ComConsult Research

Der Netzwerk Insider

Dezember 2010

Systematische Weiterbildung für Netzwerk- und IT-Professionals

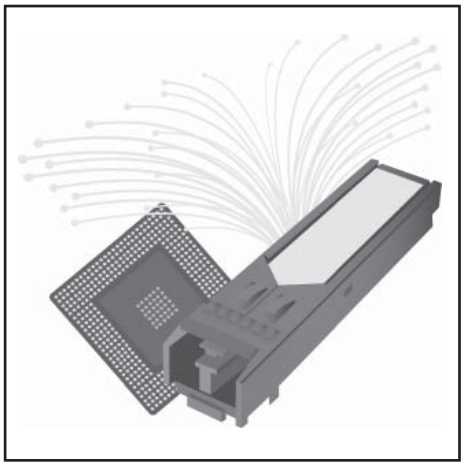
Schwerpunktthema

Low Latency Networks: Nutzungsorientierte Aspekte und Einbindung Virtueller Systeme

von Dr. Franz-Joachim Kauffels

In der November-Ausgabe des Netzwerk Insiders haben wir die unterschiedlichen Architekturen, die zur Schaffung latenzarmer Netze hinreichender Skalierbarkeit benutzt werden können, kennengelernt. In diesem Teil vertiefen wir das noch in ein paar Richtungen.

Zunächst sehen wir uns Möglichkeiten, Wirkungen und Konstruktionsformen latenzarmer Messaging-Mechanismen, die eine hervorragende Basis für die Implementierung leistungsfähiger Inter-Prozess-Kommunikation darstellen, in Theorie, Praxis und Beispiel an. Wie steht es eigentlich um die Einbindung latenzarmer Netze in die I/O von VMs im Rahmen



von Virtualisierungssoftware? Hier gibt es neue Techniken, die allerdings einen erheblichen Erklärungsbedarf nach sich ziehen. Wie kann man solche Umgebungen sicher und sinnvoll betreiben? Fabric as a Service (FaaS®) ist hier ein Stichwort. Am Ende erfahren Sie noch, dass es ein Kriterium gibt, welches RZs ohne konkrete Anforderungen aus den Anwendungen in die Richtung der Verwendung latenzarmer Netze zwingt: die Einführung masseloser Speicher: SSDs.

weiter auf Seite 9

Zweitthema

Die Service-Spezifikation - das A & O der Service-Erbringung

von Dipl.-Ing. Paul G. Huppertz

Zum Thema IT Service Management wird ITIL (IT Infrastructure Library) regelmäßig als das maßgebliche Prozess-Framework angeführt.

Dabei fehlen in ITIL wesentliche Elemente für die Erbringung von geschäftsrelevanten ICT-basierten Business Support Services (ICTBSS), vor allem

- die konsequente Ausrichtung auf die Service-Konsumenten,
- eine einheitliche Methode für die Service-Spezifizierung mitsamt einer Vorlage für Service-Spezifikationen,
- ein vollständiges Service-Erbringungsmodell mit allen beteiligten Hauptrollen,
- eine ausgearbeitete Vorlage für den Service-Erbringungsprozess.

weiter auf Seite 22

Frühbucherphase

IPv6-Forum 2011

Netzwerk-Redesign Forum 2011

ab Seite 6

Geleit

Microsoft lässt TK-Hersteller und Cisco alt aussehen, aber reicht das aus?

ab Seite 2

Sonderveranstaltung

Unified Communications - Cisco versus Microsoft

auf Seite 5

Zum Geleit

Microsoft lässt TK-Hersteller und Cisco alt aussehen, aber reicht das aus?

Microsoft hat im Rahmen des ComConsult Voice-, Video- und Kollaborations-Forums 2010 den neuen Lync-Server (früher OCS) vorgestellt und sich im Vorfeld des Forums auch der Evaluierung durch ComConsult Research und UBN/Frau Borowka gestellt. Vereinfacht formuliert ist der Lync-Server das erste ernst zu nehmende Produkt Microsofts mit Anspruch auf Vollständigkeit im Bereich Unified Communications. Die vorhergehenden Versionen, die noch unter dem Namen OCS liefen, waren aus heutiger Sicht „Work in Progress“ und Antesten der Reaktionen im Markt. Der Lync-Server greift viele der Forderungen von Benutzern und auch viele der Kritikpunkte von ComConsult Research am OCS auf und realisiert zum ersten Mal eine runde Lösung.

Mein persönliches Statement gleich zum Start dieses Geleitwortes: die traditionellen TK-Hersteller und auch Cisco sehen gegenüber einigen Funktionen des Lync-Servers alt aus. Offensichtlich hat man sich hier zu lange ausgeruht und versäumt, wirkliche Innovationen auf den Weg zu bringen.

