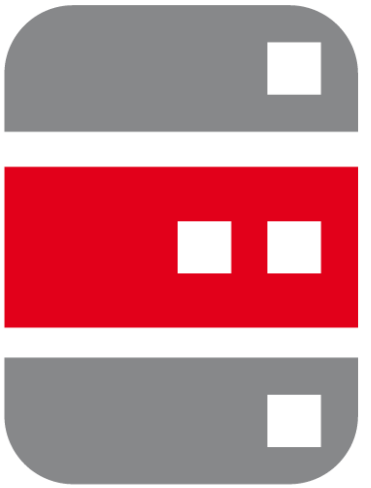
performing databases



Your reliability. Our concern.

performing databases
Your reliability. Our concern.

42 Fakten zu Oracle RAC, Grid Infrastructure und ASM

Martin Klier 

Performing Databases GmbH
Mitterteich / Germany

Speaker

performing databases
Your reliability. Our concern.

- Martin Klier
- Solution Architect and Database Expert

- My focus
  - Performance Optimization
  - High Availability
  - Architecture DBMS

- Linux since 1997
- Oracle Database since 2003

ORACLE ACE

Speaker

performing databases
Your reliability. Our concern.

- Meet & Greet

DOAG
Regionalgruppen
Fachkonferenzen
DOAG Konferenz & Ausstellung

IOUG
independent oracle users group

COLLABORATE

- Contact: martin.klier@performing-db.com
- Weblog: http://www.usn-it.de (English)

Performing Databases

performing databases
Your reliability. Our concern.

ORACLE Gold Partner

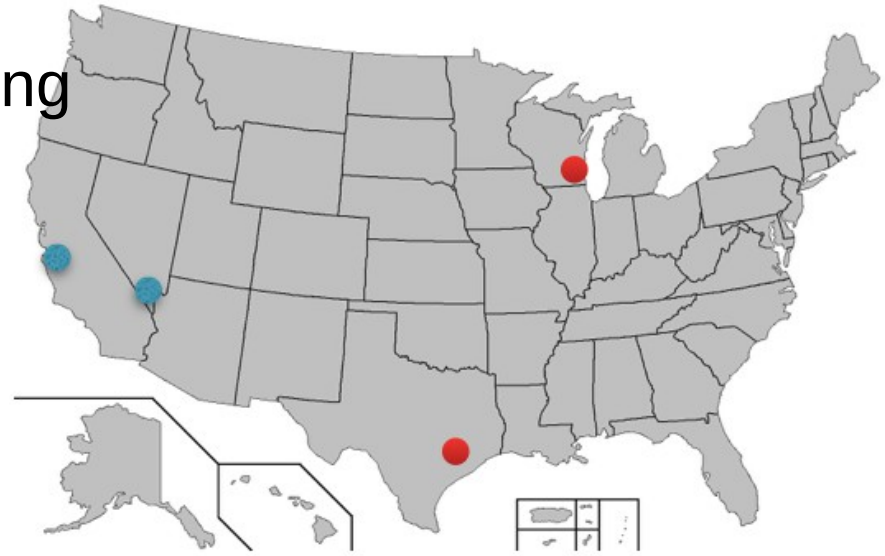
- Experts for the Oracle Database
  - Concept
  - Planning & Sizing
  - Licensing
  - Implementation and Troubleshooting
- Get in touch
  - Performing Databases GmbH
    Wiesauer Straße 27
    95666 Mitterteich, GERMANY
  - Web: http://www.performing-databases.com
  - Twitter: @PerformingDB

International

performing databases
Your reliability. Our concern.

ORACLE Gold Partner

- Design
- Licensing
- Implementation
- Tuning
- Troubleshooting
- Service
- Upgrade
- Migration



http://www.performing-databases.com

Tally-Ho!

performing databases
Your reliability. Our concern.

Seid Ihr Helden oder Memmen?

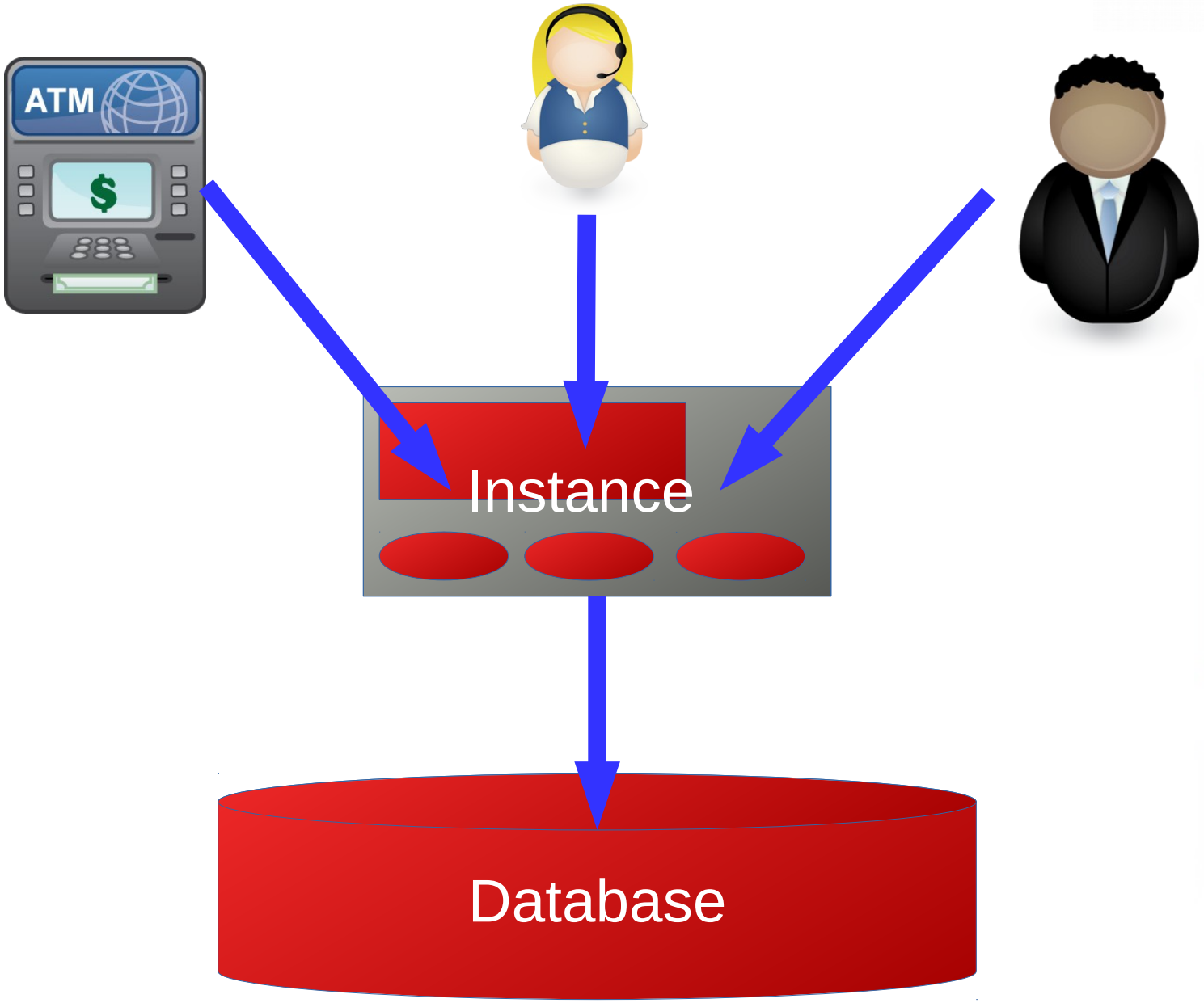
Memmen, aber ganz harte!

performing databases
Your reliability. Our concern.

Basics

Database

performing databases
Your reliability. Our concern.



performing databases
Your reliability. Our concern.

