

Konzept eines 1-Chip ATM-Kopplers für 600 Mbit/sec

R. Medow
Institut für Nachrichtenvermittlung
und Datenverarbeitung
Universität Stuttgart
Seidenstr. 36, 7000 Stuttgart 1
Tel. (0711) 121-2474

T. Schwederski
Institut für Mikroelektronik Stuttgart
Allmandring 30a, 7000 Stuttgart 80
Tel. (0711) 685-5924

Kurzfassung

Ein 4x4-Koppelement wird beschrieben, das mit 7 Pufferplätzen mit wahlfreiem Mehrfachzugriff pro Eingang eine Maximallast von 85% erreicht. Eine reguläre Chipstruktur ermöglicht die Realisierung eines 600 Mbit/s Koppelementes auf einem Chip. Ein Prototyp in Semicustom-Technologie mit 150 Mbit/s wird zur Zeit entwickelt.

Einführung

Die steigende Nachfrage nach Telekommunikationsdiensten verschiedener Art, insbesondere auch nach Breitbandkommunikation wie z.B. Übertragung von Bewegtbildern, hat zu dem Wunsch nach einem Breitbandnetz geführt, das möglichst alle öffentliche Dienste integriert. Für dieses B-ISDN hat das CCITT als Übertragungs- und Vermittlungsprinzip ATM festgelegt [1], ein paketorientiertes Verfahren, das mit asynchronem Zeitmultiplex arbeitet. Der gemultiplexte Informationsfluß ist in Blöcke fester Länge (Zellen) aufgeteilt. Jede Zelle enthält außer der Nutzinformation noch einen Kopf, der angibt, zu welcher Verbindung die Zelle gehört.

Vermittlungsstellen lassen sich wegen der hohen Datenraten (150 Mbit/s und 600 Mbit/s am Benutzerzugang, höher zwischen den Vermittlungsstellen zur Ausnutzung des Bündelungsgewinns) und wegen des ATM-Prinzips nicht mehr in der herkömmlichen Weise aufbauen. Eine ATM-Vermittlungsstelle hat vielmehr die Form eines mehrstufigen Netzwerkes aus kleinen Koppelementen. Als Netzwerktopologie werden meist Delta-Netzwerke verwendet, um die eigentliche Vermittlungsfunktion auf das ganze Netzwerk verteilen zu können [2]. Die interne Organisation der individuellen Koppelemente ist von fundamentaler Bedeutung für die Leistungsfähigkeit der gesamten Vermittlungsstelle; dabei ist die Organisation der Paketpuffer besonders wichtig. Untersuchungen der Technologie-Randbedingungen und Simulationen ergaben das hier beschriebene 4x4-Koppelement mit nahezu ideal geringen Mittelwerten und Standardabweichungen der Durchlaufzeiten sowie einfacher Realisierbarkeit auf einem Chip.

Schaltelement-Architektur

Da Delta-Netzwerke nicht blockierungsfrei sind, müssen in jedem Koppelement Pufferspeicher vorhanden sein. Deren Anordnung und Zugriffsart hat einen besonderen

Einfluß auf den möglichen Durchsatz des Koppelementes. Mögliche Anordnungen der Pufferspeicher und deren Eigenschaften sind im folgenden beschrieben [4, 5].

Eingangspuffer. Die offensichtliche Lösung des Pufferungsproblems ist durch Einfügen eines FIFO-Puffers im Eingang des Elementes gegeben. Vorteile sind geringe Geschwindigkeitsanforderungen und einfache Kontrolllogik. Ist jedoch das erste Element im Puffer blockiert, können auch die nachfolgenden nicht weitergeleitet werden, auch wenn sie an freie Ausgänge gerichtet sind. Zwei Möglichkeiten zur Umgehung dieses Problems sind wahrfreier Zugriff, bei dem beliebige Elemente im Speicher ausgelesen werden können, und Mehrfachzugriff, bei dem in jedem Takt mehrere Zellen ausgelesen werden können. Beide Möglichkeiten können auch kombiniert werden (Verallgemeinerter Eingangspuffer, Abb. 1).

Ausgangspuffer. Von einem nachrichtentheoretischen Standpunkt her haben Koppelemente mit Puffern am Ausgang eine optimale Leistungsfähigkeit. Durch die erforderliche hohe Speicherbandbreite ist eine Implementierung in verfügbarer CMOS-Technik jedoch problematisch.

