

# Linux im Einsatz

Betrieb, Installation und  
Monitoring im Datacenter  
des 21. Jahrhunderts

**CUBiT IT Solutions GmbH**  
**Ing. Peter-Paul Witta**

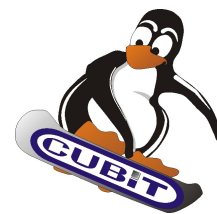
<paul.witta@cubit.at>  
<http://www.cubit.at/pres/>

Networking

Datacenter

Storage  
Systems

Redundancy



# Architektur

Networking

Datacenter

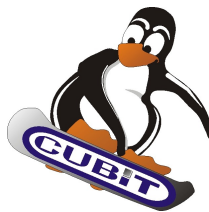
Storage  
Systems

Redundancy



## High Grade Consolidation

- hohe Packdichte
- Bladecenter
- Hardwareauswahl begrenzt
- Energieverbrauch pro Server von 1200 Watts => 300W
- Energieeinsparung
- Stichwort Lüfter mit jet-Turbinen Technologie (hp)
- Consolidation



# Building Blocks

Networking

Datacenter

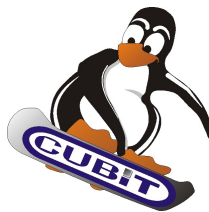
Storage Systems

Redundancy



## Enterprise Hardware x86

- 64 Bit mit viel RAM
- Serviceprozessoren: BMC, ILO & Co
- Hardware onBoard Diagnose
- Console Umlenkungen
- Sensor-Diagnostik
- IPMI 1.5 ?
- Fujitsu-Siemens, HP, IBM, Dell, Sun, intel
- Blade Technologie
- ChipKill / ECC
- AMD / Intel



# Building Blocks

Networking

Datacenter

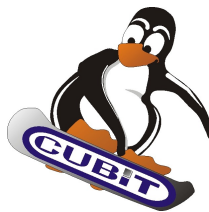
Storage  
Systems

Redundancy



## Enterprise Hardware sonstige

- zSERIES: LPAR oder VM: zLINUX
- pSERIES: LPAR oder Micropartition
- iSERIES: LPAR oder Micropartition
- HP: Linux on Itanium
- SUN: Linux on SPARC (???)



# Building Blocks

Networking

Datacenter

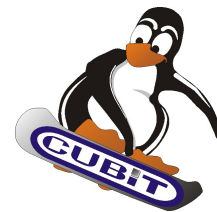
Storage  
Systems

Redundancy



## Network

- High Performance/High End Gear
- CISCO/HP
- portChannel
- VLAN
- tagged termination
- Interface nicht mehr "eth0" sondern "bond0.80:10"
- Failover/portChannel BONDING
- Link Aggregation



# Building Blocks

Networking

Datacenter

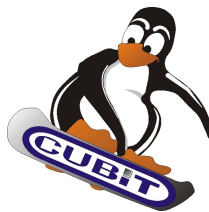
Storage Systems

Redundancy



## Buildup Network on PXEBoot

- Link resiliency must be conf'd in initrd
- dhcp on which interface
- interface/MAC binding
- high performance link utilization



# Building Blocks

Networking

Datacenter

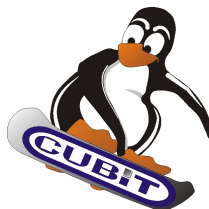
Storage Systems

Redundancy



## NetWork III

- InfiniBand
- multi-link utilization
- zero-copy 10gbE boards w/ RDMA
- FibreOptics: CWDM, Transceiver, SFP, Lasers



# Building Blocks

Networking

Datacenter

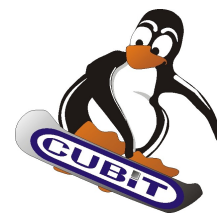
Storage  
Systems

Redundancy



## FibreChannel

- Software Protocol such as SCSI
- HBA/Drivers
- Certification issue
- QLOGIC / HP/ Driver/ Distri/ App / Storage
- SOLUTION: NAS!





# Building Blocks

Networking

Datacenter

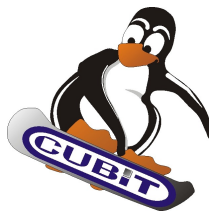
Storage Systems

Redundancy



## iSCSI/NFS

- rock-solid high performance Ethernet req'd
- ipSAN / iSAN
- specialized "SAN" on IP base
- operated by server group
- focused on high performance highly resilient links



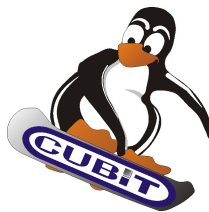
# Linux Distributionen

Networking

Datacenter

Storage Systems

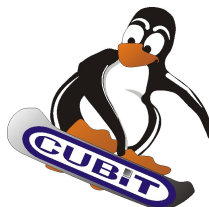
Redundancy



# Linux Distributionen

## SuSE Linux Enterprise Server

- basiert auf OpenSUSE (welche auf SuSE Prof. basiert; diese war früher Slackware)
- RPM based
- jetzt Novell
- gute Novell OES Integration
- guter Support vor Allem mit SAP
- gute Nationalisierung
- KDE/Gnome
- guter Hardware Support
- YaST based
- XeN inkludiert
- Support für SAP und Oracle



