ComConsult Research

Der Netzwerk Insider

September 2010

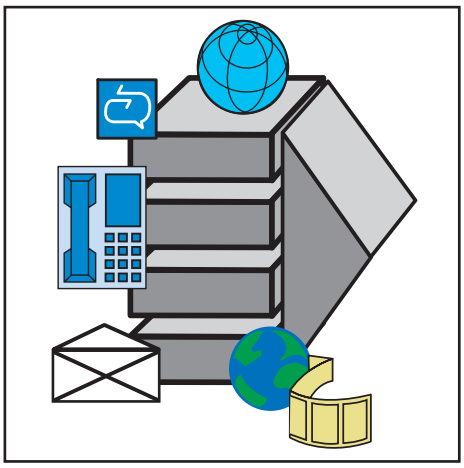
Systematische Weiterbildung für Netzwerk- und IT-Professionals

Schwerpunktthema

Virtualisierung: Unified Communications in virtuellen IT-Umgebungen

von Dipl.-Inform. Daniel Meinhold

Unified Communications und Virtualisierung waren bisher Themen, denen nicht unbedingt eine thematische Nähe zugesagt wurde. Die Anforderungen an eine Echtzeitkommunikation schienen nicht vereinbar zu sein mit einer Konsolidierung von Ressourcen, wie es z.B. die Server-Virtualisierung ermöglicht. Neben den Möglichkeiten und Grenzen einer virtualisierten UC-Umgebung beleuchtet der Artikel die Produkte gängiger TK/UC-Anbieter sowie den Einsatz von UC-Applikationen in Terminal-Server-Umgebungen speziell auf Basis von Citrix.



Virtualisierung und Verfügbarkeit

Die Verfügbarkeitsanforderungen, welche für die physischen Komponenten gelten, müssen auch in einer virtuellen Infrastruktur eingehalten werden. Dies gilt insbesondere im Bereich Telekommunikation. Die Verfügbarkeit kann unter anderem mit folgenden Mechanismen erhöht werden:

weiter auf Seite 16

Zweitthema

Einbindung von Servern in das RZ-Netz

von Dr.-Ing. Behrooz Moayeri

Die meisten Fehler in Rechenzentrumsnetzen sind nach Erfahrungen unserer Troubleshooting Teams auf Mängel in der Koordination zwischen der Konfiguration der Netzkomponenten und der netzbezogenen Konfiguration von Servern zurückzuführen. Dabei können manche solche Fehler Auswirkungen weit über den einzelnen Server haben, an dem die Konfiguration vorgenommen wurde. Ganze Subnetze oder sogar ganze RZ-Netze können von einem solchen Fehler betroffen sein.

Der vorliegende Beitrag ist eine Zusammenfassung der Erfahrungen mit solchen Fehlern und der Erkenntnisse aus den Projekten, in denen es um die abgestimmte Konfiguration der RZ-Netzkomponenten und der Server geht. Generell ist eine engere Koordination der Disziplinen erforderlich, die aus historischen Gründen insbesondere in großen Unternehmen unterschiedlichen Instanzen zugeordnet sind: Kompetenz-Center für Server und für das Netz.

Auf dem Rechenzentrum Infrastruktur-Redesign Forum 2010 der ComConsult Akademie werden aktuelle Fragestellungen im Zusammenhang mit RZ-Netzen von Referenten mit jahrelangen Erfahrungen in diesem Bereich behandelt. Unter anderem wird diskutiert, wie Fehler in Rechenzentren auch durch ein robustes Design vermieden oder zumindest in ihrer Auswirkung eingeschränkt werden können.

weiter auf Seite 26

Top-Kongress

ComConsult Voice-, Video- und Kollaborations-Forum 2010

auf Seite 5

Geleit

UC-System-Architekturen, RZ-Infrastrukturen und Virtualisierung wachsen zusammen

ab Seite 2

Neues Seminar

Service-Spezifizierung

ab Seite 25

Zum Geleit

UC-System-Architekturen, RZ-Infrastrukturen und Virtualisierung wachsen zusammen

Analysiert man die Infrastruktur-Anforderungen moderner Web-Architekturen und von zukünftigen Unified Communications-Lösungen inklusive Sprache und Video, dann stellt man schnell fest, dass diese Anforderungen fast identisch sind. Betrachtet man dazu die Entwicklung auf der Herstellerseite, dann liegt die Prognose nahe, dass UC-Architekturen innerhalb der nächsten fünf Jahre komplett virtualisiert werden.

