A Media Access Control Protocol for High-Speed Local Area Networks Integrating Asynchronous and Isochronous Services

Summary

Protocols for shared media Local Area Networks (LANs) are facing three major challenges: The growing demand for high bandwidth to the desktop, upcoming multimedia applications requiring low delay and low delay jitter at high data rates and integration of ATM traffic which will be required for interoperation with Wide Area Networks. Traditional LANs, e. g., Ethernet or Token Ring, do not guarantee the required quality of service, and they are not scaleable to rates of 1 Gbit/s and beyond. Their efficiency decreases when the medium bit rate increases since the Media Access Control (MAC) protocol allows only a single node at a time to send data.

High-Speed Local Area Networks (HSLANs) have been designed for high bit rate data traffic. For connections which require low delay, peak bit rate reservation guarantees the required quality of service. However, for bursty multimedia traffic this usually results in a waste of bandwidth.

The goal of this thesis is the development and evaluation of a new scaleable MAC protocol for HSLANs which efficiently provides high throughput for data traffic as well as low delay for delay-sensitive traffic.

Chapter One: Introduction

After a brief introduction to Local Area Networks, Chapter 1 presents the current local area network development efforts and challenges. In addition, an introduction to the ISO reference model for Open Systems Interconnection (OSI) is given.

Chapter Two: Shared Media Local Area Networks

This chapter introduces the currently most popular Local Area Networks, i. e. Ethernet and Token Ring, and their problems in a multi-media environment. These LANs are based on bus, star and ring topologies, but there are also other potential topologies which are presented in this chapter. A description of terms, definitions, and principles of MAC protocols is given as a

basis for classification and evaluation of HSLANs. The chapter concludes with an overview and evaluation of HSLAN protocol examples.

Chapter Three: ATM-based Local Area Networks

The Asynchronous Transfer Mode (ATM) was designed for public networks, but more and more effort is made towards the application and standardisation of ATM in the local area, too. Chapter 3 introduces ATM, its layered model and its Quality of Service abilities. The mechanisms required to integrate current LAN services and applications are also described in this chapter.

Chapter Four: A New Gigabit LAN Architecture

In Chapter 4 a new protocol for Gigabit LANs is presented. Requirements for a promising future HSLAN concept are derived from a comparison between shared media LANs as presented in Chapter 2 and LANs based on ATM as presented in Chapter 3. This new protocol should be scaleable for high bit rates, provide real-time connections with low delay and low jitter and utilize the available bandwidth efficiently. The new MAC protocol integrates isochronous and asynchronous traffic in a ring LAN.

Fixed-length frames are used on this ring which are further divided into slots. Each slot can hold an ATM cell. In the isochronous traffic class, slots are reserved for every connection. Thus for isochronous traffic channels the protocol provides guaranteed bandwidth with a guaranteed ATM cell delay limit. In the new approach asynchronous traffic can use the unallocated and, in addition, all the reserved slots that are currently not used within isochronous connections. It is no waste of bandwidth to reserve an isochronous channel with its peak bandwidth for a bursty connection because the bandwidth not used by the connection can be used for asynchronous traffic.

Each node has a list of „usable slots in frame“ to determine which slots in every frame are currently reserved for its isochronous connections. The node „owns“ these slots, i. e. it decides how the slots are to be used. In any case they can be used for the node‘s own isochronous traffic and the node can decide whether other nodes can use the slot for asynchronous traffic within the next ring cycle by setting the Available bit in the slot header. Incoming asynchronous cells are buffered in the node‘s insertion buffer. Unreserved slots are always available for asynchronous traffic.

To avoid an insertion buffer overflow in case of heavy asynchronous traffic load, an available counter is introduced. This counter‘s value represents the sum of all the node‘s slots that are marked as „available“ on the ring, in addition to the number of cells in the insertion buffer. If the available counter reaches a limit which is the maximum number of cells the insertion buffer can hold, the node‘s outgoing slots are marked as unavailable, i. e. the Available bit is not set.

Slots with the Available bit not set must not be used for asynchronous cells. This mechanism ensures that the insertion buffer cannot overflow.

On top of this basic medium access mechanism, a credit-based access algorithm for the asynchronous traffic is applied in order to achieve fairness. The reset interval in this algorithm may be comparably long because all time critical traffic will be sent using isochronous connections, leading to a high overall throughput in the asynchronous traffic class.

Chapter 4 continues with the description of a LAN node's architecture which could be used for implementation.

Chapter Five: Gigabit LAN Performance

In Chapter 5 the performance of the new MAC protocol is examined. Comprehensible scenarios are analysed to prove the expected behaviour of architecture. Simulation results are provided which show the performance under heavy load conditions in a realistic asymmetric traffic model combining data, video conference and video distribution applications in one local network.

Chapter Six: Gigabit LAN Evaluation

A description of a prototype node implementation in Chapter 6 adds a more practical view to the architectural description of the new MAC protocol. The prototype was implemented using Field Programmable Gate Arrays at a bit rate of approximately 1 Gbit/s. An estimation of the required resources is derived from the prototype. All of one node's functionality could be integrated on a single plug-in board. This chapter also discusses possible variations of the system, including low-cost or higher bit rate LANs or even other applications such as Controller Area Networks. The chapter concludes with a description of gateways which are required to integrate the new system into existing Ethernet environments and for access to Wide Area Networks.

Chapter Seven: Conclusion And Outlook

Chapter 7 concludes the thesis, including a summary of the new protocol and its benefits. Expectations of future LAN systems and the new MAC protocol's deployment opportunities are described in an outlook.

Inhaltsverzeichnis

Abkürzungen

AAL	ATM Adaptation Layer
AC	Available Counter
ACF	Access Control Field
ACT	Available Counter Threshold
ADU	Atomic Data Unit
AMNET	A Multimedia NETwork
ANSI	American National Standards Institute
ARP	Address Resolution Protocol
ATC	Asynchronous Traffic Class
ATM	Asynchronous Transfer Mode
ATMR	ATM-Ring
B-ISDN	Breitband-ISDN
BNC	Bayonet Norm Connector (Steckernorm)
BUS	Broadcast and Unknown Server
CAM	Content Addressable Memory
CC	Credit Counter
CCI	Credit Counter Initial Value
CLP	Cell Loss Priority
CPCS	Common Part Convergence Sublayer
CPI	Common Part Indicator
CRC	Cyclic Redundancy Check
CRMA	Cyclic Reservation Multiple Access
CS	Circuit Switching
CS	Convergence Sublayer
CSMA/CD	Carrier Sense Multiple Access with Collision Detection
D^3Q	Dynamic Distributed Dual Queue
DMA	Direct Memory Access
DQDB	Distributed Queue Dual Bus
EOP	End Of Packet
FDDI	Fiber Distributed Data Interface
FECCA	Fair and Efficient Cyclic Control Algorithm
FFOL	FDDI Follow-On LAN
FOR	Frames On Ring
GFC	Generic Flow Control

HCS	Header Check Sequence
HEC	Header Error Control
HSLAN	High-Speed Local Area Network
HTTP	Hyper Text Transfer Protocol
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
IP	Internet Protocol
ISDN	Integrated Services Digital Network
ISO	International Standards Organization
ITC	Isochronous Traffic Class
LAN	Local Area Network, Lokales Netz
LANE	LAN Emulation
LEC	LAN Emulation Client
LECS	LAN Emulation Configuration Server
LES	LAN Emulation Server
LIS	Logical IP Subnetwork
LLC	Logical Link Control
LSB	Least Significant Bit
MAC	Media Access Control
MAN	Metropolitan Area Network
MD^3Q	Multirequest Dynamic Distributed Dual Queue
MPEG	Moving Picture Experts Group
MPOA	Multi-Protocol Over ATM
MSB	Most Significant Bit
NNI	Network Node Interface
NT	Network Termination
OSI	Open Systems Interconnection
PBX	Private Branch Exchange
PAR	Parity
PCI	Peripheral Components Interconnect (Peripheriebus)
PCI	Protocol Control Information
PDU	Protocol Data Unit
PLL	Phase Locked Loop
PSR	Previous Segment Received
PT	Payload Type
QoS	Quality of Service
RAM	Random Access Memory

RFC	Request For Comments
RVCI	Ring Virtual Channel Identifier
SAP	Service Access Point
SAR	Segmentation and Reassembly
SDH	Synchronous Digital Hierarchy
SDU	Service Data Unit
SH	Segment Header
SLT	Slot Type
SOF	Start Of Frame
SOC	Start Of Cell
SP	Segment Priority
SSCS	Service Specific Convergence Sublayer
STD	Synchronous Time Division
TA	Terminal Adapter
TCP	Transmission Control Protocol
TCP/IP	Transmission Control Protocol/Internet Protocol
TE	Terminal Equipment
UDP	User Datagram Protocol
UNI	User Network Interface
UU	User-to-User Information
VCC	Virtual Channel Connection
VCI	Virtual Channel Identifier
VPI	Virtual Path Identifier
WAN	Wide Area Network
WWW	World Wide Web

Kapitel 1

Einführung

1.1 Entwicklungstendenzen der lokalen Datenkommunikation

Lokale Netze (Local Area Networks, LANs) haben sich überall dort als selbstverständlicher Teil der Rechnerinfrastruktur etabliert, wo mehrere Rechner lokal begrenzt auf gemeinsame Daten zugreifen oder verteilte Ein-/Ausgabegeräte nutzen. Die Übertragungskapazitäten im meistverwendeten lokalen Netz, dem Ethernet [34], waren mit 10 Mbit/s zu Anfang weit überdimensioniert, was sich mit dem exponentiellen Anstieg der Rechnerleistung, der Speicherkapazitäten und der zunehmenden Anzahl angeschlossener Stationen schnell änderte. Lokale Netze bilden zunehmend einen Engpaß in der Rechnertechnik. Deshalb werden große lokale Netze in logische Einheiten gegliedert, die wiederum über ein „Backbone“-Netz hierarchisch strukturiert und miteinander gekoppelt sind. Die Rechnerinfrastruktur wird derart organisiert, daß die Kommunikation möglichst innerhalb der Untereinheiten beschränkt bleibt.

Derartige organisatorische Maßnahmen zur Trennung der Datenströme reichen heute jedoch nicht mehr aus, da die Ein-/Ausgabegeschwindigkeit moderner Rechner mit bis zu 4 Gbit/s [75] um derzeit zwei Größenordnungen höher ist als die Übertragungsrate auf dem Ethernet. Benötigt werden Netze mit einer deutlich höheren Übertragungsrate, die auch dem Bandbreitenbedarf künftiger Rechnergenerationen gewachsen sind und dem gleichzeitigen Bedarf mehrerer simultan aktiver Nutzer entsprechen.

Videoverteildienste, Videokonferenzen oder Telekooperation sind neuartige Anwendungen, die gleichzeitig mit dem erhöhten Bandbreitenbedarf Dienstgütemerkmale erfordern, die in herkömmlichen lokalen Netzen nur mit zusätzlichem Aufwand oder gar nicht garantiert werden können. Videoanwendungen, insbesondere Videokonferenzen, sind in hohem Maße auf die Einhaltung einer Obergrenze der Verzögerungszeit angewiesen und werden daher derzeit über getrennte Netze geführt. Die Kooperation von Partnern in einer virtuellen Welt wird noch wesentlich zeitkritischer sein, insbesondere wenn Vorgänge modelliert werden sollen, die unmittelbare, zeitkritische Reaktionen mehrerer über ein lokales Netz verbundener Partner erfordern.

Aufgrund der zunehmenden Bedeutung des World Wide Web (WWW) und der damit einhergehenden weltweiten Datenkommunikation wird ein effizienter Übergang zu öffentlichen Netzen angestrebt. Die derzeit vorherrschenden Modem- und ISDN-Übergänge bieten eine im Vergleich zu lokalen Netzen extrem geringe Bitrate und müssen daher durch höherbitratige

Übergänge ersetzt werden. Es wird davon ausgegangen, daß bei breitbandigen Zugängen zu öffentlichen Netzen die ATM-Technik zum Einsatz kommen wird [46].

Drei Entwicklungstendenzen für künftige lokale Netze sind erkennbar:

- Hochgeschwindigkeits-Lokalnetze (HSLANs) mit gemeinsamem Übertragungsmedium sind speziell auf die Bedürfnisse der lokalen Datenkommunikation zugeschnittene Architekturen mit Medienzugriffsprotokollen, die für hohe Übertragungsraten im Bereich von 1 Gbit/s geeignet sind. Die zur Verfügung stehende Bandbreite wird nach einem Fairness-Verfahren unter den sendewilligen Stationen aufgeteilt. Wenn nur eine Station sendet, kann sie die volle Medienbandbreite nutzen. Mehrere Stationen können gleichzeitig senden. Die meisten Protokollvorschläge verwenden Datenzellen konstanter Größe, deren Format in den wesentlichen Punkten dem ATM-Zellformat entspricht [48].
- Die ursprünglich für Weitverkehrsnetze konzipierte ATM-Technik findet zunehmend Einzug in lokale Netze. Derartige ATM-LANs bestehen aus einem Vermittlungsknoten und jeweils exklusiven Übertragungsmedien zu den Stationen. Übertragungsrate und -verfahren zwischen Station und Vermittlungsknoten müssen nicht einheitlich sein und können mithin den Kommunikationsbedürfnissen der jeweiligen Station angepaßt werden. Die Mechanismen des gemeinsamen Übertragungsmediums, auf die bestehende Protokolle und Anwendungen in lokalen Netzen basieren, werden emuliert [101].
- Die bestehende, kostengünstige Architektur lokaler Netze auf Basis von Ethernet wird praktisch unverändert fortgeschrieben und zu einem Hochgeschwindigkeits-LAN mit deutlich erhöhter Bitrate weiterentwickelt. Das resultierende, aufgrund des geringen Entwicklungsaufwands und der einfachen Struktur sehr kostengünstige Netz kann jedoch aufgrund physikalischer Grenzen immer geringere Abmessungen annehmen und ist folglich nicht für größere Netze geeignet. Die Übertagungsrate kann nicht beliebig gesteigert werden, so daß nach der derzeit in der Standardisierung befindlichen Generation, dem „Gigabit-Ethernet“, eine weitere Steigerung der Übertragungsrate fraglich ist. Zudem bietet Ethernet keinerlei Möglichkeit zur Priorisierung und keine Garantie für eine maximale Verzögerungszeit, so daß es auch bei höherer Bitrate für zeitkritische Anwendungen ungeeignet ist.

Die Protokollarchitektur lokaler Netze wird entsprechend dem ISO-Basisreferenzmodell für offene Kommunikation (OSI) [32] in Schichten gegliedert. Bild 1-1 zeigt die für lokale Netze relevanten Schichten. Jede Schicht ist dabei Diensterbringer für die nächsthöhere Schicht, die weitere Funktionen hinzufügt. Die Schichten 5-7 im OSI-Modell sind anwendungsorientiert

5-7	Anwendungssystem (Session, Presentation, Application Layer)
4	Transportschicht (Transport Layer)
3	Vermittlungsschicht (Network Layer)
2b	Logische Sicherungsschicht (LLC Layer)
2a	Medienzugriffsschicht (MAC Layer)
1	Bitübertragungsschicht (Physical Layer)

Bild 1-1: Referenzmodell für lokale Netze

und für lokale Netze von geringer Bedeutung. Sie werden als Anwendungssystem zusammengefaßt.

Die Schichten 1-4 bilden ein Transportsystem für den gesicherten Nachrichtentransport zwischen Endsystemen, wobei die von der Transportschicht erbrachte Funktionalität unabhängig ist von den Realisierungsformen der darunterliegenden Schichten.

Die *Bitübertragungsschicht* (Physical Layer) definiert elektrische, mechanische, optische und funktionelle Eigenschaften, die zu einer seriellen Bitübertragung über einen ungesicherten Kanal erforderlich sind.

Die *Sicherungsschicht* (Data Link Layer) stellt eine gesicherte Nachrichtenverbindung bereit. Sie wird zur Klassifizierung lokaler Netze in zwei Subschichten eingeteilt: Die *Medienzugriffsschicht* (Media Access Control Sublayer, MAC) regelt den Zugriff auf das gemeinsame Übertragungsmedium und stellt sicher, daß keine gestörten Nachrichten empfangen werden. LAN-Protokolle werden im wesentlichen durch die MAC-Schicht charakterisiert. Die *logische Sicherungsschicht* (Logical Link Control Sublayer, LLC) ermöglicht eine Wiederholung im Fehlerfall, die Sicherung der Reihenfolge und eine Flußsteuerung.

Die *Vermittlungsschicht* (Network Layer) stellt eine Verbindung zwischen zwei Endsystemen über einzelne gesicherte Teilabschnitte her. Die Vermittlungsschicht stellt Mechanismen zum Verbindungsaufbau zwischen Endsystemen, zur Verkehrslenkung und zum Multiplexen mehrerer Verbindungen zwischen zwei Endsystemen zur Verfügung. Das „Internet Protocol“ (IP) ist die bedeutendste Realisierungsform der Vermittlungsschicht in lokalen Netzen.

Die *Transportschicht* (Transport Layer) stellt dem Anwendungssystem gesicherte Verbindungen mit wählbarer Dienstgüte zu anderen Endsystemen zur Verfügung. In lokalen Netzen sind das verbindungsorientierte „Transmission Control Protocol“ (TCP) und das verbindungslose „User Datagram Protocol“ (UDP) die wichtigsten Transportprotokolle, die auf dem „Internet Protocol“ aufsetzen.

Die Medienzugriffsprotokolle herkömmlicher lokaler Netze sind auf den effizienten Datenaustausch zwischen Rechnern und Terminals optimiert, um Daten in möglichst kurzer Zeit zu transportieren. Dabei wird akzeptiert, daß die Übertragung bei hoher Netzauslastung länger dauert und Daten mit einer größeren Verzögerung zur Verfügung stehen. Für die Übertragung von Dateien zwischen zwei Rechnern ist die Übertragungszeit von untergeordneter Bedeutung: Mehrere Stunden dauernde Übertragungen werden akzeptiert, wenn kein Nutzer interaktiv auf den Abschluß der Übertragung warten muß. Wesentlich kritischer sind die Antwortzeiten bei Terminals, da lange Reaktionszeiten Unmut bei Nutzern hervorrufen und die Produktivität senken können. Reaktionszeiten von deutlich über einer Sekunde, zum Beispiel zwischen der Eingabe eines Zeichens und dessen Darstellung auf dem Bildschirm, sind nicht akzeptabel, da sie zu unvorhersehbarem Nutzerverhalten mit ungewollten Folgen führen können. Mit der Übertragung von Videosequenzen, zum Beispiel im Zusammenhang mit dem WWW, sind herkömmliche lokale Netze überfordert.

In dieser Arbeit werden Medienzugriffsprotokolle lokaler Netze hinsichtlich der Anforderungen betrachtet, die aus neuartigen, teilweise interaktiven Anwendungen und der ständig wachsenden Leistungsfähigkeit moderner Rechner sowie dem Trend zur Diensteintegration resultieren. Der Schwerpunkt liegt in der Entwicklung und Bewertung eines neuen Protokolls für Gigabit-LANs, das optimal auf die künftigen Anforderungen und Verkehrsströme abgestimmt ist.

1.2 Gliederung der Arbeit

In Kapitel 2 werden die Grundlagen für Medienzugriffsprotokolle lokaler Hochgeschwindigkeitsnetze mit Vielfachzugriff vorgestellt. Neben möglichen Topologien werden die vielfältigen Mechanismen, die den Zugriffsprotokollen zugrundeliegen, ausführlich erörtert. Ausgewählte aktuelle Protokolle und Protokollvorschläge für lokale Hochgeschwindigkeitsnetze werden zusammengefaßt und bewertet.

Kapitel 3 geht auf die ATM-Technik ein, die auch in lokalen Netzen zunehmend Verbreitung findet und für diesen Einsatzbereich mit einem hohen Aufwand weiterentwickelt und standardisiert wird. Neben der eigentlichen ATM-Technik werden Mechanismen erklärt, die benötigt werden, um bestehende Dienste und Anwendungen auch in einem auf der ATM-Technik basierenden lokalen Netz einsetzen zu können.

In Kapitel 4 liegt der Schwerpunkt auf der Vorstellung eines neuen Protokollkonzepts für Gigabit-LANs. Zunächst wird ein Vergleich der in den Kapiteln 2 und 3 vorgestellten und an sich gegensätzlichen Netze mit Vielfachzugriff auf der einen und ATM-LANs auf der anderen Seite gezogen. Die Anforderungen, die auf ein erfolgversprechendes Konzept für ein künftiges lokales Hochgeschwindigkeitsnetz einwirken, werden dargestellt. An die Vorstellung der Protokollarchitektur des vorgeschlagenen Gigabit-LANs schließt sich als Realisierungsvorschlag die Beschreibung der Struktur eines LAN-Knotens an.

Eine Untersuchung der Leistungsfähigkeit des vorgestellten Protokolls ist Gegenstand von Kapitel 5. Neben unmittelbar nachvollziehbaren symmetrischen Verkehrsbeziehungen werden auch unsymmetrische, heterogene Szenarien durch Simulation untersucht, die eine realistischere Darstellung der Systembelastung in einem realen System ermöglichen sollen.

In Kapitel 6 wird das Gigabit-LAN nach unterschiedlichen Gesichtspunkten bewertet. Zunächst wird ein beispielhafter Laboraufbau einer Station beschrieben, der die theoretischen Erkenntnisse vervollständigen soll. Aus dem Laboraufbau und dem hierfür benötigten Aufwand werden Realisierungsaspekte für die Integration der Stationsfunktionen auf einer Einschubkarte abgeleitet. Mögliche Systemvarianten werden diskutiert, die einer Senkung der Kosten, einer weiteren Leistungssteigerung oder einer Erweiterung des Einsatzbereichs dienen. Schließlich wird dargestellt, wie das Gigabit-LAN in andere Systeme integriert werden kann.

Kapitel 7 faßt den Inhalt der vorliegenden Arbeit zusammen und gibt einen Ausblick auf weitere Entwicklungen und Fragestellungen im Zusammenhang mit künftigen Generationen lokaler Netze.

Kapitel 2

Lokale Netze mit Vielfachzugriff (Shared Media LANs)

In diesem Kapitel werden die Entwicklungen im Bereich lokaler Netze aufgezeigt. Das Hauptaugenmerk liegt dabei auf den Topologien, Technologien und konkreten Protokollvorschlägen für Hochgeschwindigkeits-Lokalnetze, die für Multimedia-Anwendungen und hohe, skalierbare Datenraten geeignet sind [64].

2.1 Geschichte

2.1.1 Ethernet

Im Jahre 1983 standardisierte das IEEE unter dem Standard IEEE 802.3 [34] ein Medienzugriffsverfahren für lokale Netze, das die Firma Xerox unter dem Namen „Ethernet" entwickelt hatte. Bis heute basieren weltweit die meisten lokalen Netze auf 10 Mbit/s IEEE 802.3 und werden - eigentlich fälschlicherweise - als „Ethernets" bezeichnet. Ursprünglich mußten teure Koaxialkabel unter der Bezeichnung „10 Base 5" verwendet werden. Die Handhabung wurde stark vereinfacht mit der Einführung von „10 Base 2", einem wesentlich dünneren und flexibleren Kabel mit BNC-Stecker, das es erlaubt, den Transceiver[1] nicht mehr auf das Kabel zu montieren, sondern in den Rechner zu verlagern. Bild 2-1 zeigt den Aufbau eines Ethernet-Segments.

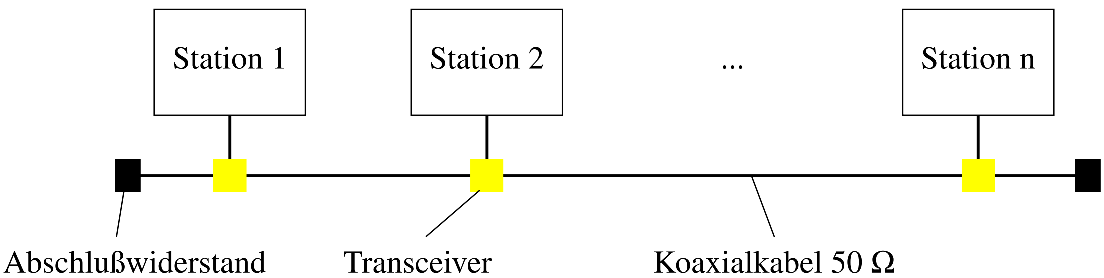


Bild 2-1: Ethernet

1. Der Transceiver ist die Anschlußeinheit an das Kabel

Mit „10 Base T“ wurde die Möglichkeit eingeführt, Ethernet über verdrillte Zweidrahtleitungen zu betreiben. Dies ermöglicht die Verwendung bestehender Telefon-Gebäudeverkabelungen: Da Telefonleitungen oft aus zwei Zweidrahtpaaren bestehen, wird eines der Leitungspaare für Telefon, das andere für Ethernet verwendet. An zentraler Stelle, neben der Telefonanlage, wird ein „Ethernet-Hub[1]“ installiert, in dem die einzelnen Zweidrahtleitungen, an die jeweils genau eine Station angeschlossen sein kann, zusammengeführt sind. Darüberhinaus bietet die daraus resultierende physikalische Sterntopologie Vorteile, auf die in Abschnitt 2.2 näher eingegangen wird.

Ethernet basiert auf dem Medienzugriffsverfahren „Carrier Sense Multiple Access with Collision Detection“ (CSMA/CD). Eine Station kann senden, wenn der Kanal frei ist („Carrier Sense“). Erkennt die Station während des Sendevorgangs eine Kollision, bricht sie den Sendevorgang zunächst ab („Collision Detection“). Der Sendevorgang wird nach einer zufällig gewählten Verzögerungszeit wiederholt. CSMA/CD zeigt ein sehr ungünstiges Verhalten bei hoher Medienauslastung, weil aufgrund des konkurrierenden, unkorrelierten Medienzugriffs die Kollisionswahrscheinlichkeit mit der Netzauslastung steigt. Bei zunehmendem Angebot nimmt der Durchsatz bis zu einem absoluten Maximum zu, um bei höherem Angebot aufgrund der dann dominierenden Kollisionen wieder abzunehmen. Daraus ist zu schließen, daß Ethernet nur dann gut funktioniert, wenn die Auslastung des Netzes gering ist.

Bei seiner Standardisierung war Ethernet weit überdimensioniert: Die maximale und am weitesten verbreitete Übertragungsrate nach IEEE 802.3 beträgt 10 Mbit/s, was in der Größenordnung der Ein-/Ausgabebandbreite damaliger Rechner entspricht. Eine niedrige Netzauslastung war die Folge und somit wirkten sich die Nachteile von Ethernet bei großem Angebot nicht aus. Dies hat sich mittlerweile geändert: Ein einfacher PC mit PCI-Bus hat eine maximale Bandbreite von ca. 1000 Mbit/s, 100 mal mehr als die Bandbreite von Ethernet. Die zur Verfügung stehende Bandbreite von 10 Mbit/s kann somit bereits von einer einzelnen Station ausgelastet werden.

Mit „Switched Ethernet“ wurde eine Möglichkeit geschaffen, die Probleme von Ethernet bei hohem Angebot zu mindern. Statt wie in einem Hub die angeschlossenen Ethernet-Leitungen einfach miteinander zu verbinden, enthält Switched Ethernet Vermittlungsfunktionen, die Kollisionen beim Vielfachzugriff reduzieren. Durch nebenläufige Verbindungen kann der theoretische Maximaldurchsatz in einem Ethernet-Switch auf ein Vielfaches von 10 Mbit/s gesteigert werden.

Neue Dienste verlangen nach noch höheren Bandbreiten für jede einzelne Station. Der Standard IEEE 802.3u [29] definiert ein „100 Base T“ Ethernet auf einer Sternstruktur mit Hub, bei dem die Bitrate auf 100 Mbit/s gesteigert wird. Das Zugriffsprotokoll ist identisch mit dem bei 10 Mbit/s, allerdings ist die maximale Ausdehnung auf 100 bis 150 m je Kabel beschränkt.

Parallel zu „100 Base T“ wurde das „100 VG-AnyLAN“ entwickelt, das als „802.12“ vom IEEE standardisiert wurde [30]. Das zugrundeliegende Medienzugriffsprotokoll ist ein zentra-

1. Ein Hub koppelt die angeschlossenen Netzsegmente auf der Bitübertragungsschicht zu einem logischen Busnetz

les Demand-Priority-Verfahren: Eine Station teilt dem Hub ihren Sendewunsch mit und darf erst dann senden, wenn sie eine Sendeberechtigung erhalten hat. Der Hub nimmt die Sendewünsche entgegen und teilt zyklisch die Sendeberechtigungen zu. „100 VG-AnyLAN“ ermöglicht zusätzlich eine Priorisierung einzelner Stationen, die dann sofort die Sendeberechtigung erhalten, sobald die laufende Übertragung beendet ist.

Die nächste Generation von Ethernet-Protokollen befindet sich derzeit in der Spezifikationsphase. Der Nachfolger von „100 Base T“ soll bei einer Datenrate von 1 Gbit/s arbeiten, für den Nachfolger von „100 VG-AnyLAN“ sind 0,5 bis 4 Gbit/s vorgesehen [83].

2.1.2 Token-LANs

Parallel zu Ethernet wurden zwei Token-LANs entwickelt: Token Bus nach IEEE 802.4 [35] und Token Ring nach IEEE 802.5 [36]. Das Token-Verfahren hat gegenüber Ethernet den Vorteil, daß keine Kollisionen auftreten können und die maximale Wartezeit für die Daten einer Station angebbar ist. Der Nachteil gegenüber Ethernet ist der Aufwand, der für die Token-Verwaltung notwendig ist sowie die höhere mittlere Wartezeit bei niedriger Last: Während eine Station bei Ethernet sofort auf das Medium zugreifen kann, wenn keine andere Station sendet, muß sie beim Token-Verfahren warten, bis das Token bei ihr angekommen ist.

Das Token stellt eine Sendeberechtigung dar, die von Station zu Station weitergereicht wird und genau einmal im Netz existiert. Eine Station, die das Token besitzt, darf Daten senden. Nach Ablauf einer vordefinierten Sendezeit muß sie das Token an die Nachfolgestation weiterleiten.

Die physikalische Struktur des Token Bus unterscheidet sich nur wenig vom Ethernet: Alle Stationen greifen auf ein gemeinsames physikalisches Medium zu. Die angeschlossenen Stationen bilden, unabhängig von ihrer physikalischen Position am Bus, einen logischen Ring. Jede Station hat einen logischen Vorgänger und einen Nachfolger, von dem sie das Token erhält bzw. an den sie das Token weitergibt. Die Bitrate kann 1, 5 oder 10 Mbit/s betragen.

Weiter verbreitet als der Token Bus ist der Token Ring, bei dem die beteiligten Stationen einen physikalischen Ring bilden. Die Datenrate beim Token Ring beträgt 4 oder 16 Mbit/s.

Mit dem Fiber Distributed Data Interface (FDDI) hat das ANSI ein Protokoll standardisiert, das als erstes Protokoll für lokale Netze eine Datenrate von 100 Mbit/s erlaubt [39, 40, 41]. Eine Weiterentwicklung des Protokolls, FDDI-II, ermöglicht neben dem Token-Verfahren für asynchronen Paketverkehr die Integration von isochronen Kanälen [42].

FDDI ist nicht beliebig skalierbar: Bei zunehmender Datenrate oder Ausdehnung nimmt die maximale Auslastung des Netzes ab [98]. Dies ist ein prinzipieller Nachteil des Token Ring-Protokolls, bei dem immer nur eine Station senden kann. Mit ihren Daten belegt die Station nur einen Teil des Rings, während die übrige Bandbreite verschwendet wird (Bild 2-2).

Isochronverkehr wird von FDDI-II zwar unterstützt, jedoch steht die für Isochronverkehr reservierte Bandbreite nicht mehr für den asynchronen Paketverkehr zur Verfügung. Während Tele-

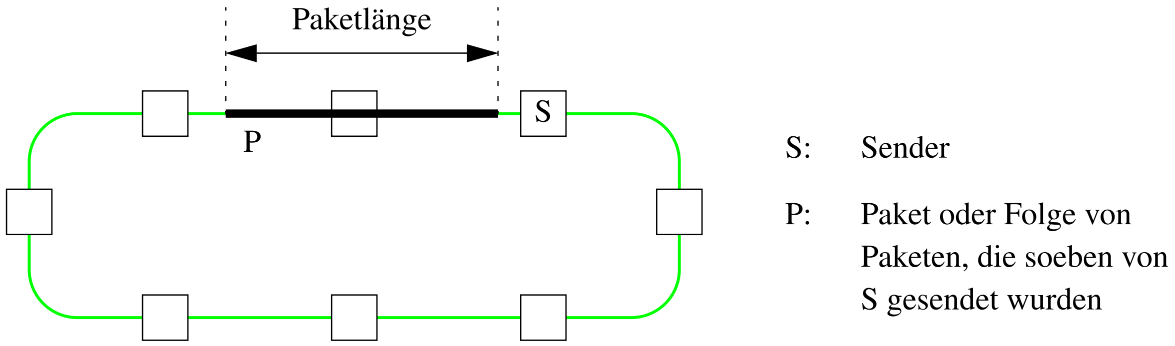


Bild 2-2: Ineffiziente Bandbreitennutzung in einem großen Token Ring

fonverkehr mit der Bitrate von ISDN nur einen Bruchteil der Gesamtbandbreite nutzt, wird der Asynchronverkehr auf dem Ring bereits durch komprimierte Video-Ströme, die eine Datenrate im Bereich mehrerer Mbit/s beanspruchen, erheblich beeinträchtigt.

2.2 Topologien für lokale Netze

Neben den bekannten Topologien wie Bus, Ring oder Stern gibt es weitere mögliche Topologien, die für lokale Netze geeignet sind. Die Topologie des physikalischen Netzes beschreibt im Gegensatz zur Topologie des logischen Netzes die Anordnung von Stationen und Leitungen. Eine logische Topologie kann auf eine andersgeartete physikalische Topologie abgebildet werden. Im folgenden werden Topologien und ihre Eigenschaften untersucht.

2.2.1 Linearer Bus

Die einfachste Struktur für ein lokales Netz ist der lineare Bus: An eine Leitung sind mehrere Stationen elektrisch angeschlossen. Der lineare Bus wird auch als „Broadcast Bus“ bezeichnet, weil ein von einer Station ausgesendetes elektrisches Signal von allen angeschlossenen Stationen empfangen wird.

Um hohe Datenraten zu erreichen, müssen die physikalischen Eigenschaften der Leitung betrachtet werden. An jeder Störstelle in einem Kabel, an der sich der Wellenwiderstand ändert, treten Reflexionen auf, die die Signalqualität mindern oder das Signal unbrauchbar machen. Aus diesem Grund müssen die Leitungsenden mit dem Wellenwiderstand abgeschlossen werden. Die Stationen müssen derart angeschlossen werden, daß sich der Wellenwiderstand an dieser Stelle möglichst wenig ändert. Auch Kabelbiegungen und Stecker verursachen Störungen, die jedoch im allgemeinen so gering sind, daß sie vernachlässigt werden können.

Wenn zwei Stationen gleichzeitig senden, kommt es zu Kollisionen. Eine Kollision entsteht durch die Überlagerung zweier oder mehrerer Signale unterschiedlicher Quellen, die sich gegenseitig stören, so daß keines der Signale mehr erkannt und korrekt empfangen werden kann. Bei Signalen kurzer Dauer ist es möglich, daß eine Kollision nicht von allen Stationen erkannt wird. Bild 2-3 zeigt die Entstehung einer Kollision auf dem Bus, wenn zwei Stationen gleichzeitig senden.

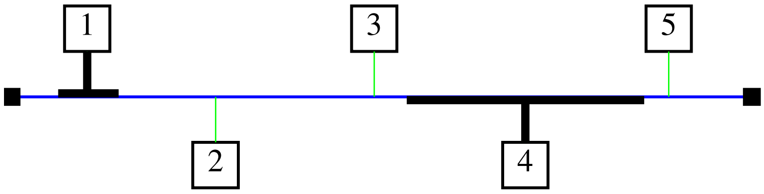


Bild 2-3a:
Station 1 und 4 senden gleichzeitig. Die Signale breiten sich jeweils in beide Richtungen aus.

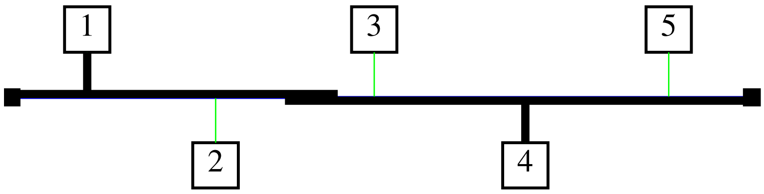


Bild 2-3b:
Kurze Zeit später: Die beiden Signale beginnen sich zwischen den Stationen 2 und 3 zu überlagern.

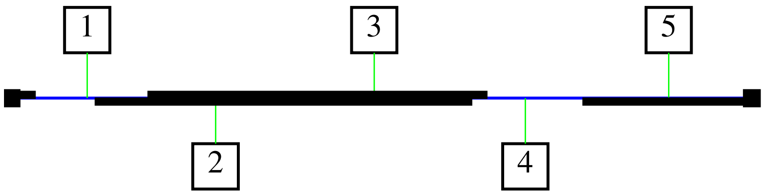


Bild 2-3c:
Die Stationen haben ihren Sendevorgang beendet. Es kommt zu einer Kollision bei den Stationen 2 und 3.

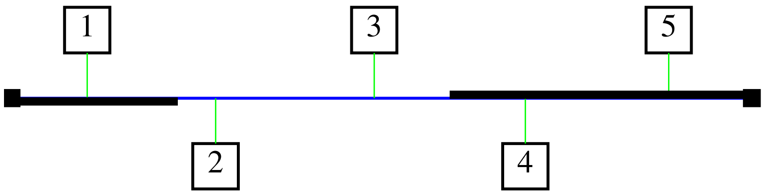


Bild 2-3d:
Die Kollision ist beendet, Station 1 empfängt das Signal von Station 4, die Stationen 4 und 5 das Signal von Station 1 fehlerfrei.

Bild 2-3: Kollisionsszenario

Ein Zugriffsprotokoll für den linearen Bus muß verhindern oder zumindest erkennen, daß beim Empfänger eine Kollision entsteht. Bei Ethernet wird durch eine Mindestpaketgröße und eine maximale Segmentlänge sichergestellt, daß ein Paket während des Sendens einmal den kompletten Bus belegt. Wenn eine oder mehrere weitere Stationen gleichzeitig senden, erkennen sie die Kollision und brechen den Sendevorgang ab. Andere Protokolle wie der Token Bus verhindern Kollisionen, indem sie sicherstellen, daß zu einem Zeitpunkt immer nur genau eine Station sendet. Ein Protokoll für einen Hochgeschwindigkeits-Bus könnte die Signallaufzeit nutzen und gleichzeitiges Senden zulassen. Wenn in Bild 2-3 zum Beispiel Station 4 an Station 1 sendet sowie Station 1 an Station 5, werden die Pakete trotz der Kollision bei den Stationen 2 und 3 korrekt übertragen.

Als Sonderfall linearer Busse können Funk-LANs betrachtet werden, bei denen sich das Signal nicht in einer Dimension entlang einer Leitung ausbreitet, sondern gleichzeitig in alle drei Raumdimensionen. Hierbei treten allerdings einige Besonderheiten auf. Da die Stationen unterschiedliche Entfernungen zueinander haben und die Signalamplitude mit dem Quadrat der Entfernung vom Sender abnimmt, werden die Signale nicht bei allen Stationen mit gleicher

Signalamplitude empfangen. Über Richtantennen können Signale außerdem stark gebündelt werden.

2.2.2 Faltbus und Dualbus

Im Gegensatz zum linearen Bus basieren der Faltbus und der Dualbus auf einer Kette von gerichteten Punkt-zu-Punkt-Verbindungen. Kollisionen durch physikalische Überlagerung mehrerer Signale treten nicht auf, jedoch kann es innerhalb der Stationen zur Blockierung zweier Datenströme kommen. Kommt ein Datenstrom am Eingang der Station an, während die Station gerade sendet, kann der ankommende Datenstrom in einem Zwischenspeicher so lange gepuffert werden, bis die Station ihren Sendevorgang beendet hat.

Ein einzelner gerichteter Bus ermöglicht Kommunikation nur in eine Richtung. Daher werden für LANs Busstrukturen wie der Falt- oder Dualbus eingesetzt, auf denen eine bidirektionale Kommunikation möglich ist.

Der Faltbus[1] besteht aus einer Kette vom Stationen, die jeweils durch zwei gegengerichtete Verbindungen gekoppelt sind (Bild 2-4). Die Kopfstation generiert die für den Datentransport nötigen Rahmenstrukturen auf dem vorwärts gerichteten Bus („Outbound Bus“). Am Ausgang der letzten Station wird der Datenstrom auf den rückwärts gerichteten Bus („Inbound Bus“) gefaltet und erreicht so die Kopfstation, in der die ankommenden Daten entfernt werden. Auf dem vorwärts gerichteten Schreibbus werden Daten geschrieben, die auf dem rückwärts gerichteten Lesebus vom Empfänger gelesen werden. Eine explizite Wegewahl ist nicht erforderlich, da jedes gesendete Datenpaket alle Stationen auf dem rückwärts gerichteten Bus erreicht.

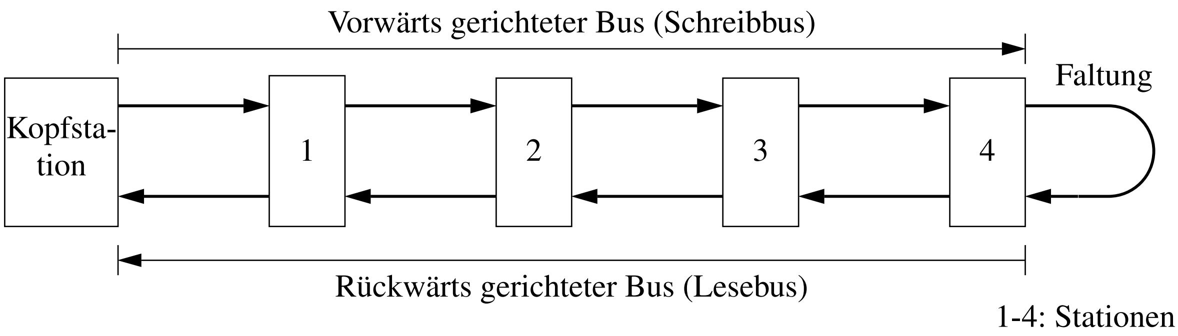


Bild 2-4: Faltbus

Der Dualbus weist im Gegensatz zum Faltbus zwei getrennte Busse auf (Bild 2-5). Die Struktur ist symmetrisch bezüglich der Busse. Die beiden Kopfstationen generieren jeweils die Rahmenstrukturen. Auf den beiden Bussen können unabhängig Daten gesendet werden, wodurch der maximal mögliche Gesamtdurchsatz über dem des Faltbusses liegt, bei dem nur auf einem Bus gesendet wird. Beim Dualbus ist allerdings eine Wegewahl erforderlich: Da das Ziel eines

1. Auch: Gefalteter Bus, englisch: Folded Bus

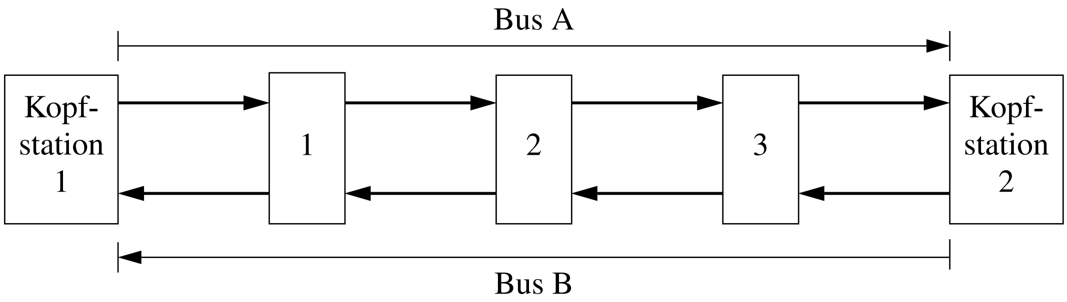
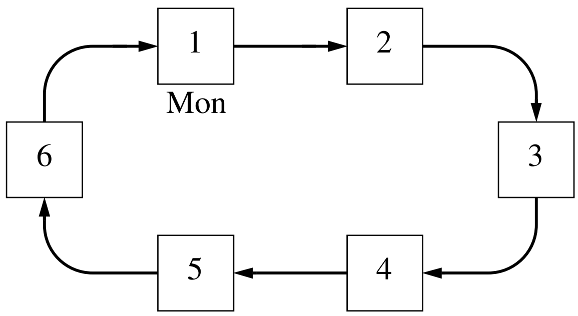


Bild 2-5: Dualbus

Datenpakets jeweils nur über einen der beiden Busse erreichbar ist, muß die Lage der Zielstation bekannt sein.

2.2.3 Ring, Doppelring und Torus

Der logische Ring kennt durch seine Drehsymmetrie zunächst keine bevorzugten oder benachteiligten Stationen (Bild 2-6). Eine Wegewahl ist nicht erforderlich: Jede Station besitzt nur einen Ausgangskanal, über den jedoch alle anderen Station erreichbar sind. Eine Ring-Monitorstation muß verhindern, daß Daten unendlich lange auf dem Ring umlaufen. Darüber hinaus übernimmt die Monitorstation Aufgaben wie Ringschluß oder die Konfiguration des Rings.



1-6: Stationen auf dem Ring
Mon: Monitorstation

Bild 2-6: Ring

Verschiedene Strategien sind möglich, um die Daten vom Ring zu entfernen. Bei der einfachsten Realisierung nimmt der Sender die Daten vom Ring. Dies bedeutet jedoch eine Verschwendung von Bandbreite, da der Ring zwischen Empfänger und Sender mit Daten belegt wird, die bereits empfangen und damit überflüssig geworden sind.

Beim „Destination Release“ entfernt der Empfänger die Daten vom Ring. Betrachtet man eine bidirektionale Verbindung zwischen zwei Stationen, dann können beide Stationen gleichzeitig mit der vollen Kanalbandbreite senden, wenn gerade keine andere Station sendet. In diesem Fall ist die Ringauslastung doppelt so hoch wie die Kanalbandbreite. Die maximale Ringauslastung beträgt bei n Stationen das n-fache der Kanalbandbreite. Dies ist dann der Fall, wenn alle Stationen jeweils an ihre Nachfolgestation senden.

Bei Punkt-zu-Mehrpunktverbindungen erreichen Daten durch die Ringstruktur alle Zielstationen, werden dort kopiert, kommen wieder zur Ausgangsstation zurück und werden von ihr entfernt. Eine etwas höhere Ringauslastung kann erzielt werden, indem die letzte Zielstation die Daten entfernt.

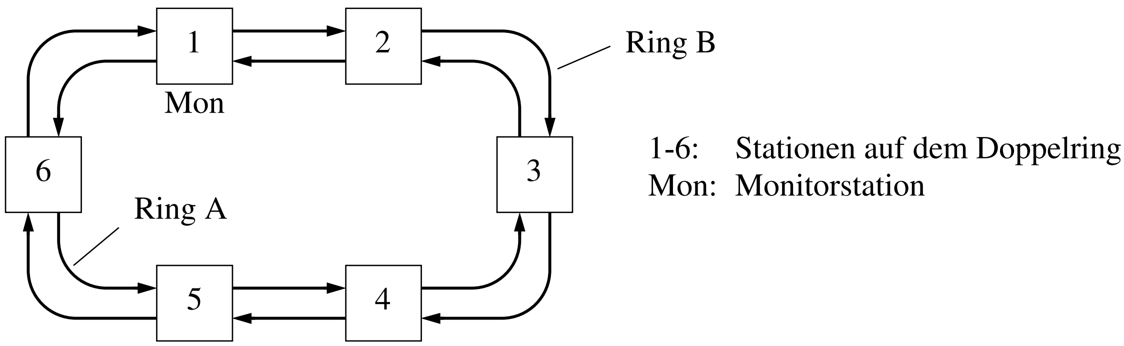


Bild 2-7: Doppelring

Der Doppelring (Bild 2-7) besteht aus zwei gegengerichteten Ringen. Diese Struktur wird in erster Linie verwendet, um die Ausfallsicherheit des Systems zu erhöhen. So wird bei FDDI der zweite Ring im Normalbetrieb nicht benutzt und dient lediglich zum Überbrücken einer ausgefallenen Leitung (Bild 2-8). Wird der zweite Ring dagegen für den Wirkbetrieb genutzt, kann die mittlere Ringauslastung bei symmetrischen Kommunikationsbeziehungen mit gleichverteiltem Verkehr gegenüber dem Einfachring verdoppelt oder, bei geeigneter Wegewahl, um den Faktor 4 gesteigert werden.

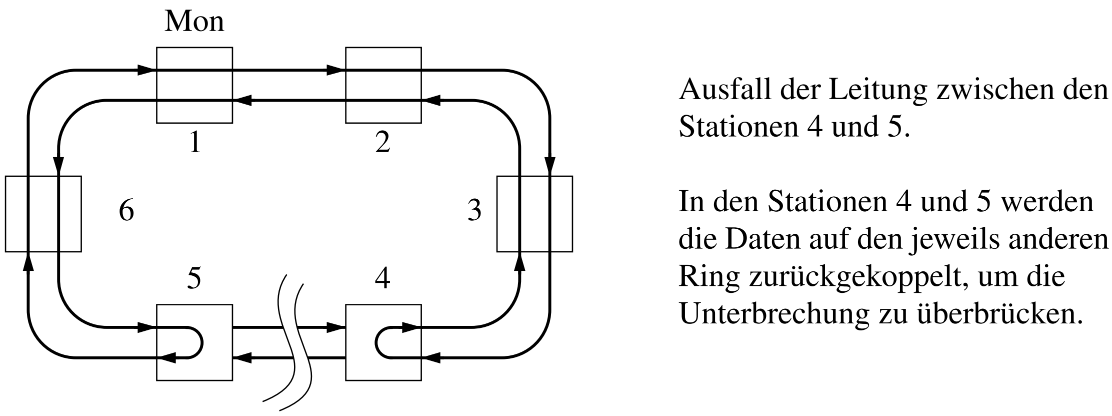


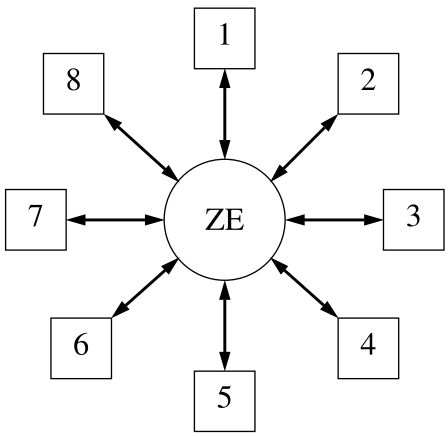
Bild 2-8: Rekonfiguration beim Doppelring im Fehlerfall

Ein Mehrfachring, der mehr als einen einfachen oder bidirektionalen Kanal zwischen den Stationen bereitstellt, wird als Torus bezeichnet. Sind alle Stationen gleichermaßen an den Torus angeschlossen, dienen die Kanäle lediglich der Erhöhung der Ringbandbreite. Daneben ist die Abbildung anderer logischer Strukturen, wie zum Beispiel von Shuffle-Netzen (Kapitel 2.2.6), auf den Torus möglich, indem die Kanäle in den einzelnen Stationen entsprechend der gewünschten Struktur vermittelt werden. Die Stationen senden und empfangen dann nur auf einem Teil der ankommenden bzw. abgehenden Kanäle, während die restlichen Kanäle unbearbeitet durchgeschaltet werden.

In Ringen mit Wellenlängen-Multiplexverfahren besteht der Torus aus Kanälen verschiedener Wellenlänge in einer Glasfaser. Ein voll- oder teilvermaschtes Netz kann darauf abgebildet werden, indem für jede gewünschte Verbindung zwischen zwei Stationen eine eigene Wellenlänge verwendet wird.

2.2.4 Stern

Die Sterntopologie mit ihrem zentralen Element und den darüber kommunizierenden Stationen zeigt Bild 2-9. Das zentrale Element kann als Sternkoppler, Hub oder Vermittlungsknoten ausgeführt sein. Diese Realisierungsformen unterscheiden sich in Funktionsumfang und Komplexität.



ZE: Zentrales Element des Sterns, über das alle Stationen kommunizieren

Bild 2-9: Sterntopologie

Der Sternkoppler ist ein passives zentrales Element und wird in der Wellenlängenmultiplex-Technik eingesetzt. Die Stationen sind über Glasfasern an den Sternkoppler angeschlossen und senden auf verschiedenen Wellenlängen. Im Sternkoppler werden die Wellenlängen überlagert. Das daraus gebildete Summensignal wird über die Glasfasern an alle angeschlossenen Stationen verteilt.

Der Hub ist ein aktives zentrales Element, das ein ankommendes Signal auf einem der Eingangskanäle erkennt, verstärkt und an alle anderen Kanäle weiterleitet. Versuchen zwei Stationen gleichzeitig zu senden, kommt es zu einer Kollision, die der Hub erkennt und den Stationen mitteilt. Eingesetzt wird der Hub zum Beispiel bei Ethernet „10 Base T“ [34] und „100 Base T“ [29].

Die Funktionalität eines Vermittlungsknotens („Switch“) geht über die reine Signalverstärkung hinaus. Ankommende Daten werden nur auf dem Ausgangskanal ausgegeben, für den sie bestimmt sind. Mehrere Datenströme an den Eingängen können, sofern sie für unterschiedliche Ausgänge bestimmt sind, gleichzeitig vermittelt und weitergegeben werden. Dieses Prinzip des Vermittlungsknotens wird seit Jahrzehnten in Telefonvermittlungsstellen angewandt.

Ein Vermittlungsknoten für paketierten Verkehr mit virtuellen Verbindungen verfügt zusätzlich über Pufferspeicher, um die statistische Verteilung von Paketankünften und Empfängern auszunutzen und mehrere virtuelle Verbindungen zwischen verschiedenen Stationen gleichzeitig auf einer Leitung zu ermöglichen. Ankommende Pakete werden im Pufferspeicher abgelegt, bis der benötigte Ausgangskanal zur Verfügung steht. Eine Vermittlungsstrategie wird eingesetzt, um eines aus mehreren gespeicherten Paketen für einen Ausgangskanal auszuwählen. Größere Vermittlungsknoten sind in Form von Koppelnetzen aufgebaut [61, 99].

2.2.5 Baum

Baumtopologien (Bild 2-10) unterstützen hierarchische Verkehrsbeziehungen mit überwiegend lokalem Verkehrsaufkommen zwischen denjenigen Stationen, die mit derselben Station der höheren Hierarchieebene verbunden sind. In einem lokalen Netz könnte ein Datei-Server einer Abteilung in der mittleren Hierarchieebene des Baumes die Abteilungsrechner in der unteren Hierarchieebene („Blätter“) bedienen. Die Abteilungsrechner wiederum greifen auf unternehmensweite Daten zu, die auf einem zentralen Rechner in der höchsten Hierarchieebene („Wurzel“) gespeichert sind.

Die ursprünglich aus Video-Verteilnetzen in Form von Arboreszensen (gerichteten Bäumen) abgeleiteten Zugangsnetze sind ebenfalls baumförmig aufgebaut. Die zusätzlich eingerichteten Rückkanäle ermöglichen die Interaktion der Teilnehmer mit den Anbietern über das Zugangsnetz.

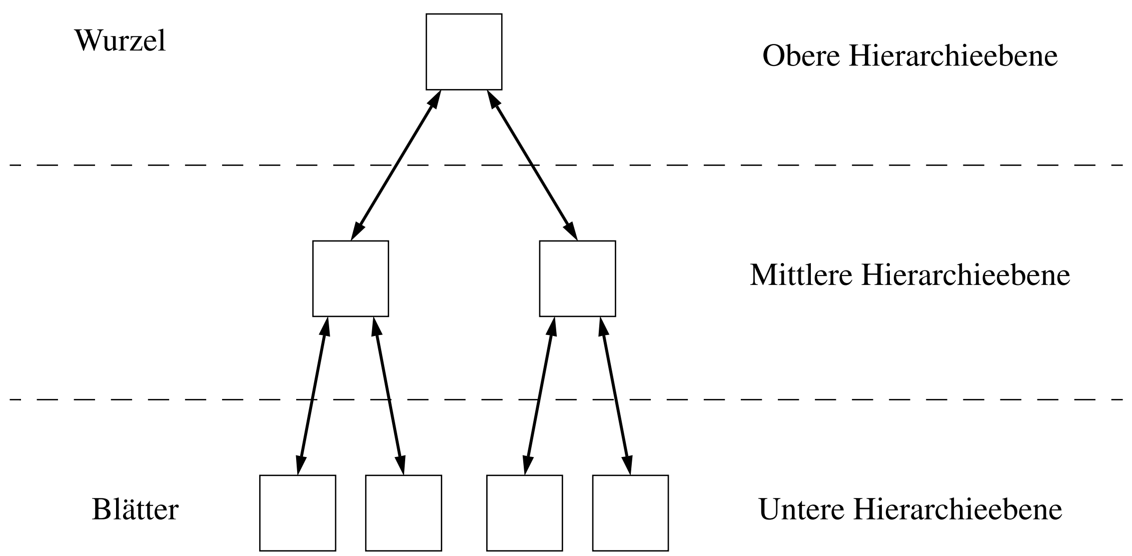


Bild 2-10: Baum

Lokale Netze mit Baumstruktur verwenden Verkehrslenkung in den mittleren und oberen Hierarchieebenen. Daten werden dabei nur so weit in Richtung Wurzel geschickt, bis der Empfänger in Richtung der unteren Hierarchieebene erreichbar ist.

