

Vom Blech zum Preis –
Technische und betriebswirtschaftliche Einflussgrößen und
Preismodelle für den Einsatz einer IaaS-Lösung

von

Jörn Seemann

Dieser Text ist unter der Creative Commons Lizenz veröffentlicht



<https://creativecommons.org/licenses/by-nc-nd/3.0/de/>

Inhaltsverzeichnis

Inhaltsverzeichnis	II
Abkürzungsverzeichnis	IV
Abbildungsverzeichnis	V
Tabellenverzeichnis	VI
1. Einleitung	1
1.1 Motivation	1
1.2 Zielsetzung	3
1.3 Aufbau und Methode	3
2. Infrastructure-as-a-Service	4
2.1 Cloud Computing	4
2.2 Positionierung im Cloud Computing	5
2.3 Historie	7
3. Technische Einflussgrößen	9
3.1 CPU	11
3.2 Arbeitsspeicher	13
3.3 Storage	14
3.4 Netzwerk	16
3.5 High-Availability	17
3.6 Elastizität	21
4. Kaufmännische Einflussgrößen	24
4.1 Geschäftsmodell und Wettbewerbsstrategie	25
4.2 Return on Invest	27
4.3 Total Cost of Ownership	27
4.4 Kostenmanagement	29
5. Der Preis und seine Ausprägungen	30
5.1 Preisstrategie	30

5.2	Traditionelle Preisfindung	31
5.3	Target Costing	32
5.4	Der Preis für die Cloud	35
6.	Schlusswort	38
6.1	Zusammenfassung	38
6.2	Fazit	41
6.3	Offene Punkte	42
	Literaturverzeichnis	44
1.	Bücher	44
2.	Internet-Quellen ohne Autor	47
3.	Internet-Quellen mit Autor	49

Abkürzungsverzeichnis

API	-	Application-Programming-Interface
BWL	-	Betriebswirtschaftslehre
CAPEX	-	Capital Expenditures
CRM	-	Customer Relationship Management
HA	-	High Availability
IaaS	-	Infrastructure-as-a-Service
ITIL	-	Information Technology Infrastructure Library
NIST	-	National Institute of Standards and Technology
OPEX	-	Operational Expenditures
PaaS	-	Plattform-as-a-Service
PAYG	-	Pay as you go
pCPU	-	Physikalische CPU
RAID	-	Redundant Array of independent Disks
RAM	-	Random Access Memory
ROI	-	Return on Invest
SaaS	-	Software-as-a-Service
TCO	-	Total Cost of Ownership
vCPU	-	Virtuelle CPU
VM	-	Virtuelle Maschine

Abbildungsverzeichnis

Abbildung 1 – Service-Modelle in der Cloud nach Majid Ali	6
Abbildung 2 – Verteilte Systeme nach Foster et al	9
Abbildung 3 – Die Industrialisierung der IT	10
Abbildung 4 – Überbuchungsfaktor für Arbeitsspeicher am Beispiel VMware	14
Abbildung 5 – Beispielhafter Aufbau einer Cloud-Lösung	18
Abbildung 6 – Kapazitätseinsparungen durch Elastizität	22
Abbildung 7 – Effekt der VM Consolidation	23
Abbildung 8 – Erweiterte Positionierungsoptionen nach Porter	25
Abbildung 9 – Direkte und indirekte Kosten nach Gadatsch	28
Abbildung 10 – Preisstrategien	31
Abbildung 11 – Häufigkeit von Preismodellen	36
Abbildung 12 – Höhere Einnahmen durch hybrides Preismodell	37

Tabellenverzeichnis

Tabelle 1 – Beispiel für die Berechnung von Reserven für Hochverfügbarkeit	20
Tabelle 2 – Berechnung der „Allowable Costs“	32
Tabelle 3 – Bedeutung von Produkteigenschaften für Kunden	33
Tabelle 4 – Auflistung der Standard-Herstellkosten	33
Tabelle 5 – Bewertung der Produkteigenschaften	33
Tabelle 6 – Aufteilung der Allowable Costs auf die Komponentengruppen	34
Tabelle 7 – Ermittlung des Korrekturbedarfs	34
Tabelle 8 – Preismodelle nach Al-Roomi	35

1. Einleitung

Der folgende Text ist als Abschlussarbeit meiner Weiterbildung zum Betriebswirtentstand. Ich möchte mich besonders bei Prof. Dr. Christa Dörge für die Unterstützung und Betreuung bedanken und bei Olga Martin, die mir die Rechtschreibung näher gebracht hat.

1.1 Motivation

Eines der größten Themen der letzten Jahre im Bereich der IT-Infrastruktur war und ist die Virtualisierung¹, bei der das Betriebssystem durch eine zusätzliche Software-Schicht von der Hardware abstrahiert und isoliert wird². Waren es anfangs Themen wie eine bessere Ausnutzung der vorhandenen Server-Ressourcen, traten danach Green-IT und Automatisierung der Prozesse in den Fokus. Heute haben Full-Service-IT-Dienstleister wie z.B. LUFTHANSA SYSTEMS einen Virtualisierungsgrad von 85%³, während Inhouse IT-Abteilungen in 2013 einen Anteil von 46%⁴ erreichten. Die Grundlage für diese Entwicklung bilden IaaS (Infrastructure-as-a-Service) Modelle, in denen die IT-Infrastruktur als fertiger Dienst den Kunden zur Miete angeboten wird. Der Anbieter setzt dabei auf Standardisierung und Automatisierung, um durch Effekte der Economies-of-Scale⁵, wie z.B. Effizienzgewinne und Fixkostendegression⁶, zu geringeren Kosten produzieren zu können, während der Kunde seine benötigten IT-Services günstiger als bisher abrufen kann⁷.

Durch die IaaS-Modelle ergeben sich auf Anbieterseite neue Schwierigkeiten. In der Zeit vor IaaS war die Verrechnung von Ressourcen und die Preisfindung für das gewünschte Produkt einfach: der physikalische Server hat feste Anschaffungskosten, welche auf die Laufzeit der Nutzung umgelegt werden. Zu den Herstellkosten kommen noch Verwaltungs- und Vertriebskosten hinzu. Zusammen mit einem Gewinnaufschlag entsteht der Preis⁸.

Heute sind die technischen Ressourcen, aus denen ein Server besteht, wie Prozessor und Arbeitsspeicher, der Virtualisierung unterworfen. Sie sind vorhanden, aber in der Regel nicht

¹ Vgl. Böckle, Regina (2012)

² Vgl. Münzl, Gerald; Przywara, Bernhard; Reti, Martin; Schäfer, Jörg; Sondermann, Karin; Weber, Mathias; Wilker, Andreas (2009), S. 25

³ Vgl. Lufthansa (2012)

⁴ Vgl. McLellan, Charles (2013)

⁵ Vgl. Weber, Jürgen; Voigt, Kai-Ingo

⁶ Vgl. Weber, Jürgen

⁷ Vgl. Heym, Michael; Montana, Constantin; Elben, Helmut (2012), S. 382

⁸ Vgl. Pepels, Werner (2006), S. 176

mehr exklusiv einer Nutzung zugeordnet. Damit einher geht der Trend zur Zentralisierung⁹, von dem auch Speicherlösungen betroffen sind. Hier spielt Virtualisierung nur eine untergeordnete Rolle, vielmehr treten Mechanismen zur Optimierung der Nutzung des vorhandenen Speicherplatzes in den Vordergrund. Techniken wie Komprimierung senken den effektiv genutzten Speicherplatz. Als Konsequenz ergeben sich auf den Ebenen der CPU, Arbeitsspeicher, Speicherplatz und Netzwerk Ressourcen, die zur Verfügung stehen, deren tatsächliche Nutzung aber nicht mehr transparent ist. Mit dem Fehlen von genauen Angaben der tatsächlich zur Verfügung stehenden technischen Ressourcen und deren Nutzungsgrad ist eine verursachungsgerechte Leistungsverrechnung nicht möglich. Wünschenswert wäre daher die Kenntnis der relevanten technischen Einflussgrößen, um eine entsprechende Bewertung durchführen zu können.

Im kaufmännischen Kontext wird die Frage nach Gesamtkapazität, Nutzungsgrad und Leistungsverrechnung wieder aufgenommen und zusätzlich die der Wirtschaftlichkeit gestellt. Gerade die gewollt hohe Flexibilität¹⁰ einer IaaS-Lösung wirft Fragen auf: welche kaufmännischen Einflussgrößen müssen im Vorfeld berücksichtigt werden, um eine belastbare Investitionsentscheidung treffen zu können, oder wie stellt ein Anbieter den häufigen Wandel des Nutzungsgrades im Rahmen der Kostenrechnung, also der Ergebnissteuerung dar und was wäre ein möglicher Steuerungsmechanismus für den wirtschaftlichen Erfolg? Gerade bei diesen Punkten ist eine genaue Kenntnis der kaufmännischen Einflussgrößen und deren Bewertung für die Messung des wirtschaftlichen Erfolges sehr wichtig.

Hier liegt die Motivation des Autors, der im Unternehmen ETEL für den Betrieb einer bestehenden Virtualisierungslösung verantwortlich ist und die Einführung einer IaaS-Lösung plant. Die Frage nach der tatsächlich zur Verfügung stehenden Kapazität einer IaaS-Lösung steht dabei schon lange im Raum, genauso wie die Frage der verursachungsgerechten Verrechnung von genutzten Ressourcen. Zwei Aspekte, ohne die eine Wirtschaftlichkeitsbetrachtung oder gar -steuerung nicht möglich ist.

⁹ Vgl. NetworkComputing (2012)

¹⁰ Vgl. Münzl, Gerald; Przywara, Bernhard; Reti, Martin; Schäfer, Jörg; Sondermann, Karin; Weber, Mathias; Wilker, Andreas (2009), S. 10

1.2 Zielsetzung

In diesem Dokument soll geklärt werden, wie technische Ressourcen in Hinblick auf Gesamtkapazität und Nutzungsgrad zu beurteilen sind und wie die daraus abgeleiteten technischen und kaufmännischen Einflussfaktoren auf die Kosten einer IaaS-Lösung Einfluss nehmen, welche Preismodelle es gibt und wofür sie sich im Einzelnen eignen. Es wird dabei kein spezieller Business-Case anhand eines fiktiven Beispiels erstellt, sondern es soll dem Leser die Möglichkeit gegeben werden, anhand der eigenen Erfordernisse eine Auswahl der Einflussgrößen zu treffen, welche für die Einführung einer solchen Lösung notwendig ist.

1.3 Aufbau und Methode

Im ersten Kapitel erfolgt die Bestimmung grundlegender Begriffe sowie eine Beschreibung der Entwicklung vom Einzelsystem über Grid-Computing zu IaaS. Eine Abgrenzung zum Thema Cloud Computing erfolgt ebenfalls.

Das anschließende Kapitel gilt der Identifizierung der technischen Einflussgrößen einer IaaS-Lösung. Es wird herausgearbeitet werden, welche Faktoren es gibt und wie die dazugehörigen Bewertungsmaßstäbe aussehen, damit Aussagen zur genutzten und verfügbaren Kapazität möglich sind.

Im vierten Kapitel stehen kaufmännische Einflussgrößen im Mittelpunkt. Dabei geht es zuerst um die Bewertung der Wirtschaftlichkeit einer geplanten Lösung, anschließend um die Beschreibung von grundlegenden Größen der Kostenrechnung, welche im operativen Betrieb für die Kostensteuerung hilfreich und für die Preisfindung notwendig sind.

Preismodellen ist der fünfte Teil gewidmet. Es gibt einen Überblick über bestehende Modelle, sowie eine Beschreibung von Methoden zur Auswahl geeigneter Modelle.

Im sechsten und letzten Abschnitt erfolgt die abschließende Betrachtung, welche Aufschluss darüber gibt, ob mit den zur Verfügung stehenden Daten das Ziel der Arbeit erreicht werden kann und ob es offene Punkte gibt.

Da es sich um eine Zusammenstellung bereits vorhandener Informationen und Daten handelt, bedient sich der Autor der Methode der Sekundärforschung¹¹, also der zusammenfassenden Arbeit und des Erkenntnisgewinns aus der zusammengetragenen Literatur.

2. Infrastructure-as-a-Service

Der Markt für Cloud-Services ist im rasanten Wachstum. Das Marktforschungsunternehmen GARTNER sagt für den Bereich der IaaS Cloud Services bis 2016 eine jährliche Wachstumsrate von 41.3% voraus¹², während das Marktforschungsunternehmen IDC von 31% berichtet¹³. Diese Daten sind mit Vorsicht zu betrachten, denn im gleichen Bericht (siehe Fußnote 12) spricht GARTNER von einer Marktgröße für IaaS-Produkte in 2013 von 9 Mrd. Dollar. Erreicht wurden laut IDC und IHS zwischen 3,6 Mrd. Dollar und 4 Mrd.¹⁴ Dollar. Trotz dieser Schwankungen in Vorhersagen und Ergebnissen lässt eine Recherche nur den Schluss zu, dass es sich hier um einen dynamischen und wachstumsreichen Markt handelt. Woher dieser Markt kommt und was seine technischen Bestandteile sind, davon handeln die nachfolgenden Abschnitte.

2.1 Cloud Computing

Eine allgemeingültige Definition für den Begriff Cloud Computing hat sich bisher nicht herausgebildet. Diese Arbeit verwendet eine Definition der US-amerikanischen Standardisierungsorganisation NIST (National Institute of Standards and Technology), da aus ihr das hier beschriebene Cloud Servicemodell IaaS hervorgeht, sie lautet¹⁵:

“Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction”.

Eine einfachere Definition stammt von der GARTNER GROUP. Für sie ist Cloud Computing „a style of computing in which scalable and elastic IT-enabled capabilities are delivered as a service using Internet technologies”¹⁶.

¹¹ Vgl. Krauß, Jens-Christoph; Schnathmann, Marina; Georgy, Ursula (2014)

¹² Vgl. Columbus, Louis (2013)

¹³ Vgl. IDC (2014)

¹⁴ Vgl. Cloudstrategymag (2014)

¹⁵ Vgl. Mell, Peter; Grance, Timothy (2011), S. 2

¹⁶ Vgl. Gartner

Das NIST geht noch einen Schritt weiter und definiert fünf essentielle Charakteristika von Cloud Computing¹⁷:

- On-Demand Self-Service: Ein Kunde kann selbsttätig Ressourcen konfigurieren (z.B. Netzwerk oder Storage), ohne dabei auf die Hilfe des Anbieters angewiesen zu sein.
- Umfassender Netzwerk-Zugriff: Die Services sind über Standard-Mechanismen zugänglich.
- Bündelung von Ressourcen: Kunden greifen über ein mandantenfähiges System auf die gebündelten Ressourcen des Anbieters zu. Sie haben dabei weder Kontrolle noch Wissen über den genauen Standort der genutzten Ressourcen, können aber durch eine Abstrahierungsschicht den Standort (z.B. Land, Provinz oder Rechenzentrum) festlegen.
- Schnelle Elastizität: Ressourcen können flexibel zur Nutzung gebracht werden. Dies kann automatisiert geschehen, um der Nachfrage nach Ressourcen angemessen hoch oder runter zu skalieren. Für den Kunden scheinen die dabei zur Verfügung stehenden Ressourcen unbegrenzt zu sein und können zu jeder Zeit und in jeder Menge genutzt werden.
- Erfasste Services: Cloud Systeme überwachen und kontrollieren automatisch ihre Ressourcen durch einen Mess-Mechanismus, um sowohl dem Anbieter als auch den Kunden Transparenz über die Nutzung zu geben.

Erfüllt eine Lösung diese Bedingungen, kann von einer Cloud-Lösung gesprochen werden.

2.2 Positionierung im Cloud Computing

Der Begriff Service ist ein zentraler Begriff im Cloud-Modell. Er bezieht sich auf die aus ITIL V2 stammende Definition eines IT Services: „A Service provided to one or more Customers by an IT Service Provider. An IT Service is based on the use of Information Technology and supports the Customer's Business Processes . An IT Service is made up from a combination of people, Processes and technology and should be defined in a Service Level Agreement“¹⁸ und trägt damit dem Umstand Rechnung, dass die Nutzung von Cloud Computing in einem Kunde-Dienstleister Kontext erbracht wird¹⁹.

¹⁷ Vgl. Mell, Peter; Grance, Timothy (2011), S. 2

¹⁸ Vgl. ITIL V2 Glossary (2006) , S. 21

¹⁹ Vgl. Waschke, Marvin (2012), S. 1

Die dabei bezogenen Dienstleistungen lassen sich auf drei Service-Modelle zurückführen²⁰:

- Software-as-a-Service (SaaS): ist Software, angeboten von einem Drittanbieter, verfügbar bei Bedarf, normalerweise via Internet konfigurierbar. Beispiele wären Google Docs und Salesforce CRM.
- Plattform-as-a-Service (PaaS): erlaubt Kunden die Entwicklung unter der Nutzung von APIs neuer Applikationen, durch Fernzugriff ausgerollt und konfigurierbar. Beispiele sind Microsoft Azure und Google App Engine.
- Infrastructure-as-a-Service (IaaS): stellt virtuelle Maschinen und andere abstrahierte Hardware zur Verfügung, sowie Betriebssysteme, welche durch APIs betrieben werden können²¹. Beispiele sind Amazon EC2 und S3, sowie Rackspace Cloud.

Um diese Klassifizierung vornehmen zu können, werden die Tätigkeiten, welche zum Betrieb von IT-Services notwendig sind, in Themenfelder aufgeteilt. Im Ergebnis ist die unterschiedliche Integrationstiefe von Cloud Service-Modellen im Vergleich zum traditionellen Betrieb ersichtlich (siehe Abbildung 1).

