                Security Label Framework for the Internet

Status of this Memo

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

Acknowledgements

   The members of the Privacy and Security Research Group and the
   attendees of the invitational Security Labels Workshop (hosted by the
   National Institute of Standards and Technology) helped me organize my
   thoughts on this subject. The ideas of these professionals are
   scattered throughout the memo.

1.0 Introduction

   This memo presents a security labeling framework for the Internet.
   The framework is intended to help protocol designers determine what,
   if any, security labeling should be supported by their protocols.
   The framework should also help network architects determine whether
   or not a particular collection of protocols fulfill their security
   labeling requirements. The Open Systems Interconnection Reference
   Model [1] provides the structure for the presentation, therefore OSI
   protocol designers may also find this memo useful.

2.0 Security Labels

   Data security is the set of measures taken to protect data from
   accidental, unauthorized, intentional, or malicious modification,
   destruction, or disclosure. Data security is also the condition that
   results from the establishment and maintenance of protective measures
   [2]. Given this two-pronged definition for data security, this memo
   examines security labeling as one mechanism which provides data
   security. In general, security labeling by itself does not provide
   sufficient data security; it must be complemented by other security
   mechanisms.

   In data communication protocols, security labels tell the protocol
   processing how to handle the data transferred between two systems.
   That is, the security label indicates what measures need to be taken
   to preserve the condition of security. Handling means the activities

performed on data such as collecting, processing, transferring, storing, retrieving, sorting, transmitting, disseminating, and controlling [3].

The definition of data security includes protection from modification and destruction. In computer systems, this is protection from writing and deleting. These protections implement the data integrity service defined in the OSI Security Architecture [4].

Biba [5] has defined a data integrity model which includes security labels. The Biba model specifies rule-based controls for writing and deleting necessary to preserve data integrity. The model also specifies rule-based controls for reading to prevent a high integrity process from relying on data that has less integrity than the process.

The definition of data security also includes protection from disclosure. In computer systems, this is protection from reading. This protection is the data confidentiality service defined in the OSI Security Architecture [4].

Bell and LaPadula [6] defined a data confidentiality model which includes security labels. The Bell and LaPadula model specifies rule-based controls for reading necessary to preserve data confidentiality. The model also specifies rule-based controls for writing to ensure that data is not copied to a container where confidentiality can not be guaranteed.

In both the Biba model and the Bell and LaPadula model, the security label is an attribute of the data. In general, the security label associated with the data remains constant. Exceptions will be discussed later in the memo, but relabeling is always the result of some network entity handling the data. Since the security label is an attribute of data, it should be bound to the data. When data moves through the network, the integrity security service [4] is generally used to accomplish this binding. If the communications environment does not include a protocol which provides the integrity security service to bind the security label to the data, then the communications environment should include other mechanisms to preserve this binding.

2.1 Integrity Labels

Integrity labels are security labels which support data integrity models, like the Biba model. The integrity label tells the degree of confidence that may be placed in the data and also indicates which measures the data requires for protection from modification and destruction.

As data moves through the network, the confidence that may be placed
in that data may change as a result of being handled by various
network components. Therefore, the integrity label is a function of
the integrity of the data before being transmitted on the network and
the path that the data takes through the network. The confidence
that may be placed in data does not increase because it was
transferred across a network, but the confidence that may be placed
in data may decrease as a result of being handled by arbitrary
network components. Entities are assigned integrity labels which
indicate how much confidence may be placed in data that is handled by
them. Thus, when data is handled by an entity with an integrity
label lower than the integrity label of the data, the data is
relabeled with the integrity label of the entity. Such relabeling
should be avoided by limiting the possible paths that data may take
through the network to those where the data will be handled only by
entities with the same or a higher integrity label than the data.

When integrity labels are used, each of the systems on a network must
implement the integrity model and the protocol suite must transfer
the integrity label with the data, if the confidence of the data is
to be maintained throughout the network. Each of the systems on a
network may have its own internal representation for a integrity
label, but the protocols must provide common syntax and semantics for
the transfer of the integrity label, as well as the data itself. To
date, no protocols have been standardized which include integrity
labels in the protocol control information.