Woran liegt das oder wie kommt dieses Statement zustande?

Der Dreh- und Angelpunkt ist die Erkenntnis, dass Unified Communications nicht gleich TK plus ein paar neue Funktionen ist. Diese Erkenntnis ist wichtig, denn sie unterstreicht, dass die Anreicherung bestehender TK-Lösungen um Video und Webkonferenzen nicht ausreicht, um eine überzeugende UC-Lösung zu bekommen.

Am deutlichsten wird das beim Thema Endgerät. Traditionelle TK-Lösungen drehen sich noch immer um das Telefon. Das kann bei UC grundsätzlich nicht funktionieren. Wer mit dieser Aussage nicht überein stimmt, der hat UC nicht verstanden! Viele Funktionsbereiche von UC lassen sich an einem traditionellen Telefon nicht bedienen. Jeder Versuch einer sinnvollen Integration wird hier scheitern. Dies bedeutet in Konsequenz, dass im Mittelpunkt jeder UC-Lösung der Software-Client auf dem PC stehen muss. Das Telefon wird zu einem ergänzenden Gerät. Noch einmal: wenn Sie damit nicht übereinstimmen, dann sind Sie nicht auf der Suche nach einer UC-Lösung.

Die große Herausforderung von UC ist die Menge von Kommunikations-Funktionen,

die nun an jeden Arbeitsplatz gebracht werden. Ein UC-Projekt kann nur dann erfolgreich sein, wenn jeder Arbeitsplatz, die Funktionen, die für ihn sinnvoll sind, auch benutzt. Im Prinzip kann eine einfache Checkliste erarbeitet werden. Zählen Sie alle Funktionsbereiche von UC auf und ermitteln Sie, wie viele davon an wie vielen Arbeitsplätzen auch tatsächlich genutzt werden. Ist die Nutzung gering, dann ist das Projekt gescheitert. So wird in Verbindung mit dem enormen Funktionsumfang auch die Herausforderung klar. Nur wenn der Benutzer alle für ihn relevanten Funktionen intuitiv und ohne Schulung direkt benutzen kann, wird er diese Funktionen auch nutzen.

Damit kommen wir direkt zum
2. Statement dieses Geleitworts

Der Erfolg eines UC-Projekts entscheidet sich am UC-Client. Gleiches gilt für die Qualität des entsprechenden Produktes. Ein Produkt kann auf dem Papier noch so gut aussehen und zum Beispiel alle Kriterien im Bereich UC-Funktionalität erfüllen, wenn dies nicht mit einem intuitiv nutzbaren Client kombiniert wird, ist das Produkt wertlos.

Ohne Frage liegt hier die Stärke des Microsoft-Ansatzes. Die Entwicklung eines Clients, der von der Grundidee von jedem bedient werden kann, der auch Office und Exchange bedienen kann und die Vermaschung der Produkte miteinander reduziert die Lernkurve für den Anwender deutlich. Allerdings entstehen auch gegebenenfalls ungewollte und teure Zwänge in der Release-Abhängigkeit dieser Software-Pakete.

Und damit kommen wir zum Problem der etablierten TK-Lösungen und auch Ciscos. Diese Lösungen sind TK-zentrisch entstanden. UC-Funktionalität wurde später angeflanscht. Das macht die Produkte automatisch komplex. Und leider muss man feststellen, dass die notwendige Entwicklung eines geeigneten Clients dabei häufig auf der Strecke blieb.

Hier setzt Microsoft mit dem Lync-Server an. Der Lync-Soft-Client kommt dem Idealzustand eines UC-Clients sehr nah. Hinzu kommt, dass Microsoft das gesamte Produkt Benutzer-zentrisch ausgelegt hat. Dies ermöglicht die nahtlose Integration in Active Directory und die Integration mit Office, Sharepoint und Exchange. In allen diesen Produkten liegt ein identisches Verständnis des Begriffs Benutzer vor. Zu jedem Benutzer gehört eine Kontaktkarte, die alle Kommunikations-Funktionen dieses Benutzers zusammenfasst. Diese Karte und ihre Nutzung, sprich UC-Kommunikation, ist in allen angesprochenen Applikationen identisch.

Mit der Leichtigkeit der Bedienung und der Integration in die üblichen Office-Anwendungen setzt Microsoft Maßstäbe.

Haben die anderen Anbieter also jetzt quasi verloren und wird Microsoft diesen Markt übernehmen?

