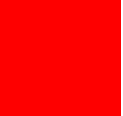
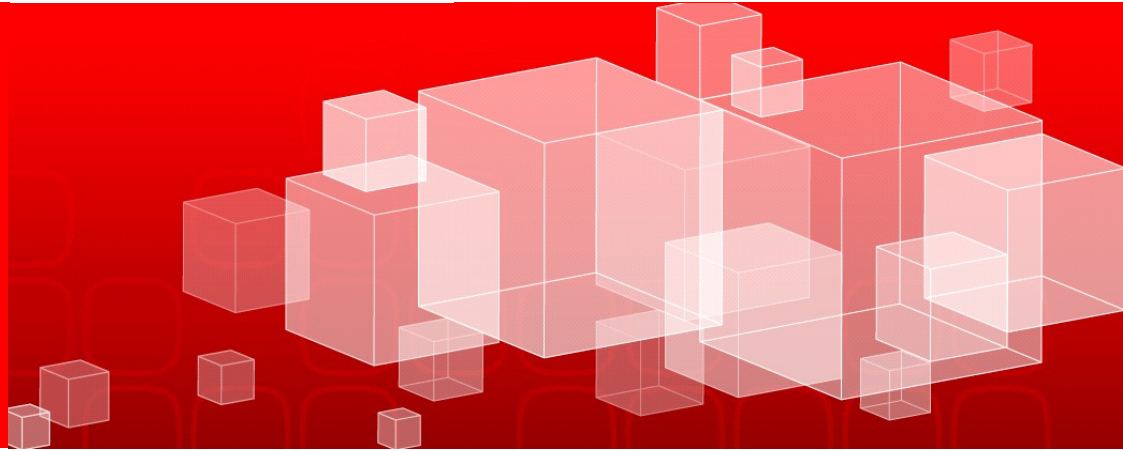


**ORACLE®**



The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.



**ORACLE®**

## **Hochverfügbarkeitslösung vs. K-Fall Absicherung**

Hartmut Streppel

Principal Sales Consultants, Server Architect Northern Europe

# Agenda

- Motivation
- Zwei Missverständnisse
- Fehlerablaufdiagramm und Fehlerklassen
- Hochverfügbarkeitslösungen
- Disaster Recovery Lösungen
- Zusammenfassung

# Motivation

- Eine spezielle deutsche Eigenschaft
  - Campus-/Metro-/“stretched“ Cluster verteilt auf 2 (oder mehr) (entfernte) Rechenzentren
- Fehlende „echte“ K-Fall Absicherung
- Nutzung block-basierter Replikationstechnologien
- Engineered Systeme
- MAA – Maximum Availability Architecture
- .....

# Agenda

- Motivation
- **Zwei Missverständnisse**
- Fehlerablaufdiagramm und Fehlerklassen
- Hochverfügbarkeitslösungen
- Disaster Recovery Lösungen
- Zusammenfassung

# Missverständnis 1

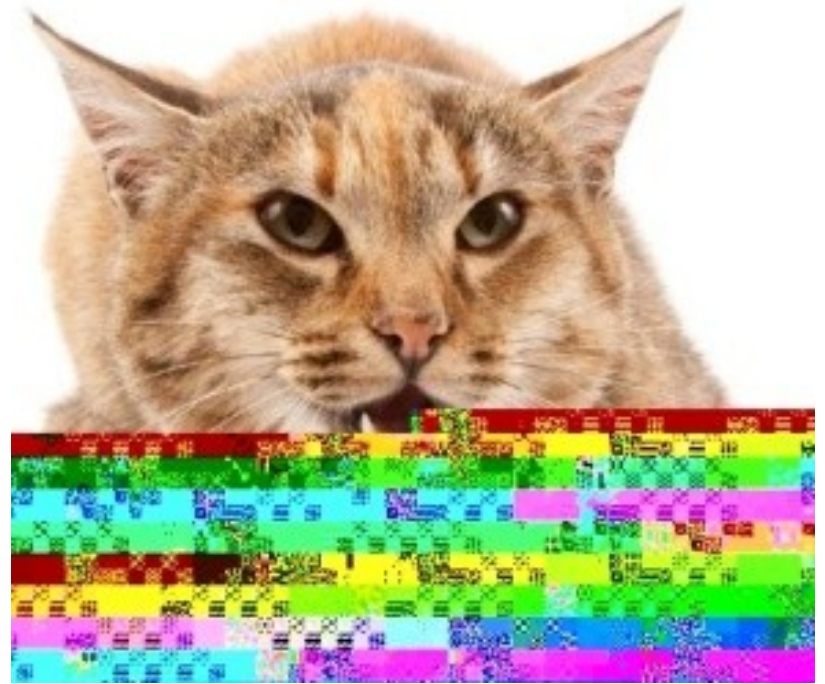
- Eine Katastrophe ist ein sehr selten auftretendes natürliches oder von Menschen erzeugtes Phänomen