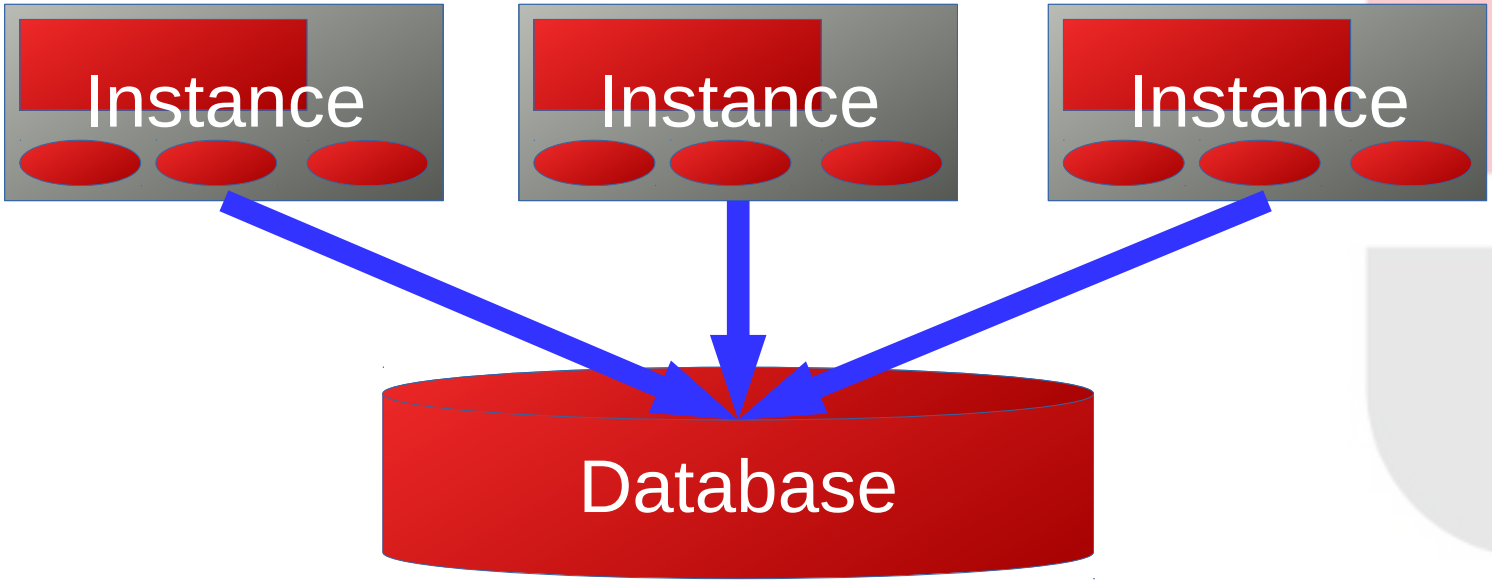
RAC

RAC

performing databases
Your reliability. Our concern.

- Real Application Clusters
  = Produkt

- Real Application Cluster
  = Grid Infrastructure + Datenbank
  mit mehreren möglichen Instanzen

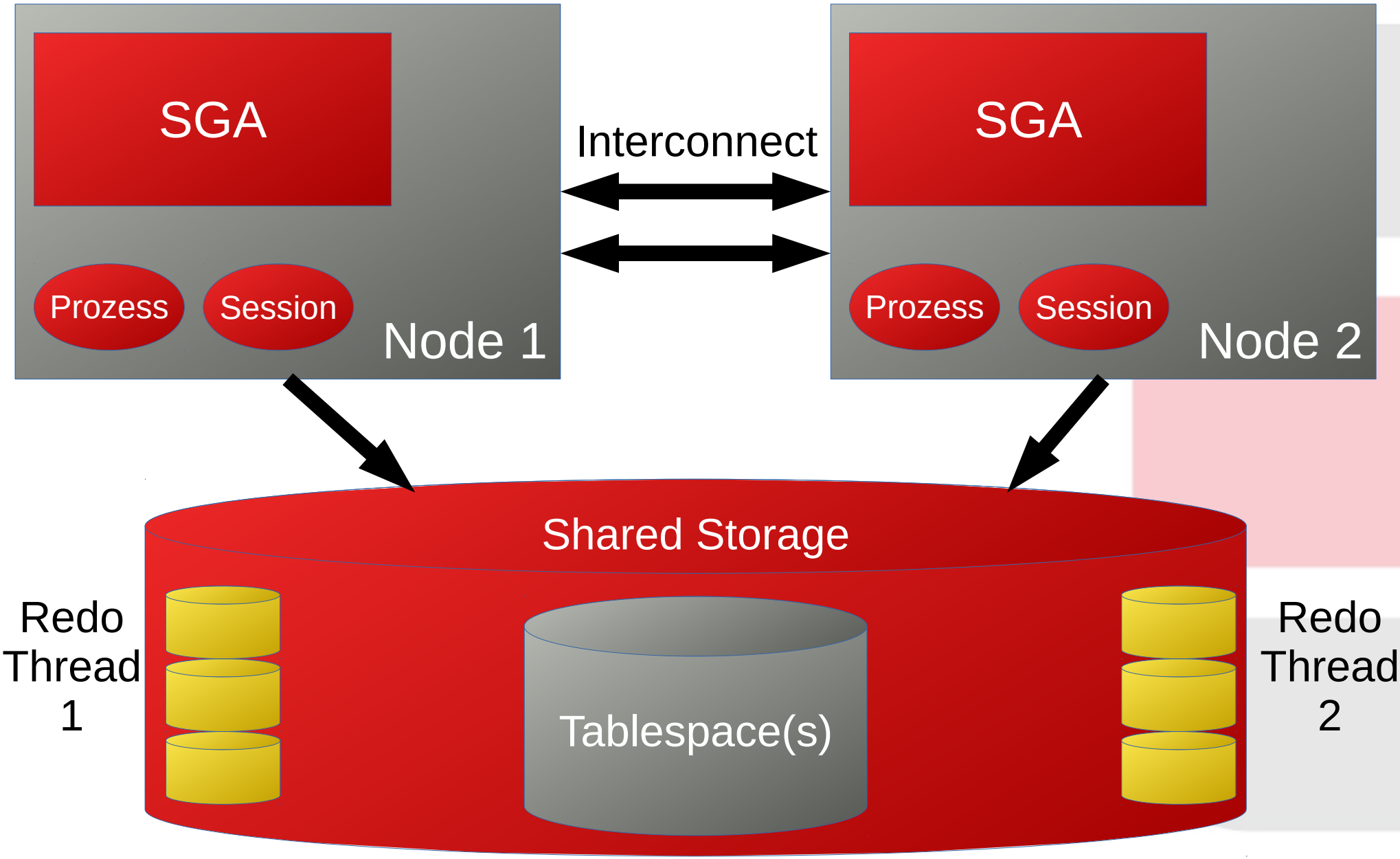


RAC

performing databases
Your reliability. Our concern.

- Shared-Everything-Architektur
  (jedes Node sieht alle Daten)
- Queries liefern auf allen Nodes gleiche Ergebnisse
- CacheFusion
  (US Patent US20060117074 A1 von Ahmet K. Ezzat)
  erzeugt virtuellen Gesamt-Cache
  (aber on demand => Block Shipping)
- Parsing erfolgt Node-spezifisch (Skalierung!)
  Ausführungspläne unterscheiden sich ggf.

