some subset of the outside world without endangering or tampering with any other internal facilities. The scheme must allow each ION participant to define the terms of liability that it and external parties must agree to. At the same time, interoperability with non-visa users must be maintained for those systems

that are globally accessible, i.e., impose no ION access control.

Another issue related to transparency is that the interconnection of two organizations may traverse other networks which may or may not be using the visa scheme. In such cases the presence of the VISA mechanism at the endpoint(s) must be transparent to the non-visa, transit gateways.

3 Visa Scheme Components

This section describes the main components of the visa scheme - hosts, visas, Access Control Servers (ACSs) and gateways (GWs). A host that wants to communicate across its organizational boundary engages in a high level authorization and authentication procedure with the ACSs on the visa networks traversed. The need for ACS communication is determined individually by the owners of each participant network. After the source-destination session has been approved by an ACS on each network, the ACSs allocate visas to their respective gateways and to the requesting host. The host uses the visa to stamp all ION packets. The gateways check all packets for appropriate stamping and pass packets until the visa expires or is terminated. If system processes are programmed to carry out the authorization procedure on behalf of the user, the entire process can be transparent to end users.

Our initial implementation of the visa scheme is based on the DOD Internet Protocol (IP). [7] IP supports connectionless datagram service between hosts. It was designed to flexibly operate over a range of network types, and to adapt to changes in topology and congestion. Both connection and connectionless transport protocols run on top of IP. Although we have designed this scheme to work within IP, the fundamental concepts could also be applied to other protocols such as X.25/X.75.

3.1 Visas

In the context of this scheme a visa is a unique value (e.g., a cryptographic key) assigned to a session between two hosts on distinct networks. Each packet that is part of an authorized session carries a special stamp value in the IP header option field that includes the VISA and packet checksum in its calculation. In our implementation, each visa packet carries two visas - one for the visa gateway that it is exiting, and one for the visa gateway that it is entering. This approach was selected because it provides flexibility in the future for different networks to employ different stamping functions (e.g., stronger functions than the simple IP checksum). The packet header format is

described further below. Initially, while we work out the protocol details, we use the IP checksum as our stamping function. Because this checksum algorithm is not secure, in future prototypes, stronger one-way functions will be employed. This option for upgrading and tailoring the mechanism is one of the features of the visa scheme.

Each host that makes use of the ION maintains an active visa list (VL). Each entry in the VL consists of a visa, the addresses of the machines involved in the session, and any restrictions that may apply (e.g. time limit). Gateways and hosts also maintain records of which ACS provided each visa. Likewise, for an ACS, the VL includes the address of the GW to which a visa was allocated. Also, an ACS associates with each entry in its VL an address of the ACS on the source or destination network.

In some cases it would be desirable to allocate visas to particular processes, not to entire hosts. However, packets do not carry process IDs, or even port numbers. Consequently, in our implementation, the gateway maintains a visa list that maps visas to host ID pairs (i.e., source and destination host ID) and relies on the source host's visa-IP implementation not to share visas among processes. Therefore, when more than one process on a host obtains visas to communicate with a common destination host, the gateway accepts packets stamped with either visa. The gateway is therefore trusting the source host's visa-IP implementation to only employ a visa for the particular process to which it was allocated. For further discussion of how finer granularity could be achieved, see section 5.4.

When a host explicitly terminates a session, visa-IP sends a special ENDING packet to its ACS. The ACS deletes the visa from its VL and forwards the packet to its gateway and to the next network's ACS. The ACS of the next network deletes the visa(s), informs its gateway(s) and sends the ENDING packet to the next network's ACS, and so on. When the ENDING packet finally arrives to the destination network's ACS, it sends the ENDING packet to the local gateway and to the local host. At that time all visas issued for that connection are invalidated by all parties involved. In addition, any GW or ACS may at any time decide to stop honoring a certain visa, e.g., timeout. In that case, it will send ENDING packets to its GWs, as well as to the local hosts and neighboring ACSs that are part of the session. The next packet bearing a stamp corresponding to the invalidated visa will be rejected. In the best possible case the ACS of the nearest network still honoring that visa will be able to recover the connection. In the worst case, the rejected packet will propagate all the way back to the source and the whole visa issuing procedure will have to be repeated. This visa expiration mechanism is also needed to terminate visas that are associated with connectionless

protocols; in this case the participating host(s) will not generate explicit ENDING packets themselves. Similarly, when topological changes cause rerouting of packets, new visas will be required to pass through any new visa GWs, or any old visa gateways that have crashed and resumed without previous state information.

