



OSL Whitepaper

SPARC-Server und hyperkonvergente x86-VM-Infrastrukturen

Copyright und Handelsmarken

Copyright © OSL Gesellschaft für offene Systemlösungen mbH 2015.

Alle Rechte vorbehalten.

Eine unveränderte Nutzung dieser Dokumentation ausschließlich für private oder interne Zwecke ist gestattet. Andere Nutzungsarten, gleich welcher Form, wie z. B. die Bearbeitung, Übersetzung oder Veröffentlichung dieses Dokumentes bedürfen einer ausdrücklichen vorherigen schriftlichen Genehmigung durch OSL.

Alle verwendete Hard- und Softwarenamen sind Handelsnamen und/oder Warenzeichen oder eingetragene Warenzeichen der jeweiligen Hersteller oder Inhaber.

Beschränkungen

OSL stellt diese Dokumentation für die vorstehend beschriebene interne oder private Nutzung unentgeltlich und »wie sie ist« («as is») bereit. Eine Garantie auf diese Dokumentation bzw. auf die durch sie beschriebene Software, auf Code-Beispiele und beschriebene Verfahren, auf eine handelsübliche Qualität oder die Eignung für einen bestimmten Zweck ist ausgeschlossen. OSL übernimmt insbesondere keine Haftung für enthaltene Fehler, unmittelbare oder mittelbare Schäden oder Schadenersatz für Aufwendungen, die durch Auslieferung, Bereitstellung, Benutzung oder Nichtbenutzung dieses Dokumentes entstehen.

Der Erhalt dieses Dokumentes begründet keine weiteren Rechte. Alle in diesem Material enthaltenen Informationen stehen unter dem Vorbehalt einer Änderung ohne vorherige Ankündigung. Weder die beschriebene Software noch die vorliegende Dokumentation stellen Programmierschnittstellen (APIs) oder Teile davon dar.

Diese Dokumentation selbst, die darin beschriebene Software und referenzierte Dokumentationen sind intellektuelles Eigentum der jeweiligen Hersteller oder Inhaber der betreffenden Rechte, das u. a. durch das Urheber-, Handels-, und Markenrecht geschützt ist. Die Benutzung, Installation, Kopie, Weitergabe oder Veräußerung solcher Software und Dokumentationen unterliegt den jeweiligen Lizenzbestimmungen.

In dieser Dokumentation enthaltene Informationen zu Produkten und Dienstleistungen Dritter sind entsprechenden Dokumentationen oder sonstigen Publikationen der jeweiligen Hersteller, sekundären oder sonstigen öffentlich zugänglichen Quellen entnommen. OSL hat diese Produkte und Dienstleistungen, Ihre Leistungsparameter und Interoperabilität - auch in Bezug auf OSL Produkte - nicht getestet und schließt folgerichtig jede Garantie oder Haftung hinsichtlich der Produkte, Dienstleistungen und Informationen Dritter aus.

Die in dieser Dokumentation enthaltenen Beispiele werden je nach Softwareständen, Hardware und sonstiger Umgebung von Ihrem System abweichen. Für die Bewertung der Korrektheit der vorliegenden Informationen, für die Auswahl und die Beurteilung der Eignung beschriebener Verfahren sowie dargestellter Hard- und Softwarekonfigurationen für einen bestimmten Zweck, für deren Anwendung oder Nichtanwendung sowie die Tauglichkeit etwaig ausgewählter Kombinationen von Hard- und Softwarekomponenten im Gesamtsystem ist allein der Anwender verantwortlich. Dies gilt auch für eine nachfolgende Installation und Konfiguration von Software, für die Nachnutzung der beschriebenen Verfahren sowie für die im Rahmen der Nutzung angestrebten Ergebnisse.

Versionen dieses Dokumentes

Version	Datum	Autor	Email	Inhalt / Änderungen
1.0	1.09.2015	MB	support@osl.eu	Erste Fassung

Inhaltsverzeichnis

1. Einleitung.....	1
2. Das OSL UVE - Funktionsweise und Komponenten.....	1
2.1. Die Unified Virtualisation Clients (UVC).....	2
2.2. Das Unified Virtualisation Network (UVN).....	2
2.3. Die Unified Virtualisation Server (UVS).....	2
2.4. Die Komponenten im Zusammenspiel.....	2
3. SPARC-Server in x86-VM-Infrastrukturen.....	4
4. Zusammenfassung.....	5



