



**ATU** AUTO TEILE UNGER

**Ein hochverfügbarer  
Kommunikationsserver**  
Martin Klier  
Systemadministration / Datenbanken



## „Ein hochverfügbarer Kommunikationsserver mit RAC, ASM, redundant shared media und DataGuard.“

- Unternehmen “A.T.U”
- Servicepaket “A.T.U-Card”
- “Datendrehscheibe” Weiden und Problemstellung
- Projektschwerpunkte
- Technik
- Lessons learned

# Das Unternehmen A.T.U



Autofahrer-Fachmärkte mit  
Integrierter Meisterwerkstatt



Moderne  
Distributionszentren



A.T.U  
„Alles Außer Teuer“



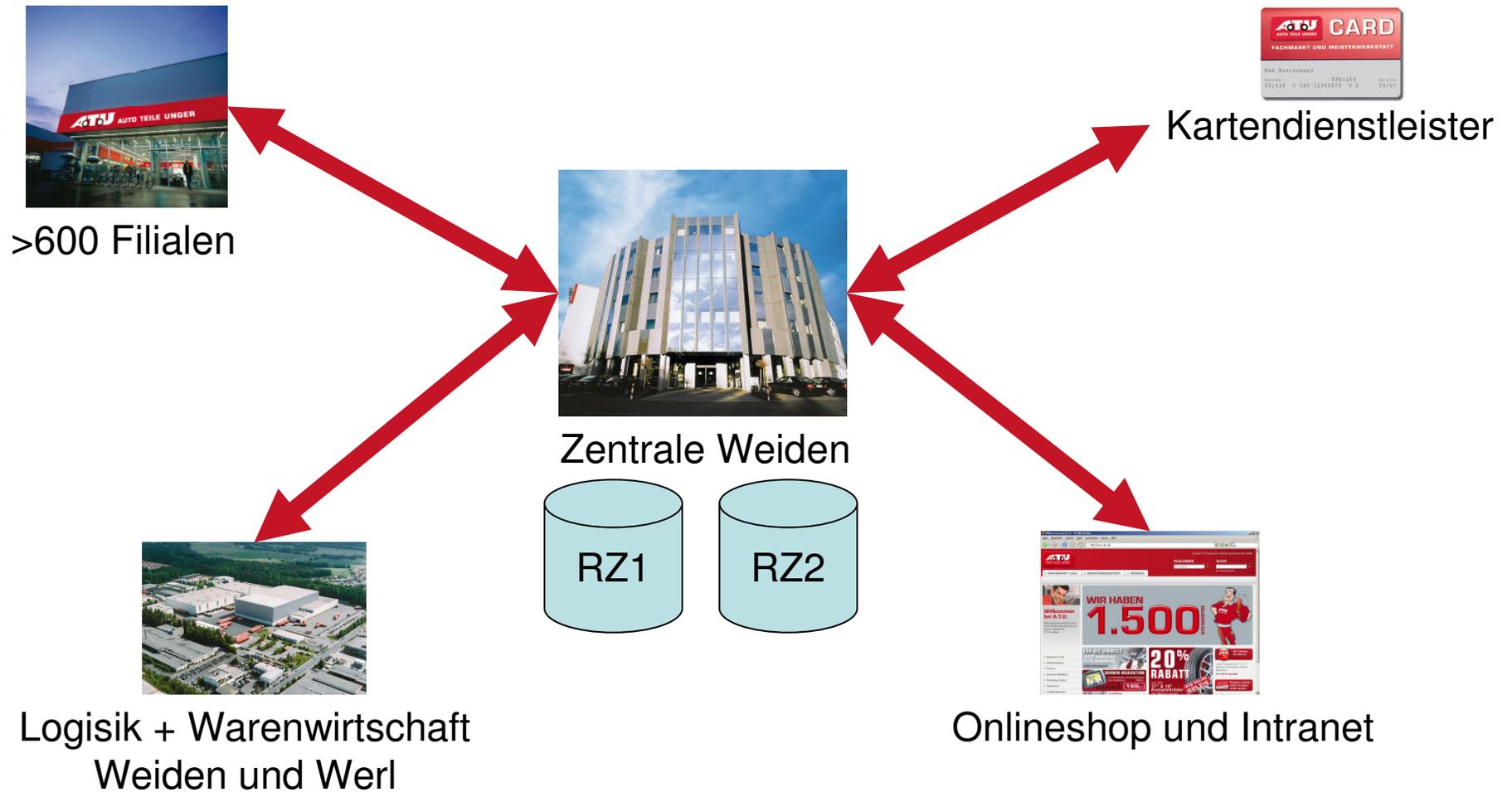
- >600 Filialen (jeweils Markt und Werkstatt)
  - Onlineshop
  - pro Jahr 40-50 Filialen Zuwachs
  - Ziel: 1000 Verkaufsniederlassungen
  - in derzeit 6, später 7 Ländern
- 
- 14.000 Mitarbeiter
  - ca. 1,4 Mrd. EUR Umsatz im Jahr
  - 12 Mio. Kunden