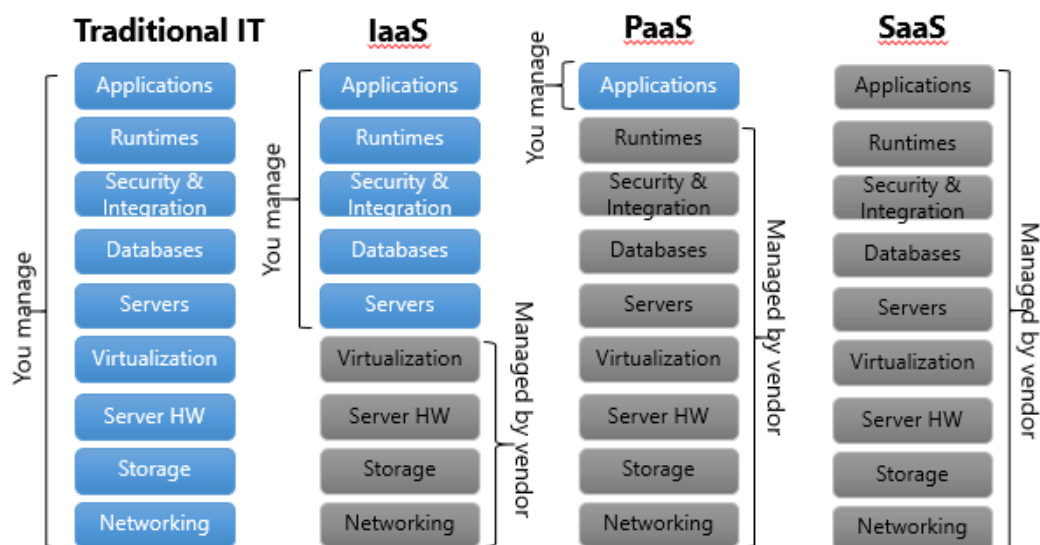


Abbildung 1 – Service-Modelle in der Cloud nach Majid Ali²²

Im Bereich des Cloud Computings beschäftigt sich das Service Modell IaaS demnach mit den Themenfeldern

- Virtualisierung
- Server Hardware

²⁰ Vgl. Mell, Peter; Grance, Timothy (2011), S. 2-3

²¹ Vgl. Mell, Peter; Grance, Timothy (2011), S. 2-3

²² Entnommen aus Ali, Majid (2014)

- Storage und
- Netzwerk

auf die im Hauptteil näher eingegangen wird. Im folgenden Text werden die Begriffe IaaS und Cloud synonym benutzt, da IaaS ein integraler Bestandteil jeder Cloud-Lösung ist. Es nutzt deren Prinzipien und der Rest der Cloud wäre ohne IaaS nicht denkbar.

2.3 Historie

Die Idee des Cloud Computings ist mittlerweile über 50 Jahre alt. Im Jahr 1961 postulierte John McCarthy: "Wenn Computer der Art, wie ich sie befürworte, sich in der Zukunft durchsetzen, wird Rechenkraft eines Tages vielleicht als Versorgungsdienstleistung organisiert sein, so wie es heute das Telefonnetz ist. Die Computerversorger könnten die Grundlage einer neuen, wichtigen Industrie sein"²³. Das Telefon steht hier für einen einfachen Zugriff und für die Abrechnung nur tatsächlich genutzter Ressourcen in Bezug auf Umfang und Zeitraum, während der Computerversorger die heutige IT-Industrie repräsentiert und damit auch das Thema Outsourcing. Versorgungsdienstleistung wiederum steht für niedrige Zugangshürden, also eine Nutzung durch Jedermann und nicht nur durch IT-Experten.

Eine genaue Zuschreibung für die Herkunft des Wortes 'Cloud' im Kontext des Cloud Computings gibt es hingegen nicht. Die am häufigsten zu findende Erklärung lehnt sich an einen Ausspruch von Kevin Marks an, der erklärte, dass der Begriff Cloud aus den frühen Tagen des Internet stamme, als das Netzwerk noch als Wolke gezeichnet wurde, weil egal war wo die Nachrichten hingingen - die Wolke versteckte sie²⁴.

Auf der technischen Ebene war das Time-Sharing der erste Schritt in Richtung Cloud. Im Jahre 1957 von Bob Bemer beschrieben, bot es eine Nutzung von freien Zeiten, in der die CPU auf Ein- und Ausgabeoperationen wartete. Die vorherrschende sequentielle Verarbeitung von Befehlen, welche von Lochkarten gelesen wurden, wurde durch direkte interaktive Eingaben an Terminals abgelöst²⁵.

Aus dem darauf folgenden Zeitalter der Mainframe stammen einige Innovationen, die sich später im Cloud Computing wiederfanden. Dazu gehören die hohe

²³ Vgl. McCarthy, John (1961)

²⁴ Vgl. Mell, Peter; Grance, Timothy (2009), S. 8

²⁵ Vgl. Bemer

Verarbeitungsgeschwindigkeit, die eingebaute Redundanz, Ausfallsicherheit und Verfügbarkeit. Auch Virtualisierung kam hier zum ersten Mal zum Einsatz. Für viele kleine Firmen war es zu teuer, eine eigene Mainframe zu kaufen und zu betreiben, das Mieten von Ressourcen eines Mainframes hingegen nicht²⁶.

Ein weiterer Schritt war das in den 1990er Jahren aufkommende Grid Computing. Allein der Name erlaubt, auf McCarthy Bezug zu nehmen, der Eingangs in seiner Vision von “Versorgern” sprach, während Grid auf das Power Grid, zu Deutsch Stromnetz, referenziert. Das Hauptthema beim Grid-Computing ist das verteilte Rechnen, wobei lose gekoppelte Computer einen virtuellen Supercomputer erzeugen²⁷. Für eine saubere Abgrenzung zu anderen verteilten Systemen definierte Foster in 2002 eine Liste von Kriterien, die ein Grid-System erfüllen muss²⁸:

- es koordiniert Ressourcen, die nicht einer zentralen Kontrolle unterliegen,
- nutzt offene, standardisierte Protokolle und Schnittstellen
- stellt nicht-triviale Dienstgütern zur Verfügung.

Als nächstes schickte sich Utility Computing an, McCarthys Vision zu erfüllen. Dabei handelt es sich nicht um neue technische Möglichkeiten, sondern um Business- und Prozess-Modelle, die Kunden mess- und damit abrechenbare technische Ressourcen zur Verfügung stellen. Das Neue daran war, dass Kunden nun Zugriff auf große technische Ressourcen hatten, ohne die dafür sonst notwendigen Investitionen tätigen zu müssen. Es wurde nur der tatsächliche Bedarf gemietet und das auch nur für einen bestimmten Zeitraum²⁹.

²⁶ Vgl. Krishnan, Sriram (2010), S. 3

²⁷ Vgl. Foster, Ian T.; Zhao, Yong; Raicu, Ioan; Lu, Shiyong (2008), S. 7

²⁸ Vgl. Foster, Ian (2002), S. 2-3

²⁹ Vgl. Foster, Ian T.; Zhao, Yong; Raicu, Ioan; Lu, Shiyong (2008), S. 2

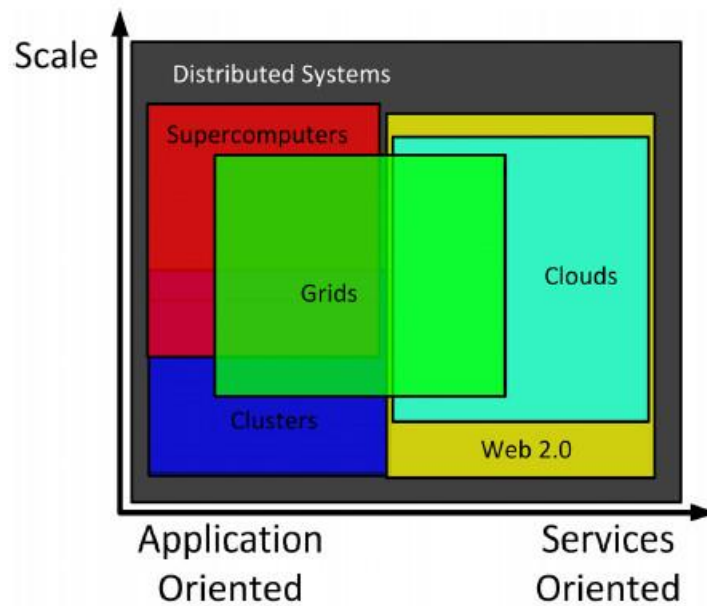


Abbildung 2 – Verteilte Systeme nach Foster et al³⁰

Zusammenfassend lässt sich sagen, dass Cloud Computing nichts revolutionär Neues ist. Es ist eine sich über Jahrzehnte andauernde Weiterentwicklung von Themen und Techniken, die daher Überschneidungen bilden, sich einander bedingen und auch nebeneinander existieren, wie in Abbildung 2 zu sehen ist.

3. Technische Einflussgrößen

Anfangs war das Thema IT in Firmen eine Randerscheinung. Es erfolgte eine evolutionäre Entwicklung über die Stationen einer handwerklich orientierten IT, die Tätigkeit als Manufaktur hin zu IT Commodities (siehe Abbildung 3). Es wird auch von der Industrialisierung der IT³¹ gesprochen.

³⁰ Entnommen aus Foster, Ian T.; Zhao, Yong; Raicu, Ioan; Lu, Shiyong (2008), S. 2

³¹ Vgl. Holtschke, Bernhard; Heier, Hauke; Hummel, Thomas (2009), S. 19-20

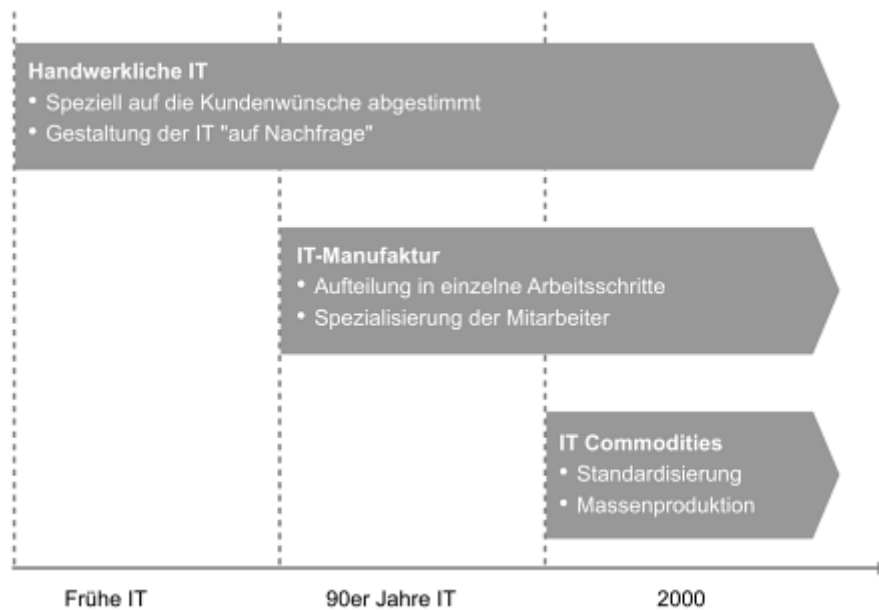


Abbildung 3 – Die Industrialisierung der IT³²

Schon die genutzten Begrifflichkeiten Industrialisierung, Commodities, zu Deutsch Verbrauchsstoffe, und Massenproduktion weisen in Richtung Betriebswirtschaft. Sie ist ein wichtiger Teil der IT geworden. Auf der einen Seite Technik, auf der anderen Seite wirtschaftliches Handeln.

Wirtschaftliches Handeln fußt dabei auf dem ökonomischen Prinzip, welches zwei Aussagen hat³³:

- Minimumprinzip: ein vorgegebenes Ziel (Erfolg, Ergebnis) mit dem geringstmöglichen Mitteleinsatz zu erreichen
- Maximumprinzip: das sich mit gegebenem Mitteleinsatz der größtmögliche Zielerreichungsgrad (Erfolg, Ergebnis) einstellt

Übertragen auf die in der Einleitung beschriebene Problemstellung geht es in den folgenden Abschnitten darum, ein Maximum an Kapazität aus den eingesetzten Komponenten zu erzielen und dabei ein Minimum an Mitteleinsatz, also Geld, zu erreichen.

Wie in Kapitel 2.2 beschrieben, besteht eine IaaS-Lösung aus den Teilbereichen

- Virtualisierung,
- Server Hardware,
- Storage und

³² Entnommen aus Holtschke, Bernhard; Heier, Hauke; Hummel, Thomas (2009), S. 20

³³ Vgl. Behrens, Christian-Uwe (2010), S. 57

- Netzwerk.

Diese Einteilung ist recht grob. Kashef et al³⁴ beschreiben in ihrer Metaanalyse insgesamt 20 Kostenfaktoren in 6 Kostentypen für den Betrieb einer Cloud. Nicht alle davon haben einen technischen Hintergrund oder sie haben keinen Bewertungsspielraum in Bezug auf ihre Endgröße.

Ein Beispiel: der Kostenfaktor Strom besteht aus den Bestandteilen Höhe des Stromverbrauchs und Strompreis. Allenfalls die Fragen nach Höhe des Strompreises und ob eine Pauschal- oder eine verbrauchsabhängige Abrechnung erfolgen soll, haben einen Einfluss auf den Kostenfaktor. Die Höhe des Stromverbrauchs und Strompreis sind feste Größen ohne Bewertungsspielraum.

In den folgenden Abschnitten geht es um Kostenfaktoren, bei denen diese einfache Bewertung nicht möglich ist, weil durch besondere Methoden oder Metriken erst die Grundlage für die Beurteilung der Gesamtkapazität einer IaaS-Lösung geschaffen werden muss. Dieser Schritt ist wichtig, weil für die Beurteilung der Wirtschaftlichkeit die Frage, wie viel Ressourcen (=Kapazität) zur Verfügung stehen, beantwortet werden muss. Denn „aus ökonomischer Sicht beschreibt Wirtschaftlichkeit das Verhältnis von monetär quantifizierbaren Kosten zur gemessenen Leistung“³⁵. Dieser Abschnitt handelt davon, die Leistung messbar zu machen. Dabei sollen die erhobenen Werte

- nachvollziehbar,
- eindeutig und
- unabhängig von Herstellern

sein.

3.1 CPU

Die CPU eines normalen Servers ist die meiste Zeit zu 10% bis 50% ausgelastet³⁶. Durch die in IaaS-Lösungen eingesetzte Virtualisierungs-Software wird eine Abstraktions-Schicht zwischen Hardware und Gast-Betriebssystemen eingeschoben. Diese Hypervisor genannte Software verfügt mit dem CPU-Scheduler über die Möglichkeit, die eigentlich nur einmal zur Verfügung stehende Ressource CPU mehrfach zur Nutzung bereit zu stellen. Der Hypervisor

³⁴ Vgl. Kashef, Mohammad Mahdi; Altmann, Jörn (2011), S. 46–60

³⁵ Vgl. Gronau, Norbert (2012)

³⁶ Vgl. Barroso, Luiz André; Hölzle, Urs (2007), S. 34

stellt dabei dem Gast-Betriebssystem eine virtuelle CPU (vCPU) zur Verfügung. Dem zur Ausführung bereitstehende Prozess im Gast-Betriebssystem werden dann über den CPU-Scheduler des Hypervisor Zeitschlitz auf einer physikalischen CPU (pCPU) zugewiesen und zur Ausführung gebracht³⁷.

Die Bestimmung der Gesamt-CPU Kapazität eines Virtualisierungs-Hosts ist mit der sog. Consolidation Ratio möglich, in manchen Publikationen auch VM Density³⁸ genannt. Es beschreibt das Verhältnis von physikalischen Servern und darauf betriebenen VMs. Wenn also 450 VMs auf 10 physikalischen Servern betrieben werden, beträgt das Consolidation Ratio 45:1³⁹. Je höher die Anzahl der VMs und je niedriger die Anzahl der physikalischen Server, desto besser das Ratio.

Das vCPU to pCPU Ratio oder vCPU per Core Ratio gibt an, wie viele virtuelle CPUs (vCPU) auf einer physikalischen CPU (pCPU oder Core) ausgeführt werden können⁴⁰. Bei einer Ratio von 15:1 können 15 vCPUs auf einer pCPU ausgeführt werden. Je höher die Anzahl der vCPUs auf einer pCPU, desto besser das Ratio.

Eine andere Methode ist das Aufaddieren von eingeplanten CPU-Ressourcen. Steht zum Beispiel eine CPU mit 4 Kernen je 3,0 Ghz zur Verfügung, ergibt dies 12 Ghz CPU Kapazität. Bei einer durchschnittlichen Arbeitslast von 80% verbleiben 9,6 GHz oder 2,4 GHz pro Core, die für einzelne VMs verplant werden können⁴¹.

Server, die über lediglich eine CPU verfügen, sind nicht mehr erhältlich, aktuell sind mehrere CPUs mit jeweils mehreren Cores Standard. Aus diesem Grund macht das Consolidation Ratio für sich allein keine Aussage über die CPU-Kapazität. Erst mit der Information über die Anzahl der CPU und deren Cores lässt sich eine konkrete Aussage machen und zusätzlich das vCPU to pCPU Ratio ableiten. Ein Beispiel: Beim oben genannten Consolidation Ratio von 45:1 verfügt ein Server über 2 Prozessoren mit jeweils 6 Cores, was einem vCPU per Core Ratio von 3,75 entspricht. Das Consolidation Ratio und vCPU per Core Ratio gehören also zusammen, um eine konkrete Aussage über die Leistungsfähigkeit eines Virtualisierungs-Hosts in Bezug auf die Virtualisierung machen zu können. Anzustreben ist ein möglichst hohes Consolidation- sowie vCPU per Core Ratio.

³⁷ Vgl. Song, Weijia; Xiao, Zhen (2013), S. 304-305

³⁸ Vgl. Li, Xinhui; Li, Ying; Liu, Tiancheng; Qiu, Jie; Wang, Fengchun (2009), S. 97

³⁹ Vgl. Guthrie, Forbes; Lowe, Scott; Coleman, Kendrick (2013), S. 405

⁴⁰ Vgl. Righini, Marco (2012), S. 5

⁴¹ Vgl. Guthrie, Forbes; Lowe, Scott; Coleman, Kendrick (2013), S. 404

Letztlich kann man zu Anfang nur Annahmen treffen, auf Herstellerangaben⁴² setzen oder Studien⁴³ nutzen um danach eigene Werte zu ermitteln. In der Praxis ist die Nutzung von aufaddierten CPU-Daten keine gebräuchliche Methode. IaaS Anbieter differenzieren ihre Produkte über die Anzahl der Cores⁴⁴, eine Nennung von CPU-Geschwindigkeiten spielt keine Rolle.

3.2 Arbeitsspeicher

Der Arbeitsspeicher (RAM) eines Computers war schon immer eine kostbare Ressource. Im Jahre 1957 hätte 1 MByte Arbeitsspeicher über 1 Milliarde Dollar⁴⁵ gekostet. Der Preis ist seitdem stark gesunken⁴⁶, andererseits benötigen Computerprogramme immer mehr Arbeitsspeicher. Über die Zeit wurden daher verschiedene Methoden entwickelt, mehr Speicher zur Verfügung zu stellen, als tatsächlich vorhanden ist. Eine ist die Nutzung von Swap. Wie der Name schon andeutet, wird der „teurere“ Arbeitsspeicher gegen „günstigen“ Festplattenspeicher ausgetauscht.

