Control of Duplicate Addresses for FDDI

Jerry Hutchison, Henry Yang

Distributed Systems Architecture and Advanced Development
Digital Equipment Corporation
550 King Street, Littleton, Ma., 01460
(internet: Hutchison @ erlang . enet . dec . com)

Abstract:

Duplicate station addresses can disrupt operation of a token ring. Failure symptoms include failure to initialize the ring (make a token), false re-initialization of the ring, and the inability of certain stations to communicate with other stations. These problems were recognized and some solutions for simple cases are included in the FDDI standards. Additional cases result from common usage of station addresses which are not solved by the standard algorithms. For example, stations may have multiple individual addresses and may use some addresses as destination addresses but not source addresses. These more complicated fault cases and solutions for them are discussed.

Introduction:

The Fiber Distributed Data Interface (FDDI) provides a high-bandwidth 100 megabits per second, general-purpose interconnection among computers and peripheral equipment. FDDI stations form a Local Area Network (LAN) based on a token-controlled access protocol. Each FDDI station has one or more addresses for identification of protocol entities within the station. These addresses are used to transmit and receive user data for LAN communications and are used for Media Access Control (MAC) protocol purposes.[1] An FDDI station is required to have an unique, individual address, called My Long Address (MLA), to identify the MAC protocol entity and the Station Management (SMT) entity.[2] Additionally, a station may transmit frames with many other addresses for clients of the MAC sublayer. This paper discusses the affect of duplicate addresses when a station has a single address (MLA) and the more general case where a station has many addresses.

Station addresses are used by the MAC to initialize and control the ring. The operation of the FDDI LAN and other similar LANs require station addresses to be unique, there cannot be duplicate use of the same address by more than one station or protocol entity. When the uniqueness requirement is violated, the FDDI LAN may incur severe failures ranging from deadlock (lack of communication service for the entire LAN) to lack of communication between a set of stations and performance degradation of the ring.

Station addresses may not be unique due to mistakes and faults. Many addresses are locally administered and are selected by network managers or address assignment protocols which operate on the LAN. A manager may miss-type an address, for instance due to confusion caused by bit order issues, causing a duplicate address for either locally administered or globally administered addresses.[3] Finally, there may be a fault in either hardware or software which causes incorrect entry or matching of an address.

Protocols have been developed to detect and control duplicate addresses. The goal is often to detect duplication so that the related stations may take actions to allow non-duplicate addresses to operate normally. SMT includes protocols which detect some cases of duplication. The SMT protocols and their limitations are discussed in this paper. Additionally, new protocols are described which extend the standard protocols to cover more complicated cases.

Duplicate address detection and control protocols are discussed in two general cases. First, stations are considered to have a single address, MLA. In this case the definition of a duplicate address condition is simply two or more stations using the same value for MLA. Duplicate MLA problems are described in detail to show how the ring may fail to initialize (reach deadlock) and may falsely re-initialize (possibly quite often). Solutions to these problems are also described. This first case is of interest as it illustrates the basic problems and represents the scope of solutions currently specified in the ANSI FDDI SMT standard.

The second general case considers stations which have many addresses, not only a MLA, and may perform some address related functions without matching to an address in a frame. A station may have a list of addresses or may have separate lists of addresses used for matching against the destination and source addresses in a frame. When a station performs an exact

Paper to be published in IEEE 16th Conference on Local Computer Networks, October, 1991

address match, the address is contained in at least one of the station's lists. A station may send and receive frames using an address not in any of these lists as address matching is not always performed. For this case we adopt a new definition of a duplicate address.

This second case represents duplicate address problems which are not resolved by ANSI FDDI standards. Solutions described are suitable for stations using many addresses or which operate for some addresses without address matching. An algorithm is presented which eliminates missed and false detection possible with the standard algorithms. These solutions provide for more robust detection of duplicate addresses in common situations. Improved detection of a duplicate addresses results in a higher reliability and performance LAN.

Operation of the token protocol:

The problems caused by duplicate addresses are the result of basic ring initialization and control protocols. Normally, access to the ring is controlled by a token and stations exchange information based on addresses. Error detection mechanisms detect absence of a token and invoke protocols which initialize the ring by creating a token. These important aspects of the MAC protocol are discussed below in terms of simple stations with a single address, called MLA. Later sections will illustrate the operational problems of a ring with duplicate addresses.

A token ring (also called ring) consists of a set of stations serially connected by segments of transmission media to form a closed loop. Information is transmitted sequentially, as a stream of symbols, from one active station to the next. Each station generates and repeats symbols. Symbols are used to form frames of information which are exchanged between stations. The frame includes fields for the address of the destination station, the originating or source station, and data to be exchanged between stations.

During normal operation, only one station has the right to transmit (originate) information onto the ring. Transmit opportunities are scheduled by a token. The ring has status called Ring Operational (RingOp) equal to "ON" when frame transmission is scheduled by a token. Information circulates from one station to the next. Stations which are not transmitting their own frames repeat the frames that arrive on the ring. The addressed destination station(s) copies the information as it passes. Finally, the station that transmitted the information removes it from the ring.

Error detection and recovery mechanisms are provided to restore ring operation in the event that transmission errors or medium transients (e.g., those resulting from station insertion or removal) cause the access method to deviate from normal operation (e.g., the loss of the token). Detection and recovery for these cases utilize fault recovery and ring initialization protocol that are distributed among the stations on the ring. In order to detect these transient errors, each FDDI station maintains a Valid Transmission timer (TVX timer) to ensure that a valid frame or token is received within a TVX period, which has a default value of greater than or equal to 2.5 milliseconds. If a valid frame or token is not received within the minimum TVX period, the TVX timer may expire causing the station to initiate fault recovery and ring initialization. The fault recovery protocols, called Beacon and Claim Token protocols, are specified in the FDDI MAC Standard and discussed in more detail later. When the ring is not operational, that is during ring initialization, the state of the ring is represented with RingOp off. After a successful ring initialization, the state of the ring is RingOp on.