# Servicepaket "A.T.U-Card"

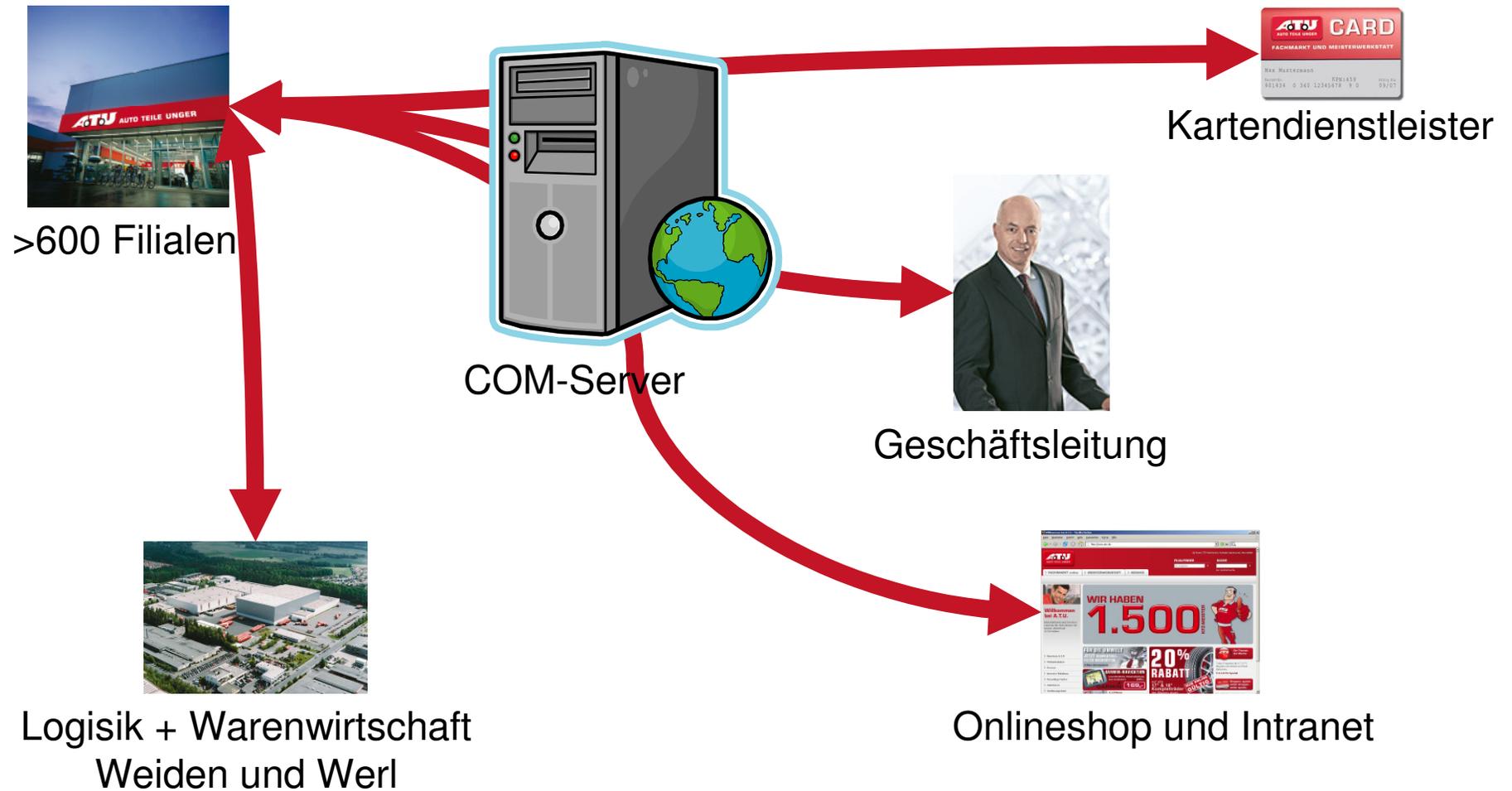


- Bonusfunktion
- Bezahlungsfunktion
- Webportal
- Gewerbe- und Großkundenservice

# Datendrehscheibe Zentrale Weiden



# Problemstellung





COM-Server

## Anforderungen:

- Leistungsfähigkeit
  - 10GB Datenzuwachs / Tag
  - 40-Tage-Historie
- Skalierbarkeit (1000 Filialen)
- Verfügbarkeit auch bei groben Störungen
  - Preislistenbereitstellung
  - Kartenabrechnung
  - Schnellfreischaltung
- Logische Absicherung der DFÜ
- DB-Gesamtgröße > 1TB
- DWH-Quellsystem
- Keine lastfreien Zeiten



COM-Server

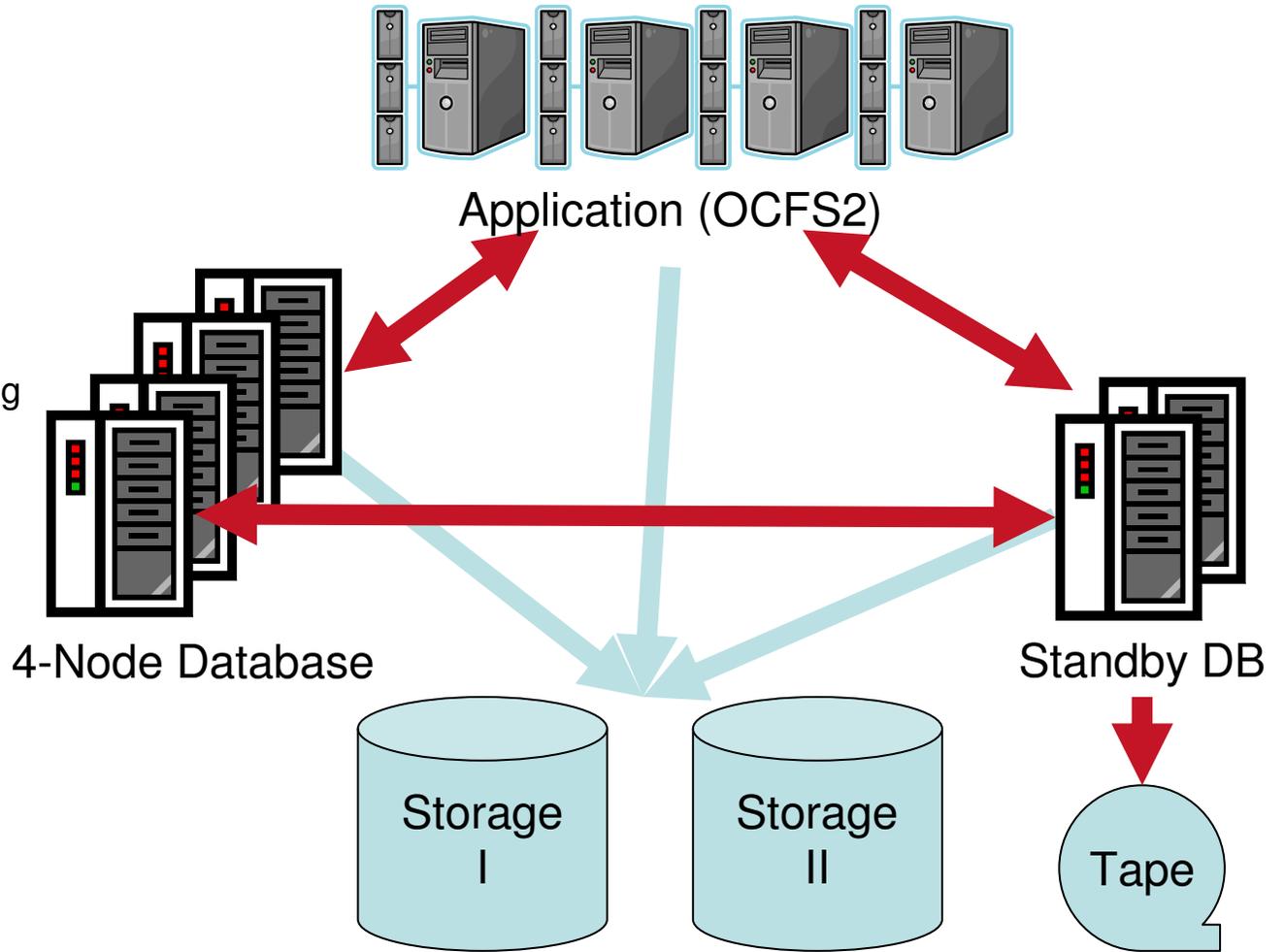
... ergeben sich aus den Anforderungen nach Verfügbarkeit, Leistung, und Skalierbarkeit:

