IT-System Engineering & Operation

Patrick Bucher

Contents

1 Das Data Center

1.1 Bestandteile Data Center

- Lüftung (Zu- und Abluft, Wärmetauscher)
- Hochwasserschutz (erhöhte Bauweise)
- Zutrittskontrolle an den Eingängen, Überwachungskameras
- Stromversorgung
  - USV: unterbrechungsfreie Stromversorgung (Energiespeicher: Batterien)
  - Dieselgenerator als Notstromaggregat (Energiespeicher: Dieseltank), mit Kühlung und Abluft
- Server in Serverracks
- Stromverteilung
- Datenleitung/Netzwerk
- Löschanlagen
- Administration/Überwachung

1.2 Klimatisierung

- optimale Temperatur: 26°C
  - keine Schäden bei leicht erhöhter Raumtemperatur (gegenüber 21°C)
  - Wärmeenergie geht von selber an die Umgebung (Heizung benachbarter Räumlichkeiten)
  - im optimalen Leistungsbereich der Klimaanlagen
  - Kondenswasser bei zu tiefen Temperaturen
- Staub und Pollen können schädlich sein
  - verstopfen Ventilatoren (gesteigerter Stromverbrauch durch erhöhte Kühlleistung)
  - Metallpartikel können Schäden an Hardware verursachen
- Probleme
  - Kondenswasser: Ablauf kann verstopfen, Kondenswasser auslaufen
  - Filterkontrolle: verstopfte Filter verursachen erhöhte Leistungsaufnahme
  - zusätzlicher Energieverbrauch
  - Luftverteilung

  - Überwachung
- Kühlluftverteilung
  1. Free-Flow-Systeme
     - Warme Luft steigt auf, kalte Luft sinkt ab
     - Gemischte Lufttemperatur
     - einfach
     - Problem: möglicher Wärmekurzschluss (warme Abluft wird als Kühlluft angesogen)
  2. Kalt- oder Warmgang-Einhausung
     - Trennung von Warm- und Kaltluft
     - dadurch bessere Energieeffizienz
     - aber teurer im Aufbau
     - Front der Racks sollten komplett abgeschlossen sein, um Warm- und Kaltluft voneinander zu trennen
- Immersion Cooling: flüssigkeitsgekühlte Systeme
  - mit Wärmetauscher und Flüssigkeit in Leitungskabel
  - oder komplett in Öl eingelegt

1.3 EDV-Einbau

- Serverracks
  - verschiedene Höhen (21-49U), Breiten (0.6-1m) und Tiefen (0.8-1.2m)
    - 1 HE = 1 U = 1.75 Zoll = 44.45 mm
  - auch mit integrierter Kühlung
  - Zuleitungen: oben, unten, seitlich
  - Standard: 19 Zoll (48.26 cm)
- Netzwerk
  - Kupfer (gegenwärtig stark verbreitet)
  - Glasfaser (löst Kupfer derzeit ab)
- Klimageräte, USV und Batterieschränke
  - Batterien sind sehr schwer, spezielle Racks/Bodenverstärkung erforderlich
- Kühlleitungen und Überwachungsgeräte

1.4 Kritische Punkte

- Einbruch, Diebstahl, Vandalismus, Sturmschäden, Trümmer
  - bauliche Massnahmen: stabile Aussenhülle
  - verschlossen mit Zaun
  - teilweise fernab von anderen Gebäuden
  - keine oder kaum Fenster
- Fremdzugriff
  - Zutrittskontrolle (biometrisch, Chip-Karten, Passwörtern)