One of the invariants of the ring is for each station to strip (i.e., remove) frames transmitted by the station after the frames have traversed the ring exactly once. One of the methods for frame stripping is for each station to continuously compare the Source Address (SA) of each frame from the ring against its MLA. If an address match is found, the rest of the frame is removed from the ring (the rest of the frame after the SA field is stripped). This method is called, "SA Match Stripping".

FDDI requires each station to maintain a MLA and to continually strip frames that has SA matching its MLA. If two or more stations have the same MLA then these stations will continually strip frames from each other. Duplicate address problems result since the error detection and recovery protocols use frames with SA to identify the source station. The stripping by a station with a duplicate address can result in ring deadlock where the ring is not operational and falsely invoke recovery protocols which interrupt service.

MLA - MLA duplicate address problems:

Duplicate addresses may affect the ring in two ways. The first problem occurs while RingOp is off and duplicates block initialization. The second problem occurs with RingOp on. Some duplicate addresses allow the ring to intialize but then cause repeated, falsely invoked re-initialization. Examples of these two types of problems are given in this section. A later section will show the solution to these cases.

While RingOp is off, the normal ring initialization or fault recovery protocol consists of a station transmitting a Claim Token frame or a Beacon frame containing its SA. The station continues to transmit such frames until it either receives its own frame back or receives a Claim Token or Beacon frame of higher

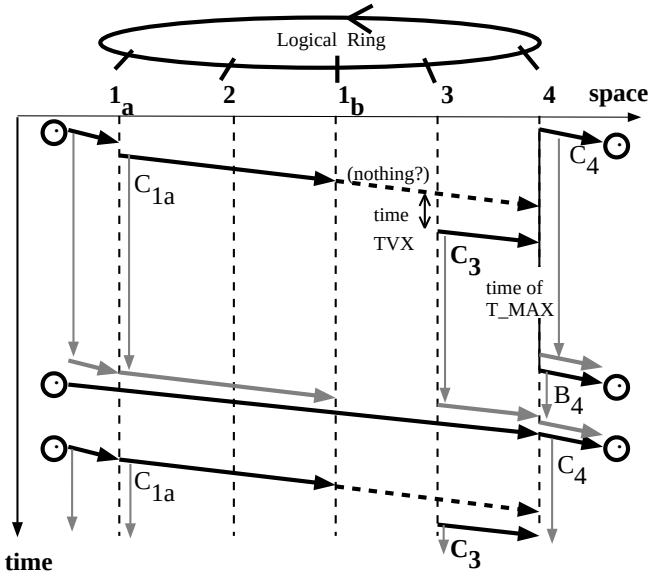
precedence sourced by another station. An example of how initialization fails (without a solution for the duplicate address problem) is shown in Figure 1. In this example, Station 1a is configured to be the claim process "winner". A normal initialization sequence would include station 1a winning the claim process and creating a token. As station 1a is duplicated by station 1b, the process fails to complete.

Figure 1 is a space-time diagram where stations and the distances between stations are shown on the horizontal axis. Time is shown along the vertical axis. Note that the right hand edge of the graph connects to the left hand edge as the topology of the ring is a physical loop. Frames or messages are indicated by arrows in the diagram as they propagate around the ring. Tokens and frames are labeled as to their type, for example a "B" for Beacon or "C" for Claim, and subscripted with the address of the sourcing station. Often, many of a type of frame are sent back-to-back. A wide, black arrow indicates the beginning of transmission of a particular frame while a shaded arrow indicates the last one sent.

In Figure 1, Claim/Beacon oscillation starts with station 4, arbitrarily selected, sending claim frames. Claim frames contain protocol information called the "bid" value which relates to winning the Claim process. Station 1a, strips station 4's frames as, based on the bid value in the frame, station 1a has the winning bid of these two stations. Station 1a then begins to transmit Claim frames which are posited to be the winning bid, overall. Station 2 repeats these Claim frames. Station 1b strips the frames from station 1a based on the value of the SA in the frames. Additionally, station 1b observes the bid received in the Claim frames to be a wining value different than the value bid by station 1b. As a result, station 1b does not win the Claim process and doesn't repeat the winning bid which would allow station 1a to win. The process then remains in this state until failure of the Claim process is detected based on a time-out value, T_max. Station 4 then initiates a Beacon process and transmits beacon frames which propagate around the ring. Beacon frames are higher precedence than Claim frames so Station 4 receives it own Beacon and restarts the Claim process. The process repeats until something breaks the deadlock. Additional modes of failure result of different initial conditions.[4] The solution to these types of cases are standardized in SMT and discussed later in the paper.

Some duplicate address cases may allow the ring initialization to complete. However, a different kind of failure mode can occur where the ring oscillates between ring initialization and TVX time-outs called, RingOp/Claim Oscillation. An example is shown in Figure 2. In this example, Station 4 is posited to have the winning bid in the Claim process such that it is possible duplicate stations will not block ring initialization. Figure 2 begins with the ring initialized, that is RingOp is on, and Station 1a capturing the token.

Duplicate station 1a holds the token and transmits data frames for longer than a TVX interval. These frames are stripped by duplicate station 1b, causing



Assume stations 1a and 1b have duplicate addresses (both have address 1). Stations 2,3,4 have unique addresses.

Values of T_Bid are such that 1a is claim winner, followed by 4, 3, 1b, 2 (in that order)

Figure 1: Claim/Beacon oscillation caused by duplicate MLAs.

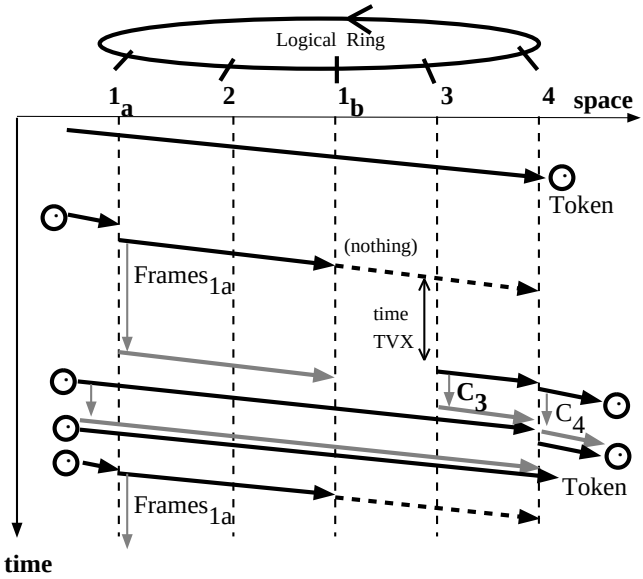


Figure 2: RingOp/Claim oscillation caused by duplicate MLAs.

station 3 to expire its TVX timer and then start a Claim process. The Claim process may then quickly complete (assuming Stations 1a and 1b do not cause a deadlock situation). Figure 2 shows Station 4 winning the Claim process and creating a new token. RingOp would then be on until the token is lost or re-initialization is again falsely triggered. The RingOp/Claim oscillation may occur at various rates depending on the traffic patterns of the involved stations. This problem could be difficult to diagnose without duplicate address detection algorithms.