- Server-Clustering
- Stabile Hardware
- Redundante Storage
- Robustheit von Jobs
- Schnelles Recovery
- Hochverfügbare, schnelle Netzwerktechnik
- Lastneutrale Datensicherung

# Grundaufbau "COM-Server"



- Server-Clustering
- Stabile Hardware
- Redundante Storage
- Robustheit von Jobs
- Schnelles Recovery
- Hochverfügbare, schnelle Netzwerktechnik
- Lastneutrale Datensicherung



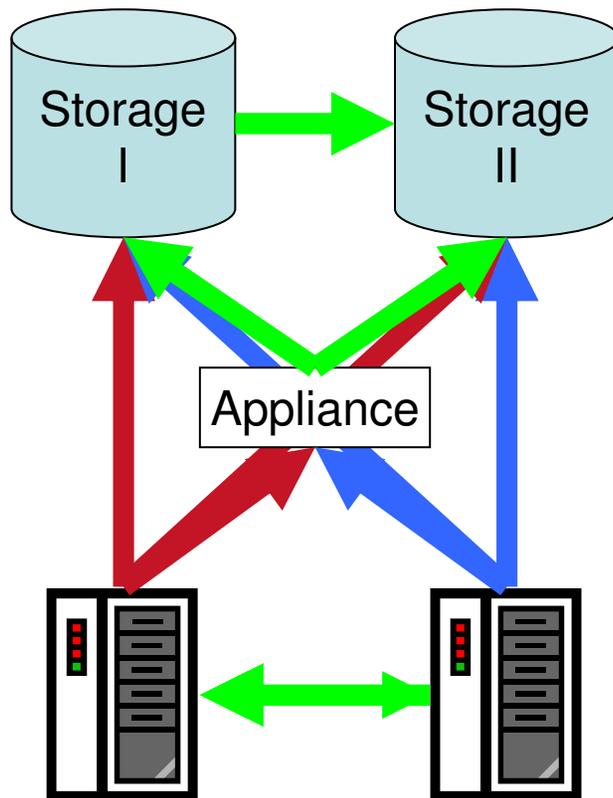
HP BL25p

# Redundant Shared Media & Clustering



„Die Sache mit der Storage“

... zur Absicherung von Transaktionen gegen RZ-Verlust



## Anforderung:

- Gebäudeübergreifende Spiegelung
- UND mehrere schreibende Peers

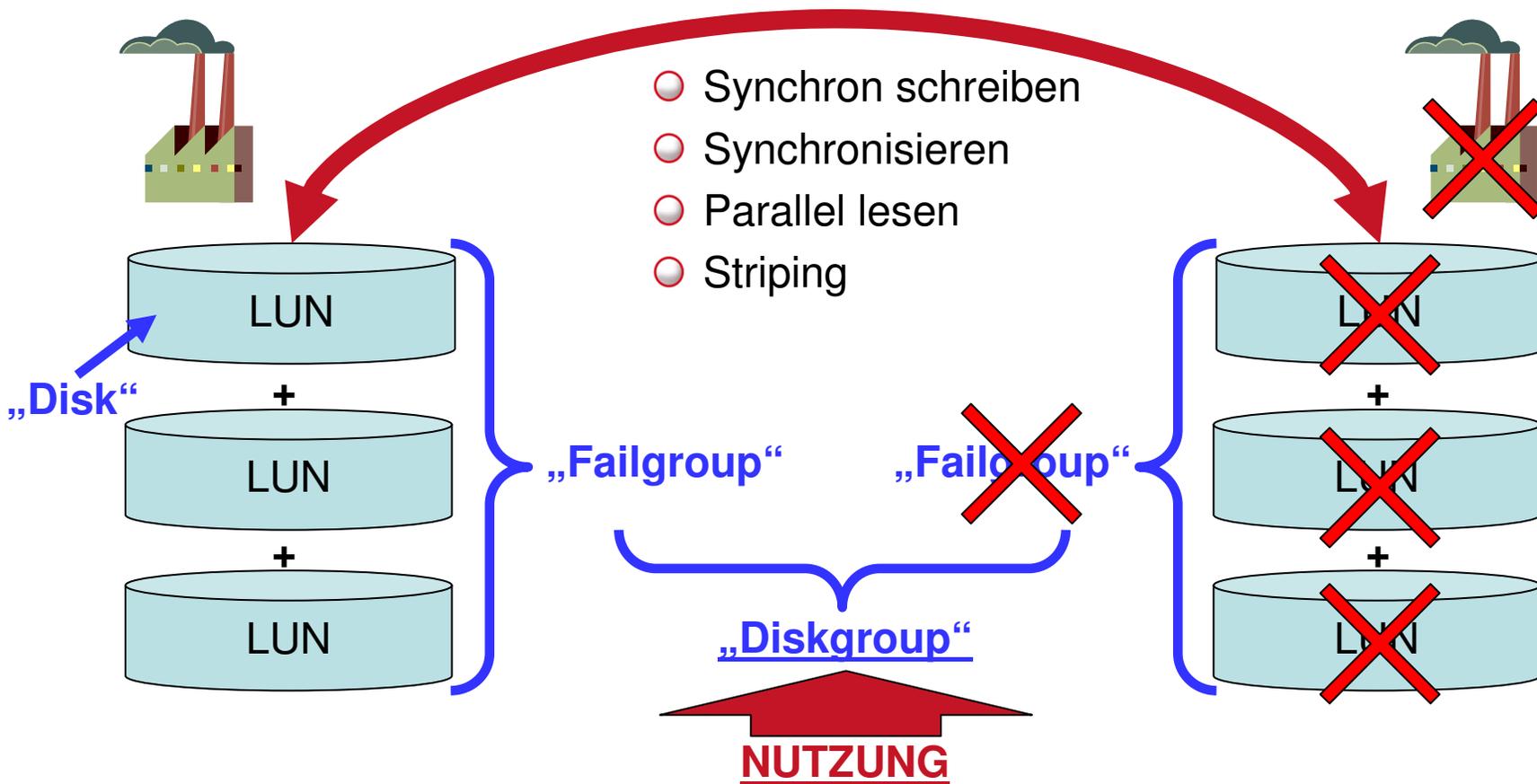
## Mögliche Lösungen:

