

Inter-Organization Networks: Implications of Access Control Requirements for Interconnection Protocols

Deborah Estrin¹
Computer Science Department
University of Southern California
Los Angeles, CA 90089-0782
estrin@usc-cse.usc.edu
(213)743-7842

Abstract

When two or more distinct organizations interconnect their internal computer networks they form an *Inter-Organization Network (ION)*. IONs support the exchange of cad/cam data between manufacturers and subcontractors, software distribution from vendors to users, customer input to suppliers' order-entry systems, and the shared use of expensive computational resources by research laboratories, as examples. This paper analyzes the technical implications of interconnecting networks across organization boundaries.

After analyzing the organization context in which IONs are used, we demonstrate that such interconnections are not satisfied by traditional network design criteria of connectivity and transparency. To the contrary, a primary high-level requirement is access control, and participating organizations must be able to limit connectivity and make network boundaries visible. We describe a scheme based on non-discretionary control which allows interconnecting organizations to combine gateway, network, and system-level mechanisms to enforce cross-boundary control over invocation and information flow, while minimizing interference with internal operations.

Access control requirements such as these impose new requirements on the underlying interconnection protocols. We demonstrate such alternative interconnection protocols that support loose coupling across administrative boundaries and that accommodate the necessary control mechanisms. Message-based gateways that support non-real-time invocation of services (e.g., file and print servers, financial transactions, VLSI design tools, etc.) are a promising basis for such loose couplings.

¹This work was conducted primarily at the M.I.T. Laboratory for Computer Science as part of the author's Ph.D. thesis.

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or specific permission.

1. Introduction

Inter-Organization Networks (IONs) support person-to-person communication via electronic mail; exchange of cad/cam data, software modules, or documents via file transfer; input to an order-entry or accounting system via a database query and update protocol; and use of shared computational resources via an asynchronous message protocol or remote login. In most such inter-organization arrangements, the set of resources that an organization wants to make accessible to outsiders is significantly smaller than the set of resources that it wants to remain strictly-internal (i.e., accessible to employees of the organization only). In addition, because the potential user is a person (or machine) outside the boundaries of the organization, the damage associated with undesired use can be high. Because of these characteristics, IONs have unique access-control requirements. These control requirements in turn place new requirements on the underlying network protocols.

This paper analyzes these control and network requirements and how they can be met. Section 2 summarizes the control requirements, and Sections 3 through 5 discuss the implications for network protocols and trade-offs in gateway design.

2. Controls for Inter-Organization Networks

Organizations frequently make their internal computing facilities accessible to off-site employees via public switched or packet networks. These external connections sometimes call for added security measures to prevent *outsiders* (i.e., persons who are not members or employees of the organization) from gaining access to the internal facilities. However, in an ION the goal is not simply to prohibit access by outsiders; some outside access is explicitly desired. The goal is to support access to certain machines, services, and processes, while preventing access to all other internal facilities. Unfortunately, the two obvious approaches—(1) physical isolation of externally-accessible systems from internally-accessible systems, and (2) increased access controls/security on all internal systems—are not

generally acceptable solutions. Namely, because the function of the internal network predates and dominates that of the ION, interconnection and associated controls must not interfere with internal operations. Therefore, it is not acceptable that ION facilities be physically isolated from all strictly-internal resources for this would interfere with internal access to information and resources that reside on ION systems. Similarly, requiring that all internal facilities adopt additional security measures to cope with an external connection may interfere with internal communication and resource sharing and furthermore assumes global knowledge of all interconnections. In general, we want to implement *logical networks* that can be isolated from one another yet share physical resources (see figure 2-1).²

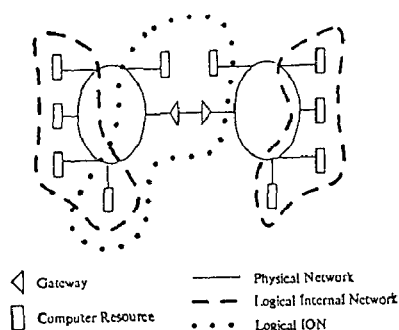


Figure 2-1: Overlapping Logical Networks: The ION shares physical resources with the two organizations' internal networks. However, at the logical level, the ION is isolated from the strictly-internal facilities.

Similarly, when two organizations interconnect, it may be inappropriate to impose a connection between the other organizations to which each was interconnected previously. In other words, the new ION may overlap physically with the existing IONs, but it must not form a transit path between those organizations that desire to remain isolated from one another (such as B and C in figure 2-2).

In [7, 8] we analyze in detail these control requirements and the design constraints that call for subtly, but significantly, different approaches to security-mechanism design and implementation. To summarize, we conclude that only the administrators of the external link (i.e., the ION gateway) and the internal resources that are made explicitly accessible should be required to take security action in response to an external connection. Owners of all other internal resources should be assured that their facilities are not accessible to outsiders. In other words, the management of a *strictly-internal*

²The term *logical network* refers to a collection of computational resources and applications that communicate with one another. Logical networks operate on top of physical networks which are composed of communication links and switches and processors.

resource should not have to rely on its own discretionary action for restriction of external access to its facilities. This requirement suggested the use of *non-discretionary* access controls in each organization's ION gateway to isolate strictly-internal resources and networks from the ION without relying on the discretion or explicit action of strictly-internal resource owners.