Anders formuliert:

- Rechenzentrums-Infrastrukturen werden und müssen sich ändern, um dem aktuellen Bedarf von Web-Anwendungen zu entsprechen.
- UC-System-Architekturen werden sich in den nächsten 5 Jahren vollständig verändern, auf der Basis von Virtualisierung werden Software-Architekturen entstehen, die mit der Vergangenheit nichts mehr zu tun haben.
- Beide Welten wachsen zusammen und werden in Zukunft auf gemeinsamen Infrastrukturen aufsetzen.

Wer ist betroffen:

- Die Planer und Betreiber von Rechenzentrums-Infrastrukturen, also Netzwerken, Speichersystemen, Servern, Klimatisierung und Stromversorgung: die hier angesprochenen Änderungen finden im Bereich UC zwar nicht sofort aber im aktuellen Invest-Zyklus statt. Investitionen, die jetzt getätigt werden, müssen diesen Trend berücksichtigen.
- Die Planer und Betreiber von UC-Lösungen. Diese müssen die sehr unterschiedlichen Ansätze der UC-Hersteller analysieren und beobachten. Neue System- und Software-Architekturen sind gefragt. Hier geht es nicht um Feigenblatt-Virtualisierung unwichtiger System-Bestandteile, hier geht es um wirklich neue Architekturen. Einige UC-Hersteller haben ihre Investitionen in Entwicklung stark reduziert, dies ist ein deutliches Indiz dafür, dass diese Hersteller in den nächsten Jahren weiter zurückfallen werden. Hier sollte gnadenlos aussortiert werden, wer als Hersteller das Geld für neue Architekturen nicht hat, der sollte von neuen Projekten ausgeschlossen werden.

Warum sind die Anforderungen von Web-Architekturen und zukünftigen UC-Lösungen so ähnlich? Warum kann heute eine RZ-Infrastruktur so aufgebaut werden, dass sie zukunftsorientiert ist und eine solide Basis für beide Bereiche bietet?

Um diese Frage zu beantworten macht es Sinn, die Kern-Eigenschaften moderner Web-Architekturen zu betrachten:

Parallelität

Web-Architekturen basieren nicht auf einem zentralen Prozess, der umso mehr Last generiert umso mehr Benutzer da sind. Die Grundidee der Web-Architekturen ist, dass sie über die Zahl paralleler virtueller Maschinen skalieren. Je mehr Benutzer eine Web-Applikation benutzen, je mehr virtuelle Maschinen werden gestartet. Die Last einer einzelnen virtuellen Maschine ist überschaubar und kann zum Beispiel einem CPU-Core zugeordnet werden.

Skalierbares Datenbank-Design

Im Rahmen eines parallelen Designs müssen Einzelprozesse vermieden werden, die zuviel Last erzeugen. Im Prinzip muss jeder Prozess parallelisiert werden können. Bei Datenbanken ist das naturgemäß bei dem schreibenden Prozess sehr schwierig und würde eine erhebliche Applikations-Logik erfordern. Ein einfacher Lösungsansatz sieht die Trennung zwischen schreibenden und lesenden Datenbanken mit einer entsprechenden Synchronisation voraus. Lesende Datenbanken können in beliebiger Zahl parallel zueinander existieren.

Prozesse sind virtuelle Maschinen

Jeder Kernprozess wird durch eine virtuelle Maschine abgebildet. Parallele Prozesse entstehen durch parallel ablaufende virtuelle Maschinen, die durch einen Broker gestartet und kontrolliert werden.

Dynamische Lasten und hohe Benutzerzahlen

In dieser Architektur spielt die Zahl der Benutzer keine Rolle. Das System skaliert über die Zahl paralleler virtueller Maschinen. Die UC-Hersteller entwickeln quasi eine Software für alle Anwendungs-Größen.