  - Abhörsicherheit (elektromagnetische Abschirmung, keine mobilen Endgeräte mit Netzwerkverbingungen zulassen, keinen WiFi-Access-Point)
  - Firewall
- Feuer und Rauch
  - Branderkennung
  - Löschanlage: CO2 (Vorwarnzeit zur Flucht nötig!), Verringerung des Sauerstoffanteils der Luft auf ca. 10% (nicht tödlich, aber das Feuer verlöscht) durch Stickstoff (gefährlicher und günstig) oder Inergen (weniger gefährlich und teurer)
  - Handfeuerlöscher: CO2
    - Feuer benötigt: Sauerstoff, Hitze und Brennstoff
  - Abschottung einzelner Zellen
  - automatische Abschaltung der Klimaanlage damit der Rauch nicht verteilt wird
  - kein PVC (bildet Salzsäure!) verwenden
- Netzausfälle, Netzstörungen
  - Netzfilter (in Netzteilen integriert)
  - vorgeschaltete USV mit Batterien
  - Diesel-Generatoren
- Elektromagnetische Störfelder
  - EMP: elektromagnetische Impulse (durch Atombomben oder spezielle Generatoren verursacht), kann Geräte zerstören
  - Abschirmung (kann teuer sein)
  - metallische Aussenfassade
  - Blitzableiter
- Staub, Schmutz, Wasser
  - Filteranlagen
  - Schmutzschleusen, spezielle Teppiche
  - erhöhte Bauweise
  - Standortwahl (nicht in Nähe von Gewässern oder mit Steinschlag und Lawinen)
  - Pumpanlagen zum Abpumpen bei Überschwemmungen

1.5 Überwachungsgebiete

- Gebäude
  - Türen (offen/geschlossen)
  - Kameras
  - Bewegungsmelder
  - Zutritte
- Räume
  - Temperatur
  - Luftfeuchtigkeit
  - Bewegung
  - Rauch
  - Brand

  - Wasserlecks
- Energieversorgung
  - Netzausfall
  - Strom, Spannung, Leistung
  - Leistungsfaktor (Kosinus Phi)
- Geräte
  - Niederspannungsverteilungen
  - Schalterstellungen (Ein/Aus)
  - Stromverbrauch einzerlner Bereiche
  - Sicherungsausfall
  - Kurzschluss
  - Überlast
- Generator
  - Kraftstoffstand (Dieseltank)
  - Funktionsbereitschaft
  - Temperatur
  - Überlast
- Klimageräte
  - Temperaturen
  - Luftfeuchtigkeit
  - Übertemperatur
  - Filterwiderstand
  - Störungen
- USV-Anlagen
  - Normalbetrieb
  - Batteriebetrieb
  - Bypass-Betrieb
  - Ladezustand
  - Batterietemperatur
- Brandmelde- und Löschanlage (Zustandsanzeigen)
  - Löschanlage ausgelöst
  - Übertragungseinrichtung ausgeschaltet
  - Störung
  - Service

1.6 Rechenzenter-Effizienz, PUE-Faktor

- PUE: Power Usage Effectiveness
- Massstab für die Effizienz eines Rechenzentrums
- PUE = gesamte vom Rechenzentrum verbrauchte Energie / Verbrauch der IT-Geräte
  - 1.0: optimal (in kalten Regionen möglich)
  - 1.2: guter Wert (normale Rechenzentren)
  - über 1.4: Optimierungsbedarf

- Stichwort “Green IT”

1.7 Repetitionsfragen

1. Notieren Sie zu 5 beliebigen Bausteinen eines Rechenzentrums die folgenden Punkte:

Baustein	Funktionen	Gefährdet durch	Abhilfe gegen Gefährdungen
Gebäude	Schutz der Server vor äusseren Einflüssen	Umweltkatastrophen	Resistente Bauweise
Klimatisierung	Schutz vor Überhitzung	Verunreinigung der Filter, Kondenswasser	Filterservice, Abpumpvorrichtung
Stromversorgung	Bereitstellung von elektrischer Energie	Stromausfälle, Netzschwankungen	USV mit Batterie, Diesel-Generatoren
Netzwerk	Verbindung der Komponenten	Ausfall, Überlastung, Überhitzung, Brand	Redundanz, Datensicherung, Lastverteilung, Kühlung, Löschanlage
Eingangskontrolle	Gewährung und Verweigerung von Einlass	unautorisierte Personen	Biometrie, Überwachungskameras, Chipkarten, Passwörter, Personenkontrolle

