
Rechnerstrukturen

Vorlesung im Sommersemester 2006

Prof. Dr. Wolfgang Karl

Universität Karlsruhe (TH)

Fakultät für Informatik

Institut für Technische Informatik



- **Kapitel 3: Multiprozessoren – Parallelismus auf Prozess/Thread-Ebene**

3.4: Verbindungsstrukturen



- **Verbindungsnetze in Multiprozessoren**
 - Ermöglichen die Kommunikation und Kooperation zwischen den Verarbeitungselementen (Knoten)
 - Zuverlässiger Austausch von Informationen
 - **Einsatz eines Verbindungsnetzwerks**
 - **Multiprozessor mit verteiltem Speicher (nachrichtenorientierter Multiprozessor)**
 - Verbinden physikalisch jeden Knoten für das Versenden von Nachrichten
 - Direkte Send/Receive-Kommunikation zwischen den Knoten
 - **Multiprozessor mit gemeinsamem Speicher**
 - Ermöglicht den Zugriff aller Knoten auf den gemeinsamen Speicher
 - Kommunikation durch Lesen und Schreiben auf gemeinsame Daten

• Charakterisierung von Verbindungsnetzwerken

– Latenz

- Übertragungszeit einer Nachricht T_{msg}
 - die Zeit, die für das Verschicken einer Nachricht von einer bestimmten Länge zwischen zwei Prozessoren benötigt wird
- Die Übertragungszeit setzt sich zusammen aus:
 - der Startzeit t_s (Message Startup Time):
 - » Die Zeit, die benötigt wird, um die Kommunikation zu initiieren
 - Transferzeit t_w pro übertragenem Datenwort:
 - » hängt von der physikalischen Bandbreite des Kommunikationsmediums ab.
 - Voraussetzung:
 - » Verbindungsnetz ist konfliktfrei, da sonst die Übertragungszeit nicht fest berechnet werden kann

- Charakterisierung von Verbindungsnetzwerken

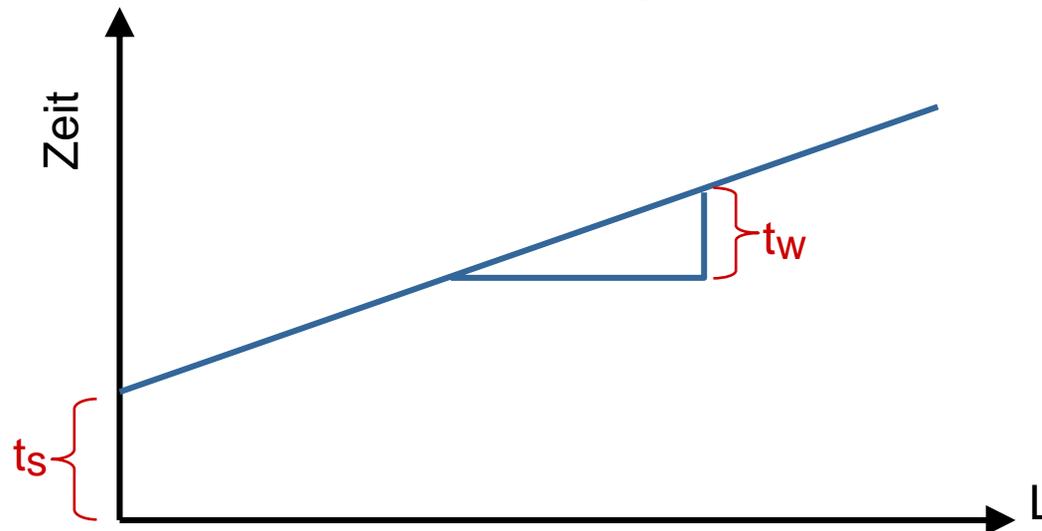
- Latenz

- Formel:

$$- T_{\text{msg}} = t_s + t_w * L,$$

mit L: Anzahl der Datenwörter

- Graphische Darstellung:



