Zürcher Hochschule
für Angewandte Wissenschaften

Cloud Computing

Modul Systemarchitektur - Software Architektur
14.12.2015
Daniel Liebhart

Software Architektur – Cloud Computing

Daniel Liebhart, 14.12.2015

Verfasser: Daniel Liebhart
3 Auflage: 2015
Version 3.0

Inhalt

1 Referenzen und Abkürzungen

Referenzen

[Beckereit et al. 2006]	F. Beckereit, T. Harrer, K.Müller, I. Wittmann, R. Zwickenpflug: Virtualisierung - Überblick und Glossar, BITKOM 2006
[Brian et al. 2012]	O. Brian, T. Brunschwiler, H. Dill, HP. Christ, B. Falsafi, M. Fischer, S.G. Grivas, C. Giovanoli, R.E. Gisi, R. Gutmann, M. Aiserswerth, M. Kündig, S. Leinen, W. Müller, D. Oesch, M. Redli, D. Rey, R. Riedl, A. Schär, A. Spichiger, U. Widmer, A. Wiggins, M. Zollinger: White Paper Cloud Computing, SATW 2012-11-6
[Buyya 2006]	R. Buyya: Computing an IT in the Next Decade, Panel Talk, 11th International CSI Computer Conference (CSICC 2006), Teheran, Iran, January 24-26, 2006
[Carvalho 2012].	L. Carvalho: Windows Server 2012 Hyper-V Cookbook, Pact Publishing, November 2012
[CAS 2013]	Cloud Security Alliance: The Notorious Nine - Cloud Computing Top Threats in 2013, February 2013
[Cearley 2010]	D.W. Cearley: Cloud Computing - Key Initiative Overview, Gartner, Inc. 2010
[Chappell 2008]	D. Chappell: Introducing the Azure Service Platform - An early look at Windows Azure, .NET Services, SQL Services, And Live Services, DavidChappell & Associates, October 2008
[Chowdhury, Boutaba 2010]	N.M.M.K. Chowdhury, R. Boutaba: A survey of network virtualization, Computer Networks, Volume 54, Issue 5, 8. April 2010
[Citrix 2012]	Citrix Systems, Inc.: Workplace of the Future: a global market research report - The workplace of the future offers mobility, bring-yourown device (BYOD) and innovative workspaces, Mobile Workstyles Survey White Paper, Citrix & Vanson Bourne, 2012
[Foster, Kesselmann 1999]	I. Foster, C. Kesselmann: The Grid – Blueprint of a New Computing Infrastructure, Morgan Kaufmann Publishers, San Francisco 1999
[Gelernter 1997]	D. Gelernter: Truth, Beauty and the Virtual Machine, Discover Magazine, September 1997
[Gens 2008]	F. Gens: Defining "Cloud Services" and "Cloud Computing", IDC eXchange Blog, September 23rd, 2008
[Hamelin et al. 2010]	R.D. Hamelin, D.C. Walden, M.E. Krueger: INCOSE Systems Engineering Handbook v3.2: Improving the Process for SE Practitioners, INCOSE, July 2010
[Höchel-Winter 2011]	C. Höchel-Winter: VXLAN (Virtual eXtensible LAN) – VMwares neuster Draft, Wissensportal 14. September 2011
[IBM 2011]	IBM: Getting Cloud Computing Right - Thought Leadership White Paper, IMB Global Technology Services, April 2011
[ISB 2012]	ISB: Cloud-Computing-Strategie der Schweizer Behörden 2012 - 2020, Verabschiedet vom Steuerausschuss E-Government am 25. Oktober 2012
[ISO/IEC 2008]	ISO/IEC: Systems and software engineering - System life cycle processes, ISO/IEC 15288, IEEE 15288-2008, Second Edition, 2008-02-01
[Krafzig et Al. 2005]	Krafzig, D., Banke, K., Slama, D.: Enterprise SOA. Service Oriented Architecture Best Practices, Prentice Hall International, 2005
[Kroker 2014]	M. Krocker: Die Amazon Cloud – das am schnellsten wachsende Software-Geschäft der IT-Geschichte, 24. Juli 2014
[Kurzidim 2014]	M. Kurzidim: Marktüberblick: Die besten Cloud-Anbieter der Schweiz, Computerworld, 10.7.2014

[Linthicum 2014]	D. Linthicum: Try again, cloud contenders: Amazon, Google, and Microsoft have won, InfoWorld, Aug 8, 2014
[Liu et al. 2011]	F. Liu, J. Tong, J. Mao, R. Bohn, J. Messina, L. Badger, D. Leaf: NIST Cloud Computing Reference Architecture - Recommendations of the National Institute of Standards and Technology - NIST Special Publication 500-292, September 2011
[Meinel et al 2011]	Ch. Meinel, Ch. Willems, S. Roschke, M. Schnjakin: Virtualisierung und Cloud Computing: Konzepte, Technologiestudie, Marktübersicht, Technische Berichte Nr. 44 des Hasso-Plattner-Instituts für Softwaresystemtechnik an der Universität Potsdam, 2011
[Mell, Grance 2011]	P. Mell, T. Grance: The NIST Definition of Cloud Computing - NIST Special Publication 800-145, September 2011
[Müller et al. 2012]	W. Müller, J. Dischl, U. Widmer: Kommentar zur Cloud-Computing-Strategie der Schweizer Behörden, 25.4.2012
[Oracle 2012]	Oracle: Cloud Reference Architecture - Oracle White Paper, Novermber 2012
[Papazoglou, Georgakopoulos 2003]	Papazoglou, M.P., Georgakopoulos, D.: Service-Oriented Computing, Communications of the ACM, October 2003 / Vol. 46, No. 10
[Ramuschkat 2012]	S. Ramuschkat: Funktionsweise und Architektur der Amazon Cloud - tecRacer 2012
[Ried et al. 2011]	S. Ried, H. Kisker, P. Matzke, A. Bartels, M. Lisserman: Sizing The Cloud – A BT Futures Report, Understanding And Quantifying The Future Of Cloud Computing , Forrester Research, April 21, 2011
[Röhl, Schmiedl 2004]	A. Röhl, S. Schmiedl: Let's Grid, Linux Enterprise 07 / 08.2004
[Schmid 2012]	M. Schmid: Blick in die Woken - Cloud Computing Architecture Blueprint, Bachelorarbeit, ZHAW - School of Engineering 1.6.2012
[Schmid et al. 2013]	M. Schmid, R. Hochuli, M. Kuendig, C. Schildknecht: SwissICT: Leitfaden Cloud-Architektur, Ausgabe vom 2.5.2013
[Sirtl 2012]	H. Sirtl: Windows Azure Referenzarchitektur, Stand November 2012
[Sirtl 2014]	H. Sirtl: Überblick über Microsoft Azure, Microsoft 8.1.2014
[Skurk 2012]	H. Skurk: Speichervirtualisierung - Leitfaden, BITKOM 2012
[Stifani et al. 2012]	R. Stifani, S. Pappe, G. Breiter, M. Behrendt: IBM Cloud Computing Reference Architecture, IBM Academy TechNotes Volume 3, Number 1, 2012
[Störchle 2007]	M. Störchle: The Mainframe and Virtualization, Technische Universität München, 2007
[Tannenbaum 2000]	A. Tannenbaum: Structured Computer Organization, 1999-2000, Prentice-Hall
[Tsai et al. 2010]	W-T. Tsai, X. Sun, J. Balasooriya: Service-Oriented Cloud Computing Architecture, 2010 Seventh International Conference on Information Technology, IEEE 2010
[Varia 2011]	J. Varia:Amazon Web Services - Architecting for The Cloud: Best Practices, January 2011
[Weber 2013]	M. Weber: Eckpunkte für sicheres Cloud Computing - Leitfaden für die Auswahl vertrauenswürdiger Cloud Service Provider, BITKOM 2013
[Ziegler 2013]	P.A. Ziegler: Der Schweizer Markt für Cloud Computing, MSM Research, Zürich 16.5.2013

Abkürzungen

AKV	Aufgabe Kompetenz Verantwortung
API	Application Programming Interface
AWS	Amazon Web Services
BCP	Business Conitnuity Planning
BITKOM	Bundesverband Informationswirtschaft Telekommunikation und neue Medien
Blob	Binary Large Objekt
BPaaS	Business Process as a Service
BPEL	Business Process Execution Language
BPM	Business Process Management
BPMN	Business Process Modelling Notation
CAS	Cloud Security Alliance
CCRA	Cloud Computing Reference Architecture
CDN	Content Delivery Network
CRA	Cloud Reference Architecture

CRM	Customer Relationship Management
DLP	Data Loss Protection
DNS	Domain Name Service
EDA	Event Driven Architecture
ESB	Enterprise Service Bus
GGF	Global Grid Forum
GUID	Global Unique Identifier
IaaS	Infrastructure as a Service
IAM	Identity and Access Management
IBM	International Business Machines
ICMP	Internet Control Message Protocol
ICT	Information & Communication Technologies
IEC	International Engineering Consortium
IEEE	Institute of Electrical and Electronics Engineers
ISB	Informatiksteuerungsorgan des Bundes
ISO	International Organization for Standardization
KMU	Kleinere und Mittlere Unternehmen
LPAR	Logical Partitioning
MEV	Monats Endverarbeitung
NIST	National Institute of Standards and Technology
OS	Operating System
OSPF	Open Shortest Path First (RFC 1983)
PaaS	Plattform as a Service oder
QoS	Quality of Service
RDID	Routing Domain Identifier
RSVP	Resource Reservation Protocol (RFC 2205)
SaaS	Software as a Service
SATW	Schweizerische Akademie der Technischen Wissenschaften
SDK	Software Development Kit
SLA	Service Level Agreement
SOA	Service Oriented Architecture
SPOC	Single Point of Contact
SW	Software
TCP	Transmission Control Protocol
TEV	Tages Endverarbeitung
TOGAF	The Open Group Architecture Framework
UDP	User Datagram Protocol
VLAN	Virtual Local Area Network
VM	Virtual Machine
VMM	Virtual Machine Monitor
VO	Virtuelle Organisation
VxLAN	Virtual eXtensible LAN

2 Einleitung

2.1 Einführung: Cloud Trends heute

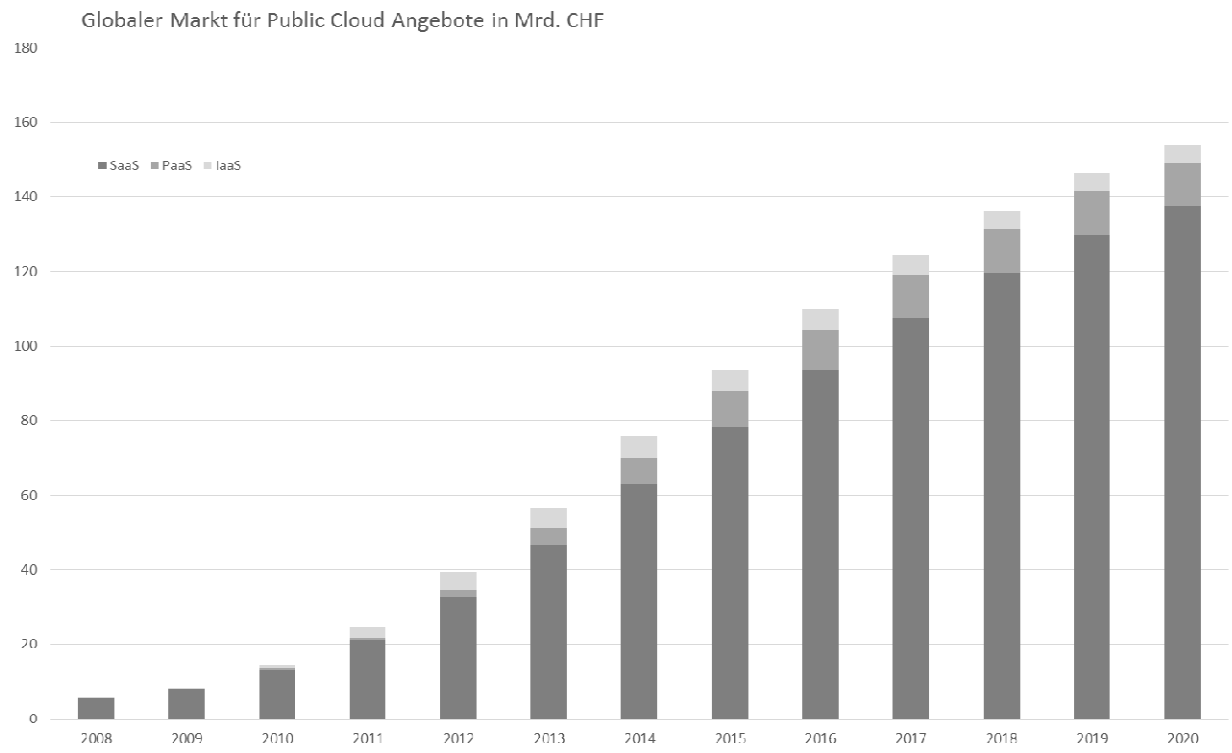

Abbildung 1: Globaler Markt für Private Cloud Angebote gemäss Forrester Research

Informatikdienste aus der Steckdose – Cloud Computing – ist seit mehreren Jahren einer der am stärksten wachsenden Bereiche des Globalen und auch des Schweizer ICT-Marktes. Und das nicht ohne Grund. Die Angebote werden immer attraktiver und versprechen gerade für Unternehmen erschwingliche Lösungen, die höchsten Ansprüchen genügen können. Die Vielzahl der Angebote und die schnelle Entwicklung macht es jedoch für viele Firmen nicht ganz einfach, das richtige Angebot zusammenzustellen.

2.1.1 Economy of Scale

Die Informationstechnologie ist für viele Unternehmen zur unverzichtbaren Ressource geworden. Kaum eine Firma kann seine Leistung ohne diese Ressource erbringen. Im Gegensatz zu Wasser, Gas oder Strom ist der Beschaffungs- und Betriebsaufwand für IT Dienste ungleich grösser. Genau das soll sich dank Cloud Computing ändern; IT Services werden als so genannte Betriebsmittel einfach bezogen und werden abhängig von der Nutzung bezahlt. Dahinter steckt eine Art firmenübergreifende „Economy of Scale“. Grosse Rechenzentren stellen Ressourcen zur

Verfügung, die Aufgrund der Vielzahl der Rechner sehr kosteneffizient betrieben werden können. Unternehmen, die Rechenleistung nur zu bestimmten Zeiten benötigen, beziehen diese nun zu einem Preis von ein paar Franken die Stunde bei Bedarf, statt einen Rechner, der im Monat im Schnitt um die 1.500 CHF kostet, ständig in Betrieb zu halten. Dasselbe gilt für Dienste wie beispielsweise E-Mail. Einen entsprechenden Server intern zu betreiben ist ungleich teurer, als sämtliche Nachrichten eines Unternehmens über einen Cloud Service abzuwickeln. Die Kosten pro User bleiben konstant, was vor allem für kleinere und mittlere Unternehmen interessant ist, da es bei internem Betrieb ungefähr gleich teuer kommt, einen E-Mail Server für 20 oder für 200 User zu betreiben. Die Kosten pro User variieren in diesem Fall sehr stark und sind bei wenigen Nutzern nicht mehr vertretbar. Seit mehr als 5 Jahren ist die Cloud Technologie marktreif und bietet eine Vielzahl verschiedenster Lösungen. Und Sie erfreuen sich zunehmender Nachfrage.

2.1.2 Die Cloud hat sich etabliert

Gemäss einer Umfrage des Markforschungsunternehmens MSM Research ist alleine im Jahr 2013 gegenüber dem Vorjahr die Nachfrage nach Cloud Services in der Schweiz um über 39% gestiegen [Ziegler 2013]. Und ein Ende dieser Entwicklung ist nicht abzusehen – es wird bis in die nächsten Jahre mit Wachstumsraten über 30% gerechnet. Dann dürfte Cloud Computing einen signifikanten Anteil an den Ausgaben Schweizer Unternehmen für die IT haben, die sich in diesem Jahr auf etwas über 16 Milliarden CHF belaufen sollen. Hinter dieser Entwicklung stecken handfeste Gründe. Gute IT Services sind bereits heute teuer und die laufende Entwicklung hin zur mobilen Arbeit mit einer Vielzahl verschiedener Geräte (PC, Laptop, Smartphone, Brille, Uhr, Sensoren und vieles andere Mehr - Analysten sprechen von 7 verschiedenen Geräte bis im Jahr 2020) verspricht keine Besserung [Citrix 2012].

Bereits heute sind Lizenzen (66 %), Personal (61.7%), Wartung (61.7%) und Umsetzung (57.4%) von IT Leistungen gemäss der MSM Umfrage „Der Schweizer Markt für Cloud Computing“ die grössten IT-Kostentreiber in hiesigen Unternehmen. Neben der zu lösenden Kostenfrage gilt es für die meisten Firmen in nächster Zeit Effizienz und Agilität zu steigern, um in einem sich stark verändernden Marktumfeld zu behaupten. Aus diesem Grund sind Cloud Computing Angebote heute unabdingbarer Bestandteil jeder Make-or-Buy Überlegung, die gemäss Eric Tveter, dem Chef der Cablecom, viele Firmen in diesen Jahren angesichts ihrer modernisierungsbedürftigen ICT-Infrastruktur machen.

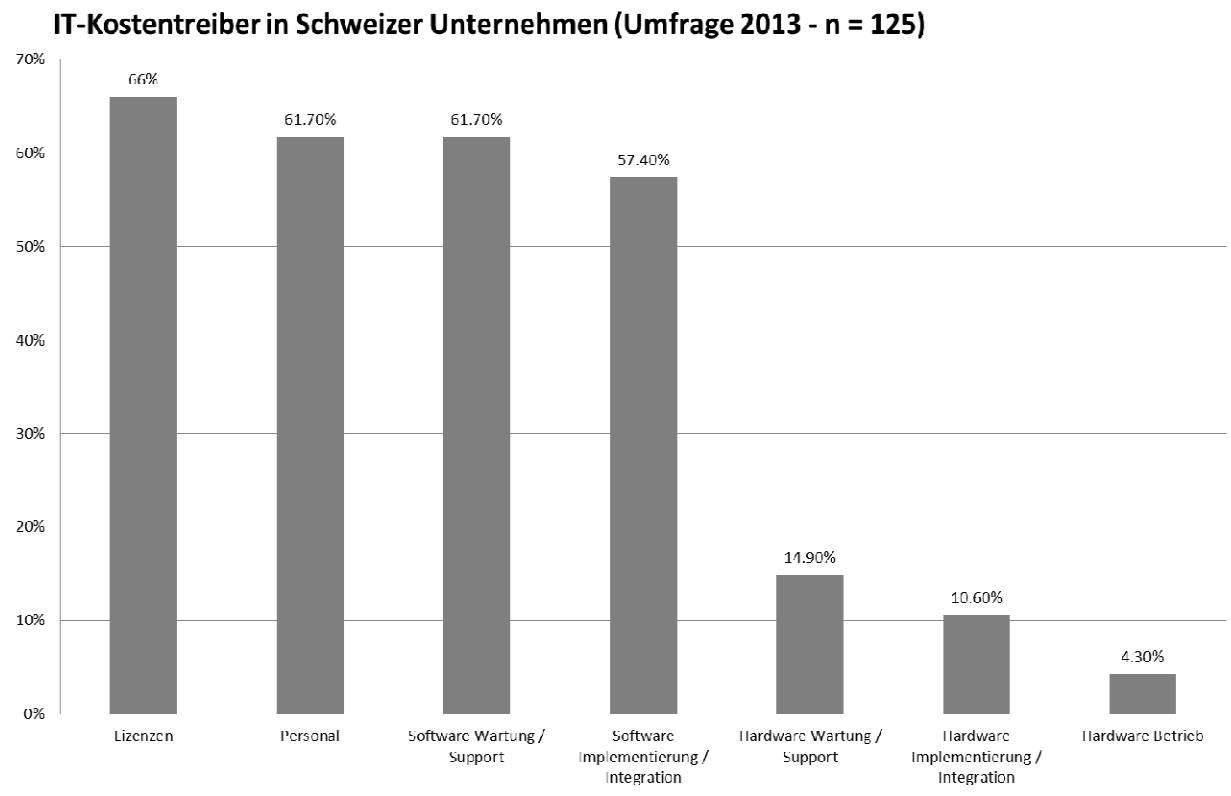

Abbildung 2: IT-Kostentreiber in Unternehmen ([Ziegler 2013]).

2.1.3 Cloud Angebote im Überblick

Heute existiert eine Vielzahl von Cloud Angeboten. Die Angebote reichen von der einfachen Bereitstellung von Rechenleistung über Mietsoftware bis hin zu spezialisierten Diensten, wie beispielsweise die 38 Online Backup Dienste, die im Mai 2014 auf der pcsupport.about.com Site in einem Vergleichstest zu finden waren.