A final problem caused by duplicate addresses is lack of communication for the stations with addresses which are duplicated. While RingOp is on, the duplicate stations will be able to communicate with some stations and not others. Using the example in Figure 2, it is easy to see that Station 1a can exchange data with Station 2 but not with Station 3. Also, duplicate addresses can cause one-way communication. In the example, Station 1a cannot send to Station 3 but will receive data from Station 3. These communication problems cause the users of particular stations to experience lack of service but do not disrupt service for non-duplicated stations. While detection of the duplicate address fault is within the scope of this paper, the solution to the communication problem for the individual duplicate station is not. The goal is detect the condition and to preserve ring operation for non-duplicate stations. Resolution of the duplicate address is left to an external authority outside the MAC and SMT protocols.

Three types of problems caused by duplicate addresses have been described: 1) failure to initialize the ring which denies service to all users, 2) repeated, falsely invoked re-initialization of the ring which reduces the performance of the LAN, 3) and lack of service to individual stations with a duplicated address. The first two of these are to be detected and resolved by duplicate address control protocols discussed herein. The third problem is detected but resolution is outside the scope of the paper.

Solutions for duplicate MLAs:

In this section, solutions are presented for the ring initialization failure and RingOp/Claim oscillation problems when caused by duplicate MLAs. Currently, SMT defines a Ring Management (RMT) protocol to resolve the MLA to MLA duplicate problem.[2] RMT controls MAC-sublayer operation when the station address is known to be duplicated. The RMT protocol requires that, once a station detects its MLA to be duplicate, that it take actions to make the ring operational for non-duplicate stations. The detection of duplicate MLAs is done by two protocols depending on when detection occurs. Detection while RingOp is off occurs in RMT, itself. Detection of duplicates with RingOp on occurs in the Neighbor Notification (NN) protocol. NN exports its status to RMT using a variable, Dup_Address_Test. After a duplicate is detected, detection of lack of duplicate is done only in the NN protocol. The detection processes in RMT and in NN are discussed below.

Duplicate addresses are detected with RingOp off, before the ring is operational, if they block completion of the ring initialization process. RMT detects a MLA to MLA duplicate problem when one of the following conditions is met[4]:

1. Reception of Beacon frames with SA matching the station's MLA for greater than the maximum ring latency period during which the station has not transmitted any Beacon frame.
2. Reception of Claim Token frames with SA matching the station's MLA for greater than the maximum ring latency period during which the station has not transmitted any Claim Token frame.
3. Reception of Claim Token frames with correct frame check sequence (FCS), with an SA matching the station's MLA but with protocol information in the Claim Token frames which differs from that station's protocol information.

The above conditions must be checked by all stations and can be difficult to observe by a station. For example, conditions 1 and 2 require strict knowledge that a particular type of frame was received but not transmitted within a particular time period. Condition 3 requires that a station copy (receive into memory) and examine the information field of Claim frames. Claim frames may arrive at a rate of about 391,000 frames/second and are not normally copied. The detection of duplicates requires the station to operate the MAC in a mode different from normal operation. As these modes of operation are not in the MAC standard, there was little hardware support for these functions and a software implementation is common. It is important that stations not need to look for duplicate conditions at all times that the ring is not operational. The RMT protocol indicates a station not look for the duplicate conditions until after initialization has failed to complete for 2 seconds. In absence of a fault or a duplicate address, initialization will complete in a fraction of this time. As such, the operation to detect a duplicate is not needed except during faults, allowing implementations to change modes and not incur overhead of the process under normal situations.

The detection of a duplicate is based on the above conditions and the RMT protocol. The actions for detection of duplicates during initialization minimize the overhead on stations needed to detect the above conditions, guarantee mutual detection by two or more duplicates and periodically re-enforces the duplicate status of previously discovered duplicate addresses.[5]

A station discovers itself to have a duplicate address and takes specific actions as outlined below.

The operation of the detection process is given in Figure 3. With RMT present to resolve the duplicate addresses, initialization will complete. The space-time diagram shows the same configuration of stations as Figure 1. As before, Station 1b is blocking the initialization process by stripping the Claim frames of Station 1a while neither station actually completes the Claim process. Based on duplicate condition 3, above, Station 1b recognizes it is duplicated: it has received Claim frames with it's own source address and with incorrect information. Station 1b then transmits beacons unconditionally for a period of time sufficient to cause station 1a to recognize the problem based on duplicate condition 1. This process of sending beacon frames is called a, JAM-beacon process in RMT. Station 1a, upon discovery of duplication, also executes the JAM-beacon process to enforce mutual discovery by all stations in case there were more than two duplicate addresses. Finally, the duplicate stations all yield to other stations and station 4 is able to complete a Claim process and create a token.

