

CUBiT IT Solutions GmbH

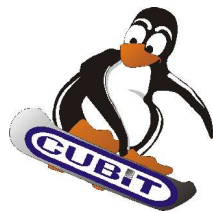


Ing. Peter-Paul Witta
paul.witta@CUBiT.at



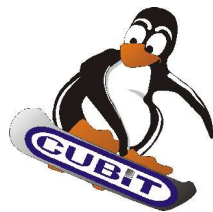
CUBiT IT GmbH

- Open Source Dienstleister
- Fokus auf Open Source Services
- Beratung bei Auswahl
- Implementierung
- Support
- Betrieb
- seit 1999 aktiv
- rein österr. Unternehmen



Einsparung mit OpenSource Lösungen

- heute hauptsächlich Infrastruktur
- noch wenig fertige Anwendungen
- hohe Flexibilität
- Anpassungsaufwand vorhanden
- keine Endlösungen von der Stange
- Aber nutzbare, ausgereifte und fertige Infrastrukturlösungen



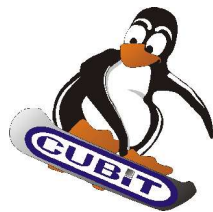
Open Source

- Kostenlos verfügbar unter GPL
- Einsatz im eigenen Unternehmen kostenlos möglich
- Weitervertrieb kaum möglich (nur mit Quelltext)
- Oft als Nebenprodukt entstanden (aus Forschung oder als konkrete Problemlösung ohne Vertriebsabsicht)
- Oft Teil einer kommerziellen Lösung
 - z.B. Inflation der Open Source DB



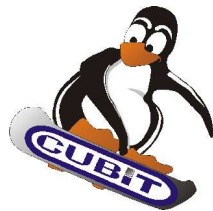
Mythen über OpenSourceSoftware

- “OpenSource unausgereift”
- Support schwierig (1999: “Manitoo aus dem IRC”)
- Zusammenarbeit Community/Unternehmen
- “Skaliert schlecht”
- “Nicht standardkonform”
- komplex und kompliziert
- aufwendig



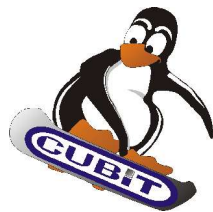
Open Source = Unix

- textbasierte Konfiguration
- Möglichkeit, direkt im System zu dokumentieren und kommentieren
- Versionierung der Konfigurationsdateien möglich
- Damit kontrolliertes Release Management
- Nachvollziehbarkeit von Aktivitäten bei Changes
- (Teil)-Automatisation von wiederkehrenden Aufgaben (Scripting)



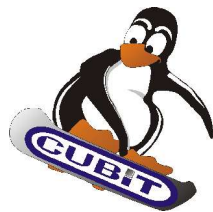
Konventionelle Lösungen

- UNIX: teure Systeme, wenig Leistung
- seltsame Funktionalität, Handling (z.B. keine Cursor Tasten)
- Features sehr teuer
- Wartung teuer
- Implementierung schwierig
- oft unterlizensiert, daher schlechter Nutzen
- Interoperabilität
- Passt zu 70%



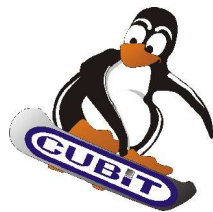
Intel/LINUX

- Preis/Performance Killer
- commodity vs. “Unix: System aus der Raumforschung”
- Hardware-Wartung
- Mit Linux ein richtiger Unix-Server (20% Leistungssteigerung durch Win2Lin Migration)
- Techniken aus der “großen Welt”: Virtualisierung, SAN, IP-Storage
- Net-, IO-Performance und Adapter gleich



Clustering mit intel/Linux

- mehrere kleine Rechner leisten gemeinsam Grosses
- günstige dual intel/AMD Server
- load sharing
- High Availability
- centralized management
- Tools:
 - heartbeat (Open Source, Alan Robertson)
 - Shared Storage (NetApp Filer)
 - optional Load Balancer
 - etwa LVS oder backhand (siehe extra Folien)