# Linux Distributionen

## RHES/RHAS/AP

- Red Hat Enterprise Server / Advanced Server
- Advanced Platform Pack
- fully GPL Distribution (siehe CentOS)
- amerikanisch
- erste Wahl für Oracle (siehe Oracle Linux)
- guter Hardware Support
- Advanced Platform: fixer Preis für virtuelle Instanzen
- XeN inkludiert
- Support für SAP und Oracle



Networking

Datacenter

Storage Systems

Redundancy



# Linux Distributionen

## CentOS

- basiert auf RHES SRPMS
- kostenfreier Build
- kostenloses frei supportetes Enterprise Linux
- Problem: Im Fehlerfall muss Problem mit RedHat reproduziert werden

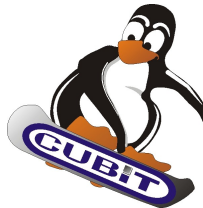


Networking

Datacenter

Storage Systems

Redundancy



# Linux Distributionen

## Oracle Linux

- basiert auf CentOS
- Oracle supportet Oracle Anwendungen am liebsten auf Oracle Linux
- technisch wirdes aber von Hardware-Herstellern nicht voll supportet

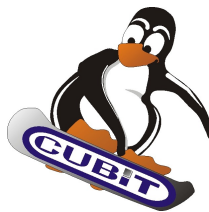
Unbreakable  
**Linux**  
ORACLE®

Networking

Datacenter

Storage  
Systems

Redundancy



# Linux Distributionen

## Debian



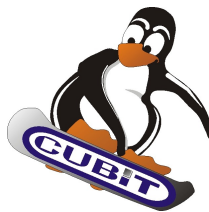
- Support von HP und Fujitsu Siemens (nicht IBM!)
- gute Releasepflege
- professionelle Distribution
- keine klagbare Firma dahinter, daher kaum supportet
- aber zB. Oracle supportet auch Systeme auf Debian
- leider aber nicht für SAP
- erste Wahl für Betrieb von OSS Anwendungen
- gute Wahl für virtuelle Systeme (kein Hardwaresupport)

Networking

Datacenter

Storage Systems

Redundancy



# Linux und NFS-Root

Networking

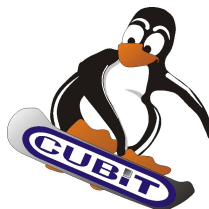
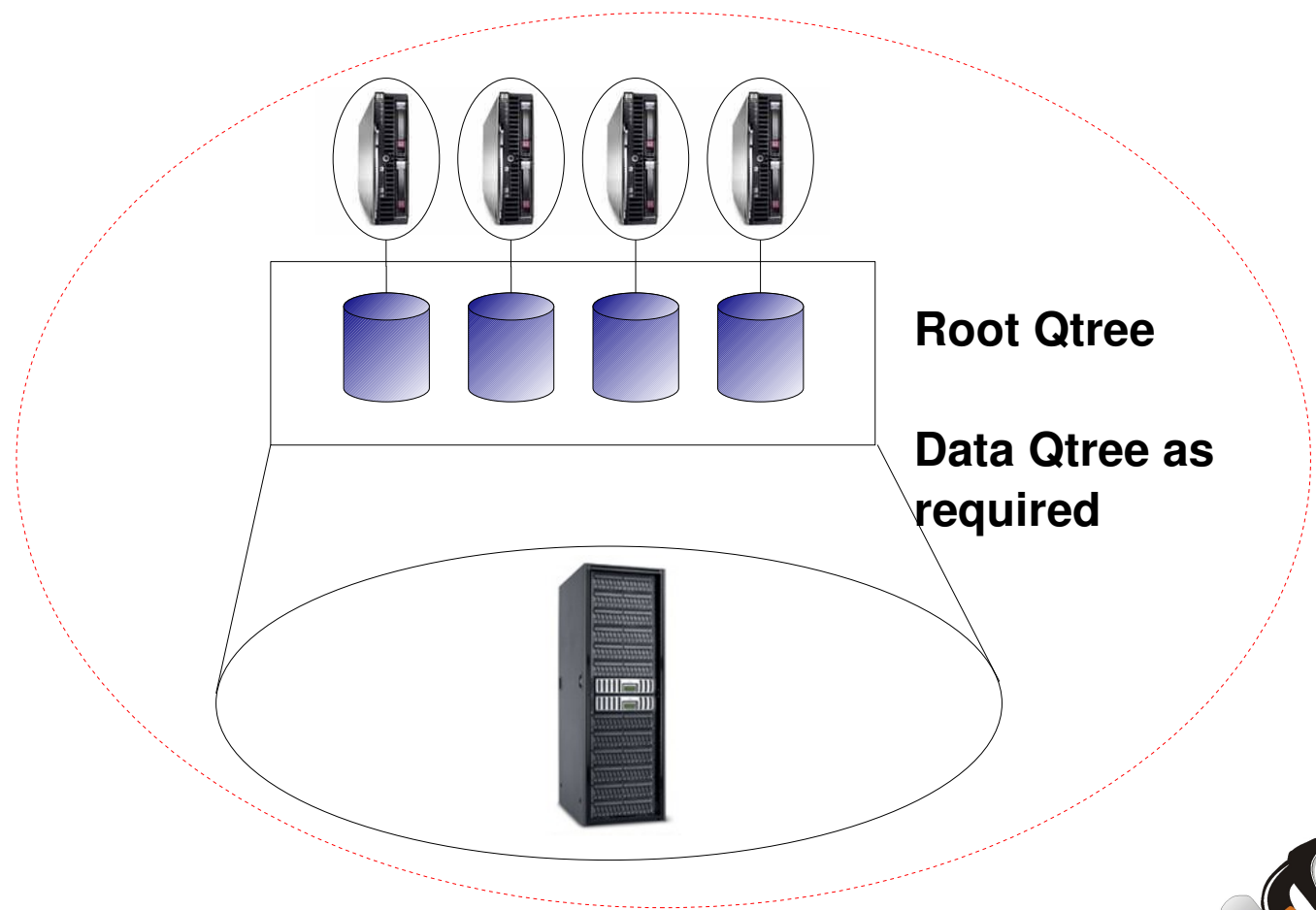
Datacenter

