INTERNATIONAL STANDARD

ISO/IEC 14476-1

First edition
2002-06-15

Information technology — Enhanced communications transport protocol: Specification of simplex multicast transport

Technologies de l'information — Protocole de transport de communication amélioré: Spécifications pour le transport «simplex multicast»

Reference number
ISO/IEC 14476-1:2002(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.ch
Web www.iso.ch

Printed in Switzerland

CONTENTS

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 3.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this part of ISO/IEC 14476 may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

ISO/IEC 14476-1 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 6, *Telecommunications and information exchange between systems*, in collaboration with ITU-T. The identical text is published as ITU-T Rec. X.606.

ISO/IEC 14476 consists of the following parts, under the general title *Information technology — Enhanced communications transport protocol*:

— *Part 1: Specification of simplex multicast transport*

— *Part 2: Specification of QoS management for simplex multicast transport*

— *Part 3: Specification of duplex multicast transport*

— *Part 4: Specification of QoS management for duplex multicast transport*

— *Part 5: Specification of n-plex multicast transport*

— *Part 6: Specification of QoS management for n-plex multicast transport*

Annexes A and B of this part of ISO/IEC 14476 are for information only.

Introduction

This Recommendation | International Standard specifies the Enhanced Communications Transport Protocol (ECTP), which is a transport protocol designed to support Internet multicast applications running over multicast-capable networks. ECTP operates over IPv4/IPv6 networks that have the IP multicast forwarding capability with the help of IGMP and IP multicast routing protocols, as shown in Figure 1. ECTP could possibly be provisioned over UDP.

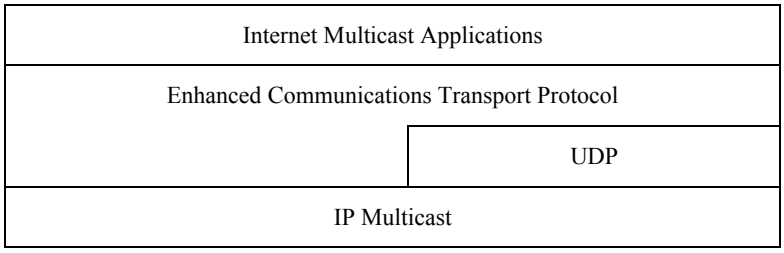

Figure 1 – ECTP Model

ECTP is designed to support tightly controlled multicast connections in simplex, duplex and N-plex applications. This Part of ECTP specifies the protocol mechanisms for reliability control in the simplex case. ECTP also provides QoS management functions for stable management of the QoS of the connection users. Such QoS management functionality can be achieved with QoS negotiation, monitoring and maintenance operations. The protocol procedures for QoS management of the simplex case will be defined in the simplex QoS management specification (ITU-T Rec. X.606.1 | ISO/IEC 14476-2), which forms an integral part of this Recommendation | International Standard. Further specifications will define control procedures and associated QoS management functions for the duplex case (ITU-T Rec. X.607 | ISO/IEC 14476-3 and ITU-T Rec. X.607.1 | ISO/IEC 14476-4) and for the N-plex case (ITU-T Rec. X.608 | ISO/IEC 14476-5 and ITU-T Rec. X.608.1 | ISO/IEC 14476-6).

In ECTP, all prospective members are enrolled into a multicast group, before a connection or session is created. Those members define an enrolled group. Each receiver in the enrolled group is referred to as an enrolled receiver. In the enrolment process, each member will be authenticated. The group information, including group key and IP multicast addresses and port numbers, will be distributed to the enrolled members during the enrolment process. An ECTP connection is created for these enrolled group members.

ECTP is targeted for tightly controlled multicast services. The sender is at the heart of multicast group communications. A single sender in the simplex multicast connection is assigned the role of the connection owner, designated as top owner (TO) in this Specification. The connection owner is responsible for overall connection management by governing connection creation and termination, connection pause and resumption, and join and leave operations.

The sender triggers the connection creation process. Some or all of the enrolled receivers will participate in the connection, becoming designated "active receivers". Any enrolled receiver that is not active may participate in the connection as a late-joiner. An active receiver can leave the connection. After the connection is created, the sender begins to transmit multicast data. If network problems (such as severe congestion) are indicated by the ECTP QoS management functions (defined in ECTP part 2), the sender suspends multicast data transmission temporarily, invoking the connection pause operation. After a pre-specified time, the sender resumes data transmission. If all of the multicast data have been transmitted, the sender terminates the connection.