Die Virtualisierungs-Software einer Cloud kennt weitere Methoden, mehr Arbeitsspeicher zur Verfügung zu stellen als physikalisch vorhanden ist. Beim Page Sharing werden mehrere gleiche Speicherseiten zu einer zusammengefasst⁴⁷. Ballooning hingegen zwingt das Gast-Betriebssystem zum Swappen. Es benötigt dann weniger Arbeitsspeicher auf den Virtualisierungs-Host, der ihn dann anderen VMs zur Verfügung stellen kann⁴⁸. Alle zuvor beschriebenen Techniken fallen unter den Begriff Memory Overcommitting⁴⁹.

Ein Virtualisierungs-Host ist dann Memory Overcommitted, wenn die zusammengerechnete zugewiesene Größe des Arbeitsspeichers aller eingeschalteten VMs größer ist als der physikalische Arbeitsspeicher des Virtualisierungs-Hosts. Aus diesen beiden Größen lässt sich dann ein Ratio bilden, mit dem eine Beurteilung der Auslastung möglich ist. Wie hoch das Ratio ausfallen darf, ist abhängig von der eingesetzten Virtualisierungs-Software. Die

⁴² <http://www.vmware.com/consolidation/overview>

⁴³ Vgl. Chang, Bao Rong; Tsai, Hsiu-Fen; Chen, Chi-Ming (2013), S. 197

⁴⁴ Vgl. Read, Jason (2014)

⁴⁵ Vgl. McCallum, John C.

⁴⁶ Vgl. ebd.

⁴⁷ Vgl. Lowe, Scott D. (2012), S. 5

⁴⁸ Vgl. ebd., S. 5

⁴⁹ Vgl. ebd., S. 5-6

Firma Red Hat empfiehlt für den KVM Hypervisor ein Ratio von 1,5⁵⁰, während VMware in einer Studie⁵¹ keine direkte Empfehlung gibt. Wie aus Abbildung 4 ersichtlich,

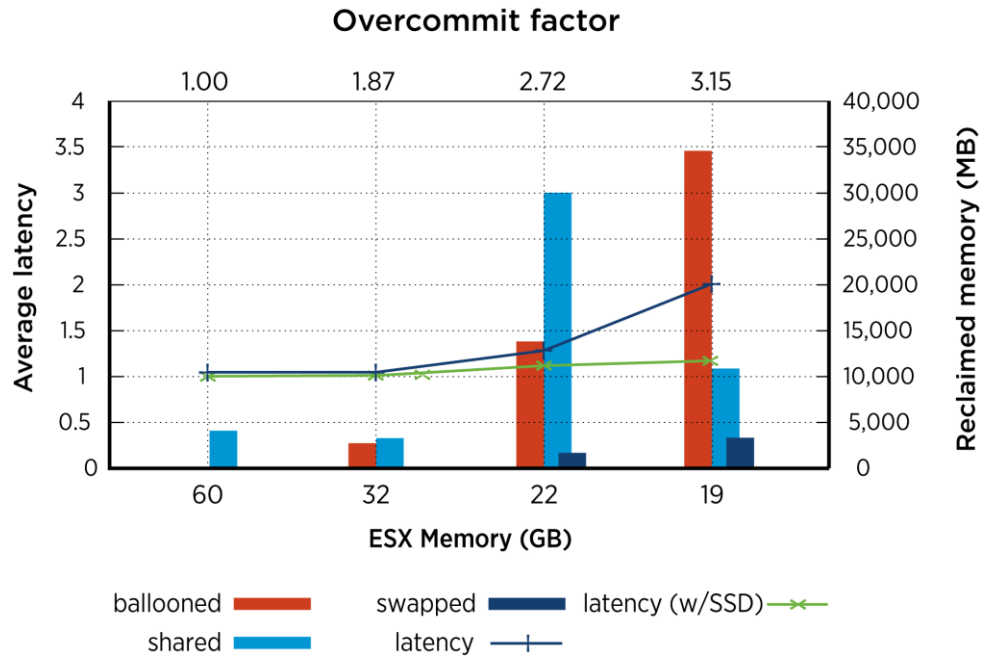


Abbildung 4 – Überbuchungsfaktor für Arbeitsspeicher am Beispiel VMware⁵²

ist ab Faktor 1,8 mit Performanceeinbußen zu rechnen. Für den Xen Hypervisor wird ein Faktor zwischen 1,05 und 1,1 empfohlen⁵³. Das Erheben eines eigenen Faktors anhand von Erfahrungswerten ist ebenfalls möglich.

3.3 Storage

Für die Bestimmung der Storage-Kapazität einer IaaS-Lösung ist es notwendig, die Kapazitätsnutzung über die gesamte Umgebung zu messen. Die Schwierigkeit ist darin begründet, dass jeder Hersteller seine eigenen Begrifflichkeiten und Methoden hat, um die Kapazitäten seiner Produkte zu beschreiben.

Es fängt im Kleinen bei den Festplatten an. Einige Hersteller geben die Größe der Festplatten auf Basis des Dezimalsystems an (1 GB = 10⁶ Byte) an, einige auf Basis des Binärsystems (1

⁵⁰ Vgl. Holzer, Jan Mark; Curran, Chris; Radvan, Scott Kapitel 33.4

⁵¹ Vgl. Banerjee, Ishan; Guo, Fei; Tati, Kiran; Venkatasubramanian, Rajesh (2013)

⁵² Entnommen aus Banerjee, Ishan; Guo, Fei; Tati, Kiran; Venkatasubramanian, Rajesh (2013)

⁵³ Vgl. Citrix Systems, S. 28

GiB = 2^{30}). Inhaltlich ist dieses Vorgehen nach IEC 60027-2⁵⁴ korrekt, solange die richtigen Begrifflichkeiten angewandt werden, macht aber aus einer 1 TB Festplatte eine mit 931 GiB.

Im nächsten Schritt folgt die Absicherung gegen Datenverlust durch Festplattenausfall. Neben allgemein bekannten RAID-Level wie 1, 3 oder 5 gibt es eine Vielzahl von Herstellerspezifischen Definitionen und Lösungen⁵⁵. Jede davon hat sein eigenes Verhältnis von Nutz- zu Rohkapazität, das nicht immer transparent dargestellt wird.

Ein weiterer Abstraktionsgrad entspringt dem Thin- bzw. Thickprovisioning. Bei Thickprovisioning wird der gesamte angeforderte Speicherplatz exklusiv zugewiesen. Nicht genutzter Speicherplatz kann nicht anderweitig eingesetzt werden. Anders beim Thinprovisioning. Hier stellt das Storage-System nur den Anschein nach den gesamten Speicherplatz zur Verfügung, es wird nur der voraussichtlich benötigte Speicher belegt, womit der nicht belegte Speicher für andere Zwecke genutzt werden kann⁵⁶.

Auf Festplatten werden Daten in 512 Byte Blöcken gespeichert. Storage-Systeme bestehen aus einer Vielzahl von Festplatten. Die Wahrscheinlichkeit, dass sich in Aber-Milliarden Datenblöcken die gleichen Daten befinden, ist sehr groß. Zum Beispiel wird nicht genutzter, zugewiesener Speicher von Dateisystemen mit Nullen beschrieben. Deduplikation setzt hier an, indem es dafür sorgt, dass Blöcke mit identischem Inhalt nur einmal gespeichert werden. Diese Ersparnis wird i.d.R. nur in internen Statistiken der Storage-Systeme dargestellt, für Nutzer ist dieser Effekt nicht sichtbar.

Es gibt eine Vielzahl weiterer Ausprägungen von Storage, deren Ausführung hier zu weit führen würde. Die bereits beschriebenen Verfahren sollen nur einen Eindruck geben, wie schnell Zahlentransparenz verloren gehen kann. Sind unterschiedliche Hersteller im Einsatz, ist die Komplexität Zahlen zu erheben und vergleichbar zu machen ungleich höher. Es ist also notwendig, Metriken einzusetzen, die gleiche Aussagen über eine heterogene und sich ständig ändernde technische Basis zulassen⁵⁷. Als Grundlage müssen daher Zahlen dienen wie z.B. Gesamt Rohkapazität, also die Summe der Kapazität aller Festplatten in einem Storage-System. Die Menge des belegten Speichers durch Nutzer, die sogenannten Nutzerdaten oder zugewiesener Speicher, also Speicher, der zur Verfügung steht, aber nicht genutzt wird.

⁵⁴ Vgl. NIST

⁵⁵ Vgl. Rieß, Ulrike (2011)

⁵⁶ Vgl. Wurm, Andreas (2010)

⁵⁷ Vgl. Intel Corporation (2011), S. 3

Aus diesen Größen ergeben sich folgende Metriken:

- Gesamt Storage Effektivität – Das Verhältnis von Nutzerdaten zur Gesamt Rohkapazität⁵⁸.
- Anteil Low-Cost Storage – Das Verhältnis von Nutzerdaten auf dem Low-Cost Storage Tier im Vergleich zu allen gespeicherten Nutzerdaten⁵⁹.
- Anteil genutzte Kapazität – Dies ist der Anteil der nutzbaren Kapazität, die von Nutzerdaten belegt werden. Diese Metrik kann nie über 100% liegen⁶⁰.
- Anteil zugewiesener Speicher – Dies ist der Anteil der nutzbaren Kapazität, die Kunden zugewiesen wurde. Mit Thin-Provisioning kann diese Metrik 100 Prozent übersteigen⁶¹.
- Auslastung der zugewiesenen Kapazität für Nutzerdaten – Dies beschreibt, wie viel von dem ihnen zugewiesenen Speicher bestimmte Kunden in Nutzung haben. Im Idealfall nutzt er 100% des zugewiesenen Speichers. Die meisten Kunden nutzen tatsächlich nur einen kleinen Prozentsatz des zugewiesenen Speichers. In diesem Fall ist ein höherer Anteil der Überprovisionierung im Storage-System möglich⁶².
- Zuweisungs-Spielraum – Anbieter müssen wissen, wie viel Spielraum nach oben sie noch haben, wenn Kunden mehr Speicherplatz anfordern. Der Zuweisungs-Spielraum ist definiert als Speicherkapazität in GB, der noch zugewiesen werden kann⁶³.

Liegen diese Zahlen nicht vor oder es gibt keine Erfahrungswerte, ist das Treffen von Annahmen notwendig. Es ist zum Beispiel möglich die durchschnittliche Größe des Speicherplatzes einer VM im IaaS-Markt zu ermitteln und mit der erwarteten Absatz-Menge im ersten Jahr zu multiplizieren, um auf eine Startgröße zu kommen.

3.4 Netzwerk

Der Datentransfer einer IaaS-Lösung lässt sich in drei Kategorien einteilen. Der Anteil Daten, welcher das Netzwerk in Richtung Internet verlässt, wird Northbound genannt, der umgekehrte Weg nennt sich entsprechend Southbound. Datenverkehr innerhalb des internen Netzwerkes ist demnach East-/Westbound⁶⁴. Diese Daten lassen sich aus den Netzwerk-

⁵⁸ Vgl. ebd., S. 4

⁵⁹ Vgl. ebd., S. 4

⁶⁰ Vgl. ebd., S. 6

⁶¹ Vgl. ebd., S. 7

⁶² Vgl. ebd., S. 7

⁶³ Vgl. Intel Corporation (2011), S. 7

⁶⁴ Vgl. Waschke, Marvin (2012), S. 110-111

Komponenten auslesen, den Verursachern zuordnen und damit abrechnen. Einen Bewertungsspielraum gibt es an dieser Stelle nicht.

Ein Blick in die Preistabellen von Microsoft Azure⁶⁵, Amazon EC2⁶⁶ oder Rackspace⁶⁷ zeigt jedoch, dass weder für Southbound noch für East-/Westbound Kosten berechnet werden. Die Frage der Bewertung ist demnach keine technische Frage, sondern eine des Geschäftsmodells (siehe Seite 25). Daten sind heutzutage eine wichtige Ressource. Provider möchten, dass Daten in ihre Cloud fließen und sie dort halten⁶⁸, andererseits beruht das Business-Modell der Cloud-Nutzer auf Datentransfer ins Internet. Dies belegt schon die Tatsache, dass das Aufrufen einer Webseite weniger eingehenden Datentransfer erzeugt als das Ausliefern der angeforderten Seite.

Die Frage nach dem jeweiligen Geschäftsmodell kann an dieser Stelle nicht beantwortet werden. Im Rahmen der Produktentwicklung muss entschieden werden, sich den großen IaaS-Providern anzuschließen oder einen eigenen Weg zu gehen. Keiner der drei oben genannten Firmen garantiert beispielsweise eine Bandbreite oder Antwortzeit für seine Produkte. Hier läge eine Möglichkeit, sich vom Markt zu diversifizieren, ein Kostenmodell für Bandbreite⁶⁹ liegt bereits vor.

3.5 High-Availability

Hochverfügbarkeit ist ein zentraler Bestandteil von Cloud- bzw. IaaS-Lösungen. Eine Definition von Hochverfügbarkeit lautet:

„Ein System gilt als hochverfügbar, wenn eine Anwendung auch im Fehlerfall weiterhin verfügbar ist und ohne unmittelbaren menschlichen Eingriff weiter genutzt werden kann. In der Konsequenz heißt dies, dass der Anwender keine oder nur eine kurze Unterbrechung wahrnimmt. Hochverfügbarkeit (abgekürzt auch *HA*, abgeleitet von engl. *high availability*) bezeichnet also die Fähigkeit eines Systems, bei Ausfall einer seiner Komponenten einen uneingeschränkten Betrieb zu gewährleisten“⁷⁰.

⁶⁵ Vgl. MicrosoftPricing

⁶⁶ Vgl. AmazonPricing

⁶⁷ Vgl. RackspacePricing

⁶⁸ Vgl. Yuan, Dong (2011), S. 26

⁶⁹ Vgl. Di Niu; Feng, Chen; Li, Baochun (2012)

⁷⁰ Vgl. Held, Andrea (2004), S. 29

Aus dieser Definition lassen sich zwei Arten der Hochverfügbarkeit ableiten. Das Wort „Anwendung“ verweist auf Applikations-Hochverfügbarkeit, während „bei Ausfall einer seiner Komponenten“ die Infrastruktur und deren Verfügbarkeit ansprechen.

In der traditionellen IT liegt der Ansatz in der Vermeidung von Fehlern, indem auf Infrastruktur-Ebene eine möglichst große Robustheit erzeugt wird. Es gibt alles, was einen Ausfall verursachen kann, mindestens doppelt und Umschalt-Mechanismen, damit die Applikation nicht mit dem Ausfall der Infrastruktur umgehen muss⁷¹.

Cloud Computing geht einen anderen Weg: „Rule of thumb: Be a pessimist when designing architectures in the cloud ; assume things will fail. In other words, always design, implement and deploy for automated recovery from failure“⁷².

Dieses Built-to-fail Paradigma überträgt sich ebenfalls auf IaaS, welches die Infrastruktur-Ebene des Cloud Computings darstellt. Der Ansatz dabei ist ein mehrstufiger: jede Hardware verfügt über Komponenten, die ausfallen können, daher gibt es immer mehrere, die in Clustern zusammengefasst werden. In einem Rechenzentrum gibt es mehrere Cluster und gegen den Ausfall eines Rechenzentrums schützt der Anbieter sich durch die Nutzung von mehreren Rechenzentren. Amazon betreibt beispielsweise zwölf Rechenzentren auf vier Kontinenten⁷³.

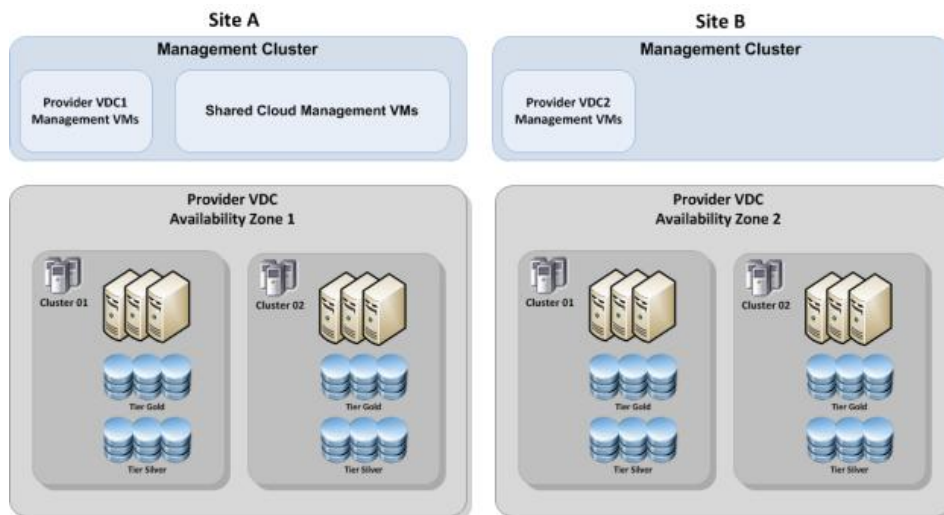


Abbildung 5 – Beispielhafter Aufbau einer Cloud-Lösung⁷⁴

⁷¹ Vgl. Leong, Lydia (2010)

⁷² Vgl. Buyya, Rajkumar; Broberg, James; Goscinski, Andrzej (2011), Abschnitt 18.4.1

⁷³ Vgl. Amazon-Regionen (2014)

⁷⁴ Entnommen aus Fojta, Tom (2014)

Abbildung 5 zeigt den beispielhaften Aufbau einer IaaS-Lösung mit mehreren Clustern pro Standort und zwei Standorten.

Zur Frage der Bewertung des Themas High-Availability hinsichtlich Kapazität lässt sich in der einschlägigen Literatur nichts finden. Entweder wurde das Thema bisher nicht als wichtig identifiziert oder es wurde übersehen. Um trotzdem zu einer Aussage zu kommen, ist es möglich selbst über das Treffen von Annahmen die benötigten Reservekapazitäten zu ermitteln (siehe Tabelle 1). Zuerst werden die Kapazität eines Virtualisierungs-Hosts und die Größe eines Clusters festgelegt, um einen Multiplikationsfaktor zu erhalten. Es folgt die Anzahl der Cluster in einem Rechenzentrum und schließlich die Zahl der Rechenzentren. Um am Ende einen Überblick der benötigten Reserven zu erhalten wird für jeden dieser Faktoren ein Wert festgelegt, der als Reserve vorgehalten werden muss.

