Stable Implementation Agreements for Open Systems Interconnection Protocols: Part 4 - Transport

Output from the December 1993 Open Systems Environment Implementors' Workshop (OIW)

SIG Chair: **Fred Burg, AT&T**
SIG Editor: **Brenda Gray**

Foreword

This part of the Stable Implementation Agreements was prepared by the Lower Layers Special Interest Group (LLSIG) of the Open Systems Environment Implementors' Workshop (OIW). See Part 1 - Workshop Policies and Procedures of the "Draft Working Implementation Agreements Document" for the charter.

Text in this part has been approved by the Plenary of the above-mentioned Workshop. This part replaces the previously existing chapter on this subject.

Future changes and additions to this version of these Implementor Agreements will be published as change pages. Deleted and replaced text will be shown as strikeout. New and replacement text will be shown as shaded.

Table of Contents

List of Figures

List of Tables

Part 4 - Transport

0 Introduction

These agreements support the integration of LANs, packet networks, and other WANs with the smallest possible set of mandatory protocol sets, in accordance with the other agreements already reached. Nothing here shall preclude vendors from implementing protocol suites in addition to the ones described in this document.

1 Scope

This part presents agreements for providing the OSI Transport layer services over both connection mode and connectionless mode services.

2 Normative References

2.1 CCITT

[**1**] Recommendation X.214 (Blue Book, 1988), *Transport Service Definition for Open Systems Interconnection for CCITT Applications*.

[**2**] Recommendation X.224 (Blue Book, 1988), *Transport Protocol Specification for Open Systems Interconnection for CCITT Applications*.

2.2 ISO

[**3**] ISO 8072, *Information processing systems - Open systems interconnection - Transport service defintion*.

[**4**] ISO 8072 Addendum 1, *Information processing systems - Open systems interconnection - Addendum 1: Transport service definition - Connectionless-mode transmission*.

[**5**]

[**5**] ISO/IEC 8073:199x,Edition 3, *Information Technology-Telecommunications and Information Exchange Between Systems - Open Systems Interconnection - Protocol for Providing the Connection-mode Transport Service, (SC6N7589 Rev)*.

[**6**] ISO 8602, *Information processing systems - Open systems interconnection - Protocol for providing the connectionless-mode transport service*.

[**7**] ISO/IEC 10736, *Information Technology - Telecommunications and Information Exchange Between Systems -Transport Layer Security Protocol*

3 Status

Completed December 1993.

4 Errata

NOTE - This clause may contain "defect report" and resolutions material, and the versions of implementor agreements to which this material applies.

5 Provision of Connection Mode Transport Service

Three connection mode protocol classes have been identified for implementation. Transport classes 0, 2 and 4 of X.224 (1988)[1] have been endorsed for use over CONS. Only Transport Class 4 of ISO 8073/Add. 2 [2] has been endorsed for use over CLNS. The following class combinations are endorsed for CONS: (0), (0,2) or (0,2,4).

5.1 Transport Class 4

5.1.1 Transport Class 4 Overview

Transport Class 4 is mandatory for communication between systems using the OSI CLNS and may also be used for systems using the OSI CONS (e.g., a private MHS, etc.).

5.1.2 Protocol Agreements

A disconnect request shall be issued in response to a connect request when the maximum number of Transport connections is reached or exceeded.

---

1 Where a CR TPDU proposing Class 2 or 4 is initiated, Class 0 shall be explicitly indicated as an alternative class except if there is already one (or several) transport connection(s) assigned to the network connection (multiplexing being possible).

2 In general, references to ISO 8073 in ISO 8073/Add. 2 should be interpreted as applying to X.224 (1988); however, the reference to Clause 14.6.a in Clause 14 of ISO 8073/Add. 2 should be interpreted as a reference to Clause 14.5.a of X.224(1988).