ECTP provides the reliability control mechanisms for multicast data transport. ECTP mechanisms are designed to keep congruency with those being proposed in the IETF. To address reliability control with scalability, the IETF has proposed three approaches: Tree based ACK (TRACK), Forward Error Correction (FEC), and Negative ACK Oriented Reliable Multicast (NORM). Each approach has its own pros and cons, and each service provider may take a different approach toward implementing reliability control. ECTP adopts the TRACK approach, because it is more similar to the existing TCP mechanisms and more adaptive to the ECTP framework.

For tree-based reliability control, a hierarchical tree is configured during connection creation. The sender is the root of the control tree. The control tree can define a parent-child relationship between any pair of tree nodes. This tree-based structure can result in local owners (parents) occurring at lower levels in the tree hierarchy as the control structure extends. Each local owner created becomes the root of its own local control tree. The connection owner will then be the root of the overall control tree. Error control is performed for each local group defined by a control tree. Each parent retransmits lost data, in response to retransmission requests from its children.

INTERNATIONAL STANDARD ISO/IEC 14476-1
ITU-T RECOMMENDATION X.606

Information technology – Enhanced communications transport protocol: Specification of simplex multicast transport

1 Scope

This Recommendation | International Standard specifies the Enhanced Communications Transport Protocol (ECTP), which is a transport protocol designed to support Internet multicast applications over multicast-capable IP networks.

This Recommendation | International Standard specifies the ECTP for the simplex multicast transport connection that consists of one sender and many receivers. This Recommendation | International Standard specifies the protocol procedures for the following protocol operations:

a) connection creation with tree creation;

b) multicast data transmission;

c) tree-based reliability control with error detection, retransmission request and retransmission;

d) late join and leave;

e) tree membership maintenance; and

f) connection termination.

2 Normative references

The following Recommendations and International Standards contain provisions which, through reference in this text, constitute provisions of this Recommendation | International Standard. At the time of publication, the editions indicated were valid. All Recommendations and Standards are subject to revision, and parties to agreements based on this Recommendation | International Standard are encouraged to investigate the possibility of applying the most recent edition of the Recommendations and Standards listed below. Members of IEC and ISO maintain registers of currently valid International Standards. The Telecommunication Standardization Bureau of the ITU maintains a list of the currently valid ITU-T Recommendations.

– ITU-T Recommendation X.601 (2000), *Multi-peer communications framework*.

– ITU-T Recommendation X.605 (1998) | ISO/IEC 13252:1999, *Information technology – Enhanced Communications Transport Service Definition*.

3 Definitions

3.1 Terms defined in ITU-T Rec. X.601

This Recommendation | International Standard is based on the definitions of the multicast groups developed in Multi-Peer Communications Framework (ITU-T Rec. X.601).

a) Enrolled group; and

b) Active group.

3.2 Terms defined in ITU-T Rec. X.605 | ISO/IEC 13252

This Recommendation | International Standard is based on the concepts developed in Enhanced Communications Transport Service (ITU-T Rec. X.605 | ISO/IEC 13252).

a) Transport connection; and

b) Simplex.

3.3 Terms defined in this Recommendation | International Standard

For the purposes of this Recommendation | International Standard, the following definitions apply:

3.3.1 application: Represents an Internet multicast application in this Specification. It corresponds to a transport service user in the OSI mode. It exchanges transport service primitives with the corresponding transport protocol entity. In the Internet, it communicates with the transport protocol entity via a socket interface.

3.3.2 packet: Represents a unit of transport data, which is equivalent to a segment in TCP/IP and a transport protocol data unit (TPDU) in OSI model. A transport entity communicates with another transport entity by transmitting packets. A transport protocol entity creates a packet, which is encapsulated into an IP datagram and then delivered to the destination entity over networks.

3.3.3 sender: Represents a transport protocol entity that sends the multicast data to the receivers.

3.3.4 receiver: Represents a transport protocol entity that receives the multicast data.