2.2.6 Vermaschte Netze

Beim vollvermaschten Netz ist jede Station mit jeder anderen direkt verbunden. Der Aufwand für diese Topologie ist proportional dem Quadrat der Stationsanzahl. Die Vollvermaschung ist folglich wenig geeignet für lokale Netze mit einer großen Anzahl von Stationen (Bild 2-11).

Teilvermaschte Netze zeichnen sich dadurch aus, daß sie mehrere mögliche Wege, jedoch nicht immer einen direkten Weg zwischen zwei Stationen bieten. Verkehrslenkungsstrategien werden angewandt, um bei gegebenen Randbedingungen wie Weglänge, Auslastung eines Kanals, Verzögerungszeit oder Ausfall einer Verbindung eine möglichst optimale Wegewahl zu treffen.

Für regelmäßige, teilvermaschte Netze existieren diverse Ansätze verallgemeinerter Verkehrslenkungsalgorithmen. Diese Verfahren lassen sich auf Netze unterschiedlicher, zumeist sogar beliebiger Größe anwenden. Derartige teilvermaschte Lokalnetze werden in der Regel nicht physikalisch aufgebaut, sondern als logische Netzstruktur unter Verwendung von Wellenlängenmultiplex zum Beispiel in einem physikalischen Ring abgebildet. Die beiden bekanntesten regelmäßigen, teilvermaschten Netzstrukturen sind das Shuffle-Netz und das Manhattan-Street-Netz.

Das Shuffle-Netz ermöglicht kurze Wege bei einfacher Verkehrslenkung. Im Beispiel des zweistufigen Shuffle-Netzes nach Bild 2-12 verfügt jede Station über jeweils zwei Ein- und Ausgangskanäle (Die erste Stufe, bestehend aus den Stationen 1-4, ist zugunsten der Übersichtlichkeit doppelt gezeichnet). Die Daten erreichen dabei über null (Beispiel: Verbindung zwischen Station 1 und 6 auf direktem Weg ohne Zwischenstation) bis maximal zwei (Beispiel: Verbindung zwischen Station 1 und 8 auf den möglichen Wegen 1-5-2-8 oder 1-6-4-8) Zwischenstationen ihr Ziel. Allgemeine Shuffle-Netze können ein- oder mehrstufig sein, wobei jede Stufe die gleiche Anzahl von Stationen aufweist.

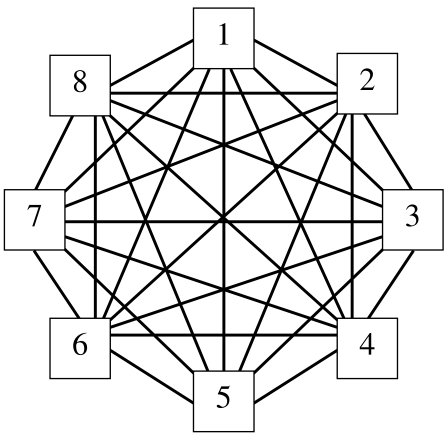


Bild 2-11: Vollvermaschtes Netz

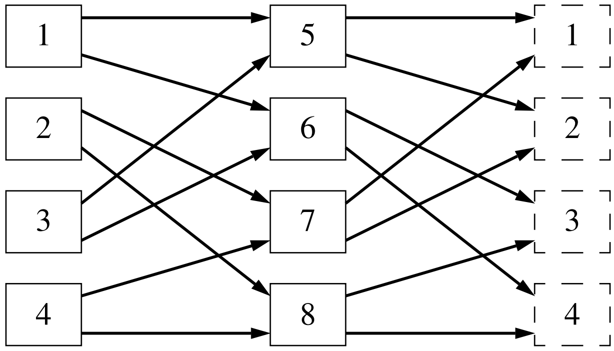


Bild 2-12: Zweistufiges gerichtetes Shuffle-Netz mit 8 Stationen

Manhattan Street-Netze bestehen aus gitterförmig angeordneten Stationen, wobei jeweils die direkt benachbarten Stationen verbunden sind. Aus Symmetriegründen können auch die gegenüberliegenden Endknoten des Gitters miteinander verbunden sein. Die Kanäle zwischen den Stationen sind bidirektional oder unidirektional (Bild 2-13).

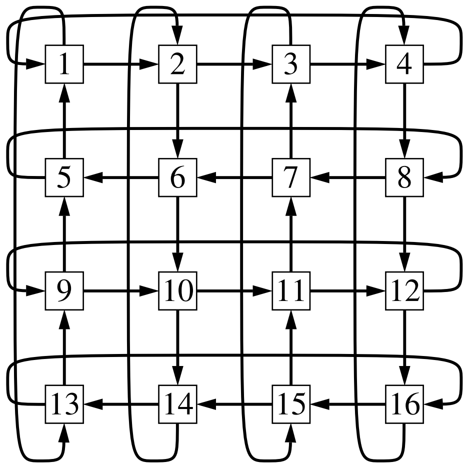


Bild 2-13: Manhattan Street-Netz mit 16 Stationen und unidirektionalen Kanälen

Für Manhattan Street-Netze existieren zahlreiche Verkehrslenkungsverfahren, die zumeist auf der einfachen und regelmäßigen Struktur des Netzes beruhen. So sind bei einem Manhattan Street-Netz mit bidirektionalen Kanälen mehrere Wege innerhalb des kleinsten Rechtecks möglich, dessen Eckpunkte die Quell- und Zielstation bilden und in dem die kürzesten Wege liegen. Bei unidirektionalen Kanälen muß dieses Rechteck teilweise verlassen werden, um das Ziel zu erreichen. In Bild 2-13 ist zum Beispiel keine Verbindung von Station 1 nach Station 7 möglich, ohne das durch diese Punkte aufgespannte Rechteck zu verlassen, da Station 7 nur von rechts oder unten erreichbar ist.

Ein Vorteil von Shuffle- oder Manhattan Street-Netzen ist, daß das Netz unabhängig von der Netzgröße mit einem einheitlichen Stationstyp aufgebaut werden kann, der über jeweils 2 Ein-und 2 Ausgangskanäle verfügt. Für eine Änderung der Stationszahl im Netz kann allerdings eine komplette Umkonfiguration des Netzes erforderlich werden.

2.2.7 Hybridstrukturen

Netze, die aus einer Kombination mehrerer verschiedener Topologien aufgebaut sind, werden als Hybrid-Netze bezeichnet. Derartige Netze entstehen entweder durch Evolution eines bestehenden Netzes oder sind in der Form geplant, um die Netztopologie den Erfordernissen optimal anzupassen.

Werden einzelne, heterogene Netze gekoppelt, kommt häufig ein „Backbone[1]"-Netz zum Einsatz. FDDI als Backbone bietet gegenüber Ethernet eine zehnfach höhere Bandbreite, so daß die Kopplung der einzelnen Netze nicht zu einem Engpaß wird. Zentrale Server-Rechner werden direkt an das Backbone-Netz angeschlossen und haben so die aus der Summe der Anforderungen aus den angeschlossenen Netzen resultierende höhere Bandbreite zur Verfügung.

Baumförmige Netze in Verbindung mit Vermittlungsknoten an den Blättern eignen sich zur Verkabelung von hierarchischen LAN-Strukturen wie Gebäuden, Abteilungen mit Unterabteilungen (Bild 2-14) sowie für Zugangsnetze.

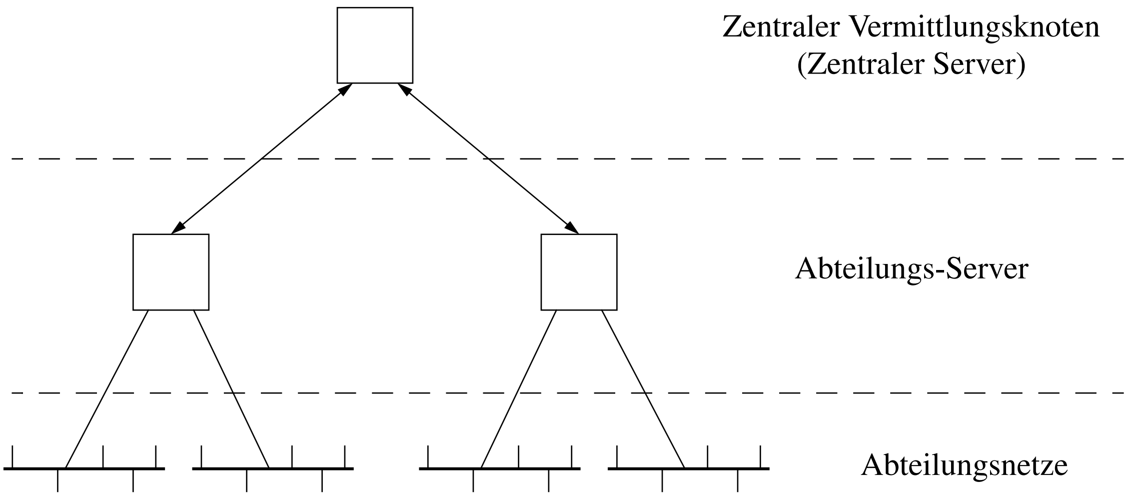


Bild 2-14: Beispiel eines Firmennetzes

2.3 Grundprinzipien von Medienzugriffsprotokollen

Medienzugriffsprotokolle regeln den Zugriff auf ein gemeinsames Übertragungsmedium. Sie stellen sicher, daß nicht zwei Teilnehmer gleichzeitig auf das Medium zugreifen und sich gegenseitig stören oder daß eine gemeinsame Nutzung und die daraus resultierende Überlagerung der Signale erkannt wird. Außerdem stellen Medienzugriffsprotokolle einen bestimmten Grad von Fairness sicher, um die zur Verfügung stehende Gesamtbandbreite „gerecht" auf die

1. „Rückgrat"

Stationen aufzuteilen sowie Dienstgüteparameter bezüglich Obergrenze und Schwankung der Verzögerungszeiten zu garantieren.

Beim Vielfachzugriff sind zwei prinzipiell verschiedene Verfahren zu unterscheiden: Der Zugriff auf ein gemeinsames physikalisches Medium oder der serialisierte, verteilte Vielfachzugriff über Multiplexer in sequentiell angeordneten Stationen. Während im ersten Fall aufgrund der Signallaufzeit auf dem Medium Kollisionen möglich sind, kann im zweiten Fall innerhalb der Stationen eine bevorstehende Kollision unmittelbar erkannt und verhindert werden.

Die nachfolgend beschriebenen Grundprinzipien von Medienzugriffsprotokollen können vielfältig kombiniert werden, um ein Medienzugriffsprotokoll zu entwickeln. Sie bilden eine Art „Grundbaukastensystem" für neue Medienzugriffsprotokolle, auf deren Grundlage neue Zugriffsverfahren eingeführt werden können.

2.3.1 Zeitkontinuierlicher und zeitdiskreter Medienzugriff

Der zeitkontinuierliche Medienzugriff erlaubt einer Station den Zugriff auf das Medium zu jedem beliebigen Zeitpunkt nach Maßgabe des Zugriffsprotokolls. Beispiele hierfür sind Ethernet oder FDDI. Ein Nachteil dieses Verfahrens besteht darin, daß ohne besondere Vorkehrungsmaßnahmen ein begonnener Sendevorgang dadurch gestört werden kann, daß eine andere Station bereits zu senden begonnen hat, diese Tatsache aber aufgrund der Laufzeit auf dem Medium beim Sender noch nicht bekannt ist.

Beim zeitdiskreten Medienzugriff darf der Sendevorgang nur zu vorbestimmten Zeitpunkten beginnen. Ferner muß die Übertragung zu Beginn des nächsten möglichen Sendezeitpunkts beendet sein.

2.3.2 Multiplexverfahren und Rahmenstrukturen

Greifen mehrere Stationen gleichzeitig auf ein physikalisches Medium zu, so ist die Aufteilung der zur Verfügung stehenden Gesamtbandbreite durch Raum-, Zeit-, Frequenz[1]-, Code- oder Wellenlängenmultiplex[2] möglich. Frequenz- und Codemultiplex werden überwiegend in Funknetzen verwendet, bei denen der freie Raum das gemeinsame physikalische Übertragungsmedium bildet. Raum- und Wellenlängenmultiplex ermöglichen die parallele Führung mehrerer unabhängiger Signale in einem Kabel.

Das in der elektronischen Digitaltechnik am leichtesten auf einer Leitung zu realisierende Multiplexverfahren ist das synchrone Zeitmultiplexverfahren. Dabei wird der zur Verfügung stehende Übertragungskanal in Zeitabschnitte eingeteilt, die periodisch wiederholt jeweils einen eigenen Kanal bilden. Einzelne Zeitmultiplex-Kanäle, die jeweils aus gleich großen,

---

1. Frequenzmultiplex bezeichnet die Überlagerung von modulierten elektrischen Signalen, wobei das Unterscheidungskriterium für die verschiedenen Kanäle die Frequenz des Trägersignals ist
2. Wellenlängenmultiplex bezeichnet die Überlagerung von amplitudenmodulierten optischen Signalen, wobei das Unterscheidungskriterium für die verschiedenen Kanäle die Wellenlänge des optischen Trägersignals ist

direkt aufeinanderfolgenden Zeitabschnitten bestehen, werden als Zeitschlitze („Slots“) bezeichnet.

Ein Rahmen faßt mehrere Zeitschlitze zu einer Einheit zusammen und ermöglicht eine flexible Aufteilung der zur Verfügung stehenden Gesamtbandbreite. Die Wiederholperiode für die einzelnen Zeitmultiplexkanäle entspricht der Rahmendauer, die häufig aus der ISDN-Technik abgeleitet wird und 125 µs beträgt. Einfache Rahmenstrukturen bestehen aus einem Rahmenkopf und daran angehängten Zeitschlitzen. Komplexe Rahmenstrukturen beinhalten zusätzliche Management-Informationen und ermöglichen ein hierarchisches Zeitmultiplexen von Datenströmen [44, 45].

Im Gegensatz zu den beschriebenen Multiplexverfahren ist die Bandbreitenzuteilung beim statistischen Multiplexen abhängig vom momentanen Verkehrsaufkommen. Es kann bei büschelhaft sendenden Verkehrsquellen verwendet werden, um die verfügbare Bandbreite effektiver auszunutzen, und setzt auf einem der beschriebenen Multiplexverfahren auf. Die Summe der mittleren Übertragungsraten der einzelnen Kanäle ist kleiner als die Medienbitrate, während jedoch die Summe der Spitzenbitraten größer als die Medienbitrate sein kann. Damit wird ein Kanal höher ausgelastet als bei der Zuteilung nach Spitzenbitraten. Aufgrund endlicher Puffergrößen können Verluste auftreten, wenn auf zu vielen Kanälen gleichzeitig mit der Spitzenbitrate gesendet wird. Statistisches Multiplexen wird angewendet im Zusammenhang mit Zeitmultiplex bei ATM (siehe Kapitel 3).

2.3.3 Erkennen des Belegungszustands

Greifen mehrere Stationen auf einen gemeinsamen physikalischen Kanal zu, muß es möglich sein, den Belegungszustand des Kanals zu ermitteln. Der Belegungszustand eines Mediums mit zeitkontinuierlichem Zugriff wird durch Beobachtung des physikalischen Pegels ermittelt („Carrier Sense“). Bei Ethernet ergibt sich durch das Manchester-Codierverfahren eine Änderung des mittleren Spannungspegels auf der Leitung, so daß mit einem geeignet dimensionierten Tiefpaß ein „Belegt“-Signal generiert werden kann.

Zeitmultiplex-Systeme sehen eine Kennung in jedem Zeitschlitz vor, die anzeigt, ob der Zeitschlitz belegt ist. Nur freie Zeitschlitze können von der Station genutzt werden. Da die Station bereits vor dem Senden ihrer Daten wissen muß, ob der Zeitschlitz belegt ist, die Information darüber andererseits erst in diesem Zeitschlitz verfügbar ist, muß in einem kausalen System der Datenfluß um die Zeit, die zwischen Beginn des Zeitschlitzes und der Erkennung des Belegungszustands liegt, verzögert werden.

Die einfachste Möglichkeit, den Belegungszustand in einem Zeitschlitz anzuzeigen, ist hierfür ein Bit im Kopf des Zeitschlitzes vorzusehen („Busy-Bit“). Weitere Möglichkeiten sind die Reservierung eines bestimmten Werts im Adressfeld (z. B. alle Adressbits = 0) für „unbelegt“ oder die Signalisierung außerhalb des Datenpfades.

2.3.4 Entfernen der Daten vom Medium

Bei Netzen mit Ringstruktur muß sichergestellt werden, daß gesendete Daten nicht unendlich lange auf dem Ring zirkulieren. Eine einfache Möglichkeit besteht darin, daß jede Station die von ihr gesendeten Daten selbst vom Ring entfernt. Bei Topologien ohne geschlossenen Datenpfad (z. B. gerichteten Bussen) endet der Datenfluß ohnehin bei der letzten Station.

Der Durchsatz des Gesamtsystems läßt sich steigern, indem die empfangende Station die Daten entfernt: Die nachfolgenden Stationen können das freigegebene Medium für ihren Verkehr benutzen. Dieser Mechanismus wird als „Destination Release" bezeichnet und kommt bei Punkt-zu-Punkt-Verbindungen zum Einsatz. Eine Ausdehnung des „Destination Release" auf Punkt-zu-Mehrpunkt-Verbindungen ist möglich, indem die letzte derjenigen Stationen, für die eine Dateneinheit bestimmt ist, diese Daten vom Medium entfernt.

2.3.5 Einfügepuffer

Ein Einfügepuffer (Insertion Buffer) ermöglicht einer Station, auch bei belegtem Medium zu senden. Aus logischer Sicht wird das Medium aufgetrennt und so weit verlängert, daß die zu sendenden Daten auf diesem eingefügten Segment Platz finden. Realisiert wird die Verlängerung des Mediums mit Hilfe eines Einfügepuffers, in den die ankommenden Daten geleitet werden, während die Station sendet. Der Füllstand des Einfügepuffers entspricht der Verlängerung des Mediums und bestimmt somit wesentlich die Stationsverzögerungszeit: Ein ankommender Datenblock wird am letzten Warteplatz eingereiht und muß warten, bis alle vorherigen Daten gesendet sind. Damit ergibt sich durch den Einfügepuffer eine variable Verzögerungszeit, die zum Füllstand proportional ist.

Das Medienzugriffsprotokoll muß sicherstellen, daß eine Station entweder nur dann sendet, wenn ihr Einfügepuffer nicht voll ist, oder daß während des Sendevorgangs nur die Datenmenge ankommt, die im Einfügepuffer Platz findet.

2.3.6 Fairness

Beim einfachsten denkbaren Zugriffsprotokoll für ein auf Zeitschlitzen basierendes System belegt eine sendewillige Station den nächsten freien Zeitschlitz. In einem System mit hoher Auslastung kann es ohne ein Fairness-Protokoll zu einer Bevorzugung einzelner Stationen kommen. Im Extremfall können manche Stationen nicht mehr senden, da sie keinen freien Zeitschlitz mehr vorfinden.

„Fairness" bezeichnet einen Mechanismus innerhalb des Medienzugriffsprotokolls, der die Aufteilung der zur Verfügung stehenden Bandbreite auf die Stationen regelt. Im einfachsten Fall besteht das Ziel darin, daß der Gesamtdurchsatz für alle sendewilligen Stationen über einen längeren Zeitraum gemittelt gleich ist. In einem Client-Server-Netz ist die systematische Bevorzugung einer Station im Rahmen des Fairness-Mechanismus für den Server sinnvoll, da dieser die Anforderungen aller Clients erfüllen muß und somit einen erhöhten Bandbreitenbedarf hat.

Man unterscheidet zwischen „globaler“ und „lokaler“ Fairness. Bezieht sich die Fairness auf alle Stationen im Netz gleichermaßen, spricht man von globaler Fairness. Dabei wird die Senderate unnötigerweise auch bei Stationen begrenzt, deren Verkehr keine direkten Auswirkungen auf die mögliche Senderate anderer Stationen hat. So wird Station 1 in Bild 2-15a durch den globalen Mechanismus begrenzt, obwohl keine andere Station das durch ihren Verkehr belegte Medium zwischen den Stationen 1 und 2 beansprucht. Lokale Fairness nutzt diese asymmetrischen Verkehrsbeziehungen derart aus, daß nur diejenigen Stationen gedrosselt werden, die einen Kanal gemeinsam nutzen (Bild 2-15b).

Die Fairness-Mechanismen unterscheiden sich ferner im Einschwingverhalten, also im Verhalten von Stationen, die zu senden beginnen. Insbesondere bei büschelhaftem Verkehr ist die Zeitdauer entscheidend, bis Fairness-Mechanismen die vorübergehende Unfairness ausgleichen. Weitere Bewertungskriterien für Fairness-Mechanismen sind die Bandbreitenausnutzung, mittlere und maximale Wartezeiten in den Stationen sowie die Transferzeitverteilung.

Im folgenden werden einige Mechanismen vorgestellt, die Fairness-Protokollen zugrundeliegen. Dabei wird davon ausgegangen, daß sich die Stationen das Medium im Zeitmultiplex-Verfahren teilen. Die Übertragung der Mechanismen auf andere Multiplexverfahren ist möglich, wird jedoch nicht näher betrachtet.

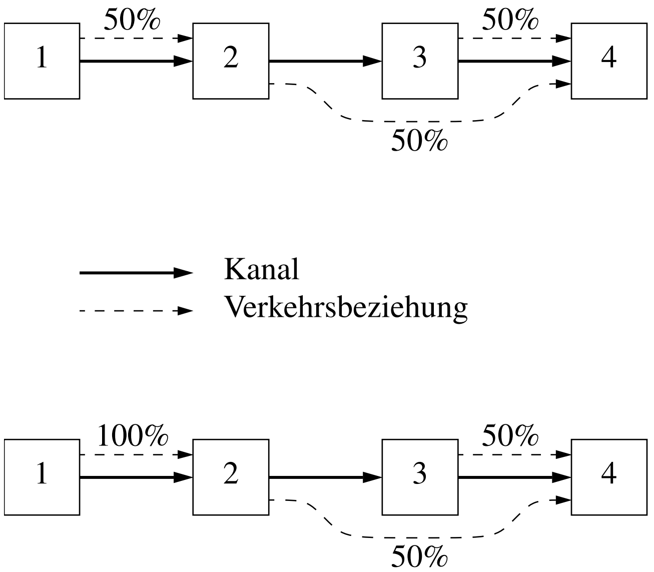


Bild 2-15a: Globale Fairness

Station 1 sendet an 2, Stationen 2 und 3 senden an 4 und teilen sich die Bandbreite auf Kanal 3-4. Alle Stationen haben gleichen Durchsatz. Kanal 1-2 ist nur zu 50% ausgelastet. Der Gesamtdurchsatz im Beispiel beträgt 150% der Kanalbandbreite.

Bild 2-15b: Lokale Fairness

Verkehrsbeziehungen wie oben, Station 1 nutzt jedoch den Kanal 1-2 zu 100%, da keine andere Station den Kanal beansprucht. Der Gesamtdurchsatz im Beispiel beträgt 200% der Kanalbandbreite.

Bild 2-15: Beispiel globaler und lokaler Fairness

2.3.6.1 Reservierung

Bei Fairness-Mechanismen mit Reservierung fordert eine sendewillige Station Bandbreite entsprechend ihrer momentan zur Übertragung anstehenden Datenmenge von einem zentralen Scheduler an. Der Scheduler teilt die verfügbare Bandbreite auf und bestätigt den Stationen die angeforderte oder eine geringere Datenmenge, die sie während des folgenden Reservierungszyklus senden dürfen.

Durch die notwendige vorherige Reservierung und das Warten auf den nächsten Zyklus wird die Zugriffsverzögerung auf das Medium selbst bei einem gering ausgelasteten System sehr hoch. Das Reservierungsverfahren wird daher meist mit anderen Fairness-Mechanismen kombiniert.

In Systemen, die auf Zeitschlitzen beruhen, kann für eine Station die angeforderte Bandbreite dadurch garantiert werden, daß die entsprechende Anzahl Zeitschlitze zu ihrer exklusiven Nutzung reserviert wird.

2.3.6.2 Sendeaufruf- und Token-Verfahren

Beim Sendeaufrufverfahren („Polling“) fragt eine zentralen Instanz in einer vorgegebenen Reihenfolge die Stationen nach ihren Sendewünschen ab. Sendewilligen Stationen wird das Übertragungsmedium für eine vordefinierte, jedoch nicht zwingend für alle Stationen gleiche maximale Zeitdauer zugeteilt, bevor die nächste Station abgefragt wird.

Das Token-Verfahren stellt eine verteilte Variante des Sendeaufrufverfahrens dar. An die Stelle der zentralen Instanz tritt das Token, das die Sendeberechtigung repräsentiert. Eine Station, die das Token erhält, kann für eine vorgegebene Zeit senden und muß das Token anschließend weitergeben.

Das Sendeaufrufverfahren erlaubt zu einem Zeitpunkt immer nur einer Station, auf dem Medium zu senden. Während das Token an die Nachfolgestation weitergegeben bzw. die Nachfolgestation aufgerufen wird, kann keine Station senden. Bei hoher Bitrate oder großer Ausdehnung des Netzes wird das Verfahren dadurch ineffizient.

2.3.6.3 Credit-Verfahren

Beim Credit-Verfahren erhalten Stationen die Berechtigung, eine bestimmte Anzahl von Zeitschlitzen zu belegen. Hierzu erhalten sie die entsprechende Anzahl von Credits. Für jeden Zeitschlitz, den sie belegt, verbraucht eine Station jeweils eines ihrer Credits. Sobald keine Station mehr senden will oder über keine Credits mehr verfügt, werden allen Stationen neue Credits vergeben. Die Zeit zwischen zwei aufeinanderfolgenden Credit-Vergaben wird als Credit-Zyklus bezeichnet.

Eine mögliche Realisierungsform des Credit-Verfahrens ist ein zentraler Mechanismus, bei dem eine Master-Station die Vergabe neuer Credits einleitet. Ein von der Masterstation ausgehendes, zirkulierendes Anforderungssignal wird von einer Station genau dann modifiziert, wenn sie noch über Credits verfügt und Daten zu senden hat („aktive“ Station). Stellt die Master-Station fest, daß keine Station mehr aktiv ist, leitet sie einen neuen Credit-Zyklus ein, indem sie die Credit-Zähler aller Stationen auf ihren Anfangswert zurücksetzt („Reset“).

Das Credit-Verfahren ist um so effizienter, je größer das Intervall zwischen zwei Credit-Zuteilungszeitpunkten bzw. je größer die Anzahl der zugeteilten Credits im Verhältnis zur Ausdehnung des Mediums ist. Ein verteiltes Verfahren zur Reset-Auslösung, bei dem die letzte aktive Station den Reset generiert, ist effizienter als das zentrale Verfahren, benötigt jedoch mehr Overhead für das Anforderungssignal, da die Stationskennung mit übertragen werden muß (siehe Abschnitt 2.4.3).

2.3.6.4 Gratis-Zugriff

„Gratis-Zugriff“ erlaubt den Stationen, das Medium unter Umgehung des eigentlichen Fairness-Verfahrens zu benutzen. Der Gratis-Zugriff soll im Fall von Niedriglast die Effizienz erhöhen, indem die durch das Fairness-Protokoll bedingten Warte- und Totzeiten vermieden werden. Bei durch erhöhte Last drohender Unfairness tritt das Fairness-Protokoll wieder an die Stelle des Gratis-Zugriffs.

Das Erkennen des Übergangs vom Niedriglast- zum Normalzustand ist problematisch, da er mit verteilter Verkehrsmessung erfaßt und möglichst schnell erkannt werden muß, um eine drohende lokale Unfairness abzuwenden.

2.3.6.5 Deferment

Das „Deferment“-Verfahren erlaubt zunächst allen Stationen den Gratis-Zugriff auf das Medium. Wird ein Schwellwert überschritten, werden die Stationen, die zu viel gesendet haben, nachträglich bestraft, indem ihnen vorübergehend die Sendeberechtigung entzogen wird. Die jeweilige Station darf erst wieder senden, nachdem eine bestimmte Anzahl freier Zeitschlitze ungenutzt die Station passiert hat.

2.3.6.6 Bitratenbegrenzung

Bei der Bitratenbegrenzung („Rate Control“) werden die Spitzenbitrate sowie die mittlere Bitrate einer Verbindung begrenzt. Die Stationen müssen sicherstellen, daß diese Grenzen nicht überschritten werden. Voraussetzung für die Bitratenbegrenzung ist ein expliziter Verbindungsaufbau, wobei eine zentrale Instanz das Wissen über die noch verfügbare Bandbreite besitzt. Die Vergabe der Bitraten erfolgt durch die zentrale Instanz beim Verbindungsaufbau aufgrund der Information über die Art der gewünschten Verbindung und die bestehenden Verbindungen.

Bei der Verbindungsannahme kann das statistische Multiplexverfahren zugrundegelegt werden, bei dem die Summe der Spitzenbitraten aller akzeptierten Verbindungen die Bitrate des Kanals übersteigt. Eine Obergrenze für die Wartezeit kann dann jedoch nicht mehr garantiert werden. Ferner entstehen Verluste, deren Wahrscheinlichkeit zunimmt, je mehr sich die Summe der zugeteilten mittleren Bitraten der Medienbitrate annähert.

2.3.7 Fehlererkennung und Fehlerkorrektur

Eine hohe Betriebssicherheit eines Netzes erfordert eine rasche Reaktion auf mögliche Fehlerfälle. Neben Bitfehlern, die aufgrund von Störeinflüssen auf Leitungen entstehen können, sind die Auswirkungen zu berücksichtigen, die durch die Unterbrechung einer Leitung, das Ausschalten einer Station oder eine Fehlkonfiguration verursacht werden.

Bitfehler können durch Hinzufügen redundanter Bitstellen und geeignete Codierverfahren (z. B. mit Hilfe zyklischer Binärcodes) erkannt werden [60]. Eine Fehlerkorrektur erfolgt entweder durch Vorwärtskorrektur oder durch erneutes Übertragen der gestörten Dateneinheit.

Eine automatische Fehlerkorrektur im Falle der Unterbrechung einer Leitung setzt einen alternativen Weg voraus. Bei FDDI und anderen Protokollvorschlägen mit Ringstruktur dient ein zweiter Ring als Ersatzweg (siehe Bild 2-8). Bei vermaschten Netzen stehen ohnehin meist mehrere Wege zur Verfügung.

In einem Ring muß verhindert werden, daß Daten, die aufgrund eines Fehlers von keiner Station vom Ring entfernt werden, nicht für alle Zeiten zirkulieren. Hierzu wird in jeder Dateneinheit ein „Monitor“-Bit eingeführt, das vom Sender der Dateneinheit zurückgesetzt wird. Eine Master-Station im Ring setzt das Bit in jeder passierenden Dateneinheit. Spätestens während des folgenden Ringumlaufs muß die Dateneinheit ihren Empfänger erreichen und vom Ring entfernt werden. An dem gesetzten Monitor-Bit in einer nicht entfernten Dateneinheit erkennt die Master-Station den Fehlerfall und kann entsprechend reagieren, z. B. durch das Entfernen der Dateneinheit.

2.3.8 Bitsynchronisation

Bei Netzen mit Vielfachzugriff kann der Bittakt zwischen zwei Station auf zwei Arten generiert werden: In einem plesiochronen Netz hat jede Station ihren eigenen Systemtakt, während es in einem synchronen System einen zentralen Taktgenerator gibt, von dem der Bittakt aller Stationen abgeleitet wird.

Plesiochrone Übertragung bietet den Vorteil sehr geringer Taktschwankungen („Jitter“) auf einer Leitung, da der Takt mit einem Quarz lokal generiert wird. Der systematische und unausweichliche Taktunterschied (Schlupf) zur Vorgängerstation muß ausgeglichen werden. Hierzu werden im plesiochronen Datenstrom Stopfbits vorgesehen, die nach Bedarf eingefügt oder weggelassen werden können und somit die mittlere Nutzbit-Taktrate der des Vorgängers anpassen.

Bei synchroner Übertragung wird der Stationstakt aus dem ankommenden Bitstrom mit einer PLL-Schaltung zurückgewonnen. In jeder Station entstehen durch das Nachregeln des Takts und durch physikalische Einflüsse zusätzliche Taktschwankungen, die sich über die Stationen aufsummieren. Ein systematischer Taktunterschied zwischen den Stationen existiert im synchronen Netz nicht, jedoch führt eine Störung in der Takterzeugung oder -ausbreitung zu einem Ausfall des Gesamt- oder eines Teilsystems.

Eine Taktverteilung über ein vom Nutzkanal getrenntes Taktverteilnetz ist insbesondere bei einer Realisierung mehrerer Komponenten in einem System sinnvoll.

2.4 Übersicht aktueller HSLAN-Protokolle

Als High-Speed LANs (HSLANs) werden Netze bezeichnet, die für sehr hohe Bitraten derart skalierbar sind, daß die maximale Ausnutzung des Netzes möglichst proportional zur Medienbitrate ist. Die in den folgenden Abschnitten beschriebenen Protokollansätze bilden eine Auswahl veröffentlichter und untersuchter Medienzugriffsprotokolle für Hochgeschwindigkeits-Lokalnetze, die für eine Integration von zeitkritischem und -unkritischem Verkehr geeignet sind. Weitere Protokollvorschläge werden in Übersichtsveröffentlichungen nach Gesichtspunkten wie den zugrundeliegenden Protokollmechanismen [64] beleuchtet und kurz beschrieben [5].

2.4.1 DQDB

Das Distributed Queue Dual Bus (DQDB)-Zugriffsverfahren für lokale und öffentliche Breitbandnetze (MAN) wurde zunächst vom IEEE, später von der ISO als Standard eingeführt [37]. Als Topologie wird ein Dualbus gewählt (Bild 2-5). Für Paketverkehr realisiert DQDB das Prinzip einer verteilten Warteschlange („Distributed Queue"), in die Sendewünsche der Stationen eingetragen werden. DQDB integriert asynchronen Datenverkehr und Isochronverkehr mit unterschiedlichen Bandbreiten.

Eine Station kann auf beiden Bussen lesen und schreiben. Das Löschen von Daten auf einem Bus ist nicht möglich, da eine Station immer nur Signale vom Zustand „0" in den Zustand „1" verändern kann, nicht aber umgekehrt. Ein einmal mit einem Segment belegter Zeitschlitz („Slot") kann folglich auch nach Erreichen des Empfängers nicht ein zweites Mal verwendet werden. Bild 2-16 zeigt die Anschlußeinheit einer DQDB-Station.

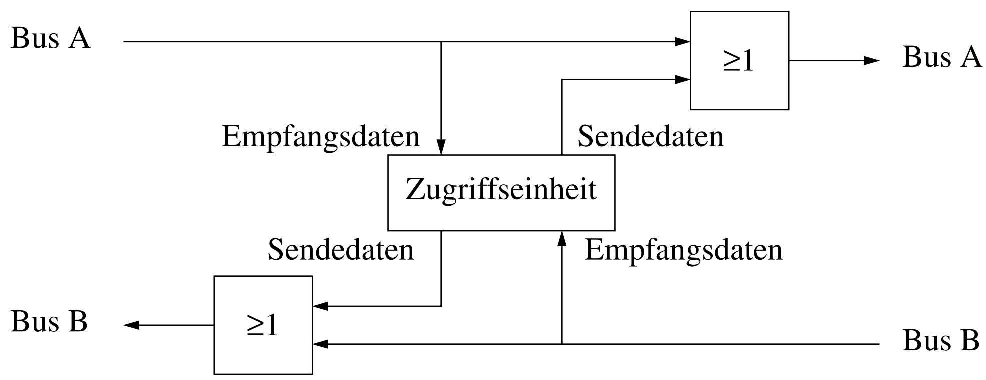
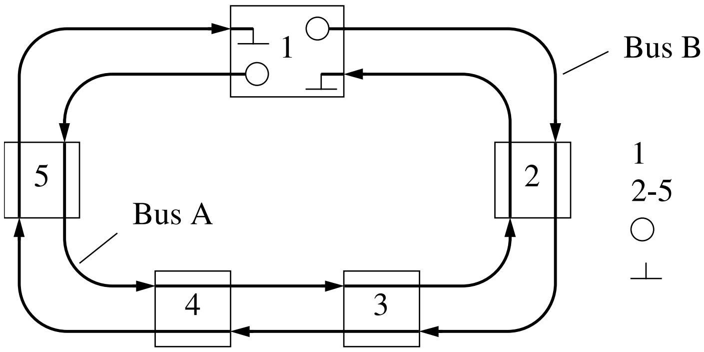


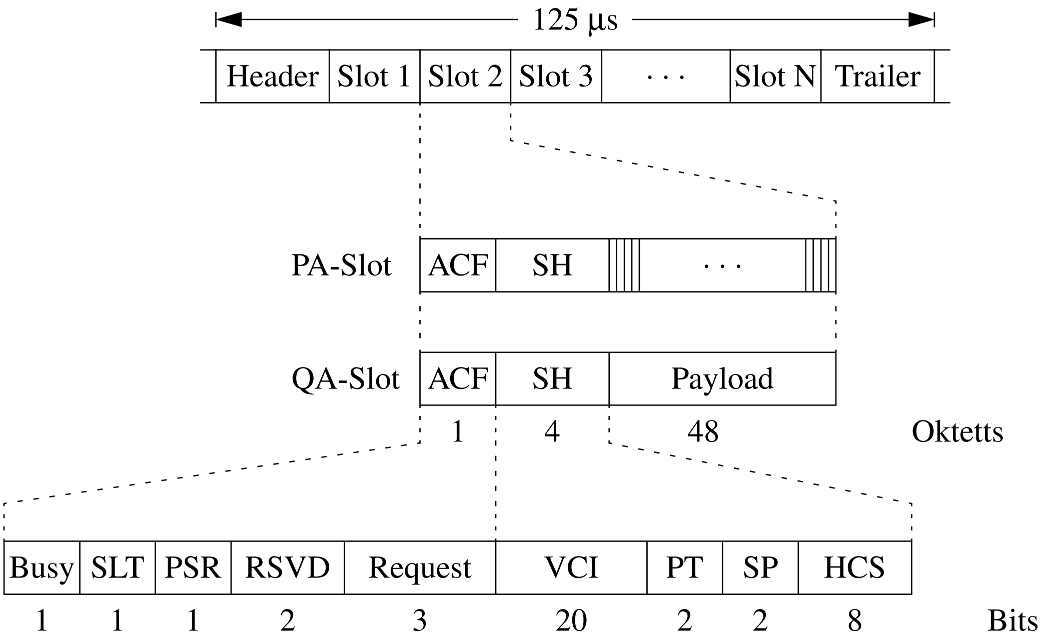
Bild 2-16: Anschlußeinheit einer DQDB-Station

Die beiden Rahmengeneratoren an den beiden Busanfängen können in einer Station vereint sein, wenn der Dualbus in Ringform aufgebaut wird (Bild 2-17). Die ausgesendeten Rahmen („Frames“) sind 125 µs lang und bestehen aus Frame Header, Frame Trailer und einer durch die Übertragungsrate vorgegebenen Anzahl von Zeitschlitzen (Bild 2-18). Zwei Arten von Zeitschlitzen sind vorgesehen und unterscheiden sich im „Payload Type“-Bit innerhalb des Access Control Field.



1	Gemeinsame Rahmengeneratorstation
2-5	Weitere Stationen auf dem Dualbus.
○	Rahmengenerator
⊥	Senke

Bild 2-17: DQDB-Konfiguration mit einer Rahmengeneratorstation



PA	Pre-Allocated	QA	Queued-Arbitrated
ACF	Access Control Field	VCI	Virtual Channel Identifier
SH	Segment Header	PT	Payload Type
SLT	Slot Type	SP	Segment Priority
PSR	Previous Segment Received	HCS	Header Check Sequence
RSVD	Reserved		

Bild 2-18: Rahmen- und Zeitschlitzformat bei DQDB

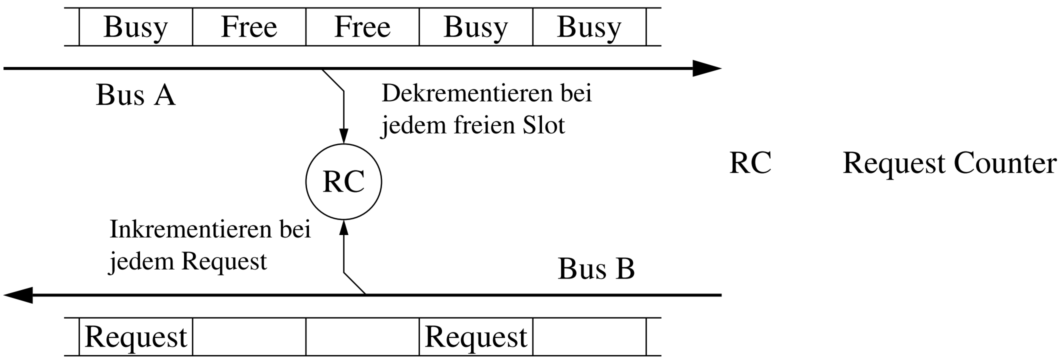


Bild 2-19: Request-Zähleroperationen einer Station für Bus A; kein Sendewunsch

Pre-Arbitrated Slots, die bereits vom Rahmengenerator reserviert werden, realisieren jeweils einen isochronen Kanal mit einer Nutzbandbreite von 3072 kbit/s, indem in jedem Rahmen jeweils der gleiche Zeitschlitz genutzt wird. Ein isochroner Kanal kann weiter in 48 isochrone Kanäle mit jeweils 64 kbit/s unterteilt werden, indem einem dieser schmalbandigen Kanäle jeweils ein Oktett innerhalb des Zeitschlitzes zugeordnet wird. Damit ist die Übertragung von Sprache oder die Integration von ISDN-B-Kanälen in DQDB möglich.

Queued-Arbitrated Slots für den verbindungslosen Datenverkehr werden von den Stationen nach dem „Distributed Queueing“- Protokoll belegt. Die Puffer aller Stationen bilden zusammen eine verteilte Warteschlange, wobei der Füllstand sowie die Position des Segments einer Station innerhalb der Warteschlange mit Hilfe von Zählern bestimmt wird. Eine Station mit Sendewunsch teilt auf dem gegengerichteten Bus den stromaufwärts liegenden Stationen diesen Sendewunsch über einen „Request“ mit und reiht ihren Sendewunsch damit in die verteilte Warteschlange ein. Alle Stationen verfügen über je einen Request Counter für die beiden Busse, der mit jedem freien Zeitschlitz dekrementiert und mit jedem Request auf dem gegengerichteten Bus inkrementiert wird. Die Stationen kennen auf diese Weise mit dem Stand des Request Counters die Anzahl der Requests von Stationen stromaufwärts, die noch nicht abgearbeitet wurden (Bild 2-19).

Eine Station, die einen Sendewunsch hat, sendet zunächst einen Request auf dem gegengerichteten Bus ab. Gleichzeitig kopiert sie den Stand des Request Counters in ihren Countdown Counter und setzt den Request Counter auf Null (Bild 2-20). Danach zählt der Request Counter nur noch die Requests auf dem gegengerichteten Bus, jedoch nicht mehr die freien Zeitschlitze, während der Countdown Counter mit jedem freien Zeitschlitz, den die Station passieren läßt, dekrementiert wird. Sobald der Countdown Counter den Wert Null erreicht hat, kann die Station den nächsten freien Zeitschlitz belegen und nach dem beschriebenen Verfahren weitere Übertragungen einleiten.

In DQDB sind drei Prioritätsklassen vorgesehen, für die jeweils ein eigenes Request-Bit im ACF-Feld genutzt wird. Die Stationen verfügen je Prioritätsklasse und Bus über ein Paar aus

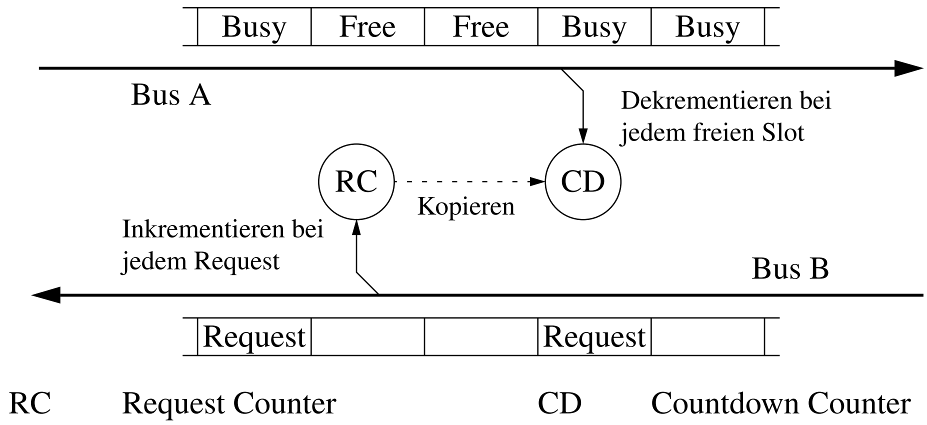


Bild 2-20: Zähleroperationen einer Station für Bus A: Sendewunsch

Request und Countdown Counter. Ein Segment niedrigerer Priorität darf erst dann gesendet werden, wenn alle Requests höherer Priorität abgearbeitet sind.

Untersuchungen der Leistungsfähigkeit von DQDB [87, 88, 107] zeigen, daß Stationen in der Mitte des Busses gegenüber den Stationen in der Nähe der Busenden benachteiligt sind. Zum Ausgleich dieser Unfairness wird der „Bandwidth Balancing"-Mechanismus [37, 25, 92] eingesetzt. Das Grundprinzip ist, daß jede Station einen kleinen Teil der ihr nach dem DQDB-Protokoll zustehenden Bandbreite zugunsten stromabwärts positionierter Stationen nicht nutzt.

Eine Erweiterung des DQDB-Protokolls sieht „Erasure Nodes" vor, die einen Zeitschlitz wieder freigeben, nachdem das darin enthaltene Segment beim Empfänger angekommen ist. Damit kann der Zeitschlitz auf dem Bus ein zweites Mal genutzt werden. Allerdings muß dann der Fairness-Algorithmus an die erhöhte Bandbreite angepaßt werden [57]. Im Dynamic Distributed Dual Queue-Protokoll (D^3Q) wird das Konzept der verteilten Warteschlangen auf einen Doppelring angewandt [15]. Multirequest-D^3Q sieht mehrere ausstehende Requests vor und kann als eine Verallgemeinerung des Konzepts der verteilten Warteschlangen gesehen werden [91].

2.4.2 Orwell

Das Orwell[1]-Protokoll [1] wurde von den British Telecom Laboratories in Ipswich (England) entwickelt. Orwell nutzt einen Slotted Ring oder einen Torus, wobei der Empfänger einer Nachricht den Zeitschlitz freigibt („Destination Release"). Das Medienzugriffsprotokoll bei Orwell basiert auf Credits (vgl. 2.3.6.3), wobei der Reset-Mechanismus mit Hilfe eines verteilten Verfahrens realisiert wird. Jede Station verfügt über vier Credit Counter für vier Prioritäten.

Bild 2-21 zeigt die Struktur eines Zeitschlitzes. Einer kompletten ATM-Zelle mit 48 Oktetts Nutzdaten und 5 Oktetts Header (siehe Abschnitt 3.1.2) wird ein Orwell-Header von 3 Oktetts vorangestellt, so daß sich insgesamt ein Zeitschlitz von 56 Oktetts ergibt. Der Start Delimiter im Orwell-Header dient der Synchronisation der Empfänger auf den Zeitschlitztakt. Eingebet-

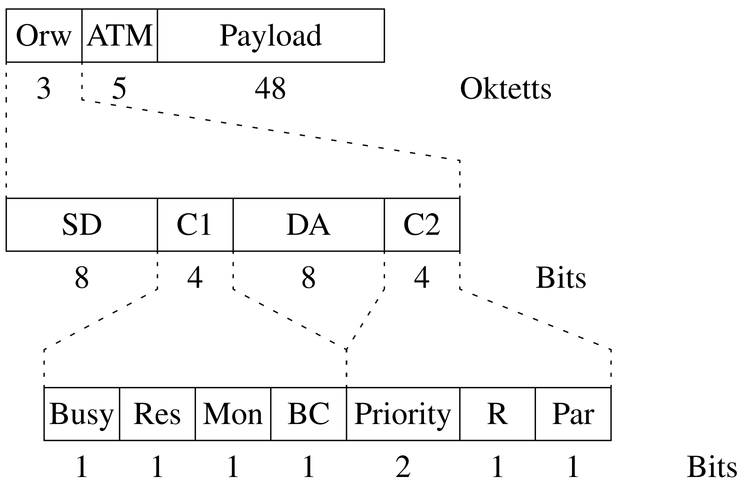
1. Benannt nach dem Fluß Orwell in der Nähe der British Telecom Laboratories

tet in die Mitte eines 8 bit großen Steuerfelds C befindet sich die Zieladresse DA. Mit dieser Adresse können 256 Stationen im Orwell-Ring adressiert werden.

Das „Busy“-Bit im Steuerfeld zeigt an, ob der Zeitschlitz belegt ist. Unbelegte Zeitschlitze werden mit Hilfe des Reset-Bits unterschieden in „Trial Slots“ und „Reset Slots“. Eine Station mit Sendewunsch darf einen „Trial Slot“ belegen, wenn ihre Credits nicht aufgebraucht sind, wenn also mindestens einer ihrer vier Credit Counter nicht Null ist. Dabei werden die vier Prioritäten unterschieden, indem immer die Sendewarteschlangen höherer Priorität zuerst bedient werden. Die sendende Station markiert die Priorität des gesendeten Zeitschlitzes mit den beiden Priority-Bits.

Das Monitor-Bit wird von einer ausgezeichneten Monitor-Station genutzt, um im Fehlerfall Zeitschlitze freizugeben, deren Empfänger nicht oder nicht mehr existiert (vgl. 2.3.7). Das Broadcast-Bit zeigt an, daß ein Zeitschlitz von allen Stationen empfangen werden soll, ungeachtet der Zieladresse DA. Außerdem wird in unbelegten Zeitschlitzen das Broadcast-Bit dazu verwendet, um den „Trial Type“ zu unterscheiden. Das „R“-Bit wird von der Monitor-Station in genau einem Zeitschlitz auf dem Ring gesetzt. Damit können alle Stationen die Ringumlaufzeit bestimmen, indem sie die Zeitdauer zwischen zwei Ankünften eines Zeitschlitzes mit gesetztem „R“-Bit messen.

Der Empfänger eines Zeitschlitzes kann den Zeitschlitz entweder sofort wieder belegen, sofern er noch über Credits verfügt, oder er gibt den Zeitschlitz frei. Auf diese Weise entsteht ein „Trial Slot“, wobei in das Zieladressfeld DA die eigene Adresse eingetragen wird. In das Broadcast-Bit (in diesem Fall verwendet als „Trial Type“-Bit) wird der „Trial State“ (0 oder 1) der Station eingetragen. Kommt ein „Trial Slot“ wieder bei der Station an, die ihn generiert



Orw	Orwell Header	Res	Reset Bit
ATM	ATM Header	Mon	Monitor Bit
SD	Start Delimiter	BC	Broadcast/Trial Type Bit
C1	Control Field 1	R	Ring Cycle Indicator Bit
DA	Destination Address	Par	Parity Bit
C2	Control Field 2		

Bild 2-21: Zeitschlitzformat bei Orwell

hat, und kann oder darf diese Station nicht senden, dann erzeugt die Station einen „Reset Slot", wenn das „Trial Type"-Bit im Slot mit dem „Trial State" der Station übereinstimmt. Gleichzeitig wechselt die Station ihren „Trial State". Damit stimmt bei nachfolgend ankommenden „Trial Slots" der „Trial Type" nicht mehr mit dem „Trial State" der Station überein. Auf diese Weise wird verhindert, daß mehrere aufeinanderfolgende „Reset Slots" generiert werden. Mit dem Aussenden des „Reset Slot" geht die Station in den „Outstanding Reset"-Zustand über, in dem sie verbleibt, bis ein „Reset Slot" ankommt oder ein stationsinterner Zeitgeber abläuft. Dieser Zeitgeber wird auf die Ringumlaufzeit gesetzt. Damit kann die Station erkennen, wenn aufgrund eines Fehlers ihr „Reset Slot" verloren gegangen ist und kann geeignet darauf reagieren.

Ein „Reset Slot" bewegt sich so lange auf dem Ring, bis er auf eine Station im „Outstanding Reset"-Zustand trifft, die ihn in einen „Trial Slot" umwandelt. Passiert der „Reset Slot" eine Station, werden ihr neue Credits zugeteilt, indem alle Credit Counter auf ihren jeweils vorgegebenen Wert gesetzt werden. Wenn zwei oder mehrere Stationen gleichzeitig einen „Reset Slot" generieren, dann werden diese Slots von der jeweils nächsten Station im „Outstanding Reset"-Zustand entfernt.

Verkehrstheoretische Betrachtungen des Orwell-Protokolls [1, 77, 108] definieren ein Reset-Intervall als die Zeitdauer zwischen zwei aufeinanderfolgenden „Reset Slots". Das Reset-Intervall steigt von einem Minimalwert im leeren System mit dem Angebot der Stationen bis zu einem Maximalwert, der sich aus der Begrenzung der Gesamtzahl zugeteilter Credits ergibt. Betrachtet man eine Station mit einem Credit je Zyklus, die in einem Reset-Intervall ihr Credit zu Beginn des Zyklus verbraucht, im nächsten Zyklus jedoch erst als letzte Station einen freien Zeitschlitz vorfindet, ergibt sich als eine obere Schranke für die Zugriffszeit dieser Station etwa das Doppelte des maximalen Reset-Intervalls.

Durch den Reset-Vorgang entstehen Phasen, in denen das System leerläuft und keine Station senden kann. Diese Totzeit ist im wesentlichen durch die Ringlänge bestimmt und unabhängig von der Auslastung des Systems. Der Durchsatz des Systems nimmt in dem Maße ab, wie das Verhältnis von Totzeit zu Reset-Intervall zunimmt. Folglich kann bei niedriger Last der Durchsatz durch die Verlängerung des Reset-Intervalls erhöht werden. Hierzu kann die Anzahl der vergebenen Credits bei einem Reset abhängig von der Dauer des vorherigen Reset-Intervalls zwei unterschiedliche Werte annehmen: Überschreitet die gemessene Zeit seit dem letzten „Reset Slot" einen Schwellwert, so wird die kleinere, ursprüngliche Anzahl von Credits zugeteilt. Eine größere Credit-Anzahl wird vergeben, wenn der Schwellwert unterschritten wird. Allerdings kann dann bei einem sprunghaften Anstieg des systemweiten Angebots die obere Grenze der Verzögerungszeit möglicherweise nicht mehr eingehalten werden.

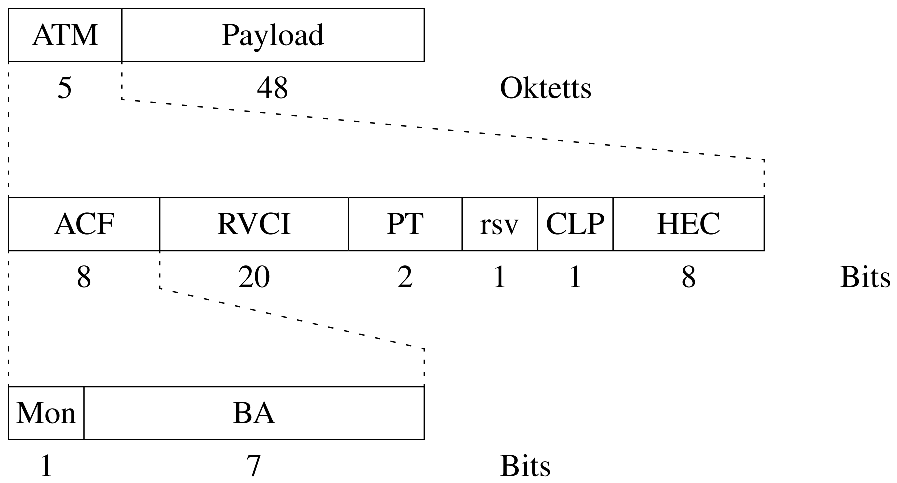
Die vier definierten Prioritätsklassen sind nicht unabhängig voneinander, insbesondere hat das Angebot niedriger Priorität direkte Auswirkungen auf die Verzögerungszeit und den Durchsatz der höheren Prioritätsklassen, da alle Klassen von einem gemeinsamen Reset-Intervall abhängen. Die Prioritätsklassen werden lediglich bei der stationsinternen Organisation im Sender und Empfänger getrennt behandelt, indem für jede Klasse eine getrennte Warteschlange und ein eigener Credit Counter bereitsteht. Damit sich eine Station mit den vier Prioritäten in ihrem äußeren Verhalten von einer Station mit einer einzigen Priorität, aber gleicher Gesamtzahl von Credits überhaupt unterscheidet, muß die Anzahl der Credits, die in den niedrigen Prioritätsklassen vergeben werden, von der Dauer des Reset-Intervalls abhängig sein.

