Dual MAC Usage in FDDI

A. Thimmel, P. Kühn
Siemens Nixdorf Informationssysteme AG, 8900 Augsburg, Germany
Universität München, Lehrstuhl für Kommunikationsnetze, 8000 München, Germany

Abstract
The FDDI Station Management (SMT) standard [01] allows the use of the primary and the secondary ring for data transmission. This dual ring operation has the ability to double the available bandwidth of FDDI, giving a maximum transmission bandwidth of 200 Mbit/s.

In this paper, two possible applications of dual ring usage in FDDI are described. First, a straightforward mechanism of load sharing on the MAC layer based on the assumption of a homogeneous dual MAC network is discussed. As a second approach, the TCP/IP sublayer DMARP (Dual MAC ARP) has been introduced. This sublayer supports dual MAC operations in inhomogeneous FDDI networks. Finally, an outlook identifies a way towards the usage of dual MAC FDDI in general network protocols.

1. Overview - Dual Ring Usage
1.1. General aspects
The FDDI dual ring operation enables a migration path towards either high bandwidth demanding FDDI applications or huge FDDI rings with numerous active FDDI nodes [04]. These two applications are targeted for different environments. In both cases, the dual ring usage in FDDI implies some common, partially still unsolved problems concerning the network operation and integrity. Therefore, this differentiation of two environments is not to be investigated here.

In order to utilize both FDDI rings for data transmission, a method for distributing traffic into different traffic classes has to be worked out. These traffic distribution methods could be based on various topics.

Some possible traffic differentiation items are summarized below:

1. transmission priorities, for instance batch versus file transfer (PRIORITY mux)
2. synchronous versus asynchronous traffic classes (SYNC/ASYNC mux)
3. FDDI I, versus FDDI II (PS/CS mux)
4. application traffic versus signaling (or network management) traffic, for instance outband signaling for higher layer protocols (ACK mux)
5. load balancing due to the number or utilization of transport or network layer connections in connection oriented layer (CONNECTION mux)
6. load balancing due to the actual total traffic on the two FDDI rings (MAC mux)

Methods 1. to 4. described above are based on relevant traffic attributes. This enables a meaningful traffic differentiation. For instance, the first method ideally supports a quality of service feature. As a tradeoff of a meaningful traffic differentiation these methods tend not to optimize the total channel utilization. This applies especially in case of very different traffic volumes in the multiplexed traffic classes.

The goal of this paper is to discuss ways of utilizing both FDDI rings in order to maximize the available FDDI bandwidth. Therefore, this paper will focus on the load balancing aspects of dual ring usage (methods 5. and 6.).
A more comprehensive discussion of the remaining traffic multiplexing methods, due to dual ring usage in FDDI, is given in [05].

1.2. FDDI topology

The FDDI topology is defined as a dual ring of trees. Figure 1 shows this structure.

![Figure 1: FDDI Topology - Dual Ring Of Trees](image)

The trees may be realized with Single Attached Concentrators (SAC) and/or Dual Attached Concentrators (DAC) connecting Single Attached Stations (SAS) and/or Dual Attached Stations (DAS) to FDDI. The DAS can be realized as single MAC (SM) or dual MAC (DM) endstations. In Figure 2, a combined SM/DM structure is shown.

![Figure 2: Inhomogeneous SM/DM FDDI Topology](image)

In an inhomogeneous SM/DM configuration a routing problem between endstations may arise as soon as at least one station is only a SM. The reachability on both rings cannot even be deduced if it is known that a target station is a DM. Consider the situation, when two dual MAC stations reside on different dual FDDI rings which are linked with only one transparent bridge.

1.3. Load sharing criteria

The process of sharing a node's load into two distinct "subloads" can be characterized by the following aspects:

- transparency to the end user
- OSI layer in which the load sharing process is to be performed
- channel (dual ring) utilization
- system behavior for reconfigurations (wrap conditions)
- adoption of changes in the overall network load distribution (fast, media traffic dependent flow control)
- atomic data unit size (smallest data unit for which load sharing is to be performed)
- PDU misordering

If there is a meaningless traffic distribution it is desirable to perform the load sharing process transparently inside the communication system without user interaction. This leads to the realization of load sharing as an OSI layer 3 or an OSI layer 2b functionality.

