



Lösungen zum Übungszettel 12

Aufgabe 1:

- a) Load Balancing fordert, dass die Last zwischen den Knoten zu jedem Zeitpunkt bezüglich eines quantitativen Maßes ausgeglichen ist. D.h., LB beschäftigt sich mit der Fragestellung: Wie verteile ich die Last (Prozesse/Aufgaben) in verteilten Computer-Systemen so um, dass die Prozesse von den schwer geladenen Systemen auf die leicht geladenen Systeme verschoben werden. Der Unterschied zwischen LB und LS ist, dass LB impliziert, dass alle Prozessoren nahezu dieselbe Last (load) haben. LS erlaubt auch Prozessoren mit höherer und niedrigerer Last ihre Last zu teilen.
- b) LB Algorithmen können je nach Informationen, die vom System den Aufgaben und Benutzer bereitgestellt werden, eine statische, dynamische oder adaptive Lastverwaltung aufweisen.
 - a. Senderinitiierte Lastverteilung
 - b. Empfängerinitiierte Lastverteilung
 - c. Hybrides System (symmetrisch initiierte Algorithmen)
- c) Es gibt verschiedene Verfahren:
 - a. DNS-Variante: Das klassische Load Balancing ist das DNS selber. Dazu wird dem DNS-Server der eigenen Domain mehrere IP-Adressen eintragen, unter der der Host-Name zu erreichen ist. Hinter jeder IP-Adresse befindet sich ein separater und eigenständiger Server. Die Anfragen der Clients bedient der DNS-Server der Reihe nach mit den eingetragenen IP-Adressen.
 - b. Round-Robin-Verfahren: Das Round-Robin-Verfahren kommt mit einer einzigen IP-Adresse aus. Anstatt des DNS-Servers übernimmt ein NAT-Proxy die Lastverteilung. Anstatt einer Liste mit den verfügbaren Servern leitet der Proxy alle Anfragen an die ihm bekannten Zielsysteme weiter. Dabei merkt er sich welche IP-Adresse mit welchem Server eine Verbindung hatte und leitet eine erneute Anfrage an diesen Server weiter.
 - c. NAT mit Feedback
 - d. URL-basiertes Verfahren
 - e. Dienst-basiertes Verfahren: Meist laufen mehrere Dienste, wie http, FTP und E-Mail auf ein und demselben Server. Unter Last erweist sich der Parallelbetrieb als Problem, wenn ein Dienst allen anderen Diensten Rechenleistung klaut. Alle Dienste benutzen unter TCP einen eigenen Port, anhand dem ein Datenpaket einer Anwendung oder einem Dienst zugewiesen wird. Wenn man die Dienste auf unterschiedlichen und eigenständigen Servern betreibt, lässt sich die Last verteilen.

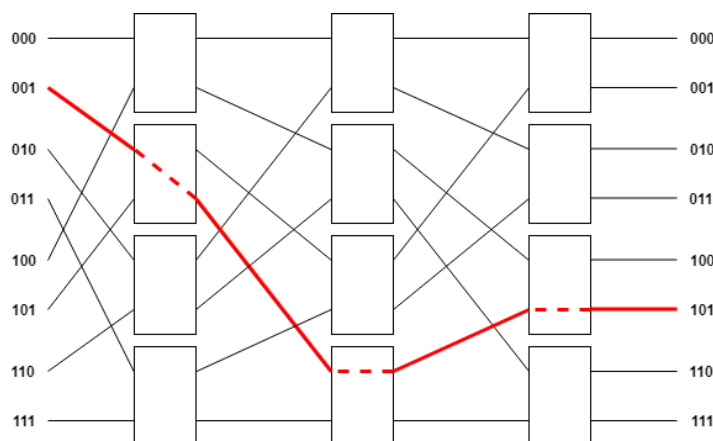
Aufgabe 2:

- a) Vermeiden von Wartezeiten, Staus und Deadlocks
- b) Deterministischer Ansatz: Wahl des Pfades nur abhängig von Start- und Zielknoten der Nachricht
Adaptiver Ansatz: Mehrere Pfade werden unter Berücksichtigung der Auslastung des Netzwerks berechnet.

Bei beiden Ansätzen gibt es zwei Varianten:

- minimale Algorithmen: wählen immer den kürzesten Weg
- nicht-minimale Algorithmen: erlauben auch Umwege, um Staus zu vermeiden

c)



Aufgabe 3:

- a) Serielle I/O verursacht ein Bottleneck, da die Berechnung stets auf die I/O warten muss. Die Leistung der CPU-Kerne wird verschwendet und wenig Arbeit verrichtet. Parallele I/O führt mehrere I/O Operationen gleichzeitig aus. Dabei greift parallele I/O gleichzeitig auf Daten auf der Festplatte zu, anstatt die I/O Anfrage seriell auszuführen. Dies ermöglicht dem System höhere Schreibgeschwindigkeiten und Bandbreite.
Bei paralleler I/O ist ein Teil der logischen Kerne auf dem Multicore-Chip für die Verarbeitung von I/O der virtuellen Maschinen vorgesehen und die restlichen Kerne bedienen die anderen Anwendungen. Dies ermöglicht es dem Prozessor, mehrere Lese- und Schreiboperationen gleichzeitig zu verarbeiten. Parallele I/O hilft, I/O-Bottlenecks zu vermeiden, die den Datenfluss stoppen oder beeinträchtigen können.
- b) Üblicherweise verwendet man Multicore CPUs in virtuellen Servern. Die umfangreiche parallele Berechnung wird meist auf mehreren virtuellen Maschinen ausgeführt, die wiederum virtuelle Speicher für parallele I/O benötigen. Bei paralleler I/O sind die Datenabhängigkeiten zu prüfen, um Wartezeiten zu reduzieren.

Hinweis: Die Lösungen sollen in PDF-Form, bzw. Code bis zum Montag (10 Uhr) der jeweils folgenden Woche per Mail an mfriedrichs@techfak.uni-bielefeld.de abgegeben werden. Zu Beginn des nächsten Übungstermins werden diese in offener Runde vorgestellt und diskutiert.