INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

Y.1711

(02/2004)

SERIES Y: GLOBAL INFORMATION INFRASTRUCTURE, INTERNET PROTOCOL ASPECTS AND NEXT GENERATION NETWORKS

Internet protocol aspects – Operation, administration and maintenance

Operation & Maintenance mechanism for MPLS networks

ITU-T Recommendation Y.1711

ITU-T Y-SERIES RECOMMENDATIONS

GLOBAL INFORMATION INFRASTRUCTURE, INTERNET PROTOCOL ASPECTS AND NEXT GENERATION NETWORKS

GLOBAL INFORMATION INFRASTRUCTURE	
General	Y.100–Y.199
Services, applications and middleware	Y.200–Y.299
Network aspects	Y.300–Y.399
Interfaces and protocols	Y.400–Y.499
Numbering, addressing and naming	Y.500–Y.599
Operation, administration and maintenance	Y.600–Y.699
Security	Y.700–Y.799
Performances	Y.800–Y.899
INTERNET PROTOCOL ASPECTS	
General	Y.1000–Y.1099
Services and applications	Y.1100–Y.1199
Architecture, access, network capabilities and resource management	Y.1200–Y.1299
Transport	Y.1300–Y.1399
Interworking	Y.1400–Y.1499
Quality of service and network performance	Y.1500–Y.1599
Signalling	Y.1600–Y.1699
Operation, administration and maintenance	**Y.1700–Y.1799**
Charging	Y.1800–Y.1899
NEXT GENERATION NETWORKS	
Frameworks and functional architecture models	Y.2000–Y.2099
Quality of Service and performance	Y.2100–Y.2199
Service aspects: Service capabilities and service architecture	Y.2200–Y.2249
Service aspects: Interoperability of services and networks in NGN	Y.2250–Y.2299
Numbering, naming and addressing	Y.2300–Y.2399
Network management	Y.2400–Y.2499
Network control architectures and protocols	Y.2500–Y.2599
Security	Y.2700–Y.2799
Generalized mobility	Y.2800–Y.2899

For further details, please refer to the list of ITU-T Recommendations.

ITU-T Recommendation Y.1711

Operation & Maintenance mechanism for MPLS networks

Summary

This Recommendation provides mechanisms for user-plane OAM (Operation and Maintenance) functionality in MPLS networks according to the requirements and principles given in ITU-T Rec. Y.1710. This Recommendation is designed primarily to support point-to-point and multipoint-to-point explicit routed LSPs (ER-LSPs) with limited applicability to LSPs that employ penultimate hop popping (PHP).

The OAM mechanisms defined in this Recommendation assume common forwarding of the LSP payload and Y.1711 PDUs. In some situations this may not be true, such as when the LSP payload is load balanced across a plurality of parallel paths while still appearing as a single trail to the ingress and egress. LSRs introducing variations in connectivity are responsible for ensuring that the availability behaviour of Y.1711 per ingress-egress pair is preserved.

Source

ITU-T Recommendation Y.1711 was approved on 12 February 2004 by ITU-T Study Group 13 (2001-2004) under the ITU-T Recommendation A.8 procedure.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications. The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure e.g. interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementors are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database.

CONTENTS

ITU-T Recommendation Y.1711

Operation & Maintenance mechanism for MPLS networks

1 Scope

This Recommendation provides mechanisms for user-plane OAM (Operation and Maintenance) functionality in MPLS networks according to the requirements and principles given in ITU-T Rec. Y.1710. This Recommendation is designed primarily to support point-to-point and multipoint-to-point explicit routed LSPs (ER-LSPs) with limited applicability to LSPs that employ penultimate hop popping (PHP).

The OAM mechanisms defined in this Recommendation assume common forwarding of the LSP payload and Y.1711 PDUs. In some situations this may not be true, such as when the LSP payload is load balanced across a plurality of parallel paths while still appearing as a single trail to the ingress and egress. LSRs introducing variations in connectivity are responsible for ensuring that the availability behaviour of Y.1711 per ingress-egress pair is preserved.

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

2.1 Normative references

[1] ITU-T Recommendation I.610 (1999), *B-ISDN operation and maintenance principles and functions*.

[2] ITU-T Recommendation G.805 (2000), *Generic functional architecture of transport networks*.

[3] ITU-T Recommendation M.20 (1992), *Maintenance philosophy for telecommunication networks*.

[4] ITU-T Recommendation Y.1710 (2002), *Requirements for Operation & Maintenance functionality in MPLS networks.*

[5] IETF RFC 3031 (2001), *Multiprotocol Label Switching Architecture. Category: Standards Track.*

[6] IETF RFC 3032 (2001), *MPLS Label Stack Encoding. Category: Standards Track.*

[7] IETF RFC 2373 (1998), *IP Version 6 Addressing Architecture. Category: Standards Track.*

[8] IETF RFC 3270 (2002), *Multi-Protocol Label Switching (MPLS) Support of Differentiated Services. Category: Standards Track.*

[9] IETF RFC 3209 (2001), *RSVP-TE: Extensions to RSVP for LSP Tunnels. Category: Standards Track.*

[10] IETF RFC 3212 (2002), *Constraint-Based LSP Setup using LDP. Category: Standards Track.*

[11] IETF RFC 3429 (2002), *Assignment of the 'OAM Alert Label' for Multiprotocol Label Switching Architecture (MPLS) Operation and Maintenance (OAM) Functions*.

[12] ITU-T Recommendation Y.1712 (2004), *OAM functionality for ATM-MPLS interworking.*

2.2 Informative references

[13] IETF RFC 1930 (1996), *Guidelines for creation, selection and registration of an Autonomous System (AS).*

3 Definitions

This Recommendation introduces some terminology, which is required to discuss the functional network components associated with OAM. These definitions are consistent with G.805 terminology.

3.1 backward direction: The backward direction is opposite to the forward direction.

3.2 client/server (relationship between layer networks): [2] A term referring to the transparent transport of a client (i.e., higher) layer link connection by a server (i.e., lower) layer network trail.

3.3 defect: [3] Interruption of the capability of a transport entity (e.g., network connection) to transfer user or OAM information. This Recommendation defines seven types of defects (the detailed specifications of which are given later):

– **dLOCV**: Loss of Connectivity Verification defect.

– **dExcess**: Receiving excess rate of CV.

– **dPeerME**: Peer network maintenance entity defect.

– **dServer**: Server layer defect. Any server layer defect arising from a non-MPLS layer technology below the lowest MPLS layer network.

– **dTTSI_Mismatch**: Trail Termination Source Identifier Mismatch defect.

– **dTTSI_Mismerge**: Trail Termination Source Identifier Mismerge defect.

– **dUnknown**: Unknown defect in the MPLS network.

3.4 failure: [3] Termination of the capability of a transport entity to transfer user or OAM information. A failure can be caused by a persisting defect.

3.5 forward direction: The forward direction is the direction that traffic and the OAM PDUs directed to the near end trail termination packets flow on an LSP.

3.6 link connection: A partition of a layer N trail that exists between two logically adjacent switching points within the layer N network.

3.7 subnetwork: [2] A subnetwork is a contiguous topological region of a network delimited by its set of peripheral access points, and is characterized by the possible routing across the subnetwork between those access points. A network is the largest subnetwork and a node is the smallest subnetwork (at least in practical physical terms, though there are smaller subnetworks within nodes).

3.8 trail: [2] A generic transport entity at layer N which is composed of a payload field (which can carry a packet from a client higher layer N – 1 trail entity) with specific overhead added to ensure the forwarding integrity of the trail transport entity at layer N.

3.9 trail termination point: [2] A source or sink point of a trail at layer N, at which the trail overhead is added or removed respectively. A trail termination point must have a unique means of identification within the layer network.

3.10 user-plane: This refers to the set of traffic forwarding components through which traffic flows. CV OAM (or alternately FFD OAM) PDUs are periodically inserted into this traffic flow to monitor the health of those forwarding components. The user-plane is also sometimes called the data-plane (especially in IETF). Note that control-plane protocols (e.g., for signalling or routing) and management-plane protocols will require their own user-plane, and their user-plane may or may not be congruent (to varying degrees) with the traffic bearing user-plane.

4 Symbols and abbreviations

This Recommendation uses the following abbreviations:

AIS	Alarm Indication Signal
AS	Autonomous System
ASM	Availability State Machine
BDI	Backward Defect Indicator
BIP	Bit Interleaved Parity
CR-LDP	Constraint-based Routing Label Distribution Protocol
CV	Connectivity Verification
DL	Defect Location
DoS	Denial of Service
DT	Defect Type
E-LSP	EXP-Inferred-PSC LSP
FDI	Forward Defect Indicator
FFD	Fast Failure Detection
FRR	Fast Re-Route
LB	Loopback
LB-Req	Loopback Request
LB-Rsp	Loopback Response
LSP	Label Switched Path
L-LSP	Label-Only-Inferred-PSC LSP
LSR	Label Switching Router
MPLS	Multi-Protocol Label Switching
NMS	Network Management System
OAM	Operation and Maintenance
PHB	Per Hop Behaviour
PHP	Penultimate Hop Popping
PSC	PHB Scheduling Class
RSVP	Resource reSerVation Protocol
SDH	Synchronous Digital Hierarchy
SLA	Service Level Agreement

TTL	Time to Live
TTSI	Trail Termination Source Identifier

5 Introduction

This Recommendation provides OAM techniques to meet the OAM requirements defined in ITU-T Rec. Y.1710. It is based on the following OAM packets that are defined in detail later, but in summary:

Connectivity Verification: Two connectivity verification probes are defined, CV and FFD. The role of CV is to instrument the LSP availability model described in clause 7. FFD is used for applications such as protection switching that require much faster response. Use of connectivity verification for applications requiring different insertion intervals is for further study.