Anzahl Virtualisierungs-Hosts in einem Cluster	Reserve	Reserve-Einheiten
10	20%	2
Anzahl Cluster in einem Rechenzentrum	Reserve	Reserve-Einheiten
100	20%	20
Anzahl Rechenzentren	Reserve	Reserve-Einheiten
10	20%	2

Tabelle 1 – Beispiel für die Berechnung von Reserven für Hochverfügbarkeit (eigene Darstellung)

Auch wenn es in diesem Abschnitt nur um technische Größen gehen soll, ein kurzer Hinweis auf die Betriebswirtschaft: durch dieses High-Availability Modell werden mehr Ressourcen vorgehalten, als für den Betrieb notwendig sind. Anhand Tabelle 1 ist die Größenordnung zu sehen. Reserven in diesen Dimensionen vorzuhalten kostet viel Geld und es stellt sich die Frage, ob Kunden bereit sind, für dieses Niveau an Sicherheit zu bezahlen.

Cloud-Anbieter haben einen anderen Ansatz: „The cloud philosophy is generally that you buy dirt-cheap things and expect they’ll fail. Since you’re scaling out anyway, you expect to have a bunch of boxes, so that any box failing is not an issue. You protect against data center failure by being in multiple data centers”⁷⁵. Es sind demnach aufgrund der Skalierungsfähigkeit (Elastizität – siehe nächster Abschnitt) der Infrastruktur immer so viele Ressourcen vorhanden, dass der Ausfall von Teilen der Infrastruktur keine Rolle spielt. Sollte dieses System einmal versagen, ist der betriebswirtschaftliche Schaden für den Anbieter überschaubar. Weder Microsoft⁷⁶, Amazon⁷⁷ noch Rackspace⁷⁸ zahlen im Falle eines Verstoßes gegen den vereinbarten Service-Level Schadensersatz, sondern bieten (manchmal sogar nur anteilig) eine Gutschrift auf das geleistete Service-Entgelt.

Welches Modell für die eigene IaaS-Lösung passt, ist keine technische Frage, sondern eine der Wettbewerbsstrategie, auf die im kaufmännischen Teil eingegangen wird.

⁷⁵ Vgl. Leong, Lydia (2010)

⁷⁶ Vgl. MicrosoftSLA (2014)

⁷⁷ Vgl. AmazonSLA (2013)

⁷⁸ Vgl. RackspaceSLA (2014)

3.6 Elastizität

Das Wort elastisch steht laut Duden für die „Fähigkeit eines Körpers oder eines Stoffes, eine durch äußere Einwirkung hervorgerufene Formänderung aus eigener Kraft wieder rückgängig zu machen“⁷⁹. Körper oder Werkstoffe mit diesen Eigenschaften lassen sich sofort verändern, kehren zum Ausgangszustand wieder zurück, außerdem ist der Vorgang wiederholbar⁸⁰, so die Definition von Elastizität in der Physik.

In der traditionellen IT sind Architekturen nicht elastisch, sie sind skalierbar. Allerdings wird hoch und runter skalieren nicht sehr oft gemacht. Ursache ist die Zeit und das Geld, welche aufgewandt werden müssen, um zusätzliche Server zu kaufen, einzubauen und in die Umgebung einzubringen. Der Anbieter tätigt mit dem Kauf ebenfalls eine längerfristige Investition⁸¹. Will er nur Lastspitzen abfedern, ist dieses Vorgehen nicht besonders effektiv, es ist sehr statisch.

In der Cloud werden Lastspitzen abgefedert, indem mehr Ressourcen hinzugefügt werden, diese treten aber in Form von zusätzlichen VMs in Erscheinung. Die Last verteilt sich so nicht mehr auf einen Server, sondern auf beliebig viele. Im Zusammenspiel mit einer rein nutzungsabhängigen Abrechnung und der Fähigkeit einer Cloud über seine Elastizität sofort zusätzliche Ressourcen zur Verfügung zu stellen ist eine große Flexibilität möglich. Die Elastizität einer Cloud lässt sich dabei mit einem Gummiband vergleichen, die Kapazität lässt sich dehnen und zieht sich aber auch wieder zusammen, wenn die Ressourcen freigegeben werden⁸². Auch wenn es eines der fünf Grundcharakteristika der Cloud (siehe Seite 5) ist, hat sich keine allgemeingültige Definition der Elastizität einer Cloud herausgebildet. Herbst et al. hat verschiedenste Definitionen zusammengetragen und diese Formulierung vorgeschlagen: “Elasticity is the degree to which a system is able to adapt to workload changes by provisioning and deprovisioning resources in an autonomic manner, such that at each point in time the available resources match the current demand as closely as possible“⁸³.

Elastizität entsteht durch Ressourcen-Pooling (siehe Seite 5). Mit Hilfe der in den vorherigen Abschnitten besprochenen Techniken wie CPU- sowie Arbeitsspeicher-Virtualisierung, Thin-Provisioning von Storage-Systemen und Überbuchung wird ein Pool von Ressourcen

⁷⁹ Vgl. Duden (2013)

⁸⁰ Vgl. Berliner, Arnold (2013), S. 135

⁸¹ Vgl. Schouten, Edwin (2012)

⁸² Vgl. ebd.

⁸³ Vgl. Herbst, Nikolas Roman; Kounev, Samuel; Reussner, Ralf (2013), S. 2

geschaffen, den Kunden sich teilen und auf den sie gemeinsam zugreifen können. Die Idee dahinter ist, dass die Auslastung der von Kunden genutzten Ressourcen sehr unterschiedlich ausfällt. Nicht jeder nutzt zur gleichen Zeit seine Ressourcen voll aus. Die Lastspitze des einen Kunden wird dabei durch den Leerlauf anderer Kunden kompensiert. Je größer der Pool, desto größer die Stabilität des Pools gegen große Lastschwankungen⁸⁴, dieses Verfahren wird statistisches Multiplexing genannt⁸⁵.

Das statistische Multiplexing hat für Anbieter und Kunde unterschiedliche Auswirkungen. Für den Kunden steht eine scheinbar unerschöpfliche Kapazität zur Verfügung, während der Anbieter mehr Ressourcen anbieten kann als physikalisch vorhanden sind. Letztlich entsteht hier die Wertschöpfung im Cloud Computing: “Since the cost efficiency of the cloud model critically depends on ability of the cloud providers to leverage statistical multiplexing, the clouds often operate in the over commit mode by design“⁸⁶.

Die Ersparnis durch statistisches Multiplexing auf Anbieterseite ist nicht unerheblich – bis zu 25%, wie Abbildung 6 zeigt:

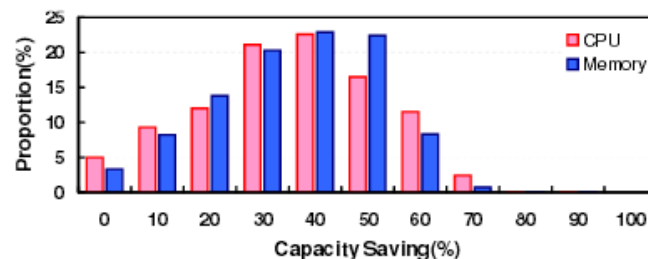


Abbildung 6 – Kapazitätseinsparungen durch Elastizität⁸⁷

Natürlich ist die Kapazität einer Cloud-Lösung nicht unendlich. Sie ist trotz allem begrenzt durch die zur Verfügung stehenden physikalischen Komponenten und deren Leistungsfähigkeit. Hinzu kommen noch die bisher erarbeiteten Überbuchungsfaktoren für CPU, Arbeitsspeicher, Netzwerk, Storage und schließlich statistisches Multiplexing, welches eine weitere Ebene der Überbuchung einführt.

Ein weiterer Faktor der Elastizität ist das schnelle Provisionieren und Deprovisionieren von VMs. Neue VMs können dort installiert werden, wo noch freie Kapazitäten vorhanden sind

⁸⁴ Vgl. Breitgand, David; Dubitzky, Zvi; Epstein, Amir; Glikson, Alex; Shapira, Inbar (2012), S. 1

⁸⁵ Vgl. Song, Weijia; Xiao, Zhen (2013), S. 315

⁸⁶ Vgl. Breitgand, David; Dubitzky, Zvi; Epstein, Amir; Glikson, Alex; Shapira, Inbar (2012), S. 73

⁸⁷ Entnommen aus Meng, Xiaoqiao; Isci, Canturk; Kephart, Jeffrey O.; Zhang, Li; Bouillet, Eric; Pendarakis, Dimitrios E. (2010), S. 3

(VM Placement) oder der Anbieter fasst kleinere freie Bereiche durch Verschiebungen in der Cloud zusammen (VM Migration).

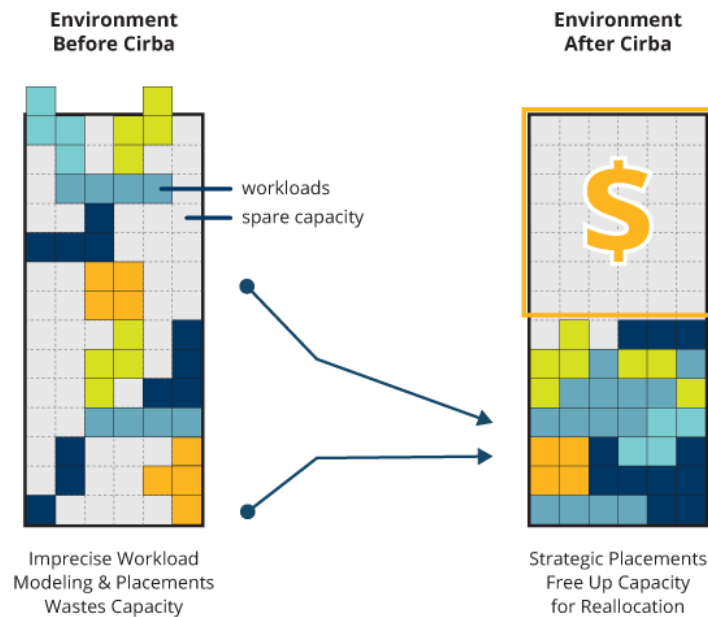


Abbildung 7 – Effekt der VM Consolidation⁸⁸

VM Placement und VM Migration werden unter dem Thema VM Consolidation zusammengefasst⁸⁹, deren Auswirkung in Abbildung 7 zu sehen ist.

VM Consolidation fügt eine weitere Ebene von Komplexität hinzu, umso wichtiger ist es für den Anbieter, ein effizientes Management zu etablieren. Es muss die richtige Balance zwischen vorgehaltenen und tatsächlich benötigten Ressourcen gefunden werden. Galante et al.⁹⁰ verweist darauf, dass Anbieter hier Probleme haben. Sie reagieren darauf, indem sie den Kunden Regeln auferlegen mit denen sie nur in begrenzten Rahmen neue Ressourcen hinzuzufügen⁹¹.

Eine allgemein gültige Methode zur Messung und Bewertung der Elastizität hat sich bisher nicht herausgebildet. Andreas Wolf beschreibt in seiner Master Thesis⁹² aus dem Jahr 2014 den Stand der Forschung und kommt zum Schluss, dass alle bisherigen Ansätze das Thema nur begrenzt abdecken⁹³ und alle ihre spezifischen Probleme haben. Seine Arbeit versucht, die bisherigen Ansätze weiter zu führen, und ist trotz Evaluation durch ein von ihm

⁸⁸ Entnommen aus Cirba (2014)

⁸⁹ Vgl. Meng, Xiaoqiao; Isci, Canturk; Kephart, Jeffrey O.; Zhang, Li; Bouillet, Eric; Pendarakis, Dimitrios E. (2010), S. 11

⁹⁰ Vgl. Galante, Guilherme; Bona, Luis Carlos Erpen De (2012)

⁹¹ Vgl. ebd., S. 268

⁹² Vgl. Weber, Andreas (2014), S. 19-22

⁹³ Vgl. ebd., S. 23

entwickeltes Programmier-Framework von theoretischer Natur, da eine Anwendung für Anbieter nicht ohne eigenen Entwicklungsaufwand möglich ist. Gleiches gilt für die von ihm besprochenen vorhergehenden Ansätze.

Da kein gültiger und praktisch anwendbarer Ansatz für die Bewertung von Elastizität im Allgemeinen und statistischen Multiplexing im speziellen erkennbar ist, bleibt lediglich die Ermittlung und Deutung eigener Daten, worauf schon das Wort ‚Statistik‘ im Namen der Methode ‚statistisches Multiplexing‘ hinweist. Diese Daten müssen auf drei Zustände hin ausgewertet werden:

- Unterprovisioniert – Die Summe aller provisionierten Ressourcen ist kleiner, als die zur Verfügung stehenden physikalischen Ressourcen.
- Optimaler Betriebsmodus – Es sind mehr Ressourcen provisioniert als physikalisch vorhanden und es gibt keine Performance-Einbußen.
- Überprovisioniert – Der Punkt, an dem Performance-Einbußen durch Überlast entstehen.

Über die Dauer der Zeit gilt es, den Punkt des optimalen Betriebsmodus zu finden, auch über Kapazitätsveränderungen hinweg und immer möglichst nah am Status ‚Überprovisioniert‘ zu halten.

Das Thema der schnellen Provisionierung und Deprovisionierung lässt sich über ein Ratio darstellen:

$$\text{Stress einer Cloud} = \frac{\text{Angeforderte Kapazität}}{\text{Zugewiesener Kapazität}}$$

Shawky und Ali benennen es schlicht Stress einer Cloud⁹⁴, denn die angeforderte Kapazität in zugewiesene Kapazität umzuwandeln bedeutet intern Aktionen auszulösen, die den Zustand der Cloud verändern. Ist das Ratio positiv, dehnt sich die Cloud, ist sie negativ, zieht sich wieder zusammen, um auf die Initialen Definitionen zurückzukommen. Der Stress einer Cloud sollte dabei möglichst nahe Null gehalten werden.

4. Kaufmännische Einflussgrößen

Die Betriebswirtschaftslehre (BWL) ist, als Teil der Wirtschaftswissenschaften, ein umfangreiches Themengebiet. Neben der allgemeinen BWL, die branchen- und

⁹⁴ Vgl. Shawky, Doaa M.; Ali, Ahmed F. (2012), S. 2

funktionsübergreifend ausgerichtet ist, gibt es mit der speziellen BWL weitere Teilbereiche, die sich mit wirtschaftlichen, organisatorischen, technischen sowie finanziellen Abläufen in Unternehmen beschäftigen⁹⁵.

Das folgende Kapitel behandelt die Positionierung im Wettbewerb, die Bewertung von Investitions- sowie Betriebskosten und das Kostenmanagement während der Betriebsphase.

4.1 Geschäftsmodell und Wettbewerbsstrategie

Auf Seite 17 im technischen Teil ist bereits das Wort Geschäftsmodell angesprochen worden. Nach Bieger ist ein Geschäftsmodell „eine Darstellung der Art und Weise, wie ein Unternehmen (...) am Markt Werte schafft“⁹⁶, also eine vereinfachte Abbildung der Realität, das Geschäftsmodell kann auch als „geistige Landkarte“ zur Orientierung dienen⁹⁷.

Eine Wettbewerbsstrategie hingegen beschreibt die Positionierung im Markt gegenüber dem Wettbewerb und wird durch die Branchenattraktivität und die Wettbewerbsposition bestimmt, sie zielt darauf ab, eine profitable Position im Markt durch Wettbewerbsvorteile zu erreichen.

Erfolgsstrategien	Problemlösung	Preisführerschaft
Bearbeitung großer Märkte / Segmente	Differenzierung im engeren Sinn (Problemlösung in großen Segmenten mittels eines USP)	Kosten- und Preisführerschaft
Bearbeitung gezielter Marktnischen / kleine Segmente	Problemlöser in der Nische	Preisführer in der Nische

Abbildung 8 – Erweiterte Positionierungsoptionen nach Porter⁹⁸

M.E. Porter hat über die Jahre sein Modell der Positionsoptionen ständig weiterentwickelt, so dass sich heute vier Optionen (siehe Abbildung 8) für eine Wettbewerbsstrategie ergeben⁹⁹:

- Eine *Differenzierung im engeren Sinn* wäre zum Beispiel Qualität. In der IT-Branche steht Apple für eine Wettbewerbsstrategie, die einen Massenmarkt beliefert, sich aber

⁹⁵ Vgl. Husmann, Sven (2012)

⁹⁶ Vgl. Belz, Christian; Bieger, Thomas (2006), S. 397

⁹⁷ Vgl. Umbeck, Tobias (2009), S. 51

⁹⁸ Entnommen aus Mussnig, Werner (2013), S. 235

⁹⁹ Vgl. Porter, Michael E. (2013)

aufgrund des Qualitätsanspruchs von den Wettbewerbern unterscheidet. Bei den Cloud-Anbietern haben Firmen wie HP, IBM oder Oracle einen ähnlichen Anspruch.

- Der *Problemlöser in der Nische* hat ein branchenspezifisches Angebot wie die Firma Datev mit Software für Steuerberater und Wirtschaftsprüfer, sei es nun im Hause oder in der Cloud.
- Der Fokus bei Hardware-Herstellern wie Lenovo, Acer oder Medion liegt klar auf dem Massengeschäft und damit bei *Kosten- und Preisführerschaft*. Hier ist Amazon oder Google für den Bereich des Cloud Computing zu nennen.
- Einen *Preisführer in der Nische* der Heim-PCs ohne Betriebssystem ist nicht so einfach auszumachen, allerdings dürften Firmen wie Notebookbilliger oder Snogard dies sicher anstreben. Cloud-Anbieter wie DigitalOcean oder CloudSigma konkurrieren zwar untereinander über den Preis, mit den Massen-Anbietern allerdings nicht. Hier differenzieren sich die Anbieter über Features, die nicht so oft nachgefragt werden, dass sie sich für ein Massengeschäft lohnen würden.