UA\_Flight\_175\_hits\_WTC\_south\_tower\_9-11.jpeg: Flickr user TheMachineStops  
[http://en.wikipedia.org/wiki/File:UA\\_Flight\\_175\\_hits\\_WTC\\_south\\_tower\\_9-11\\_edit.jpeg](http://en.wikipedia.org/wiki/File:UA_Flight_175_hits_WTC_south_tower_9-11_edit.jpeg)

# Missverständnis 1

- Eine Katastrophe ist ein sehr selten auftretendes natürliches oder von Menschen erzeugtes Phänomen
- Nein; es gibt viel häufigere, kleinere, aber für die IT nicht weniger katastrophale Ereignisse

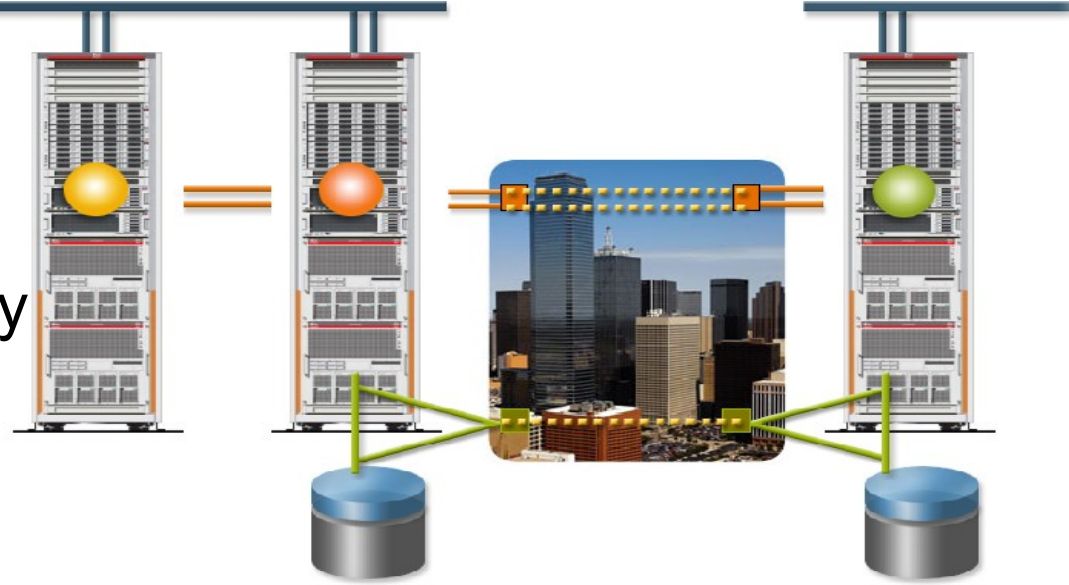


Datenkorruption



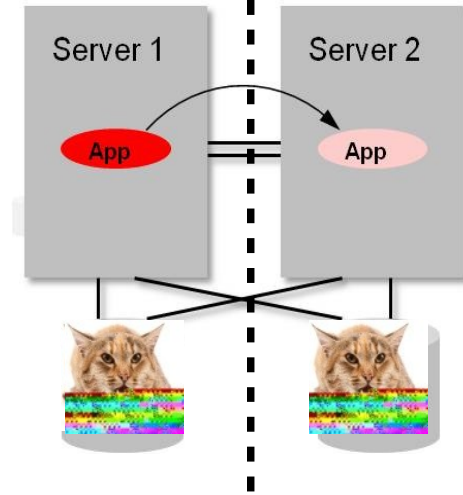
# Missverständnis 2

- Ein Cluster über zwei entfernte Standorte ist eine Disaster Recovery Lösung



# Misverständnis 2

- Ein Cluster über zwei entfernte Standorte ist eine Disaster Recovery Lösung
- Nein! Es überlebt die meisten Katastrophen nicht!
  - Definition einer Katastrophe folgt