2.4.3 ATM-Ring

Der ATM-Ring (ATMR) wurde von den NTT Communication Switching Laboratories in Tokyo, Japan, ausgehend vom Orwell-Protokoll entwickelt und vom Japanese National Body veröffentlicht [56]. ATMR ist als Slotted Ring oder Doppelring spezifiziert. Das Format der Zeitschlitze (Zellen) ist sehr stark an das ATM-Zellformat (siehe Abschnitt 3.1.2) angelehnt, um einen einfachen Übergang zum Breitband-ISDN zu ermöglichen. Mehrere Prioritätsklassen sind realisierbar, wobei höhere Prioritäten bei Überlast die niedrigeren Prioritäten verdrängen.

Bild 2-22 zeigt das Zeitschlitzformat bei ATMR. Der Zeitschlitz besteht aus einem Header mit 5 Oktetts sowie einer Nutzdateneinheit von 48 Oktetts. Der Header wird durch eine Prüfsequenz mit einer Länge von 8 bit gegen Fehler gesichert. Die Felder PT und CLP entsprechen den Definitionen im B-ISDN und werden in Abschnitt 3.1.2 beschrieben. Sie geben den Typ der Nutzdaten an und enthalten Informationen für das Verwerfen von Zellen in einem Vermittlungsknoten bei Überlast.

Das ACF ist aufgeteilt in ein Monitor-Bit zum Entfernen fehlerhafter Zellen (vgl. 2.3.7) und ein Adressfeld BA, das jedoch nicht für die Adressierung, sondern für den Credit-Mechanismus genutzt wird. Tabelle 2-1 enthält die möglichen Werte des BA-Felds bei drei Prioritäts-



ATM	ATM Header	CLP	Cell Loss Priority
ACF	Access Control Field	HEC	Header Error Check
RVCI	Ring Virtual Channel Identifier	Mon	Monitor Bit
PT	Payload Type	BA	Busy Address
rsv	Reserved Field for future use		

Bild 2-22: Zeitschlitzformat beim ATM-Ring

klassen. Die Adressierung der Stationen erfolgt über das RVCI-Feld mit 20 bit Breite. Ein leerer Zeitschlitz wird durch einen RVCI-Wert von Null gekennzeichnet.

BA	Bedeutung
0000000	Common Busy Address
0000001 ... 1111011	Busy Address
1111100	Reset cell (level 4)
1111101	Reset cell (level 3)
1111110	Reset cell (level 2)
1111111	Reset cell (level 1)

Tabelle 2-1: Werte des BA-Felds bei ATMR mit 3 Prioritäten

In das BA-Feld trägt jede sendefähige Station in allen Zeitschlitzen ihre eigene Stationsadresse ein. Sendefähig ist eine Station, die über Credits verfügt und Daten in ihrem Sendepuffer hat. Wenn eine Station, die nicht mehr sendefähig ist, ihre eigene Stationsadresse im BA-Feld vorfindet, erkennt sie daraus, daß keine Station im Ring mehr sendefähig ist. Sie sendet dann eine Reset-Zelle aus und geht in den „Hunting“-Zustand, in dem sie auf die Rückkehr der ausgesendeten Reset-Zelle wartet. Sobald eine Reset-Zelle ankommt, wird diese entfernt und der „Hunting“-Zustand verlassen. Mit einem Zeitgeber wird die Zeit bis zum Eintreffen der Reset-Zelle überwacht, um ein Blockieren des Systems bei Verlust der Reset-Zelle zu verhindern. Mit der Reset-Zelle werden den Stationen neue Credits zugeteilt, indem sie ihren Credit Counter auf einen vordefinierten Initialwert setzen.

Die Anzahl der Stationen auf dem Ring ist durch die Breite des BA-Feldes beschränkt. Bei 3 Prioritäten können maximal 124 Stationen ihre eigene Adresse in das BA-Feld eintragen und die Bedingung für einen Reset erkennen. Weitere Stationen können angeschlossen werden, die jedoch nur passiv am Reset-Mechanismus teilnehmen und nicht ihre eigene Stationsadresse, sondern die Common Busy Address in das BA-Feld eintragen. Die nächste Station, die aktiv am Reset-Mechanismus teilnimmt, ersetzt diese Common Busy Address durch ihre eigene Adresse. Mit diesem Mechanismus ist es optional möglich, die Länge des ACF-Feldes auf 4 bit zu begrenzen, um gleichzeitig das RVCI-Feld auf 24 bit zu erweitern [31].

Zur Realisierung von Prioritäten wird in allen Stationen für jede Priorität ein eigener Sendepuffer und Credit Counter sowie für alle Prioritäten außer der niedrigsten ein Zeitgeber („Queue Timer“) benötigt. Ferner wird ein „Transmission Level“ eingeführt, der bestimmt, in welcher Prioritätsklasse eine Station senden kann. Transmission Level 1 bedeutet, daß die Station nur in der höchsten Prioritätsklasse 1 senden darf, während Tansmission Level i das Senden in allen Prioritätsklassen 1 bis i, also i und höher erlaubt. Somit können niedrige Prioritätsklassen vorübergehend vom Senden ausgeschlossen werden, wenn die QoS-Parameter höherer Prioritätsklassen aufgrund des Verkehrs niedrigerer Prioritätsklassen und der daraus resultierenden Verzögerungszeit für die höheren Prioritätsklassen sonst nicht mehr einzuhalten wären.

Bei N_{Pri} Prioritäten werden $N_{Pri}+1$ verschiedene Reset-Zellen benötigt. Tabelle 2-2 zeigt die Funktion der jeweiligen Reset-Zellen im Fall von $N_{Pri}=3$ Prioritäten, der im folgenden erläutert wird. Priorität „3“ bedeutet dabei niedrigste, „1“ höchste Priorität. Nach der Initialisierung des Rings befinden sich alle Stationen im Transmission Level 3, in dem alle Prioritäten zugelassen sind. Abhängig vom Transmission Level j werden gemäß des oben beschriebenen Mechanismus Reset-Zellen mit Level j+1 verwendet.

Reset-Zelle	Neuer Transmission Level	Senden erlaubt in Priorität	Credit Counter zurücksetzen	Queue Timer zurücksetzen
Reset_1	1	1	keinen	keinen
Reset_2	2	1 und 2	1	1
Reset_3	3	1, 2 und 3	1 und 2	1 und 2
Reset_4	3	1, 2 und 3	1, 2 und 3	1 und 2

Tabelle 2-2: Aktionen bei Ankunft von Reset-Zellen in ATMR (3 Prioritäten)

Die Sendepuffer der Stationen werden in der Reihenfolge von höchster zu niedrigerer Priorität bedient. In den Prioritätsklassen werden mit steigender Priorität geringere Wartezeiten und damit kleinere Reset-Intervalle zwischen zwei Rücksetzvorgängen der entsprechenden Credit Counter verlangt. Hierzu werden die Queue Timer eingesetzt, die rechtzeitig eine Erhöhung der verfügbaren Bandbreite in der zugehörigen Prioritätsklasse einleiten. Wenn ein Zeitgeber der Prioritätsklasse i abläuft, werden zwei Fälle unterschieden:

- Fall 1: Die Station darf und will Zellen der Priorität i senden, findet jedoch aufgrund des Verkehrs niedrigerer Priorität der Vorgängerstationen keinen freien Zeitschlitz vor
- Fall 2: Die Station hat bereits alle Credits der Priorität i aufgebraucht und wartet auf einen Reset des Credit Counters i, der jedoch aufgrund des Verkehrs niedrigerer Prioritäten nicht ausgelöst wird

Im Fall 1 muß die Station freie Zeitschlitze erhalten, um ihre Daten senden zu können. Dies wird erreicht, indem die niedrigeren Prioritäten (>i) vom Senden ausgeschlossen werden: Ein Reset mit Level i wird generiert. Im zweiten Fall muß der Credit Counter der Priorität i zurückgesetzt werden, damit die Station sofort wieder Zellen der Priorität i senden kann. Dies wird mit einem Reset mit Level (i+1) erreicht. Tabelle 2-3 zeigt die möglichen Fälle für die Generierung von Resets bei 3 Prioritätsklasen in Abhängigkeit vom augenblicklichen Transmission Level, Ereignissen und dem Stand von Credit Countern.

Das zeitgesteuerte Aussenden von Resets bringt eine Verzerrung des Credit-Mechanismus in den höheren Prioritäten mit sich, da Credit Counter vorzeitig, also solange noch nicht alle Credits verbraucht sind, aufgefrischt und gleichzeitig Zeitgeber zurückgesetzt werden.

Transmission Level	Timer 1, CC1 > 0	Timer 1, CC1 = 0	Timer 2, CC2 > 0	Timer 2, CC2 = 0	Keine Station aktiv
1	-	-	-	-	Reset_2
2	Reset_1	Reset_2	-	-	Reset_3
3	Reset_1	Reset_2	Reset_2	Reset_3	Reset_4

Timer j Ablauf des Zeitgebers der Priorität j
CCj Credit Counter der Priorität j

Tabelle 2-3: Reset-Generierung bei ATMR mit 3 Prioritätsklassen

Untersuchungen der Leistungsfähigkeit des ATMR-Protokolls ergaben eine gegenüber Orwell verkürzte Wartezeit aufgrund des verbesserten Reset-Mechanismus [31]. Die Wartezeit höherer Prioritäten ist abhängig vom Angebot in niedrigeren Prioritäten. Durch geeignete Wahl der Parameter kann für zeitkritischen Verkehr eine obere Schranke der Zugriffszeit eingehalten werden [26].

FECCA[1] [86] ist eine Weiterentwicklung des ATM-Ring und basiert auf einem Doppelring. Dabei bezieht sich der Credit-Mechanismus auf den gegenläufigen Ring. Im ACF des einen Rings steht die Adresse der nächstgelegenen aktiven Station des anderen Rings, also die Adresse der Station, bis zu der der Ring momentan ungenutzt ist. Mit dieser Information kann die Station das ungenutzte Ringsegment bis zu dieser nächsten aktiven Station unter Umgehung des Credit-Mechanismus belegen, um die Gesamtauslastung des Doppelrings zu erhöhen. Unter der Voraussetzung, daß die Station Kenntnis über die Anordnung der Stationen hat, kann sie an eine Station senden, die zwischen ihr und der nächsten aktiven Station am ungenutzten Ringsegment positioniert ist. Ein lokaler Fairness-Mechanismus regelt die Aufteilung dieser zusätzlichen Bandbreite zwischen den Stationen auf dem betroffenen Ringsegment.

2.4.4 S++

Das Medienzugriffsprotokoll S++ wurde von den Hewlett-Packard Laboratories in Bristol, England, entwickelt. Das Protokoll ist abgeleitet vom Simple- oder S-Protokoll [65], das wiederum auf dem Fasnet-Protokoll [66] basiert. Diese Protokolle verwenden Zeitschlitze und einen Credit-Mechanismus für ihr Medienzugriffsverfahren.

Das Simple-Protokoll kann realisiert werden als gefalteter Bus oder Dualbus (vgl. 2.2.2). Beim Faltbus darf eine Station, die über Credits verfügt, den nächsten freien Zeitschlitz auf dem Schreibbus belegen und verbraucht damit eines ihrer Credits. Dieser Zeitschlitz kehrt unverändert, also in belegtem Zustand, auf dem Lesebus zurück. Wenn keine Station mehr senden kann, wird ein unbelegter Zeitschlitz die letzte Sendestation passieren und daraufhin auf dem Lesebus erscheinen. Mit jedem freien Zeitschlitz auf dem Lesebus setzen die Stationen ihre Credit Counter zurück und der nächste Zyklus beginnt. Beim Dualbus senden die Stationen auf beiden Bussen und die Zellen kehren nicht auf dem anderen Bus zurück. Der Belegungszu-

1. Fair and Efficient Cyclic Control Algorithm

stand ankommender Zeitschlitze wird daher von der Kopfstation in ein „End“-Bit innerhalb der Zeitschlitze auf dem abgehenden Bus kopiert. Das „End“-Bit auf einem Bus wird wiederum als Reset-Signal für den Credit-Mechanismus des gegenüberliegenden Busses interpretiert.

S++ ist eine Erweiterung von Simple zur Erhöhung des Gesamtdurchsatzes auf der Basis eines Faltbusses [102]. Die Totzeit des Systems zwischen den Credit-Zyklen soll vermieden und dadurch der Durchsatz erhöht werden. Die Zeitschlitze enthalten ein zusätzliches Bit „Restart“, das von der Kopfstation in jedem generierten Zeitschlitz gesetzt wird. Jede Station verfügt über einen Credit Counter P, ein Register P_{max}, das die Anzahl der Credits für einen Zyklus enthält, einen Sendepuffer mit Belegungszustand Q und ein Flag F (Bild 2-23).

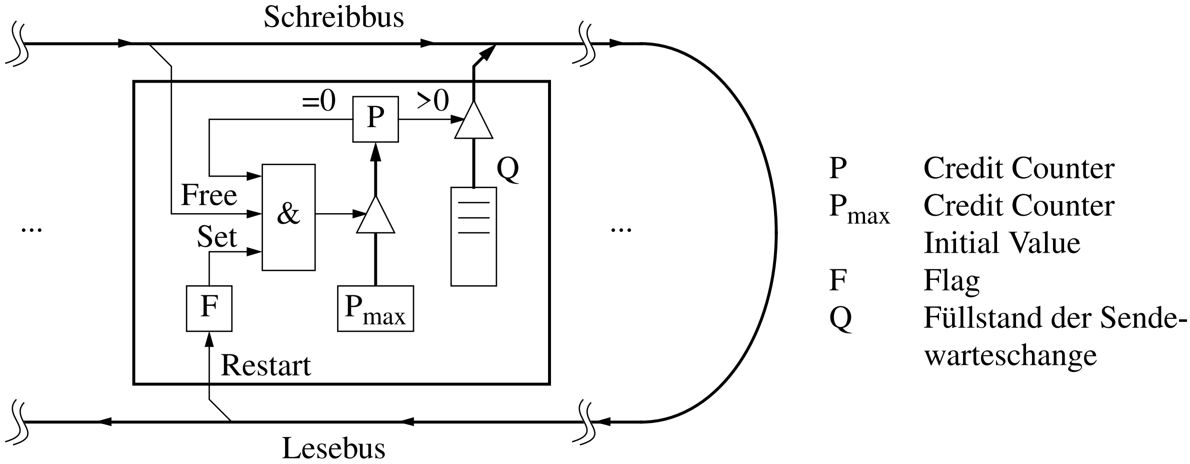


Bild 2-23: Stationsstruktur bei S++ (Faltbus)

Eine Station darf freie Zeitschlitze auf dem Sendebus belegen, solange P > 0 und Q > 0 sind, die Station also noch über Credits verfügt und Daten im Sendepuffer bereitstehen („ACTIVE“, Tabelle 2-4). Sind diese Bedingungen erfüllt, wird außerdem in jedem Zeitschlitz auf dem Sendebus das Restart-Bit zurückgesetzt. Somit ist sichergestellt, daß am Ende des Sendebusses das Restart-Bit genau dann gesetzt ist, wenn die beiden Bedingung (P > 0 und Q > 0) an keiner der Stationen gleichzeitig erfüllt sind und somit keine Station sendeberechtigt oder sendewillig ist. Der Zeitschlitz mit gesetztem Restart-Bit kommt auf dem Lesebus zurück und setzt bei allen Stationen das Flag F.

Credit Counter P	Pufferfüllstand Q	Zustand	Bemerkung
beliebig	0	IDLE	Keine Sendedaten vorhanden
0	> 0	DEFER	Keine Credits vorhanden
> 0	> 0	ACTIVE	Station darf senden

Tabelle 2-4: Mögliche Zustände einer Station in S++

Der Credit Counter P wird immer dann mit dem Wert aus dem Register P_{max} geladen, wenn drei Bedingungen erfüllt sind: Ein Zeitschlitz mit gesetztem Restart-Bit kommt bei der Station auf dem Sendebus an, der Credit Counter P ist Null und das Flag F ist gesetzt. Gleichzeitig mit dem Laden von P wird das Flag F zurückgesetzt. Tabelle 2-5 zeigt die Aktionen einer Station.

Bedingungen				Aktionen			
Zustand[a]	Zeitschlitz auf dem Sendebus		F	P	F[b]	Zeitschlitz auf dem Sendebus	
	Busy	Restart				Belegen	Restart[c]
IDLE	X	X	X				
DEFER	X	X	0				
DEFER	X	0	1				
DEFER	0	1	1	P_{max}-1	0	ja	
DEFER	1	1	1	P_{max}	0		
ACTIVE	0	X	X	P-1		ja	0
ACTIVE	1	X	X				0

X: Beliebiger Wert (don‘t care)
leeres Feld: unveränderter Wert bzw. keine Aktion

Tabelle 2-5: Verhalten einer Station bei S++

a. siehe Tabelle 2-4
b. Flag F wird gesetzt (=1), wenn auf dem Lesebus ein Zeitschlitz mit Restart = 0 ankommt
c. Restart wird von der Kopfstation in jedem generierten Zeitschlitz gesetzt (=1)

In S++ ist aufgrund seiner Ausrichtung auf den Datenverkehr in einem lokalen Netz [103] keine Priorisierung einzelner Verbindungen vorgesehen. Prioritäten könnten jedoch ähnlich dem ATM-Ring (siehe 2.4.3) eingefügt werden. Durch die auf Zeitschlitzen basierende Übertragung wäre außerdem die Einführung isochroner Kanäle möglich. Für dringende Nachrichten wird ein Request-Response-Mechanismus vorgeschlagen, bei dem eine Station einen Zeitschlitz für ihre dringende Nachricht auf dem Lesebus durch Setzen eines dafür vorgesehenen Bits anfordert. Die Kopfstation reserviert daraufhin einen Zeitschlitz für diese dringende Nachricht [102].

Im Vergleich zu DQDB bietet S++ geringere Wartezeiten und größere Fairness, da der Durchsatz einer Station nicht abhängig ist von ihrer Position auf dem Medium [104]. Die Varianz der Zugriffszeit ist höher als bei DQDB, da bei S++ eine einzelne Station im Extremfall P_{max} aufeinanderfolgende Zeitschlitze belegen kann.

2.4.5 CRMA

Das Cyclic Reservation Multiple Access Protokoll (CRMA) wurde von IBM in Zürich entwikkelt. Die ursprüngliche Version [72] basiert auf einem Doppel- oder Faltbus, während die Weiterentwicklung CRMA-II für einen Ring oder Doppelring konzipiert wurde. CRMA basiert auf der Reservierung eines Kontingents von Zeitschlitzen nach Bedarf für jede Station innerhalb eines Reservierungszyklus. Betrachtet werden soll der Zugriffsmechanismus für Bus A in einem Dualbussystem (Bild 2-5); aufgrund der Symmetrie dieser Topologie wird auf Bus B dasselbe Verfahren angewendet.

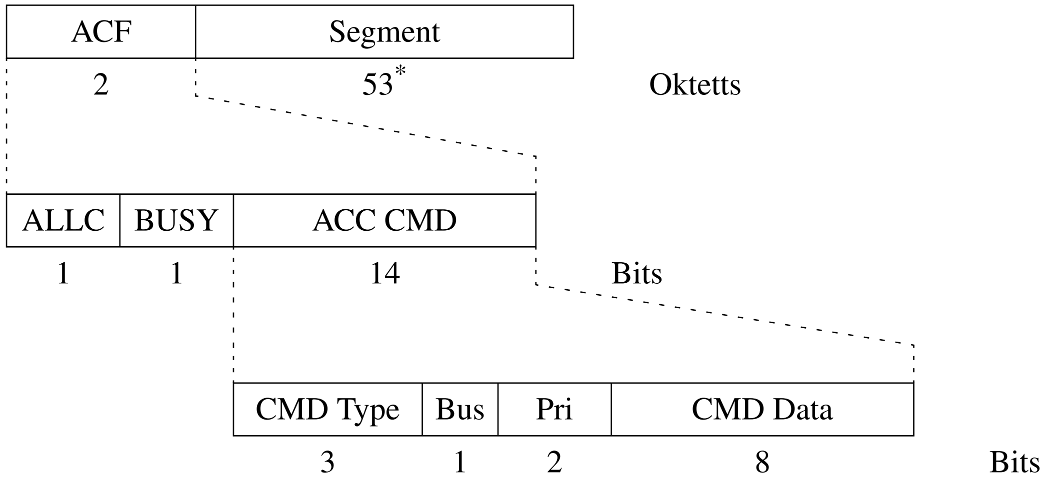
Die Kopfstation sendet auf Bus A zyklisch Reservierungs-Kommandos aus, die fortlaufend numeriert sind und eine Variable CL für die Anzahl der gewünschten Reservierungen beinhalten. Diese Reservierungskommandos werden zunächst unverändert zur gegenüberliegenden Kopfstation weitergeleitet, die sie wiederum auf Bus B zurücksendet. Stationen mit Sendewunsch auf Bus A addieren jetzt zu der Variablen CL die Anzahl der Zeitschlitze auf, die sie im zugehörigen Zyklus zu belegen wünschen. Gleichzeitig wird diese Zahl zusammen mit der Zyklusnummer in eine lokale Reservierungswarteschlange für Bus A eingereiht. Die Anzahl der reservierten Zeitschlitze darf eine für jede Station festgelegte Zahl nicht überschreiten. Die Kopfstation reiht Zykluslänge CL und die zugehörige Zyklusnummer der zurückkommenden Reservierungskommandos in ihre globale Reservierungswarteschlange ein.

Zu Beginn eines Reservierungszyklus nimmt die Kopfstation einen Eintrag mit Zyklusnummer und Zykluslänge CL aus ihrer globalen Reservierungswarteschlange und sendet ein Start-Kommando mit dieser Zyklusnummer. Nach CL Zeitschlitzen beginnt der nächste Reservierungszyklus. Die Stationen erkennen aus dem Eintrag in ihrer lokalen Reservierungswarteschlange, ob sie für diesen Zyklus Zeitschlitze reserviert haben, und belegen gegebenenfalls den ersten freien und nachfolgende Zeitschlitze des mit dem Confirm-Kommando eingeleiteten Reset-Zyklus entsprechend dem reservierten Kontingent.

Im Fall von Überlast ist der Reservierungszyklus im Mittel länger als die Zeitdauer zwischen den Reservierungskommandos. Die Summe der reservierten Zeitschlitze in der globalen Reservierungswarteschlange steigt kontinuierlich an. Zur Überlastabwehr sendet die Kopfstation ein Reject-Kommando aus, sobald die Summe der reservierten Zeitschlitze einen Schwellwert überschreitet. Dieses Reject-Kommando teilt den Stationen mit, daß die letzten, nicht bestätigten Reservierungen nicht angenommen wurden. Andererseits werden Reservierungen mit einem Confirm-Kommando und zugehöriger Zyklusnummer bestätigt, wenn der Schwellwert nicht überschritten ist.

Bei geringer Last sendet die Kopfstation dann, wenn keine Reservierungen mehr in der globalen Reservierungswarteschlange vorliegen, weiterhin Zeitschlitze aus, die sie als unreserviert markiert („Unallocated"). Diese Zeitschlitze kann jede Station belegen, sofern sie nicht bereits belegt sind. Die Übertragung eines Pakets wird abgebrochen und muß später wiederholt werden, sobald ein belegter Zeitschlitz ankommt oder ein neuer Zyklus beginnt, für den die Station keine Zeitschlitze reserviert hat. Hat eine Station Zeitschlitze für ein Paket reserviert, das Paket jedoch in als „Unallocated" markierten Zeitschlitzen übertragen, läßt sie die reservierten Zeitschlitze unbelegt passieren und markiert sie als „Unallocated", damit nachfolgende Stationen sie nutzen können.

Bild 2-24 zeigt das Format eines Zeitschlitzes in CRMA. Die Größe des Datenfelds kann beliebig gewählt werden. In einer ATM-Umgebung soll das Datenfeld eine ATM-Zelle (siehe



ACF	Access Control Field
ALLC	Allocation Bit, ermöglicht freien Buszugriff bei Niedriglast
ACC CMD	Access Command
CMD Type	Command Type
Pri	Priority
CMD Data	Command Data, Zyklusnummer oder Zykluslänge CL

* Die Segmentgröße von 53 Oktetts entspricht einer ATM-Zelle. Es kann jedoch auch je nach Anforderung eine beliebige andere konstante Größe gewählt werden.

Bild 2-24: Zeitschlitzformat bei CRMA

Abschnitt 3.1.2) aufnehmen und umfaßt dann 53 Oktetts oder ein ganzzahliges Vielfaches davon. Das Access Control Field beinhaltet das Zugriffskommando („Access Command“), ein Busy-Bit, das den Belegungszustand des Zeitschlitzes anzeigt, und das Allocation-Bit, das anzeigt, ob der Zeitschlitz reserviert ist oder frei belegt werden kann. Das Feld „Command Type“ innerhalb des Access Command legt die Art des Kommandos fest. Ein Bit „Bus“ zeigt an, zu welchem Bus (A oder B) das Kommando gehört, da das Kommando durch die Spiegelung an der gegenüberliegenden Kopfstation auf beiden Bussen unterwegs ist. Das Feld „Pri“ bestimmt die Priorität, zu der das Kommando gehört. Ein Datenfeld „CMD Data“ enthält entweder die Zyklusnummer oder einen von drei in aufeinanderfolgenden Zeitschlitzen übertragenen Teilen der Variablen CL. Tabelle 2-6 enthält die möglichen „Command Types“ sowie die zugehörige Bedeutung der Datenfelder.

Der oben beschriebene Medienzugriffsmechanismus wird für jede Priorität getrennt durchgeführt. Höhere Prioritäten haben unterbrechenden Vorrang, so daß Verkehr niedriger Priorität erst dann übertragen werden kann, wenn alle Reservierungen für die höheren Prioritäten abgearbeitet sind. Ein Reservierungszyklus niedriger Priorität wird durch ein Start-Kommando höherer Priorität unterbrochen und nach Beendigung des höherprioren Zyklus durch ein erneutes Start-Kommando wieder aufgenommen. Damit Pakete innerhalb einer Prioritätsklasse nicht unterbrochen werden, wird in den Stationen für jede Prioritätsklasse jeweils ein eigener Empfangspuffer benötigt.

Kommando	CMD Type	CMD Data
RESERVE_CN	011	Cycle Number
RESERVE_CL0	010	Cycle Length Byte 0
RESERVE_CL1	010	Cycle Length Byte 1
RESERVE_CL2	010	Cycle Length Byte 2
CONFIRM	100	Cycle Number
START	001	Cycle Number
REJECT	101	-
NOOP	000	-

Die Kommandos RESERVE_CN, RESERVE_CLi (i=0..2) werden in vier aufeinanderfolgenden Zeitschlitzen übertragen

Tabelle 2-6: Command Types bei CRMA

Untersuchungen der Leistungsfähigkeit von CRMA zeigen, daß die Auswirkungen hoch ausgelasteter Stationen auf die Dienstgüte der übrigen Stationen durch den Schwellwert-Mechanismus verringert wird [100]. CRMA verteilt die zur Verfügung stehende Bandbreite fair auf die Stationen, vorausgesetzt die Sendepuffer in den Stationen sind ausreichend dimensioniert [4]. Allerdings sind die Wartezeiten der Stationen abhängig von deren Position auf dem Bus [89].

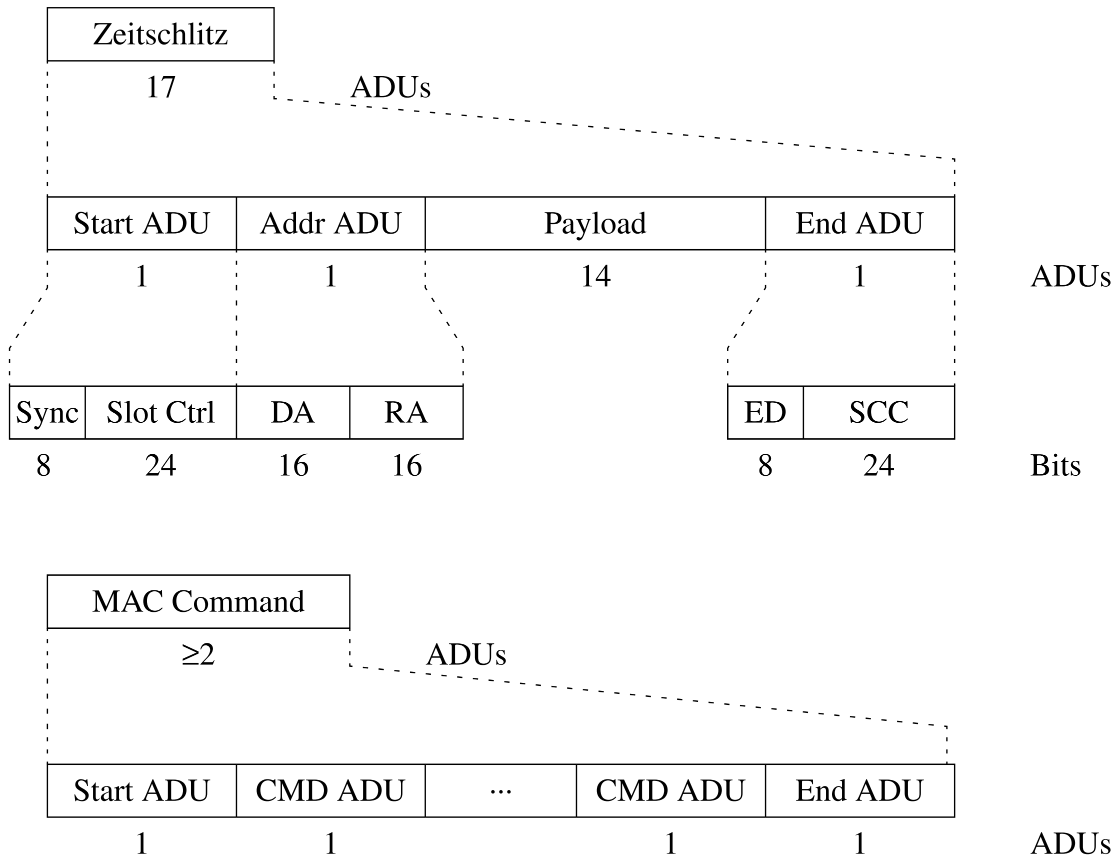
CRMA-II

CRMA-II [7, 9] erweitert das ursprüngliche CRMA-Protokoll auf Slotted Ring- oder Doppelringtopologien. Zusätzlich erlaubt CRMA-II den Stationen einen sofortigen Medienzugriff bei niedriger Last sowie optional das Versenden von Paketen in aufeinanderfolgenden Zeitschlitzen. Die Stationen teilen ihre Sendewünsche einem „Scheduler[1]“ mit, der als zentrale Instanz für eine faire Verteilung der Bandbreite sorgt.

Die kleinste Dateneinheit in CRMA-II ist eine „Atomic Data Unit[2]“ (ADU) von 32 Bits. Zwischen die Zeitschlitze für den Datentransport schiebt der Scheduler nach Bedarf „Access Commands“ ein (Bild 2-25) . Die Zeitschlitze bestehen aus einer Start ADU, die wiederum Steuerinformationen über Art und Inhalt des Zeitschlitzes enthält, einer Address ADU für die Adressierung innerhalb des Rings, der Dateneinheit sowie einer End ADU, die zur Fehlersicherung eine Kopie der Steuerinformation enthält. Die Dateneinheit hat eine Größe von 56 Oktetts und kann somit eine ATM-Zelle aufnehmen. Beim Belegen eines freien („free“) Zeitschlitzes markiert die sendende Station den Zeitschlitz als „gratis“ und „busy“.

Die „Access Commands“ werden vom Scheduler als „Reserve Command“ und „Confirm Command“ ausgesandt. Ein Reserve Command, bestehend aus einer Start und End ADU, geht

1. „Planer“; Zentralinstanz für die Verteilung von Zeitschlitzen
2. Atomare Dateneinheit



ADU	Atomic Data Unit, kleinste Dateneinheit (32 bit)
Addr ADU	Addressing ADU
Sync	Synchronization Part
Slot Ctrl	Slot Control: Slot Type, Priority, Busy/Free, Gratis/Reserved, First/Middle/Last/Only, Other
DA	Destination Address
RA	Removal Address
ED	End Delimiter
SCC	Slot Control Copy
CMD ADU	Command (Parameter) ADU, eine ADU je Station

Bild 2-25: Zeitschlitzformat bei CRMA-II

jedem Reservierungszyklus voraus. Jede sendewillige Station fügt eine ADU, die ihre Reservierungsparameter enthält, unmittelbar nach der Start ADU ein. Nach einem Ringumlauf kommt das Reserve Command beim Scheduler an, der aufgrund der Parameter die Anzahl der Zeitschlitze für den folgenden Zyklus den Stationen zuteilt. Die sich hieraus ergebenden Parameter werden den Stationen wiederum in einem Confirm Command mitgeteilt, aus dem die Stationen ihre Parameter-ADU entfernen und auswerten. Ein Reservierungszyklus beginnt mit dem Aussenden des zugehörigen Confirm Commands durch den Scheduler, der anschließend genau so viele vormals „gratis“ freigegebene Zeitschlitze als „reserved“ markiert, wie er den Stationen zugeteilt hat. Anschließend sendet er das nächste Reserve Command aus.

Als Parameter erhält der Scheduler von jeder Station die Anzahl der Zeitschlitze, die die Station noch belegen will, sowie die Anzahl der Zeitschlitze, die sie bereits durch den vorherigen Zugriff auf als „gratis“ markierte Zeitschlitze belegt hat. Aus den Parametern berechnet der Scheduler einen Schwellwert T als Referenz für den Reservierungszyklus: Einer Station i, die bereits A_i als „gratis“ markierte Zeitschlitze belegt hatte, werden T - A_i Zeitschlitze zugeteilt. Ist A_i > T, muß die Station A_i - T freie „gratis“-Zeitschlitze unbelegt passieren lassen, was als „Deferment“ bezeichnet wird. Für die Gesamtzahl R der durch den Scheduler als „reserviert“ zu markierenden Zeitschlitze in einem Zyklus gilt:

$$R = \sum_{i:\, A_i < T} T - A_i \tag{1}$$

Der Schwellwert wird so gewählt, daß die Gesamtzahl der beeinflußten Zellzugriffe Z (Reservierungen und Deferments) möglichst gering ist, daß also Z gemäß Gleichung (2) ein Minimum erreicht [63].

$$Z = \sum_{i} |T - A_i| \tag{2}$$

Die Adressierung im Ring erfolgt mit der Addressing ADU, die eine Zieladresse DA und eine „Removal Address“ RA enthält. Mit der Zieladresse kann entweder eine einzelne Station oder alle Stationen gleichzeitig erreicht werden (Broadcast). Mit RA wird die Station adressiert, die den Zeitschlitz entfernt. Bei Einzeladressen ist RA identisch mit DA, bei Broadcast ist RA die Adresse der Sendestation.

Prioritäten werden realisiert, indem der beschriebene Mechanismus für jede Priorität getrennt durchgeführt wird, wobei der Scheduler zuerst Zeitschlitze mit höherer Priorität als „reserved“ markiert und Reservierungszyklen niedriger Priorität zugunsten höherer Prioritäten unterbricht. Voraussetzung hierfür ist, daß die Stationen für jede Priorität einen eigenen Sende- und Empfangspuffer besitzen.

Pakete, die größer sind als die Dateneinheit eines Zeitschlitzes, können in aufeinanderfolgenden Zeitschlitzen übertragen werden, wenn die sendende Station über einen Einfügepuffer verfügt. Während einzelne Zeitschlitze als „only“ markiert werden, erhält der erste Zeitschlitz einer solchen Paketübertragung die Kennzeichnung „first“, der letzte „last“ und alle dazwischenliegenden „middle“. Während sie ein Paket aussendet, leitet die sendende Station ankommende Zeitschlitze in den Einfügepuffer um.

Bei CRMA-II erhalten alle sendewilligen Stationen den gleichen Durchsatz [8]. Bei hoher Last steigt der Variationskoeffizient der Wartezeit sowie die Wartezeit stark an, was zu einem enormen Jitter führt [93]. Insgesamt ist die mögliche Mediumauslastung bei CRMA-II sehr hoch, da das Protokoll nur im Überlastfall drosselnd auf die Stationen einwirkt.

Einige Erweiterungen zu CRMA-II wurden vorgeschlagen. In einer ATM-Umgebung kann das Zeitschlitzformat dem ATM-Zellformat angepaßt werden, wobei für die Kommandos eine oder mehrere Zellen verwendet werden [7]. Mit „Distributed Marking“ wird der Durchsatz in einem hoch ausgelasteten System weiter erhöht und die mittlere Wartezeit verringert [68]. Die Integration isochroner Kanäle ist nach Einführung einer Rahmenstruktur möglich [6].

2.4.6 Metaring

Die Metaring-Architektur wurde am IBM T. J. Watson Research Center entwickelt [17]. Sie basiert auf einem gegenläufigen Doppelring mit zwei möglichen Ausprägungen: Bei Verwendung eines Einfügepuffers können Pakete variabler Länge transportiert werden, während in einem System mit Zeitschlitzen kein Einfügepuffer benötigt wird, jedoch die Dateneinheiten eine konstante Länge aufweisen müssen. Die Bedingungen „Einfügepuffer ist leer“ und „Zeitschlitz ist frei“ sind äquivalent; ebenso die Aktionen „Paket senden“ und „Zeitschlitz belegen“. Im folgenden wird nur das System mit Zeitschlitzen betrachtet.

Eine Station darf jeden freien Zeitschlitz solange belegen, wie ihr Zähler COUNT unter einer Schwelle M liegt. Mit jedem belegten Zeitschlitz wird COUNT um 1 inkrementiert. Auf dem gegengerichteten Ring läuft ein Signal „SAT[1]“ um, repräsentiert durch ein Bit im Kopf jedes Zeitschlitzes. Bei einem ankommenden SAT vergleicht eine sendewillige Station ihren Zähler COUNT mit einer zweiten Schwelle N, wobei gilt: M > N. Das SAT-Signal wird weitergeleitet, sobald COUNT $\geq$ N ist oder der Sendepuffer leer ist. Mit dem Weiterleiten des SAT-Signals wird COUNT auf 0 zurückgesetzt. Mit diesem Mechanismus wird sichergestellt, daß mit jedem Umlauf des SAT-Signals alle Stationen mindestens M und höchstens N Zeitschlitze belegen können. Eine globale Fairness, bei der alle Stationen gleichen Durchsatz haben sollen, wird erreicht durch M = N. Falls M > N gewählt wird, ergibt sich eine mögliche Unfairness, da die Differenz der durch zwei Stationen belegten Zeitschlitze in einem SAT-Umlauf N - M beträgt. Ferner können die Schwellen M und N für jede Station individuell festgelegt werden, um einzelnen Stationen mehr Zeitschlitze je SAT-Umlauf zuzuteilen.

Ähnlich einem Token-System müssen vorbeugend Maßnahmen getroffen werden für den Fall, daß das SAT-Signal verloren geht oder ein weiteres SAT-Signal generiert wird. Hierzu verfügt jede Station über einen Zeitgeber, der die Zeit zwischen zwei SAT-Ankünften mißt. Beim Überschreiten der maximalen SAT-Umlaufzeit generiert die Station ein neues SAT-Signal. Wenn bei einer Station ein zweites SAT-Signal eintrifft, bevor sie ein vorher eingetroffenes SAT-Signal weiterleiten konnte, wird das zweite SAT entfernt.

Prioritäten werden eingeführt, indem dem SAT-Signal eine Prioritätsstufe zugeordnet wird. Die Stationen können nur Daten senden, deren Priorität höher oder gleich der Prioritätsstufe des letzten SAT-Signals ist. Eine Station, die Daten höherer Priorität zu senden hat, kann die Prioritätsstufe des SAT-Signals erhöhen, wobei sie die vorhergehende Prioritätsstufe in einem ihrer Register sichert. Kommt das SAT-Signal mit gleicher Prioritätsstufe zurück und hat die Station keine Daten in dieser oder einer höheren Priorität mehr zu senden, erniedrigt sie die SAT-Prioritätsstufe. Durch die Einführung eines weiteren Parameters für das SAT-Signal und die Erweiterung des Mechanismus wird verhindert, daß die Prioritätsstufe zu weit abgesenkt wird, obwohl in anderen Stationen Daten höherer Priorität warten [17].

1. von SATurated (gesättigt, befriedigt).

Synchroner Verkehr wird integriert mit der Einführung eines weiteren Signals „ASYNC-EN[1]“. Dieses Signal wird von Station zu Station entgegen der Datenrichtung weitergereicht und setzt jeweils einen Zeitgeber zurück, der die Zeitspanne zwischen zwei Ankünften von ASYNC-EN mißt. Eine Station, die synchrone Daten senden will, hält ASYNC-EN so lange auf, bis diese Daten gesendet sind. Durch das Aufhalten von ASYNC-EN läuft bei der Vorgängerstation der Zeitgeber ab, was bedeutet, daß diese Station keinen asynchronen Verkehr mehr senden darf und somit auf Bandbreite zugunsten der synchronen Verkehr sendenden Station verzichtet. Eine Erweiterung des Integrationsmechanismus ordnet ASYNC-EN drei Attribute (grün, gelb und rot) zu, um die Durchsatzeinbußen bei Ringen größerer Ausdehnung zu mindern [73].

Weiterführende Untersuchungen behandeln die Ausdehnung des Metaring-Protokolls auf vermaschte Ringe und andere Topologien sowie den Übergang zu Weitverkehrsnetzen [73]. Ein Ansatz mit lokaler Fairness erhöht die Auslastung einzelner Ringsegmente und ist auf andere Protokolle übertragbar [16].

2.4.7 AMNET

Die Integration von Sprache und Bewegtbild in ein lokales Rechnernetz bei relativ geringen Kosten ist das Ziel von AMNET[2], einem unidirektionalen Slotted Ring mit Destination Release, entwickelt an der University of New South Wales in Sydney, Australien [24]. AMNET verwendet Einfügepuffer für Asynchronverkehr und einen Reservierungsmechanismus für synchrone Zeitschlitze. Ein Zeitschlitz ist 54 Oktetts lang und besteht aus einem Header von 6 Oktetts und einem 48 Oktetts umfassenden Informationsfeld. Insgesamt können 256 Stationen am Ring angeschlossen sein.

Will eine Station eine synchrone Verbindung aufbauen, teilt sie der Monitor-Station, die zentral die Verbindungen verwaltet, ihren Reservierungswunsch mit. Die Monitor-Station akzeptiert den Verbindungsaufbauwunsch, wenn die gesamte reservierte Bandbreite inklusive der neuen Verbindung nicht größer ist als die halbe Ringbandbreite, wenn also maximal jeder zweite Zeitschlitz reserviert werden muß.

Die Reservierung eines Zeitschlitzes für einen Reservierungszyklus durch die Master-Station erfolgt immer zur einmaligen Nutzung für Synchronverkehr durch eine Station. Dabei werden zwei Phasen unterschieden. In der ersten Phase markiert die Master-Station einen zu reservierenden Zeitschlitz als „booked“. Der Inhalt des Zeitschlitzes kann während des folgenden Ringumlaufs die Zielstation erreichen und entfernt werden, jedoch darf ein als „booked“ markierter Zeitschlitz nicht mit neuem Synchronverkehr belegt werden. Kommt der Zeitschlitz nach einem Ringumlauf bei der Master-Station an, markiert diese ihn jetzt in der zweiten Phase als „owned“, wobei sie die Adresse der Station, für die der Zeitschlitz reserviert ist, in das dafür vorgesehene Feld im Header schreibt. Die Station, die den Zeitschlitz „besitzt“, nutzt ihn für ihren Synchronverkehr. Durch geeignete Wahl des Vergabealgorithmus in der Master-

1. Asynchronous Traffic Enable
2. A Multimedia NETwork

Station kann die maximale Wartezeit einer Station auf einen Zeitschlitz für den Synchronverkehr minimiert werden.

Der Asynchronverkehr ist von diesem Mechanismus unabhängig und nutzt die gerade nicht von Synchronverkehr belegten Zeitschlitze. Jede Station verfügt über einen Einfügepuffer für den Asynchronverkehr. Wenn ein Zeitschlitz für Synchronverkehr der Station genutzt werden soll, aber mit asynchronen Daten belegt ist, werden die Daten im Einfügepuffer zwischengespeichert. Synchrone Zeitschlitze werden immer am Einfügepuffer vorbei direkt weitergeleitet. Solange der Füllstand des Einfügepuffers einen Schwellwert („soft buffer“) nicht überschreitet, kann die Station einen freien Zeitschlitz für ihren Asynchronverkehr verwenden, während sie bei überschrittenem Schwellwert zuerst Daten aus dem Einfügepuffer senden muß.

Eine faire Aufteilung der asynchronen Bandbreite zwischen den Stationen garantiert ein Credit-Verfahren (siehe 2.3.6.3), mit dessen Hilfe gleichzeitig sicherstellt wird, daß der Einfügepuffer nicht überläuft. Zu Beginn und Ende eines Credit-Zyklus sind alle Einfügepuffer leer. Da die Gesamtzahl der Zeitschlitze mit Asynchronverkehr während eines Zyklus begrenzt ist, kann unter Berücksichtigung der für den Asynchronverkehr zur Verfügung stehenden Bandbreite vor und nach der Station eine Obergrenze für die Anzahl der im Einfügepuffer zwischenzuspeichernden Zeitschlitze angegeben werden. Umgekehrt ist die Anzahl der Credits je Zyklus durch die Größe des Einfügepuffers begrenzt. In einem Beispielnetz mit 256 Stationen und 50% minimal nutzbarer Bandbreite für den Asynchronverkehr, einem Anfangswert der Credit Counter von lediglich 10 sowie einem Schwellwert von 10 muß der Einfügepuffer maximal 1290 Zellen oder 69660 Oktetts fassen [24], was einerseits einen hohen Aufwand bedeutet, andererseits die maximale Durchlaufzeit im System extrem erhöht. Alternativ kann über die Einstellung von Senderaten für Asynchronverkehr verhindert werden, daß der Einfügepuffer überläuft.

Der Mechanismus für die Vergabe der synchronen Zeitschlitze ist unabhängig vom Asynchronverkehr, wodurch es bei geeigneter Wahl des Reservierungsalgorithmus in der Master-Station möglich ist, Isochronverkehr zu integrieren. Allerdings ist das Protokoll nicht für beliebige Bitraten oder Ausdehnungen skalierbar, da die minimale Größe des Einfügepuffers und die Effizienz des Credit-Mechanismus stark von diesen Parametern abhängt. Der „Buffered Access Cycle“-Mechanismus für den Asynchronverkehr kann auf andere Protokollarchitekturen übertragen werden [23].

Kapitel 3

Lokale Netze mit ATM-Vermittlungsknoten (ATM-LANs)

Der Asynchrone Transfermodus (ATM) ist die Übermittlungstechnologie, die dem Breitband-ISDN (B-ISDN) zugrunde liegt. B-ISDN wird von der ITU als breitbandige Alternative zu ISDN im Weitverkehrsbereich standardisiert. Gleichzeitig haben sich eine große Anzahl von Firmen im ATM-Forum zusammengeschlossen, um die ATM-Technologie auch im Bereich der lokalen Netze voranzutreiben und spezifische LAN-Komponenten zu entwickeln. Der Begriff „ATM“ wird im Zusammenhang mit lokalen Netzen als Synonym für ein Netz gebraucht, das auf den Spezifikationen des ATM-Forum sowie ITU-Standards für B-ISDN basiert. In Abschnitt 3.1 wird eine Einführung in die ATM-Technik gegeben.

Ein „ATM-LAN“ besteht aus Endgeräten mit ATM-Schnittstelle und Vermittlungsknoten und ist grundsätzlich verbindungsorientiert. Protokolle für herkömmliche lokale Netze sind überwiegend verbindungslos und basieren auf Mechanismen wie Broadcasts, die in einem ATM-Vermittlungsknoten nur mit zusätzlichem Aufwand zu realisieren sind. Auf lange Sicht wird es Protokolle geben, die für ATM-Netze konzipiert sind. Mit den in Abschnitt 3.2 beschriebenen Möglichkeiten zur Einbindung bestehender LAN-Protokolle kann die ATM-Technologie stufenweise Einzug finden in ein bestehendes lokales Netz.

3.1 ATM-Technik

Bei ATM wird die zu übertragende Information in Zellen konstanter Größe unterteilt oder zusammengefaßt. Die Übertragung der Zellen erfolgt asynchron, es wird also keine feste Zeitbeziehung der Zellen untereinander vorausgesetzt. Durch die Anwendung des statistischen Multiplexverfahrens ist es möglich, die Kanalkapazität zu „überbuchen“, so daß die Summe der Spitzenbitraten aller Verbindungen auf einem Kanal die Kanalkapazität übersteigt. Für kurzzeitige Überlast stehen in den Vermittlungsknoten Puffer geringer Kapazität zur Verfügung, bei deren Überlauf Zellverluste auftreten können.

Die Einordnung der ATM-Technologie und des B-ISDN Referenzmodells [47] in das OSI-Referenzmodell [32, 54] ist insbesondere im LAN-Bereich nicht eindeutig, da hier mit der Adaptionsschicht einerseits ein Transportdienst bereitgestellt wird, andererseits MAC-PDUs von LAN-Protokollen diesen Transportdienst nutzen.

Bild 3-1 zeigt eine vereinfachte Darstellung der Protokollschichten im B-ISDN. Die Bitübertragungsschicht besteht aus einer großen Vielfalt möglicher Übertragungsverfahren und Übertragungsmedien, die von der ITU [52] und vom ATM-Forum [10] definiert wurden. Die ATM-

Höhere Schichten
ATM Adaptionsschicht (AAL)
ATM-Schicht
Bitübertragungsschicht

Bild 3-1: Vereinfachtes B-ISDN Referenzmodell

Schicht stellt einen transparenten, ungesicherten Übermittlungsdienst für SDUs konstanter Größe dar. Die Umsetzung von PDUs höherer Protokollschichten in SDUs der ATM-Schicht erfolgt in der ATM-Adaptionsschicht.

ATM ist verbindungsorientiert, wodurch jeder Verbindung eine Aufbauphase vorangehen und am Ende eine Verbindungsabbauphase folgen muß. Signalisierfunktionen realisieren den Verbindungsaufbau und -abbau, wobei zusätzliche Parameter wie Dienstgüte (QOS), mittlere und maximale Bitrate, Fehlerrate und Art der Verbindung ausgehandelt werden. Die Signalisierfunktionen, die das ATM-Forum definiert [10], beruhen auf dem ITU-Standard Q.2931 [53].

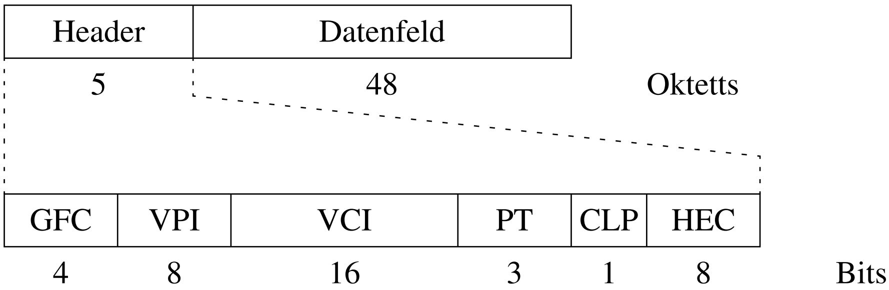
Obwohl nicht zwingend, wird bei einem ATM-LAN üblicherweise von einer hierarchisch-sternförmigen Netztopologie ausgegangen, die in höheren Hierarchieebenen zusätzlich teilvermascht sein kann, um alternative Wege im Fall von Überlast oder Ausfällen zu ermöglichen. Die Übertragung der ATM-Technologie auf andere Topologien wie einen Ring ist jedoch prinzipiell möglich.

3.1.1 ATM UNI

Das User-Network Interface (UNI) spezifiziert die Schnittstelle zwischen einem ATM-Endgerät und der Vermittlungseinrichtung. Die ITU spezifiziert ein UNI der Bitübertragungsschicht [52], das auf der SDH-Übertragungstechnik beruht [44, 45]. Ferner gehören zur UNI-Definition die in den folgenden Abschnitten beschriebenen Schichten „ATM“ und „AAL“. Im lokalen Bereich gilt jedoch die UNI-Spezifikation des ATM-Forum in der derzeitigen Version 3.1 [13] als de facto-Standard, in dem von der Bitübertragungsschicht bis zu Management und Signalisierung alle Eigenschaften der Schnittstelle enthalten sind.

Die Definition der Bitübertragungsschicht im UNI 3.1 beschreibt Übertragungsmedien mit ihren elektrischen oder optischen Eigenschaften sowie die Kodierung. Die Vielfalt an Übertragungsmedien, die durch den Zusammenschluß der einzelnen Firmen entstanden ist, nimmt weiter zu, da durch die Weiterentwicklung der Übertragungstechnik höhere Bitraten möglich werden und daneben die drahtlose Übertragung an Bedeutung gewinnt.

Die ATM-Schicht im UNI 3.1 enthält neben der in Abschnitt 3.1.2 beschriebenen Zelldefinition eine „Layer Management Specification“ sowie Funktionen zu „Traffic Control and Congestion Control“. Weiter enthält die UNI-Spezifikation eine „Interim Local Management Interface Specification“ sowie die UNI-Signalisierung, die auch in einer neueren Version 4.0 vorliegt [10].



GFC	Generic Flow Control
VPI	Virtual Path Identifier
VCI	Virtual Channel Identifier
PT	Payload Type
CLP	Cell Loss Priority
HEC	Header Error Check

Bild 3-2: Zellformat am UNI

3.1.2 ATM-Schicht

Die ATM-Schicht [48] definiert das in Bild 3-2 dargestellte Zellformat am UNI, bestehend aus einem Header von 5 Oktetts und einem 48 Oktetts großen Datenfeld. Das GFC-Feld umfaßt 4 Bits und dient einer möglichen Datenflußsteuerung zwischen dem ATM-Endgerät und der Vermittlungseinrichtung, wird jedoch selten genutzt und meist auf 0 gesetzt. Das Adressfeld besteht aus dem Virtual Path Identifier (VPI) und einem Virtual Channel Identifier (VCI). Der VPI definiert virtuelle Pfade über einen physikalischen Kanal, während der VCI virtuelle Kanäle innerhalb eines virtuellen Pfades bezeichnet [46].

Der Payload Type (PT) wird verwendet, um verschiedene Typen von Zellen zu unterscheiden. Neben der Unterscheidung zwischen Nutzerdaten und Management-Zellen und der Signalisierung einer Überlastsituation kann 1 Bit für die User-to-User Information verwendet werden. In der Adaptionsschicht (AAL-5) wird dieses Bit genutzt, um das Ende einer Paketübertragung zu kennzeichnen. Bestimmte Kombinationen aus VPI/VCI und PT sind reserviert und werden zum Beispiel verwendet, um Leerzellen zu kennzeichnen.

Zellen mit gesetztem Cell Loss Priority (CLP)-Bit werden in einem Vermittlungsknoten bei Überlast bevorzugt verworfen [51]. Die Header Error Control (HEC) realisiert eine Fehlerkorrektur über die 5 Oktetts des Header [52]. In einigen Varianten der Bitübertragungsschicht wird mit dem HEC gleichzeitig die Zellgrenzenerkennung (Cell Delineation) realisiert.

3.1.3 Adaptionsschicht

Die ATM-Adaptionsschicht (AAL) [49, 50] erweitert den Dienst der ATM-Schicht um Funktionen, die von höheren Schichten benötigt werden. Da die Aufgaben von den Anforderungen der höheren Schichten abhängen, wurden fünf verschiedene Adaptionsschichten AAL-1 bis

AAL-5 definiert, die diesen Anforderungen jeweils Rechnung tragen sollen. Das Aufteilen und Vereinigen (Segmentation and Reassembly, SAR) sowie Anpassungsfunktionen bilden jeweils eigene Subschichten innerhalb des AAL.