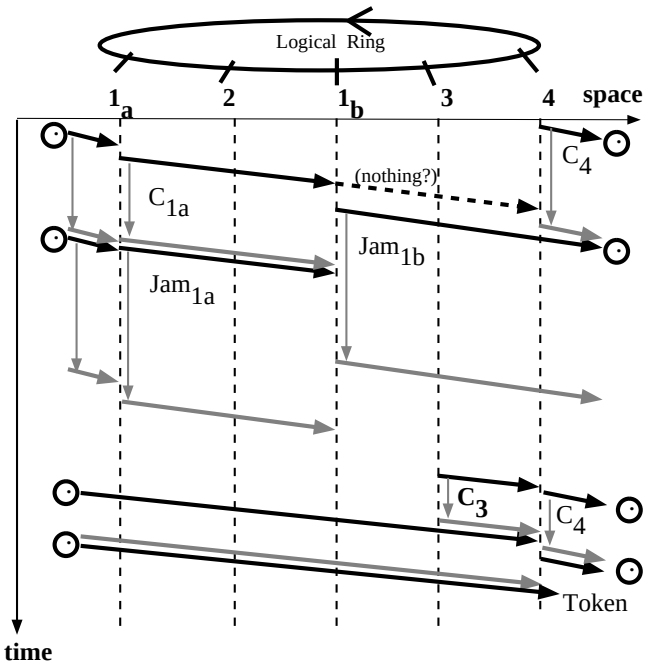


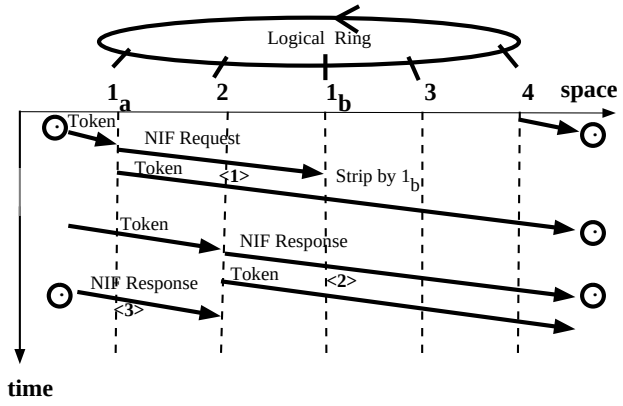
Figure 3: RMT eliminates Claim/Beacon oscillation..

In addition to RMT, the SMT standard defines a protocol to detect the MLA duplicates during RingOp on state. The SMT standard specifies a process to detect the presence and absence of duplicate addresses called, Neighbor Notification (NN). The duplicate address test outcome of the NN process is a variable called, Dup_Address_Test, which is exported to RMT. The NN process defines two tests which are: the NN Request/Response (NRR) test and the Transmit-to-Self (TS) test. They are described in more detail, below.

The NRR test requires that a station periodically transmit a Neighbor Information (NIF) Request frame to its immediate down-stream neighbor. This is accomplished without knowing the neighbor's MLA using the Next Station Addressing facility provided by MAC. After receiving the NIF Request frame, the immediate downstream station transmits a NIF Response frame addressed to the individual address of the requester based on the SA field of the request frame. The NIF response frame contains as a Destination Address (DA) the address of the station running a NNR test. The station completes the NNR test when a response frame is received as it may ascertain if one or more other stations match the DA value.

Duplicates are detected using the NIF response frame using the A indicator in the frame. The A-indicator is part of a frame that is transmitted as reset by the station that originates the frame. If a repeating station recognizes the DA in a frame as its own address, it sets the A-indicator as it repeats the frame, otherwise it repeats the A-indicator as received. When the station performing the NRR test receives a NIF Response frame it examines the A-indicator. If the A-indicator is set then the station declares a duplicate address test failure. If the A-indicator is reset then it declares a success.

As shown in Figure 4, the example of the NRR test for the simple duplicate address problem starts when station 1a captures the token and sends a NIF Request frame. To simplify the illustration, the example shows only station 1a's NRR test operation. For the simple duplicate address problem example, station 1a and station 1b have the same MLA. The NIF Request frame is received and repeated by station 2, and the frame is stripped by station 1b due to the SA match. Some time later, station 2 sends a NIF Response frame addressed to station 1a's MLA, in response to the NIF Request frame. Station 1b repeats the NIF Response frame. Since the frame contains the duplicate MLA in the DA field, station 1b sets the A-indicator as the frame is repeated. The frame with the A-indicator set is then received by station 1a. Thus, station 1a's NRR test detects the duplicate address problem.



Station 1a and 1b have duplicate MLA.

<1> Station 2 receives the NIF Request frame from station 1a.

<2> Station 1b repeats the NIF Response frame and sets the A-indicator.

<3> Station 1a receives the NIF Response frame with the A-indicator set. The NRR Test in station 1a declares a duplicate address failure.

Figure 4: NRR Test and MLA Duplicate Address Case

The TS test may be used instead of NNR to verify a stations' MLA is not duplicated by another station. The TS test requires that a station periodically transmit a NIF Announcement frame with the Destination Address set to the station's MLA. The NIF Announcement frame contains a SMT address and a transaction identifier to ensure uniqueness of the frame. If the station receives its own NIF Announcement frame then it declares that there is no duplicate address problem (a duplicate station would have stripped the frame based on SA match). In absence of a duplicate problem this tests completes relatively quickly.

If there is a duplicate address problem then the TS test result is a "time-out." The NIF Announcement frames are stripped by the duplicate address station based on the SA match. Multiple NIF Announcement frames are sent within a time-out period and at least one of the frames should be received within the time-out period. When no frames are received the station declares a time-out and reports an error event so that appropriate actions can be taken to resolve the problem. Although the TS test time-out does not positively indicate a duplicate address problem, it does identify a serious enough problem that requires corrective actions.

Figure 5 shows how the TS test fails to complete for duplicate station 1a due to a duplicate address problem. The example starts when station 1a captures the token and transmits a NIF Announcement frame. To simplify the illustration the example shows only station 1a's TS test operations. Given that stations 1a and 1b have the same MLA, the frame is stripped by station 1b due to the SA match. Station 1a then periodically transmits the NIF Announcement frame. Station 1b repeatedly strips the NIF Announcement frames. After failing to receive a single NIF Announcement frame during the TS test period, station 1a declares a TS test time-out event.

