

linuxwochen

PETER-PAUL WITTA



Clustering mit Shared Storage

Ing. Peter-Paul Witta
paul.witta@CUBiT.at

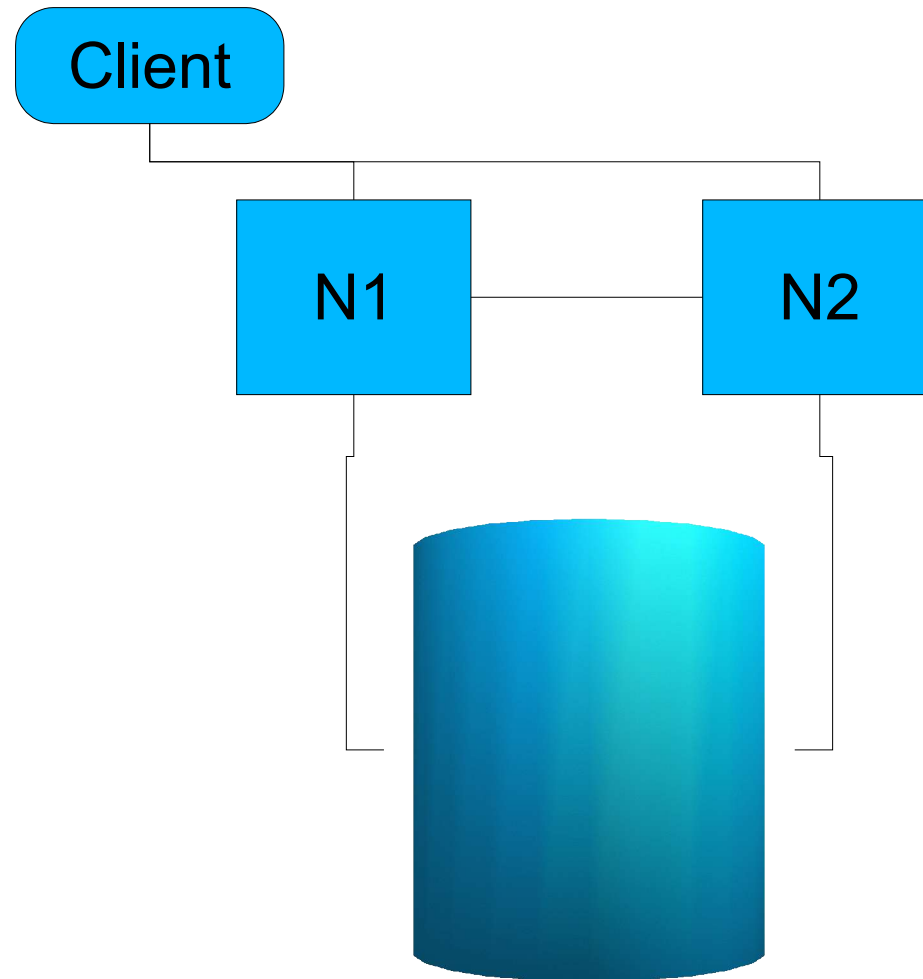


Clustering

- mehrere kleine Rechner leisten gemeinsam Grosses
- günstige dual intel/AMD Server
- load sharing
- High Availability
- combined architecture
- centralized management



Clustering (2)