2. Versuchen Sie den Kostenanteil pro Baustein am gesamten RZ abzuschätzen.
   - Gebäude: ca. 10 Millionen CHF (92%)
   - Klimatisierung: ca. 250’000 CHF (2.3%)
   - Stromversorgung: ca. 100’000 CHF (1%)
   - Netzwerk: ca. 500’000 CHF (4.6%)
   - Eingangskontrolle: 25’000 CHF (0.2%)
   - Summe: 10’875’000 CHF (100%)
3. Was ist der PUE Faktor und was sind die erreichbaren und effektiv erreichten Werte?
   - PUE bedeutet Power Usage Effectiveness und Massstab für die Effizienz eines Rechenzentrums. Er errechnet sich aus der gesamthaft durch das Rechenzentrum verbrauchten Energiemenge geteilt durch die gesamthaft von den IT-Geräten verbrauchte Energie.
     - 1.0: optimal (in kalten Regionen möglich)
     - 1.2: guter Wert (normale Rechenzentren)
     - über 1.4: Optimierungsbedarf

1.8 Verfügbarkeitsverbesserungen

Kosten der Downtime pro Stunde:

- Fertigung: 28'000.-
- Logistik: 90'000.-
- Einzelhandel: 90'000.-
- Home-Shopping: 113'000.-
- Medien (pay per view): 1'100'000.-
- Bank (Rechenzentrum): 2'500'000.-
- Kreditkartenverarbeitung: 2'600'000.-
- Broker: 6'500'000.-

1.8.1 Service Level Agreement

Verfügbarkeit bei `7*24h`:

Uptime in %	Downtime pro Jahr
90%	876 h (36.5 d)
95%	438 h (18.25 d)
99%	87.6 h (3.65 d)
99.9%	8.76 h
99.99%	52.56 min
99.999%	5.256 min
99.9999%	31.536 sec

Verfügbarkeit bei `5*9h` (zu Bürozeiten):

Uptime in %	Downtime pro Jahr
90%	234.90 h (26.1 d)
95%	117.45 h (13.05 d)
99%	23.49 h (2.61 d)
99.9%	2.35 h
99.99%	14.09 min
99.999%	1.14 min
99.9999%	8.46 sec

Massnahmen zur Erhöhung der Verfügbarkeit:

- Spiegelung (inkl. Synchronisation)
- Failover Cluster: Ausfall eines Hosts, der vom Client nicht bemerkt wird

  - zwei Hosts, die sich gegenseitig über Heartbeat kontrollieren
  - Client spricht zu einem vorgeschalteten Virtual Host
  - bei Ausfall springt der eine Host für den anderen ein
  - Automatischer Lastausgleich bei vielen Hosts
- Failover Datacenter: Ausfall eines ganzen Datacenters
  - Spiegelung der Datacenters
  - Backup und Produktivdaten über Kreuz, sodass bei einem Ausfall beide Datenbestände an einem Ort sind
- Asynchrone und synchrone Replikation
  - in Rechenzenter A wird eine Datenbank synchron auf eine lokale Instanz gespiegelt
  - in Rechenzenter B wird die Datenübertragung dann asynchron vorgenommen
  - Rechenzenter B ist im Standby-Betrieb

1.8.2 Availability Environment Classification

- AEC-0: Conventional
  - Funktion kann unterbrochen werden
  - Datenintegrität nicht essentiell
- AEC-1: Highly Reliable
  - Funktion kann unterbrochen werden
  - Datenintegrität muss gewährtleistet sein
- AEC-2: High Availability
  - Funktion darf nur zu festgelegten Zeiten unterbrochen werden
  - Zu Hauptbetriebszeiten sind minimale Unterbrüche zulässig
- AEC-3: Fault Resilient
  - Funktion muss zu Hauptbetriebszeiten ununterbrochen aufrecht erhalten werden
- AEC-4: Fault Tolerant
  - Funktion muss ununterbrochen (24/7) aufrecht erhalten werden
- AEC-5: Disaster Tolerant
  - Funktion muss unter allen Umständen verfügbar sein

1.8.3 Repetitionsfragen

1. Welche Verfügbarkeit muss im SLA festgehalten werden, wenn ich 1h Ausfallzeit während den Bürozeiten nicht überschreiten will?
   - 99.9572% Verfügbarkeit, 8-17 Uhr von Mo-Fr CET
2. Was versteht man unter Failover-Cluster-Services?
   - Eine automatische und für den Client transparente Umschaltung eines redundanten Hosts bei Störungen.
3. Wenn z.B. das SAN gespiegelt werden soll, wie erhöhen sich die Kosten? 50%/100%/mehr als 100% und warum?