5.1.2.1 General Rules

The rules are as follows:

a) All implementations shall request "use of extended formats" in the CR TPDU. Implementations shall accept the "use of extended formats" in the CC TPDU if it was proposed in the CR TPDU. Implementations shall accept "use of normal formats" if it was proposed in the CR TPDU;

b) Negotiation of protection is outside the scope of these agreements. If negotiation of protection is not supported, receipt of the protection parameters in CR TPDU and CC TPDU shall be ignored;

c) Implementations shall be capable of proposing and accepting the non-use of checksums;

NOTE - See clause 8.2 for more information on checksums when the Transport Protocol and the Transport Layer Security Protocol are both implemented.

d) Use of the acknowledgment time parameter is optional. If an implementation is operating any policy which delays the transmission of AK TPDUs, the maximum amount of time by which a single AK TPDU may be delayed shall be indicated to the peer Transport service provider using the acknowledgment time parameter. The value transmitted should be expressed in units of milliseconds and rounded up to the nearest whole millisecond;

e) QoS negotiation is outside the scope of these agreements. If QoS negotiation is not supported, receipt of the parameters "throughput," "residual error rate," "priority," and "transit delay" in the CR and CC TPDUs shall be ignored;

f) It is recommended that implementations not send user data in the CR TPDU or the CC TPDU. The disposition of any user data received in a CR TPDU or CC TPDU is implementation dependent;

g) It is recommended that implementations not send user data in the **DR TPDU**. The disposition of any user data received in a **DR TPDU** is implementation dependent;

h) An unknown parameter in any received CR TPDU shall be ignored;

i) A Transport entity shall accept a DR TPDU and a corresponding DC TPDU with or without a checksum in response to a CR or CC TPDU;

j) Transmitted DR TPDUs shall carry a disconnect reason code which pertains to the actual cause of the disconnect. A DR TPDU may carry a reason code of "0" (unspecified) if an appropriate reason code is not defined;

k) Known parameters with valid lengths but with invalid values in a CR TPDU shall be handled as follows:

1) Parameter;	2) Action:
a) TSAP id	a) Send DR TPDU

b) TPDU size	b) ignore parameter, use default
c) Version	c) ignore parameter, use default
d) Checksum	d) discard CR TPDU
e) Alternate Protocol Classes	e) Protocol Error

l) Unrecognized or not applicable bits of the Additional Options parameter shall be ignored.

m) It is recommended that the capability of request acknowledgments be supported and proposed in CR TPDUs. If request acknowledgments are supported, then if the implementation delays acknowledgments it shall:

1) request use of request acknowledgments in the CR TPDU;

2) accept the use of request acknowledgments in the CC TPDU if it was proposed in the CR TPDU.

n) It is recommended that implementations send both the preferred and existing TPDU size parameters in the CR TPDU.

o) It is recommended that inactivity timer values be exchanged during connection establishment. This may be mandatory in the future. If the "exchange of inactivity timers" capability is supported, the implementation shall send its minimum inactivity timer in the CR TPDU. If a CR TPDU is received with this timer value and the capability is supported, the responding CC TPDU shall contain the inactivity time.

If the Inactivity time is received and the capability is supported, the following shall be used as an upper bound for W:

$$I_R/N > W \qquad N \geq 2$$

5.1.2.2 Transport Class 4 Service Access Points or Selectors

If present, the TSAP Id. field in the CR and CC TPDUs shall be encoded as a variable length field and will be interpreted as an octet string. The length of the string cannot exceed 32 octets.

5.1.2.3 Retransmission Timer

It is recommended that the value used for the retransmission timer be based upon the round-trip delay experienced on a transport connection. The implementation should maintain, and continually update, an estimate of the round-trip delay for the TC. From this estimate, a value for the retransmission timer is calculated each time it is started. Example techniques for maintaining the estimate and calculating the retransmission timer are described below. Example 1 represents a simple retransmission strategy and example 2 is particularly suitable for networks subject to high traffic loads.

Example 1

The value of the retransmission timer may be calculated according to the following formula:

$$T1 \longleftarrow kE + AR.$$

In this formula, E is the current estimate of the round-trip delay on the transport connection, AR is the value of the acknowledgment time parameter received from the remote transport service provider during connection establishment, and k is some locally administered factor.

A value for k should be chosen to keep the retransmission timer sufficiently small such that lost TPDUs will be detected quickly, but not so small that false alarms are generated causing unnecessary retransmission.

