Submitted to IFIP 3rd Conference on High Speed Networking, Berlin March 1991

Architectural Concepts for Dual Ring Operation in FDDI

Martin Siegel
Institut für Nachrichtenvermittlung und Datenverarbeitung
Universität Stuttgart
Seidenstr. 36, D-7000 Stuttgart 1
E-Mail:SIEGEL@nvdv.e-technik.uni-stuttgart.dbp.de

ABSTRACT
The growing demand for more bandwidth in inhouse computer networks can be satisfied by the emerging high speed local area network standard FDDI (Fibre Distributed Data Interface). In order to integrate different telecommunication services, an evolution towards the hybrid FDDI-II is specified. Both LANs use dual counter-rotating optical fibre rings. Normally one ring is active and the second ring stands by, to allow a reconfiguration in case of single ring breaks. This paper describes the required hardware and software functions which are necessary to use both rings under normal working conditions without losing the fault-tolerance provided by ring reconfiguration around ring breaks.

1. Introduction

The steadily increasing demand for information interchange over inhouse

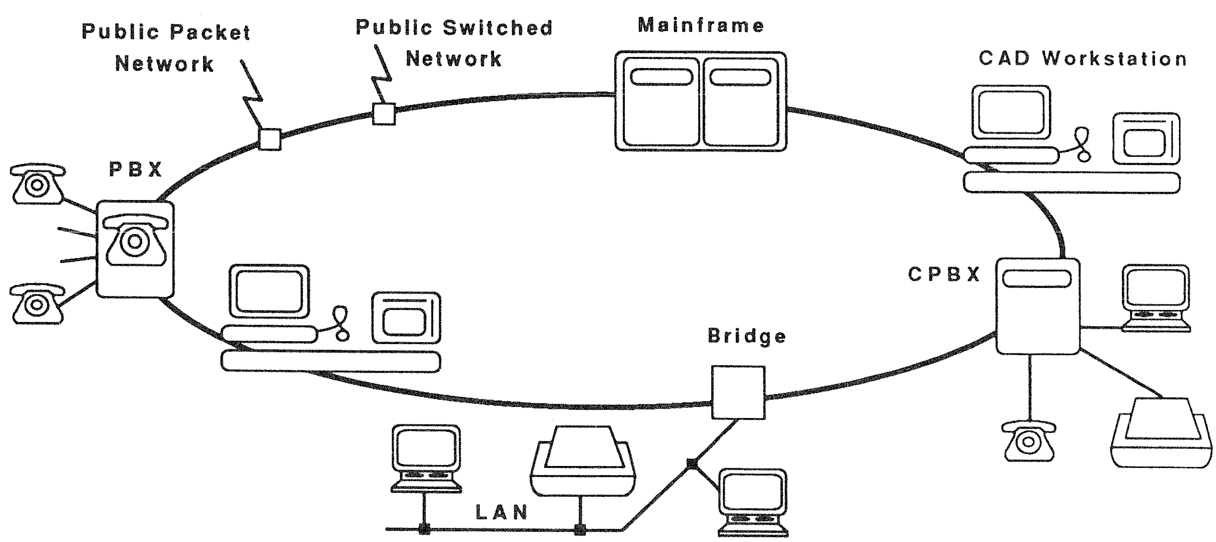

Figure 1: Typical high-speed LAN applications

computer networks is driven by the proliferation of workstations and other distributed computer resources as well as by the current tendency to integrate all kinds of digital traffic into one network (see Figure 1).

Due to the heterogeneous environment in such inhouse networks, standardization is more than a 'nice-to-have' feature. Currently, there are two main international standardization efforts in the area of MANs and HSLANs. These are the IEEE 802.6 activity for defining a metropolitan area network standard, based on the distributed queue, dual bus (DQDB) protocol [1], and the ANSI X3T9.5 committee activities. The latter group has finalised the FDDI-I standard [2-5] and has issued draft proposals for the FDDI-II enhancements [6]. This paper concentrates on the two FDDI versions.

Both are specified for 100 Mb/s user data transmission rate (a 4B/5B coding leads to 125 Mb/s transmission rate on the fibre), a total maximum ring length of 100 kilometers and a maximum of 500 dual attached stations. FDDI-I supports one synchronous traffic class, which guarantees a minimum bandwidth and delay to the user, and up to eight asynchronous priority classes. FDDI-II additionally provides the transmission of delay sensitive isochronous, circuit switched traffic such as digitized voice or video (see [7,8]).

