INTERNATIONAL STANDARD

ISO/IEC 14476-5

First edition
2008-04-15

Information technology — Enhanced communications transport protocol: Specification of N-plex multicast transport

Technologies de l'information — Protocole de transport de communications amélioré: Spécification pour le transport N-plex en multidiffusion

Reference number
ISO/IEC 14476-5:2008(E)

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 2.

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 document 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 14476-5 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.608 (02/2007).

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 5: Specification of N-plex multicast transport*

Introduction

ECTP is designed to support tightly controlled multicast connections in simplex, duplex and N-plex applications. This part of ECTP (Part 5: ITU-T Rec. X.608 | ISO/IEC 14476-5) specifies the protocol mechanisms for the N-plex multicast data transport.

In the N-plex multicast connection, the participants include one TC-Owner and many TS-users. TC-Owner will be designated among the TS-users before the connection begins. TC-Owner is at the heart of multicast group communications. It is responsible for overall connection management by governing the connection creation and termination, multicast data transport, and the late join and leave operations.

In the N-plex multicast connection, the multicast data transmissions are allowed by TS-users as well as TC-Owner. Each TS-user is allowed to send multicast data to the group only if it gets a token from the TC-Owner. That is, the multicast data transmissions of TS-users are controlled by TC-Owner.

The N-plex multicast connection specified in this Recommendation | International Standard targets the many-to-many multicast applications in which many participants (TS-users) may want to transmit the multicast data to all the other TS-users. Typical examples of such applications include 'teleconferencing' and 'multi-users on-line game', etc. In the teleconferencing application, TC-Owner may act as a 'conferencing server', and all the other participants (TS-users) may send multicast data, such as voice, text and image, to the other participants.

INTERNATIONAL STANDARD
ITU-T RECOMMENDATION

Information technology – Enhanced communications transport protocol: Specification of N-plex multicast transport

1 Scope

This Recommendation | International Standard specifies the N-plex multicast transport connection in which all participants are TS-users and one of them is TC-Owner. The N-plex multicast transport connection allows TS-users to send the multicast data to all the group members. It is noted that a TS-user is allowed to send the multicast data to the group, only if it gets a token from TC-Owner.

This Specification describes the protocol for supporting the N-plex multicast transport, which includes the connection management (establishment, termination, user join and leave) and the reliability control mechanisms for the multicast data transport.

2 References

2.1 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 currently valid ITU-T Recommendations.

– ITU-T Recommendation X.601 (2000), *Multi-peer communications framework*.
– ITU-T Recommendation X.602 (2004) | ISO/IEC 16513:2005, *Information technology – Group management protocol*.
– ITU-T Recommendation X.605 (1998) | ISO/IEC 13252:1999, *Information technology – Enhanced communications transport service definition*.
– ITU-T Recommendation X.606 (2001) | ISO/IEC 14476-1:2002, *Information technology – Enhanced communications transport protocol: Specification of simplex multicast transport*.
– ITU-T Recommendation X.606.1 (2003) | ISO/IEC 14476-2:2003, *Information technology – Enhanced communications transport protocol: Specification of QoS management for simplex multicast transport*.
– ITU-T Recommendation X.607 (2007) | ISO/IEC 14476-3:2008, *Information technology – Enhanced communications transport protocol: Specification of duplex multicast transport*.

2.2 Informative references

– ITU-T Recommendation X.607.1 (draft) | ISO/IEC 14476-4, *Information technology – Enhanced communications transport protocol: Specification of QoS management for duplex multicast transport*.

3 Definitions

This Recommendation | International Standard is based on the following definitions, which were specified in Enhanced Communications Transport Service (ITU-T Rec. X.605 | ISO/IEC 13252).

a) Transport connection: Simplex, Duplex and N-plex.

