

**Heute im Supercomputer -
morgen im Alltagscomputer
(Entwicklungen in der Hochleistungskommunikation)**

Thomas M. Stricker



Institut für Computersysteme
ETH Zentrum
CH-8092 Zürich

Antrittsvorlesung

April 22, 1997

Supercomputer - Alltagscomputer



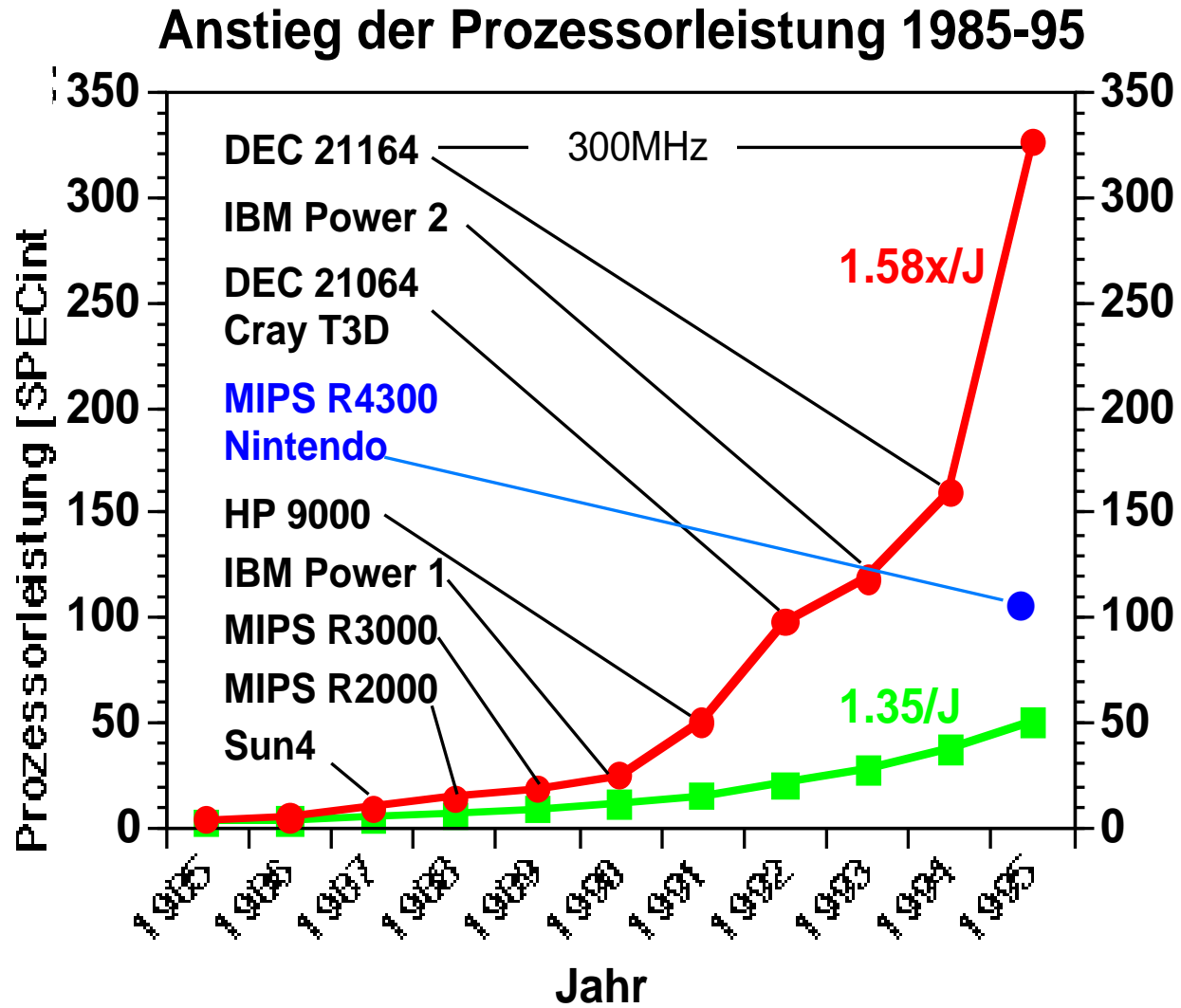
Was macht Super Mario möglich?

- **Nintendo 64-bit Video Spiel (1997)**
 - MIPS R4300 64-bit Mikroprozessor
 - Leistung: **ca. 100 SPECInt92** (bei 93.5 MHz)
 - 4 + “x” MB Speicher (500 MByte/s Bandbreite)
 - Video/Audio Co-Prozessoren
- **Einzelner Rechnerknoten der Cray T3D (1993)**
 - DEC 21064 64-bit Mikroprozessor
 - Leistung: **ca. 100 SPECInt92** (bei 150 MHz)
 - 64 MB Speicher (320 MByte/s Bandbreite)
 - Kommunikations Co-Prozessoren

Vom Supercomputer zum Alltagscomputer

- **Beispiel aus der Prozessorarchitektur**
- **Thesen zur Hochleistungs-Kommunikation**
- **Netzwerk mit vielen PCs (Networks of PCs)**
 - Kommunizieren wie in einem Supercomputer?
- **Vom Labor zum Markt (Rückschläge - Erfolge)**
- **Beispiel aus der Hochleistungskommunikation**
- **Suche nach geeigneten Anwendungen**
- **Schlussfolgerungen**

Rechenleistung wächst explosionsartig!



Ursachen für dieses Wachstum

Fortschritt im Halbleiter-Technologie Bereich

- höhere **Taktraten**, Miniaturisierung der **Bauelemente**, schnellere **Schaltkreise**

Entwicklungen im Technologie Bereich...

...gehen vom Alltagscomputer zum Supercomputer.

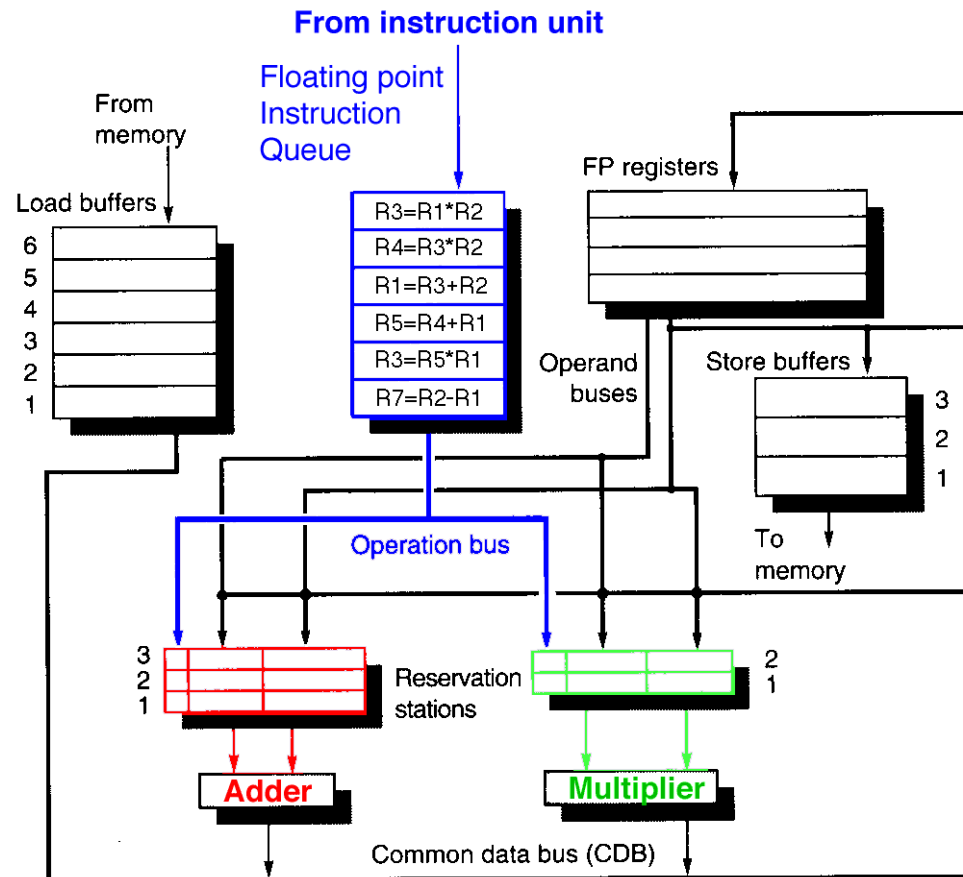
Fortschritt im Computer-Architektur Bereich

- **Speicherhierarchien**, **Pipeline-** und **Superskalar-**Verarbeitung von Instruktionen

Entwicklungen im Architektur Bereich...