Zentralpuffer. Die Anforderungen an die Speichergöße kann durch einen für alle Puffer gemeinsamen physikalischen Speicher minimiert werden [6]. Der Nutzung dieses Konzeptes stehen eine sehr hohe Speicherbandbreite und komplexe Kontrolllogik entgegen.

Kreuzpunkt-Puffer. Werden Puffer an die Kreuzpunkte der Schaltmatrix platziert (ein Puffer pro Eingangs/Ausgangs-Paar), so ergibt sich ein sehr einfaches Element, das jedoch sehr hohe Speicheranforderungen hat. Wegen der Modularität ist dies zwar eine attraktive Lösung, läßt sich jedoch praktisch wegen des Speicherbedarfs nicht als größeres Element in VLSI-Schaltungen integrieren.

Eingangs- und Ausgangspuffer. Werden nur maximal zwei gleichzeitige Speicherzugriffe auf einen Ausgangspuffer zugelassen, so verringert sich die erforderliche Speicherbandbreite bei nur unwesentlichem Leistungsverlust. Es sind allerdings zusätzliche Eingangspuffer erforderlich, um die Paketverlustrate zu begrenzen. Eine solche Anordnung ist ein Kompromiß zwischen Komplexität der Kontrolllogik, interner Geschwindigkeit und Leistung.

Untersuchungen der Entwurfsalternativen

Entwicklungsziel sind eine Übertragungsgeschwindigkeit von 600 Mbit/s und eine Zeitverlustrate von 10^{-10} bei einer Last von 80%, bei möglichst geringem Mittelwert und Standardabweichung der Zelldurchlaufzeit. Das Schaltelement sollte auf einem einzelnen Chip Platz finden. Technologische Randbedingungen wie Chipgröße und Simulationen wurden verwendet, um einen sinnvollen Kompromiß zwischen Leistungsfähigkeit und Realisierbarkeit zu finden; die Platzierung der Puffer, die Pufferlänge und Arbitrierungsstrategien wurden untersucht. Eine ereignisgesteuerte Simulation auf Verhaltens-

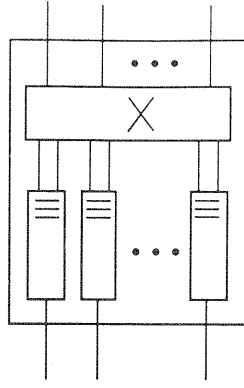


Abb. 1: Koppelement mit verallgemeinerten Eingangspuffern.

ebene wurde verwendet, und ein Netzwerk mit 64 Eingängen, konstruiert aus 4x4-Schaltenelementen, damit untersucht. Die Paket-Ankunft entsprach einem Bernoulli-Prozess, und es wurde symmetrische Verkehrslast eingesetzt. Backpressure wurde genutzt, um Zellverluste zu vermeiden.

Bei der Pufferstruktur konnten einige Entwurfsalternativen aufgrund der obigen Anforderungen unmittelbar ausgeschlossen werden. So haben FIFO-Puffer eine zu geringe Leistungsfähigkeit, werden Ausgangspuffer und Zentralspeicher wegen der hohen Speicherbandbreite nicht genutzt, und Kreuzpunktpuffer sind wegen des hohen Platzbedarfs ungeeignet. Verallgemeinerte Eingangspuffer zeigten sich als beste Alternative. Dabei sollten zwei Elemente gleichzeitig aus einem Eingangspuffer entnommen werden können. Abb. 2 zeigt für diesen Fall die Durchlaufzeit durch ein dreistufiges Netzwerk mit vorgeschalteten Puffern und Backpressure-Mechanismus in Abhängigkeit von der relativen Last bei verschiedenen Längen der Puffer in den Koppellementen. Zum Vergleich dient das ideale Verhalten nach den Formeln in [3].