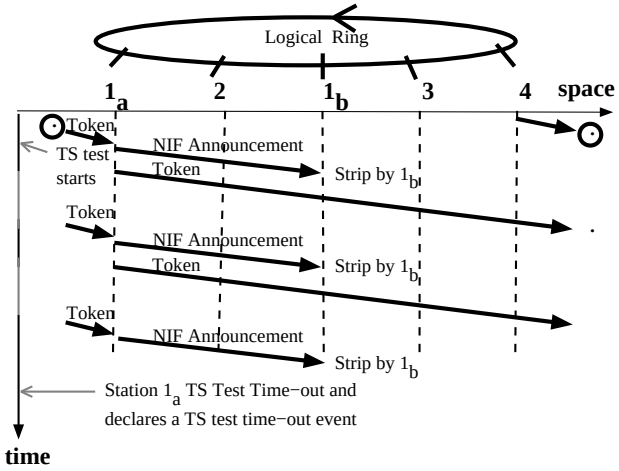


Figure 5: TS Test and MLA Duplicate Address Case

We have shown that the NRR test provides an accurate detection of the MLA duplicate problem. The TS test provides a quick and adequate confirmation of the absence of duplicate address. While these tests are sufficient to detect the MLA duplicate, they are inadequate to cover the more complicated fault cases, which can result when more general addressing schemes are considered.

More complicated addressing schemes:

Previous discussion has focused on the simple cases of duplicate addresses based on stations with a single address called, MLA. A duplicate address was defined as two stations with the same MLA. The duplicate address problem will now be expanded to include the more real-world situations which result of the addressing schemes commonly employed within FDDI implementations. Some stations have need to use many addresses and for some of addresses may not perform exact address matching to the fields of a frame. Also, in the prior discussion it has been assumed that whenever a station did an exact match on an address in the DA field of a frame, it would also have had an exact match for the same address in the

SA field of a frame. In reality, this assumption is not true. As a result, the protocols described previously may fail to detect duplicate addresses. In this section, additional modes of address-related operation and more complex addressing schemes are described. More realistic models for addressing are given. Finally, a new, more concise definition of a duplicate address is adopted.

FDDI stations may have additional addresses besides the MLA. The data link user or MAC sublayer client may use the station's MAC's address, which is the MLA, or may use another address which is then called an alias address. Additionally, a MAC entity may provide the communication service access to ring for other stations and protocol entities external to the station. A popular example of this is a "MAC-layer relay" or bridge.[6] Alias addresses are the addresses used to identify other protocol entities within the station or other stations and protocol entities external to the station. For example, a bridge station has a single MLA and a list of alias addresses, where the alias addresses are addresses of stations and protocol entities that the bridge forwards frames to and from the ring.

Several important modes of address related operation are explained next. Stations using multiple addresses often deal with addresses differently depending on whether or not the address is a DA or a SA in a frame. In the MAC sublayer, the primary problem related to SA's is that of frame stripping while for DA's it is a decision to copy the frame.

A station with several addresses can implement SA-Match Stripping with a Content Addressable Memory (CAM) to contain the set of alias addresses for frame stripping. A content addressable memory contains the set of addresses used by the MAC and does exact matching in real time as the frame is received and repeated.[7] The design of a CAM allows relatively few transistors to be used for each address relative to the number needed to implement a simple comparator. Sometimes a station uses so many source addresses that the CAM would be impractical. Other methods for frame stripping avoid matching addresses completely by keeping track of the frames which are transmitted during each token access opportunity. These schemes involve the use of a count, to count the number of frame transmitted and stripped, and/or the use of a special frame to end the stripping. Within the context of this paper, these other methods will be called "Frame content independent stripping", or FCIS.[8] Using the methods which are not SA Match based, the station immediately starts stripping from the time it begins its transmission. The count and/or the reception of the special frame are mechanisms for the station to determine when to stop stripping. The Non-SA Match Stripping methods can only be used during RingOp-on state. The SA Match Stripping method is used for MLA at all times.

A station receiving frames for many destination addresses may again use a CAM or may replace DA-matching in the MAC-sublayer with a "promiscuous receive" mode. In promiscuous receive mode exact matching to the DA field in a frame in real time is eliminated. All frames which might be received based on the DA address are copied to a packet memory (MAC specifies that some frames are never received). Later, a client outside the MAC-sublayer may examine addresses in packet memory for DA matches. As this is done outside the MAC sublayer and after the frame was repeated, the A indicator is never modified due to a DA match. The hardware complexity of the promiscuous receive mode scales much better for large numbers of addresses than schemes which try to match a DA prior to copying the frame.

Above, the use of alias addresses and the various modes of address related operation are discussed. Next, a model for a station with many addresses used for exact matching is given. First, the model is given, below. Second, it is shown that often the set of addresses used by a station for exact matching need to be kept in two separate lists for use by the MAC. One list contains only the addresses used as source addresses and another list contains only the addresses used as destination addresses. Third, it is seen that the two lists are not identical for many reasons.

A model for station addresses is shown in Figure 6. In this model, the MLA is treated as a separate address used by the MAC and not stored in either list of additional addresses. Alias addresses would be placed in the DA list or the SA list depending on the application. Addresses which are stored in the MLA, DA list and SA lists are those for which exact addresses matches will be performed for protocol steps described earlier such as frame stripping and frame reception with the setting of the A-indicator. Addresses which are stripped using FCIS mode or received using promiscuous mode are not considered to be in these lists. The MLA is the only required address for a station and is assigned special significance in later discussion. Thereby, MLA is shown separately from other addresses in the model.

The second aspect of the model is the need for two separate lists for DA and SA addresses. Examples of addresses which would appear in a DA list which cannot appear in the SA list include multicast addresses. A multicast address identifies a group of stations. As such they are not useful as source addresses (they may cause havoc when used as source addresses). Additionally, a multicast addresses cannot be in a list used to match against the SA field of a frame even if measures are taken to insure MAC never sources a

frame with a multicast address in the SA field. Multicast addresses in the SA list will cause incorrect stripping of frames. The problem arises due to the multiple uses of the multicast bit in the frame format. Stations which operate a protocol called source routing set a bit in the frame format which turns an individual address in the SA field into a multicast address.[9] As a result, if the DA list and SA list were combined or contained all the same addresses, then false address matches cause false stripping of source routed frames. For this reason alone, the DA list must be separate and will often be included in a station.