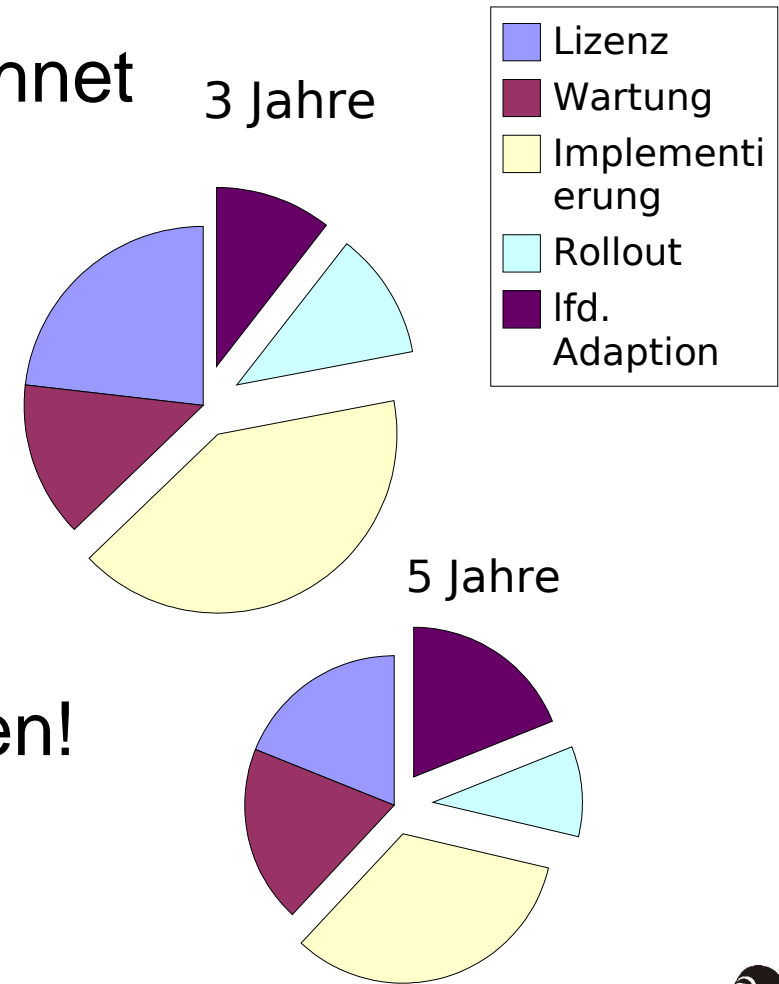
OpenSource/kommerz. Lösungen

- offene Lösung
- anpassbar
- Implementierungsaufwand für große Systeme konstant gleich
- Im **Infratrakturbereich** ausgereift
- Egal ob Client oder Server (Thin-Client, Server, Cluster, Netzwerk)
- Im Anwendungsbereich sind kommerzielle Lösungen oft noch zu überlegen



Kosten: Was kostet wieviel?

- Auf 3 oder 5 Jahre gerechnet
- Produktpreis
- Produktwartung
- Implementierung
- Handling und Wartung
- Keine skalierenden Kosten!



Preistendenzen

- Produktkosten sinken: Reaktion auf OpenSource Preisdruck
- z.B. DB-Markt: Firebird, Ingres, SAPdb,...?
- Installationskosten
- Handlingkosten konstant bis steigend
- Handlingaufwand, "Ease Of Use"



Handlingkosten

- Wie leicht ist Produkt implementierbar?
- Wie leicht handhabbar?
- Cursortasten?
- Wie lässt sich Produkt integrieren?
- Wie weit verbreitet?
- Supportbar?
- commodity vs. elitärem Spezialtool



Einsparungspotential

- weniger Lizenzkosten
- weniger Rollout-Aufwand
- weniger Wartungsaufwand
- weniger Anpassung, laufende Maintenance
- weniger Integrationsaufwand
- schnellere Abwicklung
- weniger Kosten, weniger Budget-Aufwand intern



Mythos: Open Source ist kompliziert

- Open Source oft klarer verständlich
- Best-Of-Breed
- kein Marketing-"bla bla"
- klare technisch bedingte Produktvorteile
- beste Technologie (Evolution)
- leicht und durchgängig integrierbar
- leicht einsehbar



Einheitliche Technologie

- Open Source Lösungen passen zusammen
- flexibel durch Glue-Code
- einheitliche Basistechnologie (Unix, Regexp, TCP/IP, Perl, PHP, Java,...)
- einheitliche Konfiguration (z.B. Apache-Style: ProFTP, Apache, CUPS,...)
- oder: dhcpd, bind, xinetd,...