- Charakterisierung von Verbindungsnetzwerken
 - Latenz (Übertragungszeit einer Nachricht, latency)
 - Software-Overhead
 - in Verbindung mit dem Senden und Empfangen von Nachrichten

- Charakterisierung von Verbindungsnetzwerken
 - Beispiel für Software-Overhead

Erzeugerprozess:

```
send(proci, processi, @sbuffer, num_bytes)
```

Sender

Systemaufruf
Prüfe Schutzbed.
DMA Init.
DMA nach NI

Verbraucherprozess:

```
receive(@rbuffer, max_bytes)
```

Empfänger

DMA vom Netzwerk in den Puffer
BS Interrupt und Dekodierung der Nachr.
BS kopiert Systempuffer in Userpuffer
Reschedule Benutzerprozess
Lesen der Nachricht

Zeit

NI: Netzwerkschnittstelle
DMA: Direktspeicherzugriff



- Charakterisierung von Verbindungsnetzwerken
 - Latenz (Übertragungszeit einer Nachricht, latency)
 - Kanalverzögerung (channel delay)
 - Dauer Belegung eines Kommunikationskanals durch eine Nachricht
 - Kanal:
 - » Physikalische Verbindung zwischen Schalterelementen oder Knoten mit einem Puffer zum Halten der Daten während ihrer Übertragung
 - Verbindung (link)
 - » Menge von Leitungen

- Charakterisierung von Verbindungsnetzwerken
 - Latenz (Übertragungszeit einer Nachricht, latency)
 - Schaltverzögerung, Routing-Verzögerung (switching delay, routing delay)
 - Zeit, einen Weg zwischen zwei Knoten aufzubauen
 - Pfadberechnung oder Wegefindung (Routing)
 - » die Art, wie der Weg einer Nachricht vom Sender zum Zielknoten berechnet wird
 - » Zu einer Verbindungsstruktur kann es mehrere Wegefindungsalgorithmen geben
 - » einfache Implementierung in Verbindungselementen mit Hilfe eines schnellen Hardware-Algorithmus

- Charakterisierung von Verbindungsnetzwerken
 - Latenz (Übertragungszeit einer Nachricht, latency)
 - Blockierungszeit (contention time)
 - Wird verursacht, wenn zu einem Zeitpunkt mehr als eine Nachricht auf eine Netzwerkressource zugreifen
 - Blockierung (contention)
 - Ein Verbindungsnetzwerk heisst blockierungsfrei, falls jede gewünschte Verbindung zwischen Prozessoren oder zwischen Prozessoren und Speichern unabhängig von schon bestehenden Verbindungen hergestellt werden kann

- **Charakterisierung von Verbindungsnetzwerken**
 - **Durchsatz oder Übertragungsbandbreite (bandwidth)**
 - Maximale Übertragungsleistung des Verbindungsnetzwerkes oder einzelner Verbindungen, meist in Megabits pro Sekunde (MBit/s) oder Megabytes pro Sekunde (MB/s)
 - **Bisektionsbandbreite (bisection bandwidth)**
 - Maximale Anzahl von Megabytes pro Sekunde, die das Netzwerk über die Bisektionslinie, die das Netzwerk in zwei gleiche Hälften teilt, transportieren kann

- **Charakterisierung von Verbindungsnetzwerken**
 - **Diameter oder Durchmesser r (*diameter*):**
 - maximale Distanz für die Kommunikation zweier Prozessoren, also die Anzahl der Verbindungen, die durchlaufen werden müssen. Man spricht auch von der maximalen Pfadlänge zwischen zwei Knoten.
 - **Verbindungsgrad eines Knotens P (node degree, connectivity)**
 - ist definiert als die Anzahl der direkten Verbindungen, die von einem Knoten zu anderen Knoten bestehen.
 - **Mittlere Distanz d_a (average distance) zwischen zwei Knoten**
 - Anzahl der Links auf dem kürzesten Pfad zwischen zwei Knoten
 - P/d_a ist die maximale Anzahl neuer Nachrichten, die von jedem Knoten in einem Zyklus in das Netzwerk eingebracht werden können



