

# MySQL Cluster

Kai Voigt  
MySQL AB  
kai@mysql.com

MySQL Meetup Day  
Hamburg, 6. März 2006

# Agenda

- Warum?
- Wie?
- Wie genau?
- Was sonst?

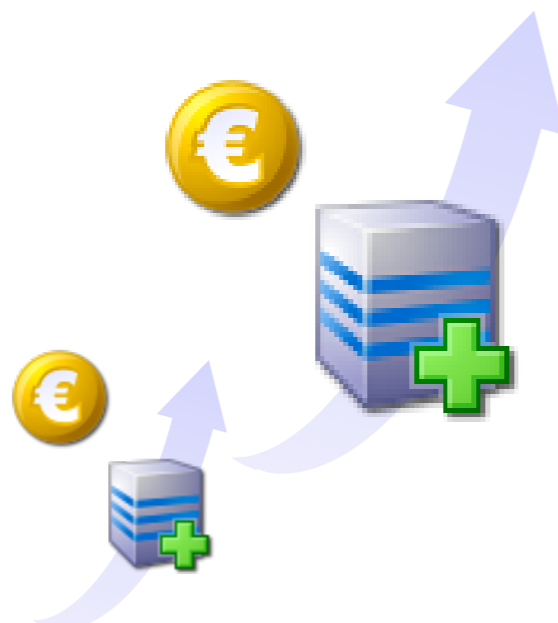
# Warum?

# Kosten runter

- Hochverfügbarkeit (99,999%)
  - Redundante Daten und Systeme
  - Wiederherstellung
- Skalierbarkeit
  - Viele Anfragen pro Sekunde
  - Standard-Hardware

# Skalierungen

- Vertikale Skalierung
  - teure SMP-Hardware
  - komplizierte Erweiterung
- Horizontale Skalierung
  - Standard-Hardware
  - Einfache Erweiterung



# Einsatzbereiche

- Web-Applikationen mit hoher Last
  - Portale
  - eCommerce
- Zeitkritische Anwendungen
  - Banken
  - Telekommunikation

**Wie?**

# Architektur

- Verbindungsebene
- SQL-Ebene
- Storage-Ebene
  
- Zukunft: Pluggable Modules



# Architektur

## MySQL-Server




### Verwaltungs- dienste & Werkzeuge

Datensicherung & Wiederherstellung, Sicherheit, Replikation, Cluster, Administration, Konfiguration, Migration & Metadaten


### Connection Pool

Authentifizierung, Wiederverwendung, Connection Limits, Speichertest, Caches




### SQL-Schnittstelle

DML, DDL, Stored Procedures, Views, Triggers, usw.