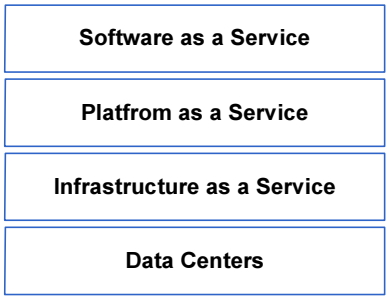

Abbildung 3: Cloud Computing Hierarchie ([Tsai et al. 2010]).

Cloud Angebote können heute anhand dem Leistungsort (Public, Private, Hybrid) oder der Leistungsart (IaaS – Infrastructure as a Service, PaaS – Plattform as a Service oder SaaS – Software as a Service) unterschieden werden. Ein Public Cloud Angebot erbringt Leistungen durch ein öffentliches Rechenzentrum, während eine Private Cloud sie durch ein firmeninternes Rechenzentrum bereitstellt. Als Hybride Cloud wird eine Kombination aus Private und Public Cloud bezeichnet. Die Leistungsart IaaS stellt Infrastrukturdienste wie beispielsweise Virtuelle Maschinen, sicheres Speichern von Daten, Backup oder Archivierung zur Verfügung. Unter PaaS werden Cloud Plattformen verstanden, die sich zur Entwicklung und für den Betrieb firmeneigener Lösungen eignen. SaaS Anbieter stellen Software über das Netz bereit. Neben dem Leistungsort und der Leistungsart gewinnt die Betriebsart zunehmend an Bedeutung. Es wird zwischen Self-Service Cloud und Manged Cloud unterschieden. Self-Service bedeutet, dass der Kunde das Angebot mit den Rahmenbedingungen des Anbieters nutzt, während Managed Cloud Services die Definition der Nutzungsbedingungen in Form eines so genannten SLA's (Service Level Agreement) erlaubt, welche Verfügbarkeit, Sicherheit und Performanz aufgrund Firmenvorgaben vertraglich absichert.

2.1.4 Was braucht ein Unternehmen?

Die alles entscheidende Frage jeder möglichen Cloud Nutzung ist diejenige, was eine Firma überhaupt braucht. Allen Cloud Angeboten ist gemeinsam, dass IT Dienste dynamisch bereitstellt, bedarfsorientiert abrechnet und vereinheitlicht werden können. Darüber hinaus gibt es Cloud Angebote, die sich geradezu aufdrängen, wenn Kosten gespart werden sollen. Ein Beispiel ist Office 365 von Microsoft, welches je nach Einsatz von 4.90 bis zu 25.10 CHF pro Monat und Mitarbeitenden kostet und im Leistungsumfang gegenüber einer gängigen Kombination aus Exchange und lokalen Office Programmen nichts zu wünschen übrig lässt. Da ist die Kosten-Nutzen Rechnung gegenüber den sprungfixen Kosten des internen Betriebs relativ schnell durchgeführt und ein Einsatz kann sich lohnen. Ein weiteres solches Beispiel sind Backup Lösungen. Swissbackup verlangt zum Beispiel 150 CHF pro Monat für ein Backupvolumen von 300 GB, was für viele KMU's durchaus reichen dürfte. Also rund 10% der Kosten, die ein interner Backupserver mit sich bringt. Allerdings gilt es in beiden Fällen, den Sicherheitsbedarf zu klären und einen guten Netzzugang zur Verfügung zu haben. Kosten sparen ist jedoch nur ein Aspekt. Falls Effizienz und Agilität durch den Einsatz von Cloud Angeboten gesteigert werden sollen, sind umfangreichere Abklärungen notwendig. Idealerweise durch vertrauenswürdige, stabile und spezialisierte Dienstleister, die eine massgeschneiderte Kombination zusammenstellen können, die definierten Anforderungen wie beispielsweise hohe Verfügbarkeit und Sicherheit sowie garantierte Antwortzeiten bei unterschiedlichem Lastverhalten erfüllen kann. Der Trend geht in diese Richtung; Cloud Anbieter orientieren sich zunehmend an den Bedürfnissen der Unternehmen und ihre Angebote erlauben immer flexiblere und besser kombinierbare Lösungen.

2.2 Definitionen

Cloud Computing steht für eine Art und Weise, wie die Informationstechnologie genutzt werden wird. Als global und universell verfügbare Ressource, die jederzeit und überall abgerufen werden kann. Australische Wissenschaftler des CLOUDS (Cloud Computing and Distributed Systems) Laboratory der University of Melbourne glauben, dass mit

Cloud Computing IT Services als so genannte Betriebsmittel wie etwa Wasser oder Strom geliefert werden können [Buyya 2006].

2.2.1 Analysten

Analysten definieren Cloud Computing wie folgt:

- **IDC:** Cloud Computing ist ein neu entstehendes IT Entwicklungs-, Einsatz- und Bereitstellungs-Modell, welches die Bereitstellung von Diensten und Lösungen in Echtzeit über das Netz erlaubt. Diese Dienste und Lösungen werden Cloud-Services genannt [Gens 2008].
- **Gartner:** Cloud Computing ist eine Art von EDV, die dem externen Nutzer skalierbare und elastische Leistungen mittels Internet-Technologien bereitstellt [Cearley 2010].
- **Forrester:** Cloud Computing ist eine standardisierte IT-Leistung (Dienst, Software oder Infrastruktur) die über das Internet als „Pay-as-You-Use“ und „Self-Service“ bereitgestellt wird [Ried et al. 2011]

2.2.2 NIST

Die am besten etablierte Definition von Cloud Computing stammt vom National Institute of Standards and Technology (NIST) [Mell, Grance 2011]:

> **Cloud Computing ist ein Modell, welches den allgegenwärtigen und komfortablen „on-demand“ netzwerkbasierten Zugriff auf einen Pool konfigurierbarer Rechenleistung (Netzwerke, Server, Speicher, Anwendungen und Dienste), die einfach bereitgestellt und mit minimalem Verwaltungsaufwand seitens des Servicegebers freigeschaltet werden kann, erlaubt.**

Das Modell hat fünf zentrale Charakteristiken (On-Demand Self-Service, Broad Network Access, Resource Pooling, Rapid Elasticity, Measured Service), drei Servicemodelle (Iaas, PaaS, SaaS) und vier Bereitstellungsmodelle (Private, Community, Public und Hybrid Cloud).

2.3 Charakteristika & Bereitstellungsmodelle

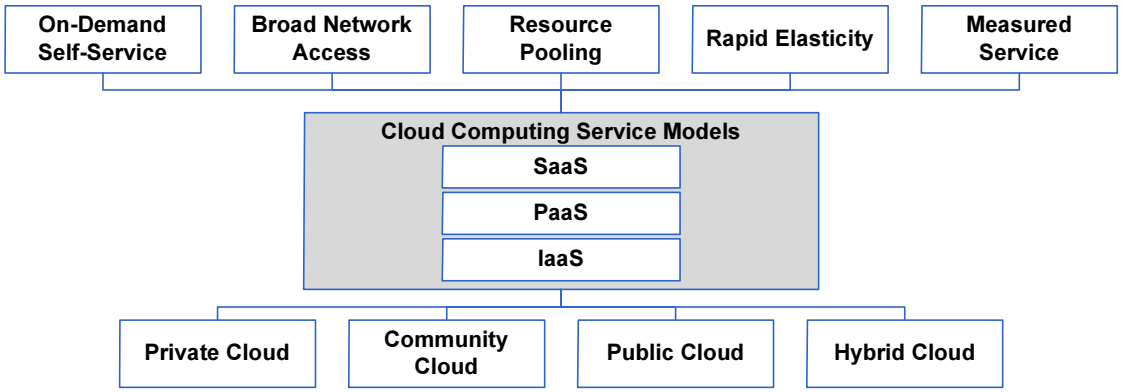

Abbildung 4: Charakteristika und Bereitstellungsmodelle des Cloud Computings

Die wichtigsten Charakteristika des Cloud Computings sind:

- **On-Demand Self-Service:** Ein Cloud Service nutzendes Unternehmen kann sich einseitig Leistungen wie beispielsweise Serverzeit oder Netzwerkspeicher beschaffen, ohne mit einer Person auf der Anbieterseite interagieren zu müssen.
- **Broad Network Access:** Auf Ressourcen können via standardisierte Mechanismen über ein Netzwerk zugegriffen werden.
- **Resource Pooling:** Ressourcen des Anbieters sind in einem Pool verfügbar, der über ein mandantenfähiges Betriebsmodell verschiedenen nutzenden Unternehmen zur Verfügung gestellt werden. Sie können dynamisch entsprechend dem Bedarf der Kunden zugeordnet werden. Der Kunde hat keine Kontrolle über den genauen Standort der Ressourcen, auch wenn er beispielsweise Land, Ort oder gewünschtes Rechenzentrum spezifizieren kann.

- **Rapid Elasticity:** Die Leistung kann dynamisch erweitert oder verringert werden. In manchen Fällen sogar automatisch, um sehr schnell auf wechselnde Nachfragesituation reagieren zu können. Aus Sicht des Kunden sind Ressourcen scheinbar unendlich verfügbar und können in jeder beliebigen Menge zu jedem beliebigen Zeitpunkt allokiert werden.
- **Measured Service:** Cloud Computing Systeme optimieren und kontrollieren die Nutzung der Ressourcen automatisch auf der zu einem Servicetyp passenden Abstraktionsebene. Die Nutzung kann überwacht, kontrolliert und rapportiert werden, um grösstmögliche Transparenz auf Kunden- und Lieferantenseite zu ermöglichen.

Die Cloud Bereitstellungsmodelle sind:

- **Private Cloud:** Die Cloud Infrastruktur steht ausschliesslich innerhalb einer Organisation zur Verfügung und wird beispielsweise von verschiedenen Abteilungen genutzt. Sie kann durch die Organisation, einem Dritten oder einer Kombination aus beidem betrieben, verwaltet oder gehalten werden.
- **Community Cloud:** Die Infrastruktur wird von einer Community gemeinsam genutzt. Sie kann durch eine oder mehrere Organisationen der Community, einem Dritten oder einer Kombination aus beidem betrieben, verwaltet oder gehalten werden.
- **Public Cloud:** Die Cloud Infrastruktur wird für eine öffentliche Nutzung ausgelegt. Sie wird durch den Cloudanbieter betrieben, verwaltet und gehalten.
- **Hybrid Cloud:** Sie ist eine Kombination aus Private, Community und Public Cloud Infrastrukturvarianten. Die Ressourcen sind jedoch so organisiert, dass die als ein standardisiertes Cloudangebot genutzt werden können.

2.4 Business Cases für die Cloud

Anwendungsfälle für die Cloud sind geprägt durch eine Reihe von wirtschaftlichen Faktoren, die durch die Eigenschaften einer Cloud Infrastruktur gegeben sind. Es sind dies die Faktoren „Economy of Scale“, „Pay-Per-Use“, „Self-Service“ und die Elastizität der Dienste:

- **Economy of Scale:** Grosse Rechenzentren stellen Ressourcen zur Verfügung, die Aufgrund der Vielzahl der Rechner sehr kosteneffizient betrieben werden können.
- **Per-Per-Use:** Rechenleistungen und Speicherplatz werden auf Basis der tatsächlichen Nutzung verrechnet.
- **Self-Service:** Rechenleistungen und Speicherplatz können sehr schnell und automatisiert bezogen werden.
- **Elastizität:** Unternehmen, die Rechenleistung nur zu bestimmten Zeiten zusätzliche Rechenleistung benötigen, beziehen diese nach Bedarf, statt die maximal notwendige Leistung ständig in Betrieb zu halten.

Zusätzlich sind Faktoren wie Time2Market, Sicherheit, Expertise und hohe Verfügbarkeit wichtige Treiber, um ein Cloud Angebot in Erwägung zu ziehen. Der deutsche IT-Branchenverband BITKOM (Bundesverband Informationswirtschaft Telekommunikation und neue Medien) hat diese Faktoren zusammengestellt [Weber 2013]:

Faktor	Erklärung
Economy Of Scale	■ Bündelung des Kundenbedarfs ■ Günstige Einkaufspreise aufgrund des höheren Einkaufsvolumens ■ Effektiverer Personaleinsatz, beispielsweise im Bereich Sicherheit ■ Hoher Automatisierungsgrad - niedrigere Personalkosten ■ Zugang zu kostengünstigen Energieressourcen
Pay-Per-Use	■ Grundsätzlich verbrauchsgesteuerte Abrechnung von Cloud Services ■ Bezugsgrößen bei Abrechnung variieren je nach Service-Ebene
Self-Service	■ Kunden nutzen Cloud Services im Self-Service
Elastizität der Dienste	■ Elastizität der Bedarfsdeckung durch Pay-Per-Use und Self-Provisioning ■ Kapazitätsrisiko auf den Cloud Service Provider verlagert ■ Abbau von Überkapazitäten möglich
Time2Market	■ Cloud Service Provider bieten gängige Konfigurationen ■ Inbetriebnahme von RZ-Kapazitäten innerhalb von Stunden oder weniger Tage ■ Freischalten zusätzlicher User im Minutenbereich

IT-Sicherheit und Datenschutz	■ Management der Daten in großen Rechenzentren und durch professionell ■ Ausgebildetes Personal ■ Professionelle IT-Sicherheit
Technologieexpertise	■ Cloud Service Provider sorgt für neuesten Stand der Technologie
Downtime	■ Cloud-Nutzer vermeiden Doppelkapazitäten ■ Keine Bevorratung von Ersatzteilen

2.4.1 Klassische Einsatzbeispiele

Die beiden wichtigsten etablierten Anwendungsfälle für Cloud Computing sind Lösungen mit sprunghaften Ressourcenanforderungen und Anwendungen mit grossen Mengen unkritischer (Archiv-)Daten.

> **Beispiel Animoto:**
> **Die Firma Animoto, stellt aus Bildern und Musik Filme und Slideshows) her. Im Jahr 2008 hat dieses Angebot durch eine Einbettung in Facebook derart grosse Nachfrage gefunden, dass die Infrastruktur innert Tagen von 50 auf 3500 Server aufgestockt werden musste. Dies ist in einem konventionellen Rechenzentrum kaum möglich. Diese Flexibilität kann nur eine Cloud bieten.**

Die sprunghafte Ressourcenanforderung (Elastizität) durch ein schnelles Ansteigen der Nachfrage ist zwar wünschenswert aber leider seltener als sich die Cloud Anbieter erhoffen. Realität ist jedoch, dass bestimmte Anwendungen nur periodisch genutzt werden. So ist die Nutzung vieler Online-Shops saisonal und es gibt eine Vielzahl von Tages-, Monats- oder Jahres-Endverarbeitungen, die von der Flexibilität und der Skalierung einer Cloud profitieren können.

> **Beispiel New York Times:**
> **Die New York Times speichert ihr Bildarchiv der letzten 60 Jahre – immerhin 4 Terabytes – in einer Cloud. Die Kosten der Speicherung in der Cloud sind um ein vielfaches kleiner als die einer internen Infrastruktur.**

Der zweite Anwendungsfall – (Archiv-)Daten – ist wesentlich verbreiteter. Es gibt kaum ein Unternehmen, welches über eine grosse Menge meist unstrukturierter Informationen verfügt, auf die nur selten zugegriffen wird, die jedoch trotzdem ab und an wichtig für das tägliche Geschäft sind. Der Ersatz aufwändiger Infrastrukturen durch eine Cloud liegt aus wirtschaftlichen Gründen auf der Hand.

3 Basis 1: Virtualisierung

3.1 Einleitung

Dank Virtualisierung ist es möglich, dass IT-Ressourcen sehr schnell – elastisch – und scheinbar unendlich zur Verfügung zu stellen. Im Prinzip erlaubt die Virtualisierung Maschinen auf Maschinen laufen zu lassen, die wiederum auf Maschinen laufen und diese logisch voneinander zu trennen. Das bedeutet, dass Kapazitäten geteilt oder zusammengefasst werden können. Darüber hinaus können verschiedenste Technologien flexibel kombiniert werden. Die Server-, Speicher- und Netzwerk-Virtualisierung sind die Grundlage dafür, warum sich der Einsatz einer Cloud Lösung überhaupt lohnt.

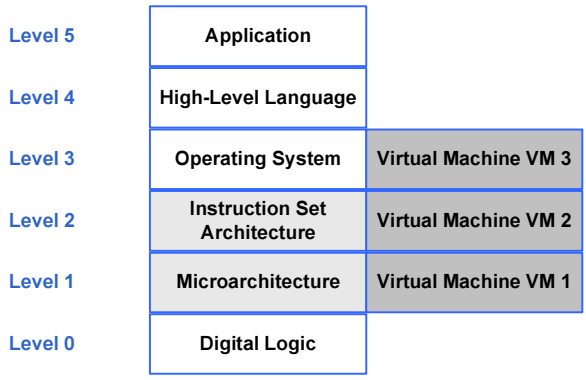

Abbildung 5: Das Konzept der Virtualisierung ([Tannenbaum 2000] und [Gelernter 1997])

Der entscheidende Punkt ist die Tatsache, dass eine Virtuelle Maschine wiederum auf einer anderen Virtuellen Maschine aufsetzen kann. Eine Virtuelle Maschine bildet immer das logische Modell eines Rechners ab.

Als Virtualisierung im Rahmen von Cloud Computing können sämtliche Techniken kombinieren, die Ressourcen eines Rechners aufteilen oder zusammenfassen und dabei das Funktionsverhalten einer realen Maschine kapseln. In Cloud Infrastrukturen kommen hauptsächlich Server-, Storage- und Netzwerk-Virtualisierung zum Einsatz.

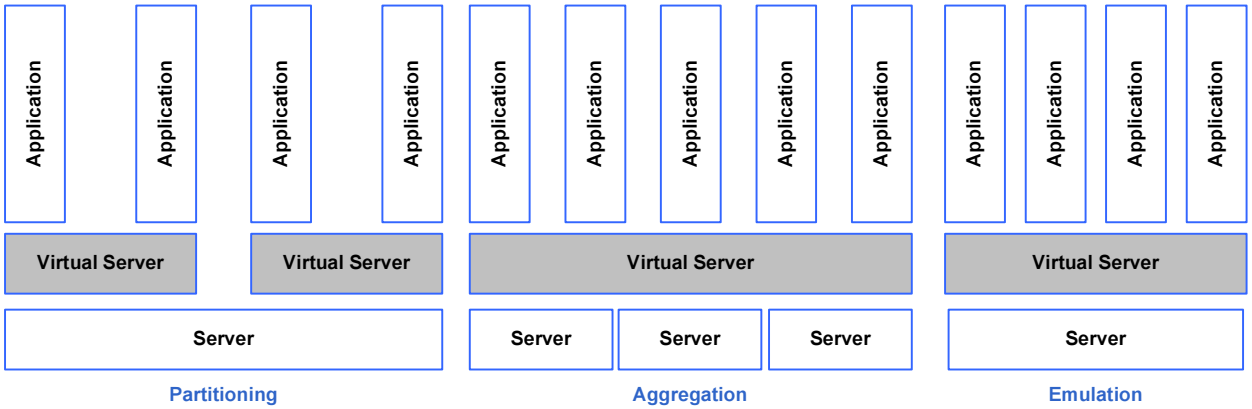

Abbildung 6: Virtualisierungklassen

Es lassen sich drei Klassen der Virtualisierung definieren:

- **Partitionierung:** Aufteilung einzelner physischer Systeme in mehrere logische Systeme
- **Aggregation:** Verbindung mehrerer physischer Systeme zu grösseren logischen Systemen
- **Emulation:** Abbildung unterschiedlicher Systemarchitekturen aufeinander

3.2 Server-Virtualisierung

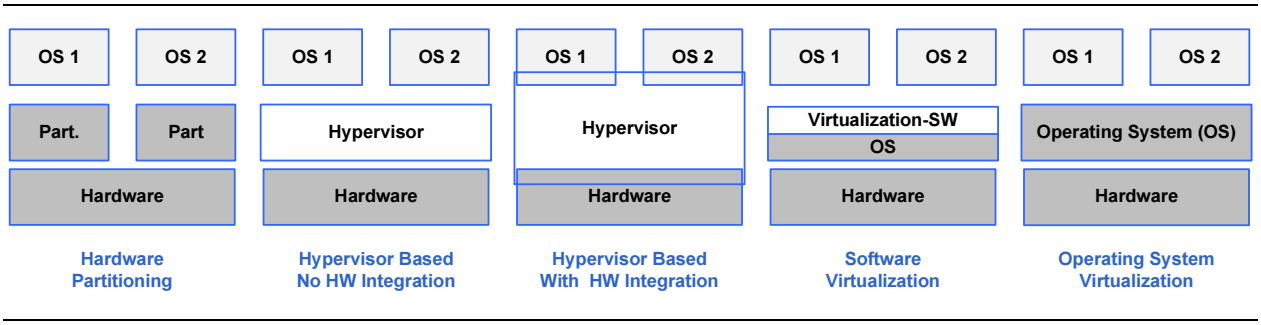

Abbildung 7: Die fünf Architekturen der Server-Virtualisierung

Die Virtualisierung von Servern ist in den 60er Jahren des letzten Jahrhunderts durch die IBM auf Grossrechnertechnologie eingeführt worden [Störchle 2007]. Sie bildet heute die Grundlage für sämtliche Cloud Infrastrukturen [Meinel et al 2011].