3.2 Access Control Server

An ACS is assumed to be a host on the network (usually dedicated to ACS functions for security reasons), whose primary concern is access control. Each ACS knows of a number of local GWs to which it issues visas. Its presence, however, is not mandatory and levels of control can vary across organizations. If a participant network does not have an ACS, the scheme will still work; although the

The headers of visa related packet are illustrated in the following figures.

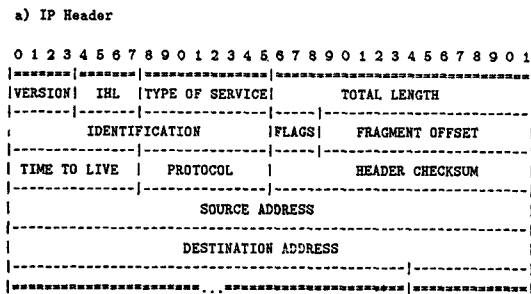
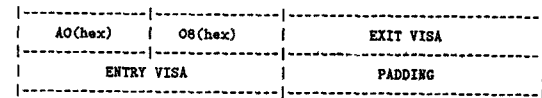


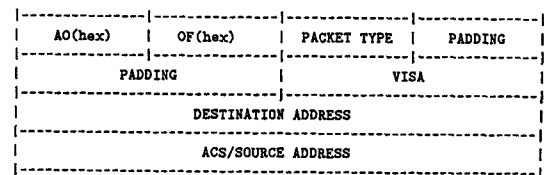
Figure 1: Standard IP Header. OPTIONS: One byte options type and one byte options length, followed by options data.



EXIT VISA: for exiting the current network

ENTRY VISA: for entering the next network

(a)



PACKET TYPE: VISA, LAST, REJECT, REQUEST.

ACS/SOURCE ADDRESS: reflects the address of the ACS in a REJECT packet, and the address of the source in a VISA, REQUEST or LAST packet.

VISA: only used in LAST and VISA packets.

(b)

Figure 2: Visa Scheme packet headers. (a) Visa option as it occurs in data packet. (b) Visa option as it occurs in a control packet.

network in question will be subject to risk associated with uncontrolled access. ACSs are trusted and assumed to be defensive against attempted abuse from external entities. This assumption is critical because visa-gateways allow any packet to flow to or from trusted internal ACSs.

The choice of the authorization and authentication procedures used by an ACS is the decision of each individual organization. The procedure may involve establishing a high-level conversation with the host, in which a password, biographical data, or other authenticating information is requested. Some ACSs will require end-user provided data, others will require information that the user's system can provide on its behalf. As described in [1], access control decisions may be most appropriately made according to group or class affiliation and associated category sets that determine access rights. The visa scheme itself does not dictate or constrain the particulars of the authorization schemes. One example of an ACS that could serve this function is the Kerberos Authentication Server developed at MIT [4]. Regardless of the approach used, the visa scheme assumes only that a YES/NO decision is passed to the visa software. In this paper we describe the visa interface of the ACS, not the ACS design itself. Finally, significant application-specific access control is left to the end-point hosts and applications; our scheme addresses only control of access to the hosts on a network.

The ACS' functions can be summarized as follows:

- On receipt of a request-to-connect from a host, authenticate and authorize that host.
- Issue new visas and send visa packets to participating GWs and hosts.
- Expire visas upon termination request by participant host or ACS, or upon timeout, and notify all parties involved - hosts, other ACSs and gateways.

Regardless of the authentication and authorization procedure used, when an ACS carries out a higher-level protocol via which it authorizes a host, it must have access to more than just the network addresses or ids. This does require for each participant host to understand the higher-level protocol used by a particular ACS whose gateway, that host wants to traverse. There are two options for dealing with this requirement: either the source host itself must have the ability to "speak" the higher-level protocols, or a local ACS must act on behalf of the source. In other words, one of the necessary, and unfortunately constraining, conditions for visa scheme implementation is that the ION participants' ACSs must satisfy one another's idiosyncratic higher level protocols or must have agreed upon a common mechanism a priori (e.g., a public key scheme).