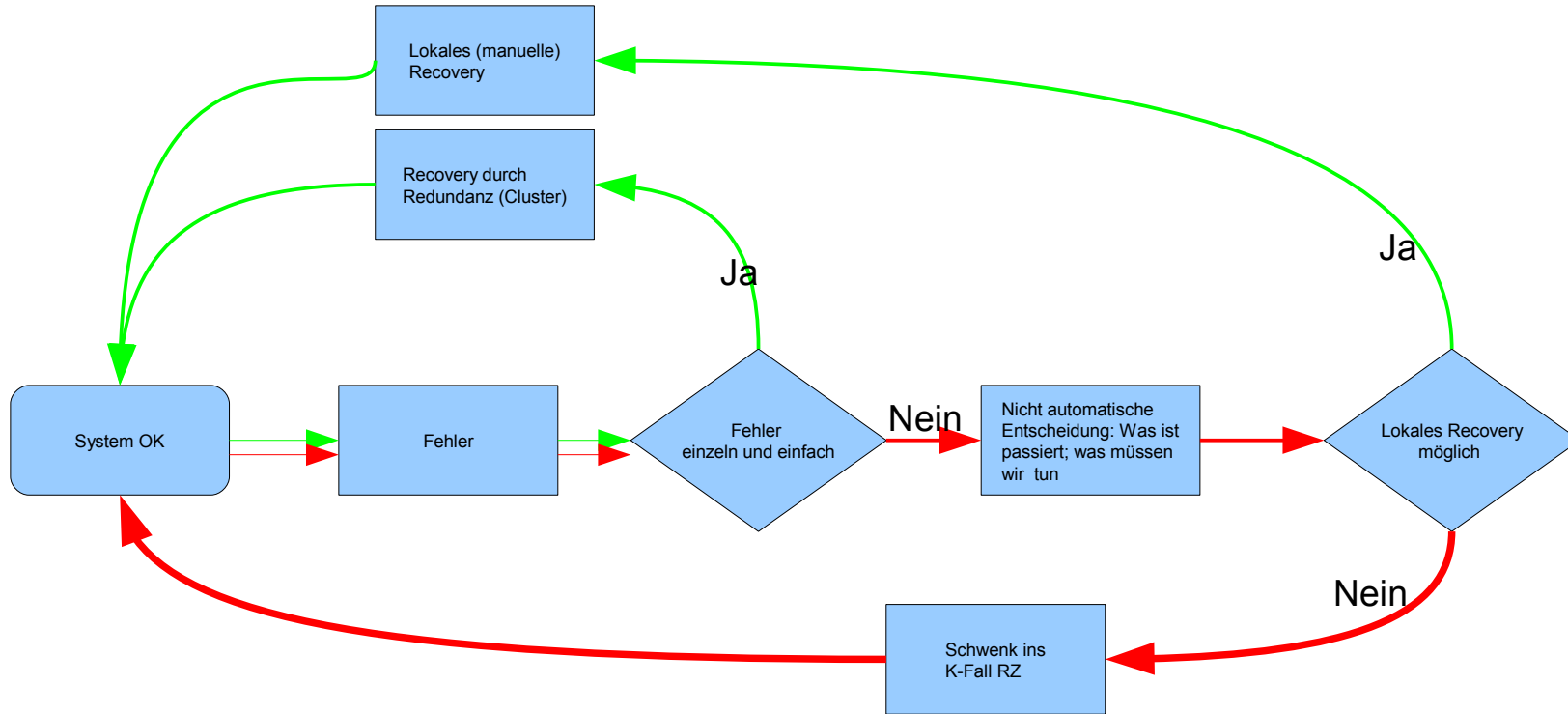


```
Starting cluster:
  Checking if cluster has been disabled at boot... [ OK ]
  Checking Network Manager... [ OK ]
  Global setup... [ OK ]
  Loading kernel modules... [ OK ]
  Mounting configfs... [ OK ]
  Starting cman... [ OK ]
  Waiting for quorum... Timed-out waiting for cluster
[FAILED]
cluster not ready - no quorum?
```

# Agenda

- Motivation
- Zwei Missverständnisse
- Fehlerablaufdiagramm und Fehlerklassen
- Hochverfügbarkeitslösungen
- Disaster Recovery Lösungen

# Ein Fehler tritt auf; was tun?



# Fehlerklassen

## Definition



# Fehlerklassen

## Definition

### F I - Einfache Einzelfehler

Vitale (HW-)Komponenten sind redundant.