- ClusterFS + Replikation auf Storageebene
- Replikation auf Hostebene (RAID)
- RAID controller head
- **ASM** im RAC

# Automatic Storage Management ASM



... eingesetzt zur Absicherung und Performanceverbesserung





## **Vorteile (aka Magic):**

- Abstrakter Sublayer
- Selbstorganisation
- Selbstheilung
- Dynamische Reorganisation
- Gute Integration (DB, RAC)

z.B.

- Problemlose, schnelle Migration bei Backendwechsel

## **Nachteile (aka Pitfalls):**

- Kann Fehlerquelle sein
- Wenig Einblick in interne Vorgänge
- Manchmal (zunächst) unerklärliche Vorgänge
- Nicht-intuitives Fehlerhandling

z.B.

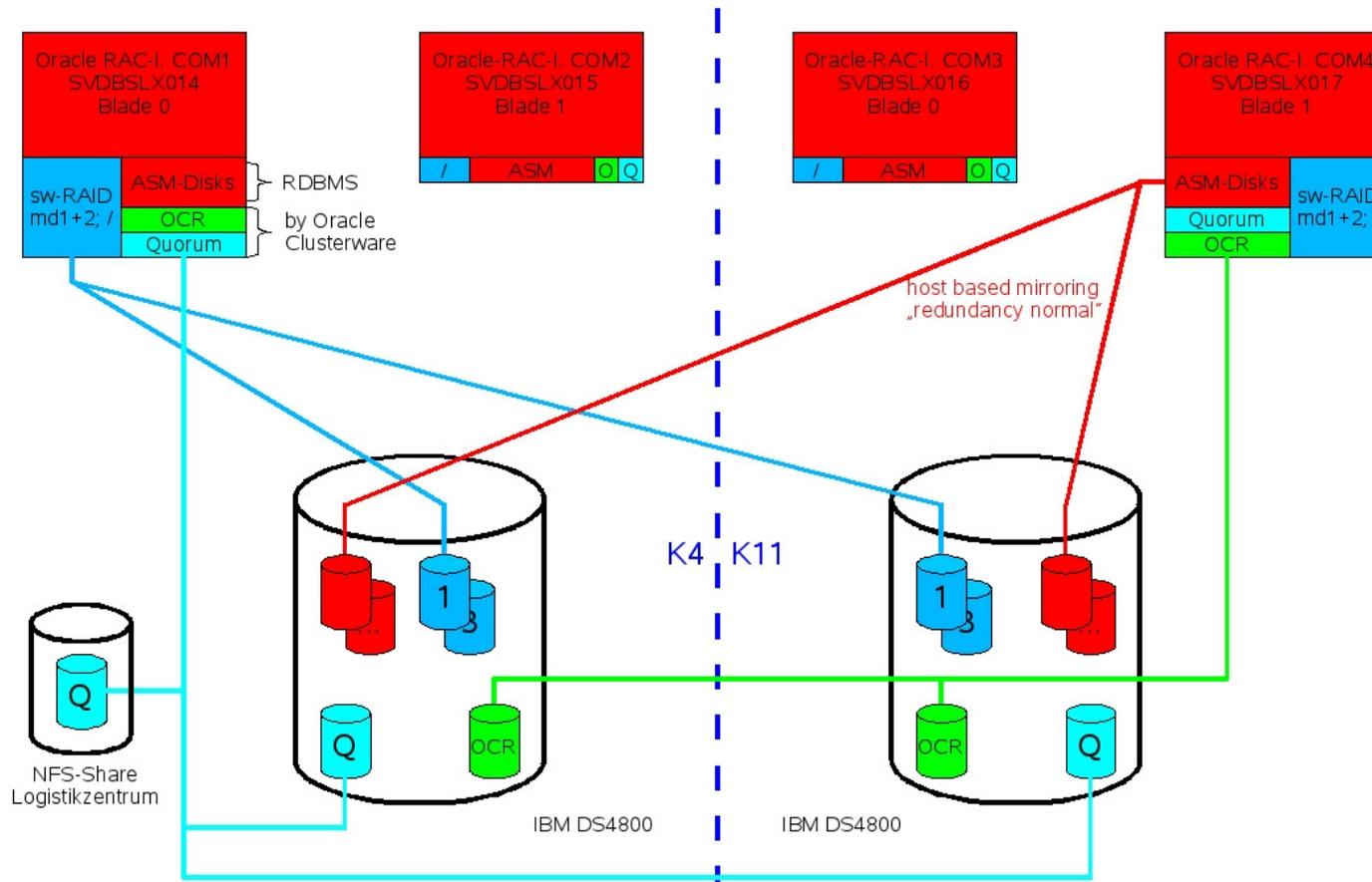
- Tri-State bei Sichtverlust einer SAN-LUN