### Parser

Abfrage-Übersetzung, Objektprivilegien




### Optimizer

Zugriffspfade, Statistiken



### Caches & Buffers

Global und spezifisch für Engines



### Austauschbare Speicher-Engines

Verwaltung von Arbeits- und Festplattenspeicher sowie Indizes




### Datei-System

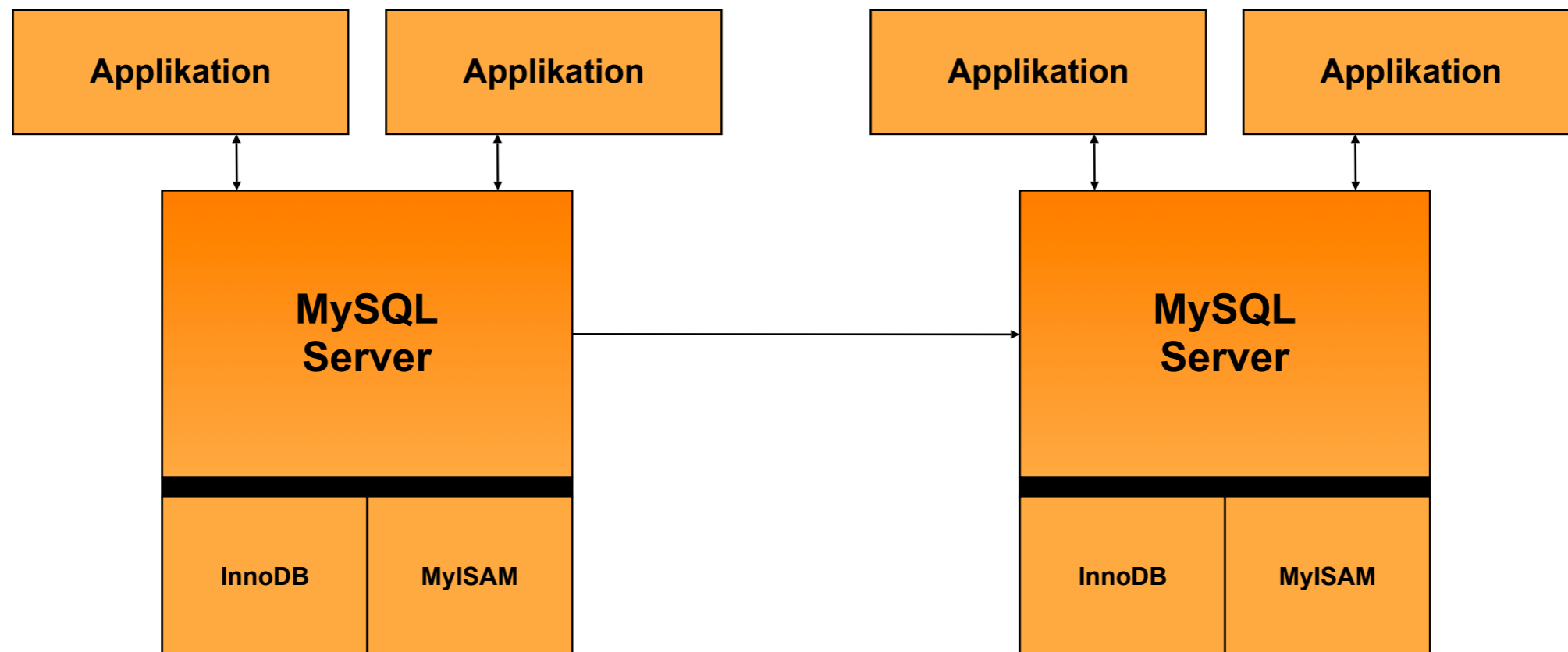
NTFS - NFS  
SAN - NAS