Fangen wir bei den TK-Herstellern und auch Cisco an. Aus meiner Sicht wurde hier ein wichtiger Trend verschlafen. Zwar reden viele der TK-Hersteller seit Jahren über neue Clients, aber passiert ist nicht viel. Siemens hat als einer der ersten im Rahmen von OpenScape damit begonnen, aber die Entwicklung wurde nicht konsequent zu Ende geführt und nicht auf Bedienerfreundlichkeit optimiert. Zusätzlich erschwert die notwendige Integration externer Produkte (Video und Webkonferenz) eine einheitliche Bedienung. Auch die vergleichsweise guten Clients wie der von Cisco sind in der Bedienung zu umständlich, wenn es zum nahtlosen Übergang zwischen Diensten im Rahmen einer laufenden Kommunikation kommt. Cisco hat aber im Vergleich zu den anderen TK-Anbietern den Vorteil, keine Software anderer Anbieter integrieren zu müssen (wobei die vielen zugekauften Teilprodukte inklusive Webex und Tandberg sicher genug Integrations-Probleme generieren, man kann bei Cisco nicht wirklich von einem monolithischen UC-Produkt sprechen).

Schwerpunkthema

Low Latency Networks: Nutzungsorientierte Aspekte und Einbindung Virtueller Systeme

Dr. Franz-Joachim Kauffels ist einer der erfahrensten und bekanntesten Referenten der gesamten Netzwerkszene (über 20 Fachbücher und unzählige Artikel) und bekannt für lebendige und mitreißende Seminare.

Fortsetzung von Seite 1

1. Wirkung von Low Latency-Komponenten auf die Kommunikation Virtueller Maschinen

Nach den ganzen Ausführungen über ULL-Komponenten und den Aufbau entsprechend latenzarmer Netze wollen wir uns der Frage zuwenden, welchen Nutzen das denn insgesamt alles hat.

Es wäre schade, dies auf die Unterstützung allgemein bekannter Anwendungen wie Finanztransaktionen, Wettermodelle, umfangreiche Simulationen und weitere aus dem Umfeld des HPC zu beschränken. Das kann der Interessent im Internet selbst nachlesen.

Es gibt nämlich einen viel nahe liegenderen Bereich, der auch wesentlich mehr Rechenzentren betrifft und in Zukunft in weit höherem Maße beschäftigen wird als bisher: die Kommunikation Virtueller Maschinen.

Blickt man auf die Entwicklungen der letzten Jahre zurück, muss man leider sagen, dass der Beitrag der Netzwerkwelt dazu irgendwo zwischen verzweifeltem Basteln und hochgradigem Unsinn anzusiedeln ist.

Letzterer ist die Verwendung von Virtuellen Switches wie dem Nexus 1000 V von Cisco. Ein Virtueller Switch wird den Hypervisor immer so stark belasten, dass keine Performance zu erwarten ist. Die Hardware-Unterstützung der I/O von VMs durch SR/IOV geht absolut in die richtige Richtung, aber systemtechnisch fehlte lange noch das geeignete Bindeglied. So ist denn das Aufkommen von Bemühungen wie VEPA oder spezieller Treiberlösungen einzelner Hersteller zu erklären. Während letztere als proprietäre Lösung definitiv grundsätzlich zu verwerfen sind, ist VEPA wenigstens ein Standard, den man einsetzen kann, solange Alternativen fehlen.

Die erste wirkliche sinnvolle Kombination ist SR-IOV mit DirectPath, wie es ab VMware 4.1 angeboten wird.

Die Hersteller von Virtualisierungssoftware wie VMware oder Citrix warten jedoch eigentlich auf etwas anderes: schnelle Inter Prozess Kommunikation. IPC ist in Betriebssystemen schon lange ein wesentliches grundlegendes Konzept, leider aber meist mangelnder Netzwerkunterstützung auf eine einzige Maschine festgelegt.

VMs sollen aber flexibel zugeordnet werden können und auch wandern dürfen.

2. Low Latency Messaging Systeme

Völlig unabhängig davon, ob Anwendungen virtualisiert sind oder nicht, kann es notwendig sein, dass sie ein zuverlässiges Messaging-System für die IPC benutzen können. Viel Öffentlichkeit haben in diesem Zusammenhang Anwendungen für das Finanzwesen bekommen, aber es gibt auch andere.

Geringe Latenz ist eine verschärfte Anforderung an derartige Lösungen. Insgesamt müssen sie Folgendes können:

- Hohe Leistung mit extrem geringer Latenz im Sub-Millisekundenbereich
- Unterstützung einer extrem hohen Nachrichtenrate von mehreren Millionen Nachrichten pro Sekunde
- Flexibilität beim Ausliefern von Nachrichten: PtP Unicast Messaging, PtMP sowie MPtoMP Broadcasting
- Zuverlässige Auslieferung von Nachrichten mit granulierter Kontrolle der tatsächlichen Message-Auslieferung
- Hohe Verfügbarkeit zur Bereitstellung von System Service Leveln und zum Schutz der Integrität des Nachrichtenstroms beim Ausfall von Komponenten
- Schutz der Integrität der Nachrichten in Übertragungsgeschwindigkeit für das Wiederherstellen von Nachrichten und das Auditing, Realisierung üblicherweise durch eine geeignete Verschlüsselungstechnik
- Monitoring und Congestion Control, um mögliche Engpässe frühzeitig erkennen zu können
- Filterungsmöglichkeiten für Messages in hoher Geschwindigkeit, um Multiplexing und effiziente Segmentierung von Daten zu ermöglichen