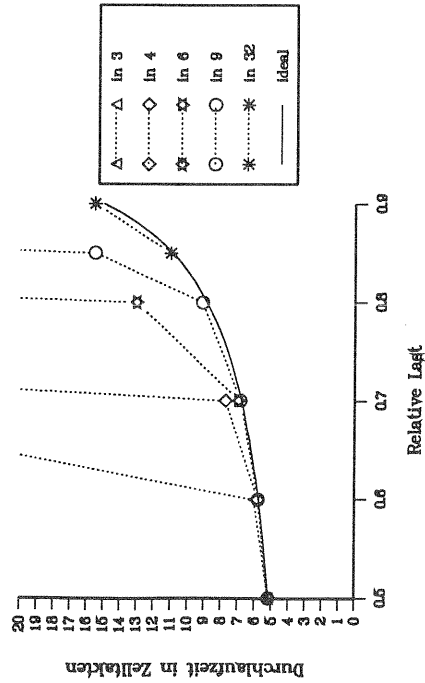


Abb. 2: Simulierte Durchlaufzeit in Abhängigkeit der Last bei unterschiedlicher Pufferanzahl.

Bei der Konfliktlösung wurden Methoden wie globale Ankunftszeit wegen der Komplexität nicht berücksichtigt. Drei Strategien (längste Eingangsschlange, lokale Zeit und zufällig) wurden simuliert. Die Auswahl der Zelle in der längsten Eingangsschlange zeigte sich als beste Lösung, denn im Vergleich zur zufälligen Konfliktlösung und lokaler Zeit führt dies zu einer verringerten Standardabweichung der Durchlaufzeit und zu deutlich kürzeren Puffern.

Um die Anzahl der Pufferspeicher bei der geforderten Zellverlustrate von weniger als 10^{-10} im einzelnen Koppellement klein halten zu können, wird Backpressure verwendet. Wie in Abb. 2 gezeigt, sind 7 Plätze für eine Maximallast von 85% ausreichend. Größere Warteschlangen vor den Eingängen des Netzwerkes sind allerdings nötig, um die Zellenverlustrate am Netzwerkeingang klein zu halten. Die Größe des

Pufferelementes wird auf 4 Ein- und Ausgänge festgelegt, um eine Implementierung auf einem Chip zu ermöglichen.

Realisierungs-Aspekte

Mit der im folgenden spezifizierten Struktur ergeben sich nahezu ideal geringe Mittelwerte und Standardabweichungen der Durchlaufzeiten sowie einfache Realisierbarkeit auf einem Chip (siehe Abb. 3)

- Pufferspeicher mit wahlfreiem Mehrfachzugriff befinden sich an den Eingängen.
- Im Konfliktfall werden Zellen aus der längeren Eingangswarteschlange bevorzugt.
- Ein Backpressure-Mechanismus wird verwendet, d.h. das Koppellement schickt keine Zellen an seinen Nachfolger, wenn dieser signalisiert, daß seine Puffer voll sind.

Die Puffer werden als RAM implementiert. Da aus den Eingangspuffern Zellen von beliebigen Plätzen entnommen werden können, lassen sich die Puffer nicht mehr als FIFO verwalten, sondern es ist eine komplexere RAM-Steuerung notwendig, die jeden freien und belegten Platz einzeln verwaltet, die Übersetzung von Warteschlangenposition auf RAM-Adresse vornimmt und für ankommende Zellen die RAM-Adresse eines freien Platzes zur Verfügung stellt. Eine solche Steuerung läßt sich nach Art eines Schieberegisters mit einer Stufe pro Pufferplatz aufbauen. Jede Stufe enthält Flipflops mit vorgeschalteten Multiplexern und Kontrolllogik, die bestimmt, ob die Stufe sich wie eine Schieberegisterstufe verhält, den gegenwärtigen Zustand beibehält oder mit Informationen von außerhalb geladen wird. Die Flipflops jeder Stufe enthalten u.a. eine

