ISO/IEC JTC 1/SC6 N 7531

Project: JTC1.06.41.04
Date: 1992-07-21

ISO/IEC JTC 1/SC6
TELECOMMUNICATIONS AND INFORMATION EXCHANGE BETWEEN SYSTEMS

Secretariat: U.S.A. (ANSI)

Title: Defect Report ISO 10589/001 for Circulation and Ballot

Source: SC6/WG2

Address reply to:
Secretariat ISO/IEC JTC 1/SC 6 – American National Standards Institute, 11 West 42nd Street, New York, NY 10036

Tel: 212 642-4931; TX 42 42 96 ANSI UI; FAX 212 302-1286

DEFECT REPORT

The submitter of a defect report shall complete items 2 to 4 and 7 to 10 and, optionally, item 11 and shall send the form to the convener or secretariat of the WG with which the relevant editor's group is associated. The WG convener or secretariat shall complete items 1,5 and 6.

1 Defect Report Number: **10589/001**

2 Submitter: SC6/WG2 Routeing Editing Group

3 Addressed to: JTC 1/SC 6/WG 2
editor's group on ISO10589

4 WG secretariat: S. Besbrode

5 Date circulated by WG secretariat:

6 Deadline on response from editor: 1

7 Defect Report concerning (number and title of International

ISO 10589 - Intermediate system to intermediate system intra-domain Routeing Information Exchange Protocol for use in conjunction with the protocol for providing the connectionless network service (ISO 8473)

8 Qualifier (e.g. error, omission, clarification required):

Error

9 References in document (e.g. page, clause, figure and/or tab

See attached

10 Nature of defect (complete, concise explanation of the perce

This is a consolidated defect covering all errors discovered in ISO 10589 since the final text was submitted to ITTF for publication. They reflect errors discovered by implementers of the protocol, errors uncovered by the editor, and national body contributions to the San Diego meeting of SC6/WG2 in July 1992.

The attached document reflects the definition of these errors and the dispositions initally agreed to by the Routeing editors group of SC6/WG2 as the result of discussion and analysis at their meeting

11 Solution proposed by the submitter (optional):

See Attached

12 Editor's response (any material proposed for processing as a technical corrigendum to, an amendment to, or a commentary on the International Standard or DIS final text is attached separately to this completed report)

See attached.

Errors in ISO 10589 and proposed corrections through Defect Reporting Procedures

1 Use of partitionAreaAddresses versus areaAddresses

In a few places in the standard, references to partitionAreaAddresses should be to areaAddresses or vice versa. this confusion arises from the fact that the partition repair capability is optional in the standard and the interpretation of what partitionAreaAddresses constitutes for a system that does not implement partition repair. For instance:

- Clause 7.3.9 (Generation of L2 LSPs (non-pseudonode) states “In the Area Addresses option - the set of areaAddresses for this intermediate system as described in 7.2.11”

For partition repair to work correctly, this clause *should* refer to partitionAreaAddresses and not areaAddresses.

The solution to this problem is to correct 7.3.9 and to clearly state the relationship between partitionAreaAddresses and areaAddresses in clause 7.2.10.3 (Computation of Partition Area Addresses). At the beginning of this clause the following paragraph should be added

> “For systems which do not implement partition repair , the value of partitionAreaAddresses is identical to the value computed for areaAddresses as described in 7.2.11. For systems which do implement partition repair, the value of partitionAreaAddresses is determined as follows.”

2 Receipt of Partial Sequence Numbers PDUs

There is a discrepancy between the detailed description of 7.3.15.2 - Action on Receipt of a Sequence number PDU and 7.3.17 Making the update reliable. Towards the end of 7.3.17 (just above note 34) it explains that although the PSNP is multicast, only the designated IS should respond to it. However, 7.3.15.2 does not say that non Designated ISs should ignore PSNPs on broadcast circuits. It would be clearer if this were done.

In 7.3.15.2, list item a), insert a new sub-item 2) after sub-item 1) and renumber the following sub-items:

> “If this is a level 2 PSNP, and this level 2 IS is not the designated IS for the circuit over which the PSNP was received, then the IS shall ignore the PDU”.