Storage Systems

Redundancy



## PXE Boot





# Linux und NFS-Root (2)

Networking

Datacenter

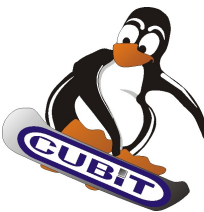
Storage  
Systems

Redundancy



## Boot Process

- 1) Power up
- 2) Start PXE BIOS extension per adapter  
(tries DHCP request)
- 3) DHCPACK, set „next-server“
- 4) Try TFTP of boot loader at „next-server“
- 5) Load PXELinux
- 6) Boot Kernel, run initrd
- 7) In initrd, activate redundant network links
- 8) Reconfiguration will break links
- 9) pivot root, run init



# Linux und NFS-Root (3)

Networking

Datacenter

Storage Systems

Redundancy

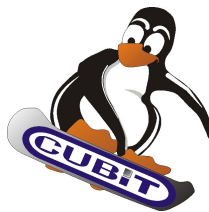


## Advantages

- Minimal storage resources required
- Snapshot backup able
- For root & Qtree-config
- Failover & hardware swap
- Based on requirements from Workshop
- Role change possible
- High compression / deduplication

## Disadvantages

- Clustered Storage required, otherwise single point of failure
- Performance on clustered storage critical
- Requires NetApp



# Storage Systeme

Networking

Datacenter

Storage Systems

Redundancy



## STORAGE SYSTEMS for Enterprise Linux

Linux: Filesystem Problem

XFS/Reiser/EXT4

Snapshot

Accelerator Cards

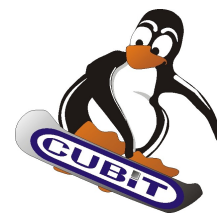
Deduplication

Proactive RAID health checks

NFS is faster than block access (more intelligent storage servers)

NFS, iSCSI, FCP compared

dedicated storage servers provide better reliability and performance



# Resiliency

Networking

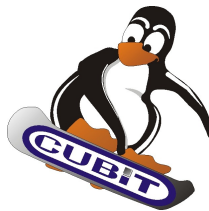
Datacenter

Storage  
Systems

Redundancy



- 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(?)



# Resiliency (2)

Networking

Datcenter

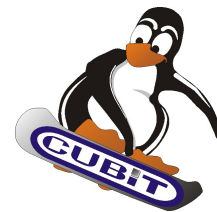
Storage  
Systems

Redundancy



## Heartbeat Cluster Ressource Manager

- “normaler” Service
- Konfiguration via ha.cf
  - keepalive
  - deadtime
  - multicast
  - serial
  - nice\_failback
  - stonith
  - node Definitionen