There are several differences between the non-discretionary access controls called for here, and those traditionally employed in military computer systems; the later being the environment for which non-discretionary controls were developed initially: [11, 12]

- Computer-based *services*, as well as *information*, may be accessible via an ION. Consequently, security mechanisms must control *invocation*, as well as *information flow*. Traditionally, non-discretionary controls have been applied only to the control of information flow.
- Each participating organization is responsible for controlling access to its own internal network. Therefore, the ION gateway's primary control function is to protect the internal services and facilities that may be *invoked*. In contrast, most existing systems that apply non-discretionary controls to *invocation* have been designed to protect the *invoker*. [2]
- An organization may want to allow a single internal machine or process to be accessible in multiple, internal and external logical networks. As a result, the control mechanisms in the ION gateway may not enforce strict confinement. Existing systems that support non-discretionary controls do enforce strict confinement and can not be used to implement overlapping logical networks.

These differences are explained in detail in [7, 8].

Before going on to discuss the implications of adopting this approach for underlying network interconnection protocols, we illustrate the use of non-discretionary controls in an ION with the following example of a university and two industrial research laboratories.

2.1. Example

The example illustrated in figure 2-3 is a hypothetical, not actual, case. It is representative of existing activities; however the details have been changed somewhat to illustrate several points in a single example.

The subject of this example is a university computer science department, such as MIT's, that is connected to the Arpanet. In the past, the user population that could access the department's Arpanet-connected machines was small and the department required no special measures in order to adequately comply with the Arpanet policy that the Arpanet be used only by computer science researchers. However, as MIT extends its computer networks out from the computer science and engineering department to the rest of

the campus, the user population that can access the Arpanet-connected machines is no longer small nor homogeneous. In addition, the university has established external network connections with local industry. In this case, the potential user population of the computer science department's facilities includes not only members of other departments, but members of other organizations altogether. In this new environment the computer science department may have to introduce control mechanisms to restrict access to the Arpanet gateway or Arpanet-connected hosts in order to adequately comply with Arpanet policy.³ Such control mechanisms have to discriminate between various segments of the user population; in this case these segments are logical groupings of users or hosts according to organization or department affiliation.

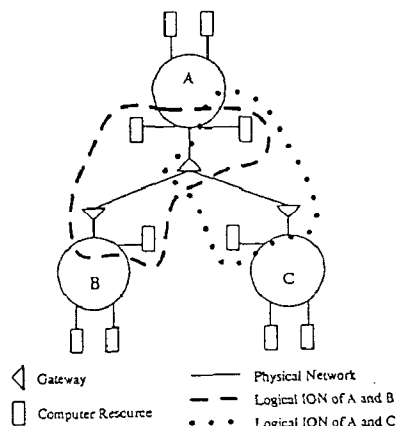


Figure 2-2: Overlapping IONs: The ION between A and B shares physical resources with the ION between B and C. However, at a logical level the two IONs are not connected to one another, i.e., B cannot communicate with C via A.

To illustrate how our general approach to access control in IONs addresses this environment we will examine MIT's connections to two of its industrial sponsors, ABC and XYZ, who happen to be competitors of one another (see figure 2-3). Each connection is intended to support exchange of software modules, access to some unique computational resources, and electronic mail. ABC has access to a host on which an MIT research group is developing educational software. XYZ has access to a different host on which another MIT research group is developing network software. XYZ also has access to a design simulation program developed by yet another group of researchers at MIT. In addition, both ABC and XYZ have access to electronic mail communications. For the sake of this example, we assume that both ABC and XYZ invoke the various services (i.e., file servers for software distribution, simulator, and mail distribution) by sending appropriately-structured messages through the gateway, and

³A related policy requirement with similar technical implications is that the Arpanet not be used as a transit path between two points, neither of which is itself a legitimate Arpanet node.

the servers return the requested data via the same gateway to the requesting organization. In addition to these ION resources, MIT has other strictly-internal computer-based facilities: administration, student accounts, other research projects, gateways to other networks, etc.

In this example there are five logical networks that need to be isolated from one another; where logical network refers to a set of computer resources that are intended to communicate and interwork. The two logical IONs are shown in figure 2-3, one between MIT and ABC, and the second between MIT and XYZ. In addition, each of the three organizations has a logical internal network which each organization should be able to isolate from the IONs. Note that there is no logical ION between ABC and XYZ because none of their facilities are intended to communicate or interwork. According to our design, in order to isolate ION from strictly-internal facilities, and the XYZ ION facilities from the ABC ION facilities, MIT could implement the controls listed below:⁴

1. Implement a single ION gateway and prohibit direct connection of all internal machines to outside organizations. Equip the gateway with an authentication mechanism to certify the source of each message.
2. Assign appropriate category sets to each of the ION facilities, and no category sets to strictly-internal ones. MIT assigns the category set {Educational-research} to the host used for development of educational software, the set {Network-research} to the host used for development of network software, {Architecture} to the design simulator, {*} to its electronic mail system to indicate all, and {Strictly Internal} to all other internal systems. See figure 2-3. If an internal facility is not registered at all the gateway assumes that it is not accessible via this entry exit point. The category information is assigned to internal resources but the information is maintained in MIT's gateway, see figure 2-4.
3. The ION gateway checks the category set of the source, $\{Ci\}_s$, and of the destination, $\{Ci\}_d$, of each message and forwards the message to the intended destination *If and only If* $\{Ci\}_s \text{ Intersect } \{Ci\}_d$ does not equal nullset, {} (referred to as the *Intersect rule*).
4. Equip the internal ION facilities (software distribution servers, electronic mail server, and design simulator) with discretionary or non-discretionary controls to enforce application-specific controls (e.g., restrictions based upon the dollar amount of a purchase order or the filename of a cad/cam file request), isolate non-ION files and processes, and prevent transit between the ABC ION and the XYZ IONs.

