

IT-Systemengineering & -Operations

Das Data Center

Kapitel 2

Vers. 1.0

Markus Waldmann



Inhalt

- Building Blocks eines Datacenters
- Kritische Punkte
- Verfügbarkeitsverbesserungen
- Tier Levels im Datacenter
- Information Lifecycle Management
- Tiered Storage
- Übung „allocation efficiency“

Lernziele

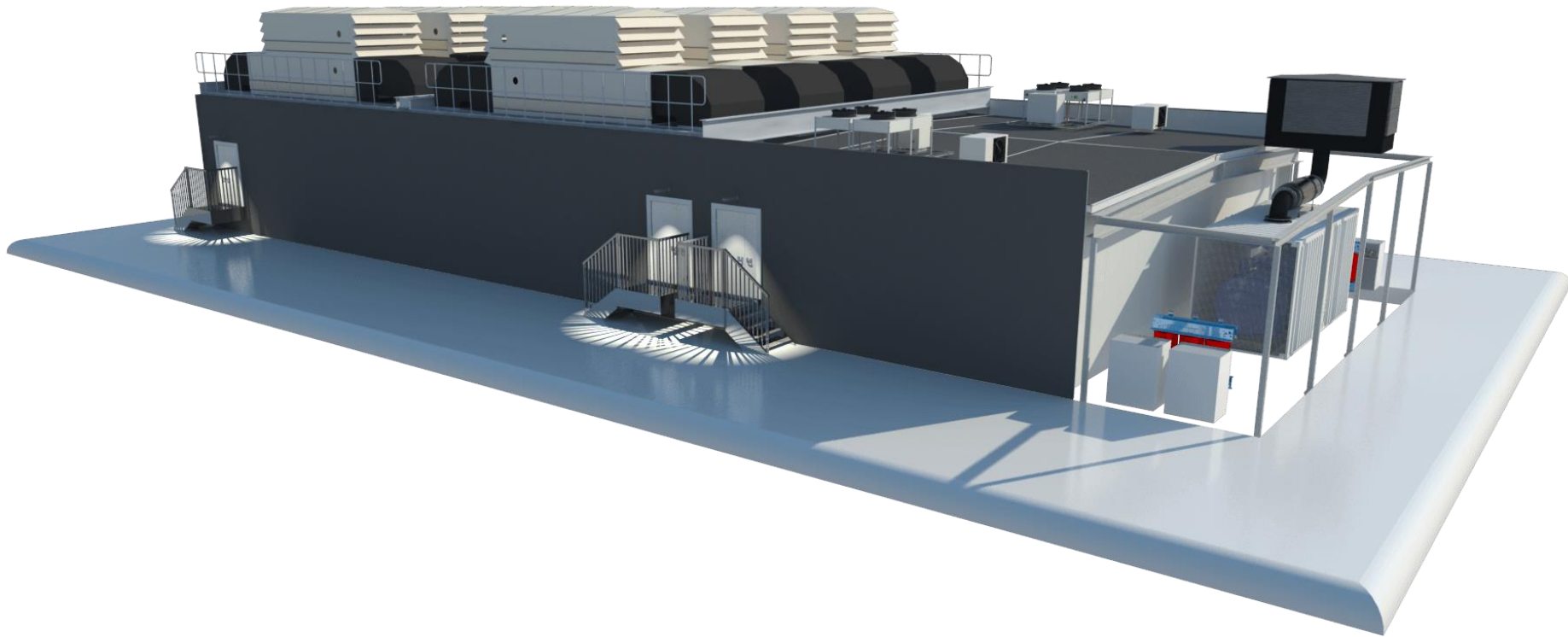
- Sie kennen die Building Blocks eines Datacenters.
- Sie sind fähig die kritischen Punkte eines Datacenters zu adressieren und Massnahmen vorzuschlagen.
- Sie kennen die Tier Level von Datacentern.
- Sie wissen wie man Daten klassifiziert und sie in SLAs einbindet.
- Sie sind in der Lage Data-Tiers und deren Aufgaben zu erklären.

Building Blocks eines Datacenter



Ein modernes Data Center

Welche Blöcke eines modernen Data Centers kennen oder erkennen Sie?



- <http://www.ictroom.com/en/modular>

Building Blocks

- Gebäude
- Zutrittskontrolle
- Brandschutz / Hochwasserschutz
- Klimatisierung
- Serverracks mit Rechnern
- Stromversorgung
- Datenleitungen / Netzwerk
- Administration und Überwachung

Klimatisierung

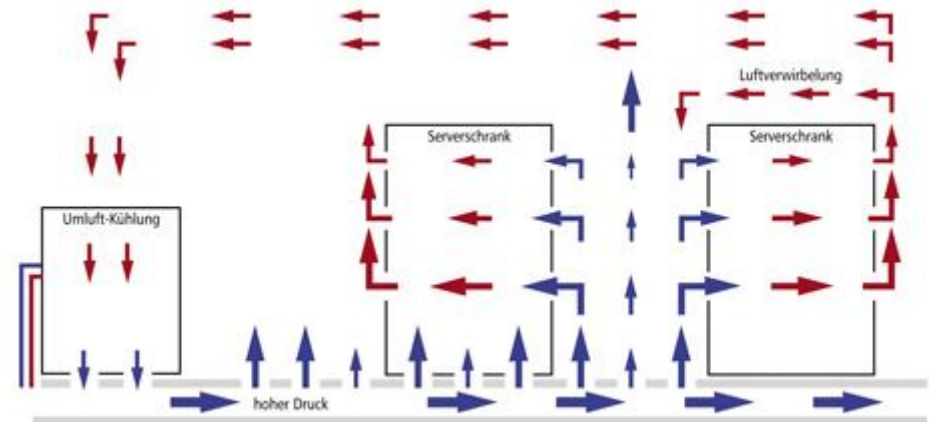
- Abfuhr von Wärme
- Konstante Luftparameter (Temperatur und Feuchtigkeit)
 - ca. 26 °C als Optimum empfohlen (siehe Zusatzdokument ILIAS)
- Reine Luft (Staub, Pollen, ...)
- Probleme
 - Kondenswasser
 - Filterkontrolle
 - Braucht zusätzliche Energie
 - Luftverteilung
 - Überwachung



Kühlluftverteilung

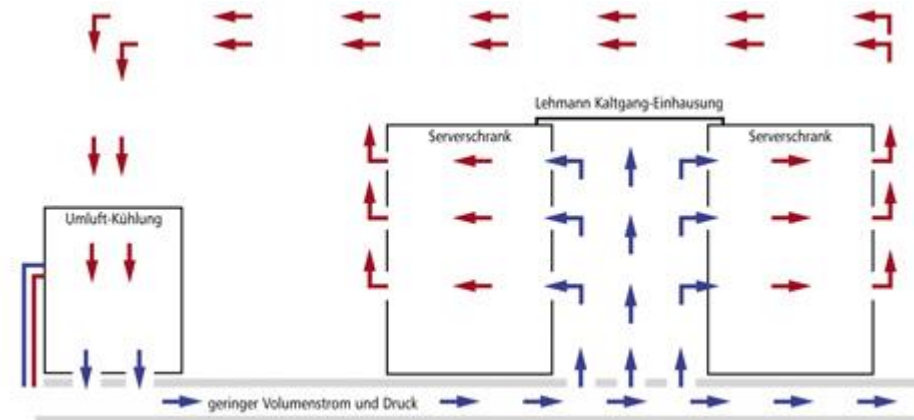
Free Flow Systeme

- + Einfacher Aufbau
- + Keine zus. Hindernisse
- Kalt-/Warmluftvermischung
- Unterschiedliche Temperaturen