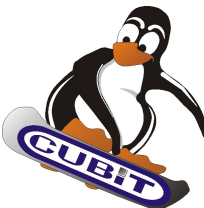
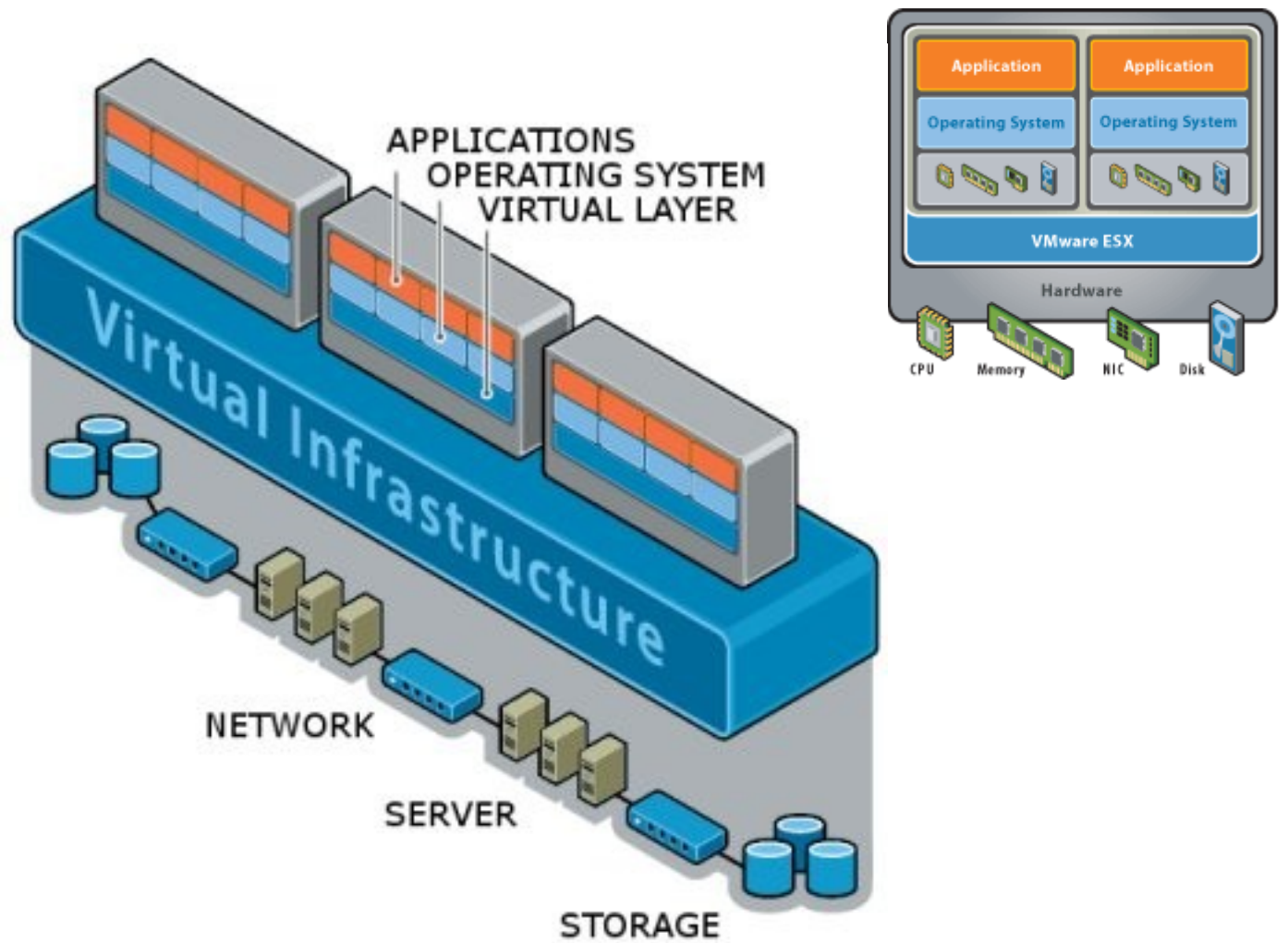
# VMWare

Networking

Datacenter

Storage Systems

Redundancy



# Monitoring

Networking

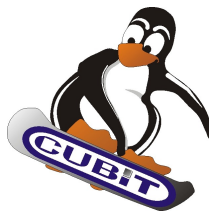
Datacenter

Storage  
Systems

Redundancy



- Information wenn Dienste ausfallen
- über Systemstatus informieren und protokollieren (Verfügbarkeit)
- langfristige Statistiken als Grundlage für Entscheidungen (Aufrüstung bei Leistungsbedarf)
- Überprüfung von externen Dienstleistern (ISP, Telekom, Outsourcer) und deren SLA
- zentrale Informationsstelle
- automatisiertes Reagieren auf Probleme
- automatisierte Behebung
- Umbrella Management



# Monitoring Strategien

Networking

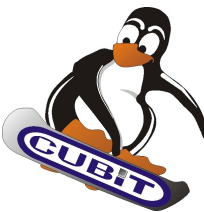
Datacenter

Storage  
Systems

Redundancy



- Blackbox Monitoring -- von außen zugreifen wie ein Anwender
- Whitebox Monitoring -- von innen alle Komponenten einzeln Funktionsprüfen
- Schwellwert Monitoring: Überwachen von Messwerten
- richtige Eskalation der Notifizierung
- ggf. automatic response („self-repairing“)
- Compound Checks
- Statistiken: Correlation, SLA-Auswertung
- Umbrella-System
- Monitoring-Netzwerk





# Strategien – Blackbox

Networking

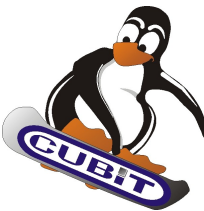
Datacenter

Storage  
Systems

Redundancy



- Blackbox Monitoring -- von außen zugreifen wie ein Anwender
- für Standard-Protokolle mit vorhandenen Plugins für FTP,HTTP,NFS, SMB (Samba/CIFS), Citrix,DNS und viele andere
- für eigene Anwendungen durchaus auch automatische Überprüfung der Business-Logik
- zB: Webshop: automat. Einkaufen, erzeugen eines speziell markierten Auftrages, der nicht weiterverarbeitet wird



# Strategien – Whitebox

Networking

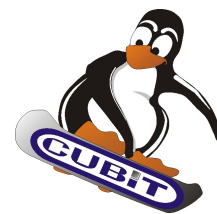
Datacenter

Storage  
Systems

Redundancy



- Alle Teilkomponenten der Anwendung getrennt prüfen
- Notwendige Datenbanken, Anwendungen, Netzwerkequipment, Frontendserver, Netzwerke, Subsysteme,... ständig jeden einzeln prüfen
- Notwendig auch zur Problemlokalisierung
- liefert aber nicht gleiche Sicht wie Anwender sieht
- System- nicht Lösungsbezogen