Viele Leser, die sich noch nicht näher mit RDMA befasst haben, sind sicher der Ansicht, dass es sich dabei lediglich um eine weitere Art einfacher Schnittstelle handelt. Vom Prinzip her ist das korrekt, die Frage ist aber, was man daraus machen kann. Üblicherweise setzt man auf RDMA einen Messaging-Service. Das ist ein anwendungsorientierter Kommunikationsmechanismus, der streng genommen in die Schicht 7 des ISO-OSI-Modells gehört, genau wie File Transfer. Ein guter Messaging Service hat daher eine Reihe von Funktionen, die weit über das hinausgehen, was z.B. ein TCP/IP-Stack kann. Letzterer kann allerdings von einem Messaging Service auch mitbenutzt werden. Es gibt natürlich bereits viele Messaging Services, die Frage ist aber, wie sie es mit der Latenz halten.

Statt aller Theorie sehen wir uns ein konkretes Produkt mit großem Leistungsumfang an.

2.1 IBM WebSphere MQ Low Latency Messaging®

Das Produkt wurde entwickelt, um den erhöhten Anforderungen in verschiedenen Bereichen, besonders bei Finanztransaktionen, Rechnung zu tragen. Als Peer-to-Peer-Lösung unterstützt es die Verbindungsformen PtP, PtMP und MPtMP. Außerdem wird die IP Multicast-Infrastruktur letztlich um Realzeitfähigkeiten erweitert. Das bedeutet, dass MQLLM nicht nur auf unterschiedlichen Arten von LANs (GbE, IB), sondern auch im Weitverkehrsbereich eingesetzt werden kann, wobei natürlich dort die Latenz maßgeblich durch die Signalverzögerung bestimmt wird. Die so genannte Open Fabrics Enterprise Distribution liefert eine Anzahl von Librarys für die Implementierung von RDMA auf standardbasierten Industriekomponenten für 1GbE, 10 GbE und IB. Neben hoher Flexibilität wurde bei der Entwicklung Wert darauf gelegt, alles mit einer möglichst geringen CPU-Belastung zu realisieren.

MQLLM kann alleine benutzt werden oder in eine weitere Architektur, wie z.B. die IBM WebSphere DataPower® Appliance eingebettet werden. DataPower® ist im Grunde genommen eine unternehmensweite Service-Busarchitektur, also eine erweiterte Form des „Netz als Systembus“ Gedankens. Das ist insgesamt sehr interessant, wir konzentrieren uns hier aber weiter auf MQLLM.

MQLLM benutzt eine proprietäre Batching-Technologie und ein kompaktes Header-Format, um die anfallenden Daten schnell, zuverlässig und mit geringem Aufwand in einen Message-Strom zu überführen. Leistungstests für IB haben einen möglichen Durchsatz von bis zu 91 Millionen Messages pro Sekunde bei einer durchschnittlichen Latenz von unter 5 μs gezeigt. WebSphere MQLLM ® kann Shared Memory als Transportsystem benutzen. Das nutzt die Möglichkeiten der neueren Multi-Core Prozessor Chips und die Strategien zur so genannten. „Co-Location“ aus. Diese Strategien beziehen sich zusammenfassend ausgedrückt auf die Vermeidung unnötig langer Wege zwischen Komponenten. Tests haben bewiesen, dass man auf diese Weise die Latenz sogar unter 1 μs senken kann.

Nicht alle Anwendungen werden immer die allerhöchsten Anforderungen haben. Neben dem PtP-RDMA bietet MQLLM zunächst einen schnellen Multicast-Transport mit UDP und Receiver Feedback. Neben dem zuverlässigen Multicast-Mechanismus gibt es noch einen weiteren lightweight-Mechanismus für PtP. Dieser basiert ebenfalls auf UDP mit Feedback und Möglichkeiten der Verkehrskontrolle. Um das Design der Messaging-Infrastruktur zu vereinfachen, können mehrere Unicast- oder Multicast-Destinationen zu einem so genannten „Transmitter Topic“ zusammengefasst werden. Schließlich gibt es noch einen zuverlässigen PtP-Unicast-Kommunikationsmechanismus über TCP/IP, wobei Zuverlässigkeit und Verkehrskontrolle hauptsächlich von TCP gewährleistet werden. So bekommt man einen Datenstrom auch zuverlässig durch ein WAN oder durch Firewalls. Die Abbildungsmöglichkeiten der Datenströme auf die TCP-Kommunikation sind vielfältig. Allerdings sind die TCP-Verbindungen „schwergewichtige“ Objekte, so dass man sie nur sparsam einsetzen sollte.