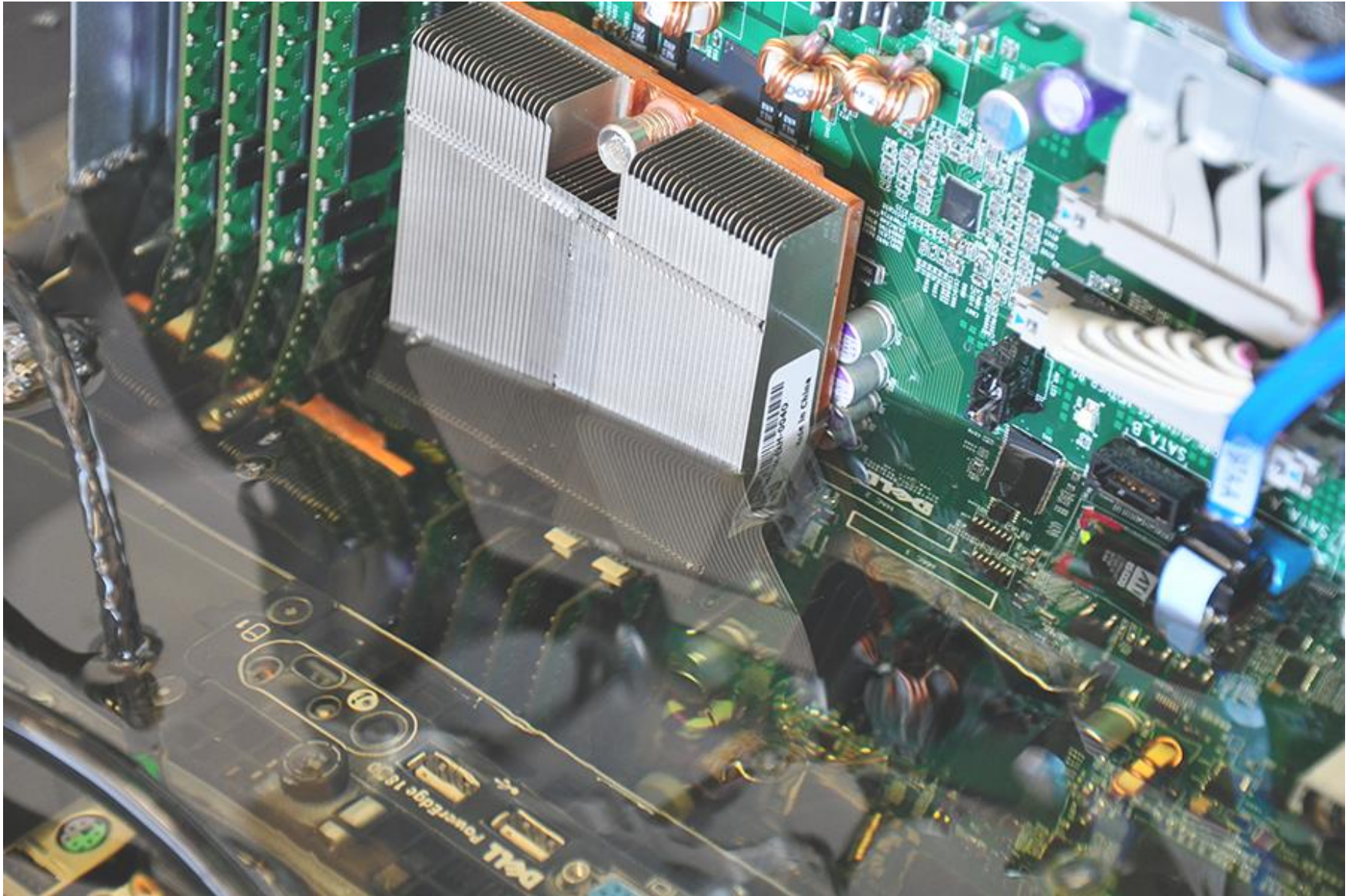


Kalt- oder Warmgang-Einhausung

- + Energieoptimiert durch Trennung von Kalt- und Warmluft
- + Geringere Energiekosten
- + Kleine vertikale Temperaturdifferenzen
- Zusätzliche Kosten im Aufbau
- Erhöhter Abschottungsbedarf



Immersion Cooling



EDV-Einbau

- **Serverracks**
 - Höhe 21-49U / Tiefe 0,8-1,2 m / Breite 0,6-1m
 - Integrierte Kühlung möglich
 - Zuleitungen oben / unten / seitlich
 - 19" als Standard (48.26 cm)
 - 1 HE = 1 U = 1.75" = 44.45 mm
- **Netzwerk (Kupfer, Glas)**
- **Stromverteilung (Trasse oben/unten)**
 - Kabelführung über Hohlboden oder Deckentrasse
- **Klimageräte, USV, Batterieschränke**
- **Ev. Kühlleitungen und Überwachungsgeräte**



Kritische Punkte



Data Center: Gefährdungen und Abhilfe

- Einbruch/Diebstahl/Vandalismus/Sturmschäden/Trümmer
 - bauliche Massnahmen
- Fremdzugriff
 - Zutrittskontrolle, Abhörsicherheit, Firewall, ...
- Feuer / Rauch
 - Brandfrüherkennung, Löschung, Abschottung
 - Problem: Brandgasverteilung durch Klimaanlage
- Netzausfälle und Netzstörungen
 - Netzfilter, USV mit Batterien, Generatoren
- elektromagnetische Störfelder
 - Abschirmung
- Staub / Schmutz / Wasser
 - Filteranlagen, Schleusen, Standortwahl, Abschottungen, Pumpenanlagen, ...

Überwachungsgebiete

- Gebäude
 - Türen offen – zu, Türen verschlossen, Kameras, Bewegungsmelder, Zutritte
- Räume
 - Temperatur, Luftfeuchte, Bewegung, Rauch, Brand, Leckage und Wasser
- Energieversorgung
 - Netzausfall, Strom, Spannung, Leistung (kW), Leistungsfaktor ($\cos \phi$)
- Einzelne Geräte
 - Niederspannungsverteilungen, Schalterstellungen 0/1, Stromverbrauch einzelner Bereiche, Sicherungsausfall, Kurzschluss, Überlast etc.