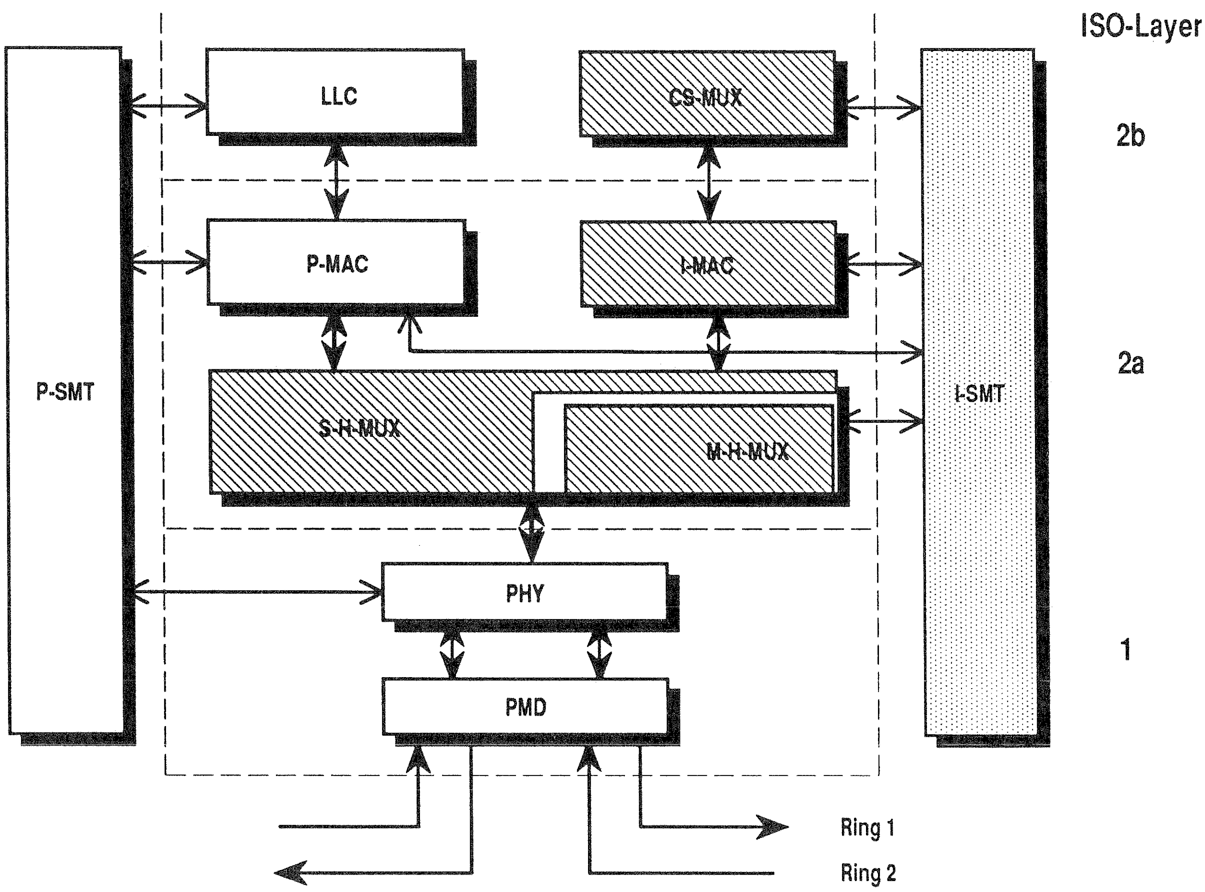

Figure 2: Block Diagram of FDDI-II Stations

Figure 2 shows the major blocks which have to be in every FDDI-II station with relation to the well known OSI communication layers [16]. The unmarked blocks PMD, PHY, P-MAC and P-SMT represent the FDDI-I parts as standardized in [2-5]. Hardware implementing these functions is available from a number of vendors (see e.g. [12]). The interface above the MAC-layer 2a is chosen according to the IEEE 802.2 conventions [17], thus allowing the use of logical link control software available on the market.

The blocks used to include the isochronous traffic are shaded. The H-MUX (including master and slave), I-MAC and CS-MUX are specified in [6], the isochronous station management I-SMT is currently under development. Due to the ongoing and therefore partially still unstable standardization procedure, no hardware to implement FDDI-II is offered by any manufacturer yet. In [9] and [10] proposed designs for the needed hardware components and the main functions of each block are described in detail.

Based on that hardware, this paper summarizes the standard-conformant variants of ring modes and station types in Section 2. The principal problems of a full dual ring usage, without losing the mandatory reconfiguration capability, and possible solutions are discussed in Section 3. Finally the actual state of the project, current aspects of interest and an outlook will be given.

2. Standard-conformant FDDI Variants

To support the heterogeneous environment for which the high speed LANs are intended, the FDDI standards are specified in an implementation-independent manner. Thus no specific combination of working modes on the two rings is prescribed, nor is one unique hardware configuration specified for FDDI. The principal intention of the standard is the interworking of even differently equipped stations. The stations are allowed to be attached to the two fibre rings, which ideally could each operate in an arbitrary mode.

The following subsections summarize the mode combinations and station hardware equipment types separately, although these aspects are interdependent (e.g. a pure FDDI-I station forces basic mode). Section 2.3 introduces the management entities which are necessary to control a fully equipped dual attached FDDI-II station.