The value of E may be calculated using an exponentially weighted average based upon regular sampling of the interval between transmitting a TPDU and receiving the corresponding acknowledgment. Samples are taken by recording the time of day when a TPDU requiring acknowledgment is transmitted and calculating the difference between this and the time of day when the corresponding acknowledgment is received. New samples are incorporated with the existing average according to the following formula:

$$E \longleftarrow E + (1 - \alpha)(S - E).$$

In this formula, S is the new sample and α is a parameter which can be set to some value between 0 and 1. The value chosen for α determines the relative weighting placed upon the current estimate and the new sample. A large value of α weights the old estimate more heavily causing it to respond only slowly to variations in the round-trip delay. A small value weights the new sample more heavily causing a quick response to variations. (Note that setting α to 1 will effectively disable the algorithm and result in a constant value for E, being that of the initial seed.)

If α is set to $1\text{-}2^{-n}$ for some value of n, the update can be reduced to a subtract and shift as shown below:

$$E \longleftarrow E + 2^{-n} (S - E).$$

When sampling, if an AK TPDU is received which acknowledges multiple DT TPDUs, only a single sample should be taken being the round-trip delay experienced by the most recently transmitted DT TPDU. This attempts to minimize in the sample any delay caused by the remote transport service provider withholding AK TPDUs.

Example 2

As network load increases, the variability of round-trip delay also increases. In environments where load fluctuates widely, it is therefore useful to estimate the variability of round-trip delay measurements and use this estimate in the calculation of retransmission timer values. An estimate of the variability of round-trip delay measurements can be efficiently calculated as an exponentially weighted average of the differences between round-trip delay measurements and the average round-trip delay. This represents the mean deviation of the round-trip delays, which is a useful approximation of the standard deviation and can be much more efficiently computed. The formula is

$$D \text{ <-- } D + (1 - a)(|S - E| - D)$$

where *D* is the estimate of variability in round-trip delays. *S, E,* and *a* are as defined for the preceding formula. As before the value of *a* must be between 0 and 1 and the choice of a value of $1 - 2^{-N}$ allows for efficient update of the average. The value of *a* for the variability estimation, though, does not need to be the same as that used for the round-trip delay estimate. A smaller value for *a* is useful in the variability estimation to cause a more rapid response to changes in round-trip delays. *D* can then be used to calculate retransmission timer values according to the formula:

T1 <-- E + AR + kD

where *T1* is the retransmission timer value, *E* is the estimated average round-trip delay, *AR* is the value of the acknowledgment timer parameter received from the remote transport service provider during connection establishment, and *k* is a locally administered factor. Since *D* approximates the standard deviation of the round-trip delays, but is greater than or equal to the standard deviation, round-trip delays within *k* standard deviations of the mean would be accounted for by the retransmission timer value (e.g., *k* = 2, if round-trip delays were normally distributed, would account for 95% of the variability).

Round-trip time measurements based on acknowledgment of any retransmitted data should not be used to update the round-trip delay estimate or the estimate of variability. Such measurements are not reliable since it is ambiguous which transmission of the data is being acknowledged.

One strategy for handling a retransmission timeout is to retransmit the PDU and reset the timer with a value that is twice the previous value. In this case, a new roundtrip delay estimate and estimate of variability should be calculated only when an acknowledgment of data is received where none of the acknowledged data has been retransmitted. This calculation uses the new round-trip delay measurement and the last estimate before the retransmission timeout(s).

5.1.2.4 Keep-Alive Function

The Class 4 protocol detects a failed Transport connection by use of an "inactivity timer." This timer is reset each time a TPDU is received on a connection. If the timer ever expires, the connection is terminated.

The Class 4 protocol maintains an idle connection by periodically transmitting an AK TPDU upon expiration of the "window timer." Thus, in a simple implementation, the interval of one transport entity's window timer must be less than that of its peer's inactivity timer, and vice versa. The following agreements permit communicating transport entities to maintain an idle connection without shared information about timer values:

a) In accordance with ISO 8073/X.224, Clause 12.2.3.9.a, all implementations <u>must</u> respond to the receipt of a duplicate AK TPDU not containing FCC by transmitting an AK TPDU containing the "flow control confirmation" parameter;

b) Implementations <u>must</u> always transmit duplicate AK TPDUs without FCC on expiration of the local window timer (see ISO 8073/X.224, Clause 12.2.3.8.1). Receipt of this TPDU by the remote Transport entity will cause it to respond with an AK TPDU containing the "flow control confirmation" parameter. When this is received by the local transport entity, it will reset its inactivity timer. See figure 1;

c) It is a local matter for an implementation to set the intervals of its timers to appropriate relative values. Specifically:

1) The window timer must be greater than the round-trip delay. See 5.1.2.3;

2) The inactivity timer must be greater than two times the window timer; and should normally be an even greater multiple if the Transport connection is to be resilient to the loss of an AK TPDU.

