

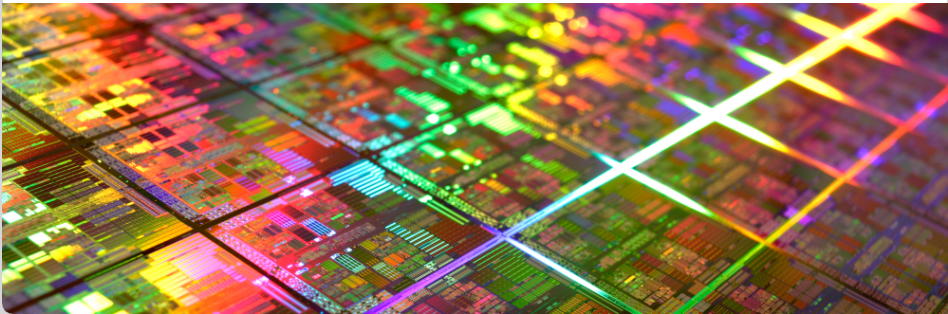
Zentralübung Rechnerstrukturen im SS 2013

Verbindungsstrukturen

Martin Schindewolf, Wolfgang Karl

Lehrstuhl für Rechnerarchitektur und Parallelverarbeitung

27. Juni 2013



Verbindungsstrukturen:

- Charakterisierung von Verbindungsnetzen
- Statische Verbindungsnetze (Aufgabe 1)
- Dynamische Verbindungsnetze (Aufgabe 2)

Vergleich von Parallelrechnern:

- Motivation
- Aufgabe 3

Sonstiges:

- Klausuraufgabe 2011/12-WS
- Klausuraufgabe 2011-SS

Verbindungsgrad eines Knoten P

Anzahl der Kanten von einem Knoten zu anderen Knoten

Durchmesser

- Maximale Distanz zwischen zwei Knoten
- Maximale Pfadlänge
- Keine Aussage über die realen Leitungslängen $\Leftrightarrow$ Latenz

Blockierung

→ blockierungsfrei?

Erweiterbarkeit

- begrenzt
- stufenlos
- z.B. durch Verdoppeln der Knoten

Skalierbarkeit des Verbindungsnetzes

- Fähigkeit, die wesentlichen Eigenschaften des Verbindungsnetzes auch bei beliebiger Erhöhung der Knotenzahl beizubehalten.
- ⇒ Vergrößerung möglich ohne die wesentlichen Eigenschaften des Netzwerks zu verlieren
- **Achtung:** Nicht verwechseln mit der Skalierbarkeit eines Parallelrechners! (vgl. Übung #5, Folie 15)

Erweiterbarkeit

- begrenzt
- stufenlos
- z.B. durch Verdoppeln der Knoten

Skalierbarkeit des Verbindungsnetzes

- Fähigkeit, die wesentlichen Eigenschaften des Verbindungsnetzes auch bei beliebiger Erhöhung der Knotenzahl beizubehalten.
- ⇒ Vergrößerung möglich ohne die wesentlichen Eigenschaften des Netzwerks zu verlieren
- **Achtung:** Nicht verwechseln mit der Skalierbarkeit eines Parallelrechners! (vgl. Übung #5, Folie 15)

Minimale Bisektionsbreite:

Schneidet man einen Graphen in zwei gleich große in sich zusammenhängende Teile und betrachtet die Menge der Kanten, die diesen Schnitt kreuzen, so bezeichnet man die Kardinalität der kleinsten Kantenmenge – über alle möglichen Schnitte – als minimale Bisektionsbreite.

Bisektionsbandbreite

Maximale Datenmenge, die das Netzwerk über die Bisektionslinie, die das Netzwerk in zwei Hälften teilt, pro Sekunde transportieren kann.

Bisektionsbandbreite

Maximale Datenmenge, die das Netzwerk über die Bisektionslinie, die das Netzwerk in zwei Hälften teilt, pro Sekunde transportieren kann.

Übertragungsbandbreite / Durchsatz:

- Die maximale Übertragungsleistung des Verbindungsnetzes oder einzelner Verbindungen
- Meist theoretisch errechnet.

Latenz (latency)

Übertragungszeit einer Nachricht

- Kanalverzögerung
- Schalt-/Routing-Verzögerung (switching/routing delay)
- Blockierungszeit (contention time)

Ausfalltoleranz (Redundanz)

- Verbindungen zwischen Knoten sind selbst dann noch zu schalten, wenn einzelne Elemente des Netzes ausfallen
- minimale Bisektionsbreite
- Diskonnektivität

Latenz (latency)

Übertragungszeit einer Nachricht

- Kanalverzögerung
- Schalt-/Routing-Verzögerung (switching/routing delay)
- Blockierungszeit (contention time)

Ausfalltoleranz (Redundanz)

- Verbindungen zwischen Knoten sind selbst dann noch zu schalten, wenn einzelne Elemente des Netzes ausfallen
- minimale Bisektionsbreite
- Diskonnektivität

Ausfalltoleranz (Redundanz):

Kriterien:

- minimale Bisektionsbreite
- Wurzelknoten
- Flaschenhals
- Ausweichverbindungen
- ...

Interpretation der Daten:

- Vergleich mit anderen Netzen
- Beachtung von Teilnetzen (z.B. Ring-Würfel-Netzwerk)
- ...

Ausfalltoleranz (Redundanz):

Kriterien:

- minimale Bisektionsbreite
- Wurzelknoten
- Flaschenhals
- Ausweichverbindungen
- ...

Interpretation der Daten:

- Vergleich mit anderen Netzen
- Beachtung von Teilnetzen (z.B. Ring-Würfel-Netzwerk)
- ...

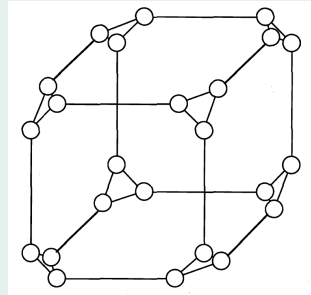
Ausfalltoleranz (Redundanz):

Kriterien:

- minimale Bisektionsbreite
- Wurzelknoten
- Flaschenhals
- Ausweichverbindungen
- ...

Interpretation der Daten:

- Vergleich mit anderen Netzen
- Beachtung von Teilnetzen (z.B. Ring-Würfel-Netzwerk)
- ...