CV: The CV flow is generated at the LSP's source LSR with a nominal frequency of 1/s and terminated at the LSP's sink LSR. The CV packet contains a network-unique identifier (TTSI) so that all types of defects can be detected.

FFD: FFD provides failure detection option for a P2P LSP independent of the CV based availability model and is not tied to the CV insertion rate. Insertion rates at 1/s or faster may also be used to instrument the availability model. Slower insertion rates will detect persistent MPLS layer defects but will not correctly instrument the availability model. Slower insertion with corresponding slower detection will impact monitored client layers as mismatches in time to detect defects that can result in additional alarms and possibly misleading alarm cycling and/or alarm persistency in the client layers. The effect of slower insertion rates on defect exit criteria (especially for mismerge and misbranch defects) is for further study.

The provisioning of FFD on an LSP will provide capability of fast failure detection (by default FFD OAM is not provisioned). It is recommended that FFD will be generated at the LSP ingress at a rate of 20 per second. This will provide failure detection on the LSP in order of 100s of milliseconds. When FFD is provisioned on an LSP CV PDU generation is disabled.

The FFD PDU contains the same information as CV. PDU processing at the egress is similar for CV and FFD with adjustments for the higher frequency of FFD PDU arrival and potential of misbranching and mismerging defects between LSPs using either FFD or CV.

It is recommended that FFD should only be used as the lowest layer LSP (Level 1), i.e., there should never be plain CV LSPs serving a FFD LSP. (Note the reason for this recommendation is that if there is a lower layer LSP running plain CV below the FFD LSP, then failures within or below the plain CV LSP may not send FDI upwards into the FFD LSP fast enough to stop the FFD thinking the true defect is in the FFD LSP. This is because it may take the plain CV LSP up to 3 s to detect a defect and insert FDI (at 1/s), whereas the FFD LSP will detect the defect in ≤ 150 ms and thus take action faster.

FDI: The FDI flow is generated in response to detecting defects (e.g., from the CV flow). Its primary purpose is to suppress alarms in layer networks above the level at which the defect occurs. It is generated at either:

i) the LSR, which first detects a dServer/dUnknown defect; or

ii) the LSP terminating LSR for all MPLS layer defects.

BDI: The BDI flow is injected on a return path (such as a return LSP) to inform the upstream LSR (which is the source of the forward LSP) that there is a defect at the downstream LSP's LSR sink point. BDI therefore tracks FDI in terms of its period of generation. BDI packets may be useful in 1:1/N instances of protection switching.

Performance: These are FFS. However, the intention is to have an on-demand method of determining packet/octet loss on an LSP in order to aid trouble-shooting. They are not intended to be used as a temporally permanent OAM function (unlike the CV flow), but they could be.

Note that path trace, performance monitoring and loopback functions are FFS.

BDI and loopback transactions use a return path. A return path could be:

- A dedicated return LSP;
- A shared return LSP, which is shared between many forward LSPs;
- A non-MPLS return path, such as an out-of-band IP path. This option has potential security issues. For example, the return path could be terminated on a different LSR interface, and potentially a malicious user could generate a BDI and send it to the ingress LSR. Therefore, due to the possibility of DoS attack, additional security measures must be taken. Operators should use the optional TTSI field in BDI packets in order to assure authentication of such packets so that the receivers of BDI OAM packets may verify that the sender of the packet is valid.

All OAM packets are identified within a LSP traffic stream by the use of globally well-known and reserved label codepoint (of 14). Further details on the encoding of OAM packets are provided later.

It is strongly recommended that CV OAM or FFD OAM packets be generated on each LSP (in order to detect all defects and potentially provide protection against traffic leakage both in and out of LSPs). It is also recommended that FDI OAM packets be used to suppress alarm storms. BDI packets are a useful tool for single-ended monitoring of both directions and also, in some protection-switching cases. However, these are only recommendations, and operators can choose to use some or all of the OAM packets as they see fit. Appendix I discusses some of the options for generating and processing CV flows.

OAM techniques are applied on a per LSP basis. If a segment of a given LSP at layer N is to be monitored for some reason (e.g., via a CV or P flow say), one way to do this is by creating a new server layer LSP (i.e., at layer N + 1) to cover the segment at layer N.

5.1 Overview of functionality

The OAM defect detection function is based on the periodic transmission of CV or FFD packets from ingress to egress of an LSP. CV packet generation rate is 1 packet per second whereas recommended packet generation rate for FFD packets is 20 per second. Each CV and FFD packet carries a unique TTSI (Trail Termination Source Identifier), which is composed of the source LSR identifier, and the LSP identifier.

An LSP enters a defect state when one of the defects noted in clause 3 occurs (these are defined in detail later in terms of precise entry/exit criteria and consequent actions).

In addition to the CV packet, there are other OAM packet types defined that provide consequential fault handling or performance monitoring functions. These will be defined later. All OAM packets are identified in terms of a function type by the first octet of the OAM packet payload as follows (see Table 1):

Table 1/Y.1711 – OAM function type codepoints

OAM function type codepoint (Hex)	First octet of OAM packet payload function type and purpose
00	Reserved
01	CV (Connectivity Verification)
02	FDI (Forward Defect Indicator)
03	BDI (Backward Defect Indicator)
04	Reserved for Performance packets
05	Reserved for LB-Req (Loopback Request)
06	Reserved for LB-Rsp (Loopback Response)
07	FFD (Fast Failure Detection)

All other OAM Function Type codepoints are reserved for possible future standardization.

5.2 Identification of OAM packets from normal user-plane traffic

The label structure defined in [6] indicates a single label field of 20 bits. Some label field values have already been reserved for special functions [6].

This Recommendation introduces a new globally reserved label value, herein referred to as the "OAM Alert Label". The recommended numerical value for the OAM Alert Label is 14 [11].

5.3 OAM payload

The payload of an OAM packet is composed of the OAM Function Type, the specific OAM function type data and a common BIP16 error detection mechanism.

All OAM packets must have a minimum payload length of 44 octets to facilitate ease of processing and to support minimum packet size requirements of current L2 technologies (e.g., Ethernet). This is achieved by padding the specific OAM type data field with all 0s when necessary. All padding bits are reserved for possible future standardization.

The order of transmission is from left to right, most significant bit (MSB) to least significant bit (LSB).

5.4 Handling of errored OAM packets

Each OAM packet uses a BIP16 (in the last two octets of the OAM payload area) to detect errors. The BIP16 remainder is computed over all the fields of the OAM payload, including the Function Type and the BIP16 bit positions (which are all pre-set to zero for initial calculation purposes).

The BIP16 generator polynomial is $G(x) = x^{16} + 1$.

BIP16 processing must be performed on all OAM packets prior to being able to reliably pass their payload for further processing. Any OAM packets that show a BIP16 violation upon reception processing should be discarded.

In the case of the CV or FFD packet flow, persistent BIP16 violations will cause a Loss of Connectivity Verification (dLOCV). This behaviour is consistent with the nature of the actual defect being experienced. However, it is recommended that at a local equipment level some notification is given to the Network Management System to indicate when any BIP16 discards are occurring, especially if these give rise to an associated dLOCV.

In the case of the other OAM packet types, i.e., the FDI, BDI and P packets, it is again recommended that at a local equipment level, some indication is given to the Network Management System that BIP16 discards are occurring. The threshold to be used for recording/reporting such BIP16 discard activity for these OAM packets should be programmable, and is outside the scope of this Recommendation.

5.5 Engineering cost/risk considerations

Operators must consider the impact of OAM functions on nodal processing resources and network traffic overhead vs an ability to detect all MPLS user-plane failures.

In the case of the CV or FFD OAM packet, which forms the basis for defect detection, a clear distinction can be made between the source generation implications and the sink processing implications. This aspect is further discussed in Appendix I.

5.6 Backward compatibility considerations

LSRs that do not support OAM functionality will drop OAM packets (because the OAM Alert Label is not recognized) and therefore will not have an adverse impact on user-plane traffic.

6 OAM mechanisms

6.1 Features common to all OAM packets

Some fields in the OAM packet header have a common treatment in all OAM packets. These are explained below.

6.1.1 Stack encoding

OAM packets are differentiated from normal user-plane traffic by an increase of one in the label stack depth at a given LSP level at which they are inserted. Therefore, they maintain this label stack difference of one (from normal user-plane traffic) as they traverse any lower layer server LSPs.

Label

The OAM Alert Labelled header is added before (i.e., below) the normal user-plane forwarding labelled header at the LSP trail source point.

EXP

The OAM packets can be used on both E-LSPs and L-LSPs. The coding of the EXP field should be set to all 0s in the OAM Alert Labelled header and to whatever is the "minimum loss-probability PHB" in the preceding normal user-plane forwarding header for that LSP. This is to ensure the OAM packets have a PHB which ensures the lowest drop probability [5]. OAM capabilities defined in the future may require different encoding of the EXP field.

S bit

The S bit is set only in the OAM Alert Labelled header.

TTL

The TTL field should be set to 1 in the OAM Alert Labelled header. The reasons for this are:

- OAM packets should never travel beyond the LSP trail termination sink point at the LSP level they were originally generated (noting that they are not examined by intermediate label-swapping LSRs, and are only observed at LSP sink points).
- the TTL of the immediately prior normal user-plane forwarding header is used to mitigate against damage from looping packets.