2.1 Ring Modes

The left branch of Figure 3 represents the usual understanding of the dual ring usage: one ring normally works in either basic (FDDI-I) or hybrid (FDDI-II) mode, the second ring is not used for transmission of data, but remains in standby state awaiting the - relatively improbable - reconfiguration case.

The right branch contains all combinations of ring modes if both rings are used under normal working conditions. The number of meaningful combinations is increased by taking different parameter settings for the two rings into account. With the same modes and parameters on both rings, the user sees "only" a network with double bandwidth.

Mixing the modes is an economically important variant for enabling a step-by-step migration from FDDI-I installations to a final full hybrid solution. Pure FDDI-I stations can still be connected to one of the two rings, while the other ring operates in hybrid mode. In case of a necessary reconfiguration either all stations can be forced to operate in basic mode, or all FDDI-I stations can be disabled, allowing FDDI-II stations to continue in hybrid mode.

With the same modes on both rings, but different parameter settings, the two rings can be optimized for different traffic characteristics, e.g. long TTRT to increase throughput (e.g. for file transfer) on one ring, and very short TTRT to minimize delay (e.g. for control messages) on the other ring (see [11]).

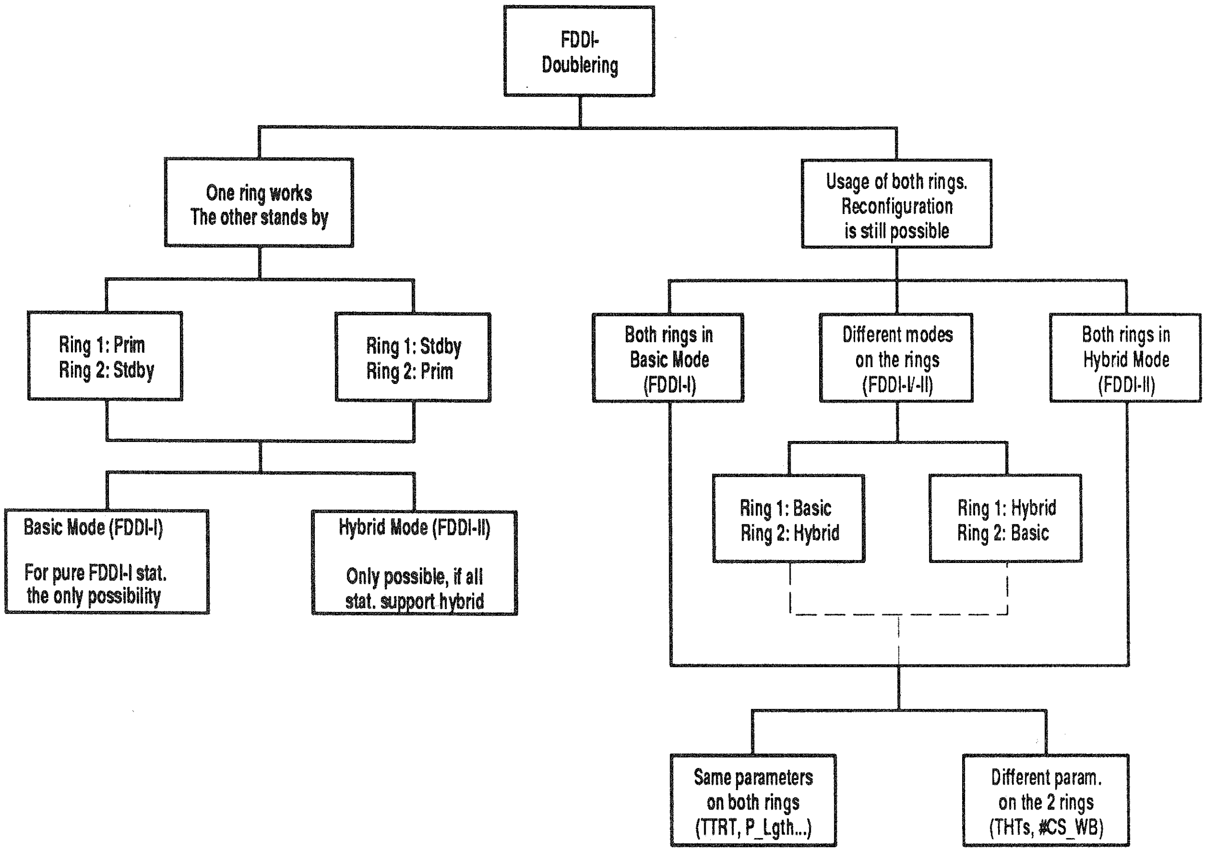

Figure 3: Possible modes on an FDDI dual ring network

2.2 FDDI Station Hardware

The obvious classification of FDDI stations into FDDI-I and FDDI-II stations can be seen at the top of Figure 4.

The well known subdivision into single and dual attached stations (SAS and DAS) follows. The third general type of FDDI station, the concentrator, is mentioned for completeness. For the further discussion however, the concentrator is not handled separately.