- **Charakterisierung von Verbindungsnetzwerken**
 - **Komplexität oder Kosten:**
 - Kosten für die Implementierung einer Hardware
 - Aufwand für das Verbindungsnetz gemessen in der Anzahl und der Art der Schaltelemente und Verbindungsleitungen.
 - **Erweiterbarkeit:**
 - Multiprozessoren können begrenzt, stufenlos oder nur durch Verdopplung der Anzahl der Prozessoren erweiterbar sein.
 - **Skalierbarkeit:**
 - Fähigkeit, die wesentlichen Eigenschaften des Verbindungsnetzes auch bei beliebiger Erhöhung der Knotenzahl beizubehalten.

- Charakterisierung von Verbindungsnetzwerken

- Ausfallstoleranz oder Redundanz :

- Verbindungen zwischen Knoten sind selbst dann noch zu schalten, wenn einzelne Elemente des Netzes (Schaltelemente, Leitungen) ausfallen.
- Ein fehlertolerantes Netz muss also zwischen jedem Paar von Knoten mindestens einen zweiten, redundanten Weg bereitstellen.
Die Eigenschaft eines Systems, bei Ausfall einzelner Komponenten unter deren Umgehung funktionstüchtig zu bleiben, wenn auch mit verminderter Leistung, wird als **Graceful degradation** bezeichnet.

- Charakterisierung von Verbindungsnetzwerken
 - Art der Adressierung:
 - zielbasiert (destination-based routing):
 - Kopfteil eines Pakets (oder einer Nachricht) wird mit einer systemweit eindeutigen Empfängeradresse versehen, die bei der Wegefindung von jedem Zielknoten zur Auswahl eines Übertragungskanals genutzt wird
 - quellenbasiert (source-based routing):
 - Paket wird mit allen Informationen versehen, um über die Zwischenknoten zum Empfänger zu gelangen. Für jeden Zwischenknoten wird im voraus die Abzweigung bestimmt, die das Paket nehmen muss.

- Charakterisierung von Verbindungsnetzwerken

- Art des Datentransfers:

- Durchschalte- oder Leitungsvermittlung (circuit switching):

- Eigenschaft eines Netzes eine direkte Verbindung zwischen zwei oder mehreren Knoten eines Netzes zu schalten. Die physikalische Verbindung bleibt für die gesamte Dauer der Informationsübertragung bestehen.

- » Blockierungsfreie Kommunikation

- » Kurze Latenz

- » Gut geeignet für lange Nachrichten, da die Zeit zum Aufsetzen einer Nachricht im Verhältnis zur Übertragungszeit kurz ist

- » Übertragungszeit einer Nachricht der Länge L über eine Distanz d beträgt: $L/b + d\delta$, mit individueller Schaltverzögerung δ und der Bandbreite b eines Kanals



- Charakterisierung von Verbindungsnetzwerken
 - Art des Datentransfers:
 - Paketvermittlung (packet switching):
 - Datenpakete fester Länge oder Nachrichten variabler Länge werden entsprechend einem Wegefindungsalgorithmus (routing) vom Absender zum Empfänger geschickt
 - Nachrichten mit Adresse und Daten werden durch das Netzwerk verschickt
 - » Adresse wird in jedem Knoten gelesen und die Nachricht wird zum nächsten Knoten weitergeleitet, bis die Nachricht das Ziel erreicht
 - » Günstig für kurze Nachrichten

- Charakterisierung von Verbindungsnetzwerken

- Art des Datentransfers:

- Paketvermittlung (packet switching):

- Übertragungsmodi: Store-and-forward-Modus

- » Jeder Knoten enthält einen Puffer zum Aufnehmen der vollständigen Nachricht
- » Nachricht wird von jedem Zwischenknoten in Empfang genommen, vollständig zwischengespeichert und dann weiter übertragen
- » Nachfolgende Pakete werden nacheinander verschickt
- » Gegenüber Circuit Switching: höhere Bandbreite, aber auch höhere Latenz
- » Übertragungszeit einer Nachricht der Länge L über eine Distanz d von einer Quelle zum Ziel beträgt $d(L/b + \delta)$

