

Schwerpunktthema

Virtualisierung: Virtualisierungs- bewusste RZ-Netze

Fortsetzung von Seite 1



Dipl.-Inform. Matthias Egerland ist seit 2005 Mitarbeiter der ComConsult Beratung und Planung GmbH. Er ist Leiter des Competence Centers Virtuelle IT und arbeitet als Berater in den Competence Centern IT-Sicherheit und Netze. Neben den Schwerpunkten Desktop-, Server- und Infrastruktur-Virtualisierung beschäftigt sich Herr Egerland insbesondere mit der Sicherheit in virtualisierten Umgebungen.



Dipl.-Inform. Daniel Meinhold ist Consultant bei der ComConsult Beratung und Planung GmbH. Dort ist er in den Bereichen Telekommunikationssysteme, Virtualisierung und IT-Sicherheit tätig. Neben diesbezüglichen Praxiserfahrungen in zahlreichen Projekten ist er für die Planung und Durchführung entsprechender Testsznarien im ComConsult-eigenen Labor zuständig.

1. Hintergrund: Dynamik virtueller Maschinen

Das Grundprinzip der Virtualisierung beruht auf der Idee, Komponenten der Informationstechnologie anstelle in Hardware in Software zu realisieren. Dieser Abstraktionsprozess von dieser physischen Basis in Richtung logischer Teilsegmente bedeutet einen hohen Grad an Flexibilität und eine enorme Beschleunigung von Prozessen. Dauert es in klassischen, d.h. nicht-virtualisierten Umgebungen Tage bis Wochen, bis ein neuer Server vom Bestellprozess über den Lieferungsvorgang, seinen Einbau sowie die Installation und Konfiguration beschafft wurde, kann dies in virtualisierten Umgebungen mittels weniger Mausklicks binnen Minuten erledigt werden. Auch der Umzug eines virtualisierten Servers - im Folgenden: einer virtuellen Maschine (VM) - geht deutlich schneller von statten, da keine physische Hardware mehr ihren geografischen Aufstellort verlassen muss, um an ihrem neuen Ziel neu verkabelt in Betrieb genommen zu werden.

Eine virtuelle Maschine besteht im laufenden Betrieb aus einem Festplatten-Image (Abbildung 1, Punkt 1), einer Konfigurationsdatei mit Informationen zu ihrer „Hardware“-Ausstattung sowie dem von ihr tatsächlich adressierten Hauptspeicherbereich (Abbildung 1, Punkt 2). Die Migration einer virtuellen Maschine bedeutet daher im Wesentlichen das Kopieren

ihres Hauptspeichers von einem physischen Host auf einen anderen physischen Host (Abbildung 1, Punkt 3). Da die VM währenddessen weiterläuft, wird auf den Hauptspeicher nicht nur lesend sondern

auch schreibend zugegriffen. Insofern ist es erforderlich, die Speicheradressen dieser Schreibzugriffe während des Kopiervorgangs in einer Matrix festzuhalten (Abbildung 1, Punkt 4), damit diese Ände-

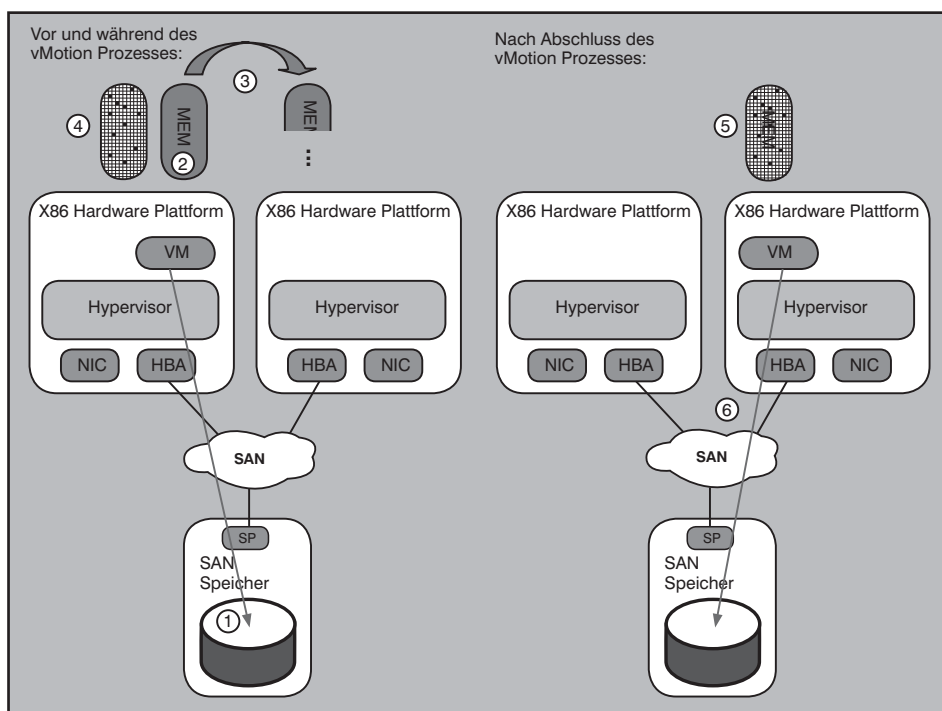


Abbildung 1: Verschieben einer virtuellen Maschine (VM) im laufenden Betrieb (vMotion)
(1): Festplatten-Image der VM; (2): Hauptspeicher der VM; (3): der Hauptspeicherinhalt wird auf das Ziel-Host-System kopiert; (4): in einer Matrix werden währenddessen die Änderungen im Hauptspeicher festgehalten; (5): nach Abschluss des Kopiervorgangs werden diese Änderungen nachgezogen; (6): der Verweis auf das Festplatten-Image wird aktualisiert

Virtualisierung: Virtualisierungsbewusste RZ-Netze

rungen abschließend auf das Zielsystem übertragen werden können (Abbildung 1, Punkt 5). Nur während dieses Transfers der Speicheränderungsinformationen muss die VM eingefroren werden, um weitere Speicherzugriffe und damit verbundene Änderungen zu verhindern. Dies führt zu einem Aussetzer von ca. 1-2 Sekunden, bis die VM wieder erreichbar ist und im zuletzt aktuellen Zustand weiterarbeitet. Sofern auf Applikationsebene diese Unterbrechung nicht so lang war, dass sie zu einem Timeout geführt hat, werden alle laufenden Verbindungen für den Anwender transparent fortgesetzt.

Auf der Seite des Festplattenspeichers besteht die Migration einer VM lediglich aus der Aktualisierung des Verweises auf das Festplatten-Image (Abbildung 1, Punkt 6). Wird vor der Migration vom ersten Host-System aus auf das Image verwiesen, muss nach der Migration vom Ziel-Host-System aus auf das Image verwiesen werden. Der Speicherort selbst bleibt hingegen konstant. Aus diesem Grund ist eine der Grundvoraussetzungen für das Verschieben einer VM im laufenden Betrieb, dass ein gemeinsam genutzter Plattenspeicher („Shared Storage“) das Festplatten-Image trägt und dieses nicht auf einem der Host-Systeme lokal abgelegt ist (Direct Attached Storage, DAS).

Neben dem manuellen Auslösen einer VM-Migration sind unterschiedliche Mechanismen zur Automatisierung dieses Prozesses zur Marktreife gelangt, um die Effizienz des Rechenzentrumsbetriebs weiter zu steigern. Der Distributed Resource Scheduler innerhalb der VMware-Umgebung verfolgt beispielsweise das Ziel, die Host-Systeme möglichst gleichmäßig auszulasten (siehe Abbildung 2). „Gleichmäßig“ kann in diesem Zusammenhang möglichst einheitliche CPU-Auslastung, Hauptspeicherbelegung und seit der aktuellen Version vSphere 4 auch Netzwerkauslastung bedeuten. Ein vergleichbares Produkt bietet Citrix mit der Funktion „Dynamic Workload Balancing“ im Rahmen der „Essentials for XenServer“. Durch diese Verteilung sollen wachsende Leistungsanforderungen und plötzliche Lastspitzen bedient werden, denn bei einem konstanten Ressourcenbedarf der VMs wäre es prinzipiell egal, auf welchem Host-System dieser Bedarf abgedeckt wird. (Ein zu 80% ausgelasteter Host und ein zu 40% ausgelasteter Host weisen in der Gesamtbetrachtung die gleiche Effizienz auf wie zwei zu jeweils 60% ausgelastete Hosts.)

VMware Distributed Power Management (DPM) ist gewissermaßen ein Spezialfall von DRS mit der Zielsetzung einer erhöh-

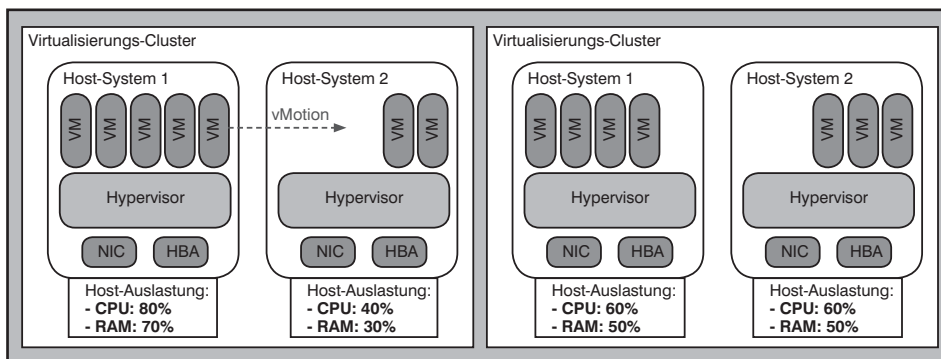


Abbildung 2: Der VMware Distributed Resource Scheduler (DRS) löst auf einem ungleichmäßig ausgelasteten Virtualisierungs-Cluster (oben links) einen vMotion-Vorgang aus. Ziel dieser Umverteilung virtueller Maschinen (VM) ist eine möglichst gleichmäßige Auslastung aller beteiligten Host-Systeme hinsichtlich ihrer CPU- und RAM-Ressourcen (oben rechts).

ten Energieeffizienz. Die Idee von DPM ist, in betriebsarmen Zeiten - z.B. nachts oder am Wochenende - die gegenüber dem Alltagsbetrieb verringerte Anzahl erforderlicher VMs und deren verringerte Leistungsanforderungen auf einem Minimum von physischen Hosts zu konzentrieren. Gelingt es mit diesem Mechanismus, eines oder mehrere Host-Systeme vollständig von VMs zu befreien, können diese strom- und klimatisierungssparend heruntergefahren werden (siehe Abbildung 3).

War DPM in der VMware Virtual Infrastructure 3 noch als „experimentell“ eingestuft,

wird es in vSphere 4 auch in produktiven Umgebungen unterstützt. Allerdings erfordert dieses Leistungsmerkmal die „Enterprise“ Lizenzierungsstufe und schlägt somit mit entsprechenden Zusatzkosten zu Buche.

1.1 Voraussetzung für wandernde VMs: einheitliche Netzkonnektivität

Virtuelle Maschinen kommunizieren über ihre virtuelle Netzwerkschnittstelle mit dem virtuellen Switch der Virtualisierungslösung. Dieser virtuelle Switch stellt die Brücke zur physischen Netzwerkschnittstelle her, die die Kommunikation mit der

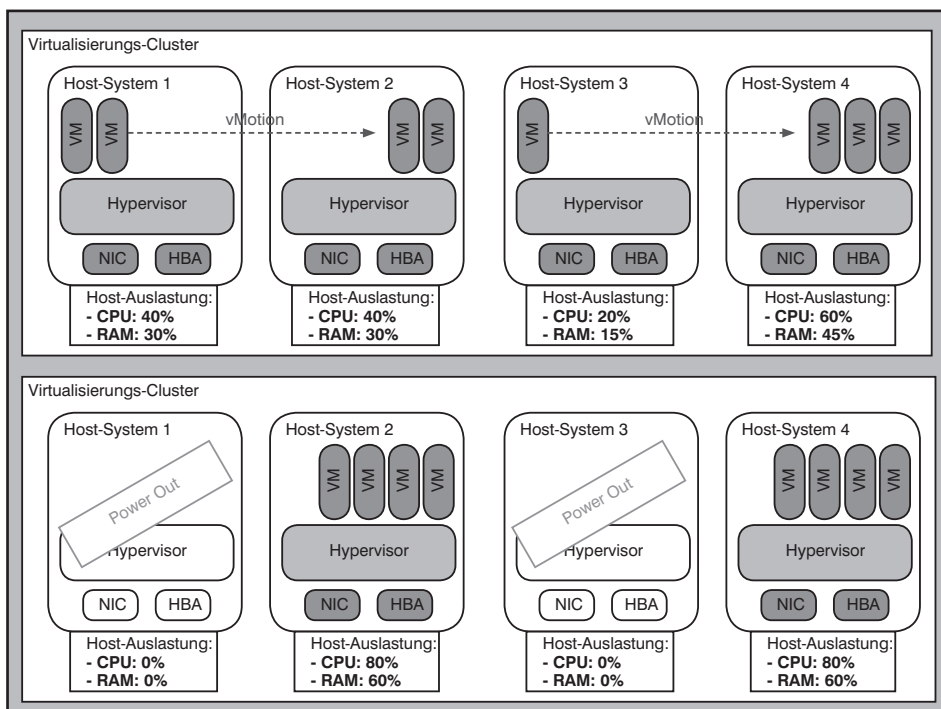


Abbildung 3: Der VMware Distributed Power Manager (DPM) verschiebt in betriebsarmen Zeiten mit Hilfe von vMotion die virtuellen Maschinen (VM) auf möglichst wenige Host-Systeme (obere Hälfte), um komplett freigewordene Host-Systeme herunterfahren zu können und auf diese Weise den Energieverbrauch des Virtualisierungs-Clusters zu reduzieren (untere Hälfte).