Die fünf Architekturen der Server-Virtualisierung sind [Beckereit et al. 2006]:

- **Hardware Virtualisierung:** Die einzelnen Betriebssysteme (OS 1 und OS 2) sind durch eine Partitionierung auf Hardware-Ebene vollständig getrennt. Bausteine, wie Motherboards können in der Regel ohne Neustart des OS hinzugefügt oder entfernt werden.
- **Hypervisor ohne Hardware-Integration:** Ein Hypervisor oder auch Virtual Machine Monitor (VMM) ist ein spezielle Software, die ausschliesslich die Funktionen für die Zuordnung physischer Ressourcen zu den virtuellen Servern realisiert.
- **Hypervisor mit Hardware-Integration:** Bestimmte Hardware unterstütz die Integration eines Hypervisors, sodass eine sehr gute Isolation virtueller Server erreicht werden kann und es können Ressourcen wie Prozessoren, Speicher oder I/O Geräte im laufenden Betrieb ausgewechselt werden. IBM nennt diese Technologie LPAR (Logical Partitioning). Häufig wird zusätzlich Aggregation unterstützt und es kann auf einen Pool von Hardware zugegriffen werden und ein automatischer oder gesteuerter Lastausgleich ist möglich.

- **Software Virtualisierung:** Instanzen eines Betriebssystems laufen auf einer speziellen Virtualisierungs-Software, die wiederum auf einem Betriebssystem abläuft. Diese Virtualisierung-Schicht implementiert ebenfalls die virtuellen Server. Die Isolation der virtuellen Betriebssysteme ist nicht so ausgeprägt wie im Falle des Einsatzes eines Hypervisors.
- **Operating System Virtualisierung:** Instanzen eines Betriebssystems laufen direkt auf einem Betriebssystem. Die virtuellen Betriebssysteme können dabei Funktionen des darunter liegenden Betriebssystems erben.

3.3 Storage-Virtualisierung

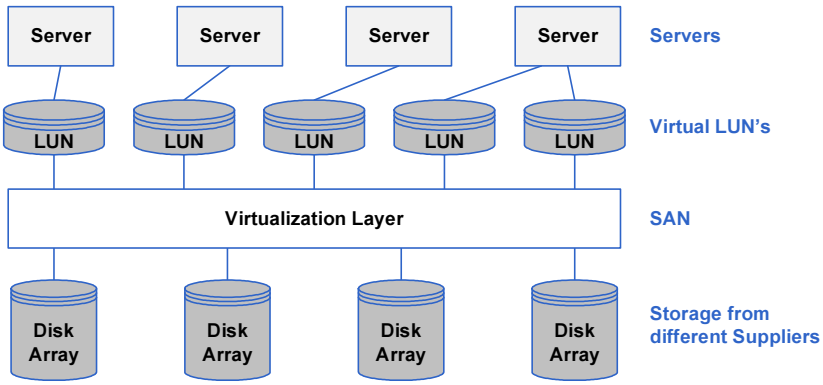

Abbildung 8: Prinzip der Storage-Virtualisierung

Die Virtualisierung von Speicher ist in vielen Unternehmen das Instrument zur umfassenden Verwaltung von Unternehmensdaten, deren Menge sehr schnell zunimmt. Dabei gilt ein einfaches Prinzip; Anwendungen greifen nicht mehr direkt auf eine bestimmte Festplatte oder einen anderen Speicher zu, sondern sie verwenden eine Virtualisierungsschicht. Diese Schicht lässt Speicher als grosse Einheit in einer logischen Form erscheinen. Sie kann flexibel aufgeteilt werden und sie ist nicht an physische Grenzen gebunden. Die tatsächlich verwendeten Speicherressourcen können dadurch sehr gut ausgenützt werden.

Virtualisierung von Storage in einem Netzwerk kann aufgrund von drei verschiedenen Architekturen erfolgen [Skurk 2012]:

- **In-Band Architektur:** Nutzdaten und Kontrolldaten werden auf demselben Weg übertragen. Die Virtualisierung findet über eine zusätzliche Appliance statt. Es müssen keine zusätzlichen Komponenten auf dem Server installiert werden – allenfalls ist eine MPIO (Multipath I/O) Lösung für redundante Datenpfade notwendig.
- **Out-of-Band Architektur:** Nutzdaten und Kontrolldaten gehen in dieser Variante getrennte Wege. Die eigentlichen Daten werden direkt über das SAN geschickt. Die Virtualisierung wird auf einer Appliance realisiert, die ausschliesslich Kontrolldaten verarbeitet und die Virtualisierung steuert. Es muss auf dem Server zusätzliche Software installiert werden, die aufgrund der von der Appliance gelieferten Adressen der Disk-Arrays die zu übertragenden Datenblöcke entsprechend auf das SAN schickt.
- **Split-Path Architektur:** Nutzdaten und Kontrolldaten werden in dieser Architektur ebenfalls getrennt. Allerdings ist keine zusätzliche Software auf dem Server notwendig, da die Trennung erst in einem intelligentem Switch-Port oder auch Fast Port erfolgt. Die Kontrolldaten werden interpretiert, die Anfrage an virtuelle Volumes wird in eine Anfrage au ein oder mehrere Disk-Arrays übersetzt und an das reale Disk-Array gesendet. Die Verwaltung der Virtualisierung wird von einem speziellen Server wahrgenommen, der mit der Appliance kommuniziert.

Andere Techniken der Speicher-Virtualisierung wie beispielsweise die Verwendung von RAID innerhalb von Plattensystemen oder der Einsatz von Storage-Kontroller sind im Rahmen von Cloud Infrastrukturen weniger relevant.

3.4 Netzwerk-Virtualisierung

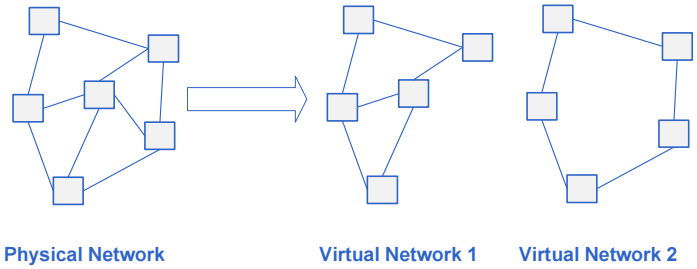

Abbildung 9: Prinzip der Netzwerk-Virtualisierung

Auf den ersten Blick ist Netzwerk-Virtualisierung eine Technik, die seit vielen Jahren in Unternehmen im Einsatz ist. Es ist die Netzwerk-Virtualisierung im Sinne des IEEE VLAN Standards 802.1q gemeint, die eine logische Unterteilung physischer Netzwerke in Virtuelle Netzwerke erlaubt. Dieses Konzept ist jedoch auf die Verwendung innerhalb einer Organisation ausgerichtet und erlaubt die Vergabe von maximal 4094 verschiedenen VLAN (Virtual Local Area Networks). Was für die meisten Unternehmen durchaus genügt, da logische Netze in den meisten Fällen aufgrund organisatorischer oder produktspezifischer Ordnungskriterien aufgebaut werden.

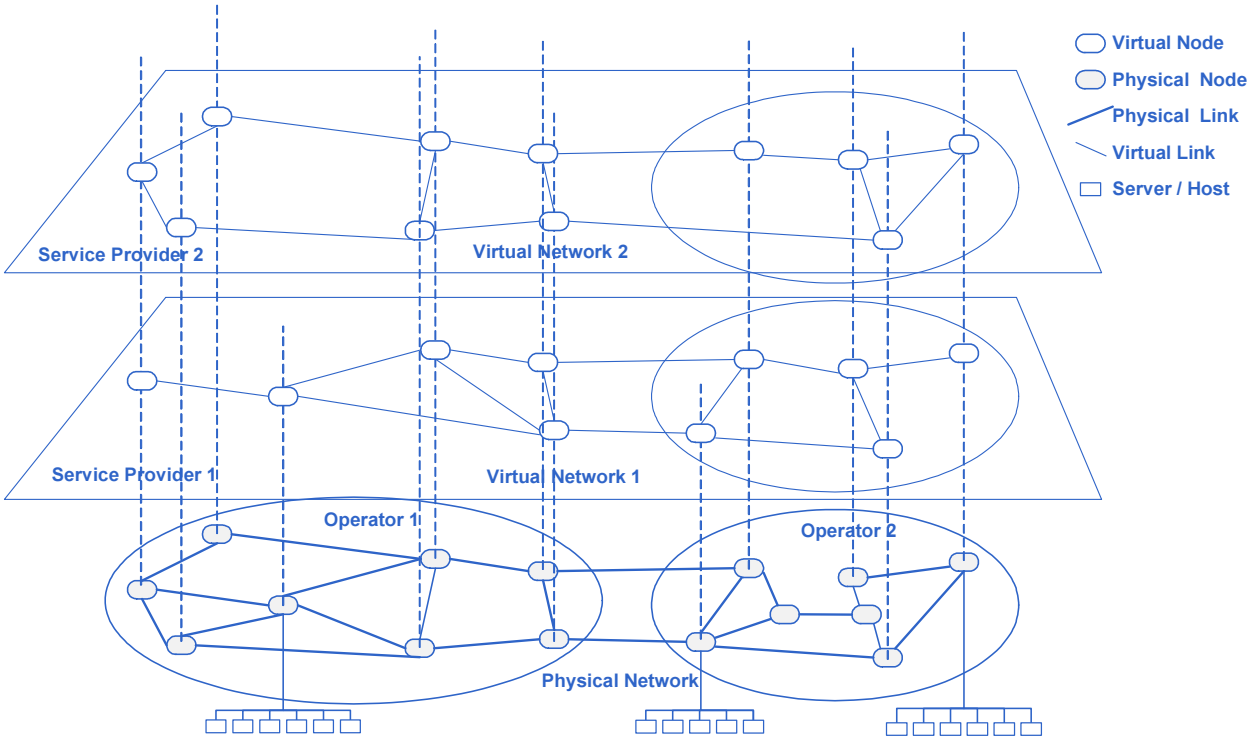

Abbildung 10: Umfassende Netzwerk-Virtualisierung

In einer vollständig virtualisierten Umgebung, wie sie der Einsatz von Cloud Technologie erfordert, ist jedoch eine Erweiterung des traditionellen VLAN-Begriffes notwendig, beispielsweise wie es Mosharaf Chowdhury, Forscher am amplab der Universität Berkeley, beschreibt:

> **Netzwerk-Virtualisierung ist eine Netzwerkumgebung, die es mehreren Dienstanbietern ermöglicht, dynamisch eine Vielzahl heterogener koexistierender virtueller Netze zusammenzustellen, die vollständig voneinander isoliert sind. Diese logischen Netzwerke erlauben es auf der Basis bestehender**

physischer oder logischer Netzwerke, die von verschiedenen Netzwerk Operatoren bereitgestellt werden, individuelle Punkt zu Punkt Dienste für den Kunden im laufenden Betrieb anzubieten [Chowdhury, Boutaba 2010].

Die wichtigsten Techniken der Netzwerk-Virtualisierung sind:

- **Portbased VLAN:** Die Zuordnung der VLAN's erfolgt über die Ports eines Netzwerkgerätes (Switches). Sämtliche Geräte, die am entsprechenden Port angeschlossen sind, befinden sich in demselben VLAN.
- **Tagged VLAN:** Die Ports eines Netzwerkgerätes befinden sich im Trunk-Modus, der es erlaubt über denselben Anschluss mehrere VLAN's zu führen. Die Identifikation des VLAN's erfolgt dabei über eine ID, den so genannten Tag. Die Zuordnung kann dynamisch aufgrund verschiedenster Kriterien erfolgen. Beispielsweise die MAC-Adresse des Gerätes oder die Art des Netzverkehrs.
- **Label Switching:** Label Switching erlaubt es einer Netzwerk Infrastruktur bestimmte Pfade (Routen) durch das Netzwerk vorzugeben. Diese Vergabe funktioniert ähnlich der Nummernvergabe in der Telefonie. Diese Technik erlaubt den Aufbau virtueller Punkt zu Punkt und Multipoint Verbindungen zwischen virtuellen Servern. Und hat gegenüber VLAN's den Vorteil, dass die Vergabe von Pfaden dynamischer und damit flexibler ist.
- **Proprietäre Techniken:** Hyper-V von Microsoft und VxLAN von VMware und Cisco sind proprietäre Erweiterung des IEEE VLAN Standards. Hyper-V arbeitet mit so genannten Routing Domänen ID's (RDID), um die Beschränkung von 4064 VLAN ID's zu umgehen [Carvalho 2012]. Als RDID werden Windows-GUID verwendet, also globale Identifier. VxLAN (Virtual eXtensible LANs) basiert auf einer so genannten Overlay Technologie, die es erlaubt durch die Einführung eines VxLAN Frames virtuelle Netze über virtuelle Netze zu legen [Höchel-Winter 2011].

3.4.1 Virtualisierung von Netzwerk-Komponenten

Netzwerke in Rechenzentren werden mit Hilfe einer Reihe von typischen Geräten, wie beispielsweise Switches, Router, Load-Balancer, Firewalls und anderen aufgebaut. Die Virtualisierung dieser Geräte ist eine Voraussetzung, dass Virtuelle Netze überhaupt betreiben werden können. In einer Cloud Umgebung werden Virtuelle Switches, Virtuelle Router, Virtuelle Firewalls und Virtuelle Load-Balancer eingesetzt. Sie werden im Normalfall als reine Softwarelösung realisiert und können damit auf einer VM eingesetzt werden.

- **Virtuelle Switches:** Ein virtueller Switch ist eine intelligente Netzwerkeinheit, die den Netzwerkverkehr zwischen verschiedenen Netzwerksegmenten durch Paketbasierte Verteilung an verschiedene Ports regelt. Er sorgt dafür, dass Datenpakete aufgrund ihrer Zieladresse in das richtige Netzwerksegment oder an die richtige Zieladresse weitergeleitet werden.
- **Virtuelle Router:** Virtuelle Router leiten Datenpakete auf Basis einer Routingtabelle weiter (Routing) oder blockieren eine Weiterleitung aufgrund bestimmter Regeln.
- **Virtuelle Firewalls:** Aufgabe einer virtuellen Firewall ist es, einen sicheren Übergang von Netzen mit unterschiedlichen Vertrauensstufen zu realisieren. Sie überprüft zu diesem Zweck Datenpakete im laufenden Betrieb und verwendet Sicherheitsrichtlinien, um unzulässige Kommunikation zwischen VMs zu unterbinden.
- **Virtuelle Load-Balancer:** Ein virtueller Load-Balancer ist für die Lastverteilung in einer virtuellen Umgebung zuständig. Er verteilt Anfragen aufgrund gegebener Lastsituationen an verschiedene VMs.

3.5 Übersicht Virtualisierung-Techniken

Für Cloud Infrastrukturen werden drei von fünf gängigen Virtualisierung-Arten (Client, Server, Storage, Network, Application) eingesetzt. Die Vorteile, Technologien und Auswirkungen sind in der nachfolgenden Tabelle zusammengefasst (nach [Skurk 2012]).

Eigenschaft	Server	Storage	Network
Vorteile	■ Einsparung von Hardware und Energie ■ Höhere Verfügbarkeit von Systemen ■ Größere Flexibilität ■ Einfachere Automatisierung	■ Effektivere Nutzung von ■ Speicherressourcen ■ Unterbrechungsfreie ■ Datenmigration ■ Einfacheres Management	■ Einfaches Ressourcensharing ■ Reduktion der Verkabelung ■ Größere Flexibilität ■ Einfacheres Management

Technologien	■ Hardware Virtualisierung ■ Hypervisor Partitionierung (mit und ohne Hardwareintegration) ■ Software Virtualisierung ■ Operating System Virtualisierung	■ In-Band Architektur ■ Out-Of-Band Architektur ■ Split-Band Architektur	■ Tagged VLAN ■ Label Switching ■ Hyper-V, VxLAN
Auswirkungen	■ Entkopplung Lebenszyklen von Hardware und Betriebssystem, dadurch maximaler Investitionsschutz ■ Minimierung der Downtimes ■ Geringere Kosten für ardware und Energie ■ Schnellere Umsetzung von Businessanforderungen.	■ Investitionsschutz ■ Geringere Betriebsaufwände ■ Schnellere Umsetzung von Businessanforderungen	■ Reduktion von Betriebsaufwänden ■ Schnellere Umsetzung von Businessanforderungen

3.6 Virtualisierung und Cloud Computing

Die Server-, Speicher- und Netzwerk-Virtualisierung machen die Wirtschaftlichkeit der Cloud auf der Ebene der Rechenzentren überhaupt aus. Nur so können physische Ressourcen zu logischen Leistungseinheiten gruppiert werden. Darüber hinaus ist es sehr viel Einfacher, die notwendige Mandantenfähigkeit (Mulltitenancy) einer Cloud Intrastruktur zu gewährleisten, wenn virtualisierte Systeme eingesetzt werden.

4 Basis 2: Grid Computing

4.1 Einleitung

Dank den Prinzipien von Grid Computing können auch sehr grosse Informationsmengen durch global vernetze Infrastrukturen schnell und sehr effizient verarbeitet werden. Da sich Cloud Computing aus dieser Technologie heraus entwickelt hat, sind heute entsprechende Lösungen wie beispielsweise „Big-Data Processing“ ein fester Bestanteil vieler Cloud Plattformen.

Grid Computing wurde ursprünglich für die Bewältigung sehr umfangreicher und rechenintensiver Aufgabenstellungen der Wissenschaft entwickelt. Es wurde Ende der 90er Jahre des letzten Jahrhunderts als Lösungsansatz für die so genannten „Grand Challenge Problems“ – einer Liste von Problemstellungen, zu deren Lösung im Rahmen des High-Performance Computing Act der Amerikanische Kongresses umfangreiche Forschungsgelder bereitgestellt hatte - entworfen worden.

Der damalige Präsident George Bush sagte 1991 beim Start dieses Programmes:

> **„Dieses Programm wird Forschern helfen, die grossen Herausforderungen der Wissenschaft zu lösen: Die DNA zu entschlüsseln, schwere Unwetter vorher zu sagen und neue supraleitende Materialien zu finden“.**

Die Lösung solcher Probleme erforderte eine Zusammenarbeit verschiedener Forschungsteams über verschiedene Standorte hinweg. Bereits 1995 wurde mit I-WAY (Information-Wide-Area-Year) an der Supercomputing Woche in San Diego eine Testumgebung bestehen aus 17 Standorten demonstriert. Diese Umgebung sollte aufzeigen, wie verteilte Anwendungen mit Leistungen im Teraflopbereich aufgebaut werden können. I-WAY gilt als der Startschuss zur Entwicklung von Grid Software, die an einer Vielzahl von Universitäten vorangetrieben wurde. Heute noch bestehende Beispiele dieser Entwicklungen sind das Globus Toolkit oder das Condor Projekt aus Berkley. Erst Ende der 90er Jahre ist das Global Grid Forum (GGF) entstanden, welches heute die wichtigste Standardorganisation für Grids ist. Eine typische Grid Anwendung dieser Tage war Telescience am San Diego Supercomputer Center. Der Neurowissenschaftler Mark Ellisman kombinierte Datenerfassung, grosse Datenbestände, komplexe Analysesysteme, Hochleistungsrechner und Spezialdisplays zu medizinischen Forschungszwecken.

Heute hat sich Grid Computing zum globalen Instrument der Zusammenarbeit verschiedenster Forschungsinstitutionen in Amerika, Europa und Asien etabliert und wird in Unternehmen für spezielle Aufgabenstellungen genutzt.

4.2 Definitionen

> **Das Wort “Grid” ist in Analogie mit dem Stromnetz gewählt worden, welches einen überall vorhandenen Zugriff auf Strom bereitstellt und einen wichtigen Einfluss auf die Leistungsfähigkeit der Menschheit hat, wie der beispielsweise der Computer und ein paar andere fortschrittliche Dinge auch hat [Foster, Kesselmann 1999]**

Grid Computing (englisch grid computing = Gitterberechnung) bezeichnet alle Methoden, die Rechenleistung vieler Computer innerhalb eines Netzwerks so zusammenzufassen, dass über den reinen Datenaustausch hinaus die (pa-

rallele) Lösung von rechenintensiven Problemen ermöglicht wird (verteiltes Rechnen). Jeder Computer in dem "Gitter" ist eine, den anderen Computern gleichgestellte Einheit. Damit kann, zu deutlich geringeren Kosten, sowohl die Kapazität als auch die Rechenleistung heutiger Supercomputer übertroffen werden. Grid-Systeme skalieren sehr gut: durch Hinzufügen von Rechnern zum Netz oder Zusammenfassen von Grids zu Meta-Grids erhöht sich die Rechenleistung in entsprechendem Masse.