6.1.2 Intermediate/penultimate processing

OAM packets are transparent to intermediate LSRs, including the penultimate LSRs.

6.1.3 Server/client relationships

OAM packets within a given LSP are not synchronous to any other OAM packets in any other LSP (this includes all nested LSPs, and OAM packets from the remote end of an LSP at level N but in the other direction when bidirectional LSPs at level N are being used).

6.1.4 TTSI (Trail Termination Source Identifier) structure

The structure of the LSP Trail Termination Source Identifier (TTSI) is defined by using a 16-octet LSR ID IPv6 address followed by a 4-octet LSP Tunnel ID. Note that the first 2 octets (MSB octets) of the LSP Tunnel ID are currently padded with all 0s to allow for any future increase in the Tunnel ID field.

LSR ID	LSP ID
16 octets	4 octets

Figure 1/Y.1711 – TTSI structure

For nodes that do not support IPv6 addressing, an IPv4 address can be used for the LSR ID using the format described in RFC 2373 [7]. That is:

All 00Hex padding	All FFHex padding	IPv4 address
10 octets	2 octets	4 octets

Figure 2/Y.1711 – LSR ID structure using IPv4 address

The TTSI provides the network-unique access point identifier of an LSP. It belongs to the traffic data-plane, and must be consistent and independent of which applications use the LSP and/or how the LSP was instantiated, e.g., by management provisioning or signalling.

6.1.5 Provisioning and signalling of expected TTSI at LSP sink points

On LSP establishment the LSP trail termination sink point should be configured with the expected TTSI. Although it could be configured manually, ideally this should be done automatically via LSP signalling at LSP set-up time (e.g., via a CR-LDP or RSVP control-plane mechanism). TTSI (Trail Termination Source Identifier) forms a network unique access point identifier constructed from the source LSR identifier and the LSP identifier. Mechanisms for automatic signalling/configuration of TTSI for near-end monitoring of CV and FFD are for further study.

NOTE – Attention is drawn to the fact that the activation/deactivation of CV/FFD/FDI/BDI functionality needs to be closely tied to the conscious set-up/tear-down of LSPs. This is necessary in order to ensure that consequent actions (especially alarms) are enabled/disabled as appropriate. For example, it is obvious that CV or FFD processing should be activated (deactivated) after (before) an LSP is set up (taken down).

6.2 Connectivity Verification (CV)

The Connectivity Verification function is used to detect/diagnose all types of LSP connectivity defects (sourced either from below or within the MPLS layer networks).

Payload Structure

Function type (01Hex)	Reserved (all 00Hex)	LSP Trail Termination Source Identifier	Padding (all 00Hex)	BIP16
1 octet	3 octets	20 octets	18 octets	2 octets

Figure 3/Y.1711 – CV payload structure

More information is given in Appendix I about various ingress/egress processing options.

6.3 Fast Failure Detection (FFD)

The fast failure detection function is used to detect connection verification at a smaller time-scale compared to CV functionality. In addition, like CV OAM, it detects all types of LSP connectivity defects (sourced either from below or within the MPLS layer networks).

Payload Structure

Function type (07Hex)	Reserved (all 00Hex)	LSP Trail Termination Source Identifier	Frequency	Padding (all 00Hex)	BIP16
1 octet	3 octets	20 octets	1 octet	17 octets	2 octets

Figure 4/Y.1711 – FFD payload structure

Where Frequency is one of:

00 reserved

01 10 ms

02 20 ms

03 50 ms (default value)

04 100 ms

05 200 ms

06 500 ms

07-255 reserved

If a reserved value is received, the egress cannot determine the frequency of probe injection by the ingress; therefore dLOCV is not a valid defect state.

6.4 Forward Defect Indication (FDI)

Forward Defect Indication is generated by a LSR detecting any defect (defined later) and inserted into affected client layers. FDI OAM packets are generated on a nominal 1 per second basis. The FDI packet traces forward and upward through any nested LSP stack. Its primary purpose is to suppress alarms being raised at affected higher level client LSPs and (in turn) their client layers (where the higher layer clients may not be in the same management domain as the initial defect source). It includes fields to indicate the nature of the defect and its location.

Near-end defect processing for MPLS LSPs or serving layer links will suppress FDI generation for MPLS client layers/levels that may have downstream merge points prior to the LSP egress. At the present time this includes LSPs set-up using LDP and LSPs set-up with RSVP-TE enhanced with FRR.

The FDI is sent downstream from the first node detecting the defect. In the case of MPLS server layer failures (i.e., in a lower layer technology such as SDH), this would be the first LSR downstream of the server layer failure (as a consequence of the appropriate client/server adaptation of the server FDI signal). In the case of MPLS layer failures (i.e., failures within the MPLS fabric), this would be the first LSR LSP trail termination sink point at the same LSP level as the failure.

Payload Structure

Function type (02Hex)	Reserved (00Hex)	Defect type	TTSI (optional, if not used set to all 00Hex)	Defect location	Padding (all 00Hex)	BIP16
1 octet	1 octet	2 octets	20 octets	4 octets	14 octets	2 octets

Figure 5/Y.1711 – FDI payload structure

The Defect Type field is set at 2 octets, and the values are given later.

The inclusion of the TTSI in the FDI packet is optional, and may have use in case of penultimate hop popping (PHP) as discussed in Appendix II. If PHP is not used the TTSI field could be encoded as all zeros.

Defect location is a 4-byte field. The identity of the network in which the defect has been detected should be encoded in the Defect Location (DL) in the form of an Autonomous System (AS) number. RFC 1930 [13] defines the AS number as being 2 bytes long. However, a 4-octet field is allocated to the defect location field to allow for larger AS number to be defined in the future. To encode a 16-bit AS number the following procedure is used:

The 16-bit AS number is encoded in the LSB half of the DL field and the MSB half of DL field is set to zero.

The FDI OAM packet is recursively mapped upwards, through a client/server adaptation process at LSP trail termination sink points, into any further affected higher client layer LSPs. When this arrives at the top LSP it needs to be mapped into an equivalent FDI for whatever client layer is then being carried. In the case of IP (or indeed any other client layer), this mapping is outside the scope of this Recommendation.

Note that higher level LSPs that straddle a lower level defect will also detect defects (as a result of corruption of their own CV or FFD flow) but they will also see an incoming FDI OAM packet flow from the lowest level LSP where the defect was initially detected. This dynamic behaviour allows for correct identification of the true source of the defect and is explained in more detail later. But for now, it is sufficient to note that the incoming FDI is needed to:

- Suppress unnecessary alarms in the affected higher layer LSPs.
- Give an indication to affected higher level LSPs that the defect is at a lower level LSP.
- Allow the appropriate BDI coding at the affected higher layer.

It is important that the LSP sink point knows (for the duration that the LSP is in service) any server→client LSP label mappings that were in existence prior to the defect. Although the exact means for achieving this are outside the scope of this Recommendation, some examples of how these server→client layer label mappings could be configured are as follows:

- manually, e.g., via the NMS;
- automatically on LSP set-up via extensions to LSP signalling;

- by an automatic "learning process", i.e., if, during the establishment of the client LSPs, the signalling is tunnelled through the server layer, then the server trail terminating node could keep the information about the established LSPs in memory as they occur.

When a FDI is to be passed from a server layer LSP to its client layer LSP(s) (i.e., at the client/server adaptation function following the server layer LSP trail termination sink point), the Defect Location and Defect Type field should be copied from the server layer LSP FDI into the client layer LSP(s) FDI.

6.5 Backward Defect Indication (BDI)

The purpose of the BDI OAM function is to inform the upstream end of an LSP of a downstream defect. However, to do this it requires a return path. Backward Defect Indication is generated at a return path's trail termination source point in response to a defect being detected at a LSP trail termination sink point in the forward direction. The functionality provided by the BDI is useful for applications such as single-ended measurements of the short-break/availability/network performance of both directions, or for providing an indication for certain types of protection switching.

Payload Structure

Function type (03Hex)	Reserved (00Hex)	Defect type	TTSI (optional, if not used set to all 00Hex)	Defect location	Padding (all 00Hex)	BIP16
1 octet	1 octet	2 octets	20 octets	4 octets	14 octets	2 octets

Figure 6/Y.1711 – BDI payload structure

The BDI is sent from the LSP trail source point of the return path as a mirror of the appropriate (see Note) FDI at the LSP trail sink point of the other direction. The Defect Location and Defect Type fields are a direct mapping of those obtained from the appropriate (see Note) FDI and have identical formats as described previously for the FDI OAM packet.

NOTE – The word "appropriate" here signifies that any incoming FDI (i.e., from a lower layer) takes precedence over any FDI that would have been generated at the layer being considered due to detecting defects at this layer (where these defects are only consequential as a result of a lower layer defect).

The inclusion of the TTSI in the BDI packet is optional, and may have use in case of shared or out-of-band return paths for BDI as discussed in Appendix II. For bidirectional LSPs, this field could be set to all zeros. If used, the TTSI is set to the expected TTSI of the forward LSP that the BDI corresponds to.

6.6 On-demand diagnostic OAM mechanisms

A number of on-demand diagnostic OAM mechanisms can be defined. These are intended to provide Operational people with additional OAM tools, e.g., to help diagnose network problems.

6.6.1 Performance measurements

These are for further study.

6.6.2 Loopback transactions

These are for further study.

6.7 Defect type codepoints

The defect type code is encoded in two octets. The first octet indicates the layer and second octet indicates the nature of the defect.