Routing with the related flow control is a natural OSI network layer task. Therefore, it seems to be obvious to locate the investigated load balancing functionality into the network layer. However, this is only possible with extensions to the existing network layer [06], [07].

Because data traffic is normally bursty, the traffic distribution of both rings may change very rapidly. To obtain an effective load balancing it is important to have close interaction with the actual channel load condition. This can easily be achieved with an OSI layer 2b load balancing.

Moreover, it is possible to have layer 2b load balancing which is compatible with existing network protocols. Therefore, for initial applications at least, it is easier to use the Data Link Layer for load balancing.

A simple load sharing method called MAC mux is described in the following.
2. MAC Layer Multiplexing Solution

2.1. MAC multiplexing description

The proposed MAC mux method is based on the following assumptions:

1. The network under consideration is a homogeneous dual MAC network.
2. Each PDU may be transferred via either FDDI ring.
3. It is assumed that the PDU length distribution is time independent. This makes it possible to observe the number of waiting customers in the transmit queues instead of observing the total amount of data to be transmitted.
4. The length of the transmit queues in an investigated station can be used as an indicator of the total ring utilization for the corresponding FDDI ring.
5. In case of wrap conditions, it is assumed that all MACs remain connected to the (logically unified) FDDI rings. This applies, in particular, to the two stations in the wrap state¹.

Based on the assumptions above, the operation of the MAC mux can be characterized in the following way:

- Each station has two consecutive MAC addresses. This means that the least significant bit of the MAC address is to be extended by the MAC mux with the ring index. On the interface to the network layer this bit is meaningless and will always be set to a certain predefined value.
- A PDU is entered in the smallest transmit queue. If both transmit queues are empty or of equal size the selected transmit queue is alternated. This avoids a default preference for one ring.
- With a single FDDI wrap condition, no further action is required. Several wrap conditions will cause several isolated rings, as in standard SM FDDI operation.
- On the receive side a common LSAP is addressed by both receive queues. Misordering of PDUs is possible and has to be fixed as usual in the Network Layer.

This MAC mux can effectively utilize both FDDI rings. The system adapts very quickly to the real, dual ring load conditions by detecting the actual transmit queues lengths.

In Figure 3, a homogeneous dual MAC FDDI ring is shown in unwrapped and in wrapped mode.

![Figure 3: Homogeneous DM FDDI Ring](image)

It can be seen that the logical operation of the MAC mux method remains unaffected when there is a single wrap condition. Although the available bandwidth is reduced to the single FDDI ring bandwidth, this topology change is transparent to the LSAP service user.

2.2. MAC Mux - applications

As mentioned already, the proposed MAC mux is restricted to homogeneous DM FDDI rings. This is acceptable only in special, dedicated applications like point to point links.

For more general use, it is desirable to extend the targeted network scenario from homogeneous DM FDDI rings to inhomogeneous SM/DM FDDI rings.

3. TCP/IP Sublayer Solution

There are several different protocol suites to investigate for their possibilities of dual MAC FDDI support. The Load Sharing Sublayer solution presented here is targeted for TCP/IP.

---

¹ The FDDI SMT standard version 6.2 allows only one MAC to remain connected in a wrapped FDDI station. This is under discussion for future SMT standards.
3.1. Assumptions

The proposed solution is based on the following assumptions:

- The network under consideration is an inhomogeneous FDDI network.
- With wrap conditions, it is assumed that all MACs remain connected to the (logically unified) FDDI rings as described in the MAC Multiplexing description.
- Each DM has two different MAC addresses.
- SM are restricted to the primary ring. In inhomogeneous SM/DM configurations the MAC of SM may theoretically reside on either FDDI ring. However, additional routers (or bridges) are required to enable connectivity between SM on different rings. This paper concentrates on applications of dual MAC FDDI operations where no additional internetworking devices are required.
- The use of transparent bridges between the primary and the secondary ring is not supported by the introduction of the new DMARP protocol [08].
- The same IP network address is used for both MACs in a DM.