Grid ist Infrastruktur zur integrierten, kollaborativen Nutzung von Ressourcen, die verschiedenen Organisationen gehören und von diesen verwaltet werden

Eigenschaften:

- überall verfügbar (persuasive)
- zuverlässig (dependable)
- konsistente Schnittstellen
- preiswert

Verwendete Ressourcen:

- Rechenkapazität, insbesondere Hochleistungsrechner
- Datenbanken, Files
- Messinstrumente, Geräte (z.B. Teilchenbeschleuniger, Windtunnel)
- Software (auf entferntem Rechner installiert)
- Menschen

4.3 Anwendung

Art	Beispiel	Eigenschaften
Distributed Supercomputing	DIS, Dynamik des Weltraums, Molekular-Chemie	Komplexe und umfangreiche Problemstellungen, die sehr viel Computing Power benötigen
Hight Throughput	Chip Design, Kryptographie, Studium von komplexen Systemen	Dynamischer Einbezug vieler Geräte, um den Durchsatz zu steigern
On Demand	Telemedizin, Komplexe Berechnungen, Wolken-Forschung	Lokales Computing wird verstärkt durch temporär eingebundene Remote Infrastructures
Data Intensive	Wettervorhersage, Physikalische Daten, Datensammlung	Synthese neuer Information aus einer sehr grossen Datenmenge
Collaborative	Collaborative Design, Schulung, Datenexploration	Unterstützung der Zusammenarbeit sehr grosser Gruppen

Andere Anwendungsbeispiele:

- Anwendungen, die nur gelegentlich hohe Rechenkapazität benötigen, z.B. Reaktion auf Krisensituationen, ungleichmäßig ausgelastete Server
- Bessere Auslastung der vorhandenen Rechner
- Zugriff auf entfernte Software, z.B. Rechendienst zum Lösen von Gleichungssystemen
- Anwendungen mit extrem hohem Rechenzeitbedarf, z.B. Klimamodellierung
- Kollaboratives Rechnen, d.h., gemeinsame Bearbeitung eines Projekts durch Teammitglieder, die sich an verschiedenen Orten befinden, aber auf gemeinsame Daten zugreifen
- Nutzung von Softwarekomponenten, die in verschiedenen Firmen installiert sind, zur gemeinsamen Lösung einer Aufgabe, z.B. Konstruktion eines Flugzeugs
- Verbinden von wissenschaftlichen Geräten, die Daten erzeugen mit Hochleistungsrechnern, die Daten auswerten. Insbesondere für große Datenmengen mit Zugriff auf gleiche Daten durch verschiedene Nutzer z.B. Ergebnisse physikalischer Experimente, Messdaten für Wettervorhersage (Bezeichnung: Data Grid)
- Anwendungen mit geringen Sicherheitsanforderungen

4.4 Der Begriff der Virtuellen Organisation (VO)

Im Zusammenhang mit Grid Computing ist der Begriff der Virtuellen Organisation (VO) zentral [Röhl, Schmiedl 2004]:

Schließen sich mehrere Individuen oder Institutionen zu einem Rechnerverbund zusammen, werden sie als virtuelle Organisation bezeichnet.

Sie zeichnen sich aus durch:

- Sehr flexible Sharing-Beziehungen, die sich schnell ändern können: Client-Server- oder Peer-To-Peer-Strukturen werden je nach Bedarf eingerichtet.
- Finden und Charakterisieren von Ressourcen.
- Genaue Kontrolle über die geteilten Ressourcen.
- Delegationsmechanismen zur lokalen oder globalen Ausführung von Programmen mit den jeweiligen Benutzerrechten.
- Generelle lokale und globale Policies.
- Kostenkontrolle und Accounting der verfügbaren und benutzten Ressourcen.
- Regelungsmechanismen zwischen hohen Performanceansprüchen und der Vermeidung hoher Kosten.
- Quality of Service.
- Scheduling.

4.5 Grid Basismodell

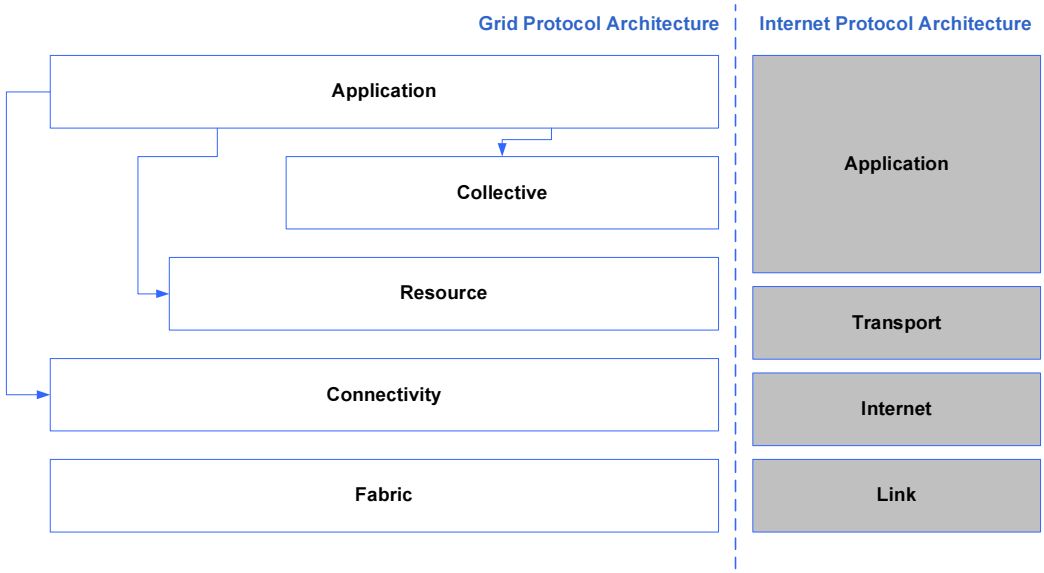

Abbildung 11: Grid Protocol Architektur im Vergleich mit der Internet Protocol Architektur

Die Grid Architektur erlaubt den universellen Zugriff auf verschiedene Grid Ressourcen und besteht aus den Schichten Fabric, Conncectivity, Resouce, Collective und Application.

- **Application:** Grid Applikationen nutzen Grid Infrastrukturen und operieren innerhalb einer Virtuellen Organisation. Auf jeder Schicht wird ein Dienst zur Verfügung gestellt, der wiederum über Programmierschnittstellen und Entwicklungsumgebungen verwendet werden kann.
- **Collective:** Der Collective Layer umfasst Protokolle zur globalen Verwaltung und zum globalen Zugriff unabhängig von der konkreten Realisierung einer bestimmten Fabric Layer Resource. Collective Funktionen können als persistente Dienst mit den zugehörigen Protokollen oder auch als System Development Kits (SDK's) mit den zugehörigen API's eingesetzt werden, um Grid Applikationen zu erstellen.

- **Resource:** Der Resource Layer ist für das sichere Aushandeln, die Initiierung, das Monitoring, die Kontrolle, die Verrechnung und die Bezahlung der gemeinsamen Benützung von Operationen zuständig. Diese Funktionen werden über Protokolle zur Verfügung gestellt. Der Resource Layer ruft Fabric Layer Funktionen auf und verwaltet lediglich einzelne Ressourcen.
- **Connectivity:** Der Connectivity Layer definiert die zentralen Kommunikations- und Authentisierungs-Protokolle für Grid Netzwerk Transaktionen. Die Kommunikations-Protokolle sind für den Datenaustausch zwischen Ressourcen des Fabric Layer zuständig. Die Authentisierungs-Protokolle sind für sicheren und verschlüsselten Datentransfer sowie für die Verifikation von Usern und Ressourcen verantwortlich.
- **Fabric:** Die Grid Fabric enthält alle Ressourcen, auf die über Grid Protokolle zugegriffen werden kann. Eine Ressource ist eine logische Entität, wie beispielsweise ein verteiltes Dateisystem, ein Computer Cluster oder ein verteilter Rechnerverbund.

4.5.1 Fabric: Schnittstellen für die lokale Kontrolle

Die Grid Fabric enthält alle Ressourcen, auf die über Grid Protokolle zugegriffen werden kann. Eine Ressource ist eine logische Entität, wie beispielsweise ein verteiltes Dateisystem, ein Computer Cluster oder ein verteilter Rechnerverbund. In diesem Fall enthält die Realisierung einer Ressource eine Reihe von internen Zugriffprotokollen, die jedoch aus Sicht Grid Protocol nicht wichtig sind.

The Komponenten einer Farbric implementieren lokale Operationen, die auf bestimmte Ressourcen als Resultat von "Sharing Operations" auf höherer Ebene zugreifen. Es existiert eine gegenseitige Abhängigkeit zwischen Funktionen auf lokaler Ebene (Fabric Level) und denjenigen auf einer höheren Ebene. Beispielsweise ist die Belegung im Voraus für lokale Ressourcen eine Voraussetzung dafür, dass high-level Services Ressourcen aggregieren können.

Die minimalen Services für Fabric Resources sind:

- **Abfrage (Enquiry) Services:** Sie erlauben die Abfrage der Struktur, des Status und der Leistungsfähigkeit einer Resource.
- **Management Services:** Diese Dienste umfassen alle Kontrollmechanismen, wie beispielsweise QoS (Quality of Service).

4.5.2 Fabric Resource Arten

Resource Type	Bemerkung	Enquiry Service	Management Service
Computational	Starting / Stoping, Execution Control	Hardware- & Software Characteristics, State Information	Process Control, Advance Reservation Mechanism
Storage	Putt / Get File, Striped Transfer, Read / Write, Reduce	Hardware- & Software Characteristics, Load Information (Space, Bandwidth)	Controls for Space, Disk Bandwidth, Network Bandwidth, CPU, Advance Reservation Mechanism
Network		Network Charateristics, Load	Proritization, Reservation
Code Repositories	Specialized Form of Storage Resources	State, Version	Version Control, Object Control
Catalogs	Specialized Form of Storage Resources	Historisation	Query, Update

4.5.3 Connectivity

Der Connectivity Layer definiert die zentralen Kommunikations- und Authentisierungs-Protokolle für Grid Netzwerk Transaktionen. Die Kommunikations-Protokolle sind für den Datenaustausch zwischen Ressourcen des Fabric Layer zuständig. Die Authentisierungs-Protokolle sind für sicheren und verschlüsselten Datentransfer sowie für die Verifikation von Usern und Resourcen verantwortlich.

Die Anforderungen an die Connectivity umfassen Transport, Routing und Naming. Dabei werden die gängigen Internet-Protokolle wie IP ICMP (Internet), TCP, UDP (Transport) und DNS, OSPF, RSVP (Applikation) verwendet.

Authentisierungs-Lösungen für Grid Computing sollten folgende Charakteristika aufweisen:

- **Single Sign-On:** Ein User muss nur einmal in ein System einloggen, um dann Zugriff auf mehrere Fabric Resourcen unabhängig vom Standort zu haben.
- **Delegation:** Ein User muss ein Programm mit seinen Rechten ausstatten können. Diese Rechte müssen gegebenenfalls auf andere Programme weiter übertragen werden können, die in ein Grid Computing Environment eingebunden sind.
- **Integration mit verschiedenen lokalen Security Lösungen:** Jede Site oder jeder Resoucen-Lieferant nutzt möglicherweise eine Reihe von Sicherheits-Mechanismen, wie beispielsweise Kerberos und Unix Sicherheit. Grid Security muss mit diesen lokalen Mechanismen interagieren können.
- **User-Based Trust Relationships:** Damit verschiedene Resourcen von demselben User verwendet werden können, darf nicht eine Autorisierung pro Resource notwendig sein.

4.5.4 Resource

Der Resource Layer ist für das sichere Aushandeln, die Initiierung, das Monitoring, die Kontrolle, die Verrechnung und die Bezahlung der gemeinsamen Benützung von Operationen zuständig. Diese Funktionen werden über Protokolle zur Verfügung gestellt. Der Resource Layer ruft Fabric Layer Funktionen auf und verwaltet lediglich einzelne Ressourcen.

Es existieren zwei Ressourcen Protokoll Klassen:

- **Information Protocols:** Sie werden für die Gewinnung von Informationen über die Struktur und den Status einer Resource verwendet (Status, Load, Usage Policy).
- **Management Protocols:** Diese Protokolle werden zum Aushandeln von verteilten Zugriffen, der Definition von Anforderungen an Ressourcen, QoS sowie die Steuerung der Ressource verwendet.

4.5.5 Collective

Dieser Layer ist für die Abstraktion von Services zuständig. Er vereinheitlicht beispielsweise verschiedene Fabric Layer Ressourcen, um einen universellen Zugriff auf eine grosse Menge verschiedenster Ressourcen zuzulassen.

4.5.6 Application

Grid Applikationen operieren innerhalb einer Virtuellen Organisation.

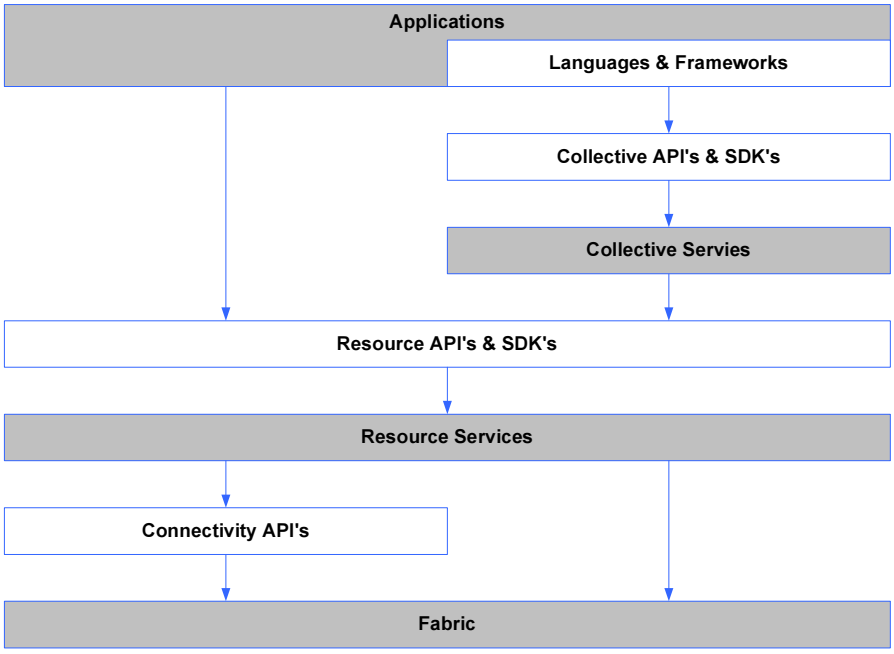

Abbildung 12: Grid Computing aus Sicht der Software Entwicklung [ANATOMY]

Auf jedem der Layer wird ein Dienst zur Verfügung gestellt, der wiederum über API's und SDK's verwendet werden kann.

Die API's greifen mittels Protokolle auf die unterliegende Schicht zu:

- Die Connectivity API's über die Connectivity Protocols auf die Fabric Resourcen.
- Die Resource API's und SDK's über Resource Service Protocols auf die Resource Services.
- Die Collective API's und SDK's über Collective Service Protocols auf die Collective Services.

4.5.7 Hourglass Model

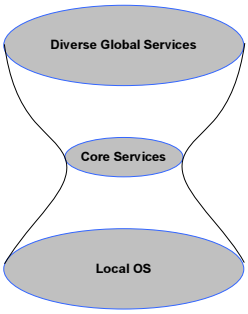

Abbildung 13: Hourglass Model

Das Hourglass Model sieht ein klar definiertes Set von Core Services als Basis-Infrastruktur vor. Dieses Basis-Set wird verwendet, um high-level domainspecific Services zu erstellen.

Die Design Prinzipien sind:

- Geringe Beteiligungskosten
- Erlaube lokale Kontrolle
- Unterstützung für Adaption

4.6 Aufgabenstellungen und Anwendungen

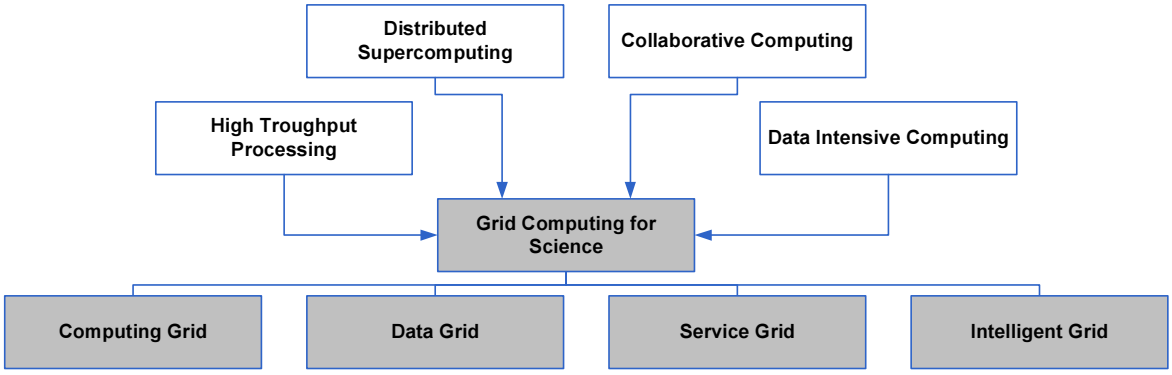

Abbildung 14: Aufgabenstellungen für GRID Computing

Die typischen Probleme bei denen sich Grid Computing als Strategie anbietet sind solche, die die Leistung einzelner Computer überfordern. So beispielsweise komplexe und umfangreiche Problemstellungen im Bereich Distributed Supercomputing, die sehr viel Computing Power benötigen oder jedoch die Synthese neuer Informationen aus einer sehr grossen Datenmenge im Bereich Data Intensive Computing. Weitere Einsatzgebiete sind so genannte On Demand Probleme, was nichts anderes bedeutet als die Verstärkung lokaler Rechenleistung durch temporär eingebundene Remote Infrastrukturen. Oder jedoch Aufgabenstellungen, die sehr hohen Durchsatz benötigen. Schließlich

werden Grids noch für die Unterstützung der Zusammenarbeit sehr großen Gruppen im Bereich Collaborative Computing eingesetzt. Grids werden je nach Anwendungsgebiet im wissenschaftlichen Bereich in Rechen-, Daten-, Service oder auch Intelligente Grids eingeteilt. Rechengrids entsprechen der virtualisierten Infrastruktur bestehend aus einer Vielzahl von einzelnen Nodes, während Datengrids dazu dienen, verteilte Speicherkapazitäten gemeinsam zu nutzen. Servicegrids stellen eine Umsetzung verteiler Anwendungen dar und Intelligente Grids sind eine Kombination aus allen anderen Kategorien, die sich zusätzlich selbst verwalten.

Anwendungen sind:

Art	Beispiel	Eigenschaften
Distributed Supercomputing	DIS, Dynamik des Weltraums, Molekular-Chemie	Komplexe und umfangreiche Problemstellungen, die sehr viel Computing Power benötigen
Hight Throughput	Chip Design, Kryptographie, Studium von komplexen Systemen	Dynamischer Einbezug vieler Geräte, um den Durchsatz zu steigern
On Demand	Telemedizin, Komplexe Berechnungen, Wolken-Forschung	Lokales Computing wird verstärkt durch temporär eingebundene Remote Infrastructures
Data Intensive	Wettervorhersage, Physikalische Daten, Datensammlung	Synthese neuer Information aus einer sehr grossen Datenmenge
Collaborative	Collaborative Design, Schulung, Datenexploration	Unterstützung der Zusammenarbeit sehr grosser Gruppen

4.6.1 Grid in Unternehmen

Während die klassischen Grid Infrastrukturen auf den weltweiten Einsatz für die Wissenschaftsgemeinde hin ausgelegt sind, kommen nun zusehends kommerzielle Produkte auf den Markt, die Grid Computing für Unternehmen erlauben. Somit werden Grids zu einem Baustein, mit dem betriebliche Informationssysteme realisiert werden können. Sie eignen sich für Anwendungen in performance-kritischen und verteilten Umgebungen und für die Optimierung vorhandener Ressourcen. Wenn es also darum geht, Daten, Prozesse und Funktionen transparent zu verteilen respektive darauf zuzugreifen oder wenn es darum geht, virtuelle Infrastrukturen optimal zu nutzen.

Die konkreten Einsatzgebiete und deren Platz in einer betrieblichen Gesamtarchitektur lassen sich daraus ableiten; Von der Realisierung einer Realtime Business Intelligence Infrastruktur über den Einsatz unternehmensweiter SOA-Grids oder als Technologie für die Datenvirtualisierung und für verteilten Caches bis hin zu High Perfomance Backup & Recovery werden eine Vielzahl von Anwendung möglich, die bis anhin nur mit sehr grossem Aufwand realisiert werden konnten.

Complex Realtime Intelligence kombiniert die Funktionalitäten von CEP (Complex Event Processing) und Datengrids und bietet so die Voraussetzung, sehr gut skalierbare Analyseanwendungen für komplexe Mustererkennungen in Echtzeitszenarien dem Business zur Verfügung zu stellen. Die Daten-Virtualisierung erlaubt über ein Datengrid den virtualisierten Zugriff auf verteilte Informationen in einem Cluster. So können beispielsweise die Zugriffe auf Datenspeichersysteme transparent für die Applikation gekapselt und gepuffert werden. Auf diese Weise können auch eventuelle Aktionen zur Unterstützung der Ausfallsicherheit des Zielsystems und notwendige Reaktionen von der Applikation entkoppelt werden. Eine Applikation muss somit nicht mehr auf Ereignisse zur Unterstützung der Ausfallsicherheit diverser Zielsysteme reagieren können, da dies auf Ebene des Grids erfolgt.