Table 2/Y.1711 – Defect type codepoints in FDI/BDI OAM packets

Defect type	DT code (Hex)	Meaning
dServer	01 01	Any server layer defect arising below the MPLS layer network. It is not suggested that these are individually identified and defined for each type of server layer, since this function is only appropriate to the server layer itself. Hence, only an indication is needed that it is the server layer and not the MPLS layer. Note that this defect is not generated by MPLS OAM mechanisms, rather it is an input to MPLS OAM from server layer.
dPeerME	01 02	Any peer maintenance entity defect arising outside the MPLS subnetwork. It is not suggested that these are individually identified and defined for each type of peer network. Hence, only an indication is needed that it is not a defect in the MPLS subnetwork. Note that this defect is not generated by MPLS OAM mechanisms, rather it is an input to MPLS OAM from a layer network interworking function.
dLOCV	02 01	Simple Loss of Connectivity Verification due to missing CV or FFD OAM packets with expected TTSI. Note that: 1) If the cause of dLOCV is the server layer (i.e., there is also an incoming FDI signal from the server layer), then the DT codepoint 01 01Hex is used. 2) If the cause of dLOCV is the peer maintenance entity (i.e., there is also an incoming FDI signal from the IWF with a DT codepoint of 01 02), then the DT codepoint 01 02Hex is used. The dLOCV codepoint 02 01Hex is only used for MPLS layer simple connectivity failures only.
dTTSI_Mismatch	02 02	Trail Termination Source Identifier Mismatch defect is due to unexpected and no expected TTSI observed in the incoming CV or FFD OAM packets. This detects misconfigured (e.g., swapped) connections. Note that the dTTSI_Mismatch defect condition takes priority over the dLOCV defect condition which is also present.
dTTSI_Mismerge	02 03	Trail Termination Source Identifier Mismerge defect is due to both unexpected and expected TTSI observed in the incoming CV or FFD OAM packets. This detects both mis-branching and unintended replication failures. Note that unlike dTTSI_Mismatch there is no dLOCV defect condition also present. However, this defect must not be confused with dExcess; since although the incoming CV or FFD rate will be increased, the increase will not solely be due to CV or FFD packets with expected TTSI.
dExcess	02 04	A dExcess defect is detected by observation of an increased rate of CV or FFD OAM packets with the expected TTSI above the nominal rate of 1/s for CV OAM packets and 20/s for FFD OAM packets. This could be due to self-mismerging, a faulty source LSR, DoS attack, etc.

Table 2/Y.1711 – Defect type codepoints in FDI/BDI OAM packets

Defect type	DT code (Hex)	Meaning
dUnknown	02 FF	Unknown defect detected in the MPLS layer. This is expected to be used for MPLS nodal failures which are detected within the node (probably by proprietary means) and affect user-plane traffic. Note that this defect is not detected by MPLS OAM, rather it is an input to MPLS OAM.
None	00 00	Reserved
None	FF FF	Reserved

6.8 Defect type entry/exit criteria and consequent actions

There are 4 MPLS layer user-plane defects, i.e., dLOCV, dTTSI_Mismatch, dTTSI_Mismerge and dExcess, which are defined in more detail.

NOTE 1 – Since OAM packet flows are not synchronized in LSPs at different hierarchical levels (i.e. when LSPs are nested), there is a possibility that a client layer LSP detects a defect before its server layer LSP. This error could be up to 1 s due to CV packet arrival time differences plus some additional uncertainty due to network delay effects. This could result in an error of judgment as to the type of defect that is present and hence which consequent actions are appropriate; especially whether the raising of a local alarm is appropriate and the correct setting of the DL and DT codepoints in FDI/BDI OAM packets. To mitigate this effect, it is recommended that the raising of an alarm is deferred for at least 2 seconds after a defect state is detected (the exact value is FFS). This will also allow the network to settle into a stable state as regards defect detection behaviour.

NOTE 2 – The starting/stopping of aggregation of any LSP user-plane network performance measurements, e.g., packet/octet loss metrics (e.g., if using the P OAM packet), is dependent on whether the LSP is in the available or unavailable state.

NOTE 3 – If more than one defect is simultaneously present for a given LSP, then the suggested order of priority is as follows: dTTSI_Mismatch, dTTSI_Mismerge, dLOCV and dExcess.

6.8.1 dLOCV entry criteria

Entry to the dLOCV condition, and hence entry to the LSP Trail Sink Near-End Defect State, occurs when:

- For an LSP not provisioned with fast failure detection (i.e., not provisioned with FFD OAM packets), there are no expected CV OAM packets observed in any period of 3 consecutive seconds.
- For an LSP provisioned with fast failure detection (i.e., provisioned with FFD OAM packets), there are no expected FFD OAM packets observed in any period of 3 consecutive x intervals, where $x < 1$s ($x > 1$ s is for further study) is the insertion interval for FFD OAM packets in the FFD PDU frequency field or locally provisioned. (Note that the 3x window moves forward with x second granularity. Further note that the recommended value for periodic insertion interval for FFD OAM packets x is 50 ms in accordance with the recommended frequency of 20 per seconds.)

In terms of consequent actions

- If there is an incoming FDI signal from a server layer below the MPLS network, then this is mapped to the DT codepoint 01 01 Hex in the FDI OAM packets sent forwards and any BDI OAM packets sent backwards. The local DL codepoint is also inserted in these FDI and BDI OAM packets. There are no alarms associated with the MPLS layer itself but only the server layer which sourced the FDI signal.

Else

- If there is an incoming FDI signal from a lower level LSP within the MPLS network, then that FDI signal's DL/DT codepoints are mapped into the FDI sent to any further client layers (i.e., suppresses generation of FDI DL/DT codepoints from this point) and the BDI OAM packet sent backwards. There are no alarms generated regarding this LSP (the alarm will be associated with the lowest layer LSP within which the defect originated).

Else

- If there is no FDI signal incoming from the server layer or a lower level LSP AND there are no CV or FFD OAM packets observed with an unexpected TTSI, then the DT codepoint 02 01 Hex is inserted in the FDI OAM packets sent downstream and any BDI OAM packets sent upstream. The local DL codepoint is also inserted in these FDI and BDI OAM packets. A local alarm is raised relevant to this defect condition.

6.8.2 dTTSI_Mismatch entry criteria

Entry to the dTTSI_Mismatch condition, and hence entry to the LSP Trail Sink Near-End Defect State, occurs when:

- For an LSP not provisioned with fast failure detection (i.e., not provisioned with FFD OAM packets), there are any CV or FFD OAM packets observed in any period of 3 consecutive seconds each with an unexpected TTSI and there are no CV OAM packets observed with an expected TTSI in the same period.
- For an LSP provisioned with fast failure detection (i.e., provisioned with FFD OAM packets), there are any CV or FFD packets observed in any period of 3 consecutive *x* intervals each with an unexpected TTSI and there are no FFD OAM packets observed with an expected TTSI in the same period.

It should be noted that the dTTSI_Mismatch defect overrides the dLOCV defect (as would be the case, for example, with swapped LSPs). The DT codepoint 02 02 Hex is inserted in the FDI OAM packets sent forwards and any BDI OAM packets sent backwards. The DL is also inserted in these FDI and BDI OAM packets. A local alarm is raised relevant to this defect condition and the unexpected TTSI captured locally (this may also be optionally sent to the NMS as an exception report). The downstream traffic signal must also be suppressed.

6.8.3 dTTSI_Mismerge entry criteria

Entry to the dTTSI_Mismerge condition, and hence entry to the LSP Trail Sink Near-End Defect State, occurs when:

- For an LSP not provisioned with fast failure detection (i.e., not provisioned with FFD OAM packets), there are any CV or FFD OAM packets each with an unexpected TTSI and there are any CV OAM packets each with an expected TTSI observed in any period of 3 consecutive seconds.
- For an LSP provisioned with fast failure detection (i.e., provisioned with FFD OAM packets), there are any CV or FFD OAM packets each with an unexpected TTSI and there are any FFD OAM packets each with an expected TTSI observed in any period of 3 consecutive *x* intervals.

It should be noted that dTTSI_Mismerge must not be confused with dExcess since, in the former case, the increase in received CV or FFD packets is due to both CV or FFD packets with expected and unexpected TTSI, whilst in the latter case this is only due to CV or FFD packets with expected TTSI. The DT codepoint 02 03 Hex is inserted in the FDI OAM packets sent forwards and any BDI OAM packets sent backwards. The DL codepoint is also inserted in these FDI and BDI OAM packets. A local alarm is raised relevant to this defect condition and the unexpected TTSI captured locally (this may also be optionally sent to the NMS as an exception report say). The downstream traffic signal could optionally be suppressed.

6.8.4 dExcess entry criteria

Entry to the dExcess condition, and hence entry to the LSP Trail Sink Near-End Defect State, occurs when:

- For an LSP not provisioned with fast failure detection (i.e., not provisioned with FFD OAM packets), there are ≥ 5 CV OAM packets observed in any period of 3 consecutive seconds each with an expected TTSI.
- For an LSP provisioned with fast failure detection (i.e., provisioned with FFD OAM packets), there are ≥ 5 FFD OAM packets observed in any period of 3 consecutive *x* intervals each with an expected TTSI.

The DT codepoint 02 04 Hex is inserted in the FDI OAM packets sent forwards and any BDI OAM packets sent backwards. The local DL codepoint is also inserted in these FDI and BDI OAM packets. A local alarm is raised relevant to this defect condition.