Die SAR-Subschicht bildet PDUs höherer Schichten auf SDUs konstanter Größe in der ATM-Schicht ab, wobei abhängig vom AAL-Typ einige Oktetts des Informationsfeldes in jeder Zelle für AAL-Funktionen benötigt werden. Lediglich bei AAL-5 stehen alle 48 Oktetts für die PDU bereit. PDUs, die größer sind als das verbleibende Informationsfeld, werden segmentiert, während bei AAL-1 kleinere PDUs zusammengesetzt und in einer ATM-Zelle übertragen werden[1].

Die CS-Subschicht[2] stellt Anpassungsfunktionen bereit, die die Funktionalität der SAR-Schicht um zusätzliche Funktionen für höhere Schichten erweitert. Die CS-Schicht ist weiter unterteilt in einen für alle auf AAL-5 basierende Dienste vorhandenen gemeinsamen Teil „Common Part Convergence Sublayer“ (CPCS) und einen dienstspezifischen Teil „Service Specific Convergence Sublayer“ (SSCS)[3]. Mögliche Funktionen sind die Bereitstellung einer gesicherten Verbindung, der Ausgleich von Takt- und Zellschwankungen und die Einführung einer Flußsteuerung. Die Varianten der Adaptionsschicht AAL-1 bis AAL-5 enthalten verschiedene Kombinationen der möglichen Subschicht-Funktionen. Ein „AAL-0“ als leere Adaptionsschicht ermöglicht den direkten Zugriff auf die Dienste der ATM-Schicht durch Protokolle höherer Schichten.

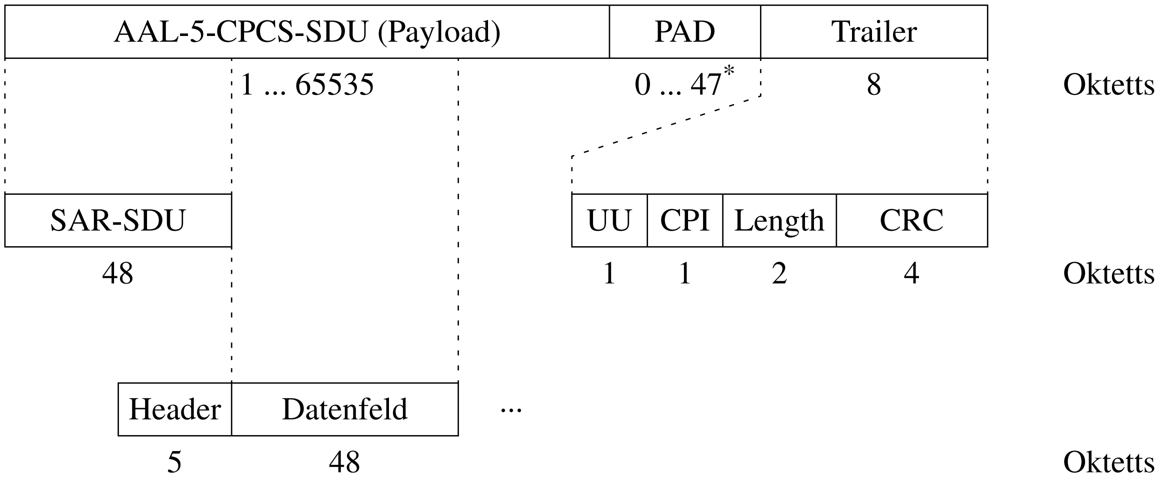
AAL-1 stellt einen synchronen Dienst für Quellen mit konstanter Bitrate bereit. Synchronisierinformation zur Taktrückgewinnung wird mit übertragen. Die optionale Sicherung beschränkt sich auf das Erkennen von Zellverlusten und fehlerhaften Zellen. Geeignet ist AAL-1 für die Übertragung von Telefon- oder konstantbitratigem Videoverkehr.

Die Dienste des AAL-2 entsprechen denen des AAL-1, jedoch für synchrone Quellen mit variabler Bitrate, in erster Linie also für komprimierten Videoverkehr[4].

AAL-3 und AAL-4 waren ursprünglich getrennte Adaptionsschichten, wurden dann aber zu einem gemeinsamen AAL-3/4 erweitert. Es wird ein gesicherter Dienst für verbindungslosen oder verbindungsorientierten Datendienst bereitgestellt. Eine AAL-3/4-SDU kann 1-65535 Oktetts umfassen. Jede SDU wird unterteilt in AAL-3/4-PDUs, die maximal 44 Oktetts Nutzdaten enthalten. Vier Oktetts in der PDU werden zur Reihenfolge- und Fehlersicherung sowie für Längenangaben genutzt. Wird eine AAL-3/4-PDU fehlerhaft übertragen oder geht verloren, wird die gesamte AAL-3/4-SDU weggeworfen.

AAL-5 hat eine ähnliche Funktionalität wie AAL-3/4 bei wesentlich einfacherer Struktur. Bild 3-3 zeigt das Datenformat der CPCS-PDU und die Segmentierfunktion im AAL-5. Die CPCS-PDU besteht aus dem Informationsteil, der 1 bis 65535 Oktetts lang sein kann, dem

1. AAL-1-PDUs bestehen in der Regel aus einem Oktett
2. Convergence Sublayer
3. in AAL-1 bis AAL-5 ist nur CPCS spezifiziert
4. Dieser ursprüngliche AAL2-Dienst wurde 1997 neu definiert und wird speziell für den Transport kleinerer Dateneinheiten an der Schnittstelle von Mobilkommunikationsnetzen und ATM-Kernnetzen diskutiert.



PAD	Padding-Feld
UU	CPCS User-to-User Information
CPI	Common Part Indicator
CRC	Cyclic Redundancy Check

* Die Länge des PAD-Feldes wird so gewählt, daß die Gesamtlänge der AAL-5-CPCS-PDU ein ganzzahliges Vielfaches von 48 beträgt.

Bild 3-3: Format und Segmentierung einer AAL-5-CPCS-PDU

PAD-Feld und einem Trailer. Die CPCS-PDU wird durch PAD-Oktetts auf ein ganzzahliges Vielfaches von 48 aufgefüllt.

Der Trailer ist 8 Oktetts lang und besteht aus einem User-to-User-Informationsfeld, einem Common Part Indicator, dessen Inhalt noch nicht spezifiziert und deshalb auf Null gesetzt ist, einer Längenangabe, die sich auf die AAL-5-CPCS-SDU bezieht, sowie einem CRC-Prüffeld, das zur Sicherung der gesamten CPCS-PDU dient.

Die Funktion der SAR-Subschicht beschränkt sich auf das Segmentieren der CPCS-PDU in Einheiten zu je 48 Oktetts, die an die ATM-Schicht weitergereicht und als SDU jeweils einer Zelle versandt werden. Um die Paketgrenzen beim Empfänger richtig erkennen zu können, wird die User-to-User-Signalisierung im PT-Feld des Zellkopfes verwendet, um die letzte Zelle einer CPCS-PDU zu kennzeichnen. Nach Erhalt der letzten Zelle setzt der Empfänger die CPCS-PDU wieder zusammen. Ist die PDU fehlerhaft, weil eine Zelle verloren oder gefälscht wurde, wirft der Empfänger die gesamte PDU weg.

3.2 Einbindung bestehender LAN-Protokolle

Zur Einführung der ATM-Technologie in ein bestehendes LAN gibt es zwei Möglichkeiten: Entweder sämtliche bestehende LAN-Technologie einschließlich darauf beruhender Software wird gleichzeitig durch neue, auf ATM basierende Technologie und Software ersetzt, oder aber es wird ein Migrationspfad geschaffen, der die ATM-LAN-Technologie in ein bestehendes

Netz mit seiner Infrastruktur, Protokollen und Anwendungssoftware integriert. Da die erste Möglichkeit unter anderem aus organisatorischen und finanziellen Gründen meist ausscheidet, müssen zwangsläufig bestehende LAN-Protokolle in ein integriertes ATM-LAN eingebunden werden.

Die meisten lokalen Netze basieren wie das weltweit zugängliche Internet mit seinem Informationsdienst World Wide Web (WWW) auf dem Internet-Protokoll (IP) [79] und den darauf aufsetzenden Transportprotokollen Transmission Control Protocol (TCP) [80], User Datagram Protocol (UDP) [78] sowie dem Address Resolution Protocol (ARP) [81] zur Umsetzung der Internet- in MAC-Adressen [18]. Ziel der Integration muß sein, für die auf dem Internet-Protokoll basierenden Anwendungen mit ATM eine transparente Übertragungsschicht zur Verfügung zu stellen, wobei die Internet-Adressen in B-ISDN-Adressen nach dem ITU-Standard E.164 [43] oder ATM-Adressen nach ATM-Forum UNI 3.1 [13] mit seinen verschiedenen Varianten umgesetzt werden müssen. Ein möglicher Mechanismus hierfür, „Classical IP and ARP over ATM“, ist in Abschnitt 3.2.1 beschrieben.

Existieren in einem lokalen Netz neben dem Internet-Protokoll noch andere Protokollarchitekturen, dann muß die Integration auf einer niedrigeren Protokollschicht aufsetzen. Die in Abschnitt 3.2.2 beschriebene LAN-Emulation des ATM-Forum stellt einen transparenten Übertragungsdienst für MAC-Pakete eines „virtuellen LAN“ bereit [85, 55].

3.2.1 Internet-Protokoll über ATM

Im „Classical IP and ARP over ATM“ nach RFC 1577 [82] wird ein logisches Subnetz (Logical IP Subnet, LIS) definiert, in dem LAN-Segmente durch ein ATM-Netz ersetzt werden. Dabei bleibt die eigentliche Struktur des LAN mit auf IP basierenden Endgeräten und Routern innerhalb des LIS bestehen. Die Adressen, die im IP hierarchisch gegliedert sind, bleiben im IP-Subnetz innerhalb der niedrigsten Hierarchiestufe. Verbindungen nach außen benötigen einen IP-Router.

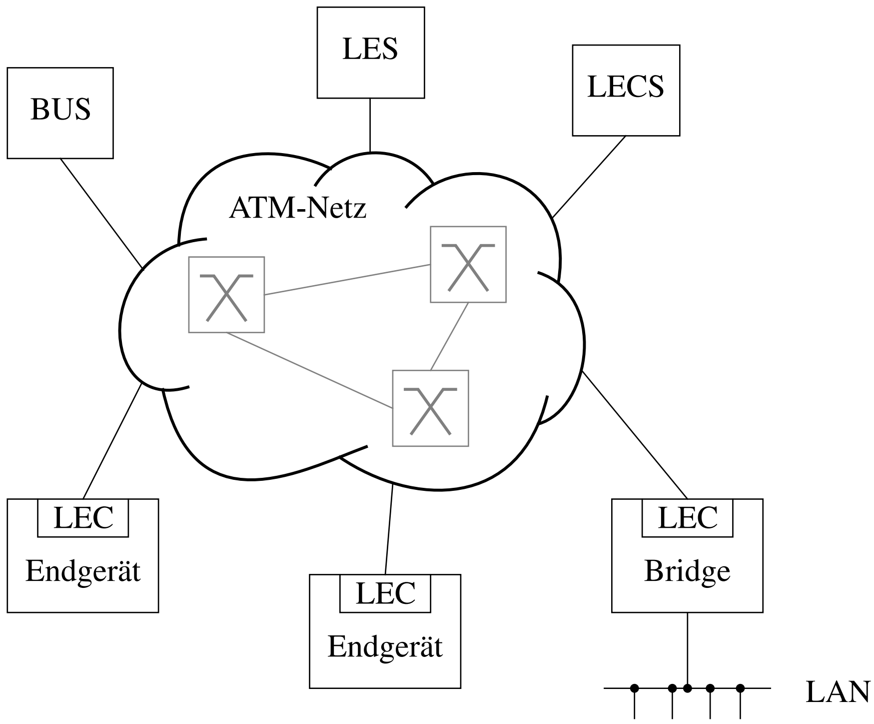
Jeder ATM-Zugang eines Endgeräts hat eine eindeutige ATM-Adresse, die im RFC 1577 als „ATM Hardware Address“ bezeichnet wird. Außerdem kennt jede Station die ATM-Adresse eines ATMARP-Servers innerhalb des Subnetzes, der die IP-Adressen in ATM-Adressen umsetzt und hierfür eine Tabelle führt. Das Address Resolution Protocol (ARP) ist in jeder Station derart modifiziert, daß es eine ARP-Anfrage an den ATMARP-Server adressiert, statt sie per Broadcast im LAN zu verteilen. Der Dienstzugang für Protokolle höherer Schichten bleibt jedoch unverändert.

Eine Station, die eine Verbindung zu einer anderen Station im LIS aufbauen will, deren ATM-Adresse sie nicht kennt, baut zunächst eine ATM-Verbindung zum ATMARP-Server auf. Von diesem erhält sie die ATM-Adresse der Zielstation. Anschließend wird die Verbindung zum ATMARP-Server wieder ab- und die Verbindung zur Zielstation aufgebaut. Wenn die Adressierung mit vollständigen Adressen nach E.164 erfolgt, ist das LIS nicht lokal begrenzt und kann prinzipiell jede Station beinhalten, die über ATM erreichbar ist. Allerdings ist die Anzahl der Stationen im LIS durch die Subnetz-Struktur beschränkt. Broadcast und Multicast, wie im IP vorgesehen, sind nicht spezifiziert.

Die IP-PDUs werden zunächst in Pakete der Logical Link Control (LLC)-Schicht [33] eingepackt und als AAL-5-SDUs an die ATM-Protokollsoftware übergeben. Diese prüft, ob eine Verbindung zur Zielstation besteht und baut gegebenenfalls eine Verbindung auf. Die maximale Größe eines IP-Pakets wurde auf 9180 Oktetts festgelegt, wobei durch die LLC-Schicht weitere 8 Oktetts hinzukommen. Die maximale Größe einer AAL-5-SDU ist somit 9188 Oktetts, so daß ein IP-Paket in maximal 192 Zellen aufgeteilt wird.

3.2.2 LAN Emulation

LAN Emulation [101] wurde vom ATM-Forum standardisiert [11] und ermöglicht die Emulation eines physikalischen LAN nach IEEE 802.3 (Ethernet) [34] oder 802.5 (Token Ring) [36] über ein ATM-Netz. Ein emuliertes LAN besteht aus einem LAN Emulation Server (LES) und mehreren LAN Emulation Clients (LEC) sowie einem Broadcast and Unknown Server (BUS) und einem LAN Emulation Configuration Server (LECS). Ein LEC ist ein ATM-Endgerät, das als reines Endgerät oder als Bridge [38] ausgeführt sein kann und dementsprechend eine oder mehrere MAC-Adressen repräsentiert. Bild 3-4 zeigt eine Beispielkonfiguration für ein emu-



LES	LAN Emulation Server
LEC	LAN Emulation Client
BUS	Broadcast and Unknown Server
LECS	LAN Emulation Configuration Server

Bild 3-4: Beispielkonfiguration eines emulierten LAN

liertes LAN. Ein emuliertes Ethernet kann eine Kombination aus Ethernet-Segmenten und ATM-Endgeräten enthalten.

In einem ATM-Netz können mehrere emulierte LANs koexistieren, die jeweils durch ihren eigenen LES repräsentiert werden. Emulierte LANs sind disjunkt: Jedem LEC ist genau ein LES zugeordnet. Soll ein Datenaustausch zwischen zwei emulierten LANs stattfinden, muß zwischen den beiden LANs ein Router eingesetzt werden, der die emulierten LANs auf einer höheren Protokollschicht koppelt.

3.2.2.1 Adressumsetzung im LAN Emulation Server

Die Kommunikation zweier Endgeräte über das ATM-Netz erfolgt mit einer virtuellen Verbindung. Vor dem Aufbau einer virtuellen Verbindung über das ATM-Netz muß der initiierende LEC mit Hilfe des LES die zu der MAC-Zieladresse gehörende ATM-Zieladresse ermitteln. Der LES ersetzt die in einem lokalen Netz üblichen Adressauflösungsprotokolle, die auf Broadcasts an alle Stationen basieren, durch eine zentralisierte Adressauflösung.

Der LES führt eine Tabelle der MAC-Adressen und zugehörigen ATM-Adressen aller in seinem emulierten LAN angemeldeten Endgeräte. Ein neues oder gerade angeschaltetes ATM-Endgerät meldet sich mit seiner MAC-Adresse beim LES an, woraufhin dieser die Adresse mit der korrespondierenden ATM-Adresse in die Tabelle einträgt.

Soll eine neue Verbindung zwischen zwei Endgeräten aufgebaut werden, sendet die initierende Station eine Anforderung LE_ARP mit der MAC-Adresse der Zielstation an den LES. Dieser sucht in seiner Tabelle nach der zugehörigen ATM-Adresse und teilt sie der anfragenden Station mit. Ist für die angefragte MAC-Adresse kein Eintrag in der Tabelle vorhanden, gibt der LES die Anforderung an alle registrierten LEC weiter. Ein Endgerät beantwortet die Anfrage mit seiner ATM-Adresse, sofern seine MAC-Adresse mit der angefragten Adresse übereinstimmt. Eine Bridge prüft, ob die angefragte MAC-Adresse in ihrem LAN-Segment vorhanden ist [38]. Falls ja, beantwortet sie die Anfrage mit ihrer eigenen ATM-Adresse. Der LES ergänzt seine Tabelle und gibt der anfordernden Station die angefragte ATM-Adresse weiter.

3.2.2.2 LAN Emulation Datenformat

Für Datenpakete sind in LAN Emulation zwei Formate für Ethernet und Token Ring vorgesehen. Bild 3-5 zeigt das Datenformat für Ethernet-Pakete, das als AAL-5-SDU interpretiert wird. Das Datenformat entspricht dem MAC-Format nach IEEE 802.3 [34], wobei Präambel, Start Delimiter und Frame Check Sequence entfallen, mit einem vorangestellten LAN Emulation Header. Entsprechend sind die Datenformate für Token Ring und für Steuerdaten definiert.

3.2.2.3 Broadcast and Unknown Server

Die virtuelle Verbindung im ATM-Netz ist eine Punkt-zu-Punkt-Verbindung. Der Broadcast and Unknown Server (BUS) stellt einen Dienst bereit, um die in einem LAN auftretenden Broadcast- und Multicast-Pakete zu kopieren und an die Zielstationen im emulierten LAN wei-

Header	DA	SA	Len	Datenfeld	
2	6	6	2	46...1500	Oktetts

Header	LAN Emulation Header
DA	Destination MAC Address
SA	Source MAC Address
Len	Length/Type Field

Bild 3-5: LANE-Datenformat für Ethernet-Pakete

terzuleiten. Bei einer Anfrage LE_ARP durch einen LEC nach der Broadcast-Adresse antwortet der LES mit der ATM-Adresse des BUS.

Eine weitere Funktion des BUS ist die Weiterleitung von Paketen, deren ATM-Zieladresse noch nicht bekannt ist, während der Verbindungsaufbauphase. Will ein LEC ein Paket zu einer bisher nicht bekannten MAC-Adresse senden, muß er zunächst die ATM-Adresse des Ziels ermitteln und anschließend eine Verbindung zum Ziel aufbauen. Da dieser Vorgang geraume Zeit in Anspruch nehmen kann und möglicherweise zu einem Zeitablauf in der sendenden Station führt, werden alle Pakete, deren ATM-Zieladresse nicht bekannt ist, zum BUS geschickt, der sie an alle Stationen im emulierten LAN weiterleitet. Damit wird die Shared-Media-Eigenschaft im lokalen Netz emuliert und das Paket kommt rechtzeitig bei der Zielstation an. Sobald die Verbindung aufgebaut ist, werden die Pakete dann statt zum BUS über diese Verbindung übertragen.

3.2.2.4 LAN Emulation Configuration Server

Existieren in einem ATM-Netz mehrere emulierte LANs, muß beim Initialisieren eines LEC die ATM-Adresse des für ihn zuständigen LES ermittelt werden, um dorthin eine Verbindung aufbauen zu können. Die Adresse wird mit Hilfe des LAN Emulation Configuration Server (LECS) ermittelt. Zunächst wird über eine allgemeingültige ATM-Adresse oder einen vergleichbaren Mechanismus eine Verbindung zum LECS aufgebaut. Dieser verfügt über eine Konfigurationsdatei, aus der er den für die anfragende Station zuständigen LES entnimmt.

3.2.3 MPOA

Sollen bei LAN Emulation zwei logische Subnetze gekoppelt werden, wird ein herkömmlicher Router benötigt, der die Subnetze auf der Vermittlungsschicht koppelt. Hierzu müssen die ATM-Zellen zu LAN-Paketen vereinigt werden, um die Zieladresse festzustellen. Anschließend werden die Pakete wieder in ATM-Zellen aufgeteilt. Da jedes Paket der gleichen Verbindung diese Prozedur durchläuft, muß der Router einen hohen Durchsatz bewältigen können.

Das ATM-Forum entwickelt mit „Multi-Protocol Over ATM“ (MPOA) [12] einen alternativen Ansatz, der den Aufwand wesentlich reduziert und als „Virtual Routing“ bezeichnet wird. Dabei wird die einfache Weiterleitung von Paketen („Forwarding“) von der eigentlichen Routing-Funktion, also dem Austausch von Wegeinformationen, getrennt. Gesteuert vom „Route Server“ übernehmen „Forwarder“ die Weiterleitung von Paketen der Vermittlungsschicht, solange noch keine direkte Verbindung innerhalb der ATM-Schicht besteht. Sobald eine ATM-Verbindung zur Zielstation aufgebaut ist, werden die ATM-Zellen direkt, also ohne Routerfunktion, zur Zielstation übermittelt.

Im Gegensatz zur LAN Emulation setzen bei MPOA die Protokolle der Vermittlungsschicht, z. B. das Internet Protocol (IP), direkt auf der ATM-Schicht auf. Damit entfällt die Beschränkung der Paketlänge, die bei LAN Emulation durch das zugrundeliegende Medienzugriffsprotokoll (Ethernet oder Token Ring) vorgegeben ist. Stationen, die mit ATM-Karten ausgestattet sind („ATM-attached Hosts“) können somit IP-Pakete austauschen, deren Größe nur durch das IP-Protokoll begrenzt ist.

Kapitel 4

Ein Architekturvorschlag für Gigabit-LANs

In diesem Kapitel wird das Medienzugriffsverfahren eines neuen Gigabit-LAN-Protokolls vorgestellt. Zunächst wird in Abschnitt 4.1 ein Vergleich zwischen ATM- und Shared Media LANs gezogen. Abschnitt 4.2 definiert die Anforderungen, die das Gigabit-LAN-Protokoll erfüllen soll. Die hieraus resultierende Protokollarchitektur ist in Kapitel 4.3 und die Struktur eines Knotens in Kapitel 4.4 beschrieben.

4.1 Vergleich von ATM-LANs mit Shared Media LANs

ATM-LANs unterscheiden sich in ihrer Struktur grundsätzlich von herkömmlichen lokalen Netzen, die auf einem gemeinsamen Übertragungsmedium (Shared Medium) beruhen. Da die Funktionalität für Protokolle höherer Schichten zumindest bei LAN Emulation mit der von Shared Media LANs identisch ist, kann sich ein Vergleich lediglich auf Parameter wie Kosten, Effizienz, Sicherheit und Strukturen beziehen [26]. Tabelle 4-1 zeigt beispielhaft eine Gegenüberstellung von Ethernet und einem im ATM-LAN emulierten Ethernet.

Parameter	„klassisches“ Ethernet	emuliertes Ethernet
Kosten	Ethernet-Adapterkarten, Ethernet-Software, Kabel, evtl. Hub oder Ethernet-Switch	ATM-Adapterkarten, Ethernet-Software, LEC-Software, AAL/ATM-Software, Signalisiersoftware, Kabel, ATM-Switch, LES, BUS, LECS
Initialisierung	Keine	Verbindungsaufbau zum LECS, Bestimmen der LES-ATM-Adresse, Verbindungsaufbau zum LES, Anmeldung am LES, Bestimmen der BUS-ATM-Adresse, Verbindungsaufbau zum BUS
Verbindungsaufbau	Implizit durch Zieladresse	LE_ARP zum LES, Pakete zum BUS, Bestimmen der Ziel-ATM-Adresse, Verbindungsaufbau zum Ziel, Umleiten der Pakete über neue Verbindung

Tabelle 4-1: Gegenüberstellung „klassisches“ Ethernet und emuliertes Ethernet

Parameter	„klassisches“ Ethernet	emuliertes Ethernet
Übertragungsrate	10 Mbit/s (neuer Standard mit 100 Mbit/s verfügbar)	Implementierungsabhängig, z.B. 155 Mbit/s
Durchsatz-spitzenwert	10 Mbit/s; bei Switched Ethernet und n disjunkten Verkehrsbeziehungen $n \cdot 10\text{Mbit/s}$	Beliebig skalierbar, abhängig vom Vermittlungsknoten
Konfiguration	Physikalisches Netz	Virtuelles Netz
Flexibilität	Umkonfiguration erfordert Umstecken von Kabeln	Umkonfiguration eines emulierten LANs durch Ändern der LECS-Datei
Ausdehnung	Beschränkte Ausdehnung	Beliebige Ausdehnung, möglicherweise weltweit verteiltes Netz
Parallelität	Maximal 1 Paket auf dem Netz[a]	Mehrere Pakete auf disjunkten Verbindungen gleichzeitig
Pufferbedarf beim Empfänger	1 Paket	1 Paket je virtueller Verbindung
Pufferbedarf im Netz	Bei Switched Ethernet sowie bei Bridges und Routern geringer, sonst kein Pufferbedarf	Großer Pufferbedarf in jedem Vermittlungsknoten
Punkt-zu-Mehr-punkt-Verbindungen	Pakete erreichen ohne Mehraufwand alle möglichen Empfänger auf dem gleichen Medium	Dateneinheiten müssen für jeden Empfänger explizit kopiert werden
Dienstgüte	Keine garantierte Dienstgüte	Dienstgüte wird bei jedem Verbindungsaufbau ausgehandelt

Tabelle 4-1: Gegenüberstellung „klassisches“ Ethernet und emuliertes Ethernet

a. Bei „Switched Ethernet“ können auch mehrere Pakete gleichzeitig auf dem Netz sein, wenn ihre Wege disjunkt sind.

Aufgrund der hohen Anfangskosten für ein ATM-LAN ist der Einsatz der LAN Emulation eher für sehr große Netze oder die Vernetzung mehrerer Standorte einer Firma geeignet. ATM-LANs bieten bei gleicher Bitrate einen höheren Gesamtdurchsatz im Netz gegenüber Shared Media-LANs, wenn auf mehreren Verbindungen gleichzeitig Pakete übertragen werden können. Allerdings setzt dies voraus, daß die Verkehrsbeziehungen im Netz relativ homogen sind und sich nicht auf einen einzelnen Server im Netz konzentrieren.

Anwendungen, die auf ATM basieren, können das Aushandeln der Dienstgüte beim Verbindungsaufbau nutzen, um neuartige Dienste zu implementieren, die auf die Einhaltung von Dienstgüteparametern wie Mindestdurchsatz oder Maximalverzögerung angewiesen sind. Ein vollständiger Ersatz der auf bestehenden LAN-Protokollen beruhenden Software durch ATM-basierte Lösungen ist in absehbarer Zukunft kaum zu erwarten, da herkömmliche LAN-Stationen nicht mehr in der Lage wären, diese Software zu nutzen. In diesem Fall müßte für einen gleitenden Übergang hin zu ATM-Anwendungen ein Protokoll implementiert werden, das ATM-Zellen in LAN-Paketen überträgt – mit der Folge, daß gerade der Hauptvorteil von ATM, nämlich eine garantierte Dienstgüte, nicht mehr bereitgestellt werden kann. Dies stellt aber den Sinn einer solchen Lösung in Frage.

Insgesamt ist die Frage „ATM- oder Shared Media-LAN“ von vielen Umgebungsparametern abhängig und läßt derzeit keine eindeutige und allgemeingültige Antwort zugunsten eines der Systeme zu.

4.2 Anforderungen an Gigabit-LANs

Die Entwicklung im Bereich der lokalen Netze zeigt in erster Linie einen Bedarf an höheren Übertragungsraten auf, dem Protokollvorschläge für Gigabit-LANs Rechnung tragen. Daneben gibt es weitere Kriterien, nach denen LAN-Vorschläge zu bewerten sind. Insbesondere sollen multimediale Anwendungen und eine einfache Anbindung an ein Weitverkehrsnetz unterstützt werden. Ein neues Protokoll sollte nach Möglichkeit die Vorteile der beiden „Welten“ ATM-LAN und LAN mit Vielfachzugriff integrieren und die Nachteile vermeiden. Im folgenden werden einige Kriterien für neue LAN-Protokolle vorgestellt sowie bestehende und vorgeschlagene Netzarchitekturen hinsichtlich dieser Kriterien verglichen und bewertet.

4.2.1 Multimedia-Anwendungen

Im „multimedialen Zeitalter“ gewinnen zunehmend zeitkritische Anwendungen an Bedeutung. Dazu gehören insbesondere Bewegtbild-Anwendungen (Verteil- und Abrufdienst für Video[1]) von einem Video-Server zu einem Terminal oder interaktiv zwischen Terminals in Form einer Videokonferenz. Video-Datenströme, bei denen in äquidistanten Abständen Bilder konstanter Auflösung und Farbtiefe digitalisiert und übertragen werden, benötigen eine hohe konstante Bitrate[2]. Bei Redundanz- oder Irrelevanzreduktion kann die mittlere Bitrate drastisch gesenkt werden. Der resultierende Datenstrom für eine derart komprimierte Video-Übertragung hat eine variable Bitrate, wobei die Spitzenbitrate unter der Datenrate im nicht komprimierenden System liegt.

Das ideale Übertragungssystem für derartige Video-Dienste ist ein isochroner Kanal, dessen Bandbreite mit der vom Video-Dienst benötigten Spitzenbitrate identisch ist. Die Komprimierung des Ausgangssignals führt in einem derartigen System dazu, daß der Kanal – abhängig vom Grad der Komprimierung – nur zum Teil ausgenutzt wird, was zu einer Verschwendung von Bandbreite führt, sofern die Bandbreite nicht anderweitig genutzt werden kann. Der Vorteil ist eine sehr geringe, konstante Verzögerungszeit für jedes Bild.

Bei einem System, in dem sich mehrere Verkehrsströme gemeinsam genutzte Bandbreite asynchron teilen, ist die Verzögerungszeit nicht konstant. Zum Ausgleich der Verzögerungsschwan-

1. Video ist eine Kombination aus Bewegtbild und dazu synchronem, ein- oder mehrkanaligem Ton
2. Übertragungsrate = Anzahl der Bits pro Bild · Bildwiederholrate

kungen wird ein Ausspielpuffer benötigt, der das ankommende Signal verzögert, also eine zusätzliche Verzögerung einfügt. Ist eine Obergrenze für die Verzögerungszeit im Netz angebbar, kann der Ausspielpuffer auf Kosten der Verzögerungszeit so groß gewählt werden, daß er die Differenz zwischen minimaler und maximaler Verzögerungszeit ausgleichen kann. Ist die Obergrenze nicht bekannt, muß der Ausspielpuffer derart dimensioniert werden, daß möglichst wenig Verluste dadurch entstehen, daß der Puffer leerläuft. Verluste durch zu große Verzögerung führen bei Video-Diensten zu einer verringerten Dienstqualität. Um die Verzögerungszeiten für Video-Dienste in einem heterogenen Netz mit Vielfachzugriff gering zu halten, wird ein Prioritätensystem eingeführt, das zeitkritischen Verkehr bevorzugt.

Unter den Protokollen und Protokollvorschlägen für LANs mit Vielfachzugriff nach Kapitel 2 sind einige mit integrierten isochronen Kanälen zu finden. Den Protokollen ist, abgesehen von AMNET, gemein, daß die reservierte Bandbreite exklusiv für isochrone Kanäle genutzt wird und nicht für asynchronen Paketverkehr zur Verfügung steht, selbst wenn ein Kanal nicht voll ausgelastet ist. AMNET ermöglicht die Nutzung der reservierten, aber momentan nicht genutzten Bandbreite, jedoch auf Kosten sehr großer Einfügepuffer mit den damit verbundenen sehr großen Schwankungen der Übertragungszeit für den Asynchronverkehr.

Die meisten LAN-Protokolle mit Vielfachzugriffsverfahren ermöglichen die Priorisierung von zeitkritischem Verkehr gegenüber dem asynchronen Datenverkehr. Allerdings ist die Obergrenze für die Verzögerungszeit zwischen Sender und Empfänger meist relativ hoch, so daß die Qualität von Videokonferenzen, die mit steigender Verzögerungszeit abnimmt, nicht optimal sein kann. Auch bei lokalen Netzen, die auf der ATM-Technik beruhen, ist die Verzögerungszeit nicht konstant und es kann Verluste bei der Übertragung geben, die sich negativ auf die Qualität des Dienstes auswirken.

Ein optimales Medienzugriffsprotokoll für Gigabit-LANs sollte die Reservierung isochroner Kanäle ermöglichen, dabei aber die gerade nicht genutzte Restbandbreite weitestgehend für Asynchronverkehr zur Verfügung stellen. Das Protokoll darf keine unerfüllbaren Anforderungen an die Systemressourcen stellen, insbesondere an die Anzahl und Größe von Puffern. In Abschnitt 4.3 wird ein neues Protokoll vorgestellt, dessen Medienzugriffsverfahren erstmals diese Vorgaben erfüllt.

4.2.2 Skalierbare Architektur für hohen Durchsatz

Der Bedarf an höheren Bitraten wächst in einem lokalen Netz mit der Leistungsfähigkeit der angeschlossenen Terminals und Rechner. Ein für höhere Bitraten skalierbares Protokoll nutzt die zur Verfügung stehende Bandbreite bei höheren Bitraten ähnlich effizient wie bei niedrigen Bitraten: Die maximale Auslastung ρ_{max} weist keine oder nur eine geringe Abhängigkeit von der Medienbitrate auf.

Maßgeblich für die Leistungsfähigkeit eines Medienzugriffsprotokolls ist das Produkt aus Medienausdehnung und Bitrate. Der a-Faktor[1] gibt die Anzahl der Dateneinheiten an, die

1. $a = \frac{\text{äquivalente Medienlänge in bits}}{\text{Blocklänge in bits}}$

gleichzeitig auf dem Medium Platz finden. Für herkömmliche Protokolle wie Token Ring oder Ethernet gilt a « 1, während für Gigabit-LANs der a-Faktor im Bereich zwischen 100 und 1000 liegt. Skalierbar ist ein Protokoll, wenn der Durchsatz unabhängig von a ist, wie in Bild 4-1 dargestellt.

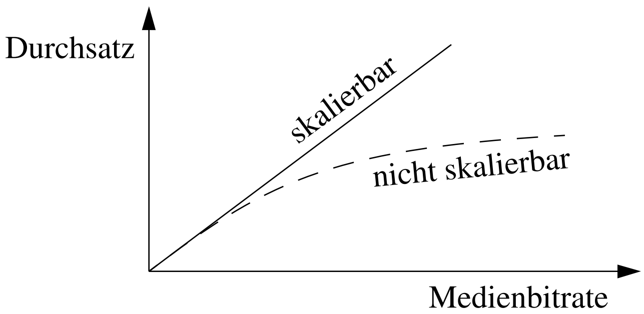


Bild 4-1: Skalierbarkeit von Medienzugriffsprotokollen

4.2.3 Verteildienst

Der überwiegende Anteil des Verkehrsaufkommens in einem lokalen Netz ist Teil von Punkt-zu-Punkt-Verbindungen. Punkt-zu-Mehrpunkt- oder Rundsendeverbindungen werden für interaktive Konferenzdienste, Verteildienste und zur Adressauflösung benötigt. Die Daten müssen entweder kopiert werden oder stehen, bei einem lokalen Netz mit Vielfachzugriff, ohnehin bei allen Stationen zur Verfügung. Mit einer geeigneten Stationsstruktur kann verhindert werden, daß die Medienzugriffseinheit einer Station Daten an die Station zur Verarbeitung weiterreicht, die nicht für sie bestimmt sind.

Verteildienste sind in lokalen Netzen mit Vielfachzugriff einfach realisierbar. Die ATM-Technik sieht jedoch nur Punkt-zu-Punkt-Verbindungen vor, so daß ein Verteildienst mit erheblichem Zusatzaufwand verbunden ist, weil die Daten jeweils für jeden Empfänger kopiert werden müssen. Für ein lokales Netz, in dem Verteildienste ein wesentlicher Bestandteil sind, ist das zentralisierte ATM-Vermittlungsverfahren weniger gut geeignet.

4.2.4 Fairness

Ein Fairness-Mechanismus, der bei lokalen Netzen mit Vielfachzugriff den Zugang der Stationen zum Medium regelt, garantiert allen Stationen eine bestimmte Dienstgüte, insbesondere einen bestimmten Mindestdurchsatz. Ein effizienter Fairness-Mechanismus sollte nicht allen Stationen gleichen Durchsatz zuteilen, sondern die Bandbreite der einzelnen Stationen den Erfordernissen anpassen können. So ist es sinnvoll, einem Server eine höhere Bandbreite zur Verfügung zu stellen als den daran angeschlossenen Terminals.

Die Optimierung eines Fairness-Mechanismus ist ein Abwägen zwischen dem erreichbaren Gesamtdurchsatz des Systems und der Verzögerungszeit sowie deren Schwankung. Bei LANs, die einen separaten Mechanismus für zeitkritischen Verkehr aufweisen, ist die Verzögerungszeit weit weniger kritisch als bei Netzen, die auch den zeitkritischen Verkehr mit demselben Fairness-Verfahren integrieren.

Ein Fairness-Protokoll soll Fairness innerhalb eines Zeitintervalls sicherstellen. Bei höheren Bitraten kann dieses Intervall in einem skalierbaren Protokoll konstant bleiben, lediglich die innerhalb des Intervalls verfügbare Bandbreite einer Station nimmt um den gleichen Faktor wie die Bitrate zu. Damit kann in einer interaktiven Anwendung bei gleicher Antwortzeit mehr Information übertragen werden.

4.2.5 Effizienz

Neben der Optimierung des Fairness-Mechanismus gibt es weitere Möglichkeiten, die Effizienz zu optimieren. Eine direkte Adressierung, bei der eine kurze, innerhalb des lokalen Netzes eindeutige Stationsnummer als Adresse verwendet wird, ermöglicht verbindungslose Dienste, die die Basis für die klassischen LANs bilden. Ein Teil des Adressraums kann für eine solche direkte Adressierung verwendet werden. Einige der in Kapitel 2 beschriebenen Protokollvorschläge verwenden die direkte Stationsadressierung, allerdings in einem zusätzlichen Adressfeld, was zu zusätzlichem Overhead und damit einhergehend einer geringeren Effizienz des Protokolls führt.

Die Übertragung ganzer Pakete ohne Unterbrechung erhöht zwar nicht direkt die Effizienz des Medienzugriffsprotokolls, ermöglicht aber eine vereinfachte Empfänger-Struktur. Paketübertragung bedeutet, daß die Fragmente eines Pakets beim Empfänger nicht durch Fragmente anderer Pakete unterbrochen werden. Ist eine solche Paketübertragung im Netz garantiert, dann wird das Zusammenfügen des Pakets beim Empfänger wesentlich vereinfacht. Nicht für jede Verbindung muß bei Paketübertragung ein Paketpuffer bereitgehalten werden, sondern nur einer insgesamt.

4.2.6 Netzübergänge

Zwei Arten von Netzübergängen sind zu unterscheiden: Der Zugang zu öffentlichen Netzen sowie die Integration in bestehende lokale Netze. Besonders effizient ist ein Netzübergang, wenn er die Konzepte des jeweiligen Netzes weitgehend auf die eigene Protokollstruktur abbilden kann, so daß praktisch keine Konvertierung im Netzübergang erfolgen muß.

Ein einfacher Übergang in ein öffentliches Netz, das auf ATM beruht, setzt voraus, daß das lokale Netz auf Zellen gleicher Größe beruht. Ein öffentliches ATM-Netz ist verbindungsorientiert mit virtuellen Kanälen, die das lokale Netz unterstützen muß. Alle Protokollvorschläge in Kapitel 2 erfüllen diese Vorgaben bereits oder können dahingehend modifiziert werden. Neue Protokolle für Gigabit-LANs müssen sich daran messen.

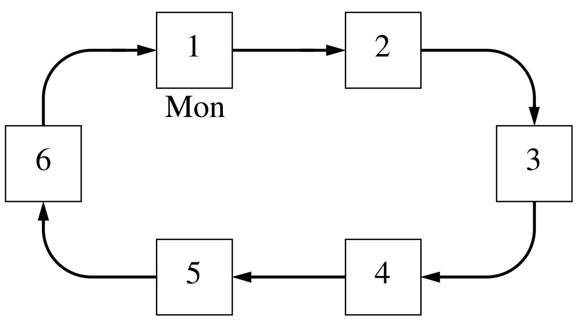
Da neue Netze mit höheren als den bisher verfügbaren Übertragungsraten häufig als Backbone in einem bestehenden Netz eingesetzt werden, muß ein Übergang zu bestehenden lokalen Netzen möglich sein. Versteht man die ATM-Technik als Transportmedium für LAN-Datenpakete ähnlich der LAN-Emulation (siehe Abschnitt 3.2.2), dann unterscheidet sich der Medienzugang für einen Netzübergang nicht von dem einer Station. Der Übergang ist folglich auf eine Kombination herkömmlicher Netzübergänge mit dem Netzzugangsverfahren einer Station reduziert. Der Netzübergang wird in einem auf Zeitschlitzen basierenden LAN wesentlich vereinfacht, wenn das LAN die Übertragung von Paketen in unmittelbar aufeinanderfolgenden Zeitschlitzen ermöglicht, da im Gegensatz zu einer gewöhnlichen Station ein Netzübergang ein Vielfaches an gleichzeitigen Verbindungen unterstützt, für die sonst jeweils ein eigener Empfangspuffer bereitgestellt werden muß. Der Verwaltungsaufwand für derart viele Verbindungen steigt aufgrund der notwendigen Suchvorgänge zur Pufferverwaltung mit der Anzahl der Verbindungen.

4.3 Entwurf der Protokollarchitektur

Dieser Abschnitt enthält die Beschreibung des neuen HSLAN-Protokollvorschlags. Eine Möglichkeit der Realisierung wird in Abschnitt 4.4 vorgestellt, Varianten werden in Abschnitt 6.2 beschrieben. Das Hauptaugenmerk liegt auf der optimalen Integration von isochronen und asynchronen Kanälen, während für das Fairness-Protokoll praktisch jeder auf Zeitschlitzen basierende Fairness-Mechanismus angewendet werden kann. Exemplarisch wird auf ein modifiziertes Orwell-Protokoll nach Abschnitt 2.4.2 zurückgegriffen, weil es Fairness mit vergleichbar geringem Aufwand garantiert.

4.3.1 Topologie

Als Topologie für das Gigabit-LAN wird ein Ring gewählt (Bild 4-2). Der Ring kennt aufgrund seiner Drehsymmetrie keine ausgezeichnete Station, es wird also keine Station aufgrund ihrer Position im Ring bevorzugt oder benachteiligt. Im Gegensatz zur Sterntopologie mit Vermittlungsknoten ist der Ring inhärent broadcast-fähig, da eine Nachricht in der Regel ohnehin über mehrere Stationen weitergereicht wird. Ein Broadcast oder Multicast muß somit nicht kopiert, sondern lediglich bis zur letzten Empfängerstation weitergereicht werden.



1-6: Stationen auf dem Ring
Mon: Monitor- und Masterstation

Bild 4-2: Ring-Topologie für das Gigabit-LAN

In einem einfachen Ring wird für ein zu sendendes Paket keine Routing-Entscheidung benötigt, da alle Stationen die Zieladresse des Pakets mit ihrer eigenen Adresse vergleichen. Damit ist der Aufwand für die verteilte Vermittlungsfunktion im Ring minimal und beschränkt sich auf drei Funktionen:

- Kopieren der Daten, die an die Station adressiert sind, vom Ring in den Empfangspuffer
- Multiplexen des Transit-Datenstroms und des Sendedatenstroms zum Ausgangsdatenstrom
- Entfernen der Daten vom Ring, die nicht mehr benötigt werden

Für den Fehlerfall, insbesondere beim Versagen der letztgenannten Funktion, wird eine einfache Monitor-Funktionalität im Ring benötigt. Diese Funktion übernimmt eine Monitor- und Masterstation neben weiteren Überwachungs- und Initialisierungsfunktionen. Wird der Ring in logischer Sternform aufgebaut (siehe Abschnitt 6.2.2.3), können in einem zentralen Schaltelement einzelne Stationen elektrisch oder optisch überbrückt werden. Ein aktives zentrales Schaltelement kann auch die Monitor- und Masterfunktionalität erbringen.

4.3.2 Rahmenstruktur

Die Struktur eines Rahmens ist in Bild 4-3 dargestellt. Der Header umfaßt vier Oktetts. Ihm folgen N Zeitschlitze mit jeweils 56 Oktetts. Den Abschluß des Rahmens bildet ein Padding-Feld mit Füllworten ohne Bedeutung. Die Anzahl der Füllworte wird so gewählt, daß die Rahmendauer genau 125 µs beträgt. Die Anzahl der Zeitschlitze N hängt ab von der Übertragungsrate auf dem Medium.

Header	Zeitschlitz 1	Zeitschlitz 2	···	Zeitschlitz N	PAD	
4	56	56		56	p	Oktetts

Rsv	Rdy	Run	Frame#	FOR	
10	1	1	12	8	Bits

Header	Rahmen-Header
PAD	Padding, um eine Rahmendauer von 125 µs zu erreichen
N	Anzahl der Zeitschlitze je Rahmen; implementierungsabhängig
Rsv	Reserviert
Rdy	Ready, alle Stationen sind bereit
Run	Running, Rahmen kann mit Daten belegt werden
Frame#	Rahmennummer, fortlaufender Modulo-2^{12}-Zähler
FOR	Frames On Ring = Anzahl der Rahmen auf dem Ring

Bild 4-3: Rahmenformat

Der Header enthält die Anzahl der Rahmen auf dem Ring (FOR) sowie eine modulo 2^{12} fortlaufende Rahmennummer (Frame#), aus der die Monitor- und Masterstation die Größe FOR ermittelt. Solange sich das System in der Initialisierungs- oder einer Fehlerbehebungsphase befindet, wird ein „Run“-Bit gesetzt, das den Stationen signalisiert, daß keine Zeitschlitze belegt werden können.

Die Monitor- und Masterstation setzt in allen Rahmen das „Rdy“-Bit, um festzustellen, ob alle Stationen auf dem Ring bereit sind. Eine Station, die noch nicht bereit ist, setzt dieses Bit zurück, alle anderen Stationen lassen es unverändert. Sobald ein Rahmen mit gesetztem „Rdy“-Bit bei der Monitor- und Masterstation ankommt, sind alle Stationen bereit und die Initialisierungsphase kann abgeschlossen werden.

4.3.3 Zeitschlitz- und Zellstruktur

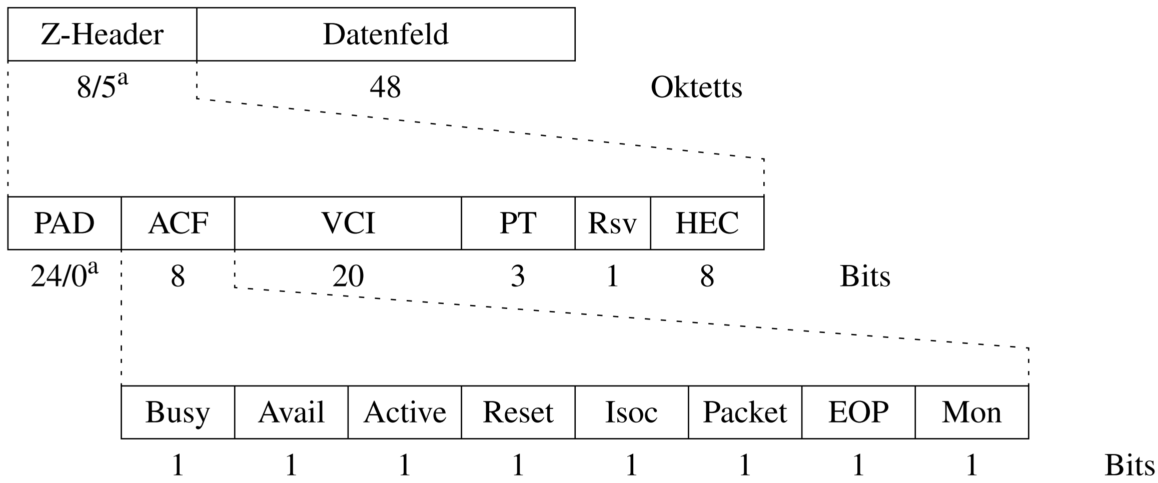
Die Rahmen sind in Zeitschlitze konstanter Länge eingeteilt, die wiederum Zellen enthalten können. Zeitschlitze strukturieren den Rahmen, Zellen sind die Träger von Nutzlast. Ein Zeitschlitz kann leer sein oder eine Zelle enthalten, während eine Zelle immer Nutzdaten enthält. Zeitschlitz und Zelle sind nicht strikt voneinander trennbar, da sie sich Bits im Zeitschlitz-/Zell-Header teilen. Die für eine Zelle spezifischen Bits sind in einem leeren Zeitschlitz ohne Bedeutung.

Bild 4-4 zeigt die Struktur eines mit einer Zelle belegten Zeitschlitzes. Die Länge eines Zeitschlitzes beträgt 56 Oktetts, wobei 3 Oktetts davon lediglich zum Erreichen einer 32-Bit Wortgrenze dienen und sonst keine Bedeutung haben. Die eigentliche Zellgröße ist 53 Oktetts. Falls die zugrundeliegende Übertragungsschicht es erfordert, kann die Zeitschlitzgröße auf Kosten der dann nur noch mit größerem Aufwand realisierbaren 32-bit-Verarbeitung auch direkt 53 Oktetts betragen.

Eine Zelle setzt sich zusammen aus einem Header und einem 48 Oktetts langen Datenfeld. Der Header besteht aus einem optionalen PAD-Feld, dem Access Control Field (ACF), einem Virtual Channel Identifier (VCI), dem Payload Type (PT) sowie einem reservierten Bit und einer 8 bit großen Prüfsequenz, die sich über den Header ohne PAD erstreckt. Mit diesem an ATM orientierten Zellformat ist ein einfacher Übergang zu ATM-Netzen möglich.

Das ACF enthält acht Bits, wovon die ersten vier dem Zeitschlitz und die letzten vier der Zelle zugeordnet sind. In einem leeren Zeitschlitz sind somit nur die ersten vier ACF-Bits von Bedeutung. Das Busy-Bit zeigt an, ob der Zeitschlitz belegt ist. Das Available Bit wird im Zusammenhang mit der Integration von asynchronen und isochronen Kanälen benötigt (siehe Abschnitt 4.3.5). Für den Fairness-Mechanismus werden die beiden Bits „Active“ und „Reset“ benötigt (siehe Abschnitt 4.3.6). Das Bit „Isochronous Cell“ kennzeichnet die Zugehörigkeit einer Zelle zur isochronen oder asynchronen Verkehrsklasse. „Packet“ und „EOP“ ermöglichen die Übertragung ganzer Pakete ohne Unterbrechung (siehe Abschnitt 4.3.7). Das Monitor-Bit wird von der Monitor- und Masterstation genutzt, um Zellen vom Ring zu entfernen, die bereits ein zweites Mal bei ihr ankommen, weil sie nicht von der zuständigen Station entfernt wurden („Zombie“-Zellen). Hierzu wird beim Senden einer Zelle das Monitor-Bit gelöscht, während die Monitor- und Masterstation es in jeder passierenden Zelle setzt. Wird die Zelle während des folgenden Ringumlaufs nicht entfernt, kommt sie mit gesetztem Monitor-Bit bei der Monitor- und Masterstation an und wird entfernt.

Das VCI-Feld umfaßt 20 bits und kann als Adresse aufgefaßt werden. Da das Kopieren und das Entfernen von Zellen zwei getrennte Aufgaben einer Station sind, hat jede Station einen



Z-Header	Zeitschlitz-/Zell-Header
PAD	Padding, Füllbits ohne Funktion
ACF	Access Control Field
VCI	Virtual Channel Identifier; Adressfeld
PT	Payload Type
Rsv	Reserved
HEC	Header Error Check
Avail	Available Bit
Active	Active Node Bit
Isoc	Isochronous Cell
EOP	End Of Packet
Mon	Monitor Bit

a. Die Größe des Zeitschlitz-Headers kann zur einfacheren 32-bit-Verarbeitung des Zeitschlitzes auf 8 Oktetts verlängert werden

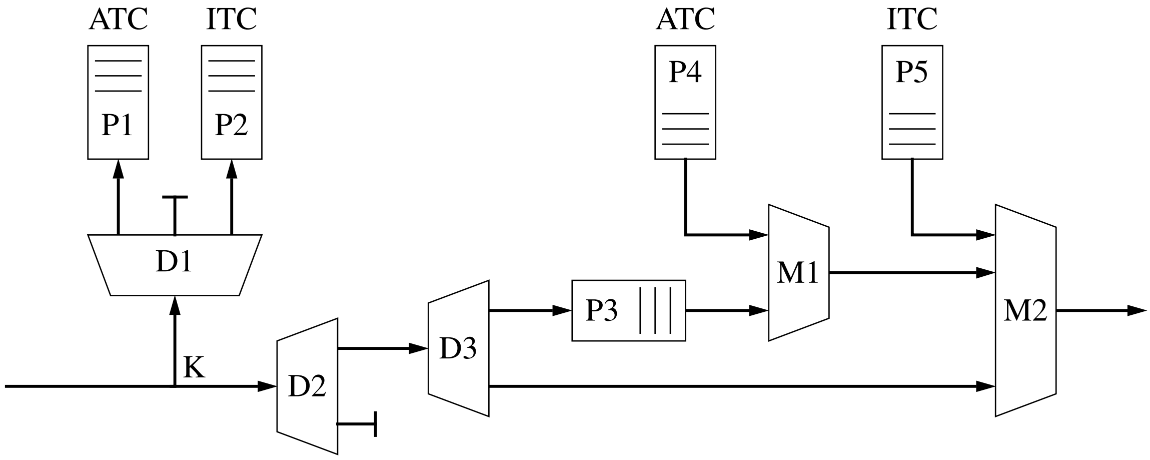
Bild 4-4: Zeitschlitz- und Zellformat

Assoziativspeicher mit je einer Liste für die VCI-Werte, unter denen sie als Ziel angesprochen ist, sowie einen für die Werte, bei denen sie für das Entfernen der Zelle zuständig ist. Abschnitt 4.3.8 enthält einen Vorschlag für die Nutzung des VCI-Raumes von 2^{20} Adressen im lokalen Netz.

Das PT-Feld wird zur Signalisierung entsprechend ITU-T Recommendation I.361 verwendet [48] und dient unter anderem zur Kennzeichnung der letzten Zelle einer AAL-5-CPCS-PDU (siehe Abschnitt 3.1.3). Die Berechnung und Auswertung der Prüfbits im HEC-Feld erfolgt gemäß ITU-T Recommendation I.432 [52] und sichert die vier relevanten Oktetts im Header.

4.3.4 Stationsaufbau

Der Aufbau der Datenpfade einer Station ist in Bild 4-5 dargestellt. Eine ankommende Zelle wird, sofern die Station Empfänger der Zelle ist, in den ATC-Empfangspuffer P1 oder, bei gesetztem Isoc-Bit, in den ITC-Empfangspuffer P2 kopiert. Der Demultiplexer D2 entfernt



K	Kopierfunktion
D1	Demultiplexer, Auswahl eines Empfangspuffers oder Verwerfen der kopierten Zelle
P1	Empfangspuffer für ATC-Zellen
P2	Empfangspuffer für ITC-Zellen
D2	Demultiplexer zum Entfernen von Zellen aus dem Ring
D3	Demultiplexer, leitet ATC-Zellen in den Einfügepuffer P3, ITC-Zellen und die Rahmenstruktur nach M2
P3	Einfügepuffer für ATC-Zellen
P4	Sendepuffer für ATC-Zellen
M1	Multiplexer für ATC-Verkehr
P5	Sendepuffer für ITC-Zellen
M2	Ausgangsmultiplexer

Bild 4-5: Datenpfade einer Station

Zellen vom Ring, wenn die Station der letzte Empfänger dieser Zelle war oder als Monitor-Station eine „Zombie"-Zelle erkannt hat.