SPARC-Server und hyperkonvergente x86-VM-Infrastrukturen

1. Einleitung

Webzentrierte Applikationen und Cloud-Services setzen Maßstäbe, denen sich das eigenbetriebene Rechenzentrum trotz typischerweise klassischer Applikationsstacks heute stellen muß. Allein mit softwaredefinierten Infrastrukturen ist die gebotene Schnelligkeit, Flexibilität und Effizienz erreichbar. OSL bietet mit seinem Unified Virtualisation Environment eine hyperkonvergente VM-Infrastruktur als Client-Server-Lösung, die vorhandene Komponenten vom Enterprise Storage bis hin zu Solaris-Servern integrieren und so auch mit modernen SPARC-Systemen eine höchst nutzbringende Symbiose eingehen kann.

Während klassische VM-Umgebungen vorrangig die Themen Servervirtualisierung und Konsolidierung adressieren, sollten hyperkonvergente Infrastrukturen *alle* relevanten Virtualisierungsebenen für das softwaredefinierte Rechenzentrum abdecken können. OSL beschreibt dies für sein Unified Virtualisation Environment mit der Formel $V3 = \text{Virtual} \times (\text{Server} + \text{Storage} + \text{Network})$. Virtualisierung und Hardwareabstraktion gehen dabei Hand in Hand. Weil sämtliche Beschreibungen für das softwaredefinierte Rechenzentrum abstrakt und außerhalb der Hardwarekomponenten erstellt und beschrieben werden, entfallen komplexe Konfigurationsvorgänge beispielsweise an Switches, Speichersystemen oder Servern. Hardwarekomponenten werden so auf einfachste Weise austauschbar und der Aufbau neuer Systemlandschaften per Software ist – leicht nachvollziehbar – deutlich einfacher und schneller zu bewerkstelligen, als wenn physische Komponenten konfiguriert und verkabelt werden.

Neben dem Gedanken „Infrastruktur aus Software“ steckt noch eine wichtige Neuerung im OSL Unified Virtualisation Environment: Das Converged Network. Sowohl das klassische LAN als auch der Block-I/O (sonst im RZ zumeist FC) nutzen die gleichen Wege, was nicht die Infrastruktur radikal vereinfacht (und kostengünstiger macht). Auch die Administration wird radikal anders verstanden: Bestimmte Aufgabenfelder wie die SAN-Administration, LUN-Masking u. ä. entfallen beim Aufbau der VM-Infrastrukturen komplett.

2. Das OSL UVE - Funktionsweise und Komponenten

Das OSL Unified Virtualisation Environment (UVE) stellt sich schematisch wie folgt dar:

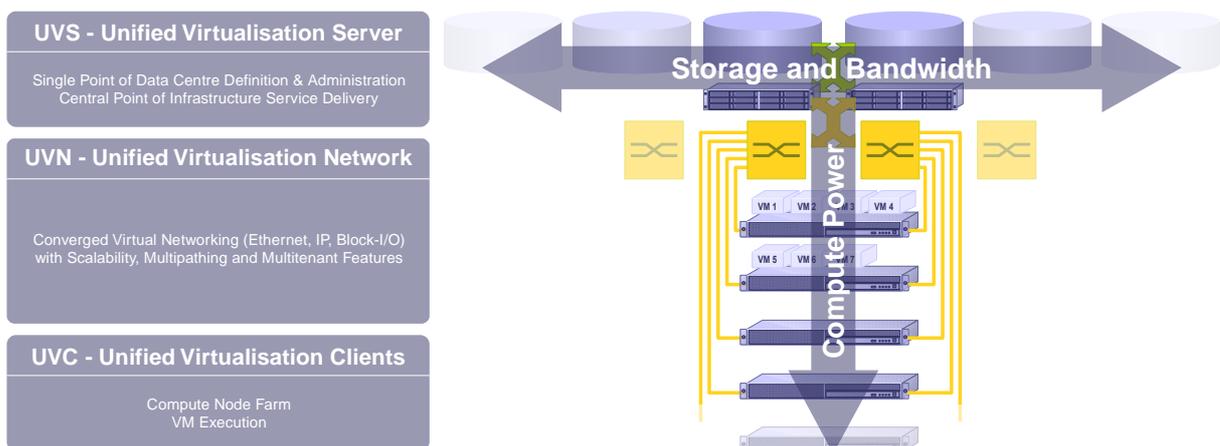


Abbildung 1: Die Funktionsblöcke im OSL Unified Virtualisation Environment (links) und die Abbildung auf bekannte physische RZ-Komponenten.