# Storagezugriff mit RAC

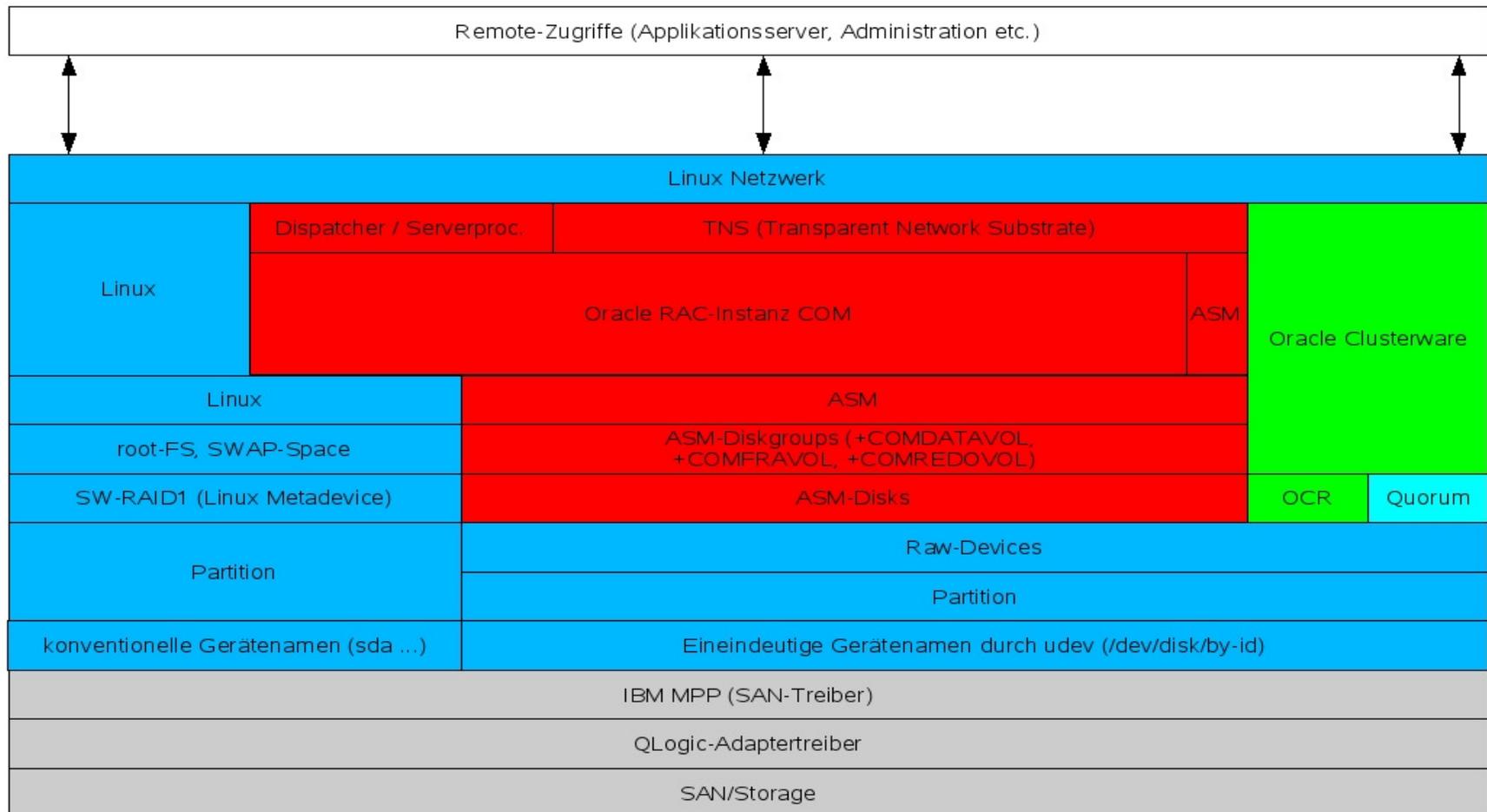


## A.T.U – Com-Server-Cluster, Storagezugriff Primary DB logisch Exemplarisch je Verbindung ein Knoten im Detail

M. Klier  
14.08.2006



# Storagezugriff - Schichtenmodell

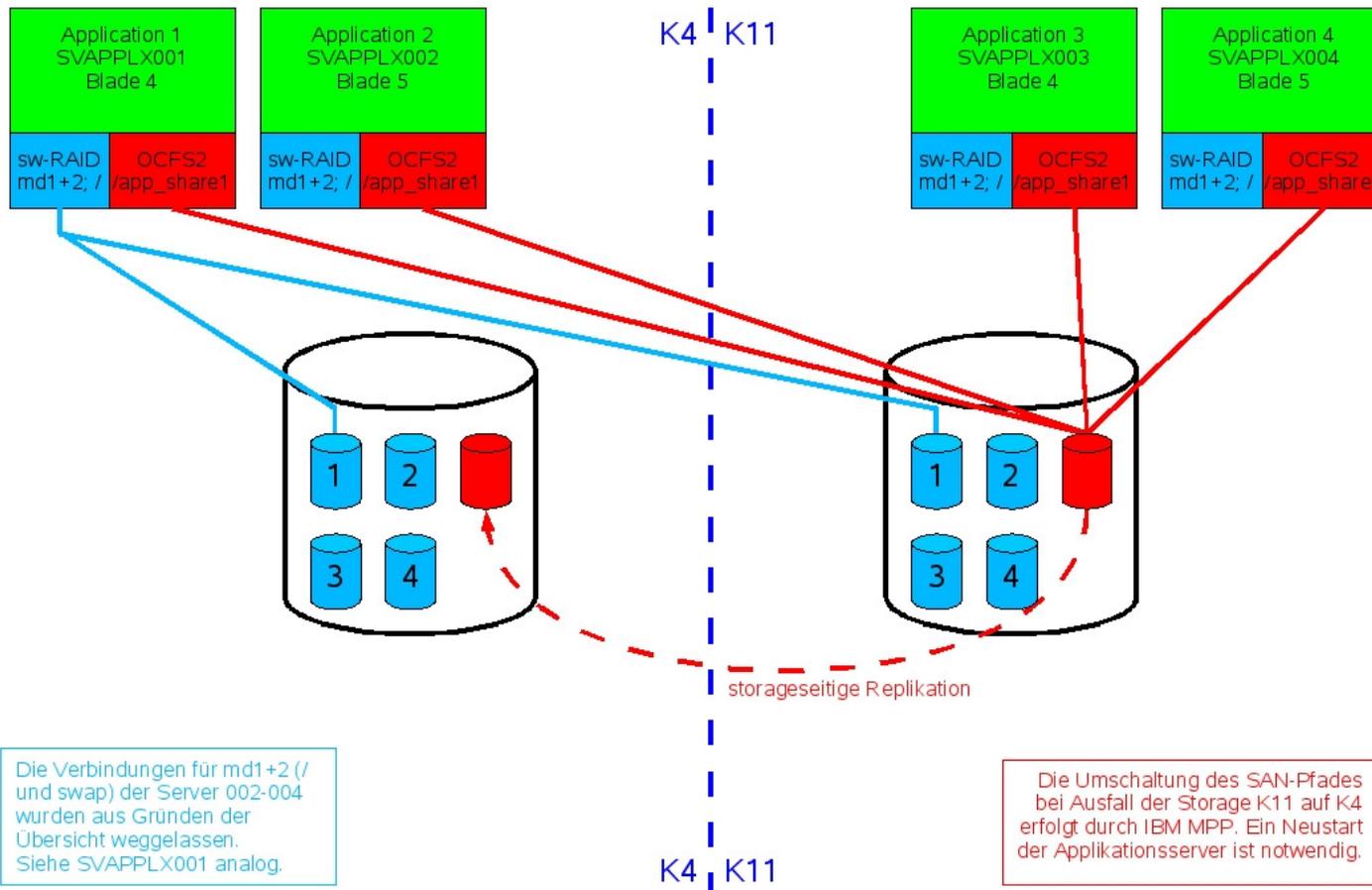


# Storagezugriff Application Server



## A.T.U – Com-Server Applikations-Cluster, Storagezugriff logisch

M. Klier  
14.08.2006

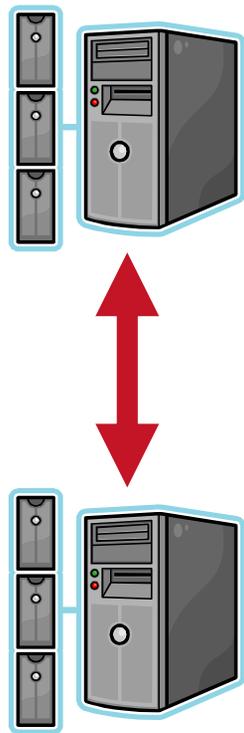