Überwachungsgebiete (2)

- Generator
 - auf Kraftstoffmangel, Funktionsbereitschaft, Temperatur, Überlast
- Klimageräte
 - Temperaturen, Luftfeuchtigkeit, Übertemperatur, Filterwiderstand, Störungen
- USV-Anlagen
 - Normalbetrieb, Batteriebetrieb, Bypass- Betrieb, Ladezustand, Batterietemperatur
- Brandmelde- und Löschanlage
 - Zustandsanzeigen wie z.B. Löschanlage ausgelöst, Übertragungseinrichtung ausgeschaltet, Störung, Service

RZ Effizienz, z.b. PUE Faktor

- PUE-Wert = Power Usage Effectiveness)

- $$\text{PUE} = \frac{\text{Gesamte vom RZ verbrauchte Energie}}{\text{Verbrauch der IT-Geräte}}$$

- Massstab für die Effizienz eines RZs
- Im Umfeld von «Green IT» immer wieder benutzt
- Größenordnungen:
 - 1.0: optimal, kein Energieaufwand für Kühlung nötig
 - 1.2: guter Wert, Bereich der normalen RZs
 - >1.4: Optimierungsbedarf?

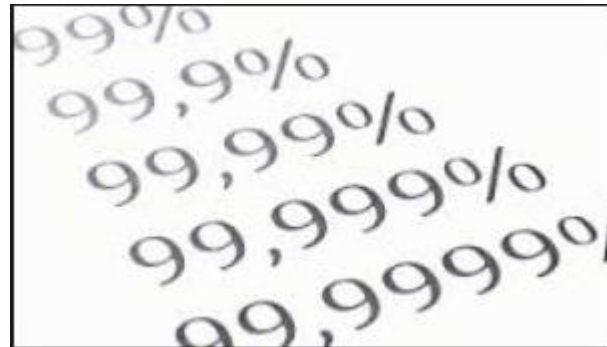
Repetitionsfragen

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

Baustein	Funktionen	Gefährdet durch?	Abhilfe gegen Gefährdungen
1)	a) b)	i) ii) ...	- - ...
2)	a)

- Versuchen Sie den Kostenanteil pro Baustein am gesamten RZ abzuschätzen
- Was ist der PUE Faktor und was sind die erreichbaren und effektiv erreichten Werte?

Verfügbarkeitsverbesserungen



Kosten für Downtime

Branche	Kosten pro Stunde (\$) (Richtwerte)
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

Das Rätsel der Neunen

Servicelevel: 7 * 24h

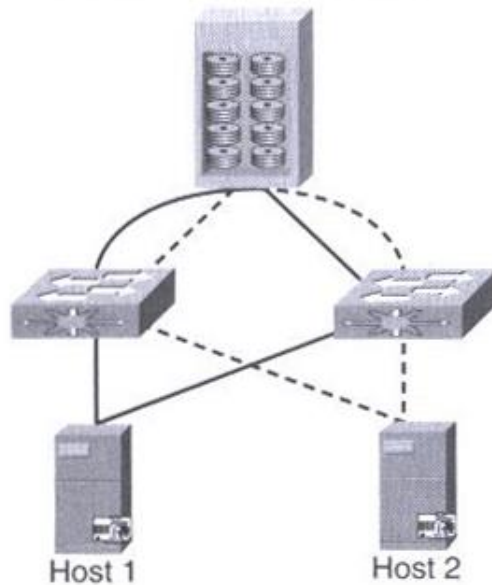
Uptime (%)	Downtime / Jahr
90%	876 Stunden (36,5 Tage)
95%	438 Stunden (18,25 Tage)
99%	87,6 Stunden (3,65 Tage)
99,9%	8,76 Stunden
99,99%	52,56 Minuten
99,999%	5,256 Minuten
99,9999%	31,536 Sekunden

9er-Vergleich bei verschiedenen SLA

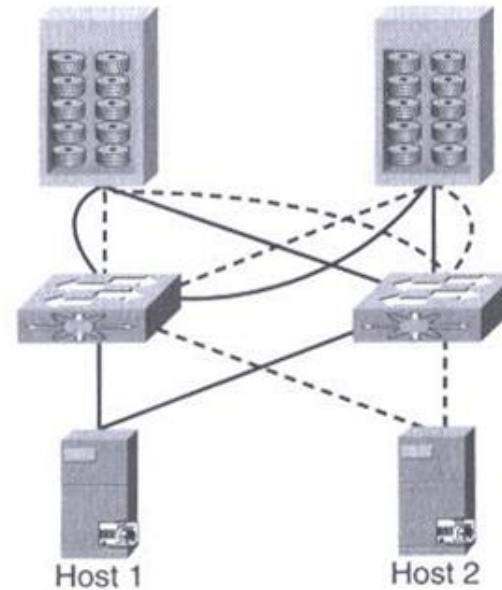
7 * 24	365	24	60	60
	365	8760	525600	31536000
	Tage	Stunden	Minuten	Sekunden
90.0000%	36.50	876.00	52560.00	3153600.00
95.0000%	18.25	438.00	26280.00	1576800.00
99.0000%	3.65	87.60	5256.00	315360.00
99.9000%	0.37	8.76	525.60	31536.00
99.9900%	0.04	0.88	52.56	3153.60
99.9990%	0.00	0.09	5.26	315.36
99.9999%	0.00	0.01	0.53	31.54
5 * 9	365	9	60	60
	261	2349	140940	8456400
	Tage	Stunden	Minuten	Sekunden
90.0000%	26.10	234.90	14094.00	845640.00
95.0000%	13.05	117.45	7047.00	422820.00
99.0000%	2.61	23.49	1409.40	84564.00
99.9000%	0.26	2.35	140.94	8456.40
99.9900%	0.03	0.23	14.09	845.64
99.9990%	0.00	0.02	1.41	84.56
99.9999%	0.00	0.00	0.14	8.46