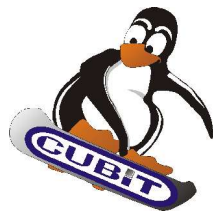
IT Felder, IT im Betrieb

- Infrastruktur
 - Betriebssystem
 - Datenbanken
 - Netzwerk
 - Security
- Anwendungen
 - ERP
 - Groupware
 - ...



IT “OpenSource Style”

- offene Protokolle
- offene Systeme
- Systeme passen zusammen
- Database: MySQL, PostgreSQL
- Application Server: Apache, Tomcat, PHP
- Email: Sendmail, Postfix, Exim (mit ClamAV und SpamAssassin)
- System-Management: Nagios
- Snort



Infrastruktur mit OpenSource



PostgreSQL



gcc

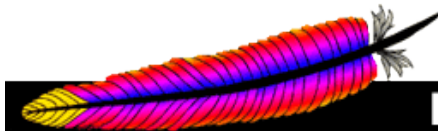


Nagios™



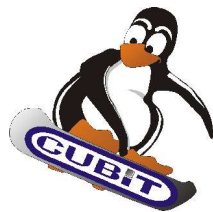
Apache

HTTP SERVER PROJECT



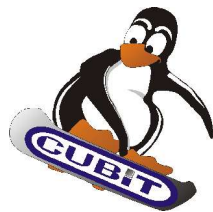
Konsolidierung und Virtualisierung

- vmWARE
 - auf Linux Basis
 - ermöglicht stabiles Handling mehrerer Win-Apps auf einem phys. Server
- User Mode Linux
 - echte virtuelle Linux Server
 - jeder mit eigenem "Root", IP, MAC,...
- virtual Hosting
 - Anwendungskonsolidierung



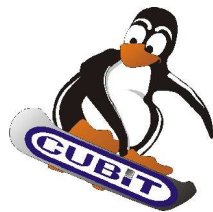
Open Source Highlights, Trends

- VmWare, UML
- Konsolidierung, Integration
- Security (IDS, IPS, Spam, Firewall, Authentication, Virus)
- Load Balancing, Clustering
- Durchsatzsteigerung
- Betriebsautomatisation
- Service und SLA Management (Nagios)



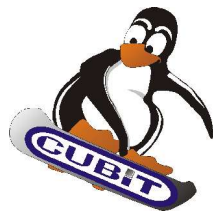
Client Trends

- Linux based
- Anwendungen per RDP oder Citrix
- Web-basierte Anwendungen
- “Beschnuppern” von End-User Anwendungen
- Unbegründete Angst vor Akzeptanz
- Mutige Versuche
- unterschiedl. Ergebnisse je nach Projekt-Setup



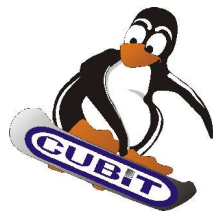
Referenzen

- Österreich und Umgebung
 - CUBiT ist in D,A,CH aktiv
 - Consulting auch in anderen Ländern
 - Services werden oft remote geliefert