- Mehr als 100%, weil die Anzahl der Verbindungen zwischen den einzelnen Komponenten sich mehr als verdoppelt.

1.9 Tier-Levels

- Tier I: Redundanz N: keine Redundanz, 28.8h Ausfallszeit pro Jahr
- Tier II: Redundanz N+1: ein zusätzlicher Server, 22h Ausfallszeit pro Jahr
- Tier III: Redundanz N+1: ein zusätzlicher Server, weitere Redundanzen in der Infrastruktur, 1.6h Ausfallszeit pro Jahr
- Tier IV: Redundanz 2(N+1): ein zusätzlicher Server, alle Server doppelt, 0.8h Ausfallszeit pro Jahr

1.9.1 Rechercheaufgaben

1. Suchen sie RZ-Services Anbieter mit Level 3, 3.5 und Level 4 Rechenzentren.
   - Tier III: infomaniak
   - Tier 3+: greeendatacenter
   - Tier IV: greendatacenter
2. Versuchen sie die Kosten für den Service zu bestimmen (pro Rack, pro Server, ...).
   - monatlich CHF 450.- pro Monat und Rack (10 Server)
   - monatlich CHF 650.- pro Monat und Rack (? Server)
   - monatlich CHF 1250.- pro Monat und Rack (? Server)

1.10 Informaton Lifecycle Management

Datenzyklus:

- Create (erstellen)
- Store (abspeichern)
- Use (verwenden: einsehen, anpassen)
- Share (weitergeben)
- Archive (archivieren)
- Destroy (löschen)

ILM: Speicherstrategie zur Speicherung von Informationen *entsprechend ihrem Wert* auf dem jeweils günstigsten Speichermedium.

- Verwaltung und Speicherung orientieren sich an *Wichtigkeit*, *Wertigkeit* und *Kosten* der Daten.
- Daten, Quellen und Speichersysteme werden *klassifiziert* (Speicherhierarchie).
  - Tier 1: SSD, Server-Festplatten (Fiber Channel Disc 15k rpm)
  - Tier 2: HDD,
  - Tier n: SATA-Festplatten

  - spezialisiert: Archivierung, Tapes, langsame Festplatten (disc to disc)

ILM-Management:

- Storage Management
- Document Lifecycle Management
- Content Lifecycle Management
- Records Management

Regeln legen fest, wie und wo Daten gespeichert werden:

- Änderungshäufigkeit
- Zugriffsgeschwindigkeit
- Zugriffshäufigkeit
- Kosten
- ökonimischer Wert
- gesetzliche Bestimmungen

Inaktive Daten:

- konsumieren Speicherplatz
- werden gepflegt, gesichert, repliziert usw.
- unterliegen rechtlichen und Datenhaltungsansprüchen
- müssen im Katastrophenfall wiederhergestellt werden

1.10.1 Repetitionsfragen

1. Was versteht man unter Records Management?
   - Was soll wo und wie lange gespeichert und von wem eingesehen oder bearbeitet werden (rechtliche Aspekte).
2. Welche (gesetzlichen) Vorschriften für die Datenaufbewahrungszeit kennen sie?
   - 3 Jahre für intern relevante Daten
   - 10 Jahre für Rechnungen, Geschäftsabschlüsse
   - 20 Jahre bei börsenkotierten Unternehmen
3. Wer soll das Records-Management (RM) anordnen und durchsetzen?
   - Rechtsabteilung: sanktionieren
   - Geschäftsleitung: anordnen
   - Abteilungsleiter: durchsetzen
   - Administratoren: ausführen
4. Kennen sie aus der eigenen Umgebung Beispiele?
   - keine Positivbeispiele

1.11 Tiered Storage

Daten können nach Nutzung auf verschiedener Hardware gespeichert werden:

- SSD: fast
- DAS/SAN: active
- HDFS/NAS: historical
- Amazon S3/HDFS/NAS: archived (on- or offline)

Verschiedene Datenklassen werden auf verschiedene Speicherklassen gespeichert:

1. mission critical data (z.B. Online-Datenbank mit Bestellwesen)
2. business critical data (Daten, die immer zur Verfügung stehen müssen)
3. nearline or historical data (z.B. alte Pläne in einem Architekturbüro zum gelegentlichen Nachschauen)
4. offline data (z.B. Backups und Daten, die nur aufgrund gesetzlicher Bestimmung aufbewahrt werden)

TODO p.41-44

MAID: massive array of idle disks (Festplatten können zum Stromsparen heruntergefahren werden und/oder langsamer laufen)

RTO: recovery time objectives (Ziele betreffend Dauer der Rückspielbarkeit einer Datensicherun einer Datensicherung) RPO: recovery point objectives (Ziele betreffend Dauer zwischen Datensicherungen)

1.11.1 Übung Allocation Efficiency

Alloziert: Allokation von 80% bedeutet, dass bei 5 RAID-Platten 4 verwendet und 1 für Redundanz benötigt wird

Je höher die Belegung, desto höher der Yield (die Ausbeute). TODO: p.47

[Siehe Tabelle mit Berechnung]

2 Netzwerke im Rechenzentrum

Netzwerktopologie:

- Provider
  - Angebot
  - Speed
  - Technik
- Grenze
  - Router
  - Firewall
  - IDP (Intrusion Detection and Protection)
  - Redundanz

- DMZ (Demilitarized Zone)
  - Web-Services
  - Authentifizierung
  - Dienste
- Lokales Netzwerk (LAN)
  - Topologien
  - Speed
  - Trennungen
  - Services

Private IP-Bereiche:

- 10.0.0.0/8
- 172.16.0.0/12
- 192.168.0.0/16
- Router
  - oft in Firewall oder Layer-3-Switches eingebaut
  - Arbeitet mit IP-Paketen, setzt öffentliche in private Adressen um (NAT)
  - Anbendung verschiedener Netze
  - optionale Weiterleitung bei redundanten Leitungen
  - Aufrechterhaltung von QOS durch spezifische Weiterleitungen (aufgrund Pakettyp, Protokoll)
  - kann VPN-Endpunkt sein
  - Anpassung an unterschiedliche Netzwerktechniken
- Firewall
  - Sicherheitsbaustein, Teil des Sicherheitskonzepts
  - Übernimmt meist auch Routing-Funktionen
  - regelbasierte Sperrung/Weiterleitung von Paketen
  - VPN-Endpunkt mit Authentifizierung
  - kann auf allen OSI-Schichten arbeiten
- IDP
  - IDS (Intrusion Detection System) + IDP (Intrusion Prevention System) = IDP (Intrusion Detection and Protection)
  - Eindringungsversuche erkennen (Muster, DOS, Fakes, Portscan, IP-Spoofing)
  - Rechenintensiv, oft separate Hardware
  - über 1Gbit/s-Netzwerke können nicht komplett überwacht werden (mehrere Geräte oder Heuristik)
- DMZ
  - demilitarized zone
  - Geschützer Bereich, in dem bestimmte Zugriffe (www, Mail, FTP) erlaubt weden

  - 1 oder 2 Firewalls
  - Model 1: eine Firewall, an der Internet, DMZ und Inside-Zone hängen
  - Model 2: zwei Firewalls: Internet, Firewall, DMZ, Firwall, Inside Zone
  - bei Banken müssen die beiden Firewalls verschiedener Hersteller sein
- Netzwerk-Redundanz
  - 1 Provider/2 Zugänge
  - 2 Provider/je 1 Zugang
  - Ausfallüberwachung nötig
  - Getrennte Wege
  - verschiedene Medien
  - Loadbalancing
- LAN-Strukturen (physisch)
  - TOR: top of rack (jedes Rack hat seine eigenen Switches)
    - neue Switches für jedes neue Rack notwendig
  - EOR: end of row (Racks nur über Patchpanels verbunden)
    - einfacher, solange genügend Ports vorhanden sind (aber mehr Verkabelungsaufwand nötig)
- LAN-Strukturen
  - mehrere Switch-Ebenen
  - peering/peripherie
  - backbone/spine/core
  - leaf (Anschluss für Server)
- MPLS: multiprotocol label switching
  - VPN-ähnliche Strukturen zur Verbindung zusammengehöriger Netzwerke ohne Rücksicht auf IP-Segmente
  - Paketvermittlung durch den Provider aufgrund von Labels in den Paketen
    - QOS (quality of service)
    - COS (class of service)
      - realtime (voice)
      - best effort
      - bulk (grosse Mengen)
      - business critical
      - video
- SDN: Software Defined Networking
  - zwei Ebenen
    - Control Plane
    - Data Plane
  - Netzwerke werden via Management-Software erstellt

