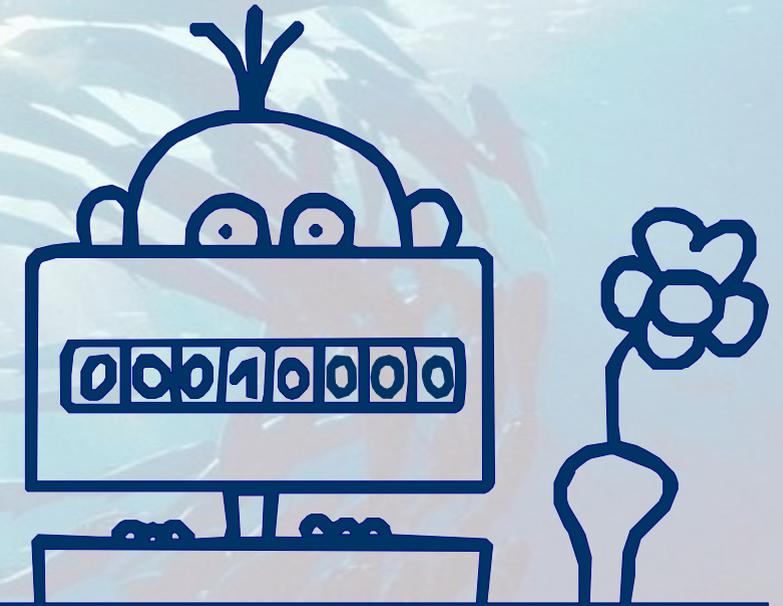


Das softwaredefinierte RZ für heute und morgen



OSL Technologietag

Schöneiche / Berlin • 14. November 2018

Cloud und Digitalisierung ohne Ende

Wo stehen wir heute?

Versuche einer Bewertung



Das Thema Cloud ist für Europa (und andere) im Wesentlichen verloren

aber

Das eigenbetriebene RZ und der Hoster vor Ort haben eine Perspektive:

- wenn agiler, flexibler, günstiger als die Cloud
- wenn Business Intelligence geboten wird → KnowHow und Verfahren kundenspezifisch
- wenn eigene Applikationen betrieben werden bzw. Standardapplikationen veredelt werden
- als Vertrauensinstanz vor Ort

Denken Sie selbst! Sonst tun es andere für Sie ... (Vince Ebert)

das beste Komplettpaket für eigenbetriebene IT-Infrastrukturen

- Virtual Storage, Virtual Server, Virtual Network, Clustering, Betriebsführung
- Vollkommen neue Möglichkeiten
- Einfachste Bedienung
- Eigene Technologie ergänzt um langfristig stabile OpenSource-Komponenten
- Open Systems (Linux, Solaris ...?)
- Klare Fokussierung auf Entwicklung
- Langfristiger Support

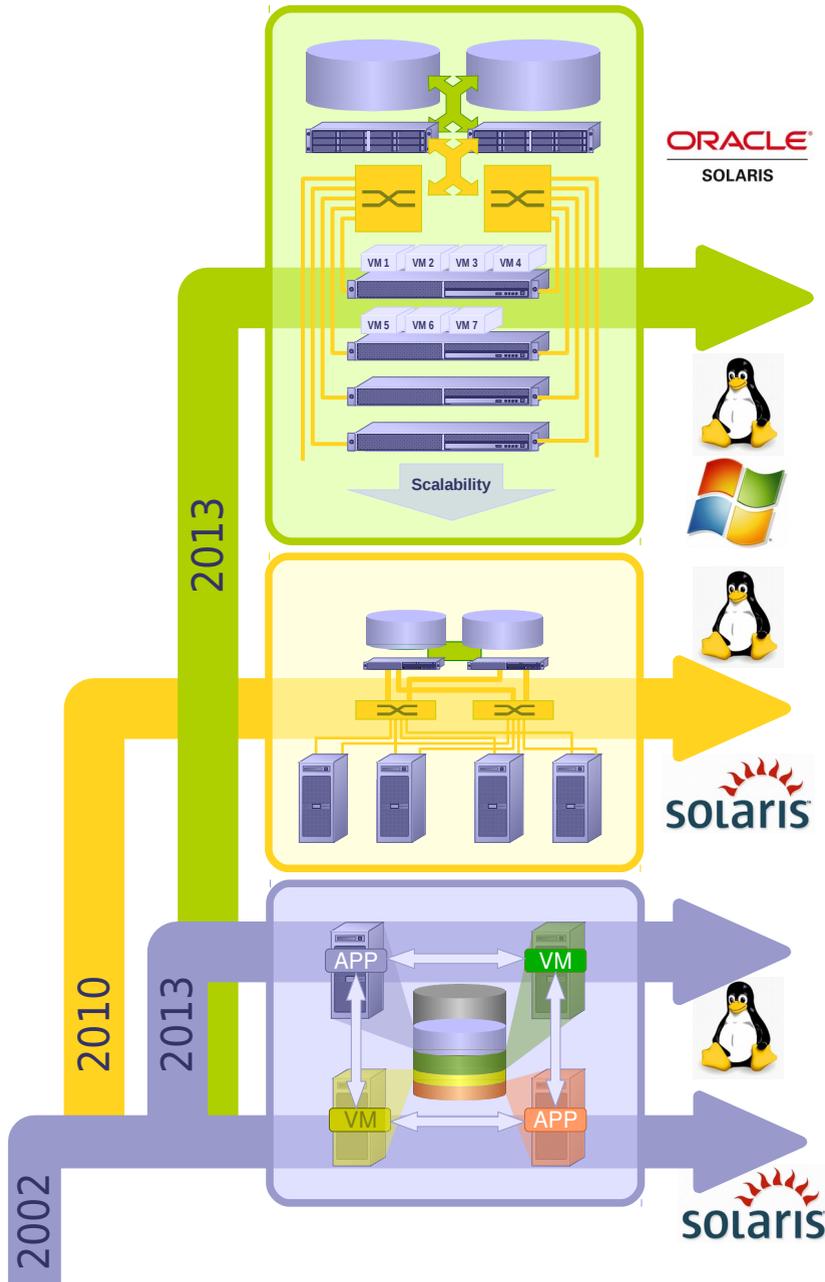
Damit

- Zukünftig: Konsequente Konzentration auf Technologieplattform
- Bedeutung OSL-Community / Anwenderaktiv wird für Tools und Utilities zunehmen
- Wir wollen die Zukunft *gemeinsam* mit Anwendern und Partnern gestalten

OSL 4.x - unser Portfolio

Unsere Produktfamilie

IT-Infrastrukturen "im Griff" haben



OSL Unified Virtualisation Environment:

V³ = Storage • Network • Server virtualisiert über ein Unified Network

Softwaredefinierte, flexible, hochverfügbare VM-Infrastruktur

OSL RSIO:

Data Center Block I/O over Ethernet

OSL Storage Cluster:

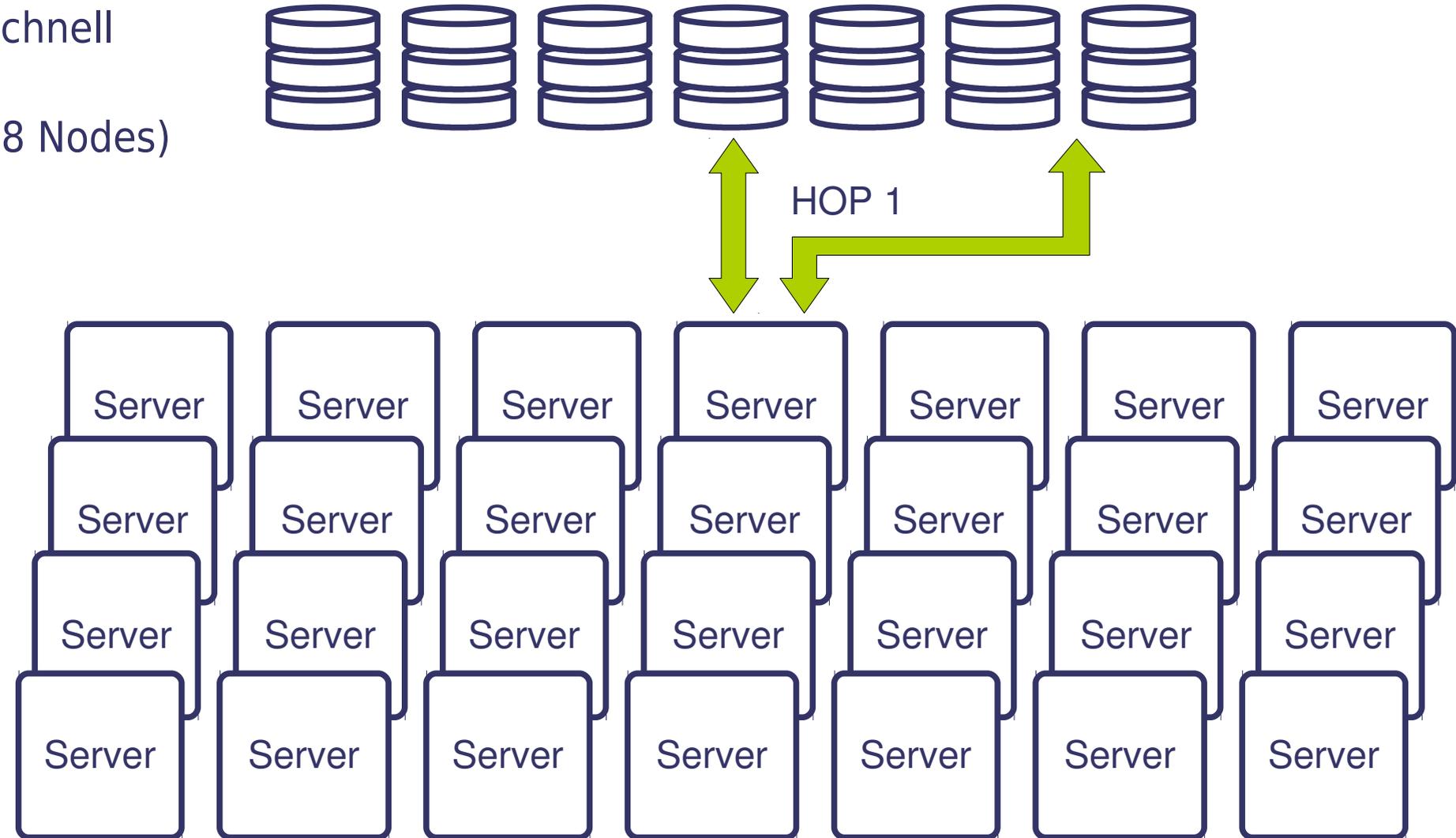
Hostbasierte Speichervirtualisierung und Clustering sowie viele Lösungen für flexible RZ-Prozesse in einem Produkt