# Strategien – Messwerte

Networking

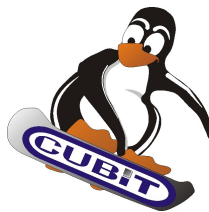
Datacenter

Storage  
Systems

Redundancy



- laufende Überwachung von Leistungsdaten
- CPU, Netz, Plattenauslastung
- Überwachen von Tuningmaßnahmen, wie zB. Cache-Hit-Ratio
- Alarm bei nicht optimaler Leistung
- Alarm bei bedrohlichem Zustand (Disk Full 90%)
- Alarm bei Aufrüstungsbedarf (80% Leitungsauslastung im Tagesmittel)



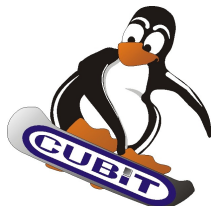
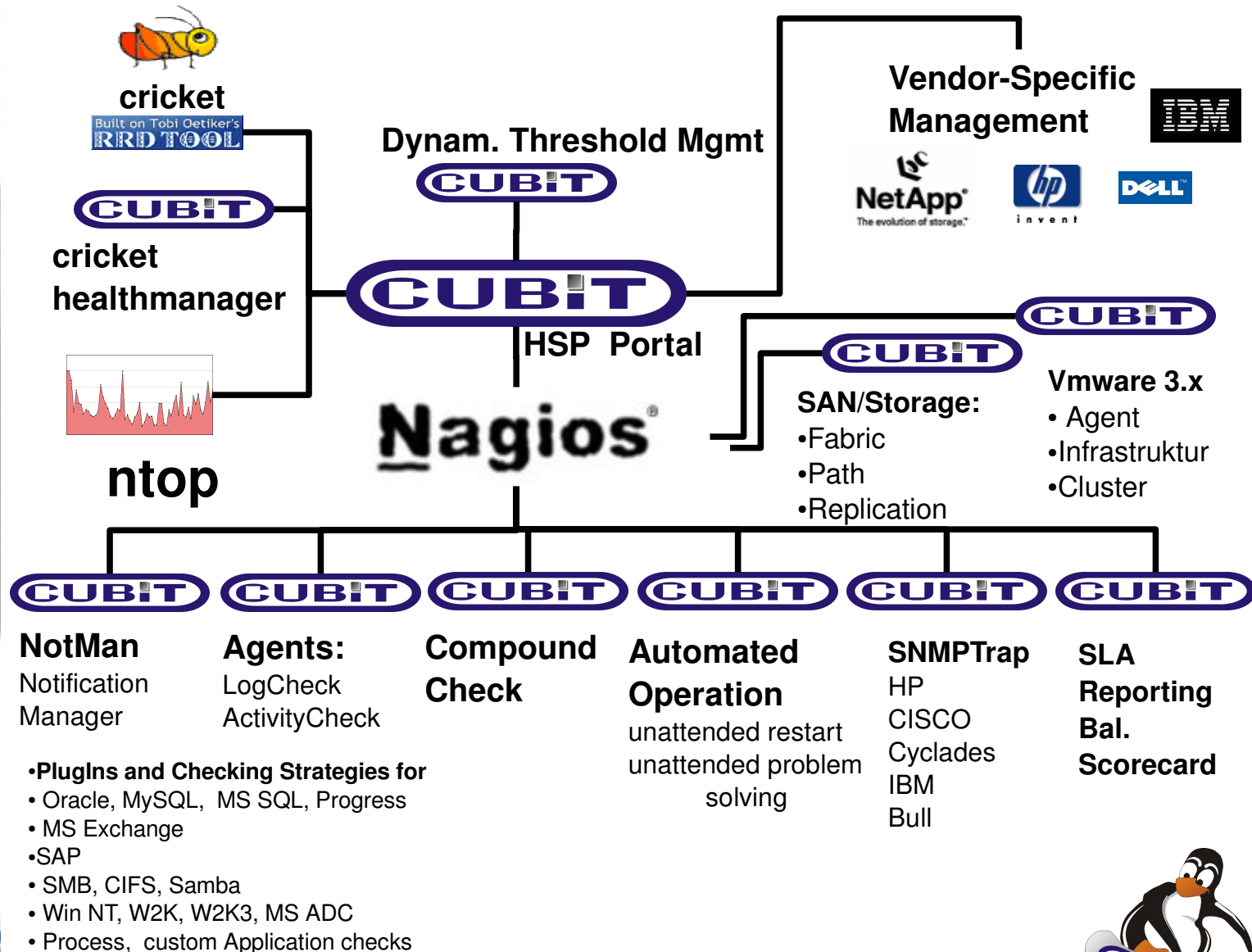
# Strategien – Positionierung