### Dateien & Log-Dateien

Redo, Undo, Daten, Index, Binary, Error, Query und Slow

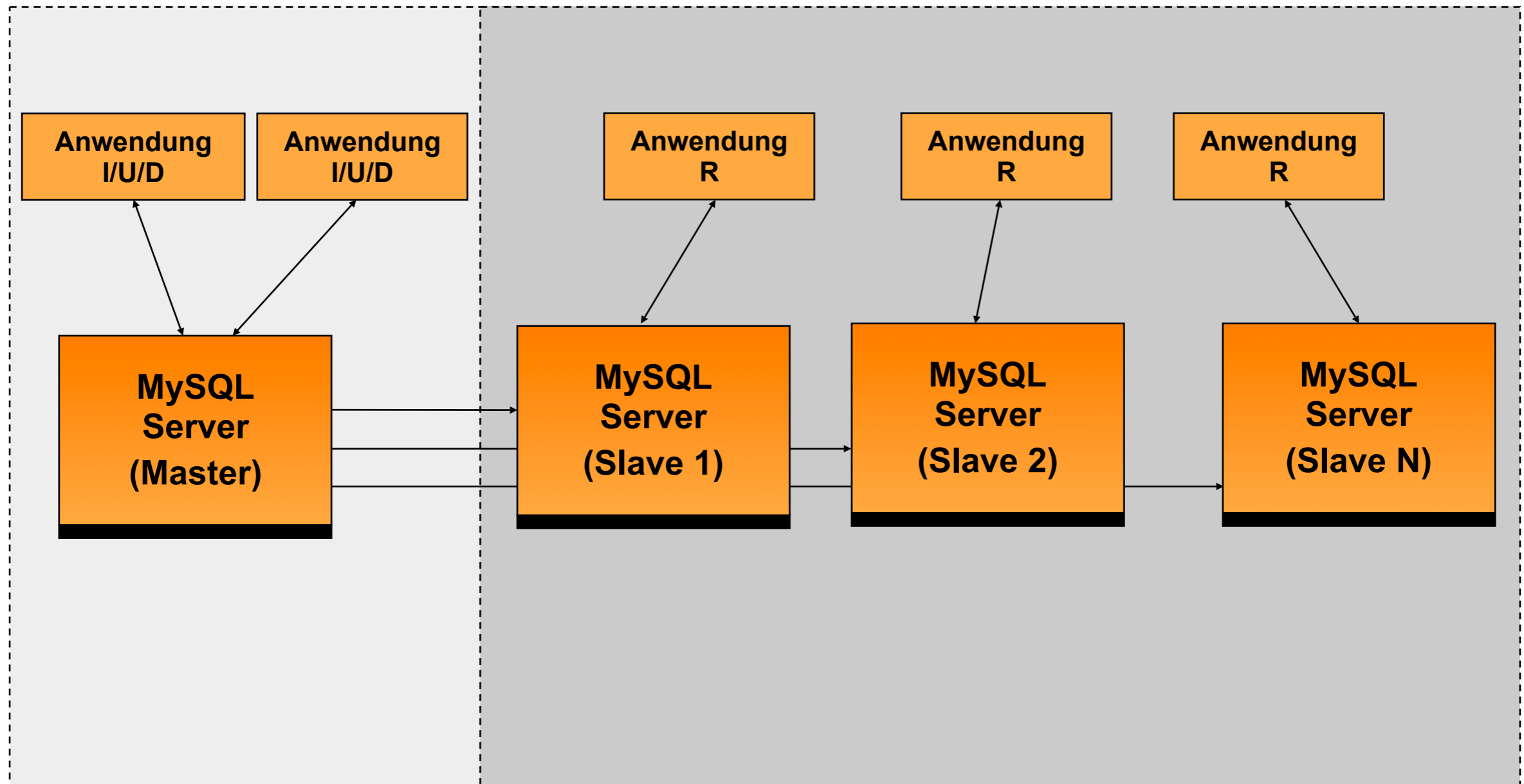


# Replikation

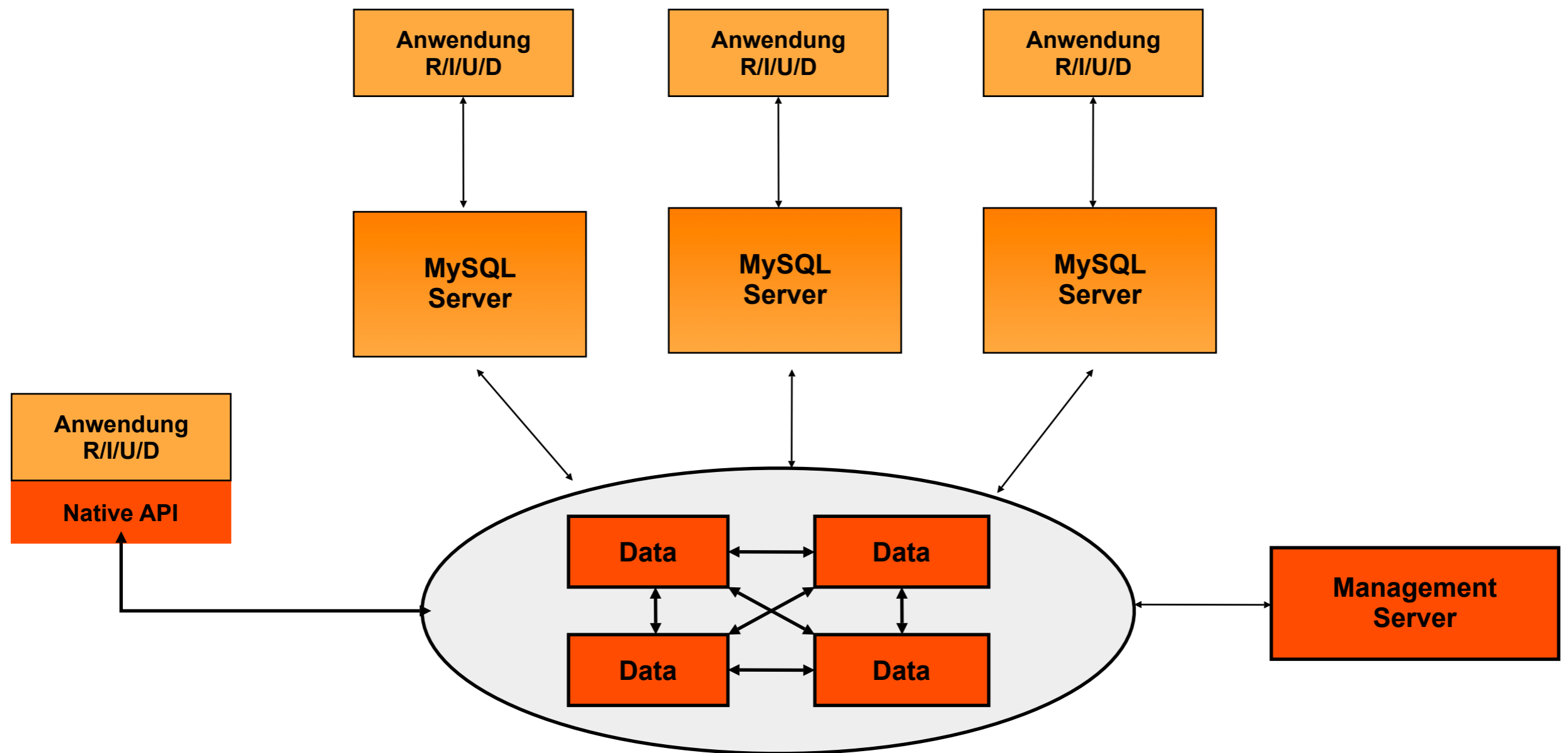


- Asynchrone Datenkopien

# Replikation

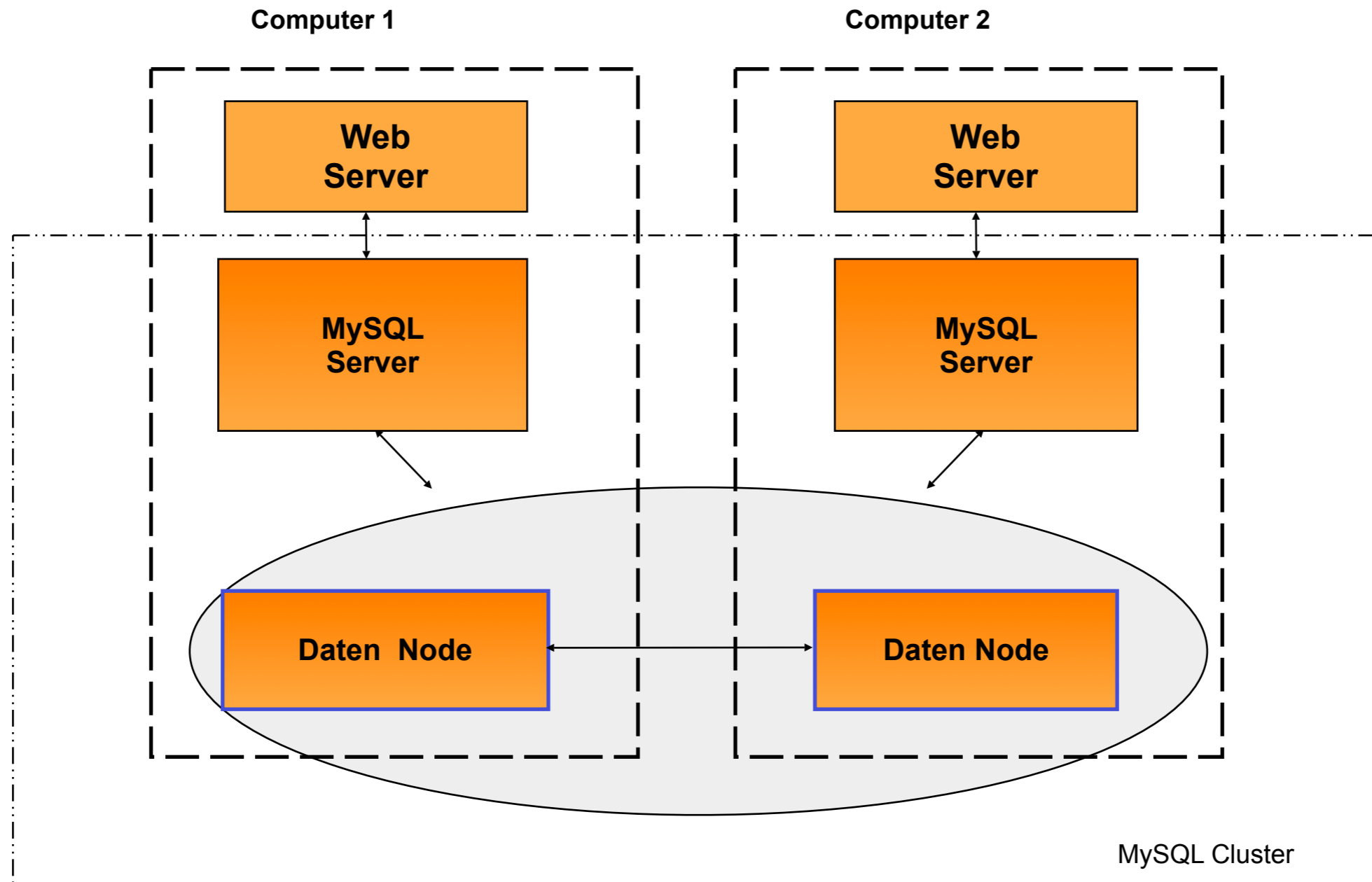


# Cluster-Architektur

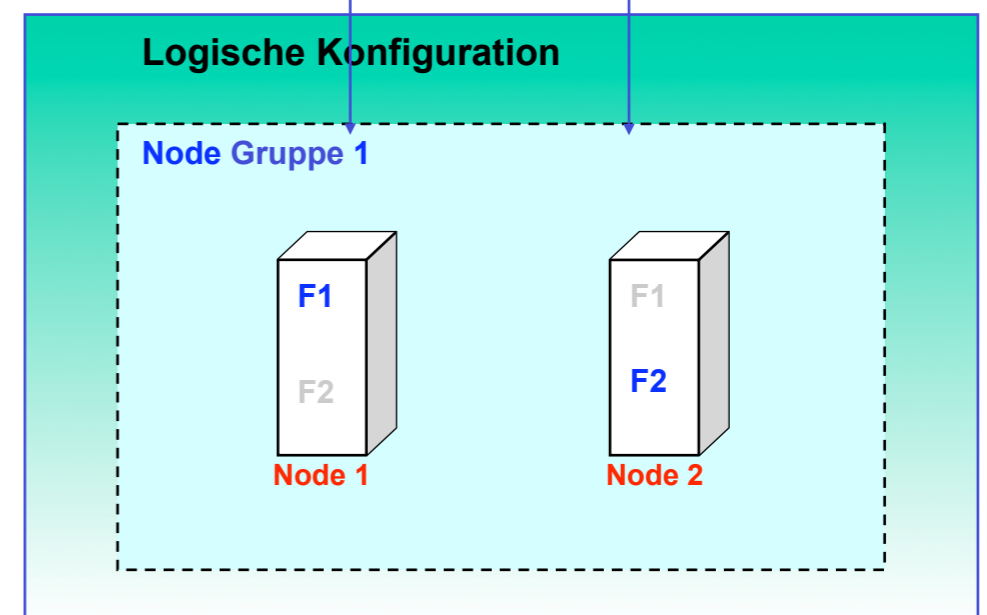
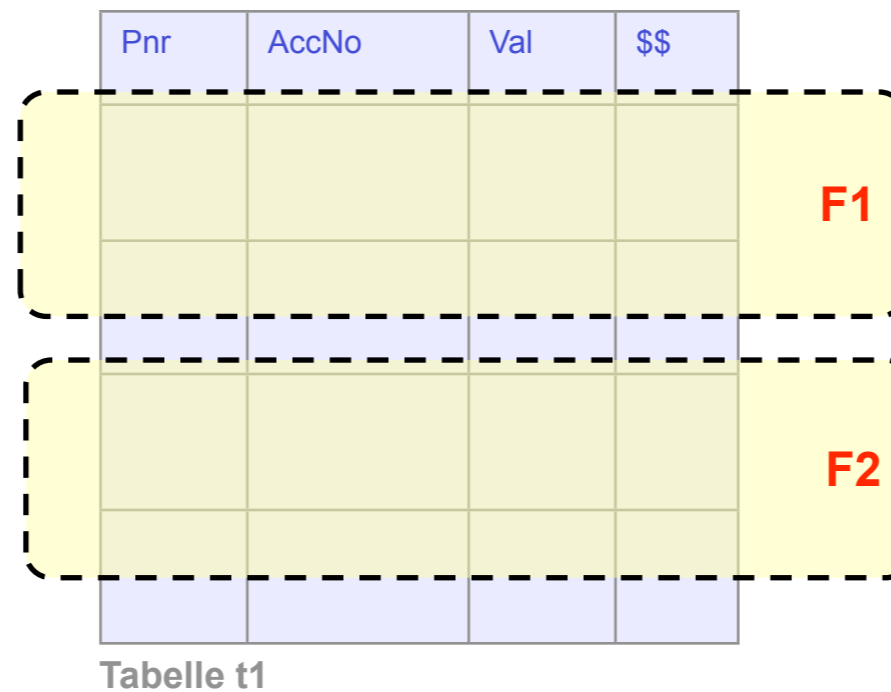


