

Hans Dieter Hellige

**Cloud Computing versus Crowd Computing.
Die Gegenrevolution in der IT-Welt
und ihre Mystifikation in der Cloud**

artec-paper Nr. 184

November 2012

ISSN 1613-4907



Universität Bremen

artec - Forschungszentrum Nachhaltigkeit
Enrique-Schmidt-Str. 7
Postfach 330 440
28334 Bremen
<http://www.artec.uni-bremen.de>

Das Forschungszentrum Nachhaltigkeit ist eine Zentrale Wissenschaftliche Einrichtung der Universität Bremen. Es wurde 1989 zunächst als Forschungszentrum Arbeit und Technik (artec) gegründet. Seit Mitte der 90er Jahre werden Umweltprobleme und Umweltnormen in die artec-Forschung integriert. Das Forschungszentrum bündelt heute ein multi-disziplinäres Spektrum von – vorwiegend sozialwissenschaftlichen – Kompetenzen auf dem Gebiet der Nachhaltigkeitsforschung. „artec“ wird nach wie vor als ein Teil der Institutsbezeichnung beibehalten.

Das Forschungszentrum Nachhaltigkeit gibt in seiner Schriftenreihe „artec-paper“ in loser Folge Aufsätze und Vorträge von MitarbeiterInnen sowie ausgewählte Arbeitspapiere und Berichte von durchgeföhrten Forschungsprojekten heraus.

Impressum

Herausgeber:

Universität Bremen
artec Forschungszentrum Nachhaltigkeit
Postfach 33 04 40
28334 Bremen
Tel.: 0421 218 61800
Fax.: 0421 218 98 61800
Web: www.uni-bremen.de/artec

Kontakt:

Andrea Meier
E-Mail: andrea.meier@artec.uni-bremen.de

Hans Dieter Hellige

Cloud Computing versus Crowd Computing. Die Gegenrevolution in der IT-Welt und ihre Mystifikation in der Cloud¹



„Leviathan“ von Thomas Hobbes als Sinnbild der Aufgabe individueller Souveränität zugunsten von allgemeiner Sicherheit, Ordnung und Effizienz (Titelbild der Erstausgabe des von 1651, Ausschnitt)

„Cloud computing is, in many ways, a return to the centrally coordinated integration of the mainframe time-share era: The personal computer gave us the opportunity to strike back against the ‚glass house‘ and elitism common to that day instead of allowing more cooperative integration of the new resource. Cloud computing is a counterrevolution that brings with it the possibility of some very real performance wins for us and their organizations.“ (Chee, Franklin 2010, S. 2 f.)

„This is a hard lesson of cold mathematics, but it throws a certain light on the adventure of our century: hesitation between an indefinite turbulence of human affairs and the rise of a prodigious Leviathan. In comparison with this, Hobbes’ Leviathan was nothing but a pleasant joke. We are running the risk nowadays of a great World State, where deliberate and conscious primitive injustice may be the only possible condition for the statistical happiness of the masses: a world worse than hell for every clear mind.“ (Norbert Wiener 1950, S. 180)

¹ Eine thesenartige Zusammenfassung wurde vorgetragen unter dem Titel „Cloud oder Crowd im Internet der Informationen und der Energien. Verantwortung abgeben oder übernehmen?“ beim Symposium des Fachbereiches Informatik und Gesellschaft der „Gesellschaft für Informatik“, Berlin, 2.-3.11.2012 »Verantwortung übernehmen – Mitten in der Turing Galaxis«.

GLIEDERUNG

1. Einführung: Von der Dezentralisierung zur Rezentralisierung des Computing 3
2. Die Anfänge des Cloud-Begriffs als Metapher im Networking und Computing 9
3. „Computer Utilities“ und „Digital Libraries“ als historische Vorbilder eines zentral organisierten Computing 16
4. Die dezentralen Vorläufer des Cloud Computing: „Metacomputing“ und „Grid Computing“ 28
5. Die Anfänge der Rezentralisierung: „Network Computing“ und „Utility Computing“ 33
6. Die Startup-Firmen des Cloud Computing als Service-Anbieter für E-Commerce 36
7. Die Grid-/ Utility-Konzepte von IBM, SUN, HP und Microsoft zur Rückgewinnung der Kontrolle über die PC-Internet-Welt 40
8. Amazons und Googles Cloud-Strategien zur Monopolisierung der Web 2.0 Kulturen 52
9. Legitimationsstrategien für das Cloud Computing: Die Inszenierung einer Epochenwende 63
10. Literatur 74

1. Einführung: Von der Dezentralisierung zur Rezentralisierung

Die Geschichte des Computing und der Computerkommunikation erschien lange Zeit als ein weitgehend linearer Prozess des Downscaling und der Dezentralisierung. Nach der Ablösung des Gersch's Law als Treiber der Größensteigerung und Zentralisierung von Rechnersystemen ermöglichte das Moore's Law als strategischer Schrittmacher der Miniaturisierung immer kleinere und zugleich leistungsfähigere Computer und IT-Geräte und mit ihnen eine Befreiung von der Kontrolle allmächtiger Rechenzentren. So kam gemäß dem von Gordon C. Bell formulierten „Bell's Law of Computer Classes“, einer Erweiterung des Moore's Law, alle 10 Jahre eine neue, kleinere Computergeneration auf den Markt: nach den Mainframes die Minicomputer, Workstations, PCs und Notebooks, denen seit 2000 Handheld-Computer, Smartphones, Wearables, intelligente Sensoren usw. folgen.² Die Mainframe-Basis verlor zwar ihre beherrschende Stellung in der IT-Welt, doch sie verschwand nicht, wie die PC-Counterculture der 70er Jahre und auch Bell erwartet hatten, im Gegenteil, sie gewann gerade durch die Ausbreitung der Client-Server-Architektur und Lokalen Netze sowie von Internet- und Web-Diensten seit den 80igern wieder an Bedeutung. Die Ausnutzung der Vorteile von Standard-ICs in massiv-parallelen Computerarchitekturen führte seit den 90iger Jahren zu einem neuen Rechenzentrumstyp, den weiträumig vernetzten Clustern von „multi-datacenters“.³ Sie wurden zusammen mit den High-performance Supercomputern die Basis einer regelrechten Renaissance der Mainframes und, verdeckt durch die in neuen Endgeräten ständig präsente Moore's Law-Dynamik, einer Revitalisierung des Gersch's Law.

Gestützt auf immer leistungsfähigere zentrale Server- und Speicher-Cluster kam es in der IT-Welt zu einer Trendwende vom unkontrollierten „client-side computing“ zum kontrollierten „server-side-computing“.⁴ Insbesondere die großen Software- und Contentanbieter versuchten den durch die PC-Revolution verlorenen Einfluss auf die Computernutzung wieder zurückzugewinnen. Eine weitere Verschiebung der Machtgewichte in der sozialen Architektur des Internet ergab sich aus der fortschreitenden Kommerzialisierung der Social Networks und Web 2.0-Netzdienste, durch die Internet und Web zu einem zunehmend transparenten Nutzerbeobachtungsfeld der IT-Konzerne und zur Basis von neuartigen, die Eigenarbeit der User und deren kollektive Netzintel-

² Hellige (2009), bes. S. 136-153.

³ Barroso, Hözle (2009), S. 1.

⁴ Barroso, Hözle (2009), S. 1.

ligenz einbeziehenden Geschäftsmodellen wurden. Schon der Zugang zum World Wide Web erfolgt zum allergrößten Teil über die Serverfarmen der Betreiber von Suchmaschinen und kommerziellen sozialen Netzwerken und ermöglicht diesen schon jetzt eine feinstrukturierte Erfassung des Nutzerverhaltens.⁵ Zur ständigen Optimierung des „targeted advertising“ sammeln sie, der Vision eines allwissenden Suchdienstes folgend, mithilfe immer ausgefeilterer (Ever-)Cookie-Techniken sowie Tracking-, Registrierungs- und Identifikationsmethoden von den Benutzern wirklich alles. Google und Facebook führen die von ihnen erhobenen Userdaten sogar zu Megaprofilen bzw. zu Dossiers zusammen. Diese unterliegen gestützt auf den „Patriot Act“ von 2001 dem Zugriff von Geheimdiensten, FBI usw. und werden nach sich verdichtenden Hinweisen auch bereits als Informationsbasis und Instrument der Personen-Filterung genutzt. Es besteht somit die begründete Gefahr, dass die in den Stichwort- und Dossier-basierten Such-, Filterungs- und Auswertungsmethoden, im Data-Mining, in der „Big-Data-Analytik“, Lokalisierung und sogar in biometrischen Erkennungsverfahren einander annähernde kommerzielle und geheimdienstliche Ausforschung und Überwachung der Internet-Nutzer sich auch wechselseitig immer mehr ergänzen. Durch diese sich allmählich herausbildende Arbeitsteilung bei der *Erfassung der Welt* könnte sich das Internet am Ende zu einer „idealen Überwachungsplattform“ entwickeln.⁶

Doch die Rückgewinnung der Kontrolle über die User beschränkt sich nicht auf die Speicherung von MAC-Adressen, Metadaten und die Registrierung von Navigationsverläufen, es landet inzwischen auch immer mehr Content in den Rechenzentren von IT-Konzernen. Gelockt von kostenlosen werbefinanzierten Angeboten, verlagern viele private Internetnutzer bereits seit längerem neben Mails auch Teile ihrer Fotos, Videos und persönlichen Dateien zum Austausch oder zur längeren Lagerung auf Serverfarmen von spezialisierten Speicherdielen und kommerziellen Web-Dienstleistungsportalen wie YouTube, Flickr, Picasa, Dropbox u.a. Weitere User-Datenbestände landen in zentralen Speichersystemen aufgrund der Ausweitung der Software-Portefeuilles von Google, Facebook und Microsoft, die vom Browser aus den mobilen Zugriff auf Textverarbeitungs-, Spreadsheet- und Präsentations-Programmpakete im Netz gestatten und sich auf diese Weise zu kompletten Software-as-a-Service-Plattformen entwickeln, die die Benut-

⁵ Zu den Techniken des Tracking und der Registrierung siehe u.a. Röhle (2010), S. 167-228, zum „Behavioural Targeting“ und zur Personalisierung S. 211 ff.

⁶ Zeger (2012), S. 5.

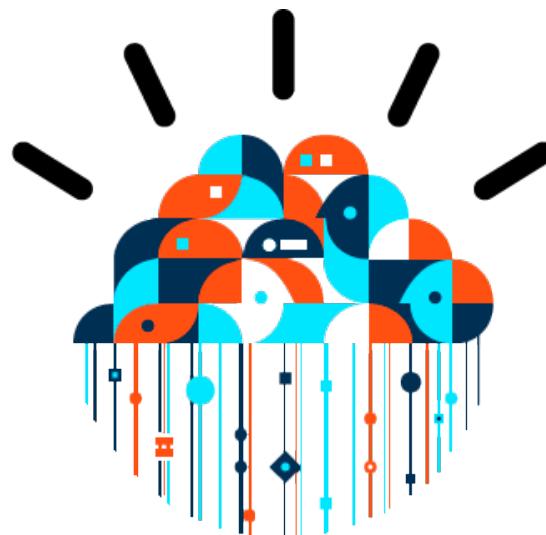
zer nach Möglichkeit gar nicht mehr zu verlassen brauchen: Der Übergang zum sozial zentrierten „Cloud Computing“ hat somit als Begleiterscheinung der permanenten Ausweitung des „advertising space“ im Internet längst eingesetzt, doch wird dieser schrittweise Prozess von den Usern bislang kaum als Problem wahrgenommen.⁷

Ungeachtet dieser schlechenden Verlagerung bislang dezentraler Verarbeitungs- und Speicherkapazitäten in zentrale Serverfarmen bewahrten sich die User dank eigener Rechnerkapazitäten, Festplatten und externer Speicher-Ressourcen bislang aber noch ihre Wahlmöglichkeiten und ihren Spielraum für einen nicht kommerziell überwachten selbstorganisierten Rechnergebrauch. So blieb trotz der beträchtlichen Machtverschiebung die Arbeitsteilung zwischen den Netzinstanzen im Prinzip noch unangetastet. Mit dem Cloud Computing ist nun aber im letzten Jahrzehnt ein paradigmatisches Architekturkonzept erschienen, das viel radikaler eine Rückverlagerung der *gesamten* Datenverarbeitung und -speicherung von den dezentralen Endsystemen in zentrale Serverfarmen als das zukünftige Leitkonzept der IT-Welt propagiert und mit der Devise des Endes des PC-Zeitalters einen grundlegenden Wandel in den Beziehungen der Netzinstanzen und in der Arbeitsteilung zwischen verschiedenen Computing-Ressourcen anstrebt. Denn bei einer vollständigen Zusammenfassung von Hard- und Softwareressourcen, Metadaten und Nutzerdaten in zentralen kommerziellen Daten-Sammel-, Speicherungs und -Auswertungskonglomeraten wäre ein radikaler Wandel der sozialen Architektur von Computing und Internet die Folge, durch den die User auf Dauer die Kontrolle über die Verarbeitung und Speicherung ihrer Computing-Aktivitäten und zugleich ein entscheidendes Druckmittel gegen mögliche Preisdictate und Nutzungsvorschriften der Provider und Anbieter von Computingdiensten verlören.

Die in der IT-Community diskutierten Cloud Computing-Szenarien weisen zwar eine große Spannweite von eher additiven bzw. selektiven Outsourcing-Strategien bis zu radikalen Konzepten einer Komplettlauslagerung des Computing zu externen Cloud-Providern auf. Doch nur für relativ wenige Autoren stellt Cloud Computing lediglich eine inkrementelle Innovation dar, die vor allem der Abdeckung von Spitzenlast und selten gebrauchten Diensten dient. Für die Mehrheit der Fachleute und Analysten handelt es sich vielmehr um eine radikale Innovation und generelle Umstiegstechnologie. Dabei geht man davon aus, dass kleine und mittlere Unternehmen völlig auf eigene Server und

⁷ Beckmann, Horowitz 2010, S. 21 f.; Levy 2011a, S. 211.; Sitaram, Manjunath 2012, S. 3.

„Fat Clients“ verzichten, dass Behörden, Hochschulen und öffentliche Institutionen ihre DV-Kapazitäten wieder durchgängig in Ressourcen-Pools und Community Clouds zentralisieren und dass größere Unternehmen mithilfe von Private Clouds den teuren PC- und Lokalserver-Betrieb durch „Thin Clients“ und Intranet-Dienste ablösen werden. Neben den von einer Institution bzw. einem Unternehmen betriebenen Clouds gibt es auch bereits Planungen für Cloud-Plattformen ganzer Verwaltungsbereiche, Kommunen und Branchen. Schließlich existiert auch eine ganze Reihe von Vorschlägen für „volunteer cloud infrastructures“, kooperative „cloud federations“ und „P2P Clouds“, die sich mit dezidiert verteilten Architekturen und Peer-to-Peer-Organisationstrukturen als demokratische Alternativen zum herrschenden Cloud-Zentralismus verstehen.⁸



IBM SmartCloud

Logo der IBM SmartCloud von 2011: Symbolische Vereinigung der User in der Cloud⁹

Während bei letzteren Cloud-Formationen die volle Verantwortung und Kontrolle bei den sie tragenden Institutionen und Firmen verbleiben, weisen die Public Clouds eine grundsätzlich andere soziale Architektur auf. Die Benutzer begeben sich hier völlig in die Obhut eines externen Cloud-Service-Providers, der die technischen, ökonomischen und vor allem rechtlichen Nutzungsbedingungen bestimmt. Es gibt in diesem Marktsegment zwar eine große Zahl von „small und medium scale providers“, doch nur wenige

⁸ Siehe u.a. Babaoglu, Jelasity u.a. 2006; Anderson, Fedak 2006; Anderson, David P. 2010; Distefano, Puliafito 2011, 2012; Babaoglu, Marzolla, Tamburini 2012.

⁹ Abb. nach <http://blogs.voanews.com/tedlandphairsamerica/files/2011/10/06-smarter-cloud-Michigan-Municipal-League-MML.jpg>

IT-Giants dominieren im weltweiten Cloud-Markt. Es sind hier neben den meist spezialisierten Cloud-Plattformen der Telcos in Europa vor allem die in den USA beheimateten „mega-scale providers“ wie Google, Amazon, IBM, Microsoft, Yahoo und Apple, die über ein globales Netz von Serverzentren verfügen und die die komplette Palette von Cloud-Diensten anbieten.¹⁰ Auch wenn bei diesen das Computing *technisch* über mehrere Server-Standorte, also quasi *dezentral* abgewickelt wird, geschieht dies in einer *sozialen zentralen* Architektur, denn alle Prozesse bleiben unter der „single-organization control“ und vorher dezentrale Datenbestände konzentrieren sich nun in wenigen Datenaggregationen.¹¹ Von diesem bereits jetzt marktbeherrschenden Typ werden im IT-Sektor die größten Veränderungspotentiale erwartet. Die „IT-Fabrics“ würden zum Haupttreiber des Cloud Computing, da die „construction and operation of extremely large-scale commodity-computer data centers at low-cost locations“ durch die Konzentration in unmittelbarer Nähe billiger Wasserkraft und über eine Ausschöpfung von „very large economies of scale“ eine beispiellose Senkung der Kosten für Software und Hardware, Elektrizität, Übertragung, und Betrieb bewirkten.¹² Mit ihren zentralen Rechen-, Speicher-, Software- und Kommunikations-Dienstleistungen sollen sie daher künftig den bisherigen privaten PC-Massenmarkt ersetzen und sogar, wie vielfach angenommen, durch volle Ausspielung der skalenökonomischen Kostenvorteile nach und nach auch die Wissenschafts- und Unternehmens-Clouds in sich aufnehmen. Es wird deshalb vielfach auch bereits davor gewarnt, dass ein „*techno-utility complex*“ weniger großer Cloud-Provider den gesamten Markt dominieren könnte.¹³ Als künftige Alternative zu den „monolithic mega providers“ und dem „concept of a cloud operated by one service provider or enterprise“ fordern daher, angeführt von Kevin Kelly, Vinton G. Cerf und einem Cisco-Team immer mehr Autoren den Zusammenschluss aller Clouds in einem globalen Verbundsystem, einer „cloud of clouds“ bzw. „Intercloud“, die nach dem Vorbild des Internet föderativ organisiert wird.¹⁴

¹⁰ Celesti, Tusa, Villari 2012, S. 169 f.; Wang, Ranjan, Chen 2012, S. 6 f.

¹¹ Barroso, Hölzle 2009, S. 2.

¹² Armbrust, Fox, Griffith 2010, S. 52. Den Sieg der skalenökonomischen Vorteile erwarten ebenfalls: Barroso, Hölzle 2009, S. 2, 91; Chorafas 2010, S. 186 f. Dagegen glauben Hwang, Fox, Dongarra 2012, S. 196, dass trotz der Übermacht der Public Clouds längerfristig die Private Clouds überwiegen werden, da sie den Firmen und Institutionen größere Sicherheit und Vertrauenswürdigkeit bieten.

¹³ Nicholas Carr's Artikel „The techno-utility complex“ in seinem Blog „Rough Type vom 13.12.2007; Wainewright 2007 (the techno-utility complex, in which vested economic and political interests conspire together to build huge technology-based utility industries that preserve and reinforce their power bases.“; Cooke 2010, S. 4; Distefano, Puliafito 2011; 2012, S. 27 f.

¹⁴ Kelly 2007 („Eventually we'll have the intercloud, the cloud of clouds. This intercloud will have the dimensions of one machine comprised of all servers and attendant cloudbooks on the planet.“); Cerf 2009;

Die folgende Rekonstruktion des Cloud Computing legt dagegen den Schwerpunkt auf die Genese und die Entwicklung des bis heute bestimmenden Typus der Mega Cloud. Im Zentrum der Analyse steht dabei der Zusammenhang zwischen Akteurskonstellationen und sozialen Architekturen, d.h. der jeweiligen Ausstattung der System- und Netzinstanzen mit Ressourcen und Aktivitätsspielräumen. Die *soziale Architektur* entscheidet über die Machtbalance zwischen den beteiligten Akteuren im technischen System, also über die Verteilung der Einflussmöglichkeiten und damit letztlich über die Systemführerschaft. Das erste Kapitel widmet sich der Entstehung der sprachlichen bzw. metaphorischen Mystifikation der „Cloud“, das zweite den Vorläufern eines zentral organisierten Computing und die folgenden Kapitel den einzelnen Stadien der Genese des sozial zentrierten Cloud Computing.



Mystifizierende Darstellung der Cloud Computing-Genese in einem Artikel in der Software Development Times vom 12.8.2011¹⁵

2. Die Anfänge des Cloud-Begriffs als Metapher im Networking und Computing

Als Begriff kam „Cloud-Computing“ erst 1997 auf, sein kometenhafter Aufstieg begann sogar erst 2006/7, doch die Cloud-Metapher hat eine bis in die 60iger Jahre zurückreichende Vorgeschichte. Deren Ursprünge liegen zum einen in der Wolken-Metapher der Systemtheorie, die wohl auf Karl Popper zurückgeht, der in seiner Compton-Vorlesung von 1965 *Wolken* als Sinnbild chaotischer, nicht-determinierter und deshalb nicht vorhersehbarer Zufallsprozesse einführte und diesen die *Uhr* als Symbol des berechenbaren mechanistisch-deterministischen Systemparadigmas gegenüberstellte.¹⁶ Sehr bald verwendete man Wolkengebilde auch in systemtheoretischen und kybernetischen Schaubildern zur Kennzeichnung von ausgeblendeten oder nicht spezifizierten Systemumfeldern. Zum anderen geht die Cloud-Metapher auf unregelmäßige runde Graphen zurück, mit denen Nachrichtentechniker in Netzwerkdarstellungen im Detail unbekannte oder nicht spezifizierte Netzinfrastrukturen sowie seit den 60iger Jahren nicht-deterministische Vermittlungsstrukturen in Store-and-Forward-Netzen symbolisch ausdrückten.¹⁷ So verwendete das Team um Donald Davies am National Physical Laboratory bereits seit 1967/68 wolkenartige Graphen für Knotennetzwerke mit zuvor nicht festge-

¹⁵ Morales 2011.

¹⁶ Popper, Wolken und Uhren (1973).

¹⁷ Aljawarneh 2012, S. 268 („Cloud computing originates from telephone network, the industry called the transparent and dark telephone network cloud.“)

legten Teilstrecken bzw. Übertragungsrouten. Auch der zweite Pionier von Packet-Switching Networks, Leonard Kleinrock, der 1961 noch einen Kreis als Darstellungssymbol für Subnetworks mit virtuellen Routen verwendet hatte, ging in seinem theoretischen Hauptwerk von 1974/76 zu wolkenartigen Graphen über.¹⁸

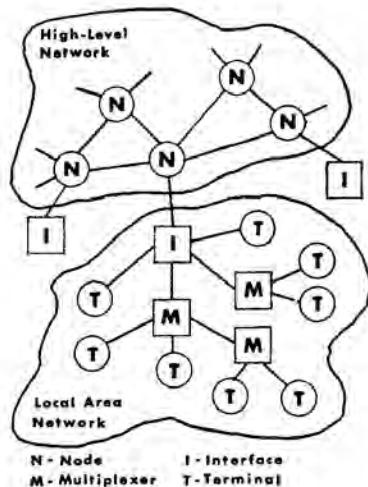


Fig. 1.

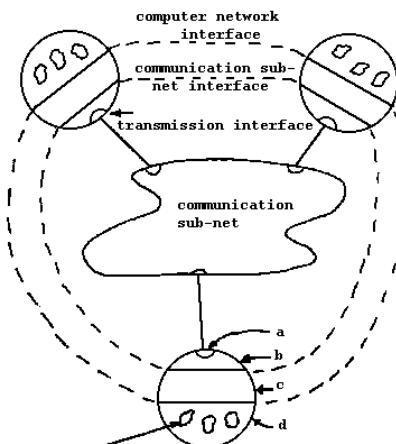


Fig 1. Levels of hierarchy

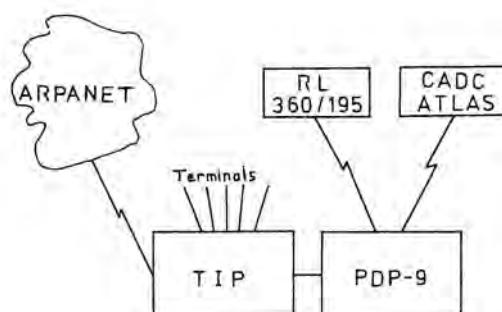
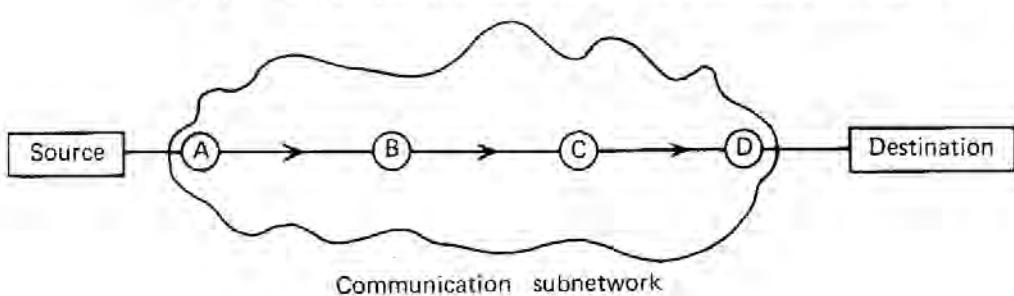
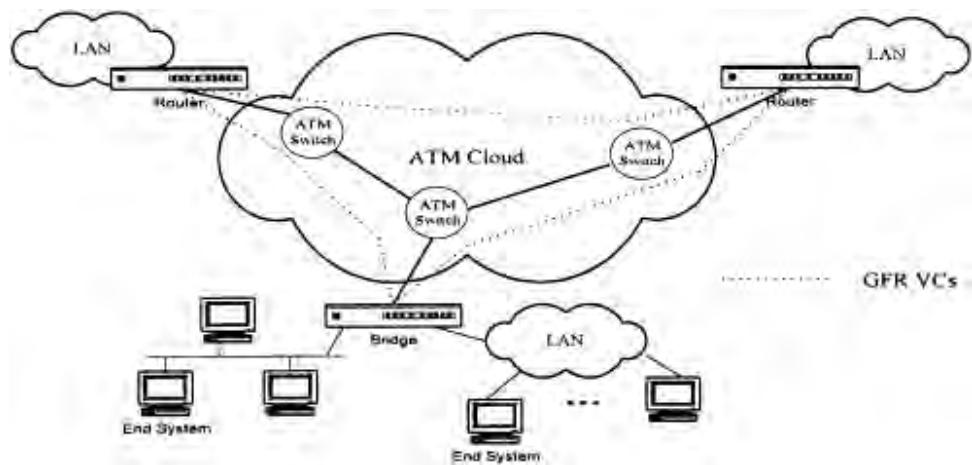


Figure 14: UCL Computer Configuration.