2.2 Sensitivity Labels

Sensitivity labels are security labels which support data
confidentiality models, like the Bell and LaPadula model. The
sensitivity label tells the amount of damage that will result from
the disclosure of the data and also indicates which measures the data
requires for protection from disclosure. The amount of damage that
results from unauthorized disclosure depends on who obtains the data;
the sensitivity label should reflect the worst case.

As data moves through the network, it is processed by various network
components and may be mixed with data of differing sensitivity. If
these network components are not trusted to segregate data of
differing sensitivities, then all of the data processed by those
components must be handled as the most sensitive data processed by
those network components. For example, poor buffer management may
append highly sensitive data to the end of a protocol data unit that
was otherwise publicly releasable. Therefore, the sensitivity label
is a function of the sensitivity of the data before being transmitted
on the network and the most sensitive data handled by the network
components, and the trustworthiness of those network components. The

amount of damage that will result from the disclosure of the data
does not decrease because it was transferred across a network, but
the amount of damage that will result from the disclosure of the data
may increase as a result of being mixed with more sensitive data by
arbitrary network components. Thus, when data is handled by an
untrusted entity with a sensitivity label higher than the sensitivity
label of the data, the data is relabeled with the higher sensitivity
label. Such relabeling should be avoided by limiting the possible
paths that data may take through the network to those where the data
will be handled only by entities with the same sensitivity label as
the data or by using trustworthy network components. Entities with
lower sensitivity labels may not handle the data because this would
be disclosure.

When sensitivity labels are used, each of the systems on a network
must implement the sensitivity model and the protocol suite must
transfer the sensitivity label with the data, if the protection from
disclosure is to be maintained throughout the network. Each of the
systems on a network may have its own internal representation for a
sensitivity label, but the protocols must provide common syntax and
semantics for the transfer of the sensitivity label, as well as the
data itself. Sensitivity labels, like the ones provided by the IP
Security Option (IPSO) [7], have been used in a few networks for
years.

3.0 Security Label Usage

The Internet includes two major types of systems: end systems and
intermediate systems [1]. These terms should be familiar to the
reader. For this discussion, the definition of intermediate system
is understood to include routers, packet switches, and bridges. End
systems and intermediate systems use security labels differently.

3.1 End System Security Label Usage

When two end systems communicate, common security label syntax and
semantics are needed. The security label, as an attribute of the
data, indicates what measures need to be taken to preserve the
condition of security. The security label must communicate all of
the integrity and confidentiality handling requirements. These
requirements can become very complex.

Some operating systems label the data they process. These security
labels are not part of the data; they are attributes of the data.
Some database management systems (DBMSs) perform similar labeling.
The format of these security labels is a local matter, but they are
usually in a format different than the one used by the data
communication protocols. Security labels must be translated by these

operating systems and DBMSs between the local format and the format
used in the data communication protocols without any loss of meaning.

Trusted operating systems that implement rule-based access control
policies require security labels on the data they import [8,9].
These security labels permit the Trusted Computing Base (TCB) in the
end system to perform trusted demultiplexing. That is, the traffic
is relayed from the TCB to a process only if the process has
sufficient authorization for the data. In most cases, the TCB must
first translate the security label into the local syntax before it
can make the access control decision.

3.2 Intermediate System Security Label Usage

This section discusses "user" data security labels within the
intermediate system. The labeling requirements associated with
intermediate system-to-end system (IS-ES) traffic, intermediate
system-to-intermediate system (IS-IS) traffic, and intermediate
system-to-network management (IS-NM) traffic are not included in this
discussion.

Intermediate systems may make routing choices or discard traffic
based on the security label. The security label used by the
intermediate system should contain only enough information to make
the routing/discard decision and may be a subset of the security
label used by the end system. Some portions of the label may not
effect routing decisions, but they may effect processing done within
the end system.

In the Internet today, very few intermediate systems actually make
access control decisions. For performance reasons, only those
intermediate systems which do make access control decisions should be
burdened with parsing the security label. That is, information
hiding principles apply. Further, security labels which are to be
parsed only by end systems should not be visible to physical, data
link, or network layer protocols, where intermediate systems will
have to examine them.