This Recommendation | International Standard redefines the following definitions specified in Enhanced Communications Transport Service (ITU-T Rec. X.605 | ISO/IEC 13252).

a) **TC-owner (TCN)**: TCN manages overall operations of an N-plex multicast connection.

b) **transport service user (TS-user)**: TS-users can send and receive multicast data in the N-plex multicast connection.

c) **sending TS-user (SU)**: A TS-user who gets a token from TCN. Only the SU is allowed to send multicast data to the group. In other words, before sending multicast data, each TS-user must request a token to TCN.

This Recommendation | International Standard redefines the following terminologies specified in Enhanced Communications Transport Protocol: part 1 (ITU-T Rec. X.606 | ISO/IEC 14476-1) to accommodate to N-plex multicast connection.

a) **local group**: A set of nodes in vicinity which has network-layer correlation in terms of packet loss and delay.

b) **local owner (LO)**: LO is a representative node of a local group and designated statically. It is responsible for maintaining an intra-group tree of the group and control trees for all SUs in its local group. Each LO is also connected to the other LOs along inter-group trees. It also generates test traffic periodically for logical tree adaptation.

c) **multicast data channel**: TCN or SU can send multicast data to all the other group members over IP multicast address.

This Recommendation | International Standard newly defines the following terminologies:

a) **logical tree**: A tree that spans all TS-users and one or more control trees are derived from it.

b) **inter-group tree**: A per-source logical tree of the LOs.

c) **intra-group tree**: A shared logical tree of each local group.

d) **control tree**: A tree along which control packets for error control are traversed.

e) **token**: It represents the right for a TS-user to transmit multicast data. The TS-user who has a token is called SU. The tokens are managed by TCN.

4 Abbreviations

The following acronyms for ECTP protocols are used in this Recommendation | International Standard:

ECTP-1	ECTP part 1 (ITU-T Rec. X.606 | ISO/IEC 14476-1)
ECTP-2	ECTP part 2 (ITU-T Rec. X.606.1 | ISO/IEC 14476-2)
ECTP-3	ECTP part 3 (ITU-T Rec. X.607 | ISO/IEC 14476-3)
ECTP-4	ECTP part 4 (ITU-T Rec. X.607.1 | ISO/IEC 14476-4)
ECTP-5	ECTP part 5 (ITU-T Rec. X.608 | ISO/IEC 14476-5)
ECTP-6	ECTP part 6 (ITU-T Rec. X.608.1 | ISO/IEC 14476-6)

The following acronyms for ECTP-5 packets are used in this Recommendation | International Standard:

ACK	Acknowledgment
CC	Connection Creation Confirm
CCC	Control Tree Change Confirm
CCR	Control Tree Change Request
CR	Connection Creation Request
CT	Connection Termination Request
DT	Data
JC	Late Join Confirm
JR	Late Join Request
LR	User Leave Request
NACK	Negative Acknowledgement
PB	Probe

PBACK	Probe Acknowledgment
RD	Retransmission Data
TC	Tree Join Confirm
TCC	Tree Change Confirm
TCR	Tree Change Request
TDC	Tree Delegation Confirm
TDR	Tree Delegation Request
TGC	Token Get Confirm
TGR	Token Get Request
TJ	Tree Join Request
TLC	Tree Leave Confirm
TLR	Tree Leave Request
TNC	Tree Change Notification Confirm
TNR	Tree Change Notification Request
TRC	Token Return Confirm
TRR	Token Return Request
TSR	Token Status Report
TSRR	Token Status Report Request

5 Conventions

In this Recommendation | International Standard, packets of ECTP-5 are represented as words with all capital characters (e.g., CR for Connection Creation Request packet), and system parameters are represented as words with all italic capital characters (e.g., *TD_PACKET_INT* for test data packet interval).

6 Overview

The Enhanced Communications Transport Protocol (ECTP) is a transport protocol designed to support Internet multicast applications. 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.