Geschäftsmodell und die Wettbewerbsstrategie bauen aufeinander auf. Während das Geschäftsmodell essentielle Prinzipien (siehe Seite 17) definiert, lassen sich aus der Wettbewerbsstrategie Erkenntnisse über Wettbewerber und damit deren Produkte und Preismodelle ableiten. Hier raus lässt sich die eigene Positionierung im Markt ableiten, Potential für die Entwicklung eigener Produkte identifizieren und die sich daraus ergebenden Wettbewerbsvorteile entwickeln. Ob dies über eine Kostenführerschaft oder durch Differenzierung vom Wettbewerb geschieht, ist eine der Abwägungen, die zu treffen sind. Letztlich geht es darum, vom Kunden wahrgenommen zu werden¹⁰⁰.

Das Geschäftsmodell als auch die Wettbewerbsstrategie bilden die Grundlage allen Handelns. Für die Einführung einer Cloud-Lösung am Markt ist es damit wichtig, beides im Strategie-Prozess festzulegen. An den Themen High-Availability (siehe Kapitel 3.5) und auch Elastizität (siehe Kapitel 3.6) ist zu sehen, dass an diesen Stellen Weichen gestellt werden, die bis in die technischen Themen ausstrahlen. Denn ein qualitätsorientierter Anbieter wird sich technisch anders aufstellen, als ein Anbieter, der auf möglichst günstige Preise setzt und damit einen Massenmarkt anspricht.

¹⁰⁰ Vgl. Kirchgeorg, Manfred (2014)

4.2 Return on Invest

Bevor es mit dem eigentlichen Thema los geht, gilt es zwei Begrifflichkeiten zu klären. Da wären zum einen die *capital expenditures* (CAPEX), mit denen generell Investitionsausgaben gemeint sind, und zum anderen die *operational expenditures* (OPEX). Sie repräsentieren die Betriebskosten, also Miete, Personalkosten, Schulung, Serviceverträge etc¹⁰¹.

Beide Werte sind Bestandteil der Berechnung des Return on Investment (ROI), der den Quotienten aus erwirtschafteten Gewinn und dem dafür eingesetztem Kapital angibt. Der ROI kann aber auch generischer ausgedrückt werden, indem eine Ergebnisgröße auf das eingesetzte Gesamtkapital bezogen wird¹⁰².

Der ROI für eine Cloud ergibt sich damit so:

$$\frac{CAPEX + OPEX + \text{Kosten für Downtimes}}{\text{Anschaffungskosten}}$$

Die Anschaffungskosten bestehen aus Kosten für neue Server, Storage, Software Lizenzen & Support sowie Schulungskosten für Mitarbeiter¹⁰³. Die Kosten für Downtimes werden angesetzt, weil es sich hier um Opportunitätskosten handelt, also entgangener Gewinn¹⁰⁴.

Ein positiver ROI ist demnach anzustreben. Es gilt allerdings zu bedenken, dass ein positiver ROI keine Aussagen über die Risiken einer Investition macht, daher sollte er nie alleiniger Faktor für eine Investitionsentscheidung sein¹⁰⁵.

4.3 Total Cost of Ownership

Eine Investition verursacht direkte und indirekte Kosten. Während direkte Kosten im Rechnungswesen einer Firma sichtbar und damit beeinflussbar sind, handelt es sich bei den indirekten Kosten um Produktivitätsverluste, die im Rahmen des Betriebes anfallen (siehe Abbildung 9).

¹⁰¹ Vgl. Bachmann, Werner; Tiemeyer, Ernst (2013), S. 408

¹⁰² Vgl. Wöhe, Günter; Döring, Ulrich (2013), S. 863

¹⁰³ Vgl. Chang, Bao Rong; Tsai, Hsiu-Fen; Chen, Chi-Ming (2013), S. 195

¹⁰⁴ Vgl. Wöhe, Günter; Döring, Ulrich (2013), S. 893

¹⁰⁵ Vgl. Ebel, Nadin (2009), S. 156

Direkte Kosten (im Rechnungswesen sichtbar)	Indirekte Kosten (im Rechnungswesen unsichtbar)
<ul style="list-style-type: none"> • Hardware (Anschaffung, Leasing) • Software (Lizenzen, Updates) • IT-Infrastruktur (Netzwerk, Telefongebühren) 	<ul style="list-style-type: none"> • Versteckte dienstliche Endbenutzer-Kosten (Arbeitszeitverlust durch Kollegenschulung [Hey Joe-Effekt], Trial-and-Error-Schulung)
<ul style="list-style-type: none"> • IT-Entwicklung von Firmen-Add Ons (z. B. Schriftarten, Makros für Geschäftsbriefe, Funktionstest, Anwenderdokumentation) • Schulung und Support (Grundlagenkurse, Telefonhotline, Individualtraining) 	<ul style="list-style-type: none"> • Produktivitäts- und Arbeitszeitverluste durch technische Probleme (Zusammenbruch des Netzwerks, nicht nutzbarer Endbenutzerarbeitsplatz, Druckerprobleme, Serverausfall etc.)
<ul style="list-style-type: none"> • Verwaltung- und Wartung (Eigene Mitarbeiter, Fremdfirmen bei Outsourcing) 	<ul style="list-style-type: none"> • Versteckte private Endbenutzerkosten (Arbeitszeitverlust durch private Internetnutzung, sog. Futzing)

Abbildung 9 – Direkte und indirekte Kosten nach Gadatsch¹⁰⁶

Diese Kosten hingegen sind nicht transparent. Durch diese Unsichtbarkeit können sie nicht beeinflusst werden und dabei ein nicht unerhebliches Ausmaß annehmen (45% transparent vs. 55% intransparent)¹⁰⁷.

Mit der TCO-Analyse (Total Cost of Ownership) geht der Versuch einher, möglichst umfangreich alle Kosten zu identifizieren und quantifizieren, die für den Lebenszyklus einer Cloud-Lösung eine Rolle spielen. Die TCO-Analyse nach Gadatsch „umfasst nicht nur die Kosten der Anschaffung, der Installation der Hard- und Software, die Wartung und den Betrieb, sondern auch die Anschaffung und Wartung von Servern und Netzwerken, Benutzersupport, Schulung und Training, Entwicklung spezieller Anwendungen sowie die Kosten für den Systemausfall“¹⁰⁸. Mit diesem Ansatz werden vorher unbekannte Kosten sichtbar und auch steuerbar gemacht.

In Li et al.¹⁰⁹ ist eine Auflistung aller zu berücksichtigten Komponenten sowie die Berechnung der Kosten zu finden. Zudem erweitert sie das TCO-Modell um den Punkt Utilization Cost¹¹⁰ (Nutzungskosten). Sie beschreibt diese Kosten als Teilmenge der TCO und verrechnet dabei die Parameter VM Density (siehe Kapitel 3.1), Anzahl der phys. Server, sowie Anzahl der benötigten Server-Racks, um eine Reihe von Kennzahlen zu erhalten, die näheren Aufschluss darüber geben, welche Nutzungskosten in verschiedenen Situationen entstehen.

¹⁰⁶ Entnommen aus Gadatsch, Andreas; Mayer, Elmar (2010), S. 111

¹⁰⁷ Vgl. Gadatsch, Andreas; Mayer, Elmar (2010), S. 108-111

¹⁰⁸ Vgl. ebd., S. 111

¹⁰⁹ Vgl. Li, Xinhui; Li, Ying; Liu, Tiancheng; Qiu, Jie; Wang, Fengchun (2009)

¹¹⁰ Vgl. ebd., S. 1

Ein ähnliches Konzept wählt Iglesias et al.¹¹¹. Er setzt die Nutzungskosten in Relation zu den erzielbaren Kapazitäten eines Virtualisierungs-Hosts und ist damit in der Lage, die Nutzungskosten einer bestimmten VM auf einem bestimmten Virtualisierungs-Host an einen bestimmten Standort zu ermitteln. Durch diese Transparenz ist es möglich, die VM dort zu betreiben, wo die niedrigsten Nutzungskosten anfallen.

Allen TCO-Analysen ist gemein, dass sie keine Aussagen über den Nutzen einer Investition treffen, sondern nur deren Kosten beleuchten. Daher sollte die TCO-Kennziffer nicht alleiniger Grund für eine Investitionsentscheidung sein, sondern noch andere Kennziffern herangezogen werden¹¹², wie z.B. den ROI. Andererseits ist eine gründliche TCO-Analyse notwendig, weil sie die Vollkosten der geplanten Investition erhebt¹¹³ und damit eine Grundlage für die Preiskalkulation (siehe Kapitel 5.2) bildet.

4.4 Kostenmanagement

Wenn die Investitionsentscheidung getroffen und die Cloud-Lösung sich in Betrieb befindet, beginnt eine Phase, in der sich der Anbieter intensiv mit den nun entstehenden Kosten auseinandersetzen muss.

Fixkosten ist ein sehr wichtiger Punkt, weil sie die Kosten der Betriebsbereitschaft darstellen. Im Bereich der produzierenden IT ist dies der größte Kostenblock¹¹⁴, weil sie unabhängig von der Ausbringungsmenge sind, denn bevor auch nur ein Bit verarbeitet werden kann, muss die entsprechende Ausrüstung angeschafft, Personal und Räumlichkeit zur Verfügung gestellt werden usw.¹¹⁵. Dieser Fixkostensockel ist in der Regel mittel- bis langfristig gebunden und kann nicht ohne weiteres einer schwankenden Nachfrage angepasst werden. Ein Mittel der Steuerung ist die Aufspaltung in Nutz-, Leer- und Remanenzkosten. Wobei Nutzkosten den Anteil der tatsächlich genutzten Kapazität einnimmt, während Leerkosten den anderen Teil darstellen. Remanenzkosten beschreiben den Nutzungsrückgang, also den Übergang von Nutz- zu Leerkosten¹¹⁶.

Der Anteil der variablen Kosten bei einer Cloud-Lösung hingegen ist recht gering. Bei einer Abrechnung nach Volumina sind Kosten für Strom und der Datentransfer ins Internet

¹¹¹ Vgl. Iglesias, Jesus Omana; Perry, Philip; Stokes, Nicola; Thorburn, James; Murphy, Liam (2013)

¹¹² Vgl. Hill, Richard; Hirsch, Laurie; Lake, Peter; Moshiri, Siavash (2012), S. 194

¹¹³ Vgl. Brugger, Ralph (2009), S. 263

¹¹⁴ Vgl. Kütz, Martin (2010), S. 6-7

¹¹⁵ Vgl. Wöhe, Günter; Döring, Ulrich (2013), S. 298-299

¹¹⁶ Vgl. Kütz, Martin (2010), S. 7

abhängig von der Ausbringungsmenge. Die Kostenhöhe ist Produkt der Menge und des Preises. Bei variablen Kosten bestimmt die Nachfrage die Menge, womit der Beschaffungspreis als Steuerungskomponente übrigbleibt¹¹⁷.

Die Steuerungsmöglichkeiten im Kostenmanagement liegen, wie zu sehen ist, hauptsächlich im Bereich der Fixkosten und dort bei den Leer- bzw. Remanenzkosten. Sie sind im Sinne des ökonomischen Prinzips (siehe Seite 10) zu reduzieren, aber nie ganz zurückzuführen, da Reserven für plötzliche Kapazitätsschwankungen notwendig sind.

5. Der Preis und seine Ausprägungen

Das Ziel eines jeden Unternehmens ist die Gewinnerzielung und das nach außen sichtbarste Zeichen hierfür ist der Preis des Produktes. Nach Wöhe¹¹⁸ steht für ein Unternehmen die Suche nach dem gewinnmaximalen Preis im Zentrum der Preispolitik. Es gilt dabei den richtigen Preisansatz zu finden:

- Ist er zu niedrig, verzichtet der Anbieter auf einen Teil der Gewinnmarge.
- Ist er zu hoch, wandern die Nachfrager in Scharen zur Konkurrenz ab.

Der Weg führt dabei über die Preisstrategie und den Preisfindungsmodellen.

5.1 Preisstrategie

Ein erster Schritt zur Preisfindung greift auf das Geschäftsmodell und die Wettbewerbsstrategie (siehe Kapitel 4.1) zurück. Die Positionsoptionen im Wettbewerb bilden die Grundlage der strategischen Preispolitik. Ein Qualitätsanbieter wird eine Prämienpreis-Strategie anstreben, genauso wie im Massenmarkt eher Promotionspreis-Strategien anzutreffen sind (siehe im unteren Abschnitt der Abbildung 10).

¹¹⁷ Vgl. Gadatsch, Andreas; Mayer, Elmar (2010), S. 195

¹¹⁸ Vgl. Wöhe, Günter; Döring, Ulrich (2013), S. 430

Markteinführung	
Penetrationspreis	Abschöpfungspreis
Mit einem niedrigem Preis in den Markt gehen, um schnell einen hohen Bekanntheitsgrad zu erzielen. Später den Preis erhöhen.	Mit hohem Preis in den Markt gehen, um die Kaufkraft der Gutverdiener abzuschöpfen. Später nach und nach den Preis senken.
Qualität	
Prämienpreis	Promotionspreis
Ein hoher Preis signalisiert dem Käufer eine hohe Qualität.	Ein niedriger Preis signalisiert „günstig“, Qualität muss stimmen.

Abbildung 10 – Preisstrategien¹¹⁹

In der Phase der Markteinführung gelten hingegen andere Regeln (siehe im oberen Abschnitt der Abbildung 10). Für den etablierten Markt empfiehlt sich der Penetrationspreis, eröffnet der Anbieter mit einer Produktinnovation einen neuen Markt, ist die Abschöpfungspreis-Strategie zu wählen¹²⁰.

5.2 Traditionelle Preisfindung

Nach Meffert¹²¹ gibt es drei wesentliche Methoden, den gewinnmaximalen Preis zu ermitteln:

Bei kostenorientierter Preisfindung ermittelt der Anbieter den Preis als Summe aus Selbstkosten und Gewinnzuschlag¹²². Die Selbstkosten ergeben sich dabei als Summe aller durch den betrieblichen Leistungsprozess entstandenen Kosten¹²³, sie werden auch Vollkosten genannt und können im Rahmen einer TCO-Analyse (siehe Kapitel 4.3) erhoben werden. Die Art dieser Preisfindung ist allerdings nicht praxisrelevant, da nur in Planwirtschaften der Preis von den Selbstkosten abgeleitet wird. Im marktwirtschaftlichen Umfeld hingegen bestimmt Angebot und Nachfrage den Preis. Die Selbstkosten können aber als Preisuntergrenze herangezogen werden. Sie beschreiben den Punkt, an dem weder Gewinn noch Verlust gemacht wird¹²⁴.

Die Kundenbefragung steht im Mittelpunkt der nachfrageorientierten Preisfindung. Über geeignete Fragemodelle wird ermittelt, in welchem Korridor sich der Preis für ein Produkt

¹¹⁹ Entnommen aus Wünsche, Manfred (2010), S. 126

¹²⁰ Vgl. Wöhe, Günter; Döring, Ulrich (2013), S. 432-433

¹²¹ Vgl. Meffert, Heribert; Burmann, Christoph; Kirchgeorg, Manfred (2012), S. 513ff

¹²² Vgl. Wöhe, Günter; Döring, Ulrich (2013), S. 434

¹²³ Vgl. ebd., S. 921 Abb. 269

¹²⁴ Vgl. ebd., S. 434

bewegen müsste, den der Kunde bereit wäre zu zahlen¹²⁵. Orientiert sich der Anbieter beim Preis an Wettbewerbsprodukte, wird von konkurrenzorientierter Preisfindung gesprochen¹²⁶.

5.3 Target Costing

Die bisher behandelten Themen wie Preisstrategie und Preisfindung sind in ihrer Aussage abstrakt und ein konkreter Preis für ein Produkt lässt sich mit ihnen nicht ermitteln. Hierfür werden normalerweise Kostenrechnungssysteme eingesetzt, die aus den fünfziger und sechziger Jahren des letzten Jahrhunderts stammen und damit auf produzierendes Gewerbe der traditionellen Art ausgerichtet sind.

Das Target Costing setzt auf frühzeitige Kostenbeeinflussung, weil die meisten Unternehmen in Käufermärkten mit hoher Wettbewerbsintensität unterwegs sind. Die Fragestellung wird dabei von

- „Was wird ein Produkt kosten?“, hin zur Frage
- „Was darf ein Produkt kosten?“

umgedreht¹²⁷. Es gehört damit zu den nachfrageorientierten Methoden der Preisfindung (siehe Kapitel 5.2).

Eine Quellenlage zu Target Costing in Bezug auf Cloud Computing ließ sich nicht ermitteln, allerdings lässt sich das Modell anhand eines Beispiels mit fiktiven Zahlen erklären:

Für den Preis einer VM einer IaaS Cloud hat das Marketing des Anbieters, im Rahmen einer Kundebefragung, einen Zielpreis von 50€ pro Monat festgestellt und, da sich der Anbieter auf eine Umsatzrentabilität von 15% festgelegt hat, ergibt sich folgende Rechnung¹²⁸, die in der Tabelle 2 dargestellt ist.

Zielpreis für die VM	50,00€
- geplante Umsatzrendite	7,50€
= „Allowable Costs“	42,50€

Tabelle 2 – Berechnung der „Allowable Costs“

¹²⁵ Vgl. Wöhe, Günter; Döring, Ulrich (2013), S. 435-436

¹²⁶ Vgl. ebd., S. 436-437

¹²⁷ Vgl. ebd., S. 963

¹²⁸ Angelehnt an Strautenrauch, Thomas (2014)

Das Ergebnis sind die *Allowable Costs*, also die Kosten, welche für die Herstellung des Produktes entstehen dürfen. Die Kundenbefragung des Anbieters hat aber auch die Bedeutung der einzelnen Produkteigenschaften ermittelt, wie die Tabelle 3 zeigt.

Produkteigenschaft	Bedeutung (in Prozent)
Preis	30%
Performanz	30%
Datensicherheit	25%
Funktionsumfang	15%

Tabelle 3 – Bedeutung von Produkteigenschaften für Kunden

Im Rahmen einer Abfrage hat das Produktmanagement des Anbieters die Herstellkosten zur Verfügung gestellt, die sogenannten *Drifting Costs*, welche in Tabelle 4 dargestellt werden.