Intermediate systems do not usually translate the security labels to
a local format. They use them "as is" to make their routing/discard
decisions. However, when two classification authorities share a
network by bilateral agreement, the intermediate systems may be
required to perform security label translation. Security label
translations should be avoided whenever possible by using a security
label format that is supported by all systems that will process the
security label. Since end systems do not generally know which
intermediate systems will process their traffic, security label
translation cannot always be avoided.

   Since security labels which are to be parsed only by end systems
   should not be carried by protocols interpreted by intermediate
   systems, such security labels should be carried by upper layer
   protocols, and end systems which use different formats for such
   security labels cannot rely on an intermediate systems to perform
   security label translation. Neither the Internet nor the OSI
   architecture includes such transformation functions in the transport,
   session, or presentation layer, which means that application layer
   gateways should be used to translate between different end system
   security label formats. Such application gateways should be avoided
   because they impinge on operation, especially when otherwise
   compatible protocols are used. This complication is another reason
   why the use of a security label format that is supported by all
   systems is desirable. A standard label syntax with registered
   security label semantics goes a long way toward avoiding security
   label translation [10].

4.0 Approaches to Labeling

   There are several tradeoffs to be made when determining how a
   particular network will perform security labeling. Explicit or
   implicit labels can be used. Also, security labels can either be
   connectionless or connection-oriented. A combination of these
   alternatives may be appropriate.

4.1 Explicit Versus Implicit Security Labels

   Explicit security labels are actual bits in the protocol control
   information (PCI). The IP Security Option (IPSO) is an example of an
   explicit security label [7]. Explicit security labels may be either
   connectionless or connection-oriented. The syntax and semantics of
   the explicit security label may be either tightly or loosely coupled.
   If the syntax and semantics are tightly coupled, then the explicit
   security label format supports a single security policy. If the
   syntax and semantics are loosely coupled, then the explicit security
   label format can support multiple security policies through
   registration. In both cases, software enforces the security policy,
   but the label parsing software can be written once if the syntax and
   semantics are loosely coupled. Fixed length explicit security label
   format parsers are generally faster than parsers for variable length
   formats. Intermediate systems suffer less performance impact when
   fixed length explicit security labels can be used, but end systems
   often need variable length explicit security labels to express data
   handling requirements.

   Implicit security labels are not actual bits in the PCI; instead,
   some attribute is used to determine the security label. For example,
   the choice of cryptographic key in the SP4 protocol [11] can

determine the security label. Implicit security labels may be either
connectionless or connection-oriented.

4.2 Connectionless Versus Connection-oriented Security Labels

When connectionless security labels are used, the security label
appears in every protocol data unit (PDU). The IP Security Option
(IPSO) [7] is an example of connectionless labeling. All protocols
have limits on the size of their PCI, and the explicit security label
cannot exceed this size limit. It cannot use the entire PCI space
either; the protocol has other fields that must be transferred as
well. This size limitation may prohibit explicit connectionless
security labels from meeting the requirements of end systems.
However, the requirements of intermediate systems are more easily
satisfied by explicit connectionless security labels.

Connection-oriented security labels are attributes of virtual
circuits, connections, and associations. For simplicity, all of
these are subsequently referred to as connections. Connection-
oriented security labels are used when the SDNS Key Management
Protocol (KMP) [12] is used to associate security labels with each of
the transport connection protected by the SP4 protocol [10,11] (using
SP4C). The security label is defined at connection establishment,
and all data transferred over the connection inherits that security
label. This approach is more compatible with end system requirements
than intermediate system requirements. One noteworthy exception is
X.25 packets switches; these intermediate systems could associate
connection-oriented labels with each virtual circuit.

Connectionless security labels may be used in conjunction with
connectionless or connection-oriented data transfer protocols.
However, connection-oriented security labels may only be used in
conjunction with connection-oriented data transfer protocols.

5.0 Labeling Within the OSI Reference Model

This section examines each of the seven OSI layers with respect to
security labels.