Verteilte Caches können Anwendungsdaten, Objekte und Prozesse in einem verteilten Cache, linear skalierbar und transaktionsgesichert vorhalten. Ein SOA-Grid ist nichts anderes als eine Umsetzung eines solchen Caches. So können beispielsweise Prozesse in serialisierter Form (Hydration) clusterweit verteilt und anschliessend durch Deserialisierung (Dehydration) auf einem anderen Server weiter ausgeführt werden. Dies bedeutet die Umsetzung sehr gut skalierbarer SOA Infrastrukturen.

Aber auch im Betrieblichen Umfeld ergeben sich neue Anwendung wie beispielsweise die Möglichkeit Backup und Recovery in Echtzeit durchzuführen und die sonst so leidigen Wartungsfenster vollständig zu eliminieren.

4.7 Grid und Cloud Computing

Während Grid Computing seit 15 Jahren ein fester Begriff ist, hat sich Cloud Computing als Paradigma für die Bereitstellung von Computing Power als global und universell verfügbare Ressource, die jederzeit und überall abgerufen werden kann, erst in den letzten Jahren etabliert. IBM sieht sogar eine direkte Entwicklungslinie von Grid Computing über Utility Comping und Software as a Service zu Cloud Computing. Tatsächlich versuchen Grid und Cloud Computing dieselbe Vision von John McCarthy – inzwischen emittierter Standford Professor – aus den 60er Jahren wahr werden zu lassen: **„Eines Tages könnte Rechenleistung als öffentliche Versorgung organisiert werden“**, wie er in seiner Rede zum hundertjährigen Jubiläum des MIT 1961 formulierte.

Aus heutiger Sicht sind die Unterschiede offensichtlich. Grid Computing geht von der Abstrahierung komplexer Ressourcen, die grosse Datenbestände, umfangreiche und komplexe Funktionalität oder ganze Verbände von Rechenzentren aus, während Cloud Computing von der Virtualisierung kleinerer Einheiten ausgeht. Grid Computing bewältigt einzelne aber sehr rechen- und datenintensive Aufgabenstellungen während Cloud Computing auf viele gleichzeitige auszuführende einfache Anfragen ausgerichtet ist. Beide nutzen verteilte Ressourcen aber in einer anderen Art und Weise.

Allerdings gibt es bereits heute Grid Computing Infrastrukturen, die Cloud Computing Angebote – wie etwa die Elastic Cloud von Amazon – nutzen. Und es gibt eine Reihe von Problemstellungen wie beispielsweise die Notwendigkeit eine Vielzahl einzelner Ressourcen zu Verwalten oder Mechanismen zum Auffinden, Anfragen und Nutzen dieser Ressourcen, die beiden Technologien gemeinsam sind.

5 Cloud Computing Referenzarchitekturen

5.1 Einleitung

Cloud Computing Referenzarchitekturen abstrahieren den konkreten Aufbau einer bestimmten Umsetzung, um die zentralen Elemente und die Schichtung, die eine Cloud ausmachen, zu illustrieren. Zusätzlich definieren viele Referenzarchitekturen noch die Aktoren und deren Rollen, um klarzustellen, wie die Aufgaben, Kompetenzen und Verantwortungen (AKV) der beteiligten Personen oder Organisationen (Unternehmen, Behörden) verteilt sind.

Eine Cloud Computing Referenzarchitektur in seiner einfachsten Form kann als Kombination aus standardisierten Services – aufgebaut wie eine SOA – und einer virtualisierten Infrastruktur gesehen werden, die mit Mechanismen für die Lieferung, Überwachung und Verrechnung erweitert wird.

Eine Cloud Computing Referenzarchitektur kann jedoch auch als Standard angesehen werden wie es die Behörden oder Hersteller tun. Ein solcher Standard definiert die Aktoren und die Bausteine und erlaubt damit eine genaue Verteilung von organisatorischen und Technischen AKV.

Referenzarchitekturen sind in jedem Fall eine Orientierungshilfe, wenn der Aufbau eines Cloud Angebots genau verstanden werden soll. Und sie helfen ein bestehendes oder geplantes Cloud System auf seine Vollständigkeit hin zu prüfen.

5.2 Einführendes Beispiel

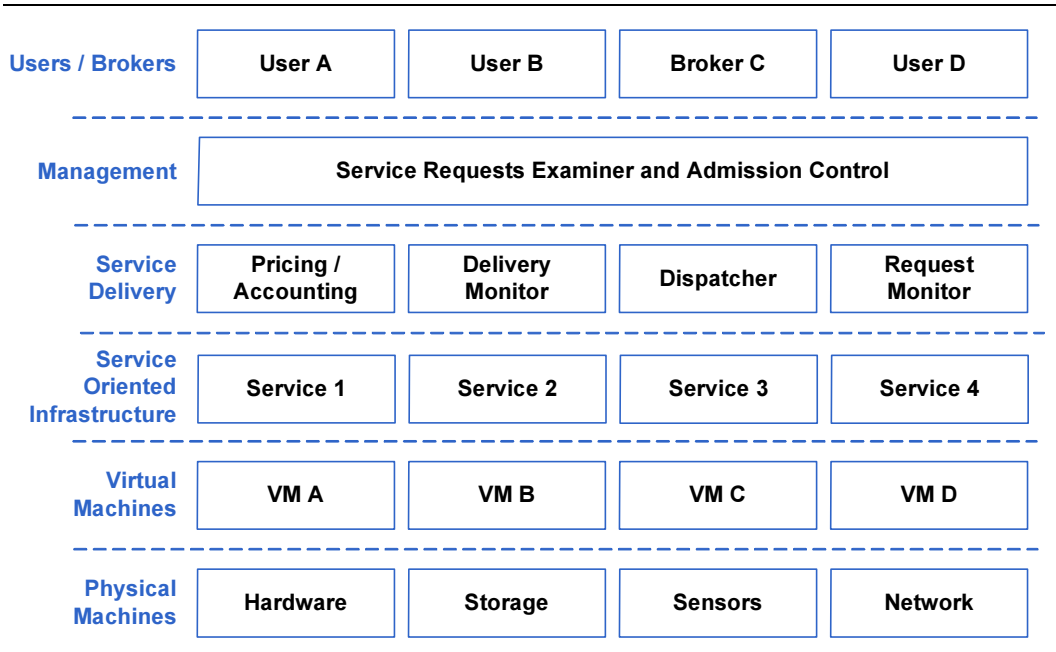

Abbildung 15: Grid Computing Architecture von CLOUDS

Eine einfache Cloud Computing Architektur direkte Kombination von Konzepten aus der Virtualisierung und SOA hat das CLOUDS (Cloud Computing and Distributed Systems) Laboratory der University of Melbourne entwickelt. Es besteht aus den Ebenen Physical Machines, Virtual Machines, Service Oriented Infrastructure, Service Delivery und Cloud Management:

- **Management:** Die Management Ebene ist für die Prüfung der Anfragen auf Zulässigkeit sowie für die Gesamtsteuerung zuständig.
- **Service Delivery:** Die Service Delivery Ebene stellt Mechanismen für die Verrechnung, die Verteilung und die Überwachung zur Verfügung.
- **Service Oriented Infrastructure:** Die Ebene der Service Oriented Infrastruktur stellt Services zur Verfügung, die neben der technischen Schnittstelle und der Service Implementierung auch die Service Infrastruktur, den SLA und die entsprechende betriebliche Serviceorganisation vorsehen.
- **Virtual Machines:** Die Ebene der virtuellen Maschinen ist für die situative Bereitstellung der angeforderten Ressourcen verantortlich.
- **Physical Machines:** Die Ebene der physischen Maschinen enthält die Datenzentren oder Sensoren als Basisressourcen-

Nutzer und / oder Broker der Cloud setzen Cloud Service Anfragen ab, die von der Cloud Infrastruktur abgearbeitet werden. Die Schichtung ist wie die Schichtung einer SOA nicht streng hierarchisch zu verstehen. Dies bedeutet, dass ein Service aus der „Schicht Service Oriented Infrastructure“ beispielsweise den Dispatcher aus der darüber liegenden Schicht „Service Delivery“ Aufrufen kann.

5.3 NIST Referenzarchitektur

Die NIST Referenzarchitektur ist im Jahr 2011 als so genannte „Recommendations“ der amerikanischen Behörden – den U.S. Departement of Commerce, also der Handelskammer - veröffentlicht worden. Damit kommt diesem Standard weitreichende Bedeutung zu und gehört zu den wichtigsten Cloud Referenzarchitekturen überhaupt. So basieren beispielsweise die Referenzarchitekturen von IBM und Oracle darauf.

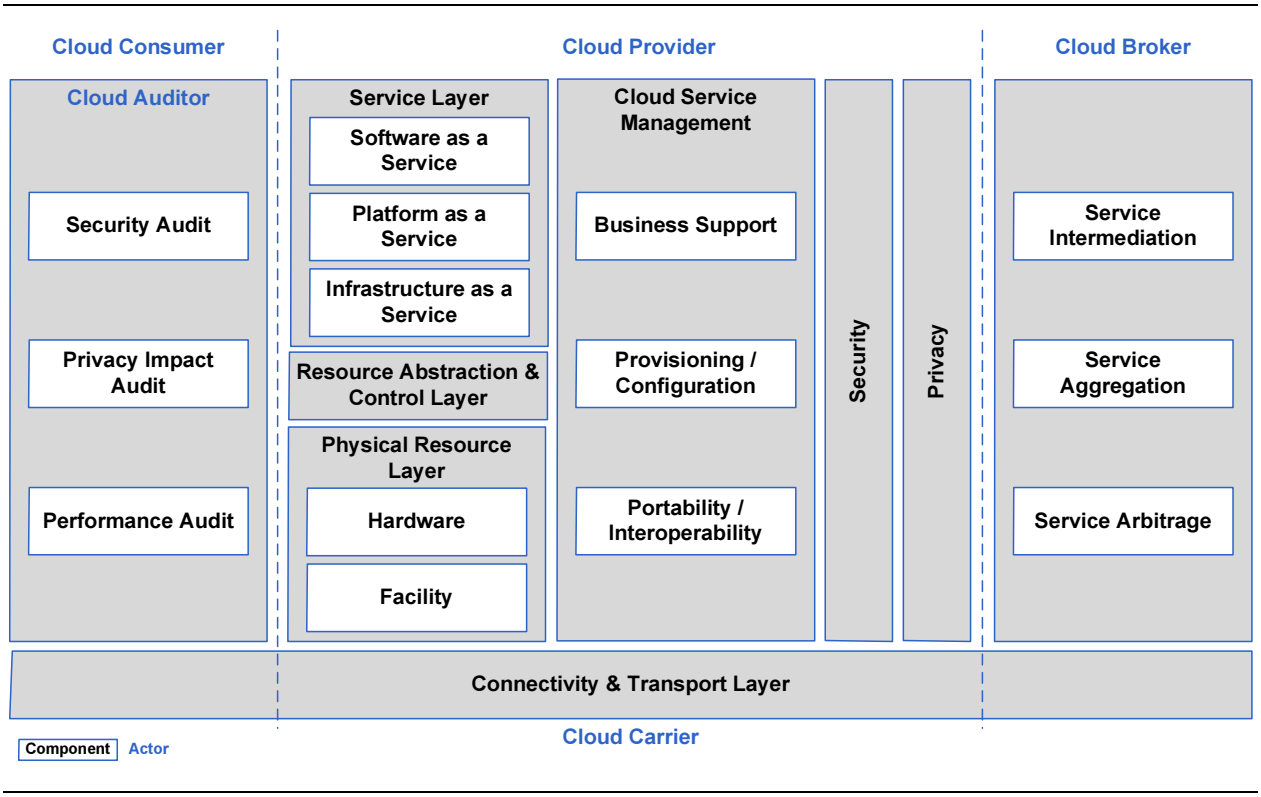

Abbildung 16: NIST Cloud Computing Reference Architecture (leicht angepasst)

Die NIST Cloud Computing Reference Architecture beschreibt den logischen Aufbau eines Cloud Angebotes bestehend aus den Bausteinen und den beteiligten Aktoren wie beispielsweise dem Anbieter eines Cloud Services (Cloud Provider) und dem Kunden dieses Dienstes (Cloud Consumer) [Liu et al. 2011]

- **Component:** Ein einzelner Baustein beschreibt einen technischen oder organisatorischen Dienst, der durch einen Aktor erbracht wird.
- **Actor:** Ein Aktor nimmt im Rahmen eines bestimmten Szenarios eine definierte Rolle ein.

Der Standard definiert für die Bereitstellung von Cloud Diensten durch den Anbieter die verschiedenen Service Modelle Private Cloud (On-site und Out-Sourced), Communicty Cloud (On-site und Out-Sourced) und Hybrid Cloud.

5.4 Die Bausteine der NIST Referenzarchitektur

Die Bausteine der Architektur haben in erster Linie eine Ordnungsfunktion. Das bedeutet, dass bestimmte Funktionen bestimmten Bausteinen zugeordnet werden müssen, damit die gesamte Infrastruktur in einer für alle Beteiligten einheitlichen Art und Weise funktioniert. Die logischen Bausteine definieren nicht im Detail, wie genau die Umsetzung zu realisieren ist. Aber sie sind im Sinne einer SOA aufgebaut und bestehen aus einer Service Schnittstelle (Service Interface), einer Service Beschreibung (Service Contract) und einer Service Realisierung (Service Implementation). Sie können also als Software – im Idealfall – oder als organisatorischer Dienst oder als Kombination von Beidem umgesetzt werden.

5.4.1 Service Orchestration

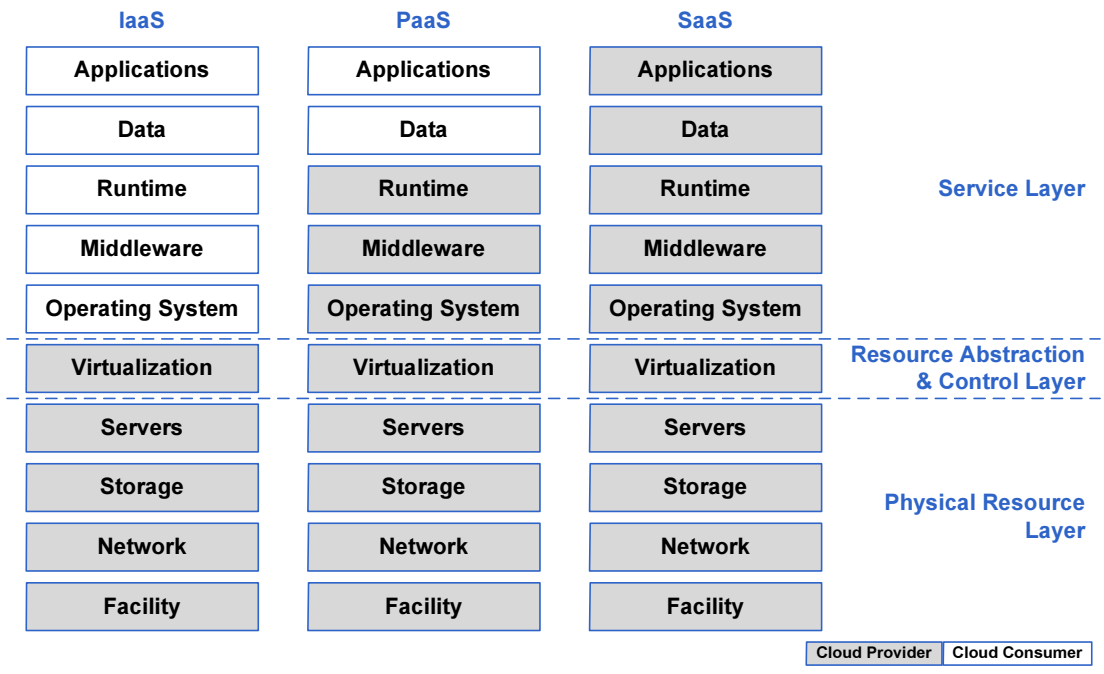

Abbildung 17: Cloud Service Orchestration Varianten

Die zentralen Bausteine jeder Cloud Lösung werden durch den Cloud Provider (Anbieter) bereitgestellt und basieren auf den drei Schichten Service Layer, Resource Abstraction & Control Layer und Physical Resource Layer. Da jedes Cloud Angebot aus den logischen Bausteinen dieser drei Schichten aufgebaut sein muss, werden diese Schichten Service Orchestration genannt.

- **Service Layer:** Diese Schicht stellt die Schnittstelle zur Nutzung der Cloud als **IaaS** (Infrastructure as a Service), **PaaS** (Plattform as a Service) oder **SaaS** (Software as a Service) dar.
- **Resource Abstraction & Control Layer:** Diese Schicht enthält Mechanismen für die Kontrolle und den Zugriff auf physische Ressourcen (des Physical Resource Layers) zuständig sind. „Resource Abstraction“ bedeutet in diesem

Fall die Nutzung verschiedener Virtualisierungs-Techniken wie in Kapitel 3 beschrieben. Kontrolliert werden die Ressourcen durch die Steuerung der Allokation, der Kontrolle der Zugriffe und dem Monitoring der Nutzung.

- **Physical Resource Layer: Hardware** – Computer, Netzwerk - und andere Bestandteile eines Rechenzentrums (**Facitily**) sind auf diesem Layer zu finden.

5.4.2 Service Management

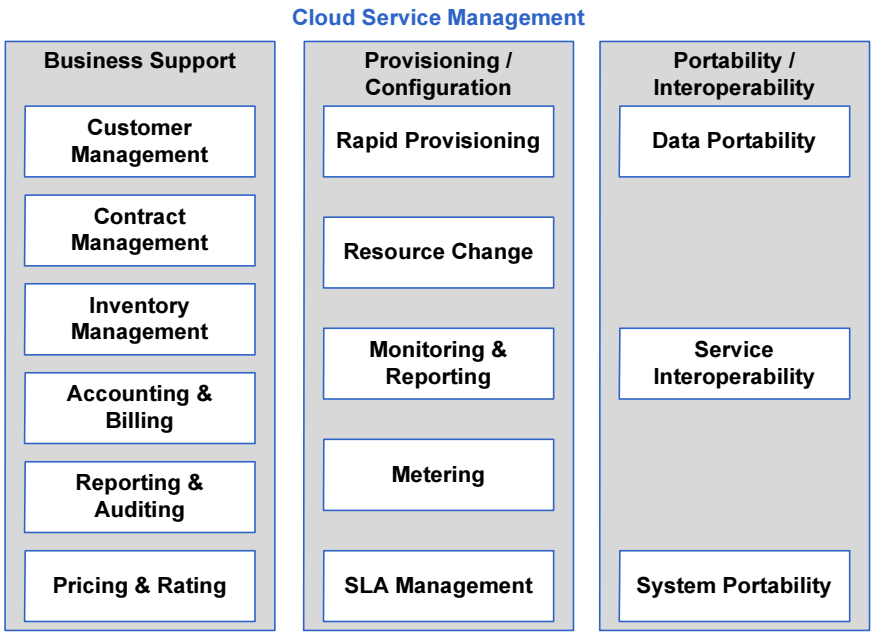

Abbildung 18: Cloud Service Management

Die Verwaltung und der Betrieb eines Angebotes, welches aus einer bestimmten Konstellation von Bausteinen im Rahmen einer Service Orchestration besteht, erfolgt im Rahmen des Service Management. Sämtliche Dienste können in die Bereiche Business Support (Geschäftsunterstützung), Provisioning / Configuration (Bereitstellung und Konfiguration) und Portability / Interoperability (Portabilität und Interoperabilität) unterteilt werden.

- **Customer Management:** Die Kundenverwaltung umfasst Tätigkeiten wie beispielsweise CRM (Customer Relationship Management), SPOC (Single Point of Contact) und anderen.
- **Contract Management:** Die Vertragsverwaltung vom Angebot über die Verhandlung, der Unterschrift bis hin zur Auflösung eines Vertrages.
- **Inventory Management:** Die Bereitstellung und Verwaltung des Angebots in Form eines Servicekatalogs.
- **Accounting & Billing:** Die Verrechnung der Leistungen auf Basis der Vertragskonditionen inklusive der Erbringung eines Leistungsnachweises in Form von Reports.
- **Pricing & Rating:** Die Preisgestaltung und eventuelle Rabattierung von Diensten.
- **Rapid Provisioning:** Die schnelle und automatisierte Bereitstellung von Cloud Diensten.
- **Resource Changing:** Anpassung der Konfiguration aufgrund von Reperaturarbeiten, Upgrades oder Ausbau der Infrastruktur.
- **Monitoring & Reporting:** Überwachung der virtuellen Ressourcen im laufenden Betrieb.
- **Metering:** Bereitstellung von Messmöglichkeiten basierend auf gegebenen Servicetypen wie beispielsweise verwendeter Speicher, genutzte Bandbreite oder Anzahl User.
- **SLA Management:** Die Verwaltung der SLA (Service Level Agreement) und der entsprechenden QoS (Quality of Service)) Paramenter
- **Data Portablity:** Die Möglichkeit Datenobjekte in die Cloud respektive aus der Cloud heraus zu kopieren sowie die Nutzung von Mechanismen für den Transfer von Massendaten.
- **Service Interoperability:** Einheitliche Nutzung von Cloud Services verschiedenster Anbieter.