6.8.5 Defect exit criteria

Exit of the dLOCV, dTTSI_Mismatch, dTTSI_Mismerge or dExcess defect condition and, hence, exit of the LSP Trail Sink Near-End Defect State, occurs when in an aggregate period of 3 consecutive seconds:

- For an LSP not provisioned with fast failure detection (i.e., not provisioned with FFD OAM packets), there are:
  - ≥ 2 but ≤ 4 CV OAM packets observed each with an expected TTSI; and
  - no CV or FFD OAM packets observed with an unexpected TTSI.
- For an LSP provisioned with fast failure detection (i.e., provisioned with FFD OAM packets):
  - a sliding window of 3*x* second observed ≥ 2 and ≤ 4 FFD OAM packets with expected TTSI; and
  - no CV or FFD OAM packets observed with an unexpected TTSI.

All the consequent actions invoked when entering the LSP Trail Sink Near-End Defect State (i.e., sending of FDI and BDI OAM packets, the raising of local alarms and any suppression of traffic) are stopped when the LSP Trail Sink Near-End Defect State is exited.

7 Available and unavailable state processing

The main purpose of defining harmonized defect entry/exit criteria as noted above, is in order to significantly simplify:

- near-end/far-end LSP Trail Sink Defect State processing;
- near-end/far-end LSP Available State processing;
- the decision point at which any LSP user-plane traffic network performance metrics (if being collected) are stopped/started with respect to aggregation into long-term registers.

In all sections where the evaluation of events is described, the measurement technique is based on a sliding-window with a 1 second (and *x* second for FFD OAM option) granularity of advance. Note that the datum for the commencement of the sliding window is an arbitrary point in time decided by each node independently and is not synchronized to OAM packet arrival events on any LSP. This is deemed acceptable to allow simpler nodal processing.

It should be noted that this Recommendation uses the traditional functional dependency relationship between network performance and availability. That is:

- Network performance is a unidirectional metric, i.e., if network performance metrics are being measured then each direction is measured independently.
- Availability is a bidirectional metric in the case of bidirectional LSPs, in the sense that if any direction enters the unavailable state (defined later) then both directions are deemed to be unavailable. In the case of unidirectional LSPs, then availability can only have unidirectional significance.
- Network performance measurements must be suspended (as regards aggregation, into long-term available state registers) if an LSP enters the unavailable state; noting that this means the network performance measurements of both directions from the definition of the availability metric above in the case of bidirectional LSPs.

However, it should also be noted that (for both pragmatic reasons and to preserve their statistical significance) network performance metric aggregation is actually suspended after detecting a short-break event.

The LSP Timer (T1) defines the length of time the LSP is in the Near-End Defect State before declaring an LSP Unavailable or returning to defect free state. It starts when the LSP enters the Near-End Defect State and stops when the LSP exits the Near-End Defect State or enters Near End Unavailable state. It has a maximum value of 10 seconds.

The LSP Timer (T2) defines the length of time the LSP is in the Near-End Unavailable State minus 10 s (measured as maximum value of T1). It starts when the LSP enters the Near-End Unavailable State and stops when the LSP exits the Near-End Unavailable State plus 10 s (confirmation period for being defect free).

The LSP Timer (T3) defines the length of time the LSP is in the Far-End Defect State before declaring an LSP Unavailable or returning to Far-End defect free state. It starts when the LSP enters the Far-End Defect State and stops when the LSP exits the Far-End Defect State or enters Far-End Unavailable state. It has a maximum value of 13 seconds.

The LSP Timer (T4) defines the length of time the LSP is in the Far-End Unavailable State minus 13 s (measured as maximum value of T3). It starts when the LSP enters the Far-End Unavailable State and stops when the LSP exits the Far-End Unavailable State plus 10 s (confirmation period for being defect free).

7.1 Short-breaks

A short-break event is an event in which the exit from any of the previously defined defect conditions occurs before the LSP Unavailable Timer expires.

The start of the short-break occurs at the end of the 3 s period in which defect state is entered and the end of the short-break occurs at the end of the 3 s period in which the defect state is exited. The short-break is a transient state that can only exist when the LSP is in the available state.

7.2 Available/unavailable state definition

If the LSP Timer (T1) expires then the LSP enters the Unavailable State. The start point of the Unavailable State is deemed to be at the entry into the LSP Trail Sink Near-End Defect State.

An LSP re-enters the Available State after first exiting the LSP Trail Sink Near-End Defect State and there has been an aggregate period of 10 consecutive seconds in which:

- For an LSP not provisioned with fast failure detection (i.e., not provisioned with FFD OAM packets), there have been:
  - ≥ 9 and ≤ 11 CV OAM packets each with an expected TTSI, and
  - no CV or FFD OAM packets with an unexpected TTSI.
- For an LSP provisioned with fast failure detection (i.e., provisioned with FFD OAM packets):
  - a sliding window of 10*x* second observed ≥ 9 and ≤ 11 FFD OAM packets with expected TTSI, and
  - no CV or FFD OAM packets observed with an unexpected TTSI.

The start point of the Available State is deemed to be at the beginning of these 10 consecutive seconds.

7.3 Near-end and far-end measurements of availability

All of the above discussion is strictly only relevant to the near-end processing when the LSP trail termination sink point is in the LSP Trail Sink Near-End Defect State as discussed previously. The far-end availability behaviour can also be measured (useful when only a single end is accessible for measurement) by using the BDI signal (when bidirectional LSPs are being used) since this is a reflected upstream mirror of the duration over which FDI is sent downstream.

The LSP Trail Sink Far-End Defect State is therefore defined to be the period over which BDI OAM packets are observed, subject to the following entry and exit criteria:

- Entry of the LSP Trail Sink Far-End Defect State occurs on the first BDI OAM packet observed.
- Exit of the LSP Trail Sink Far-End Defect State occurs after a period of 3 consecutive seconds in which no BDI OAM packets have been received.

Note that this 3 s processing delay on exit is to cater for cases in which perhaps a single BDI is lost (say due to congestion or errors). Its effect must be catered for in the far-end processing state machine as discussed later.

Since the temporal duration of the far-end state is directly related to the near-end state (albeit with a +3 s exit checking period), a single end can measure both short-breaks and unavailability of both directions (on the assumption that bidirectional LSPs are being used).

≥≤

7.4 Near-end state processing flow-chart

Figure 7 summarizes many of the key points regarding the near-end state processing algorithm for a given LSP provisioned with CV flow. Near-end state processing for an LSP with FFD flow (instead of CV flow) differs in the entry criteria into the defect state, exit criteria from the defect state, and the re-entry criteria into the available state.

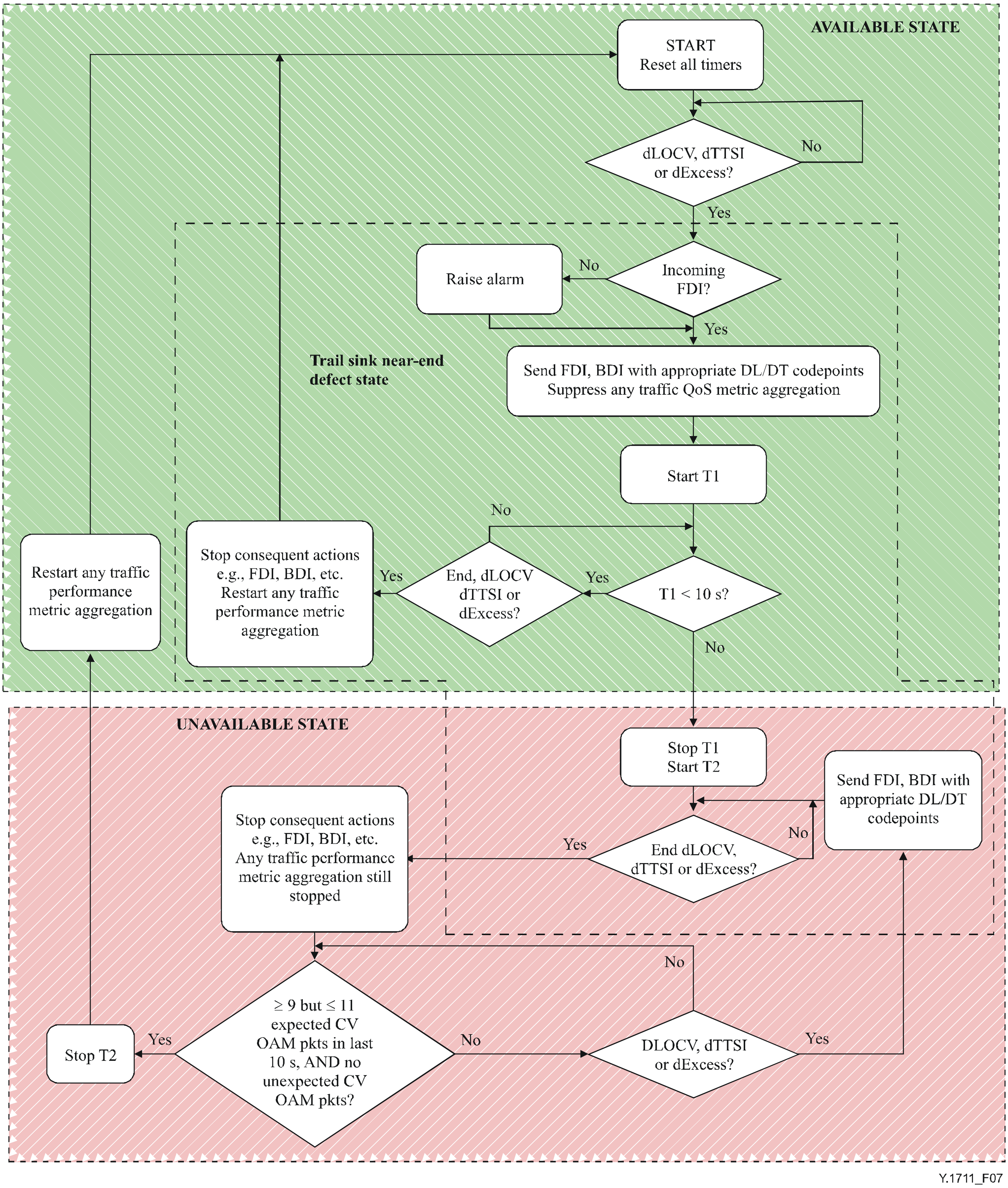


Figure 7/Y.1711 – LSP near-end state processing algorithm

NOTE – The above diagram is not meant to imply a physical realization (this is down to the vendor to interpret), but rather the logical flow of processing and consequent actions at key points. Further, for clarity it is not possible to provide all the details in such a diagram and one needs to refer to the accompanying text, e.g., dTTSI in the above diagram is meant to imply covering both Mismatch and Mismerge defect types.