Ausfallsicherheit

Neben den vielen Möglichkeiten, die Ausfallsicherheit in virtuellen Umgebungen zu steigern, so wie sie auch im Leitartikel besprochen werden, ist natürlich die Erzeugung paralleler virtueller Maschinen ein zentrales Instrument der Ausfallsicherheit. Das passt durchaus zusammen. Man nutzt HA und FT so wie im Artikel beschrieben, um die Prozesse abzusichern, die sich nicht parallelisieren lassen. Dazu zählt zum Beispiel die schreibende Datenbank, die Verbindungs-Information speichert (sofern Status-basierte Protokolle zum Einsatz kommen). Auch der Broker, der parallele Instanzen erzeugt und steuert, gehört dazu.

Schnittstellen und Latenz

In einer Architektur paralleler virtueller Maschinen gibt es zwei zentrale Werte, die die Leistung der Architektur bestimmen. Die Latenz in der Kommunikation zwischen virtuellen Maschinen und die Zeit zum Verlagern einer virtuellen Maschine auf ein anderes System. In der Latenz gehen die Anforderungen in den Bereich von Millisekunden (kleiner 1 ms). Zur schnellen Verlagerung des RAM-Bereichs einer virtuellen Maschine muss der aktive virtuelle Speicher verlagert werden. Je langsamer dies geschieht, desto größer kann die zu verlagernde Datenmenge werden (siehe Betriebssysteme, Virtuelle Speicherkonzepte). 10 Gigabit-Ethernet kann man hier als das notwendige Minimum anse-

Schwerpunktthema

Virtualisierung:

Unified Communications in virtuellen IT-Umgebungen

Fortsetzung von Seite 1

Dipl.-Inform. Daniel Meinhold ist Consultant bei der ComConsult Beratung und Planung GmbH. Dort ist er im Competence Center Data Center tätig und spezialisiert auf die Themen Virtualisierung und IT-Sicherheit.

- Live-Migration (Abbildung 1): Virtuelle Maschinen (VMs) können im laufenden Betrieb von einem Host auf einen anderen verschoben werden. Nur der RAM-Inhalt wird über das Netz übertragen, die virtuelle Festplatte der VM bleibt im Shared Storage.
- HA-Cluster (Abbildung 2): Fällt ein Host aus, werden die darauf laufenden VMs umgehend auf einem anderen Host neu gestartet.
- Fault Tolerance (FT): Hierbei wird eine VM gespiegelt und alle CPU-Instruktionen der primären VM werden auf der sekundären VM ausgeführt. Fällt die primäre VM aus, kann die sekundäre VM deren Aufgaben ohne Unterbrechung übernehmen.
- Cluster auf VM-Ebene (Abbildung 3): Unabhängig vom Host können die Cluster-Funktionen innerhalb der jeweiligen VM genutzt.

Vorteile der UC-Virtualisierung

Während die klassische Telekommunikation auf dedizierter und häufig dezentraler Infrastruktur basiert, ist heute eine IP-basierte und zentrale Kommunikationsplattform gängige Praxis. Mit dem Einzug von IP-PBXen und weiteren UC-Komponenten in die Rechenzentren der Unternehmen stellt sich jedoch die Frage, inwiefern hier eine Integration in die RZ-Infrastruktur erfolgen kann. Server-Virtualisierung ist mittlerweile Usus in den meisten Rechenzentren, so dass die Option zu virtualisieren in Betracht gezogen werden sollte, um die bestehende Infrastruktur und zugehörige Prozesse optimal nutzen zu können. Durch den Wandel in der TK von proprietärer Hardware zu Standard-Komponenten und Software-basierten Systemen ist auch hier eine Server-Virtualisierung prinzipiell möglich. Die Kopplung zu anderen Systemen (z.B. Media Gateways) erfolgt ohnehin oftmals IP-basiert, wenn nicht bereits sogar SIP Trunking eingesetzt wird. Neben der IP-PBX umfasst eine UC-Umgebung weitere Produkte, wie z.B. Voicemail, Instant Messaging, Web und Video Conferencing sowie zugehörige Frontend-, Backend- und Management-Systeme. Aufgrund der Vielzahl an Komponenten ist Unified Communications auf den ersten Blick ein dankbarer Kandidant für eine Virtualisierung. Warum sollte man diese Systeme überhaupt virtualisieren wollen? Dafür sprechen u.a.:

- Einheitliche Infrastruktur: Die im RZ bereitgestellte Infrastruktur sowie zugehörige Prozesse und Erfahrungen können

