

### 3.15 Konzept eines Kommunikationsprozessors für massiv parallele Rechner (NP-COMPAS)

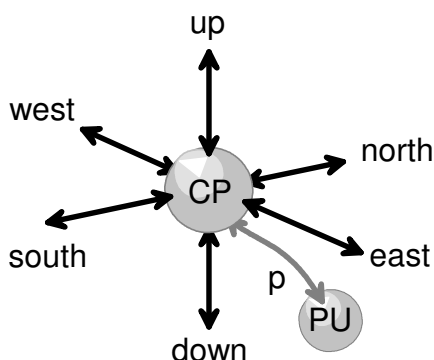
(Projektzeitraum: 01.07.1993 - 31.12.1993)

Dr.habil. Vesselin Jossivof (IIEF), Dr.Ing. Gerd Karl Heinz, Stephan Dengel

Im Projekt wird ein Routing-Schaltkreis entwickelt, der die Vermittlung und Durchschaltung von Daten in Rechnernetzen mit höchster Effizienz vornimmt.

Abgeleitet vom e-Cube Routing Algorithm (Basis Intel iPSC/2) wurden effiziente, parallele Routing-Algorithmen für deterministische und adaptive, parallele Paket-Kommunikation in Rechnernetzen des Typs 2D-Torus, 3D-Torus, Hypercube für Transputernetze T800-MTM2 (Parsytec) untersucht. Ausgehend von Hardwarebeschreibungen in VHDL wurden Schaltungsvarianten mit Mentor-Autologic synthetisiert und simuliert. Die Funktion konnte anhand verschiedener Simulationen nachgewiesen werden. Erfahrungen konnten mit der Implementierung der Funktion eines 3D-Hypercube auf programmierbare Logikschaltkreise vom Typ XILINX XC4005 gesammelt werden.

Die wissenschaftliche Aufgabenstellung des Vorhabens bestand in der Entwicklung einer Methode der Zusammenführung von parallelen Automatengraphen mit vergleichbarer Konnektivität, deren Kantengewichte und Zustandsfunktionen programmierbar geladen werden können.



CP-Knoten im Detail (3D-Netz).  
Bezeichnung der sechs Richtungen  
am CP. Kanal p kennzeichnet eine  
Verbindung zum Knotenprozessor (PU).

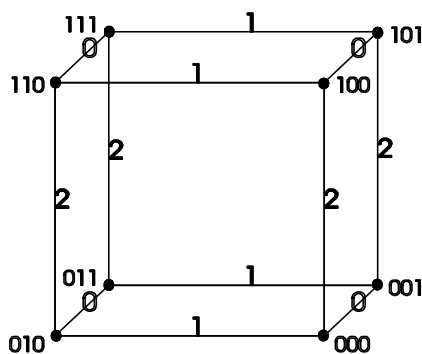
CP: Kommunikationsprozessor  
PU: Processing Unit

Grundlage des Vorhabens NP-COMPAS bilden drei Verhaltensmodelle, die durch die Beschreibung mit Automatennetzen eindeutig in Schaltungsstrukturen mit VHDL-Werkzeugen in einem Netzwerkprozessor für massiv parallele Computer umgesetzt werden können.

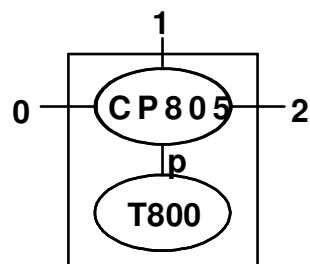
Die innovativen Aspekte des Kommunikationsprozessors bestehen in folgendem:

1. Die Funktion des Netzwerkprozessors ist eindeutig analytisch mit parallelen Automatennetzwerken beschrieben, die besonders zur Formalisierung von parallelen Schaltsystemen geeignet sind. Sie haben ihren Ursprung in der Turing-Maschine.