Similarly, ABC and XYZ each label their own research hosts and inventory systems with the category set {MIT} only, and implement gateways with message authentication and the *Intersect* rule. Note that each organization assigns category labels to incoming messages for interpretation by its own internal facilities. Therefore,

although labeling must be consistent within each organization, it need not be consistent throughout the ION as a whole.

2.2. Conclusions

The intention of the approach described is to minimize interference with the organizations' internal facilities and operations. However, the approach *does impose new requirements* on the network interconnection mechanisms—namely the ION gateway and the communication protocols used to communication with and through it. In particular, this approach has implications for the *level of interconnection*. The remainder of this paper investigates the implications of inter-organization connections, and the proposed control mechanisms, for network interconnection protocols. Further research is needed also to understand the range of applications for which the proposed modifications might be suited, the implications for non-discretionary security models, and appropriate authentication schemes.

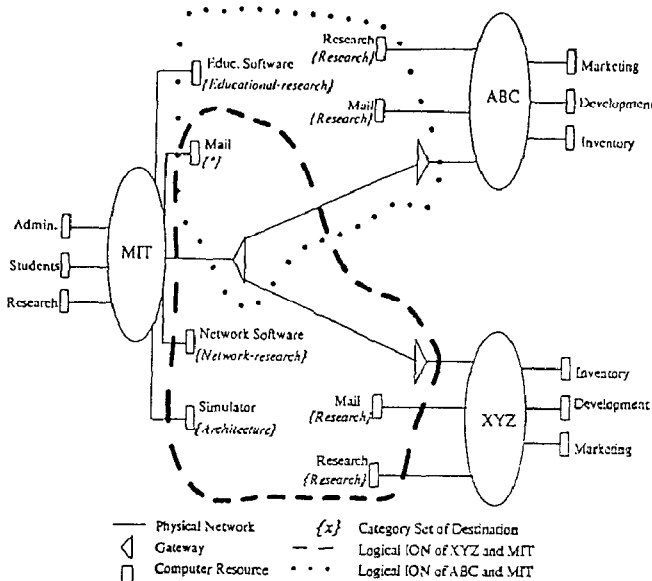


Figure 2-3: Example of an Inter-Organization Network: One ION exists between MIT and XYZ and another ION exists between MIT and ABC. Both IONs overlap physically yet are isolated logically from the internal networks of the three organizations.

Entity	Category Set
ABC	{Educational-computing}
XYZ	{Network-research, Architecture}
<hr/>	
Educ. Software	{Educational-computing}
Network Software	{Network-Research}
Mail	{*}
Simulator	{Architecture}

Figure 2-4: MIT's gateway table containing category set information.

3. Implications for Network Interconnection

In the previous section we described an approach to the problem of access control in IONs based on non-discretionary controls in all entry and exit points to each organization's internal network, i.e., all ION gateways. The primary difference between such ION gateways and traditional, uncontrolled gateways is that the non-discretionary control mechanisms in the ION gateway must have access to additional information about the characteristics of gateway traffic, e.g., organization affiliation of source and destination, type of service, amount of resource requested, etc. According to this information the gateway determines which categories of internal information or resources the external entity may access. In other words, in addition to the traditional bindings between user or service and node, node and network attachment point, and network points and path, [17] the ION gateway needs a binding between user or service and organization affiliation.

If the logical information required for the access control decisions is available, then the non-discretionary controls can be implemented by assigning category sets to incoming and outgoing traffic, according to logical characteristics of the traffic, and enforcing invocation and information flow controls accordingly. [7] In the remainder of this paper we describe what is required of the underlying network protocol in order to make this logical information available to the gateway. We evaluate the alternatives and tradeoffs associated with designing ION gateways to implement these controls. Although most of our examples focus on the issue of controlling ION flows to meet security policy concerns, similar mechanisms and issues can be employed to support concerns related to resource management. An example is given in the following section, after which we proceed with the discussion of network requirements. For example, broadcast protocols used for updating databases or locating resources⁵ often should be confined to the organization's internal network only and not forwarded across ION gateways.

3.1. Level of Interconnection

As with any gateway, an ION gateway can be designed to operate at one of several levels. For the sake of simplicity, we classify gateways as either high or low level. A high-level gateway is an end-point in a message- or connection-based communication session, such as file transfer, remote login, or electronic mail. A low-level gateway forwards packets between machines that are the endpoints of higher-level message or connection-based communication sessions, but the gateway itself is not an endpoint.⁶

⁵For example, the Address Resolution Protocol. [14]

A major difficulty of traditional interconnection methods when applied to inter-organization connections is that the traditional connections operate typically at lower protocol levels. Most low-level gateways do not have access to the information needed to make ION policy decisions. This is not necessarily inherent to this level of connection, but is a result of the competing requirements that constrain the design of low-level protocols.

As illustrated in section 2, the organization affiliation of source and destination is fundamental to many, if not most, conceivable usage control requirements for IONs. Given this information, the ION gateway should be able to assign categories and determine the rights of the source and destination.⁷ Consequently, an ION gateway must be able to identify the organization affiliation of the traffic destination and source. In low-level protocols, information about the source and destination is carried in the packet header, in the form of network numbers. The following paragraphs describe some of the problems of relying on these numbers, and therefore traditional low-level protocols—for identification of organization affiliation.