- **System Portability:** Die Möglichkeit, eine bestimmt VM von einem Anbieter weg zu einem anderen Anbieter hin zu transferieren.

5.4.3 Sicherheit und Datenschutz

Der Sicherheit und dem Datenschutz kommt in der NIST Referenzarchitektur eine besondere Rolle zu. Auf der einen Seite sind sämtliche Aktoren dafür verantwortlich, dass die Sicherheit ihrer Bausteine gewährleistet werden kann. Das bedeutet, dass jeder Service Mechanismen für Authentifizierung und Authorisierung und die Gewährleistung von Verfügbarkeit, Vertraulichkeit, und Integrität bereitstellen müssen. Im Weiteren muss die Verwaltung von Identitäten und Sicherheitsregeln garantiert werden und es muss möglich sein, die Sicherheit zu überwachen und zu prüfen. Der Datenschutz wird durch den Schutz von persönlichen und von Persönlichkeit identifizierenden Informationen gewährleistet.

5.4.4 Aktoren: NIST Cloud Definitionen

Das amerikanische National Institute of Standards and Technology (NIST) definiert die Aktoren und Bausteine sehr genau.

Aktor	Beschreibung
Cloud Consumer	Der Nutzer oder Kunde einer Cloud ist eine Person oder eine Organisation, welche eine Geschäftsbeziehung zu einen Anbieter hat oder welcher dessen Dienste nutzt.
Cloud Provider	Ein Anbieter einer Cloud ist eine Person oder Organisation, die Cloud Dienste Nutzern zugänglich macht.
Cloud Carrier	Ein Mittelsmann – typischerweise ein oder mehrere Netzwerkbetreiber – der die Konnektivität und Transport von Cloud Services zwischen Kunde und Anbieter garantiert.
Cloud Broker	Ein Mittelsmann – typischerweise eine bestimmte Art von Anbieter oder eine interne Organisationseinheit – die zwischen Anbieter und Kunde steht und das Aushandeln und die Verwaltung von Cloud Dienste übernimmt.
Cloud Auditor	Eine Person oder eine Organisation, die eine unabhängige Beurteilung von Cloud Services durchführen kann.

5.5 IBM CCRA

Die IBM Cloud Computing Reference Architecture (CCRA) ist gemäss IBM ein technischer Blueprint, der ein Cloud Computing Angebot abstrakt beschreibt [Stifani et al. 2012]. Die Architektur basierte auf den bestehenden Angeboten der IBM und ist in den Jahren 2009 bis 2012 bis hin zu einer Version 3.0 entwickelt worden, die als Vorschlag für einen Standard an die Open Group, einem Konsortium für Industriestandards in der IT welches unter anderem hinter dem Architekturstandard TOGAF steht, eingereicht wurde.

Im Unterschied zur NIST Referenzarchitektur sieht die IBM lediglich drei verschiedene Aktoren vor. Den Anbieter (Cloud Service Provider) und den Kunden (Cloud Service Consumer) und zusätzlich den Cloud Service Creator. Der Ersteller eines Cloud Services wird mit speziellen Werkzeugen ausgestattet, um Cloud Anwendungen zu entwickeln.

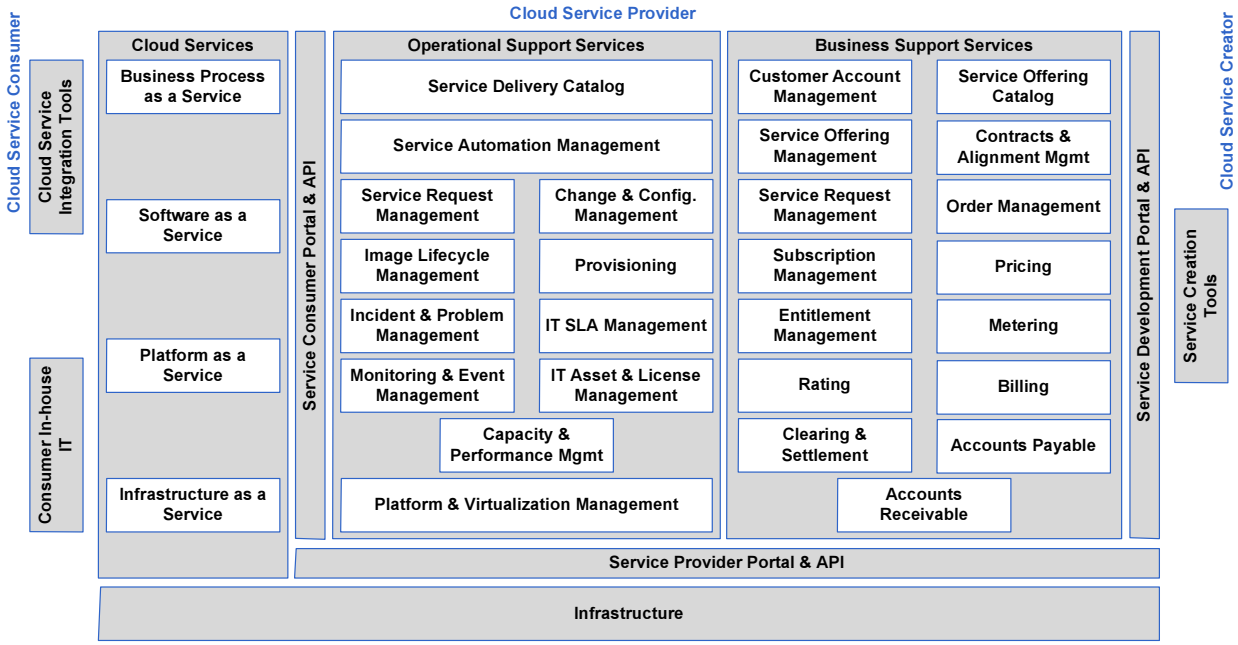

Abbildung 19: IBM CCRA

Die IBM Referenzarchitektur ist auf den Anbieter fokussiert und definiert die verschiedenen Dienste, die für die Bereitstellung eines Angebotes notwendig sind, sehr detailliert. Dabei werden zwischen Cloud Services, Operational Support Services und Business Support Services, der eigentlichen Infrastruktur und den unterschiedlichen Portalen für den Zugriff unterschieden.

- **Cloud Services:** Umfasst die verschiedenen auf einander aufbauenden Cloud Angebotsvarianten IaaS, PaaS, SaaS und zusätzlich noch Business Process as a Service (BPaaS).
- **Service Consumer Portal & API:** Die Schnittstelle und das Portal vom Kunden zum Anbieter eines Cloud Services.
- **Operational Support Services:** Dienste die ein Anbieter für den Betrieb seiner Plattform benötigt.
- **Business Support Services:** Geschäftsunterstützende Dienste des Cloud Anbietes.
- **Service Provider Portal & API:** Die Schnittstelle und das Portal vom Anbieter eines Cloud Services zum Anbieter von Rechenzentrumsinfrastrukturen mit Servern, Speichersystemen und Netzwerkkomponenten.
- **Infrastructure:** Server, Storage, Netzwerk, Gebäudeinfrastrukturen.
- **Service Deleopment Portal & API:** Die Schnittstelle und das Portal vom Anbieter von Cloud Services für die Entwicklung von Cloud basierten Lösungen.

Aus Sicht der IBM sind Sicherheit, Elastizität und Geschwindigkeit übergreifende Elemente, die alle Aktoren der Referenzarchitektur betreffen [IBM 2011]. Und damit das Gesamtbild auch vollständig ist, fehlt im Modell auch nicht die Governance. Auch sie umfasst sämtliche Aktoren und Bausteine der Architektur.

5.6 Oracle CRA

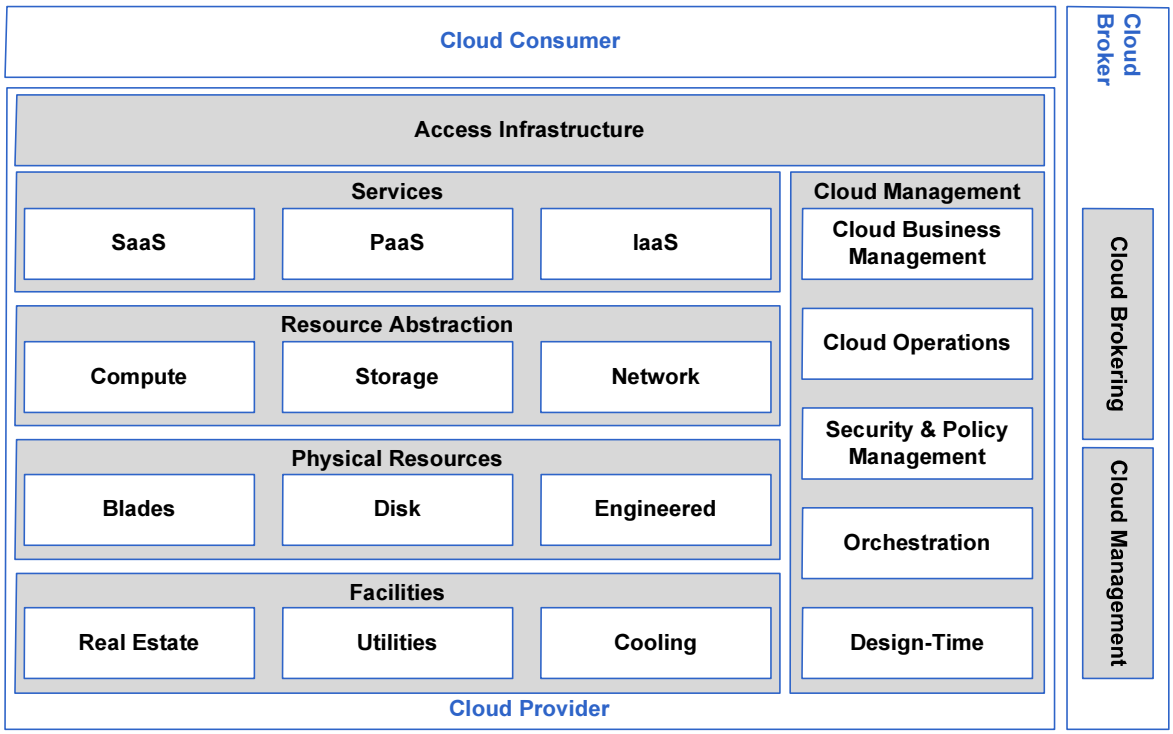

Abbildung 20: Oracle CRA

Die Oracle Cloud Reference Architecture (CRA) geht von den drei Aktoren Kunde, Anbieter und Broker aus und lehnt sich im Aufbau seiner Bausteine an die NIST Referenzarchitektur an.

Die Referenzarchitektur definiert darüber hinaus eine Reihe von typischen Services, die im Rahmen von IaaS, PaaS oder SaaS Angeboten bereitgestellt werden können [Oracle 2012]:

- **IaaS Bausteine:** Server, Storage, Network, Virtualization, Solaris, Linux
- **PaaS Bausteine:** Database, Data Integration, Business Analytics, SOA, BPM (Business Process Management), EDA (Event Driven Architecture), User Interaction, Management, IAM
- **SaaS Bausteine:** Financials, Sales & Marketing, Supply Chain Management, Human Capital Management, Governance, Risk & Compliance, Product Lifecycle Management, Procurement, Project Portfolio Management, Service Management.

5.7 Microsoft Windows Azure Referenzarchitektur

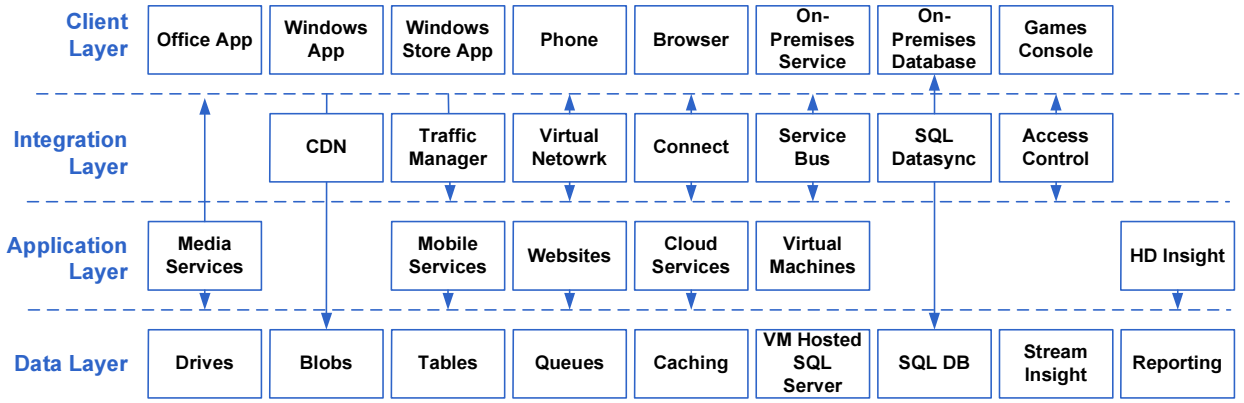

Abbildung 21: Microsoft Azure Referenzarchitektur [Sirtl 2012]

Die Microsoft Azure Referenzarchitektur sieht eine Schichtung von Cloud Services in Data Layer, Application Layer, Integration Layer und Client Layer vor und orientiert sich am Aufbau einer mehrschichtigen Anwendung.

- **Data Layer:** Dienste zur sicheren und verlässlichen Speicherung von Daten.
- **Application Layer:** Services, die skalierbare und hochverfügbare Ausführungsumgebungen für Anwendungslogik bieten.
- **Integration Layer:** Dienste zur Integration von Azure Cloud Services mit Anwendungen oder Infrastrukturkomponenten der internen IT eines Unternehmens.
- **Client Layer:** Anwendungen und Services, die in der unternehmensinternen IT ausgeführt werden.

5.8 Swiss ICT: Cloud Architecture Blueprint

Die Swiss ICT, der Dachverband der schweizerischen ICT Branche, hat im Jahr 2013 den „Leitfaden Cloud-Architektur“ herausgebracht, der sich an „Führungskräfte und Entscheidungsträger mit technischem Verständnis, welche direkten Einfluss auf die IT-Strategie ihres Unternehmens haben“ richtet [Schmid et al. 2013].

Motivation des Leitfadens:

> **Das Thema Cloud-Computing-Architektur in kompakter und prägnanter Art aufbereiten und dies mit Blick auf den Schweizer Markt. Dieser Leitfaden, welcher insbesondere auch den technischen Bereich des Themas abdeckt, soll dazu beitragen, Unsicherheiten abzubauen und damit die Adaptierung von Cloud-Computing zu fördern.**

Die Cloud Architektur, wie sie im Leitfaden definiert ist, definiert zwar keine Aktoren im Sinne der anderen Referenzarchitekturen, es werden jedoch sämtlicher Komponenten einer klaren Schichtung zugewiesen, was den Aufbau einer Cloud-basierten Anwendung sehr gut illustriert.

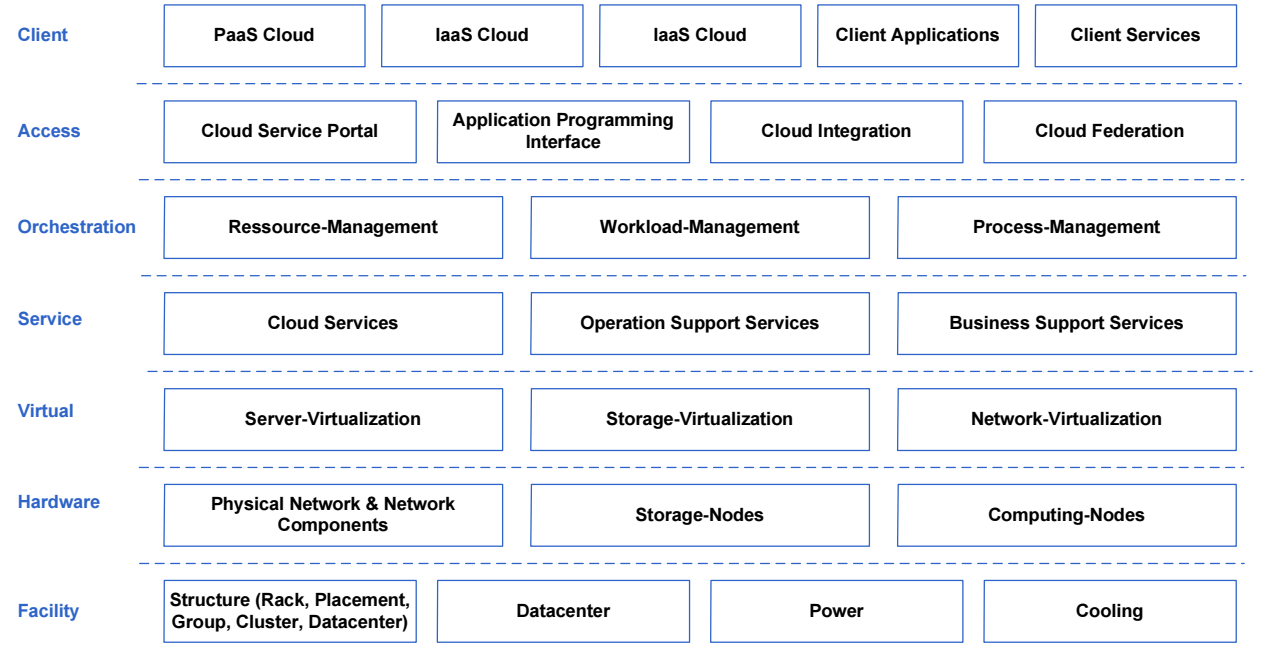

Abbildung 22: Der Cloud Architecture Blueprint [nach Schmid 2012].

Der Cloud Architecture Blueprint ist von Marco Schmid im Jahr 2012 im Rahmen der Bachelorarbeit an der ZHAW entwickelt worden. Auf diesem Blueprint basiert der Leitfaden und er sieht 7 Schichten vor.

- **Client Layer:** Diese Schicht wird durch den Kunden der Cloud realisiert. Die Information über die Nutzungsart ist jedoch für den Anbieter von Services sehr wohl relevant, da die angebotenen Cloud Services entsprechend optimiert werden können.
- **Access Layer:** Der Zugriff auf Cloud Services ist über ein Portal, über ein API, über eine spezielle Cloud Integration oder jedoch über Cloud Federation Mechanismen möglich.
- **Orchestration Layer:** Diese Schicht ist für den Aufbau eines bestimmten Cloud Angebotes für einen oder mehrere Kunden zuständig. Services und physische und virtuelle Ressourcen werden zu einem funktionierenden Ganzen zusammengestellt. Ausserdem werden die Abläufe und Regeln für den laufenden Betrieb definiert.
- **Service Layer:** Der Blueprint sieht drei Gruppen von Services auf dieser Ebene vor: Die Cloud Services, die Operation Support Services und die Business Support Services.
- **Virtual Layer:** Diese Schicht stellt virtuelle Ressourcen zur Verfügung. Sie abstrahiert Hardware und Facility zu skalierbaren Einheiten, die je nach Bedarf alloziert werden können.
- **Hardware Layer:** Der Hardware-Layer enthält Hardware-Komponenten, die virtualisierbar sind. Es kann zwischen Server Nodes, Storage Nodes und Network Components unterschieden werden.
- **Facility Layer:** Der Faciltiy Layer enthält die typischen Komponenten eines Datacenter-Bauwerks sowie die Art und Weise der Strukturierung innerhalb des oder der Gebäude.

6 Cloud Computing Anbieter

6.1 Einleitung

Alleine in der Schweiz gibt es im Jahr 2014 gemäss Computerwoche über 150 Anbieter, die in diesem Markt tätig sind [Kurzidim 2014]. Sehr viele Angebote basieren auf Cloud Infrastrukturen der drei wichtigsten Anbieter Amazon, Google und Microsoft, Ihnen allen ist gemeinsam, dass sie ein globales Netzwerk von sehr grossen Rechenzentren betreiben. Es kann davon ausgegangen werden, dass jeweils mehr als 1 Million Server im Einsatz sind.

Amazon, Google und Microsoft verfügen also über die notwendige Kapazität, sämtliche Anforderungen einer Cloud Anwendung vollständig abzudecken. Sie sind die drei grossen Player, die den Markt für Cloud Computing in den nächsten Jahren mit aller Wahrscheinlichkeit bestimmen werden.

David Linthicum, einer der führenden Spezialisten in diesem Bereich, bringt es auf den Punkt: „**Try again, cloud contenders: Amazon, Google, and Microsoft have won - IaaS and PaaS markets will no longer support smaller providers, which now need to find new specialties or call it quits**“ [Linthicum 2014].

Die Konsequenzen liegen auf der Hand; Wer eine Cloud Lösung entwickeln, bereitstellen und betreiben möchte, kommt nicht darum herum, den Aufbau und die Funktionsweise der drei Plattformen von Amazon, Google und Microsoft zu kennen.