OSL Storage Cluster 4.x

Speichervirtualisierung und Clustering in Einem



- Speichervirtualisierung
hostbasiert – clusterfähig – schnell
- Hochskalierbarer Cluster (128 Nodes)

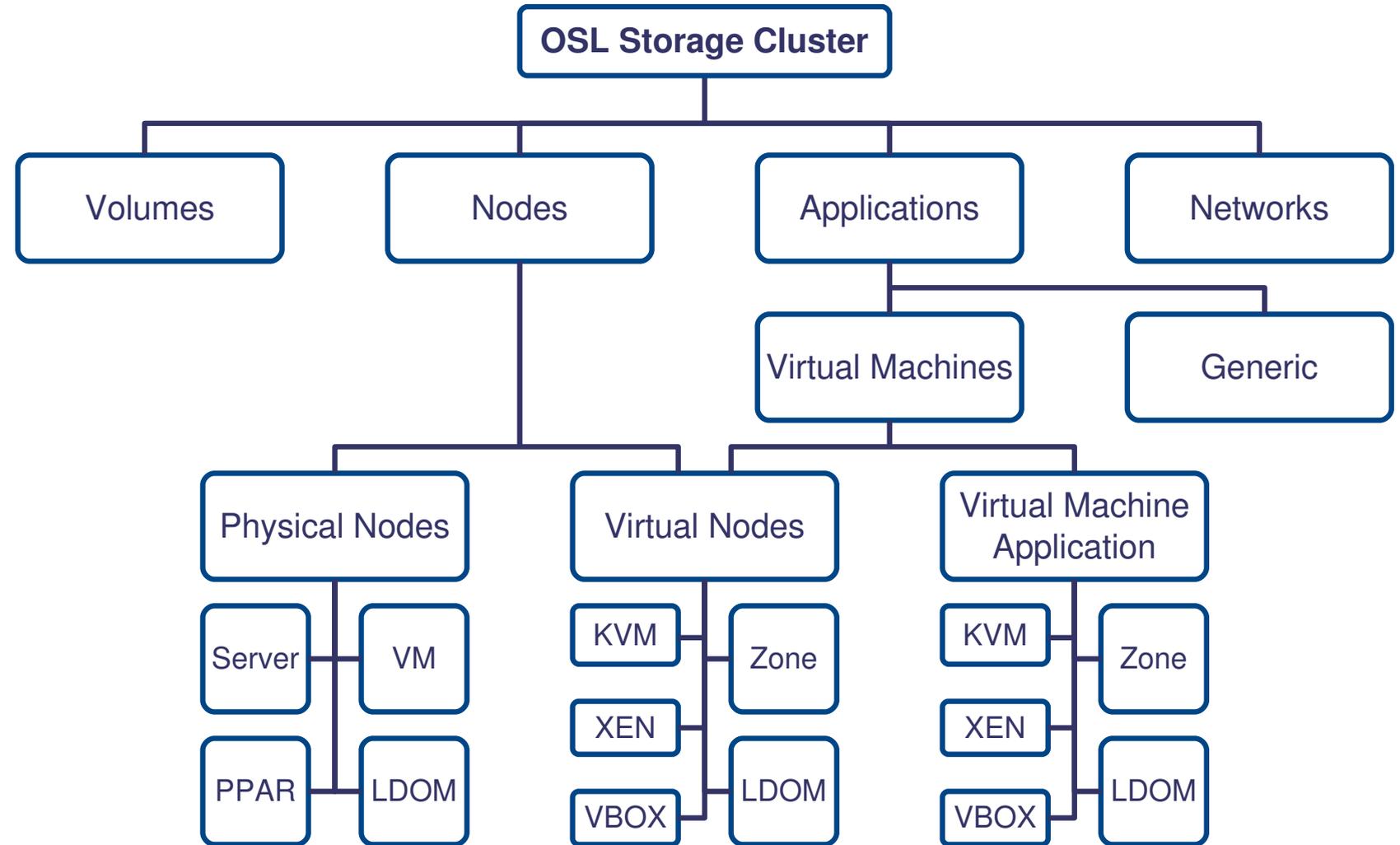


OSL Storage Cluster 4.x

Speichervirtualisierung und Clustering in Einem



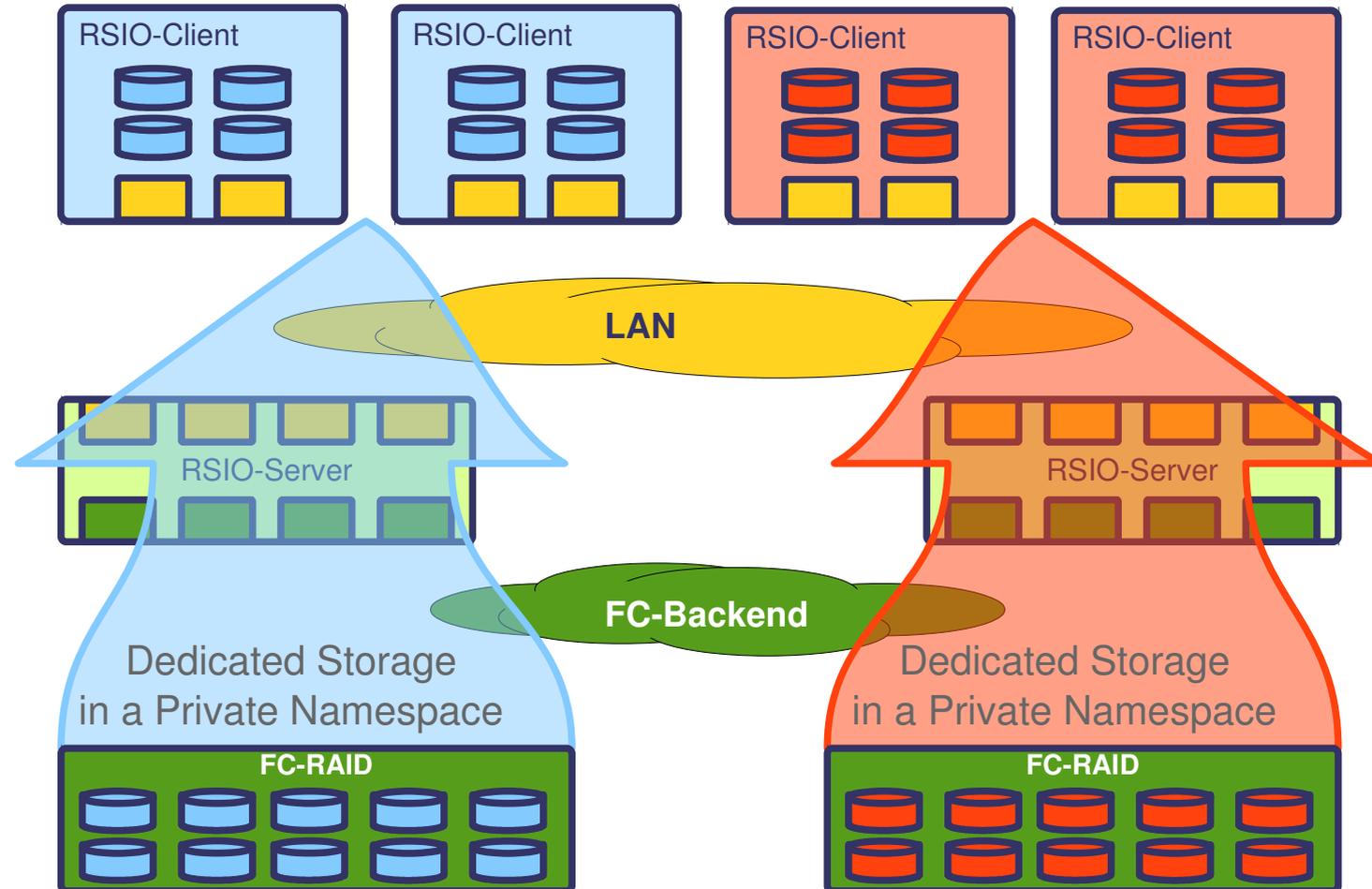
- Speichervirtualisierung
hostbasiert – clusterfähig – schnell
- Hochskalierbarer Cluster (128 Nodes)
- HA/Selbstmanagement eingebaut
- Unified Networking / Plattform-Mix
- Applikationsbewußtsein
(Storage, Ressourcen, VNO)
- Integration virtueller Maschinen