# Wie genau?

# Minimal-Konfiguration



# Daten-Verteilung



2 Kopien der Daten

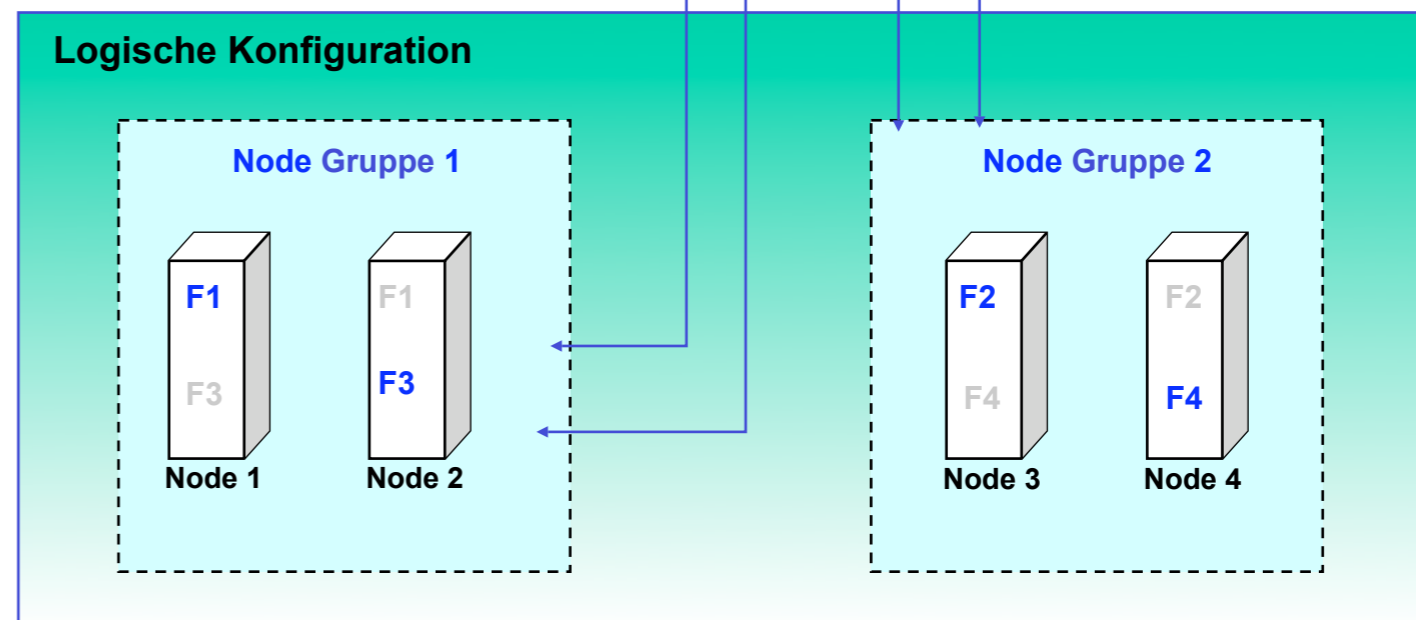
Fx – primäre Replik

Fx – sekundäre Replik

# Daten-Verteilung

Pnr	AccNo	Val	\$\$	
				F1
				F2
				F3
				F4

Tabelle 1



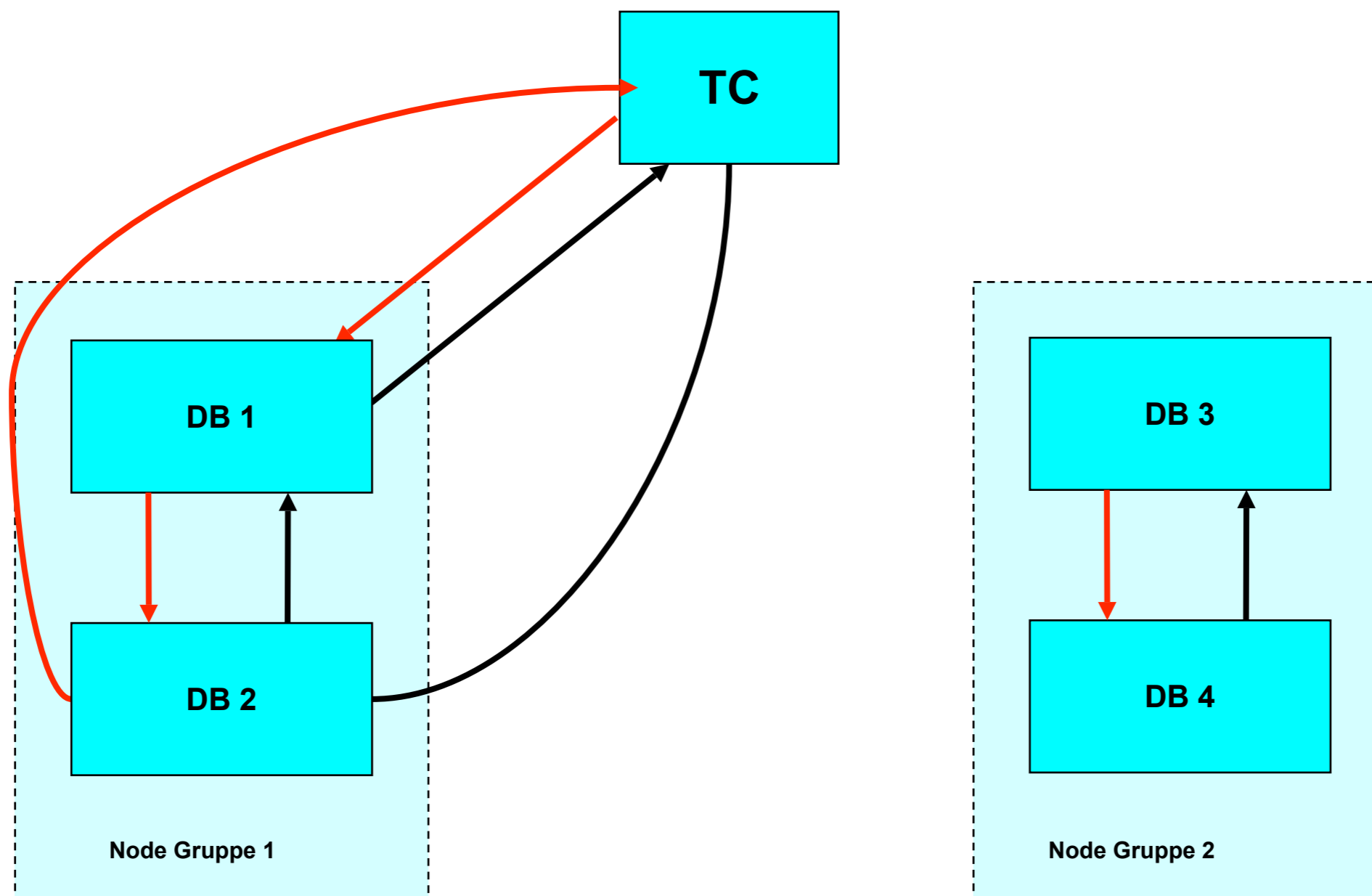
2 Kopien der Daten

Fx – primäre Replik