3 Partition Designated IS reporting in L2 LSPs

Clause 9.9, Level 2 Link State PDUs states that the Partition Designated Level 2 IS (PDL2IS) field shall be present in all non-pseudonode LSPs with LSP number zero. If no L2 IS in an area is ATTACHED the procedure for selecting a PDL2IS clearly fails (and are meant to fail!). In this case it is not obvious from the text what to do. The correct behavior would be too leave out the field in systems whose decision process has not successfully elected a PDL2IS. This, however, is conflict with the text which says that the field should unconditionally be present.

Modify the description of the Partition Designated Level 2 IS in clause 9.9, page 58, second sentence, to read:

> "For non-pseudonode LSPs issued by intermediate systems which support the partition repair optional function and which are currently ATTACHED, this option shall always be present..."

4 Conflict in the handling of the QoS option for encapsulation

Clause 7.4.3.2 contains text which is largely duplicated in 7.2.10.4. The requirements on the QoS-parameter of the outer header are however in open conflict: 7.2.10.4 says copy from inner header when present else set to default, 7.4.3.2 says set unconditionally to default.

To fix this conflict, modify the the text of item d) 1) in clause 7.2.10.4 by adding the following sentence after the first sentence:

> "Change the field to indicate routeing by the Default routeing metric if necessary."

5 Serious error in the partition repair function

There is a serious error in the description of partition repair, which essentially prevents partition repair from working. This error crept in due to two oversights:

- When partition repair was made optional, insufficient care was taken in looking at the decision process to make sure that systems which did not implement partition repair still were required to do all the functions necessary to allow proper routeing in the presence of systems which did implement partition repair

- Some essential computations expressed in the decision algorithm in Annex c (non-normative) were never clearly spelled out in normative text.

The essence of the problem can be described as follows:

Let's say area FOO is partitioned. R1, R2, and R3 are level 2 ISs. R1 and R2 are in one partition of FOO, R3 is in the other. R1 and R3 are the partition designated ISs The level 2 path from R1 to R3 travels past R2. The problem is how does R2 forward packets. If it gets a packet for R3, does it route via level 2 (which sends it towards R3), or via level 1 (which routes it via level 1, because R1 claims a "virtual level 1 path" to R3. Clearly the latter strategy (routing via level 1) won't work.

One reasonable way to correct this problem is to observe that destinations reachable via level 1, but via virtual links, are suspect — metrics get truncated, for instance. A level 2 IS (Unfortunately ALL level 2 ISs, not just ones that do partition repair), must first calculate its level 1 database, marking entries reachable only via virtual links (they do this anyway in order to calculate the partition designated level 2 IS). Then they compute their level 2 database. For NETs that are reachable via the level 2 net, and are in their area (which will be the level 2 ISs in their area), they should overwrite any entries in the level 1 forwarding database if the path is via virtual links. In other words, R2 will notice that R3 is reachable via level 1, but only via virtual links. When R2 calculates its level 2 forwarding database, it notices R3 is in its area and is reachable via level 2. Then R2 changes its level 1 forwarding database for R3 to route via level 2. This means that if R2 happens to get a packet addressed to R3, it will not travel via R1, and will not get encapsulated — it will just get forwarded by level 2.

In order to avoid expressing the solution as a specific implementation mechanism, the standard should be ammended to make clear that routes traversing "real" adjacencies are preferred over routes which traverse a virtual link. Make the following changes to the standard:

Add a new subclause between 7.2.12.2 and 7.2.12.3 as follows:

"**7.2.12.x** If an Intermediate system takes part in level 2 routeing and the IS determines (by looking at the area address) that the destination is reachable within its area, then the destination will be reached either by level 1 or level 2 routeing, as follows:

a) Level 1 routeing is always based on internal metrics.

b) A Level 2 route, if it exists, shall be preferred over a level 1 route that traverses a virtual link. Otherwise, the level 1 route shall be preferred.

c) [copy from item b) of 7.2.12.1]

d) [copy from item c) of 7.2.12.1]"

6 Initialization of pt-pt Links

There seems to be a problem with the 10589 description of the initialization of pt-pt circuits in clause 8.2. In ISO 9542 the IS starts off by sending an ISH PDU and only when the IS has detected that the other end is a IS (by receipt of its ISH PDU) should the IS start sending IIH PDUs. In 10589 there seems to be an anomaly. 8.2.2. says what you do when you receive an ISH, but nothing tells you to send them. 8.2.3 tells you to send IIH PDUs all the time. It looks as if some attempt has been made to remove the initial ISH exchange step, but it hasn't been completed. The standard needs to specify that upon circuit initialization the IS sends an ISH PDU, and sends IIH PDUs in response to receipt of ISH PDUs.