Datentransfer: Durchschalte- oder Leitungsvermittlung

Circuit switching

- direkte Verbindung zwischen zwei oder mehreren Knoten zu schalten
 - Blockierungsfreie Kommunikation
 - Kurze Latenzen
 - Startup-Time lange
- ⇒ Besonders geeignet für lange Nachrichten

Datentransfer: Paketvermittlung

Packet switching

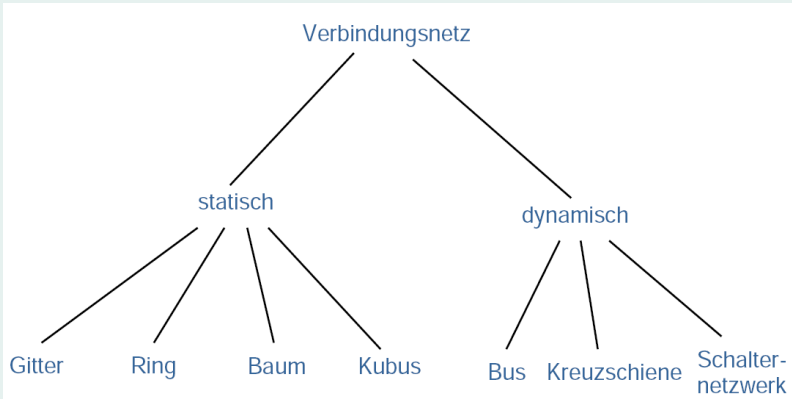
- Datenpakete fester Länge und Nachrichten variabler Länge werden verschickt
 - Wegefindungsalgorithmus (routing) notwendig
 - Absender und Empfänger
 - In jedem Knoten auf dem Weg erfolgt eine Verarbeitung zur Weiterleitung
- ⇒ Günstig für kurze Nachrichten

Datentransfer: Paketvermittlung (Packet switching)

Verschiedene Übertragungsmodi:

- Store and forward
- Cut through
- Wormhole
- Virtual cut through
- Wormhole routing
- Buffered wormhole routing
- ...

Klassifizierung von Verbindungsnetzen:



Statische Verbindungsstrukturen

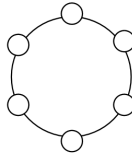
- In statischen Netzen existieren fest installierte Verbindungen zwischen Paaren von Netzknoten
- Steuerung des Verbindungsaufbaus ist Teil der Knoten

Dynamische Verbindungsstrukturen

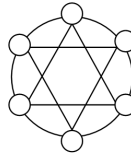
- Dynamische Netze enthalten eine Komponente „Schaltnetz“, an die alle Knoten über Ein- und Ausgänge angeschlossen sind.
- Direkte, fest installierte Verbindungen zwischen den Knoten existieren nicht.
- Alle notwendigen Steuerungsfunktionen sind im Schaltnetz konzentriert

Statische Verbindungsstrukturen

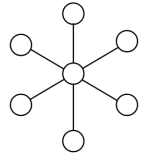
- Kette
- Ring
- Stern
- Baum (event. mit Fat-Tree)
- Chordaler Ring
- Gitter



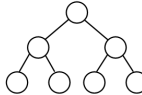
Ring



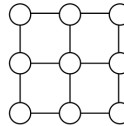
Chordaler Ring



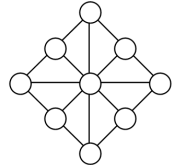
Stern



Baum



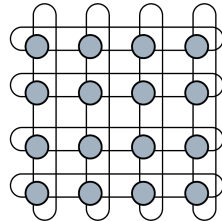
Gitter mit vier
Nachbarknoten



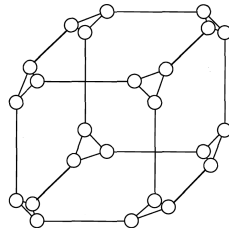
Gitter mit acht
Nachbarknoten

- Torus
- Pyramide
- Würfel
- n-dimensionaler Hyperwürfel
- Ring-Würfel-Netzwerk
Cube-Connected-Cycle (CCC)

2-dim. Torus



CCC

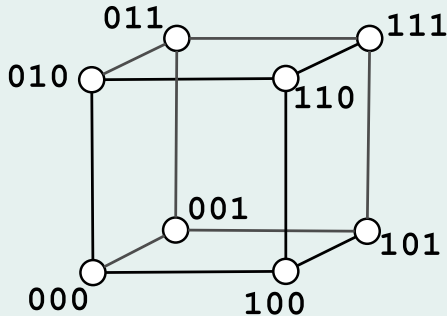


K-ärer n-Kubus (Cubes, Würfel)

- Allgemeine Form eines Kubus-Verbindungsnetzwerkes
- Ringe, Gitter oder Hyperkubi sind eine Teilmenge der Klasse der K-ären n-Kubus-Netzwerke
 - n ist die Dimension
 - Der Radius K ist die Anzahl der Knoten, die einen Zyklus in einer Dimension bilden (Rückwärtskanten)
- Enthält $N = K^n$ Knoten
- Die Knoten werden über eine n -stellige K-äre Zahl der Form $a_0, a_1, \dots, a_{n-1}$ adressiert
 - Jede Stelle a_i mit $0 \leq a_i < K$ stellt die Position des Knotens in der entsprechenden i -ten Dimension dar mit $0 \leq i \leq n - 1$
 - Von einem Knoten mit Adresse $a_0, a_1, \dots, a_{n-1}$ kann ein Nachbarknoten in der i -ten Dimension mit $a_0, a_1, \dots, (a_i \pm 1) \bmod k, \dots, a_{n-1}$ erreicht werden
- Knotengrad ist $2n$ und der Diameter ist $n \lfloor \frac{k}{2} \rfloor$

K-ärer n-Kubus – Beispiele

- **K=2, n=3** (hier vereinfacht ohne Rückwärtskanten)
- Adresse: 3-stellige 2-äre (binäre) Zahl $a_0 a_1 a_2$
- a_i mit $0 \leq a_i < 2$