3.2. Standard IP stack model

In a standard FDDI station the Internet Protocol (IP) and the Address Resolution Protocol (ARP) [10] work on top of the Subnetwork Access Protocol (SNAP) [11].

IP is the well known network protocol of the TCP/IP protocol stack and ARP converts Internet Protocol addresses to 48 bit MAC addresses. SNAP provides a method of running protocols which are not standard ISO protocols over 802.2 LLC [07], [11].

The services of the Data Link Layer are assumed to be provided by the IEEE 802.2 Type 1 Logical Link Control (LLC) protocol [09].

A simplified model of an IP/ARP stack and the ARP layer is shown in Figure 4. The Address Conversion Unit actually converts the IP addresses to hardware addresses. Multicast and broadcast IP addresses are directly converted, individual addresses are converted by the use of the address cache. If an individual IP address cannot be found in the address cache, an ARP request is transmitted.

The cache administration communicates with other stations through ARP and creates, deletes and renews the entries in the address cache.

![Figure 4: Standard IP Stack Model](image)

3.3. Protocol stack extensions

To use dual MAC stations on FDDI networks, an extension to the TCP/IP on FDDI protocol suite in these stations is necessary. This is done by introducing a new load sharing sublayer.

The extension to the protocol suite has the following characteristics:

- The protocols above the new load sharing sublayer may reside in either a SM or DM. The new sublayer holds the property DM transparent to ARP, IP and the higher protocol layers (UDP, TCP, ...).
- It is transparent to ARP, IP and the higher layer protocols that other stations on the network can be realized as DM.
- The ARP in the SM and DM only knows the primary MAC address as the hardware address.
for other stations on the network. This is true whether or not a counter station uses an additional MAC on the secondary ring. The addresses of MACs on the secondary ring should only be known by the new sublayer.

3.4. Load Sharing Sublayer

To provide against modifications to existing standard protocols, the required extension to the TCP/IP on FDDI protocol suite in a DM is done by the introduction of a new sublayer, called the Load Sharing Layer.

Figure 5 shows the proposed location for the new sublayer.

![Figure 5: DMARP IP Stack Model](image)

According to the OSI Reference Model, the new sublayer is part of the network layer because it requires an address cache which contains address and reachability information.

The new sublayer only needs to be introduced in the network layer of dual MAC stations. The protocol suite of single MAC stations needs no modification.

3.5. Load Sharing Sublayer structure

The Load Sharing Layer consists of a multiplexer, an address cache and a cache administration.

The multiplexer multiplexes IP frames onto the MACs. To transmit individual frames, the contents of the address cache and the load sharing criteria are used to decide on which of the rings an IP frame has to be transmitted.

If an individual IP frame is transmitted on the secondary ring, the source and destination hardware address must be converted from primary MAC addresses to secondary MAC addresses.

Broadcast and multicast frames are always sent on the primary ring. Multicast and broadcast IP frames which are received on the secondary ring are discarded so that they are not duplicated when a wrapped dual FDDI ring is being used.

The address cache contains the pair of primary MAC and secondary MAC addresses for each known DM on the network. There are no entries for SM. In addition to the MAC addresses for another DM, information indicating whether the other DM can be reached via both rings or only via the primary ring must be saved in control bits.

The cache administration creates, deletes and reviews the entries in the address cache. For this purpose, it communicates with the other DM through a new protocol, called the Dual MAC Address Resolution Protocol (DMARP) [08],[15].

Also, the cache administration is responsible for transferring ARP frames to the SNAP layer for transmission. ARP frames are always sent on the primary ring. Broadcast ARP frames which are received on the secondary ring are discarded.