Der Demultiplexer D3 leitet alle ATC-Zellen in den Einfügepuffer P3 um, während das Rahmengerüst mit den Zeitschlitzen sowie alle ITC-Zellen direkt an den Ausgangsmultiplexer M2 weitergeleitet werden. Somit bleibt die Rahmenstruktur am Ausgang erhalten. ITC-Zellen erfahren die geringstmögliche, konstante Verzögerung durch die Station.

Mit dem Multiplexer M1 wird beim Senden einer ATC-Zelle ausgewählt, ob diese aus dem Einfügepuffer P3 oder dem ATC-Sendepuffer P4 der Station entnommen werden soll. Die Strategie, immer zuerst Zellen aus P3 und nur dann aus P4 zu entnehmen, wenn P3 leer ist, führt zum niedrigst möglichen mittleren Füllstand von P3 und somit zur geringsten Verzögerung für ATC-Zellen durch die Station.

Der Ausgangsmultiplexer M2 fügt gemäß dem im folgenden Abschnitt beschriebenen Medienzugriffsprotokoll und dem Fairness-Mechanismus (Abschnitt 4.3.6) in den von D3 kommenden und den Rahmen enthaltenden Transitdatenstrom Zellen aus dem ITC-Sendepuffer P5 oder ATC-Zellen aus dem Multiplexer M1 ein.

4.3.5 Medienzugriffsprotokoll

Für den Zugriff auf die Zeitschlitze werden zwei Verkehrsarten unterschieden: In der isochronen Verkehrsklasse (ITC) werden Verbindungen aufgebaut und Zeitschlitze explizit zugeteilt, während in der asynchronen Verkehrsklasse (ATC) freie Zeitschlitze belegt werden können. Der ITC-Verkehr hat absoluten Vorrang und erfährt eine konstante Verzögerung vom Sender zum Empfänger, während für ATC-Verkehr alle nicht reservierten sowie reservierte, aber momentan ungenutzte Zeitschlitze zur Verfügung stehen.

4.3.5.1 ITC-Verkehr

Um Zeitschlitze für ITC-Verbindungen reservieren zu können, sind sie durch eine eindeutige Nummer gekennzeichnet, wobei die Nummer der Position im Rahmen entspricht. Ist ein Zeitschlitz für eine ITC-Verbindung einer Station reserviert, gilt diese Reservierung in jedem Rahmen so lange, bis die Verbindung wieder abgebaut wird. Die Monitor- und Masterstation vergibt, verwaltet und überwacht die Reservierungen.

Die Anzahl der zu reservierenden Zeitschlitze für eine Verbindung ergibt sich aus der Spitzenbitrate B_{max}, die innerhalb der Verbindung auftreten kann. Bei 48 Oktetts Nutzdaten je Zelle und einer Rahmendauer von 125 µs ergibt sich eine Bitrate je reservierter Zelle und damit eine Granularität der reservierbaren Bandbreite von B_{ITC} = 3072 kbit/s. Die Anzahl der zu reservierenden isochronen Kanäle für eine ITC-Verbindung ist der zur nächstgrößeren ganzen Zahl aufgerundete Quotient B_{max}/B_{ITC}.

Beim Verbindungsaufbau sendet die initiierende Station eine Anforderung an die Monitor- und Masterstation, wobei die benötigte Spitzenbitrate einen Teil der auszuhandelnden Verbindungsparameter darstellt. Die Monitor- und Masterstation bestätigt, sofern ausreichend unreservierte Zeitschlitze zur Verfügung stehen, die Reservierung und teilt der initiierenden Station die Nummern der Zeitschlitze mit, die sie für die Verbindung reserviert hat. Anschließend kann die initiierende Station über die für sie reservierten Zeitschlitze verfügen, bis die Verbindung wieder abgebaut wird. Insbesondere kann die Station jeden für sie reservierten Zeitschlitz unabhängig von dessen Belegungszustand mit einer ITC-Zelle aus ihrem Sendepuffer belegen oder ihn gemäß Abschnitt 4.3.5.2 für ATC-Verkehr freigeben.

4.3.5.2 Freigabe reservierter Zeitschlitze für ATC-Verkehr

Für den ATC-Verkehr stehen zunächst alle nichtreservierten Zeitschlitze zur Verfügung. Hierzu markiert die Monitor- und Masterstation alle nichtreservierten Zeitschlitze als „Available“. Auch eine Station, die Zeitschlitze für ITC-Verkehr reserviert hat, kann „ihre“ Zeitschlitze für ATC-Verkehr freigeben, indem sie das „Available“-Bit in ihren Zeitschlitzen setzt und damit signalisiert, daß der jeweilige Zeitschlitz von einer beliebigen Station mit einer ATC-Zelle belegt werden kann. Werden möglichst viele reservierte Zeitschlitze auf diese Art freigegeben, erhöht sich in einem System mit hohem Anteil reservierter Kanäle der Gesamtdurchsatz entsprechend, insbesondere wenn die ITC-Kanäle mit stark büschelhaftem Verkehr belegt sind.

Gibt eine Station einen ihrer reservierten Zeitschlitze für Asynchronverkehr frei, kann der Zeitschlitz bei der Rückkehr nach einem Ringumlauf mit einer ATC-Zelle belegt sein. Damit die Station ihren reservierten Zeitschlitz für eine ITC-Zelle nutzen kann, wird die ankommende Zelle in den Einfügepuffer (P3 in Bild 4-5) geleitet und der Zeitschlitz auf diese Weise geleert. Damit die Zellreihenfolge innerhalb der asynchronen Verkehrsklasse erhalten bleibt, durchlaufen alle von der Station weitergeleiteten ATC-Zellen den Einfügepuffer.

Da der Einfügepuffer nur eine endliche Kapazität besitzt, müssen Maßnahmen gegen ein Überlaufen im Falle hoher Last getroffen werden. Aus diesem Grund kann eine Station gerade so viele ihrer reservierten Zeitschlitze als „Available" markieren, wie der Einfügepuffer fassen kann. Die Summe aus Pufferfüllstand und Anzahl der als „Available" markierten Zeitschlitze der Station auf dem Ring darf das Fassungsvermögen des Puffers nicht überschreiten. Hierzu wird ein „Available Counter" eingeführt, dessen Zählerstand dieser Summe entspricht.

Der Anfangswert des Available Counters ist 0. Der Available Counter wird inkrementiert, wenn ein Zeitschlitz, der für die Station reserviert ist, als „Available" markiert die Station verläßt. Er wird dekrementiert, wenn ein derartiger Zeitschlitz bei der Station ankommt. Beim Einschreiben einer Zelle in den Einfügepuffer inkrementiert, beim Entnehmen einer Zelle aus dem Einfügepuffer dekrementiert die Station ihren Available Counter. Die Station kann jetzt so lange alle ihre Zeitschlitze als „Available" markieren, wie der Stand des Available Counters kleiner oder gleich dem Fassungsvermögen des Einfügepuffers ist.

Beim Aufbau einer neuen ITC-Verbindung wird zum Available Counter die Anzahl der zusätzlich für die Station reservierten Zeitschlitze addiert, da diese bis zu diesem Zeitpunkt von der Monitor- und Masterstation verwaltet und als „Available" markiert wurden. Diese Zahl entspricht dem Produkt aus Anzahl der reservierten Zeitschlitze je Rahmen und der Anzahl der Rahmen auf dem Ring. Übersteigt durch diese Aktion der Available Counter seine Grenze, werden für die Station reservierte Zeitschlitze als „Unavailable" markiert, bis die Grenze wieder unterschritten wird. Bis zu diesem Zeitpunkt darf die Station die neue Verbindung nicht nutzen. Dadurch wird sichergestellt, daß der Einfügepuffer nicht überläuft. Entsprechend gibt die Station beim Verbindungsabbau die bisher reservierten Zeitschlitze als „Available" frei und subtrahiert vom Available Counter die Anzahl der zurückgegebenen Zeitschlitze auf dem Ring.

Eine Station kann auch die von ihr mit einer ITC-Zelle belegten Zeitschlitze als „Available" markieren. Diese Zeitschlitze können dann nach der Freigabe durch den Empfänger der Zelle für ATC-Verkehr genutzt werden.

4.3.6 Fairness-Mechanismus

Zur Sicherstellung von Fairness, also der Verteilung der Bandbreite entsprechend einer Vorgabe, wird ein vom Orwell-Verfahren (siehe Abschnitt 2.4.2) abgeleiteter Credit-Mechanismus verwendet. Prinzipiell kann jeder Fairness-Mechanismus eingesetzt werden, sofern er mit der durch die Restbandbreitennutzung entstehenden variablen Bitrate zurechtkommt, also nicht von einer konstanten Bitrate auf jedem physikalischen Kanal ausgeht. Das Credit-Verfahren ist einfach zu realisieren und bietet zudem den Vorteil, daß es durch den zyklischen Neustart äußerst robust und fehlertolerant ist.

Zu Beginn eines Zyklus wird der Credit Counter jeder Station auf seinen Anfangswert gesetzt. Der Anfangswert entspricht der Anzahl von Zeitschlitzen, die die Station im folgenden Zyklus für ihren ATC-Verkehr belegen darf. Wenn die Station einen Zeitschlitz mit einer Zelle aus ihrem ATC-Sendepuffer (P4 in Bild 4-5) belegt, wird der Credit Counter um 1 vermindert. Sobald der Credit Counter den Wert Null erreicht, darf die Station nur noch Zellen aus dem Einfügepuffer (P4 in Bild 4-5) senden, nicht jedoch aus ihrem ATC-Sendepuffer.

Die Monitor- und Masterstation überprüft den Zustand der Stationen und startet einen neuen Zyklus, wenn keine Station mehr sendet. Hierzu werden die Bits „Active“ und „Reset“ im ACF-Feld des Zeitschlitz-Headers verwendet. „Active“ wird von der Monitor- und Masterstation in jedem Zeitschlitz zurückgesetzt. Jede Station, die gemäß Tabelle 4-2 aktiv ist, setzt das

Einfügepuffer	Credit Counter	ATC-Sende-puffer	Stationszustand
nicht leer	beliebig	beliebig	aktiv
leer	> 0	nicht leer	aktiv
leer	=0	beliebig	inaktiv
leer	beliebig	leer	inaktiv

Tabelle 4-2: Stationszustand im Sinne des Fairness-Mechanismus

„Active“-Bit in jedem Zeitschlitz. Solange noch eine Station aktiv ist, kommen nur Zeitschlitze mit gesetztem „Active“-Bit bei der Monitor- und Masterstation an. Ist keine Station mehr aktiv, erreicht die Monitor- und Masterstation ein Zeitschlitz ohne gesetztes „Active“-Bit, woraufhin diese einen Reset-Zyklus startet. Hierzu setzt sie das Reset-Bit in einem Zeitschlitz und startet gleichzeitig einen Zeitgeber, um die Umlaufzeit des Reset zu überwachen. Der Zeitgeber wird so eingestellt, daß er nach etwas mehr als einer Ringumlaufzeit abläuft. Die Ringumlaufzeit beträgt 125 µs multipliziert mit der Anzahl der Rahmen auf dem Ring.

Die Stationen interpretieren das gesetzte Reset-Bit als Start eines neuen Zyklus und setzen ihren Credit Counter auf den Anfangswert. Nach einem Ringumlauf setzt die Monitor- und Masterstation das Reset-Bit sowie ihren Zeitgeber zurück. Läuft der Zeitgeber vor der Rückkehr des Reset-Bits ab, ist das Reset-Bit aufgrund eines Fehlers verloren gegangen. In diesem Fall startet die Monitor- und Masterstation den Reset-Zyklus erneut. Die Monitor- und Masterstation setzt ein eventuell fälschlicherweise gesetztes Reset-Bit zurück und beendet somit einen fehlerhaften Reset-Zyklus. Die Fairness ist bei einem verlorenen oder zusätzlich generierten Reset für die Dauer eines Zyklus nicht gewährleistet.

4.3.7 Paketübertragung

In einem lokalen Netz werden Datenpakete übertragen, deren Größe 48 Oktetts übersteigen kann [97]. Der ATM Adaptation Layer 5 (AAL-5), der für die Übertragung in einem zellbasierten LAN vorwiegend verwendet wird, definiert Paketlängen bis zu 65535 Oktetts [50]. Eine Station, die auf mehreren Verbindungen gleichzeitig empfängt, muß für jede dieser Verbindungen einen Empfangspuffer bereithalten, da der Zellstrom einer Verbindung durch Zellen anderer Verbindungen unterbrochen werden kann.

Der Aufwand kann auf einen einzigen Paketpuffer reduziert werden, wenn sichergestellt ist, daß der Zellstrom eines Pakets nicht durch Zellen einer anderen Verbindung unterbrochen werden kann. Eine unterbrechungsfreie Paketübertragung wird mit dem im folgenden beschriebenen Mechanismus sichergestellt. Eine Unterbrechung des ATC-Zellstroms durch ITC-Zellen ist möglich, weil sie durch die getrennten Empfangspuffer für ITC- und ATC-Verkehr keine negativen Auswirkungen hat.

Eine Station, die ein Paket aussenden will, muß in ihrem Einfügepuffer zusätzlichen Speicher bereitstellen, der so viele Zellen faßt, wie bei der Segmentierung des längsten Pakets entstehen[1]. Dies ist notwendig, um während der Paketübertragung eintreffende ATC-Zellen puffern zu können.

Eine Station beginnt die Paketübertragung mit dem Senden der ersten Zelle gemäß den oben beschriebenen Protokollvorschriften. Voraussetzung für die Paketübertragung ist, daß die Anzahl der zum Paket gehörenden Zellen kleiner oder gleich der Anzahl der Credits ist und der für die Paketübertragung reservierte Platz im Einfügepuffer unbelegt ist. Alle zum Paket gehörenden Zellen werden mit einem gesetzten „Packet“-Bit im ACF-Feld des Zell-Headers gekennzeichnet. Nach dem Senden der ersten Zelle haben alle nachfolgenden zu dem Paket gehörenden Zellen Vorrang gegenüber den Zellen im Einfügepuffer. Die letzte Zelle des Pakets wird gekennzeichnet durch ein zusätzlich gesetztes „EOP“-Bit im ACF-Feld des Zell-Headers.

Die maximale Paketgröße für eine Übertragung ohne Unterbrechung hängt ab von der bereitgestellten zusätzlichen Kapazität des Einfügepuffers und dem Anfangswert des Credit Counters. Soll eine maximale Paketgröße von 65535 Oktetts übertragen werden, entspricht dies knapp 1400 Zellen. Der Einfügepuffer muß folglich zusätzlich 1400 Zellen fassen und der Anfangswert des Credit Counters mindestens 1400 betragen.

4.3.8 Adressierung und Verbindungssteuerung

In einem lokalen Netz wie Ethernet erfolgt die Adressierung der Stationen auf der MAC-Schicht über eine eindeutige, dem Netzzugang fest zugeordnete Adresse. Der Header eines Pakets enthält die Adressen von Sender und Empfänger mit einer Länge von jeweils 6 Oktetts [34]. Eine ATM-Zelle hingegen hat keine Zieladresse, sondern verfügt über ein VCI-Feld von 16 bit sowie ein VPI-Feld von 8 bit [48]. Für das Ring-Protokoll wurde in Abschnitt 4.3.3

1. Abzüglich einer Zelle, da die Paketübertragung in einem leeren Zeitschlitz beginnt

lediglich ein VCI-Feld mit einer Größe von 20 bit definiert, das als Adresse aufgefaßt werden kann.

Die Nutzung des VCI-Feldes kann gemäß den Spezifikationen für ein ATM-LAN erfolgen, wie in Kapitel 3 beschrieben. Dabei ist lediglich der VPI/VCI-Raum durch die Einführung des ACF von 24 auf 20 bit eingeschränkt, durch die begrenzte Anzahl von Stationen in einem lokalen Netz jedoch ausreichend.

Eine Alternative hierzu ist eine gemischte Nutzung des VCI-Feldes, die sowohl Verbindungen mit virtueller Adresse, als auch direkte Adressierung zuläßt. Das 20 Bits umfassende VCI-Feld wird unterteilt in ein 4 bit und ein 16 bit großes Feld, wobei die 4 Bits als VCI-Typ des 16 bit-Feldes verstanden werden sollen. Einer der 2^4 möglichen VCI-Typen wird reserviert für direkte LAN-Adressierung innerhalb der asynchronen Verkehrsklasse, wobei die 16 verbleibenden Bits gemäß Bild 4-6 als Quell- und Zieladresse innerhalb der MAC-Schicht interpretiert werden. Diese 8 bit-Adressen werden in den Stationen auf 48 bit-Adressen umgesetzt, so daß Adressierung und Übertragung von Paketen für herkömmliche ATM-Anwendung transparent erfolgt. Die Adressierung innerhalb isochroner Verbindungen, die ohnehin eines Verbindungsaufbaus zur Resevierung der Ressourcen bedürfen, sowie auf ATM basierender Anwendungen erfolgt wie bei ATM gemäß Kapitel 3.

VCI-Typ	Zieladresse	Quelladresse	
4	8	8	Bits

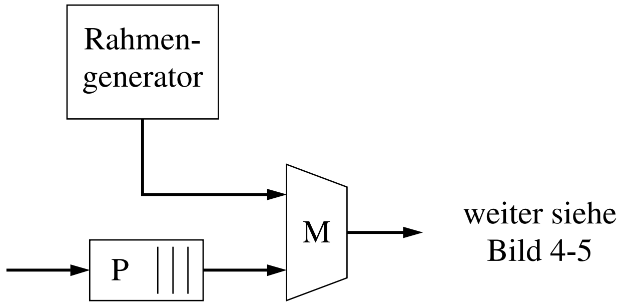
Bild 4-6: Aufteilung des VCI für direkte Adressierung

Bei Nutzung der 8 bit-Adressen ergibt sich innerhalb des Rings ein Raum von 2^8 Adressen. Damit können 255 Stationen adressiert werden, wenn eine Adresse als Broadcast-Adresse reserviert wird. Die Stationsadressen können statisch durch manuelle Konfiguration oder dynamisch bei der Initialisierung des Rings vergeben werden. Beim Senden trägt eine Station als Zieladresse die Adresse der Zielstation oder bei einem Broadcast die Broadcast-Adresse ein, während sie als Quelladresse ihre eigene Adresse einträgt. Die Zielstation kopiert die Zelle in den Empfangspuffer und markiert den Zeitschlitz als unbelegt. Bei einem Broadcast entfernt die Station die Zelle, die sie gesendet hat.

4.3.9 Monitor- und Master-Funktionen

Die Monitor- und Masterstation übernimmt alle Funktionen, die nur einmal im Ring benötigt werden. Die Funktionalität dieser Station ist in allen Stationen vorhanden, wobei entweder während der Initialisierungsphase oder durch manuelle Konfiguration die Monitor- und Masterstation bestimmt wird. Alternativ kann die Funktionalität von einer zentralen Instanz erbracht werden. Die Ermittlung der Monitor- und Masterstation entfällt in diesem Fall, jedoch sind die Stationen nicht mehr einheitlich aufgebaut.

Bild 4-7 zeigt die zusätzlichen Elemente im Datenpfad der Monitor- und Masterstation. Mit dem Rahmengenerator wird zu Anfang oder nach einem Fehler der Rahmen erzeugt. Der Elastizitätspuffer P wird benötigt, um den Anfang eines ankommenden Rahmens genau dann über den Multiplexer M weiterzuleiten, wenn das Ende des zuletzt generierten Rahmens erreicht ist, um die Ringlänge auf ein ganzzahliges Vielfaches der Rahmenlänge zu erweitern. Damit wird der Ring geschlossen und beinhaltet einen oder mehrere vollständige Rahmen.



P Elastizitätspuffer für Ringschluß
M Multiplexer für Auswahl der Rahmenstruktur

Bild 4-7: Zusätzliche Datenpfade der Monitor- und Masterstation

4.3.9.1 Automatisches Bestimmen der Monitor- und Masterstation

Eine einfache Möglichkeit zur automatischen Ermittlung der Monitor- und Masterstation setzt die Vergabe einer eindeutigen, fortlaufenden Identifikationsnummer für jede Netzschnittstelle voraus. Nach dem Einschalten verhält sich jede Station wie die Monitor- und Masterstation und generiert Rahmen. Darin sendet sie periodisch eine Anforderung mit ihrer Identifikationsnummer in einer Management-Zelle. Eine Station, die eine derartige Anforderung erhält, vergleicht die Identifikationsnummer in der Anforderung mit ihrer eigenen. Ist die eigene Nummer größer, wird die angekommene Datenstruktur ignoriert. Ist sie kleiner, so gibt es eine Station mit größerer Identifikationsnummer. In diesem Fall wird die angekommene Datenstruktur unverändert weitergeleitet und die Station verhält sich ab diesem Zeitpunkt wie eine gewöhnliche Station; Rahmen werden nicht mehr generiert, sondern ankommende Rahmen werden weitergeleitet. Bei Gleichheit der Anforderungsnummern hat die eigene Anforderung unverändert den Ring durchlaufen, woraufhin die Station als neue Monitor- und Masterstation den Konfigurationsvorgang beendet. Auf diese Weise wird die Station mit der höchsten Nummer, also die jüngste Station mit der neuesten Betriebssoft- und Hardware, zur Monitor- und Masterstation bestimmt.

4.3.9.2 Initialisierung des Rings

Die Monitor- und Masterstation initialisiert den Ring. Alle Stationen generieren nach dem Einschalten eine Rahmenstruktur, wobei gemäß Abschnitt 4.3.9.1 nur die Rahmenstruktur der Monitor- und Masterstation von allen anderen Stationen weitergeleitet wird. Wenn alle Stationen im Ring eingeschaltet sind, kommt der erste von der Monitor- und Masterstation generierte Rahmen wieder bei ihr an. Um den Ring zu schließen, kann dieser Rahmen erst dann wieder ausgesendet werden, wenn der momentan zu diesem Zeitpunkt generierte Rahmen beendet ist. Aus der Differenz zwischen der fortlaufenden Nummer eines Rahmens beim Senden und beim Empfangen bestimmt die Monitor- und Masterstation die Anzahl der Rahmen auf dem Ring und trägt diese Zahl in das FOR-Feld des Rahmen-Headers ein.

Zur automatischen Konfiguration ihrer Stationsadresse wählt eine Station eine beliebige Adresse aus und sendet eine Anforderung mit dieser Stationsadresse auf dem Ring. Die weiteren Stationen auf dem Ring vergleichen diese Adresse mit ihrer eigenen Stationsadresse und vermerken in der Anforderung eine eventuelle Kollision der Adressen. Wenn die Anforderung bei der initiierenden Station wieder ankommt, wird die gewählte Adresse zur Stationsadresse, falls keine Kollision aufgetreten ist. Ansonsten wird der Vorgang mit einer anderen Adresse wiederholt [18].

4.3.9.3 Verbindungs-Management

Die Monitor- und Masterstation ist zuständig für die Vergabe isochroner Kanäle und die damit verbundene Reservierung von Zeitschlitzen im Rahmen sowie für die Vergabe von virtuellen Kanälen. Für die Reservierungsanforderung sind spezielle VCI-Werte reserviert, über die eine Station mit Verbindungsaufbauwunsch mit der Monitor- und Masterstation kommuniziert. Die Signalisierung zum Verbindungsauf- und Abbau erfolgt nach ITU-T Standard Q.2931 [53], wobei zusätzliche Funktionen für die Reservierung und Bekanntgabe der isochronen Kanäle benötigt werden.

Die Monitor- und Masterstation führt eine Liste der vergebenen Zeitschlitze. Fordert eine Station eine isochrone Verbindung an, werden so viele freie Zeitschlitze in dieser Liste gesucht, wie für die Verbindung benötigt werden. Die Zeitschlitze sollen einen möglichst großen Abstand voneinander haben, um die Zugriffszeit auf das Medium zu minimieren. Die gewählten Zeitschlitze werden registriert und der anfordernden Station zusammen mit dem für diese Verbindung relevanten VCI-Wert mitgeteilt.

Der Verbindungsabbau erfolgt entweder explizit durch die sendende Station oder implizit durch Zeitüberwachung. Die Monitor- und Masterstation fragt hierzu die Stationen zyklisch ab, ob die registrierten Verbindungen noch bestehen. Nicht mehr bestehende Verbindungen werden aus der Registrierung entfernt und die Reservierung der Zeitschlitze dadurch aufgehoben.

4.3.9.4 Fehlerbehebung

Eine weitere Aufgabe der Monitor- und Masterstation ist die Beseitigung von Fehlern aufgrund fehlerhafter Zellen oder Rahmen. Erkennt die Monitor- und Masterstation einen fehlerhaften Rahmen, so wird der Rest des Rahmens verworfen und der Rahmen ab der fehlerhaften Stelle durch einen neu generierten Rahmenteil ersetzt. Fehlende Rahmen, die zum Beispiel von einer Unterbrechung des Rings herrühren, werden ebenfalls durch neu generierte Rahmen ersetzt. Eine nach der Unterbrechung geänderte Ringlänge, die aus einer Umkonfiguration des Rings resultieren kann, wird von der Monitor- und Masterstation erkannt. Eine möglicherweise geänderte Anzahl der Rahmen auf dem Ring wird den Stationen über das FOR-Feld im Rahmen-Header mitgeteilt.

Zellen, die nach einem Ringumlauf keine Station vom Ring genommen hat, werden von der Monitor- und Masterstation unter Nutzung des Monitor-Bits im Zell-Header nach dem in Abschnitt 4.3.3 beschriebenen Verfahren erkannt und entfernt. Gleichzeitig wird das Ereignis protokolliert, um einen möglicherweise defekten Empfänger identifizieren zu können.

Bei einem aufgrund eines Fehlers neu generierten Rahmen hat die Monitor- und Masterstation keine Kenntnis über den Zustand des Available-Bits vor dem Fehler. Sie setzt daher das Available-Bit aller nicht reservierten Zeitschlitze auf „available“, das der reservierten Zeitschlitze hingegen auf „nicht available“, um zu verhindern, daß der Einfügepuffer einer Station überläuft. Dies führt möglicherweise dazu, daß der Available Counter einer Station einen zu großen Wert aufweist, da die Station von einer größeren Anzahl als „available“ markierter Zeitschlitze ausgeht.

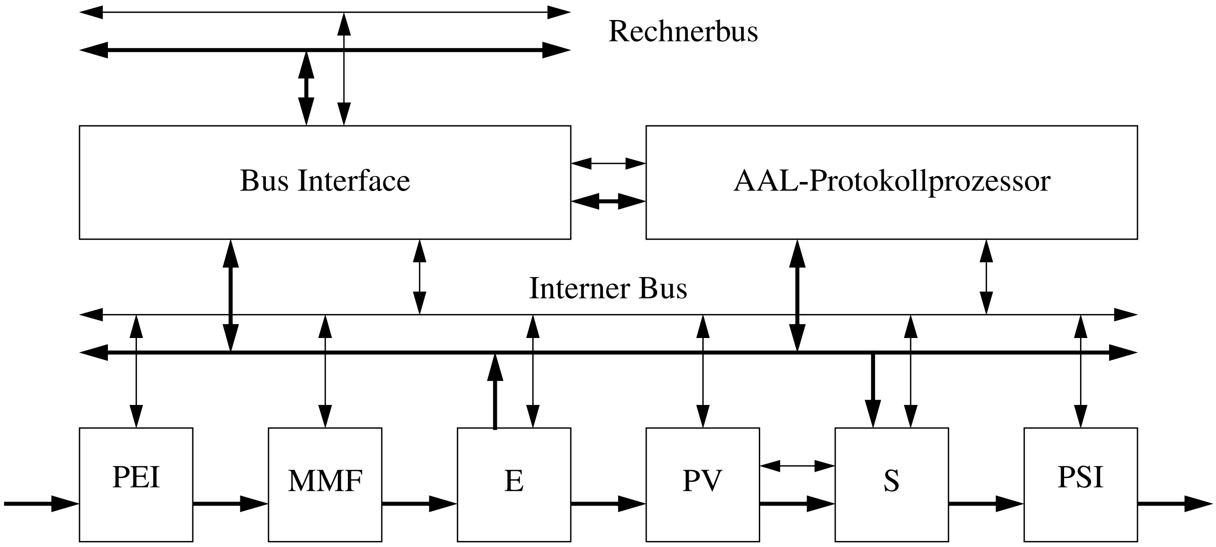
Eine mögliche Gegenmaßnahme ist ein zusätzlicher Speicher in jeder Station, in dem für jeden reservierten Zeitschlitz auf dem Ring der Zustand des Available-Bits vermerkt wird. Relevant für die Änderung des Available Counters ist dann nicht mehr das ankommende Available-Bit im Zell-Header, sondern das zugehörige Bit im Speicher.

Alternativ kann die zyklische Funktion des Credit-Mechanismus ausgenutzt werden, um bei jedem Reset des Credit Counters auch den Available Counter neu zu initialisieren. Ein Reset wird nur ausgelöst, wenn gerade keine Station mehr eine Zelle in ihrem Einfügepuffer hat. Hat eine Station i in einem Rahmen Z_i Zeitschlitze reserviert, wobei FOR Rahmen auf dem Ring zirkulieren, kann sie $FOR \cdot Z_i$ Zeitschlitze als „available“ markieren und den Available Counter auf dieses Produkt initialisieren. Allerdings ist es möglich, daß eine Station, die lediglich aufgrund fehlender Daten im Sendepuffer inaktiv war, bereits vor dem Reset zu senden beginnt. Für diesen Fall muß der Einfügepuffer zusätzlich eine ausreichende Anzahl von Zeitschlitzen bereit halten. Die Obergrenze dieser zusätzlichen Pufferkapazität ist wiederum $FOR \cdot Z_i$ Zeitschlitze.

4.4 Entwurf einer Netzknotenstruktur

Eine Übersicht der Datenflußarchitektur eines Netzknotens ist in Bild 4-8 dargestellt. Das Bus-Interface stellt die Schnittstelle zu einem Standard-Rechnerbus dar, über das die Sende- und Empfangsdaten des Rechners sowie Steuer- und Management-Information umgesetzt werden. Ankommende Datenrahmen durchlaufen eine Reihe logischer Einheiten, die vom Bus-Interface über den internen Bus gesteuert werden:

- Das physikalische Empfangsinterface realisiert die Bitübertragungsschicht mit optisch/elektrischer Wandlung, Taktrückgewinnung, Seriell/Parallel-Wandlung und Leitungsdekodierung. Je nach Entfernung zwischen zwei Stationen sind verschiedene Ausführungen möglich.



Steuerpfad

Datenpfad

PEI	Physikalisches Empfangsinterface
MMF	Monitor- und Masterfunktion
E	Empfänger
PV	Protokollverarbeitung
S	Sender
PSI	Physikalisches Sendeinterface

Bild 4-8: Architektur eines Netzknotens

- Die Monitor- und Masterfunktionen werden nur einmal im Ring benötigt und beinhalten die Initialisierung des Rings, einen Elastizitätspuffer für den Ringschluß, Fehlerbehebung und weitere Master-Funktionen. In allen anderen Stationen beschränkt sich die Funktion dieser Einheit auf die Fehlererkennung.

- Der Empfänger stellt Empfangspuffer für ITC- und ATC-Zellen bereit, kopiert Empfangsdaten in die Empfangspuffer und gibt Zeitschlitze nach Bedarf frei.

- Die Protokollverarbeitung beinhaltet die Realisierung des Medienzugriffsprotokolls, den Einfügepuffer für den ATC-Verkehr, die Implemetierung des Fairness-Protokolls und die Steuerung des Sendevorgangs.

- Der Sender verfügt über Sendepuffer für ITC- und ATC-Zellen, multiplext auf Anforderung der Protokollverarbeitungseinheit die Sendedaten in den Rahmen und berechnet den Wert des HEC-Feldes.

- Das physikalische Sendeinterface stellt das Gegenstück zum physikalischen Empfangsinterface dar und übernimmt die Leitungskodierung, die Parallel/Seriell-Wandlung und die elektrisch/optische Wandlung.

Für die effiziente Implementierung der AAL-5-Protokollfunktionen ist eine eigenständige Einheit vorgesehen, die insbesondere das Aufteilen und Vereinigen von AAL-5-PDUs mit Hardware-Unterstützung realisiert. Dieser AAL-Protokollprozessor kann mit dem Bus-Interface über einen dedizierten Kanal kommunizieren und gleichzeitig über den internen Bus Daten senden oder empfangen.

4.4.1 Datenformat des Ringsegments

Das Medienzugriffsprotokoll ist annähernd unabhängig von der verwendeten Übertragungsrate, da bei konstanter Rahmendauer lediglich die Anzahl der Zeitschlitze je Rahmen von der Übertragungsrate abhängt. Eine Beispielrealisierung mit einer Datenrate in der Größenordnung von einem Gbit/s ist nur durch schnelle Parallelverarbeitung mit der verfügbaren Standard-Technologie machbar [95]. Eine Taktfrequenz von 30 MHz sowie eine Datenbusbreite von 32 bit ist die Grenze dessen, was mit programmierbarer Logik in Verbindung mit maximal vierlagigen Standard-Platinen und gängigen Steckverbindern handhabbar ist [14]. Für den Datenpfad im Netzzugang ergibt sich daraus eine maximale Datenrate von 960 Mbit/s. Somit beträgt die Anzahl Bits B je Rahmen bei einer Rahmendauer von 125 µs

$$B = 9,6 \cdot 10^8 \frac{\text{bit}}{\text{s}} \cdot 1,25 \cdot 10^{-4}\text{s} = 1,2 \cdot 10^5\text{bit}$$

Die Größe eines Zeitschlitzes beträgt 56 Oktetts oder 448 bit. Für die maximale Anzahl Zeitschlitze je Rahmen Z_{max} gilt unter Vernachlässigung des Rahmen-Headers

$$Z_{max} = \frac{B}{448\frac{\text{bit}}{\text{Zeitschlitz}}} = 267,86 \text{ Zeitschlitze}$$

Einige Bits werden für die Synchronisation und den Ausgleich von Taktschwankungen zwischen den Stationen benötigt. Für die Anzahl der Zeitschlitze pro Rahmen kann daher z. B. die Zweierpotenz 256 gewählt werden.

Tabelle 4-3 beschreibt die zum Ringsegment innerhalb der Station gehörenden Leitungen, die alle unidirektional vom Eingang in Richtung des Ausgangs der Station gerichtet sind.

Signal	Beschreibung
D[31..0]	Datenleitungen
SOF	Start Of Frame, markiert das erste Datenwort des Rahmens, also den Rahmen-Header
SOC	Start Of Cell, markiert das erste Datenwort eines Zeitschlitzes/einer Zelle
Idle	Kennzeichnet ein ungenutztes Datenwort ohne Inhalt (Pad)
CLK	Taktsignal, zu dem Daten- und Steuerleitungen synchron sind

Tabelle 4-3: Leitungen im Datenpfad des Rings innerhalb der Station

4.4.2 Bus-Interface

Das Bus-Interface bildet die Schnittstelle zwischen einem internen Bus, über den die Einheiten der Station kommunizieren, und einem externen Rechnerbus (der zum Beispiel ein PCI-Bus sein kann [75]). Zunächst wird der interne Bus, danach die Struktur der Buskopplung betrachtet.

4.4.2.1 Interner Bus

Die in Bild 4-8 vorgestellten Einheiten einer Station kommunizieren über einen internen Bus mit 32 bit breitem Datenbus. Bus-Master kann alternativ das Bus-Interface oder der AAL-Protokollprozessor sein, die sich die Busanforderungen teilen. Jegliche interne Kommunikation wird vom jeweiligen Bus-Master gesteuert und erfolgt zwischen ihm und einer weiteren Einheit. Die Kommunikation erfolgt über ein einfaches asynchrones Handshake-Protokoll. Jede Einheit hat einen festgelegten Adressbereich, in dem sie auf eine Anforderung durch den Bus-Master reagiert.

Ausnahmezustände, Fehler sowie Ereignisse, die einer Reaktion durch den angeschlossenen Rechner bedürfen, können mit dem Aktivieren einer Interrupt-Leitung angezeigt werden. Jede Einheit kann diese Leitung durch eine logische Oder-Verknüpfung aktivieren. Wenn das zugehörige Interrupt-Register einer Einheit abgefragt wird, setzt sie ihre Interrupt-Anforderung zurück.

Die Datenübertragung von und zu den insgesamt vier Dateneingangs- und Datenausgangspuffern ist synchron zum Takt des externen Busses, um eine möglichst hohe Datenrate zu erzielen. Der Takt des internen Busses wird aus dem des externen Busses abgeleitet.

Jeder der vier Puffer verfügt über eine eigene Anforderungsleitung, über die angezeigt wird, ob Daten vom oder zum Puffer transferiert werden können. Im Falle eines Empfangspuffers wird damit signalisiert, daß mindestens eine vollständige Zelle im Puffer vorhanden ist. Bei einem Sendepuffer bedeutet eine aktive Anforderungsleitung, daß im Puffer Platz für mindestens eine vollständige Zelle vorhanden ist.

Das Zurücksetzen der Einheiten erfolgt über zwei Reset-Leitungen. Der lokale Reset setzt alle Register der Station zurück, die keine unmittelbare Auswirkung auf den Datenfluß im Ring haben. Bei einem lokalen Reset wird zwar die Station zurückgesetzt, der Datenfluß im Ring aber nicht gestört. Der globale Reset setzt die Station komplett zurück und unterbricht auch den Datenfluß im Ring. Nach dem Einschalten wird ein globaler Reset ausgelöst, während ein Zurücksetzen des angeschlossenen Geräts am externen Bus einen lokalen Reset zur Folge hat.

4.4.2.2 Buskopplung

Die Aufgabe des Bus-Interface ist die Kopplung zwischen externem und internem Bus sowie dem AAL-Protokollprozessor. Beispiele externer Busse, die synchrone Datenübertragung erlauben und damit Spitzendatenraten in der Größenordnung von einem Gbit/s erreichen können, sind der SBus [67] oder der PCI-Bus [75].

Bild 4-9 zeigt die Blockstruktur des Bus-Interface. Daten- und Steuerschnittstellen existieren zum externen Bus, dem AAL-Protokollprozessor und dem internen Bus. Der Block „Datenauf-

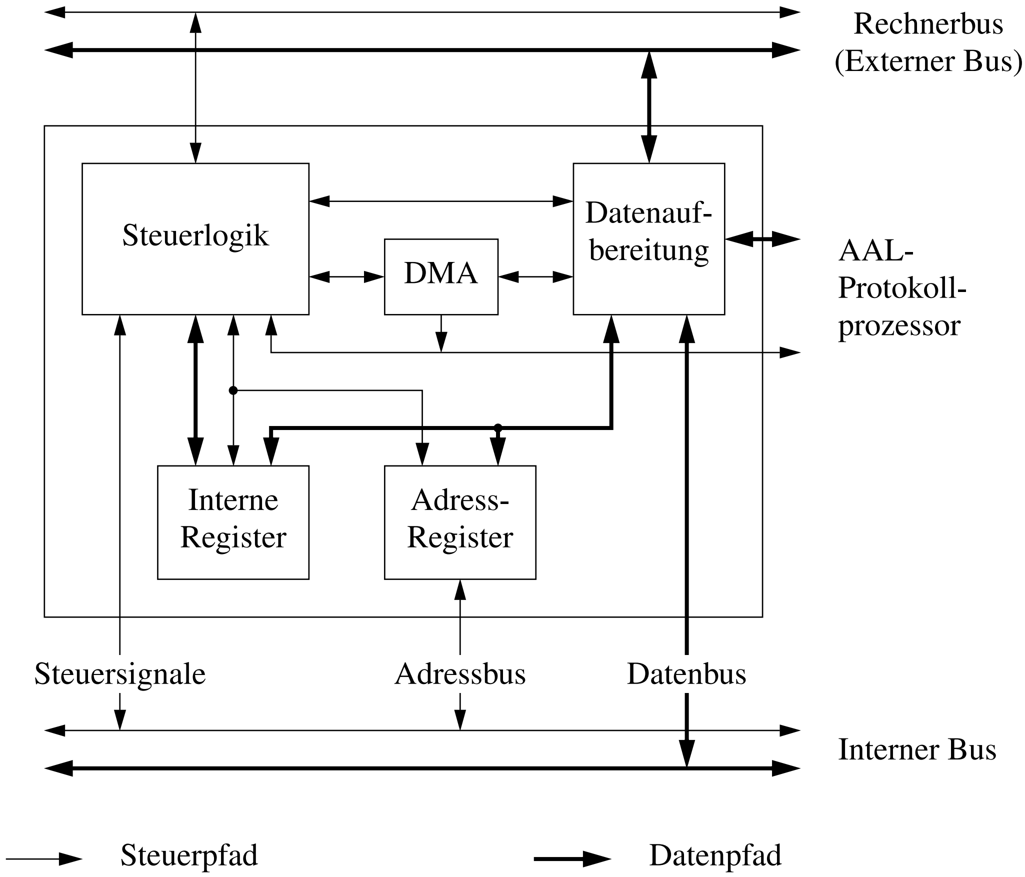


Bild 4-9: Bus-Interface

bereitung“ beinhaltet Datentore, Multiplexer und Datenpuffer, um die Daten jeweils im richtigen Format rechtzeitig auf einem der Busse auszugeben.

Um den Adressbereich auf dem externen Bus einzuschränken, werden die Adressen des internen Busses in einem Adressregister gespeichert. Somit werden auf dem externen Bus lediglich zwei Adressen benötigt: Eine für den Zugriff auf den Datenbus sowie eine für den Zugriff auf das Adressregister, um die Adresse des internen Busses festzulegen. Für den Zugriff auf ein Register in einer der Einheiten werden mit diesem Verfahren zwei externe Buszugriffe benötigt, um die Adresse des Registers einzustellen und die Daten zu transferieren. Der Zugriff auf ein Datenpaket erfordert jedoch nur einmalig das Einstellen einer Adresse, danach können komplette Datenpakete synchron im Burst-Modus übertragen werden.

Die DMA-Einheit steuert den synchronen Buszugriff auf dem externen Bus. Die Einheit ist während des Datentransfers der Bus-Master des externen Busses. Der DMA-Einheit können ein oder mehrere Aufträge für eine Datenübertragung gegeben werden, die sie in einer Warteschlange zwischenspeichert und dann sequentiell abarbeitet.

Die Steuerlogik bestimmt den Ablauf der Funktionen innerhalb des Bus-Interfaces. Abhängig vom Zustand der internen Register und den Anforderungen durch den externen Bus, den internen Bus, die DMA-Einheit sowie des AAL-Protokollprozessors werden Steuersignale generiert oder interne Register modifiziert.

4.4.3 Physikalisches Sende- und Empfangs-Interface

Das physikalische Sende-Interface, die Lichtwellenleiterstrecke und das dazugehörige Empfangs-Interface stellen eine transparente Übertragungsstrecke zwischen zwei Stationen bereit. Sie können je nach Anforderung unterschiedlich realisiert sein. Zu Testzwecken im Labor, wo zwei Stationen unmittelbar benachbart aufgebaut sind, kann das Interface-Paar aus parallelen Leitungen bestehen, die die interne Datenstruktur des Ringsegments lediglich wie in Bild 4-10 dargestellt als differentiell verstärkte Signale weitergeben. Zur Synchronisation und Regeneration der Daten wird am Empfänger ein Register eingefügt.

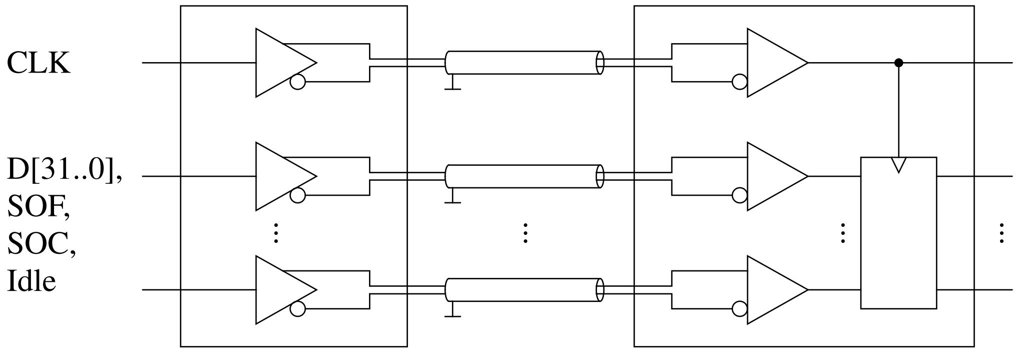


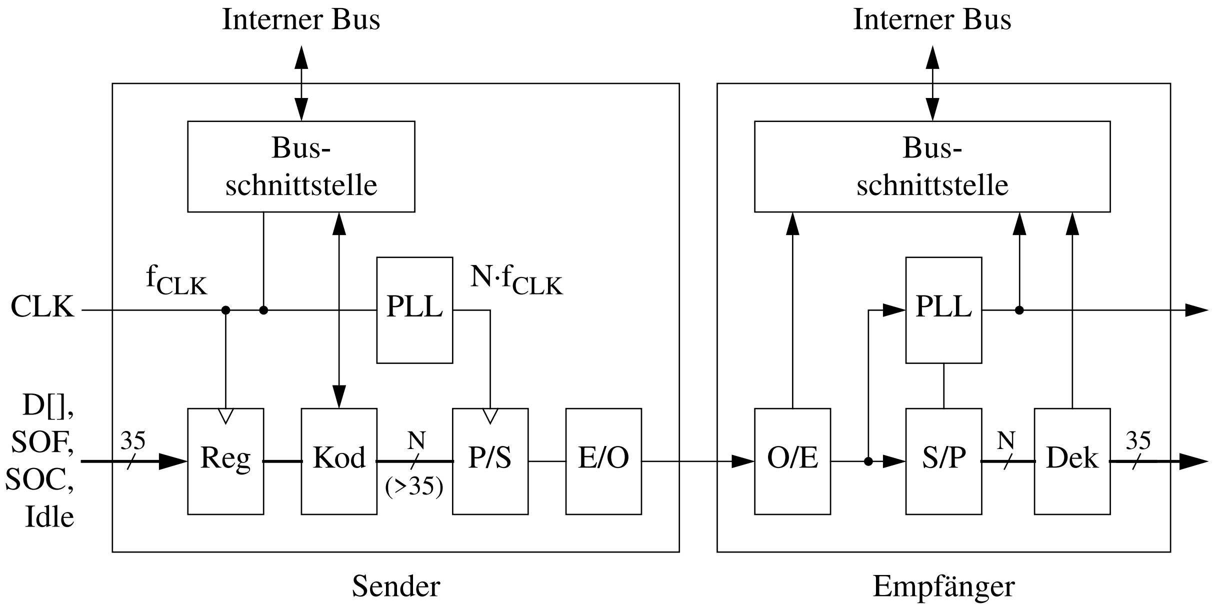
Bild 4-10: Physikalisches Sende- und Empfangs-Interface: Parallelkabel

Eine serielle optische Datenübertragung kann mit Standardbaugruppen erfolgen, die auf der Senderseite bereits die Taktvervielfachung, Leitungskodierung, Parallel/Seriell-Wandlung und elektrisch/optische Wandlung integriert haben. Auf der Empfängerseite werden demnach optisch/elektrische Wandlung, Taktrückgewinnung, Seriell/Parallel-Wandlung und Leitungsdekodierung benötigt. In Bild 4-11 wird von einem Baugruppenpaar ausgegangen, das synchron zum Eingangstakt das 35 bit breite Datenwort kodiert und seriell transparent überträgt, so daß am Ausgang des Empfangs-Interface der ursprüngliche Datenwort-Strom zusammen mit dem zurückgewonnenen Takt zur Verfügung steht.

Für die Leitungskodierung wird ein Kodierverfahren mit Redundanz verwendet, das eine sichere Taktrückgewinnung erlaubt, das Erkennen mindestens eines Bitfehlers ermöglicht und zusätzliche Steuercodes in den Datenstrom einfügen kann, wenn das Eingangsdatenwort als „Idle“ gekennzeichnet ist. Über die Busschnittstelle beim Sender kann die Leitungskodierung beeinflußt werden. Das optisch/elektrische Wandlungsmodul beim Empfänger signalisiert der Busschnittstelle, ob am Eingang ein optisches Signal anliegt. Fehlerfälle und eventuell beim Sender eingefügte Steuerworte werden vom Leitungsdekodierer an die Busschnittstelle weitergegeben.

4.4.4 Monitor- und Mastereinheit

Die Funktionalität der Monitor- und Mastereinheit [27] teilt sich auf in die eigentlichen Monitor- und Masterfunktionen, die nur einmal im Ring benötigt werden, und die Mechanismen zur



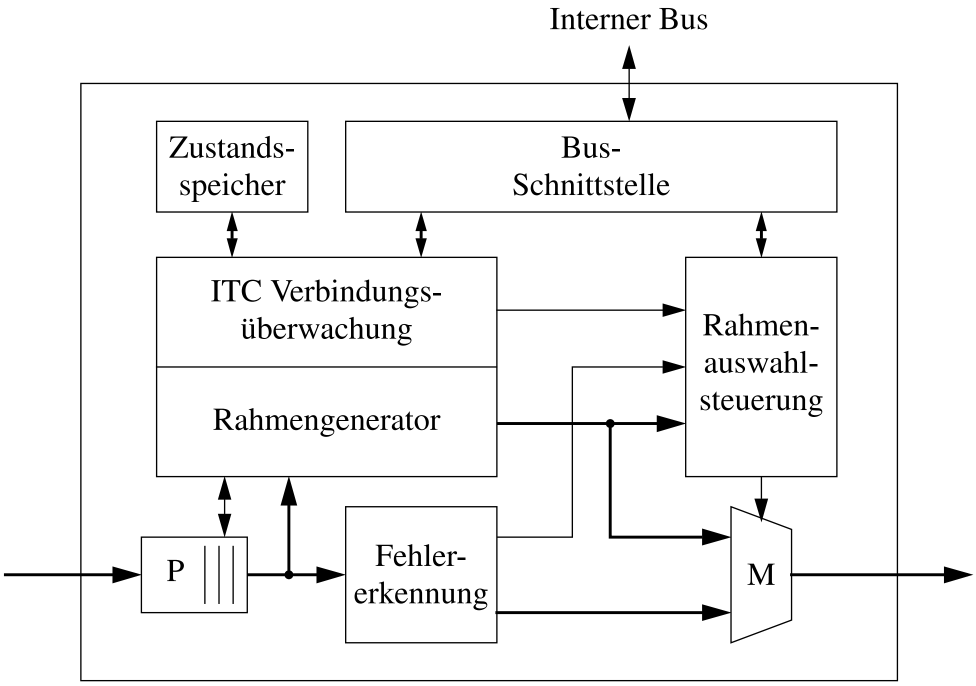
D[]	Datenleitungen D[31..0]
Reg	Eingangsregister
Kod	Leitungskodierung
P/S und S/P	Parallel-/Seriellwandlung und umgekehrt
E/O und O/E	Elektrisch-/optische Wandlung und umgekehrt
PLL	Phase Locked Loop-Schaltung zur Taktregeneration, Taktvervielfachung, Taktteilung
Dek	Leitungsdekodierung

Bild 4-11: Physikalisches Sende- und Empfangs-Interface: Optische Übertragung

Synchronisierung und Fehlerbehebung, die in jeder Station vorhanden sind. Um Flexibilität und ein einheitliches Stationskonzept sicherzustellen, ist die gesamte Funktionalität in allen Stationen vorhanden, jedoch automatisch über ein geeignetes Auswahlprotokoll oder manuell per Schalter oder Software konfigurierbar.

Bild 4-12 zeigt den Aufbau der Monitor- und Mastereinheit. Die Blöcke zur Eingangssynchronisation und Fehlererkennung sind diejenigen, die in jeder Station vorhanden sind. Abhängig vom verwendeten Übertragungssystem kann die Synchronisation bei einem synchronen System entfallen oder bei einem plesiochronen System mit einem Elastizitätspuffer realisiert werden. Die Fehlererkennung und -behebung beschränkt sich bei einer einfachen Station auf Bitfehlererkennung und -behebung im Zeitschlitz-Header und auf Rahmenfehler, also Inkonsistenzen bei Rahmennummer oder Zeitschlitzanordnung.

Der Rahmengenerator erzeugt laufend ein Rahmengerüst, das als Basis für den Datenstrom am Ausgang der Einheit und damit im gesamten Ring dient. In der Monitor- und Mastereinheit erfüllt die Eingangssynchronisation zusätzlich die Aufgabe, den Ring zu schließen, so daß die Rahmendaten am Ausgang der Eingangssynchronisation synchron zum Ausgangsstrom des Rahmengenerators sind. Somit kann der Multiplexer M im vom Rahmengenerator erzeugten Rahmengerüst die Zeitschlitze durch ankommende Zeitschlitze ersetzen, sofern diese fehler-



P Eingangssynchronisierung (Puffer)
M Multiplexer zur Auswahl der Rahmendaten

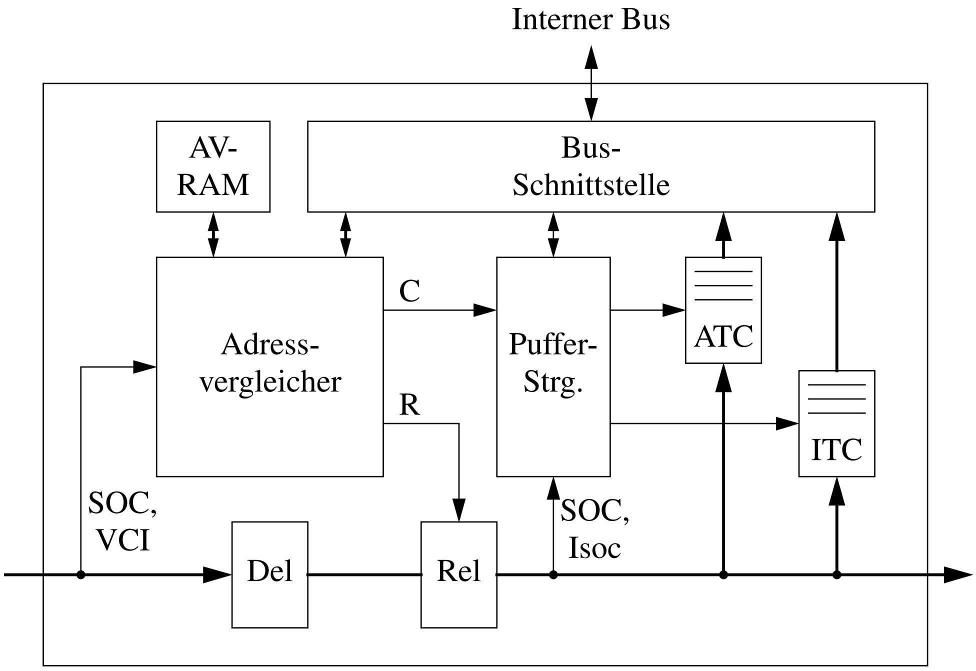
Bild 4-12: Monitor- und Masterfunktionen

frei sind. Im Fehlerfall werden die generierten Zeitschlitze in den Datenstrom übernommen, so daß am Ausgang der Monitor- und Mastereinheit in jedem Fall ein gültiger, fehlerfreier Datenstrom zur Verfügung steht.

Der Rahmengenerator muß reservierte Zeitschlitze im Fehlerfall als „unavailable“ markieren, während nicht reservierte Zeitschlitze als „available“ gekennzeichnet werden. Im Zustandsspeicher, auf den der Rahmengenerator Zugriff hat, existiert hierzu ein Abbild des Reservierungszustands jedes Zeitschlitzes. Im Rahmen-Header, der nur von der Monitor- und Masterstation modifiziert wird, trägt der Rahmengenerator die fortlaufende Rahmennummer und die Anzahl der Rahmen auf dem Ring ein, die er aus der Differenz zwischen der Nummer des gerade generierten Rahmens und der des ankommenden Rahmens ermittelt.

Über die Busschnittstelle wird das RAM beschrieben und ausgelesen. Außerdem können Zustände abgefragt, Fehler signalisiert oder Fehlerstatistiken generiert und über die Busschnittstelle ausgelesen werden.

Die Annahme, Bereitstellung, Verwaltung und Überwachung von ITC- und explizit aufgebauten ATC-Verbindungen wird in Software auf dem an die Monitor- und Masterstation angeschlossenen Rechner realisiert.



AV-RAM	Adressvergleicherspeicher
Del	Verzögerungseinheit (Pipeline-Architektur)
Puffer-Strg.	Steuereinheit für Empfangspuffer
Rel	Zeitschlitzfreigabe, setzt „Busy“-Bit zurück
R	Release-Signal zur Zeitschlitzfreigabe
C	Kopiersignal für Empfangspuffer

Bild 4-13: Empfänger

4.4.5 Empfänger

Für den Empfang von Daten sowie das Freigeben von Zeitschlitzen ist die Empfänger-Einheit [69] zuständig. Sie liest das VCI-Feld jeder Zelle aus und vergleicht den Inhalt mit allen Einträgen einer Liste von VCI-Werten, für die die Station als Empfänger konfiguriert ist sowie einer Liste von VCI-Werten, bei denen die Station die Zelle entfernt. Abhängig von den Tabelleneinträgen gibt die Station den Zeitschlitz frei und/oder kopiert die Zelle in einen der beiden Empfangspuffer. Bild 4-13 zeigt eine mögliche Realisierung der Empfänger-Einheit.