The FDDI-I branch does not provide many hardware alternatives. Only the number of available P-MACs, and thus the ability to use one or both rings, may differ. Dual attached FDDI-II stations on the other hand, offer a host of implementation possibilities, which can be customized to the intended ring modes! The number of H-MUXes is the main characteristic, followed by the number of P-MACs and I-MACs as well as their relative configuration. The dotted branches concerning single sttachment stations (SAS) and the SAS concentrators are possible from the hardware side, but violate the rules of the standard, which only allows SAS connected to DAS concentrators. Regular FDDI stations have to be DAS, anyway!

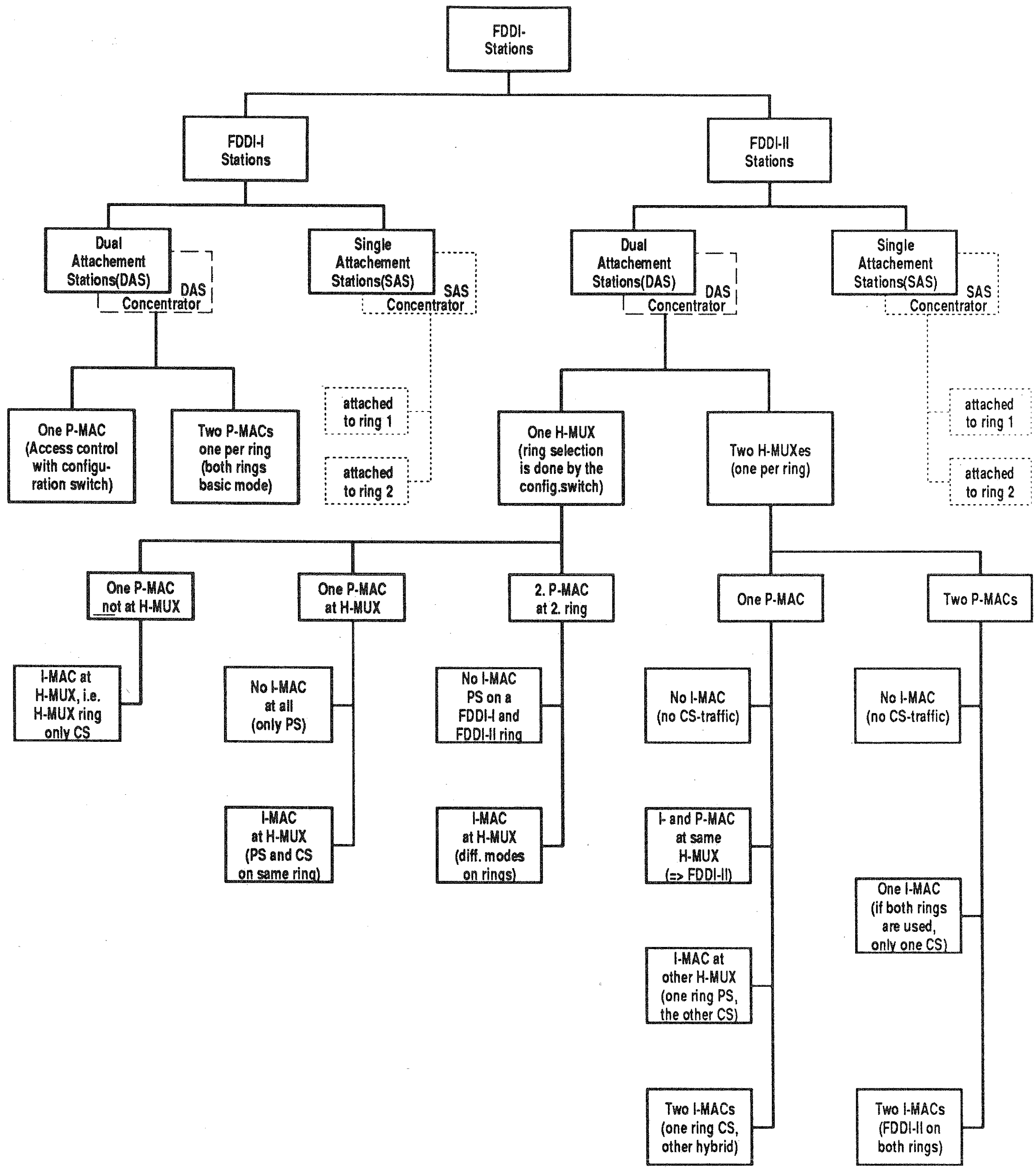

Figure 4: FDDI Station Types

2.3 Management Concept

The previous sections have shown a variety of possible ring modes and station hardware configurations. Handling all those physical variants alone would require some management effort. In fact, *normal* protocol-conformant operation of an FDDI network (e.g. hardware monitoring and configuration, unique identification, manipulating protocol parameters etc.) also can not be performed by the station's hardware itself. Any kind of error handling also introduces lots of external monitor and control effort.