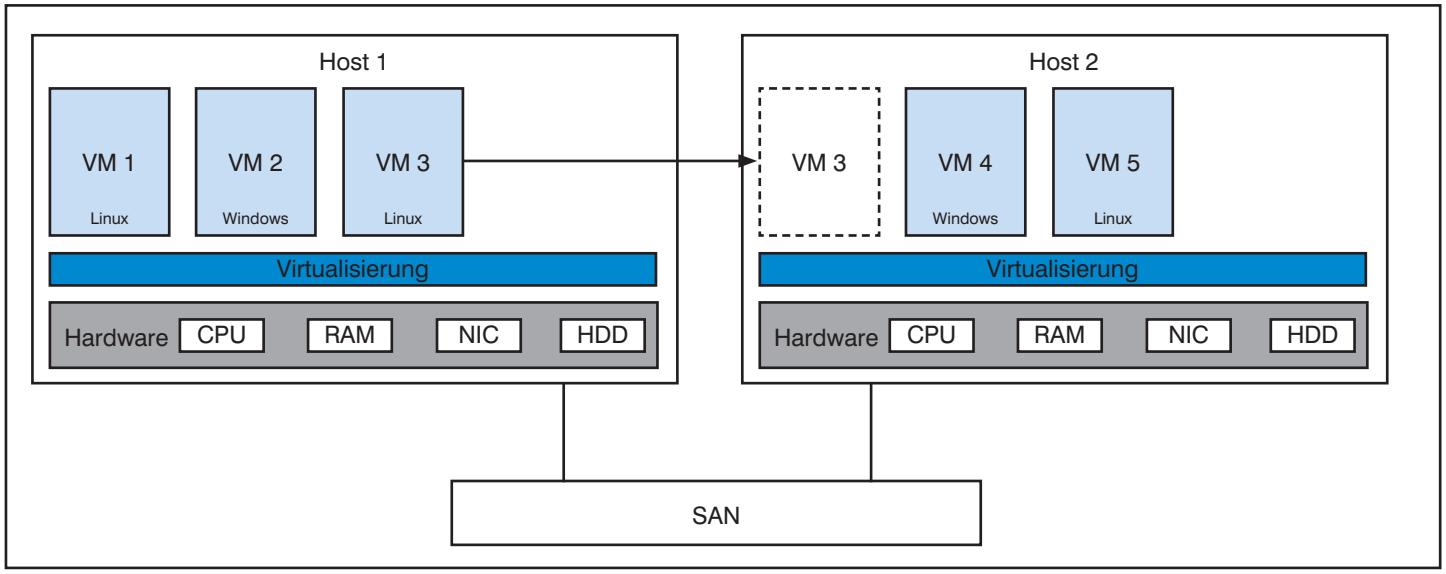


Abbildung 1: Server-Virtualisierung und Live-Migration

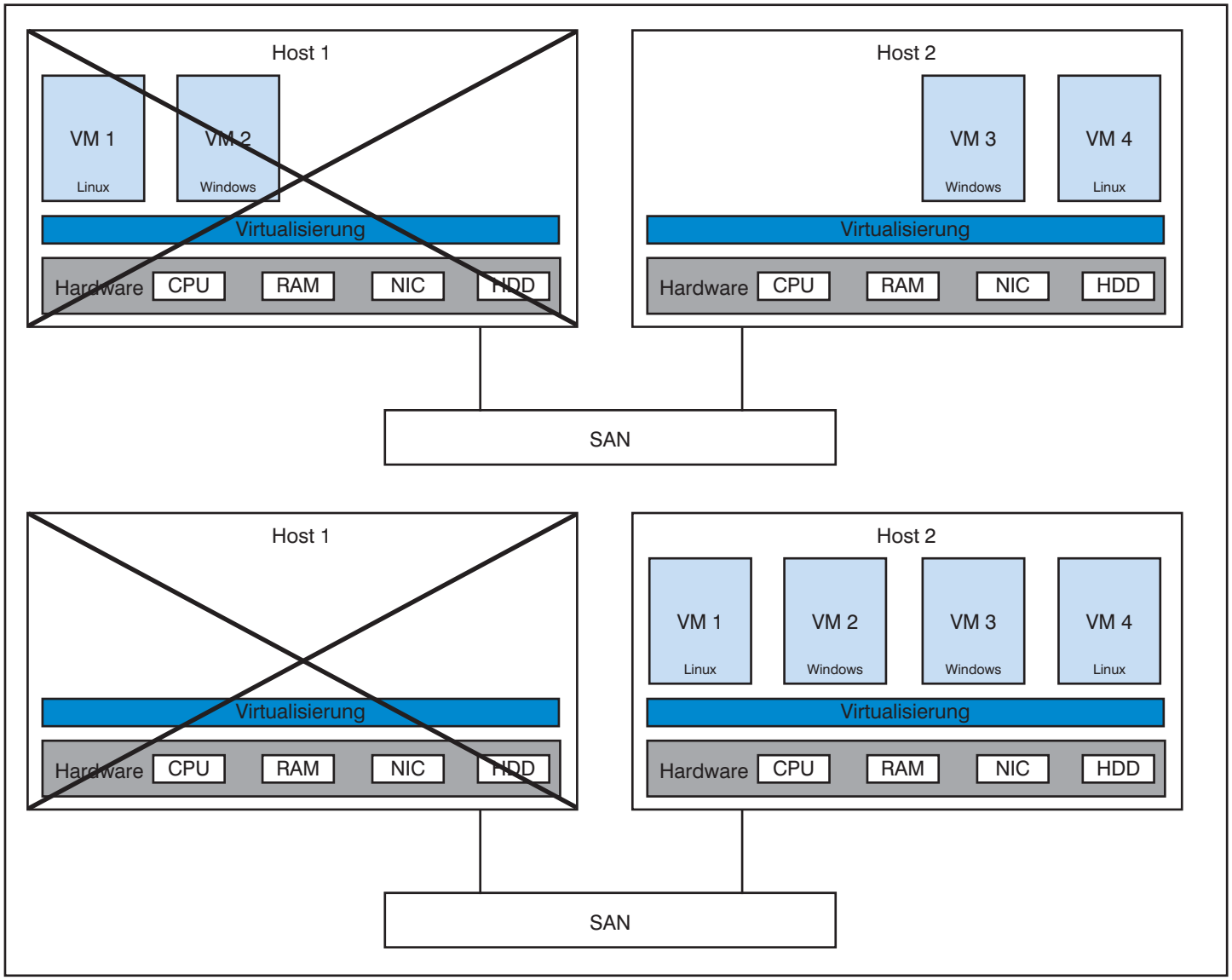


Abbildung 2: HA-Cluster auf Host-Ebene

auch für UC-Systeme genutzt werden. Die Anzahl der „Sonderlocken“ wird minimiert.

- Verfügbarkeit: Während die IP-PBX in der Regel als HA-Cluster realisiert wird, werden andere Systeme wie z.B. Voicemail einfach ausgelegt. Mittels der HA-Funktionen der Virtualisierungsplattform können bei Ausfall der Hardware die VMs automatisch auf einem anderen Host neu gestartet werden. Ist diese Funktion einmalig lizenziert, steht diese allen VMs zur Verfügung. Zudem können VMs, zwecks Wartungsarbeiten an der Hardware des Hosts, mit minimaler Ausfallzeit auf andere Hosts umgezogen werden (Live-Migration). Insgesamt kann die Verfügbarkeit der angebotenen Dienste somit erhöht werden.

- Backup: VMs bestehen meist nur aus einer handvoll Dateien, welche einfach gesichert werden können. Dies ermöglicht ein einfaches Backup und Restore.

- Geringerer Platz- und Energiebedarf: Für kleine und mittlere Unternehmen aber auch für Filialen bietet die Konsolidierung mittels Virtualisierung Vorteile, da die Anzahl der erforderlichen physischen Server deutlich reduziert wird.

- Testumgebung: Mittels Server-Virtualisierung kann auf einfache Art und Weise eine UC-Testumgebung aufgesetzt werden. Zusätzliche Hardware sowie deren Inbetriebnahme entfällt. Im Idealfall entspricht die Testumgebung dabei zugleich der Produktivumgebung.