4.4.5.1 Adressvergleich

Für die Realisierung der VCI-Listen wird jeweils ein Assoziativspeicher benötigt, der zu jedem möglichen VCI-Wert einen zugehörigen Tabelleneintrag erfassen kann. Im einfachsten Fall beschränkt sich die Information auf ein Bit, das durch das Vorhandensein des Tabelleneintrags im Assoziativspeicher realisiert wird: ist ein Tabelleneintrag vorhanden, wird die Zelle kopiert oder gelöscht, ansonsten erfolgt keine Aktion.

Marktübliche Assoziativspeicher sind in der Anzahl der Tabelleneinträge auf eine Größenordnung von 1000 beschränkt [70]. Eine Alternative zum Assoziativspeicher ist ein RAM-Speicher mit vollständig auskodiertem Adressraum. Da die Adressierung mit dem 20 bit breiten VCI-Wert erfolgt, muß ein RAM entsprechend 2^{20} Speicherstellen umfassen. Für die beiden benötigten „Assoziativspeicher“ genügen folglich zwei RAM-Bausteine mit jeweils 1 Mbit Kapazität oder ein RAM-Baustein mit 2 Mbit und einer 2 bit breiten Datenorganisation. Ein 24 bit breiter VCI-Wert würde ein RAM mit insgesamt 32 Mbit erfordern. Im Gegensatz zu Assoziativspeichern müssen sämtliche Speicherstellen des RAM mit dem neutralen Wert (nicht kopieren, nicht freigeben) vorbelegt werden.

Die maximal erlaubte Zugriffszeit der RAM-Bausteine ist dadurch beschränkt, daß einerseits mindestens einmal je Zeitschlitz der Speicher gelesen wird und andererseits ein Zweitorspeicher im Zeitmultiplex nachgebildet werden muß, da eine Möglichkeit zum Beschreiben des Speichers benötigt wird. Insgesamt ist es ausreichend, wenn zwei RAM-Zugriffe während der Dauer eines Zeitschlitzes erfolgen können. Die Zeitschlitzdauer beträgt bei einem 32 bit-Worttakt von 30 MHz und 14 Worten je Zeitschlitz 467 ns. Somit stehen für jeden Speicherzugriff 233 ns zur Verfügung, was selbst für langsame dynamische Speicher ausreicht [71].

Das RAM kann nicht innerhalb einer Taktperiode auf eine ankommende Zelle reagieren, da es durch einen Zugriff durch die Bus-Schnittstelle belegt sein kann und der RAM-Zugriff länger als einen Takt dauert. Ferner müssen die mit Hilfe des RAM generierten Freigabe- und Kopiersignale gleichzeitig mit dem Zeitschlitzbeginn vorliegen, sind jedoch abhängig vom zweiten Datenwort im Zeitschlitz, das den VCI-Wert enthält. Daher muß eine Verzögerung eingefügt werden, deren Dauer der maximalen Reaktionszeit der RAM-Steuerung zuzüglich einem Takt entspricht. Realisiert wird eine solche Verzögerung durch eine Pipeline-Struktur.

4.4.5.2 Funktionen der Bus-Schnittstelle

Die Bus-Schnittstelle realisiert die Kommunikation zum internen Bus. Folgende Funktionen werden benötigt:

- Schreiben und Lesen der Adressvergleichertabelle
- Pufferfüllstandssignale, die anzeigen, ob eine oder wie viele komplette Zellen im Empfangspuffer stehen
- Synchrones Lesen der Daten aus den Empfangspuffern
- Signalisierung und Behebung möglicher Fehlerzustände
- Initialisierungsfunktionen

Für die beiden Empfangspuffer wird jeweils ein Signal auf einer dedizierten Leitung des internen Bus ausgegeben, wenn mindestens eine komplette Zelle im Puffer steht. Die Puffersteuerung zählt hierzu die vollständig in den Empfangspuffer geschriebenen Zellen und subtrahiert hiervon die vom internen Bus aus dem Puffer gelesenen Zellen. Ist diese – über den internen Bus abfragbare – Zahl größer als 0, wird die entsprechende dedizierte Leitung gesetzt. Damit wird erreicht, daß zum frühest möglichen Zeitpunkt der angeschlossene Rechner über angekommene Zellen informiert wird, um einem Pufferüberlauf vorzubeugen.

4.4.5.3 Management

Nach dem Einschalten muß der Adressvergleicherspeicher vorbelegt werden, bevor die Station damit beginnt, Zeitschlitze freizugeben oder Zellen zu kopieren. Zunächst wird der komplette Speicher mit „nicht freigeben“ und „nicht kopieren“ vorbelegt. Bei Anwendung des in Abschnitt 4.3.8 beschriebenen impliziten Adressierungsverfahrens muß die Station innerhalb des dafür reservierten Adressbereichs die folgenden abweichenden Einträge vornehmen:

- Broadcast-Zieladresse mit beliebiger Senderadresse: „Kopieren“
- Eigene Senderadresse mit beliebiger Zieladresse: „Freigeben“
- Eigene Zieladresse mit beliebiger Senderadresse: „Kopieren und Freigeben“

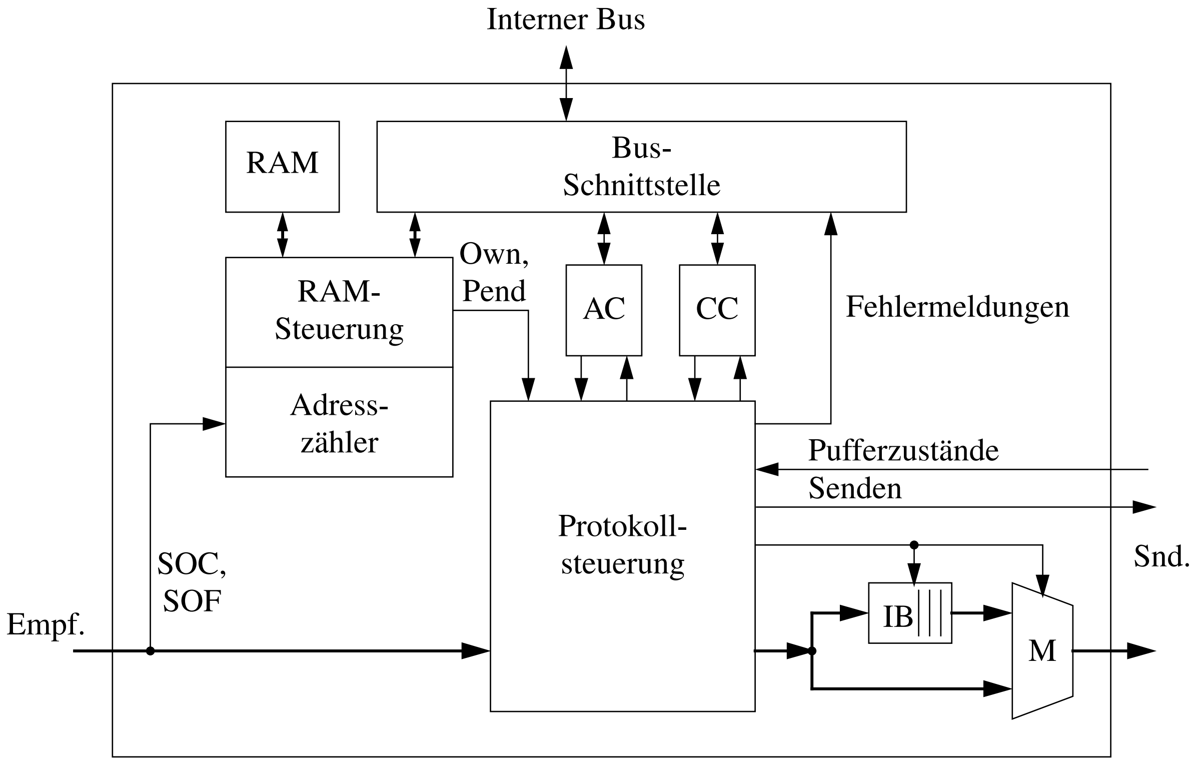
Da der VCI-Wert immer vollständig dekodiert wird, entsprechen der „beliebigen“ Adresse jeweils 256 einzelne Tabelleneinträge. Neben diesen Einträgen muß die Station ferner auf mögliche vordefinierte VCI-Werte reagieren, die für Verbindungsaufbau und andere Management-Aufgaben benötigt werden.

Im Normalbetrieb werden ankommende Zellen, die nicht mehr im Empfangspuffer Platz finden, vom Empfänger verworfen. Läuft jedoch ein Puffer aufgrund eines Fehlers dennoch über, dann müssen der übergelaufene Puffer und die mit ihm assoziierten Empfangsprozesse zurückgesetzt werden, da die Zellgrenze im Puffer durch einen möglichen teilweisen Zellverlust nicht mehr eindeutig ist.

4.4.6 Protokollverarbeitung

Die Protokollverarbeitungseinheit übernimmt die Teilnahme der Station am Medienzugriffsprotokoll und verfügt über einen Einfügepuffer für den ATC-Verkehr [90]. In Abhängigkeit von den Pufferfüllständen der Sendeeinheit (Abschnitt 4.4.7) und den Kriterien, die durch das Medienzugriffsprotokoll gegeben sind, entscheidet die Protokollverarbeitungseinheit, wie ein Zeitschlitz belegt wird. Bild 4-14 zeigt die Struktur dieser Einheit.

Wenn ein Zeitschlitz von der Empfänger-Einheit ankommt, muß die Station zunächst prüfen, ob dieser Zeitschlitz für sie reserviert ist („Own Slot“). Dazu führt die Station eine Tabelle in einem RAM, das über die Bus-Schnittstelle administriert werden kann. Für jeden Zeitschlitz in einem Rahmen existiert ein Tabelleneintrag entsprechend seiner Position innerhalb des Rahmens. Der Tabelleneintrag kann, neben der Kennzeichnung als „Own“ und „Pending“, für Erweiterungen weitere zeitschlitzspezifische Parameter umfassen. Der Adresszähler wird um 1 erhöht, wenn ein Zeitschlitz ankommt (SOC gesetzt), und er wird zurückgesetzt, wenn der Beginn eines Rahmens signalisiert wird (SOF gesetzt). Da der Zählvorgang, abgesehen vom Rahmenanfang, deterministisch ist, kann der nächste Tabelleneintrag bereits dann gelesen werden, wenn der letzte Eintrag verarbeitet wurde. Am Rahmenanfang muß der Eintrag für den ersten Zeitschlitz im Rahmen sofort parallel zum Anfang des ersten Zeitschlitzes bereitstehen und ist daher in einem Register innerhalb der RAM-Steuerung gespiegelt. Die RAM-Steue-



Empf.	Datenstrom von der Empfänger-Einheit
SOC, SOF	Start of Cell, Start of Frame
RAM	Verbindungsspeicher (Liste), in dem der Besitz von Zeitschlitzen abgebildet wird
AC	Available Counter
CC	Credit Counter
Own	Signalisiert, daß der aktuelle Zeitschlitz der Station „gehört“ (Own Slot)
Pend	Pending: Der Zeitschlitz wird gerade freigegeben
Del	Verzögerungseinheit (Pipeline)
M	Multiplexer
IB	Einfügepuffer (Insertion Buffer)
Snd.	Datenstrom und Steuerleitungen zu/von der Sender-Einheit

Bild 4-14: Protokollverarbeitung

rung löst Zugriffskonflikte zwischen Bus-Schnittstelle und dem Protokollmechanismus durch Serialisierung auf.

In der Protokollsteuerung werden die notwendigen Entscheidungen bezüglich des aktuellen Zeitschlitzes getroffen: Die Modifikation von Bits im Zeitschlitz-Header, das Kopieren der Zelle in den Einfügepuffer, Auswahl der Zelle aus dem Einfügepuffer oder dem direkten Datenstrom durch Steuerung des Multiplexers M, die Modifikation von Available- und Credit Counter sowie die Steuerung der Datenmultiplexer in der Sender-Einheit (Abschnitt 4.4.7).

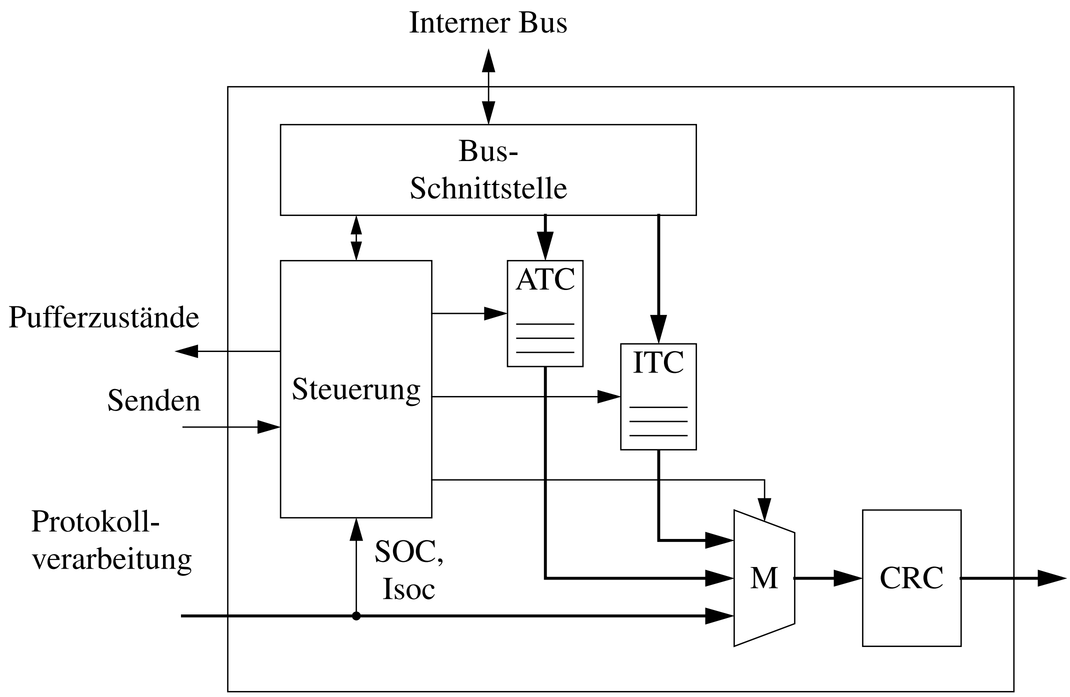
Über die Bus-Schnittstelle können der Available- und der Credit Counter sowie das RAM ausgelesen und beschrieben werden. Ferner können Fehlerzustände, insbesondere verbotene Bitkombinationen im Zell-Header, gezählt und ausgewertet werden.

Die Funktionen, die für die Realisierung der Protokollverarbeitungseinheit benötigt werden, wurden bereits in Abschnitt 4.3 vorgestellt. Im Anhang A wird auf die Details einer möglichen Realisierung ausführlich eingegangen.

4.4.7 Sender

Die Sender-Einheit [59] multiplext zu sendende Zellen in den Rahmen, wobei der Multiplexer von der Protokollverarbeitungseinheit gesteuert wird. Ferner generiert die Einheit die Prüfbits im Zeitschlitz-Header. Bild 4-15 zeigt die Funktionseinheiten des Senders.

Für die Sendepuffer für ATC- und ITC-Verkehr wird jeweils ein Signal auf einer dedizierten Leitung des internen Bus ausgegeben, wenn mindestens für eine komplette Zelle freier Platz im Puffer vorhanden ist. Die Schnittstelle zum internen Bus ist vergleichbar mit der Empfänger-Einheit, lediglich mit umgekehrtem Datenfluß. Ein Pufferüberlauf ist ausgeschlossen, da für den Datenfluß zum Sendepuffer mit den beschriebenen Leitungen eine effiziente Flußsteuerung erfolgt. An die Protokollverarbeitung werden Pufferzustände signalisiert. Für jeden Sendepuffer getrennt wird angezeigt, wenn eine vollständige Übertragungseinheit[1] im Puffer zum Senden bereitsteht.



M	Multiplexer
CRC	CRC-Generator, berechnet die Prüfsumme im Zeitschlitz-Header

Bild 4-15: Sender

1. Eine einzelne Zelle oder eine komplette AAL-5-CPCS-PDU, die ohne Unterbrechung übertragen werden soll

Für die Paketübertragung in der asynchronen Verkehrsklasse werden bereits beim Schreiben der Zellen in den ATC-Sendepuffer die Bits „Packet“ und „EOP“ entsprechend dem Zellstatus gesetzt, so daß die Steuerung Kenntnis über den Status jeder Zelle im ATC-Puffer hat. Die Steuerung informiert über eine Statusleitung die Protokollverarbeitung über eine laufende Paketübertragung.

Die CRC-Berechnung über den 5 Oktetts umfassenden Zeitschlitz-Header basiert auf einem rückgekoppelten Schieberegister mit dem Generatorpolynom $G(u) = u^8 + u^2 + u + 1$ und geht von einem seriellen Bitstrom aus [48]. Für die parallele Berechnung der CRC-Prüfbits für den gesamten Header in einem kombinatorischen Netzwerk wird das serielle Einschreiben in das rückgekoppelte Schieberegister zunächst schrittweise nachvollzogen, wobei für jedes der eingeschobenen Bits i eine Variable M_i eingeführt wird. Unter Ausnutzung der Modulo-2-Rechnung, bei der $M_i + M_i = 0$ gilt, ergibt sich für jedes Prüfbit eine Anzahl von Variablen M_i, die jeweils modulo 2 miteinander verknüpft werden. Maximal müssen bei dem gewählten Generatorpolynom 19 Terme modulo 2 verknüpft werden [59]. Der Datenstrom wird mit einer Pipeline verzögert, so daß das Ergebnis der CRC-Berechnung nach der maximalen kombinatorischen Durchlaufzeit in den Datenstrom eingefügt werden kann.

4.4.8 AAL-Protokollprozessor

Die Funktion des AAL-Protokollprozessors ist es, den am externen Bus angeschlossenen Rechner von einem Teil der AAL-Protokollfunktionen zu entlasten. Insbesondere das Segmentieren und Zusammensetzen sowie die Berechnung der Prüfsequenz im AAL-5 ergeben einen erheblichen Aufwand, den der Protokollprozessor mit Hardware-Unterstützung übernehmen kann. Die Schnittstelle des Protokollprozessors über das Bus-Interface zum externen Bus ist zweigeteilt: Sie kann ATM-Zellen transparent übertragen und damit lediglich die Funktionalität der ATM-Schicht bereitstellen, oder sie kann einen Teil der AAL-5-Funktionalität bereitstellen und AAL-5-CPCS-PDUs übertragen. Die Integration des Protokollprozessors in die Schnittstellenkarte einer Station ist nicht zwingend notwendig und kann entfallen, wobei dann die Station selbst die komplette AAL-Funktionalität übernehmen muß [19].

4.4.8.1 Struktur des Protokollprozessors

Die wesentlichen Elemente des AAL-Protokollprozessors sind ein Mikroprozessor, eine CRC-Berechnungseinheit sowie ein Zweitorspeicher, wie in Bild 4-16 dargestellt. Der Mikroprozessor verfügt über zwei getrennte Datenbusse, wovon der eine zur Kommunikation mit dem Bus-Interface genutzt wird, während der andere den internen Bus und die CRC-Berechnungseinheit bedient. Die vier Status-Leitungen der Sende- und Empfangspuffer können im Mikroprozessor Unterbrechungen auslösen, um die Rechenleistung effektiv zu nutzen und Pufferüberläufe zu verhindern. Der Mikroprozessor steuert die Tore für Daten- und Steuerleitungen zur CRC-Berechnungseinheit und zum internen Bus.

Der Speicher des Mikroprozessors ist mit einem RAM realisiert, auf das der Mikroprozessor selbst sowie das Bus-Interface gleichzeitig zugreifen können. Mögliche Zugriffskonflikte auf

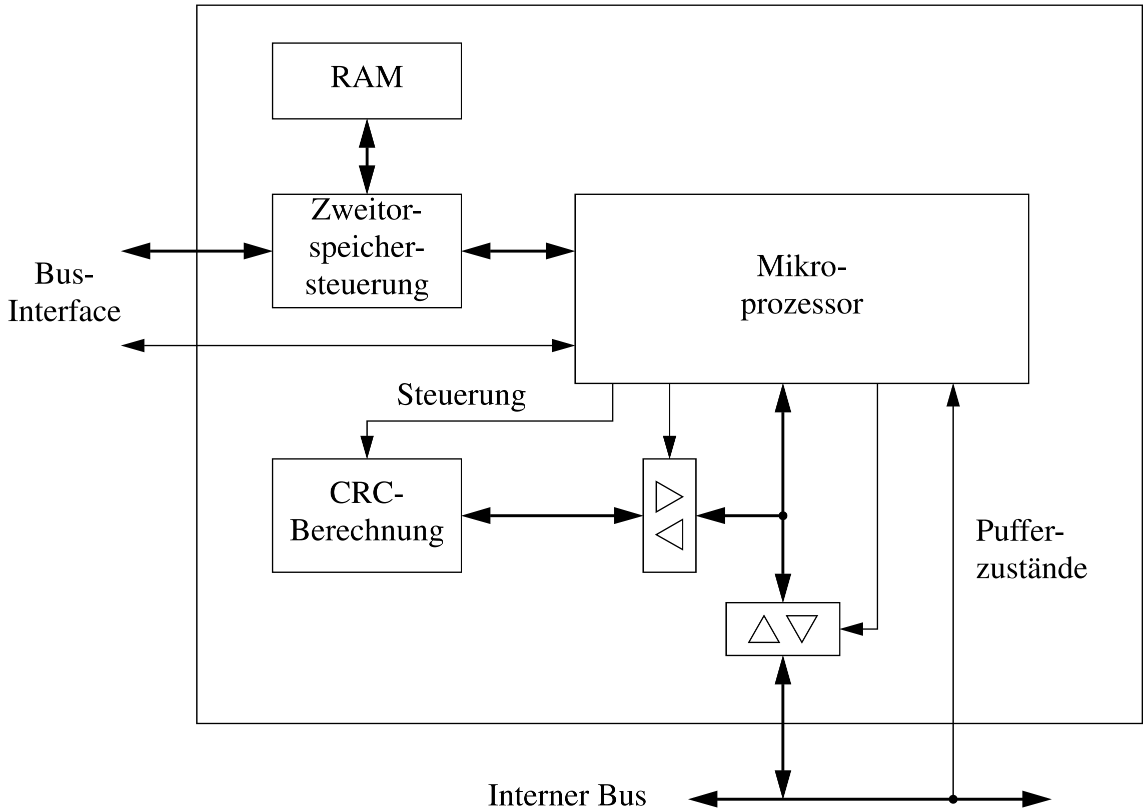
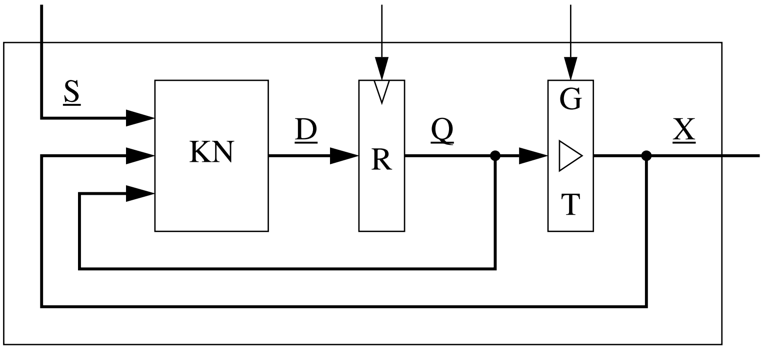


Bild 4-16: AAL-Protokollprozessor

das RAM löst die Zweitor-Speichersteuerung auf. Damit kann über das Bus-Interface vom externen Bus direkt auf den Speicher des Mikroprozessors zugegriffen werden, ohne den Mikroprozessor zu belasten.

Die CRC-Berechnung im AAL-5 übernimmt die CRC-Berechnungseinheit, die in Bild 4-17 detailliert dargestellt ist, parallel zum Datenstrom vom oder zum internen Bus in Echtzeit. Vor dem Start einer AAL-5-PDU wird die CRC-Einheit initialisiert. Die Realisierung der CRC-Berechnung ist äquivalent zu einem bitweise rückgekoppelten Schieberegister, erfolgt jedoch für 32 Bit parallel nach einer Funktion $\underline{Q}^{n+1} = \underline{F}(\underline{Q}^n,\underline{X}^n)$. Das kombinatorische Netzwerk KN kann in Abhängigkeit vom Steuervektor $\underline{S}$ neben der Funktion $\underline{F}$ weitere Funktionen realisieren, insbesondere die Initialisierung des Registers mit einem Anfangswert $\underline{Q}^{n+1} = \underline{X}^n$ sowie die Erhaltung des aktuellen Registerinhalts $\underline{Q}^{n+1} = \underline{Q}^n$. Am Ende einer AAL-5-PDU wird der berechnete CRC-Wert bei zu sendenden PDUs als letztes Datenwort an den Datenstrom angefügt oder kann vom Mikroprozessor ausgelesen und ausgewertet werden.

Die CRC-Berechnung kann unterbrochen werden, um eine andere Verbindung zu bearbeiten. In diesem Fall liest der Mikroprozessor den aktuellen Inhalt des Registers R aus und speichert ihn in einer der aktuellen Verbindung zugeordneten Datenstruktur zwischen. Soll die Übertragung der PDU wieder aufgenommen werden, dann wird das Register R vorher mit dem ursprünglichen Wert beschrieben.



KN	Kombinatorisches Netzwerk
R	32-Bit-Register (D-Flip-Flops)
T	Datentor
G	Steuereingang am Datentor
$\underline{S}$	Steuervektor
$\underline{D}$	Eingangsvektor des Registers
$\underline{Q}$	Ausgangsvektor des Registers, $\underline{D}^n = \underline{Q}^{n+1}$
$\underline{X}$	Eingangsvektor bei geschlossenem Datentor T, $\underline{X} = \underline{Q}$ sonst

Bild 4-17: CRC-Berechnungseinheit

Für die Bearbeitung der diversen Sende- und Empfangsfunktionen werden im Mikroprozessor Prozesse definiert, die logisch parallel ablaufen. Eine Ablaufsteuerung übernimmt die Serialisierung der Prozesse, wobei einzelne Prozesse gegenüber anderen priorisiert sein können. Insbesondere muß die Ablaufsteuerung dafür Sorge tragen, daß Daten in einem Empfangspuffer, der überzulaufen droht, möglichst rasch bearbeitet werden.

4.4.8.2 Senden von AAL-5-CPCS-PDUs

Beim Senden wird die zu sendende AAL-5-CPCS-PDU über den externen Bus in den Speicherbereich des Protokollprozessors geschrieben, der für das Senden reserviert ist. Parameter sind hierbei die Länge der PDU, der zugeordnete VCI sowie die Verkehrsklasse. Sobald 48 Oktetts der PDU oder aber die komplette PDU im Speicher stehen, kann der Protokollprozessor den Header der ersten SAR-PDU generieren und in den der CPCS-PDU-Verkehrsklasse entsprechenden Sendepuffer transferieren. Anschließend überträgt er die 48 Oktetts, wobei er parallel die CRC-Berechnung durchführt. Ist das Ende der AAL-5-CPCS-PDU erreicht, berechnet der Protokollprozessor die Anzahl der benötigten Füll-Oktetts. Der Trailer wird während des Übertragens der letzten Zelle generiert und am Ende der Zelle eingefügt. Damit ist die Übertragung abgeschlossen. Der Speicher wird freigegeben und das erfolgreiche Absenden des Pakets mit einer möglichen Unterbrechungsanforderung über den externen Bus gemeldet.

Um die beiden Verkehrsklassen zu unterscheiden, kann für jede Klasse ein eigener Speicherbereich reserviert werden. Der Protokollprozessor bearbeitet dann eine PDU der isochronen Verkehrsklasse immer bevorzugt, um eine möglichst geringe Verzögerung in dieser Verkehrsklasse zu gewährleisten. Wird von einem der Sendepuffer signalisiert, daß er keine komplette Zelle mehr aufnehmen kann, muß der Protokollprozessor die Übertragung so lange aufschieben, bis wieder Platz im Puffer vorhanden ist.

Verbindungen, die nicht auf AAL-5 beruhen, werden bereits von dem am externen Bus angeschlossenen Rechner verarbeitet, der ATM-Zellen an den Protokollprozessor übergibt, die dieser transparent in den entsprechenden Sendepuffer weiterleitet. Alternativ könnte eine Zelle direkt unter Umgehung des Protokollprozessors in den Sendepuffer transferiert werden, wobei dann sichergestellt werden muß, daß sich Sendepuffer und externer Bus beim Schreiben einer Zelle in einen Sendepuffer nicht gegenseitig unterbrechen.

4.4.8.3 Empfangen von AAL-5-CPCS-PDUs

Der Empfangsprozeß für AAL-5-CPCS-PDUs ist ungleich komplexer als das Senden, weil nicht sichergestellt ist, daß ein Zellstrom, der eine AAL-5-PDU darstellt, nicht durch Zellen anderer Verbindungen unterbrochen wird. Durch den in Abschnitt 4.3.7 beschriebenen Mechanismus, innerhalb des LANs Pakete immer ohne Unterbrechung zu senden, wird die Anzahl der gleichzeitig zu empfangenden PDUs innerhalb des LANs auf 1 reduziert. Bei Verbindungen über einen Router von außerhalb des LANs ist die Sequenz der Zellen einer AAL-5-PDU jedoch durch Zellen anderer Verbindungen unterbrochen.

Wird von einem der Empfangspuffer signalisiert, daß eine komplette Zelle angekommen ist, wertet der Protokollprozessor zunächst den Zellkopf aus. Mit dem VCI kann die Zelle der richtigen Verbindung zugeordnet werden. Die möglichen VCI-Werte lassen sich in drei Arten einteilen:

a) Der VCI entspricht einer direkten Adresse nach Abschnitt 4.3.8 und gehört somit zu einer AAL-5-PDU, die nicht von einer weiteren AAL-5-PDU unterbrochen wird, deren Quelle innerhalb des Rings liegt.

b) Der VCI entspricht einem Eintrag in der Tabelle des Protokollprozessors, in der explizit aufgebaute AAL-5-Verbindungen eingetragen sind, die mit Unterstützung des Protokollprozessors verarbeitet werden.

c) Der VCI entspricht keinem der beiden Fälle a) und b); die Zelle wird also unverarbeitet als ATM-Zelle vom Protokollprozessor zum externen Bus weitergereicht.

Für jede der drei Klassen wird ein separater Empfangsprozeß definiert. Der Hauptverkehrsanteil innerhalb eines lokalen Rechnernetzes ist Datenverkehr nach a), der daher getrennt und möglichst effizient behandelt wird. Bei explizit aufgebauten Verbindungen nach b) muß für jede Verbindung mindestens ein Puffer für eine AAL-PDU maximaler Größe bereitgehalten werden, da sich derartige PDUs gegenseitig unterbrechen. Auch Zellen nach c) können nicht direkt über den internen Bus und das Bus-Interface unter Umgehung des Protokollprozessors aus dem Empfangspuffer gelesen werden, da erst nach Auswertung des VCI feststeht, welcher der drei Verbindungsarten die Zelle zugeordnet ist.

a) Zellen von AAL-5-PDUs impliziter Verbindungen innerhalb des LAN

Die Zugehörigkeit einer Zelle zu einer AAL-PDU nach a) wird am gesetzten „Packet“-Bit erkannt, so daß der Protokollprozessor den VCI-Wert nicht auswerten muß. Die erste Zelle einer AAL-PDU wird in einen neuen Speicherbereich abgelegt. Der Protokollprozessor legt jetzt das Datenfeld jeder folgenden Zelle aus dem Empfangspuffer der gleichen Verkehrsklasse in diesem Puffer ab. Die letzte Zelle der PDU erkennt der Protokollprozessor am gesetzten EOP-Bit, woraufhin er bei einer fehlerfreien PDU die Weiterleitung der AAL-5-SDU an den angeschlossenen Rechner initiiert.

Mögliche Fehler, die der Protokollprozessor erkennen muß, sind Fehler der Prüfsequenz, also fehlerhafte Zellen oder Zellverluste, sowie ein fehlendes EOP-Bit, das zu einem Pufferüberlauf führen kann. Hierzu prüft der Protokollprozessor die Länge der bereits empfangenen PDU und erkennt den Fehler, sobald die PDU-Länge die maximale Länge überschreitet. Eine weitere Möglichkeit der Fehlersicherung ist die Überprüfung des VCI-Wertes auf den erwarteten Wertebereich.

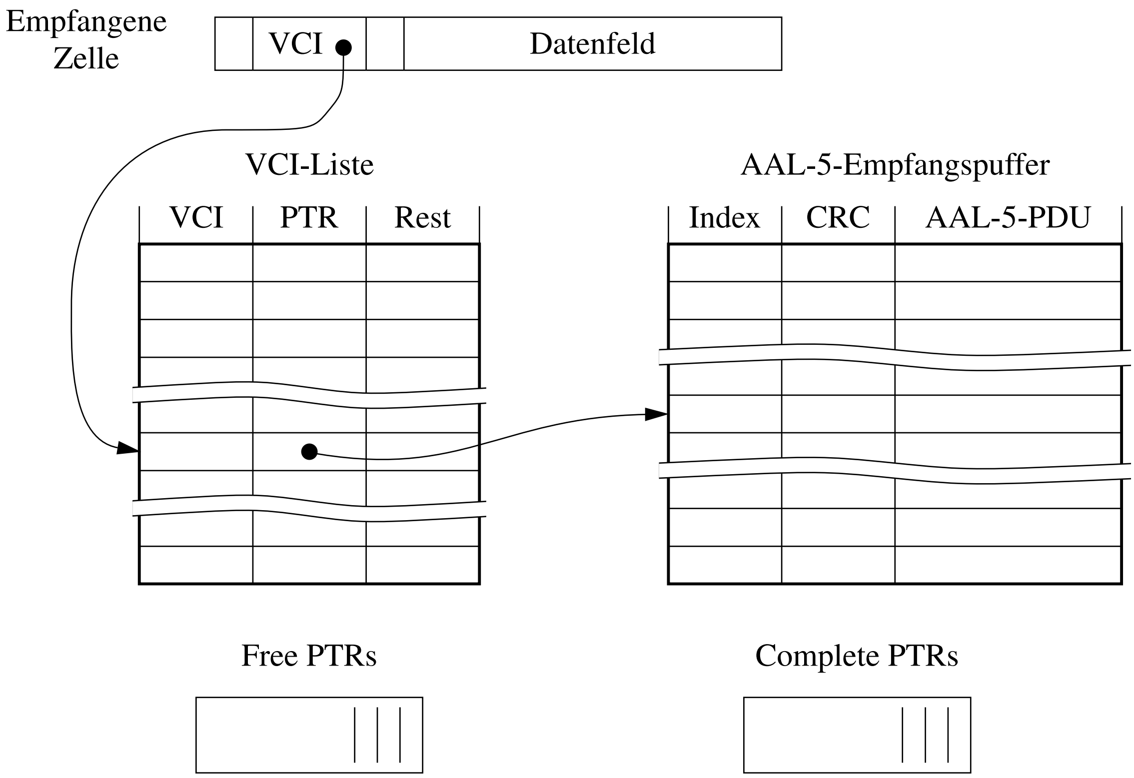
Die Anzahl der bereitzustellenden Puffer ist abhängig von der Zeitspanne, die das Lesen der PDU aus dem Speicher des Protokollprozessors dauert. Der Protokollprozessor muß in der Lage sein, sofort nach dem vollständigen Empfangen einer AAL-PDU die nächste PDU zu puffern. Er benötigt also minimal zwei Pufferplätze.

b) Zellen von AAL-5-PDUs explizit aufgebauter Verbindungen

Die Verwaltung der Empfangsfunktion explizit aufgebauter Verbindungen ist vergleichsweise komplex. Bild 4-18 zeigt die erforderliche Datenstruktur im Speicher des Protokollprozessors. Beim Verbindungsaufbau wird der gewählte VCI in die VCI-Liste eingetragen. Da in dieser Liste für jede empfangene Zelle gesucht werden muß, sollte die Suche möglichst effizient sein. Mögliche Realisierungsformen [106] reichen von der linearen Suche in der nicht sortierten Liste über das binäre Suchen, das eine sortierte Liste voraussetzt, und ein Hashing-Verfahren bis zu einer Realisierung mit einem separaten Assoziativspeicher [70]. Zu dem Listeneintrag gehört neben der VCI-Nummer ein Zeiger auf den Puffer, in dem die aktuelle AAL-5-PDU zusammengesetzt wird. Weitere verbindungsabhängige Einträge können ebenfalls in die Liste integriert werden.

Nachdem der VCI-Wert in die Liste eingetragen oder, falls das gewählte Suchverfahren es erfordert, einsortiert ist, muß ein Puffer für die erste AAL-5-PDU bereitgestellt werden. Hierzu wird aus dem Vorrat der Zeiger auf freie Puffer ein Zeiger entnommen und in die Liste eingetragen. Der zugehörige Pufferplatz wird initialisiert.

Wenn der Header einer Zelle eintrifft, wird der VCI-Wert der Zelle in der VCI-Liste gesucht. Ist ein Eintrag vorhanden, ist über den zugeordneten Zeiger der Pufferplatz für die Zelle definiert. Mit dem darin befindlichen CRC-Wert wird die CRC-Berechnungseinheit geladen und anschließend das Datenfeld der Zelle aus dem Empfangspuffer in die durch den Index bezeichnete Stelle innerhalb des Pufferplatzes transferiert. Die parallel berechnete Prüfsumme wird anschließend von der CRC-Berechnungseinheit zurück in den Pufferplatz geschrieben. Der



VCI	Virtual Channel Identifier
PTR	Zeiger auf den aktuellen AAL-5-Empfangspuffer des VCI
Rest	Weitere verbindungsabhängige Information
Index	Aktuelle Position des Schreibzeigers innerhalb der AAL-5-PDU
CRC	Aktueller Wert der CRC, aus der CRC-Berechnungseinheit
Free PTRs	Vorrat der Zeiger auf freie Puffer
Complete PTRs	Warteschlange der Zeiger auf komplette, abzuarbeitende PDUs

Bild 4-18: Datenstruktur für den AAL-5-Empfang bei explizit aufgebauten Verbindungen

Index wird um die Länge des Zell-Datenfeldes erhöht. Damit ist der Empfang der Zelle abgeschlossen.

Ist eine AAL-5-PDU komplett empfangen, wird der Zeiger auf den Pufferplatz, in dem die PDU steht, von der VCI-Liste in die Warteschlange der Zeiger auf komplette, abzuarbeitende PDUs transferiert und ein neuer Pufferplatz dem VCI-Wert zugewiesen. Ein Prozeß arbeitet die „Complete PTRs“-Warteschlange ab und stößt das Kopieren der PDU in den angeschlossenen Rechner über das Bus-Interface und den externen Bus an. Wenn der Kopiervorgang beendet ist, wird der Zeiger auf den Pufferplatz der Warteschlange entnommen und dem Vorrat an Zeigern auf freie Puffer zugefügt.

c) Sonstige Zellen

Wird eine Zelle empfangen, deren VCI-Wert weder der Adressierung nach a), noch einem Eintrag in der VCI-Liste nach b) entspricht, wird sie in eine Ringspeicherstruktur innerhalb des Speichers transferiert, ohne den Zellinhalt zu ändern oder eine Prüfsumme zu berechnen. Unmittelbar nachdem die Zelle komplett im Speicher steht, wird das Transferieren der Zelle an den angeschlossenen Rechner über das Bus-Interface und den externen Bus angestoßen.

Kapitel 5

Bewertung der Verkehrsleistung

Dieses Kapitel beschreibt eine Untersuchung der Leistungsfähigkeit des vorgestellten Medienzugriffsprotokolls durch Simulation. Zunächst wird in Abschnitt 5.1 das dem Simulationsprogramm zugrundeliegende Modell der Stationen vorgestellt. Die Simulationsergebnisse verschiedener Szenarien folgen in Abschnitt 5.2.

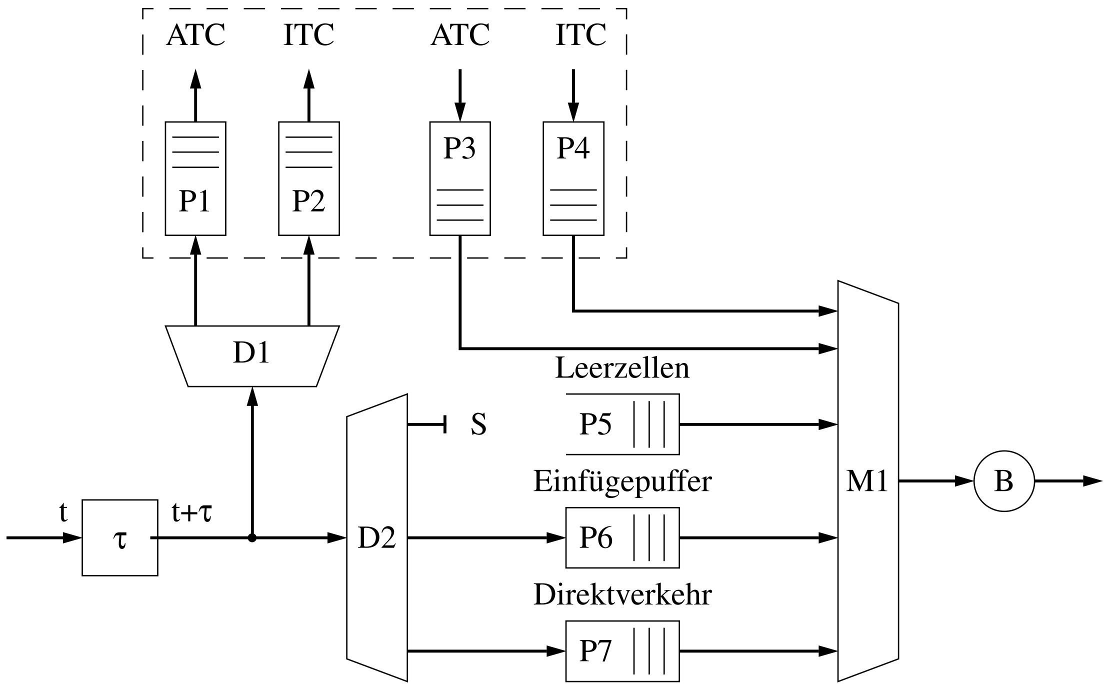
5.1 Simulationsmethode

Die Simulation des vorgestellten Protokolls erfolgt mit einem Simulationsprogramm durch zeittreue, ereignisgesteuerte Simulation [62]. Dabei wird eine Simulationsbibliothek verwendet, die auf objektorientierten Methoden beruht [58]. Das zu simulierende lokale Netz mit seinen physikalischen Eigenschaften wird auf ein für die Simulation verwendbares Modell abstrahiert [76, 105]. Die Simulation erfolgt auf Zellebene, das heißt die Datenrate einer Quelle wird durch die Zellankunftsabstandsverteilung bestimmt. Zeitschlitze und Rahmen werden durch die vorgegebenen Zeitbeziehungen zwischen aufeinanderfolgenden Zellen, leere Zeitschlitze durch Leerzellen modelliert. Das Simulationsprogramm ist in den wesentlichen Netz- und Simulationsparametern konfigurierbar.

Bild 5-1 zeigt das Simulationsmodell einer Station. Auf dem Ring ankommende Zellen erfahren eine konstante Verzögerung am Eingang, um die Verzögerungszeiten von Vorgängen wie Adressvergleich oder Fehlererkennung zu berücksichtigen. Der Demultiplexer D1 leitet für die Station bestimmte Zellen an den entsprechenden Empfangspuffer weiter. Dort werden sie statistisch erfaßt und anschließend entfernt. Der Demultiplexer D2 kann Zellen entfernen, sie über den Puffer P7 direkt an den Ausgangsmultiplexer weiterleiten oder sie in den Einfügepuffer P6 übergeben. Der Ausgangsmultiplexer M1 wählt die nächste zu sendende Zelle gemäß dem Protokollalgorithmus aus. Die Sendebedieneinheit B schließlich stellt mit ihrer Bedienzeit die Zeitbeziehung zwischen zwei aufeinanderfolgenden Zellen her.

Der gesamte Ring im Simulationsmodell (Bild 5-2) besteht aus Stationen nach Bild 5-1 und Leitungen, die als Verzögerungsglieder modelliert sind. Die Summe aller Verzögerungszeiten im Ring wird so gewählt, daß sie ein Vielfaches der Rahmendauer von 125 µs beträgt.

Das Simulationsprogramm instanziiert zunächst die notwendigen Objekte und deren Beziehungen untereinander. Danach beginnt die Sendebedieneinheit B der Monitor- und Masterstation mit dem Aussenden von Leerzellen. Mit der Ankunft der ersten Zelle am Demultiplexer D2 ist die Initialisierungsphase beendet und die statistische Verkehrsgenerierung beginnt.



τ	Verzögerungsglied mit konstanter Verzögerungszeit τ (Adressvergleich, Fehlererkennung)
D1	Demultiplexer, Auswahl des Empfangspuffers
P1	Empfangspuffer für ATC-Zellen
P2	Empfangspuffer für ITC-Zellen
D2	Demultiplexer
S	Senke für zu entfernende Zellen
P3	Sendepuffer für ATC-Zellen
P4	Sendepuffer für ITC-Zellen
P5	Leerzellen-Vorrat mit unendlicher Anzahl Leerzellen
P6	Einfügepuffer für ATC-Zellen
P7	Puffer für Direktverkehr
M1	Ausgangsmultiplexer
B	Sendebedieneinheit

Bild 5-1: Simulationsmodell einer Station

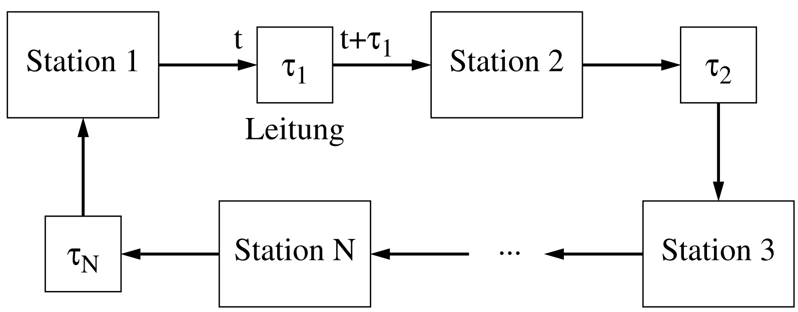


Bild 5-2: Simulationsmodell des Rings

5.2 Simulation des Gigabit-LAN

5.2.1 Symmetrisches Szenario

Zur Ermittlung des grundsätzlichen Systemverhaltens wird zunächst ein symmetrisches Szenario untersucht, bei dem sich der Verkehr jeder Station zu gleichen Teilen auf alle möglichen Zielstationen verteilt. Hierbei kann eine Obergrenze des erreichbaren Gesamtdurchsatzes abgeschätzt und ein Vergleich zu einem System mit exklusiver Bandbreitenreservierung gezogen werden. Es wird bei allen Simulationen eine negativ-exponentielle Verteilung der Zellankünfte zugrundegelegt.

5.2.1.1 Durchsatzbetrachtung

In einem symmetrischen System sind die Zielstationen gleichverteilt. Zur Ermittlung der Obergrenze des Durchsatzes wird angenommen, daß das System zu jeder Zeit voll ausgelastet ist. Der Durchsatz ist im Mittel für alle Verkehrsbeziehungen gleich. Bild 5-3 zeigt ein Beispiel mit 3 Stationen.

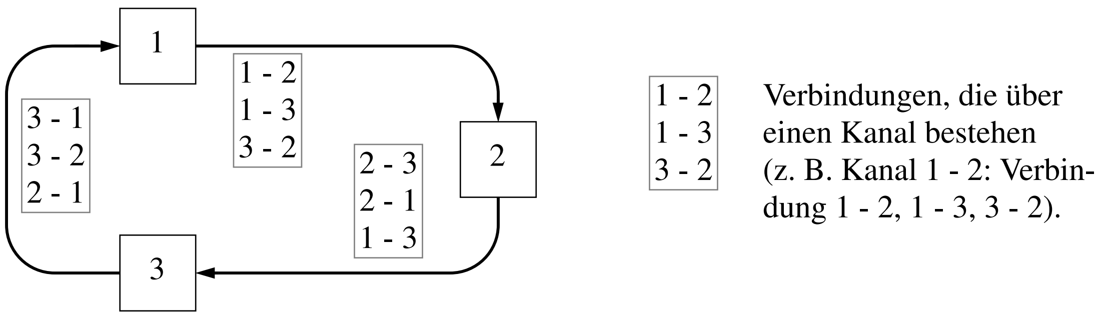


Bild 5-3: Symmetrisches Szenario

Die Gesamtzahl der Stationen sei S. Jede Station sendet an ihre S-1 Nachbarn mit gleichem Durchsatz. Damit ergibt sich eine Gesamtzahl aller Verkehrsbeziehungen von $V = S \cdot (S-1)$. Die Anzahl der Verbindungen L, die *über eine einzelne physikalische Übertragungsstrecke* bestehen, ist im Beispiel von Bild 5-3 mit S=3 Stationen L=3 oder allgemein

$$L = \sum_{i=1}^{S-1} i = \frac{S \cdot (S-1)}{2} = \frac{V}{2} \tag{3}$$

Für jede physikalische Übertragungsstrecke zwischen zwei Stationen wird die Übertragungsrate r_{link} auf L Verbindungen aufgeteilt. Für den maximalen Gesamtdurchsatz D_{gesamt} folgt

$$D_{gesamt} = V \cdot \frac{r_{link}}{L} = (S \cdot (S-1)) \cdot \frac{r_{link}}{\frac{S \cdot (S-1)}{2}} = 2 \cdot r_{link} \tag{4}$$

Wird ein Teil der Ringbandbreite für isochronen Verkehr reserviert, dann sinkt die Obergrenze für den Durchsatz des Asynchronverkehrs um die mittlere tatsächlich innerhalb der isochronen

Kanäle genutzte Bandbreite. Im symmetrischen Szenario hängt somit der Durchsatz des Asynchronverkehrs vom Anteil isochroner Kanäle und deren Auslastung ab.

Sei Z_{max} die Anzahl der Zeitschlitze je Rahmen und Z_{iso} die Anzahl der für isochronen Verkehr reservierten Zeitschlitze, so gilt für den Anteil isochroner Kanäle an der Übertragungsrate $\gamma = Z_{iso}/Z_{max}$. Der Durchsatz innerhalb der isochronen Kanäle D_{iso} ist bis zur Maximalauslastung näherungsweise proportional zur Ankunftsrate innerhalb der isochronen Verkehrsklasse A_{iso}, bei höherer Ankunftsrate konstant gemäß Gleichung 5.

$$D_{iso} = r_{link} \cdot \frac{Z_{iso}}{Z_{max}} \cdot \begin{cases} A_{iso} & A_{iso} < 1 \\ 1 & A_{iso} \geq 1 \end{cases} \tag{5}$$

Für den maximalen Durchsatz in der asynchronen Verkehrsklasse gilt folglich Gleichung 6.

$$D_{asyn} = D_{gesamt} - D_{iso} \tag{6}$$

Zum Vergleich ergibt sich in einem System, das Zeitschlitze während eines Ringumlaufs exklusiv für jeweils einen isochronen Kanal reserviert, ein Maximaldurchsatz für den Asynchronverkehr gemäß Gleichung 7.

$$D_{asyn'} = D_{gesamt} \cdot \left(1 - \frac{Z_{iso}}{Z_{max}}\right) \tag{7}$$

Bild 5-4 zeigt den Durchsatz der asynchronen Verkehrsklasse in Abhängigkeit von der Ankunftsrate in der isochronen Verkehrsklasse. Dabei wurde vorausgesetzt, daß die Hälfte der Zeitschlitze für isochrone Kanäle reserviert sind und die reservierten Zeitschlitze einer Station annähernd äquidistant innerhalb eines Rahmens verteilt sind. Verglichen werden die berechnete Obergrenze, das Simulationsergebnis[1] und ein System mit exklusiver Reservierung von Zeitschlitzen mit den Parametern nach Tabelle 5-1. Das Simulationsergebnis liegt sehr nahe an

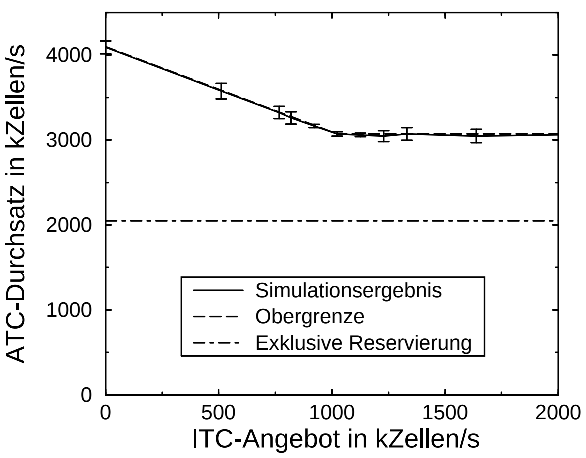


Bild 5-4: ATC-Durchsatz

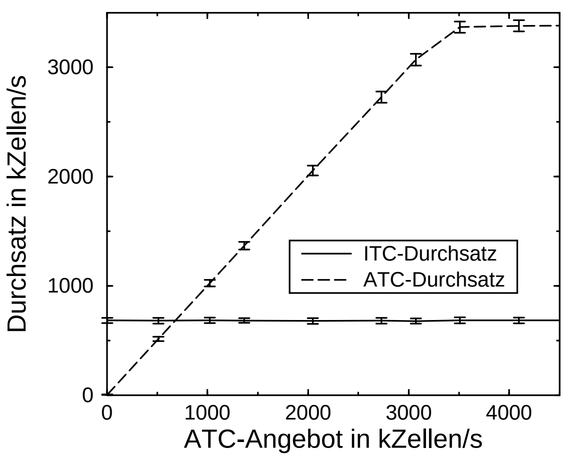


Bild 5-5: ITC-Durchsatz

1. Bei allen Simulationen werden 95%-Vertrauensintervalle und Verkehrsquellen mit negativ-exponentiellem Zellankunftsabstand verwendet

Parameter	Wert
Rahmendauer	125 µs
Zeitschlitze je Rahmen	256
Stationen	16
Rahmenzahl auf dem Ring	1
Zielverteilung	gleichverteilt
Anzahl reservierter Zeitschlitze	128
Fassungsvermögen der Sendepuffer	80 Zellen
Fassungsvermögen des Einfügepuffers	50 Zellen
Credits je Zyklus	500

Tabelle 5-1: Allgemeine Parameter

der berechneten Obergrenze, wobei sich bei den gewählten Parametern der Durchsatz gegenüber der exklusiven Reservierung verdoppeln kann. Bei Vollast innerhalb der isochronen Kanäle ergibt sich ein um den Faktor 1,5 höherer Durchsatz dadurch, daß die reservierten Zeitschlitze im Mittel nur auf dem halben Ring belegt sind und für den Rest des Ringumlaufs durch asynchronen Verkehr belegt werden können.

Der Durchsatz der isochronen Verkehrsklasse ist wie erwartet unabhängig von der Ankunftsrate in der asynchronen Verkehrsklasse, wie Bild 5-5 zeigt. Aufgetragen sind ITC- und ATC-Durchsatz in Abhängigkeit von der ATC-Ankunftsrate, wobei wiederum jeder zweite Zeitschlitz reserviert ist und die Ankunftsrate innerhalb der isochronen Verkehrsklasse mit 670 kZellen/s die reservierten Kanäle zu 2/3 auslastet. Der ATC-Durchsatz folgt der ATC-Ankunftsrate bis zur Auslastungsgrenze.

5.2.1.2 Wartezeit

Die Wartezeit bezeichnet die Zeitdauer zwischen dem Eintreffen einer Dateneinheit bei der Station und dem Sendebeginn auf dem Medium und berücksichtigt die Verzögerungen durch den Sendepuffer sowie die Zugriffsverzögerung durch das Warten auf einen nutzbaren Zeitschlitz.

Bild 5-6 zeigt die Wartezeit in der asynchronen Verkehrsklasse in Abhängigkeit von der Ankunftsrate in der isochronen Verkehrsklasse mit der ATC-Ankunftsrate als Parameter, wobei die Parameter nach Tabelle 5-1 und Abschnitt 5.2.1.1 zugrundeliegen. Die absoluten Zeiten liegen, solange der Ring nicht an seine Durchsatzgrenze stößt, im Bereich weniger Mikrosekunden. Eine größere Ankunftsrate in der isochronen Verkehrsklasse bedeutet weniger zur Verfügung stehende Bandbreite für die asynchrone Verkehrsklasse und damit eine erhöhte Wartezeit für den Asynchronverkehr. Die Begrenzung der Wartezeit ist durch die Begrenzung der Bandbreite für den ITC-Verkehr und dadurch einsetzende Verluste in der isochronen Verkehrsklasse begründet.