ACSs play a critical role in this proposed scheme. Consequently, the availability of network service is a direct function of the availability of the ACS service. It therefore becomes worthwhile to designate backup ACSs within a single organization. In this case, each gateway would be initialized with the address of backup ACSs in case the primary ACS becomes unavailable. Similarly, the security of the scheme is dependent upon the security of the ACS. This suggests that the ACS reside on a dedicated trusted machine, and that the ACS employ a secure mechanism for communication with hosts; see subsection 5.2 for further discussion. In addition, as is described in section 5.2, ACSs should employ mechanisms to insure secure distribution of keys, i.e., visas.

3.3 Gateway

An ION GW is assumed to be a host on the network (usually dedicated for performance, and in this case security, reasons) concerned primarily with packet forwarding. Each GW knows of some number of trusted, local ACSs. By trusted we mean that the GW is willing to accept visa assignments from these ACSs and thereby trusts their decisions about authorizing sessions. Moreover, the GW allows any external party to communicate with (send packets to and receive packets from) any registered, internal ACS; similarly the GW allows all registered, local ACSs to communicate with any external party. In other words the GW trusts the ACS to protect itself from any external access and to not abuse the ION connection. This trust is reasonable because ACSs are special machines explicitly designed to be defensive and to enforce organization policy.

The gateway's functions include:

- Trap all packets, extract visa-stamp, search for source, destination, and visa in VL.
- Reject packets not possessing a valid stamp and return them to source along with the address of a local Access Control Server (ACS). If a packet does not possess any stamp option field, the gateway knows that the packet originated from a host that is not equipped to participate in the visa protocol. In such a case, the gateway simply drops the packet and leaves it to the source to time out and diagnose why the connection was not established.
- Forward packets bearing a valid visa through.
- Accept special VISA packets from the trusted ACS and add new visa entries to the VL.
- Accept special ENDING packets from trusted ACS and delete visa entries from the VL.

- Upon visa expiration, notify the corresponding ACS.

3.4 Network Environment

The particular visa scheme described here is designed to operate on IP networks such as the DARPA Internet as well as privately operated internets.[7] The general approach is applicable to other protocols but implementation is protocol dependent. Visa software is being integrated into IP code. We chose to implement the visa mechanisms at the IP level to exploit the use of this protocol for efficient network interconnection. In the future, to evaluate the relative value of this approach, we will compare its performance to that of transport and higher level gateways.

4 Illustration

The following example illustrates how the visa scheme is applied in a sample Inter-Organization Network. This example illustrates a pairwise connection. The scheme conceptually works in a multinet case where intermediate networks also employ visa gateways. However, due to the overhead per visa gateway transited, we suggest the scheme is most practical for the end-to-end case (i.e., only gateways on the source and destination network enforce visa requirements). See section 4.1 for further discussion.

Figure 3 shows the interconnection of a university department and a research division of a manufacturing company. Suppose, that department A was contracted to do some research for company B. Furthermore, B is allowing a certain number of faculty to use some of its resources in order to assist with ongoing research. However, being understandably protective about its assets, B is very much concerned with security and requires restricted access to internal systems. At the same time,

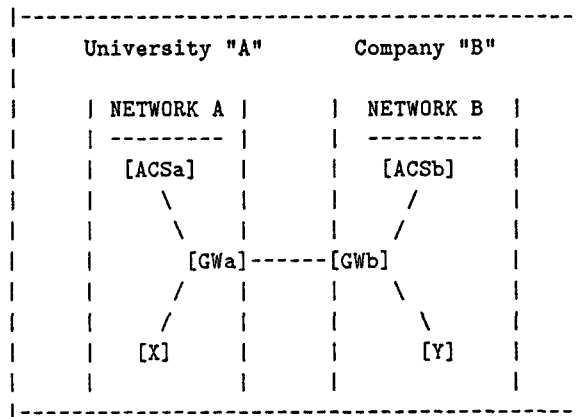


Figure 3: Example ION between a university and company.

physical isolation is not an acceptable solution because it limits the functionality of the connection by preventing communication between ION-accessible and strictly-internal machines. Instead, B "screens" all incoming and outgoing connections and imposes time limits on sessions. A, on the other hand, is only concerned with the appropriate usage of its gateway and external machines (i.e., A is more concerned with liability than with protection) and requires anyone requesting a remote connection be authorized to do so.