Networking

Datacenter

Storage Systems

Redundancy



# Monitoring Komponenten

Networking

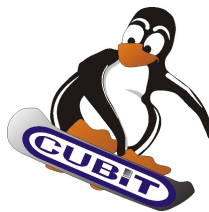
Datacenter

Storage  
Systems

Redundancy



- Nagios
- Cricket
- Integration von lokal laufenden Agenten: logcheck, activitycheck
- Integration von Syslog, SMTP, und anderen Diensten mit NSCA
- Einbindung SNMP Traps: trapreceiver
- Compound Checks
- HSP Portalsystem
- Alerting via SMS, Email, ICQ über Nagios
- ntop und rrd für Statistikdaten



# Nagios Features

Networking

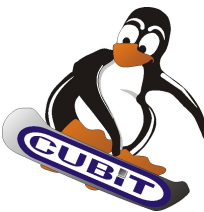
Datacenter

Storage  
Systems

Redundancy



- Nagios Core Process zentral
- führt regelmäßig Plugins aus und wertet Ergebnis aus
- empfängt passive Alerts
- Status-Änderungen lösen Events aus
- Events können gehandelt werden (default ist Notify)
- Notifizierungen werden wiederholt solange sie lt. Konfiguration wiederholt werden sollen
- Eskalation bei Notifizierung möglich



# Nagios PlugIn Strategy

Networking

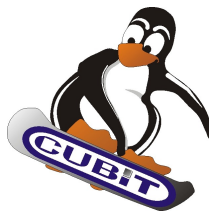
Datacenter

Storage  
Systems

Redundancy



- führt regelmäßig Plugins aus und wertet Ergebnis aus
- Plugin bewertet Situation
- Nagios routet nur PlugIn Output, keine Umformung
- -> PlugIn muss Fähigkeit haben Situationen zu bewerten!
- kaum eine Installation ohne Custom-made PlugIn
- PlugIn-Entwicklung nicht Sonderfall sondern Standard
- PlugIn entspricht Parametrisierung anderer Systeme
- keine Angst vor PlugIns!



# Komponenten: Nagios (2)

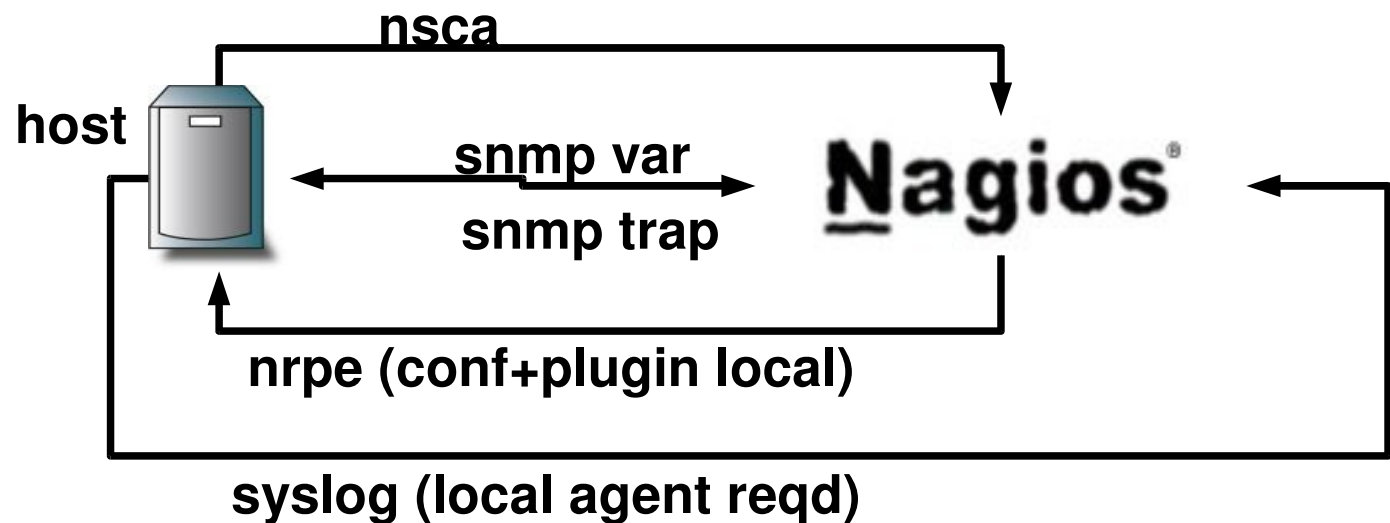
## Schnittstellen

Networking

Datacenter

Storage Systems

Redundancy



- NSCA: Schnittstelle mit der anderes Programm einen Passive Alert ins Nagios zur Weiterverarbeitung senden kann. Wird extern angestoßen.
- NRPE: Schnittstelle, mit der Nagios Plugins (zur Feststellung der Systemverfügbarkeit) auf einem entfernten System gestartet werden können. Die Ausgabe und Prüfung erfolgt zentral im Nagios Core; wird von Nagios aus gestartet





# Komponenten: Cricket

Networking

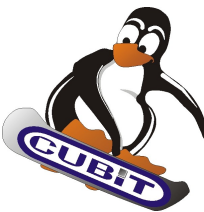
Datacenter

Storage  
Systems

Redundancy



- altbekannter Vorgänger: MRTG
- Trennung Datenbank RRDTool und Präsentation (Cricket)
- entwickelt von Tobias Oetiker
- Speichern und Anzeigen von Messwerten; je weiter zurückliegend umso geringere Auflösung
- Messwernerfassung per SNMP oder anders
- Bsp Apache server-status
- Echtzeiterfassung notwendig!
- dynamische Schwellwert - Monitore