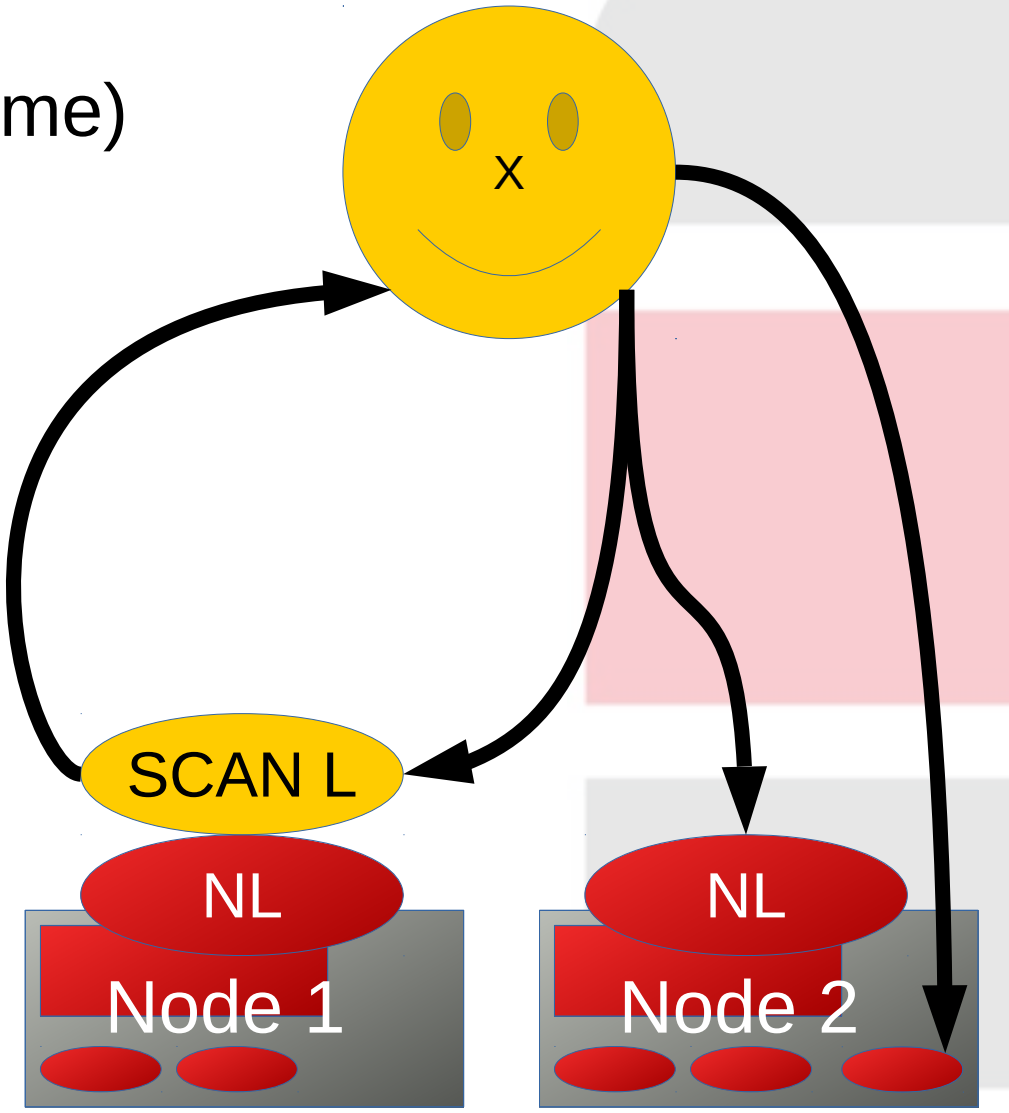
Grundfunktion



Listener

performing databases
Your reliability. Our concern.

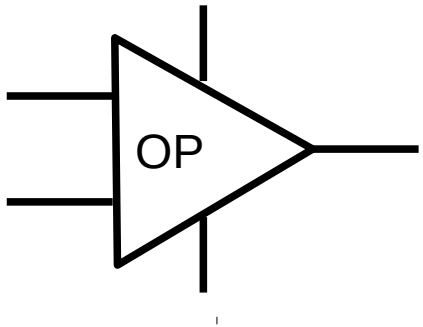
- SCAN Listener (11.2)
  (Single Cluster Access Name)
- Standard: 3x SCAN
  Praxis: 1x reicht zu 95%
- Node-Listener



RAC Performance

performing databases
Your reliability. Our concern.

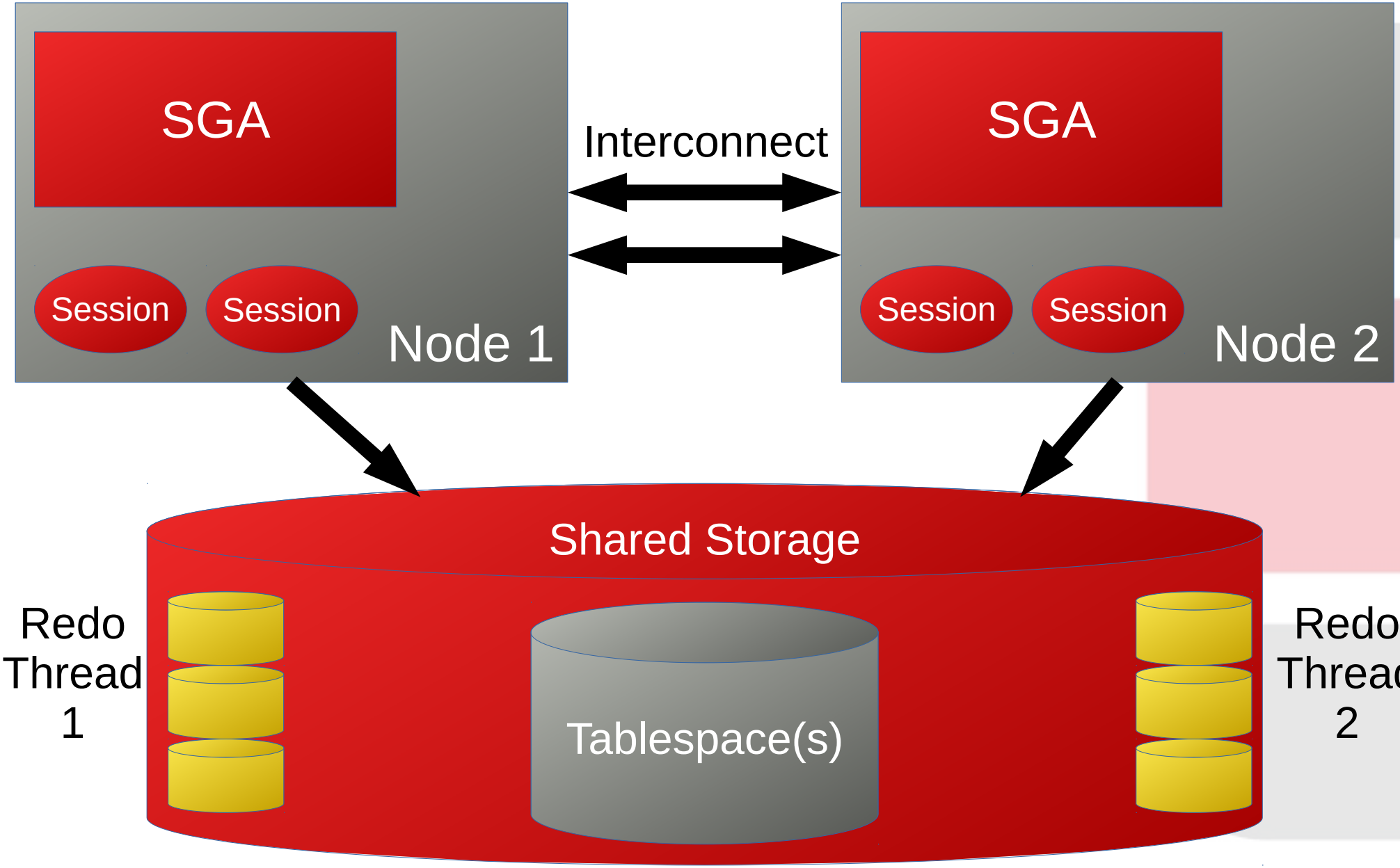
- RAC = Verstärker :)



- Sequenzen
  - ordered
  - cached
- 1 Query nicht „größer“ als 1 Node
- Interconnect-Überlastung vermeiden

RAC Performance

performing databases
Your reliability. Our concern.



Diverses

performing databases
Your reliability. Our concern.

- Auto-DOP (Degree of Parallelism) im RAC
  => Verteilung PX Worker auf die Nodes
  Interessant für InMemory ohne Exa!

- InMemory IMCUs ohne Exa nicht redundant
  Grund: Implementiert mit RDMA
  => Geht nur mit Infiniband
  (Mellanox gibt nur für definierte Umgebung frei)

performing databases
Your reliability. Our concern.