If a professor operating machine X located on the network A wants to query a database (host Y) located on the network B, the following procedure takes place:

1. X sends a packet addressed to Y.
2. The packet is trapped by GWA. The packet does not have a valid stamp. GWA sends a REJECT packet to X along with the address of the local ACS, ACSa.
3. X sends a REQUEST packet to ACSa. ACSa carries out an authorization and authentication procedure with X, the particulars of which will vary across organizations (and across different ACSs within an organization). The procedure may be executable by X's local ACS or operating system, or may require X's direct input.
4. (a) If the ACS decides that X is not authorized to communicate with Y then the packet is dropped and it is left to the higher level protocol to time out and diagnose the problem.
(b) If ACSa does not reject X it sends a REQUEST packet to Y (on behalf of X). GWA passes the packet since it originated from a local ACS. But the packet is trapped by GWB and as in step 2. a REJECT packet is sent to the source, this time ACSa, along with the destination's local ACS address, in this case ACSb.
5. On receipt of a REJECT, ACSa sends a REQUEST packet to ACSb. That packet passes through both GWA and GWB and gets to ACSb, because both gateways are passing packets to or from recognized ACSs. Upon receipt of this packet ACSb knows that someone wants a session with Y.
6. ACSb initiates its own authentication and authorization procedures with the requesting source, X, just as ACSa did. The conversation is carried out via ACSa, since GWA will only accept unstamped packets destined for an ACS.
7. After ACSb has authorized and authenticated X, it issues visaXYb and sends a special VISA packet to

GWb and ACSa. The gateways store the visa and associated information in their VLs.

8. When ACSa receives visaXYb it issues visaXYa and sends it to GWa. Then, it sends visaXYb and visaXYa to X. Now, X is armed with a visa for exporting packets from A to B.
9. X sends its first properly-stamped packet (with XYa and XYb) which passes through GWa and GWb and arrives at Y.

If ACSa and/or ACSb deploy symmetric policies regarding communication between X and Y (i.e., if X is authorized to send packets to Y then Y is authorized to send packets to X), then they can allocate two-way visas during the procedure described above. If ACSa or ACSb does not allocate two-way visas, then when Y attempts to reply to X's communication, its first packet triggers the same procedure as was just described for X. This time, the first gateway and ACS involved is B's, followed by A's. During this process if all participant ACSs authenticate and authorize Y, then they allocate visas to their respective GWs, and to Y, for the Y to X path.

In conformance with the spirit of IP, should any intermediate gateway or network go down, the session will resume automatically (albeit with additional overhead) just as when a gateway or ACS decides that a visa has expired or become suspect (see previous discussion).

4.1 Transit Case

For communication between X and Y when the networks of X and Y are not directly connected (see figure 4) the procedure may involve an additional set of steps.

If the intermediate network (e.g., belonging to an organization, "C") does not employ visa gateways, then the procedure would not change. The packets would simply be routed via an additional network before being processed by the participants' visa gateways; i.e., C's gateways would route the visa packets just as it does regular IP packets since the packets are not detectably different to a regular IP gateway (or host). If C employs visa gateways, it can elect to require visas for transit packets, or to allow transit packets without special visas. In the former case, C may use the visa mechanism to discriminate in its provision of transit service. In the latter case, C is agreeing to a policy whereby it will either allow or restrict ALL transit packets, independent of the source and destination, etc. In the latter case, C's gateways would recognize that the packets are transit packets and would pass them on without adding any steps to the visa set-up phase. However, if C chooses to implement controls on transit traffic also, several additional steps are added to

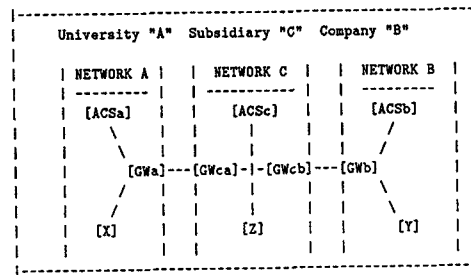


Figure 4: Example ION between a university and company when the networks are connected indirectly. Network C operates as a transit network.