Es gibt in MQLLM verschiedene Methoden zur Gewährleistung der Zuverlässigkeit, die wir hier nicht weiter besprechen wollen. Zwischen einem „Fire-and-Forget-Ansatz“ und einer Zuverlässigkeit, die über die Möglichkeiten eines normalen TCP hinausgeht, gibt es ein hinreichendes Spektrum. Insgesamt ist es auch ohne TCP in der Lage, verlorengegangene Pakete oder Pakete, die sich der Reihenfolge entzogen haben, erneut zu beschaffen bzw. richtig einzusortieren.

Eine fundamentale Grenze für die Zuverlässigkeit eines Messaging Systems ist die Größe des „History Buffers“, der dazu benötigt wird, Pakete, die vom Empfänger vermisst wurden, wieder erneut auszuliefern. Langsame Messaging Systeme können den Puffer durch Swappen auf eine Festplatte vergrößern. Bei Low Latency ist das aber nicht möglich, da die Verarbeitungsgeschwindigkeit der Nachrichtenübertragungsgeschwindigkeit entsprechen muss. Der Speicher kann auch für das Wiederaufsetzen kleinerer Prozeduren, die abgestürzt sind, verwendet werden. Der Message Store bei WebSphere MQLLM ® entspricht prinzipiell dem, was in Kapitel 2 über die Speicher, die in einem Switch-ASIC verwendet werden, gesagt wurde. Der Speicher lässt sich in vielfacher Weise für Wiederherstellungs-, Kontroll- und Multiplexing-Aufgaben verwenden.

Die Zuverlässigkeit des Messaging-Dienstes kann bei bestimmten Anwendungen oder Unternehmen Bestandteil des Auditings sein. Der WebSphere MQLLM ® Message Store unterstützt eine Anzahl von Hochverfügbarkeitskonfigurationen zum Schutz gegenüber Komponenten- und Anwendungsfehlern, um definierte Anforderungen hinsichtlich der Zuverlässigkeit nachweisbar durchzusetzen. Viele Anwendungen bekommen Probleme, wenn das Volumen der zu bewältigenden Messages in einem kurzen Zeitraum zu hoch wird. Congestion Facilities in WebSphere MQLLM® helfen dabei, die Infrastuktur auch dann zuverlässig weiter laufen zu lassen, wenn die Anwendungen überbeschäftigt sind. Sowohl Multicast- als auch Unicast-Transporte beinhalten Methoden, um den Verkehr zu beobachten (Übertragungsrate, Verluste, Retransmissionen, Latenz), die Anwendungen von bevorstehenden Problemen zu informieren und die Probleme durch Umgehung langsamer Empfänger oder das Herabregeln von Sendern zu umgehen.

Normalerweise ist die Bestimmung der Latenz von Ende-zu-Ende-Verbindungen problematisch, weil die Maschinen nicht über eine einheitliche Synchronisation verfügen. WebSphere MQLLM® hat aber ab Version 2.3 eine nachrichtenbasierte Synchronisationstechnologie. So kann man innerhalb der Reichweite des Produkts präzise Messungen vornehmen. Es gibt noch eine Reihe weiterer Funktionen. Normalerweise arbeitet MQLLM Übertragungsanfragen der Reihe nach so schnell wie möglich ab. Man kann aber auch Regeln eingeben, die letztlich einer Priorisierung entsprechen. Genauso gut ist es möglich, bestimmte Grenzen anzugeben, wie z.B. der maximale Anteil einer Nachricht am MessageStore, ACK/NACK-Limits und weitere Schwellwerte, dynamische Akkumulationstechnologie mit Label-Switching und Filter für alles Mögliche.

Ein Nachrichtenstrom wird dazu in mehreren Schichten bearbeitet, bevor er die Anwendung erreicht. WebSphere MQLLM ® unterstützt einen weit reichenden Filtermechanismus, der dem in Java Message Service (JMS) nicht unähnlich ist.

Stream Failover und weitere Hochverfügbarkeitsmechanismen runden das Funktionsspektrum ab, siehe Abbildung 1.

Ein weiteres Element in diesem Zusammenhang ist RCMS (Reliable and Consistent Message Streaming), welches die gesamten Möglichkeiten angewandt auf Applikationsbeziehungen systematisch nutzt, siehe Abbildung 2.

Die Lösung ist für eine große Anzahl von Plattformen verfügbar, darunter MS Windows, Linux, zLinux, HP-UX Itanium und verschiedene Sun Plattformen mit Solaris und x86.