ASM

ASM

- Automatic Storage Management
  = Produkt
- Logical Volume Manager
- Software-RAID
- Cluster - ~~Dateisystem~~
- **Cluster - LVM**

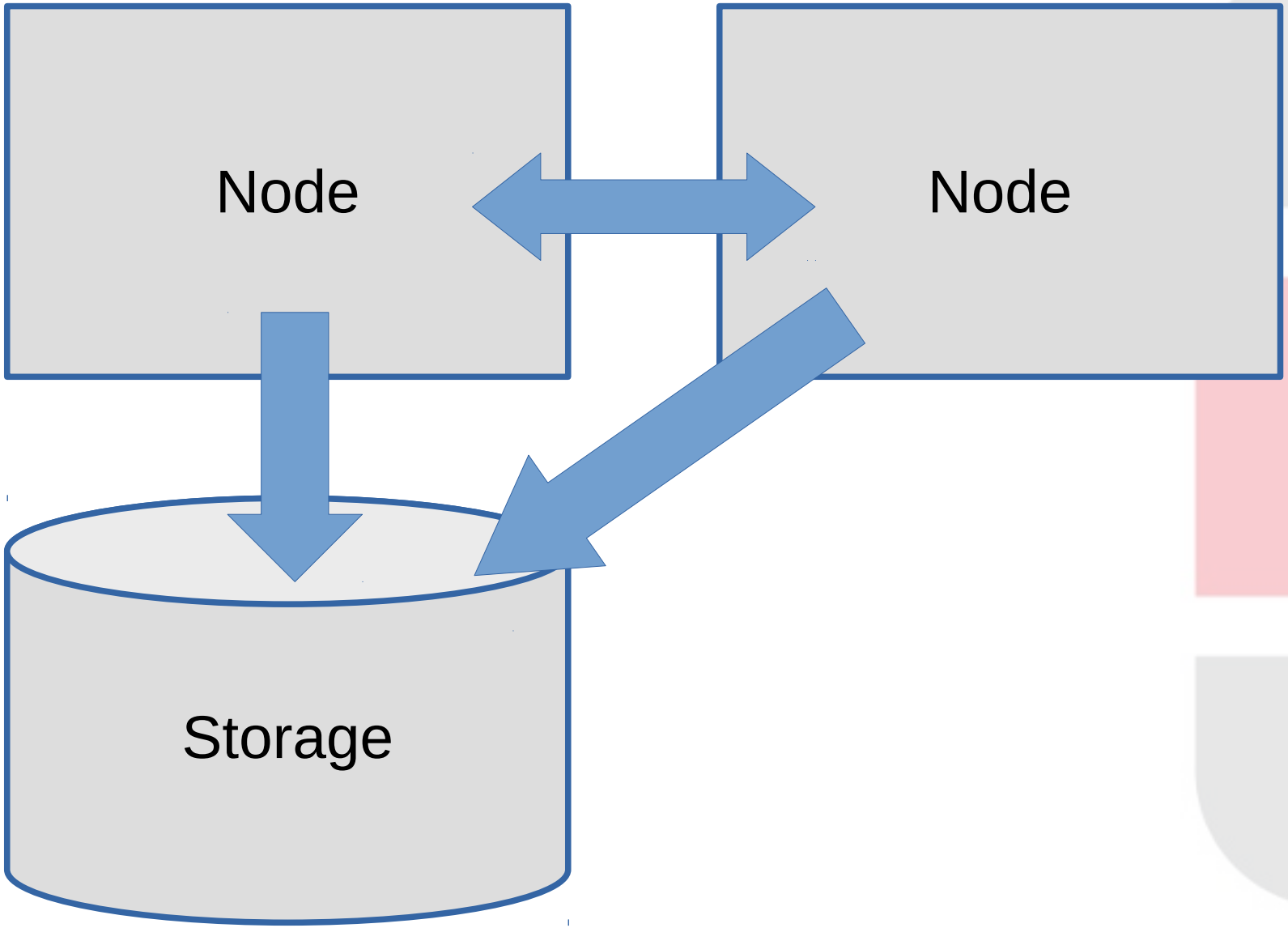
ASM

performing databases
Your reliability. Our concern.

- Verbindet Vorteile
  von Raw Devices (wenig Overhead)
  mit den von Dateisystemen (einfache Verwaltung)
- DB greift direkt auf Volumes zu
- Schlechter Ruf: Ich hätte gerne meine Datafiles <...>!
- „Menschenfreundliche“ Administration
  über diverse UIs
  (SQL*plus, asmcmd -p, EM CloudControl)

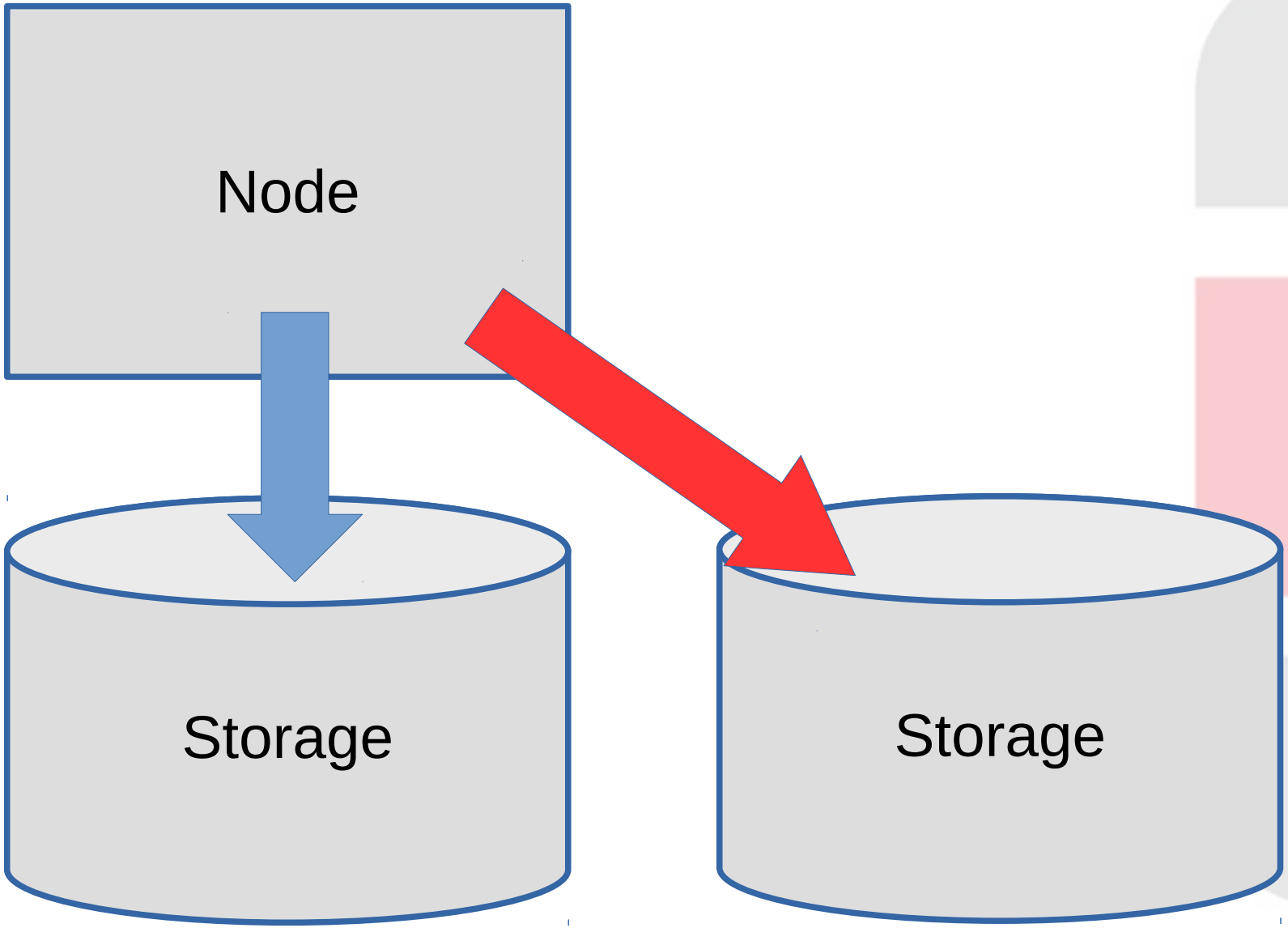
Classic I

performing databases
Your reliability. Our concern.