Gut zu erkennen ist, dass sich die netzseitige Anbindung der Compute Nodes deutlich vereinfacht. FibreChannel entfällt und Mehrfachanschlüsse sind in der Regel nur aus Verfügbarkeitsgründen vorhanden, die Fähigkeiten der Software ersparen dem Administrator allerdings in jedem Fall die Konfiguration von Multipfadkonstrukten sowohl für Block-I/O als auch für Ethernet/IP.

2.1. Die Unified Virtualisation Clients (UVC)

Die UVC stellen die Rechenleistung zur Verfügung und sind der Ausführungsort für die virtuellen Maschinen. OSL unterstützt mit einer einheitlichen Handhabung derzeit für x86-Plattformen KVM, Xen und VirtualBox, eine Integration von Solaris-Technologien wie (Kernel) Zones und LDoms steht für das Schwesterprodukt OSL Storage Cluster bereits zur Verfügung und ist für das OSL Unified Virtualisation Environment ebenfalls auf der Roadmap. Die kleinstmögliche Konfiguration beginnt mit einem UVC, weitere Virtualisation Clients können bei Bedarf on-the-fly ergänzt werden. Für die redundante Anbindung über das Converged „Unified Virtualisation Network“ werden typischerweise zwei, ggf. auch vier 1GE-Ports genutzt, womit sich mehr als 200 bzw. 400 MBytes/s an I/O-Bandbreite je UVC kostengünstig und einfach darstellen lassen. Zwei preiswerte 24-Port-Switches können so 5 bzw. 10 UVC bedienen. Natürlich sind auch größere Switches ggf. mit 10GE-Uplink zu den UVS oder eine durchgängige Verwendung von 10GE (ggf. auch 40GE / Infiniband) möglich. Softwareseitig kommt auf den linuxbasierten UVC mit dem gewünschten Hypervisor das OSL UVC-Paket zum Einsatz, das die komplette Steuerung des Hypervisors unter Einbeziehung des Virtual Networking und des virtualisierten Block-I/O übernimmt.

2.2. Das Unified Virtualisation Network (UVN)

Über das UVN werden sowohl LAN- als auch Block-I/O abgewickelt. Anders als bei (lossless) FCoE können hier konventionelle Ethernet- Komponenten verwendet werden. In vielen Fällen ist, wie oben dargestellt, 1GE in redundanter Auslegung vollkommen ausreichend, zudem besteht über die Zahl der Controller an den UVC die Möglichkeit einer Skalierung. Für gehobene Ansprüche wird typischerweise auf 10GE zurückgegriffen.

Ohne eine Rekonfiguration der Switches werden die für das virtuelle Rechenzentrum notwendigen Netzwerkstrukturen per Software erzeugt. Dabei kann der Zugriff auf Block-I/O und Netzwerkressourcen granular gesteuert werden, so dass verschiedene virtuelle RZ-Infrastrukturen zuverlässig voneinander getrennt auf der gleichen Physik laufen können. Wesentliche Technologien hinter den Unified Virtualisation Network sind VLANs, die UVE-interne Network Monitoring and Configuration Engine sowie das RSIO-Protokoll, das bei gerinstem CPU-Bedarf skalierbaren I/O für virtualisierte Blockdevice-Ressourcen über alle zur Verfügung stehenden Interfaces anbietet.

2.3. Die Unified Virtualisation Server (UVS)

Diese Maschinen sind Server für Virtual Storage, Virtual Networking und VM Control in einem und bieten damit die wesentlichen Infrastrukturservices aus nur einer Appliance heraus. Möglich sind derzeit Single- und Dual-Server-Konfigurationen, der Wechsel zwischen beiden Varianten kann ohne Betriebsunterbrechung erfolgen.

Storage-Ressourcen können beispielsweise interne (SSD-) Disks aber auch externe Plattensysteme sein. Die unter Solaris (aktuell 11.2) laufende UVS-Software übernimmt dabei alle Funktionen einer Speichervirtualisierung und kann Daten sowohl selbständig spiegeln als auch autonome Spiegel externer RAID-Systeme integrieren und so auch höhere Anforderungen bis hin zum Disaster Recovery bedienen. Die Speichervirtualisierung arbeitet dabei eigenständig und unabhängig von Solaris-Technologien wie ZFS oder dem Solaris Volume Manager. Sie verfügt über eine für die Interoperabilität mit anderen Plattformen und netzwerk- und clusterzentrierten I/O (RSIO) optimierte Block-I/O-Schnittstelle.