Virtualisierung: Virtualisierungsbewusste RZ-Netze

physischen Netzwerke herstellt. In der Virtualisierungslösung können bei allen Herstellern mehrere virtuelle Switches angelegt und mit jeweils unterschiedlichen physischen Netzwerkschnittstellen assoziiert werden. Alle marktgängigen Virtualisierungslösungen ermöglichen außerdem die Nutzung von VLANs mittels virtueller Switches. Einzig die Terminologie und die Umsetzung unterscheiden sich bei den einzelnen Anbietern.

Während die VLAN-Zuordnung bei Citrix über die virtuellen Switches („Netzwerke“) erfolgt, sind es bei Microsofts Hyper-V beispielsweise die virtuellen Schnittstellen der VMs oder alternativ die virtuellen Switches, die eine VLAN-ID zugewiesen bekommen. Bei VMware müssen zunächst einzelne Portgruppen auf dem virtuellen Switch angelegt werden, die ihre jeweiligen VLAN-IDs erhalten. Die Netzwerkschnittstellen der VMs werden einfach mit diesen Portgruppen verbunden, um einem VLAN zugeordnet zu werden. Auf Seiten der physischen Schnittstelle des Host-Systems müssen in diesem Fall keine weiteren Einstellungen vorgenommen werden, um diese VLANs zu transportieren, da dies implizit konfiguriert ist (siehe Abbildung 4).

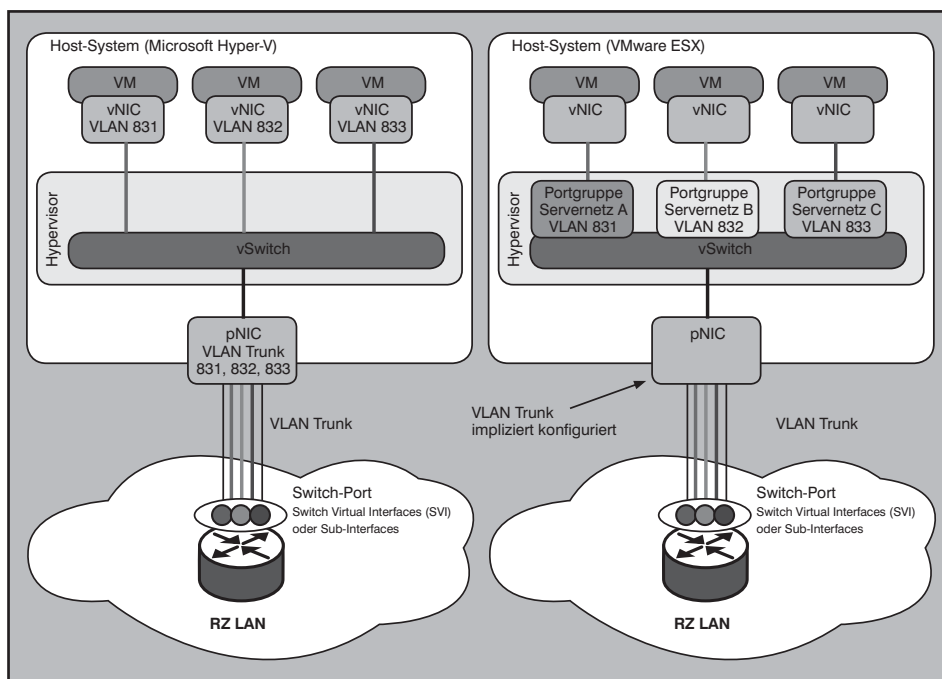


Abbildung 4: Unterschiedliche Servernetze auf einer physischen Netzwerkschnittstelle (pNIC) des Host-Systems. Bei Microsoft Hyper-V müssen die virtuellen Netzwerkschnittstellen (vNIC) der virtuellen Maschinen (VM) mit VLAN IDs konfiguriert werden (links). Bei VMware sind auf den virtuellen Switches (vSwitch) unterschiedliche Portgruppen inkl. VLAN ID anzulegen.

Voraussetzung dafür, dass die Virtualisierungslösung das Verschieben einer virtuellen Maschine zulässt, ist eine einheitliche Konfiguration der virtuellen Switches sowohl auf dem Quell- als auch auf dem Ziel-Host-System. D.h. die Virtualisierungslösung selber achtet darauf, dass eine VM nach der Migration nicht „in der Luft hängt“ oder mit einer anderen Port-Gruppe verbunden wird.

Worauf die Virtualisierungslösung nicht achtet, ist, ob die Sicherheitsrichtlinien der jeweiligen Portgruppen auf dem Quell- und dem Ziel-Host-System identisch konfiguriert sind. So kann es durchaus vorkommen, dass eine VM nach ihrer Migration innerhalb der virtualisierten Umgebung anderen Sicherheitsrichtlinien und Quality-of-Service-Parametern (QoS) unterliegt als zuvor, auch wenn sie netzwerkseitig ihren Kontext nicht verlassen hat. Die Gefahr besteht also weniger darin, dass eine VM durch Migrationen ihre Sicherheitszone verlässt, sondern dass die Sicherheitszonen unterschiedliche Parameter auf den einzelnen Host-Systemen aufweisen. Die Frage stellt sich also, wie eine einheitliche Konfiguration der Zonen-Merkmale gewährleistet werden kann und wer hierfür verantwortlich ist.

Ein erster Schritt in Richtung einer Host-System-übergreifend einheitlichen Konfiguration ist das Konzept der Distributed

Virtual Switches (DVS), die seit VMware vSphere 4 verfügbar sind und auch von Citrix XenServer bereits angekündigt wurden. Hier ist es so, dass sich ein virtueller Switch über mehrere Host-Systeme erstreckt und insofern die Konfiguration von Portgruppen inkl. aller Parameter nur einmalig erfolgt und von der Virtualisierungslösung automatisch auf allen beteiligten Host-Systemen einheitlich verfügbar gemacht wird.

Allerdings sind die Leistungsmerkmale des von Hause aus mitgelieferten DVS von VMware in puncto Sicherheit und QoS auf wenige Aspekte beschränkt. Auch obliegt die Konfiguration dieses Switches demjenigen, der die Hoheit über das Werkzeug zur Administration der Virtualisierungslösung besitzt - in diesem Fall den vCenter Server. Die Konfiguration dieser virtuellen Switches unterscheidet sich grundsätzlich von derjenigen einer traditionellen Netzwerkkomponente. Auch hat die Netzwerkeite keinerlei Möglichkeit die Korrektheit dieser Konfiguration und insbesondere ihrer Sicherheitsrichtlinien zu überprüfen, ohne auf die Management-Umgebung der Virtualisierungslösung zurückzugreifen.

Hier sind also andere Mechanismen gefordert, wenn man die Administration der Netzanbindung virtueller Maschinen wieder in die Hand der Netzwerker legen

möchte, wie es in der physischen Welt bewährte Praxis ist. Die folgenden Abschnitte sollen zwei unterschiedliche Ansätze hierfür näher beleuchten.

2. Enterasys Network Access Control

Enterasys bietet im Bereich virtueller RZ-Netze ein Produkt aus dem Bereich Netzzugangskontrolle (Network Access Control, NAC) an: Enterasys NAC. Dabei werden die Elemente Erkennung, Authentisierung, Prüfung und Autorisierung von Systemen einschließlich Überwachung und Remediation z.B. per Captive Portal (Abbildung 5) abgedeckt.

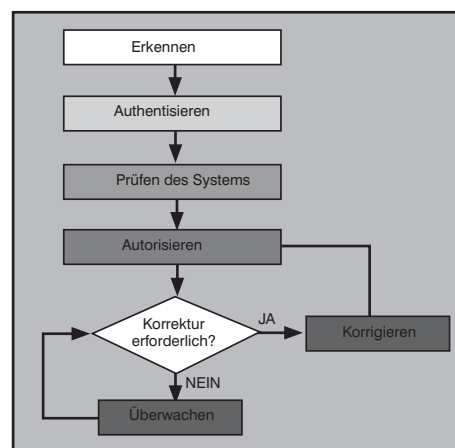


Abbildung 5: Elemente einer Netzzugangskontrolle

Virtualisierung: Virtualisierungsbewusste RZ-Netze

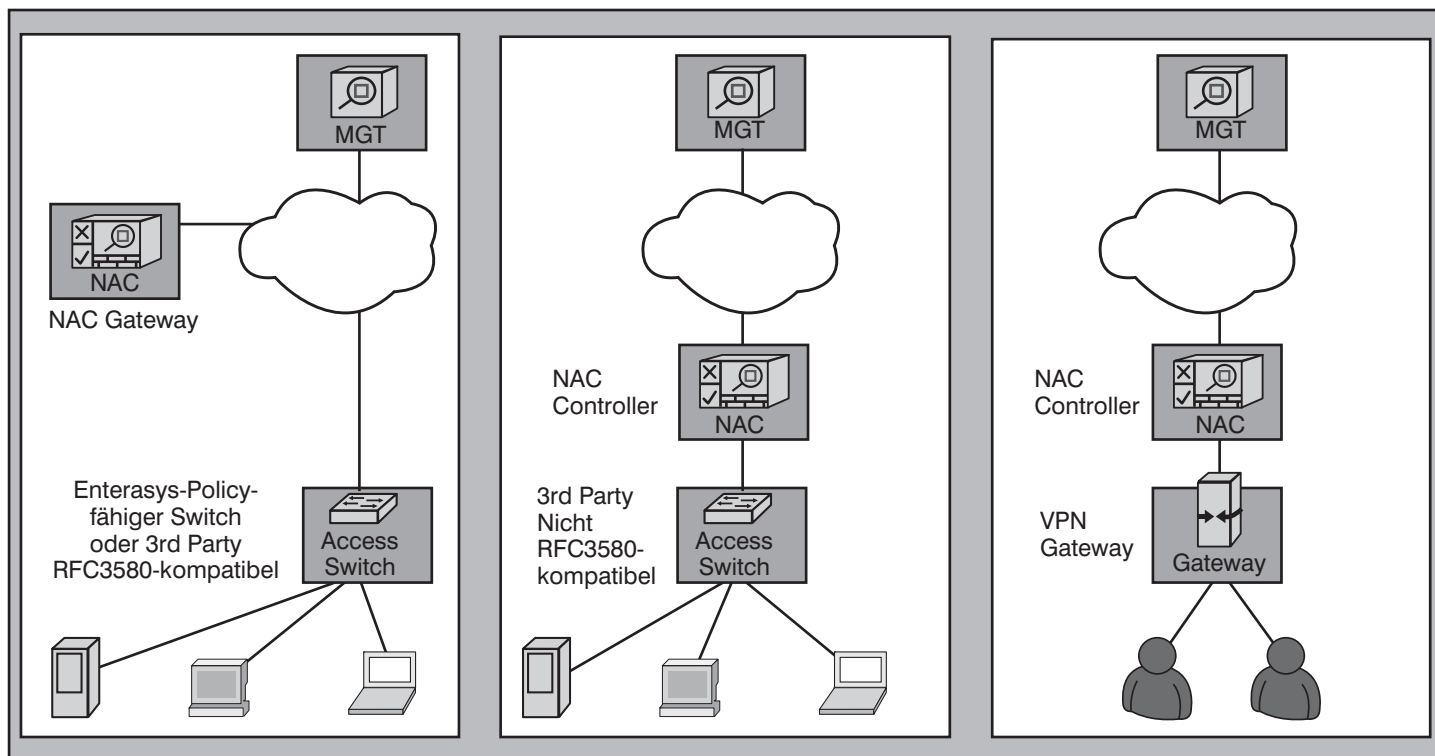


Abbildung 6: Auswahl möglicher Szenarien in einer Enterasys-NAC-Umgebung

Die Systeme beziehen sich in der Regel auf Endgeräte wie PCs oder Notebooks. Anders verhält es sich bei der Servervirtualisierung: Hier sind es die VMs, deren Zugriff auf die Infrastruktur kontrolliert und in Einklang mit den Sicherheitsrichtlinien gebracht werden soll. Insbesondere folgende Punkte werden hierbei seitens Enterasys hervorgehoben:

- Erkennung von verschiedenen VMs an einem physischen Switch-Port
- Separation verschiedener VMs anhand von Regelwerken (z.B. ACLs / VLANs)
- Unterschiedliche Dienstgütern (z.B. Bandbreite, QoS) für jede VM
- Bereitstellung dieser Regelwerke und Dienstgütern unabhängig vom physischen Aufenthaltsort der VM (z.B. aufgrund einer Migration im laufenden Betrieb)
- Einsatz unabhängig von der Virtualisierungsplattform
- VLAN-Trunking zwischen Host und Switch nicht erforderlich

Nachfolgend werden Architektur und Komponenten der Lösung näher beschrieben sowie Hinweise und Einschränkungen zur Integration in eine virtuelle Serverlandschaft gegeben. Untersucht wurde dies im Rahmen eines Testaufbaus im ComConsult-eigenen Virtualisierungs-Labor auf Basis von Citrix XenServer 5.5.