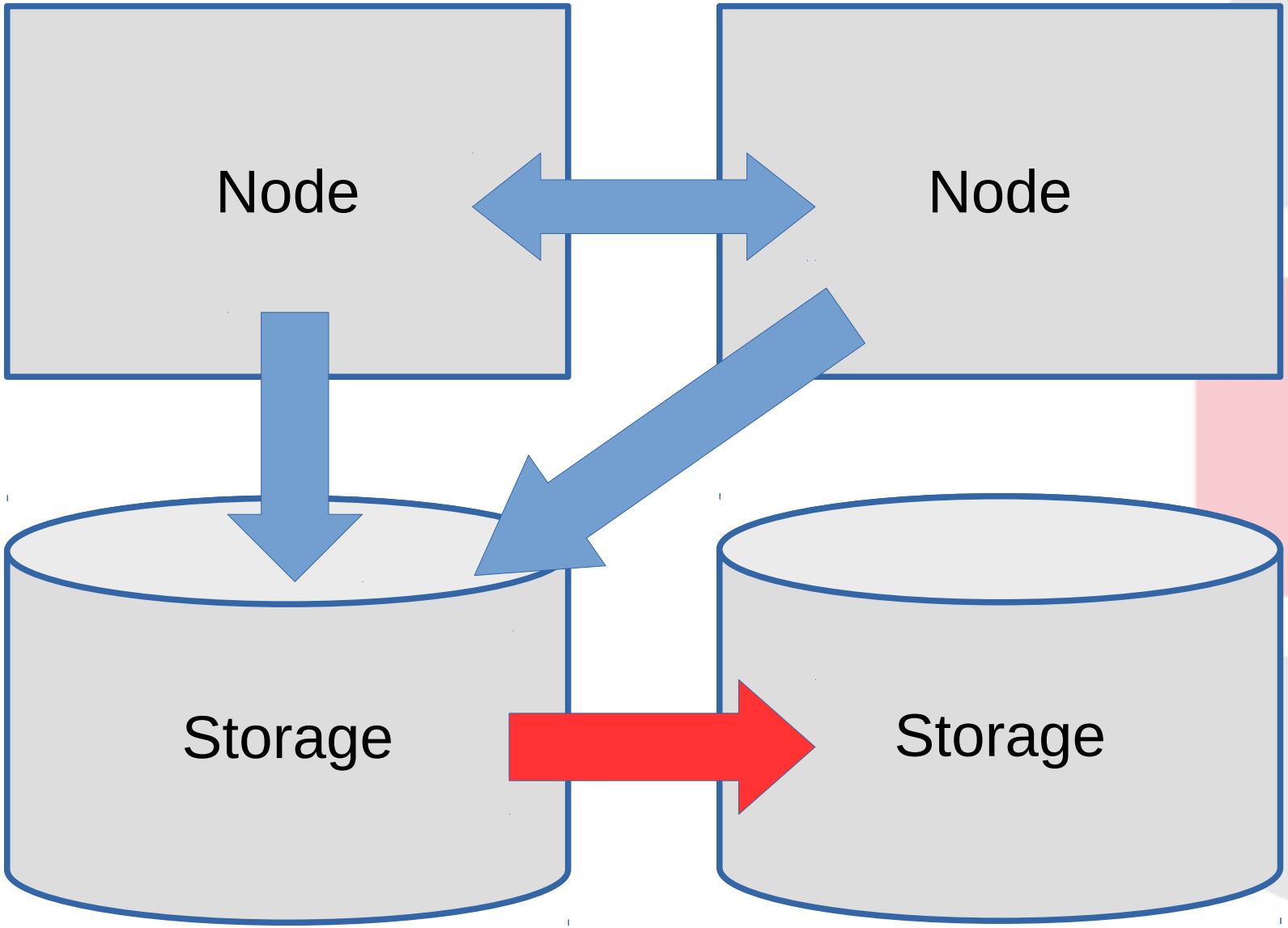
Classic II

performing databases
Your reliability. Our concern.



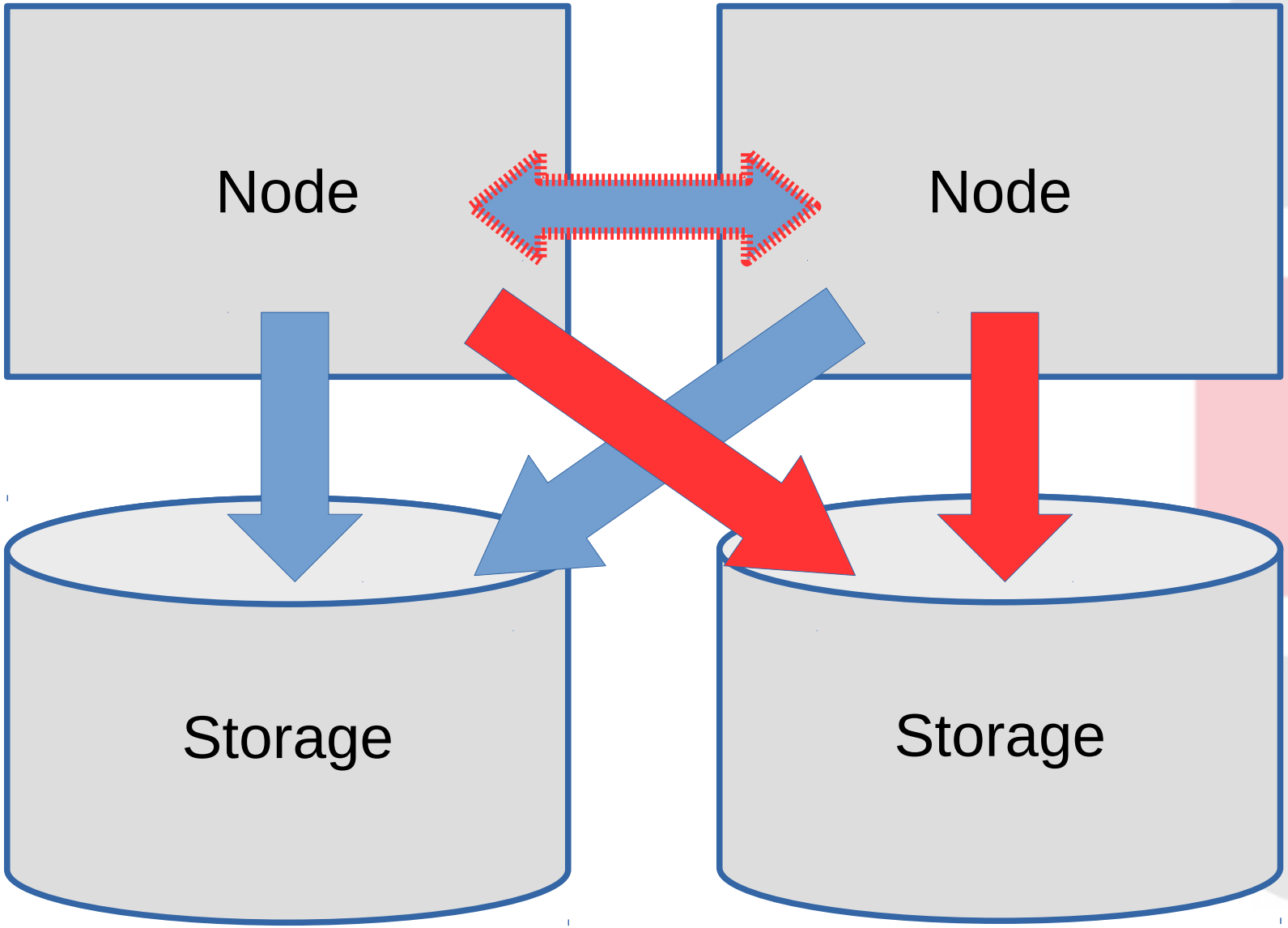
Classic III

performing databases
Your reliability. Our concern.