3.3.5 tree: Is a hierarchical logical tree employed for providing scalable reliability control. A tree defines a parent-child relationship between a pair of tree nodes. Sender and receivers are organized into a tree. In the tree hierarchy, a tree node is designated as TO (Top Owner), LO (Local Owner) or LE (Leaf Entity). TO is a single ECTP sender. All the receivers are designated as LOs or LEs.

3.3.6 TO (top owner): Is a single sender in the ECTP simplex multicast connection. TO is the root of the tree and manages the overall protocol operations for the connection.

3.3.7 LO (local owner): Is a receiver that manages a local group. An LO is responsible for the overall protocol operations for its local group defined by the control tree. For error recovery, it retransmits the multicast data that have been lost by its children. For flow and congestion control, it aggregates the control information for all of its children and then delivers the aggregated information toward TO. In terms of the reliability control operations, TO is also an LO.

3.3.8 LE (leaf entity): Is a receiver that has not been designated as an LO. An LE cannot have any children. It is located as a leaf node on the tree.

3.3.9 local group: Consists of a parent and its children in the tree hierarchy.

3.3.10 parent: Is a parent node for a local group. TO or an LO can be a parent.

3.3.11 child: Is a child node for a local group. An LO or LE can be a child.

4 Abbreviations

For the purposes of this Recommendation | International Standard, the following abbreviations apply:

4.1 Packet types

ACK	Acknowledgment
CC	Connection Creation Confirm
CR	Connection Creation Request
CT	Connection Termination
DT	Data
HB	Heartbeat
JC	Late Join Confirm
JR	Late Join Request
LR	Leave Request
ND	Null Data
RD	Retransmission Data
TC	Tree Join Confirm
TJ	Tree Join Request

4.2 Miscellaneous

ECTP	Enhanced Communications Transport Protocol
ECTS	Enhanced Communications Transport Service
IETF	Internet Engineering Task Force
IGMP	Internet Group Management Protocol
IP	Internet protocol
QoS	Quality of Service
RFC	Request for Comments
RMT	Reliable Multicast Transport
SAP	Session Announcement Protocol
SDP	Session Description Protocol
TCP	Transmission Control Protocol
UDP	User Datagram Protocol

5 Conventions

In this Recommendation | International Standard, the key words "MUST", "REQUIRED", "SHALL", "MUST NOT", "SHALL NOT", "SHOULD", "SHOULD NOT", "MAY", and "OPTIONAL" are to be interpreted as described in IETF RFC 2119, and indicate requirement levels for compliant ECTP implementations. Those key words are case-sensitive.

6 Overview

The ECTP is a transport protocol designed to support Internet multicast applications. ECTP operates over IPv4/IPv6 networks that have IP multicast forwarding capability.

This Specification describes the ECTP protocol for the simplex multicast transport connection that consists of one sender and many receivers. ECTP supports the connection management functions, which are based on ITU-T Rec. X.605 | ISO/IEC 13252. The connection management functions include connection creation and termination, connection pause and resumption, and late join and leave. For reliable delivery of multicast data, ECTP also provides the protocol mechanisms for error, flow and congestion controls. To allow scalability to large-scale multicast groups, tree-based reliability control mechanisms are employed which are congruent with those being proposed in the IETF RMT WG.

Figure 2 shows an overview of the ECTP operations.

As shown in the figure, the QoS management operations such as QoS negotiation, monitoring and maintenance will be specified in ITU-T Rec. X.606.1 | ISO/IEC 14476-2. In particular, QoS maintenance includes the operations for connection pause and resume, and the flow and congestion controls.

Before an ECTP transport connection is created, the prospective receivers are enrolled into the multicast group. Such a group is called an enrolled group (see 8.1). During enrolment, authentication processes may be performed together with group key distribution. The IP multicast addresses and port numbers must be announced to the receivers. These enrolment operations may rely on the well-known SAP/SDP, HTTP (Web Page announcement) and SMTP (E-mail) protocols. The specific enrolment mechanisms are outside the scope of this Specification.

An enrolled receiver will be connected to the multicast-capable network with the help of the IGMP and IP multicast routing protocols. Those IGMP and multicast routing protocols will refer to the announced multicast addresses. An ECTP transport connection is created for the enrolled receivers.