...gehen vom Supercomputer zum Alltagscomputer.

Beispiel: Superskalare Verarbeitung



Scoreboard

FPU 1	$R3=R1*R2$	✓	
FPU 3	$R4=R3+R2$	✓	
FPU 2	$R1=R3*R2$		
FPU 1	$R7=R2-R1$		

Abkehr von der starren Abarbeitung einer einzigen Instruktion pro Takt.

IBM 360



Superskalare Verarbeitung

Idee:

- R.Tomasulu: IBM Research Laboratories 1967!

Auswertung:

- in der IBM 360/91, Baujahre ca. 1969-1973
 - Modell war kein grosser Erfolg
 - Idee war im Prinzip gut, aber die Komplexität war zu hoch und damals viel zu teuer.
 - Nachfolgemodelle:
Idee vorerst wieder fallengelassen.

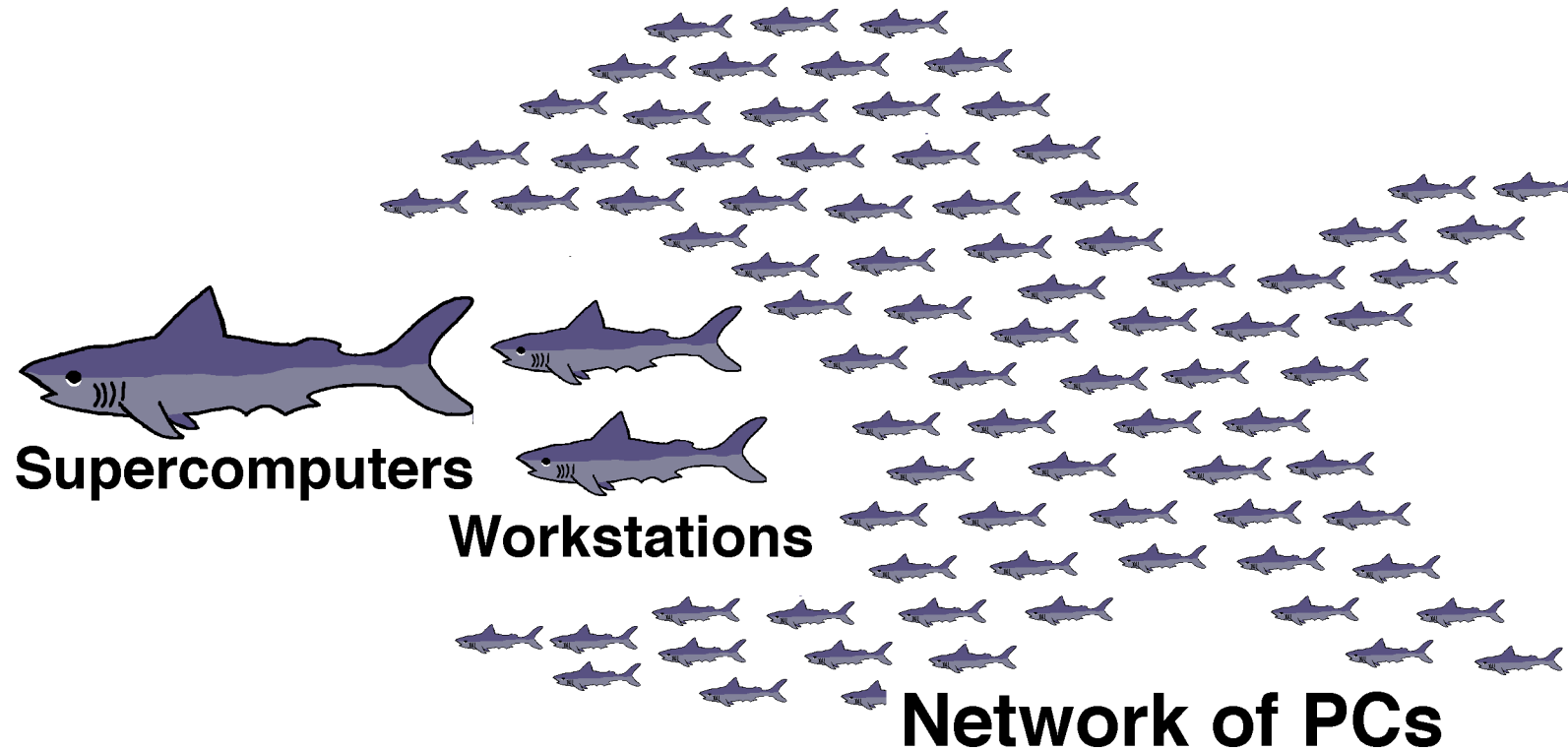
Als Idee ***nach 25 Jahren*** mitverantwortlich für die spektakulären Rechenleistungen auf den heutigen Chips.

Thesen zur Hochleistungskommunikation

Massiv parallele *Supercomputer* verfügen bereits heute über ein **integriertes** Hochleistungskommunikationssystem (dient zur Verbindung der einzelnen Prozessoren).

Gelingt es uns **kostengünstige** Kommunikationssysteme **mit hoher Leistung** zu bauen, so können auch *Alltagscomputer* in Zukunft effizient und auf natürliche Weise miteinander kommunizieren.

Networks of PCs



Neue “Ökologie” der Hochleistungs-Rechner?

Networks of PCs:

(viele kleine Fische fressen den Grossen?)

Spricht ganz eindeutig dafür:

- CPU Preis/Leistung - vgl. Nintendo.
- Grosse Stückzahlen - vgl. Preisvorteile der PCs.

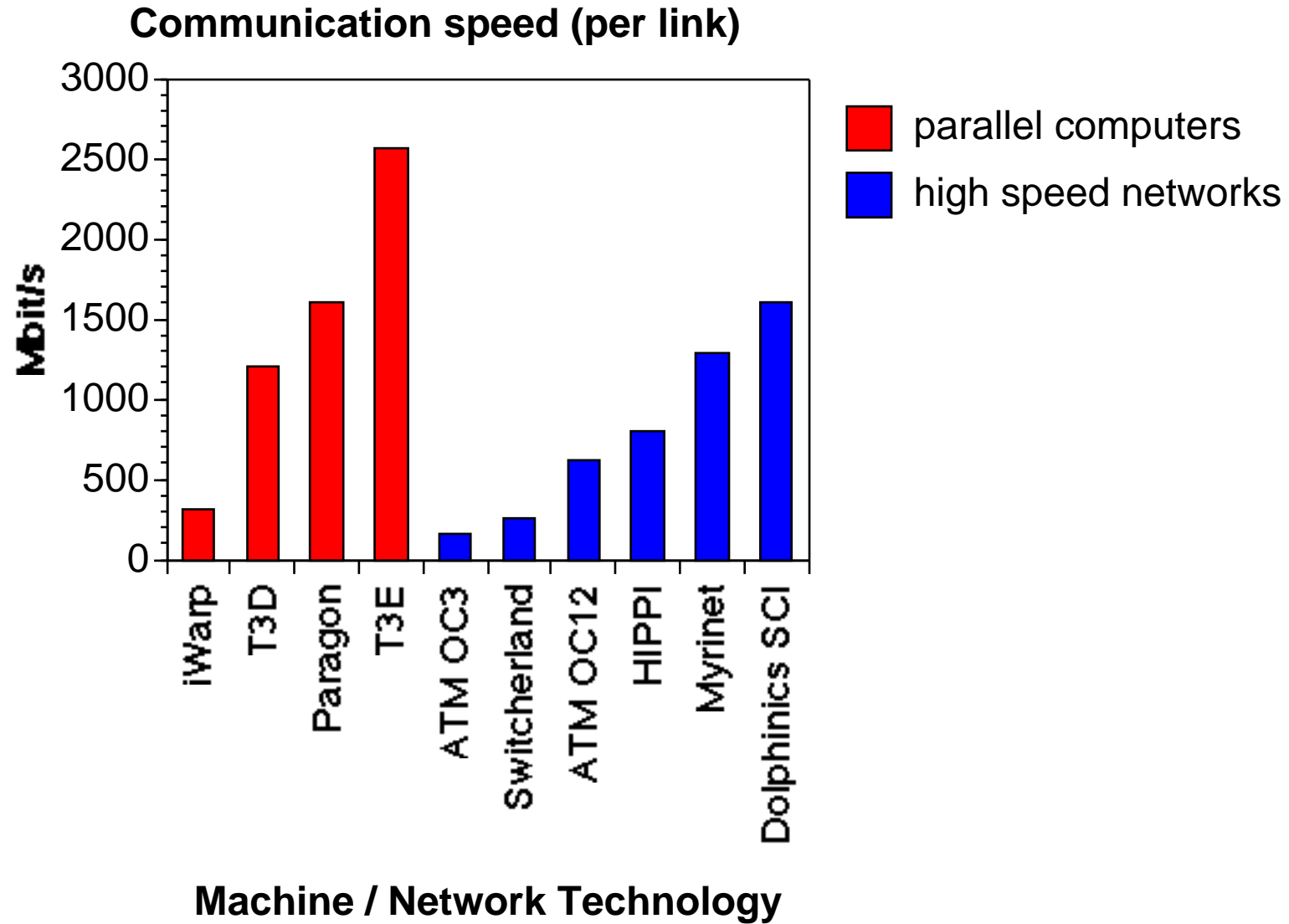
Mehr Forschung & Entwicklung nötig:

- Hochleistungskommunikation:
 - Hardware noch zu teuer!
 - System Software hat Probleme mit der Effizienz!

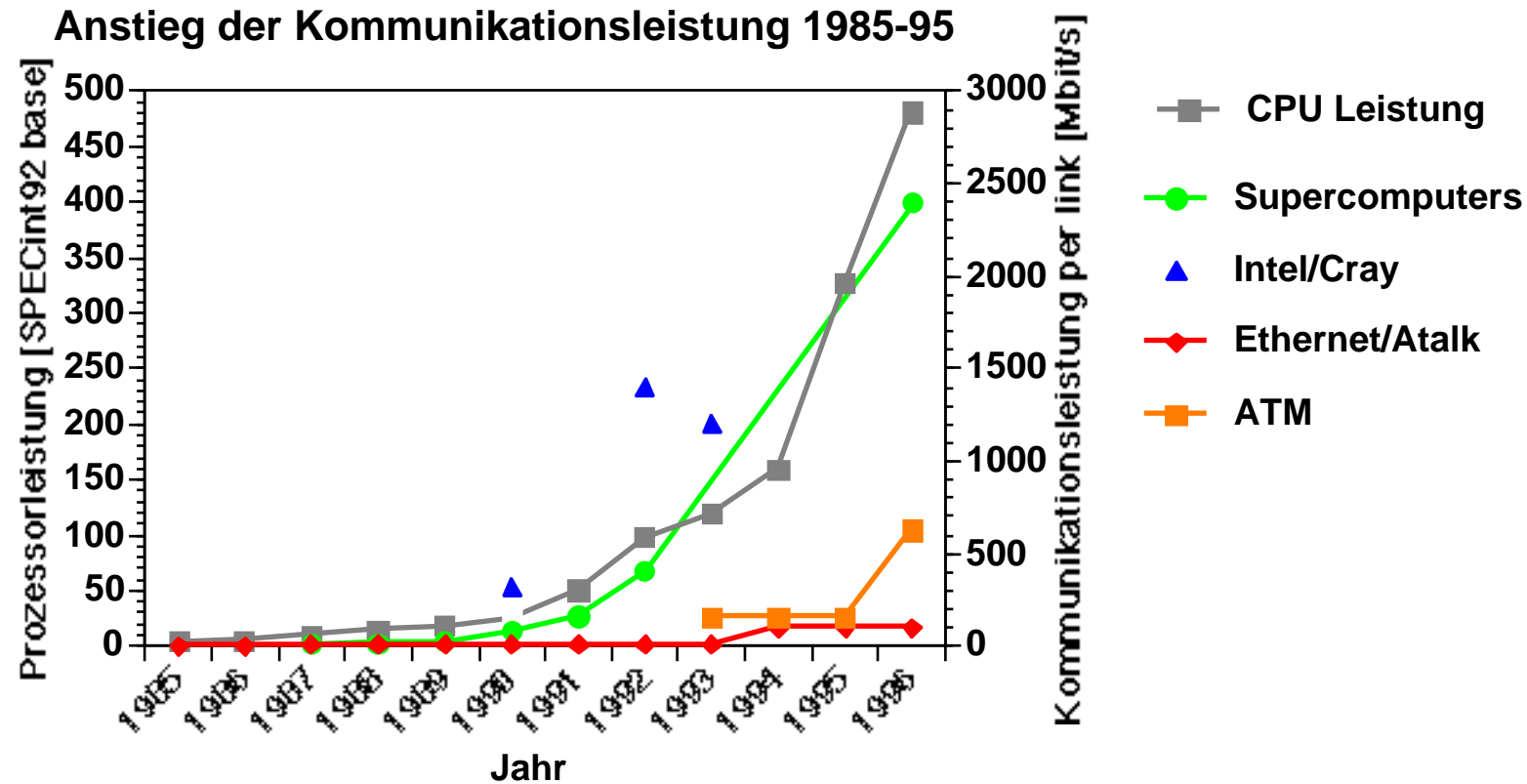
Die grosse Unbekannte:

- Betriebssystem - Ist WindowsNT dafür brauchbar?

Verfügbare Kommunikationstechnologien



Wachstum Kommunikationsleistung



Kommunizieren wie im Supercomputer

Telekommunikations Technologien:

- Durch Standards bestimmt
- Komplexität/Kosten unwichtig

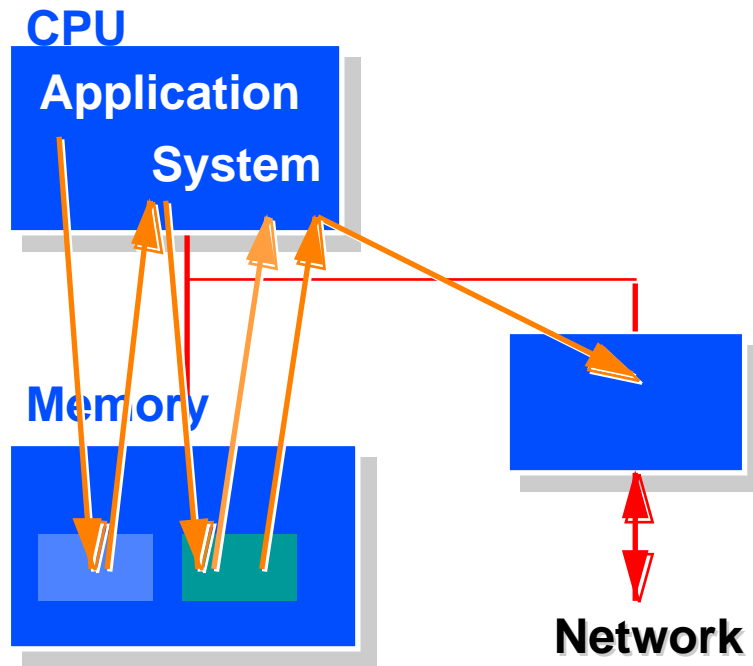
Parallel Computing Technologien:

- Durch Leistung bestimmt
- Komplexität/Kosten optimiert
- Oftmals geschlossenes System
(Technologische Abkürzungen sind erlaubt!)