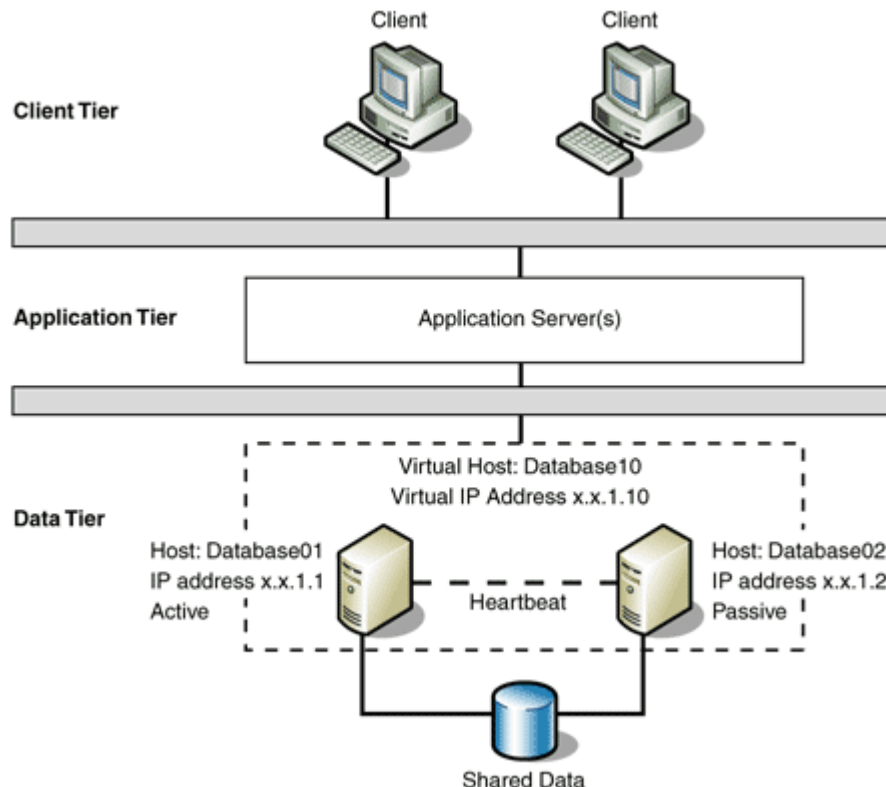
Beispiel: erhöhte Verfügbarkeit im SAN

SAN Configuration Example 1



SAN Configuration Example 2

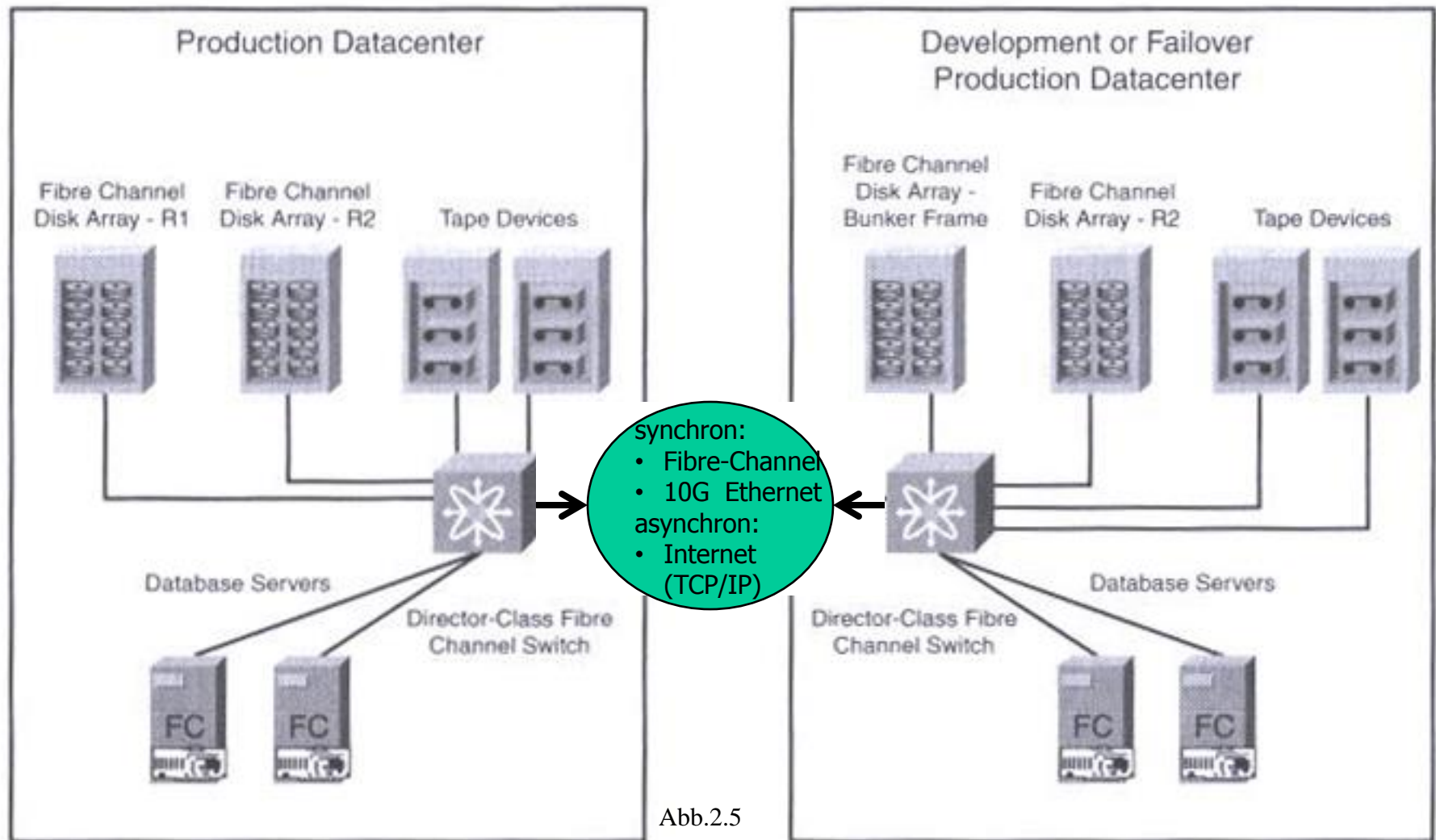




Der Zugriff auf die Daten erfolgt über eine virtuelle IP Adresse, welche zusammen mit einem Service (Datenbank, Mailserver, Fileshares, Druckserver, ...) im Fehlerfall eines physischen Hosts automatisch auf den zweiten Host verschoben wird.

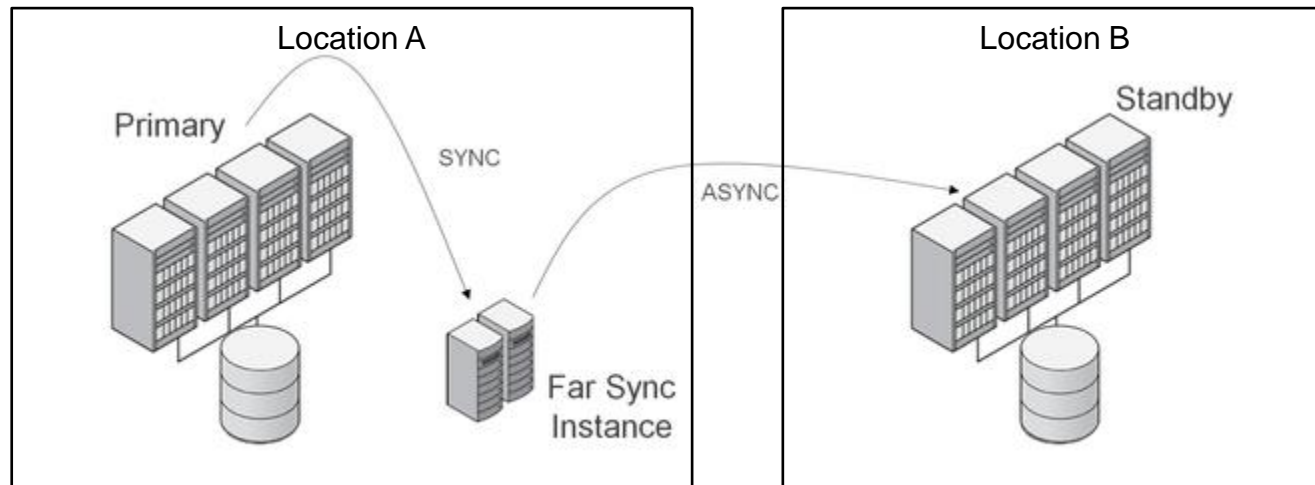
Es sind auch Lösungen mit mehr als 2 Servern und automatischem Lastausgleich möglich.