Aber: Es gibt noch Single Points of Failure! (z.B: OS, Backplane...

Single  
System

HW | HW  
SW

# Fehlerklassen

## Definition

### F II – Nicht triviale Einzelfehler

Redundante Systeme: Wenige SPOF

### F I - Einfache Einzelfehler

Vitale (HW-)Komponenten sind redundant.

Aber: Es gibt noch Single Points of Failure! (z.B: OS, Backplane...)

Single  
System

HW | HW  
SW

Single  
System

HW | HW  
SW

Single  
System

HW | HW  
SW

# Fehlerklassen

## Definition

### F III - Doppelfehler und komplexe Einzelfehler

Mehrfachredundante, unabhängige Systeme (u.U. mehrere Sites)  
Disaster Recovery möglich.

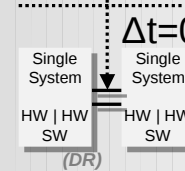
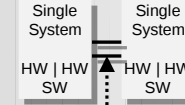
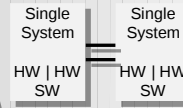
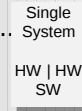
### F II – Nicht triviale Einzelfehler

Redundante Systeme: Wenige SPOF

### F I - Einfache Einzelfehler

Vitale (HW-)Komponenten sind redundant.

Aber: Es gibt noch Single Points of Failure! (z.B: OS, Backplane...)





# Fehlerklassen

## Definition

### F IV - Datenkorruption

Mehrfachredundante Systeme über mehrere Sites!

DR, plus Absicherung gegen logische Fehler durch z.B.: Nachlauf, **unterschiedliche** Architekturen

### F III - Doppelfehler und komplexe Einzelfehler

Mehrfachredundante, unabhängige Systeme (u.U. mehrere Sites)

Disaster Recovery möglich.

### F II – Nicht triviale Einzelfehler

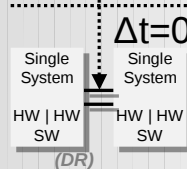
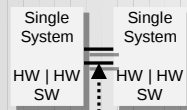
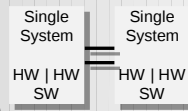
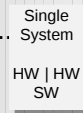
Redundante Systeme: Wenige SPOF

### F I - Einfache Einzelfehler

Vitale (HW-)Komponenten sind redundant.

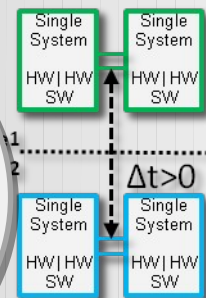
Aber: Es gibt noch Single Points of Failure! (z.B: OS, Backplane...)

## Katastrophe



$\Delta t = 0$

(DR)



DR

# Fehlerklassen

## Definition

### F IV - Datenkorruption

Mehrfachredundante Systeme über mehrere Sites!

DR, plus Absicherung gegen logische Fehler durch z.B.: Nachlauf, **unterschiedliche** Architekturen

### F III - Doppelfehler und komplexe Einzelfehler

Mehrfachredundante, unabhängige Systeme (u.U. mehrere Sites)

Disaster Recovery möglich.

### F II – Nicht triviale Einzelfehler

Redundante Systeme: Wenige SPOF

### F I - Einfache Einzelfehler

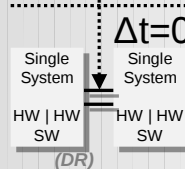
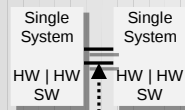
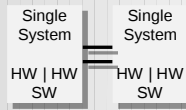
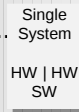
Vitale (HW-)Komponenten sind redundant.

Aber: Es gibt noch Single Points of Failure! (z.B: OS, Backplane...)

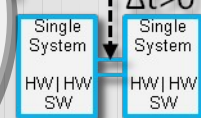
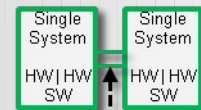
### F0 – kein Fehler

Ausfall einer Komponente führt zum Ausfall des Gesamtsystems

Single System



$\Delta t=0$



$\Delta t>0$

DR

# Theoretischer Überbau

## Disaster Recovery – Business Continuity – RPO - RTO

- „Disaster recovery (DR) is the process, policies and procedures that are related to preparing for recovery or continuation of technology infrastructure which are vital to an organization after a natural or human-induced disaster.“
  - „Disaster recovery is a subset of business continuity. (BC)“
  - „Disaster recovery focuses on the IT or technology systems that support business functions, as opposed to business continuity, which involves planning for keeping all aspects of a business functioning in the midst of disruptive events.“  
([http://en.wikipedia.org/wiki/Disaster\\_recovery](http://en.wikipedia.org/wiki/Disaster_recovery))
- RTO = Recovery Time Objective: Maximale Zeit, die nach einer Serviceunterbrechung verstreichen darf bis der Service wieder verfügbar ist
- RPO = Recovery Point Objective: Maximale Dauer einer Unterbrechung, in der Daten verloren gehen dürfen.