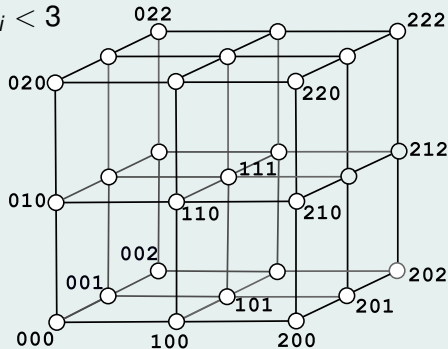
K-ärer n-Kubus – Beispiele

■ **K=3, n=3**

⇒ 3D-Torus (hier vereinfacht ohne Rückwärtskanten)

■ Adresse: 3-stellige 3-äre Zahl $a_0 a_1 a_2$

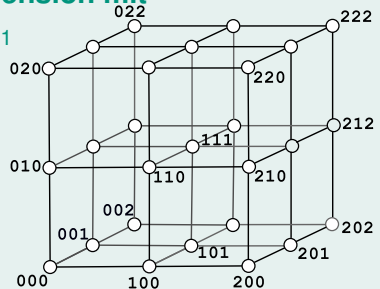
■ a_i mit $0 \leq a_i < 3$



K-ärer n-Kubus – Beispiele

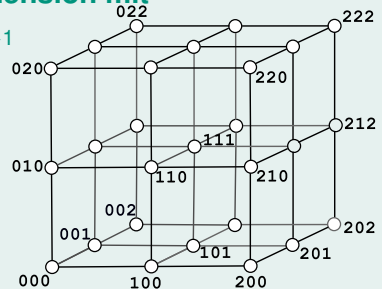
- **K=3, n=3**, 3D-Torus (im Bild vereinfacht o. Rückwärtskanten)
- Adresse: 3-stellige 3-äre Zahl $a_0 a_1 a_2$
- a_i mit $0 \leq a_i < 3$
- Von einem Knoten mit Adresse $a_0, a_1, \dots, a_{n-1}$ kann ein Nachbarknoten in der i -ten Dimension mit $a_0, a_1, \dots, (a_i \pm 1) \bmod k, \dots, a_{n-1}$ erreicht werden

- Von $110 \Rightarrow 100$
 $a_1 = 1 \Rightarrow (a_1 - 1) \bmod 3 = 0$
- Von $210 \Rightarrow 010$
 $a_0 = 2 \Rightarrow (a_0 + 1) \bmod 3 = 0$



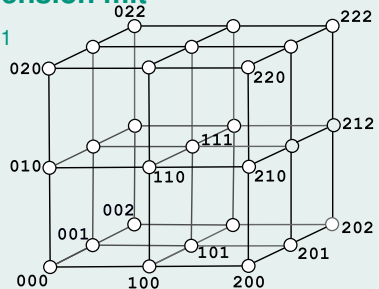
K-ärer n-Kubus – Beispiele

- **K=3, n=3**, 3D-Torus (im Bild vereinfacht o. Rückwärtskanten)
- Adresse: 3-stellige 3-äre Zahl $a_0 a_1 a_2$
- a_i mit $0 \leq a_i < 3$
- Von einem Knoten mit Adresse $a_0, a_1, \dots, a_{n-1}$ kann ein Nachbarknoten in der i -ten Dimension mit $a_0, a_1, \dots, (a_i \pm 1) \bmod k, \dots, a_{n-1}$ erreicht werden
- Von 110 $\Rightarrow$ 100
 $a_1 = 1 \Rightarrow (a_1 - 1) \bmod 3 = 0$
- Von 210 $\Rightarrow$ 010
 $a_0 = 2 \Rightarrow (a_0 + 1) \bmod 3 = 0$



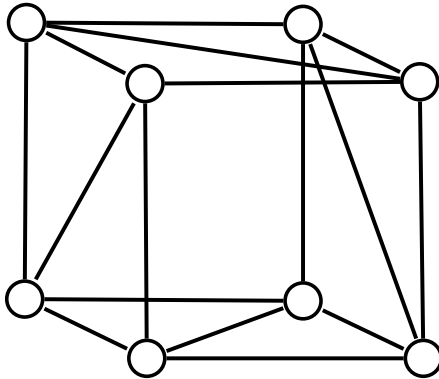
K-ärer n-Kubus – Beispiele

- **K=3, n=3**, 3D-Torus (im Bild vereinfacht o. Rückwärtskanten)
- Adresse: 3-stellige 3-äre Zahl $a_0 a_1 a_2$
- a_i mit $0 \leq a_i < 3$
- Von einem Knoten mit Adresse $a_0, a_1, \dots, a_{n-1}$ kann ein Nachbarknoten in der i -ten Dimension mit $a_0, a_1, \dots, (a_i \pm 1) \bmod k, \dots, a_{n-1}$ erreicht werden
- Von $110 \Rightarrow 100$
 $a_1 = 1 \Rightarrow (a_1 - 1) \bmod 3 = 0$
- Von $210 \Rightarrow 010$
 $a_0 = 2 \Rightarrow (a_0 + 1) \bmod 3 = 0$

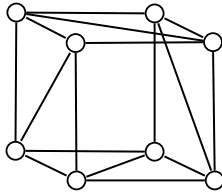


Aufgabe 1 – Statische Verbindungsstrukturen

Gegeben sei ein Verbindungsnetzwerk mit der nachfolgend dargestellten Topologie:



- a) Bestimmen Sie den Verbindungsgrad, den Diameter und die minimale Bisektionsbreite.



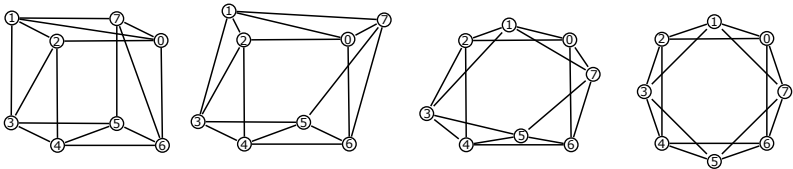
Verbindungsgrad: 4
Durchmesser: 2
min. Bisektionsbreite: 6

- b) Um welche Art eines Verbindungsnetzwerkes handelt es sich in diesem Fall?

Chordaler Ring mit Knotengrad 4

b) Um welche Art eines Verbindungsnetzwerkes handelt es sich in diesem Fall?