the visa set up phase. Steps 1 through 6 would continue as before, although this time between ACSa and ACSc. However, instead of issuing a visa as ACSb did in step 7 in the previous example, ACSc would continue the set up chain by attempting to send an REQUEST packet on to Y in network B. At that point ACSb would get into the act as ACSc did before. Only after ACSb had assurance of authorization and authentication (via ACSa and ACSc), would it then issue a visa. The visa issuing process would propagate back through C and A just as it did from B to A in the previous example. In this way, conceptually the visa schemes is extendible to an internet in which any number of participating gateways and networks employ visa based access control. However, the greater the number of visa gateways on the path between two points, the greater the overhead for that particular conversation. Consequently we expect implementation to be practical when visa gateways treat transit packets differently than packets that are destined or originating from hosts on their network. This is accomplished by having each visa gateway on a network know about the other visa gateways on a network and allow transit packets to pass unchecked.

5 Implementation Issues

This section is devoted to the issues involved in implementation of the visa mechanism.

5.1 Performance

Our first and foremost goal in implementing the visa scheme is to analyze and evaluate trade-offs between performance, flexibility, and security. The extent to which we can actually meet our goals of transparency and flexibility and, yet, incur relatively low performance overhead will determine the usefulness of this security mechanism in a dynamic ION environment.

For the illustration given above, a minimum of 15 ex-

tra packets are generated before the first user packet gets through; not including the packets that comprise the authorization/authentication conversations between hosts and ACSs. Note that all of these control packets will be of minimal length. These ACS conversations may involve as few as 2 packets, but may involve many more depending upon the particular ACS design. Once the visa is allocated, successive user packets do not entail additional overhead (other than the added IP option field containing the stamp) unless a visa is expired or lost or the network state changes and a new gateway must be used.

There are several short cuts that organizations can take that tradeoff trust for performance. For example, an organization may choose to allocate two-way visas automatically so that Y would not have to go through an explicit visa-allocation process. Although this assumes greater trust in the remote organization, it would eliminate several steps and corresponding overhead. Another widely-applicable example is passing transit packets without visas, as described earlier.

In the future, the performance of this scheme must be compared to equivalent access control functions implemented in transport and higher level gateways.

5.2 Security

As mentioned previously, there are three points of potential vulnerability in the proposed scheme. The first is in the distribution of visas. If visas are distributed in the clear then packets emanating from a local ACS can be monitored by an attacker on the local network and visas can be illicitly acquired. Assuming the attacker can modify its network address, the stolen visa could be used to send and receive unauthorized ION packets. We assume that in the future most ACSs will have to carry out various kinds of key distribution functions and therefore will have an existing, local, mechanism by which to pass private information to hosts on the network, i.e., via encryption with the host's private key (e.g., as described in [6] and [4]).

The second point of vulnerability is in the storage of visa lists by hosts and gateways. Once again, the vulnerability depends upon the level of security mechanism available on particular hosts within an organization. If an organization does not trust a particular host or gateway to have adequate protection mechanisms, the ACS would be programmed not to allocate visas to that host or gateway. Similarly, the gateway must trust the visa-IP software belonging to a particular host to not use a visa belonging to an authorized process for stamping a visa belonging to an unauthorized process when both processes are communicating with a common destination.

The third point of vulnerability is the stamp itself. The stamping function used must not allow a wiretapper to obtain the visa through analysis of the stamp and other packet data. Therefore, implementations should employ a strong one-way function for computing the packet stamp as a function of the visa and packet data or checksum. The function we are currently using is quite vulnerable to such attacks. Our rationale for beginning with a simple checksum is to investigate the other performance issues associated with our general protocol design. Future versions will experiment with more sophisticated stamping functions. The range of possibilities is wide and is dealt with in some depth in existing literature so we do not elaborate here. In general, the more secure the scheme, the greater the computational overhead and the greater the need to employ special hardware. This might result in visa gateways being more expensive than traditional gateways. However, the relative number of ION gateways to internal gateways should be small and the expense justifiable.

Once a host is registered as being accessible via the visa gateway, it is then up to that host to protect itself from abuse and to not allow transit traffic to other internal, non-ION, hosts.

5.3 Implications for IP

There are two significant implications for the use of the IP and other datagram protocols. The first is that this scheme imposes a kind of single path behavior on IP. Packets can travel via multiple paths only if the gateways coordinate sharing of visas. Therefore we make use of the IP strict source routing option. The second issue is that fragmentation is a problem since the packet stamp is a function of the packet data (i.e., checksum). Consequently stamps would have to be recalculated at all fragmenting gateways.