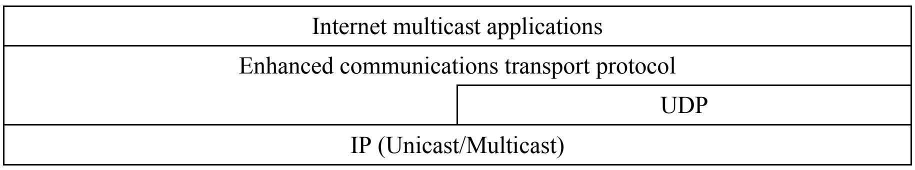

Figure 1 – ECTP model

This Recommendation | International Standard describes the protocol specification of the ECTP part 5 (ECTP-5) for the N-plex multicast connection. The N-plex multicast connection is used for supporting multicast data transport between the participants (TS-users). In the N-plex multicast connection, TS-users can send multicast data packets to the group over multicast data channel. A TS-user who is sending multicast data in the N-plex multicast connection is called SU (Sending TS-user). Any SU must have a token for multicast data transmission. In other words, the TS-user who gets a token from TCN is called an SU.

Figure 2 illustrates the multicast data transport channel in the N-plex multicast connection. As shown in the figure, TCN and SU can transmit multicast data to the other session members over IP multicast (group) address.

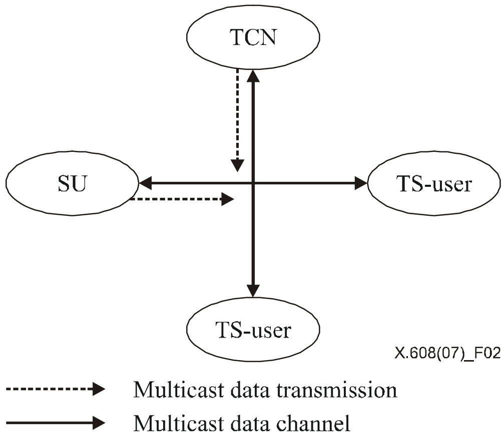

Figure 2 – Multicast data transport in N-plex multicast connection

To establish an N-plex multicast connection, TCN should be activated to manage session information and tokens.

If the TCN is informed of a participant list by out-of-band signalling prior to starting the connection, it should start a connection creation phase by transmitting a CR packet to the group. The CR packet contains the connection information including general characteristics of the connection. Each TS-user who is contained in the participant list should respond to the TCN with a CC packet. The connection creation operation will be completed when the TCN receives CC packets from all of the TS-users in the participant list, and the data transmission phase starts. If there is no predetermined participant prior to starting the session, TCN starts the data transmission phase without connection creation operations.

In the middle of data transmission phase, the prospective TS-users should join the connection as late-joiners. The late-joining TS-user participates in the connection by sending a Late Join Request (JR) message to TCN. In response to the JR message, TCN sends a Late Join Confirm (JC) message to the TS-user. After a TS-user is confirmed to join the session, it should join a logical tree by exchanging the Tree Join Request (TJ) and Tree Join Confirm (TC) messages with an appropriate LO. The logical tree is used for error control.

An N-plex multicast connection builds a two-layer logical tree, which consists of intra-group shared trees and inter-group per-source trees, as shown in Figure 3.

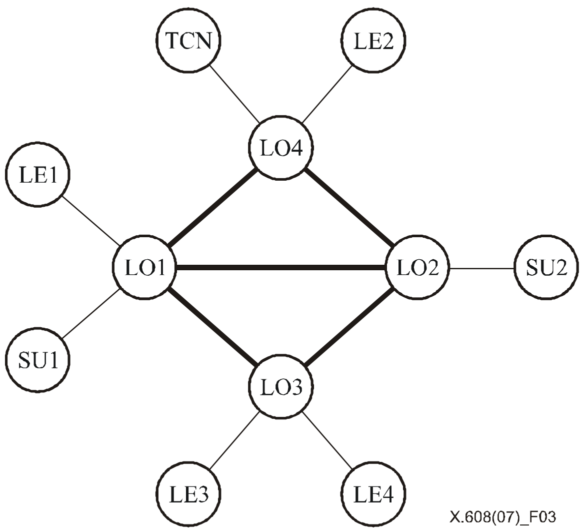

Figure 3 – A two-layer tree for N-plex multicast connection

At the bottom layer, each LE in a local group joins an LO-rooted shared logical tree (*intra-group tree*). At the top layer, LOs constitute logical trees for each SU (*inter-group tree*). It is noted that the control tree for each SU is derived by grafting these inter-group and intra-group trees.