Networks interconnected at the packet level (e.g., Internet Protocol (IP) level in the DARPA TCP/IP family of protocols) must coordinate the assignment of network numbers in order for packet addresses to be meaningful throughout the internet. In addition, network numbers provide information about efficient routing of a packet to its destination, e.g., which subnet on which network a particular host sits. This routing information pertains to the physical location of the destination. When networks cross organization as well as geographic boundaries, *logical* information is desired in addition to *topological* information. In other words, policy control mechanisms need to know the organization domain to which a message is being sent, and from whence it came, in addition to the physical locations. One possibility is to interpret the physical address as a logical address. For an example of why this is not generally practical, we consider the case of the DARPA Internet.

Currently, network numbers in the Internet are allocated to sites by a centralized number czar. Each site may then allocate numbers to hosts and even subnets that lie within its topological network. Most of these hosts and subnets are within the confines of a single organization, but some are not. For example, MIT has direct

network connections to several local companies; see figure 3-1. The network numbers of destinations in these companies look like the network numbers of other MIT subnets because they contain topological information for routing purposes. In order to discriminate between subnets and hosts that are part of MIT's logical network (i.e., actually belong to MIT) and those that lie outside of the logical network (i.e., facilities in the local companies which are accessible but do not belong to MIT), the gateway must be able to bind the source and destination network numbers in the packet header to the organization affiliations.

These issues were not among the many considered during design of the Arpanet/Internet protocols. At that time, the primary concern was to achieve connectivity and transparency and make

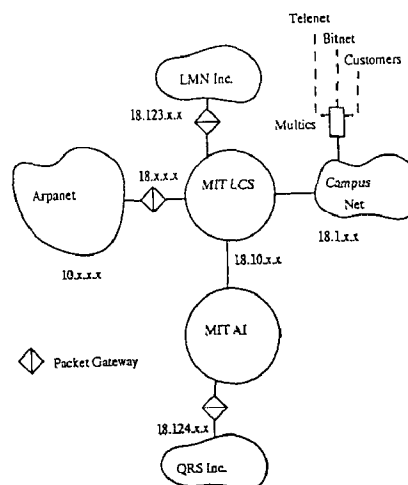


Figure 3-1: A simplified depiction of MIT internal networks and several external networks to which it is connected. The Internet network numbers are listed near the gateway to each network.

network boundaries disappear.⁸ Therefore, it makes sense that providing information needed to enforce organization boundaries was not a design requirement. Even if it had been a consideration, the number of competing requirements and constraints on the low-level protocol would probably have led the designers to leave such application-specific information to higher levels. In particular, because routing table size is limited, there is pressure to be able to make routing decisions on the basis of a packet's destination subnet number.

One might nevertheless try to use the network and subnetwork

⁶Low-level gateways may operate on individual packets (datagrams) or virtual circuits, depending upon the protocol design of the interconnected networks. In terms of the International Standards Organization, Open Systems Interconnect (ISO-OSI) reference model [10], high-level gateways operate at *session*, *presentation*, or *application* layers, whereas low-level gateways operate at *transport* or *network* layers.

⁷Many policies may require more information than just source and destination affiliation. But for simplicity I focus on this information to illustrate the argument.

⁸In many ways the Arpanet is not a typical ION. In the past, the Arpanet was intended to encompass the internal networks of the participating organizations. Only recently, as the participating organizations have extended their internal networks to other internal communities, is the Arpanet community manifesting many of the issues attributed here to IONs.

numbers as a *hint* to organization affiliation. However, because of the decentralized manner in which networks and subnetworks may establish their own interconnections, over time these topological numbers may not map into meaningful logical groupings. For example, MIT might implement a filter in the Arpanet gateway to reject outgoing messages with source addresses other than MIT's Arpanet network number, 18. However, if an MIT department or laboratory connects some local company to the department or laboratory local network, according to current practices, that company is assigned a subnet number within net 18. Therefore, the filter would not catch transit traffic sent from that company to non-MIT Arpanet sites. In addition, at different times, several MIT hosts have been located on networks other than network 18.

Of course it is possible to identify individually the various subnet numbers that are assigned to non-MIT entities and add such information to the gateway filter. However, this is not a *general* solution because such interconnections are established in an incremental and decentralized manner, and therefore there is no good way of tracking these exception cases without centralizing the interconnection and number allocation process in some way. One approach might be to establish guidelines that set aside blocks of numbers to be used for non-MIT sites. Unfortunately, because of the nature of the namespace, it is hard to know a priori how many such numbers to set aside, and exactly what groupings one will want to be able to distinguish between, i.e., MIT/non-MIT is only one relevant distinction. If such guidelines do not exist, and the connection is not centrally managed, it is not feasible for the gateway to maintain a list of allowable host and/or subnet addresses with which to implement packet-level controls.

An example of a packet-level ION gateway that implements usage controls is the University College London (UCL) network connection to the Arpanet. [3] The UCL network employs two gateways to the Arpanet. One connection forwards packets via a private satellite network to the Arpanet. The second connection forwards packets via an X.25 connection over public packet-switched networks. The two separate gateways are needed because of the different protocols used, and the division satisfies policy requirements. Due to PTT regulations, only Ministry of Defense traffic can be sent via the private satellite path, while civilians (such as many university researchers) must send traffic via the public-network path. Because only routing information is available at the IP level, the restriction is enforced by making UCLnet appear as two separate networks, UCLnet and PSSnet. This is achieved by splitting the namespace in two and assigning addresses to MOD and civilian hosts accordingly. Because there is a small and fixed number of user groups (i.e., two) the mechanism works. In addition, higher-level

controls enforce restrictions on invocation of higher-level network applications, i.e., mail, file transfer, and remote login (see the description in section 3.1). A similar mechanism could be employed by MIT to restrict access to the Arpanet. For example, most student access to computational facilities and the MIT network will occur via the MIT subnets that belong to project Athena.⁹ The Arpanet gateway could simply reject all packets originating from those subnets, but in so doing it would preclude transit by legitimate research hosts that are physically located on Athena subnets.