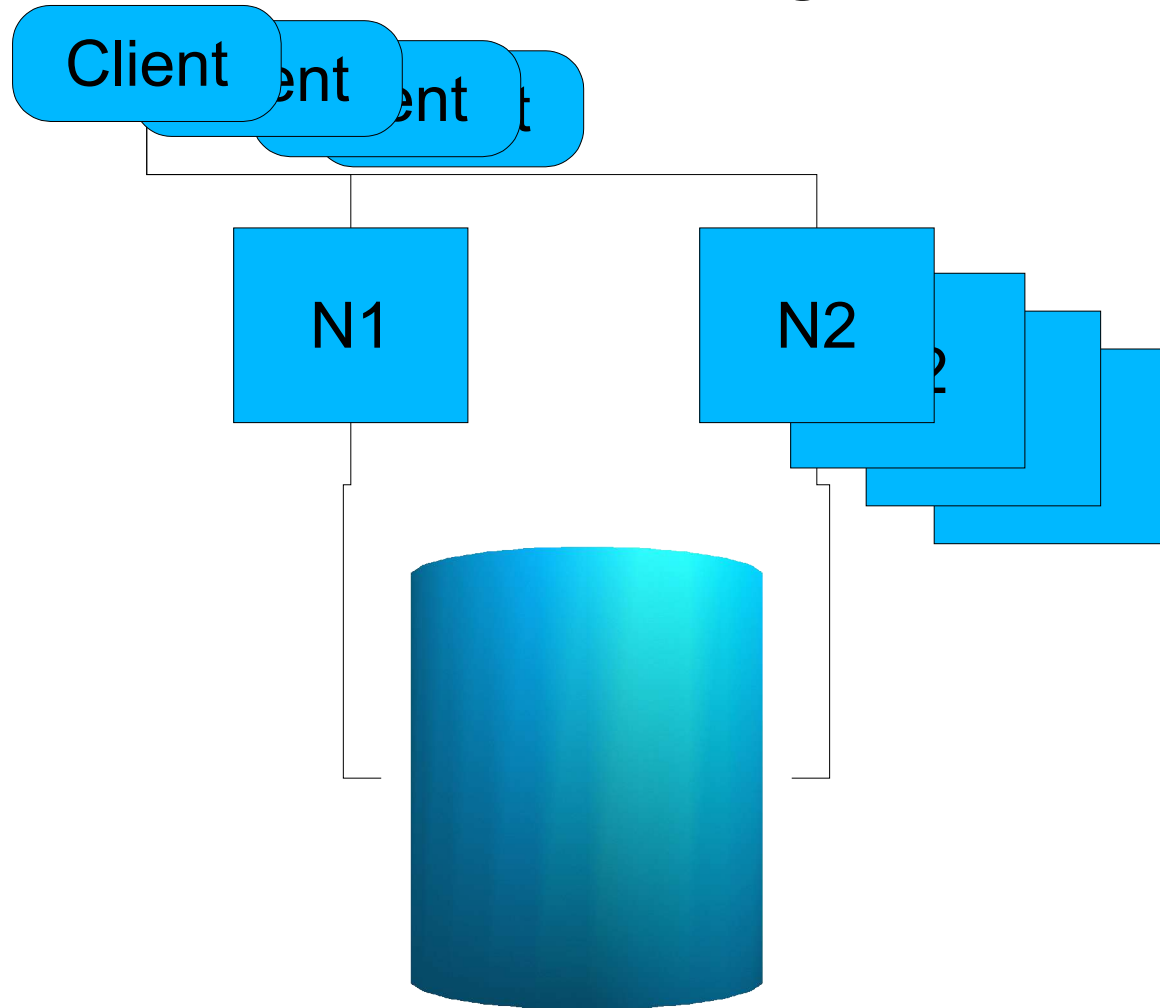


High Availability

- Ausfallsicherheit durch Redundanz
- duale Server
- jeder “backt” den Anderen
- pro Dienst nur ein Server aktiv
- hohe Leerkosten durch Mehr-Hardware
- Bei Ausfall Migration der Dienste auf Ersatzsystem
- cold standby
- hot standby
- active/active
- Umschaltzeiten
- Session-Verlust(?)