2.1 Grundarchitektur

Die Grundarchitektur der Enterasys-NAC-Lösung setzt sich aus den folgenden Elementen zusammen:

- Access Layer (drahtlos oder kabelbasiert)
- NAC Appliance (In-band oder Out-of-band)
- Management-Software zur Konfiguration der NAC Appliance und Enterasys-Switches

Der Access Layer kann bei Enterasys NAC wie folgt unterschieden werden:

- Enterasys-Policy-fähige Switches, z.B. Enterasys Matrix N-Series
- Switches / WLAN Access Points / WLAN Controller, welche als IEEE 802.1X Authenticator fungieren, RFC3580-kompatibel sind und eine dynamische VLAN-Zuweisung per RADIUS ermöglichen (z.B. für ein Quarantänenetz).
- Switches / WLAN Access Points ohne IEEE-802.1X-Funktionalität

Hinzu kommen Szenarien vom Typ Virtual Private Networks (VPNs), wie z.B. Remote Access VPNs, in denen die Authentisierung der Clients über VPN-Mechanismen erfolgt. Über eine NAC Appliance kann beispielsweise die Integrität der Endsysteme hinsichtlich Firewall-Konfiguration, Virenschutz und Software-Updates geprüft

werden. (siehe Abbildung 6)

Bei Nutzung einer Servervirtualisierung ist die erste Switching-Ebene jedoch der virtuelle Switch (vSwitch) und damit in der Regel ein Switch ohne IEEE 802.1X (Abbildung 7). Szenarien, in denen die erste Switching-Ebene nicht zur Authentisierung geeignet ist, können bei Enterasys durch Platzierung eines Enterasys-Switches in einer höheren Switching-Ebene realisiert werden.

Verwaltet wird die gesamte Lösung über die Enterasys-Management-Software (Enterasys NetSight).

2.2 Auswahl der NAC Appliance (In-band oder Out-of-band)

Der Access Layer ist für die Wahl der NAC Appliance mitbestimmend. Allgemein handelt es sich bei einer NAC Appliance um dedizierte Hardware, welche die eigentliche Zugangskontrolle realisiert. NAC Appliances können in zwei Klassen unterteilt werden:

- In-line Appliances: Diese Komponenten werden in den Datenpfad der Geräte eingebracht, deren Netzzugang zu kontrollieren ist..
- Out-of-band Appliances: Diese Komponenten können auch außerhalb des Datenpfads der Geräte, deren Netzzugang zu kontrollieren ist, angeschaltet wer-

Virtualisierung: Virtualisierungsbewusste RZ-Netze

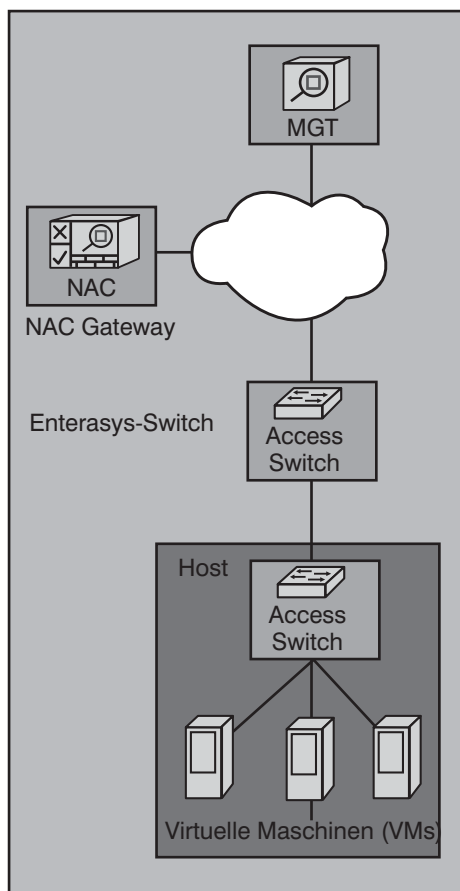


Abbildung 7: Authentisierung in einer höheren Switching-Ebene (hier: am Enterasys-Switch)

den. Hier wird beispielsweise nur der Verkehr für die Authentisierung und Autorisierung verarbeitet, während der normale Datenverkehr die Out-of-band Appliance nicht durchläuft.

Enterasys bietet sowohl eine In-line Appliance, den sogenannten Enterasys NAC Controller, als auch eine Out-of-band Appliance - das Enterasys NAC Gateway; letzteres fungiert als reiner RADIUS-Proxy bzw. RADIUS-Server.

Während der NAC Controller den Datenfluss kontrolliert und Systeme anhand ihrer MAC- (Layer-2-Modus) oder IP-Adresse (Layer-3-Modus) erkennt, erfordert das NAC Gateway eine explizite RADIUS-Anfrage durch einen IEEE 802.1X Authenticator.

Dies bedeutet jedoch auch, dass der Access Layer über eine entsprechende „Intelligenz“ verfügen muss, d.h. mindestens IEEE 802.1X und RFC3580 unterstützt. Funktionen, die über die dynamische VLAN-Zuweisung nach RFC3580 hinausgehen, wie die Zuweisung von ACLs, QoS-Parametern oder Bandbreite über

das Enterasys-Management erfordern hingegen einen Enterasys-Policy-fähigen Switch. Optional können Switches von Drittherstellern beliebige RADIUS-Attribute zugeordnet werden.

Da der NAC Controller direkt in den Datenpfad integriert wird, sind der Aufstellungsort sowie die Anzahl der NAC Controller abhängig vom Netzdesign. Sofern der NAC Controller die Authentisierung per IEEE 802.1X oder MAC-Adresse realisieren soll, muss dieser im Layer-2-Modus arbeiten und damit vor der ersten Routing-Instanz platziert werden. Für Remote-Access-VPN-Szenarien (NAC Controller im Layer-3-Modus) oder vergleichbare Szenarien, in denen die Authentisierung durch Drittsysteme erfolgt, gibt es die Möglichkeit die Authentisierung durch den NAC Controller zu deaktivieren.

Demgegenüber benötigt das NAC Gateway nur eine IP-Verbindung zu den authentisierenden Switches und ist damit generell standortunabhängig. Dies ist ein wesentlicher Vorteil hinsichtlich der Skalierbarkeit gegenüber dem NAC Controller (sowohl in technischer als auch in finanzieller Hinsicht).

Aus Perspektive der Sicherheit bietet eine Switch-basierte NAC-Lösung am Endgeräteanschluss in der Regel ein höheres Sicherheitsniveau als eine Lösung, welche im Distribution oder Core Layer realisiert wird, da die Netzzugangskontrolle direkt an der Quelle erfolgt und der Datenverkehr nicht erst in den Distribution oder Core Layer vordringen kann.

2.3 Integration in die virtuelle Umgebung

Nachfolgend wird ein Szenario betrachtet, welches die in Kapitel 2 aufgeführten Vorteile der Enterasys-Lösung abbilden soll. Ziel soll es sein, drei Sicherheitszonen zu realisieren, ohne ein VLAN-Trunking zwischen Host und Server Switch zu konfigurieren. Die Sicherheitszonen sollen analog den VLANs 811, 812 und 813 zugeordnet werden (Abbildung 8).

Die Enterasys-spezifischen Komponenten des Szenarios sind:

- Enterasys-Policy-fähiger Switch (Enterasys Matrix N-Series) als Server Switch
- Enterasys NAC Appliance: NAC Gateway (Out-of-band)

Da der virtuelle Switch innerhalb des Xen-Server-Hosts derzeit kein IEEE 802.1X beherrscht, müssen die NAC-Prozesse der Authentisierung und Autorisierung in einer höheren Switching-Ebene abgebildet werden (vgl. Abbildung 7). Dies kann

also frühestens am ersten physischen Server Switch geschehen (hier: Enterasys Matrix N).

Grundlage hierfür ist die Enterasys-Funktion „Multi-User Authentication“ (MUA), welche eine Authentisierung von unterschiedlichen Systemen bzw. VMs an einem physischen Switchport ermöglicht (Abbildung 9). Daher muss zur Erkennung, Authentisierung und Autorisierung von VMs mindestens ein Authentisierungsverfahren konfiguriert werden. Hierfür stehen IEEE 802.1X oder eine MAC-Adresse-Authentisierung zur Auswahl. Weitere hier nicht relevante Authentisierungsmethoden sind Convergence Endpoint (CEP) für IP-Telefone sowie Port Web Authentication (PWA) z.B. zur Realisierung eines Gastzugangs per Web-Authentisierung. In unserem Beispiel wird eine MAC-Adresse-Authentisierung genutzt, welche lokal auf dem NAC Gateway realisiert werden kann. Für IEEE 802.1X oder eine Webauthentisierung können beispielsweise Cisco ACS, FreeRADIUS, Funk Steelbeltd RADIUS, Microsoft IAS oder andere RADIUS-kompatible Server eingesetzt werden.

Für die Autorisierung können ACLs und VLANs (hier 811, 812 und 813) über den NetSight Policy Manager definiert und per SNMP auf dem Enterasys-Switch aktiviert werden.

Die Konfiguration des virtuellen Switches beschränkt sich auf die Wahl des Switch-Typs „External Network“ zur Kommunikation mit der Außenwelt sowie der zu nutzenden Netzwerkkarte bzw. logischen Netzwerkkarte bei Verwendung von NIC Bonding (auch als NIC Teaming bekannt).

2.4 Anbindung der Außenwelt

Die Anbindung der Außenwelt erfolgt über den zuvor konfigurierten virtuellen Switch des Xen-Server-Hosts. Die Anbindung zwischen Host und Enterasys-Switch ist ohne VLAN-Trunking konfiguriert. Über eine Policy wurde zuvor festgelegt, welche VM in welche Sicherheitszone und welches VLAN eingeordnet wird. Für noch unbekannte und nicht authentisierte VMs kann beispielsweise ein Gästenetz konfiguriert werden, welches nur eine eingeschränkte Kommunikation erlaubt.