# Data Guard

Physical Standby Database

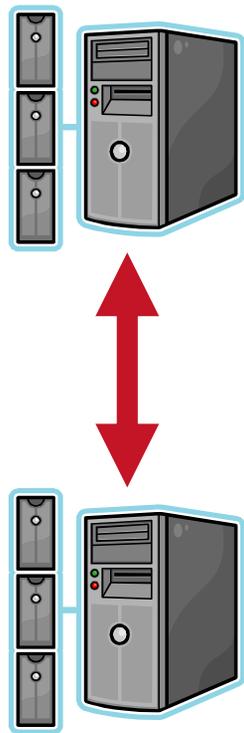


# Anforderungen Data Guard



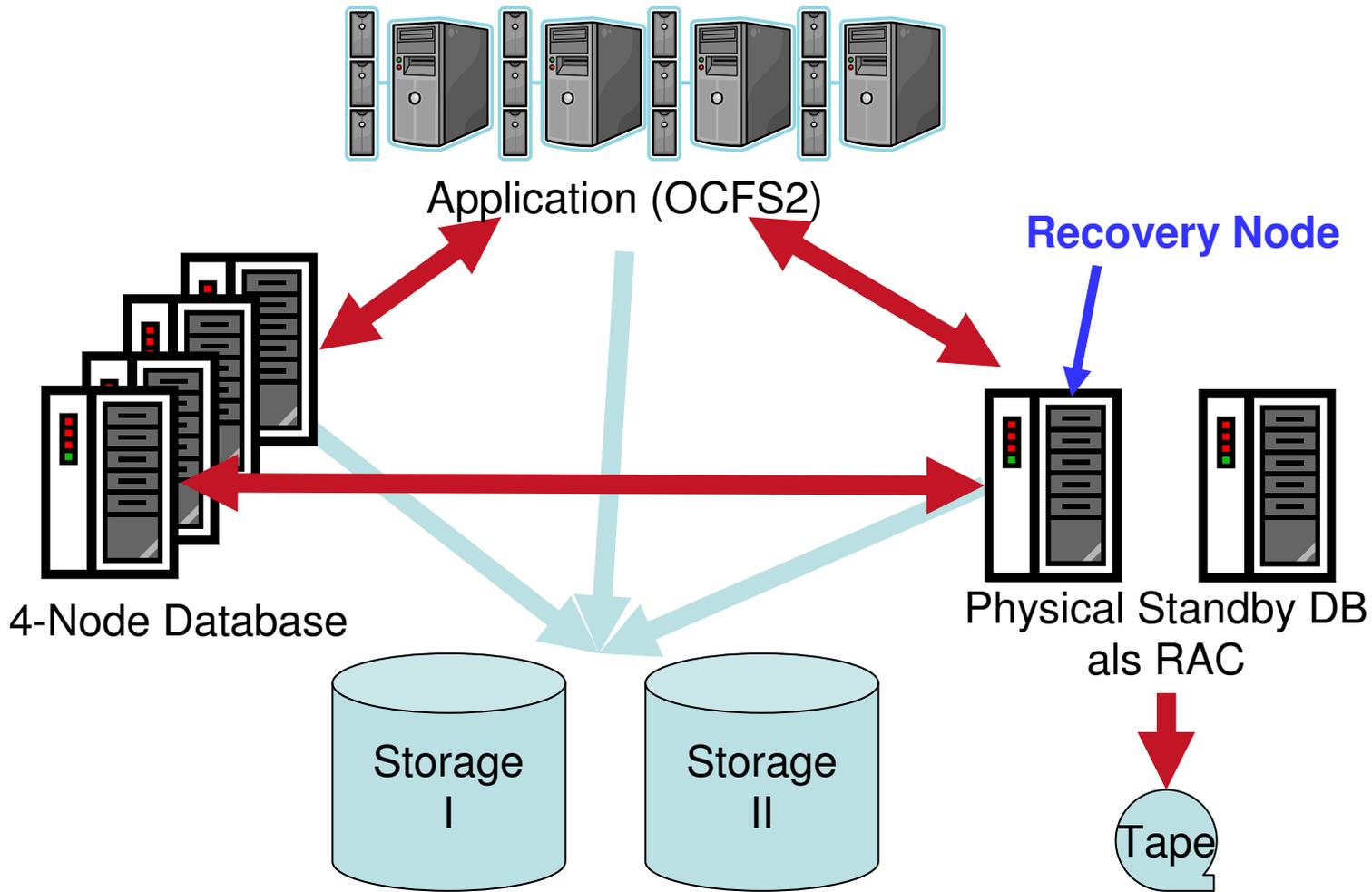
- Schnelles Recovery
- Lastneutrale Datensicherung
- Nur Grundlastdeckung im Failover-Betrieb

# Quick Facts Standby-Database



- 2-Knoten-RAC
- Physical Standby Database
- Real Time Apply
- Logshipping mit TNS Load Balancing
- KEIN automa(t/g)isches Failover zur Standby
  