Load Sharing





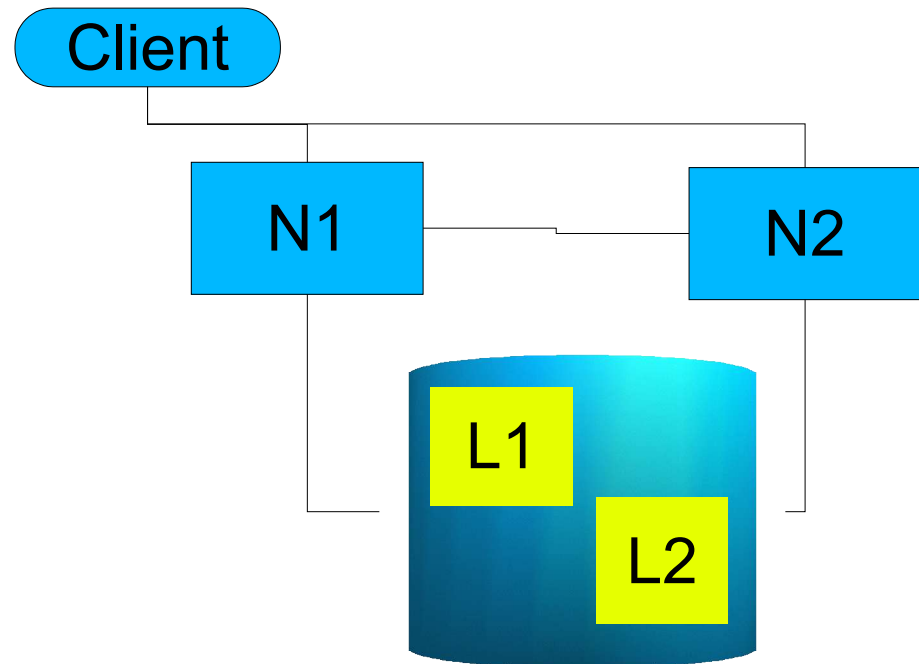
Architecture (lg concept)

- load group=(service+Ressourcen)
- = (service+network+disk)
- enthält alle Ressourcen für Service
- unteilbar, kann nur als Ganzes einem Server zugeordnet werden
- mehrere lg pro Server möglich
- mehrere Services pro lg möglich



Combined (virtual image)

- Dienste nur in Ig
- Zuordnung problemlos verschiebbar bei Performance-Problemen
- Failover problemlos möglich
- Shared-Storage active/active möglich





Storage

- Multihost-fähig
- FC/AL vs. SCSI
- Terminationsprobleme bei SCSI
- Achtung: Multipath ist schön, aber schwierig
- Qlogic-FC/AL Probleme
- Disk-Probleme, Locking, Dual-Mount
- FC-Switching



Applications

- Anwendungen teilweise Cluster-fähig
- mit heartbeat: Konfiguration
- Metadaten mitnehmen bei Failover
- Ziel: lokalen Restart simulieren
- Anwendung muss richtig “aufräumen”
- oder selbst Cleaner herrichten



Storage: The File Problem

- Für Device paralleler Zugriff kein Problem
- Filesystem kann das nicht
- nichtmal R/O Zugriff mit ext2, wegen cache-Kohärenz
- Filesystem nur einmal pro Filesystem mountbar
- sonst teure, sehr junge Erweiterungen: Sistica GFS, Veritas, XFS(?)
- DLM notwendig zum Synchronisieren
- Write-Performance penalty



Solution 1: Block Access

- Raw Block Access
- DLM in Anwendung
- Beispiel: Oracle RAC
- Serialisierung der Zugriffe via DLM
- kein Linux DIM verfügbar
- DLM von Anwendung
- Nur von Oracle derzeit verwendet



Solution 2: File Access

- Raw Block Access
- DLM in Anwendung
- Beispiel: Oracle RAC
- Serialisierung der Zugriffe via DLM
- kein Linux DIM verfügbar
- DLM von Anwendung
- Nur von Oracle derzeit verwendet



Open Source Tools

- heartbeat
 - Alan Robertson (IBM)
 - Load Group orientierter Service Manager
- drbd: Shared Storage Simulator
- div. Scripts für heartbeat
 - LinuxSCSI
 - Ipaddr
 - Device-locking ist ein Thema um Fehlbedienung zu vermeiden