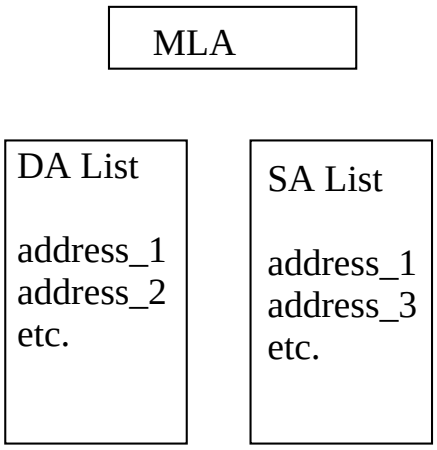


Figure 6: Addressing model showing "lists" and MLA.

Thirdly, it is seen that often the lists do not contain the same individual addresses. Many implementations of FDDI stations result in situations where the an SA list contains individual addresses not in the DA list. The most common application which results in this situation are bridges designed based on VLSI chips which did not offer a FCIS mode of stripping. Said vendors could implement an SA-CAM for stripping purposes while avoiding the cost of a CAM for DA-matching through the promiscuous receive mode.

The appearance of an individual address in the DA list which is not in an SA list may also occur but for different reasons. A station designed based on VLSI chips containing a FCIS support would be very unlikely to include any SA list at all. (The MLA is sufficient for all SA processing). At the same time, said station may contain a DA CAM for exact matching of addresses. Obviously, this would be useful for multicast addresses but individual addresses can appear in the list as well. An application which used the FDDI Frame Status indicators would need to place individual addresses in the DA list. As an additional example, a network management station may be set up to monitor traffic patterns for a set of addresses by loading those individual addresses into a DA CAM.

While these more complicated addressing schemes are common in the industry, there has been little discussion of the detection of duplicate addresses in these cases. Of these new cases, the most important involve Figure 6 and those addresses in addition to MLA for which a station does exact matching for protocol within the MAC sublayer. The RMT protocol does not explicitly solve the duplicate problems involving additional addresses in an address list. Detection of duplicate addresses involving MLA or an address list will be the focus of the remainder of this paper. Algorithms presented later in the paper solves these problems in a simple and elegant way.

First, a better definition of a duplicate address is needed given the address structure described above in Figure 6. A duplicate address was initially described as two stations with the same address for MLA. There are actually seven forms of duplicate address, described by the list in which the address values appear:

1. A station's MLA duplicates a second station's MLA.
2. A station's MLA duplicates a second station's SA-list address.
3. A station's MLA duplicates a second station's DA-list address.
4. a station's SA-list address duplicates a second station's MLA (used as a SA) or a SA-list address.
5. a station's SA-list address duplicates a second station's MLA (used as a DA) or a DA-list address.
6. a station's DA-list address duplicates a second station's MLA (used as a SA) or a SA-list address.
7. a station's DA-list address duplicates a second station's MLA (used as a DA) or a DA-list address.

The original motivation to detect duplicate addresses was to preserve network operation for non-duplicate stations. The actual stations with duplicate addresses need not be made operable but they must be prevented from disrupting the ring. Examination of the situations which caused the network disruptions described in the earlier sections reveals that the primary problem was related to false stripping of another stations frames. As a result, duplicate addresses which are matched against the SA of a frame is the important case to detect to preserve ring operation. The following is adopted as a more exact definition of duplicate addresses to be detected:

1. MLA duplicated by another station's MLA or address in an SA-List.

2. SA-List address duplicated by another station's MLA or an address in an SA-List.

This definition determines when RMT should cause a station to take actions to protect the ring from a duplicate address. As these actions are disruptive to users of the affected station, the scope of the definition is carefully limited. For instance, an address in the DA-list of one station that is duplicated in another stations SA-list is not called an "duplicate address" for the purpose of RMT. This case may affect communication for users of the duplicated address but it should not disrupt the operation of the LAN for non-duplicate addresses. (This condition is of interest to network managers and should be reported as an error).

More complicated RingOp-off case:

The more complicated address structure described above requires additional steps to detect duplicate addresses and to preserve ring operation. The issues will be considered for the RingOp-off case next. In a straight forward fashion, the RMT protocols described earlier can be operated for each source address. The discussion of how RMT is implemented shows this to be impractical for many addresses. A simpler solution is described. Note that a station which has additional addresses in an SA list and which doesn't extend RMT coverage to those addresses can cause a non-operational ring.

A station may be designed to utilize only a particular address during ring initialization as a source address in Claim and Beacon frames. The natural choice for this purpose is MLA, the only required address. Using this convention, a station with addresses in an SA-list can identify presence of a duplicate address during RingOp off as,

1. Detection by RMT protocol of duplicate for the MLA,
2. an SA-list match with the SA of beacon or Claim frame with a correct frame check sequence for that frame.

The conditions for item two would replace the 3 conditions given earlier for RMT and are much simpler to implement. A common implementation of an SA list would use a CAM and the "address match=true" event(s) would be basis for detection of a duplicate address during ring initialization. Once the duplicate situation is detected, a station would need to ascertain which address in the SA list was at fault and take the actions as described in RMT for duplicate addresses. For example, the station could remove the duplicate address from the CAM, report the event to the local management interface, and cause that address not to be used (set the "Dup-ID flag" for that address in RMT) until the the duplicate condition is eliminated.

If RMT is not extended in some way the presence of an SA list can cause serious problems due to SA stripping. It should be noted that elimination of the SA list entirely through use of FCIS mode stripping is a viable way to prevent deadlocks in the initialization process. Note also that RMT must operate for the MLA in either case.