# Komponenten: Cricket

Networking

Datacenter

Storage Systems

Redundancy



## Graphs for Proc-Counter (apache-ssl) (/server/cube1/proc-apache-ssl)

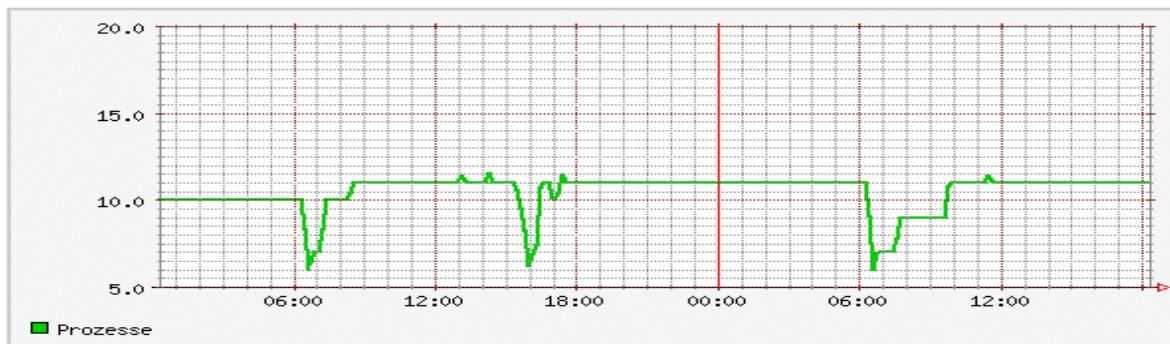
### Summary

Values at last update:  
**Prozesse** (for the day):  
**Cur:** 11.00  
**Avg:** 10.63  
**Max:** 11.79  
 Last updated at Thu Apr 24 18:16:06 2003

### Time Ranges:

- [ Daily ]
- [Weekly](#)
- [Monthly](#)
- [Yearly](#)
- [Short-Term](#)
- [Long-Term](#)
- [All](#)

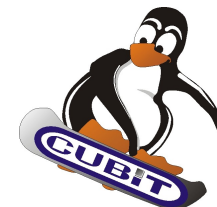
### Daily graph



**Cricket**  
 Version 1.0.3

Bei Fragen zu den Grafiken kontaktieren Sie bitte  
[support@cubit.at](mailto:support@cubit.at)

Built on **Tobi Oetiker's**  
**RRDTOOL**



# Komponenten: Cricket

Networking

Datacenter

Storage Systems

Redundancy



## Graphs for Overlay (/webserver/cubit-cluster)

### Summary

Values at last update for www-lga.cubit.at:

**Bearbeitete Zugriffe** (for the day):

**Cur:** 3.01 Zugriffe/s

**Avg:** 3.32 Zugriffe/s

**Max:** 8.66 Zugriffe/s

[ ? ]

Last updated at Thu Apr 24 18:17:08 2003

Values at last update for www-lgb.cubit.at:

**Bearbeitete Zugriffe** (for the day):

**Cur:** 0.09 Zugriffe/s

**Avg:** 0.00 Zugriffe/s

**Max:** 0.00 Zugriffe/s

[ ? ]

Last updated at Thu Apr 24 18:17:08 2003

*Time Ranges:*

[ Daily ]

[Weekly](#)

[Monthly](#)

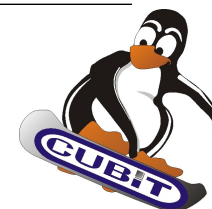
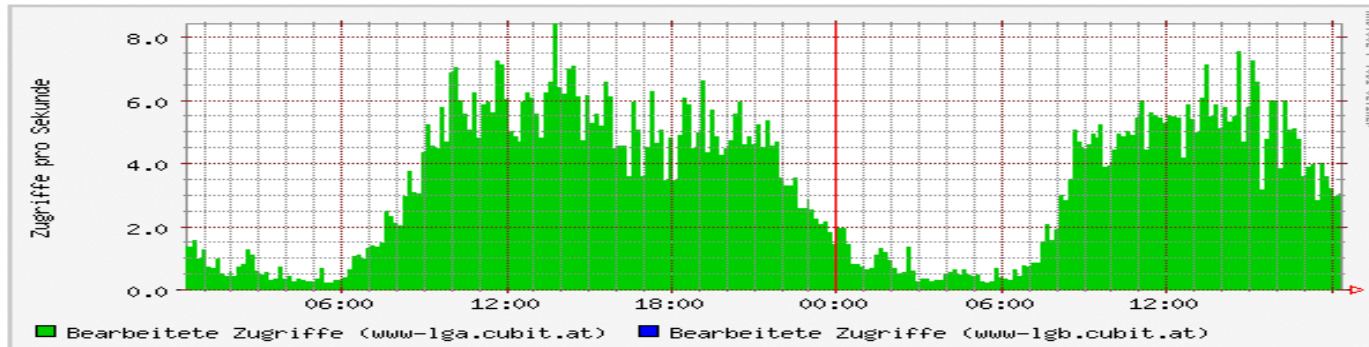
[Yearly](#)