A duplicate AK TPDU (see figure 1) is one which contains the same values for YR-TU-NR, credit, and subsequence number as the previous AK TPDU transmitted. A duplicate AK TPDU does not acknowledge any new data, nor does it change the credit window.

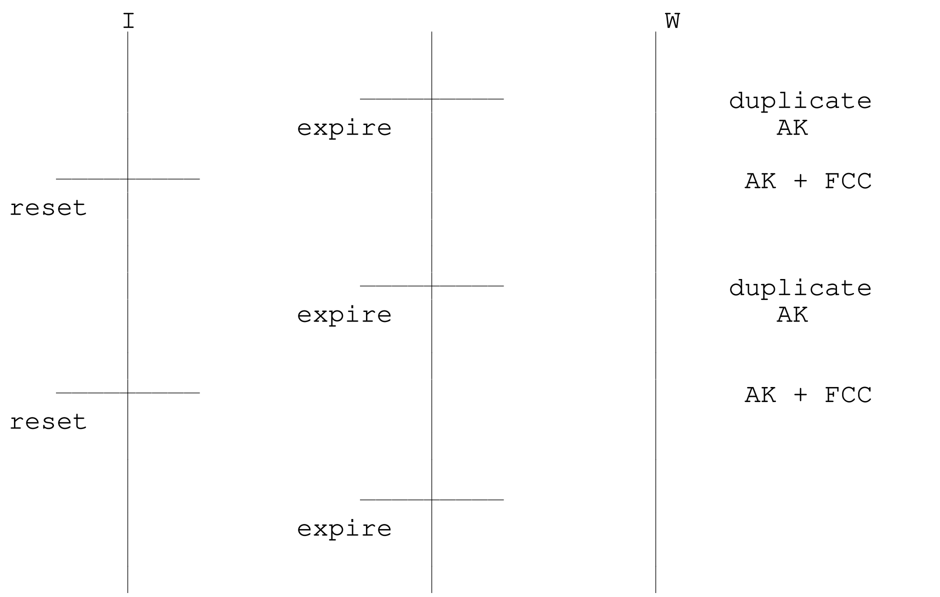

Figure 1 - AK exchange on idle connection.

5.1.2.5 Congestion Avoidance Policies

This clause defines both mandatory and optional requirements relating to avoiding congestion in OSI networks and recovering from it when it is experienced. The mandatory requirements specify a minimum approach to congestion avoidance/recovery which can be tuned based upon the specific requirements of the network. The optional requirements specify a dynamic window sizing scheme which, if implemented, will contribute further to the avoidance of congestion in the network.

Mandatory Requirements are as follows:

a) A maximum size for the "receive credit window," the value of which is locally configurable, should be provided. A "receive credit window" reflects the number of credits sent by a Transport entity for a Transport connection. The maximum size of the "receive credit window" shall be referred to as WR_1;

b) A maximum size for the "sending credit window," the value of which is locally configurable, shall be provided. A "sending credit window" reflects the number of data TPDUs that a Transport entity is willing to send on a Transport connection. The maximum size of the "sending credit window" shall be referred to as WS_1. As specified in ISO 8073, the "sending credit window" shall also be less than or equal to the remote "receive credit window" as conveyed in the last CDT field;

c) It is strongly recommended that an implementation use a retransmission timer per Transport connection. If, upon expiration of the retransmission timer, an implementation allows more than "1" TPDU to be transmitted a means to locally adjust the maximum number shall be provided;

d) All implementations shall have the capability of operating without delaying ACKs of data TDPUs received in-sequence (i.e., A_L essentially equals zero). If an implementation optionally chooses to explicitly delay ACKs, a means to locally adjust A_L shall be provided.