Hat sich eine VM erfolgreich z.B. über die MAC-Adresse authentisiert, erfolgt die Autorisierung durch ein RADIUS-Attribut („Filter-ID“) des NAC Gateway. Inhalt dieses Attributs ist die zuvor konfigurierte Policy einschließlich ACL- und VLAN-Konfiguration auf dem Enterasys-Switch. Kommuniziert diese VM mit der Außenwelt oder einer anderen Sicherheitszone bzw. alle-

Virtualisierung: Virtualisierungsbewusste RZ-Netze

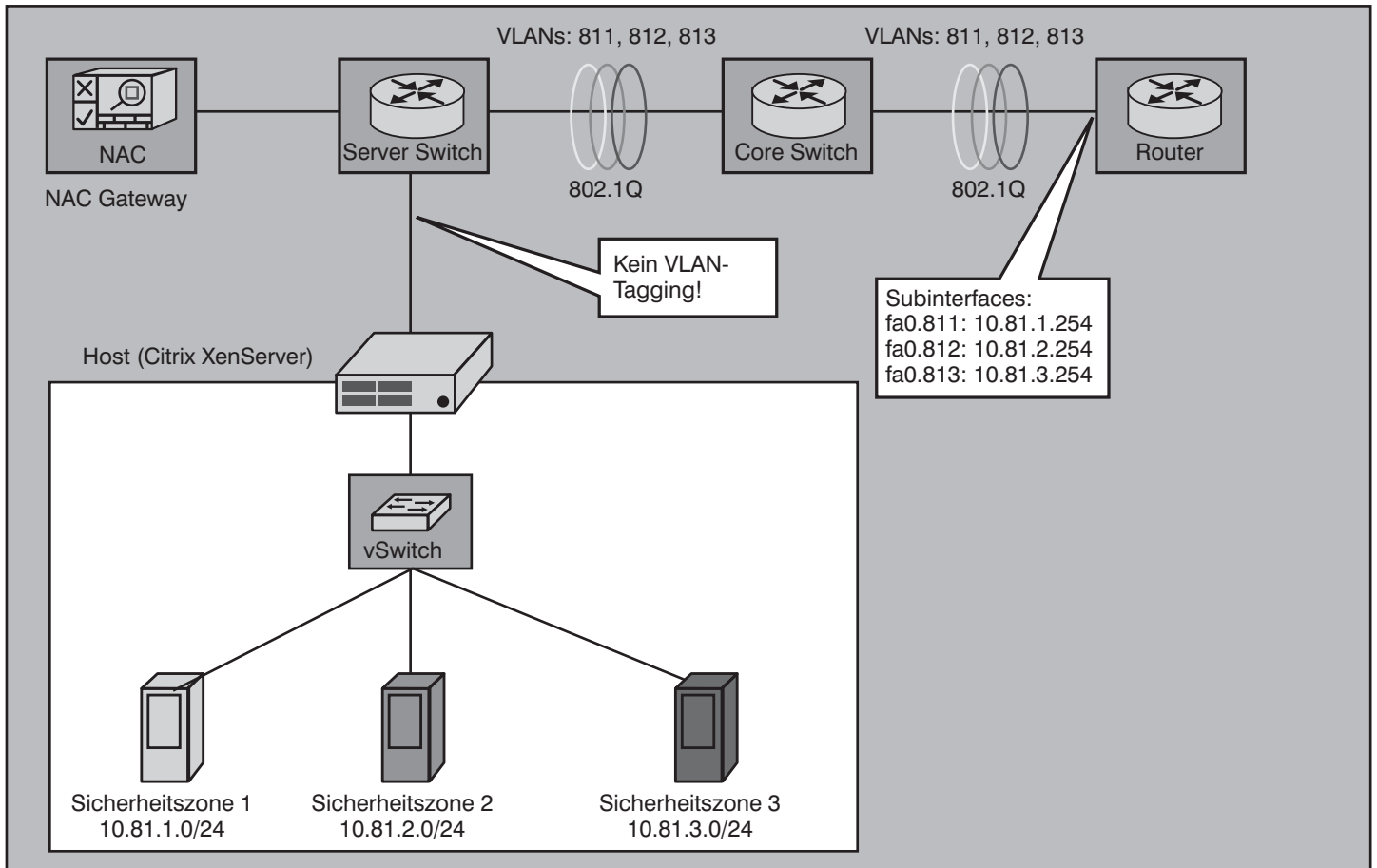


Abbildung 8: Beispielszenario Enterasys NAC und Citrix XenServer

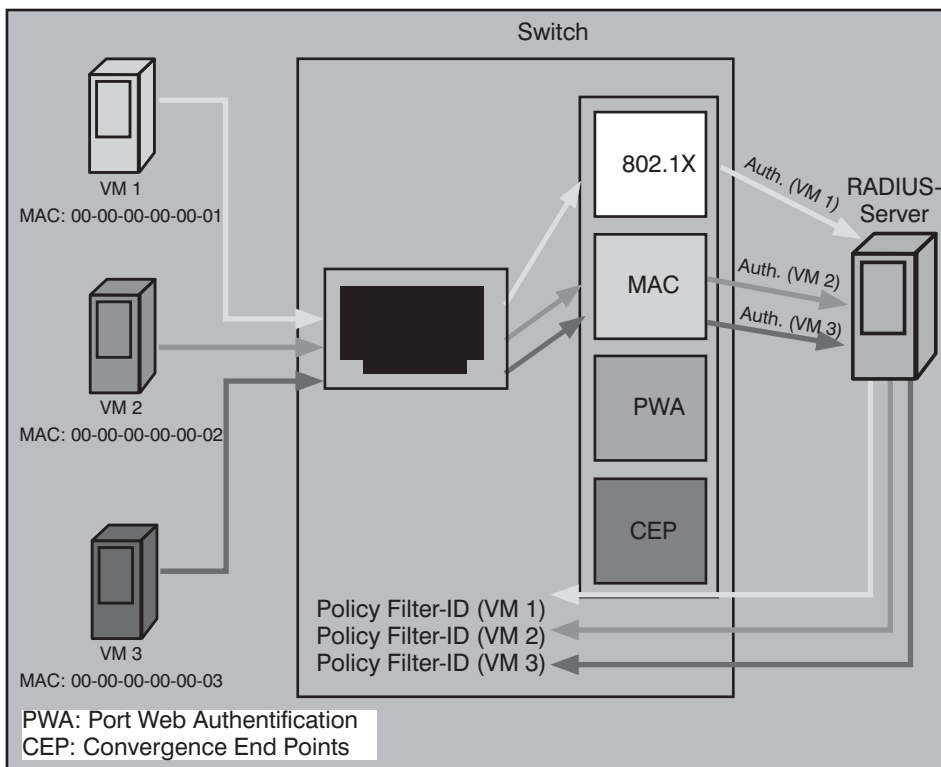


Abbildung 9: Authentisierung unterschiedlicher VMs an einem physischen Switchport

meinen einem anderen IP-Netz, wird der Verkehr erst am Enterasys-Switch mit einem VLAN Tag versehen (Egress VLAN Tagging) und in Richtung Router weitergeleitet (Abbildung 10).

2.5 Interne Kommunikation

Ein anderes Bild ergibt sich für Verkehr innerhalb des Host-Systems (Abbildung 11) bzw. innerhalb einer Sicherheitszone. Dieser Datenverkehr ist für externe physische Komponenten wie beispielsweise den Enterasys-Switch unsichtbar, da der Verkehr nur den virtuellen Switch des Host-Systems passiert.

Aus Perspektive der Sicherheit besteht folgendes Problem: Alle Sicherheitszonen und damit VMs sind in einer Layer-2-Domäne. Dies kann ein Angreifer für Layer-2-Angriffe vom Typ Denial of Service (DoS) oder Man-in-the-Middle (MITM) ausnutzen. Ein einfaches Umkonfigurieren der Netzwerkeinstellungen genügt, um mit anderen VMs des gleichen IP-Netzes zu kommunizieren. Mittels ARP Spoofing kann auf diese Weise jeglicher Verkehr zwischen allen VMs aller Sicherheitszonen aufgezeichnet oder manipuliert werden. Die Konfiguration von Dynamic ARP

Virtualisierung: Virtualisierungsbewusste RZ-Netze

Name	Description	NIC	VLAN	Auto	Link Status	MAC
bond-mgt	Bond 0+1	-	-	No	Connected	00:22:19:57:df:31
bond-vm	Bond 2+3	-	-	Yes	Connected	00:15:17:ae:83:5e
iscsi-network871		NIC 4	-	No	Disconnected	00:15:17:ae:83:64
iscsi-network872		NIC 5	-	No	Connected	00:15:17:ae:83:65

Abbildung 10: Auswahl des vSwitch zur Kommunikation der VMs mit der Außenwelt (hier: bond-vm)

Inspection auf dem physischen Server Access Switch verhindert in diesem Fall nur ein ARP Spoofing des Router.

Sofern keine proprietären Funktionen wie Private VLANs¹ (z.B. mit VMware vSphere 4) genutzt werden, sollte daher die Bildung von Sicherheitszonen je nach erforderlichem Sicherheitsniveau mindestens per VLAN, separater Netzwerkkarte oder Host je Sicherheitszone realisiert werden.

2.6 Sicherheitsrichtlinien und QoS-Merkmale (Policies)

Über das Enterasys-Management, genauer den NetSight Policy Manager, können die Policies für Enterasys-Switches zentral konfiguriert werden. Über die Policies werden die verschiedenen Nutzergruppen (z.B. Entwicklung, Finanzen, Geschäftsleitung) modelliert. Hierzu werden

sogenannte Rollen definiert und den VMs bzw. beliebig konfigurierbaren Gruppen zugewiesen. Dies beinhaltet beispielsweise die Konfiguration von:

- VLANs
- ACLs
- Parametern bzgl. CoS/QoS mittels IEEE 802.1p und DiffServ
- Beschränkung der Bandbreite (Inbound Rate Limit, IRL)

Diese Rollen (Abbildung 12) können den VMs als Ergebnis der Authentisierung dynamisch zugewiesen werden. Wird eine VM beispielsweise von einem Host auf einen anderen Host migriert (Abbildung 13), gilt für die VM weiterhin dasselbe Regelwerk, unabhängig von ihrem physischen Aufenthaltsort. Grundvoraussetzung ist, dass sich beide Systeme in derselben Layer-2-Domäne befinden und

für die Switchports der Host-Systeme die Authentisierung aktiviert wurde.

2.7 Zusammenfassung Enterasys NAC

Das Thema Servervirtualisierung und Sicherheit im RZ wird bei Enterasys mit dem Produkt Enterasys NAC adressiert. Die Grundlage der Enterasys-NAC-Lösung ist eine Authentisierung von mehreren VMs an einem physischen Switch-Port. Dies ist speziell bei Verwendung von IEEE 802.1X eine herstellerspezifische Funktion, da der Standard von 2004 eine solche Nutzung nicht vorsieht. Bei Enterasys wird diese Funktion Multi-User Authentication (MUA) genannt und ermöglicht die Identifizierung von bis zu 2048 Systemen an einem physischen Switchport (Enterasys Matrix N-Series). Jeder einzelnen VM kann hierbei ein Regelwerk (Enterasys Policy) zugewiesen werden, welches ACLs, Bandbreitenbeschränkungen, Priorisierungsmechanismen sowie ein VLAN enthalten kann. Dieses Regelwerk gilt unabhängig vom Aufenthaltsort der VM (z.B. aufgrund einer Migration im laufenden Betrieb).

Verschiedene Sicherheitszonen an einem einzigen virtuellen Switch (und damit einer Layer-2-Domäne) ohne weitere Sicherheitsmerkmale (z.B. IEEE 802.1X) anzubinden und erst am physischen Server Switch beispielsweise per Enterasys Policy zu separieren, ist aufgrund von Gefährdungen wie z.B. Man-in-the-Middle oder Denial of Service - ebenso wie in der phy-