1) Assume the endpoint start in the available state in the box marked "Start". All timers (shown later) can conceptually be assumed reset at this point. If there are any network performance metrics being collected (e.g., packet/octet loss measurements from the P OAM packet), then this is assumed to be active at this time.

2) The first decision box is "dLOCV, dTTSI or dExcess?". These defects were defined previously. If none of these defects are present, this condition is kept to be checked and stay in the available state. However, if one of these defects is present, the endpoint enters the Trail Sink Near-End Defect State.

3) The consequent actions then required depend on the nature of the defect observed, and whether there is any incoming FDI from a lower layer, and should follow the rules given previously. But note that any network performance metrics which are being collected are suppressed from aggregation into the long-term registers against available time. The registers are effectively backdated 3 s to allow for the defect detection time (at this stage it cannot be judged whether the event will be a Short-Break, and hence the LSP remains in the Available State, or whether the LSP will enter the Unavailable State).

4) Timer T1 is started. This timer is used to determine the duration of the Trail Sink Near-End Defect State, and if this persists for a sufficient time (i.e., a further 10 s), then this timer is used to branch the flow-chart into the Unavailable State processing region.

5) Below (timer) T1, it loops round the decision boxes "T1 < 10 s?" and "End dLOCV, dTTSI or dExcess?". It can exit this loop if the defect state ends (in accordance with criteria given previously) before T1 reaches 10 s. Since the endpoint is still in the available state, any network performance metric aggregation into the long-term registers is restarted (noting the last 3 s must be accounted for), FDI/BDI OAM packet generation is stopped and the short-break event in the local registers are captured. Additionally, if the event was due to a dTTSI, then the endpoint should also capture the TTSI of the offending LSP and cease the suppression of traffic (if appropriate). The timestamp of the event should be related to the onset of the defect which caused it. If, however, T1 reaches 10 s, the endpoint enters the Unavailable State. Note that it is not possible to enter the Unavailable State unless the Trail Sink Near-End Defect State has persisted for at least 10 s in the Available State.

6) Then a date/timestamped Unavailable State entry event in the local endpoint is registered together with information on the nature of the defect which caused it. Note that the date/timestamp must be backdated 13 s. Optionally, an exception report may also be sent to the NMS with the Unavailable State entry date/timestamp noted above, together with any other relevant information about the defect which caused it, e.g., in the case of dTTSI, this should include the TTSI of the offending LSP. Then timer T1 is stopped and timer T2 is started, whose purpose is to record the duration of the Unavailable State. Note that, when the endpoint enters the Unavailable State, it also remains in the Trail Sink Near-End Defect State.

7) It then runs round a decision box "End dLOCV, dTTSI or dExcess?", which is just below the point where timer T2 is started, which checks for the end of the defect state. When the defect ends (in accordance with the criteria given previously), FDI/BDI OAM packet generation is stopped and the endpoint exits the Trail Sink Near-End Defect State. Any network performance metric aggregation is still inhibited.

8) It then runs round the decision loop comprised of the two boxes "≥ 9 but ≤ 11 expected CV OAM packets in last 10 s AND no unexpected CV OAM packets" and "dLOCV, dTTSI or dExcess?". If a further defect occurs before the exit criteria of the former decision box are met, the endpoint re-enter the Trail Sink Near-End Defect State and hence the generation of FDI/BDI OAM packets is restarted (with DL/DT codepoints and other consequent actions relevant to the specific defect observed). Any network performance metric aggregation continues to be inhibited. In this case, it goes back to point 7) above of state processing, and recommences checking for the end of the defect. Note that timer T2 continues to run.

9) To get out of the Unavailable State, the endpoint must first have exited the Trail Sink Near-End Defect State as noted in 7) above, and then have met the criteria of the decision box "≥ 9 but ≤ 11 expected CV OAM packets in last 10 s AND no unexpected CV OAM packets?", as noted in 8) above. Note that the "last 10 s" referred to here includes the 3 s interval required to check for the end of the Trail Sink Near-End Defect State as noted above in item 7).

10) Timer T2 is then stopped, and the duration of the unavailability event in the local registers is recorded. Any network performance metric aggregation into the local registers is recommenced and all consequent actions associated with the Unavailable State are ceased. Note that T2 will record an Unavailable State duration which is 3 s less than the true unavailability event. Note also that the last 10 s belong to the Available State and so any network performance metric aggregation will need to take these 10 s into account. Optionally, an exception report may also be sent to the NMS with the Unavailable State exit date/timestamp suitably corrected, as noted above.

11) Then the process goes back to the starting point in the Available State.

7.5 Far-end state processing flow-chart

Figure 8 summarizes many of the key points regarding the far-end state processing algorithm for a given LSP. Note Far-End Defect state as well as unavailable state processing remains the same for CV or FFD provisioned LSP.

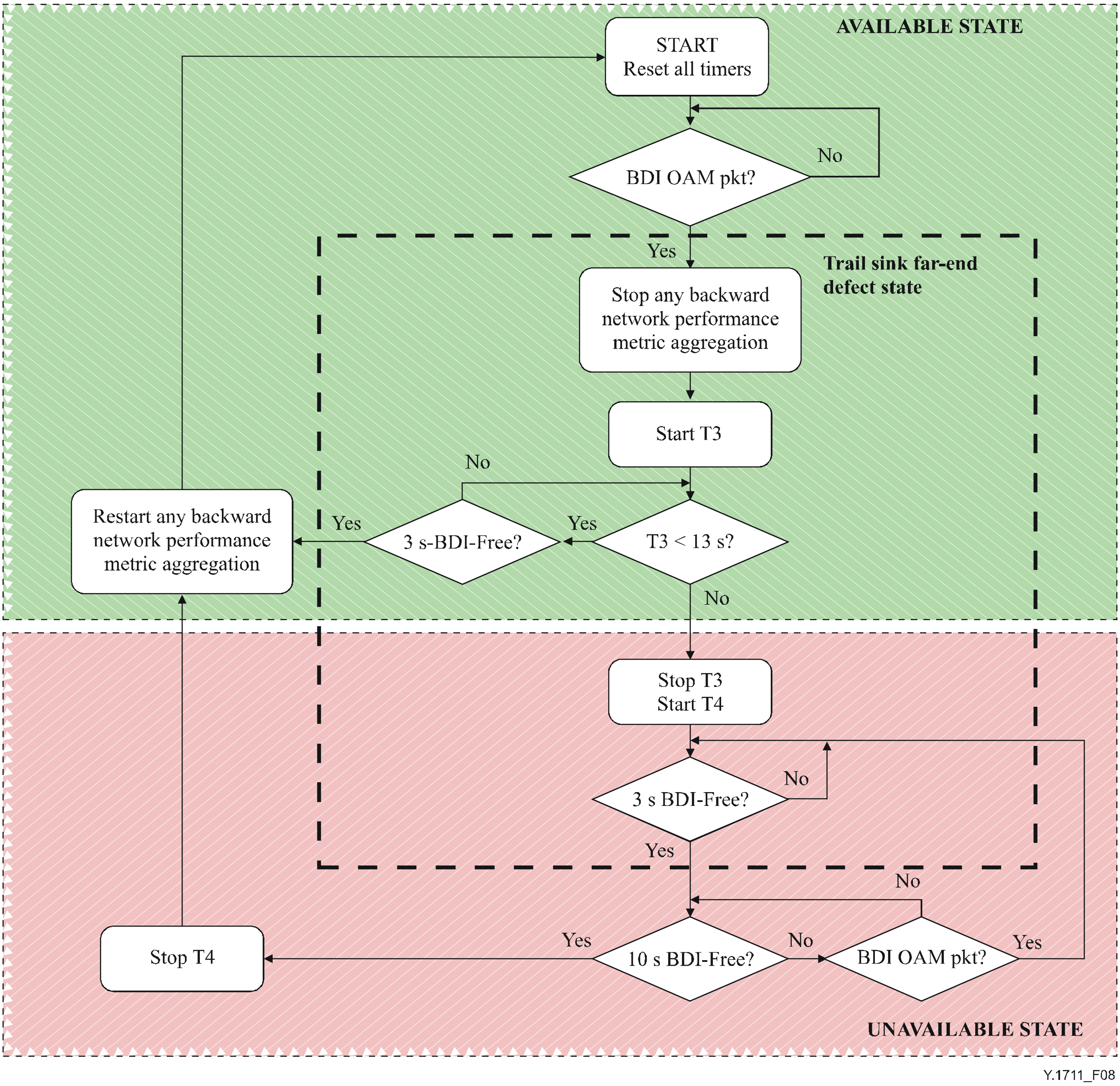


Figure 8/Y.1711 – LSP far-end state processing algorithm

NOTE – The diagram is not meant to imply a physical realization (this is down to the vendor to interpret), but rather the logical flow of processing and consequent actions at key points. Further, for clarity, it is not possible to provide all the details in such a diagram and so one needs to refer to the accompanying text.