Server-Virtualisierung und Echtzeitkommunikation

Das entscheidende Kriterium, warum UC-Systeme schwierig zu virtualisieren sind, liegt in der Verarbeitung der Medienströme für Audio und Video. Eine gute Sprach- und Bildqualität wird z.B. bei einer maximalen Verzögerung von ca. 150 ms erzielt. Kleinste Unterbrechungen gelten dabei bereits als störend. In einer nicht geeigneten virtuellen Umgebung äußert sich dies dann durch abgehackte Gespräche oder ruckelige Videodarstellungen. Dennoch ist anzumerken, dass der Medienstrom in der Regel nicht die IP-PBX passiert, sondern direkt zwischen den Endgeräten ausgetauscht wird. Die IP-PBX ist primär für die Signalisierung zuständig, kann jedoch in gewissen Situationen auch

Zweitthema

Einbindung von Servern in das RZ-Netz

Fortsetzung von Seite 1

Dr.-Ing. Behrooz Moayeri hat viele Großprojekte mit dem Schwerpunkt RZ-Redesign geleitet. Er gehört der Geschäftsleitung der ComConsult Beratung und Planung GmbH an und betätigt sich als Berater, Autor und Seminarleiter.

Zunehmende Vielfalt von Servern in Rechenzentren

In den Rechenzentren sind vielfältige Typen von Servern im Einsatz (siehe Abbildung 1).

Nur wenige Rechenzentren werden auf grüner Wiese gebaut und ausschließlich mit neuer Serverhardware ausgestattet. Die meisten Rechenzentren sind historisch gewachsen. Obwohl Serverhardware nach 3-4 Jahren als veraltet gilt, gibt es viele Gründe, auch ältere Serverhardware weiter zu betreiben, vor allem motiviert durch die Vermeidung von Risiken bei der Umstellung einer Anwendung auf eine neue Plattform, wenn die Anwendung die hohe Leistung neuer Hardware nicht unbedingt braucht.

So ist in den meisten RZs eine Vielfalt von Servern mit unterschiedlicher Technologie im Einsatz, die weiter anzubinden sind:

- Ältere Server, die teilweise nachträglich mit Netzadaptern ausgestattet wurden,
- Rack Server mit eingebauten Netzadaptern,
- Blade Enclosures mit diversen Typen von Connectivity-Modulen und
- Virtuelle Server, die über virtuelle Switches (Funktionalität in den Hosts) mit dem Netz verbunden sind.

Unabhängig vom Servertyp haben sich einige Verfahren und Mechanismen wie Adapter Teaming und Lastverteilung in den letzten Jahren bei allen Herstellern ähnlich entwickelt. Wir gehen im Folgenden zunächst auf diese Verfahren ein, bevor besondere Aspekte bei Blade- und virtuellen Servern diskutiert werden.

Adapter Teaming

Mit Adapter Teaming wird erreicht, dass alle Ausfallszenarien bis auf den Ausfall des gesamten Servers durch automatische Redundanzmechanismen abgedeckt werden. Auch bei Einsatz von Serverclustern und hochverfügbaren virtualisierten Umgebungen bevorzugen viele RZ-Betreiber die redundante Anbindung jedes physikalischen Servers an das Netz. Diese Präferenz wird in der Regel wie folgt begründet:

- Server-Cluster-Mechanismen können den redundanten Anschluss jedes einzelnen Servers nicht voll ersetzen. Failover-Mechanismen bei Server-Clustern greifen wesentlich langsamer als Redundanzmechanismen bei Adapter Teaming. Wenn auf Adapter Teaming verzichtet wird, führt der Ausfall der Netzverbindung eines Server-Cluster-Knotens zum Failover von einem zum anderen Server mit deutlich längerer Ausfallzeit als bei einer reinen Adapterumschaltung.
- Server, die nicht per Server Cluster abgesichert werden, können immerhin von Adapter Teaming profitieren.
- Mit Adapter Teaming können die Adapter so eingestellt werden, dass sie regelmäßig Multicasts senden, so dass Pakete an die Adressen der Server nicht geflutet werden.

Das grundsätzliche Netzkonzept für Adapter Teaming geht aus der Abbildung 2 hervor.

Dieses Netzschema ist Grundlage der heute meistens angewandten Verfahren für den redundanten Netzanschluss eines einzelnen Servers. Dieses Konzept sieht vor, dass beide Netzadapter des Servers an das selbe IP-Subnetz (und somit an die selbe Layer-2-Broadcast-Domäne, d.h. auch das selbe virtuelle LAN – VLAN) angeschlossen werden.