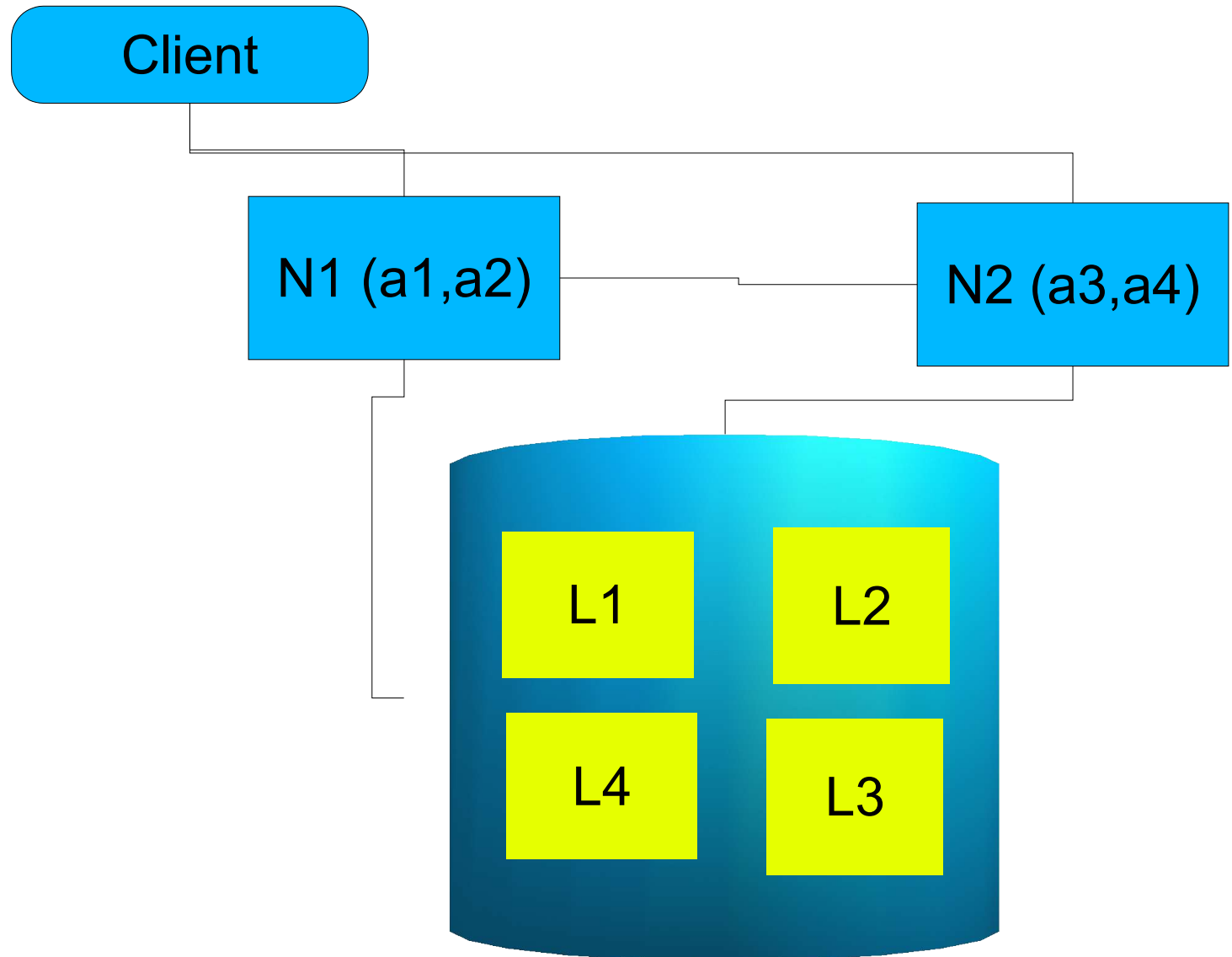


heartbeat

- “normaler” Service
- Konfiguration via ha.cf
 - keepalive
 - deadtime
 - multicast
 - serial
 - nice_failback
 - stonith
 - node Definitionen
- haressources:
 - von links nach rechts
 - isis 10.0.0.3 LinuxSCSI::1:0:0:0 LinuxSCSI::1:0:0:1:0 samba