In conclusion, traditional gateways, and the packet-level protocols that they speak, do not carry the information in each packet header that an ION gateway needs to make policy decisions. Provided with access to ad hoc mappings of network number to policy related information, specific ION gateway requirements can sometimes be met. However, the following section describes two approaches to network interconnection that are better suited to interconnection across organization boundaries.

4. Alternative Network Interconnection Schemes

The first part of this section describes a scheme for augmenting a packet level protocol in order to accommodate policy controls. The second part describes alternative approaches that involve interconnecting at higher protocol levels.

4.1. Visa Scheme

Given that logical information generally is not deducible from packet headers alone, one alternative is to adopt an approach first suggested by D. Reed and documented by J. Mracek. [15] This scheme, depicted in figure 4-1, requires that the source carry out a higher-level dialog with a policy server in the destination network in order to authorize a particular conversation (e.g., mail, file transfer, etc.).

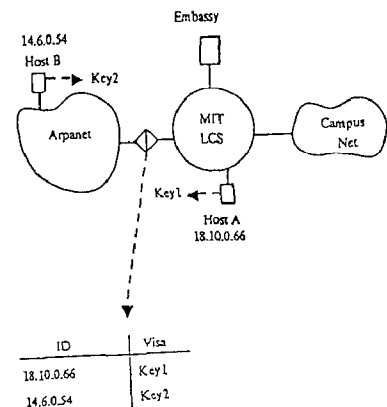


Figure 4-1: Example of a packet-level ION gateway using a visa scheme.

The policy server passes the authorization information to the packet-information and a way of mapping traffic information to that category information. In effect, by granting a visa to a source or source-destination pair, and informing the gateway of the granting, the ACS wraps a set of individual packets into a logical unit which is then subjectable to policy control in the packet-level gateway.

For example, a visa scheme could be used to control incoming traffic to a dial-up, packet-forwarding gateway to the MIT network. This gateway is connected on one side to the public switched telephone network, and on the other to an MIT local network. Although a single physical gateway is used, MIT would like to apply different access policies to the different groups that use it. Some MIT resources are intended for access by members of the MIT community only (e.g., gateways to other networks, a New York Times clipping service, high-speed printers, etc.). Other resources are intended for access by some non-MIT users as well. In order to implement non-discretionary controls as proposed in [7], the gateway would operate as follows. When a user calls the gateway, the gateway associates the call with a particular port and accepts packets from the user only if they are addressed to an ACS. The external user carries out a high-level dialog with the ACS and authenticates itself. After authenticating the user and identifying the internal facilities that the user is authorized to access, the ACS sends the gateway a key and a list of destination addresses to which the particular user should be allowed to send packets. In addition, the ACS sends the same key to the user. The gateway associates both the list of destinations and the key with the port assigned to the user. The key is used as a *connection-authenticator* for the duration of the connection. In order for the gateway to accept a packet through the port, the checksum of the packet must have been calculated including the connection-authenticator that is currently associated with the port. When the user first dials up the connection-authenticator is zero and the user can send packets to only a single destination, the ACS.

If access policies are relatively stable, an even simpler visa scheme can be used which would reduce the need for frequent higher-level dialogs with the ACS, and would therefore reduce the overall performance requirements of the ACS itself. For example, assume that an ION gateway needs only distinguish between those hosts that belong to a particular research community and those hosts that do not. In this case it would be adequate to regularly distribute a key to all eligible research hosts and the ION gateway. The gateway could use the key to identify packets belonging to authorized research

hosts, as in the above description. The research hosts would no longer require a dialog with the ACS for each external communication. Consequently, the ACS could even operate offline or via electronic mail.

A packet-level gateway together with an ACS can effect higher-level controls. However, if the ION is intended to support only a small number of higher-level applications, if the policies can not be expressed in terms of information available in packet headers, or if performance benefits of packet-level interconnection are not significant, it makes sense to consider interconnection at a protocol level at which the policy-related information is available directly. Therefore, the following section describes higher-level interconnection. The schemes are compared in more detail in Section 5.

4.2. Higher-Level Interconnection

The introduction presented motivations for treating entities on the other side of an ION gateway differently from those within an organization's internal network. First, policy concerns may require that non-local users be restricted from using some internal resources and other gateways. In addition, efficient network use suggests that information needed for local, tightly-coupled applications, not be broadcast through an ION gateway across which applications are more loosely coupled. In the discussion of low-level gateways I explained that the logical information needed to implement intelligent filtering in an ION gateway is not available at the packet level. The visa scheme described above can effect higher-level control for many simple usage control policies. However, when policy decisions are dependent on higher-level information that cannot easily be bound to packet-level information and represented in the form of a key, higher-level connections may be more practical. In addition, the visa scheme requires modification of lower-level communication protocols because of the need to alter calculation of the packet checksum. In general, the cost and inconvenience of modifying lower-level protocols is greater than that of modifying higher-level protocols.