Failover Datacenter



Asynchrone und Synchron Replikation

In der Praxis wird oft eine Kombination aus Asynchronität und Synchronität angewendet um Geschwindigkeit und Zuverlässigkeit zu erhöhen.



Damit die primäre DB nicht durch die Laufzeiten bei der Spiegelung belastet wird, werden die Daten synchron auf eine lokale Instanz gespiegelt.

Die zeitintensivere Datenübertragung auf das entfernte Datacenter wird danach asynchron vorgenommen.

Das off-site Rechenzentrum befindet sich im Standby Betrieb.

Availability Environment Classification

Harvard Research Group (HRG) teilt Hochverfügbarkeit in ihrer *Availability Environment Classification* (AEC) in sechs Klassen ein.

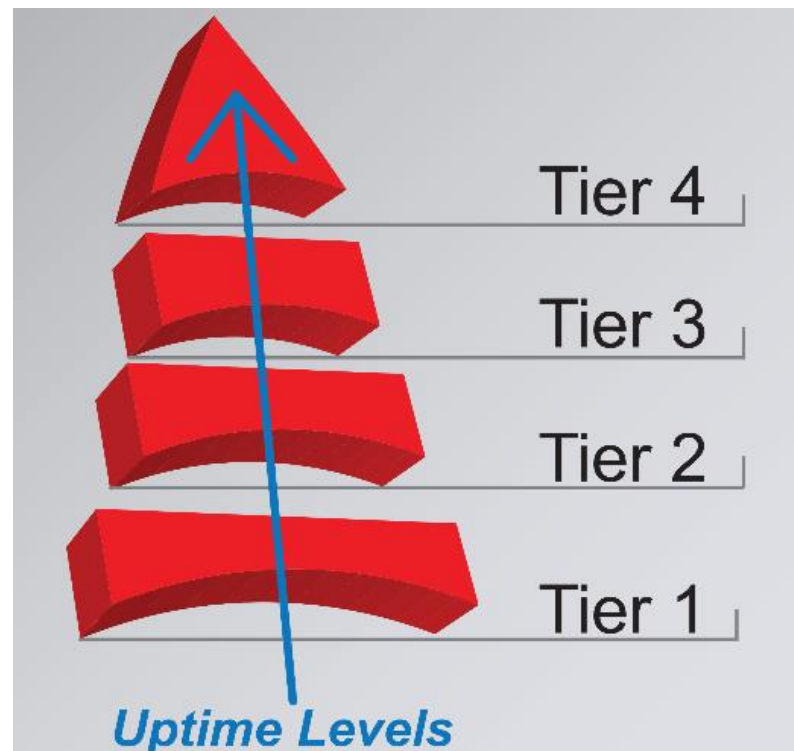
[\[https://de.wikipedia.org/wiki/Hochverfügbarkeit\]](https://de.wikipedia.org/wiki/Hochverfügbarkeit)

HRG-Klasse	Bezeichnung	Erklärung
AEC-0	<i>Conventional</i>	Funktion kann unterbrochen werden, Datenintegrität ist nicht essentiell.
AEC-1	<i>Highly Reliable</i>	Funktion kann unterbrochen werden, Datenintegrität muss jedoch gewährleistet sein.
AEC-2	<i>High Availability</i>	Funktion darf nur innerhalb festgelegter Zeiten oder zur Hauptbetriebszeit minimal unterbrochen werden.
AEC-3	<i>Fault Resilient</i>	Funktion muss innerhalb festgelegter Zeiten oder während der Hauptbetriebszeit ununterbrochen aufrechterhalten werden.
AEC-4	<i>Fault Tolerant</i>	Funktion muss ununterbrochen aufrechterhalten werden, 24/7-Betrieb (24 Stunden, 7 Tage die Woche) muss gewährleistet sein.
AEC-5	<i>Disaster Tolerant</i>	Funktion muss unter allen Umständen verfügbar sein.

Repetitionsfragen

- Welche Verfügbarkeit muss im SLA festgehalten werden, wenn ich 1h Ausfallzeit während den Bürozeiten nicht überschreiten will?
- Was versteht man unter Failover-Cluster-Services?
- Wenn z.B. das SAN gespiegelt werden soll, wie erhöhen sich die Kosten?
50% / 100% / mehr als 100% und warum?

Tier Levels im Datacenter



<http://www.thedatacave.com/wp-content/uploads/2014/02/data-center-tier-graphic.jpg>

Data Center Tiers

Tier Klasse	Anforderungen
Tier I	Redundanz: N, Jährliche Ausfallzeit: 28,8 h, 99,67 % Verfügbarkeit
Tier II	Redundanz: N+1, Jährliche Ausfallzeit: 22 h 99,75 % Verfügbarkeit
Tier III	Redundanz: N+1, Jährliche Ausfallzeit: 1,6 h 99,98 % Verfügbarkeit
Tier IV	Redundanz: 2 x (N+1), Jährliche Ausfallzeit: 0,8 h 99,99 % Verfügbarkeit

Quelle: Uptime Institute

DC Tiers in details

Features	1	2	2+	3	3+	4
Dual Electric Utility Feed		✓	✓	✓	✓	✓
Dual Power Path Above 600V		✓	✓	✓	✓	✓
UPS System		N	N	N+1	N+1	2N
Standby Generators Provided For Critical Loads				N	N+1	N+1
Fuel Systems				N+1	N+1	N+1
Dual Power Path to PDU					✓	✓
Dual Feed CAC Units				✓	✓	✓
Dual Power Path For Dual Power Cord Loads				✓	✓	✓
Chiller Plant - N+1 Redundancy	✓	✓	✓	✓	✓	N+2
Redundant Tower Water Make up Supply	✓	✓	✓	✓	✓	✓
Raised Floor - CAC Units - N+25%	✓	✓	✓	✓	✓	✓
Critical Load - Cooling Systems - N+1	✓	✓	✓	✓	✓	✓
Automation & Monitoring	✓	✓	✓	✓	✓	✓
Single Points of Failure Eliminated						✓
7 x 24 Facility Operator Coverage				✓	✓	✓
Mechanical Systems - Concurrent maintenance	✓	✓	✓	✓	✓	✓
Electrical Systems - Concurrent Maintenance			✓		✓	✓

Quelle: Uptime Institute

Rechercheaufgabe

- Suchen sie RZ-Services Anbieter mit Level 3, 3.5 und Level 4 Rechenzentren.
- Versuchen sie die Kosten für den Service zu bestimmen (pro Rack, pro Server, ...)
- Die Resultate können sie im ILIAS in den Ordner «Datenaustausch» legen und so ihren Mitstudierenden zur Verfügung stellen.