- Backup mit RMAN

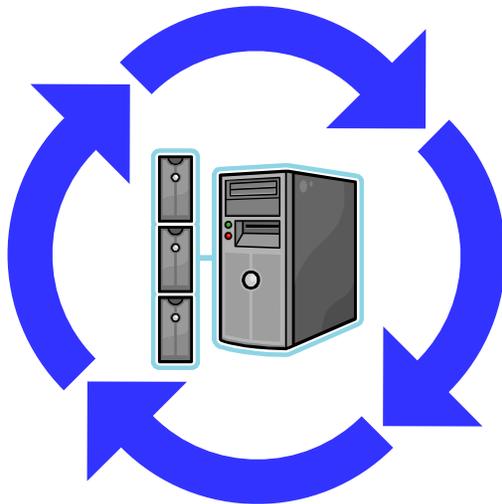
# Überblick Data Guard



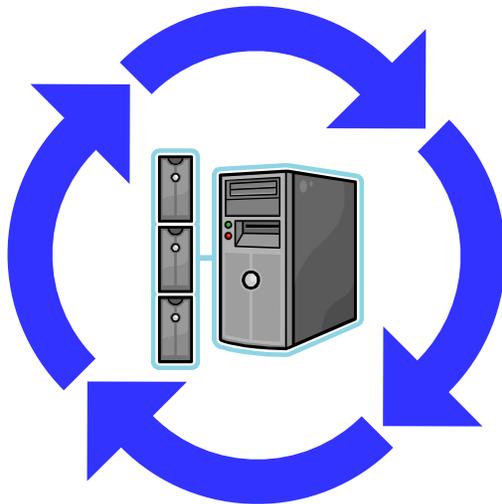
# Application Layer



Fehlertolerante Jobs und Kommunikation

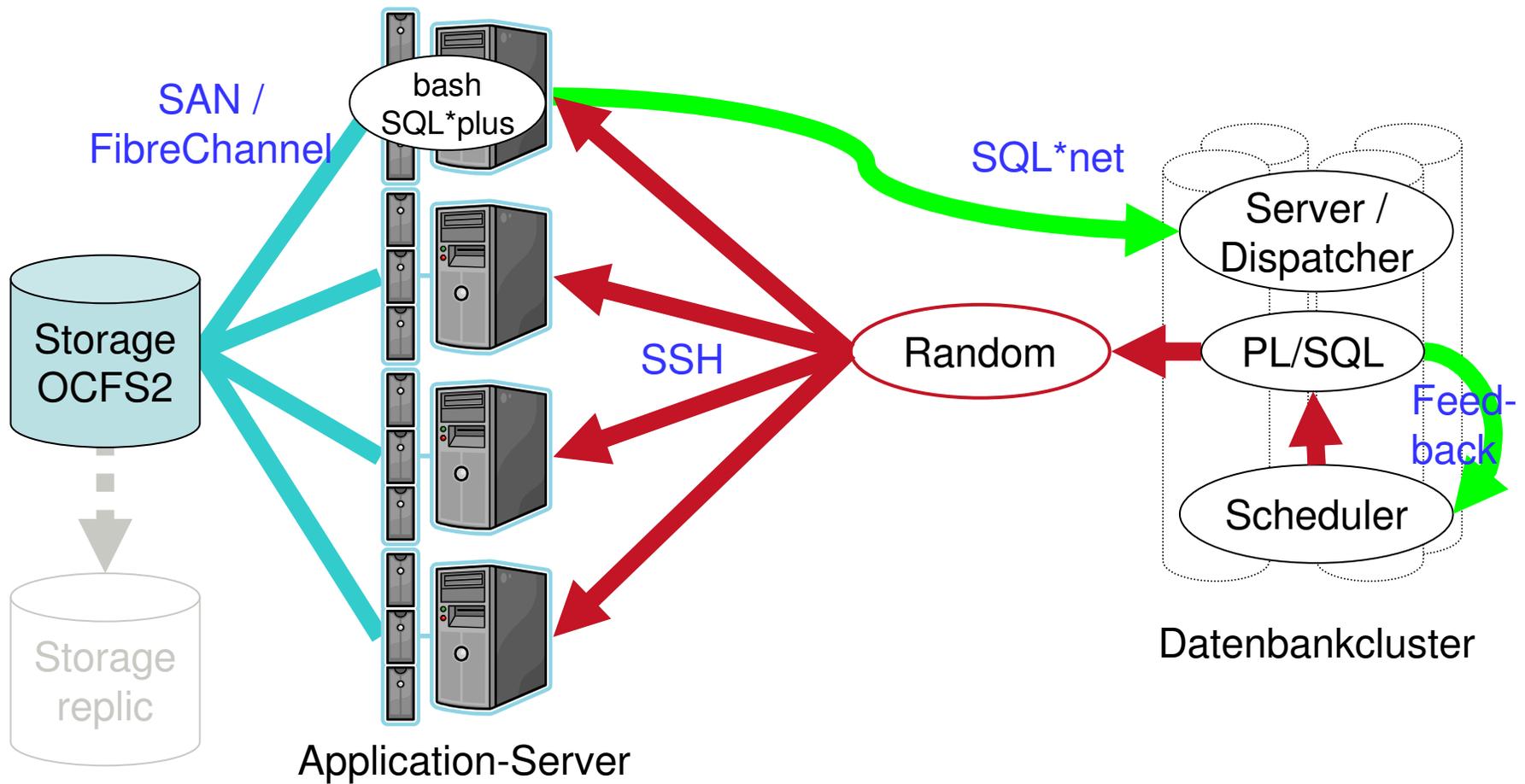


- Hohe Verfügbarkeit
- Wiederanlauf der Jobs nach Rechnerausfall:
  - App-Server
  - DB-Server
- Flexibilität
- Skalierbarkeit
- Sicherheit



- DBMS\_SCHEDULER – Paket
- Registrierte Programme vom Typ
  - EXECUTABLE (extern)
  - STORED PROCEDURE (intern)
- bash
- PL/SQL

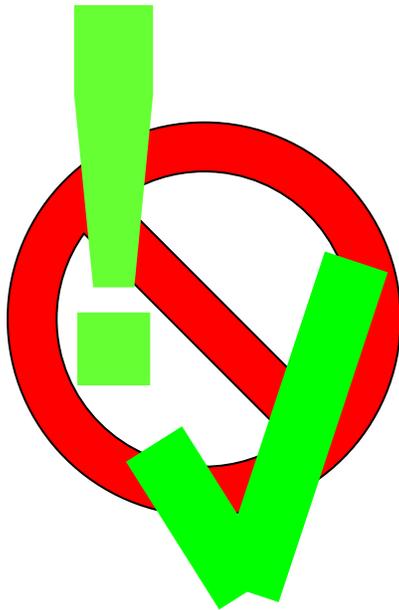
# Überblick Application Layer



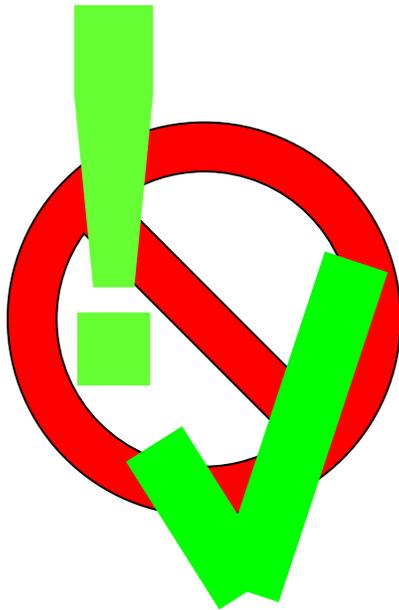
# Lessons Learned

Pitfalls & Co





- HDDs: 15.000rpm  
KEINE gemeinsame Plattennutzung mit anderen Systemen (6.000 - 9.000 IO/s!)
- /dev/raw/raw\* zeigen auf Partitionen
- udev bzw. device mapper nutzen
- elevator=deadline + Parameter
- SAME
- 2x Gigabit-Interconnect
- parallel\_execution\_message\_size=65536
- Entwicklungssystem mit mindestens 1/2 Leistung als RAC
- Standby DB nicht in SAN (lokale Disks)