Anfänge der Cloud-Metapher bei den ersten Paketvermittlungsnetzen
Oben: das NPL-Data Communication Network (links: Davies 1968, S. 710;
rechts: Scantlebury, Wilkinson 1974, S. 224)
Mitte: Cloud-Darstellung des Arpanet : (Kleinrock 1976, S. 295)
Unten: Systemarchitektur des University College London
(Stokes, Higginson 1975, S. 27)

¹⁸ Davies, 1968, S. 710; Kleinrock, 1961, S. 14; Kleinrock, 1974/76, Bd. 2, S. 295.

Seit Ende der 70er Jahre wurden Clouds zum Standard-Bildsymbol für alle Arten von Paketvermittlungsnetzen (X.25 Standard von 1976). Sie bürgeren sich vor allem in der Arpanet-/Internet Community ein, so dass „cloud“ hier bald zum *visuellen Synonym* für das Internet wurde. Als *sprachlicher Terminus* tauchte „Cloud“ aber erst 1993/94 im Zusammenhang mit den asynchronen und synchronen Breitbandnetzen ATM (Asynchronous Transfer Mode) und SDH (Synchrone Digitale Hierarchie) bzw. Sonet (Synchronous Optical Network) auf. Die "ATM-Cloud" bzw. die "SDH-Cloud" bezeichnen hier die im Unterschied zu den im Verbindungsaufbau zuvor festgelegten Durchschaltevermittlungen die in ihrem Verhalten nicht-determinierten virtuellen Verbindungsnetzwerke. Im Zusammenhang mit dem Interworking zwischen ATM-Netzen und dem Internet ist dann in der zweiten Hälfte der 90iger Jahre auch immer häufiger von der „Internet-Cloud“ und von „IPv4-Clouds“ die Rede. Infolge des schnellen Endes der Universalnetzaspirationen und der Rückstufung der ATM-Netze zu vermittlungstechnischen Infrastrukturen wurde dann ab 2000 „die Cloud“ weitgehend identisch mit dem Internet.



ATM-Cloud als virtuelles IP/ATM Internetwork zur Verbindung lokaler Netzwerke
(Wang, Chan, Chen 2000, S. 291)

Im Unterschied zur starken Präsenz bei den Netzwerkarchitekturen spielte die Cloud-Metapher im Computing lange Zeit nahezu keine Rolle. Bis zum Ende der 90er Jahre finden sich nur sehr vereinzelte Belege, so das 1985-90 am Georgia Institute of Technology entwickelte objektorientierte Betriebssystem „Clouds Distributed Operating System“. Zusammen mit einer Programmumgebung und einem „distributed shared memory“ sollte es Workstations eine fehlertolerante Kommunikation und Zusammen-

arbeit über Rechnernetze ermöglichen: „This control unifies the different computers into a single integrated compute and storage resource.“ Als ein „distributed shared global virtual space“ könnte das Clouds OS in Zukunft das Paradigma für eine verteilte Kooperation im Netz werden.¹⁹ Doch angesichts der gleichzeitigen Konkurrenz einer Reihe weiterer verteilter Betriebssysteme konnte sich das Clouds OS nicht durchsetzen, so gab es zu den um 2010 wieder einsetzenden Entwicklungen von „Cloud Operation Systems“ keinerlei Kontinuität.²⁰ Dadurch fand auch der Cloud-Begriff im Computing zunächst keine Nachahmer, die Begriffe Grid-, Utility- und Network Computing bestimmten die Debatten über Alternativkonzepte zur PC-Client-Server-Architektur.

Erst 1996/97 finden sich die ersten Erwähnungen des Begriffs *Cloud Computing* und zwar gleich an zwei Stellen in ganz unterschiedlichen Kontexten. Zuerst taucht er in einem vertraulichen Planungspapier des PC-Herstellers Compaq vom 14.11.1996 auf als Synonym für das Internet-zentrierte IT-Business der Zukunft.²¹ Im folgenden Jahr verwendete der Informations-Management- und E-Commerce-Spezialist *Ramnath K. Chellappa* von der University of Texas in Austin den Begriff erstmals auf einer wissenschaftlichen Konferenz. Dieses Faktum wird in diversen historischen Abrissen des Gebietes in Fachbüchern als der offizielle Startpunkt von Cloud Computing gewertet. Fixiert auf den Namen, übersah man dabei freilich, dass sich Chellappas Begriff überwiegend auf neue Markt- und Verwendungsformen und kaum auf informationstechnische Architekturen und Infrastrukturen bezieht und dass er in den Folgejahren kaum rezipiert, sondern erst nachträglich in die Geschichtskonstruktion einbezogen wurde. Chellappa, der gerade eine Dissertation über Anbieterstrukturen und Vermarktungsformen digitaler Produkte²² abgeschlossen hatte und sich intensiv mit neuen E-Publishing- und E-Learning-Ansätzen in „virtuellen Universitäten“ beschäftigte, wollte mit seinem Vorstoß die nach dem Ende der „mainframe-based structure“ in der „network-based architecture“ entstandene Vielfalt konkurrierender technischer Begriffe überwinden.

Das neue „dynamic computing paradigm“ erschien ihm ohnehin weniger von technischen Restriktionen als von ökonomischen und sozialen Kriterien bestimmt. „While

¹⁹ Dasgupta, Le Blanc u.a. 1991, S. 35 f.

²⁰ Pianese, Bosch u.a. 2010.

²¹ Siehe dazu unten Kapitel 6.

²² Chellappa 1997a (Im Zentrum standen hier „electronic intermediaries“ wie das „Electronic Mall Environment (EME), „digital warehouses“ und „payment systems“).

many terms have appeared to describe these new forms, the advent of electronic commerce has led to the emergence of 'cloud computing.' [...] this would be a new computing paradigm where the boundaries of computing will be determined by economic rationale rather than technical limits alone.²³ Chellappa entschied sich deshalb aus seiner E-Commerce-Perspektive für einen von allen technischen Architekturen und Lösungsalternativen abstrahierenden Begriff „cloud“, unter dem er recht vage den „computing space between the provider and the enduser“ verstand. Ihn interessierten vorrangig die finanziellen und organisatorischen Erwägungen, Sicherheitsfragen sowie die neuen Akteurskonstellationen und Marktformen elektronischer Produkte und Dienstleistungen einschließlich der Probleme der Gewichtung der vielfältigen Preisinformationen. Denn künftig seien finanzielle Aspekte ausschlaggebend dafür, ob die Datenverarbeitung in Fremd- oder Eigenversorgung erfolgen sollte: „foreign computing can be a computing paradigm the location where the limitations involving computing will be determined by financial explanation as opposed to specialized boundaries.“ Auf die allgemeinen Erwägungen Chellappas über Outsourcing-Bedingungen und neue Verwendungsformen der Internet-Kommunikation in Märkten und im Bildungswesen wurde später zwar immer wieder Bezug genommen, doch im eigentlichen Entstehungsprozess des ‚Cloud Computing‘-Konzeptes spielten sie keine erkennbare Rolle. Hier bestimmten zunächst modifizierte Grid- und Utility Computing-Begriffe den fachlichen Diskurs, zu denen noch „Pervasive“, „On Demand“- und „Service-Computing“ hinzukamen.

Es ist üblich, dass neue Computing-Paradigmen in der Entstehungsphase unter verschiedenen mit einander konkurrierenden Begriffen und Metaphern in Erscheinung treten und dass sich hinter einem neuen Leitbegriff anfangs recht divergierende Konzepte und soziale Architekturen verbergen. Eine rein begriffs- bzw. metaphorengeschichtlich angelegte Paradigmen-Chronologie führt daher meist zu problematischen Epochen-Konstruktionen, die Diskurshistorie muss deshalb die begrifflichen und metaphorischen Sprachspiele der Zunft in der Genesephase genauer rekonstruieren und dabei die jeweiligen sozialen und kulturellen Bedeutungen näher untersuchen. Hierbei fällt nämlich auf, dass die Cloud-Metapher und der Cloud Computing-Begriff von dem für die Informatik typischen Entwicklungsmuster abweichen, bei dem anfangs technische, naturwissenschaftliche und speziell anthropomorphe Metaphern (Organe, Neuronen) vorherrschen,

²³ Chellappa, 1997a,b; Chellappa, Gupta, 2002, S. 118.

die mit zunehmendem Entwicklungsgrad kulturellen (Sprache, Bibliotheken) und vor allem soziomorphen Struktur- und Organisations-Metaphern (Hierarchien, Pyramiden, Fließbänder, Fabriken) Platz machen, um dann vielfach wieder zu erneuten biologischen bzw. anthropomorphen Metaphern und Begriffen überzugehen.²⁴ Mit Cloud rückt nun in einem sehr fortgeschrittenen Stadium des Computing und der Computervernetzung ein Gestaltmuster der *Natur* in den Mittelpunkt, das gerade die Unbestimmtheit und Intransparenz in den ansonsten so logisch durchorganisierten Techniksstrukturen verbildlicht. Diese Unbestimmtheit bildete dann auch den Ansatzpunkt für etliche postmoderne Sprach- und Metaphernspielereien in den frühen kultur- medien- und techniksoziologischen Betrachtungen zum Cloud Computing.²⁵ Vor allem aber ist die Wolkenmetapher bis heute Anknüpfungspunkt für zahllose Visualisierungen in der Fachpresse und besonders in der Werbung von Cloud-Serviceprovidern.



Intransparente Cloud-Darstellung von 2012: Die Leichtigkeit des Seins in der Cloud²⁶

In ihrem Bedeutungskern zielt die Wolken-Metapher auf das Abstrahieren oder Verbergen von konkreten Strukturen, sozialen Architekturen und vor allem von Komplexität durch Virtualisierungs-Technologien. Cloud Computing korrespondiert darin mit den

²⁴ Siehe dazu Busch 1998 und Hellige 2003.

²⁵ Siehe hierzu z.B. die Beiträge auf der HyperKult 2009 (abrufbar unter <http://www.leuphana.de/institute/icam/forschung-projekte/hyperkult/hyperkult-videoarchiv.html>), und kritische Forschungsskizze in Heilmann 2010.

²⁶ Abbildung in einem Bericht über die Messe SPS/IPC/DRAVES 2012, Quelle: Enisa, abgebildet in Scharf 2012.

zeitgleich entstandenen Begriffen „Invisible bzw. Calm Computing“, „Disappearing Computer“ und „Ubiquitous bzw. Pervasive Computing“ sowie auch mit „Autonomous Computing“ im Sinne eines stets aktiven aber unsichtbaren „autonomen Nervensystems“.²⁷ Alle diese Verflüchtigungsmetaphern sind nur Variationen des „Magic-Behind-the Scenes“-Gedankens, der gerade durch seine Unsichtbarkeit die User beglücken möchte. Diese Metaphern-Intention hat sich dabei so gründlich in dem Cloud Computing-Begriff niedergeschlagen, dass der zugrundeliegende *soziale* Inhalt verborgen wird: die technische und gesellschaftliche Neuverteilung der Computing-Ressourcen, die vor allem in den Mega Clouds mit einer Verschiebung der Macht- und Wissensbalance zwischen den beteiligten Netzakteuren einhergeht. Im Folgenden sollen deshalb die in der Metaphorik und der von ihr geprägten Begrifflichkeit verdrängten Akteurskonstellationen und sozialen Architekturen sichtbar gemacht und die Cloud Computing-Genese in ihren gesellschaftlichen und geistesgeschichtlichen Kontexten rekonstruiert werden.



Transparente Cloud-Darstellung von 2010: IT-Konzerne als dominante Cloud-Provider²⁸

²⁷ Siehe dazu Hellige 2008b, Kap. 9; Hellige 2010, S. 4-14.

²⁸ Quelle: Abbildung in Hall 2010.

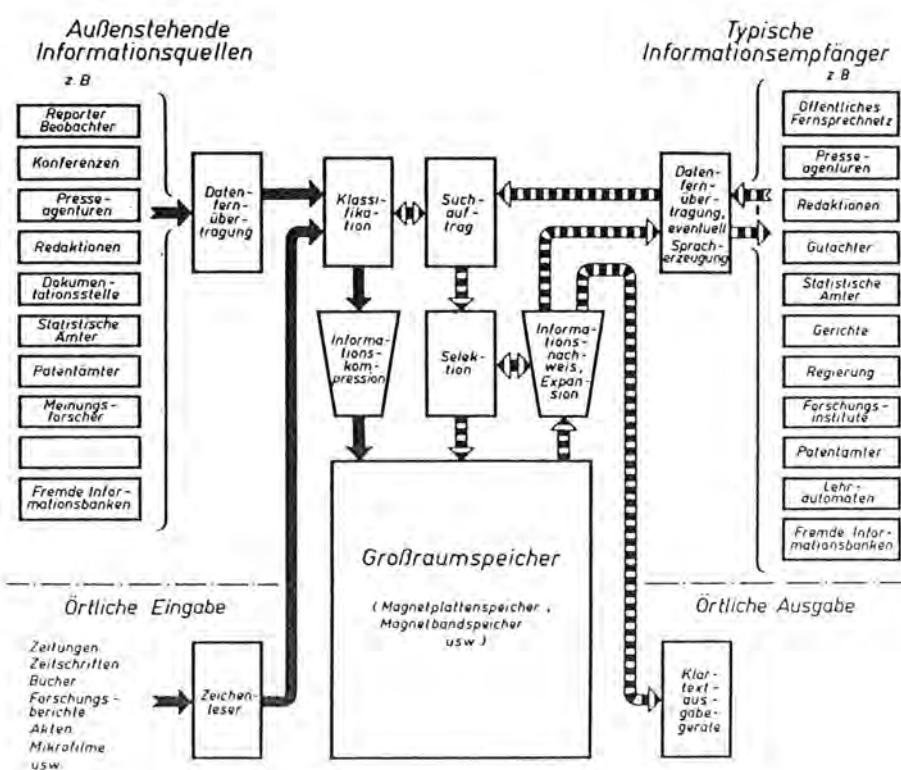


Abb. 4. Hypothetische Informationsbank.

Vision eines „informationellen Verbundnetzes“ von Informationsbanken nach dem Vorbild amerikanischer „Information Utilities“ bzw. „Libraries of the Future“ (Karl Steinbuch 1968b, S.13 f.)

3. „Computer Utilities“ und „Digital Libraries“ als historische Vorbilder eines zentral organisierten Computing

Zentralistische Systemarchitekturen prägen die Frühgeschichte des modernen Computing und der Computerkommunikation, sie bilden daher den maßgeblichen Ausgangspunkt für die schon in der Entstehungsphase des Cloud Computing einsetzenden Bestrebungen, dem durchzusetzenden neuen Paradigma eine historische Ahngalerie zu verschaffen. Etliche der stereotypen Geschichtskonstrukte sehen die Vorläufer des Cloud Computing in den Visionen und Konzepten hochzentralisierter Computer- bzw. Information Utilities der Time-Sharing-Ära. Eine ganze Reihe historischer Abrisse in Büchern und Artikeln über Cloud Computing nennen *Herbert Grosch* als den Urheber der frühesten Cloud Computing-Vision, da er aus dem von ihm Ende der 1940er Jahre entdeckten und 1953 publizierten „Grosch's Law“ eine „natural technological evolution“ zum „super computing“ und „dumb terminals“ als die künftige Norm abgeleitet habe.²⁹ Grosch habe

²⁹ Ryan, Merchant, Falvey 2011, S. 7.

daraus bereits während der 50er Jahre gefolgert, „that the entire world would operate on dumb terminals powered by about 15 large data centers“.

Die Quelle für diese Information sind, das zeigt der immer gleiche Wortlaut in den Darstellungen, die Artikel „Cloud Computing“ und „Internet Area Network“ der englischen Wikipedia. Doch die hierzu in den Artikeln aufgeführten Belege enthalten nichts über die angebliche „prediction“, ebensowenig Groschs eigene Papers, Interviews und Memoiren. Vielmehr prophezeite er um 1948/50 als „wave of the future“: „that were *hundreds, and thousands of computing centers waiting to be born*“. Selbst seine Prognose „The Future of Computing“ aus dem Jahre 1958, in der er Computer als „tool or technology for every-day living“ mit radikalen Folgen für die Gesellschaft vorhersagte, enthält keinerlei Aussagen über eine künftige Zentralisierung der Computer-Landschaft.³⁰ Das Grosch's Law und skalenökonomische Argumentationen spielten zwar seinerzeit als Begründung für eine kontinuierliche Größensteigerung von Computern und Rechenzentren immer eine wesentliche Rolle, doch angeregt wurden derartige Szenarien vielmehr durch zentralistische Architekturmodelle in der Folge des militärischen „Central Command and Control-System“ SAGE, dessen Leitbild eines nationalen Informationssammel-, verarbeitungs- und -verteilzentrums auch nicht-militärische Informationszentralen und große Time-Sharing-Systeme folgten.³¹

Das früheste Beispiel für einen solchen Spinoff ist der Vorschlag des Leiters des Whirlwind-Projektes *Jay Forrester* von 1948 für den Aufbau eines „problem solving network“, bei dem eine Vielzahl von Nutzern über Fernmeldeleitungen für die Lösung komplexer mathematischer Aufgaben auf die Ressourcen eines leistungsstarken Echtzeitcomputers zugreift.³² Die erste publizierte Vision eines Zentralrechners mit vielen angeschlossenen z.T. weit entfernten Stationen als Rechner der Zukunft stammt von dem Bell Lab Systems Engineer *Richard C. Matlack* aus dem Jahre 1955: „I am sure that we have all dreamed of the future possibility of having quick and easy access to an electronic computer in order to obtain from its memory a fast and painless answer to a question.“ Mit einem „centralized data processing center“ könnten alle Daten einer „large number of remote locations“ zusammengeführt und dabei stets auf aktuellem Stand gehalten

³⁰ Grosch, Memoirs, S. 130 ff.; Grosch 1958, S. 1664 (meine Hervorhebung).

³¹ Vgl. Hellige 1992, Hellige 1996, S. 210 ff.

³² Friedewald 1999, S. 88 f.

werden. Dadurch könne auch das Management immer ein genaues Lagebild über das Gesamtgeschehen erhalten und so die infolge von Centralisierungsprozessen in der Industrie verloren gegangene Kontrolle wieder zurückgewinnen. Die „crisis of control“ (Beniger) war so von Beginn an ein Leitmotiv für die Centralisierung von Systemarchitekturen im Computing.³³

Kontrollaspekte zusammen mit marktstrategischen Überlegungen bildeten auch den Hintergrund der schon wesentlich konkreteren, 1957 entstandenen Konzeption von *Robert Bemer*, der bei IBM für die Standardisierung von Zeichensätzen und Programmiersprachen zuständig war. Bemer schlug einen zentralen Computer-Service vor, bei dem ein „huge central computer“ eine Vielzahl weit verstreuter Nutzer mit Computer-Leistungen versorgt und dank der höheren Wirtschaftlichkeit der „next generation of super-computers“ die Massenausbreitung wenig leistungsfähiger Kleincomputer ab-blockt: „Assuming the availability of practical micro-wave communications systems, it is conceivable that one or several computers, much larger than anything presently contemplated, could service a multitude of users. They would no longer rent a computer as such; instead they would rent input-output equipment, although as far as the operation will be concerned they would not be able to tell the difference. This peripheral equipment would perhaps be rented at a base price plus a variable usage charge on a non-linear basis.“³⁴ Dieser Computer-Service wäre „self-scheduling, self-regulating, and self-billing to the customer on the basis of use of the input-output device“. Von einer derartigen Centralisierung des Computing versprach sich Bemer eine Vereinheitlichung der „entire structure of usage“, eine Kontrolle von „entire plant operations“ durch einen „super-speed computer“ und die Durchsetzung eines „advanced common language system“, „so all users can integrate their particular operations into the complex of control demanded by an automated future [...].“ Bemer weitete, ohne hierbei den Begriff zu verwenden, das bis dahin dominierende Hardware-Konzept des „Time-Sharing“ zu einem gewerblichen Teilnehmerbetrieb aus, der die Tradition von zentralen Lochkarten-Services und der IBM-Kontrollphilosophie mit der neuen Computer-Technologie fortführte. Sein konsequent auf Centralisierung setzendes Zukunfts-Szenario nahm damit zugleich bereits 1957/58 einige Momente des späteren Cloud Computing vorweg.

³³ Matlack 1955, S. 83; Beniger 1986.

³⁴ Bemer 1957, auch zum Folgenden, zit. nach Bemer undatiert.

Unabhängig von Bemer entwarf *Walter F. Bauer*, der in der militärnahen Computerfirma Thompson Ramo-Woolridge die Entwicklung von „large-scale data processing systems“ und industriellen Softwaresystemen leitete, an der Jahreswende 1957/58 das Ideal-System eines "large-scale computers of the future". Sein von ihm "Ultradatic" genannter Service-Computer sollte eine große Benutzerpopulation versorgen und sich dabei mit Hilfe von Überwachungs- und Kontrollmechanismen selber steuern und sich automatisch an den Wechsel der "priorities of problems" anpassen. Die Kosten der sich selber organisierenden Großanlage wollte er durch den Anschluß von externen Stationen über Telefonleitungen auf möglichst viele Nutzer verteilen: „The central idea here is that each large metropolitan area would have one or more of these super computers. The computers would handle a number of problems concurrently. Organizations would have input-output equipment installed on their own premises and would buy time on the computer *much the same way that the average household buys power and water from utility companies.* [...] The user would be charged for that time used and only that time used, and the accounting procedure would reflect the customers' detailed use.“³⁵ Hiermit formulierte Walter F. Bauer als erster das Leitbild einer "*Computer utility*" und stellte diese den traditionellen „public utilities“ der Strom- und Wasserversorgung an die Seite. Die American Management Association griff Bauers Utility-Konzept 1959 auf, da sie hoffte, dass "computer power as a public utility" wie einst in der Stromversorgung die Rentabilitätsnachteile kleiner privater Versorgungsgesellschaften beseitigen würde, weil auch beim Computing die gemeinsame Anlagennutzung und das "sharing of computer programs" die Kosten drastisch senken könne.³⁶ Bauers "*Computer utility*"-Gedanke führte damit das *Resource-Sharing-Prinzip* in die Computer Community ein, das im folgenden Jahrzehnt das vorherrschende Leitprinzip werden sollte.

Ende der 50er Jahre wurde das Utility-Konzept vom Massachusetts Institute of Technology (MIT) und anderen Hochschulen aufgriffen und grundlegend neu ausgerichtet. Bei ihnen standen nicht mehr die Absicherung der Mainframe-Ökonomie und die Größensteigerung im Vordergrund, sondern die Informationsbeschaffung und der Informationsaustausch einer Community sowie die Wirtschaftlichkeit des Computing insgesamt. Der KI-Pionier *John McCarthy* vom MIT war es, der die „dreamy visions“ 1959/60 in einem konsistenten Leitbildkomplex bündelte und den Plan für einen "quick service

³⁵ Bauer 1958, S. 46, 49 (meine Hervorhebung); siehe auch Bemer undatiert und Hellige 1996, S. 209 ff.

³⁶ Mann 1959, S. 11 ff.

computer" ausarbeitete, der verschiedenen akademischen Nutzerkreisen einen kosten-günstigen Zugriff auf Rechenzeit und Informations-Pools ermöglichen sollte: „There are a large number of users. Each user has his own console, gets service from the computer whenever he desires it, and has the computer maintain his files for him.“³⁷ Bereits 1960/61 sahen McCarthy und seine MIT-Kollegen die entstehenden Campus-Utilities als den Kern für lokale und regionale Informationsversorgungs- und -austauschsysteme. Am Ende sollten Timesharing-Zentren neben Telefon, Gas und Wasser eine zentrale Infrastruktur des gesamten Gemeinwesens werden, die jedem interessierten Nutzer Rechenkapazitäten, Anwendungsprogramme und Informationsdienstleistungen (z.B. Wetter- und Wirtschaftsprognosen) zur Verfügung stellte. Unternehmen und Institutionen könnten dann darauf verzichten, eigene Rechner zu erwerben: „If computers of the kind I have advocated become the computers of the future, then computation may someday be organized as a *public utility*, just as the telephone system is a public utility. We can envisage computing service companies whose subscribers are connected to them by telephone lines. [...] The *computing utility* could become the basis for a new and important industry.“³⁸

Das bei der Centenarfeier des MIT im Jahre 1961 entwickelte Infrastruktur-Leitbild einer „*public computer utility*“, das vielen historischen Abrissen des Cloud Computing als früheste Ausprägung des Cloud-Konzeptes gilt, wurde Ausgangspunkt für eine Vielzahl immer weiter ausgreifender Anwendungszonenarien. McCarthy selber beschrieb 1966 ein gegenüber den ursprünglichen Ideen wesentlich erweitertes Leistungsspektrum der „öffentlichen Computerzentrale“, die nun zur für alle zugänglichen Staatsbibliothek, zur zentralen Dienstplattform für Handel und Gewerbe, zum allgemeinen Programm- und Wissensspeicher und damit zum eigentlichen Zentrum der Gesellschaft avancierte. Über in jedem Heim installierte Computerkonsolen sollte dann das Tele-Learning, Tele-Working, Tele-Banking und sogar der Ferneinkauf abgewickelt werden: „Da die Gebühren von der Benutzung abhängen, lassen sich die verschiedenartigsten Bedürfnisse befriedigen. Der öffentliche Zentralcomputer steht jedem Menschen als sein ‚zweites Gedächtnis‘ zur Verfügung. Seine ein- und ausgehenden Nachrichten können genau registriert und Termine zu den vorbestimmten Zeiten angezeigt werden. [...] Das öffentliche Informationssystem schließt zum Beispiel auch die Fenster, wenn es anfängt zu regnen, vorausge-

³⁷ Teager, McCarthy 1959, S. 1f. (meine Hervorhebung).

³⁸ McCarthy, 1962, S. 229, 231, 236 (meine Hervorhebungen).

setzt, man hat in seiner Wohnung die erforderlichen Meßfühler und Betätigungsseinrichtungen installiert.“ Ja durch den Anschluss an ein „automatisiertes Konstruktionsystem“ könnten die Benutzer die gesamten Entwurfs- und Planungsarbeiten für den Eigenheimbau selber durchführen.³⁹ Für jedes Bedürfnis der Teilnehmer werde sich ein Unternehmen finden, das entsprechende Service-Leistungen anbieten würde. Die Vision einer „public computer utility“ ging damit weit über das Leistungsspektrum klassischer Infrastrukturen hinaus, sie hatte konzeptionell bereits den Charakter einer allgemeinen Plattform für alle Arten von Dienstleistungen.