- Integriert mit OSL Storage Cluster und dessen Speichervirtualisierung

oder als

- Simple RSIO (standalone) ohne Lizenzkosten



Das OSL Unified Virtualisation Environment

Unsere HCI in Client-Server-Architektur



Unified
Virtualisation

UVS

Server

**Single Point of
Data Centre Definition & Administration**

**Central Point of
Infrastructure Service Delivery**

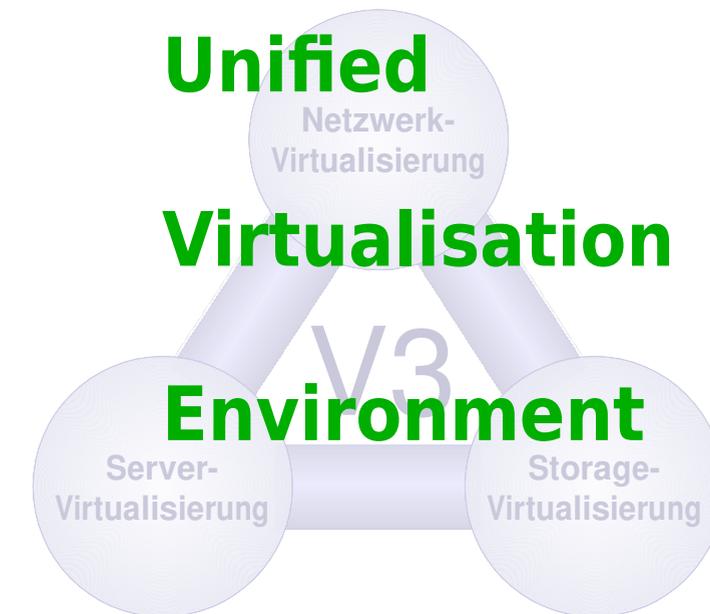
Converged Networking

Unified
Virtualisation

UVC

Client

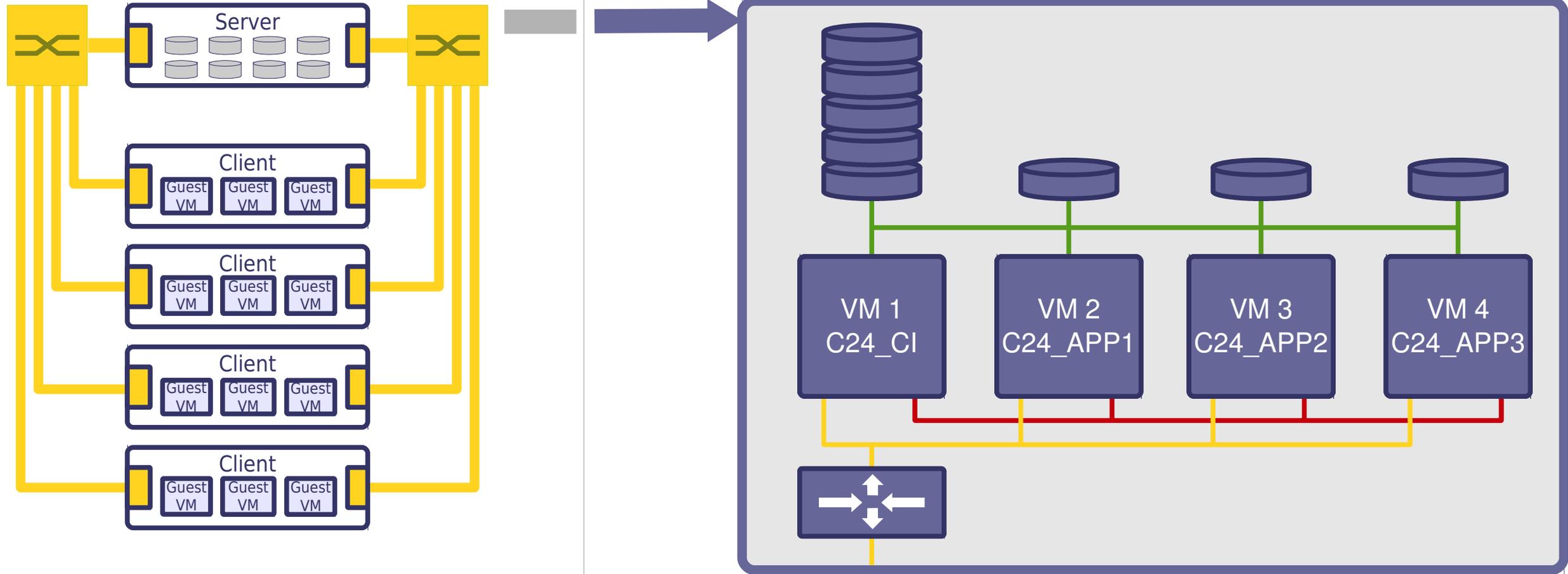
**Compute Node Farm
VM-Execution**



SDDC as a hyper-converged
client-server infrastructure

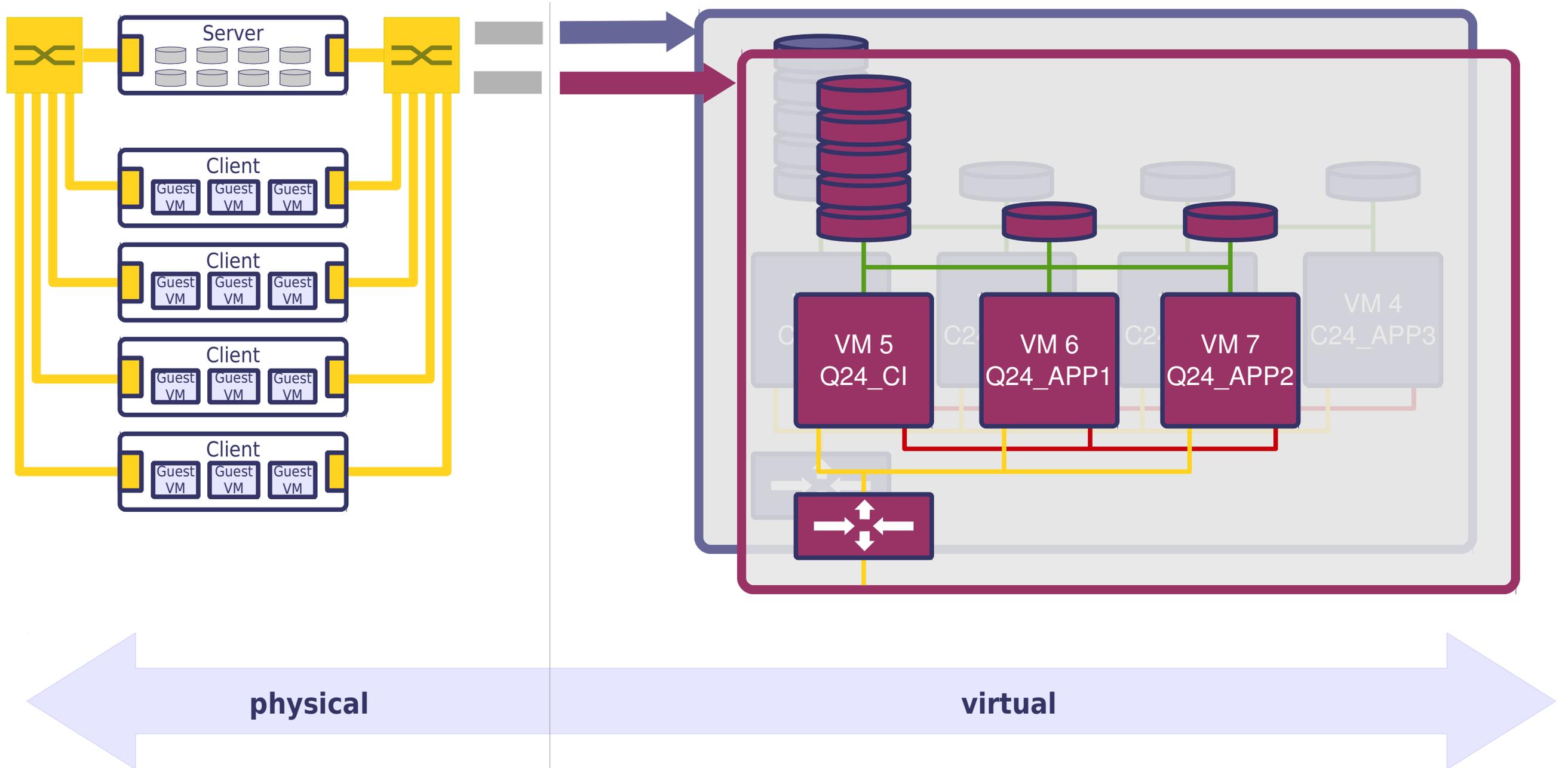
Was ist anders mit OSL UVE ?