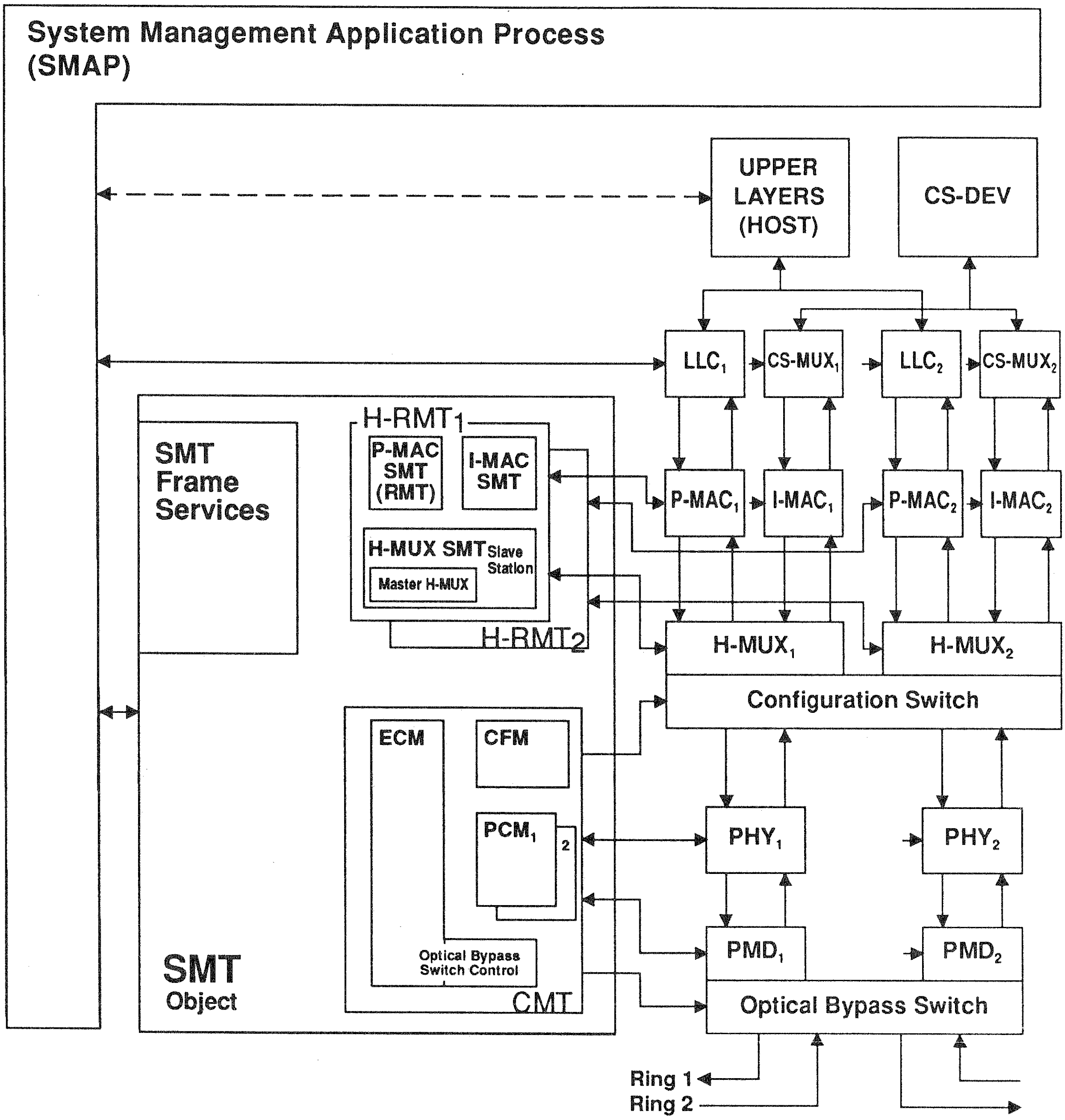

Figure 5: FDDI Station Management

The FDDI standard committee decided, therefore, that although the FDDI standard should leave management issues to the individual implementors, there should be a uniform management concept (see [5]). Most current implementations and also the examples in [5], however, cover only single ring operation. Due to the lack of the isochronous management standard, naturally only management structures for FDDI-I stations are given. Figure 5 enhances the station management concept for stations which are fully equipped to use both rings, as well as including additional components to enable hybrid mode. The latter enhancements lead to the box hybrid ring management H-RMT and the former simply results in a doubling of the H-RMT (the doubled physical configuration management PCM is needed in any dual attached station anyway).

The elements on the right represent the managed objects assigned to the same functional blocks as in Figure 2. The SMT-box represents the FDDI specific management entities. In the standardized ISO management nomenclature the SMT is a layer management entity (LME). The SMT provides an interface to the system management application process (SMAP) for the whole network on the leftside of Figure 5. The internal complexity of the whole connection management CMT is of course, increased by the change from single to double ring operation. In particular, the entity coordination management ECM, which controls the configuration management CFM must perform more functions. Anyhow, it is important to remark that all parts of the FDDI-I station management are used also in the FDDI-II environment. A division into P-SMT and I-SMT, as used in Figure 2, is only valid in a logical sense!