3.6. The Dual MAC Address Resolution Protocol (DMARP)

IP and higher layer protocols only use the network-independent IP addresses. The ARP layer converts IP addresses to the appropriate hardware (MAC) addresses.

The ARP layer in all stations only knows the addresses of MACs on the primary ring. For this purpose, the Load Sharing Layer in a DM always transfers ARP frames to the primary MAC for transmission. Communication through ARP is performed via the primary ring in the normal state. It is sufficient, that the ARP layer in a DM knows the primary MAC address at the station's hardware address. In this way, the IP address / primary MAC address assignment for the SM and DM on the network can be stored in the ARP cache.
The Load Sharing Layer in the DM must know both MAC addresses of the other DM on the network and which of the other DM can be reached on both FDDI rings to be able to perform load sharing.

DMARP was created for this purpose. DMARP is the protocol which is used by the Load Sharing Layer to exchange address and reachability information between the DMs. Thanks to DMARP, two DM can determine whether they can perform load sharing on the two rings or whether they can only use the primary ring to exchange frames.

It is not possible to give a complete description of DMARP in this paper. A comprehensive and formal description of DMARP and a complete discussion of the Load Sharing Layer can be found in [08] and [15].

Now, one problem which can be solved by DMARP will be discussed. Consider two DMs which reside on different FDDI rings. The primary rings are linked by a transparent bridge, but there is no connection between the secondary rings. These DM stations cannot perform load sharing of frames. Like ARP, DMARP uses reply frames, which contain both MAC addresses and the IP address of the sending station to broadcast these addresses to other DM. These frames are sent by the secondary MAC. Normally, a DM can deduce from the reception of a DMARP reply frame by its secondary MAC, that the sending station can be reached on both rings and load sharing can be performed with it.

The configuration under consideration runs into a problem when both dual FDDI rings are wrapped. Without further precautions, the DM would assume that they can perform load sharing on receiving DMARP reply frames. As long as both rings remain wrapped, no problem occurs. However, a "black hole" in the routing of frames among these DMs occurs when one of the dual rings returns to the normal state.

A possible approach to overcome this problem is to observe whether a DMARP reply frame was also received by the primary MAC. No load sharing will be performed with a station when a DMARP reply frame has been received with the primary MAC of it. In this case, DM stations which reside on the same dual FDDI ring would stop the load sharing among them when the common ring wraps once.

The solution of this problem using DMARP involves observing whether a DMARP reply frame was received by the primary MAC, but the reception of such a frame starts a very refined algorithm for exchanging additional types of DMARP frames. This algorithm can also handle lost-frame errors.

For example, all DMs check if they have received their own DMARP replies with their primary MAC. They send an error frame to the target station(s) as soon as they detect this event.

If a DM has received a DMARP reply frame with its primary MAC, it sends back requests periodically to the counter station to initiate further transmission of reply frames. By this way the DM determines if it really can perform load sharing with this station or if an abnormal network state has occurred for some time.

If two DM cannot solve their reachability problem by themselves, they finally can, as a last resort, request help from other stations by the transmission of DMARP search frames and use the routing information in the Load Sharing Layer of other stations.

3.7. Applications

The widespread use of inhomogeneous dual FDDI rings is made possible by the introduction of a new sublayer below the network layer of the TCP/IP protocol suite.

Other IEEE 802 LANs, or further inhomogeneous dual FDDI rings, can be linked with transparent bridges or routers. Two inhomogeneous dual FDDI rings can be connected with one or two transparent bridges, whereas one bridge links the primary rings while the other links the secondary rings. It is still assumed that the MAC in single MAC stations resides on the primary ring. Figure 6 illustrates an inhomogeneous SM/DM network configuration supported by the DMARP protocol.

Figure 6: Example Illustrating DMARP Topology

As mentioned before, using SMs with their MAC on the secondary ring is also possible. In order to provide connectivity between stations, additional routers
(bridges) have to be used in these SM/secondary ring configurations. A significant problem arises with the use of transparent bridges which link the primary and the secondary ring. The DMARP protocol's ability to determine whether a dual ring is wrapped is lost under these circumstances.