Chordaler Ring mit Knotengrad 4



c) **Liegt Redundanz vor? Wenn ja, wieviele Verbindungsleitungen können ausfallen bevor eine Verbindung zwischen zwei beliebigen Knoten nicht mehr geschaltet werden kann?**

- Es liegt Redundanz vor.
- Da der Verbindungsgrad jedes Knotens 4 ist und bidirektionale Leitungen verwendet werden, können bis zu drei Leitungen ausfallen und dennoch jeder Knoten von einem anderen erreicht werden.
Anmerkung: Hier ist die minimale Anzahl von Kanten gesucht, die ausfallen dürfen, bevor ein Knoten nicht mehr erreichbar ist.
- Allerdings kann beim Ausfall einer Kante der Durchmesser steigen, das heißt es könnten längere Wege notwendig sein.

- d) **Vergleichen Sie diese Netzwerktopologie mit den Topologien (unidirektionaler) Ring, 2D-Gitter, (binärer) Baum und Hyperkubus in den Punkten Verbindungsgrad, Durchmesser und minimaler Bisektionsbreite.**

$N = \#$ Knoten

	Aufgabe a)	Ring	2D-Gitter	Baum	Hyperkubus
Knotenzahl	N	N	$N = n^2$	N	$N = 2^n$
Verbindungsgrad	4	2	$2 - 4$	1 – 3	$\log_2 N$
Durchmesser	$\lfloor \sqrt{N} \rfloor$	$N - 1$	$2(\sqrt{N} - 1)$	$2(\lceil \log_2 N \rceil - 1)$	$\log_2 N$
min. Bisektionsbreite	6	2	$\sqrt{N}$	1	$N/2$

- e) **Lange Zeit war ein Hyperkubus die häufigste Verbindungsstruktur bei nachrichtengekoppelten Multiprozessorsystemen. Wie viele Knoten müssen bei einem Hyperkubus für eine Erweiterung hinzugefügt werden? Was stellen Sie dabei für den Verbindungsgrad fest und was hat das für Auswirkungen auf den Aufbau und die Erweiterbarkeit des Rechners?**
- Jede Erweiterung benötigt die Verdopplung der Prozessorenanzahl
 - Der Verbindungsgrad der Knoten steigt bei jeder Erweiterung um 1
 - Rechner sind deshalb aus räumlichen Anordnungsgründen begrenzt

Statische Verbindungsstrukturen

- In statischen Netzen existieren fest installierte Verbindungen zwischen Paaren von Netzknoten
- Steuerung des Verbindungsaufbaus ist Teil der Knoten

Dynamische Verbindungsstrukturen

- Dynamische Netze enthalten eine Komponente „Schaltnetz“, an die alle Knoten über Ein- und Ausgänge angeschlossen sind.
- Direkte, fest installierte Verbindungen zwischen den Knoten existieren nicht.
- Alle notwendigen Steuerungsfunktionen sind im Schaltnetz konzentriert

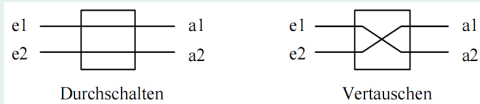
- **Bus**, Mehrfachbus
- **Kreuzschienenverteiler** (Crossbar Switch)
Alle angeschlossenen Prozessoren und Speicher können paarweise disjunkt gleichzeitig und blockierungsfrei miteinander kommunizieren.
- **Schalternetzwerke** aus Zweierschaltern
- **Permutationsnetze**

- **Bus**, Mehrfachbus

- **Kreuzschienenverteiler** (Crossbar Switch)

Alle angeschlossenen Prozessoren und Speicher können paarweise disjunkt gleichzeitig und blockierungsfrei miteinander kommunizieren.

- **Schalternetzwerke** aus Zweierschaltern



- **Permutationsnetze**

■ Permutationsnetze

- p Eingänge des Netzes können gleichzeitig auf p Ausgänge geschaltet werden

⇒ Es wird eine Permutation der Eingänge erzeugt

- **Einstufige und mehrstufige Permutationsnetze** enthalten eine bzw. mehrere Spalten von Zweierschaltern

■ Reguläre Permutationsnetzwerke

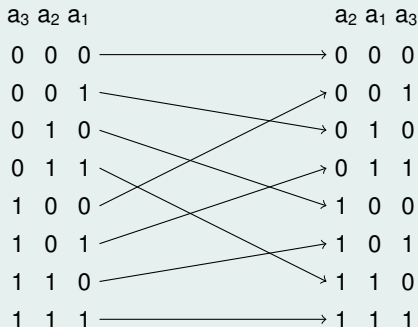
- p Eingänge
- p Ausgänge
- k Stufen mit je $p/2$ Zwischenschaltern
- p normalerweise eine Zweierpotenz

- **Irreguläre Permutationsnetzwerke** weisen gegenüber der regulären Struktur Lücken auf

Mischpermutation M (Perfect Shuffle)

Kreisverschiebung der Adressbits

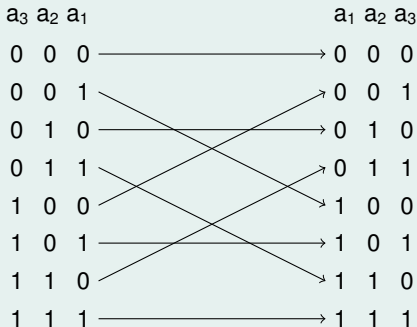
$$M(a_n, a_{n-1}, \dots, a_2, a_1) = (a_{n-1}, \dots, a_2, a_1, a_n)$$



Kreuzpermutation K (Butterfly)

Vertauschen des hochwertigsten mit dem niedrigwertigsten Adressbit

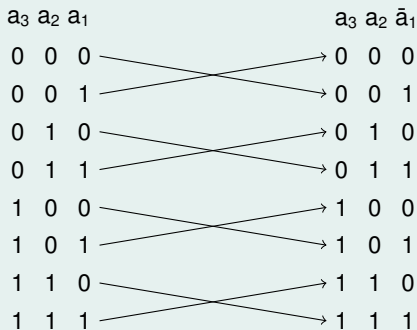
$$K(a_n, a_{n-1}, \dots, a_2, a_1) = (a_1, a_{n-1}, \dots, a_3, a_2, a_n)$$



Tauschpermutation T

Negation des niedrigwertigsten Adressbits

$$T(a_n, a_{n-1}, \dots, a_2, a_1) = (a_n, a_{n-1}, \dots, a_2, \bar{a}_1)$$



Umkehrpermutation U

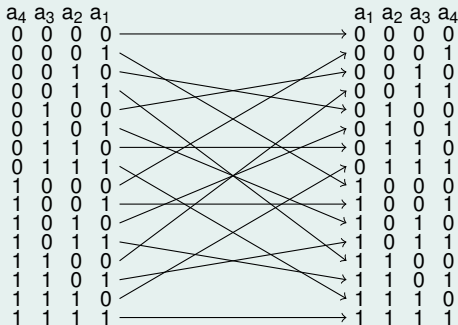
Spiegelung aller Adreßbits um die Mitte der Adressbitfolge:

$$U(a_n, a_{n-1}, \dots, a_2, a_1) = (a_1, a_2, \dots, a_{n-1}, a_n)$$

Für $n = 2$ und $n = 3$ ergibt sich dasselbe Grundmuster wie bei der Kreuzpermutation.