2. Der Netzwerkprozessor wird selbst als ein paralleles Schaltsystem mit asynchronem Verhalten realisiert. Der Vorteil dieser Konstruktion besteht in der konstanten Zeit, die für jede beliebige Kommunikationsentscheidung notwendig ist. Sie ist gleich der Summe der Zustandssequenzen von nur zwei Automaten, die die Steuerung eines jeden Datenkommunikationskanals bilden, unabhängig vom Grad des Kommunikationsnetzwerkes der MPC.
3. Für Hypercubusnetze ist die Anzahl der parallelen Automaten in Schaltnetzwerken linear skalierbar mit  $O(2n+2)$ ,  $n$  Grad des Hypercubus mit  $N=2^n$  Prozessorknoten. Für 2D und 3D-Netze ist



Nummerierungsschema  
am Beispiel 3D-Hypercube:  
- Knotennr. 000 bis 111 (bin)  
- Achsenrichtung 0...2 (dez)



Achsenrichtungen am  
Netzwerkknoten,  
Achsen 0,1,2,p

Beispiel 3D-Hypercube.

die Anzahl der Automaten in diesem Netzprozessor mit  $O(n+1)$  konstant.

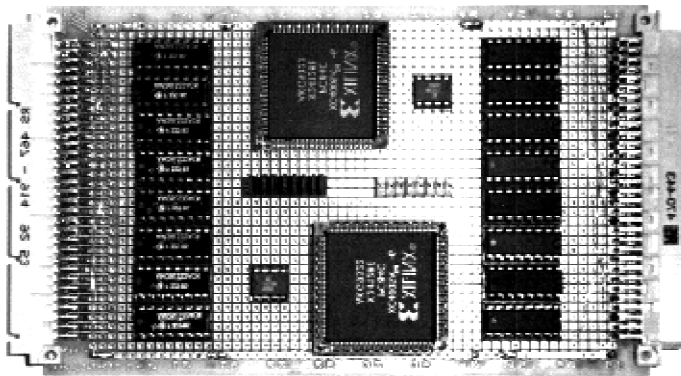
Für alle drei Netzwerktypen ist trotz der differenzierten Prioritätszuordnung der Knoten im benutzten wormhole-routing-Algorithmus der Netzprozessor ohne strukturelle Unterschiede universal instanzierbar. Das bedeutet, daß das PLD in jedem Knoten ohne zusätzliche Verschaltung einsetzbar ist.

Folgende Eigenschaften können zur Charakteristik der untersuchten Algorithmen angegeben werden:

1. Es sind bis zu 64k Knoten in Single-Stage Netzen (Backplane) unterstützbar.
2. Der Hardwareaufwand ist minimal (etwa 7 k Gatter für einen 3D-HC, ca. 50 k Gatter für einen 16D-HC). (Vergleich zum Schaltkreis C104 : 1,2 Mio. Trs.)
3. Es besteht eine flexible Anpassbarkeit an die Netzwerktypen 2D-Torus, 3D-Torus und nD-Hypercube.
4. In 2D/3D-Netzen erfolgt die Adreßrechnung ganzzahlig (integer), aber nach denselben Prinzipien wie beim Hypercube, beim Hypercube ist die Adreßrechnung binär.
5. Prozessornetze sind variierbar und erweiterbar ohne Prinzipwechsel.
6. Die verwendeten Algorithmen gestatten deterministisches ebenso wie adaptives Routing.
7. Das Prinzip der Knotennummerierung ist variierbar.
8. Die Algorithmen wurden für absolute Adressierung (mit Knotennummern) und für relative Adressierung (knotennummernfrei, implizite Nummerierung) untersucht.
9. Das entwickelte, innovative Konzept impliziter Nummerierung gestattet den parallelen Betrieb virtueller Netze unterschiedlichsten Typs auf einem Rechner. Entsprechende Adreßrechnungen wurden in den Varianten erprobt.
10. Die Adressierung eines Zieles kann sowohl mit dynamisch sich wandelndem, als auch mit statischem Adressvektor erfolgen.

11. Über Simulationen an den Beispielen 2D\_16, 3D\_9 und HC\_8 konnten wesentliche Funktionen prinzipiell bestätigt werden.
12. Die Prinzipbindung an eine statische Leitungsbelegung (T80x) kann bei HC zu Dead Locks führen. Als Ausweg bietet sich Paketbetrieb mit Längenbegrenzung der Messages auf den minimalen Diagonaldurchmesser des Netzes, gemessen in Input-Buffer-Verzögerungen an.
13. Arithmetische Schaltungen sind bei dem Verfahren der Knotennummerierung vermeidbar (auch bei 2D/3D !).