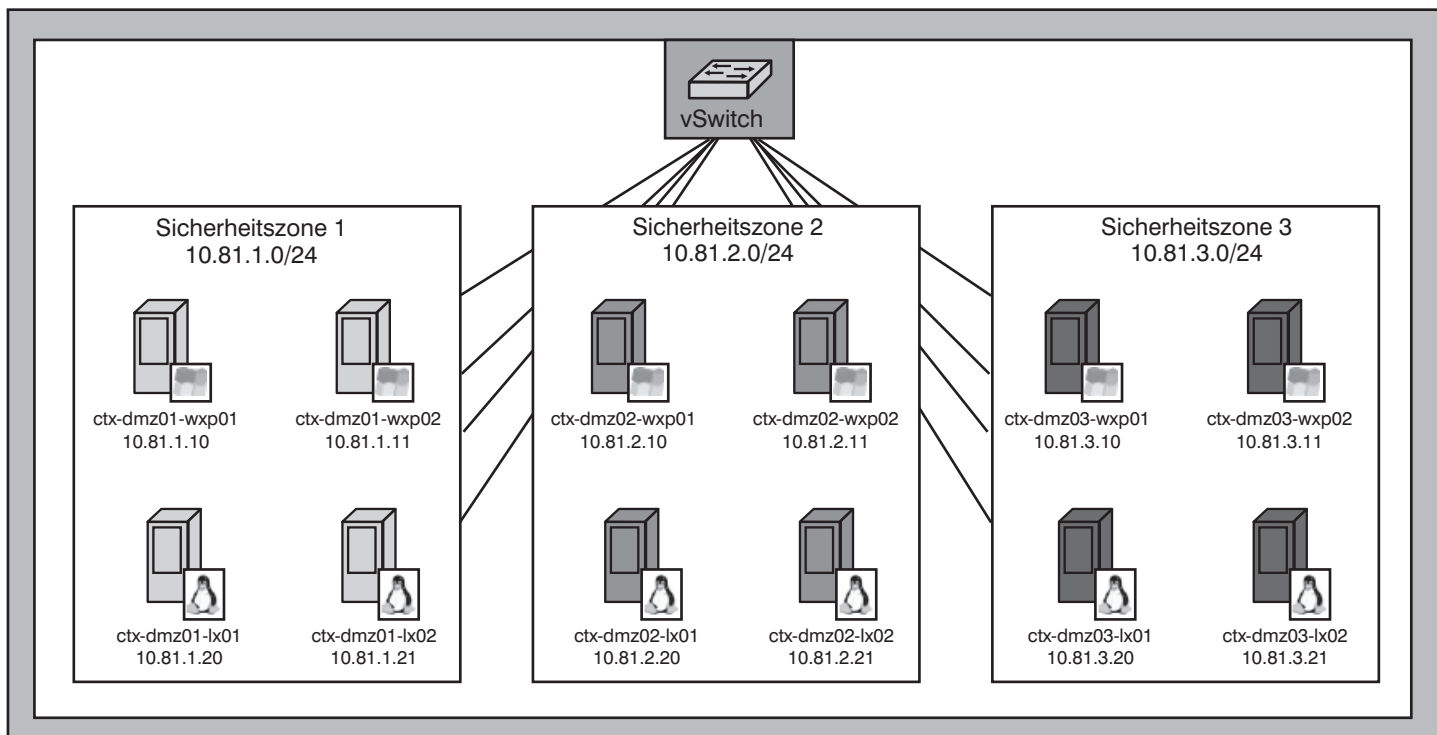


Abbildung 11: Unterschiedliche Sicherheitszonen an einem vSwitch eines Host-Systems

¹ Mittels Private VLANs (PVLANS) ist es möglich, auch die Kommunikation innerhalb eines VLAN einzuschränken, so dass beispielsweise einzelne Systeme isoliert werden können oder nur bestimmten Systemen innerhalb eines VLAN die Kommunikation gestattet wird.

Virtualisierung: Virtualisierungsbewusste RZ-Netze

Contents of 'Roles'			
Name	Access Control	CoS	Number of Rules
Administrator	Permit Traffic	None	0
Assessing	Deny Traffic	None	25
Citrix_VM_811	Contain to VLAN (811[ctx-dmz01])	None	0
Citrix_VM_812	Contain to VLAN (812[ctx-dmz02])	None	0
Citrix_VM_813	Contain to VLAN (813[ctx-dmz03])	None	0
Deny Access	Deny Traffic	None	20
Enterprise Access	Permit Traffic	Critical Data [802.1p: 3]	49
Enterprise User	Permit Traffic	Network Control [802.1p: 4]	50
Failsafe	None	None	0
Guest Access	Permit Traffic	Best Effort [802.1p: 1]	91
Quarantine	Deny Traffic	None	24
Unregistered	Deny Traffic	None	24

Abbildung 12: Definition der Nutzergruppen/Rollen mittels NetSight Policy Manager

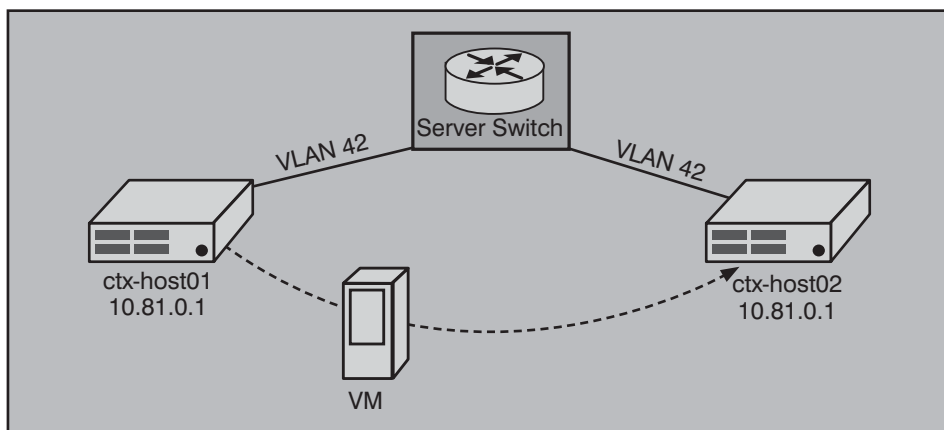


Abbildung 13: Migration einer VM unter Beibehaltung von VLAN, ACLs und CoS/QoS-Parametern

sischen Welt - nicht zu empfehlen. Hier gilt weiterhin die Empfehlung, auf einem Host nur VMs einer Sicherheitszone zu betreiben und auf proprietäre Funktionen weitestgehend zu verzichten.

Eine weitergehende „Verzahnung“ mit der virtuellen Infrastruktur ist derzeit nicht gegeben, wie beispielsweise ein Management auf Ebene des virtuellen Switches. Enterasys plant in diesem Bereich eine Integration in die virtuelle IT-Infrastruktur über die jeweiligen APIs von Citrix und VMware, um Einblicke in die virtuelle Vernetzung zukünftig über ein einheitliches Management realisieren zu können. Hierdurch wäre eine klare Trennung zwischen Netz- und Serverbetrieb möglich.

3. Cisco Nexus 1000v

Der Nexus 1000v ist ein virtueller Switch, der von Cisco als reine Software-Lösung für die VMware vSphere 4 Virtualisierungslösung angeboten wird. Der Switch stellt sich für die Netzwerkseite wie ein physisches System dar und wird in gleicher Weise administriert, da er ebenfalls mit dem Nexus-OS als Betriebssystem ausgestattet ist. Der Switch macht sich die vNet-

work API zu nutze, um sich in den Hypervisor von VMware zu integrieren. Daher ist dieser Switch derzeit nicht für andere Virtualisierungslösungen einsetzbar. Ob es zukünftig auch eine Integration in Citrix XenServer geben wird, bleibt fraglich, da es diesbzgl. von den beteiligten Herstellern unterschiedliche Aussagen gibt. Die folgenden Abschnitte stellen die Grundarchitektur des Nexus 1000v, seine Integration in die virtuelle Umgebung sowie die konsistente Administration von Ende-zu-Ende Sicherheitsrichtlinien durch die Netzwerker dar.

Untersucht wurde der Nexus 1000v im Rahmen eines Testaufbaus im ComConsult-eigenen Virtualisierungs-Labor auf Basis von VMware ESXi 4.0.0.

3.1 Grundarchitektur

Analog zu zahlreichen modularen Switches der physischen Welt wird auch der Cisco Nexus 1000v in Control Plane und Data Plane Komponenten unterschieden. Als Control Plane fungieren sogenannte Virtual Supervisor Module (VSM). Stellt man sich den Nexus 1000v mit einem virtuellen Chassis vor, kann dieses Gehä-

se in den ersten beiden Slots bis zu zwei VSMs aufnehmen. Aus Redundanzgründen wird empfohlen, zwei VSM im Active/Standby-Modus zu betreiben. Die Data Plane des Switches wird durch sog. Virtual Ethernet Module (VEM) realisiert. Jeder Nexus 1000v kann bis zu 64 solcher Linecards in den Slots 3 bis 66 aufnehmen. Da jede dieser Linecards bis zu 256 Virtual Interfaces (VIF) zur Anbindung von VMs besitzen kann, ist theoretisch eine Gesamtdichte von 16384 Ports pro Nexus 1000v denkbar (siehe Abbildung 14).

3.2 Integration in die virtuelle Umgebung

Die VSM des Nexus 1000v werden als virtuelle Maschinen (VM) in der virtualisierten Umgebung gestartet. Zur Installation kann entweder auf ein .OVF-Template zurückgegriffen werden oder es wird eine VM manuell mit den erforderlichen Hardware-Eigenschaften angelegt und das VSM von einem ISO-Image auf dieser VM installiert. Da die beiden VSMs Redundanz herstellen sollen, sollten sie niemals auf dem gleichen physischen Host-System laufen. Um eine manuelle oder automatische Migration der VSMs auf den gleichen Host zu verhindern, sollte DRS für diese VMs abgeschaltet und eine negative Affinität zwischen den VMs definiert werden. Durch diese Konfiguration wird auch ein versehentliches manuelles Zusammenführen der VMs verhindert. Zukünftig soll das VSM auch als separate physische Appliance verfügbar sein, um die flexible Nutzung des Host-Systems nicht einzuschränken und die dortigen Ressourcen

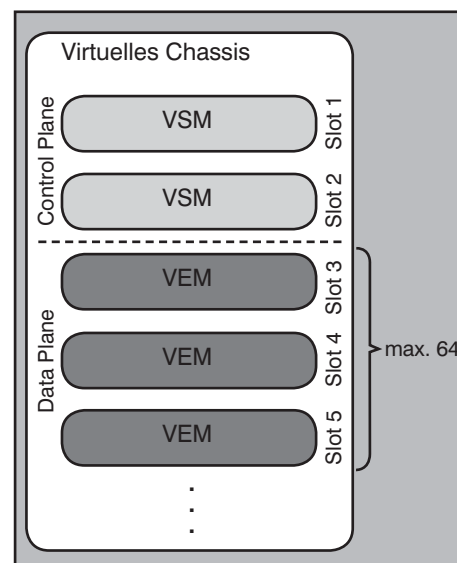


Abbildung 14: Das virtuelle Chassis eines Cisco Nexus 1000v. Die Control Plane besteht aus 2 Virtual Supervisor Modules (VSM) in den Slots 1 und 2. In den Slots 3 bis 66 können bis zu 64 Virtual Ethernet Modules (VEM) – vergleichbar den Linecards eines physischen Switches – installiert werden.

Virtualisierung: Virtualisierungsbewusste RZ-Netze

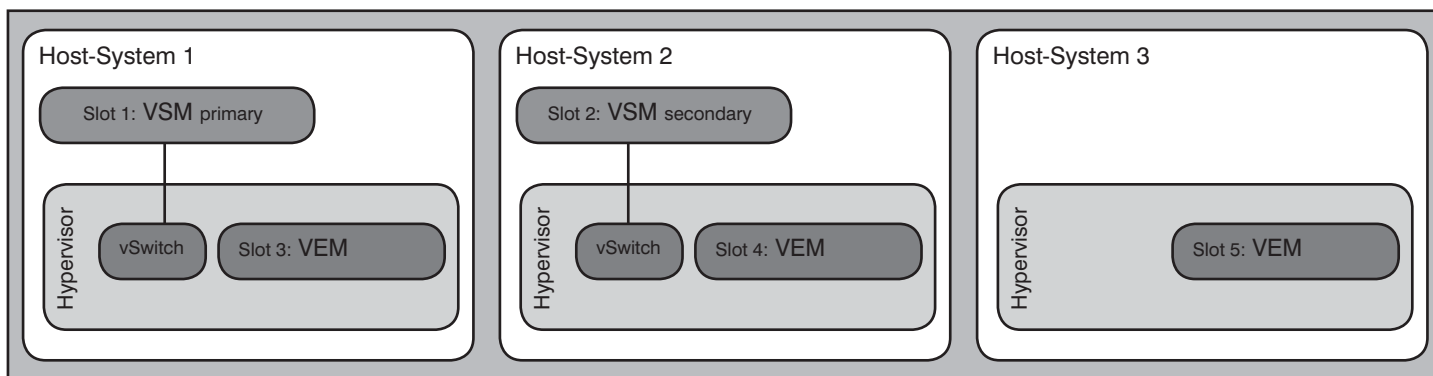


Abbildung 15: Die einzelnen Komponenten des Cisco Nexus 1000v verteilt auf 3 Host-Systeme: die Virtual Supervisor Modules (VSM) werden in Form von virtuellen Maschinen betrieben. Dabei nimmt ein VSM die Primärrolle ein, während das zweite VSM die Redundanzfunktion übernimmt. Die Virtual Ethernet Modules (VEM) werden durch einen VMkernel Patch in den Hypervisor von VMware integriert und bilden dort einen virtuellen Switch.

nicht zu beanspruchen. Andererseits widerspricht eine separate physische Komponente dem Konsolidierungsgedanken der Virtualisierung.

Das VEM integriert sich über die vNetwork API in den Hypervisor und nimmt dort die Rolle eines Distributed Virtual Switch (DVS) ein. Diese Integration wird ähnlich einem Hypervisor-Patch vorgenommen und kann insofern vom VMware Update

Manager (VUM) durchgeführt werden. Theoretisch können mehrere VEMs in einem ESX Host installiert werden. Aufgrund der oben erwähnten Portdichte ist es jedoch in der Regel weder sinnvoll noch erforderlich, mehr als ein VEM gleichzeitig zu betreiben.