Umkehrpermutation

Für $n \geq 4$ unterscheiden sich jedoch Umkehr- und Kreuzpermutation

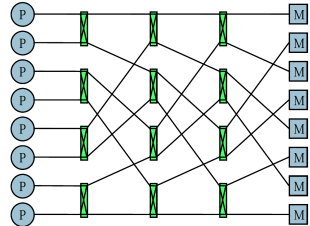


Mehrstufige Permutationsnetzwerke:

- jeweils aus einem bestimmten Grundmuster aufgebaut
- oft mit einer der eben vorgestellten Permutationen

Beispiele:

- Omega-Netzwerk
 - Mischpermutation
- Switching-Banyan-Netzwerk
 - Kreuzpermutation
- Benes-Netzwerk
 - rekursiver Aufbau

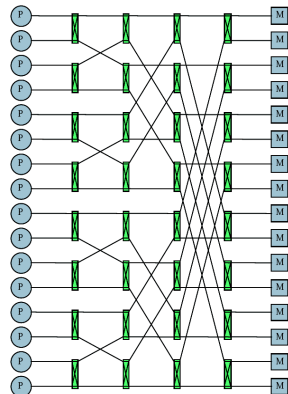


Mehrstufige Permutationsnetzwerke:

- jeweils aus einem bestimmten Grundmuster aufgebaut
- oft mit einer der eben vorgestellten Permutationen

Beispiele:

- Omega-Netzwerk
 - Mischpermutation
- Switching-Banyan-Netzwerk
 - Kreuzpermutation
- Benes-Netzwerk
 - rekursiver Aufbau



Mehrstufige Permutationsnetzwerke:

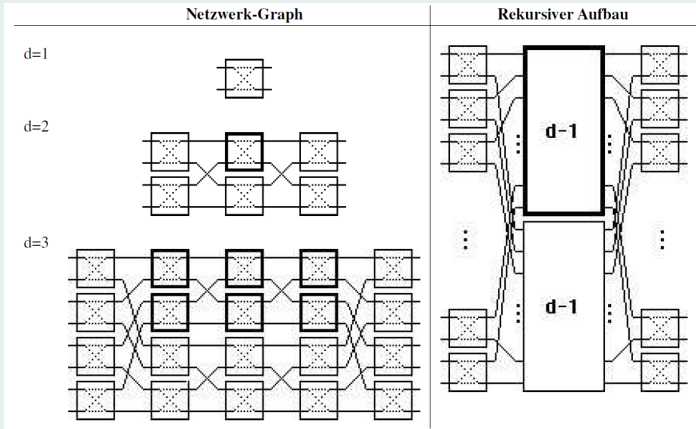
- jeweils aus einem bestimmten Grundmuster aufgebaut
- oft mit einer der eben vorgestellten Permutationen

Beispiele:

- Omega-Netzwerk
 - Mischpermutation
- Switching-Banyan-Netzwerk
 - Kreuzpermutation
- Benes-Netzwerk
 - rekursiver Aufbau

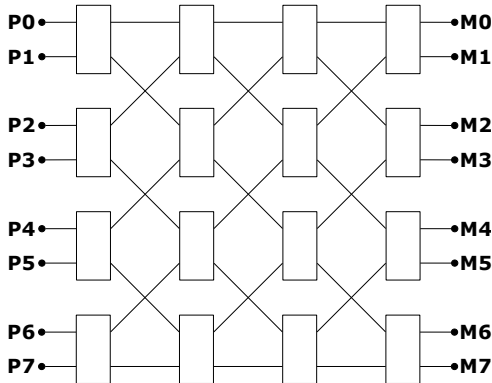
Benes-Netzwerk

■ rekursiver Aufbau



Aufgabe 2 – Dyn. Verbindungsstrukturen

Gegeben sei ein dynamisches Verbindungsnetzwerk, das 8 Prozessoren (P0 – P7) mit 8 Speichern (M0 – M7) wie folgt über einen Verbund von Zweierschaltern verbindet:



- a) **Kann zwischen jedem Prozessor- und Speicherpaar eine Verbindung hergestellt werden?**

Ja!

Hier war nicht nach gleichzeitig möglichen Verbindungen gefragt (vgl. Teilaufgabe b über alle Permutationen)

b) Kann jede Permutation generiert werden? Begründen Sie Ihre Antwort!

Nein! Beweis durch Widerspruch.

Annahme: jede Permutation kann generiert werden.

Gesucht: mindestens eine Permutation, für die die Annahme nicht gilt.

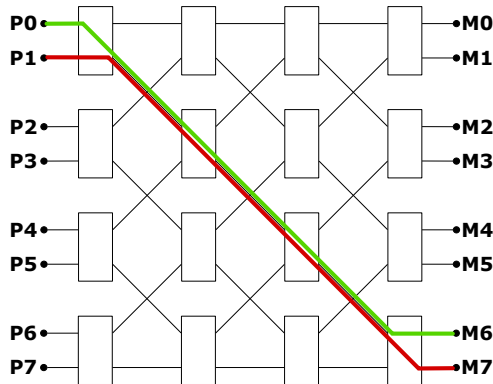
- Bei einer paarweisen Mischpermutation (Kreisverschiebung), hier also Verbindung von P0 und P1 mit M6 bzw. M7 gibt es nur einen möglichen Verbindungsweg, der gleichzeitig für beide Verbindungen benutzt werden müßte

⇒ Blockierung

Aufgabe 2 – Dyn. Verbindungsstrukturen

- c) Was ist die minimale Verbindungszahl ab der eine Blockierung auftritt? Geben Sie ein Beispiel an.