6.2 Amazon

Amazon gilt als einer der ersten Cloud Computing Anbieter überhaupt. Bereits seit Ende 2006 bietet Amazon IT-Kapazitäten anderen Firmen an. Die AWS (Amazon Web Services) genannten Services sind ein sehr schnell wachsendes Geschäft. Bereits im nächsten Jahr erwartet Amazon knapp sieben Milliarden Dollar Umsatz [Kroker 2014]. Amazon stellt eine Vielzahl von Referenzarchitekturen zur Verfügung, die den jeweiligen Einsatz vorgeben.

Einsatzgebiete sind beispielsweise:

- **Hosting von Webanwendungen:** Bereitstellung einer skalierbaren Infrastruktur bestehend aus Datenbank Servern, Application Server und Web Servern mit Load Balancing und der Möglichkeit ein Content Delivery Network aufzubauen.
- **Bereitstellung von Inhalten und Medien:** Eine Kombination aus Streaming Plattform, skalierbarem Speicher und Serverinfrastruktur für Film, Ton, Bild und Text.
- **Batchverarbeitung:** Die Möglichkeit sehr grosse Datenmengen zu verarbeiten, wie sie beispielsweise im Rahmen von TEV- (Tages Endverarbeitung) oder MEV- (Monats Endverarbeitung) Prozessen üblich sind.
- **Fehlertoleranz und hohe Verfügbarkeit:** Eine redundante Infrastruktur für Anwendungen mit hohen Anforderungen an die Verfügbarkeit und die Fehlertoleranz .
- **Umfangreiche Verarbeitung und riesige Datenmengen:** Siehe Batchverarbeitung
- **Ad-Serving:** Eine Kombination aus Mechanismen sehr schnellen Lieferung von Werbeinhalten auf Basis von kontextbasierten Nutzerprofilen.
- **Notfallwiederherstellung für lokale Anwendungen:** Eine Lösung für BCP (Business Continuity Planning) Fragestellungen.

Darüber hinaus existieren Lösungen für die Dateisynchronisierung, die Freigabe von Medien, Online-Spiele, Protokollanalyse, das Grid Computing für Finanzdienstleister, die Bereitstellung von E-Commerce-Websites und die Zeitreihenverarbeitung.

Sämtliche Angebote werden durch unterschiedliche Zusammenstellung von Amazon Web Services zur Verfügung gestellt.

Wichtige Produkte sind die Elastic Compute Cloud (EC2) und der Simple Storage Service (S3). EC2 stellt eine virtuelle Maschine zur Verfügung, deren Kapazität sehr schnell verändert werden kann. S3 ist ein redundanter und damit hochverfügbarer Speicher, der Nutzungsabhängig verrechnet wird. Beide Produkte sind als Services ausgelegt und können über definierte Schnittstellen automatisch bestellt, bereitgestellt, betrieben und wieder abgestellt werden.

6.2.1 Aufbau der Amazon Web Services

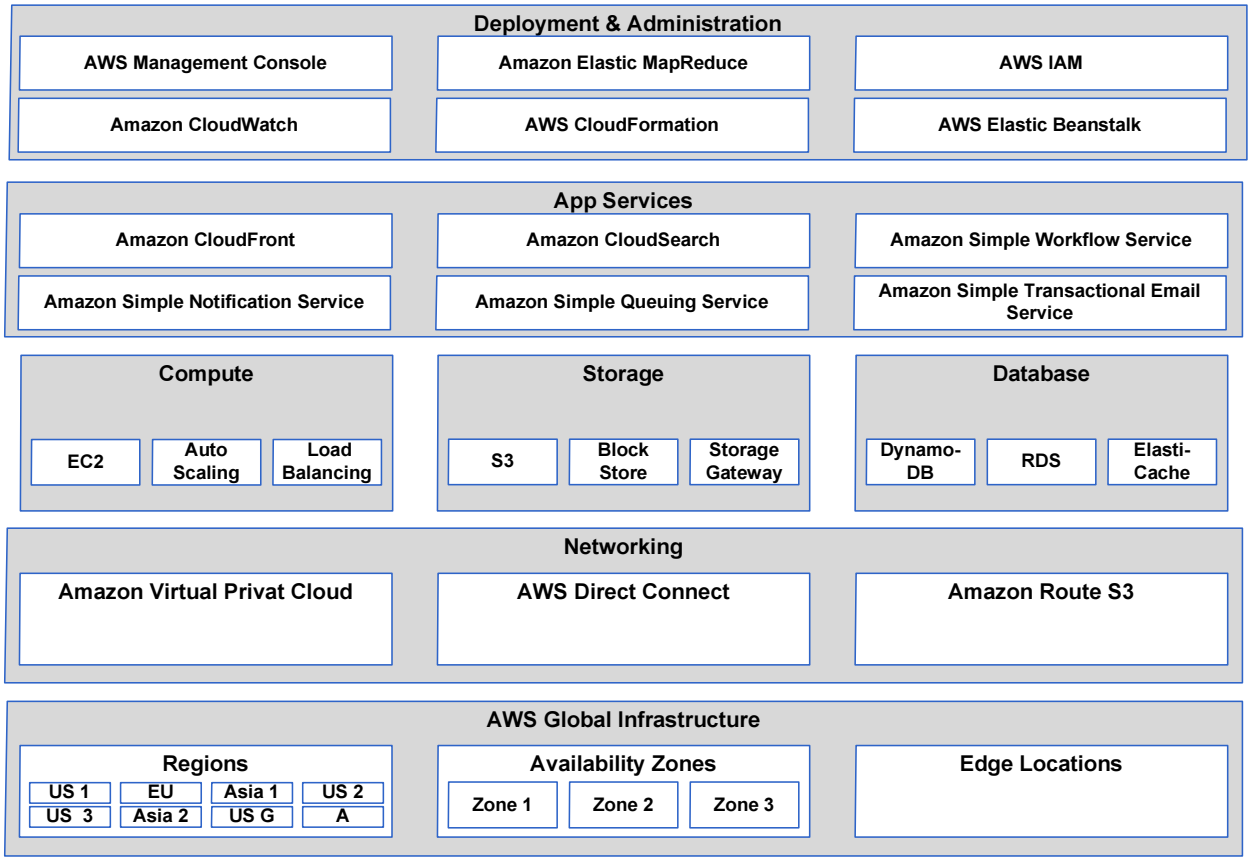

Abbildung 23: Amazon Web Services (nach [Ramuschkat 2012]).

Die Amazon Web Services sind auf sieben logischen Bausteinen aufgebaut [Varia 2011]

- **AWS Global Infrastructure:** Die Infrastruktur umfasst mehr als eine halbe Million Server, die in mindestens 8 Regionen der Welt in verschiedenen Datacentern laufen. Jedes Rechenzentrum verfügt über mindestens eine der drei Verfügbarkeitszonen, die durch unabhängige Strom- und Datenversorgung von anderen Zonen getrennt sind. Die Verteilung von Daten erfolgt über so genannte Edge Locations. Diese Standorte sind weltweit verteilt, was den Aufbau von Content Delivery Networks erleichtert.
- **Networking:** Das Netzwerk verfügt über drei Technologien. Die Virtual Private Cloud unterstützt den Aufbau und den Betrieb virtueller Netzwerke, AWS Direct Connect stellt einen direkten Zugriff für Services zur Verfügung während Amazon Route S3 einen virtuellen DNS (Domain Name Service) bereitstellt.
- **Compute:** Die Computing Services sind EC2 (Elastic Cloud 2), ein Service für die Bereitstellung von Virtuellen Maschinen und Services für die Skalierbarkeit und die Lastverteilung.

- **Storage:** Die drei Storage Services sind Amazon Simple Storage Service (S3, ein universeller redundanter Speicher), Elastic Block Store (Speicher für EC2) und Storage Gateway (für Backups auf S3).
- **Database:** Database Services stehen in drei Varianten zur Verfügung. Als Nicht-Relationale DB (DynamoDB) als Relationale DB (MySQL, Oracle) und als Cache.
- **App Services:** Die Application Services sind CloudFront (Content Delivery), CloudSearch und eine Reihe von einfachen Basisdiensten für Workflows, Meldungsverarbeitung, Queuing und Email.
- **Deployment & Administration:** Die Services für Deployment und Adminstration umfassen eine Management Konsole, einen Big Data Analytics Service (MapReduce), Identity und Access Management, Monitoring, Provisionierungs- und Deployment-Dienste.

6.3 Google

Google hat im Jahr 2008 mit der Google App Engine (GAE) einen Cloud Services auf den Markt gebracht, der auf die Entwicklung und den Betrieb von performanten und skalierbaren Webseiten ausgerichtet ist. Diese Cloud Computing Plattform ist vollständig transparent. Dies bedeutet, dass der Kunde keinen Einfluss darauf hat, wie viele Ressourcen für eine Anwendung zur Verfügung stehen. Die Skalierung und die Lastverteilung erfolgt durch Google aufgrund der jeweiligen Lastsituation im laufenden Betrieb.

Neben der GAE hat Google eine Reihe von weiteren Produkten im Angebot. Die Google Cloud Plattform enthält Instrumente für Entwickler und für die Cloud Service Verwaltung sowie eine Vielzahl von Computing, Storage und anderen Services. Sie ist nach wie vor auf die Entwicklung von Cloud Lösungen ausgerichtet.

Einsatzgebiete sind:

- **Entwicklung mobiler Back-Ends:** Mobile Anwendungen brauchen eine gute Back-End Integration, wenn sie über einfache Funktionen hinausgehen. Google stellt spezielle Mechanismen für die Entwicklung und den Betrieb solcher Plattformen zur Verfügung.
- **Big Data Anwendungen:** Die Verarbeitung sehr grosser Datenmengen aus unterschiedlichsten Datenquellen wird durch spezialisierte Dienste unterstützt.
- **Spiele:** Spiele mit sehr vielen gleichzeitig teilnehmenden Playern haben besondere Anforderungen an die Skalierbarkeit und die Antwortzeiten.

6.3.1 Aufbau der Google Cloud Plattform

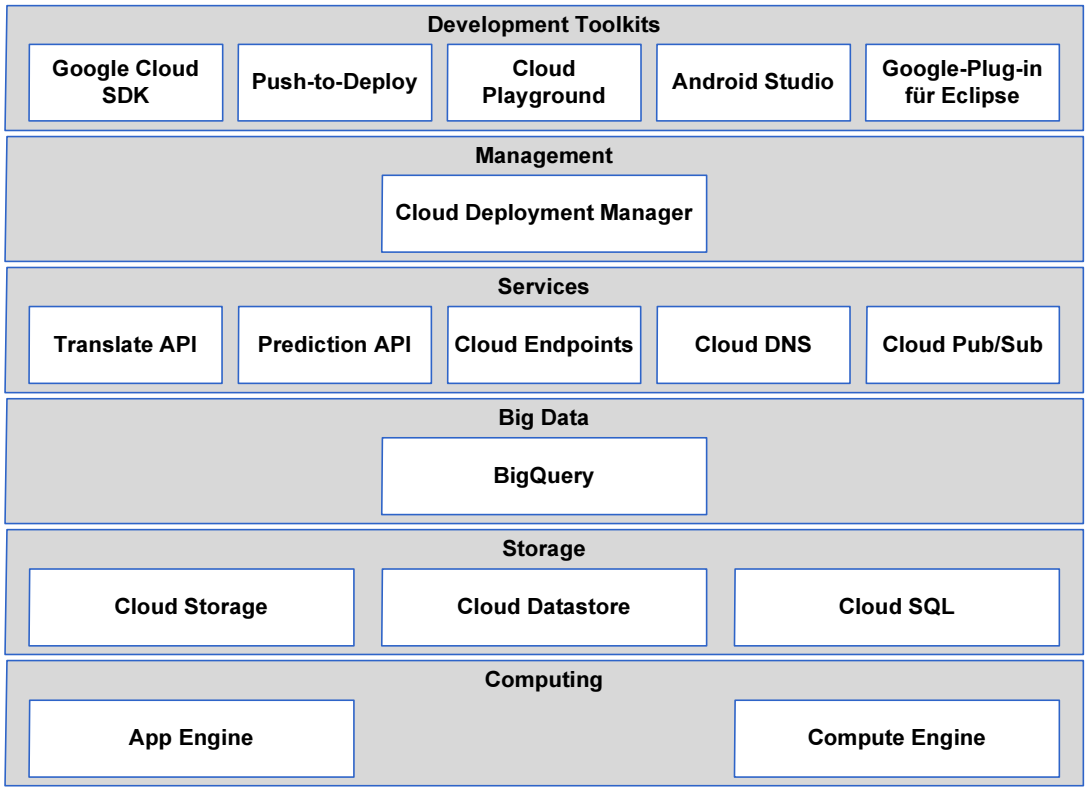

Abbildung 24: Google Cloud Plattform

Die Google Cloud Service Plattform besteht aus 17 verschiedenen Diensten, die den Ebenen Computing, Storage, Big Data, Services, Management und Development Kits zugeordnet werden können.

- **Development Kits:** Die Entwicklung einer Cloud Anwendung wird durch eine Reihe von Werkzeugen unterstützt. Das Google SDK (Software Development Kit), Android Studio und ein Plug-In für Eclipse sind Entwicklungsumgebungen, während Push-to-Deploy und Cloud Playground Instrumente für die Versionierung und das Testing sind.
- **Management:** Die Bereitstellung und Verwaltung einer Lösung wird durch den Cloud Deployment Manager vereinfacht.
- **Services:** Die Dienste umfassen die automatische Übersetzung für mehrsprachige Anwendungen (Translate API), ein API für die Daten-Analyse und die Vorhersage von Ereignissen (Prediction API), ein Messaging Service (Cloud Pub/Sub), die Bereitstellung von sicheren Dienst-Schnittstellen (Cloud Endpoints) und ein DNS Service (Domain Name Service).
- **Big Data:** Ein Dienst, der die Analyse sehr grosser Datenmengen im Terabyte-Bereich unterstützt (BigQuery).
- **Storage:** Services für die Speicherung von Daten als einfacher Datenspeicher (Cloud Storage) oder als NoSQL-Datenbank (Cloud Datastore) oder jedoch als MySQL-Datenbank (Cloud SQL).
- **Computing:** Neben der sehr weit verbreiteten App Engine, dem PaaS-Dienst von Google steht ein IaaS-Dienst (Compute Engine) für VM's zur Verfügung.

6.4 Microsoft

Die Cloud Computing Plattfrom Microsoft Azure ist im Jahr 2010 als Ausweitung des Betriebssystemes Windows in die Cloud auf den Markt gekommen. Die beiden zentralen Bestandteile waren SQL Azure und Windows Azure. Und die Grundidee hinter Windows Azure und SQL Azure war, sämtliche Funktionen des Betriebssystems Windows in die Cloud zu verlagern und damit auch eine Verlagerung sämtlicher Microsoft-Produkte in die Cloud zu ermöglichen [Chappell 2008]. Allen voran war natürlich Microsoft Office im Fokus. Was mit der Lancierung von Microsoft Office 365 erfolgte.

Seit dem 25.3.2014 heisst die Plattform neu Microsoft Azure und stellt Funktionen zur Verfügung, die die Ausführung eigener Anwendungslogik, die Speicherung und Auswertung von Daten, die Sicherheit Azure-basierter Anwendungen, Anwendungsdienste zur Verwendung in eigenen Apps und die Integration und Kommunikation unterstützen [Sirtl 2014].

Microsoft definiert 4 primäre Modelle für die Erstellung und den Betrieb von Anwendungen:

- **Virtual Machines:** Virtuelle Maschinen als Cloud Basisbausteine erlauben die volle Kontrolle über die Bereitstellung und den Betrieb einer Anwendung. Die Funktionalität der VM entspricht der Funktionalität von realen Server. Entsprechend kann eine Anwendung aufgebaut werden. Dies ist die einfachste Variante, Anwendungen in die Cloud zu verschieben.
- **Cloud Services:** Cloud Services unterstützen Anwendungen mit einem Web-basierten Frontend und unterscheiden zwischen zwei Arten von Virtuellen Maschinen; Cloud Service mit einem vorinstallierten ISS (Internet Information Service – der Microsoft Web Server) und solche ohne vorinstallierten ISS.
- **Web Sites:** Für reine Web Anwendungen basierend auf vorinstallierten Cloud Services (z.b. WordPress oder Drupal).
- **Mobile Services:** Bereitstellung von Cloud Services als Back-End für mobile Anwendungen.

Unterstützt werden alle vier Modelle durch einen Servicekatalog, der eine Vielzahl von Diensten auf Anwendungs-, Daten-, Berechnungs- und Netzwerk-Ebene enthält.

6.4.1 Aufbau von Microsoft Azure

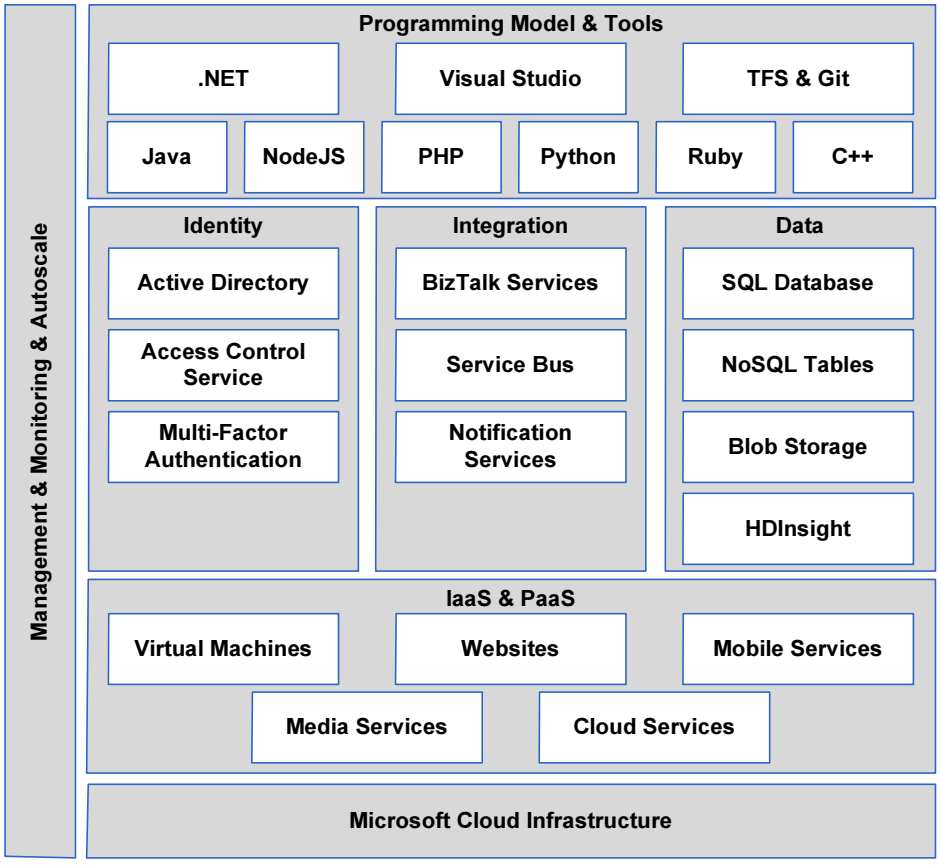

Abbildung 25: Der logische Aufbau von Microsoft Azure

Der logische Aufbau von Microsoft Azure ordnet die verschiedenen zur Verfügung stehend Azure Cloud Services den Bereichen Microsoft Cloud Infrastructure, IaaS & PaaS, Identity, Integration, Data, Programming Model & Tools und

Management & Monitorung & Autoscale zu. Eine bestimme Anwendung ist immer eine Kombination aus Services verschiedenster Bereiche.

- **Microsoft Cloud Infrastructure:** Die Basis für sämtliche Dienste besteht aus Server- und Netzwerkkomponenten, die in Microsoft Rechenzentren stehen.
- **IaaS & PaaS:** Dienste auf Plattform oder Infrastruktur-Ebene können automatisiert übernommen warden. Sie reichen von Virtuellen Maschinen mit Windows- oder Linux-Betriebssystem bis hin zu Services für Mobile Anwendungen, Medienplattformen oder anderen Cloud Services.
- **Identity:** Dienste für das IAM (Identity & Access Management)
- **Integration:** Services für die Integration von Anwendungen und / oder Services
- **Data:** Verschiedene Arten der Datenspeicherung (SQL, NoSQL oder Blob – Binary large object) und einen Service für Big Data Berechnungen
- **Programming Model & Tools:** Services für die Entwicklung von Cloud Services in verschiedensten Programmiersprachen mit einer Vielzahl von Microsoft Entwicklungs-Werkzeugen und anderen Tools.
- **Management & Monitoring & Autoscale:** Die Verwaltung, Überwachung und die automatisierte Skalierung von Diensten.

6.4.2 Der Azure Servicekatalog

Der Azure Servicekatalog entspricht der Aufteilung in der Referenzarchitektur aus Kapitel 5.7.