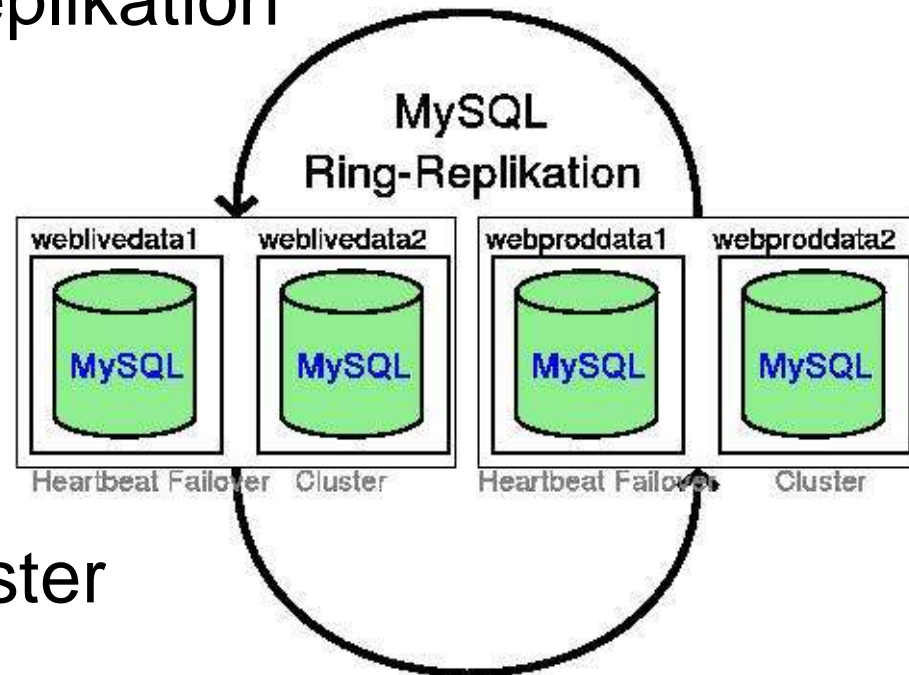
Schloß Schönbrunn Zentralisierung

- DB Server Progress
- Konsolidierung vieler Server auf 1 Cluster
- Cluster mit DRBD (geograph. verteilt)
- Volle Datenredundanz ohne Hardware
- Spezielle SW Lösung für Kassensysteme
- zeigt Verfügbarkeitsprognose an
- redundanter Backbone mit Linux Routern



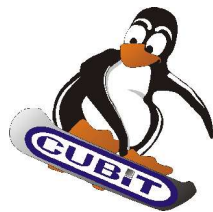
ÖAMTC MySQL Cluster

- DB Cluster mit Ringreplikation
- Metrocluster aus 2 lokalen Subclustern
- für viele interne und externe Datenbanken
- zwei lokale Linux Cluster mit Shared Storage
- MySQL MaxDB mit MySQL Replikation



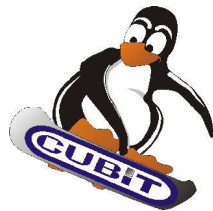
SAP mit Linux

- Ablöse Tru64 Mainframe
- durch NetApp Filer und intel Linux Server
- höhere Leistung
- Finanzierung in Grössenordnung des Jahreswartungsbudgets der Altsysteme



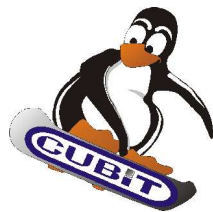
Netzwerkmanagement Nagios

- mehrere Projekte mit renomierten österr. und internationalen Kunden
- Beratungsintensiv
- keine Lizenzkosten
- Rollout-Aufwand vergleichbar mit kommerziellem Produkt
- flexibel, leichter handhabbar
- Unix-affin



ÖAMTC Nagios

- Überwacht werden:
- Samba und nmbd (File- und Printservices)
- lpd (print spooler)
- winbind (Benutzerauthentifizierung und Authorisierung gegen das Active Directory)
- Systemdaten (Unix Load, Memory, Swap)
- Prozessstati der Serverprozesse



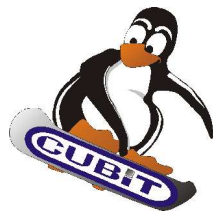
Center Systems

- Fileserver Cluster
- stabil und nachvollziehbar (vgl. Windows File Server)
- ACL Integrtation
- hohe Leistung
- günstig und skalierbar
- Samba