Information Lifecycle Management



<https://www.smartfile.com/blog/wp-content/uploads/2016/02/lifecycle-management.png>

Information Life Cycle Management

- ILM beschreibt eine Speicherstrategie. Ziel dieser Strategie ist die Speicherung von Informationen entsprechend ihrem **Wert** auf dem jeweils **günstigsten Speichermedium** einschließlich der Regeln und Prozesse, wie Information auf die geeigneten Speichermedien verschoben wird.
- Die Steuerungsmechanismen der Verwaltung und der Speicherung orientieren sich an **Wichtigkeit, Wertigkeit und Kosten** der elektronischen Information.
- Hierfür wird eine **Klassifizierung der Daten**, Quellen und Speichersysteme vorgenommen, die innerhalb einer Speicherhierarchie die automatisierte Bereitstellung erlauben.

(Auszug aus Wikipedia)

Information Lifecycle Management (ILM)



Abb. 3.6

ILM - Management

Das ILM wird getrieben durch die Anforderungen von:

- Storage Management
- Document Lifecycle Management
- Content Life Cycle Management
- Records Management

Regeln

Regeln legen fest wie und wo die Daten gespeichert werden:

- Änderungshäufigkeit
- Zugriffsgeschwindigkeit
- Zugriffshäufigkeit
- Kosten
- Ökonomischer Wert
- Relevante gesetzliche Bestimmungen

Danach werden die Daten auf die angemessenen Speichermedien verschoben und am Ende des Zyklus auch gelöscht.

Active vs. inactive Data

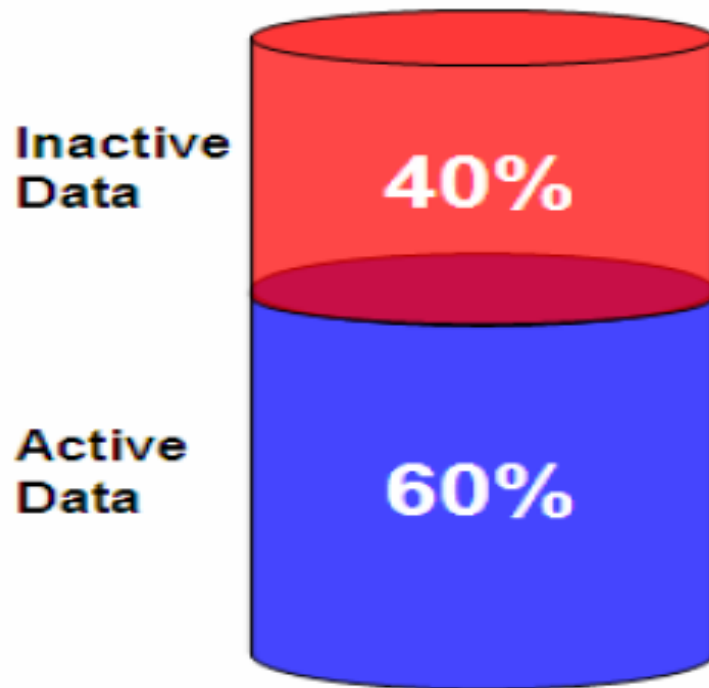
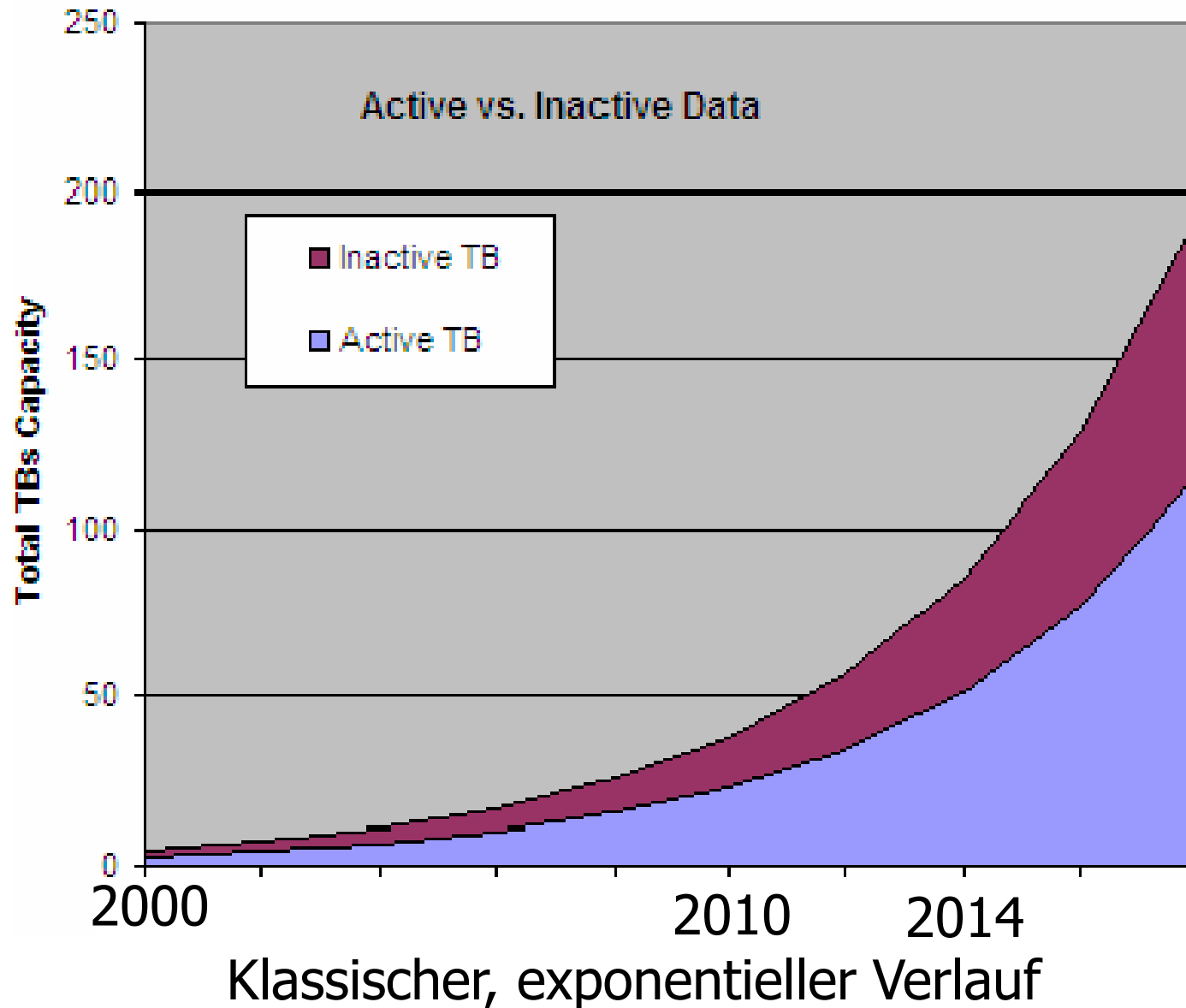