5.1 Layer 1, The Physical Layer

Explicit security labels are not possible in the Physical Layer. The
Physical Layer does not include any protocol control information
(PCI), so there is no place to include the bits which represent the
label.

Implicit security labels are possible in the Physical Layer. For
example, all of the data that comes in through a particular physical

port could inherit one security label. Most Physical Layer
communication is connectionless, supporting only bit-at-a-time or
byte-at-a-time operations. Thus, these implicit security labels are
connectionless.

Implicit security labels in the Physical Layer may be used to meet
the requirements of either end systems or intermediate systems so
long as the communication is single level. That is, only one
security label is associated with all of the data received or
transmitted through the physical connection.

5.2 Layer 2, The Data Link Layer

Explicit security labels are possible in the Data Link Layer. In
fact, the IEEE 802.2 Working Group is currently working on an
optional security label standard for the Logical Link Control (LLC)
protocol (IEEE 802.2) [13]. These labels will optionally appear in
each LLC frame. These are connectionless security labels.

Explicit connection-oriented security labels are also possible in the
Data Link Layer. One could imagine a security label standard which
worked with LLC Type II.

Of course, implicit security labels are also possible in the Data
Link Layer. Such labels could be either connectionless or
connection-oriented. One attribute that might be used in IEEE 802.3
(CSMA/CD) [14] to determine the implicit security label is the source
address of the frame.

Security labels in the Data Link Layer may be used to meet the
requirements of end systems and intermediate systems (especially
bridges). Explicit security labels in this layer tend to be small
because the protocol headers for data link layer protocols are
themselves small. Therefore, when end systems require large security
labels, a higher protocol layer should used to carry them. However,
when end systems do not require large security labels, the data link
layer is attractive because in many cases the data link layer
protocol supports several protocol suites simultaneously. Label-
based routing/relay decisions made by bridges are best supported in
this layer.

5.3 Layer 3, The Network Layer

Explicit security labels are possible in the Network Layer. In fact,
the IP Security Option (IPSO) [7] has been used for many years.
These labels optionally appear in each IP datagram. IPSO labels are
obviously connectionless security labels.

Explicit connection-oriented security labels are also possible in the Network Layer. One could easily imagine a security label standard for X.25 [15], but none exists.

Of course, implicit security labels are also possible in the Network Layer. These labels could be either connectionless or connection-oriented. One attribute that might be used to determine the implicit security label is the X.25 virtual circuit.

Security labels in the Network Layer may be used to meet the requirements of end systems and intermediate systems. Explicit security labels in this layer tend to be small because the protocol headers for network layer protocols are themselves small. Small fixed size network layer protocol headers allow efficient router implementations. Therefore, when end systems require large security labels, a higher protocol layer should used to carry them. Alternatively, the Network Layer (especially the Subnetwork Independent Convergence Protocol (SNICP) sublayer) is an excellent place to carry a security label to support trusted demultiplexing, because many implementations demultiplex from an system-wide daemon to a user process after network layer processing. The SNICP is end-to-end, yet it is low enough in the protocol stack to aid trusted demultiplexing.

Label-based routing/relay decisions made by routers and packet switches are best supported in the Network Layer. Routers can also add labels at subnetwork boundaries. However, placement of these security labels must be done carefully to ensure that their addition does not degrade overall network performance by forcing routers that do not make label-based routing decisions to parse the security label. Also, performance will suffer if the addition of security labels at subnet boundaries induces fragmentation/segmentation.

5.4 Layer 4, The Transport Layer

Explicit security labels are possible in the Transport Layer. For example, the SP4 protocol [10,11] includes them. These labels can be either connectionless (using SP4E) or connection-oriented (using SP4C). SP4 is an addendum to the TP [16] and CLTP [17] protocols.

Implicit security labels are also possible in the Transport Layer. Such labels could be either connectionless or connection-oriented. One attribute that might be used to determine the implicit label in the SP4 protocol (when explicit labels are not used as discussed above) is the choice of cryptographic key.