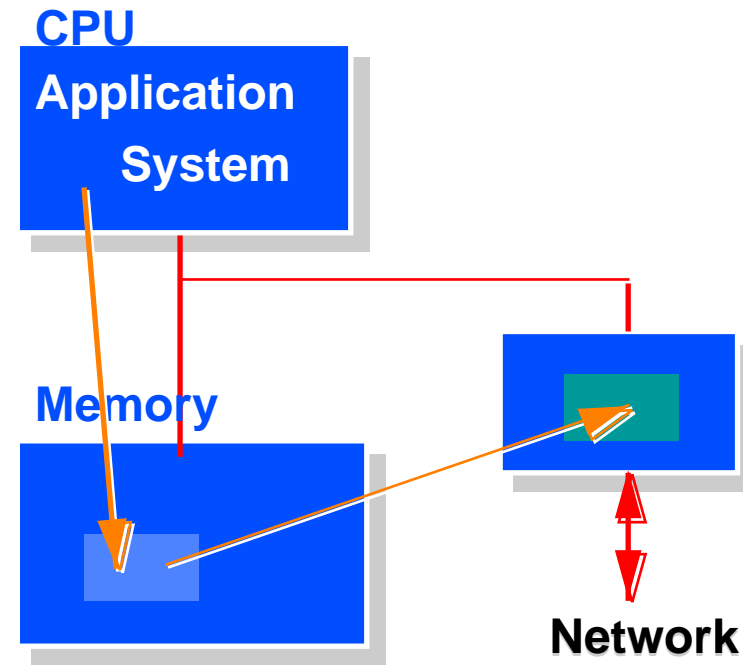
Probleme der Software Effizienz

- **Hohe Bandbreite** - mit PCs noch nicht möglich:
 - Daten werden oft kopiert
 - Kommunikation schneller als Speichersysteme
- **Kurze Latenzzeiten** - (ping/pong Zeit)
 - Aufwendige Verarbeitung jedes Daten-Paketes
 - Schutzmechanismen im Betriebssystem
- **Prozessorzuteilung / Synchronisation**
 - Senden / Rechnen / Empfangen gleichzeitig
 - Unterbrechungen beim Eintreffen von Daten

Problem der Software Effizienz



Traditional



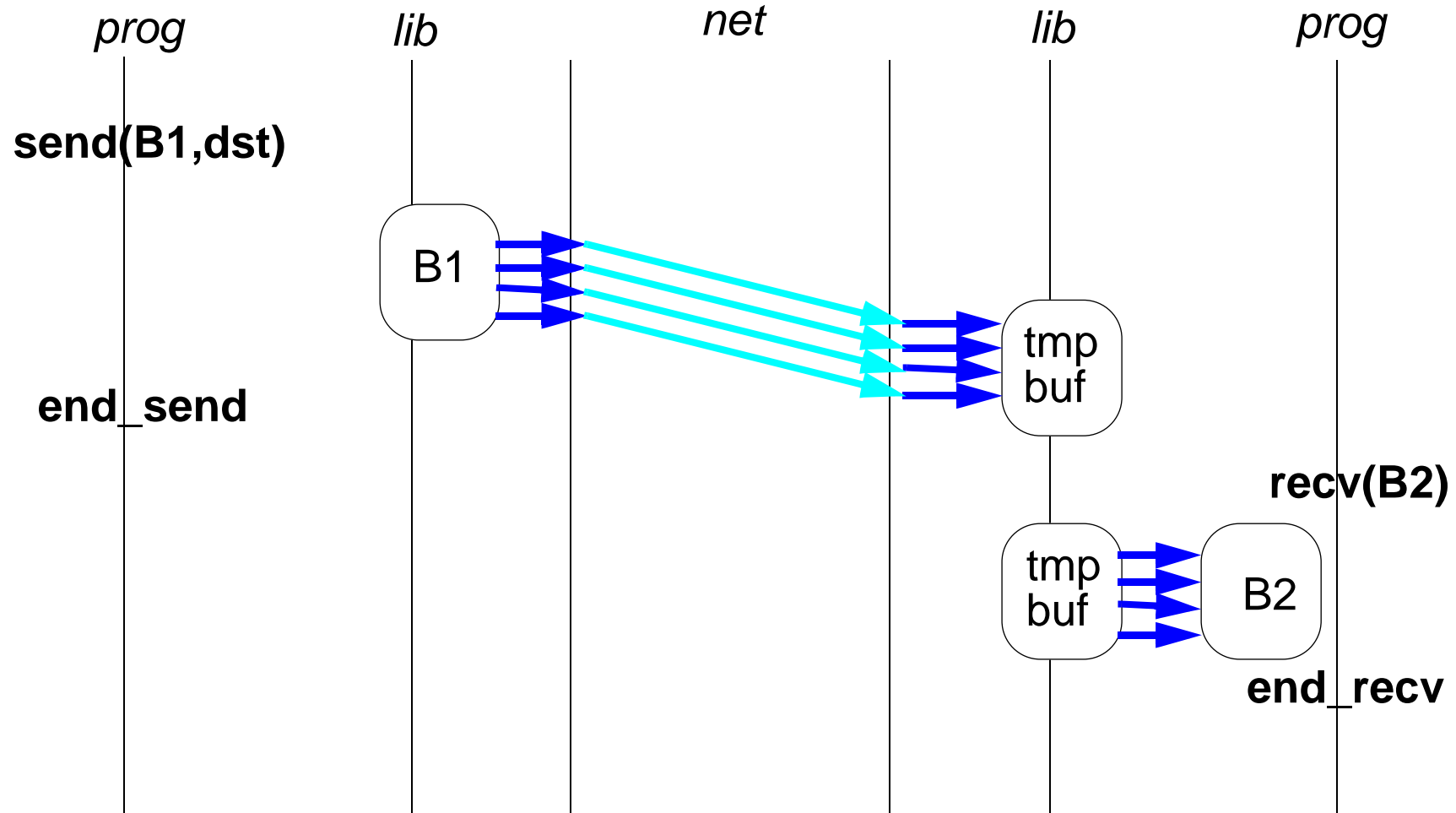
Single copy

Modelle der Kommunikation

Synchronisations-Semantiken:

- Meldungen (messages) bisher
 - **Post Modell**
(PVM, MPI, NX, SUNMOS)
- Verbindungen (connections) bisher
 - **Rendez-vous Modell**
(TCP/IP, CSP, Ada)
- Fernzugriffe (shared memory get/put) neu
 - **Deposit Modell**
(Cray T3D, Active Messages)

Post Modell



- **Meldungsvermittlung mit Zwischenspeicher**

Warum ein neues Modell?

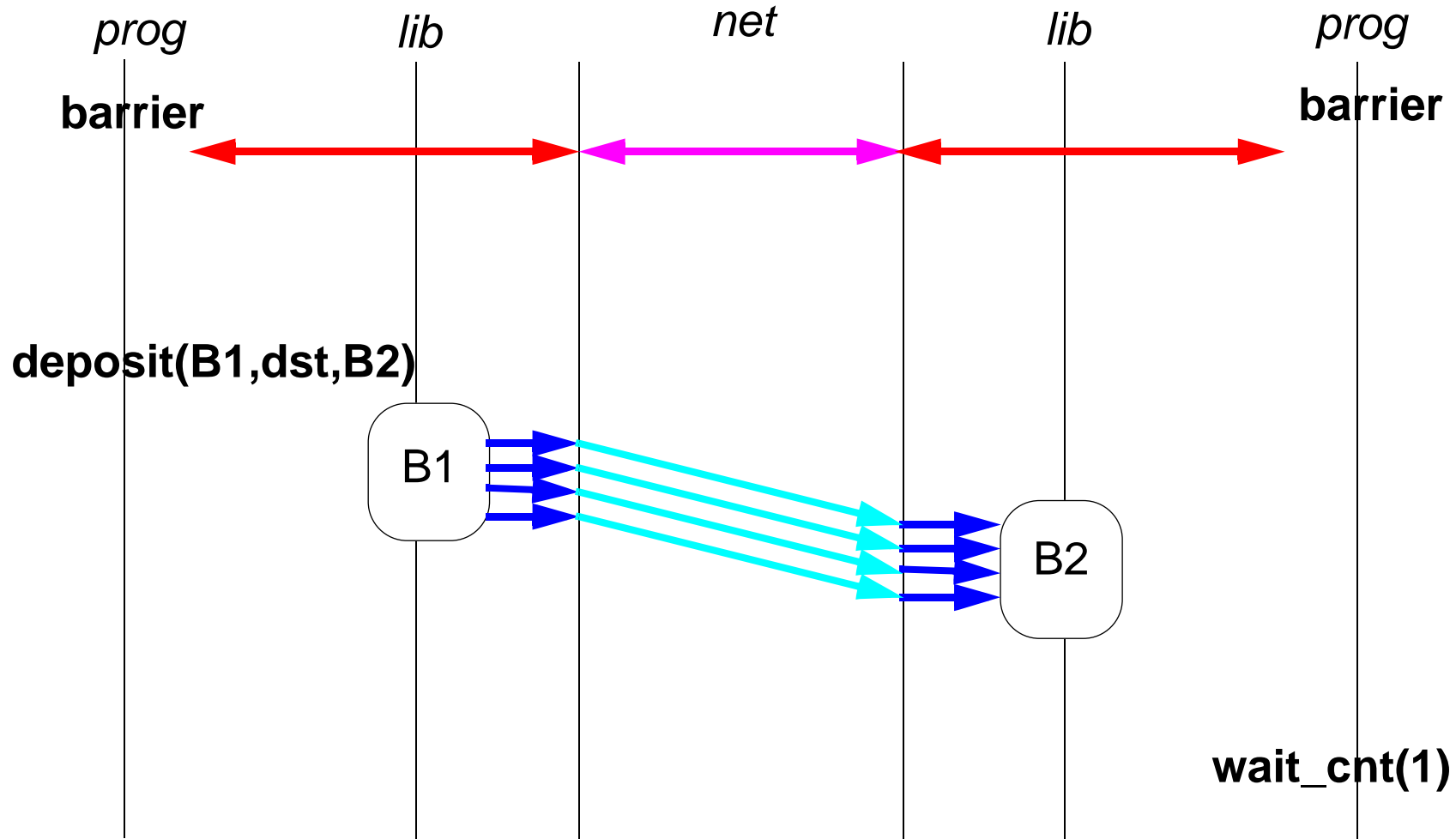
Zwei Arten von Meldungen im Netz:

- **Meldungen zum “Control Transfer” (Synchronisation)**
- **Meldungen zum “Daten Transfer”**

Aus unterschiedlicher Charakteristik der Meldungen folgt...

- **Trennung von Kontroll- und Daten-Transfer** im Modell und im System

“Direct Deposit” Modell



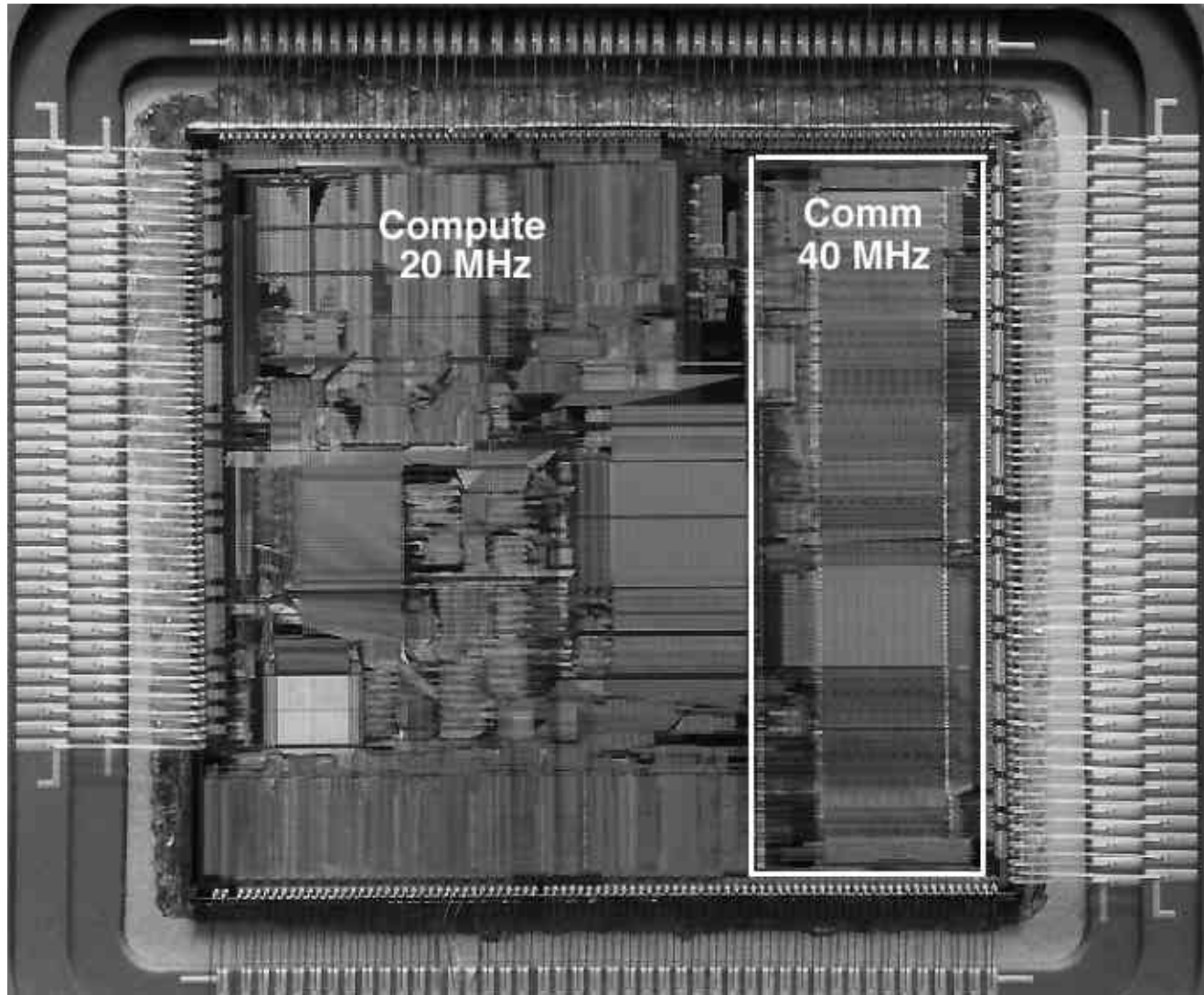
Vorteile von “direct deposit”

- **Aus beiden semantischen Modellen das Beste:**
 - Uneingeschränkten Meldungsfolge (Post Modell)
 - Direkte und effiziente Datentransfers (Rendez-vous Modell)
- **Deposit Modell**
 - Macht anwendungsspezifische Eigenschaften ausnutzbar.
 - Effiziente Mechanismen dafür sind in vielen Supercomputer bereits vorhanden (get/put).