Abb. 3.4

Inaktive Daten

- Konsumieren teuren Speicherplatz
- Werden gepflegt, gesichert, repliziert, etc.
- Unterliegen trotzdem Rechtlichen- und Datenhaltungs-Ansprüchen
- Müssen im Katastrophenfall wiederhergestellt werden

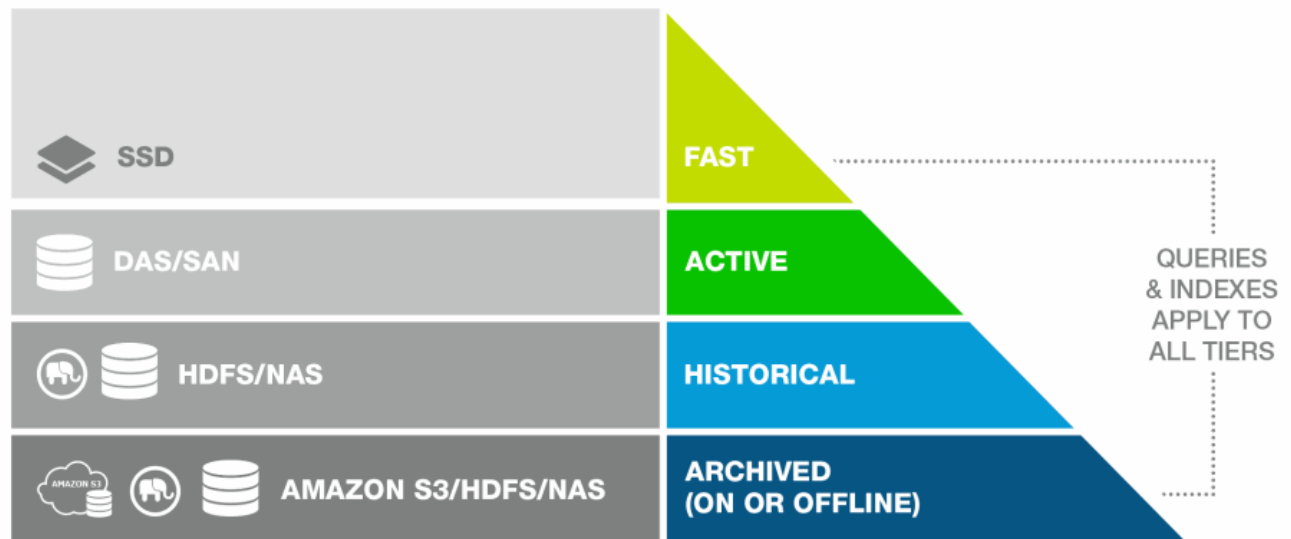
Active vs. Inactive Data (2)



Repetitionsfragen

- Was versteht man unter Records Management?
- Welche (gesetzlichen) Vorschriften für die Datenaufbewahrungszeit kennen sie?
- Wer soll das Records-Management (RM) anordnen und durchsetzen?
 - Administratoren
 - Abteilungsleiter
 - Rechtsabteilung
 - Geschäftsleitung
 - Polizei?
- Kennen sie aus der eigenen Umgebung Beispiele?

Tiered Storage



Heisst soviel wie:

„Verschiedene Datenklassen gespeichert auf verschiedenen Speicherklassen“

Daten Charakteristik:

1. Mission Critical Data
2. Business Critical Data
3. Nearline or Historical Data
4. Offline Data

Business Modell

- Nicht alle Daten müssen auf der gleichen Klasse von Speichern liegen
- Eine Klassifizierung der Daten und Speicherung auf dem entsprechenden Medium senkt die Kosten!
- Records Management ([ISO 15489](#)) steuert Datenvolumen:
 - Aufbewahrungszeiten (gesetzlich / intern)
 - Ablageordnung, Namensgebung, Versionen
 - Automatische Routinen zum Löschen

Vergleich der Speichertypen





Attribut	SSD	FC/SAS	SATA	MAID *	Tape
Zugriffszeit	r:0,2 ms w:0,4 ms	~4 ms	~12 ms	ms-Sekunden	Sekunden bis Stunden, abhängig vom Aufbewahrungsort
Speicherdichte	moderat	moderat	moderat-hoch	hoch	hoch
Zuverlässigkeit	hoch bis sehr hoch	hoch	moderat	moderat-hoch, grosse Serviceabstände	Niedrig-moderat, grosse Serviceabstände
Datenrate sequentiell r/w	300-500 MB/s	160 MB/s	100 MB/s	n.a	Durch Einsatz von D2T und Verbesserungen der Speicherdichte immer steigend.
Datenrate random r/w	20'-100' kIOPS	140 IOPS	90 IOPS	n.a	gering
Energie	hoch	hoch	moderat-hoch	niedrig-moderat	Sehr niedrig
Speicherkosten	sehr hoch	hoch	moderat	niedrig	Sehr niedrig, abhängig von Anzahl Tapes und Lagerraum

* MAID: Massive Array of Idle Disks (sehr viele Disk, powersafe oder offline, nur aktiv wenn benötigt), siehe Zusatzdoku

Festplatten / Einsatzbereiche

Requirement	Desktop	Enterprise
Operational Availability	8 hours/day - 5 days/week	24 hours/day - 7 days/week
Work Load	10 - 20%	100%
Cost Sensitivity	Very sensitive to low cost	Moderately sensitive, balanced with requirements for reliability, availability, and data integrity.
Performance	Low to Moderate	High
Reliability	Moderate: <ol style="list-style-type: none"> 1. Outage affects only one user <ol style="list-style-type: none"> a. Critical data is not usually stored locally b. Higher Tolerance for long error recovery timeout 2. Lower Mean Time Between Failure Acceptable	High: <ol style="list-style-type: none"> 1. Outage affects multiple users 2. Higher Mean Time Between Failure 3. Intolerance for long Error Timeout
Data Integrity	Moderately desirable (a bit corruption may result in system lockup or critical data loss on one system)	Highly Desirable (a bit corruption may result in catastrophic critical data loss to multiple clients)

Klassifizierung der Speicher Tiers

	Tier 1 	Tier 2 	Compliance 	Tape vault 
Charakteristik	Hohe Geschwindigkeit Hohe Verfügbarkeit	Mittlere Geschwindigkeit, Hohe Speicherdichte	Einmalbeschreibbar, Sicher, unveränderlich	Geringe Geschwindigkeit, Entfernbare
Gebrauch	Produktionsdaten	Referenzdaten	Compliance Data (vom Gesetzgeber vorgeschrieben)	Offsite Disaster Recovery
% der gesamten Daten	20%	35%	30%	15%