In einem weiteren Home-Computing-Szenario entwarf McCarthy 1970 sogar bereits eine Art „Industrie 4.0“-Vision, durch die über eine Verkopplung von „Home Information Terminals“, Public Utility und Unternehmens-Computerzentralen die bisherige Massenproduktion individualisiert werden könnte: „The idea is that automated design programs can produce designs for articles meeting individual specifications. Either by himself or in consultation with an expert, an individual would use the system to produce a design and display how it would look and possibly how it would perform. Candidates for individual design include clothing, furniture, boats, electronic equipment, houses, and even cars. The system would then produce the instructions for controlling machine tools, fabric cutters, and also printed instructions for the hand parts of the operation. In general, it should be possible to make single objects at little more cost than present mass produced objects.“⁴⁰ Doch diese Vorwegnahme eines „Internet der Dinge und der Dienste“ blieb angesichts des tatsächlich vorhandenen und damals erreichbaren Standes der Technik ein hochspekulatives Szenario, das schnell wieder in der Versenkung verschwand und erst 2000 als visionäres Vorbild und als „Vorgeschichte“ wiederentdeckt wurde.⁴¹

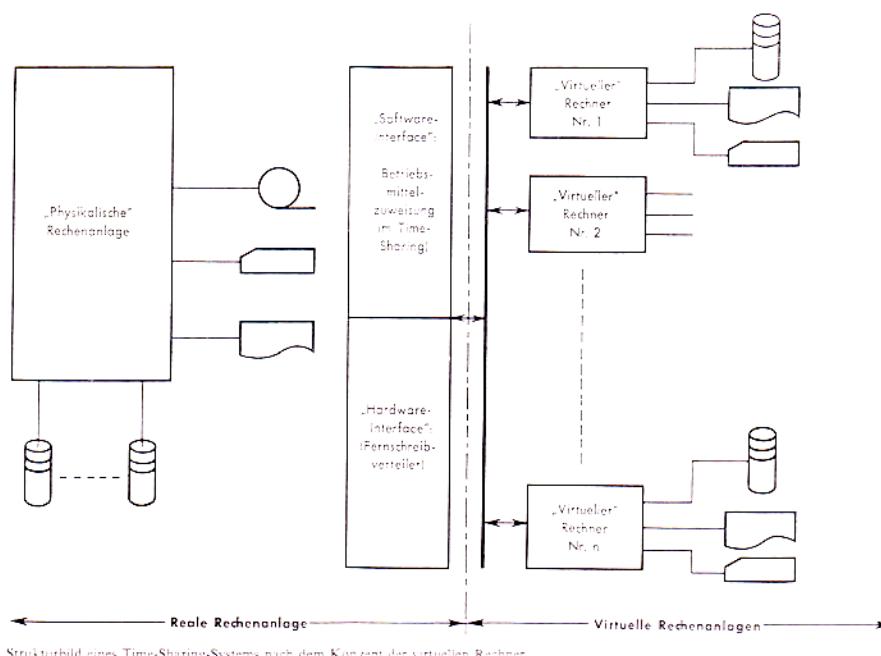
Im Unterschied zu McCarthys Konzeption einer kommerziellen Service-Plattform auf Basis einer öffentlich-rechtlichen Time-Sharing-Infrastruktur betonten andere Public Utility-Szenarien stärker den Charakter eines „Informationspools“ des ganzen Gemeinwesens. So sah *Robert Fano*, der am MIT das berühmte „MAC-Project“ leitete, in der

³⁹ McCarthy 1966, S. 22 f.; Hellige 1996, S. 227 f.

⁴⁰ McCarthy 1970, S. 5.; vgl. die Bekanntmachung des Industrie 4.0-Szenarios durch Kagermann, Lukas 2011.

⁴¹ Siehe dazu McCarthys Kommentierung vom Juni 2000 in der Internet-Version seines Papers (abrufbar unter: <http://www-formal.stanford.edu/jmc/hoter.html>).

Time-Sharing-Zentrale vorrangig eine Einrichtung der zugleich öffentlichen und privaten Informationsversorgung, ein „Gegenstück zur öffentlichen Bibliothek“.⁴² Nach der Realisierung der Time-Sharing-Zentrale am MIT 1962/63, bei der eine Vielzahl von „virtual machines“ Zugriff auf die Hardware- und Software-Ressourcen des Großrechners erhielten, setzte der Boom des neuen Schlagwortes ein. In nur wenigen Jahren wurden "computer- bzw. information-utilities", z. T. auch "data-utilities", überall in den USA Leitbegriffe der fachlichen und öffentlichen Debatte über die künftige DV- bzw. Informationsversorgung. Die für ein zentral organisiertes Computer-Utility-Konzept ins Feld geführten Argumente erinnern dabei bereits sehr stark an heutige Legitimationsstrategien für Public Clouds: Die User könnten dank voller Ausschöpfung von Skaleneffekten ein wesentlich besseres Preis-Leistungsverhältnis erzielen als mit Eigenanlagen; sie erhalten über dieses frühe Virtualisierungs-Konzept Zugang zu Hochleistungs-Computern modernster Bauart; ihnen steht ein breites Spektrum von Software-Tools und Programmiersprachen zur Verfügung; sie ersparen sich teure Investitionen in eigene Anlagen, die nur schlecht ausgelastet werden; sie haben viel geringere Wartungs- und Betriebskosten; sie zahlen nur für tatsächlich in Anspruch genommene Leistungen; Hard- und Software-Ressourcen sind flexibel skalierbar.⁴³



Strukturbild eines Time-Sharing-Systems nach dem Konzept der virtuellen Rechner.

Erste Realisierung des Virtualisierungs-Konzeptes in Time-Sharing-Systemen⁴⁴

⁴² Fano, Corbató 1966, S. 83, 99; Fano 1967

⁴³ Siehe die Zusammenfassung der „positive cost-effectiveness arguments“ bei Parkhill 1966, S. 123-142.

⁴⁴ Abb. in Giloi 1968, S. 132.

Es zeigte sich aber im praktischen Time-Sharing-Betrieb sehr bald, dass die Technologie für eine nationale Informationsinfrastruktur noch nicht reif war und dass insbesondere hochzentralisierte Systeme stärker unter Warteschlangeneffekten, hohem Overhead und Verfügbarkeitsproblemen zu leiden hatten und dass diese vor allem weitaus anfälliger für Datenschutzverletzungen waren. Die Konzentration aller Nutzerdaten in einem Riesenspeicher oder gar in einem „National Data Center“, die in den frühen, von der akademischen „club-house mentality“ getragenen Computer-Utility-Szenarien völlig unbedenklich schien, wurde Mitte der 60er Jahre nach ersten praktischen Erfahrungen im MAC-Project als gravierendes Problem erkannt. Die Time-Sharing-Technologie wurde so noch vor ihrer eigentlichen Ausbreitung Anlass für eine intensive Privacy-Debatte. Vor dem Hintergrund der Massenüberwachung durch Geheimdienste in der noch nicht lange zurückliegenden McCarthy-Ära⁴⁵ erschienen die avancierten Computersysteme nun als mögliche ernste Bedrohung der Privatshäre.⁴⁶ Die Pioniere des Time-Sharing und der Computer Utility erwiesen sich damit im Unterschied zu denen des Cloud Computing als weitaus sensibler für die informationelle Verletzlichkeit und den hohen rechtlichen Schutzbedarf zentraler Computing- und Informationssysteme.

Doch am Ende waren es nicht die massiven Datenschutz- und Datensicherheitsprobleme, die einer breiten Diffusion von „large computer utilities“ im Wege standen, sondern in erster Linie die unzureichende Wirtschaftlichkeit großer Time-Sharing-Systeme. Denn wegen des exponentiellen Wachstums des Overheads und der sich daraus ergebenden starken Zunahme von Delays konnten weitaus weniger Terminals angeschlossen werden, als man beim Einsatz größerer und schnellerer Computer erwartet hatte. Vor allem blieb die fest eingeplante Senkung der Übertragungskosten aus, da sich die Bandbreiten der Übertragungsstrecken nur langsam erweiterten. Insgesamt führten das deutlich schlechtere Preis-Leistungsverhältnis und mannigfache Probleme im Time-Sharing-Betrieb dazu, dass eine radikale Zentralisierung der Datenverarbeitung und Informationsversorgung ausblieb. Der Aufstieg von Minicomputern und PCs seit der zweiten Hälfte der 60er Jahre und mit ihm das Ende des kurzen Hypes des „Universal Utility“-Leitbildes ließ sich nicht mehr verhindern. Erst im Laufe der Entstehung des Cloud

⁴⁵ Nach dem US-Senator Joseph McCarthy benannte, von 1947-1956 reichende Ära hysterischer „Kommunistenjagden und Denunziationen.“

⁴⁶ David, Fano 1965, S. 245; Fano, Corbatò 1966, S. 99 f.; vgl. auch Sprague 1969, Kapitel 11.

Computing besann man sich während der 90er Jahre wieder auf das „utility computing paradigm“ und die System-Analogien zwischen Strom-, Telefon- und Computernetzen.⁴⁷

Demgegenüber spielte das zweite herausragende informationelle Infrastruktur-Leitbild der 60er/70er Jahre, die „*Digital Library*“ bzw. „*Library of the Future*“, in der Cloud Computing-Genese keine Rolle mehr. Auch dieses Leitbild zielte auf ein öffentlich-rechtliches bzw. staatlich reguliertes Infrastruktursystem, doch lag hier im Gegensatz zur Computer Utility der Schwerpunkt auf einer Rationalisierung der *Informationsversorgung*. Durch die Zusammenfassung aller Informationen in einem Informationssammel- und -verteilzentrum sollte das Hardware- und Software-Resource-Sharing der Computer Utility zu einem gesamtgesellschaftlichen Information-Sharing erweitert werden. Auch hierbei lässt sich eine Entwicklung von monolithischen hochzentralisierten Systemstrukturen über hierarchisch-verteilte zu stärker dezentralisierten Architekturen beobachten. Die erste Phase der „*Digital Library*“-Konzepte stand noch ganz unter dem Einfluss von Zentralbibliotheks-Modellen und Mikroform-basierten Informationszentralen sowie militärischen Central Command and Control-Systems. Den Anstoß gab hierbei die Zukunftsvision des führenden Informations- bzw. Dokumentationswissenschaftlers *Calvin N. Mooers* von 1959, nach der im Jahre 1980 das gesamte Wissen in einem Netz von großen und kleinen „machine information centers“ gespeichert und den Nutzern über „information“ bzw. „reading machines“ zugänglich gemacht werden sollte.⁴⁸

Unter dem Einfluss der Time-Sharing-Technologie begann im Laufe der 60er Jahre ein Wandel der Digital Library-Konzepte von zentralistischen zu hierarchisch verteilten Architekturen. Leidbild-prägend wurden hierbei die Bibliotheks-Visionen von *Joseph Licklider*, dem wichtigsten Promotor des Interactive Computing und der Computernetz-Entwicklung. Seine Szenarien für ein System von „thinking centers“ und eine "Library of the Future" blieben zwar weiterhin auf zentrale Großcomputer und Rechenzentren angewiesen, doch diese wandelten sich mehr und mehr von den Sammel-, Verarbeitungs- und -Verteil-Zentren der SAGE-Ära zu Servern für Informationsaustausch und dezentrale Wissensproduktion an den Konsolen.⁴⁹ Wie die Time-Sharing-Systeme an den Universitäten wurden die digitalen Bibliotheken als eine kollektive Infrastruktur

⁴⁷ Siehe dazu die folgenden Kapitel.

⁴⁸ Mooers 1959

⁴⁹ Licklider 1960; Licklider 1965.

mit implizitem Open-source-Charakter verstanden. Man betrachtete sie als "public utilities", bei denen entweder wie in der Strom- und Gasversorgung der Versorgungsaspekt betont wurde oder wie bei der Telefonzentrale die reine Vermittlungsfunktion im Vordergrund stand.

Doch in der zweiten Hälfte der 60er Jahre entstand in Teilen der Computer Science unter dem Eindruck der konkreten Erfahrungen mit Time-Sharing-Systemen und der Herausbildung von lokalen und regionalen „user communities“ ein erneuter Dynamisierungs- und Dezentralisierungsschub in den Modellen der informationstechnischen Wissensorganisation und Wissensordnung. Die „Library of the Future“-Szenarien lösten sich nun weitgehend vom Zentralbibliotheksmodell und gingen zu einem Netzmodell der „online communities“ über, die mehr dem späteren Grid Computing als dem Cloud Computing ähnelten. In dieser gewandelten Architektur wurden Lickliders digitales Bibliothekskonzept nun zu einem entscheidenden Auslöser und wichtigen Leitbild der frühen Arpanet / Internet-Entwicklung. Sie entzogen damit aber auch der Big Library-Metapher den Boden, das *Netz* wurde nun der Computer. In der Folgezeit repräsentierten daher die Metaphern „Future Library“, „Neolibrary“, „Digital Library“ usw. bis auf wenige Ausnahmen nicht mehr das informationelle Gesamtsystem, sondern nur noch den engeren Bereich der Digitalisierung des traditionellen Bibliothekswesens. Doch Ende der 80er Jahre wurde die Big Library-Metapher noch einmal Ausgangspunkt für eine Vision der Designer der ersten Version der TCP/IP-Protokolle Kahn und Cerf für das Internet der Zukunft.

Bereits in dieser ersten umfassenden Ideenkonzeption für eine Art Cloudsystem aus dem Jahre 1988 wurde ein enger Zusammenhang zwischen skalenökonomisch motivierter Zentralisierung, Kommerzialisierung von Internetdiensten, Bequemlichkeitsversprechen und einer kaum wahrnehmbaren Beobachtung und Kontrolle der Nutzeraktivitäten hergestellt. Darin entwickelten sie einen Stufenplan für den Aufbau einer „*Digital Library System*“ genannten „National Information Infrastructure“, in der alle öffentlichen, privatwirtschaftlichen und persönlichen Datenbanken und Informationsnetze zu einer „seamless structure of access to information“ zusammengeführt werden sollten, ähnlich wie es früher bei den klassischen Infrastrukturen von Wasser, Elektrizität, Telefon und Highways geschehen sei. Aus der Geschichte der Infrastrukturen entwickelten sie ihre Projektion der Informationsversorgung im Jahr 2000: „*Digital Libraries*

have now become such a pervasive part of everyday living that it's hard to remember what life was like without them. Like other infrastructure, one never really thinks about how it works, how it evolved or, how it is maintained, any more than one thinks about water, electricity, telephones and highways when they are readily available".⁵⁰ Damit brachten Kahn und Cerf das spätere Leitmotiv der in den Hintergrund tretenden Informationsinfrastruktur in den Diskurs über das Internet der Zukunft ein.

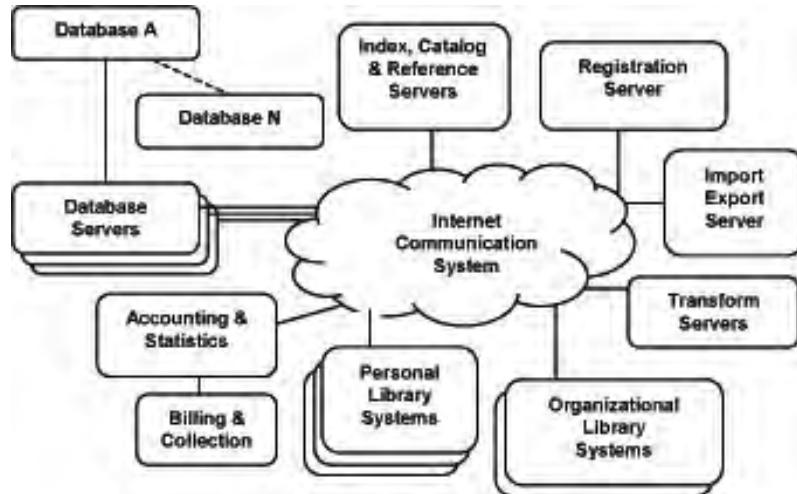


Figure 2 Structure of the Digital Library System

Cloud-artige Architektur-Konzeption für ein Internet der Zukunft von Robert Kahn und Vinton G. Cerf aus dem Jahre 1988⁵¹

Ihr integrales "Digital Library System", das auch aus einem System von interagierenden „autonomous independent Digital Library Systems“ bestehen konnte, knüpfte sowohl an das Utility-Denken der Time-Sharing-Ära wie auch an Lickliders Pläne für eine digitale Weltbibliothek und ein Netz von „Thinking Centers“ an.⁵² In der Technikausstattung gingen Kahn und Cerf aber weit darüber hinaus, denn um 2000 sollten alle Sprachdienste und breitbandigen Informationsmedien einschließlich HDTV einbezogen sein. Dazu sollten der natürliche Dialog mit den Rechnern und ein komplexes verteiltes System von Software-Agenten, die sie „World of Knowbots“ nannten, alle Probleme der Bedientechnik, der Informationssuche und der Ressourcenallokation lösen. Als Grundstruktur sahen sie ein „distributed system“ vor, in dem die User Dienste von „powerful, geographically distributed and often locally networked workstations“ bezögen. Doch

⁵⁰ Kahn, Cerf 1988, S. 9.

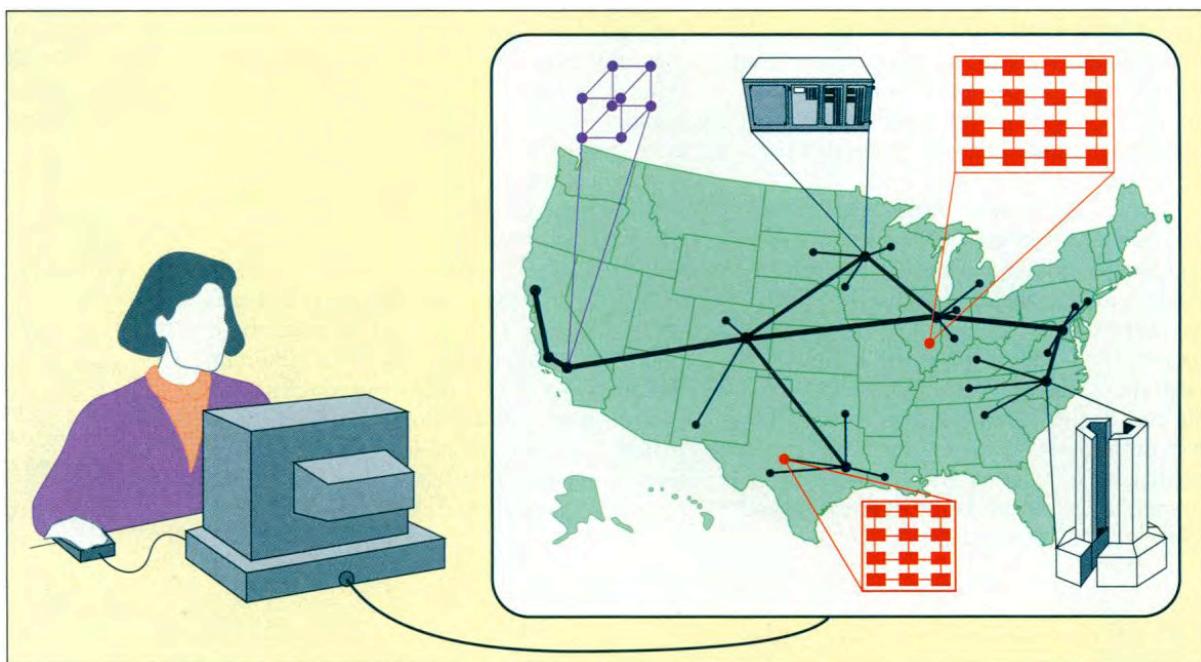
⁵¹ Kahn, Cerf (1988), S. 17.

⁵² Siehe Kahn, Cerf (1988), S. 41 und die Belege in Hellige (2008c), S. 11 ff.

dort, wo es ökonomisch angebracht ist, würden die Ressourcen in größeren Einheiten konzentriert, wobei sie sogar eine Zentralisierung des gesamten „data content“ für möglich hielten.

Deutlich in Richtung „Cloud Architecture“ wies auch der Vorschlag eines aus acht Hauptkomponenten bestehenden verteilten Serversystems, das das Internet-Knotennetz erweitern und das über Standardprozeduren und Knowbots die Dokumentenbeschaffung, die Formatwandlung, die Nutzer- und Objekt-Registrierung, die Indexierung und Katalogisierung sowie die automatische Verwaltung der Urheberrechte und Abrechnung übernehmen sollte. Alle wissenschaftlichen, ökonomischen, gesellschaftlichen und persönlichen Transaktionen würden dabei ständig durch „tireless knowbots“ sortiert, analysiert und „mined“, um proaktiv den Benutzern, Providern und Anbietern die von ihnen benötigten Informationen zur Verfügung zustellen.⁵³ Sie stellten damit bereits Ende der 80er Jahre einen konzeptionellen Zusammenhang zwischen einem KI-basierten Data Mining, Datenanalytik und Proaktiven Computing her. Die Grenze zwischen beauftragten User-Agents und dem Ensemble von eigenständig agierenden Knowbots war im Interesse einer optimalen Informationsversorgung dabei bewusst unscharf gelassen, was hohe Datenschutzrisiken in sich barg. Generell gingen Kahns und Cerfs Pläne vom Modell sich selbst organisierender On-line-Communities der frühen Netzkulturen zu einem stark privatwirtschaftlich geprägten Versorgungs- und Vertriebskonzept über. An die Stelle der Open-source Information trat hier der KI-unterstützte offene Markt der Informations-Anbieter mit automatischem Abrechnungssystem. Was hier noch an einer Cloud fehlte, war die völlige Auslagerung von Verarbeitungs- und Speicherfunktionen per Virtualisierung sowie der Begriff selber.

⁵³ Kahn, Cerf 1988, S. 34 ff.



Grimshaws Vision eines „Enterprise-wide Computing“ auf der Basis eines nationalen „Metacomputers“⁵⁴

4. Die dezentralen Vorläufer des Cloud Computing: „Meta-Computing“ und „Grid Computing“

Entstehungsmilieu des *Grid Computing* war die US-amerikanische Parallelrechner- und Supercomputing-Forschung und Entwicklung, die vor allem auf die Nutzung in der (halb-) militärischen Hochenergiephysik, der Luft- und Raumfahrt, der Mikroelektronik sowie der Klimaforschung, Medizin und der naturwissenschaftlichen Grundlagenforschung ausgerichtet war. Ein früher Vorläufer eines Grid-Konzeptes entstand bereits 1957/58 am U.S. National Bureau of Standards. Die sehr unregelmäßige Rechnernutzung war hier der Anlass zu einer Verkopplung mehrerer Computer zu einem „interconnected data-processing network“, in dem größere Jobs in kleinere Teile aufgesplittet und mit Hilfe eines „workload-sharing scheme“ parallel abgearbeitet wurden. Dieses Multi-computer-System PILOT wurde Ausgangspunkt für die Vision, dass künftig Usergruppen anspruchsvolle Computingprozesse selber über ein weiträumiges Netz von „collaborating Computers“ organisieren.⁵⁵ In der Parallerechner-Community und der Computerkommunikation gab es auch in der Folgezeit immer wieder Ansätze und Pläne für ein weiträumiges Teamwork von miteinander kommunizierenden Prozessoren und Compu-

⁵⁴ Grimshaw 1994b, S. 893.

⁵⁵ Leiner, Notz, Smith, Weinberger 1957, S. 115, 128.

tern, doch es entwickelte sich angesichts der Dominanz zentralistischer Architekturen daraus kein Grid-Konzept.

Erst Anfang der 90iger Jahre ließen extrem rechenintensive Simulationen, die selbst die Kapazität leistungsfähigster Rechner überforderten, in größerem Ausmaß in US-Großforschungseinrichtungen informelle „High Performance Computing Cluster“ auf Zeit entstehen, die über die bereits bestehenden, dem Back-up und der besseren Auslastung von Rechner-Ressourcen dienenden „High Availability-Cluster“ und „Load-Sharing Cluster“ hinausgingen. An der University of Virginia entwickelte *Andrew S. Grimshaw* Anfang der 90iger Jahre die Vision von „large *meta-systems* comprised of a variety of interconnected high-performance architectures“, die durch die Kombination von Gigabit-Übertragung, Parallelrechentechnik und Objektorientierung die Computer-User entlasten und miteinander verbinden. Das Leitbild war getragen von der ‚basisdemokratischen‘ Intention „to make the power of these architectures available to ‚the masses‘.“ Die Bedienung der heterogenen „Cluster-Clusters“ sollte mithilfe der „principles of the object-oriented paradigm“ vereinfacht und die Komplexität der Struktur des jeweiligen „meta-system“ hinter einem „logical ‚window‘“ bzw. einer „virtual machine“ verborgen werden.⁵⁶ Realisiert wurde das Vorhaben ab 1993 in dem Projekt „Legion“ in Form eines prototypischen „Campus-Wide Virtual Computer“ (CWVC), der zugleich als Muster für ein „Enterprise-wide Computing“ dienen sollte. Ihre Vision zielte auf die Ablösung der bisher isolierten Supercomputer und Workstations durch ein „nation-wide meta-system“, mit dem Grimshaws Team demonstrieren wollte „that a nation-wide virtual computer can out-perform the resources available at any single site.“⁵⁷ Fernziel war die „Legion Vision of a Worldwide Virtual Computer“, d.h. „a seamless metasystem or metacomputer“ für ein weltweites Resource Sharing: „Our vision of Legion is a system consisting of millions of hosts and billions of objects co-existing in a loose confederation united through high-speed links. Users will have the illusion of a very powerful desktop computer through which they can manipulate objects“.⁵⁸ Mit der sozialen Architektur einer losen Konföderation selbständiger Computing-Ressourcen bildete es nahezu eine Hardware-Entsprechung zum World Wide Web von Berners Lee.

⁵⁶ Grimshaw, 1992, S. 54 f., 58.; Gentzsch 2001a, S. 3 ff. (meine Hervorhebung)

⁵⁷ Grimshaw, Wulf, French 1994, S. 21; Grimshaw 1994.

⁵⁸ Grimshaw, Wulf, French 1994a; Grimshaw, Wulf 1997, S. 40 (meine Hervorhebung; ab 1999 verwendete Grimshaw zunehmend auch den „Grid“-Begriff).