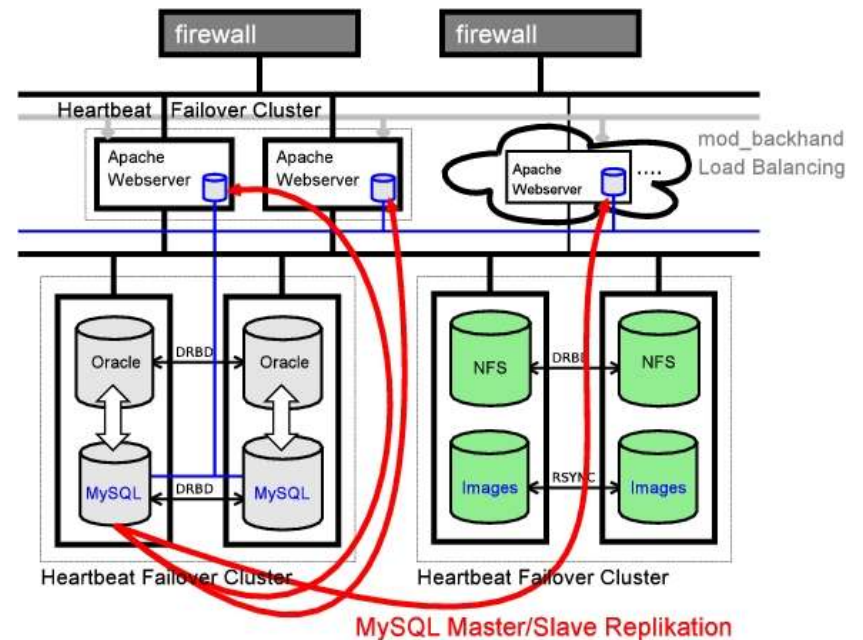
Ebewe Pharma DHCP Lösung

- Redundanter Load Balancing DHCP Cluster
- Hohe Leistung
- DynDNS
- vgl. MSAD
- Integriert sich gut in Open Source Umgebung



MC Retail Hosting

- 5 Node MySQL Cluster mit Replikation
- schnelle Replikation
- leicht anwendbar
- Lizenzkosten vernachlässigbar (vgl. EUR 15k/CPU)
- Suchzeiten
180sec =>0.8sec



Danke

- Wir freuen uns auf ein Gespräch
- Folien im Web: <http://www.cubit.at/>



High Availability

- Ausfallsicherheit durch Redundanz
- duale Server
- jeder “backt” den Anderen
- pro Dienst auch mehrere Server aktiv
- hohe Leerkosten durch Mehr-Hardware
- Bei Ausfall Migration der Dienste auf Ersatzsystem
- Umschaltzeiten beachten und tunen!



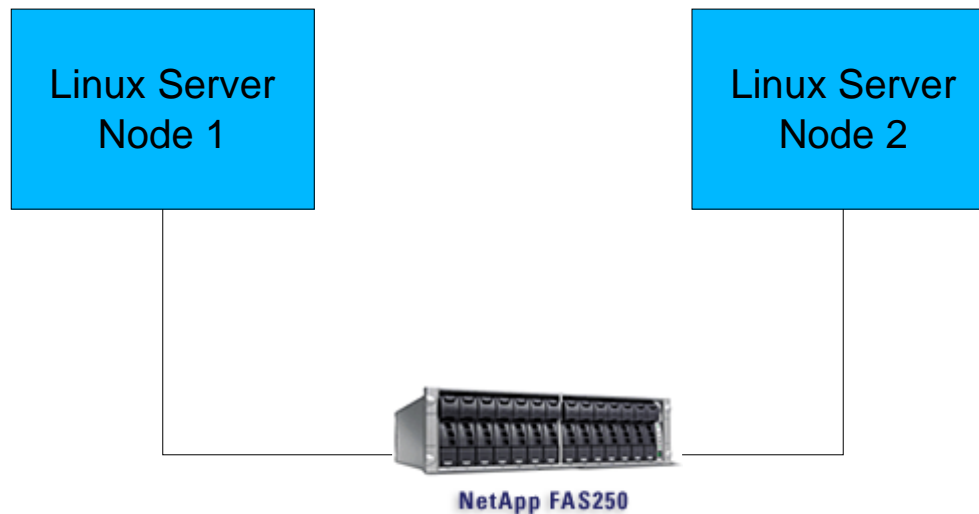
System Architecture - Simple



- Datenhaltung synchron
- nur ein System kann gleichzeitig Daten am FC/AL Disk nutzen
- bei Failover Daten sofort ansprechbar (kein fsck notwendig!)
- kein Loadsharing möglich
- massive Fibrechannel-Probleme (Qlogic-Firmware, Switches,...)
- Standard-Linux-Tools wie heartbeat begrenzt nutzbar
- Achtung: Datenverlust bei parallel Mount(!)
- Achtung: limitierter Support von proprietären Lösungen (Steeleye)
- Achtung: OpenSource-nahe Anwendungen supporten nur heartbeat



System Architecture - Simple



- Datenhaltung synchron
- jedes System kann Daten am Filer nutzen
- bei Failover Daten sofort ansprechbar (kein fsck notwendig!)
- durch NFS können beide Systeme Filer mounten (aktiv/aktiv)
- so auch Loadsharing möglich
- keine Fibrechannel-Probleme (Qlogic-Firmware, Switches,...)
- Standard-Linux-Tools wie heartbeat nutzbar
- Achtung: Filer Single Point Of Failure