3. Dual Ring Operation

The main arguments for a dual ring operation are the available, but mostly unused, technologically complex - and thus expensive - resources of the complete second physical layer in each dual attached station. If VLSI components for the H-MUX, P-/I-MAC, CS-MUX and LLC support are assumed (see [9-12]), the additional hardware effort to support dual ring operation is small. As shown in Subsection 2.3, the required changes inside SMT are only marginal, and already included in the standard [5].

On the other hand, two main problems occur with full dual ring operation:

1. Distributing the load onto the two available rings.
2. Guaranteeing the required reconfiguration capability of the network. This especially includes the problems of how to handle the resulting overload and which traffic can in this case be discarded.

The two problem areas and a proposed architecture are discussed in the following subsections.

3.1 Traffic Acceptance and Distribution

The first general task is to decide if there are enough network resources (of the requested type). If such resources are available, the second task has to select one of the rings as the transmission medium. These tasks have to be performed for PS and CS traffic. The main relevant aspects within the two tasks are:

- Is the decision hidden from the user (i.e. does the communication system decide 'automatically' or is user control needed)? In other words: is the user aware that there are two rings, or does it simply seem as if there is one ring with enhanced performance?
- Which parameters are relevant for the decision?
- Where (i.e. on which OSI layer) should the selection be performed?
- What is the smallest unit of data for which an independent ring selection should be performed?
- Does the selection influence the receiver end?

- How does the selection strategy fit into the overall environment with partially standardized interfaces and strict timing constraints?

Discussing all these aspects and their correlations exhaustively, is a seperate research area, and is far beyond the scope of this paper. Some principal possibilities with their main pros and cons are summarized below in order to make the decision for the proposed solution more transparent to the reader.

The effort for the network is obviously minimized if the user is responsible for ring selection. In that case, the users can determine the parameters of the rings according to their wishes. Besides the loss of ring operation transparency, however, this method has several severe disadvantages:

The standard makes no provision for access to the low-level information needed to make this decision, and hence existing hardware provides insufficient or no support.

In case of trouble the users are directly involved, i.e. no global overload strategies could be supplied.

In contrast to the OSI goals, a functionality which is normally located inside the OSI stack, is performed by the application.

To allow the synchronous traffic class, additional reservation routines are required.

As a consequence of the mentioned disadvantages,
a network-internal ring selection mechanism should be used!

A distribution below the MAC layer on a bit, symbol or octet basis - similar to processor busses - fails due to the counter-rotating traffic flow directions on the two rings. But even with same traffic flow direction on the rings (which is not allowed in the standard!), the capacity doubling below the MAC layer would also fail, because it is nearly impossible to synchronize data streams at 100 (125) Mb/s over distances up to 100 km.

The main criterion for the acceptance and distribution of CS traffic is the required bandwidth. Besides the bandwidth request, some kind of user defined priority, could be a further criterion.

The criteria for PS depend on the ring modes. If both rings are operating with the same parameter settings three principal strategies are possible:

1. The selection is done just above the P-MACs, i.e. at the entrance to layer 2a. An incoming data unit takes that ring, on which the next token arrives. The load thus is balanced equally between the two rings and the delay characteristic is optimized. An implementation has to provide one single data buffer, to which both P-MACs have access. A disadvantage is the possible loss of data unit sequence at the receiver end. The strategy also works after a reconfiguration; the next arriving token is the only existing one, and the user will only see degraded performance (and overflowing buffers!).
2. To guarantee correct sequence up to the corresponding receiver entity, whole packets have to take the same ring. Packet here is meant in the sense PDU - protocol data unit, and need not necessarily be identical to user packets. The selection in that case is done before a packet is copied into a ring specific buffer, i.e. layer 2b. The criterion is the actual filling level of each buffer. The less full buffer is chosen. After a

reconfiguration one buffer could be marked as permanently full, or could be logically added to the other buffer, to increase the available buffer space.

3. The user could divide traffic into classes, and assign each class to a ring. These classes need not be equal to the FDDI priorities, but such relations can be used to minimize the implementation effort. The classes could even be defined in higher OSI layers. As a consequence, the load situation on the two rings can differ greatly if the traffic profile is assymetric. The behaviour in case of error also depends on the implicit priorities, and thus does not introduce much additional effort. The synchronous traffic can be transmitted either over one fixed ring or arbitrarily over both rings (due to the reservation overhead, it is more likely, that synchronous traffic will be restricted to only one specific ring!).

With *different operation parameters on the two rings* the load dependent criteria 1. and 2. cannot be applied, e.g. because an earlier arriving token does not necessarily indicate lower load, or because a longer queue on the faster ring is perhaps served more quickly than an empty queue on the slower ring. Due to the possibly extremely divergant delay characteristics, all parts of one PDU must use the same ring.