More complicated RingOp-on case:

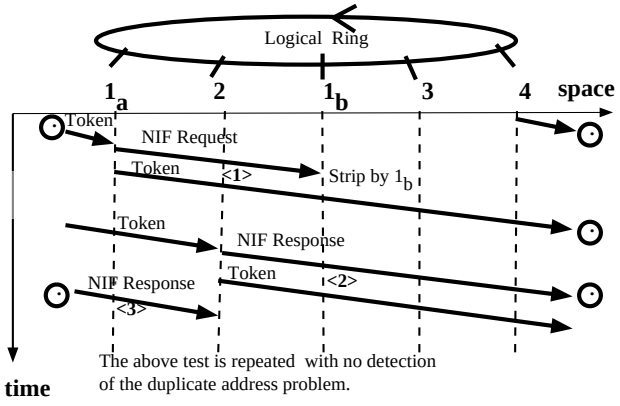
The goals of the duplicate address test during RingOp-on state are to ensure that the ring continues to operate with minimal impact and that the duplicate address is detected and resolved. Failures in duplicate address test can be classified into two general classes. They are: false detection and missed detection.[10] A false detection means that the test concludes that the duplicate address is present when it is not. A missed detection means that the test concludes that the duplicate address is absent when it is present. A false detection is costly to the station as it impacts the operation of the station. A missed detection is costly to the entire ring as it can cause RingOp oscillations. Also, a missed detection may cause the affected stations to have communication problems ranging from a lack of communication to a one way communication problem.

As described earlier, the SMT protocols define the NRR and the TS tests for duplicate address tests during the RingOp-on state. We examine the properties of the two tests and show that neither test is sufficient to cover the more general usages of address as modeled in Figure 6. Later in this section, we present a new algorithm which combines the results of the two tests using a decision table. The new algorithm provides better and more reliable detection of the duplicate address for the more general usages of address. The new algorithm minimizes the likelihood of missed detection and of false detection.

The NRR test based its detection criteria on the A-indicator of the NIF Response frame. Therefore, the test provides a direct indication when a station's MLA is duplicated by another station's MLA or an address in the DA list. In order for the test to provide correct detection, all stations on the ring must comply to the simple address usage. If stations on the ring deviate from the simple address usage then the test may exhibit both missed detection and false detection failures.

As shown in Figure 7, the algorithm has missed detection if a station's SA list contains the duplicate address but the DA list does not. In the example, station 1b contains in its SA list (but not it's DA list) the value

of station 1a's MLA. Station 1a transmits the NIF Request frame which is received by station 2. However, station 1b strips the frame due to the SA list match. Some time later, station 2 sends a NIF Response frame as a response to station 1a's NIF Request frame. Since there is no DA match, station 1b repeats the NIF Response frame without setting the A-indicator. The frame with the A-indicator reset is received by station 1a. The NRR test in station 1a declares that there is no duplicate address problem. Thus, a missed detection of the duplicate address problem results. Also, a missed detection can result if the immediate downstream neighbor station cannot respond to NIF Request frames due to an error condition in the station (e.g., temporary congestion).



Station 1_b has station 1_a's MLA in its SA list but station 1_a's MLA is not in its DA list.

<1> Station 2 receives the NIF Request frame from station 1_a.

<2> Station 1_b repeats the NIF Response frame and does not set the A-indicator.

<3> Station 1_a receives the NIF Response frame with the A-indicator reset. The NRR Test in station 1_a does not detect the duplicate address problem.

Figure 7: Example of Missed Detection with NRR Test

Several other conditions may cause the NRR test to exhibit a false detection. A false detection results when a station's DA list contains the duplicate address but the address is not the MLA or in the SA list. Also, a false detection can result if the immediate downstream neighbor does not strip the NIF Response frame, due to a frame stripping failure. When the station performing the test receives the NIF Response frame the first time, it sets the A-indicator as it repeats the frame. When the station receives the same frame the second time, it will falsely declare a duplicate address problem.

The TS test detects the absence of duplicate address based on the fact that if a frame can traverse successfully around the entire ring then the SA of the frame is unique with respect to the MLA and the SA list of stations on the ring. The test covers all cases of duplicate involving MLA or SA list addresses. It does not cover any duplicate involving another station's DA list. In the absence of a duplicate address problem the test can complete in one round trip delay (e.g., less than about 1.7 milliseconds). After receiving its own test frame, the station can conclude that there is no duplicate address problem for its MLA. Therefore, the test is quick and adequate to cover the absence of duplicate address. In the presence of a duplicate address problem the test waits for a time-out before concluding that the test has failed. Since the test indicates failure based on time-out, it does not provide a direct indication of the duplicate address problem. Additional faults may cause the TS test frames to be repeatedly lost during the test period.

When considering the more complicated addressing scheme, the NRR test has been shown to have false detection and missed detection. In the presence of a duplicate address problem, the TS test detects a failure. However, it does not provide a positive indication of the duplicate address problem. By combining the strength of each of the tests, we arrive at a new algorithm. We present the following new algorithm that has the desirable and improved reliability in the detection of the duplicate address problem.

The new algorithm performs both the NRR test and the TS test in parallel as two concurrent and independent tests. The results of the tests are combined by a decision table, shown in Table 1, to determine if there is a duplicate address problem or not. Each time the NRR test or the TS test changes state, the decision table is consulted to update the result of the duplicate address test. The result of Table 1 determines the value of Dup_Address_Test which is exported to the RMT.

The SMT standard allows either the NRR test or the TS test to be run as part of the duplicate address test. The output of the NRR test can change the Dup_Address_Test directly. The use of the TS test is optional. If the TS test indicates a Pass then the Dup_Address_Test is set to Pass. However, if the TS test indicates a Time-out then the Dup_Address_Test is set to Unknown (also called, NONE). The following summarizes the results of the NRR and TS tests.

The NRR test has the following results:

- Unknown - the test has not concluded if a duplicate address is present or not. NOTE: ANSI SMT uses NONE to indicate this result.

- Pass - the test has concluded that there is no duplicate address.

- Fail - the test has concluded that there is duplicate address.

The TS test has the following results:

- Starting - the test was just started. This is a transitory state that occurs right after a station reset or an new address is installed.
- Pass - the test has concluded that there is no duplicate address.
- Time-out - the test has concluded that it has failed to received its test frame within a time-out period.

Case	Inputs :		Outputs :
	RESULTS OF NRR	RESULTS OF TS (note 1)	DUP_ADDRESS_TEST (note 2)
1	Unknown	TIMEOUT (note 3)	UNKNOWN
2	Unknown	PASS	PASS
3	PASS	TIMEOUT (note 3)	UNKNOWN
4	PASS	PASS	PASS
5	FAIL	TIMEOUT	FAIL
6	FAIL (note 3)	PASS	PASS

Notes:
1. To simplify the table, the STARTING input is not shown.
2. ANSI SMT uses NONE to indicate the fact that the result of the test is neither PASS nor FAIL.
3. An error event is reported.