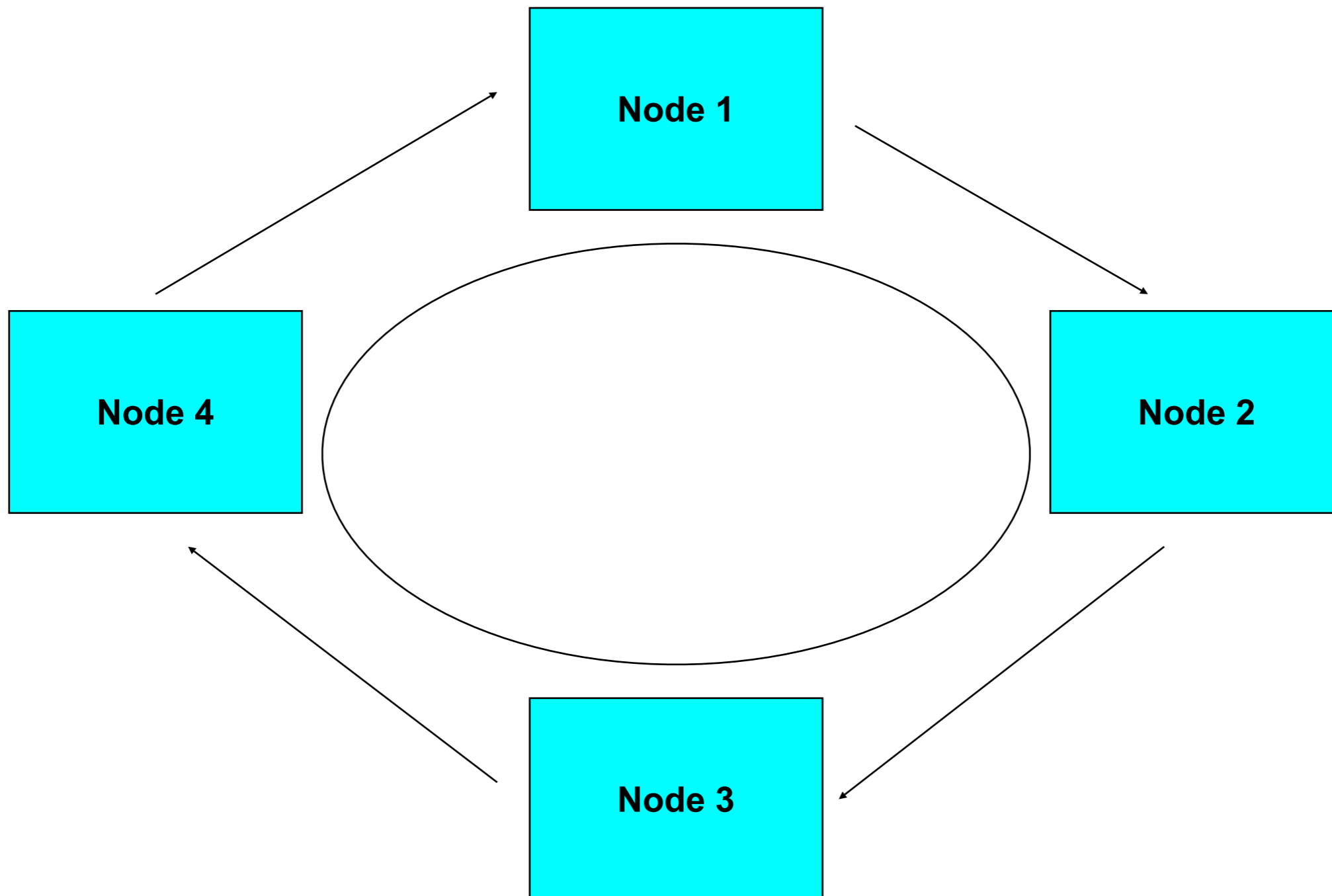
Fx – sekundäre Replik



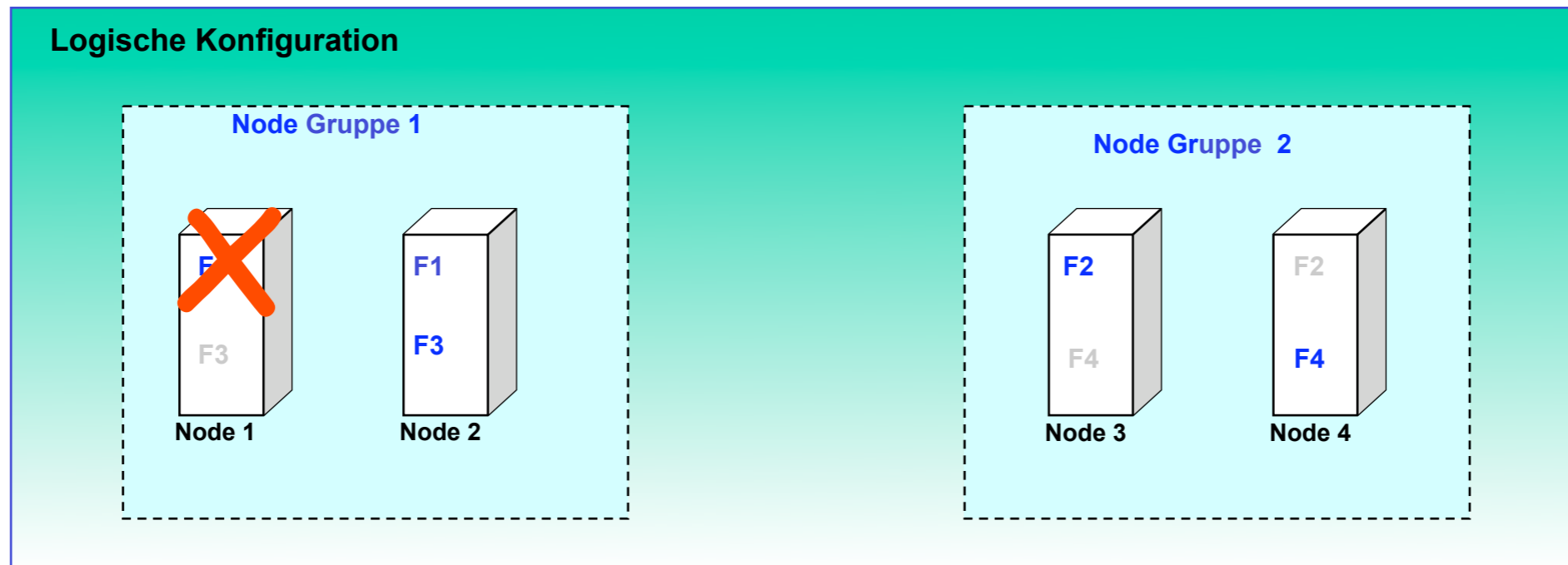
# Two-Phase Commit



# Heartbeat

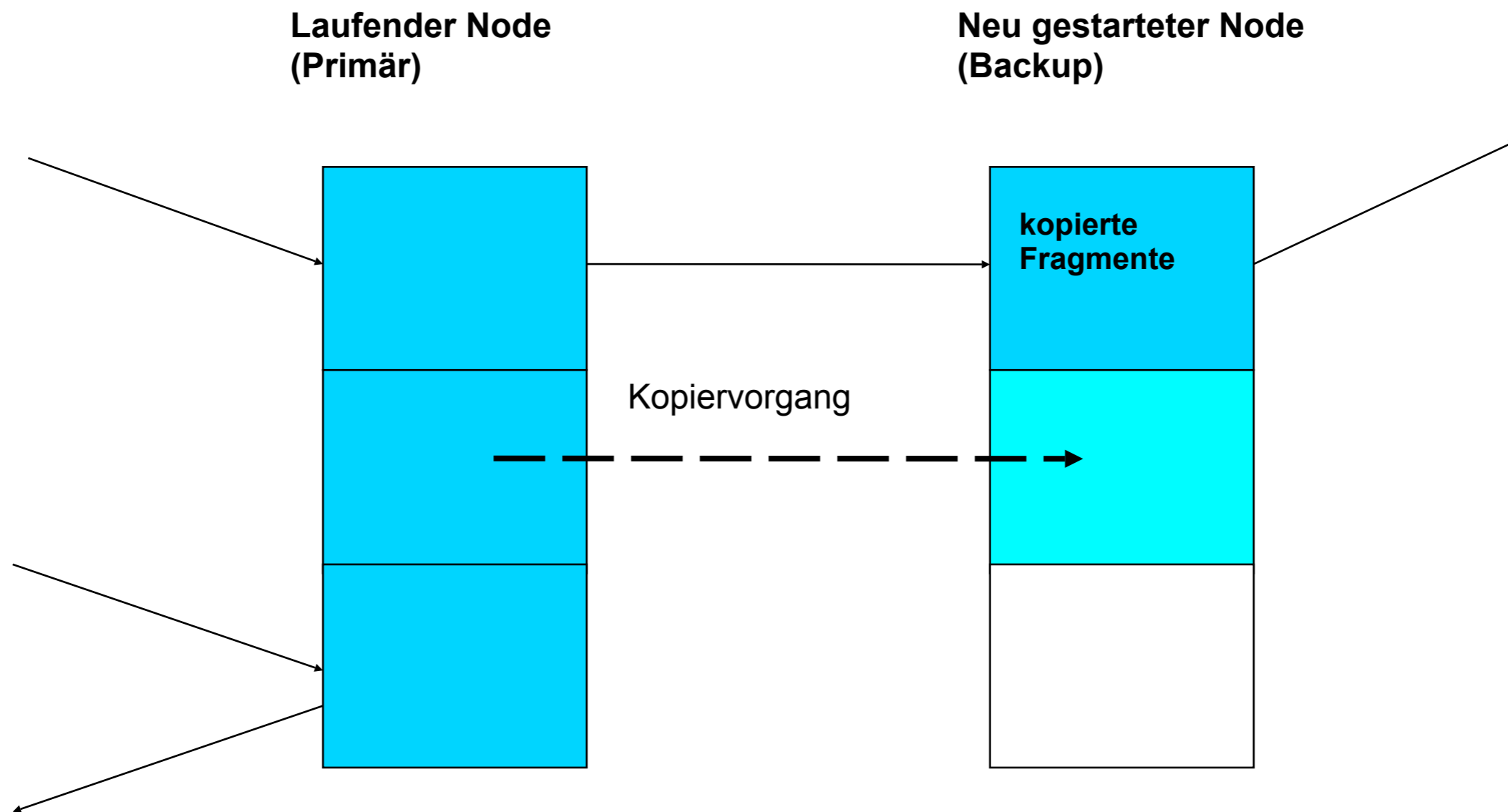


# Knoten-Ausfall

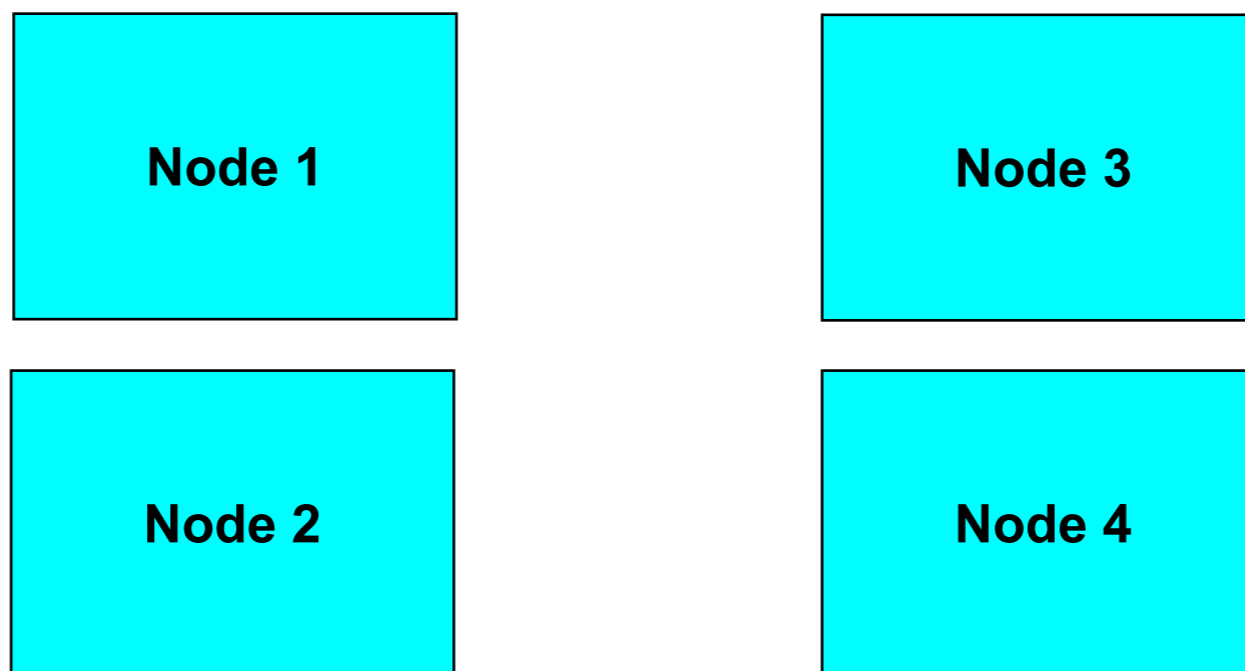


- Erkennung durch Heartbeats
- Übernahme durch Sekundärfragment
- Automatische Wiederherstellung