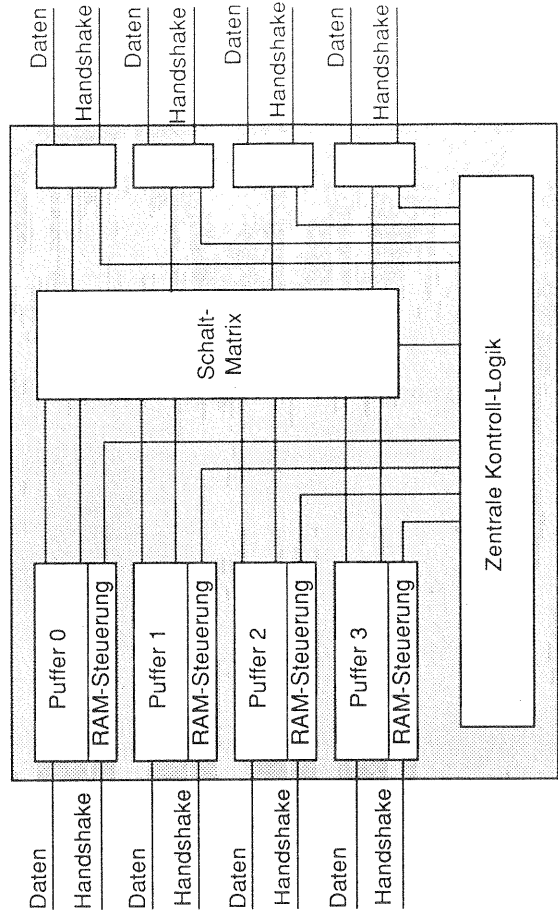


Abb. 3: Blockdiagramm des Koppelfeldes.

RAM-Adresse und ein Belegbit. Beim Reset werden alle Belegbits zurückgesetzt und die Stufen mit verschiedenen RAM-Adressen geladen. Wird eine Zelle in den Speicher eingetragen, so legt die niedrigste Stufe, deren Belegbit gelöscht ist, ihre RAM-Adresse auf einen Bus und setzt ihr Belegbit. Beim Auslesen werden zyklisch geschoben und die Belegbits mit 0 aufgefüllt. Eine solche RAM-Steuerung ist auf der Ebene von Flipflops und logischen Gleichungen beschrieben und simultan verifiziert worden.

Damit bei der Implementierung als integrierter Baustein eine hohe Geschwindigkeit erzielt werden kann, müssen kurze Leitungslängen, d.h. eine möglichst kompakte Anordnung von RAM und Koppelfeld, gewährleistet sein. Erschwert wird diese Randbedingung durch den wahlfreien Mehrfachzugriff. Eine optimierte Anordnung zeigt Abb. 4; hierbei werden die RAM-Zellen, die Teile einer Zelle aufnehmen sollen, untereinander in einer Spalte angeordnet. Die Anzahl der Spalten entspricht der Anzahl der Pufferplätze pro Eingang (7 in Abb. 4). Am unteren Ende des RAM-Blockes stehen die Informationen der eingespeicherten Zeilen zur Verfügung. Sie werden über ein verteiltes Koppelfeld ausgewählt und an die gewünschten Ausgänge weitergeleitet. Eingangsseitig werden die eintreffenden Daten an die Spalten des entsprechenden Eingangs-RAM angelegt und in die von der RAM-Steuerung bestimmten Spalte übernommen, indem die Schreibfreigabe nur für diese Spalte aktiviert wird. Diese Anordnung ergibt eine reguläre Struktur, die zusätzlich die Datenein- und -ausgänge gleichmäßig über die Chip-peripherie verteilt.

Bei einem vollständigen Koppellement ist eine Pufferlänge von 53 Bytes plus 3 Bytes für das Routing Tag erforderlich. Aufgrund der niedrigen Pufferanzahl und der regulären Struktur kann ein 4x4-Koppellement auf einem einzelnen Chip sowohl in Full-Custom als auch in Semicustom-Technologie gefertigt werden. Bei einer 40 MHz-Taktrate müssen die Eingangs- und Ausgangssignale 16 Bit breit sein, um die erforder-