Komponentengruppe	Standard-Herstellkosten
CPU	12,00€
Arbeitsspeicher	14,00€
Festplattenplatz	10,00€
Netzwerk	14,00€
Summe	50,00€

Tabelle 4 – Auflistung der Standard-Herstellkosten

Aus der Produktentwicklung des Anbieters kommt die Bewertung der Abhängigkeit der Produkteigenschaften von den Komponentengruppen:

Produkteigenschaft	Preis	Performanz	Datensicherheit	Funktionsumfang
CPU	30%	70%	5%	5%
Arbeitsspeicher	20%	5%	10%	5%
Festplattenplatz	20%	5%	80%	70%
Netzwerk	30%	20%	5%	20%

Tabelle 5 – Bewertung der Produkteigenschaften

Für die Aufteilung der *Allowable Costs* (siehe Tabelle 2) auf die Komponentengruppen werden das Bedeutungsmaß der Komponentengruppen mit dem der Produkteigenschaften multipliziert, woraus die Nutzenanteile je Komponentengruppe ergeben:

Produkteigenschaft	Preis	Performanz	Datensicherheit	Funktionsumfang	Nutzenanteil
CPU	9% (=0,3*0,3)	21%	1,5%	1,5%	33%
Arbeitsspeicher	3%	1,5%	3%	1,5%	12%
Festplattenplatz	5%	1,25%	20%	17,5%	43,75%
Netzwerk	4,5%	3%	0,75%	3%	11,25%
Nutzenanteil der Eigenschaft	24,5%	26,75%	25,25%	23,50%	100%

Tabelle 6 – Aufteilung der Allowable Costs auf die Komponentengruppen

Im letzten Schritt werden die im vorherigen Schritt ermittelten Nutzenanteile mit den *Allowable Costs* (siehe Tabelle 2) multipliziert und damit der Korrekturbedarf ermittelt:

Komponentengruppe	Standard-Herstellkosten	Zielkosten	Korrekturbedarf
CPU	12,00€	14,03€ (=33%*42,5€)	-2,03€
Arbeitsspeicher	14,00€	5,10€	8,90€
Festplattenplatz	10,00€	18,59€	-8,59€
Netzwerk	14,00€	4,78€	9,22€
Summe	50,00€	42,50€	7,50€

Tabelle 7 – Ermittlung des Korrekturbedarfs

Wie ersichtlich, ergeben sich aufgrund des Target Costing zwei Anpassungsbereiche: bei CPU und Festplattenplatz ist Spielraum nach oben und Arbeitsspeicher sowie Netzwerk müssen günstiger werden.

Durch eine Umkehrung der Fragestellung lässt sich mit Hilfe des Target Costing der Preis eines Produktes ermitteln. Dabei steht der Kunde mit seiner Preisvorstellung und Erwartungen an die Produktqualität im Mittelpunkt. Das Target Costing liefert somit konkrete Zahlen, wie die Technik ausgestaltet werden muss, um dies sicher zu stellen.

5.4 Der Preis für die Cloud

Ein Überblick der Preisfindung und anzufindenden Preismodelle in Cloud-Umgebungen geben die Metanalysen von Al-Roomi et al.¹²⁹ aus 2013 sowie Lampe et al.¹³⁰ aus 2014.

Al-Roomi et al. identifiziert drei dabei grundlegende Preismodelle mit folgenden Attributen:

Preismodell	Attribut(e)
Fixed	Preis bleibt im Verlauf der Zeit gleich
Dynamic	Preis ändert sich Aufgrund von gewählten Features, Kundenwünschen und/oder Abnahmemengen
Market-Dependent	Reagiert in Echtzeit auf Marktverhältnisse mit Angebotspreisen und Auktionen. Auch Kundennachfrage und Ertragsmanagement ¹³¹ können hier Einfluss nehmen.

Tabelle 8 – Preismodelle nach Al-Roomi (eigene Darstellung)

Lampe et al. hingegen verweist auf insgesamt sechs Preismodelle in zwei Kategorien, die sich allerdings nach der Logik von Al-Roomi et al. in die Preismodelle der Tabelle 8 einsortieren lassen. Beide bieten einen Vergleich der besprochenen Preismodelle sowie die Beschreibung der jeweiligen Vor- und Nachteile. Al-Roomi et al. stuft die identifizierten Preismodelle zusätzlich in die Kategorien „allgemein umgesetzt“, „umgesetzt“ und „theoretischer Ansatz mit Simulation“ ein, im Gegensatz dazu weist Lampe et al. auf die Herausforderungen hin, die ein Preismodell an den Anbieter stellt. Einig sind sich beide, dass das Pay-as-you-go (PAYG) Preismodell das am häufigsten verwendete ist.

PAYG gehört zu den Fixed-Preismodellen, in denen der Preis über die Zeit der Nutzung immer gleich bleibt. Al-Roomi et al. beschreibt es folgendermaßen: „Customers pay a fixed price per unit of use. Amazon, considered the market leader in cloud computing, utilizes such a model by charging a fixed price for each hour of virtual machine usage“¹³². Lampe et al. schreibt diesem Modell eine perfekte Übereinstimmung mit dem Kundenerwartungen an die Elastizität einer Cloud zu. Auch ist dieses Preismodell für Kunden einfach zu verstehen. Andererseits erfordert es eine umfangreiche Analyse der Nutzung und es ist schwierig, den

¹²⁹ Vgl. Al-Roomi, May; Al-Ebrahim, Shaikha; Buqrais, Sabika; Ahmad, Imtiaz (2013), S. 93-106

¹³⁰ Vgl. Lampe, Ulrich; Hans, Ronny; Seliger, Marco; Pauly, Michael; Schiefer, Michael (2014), S. 1-10

¹³¹ Konzept zur Steuerung des Gewinns eines Unternehmens

¹³² Vgl. Al-Roomi, May; Al-Ebrahim, Shaikha; Buqrais, Sabika; Ahmad, Imtiaz (2013), S. 98-99

Geldfluss vorherzusagen, da die Nutzung durch den Kunden schwankt und damit auch die Einnahmen des Anbieters.

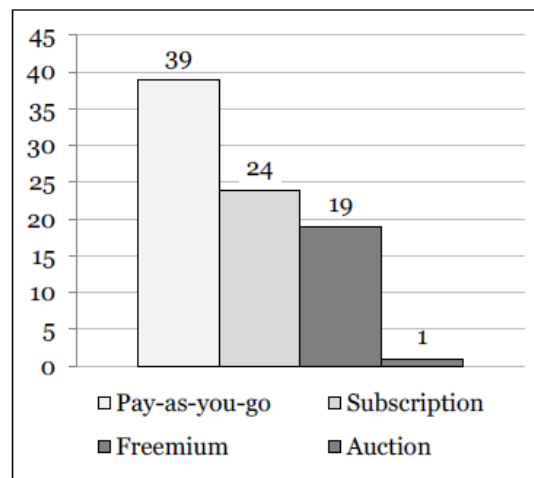


Abbildung 11 – Häufigkeit von Preismodellen¹³³

Am Zweithäufigsten nach Al-Roomi et al. vertreten (siehe Abbildung 11) ist das Subskriptions-Modell. Der Kunde abonniert in diesem Fall die von ihm benötigten Ressourcen für einen festen Preis über eine längere Zeit. Es gehört damit ebenfalls zu den Fixed-Preismodellen und macht dem Anbieter die Voraussage von Nutzungsgrad und Geldfluss einfach. Ebenfalls ist dieses Modell für Kunden einfach zu verstehen und sie wechseln nicht so häufig, da sie für längere Zeit gebunden sind.

Beim Freemium-Modell wird das Basisprodukt gratis angeboten, während das Vollprodukt kostenpflichtig ist. Ein Beispiel für dieses Fixed-Preismodell sind die kostenlosen Kontingente von Amazon¹³⁴. Dieses Modell kann Vorteile bei der Bewerbung der Produkte haben, da der Kunde anfangs kostenlos Erfahrungen mit dem Angebot sammeln kann. Auch ist der schnelle Aufbau einer Kundenbasis beim Markteintritt möglich. Negativ zu bewerten ist, dass der Wechsel von der kostenlosen zur kostenpflichtigen Produktversion und damit auch die Einnahmen ungewiss sind. Als besonders komplex wird hervorgehoben, ein angemessenes Produkt-Design zu schaffen, den Kunden einen Anreiz zum Wechsel auf das Vollprodukt zu geben.

Neben den Beschriebenen gibt es noch eine Vielzahl von Preismodellen und Variationen von eben diesen. Sie sind allerdings wenig verbreitet, wie die Abbildung 11 zeigt, oder von rein theoretischer Natur, es würde an dieser Stelle zu weit führen, sie alle zu beschreiben.

¹³³ Entnommen aus Lampe, Ulrich; Hans, Ronny; Seliger, Marco; Pauly, Michael; Schiefer, Michael (2014), S. 7

¹³⁴ Vgl. AmazonFree (2014)

Grundsätzlich lässt sich aber sagen, dass Kunden Preismodelle bevorzugen, die ihnen eine gewisse Kostensicherheit geben. Spot-Modelle haben für den Kunden den Reiz, ein Schnäppchen zu erwischen, allerdings sind sie in der Laufzeit begrenzt und der Kunde weiß nicht, wie der nächste Preis ausfällt. Die Fähigkeit, mit dieser Dynamik umgehen zu können, muss beim Kunden gegeben sein, sie fehlt aber in vielen Unternehmen¹³⁵.

Auktionsbasierte Preismodelle spielen derzeit nach Lampe keine Rolle, lediglich Amazon bietet derartiges an. Das Preismodell an sich ist recht komplex, sowohl was die Umsetzung beim Anbieter angeht als auch für den Kunden zu verstehen¹³⁶. Abhishek et al.¹³⁷ merkt aber richtigerweise an, dass Fixed-Preismodelle schon aus statistischen Gründen zu ungenutzten Ressourcen führen werden. Der Anbieter kann dies hinnehmen oder versuchen, die hier entstehenden Leerkosten (siehe Kapitel 4.4) über ein Auktions-Modell zu minimieren. In ihrer Studie gehen sie der Frage nach, ob sich der Unterhalt eines PAYG- und eines Auktionsbasierten Preismodelles in höheren Erträgen niederschlägt. Das Ergebnis, ein Fixed-Price Modell mit gut ausgesuchten Preis ist einem hybriden Modell überlegen, hilft nicht weiter, da ein monopolistischer Markt als Grundlage gewählt wurde. Gohad et al.¹³⁸ weisen für ihr Preismodell bis zu 23% höhere Einnahmen nach (siehe Abbildung 12). Für eine Implementierung fehlt allerdings eine nachvollziehbare Beschreibung der Kriterien für Entscheidungen, die innerhalb des Modells getroffen werden.

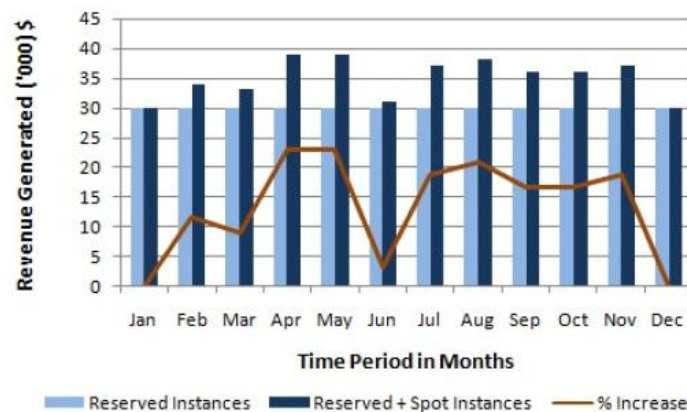


Abbildung 12 – Höhere Einnahmen durch hybrides Preismodell¹³⁹

¹³⁵ Vgl. Technoconsult (2013)

¹³⁶ Vgl. Lampe, Ulrich; Hans, Ronny; Seliger, Marco; Pauly, Michael; Schiefer, Michael (2014), S. 4

¹³⁷ Vgl. Abhishek, Vineet; Kash, Ian A.; Key, Peter (2012), S. 1

¹³⁸ Vgl. Gohad, Atul; Narendra, Nanjangud C.; Ramachandran, Parathasarthy (2013)

¹³⁹ Entnommen aus Gohad, Atul; Narendra, Nanjangud C.; Ramachandran, Parathasarthy (2013), S. 7

Es lässt aber den Schluss zu, dass ein hybrides Preismodell günstig für den Anbieter sein kann. Ein Fixed-Preismodell sorgt dabei für die Grundlast und damit eine stabile Einnahmenbasis, während ein auktionenbasiertes Modell zusätzliche Einnahmen generiert.

Zusammenfassend ergibt sich für den Anbieter damit die Erkenntnis, möglichst einfache Preismodelle anzubieten, um die Kunden nicht zu überfordern und gleichzeitig Planungssicherheit bezüglich der Leistung zu geben. Dies gilt ebenfalls für den Anbieter: einfach um die Implementierungskosten klein zu halten und Planungssicherheit mit Blick auf planbare Einnahmen. Aus Sicht des Anbieters erfüllt zuerst das Subskriptions-Modell diese Kriterien. Hier sind Laufzeit, Preise und Leistungsumfang für beide Seiten fest definiert und die Integration ist für den Anbieter trivial, da keine Nutzungsdaten erhoben und verarbeitet werden müssen. Kunden erwarten hingegen, wegen der oft auf Stunden basierenden Abrechnung¹⁴⁰, das PAYG-Modell. Auch wenn es daher für den Anbieter ungünstig ist, gehört es zum Angebot eines jeden Cloud-Anbieters. Auktionenbasierte Modelle spielen am Markt zurzeit keine Rolle, können aber für zusätzliche Einnahmen sorgen.

6. Schlusswort

Die vorangegangenen Seiten haben sich eingehend mit den technischen und kaufmännischen Einflussgrößen, sowie der Preisfindung, einer IaaS Cloud beschäftigt. Daher abschließend eine Zusammenfassung und ein Fazit.

6.1 Zusammenfassung

Über 50 Jahre nach Postulierung der Idee, Rechenkraft wie eine öffentliche Versorgungsdienstleistung zu beziehen, ist Cloud Computing der aktuellste Schritt in diese Richtung. Es benötigte erst einen gewissen technischen wie kaufmännischen Entwicklungsprozess, um die sich hauptsächlich in Privatbesitz befindliche Rechenleistung in Service-Angebote mit rein nutzungsbasierten Preismodellen und definierter Service-Qualität zu wandeln.

Cloud Computing gliedert sich dabei in drei Service-Modelle, die eine unterschiedliche Tiefe der Integration darstellen. IaaS bildet dabei die Grundlage, indem es CPU, Arbeitsspeicher, Festplattenplatz in Form von virtuellen Maschinen zur Verfügung stellt. In der nächsten Stufe PaaS können zusätzliche Software-Komponenten wie Datenbanken oder Applikations-Server

¹⁴⁰ Vgl. Lampe, Ulrich; Hans, Ronny; Seliger, Marco; Pauly, Michael; Schiefer, Michael (2014), S. 7

direkt vom Anbieter genutzt werden. Die höchste Integration bietet SaaS. Hier liegt die komplette Verantwortung für den Betrieb, von der Hardware bis zur Applikation, beim Anbieter.

Um die Kapazitäten im Bereich der technischen Einflussgrößen abzubilden, bedient sich ein Anbieter Kennzahlen, die eine Überbuchung der tatsächlich vorhandenen Komponenten beschreiben. Letztlich ist für CPU, Arbeitsspeicher, Storage und Elastizität ein möglichst hohes Verhältnis zwischen Überbuchungsfaktor und tatsächlich vorhandenen Kapazitäten anzustreben. Elastizität ist dabei ein Sonderfall. Sie stellt einen Überbuchungsfaktor oberhalb der drei anderen Überbuchungsfaktoren dar. Durch die Nutzung von Lastschwankungen und Konsolidierung der sich im laufenden Betrieb befindlichen Ressourcen schafft sie Freiräume, um Kapazitätsänderungen in der Cloud zu ermöglichen. Jedes der vier angesprochenen Komponenten hat dabei seine eigene Bewertungsbasis. Allen gemein ist, dass die Überbuchung nicht zu Performanceproblemen führen darf.

Den Rahmen für die betriebswirtschaftliche Sicht auf die Anschaffung und den Unterhalt einer Cloud-Lösung bilden das Geschäftsmodell und die Wettbewerbsstrategie. Gibt ersteres grobe Leitplanken vor, so sorgt die Wettbewerbsstrategie für die Positionierung im Markt und gegenüber den Marktteilnehmern. Durch Differenzierung von ebendiesen, sei es durch eine besondere Qualität, einen günstigen Preis oder besonderer Produktgestaltung, wird die Wahrnehmung durch den Kunden sichergestellt.

Für die Phase der Investitionsentscheidung steht die Frage nach der Wirtschaftlichkeit im Vordergrund. Das einfache Verhältnis zwischen eingesetzten Geld und erwarteten Gewinn wird über den ROI ausgedrückt. Er spiegelt also die Verzinsung des Kapitals wieder. Geht es um Kosten-Transparenz, kommt der TCO zum Einsatz. Diese Analyse teilt die Anschaffungskosten in direkte, steuerbare Kosten und indirekte, nicht steuerbare Kosten ein. Die Idee dahinter ist, die indirekten Kosten möglichst weit zu minimieren um am Ende nur steuerbare Kosten zu erhalten. Beiden Modellen gemein ist die geringe Aussagekraft, beide zusammen bieten dagegen eine solide Entscheidungsbasis.

Das Kostenmanagement in der Betriebsphase arbeitet hauptsächlich an der Minimierung der Fixkosten, denn diese Kosten der Betriebsbereitschaft entstehen auch dann, wenn kein Kunde eine Leistung nachfragt. Die Nutz-, Leer- und Remanenzkosten beschreiben dabei die unterschiedlichen Zustände der zur Verfügung gestellten Ressourcen. Ziel dabei ist, unter

Berücksichtigung von Lastschwankungen die Leer- und Remanenzkosten möglichst gering zu halten.

Das Aushängeschild eines jeden Unternehmens ist der Preis. Auch hier gilt es zuerst eine Strategie festzulegen, die auf der Wettbewerbsstrategie fußt. Ein Qualitätsanbieter wird eine andere Strategie wählen, als jemand im Massenmarkt. Auch die Frage, ob man sich in einem etablierten Markt befindet oder einen neuen eröffnet, hat seine Auswirkung auf den Preis.

Die Preisfindung der traditionellen Art bietet drei Modelle:

- kostenorientiert,
- nachfrageorientiert und
- konkurrenzorientiert.

Ersteres spielt in der Marktwirtschaft keine Rolle, bietet aber die Möglichkeit, die Preisuntergrenze für ein Produkt zu bestimmen, während die beiden anderen auf Marktsituationen reagieren.