Zugleich sind die UVS der zentrale Administrationspunkt, wobei das Management des virtualisierten Rechenzentrums sowohl über ein leistungs- und scriptfähiges CLI als auch über eine Web-GUI oder alternativ ein curses-basiertes Menüsystem erfolgen kann.

Im Backend werden Massenspeicherressourcen typischerweise über FC eingebunden, in Richtung Unified Virtualisation Network bietet der UVS seine Dienste typischerweise über 1GE oder 10GE an.

2.4. Die Komponenten im Zusammenspiel

Im Zusammenspiel der vorstehend beschriebenen Funktionsblöcke des OSL Unified Virtualisation Environment ist es möglich, komplette – auch sehr unterschiedliche RZ-Infrastrukturen per Software darzustellen. Hosts und Netzwerkkomponenten realisieren sich in entsprechenden Softwareobjekten, auf der Block-I/O-Seite entfällt durch das Zusammenspiel von Speichervirtualisierung und RSIO der klassische Administrationsstack von LUNs über das SAN bis zum Blockdevice der nutzenden OS-

Instanz. Entscheidend ist, dass neue Aufgaben und Architekturen per Software und ohne Änderungen am physischen Setup umgesetzt werden können.

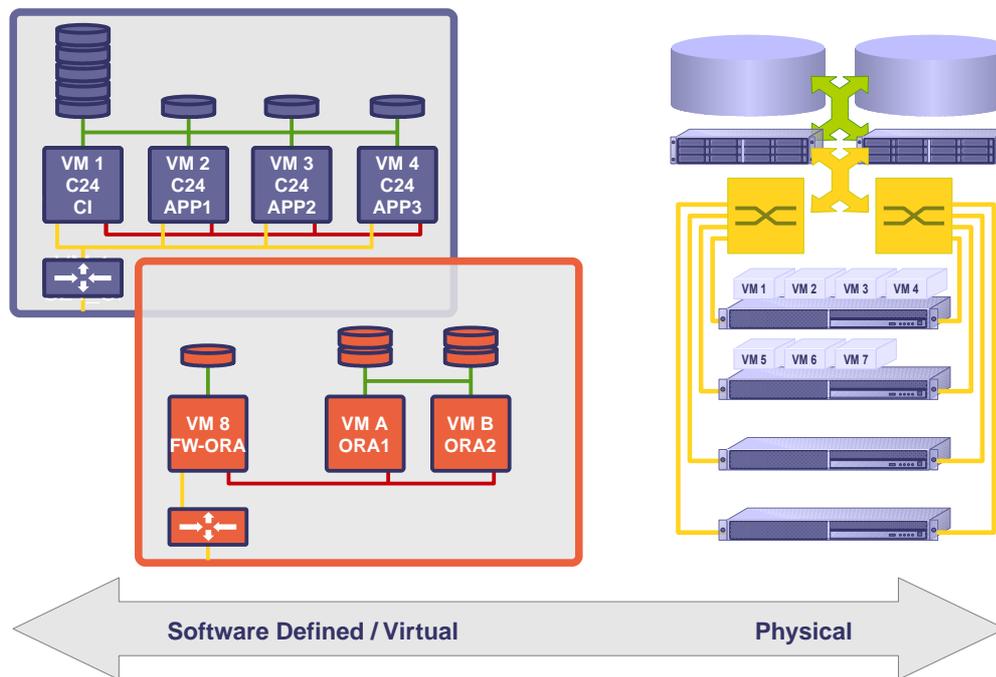


Abbildung 2: Neue Infrastrukturen werden im Virtualisierungslayer per Software auf einer immer gleichen physischen Hardwarekonfiguration erzeugt

Neben der Schnelligkeit spielt auch die Automatisierungsfähigkeit im RZ-Betrieb eine große Rolle. Hier kann das UVE durch die Zentralisierung aller Infrastrukturdienste in Verbindung mit einem leistungsstarken Command Line Interface und besten Script-Fähigkeiten punkten. Natürlich ist dies auch für automatisierte Überwachungen von Vorteil. In der interaktiven Bedienung bieten ein Curses-basiertes System und ein modernes, im Browser lauffähiges Web-Interface zusätzlichen Komfort.