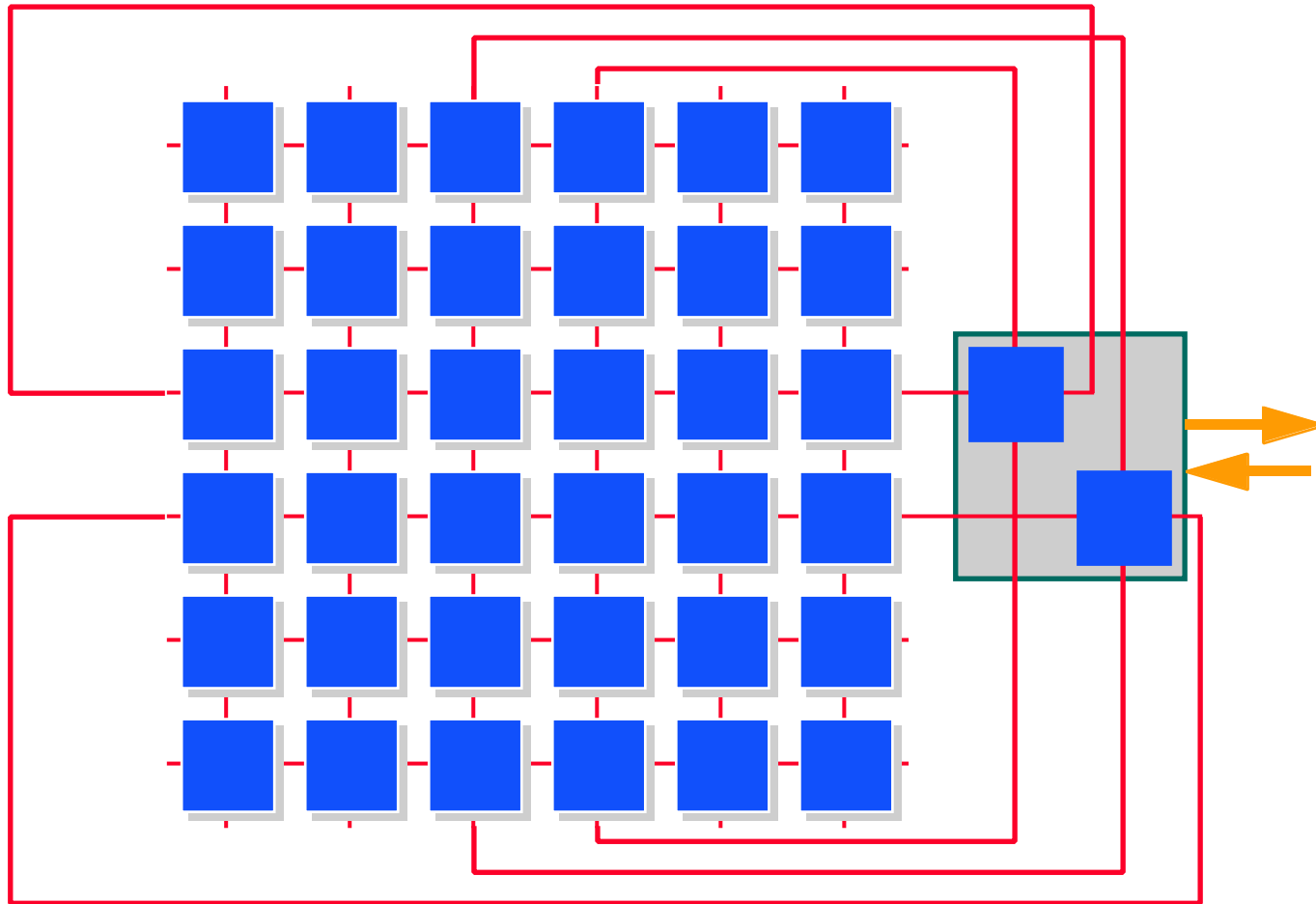
Vom Labor zum Markt

- **Von iWarp/Transputer zum idealen Parallelrechner**
 - Wissenschaftlich hervorragende Idee!
 - Enthielt aber technologische Fehler
 - 1990 noch kein genügender Markt vorhanden
- **Von Nectar zu ATM [FORE Systems Inc.]**
 - Pragmatische Technologie:
keine Chips - nur Standardkomponenten!
 - Technologie Transfer vom Supercomputing in
den Bereich Networking
 - Kommerzieller Grosserfolg im Networking

iWarp [Kung, Gross 1987-1993]



Grosses Potential in der Hochleistungskommunikation



Warum wurde aus dem “iWarp” kein Alltagscomputer?

- **Spezieller Mikroprozessor** (20-30 Designer)
 - Zweitrangige FPU Designer fürs Projekt
 - Keine Erfahrung = zwei Jahre Verspätung
 - Compiler blieb immer ein Problem
- **... dennoch ein Nachfolgeprojekt:**
 - Intel TriStar/Cavallino Chipset mit 6 x 400MB/s.
 - Co-Prozessor zum Intel 80686, PentiumPro
 - Mangelnde Vision - Intel verweigert Marketing

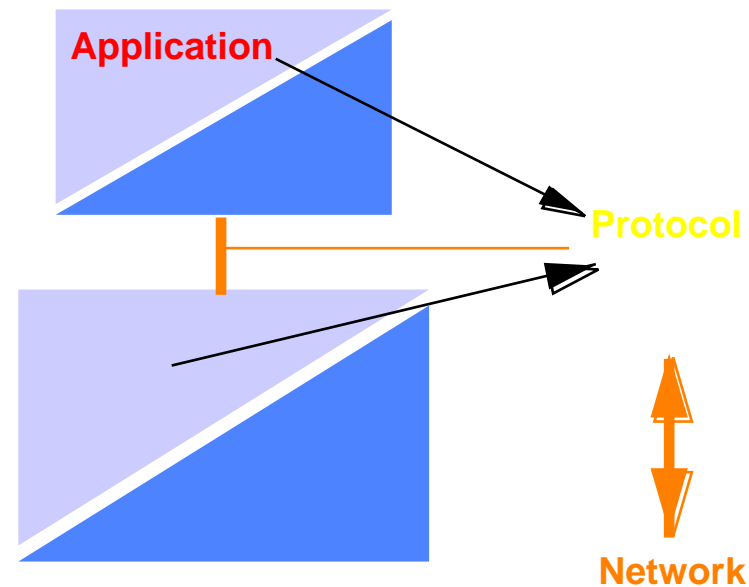
Intel: wenig Erfahrung mit Kommunikationsprozessoren - bessere Geschäfte - Zurückhaltung verständlich.

Nectar [Kung, Steenkiste 1987-1992]

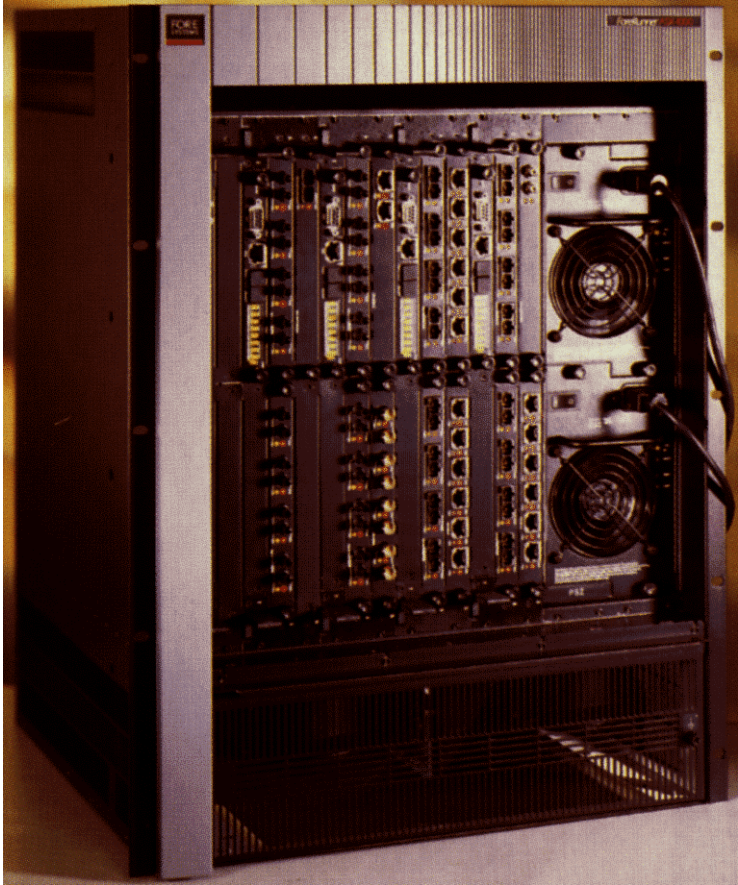
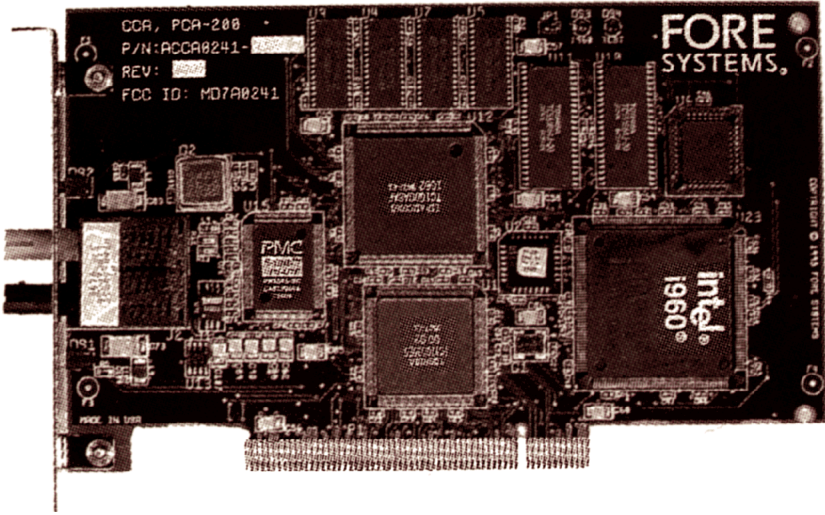
- Prototype eines Networks of Workstations
 - Ein Dutzend SUN4/300 Workstations
 - 100 MBit/s - Fiberoptisches Netz
- Umstellung von Mesh Interconnects auf Switch basierte Netzwerke (crossbar switches).
- Intelligente Interfaces im Hostrechner (CABs Communication Acceleration Boards)