Make the following modifications to the text of ISO 10589:

Insert a new clause between clauses 8.2.2 and 8.2.3 (renumbering 8.2.3 and subsequent clauses) to state that the IS must send ISH PDUs:

"**8.2.3 Sending ISH PDUs by an Intermediate system**

An Intermediate system shall cause ISO 9542 to send an ISH PDU whenever a point-to-point circuit is first enabled."

Replace the text of item a) following the first paragraph of clause 8.2.3 (now renumbered to 8.2.4) with the folowing:

"a) the IS receives an ISH PDU"

Add the following paragraph immediately after item b):

"The iSISHelloTimer shall be (re)started upon transmission of the IIH PDU."

7 LSPs with no Adjacencies

The case is when you have multiple LSPs for a node and one of them no longer has any adjacencies to report. What are you supposed to do. Clearly you have to do something to indicate that the adjacency is no longer relevant. You *could* send out the LSP as normal but with no options. However, this would take up space in the database; permanently if you continued to regenerate the empty LSP, or for at least 20 minutes if you didn't regenerate it subsequently and just let it die.

However, the correct thing to do is to purge the original LSP. Indeed, this is what would happen by default if you just removed the original LSP from you own database, since receipt of a CSNP mentioning the old LSP would cause a purge to be issued. However, you don't really want to have to rely on CSNP's (you may be on a pt-pt link).

We should add some words which say that when an LSP is no longer required it should be purged. We currently have words which say this for a no longer required pseudonode LSP, but they don't mention the possibility that any old LSP could become redundant.

Solution is to add a clause 7.3.4.6 which says:

> "If an LSP becomes empty because all the adjacencies reported in that LSP no longer exist, an IS may purge that LSP instead of re-issuing it as an empty LSP (i.e. with no options) or no reissuing it at all and allowing the LSP to time out in the other ISs. This saves space in all of the LSP databases and hence is highly recommended."

8 Confusing wording in the description of path selection by the Decision process

There is potential confusion over the interpretation of the IS preference section of 10589 (7.2.12.3).

It is not clear what is meant by the term "the next hop's area address", and the next hop NET. This was rewording to meet the (I think UK) comment into Berlin that you couldn't have a route matching an area address. However the replacement text appears to be even more obscure. What it is trying to say is that information from the area address field in an LSP has preference over that from an internal neighbor prefix field, which in turn has preference over an external neighbor prefix. This sort of makes sense, in that it stops you from having mis-routeing within the local domain by incorrect manual (or externally supplied) information, but it is still unclear how it is supposed to work

Suppose you have an area 490001 in your routeing domain which is reachable at cost 10 from a particular IS. Suppose you have at that IS an RA pointing some other direction for 490001 at cost 5. Presumably the RA is ignored. Now suppose you had an RA for 49000102? Is it the intention that that should be ignored as well, even though it is a longer match?

The solution is to express the precedence directly in terms of what information in what LSP implied the availability of a particular route, instead of in terms of the forwarding information derived from that information. Therefore, make the following changes to 7.2.12:

Rewrite sub-items 1-3 of item b as follows:

1) Highest precedence: routes constructed from the Area Addresses information in an LSP (i.e. the path does not go outside the routeing domain)

2) Medium precedence: routes constructed from the Reachable Address Prefix information in an LSP which indicates an internal metric. In the case of multiple prefixes which match a given destination address which all have internal metrics, then the longest prefix shall be preferred.

3) Lowest precedence: routes constructed from the Reachable Address Prefix information in an LSP which indicates an external metric. In the case of multiple prefixes which match a given destination address which all have external metrics, then the longest prefix shall be preferred.

9 Alternative multicast addresses for Token Ring LANs

In 10589 there are two sets of address to be used for allL1ISs, AllL2ISs, AllintermediateSystems and AllEndSystems. The first set are the normal ones and these are required for all except 802.5. The second set are permitted for 802.5 only, but there is a strong recommendation to use the former set even on 802.5. In the 802.5 set there are in fact only two distinct addresses, one for ESs and one for all three functions of ISs.