Um zu einer konkreten Aussage bezüglich des Preises zu kommen, bietet sich das nachfrageorientierte Target Costing an. Es geht der Frage nach, was ein Produkt im Markt kosten darf, und arbeitet sich durch Anpassungen am Produktionsprozess schrittweise an die optimale Ausprägung der Komponenten oder Funktionen des Produktes heran. Auch wenn es keine Beispiele in Bezug auf die Anwendung im Cloud Computing gibt, ist diese Methode einfach adaptierbar.

Einteilen lassen sich die vorhandenen Preismodelle in drei Arten: Fixed hält seinen Preis über die Zeit konstant, das dynamische Modell passt sich Kundenwünschen und Abnahmemengen an und das marktorientierte reagiert in Echtzeit auf Angebot und Nachfrage. Der genauere Blick in die Preismodelle zeigt, dass lediglich zwei eine Rolle spielen. Während Kunden das PAYG-Modell bevorzugen, weil es eine stark nutzungsorientierte Abrechnung bietet, ist für Anbieter das Subskriptions-Modell vorteilhaft. Es gibt, im Gegensatz zum PAYG-Modell, keine Integrationskosten und die Einnahmen sind planbar. Auktionsbasierte Modelle spielen bei Anbietern und Kunden derzeit keine Rolle. Allerdings kann ein hybrides Preismodell, also eine Mischung aus Fixed- und Marktorientiert, zu höheren Einnahmen führen.

6.2 Fazit

Die Zielsetzung dieses Dokuments auf Seite 3 spricht davon „wie technische Ressourcen in Hinblick auf Gesamtkapazität und Nutzungsgrad zu beurteilen sind“. Erreicht wird dies für CPU, Arbeitsspeicher, Storage und statistisches Multiplexing über die Definition von Kennzahlen. Sie repräsentieren jeweils einen Überbuchungsfaktor, der das Verhältnis zwischen tatsächlich vorhandener und zur Nutzung freigegebener Kapazität beschreibt und entsprechen damit den Kriterien der Nachvollziehbarkeit, Eindeutigkeit und Unabhängigkeit von Herstellern (siehe Seite 11).

Für das Thema Elastizität gelten diese Kriterien nur begrenzt. Es fehlt an Eindeutigkeit und Nachvollziehbarkeit, weil sich keine allgemein anerkannte Definition herausgebildet hat. Jede Studie hat eine andere Definition, nur die Sicht auf die Lösung der eigenen Problemstellung und bewertet daher anders. Elastizität fügt auch eine neue Dimension der Komplexität hinzu. Gelten für einzelne Virtualisierungs-Hosts die Rahmenparameter der Überbuchung für CPU, Arbeitsspeicher und Storage, so koordiniert das Management-System der Cloud eine Vielzahl von Virtualisierungs-Hosts, um über statistisches Multiplexing eine weitere Überbuchung hinzuzufügen und gleichzeitig über VM Placement das Dehnen und Zusammenziehen der Kapazitäten zu steuern. Eine Beschreibung, wie dieser Vorgang konkret abläuft, welche Kriterien dabei berücksichtigt werden und dies auch noch in allgemeiner Gültigkeit, gibt die Studienlage nicht her.

Während für VM Placement hinreichend Modelle vorliegen, bleibt statistisches Multiplexing ein offenes Thema. Es wird vielfach erwähnt, eine tiefere Recherche führt zu Studien aus dem Bereich der Netzwerktechnik, wo das Thema seinen Ursprung hatte. Eine konkrete Beschreibung und auch Kriterien für die Reaktion auf Systemzustände fehlen auch hier. Es bleibt daher nur die Möglichkeit, im Allgemeinen zu bleiben und sich über Erfahrungen dem Thema zu nähern.

Allgemein beginnt auch der betriebswirtschaftliche Teil. Geschäftsmodell und Wettbewerbsstrategie sind Bereiche, die jeder Anbieter aufgrund der eigenen Rahmenbedingungen wie Wirtschaftskraft, vorhandenes Know-how und allgemeine Unternehmensstrategie selbst festlegen muss. Es ist für Anbieter wichtig, hier eine eigene Sicht auf den Markt, die Marktteilnehmer und die eigene Stellung im Markt zu entwickeln, denn die Reaktion darauf, also das eigene Geschäftsmodell, die eigene Wettbewerbsstrategie

sorgen für die Differenzierung und damit für die Wahrnehmung durch den Kunden. Hier gilt es, die richtigen Hinweise auf grundsätzliche Modelle zu geben.

ROI und TCO sind in der Aussage klarer und es gibt eine Reihe von Studien, gerade in Bezug auf Cloud Computing, die sich näher damit beschäftigen. Die TCO-Modelle greifen allerdings zu kurz, da sie keine Antworten auf Kapazitätsfragen bieten. Das jeweilige TCO-Modell liefert somit einen Preis, ohne die erhaltene Leistung konkret benennen zu können. Eine seriöse wirtschaftliche Planung ist unter diesen Umständen nicht möglich. Was bleibt, ist die Faustformel von einem möglichst hohen ROI und niedrigen TCO.

Einige betriebswirtschaftliche Themen sind in Bezug auf Cloud Computing durch Studien nicht besetzt, das Kostenmanagement (siehe Kapitel 4.4) ist ein Beispiel hierfür. An dieser Stelle bietet aber der Rückgriff auf die Mittel des IT-Controllings eine solide Grundlage für die Kostensteuerung. Ebenso verhält es sich beim Target Costing, dessen Modell sich, wie in Kapitel 5.3 beschrieben, auf Cloud Computing anwenden lässt.

Generell lässt sich damit sagen, dass die Aussagen für den technischen Teil dieser Arbeit konkreter sind als für den betriebswirtschaftlichen Teil. Die Ursache liegt in der Natur der Dinge: die Beurteilung der technischen Größen gilt für alle Firmen gleichermaßen. Ein Byte ist ein Byte. Anders sieht es bei den betriebswirtschaftlichen Größen aus. Hier beurteilt jede Firma die eigene wirtschaftliche Situation anders, womit sich die geringere Bestimmtheit erklären lässt. Ein beispielhafter Business-Case hätte hier Klarheit bringen können. Weil seine Rahmenparameter aber nicht übertragbar sind und damit immer nur ein Beispiel sein kann, war dies kein Ziel dieser Arbeit. Ein gewisser Grad an Unbestimmtheit war demnach gewollt. Trotzdem bietet die Arbeit für den Einstieg genügend Informationen und gibt Denkipulse, sich eingehender mit den Themen auseinanderzusetzen.

6.3 Offene Punkte

Die Dynamik der Cloud ist eine Herausforderung. Auf der technischen Ebene ist die fehlende allgemeingültige Definition des statistischen Multiplexing und damit der Elastizität ein offener Punkt. Ohne dies ist die Beurteilung der technischen Leistungsfähigkeit nicht möglich. Diese Lücke reicht aber noch tiefer. Es sei noch einmal auf Breitgand et al. verwiesen: "Since the cost efficiency of the cloud model critically depends on ability of the cloud providers to leverage statistical multiplexing, the clouds often operate in the over

commit mode by design¹⁴¹. Ohne Elastizität demnach keine Effizienz. Eine geringe Effizienz schlägt sich in niedriger Profitabilität nieder. Es ist also nicht nur eine Lücke im technischen Bereich, sondern ist auch für den betriebswirtschaftlichen Teil von Bedeutung.

Für den betriebswirtschaftlichen Teil stellen die Vorhersagbarkeit der Einnahmen und der Umgang damit einen offenen Punkt dar. Prognoseprozesse in Firmen laufen mit Zeithorizonten von Monaten, Quartalen und/oder Jahren ab. Cloud Computing mit seinem PAYG Preismodell basiert dagegen auf sehr kurzen Vertragslaufzeiten im Stundenbereich. Einnahmen können daher, selbst im Tagesverlauf, stark schwanken. In welchen Zyklen dies stattfindet, welche Ursachen hierfür verantwortlich sind und wie darauf zu reagieren ist, bleibt demnach offen. Lösungsmodelle werden aber nicht auf sich warten lassen, schließlich ist Cloud Computing angetreten, Computernutzung als Versorgungsdienstleistung anzubieten (siehe Seite 7). Die bisherigen Versorgungsdienstleister für Gas, Wasser oder Strom leben mit diesem Modell seit langem und haben damit bewiesen, dass Firmen mit schwankenden Einnahmen leben können.

Das Thema der Kostenrechnung für die Betriebsphase der Cloud ist ebenfalls ein offener Punkt. Kapitel 4.4 ist auf Basis betriebswirtschaftlicher Standardliteratur entstanden, da keine Literatur bezüglich der Anwendbarkeit der klassischen Modelle wie Kostenarten-, Kostenträger- und Kostenstellenrechnung vorliegt. Gleiches gilt für Prozesskostenrechnung und dessen Vorläufer des Activity-Based-Costings.

141 Vgl. Breitgand, David; Dubitzky, Zvi; Epstein, Amir; Glikson, Alex; Shapira, Inbar (2012), S. 73

Literaturverzeichnis

1. Bücher

- Abhishek, Vineet; Kash, Ian A.; Key, Peter** (2012): Fixed and market pricing for cloud services. In: 2012 IEEE Conference on Computer Communications workshops (INFOCOM WKSHPs 2012). Orlando, Florida, USA, 25 - 30 März 2012. Institute of Electrical and Electronics Engineers; IEEE Conference on Computer Communications; INFOCOM. Piscataway, NJ: IEEE, S. 157–162. Online verfügbar unter <http://dblp.uni-trier.de/db/conf/infocom/infocom2012w.html#AbhishekKK12>.
- Al-Roomi, May; Al-Ebrahim, Shaikha; Buqrais, Sabika; Ahmad, Imtiaz** (2013): Cloud Computing Pricing Models: A Survey. In: IJGDC 6 (5), S. 93–106. DOI: 10.14257/ijgdc.2013.6.5.09.
- Bachmann, Werner; Tiemeyer, Ernst** (2013): Handbuch IT-Management. Konzepte, Methoden, Lösungen und Arbeitshilfen für die Praxis ; [neu: IT-Management, Informations- und Datenmanagement, Cloud-Computing]. 5., überarb. und erw. Aufl. München: Hanser.
- Barroso, Luiz André; Hölzle, Urs** (2007): The Case for Energy-Proportional Computing. In: *Computer* (12/2007), S. 33–37, zuletzt geprüft am 26.12.2014.
- Behrens, Christian-Uwe** (2010): Grundlagen der Volkswirtschaftslehre. Einführung. Unter Mitarbeit von Matthias Kirspel. München: Oldenbourg Wissenschaftsverlag.
- Belz, Christian; Bieger, Thomas** (2006): Customer-Value. Kundenvorteile schaffen Unternehmensvorteile. 2., aktualisierte Aufl. Landsberg am Lech: mi-Fachverl.
- Berliner, Arnold** (2013): Lehrbuch Der Physik. Norderstedt: BoD – Books on Demand.
- Breitgand, David; Dubitzky, Zvi; Epstein, Amir; Glikson, Alex; Shapira, Inbar** (2012): SLA-aware resource over-commit in an IaaS cloud. In: Jorge Lobo (Hg.): 2012 8th International Conference on Network and Service Management. International Conference on Network and Service Management; CNSM; Workshop on Systems Virtualization Management; SVM. Las Vegas, Nevada, USA, 22 - 26 Oktober 2012. Piscataway, NJ: IEEE, S. 73–81. Online verfügbar unter <http://dblp.uni-trier.de/db/conf/cnsm/cnsm2012.html#BreitgandDEGS12>.
- Brugger, Ralph** (2009): Der IT Business Case. Kosten erfassen und analysieren, Nutzen erkennen und quantifizieren, Wirtschaftlichkeit nachweisen und realisieren. 2., korrigierte und erw. Aufl. Berlin: Springer Berlin (Xpert.press).
- Buyya, Rajkumar; Broberg, James; Goscinski, Andrzej** (Hg.) (2011): Cloud computing. Principles and paradigms. Hoboken, NJ: Wiley (Wiley series on parallel and distributed computing).
- Chang, Bao Rong; Tsai, Hsiu-Fen; Chen, Chi-Ming** (2013): Evaluation of virtual machine performance and virtualized consolidation ratio in cloud computing system. In: *Journal of Information Hiding and Multimedia Signal Processing* 4 (3), S. 192–200.
- Di Niu; Feng, Chen; Li, Baochun** (2012): Pricing cloud bandwidth reservations under demand uncertainty. In: ACM SIGMETRICS/PERFORMANCE Joint International Conference on Measurement and Modeling of Computer Systems, SIGMETRICS '12, London, United Kingdom, June 11-15, 2012, S. 151–162. Online verfügbar unter <http://dblp.uni-trier.de/db/conf/sigmetrics/sigmetrics2012.html#NiuFL12>.

- Ebel, Nadin** (2009): ITIL V3 Basis-Zertifizierung. Grundlagenwissen und Zertifizierungsvorbereitung für die ITIL-Foundation-Prüfung. München: Addison Wesley.
- Foster, Ian**: What is the Grid? A Three Point Checklist. In: GRIDToday. Online verfügbar unter <http://www.bibsonomy.org/api/users/diego/posts/0c85d01579d955d36cf6653c89c56861>.
- Foster, Ian T.; Zhao, Yong; Raicu, Ioan; Lu, Shiyong** (2008): Cloud Computing and Grid Computing 360-Degree Compared. In: 2008 Grid Computing Environments Workshop. GCE '08 ; 12 - 16 Nov. 2008, [Austin, Texas], abs/0901.0131. Grid Computing Environments Workshop; GCE. Piscataway, NJ: IEEE. Online verfügbar unter <http://www.bibsonomy.org/api/users/dblp/posts/4da5a6535b024a4b86b93a627b7b28b5>.
- Gadatsch, Andreas; Mayer, Elmar** (2010): Masterkurs IT-Controlling. Grundlagen und Praxis für IT-Controller und CIOs - Balanced Scorecard - Portfoliomanagement - Wertbeitrag der IT - Projektcontrolling - Kennzahlen - IT-Sourcing - IT-Kosten- und Leistungsrechnung. 4., erw. Aufl. Wiesbaden: Vieweg + Teubner (Studium).
- Galante, Guilherme; Bona, Luis Carlos Erpen De** (2012): A Survey on Cloud Computing Elasticity. In: 2012 IEEE Fifth International Conference on Utility and Cloud Computing (UCC 2012). International Workshop on Green and Cloud Computing Management (GCM). Chicago, Illinois, USA, 5-8 November 2012. Institute of Electrical and Electronics Engineers; Association for Computing Machinery. Piscataway, NJ: IEEE, S. 263–270. Online verfügbar unter <http://dblp.uni-trier.de/db/conf/ucc/ucc2012.html#GalanteB12>.
- Gohad, Atul; Narendra, Nanjangud C.; Ramachandran, Parathasarthy** (2013): Cloud Pricing Models: A Survey and Position Paper. In: 2013 IEEE International Conference on Cloud Computing in Emerging Markets (CCEM). Bangalore, India, 16-18 October 2013, S. 1–8.
- Guthrie, Forbes; Lowe, Scott; Coleman, Kendrick** (2013): VMware vSphere design. 2nd Edition. Indianapolis: Sybex.
- Held, Andrea** (2004): Oracle 10g Hochverfügbarkeit. Die ausfallsichere Datenbank mit RAC, Data Guard und Flashback. München: Pearson Deutschland; Addison-Wesley (Edition Oracle).
- Herbst, Nikolas Roman; Kounev, Samuel; Reussner, Ralf** (2013): Elasticity in Cloud Computing: What it is, and What it is Not. In: 2013 19th International Conference on Automation and Computing (ICAC): IEEE / Institute of Electrical and Electronics Engineers Incorporated. Online verfügbar unter <https://www.usenix.org/conference/icac13/elasticity-cloud-computing-what-it-and-what-it-not>.
- Heym, Michael; Montana, Constantin; Elben, Helmut** (2012): Paradigmenwechsel im IT-Sourcing – Die Sicht des CFO. Vom 1 st und 2 nd Generation IT-Outsourcing zum Cloud Computing – Aktueller Praxisbericht. In: Controlling & Management : ZfCM ; Zeitschrift für Controlling und Management 56 (5), S. 380–386.
- Hill, Richard; Hirsch, Laurie; Lake, Peter; Moshiri, Siavash** (2012): Guide to cloud computing. Principles and practice. Heidelberg: Springer Science & Business Media (Computer communications and networks).
- Holtchke, Bernhard; Heier, Hauke; Hummel, Thomas** (2009): Quo vadis CIO? Berlin, Heidelberg: Springer. Online verfügbar unter <http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10258004>.

Iglesias, Jesus Omana; Perry, Philip; Stokes, Nicola; Thorburn, James; Murphy, Liam (2013): A Cost-Capacity Analysis for Assessing the Efficiency of Heterogeneous Computing Assets in an Enterprise Cloud. In: 2013 IEEE/ACM 6th International Conference on Utility and Cloud Computing (UCC 2013). Dresden, 9 - 12 Dezember 2013. Institute of Electrical and Electronics Engineers; Association for Computing Machinery. Piscataway, NJ: IEEE, S. 107–114.

Kashef, Mohammad Mahdi; Altmann, Jörn (2012): A Cost Model for Hybrid Clouds. In: Kurt Vanmechelen, Jörn Altmann und Omer F. Rana (Hg.): Economics of grids, clouds, systems, and services. International Workshop on the Economics of Grids, Clouds, Systems, and Services, Bd. 7150. Paphos, Cyprus, 5. December 2011. GECON. Berlin: Springer (7150), S. 46–60. Online verfügbar unter <http://dblp.uni-trier.de/db/conf/gecon/gecon2011.html#KashefA11>.

Krishnan, Sriram (2010): Programming Windows Azure. Sebastopol: O'Reilly Media Inc. Online verfügbar unter <http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10766923>.

Kütz, Martin (2010): Kosten- und Leistungsrechnung für IT Services. In: Oliver Bartsch und Markus Lindinger (Hg.): IT-Servicemanagement. Stand März 2010. Köln: TÜV Media, S. 1–48.

Lampe, Ulrich; Hans, Ronny; Seliger, Marco; Pauly, Michael; Schiefer, Michael (2014): Pricing in Infrastructure Clouds - An Analytical and Empirical Examination. In: Smart sustainability: the information systems opportunity. 20th Americas Conference on Information Systems (AMCIS 2014). Savannah, Georgia, USA, 7 - 9 August 2014. Association for Information Systems. Red Hook, NY: Curran.