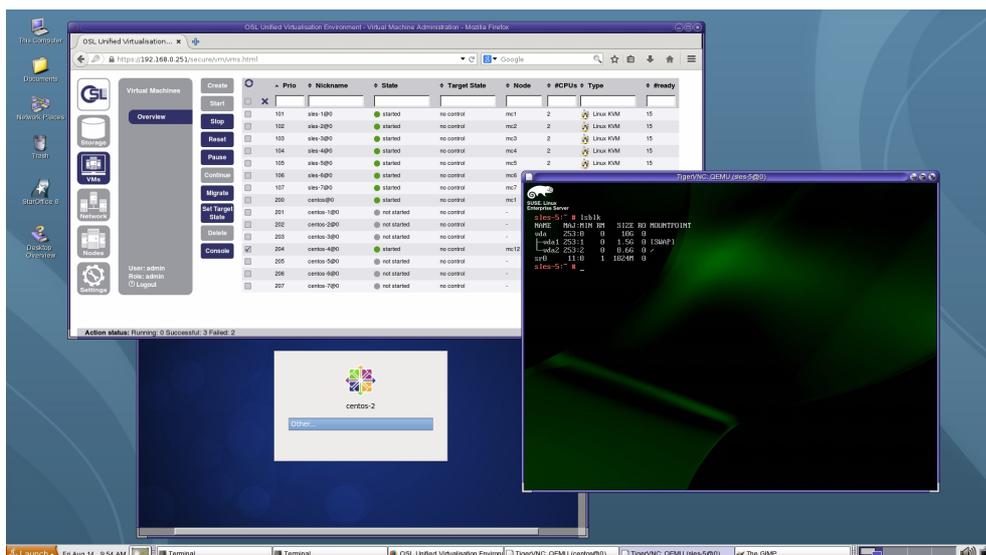


Abbildung 3: Desktop mit der UVE-WebGUI nebst Konsolen virtueller Maschinen.

3. SPARC-Server in x86-VM-Infrastrukturen

Die enorme Leistungsfähigkeit aktueller SPARC-Systeme (z. B. T5 und M10) ermöglicht heute eine Konsolidierung vorhandener Landschaften, wie es noch vor 5 Jahren unvorstellbar erschien. Dabei gibt es auch extreme Konstellationen, in denen ein halbes Server-Rack auf nurmehr eine Höheneinheit einer M10-1 schrumpft. Herausragende Faktoren, die diese Entwicklung ermöglichen sind:

- die große Zahl von Cores und Threads in einem System
- die gegenüber früheren Systemen deutlich erhöhte Taktrate
- die Reife von LDom als überlegener, feingranularer Virtualisierungstechnologie
- überzeugende Möglichkeiten zum I/O-Sharing (Direct I/O und SR-IOV)

Dass sich so viele physische Server in einem einzigen zusammenfassen lassen, ist leicht zu erkennen. Auf den zweiten Blick eröffnen sich aber weitere Möglichkeiten, die nicht weniger begeistern und noch weit größere Effekte für den RZ-Betrieb zeitigen können.

Beginnen wir im Storage-Backend: Für kleinere und mittelgroße Rechenzentren ist der Verzicht auf ein SAN heute realistisch. Moderne Enterprise-RAID-Systeme verfügen mindestens über 4 FC-Ports (oft deutlich mehr), an die sich zwei oder mehr konsolidierte SPARC-Server problemlos direkt und redundant anschließen lassen. So ist auf einfachste Weise ein beeindruckendes Spektrum an Performance und Verfügbarkeit abzudecken. Zudem verfügen die SPARC-Server über ausreichend Performance, um in einer LDom einen Unified Virtualisation Server laufen zu lassen, der seinerseits über das Unified Virtualisation Network eine komplette x86-VM-Infrastruktur bedienen kann. Somit lassen sich, wie Abbildung 4 verdeutlicht, nicht nur eine große Zahl von virtualisierten Servern, sondern auch die dafür erforderliche physische Infrastruktur auf einfachste Weise und dennoch skalierbar aufbauen.