The problem arises with mixed LANs connected by bridges.

It is relatively straightforward to arrange for bridges to map from the four multicast addresses down to the 2 functional address since that only involves looking at the addresses, but it is very difficult to map back the other way, since the only way one could distinguish between the 3 multicast addresses is to look into the type field of the individual PDUs. This is therefore not considered an option by the bridge implementers.

The OIW proposal is that 10589 be modified to accept all three types of traffic (ES, L1 IS and L2 IS) on the allIS multicast (as well as on the proper L1 and L2 multicasts). Thus the mapping from 802.5 to the real world could just map the IS functional address to the AllIS multicast.

The problems are:

a) ISO 10589 currently has words like "On receipt of a level 1 LAN IIH PDU on the multi-destination address AllL1ISs..." which implies that you don't do anything if you receive it on some other address. They would therefore want to propose a defect report to 10589 to fix this. It's not within the scope of an ISP to make such a change. (Incidentally there doesn't seem to be any similar wording around L1 and L2 LSP and SNP reception.)

b) If this was accepted it would potentially require all IS implementations to change, just to accommodate the possibility that a token ring is bridged to your network.

c) ISO 10589 deliberately used the separate multicasts so that if you wanted to you could have different adapters for receiving ISHs and IIHs, or that you could distinguish different queues based on the reception address. This would break all that.

The EWOS position is that token ring ISs be required to use the proper multicast addresses.

The conclusion of the routeing editors group meeting at the SC6/WG2 meeting is that no action be taken on this item until the functional profile work for ISO 10589 has been harmonized among OIW, EWOS and AOW.

10 Behavior clause for Resetting Timers

When applying resettingTimer-B to the dRISISHelloTimer and iSISHelloTimer errors may be eperienced on circuits with more than one IS attached to it.

The nature of this error is that other ISs will lose their adjacency with an ?IS if the value of one of these two timers is increased by more than ISISHoldingMultiplier. The reason for this is that the system in which a timer was increased will immediately start transmitting IIH PDUs at the new (reduced) rate while the previous (low) HoldingTimer is still in progress at the other system(s).

To correct this error, make the following changes to the standard.

Add a new behaviour definition resettingHoldingTimer-B under clause 11.2.1.1 on page 69 which is identical to the resettingTimer-B but with the words:

> "...the specified timer after the...whichever is later..."

replaced with the words:

> "no later than the expiration of the interval in progress or the specified interval, whichever is sooner..."

Replace "resettingTimer-B" with "resettingHoldingTimer-B" in the definitions of the dRISISHelloTimer in clause 11.2.4 on page 77, and the iSISHelloTimer in clause 11.2.5.9 on page 86.

11 Interpretation of the encoding of NSAP addresses

This problem was raised in a UK contribution (2S22) to the San Diego SC6/WG2 meeting.

According to ISO 8348/Add.2 it is possible that when encoding an NSAP address (into NPAI) using the preferred binary encoding scheme, the final octet of NPAI will be padded with an ‘F’ in the least significant semi-octet

For example, if the last three digits using the decimal abstract syntax are:

.......123

then the final pair of binary encoded digits may become

....-12-3F

When considered for forwarding by an IS, the final octet of NPAI is considered to be the SEL. Thus the IS would take the SEL to be 3F and not 23, and the 2 would be interpreted as part of the ID field.

The solution is to make the following changes to clause 7.1.3 of ISO 10589:

Add the following paragraph to the begining of clause 7.1.3 (NPAI of systems within a routeing domain) and followed by a new subclause 7.1.3.1 incorporating new text before the existing text:

> “This clause first defines how the NPAI corresponding to NSAP addresses and Network Entity titles of systems deployed in a routeing domain is constructed and second how the NPAI is structured for use by the protocol.
>
> **7.1.3.1 Construction of NPAI from network addresses**
>
> NPAI is derived from NSAP addresses and NETs accoding to ISO/IEC 8348/Add.2. The NETs and NSAP addresses are obtained from the appropriate addressing authorities.
>
> For these addresses to be correctly interpreted by the protocol in this International Standard it is mandatory that the routeing domain authority ensure that whenever the Network address includes a DSP whose syntax is decimal digits, then:
>
> - if the IDI is in the ISO DCC or E.164 format, the decimal syntax DSP must be an odd number of digits.
> - for the other IDI formats the decimal synta DSP must be an even number of decimal digits.
>
> **7.1.3.2 Structure of the NPAI**
>
> [Insert exsiting text of clause 7.1.3]
>
> NOTE - The SEL field is always the last octet of the NPAI, since the rules enforced in 7.1.3.1 guarantee that there will be no pad at the end of the NPAI.”