Parallel dazu liefen seit 1990 mehrere vom National Center for Supercomputing Applications (NCSA) initiierte Projekte für „National Collaboratories“, bei denen man sich ebenfalls auf die Resource-Sharing-Philosophie der 60iger Jahre besann.⁵⁹ Als Leitbegriff setzte sich hierbei zunächst neben „High-performance Distributed Computing“ und „Wide-area Computing“ vor allem „*Metacomputing*“ durch, das bereits 1887 vom Koordinator *Larry Smarr* verwendet worden war. Erst auf der Supercomputer Conference am Argonne Laboratory 1997, d.h. nach Abschluss der NCSA-Projekte, wurde von Smarr der Begriff „National Technology Grid“ als Bezeichnung für die vernetzten Supercomputer Centers der im Jahr zuvor gegründete „National Computational Science Alliance“ eingeführt: „The Grid is a persistent evolving prototype of the early 21st century computational and information infrastructure. The goal is to make this national-scale metacomputer as usable as standalone supercomputers were during the past decade.“⁶⁰ Die oft als „Fathers of the Grid“ apostrophierten *Ian Foster* und *Carl Kesselman*, die seit 1995 am Argonne National Laboratory eine Reihe von Adhoc-Verkopplungen von Hochleistungsrechnern zu „virtual supercomputers“ realisierten, bezeichneten die „distributed computing infrastructure for advanced science and engineering“ noch bis 1997 meist als „Metacomputing“: „We use the term *metacomputer* to denote a networked virtual supercomputer, constructed dynamically from geographically distributed resources linked by high-speed networks.“⁶¹.

Erst seit der von ihnen geleiteten Supercomputer Conference von 1997, die auch ihrem Grid Manifesto-Buch von 1999 zugrundelag, verwendeten auch sie für ihr Konzept konsequent *Grid Computing* und seltener *Utility Computing* als Leitbegriffe. Sie stellten nun erst einen Bezug zwischen dem „Computing Grid“ und dem „Electrical Grid“ her, wobei sie immer wieder auf den in Analogie zu den „public utilities“ Strom- und Telefonnetz geschaffenen Computer Utility-Ansatz der Time-Sharing-Pioniere Corbatò, McCarthy und Kleinrock verwiesen.⁶² Nicht zuletzt aus dieser Analogie leiteten sie wie vor ihnen schon Larry Smarr die Erwartung ab, dass die Grid-Technologie sich nicht auf die temporäre Ressourcenbeschaffung für rechenintensive wissenschaftliche Anwendungen

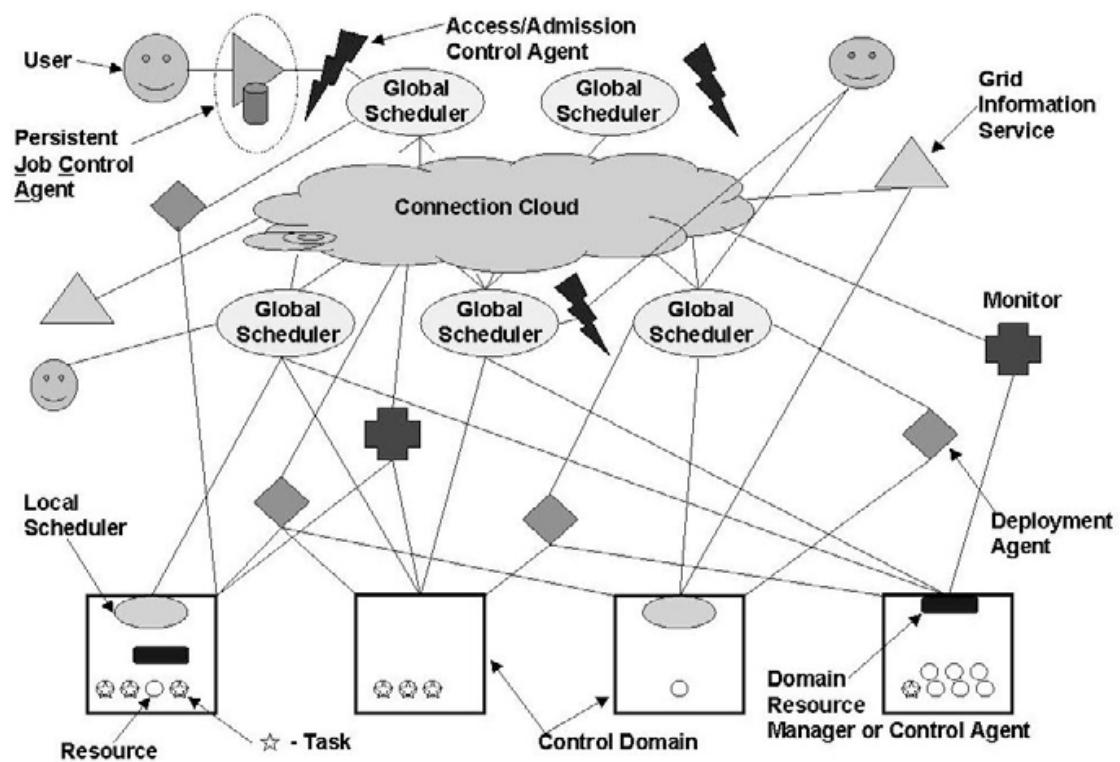
⁵⁹ Siehe Smarr, Catlett 1992.

⁶⁰ Smarr 1997, S. 30; siehe auch Stevens, Woodward u.a. 1997.

⁶¹ Foster, Kesselman 1997.

⁶² Foster 2002a: S. 56, 52 („The concept of sharing distributed resources is not new. In 1965, MIT's Fernando Corbató and the other designers of the Multics operating system envisioned a computer facility operating 'like a power company or water company'.“ Foster 2002b; zur Entstehung siehe bes. auch Braverman 2004.

beschränke, sondern in der Erweiterung zu einem allgemeinen „utility computing environment“ die „next revolution after the Internet and the Web“ darstelle und als allgemeine Infrastruktur im 21. Jahrhundert so umwälzend wirken werde wie im 20. Jahrhundert die Elektrizitätsnetze.⁶³



Hierarchisches Architekturmodell für das Grid-Resourcen-Management mit Cloud-Begriff für das Verbindungsnetz von 2000⁶⁴

Die Grid-Pioniere konzipierten das neue Paradigma als einen föderativen Verbund von physikalischen Rechner-, Speicher- und Netzwerkressourcen (Grid Fabric), von Betriebssystem- und Anwendungssoftware-Plattformen sowie „Service Families“, über die selbständig bleibende Workstations, Server oder Supercomputer für rechenintensive Aufgaben auf Zeit zu einem von einer virtuellen Organisation getragenen kooperativen Cluster zusammenschlossen werden.⁶⁵ Die Ressourcen-Nutzung erfolgt dabei im „Selbstmanagement“: „Die Teilnehmer am Grid-Markt vereinbaren und kontrollieren eigenständig die (Spiel-)Regeln der Kooperation. Bei veränderten Umfeldbedingungen entwi-

⁶³ Siehe dazu Foster, Kesselman 1999, S. 15 ff.; Baker, Buyya, Laforenza 2002, S. 1464; Chetty, Buyya 2002.

⁶⁴ Buyya, Chapin, DiNucci 2000, S. 23.

⁶⁵ Zu Grid-Architekturen und Grid-Service-Protokollstacks siehe Cafaro, Aloisio 2011, S. 5 f.; Hwang, Fox, Dongarra 2012, Kapitel 7.

ckeln die Beteiligten in einem selbstorganisierten Prozess ihr Regelsystem weiter.“⁶⁶ Die Folge war jedoch, dass die meist nur lose gekoppelten und geographisch weit verteilten Grids sich durch starke Heterogenität und Fluktuation ihrer Ressourcen, Dienste und Protokolle auszeichneten.⁶⁷ Daher übertrug man die Ressourcenverwaltung und zeitliche Koordination der „computational-, data-, application- and information services“ bald einem „Grid Broker“. Über das Internet weiteten sich die Reichweiten von Grids nach und nach aus, am Ende entwickelten sich daraus die Visionen eines globalen virtuellen Supercomputers.

Ausschlaggebend für die soziale Architektur des Grid bzw. von Zusammenschlüssen mehrerer Grids zu einer „Grid Family“ bzw. einem „Intergrid“ ist die föderative kollaborative „Governance“. Seinem Wesen nach ist das Grid somit ein „resource pool for some purpose, that are *not subject for centralized control*.“⁶⁸ Gegenüber dem späteren Cloud Computing fehlte jedoch noch eine durchgängige Virtualisierung der Ressourcen und vor allem ein von einem „centralized service hub“ verwaltetes „software environment“.⁶⁹ Die Folge war, dass „Grids meet Too much Computing, Too much Data and never Too much Simplicity“.⁷⁰ Wegen des geringeren Virtualisierungs-Potentials mussten Grid-Infrastrukturen zudem auf Spitzenlast hin dimensioniert werden und waren daher gegenüber den späteren Cloudsystemen unwirtschaftlicher.⁷¹ Doch mit der Adoption der Grid-Technologie im Business-Sektor traten um 2000 sehr bald spezialisierte „grid provider“ auf, die von „utility data centers“ aus in kommerziellen „Service Grids“, „Utility Grids“ oder „e-Utilities“ standardisierte Infrastruktur-, Plattform- und Anwendungs-Funktionen als Dienstleistungen anboten. Damit deutete sich noch innerhalb des Grid-Konzepts ein Übergang zu der von einem Platform-Provider zentral gemanagten Cloud Computing-Architektur an.

⁶⁶ Eymann 2006, S. 69.

⁶⁷ Cafaro, Aloisio 2011, S. 2. ff. nennen folgende Probleme des Metacomputing: „unpredictable structure; heterogeneity; dynamic behavior; multiple administrative domains“.

⁶⁸ Siehe die „Grid Checklist“ in Foster 2002b, S. 2, meine Hervorhebung.

⁶⁹ Armbrust, Fox, Griffith 2010, S. 51; Rittinghouse, Ransome 2010, S. XXXI.

⁷⁰ Fox, Pierce 2007.

⁷¹ Youseff, Butrica, Da Silva 2008, S. 8.

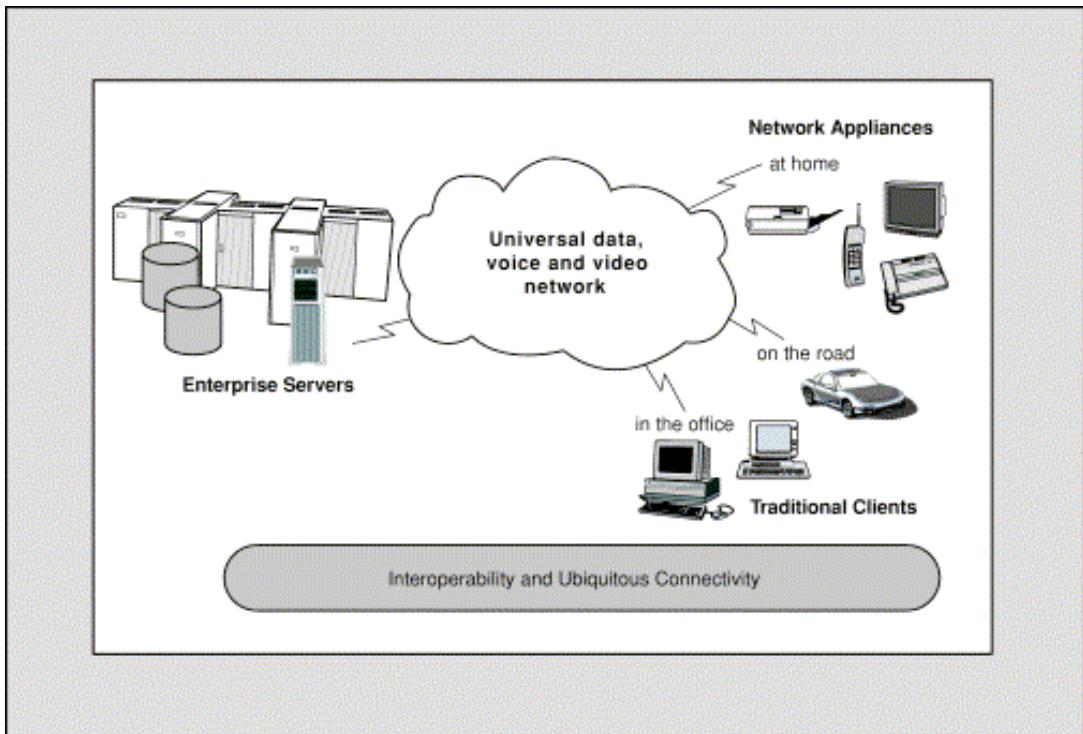


Figure 12. Network computing architecture.

Das zentral von Mainframes gesteuerte „Network computing paradigm“ der IBM von 1998⁷²

5. Die Anfänge der Rezentralisierung: „Network Computing“ und „Utility Computing“

Im Gegensatz zur charakteristischen dezentral verteilten Architektur des ursprünglichen Grid Computing wies das 1995 von *Larry Ellison*, dem Gründer und Leiter der führenden Firma für Datenbank- und Unternehmens-Software Oracle, und dem Oracle Senior Vice President of New Media *Farzad Dibachi* konzipierte „Network Computing“ bereits deutlich in Richtung Rezentralisierung. Bei der NC-Initiative waren nicht mehr akademische Forschungsinstitutionen federführend, sondern vorrangig eine Allianz der IT-Konzerne Oracle, IBM, Sun Microsystems und Apple, die die Definitionsmacht von Intel und Microsoft im PC-zentrierten Computing und damit die Vorherrschaft des „Wintel-Imperiums“ auf dem Computermarkt durch eine Neuordnung der sozialen Architektur der Rechnerlandschaft brechen wollten. Sie hofften die mit jeder neuen Prozessorgeneration in seinem Leistungsumfang erweiterten und dadurch immer komplexer werdenden Universal-PCs durch eine ganze Familie von funktionsreduzierten „Internet appliances“ zu ersetzen, deren minimale Grundfunktionalität sie als „network

⁷² Sweeny 1998, S. 682, Abb. 12.

computer“ (NC) definierten und standardisierten.⁷³ Das neue Paradigma zielte damit auf eine neue Akteurs- und Medienkonstellation, die den universellen Personal Computer in ein Ensemble neuer dedizierter Medien auflöste, dessen aktives Funktionszentrum hinter dem Rücken des Nutzers von einem Provider organisiert wird. Ellison und Dibachi integrierten damit die dezentralen Konzepte von Donald Normans „Information Appliances“ und Mark Weisers „Ubiquitous Computing“, die den konkreten Akteurskontext noch weitestgehend ausgeblendet hatten, in eine zentralistische soziale Architektur des Gesamtsystems und machten so aus dem „Invisible Computer“ einen kommerziell betriebenen ‚Invisible Mainframe‘.

Wie die späteren „Cloud clients“ sollten diese Netzwerkrechner von zentralen „application and service providers“ über das Netz konfiguriert und gewartet und mit Java-basierten „fine-grained software components“ versorgt werden, die nach dem Pay-per-Use-Ansatz abzurechnen waren. Als Bedienkonzept sah das „Thin-client networking“ vor, dass für den Benutzer nur das über den Netscape-Browser Navigator bzw. über einen graphischen Oracle-Browser aufgerufene Interface der Anwendungssoftware das System darstellt und dass alles andere vor ihm verborgen wird.⁷⁴ Die Benutzer dieser „Thin Clients“ waren so zwar von den Problemen der „Fat Client-PCs“ befreit, doch sie gerieten nun mangels ausreichender eigener Rechen- und Speicherkapazitäten in Abhängigkeit von einem „Server-Based Computer System“, in dem die Betreiber von Rechenzentren, Datenbanken sowie Software- und Diensteanbieter die Systemführung beanspruchten. Das Network- bzw. Thin-Client-Computing zielte über den Business-Bereich hinaus besonders auf den Home-Markt, im Mittelpunkt der NC-Vermarktung standen daher neben den NCs die mobilen Handhelds und Set-Top-Boxen für den Fernsehempfang. Die Funktionalität und der Content sollte bei diesen „consumer appliances“ vor allem im Hinblick auf die vorrangig kommunikative Nutzung soweit wie möglich auf zentrale Server bzw. Provider verlagert werden.

Der NC-Pionier Larry Ellison propagierte ähnlich wie die späteren Cloud-Promotoren das „network-centric computing“ als eine weltverändernde Revolution, die in logischer Konsequenz als drittes Stadium auf das „host-centric computing“ der 60er/70er Jahre und das „desktop-centric computing“ der 80er/90er Jahre folgen werde. Als Begründung

⁷³ Revett, Knul, Stephens 1997, S. 172; Roth 2009.

⁷⁴ Gollick 1999, S 33.

diente ihm auch bereits die Analogie mit dem Strombezug bei einer „public utility“ und der einfachen, fest in den Alltag integrierten Stromnutzung : “A PC is a ridiculous device; the idea is so complicated and expensive. What the world really wants is to plug into a wall to get *electronic power*, and plug in to get data”.⁷⁵ Vom Anspruch her handelte es es bereits um Cloud Computing, doch zu einer Verwirklichung der Konzeption fehlten nahezu alle Voraussetzungen, insbesondere die Clusterung von massiv-parallelen Servern, elaborierte Virtualisierungs-Methoden, standardisierte Service-Plattformen und und ausgereifte nicht-prorietäre Anwendungssoftware. Die Anti-PC-Offensive scheiterte zudem schon Ende der 90er Jahre an den erheblichen Nutzungsmängeln der angebotenen NCs, Pen-Computer und Set-Top-Boxen, an den zu geringen Übertragungsgeschwindigkeiten, fehlenden Geschäftsmodellen und vor allem an dem durch den NC aus ausgelösten drastischen Preisverfall bei PCs. Der Fehlschlag der so stark ins Zentrum gerückten Alternativmedien zog dann um 2000 auch den schnellen Niedergang des „Network Computing“ als Nachfolge-Paradigma nach sich, so dass der Begriff in der Entstehungsphase des Cloud Computing keine Rolle mehr spielte.

⁷⁵ Ellison in seiner berühmten Attacke auf Bill Gates und Windows-95 beim European IT-Forum der International Data Corporation am 4.9.1995 in Paris, siehe Roth 2009 (meine Hervorhebung); Ferranti, Marc, 1995; Comerford 1997, S. 21.



Firmenlogo der Gründerfirma Salesforce mit der „End of Software“-Message und später hinzugefügtem Cloud-Symbol

6. Die Startup-Firmen des Cloud Computing als Service-Anbieter für E-Commerce

Die seit dem Ende der 90er Jahre entstandenen ersten Gründungen von Cloud-Service-Unternehmen zielten im Unterschied zur „Network Computing“-Offensive der großen IT-Konzerne nicht auf eine marktstrategische Übernahme der PC-Welt. Dem ehemaligen Oracle Senior Vice President *Marc Benioff* ging es als einem Anhänger humanitärer, philanthropischer Social Business-Ideen mit seiner ab 1996 geplanten und 1999 gegründeten Cloud-Pionierfirma *Salesforce.com* primär um eine Vereinfachung der Computerbedienung und eine Demokratisierung der Software durch „a really clean, easy interface that could be customized for each user“. Seine Grand Vision war es, „to make software easier to purchase, simpler to use, and more democratic without the complexities of installation, maintenance and constant upgrades.“ Missionarisch verkündete er bei der Unternehmensgründung „*The End of Software*“ und beantragte für diesen Slogan sogar ein Warenzeichen. Sein Ziel eines „Extremely easy to use“-Interface und -Service realisierte er mit einem Plattformkonzept und die Software-Versorgung per Browser: „This model made software similar to a utility, akin to paying a monthly electric bill.“⁷⁶ Mit seiner *Utility-Computing*-Konzeption einer zentralen „shared infrastructure“ und dem „On-demand“ Business-Software-Vertriebsmodell wurde Benioff ein erfolgreicher

⁷⁶ Benioff, Adler 2007, S. XXI, 3 f., 12.; Carr 2008a, S. 68-71.

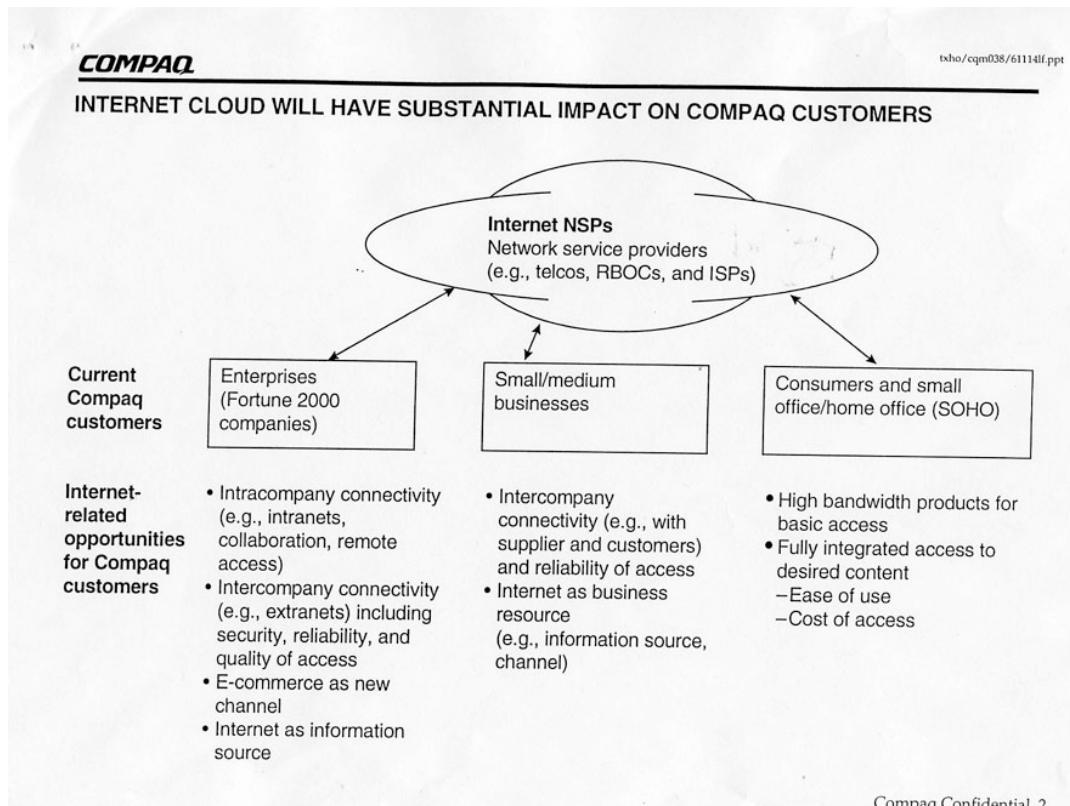
Pionier von „*Software-as-a-service*“ und z. T. auch von dem Dienst, dem er den Namen *Platform-as-a-service* gab. Seinem Vorbild folgten viele weitere Gründerunternehmen, doch nur wenige hatten Bestand. Salesforce.com gelang als einziger der frühen Startup-UP-Firmen, sich neben den späteren Marktführern im Cloud Computing-Sektor behaupten, ohne jedoch je zu ihnen aufzuschließen.

Ein weiteres Pionierunternehmen, die 1997 von dem Electronic Computer Design-Ingenieur *John P. Gorst* gegründete „*Internet Utility Company InsynQ, Inc*“, folgte ebenfalls dem „*Software-as-a-service*“-Ansatz, vertrat aber von Beginn an ein stärker auf Business-Kunden zugeschnittenes „*always-on utility computing model*“. Danach sollten kleine und mittlere Firmen ihre Anwendungssoftware und IT-Services zu einer „*shared computing platform*“ bzw. einer „*computing power plant*“ outsourcen, um dadurch ohne Rücksicht auf Betriebssysteme von jedem webfähigen Computer in der Welt aus Zugriff zu haben. Während Gorst für mediale Anwendungen im Privatbereich PCs noch für sinnvoll hielt, erwartete er in der Unternehmens-IT „*a migration away from the desktops*“. Auch dieses Unternehmen bereitete der „*Shared utility computing services*“-Philosophie den Boden, kam jedoch wie die meisten kleineren Cloud-Gründerfirmen nicht über eine Nischenposition hinaus.⁷⁷

In den Jahren 1996/97 begann auch einer der führenden PC-Hersteller, die Firma Compaq, mit konkreten Planungen für das Internet-zentrierte IT-Business der Zukunft und verwendete in diesem Zusammenhang erstmals den Begriff „*Cloud Computing*“. In einem vertraulichen Planungspapier vom 14.11.1996 entwickelte der Compaq marketing executive *George Favaloro* in Zusammenarbeit mit *Sean O'Sullivan* eine „*Internet Solutions Division Strategy for Cloud Computing*“, in der sie für die abzusehende „*evolution to cloud computing*“ die künftigen Compaq-Geschäftsfelder skizzierten. Demgemäß müssten sich die Network und Internet Service Provider auf zusätzliche „*cloud services*“ wie „*consumer file storage*“ und „*cloud computing-enabled applications*“ einstellen, für die Compaq die entsprechenden Server liefern solle. Der Schwerpunkt lag zwar insgesamt noch ganz klar bei Telekommunikationsdiensten von Unter-

⁷⁷ Siehe Gorst 2009. Siehe auch das Nicholas Carr inspirierende Beispiel der 1899 von Michael Sullivan gegründeten Firma VeriCenter, die Datacenter-Services nach dem „utility model“ der Elektrizitätsversorgung anbot, Carr 2008a, Prolog.

nehmen, doch das anvisierte Geschäftsmodell und wie auch die äußerst ehrgeizigen strategischen Zielsetzungen ähneltem schon stark dem späteren Cloud Computing.⁷⁸



Überblick der „Cloud-Services“ der Compaq „Internet Solutions Division Strategy for Cloud Computing“ vom 14.11.1996 mit der ersten bisher nachgewiesenen Verwendung des Begriffs „Cloud Computing“⁷⁹

Das vierte markante Pionierunternehmen war die Firma *Loudcloud Inc*, die als erste die Cloud-Metapher in ihrem Namen verwendete. Sie war 1999 von *Marc Andreessen* nach dem Verkauf seiner Browser-Firma Netscape aufgebaut und im Februar 2000 in Betrieb genommen worden. Mit dieser Gründung, die darauf abzielte, „to democratize the Web“, wollte er jungen E-Commerce-Startups eine Software- und Hardware-Plattform bieten, um ihnen den Aufbau einer eigenen Web-Präsenz neben der eigentlichen Gründung zu ersparen. Sie konnten aus einem „menu of preexisting clouds“ auswählen und jeweils einzelne oder gebündelte Dienste subskribieren, die dann je nach Anspruch genommenen Leistungen monatlich abgerechnet wurden. Die Palette von „Smart Cloud Services“ umfasste Anwendungs-Dienste („data base cloud“, „application server cloud“,

⁷⁸ Siehe den Draft 3 des Dokuments aus dem Compaq-Firmenarchiv unter der URL: <http://de.scribd.com/doc/70929365/Draft-3-2> (meine Hervorhebung).