In N-plex multicast connection, intra-group trees can be organized differently depending on the roles of LEs, which is determined by TCO (Tree Configuration Option). One option of TCO is that all LEs are direct children of LOs and do not participate in the repair process and ACK aggregations. Thus, intra-group trees will be an LO-rooted one-level tree in this option. The other option is that LEs can be children of other LEs. In this option, LEs are responsible for reliability support of its children LEs and intra-group trees can be multi-level trees. For the sake of efficiency of reliability support, the multi-level intra-group trees should be as close as possible to underlying multicast routing tree. The N-plex multicast connection adopts logical tree adaptation mechanism for this option. For this, error bitmaps of TS-users are used. An error bitmap represents packet delivery status, which indicates the loss pattern of multicast packets. Each TS-user sends its error bitmap to its parent with respect to multicast data from the root node of its intra-group tree with periodic ACK messages. By comparing error bitmaps of itself and those of its children, a node decides whether each child is likely to be its actual child in the underlying multicast routing tree or not. If the child node is determined not to be its actual child, a node changes the logical tree by delegating the child node to its parent or one of its other children. After recursive tree changes, the intra-group tree will converge to a multi-level tree that is closer to underlying multicast routing tree.

The error control will be performed based on the ECTP control tree which is constructed as described above. If packet loss is detected by a gap of the packet sequence numbers, a child node sends a NACK packet to its parent immediately via unicast. The parent LO or SU, that receives the NACK packet, will retransmit data packet (RD) to the requestor via unicast. Each child generates an ACK packet every *ACK_GENERATION_NUM* (*AGN*) data packets.

In the multicast data transmission, TCN and SUs can begin the multicast data transmission to the group by using the IP multicast address and group port number. The TS-users will deliver the received DT packets to the upper-layer application in the order transmitted by SU or TCN.

For the multicast data transport, a TS-user in the connection may get a token from TCN by sending a TGR message. The TCN will then respond to the TS-user with the TGC message that contains a *Token ID*. Accordingly, the total number of tokens in the connection is controlled by TCN. The TS-user who has a token is called Sending TS-user (SU). In a certain case, the TCN can first request a TS-user to be an SU by sending the TGR message to the promising SU, which is called Token Give.

After completing the multicast data transmission, the SU will return the token to the TCN by sending a TRR message. TCN will respond to the SU with a TRC message. In a certain case, the TCN first requests the SU to return the token, which is called Token Withdrawal. TCN announces the overall status of the Token IDs valid in the connection to the group by sending the TSR packets.

TCN manages the connection for user leave. In the User Leave operation, a participating TS-user may leave the connection by sending an LR message to the parent. In a certain case, the TCN may enforce a specific TS-user to leave the connection by sending the LR message, which is called the troublemaker ejection.

TCN may terminate the N-plex connection by sending a CT message to the group.

7 Considerations

7.1 Participants

All participants to an N-plex multicast connection are TS-users and one of them is TCN (TC-Owner).

TCN (TC-Owner):

An N-plex multicast connection has a single TCN. The TCN is responsible for connection management including connection creation/termination, late join, connection maintenance, and token management.

For example, in the teleconferencing applications, the TCN may act as the 'conference server', which may be used for control of the conferencing without sending multicast data. In the example of 'multi-users on-line game' application, the TCN may act as the 'game-control server'.

TS-user (Transport Service User):

An N-plex multicast connection has one or more TS-users. Each of them sends and receives multicast data in the connection.

A TS-user can become LO or LE depending on its role.

LO (Local Owner):

LO is a representative node of a local group and designated statically. It is responsible for maintaining an intra-group tree of the group and control trees for all SUs in its local group. Each LO is also connected to the other LOs with inter-group trees. It also generates test traffic periodically for logical tree adaptation.