- Charakterisierung von Verbindungsnetzwerken
 - Art des Datentransfers:
 - Paketvermittlung (packet switching):
 - Übertragungsmodi: Cut-through oder wormhole
 - Phit und Flusskontrolle
 - » Eine Nachricht selbst wird in eine Anzahl von Übertragungseinheiten (phits – physical transfer units – oder auch flits – flow control digits – genannt) zerlegt.
 - » Ein Phit ist dabei die Datenportion, die zu einem Zeitpunkt zwischen zwei Knoten übertragen werden kann.
 - » Bei der Nachrichtenübertragung zwischen nicht benachbarten Sender- und Empfängerknoten sind Puffer nötig.

- Charakterisierung von Verbindungsnetzwerken
 - Art des Datentransfers:
 - Paketvermittlung (packet switching):
 - Übertragungsmodi: Virtual-cut-through-Modus:
 - » Nachricht wird aufgeteilt in Zellen (Flits) fester Größe
 - » Der Kopfteil der Nachricht enthält die Empfängeradresse und bestimmt den einzuschlagenden Weg. Flits mit Daten folgen dem Kopf auf dem Pfad von der Quelle zum Ziel
 - » Bei Ankunft des Kopfs einer Nachricht wird dieser dekodiert . Nachfolgende Flits werden automatisch an den nächsten Knoten auf dem ausgewählten Pfad weitergeleitet , ein Flit pro Zeiteinheit gemäß einer Pipeline-Verarbeitung
 - » Kopf-Information wird festgehalten bis letztes Flit angekommen ist.
 - » ankommende Daten werden nur im Konfliktfall im Knoten vollständig zwischengespeichert.
 - » In jedem Knoten werden Puffer bereit gehalten, die auch ein maximal großes Nachrichtenpaket zwischenspeichern können

- Charakterisierung von Verbindungsnetzwerken
 - Art des Datentransfers:
 - Paketvermittlung (packet switching):
 - Übertragungsmodi: Wormhole-routing-Modus:
 - » solange keine Übertragungskanäle blockiert sind, mit den Virtual-cut-through-Modus identisch.
 - » Falls der Kopfteil der Nachricht auf einen Kanal trifft, der gerade belegt ist, wird er abgeblockt. Alle nachfolgenden Übertragungseinheiten der Nachricht verharren dann ebenfalls an ihrer augenblicklichen Position, bis die Blockierung aufgehoben ist. Durch das Verharren werden die Puffer nachfolgender Kanäle auch für weitere Nachrichten blockiert.

- Charakterisierung von Verbindungsnetzwerken
 - Art des Datentransfers:
 - Paketvermittlung (packet switching):
 - Übertragungsmodi: Buffered wormhole routing:
 - » Kompromisslösung zwischen Virtual-cut-through- und Wormhole-routing-Modus eingesetzt,
 - » begrenzter Puffer zur Aufnahme kleinerer Pakete vorhanden
 - » größere Pakete werden im Blockierungsfall – ähnlich dem Wormhole-routing-Modus – in den Puffern mehrerer Knoten zwischengespeichert.

- **Topologie**

- **Statische Verbindungsnetzwerke**

- Die Prozessoren sind mit jedem anderen Prozessor direkt verbunden
- Fest installierte Verbindungen zwischen Paaren von Netzknoten
 - Die Knoten sind mit Punkt-zu-Punkt-Verbindungen verbunden, die sich während der Programmausführung nicht ändern

- **Dynamische Verbindungsnetze**

- Die Knoten sind über Schaltelemente miteinander verbunden, die konfiguriert werden können, um den Kommunikationsanforderungen des ausführenden Programms zu genügen
 - Direkte fest installierte Verbindungen zwischen den Knoten existieren nicht