⁷⁹ Siehe den Link zum Dokument in Regaldo 2011.

„full web cloud“ und „security cloud“), das Betriebssystem „Opsware“ für die standariserte Dienste-Konfiguration, Bandbreitenzuteilung und das „application scaling“ sowie ein „operational environment“ für das Management der Hardware und Zusatzdienste wie eine „mail cloud“.⁸⁰ Der Cloud-Begriff war hier relativ eng gefasst, er bezeichnete die unterschiedlichen Servicebereiche und bezog sich nicht auf ein neues Computing-Paradigma. Gleichwohl wiesen Andreessens „Smart Cloud Services“ doch in Richtung auf das spätere „Cloud Computing“, war es doch seine Vision „to build the web's next power play: custom-designed, infinitely scalable sites that blast off a virtual assembly line“.⁸¹ Doch trotz eines beachtlichen Startkapitals und einer strategischen Allianz mit Microsoft musste die Firma von einem geleasten Datacenter aus mit geleasten Computern operieren. Sie litt, da der Börsengang im April 2001 mitten in die Dot.com-Krise geriet, schnell an Kapitalmangel und musste nach großen Verlusten bereits im Juni 2002 das operative Geschäft an die etablierte Datenbank-Plattform „Electronic Data Systems“ verkaufen.

Wie Loudcloud wurden auch die anderen Cloud-Pioniere von dem Ende der Dot.com-Blase schwer getroffen. Die erst 1997 begonnene Gründerwelle endete so schon nach fünf Jahren und mit ihr verschwand auch der Cloud-Begriff erst einmal wieder von der Bildfläche. Charakteristisch für diese Cloud-Anbieter der ersten Stunde waren meist recht begrenzte Zielsetzungen, eine starke Konzentration auf das Software-Segment („Software-as-a-service“) und oft auch eine bescheidene Kapitalausstattung. Insgesamt waren sie noch stark in der Web 1.0-Kultur bzw. in der klein- und mittelbetrieblichen eBusiness-Ära verankert und entsprechend auf eine Demokratisierung des Internet bedacht. Die Visionen und Szenarien waren hier nicht auf eine Zentralisierung der sozialen Architektur des Computing gerichtet, sondern vor allem auf Userinteressen und Nutzungsqualitäten. Mit ihrem Niedergang minderte sich daher auch die Chance für stärker dezentral und kooperativ ausgerichtete Cloud Computing-Architekturen.

⁸⁰ Vencat, Girish 2009; Gerald 2000.

⁸¹ Sheff 2003; Elgin 2001.



Computing Utility Vision der IBM nach Vorbild klassischer Infrastrukturnetze (Saracini 2010)

7. Die Grid-/ Utility-Konzepte von IBM, SUN, HP und Microsoft zur Rückgewinnung der Kontrolle über die PC-Internet-Welt

In der zweiten Hälfte 90er Jahre verschob sich nach dem Scheitern von Ellisons Network Computing-Initiative bei den Monopolisierungs-Bestrebungen der Fokus der Bemühungen um ein neues Computing-Paradigma von einem brancheninternen Konflikt um die Dominanz im Hardware- und Software-Markt hin zu einer mehr oder weniger koordinierten Gesamtstrategie der etablierten IT-Giganten IBM, SUN, HP und Microsoft. Es ging ihnen nun primär um eine Rückgewinnung des ihnen durch den Aufstieg des World Wide Web entglittenen Einflusses auf die sich weitgehend selbst organisierende vernetzte PC-Welt. Die intensivsten Anstrengungen zu einer Umwälzung der sozialen Architektur des Computing gingen dabei in der zweiten Hälfte der 90iger Jahre eindeutig von der *IBM* aus. Die Federführung hatte dabei die 1995 neu geschaffene „Internet Division“ unter ihrem General Manager *Irving Wladawsky-Berger*. Als „IBM's technology strategist“ war es, der auf eine Bündelung der unterschiedlichen Ansätze und Strategien im Konzern für eine Ablösung des „PC-Centric Computing“ durch eine „network-centric computing strategy“ hinwirkte, die, indem sie sich den „new technology ideas from

external communities“ öffnete, das Internet und seine Kulturen zu umarmen trachtete.⁸² Mit der wesentlich von ihm initiierten und geleiteten „IBM e-business strategy“ von 1997 und dem „On Demand Computing“-Programm von 2002, die das traditionelle IBM-Mainframe-Geschäft nun konsequent auf Internet-Technologien ausrichten sollten, wurde er einer der maßgeblichen Architekten des zentralistischen Cloud Computing, ohne dabei jedoch den Cloud-Begriff zu verwenden. Es gab in der Integrationsphase überhaupt noch keine einheitliche Bezeichnung für das „Post-PC-Era Paradigm“, man arbeitete bei der IBM vielmehr mit einem Bündel von Leitmetaphern, die sich anfangs stärker auf die Hardware-Konfiguration fokussierten, später auch die Software-Architektur und die Geschäftsmodelle einbezogen.

Von der durch IBM angestrebten Anbindung der unüberschaubaren Vielfalt der Clients an „central servers“ versprach sich Wladawsky-Berger endlich die Rückkehr der Ordnung in die IT-Welt, denn „what has been built is the computing equivalent of the Tower of Babel“.⁸³ Als ordnende Instanz im Grid waren die „future data centers“ ausersehen, die sich im Unterschied zu den bisherigen isolierten Mainframe-Monolithen zu vernetzten Parallelrechner-Clustern weiterentwickeln und ein weltumspannendes Hochleistungsnetz von Supercomputern mit jeweils der Rechenkapazität von IBMs berühmtem „Big Blue“ bilden sollten. Gerade die ubiquitäre Verbreitung von Computerleistungen würde die Bedeutung und Größe dieser neuen „data centers“ im Grid immer mehr wachsen lassen und schon in wenigen Jahren werde die Zentralisierung zu einer globalen „infrastructure with *giant server farms*“ führen. Mit einem “massively large power ,utility’ grid, das Unternehmen und Privathaushalten jederzeit weltweit zur Verfügung steht, wollte man endlich auch die bestehende Organisationslücke beim Ressourcenmanagement im Grid schließen.⁸⁴

Die zum Grid hinzugefügte Zentralinstanz sollte den Benutzern auch die Mühe abnehmen, die Ressourcenaquisition im Grid selber zu organisieren und stattdessen die Rechenzentrumsfunktion in Gestalt von im Superrechner-Server virtuell erzeugten kundenspezifischen Hardware-Plattformen als Infrastruktur-Service anbieten. Erreicht wurde dies durch aus Softwarekomponenten bestehende virtuelle Maschinen (VM), die als

⁸² Zur Strategie des „Embracing the Internet“, „Embracing the Internet “Culture” siehe Wladawsky-Berger (2005c, 2005d).

⁸³ Wladawsky-Berger, zit. nach Economist vom 12.9.1998.

⁸⁴ Joseph, Fellenstein (2004), S. 10 ff.

eine Abstraktionsschicht über die physikalischen Rechner gelegt wurden, und die nun die Eigenschaften von Prozessoren, Festplatten usw. simulierten. Über einen zentralen Hypervisor wurde dann der Zugriff mehrerer virtueller Maschinen auf die physikalischen Rechner mit dem Ziel optimaler Auslastung gesteuert. Ziel des asymmetrischen Grid-Konzeptes war es, durch die Virtualisierung aller Host- und Netzressourcen die Unterschiede physikalischer Infrastrukturen vor den Benutzern zu verbergen und ihnen den Anschein des Zugangs zu einem homogenen, dynamisch rekonfigurierbaren und leicht bedienbaren virtuellen Großrechner zu vermitteln.

Die Firma konnte dabei auf die eigene Tradition von Virtualisierungstechniken rekurren, die bereits in die frühen Time-Sharing-Projekte der 60er Jahre zurückreichte. So war schon mit dem IBM „Control Program/Monitor-System“ (CP/CMS) von 1964, das sich an das von Fernando Corbató am MIT entwickelte „Compatible Time-Sharing System“ (CTSS) angelehnt hatte, ein Steuerprogramm für eine virtuelle Maschinenumgebung entwickelt worden, d.h. ein „multiprogramming system that uses virtual machines to organize independent job streams“. Hinzu kamen mit dem „Conversational Monitor System“ zur „desktop virtualization“ vereinfachte Benutzeroberflächen, über die viele User ‚gleichzeitig‘ interaktiv auf Großrechner zugreifen konnten. Fortgeführt in der „IBM Virtual Machine Facility“, sorgten diese Hardware- und Software-Abstraktionen für eine bessere Auslastung der IBM 360- und IBM 370-Rechnerfamilie.⁸⁵ Aufgrund dieser Erfahrungen mit logischen Architekturen vermochte die IBM nun auch wichtige Anstöße zu geben für die Entwicklung einer kompletten Palette von Virtualisierungs-Tools für Infrastrukturen, Betriebssysteme, Plattformen, Anwendungen, Dienste, Daten und Netzwerk, die zusammen erst den Aufstieg des Grid-, Utility- und Cloud Computing ermöglichten. Denn Computing-Prozesse waren nun nicht mehr an bestimmte Hardware-Ausstattungen und Betriebssysteme und Anwendungssoftware gebunden.

Mit dem Festhalten am Grid-Begriff verdeckte das IBM-Strategiekonzept, dass mit dem Hinzutritt von kommerziellen Infrastrukturanbietern, die die Bereitstellung und das Management von IT-Ressourcen übernahmen, die soziale Architektur des Grid grundlegend verändert wurde. Denn eine „service-oriented computing platform“ und „commercial on-demand utility grid services“ lösten nun die bisherigen „grid commu-

⁸⁵ Zur Entstehung der Virtual Machines der IBM siehe die Darstellung des CP/CMS-Entwicklers Robert Creasy 1981, S. 483 ff., 487.

nities“ und „peer-to-peer utilities“ ab.⁸⁶ Das föderative Ressource-Sharing durch gleichberechtigte, mit eigenen Ressourcen ausgestattete Teilnehmer im Scientific Grid Computing wichen damit im kommerziellen „grid utility concept“ der Nutzung großskaliger Zentralressourcen, über die externe Provider oder gar mächtige IT-Fabrics verfügten, die mehr oder weniger abhängigen „Clients“ virtuelle Ressourcen auf Zeit zur Verfügung stellten. Damit aber entstand im Grid eine „Netzführerschaft“, bei der große IT-Dienstleister jeweils eine „single-organization control“ ausübten und dabei die Spielregeln der Ressourcennutzung vorrangig nach ihren ökonomischen Kriterien bestimmen und kontrollieren konnten.⁸⁷

Während das modifizierte Grid-Konzept im IBM-Strategieplan auf eine Neuordnung der Ressourcenorganisation und daraus folgend auf einen Wandel der Governance im Internet abhob, zielte das eng zusammenhängende Begriffspaar „On-Demand Computing“ und „Utility Computing“ mehr auf die veränderten Nutzungsphilosophien und Geschäftsmodelle, mit denen die bisherige freie Aushandlung der Hardware-/Software-Voraussetzungen, der Anwendungs- und Abrechnungsmodalitäten durch eine konfektionierte Dienstpalette und ein entsprechendes Geschäftsmodell abgelöst werden konnte. Der Begriff „Utility Computing“ schließlich bündelte die Zielsetzungen des IBM-Strategiekonzeptes für ein neues Computing-Paradigma und ein „Second Generation Internet“ in einer suggestiven Großvision, die sowohl das grundlegend veränderte IBM-Geschäftsmodell auf einen prägnanten Begriff brachte als auch die angestrebte weitreichende Umwälzung der IT-Welt als logischen, notwendigen Evolutionsschritt legitimierte.

Wie mit diesen strategischen Zielsetzungen so nahm das neu fokussierte Utility-Konzept der IBM auch in der metaphorischen argumentativen Ausgestaltung wesentliche Momente des Cloud Computing-Leitbilddiskurses vorweg. So fungierte die nach 2000 häufiger herangezogene *Metapher der Informationssteckdose* als zentrales Werbeargument für das Outsourcing von Rechnerinfrastruktur, Datenbankfunktion, Betriebssystem und Anwendungssoftware an externe „service providers“. Die suggestive Analogie von „electric“ und „information utility“ sollte die Unsinnigkeit des Besitzes eigener Computing-Ressourcen und die Selbstverständlichkeit des neuen Geschäftsmodells

⁸⁶ Joseph, Ernest, Fellenstein (2004), 624 f.

⁸⁷ Smith (2004), S. 5; Eymann (2006), S. 69. Zum Begriff „single-organization control“ siehe Barroso, Hölzle 2009, S. 2.

„On-Demand-Computing“ plastisch vor Augen führen. Die Parallelisierung von Strom, Gas Wasser und Computing ebnete zugleich die großen strukturellen Unterschiede zwischen energetischen und informationellen Infrastrukturen ein und verdeckte damit, wie tief das neue Computing-Paradigma in Wirtschaft, Gesellschaft und Alltagswelt eingreifen würde. So entsteht beim „utility computing approach“ wegen der im Vergleich zum Strom bedeutend schwierigeren Transformierbarkeit von Daten und Informationen eine weitaus größere Abhängigkeit vom jeweiligen Provider. Vor allem aber verdeckt die Utility-Metapher infolge der impliziten Fixierung auf Großanlagen die aus der Zentralisierung der Informationsverarbeitung und Datenspeicherung resultierende gesellschaftliche Machtverschiebung zugunsten führender IT-Konzerne und Betreiber von Mega-scale Data Centers und Internet-Plattformen, die Bedingungen und Preise der Infrastruktturnutzung diktieren und die sich dadurch zum Machtzentrum in einem „Monolithic Internet“ entwickeln könnten.

Die IBM hatte damit zwischen 1996 und 2002 die Bausteine für ein zentrales Cloud-Computing-Konzept entwickelt und sich bemüht, diesem als dem künftigen Paradigma große Publicity zu verschaffen, wobei man ständig zwischen den strategischen Leitbildern „Great Global Grid“, „neue Ära des On-Demand Computing“, und „Next Utility“ wechselte. Grid und Utility, oft zusammengezogen zu „Grid Utility“ wurden bis 2006/07 die Standardbezeichnungen für alle Arten von ‚Cloud‘-Architekturen bzw. deren Vorstufen. Obwohl die IBM wesentliche Anstöße gegeben hatte, zögerte das Management eine volle Realisierung des Programms aber noch hinaus, denn für ein Komplettangebot seien noch etliche Hürden zu überwinden und müsse erst die erforderliche „infrastructure with giant server farms“ errichtet werden, bevor die Diensteplattform starten könne und ehe das „utility-based computing becomes mainstream“. Auch waren wegen der traditionellen „IBM culture“ im Konzern selber noch manche Hürden zu überwinden, sodass das 2002 vom Firmenchef Samuel J. Palmisano angekündigte 10 Mrd. \$ Investitionsprogramm für die “On Demand”-Offensive nur zum kleinen Teil umgesetzt wurde.⁸⁸ Insgesamt gelang es der IBM mit dem Fokus Cloud Computing-Hardware und -Systemsoftware nicht, einen kompletten Cloudstack und eine breite Palette von Clouddiensten

⁸⁸ Lyons 2005. Noch 2008 erwartete der inzwischen viel skeptischere Wladawsky-Berger, dass hauptsächlich kleinere Firmen dem „pattern of electricity's evolution a century ago“ folgen und ihre Rechenzentren schließen würden, während die großen Unternehmen eher an der Eigenversorgung festhielten. Auch gegenüber Carrs sehr weitgehender Parallelisierung von „electricity and computer utilities“ hatte er nun wegen der grundlegenden Unterschiede beider Infrastrukturen Vorbehalte: „So it is too early to tell how IT clouds will evolve.“ Wladawsky-Berger 2008a, 2008b.

zu entwickeln und damit die erhoffte Führung im Cloud-Plattform-Markt zu erlangen. So kam es, dass der Vorreiter bei der konzeptionellen Entwicklung des Cloud Computing zwar ab 2003/04 eine Reihe von Grid-Diensten von seinen zu einem Verbund zusammengeschlossenen „BlueGene“-Rechnern sowie Business-on-Demand-Dienste und Toolboxes für die zentrale Verwaltung von Grids anbot, doch beim eigentlichen Aufbau von Clouds den Newcomern Amazon und Google das Feld überließ und erst ab 2007 eine Cloud-Strategie und ab 2008 Services für Geschäftskunden entwickelte und so erst mit Verspätung einer der ganz großen Cloud-Provider für den Business-Bereich wurde.⁸⁹ Ähnlich erging es auch den anderen großen Cloud-Pionieren Sun, HP und Microsoft, die bereits um 2000 Grid-Utility-Programme starteten, aber erst Jahre später voll in den Cloudmarkt eintraten.

Auch die Firma *Sun Microsystems*, die 1984 den berühmten Slogan „The Network is the Computer“ geprägt hatte, ging ab 1999 mit der „*Commercial Grid Strategy*“ von der früheren dezentralen Philosophie zu einem „more central management control“-Ansatz über. Nachdem sie sich mit der JavaStation bereits am Network Computing Hype beteiligt hatte, setzte sie nun auf eine „strong platform technology“, bei der das „thin client server model“ mehr oder weniger auf ein bloßes Zugangsmedium reduziert wurde. Längerfristig setzte man auf eine Strategie des „diskless user“ und der „datacenter-less company“⁹⁰ Die Pionier-Firma im Network Computing Sun wollte dabei mit der JavaStation die Führungsrolle im Umsturz gegen den den obsoleten PC übernehmen: „Fifty years from now, people are going to look back and say: 'Did you really have a computer on your desk? How weird.'“⁹¹ Geleitet wurde die frühe Grid-Computing- bzw. „Service-on-Demand“-Entwicklung bei Sun von *Wolfgang Gentzsch*, nachdem Sun 2000 dessen „Gridware“ erworben und zur Basis der 2001 vorgestellten „Sun ONE Grid Engine“ gemacht hatte.

Ziel der Anstrengungen für „Computational Grid Environments“ war es, auf der Basis der Virtualisierungs-Tools der plattformunabhängigen Programmiersprache Java das Grid zu einer offenen „platform for commercial computing“ zu machen und langfristig die

⁸⁹ Wladawsky-Berger-Interview, Hoffman 2002; Miller 2008a. Zu Cloudstacks siehe u.a. Lenk, Klems, Tai, Sandholm u.a. 2009.

⁹⁰ Die JavaStation war 1996 noch von dem damaligen Chief Technical Officer und späterem Google-Chef Eric Schmidt initiiert worden, siehe dazu Hsin 1996 („corporate policy to make it diskless“); Malik 2010.

⁹¹ McNealy in Hof 1999.

wildwüchsigen entstandenen wissenschaftlichen, institutionellen und kommerziellen Grids (Campus Grids, Enterprise Grids, Research Grids, Entertainment Grids, Community Grids) zu einer weltumspannenden „Grid Community“ zu vereinigen. Dieser würde dann eine alle Rechenzentren, „Clients“, Anwendungen und Dienste integrierende Utility-Infrastruktur zur Verfügung stehen, wobei Sun die Rolle eines „category leader in Grid Computing“ beanspruchte.⁹² Auf die „Grid Utility“ mit einheitlicher Architektur könne jeder User wie auf das „power grid“ von jedem Ort der Welt zu jeder Zeit transparent auf Ressourcen und „pay-for-use-grid computing services“ zugreifen, „ohne zu wissen, wo die Rechner stehen, welche Eigenschaften sie haben oder wem sie gehören.“ Die im Hintergrund agierenden Data Centers werden extrem sicher, verfügbar und nahezu immer und von überall von Geräten jeder Art erreichbar sein. Gesteuert wird dies alles durch „eine zentrale Überwachungsinstanz, die jederzeit nachvollziehen könne, welcher Teilnehmer an welchem Ort welche Aufgaben berechnen lässt oder wer welche Leistungen in Anspruch nimmt. Schließlich sah man auch bei Sun die „Grid Engine“ als den Wegbereiter einer neuen „Utility“, die Unternehmen und Privatkunden künftig den Erwerb eigener Rechnerressourcen erspare: „The name “grid computing” refers to the goal of providing computer services to users in the same way that water and electricity are provided. Customers need not own the means to produce their own water or electricity; instead, they outsource that capability to a utility and purchase only as much as they need.“⁹³

Mit dem neuen strategischen Ziel "unifying all the resources in a *compute fabric*", wurde der berühmte Sun-Slogan von 1984 „The Network is the Computer“ nun einfach umgedreht: „If the network is the computer, what is the computer that is the network?“⁹⁴ Die Pionierfirma für das „distributed computing“ hatte sich damit zu einem der extremsten Vertreter des Hochzentralsimismus gewandelt, und dies obwohl ein führender Sun-Ingenieur bereits davor warnte, dass eine Massierung des Computing in wenigen Großtaggregaten ein „huge single-point-of-failure risk“ darstelle und wegen des extrem hohen Stromverbrauches das Vulnerabilitäts-Niveau der „new era of web-based computing“ generell erhöhten.⁹⁵ Doch die hochtrabenden Visionen des „Sun Future Scenario“ zer-

⁹² Gentzsch 2001a, 2001b, 2002, S. 6 ff.

⁹³ Gentzsch 2002, S. 8 f.; Fordahl 2005.

⁹⁴ Kraemer, Dedrick 1999, S. 9; Papadopoulos 2002 („N1 is, simply put, the first architecture to treat the network as a computer“); Moody 2002.

⁹⁵ Siehe Carr 2007.

schlugen sich ohnehin infolge der Auswirkungen der DotcomKrise und der Finanzkrise auf das Unternehmen, die die Übernahme durch Oracle im Jahre 2009 zur Folge hatten. Obwohl Sun mit seinen Tools für eine durchgängige Virtualisierung und seinem modularen Datacenter-Design über die Bausteine für einen kompletten Cloudstack verfügte, wurden die weitreichenden Grid- bzw. Cloud-Pläne nur in kleinem Maßstab realisiert, ein vollständiger Cloud-Service kam sogar erst 2010 zustande. Dem mehrjährigen Übergangsstadium vom Grid- zum Cloud Computing entsprach der zögerliche Wechsel zum Begriff *Cloud Computing*, den man bei Sun lange Zeit für ein bloßes „re-labeling of ‚grid‘, ‚utility‘ and ‚network computing‘“ hielt und daher lange Zeit mied.⁹⁶

Nicht viel anders erging es *Hewlett-Packard* mit dem 1998-2003 entwickelten „*HP Utility Computing*“, das im Gegensatz zum verteil operierenden „*social model of grid computing*“ ein vom Data Center aus organisiertes „*commercial model*“ anstrebte. Bereits 1997 entwarf der langjährige HP-Forschungs-Direktor *Joel S. Birnbaum* seine Vision einer „*Digital Information Utility*“, die analog zu den traditionellen Public Utilities für Wasser und Strom die Gesellschaft über zentrale Infrastrukturen mit Informationen versorgt. Die User sollten ganz im Sinne von Donald Normans „*Design of Everyday Things*“ vereinfachte Computer und „*pervasive information appliances*“ mithilfe von „*intuitive point-and-click-style browsers*“ so einfach wie herkömmliche Elektrogeräte bedienen können. Die kommende Ära der „*Information Utility*“ und „*Information Appliances*“ sah er dabei als das vorerst letzte in der Entwicklungskette Computing-Paradigmen.⁹⁷ Zur Realisierung des „*new paradigm*“ gründete HP 1998 die „*Utility Computing Division*“ und beauftragte ehemalige Computer Scientists aus den Bell Labs unter der Leitung von Frank Barker mit der Entwicklung einer „*computing power plant*“ und eines Software Stacks für eine „*e-services-on-tap*“ genannte Dienste-Plattform. Die Entwicklung eines Cloud-Computing-Vorläufers bei HP beruhte so bis in die Begriffe hinein auf historischen und technologischen Metaphern und Analogien.

Das 1999 angekündigte und Ende 2001 vorgestellte „*HP Utility Data Center*“ war eine Komplettlösung zur Virtualisierung und Automatisierung von Rechenzentrums-Umbgebungen und bildete damit bereits einen Übergang von einem zentral gesteuerten „*grid environment*“ zu einer privaten Cloud. Das „*HP Utility Data Center*“ ermöglichte über

⁹⁶ Martin 2007; Carr 2006; Papadopoulos 2009; Sun Microsystems, Inc (2009), S. 31 ff.

⁹⁷ Birnbaum 1997, S. 6 ff.

eine Management-Plattform die Server-, Speicher- und Netzwerk-Virtualisierung und damit eine flexible Resourcenverwaltung.

Auch bei den HP Labs verfolgte man eine „*Planetary Computing Vision*“: „We envision a world where distributed services execute on a utility that dynamically and securely allocates globally connected server and storage resources on demand ... a *global commercial Grid*“⁹⁸ Doch schon im Herbst 2004 gab die Firmenleitung das „HP Utility Data Center“-Programm wieder auf, da die versprochene Kosten- und Komplexitäts-Reduktion durch die Center-Utility-Lösung nicht eintrat und da man die Nachfrage nach kommerziellen Grids stark überschätzt hatte.⁹⁹ Von 2005 bis 2008 konzentrierte sich HP unter dem neuen Chief Information Officer Randy Mott erst mal um eine Konsolidierung der eigenen IT-Infrastruktur. Während die Firma in dieser Phase lediglich einzelne „HP Utility Services“ anbot, begann sie nach dem Aufbau der eigenen Cloud 2009 unter dem Begriff „Cloud Assure“ mit einem Software-as-a-Service und Beratungsleistungen.¹⁰⁰ Erst mit der 2011 angekündigten, aber erst Ende 2012 voll verfügbaren „HP Cloud-Factory“ und der „HP Public Cloud“ gelang es der Firma, sich als einer der großen Cloud-Plattform-Betreiber zu etablieren.

Angesichts der sich Ende der 90er Jahre intensivierenden Debatten über ein „Ende der PC-Ära“ trat im Jahr 2000 auch Microsoft mit einem ‘Cloud Computing’-Programm hervor. Die Firma hatte neben Intel am meisten von der PC-Massenausbreitung profitiert und zögerte daher einen Strategiewechsel in Richtung Rezentralisierung lange hinaus. Seit 1998 arbeiteten Microsoft-Entwickler an Standards für die Realisierung von ‘cloud-computing ideas’, mit denen die Firma mit einem “big step” das Heft des Handelns im Internet und World Wide Web zurückgewinnen wollte, deren Bedeutung sie so gründlich verkannt hatte.¹⁰¹ Die „*.NET Platform*“ (Dot.NET) sollte bei der “long-term transformation” des Internet in eine “services-rich platform” die Schlüsselrolle übernehmen. Mit der weitgehend von *Paul Maritz* entwickelten Vision und “long-term roadmap” peilte die Firma einen Wandel des bisherigen Geschäftsmodells an, das weitgehend auf der “one body, one computer metaphor” beruht hatte.¹⁰² Statt wie bisher Windows als

⁹⁸ Vickers (2003), S. 10.