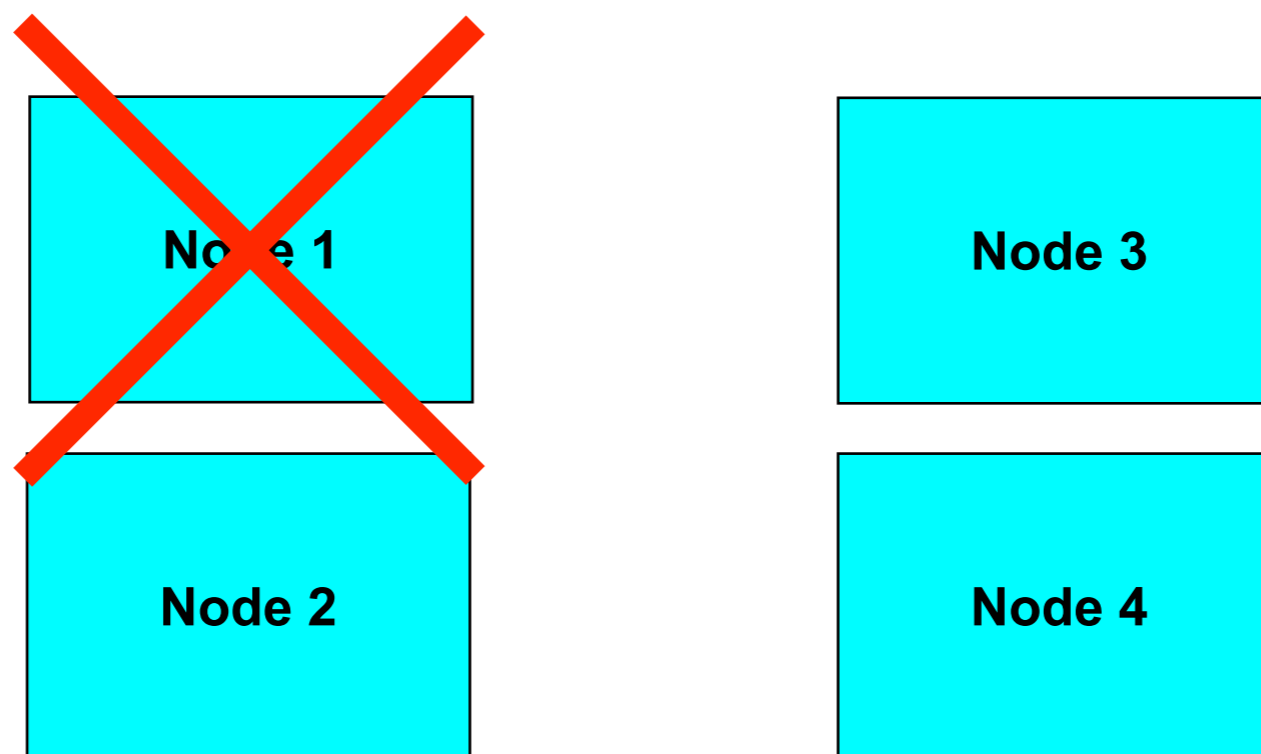
# Recovery



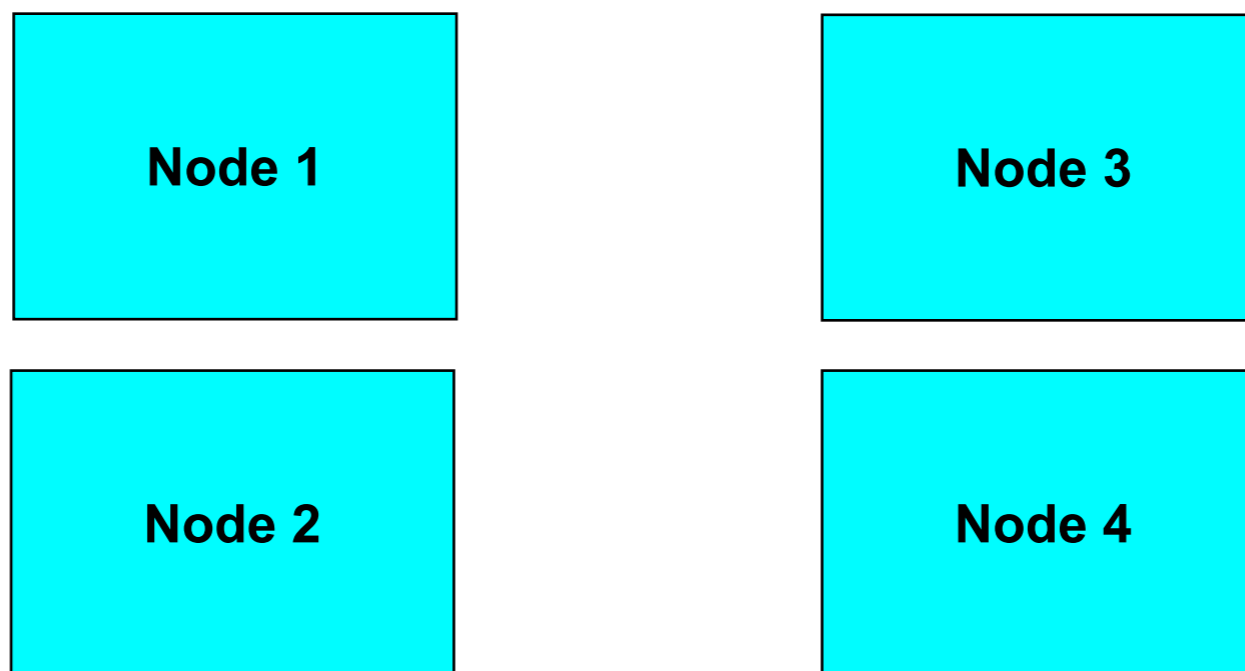
# Network Partitioning



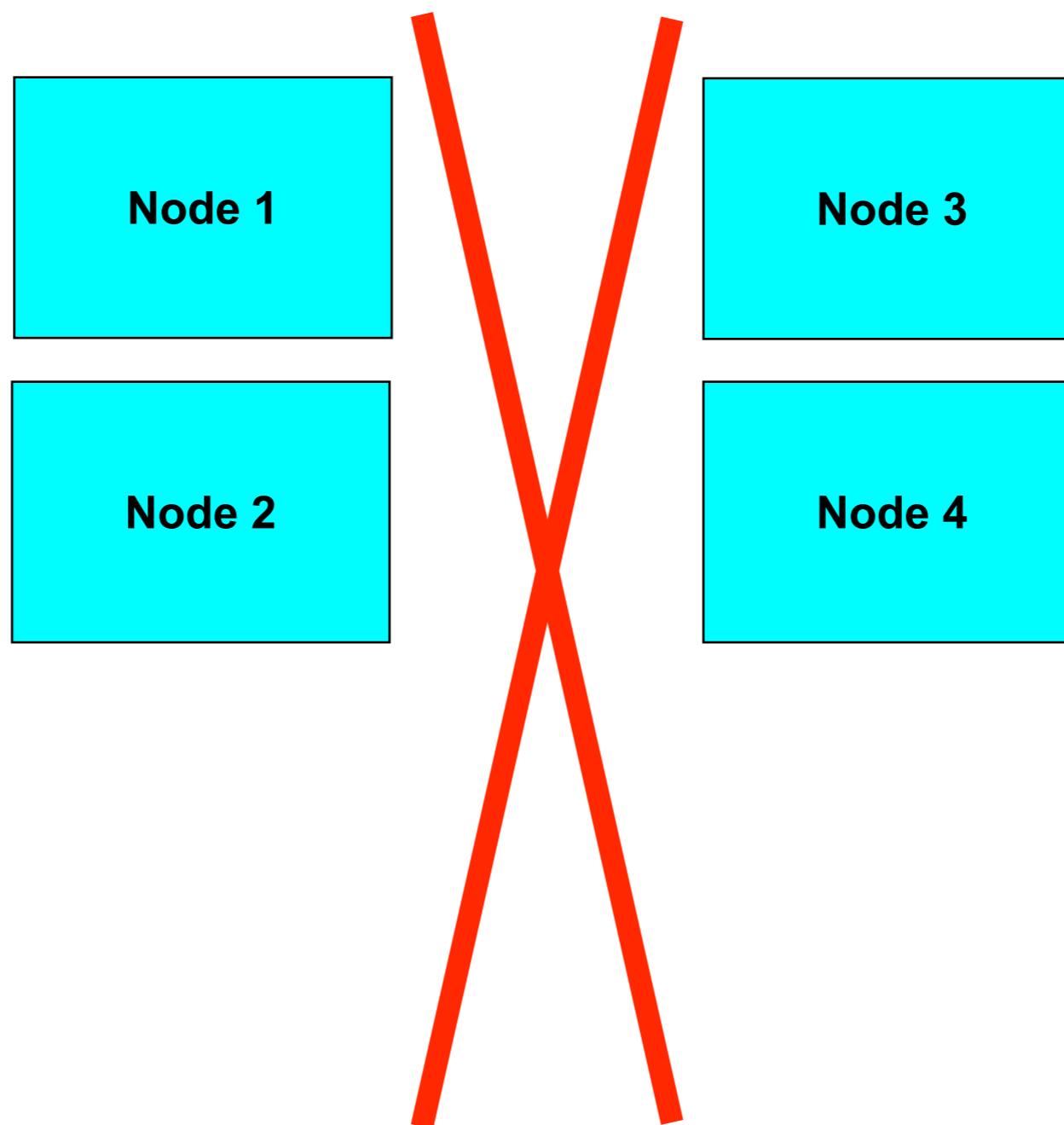
# Network Partitioning



# Network Partitioning

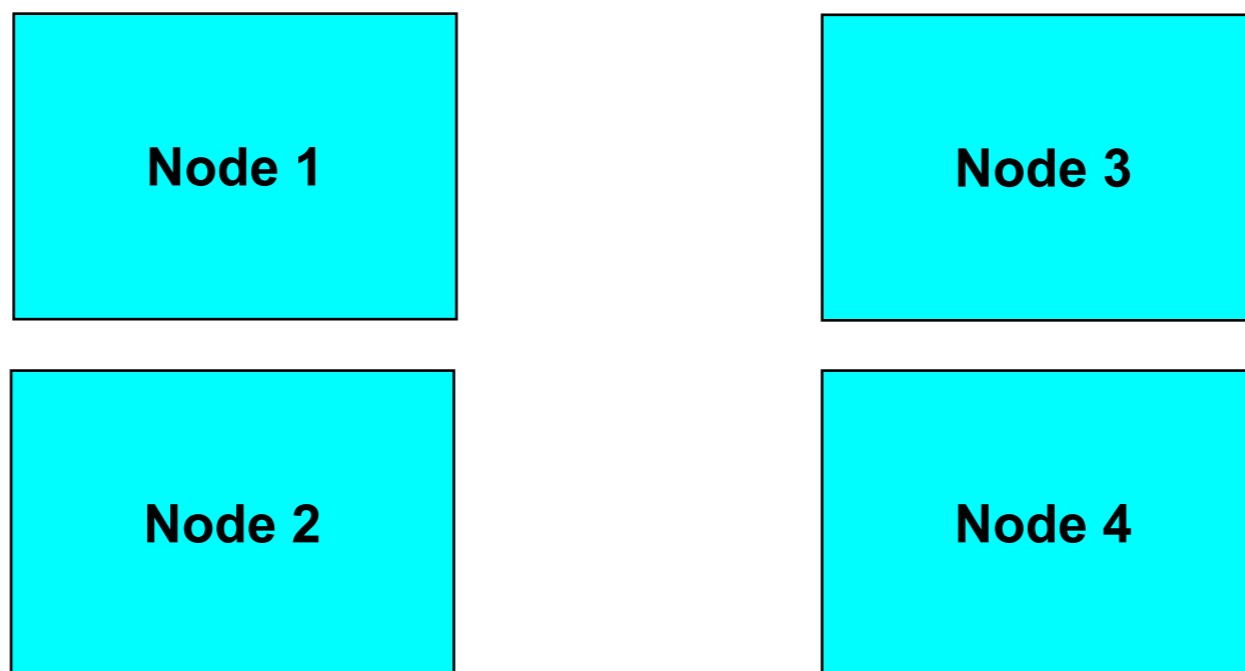


# Network Partitioning

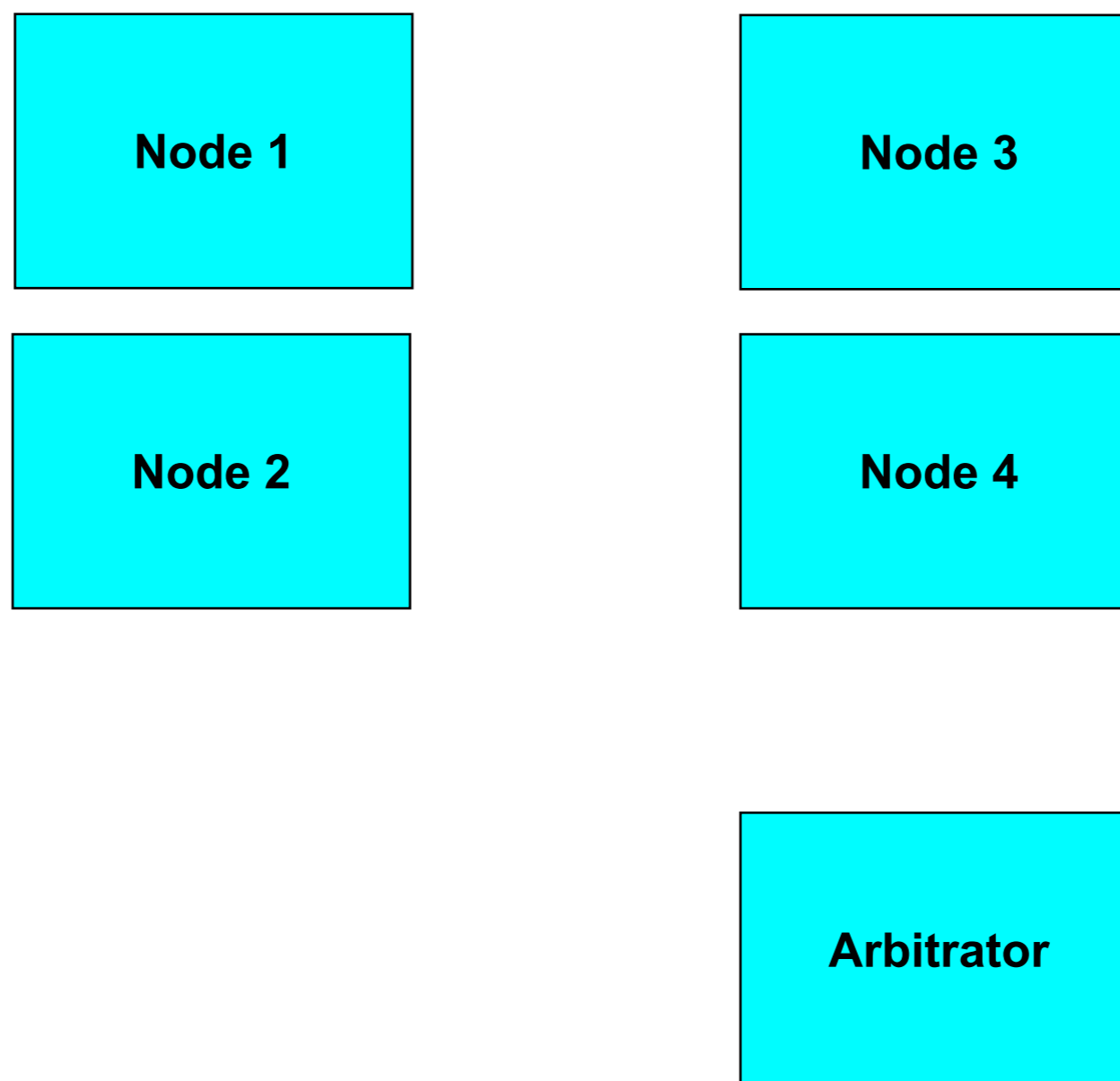




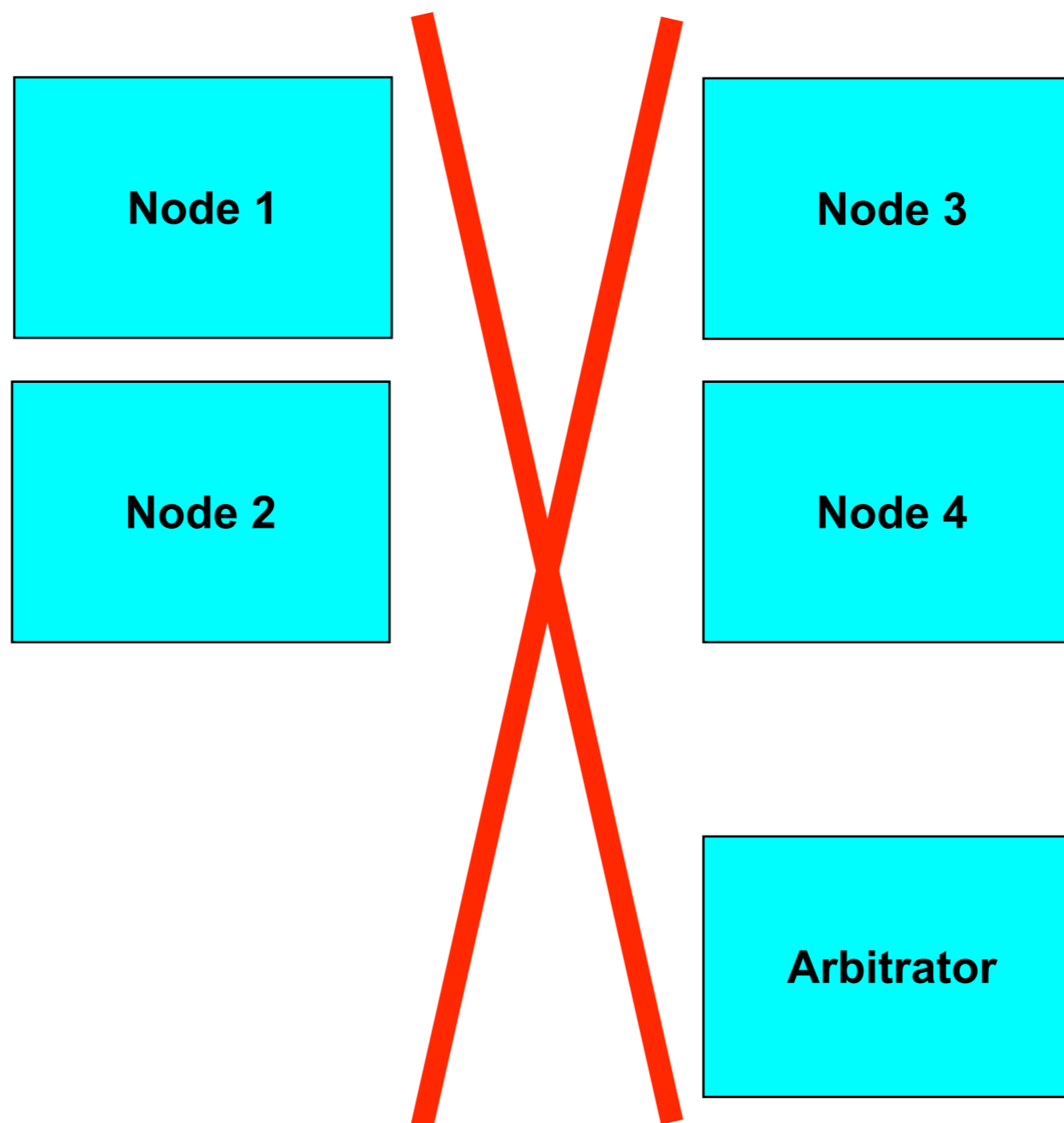