# Keine Integration von DR in einen BC Plan

## Ein beispielhaftes Disaster

- Katastrophe: Feuer im RZ
  - Erfolgreich gelöscht; IT Systeme schalten um in 2. Brandabschnitt im gleichen Gebäude
  - Aber: Gebäude muss geräumt werden
  - Folge: Keine Arbeitsplätze; ungenügende Bandbreite ins RZ; keine FAX-Umleitung, Aufträge landen 24h im evakuierten RZ
- DR-Projekt muss immer im Einklang mit einem BC Projekt durchgeführt werden



# Agenda

- Motivation
- Zwei Missverständnisse
- Fehlerablaufdiagramm und Fehlerklassen
- Hochverfügbarkeitslösungen
- Disaster Recovery Lösungen
- Zusammenfassung

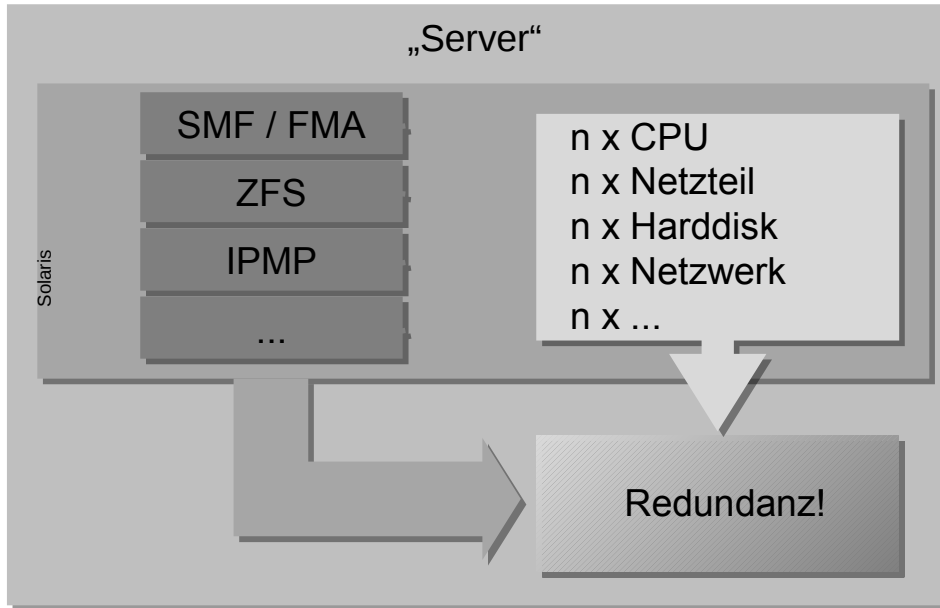
# Hochverfügbarkeit – Die vielen Neunen

- Hochverfügbarkeit ab ~99,95% Verfügbarkeit
  - d.h. ca. 4h Downtime pro Jahr
- Allein durch manuelle Überwachung und Eingriffe schwer zu erreichen
- Notwendig
  - Hohe Redundanz auf allen Ebenen
  - Clustertechnologie
- Wie berechnet man die Gesamtverfügbarkeit
  - Wartungsfenster ?

# Redundanz in Einzelsystemen

Systemverfügbarkeit - Hardware & Software

Single System  
HW | HW  
SW



Systemverfügbarkeit durch:

- Redundante Systemarchitektur
- Systemsoftware

Ziel:

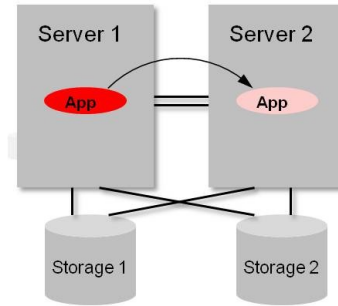
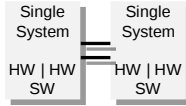
Möglichst keine SPOFs im System!