⁹⁹ Farber (2004).

¹⁰⁰ Morgan 2008; Carr 2008c.

¹⁰¹ Über Gespräche darüber berichtet der Software-Pionier David Winer, siehe Markoff 2001.

¹⁰² Sándor 2012.

Produkt, als “nice little box for one place” zu vermarkten, sollte künftig “software-as-a-service” angeboten werden.

Im Zuge dieser Erweiterung des Endgerätespektrums wollte man die Bedienung von PCs stark vereinfachen und PC-User-Interface und Internet-Browser integrieren, doch sollten PC, Client-Server Computing und Peer-to-Peer- Netzwerke weiterhin als eigenständige Kulturen bewahrt bleiben. Strategisches Ziel von Microsoft war nämlich kein radikaler Bruch mit dem PC-Internet-Computing, sondern eine Kompromissstrategie, die den Usern die freie Wahl zwischen zentral gesteuerten mobilen “cloud services” und traditioneller Desktop-Nutzung ließ: “We're betting that the whole world does not move back to a mainframe model of centralized computing, that rich servers, rich clients, and rich *cloud services* are important. We're betting that user interface matters, that it's not going to be just a world of browsers, but a world in which people want to locally store, and author, and shape the presentation and data they work with in rich ways.”¹⁰³ Microsoft griff damit als erste der großen IT-Firmen den Cloud-Begriff auf, jedoch nicht im Sinne eines „post-PC paradigm“ sondern eines Bündels von Internet-basierten Software- und Platform-Services. Microsoft machte sich hiermit die Auffassung von Marc Andreessen zueigen, mit dessen Loudcloud Inc. kurz zuvor eine strategische Partnerschaft verabredet worden war.

Dot.NET wurde auf diese Weise ein äußerst ambitioniertes All-in-one-Projekt, eine “far-reaching and many tentacled strategy and vision”, vor deren Komplexität und Risiken die Unternehmenspitze schon sehr bald zurückschreckte.¹⁰⁴ Nach nur 49 Tagen wurde die große Integrationslösung zugunsten einer “non-cloud version of .NET” fallengelassen, d.h. einer fortschrittlich-konservativen Innovation, die dem PC und dem gewohnten Softwarepakete-Verkauf wieder Vorrang einräumte. Microsoft, das im Jahr 2000 im Begriff gewesen war, die Führungsrolle beim Einstieg in das Cloud Computing zu übernehmen, verfiel so mit seinem Rückzug in eine “lost decade”, die von Bürokratisierung und “software-design by committee” geprägt war.¹⁰⁵ Die führende Softwarefirma wurde damit neben IBM, Sun und HP ein weiterer zaudernder und dadurch zunächst scheiternder Cloud-Pionier, der die Initiative den Startups der Web 2.0-Welt Amazon

¹⁰³ Ballmer 2000 (meine Hervorhebung).

¹⁰⁴ Weiss 2001.

¹⁰⁵ Markoff 2007; Sándor 2012; Eichenwald 2012.

und Google überließ. Im Jahre 2005/06 nahm Microsoft dann mit dem Engagement von Ray Ozzie als "Chief Software Architect" einen neuen Anlauf für eine Transformation vom proprietären Softwarelieferanten zu einem offenen Plattform- und Service-Provider. Bill Gates begründete den neuerlichen Kurswechsel in einem Memorandum an den Führungsstab vom Oktober 1905 mit der durch das Internet ausgelösten „services wave“, die Microsoft nun zwinge, seine Strategie auf Plattform-basierte Dienste auszurichten: „We will build our strategies around Internet services and we will provide a broad set of service APIs and use them in all of our key applications. This coming „services wave“ will be very disruptive.“¹⁰⁶



Die Cloud-Metapher in einem Artikel über Microsoft's Dot.NET: „Platform in the clouds“ von 2000
(Weiss 2001, S. 29)

Doch die Firma blieb befangen im Dilemma des Innovators, der das eigene erfolgreiche Kerngeschäft nicht schädigen wollte, und setzte weiterhin auf eine dilatorische Doppelstrategie. So konnte Ozzie seine Pläne für einen Drei-Ebenen Cloudstack nur stufenweise verwirklichen: Zunächst wurde ab 2006 die "physical layer" von massiv-parallelen Mega Data Centers errichtet, eine „Storage Cloud“ vorbereitet, dann ab 2007 die „cloud infrastructure services layer“ und schließlich die „Live platform services layer“.¹⁰⁷ Doch

¹⁰⁶ Gates 2005. Siehe hierzu Ozzies Memorandum „The Internet Services Disruption“ vom 28.10. 2005 sowie Nicholas Blog-Artikel in Rough Type vom 9.11.2005, 27.10.2006; Carr 2008a, S. 63 f, 81 f.

¹⁰⁷ Siehe die Artikel vom 5. 4. 2007, 27.7.2007, 1.10.2007, 21.11.2007 in Nicholas Blog Rough Type. 2010 verließ Ozzie Microsoft schon wieder; Li 2006.

obwohl Microsoft sogar den Bau eines der Mega Data Center-Systeme plante, hielt man daran fest, dass die Cloud auf absehbare Zeit mit den bestehenden Formen des Personal Computing koexistieren sollte. Erst bei seinem Rückzug von der Firmenleitung im Jahre 2008 entwickelte Bill Gates noch einmal ein radikales Umstiegs-Szenario: Die bisherigen und viele neuen Services “will be hosted in our cloud with the big, big data center, and geo-distributed automatically. This is kind of fascinating because it's getting us to think about data centers at a scale that never existed before.“ Das Computing werde sich in wenigen Großsystemen von „Mega Data Centers“ mit Millionen von Servern konzentrieren, bei denen Microsoft eine führende Rolle spiele.¹⁰⁸

Am Ende schaffte keiner der IT-Giganten den Aufbruch in die neue Ära eines hochzentralisierten Computing, die sie in Leitbildern, Paradigmenmodellen und Entwicklungsszenarien verkündet und medienwirksam inszeniert hatten. Der Monopolist der Mainframe-Ära IBM entwickelte zwar wesentliche Ideenkonzepte und Technologien für die Renaissance der Datacenter, war jedoch und noch zu sehr auf sein Hardware- und System-Geschäftsmodell fixiert und zu zögerlich, die Führung bei der Web-Zentralisierung zu übernehmen. Die Promotoren des Distributed Computing und PC-Client-Server-Computing Sun und HP steigerten sich in Visionen und Großszenarien der Rezentralisierung hinein, die ihre Kräfte bei weitem überforderten und blieben letztlich doch mehr an dem Hardwareverkauf als an Plattformdiensten interessiert. Der Monopolist der PC-Ära Microsoft schließlich scheute vor dem radikalen Rollenwechsel vom Software- zum Plattform-Vermarkter zurück und traute sich angesichts erwachender Kritik eine Monopolisierung der Web 2.0-Kulturen auch noch nicht zu. Der entscheidende Anstoß für den breiten Einstieg in die Rezentralisierung des Computing und des Web ging nicht von den führenden PC-, Mainframe- und Software-Herstellern oder den mittelständischen Cloud-Pionierfirmen aus, die seit Ende der 90er Jahre erste Software- und Platform-Services anboten. Ebensowenig von den akademischen Supercomputing- und Grid-Communities, es waren vielmehr die absoluten Außenseiter *Amazon* und *Google*, die der der Cloud-Technologie und dem Cloud-Begriff zum Durchbruch verhalfen

¹⁰⁸ Siehe Bill Gates Keynote: Microsoft Tech•Ed 2008 – Developers, 3.6.2008; Waters 2009.



Amazon und Google („Amazonia“ und „Google Earth“) als die beherrschenden Cloud-Imperien im Internet:
Cartoon im Economist vom 2012 über die konfigierenden „Enclosure“-Strategien der Konzerne¹⁰⁹

8. Amazon und Googles Cloud-Strategien zur Monopolisierung der Web 2.0 Kulturen

Während Oracle, IBM, Sun und HP große Visionen für das Post-PC-Zeitalter verkündeten und Strategien für den Übergang zum Network-, Grid-, Utility- und On-Demand-Computing entwarfen, arbeiteten der führende Online-Händler Amazon und der Suchmaschinen-Betreiber Google ganz im Stillen an der Entwicklung einer ganzen Palette von zentral über das Internet angebotenen Infrastruktur-, Plattform- und Software-Services.¹¹⁰ Das Konzept einer cloudartigen Webplattform ergab sich bei ihnen nicht aus Überlegungen für ein neues Computing-Paradigma oder einer Geschäftsidee, Computing-, Speicher- und Software-Ressourcen zentral als Dienstleistung anzubieten, es entwickelte sich bei ihnen vielmehr unmittelbar aus der Notwendigkeit, für das eigene Geschäftsmodell leistungsfähige rechnergestützte Transaktionszentren zu schaffen. Die dabei errichteten Server-Konglomerate und Webplattformen wurden erst im Nachhinein über

¹⁰⁹ Economist, 1.12.2012.

¹¹⁰ LaMonica 2006b.

eine ganze Reihe von Webservices für Geschäfts- und Privatkunden geöffnet, wodurch sich die anfänglichen Private Clouds stufenweise zu Public Clouds wandelten. Entsprechend spät fanden Amazon und Google Anschluss an die seit den 90er Jahren laufenden Leitbilddiskurse zum Grid-, Utility- und Cloud-Computing.

Auslöser für die Cloud Genese war nicht, wie ein „Economist“-Artikel vermutete, die Absicht, die außerhalb der Hauptstoßzeiten im Weihnachtsgeschäft oft nur zu 10% ausgelasteten Kapazitäten der Amazon-Rechenzentren zu vermarkten, eine Version, die in der Folgezeit viele Autoren übernahmen.¹¹¹ Die Nutzung von Amazon-Rechnerressourcen durch Externe stand anfangs überhaupt noch nicht zur Debatte, es waren in erster Linie massive Probleme der Firma mit der eigenen unflexiblen, nur begrenzt skalierbaren Computing-Infrastruktur, die mit dem schnellen Wachstum des Kerngeschäfts Online-Buchvertrieb und noch weniger mit der Expansion in neue Geschäftsbereiche des Medien- und Warenvertriebs Schritt hielten. Die Rechner- und Software-Ausstattung der ersten Jahre bestand noch aus einem starr gekoppelten System mit einem Web Server, einer Datenbank auf dem Backend und der monolithischen Anwendungssoftware „Obidos“, die alle Listen, Artikel und Kundendaten verwaltete. Die bei wachsendem Geschäftsumfang fortlaufend erfolgte Erweiterung von Datenbank- und Hardwareressourcen im Backend konnte aber bald schon die zunehmenden Probleme bei der „frond-end application“ nicht mehr lösen.¹¹²

Die Firmenleitung sah sich daher um 2001 zu einem grundlegenden Systemwechsel der gesamten IT-Infrastruktur veranlasst. So wurden die bisherigen „monolithic systems“ in eine zentral organisierte, von allen Abteilungen gemeinsam genutzte Amazon Plattform überführt, die hardwaremäßig aus einem virtualisierten Verbund von Datacenterclustern und softwaremäßig aus Hunderten von modularen „primitiv services“ bestanden. Dieses nach *Prinzipien der Serviceorientierten Architektur* entwickelte Baukastensystem ermöglichte die Wiederverwendung, Neukombination und Orchestrierung immer neuer Dienstmodule bzw. Dienste und erlaubte durch Verbergung von Software-Komplexität in einfachsten Bedienschnittstellen auch die Ausschöpfung von Kunden-

¹¹¹ Economist 5.10.2006; z.T. auch Carr 2008a, S. 74; dagegen Bezos in Roush 2006; Selipsky 2010, S. 22 („There's this myth that we created AWS to sell underutilized capacity at Amazon.com, but this had absolutely nothing to do with why we built AWS.“)

¹¹² Gray, Vogels 2006, S. 14 f.

arbeit für die Amazon-Geschäftsprozesse.¹¹³ Die wegen der unterschiedlichen Nutzergruppen erforderliche hohe Flexibilität erreichte man über eine durchgängige Virtualisierung, die das entscheidende Charakteristikum der Amazon-Infrastruktur bildete.

Als Amazon ab 2002/03 die eigene Web-Handelsplattform für andere Online-Händler öffnete, wurde aus der bis dahin nur unternehmensinternen ‚IT utility‘ ein vielen „IT professionals and developers“ zugänglicher Infrastruktur- und Plattform-Dienst: Der „Infrastructure as a Service“ (IaaS) virtualisierte Amazons eigene Cloud-Infrastruktur für die externen Kunden und der „Platform as a Service“ stellte dafür Betriebssystem sowie Virtualisierungs- und Entwicklungs-Tools (APIs) bereit, die die Mitnutzung von Amazons Rechenzentren und Dienstpalette für Entwicklung und Betreiben von Web-Diensten und -Anwendungen ermöglichten. Der große, in diesem Ausmaß völlig unerwartete Erfolg der Amazon-Services motivierte die leitenden AWS-Entwickler *Chris Pinkham* und *Benjamin Black* zu weitergesteckten Geschäftsmodellen und ersten Leitbildern. Ihre positiven Erfahrungen mit der Virtualisierung von Hardware- und Software-Ressourcen verallgemeinerten sie in der „virtual server as a service idea“ und der noch vagen Vision eines „*universe of virtuals*“.¹¹⁴ Aber trotz der Vorwegnahme des „everything-as-a-service“-Konzeptes verzichteten sie noch immer auf einen paradigmatischen Leitbegriff: „Again, there was no term ‚cloud computing‘. Really, when we started this, it was called Amazon Web Services“.¹¹⁵

Ungeachtet der erkennbaren Ausweitung des Entwicklungshorizontes erfolgte der Aufbau der ‚Cloud‘ weiterhin in Form eines schrittweisen Ausbaus von „Amazon’s internal infrastructure“. Hatten zunächst Software- und Platform-Services im Vordergrund gestanden, wie sie auch andere IT-Konzerne anboten, schuf Chris Pinkham 2004-2006 einen Speicher- und einen Computingdienst: den „*Simple Storage Service*“ (S3), der Unternehmen und Entwicklern den Zugriff auf ‚unbegrenzte‘ Speicherkapazitäten eröffnete.

¹¹³ Gray, Vogels 2006, S. 16 („For us service orientation means encapsulating the data with the business logic that operates on the data, with the only access through a published service interface.)

¹¹⁴ Der Amazon CTO Werner Vogels formulierte die große Bedeutung der Virtualisierung für Amazon folgendermaßen: „Virtualization is a critical factor in the success of IaaS. Virtualization in this regard is not simply the virtualization of compute resources. It is the virtualization of compute, storage, network, and I/O. We must be able to manage each of these resources discretely in order to meet SLAs and economies of scale needed by IaaS.“ (Vogels 2009).

¹¹⁵ Selipsky 2010, S. 24 (meine Hervorhebung); Black 2009.

nete, und bis 2008 im Testmodus, die „*Elastic Computing Cloud*“ (EC2), die beliebig skalierbare Computing-Leistungen zu Discountpreisen anbot.¹¹⁶

Es stellte sich schon sehr bald heraus, dass es gerade die Kombination von Software-, Platform- und Hardware-Services und „Human Resources“ zu einer umfassenden, selektiv nutzbaren, stundenweise abzurechnenden Dienstpalette war, die zu einer „Verkettung kritischer Massen“ und dadurch zu einem in diesem Ausmaß nicht erwarteten Take-off des Amazon-Cloudkonzeptes führte.¹¹⁷ Denn allein über Browser und Web APIs erhielten Firmenkunden Zugang zu der mit 2 Mrd. \$-Aufwand errichteten Amazon-Infrastruktur, d.h. dem Firmencluster von „multiple data centers“, der Speicher-Cloud und den Softwarediensten einschließlich einem „vast data warehouse“ für die Datenanalytik. Indem Unternehmen Computing und Datenspeicherung in die Cloud verlegten, könnten sie, so Bezos' stereotyp wiederholtes Argument, die 70% fixen Kosten einsparen, die im E-Commerce für die Computing- und Web-Infrastruktur aufgewendet werden müssten. Der größte Kostenfaktor entfiel dabei nicht einmal auf Server, Personal und Energie, sondern auf die chronische Unterauslastung der eingesetzten Ressourcen.¹¹⁸

Obwohl Bezos jetzt häufig von „storage in the cloud“ und „compute capacity in the cloud“ sprach, ja der Rechnerservice „Cloud“ sogar im Namen führte, war während Jahre 2006-2009 in der Firmenleitung nicht „Cloud Computing“, sondern „Utility“ der mit Abstand am häufigsten verwendete Leitbegriff. Und unter ihm verstand man das kommende Computing-Paradigma, in dem die AWS mit gutem Recht eine, wenn nicht gar die Führungsstellung beanspruchen konnten, denn: „Amazon is a natural in the emerging utility computing world.“¹¹⁹ Der Stilisierung der „Utility“ zum epochalen Paradigma entsprach auch der immer häufigere Rekurs von Bezos, Jassy und Selipsky auf die *Computing-Electricity-Analogy*, wobei sie sich nun eng an das historische Narrativ von Nicholas G. Carr und dessen „IT utility model“ anlehnten. Die Computing-Electricity-Metapher entwickelte sich sehr bald von einer vagen Analogie, mit der ein aktueller Prozess in seiner welthistorischen Dimension begriffen werden sollte, zu einer auto-

¹¹⁶ Bezos, O'Reilly 2006; LaMonica 2006a; Selipsky 2010; Black 2009; siehe auch Brandt 2012, S. 167 ff.

¹¹⁷ Bezos, O'Reilly 2006 („We're a little bit surprised by how excited people seem to be by these services“); Clark 2012 (meine Hervorhebung); Pattay 1993, S. 170ff.

¹¹⁸ Bezos 2006; Bezos, O'Reilly 2006.

¹¹⁹ Bezos O'Reilly 2006; LaMonica 2006b (meine Hervorhebung).

suggestiven „self fulfilling prophecy“ und Inspirationsquelle für allerlei Zukunfts-szenarien und Verheißungen.

Aus dem Infrastrukturengpass IT-Versorgung waren so im Laufe weniger Jahre ein eigener Geschäftsbereich und schließlich ein weiteres Amazon-Weltprojekt hervorgegangen. Auch dieses verfolgt wie bereits die früheren, der ‚Welt-Online-Buchladen‘ mit dem Fernziel Universalbibliothek¹²⁰, der ‚Welt-Mediendiscounter‘ und das ‚Weltwarenhaus‘, monopolistische Ansprüche. Langfristig möchte Amazon nun die Versorgung der Welt mit Content und die globale Cloud Infrastruktur in einem Geschäftsmodell vereinigen und damit über eine universelle Virtualisierung von Hardware, Software und Content im eigenen Cloud-System die soziale Architektur des Web grundlegend verändern. In einem künftigen „fully integrated media service“ wird der Content nämlich nicht mehr als Produkt über das Internet dem User verkauft und damit übereignet, er wird vielmehr nur noch als Lizenz zum Online-Gebrauch vergeben und über Streaming-Dienste aus der Cloud bezogen. Die gesamte Wertschöpfungskette von den Contentproduzenten bis zu den Kunden soll in ihren unterschiedlichen ‚Aggregatzuständen‘ aus der Amazon-Cloud heraus gesteuert werden.

Mit der Rückverlagerung der gesamten Datenverarbeitung und –speicherung von den dezentralen Endsystemen in Amazons Multiple Datacenters geraten die Benutzer allerdings in ein mehrfaches Abhängigkeitsverhältnis. Denn Amazon bestimmt nun wesentlich die technischen, ökonomischen und vor allem rechtlichen Bedingungen des Mediengebrauchs. So kann die Ausleihe unterbunden oder stark eingeschränkt und sogar über die Software die Lebensdauer und Nutzungsrechte von Geräten und Content zeitlich begrenzt werden. Zudem muss der User die detaillierte Kontrolle seiner Mediennutzung zulassen, ohne seinerseits Einblicke in die Weiterverwendung seiner vom Provider gespeicherten Nutzerdaten zu erhalten. Dem Cloud-Anbieter erwächst so ein Kontrollpotenzial, für das es in der traditionellen Elektrizitätsversorgung keine Entsprechung gibt. Der Kunde hingegen verliert durch den Verzicht auf eigene Speichermedien und IT-Produktionsmittel seine Souveränität, er tritt in ein quasi-feudales Verhältnis zum Cloud-Service-Provider und wird Bestandteil eines geschlossenen Bereiches im Internet. Das Resultat von Amazons Aufbau eines vertikal integrierten Cloud-Computing-Stack und einer flexibel kombinierbaren Cloud-Dienste-Palette war so eine recht erfolgreiche

¹²⁰ Zur „Alexandrian fantasy“ der „Great Library of Amazonia“ siehe Wolf 2003.

„Enclosure-Strategie“ im Web, der bald auch andere „cloud-centric“ Internetprovider folgten.¹²¹ Dies zeigt sich besonders deutlich bei dem zweiten Megaplayer des Cloud Computing, der Firma Google, obwohl die Cloud-Genese hier in vielen Punkten von der Amazons abwich.

Während sich Amazon bei der Konstituierung der AWS von den Hardware-nahen, unteren zu den höheren Schichten vorarbeitete, entwickelte Google seinen Cloudstack genau umgekehrt von den Anwendungs- und Software-Diensten abwärts zu den Infrastruktur- und Hardware-Services. Google peilte zunächst auch nicht Geschäftskunden als Zielmarkt an sondern begann als ein auf Privatnutzer ausgerichteter Mainstream-Cloudanbieter und zog, getrieben von anderen Cloud-Providern, erst nachträglich mit Plattform- und Infrastrukturdiensten für den Business-Bereich nach. Generell werden bei Google neue Cloud-Dienste zunächst als vorläufige Versionen für den Privatbereich eingeführt und erst nach der Testphase in erweiterter Form Firmenkunden angeboten.¹²² Die Google-Cloud-Genese nahm ihren Ausgang in einer völlig anderen professionellen Kultur, denn hier ging es nicht wie bei Amazon um eine Rechnerinfrastruktur für den datengesteuerten Buch- bzw. Medienvertrieb und die Warenversorgung über das Web, sondern um die allgemeine Informationssammlung und -erschließung mithilfe effizienter Algorithmensysteme. Die beiden Google-Gründer *Sergey Brin* und *Larry Page*, zwei Söhne von Informatik- bzw. KI-Professoren, entwickelten 1995-98 als Stanford-Postgraduates in einem Forschungsprojekt des *Digital Libraries Initiative*-Programms von NSF, DARPA und NASA ein „new model for information retrieval“ für das „searching and organizing all the world's information“.¹²³ Ihre Forschungsintention, „to crawl the whole web“ und „to collect a database of links“, erweckte in ihnen schon bald einen informationellen Universalanspruch und eine noch sehr vage Geschäftsidee, die privatwirtschaftliche Aneignung des Webinhalts in einem zentralen Rechner- und Speichersystem, die dann später auch der Anlass für das Google Cloud-Computing-Konzept werden sollte.

Die Anforderungen einer Web-Suchmaschine führten dann nahezu zwangsläufig zu einer Strategie der „massive centralization“ und des „massive parallelism“, denn es galt

¹²¹ Zur „feudalization of public space and the art of enclosure“ siehe Meinrath, Losey, Pickard 2011, S. 432 f, 427-431.

¹²² Brandt, Schürmann 2009, S. 6.

¹²³ Brin, Page 2004, S. B 11.

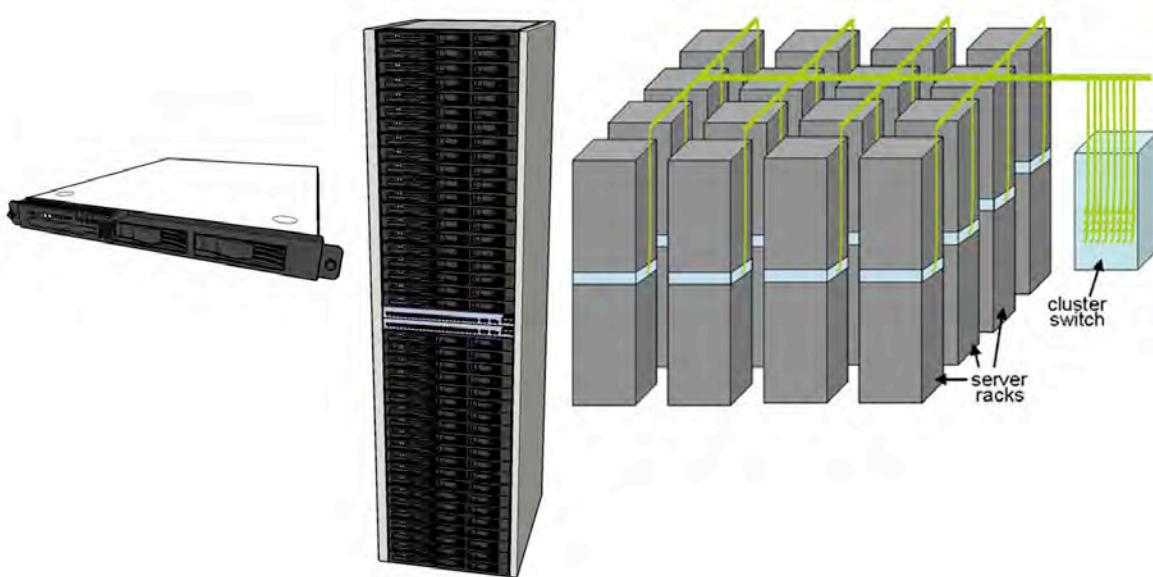
für die zahllosen Zugriffe auf die Suchmaschine ein geeignetes Rechnersystem zu schaffen, das den gleichzeitigen Anfragen einer rasch zunehmenden Nutzerzahl gewachsen war. Im Jahre 2003 begann der Bau eigener Rechenzentren, für den Hölzle zusammen mit dem Spezialisten für Multicore-Prozessoren *Luiz Barroso* und dem Softwareingenieur *Jeffrey Dean* die zellulare „*Google Cluster Architecture*“ entwickelte. Diese ersetzte das bis dahin übliche Parallel-Supercomputer-Design mit wenigen „high-end servers“ durch Servercluster, die jeweils aus Pools von bis zu 15.000 massiv-parallelen, in Racks von jeweils 40-80 gebündelten „commodity-class PC's“ bestanden, die allesamt zu einer einheitlichen „computing Resource“ integriert wurden.¹²⁴ Die massiv-parallele Server-Architektur wurde softwareseitig ergänzt durch Googles Standard File System mit „shared data models“ und das neue, von Dean und Sanjay Ghemawat entwickelte Programmiermodell „MapReduce“ für nebenläufige Berechnungen. Dieses koordinierte die Server bei der automatischen Verteilung der Workload und organisierte die Aufteilung der „query streams“ in „multiple streams“ und ermöglichte dadurch die arbeitsteilige Abarbeitung aller Anfragen durch eine Vielzahl von Prozessoren und Servern bei geringer Latenz und maximalem Durchsatz. Google schuf so mit dem durchgängigen Parallelismus die Infrastruktur für die Dekomposition, Sortierung, Verarbeitung und Speicherung der globalen Information und die Bewältigung von Big Data-Analysen. Die parallele Ausführung von Jobs auf mehreren Servern förderte darüber hinaus die Fehlererkennung und -behebung und löste mit dem so erzielten höheren Redundanzniveau zugleich das eminente Vulnerabilitätsproblem der angestrebten Zentralisierung des Computing.¹²⁵

Mit der „*Google Cluster Architecture*“ und der Errichtung eines Netzes von einem Dutzend Multiple Datacenters war 2000-2004 der Grundstock von Googles Cloud-Infrastruktur geschaffen worden, die bis zu diesem Zeitpunkt im Wesentlichen dem Eigenbedarf der Firma für die Erfüllung des Kerngeschäftes diente, d.h. der Errichtung einer globalen Information Retrieval Utility, der Durchführung eines werbefinanzierten Suchmaschinenbetriebs sowie dem Data Mining und Management der Kundendaten. Obwohl es sich damit faktisch schon um ein massiv-paralleles, massiv-zentralisiertes

¹²⁴ Hierzu und zum Folgenden Barroso, Dean, Hölzle 2003, S. 22 f., 25 ff., Zitat auf S. 27.