1) Assume the endpoint start in the available state at the box marked "Start". All timers shown later in the flow chart can conceptually be assumed to be reset at this point. If there is any backward network performance aggregation activated on the return direction LSP, then this will be via a separate P OAM packet flow on the return LSP.

2) The first decision box is "BDI OAM packet?". If the answer is "No", then it keeps looping this check condition and stays in the Available State. If the answer is "Yes", then this implies that the near-end processing at the other end of the (outgoing) LSP has entered the Trail Sink Near-End Defect State. Note that this also implies that the defect has already existed for 3 s at the other end of this LSP.

3) The endpoint then enters the Trail Sink Far-End Defect State and any backward network performance metric aggregation is prohibited. The network performance registers will need to be corrected for the previous 3 s which should not be aggregated into the long-term Available State counts.

4) Then timer T3 is started, and it runs round the loop composed of the decision boxes "T3 < 13 s?" and "3 s BDI-Free?". T3 is used to check the duration of the Trail Sink Far-End Defect State. If T3 does not reach 13 s and the endpoint gets 3 s which are BDI-Free, then any performance measurement aggregation is restarted. Note that the last 6 s must be accounted for in any backward network performance metric aggregation registers. This arises since it takes the near-end processing 3 s to declare the end of the defect at the other end of the (outgoing) LSP, and a further 3 s to declare the end of the Trail Sink Far-End Defect State at this end of the (return) LSP, and all this time should count towards the Available State at this end of the LSP to ensure correct network performance metric aggregation. A Short-Break date/time stamped event should also be recorded in the local registers, together with the DL/DT information of the defect as given in the BDI OAM packet. This Short-Break event must be date/timestamped relative to 3 s before the time at which the first BDI OAM packet was observed. The endpoint goes back to the initial start position. If, however, T3 reaches 13 s, the endpoint enters the far-end Unavailable State. Note that it is not possible to enter the Unavailable State unless the Trail Sink Far-End Defect State has effectively persisted for at least 13 s (which means that, at the other end of the (outgoing) LSP, the Trail Sink Near-End Defect State has persisted for at least 10 s) in available time.

5) Optionally, a date/timestamped unavailability entry exception report may then be sent to the NMS which includes the relevant BDI OAM packet DL/DT information. Note that the date/timestamp of any such exception report should be backdated by 16 s (i.e., 3 s prior to the first BDI OAM packet being observed for this event) to align the far-end processing with that of the near-end processing at the other end. Timer T3 is then stopped and a timer T4 is started, whose purpose is to record the duration of this unavailability event. Note that when the endpoint enters the Unavailable State, it also remains in the Trail Sink Far-End Defect State.

6) It then runs round a loop that checks for 3 s which are BDI-Free. This is used to take the endpoint out of the Trail Sink Far-End Defect State. Note that this is not strictly necessary, and this check condition could have been omitted and the following one, which checks for a continuous (i.e. overall) 10 s of BDI-Free behaviour, could just have been shown. However, it has been shown like this to harmonize the "look" of the near-end and far-end Trail Sink Defect State processing.

7) If the endpoint observes 3 s which are BDI-Free, then it exits the Trail Sink Far-End Defect State and runs a loop which checks if the endpoint has had an overall continuous period of 10 s which are BDI-Free. If any further BDI OAM packets appear within this overall 10 s checking period, it then re-enters the Trail Sink Far-End Defect State and needs to repeat the process from step 6) above. If, however, no further BDI OAM packets appear within the 10 s checking period, it exits the far-end Unavailable State.

8) Timer T4 is stopped and the duration of the unavailability event is recorded. T4 will record a time which is 3 s less than the true unavailability event. A date/time stamped unavailability exit event, backdated 13 s, together with the unavailability duration, should then be recorded in the local registers. Optionally, this information may also be sent to the NMS as an exception report.

9) Any backward network performance metric aggregation can then be restarted, noting that the last 13 s belong to available time and so the aggregate registers should be corrected accordingly.

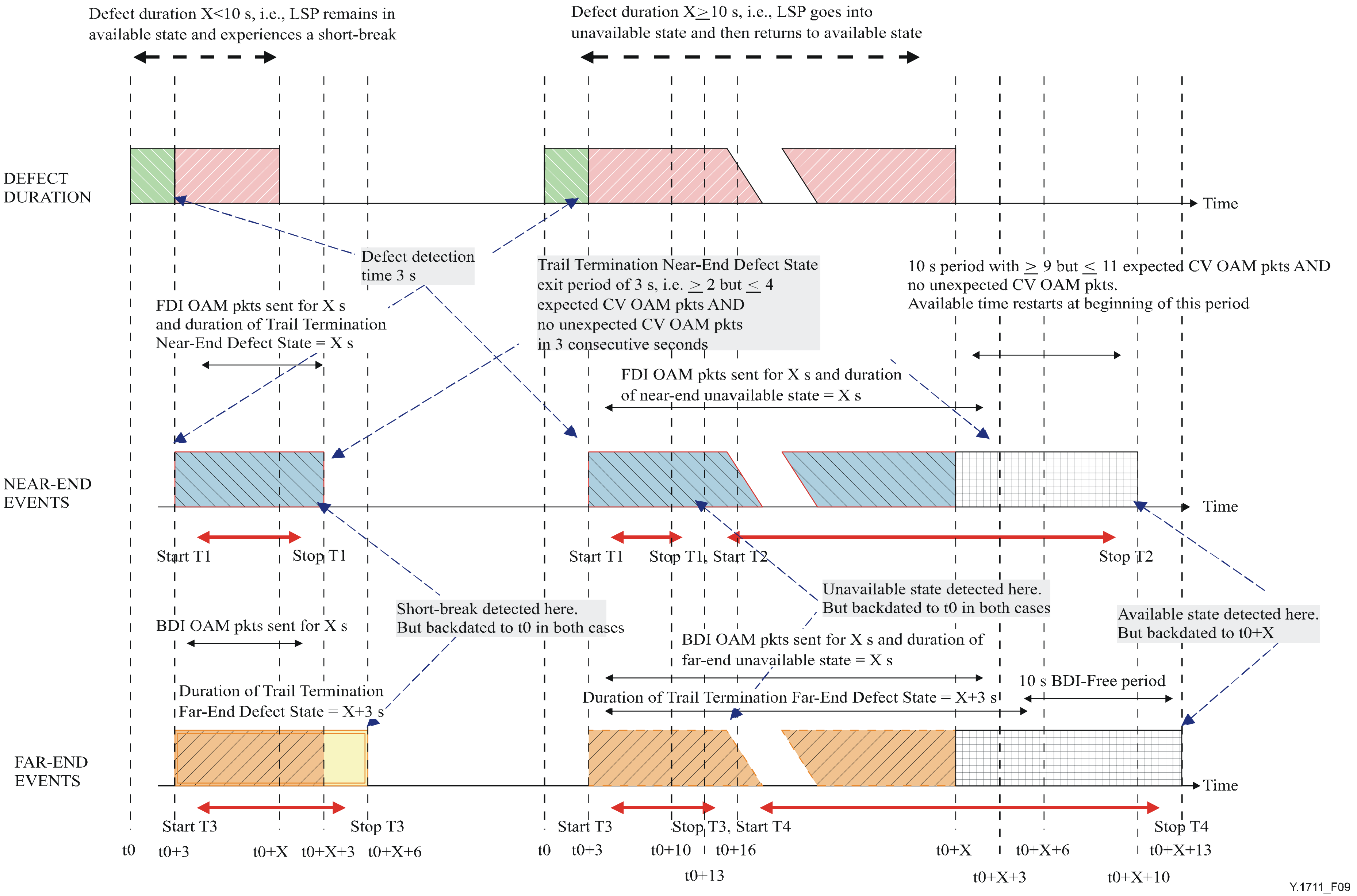
7.6 A pictorial view of near-end and far-end state processing for a short-break and an unavailability event

Figure 9 is given to help clarify the temporal relationships between the near-end and far-end state processing given in the previous flow-charts for short-break event and an unavailability event.

8 Security aspects

This Recommendation does not raise any security issues that are not already present in either the MPLS architecture or in the architecture of its client layer protocols.

OAM functions could enhance the security of MPLS networks. For example, Connectivity Verification (CV) functions defined in this Recommendation can detect misconnections, and therefore can prevent customers' traffic being exposed to other customers.