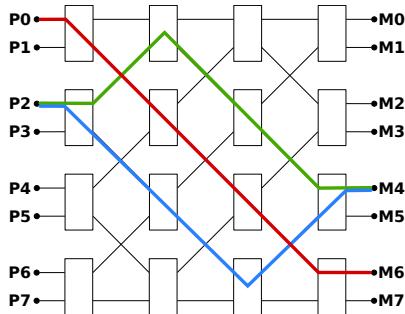
Schon bei zwei Verbindungen kann eine Blockierung auftreten:
z.B. bei $P0 \rightarrow M6$ und $P1 \rightarrow M7$



d) Ist das Netzwerk redundant? Begründen Sie Ihre Antwort.

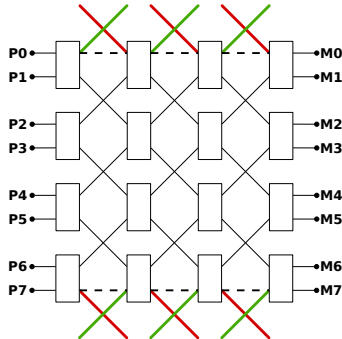
Nein!

Auf der einen Seite gibt es für bestimmte Paare mehrere Möglichkeiten (vgl. $P2 \rightarrow M4$), aber ebenso gibt es Paare, bei denen schon der Ausfall einer Verbindung die Weiterleitung ausschließt (z.B. $P0 \rightarrow M6$).



■ Wie könnten manche der Problem mit diesem Netzwerk vermeiden/verringert werden?

⇒ Verbindungen vom oberen Rand nach unten (vgl. n-dim Torus)



■ Achtung, durch das Weglassen der gestrichelten Verbindungen wäre z.B. auch keine Verbindung von P0 zu M0 mehr möglich!

Motivation

- Top500-Liste
 - halbjährlich neu
 - Kommentare bei Golem/Heise (Mehr Frames beim Zocken,...)
- Welchen Rechner kaufen?
- Was für eine Leistung ist notwendig?
 - ⇒ Wetterbericht vgl. Ocean
 - Abstand zwischen den berechneten Punkten
 - Bei Ocean: Gitter mit 100 x 100 Punkte, Abstand 20 km
- Unterstützte Programmiermodelle
- Skalierbarkeit
- Abhängigkeiten vom Verbindungsnetz

FLOPS

$$\text{MFLOPS} = \frac{\text{Anzahl der ausgeführten Gleitkommainstruktionen}}{10^6 \times \text{Ausführungszeit}}$$

- Maßzahl für die Operationsleistung (Gleitkomma-Verarbeitung)
- MFLOPS, GFLOPS, TFLOPS, PFLOPS,...
- Abkürzung FLOPS vs. Flops vs. Flop/s
 - FLOPS: **F**loating-point **O**perations **P**er **S**econd
 - FLOP/s: Floating Point Operations per Second
 - FLOP: **F**loating-point **O**peration
 - FLOPs: **F**loating-point **O**perations
 - 1 MFLOPS $\Leftrightarrow$ 1 MFLOP/s $\Leftrightarrow$ 1 MFLOP

Aussagekraft der Top 500-Liste

- Leistungsfähigkeit von vielen Faktoren abhängig
 - Linpack-Benchmark (optimiert für Listenplatz)
 - Rechenleistung $\Leftrightarrow$ Energieverbrauch
 - Auslastung im regulären Betrieb
 - Programmierbarkeit
 - Real zu berechnende Probleme

Frage

- Welches ist der bessere Rechner?

⇒ **Oft keine allgemeingültige Antwort möglich!**

Zu vergleichende Parallelrechner:

- BlueGene/L am LLNL
 - siehe Foliensatz 4, Folien 2-4
 - 596 TFLOPS, 212.992 Prozessoren

- HP XC6000 am KIT
 - 1,9 TFLOPS, 282 Prozessoren
 - Netzwerk siehe Foliensatz 5, Folien 2-34ff

- K computer am RIKEN
 - siehe Foliensatz 4, Folien 2-12
 - 8.162 TFLOPS, 548.352 CPU-Kerne

a) Wieviel GFLOPS trägt jeder einzelne CPU-Kern zur theoretischen Spitzenleistung bei?

■ BlueGene/L:

$$596.000 \text{ GFLOPS} / 212.992 \text{ Proz} = 2,80 \text{ GFLOPS/Proz}$$

■ HP XC6000:

$$1.900 \text{ GFLOPS} / (101 * 2 + 10 * 8) \text{ Proz} = 6,74 \text{ GFLOPS/Proz}$$

■ K computer:

$$8.162.000 \text{ GFLOPS} / 548.352 \text{ Kerne} = 14,9 \text{ GFLOPS/Kern}$$

aktuelle Konfiguration:

$$10.510.000 \text{ GFLOPS} / 705.024 \text{ Kerne} = 14,9 \text{ GFLOPS/Kern}$$

$$10.510.000 \text{ GFLOPS} / 88.128 \text{ Proz} = 119,2 \text{ GFLOPS/Proz}$$

■ Achtung, dies sind sehr theoretische und vereinfachte Werte!

b) Was für ein Netzwerktyp/-struktur wird verwendet? (Topologie, Hersteller, statisches oder dynamisches Netz,...)

- BlueGene/L:
3-dimensionaler Torus, Eigenentwicklung von IBM, statisches Netz
- HP XC6000:
Fat-Tree (Baumstruktur), Quadrics QsNet II Interconnect, dynamisches Netz, Rechnerknoten sind nicht im Netzwerk auf verschiedenen Ebenen verteilt

c) Wie groß ist der Durchmesser, d.h. die längste Verbindung zwischen zwei Knoten?

■ BlueGene/L:

Kantenlänge eines 3-dim. Würfels: $\sqrt[3]{212992} \approx 60$

⇒ Durchmesser von 3D-Torus: $3 * 60 / 2 \approx 90$

Bei einem Torus wird im Vergleich zu einem Gitter der Durchmesser auf Grund der Rückwärtskanten halbiert.

Achtung, dies ist eine Schätzung ohne Berücksichtigung des tatsächlichen Netzwerkaufbaus!