Erreichte Verzögerungszeiten reichen für ein maximales Adressfeld von 3Byte nicht über 50 Gate-Delays hinaus. Alle Baugruppen sind weitgehend unifizierbar und für beide Adreßtypen (integer, binary) geeignet, wesentliche Verschiedenheiten liegen in der fixierten Kanalzahl pro Netztyp (2D: 4; 3D: 6; n-dimensionaler Hypercube: n). Der Hardwareaufwand ist so minimal, daß im Gegensatz zu Referenzlösungen der komplette Kommunikationsprozessor für sechs Links (physisch unterlegter 3D-Torus) auf einem Gatearray statt auf einem PCB Platz finden kann. Es ist mit diesem neuartigen Konzept möglich, virtuell beliebig viele, verschiedene Netztypen (2D-Torus, 3D-Torus, Hypercube) auf ein- und demselben Netz parallel zueinander und unabhängig voneinander zu routen und zu betreiben, wengleich das real vorhandene Netz stets von ein und demselben Typ ist und bleibt (zB. 3D).



CP-Doppelmodul  
CP805-2-05. Versuchsplatine zur Erprobung zweier CP am MTM2-12 Board von Parsytec.

Links- und rechtsseitig sind die Schnittstellenkonverter RS422 zu erkennen. Die logische Funktion des Netzprozessors erfüllen die

Die Arbeiten werden 1994 voraussichtlich mit der Entwicklung eines Single-Stage-Knotenprozessors für Transputermodule mit sechs Links zur Unterstützung eines Hochleistungsprozessors fortgesetzt.

Dazu wird die Routingadresse dem einlaufenden, seriellen Datenstrom entnommen, die Kanalprüfung und -arbitrierung erfolgt online. Wenn der Datenstrom das Ende des Eingangsbuffers erreicht hat, ist der Vermittlungsvorgang bereits abgeschlossen, das Datenpaket geht ohne Stop auf eine in Richtung des Ziels liegende Leitung. Das gewählte Verfahren zeigt höchste Einfachheit und Transparenz bei Deadlockfreiheit sowie maximalem Datendurchsatz. Es toleriert auf physischer Ebene unterbrochene und belegte Leitungen, und gestattet Verbindungsstrukturen, deren real vorhandene, physische Struktur vollständig von der logischen Netzstruktur abweicht. Alle physischen Datenkanäle können entsprechend ihrer momentanen Belegtheit von beliebig vielen logischen Kanälen parallel benutzt werden (Virtual Channel Konzept).