Funktionsgruppe	Service Name	Beschreibung
Ausführung eigener Anwendungslogik	Websites	Bereitstellung und Betrieb kleinerer Webanwendungen
	Cloud Services	Bereitstellung und Betrieb hoch-skalierbarer Anwendungen
	Virtual Machines	Bereitstellung und Betrieb von virtuellen Maschinen
	Mobile Services	Dienste für das Back-End von Mobile Apps
Speicherung und Auswertung von Daten	Tables	NoSQL-Datenbank
	Blobs	Speicherung grosser und unstrukturierter Daten (Blob – Binary large Object)
	Drives	Persistente Harddisk Drives für Cloud Services
	SQL Database	Relationale Datenbank
	HD Insight	Service für die Analyse und die Berechnung sehr grosser Datenmengen
	Cache Service	Zentral bereitgestellter Cache
	In-Role Cache	Cache innerhalb von Cloud Service VMs
	CDN (Content Delivery Network)	Globales Caching von grossen und unstrukturierten Daten (Content – Film, Bild, Ton)
	Backup Service	Sicherung virtueller Maschinen
	Recovery Service	Wiederherstellung virtueller Maschinen
Sicherheit Azure-basierter Anwendungen	Active Directory	Benutzerverzeichnis
	Multi-Factor Authentication	Multi-Faktor Authentifizierung
Anwendungsdienste zur Verwendung in eigenen Anwendungen	Media Services	Medien-Workflows in der Cloud
	Notification Hubs	Benachrichtigungen an viele Clients
	Scheduler	Zeitgesteuerte Ausführung von Services
	Automation	Automatisiertes Management von Diensten
Integration und Kommunikation	SQL DataSync	Synchronisation von relationalen Datenbanken
	Queues	Nachrichtenversand
	Service Bus	Enterprise Service Bus (ESB)
	BizTalk Services	Integration von Backend-Systemen
	Express Route	Exklusive Netzwerkverbindungen

	Virtual Network	Netzwerk-zu-Netzwerk VPN für virtuelle Netzwerke
	Traffic Manager	Routing von Service-Zugriffen

7 Einsatzgebiete und Risikofaktoren von Cloud Lösungen

7.1 Einleitung

Der Einsatz und die Planung von Cloud Lösungen unterscheiden sich nicht grundsätzlich von der Planung und Einführung konventioneller Informationssysteme. Dies bedeutet, dass – wie in jedem IT-Projekt – nach der Formulierung des Projektauftrages eine Wirtschaftlichkeitsbetrachtung gemacht werden muss. Anschliessend sind - je nachdem, welche Vorgehensweise gewählt wird – die entsprechenden Projektschritte von der Analyse über das Design bis hin zur Implementation oder Konfiguration, dem Testing, der Abnahme und der Einführung sämtliche Schritte zu durchlaufen.

Die Unterschiede zwischen Cloud Computing Projekten und anderen IT Projekten liegen im Detail und betreffen die Art und Weise der Interaktion mit dem Anbieter und eine Reihe von Sicherheitsaspekten, die zu berücksichtigen sind. Und es ist eine Tatsache, dass die Zuordnung der Einsatzgebiete, also die Prüfung, ob sich eine bestimmte Aufgabenstellung besonders für eine Cloud Lösung eignet, noch nicht zum Alltagsgeschäft der IT-Verantwortlichen gehört.

7.2 Einsatzgebiete

7.2.1 Einsatzgebiet 1: Ein einfaches Raster zur Auswahl

Die Fragestellung, ob sich eine bestimmte Aufgabenstellung für eine Lösung in der Cloud eignet, kann mit einem einfachen Fragenkatalog und ein paar grundlegenden Überlegungen beantwortet werden.

Die Kandidaten für eine Cloud sind gegeben, wenn sich mindestens eine der nachfolgenden Fragen mit Ja beantworten lassen:

Nr	Fragestellung
1.	Existieren Anwendungen, die nur sehr selten verwendet werden, die aber dennoch unbedingt am Leben erhalten werden müssen?
2.	Existieren umfangreiche Bild- oder Dokumentarchive?
3.	Gibt es Systeme, die nur sehr selten oder nur periodisch unter Voll-Last laufen, aber auf sehr teuren Infrastrukturen betrieben werden müssen?
4.	Ist die Entwicklung einer Individualanwendung basierend auf Web Technologie mittels einer Cloud Entwicklungsumgebung eine Alternative?
5.	Gibt es vom Hersteller einer Standard-Software ein entsprechendes Cloud Angebot?

Weitere grundlegende Überlegungen, die bei der Auswahl helfen können sind:

Nr	Grundlegende Überlegung
1.	Lösungen, die meist mit Standard-Software realisiert werden, setzen voraus, dass ein entsprechendes SaaS Angebot eines etablierten Herstellers vorliegt.
2.	Lösungen, die auf Individualentwicklungen basieren, eignen sich nur dann, wenn sie mit der entsprechenden PaaS Infrastruktur neu gebaut werden können.
3.	Lösungen, die mit grossen und sich nicht verändernden unstrukturierten Daten arbeiten, eigenen sich mehr als solche, die mit strukturierten oder kritischen Daten oder mit grossen Bewegungsdaten arbeiten.

7.2.2 Einsatzgebiete 2: Was Anbieter empfehlen

Die Einsatzgebiete, die die grossen Anbieter von Cloud Plattformen empfehlen, sind eine sehr gute Grundlage für den Entscheid, eine Anwendung als Cloud Lösung zu konzipieren.

Einsatzgebiet	Beschreibung	Leistung des Cloud Anbieters
Web	Websites, Webanwendungen und E-Commerce	Hosting einer skalierbaren Infrastruktur bestehend aus Datenbank Servern, Application Server und Web Servern mit Load Balancing
Content Streaming	Globales Publishing von Inhalten und Medien	Bereitstellung und Betrieb hoch-skalierbarer Content Delivery Networks für sehr schnelles Streaming
Big Data	Verarbeitung und Analyse sehr grosser Datenmengen	Spezialisierte Dienste, die die Verarbeitung und Analyse sehr grosser Datenmengen erlauben
Mobile Back-Ends	Services für mobile Anwendungen	Dienste für das Back-End von Mobile Apps
Datensynchronisierung	Backup und BCP Lösungen	Einfache und sicheres Backup kritischer Daten sowie Mechanismen für die Notfallwiederherstellung ganzer Anwendungen
Spiele	Blobs	Speicherung grosser und unstrukturierter Daten (Blob – Binary large Object)

7.2.3 Einsatzgebiete 3: Empfehlungen des SATW

Die Schweizerische Akademie der Technischen Wissenschaften (SATW) hat in einem Whitepaper zum Thema Cloud Computing die Vor- und Nachteile der verschieden Cloud Angebote gegenüber dem internen Betrieb von entsprechenden Diensten auf der Ebene SaaS, PaaS und IaaS herausgearbeitet [Brian et al. 2012].

Cloud Ebene	Vorteil gegenüber internem Betrieb	Nachteil gegenüber internem Betrieb
Software as a Service	■ Trennbarkeit / Mandantenfähigkeit der Applikationen ■ Schnell einsatzfähig/schnellere Projekteinführung (time to market) ■ Keine Maintenance für den Betrieb der Business-Funktionalitäten ■ Niedrigere Gesamtkosten (TCO) ■ Mobilität/Standortunabhängigkeit	■ Auswahl des richtigen Providers ■ Fehlende Portabilität ■ Geringere Integrierbarkeit in bestehende Applikationslandschaften ■ Geringere Anpassungsmöglichkeiten, da vorgegebene Standardisierung ■ Eventuell höhere Antwortzeiten ■ Auswirkung von Sicherheitslücken beim Einsatz gemeinsamer SaaS-Lösungen ■ Keine Nutzung ohne Internetzugang
Plattform as a Service	■ Weniger Administrationsaufwand, da die notwendige Infrastruktur nicht selbst implementiert und bereitgestellt werden muss ■ Entwicklung im Team (auch geographisch verteilt) ■ Eine einzige Plattform mit minimalen Kosten (Standardisierung) ■ Keine Maintenance für Einrichtung und Betrieb der Plattform und deren Tools	■ Vendor Lock-in ■ fehlende Portabilität ■ fehlende Interoperabilität ■ keine standardisierten Technologien ■ Mangelnde Flexibilität ■ Anforderungen von proprietären Anwendungen oder Entwicklungsumgebungen
Infrastructure as a Services	■ Hohe Skalierbarkeit der benötigten Systeme, je nach benötigtem Bedarf ■ Redundante Datenspeicherung ■ Physisch getrennte Aufbewahrung und Nutzung von Daten	■ Standort der Daten bei Public wie auch bei Private Clouds nicht immer erkennbar ■ Stark abhängig von der Verfügbarkeit der Infrastruktur und Netzwerke ■ Fehlende oder mangelnde Abgrenzung/Isolierung der Datenbearbeitungen

	■ Keine Maintenance für Einrichtung und Betrieb der Infrastruktur	■ Unberechtigter Datenzugriff auf Grund einer Fehlkonfiguration. In der Annahme, dass Rechenzentren oder Netzwerke geteilt werden, ist dieses Risiko anders einzuschätzen als im Eigenbetrieb. ■ Gewährleistung und Haftung bei Verletzung der Vertraulichkeit, Sicherheit und Integrität der Daten
Sämtliche Ebenen	■ OPEX statt CAPEX ■ Pay as you go	

7.2.4 Einsatzgebiete 4: Die Cloud Strategie der Schweizer Behörden

Der Steuerausschuss E-Government hat am 25. Oktober 2012 die „Cloud-Computing-Strategie der Schweizer Behörden 2012 – 2020“ verabschiedet [ISB 2012]. Darin sind die strategischen Grundsätze und die strategischen Stossrichtungen für den Einsatz von Cloud Computing Lösung in der öffentlichen Verwaltung festgelegt.
Die strategischen Grundsätze umfassen bemerkenswerte Aussagen wie Beispielseise diejenige „Cloud first“:

> **Bei Neuentwicklungen und Anschaffungen wird systematisch geprüft, ob geeignete Cloud-Angebote vorhanden sind. Eine Cloud-Lösung wird gewählt, wenn sie die Gesamtheit der Anforderungen insbesondere an Funktionalität, Wirtschaftlichkeit und Sicherheit über alles betrachtet am besten abdeckt. Bestehende Lösungen werden dann vorzeitig durch eine Cloud-Lösung ersetzt, wenn dies wirtschaftlich sinnvoll ist.**

Sie durch die folgenden 5 strategischen Stossrichtungen unterstützt:

1. Förderung des verantwortungsvollen Cloud-Einsatzes
2. Anpassung der rechtlichen Grundlagen
3. Aufbau von dedizierten Cloud-Angeboten für die Behörden
4. Aufbau von Cloud-Angeboten für Private und Wirtschaft
5. Zusammenarbeit mit Wirtschaft und dem internationalen Umfeld

Im „Kommentar zur Cloud-Computing-Strategie der Schweizer Behörden“ wird genauer beschrieben und / oder eingeschränkt, wie denn die Umsetzung genau erfolgen soll [Müller et al. 2012]. Dabei ist bemerkenswert, dass die Behörden gehalten sind für Führungs-, Kern- und Unterstützende Aufgaben SaaS & BaaS Angebote zu nutzen. Im Bereich der Kernaufgaben mit Vollzugscharakter wie beispielsweise Steuern, Arbeit, Bauen, Bildung, etc. sollen die Behörden SaaS Angebote bereitstellen.

7.3 Risikofaktoren der Cloud

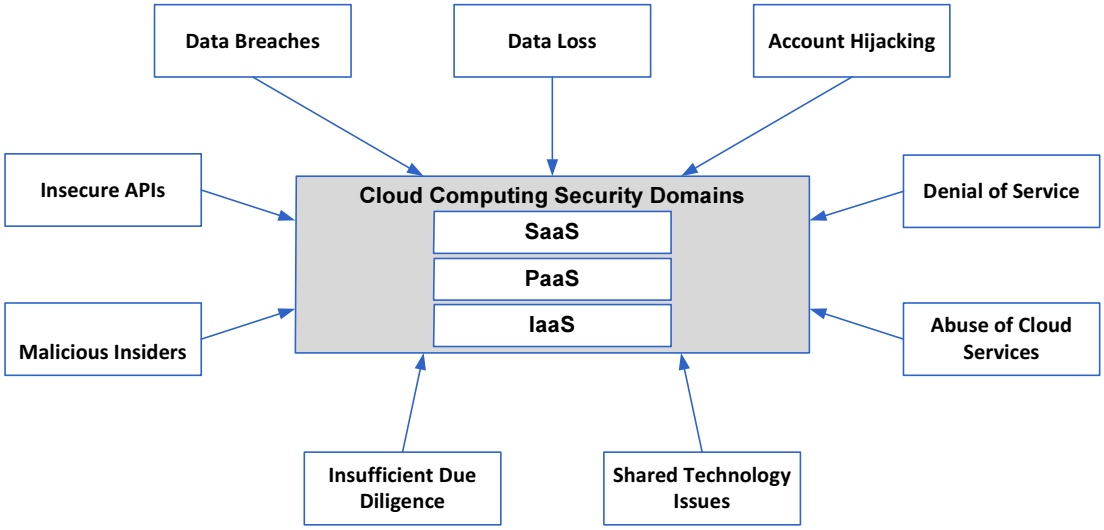

Abbildung 26: Risikofaktoren in der Cloud

Will ein Unternehmen Cloud Services als Lösung einführen, bereitet den Verantwortlichen neben der Sicherheit der Anwendungen und meist mobilen Endgeräte vor allem die Absicherung der Cloud Kopfzerbrechen. Der Grund: Die Einhaltung der Sicherheitsvorschriften muss vom Cloud-Anbieter gewährleistet werden. Der Security Officer des Unternehmens hat darüber keine Kontrolle –wenn auch nur scheinbar. Nach Angaben der Top Threats Working Group der Cloud Security Alliance (CAS), einer Vereinigung von Anbietern von Cloud Services, sind es neun Risiken, die einer breiten Akzeptanz von Cloud-Angebote im Wege stehen. Im Report „The Notorious Nine: Cloud Computing Top Threats in 2013“ werden diese nach Sicherheitsrelevanz geordnet aufgelistet und entsprechende Gegenmaßnahmen formuliert [CAS 2013]. Datendiebstahl, Datenverlust, Missbrauch von Nutzerprofilen, unsichere Schnittstellen (API) und die Angst vor Denial-of-Service-Attacken gehören dabei zu den wichtigsten.

7.3.1 Risiko Nummer 1: Datendiebstahl

Der Datendiebstahl gilt als größtes Risiko, welches sich nach sich aus Sicht der Unternehmen in den letzten drei Jahren stark erhöht hat. Das kritischste Szenario ist dabei, dass sensible Informationen unberechtigt in die Hände von Mitbewerbern fallen. Nach Angaben der Deutschen Handelskammer entsteht allein durch daraus resultierende Plagiate ein jährlicher Schaden von über 30 Milliarden Euro. Im Extremfall kann das geschädigte Unternehmen durch Datendiebstahl sogar vollständig vom Markt verdrängt werden. Das Marktforschungsinstitut Valid Research hat letztes Jahr im Auftrag der Firma Ernst & Young 400 Führungskräfte deutscher Unternehmen zum Thema Datendiebstahl befragt und ermittelt, dass das Risiko als stark steigend empfunden wird. Über 90 Prozent aller Unternehmen rechnen damit, dass konkurrierende Unternehmen (42 %), Geheimdienste (17 %), ehemalige oder eigene Mitarbeitende (15 %) oder Online-Plattformen (15 %) als Täterschaft in Frage kommen. Um dem Datendiebstahl zuvor zu kommen, schlägt die Cloud Security Alliance ein Paket von nicht weniger als elf verschiedenen Maßnahmen vor. Sie sind Bestandteil der so genannten „CSA Cloud Control Matrix“, einer detaillierten Checkliste, die eine systematische Vermeidung und Absicherung von Risiken im Cloud Umfeld erlaubt. Sie liegt heute in der Version 3 vor und kann über die CAS Website bezogen werden.

7.3.2 Risiko Nummer 2: Datenverlust

An zweiter Stelle steht laut CAS das Risiko des direkten Datenverlustes in der Cloud. Ganz gleich ob versehentliches Löschen durch den Anbieter, ausgelöst durch Naturkatastrophen, Ausfälle in Rechenzentren, verlorene Kryptografie-Schlüssel oder absichtliche Löschung: Wichtige Daten gehen verloren. Dagegen existieren heute gut etablierte und durchgängige Technologien, die eine vollständige und automatische Überwachung der Datenmanipulation in- und außerhalb des Unternehmens unter Einhaltung der gesetzlichen Vorgaben erlauben. Diese Technologien werden unter dem Begriff Data Loss Prevention (DLP) zusammengefasst. Die grundlegende Idee des Ansatzes besteht darin, sämtliche Daten so zu beobachten, dass sie keine unkontrollierten Wege gehen können. Und dies auf den drei

Ebenen Bewegungsdaten (über das Netzwerk aus dem Unternehmen heraus via Internet transferierte Daten), gespeicherte Daten (in Dateisystemen, Datenbanken und mittels anderen Speichermethoden abgelegte Daten.) und Daten auf dem Endgerät (auf Endgeräten gespeicherte Daten - Laptop, USB Stick, MP3 Player, Smartphones, etc.). In einer Cloud-Umgebung sind die beiden Aspekte Bewegungsdaten und gespeicherte Daten zentral. Aus diesem Grund spielt beispielsweise auch der Speicherort eine entscheidende Rolle. Deutsche Unternehmen wollen ihre Daten am liebsten in Deutschland aufbewahrt wissen und stehen der Nutzung von Servern ausländischer Rechenzentren sehr kritisch gegenüber.

7.3.3 Risiko Nummer 3: Missbrauch von Nutzerprofilen

Kaum eine Woche vergeht, in der nicht über einen Diebstahl von Passwörtern berichtet wird. Obwohl in vielen Fällen vorwiegend private E-Mail oder einfache Online-Nutzerkonten betroffen sind, ist dieses Risiko im Bereich unternehmensrelevante Anwendungen und Daten zu berücksichtigen. Die Anzahl der Betrugsdelikte mit Zugangsberechtigungen nehmen in der Statistik des Cybercrime zwar nicht den ersten Platz ein, doch sind die Konsequenzen aufgetretener Fälle weitreichend. So können beispielsweise Nutzerprofile missbraucht werden, um einen unbemerkten Informationsdiebstahl zu begehen oder absichtlichen Datenverlust herbeizuführen. Deshalb schlägt die Cloud Security Alliance acht Cloud Control Matrix Maßnahmen vor, die vom einfachen Zugriffsschutz bis hin zur weitreichenden Autorisierung-Strategie reichen.

8 Fragen und Übungen

8.1 Fragen zum Kapitel

Nr	Frage
1	Welches sind die wichtigsten wirtschaftlichen Faktoren, die für die Einführung einer Cloud sprechen?
2	Was sind die wichtigsten Eigenschaften einer Cloud
3	Welche Anwendungsfälle halten sie für die wichtigsten? Erstellen sie eine Liste mit Prioritäten.
4	Welche Basis-Technologien verwendet Cloud Computing
5	Aus welchem Grund ist die Virtualisierung für Cloud Computing wichtig?
6	Aus welchem Grund werden Virtuelle Netzwerk-Komponenten in einer Cloud eingesetzt?
7	Aus welchen drei Komponenten muss ein Service im Minimum bestehen?
8	Für welche Aufgabenstellungen eignet sich Grid Computing?
9	Aus welchem Grund sind Referenzarchitekturen wichtig?
10	Was ist der Unterschied zwischen Service Management Diensten und Business Support Diensten?
11	Welche Referenzarchitektur halten Sie für die Wichtigste? Begründen Sie Ihre Wahl.
12	Was ist der grundlegende Unterschied zwischen Microsoft Azure und anderen Cloud Plattformen
13	Sie müssen ein Mobiles Back-End in der Cloud realisieren. Welche Plattform würden Sie einsetzen? Begründen Sie Ihren Entscheid
14	Was halten Sie von der Cloud Strategie der Schweizer Behörden?
15	Welches sind die wichtigsten Risikofaktoren der Cloud?

8.2 Übungen zum Kapitel

8.2.1 Konzipierung eines Systems mit der Amazon Referenzarchitektur

Wählen Sie ein Angebot aus den Amazon AWS-Referenzarchitekturen aus, welches für eine Anwendung Ihrer Firma oder Ihres Kunden passen würde. Versuchen Sie abzuschätzen, welche Bestandteile Sie verwenden würden und wie viele davon notwendig wären. Erstellen Sie eine entsprechende Zeichnung und errechnen Sie den monatlichen Preis, den Sie Amazon bezahlen müssten.

8.2.2 Kostenvergleich: Speicherung und Bereitstellung eines Bildarchives mit 4 Terrabytes in der Cloud

Sie müssen ein Bildarchiv mit 4 Terrabytes Daten als einfache Webplattform zugänglich machen. Konzipieren Sie das notwendige Systemen (Speicher und Web Server) und suchen Sie die Preise der drei grossen Cloud Anbieter Amazon, Google und Microsoft heraus. Welches Angebot ist das kostengünstigste?