Wichtig zu verstehen ist bei dieser Architektur, dass ein Nexus 1000v nicht pro Host-System installiert wird, sondern dass

dieser Switch das Host-übergreifende, verteilt agierende Konstrukt, bestehend aus zwei VSMs und diversen VEMs darstellt.

Damit wird auch klar, dass die ebenfalls oben erwähnte Gesamtanzahl von 64 VEMs pro Nexus 1000v eher akademischer Natur ist. Schließlich können Stand heute nur maximal 32 ESX-Hosts in einem gemeinsamen Cluster betrieben werden. (siehe Abbildung 15)

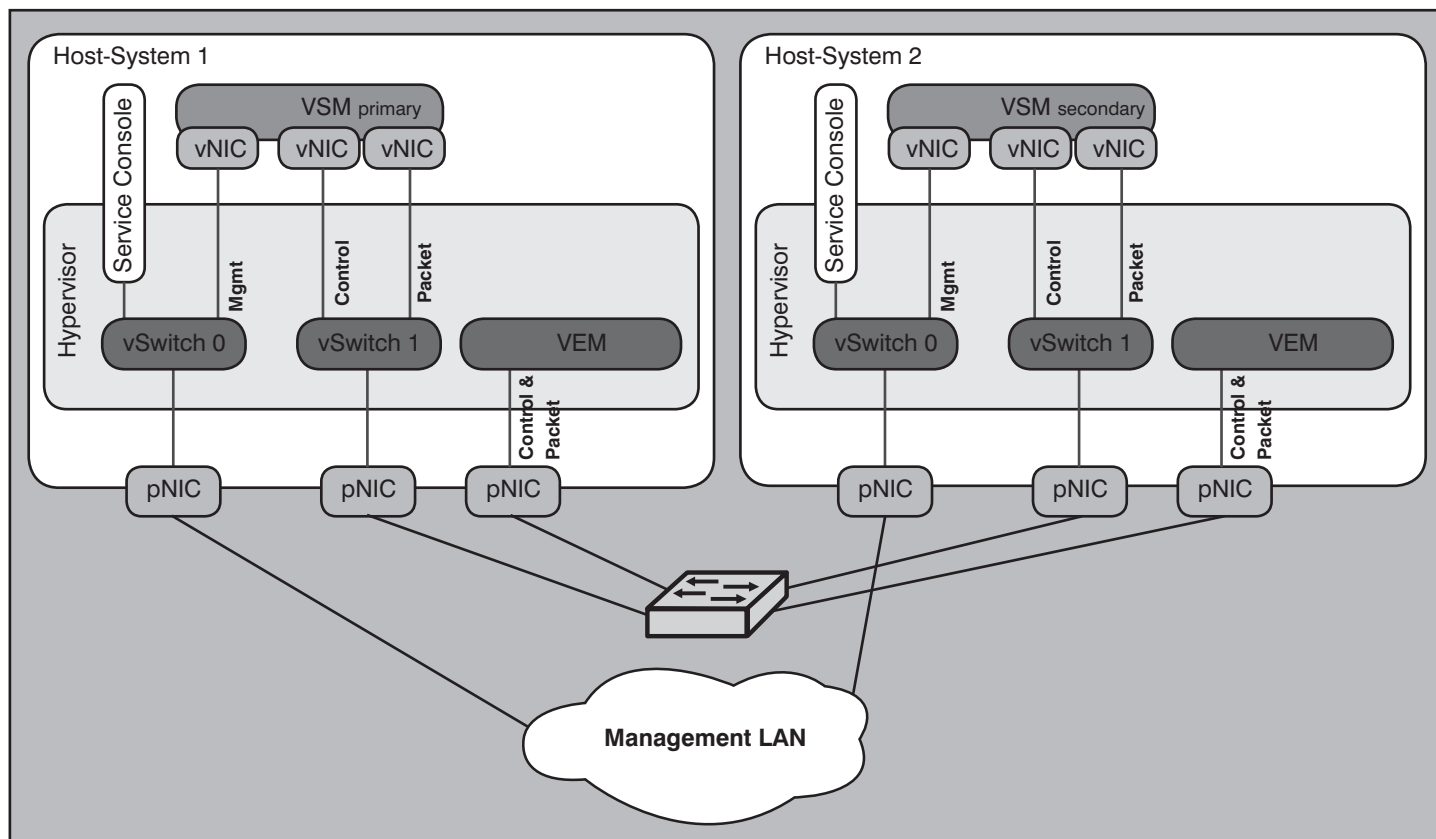


Abbildung 16: Die Konfiguration des Cisco Nexus 1000v erfolgt über die Management-Schnittstelle des primären Virtual Supervisor Modules (VSM). Über zwei weitere Schnittstellen tauschen die VS-Module Statusinformationen aus. Die „Control“-Schnittstelle übermittelt u.a. Heartbeat-Informationen an das sekundäre VSM. Über die „Packet“-Schnittstelle, werden CDP- und IGMP-Pakete - soweit erforderlich - ausgetauscht. Beide Schnittstellen kommunizieren auch mit den Virtual Ethernet Modulen (VEM), so dass hierfür eine weitere physische Netzwerkschnittstelle (pNIC) oder ein geeigneter VLAN-Trunk auf vorhandenen pNICs erforderlich wird. Die physischen und virtuellen Schnittstellen, über die das VEM die Anbindung virtueller Maschinen an die Außenwelt realisiert, sind in dieser Darstellung nicht enthalten.

Virtualisierung: Virtualisierungsbewusste RZ-Netze

3.3 Kommunikation der Nexus Komponenten untereinander

Wie auch bei physischen Switches ist zwischen der Control Plane (VSM) und der Data Plane (VEM) des Nexus 1000v ein reger Informationsaustausch erforderlich. Für den Austausch von Kontrollinformationen zwischen den beiden redundanten VSMs, für deren Management sowie für die Kommunikation zwischen VSM und VEM sind spezielle Netze in der virtualisierten Umgebung zu konfigurieren.

Das Architekturmodell des Nexus 1000v sieht für den Informationsaustausch ein Control VLAN und ein Packet VLAN vor. Ersteres dient u.a. der Übermittlung von Heartbeat-Signalen, während letzteres mittels CDP und IGMP Daten übermittelt. Der Bandbreitenbedarf dieser VLANs ist nach Angaben des Herstellers eher gering. Eine Gigabit-Ethernet-Verbindung stellt hierfür eine mehr als ausreichende Bandbreite zur Verfügung, so dass zur Einsparung physischer Netzwerkschnittstellen auf dem Host-System diese VLANs in Form eines VLAN-Trunks über noch anderweitig genutzte Schnittstellen geführt werden können.

Für das Management des Nexus 1000v ist jedes VSM mit einem Management VLAN auszustatten. Da die Managebarkeit des Switches die gleiche Bedeutung hat, wie das Management des Host-Systems insgesamt, kann dieses Management VLAN

über den gleichen virtuellen Switch und die gleichen redundanten physischen Interfaces realisiert werden, wie der Service Console Zugang. In Zeiten knapper Netzwerkschnittstellen können hierdurch ebenfalls weitere physische Ports eingespart werden. (siehe Abbildung 16)

Theoretisch können nach erfolgreicher Installation des Nexus 1000v die Control- und Packet-Verbindungen eines VSM auf das VEM umgezogen werden. Sobald ein drittes Host-System eingesetzt wird, das keines der VSMs hosted (wie in Abbildung 15), bleibt jedoch die Notwendigkeit nach Außen geführter Control- und Packet-VLANs bestehen, um die Kommunikation mit dem externen VSM zu ermöglichen. Insofern bietet es sich im Sinne einer reduzierten Komplexität an, diese Netze auf einem virtuellen Standard Switch zu belassen. Einzig zu bedenken gilt es in diesem Zusammenhang, dass diese Netze, i.e. Portgruppen, auf allen Host-Systemen vorliegen müssen, auf die ein VSM migriert werden sollen. In der Regel empfiehlt es sich jedoch ohnehin, die VSM statisch einzurichten und an keinen dynamischen Migrationenvorgängen (z.B. DRS, DPM) teilnehmen zu lassen.

3.4 Anbindung der Außenwelt und virtueller Maschinen

Die VEM können i.W. mit zwei verschiedenen Port-Typen konfiguriert werden: Uplink-Ports und Access-Ports. Die Uplink-

Ports stellen über die physischen Netzwerkschnittstellen des Host-Systems – in der VMware Terminologie als vNIC bezeichnet – die Konnektivität mit den physischen RZ-Netzkomponenten her. VEM-seitig werden Ethernet-Ports als Uplink-Ports definiert. Die Bezeichnung dieser Ethernet-Ports entspricht der bei modularen Switches üblichen Konvention: so wird über ETH3/4 der vierte Port auf dem ersten VEM adressiert, das sich in Slot 3 des virtuellen Nexus 1000v Chassis befindet.

Die Anbindung virtueller Maschinen erfolgt über deren virtuelle Netzwerkschnittstelle (vNIC) an die Virtual Ethernet Ports (vETH-Ports), die damit als Access Ports fungieren. Das Besondere an diesen vETH-Ports ist, dass deren Bezeichnung eindeutig mit der VM gekoppelt ist, die mit ihm verbunden ist. Innerhalb des Distributed Virtual Switch vom Typ Nexus 1000v ist diese Port-Nummer immer gleich, egal an welchem VEM und damit auch egal auf welchem Host-System die VM läuft. Ist eine VM also beispielsweise mit dem VEM auf Host 1 an vETH-Port 74 verbunden und wird per vMotion auf Host 2 migriert, wird sie auch dort am vETH-Port 74 zu finden sein (siehe Abbildung 18). Die vETH-Port IDs werden demzufolge fortlaufend durchnummeriert, da sie keine Referenz zur Slot-ID eines VEM erfordern. Damit sind es die vETH-Ports, die den Host-System-übergreifenden Teil des Distributed Virtual Switches ausmachen.

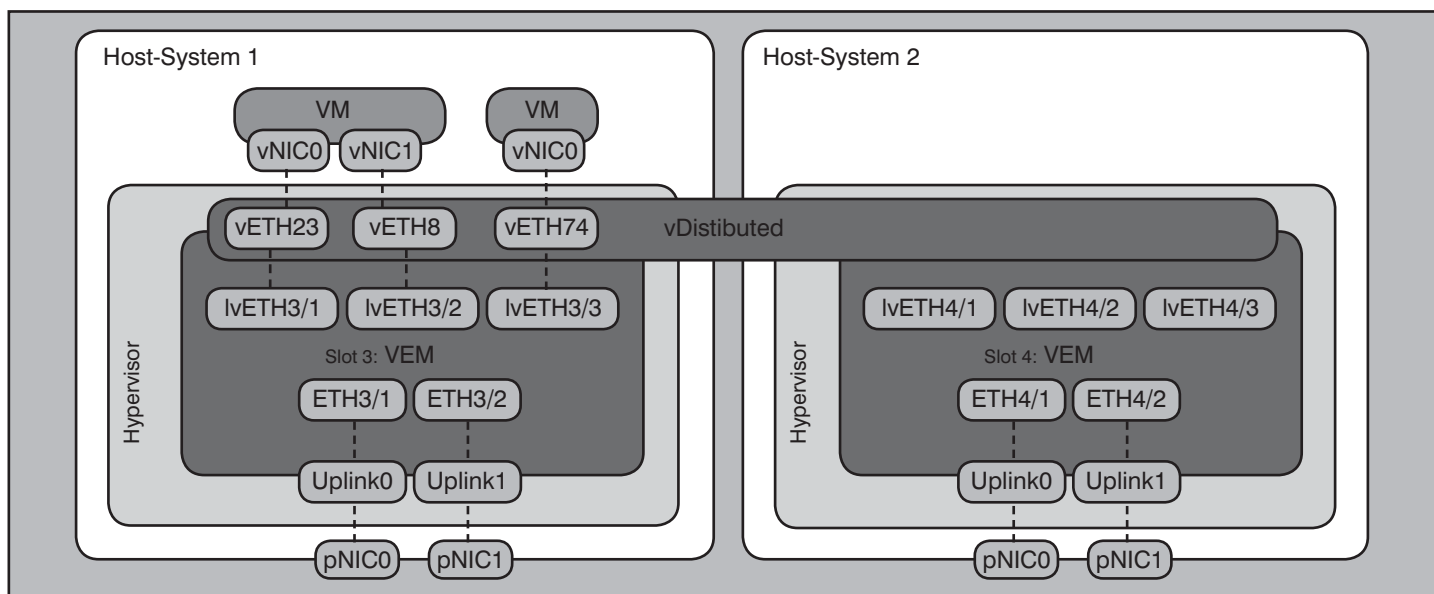


Abbildung 17: Port-Struktur der Virtual Ethernet Module (VEM): intern wird zwischen Ports in Richtung physischer Infrastruktur und in Richtung virtueller Maschinen (VM) unterschieden. Mit „ETH<SlotNr>/<PortNr>“ werden Ports bezeichnet, die über einen „Uplink“ mit den physischen Netzwerkschnittstellen (pNIC) des Host-Systems verbunden sind und damit den Übergang in die RZ-Infrastruktur realisieren. Local Virtual Ethernet (lvETH) Schnittstellen stellen die Ports in Richtung VMs zur Verfügung und sind durch ihr Nummerierungsschema „<SlotNr>/<PortNr>“ ebenfalls dem VE-Modul eindeutig zuzuordnen. Eine lvETH-Schnittstelle wird jedoch mit einer Virtual Ethernet (vETH) Schnittstelle assoziiert, die über den gesamten Nexus 1000v Switch fortlaufend nummeriert wird. Erst an diesen vETH-Schnittstellen werden die virtuellen Schnittstellen (vNIC) der VMs konnektiert. Da eine 1:1-Zuordnung zwischen vNIC einer VM und einer vETH-Schnittstelle im Host-System-übergreifend installierten Nexus 1000v Switch besteht, sind es genau diese vETH-Schnittstellen, die das Host-System-übergreifende „verteilte“ Switching ermöglichen (Distributed Virtual Switch, vDistributed).