Security labels in the Transport Layer may be used to meet the requirements of end systems. The Transport Layer cannot be used to

meet the requirements of intermediate systems because intermediate
systems, by definition, do not process protocols above the Network
Layer. Connection-oriented explicit security labels in this layer
are especially good for meeting end system requirements where large
labels are required. The security label is transmitted only at
connection establishment, so overhead is kept to a minimum. Of
course, connectionless transport protocols may not take advantage of
this overhead reduction technique. Yet, in many implementations the
Transport Layer is low enough in the protocol stack to aid trusted
demultiplexing.

5.5 Layer 5, The Session Layer

Explicit security labels are possible in the Session Layer. Such
labels could be either connectionless or connection-oriented.
However, it is unlikely that a standard will ever be developed for
such labels because the OSI Security Architecture [4] does not
allocate any security services to the Session Layer, and the Internet
protocol suite does not have a Session Layer.

Implicit security labels are also possible in the Session Layer.
These implicit labels could be either connectionless or connection-
oriented. Again, the OSI Security Architecture makes this layer an
unlikely choice for security labeling.

Security labels in the Session Layer may be used to meet the
requirements of end systems, but the Session Layer is too high in the
protocol stack to support trusted demultiplexing. The Session Layer
cannot be used to meet the requirements of intermediate systems
because intermediate systems, by definition, do not process protocols
above the Network Layer. Security labels in the Session Layer do not
offer any advantages to security labels in the Transport Layer.

5.6 Layer 6, The Presentation Layer

Explicit security labels are possible in the Presentation Layer. The
presentation syntax may include a security label. This approach
naturally performs translation to the local label format and supports
both connectionless and connection-oriented security labeling.

Implicit security labels are also possible in the Presentation Layer.
Such labels could be either connectionless or connection-oriented.

Security labels in the Presentation Layer may be used to meet the
requirements of end systems, but the Presentation Layer is too high
in the protocol stack to support trusted demultiplexing. The
Presentation Layer cannot be used to meet the requirements of
intermediate systems because intermediate systems, by definition, do

not process protocols above the Network Layer. To date, no
Presentation Layer protocols have been standardized which include
security labels.

5.7 Layer 7, The Application Layer

Explicit security labels are possible in the Application Layer. The
CCITT X.400 message handling system includes security labels in
message envelopes [18]. Other Application Layer protocols will
probably include security labels in the future. These labels could
be either connectionless or connection-oriented. Should security
labels be incorporated into transaction processing protocols and
message handling protocols, these will most likely be connectionless
security labels; should security labels be incorporated into other
application protocols, these will most likely be connection-oriented
security labels. Application layer protocols are unique in that they
can include security label information which is specific to a
particular application without burdening other applications with the
syntax or semantics of that security label.

Store and forward application protocols, like electronic messaging
and directory protocols, deserve special attention. In terms of the
OSI Reference Model, they are end system protocols, but multiple end
systems cooperate to provide the communications service. End systems
may use security labels to determine which end system should be next
in a chain of store and forward interactions; this use of security
labels is very similar to the label-based routing/relay decisions
made by routers except that the security labels are carried in an
Application Layer protocol. Also, Application Layer protocols must
be used to carry security labels in a store and forward application
when sensitivity labels must be concealed from some end systems in
the chain or when some end systems in the chain are untrustworthy.

Implicit security labels are also possible in the Application Layer.
These labels could be either connectionless or connection-oriented.
Application title or well know port number might be used to determine
the implicit label.

Security labels in the Application Layer may be used to meet the
requirements of end systems, but the Application Layer is too high in
the protocol stack to support trusted demultiplexing. The
Application Layer cannot be used to meet the requirements of
intermediate systems because intermediate systems, by definition, do
not process protocols above the Network Layer.

6.0 Summary

   Very few hard rules exist for security labels. Internet architects
   and protocol designers face many tradeoffs when making security label
   placement decisions. However, a few guidelines can be derived from
   the preceding discussion:

   First, security label-based routing decisions are best supported by
   explicit security labels in the Data Link Layer and the Network
   Layer. When bridges are making the routing decisions, the Data Link
   Layer should carry the explicit security label; when routers are
   making the routing decisions, the Network Layer should carry the
   explicit security label.