¹²⁵ Dean, Ghemawat 2004; für „System and method for networking computing clusters“ beantragte Google Anfang 2007 ein Patent, das 2012 erteilt wurde (US 8144697 B2); das für das bald Schule machende MapReduce-Verfahren 2004 beantragte Patent wurde im Januar 2010 erteilt (US 7.650.331); Levy 2011, S. 182 ff.; Di Justo, Patrick 2008.

„cloud model“ handelte, sprach man in dieser Phase weder von Cloud noch von Grid oder Utility, sondern verwendete meist den von Hölzle mit Blick auf die verwendete Cluster-Architektur eingeführten Begriff „*Cluster Computing*“. Der Cloud-Begriff war bei Google ursprünglich überhaupt nicht populär, er taugte nach Hölzle bestenfalls als Markenname, denn „technically speaking, it’s *cluster computing* that you do.“¹²⁶ Dies änderte sich erst, als es 2004-2006 zu einer deutlichen Ausweitung der Anwendungsdienste und Plattformfunktionen und in diesem Kontext zu einer weiteren Skalierung des Google-Datacenter-Designs kam.



Die Google Cluster Architecture als massiv-parallele Grundstruktur der „warehouse-scale systems“:
Links: Servers (CPUs, DRAM, Disks); Mitte: Racks à 40-80 Servers; Rechts: Clusters aus Server Racks¹²⁷

Der Google-Chef erkannte nämlich, dass die Verlagerung der bislang auf PCs laufenden Computing- und Speicherfunktionen in zentrale Serverfarmen noch einer gewaltigen Erweiterung der Google-Hardware-Ressourcen bedurfte und er beauftragte daher Hölzle und Barroso mit der Errichtung eines Mega-Datacenters. So entstand 2005/06 in The Dalles in unmittelbarer Nähe billiger Wasserkraft am Columbia River das weltgrößte Rechner-Cluster, das alle bisherigen innovativen architektonischen Ansätze Googles in der von Hölzle als „*Warehouse Scale Computers*“ (WSC) bezeichneten neuen Klasse von Computersystemen integrierte. Google hatte damit einen auf große Public Clouds

¹²⁶ Hölzle nach Levy 2011, S. 180 (Unter „cluster“ verstand man bei Google „a large number of servers – well into the thousands – usually representing the minimum number of machines needed to serve search results from a query.“)

¹²⁷ Barroso, Dean, Hölzle 2009, S. 6.

zugeschnittenen Typ von Datacenter entwickelt, den sich wegen seiner deutlich verbesserten Skalenökonomie und seiner höheren energetischen Effizienz auch andere führende Cloud-Provider wie IBM, Microsoft und Apple schon nach kurzer Zeit zum Vorbild nahmen. Mit Blick auf die Vorbildfunktion für das Design von Cloud-Infrastrukturen hat man die Spitzengruppe der Google-Hard- und Software-Entwickler gelegentlich sogar als das „Bell Lab“ bzw. „Xerox PARC“ des Cloud-Zeitalters bezeichnet.¹²⁸

Die forcierte Ausrichtung der Geschäftspolitik auf die Google-Cloud kommunizierte die Firma nun auch nach innen und außen mit dem Einschwenken auf den Begriff „Cloud Computing“, mit explizierten Leitbildern und Paradigmen-Konstrukten sowie Entwicklungsszenarien, die den strategischen Schwenk vom dezentralen zum zentralistischen Computing logisch bzw. historisch untermauerten. Schon bald nach Abschluss der Großinvestition in The Dalles schwor Eric Schmidt bei der jährlichen „Search Engine Strategies Conference“ im August 2006 die Google-Entwickler auf das neue technische und soziale Architekturmodell ein, indem er „*cloud computing*“ zum offiziellen Leitbegriff erhob und diesen in den historischen Kontext des Epochentwechsels vom PC-Client-Server- zum Data-Center-Cloud-Konzept stellte. Dabei verhinderte er jeden Bezug zu bereits bestehenden Verwendungen des Begriffs, insbesondere durch Amazon, ja er erweckte sogar den Anschein, als sei der Begriff „*cloud computing*“ eine Google-Schöpfung: „What's interesting [now] is that there is an *emergent new model*, and you all are here because you are part of that new model. I don't think people have really understood how big this opportunity really is. It starts with the premise that the data services and architecture should be on servers. *We call it cloud computing – they should be in a 'cloud' somewhere.* [...]“¹²⁹ Der Google-Chef wählte damit gerade den von den Google-Spitzenentwicklern am wenigsten geschätzten Begriff.

Cloud Computing zielte nicht wie „utility computing“ auf den Charakter einer öffentlichen Infrastruktur, sondern stellte die Internet-Cloud in den Mittelpunkt. Zugleich verbarg er die neue soziale Architektur der Zentralisierung sowie die kommerzielle Aneignung von Content, Meta- und Nutzerdaten in Googles Server-Park hinter einer diffusen Cyberspace-Metapher. So ließ sich auch nach außen der Anschein aufrechterhalten, dass das Netzwerk der Computer ist, obwohl man intern der Auffassung war „The data center

¹²⁸ Siehe u.a. Metz, Cade 2012.

¹²⁹ Eric Schmidt in Sullivan 2006, S. 2 (meine Hervorhebung).

is the computer".¹³⁰ Schmidts Adaptation des Cloud-Begriffs stellte von Beginn an auch ein engen Zusammenhang zwischen dem neuen Computing-Paradigma und dem Google-Geschäftsmodell her. Die User sollten weltweit von jedem beliebigen Endgerät auf die Google-Cloud-Dienste zugreifen können und zwar wie bisher bei den Suchmaschinenfunktionen kostenlos, dafür mussten sie aber ihre Nutzer- und Metadaten dem Cloud-Provider für Data Mining- und Werbezwecke voll überlassen, und zwar ohne jegliche Kontrolle über deren Verwendung. Google strebte sogar die Zusammenführung aller Nutzerdaten aus den verschiedenen genutzten Services, ja sogar die Speicherung von „100% of a user's data“ an, da nur so die für eine zielgenaue Plazierung von Werbung erforderliche „transparent personalization“ gelingen kann.¹³¹ Das werbungsbasierte Geschäftsmodell verstärkte so den Drang der Firma, durch immer mehr Dienste immer mehr Nutzerdaten zu erlangen und machte auf diese Weise die Cloud zum infrastrukturellen Fundament für Googles Big Data-Aspirationen.

Getrieben vom eigenen „*Infrastructural Imperialism*“¹³² und der *Mimicry* an die Portefeuilles der Cloud-Konkurrenten arbeitet Google so darauf hin, sich mit jedem neuen Dienst weitere IT-Bereiche zu erschließen, um so zur Universal-Datenbank und Universal-Webplattform für jedermann zu werden. Durch die datenmäßige Einverleibung vieler Social Web-Initiativen und Apps in die Google-Plattform und die Zusammenführung aller „*Google Smart Services*“ entsteht am Ende eine geschlossene Informationswelt, die über die Erfassung und Speicherung aller Nutzerdaten die User an sich bindet und von diesen schließlich mit dem Internet gleichgesetzt wird. Schmidt stilisierte das zentral organisierte Cloud Computing und die Weltprojekte Googles schließlich zu einer universalhistorischen utilitaristischen Mission, die eine weitere Etappe des Internets anstoße, das „*Web 4.0: web as artificial intelligence complementing human race*“. Indem die globale Information in der Google-Cloud gesammelt und die so entstehende „*collective intelligence*“ jedermann verfügbar gemacht werde, entstehe die Grundlage für die Lösung aller Weltprobleme und das Glück aller. Über die gewaltigen Vorzüge der künftigen Zentralisierung des Computing für die gesellschaftliche Informationsversorgung und Rationalisierung hinaus wurden bei Google zur Legitimierung eines hochzentralisierten Cloud Computing auch noch ökologische Argumente ins Feld geführt. Ja das

¹³⁰ Siehe dazu weiter unten.

¹³¹ Carr 2008a, S. 161.

¹³² Vaidhanathan 2011, S. 107 ff.

„Green Computing“ rückte zeitweise an die vorderste Stelle der Werbung für die Migration von PC und dezentralen Rechenzentren in die Cloud, während die besonders von Amazon favorisierte Electricity-Metapher bei Google so gut wie keine Rolle spielte.¹³³

So wie beim Green Computing wurde auch der allgemeine Cloud Computing-Diskurs keinesfalls von den Unternehmen geprägt, die dem neuen Computing-Paradigma zum entscheidenden Durchbruch verholfen hatten. Denn sowohl Google wie auch Amazon hatten sich erst nach vollzogener Innovation und beginnender Diffusion in die Debatte eingeschaltet und übernahmen demgemäß in ihren Legitimationen für eine radikale Zentralisierung des Computing Argumente und Begründungsmuster des schon seit längerem laufenden fachinternen und fachöffentlichen Diskurses über einen notwendigen Paradigmenwechsel. Die Konstituierung der Legitimationsbasis für die rezentralisierte Computing-Architektur ergab sich so aus einer Wechselbeziehung zwischen unternehmensinternen und öffentlichen Leitbilddiskursen. Auch die Durchsetzung von „Cloud Computing“ als Leitbegriff war das Resultat einer komplizierten Wechselwirkung. Hier kam es zu dem erstaunlichen Sachverhalt, dass gerade die beiden Unternehmen, die den Cloud-Begriff lange zugunsten von „Utility Computing“ bzw. „Cluster Computing“ gemieden hatten und ihn erst im Rahmen des eigenen Eintritts in den Cloud Computing-Markt im Jahre 2006 aufgriffen, den Hype um das neue Catchword auslösten. Denn trotz mehrfacher und ab 2004 zunehmender Verwendung der Begriffe Cloud und Cloud Computing in der Vergangenheit, verschafften erst die Einführung der „Elastic Computing Cloud“ (EC2)“ durch Amazon und die öffentliche Inszenierung von Googles „Cloud Computing“-Programm anlässlich des Betriebsbeginns des Mega-Datacenter in The Dalles dem Cloudbegriff eine derart große publizistische Resonanz in der IT-Welt und darüber hinaus, dass die bis dahin die Debatte prägenden Begriffe „Utility Computing“ und „On-Demand Computing“ in der Folgezeit schnell zurückgedrängt wurden.

¹³³ Siehe besonders Barroso 2005; Barroso, Hölzle 2007 und Hölzle, Hanselmann 2009 („es gibt Voraussagen, die vorhersagen, dass der Gesamtkonsum von IT weltweit von 2007 bis 2020 um fast 80, 90 Prozent steigen wird. Ich bin fest überzeugt, dass wenn wir die Technologie ausnutzen, die es schon gibt, dass wir statt einer 80-Prozent-Steigerung eine Verminderung erreichen können.“)



Cloud Computing als Heilsbotschaft¹³⁴

9. Legitimationsstrategien für das Cloud Computing: Die Inszenierung einer Epochenwende

Der von mächtigen Akteuren forcierte Übergang vom verteilten PC-basierten Computing zu einer sozialen Architektur, die sich auf weltweit agierende vernetzte Datacenter-Konglomerate und hochzentralisierte Service-Plattform-Betreiber stützt, erzeugte einen besonderen Legitimationsbedarf. Während für die schleichende Verlagerung von dezentralen Datenbeständen und Rechenprozessen zu zentralen Diensteanbietern keine weitergehende Rechtfertigung erforderlich war, da hierbei die Cloud-Mystifikation die soziale Rezentralisierung hinter globalen Datacenter-Networks verbarg, musste die IT-Welt und die IT-bezogene Technologiepolitik von der Notwendigkeit einer Rückkehr zu einer „newly recentralized computing architecture“¹³⁵ überzeugt werden. Hierzu reichten die Werbebotschaften der Leitbilder und –Szenarien der Cloud-entwickelnden Unternehmen und Forschungsinstitute nicht aus, für einen radikalen Paradigmenwechsel bedurfte es einer breiteren Diskursallianz über die enge Fachöffentlichkeit hinaus.

¹³⁴ Abb. in der Online-Fassung von Davis 2012.

¹³⁵ Gilder 2006.

Die öffentliche Debatte über radikale Outsourcing-Strategien war 2003/04 durch eine Artikelserie des Chefredakteurs des Harvard Business Review Nicholas Carr angestoßen worden. Carr brachte die Utility-Grid-Pläne und von IBM, Sun, HP und Microsoft auf die provozierende These, dass die firmeneigene Datenverarbeitung für das eigentliche Unternehmensziel irrelevant sei und daher sehr bald auf externe Dienstleister übergehen werde. Als Argumentationsfigur für das bevorstehende Ende der Unternehmens-Rechenzentren diente ihm dabei auch die bereits von Wladawsky-Berger und Gentzsch vertretene historische Analogie, dass nun auch die IT in das Zeitalter der Industrialisierung komme, wobei er die ebenfalls schon von ihnen propagierte Electricity-Computing-Metapher nun ganz in das Zentrum stellte. In der 2004 in Buchform erschienenen Darlegung seiner Thesen entwickelte Carr erstmals ausführlich die „power plant analogy“ als Modell für die historisch notwendige Transformation des proprietären „corporate computing“ zu einer „shared and standardized infrastructure“.¹³⁶ Das Buch wurde zwar schon viel besprochen und kommentiert, doch eine wirklich große Resonanz erzielte Carr erst mit einem weiteren, seine Thesen nochmals zuspitzenden Artikel vom April 2005 im MIT Sloan Management Review über das Ende des Corporate Computing.¹³⁷ Angesichts der im selben Jahr bekannt gewordenen Projekte bzw. Pläne von Google, Amazon, Microsoft, HP und Sun für den Bau großer Datacenter brachten nur wenige Wochen nach Carrs Artikel die IT-Management-Beratungen McKinsey und Gartner erste Trendreports über die bevorstehende „Industrial Revolution“ in der IT-Welt.¹³⁸

Besonderes Aufsehen erregte auch eine Reihe von Vorträgen und Artikel des renommierten IT-Beraters und Rationalisierungsspezialisten großer US-Unternehmen und vor allem des Department of Defense *Paul Strassmann* an der Jahreswende 2005/06. Aufgrund von genaueren Einblicken in Googles Multicloud, das er schon Anfang Dezember 2005 als „world's largest computer“ rühmte, stand für Strassmann fest, dass Google mit seiner massiv-parallelen Cluster-Architektur und seiner standardisierten, containerisierten Software-Architektur „the Model for the Systems Architecture of the Future“ geschaffen habe. Google werde, so seine Prognose auf der Basis eines Dreiphasenmodells der Computing-Entwicklung, der herausragende „harbinger of change“, der ähnlich wie IBM die „Data-centric Era“ von 1950-80 und Microsoft und Intel die „Workgroup-

¹³⁶ Carr 2005, S. 68; Carr 2003 S. 7 f., 11 f.; Carr 2004, Kap 2; Carr 2008a S. 22 ff., 57 ff.

¹³⁷ Carr 2005.

¹³⁸ Kanakamedala, Krishnakantan, Roberts 2006; Gartner Press Release 25.4.2005 (<http://www.gartner.com/newsroom/id/492195>).

centric Era“ ab 1980 die nun folgende „Network-centric Era“ ab 2010 bestimmen werde. Mit seinem Hinweis auf die Kontroll- und Resilienz-Potentiale von Utility-Systemen gerade im Kontext der Terroristengefahren stellte der äußerst einflussreiche Strassmann einen direkten Zusammenhang her zwischen dem Cloud Computing-Diskurs und der „Surveillance, Control and Security-Debate“ in der US Computer Science nach dem 11. September.¹³⁹

An Strassmanns historische Szenarien und Argumentationsmuster knüpfte Ende 2006 auch der neokonservative Trendanalytiker, Redenschreiber und Technologieberater mehrerer republikanischer US-Präsidenten *George Gilder* mit einem Aufsehen erregenden Artikel im ‚Zentralorgan‘ der Internet-Communities „Wired“ an, der den bislang vorwiegend internen Fachdiskurs der Computer Community zu einem öffentlichen technologiepolitischen Diskurs über das Cloud Computing erweiterte. Er stellte schon mit dem Titel das Warehouse-scale Datacenter in den historischen Kontext des Übergangs zu einer neuen Industrialisierungsepoke, der Industrialisierung der Information: „*The Information Factories. The desktop is dead. Welcome to the Internet cloud, where massive facilities across the globe will store all the data you'll ever use.*“¹⁴⁰

In den Jahren 2007/08 erlangte Cloud Computing in der IT-Welt seinen Durchbruch als „The Next Big Thing“. Auch der Cloud-Begriff begann sich nun gegenüber den nachwievor konkurrierenden Begriffen „Utility Computing“, „On-Demand Comuting“ und „Grid Computing“ durchzusetzen. Er wurde jetzt auch bereits als ein etablierter „generischer Begriff“ gewertet, so dass der Versuch des PC-Herstellers Dell im März 2007, sich „Cloud Computing“ für sein Service-Palette Design und Anpassung von Hardware, Networks und Services „for data centers and mega-scale computing environments“ als Warenzeichen zu sichern, ein Jahr später vom US Patent and Trademark Office abgelehnt wurde.¹⁴¹ In der Computer Science tauchte Cloud Computing als Konferenzthema an der Jahreswende 2007/08 auf, der erste Artikel zu diesem Thema erschien in den „Communications of the ACM“ im Juli 2008. Brian Hayes betonte darin aber noch, dass neben dem Cloud-Begriff nachwievor auch noch „on-demand computing, software as a service,

¹³⁹ Strassmann 2005 a/b/c; Strassmann 2006.

¹⁴⁰ Gilder 2006; im April 2006 hatten schon Markoff, Hansell in der New York Times das Geheimnis um das The Dalles-Datacenter gelüftet und es als „information-age factory“ bezeichnet.

¹⁴¹ Siehe die Application vom 23.3.2007 im US- Trademark/Service Mark Application, Principal Register (<http://tsdr.uspto.gov/documentviewer?caseId=sn77139082&docId=APP20070327102751#docIndex=1&page=1>); Hoover 2008; Savvas 2008.

or the Internet as platform“ üblich seien. Deren Gemeinsamkeit sah er in der Verlagerung des Computing vom lokalen PC in die „compute cloud“, womit das „new model of remote computing seems to reverse the 1980s *'liberation' movement* that gave individual users custody over programs and data.“ Wegen des möglichen Kontrollverlustes bedürften die Gefahren für „privacy, security, and reliability“ besonderer Beachtung.¹⁴²

Der Schwerpunkt in den Debatten über das neue Paradigma lag in dieser Zeit ganz eindeutig in einer Rezentralisierung der künftigen sozialen Architektur des Computing. Gerade in der Take-off-Phase des Cloud Computing erlebten Hochzentralismus-Visionen im Computing einen Höhepunkt. Beflügelt von informatischen Allmachtsphantasien, sahen einige IT-Strategen die Zukunft gar wieder ähnlich wie einst in den Anfängen der Mainframe-Ära in nur fünf Super Computer-Cloud-Konglomeraten weltweit bzw. wie im IBM-Kittyhawk-Programm gar in einem einzigen massiv-parallelen Exascale-Serversystem, das die gesamte Information und Kommunikation des Internet speichern, verarbeiten und kontrollieren könnte.¹⁴³ Es gab zwar weiterhin eine ganze Reihe von Peer-to-Peer-Konzepten und konföderativen Grid-Ansätzen, doch die Diskussion wurde nun von den zentral organisierten „Megascale Clouds“ beherrscht. Deren Diskursdominanz innerhalb und zunehmend auch außerhalb der IT-Community wurde Anfang 2008 mit dem Erscheinen des schnell zum Bestseller aufgestiegenen Buches „The Big Switch“ von Nicholas Carr noch einmal deutlich gestärkt.

Carrs eloquentes Plädoyer für eine „epochal transformation“¹⁴⁴ bündelte die bisherige Cloud-Computing-Debatte in einem geschlossenen Entwicklungsszenario, das die verschiedenen historischen Analogien, Narrative und Entwicklungs'gesetze' zu einer stimmigen Legitimationsbasis für das neue Computing-Paradigma integrieren sollte. Obwohl das Buch überwiegend als technik- bzw. wirtschaftshistorische und technologiepolitische Darstellung angelegt war, kaum auf die Cloud-Technologie selber einging, sondern vorwiegend auf der Ebene von historischen Entwicklungslogiken und sozialen Architekturen argumentierte, wurde es zum Leitszenario des Cloud Computing und zum meist zitierten Titel in der Cloud-Computing-Literatur. Carrs „Big Switch“-Szenario lenkte die Aufmerksamkeit von Computer Science, IT-Welt und Technologiepolitik einseitig auf

¹⁴² Hayes 2008 (meine Hervorhebung).

¹⁴³ Siehe die Szenarien von Sun (Papadopoulos 2006) und von IBM (Appavoo u.a. 2008 und 2010).

¹⁴⁴ Carr 2008a, S. 12.

eine radikal zentralisierte Computing-Infrastruktur nach dem Vorbild der elektrischen Großkraftversorgung. Durch seine spektakuläre Inszenierung des Paradigmenwechsels als neokonservative Revolution und einschneidende Epochenwende erhielten historische Analogien und Narrative von Beginn an einen zentralen Platz in den Legitimationsstrategien für das neue Computing-Leitbild.

So folgen die meisten der ab 2009/10 zum Thema Cloud Computing erscheinenden Grundsatzartikel und Fachbücher dem von den Cloud-Pionieren und Cloud-Promotoren erzeugten „Sound of Inevitability“.¹⁴⁵ Für einen radikalen Rezentralierungs-Kurs werden zwar mannigfache wirtschaftliche, EDV-praktische und Bequemlichkeits-Gründe ins Feld geführt, doch werden diese von szenarienlastigen Argumentationen überfrachtet, die der neuen Option einer zentralistischen Architektur des Computing apriori einen entwicklungslogischen Charakter verleihen. An vorderster Stelle der ökonomischen Vorteile für gewerbliche, intitutionelle und private Nutzer steht dabei die Einsparung von kostspieligen Investitionen für Hardware-, Software-Beschaffung und -Pflege sowie für die Transaktionskosten des Aufbaus einer kommerziellen Web-Präsenz. Der Zugriff auf beliebig skalierbare Rechen- und Speicherkapazitäten als On-Demand-Infrastruktur und standardisierte integrierte Software als On-Demand-Software führe durch Einsparung lokaler Ressourcen zu einer neuen Dimension der Kosteneffizienz. Der drastische Rückgang des Hard- und Software-Preisniveaus als Folge skalenökonomischer Zentralisierung und nutzungsabhängiger Abrechnung ist denn auch eine der häufigsten Werbebotschaften der Cloudanbieter. Dazu beseitige Cloud Computing für alle User das bislang unlösbare Problem der dynamischen Ressourcenallokation und schaffe so unbegrenzte, beliebig anpassbare Computerkapazitäten.

Doch noch vor unbegrenzter Ressourcenverfügbarkeit, drastischen Kosteneinsparungen und ortsunabhängigem Zugriff bilden Bedienungsvereinfachung und Bequemlichkeit das Hauptargument für die Auslagerung von Rechen- und Speicherkapazitäten in die Cloud. Wie man bei der Hardware generelle Plug-and-Play-Qualität verspricht, so sollen mit dem Cloud Computing Software-Probleme der Vergangenheit angehören. Die User sollen ihre Verantwortung für die ohnehin nur schlecht bewältigte Organisation der Speicherung und Softwarepflege aufgeben und sich in die Obhut der Cloud-Service-Plattformen begeben. Diese gewähren ihnen mit ihrer standardisierten, modularisierten und beliebig

¹⁴⁵ Zum „Sound of Inevitability“ siehe das Eingangskapitel in Marks, Lozano 2010.

kompositionsfähigen Dienstespalette einen „transparenten Zugriff“ auf alle Anwendungen, ohne dass sie sich um die Auswahl, Installation und die Funktionsweise der jeweiligen Software kümmern müssten. In der Fachpresse wird das vermeintlich sorgenbefreiende Computing in der Cloud sehr häufig mit Ausmalungen der Wolken-Metaphorik visualisiert.



Visualisierung der Verbergung der DV-Komplexität durch Cloud Computing in einer SAP-Werbung¹⁴⁶

Das Standardmotiv der Vereinfachung und Alltagserleichterung für die User geht dabei vielfach einher mit dem Argument einer Reprofessionalisierung des Computing. Technische Infrastruktur, Datenverarbeitung und –speicherung gehörten in die Obhut der dazu berufenen Eliten, Experten und effizientesten Datacenter und Service-Provider. Diese sollen die laienhafte Selfmade-IT der PC-Web-Ära wie auch die völlig unökonomische DV-Eigenversorgung der Betriebe zugunsten einer Fremdversorgung durch hocheffiziente und kompetente „Information Factories“ überwinden. Als Vorbilder für die Reprofessionalisierung des Computing durch Industrialisierung dienen nicht selten die Ablösung der handwerklichen durch die industrielle Massenfabrikation und vor allem der Übergang zu den großen elektrischen Kraftzentralen, die Verbrauchern und Betrieben alle Lasten und Sorgen der Eigenversorgung abgenommen hätten. Derartige Vergleichsprozesse bilden die argumentive Brücke zu historischen Narrativen, mit denen über betriebswirtschaftliche und praktische Opportunitätsgründe hinaus einem

¹⁴⁶ Bild nach SAP Info vom 27.9.2012, Quelle: Fotolia.com (<http://en.sap.info/cloud-from-zero-to-hero/80438>).

fundamentalen Richtungswechsel in der gesellschaftlichen Organisation des Computing mehr Nachdruck verliehen werden soll.