Ein Beispiel: Neue Infrastrukturen per Software



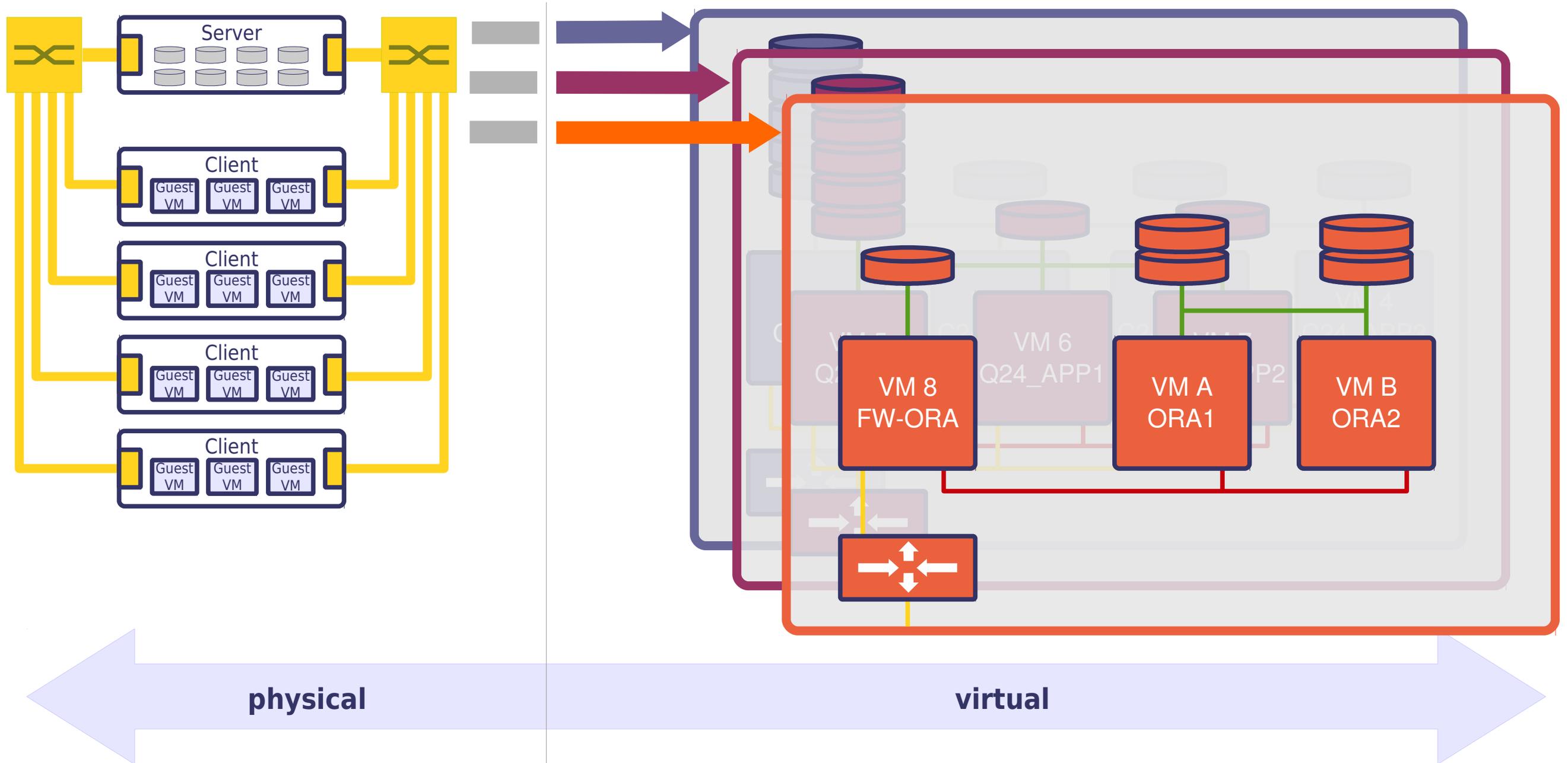
Was ist anders mit OSL UVE ?

Ein Beispiel: Neue Infrastrukturen per Software



Was ist anders mit OSL UVE ?

Ein Beispiel: Neue Infrastrukturen per Software



Wie stellt sich die Handhabung dar ?

Wie ein einziges System – Beispiel OSL Unified Virtualisation Environment



The screenshot displays the OSL Unified Virtualisation Environment interface. The main window shows a list of virtual machines with columns for Prio, Nickname, State, Target State, Node, #CPUs, Type, and #ready. The VMs are listed as follows:

Prio	Nickname	State	Target State	Node	#CPUs	Type	#ready
101	sles-1@0	started	no control	mc1	2	Linux KVM	15
102	sles-2@0	started	no control	mc2	2	Linux KVM	15
103	sles-3@0	started	no control	mc3	2	Linux KVM	15
104	sles-4@0	started	no control	mc4	2	Linux KVM	15
105	sles-5@0	started	no control	mc5	2	Linux KVM	15
106	sles-6@0	started	no control	mc6			
107	sles-7@0	started	no control	mc7			
200	centos@0	started	no control	mc1			
201	centos-1@0	not started	no control	-			
202	centos-2@0	not started	no control	-			
203	centos-3@0	not started	no control	-			
204	centos-4@0	started	no control	mc12			
205	centos-5@0	not started	no control	-			
206	centos-6@0	not started	no control	-			
207	centos-7@0	not started	no control	-			

The interface also includes a sidebar with navigation options (Storage, VMs, Network, Nodes, Settings) and a control panel with buttons for Create, Start, Stop, Reset, Pause, Continue, Migrate, Set Target State, Delete, and Console. The status bar at the bottom indicates "Action status: Running: 0 Successful: 3 Failed: 2".

In the foreground, a terminal window titled "TigerVNC: QEMU (sles-5@0)" shows the output of the command `lsblk` on a SUSE Linux Enterprise Server:

```
sles-5:~ # lsblk
NAME MAJ:MIN RM SIZE RO MOUNTPOINT
vda  253:0  0  10G  0
├─vda1 253:1  0  1.5G  0 [SWAP]
└─vda2 253:2  0  8.6G  0 /
sr0  11:0  1 1024M  0
sles-5:~ #
```

Below the terminal, a window titled "centos-2" shows a graphical installer or boot screen with the OSL logo and a text input field labeled "Other...".

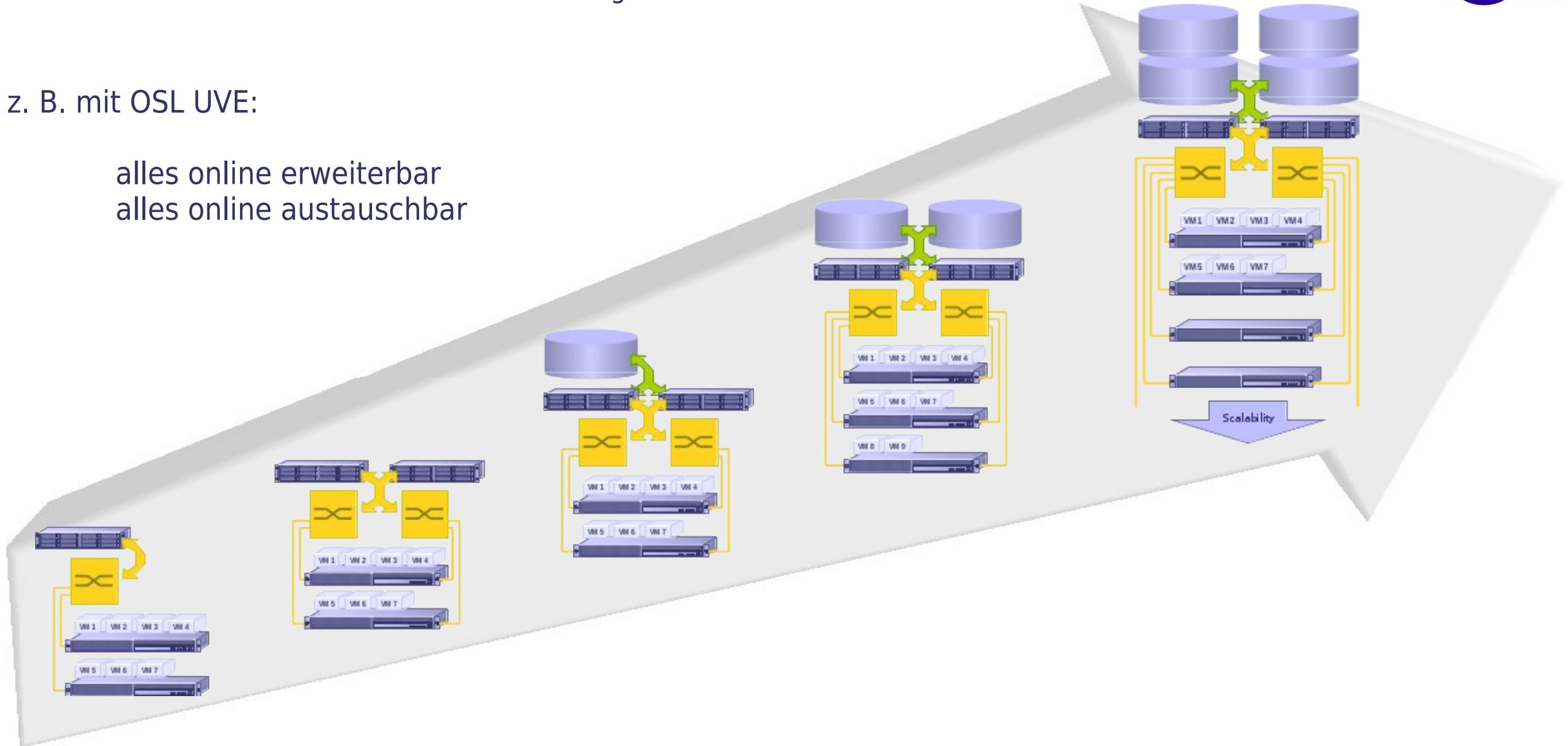
Nicht nur Vereinfachung und Kostensenkung