Forschungsergebnisse

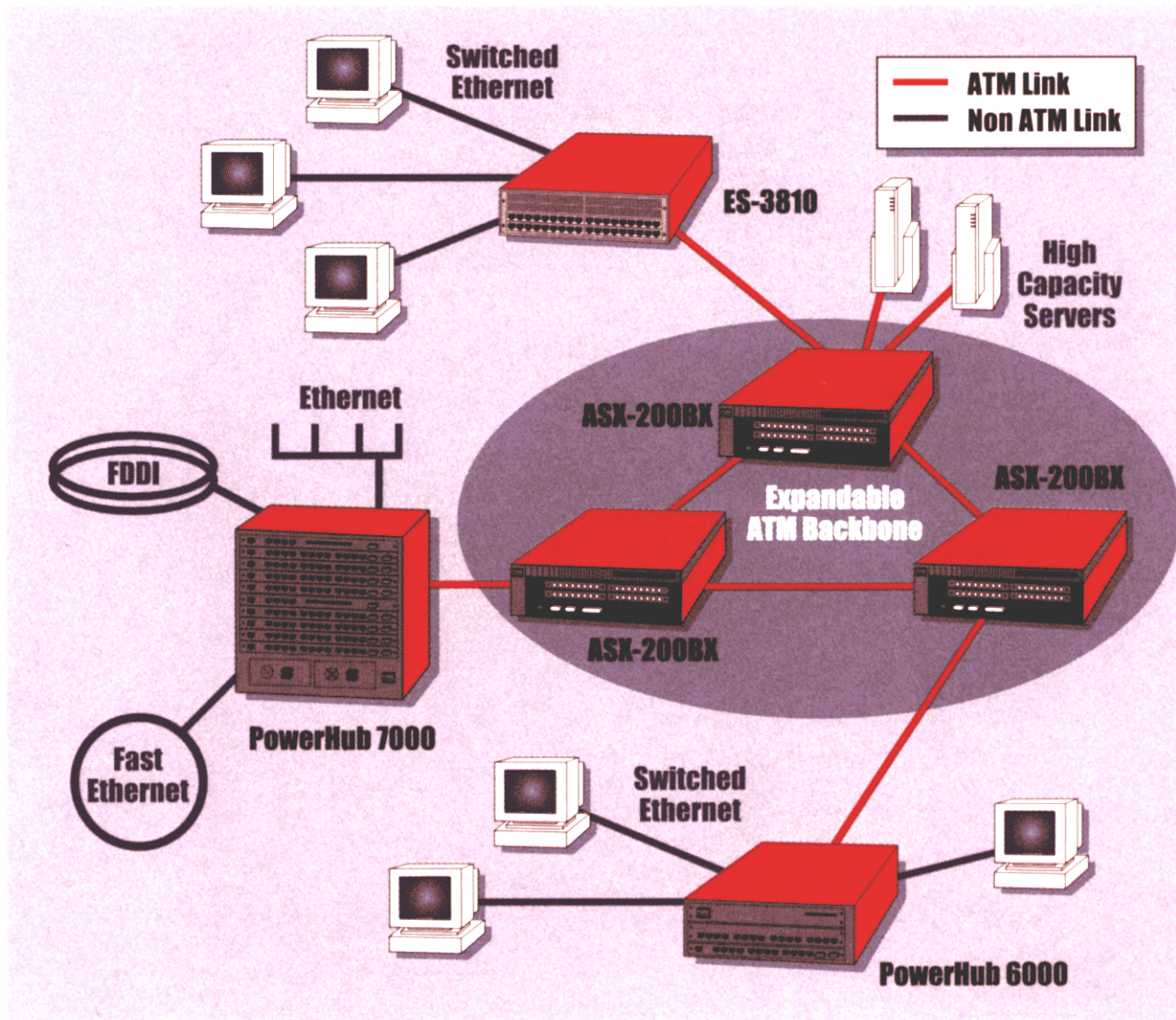
- **Shared-buffer API.**
- **Powerful adapter.**
 - » outboard buffering
 - » protocol processing
 - » application device access
- **0-copy communication**
 - » application places data directly on adapter
- **No protocol overhead on host.**
- **“Optimum” efficiency.**



ATM Networking Technologie



Typisches LAN/MAN Netzwerk



Wie fand “Nectar” den Weg in den Kommunikationsalltag (ATM)?

- **Loslösung von dem Parallelen Rechnen**
 - Transfer der Technologie in anderen Bereich
 - Switch Technologie wurde eine Notwendigkeit.
 - Interface Board statt Chipset
- **Standardisierung/Inter-Operability** (ATM Forum)
 - Reihe von Startup Companies, e.g. FORE Sys.
 - Networking Euphorie des Internet
- **Komplexität und Kosten von ATM [1997].**

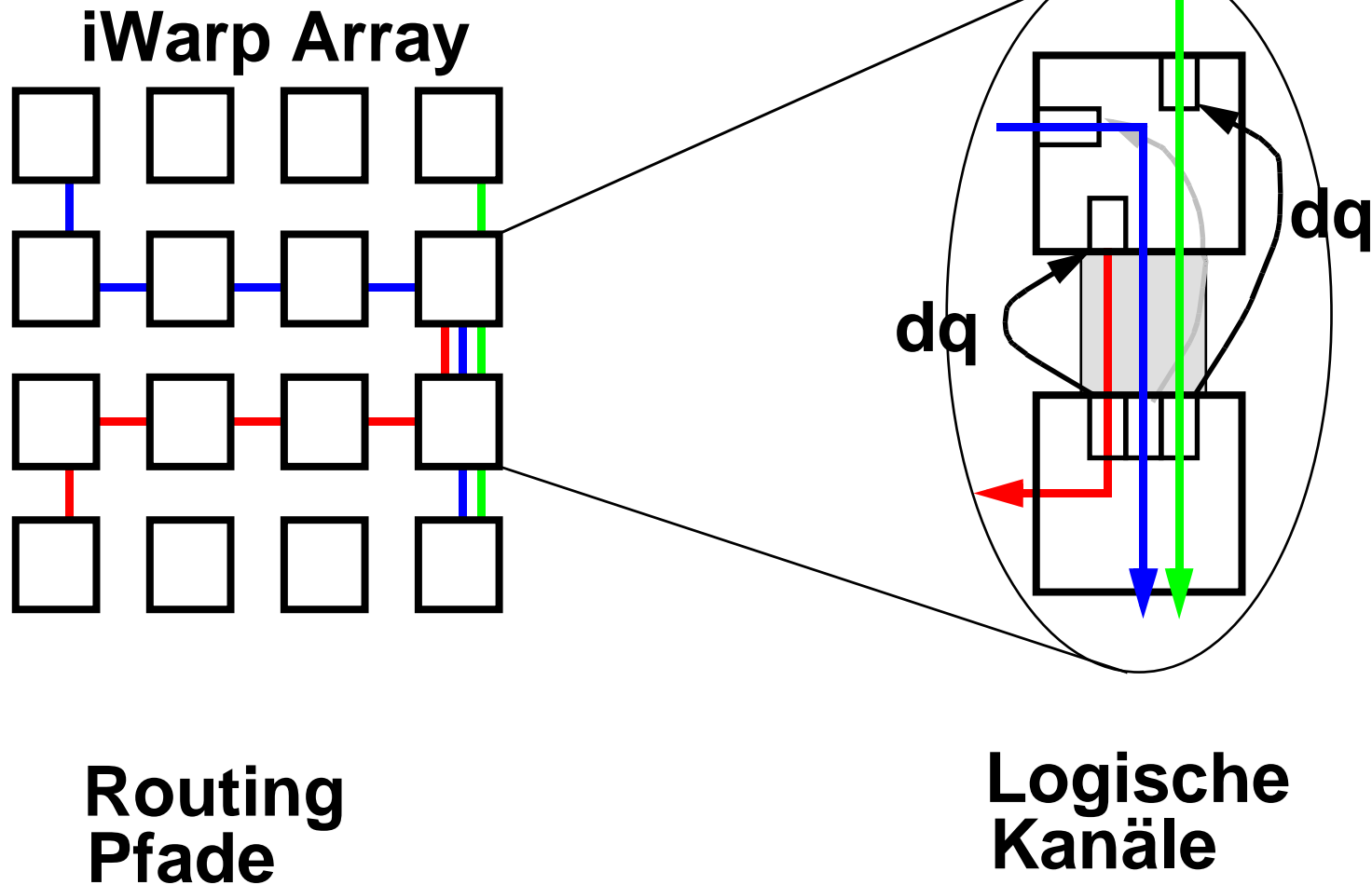
Hat der Markt an der Technologie vorbeigeschossen?