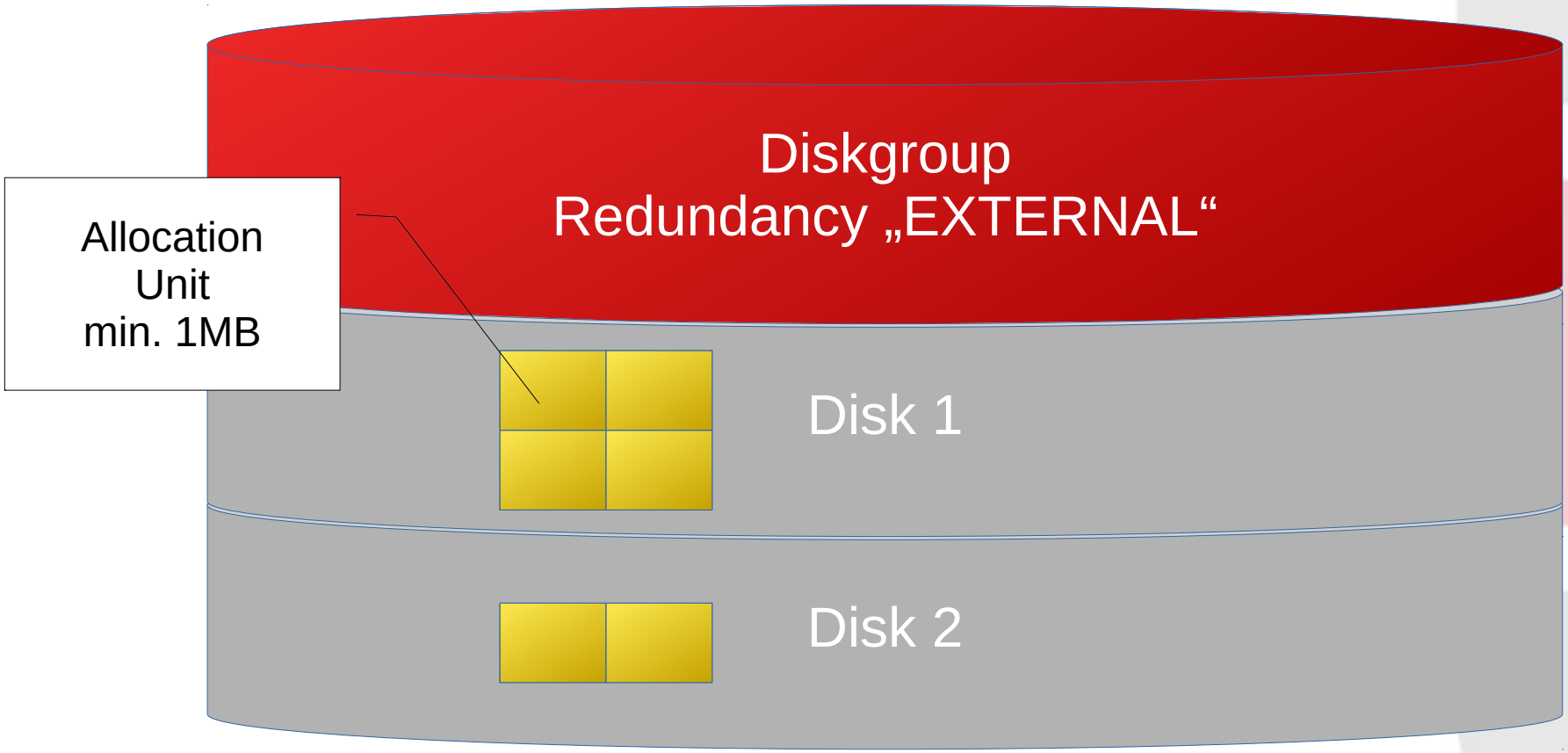
ASM++

performing databases
Your reliability. Our concern.



ASM Diskgroups

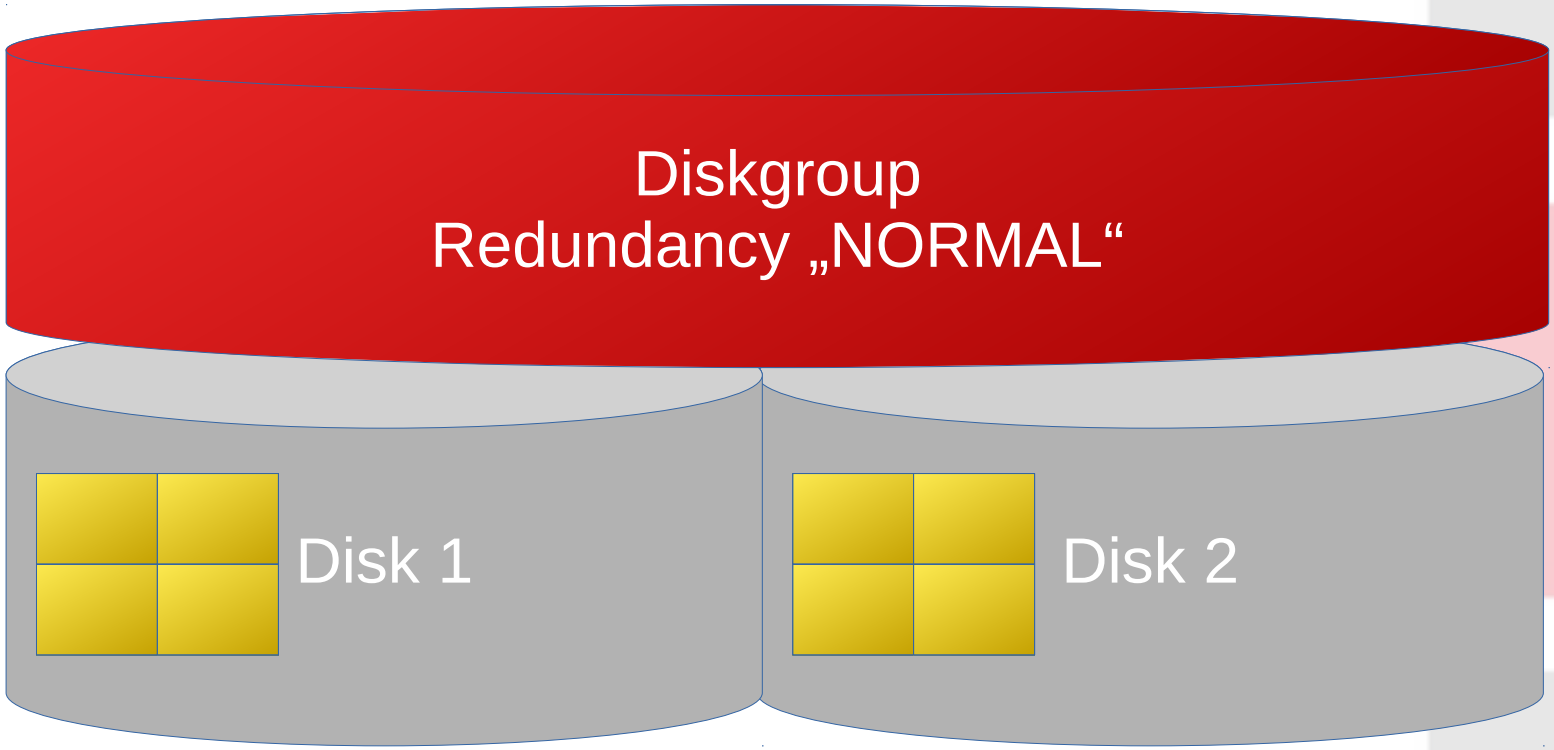
performing databases
Your reliability. Our concern.



Disk 1 und 2 gleich groß wählen!