Softwarezentrierter Ansatz resultiert in neuen Möglichkeiten



z. B. mit OSL UVE:

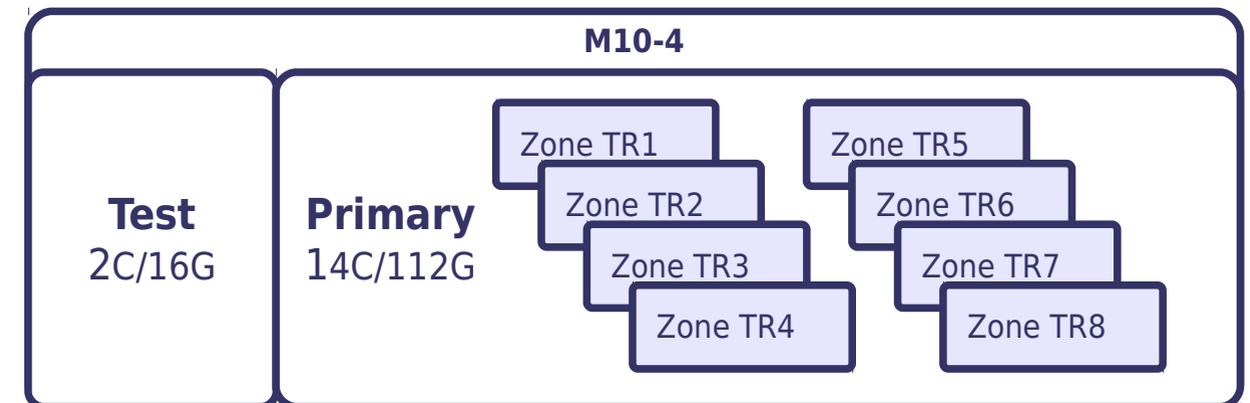
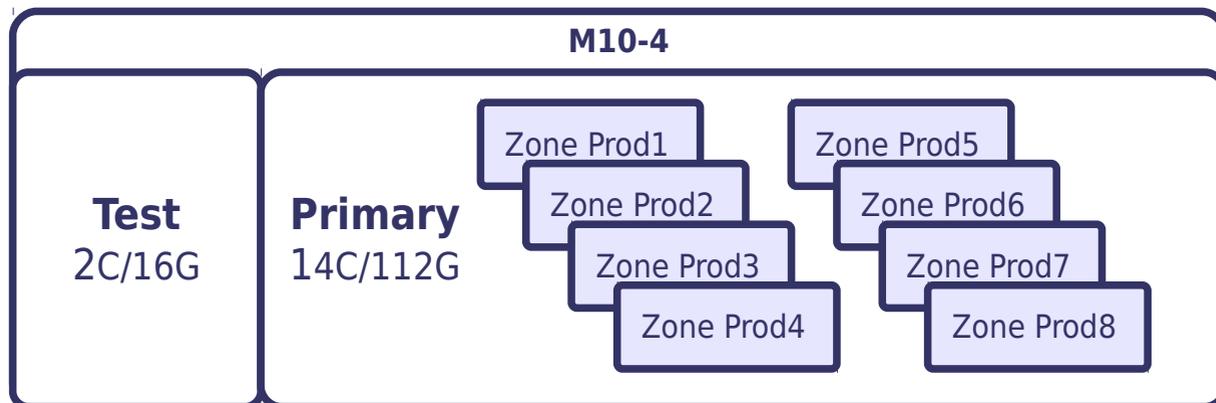
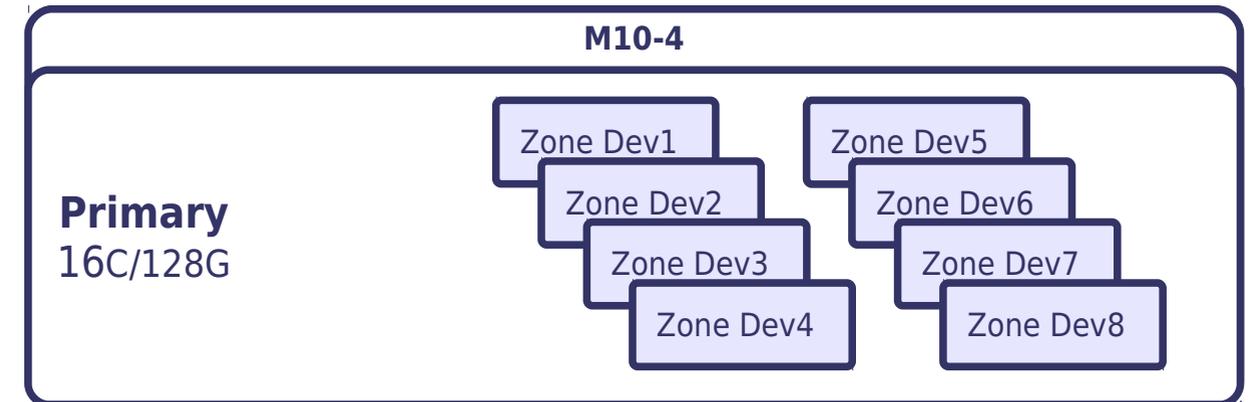
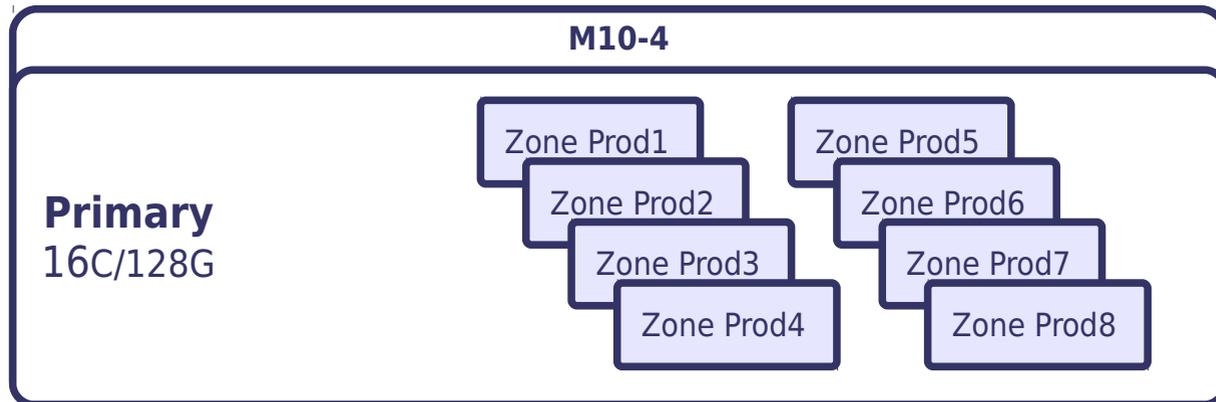
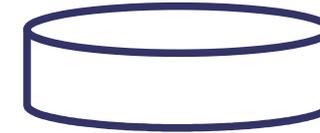
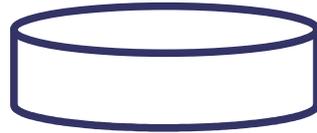
alles online erweiterbar
alles online austauschbar



Projekte

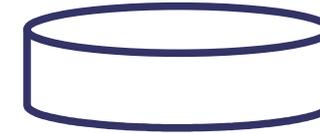
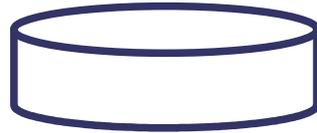
Projekt 1: Der Klassiker - OSL Storage Cluster

Hochverfügbare Applikationen mit DR-Fähigkeit (2 Standorte)



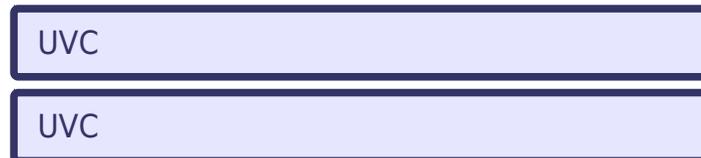
Projekt 2: Shared System für SAP und HCI