Optional Requirements are as follows:

For systems implementing the dynamic window sizing scheme the following rules apply as described below:

1. RECEIVING TRANSPORT ENTITY (RTE) RULES:

a) Rule 1 - Initialization of Window:

1) The initial value of WR (known as WR_0) shall have a locally configurable upper bound. This window is sent to the sending transport entity (STE) in the next CDT field transmitted;

a) Rule 2 - Required Sampling Period:

1) All RTEs shall maintain a fixed value for WR until the next 2WR DT TPDU arrive since the last CDT field was transmitted by the RTE;

b) Rule 3 - Required Counting of Received TPDUs in a Sampling Period:

1) All RTEs shall maintain a count, N, equal to the total number of TPDUs received and a count, NC, equal to the total number of TPDUs received which had the CE Flag set. All types of TPDUs are included in the counts for N and NC, not just DT TPDUs;

c) Rule 4 - Required Action upon the end of a Sampling Period: All RTEs shall take the following actions at the end of each sampling period:

1) If the count NC is less than 50 percent of the count N, the RTE shall increase WR by adding 1 up to a maximum, WR_1, (that is based on the local buffer management policy); otherwise, it shall decrease WR by multiplying by 0.875 (a minimum of 1);

2) Reset N and NC to zero;

3) Transmit the new window WR in the next CDT field sent to the sending

transport entity;

2) SENDING TRANSPORT ENTITY (STE) RULES:

a) Rule 1: Initialization of Window:

1) All STEs shall maintain a sending window size (WS). Initially and also as long as there is no loss, WS is set equal to the receiving window value WR received from the remote RTE in the last CDT field;

b) Rule 2: Required Action on a Timeout;

1) All STEs shall reset WS to one when the retransmissions timer expires and indicates a lost TPDU. WS now limits the number of DT TPDUs that may be transmitted or retransmitted without further acknowledgments;

c) Rule 3: Required Counting of Acknowledged TPDU:

1) All STEs shall maintain a count, ACKRCVD of the number of DT TPDUs acknowledged, by the RTE, since WS was last adjusted. Therefore each time WS is adjusted, the count ACKRCVD shall be reset to zero;

d) Rule 4: Increase Window Policy:

1) All STEs shall increase WS by one each time ACKRCVD is equal to or greater than the current value of WS, unless WS exceeds the window permitted by the remote RTE.

5.1.2.6 Use of Priority

(Refer to the Working Implementation Agreements).

5.2 Transport Class 0

5.2.1 Transport Class 0 Overview

Transport Class 0 over X.25 is mandatory (see X.400) for use in communicating with public MHS systems operating in accordance with the CCITT X.400 series recommendations. The purpose of the agreements concerning Transport Class 0 is to allow connection to these public services. Transport Class 0 over X.25 can also be used in communicating between PRMDs (this choice is prevalent outside North America).

5.2.2 Protocol Agreements

5.2.2.1 General Rules

Transport Class 0 agreements are as follows:

a) The Error (ER) TPDU may be used at any time and upon receipt requires that the recipient disconnect the network connection, and by extension the transport connection;

b) The allowed values for the maximum TPDU size are 128, 256, 512, 1024, and 2048 octets;

c) The Class 0 protocol does not support multiplexing. At any instant, one Transport connection corresponds to one Network connection;

d) It is recommended that the optional timers TS1 and TS2, if implemented, be settable by local system management. Values in the order of minutes should be supported;

e) An unlimited TSDU length must be supported.

f) It is recommended that implementations send both the preferred and existing TPDU size parameters in the CR TPDU.

5.2.2.2 Transport Class 0 Service Access Points

For communicating with public MHS systems, section 5 of X.410 specifies the use and format of TSAP identifiers.

5.2.3 Rules for Negotiation

The rules for class negotiation shall be used.

5.3 Transport Class 2

5.3.1 Transport Class 2 Overview

Transport Class 2 is applicable in OSI end systems which provide the Connection-mode Network Service.

5.3.2 Protocol Agreements

Transport Class 2 agreements follow:

a) The values of the TS1 and TS2 timers shall be configurable. The recommended timer values are:

1) TS1: 60 seconds;

2) TS2: 60 seconds;

b) If present, the TSAP-id field in the CR and CC TPDUs shall be encoded as a variable length field and will be interpreted as an octet string. The length of the string cannot exceed 32 octets;

c) The rules for class negotiation shall be used;

d) QoS negotiation is outside the scope of these agreements. If QoS negotiation is not supported, receipt of the parameters "throughput," "residual error rate," "priority," and "transit delay" in the CR and CC TPDU shall be ignored.

NOTE - If Class 0 is indicated in the Alternative Protocol Class field and QoS parameters are conveyed and the responding end system chooses Class 0, then the QoS parameters have been ignored by the responding system.

e) It is recommended that implementations send both the preferred and existing TPDU size parameters in the CR TPDU.