Virtualisierung: Virtualisierungsbewusste RZ-Netze

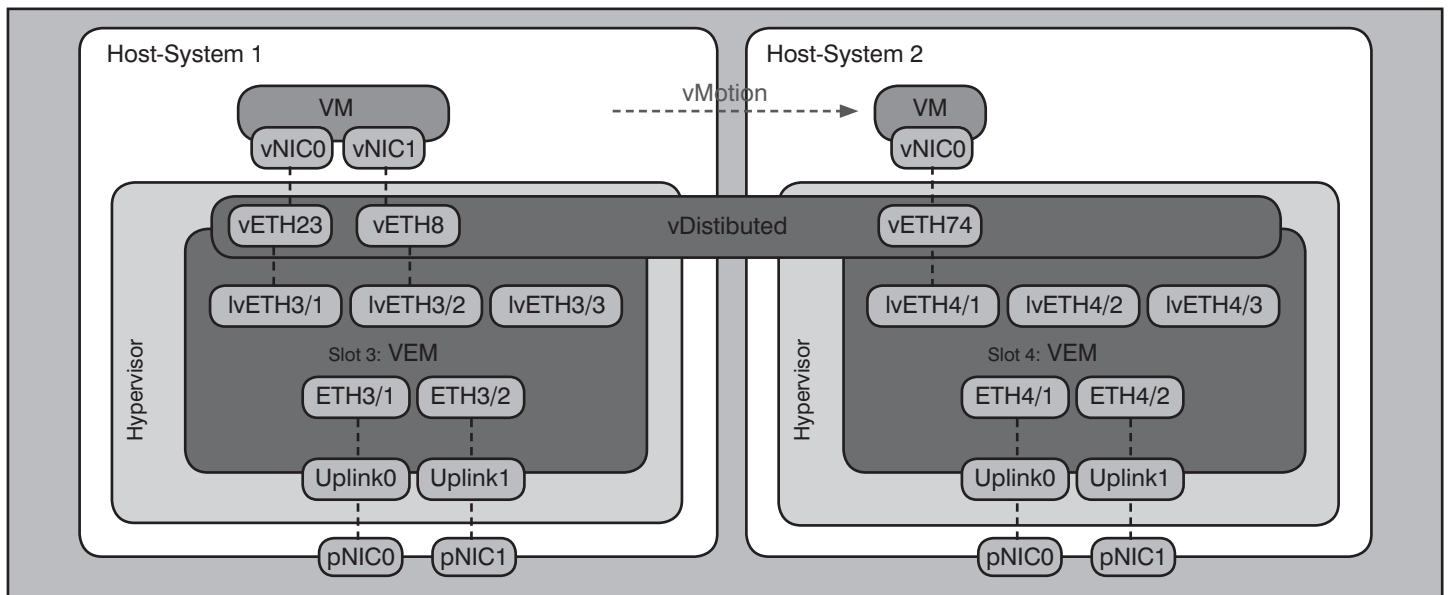


Abbildung 18: Nach der Migration einer virtuellen Maschine (VM) auf ein anderes Host-System wird die VM zwar an dem dortigen Local Virtual Ethernet (lvETH) Port des Virtual Ethernet Modules (VEM) angeben, die Virtual Ethernet (vETH) Portnummer ist aber gegenüber dem ersten Host-System unverändert geblieben, da die vETH-Schnittstellen innerhalb des Host-System-übergreifend installierten Nexus 1000v Switches eindeutig sind.

Eine eindeutige Bezeichnung der Access-Ports eines VEM erfolgt dadurch, dass jeder vETH-Port mit einem Local Virtual Ethernet Port (lvEth-Port) assoziiert wird. Dieser lvEth-Port ist wiederum für jedes VEM eindeutig identifizierbar, da er entsprechend der Bezeichnungskonvention für Ports eines physischen Switches die Slot-ID beinhaltet, in der sich das VEM innerhalb des virtuellen Nexus Chassis befindet. Der erste Access-Port des ersten VEM in Slot 3 wird also über lvEth 3/1 adressiert, während der dritte Access-Port des dritten VEM in Slot 5 die Bezeichnung lvEth 5/3 trägt (siehe Abbildung 17 und 18).

3.5 Sicherheitsrichtlinien und QoS-Merkmale (Policies)

Mit Hilfe des im letzten Abschnitt beschriebenen Nummerierungsschemas ist es nun gelungen, den virtuellen Port, mit dem eine VM verbunden ist, über die gesamte Virtualisierungsinfrastruktur hinweg eindeutig zu identifizieren. Egal auf welchem Host-System diese VM im Laufe der Zeit betrieben wird, ist sie immer über diese vETH-Port ID zu adressieren.

Werden nun eine Sicherheitsrichtlinie oder QoS-Merkmale definiert, denen diese Port ID zugeordnet ist, gelten diese Leistungsmerkmale für die betroffene VM in der gesamten Virtualisierungslösung. Der Nexus 1000v Switch gestattet gegenüber dem Standard Distributed Virtual Switch die Definition zahlreicher zusätzlicher Sicherheits- und QoS-Funktionen, wie z.B.:

- IP Source Guard
- Dynamic ARP Inspection
- Quality of Service über Differentiated Services Code Point (DSCP), Type of Service (ToS) und Class of Service (CoS) Werte

Darüber hinaus bietet der Nexus 1000v erweiterten Multicast Support, zusätzliche Lastverteilungsalgorithmen, LACP-Support sowie zusätzliche Netzmanagementfunktionen in Form von SPAN, ERSPAN, NetFlow v9 sowie Paketanalyse-Funktionen an.

All diese Leistungsmerkmale können in geeigneten Portprofilen gemeinsam mit einer VLAN-ID zusammengestellt werden. Diese Portprofile werden über die vNetwork API und die vCenter API an die Virtualisierungslösung übergeben und dort in Form von Portgruppen zur Verfügung gestellt. Der Serveradministrator verbindet die virtuellen Netzwerkschnittstellen seiner virtuellen Maschinen mit diesen Portgruppen. Die Zuordnung zu einer virtuellen Ethernet-Schnittstelle (vETH) des Nexus1000v erfolgt durch den virtuellen Switch automatisch (siehe Abbildung 19).

Wie eingangs beschrieben wird aus Sicht des Serveradministrators die virtuelle Netzwerkschnittstelle einer VM mit einem virtuellen Switch oder einer dort konfigurierten Portgruppe verbunden. Damit wird diese VM einem speziellen Netzsegment zugeordnet und unterliegt der dortigen Sicherheits- und QoS-Richtlinie. Um das Zusammenspiel von Server- und Netzwerkadministration nun zu vereinfachen, reicht

- Access Control Lists
- DHCP Snooping

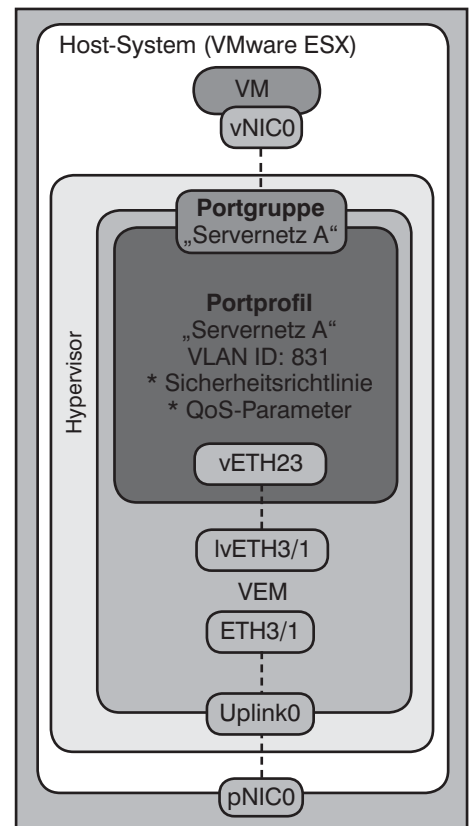


Abbildung 19: Zuordnung und Zuständigkeiten zwischen Virtual Ethernet Module (VEM) und der Servervirtualisierungslösung: im Nexus 1000v: die Portprofile des virtuellen Switches werden an den Hypervisor der Virtualisierungslösung durchgereicht und dort in Form von Portgruppen dargestellt. Der Serveradministrator verbindet die virtuelle Netzwerkschnittstelle (vNIC) seiner virtuellen Maschine (VM) mit dieser Portgruppe. Die zugehörige virtuelle Ethernetschnittstelle (vETH) legt der Nexus 1000v auf dem VEM automatisch an.

Virtualisierung: Virtualisierungsbewusste RZ-Netze

```
VIT-Nexus1000v# configure terminal
VIT-Nexus1000v(config)# port-profile VIT-Servernetz_A
VIT-Nexus1000v(config-port-prof)# switchport mode access
VIT-Nexus1000v(config-port-prof)# switchport access vlan 831
VIT-Nexus1000v(config-port-prof)# vmware port-group VIT-Servernetz_A
VIT-Nexus1000v(config-port-prof)# no shut
VIT-Nexus1000v(config-port-prof)# state enabled
VIT-Nexus1000v(config-port-prof)# exit
```

Abbildung 20: Anlegen eines Portprofils unter Nexus-OS auf dem Nexus 1000v. Über den „vmware port-group“ Befehl wird der Name definiert, unter dem das Portprofil als Portgruppe in der Virtualisierungslösung angelegt wird. Das „state enabled“ Kommando aktiviert diese neue Portgruppe indem der Nexus-Switch mit dem Hypervisor über die vNetwork API und die vCenter API kommuniziert.

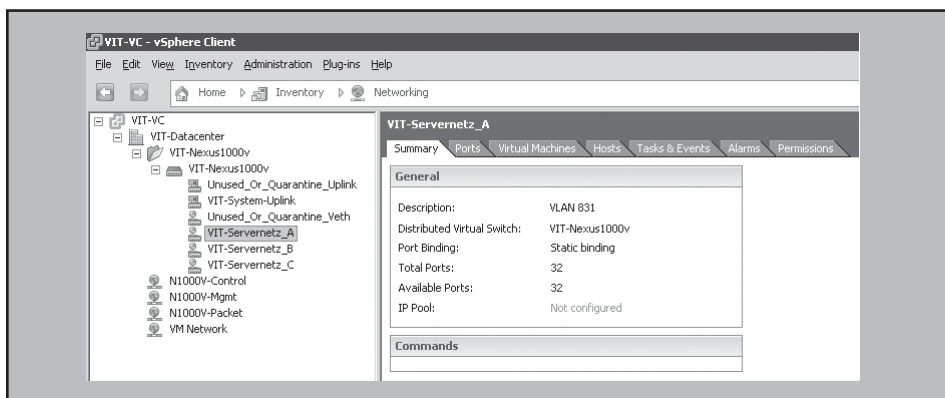


Abbildung 21: Darstellung der neu angelegten Portgruppe „VIT-Servernetz_A“ in der vCenter Managementumgebung.