Hochverfügbarkeitscluster mit OSL SC und HCI mit OSL UVE in Einem

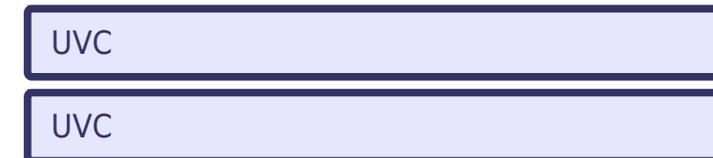


M10-4			
SAP 2C/32G		SAP 2C/48G	CT 4C/16G
NAS 1C/16G	UVS 1C/16G	SAP 4C/16G	

M10-4				
CT 4C/16G		SAP 2C/48G	SAP 2C/32G	
SAP 4C/16G			UVS 1C/16G	NAS 1C/16G

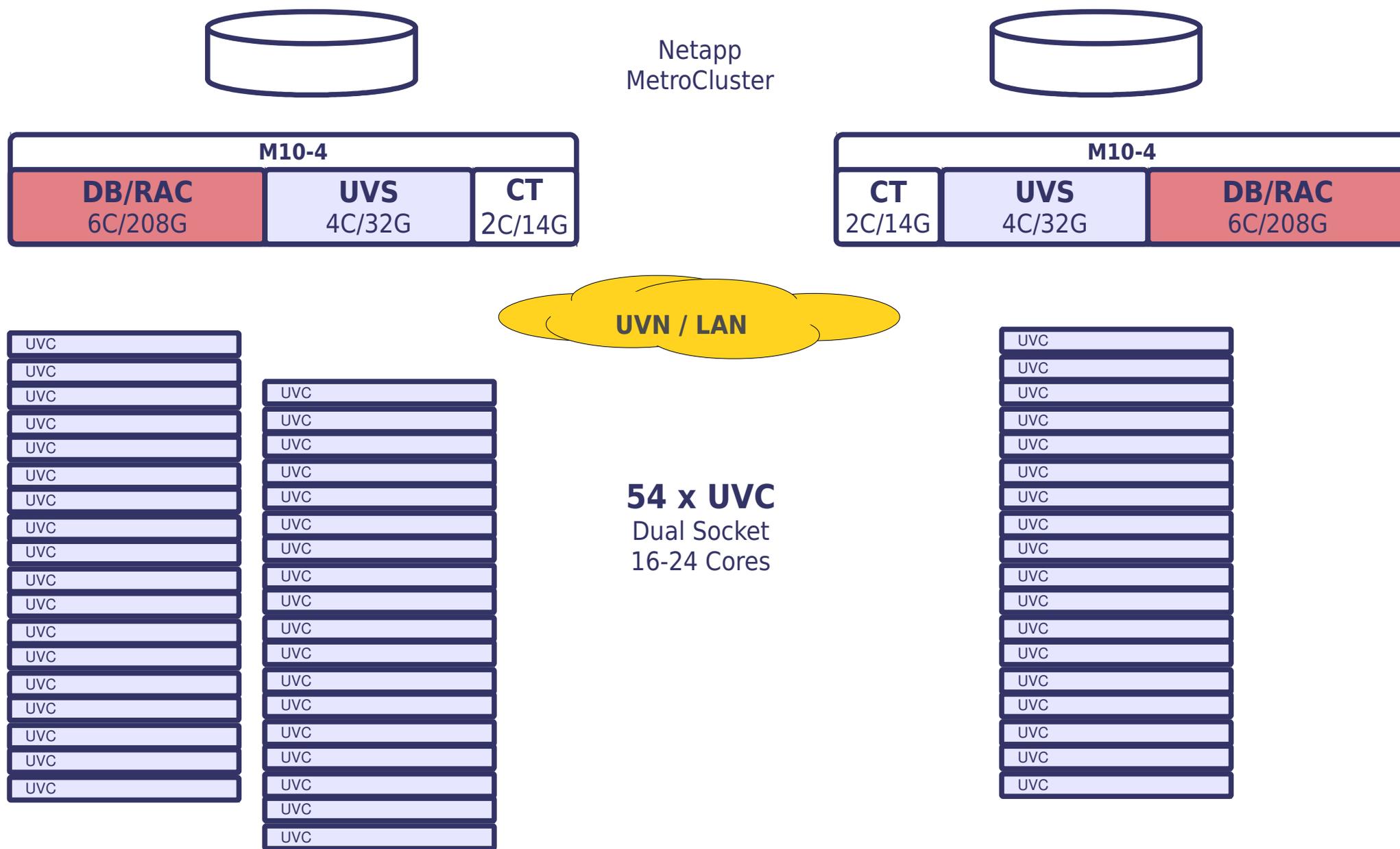


4 x UVC
Dual Socket



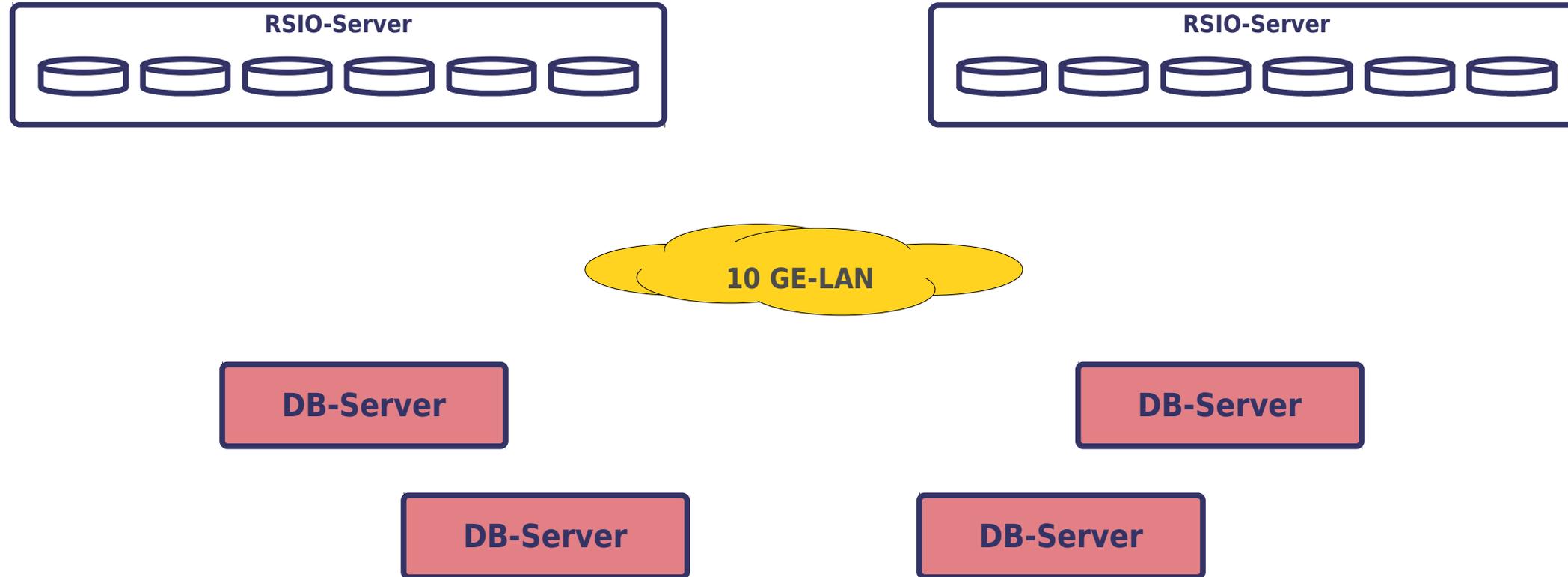
Projekt 3 - Dual-Site Non-Stop HCI-Cluster

Von Solaris bis Linux, von SPARC bis x86, von Storage bis Server - alles virtualisiert, alles integriert



Projekt 4 - Linuxbasierter RSIO-Server

Ablösung von RAID-Systemen durch RSIO-Server mit internen SSDs



Steigerung der IOPS in allen Lastprofilen bei niedrigerem TCO
DR-Tauglichkeit in Verbindung mit OSL Storage Cluster auf den DB-Servern

Neuheiten

Was ist im letzten Jahr passiert ?

Reaktion auf Kundenanforderungen und Projekte



- Neues, durchgängiges Versionierungsschema MAJOR.MINOR.LOAD z. B. 4.3.08
- Solaris-spezifische Anpassungen
 - Anpassungen an striktes devfs-Management im neuen Solaris 11.3, devlnktb-Utility
 - deny-zbe-clone, wenn möglich
 - Shared R pools für Solaris-Zonen
- Portierung auf SLES 12.0, 12.1, 12.2 und Unterstützung neuer Hardware
 - CPU-Typen (IvyBridge, Haswell, Broadwell, Skylake) und Systeme mit > 100GByte RAM
- Unterstützung sehr großer LUNs (EFI) durch RSIO
 - benötigt für RSIO Clients und UVE-Kontexte (Remote PVs im Slim UVS)
- Verborgene NPC-Paßwörter
- UVE-spezifische Änderungen bzw. Erweiterungen
 - uvcmanage durch uvc ersetzt, selektives Routing und Firewall-Funktionen, bis zu 96 Volumes je VM
- UVE auf Infiniband-Infrastruktur (IPoIB, VXLAN)
- pvimport-Utility
 - Für Übernahme von Volumes aus alten Versionen und für das Verschieben zwischen Clustern

Was ist im letzten Jahr passiert ? (Fortsetzung)