Zwischen den beiden Ports, die für den Anschluss des Servers verwendet werden, muss eine Layer-2-Verbindung bestehen. Im dargestellten Netz erfolgt dies mithilfe zweier Layer-2-Switches, die über eine direkte Verbindung miteinander verbunden sind. Beide Layer-2-Switches befinden sich in derselben Layer-2-Broadcast-Domäne, auf die das IP-Subnetz A abgebildet ist. Sehr wohl können an dieses IP-Subnetz verschiedene Layer-3-Switches angeschlossen werden, wie in der Abbildung 2 auch dargestellt ist.

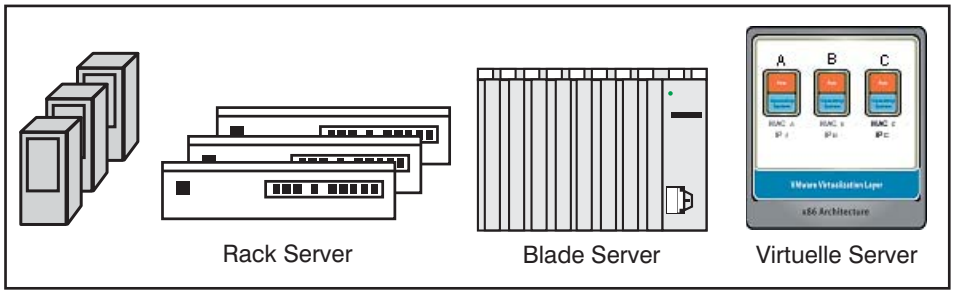


Abbildung 1: Vielfalt der Servertypen

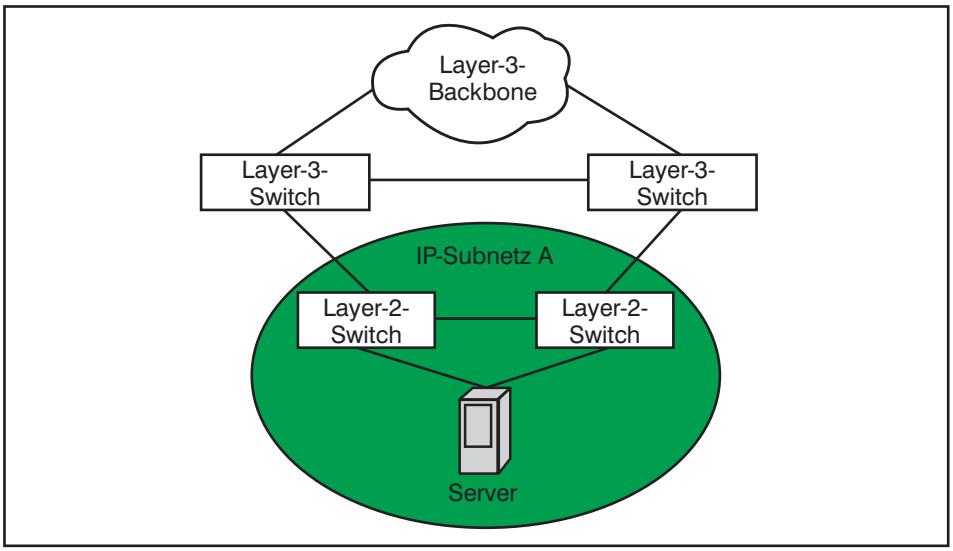


Abbildung 2: Redundanter Serveranschluss in einem IP-Subnetz

Beispiel HP Network Fault Tolerance

Exemplarisch für alle anderen ähnlichen Mechanismen sei hier Network Fault Tolerance (NFT) von HP erwähnt. Das Verfahren ist in der Abbildung 3 dargestellt.

Im Normalfall ist die Network Interface Card (NIC) mit der Nummer 1 der aktive Netz- adapter des Servers. Fällt der Layer-2-Switch, an den dieser Adapter angeschlossen ist, oder die Verbindung des Adapters zu diesem Switch aus, wird dieser Ausfall vom Teaming-Treiber bemerkt, der den zweiten physikalischen Adapter (NIC 2) aktiviert. Eine Layer-2-Verbindung zwischen den beiden Layer-2-Switches ist notwendig. Beide Adapter befinden sich in derselben Broadcast-Domäne bzw. VLAN.

Bei NFT und anderen ähnlichen Teaming-Modi ist immer nur eine Karte aktiv (Primary Adapter), und die anderen Karten sind standby. Eine Lastverteilung findet nicht statt. Alle Adapter im Team verwenden dieselbe IP-Adresse. Die Switches bemerken nichts von NFT und bedürfen keiner speziellen Einstellung.