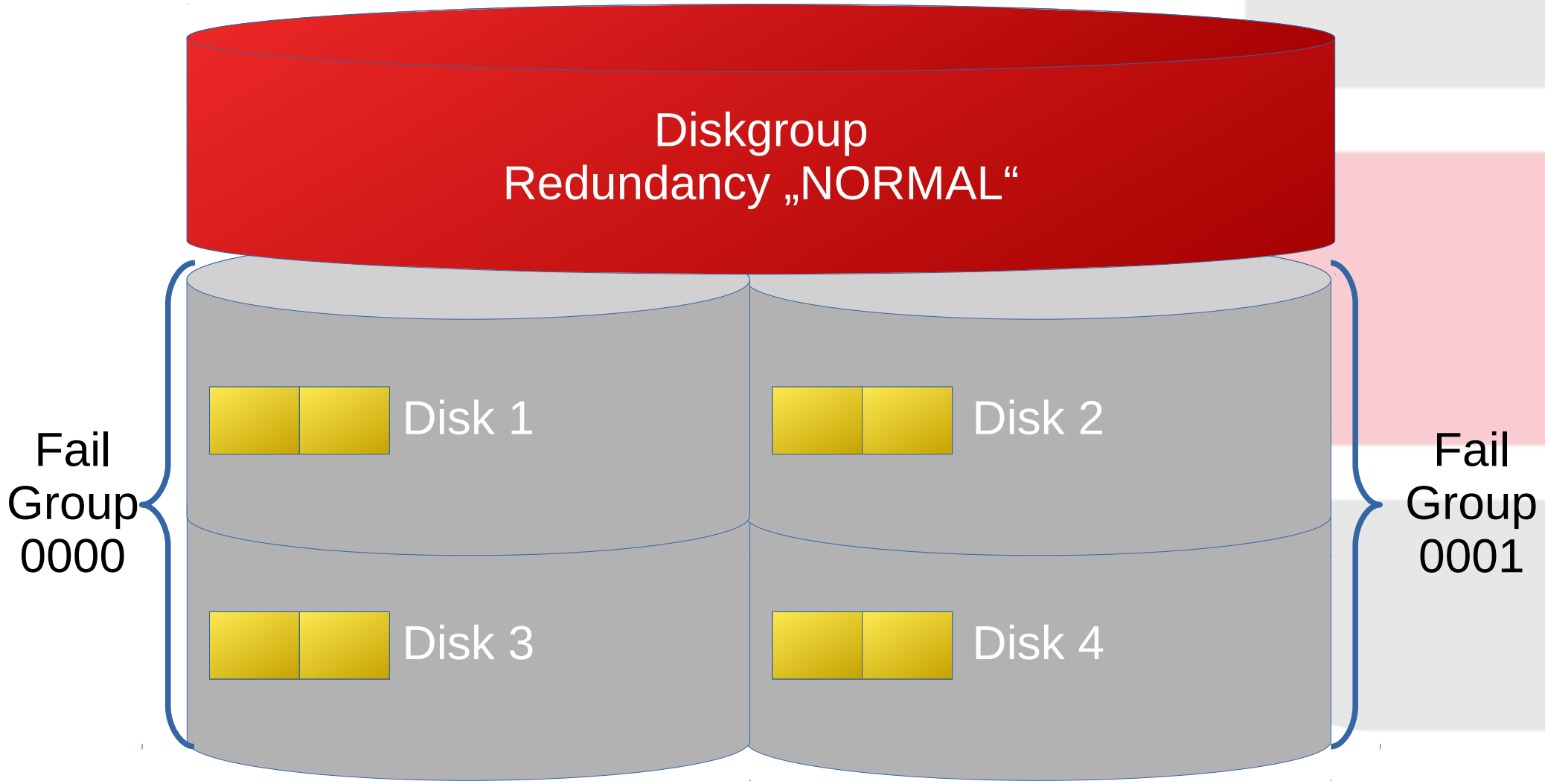
ASM Diskgroups

performing databases
Your reliability. Our concern.



ASM Diskgroups

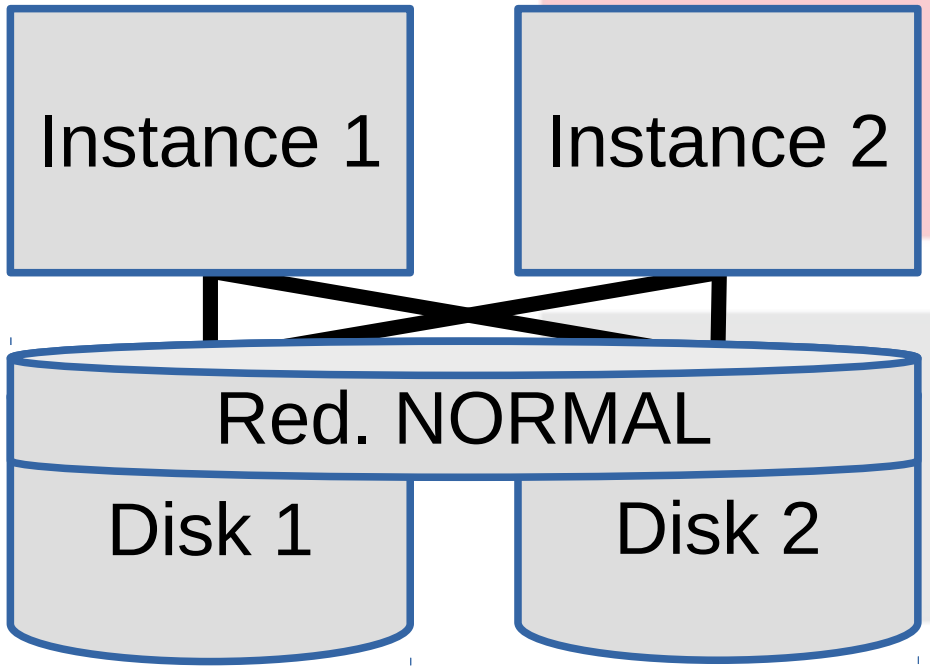
performing databases
Your reliability. Our concern.



ASM

performing databases
Your reliability. Our concern.

- Preferred Read Node (11g)
  erlaubt „local first“
- Even Reads (12c)
  liest regelmäßig von allen
  Disk-Spiegeln der DG



ASM Erfahrung

performing databases
Your reliability. Our concern.

- In >12 Jahren kein Datenverlust
  (aber es war knapp)
- Größtes Problem in 10.x
  Kurz „abwesende“ LUNs nicht dropped + rebalanced
- „Lösung“ mit 11.2
  - Event verbessert
  - DISK_REPAIR_TIME (drop delay)
- Größte Herausforderung
  Wiederkehrende LUNs erfordern manuellen Eingriff
  => Monitoring!

ASM 12.x

performing databases
Your reliability. Our concern.

- Flex Diskgroup
  - erlaubt Gruppierung nach Datenbank innerh. DG
  - ermöglicht Split-Mirror Sicherung pro DB innerh. DG
- ASM Service
  - zentralisiert das Management von Shared Disks für alle RACs einer Domäne
  - Zentraler Shared Disk Service („ASM über TNS“)
- Hoffentlich nicht nur für Exa!

performing databases
Your reliability. Our concern.

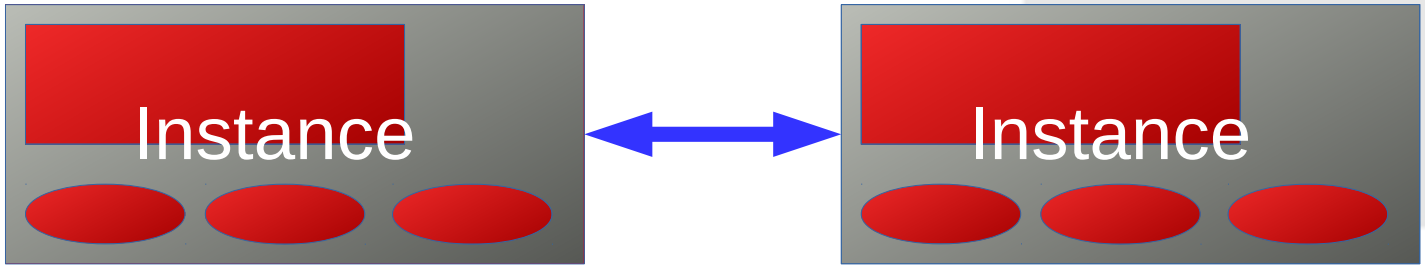
Clusterware

Oracle Clusterware

- alias OCW
  Grid Infrastructure (GI)
  Cluster Ready Services (crs)

- Clusterware, z.B. Failover von Services

- APIs und Infrastruktur für
  - Datenbank (CacheFusion, …)
  - ASM (FlexASM, …)
  - ACFS
  - ...