Baustein für hochverfügbare Architekturen!

ORACLE

# Clustertechnologien

## Implementierungen



### Entfernung

- d=0: „cluster in a box“
- near - 1 Raum
- near - 2 Räume
- far - 2 Sites (typ.: ~10 km)

### Daten

- 1-fach
- Redundant
  - RAID 1-Site
  - RAID 2-Sites
- Replikation
  - synchron
  - asynchron
- Konsistenz
  - logisch (Inhalte)
  - physisch (Daten)

### Verfahrensweise / Entscheidungslogik

- Manueller Cluster
- Automatischer Cluster
  - Physik
  - Logik
- Entscheidungslogik
  - z.B. Quorum

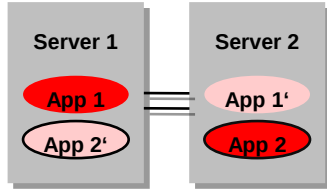
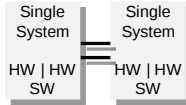
### Topologie

- 1:1
- n:1

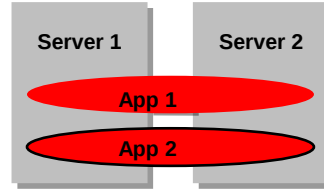


# Clustertechnologien

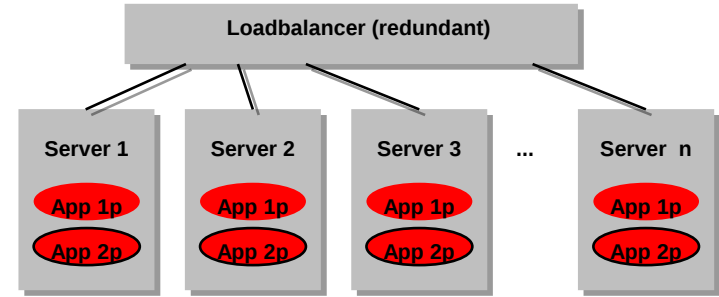
## Clusterverfahren



Failoveranwendung  
z.B. SAP CI



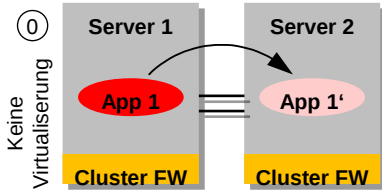
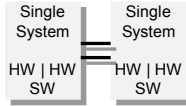
Parallele Anwendung  
z.B. Oracle RAC



Parallele Architektur (Grid)  
z.B. Webserver

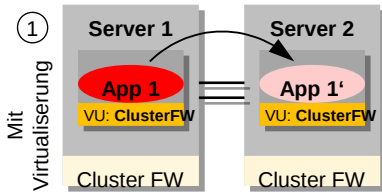
# Clustertechnologien

## Failover Cluster Typen



### Failover Cluster

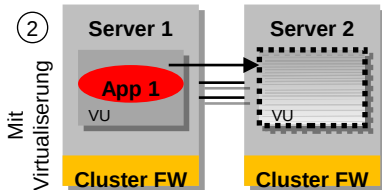
- Applikation schwenkt
- Applikationsüberwachung mit HA Agenten



### Failover Cluster

- Applikation schwenkt
- Applikationsüberwachung mit HA Agenten
- Failover Cluster in einer VU

Z.B. 1a) Zonen Cluster, LDOM Cluster, 1b) Zone Nodes (nur mit Solaris 10 / OSC 3.x)



### Failover Cluster

- VU schwenkt
- Applikationsüberwachung mit SMF Scripten
- Clustersteuerung in der globalen VU

Z.B: Failover Zonen / ‚Flying zones‘, Failover LDOMs

VU: Virtuelle Umgebung / Virtuelle Instanz  
Cluster FW: Cluster Software Framework

# Anforderungen an Cluster

- „Auch eine sehr unzuverlässige Airline bringt Sie an 364 Tagen im Jahr sicher(!) nach Hause“
- Monitoring aller Ressourcen, auch der Anwendungen
- Doppelsicherung durch Failure Fencing
- Abdeckung der Extremfälle
- Integrierbarkeit eigener Anwendungen
- Integrierbarkeit in Disaster Recovery Lösungen
- z.B. Oracle Solaris Cluster, (Oracle Grid Infrastructure)