# Network Partitioning



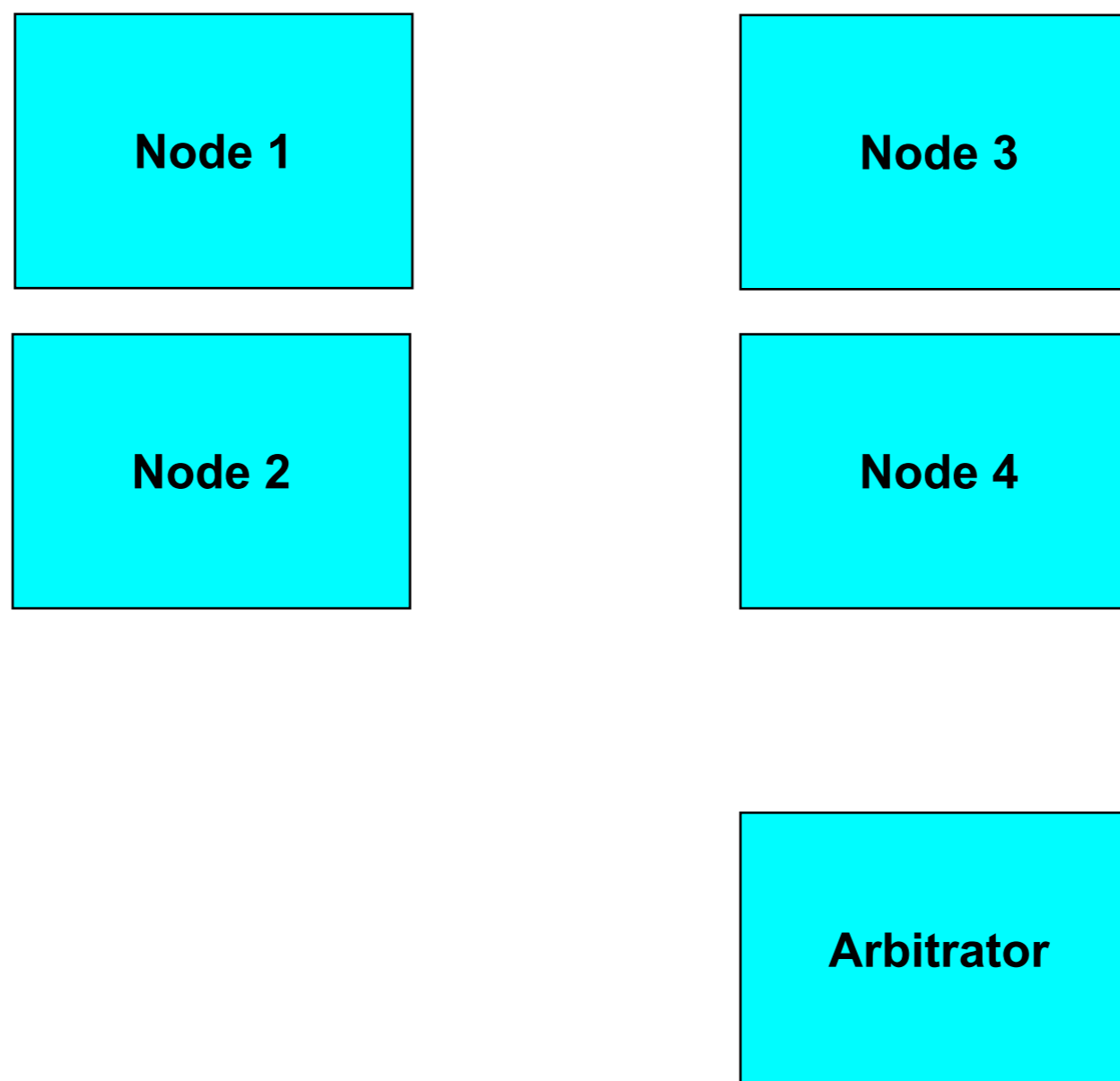
# Network Partitioning



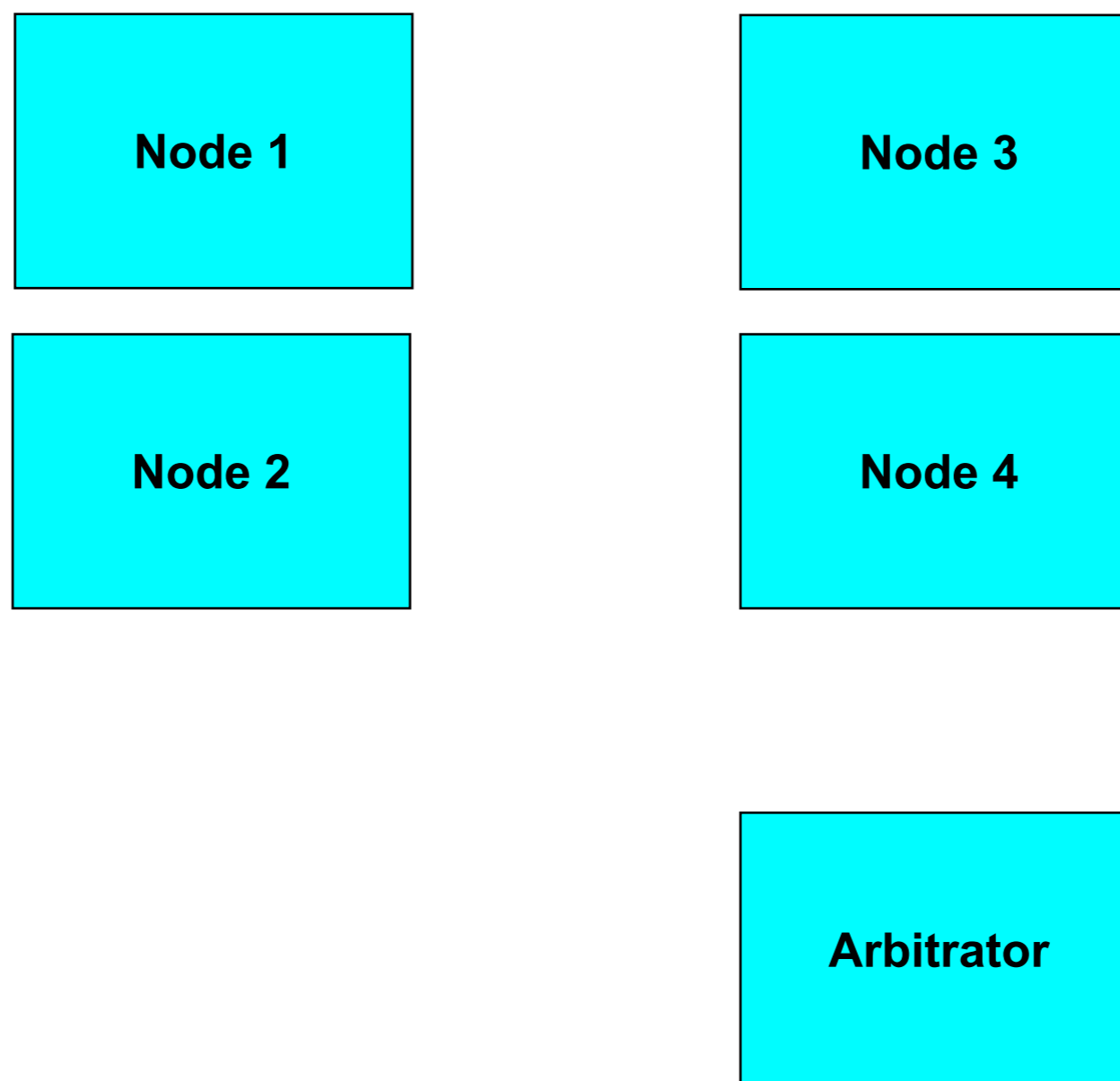
# Network Partitioning



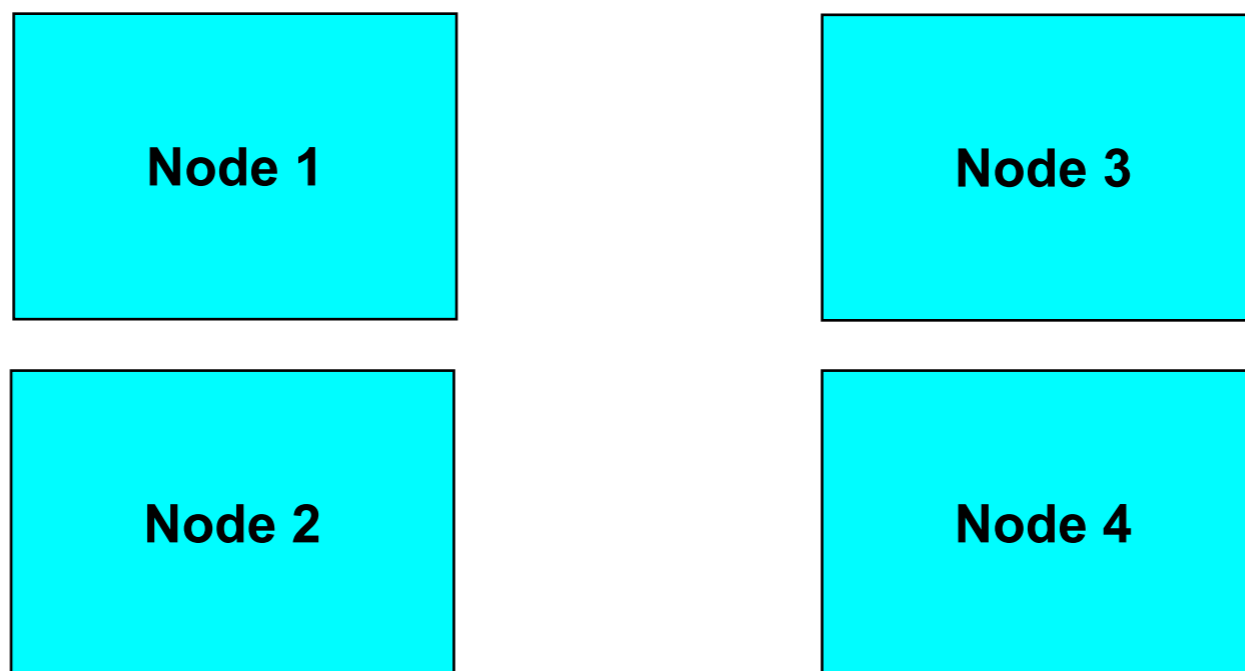
# Network Partitioning



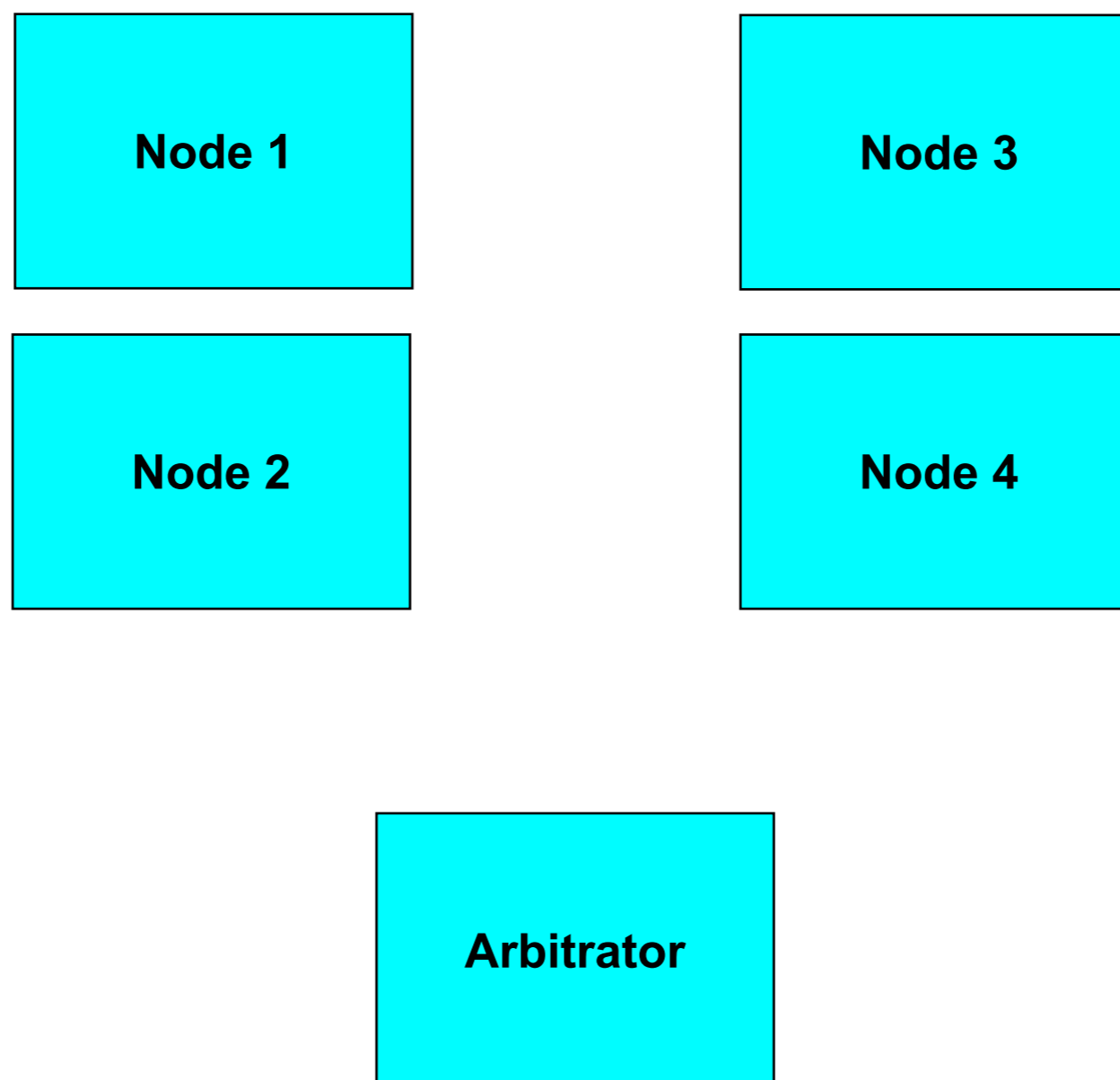
# Network Partitioning



# Network Partitioning



# Network Partitioning



# Was sonst?



# Weitere Funktionen

- Management-Client
- Online-Backup
- Online-Restore
- Online-Upgrade

# Begrenzungen

- Daten müssen ins RAM passen
- Keine Online-Skalierung
- Kein VARCHAR

# Neu in MySQL 5.1

- Daten auf Festplatte
- Definierbare Partitionierung
- Variable Zeilengrößen
- Integration in Replikation

# Mehr Infos

- White Papers
- Downloads: MySQL 5.0.16-max
- Dokumentation
- Mailing-Listen
- Forum

# Danke

<http://www.mysql.com/>