Dienste

performing databases
Your reliability. Our concern.

- OHASD - Oracle High Availability Service Deamon
  Starten und Überwachen der Clusterprozesse
  Schnittstelle zum Dienstemanagement des OS

- EVM (evmd) - Event Manager
  Ein- und ausgehende Benachrichtigungen zwischen Nodes und Diensten
  (denke: Message Bus)

Dienste

performing databases
Your reliability. Our concern.

- CSS (ocssd) - Cluster Synchronization Service
  Verwaltung Member Nodes / Erreichbarkeit
  Disk/Network Heartbeat
  Logging: ocssd.log / trc (12c)

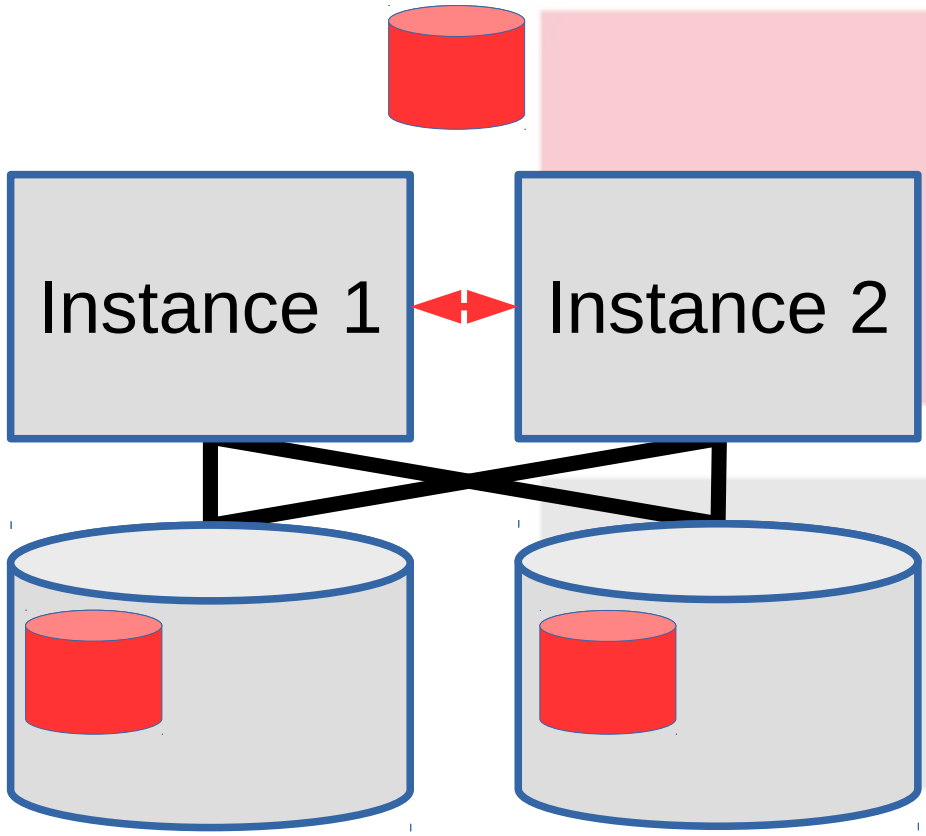
- CRS (crsd) - Cluster Ready Service
  Steuern und Verwalten der „User“-seitigen Dienste (z.B. DB, Listener)
  Logging: crsd.log / trc (12c)
  - OraRootAgent - steuert privilegierte Dienste
  - OraAgent - steuert andere Dienste
    (Vorsicht: beide Namen auch beim OHASD)
  - Agent Logs liefern Applikations-Output

H.A.

performing databases
Your reliability. Our concern.

- Heartbeat Node Fencing
  - Bisher: Niedrigste Node-Nummer
  - 12.2: Niedrigste Anzahl Services oder Prio-Service

- Split-Brain-Problem
- Quorum: 1,3,5,7 …



Planung vs. Trouble

performing databases
Your reliability. Our concern.

- Infrastruktur zur Zeitsynchronisation LEBENSWICHTIG (Node Evictions)

- bevorzugt NTP (-x)
- Notfalls CTSS nicht als Observer
  => (Node mit Node)

- Interconnect
  - Latenz minimieren - jedoch mindestens Switched
  - Jumbo Frames
  - Multicast
  - Isolieren!

Cool & CLI

performing databases
Your reliability. Our concern.

- srvctl -h | grep xyz
  - Möglichst alles mit srvctl
  - UNBEDINGT aus GRID-Home starten! ($OH+$PATH)
  - Kein ADD für Diskgroup Services wird beim ersten Mount angelegt ;)
  - Volle Unterstützung für Dataguard (mit DG Broker)
- crsctl status resource -t
- crsctl -unsupported (z.B. entfernen DB Service ohne $OH)

OCW und SE*

performing databases
Your reliability. Our concern.

- Grid Infrastructure als Clusterware für Failover Cluster DB Failover ohne RAC One Node!
- Kein Linken von DB RAC Binaries
  => Standard Edition, SE1, SE2
  => 10-Tage-Regel möglich (1 Storage)
- Kein Linken von DB RAC Binaries
- 1-Node-Serverpools einrichten
- DB Service für Single Instance mit Server_pools=* anlegen (crsctl -unsupported, but works!)

performing databases
Your reliability. Our concern.

Non-RAC

Schon gewusst?

performing databases
Your reliability. Our concern.

- $TWO_TASK konkatiniert Inhalt mit @ an sqlplus
- Customized SQL prompt (glogin.sql) stört Utilities (dbca, asmca)
  z.B. *+SQL+>*
- ORACLE_HOME darf nicht mit / enden (hash Vergleich)
- DB_DOMAIN darf keinen Bindestrich enthalten (div. Probleme mit DB Links)

RMAN & Service

performing databases
Your reliability. Our concern.

- RMAN restore from service name (12c)

  RMAN> restore datafile 5 from service STANDBY;

- Funktioniert mit Datafiles, Control Files
- Leider NICHT mit Archived Redo Logs :(

Umgebung

performing databases
Your reliability. Our concern.

- Oracle: Unterschiedliche OS User nutzen
- Mein Tip: Vermeiden.
  Umgebungsskripte für installierte Produkte anlegen
- Vorschlag: /home/oracle/bin
- Setzen Umgebung UND Prompt
  z.B. [oracle@myhost ~] (MYSID) $
- Einbinden nach logon über source
  . db oder . grid

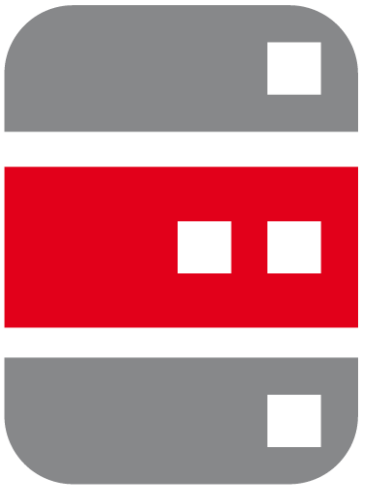

Heads in and not LOOKING OUT...
ETPS
"I say, that's a jolly interesting reading"
SPARKY
BONES
...can seriously damage your health
Maintain constant LOOK-OUT at all times
However interesting the data may be!
Defence Aviation Safety Centre Poster No 200513

performing databases
Your reliability. Our concern.

Q & A

Download my Presentations and Whitepapers
http://www.performing-databases.com

performing databases



Your reliability. Our concern.