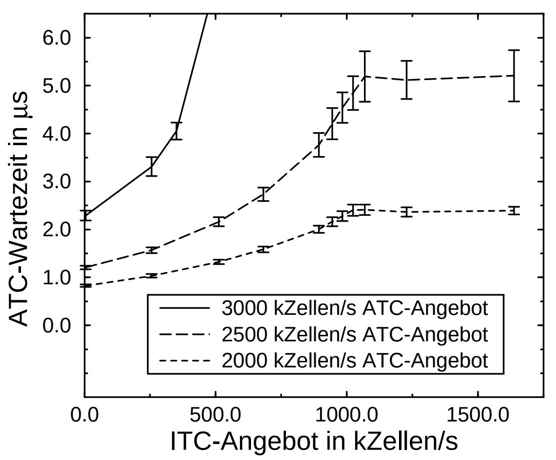


Bild 5-6: ATC-Wartezeit

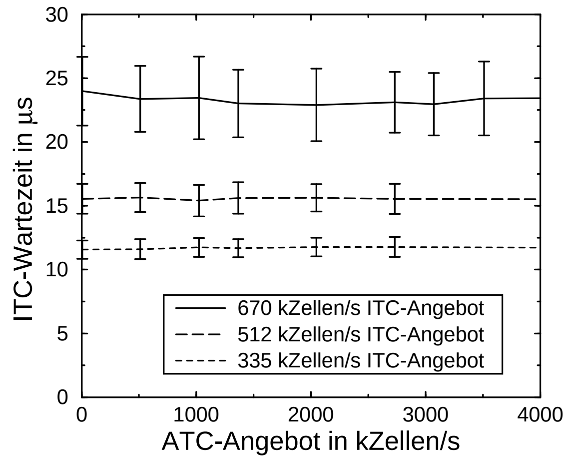


Bild 5-7: ITC-Wartezeit

Die Wartezeit in der isochronen Verkehrsklasse ist erwartungsgemäß nicht abhängig von der Ankunftsrate in der asynchronen Verkehrsklasse, wie Bild 5-7 zeigt. Da der ITC-Verkehr nur auf die reservierten Zeitschlitze zurückgreifen kann, ergibt sich eine erhöhte Wartezeit im wenig belasteten System gegenüber dem ATC-Verkehr. Nach den Werten von Tabelle 5-1 ist für eine Station jeder 32. Zeitschlitz reserviert. Daraus ergibt sich eine systembedingte mittlere Wartezeit auf den nächsten reservierten Zeitschlitz von 16 Zeitschlitzen oder 7,8 µs unter der Bedingung, daß keine Zelle im Sendepuffer wartet.

5.2.1.3 Dimensionierung des Credit Counters

Der Anfangswert des Credit Counter stellt eine Optimierungsaufgabe dar, bei der zwischen Durchsatz des Gesamtsystems und maximaler Wartezeit in der asynchronen Verkehrsklasse abzuwägen ist. Ein Maß für die maximale Wartezeit ist das Reset-Intervall, also die Zeitdauer zwischen zwei Resets im Fairness-Mechanismus. Durchsatz und Reset-Intervall in Abhängig-

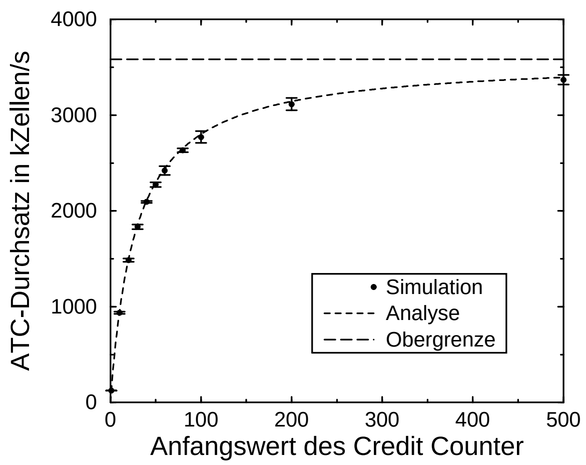


Bild 5-8: Durchsatzoptimierung

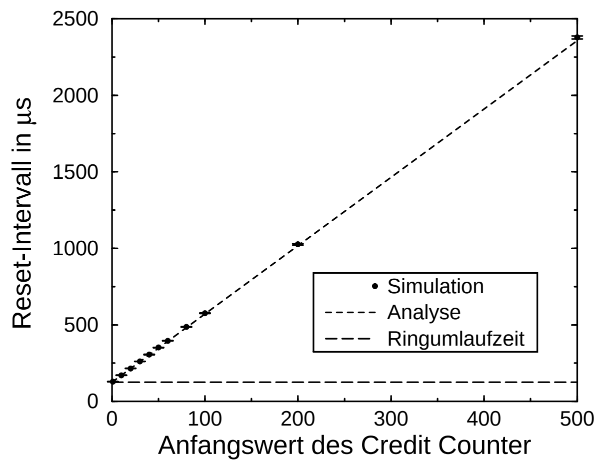


Bild 5-9: Reset-Intervall

keit vom Anfangswert des Credit Counter sind für ein symmetrisches System in Bild 5-8 und Bild 5-9 gegenübergestellt.

Simuliert wurde ein System in Überlast in der ATC-Verkehrsklasse mit den Parametern nach Tabelle 5-1, wobei der Anfangswert des Credit Counter variiert wurde. Die ITC-Ankunftsrate beträgt 512 kZellen/s, wodurch sich die in Bild 5-8 dargestellte obere Schranke für den ATC-Durchsatz von 3584 kZellen/s ergibt.

Das minimale Reset-Intervall entspricht der Dauer eines Ringumlaufs, minimal also einer Rahmendauer oder 125 µs.

Die Kurve für den ATC-Durchsatz D_{ATC} in Abhängigkeit des Credit Counter-Anfangswertes CCI bei symmetrischem ATC-Verkehr unter Überlast läßt sich herleiten. Station 1 von S sei die Monitor- und Masterstation. Unmittelbar nach Aussenden des Reset-Signals sendet sie ihr gesamtes Kontingent an Zellen. Anschließend wird „Active“ zurückgesetzt.

Station 2 erhält das Reset-Signal und mit diesem CCI Zellen von Station 1, von denen im Mittel $(CCI)/(S-1)$ für sie bestimmt sind. Die restlichen Zellen läßt sie passieren, weil sie im Einfügepuffer stehen. Ihre eigenen CCI Zellen fügt sie diesem Strom an, danach wird jeder Zeitschlitz als „Inactive“ weitergereicht. Insgesamt verlassen $2 \cdot CCI - ((CCI)/(S-1))$ Zellen Station 2. Nach Station i sind demnach Z_i Zellen gemäß Gleichung 8 auf dem Ring, nach der letzten Station Z_S Zellen gemäß Gleichung 9.

$$Z_i = \sum_{j=0}^{i-1} CCI\left(1 - \frac{j}{(S-1)}\right) = CCI\left(i - \frac{1}{(S-1)} \cdot \frac{i(i-1)}{2}\right) \tag{8}$$

$$Z_S = CCI\left(S - \frac{1}{(S-1)} \cdot \frac{S(S-1)}{2}\right) = CCI \cdot \frac{S}{2} \tag{9}$$

Nach 125 µs Rahmendauer kommt das Reset-Bit wieder bei Station 1 an und wird entfernt. Anschließend erreichen im Mittel Z_S Zellen Station 1 mit der maximalen Zellrate, die innerhalb der asynchronen Verkehrsklasse zur Verfügung steht. Danach hat keine Station mehr eine Sendeberechtigung und der nächste Zyklus beginnt. Das Reset-Intervall beträgt somit 125 µs zuzüglich der Dauer von Z_S Zellen. Bei einer Zellrate innerhalb der ATC von 1792 kZellen/s und 16 Stationen ergibt sich für das Reset-Intervall die in Bild 5-9 dargestellte Abhängigkeit von CCI gemäß der folgenden Gleichung:

$$t_R(CCI) = 125\mu s + \frac{CCI \cdot 8}{1792 \cdot 10^3} s = (125 + 4,46 \cdot CCI)\mu s \tag{10}$$

Hieraus läßt sich der ATC-Durchsatz in Abhängigkeit von CCI bestimmen. In einem Reset-Intervall werden $S \cdot CCI$ Zeitschlitze gesendet. Der ATC-Durchsatz in Abhängigkeit von CCI beträgt somit:

$$D_{ATC}(CCI) = \frac{S \cdot CCI}{(125 + 4,46 \cdot CCI) \cdot 10^{-3}} \frac{\text{kZellen}}{\text{s}} \tag{11}$$

Da der zeitkritische Verkehr in der isochronen Verkehrsklasse übertragen wird, ist die Dimensionierung des Credit Counter-Anfangswerts unkritisch, solange sich nur einer oder wenige Rahmen auf dem Ring befinden, wovon in einem lokalen Netz mit seiner beschränkten Ausdehnung auszugehen ist. Bei einem CCI von 500 ergibt sich ein Durchsatz von 3394 kZellen/s, entsprechend 95% des Maximaldurchsatzes. Das Reset-Intervall beträgt dann 2,36 ms.

5.2.2 Asymmetrischer ITC-Verkehr

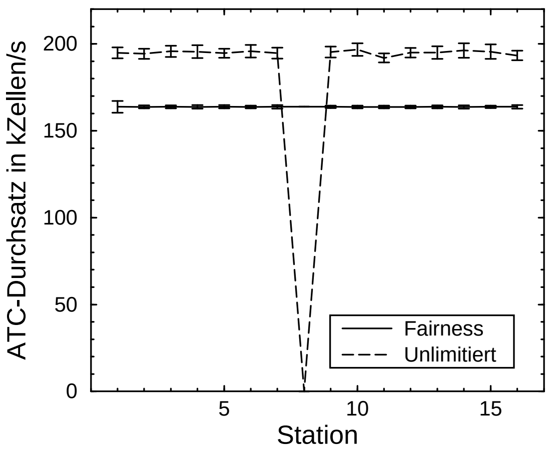
Ein asymmetrisches Verkehrsprofil in der isochronen Verkehrsklasse, bei dem die ITC-Ankunftsrate nicht bei allen Stationen gleich ist, hat Auswirkungen auf den Asynchronverkehr. Als Extrembeispiel kann ein Ring angesehen werden, bei dem nur eine einzige Station ITC-Verkehr sendet, während gleichzeitig alle Stationen symmetrisch Asynchronverkehr senden. Wenn die Zielverteilung der isochronen Zellen dieser einen Station einer Gleichverteilung entspricht, steigt die Kanalkapazität für den Asynchronverkehr nach dieser Station kontinuierlich von Station zu Station, da der Isochronverkehr abnimmt. Die Auswirkungen dieser Asymmetrie werden nachfolgend untersucht.

Betrachtet wird ein Ring aus 16 Stationen, wobei alle Stationen Asynchronverkehr senden. Die Parameter wurden, abgesehen vom ITC-Verkehr, wie in Abschnitt 5.2.1 gewählt. Der Einfügepuffer von Station 8 wurde auf 256 Zellen vergrößert. Station 8 ist die einzige Station, die ITC-Verkehr sendet. Für diesen Isochronverkehr ist die Hälfte der Bandbreite reserviert, wobei die Ankunftsrate einem Drittel der Ringbandbreite, insgesamt also 680 kZellen/s entspricht. Dieser Verkehr ist auf alle möglichen Zielstationen gleichmäßig verteilt.

Bild 5-10 zeigt den Durchsatz in der asynchronen Verkehrsklasse aufgetragen über den Stationen. Zur übersichtlicheren Darstellung sind die einzelnen Werte miteinander verbunden. Der Durchsatz in der asynchronen Verkehrsklasse bei Überlast[1] ist für alle Stationen gleich. Hier wirkt sich das Fairness-Protokoll aus, das eben diese Wirkung erzielen soll und die Stationen gleichmäßig drosselt. Wird jedoch der Fairness-Mechanismus abgeschaltet, ergibt sich ein insgesamt höherer ATC-Durchsatz, allerdings kann Station 8 keinen ATC-Verkehr mehr senden, da ihr Einfügepuffer nie leer wird und dieser immer Vorrang vor ihrem Sendepuffer hat.

Auch bei Anwendung des Fairness-Mechanismus hat die Asymmetrie Auswirkungen auf die Wartezeit der ATC-Zellen, wie in Bild 5-11 gezeigt. Aufgetragen ist der Mittelwert der Wartezeit aller Stationen bei einer Gesamtankunftsrate von 2000 kZellen/s in der asynchronen Verkehrsklasse, bei der das System noch nicht in Überlast betrieben wird. Erwartungsgemäß tritt bei Station 8 die größte Wartezeit auf, da zwischen Station 8 und Station 9 die verbleibende Kanalkapazität für den Asynchronverkehr am geringsten ist. Entsprechend ist der Kanal zwischen Station 7 und Station 8 frei von ITC-Verkehr und somit, sofern keine Zeitschlitze „unavailable“ sind, vollständig für ATC-Verkehr nutzbar.

1. Die ATC-Sendepuffer sind nie leer



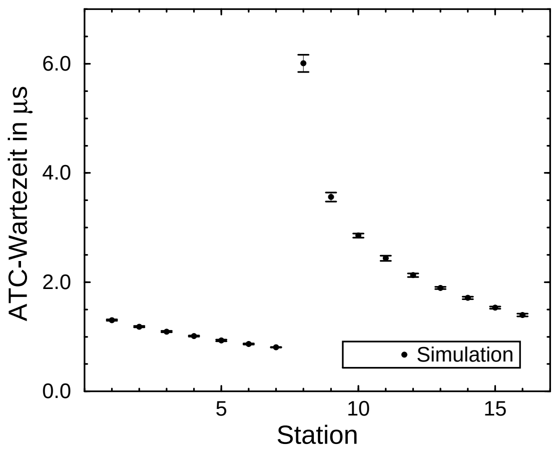
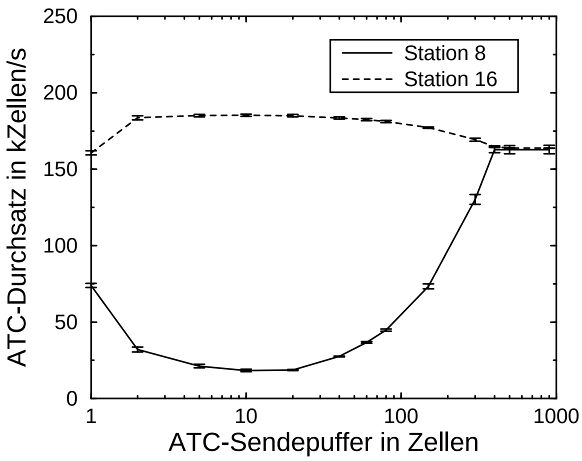


Bild 5-10: Durchsatz (ITC nur von Station 8) **Bild 5-11:** Wartezeit (ITC nur von Station 8)

5.2.2.1 Dimensionierung der ATC-Sendepuffer

Durch die Asymmetrie wirkt sich die Größe des ATC-Sendepuffers auf den Durchsatz der Station 8 im Vergleich zu anderen Stationen aus. Dabei liegt die Gesamtankunftsrate in der asynchronen Verkehrsklasse bei 3277 kZellen/s, also beim 1,6-fachen der Ringbandbreite, womit das System an der Grenze zur Überlast in der asynchronen Verkehrsklasse betrachtet wird.

Bild 5-12 zeigt die Abhängigkeit des ATC-Durchsatzes der Stationen 8 und 16 in Abhängigkeit von der Größe des ATC-Sendepuffers. Bei sehr kleinen Sendepuffern kommt es in allen Stationen zu Verlusten, so daß der Ring weniger stark ausgelastet ist. Der Durchsatz von Station 8 nimmt daher zunächst mit wachsendem Sendepuffer ab und liegt trotz des Fairness-Mechanismus unterhalb dem der übrigen Stationen. Da der Einfügepuffer Vorrang vor dem Sendepuffer hat, kann Station 8 bei Überlast erst am Ende des Reset-Zyklus Zellen aus ihrem Sendepuffer senden. Dieser wird mit einer Rate von 1365 kZellen/s geleert, während die Füllrate 205 kZellen/s beträgt. Der Sendepuffer läuft also sehr schnell leer, noch bevor die Station ihr



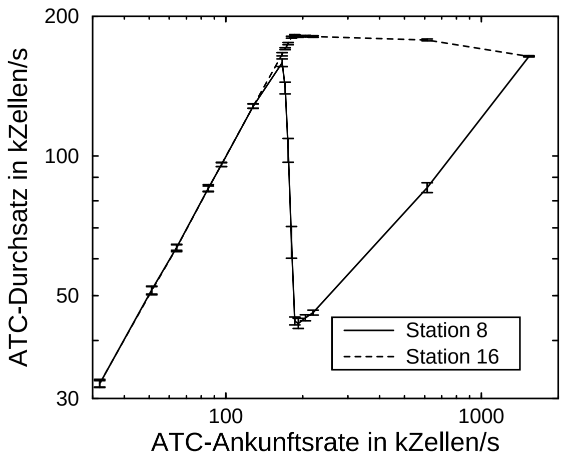


Bild 5-12: Sendepuffer-Dimensionierung **Bild 5-13:** ATC-Durchsatz bei Überlast

Kontingent von 500 Zellen senden konnte. Sobald der Sendepuffer aber keine Zelle mehr enthält, ist die Station inaktiv, was zu einem neuen Credit-Zyklus und damit zu einer dauerhaften Benachteiligung der Station 8 führt.

Die Bandbreite, die Station 8 aufgrund des beschriebenen Effektes nicht nutzen kann, steht den anderen Stationen zur Verfügung. Bei weiter steigender Größe des Sendepuffers nähern sich die Durchsatzkurven der Stationen einander an. Spätestens bei einem Sendepuffer mit einem Fassungsvermögen von etwa 500 Zellen, also dem Kontingent einer Station für einen Zyklus, ist der ATC-Durchsatz aller Stationen identisch.

Die Auswirkungen des zu gering dimensionierten Sendepuffers werden auch deutlich, wenn man den stationsspezifischen Durchsatz in Abhängigkeit von der Ankunftsrate betrachtet. Wie in Bild 5-13 dargestellt, bricht der Durchsatz der Station 8 bei beginnender Überlast stark ein und steigt dann bei zunehmender Ankunftsrate wieder an. Erreicht die Ankunftsrate die Medienrate von 2048 kZellen/s, dann ist diese durch den zu klein dimensionierten Sendepuffer verursachte Unfairness beseitigt, da der Sendepuffer nicht mehr leer wird.

5.2.2.2 Dimensionierung des Einfügepuffers

Die verfügbare Bandbreite für den ATC-Verkehr hängt davon ab, inwieweit für ITC-Verkehr reservierte Zeitschlitze für Asynchronverkehr freigegeben werden. Die Größe des Einfügepuffers in Station 8 hat direkten Einfluß auf die Anzahl der freigegebenen („Available“) Zeitschlitze. Der ATC-Durchsatz steigt folglich mit der Größe des Einfügepuffers an. Liegt das Fassungsvermögen des Sendepuffers (200 Zellen) unter dem Anfangswert des Credit Counters (500 Zellen), so ergibt sich die in Bild 5-14 dargestellte unterschiedliche Abhängigkeit des ATC-Durchsatzes der Stationen 8 und 16 von der Größe des Einfügepuffers. Während Station 16 den erwarteten Anstieg zeigt, verläuft die Durchsatzkurve der Station 8 bei kleinem Einfügepuffer ähnlich, bricht jedoch bei einer Grenze von etwa 45 Zellen deutlich ein, so daß sich erst hier der zu kleine Sendepuffer auswirkt.

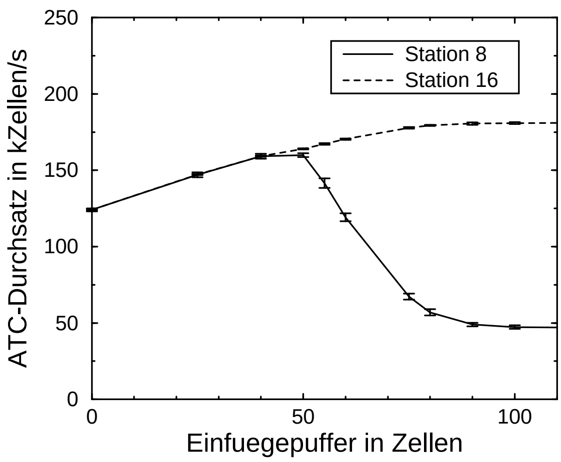


Bild 5-14: Einfügepuffer-Dimensionierung

Der Effekt ist dadurch zu erklären, daß bei sehr kleinem Einfügepuffer zu viele Zeitschlitze als „unavailable" markiert sind, die somit leer wieder bei Station 8 ankommen, nicht für ITC-Verkehr gebraucht werden und in der Folge dazu verwendet werden können, den Einfügepuffer zu leeren. Insgesamt sind in dem einen Rahmen auf dem Ring 128 Zeitschlitze für Station 8 reserviert, von denen sie im Mittel 85 in einem Umlauf nutzt, während 43 reservierte Zeitschlitze ungenutzt bleiben. Ein Einfügepuffer mit weniger als 43 Pufferplätzen verhindert, daß mehr als die 43 für ITC-Verkehr nicht genutzten Zeitschlitze für ATC-Verkehr freigegeben werden, so daß der Einfügepuffer ständig annähernd leer bleibt. Bei Überschreiten der Grenze wird anfangs mehr Bandbreite freigegeben, als verfügbar ist, so daß der Füllstand des Einfügepuffers rasch ansteigt und erst wieder am Ende eines Credit-Zyklus auf Null zurückgehen kann.

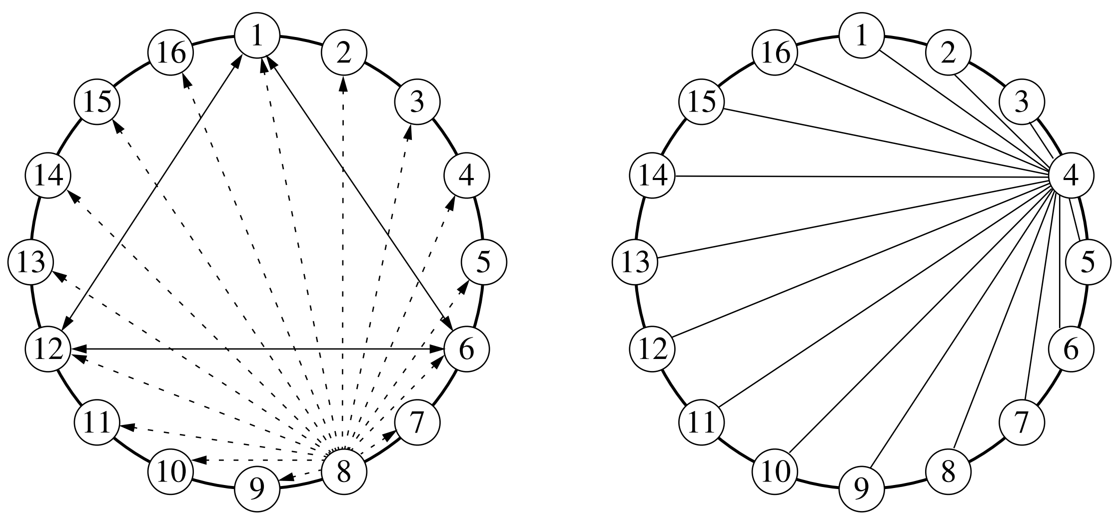
5.2.3 Heterogenes Szenario

Ein heterogenes Szenario, das mehrere Dienste in einem zukünftigen lokalen Netz beinhaltet, ist in Bild 5-15 dargestellt. Station 4 stellt einen Server dar, der die übrigen 15 Stationen bedient. Die Asynchronverbindungen sind somit als bidirektionale Verbindungen zwischen einer Station i ($i \neq 4$) und Station 4 unter Verwendung von Verkehrsquellen mit negativ-exponentieller Zellankunftsabstandsverteilung modelliert. Die Ankunftsrate für alle Asynchronverbindungen ist gleich. Der Anfangswert des Credit Counter von Station 4 beträgt 7500, ist also um einen Faktor 15 größer als der Anfangswert der übrigen Stationen.

ITC-Verbindungen bestehen für eine Dreierkonferenz zwischen Station 1, 6 und 12 als Videokonferenz mit hoher Datenrate, sowie von Station 8 als Video-Server zu allen anderen Statio-

Verkehrsbeziehungen isochron

Verkehrsbeziehungen asynchron



Video-Verbindung vom Video-Server (Station 8) zu allen anderen Stationen; Videokonferenz mit hoher Bitrate zwischen Station 1, 6 und 12.

Station 4 ist Server-Station, direkte Verkehrsbeziehungen zwischen den übrigen Stationen werden vernachlässigt.

Bild 5-15: Heterogenes Beispielszenario nen bei geringerer Datenrate. Für einen Kanal vom Video-Server zu einem Video-Terminal ist eine Bandbreite von 27 kZellen/s[1] reserviert, wobei von einem komprimierten Datenstrom mit einer mittleren Bitrate von 5 Mbit/s ausgegangen wird [20]. Für die Dreierkonferenz ist eine Bandbreite von $3 \cdot 408$ kZellen/s reserviert, was für unkomprimierte Videosignale mit jeweils 140 Mbit/s ausreicht.

Die Belegung der Kanäle mit ITC-Verkehr jeweils am Ausgang der Stationen zeigt Bild 5-16. Vor dem Video-Server, Station 8, ist die Auslastung mit ITC-Verkehr am geringsten, nach Station 8 am höchsten. Die Teilnehmer an der Video-Konferenz unterscheiden sich hier nicht von den Stationen ohne ITC-Verkehr, da eine Station genau so viele ITC-Zellen entfernt wie sie sendet. Die Videokonferenz ergibt einen gleichmäßigen Belag von ITC-Zellen auf dem Ring, der somit lediglich eine Reduzierung der für ATC-Verkehr zur Verfügung stehenden Bandbreite zur Folge hat.

Damit unterscheidet sich das Verhalten der Vidokonferenzteilnehmer auch in der asynchronen Verkehrsklasse nicht von den Stationen ohne ITC-Verbindungen, wie Bild 5-17 bestätigt. Solange der Ring nicht voll ausgelastet ist, erhält Station 4, die Server-Station, den erwarteten 15-fachen Durchsatz gegenüber den übrigen Stationen, deren Durchsatz gleich ist. Bei Überlast ist der Durchsatz von Station 4 nach wie vor 15 mal höher als der mittlere Durchsatz der übrigen Stationen, aber der Durchsatz von Station 8, dem Video-Server, fällt deutlich ab. Diese Benachteiligung von Stationen, die mehr ITC-Verkehr senden als empfangen, wurde bereits in Abschnitt 5.2.2.1 begründet. Bei einem Video-Server ist jedoch das Verhalten in der asynchronen Verkehrsklasse unerheblich, weil darin lediglich nicht zeitkritische Signalisierinformation von geringem Umfang übertragen wird.

Das ATC-Wartezeitprofil in Bild 5-18 zeigt eine nahezu identische mittlere Wartezeit aller Stationen bei geringer Auslastung, da für ankommende Zellen sehr schnell ein freier Zeitschlitz verfügbar ist. Bei hoher Auslastung ist die Wartezeit der Station 4 geringer als die aller ande-

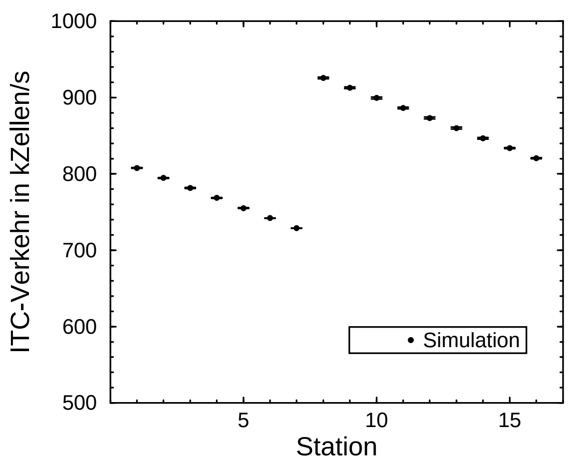


Bild 5-16: ITC-Auslastung

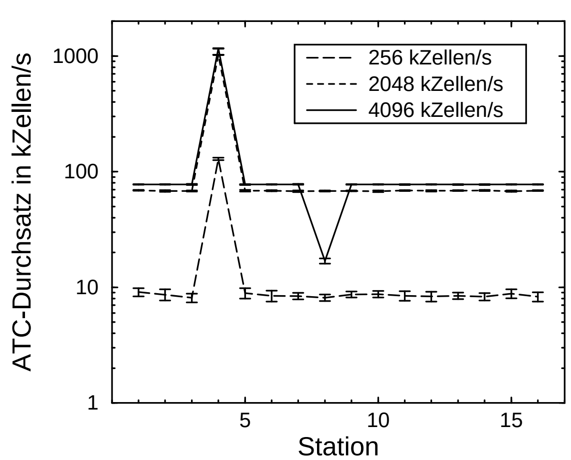


Bild 5-17: ATC-Durchsatz

1. Entspricht 10,4 Mbit/s

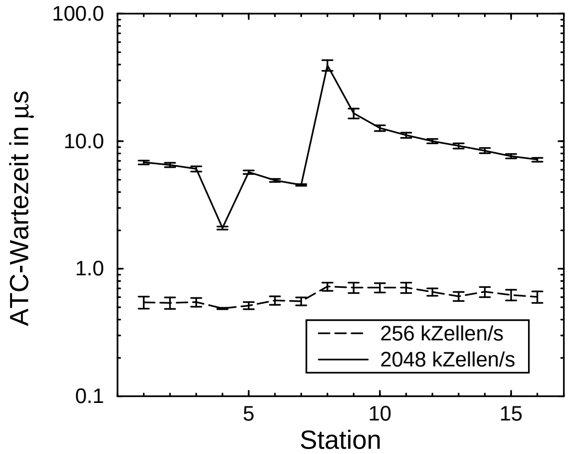


Bild 5-18: ATC-Wartezeit

ren Stationen, da die Server-Station häufiger ATC-Zellen entfernt als die übrigen Stationen und damit der Abstand der nutzbaren Zeitschlitze geringer ist. Station 8 hat die größte mittlere Wartezeit, die bei den folgenden Stationen jeweils absinkt, wie bereits in Abschnitt 5.2.2 erläutert wurde.

Zusammenfassend zeigt sich, daß das Protokoll für ein lokales Netz mit heterogenem Verkehr sehr gut geeignet ist, da die zur Verfügung stehende Bandbreite optimal ausgenutzt wird. Bei sehr hoher Auslastung entsteht eine geringe Unfairness hinsichtlich der mittleren Wartezeit in der asynchronen Verkehrsklasse, die bedingt ist durch die Belegung des Rings mit ITC-Zellen. Die absolute Wartezeit ist jedoch gering, und die Alternative wäre lediglich, die Bandbreite für ATC-Verkehr zu sperren und somit den Gesamtdurchsatz zu senken und die mittlere Wartezeit auf einen für alle Stationen gleichen Wert anzuheben. Zudem sollte ein lokales Netz, das dauerhaft an der Grenze seiner Belastbarkeit betrieben wird, durch ein leistungsfähigeres Netz ersetzt werden.

Kapitel 6

Funktionelle Bewertung des Protokollvorschlags

6.1 Realisierungsaspekte

Der Hardware-Aufwand für eine Station des in Kapitel 4 vorgeschlagenen Gigabit-LANs kann anhand einer prototypischen Implementierung abgeschätzt werden. Hierfür wird ein System betrachtet, das die wesentlichen Elemente einer Station enthält. Protokollprozessor, PCI-Busschnittstelle sowie die physikalischen Empfangs- und Sendeinterfaces sind Stand der Technik und als integrierte Bausteine oder Module verfügbar.

Bild 6-1 zeigt die Kernarchitektur der prototypischen Implementierung. Jede der 5 Einheiten in Bild 6-1 sowie die physikalischen Sende- und Empfangsinterfaces werden auf jeweils einer Baugruppe realisiert, die über ein Bussystem miteinander kommunizieren. Ziel der prototypischen Implementierung ist die effiziente Realisierung eines LAN-Knotens unter Verwendung von marktgängigen Bauelementen. Kombinatorische Logik, Register und Steuerwerke werden

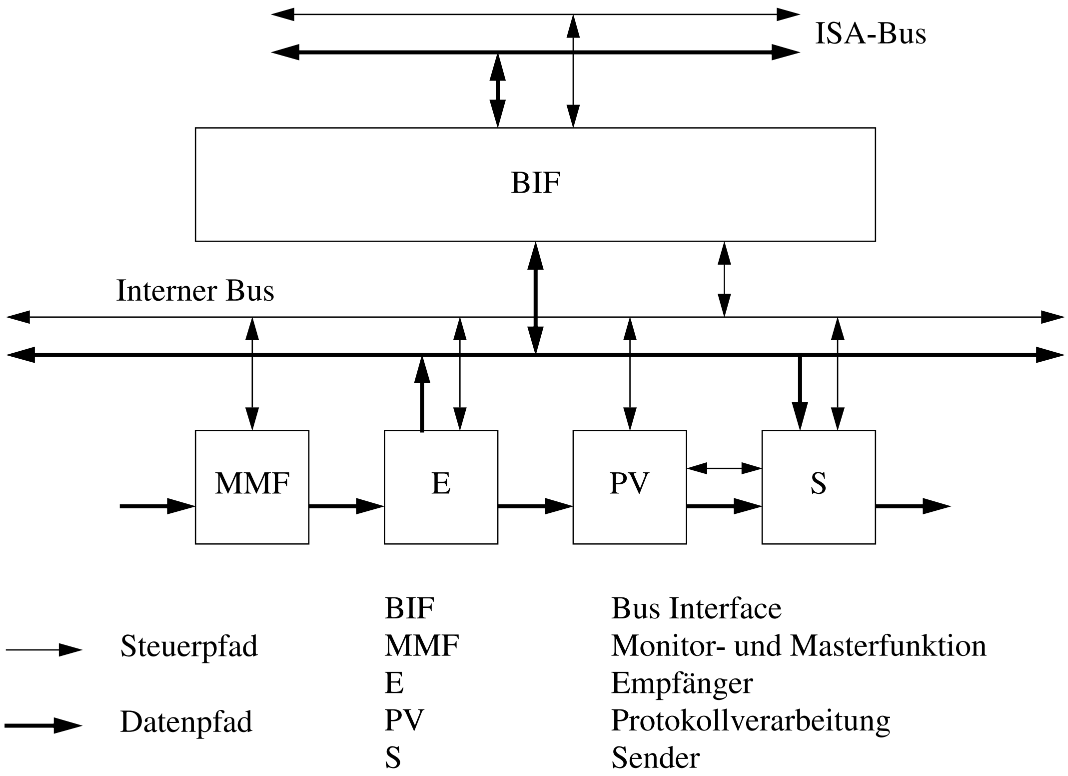


Bild 6-1: Kernarchitektur der Beispielimplementierung

dabei mit Hilfe programmierbarer Logikbausteine realisiert, deren kleinste Einheit, eine Logikzelle, vier Eingänge beliebig logisch verknüpft und am Ausgang über ein Flipflop verfügt [2]. Tabelle 6-1 listet die wesentlichen Standardbauelemente sowie die Anzahl der benötigten Logikzellen für die fünf Baugruppen nach Bild 6-1 auf.

Baugruppe	Adressierbarer Speicher (RAM)		Puffer-speicher (FIFO)	Logik-zellen	Flipflops
	statisch	dynamisch			
Monitor- und Masterfunktion	256 bit		4Kx32 bit	913	612
Empfänger		2 Mbit	2x4Kx32 bit	1086	863
Protokollverarbeitung	256 bit		4Kx32 bit	1021	685
Sender			2x4Kx32 bit	772	483
Bus Interface				206	63
Summe	512 bit	2 Mbit	24Kx32 bit	3998	2706

Tabelle 6-1: Aufwand für die prototypische Implementierung

Insgesamt werden etwa 4000 Logikzellen und 2700 Flipflops benötigt. Der leistungsfähigste Baustein innerhalb der verwendeten Bausteinfamilie verfügt über 130000 Gatter, stellt inzwischen 6656 Logikzellen und 7126 Flipflops zur Verfügung und enthält darüber hinaus 32 Kbit RAM [3], das die externen statischen RAMs vollständig ersetzen könnte. Ein Baustein mit 250000 Gattern ist für 1998 angekündigt. Die gesamten Logikfunktionen können somit bei der Integration des Gesamtsystems auf einer Einschubkarte in einem Baustein realisiert werden. Eine Übertragung der Logik- und Registerfunktionen auf ein Gate Array mit Hilfe eines Synthesewerkzeugs und der Bibliothek eines Halbleiterherstellers ergab einen Bedarf von etwa 10000 Gattern. Somit können auch die statischen RAMs und ein Teil der Pufferspeicher, insbesondere der Einfügepuffer, in ein Gate Array integriert werden.

6.2 Systemvarianten

Das in Kapitel 4 beschriebene Medienzugriffsprotokoll für Gigabit-LANs ist skalierbar: Es kann hinsichtlich der Kosten oder der Flexibilität und Leistungsfähigkeit optimiert werden. Das Protokoll kann ferner derart vereinfacht werden, daß es für Steuergerätenetze geeignet ist. Im folgenden werden die Variationsmöglichkeiten nicht vollständig beschrieben. Vielmehr soll aufgezeigt werden, wie das Protokoll an unterschiedliche Anforderungen angepaßt werden kann.

6.2.1 Kostensenkende Maßnahmen

Stationen, die keine isochronen Kanäle zum Senden benötigen, können wesentlich vereinfacht werden, indem der für die isochrone Übertragung erforderliche Hardware-Aufwand eingespart wird: Der ITC-Sendepuffer entfällt ebenso wie der Available Counter und das RAM für die Speicherung der Nummern reservierter Zeitschlitze. Der Einfügepuffer kann entfallen, wenn die Station nicht mehrere Zeitschlitze hintereinander für ein ATC-Paket belegen können muß. Insgesamt wird der Speicheraufwand damit deutlich reduziert. Trotzdem ist die Station in der Lage, Verkehr in isochronen Kanälen zu empfangen, um zum Beispiel an Video-Verteildiensten teilzunehmen. Wird auch diese Funktionalität nicht benötigt, kann noch der ITC-Empfangspuffer entfallen, was den Pufferbedarf weiter reduziert.

Die Monitor- und Master-Funktionalität ist in jeder Station vorgesehen, um eine größtmögliche Ausfallsicherheit des Netzes zu gewährleisten. Bei einer Aufteilung in zwei Stationsklassen – mit und ohne Monitor- und Masterfunktionalität – muß nicht jede Station die volle Funktionalität vorhalten, was den Gesamtaufwand des Netzes reduziert. Nicht nur der Hardware-Aufwand in den Stationen wird verringert, sondern auch die Signalisierungs- und Verbindungssteuerungssoftware muß nur in den Stationen mit voller Funktionalität vorgehalten werden. Ausfallsicherheit ist dennoch gewährleistet, wenn mehrere Stationen die Monitor- und Masterfunktionalität bereithalten.

Der AAL-Protokollprozessor ist optional. Kommt kein Protokollprozessor zum Einsatz, dann muß das Aufteilen und Vereinigen von AAL-5-CPCS-PDUs im angeschlossenen System in Software erfolgen, was bei hohem Verkehrsaufkommen eine starke Belastung für das System darstellt. Ein System, das auf niedrige Kosten optimiert sein muß, benötigt jedoch möglicherweise nur einen geringen Durchsatz bei einer sehr beschränkten Anzahl von Verbindungen, so daß die AAL-Implementierung mit einem Protokollprozessor die Mehrkosten nicht rechtfertigt.

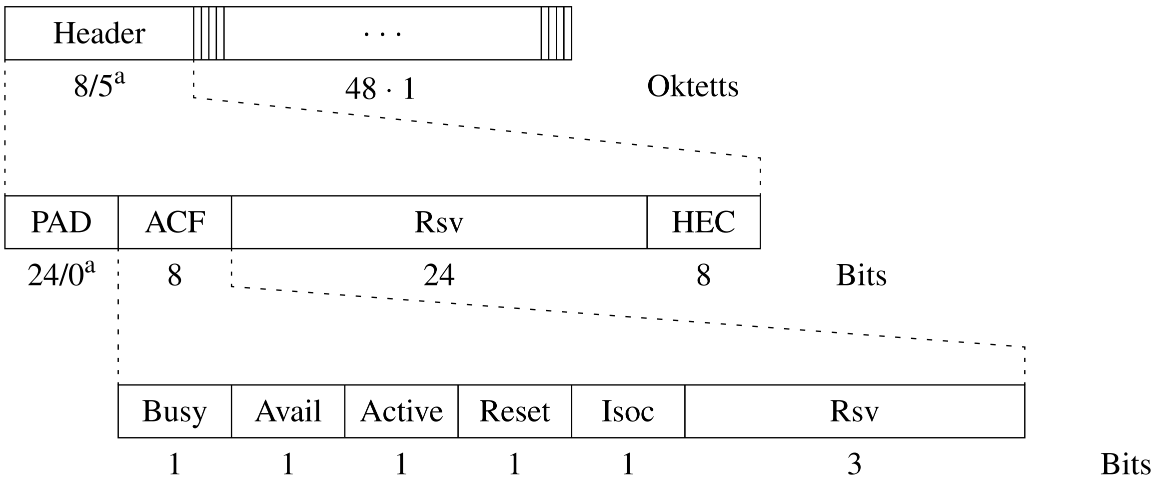
6.2.2 Erweiterungen

6.2.2.1 Integration von Telefonkanälen

Für die Integration von Telefonkanälen mit jeweils 64 kbit/s[1] kann in einem System mit sehr wenig Telefonverkehr jeweils ein Isochronkanal, also ein Zeitschlitz je Rahmen reserviert werden, wobei entweder alle 125 µs eine Zelle mit jeweils einem Oktett Inhalt gesendet wird oder die Oktetts zusammengefaßt und alle 6 ms in einer Zelle übertragen werden. Das erste Verfahren bietet eine extrem schlechte Ausnutzung des Isochronkanals mit etwa 2%, während beim zweiten Verfahren eine zusätzliche Verzögerung von 6 ms eingefügt wird.

Ein effizienteres Verfahren zur Integration von Telefonkanälen, insbesondere in einem System mit sehr viel Telefonverkehr, ist die Unterteilung von Zeitschlitzen gemäß Bild 6-2. Jedes Oktett im Datenfeld wird einem von 48 Kanälen zugeordnet. Die Monitor- und Masterstation reserviert die benötigte Anzahl von Zeitschlitzen in jedem Rahmen und setzt die Bits „Busy", „Avail" und Isoc" entsprechend Bild 6-2. Durch die Markierung als „Unavailable" ist sichergestellt, daß keine Station den Zeitschlitz mit einer Zelle belegt. Der Monitor-Mechanismus entfernt die Daten nicht, da das Busy-Bit nicht gesetzt ist.

1. Entsprechend der Bitrate eines ISDN-B-Kanals



PAD	Padding, Füllbits ohne Funktion
ACF	Access Control Field
Rsv	Reserved
HEC	Header Error Check
Busy	Busy-Bit, ist auf „Free“, also nicht belegt gesetzt
Avail	Available Bit, ist auf „Unavailable“ gesetzt
Active	Active Node Bit
Reset	Reset-Bit
Isoc	Isochronous Cell, ist gesetzt

a. siehe Abschnitt 4.3.3

Bild 6-2: Zeitschlitzformat für die Integration schmalbandiger Isochronkanäle

Ein derartiger Schmalband-Isochronkanal wird einer Station bei Bedarf von der Monitor- und Masterstation zugeteilt und wird charakterisiert durch die Nummer des Zeitschlitzes im Rahmen sowie die Nummer des Oktetts innerhalb des Zeitschlitz-Datenfelds. Schmalband-Isochronkanäle sind bidirektional, werden also während eines Ringumlaufs nacheinander von beiden Kommunikationspartnern genutzt.

Zum Senden und Empfangen von Daten in einem Telefonkanal muß eine Station zusätzlich über einen Mechanismus einschließlich der dafür notwendigen Hardware verfügen, ein oder mehrere einzelne Oktetts an durch die Verbindungssteuerungs-Software konfigurierbaren Stellen in den Datenstrom zu multiplexen und zu demultiplexen.

6.2.2.2 Erweiterung des direkten Adressbereichs

Die in Abschnitt 4.3.8 beschriebene direkte Stationsadressierung verwendet Quell- und Zieladresse, um jede mögliche Paarbeziehung[1] zwischen den maximal 255 adressierbaren Stationen durch einen eindeutigen VCI-Wert zu kennzeichnen. Dies ist notwendig, wenn Zellen

1. Insgesamt sind bei 255 Stationen 255^2 gerichtete Paarbeziehungen möglich

beim Empfänger in beliebiger Reihenfolge ankommen können. In diesem Fall muß der logische Empfangspuffer, in dem die Zellen zu einer AAL-5-PDU vereinigt werden, eindeutig identifizierbar sein.

Wenn sichergestellt ist, daß alle Stationen nur komplette AAL-5-PDUs ohne Unterbrechung durch andere ATC-Zellen senden (siehe 4.3.7), braucht nur noch ein logischer Empfangspuffer bereitgestellt zu werden, da die Zuordnung der ankommenden Zellen zu einer AAL-5-PDU eindeutig ist. Damit kann die Quelladresse bei der direkten Adressierung entfallen, wodurch die Breite der Zieladresse auf 15 bit und damit die Anzahl der adressierbaren Stationen auf 2^{15} erhöht wird. Ein Bit wird benötigt, um einen Broadcast anzuzeigen. Bild 6-3 zeigt das VCI-Format für diese Art der Adressierung.

VCI-Typ	BC	Adresse	
4	1	15	Bits

BC Broadcast-Bit
Adresse BC nicht gesetzt: Zieladresse; BC gesetzt: Quelladresse

Bild 6-3: Aufteilung des VCI für direkte Adressierung

Ist das Broadcast-Bit nicht gesetzt, so ist die Zelle Teil einer Punkt-zu-Punkt-Verbindung und die Adresse als Zieladresse zu interpretieren. Die Zielstation entfernt auch die Zelle. Bei gesetztem Broadcast-Bit wird die Zelle von allen Stationen empfangen, jedoch nicht entfernt. Die sendende Station, die ihre eigene Adresse in das Adressfeld geschrieben hat, entfernt die Zelle.

6.2.2.3 Topologie-Varianten

Neben dem Ring gibt es weitere, davon abgeleitete Topologien, für die das in Kapitel 4 vorgestellte Protokoll einsetzbar ist. Die wichtigste Variante ist der Doppelring, da er den Gesamtdurchsatz gegenüber dem einfachen Ring um einen Faktor zwei bis vier erhöht (siehe Abschnitt 2.2.3) und zudem effizientere Fairness-Mechanismen ermöglicht, die teilweise auf eine Doppelringstruktur ausgelegt sind (siehe 6.2.3). Ein Doppelring erfordert einen annähernd verdoppelten Systemaufwand, wobei einzelne Teile der Stationsstruktur wie zum Beispiel der AAL-Protokollprozessor für beide Ringe gemeinsam genutzt werden können.

Zusätzlicher Verkehrslenkungsaufwand wird benötigt, um einen der Ringe für eine zu sendende Zelle auszuwählen. Neben einem einfachen Verkehrslenkungsverfahren, das den Ring nach einem festen oder zufälligen Prinzip auswählt, ist ein intelligenteres Verfahren denkbar, das von der Last auf einem Ring oder der Entfernung zum Empfänger abhängig ist. Damit kann der Gesamtdurchsatz weiter gesteigert werden.

Eine weitere Topologievariante ist die Verkabelung des Rings in einer Sternstruktur wie in Bild 6-4 dargestellt. Das Sternelement kann passiv sein und nur aus Kabeln bestehen, die die Ringstruktur herstellen. Damit ist es möglich, an zentraler Stelle durch Umstecken von Kabeln die

Reihenfolge der Stationen im Ring zu modifizieren. Ein aktives Sternelement kann den Ring dynamisch rekonfigurieren und so im laufenden Betrieb weitere Stationen in den Ring einfügen und ausgefallene oder ausgeschaltete Stationen aus dem Ring entfernen. Wenn das Sternelement zusätzlich die Monitor- und Masterfunktionalität übernimmt, kann die Rekonfiguration, die einen zu behebenden Rahmenfehler verursacht, in kürzestmöglicher Zeit abgeschlossen sein. Zudem kann die Monitor- und Masterfunktionalität in den Stationen eingespart werden.

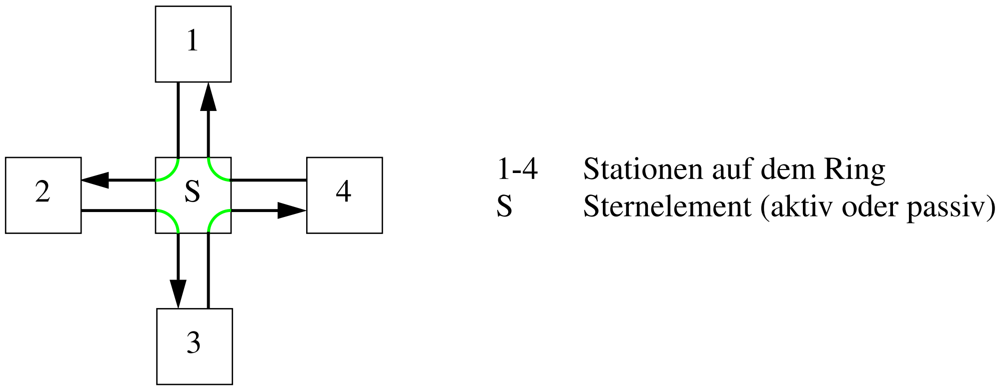


Bild 6-4: Ringnetz in Sternverkabelung

6.2.3 Erhöhung der Leistungsfähigkeit für ATC-Verkehr

Die Verzögerungszeit und deren Verteilung sowie der erreichbare Durchsatz innerhalb der asynchronen Verkehrsklasse sind abhängig vom gewählten Fairness-Mechanismus. Da für den zeitkritischen Verkehr isochrone Kanäle zur Verfügung stehen, wurde beim Fairness-Mechanismus für den Asynchronverkehr innerhalb des Protokollvorschlags der Schwerpunkt auf ein möglichst einfaches Verfahren und weniger auf hohe Effizienz gesetzt. Aufwendigere Fairness-Mechanismen können das gewählte Verfahren ersetzen, ohne die grundsätzliche Protokollstruktur in Frage zu stellen.

Eine Voraussetzung für einen Fairness-Mechanismus für den Asynchronverkehr ist, daß das Verfahren nicht von einer konstanten Kanalbitrate ausgehen darf, da die verfügbare Bandbreite für den Asynchronverkehr vom Isochronverkehr abhängt und weder zeitlich noch auf allen Kanälen zwischen den Stationen konstant oder gleich groß ist. Ein Reservierungsverfahren wie bei CRMA (siehe 2.4.5), das für einen Zyklus die Anzahl der reservierten Zeitschlitze abzählt und anschließend den nächsten Zyklus beginnt, ist somit ohne Modifikationen ungeeignet. Ein Scheduler kann nicht davon ausgehen, daß ein belegbarer („available“) Zeitschlitz für alle nachfolgenden Stationen belegbar bleibt: der Zeitschlitz kann für ITC-Verkehr reserviert sein und von einer nachfolgenden Station mit einer ITC-Zelle belegt werden, womit er nicht mehr für ATC-Verkehr zur Verfügung steht. Bild 6-5 verdeutlicht die Problematik.

Ein Beispiel für ein geeignetes Verfahren ist der Fairness-Mechanismus des Metarings (siehe Abschnitt 2.4.6), wobei allerdings ein Doppelring vorausgesetzt wird. Das „SAT“-Bit des Metarings ersetzt in diesem Fall das „Active“-Bit im Zeitschlitz-Header. Das „Reset“-Bit wird

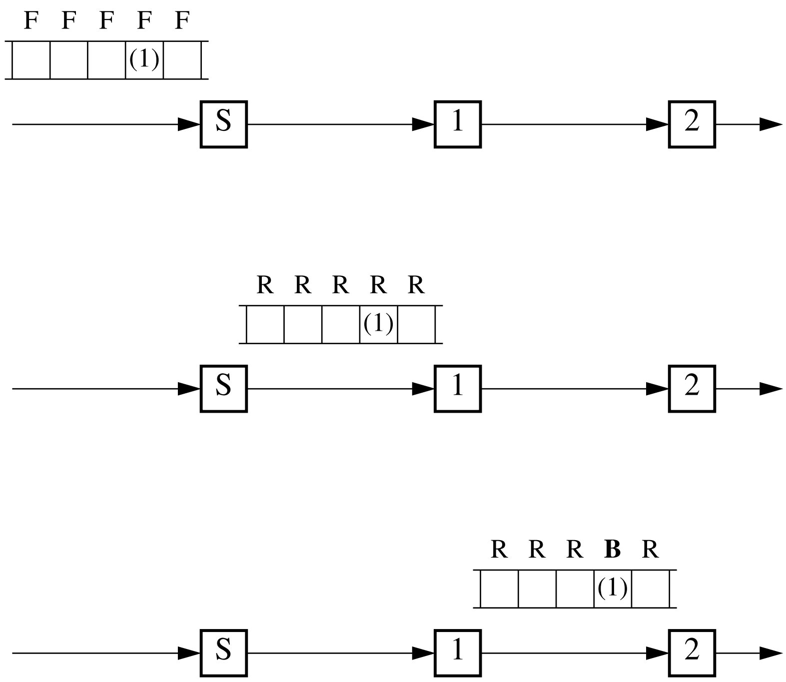


Bild 6-5a:
5 freie Zeitschlitze kommen beim Scheduler an, einer davon gehört Station 1 für ITC-Verkehr

Bild 6-5b:
Der Scheduler hat die 5 freien Zeitschlitze für ATC-Verkehr der Station 2 reserviert

Bild 6-5c:
Station 1 hat ihren Zeitschlitz belegt, der damit nicht mehr wie vom Scheduler vorgesehen für ATC-Verkehr der Station 2 zur Verfügung steht

S	Scheduler
1, 2	Stationen
(1)	Zeitschlitz gehört Station 1 für ihren Isochronverkehr
F	Freier Zeitschlitz
R	Für ATC-Verkehr reservierter Zeitschlitz
B	Belegter Zeitschlitz

Bild 6-5: Problematik bei einem Reservierungsverfahren als Fairness-Mechanismus

Reservierungsverfahren als Fairness-Mechanismus für ATC-Verkehr ähnlich CRMA; Station 2 hat für den aktuellen Zyklus 5 Zeitschlitze für ATC-Verkehr reserviert, ein Zeitschlitz gehört Station 1 für ITC-Verkehr, ist aber beim Scheduler frei und „available“. Statt der vorgesehenen 5 Zeitschlitze kann Station 2 aber im aktuellen Zyklus nur 4 Zeitschlitze belegen, weil einer der Zeitschlitze für ITC-Verkehr genutzt wurde.

nicht benötigt. Jede Station verfügt über einen Zähler für jeden Ring, der immer dann inkrementiert wird, wenn die Station einen Zeitschlitz auf dem zugeordneten Ring mit einer Zelle aus ihrem ATC-Sendepuffer belegt. Die Station darf ATC-Zellen auf einem Ring senden, solange der zugeordnete Zähler eine Schwelle M nicht überschreitet. Kommt auf dem gegengerichteten Ring ein Zeitschlitz mit gesetztem „SAT“-Bit an, wird das „SAT“ so lange bei der Station[1] aufgehalten, bis der Zähler eine Schwelle N erreicht, wobei $N \leq M$ gilt. Anschließend wird das „SAT“ weitergegeben, also im nächsten Zeitschlitz auf dem gegengerichteten Ring gesetzt, und der Zähler wird auf 0 zurückgesetzt. Das Verfahren ist fair, wenn N = M, anson-

1. Voraussetzung ist, daß die Station aktiv ist, also Daten in ihrem ATC-Sendepuffer oder im Einfügepuffer hat

sten kann eine Station mindestens N, höchstens jedoch M Zeitschlitze in jedem „SAT“-Umlauf belegen.

Durch eine andere Bedienstrategie für Einfüge- und ATC-Sendepuffer kann die Wartezeit für ATC-Verkehr einer Station mit hohem ITC-Verkehrsanteil reduziert werden. Wenn nicht mehr der Einfügepuffer, sondern zum Beispiel bei jedem zehnten Zeitschlitz der ATC-Sendepuffer Vorrang hat[1], wird die mittlere Wartezeit der Station für ATC-Verkehr reduziert, allerdings auf Kosten eines größeren mittleren Füllstands des Einfügepuffers und damit einer erhöhten Verzögerung für andere Stationen. Auf diese Weise wird die Benachteiligung einer Station reduziert, die ihre reservierten Zeitschlitze anderen Stationen für Asynchronverkehr zur Verfügung stellt.