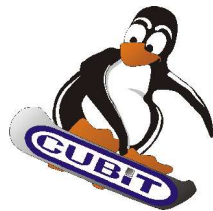
System Architecture – HA Storage



- wie mit FAS 250 aber
- Filer redundant
- bei Ausfall eines Filers (Hardwaredefekt,...)
 - kein Totalausfall sondern leichter kurzzeitiger Performanceeinbruch
- Load Sharing auch im Filer-Bereich; beide Filer aktiv/aktiv

Open Source Tools - heartbeat

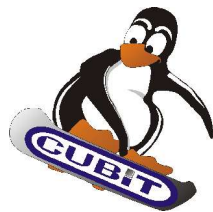
- heartbeat
 - Alan Robertson (IBM)
 - Load Group orientierter Service Manager
 - ähnlich HACMP
 - Standard-Lösung im Open Source Bereich
 - Achtung: kein Schutz vor Parallel Mount des shared Storage(!Datenverlust!)
- div. Scripts für heartbeat
 - Ipaddr – Serviceadressenmanagement
 - Support für die meisten Linux Standardservices



heartbeat

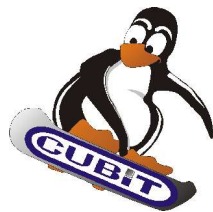
Architecture (lg concept)

- 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



Applications

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



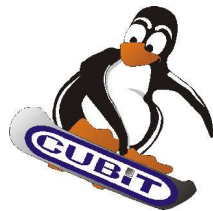
Storage ohne Filer

- FC/AL und SCSI ist theoretisch Multihost-fähig (NFS wäre noch besser)
- FC/AL / SCSI hat Locking-Probleme
- Terminationsprobleme bei SCSI
- Achtung: Multipath mit FC/AL ist schön aber schwierig, kaum zertifizierbar
- Qlogic-FC/AL Probleme
 - Disk-Probleme, Locking, Dual-Mount
 - Datenverlust droht
 - FC-Switching: Multipath problematisch
 - bei parallelem Mount Filesystem-Verlust



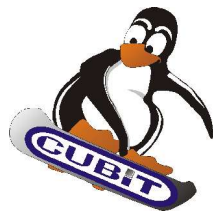
Storage ohne Filer (2): The File Problem

- 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, unerprobte Erweiterungen:
 - Sistina GFS, Veritas, XFS(?)
- DLM notwendig zum synchronisieren
- massive Write-Performance penalty
- permanente Gefahr von Datenverlust



Storage mit Filer

- Multihost-fähig (NFS besser)
- keine locking Probleme
- keine Terminierungsprobleme
- Multipath mit Ethernet-Bonding erprobt und stabil
- keine Write-Penalty
- keine Lock Contention am Cluster-Filesystem
- keine Adapter-Probleme, da Ethernet
- kein Clusterfilesystem notwendig!
- erprobte Lösung