  - Netze sind teilweise virtuell basierend auf einem Trägernetzwerk und werden mit speziellen Protokollen konfiguriert und verbunden (möglicher Standard: Openflow)

- VLAN: virtual LAN
  - Bildung getrennter Netze auf gemeinsamer Hardware
  - tagged oder untagged
  - Trunk: Verbindung zweier Switches

3 Virtualisierung im Data Center

3.1 SMP: Symmetrische Multiprozessoren

- zwei oder mehr gleichartige Prozessoren mit vergleichbaren Möglichkeiten
- Prozessoren zeilen sich das Memory und sind über einen Bus oder andere interne Verbindungen zusammengeschaltet
- Prozessoren teilen sich I/O-Geräte
- Alle Prozessoren können die gleichen Funktionen ausführen (daher “symmetrisch”)
- System wird durch ein integriertes OS (Microcode auf dem Chip) kontrolliert
  - stellt Interaktion zwischen Prozessoren und deren Programmen her
  - auf Stufe Job, Task, File und Daten-Elementen

3.1.1 Vorteile von SMP

- Performance durch parallele Ausführung
- Skalierbarkeit durch Geräte mit unterschiedlichem Preis-Leistungs-Verhältnis
- Verfügbarkeit
- TODO p.6

3.1.2 SMP-Organisation

- L1- und L2-Cache auf dem Prozessor
- L3-Cache geteilt
- Prozessoren über Bussystem verbunden

3.2 Multi-Core-Prozessoren

- bekannt als Chip-Multiprozessor
- zwei oder mehr Cores auf einem Chip
- jeder Core beinhaltet alle Komponenten eines unabhängigen Prozessors
  - auch L1- und L2-Cache

3.2.1 Architekturen

Mit steigender Anzahl Cores wird die Verbindung zwischen den Cores immer kritischer.

- Kommunikationsmodell
  - Message Passing: experimentell
  - Shared Address
- Physische Verbindung
  - Network
  - Bus: schlechte Skalierbarkeit

3.3 Speicherorganisation

- UMA: unified memory architecture
  - mehrere CPUs an ein gemeinsames Memory angeschlossen
  - alle CPUs haben Direktzugriff
  - jede CPU hat einen lokalen Adressraum (schlechte Skalierbarkeit)
- NUMA: non-unified memory architecture (shared memory architecture)
  - zwei oder mehr SMPs physisch verbunden
  - Shared Memory: ein globaler Adressraum
  - SMP kann direkt auf das Memory anderer SMPs zugreifen
  - nicht alle CPUs haben die gleiche Zugriffszeit auf das Memory
    - variiert je nach Speicheradresse
    - je weiter weg das Memory, desto langsamer der Zugriff
  - CC-NUMA: mit Cache-Kohärenz (cache coherence)

3.3.1 LLC: last level cache

TODO p. 17

3.4 Simultanes Multi-Threading

Hauptprobleme bei Multi-Core-Architekturen:

1. Memory-Stalls
2. Cache-Kohärenz

3.4.1 Memory Stalls

Zeitanteil, der für das Warten auf die Daten ins Memory benötigt wird. Beispiel:

- Transaktionsdatenbank: ca. 75%

- Web-Server: ca. 50%
- Entscheidungsunterstützte (mittels Machine-Learning, “trainierte Algorithmen”) Datenbank: ca. 10-50%

Analogie-Beispiel: betrüge ein einziger CPU-Zyklus eine Sekunde statt 0.3 Nanosekunden, so dauerte das Laden:

- aus dem L1-Cache: 3 Sekunden
- aus dem L3-Cache: 45 Sekunden
- aus dem DRAM: 6 Minuten
- von einer HD: Monate oder Jahre

Lösungsansatz: Threads (kleinste Einheit, die ein OS auf Cores verteilen kann). Während ein Thread darauf wartet, dass Daten ins Memory geladen werden, kann derweil ein anderer Thread arbeiten, statt zu warten (Multithreading oder Pipelining); Reduktion um ca. 50% der Idle-Time.