Higher-level gateways terminate the higher-level communication protocols and thereby gain access to more information about the application of the connection. These protocols deal with aggregated units of traffic that contain more semantic information in the headers and control fields (e.g., electronic mail messages, remote login or file transfer connections, etc.). Depending upon the level of connection and application, this information may include the logical affiliation of source and destination, the actual service being performed, and the amount of communication resources requested, for example. Although a key security issue arises with

⁹Athena is a university-wide experimental project in the use of computers in education.

regard to the authentication of this information, the point is that the information is available for evaluation, and authentication mechanisms can be employed as needed. Note that even with a higher-level ION gateway, some controls are best implemented in the endpoint applications themselves; in particular, controls that discriminate according to the content of a message, e.g., the size of a purchase order, or the name of a file requested. In addition, these applications-level controls may be required to isolate the ION processes and applications from the non-ION ones. In the remainder of this section I will continue to focus on higher-level, communication protocol controls, under the assumption that some application-level controls are implemented in the endpoints.

Harvard University's connection to the Arpanet via a packet-level gateway illustrates the need for higher-level gateways. Harvard would like to allow any university member with a computer account to send electronic mail via the Arpanet gateway, but at the same time it wants to provide file transfer and remote login to select groups of users only. Currently Harvard is able to control remote login use because for internal resource control purposes, it does so anyway within the internal network. Remote login is a restricted command on all internal hosts and because only certain users can use it internally, only those users can use it through the gateway. However, file transfer is not a restricted command; it is too common and useful a facility to even consider restricting internally. As a result, there are no controls on doing file transfer via the gateway. Because the gateway is a packet-forwarding gateway, and such information is not deducible directly from packet headers, controls that discriminate according to service type (i.e., remote login, file transfer, mail) and host are difficult to implement on such a gateway. If the gateway operated at a higher-level, it would be a more simple and modular task to restrict file transfer use to authorized users, because the connection would carry information about the affiliation of source and destination as well as the communication mode.

This discussion of level of interconnection is concerned most directly with what Sunshine refers to as service level and implementation approach. [18] Service level refers to the communication mode supported in the gateway, e.g., datagram, virtual circuit, file transfer, remote login, mail, etc. He distinguishes between two implementation approaches, *endpoint* (where the source and destination each act as an endpoint in the communication mode and each gateway passes lower-level information), and *hop-by-hop* (where each intermediate gateway acts as an endpoint of the

communication mode as well). I refer to the former as lower-level, or packet-level, and the latter as higher-level, interconnection.¹⁰ With respect to traditional interconnection concerns alone (not inter-organization concerns), Sunshine finds the hop-by-hop (higher-level) approach more appropriate where backward compatibility of protocols and immediate needs predominate and where user awareness of crossing network boundaries is acceptable. He finds the endpoint approach (packet-level) preferable when robustness and generality are important and there is more basis for agreement and conformance to standards. Sunshine's conclusions lend support to the argument in favor of higher-level (i.e., hop-by-hop) ION connection. When an organization interconnects to the outside, backward compatibility with internal protocols and procedures is still of primary importance. Similarly, expediency is often a key criterion for the interconnection and it is often desirable for the connection to be less than transparent so that insiders are conscious of their actions when communicating with outside entities. There are many examples of higher-level gateways in use today; for example: ISI's experimental mail gateway called *Intermail* [6]; IBM's SNA interconnection technique [1]; Cambridge university's X.25, local area to public packet-switched gateway [5]; The University College London's mail and remote login gateway to the Arpanet; The UUCP based network [9].

Even at higher levels the distinction between topological (routing) and logical (organization affiliation) information may be blurred. If an organization supports transit, and guidelines are not set for the assignment and structure of source and destination names, address and routing information can be confused with logical information. For example, if person Smith at QRS Inc. sends mail to Jones at XYZ Inc. via MIT's network, then if XYZ receives the source address as Smith.QRS%MIT, the XYZ ION gateway must figure out that the source's logical affiliation is QRS and not MIT. At the same time XYZ must be able to determine how to return a message, namely via MIT. UCLnet has experienced this problem in its mail connections to the Arpanet. As described earlier, mail from Ministry of Defense and civilian users must be treated differently. Some mail is forwarded to the Arpanet from other civilian research networks and hosts that are connected to UCL. The addresses assigned to mailboxes on these hosts are constructed so that the mailboxes appear to lie within UCL. The mail gateway relies on a list of registered users to filter mail, in part because the organization affiliation of a user is not necessarily evident from the header. [4] However, in general, textual mailbox-names carry more semantic information and are taken from a larger namespace than network numbers. Therefore textual mailbox names can be constructed in such a way that the correct affiliation can be interpreted using easy to follow guidelines.¹¹

¹⁰Application level refers to the endpoints in the application itself and therefore application-level controls are even higher than the higher-level communication controls discussed here. For example, where a high-level gateway would forward a file transfer request without looking at the content of the request, an application-level gateway would interpret the request itself.

Electronic messaging is the higher-level protocol that is most commonly supported across organization boundaries, i.e., in IONs. Although message based interconnection is used primarily for person-to-person communication, it can be used to access computer-based services that can tolerate extended response times (e.g. printing, file retrieval, some data base updates and queries, etc.).¹² For example, Arpanet users invoke a fabrication system called MOSIS (developed by Information Sciences Institute at USC) via electronic mail. [16] Clients use specially formatted messages to enter designs into the fabrication queue, request status information, or check that a design is acceptable to the system. MOSIS interprets the design messages as textual descriptions of the geometry of masks for IC fabrication and the physical products are eventually shipped to the client via air carrier. As end-users increase their computing activities and autonomy with personal hardware and software, electronic mail can provide a convenient substrate via which to offer end-user developed servers. For this reasons alone, the number of message-invokable servers within organizations is likely to increase across a broad range of organization functions.