Reaktion auf Kundenanforderungen und Projekte

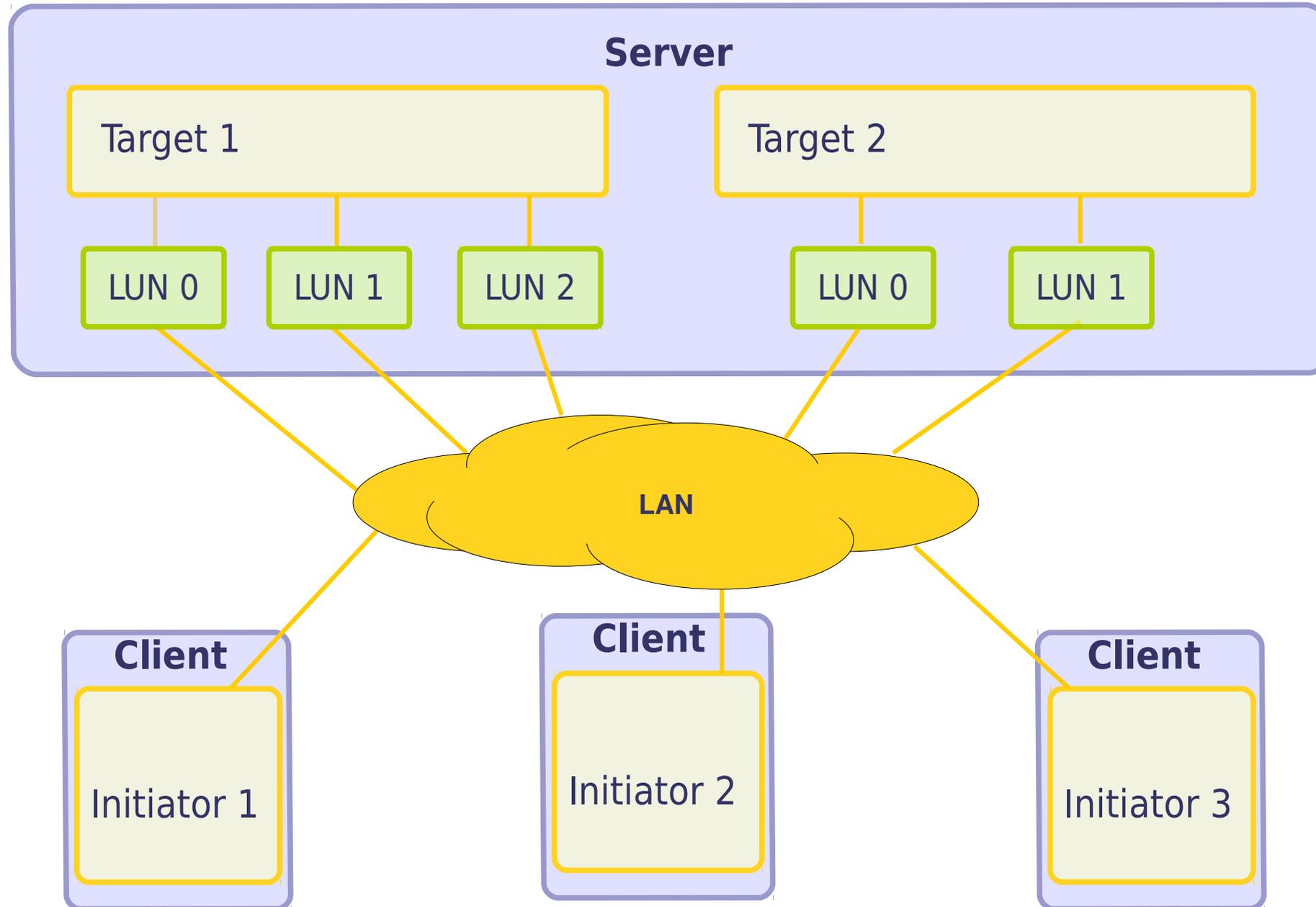


- WebGUI für OSL SC
- Abschluß Design HANA-Referenzarchitektur und Beginn der Implementierung
- Weiterentwicklung VSD
- Konzeptstudie Kernel Zones im OSL Storage Cluster – Pilotierung
- Neuimplementierung iSCSI-Server im OSL Storage Cluster für Solaris 10 und Solaris 11
- Neuimplementierung SAMBA-Server im OSL Storage Cluster (mit und ohne ADS)
- Organisationseinheiten sowie neues Rollen- und Nutzerkonzept für UVE
- Entwicklung Kommunikationsbibliotheken für IB
- Neues, portables Performancemonitoring für SC + UVE
- Applikationsjournale
- Erweiterungen der Web-Interfaces für SC und UVE
- Erweiterung Memdisk-Fähigkeiten im Simple RSIO-Server

Eine Auswahl von Neuerungen im Detail

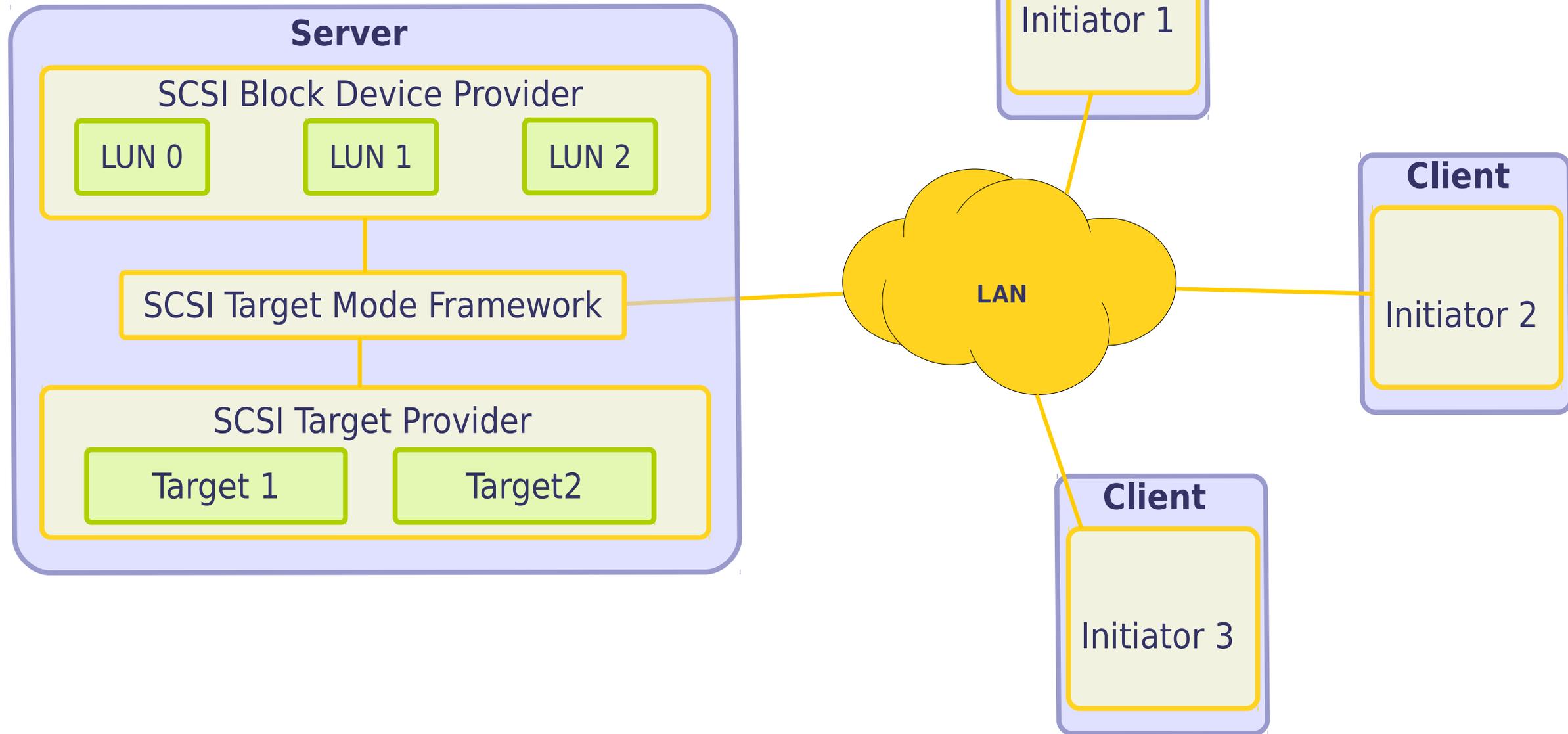
Hochverfügbarer iSCSI-Server

Die Umgebung unter Solaris 10



Hochverfügbarer iSCSI-Server

Die Umgebung unter Solaris 11 mit COMSTAR



Hochverfügbarer iSCSI-Server



Die abstrahierte Darstellung im ACO-Konfigurationsfile (einheitlich unter Solaris 10 & 11)

```
CFGDIR      /dvsc/iscsi_conf
ALIAS       iqn.1986-03.com.sun:01:fab807f0ffff.49d1f23f big-1
ALIAS       iqn.1986-03.com.sun:01:fab807f0ffff.49d1f672 big-2
ALIAS       iqn.1986-03.com.linux:99:gurke.big-3          big-3

TARGET      server1
LUN 0       server_lun0@0
LUN 1       server_lun1@1
INITIATOR   big-1
INITIATOR   big-2

TARGET      iqn.1986-03.com.sun:02:97e:1234.client1
LUN 0       client_lun0@0
INITIATOR   big-3
```

Verzeichnis für iSCSI Konfigurationsdateien

Initiatoren werden Aliase zugeordnet

aus Target-Namen automatisch IQN generieren

Zuweisung von LUNs zu einem Target

Initiatoren mittels Alias einem Target zuweisen

explizite Zuweisung einer IQN

Hochverfügbarer iSCSI-Server

Eigenschaften der Lösung



- Plattformübergreifend einfach zu verwalten über abstrakte Konfigurationsdatei
 - Einfache Zuordnung von Targets und Initiatoren (Aliase)

- Umschalten und Konfigurationsänderungen im laufenden Betrieb und ohne I/O-Verlust
 - Persistente GUIDs und IQNs
 - “wandernder” Service

Hochverfügbarer SMB/CIFS-Server

Eigenschaften der Lösung



- **SAMBA als Standalone-Konfiguration**

- Bei Nutzung des Clusterweiten Benutzer- & Gruppenmanagements hochverfügbar
- Benutzerauthentifikation über Unix-User
- Einfache Konfiguration über die Samba-Konfigurationsdatei