   Second, when security labels are specific to a particular application
   it is wise to define them in the application protocol, so that these
   security labels will not burden other applications on the network.

   Third, when trusted demultiplexing is a concern, the Network Layer
   (preferably the SNICP) or Transport Layer should be used to carry the
   explicit security label. The SNICP or transport protocol are
   especially attractive when combined with a cryptographic protocol
   that binds the security label to the data and protects the both
   against undetected modification.

   Fourth, to avoid explicit security label translation, a common
   explicit security label format should be defined for the Internet.
   Registration of security label semantics should be used so that many
   security policies can be supported by the common explicit security
   label syntax.

References

   [1] ISO Open Systems Interconnection - Basic Reference Model (ISO
       7498). International Organization for Standardization, 1981.

   [2] Dictionary of Military and Associated Terms (JCS Pub 1). Joint
       Chiefs of Staff. 1 April 1984.

   [3] Security Requirements for Automatic Data Processing (ADP) Systems
       (DODD 5200.28). Department of Defense. 21 March 1988.

   [4] Information Processing Systems - Open Systems Interconnection
       Reference Model - Security Architecture (ISO 7498-2).
       Organization for Standardization, 1988.

   [5] Biba, K. J. "Integrity Considerations for Secure Computer
       Systems", MTR-3153, The Mitre Corporation, April 1977.

   [6] Bell, D. E.; LaPadula, L. J.   "Secure Computer System: Unified
       Exposition and Multics Interpretation", MTR-2997, The MITRE
       Corporation, March 1976.

   [7] Kent, S.   "U.S. Department of Defense Security Options for the
       Internet Protocol", RFC 1108, BBN Communications, November 1992.

   [8] Trusted Computer System Evaluation Criteria (DoD 5200.28-STD)
       National Computer Security Center, 26 December 1985.

   [9] Trusted Network Interpretation of the Trusted Computer System
       Evaluation Criteria, (NCSC-TG-005, Version-1). National Computer
       Security Center, 31 July 1987.

  [10] Nazario, Noel (Chairman). "Standard Security Label for GOSIP An
       Invitational Workshop", NISTIR 4614, June 1991.

  [11] Dinkel, Charles (Editor). "Secure Data Network System (SDNS)
       Network, Transport, and Message Security Protocols", NISTIR 90-
       4250, February 1990, pp 39-62.

  [12] Dinkel, Charles (Editor). "Secure Data Network System (SDNS) Key
       Management Documents", NISTIR 90-4262, February 1990.

  [13] IEEE Standards for Local Area Networks: Logical Link Control,
       IEEE 802.2. The Institute of Electrical and Electronics
       Engineers, Inc, 1984.

  [14] IEEE Standards for Local Area Networks: Carrier Sense Multiple
       Access with Collision Detection (CSMA/CD) Access Method and
       Physical Layer Specification, IEEE 802.3. The Institute of
       Electrical and Electronics Engineers, Inc, 1985.

  [15] Recommendation X.25, Interface Between Data Terminal Equipment
       (DTE) and Data Circuit Terminating Equipment (DCE) for Terminals
       Operating in the Packet Mode on Public Data Networks.
       Consultative Committee, International Telephone and Telegraph
       (CCITT), 1984.

  [16] Information Processing Systems - Open Systems Interconnection -
       Connection oriented transport protocol specification (ISO 8073).
       Organization for Standardization, 1985. [Also ISO 8208]

  [17] Information Processing Systems - Open Systems Interconnection -
       Protocol for providing the connectionless-mode transport service
       (ISO 8602). Organization for Standardization, 1986.

   [18] Recommendation X.411, Message Handling Systems: Message Transfer
        System: Abstract Service Definition and Procedures. Consultative
        Committee, International Telephone and Telegraph (CCITT), 1988.
        [Also ISO 8883-1]

Security Considerations

   This entire memo is devoted to a discussion of a Framework for
   labeling information for security purposes in network protocols.

Author's Address

   Russell Housley
   Xerox Special Information Systems
   7900 Westpark Drive
   McLean, Virginia 22102

   Phone: 703-790-3767
   EMail: Housley.McLean_CSD@Xerox.COM