Several characteristics make message-based interconnection potentially appropriate for inter-organization communications and interchange. First, many ION applications are more loosely coupled than are internal applications and can be supported adequately by asynchronous message-based communications. Second, protocol conversion, which is often necessary between distinct organizations that have not coordinated equipment selection, is easier because it need not be achieved within real-time performance constraints. For the same reason, overall gateway implementation is less complex. Message-based communication accommodates filtering and access control more easily because it provides higher-level information than packet-based communication, and yet does not impose real-time constraints as do connection-based protocols. In addition, if internal users invoke most internal servers via connection-based protocols while external users have access to message-based protocols only, then only those internal servers that support message-based invocation will be accessible to external users.

The tradeoffs associated with visa and high-level interconnection are discussed further in the next section.

¹¹Arpanet Domain Name format is one possibility although it was not designed for this purpose and therefore may not be practical. [13]

¹²Message-based gateways operate at a higher protocol level than packet gateways, which also operate on a store-and-forward basis. A message is a complete semantic unit whose content, and address, a remote process, application, or person can interpret. Packets are limited-sized components of a connection or message. Any single packet may have no semantic meaning by itself and typically carries only low-level addressing information. To distinguish message-based gateways from lower-level packet gateways, the former are sometimes referred to as relays.

5. Tradeoffs in the Design and Implementation of ION Gateways

As described, the ION gateway's primary purpose is to associate packets, messages, or connections with access rights and either forward or reject them according to the designated policies. We therefore compare visa and high-level gateways based upon the cost and ease of implementing the following functions without violating the design constraints outlined in section 2:

- Identify logical communication characteristics, e.g., mode of access, organization affiliation of source and destination, amount of resources requested, etc.¹³ Depending upon the application, the gateway may need to employ mechanisms to authenticate the logical information that the gateway uses to make its policy decisions.
- Identify category sets that correspond to communication characteristics: i.e., manage the assignment of category sets to internal and external entities and manage the storage and lookup of this information in tables.
- Compare the identified category sets against the designated policy algorithm's definition of a permitted operation (e.g., the intersect rule).
- Leave ION application-level controls to the endpoint ION systems since that is where the information needed to make application-level decisions is available.

The two approaches differ from one another with respect to several of these ION gateway tasks. However, the most significant difference is in associating communications with logical information. The approaches also differ with respect to several performance parameters.

A high-level gateway can associate communications with logical information directly, or it can call an ACS. The visa scheme must employ a high-level dialog with an ACS to associate packets with logical information. Similarly, high-level gateways evaluate each connection or message according to programmed control policies. They may or may not apply some check to each successive packet in a connection. In addition to employing an ACS to apply a high-level control algorithm to each connection or message request, a visa-based gateway always checks each and every packet against the visa.

A more significant limitation of visas for packet forwarding gateways is that they must make decisions based on information in the packet header, which usually contains source and destination addresses only. The packet-level gateway must be able to evaluate the legitimacy of each packet based solely on the packet header and the visa. It is difficult, and sometimes impossible, to represent

¹³In many cases this is the most difficult aspect of implementing controls in an ION gateway. The solution is constrained by existing protocol and naming semantics which are not easily changed and most often were not designed with usage controls in mind.

complicated policies in this manner. For example, there is no way for a packet-level gateway to discriminate on the basis of mode of access (e.g., mail, file transfer, remote login, etc.) because no information about higher application levels is available in the packet headers.¹⁴ Consequently, even if mode of access is indicated in the visa, there is no way for the gateway to verify that a particular packet is supporting one mode of access and not another since this information is not carried in the packet header. The same problem arises if the gateway needs to discriminate on the basis of user ID. For this reason, higher-level gateways are better suited to implementation of some types of policies.

The two schemes are comparable in terms of several cost and performance criteria—storage and trusted components—but differ significantly in terms of others—end-user performance and protocol modification. Storage requirements are the same for both, although a high-level gateway may store control information locally instead of in an Access Control Server (ACS) and a visa gateway by definition stores it in an ACS. In addition, the visa gateway stores locally a small number of currently-in-use keys, whereas the high-level gateway maintains more state information about the connections passing through it. In both cases, the amount of storage used for access control information depends on the grain of control, i.e., user, host, network, organization. The two approaches are also similar in terms of the number and extent of components that must be trusted. In both cases, security depends upon the authentication of header and connection request information, the evaluation program in the gateway and ACS, and the ability to subvert the access control mechanisms used to approve connections or messages. The latter risk is somewhat higher for gateways that do not authenticate each packet.

The two schemes differ most in performance overhead and the modification required of existing protocols. Each of the methods exacts a performance cost. The visa gateway is costly because of the required dialog with the ACS and the checks applied on a per packet basis. The high-level gateway is costly because protocols are terminated and because the gateway must be programmed with each higher-level protocol that it supports. The tradeoff depends much on traffic patterns; in particular, the number of packets per session or message, the volume of traffic, and the number of communication service types. On the other hand, protocol conversion is hardest for the lower-level protocol because of the tighter real-time constraints; for the same reason protocol conversion is harder for connection-

¹⁴In the Arpanet, packet-level gateways can make a pretty good guess at the higher application because of the use of well-known ports. Specific port numbers are routinely used for specific applications throughout the Arpanet. Packet headers do contain the port number and therefore the higher level application can often be determined.

based gateways than it is for message based.