Wie schon erwähnt, läuft IBM WebSphere MQLLM ® prima auf IB mit RDMA. Wir wollen aber hier wissen, was passiert, wenn man es mit RoCE über ein latenzarmes Ethernet implementiert.

Zweitthema

Die Service-Spezifikation - das A & O der Service-Erbringung

Fortsetzung von Seite 1

Dipl.-Ing. Paul G. Huppertz hat in über 20 Jahren als Netzwerkplaner, Systemdesigner, RZ-Leiter, Projektleiter und Consultant in unterschiedlichen Aufgabenbereichen und internat. Großprojekten reichhaltige Erfahrungen gesammelt. Seit 1995 nutzt er ITIL, um Aufgabenstellungen im IT Service Management zu lösen. In seinem Pocket Book „IT-Service – Der Kern des Ganzen“ klärt er den Service-Begriff in einer klaren und kompakten Definition. Darauf basierend hat er die eindeutige und konsistente Service-Spezifikation mit den 12 Standard-Service-Attributen und ein durchgängiges Service-Erbringungskonzept entwickelt. Als Gründer und Inhaber der Firma servicEvolution unterstützt er IT-Dienstleister bei der verlässlichen und rentablen Service-Erbringung.

Diese Elemente werden durch den Ansatz der Servicialisierung ergänzt, der auf verlässliche, rationelle und rentable Service-Erbringung sowie auf die Service-Konsumenten ausgerichtet ist, d.h. auf die Mitarbeiter in den Fachabteilungen. Sie rufen bei Bedarf ICT-basierte Business Support Services (ICTBSS) ab, um ihre aktuell anstehende geschäftliche Aktivität auszuführen und auf diese Weise zur Wertschöpfung ihrer Fachabteilung beizutragen. Ein Mitarbeiter ruft z.B. über das R/3-basierte Buchhaltungssystem die Erbringung eines Fakturierungs-Service ab, um die Rechnung für einen erledigten Auftrag erstellen und versenden zu lassen. In diesem Zusammenhang agiert er aus der Sicht des ICT Service Providers als Service-Konsument.

Um die Mitarbeiter in ihrer Rolle als Service-Konsumenten durch verlässliche Service-Erbringung effizient zu unterstützen, muss der ICT Service Provider die Anforderungen an die Qualität der geschäftsrelevanten ICTBSS angemessen erfassen und so beschreiben, dass die Service-Konsumenten die entsprechende Service-Spezifikation verstehen und bestätigen können. Die mit den Service-Konsumenten abgestimmte und verabschiedete Spezifikation der erforderlichen ICTBSS ist die Grundlage für die Service-Kommissionierung durch den jeweiligen Service-Kunden, z.B. durch den Leiter der Fachabteilung. Dieser schließt mit dem unternehmensinternen ICT Service Provider ein Service Level Agreement (SLA) ab, das als Kernbestandteil die Service-Spezifikation des kommissionierten ICTBSS enthält sowie die Liste derjenigen Mitarbeiter, die berechtigt sind, derartige ICTBSS abzurufen (= berechtigte Service-Konsumenten). Das SLA ist die beiderseits verbindliche Vereinbarung für die Erbringung der ICTBSS an die berechtigten Service-Konsumenten und für die Abrechnung der erbrachten ICTBSS. Der Service Provider ist gegenüber dem auftraggebenden Service-Kunden rechenschaftspflichtig für die Erbringung aller abgerufenen ICTBSS an die berechtigten Service-Konsumenten und der Service-Kunde verpflichtet sich dazu, die spezifikationsgemäß erbrachten Service-Volumina zum vereinbarten Service-Tarif zu bezahlen bzw. intern verrechnen zu lassen.

Die Service-Spezifikation ist also die wesentliche Grundlage für die verbindliche Regelung der Beziehung zwischen dem Service-Kunden mit seinen berechtigten Service-Konsumenten auf der einen Seite und dem unternehmensinternen rechenschaftspflichtigen ICT Service Provider auf der anderen Seite. Daraus ergeben sich die folgenden Anforderungen an die Service-Spezifikation:

- sie muss in den Begriffen der berechtigten Service-Konsumenten und der auftraggebenden Service-Kunden aus den Fachabteilungen abgefasst sein
- sie muss klar und kompakt formuliert sein, so dass sie leicht lesbar und gut verständlich ist
- sie muss den jeweiligen ICTBSS im Service-Katalog prägnant und nachvollziehbar darstellen
- sie muss mehrere verschiedene Service Levels in der gleichen Form abbilden können, so dass der gleiche ICTBSS in verschiedenen Qualitätsstufen kommissioniert werden kann
- sie muss die Service-Qualität jeweils eindeutig, vollständig und konsistent beschreiben sowie eine eindeutige Bezugsgröße enthalten, so dass die Service-Erbringung objektiv geprüft werden kann
- sie muss für die spezifizierte Service-Qualität eine konkrete Angabe zum Ser-