Solution: File Access

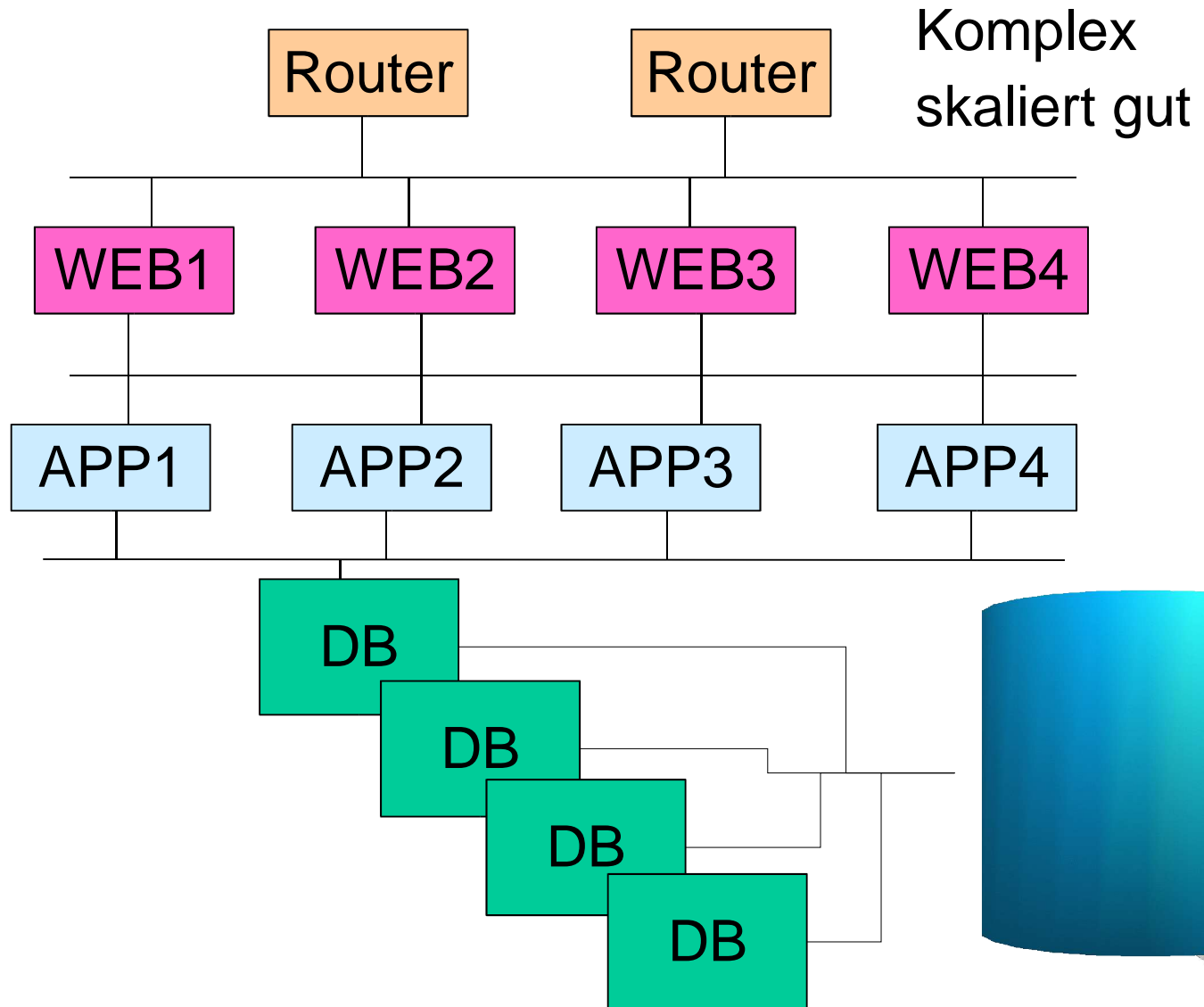
- File-Locking via NFS und Netapp
- DLM nicht notwendig
- Beispiel: Oracle RAC
- Serialisierung der Zugriffe via Filer
- kein Linux DLM notwendig
- funktioniert mit allen Anwendungen
- skaliert gut
- unterstützt load sharing mit mehreren Knoten