The third mechanism can be adapted in such a way, that there is a primary ring choice for each priority class and a secondary choice for overload situations.

Although the above mentioned strategies were presented only for PS traffic, the third mechanism, which is suitable for PS traffic on dual rings with the same characteristics as well as for dual rings with different characteristics, can be expanded to handle CS the same way! To be able to react more flexibly to the current load situation, a *logical channel concept* seems suitable. The traffic, which is assigned to these logical channels can be distributed in a load-dependant mannor. For management, further information to determine the parameters of a reconfigured ring (such as for example maximum tolerable delay or minimum acceptable mean throughput) can be related to those channels. The logical channel concept is discussed in the next subsection.

3.2 Resulting Architecture

The previous subsection made clear that the location for the traffic acceptance and distribution logic should be above layer 2.

- Choosing one way out of different alternatives, is basically a routing task, and thus a 'classical' layer 3 function (see [18]).
- The logical channel concept is reminiscent of the virtual channels of the connection oriented network service, CONS.
- An implementation below layer 3 would introduce some modifications to standard FDDI hardware blocks.

The resulting layered structure of an enhanced FDDI station can be seen in Figure 6. The FDDI Subnet Access sublayer performs the adaption to the LLC interface for one single ring. The internet sublayer is the place to implement all necessary general layer 3 functions, and to realize the standardized interface to the transport layer.

The enhanced FDDI LME block is introduced, to make clear that there are management functions for the enhanced FDDI subsystem which neither are a part of the standardized FDDI LME for the lower layers (i.e. the SMT) nor do they belong to the overall system management application process SMAP. Once more it has to be stated, that the content of the SMT need not to be changed in order to enable the full dual ring operation!

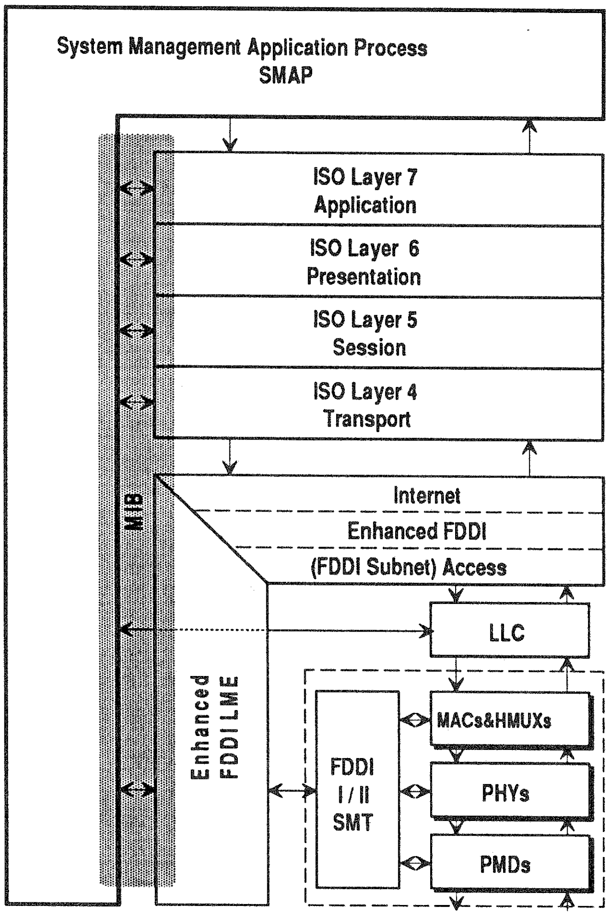

Figure 6: Layered Structure of the enhanced FDDI Station

Three further aspects also argue for a layer 3 location:

Before variants for implementing the proposed mechanisms are described, the user's point-of-view of the traffic classes is given. A first classification of the user traffic can be made according to the FDDI traffic types isochronous (CS), synchronous and asynchronous. A known relative weight of these classes alone is not sufficient for performing reconfiguration. Instead, for all three classes, additional information about the traffic's type internal 'importance' - using the word priority here would be clearer, but could lead to misunderstandings - has to be given, to allow sensible overload handling. Besides an explicitly fixed 'throw away in case possible/impossible' classification by the user, also traffic characteristics can be transferred. For CS traffic the needed bandwidth, supposed call duration, error sensitivity or ordering scheme of single channels of multichannel connection are examples. For the PS traffic types, a decision between throughput and delay optimizing has to be made. To weight asynchronous traffic in one station a relation to the FDDI priorities can also be used for traffic classification.

In requests to the network layer, generally one field for including quality of service (QoS) parameters is foreseen. This QoS field can be used for assigning the user traffic category to one of the available classes of the double ring network.