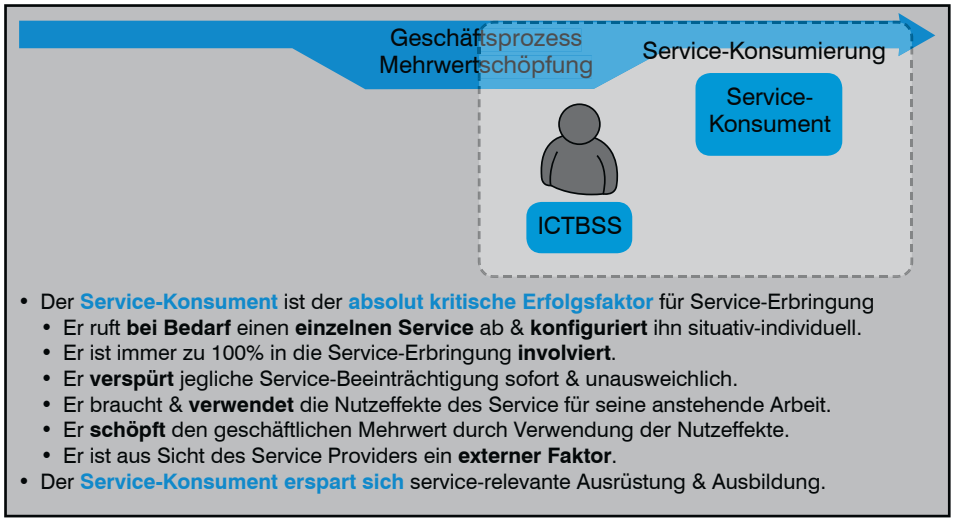


Abbildung 1: Service-Konsument als erfolgskritischer Faktor der Service-Erbringung

Die Service-Spezifikation - das A & O der Service-Erbringung

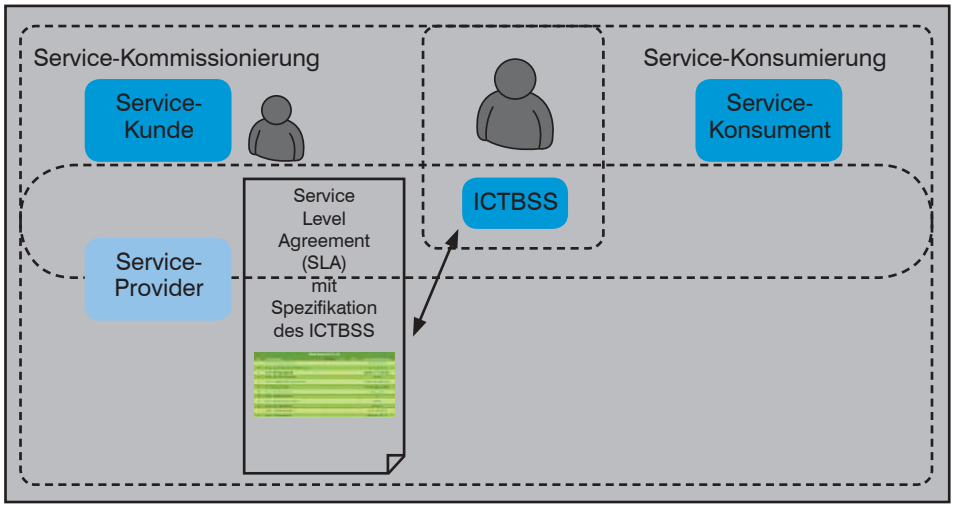


Abbildung 2: Service-Erbringungsmodell mit Hauptrollen und SLA

vice-Tarif bzw. zum Service-Preis enthalten, so dass die Service-Erbringung fundiert und nachvollziehbar verrechnet bzw. fakturiert werden kann

- sie muss einfach, eingängig und einheitlich sein, so dass sie in der gleichen Form sowohl im Service-Katalog als auch im Service Level Agreement (SLA) verwendet werden kann
- sie muss als Grundlage für die durchgängige Service-Konzipierung geeignet sein
- sie muss Basisangaben enthalten, die die fundierte Ermittlung des Service-Aufkommens ermöglichen.

Um all diese Anforderungen abzudecken und eine geeignete Service-Spezifikation für die Servicialisierung zu entwickeln, wurde die folgende Methodik mit 6 Phasen erarbeitet und angewendet: (siehe Abbildung 3)

1. Im Rahmen der Service-Identifizierung werden diejenigen Nutzeffekte herausgearbeitet, die einem Service-Konsumenten auf seinen expliziten Abruf hin jeweils erbracht werden müssen, damit er seine aktuell anstehende geschäftliche Aktivität ausführen kann. Diese Nutzeffekte werden in Form von kurzen Freitextformulierungen beschrieben und vollständig aufgezählt, denn sie bilden den Start- und Zielpunkt jeglicher Service-Erbringung.