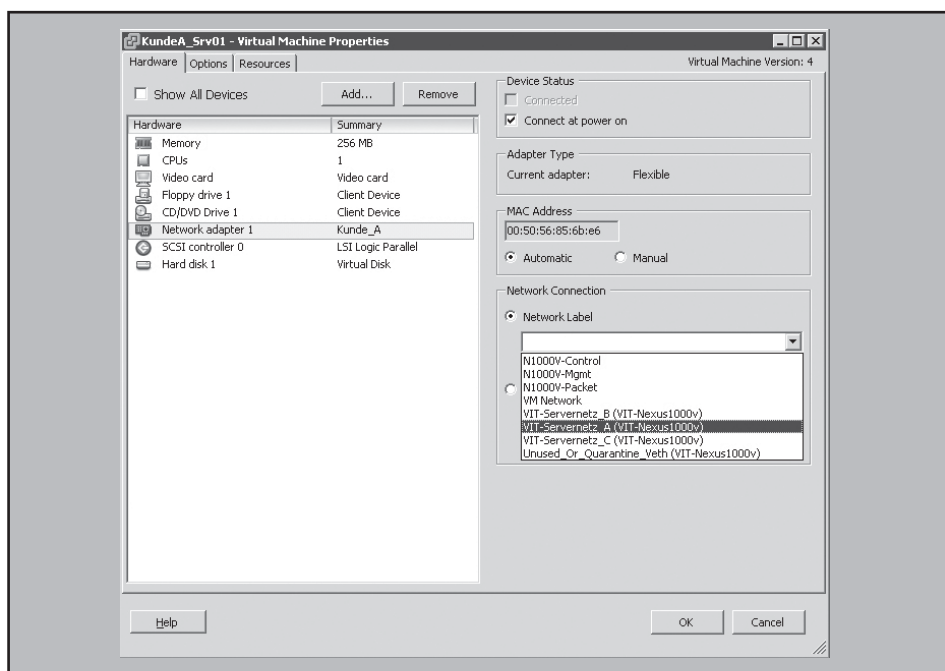


Abbildung 22: Auswahl der neuen Portgruppe „VIT-Servernetz_A“ als Netzbereich, mit dem die virtuelle Netzwerkschnittstelle der virtuellen Maschine „KundeA_Srv01“ verbunden werden soll.

der Nexus 1000v die vom Netzwerkadministrator angelegten Portprofile über die vNetwork API an die VMware-Umgebung weiter und erzeugt dort entsprechende Portgruppen. Der Serveradministrator kann nun die virtuellen Netzwerkschnittstellen seiner VMs mit diesen auf Portprofilen des Nexus 1000v basierenden Portgruppen verbinden. Abbildung 21 und Abbildung 22 zeigen die Darstellung des Portprofils als Portgruppe in der vCenter-Managementumgebung und wie der Serveradministrator die Netzwerkschnittstelle einer virtuellen Maschine dieser Portgruppe zuordnet.

Abbildung 23 zeigt die Portkonfiguration des Nexus 1000v mit der angebotenen virtuellen Maschine aus Sicht der vCenter Managementoberfläche. Die gleiche Kontrolle besitzt nun auch der Netzadministrator über seine Netzadministrationswerkzeuge. Abbildung 24 zeigt, wie mit Bordmitteln des Nexus-OS die Konnektivität einer virtuellen Maschine an einem Port des Nexus 1000v kontrolliert werden kann.

4. Ausblick: Ablösung des virtuellen Switches

Betrachtet man den Leistungsbedarf des Nexus 1000v, so muss man zwischen dem VSM (Control Plane) und dem VEM (Data Plane) unterscheiden. Dem als VM realisierten VSM müssen 1500 MHz der CPU-Ressourcen und 2 GB Hauptspeicher fest zugeordnet werden. Zusätzlich werden 3 GB Festplattenspeicher belegt. Das VEM ist über die vNetwork API mit dem Hypervisor verbunden und regelt den Kommunikationsfluss zwischen den virtuellen Netzwerkschnittstellen der VMs und den physischen Netzwerkschnittstellen des Host-Systems. Die genauen Ressourcenanforderungen dieses Moduls können ohne umfangreiche Labortests nicht näher benannt werden. Das in Abbildung 17 dargestellte Architekturmodell lässt jedoch unschwer erahnen, dass diese Funktionalität gerade bei einer hohen Anzahl von VMs und aufwendigen Portprofilen nicht ohne merklichen Overhead einhergehen wird.

Insofern müssen Lösungen für virtualisierte Umgebungen entwickelt werden, die zwar die gleichen Vorteile wie die bereits genannten bieten (einheitliche Administration, garantierte Ende-zu-Ende Richtlinien), jedoch ohne aufwendige Software-Implementierungen innerhalb der Virtualisierungslösung auskommen.

Ciscos Virtual Network Link (VN-Link) ist ein Beispiel für diesen Ansatz einer logi-

Virtualisierung: Virtualisierungsbewusste RZ-Netze

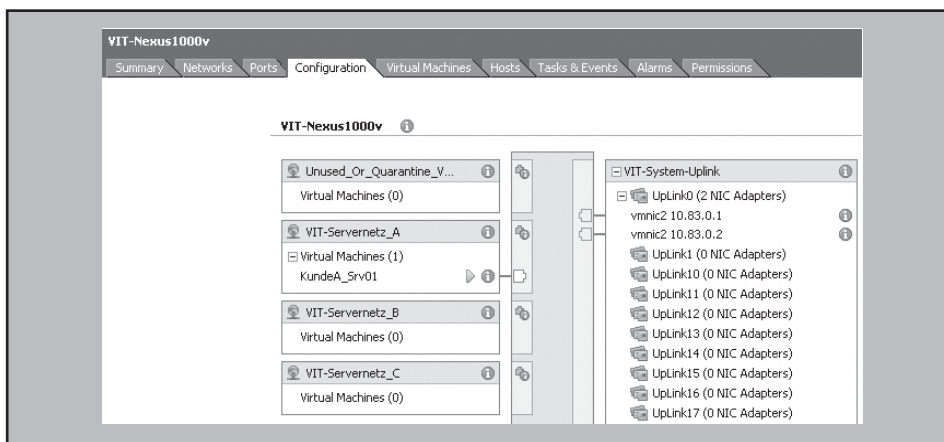


Abbildung 23: vCenter-Ansicht der Portkonfiguration des Nexus1000v inkl. der mit der Portgruppe „VIT-Servernetz_A“ verbundenen virtuellen Maschine „KundeA_Srv01“

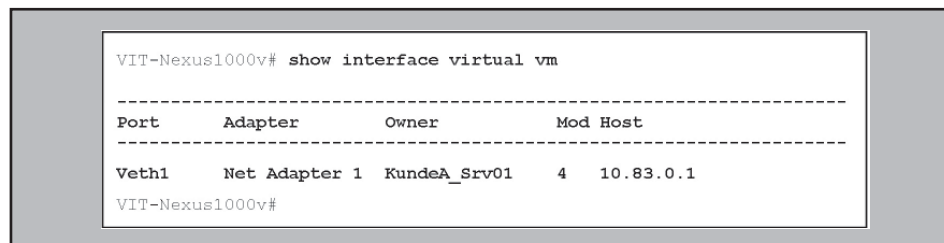


Abbildung 24: Mit Bordmitteln des Nexus-OS kann der Netzwerkadministrator kontrollieren, welche virtuelle Maschine mit welcher virtuellen Ethernet-Schnittstelle (Veth) des Nexus1000v verbunden ist.

schen Verbindung zwischen VM und einem externen physischen Switch. Hierbei wird die virtuelle Netzwerkschnittstelle der VM (vNic) über die vNetwork API direkt mit der physischen Netzwerkschnittstelle (vmNic) des Host-Systems verbunden, ohne eine mit Switching-Funktionalität oder sonstiger Intelligenz versehene dazwischen befindlichen Software-Implementierung. Auf der physischen Netzwerkkarte wird der VM ein virtueller Ethernet Port (vEth-Port) zur Verfügung gestellt, an dem die Kommunikationsdaten der VM mit einem über die gesamte Virtualisierungsumgebung hinweg eindeutigen Label - dem VN-Tag - versehen werden. Über dieses VN-Tag werden die Daten der VM auf dem nächstliegenden physischen Netzwerkschicht einem Portprofil zugeordnet.

Der gleiche Ansatz soll auch im Rahmen der Unified Computing Architektur von Cisco verfolgt werden. Stand heute sind jedoch die hierfür erforderlichen Netzwerkkarten noch nicht verfügbar.

Auf dem Intel Developer Forum 2009 wurden jedoch bereits erste Server-Hostadapter, die diese Form einer „Single-Root-I/O-Virtualisierung“ unterstützen angekündigt¹ und auch Dell hat erste Implementierungen auf dieser Basis vorgestellt².

5. Fazit

Mehrere Hersteller haben mittlerweile erkannt, dass es von hoher Wichtigkeit für die Zuverlässigkeit und Sicherheit einer virtualisierten Umgebung ist, dass die Durchsetzung von Sicherheits- und QoS-Richtlinien bei der Anbindung von VMs in einer eindeutigen Administrationshoheit liegt und dass diese Aufgabe am besten von der Netzwerkebene erfüllt wird.

Die Nexus 1000v Architektur erweitert nicht nur den Distributed Virtual Switch einer auf Basis von VMware vSphere 4 virtualisierten Umgebung um zusätzliche Leistungsmerkmale, er bietet auch eine Lösung für das Problem der einheitlichen Administration. Dabei bringt der Nexus 1000v jedoch nicht nur eine zusätzliche - und nicht unerhebliche - Komplexität in die virtuelle Infrastruktur, was erst über geeignete Mitarbeiterschulungen kompensiert werden muss. Diese Lösung erzeugt auch zusätzliche Kosten, und zwar nicht nur durch eigene Lizenzen, sondern auch durch die Notwendigkeit der höchsten Lizenzierungsstufe von VMware, deren Leistungsumfang möglicherweise sonst nicht erforderlich gewesen wäre.

Citrix hat bereits angekündigt, ebenfalls einen verteilten virtuellen Switch zu veröffentlichen. Dieser soll neben Xen/XenServer

auch Kernel-based Virtual Machine (KVM)³ unterstützen. Einen Einblick erhält man bereits anhand des Open vSwitch. Dieser soll ebenfalls Funktionen wie ACLs², 802.1X, Netflow, SPAN, RSPAN und ERSPAN anbieten.

Beide Lösungen werden nur eine Zwischenstufe auf dem Weg zu einer virtualisierten Umgebung ohne virtuellen Switch darstellen, da es aufwendig und letztlich unwirtschaftlich ist, die Intelligenz von Netzkomponenten innerhalb der Lösung zur Servervirtualisierung abzudecken. Beide hier vorgestellte Lösungen belegen Ressourcen und führen zu Komplexitäten an Stellen, an denen eigentlich höhere Effizienz und Vereinfachung das Ziel sein sollte.

Nicht anders als schon im Netzwerk Insider 05/2009 zum Thema „Mehr Sicherheit durch virtuelle Firewalls?“ gilt auch für Netzwerkkomponenten, dass man komplexe und leistungshungrige Anforderungen besser da abdeckt, wo sie technisch und personell beherrscht werden: in physischer Hardware.

Abkürzungen

ACL	Access Control List
CEP	Convergence Endpoint
CoS	Class of Service
DoS	Denial of Service
DPM	Distributed Power Management
DRS	Distributed Resource Scheduler
DSCP	Differentiated Services Code Point
ERSPAN	Encapsulated RSPAN
ETH	Ethernet Schnittstelle
HA	High Availability
HBA	Host Bus Adapter
I/O	Input / Output
IP	Internet Protocol
KVM	Kernel-based Virtual Machine
LAN	Local Area Network
lvETH	Local Virtual Ethernet Schnittstelle
MAC	Media Access Control
NAC	Network Access Control
NIC	Network Interface Card, Network Interface Controller
OVF	Open Virtual File Format
pNIC	physical NIC
QoS	Quality of Service
RAM	Random Access Memory
RSPAN	Remote SPAN
RZ	Rechenzentrum
SPAN	Switched Port Analyzer
ToS	Type of Service
VLAN	Virtual LAN
vETH	Virtual Ethernet Schnittstelle
vNIC	virtual NIC
VM	Virtual Machine, Virtuelle Maschine
VMDK	Virtual Machine Disk Format
WLAN	Wireless LAN

¹ KVM gilt noch als Neuling im Bereich Servervirtualisierung, ist jedoch im Unterschied zu Xen bereits seit 2007 im Linux Kernel vertreten und auch bei Red Hat als Nachfolger von Xen gesetzt.

² <http://openvswitch.org>