LE (Leaf Entity):

LE is a member of a local group whose representative is an LO. It should join an intra-group tree of the group and is responsible for exchanging control packets with its parent or child LEs along the control tree.

A TS-user can become SU when it obtains a token from TCN.

SU (Sending TS-user):

An SU is a TS-user who can send multicast data to the group. In an N-plex multicast connection, a TS-user becomes an SU when it has a token and it can thus transmit multicast data to the group.

7.2 Data channel and addressing

In N-plex multicast connection, SU or TCN can send multicast data packets to the session members as shown in Figure 4.

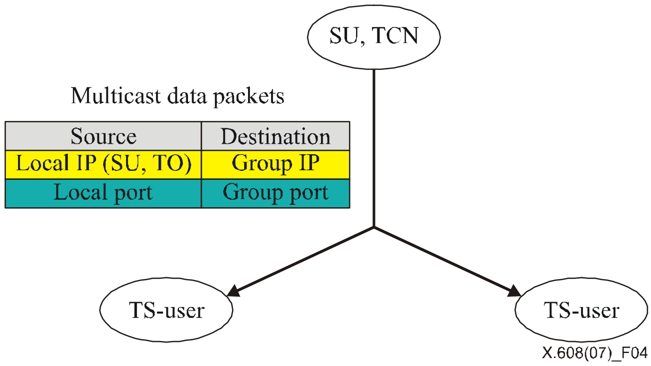

Figure 4 – Multicast data addressing in N-plex multicast connection

In an N-plex multicast connection, the multicast data packets (DTs) use the IP multicast address and the group port number as the destination address. The source address of the multicast data IP packets is the IP address of the sender of the packet. In contrast, the retransmission data packets (RDs) in response to the repair requests (NACKs) are delivered by LO or SU over control channel using unicast with the IP address of the repair requester as the destination.

7.3 Control channel and tree

In an N-plex multicast connection, there are control channels for error recovery. All members participate in one or more control trees. These trees are used as control channels for exchanging control messages among participants. A parent node acts as an agent who helps child nodes to recover from packet loss. It also aggregates the acknowledgement information from its descendant nodes. All control packets such as RD, NACK, and ACK are delivered via control channel using unicast.

An N-plex multicast connection is divided into multiple local groups of the participants for error control. Based on this group approach, participants construct and maintain control trees along which all the error control packets are exchanged. Control trees are constructed from two layers of logical trees. At the bottom layer, members in a local group join an LO-rooted shared logical tree (*intra-group tree*). At the top layer, the LOs of the groups constitute per-source logical trees (*inter-group tree*). That is, every LE joins a local group and it grafts onto LO-rooted logical tree of the group as a child (or descendents) of the LO. Every LO grafts onto logical tree of the LOs, of which root is the LO of the group that the SU belongs to. A control tree for each source is constructed by connecting the inter-group trees and intra-group trees.

Inter-group trees are generated and maintained exactly as required. Any inter-group tree rooted by an LO who has no more SU children should be removed.

Intra-group trees can evolve to multi-level trees that reflect actual multicast routing paths by logical tree adaptation mechanism, which is described in 7.5.

By traversing logical trees starting from an SU, we can get an SU-rooted tree that is the control tree for the SU.

If an SU belongs to another local group, a part of the control tree generated for current local group is the intra-group tree of the group. Control trees for the SUs belonging to the same local group are slightly different from the intra-group tree, because the intermediate nodes between LO and the SU have reversed parent-child relationships.

Assuming that the logical trees of the participants are constructed as in Figure 5, the control tree for source SU1 will be constructed as in Figure 6.