There are various ways of implementing the load sharing criteria which have been discussed and so the same approach used for the MAC mux load sharing layer can be adopted.

4. Outlook: Further Network Protocols

Aside from TCP/IP, it is desirable to enable the operation of dual MAC FDDI networks for arbitrary network protocols. For the OSI Connectionless Network Protocol [16], a very complete discussion of its application on dual MAC FDDI networks can be found in [06].

For a general sublayer solution which can be implemented for several different network protocols, the concept of an OSI Subnetwork Dependent Convergence Protocol Sublayer (SNDCP Sublayer) [18] offers an interesting perspective.

An appropriate SNDCP Sublayer can support network protocols complying with the requirements of an OSI Subnetwork Independent Convergence Protocol (SNICP Protocol). One possible SNDCP Sublayer can be developed from the Load Sharing Sublayer. Figure 7 shows the service interface and the structure of the extended Load Sharing Sublayer.

The extended Load Sharing Sublayer supports the services of the IEEE 802.2 Type 1 LLC layer to the network protocol.

A complete description of required Load Sharing Sublayer extensions to create the SNDCP Sublayer can be found in [15]. The major extensions are:

- The hardware addresses in the MAC header of received frames must be converted from secondary MAC addresses to primary MAC addresses. For this purpose, the contents of an Address Cache is used. If a secondary MAC address cannot be converted the frame will be discarded.

- When sending a frame, the multiplexer searches an entry containing the destination primary MAC address. The event of a missing address is signaled to the Cache Administration. The missing address is stored in the Address Cache. Later, the Cache Administration determines the corresponding secondary MAC address if the target station is a DM. This is done with the DMARP protocol.

- The primary MAC addresses of SM are also stored in the Address Cache.

The extended Load Sharing Sublayer supports all network protocols which can operate on top of the IEEE 802.2 LLC Type 1 protocol layer. An example is the OSI Connectionless Network Protocol (ISO 8473) in conjunction with the ES-IS protocol [17].

5. Summary

Several ways of implementing dual ring usage in FDDI have been identified.

The idea of a MAC mux has been discussed. There will definitely be a demand for the additional bandwidth of the secondary ring. The MAC mux discussed in this paper is the first step in this direction even with the restriction of a homogeneous DM network.

To allow the widespread use of DM, a solution for inhomogeneous FDDI rings has to be found. The DMARP method which has been described allows the introduction of SM/DM FDDI configurations into the TCP/IP world while at the same time being transparent to existing SM applications.
The SNDCP, described as an outlook can be used as base for a protocol stack independent solution for DM FDDI networks.

6. References

[01] ANSI; FDDI Station Management (SMT); ANSI X3T9/90-X3T9.5/84-49 Rev 6.2
[02] ISO, Fibre Distributed Data Interface (FDDI) - Media Access Control, ISO 9314-2
[03] Ross, F., Hamstra, J., Fink, R.; FDDI - A LAN Among MANs; ACM Press, Volume 20 Number 3; July 1990
[04] Strohl, M.; High Performance Distributed Computing in FDDI Networks; IEEE LTS; May 1991
[08] Kühn, P.; Thoughts on Address Resolution for Dual MAC FDDI networks; RFC 1329; March 1992
[10] Plummer, C.; An Ethernet Address Resolution Protocol --or-- Converting Network Protocol Addresses to 48 bit Ethernet Address for Transmission on Ethernet Hardware; RFC 826; November 1982
[12] Requirements for Internet Hosts - Communication Layers; RFC 1122; Internet Engineering Task Force; October 1989
[13] Ochelrree, K. B.; FDDI Ring Management; IBM Research Division, T.J. Watson Research Center, Yorktown Heights, N.Y. 10598
[17] ISO; Information processing systems - Telecommunications and information exchange between systems - End system to Intermediate system routing exchange protocol for use in conjunction with the Protocol for providing the connectionless-mode network service (ISO 8473); ISO 9542, August 1988.