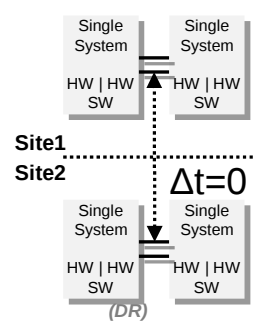
# Agenda

- Motivation
- Zwei Missverständnisse
- Fehlerablaufdiagramm und Fehlerklassen
- Einige Grundbegriffe
- Hochverfügbarkeitslösungen
- **Disaster Recovery Lösungen**
- Zusammenfassung

# K-Fall Absicherung

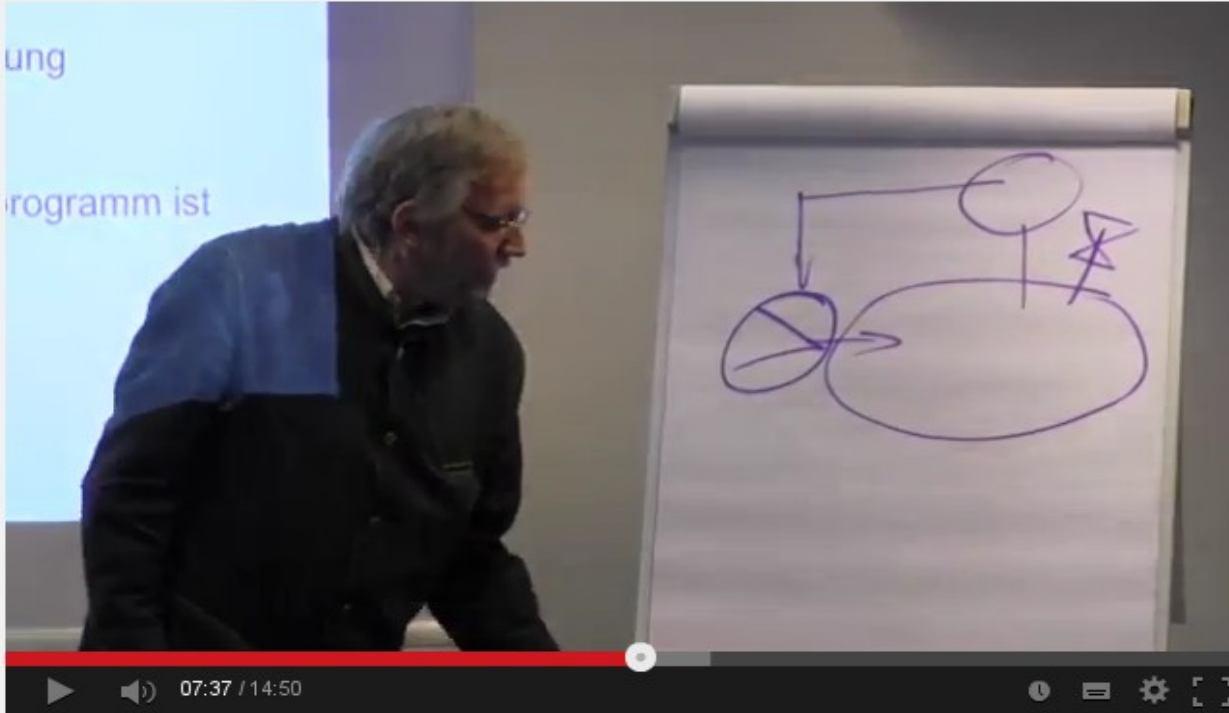
## Anforderungen an eine DR-Lösung

- Weitestgehende Unabhängigkeit
  - Kein Einfluss eines Fehlers auf die DR-Infrastruktur
- Möglichst geringe RPO
  - Weniger ist besser → kein Datenverlust
- Möglichst geringe RTO
  - Kurze Downtime ist besser: keine Service-Unterbrechung
- “Möglichst gering” definiert als Geschäftsanforderung, denn:
  - $RTO = 0$  und  $RPO = 0$  ist sehr teuer!!!

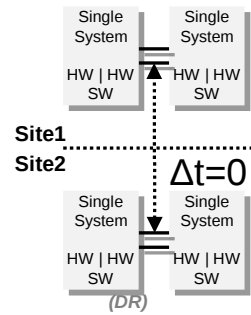


# K-Fall Absicherung

## Anforderungen an eine DR-Lösung



Hans Bonfigt - Vorsicht E-Mail!