TABLE 1 - Decision table for the new duplicate address test

Each time the NRR test or the TS test changes state, Decision Table 1 is used to decide the state of the test. The results of NRR test and TS test are combined to produce the new state for the Dup_Address_Test. Then, the Dup_Address_Test value is exported to the RMT.

The algorithm has better reliability and detection than either of the two tests used separately. The following are the key advantages:

- It minimizes missed detection which impacts the entire ring. As shown in case 3, when the result of the NRR test is PASS but the result of the TS test is TIMEOUT, the output of the test is set to UNKNOWN. In this case, an error event is reported so that the problem can be resolved. For example, case 3 can result if a station's MLA is in another station's SA list but the address is not in the DA list. This is an important duplicate address case. The NRR test will indicate a PASS but the TS test will indicate a Time-out. The output of case 3 is designed to be compliant to the SMT standard, which disallows the TS time-out to change the Dup_Address_Test to Fail.
- It minimizes false detection which impacts the operation of the station. As shown in Case 6, when the result of the NRR test is FAIL but the result of the TS test is PASS, the output of the duplicate address test remains as PASS to minimize unnecessary impact to the station. For example, a station's MLA is duplicated by an address in another station's DA list (but not in SA list). Since the station's MLA is not duplicated by another station's MLA or an address in the SA list, the station continues to operate in the ring.
- It minimizes the dependency on the immediate downstream neighbor and the A-indicator. For example, if the immediate downstream neighbor fails to respond to the NIF Requests then the NRR test result may not change. The decision table is designed such that for Cases 2, 3 and 6, the results of the TS test decide the outcome of the duplicate address test.
- It allows more general use of addresses in the ring. The decision table is designed to detect duplicate address problem involving another station's MLA or address in the SA list. It does not indicate a duplicate address problem when a station's MLA is duplicated by another station's address in the DA list. The decision table can be custom designed to suit special requirements with respect to the use of addresses and the sensitivity to the type of duplicate address.

We have demonstrated the advantages of the new duplicate address detection algorithm that is based on the ANSI SMT defined tests. The decision table is presented as an example that has improved reliability in the detection and also is compliant to the ANSI SMT. The decision table is designed to indicate duplicate address problem based on duplicate involving another station's MLA or address in the SA list. One can design the decision table to suit the application's needs.

Conclusions:

We have discussed uses of addresses in FDDI and focused on the detection of duplicate addresses which otherwise could disrupt the ring operation. The SMT standard was shown to detect and resolve simple duplicate address problems but not be sufficient for many complicated but common cases. We described a more formal and elaborate definition of the duplicate

address problem. Methods described provide reliable detection of duplicate addresses which are not covered by the SMT standard. These methods are incorporated in Digital's FDDI products to increase LAN reliability and performance.

Common usage of addresses in FDDI gives rise to more complicated problems than are covered by currently standardized algorithms. In order to avoid serious ring problems, stations must implement appropriate duplicate address detection methods that can cover the more complicated problems. Stations with additional addresses in an SA list must implement the RMT extension for the RingOp-off duplicate address detection. The RMT extension detects an SA list duplicate based on an SA list match with the SA of beacon or Claim frame. If the RMT extension is not in place then the station with the SA list should disable the SA list whenever RingOp is not on, to avoid ring initialization failure. Another method to minimize duplicate address problem is to use FCIS for frame stripping which eliminates the use of SA list.

For the RingOp-on case, the SMT defines the Neighbor Request/Response test and the Transmit-to-Self test for duplicate address detection. We showed several cases of missed detection and false detection when the two tests are used separately. We presented an improved method for the duplicate address detection which runs the two tests simultaneously and combines the results of the tests using a decision table. The improved method gives more reliable detection and minimizes failures such as missed detection or false detection.

Additional study is needed for several pathological failure modes involving duplicate address. For instance, duplicate address may cause a rapid oscillation between RingOp on and ring initialization, which prevents completion of the duplicate address test that could otherwise detect and resolve the condition. Stations with many addresses used for SA stripping, which need to complete the duplicate address tests for many addresses, serve to aggravate the likelihood of such pathology. A reduction in the number of addresses used for SA stripping through the use of FCIS may be a more effective means in minimizing ring disruption. Then, the number of addresses which present a high likelihood of ring disruption, is reduced from the number of clients of the network to the number of MAC entities in the ring.

References:

1. FDDI MAC, ANSI X3.136-1987, Token Ring Media Access Control (New York: American National Standards Institute, 1987).

2. Preliminary Draft proposed standard for FDDI Station Management (SMT), Accredited Standards Committee (ASC) X3T9.5/90-078, Rev. 6.3, (May 18, 1990).

3. IEEE 802 Overview and Architecture, IEEE802-1990, Institute of Electrical and Electronic Engineers.

4. K. Ocheltree, R. Montalvo, "FDDI Ring Management," IEEE 14th Conference on Local Computer Networks, October 1989, Minneapolis, Mn.

5. J. Hutchison, "RMT (Ring Management)," FDDI X3T9.5 Working Group on SMT, Committee Document X3T9.5/89-127 (August 22, 1989).

6. IEEE P802.1d, "MAC Bridges (Unapproved Draft)," P802.1d9, July 14, 1989.

7. Advanced Micro Devices, "AM99C10 Content Addressable Memory," Memory Products Data Book, publication 08125, Sunnyvale, Ca., February 1989.

8. H. Yang, K.K. Ramakrishnan, "Frame Content Independent Stripping for Token Rings, " Proceedings of the ACM SIGCOMM '90 (September, 1990).

9. ANSI/IEEE Draft Appendix to ANSI/IEEE Std. 802.5-1989, "Enhancement for Multiple Ring Networks," unapproved draft, P802.5D/D15 (89/25), Institute of Electrical and Electronic Engineers.

10. Harry L. Van Trees, Detection, Estimation, and Modulation Theory, John Wiley & Sons, 1968.