5.4 Transport and Higher-Level Protocols

Although the visa scheme is being implemented at IP level, the choice of a higher-level protocol is not arbitrary. At this time, the scheme is being experimented with under TCP [8]. Since TCP is a connection-oriented protocol our software can detect when a session is terminated and visas should be invalidated (check for FIN flag in TCP header). In the presence of a connectionless transport protocol (e.g., UDP), detecting the end of an application level session becomes not possible; for such applications timeouts must be used to expire visas. Further research is needed to determine the role that the visa scheme can play in support of connectionless proto-

cols. The source of the problem is that we are modifying the connectionless IP protocol to be "aware" of connections in the sense of expiring visas when transport-level connections are closed.

It is sometimes necessary to issue visas to specific users or user processes, not to entire hosts. Although this issue may not arise in a PC environment where a machine is usually associated with a single user, in a multi-user environment the internet address (common to all users on a host) is not fine-grained enough to provide process-level control. In that case, higher-level IDs are needed to distinguish among user processes. Both UDP and TCP provide such information in their headers (port numbers). Thus visas could be issued to specific user processes if the visa-IP code is programmed with knowledge of specific transport protocols (e.g., where to find the port information in the UDP header of each IP encapsulated UDP packet).

5.5 Outstanding Design Issues

We conclude our discussion with a list of several outstanding design issues.

- In our experiments we are investigating the tradeoff in implementing functions in the ACS or gateway. We need to offload as much as possible from the gateway to maximize gateway performance while not exporting so much as to degrade performance through excessive communication requirements.
- We have designed this scheme to work within IP. However, the fundamental concepts could also be applied to other protocols such as X.25/X.75. The analysis and implementation of visas in other protocols is left for future investigation.
- As mentioned above, hosts must know when to send termination packets. This is a problem because IP is a connectionless protocol. We have modified it to detect TCP ending packets but it is unclear what the correct approach is to achieve this connection-oriented function without providing IP with knowledge about higher level protocols or without modifying higher level protocols as well. In general, further analysis is required to understand the applicability of this scheme for modern, connectionless protocols.
- More experience with the protocol is needed before we can evaluate the practicality of this scheme in the transit case, i.e., where networks enforce visa-based control over transit traffic.
- Finally, there are questions associated with the in-

teraction of our modifications and existing transport and higher level protocol mechanisms such as timeouts. Our performance must allow us to operate within the timeout periods of higher level protocols.

6 Status and Acknowledgments

We are currently experimenting with a prototype implementation. In future documents we will provide further details on implementation experience and ACS design.

We thank the following USC graduate students for participation in development and implementation: Kim Loh, Dev Mazumdar, Masuma Rahman, and David Woroboff.

In addition, we thank Matt Bishop, Vint Cerf, Jeff Mogul, Barry Leiner, Jerry Saltzer, and Lixia Zhang for comments on a previous draft.

7 Bibliography

[1] Estrin, D., "Non-Discretionary Controls for Inter-Organization Networks",
Proceedings of the 1985 IEEE Symposium on Security
and Privacy,
April 1985, IEEE.

[2] Estrin, D., "Access to Inter-Organization Computer Networks",
Ph.D. Thesis, M.I.T. Dept. of Electrical Engineering and Computer
Science, August 1985.

[3] Estrin, D., "Implications of Access Control Requirements for
Inter-Organization Network Protocols", Proceedings
of Sigcomm '86,
August 1986, ACM.

[4] Miller, S., Neuman B , "Kerberos: Athena Authentication, Authorization,
and Accounting Plan", Draft 3, M.I.T. Project Athena,
July 1985.

[5] Mracek, J., "Network Access Control in Multi-Net Internet Transport",
S.B. Thesis, M.I.T., Dept. of Electrical Engineering and Computer
Science, June 1983.

[6] Needham, R., Schroeder, M., "Using Encryption for Authentication
in Large Networks of Computers", Communications of
the ACM, December 1978.

[7] Postel, J., "Internet Protocol Internet Program Protocol Specification",
RFC 791, USC/Information Sciences Institute, September 1981.

[8] Postel, J., "Transmission Control Protocol DARPA Internet Program
Protocol Specification", RFC 793, USC/Information
Sciences Institute,
September 1981.