Add the following text to the end of the first sentence of clause 7.1.4:

> “...and the rules stated above in clause 7.1.3.1”

Replace the text of NOTE 5 with the following:

> “NOTE 5 To correctly interpret the requirements given below, it is necessary to refer to the structure of the NPAI presented in 7.1.1 and 7.1.3.2, and to the concept of manual area addresses defined in 7.1.5.”

12 Setting attachedFlag on ISs with Reachable Address prefixes

In ISO 10589 a level 2 IS is considered *attached* if the IS determines from its LSP database that it can reach other areas in the routeing domain. The purpose of this mechanism is to optimize the exit of traffic from an area into the level 2 net by ensuring that level 1 ISs do not bother to forward traffic to a level 2 IS if that IS could not get the traffic out of the area.

The method for setting attachedFlag is deficient, however, because it does not deal with the possible presence of Reachable Address prefixes (which indicate routes to another routeing domain) in a level 2 IS that does not have any inter-area connectivity. For example, consider a simple (and useful) topology which consists of a single-area routeing domain with a single level 2 IS having a link to another routeing domain. That IS will have a reachable address prefix for the destinations reachable through the other routeing domain. However, according to the rules the IS is not permitted to declare itself attached since it cannot reach other areas in the RD (indeed, there are no other areas to reach!). Since there is no attached level 2 IS in the area, the level 1 ISs conclude that no destinations outside the area are reachable and black hole traffic instead of sending it to the level 2 Is with the eternal connectivity.

The solution is simply to have a level 2 IS set attachedFlag if *either* it can reach other areas, *or* it has enabled reachable address prefixes.

One consequence of this, however, is that if there in fact 2 level 2 ISs in the area, one of which is the IS described above and the other is an IS which *can* reach other areas, incorrect routeing will result. This is because level 1 will send all traffic to the attached level 2 IS. However, this IS has no way to redirect traffic destined outside the RD to the correct level 2 IS since they have no level 2 path between them.

This is not a new problem however - it is just another example of a partitioned level 2 net, which can occur whenever there is no connected level 2 path among all the level 2 ISs.

The solution can be incorporated into ISO 10589 as follows. In clause 7.2.9.2 on page 16, rewrite the first paragraph as follows:

> "When executing the level 2 decision process for each supported metric, level 2 IS shall ascertain whether or not it can reach any destinations outside its area using that metric. The IS considers itself attached if either:
>
> a) it can reach at least one other area using the corresponding routeing metric, or
>
> b) it has at least one enabled reachable address prefix with the corresponding metric defined.
>
> Otherwise the IS considers itself not attached.
>
> If the IS discovers that it is not attached and attachedFlag was previously True, it shall:"

Rewrite the paragraph immediately below NOTE 13 as follows:

> "If the IS discovers that it is attached and attachedFlag was previously False, it shall:"

13 Minor Editorial defects

13.1 References to the procedure "Forward"

In clauses 7.4.3.2 and 7.4.3.3 there are references to the procedure forward (or its parameters). However, the text of this procedure was removed during IS processing in the basis that this was over-specification of an implementation.

In 7.4.3.2, change the last sentence to read:

> "Then forward the encapsulated PDU as described in 7.4.3.3 below."

In 7.4.3.3, change the first paragraph to read:

> “Choose either the level 1 or level 2 forwarding database, depending on the destination Network Address in the NPDU. From that database, select the adjacency for the next hop to that destination. If forwarding at level 1 for a destination which is not in the area, choose the adjacency for the nearest Level 2 IS computed as described in 7.2.9.1.”

In the first sentence of the last paragraph of 7.4.3.3, replace the words “for the address dest” with “for the destination Network Address”, and delete the end of the sentence following the words “or in round robin fashion)”, but retain the footnote and its reference.

13.2 GDMO Corrections

Clause 11.2.5.7: linkageISISlevel2Broadcast-P should be linkageISISLevel2Broadcast-P to keep in line with capitalization practice.