ECTP is targeted to support tightly controlled multicast connections. The ECTP sender is at the heart of the multicast group communication. The sender, designated as connection owner (TO), is responsible for the overall management of the connection by governing connection creation and termination, connection pause and resumption, and the late join and leave operations.

The ECTP sender triggers the connection creation process by sending a connection creation message. Some or all of the enrolled receivers may respond with confirmation messages to the sender. The connection creation is completed when the sender receives the confirmation messages from all of the active receivers, or when a pre-specified timer expires (see 8.2).

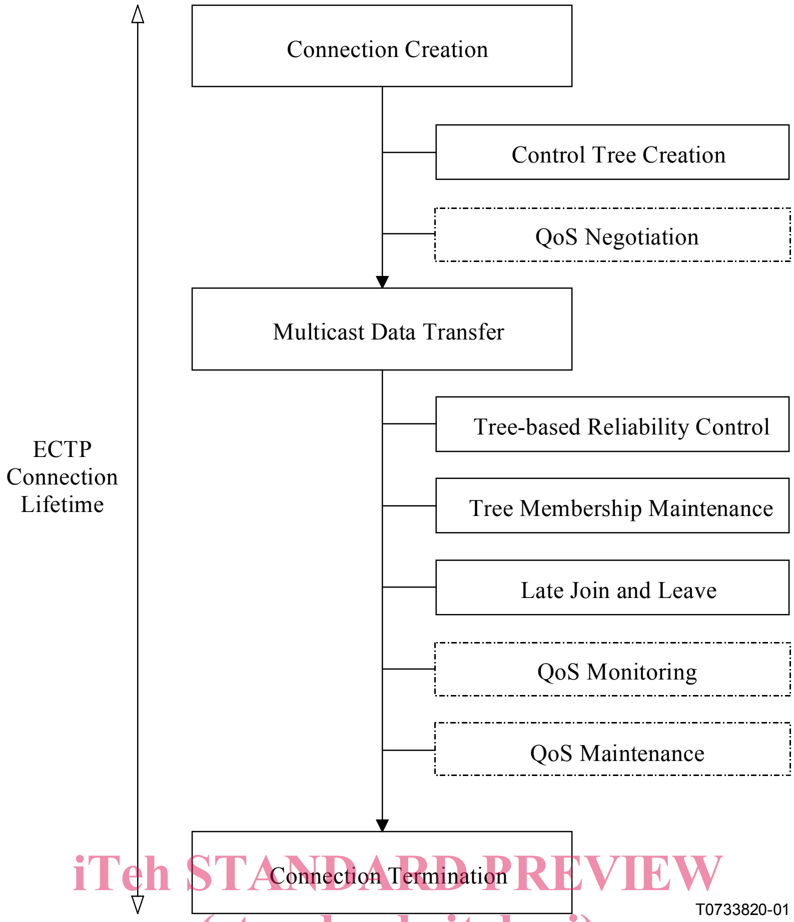

Figure 2 – ECTP Protocol Operations

Throughout the connection creation, some or all of the enrolled group receivers will join the connection. The receivers that have joined the connection are called active receivers. An enrolled receiver that is not active can participate in the connection as a late-joiner (see 8.6). The late-joiner sends a join request to the sender. In response to the join request, the sender transmits a join confirm message, which indicates whether the join request is accepted or not. An active receiver can leave the connection by sending a leaving request to the sender. A trouble-making receiver, who cannot keep pace with the current data transmission rate, may be ejected (see 8.7).

After a connection is created, the sender begins to transmit multicast data (see 8.3). For data transmission, an application data stream is sequentially segmented and transmitted by means of data packets to the receivers. The receivers will deliver the received data packets to the applications in the order they were transmitted by the sender.

To make the protocol scalable to large multicast groups, ECTP employs the tree-based reliability control mechanisms. A hierarchical tree is configured during connection creation. A control tree defines a parent-child relationship between any pair of tree nodes. The sender is the root of the control tree. In the tree hierarchy, a set of local groups are defined. A local group consists of a parent and zero or more children. The error, flow and congestion controls are performed for each local group defined by the control tree.

Figure 3 illustrates a control tree hierarchy for reliability control, in which a parent-child relationship is configured between a sender (S) and a receiver (R), or between a parent receiver (R) and its child receiver (R).