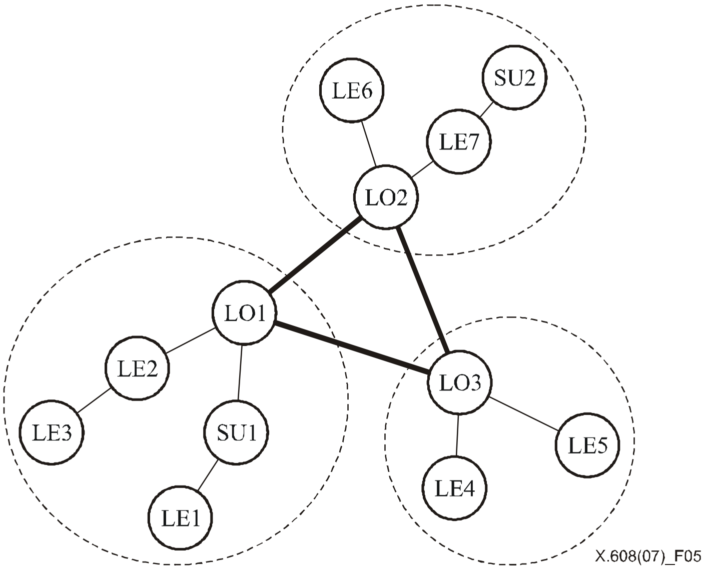

Figure 5 – Two layers of logical trees

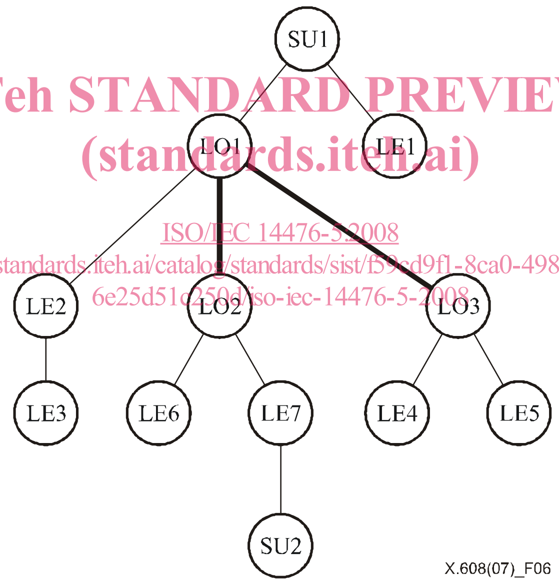

Figure 6 – Control tree when sender is SU1

In the same way, Figure 7 shows the control tree whose source is SU2.

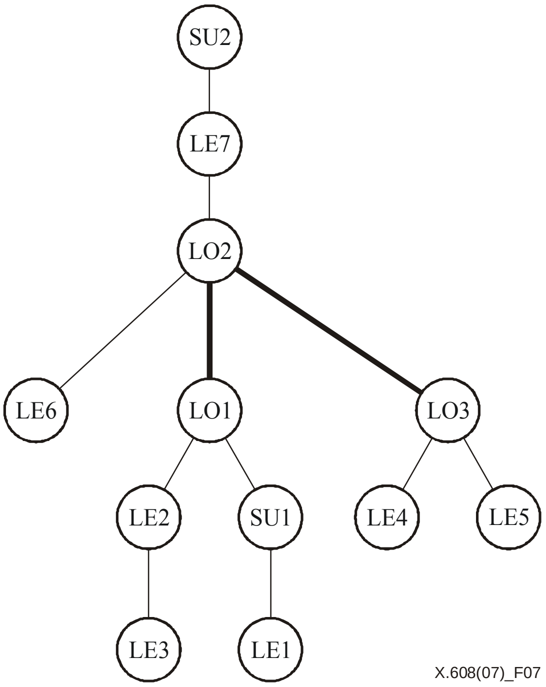

Figure 7 – Control tree when sender is SU2

Control trees are maintained by information obtained through the exchange of NACK and ACK packets. The parent can detect a child's failure by comparing its own LSN and received child's LSN. Children can detect their parent's failure if there is no response from parent for several consecutive repair requests. Then the child finds another parent by contacting LO.

7.4 Tokens

In N-plex multicast connection, a token represents the right for a TS-user to send multicast data to TCN. Before transmitting data, each TS-user must get a token from the TCN, as per the Token Control procedures of N-plex multicast connection. By this procedure, TCN can authorize a TS-user to become a sender so that TS-users can effectively filter out multicast data sent by unauthorized users. However, note that use of token does not provide any protection for IP multicast.