A second significant difference between the two methods is that visa gateways require that all internal systems that use the ION add the visa to the header or checksum calculation. This requires that each machine modify its low-level communication protocols. In contrast high-level gateways require that application-level procedures be changed; or, in some cases, only that name tables be updated. Although the latter is less transparent to the end user, the cost and inconvenience of software modification is avoided. This cost can be quite high if it implies incompatibility with existing and future equipment. On the other hand, an additional cost associated with higher-level gateways is the need to program the gateway separately for each higher-level protocol that the organization wants to support; in contrast, the packet-level gateway supports all higher-level applications.

In summary, the most difficult aspect of implementing ION gateways is the association of communications with logical information. Aside from this difficulty the major implementation decision is whether to interconnect at the packet level and employ an ACS and visa scheme, or whether to interconnect at higher levels and employ structured naming. Each approach is well suited to different environments and may be used in conjunction with one another in some cases. Finally, these controlled connections should be placed as close as possible to the administrative boundary being enforced.

6. Summary

This paper has characterized IONs and the access control and network interconnection issues raised therein. We began by characterizing a set of control requirements that are not addressed by traditional security mechanisms. Non-discretionary controls based on category sets and a simple intersect rule were integrated in a design that allows strictly-internal applications to be logically isolated from external interconnections without requiring physical isolation or universally-increased internal security. Based on this approach we identified and characterized these applications for which for which transparency, connectivity, and performance criteria alone are not adequate for selecting the interconnection method. We then proposed and evaluated higher-level and visa-based interconnections as alternatives to packet-level interconnection.

Acknowledgments

I wish to thank B. Baldwin, J. N. Chiappa, D. Clark, D. Feldmeier, C. Landwehr, S. Lipner, D. Reed, J. Saitzer, S. Sluizer, J. Sutherland, L. Zhang, and the anonymous reviewers for insightful comments and suggestions on the material presented herein.

7. Bibliography

1. Benjamin, J., Hess, M., Weingarten, R., Wheeler, W. "Interconnecting SNA Networks". *IBM Systems Journal* 22, 4 (1983), 344-366.
2. Biba, K. Integrity Considerations for Secure Computer Systems. Technical Report ESD-TR-76-372, The Mitre Corp., Bedford, MA, April, 1977.
3. Braden, R., Cole, R. Some Problems in the Inter-connection of Computer Networks. In *Pathways to the Information Society: Proceedings of the 6th International Conference on Computer Communications*, Williams, W., Ed., North-Holland, 1982, pp. 969-974.
4. Cole, R., Higginson, P., Lloyd, P., Moulton, R. "International net faces problems handling mail and file transfer". *Data Communications* (June 1983), 175-187.
5. Dallas, I. Implementation of a Gateway between a Cambridge Ring Local Area Network and a Packet Switching Wide Area Network. In *Pathways to the Information Society: Proceedings of the 6th International Conference on Computer Communications*, Williams, W., Ed., North-Holland, 1982, pp. 137-142.
6. DeSchon, A. MCI Mail/Arpa Mail Forwarding. Technical Report ISI/RR-84-141, USC Information Sciences Institute, August, 1984.
7. Estrin, D. Non-Discretionary Controls for Inter-Organization Networks. *Proceedings of the 1985 Symposium on Security and Privacy*, Silver Spring, MD, 1985, pp. 56-61.
8. Estrin, D. *Access to Inter-Organization Computer Networks*. Ph.D. Th., M.I.T., Department of Electrical Engineering and Computer Science, August 1985.
9. Horton, M. Standard for Interchange of USENET Messages. Request for Comments RFC 850, USC Information Sciences Institute, June, 1983.
10. ISO. Directives for the Technical Work of ISO. International Standards Organization, Geneva, Switzerland, 1982.
11. Karger, P. Non-Discretionary Access Control for Decentralized Computing Systems. S.M. Thesis. Massachusetts Institute of Technology, Dept. of Electrical Engineering and Computer Science, May, 1977. Also available from the M.I.T. Laboratory for Computer Science as TR-179.
12. Landwehr, C., Heitmeyer, C., McCleen, J. "A Security Model for Military Message Systems". *ACM Transactions on Computer Systems* 2, 3 (August 1984), 198-222.
13. Mockapetris, P. The Domain Name System. In *Computer-Based Message Services*, Smith, H., Ed., Elsevier Science Publishers B.V., North-Holland, 1984.
14. Mogul, J. Internet Subnets. Request for Comments RFC 917, USC Information Sciences Institute, October, 1984.
15. Mracck, J. Network Access Control in Multi-Net Internet Transport. S.B. Thesis, Massachusetts Institute of Technology, Dept. of Electrical Engineering and Computer Science, June, 1983.
16. Newell, A., Sproull, R. "Computer Networks: Prospects for Scientists". *Science* 215 (February 12 1982), 843-852.
17. Saltzer, J. On the Naming and Binding of Network Destinations. In *Local Computer Networks*, Ravisio, P.C., Hopkins, G., Naffah, N., Eds., North-Holland Publishing Company, New York, 1982, pp. 311-318.
18. Sunshine, C. "Interconnection of Computer Networks". *Computer Networks*, 1 (1977), 175-195.