[Short-Term](#)

[Long-Term](#)

[All](#)

### Daily graph



# Komponenten: Nagios Plugins

Networking

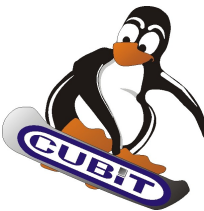
Datacenter

Storage  
Systems

Redundancy



- vielfältig im Internet vorhanden
- in definierten Zeitabständen vom Nagios Core Prozess aufgerufen
- laufen auf dem Nagios Rechner oder mittels NRPE verteilt
- Returnwert im Nagios verarbeitet:  
4 Stati: OK, Warning, Critical, Unknown
- viele Standardprotokolle (Ftp,nfs, http,...) bereits abgedeckt
- neue Plugins sehr leicht erstellbar
- Migration von MON-Scripts  
z.B. einfach möglich
- Angebot: <http://www.cubit.at/?nav=produkte>



# HSP Portalsystem

Networking

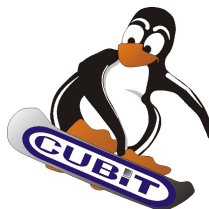
Datacenter

Storage Systems

Redundancy



- Web-basiertes Portal (Apache/MySQL)
- verlinkt zu nachgelagerten Systemen
  - Vendor-SM, Ntop, Cricket, SSH, VNC,...
- Wartungs- und Zusatzinformationen
  - Wartungsverträge
  - Kontakte
  - Logbücher
- Hardware Profile Inventory
  - MAC-Adressen, CPU, Memory, Disks



# HSP Portalsystem (2)

Networking

Datcenter

Storage Systems

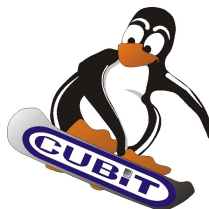
Redundancy



CPU-Informationen					
vendor_id	cpu_family	model_name	cpu_mhz	cache_kb	Optionen
GenuineIntel	15	Intel(R) Xeon(TM) CPU 2.40GHz	2399.397	512 KB	

Netzwerk-Informationen			
interface	inet_addr	mac_addr	Optionen
eth1	172.23.64.68	00:0B:CD:37:4B:0E	
eth1	172.23.64.66	00:0B:CD:37:4B:0E	
eth3	172.24.128.66	00:05:5D:7D:2B:4D	
eth0	10.7.0.1	00:0B:CD:37:4B:77	

Festplatten			
filesystem	size	mounted on	Optionen
/dev/cciss/c0d0p7	6015880	/home	
/dev/cciss/c0d0p1	197546	/boot	
/dev/cciss/c0d0p3	1521984	/	
/dev/cciss/c0d0p2	5039856	/usr	
/dev/cciss/c0d0p5	20159916	/var	
/dev/sda1	10080488	/var/lib/mysql	



# Notification Manager „NotMan(n)“

Networking

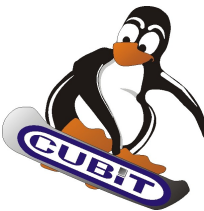
Datacenter

Storage  
Systems

Redundancy



- grafisches Abonnieren von Alarmen
- Aussendung in Echtzeit an alle abonnierten Empfänger
- Kriterien: Uhrzeit, Medium (SMS, Email, Voice),  
Wochentag, Feiertag, Alarmklasse,  
Systemgruppen
- freundlicher Assisten (auch für nicht-IT User)
- Addon Tool: keine Rekonfiguration von Nagios  
notwendig
- kein Restart von Nagios bei Änderungen notwendig
- MySQL basierend
- hohe Leistung, hoher Durchsatz
- geplant: Zeitzonen-Routing  
Voice IVR



# Event-Handling

Networking

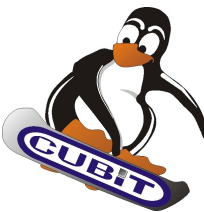
Datacenter

Storage  
Systems

Redundancy



- un-attended Operation
- automat. Reagieren auf Probleme
- Reduktion Service-Calls und Alarme um typisch 70%
- automat. Einhalten von Service-Profilen
- Event-Routing mit Notification Manager
- GUI zum Routen der Events
- Vertretungsfunktion, Schablonen, Berechtigungen





# Integration

Networking

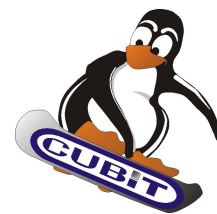
Datacenter

Storage  
Systems

Redundancy



- via HSP können Webtools eingebunden werden
- keine Ersatz für Management Tools (Switch Management, Server Management)
- zentrale Kommandozentrale
- Nagios Core Process als Information Hub
- Duale Information
- spezielle Systeme (Switch Management etc.) berichten an Nagios



# Danke

Networking

Datacenter

Storage  
Systems

Redundancy



- LINUX ist Datacenter-fähig
- Adaptionen notwendig wie bei jedem System
- System Integrator ist wichtig – er macht aus dem Rohprodukt eine betriebsfertige Lösung
- Mehr Information und Folien unter <http://www.cubit.at/>