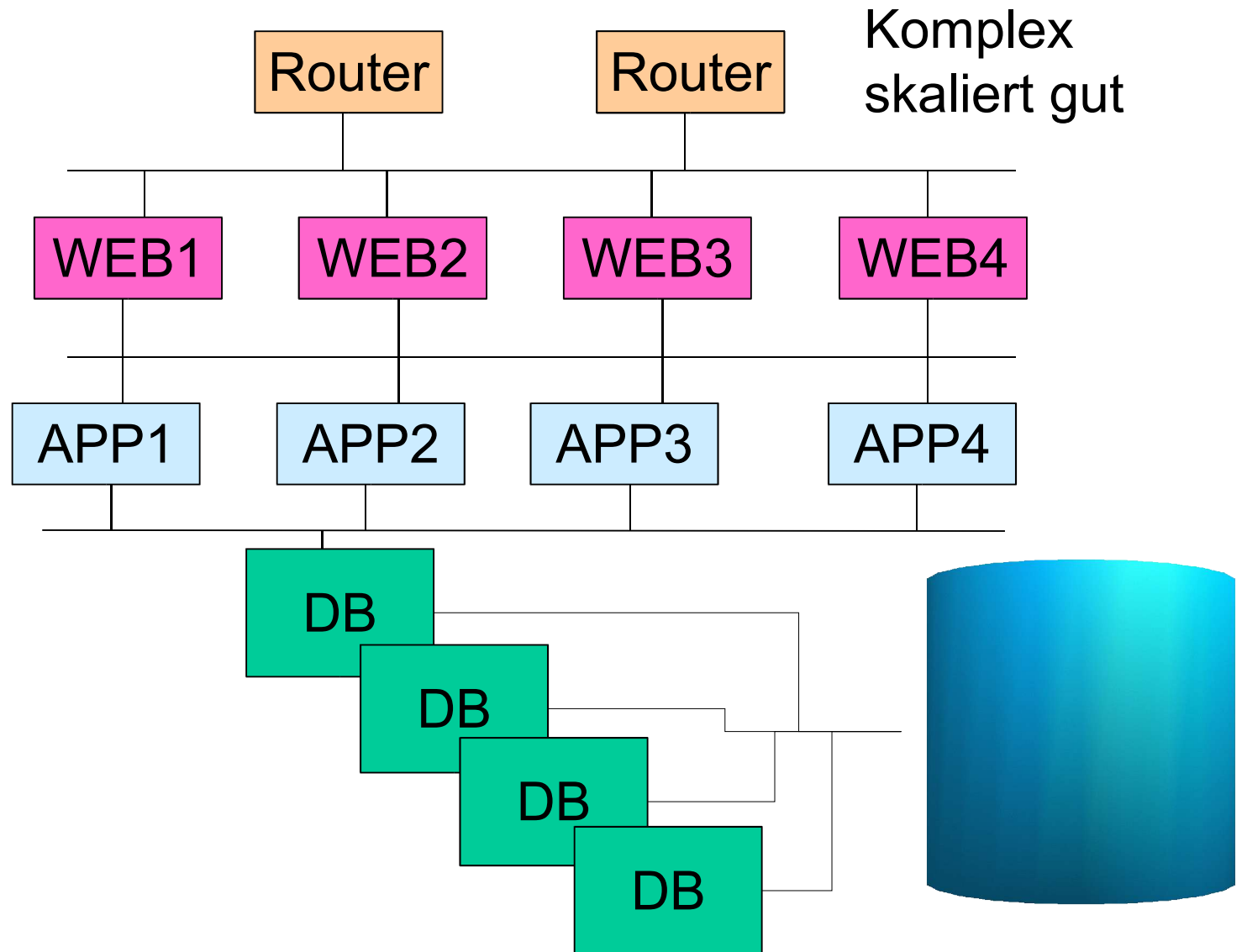


System Design



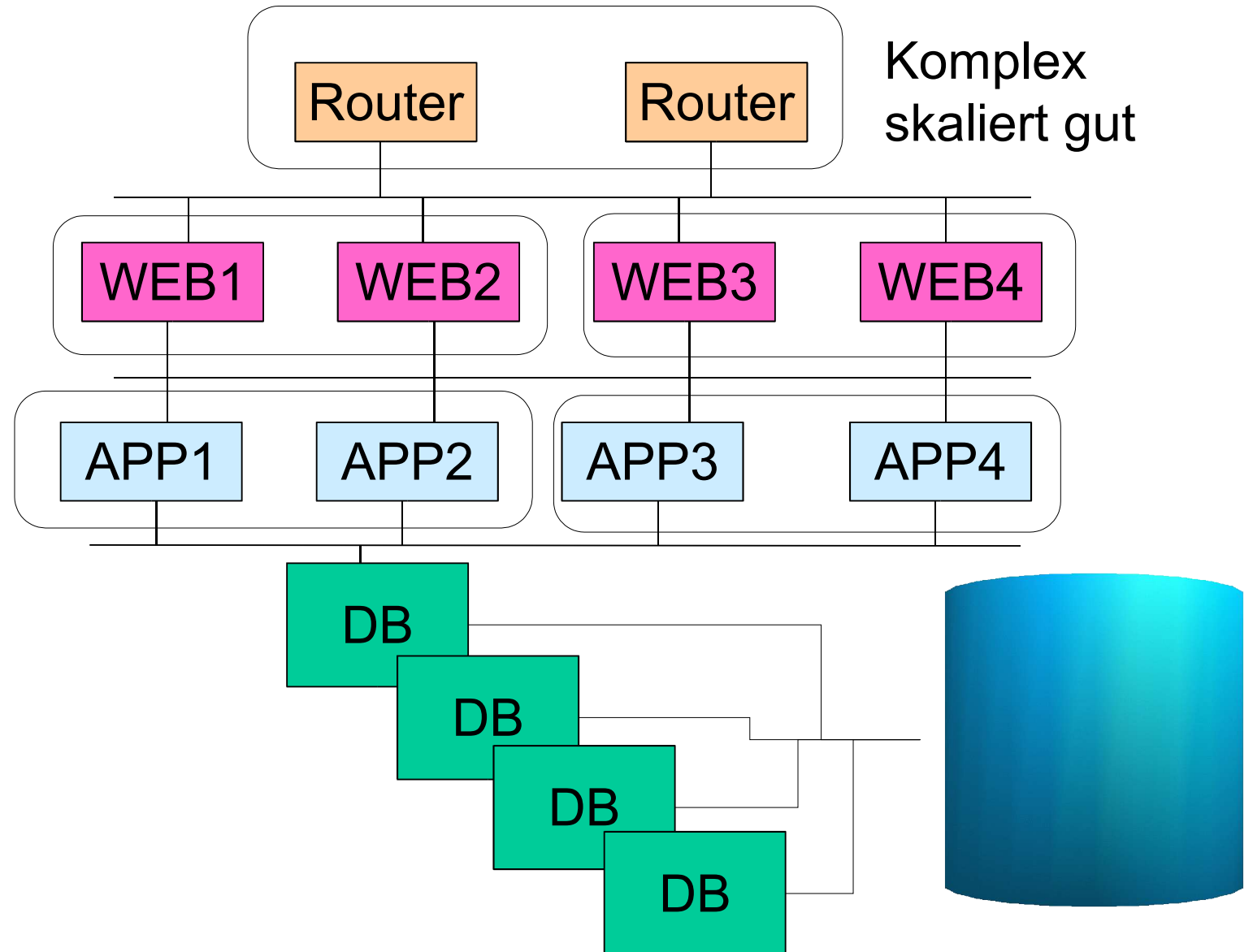


complex Architectures





complex Architectures





Management-Tools

- ex?: schnelles Ausführen von Commands auf fremden Knoten im aktuellen Pfad
- to?: Kopieren von Files auf fremden Knoten im aktuellen Pfad
- sync-machines: rsyncen von Files basierend auf Timestamps
- Sym-Links: zu synchronisierende DIR auf shared-storage linken
- etc weitestgehend synchron halten
 - für grosse Cluster: Macro-Preprocessing



Examples

- Hosting Environment CUBiT IT
- Progress Cluster Schloss Schönbrunn
- mysql-Clusternodes
- Fileserver Cluster



Building Blocks in complex environments, common Problems

- Filesystemproblem
- externe Disken teuer
- NFS-System: NetApp Filer
- eigene NFS Server: externer Storage teuer
- SAN teuer
- Cluster als Komponente für grosse Cluster
- performance steigt nicht immer
- Serialisierung kostet Leistung

linuxwochen

PETER-PAUL WITTA



Clustering mit Shared Storage

DANKE!

Ing. Peter-Paul Witta
paul.witta@CUBiT.at