Each token is represented as a 1-byte non-negative integer. Such a token number (or Token ID) will be assigned by TCN when a TS-user requests a token in the connection. Token ID is ranged between 1 and 255. The Token ID of '0' is reserved for use of TCN. At the receiver side, the Token ID can be used to authorize who can send the multicast data.

7.5 Logical tree adaptation

In N-plex multicast connection, intra-group trees may evolve to multi-level trees close to underlying multicast routing trees when TCO is '10'. Loss pattern comparison is used to estimate the underlying multicast routing trees. The estimation is made feasible because if a loss occurs at a parent node of a multicast routing tree, all of its children also experience the same loss. The root of an intra-group tree does not change by the logical tree adaptation process.

For this, each receiver in a multicast session maintains packet delivery status called error bitmap, which indicates the loss pattern of multicast packets. An error bitmap consists of two parts: sequence number (*Ns*) and an actual bitmap (*B*). *Ns* is the sequence number of the first packet in a sequence of packets represented in the bitmap. One bit of *B* indicates the reception status of the corresponding packet; '1' means a successful packet delivery to a receiver without error recovery and '0' otherwise. The bitmap includes loss patterns of multicast packets. For example, if *Ns* = 5 and *B* = 11010, the receiver has successfully received packets 5, 6, and 8. Even packets 7 and 9 have been recovered via retransmissions, these bits are recorded as '0's. Each receiver periodically feeds the error bitmap information to its parent in the tree. If a logical tree matches a multicast routing tree and a bit in the error bitmap at a node is set to 1, the corresponding bit in the error bitmap of its parent is very likely to be '1'. Note that the error bitmap of a source always consists of all '1's.

Every node that has child nodes compares the error bitmaps of itself and its children to check whether the relationships between the node and its child nodes reflect the underlying multicast routing tree closely. A parent node can infer that some child nodes need to change their locations in the logical tree. The child node would be a child of other child node. Or it would not be a descendant of the parent node.

For instance, Figure 8 shows an example of a multicast routing tree and error bitmaps of each node.

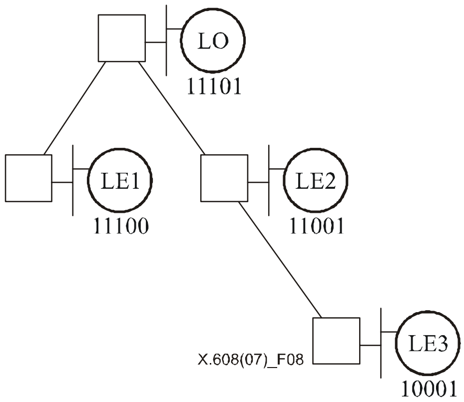

Figure 8 – An example of multicast routing tree

Figures 9 and 10 briefly describe how a logical tree close to a multicast routing tree is constructed by the error bitmap information.

First, it is assumed that a root LO is a strategically deployed server that is usually located nearby the egress point of Internet service provider. Boxes represent routers and numbers below nodes are error bitmaps of each node.

Figure 9 – Logical tree adaptation (LE2 adopts LE3)

The left side of Figure 9 shows an initial one-level intra-group tree of nodes in Figure 8. When LO becomes aware of error bitmaps of its children, it finds out that LE3 is likely to be a child or descendant of LE2 since a set of received packets of LE3 is a subset of that of LE2. Thus, it delegates LE3 to LE2 by sending a TDR (Tree Delegation Request) message. Receiving the message, LE2 judges that LE3 is its child and sends a TCR message (Tree Change Request) to LE3. Receiving the TCR message, LE3 joins as a child of LE2 by sending TJ to LE2 and leaves its previous parent, LO, by sending TLR. Note that a node receiving a TCR message is attached to the new parent before leaving its previous parent. The result is a tree on the right side of Figure 9.