The QoS field can be used either as a short type of service indicator, much as in the internet protocols [14], or it can explicitly contain all parameters which are used for the class assignment, similar to the Facility field in X.25 (see e.g. [13,15]). Regarding connectionless network service, CLNS, for which the QoS parameters must be contained in every single PDU, the latter variant obviously introduces a non negligible overhead.

For the former variant (that with a short type of service indicator) two principal realizations are possible: either the meaning of the indicator can contain a fixed predefined relation to a certain class (which has to be known by the calling user), or it can simply contain a logical channel number. In this case, the relation to a class has to be negotiated between the user and the enhanced FDDI sublayer once, perhaps via interaction with system management software.

This - only on a first view rather complex - construct has the big advantage that no specific mechanisms for CS channels have to be provided! In a CONS environment a combining of the logical channel concept with the virtual (connection) channel construct is possible. Furthermore, the dynamically and freely changeable relations provide a flexibility which is very helpful during and after a reconfiguration when overload becomes a problem.

4. Conclusion

This paper has summarized the principal standard-conformant possibilities of FDDI ring modes and principal FDDI station hardware types. The discussion of the management structure needed for FDDI-II has shown, that the FDDI-I SMT will also be used completely for FDDI-II, and that the needed effort to support full dual ring operation is only marginal.

Motivated by the availability of - partially in-house developed - highly integrated FDDI-II hardware, a project to realize a working FDDI station has been instigated. Since this hardware allows dual ring operation, alternative ring management architectures have been evaluated, and the impact on surrounding entities has been investigated. Finally the resulting enhanced layer 3 functionality, based on dynamically changeable logical channels, has been introduced.

Currently the implementation of the complete layer 3 is underway in association with studies, which are intended to generalize the layer 3 in such a way that all kinds and also an arbitrary number of underlying OSI layer 2 systems, e.g. a Ethernet and a FDDI, can be supported.

Within the overall intention to develop and implement a performance optimized communication system, a project to tune layers 2b to 4 and to investigate general architectural issues, e.g. buffer management, context switching, pipelining etc., has been started.

Acknowledgement

I would like to thank Peter Domschitz for valuable discussions and for the time consuming and difficult reading and structuring of the related standards. Further thanks go to Mark Williams, who has spent some hours to correct errors and make the paper clearer. Finally I thank an anonymous reviewer for valuable comments, which showed a very careful reading.

References

[1] *DQDB Metropolitan Area Network,* Draft Proposed IEEE 802.6 Standard D11, November 1989

[2] *FDDI Physical Layer Medium Dependent*, Draft Proposed ANSI Standard X3T9.5, Rev. 7.3, 1988

[3] *FDDI Physical Layer*, Draft International Standard 9314-1, 1987

[4] *FDDI Token Ring Media Access Control*, ANSI Standard X3T9.5, 1988

[5] *FDDI Station Management*, Draft Proposed ANSI Standard X3T9.5, Rev. 6.1, April 1990

[6] *FDDI Hybrid Ring Control*, Draft Proposed ANSI Standard X3T9.5, Rev. 5.2, April 1990

[7] Ross, F.E.; *FDDI - A Tutorial*, IEEE Communication Magazine, Vol. 24, No. 5, 1986, pp. 10-17.

[8] Ross, F.E.; *An Overview of FDDI: The Fibre Distributed Data Interface*, IEEE JSAC, vol. 7, No. 7, Sept. 1989, pp. 1043-1051.

[9] Siegel, M.; Sauer K.; Schödl W.; Tangemann M.; *Design of an enhanced FDDI-II System*, Proc. of the International Zurich Seminar 1990, Paper E9, pp. 418 - 433.

[10] Siegel, M.; *ASICs für das Hochgeschwindigkeits-LAN FDDI-II*(in german), Mikroelektronik für die Informationstechnik '90, Berlin, Okt. 1990.

[11] Tangemann, M.; Sauer, K.; *Performance Analysis of the FDDI Media Access Control Protocol*, Proceedings of the 4th International Conference on Data Communication Systems and their Performance, Barcelona, June 1990, pp. 32 - 44.

[12] *The SUPERNET Family for FDDI*; AMD databook, March 1989

[13] Sloman M.S.; *X.25 Explained*, Computer Communication, Vol. 1, No. 6, Dec. 1987, pp. 310-326.

[14] Comer D.; *Internetworking with TCP/IP Principles, Protocols, and Architectures*, Prentice-Hall International Editions, Englewood Cliffs, New Jersey, 1988.

[15] Tanenbaum, A.S.; *Computer Networks* (second Edition), Prentice-Hall International Editions, Englewood Cliffs, New Jersey, 1989.

[16] ISO 7498; *Information Processing Systems - Open Systems Interconnection - Basic Reference Model*, November 1983.

[17] ISO 8802/2; *Local Area Networks - Logical Link Control* , 1985.

[18] ISO 8348 und 8880; *Network Service Definition* and *Specifications of Protocols to Provide and Support The OSI Network Service*.

[19] ISO 7498-4; *Information Processing Systems - OSI Reference Model - Part 4: Management Framework*.