2. Ausgehend von den identifizierten Nutzeffekten wird die Service-Qualität auf der Basis weiterer Service-Attribute jeweils eindeutig, vollständig und konsistent spezifiziert. Verschiedene Service Levels werden jeweils durch eigenständige Sets von Attributwerten abgebildet.

3. Alle ausgearbeiteten und gemeinsam verabschiedeten Service-Spezifikationen werden im Service-Katalog zusammengestellt, womit der rechenschaftspflichtige Service Provider sein Service-Angebot klar und kompakt sowie prägnant und nachvollziehbar präsentiert.

4. Die Service-Spezifikation ist jeweils das Ausgangs- und Bezugsobjekt für die Ausarbeitung des durchgängigen Service-Konzepts mit Service Map und Service-Drehbuch. Die Service Map enthält alle erforderlichen Service-Beiträge, aus denen ein abgerufener ICTBSS bei Service-Abruf aggregiert wird. Im Service-Drehbuch wird dargestellt, durch welche Kombination und Sequenz von Service-Beiträgen ein abgerufener ICTBSS jeweils per Echtzeittransaktion aggregiert wird.

5. Nachdem die Service-Konzipierung abgeschlossen ist, werden die erforderlichen ICTBSS vom Service-Kunden aus der Fachabteilung per SLA beim rechenschaftspflichtigen Service Provider kommissioniert. Das SLA enthält neben den organisatorischen Rahmendaten als Kernbestandteil jeweils die Spezifikation des kommissionierten ICTBSS mit dem ausgewählten Service Level sowie die Liste der berechtigten Service-Konsumenten aus dem Verantwortungsbereich des Service-Kunden. Damit verfügt der beauftragte Service Provider über alle Basisdaten, um die verlässliche Service-Erbringung vorzubereiten und zu gewährleisten.

6. Im Rahmen des ‚Service-Triathlon' stellt der rechenschaftspflichtige Service Provider die Service-Erbringungsbereitschaft für diejenigen ICTBSS her, die die Service-Kunden aus den Fachabteilungen per SLA bei ihm kommissioniert haben. Gleichzeitig ermittelt er aus der konkreten Service-Konsumentenanzahl, die in den abgeschlossenen SLAs angegeben wird, und aus dem Service-Abrufverhalten dieser berechtigten Service-Konsumenten das absehbare Service-Aufkommen sowie dessen Verlauf in den relevanten Zeiträumen. Daraus leitet er Anforderungen ab, die einfließen in die Beauftragung der Service Supplier bzw. in das Design und Sizing der service-relevanten ICT-Systeme. Durch das

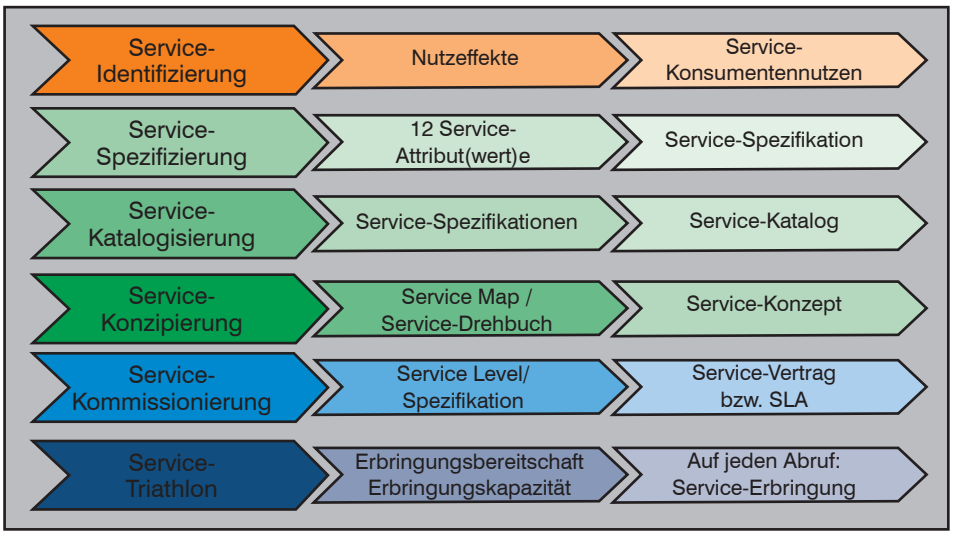


Abbildung 3: Methodik der Servicialisierung