Defect duration X<10 s, i.e., LSP remains in available state and experiences a short-break
Defect duration X≥10 s, i.e., LSP goes into unavailable state and then returns to available state
DEFECT DURATION
Time
Defect detection time 3 s
Trail Termination Near-End Defect State exit period of 3 s, i.e. ≥ 2 but ≤ 4 expected CV OAM pkts AND no unexpected CV OAM pkts in 3 consecutive seconds
10 s period with ≥ 9 but ≤ 11 expected CV OAM pkts AND no unexpected CV OAM pkts.
Available time restarts at beginning of this period
FDI OAM pkts sent for X s and duration of Trail Termination Near-End Defect State = X s
FDI OAM pkts sent for X s and duration of near-end unavailable state = X s
NEAR-END EVENTS
Time
Start T1
Stop T1
Start T1
Stop T1, Start T2
Stop T2
Unavailable state detected here. But backdated to t0 in both cases
Short-break detected here. But backdated to t0 in both cases
Available state detected here. But backdated to t0+X
BDI OAM pkts sent for X s
BDI OAM pkts sent for X s and duration of far-end unavailable state = X s
10 s BDI-Free period
Duration of Trail Termination Far-End Defect State = X+3 s
Duration of Trail Termination Far-End Defect State = X+3 s
FAR-END EVENTS
Time
Start T3
Stop T3
Start T3
Stop T3, Start T4
Stop T4
t0
t0+3
t0+X
t0+X+3
t0+X+6
t0
t0+3
t0+10
t0+16
t0+13
t0+X
t0+X+6
t0+X+13
t0+X+3
t0+X+10
Y.1711_F09

Figure 9/Y.1711 – Near-end and far-end temporal processing of a short-break and an unavailability event

Appendix I

CV source and sink processing

CV source generation and CV sink processing should be considered as independent functions. This functional decoupling allows operators the flexibility to use different degrees of LSP monitoring on a per LSP basis, e.g., say between those LSPs deemed as "important" and those LSPs deemed "less-important".

CV generation is a relatively trivial function (since it never varies) and is much simpler than CV sink processing. Hence, CV generation could be enabled on all (or most) of the LSPs, but the sink processing could be decomposed into several "degree classes" per LSP such as:

1) No CV processing. Hence no defect processing, no availability measurements and no network performance measurements.

2) A simple check of CV arrivals without examining the TTSI (though it is assumed the TTSI is still generated). This cannot provide totally reliable connectivity verification since it cannot detect certain defects, e.g., d-Mismerge/d-Mismatch.

3) Only a very simple check for arrival of CV packets with an unexpected TTSI. This could be used on less important LSPs as a simple method for detecting leakage of important LSP traffic (into the less important LSP). However, there might be no other defect processing done (e.g., dLOCV) and no availability measurements.

4) Full defect processing but no availability measurements. Note that if availability measurements are not being done, then network performance measurements are also strictly not possible (since these should only relate to when the LSP is in the available state).

5) Full defect processing and availability measurements (this then also allows the option of network performance measurements too).

Appendix II

Indexing LSP Availability State Machine (ASM)

Based on this Recommendation, each LSP requires a state machine at its terminating LSRs (both ingress and egress LSR). These state machines need to be indexed, based on the information contained in the incoming OAM packets. The following two methods can be used at the terminating LSRs to index the LSP availability state machine (ASM):

- Label;
- TTSI.

There are advantages and disadvantages associated with each of these methods.

Label

The user label that begins the OAM label could be used as an index to ASM. This is the simplest index that could be used and has the advantage of being a relatively small number (20 bits). However, if the user label is used as an index, then mp2p, PHP, out-of-band return paths and shared return path LSPs could not be supported. The reason is that in the case of mp2p, the merge of LSPs will be viewed as dExcess defect; in the case of PHP, the user-label does not exist when the packet arrives at egress LSR; in the case of out-of-band return path, there is no user-label and in the case of shared return path LSP, the user-label of the return path cannot uniquely identify the forward LSP. Also, certain types of mismerging defects, such as leaking of an LSP to another LSP that terminates in the same interface cannot be detected.

TTSI

The TTSI included in the OAM packet could be used as an index to ASM. Although not as simple as using the user label as an index, this method has the advantage of supporting mp2p, PHP, out-of-band return path and shared return path LSPs. The reason is that, in the case of mp2p LSPs, each LSP will have its own TTSI and therefore mp2p LSPs are effectively treated as a number of p2p LSPs from the MPLS OAM point of view; in the case of PHP, the TTSI is not stripped off at the penultimate node; in the case of out-of-band return path, the TTSI still exists in the OAM packet, and in case of shared return LSPs, the TTSI can uniquely identify the LSP that the OAM packet belongs to. However, the TTSI is 20-bytes long, which requires compression/hashing, before being used as an index.

When a hierarchy of multiple PHP'd LSPs terminate at a common LSR (multiple labels popped off the label stack upstream of the terminating LSR), some additional ambiguity exists in that the source of OAM messaging is lost in the simultaneous termination of multiple layers. This does not affect CV processing if the TTSI is used as a state association mechanism, however FDI PDUs must be silently discarded.

NOTE – When IPv4 address is used, TTSI value is effectively 8 bytes.

The choice between these two indexes depends on the LSP type, processing power, whether PHP is being used and the type of reverse path. For explicitly routed p2p LSPs, with no PHP, indexing on label may be sufficient. However, in the case of mp2p LSPs (with or without PHP), or p2p LSPs with PHP, indexing on TTSI is more appropriate.

Appendix III

Different possible defect scenarios when using FFD OAM

FFD OAM packets will be provisioned on certain LSPs on which fast failure detection is desired. Note that the rate of insertion of FFD OAM packets will be higher (faster) than the CV-OAM packets (impact of slower rates on defect exit criteria is for further study). This means that if *x* is the insertion interval for FFD OAM packets then *x* is less than 1-s. Note 1-s is the insertion interval for CV OAM packets. Note that the rate of insertion (or insertion interval) on all LSPs requiring fast failure detection in the network is the same.

Given these assumptions, the following are the possible failure scenarios:

i) Connectivity of an LSP provisioned with FFD OAM is broken

In this case, no FFD OAM packets are observed in 3*x* period, where *x* is the insertion interval for FFD OAM packets. This will prompt a signal fail condition and the LSP will enter the Trail Sink Near-End Defect State by declaring dLOCV defect condition.

ii) LSP provisioned with FFD OAM (say LSP-1) gets misconfigured with another LSP (say LSP-2)

Here there are two cases:

1) LSP-2 is also configured with FFD OAM; and

2) LSP-2 is not configured with FFD OAM.

In the first case, LSP-1 and LSP-2 will not observe any FFD OAM packets with the correct TTSI in their respective 3*x* intervals. However, they will observe at least one FFD OAM packet coming in from the other LSP that is with the wrong TTSI during their respective 3*x* intervals. This will result in a dTTSI_Mismatch condition on each LSP. Each LSP will enter the Trail Sink Near-End Defect State by declaring dTTSI_Mismatch defect condition.

In the second case, LSP-1 will not observe any of its FFD OAM packets in a 3*x* interval. However, during this 3*x* interval, it may or may not observe any CV OAM packet from LSP-2. Based on that the LSP-1 will either declare a dLOCV (if it did not receive any CV OAM packet during the 3*x* interval from LSP-2) or dTTSI_Mismatch (if it receives a CV OAM packet during the 3*x* interval from LSP-2). Thus LSP-1 will enter the Trail Sink Near-End Defect State by declaring either dLOCV or dTTSI_Mismatch defect condition.

In the second case, LSP-2 will not observe any of its CV OAM packets in any 3-second interval while the defect persists. However, it will observe at least one FFD OAM packet with a wrong TTSI. Therefore, LSP-2 will enter the Trail Sink Near-End Defect State by declaring a dTTSI_Mismatch condition.

iii) LSP provisioned with FFD OAM (say LSP-1) gets misbranched into another LSP-2

In the simplest case of misbranching, FFD OAM packets from LSP-1 get transferred onto LSP-2 (without any unintended replication); then LSP-1 will not observe any FFD OAM packets in a 3*x* interval. This will prompt the signal fail condition and the LSP-1 will enter the Trail Sink Near-End Defect State by declaring dLOCV defect condition. Note that it can possibly happen that LSP-1 sees its own packets (thus not observe any defect) but LSP-2 observes packets from LSP-1 due to some unintended replication in the network.

For LSP-2 there can be two cases:

1) LSP-2 is also configured with FFD OAM packets; and

2) LSP-2 is not configured with FFD OAM packets.

In the first case, LSP-2 will observe FFD OAM packets with correct and wrong TTSI in a 3*x* time interval. This will prompt LSP-2 to detect misbranching and enter the Trail Sink Near-End Defect State declaring dTTSI_Mismerge defect condition.

In the second case, LSP-2 will observe that in a 3-s time interval it is receiving CV OAM packets with correct TTSI as well as FFD OAM packets with wrong TTSI. This will also prompt LSP-2 to detect misbranching and enter the Trail Sink Near-End Defect State declaring dTTSI_Mismerge defect condition.

iv) LSP without fast failure detection (say LSP-2) gets misbranched into an LSP with FFD OAM (say LSP-1)

LSP-1 will observe along with its FFD OAM packets a CV OAM packet from LSP-2. (Note that the CV OAM packet will be observed but may not be within the 3*x* second interval after the misbranching occurrence.) This will prompt LSP-1 to detect misbranching and enter the Trail Sink Near-End Defect State declaring dTTSI_Mismerge defect condition.

BIBLIOGRAPHY

ALLAN *et al*: Framework for MPLS user-plane OAM, *draft-allan-mpls-oam-frwk-02.txt*, April 2002.

NADEAU THOMAS *et al*: OAM Requirements for MPLS Networks, *draft-ietf-mpls-oam-requirements-01.txt*, June 2003.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series B	Means of expression: definitions, symbols, classification
Series C	General telecommunication statistics
Series D	General tariff principles
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Construction, installation and protection of cables and other elements of outside plant
Series M	TMN and network maintenance: international transmission systems, telephone circuits, telegraphy, facsimile and leased circuits
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks and open system communications
Series Y	**Global information infrastructure, Internet protocol aspects and Next Generation Networks**
Series Z	Languages and general software aspects for telecommunication systems

Geneva, 2004