Li, Xinhui; Li, Ying; Liu, Tiancheng; Qiu, Jie; Wang, Fengchun (2009): The Method and Tool of Cost Analysis for Cloud Computing. In: 2009 IEEE International Conference on Cloud Computing. CLOUD 2009. Bangalore, India, 21-25 September 2009. IEEE. Piscataway, N.J.: IEEE, S. 93–100.

McCarthy, John (1961): Architects of the Information Society. Thirty-Five Years of the Laboratory for Computer Science at MIT. Edited by Hal Abelson. MIT Centennial. Massachusetts Institute of Technology, 1961.

Meffert, Heribert; Burmann, Christoph; Kirchgeorg, Manfred (2012): Marketing. Grundlagen marktorientierter Unternehmensführung. 11., überarb. und erw. Aufl. Wiesbaden: Gabler Verlag (Meffert-Marketing-Edition).

Meng, Xiaoqiao; Isci, Canturk; Kephart, Jeffrey O.; Zhang, Li; Bouillet, Eric; Pendarakis, Dimitrios E. (2010): Efficient resource provisioning in compute clouds via VM multiplexing. In: Proceedings of the 7th International Conference on Autonomic Computing. ICAC 2010. Reston, VA, USA, 7-11 Juni 2010, S. 11–20. Online verfügbar unter <http://dblp.uni-trier.de/db/conf/icac/icac2010.html#MengIKZBP10>.

Münzl, Gerald; Przywara, Bernhard; Reti, Martin; Schäfer, Jörg; Sondermann, Karin; Weber, Mathias; Wilker, Andreas (2009): Cloud Computing - Evolution in der Technik, Revolution im Business. BITKOM-Leitfaden. Hg. v. Dr. Mathias Weber. BITKOM - Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e. V. Berlin.

Mussnig, Werner (2013): Strategien entwickeln und umsetzen. Speziell für kleine und mittelständische Unternehmen. 2. aktualisierte Aufl. Wien: Linde.

Pepels, Werner (2006): Pricing leicht gemacht. Höhere Gewinne durch optimale Preisgestaltung. Heidelberg: Redline Wirtschaft.

Porter, Michael E. (2013): Wettbewerbsstrategie. Methoden zur Analyse von Branchen und Konkurrenten. 12., erweiterte und aktualisierte Auflage. Frankfurt am Main: Campus.

Shawky, Doaa M.; Ali, Ahmed F. (2012): Defining a measure of cloud computing elasticity. In: ICSCS 2012. 2012 1st International Conference on Systems and Computer Science, Bd. 2012. Villeneuve d' Ascq, France, 29-30 August 2012. Ecole PolyTech'Lille, S. 1–5.

Song, Weijia; Xiao, Zhen (2013): An Infrastructure-as-a-Service Cloud. On-Demand Resource Provisioning. In: Xiaoyu Yang und Lu Liu (Hg.): Principles, methodologies, and service-oriented approaches for cloud computing. Hershey, Pa.: Business Science Reference, S. 302–324. Online verfügbar unter http://zhenxiao.com/papers/cloud_book_chapter.pdf.

Umbeck, Tobias (2009): Musterbrüche in Geschäftsmodellen. Ein Bezugsrahmen für innovative Strategie-Konzepte. Wiesbaden: Gabler.

Waschke, Marvin (2012): Cloud Standards. Agreements That Hold Together Clouds. Berkeley, CA: Apress. Online verfügbar unter <http://dx.doi.org/10.1007/978-1-4302-4111-9>.

Weber, Andreas (2014): Resource Elasticity Benchmarking in Cloud Environment. Master Thesis. Karlsruhe Institute of Technologie, Karlsruhe. Department of Informatics Institute for Program Structures and Data Organization (IPD). Online verfügbar unter <https://sdqweb.ipd.kit.edu/publications/pdfs/Weber2014.pdf>, zuletzt geprüft am 27.12.2014.

Wöhe, Günter; Döring, Ulrich (2013): Einführung in die allgemeine Betriebswirtschaftslehre. 25.Auflage. München: Vahlen.

Wünsche, Manfred (2010): BWL für IT-Berufe. Ein praxisorientierter Leitfaden für kaufmännische Berufsfelder. 2. Auflage. Wiesbaden: Vieweg + Teubner.

Yuan, Dong (2011): Achieving the best trade-off between computation and storage in the cloud. cost model, benchmarking and strategies for datasets storage of scientific applications. Thesis (PhD). Swinburne University of Technology, Melbourne. Faculty of Information and Communication Technologies.

2. Internet-Quellen ohne Autor

Amazon-Regionen (2014). Online verfügbar unter <https://aws.amazon.com/de/about-aws/global-infrastructure/regional-product-services/>, zuletzt geprüft am 27.12.2014.

AmazonFree (2014): Kostenloses Nutzungskontingent für AWS. Hg. v. Amazon. Online verfügbar unter <https://aws.amazon.com/de/free/>, zuletzt geprüft am 27.12.2014.

AmazonPricing: Amazon EC2 - Preise. Hg. v. Amazon. Online verfügbar unter <https://aws.amazon.com/de/ec2/pricing/>, zuletzt geprüft am 28.12.2014.

AmazonSLA (2013): Amazon EC2 Service Level Agreement. Hg. v. Amazon. Online verfügbar unter <https://aws.amazon.com/de/ec2/sla/>, zuletzt aktualisiert am 01.06.2013, zuletzt geprüft am 27.12.2014.

Bibliotheksportal (2014): Sekundärforschung. Hg. v. Deutscher Bibliotheksverband. Online verfügbar unter <http://www.bibliotheksportal.de/index.php?id=655>, zuletzt aktualisiert am 06.10.2014, zuletzt geprüft am 26.12.2014.

Cirba (2014): Explore Cirba by Use Case. Hg. v. Cirba Inc. Online verfügbar unter <http://www.cirba.com/solutions/Cirba-solutions-by-use-case.htm>, zuletzt geprüft am 27.12.2014.

Citrix Systems: XenDesktop and XenApp. Best Practices – Reference Guide. Hg. v. Inc. Citrix Systems. Online verfügbar unter <https://support.citrix.com/servlet/KbServlet/download/30418-102-710697/XenDesktop%20and%20XenApp%20-%20Best%20Practices.pdf>, zuletzt geprüft am 26.12.2014.

Cloudstrategymag (2014): Report: 2013 IaaS Market Revenue Surpassed \$4 Billion. Online verfügbar unter <http://www.cloudstrategymag.com/articles/85257-report-2013-iaas-market-revenue-surpassed-4-billion>, zuletzt aktualisiert am 21.08.2014, zuletzt geprüft am 26.12.2014.

Duden (2013): Elastizität. Hg. v. Bibliographisches Institut GmbH. Online verfügbar unter <http://www.duden.de/rechtschreibung/Elastizitaet>, zuletzt geprüft am 27.12.2014.

Gartner: Cloud Computing. IT Glossary. Hg. v. Gartner Inc. Online verfügbar unter [https://www.gartner.com/it-glossary/Cloud Computing/](https://www.gartner.com/it-glossary/Cloud%20Computing/), zuletzt geprüft am 26.12.2014.

IDC (2014): Public Cloud Services Spending Is Being Driven by Enterprise Applications Solutions, According to IDC. Framingham. Online verfügbar unter <http://www.idc.com/getdoc.jsp?containerId=prUS24977214>, zuletzt aktualisiert am 07.07.2014, zuletzt geprüft am 26.12.2014.

Intel Corporation (2011): Implementing Cloud Storage Metrics to Improve IT Efficiency and Capacity Management. Hg. v. Intel Corporation. Online verfügbar unter <http://www.intel.in/content/dam/doc/white-paper/intel-it-implementing-cloud-storage-metrics-paper.pdf>, zuletzt geprüft am 14.10.2014.

ITIL V2 Glossary (2006): ITIL - IT Service Management. Glossary of Terms, Definitions and Acronyms. Hg. v. Office of Government Commerce. Office of Government Commerce. Online verfügbar unter http://www.bonneaud.net/download/ITIL_Glossary_May_v2_2007.pdf, zuletzt geprüft am 26.12.2014.

Lufthansa (2012): Kennzahlen 2012. Hg. v. Lufthansa Systems. Online verfügbar unter https://www.lhsystems.de/fileadmin/user_upload/files/de/company/keydata_2012.pdf, zuletzt geprüft am 26.12.2014.

MicrosoftPricing: Microsoft Azure - Datenübertragungen – Preisdetails. Hg. v. Microsoft. Online verfügbar unter <http://azure.microsoft.com/de-de/pricing/details/data-transfers/>, zuletzt geprüft am 28.12.2014.

MicrosoftSLA (2014): Vereinbarungen zum Servicelevel (SLAs). Hg. v. Microsoft. Online verfügbar unter <http://azure.microsoft.com/de-de/support/legal/sla/>, zuletzt aktualisiert am 12/2014, zuletzt geprüft am 27.12.2014.

NetworkComputing (2012): Consolidation, Scale-Out Architectures And Virtualization Drive 2012 Storage Trends. Online verfügbar unter <http://www.networkcomputing.com/storage/consolidation-scale-out-architectures-and-virtualization-drive-2012-storage-trends/d/d-id/1233426?>, zuletzt aktualisiert am 27.02.2012, zuletzt geprüft am 26.12.2014.

NIST: The NIST Refence on Constants, Units and Uncertainty. Prefixes for binary multiples. Hg. v. National Institute of Standards and Technology. Online verfügbar unter <http://physics.nist.gov/cuu/Units/binary.html>, zuletzt geprüft am 26.12.2014.

RackspacePricing: Die Rackspace Open Cloud. „Pay as you go“-Prinzip. Hg. v. Rackspace US, Inc. Online verfügbar unter <http://www.rackspace.com/de/cloud/pricing>, zuletzt geprüft am 28.12.2014.

RackspaceSLA (2014): Cloud Servers SLA. Hg. v. Rackspace US, Inc. Online verfügbar unter <http://www.rackspace.com/information/legal/cloud/sla>, zuletzt aktualisiert am 21.06.2011, zuletzt geprüft am 27.12.2014.

Technoconsult (2013): Wieso ist ihr Unternehmen nur unzureichend auf die Nutzung von Cloud Computing vorbereitet? Online verfügbar unter [http://www.cio.de/assets/galleries/detail-171467.html?i=RC_176_177&p=knowledge-center.outsourcing.Cloud Computing.anbieter.2942882.wie-sie-den-richtigen-anbieter-fuer-cloud.p1](http://www.cio.de/assets/galleries/detail-171467.html?i=RC_176_177&p=knowledge-center.outsourcing.Cloud%20Computing.anbieter.2942882.wie-sie-den-richtigen-anbieter-fuer-cloud.p1), zuletzt geprüft am 27.12.2014.

3. Internet-Quellen mit Autor

Ali, Majid: What is Cloud Computing Stack (SaaS, PaaS, IaaS). Online verfügbar unter <http://blogs.mazikglobal.com/cloud-computing-stack-saas-paas-iaas/>, zuletzt geprüft am 26.12.2014.

Banerjee, Ishan; Guo, Fei; Tati, Kiran; Venkatasubramanian, Rajesh (2013): Memory Overcommitment in the ESX Server. Hg. v. VMware Inc. VMware Technical Journal. Online verfügbar unter <https://labs.vmware.com/vmtj/memory-overcommitment-in-the-esx-server>, zuletzt geprüft am 26.12.2014.

Bemer, Bob: ORIGINS OF TIMESHARING. Online verfügbar unter <http://www.bobbemer.com/TIMESHAR.HTM>, zuletzt geprüft am 26.12.2014.

Böckle, Regina (2012): Händler verschenken Virtualisierungs-Chancen. Online verfügbar unter <http://www.channelpartner.de/a/haendler-verschenken-virtualisierungs-chancen,2586775>, zuletzt aktualisiert am 21.06.2012, zuletzt geprüft am 26.12.2014.

Columbus, Louis (2013): Gartner Predicts Infrastructure Services Will Accelerate Cloud Computing Growth. Hg. v. Forbes. Online verfügbar unter <http://www.forbes.com/sites/louiscolumbus/2013/02/19/gartner-predicts-infrastructure-services-will-accelerate-Cloud-Computing-growth/>, zuletzt aktualisiert am 19.02.2013, zuletzt geprüft am 26.12.2014.

Fojta, Tom (2014): Availability Zone Design in vCloud Director. Online verfügbar unter <https://fojta.wordpress.com/2014/06/27/availability-zone-design-in-vcloud-director/>, zuletzt aktualisiert am 27.06.2014, zuletzt geprüft am 27.12.2014.

Gronau, Norbert (2012): Wirtschaftlichkeit von Informationssystemen. Enzyklopaedie der Wirtschaftsinformatik. Hg. v. Oldenbourg Wissenschaftsverlag. Lehrstuhl für Wirtschaftsinformatik und Electronic Government. Online verfügbar unter <http://www.enzyklopaedie-der-wirtschaftsinformatik.de/wi-enzyklopaedie/lexikon/daten-wissen/Informationsmanagement/Information-/Wirtschaftlichkeit-von-Informationssystemen->, zuletzt aktualisiert am 22.08.2012, zuletzt geprüft am 26.12.2014.

Holzer, Jan Mark; Curran, Chris; Radvan, Scott: Red Hat Enterprise Linux Virtualization Guide. Hg. v. Red Hat Inc. Online verfügbar unter https://access.redhat.com/documentation/en-US/Red_Hat_Enterprise_Linux/5/html/Virtualization/sect-Virtualization-Tips_and_tricks-Overcommitting_with_KVM.html, zuletzt geprüft am 26.12.2014.

Husmann, Sven (2012): Betriebswirtschaftslehre. Enzyklopaedie der Wirtschaftsinformatik. Hg. v. Oldenbourg Wissenschaftsverlag. Lehrstuhl für Wirtschaftsinformatik und Electronic Government. Online verfügbar unter <http://www.enzyklopaedie-der-wirtschaftsinformatik.de/wi-enzyklopaedie/lexikon/uebergreifendes/Kerndisziplinen/Betriebswirtschaftslehre>, zuletzt aktualisiert am 01.11.2012, zuletzt geprüft am 27.12.2014.

Kirchgeorg, Manfred: Wettbewerbsstrategie. Hg. v. Springer Fachmedien Wiesbaden GmbH. Online verfügbar unter <http://wirtschaftslexikon.gabler.de/Archiv/10592/wettbewerbsstrategie-v11.html>, zuletzt geprüft am 27.12.2014.

Krauß, Jens-Christoph; Schnathmann, Marina; Georgy, Ursula (2014): Sekundärforschung. Online verfügbar unter <http://www.bibliotheksportal.de/index.php?id=655>, zuletzt aktualisiert am 25.06.2014, zuletzt geprüft am 21.09.2014.

Leong, Lydia (2010): Designing to fail. Online verfügbar unter http://blogs.gartner.com/lydia_leong/2010/12/03/designing-to-fail/, zuletzt aktualisiert am 03.12.2010, zuletzt geprüft am 26.12.2014.

Lowe, Scott D. (2012): Best Practices for Oversubscription of CPU, Memory and Storage in vSphere Virtual Environments. Online verfügbar unter <https://communities.vmware.com/servlet/JiveServlet/previewBody/21181-102-1-28328/vsphere-oversubscription-best-practices%5B1%5D.pdf>, zuletzt geprüft am 4.1.2015.

McCallum, John C.: Graph of Memory Prices Decreasing with Time (1957-2014). Online verfügbar unter <http://www.jcmit.com/mem2014.htm>, zuletzt geprüft am 26.12.2014.

McLellan, Charles (2013): Virtualizing the Enterprise: An overview. Online verfügbar unter <http://www.zdnet.com/virtualizing-the-enterprise-an-overview-7000018110/>, zuletzt aktualisiert am 01.08.2013, zuletzt geprüft am 26.12.2014.

Mell, Peter; Grance, Timothy (2009): Effectively and Securely Using the Cloud Computing Paradigm. NIST, 07.10.2009. Online verfügbar unter https://www.cs.purdue.edu/homes/bb/cs590/handouts/Cloud_NIST.pdf, zuletzt geprüft am 21.09.2014.

Mell, Peter; Grance, Timothy (2011): The NIST Definition of Cloud Computing (800-145). Online verfügbar unter <http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf>.

Read, Jason (2014): Comparing Cloud Compute Services. Online verfügbar unter <http://blog.cloudharmony.com/2014/07/comparing-cloud-compute-services.html>, zuletzt aktualisiert am 16.07.2014, zuletzt geprüft am 26.12.2014.

Rieß, Ulrike (2011): RAID-Level im Überblick. Online verfügbar unter <http://www.speicherguide.de/storage-hardware/disk-subsysteme/raid-level-im-ueberblick-13981.aspx>, zuletzt aktualisiert am 28.04.2011, zuletzt geprüft am 26.12.2014.

Righini, Marco (2012): How Processor Core Count Impacts Virtualization Performance and Scalability. Hg. v. Intel Corporation. Online verfügbar unter <http://www.intel.com/content/dam/www/public/us/en/documents/white-papers/virtualization-xeon-core-count-impacts-performance-paper.pdf>.

Schouten, Edwin (2012): Rapid elasticity and the cloud. Online verfügbar unter <http://thoughtsoncloud.com/2012/09/rapid-elasticity-and-the-cloud/>, zuletzt aktualisiert am 18.09.2012, zuletzt geprüft am 27.12.2014.

Strautenrauch, Thomas (2014): Marktorientierte Zielkostenrechnung (Target Costing) zur Preisfindung bei neuen Produkten. WEKA Business Media AG. Online verfügbar unter http://www.managementpraxis.ch/praxistipp_print.cfm?nr=94, zuletzt geprüft am 27.12.2014.

Weber, Jürgen: Fixkostendegression. Online verfügbar unter <http://wirtschaftslexikon.gabler.de/Archiv/7666/fixkostendegression-v4.html>, zuletzt geprüft am 26.12.2014.

Weber, Jürgen; Voigt, Kai-Ingo: Economies of Scale. Online verfügbar unter <http://wirtschaftslexikon.gabler.de/Archiv/54610/economies-of-scale-v6.html>, zuletzt geprüft am 26.12.2014.

Wurm, Andreas (2010): Variable Speicherzuweisung mit Thin Provisioning. Harte Diät. Hg. v. Heise Zeitschriften Verlag. Hannover, zuletzt aktualisiert am <http://www.heise.de/ix/artikel/Harte-Diaet-998462.html>, zuletzt geprüft am 26.12.2014.