3.4.2 Superskalare Architektur

- Es werden mehrere Instruktionen pro Taktzyklus abgearbeitet.
- Dazu müssen mehrere Ausführungseinheiten (z.B. ALU, Bit Shifter, Multiplier) pro Chip vorhanden sein.
- Im besten Fall sind zur gleichen Zeit alle Ausführungseinheiten beschäftigt.
- Das ist nach dem Aufbau/Füllen bzw. vor dem Abbau/Leeren der Pipeline möglich.
- Nachteil von Pipelining: bedingte Codeblöcke werden immer in die Pipeline geladen, auch wenn sie teilweise gar nicht gebraucht werden, und müssen teilweise geflusht werden.
- Hyper-Threading: Synomym für Pipelining

3.4.3 Amdahls Gesetz

- Ab einer gewissen Anzahl Cores nimmt die zusätzliche Leistung pro weiterem Core ab
  - nicht parallelisierter Anteil `a`: ca. 20%
  - parallelisierter Anteil `1-a`: ca. 80%

  a = nicht parallelisierter Anteil n = Anzahl Cores T = Ausführungszeit
  T=a+(a-a)/n

T konvergiert für grosse n gegen 0.2 für einen nicht parallelisierten Anteil von 20%.

3.4.4 Cache-Kohärenz

- gemeinsames Datenobjekt in beiden Caches und im Memory

1. write through cache
2. write back cache

TODO: p.38-39

Lösungsansätze:

1. Erweiterung der beiden genannten Algorithmen
   - Snoop-Protokoll
     - Zugriffe erfolgen über das gleiche Medium (Bus-System)
     - Cache-Controller beobachten und erkennen Datenänderungen
     - skaliert schlecht
     - TODO: p.44
   - Directory-Protokoll
     - zentrale Liste mit Status aller Cache-Blöcke (welche CPU hat was?)
     - skaliert besser
     - TODO: p.45-47
2. Verwendung eines gemeinsamen Caches
3. Unterteilung der Daten (auf Softwareebene)

3.5 Verbindungsnetzwerke

- Bei 1000 Cores bräuchte es 10^6 Verbindungen zwischen den Cores.
- Jede Verbindung müsste 64 Bit breit sein.
- Es bräuchte `6.4*10^7` Leitungen.
- Bei einer 100-lagigen Platine ergäbe das eine Platinenbreite von 6 km.
- Es ist nicht praktikabel, alle Cores miteinander zu verbinden!

TODO: p.50

Lösung: Verbindungsnetzwerke. Die Cores werden über geschaltete Verbindungen (Switches) miteinander verbunden.

TODO: p.52

3.5.1 Bewertungskriterien für Topologien

TODO: p.53-62

“Koppelnetze sind gekoppelte Netze” – “Nein” – “Doch” – “Ohh”

3.6 Skalierbare Applikationen

Wie erstellt man skalierbare Applikationen?

- skalierbare Algorithmen
- feingranulare Locking-Mechanismen verwenden
- Worker-Thread-Pools verwenden

- Skalierungs-Tests durchführen
- Beobachtung der Applikation
- Stellen mit Wartezeiten identifizieren
- Synchronisations-Locks identifizieren
- nicht-skalierbare Algorithmen identifizieren

TODO: p.65 ff

4 Glossar

- ITIL: IT Infrastructure Library, Standard für IT-Belange v.a. für Grossunternehmen, für KMU übertrieben
- PUE: Power Usage Effectiveness, Massstab für die Effizienz eines Rechenzentrums
- DNS: Domain Name System
- IPAM: IP Address Management
- DHCP: Dynamic Host Configuration Protocoll
- NAP: Network Access Protection
- NAC: Network Access Control