■ HP XC6000:

Aufsteigen im Baum bis zur Wurzel und zurück: 4

d) Vergleichen Sie Bandbreite, Latenz und Blockierungsfreiheit der beiden Netzwerke.

- BlueGene/L:
Netzwerk ist nicht blockierungsfrei, Bandbreitenengpässe können auftreten, die Latenz ist unterschiedlich je nach Verbindung
- HP XC6000:
Netzwerk ist blockierungsfrei, Bandbreite von mehr als 800 MB/s, geringe Latenz

e) Gibt es einen Flaschenhals?

- BlueGene/L:
Prinzipiell nein.
Je nach Wegewahlverfahren können aber Probleme auftreten.
- HP XC6000:
Nein, da ein „Dynamic Fat-Tree“ verwendet wird, bei dem jede Permutation geschaltet werden kann

f) **Bewerten Sie die Skalierbarkeit und Erweiterbarkeit der beiden Netzwerkvarianten.**

- BlueGene/L:
Sehr gut, das Netzwerk kann einfach um eine Ebene erweitert werden, prinzipiell unbeschränkt
- HP XC6000:
Sehr schlecht, erweiterbar um jeweils eine 2-er-Potenz, maximal 4096 angeschlossene Rechenknoten, d.h. maximal ≈ 40000 CPUs je nach Rechenknoten

g) Nehmen Sie an, die Prozessorenzahl des HP XC6000 würde an die Größenordnung der Prozessorenzahl des BlueGene/L angepasst. Welches Problem hinsichtlich der Netzwerkkommunikation ergibt sich hierbei? Insbesondere welche Veränderungen am Netzwerk müssten durchgeführt werden, damit es die Anforderungen hinsichtlich Blockierungsfreiheit weiterhin erfüllt?

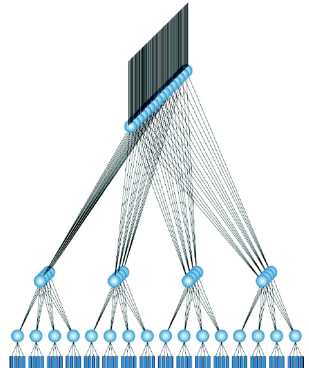
- Netzwerk besteht aus drei Schichten miteinander verknüpfter Switches
- Ebene 1 - Switche haben genausoviel Verbindungen zur nächsten Ebene wie angebundene Rechenknoten
- In jeder Ebene nimmt die Portzahl quadratisch zu
- ⇒ Netzwerkgröße limitiert durch die Größe (Portzahl) der Switche
- ⇒ Nach Erweiterung auf ~ 200000 Prozessoren müssten Switche mit mehr als 20000 Port verwendet werden.

g) Nehmen Sie an, die Prozessorenzahl des HP XC6000 würde an die Größenordnung der Prozessorenzahl des BlueGene/L angepasst. Welches Problem hinsichtlich der Netzwerkkommunikation ergibt sich hierbei? Insbesondere welche Veränderungen am Netzwerk müssten durchgeführt werden, damit es die Anforderungen hinsichtlich Blockierungsfreiheit weiterhin erfüllt?

- Netzwerk besteht aus drei Schichten miteinander verknüpfter Switches
- Ebene 1 - Switche haben genausoviel Verbindungen zur nächsten Ebene wie angebundene Rechenknoten
- In jeder Ebene nimmt die Portzahl quadratisch zu
- ⇒ Netzwerkgröße limitiert durch die Größe (Portzahl) der Switche
- ⇒ Nach Erweiterung auf ~ 200000 Prozessoren müssten Switche mit mehr als 20000 Port verwendet werden.

g) Quadrics

- Firma inzwischen insolvent
- Dieser Netzwerktyp wurde zeitweise (~ 2003) von mehreren der Top-10 der schnellsten Supercomputern der Welt genutzt.
- Quadrics QsNet III wurde nie auf den Markt gebracht



g) Quadrics

■ Weitere Netzwerktypen

- **Infiniband**
- **Gigabit Ethernet**
- **Spezielle Entwicklungen** (vgl. IBM BlueGene)
- Myrinet
- Scalable Coherent Interconnect (SCI)
- NUMalink
- HIPPI
- ...

Aufgabe 3 – Vergleich von Parallelrechnern

g) Beispiel: Infiniband-Switch mit 4096 Ports



- h) Welche Vereinfachungen im Netzwerk könnten gemacht werden, um den Aufwand für Netzwerkhardware zu verringern und was wären die Auswirkungen hiervon?
- Ausdünnung der Verbindungen in Richtung Wurzel des Baumes
 - ⇒ Keine Blockierungsfreiheit mehr
 - ⇒ Durchsatz verringert sich
 - ⇒ Redundanz geringer

Aufgabe 2

a) **Zeichnen Sie das Grundmuster der Umkehrpermutation mit 8 Eingängen und 8 Ausgängen ($n = 3$) auf.**

- Für $n = 3$ ergibt sich dasselbe Grundmuster wie bei der Kreuzpermutation.

⇒ siehe Folie 32 bzw. Folie 30!

b) **Wie ist die Bisektionsbandbreite definiert?**

- Die Bisektionsbandbreite ist die maximale Datenmenge, die das Netzwerk über die Bisektionslinie, die das Netzwerk in zwei Hälften teilt, pro Sekunde transportieren kann.

⇒ siehe Folie 6!

Aufgabe 2

a) **Zeichnen Sie das Grundmuster der Umkehrpermutation mit 8 Eingängen und 8 Ausgängen ($n = 3$) auf.**

- Für $n = 3$ ergibt sich dasselbe Grundmuster wie bei der Kreuzpermutation.

⇒ siehe Folie 32 bzw. Folie 30!

b) **Wie ist die Bisektionsbandbreite definiert?**

- Die Bisektionsbandbreite ist die maximale Datenmenge, die das Netzwerk über die Bisektionslinie, die das Netzwerk in zwei Hälften teilt, pro Sekunde transportieren kann.

⇒ siehe Folie 6!