Beispiel: Service Level Agreements

	Mission Critical	Business Critical	Business Important	Development
Requirement	SLA 1	SLA 2	SLA 3	SLA 4
Availability	99.99%	99.9%	99%	97%
Threshold based Automatisch Laden	bis 20% vom Filesystem alloziert während 1 Arbeitstag	bis 20% vom Filesystem alloziert während 2 Arbeitstagen	bis 10% vom Filesystem alloziert während 4 Arbeitstagen	bis 20% vom Filesystem alloziert während 1 Arbeitswoche
RTO	15 Minuten	1 Stunde	8 Stunden	24 Stunden
RPO	1 Stunde	12 Stunden	48 Stunden	96 Stunden
Restore Anfragen	100 Anfragen/Woche	100 Anfragen/Woche	50 Anfragen/Woche	50 Anfragen/Woche
Backup Erfolgsrate	97%	95%	90%	90%
Archivierungs - Policy	Kein Zugriff während 90 Tagen	Kein Zugriff während 30 Tagen	Kein Zugriff während 90 Tagen	Kein Zugriff während 180 Tagen
Zugriffszeit zum Archiv	Sekunden	Sekunden	Bis 4 Stunden	24-48 Stunden
Prognose	Monatlich	Vierteljährlich	Jährlich	Jährlich
Kosten	Hoch	Mittel	Niedrig	Gering

Inhaltsübersicht

- Building Blocks eines Datacenters.
- Kritische Punkte
- Verfügbarkeitsverbesserungen
- Tier Levels im Datacenter
- Information Lifecycle Management
- Tiered Storage
- Übung „allocation efficiency“



Kosten schlechter Qualität (Annahme: \$1'000'000 Einnahmen pro 1000 GB)

Speichertyp	Kosten pro MB	GB	Kosten Total	Alloziert	Gebrauch
Tier 1	\$0.05	1000	\$51'200	80%	75%
Tier 2	\$0.03	1000	\$30'720	80%	60%
Tier 3	\$0.01	1000	\$10'240	80%	90%
Speichertyp	allozierter Yield	Implementierter Yield	Realisierter Yield	Kosten durch schlechte Qualität	COPQ in % der Einnahmen
Tier 1	\$40'960	\$30'720	60%	\$20'480	2.05%
Tier 2	\$24'576	\$14'746	48%	\$15'974	1.60%
Tier 3	\$ 8'192	\$ 7'373	72%	\$ 2'867	0.29%

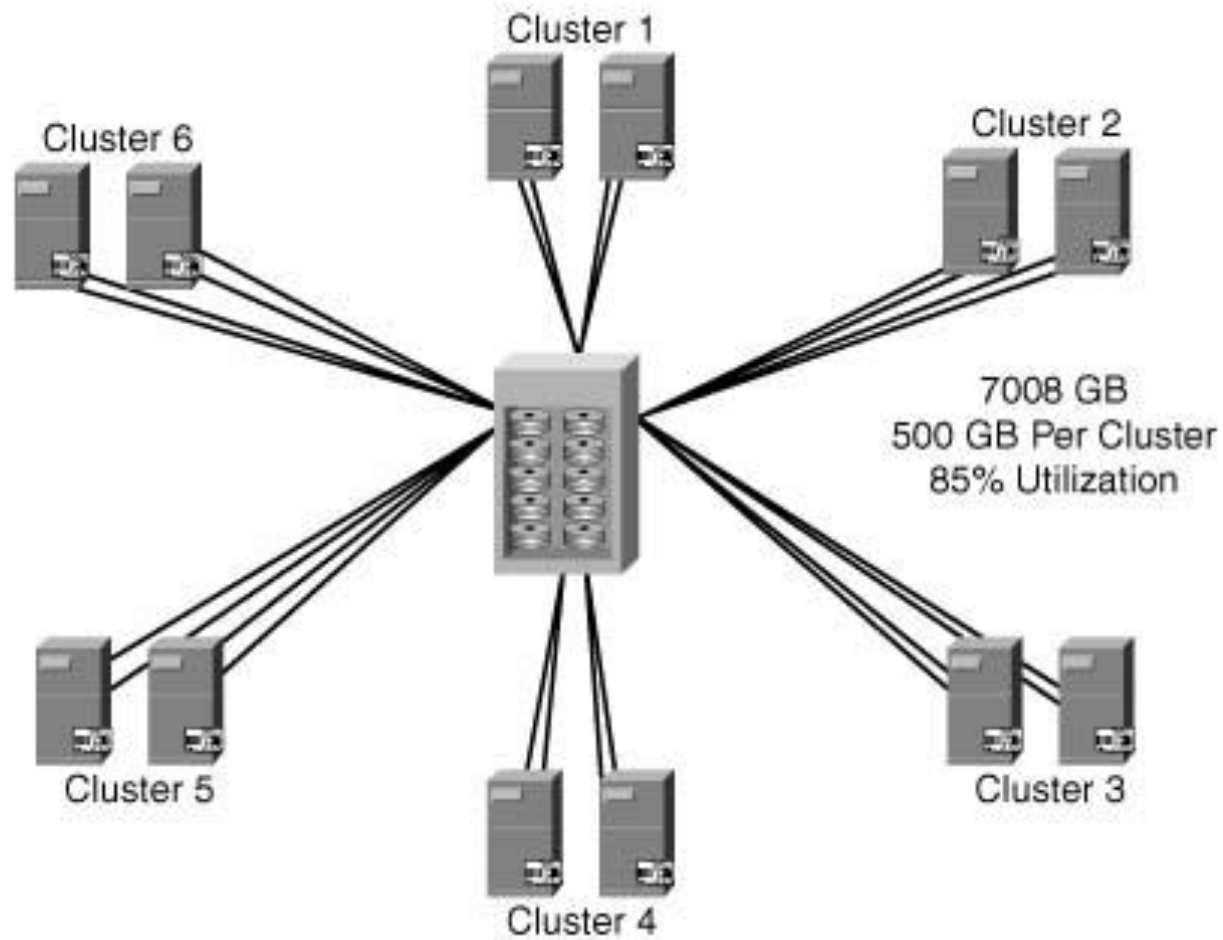
Lesen sie das Dokument „Utilization and Yield“

Übung:

Wie hoch sind die Kosten verursacht durch schlechte Qualität (COPQ)?

- 96 Platten a 73 GB
- 6 Adapterkarten (4-Port)
- Redundante Pfade zu den Servern (max. 12)
- Anschaffungspreis: \$0.1/MB
- Ähnliche Applikation pro Cluster mit 85% *allocation efficiency*

Übung



Links und Literatur

- Wikipedia RZ Portal

<https://de.wikipedia.org/wiki/Portal:Rechenzentrum>

- BITCOM RZ Guide

https://www.bitkom.org/Bitkom/Publikationen/Publikation_1349.html

- E-TEC Turnkey-Rechenzentren

http://www.e-tecpowerman.de/downloads/datacenter_rz_a4.pdf