Eine Erweiterung von NFT ist der Mechanismus „NFT mit Preference Order“. Hierbei wird der Primäre Adapter administrativ gesetzt. Ohne eine solche Reihenfolge bleibt zum Beispiel nach einem Fail-over der Adapter, der die Funktion des aktiven NIC übernommen hat, weiterhin aktiv, während bei Aktivierung der Preference Order die Wiederverfügbarkeit des administrativ gesetzten Adapters eine Umschaltung zurück zu diesem Adapter verursacht.

Transmit Load Balancing

HP Transmit Load Balancing (TLB) bietet eine Lastverteilung beim Senden vom Server. Alle Adapter in einem Team melden sich mit derselben IP-Adresse im Netz. Alle Adapter, primäre und standby-NICs, senden, aber nur der primäre Adapter empfängt. Es ist keine spezielle Konfiguration an den Switches notwendig. Auch hier müssen sich alle Netzkarten in der gleichen Broadcast-Domäne bzw. im gleichen VLAN befinden.

Switch-Assisted Load Balancing (SLB) bietet eine Lastverteilung in Sende- und Empfangsrichtung. Am Switch muss eine Link Aggregation Group (LAG) konfiguriert oder mit dynamischen Mechanismen zwischen Switch und Adaptern ausgehandelt werden.

Zusammenfassend kann exemplarisch anhand der HP-Namenskonventionen zwischen folgenden Teaming-Modi unterschieden werden:

- Network Fault Tolerance (NFT):
  - Ein Adpater ist der primäre Adapter.
  - Alle anderen Adapter im Team sind standby.
  - Der Anschluss der Adapter an verschiedene Switches ist möglich.
- NFT mit Preference Order: entspricht NFT, wobei der primäre Adapter administrativ gesetzt ist
- Transmit Load Balancing (TLB):
  - Der primäre Adapter sendet und empfängt.
  - Andere Adapter senden nur.
  - Anschluss der Adapter an verschiedene Switches ist möglich.
- TLB mit Preference Order: entspricht TLB, wobei der primäre Adapter administrativ gesetzt ist.
- Switch-Assisted Load Balancing (SLB)
  - Link Aggregation zwischen Server und Switch wird genutzt.
  - Der Anschluss der Adapter erfolgt an denselben Switch (oder einen Switch-Cluster, der Multi-Chassis Link Aggregation unterstützt).

Basic und Advanced Teaming

Bei HP wird zwischen Basis- und Zusatz-Funktionen von Adapter Teaming unterschieden. Unter die Basisfunktionen fallen die folgenden:

- NFT
- NFT with Preference
- TLB
- TLB with Preference
- SLB

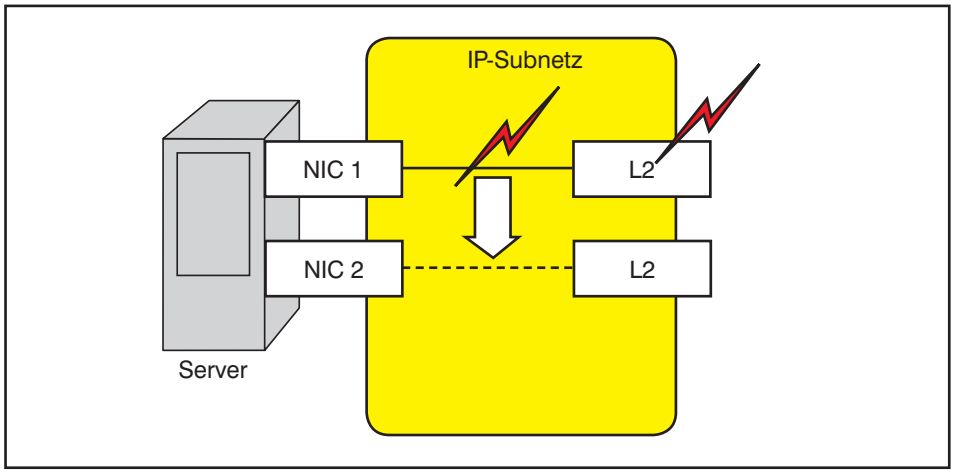


Abbildung 3: Network Fault Tolerance (NFT) von HP