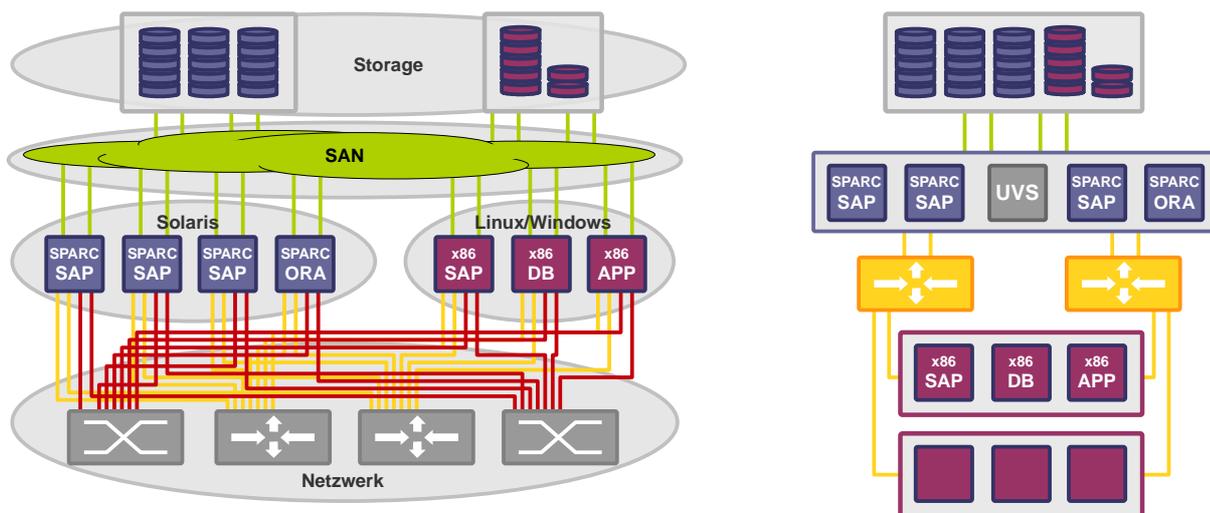


Abbildung 4: Hyperkonvergente Systeme (rechts ein konsolidiertes SPARC-System und zwei x86-basierte UVC) bedeuten im Vergleich zu einer klassischen Landschaft (links) nicht nur eine schlanke und aufgeräumte Infrastruktur sondern integrieren vielfältige administrative Aufgabenfelder in einem vereinfachten und einheitlichen Prozess.

Der UVS mit seiner Fähigkeit, komplette Infrastrukturen von nur einem Punkt aus zu managen, bringt dabei auf natürliche Weise zusammen, was sonst typischerweise in dedizierten Projekten mit dedizierter Hardware und dedizierten Teams dargestellt wird. Dieser organisatorische Vorteil ist der entscheidende Punkt bei der überlegenen Agilität der vorgestellten hyperkonvergenten Infrastruktur. Dichter an der Hardware lassen sich folgende Unterschiede erkennen:

- Eliminierung der FibreChannel-Switches.
- Erhebliche Reduzierung der Zahl von FC-Ports, was nicht nur geringere Kosten in der Anschaffung, sondern auch in der täglichen Administration bedeutet.
- Vereinfachte Netzwerk-Infrastruktur mit deutlich weniger Ports.

Die vorgenannten Punkte führen zu einer weiteren Vereinfachung im Storage-Backend. Durch die im UVS integrierte Speichervirtualisierung entfällt die sonst für die VM-Infrastruktur notwendige Storage-

Administration auf dem RAID-System im Kontext von VM-Provisionierungen. Vielmehr werden Kapazitäten nur noch anonym bereitgestellt – die Verwaltung erfolgt dann durch den UVS im natürlichen Kontext der VM-Administration. Dieser organisatorische Fortschritt erlaubt es, die sonst oft exklusiv für die SPARC-Highend-Umgebung bereitgestellten Diskressourcen gefahrlos auch für die VM-Infrastruktur zu nutzen. Auf der Hardwareseite nutzt der UVS die gleichen Controller, wie die Datenbank-Domains – eine sinnvolle Synergie.

4. Zusammenfassung

SPARC-Systeme mit ihrer Leistungsfähigkeit und ihren herausragenden Verfügbarkeitseigenschaften bilden zusammen mit einer abgestimmten Storage-Infrastruktur weiterhin das Rückgrat vieler Datenbank- und SAP-orientierter Rechenzentren. Sie eignen sich auch als zentrale Server für hyperkonvergente x86-basierte VM-Infrastrukturen und erlauben so die Transition von vielgliedrigen Organisationsformen (Solaris, Linux, Windows, Datenbank, SAP, Backup, Netzwerk, Storage, SAN) zu einem schlanken, integrierten RZ-Prozess.