Beispiel: Flusskontrolle mit “Credits”

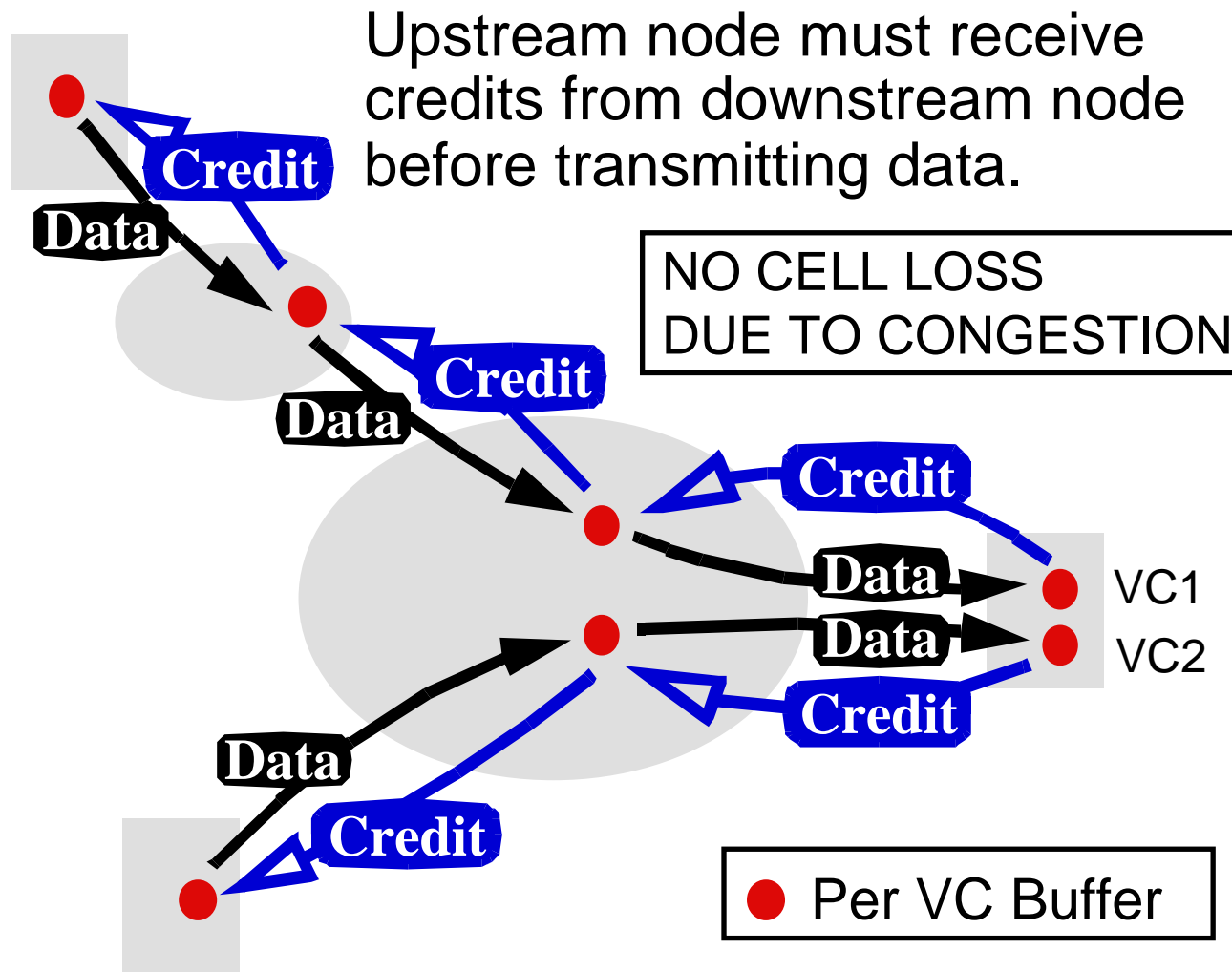
Vom Supercomputer in den Kommunikationsalltag

- **Supercomputer Kommunikation:**
 - Netzwerke sind zuverlässig (kleine Ausdehnung)
 - Flusskontrolle durch Rückstau
 - Verstopfung wird mit geschalteten Kanälen bekämpft (Virtual Channel [Seitz/Dally 1986]).
- **Allgemeine Kommunikationstechnik:**
 - Flusskontrolle durch Beseitigung von Paketen
 - Protokolle fangen allfällige Fehler auf, damit Belastung von Sender/Empfänger.

Virtual Channels im Supercomputer



Credits: Channels in ATM Netzwerken



vorerst noch Forschung: Standardisierung in Arbeit.

Suche nach geeignete Anwendungen

im parallelen Rechnen

- Grosse intellektuelle Anstrengungen versuchen die notwendige **Kommunikation zu vermindern** (nach dem Prinzip der Lokalität!).

Ausnahme: Systolische Algorithmen

in der Telekommunikation

- Alle Anstrengungen versuchen **Kommunikation zu reduzieren** (z.B. Videoströme komprimieren).

Dies steht eigentlich im Gegensatz zu den grossen technologischen Bandbreitereserven in Glasfasern und selbst in elektrischen Kabeln.

Neue Anwendungen I

Herausforderung: Bildsequenzen auf Verlangen

Gegeben: Gebäude (Betrieb, Schule) mit 500 Bildschirmarbeitsplätzen und einem Rechen bzw. Speicherzentrum im Keller. Die Daten bestehen aus kurzen Videosequenzen und hochauflösenden Bildern mit Informationen fürs “Business” (d.h. einigen Dutzend Terabytes)

Gesucht: Infrastruktur, die es erlaubt solche Videosequenzen per Mausklick einzuspielen, d.h. zu indizieren und durchzusuchen.

Wichtig: Kein Video Laden à la “Blockbuster Video”!

Neue Anwendungen II

Herausforderung: Software Updates für PCs

Gegeben: Gebäude (Betrieb, Schule) mit 1000 autonomen Arbeitsplatzrechnern mit lokalem Disk Speicherplatz von je einigen Gigabytes.

Gesucht: Infrastruktur, die es erlaubt die darauf enthaltene Software und Daten periodisch zu sichern, die Programminstallation nachzuführen und untereinander konsistent zu halten.

Viele Informatikverantwortliche können aus diesem Grund nicht auf den Zentralrechner verzichten.

Neue Anwendungen III

Herausforderung: Papierlose Hauptbibliothek

Gegeben: Traditionelle Bibliothek mit unverarbeiteten Originalquellen, die als hochauflösende Bilder gespeichert sind. Grosse CD-Jukebox oder DLT Silo (d.h. Disk oder Magnetband Roboter).

Gesucht: Wie gewohnt wollen wir blättern und mühelos auf dem Bildschirm Dokumente durchsuchen, genau wie in einem Lesesaal!

Schlussfolgerungen

- Wenn es gelingt **Kommunikationssysteme** mit **hoher Leistung** und **niedrigen Kosten** zu bauen, werden die *Alltagscomputer von morgen* kommunizieren können wie die *Supercomputer von heute*.
- **Anwendungen** der Hochleistungskommunikation sind teilweise **noch nicht vorhanden**, doch geeignete Probleme dazu sind **klar erkennbar**.
- Trotz schnellstem technologischem Fortschritt sind **Konzepte** der Computerarchitektur von **grosser Dauer**. Hochschulforschung in dem Gebiet ist **sinnvoll** und **notwendig**, auch wenn reine Entwicklungsarbeit *per se* am Standort Schweiz fraglich bleibt.