Aufgabe 2

c) **Ordnen Sie die folgenden vier, hier alphabetisch sortierten, Begriffe den statischen oder dynamischen Verbindungsnetzen zu.**

- Statische Verbindungsnetze:
 - Baum
- Dynamische Verbindungsnetze:
 - Bus
 - Kreuzschiene
 - Schalternetzwerk

d) **Geben Sie 2 Kriterien an, mit Hilfe derer die Ausfalltoleranz (Redundanz) eines Verbindungsnetzes charakterisiert werden kann.**

- minimale Bisektionsbreite, Diskonnektivität, Wurzelknoten, Flaschenhals, Ausweichverbindungen, . . .

Aufgabe 2

- c) Ordnen Sie die folgenden vier, hier alphabetisch sortierten, Begriffe den statischen oder dynamischen Verbindungsnetzen zu.
- Statische Verbindungsnetze:
 - Baum
 - Dynamische Verbindungsnetze:
 - Bus
 - Kreuzschiene
 - Schalternetzwerk
- d) Geben Sie 2 Kriterien an, mit Hilfe derer die Ausfalltoleranz (Redundanz) eines Verbindungsnetzes charakterisiert werden kann.
- minimale Bisektionsbreite, Diskonnektivität, Wurzelknoten, Flaschenhals, Ausweichverbindungen, . . .

Aufgabe 2

e) Was ist das Problem für die Skalierbarkeit eines Verbindungsnetzes mit Baumstruktur? Und wie kann es abgemildert werden?

- Ansteigende Belastung der Verbindungen in Richtung Wurzelknoten
- Kann durch Fat-Tree abgemildert werden

f) Wofür stehen die Parameter K und n beim K -ären n -Kubus?

- K = Radius: Anzahl der Knoten, die einen Zyklus in einer Dimension bilden
- n = Dimension

Aufgabe 2

- e) Was ist das Problem für die Skalierbarkeit eines Verbindungsnetzes mit Baumstruktur? Und wie kann es abgemildert werden?
- Ansteigende Belastung der Verbindungen in Richtung Wurzelknoten
 - Kann durch Fat-Tree abgemildert werden
- f) Wofür stehen die Parameter K und n beim K -ären n -Kubus?
- K = Radius: Anzahl der Knoten, die einen Zyklus in einer Dimension bilden
 - n = Dimension

Aufgabe 2

g) Wie wird ein K -ärer n -Kubus mit $K=3$ und $n=3$ auch vereinfacht bezeichnet?

- 3D-Torus

Wie viele Knoten besitzt dieses Verbindungsnetz und wie groß ist der Knotengrad (inkl. Rückwärtskanten)?

- $N = K^n$ Knoten, hier also $N = 3^3 = 27$ Knoten
- $2 * n$ Knotengrad (mit Rückwärtskanten),
hier also $2 * n = 2 * 3 = 6$

Aufgabe 1

e) Was versteht man unter der Skalierbarkeit einer Verbindungsnetzes?

- Fähigkeit, die wesentlichen Eigenschaften des Verbindungsnetzes auch bei beliebiger Erhöhung der Knotenzahl beizubehalten.
- Vergrößerung möglich ohne die wesentlichen Eigenschaften des Netzwerks zu verlieren.

f) Zeichnen Sie das Grundmuster der Kreuzpermutation mit 8 Eingängen und 8 Ausgängen auf.

⇒ siehe Folie 30!

Aufgabe 1

e) **Was versteht man unter der Skalierbarkeit einer Verbindungsnetzes?**

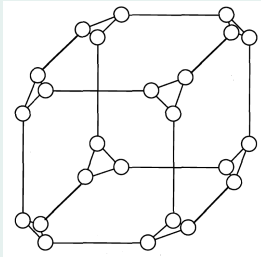
- Fähigkeit, die wesentlichen Eigenschaften des Verbindungsnetzes auch bei beliebiger Erhöhung der Knotenzahl beizubehalten.
- Vergrößerung möglich ohne die wesentlichen Eigenschaften des Netzwerks zu verlieren.

f) **Zeichnen Sie das Grundmuster der Kreuzpermutation mit 8 Eingängen und 8 Ausgängen auf.**

⇒ siehe Folie 30!

Aufgabe 1

g) Gegeben ist folgendes Ring-Würfel-Netzwerk:



Bestimmen Sie den Durchmesser sowie die minimale Bisektionsbreite.

- Durchmesser: 6
- Minimale Bisektionsbreite: 4

Aufgabe 1

h) Welche Bedingung muss bei einem fehlertoleranten Netzwerk erfüllt sein?

- Bei einem fehlertoleranten Netz muss zwischen jedem Paar von Knoten mindestens ein weiterer, redundanter Weg vorhanden sein.

i) Geben Sie für einen K -ären n -Kubus die Anzahl N der Knoten sowie den Knotengrad (mit Rückwärtskanten) der einzelnen Knoten in diesem Netzwerk an.

- Anzahl der Knoten $N = K^n$
- Verbindungsgrad: $2n$

Aufgabe 1

h) Welche Bedingung muss bei einem fehlertoleranten Netzwerk erfüllt sein?

- Bei einem fehlertoleranten Netz muss zwischen jedem Paar von Knoten mindestens ein weiterer, redundanter Weg vorhanden sein.

i) Geben Sie für einen K -ären n -Kubus die Anzahl N der Knoten sowie den Knotengrad (mit Rückwärtskanten) der einzelnen Knoten in diesem Netzwerk an.

- Anzahl der Knoten $N = K^n$
- Verbindungsgrad: $2n$

Aufgabe 2

- d) In welche Klasse bei der Klassifikation von Rechensystemen nach Flynn fallen Vektorprozessoren (Vektorrechner)? In welche Klasse fallen im Vergleich dazu nachrichtengekoppelte Parallelrechner?
- Vektorprozessoren (Vektorrechner): SIMD
 - Nachrichtengekoppelte Parallelrechner: MIMD

Vektorprozessoren und -computer und deren Programmierung

- **Klausurrelevant!**
- Thema wird in einer späteren Vorlesung sowie in einer der nächsten Übungen behandelt
- Bisher jedoch immer in einer Aufgabe zusammen mit den Themen der Übungen 5 und 6

Evaluation der RS-Übung

- In der nächsten Übung

Fragen?

Zentralübung Rechnerstrukturen im SS 2013

Verbindungsstrukturen

Martin Schindewolf, Wolfgang Karl

Lehrstuhl für Rechnerarchitektur und Parallelverarbeitung

27. Juni 2013