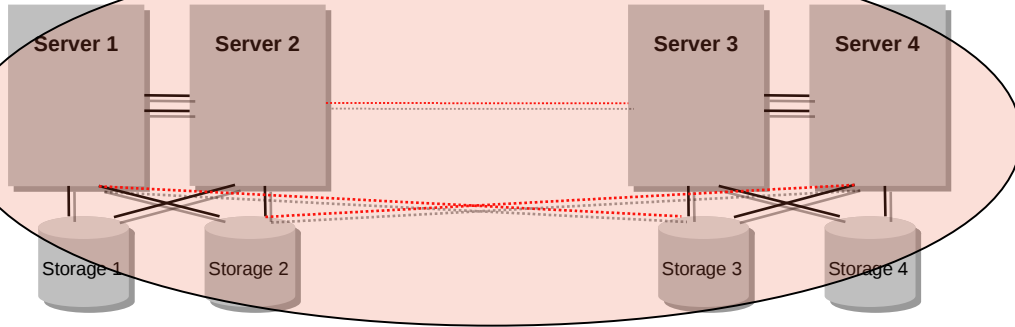
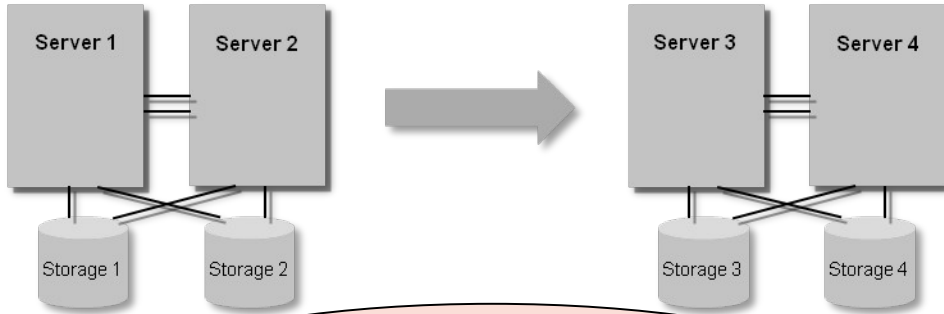
- Homogene Redundanz
- vs.
- diversitäre/dissimilare Redundanz (prinzip-verschieden)

• [http://www.youtube.com/watch?feature=player\\_embedded&v=ca6y52caaRw](http://www.youtube.com/watch?feature=player_embedded&v=ca6y52caaRw)

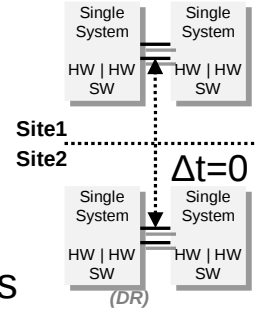
ORACLE

# K-Fall Absicherung

Weitestgehende Unabhängigkeit



- Eliminierung weiterer SPOFs
- Getrennte, unabhängige Daten
- Unterschiedliche Versionen
- Unterschiedliche Technologien???
- ....
- Kein shared SAN über Sites
- Kein Layer-2 Netz über Sites
- Kein Cluster über Sites
- ....



# DR-Lösungen

- Sollten bestehen aus
  - Datenreplikation
    - z.B. Oracle Data Guard
  - Orchestrierung
    - z.B. Oracle Solaris Cluster Geographic Edition



# Datenreplikation

- Schutz vor Doppelfehlern
  - Oracle Data Guard: maximum protection mode
  - Verschlechterung der Verfügbarkeit bei Erhöhung der Datensicherheit
- Replikation kompletter Ablaufumgebungen (z.B. virtuelle Maschinen) sichert nicht gegen Fehler in diesen Ablaufumgebungen
  - Verstoß gegen das Gebot weitestgehender Unabhängigkeit

# Agenda

- Motivation
- Zwei Missverständnisse
- Fehlerablaufdiagramm und Fehlerklassen
- Hochverfügbarkeitslösungen
- Disaster Recovery Lösungen
- Zusammenfassung

# Zusammenfassung

- Zwei Rechenzentren sind noch keine Vorsorge gegen eine Katastrophe
- Katastrophen können klein und trotzdem massiv sein
  - Datenkorruption
- MAA
  - „Lokale“ Vorsorge gegen einfache Fehler durch Redundanzen und Cluster
  - K-Fall Vorsorge durch eine echte Disaster Recovery Lösung

# Q&A

Hartmut.Streppel@oracle.com

**ORACLE®**