- **SAMBA als Active-Directory-Mitglied**

- Benutzerverwaltung über einen AD-Server in der Domäne
- Benutzer können mit Ihrer Active Directory Authentifizierung auf Unix Ressourcen zugreifen
- Freigaben & Domänenparameter werden einfach über die Samba-Konfigurationsdatei verwaltet
- Kerberos ist unter Solaris 11 in das Samba Paket integriert und benötigt keine eigene Konfiguration
- Hochverfügbar, da die Applikation unter Ihrer virtuellen IP Adresse ins Active Directory integriert wird

Organisationseinheiten I

Struktur von Ressourcen und Berechtigungen im Überblick



- Kontrollierte Delegation von bestimmten Berechtigungen
- Einschränkung der Sichtbarkeit auf bestimmte Objekte
- Hierarchien in Organisationseinheiten darstellbar
- Verknüpfung mit der Nutzerverwaltung
- Verknüpfung mit Applikationen und VMs
- Umsetzung für OSL UVE und für OSL SC

Organisationseinheiten II

Darstellung von Organisationseinheiten und OE-Muster



- Abbildung der Struktur eines Unternehmens bzw. einer Behörde
- max. 3 Ebenen mit bis zu 1023 IDs
 - Darstellung ähnlich einer Kapitelnummerierung:
 - 1.
 - 1.2
 - 1.2.435
- OEs können Attribut einer VM werden:
 - eine VM kann zu genau einer OE gehören
 - zu einer Organisationseinheit können mehrere VMs gehören
- OE-Muster können für Abfragen verwendet werden:
25.* liefert 25, 25.3, 25.1.1, 25.1023.1023
- Umsetzung der OEs für OSL UVE und für OSL SC
- Im OSL UVE erhalten Benutzer Berechtigungen für bestimmte OEs

- Berechtigungen werden über Rollen verwaltet
 - Rollen fassen einen Satz von Berechtigungen zusammen
- Ein Benutzer erhält eine Rolle für bestimmte OEs (oder Muster von OEs)
- Es werden vordefinierte Rollen ausgeliefert

Rolle	Berechtigungen	Beschreibung
uveadm	alle	berechtigt für alle Aktionen
uvemon	Monitoring	Monitoring der gesamten UVE
ndmon	Monitoring	Monitoring aller UVE-Knoten
oadm	Verwalten von OEs	Erstellen, Bearbeiten, Löschen von OEs
vmadm	Verwalten von VMs	Komplette Kontrolle über VMs (Mirror, Edit, ..)
vmoper	Operating von VMs	VM installieren, CD einlegen, Konsole, Start/Stop
vmuser	Anzeigen von VMs	für Anwender einer VM , Statusinformationen

- Rollen können auch selbst sehr detailreich definiert werden

Organisationseinheiten IV

Berechtigungen und Benutzer in der vierten Dimension



- Benutzern können Berechtigungen temporär entzogen werden (enable/disable)

- Der Zugriff auf die Funktionen der WebGUI kann zeitlich limitiert werden, z. B.:

201805-201910
!(20180301-20180331)
1-5 * * 8-17 *

Beinhaltet Mai 2018 bis Ende Oktober 2019
Zugriff nicht in der Zeit vom 1.3. bis 31.3. 2018
Montag bis Freitag von 8-17 Uhr (crontab-Logik)

- Die verschiedenen Verfahren zur Zeitspezifikation sind kombinierbar

Weitere Details: OSL-Hinweis 20180004

Benutzerverwaltung in der OSL SC-WebGUI

Ein einfacheres System mit Fokus auf simple Administrationsaufgaben und Monitoring



- Die WebGUI des OSL SC bietet eine simples Berechtigungskonzept über Benutzergruppen
- Vordefinierte Benutzergruppen
 - Monitor - Reine Überwachung des Systems
 - Operator - Starten, Stoppen, Verschieben, Spiegeln, Disaster Recovery
 - Admin - Zusätzlich zum operativen Teil auch Anlegen/Löschen von Instanzen
 - Superuser - Administrator mit Rechten zur Benutzerverwaltung

- Welche Applikation lief wann auf welchem Rechner?
 - Nachweismöglichkeit
 - Erkennbarkeit der Auslastung / Systemverwendung
 - Nachweis auch für Lizenz-Audits von Anwendungssoftware
- Kunden- bzw. anwenderorientierte Übersicht zur Verfügbarkeit
 - vereinfachte / endtechnisierte Darstellung
 - entscheidend ist die tatsächliche Verfügbarkeit
 - Aufbereiten der Statistik (99,99%...)
- Erfassung auch von Komplettausfällen
 - auch wenn der Cluster komplett aus ist, müssen diese Zeiten miterfaßt werden
- Langfristaufzeichnung
 - /var/adm/messages, journalctl ... werden irgendwann „herausaltern“
 - aggregierte Auswertungen sind nicht trivial
 - man sollte über einen längeren Zeitraum Aufzeichnungen zur Hand haben

- **Einbettung der Funktionen in das SC/VE-Framework**
 - neues DataStore in der CCF
 - Aufzeichnung in PCCF und SCCF
 - Erweiterung nahezu aller Programme in Base und ACO (Erfassen von Statusänderungen)
 - Lebendüberwachung durch "dvwatchd" ("dvamd" hat teilw. Langläuferaufgaben)
 - Zugriff nur über Funktionen der Clusterengine (schwer manipulierbar)

- **Bis zu 1024 Ereignissätze je Applikation**
 - z. B. Start, Anwendung online, Anwendung gestoppt, Anwendung ausgefallen ...
 - Differenzierung zwischen expliziten Kommandos und Steuerung durch Clusterengine
 - Knoten und Zeit identifizierbar (effektive Genauigkeit/Auflösung mdst. 60 Sekunden)
 - bei mehr als 1024 Einträgen werden die ältesten zugunsten neuerer Einträge automatisch entfernt

- **Auswertung per appadmin / vmadmin**
 - Ausgabe der Ereignissätze
 - aggregierte und Statistiksätze verfügbar

- **Empfehlung: Zeitsynchronisation im Cluster für plausible Statistiken**

Application Journals

Beispiel



```
root@sc43-11:~# appadmin -qj glldb@0
2018_10_25-09:18:48 (UTC) sc43 sc43-11 start initiated by by cluster engine
2018_10_25-09:18:49 (UTC) sc43 sc43-11 started (by cluster engine)
2018_10_25-09:34:51 (UTC) sc43 sc43-11 online (started by cluster engine)
```

summary statistics:

```
time covered by journal:          963s (0d 00:16:03)
controlled downtime:             1s (0d 00:00:01) 000.104%
failure time:                     0s (0d 00:00:00) 000.000%
online time:                       962s (0d 00:16:02) 099.896%
```

Performance-Monitoring mit ndperf

Die wichtigsten Performancedaten eines Rechnerknotens auf einen Blick



- CPU, Memory, Block-I/O und Netzwerk auf 80 Spalten
- Portable Implementierung für Solaris und Linux
- Sehr geringer Ressourcenbedarf (ca. 0,2 Promille einer CPU im Loop-Mode)

```
root@sc43-11:~# ndperf 2 5
```

	%cpu	%usr	%sys	%wio	%mem	%swp	rdMiB/s	wrMiB/s	rxMiB/s	txMiB/s
22:35:34	57.1	19.2	37.9	0.0	18.0	0.0	0.0	0.0	30.9	0.3
22:35:36	56.8	18.0	38.8	0.0	18.7	0.0	0.0	0.0	32.8	0.4
22:35:38	54.6	17.1	37.5	0.0	19.4	0.0	0.0	0.0	26.4	0.3
22:35:40	58.0	18.5	39.5	0.0	20.1	0.0	0.0	0.0	30.9	0.4
22:35:42	57.5	16.9	40.6	0.0	20.8	0.0	0.0	0.0	30.6	0.3

Der große Bruder: ndadmin & dvndperf

Die wichtigsten Performancedaten clusterweit auf einen Blick



- CPU, Memory, Block-I/O, Netzwerk uvm. clusterweit
- jetzt permanente Aktualisierung, d.h. stets Anzeige der aktuellen Werte
- hochauflösendes Live-Journal für 1 Stunde rückwärts (dvpdlogd)
- permanente Historie von Min - Avg - Max in speicheroptimierter Auflösung im Minutentakt!
- Integration in sc-wui bereits verfügbar

```
root@sc43-11:~# ndadmin -pv
```

```
nodename id state  cpu-isa  ncpu clock memory swap rip    rip_free  physfree swapped  %usr  %sys  %wio  %idle  rdMiB/s  wrMiB/s  rxMiB/s  txMiB/s
sc42-11   1 ONLINE  amd64    2   3408   8191 2047 150.21  149.46    5936     0     0.0   0.5   0.0  99.5   0.1     0.0     0.0     0.0
sc42-12   2 ONLINE  amd64    2   3408   8191 2047 156.92  156.92     0     0     0.0   0.0   0.0 100.0   0.0     0.0     0.0     0.0
```

Ausblick

Projekte

Die nächsten Vorhaben mit hoher Priorität



- HANA-Referenzarchitektur
- VSD / ovvemu
- Anwendungsreife Umsetzung SAP-HANA-Referenzarchitektur
- Neuimplementierung Kernel-Zone-Integration in OSL SC
- Native Infiniband-Unterstützung im OSL UVE
- Portierung OSL SC, OSL UVE, OSL RSIO für Solaris 11.4, SLES 15, CentOS 6/7
- Erweiterungen UVE VM-Administration (pass-through devices ...)
- OSL SC 4.9

Wir freuen uns auf das, was vor uns liegt!



virtualization and clustering - made simple