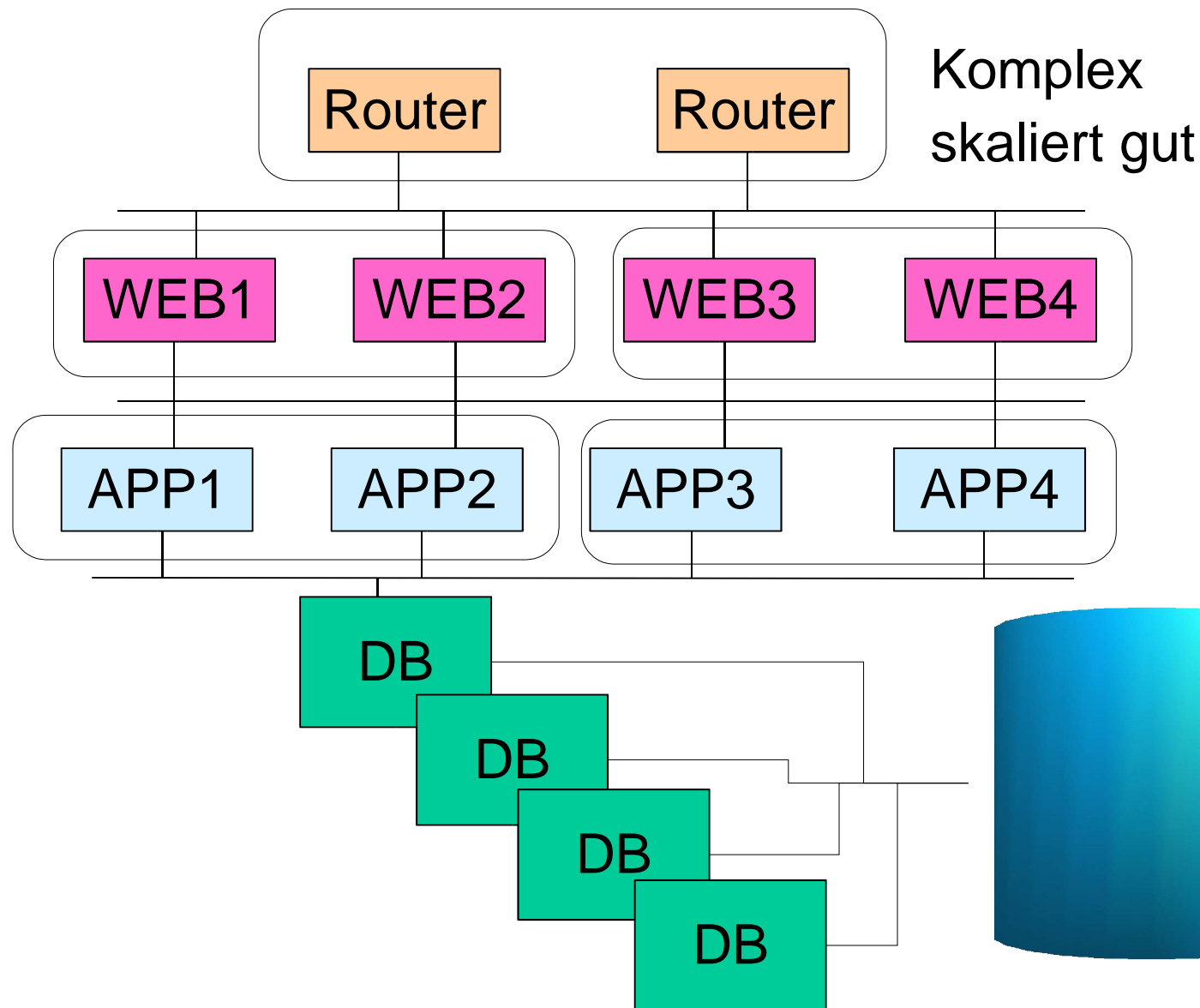
heartbeat im Detail – config “ha.cf”

- heartbeat ist ein “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

complex Architectures

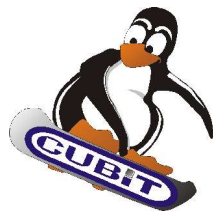


complex Architectures – build cluster pairs!



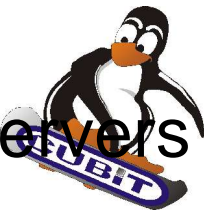
CUBiT Management-Tools

- 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



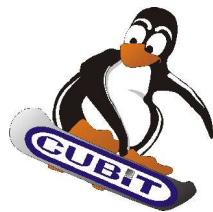
CUBiT Services

- Testen und Zertifizierung von heartbeat-Versionen
- Achtung: SuSE und Redhat liefern tw. defekte heartbeat-Version (unwanted Failover,...)
- Switch-Tuning für heartbeat oder cross-cable
- heartbeat-Network definitions
- mlm/ipfail kann bei Netzverlust Failover erzwingen auch wenn Server läuft (enhanced Stability)
- SNMP Trap und Alarm bei Failover
- Advanced Management und Training:
- 80% der Fehler in Clustern sind Bedienfehler!
- Anpassung an Health Management des Intel Servers



Examples

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



more Info

bei Rückfragen stehen wir jederzeit gern zur Verfügung

<http://www.cubit.at/>

Cluster-Referenzen unter “Referenzen”

Email: paul.witta@cubit.at

Verwandte Themen:

- Clustering mit Shared Storage
- Load Balancing Cluster Architectures
- NFS und NIS mit Linux
- Nagios Systems Management

Folien unter <http://www.cubit.at/lw04/>