ECTP specifies the protocol procedures for tree creation. In the tree creation, a control tree is gradually expanded from the sender to the receivers (see 8.2.2). This is called a top-down configuration. On the other hand, the IETF RMT WG has proposed a bottom-up approach, where the receivers initiate a tree configuration (see Annex B). Those schemes may be incorporated into the ECTP as candidate tree creation options in the future.

Tree-membership is maintained during the connection. A late-joiner is allowed to join the control tree. The late-joiner listens to the heartbeat messages from one or more on-tree parents, and then joins the best parent. When a child leaves the connection, the parent removes the departing child from the children-list. Node failures are detected by using periodic control messages such as null data, heartbeat and acknowledgement. The sender transmits periodic null data messages to indicate that it is alive, even if it has no data to transmit. Each parent periodically sends heartbeat messages to its children. On the other hand, each child transmits periodic acknowledgement messages to its parent (see 8.8).

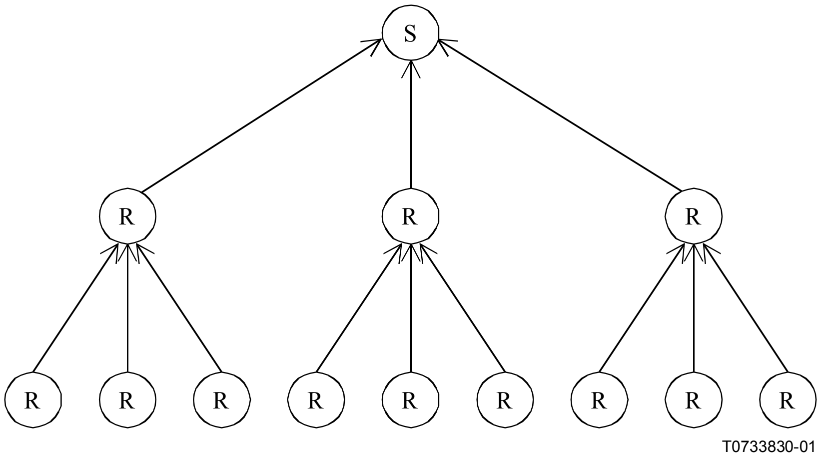

Figure 3 – Control Tree Hierarchy for Reliability Control

In ECTP, error control is performed for each local group defined by a control tree (see 8.4). If a child detects a data loss, it sends a retransmission request to its parent via ACK packets.

An ACK message contains the information that identifies the data packets which have been successfully received. Each child can send an ACK message to its parent using one of two ACK generation rules: ACK number and ACK timer. If data traffic is high, an ACK is generated for the ACK number of data packets. If the traffic is low, an ACK message will be transmitted after the ACK timer expires.

After retransmission of data, the parent activates a retransmission back-off timer. During the time interval, retransmission request(s) for the same data will be ignored. Each parent can remove the data out of its buffer memory, if those have been acknowledged by all of its children.

The flow and congestion control information is delivered from the receivers to the sender, along the control tree. The detailed description of flow and congestion control will be given in ITU-T Rec. X.606.1 | ISO/IEC 14476-2, the QoS management specification for the simplex multicast transport. Based on the monitored flow and congestion control information, the sender will adjust the transmission rate.

During the data transmission, if network problems (for example, severe congestion) are indicated by the QoS management functions specified in ITU-T Rec. X.606.1 | ISO/IEC 14476-2, the sender suspends the multicast data transmission temporarily. In this period, no new data is delivered, while the sender transmits periodic null data messages to indicate that the sender is alive. After a pre-specified time has elapsed, the sender resumes the multicast data transmission (see 8.5).

The sender terminates the connection by sending a termination message to all the receivers, after all the multicast data are transmitted. The connection may also terminate due to a fatal protocol error such as a connection failure (see 8.9)

7 Protocol components

7.1 Nodes

ECTP protocol mechanisms are based on a logical control tree, which defines a parent-child relationship between any pair of tree nodes. Each node on the tree is classified into one of three node types: TO (top owner), LO (local owner) and LE (leaf entity).

a) Top Owner (TO)

TO is the root of the control tree and also a single sender in the simplex multicast connection. TO manages the overall connection management functions including the connection creation and termination. In the connection creation phase, a control tree is configured by interactions between the sender and receivers. After the connection is created, TO sends multicast data to the receivers. TO can temporarily suspend and resume the connection. TO can admit or reject the group members who want to join the existing connection. After all the data is transmitted, TO terminates the multicast transport connection.