- ASM: Lost disks vor Wiedereinbinden mit `dd if=/dev/zero` Header leeren
- ASM: Volume-Migration von Medium zu Medium sehr einfach
- Regelmäßige Backups von OCR/Voting Disk UND deren Replay üben
- Vor Arbeiten am Subsystem ggf. CRS-Einträge in der `inittab` auskommentieren
- OCFS2: Wiederanlauf nach Vollcrash tricky



## **Oracle-Software:**

- Oracle Enterprise Edition 10.2.0.3  
64bit + Partitioning
- Real Application Cluster 10.2.0.3  
64bit
- Oracle Clusterware (CRS)
- Automatic Storage Management  
ASM
- OCFS2

## **Hardware:**

- 10x HP Blade BL25p
- 2x IBM Storage DS4800
- 128x FC-HDD 146GB 15.000rpm

## **OS / Treiber:**

- SuSE Linux Enterprise Server 9 SP3  
64bit
- IBM RDAC/MPP-Treiber
- Device Mapper
- OpenSSH
- Bourne Again Shell (bash)



Vielen Dank für Ihre  
Aufmerksamkeit

Martin Klier  
Systemadministration/DB  
A.T.U Auto-Teile-Unger

Dr.-Kilian-Straße 4  
92637 Weiden

[martin.klier@de.atu.eu](mailto:martin.klier@de.atu.eu)

Tel. +49 961 306-5663

Fax +49 961 306-5982

Unterlagen zum Vortrag:

<http://www.usn-it.de>