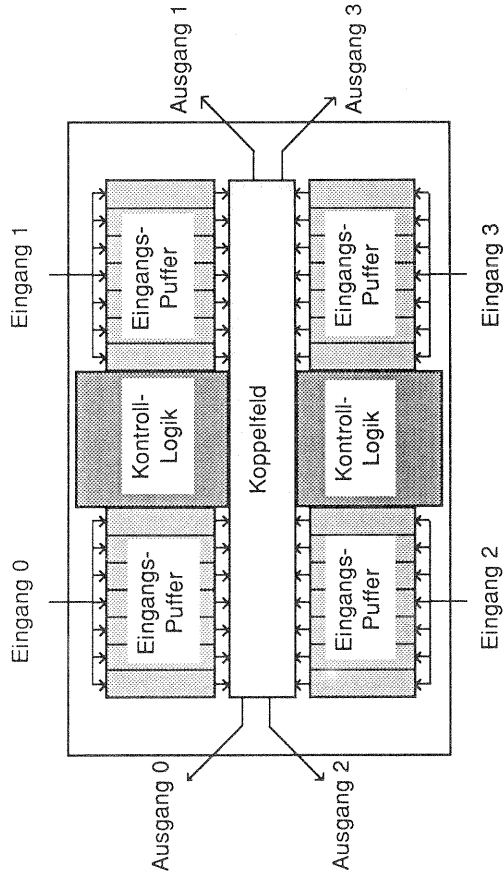


Abb. 4. Layout-Struktur des 4x4 ATM Koppelchips.

liche 600 Mbit/sec Datenrate zu erzielen; das Gehäuse hat dann ca. 200 Pins. Eine Prototyp-Implementierung wird zur Zeit mit der IMS Gate Forest Semicustom-Technologie [7] durchgeführt. Der Gate Forest ist ein fortgeschrittenes Sea-of-Gates-Konzept der zweiten Generation, das für statische und dynamische Logik sowie RAMs optimiert ist. Es erlaubt ähnlich hohe Integrationsdichten wie Full-Custom-Entwürfe [8]. Da die Semicustom-Implementierung der Konzeptverifikation dient, wird aus Kostengründen die Pufferstruktur auf 4 Puffer pro Eingang mit einer Breite von 8 Bit und einer Länge von 16 Bytes pro Paket reduziert. Ein 20 MHz-Takt ist vorgesehen, um eine Datenrate von 160 Mbit/sec zu erreichen.

Zusammenfassung

Die Realisierung eines Koppellementes als Single-Chip-VLSI-Baustein kann nicht nur unter rein nachrichtentechnischen Aspekten gesehen werden, da technologische Randbedingungen wie Chipgröße und Geschwindigkeitsgrenzen einen wesentlichen Einfluß auf die Realisierbarkeit haben. Diese Aspekte wurden beim hier vorgestellten Entwurf berücksichtigt, und ein Koppellement wurde vorgestellt, das mit sieben Eingangspuffern bei wahlfreiem Mehrfachzugriff fast optimale Eigenschaften der Durchlaufzeit und Standardabweichung besitzt. Eine optimierte VLSI-Architektur für dieses Konzept wurde vorgestellt, und ein Prototyp befindet sich am IMS in Entwicklung.

Literatur

- [1] CCITT: "Draft Recommendation I.121: On the Broadband Aspects of ISDN"; COM XVIII-R55(C); Seoul, Februar 1988
- [2] M. N. Huber, E. P. Rathgeb, T. H. Theimer: "Self Routing Banyan Networks in an ATM Environment"; 9th International Conference on Computer Communications (ICCC); Tel Aviv, 1988, pp. 167-174
- [3] C. P. Kruskal, M. Snir, A. Weiss: "The Distribution of Waiting Times in Clocked Multistage Interconnection Networks"; 1986 Intern'l Conf. on Parallel Processing; 1986, pp 12-19
- [4] E. P. Rathgeb, T. H. Theimer, M. N. Huber: "ATM-Switches - Basic Architecture and their Performance"; International Journal of Digital and Analog Cabled Systems; Vol. 23, No. 3, June 1988, pp. 728-735
- [5] K. Rotheimer, D. Seeger: "Traffic Studies of Switching Networks for Asynchronous Transfer Mode (ATM)"; Proceedings of the 12th International Telegraphic Congress (ITC), June 1988, Paper 1.3A.5
- [6] H. Kuwahara, N. Endo, M. Ogino, T. Kozaki: "A Shared Buffer Memory Switch for an ATM Exchange"; Proceedings of the 1989 International Conference on Communication (ICC); 1989, pp. 118-122
- [7] M. A. Beunder, B. Höflinger, J. P. Kernhof: "New Directions of Semicustom Arrays"; IEEE Journal of Solid-State Circuits; Vol. 23, No. 3, June 1988, pp. 728-735
- [8] M. A. Beunder: Design and Analysis of Semi-Custom Architectures; Ph.D. Thesis, Twente University / Institute for Microelectronics Stuttgart, December 1989