b) Local Owner (LO)

In the ECTP connection, some of the receivers may be designated as LOs. Each LO has its children that consist of other LOs and/or LEs. LOs are thus located as interior nodes on the tree. Each LO retransmits the multicast data that have been lost by its children. It also aggregates the information on the flow and congestion control from its children, and forwards the aggregated information toward TO. TO is also an LO in terms of the reliability control operations.

c) Leaf Entity (LE)

A receiver, which has not been designated as LO, is referred to as an LE. An LE cannot have any children. It is thus located as a leaf node on the control tree.

TO is a single sender. LOs and LEs are receivers. In the tree hierarchy, a local group consists of a parent and its children. TO or an LO can be a parent, and an LO or LE can be a child.

In the tree hierarchy, an LO retransmits lost multicast data to its children (error recovery) and forwards the flow and congestion control information to TO. Moreover, each LO has authority to eject a trouble-making child to maintain the stability of the connection. It is thus expected that LOs are given more processing power and responsibility than LEs.

In ECTP, it is presumed that some of receivers have been designated as LOs before the connection is created. This Specification does not consider the selection of LOs among flat receivers during the connection. That is, before the connection creation (or in the enrolment phase), each receiver MUST know whether it is an LO or LE. In a very small-sized group or asynchronous networks such as satellite or mobile networks, no LO may be designated. In those environments, all the receivers will be LEs.

An LO may be an end host or a dedicated server. In privately controlled networks, it is probable that dedicated servers function as LOs. In public networks, end hosts may be employed as LOs. In either case, an LO is a receiver and performs the reliability control operations for its local group as a parent.

7.2 Control tree

After a connection is created, TO transmits data to all the receivers by multicast. Each child sends status information on data reception to its parent. The information will thus be delivered to TO along the control tree. The multicast data streams flow from TO to LOs and LEs in the downward direction, while the control information is transferred from LEs to TO via LOs in the upward direction along the control tree.

Figure 4 shows a general structure of an ECTP control tree.

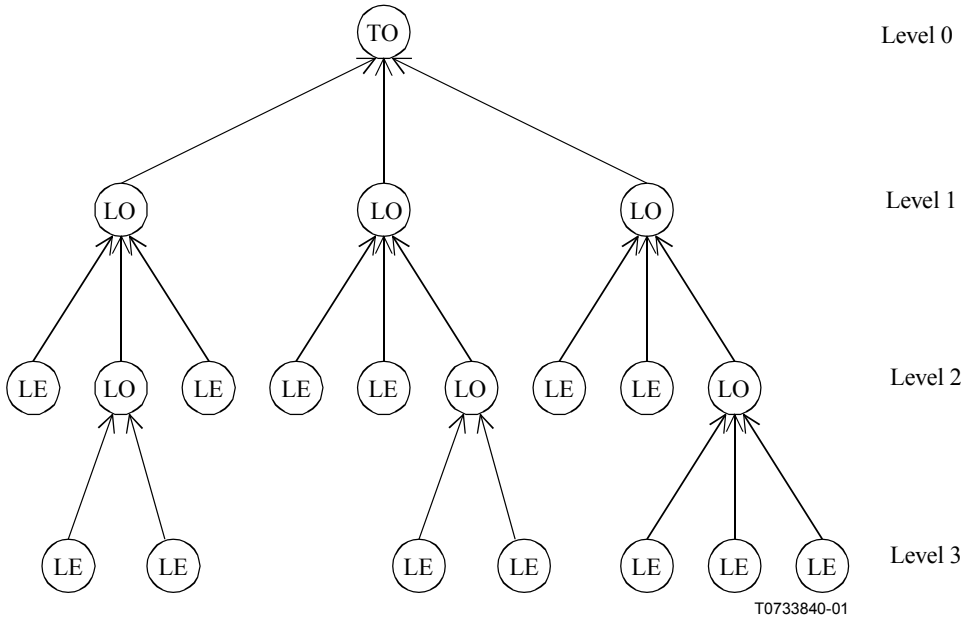

Figure 4 – An ECTP Control Tree