6 Provision of Connectionless Transport Service

ISO 8072/Add. 2 is the Transport Service Definition covering Connectionless-mode Transmission. ISO 8602 is the Protocol for providing the Connectionless-Mode Transport Service.

6.1 Connectionless Transport Overview

When providing the connectionless Transport Service, the protocol shall be implemented as specified in ISO 8602.

6.2 Protocol Agreements

6.2.1 General Rules

The connectionless Transport protocol is a relatively simple protocol providing little opportunity for conflicting interpretations. A few relevant agreements follow:

a) The optional elements of procedure for use of CLTS over CONS (i.e., clause 6.3 of ISO 8602) will not be supported;

b) A Unitdata TPDU that is received that contains a protocol error or an unknown destination TSAP ID shall be discarded.

6.2.2 Connectionless Transport Service Access Points or Selectors

The TSAP selector field in the UD TPDU shall be encoded as a variable length field and will be interpreted as an octet string. The length of the string cannot exceed 32 octets.

7 Transport Protocol Identification

The absence of Call User Data (CUD) in an X.25/ISO 8208 Call Request/Incoming Call packet indicates the operation of ISO 8073/CCITT X.224.

Protocol Identification TPDU values applicable to these agreements are given in table 1. These TPDUs, when used, are conveyed as N-connect user data.

Table 1 - Protocol Identification TPDU Values

TPDU Value	Protocol
03 01 01 00 * (see note 1)	ISO 8073/Add. 1
03 01 02 00 ** (see note 2)	ISO 8602

NOTES

1 Corresponds to an ISO 8073/Add. 1 UN-TPDU and a X.224 Annex B PI-TPDU.

2 Corresponds to an ISO 8073/Add. 1 UN-TPDU.

The following agreements apply:

a) Any additional TPDU, which follows (by concatenation) a Protocol Identification TPDU shall be ignored if ISO 8073/Add. 1 is not supported;

b) When using ISO 8208, usage of a Protocol Identification TPDU not corresponding to those listed in table 1 is outside the scope of these agreements.

8 Security

8.1 ISO/IEC 10736 Transport Layer Security Protocol (TLSP)

ISO/IEC 10736 describes both a connection oriented and connectionless security protocol that can be used in conjunction with OSI Transport Layer Protocols (ISO/IEC 8073 and ISO/IEC 8602). Before secure communication can be accomplished, a security association (in band or out of band) shall have been established with agreement on all attributes associated with this association.

Managed objects are not yet specified by this standard and therefore the security domain/administrative authority shall determine the procedures and policies that govern this information with other security information.

All mandatory functions are supported by these implementation agreements.

8.2 Services

If access control service is selected and the labels mechanism is used, then integrity shall also be selected.

The Transport (Class 4) initiator shall propose the non-use of checksums if TLSP is also invoked with connection integrity selected (as this would be redundant functionality). The integrity mechanism selected shall be one of the recommended algorithms (a signed MD5 or SHA for public key systems or DES MAC for secret key systems to name just a few) in part 12 (OS Security) of these agreements or a private algorithm that both communicating parties have agreed to use.

8.3 Mechanisms

To optimize efficiency and assist in the interoperability of secure implementations, it is useful to specify which mechanisms and algorithms apply. This specification shall allow implementations to know the exact encapsulation format used including what fields are required, their length, and order. A set of applicable profiles (mechanisms and algorithms) shall be specified within the Implementation Agreements to insure this efficient interoperability.

8.4 Protocol Constraints

Although the standard has the option of all type-length-value (tlv) fields being in any order, for efficiency, the encapsulation format depicted in the standard shall be used. If the tlv fields are not in order, undefined (type field has not been allocated a value in the TLSP Standard), or the SE TPDU fails one of the TLSP Security checks, the secure encapsulated PDU should be discarded. The reporting of this situation is a local matter. If shared knowledge of this event is required, a possible technique would be to use the system management to report the error.

The Security Association-Identification field should be no more than 20 octets.

8.5 Functional Security Sequence Ordering

If Access control is implemented using labels, the label function is first applied followed by the integrity function. If confidentiality has also been selected, then that function is perfomed after the integrity function.

If integrity and confidentiality have been selected, the integrity function is performed before the confidentiality function.