## Veröffentlichungen/ Vorträge und Referenzen

- 1) Dengel, S.: Entwurf und Simulation eines Netzwerkprozessors für massiv-parallele Computer. Diplomarbeit. Fachhochschule für Technik und Wirtschaft, Fachbereich Technische Informatik, 18. Juni 1993, 84 S.
- 2) Dengel, S., Heinz, G., Jossifov, V.: Entwurf und Simulation eines Netzwerkprozessors für massiv parallele Computer. BMWi-Abschlußbericht 613/93, Projekt NP-COMPAS, GFaI-Report vom 30. Januar 1994
- 3) Getov, V., Jesshope, C.R.: Simulation facility of distributed memory system with "mad postmen" communication work, Proc. Sec. Europ. Distrib. Memory Comp. Conference, Munich, April 1991 Workshop NP-COMPAS "Konzept eines Kommunikationsprozessors für massiv parallele Rechner" Thesys, Erfurt, 2.12.93
- 4) Heinz, G.: "Kommunikationsprozessor für PowerPC 601 - Aufgaben und Leistungsmerkmale", Gesellschaft zur Förderung angewandter Informatik e.V.. Vortrag zum GFaI-Workshop im Projekt NP-COMPAS, Thesys, Erfurt, 2.12.93
- 5) Heinz, G: Lehrgang VHDL 1076 mit Mentor Autologic VHDL V.8.2, FHTW Berlin-Karlshorst, 12./13.10.1993
- 6) Intel Corp.: iPSC/2 System Manual. 1988
- 7) Jesshope, C. R., Miller, P., Jantchev, J.: Programming with active data, Workshop PARCELLA'88, Berlin, 1988
- 8) Jesshope, C.R.: Communications architectures for finegrain parallel processing, Proc. PARCELLA'90, Berlin, 1990
- 9) Jesshope, C.R.: Dynamic load balanced active data model of parallel processing for vision, Proc. BCS, Workshop on parallel processing for vision, Oxford University, 1986
- 10) Jesshope, C.R.: Parallel processing based on active data, Proc. BIRA seminar on Parallel processing, Antwerpen, Nov., 1987
- 11) Jesshope, C.R.: High Performance Communication Architectures, Proc. Workshop WP&DP'90, Sofia, March 1990
- 12) Jossifov, V., Dengel, S.: "Inhalt und Ergebnisse des Projekts NP-COMPAS: Entwurf eines Netzwerkprozessors für massiv parallele Computer als ASIC-Prototyp". Vortrag zum GFaI-Workshop im Projekt NP-COMPAS, Thesys, Erfurt, 2.12.93
- 13) Jossifov, V., Krapp, M.: Adaptive packet routing algorithm for 3D-meshes. AdW-ZKI Berlin, technical report 6/91
- 14) Jossifov, V., Krapp, M.: Adaptive routing algorithm for 3D-mesh computer networks - formal description with automaton nets. 36. Int. Wiss.Kolloquium der TU Ilmenau, 24.-28.10.1991
- 15) Jossifov, V.: Parallel Computer architectures - theory and applications, Habilitation, Institut für Mathematik und Informatik der Bulg. Akad. der Wissenschaften, Sofia, Dez. 1990
- 16) Jossifov, V.: The challenge of the key technologies in High-Performance Computing - Some aspects, Deutsch-Französisches Kolloquium an der FHTW Berlin "Moderne Rechnerarchitekturen", Berlin, 07-09.12.1992
- 17) Krapp, M., Jossifov, V.: Formal specification of a distributed packet-router for nD-hypercubes. Workshop on parallel and distributed processing, Bulgarian Academy of Sciences, March, 1990, Sofia, Elsevier Sc. Publ., 1991, accepted for presentation also on Joint Conference on Vector and Parallel Processing - CONPAR 90, Zurich, Sep. 1990 and at Workshop on Algorithms and Parallel VLSI Architectures, Pont-a-Mousson, June 1990
- 18) Krapp, M., Jossifov, V.: Formal specification of a distributed packet-routing algorithm with automation-nets, Technical report No.5, Bulgarian Academy of Sciences, Center for Informatics and Computer technology, Sofia, Oct. 1989
- 19) Krapp, M, Jossifov, V.: Formal specification of a dynamic load balancing packet routing algorithm for binary nD-hypercubes, ZKI-Berlin Technical report 6/91
- 20) Tuazon, J., Peterson, K., Puiel, K., Liebermann, D.: Caltech/IPC Marc II Hypercube concurrent processor, Proc of ICPP, 1985, pp. 666-673





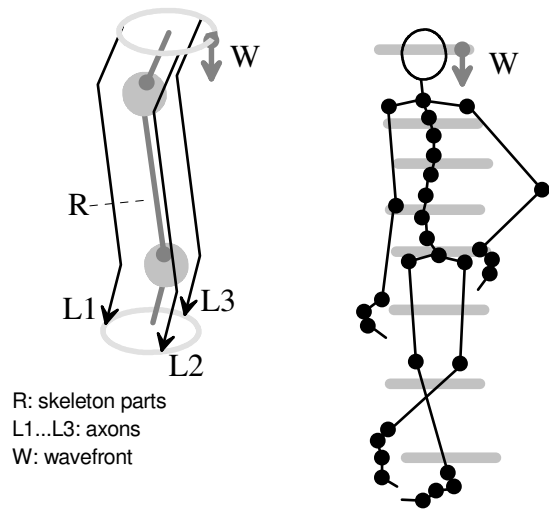


Fig. 4: Interference wavefront model of the human equilibration system.