6.2.4 Einsatz als Steuergerätenetz

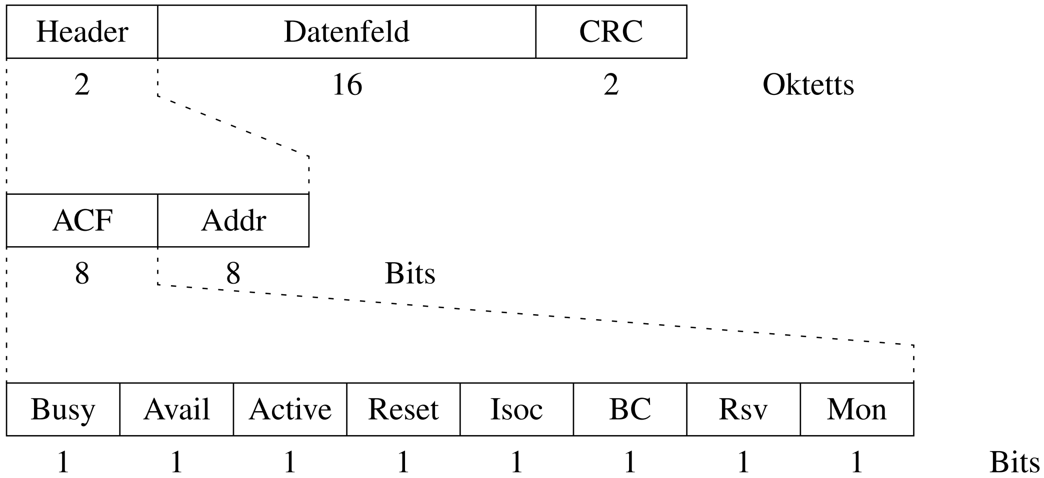
Der Protokollvorschlag kann derart modifiziert werden, daß er für ein Steuergerätenetz geeignet ist. Bei wesentlich geringerer Bitrate liegt der Schwerpunkt in einem Steuergerätenetz auf teilweise zeitkritischer Übertragung kurzer Nachrichten innerhalb kurzer Entfernungen [94].

In einem Steuergerätenetz darf die typische maximale Verzögerungszeit einer ITC-Nachricht etwa 5 ms betragen [94]. Rahmendauer T_R, Rahmengröße R und die Bitrate auf dem Medium $v_ü$ stehen in einem direkten Zusammenhang $R = T_R \cdot v_ü$. Die Rahmendauer T_R darf maximal 5 ms betragen, da aufgrund der geringen Ausdehnung die Anzahl der Rahmen auf dem Ring FOR im Steuergerätenetz gleich Eins ist. Bei einer angenommenen Bitrate von $v_ü$ = 10 Mbit/s ergibt sich für die maximale Rahmengröße R_{max} = 50000 bit.

Da die Nachrichten in einem Steuergerätenetz geringe Größe haben, gleichzeitig aber die Rahmendauer eine kritische Größe darstellt, muß das Zeitschlitzformat deutlich verkürzt werden. Bei 20 Oktetts oder 160 bit Länge für einen Zeitschlitz ergibt sich eine maximale Anzahl von Z_{max} = 312 Zeitschlitzen je Rahmen, das heißt es können bis zu 312 isochrone Kanäle vergeben werden. Weitere 10 Oktetts stehen für Rahmen-Header und Synchronisation zur Verfügung. Da die Anzahl der benötigten ITC-Kanäle in einem typischen Steuergerätenetz deutlich niedriger ist, kann die Übertragungsrate oder die Rahmendauer entsprechend verringert werden.

Bild 6-6 zeigt das Zeitschlitz- und Zellformat beim Einsatz des Protokollvorschlags als Steuergerätenetz. Die Nachrichtenlänge wurde auf 16 Oktetts verkürzt, so daß die Gesamtlänge eines Zeitschlitzes einschließlich des Headers und eines 2 Oktetts großen Prüffelds die geforderten 20 Oktetts beträgt. Der Zeitschlitz-Header ist 2 Oktetts lang, wovon 8 Bits für die Adressierung von bis zu 256 Stationen verwendet werden (siehe Abschnitt 6.2.2.2). Die übrigen 8 Bits bilden das Access Control Field (ACF), wobei die im Protokollvorschlag vorgesehenen Bits für die Paketübertragung entfallen. Stattdessen wird ein Broadcast-Bit gemäß Abschnitt 6.2.2.2 eingeführt. Ein Bit im Header bleibt ungenutzt.

1. Unter der Voraussetzung, daß die Station gemäß Fairness-Mechanismus senden darf



CRC	Cyclic Redundancy Check (umfaßt den gesamten Zeitschlitz)
ACF	Access Control Field
Addr	Address Field
Busy	Busy Bit
Avail	Available Bit
Active	Active Node Bit
Reset	Reset Bit
Isoc	Isochronous Cell
BC	Broadcast Bit
Rsv	Reserved
Mon	Monitor Bit

Bild 6-6: Zeitschlitz- und Zellformat im Steuergerätenetz

6.3 Netzübergänge

Die für den Anschluß von Stationen an das Gigabit-LAN benötigte Treiber-Software stellt sich dem Betriebssystem für den ATC-Verkehr als Ethernet-Treiber dar und verpackt die Ethernet-Pakete in AAL-5-PDUs [19]. Das Datenformat entspricht dabei dem der LAN Emulation (siehe 3.2.2.2). Die Treiber für den ITC-Verkehr sind abhängig von der Anwendung und transportieren anfallende Daten zum Beispiel direkt von einer Videokamera in den ITC-Sendepuffer.

Für die Kommunikation mit Systemen außerhalb des Rings sind zwei Netzübergänge wesentlich: Eine Schnittstelle zu bestehenden lokalen Netzen auf Ethernet-Basis und die Verbindung zu Weitverkehrsnetzen auf Basis des Breitband-ISDN.

6.3.1 Übergang zu Ethernet

Wird das Gigabit-LAN als Backbone-Netz für Ethernet-Segmente eingesetzt oder in eine bestehende Ethernet-Infrastruktur integriert, muß ein Übergang zum Ethernet bereitgestellt werden. Da das verwendete LAN Emulation-Datenformat auf dem Ethernet-Datenformat basiert, entspricht die Funktion des Netzübergangs weitgehend der eines Ethernet-Routers. Die Treibersoftware für das Gigabit-LAN innerhalb des Netzübergangs ist dabei wiederum identisch mit der Treibersoftware einer einfachen Station. Im Gegensatz zur LAN Emulation werden keine speziellen Server benötigt, da das Gigabit-LAN als transparentes Übertragungsmedium für Ethernet-Pakete dient.

Im Netzübergang muß eine Adressumsetzung zwischen den 48 bit breiten Ethernet-Adressen und den in Abschnitt 4.3.8 vorgestellten 8 bit umfassenden Gigabit-LAN-Adressen stattfinden. Dazu muß jeder Station im Gigabit-LAN eine eindeutige Ethernet-Adresse zugeordnet werden, die dem Netzübergang bekannt ist. Der Netzübergang stellt eine Tabelle bereit, in der für jede Station im Gigabit-LAN ein Eintrag existiert, der die Ethernet-Adresse der Station der Gigabit-LAN-Adresse zuordnet. Ein auf einem Ethernet-Segment ankommendes Datenpaket, das für eine Station innerhalb des Gigabit-LAN bestimmt ist, erkennt der Netzübergang daran, daß die Zieladresse in seiner Tabelle eingetragen ist. Das Paket wird dann mit der Gigabit-LAN-Adresse der Zielstation versehen und weitergeleitet.

Der Aufwand für die Adressumsetzung kann minimiert werden, wenn die Ethernet-Adressen innerhalb des Gigabit-LANs nach einem modifizierten Verfahren vergeben werden [74]. Jeder Ring erhält dafür einen Vorrat an 256 Ethernet-Adressen, die alle in den ersten 5 Oktetts übereinstimmen. Die Ethernet-Adresse einer Station ergibt sich dann aus der Kombination der Ring-Adresse und der Gigabit-LAN-Adresse der Station. Der Netzübergang leitet alle auf dem Ethernet ankommenden Pakete weiter, deren Ethernet-Adresse in den ersten 5 Oktetts mit der Ring-Adresse übereinstimmen. Die Gigabit-LAN-Adresse der Zielstation ist identisch mit dem letzten Oktett der Ethernet-Zieladresse. In umgekehrter Richtung wählt eine Station für ein Datenpaket, dessen Ziel nicht innerhalb des Gigabit-LAN liegt, die Adresse des Netzübergangs als Gigabit-LAN-Adresse.

6.3.2 Übergang zu Weitverkehrsnetzen

Da der Asynchronverkehr bereits das ATM-Zellformat verwendet, ist ein Übergang zu einem auf ATM basierenden Weitverkehrsnetz wie Breitband-ISDN sehr einfach möglich. Die Funktionalität eines Netzübergangs ist die eines Vermittlungsknotens, der Signalisierung ermöglicht und die innerhalb des Gigabit-LANs beschränkte Adressierung auf virtuelle Pfade und Kanäle im Weitverkehrsbereich umsetzt. Allerdings muß jede Verbindung ins Weitverkehrsnetz im Gegensatz zu internen Verbindungen explizit aufgebaut werden. Paketübertragung ohne Unterbrechung ist nicht möglich, da dieser Mechanismus vom Weitverkehrsnetz nicht unterstützt wird.

Weitverkehrsnetze, die auf der Synchronen Digitalen Hierarchie (SDH) [44, 45] basieren, sind geeignet für die Aufnahme von ITC-Verbindungen. Für jede ITC-Verbindung zu einem Kommunikationspartner außerhalb des Gigabit-LANs wird eine synchrone Verbindung aufgebaut, deren Bitrate mindestens der Spitzenbitrate der ITC-Verbindung entspricht. Ist der Netzübergang gleichzeitig Monitor- und Masterstation, kann das Gigabit-LAN mit dem 8 kHz Rahmentakt des externen Netzes synchronisiert werden.

Kapitel 7

Zusammenfassung und Ausblick

7.1 Zusammenfassung

Für eine optimale Integration von asynchronem und isochronem Verkehr in einem Ring wurde ein Medienzugriffsprotokoll entwickelt, das neben asynchronem Datenverkehr die Reservierung von Zeitschlitzen für isochrone Kanäle erlaubt. Zusätzlich zu den nicht reservierten Zeitschlitzen kann der Asynchronverkehr auch reservierte Zeitschlitze nutzen, wenn sie nicht von isochronem Verkehr belegt sind. Damit steht für zeitkritische, büschelhafte Verkehre ein effizienter Mechanismus zur Verfügung, der den bisher zwangsläufigen Zusammenhang zwischen Reservierung der Spitzenbitrate und der damit einhergehenden ineffizienten Nutzung der zur Verfügung stehenden Bandbreite beseitigt.

Eine Station, die Zeitschlitze für zeitkritischen Verkehr reserviert hat, kann jeden ihrer Zeitschlitze einzeln für Asynchronverkehr freigeben oder sperren. Kommt ein solcher Zeitschlitz belegt bei der Station an, wird dessen Inhalt durch einen Einfügepuffer geleitet und bei Bedarf zwischengespeichert. Der Einfügepuffer kann dabei nicht überlaufen, weil bei drohendem Überlauf Zeitschlitze nicht mehr für Asynchronverkehr freigegeben werden.

Die Bandbreite für Asynchronverkehr wird unter den Stationen mit Hilfe eines Fairness-Mechanismus flexibel aufgeteilt. Dafür kann jedes Verfahren eingesetzt werden, das auch bei variabler Kanalbandbreite Fairness garantiert. Der beispielhaft implementierte Fairness-Mechanismus auf Credit-Basis wurde aufgrund seiner einfachen Realisierbarkeit gewählt.

Der Mechanismus zur unterbrechungsfreien Paketübertragung sowie das vorgeschlagene Adressierungsverfahren ermöglichen eine effizientere Implementierung der Empfangsstufe.

Zum Nachweis der Realisierbarkeit des beschriebenen Protokolls wurde eine mögliche Stationsstruktur entwickelt und prototypisch implementiert.

Zur Beurteilung der Leistungsfähigkeit wurde ein Modell des Protokolls entwickelt und simuliert. Das dafür erstellte Simulationsprogramm ermöglicht es, den Durchsatz und die Wartezeit einzelner Datenströme zu messen. Neben symmetrischen Verkehrsbeziehungen wurden asymmetrische und heterogene Szenarien betrachtet. Dabei hat sich gezeigt, daß das vorgeschlagene Protokoll einen höheren Gesamtdurchsatz gegenüber Systemen mit exklusiver Reservierung von Zeitschlitzen bietet.

Die Simulation hat die bekannten Unzulänglichkeiten des gewählten einfachen Fairness-Verfahrens auf Credit-Basis bestätigt. Dazu gehört insbesondere die Abhängigkeit des Durchsatzes von der Zykluslänge, wobei die Zykluslänge nicht beliebig gesteigert werden kann, weil dadurch die Wartezeit erhöht wird. Soll das System in dieser Hinsicht optimiert werden, lohnt sich die Implementierung eines aufwendigeren Fairness-Verfahrens.

Bei inhomogener Verteilung des Isochronverkehrs ist die verbleibende Bandbreite für den Asynchronverkehr nicht überall im Ring identisch. Daraus resultiert eine nicht bei allen Stationen gleiche mittlere Wartezeit, die jedoch geringer ist als bei einem System mit exklusiver Reservierung von Zeitschlitzen. Für wirklich zeitkritischen Verkehr werden isochrone Kanäle verwendet.

Insgesamt hat die Simulation gezeigt, daß das Protokoll sehr gut für heterogene Verkehrsbeziehungen und damit für die Integration künftiger Dienste in einem Gigabit-LAN geeignet ist.

In einem Laboraufbau wurde eine Station des Gigabit-LAN mit vertretbarem Aufwand realisiert. Dabei kamen komplexe programmierbare Logikbausteine in einem auf mehrere Platinen verteilten System zum Einsatz. Die Übertragung der Aufwandsabschätzung auf einen anwendungsspezifischen Schaltkreis oder auf einen programmierbaren Logikbaustein der neuesten Generation ergab, daß die Funktionen einer Station auf einer Einschubkarte integrierbar sind.

Da der Protokollvorschlag auf dem ATM-Zellformat beruht, ist ein Übergang zu einem ATM-Weitverkehrs- oder Lokalnetz sehr einfach realisierbar. Für den Übergang zu Ethernet-Systemen wurden effiziente Mechanismen vorgeschlagen, die ein effizientes Adressierungsverfahren und die unterbrechungsfreie Paketübertragung umfassen. Im Gegensatz zur LAN-Emulation des ATM-Forums wird nur ein einfacher Netzübergang benötigt.

Neben Varianten des Systems, die der Leistungssteigerung oder der Kosteneinsparung dienen, wurde der Einsatz des Protokolls als Steuergerätenetz beschrieben. Bei modifizierter Zeitschlitzstruktur ist der gleiche Protokollmechanismus für die Realisierung eines Steuergerätenetzes anwendbar. Für die Übertragung sporadischer, aber zeitkritischer Nachrichten dienen isochrone Kanäle, während zeitunkritische Nachrichten asynchron übertragen werden. Die Kosten lassen sich aufgrund der wesentlich vereinfachten Struktur und geringeren Bitrate gegenüber dem Gigabit-LAN drastisch reduzieren.

Zusammenfassend läßt sich feststellen, daß mit der heutigen Technologie effiziente Medienzugriffsprotokolle für lokale Netze mit Vielfachzugriff realisierbar sind, die sowohl gegenüber einem hochbitratigen Ethernet, als auch gegenüber ATM-basierten lokalen Netzen Vorteile bieten. Für den Protokollvorschlag wurde neben bekannten Topologien und Verfahren ein neuer Protokollmechanismus für die Integration zweier Verkehrsklassen mit gegensätzlichen Anforderungsprofilen eingesetzt.

7.2 Ausblick

Die Diskussion über die Zukunft lokaler Netze ist derzeit geprägt von ATM-LANs und Fast Ethernet. Dabei wird zunehmend deutlich, daß beide Systeme wesentliche Nachteile aufweisen. Ethernet ist einfach realisierbar und relativ billig, bietet aber keine Garantie für eine bestimmte Dienstgüte und ist deshalb für multimediale Anwendungen nicht geeignet. Ein ATM-LAN ist sehr aufwendig und erfordert hohe Anfangsinvestitionen. Die Probleme bei der Integration isochroner Echtzeitdienste sind nicht gelöst. Gefordert werden preiswerte Hochgeschwindigkeitsnetze mit Dienstgütegarantie.

Wenn hoher Datendurchsatz und zeitkritische Dienste kombiniert werden sollen, ist das vorgestellte Gigabit-LAN eine leistungsfähige und mit vertretbarem Aufwand realisierbare Alternative zu Ethernet und ATM-LANs. Beim Einsatz als Backbone-Netz fügt es sich nahtlos in heterogene Umgebungen ein. Selbst ein ATM-Vermittlungsknoten, der isochrone und asynchrone Dienste integriert, kann mit Hilfe des Gigabit-LANs realisiert werden: Die einzelnen Stationen dienen dazu, je eine ATM-Schnittstelle zur Verfügung zu stellen. Auf diese Weise wird die Verbindungsmatrix mit vollständiger Erreichbarkeit auf das Gigabit-LAN abgebildet („LAN in a box“).

Die vielfältigen Einsatzmöglichkeiten von der integrierten LAN-Anwendung über Backbone-Netze bis hin zu einem ATM-Vermittlungsknoten, die einfache Realisierbarkeit sowie die effiziente Integration isochroner und asynchroner Dienste zeichnen das vorgestellte Hochgeschwindigkeits-LAN gegenüber bisherigen Lösungen aus.

Literatur

[1] Adams, J. L.: „Orwell“; Computer Networks and ISDN Systems, Volume 26, March 1994, pp. 771-784.

[2] Altera Corporation: „1996 Data Book“; San Jose, USA, June 1996.

[3] Altera Corporation: „News & Views“; San Jose, USA, First Quarter, February 1997.

[4] Anastasi, G.; Conti, M.; Gregori, E.; Lenzini, L.: „CRMA Media Access Control Protocol: A Simulative Analysis“; 5th IEEE Workshop on MAN, Session T1, Taormina, Italy, May 1992.

[5] van As, H. R.: „Media access techniques: The evolution towards terabit/s LANs and MANs“; Computer Networks and ISDN Systems, Volume 26, March 1994, pp. 603-656.

[6] van As, H. R.; Lemppenau, W. W.: „CRMA II: Dynamic Embedding of Isochronous Channels“; 12th Conference on European Fibre Optic Communicatons and Networks, Heidelberg, Germany, June 1994.

[7] van As, H. R.; Lemppenau, W. W.; Schindler, H. R.; Zafiropulo, P.: „CRMA-II: A MAC protocol for ring-based Gb/s LANs and MANs“; Computer Networks and ISDN Systems, Volume 26, March 1994, pp. 831-840.

[8] van As, H. R.; Lemppenau, W. W.; Zafiropulo, P.: „Performance of CRMA-II: A Reservation-Based Fair Media Access Protocol for Gbit/s LANs and MANs with Buffer Insertion“; Proceedings of the 10th EFOC/LAN, Paris, June 24-26, 1992, pp. 162-169.

[9] van As, H. R.; Lemppenau, W. W.; Zafiropulo, P.; Zurfluh, E. A.: „CRMA-II: A Gbit/s MAC Protocol for Ring and Bus Networks with Immediate Access Capabilities“; Proceedings of the 9th EFOC/LAN, London, UK, June 19-21, 1991, pp. 262-277.

[10] ATM Forum: „ATM User-Network Interface (UNI) Signalling Specification Version 4.0“; The ATM Forum Technical Committee, July 1996.

[11] ATM Forum: „LAN Emulation Over ATM“; Version 1.0, The ATM Forum Technical Committee, January 1995.

[12] ATM Forum: „Multi-Protocol Over ATM“; Version 1.0, Draft, Tne ATM Forum Technical Committee, February 1997.

[13] ATM Forum: „User-Network Interface (UNI) Specification Version 3.1“; The ATM Forum Technical Committee, September 1994.

[14] Baum, E.: „Entwicklung einer Testplattform für die Untersuchung der Systemeigenschaften komplexer EPLDs bei hohen Taktraten“; Semesterarbeit am Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, Mai 1993.

[15] Beeler, R.; Potts, M.; Rao, S.: „D3Q — The Dynamic Distributed Dual Queue — An Optimised Distributed Queueing Protocol“; Proceedings of the 13th ISS 90, Vol. II, May 1990, pp. 121-125.

[16] Chen, J. S.-C.; Cidon, I.; Ofek, Y.: „A Local Fairness Algorithm for Gigabit LAN‘s/MAN‘s with Spatial Reuse“; IEEE Journal on Selected Areas in Communications, Vol. 11, No. 8, October 1993, pp. 1183-1192.

[17] Cidon, I.; Ofek, Y.: Metaring — A Full-duplex Ring with Fairness and Spatial Reuse“; Proceedings of IEEE Infocom ‚90, San Francisco, CA, June 1990, pp. 969-981.

[18] Comer, D. E.: „Internetworking With TCP/IP — Vol I: Principles, Protocols and Architecture“; Second Edition, Prentice-Hall International Editions, ISBN 0-13-47321-0, 1991.

[19] Eberle, S.: „Entwicklung eines MAC-Treibers für ein High-Speed LAN“; Semesterarbeit am Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, November 1996.

[20] Enssle, J.: „Modellierung und Leistungsuntersuchung eines verteilten Video-On-Demand-Systems für MPEG-codierte Videodatenströme mit variabler Bitrate“; Dissertation, eingereicht am Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, 1996.

[21] Götzer, M.: "A New Access Control Protocol For High-Speed ATM LANs"; Proceedings of the 5th Open Workshop on High Speed Networks, Paris, March 20-21, 1996.

[22] Götzer, M.: "A New MAC Protocol For High-Speed Multimedia ATM LANs"; Eight IEEE Workshop on Local and Metropolitan Area Networks, Session 8, Berlin/Potsdam, August 25-28, 1996.

[23] Goodall, D.: „An Efficient Medium Access Scheme to Support VBR Traffic on a Slotted Ring“; International Journal of Communication Systems, Vol. 9, No. 3, May/June 1996, pp. 151-159.

[24] Goodall, D. S.; Burston, A. K.: „Medium Access Control for Asynchronous Traffic in the AMNET LAN“; International Journal of Communication Systems, Vol. 8, 1995, pp. 155-163.

[25] Hahne, E. L.; Choudhury, A. K., Maxemchuk, N. F.: „Improving the Fairness of the Distributed-Queue-Dual-Bus Networks“; Proceedings of the 9th Joint Conference of the IEEE Computer and Communications Societies (Infocom), San Francisco, California, USA, June 1990, pp. 175-184.

[26] Hartmann, H.: „Protokolle für Hochgeschwindigkeits-LANs mit ATM-Anbindung und Simulation des ATMR-Protokolls“; Semesterarbeit am Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, Oktober 1994.

[27] Hedenetz, A.: „Entwicklung und Integration der Master-Einheit für einen HSLAN-Knoten“; Diplomarbeit am Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, 1997.

[28] Herrmann, M.: „Spezifikation und Implementierung des OSEK-Netzmanagements für D2B“; Diplomarbeit am Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, Oktober 1996.

[29] IEEE Standard 802.3u for Local and Metropolitan Area Networks: Supplement to Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications: Media Access Control (MAC) Parameters, Physical Layer, Medium Attachment Units, and Repeater for 100 Mb/s Operation.

[30] IEEE Standard 802.12, IEEE Standards for Local and Metropolitan Area Networks: Demand Priority Access Method Physical Layer and Repeater Specification for 100 Mbit/s Operation.

[31] Imai, K.; Ito, T.; Kasahara, H; Morita, N.: „ATMR: Asynchronous transfer mode ring protocol“; Computer Networks and ISDN Systems, Volume 26, March 1994, pp. 785-798.

[32] ISO 7498: „Information Processing Systems — Open Systems Interconnection — Basic Reference Model“; 1984.

[33] ISO/IEC 8802-2 : „Information technology — Telecommunications and information exchange between systems — Local and Metropolitan area networks — Specific requirements — Part 2: Logical link control“; [ANSI/IEEE 802.2, 1994 Edition], 1994.

[34] ISO/IEC 8802-3: „Information technology — Local and metropolitan area networks — Part 3: Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications“; [ANSI/IEEE Std 802.3, 1995 Edition], 1995.

[35] ISO/IEC 8802-4: „Information processing systems — Local area networks — Part 4: Token-passing bus access method and physical layer specifications“; [ANSI/IEEE Std 802.4-1990], 1994.

[36] ISO/IEC 8802-5: „Information technology — Local and metropolitan area networks — Part 5: Token ring access method and physical layer specifications“; [ANSI/IEEE Std 802.5-1992], 1992.

[37] ISO/IEC 8802-6: „Information technology — Telecommunications and information exchange between systems — Local and metropolitan area networks — Specific requirements — Part 6: Distributed Queue Dual Bus (DQDB) access method and physical layer specifications“; [ANSI/IEEE Std 802.6, 1994 Edition], 1994.

[38] ISO/IEC 10038: „Information technology — Telecommunications and information exchange between systems — Local area networks — Media access control (MAC) bridges“; [ANSI/IEEE Std 802.1D, 1993 Edition], 1993.

[39] ISO/IEC 9314-1: „Information processing systems — Fibre Distributed Data Interface (FDDI) — Part 1: Token Ring Physical Layer Protocol (PHY)“; 1989.

[40] ISO/IEC 9314-2: „Information processing systems — Fibre Distributed Data Interface (FDDI) — Part 2: Token Ring Media Access Control (MAC)“; 1989.

[41] ISO/IEC 9314-3: „Information processing systems — Fibre Distributed Data Interface (FDDI) — Part 3: Physical Layer Medium Dependent (PMD)“; 1990.

[42] ISO/IEC 9314-5: „Information technology — Fibre Distributed Data Interface (FDDI) — Part 5: Hybrid Ring Control (HRC)“; 1995.

[43] ITU-T E.164: „Numbering Plan for the ISDN Era“; ITU-T Recommendation E.164, 8/91.

[44] ITU-T G.708: „Network Node Interface for the Synchronous Digital Hierarchy“; ITU-T Recommendation G.708, 3/93.

[45] ITU-T G.709: „Synchronous Multiplexing Structure“; ITU-T Recommendation G.709, 3/93.

[46] ITU-T I.150: „B-ISDN Asynchronous Transfer Mode Functional Characteristics“; ITU-T Recommendation I.150, 3/93.

[47] ITU-T I.321: „B-ISDN Protocol Reference Model and its Application“; ITU-T Recommendation I.321, 4/91.

[48] ITU-T I.361: „B-ISDN ATM Layer Specification“; ITU-T Recommendation I.361, 11/95.

[49] ITU-T I.362: „B-ISDN ATM Adaptation Layer (AAL) Functional Description“; ITU-T Recommendation I.362, 3/93.

[50] ITU-T I.363: „B-ISDN ATM Adaptation Layer (AAL) Specification“; ITU-T Recommendation I.363, 3/93.

[51] ITU-T I.371: „Traffic Control and Congestion Control in B-ISDN“; ITU-T Recommendation I.371, 11/93.

[52] ITU-T I.432: „B-ISDN User-Network Interface - Physical Layer Specification“; ITU-T Recommendation I.432, 3/93.

[53] ITU-T Q.2931: „Broadband integrated services digital network (B-ISDN) - Digital subscriber signalling system No. 2 (DSS-2) - User-Network interface (UNI) layer 3 specification for basic call/connection control“; ITU-T Recommendation Q.2931, 2/95.

[54] ITU-T X.200: „Information Technology — Open Systems Interconnection — Basic Reference Model: The Basic Model“; ITU-T Recommendation X.200, 7/94.

[55] Jeffrey, L.: „LAN-Emulation über ATM-Netzwerke“; Datacom 8/96, Vol. 13, August 1996, pp. 86-92.

[56] JTC1/SC6 WG6 N71: „Specification of the Asynchronous Transfer Mode Ring (ATMR) Protocol“; Japanese National Body, Version A1.0, 20. September 1990.

[57] Kamal, A. E.; Wong, J. W.; Hassanein, H. S.: „An algorithm for slot reuse in DQDB networks with erasure nodes“; Computer Networks and ISDN Systems, Volume 26, March 1994, pp. 695-710.

[58] Kocher, H.: „Entwurf und Implementierung einer Simulationsbibliothek unter Anwendung objektorientierter Methoden“; 59. Bericht über verkehrstheoretische Arbeiten, Universität Stuttgart, IND, 1994.

[59] Kühlewein, M.: „Entwicklung der Ausgangsstufe für ein HSLAN“; Semesterarbeit am Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, 1997.

[60] Kühn, P. J.: „Codierung zur Fehlererkennung und Fehlerkorrektur“; Manuskript zur Vorlesung, Universität Stuttgart, 1995.

[61] Kühn, P. J.: „Technische Informatik III – Kommunikationsnetze“; Manuskript zur Vorlesung, Universität Stuttgart, 1995/96.

[62] Kühn, P. J.: „Nachrichtenverkehrstheorie“; Manuskript zur Vorlesung, Universität Stuttgart, 1995/96.

[63] Lemppenau, W. W.; van As, H. R.; Schindler, H. R.: "A 2.4 Gbit/s ATM Implementation of the CRMA-II Dual-Ring LAN and MAN", Proceedings of the 11th Conference on European Fibre Optic Communicatons and Networks, The Hague, Netherlands, June 1993, pp 274-281.

[64] Lemppenau, W. W.; Götzer, M.; Seibold, W.: "Access Protocols for High-Speed LANs and MANs", Proceedings of the 11th Conference on European Fibre Optic Communicatons and Networks, The Hague, Netherlands, June 1993, pp 253-259.

[65] Limb, J. O.: „A Simple Multiple Access Protocol For Metropolitan Area Networks“; Prceedings of ACM SIGCOMM, Philadelphia, PA, USA, September 1990.

[66] Limb, J. O., Flores, C.: „Description of Fasnet - A Unidirectional Local-Area Communications Network“; The Bell System Technical Journal, Vol. 61, No. 7, September 1982, pp. 1413-1440.

[67] Lyle, J.; Harvey, M.: „SBus Specification B.0“; Revision A, Sun Microsystems Inc., Mountain View, USA, December 1990.

[68] Meuser, T.: „Enhancement of the CRMA-II protocol by distributed marking“; Proceedings of the 12th Conference on European Fibre Optic Communicatons and Networks, Heidelberg, Germany, June 1994, pp. 129-132.

[69] Mors, P.: „Entwicklung der Monitoring- und Masterfunktionen für die Eingangsstufe eines HS-LANs“; Semesterarbeit am Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, 1997.

[70] Music Semiconductors: „The MU9C1480 LANCAM Handbook“; Rev. 0, Colorada Springs, USA, February 1992.

[71] Motorola: „Dynamic RAMs and Memory Modules“; Rev. 1, Motorola Semiconductor, Phoenix, USA, March 1994.

[72] Nassehi, M. M.: „Cyclic-Reservation Multiple-Access Scheme for Gbit/s LANs and MANs based on Dual-Bus Configuration“; Proceedings of the 8th European Fibre Optic Communications and LAN Conference, Munich, Germany, June 1990, pp. 246-251.

[73] Ofec, Y.: „Overview of the MetaRing architecture“; Computer Networks and ISDN Systems, Volume 26, March 1994, pp. 817-829.

[74] Payer, W.: „Entwicklung eines TCP/IP-Gateways zwischen Ethernet und einem High-Speed-LAN“; Studienarbeit am Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, Dezember 1996.

[75] PCI SIG: „PCI Local Bus Specification“; Revision 2.1, PCI Special Interests Group, Portland, USA, June 1995.

[76] Pohl, T.: „Simulation und Optimierung eines MAN-Medienzugriffsprotokolls“; Diplomarbeit am Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, Juli 1995.

[77] Probst, U.: „Verkehrstheoretische Analyse des Medienzugriffsprotokolls von Orwell“; Diplomarbeit am Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, Juli 1990.

[78] RFC 768: „User Datagram Protocol“; Request For Comments: 768; ISI, August 1980.

[79] RFC 791: „Internet Protocol“; Request For Comments: 791; Darpa Internet Program Protocol Specification, September 1981.

[80] RFC 793: „Transmission Control Protocol“; Request For Comments: 793; Darpa Internet Program Protocol Specification, September 1981.

[81] RFC 826: „An Ethernet Address Resolution Protocol“; Request For Comments: 826; Network Working Group, November 1982.

[82] RFC 1577: „Classical IP and ARP over ATM“; Request For Comments: 1577, Network Working Group, January 1994.

[83] Roberts, E.: „It Takes Two to Tangle: Do 1-Gbit/s versions of 100Base-T and 100VG mean double trouble for net managers?“; Data Communications International, March 21, 1996, pp. 39-41.

[84] Ross, F. E.; Fink, R. L.: „Overview of FFOL - FDDI Follow-On LAN“; Computer Communiacations, Vol. 15, No. 1, January 1992, pp. 3-10.

[85] Rössel, H.: „Virtuelle Netze: Standort-Bestimmung und Bewertung“; Datacom 8/96, Vol. 13, August 1996, pp. 58-71.

[86] Rubin, I.; Wu, H.-T.: „FECCA - A New Access Algorithm for an ATM Ring Network with Destination Removal“; Proceedings of Twelfth IEEE Infocom, Paper 3d.3, San Francisco, USA, March/April 1993.

[87] Sauer, K.; Schödl, W.: „Performance Aspects of the DQDB Protocol“; Proceedings of the 7th ITC Specialists Seminar, Paper 16.3, Adelaide, Australia, September 1989.

[88] Schödl, W.: „Kopplung von DQDB-Regionalnetzen mit ATM-Weitverkehrsnetzen: Architektur, Steuerungsstrategien und Leistungsverhalten“; 58. Bericht über verkehrstheoretische Arbeiten, Universität Stuttgart, IND, 1994.

[89] Schödl, W.; Tangemann, M.: „Performance evaluation of CRMA“; Proceedings of the 10th EFOC/LAN, Paris, June 24-26, 1992, pp. 170-174.

[90] Schröder, C.: „Entwicklung der Protokollverarbeitungseinheit für ein HSLAN“; Semesterarbeit am Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, November 1995.

[91] Slosiar, R.; Potts, M.; Beeler, R.: „MD^3Q: A distributed queueing protocol with full channel capacity re-use and guarantee of bandwidth“; Computer Networks and ISDN Systems, Volume 26, March 1994, pp. 799-815.

[92] Spratt, M. P.: „Bandwidth Allocation in IEEE 802.6 Using Non Unity Ratio Bandwidth Balancing“; Proceedings of the ICC, Paper 24.1, June 1991, pp. 729-735.

[93] Stickel, J.: „Simulation des Medienzugriffsprotokolls CRMA-II und Optimierung dessen Reservierungsprotokolls“; Diplomarbeit am Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, Juni 1992.

[94] Stümpfle, M.: „Planung und Optimierung von prioritätsbasierten Steuergerätenetzen für Fahrzeuge“; Dissertation, eingereicht am Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, 1997.

[95] Stürzl, V.: „Entwicklung von komponenten für einen Hochgeschwindigkeits-LAN-Knoten“; Semesterarbeit am Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, April 1995.

[96] Stüttgen, H. J.: „Network Evolution and Multimedia Communication“; IEEE Multimedia, October 1995, pp. 42-59.

[97] Tanenbaum, A. S.: „Computer Networks“; Third Edition, Prentice-Hall International Editions, ISBN 0-13-394248-1, 1996.

[98] Tangemann, M.: „Modellierung und Analyse von lokalen Hochgeschwindigkeitsnetzen mit zeitgesteuerten Token Passing-Medienzugriffsprotokollen“; 57. Bericht über verkehrstheoretische Arbeiten, Universität Stuttgart, IND, 1994.

[99] Theimer, T.: „Vergleichende Untersuchungen an ATM-Koppelnetzstrukturen“; 56. Bericht über verkehrstheoretische Arbeiten, Universität Stuttgart, IND, 1994.

[100] Tran-Gia, P.; Dittmann, R.: „Performance Analysis of the CRMA-Protocol in High-Speed Networks“; IFIP Transactions C5, 1992, pp. 27-46.

[101] Truong, H. L.; Ellington Jr., W. W.; Le Boudec, J.-Y.; Meier, A. X.; Pace, J. W.: „LAN Emulation on an ATM Network“; IEEE Communications Magazine, May 1995, pp. 70-85.

[102] Watson, G.: „The S++ MAC protocol“; Computer Networks and ISDN Systems, Volume 26, March 1994, pp. 745-755.

[103] Watson, G.; Ooi, S.; Skellern, D.; Cunningham, D.: „HANGMAN Gb/s Network“; IEEE Network, Volume 6, July 1992, pp. 10-18.

[104] Watson, G.; Tohmé, S.: „A Performance Analysis of S++: A MAC Protocol for High Speed Networks“; Proceedings of the 3rd IFIP International Workshop on Protocols for High-Speed Networks, Stockholm, Sweden, May 1992, pp. 87-102.

[105] Weikinger, M.: „Simulation eines ATM-LAN-Protokolls“; Studienarbeit am Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, Dezember 1996.

[106] Wirth, N.: „Algorithmen und Datenstrukturen mit Modula-2“; B. G. Teubner Stuttgart, 4. Auflage, ISBN 3-519-02260-5, 1986.

[107] Wong, J.W.: „Throughput of DQDB Networks under Heavy Load“; Proceedings of EFOC/LAN, Paper 1.4.3, Amsterdam, Netherlands, June 1989.

[108] Zafirovic-Vucotic, M.; Niemegeers, I. G.: „Analytical Model and Performance Analysis of the Orwell Access Mechanism“; Proceedings of the 12th ITC, Paper 2.3B.5, Torino, June 1988.

Anhang A

Detailbeschreibung der Protokollverarbeitungseinheit

A.1 Funktionen der Protokollsteuerung

Die Protokollsteuerungsfunktionen beziehen sich immer auf den aktuellen Zeitschlitz und sind abhängig von den Eingangsgrößen, die in Tabelle A-1 dargestellt sind. Die Funktionen werden zur übersichtlicheren Darstellung zunächst für eine Station dargestellt, die nicht über die Funktion verfügt, ein Paket ohne Unterbrechung zu senden. Die für die Paketübertragung notwendige zusätzliche Funktionalität einer Station ist in Abschnitt A.3 beschrieben.

<table>
<tr><th>Funktionsblock</th><th>Größe</th><th>Werte</th></tr>
<tr><td rowspan="4">Header-Bits im aktuellen Zeitschlitz</td><td>Busy</td><td>Ja = Zeitschlitz ist belegt, Nein = Zeitschlitz ist frei</td></tr>
<tr><td>Avail</td><td>Ja, Nein</td></tr>
<tr><td>Reset</td><td>Ja = Reset wurde ausgelöst, Nein = kein Reset</td></tr>
<tr><td>Isoc</td><td>Ja = ITC-Zelle, Nein = ATC-Zelle</td></tr>
<tr><td>Einfügepuffer</td><td>IBEmpty</td><td>Ja = leer, Nein sonst</td></tr>
<tr><td rowspan="2">Sendepuffer</td><td>ATCEmpty</td><td>Ja = ATC-Puffer enthält mindestens eine komplette Zelle, Nein sonst</td></tr>
<tr><td>ITCEmpty</td><td>Ja = ITC-Puffer enthält mindestens eine komplette Zelle, Nein sonst</td></tr>
<tr><td rowspan="2">Verbindungs-steuerung</td><td>Own</td><td>Ja = Zeitschlitz ist für die Station reserviert, Nein sonst</td></tr>
<tr><td>Pend</td><td>Ja = Zeitschlitz war für die Station reserviert und wird gerade freigegeben, Nein sonst.
Zur besseren Übersichtlichkeit wird die Funktion erst in Abschnitt A.2 eingeführt.</td></tr>
<tr><td>Credit Counter</td><td>CCEmpty</td><td>Ja = Credits sind aufgebraucht, Nein sonst</td></tr>
<tr><td>Available Counter</td><td>AC, ACT</td><td>Available Counter, Available Counter Threshold (Schwellwert des Available Counter)</td></tr>
</table>

Tabelle A-1: Eingangsgrößen der Protokollsteuerung

Um das Verständnis der Grundfunktionen innerhalb der Protokollsteuerung zu erleichtern, werden die für den Normalbetrieb unerheblichen Sondermechanismen während des Verbindungsauf- und -abbaus in einem getrennten Abschnitt A.2 erläutert.

A.1.1 Steuerung des Einfügepuffers und des Multiplexers M

Beim Einfügepuffer werden zwei Funktionen benötigt: das Schreiben einer kompletten Zelle in den Einfügepuffer sowie das Lesen einer kompletten Zelle. Die Steuerung des Multiplexers M ist korreliert mit dem Lesevorgang aus dem Einfügepuffer: immer wenn aus dem Einfügepuffer eine Zelle gelesen wird, wird der Multiplexer M so gesteuert, daß das Zeitschlitzgerüst auf dem direkten Weg, der Inhalt des Zeitschlitzes jedoch aus dem Einfügepuffer zum Ausgang der Protokollverarbeitungseinheit weitergegeben wird. Tabelle A-2 enthält die Schreibfunktion des Einfügepuffers, während die Lesefunktion, die identisch ist mit der Steuerfunktion des Multiplexers M für Zellen, in Tabelle A-3 enthalten ist. Voraussetzung für das Schreiben einer ankommenden Zelle in den Einfügepuffer ist, daß der Zeitschlitz eine ATC-Zelle enthält und als „available" markiert ist. Gleichzeitiges Schreiben und Lesen in einen zuvor leeren Einfügepuffer wird vermieden, um die Verzögerungszeit in einem gering ausgelasteten System zu minimieren.

Busy	Isoc	Avail	Own	IBEmpty	ITCEmpty	Zelle schreiben
J	N	J	N	N	X	J
J	N	J	J	X	N	J
J	N	J	J	N	J	J
sonst						N

J: Ja N: Nein X: Don't Care

Tabelle A-2: Schreibfunktion des Einfügepuffers

Busy	Isoc	Avail	Own	IBEmpty	ITCEmpty	Zelle LESEN
X	X	X	J	N	J	J
N	X	J	N	N	X	J
J	N	J	N	N	X	J
sonst						N

J: Ja N: Nein X: Don't Care

Tabelle A-3: Lesefunktion des Einfügepuffers (identisch Steuerfunktion des Multiplexers M)

A.1.2 Steuerung der Sendepuffer im Sender

Die Sendepuffer der Station sind zwar Elemente der Sender-Einheit (Abschnitt 4.4.7), die Sendesteuerung ist jedoch Teil der Protokollverarbeitung. Für die beiden Sendepuffer für ITC- und ATC-Zellen werden grundsätzlich verschiedene Sendefunktionen benötigt, die in den Tabellen A-4 und A-5 gegenübergestellt sind. Während die Sendefunktion für ITC-Zellen sehr einfach ist und lediglich vom Reservierungszustand des Zeitschlitzes und dem Füllstand des ITC-Sendepuffers abhängt, ist die Sendefunktion für den ATC-Verkehr wesentlich komplexer, da der

ATC-Verkehr die Restbandbreite nutzt und somit vom ITC-Verkehr, der Systemauslastung und darüber hinaus noch vom Fairness-Protokoll abhängt. Grundvoraussetzung für das Senden einer ATC-Zelle ist, daß der Einfügepuffer leer ist, der ATC-Sendepuffer mindestens eine Zelle zum Senden bereithält und das Fairness-Protokoll das Senden erlaubt (beim gewählten Credit-Verfahren müssen noch Credits vorhanden sein). Zwei mögliche Fehlerfälle werden automatisch erkannt und korrigiert: Einerseits kann eine ITC-Zelle in einem für die Station reservierten Zeitschlitz ankommen und somit während des Ringumlaufs fälschlicherweise nicht entfernt worden sein. Andererseits darf in einem nicht als „available“ markierten Zeitschlitz keine ATC-Zelle enthalten sein. In beiden Fällen wird die ankommende Zelle verworfen und der Zeitschliz als leer betrachtet.

Own	ITCEmpty	ITC-Zelle senden
J	N	J
sonst		N

J: Ja N: Nein X: Don‘t Care

Tabelle A-4: Steuerfunktion für den ITC-Sendepuffer

Busy	Isoc	Avail	Own	IBEmpty	ATC Empty	ITC Empty	CCZero	AC	ATC-Zelle senden
N	X	J	N	J	N	X	N	X	J
N	X	X	J	J	N	J	N	< ACT	J
J	J	X	J	J	N	J	N	< ACT	J[a]
J	N	N	J	J	N	J	N	< ACT	J[b]
sonst									N

J: Ja N: Nein X: Don‘t Care

Tabelle A-5: Steuerfunktion für den ATC-Sendepuffer

a. Fehlerfall: Eine ITC-Zelle wurde während eines Ringumlaufs nicht entfernt
b. Fehlerfall: Ein nicht als „Available“ markierter Zeitschlitz der Station ist belegt

Die Steuerfunktion für die Sendepuffer wird nicht direkt, sondern kodiert an die Sende-Einheit weitergegeben. Ein Signal „Senden“ (Bild 4-14) zeigt an, daß aus einem der Sendepuffer eine Zelle zu senden ist. Die Art der Zelle und damit die Auswahl des Sendepuffers erfolgt durch Signalisierung innerhalb des Zeitschlitzes über das Isoc-Bit, das die Protokollsteuerung entsprechend der Art der zu sendenden Zelle setzt.

A.1.3 Setzen der Bits im Zeitschlitz-Header

Das Available-Bit wird nur bei Zeitschlitzen modifiziert, die für die Station reserviert sind. Tabelle A-6 enthält die Ausgangsfunktion für das Available-Bit. Solange der Stand des Available Counter unter dem Schwellwert ACT liegt, wird jeder für die Station reservierte Zeitschlitz als „available“ markiert. Ebenso wird Avail gesetzt, wenn der Schwellwert zwar überschritten ist, die Station aber dennoch eine ATC-Zelle sendet.

Busy	Isoc	Avail	Own	IBEmpty	ITCEmpty	AC	Avail neu
X	X	a	N	X	X	X	a
X	X	X	J	X	X	< ACT	J
J	N	J	J	J	J	X	J
X	X	X	J	N	J	X	J
sonst							N

J: Ja N: Nein X: Don‘t Care a: Wert von Avail am Eingang

Tabelle A-6: Steuerfunktion für das Available-Bit

Das Busy-Bit wird immer dann gesetzt, wenn ein abgehender Zeitschlitz eine Zelle enthält. Dies trifft immer dann zu, wenn eine Zelle aus dem Einfügepuffer entnommen wird (Tabelle A-3), aus dem ITC- oder ATC-Sendepuffer gesendet wird (Tabellen A-4 und A-5) oder aber am Einfügepuffer vorbei direkt die Station durchläuft. Die resultierende Funktion ist in Tabelle A-7 zusammengefaßt.

Busy	Isoc	Avail	Own	IBEmpty	ATC Empty	ITC Empty	CCZero	AC	Busy neu
N	X	J	N	J	N	X	N	X	J
N	X	J	N	N	X	X	X	X	J
N	X	X	J	J	N	J	N	< ACT	J
J	J	X	J	J	N	J	N	< ACT	J
J	N	N	J	J	N	J	N	< ACT	J
X	X	X	J	N	X	J	X	X	J
J	N	J	N	X	X	X	X	X	J
J	N	J	J	J	X	J	X	X	J
X	X	X	J	X	X	N	X	X	J
J	J	X	N	X	X	X	X	X	J
sonst									N

J: Ja N: Nein X: Don‘t Care

Tabelle A-7: Steuerfunktion für das Busy-Bit

Das Active-Bit wird dann gesetzt, wenn die Station Zellen in ihrem Einfügepuffer hat oder Zellen im ATC-Sendepuffer bereitstehen und der Credit Counter nicht Null ist. Ansonsten wird das Active-Bit, wie in Tabelle A-8 dargestellt, nicht verändert.

Active	IBEmpty	ATCEmpty	CCZero	Active neu
X	N	X	X	J
X	X	N	N	J
a	sonst			a

J: Ja N: Nein X: Don‘t Care a: Wert von Active am Eingang

Tabelle A-8: Steuerfunktion für das Active-Bit

A.1.4 Steuerfunktion des Available Counter

Der Available Counter wird bei jedem Zeitschlitz, der für die Station reserviert ist, basierend auf den Eingangsgrößen eingestellt. Dabei führt ein ankommender Zeitschlitz mit gesetztem Available-Bit sowie das Lesen einer Zelle aus dem Einfügepuffer zum Dekrementieren des Available Counters, während das Schreiben einer Zelle in den Einfügepuffer sowie ein abgehender Zeitschlitz mit gesetztem Available-Bit ein Inkrementieren des Available Counters nach sich ziehen. Die Ergebnisse der vier möglichen Vorgänge werden summiert und ergeben so den neuen Stand des Available Counters gemäß Tabelle A-9, in der alle erlaubten Kombinationen der Eingangsgrößen aufgeführt sind. Zur besseren Übersicht werden Zwischenergebnisse aus den vorherigen Protokollsteuerfunktionen als Eingangsgrößen genutzt.

Own	Avail	Einfügepuffer: Zelle schreiben (Tabelle A-2)	Einfügepuffer: Zelle lesen (Tabelle A-3)	Avail neu (Tabelle A-6)	AC neu
J	N	N	J	J	=
J	N	N	N	N	=
J	N	N	N	J	+1
J	J	N	J	J	-1

Tabelle A-9: Steuerfunktion für den Available Counter

Own	Avail	Einfügepuffer: Zelle schreiben (Tabelle A-2)	Einfügepuffer: Zelle lesen (Tabelle A-3)	Avail neu (Tabelle A-6)	AC neu
J	J	N	N	N	-1
J	J	N	N	J	=
J	J	J	J	J	=
J	J	J	N	N	=
J	J	J	N	J	+1
N	N	N	N	N	=
N	J	N	N	J	=
N	J	N	J	J	-1
N	J	J	J	J	=

J: Ja N: Nein +1/-1/= : inkrementieren/dekrementieren/konstant

Tabelle A-9: Steuerfunktion für den Available Counter

A.1.5 Steuerfunktion des Credit Counter

Der Credit Counter wird immer dann dekrementiert, wenn die Station eine ATC-Zelle sendet. Er wird auf den Anfangswert CCI gesetzt, wenn in einem Zeitschlitz das Reset-Bit im Header gesetzt ist. Tabelle A-10 enthält die Steuerfunktion für den Credit Counter.

Reset	ATC-Zelle senden (Tabelle A-5)	CC neu
N	N	=
N	J	-1
J	X	= CCI

J: Ja N: Nein
X: Don‘t Care CCI: Credit Counter Initial Value

Tabelle A-10: Steuerfunktion für den Credit Counter

A.1.6 Anzahl der Rahmen auf dem Ring

Da für den Aufbau einer ITC-Verbindung die Anzahl der Rahmen auf dem Ring FOR bekannt sein muß, wird diese Zahl in jedem Rahmen-Header übertragen. Die Protokollsteuerung übernimmt außerhalb der eigentlichen zeitschlitzbasierten Funktionen die Aufgabe, FOR in einem Register zum Auslesen über die Bus-Schnittstelle bereitzustellen.

A.2 Verwaltung isochroner Kanäle

Beim Aufbau einer ITC-Verbindung werden der anfordernden Station von der Monitor- und Masterstation Z Zeitschlitze für die neue Verbindung zugeteilt, deren Position innerhalb eines Rahmens jeweils durch ihre Nummer eindeutig gekennzeichnet ist. Die Station schreibt die Nummern der Zeitschlitze in ihren Verbindungsspeicher (Bild 4-14) und addiert zum Available Counter das Produkt aus der Anzahl der Rahmen auf dem Ring FOR und der Anzahl der Zeitschlitze je Rahmen Z. Nicht reservierte Zeitschlitze sind immer als „available" markiert, damit sie für Asynchronverkehr zur Verfügung stehen. Der Available Counter wird mit dem Produkt $FOR \cdot Z$ um die Anzahl von Zeitschlitzen korrigiert, die neu für die Station reserviert und als „available" markiert sind.

Die Station kann erst dann in der neuen Verbindung zu senden beginnen, wenn der Available Counter AC den Schwellwert ACT unterschritten hat. Ein mögliches Überschreiten der Schwelle durch die Addition von $FOR \cdot Z$ ist zwar möglich, führt jedoch nicht zu einem Überlauf des Einfügepuffers, solange in den reservierten Zeitschlitzen keine ITC-Zellen gesendet werden.

Beim Verbindungsabbau werden für alle zurückzugebenden Zeitschlitze zunächst im Verbindungsspeicher die Einträge von „Own" auf „Pending" modifiziert. Die Zeitschlitze werden nicht mehr für ITC-Verkehr genutzt. Ein „pending" Zeitschlitz wird entgegen den beschriebenen Protokollsteuerfunktionen auch dann als „available" markiert, wenn der Available Counter den Schwellwert überschreitet, und kann dann für ATC-Verkehr genutzt werden. Anschließend werden die Zeitschlitznummern aus dem Verbindungsspeicher entfernt und das Produkt $FOR \cdot Z$ vom Available Counter subtrahiert. Schließlich wird der Monitor- und Masterstation der Verbindungsabbau signalisiert.

Die Schwelle ACT muß derart festgelegt werden, daß der Einfügepuffer nicht überläuft, wenn in allen Zeitschlitzen, die für die Station reserviert und als „available" markiert sind, ATC-Zellen ankommen und in den Einfügepuffer geschrieben werden, während aufgrund einer Überlastsituation keine Zellen dem Einfügepuffer entnommen werden. Der größtmögliche Wert von ACT ist somit die Anzahl der Zellen, die der Einfügepuffer fassen kann.

Der Einfügepuffer wiederum sollte mindestens so viele Zellen fassen, wie Zeitschlitze auf dem Ring für die Station reserviert sind. In einem Beispielring mit FOR = 5 Rahmen auf dem Ring und $Z = 10$ reservierten Zeitschlitzen je Rahmen sollte der Einfügepuffer mindestens $FOR \cdot Z = 50$ Zellen oder, bei 56 Oktetts je Zelle, 2800 Oktetts fassen, damit bei leerem Einfügepuffer alle reservierten Zeitschlitze als „Available" markiert werden können.

A.3 Paketübertragung

Soll innerhalb der asynchronen Verkehrsklasse die Übertragung eines Pakets in mehreren aufeinanderfolgenden Zeitschlitzen ermöglicht werden, beschränkt sich die Signalisierung vom Sender zur Protokollverarbeitung (Bild 4-14) nicht auf die Pufferzustände ATCEmpty und ITCEmpty, sondern umfaßt ein weiteres Signal PacketTransmission, das während einer Paketübertragung aus dem ATC-Sendepuffer aktiv ist.

Ist N die Anzahl der Zeitschlitze, die für die Übertragung des größten zusammenhängenden Pakets benötigt werden, so muß der Einfügepuffer zusätzlich über N - 1 Pufferplätze für Zellen verfügen. Ferner wird ein Zustandsregister PacketMem benötigt, das die Zugehörigkeit der zuletzt gesendeten ATC-Zelle zu einem Paket registriert. Tabelle A-11 beschreibt die Zustandsübergangsfunktion von PacketMem beim Eintreffen einer neuen ATC-Zelle in Abhängigkeit von den Bits „Packet", das die Zugehörigkeit einer ankommenden Zelle zu einer Paketübertragung kennzeichnet, und „End of Packet" (EOP), mit dem die letzte Zelle einer Paketübertragung markiert ist. Das Senden aus dem ATC-Sendepuffer ist immer nur dann möglich, wenn PacketMem = 0 gilt.

Packet	EOP	PacketMem
N	N	0
J	N	1
J	J	0

J: Ja N: Nein

Tabelle A-11: Zustandsübergangsfunktion von PacketMem beim Eintreffen einer ATC-Zelle

Die Sendefunktion für den ATC-Verkehr wird ferner derart modifiziert, daß nach dem Senden einer Zelle aus dem ATC-Sendepuffer weitere ATC-Zellen gesendet werden, solange das Signal PacketTransmission aktiv ist. Für die Zeitdauer der Paketübertragung hat also nicht mehr der Einfügepuffer Vorrang, sondern der ATC-Sendepuffer, sofern der abgehende Zeitschlitz weder eine ITC-Zelle enthält, noch als „unavailable" markiert ist. Alle ankommenden ATC-Zellen werden für die Dauer des Paketsendevorgangs im Einfügepuffer gespeichert, wobei der Available Counter entgegen der Funktion nach Tabelle A-9 nicht beim Schreiben in den Einfügepuffer inkrementiert wird, wenn gleichzeitig eine zu einem Paket gehörige Zelle aus dem ATC-Sendepuffer gesendet wird.