Die bei nahezu allen neuen Technikzweigen zu beobachtende Tendenz, der neuen Technik einen historischen Ort in der Technikentwicklung zuzuweisen und sie auf diese Weise in ein legitimierendes historisches Narrativ einzubetten, ist beim Cloud Computing besonders ausgeprägt. Fast keine Monographie und kaum ein Überblicksartikel verzichtet daher auf einen historischen Eingangsabschnitt, durch den die Rezentralisierung als logische Konsequenz bisheriger Entwicklungsstufen plausibel gemacht werden soll. Die Ausgiebigkeit so mancher Geschichtskonstruktion deutet dabei auf einen erhöhten Legitimationsbedarf des propagierten Paradigmen- und Strategiewechsels hin. So schließt sich bei vielen historischen Abrissen an eine kurze meist stereotype Begriffsgeschichte die Aufzählung einer Ahnenreihe an, die bis zu McCarthy und den Time-Sharing-Utilities, gelegentlich sogar bis zu Herbert Grosch und zur frühen Mainframe-Ära zurückgeht. Explizit oder implizit lautet die Botschaft dabei, dass das Konzept zentraler DV- bzw. Informationsversorgung schon sehr alt ist, aber bislang infolge unzureichender hardware- und softwaretechnischer Voraussetzungen nicht realisiert werden konnte. Derartige Pionier-Stammbäume konstituieren zwar eine historische Kontinuität, doch liefern sie noch keine tragende Erklärung für den doppelten Wechsel der sozialen Architektur im Computing vom Zentralismus zu verteilten und zurück zu zentralistischen Strukturen, hierzu bedarf es stringenterer Formen historischer Begründung und Rechtfertigung.

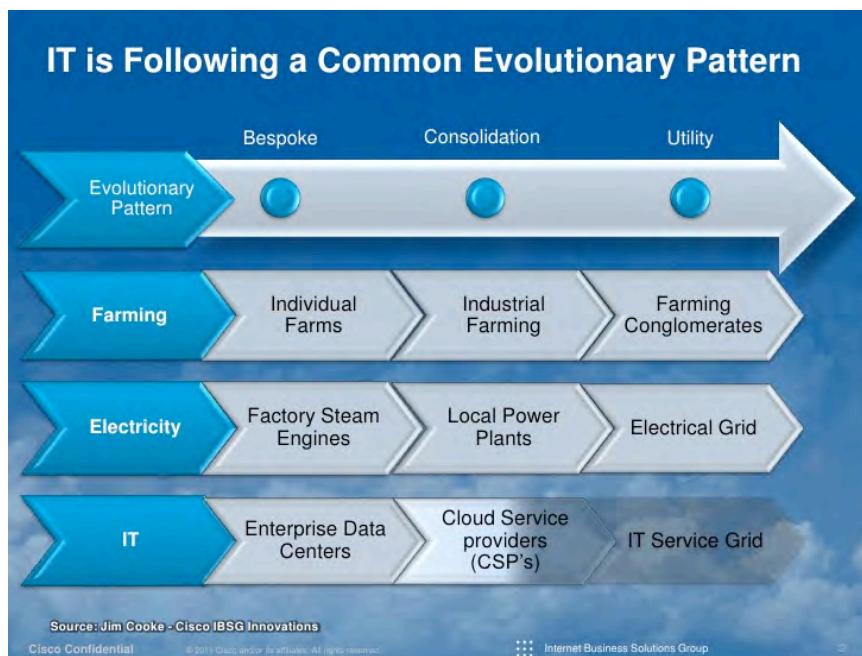
Einen recht verbreiteten Erklärungsansatz bilden scheinbar gesetzmäßige Veränderungen in den Relationen zwischen dem Leistungs- und Kostenniveau von Computing und Übertragung. Danach wurde die direkte Weiterentwicklung der Time-Sharing-Zentrallennete zur Cloud durch die eclatanten Differenzen in der Entwicklung von Prozessor- und Übertragungsgeschwindigkeit ausgebremst, so dass das Utility-Konzept von Mini-computer und PC verdrängt werden konnte. Nachdem die Übertragungstechnik durch den Aufstieg der Optoelektronik wieder Anschluss an die Leistungsentwicklung der Mikroelektronik gefunden hatte, konnte sich der jahrzehntelang zurückgehaltene Trend von der Mengenskalierung kleindimensionierter dezentraler zur Größenskalierung zentraler Computing-Aggregate wieder durchsetzen. Diese Wechselbeziehungen zwischen den Leistungsparametern bilden bei einer Reihe von Autoren die Grundlage von elabo-

rierten Entwicklungsmodellen. Die Geschichte des Computing erhält in dieser Epochenkonstruktion den Charakter eines zyklischen Prozesses mit drei Phasen:

- Die Phase des frühen, noch auf isolierten Mainframes beruhenden zentralisierten Utility Computing dessen Größensteigerungen dem Grosch's Law folgten,
- Die dezentrale Übergangsphase von Minis, PCs und Handhelds, die sich den Takt der Miniaturisierung und Mengenskalierung vom Moore's Law vorgeben ließen und
- Die kommende Phase des entwickelten hoch zentralisierten Cloud Computing mit vernetzten Serverfarmen als dem "Ubiquitous Data Center", bei dem das Moore's Law nur noch Bauelemente und Zugangsmedien betrifft, während die Zentralressourcen sich wieder nach dem Gesetz der Größenskalierung richten, dem Grosch's Law.

Doch für einen derart tief in etablierte Computing- und Internet-Kulturen eingreifenden Richtungswechsel reichen solche sich auf technisch-ökonomische ‚Gesetzmäßigkeiten‘ stützende Erklärungsansätze auch noch nicht aus. Deshalb bedient sich ein Großteil der Autoren weitaus häufiger historischer Vergleichsprozesse, die den Schritt von der dezentralen Eigenversorgung zur zentralen Fremdversorgung als selbstverständlich erscheinen lassen. So wird das um 1960 mit der Time-Sharing-Technologie aufgekommene Leitbild vom Computing als der „Fifth Utility“ nach Wasser, Gas, Strom und Telefon zum beliebtesten Rechtfertigungsargument. Mit der nun allgegenwärtigen Electricity-Computing-Analogie wird vor allem der Zentralisierungsprozess zu einem universellen Entwicklungsmuster großtechnischer Infrastruktursystem deklariert. Dabei übersieht man freilich, dass sich in der Energieversorgung im letzten Jahrzehnt gerade ein gegenläufiger Wandel der vorherrschenden sozialen Architektur ereignet hat, so dass es zu einer aufschlussreichen *Metaphernkreuzung* gekommen ist. Dem dezentralen „Internet der Energien“-Leitbild dient der erfolgreiche Dezentralisierungsprozess in der Computerkommunikation als Vorbild, während sich das zentralistische Cloud Computing-Leitbild auf das Vorbild einer radikalen Zentralisierung der elektrischen Energieversorgung beruft. Im Ringen um die soziale Architektur des künftigen Energiesystems dienen so *historische Transformationsprozesse* wechselseitig als strategische Vorbilder und Rechtfertigungsargumente, so dass es hier zu einem bemerkenswerten metaphorischen Kurzschluss kommt.

Noch weiter ausholende Narrative stellen die Entwicklung zentraler Cloud-Architekturen in einen universalhistorischen Kontext. Danach werden die großen Phasen der Zivilisationsgeschichte von einem Entwicklungsmuster bestimmt, an dessen Anfang jeweils dezentrale Produktionsformen und Selbstversorgung stehen, die dann der Industrialisierung und schließlich zentralisierten Versorgungsinfrastrukturen weichen müssen. Das skalenökonomische Optimum werde in einem alle Dienste integrierenden Universalnetz mit zentralen Hochleistungsaggregaten erreicht.¹⁴⁷ Dieses am aufführlichsten in Nicholas Carrs „Big Switch“-Szenario aufgestellte universalhistorische Erklärungsmuster findet sich in so manchen Cloud Computing-Apologien wieder, jedoch meist stark verkürzt wie in der mehrfach wiedergegebenen Eingangsfolie der Präsentation des Senior Director von Ciscos Internet Business Solutions Group Jim Cooke „Cloud Computing: The New Economics of IT“ von 2011.



Die Rezeption von Carrs „Evolutional Pattern“ in einer Präsentation von Jim Cooke, Senior Director von Ciscos Internet Business Solutions Group (Cooke 2011)

Gemeinsames Bestreben all dieser historischer Analogien ist der Nachweis der Alternativlosigkeit der epochalen Transformation zum Cloud Computing. Mit der Konstruktion von solchen Entwicklungslogiken wird ein Entscheidungzwang konstituiert, der die historische Kontingenz durch ein „Technological Fix“ beschließt und den Informatikern bzw. IT-Entwicklern die Entscheidung über die zu wählenden Pfade unter Beachtung

¹⁴⁷ Carr 2008a, S. 23 ff.

der jeweiligen gesellschaftlichen Implikationen abnimmt. Vor allem aber sollen die User mit derartigen suggestiven Apellen bewogen werden, durch den weitgehenden Verzicht auf eigene IT-Produktionsmittel ihre informationelle Eigenverantwortung an externe Dienstleistungskonzerne abzutreten. Für das Ziel einer radikalen Vereinfachung des Computing, höhere Effizienz und Sicherheit des Computing, so vor allem die Botschaft der Propagandisten der Mega Clouds, müsse eine Verschiebung der Macht- und Wissens-Balance zwischen den beteiligten Netzakteuren in Kauf genommen werden. Denn Vereinfachung und Bequemlichkeit beim Computing implizieren ein asymmetrisches Transparenz-Konzept: Transparenz in der User-Perspektive bedeutet die Verbergung der Benutzungskomplexität vor dem User. Damit die Komplexität aber in der „Wolke“ adäquat bewältigt werden kann, benötigt der Clouddienste-Anbieter seinerseits umfassendes Wissen über den Benutzer und seine Anwendungen. Damit die Verbergung der Komplexität gelingt und er zufriedenstellend umsorgt werden kann, muss er die Erkundung seiner Komplexität zulassen. Transparenz in der Cloudanbieter-Perspektive bedeutet also die Entbergung des Benutzers und seiner Kontexte, doch dies bleibt dem Benutzer verborgen, denn es geschieht ja ihm zum Nutzen und zur Entlastung in der „Wolke“.

Doch nicht nur im Interesse der eigenen Annehmlichkeiten, sondern auch für ein höheres Sicherheits- und Effizienzniveau von Staat, Wirtschaft und Gesellschaft wird ein Verzicht der Bürger auf die Souveränität über ihre eigenen Daten angeraten. Angesichts der aktuellen „crisis of control“, die das für die Weltwirtschaft strategisch so bedeutende „Net“ zu einer „surprisingly insecure infrastructure“ habe werden lassen, erhält deshalb in dem neokonservativem „Big Switch“-Szenario das ordnungsstiftende Technologiebündel von gesellschaftlich zentriertem Cloud Computing, Data Mining und Data Analytics die historische Rolle zugewiesen, die Kontroll- und Governance-Defizite in Computing und IT wieder zu beseitigen. Letztlich läuft das den Durchbruch der Cloud-Technologie legitimierende und propagierende neokonservative „Big Switch“-Szenario darauf hinaus, die Alternativlosigkeit der Transformation des bisherigen ungeordneten World Wide Web zu einem einem einzigen, durchgängig kontrollierten „World Wide Computer“ aufzuzeigen und Einsicht in die Notwendigkeit der Unterordnung aller unter den alles erfassenden, Sicherheit und Wohlstand garantierenden ‚digitalen Leviathan‘ einzufordern: „After all, we also benefit from the personalization that the Internet makes possible—it makes us more perfect consumers and workers. We accept greater control in

return for greater convenience. The spider's web is made to measure, and we're not unhappy inside it.”¹⁴⁸ Unter diesen technisch-gesellschaftlichen Bedingungen stabilisierter Verhältnisse kann sich dann im „World Wide Computer“ die in der Turingmaschine in nuce angelegte Logische Vernunft in einem Prozess fortschreitender Perfektionierung entfalten, um schließlich in einem „*artificial intelligence Leviathan*“ aufzugehen: „As the computing cloud grows, as it becomes ubiquitous, we will feed ever more intelligence into it. [...] What will most surprise us is how dependent we will be on what the Machine knows—about us and about what we want to know. [...] The more we teach this megacomputer, the more it will assume responsibility for our knowing. It will become our memory. Then it will become our identity.¹⁴⁹

Mit dieser universalhistorischen Großvision hatte die sozial zentrierte Cloud Computing-Architekturen legitimierende und propagierende Szenarien-Produktion ihren Höhepunkt erreicht. Ihrer bedienten sich nun auch zunehmend die beiden Mega Cloud-Provider, die beim eigentlichen Entwicklungsprozess der Cloud-Technologie nahezu ohne Leitbilder und Szenarien ausgekommen waren. So setzte wie schon bei der Time-Sharing-Technologie auch beim Cloud Computing erst mit der beginnenden Diffusion die eigentliche Produktion von Entwicklungsszenarien und historischen Narrativen ein und auch hier wurde das neue Pararadigma als Epochenwende inszeniert, doch diesmal nicht als eine Revolutionierung und Befreiung der User Community, sondern als eine neokonservative Konterrevolution. So feiern Brian Chee und Curtis Franklin in ihrer groß angelegten Gesamtdarstellung des neuen Gebietes Cloud Computing als eine Rückkehr zur "Centrally coordinated integration of the mainframe time-share era": „It's amazing how cyclical the IT world is [...] Cloud Computing is a counterrevolution that brings with it the possibility of some very real performance wins for users and their organizations.“¹⁵⁰ Doch die kommenden Jahre werden zeigen, ob der oligopolistische „*techno-utility complex*“ großer Cloud-Provider tatsächlich den gesamten Markt dominieren wird oder ob föderative Grid- bzw. Community-Cloud-Architekturen noch eine Chance erhalten.

¹⁴⁸ Carr 2008a, S. 209.

¹⁴⁹ Carr 2008a, S. 220, 226, 229; „*artificial intelligence Leviathan*“ knüpft an Interviewäußerungen von Sergey Brin über Googles „search engine's teleological development“ an, siehe Kibbe 2008.

¹⁵⁰ Chee, Franklin 2010, S. 2 f.

Literatur

- Aljawarneh, Shadi (2012): Cloud Computing Advancements in Design, Implementation, and Technologies. Hershey, PA
- Anderson, David P.; Fedak, Gilles (2006): The Computational and Storage Potential of Volunteer Computing. In: Cluster Computing and the Grid, 2006. CCGRID 06. Sixth IEEE International Symposium on Cluster Computing and the Grid (CCGRID 06), New York 2006, S. 73–80.
- Anderson, David P. (2010): Volunteer Computing: The Ultimate Cloud. In: ACM Crossroads 16 (2010) 3, S. 7-10.
- Appavoo, Jonathan; Uhlig, Volkmar; Waterland, Amos (2008): Project Kittyhawk: Building a Global-Scale Computer. In: ACM SIGOPS Operating Systems Review 42 (2008): 77-84.
- Appavoo, Jonathan; Uhlig, Volkmar; Stoess, Jan u.a. (2010): Providing a Cloud Network Infrastructure on a Supercomputer. In: Proceedings of the 19th ACM International Symposium on High Performance Distributed Computing, 385-394. New York: ACM, 2010.
- Armbrust, Michael; Fox, Armando; Griffith, Rean (2010): A View of Cloud Computing. In: communications of the ACM 53 (2010) 4, S. 50-58.
- Babaoglu, Ozalp; Jelasity, Mark; Kermarrec, Anne-Marie; Montresor, Alberto; Van Steen, Maarten (2006): Managing clouds: a case for a fresh look at large unreliable dynamic networks. In: ACM SIGOPS Operating Systems Review 40 (2006) 3, S. 9-13.
- Babaoglu, Ozalp; Marzolla, Moreno; Tamburini, Michele (2012): Design and Implementation of a P2P Cloud System. In: SAC'12. Proceedings of the 27th Annual ACM Symposium on Applied Computing. New York 2012, S. 412-417.
- Baker, Mark; Buyya, Rajkumar; Laforenza, Domenico (2002): Grids and Grid technologies for wide-area distributed computing. In: Software—Practice & Experience 32 (2002) 15, S. 1437 – 1466.
- Ballmer, Steve (2000): Steve Ballmer Speech Transcript - Forum 2000 (<http://www.microsoft.com/en-us/news/exec/steve/06-22f2k.aspx>).
- Barroso, Luiz André (2005): The price of performance: An economic case for chip multiprocessing. In: ACM Queue, 3 (2005) 7, S. 49-53.
- Barroso, Luiz André; Dean, Jeffrey; Hölzle, Urs (2003): Web Search for a Planet: The Google Cluster Architecture. In: IEEE Micro 23 (2003) 2, S. 22-28.
- Barroso, Luiz André; Hölzle, Urs (2007): The Case for Energy-Proportional Computing. In: IEEE Computer 40 (2007) 12, S. 33-37.
- Barroso, Luiz André; Hölzle, Urs (2009): The Datacenter as a Computer: An Introduction to the Design of Warehouse- Scale Machines (Synthesis Lectures on Computer Architecture, Lecture 6). San Rafael, Cal 2009.
- Bauer, Walter F. (1958): Computer Design from the Programmer's Viewpoint. In: In: Proceedings of the Eastern Computer Conference, AFIPS 14 (1958), S. 46-51.
- Beckman, Pete; Horowitz, Bradley (2010): The Cloud at Work. Interviews with Pete Beckman of Argonne National Lab and Bradley Horowitz of Google. In: ACM Crossroads 16 (2010) 3, S. 19-22.
- Bemer, Robert W. (1957): How to consider a computer. Data Control Section, Automatic Control Magazine, 1957 Mar, 66-69
- Bemer, Robert W. (undatiert): Origins of Timesharing. Computer History Vignettes. (<http://www.bobbemer.com/TIMESHAR.HTM>).
- Beniger, James (1986): The Control Revolution. Technological and Economic Origins of the Information Society. Cambridge, Mass., London 1986.
- Benioff, Marc; Adler, Carlye (2009): The Untold Story of How Salesforce.com Went from Idea to Billion-Dollar Company And Revolutionized an Industry, San Francisco 2009.
- Bernstein, David; Ludvigson, Erik; Sankar, Krishna u.a. (Cisco Systems) (2009): Blueprint for the Intercloud – Protocols and Formats for Cloud Computing Interoperability. In: Fourth International Conference on Internet and Web Applications and Services, 2009. ICIW '09., S. 328 - 336
- Bezos, Jeff (2006): Opening Keynote and Keynote Interview with Jeff Bezos. Emerging Technologies Conference at MIT 2006, MIT Video 27.9.2006 (<http://video.mit.edu/watch/opening-keynote-and-keynote-interview-with-jeff-bezos-9197/>).
- Bezos, Jeff.; O'Reilly, Tim (2006): Web 2.0 Podcast: A Conversation with Jeff Bezos. Web 2.0 Summit 2006. Transkript von Daniel H. Steinberg. In: O'Reilly 2012.2006 (<http://www.oreillynet.com/pub/a/net/network/2006/12/20/web-20-bezos.html>).
- Birnbaum, Joel S. (1997): Communications Challenges of the Digital Information Utility. In: Hewlett Packard Journal, 48 (1997) 12, S. 6-10.
- Black, Adam (2010): Modular services drive value in the cloud. Interview conducted by Vinod Baya, Bud Mathaisel,, Bo Parker. In: PricewaterhouseCoopers Technology Forecast 2010/4, S. 22-25

- (<http://www.pwc.com/us/en/technology-forecast/2010/issue4/interviews/interview-adam-selipsky-pg1.jhtml>).
- Brandt, Richard L. (2012): Mr. Amazon: Jeff Bezos und der Aufstieg von amazon.com. München 2012.
- Brandt, Gordon; Schürmann, Christian (2009): Analyse und Bewertung Cloud Computing Strategie der Firma Google. Facharbeit an der FOM Hochschule Essen 2009 (http://winfwiki.wi-fom.de/index.php/Analyse_und_Bewertung_Cloud_Computing_Strategie_der_Firma_Google).
- Brin, Sergey; Page, Larry (2004): Playboy Interview: Google Guys. April 2004, In: Playboy, September 2004. Abdruck in Securities and Exchange Comission, Washington, D.C. 20549, Amendment No. 7 to Form S-1 Registration Statement Under The Securities Act of 1933. GOOGLE INC, 13.8.2004. (<http://www.secinfo.com/d14D5a.148c8.htm#1ovvc>).
- Busch, Carsten (1998): Metaphern in der Informatik. Modellbildung, Formalisierung, Anwendung. Wiesbaden 1998.
- Buyyaa, Rajkumar; Chapin, Steve; DiNucci, David (2000): Architectural Models for Resource Management in the Grid. In: Buyyaa, Rajkumar; Baker, Mark (Hrsg.): Grid computing : GRID 2000. First IEEE/ACM international workshop, Bangalore, India, December 17, 2000. Proceedings (Lecture Notes in Computer Science; Vol. 1971), Berlin, Heidelberg, NewYork 2000, S. 18-35.
- Cafaro, Massimo; Aloisio, Giovanni (2011): Grids, Clouds, and Virtualization. In: Dies. (Hrsg.): Grids, Clouds, and Virtualization. London 2011, S. 1-22.
- Carr, Nicholas G. (2003): IT Doesn't Matter. In: Harvard Business Review 81 (2003) 5, S.41-49.
- Carr, Nicholas G. (2004): Does IT Matter? Information Technology and the Corrosion of Competitive Advantage. Boston, MA 2004.
- Carr, Nicholas G. (2005): The End of Corporate Computing. In: MIT Sloan Management Review, Spring 2005, S. 67-73.
- Carr, Nicholas G. (2006): Trailer park computing. In: Rough Type, 17.10. 2006 (http://www.roughtype.com/archives/2006/10/trailerpark_com.php).
- Carr, Nicholas G. (2007): Sun guru foresees data center disaster in '08. In: Rough Type 7.12.2008 (http://www.roughtype.com/archives/2007/12/sun_guru_forese.php).
- Carr, Nicholas G. (2008a): The Big Switch: Rewiring the World, from Edison to Google. New York, London 2012.
- Carr, Nicholas G. (2008b): How many computers does the world need? Fewer than you think. In: The Guardian 21.2.2008 (<http://www.theguardian.com/technology/2008/feb/21/computing.supercomputers>).
- Carr, Nicholas G. (2008c): IT in 2018: From Turing's Machine to the Computing Cloud. In: An Internet.com IT Management eBook, 2008.
- Celesti, Antonio; Tusa, Francesco; Villari, Massimo (2012): Intercloud: The Future of Cloud Computing. Concepts and Advantages. In: Wang, Ranjan, Chen, (Hrsg.) Cloud Computing, S. 167-194.
- Cerf, Vinton G. (2009): Cloud Computing and the Internet. In: Research Blog The latest news from Research at Google, 28.4.2009 (URL: <http://googleresearch.blogspot.no/2009/04/cloud-computing-and-internet.html>).
- Chee, Brian J.S.; Franklin, Jr., Curtis (2010): Cloud Computing: Technologies and Strategies of the Ubiquitous Data Center, Boca Raton, F, London, New York 2010.
- Chellappa, Ramnath K. (1997a): Electronic commerce for digital product companies.
- Chellappa, Ramnath K. (1997b): Intermediaries in Cloud-Computing: A New Computing Paradigm, INFORMS Annual Meeting, Dallas, TX, October 26-29, 1997.
- Chellappa, Ramnath K., Gupta, Alok (2002): Managing computing resources in active intranets. In: International Journal of Network Management 12 (2002) 2, S. 117-128.
- Chetty, Madhu; Buyya, Rajkumar (2002): Weaving Computational Grids: How Analogous Are They with Electrical Grids? In: Computing in Science and Engineering, 4 (2002) 4, S. 61-71.
- Chorafas, Dimitris N. (2010): Cloud Computing Strategies. Boca Raton, FL, London, New York 2010.
- Clark, Jack (2012): Amazon Web Services. A guide to Amazon's Cloud. ZDNet 2012 (Neudruck von vier Beiträgen in ZDNet Juni/Juli 2012).
- Comerford, Richard (1997): The battle for the desktop. In: IEEE spectrum 34 (1997) 5, S. 21-28.
- Cooke, Jim (2010): The Shift to Cloud Computing: Forget the Technology, It's About Economics. Cisco Internet Business Solutions Group (IBSG) Dezember 2010 (http://www.cisco.com/web/about/ac79/docs/pov/Shift_to_Cloud_Computing_POV_IBSG.pdf).
- Cooke, Jim (2011): Cloud Computing: The New Economics of IT (http://www.cisco.com/web/strategy/docs/gov/GSF11_JimCooke_1-1_web.pdf).
- Creasy, Robert (1981): The Origin of the VM/370 Time-sharing System. In: IBM Journal of Research & Development, 25 (1981) 5, S. 483-490

- Dasgupta, Partha; LeBlanc, Jr., Richard J.; Ahamad, Mustaque; Ramachandran, Umakishore (1991): The Clouds Distributed Operating System. In: Computer 24 (1991) 11, S. 34-44.
- David, Edward E.,Jr.; Fano, Robert M. (1965): Some Thoughts About the Social Implications of Accessible Computing. In: AFIPS, Bd. 27/I (FJCC 1965), S. 243-247.
- Davies, Donald W. (1968): The principles of a data communication network for computers and remote peripherals. In: Information Processing 68, Proceedings of IFIP Congress, 2 Bde, Amsterdam 1969, Bd. 2 S. 709-715.
- Davis, Robert (2012): All our Eggs in One Cloud: The International Risk to Private Data and National Security, A Study of United States' data protection law using the International Communications Union Legislative Toolkit. In: Minnesota Journal of International Law, 15.4.2012 (<http://www.minnjil.org/wp-content/uploads/2012/04/Cloud-Computing.jpg>).
- Dean, Jeffrey; Ghemawat, Sanjay (2004): MapReduce: Simplified Data Processing on Large Clusters. In: OSDI'04 Proceedings of the 6th conference on Symposium on Operating Systems Design & Implementation, Berkeley 2004, Bd. 6, S. 10-10, auch in Communications of the ACM 51 (2008) 1, S. 107-113.
- Distefano, Salvatore; Puliafito, Antonio (2012): Cloud@Home: Toward a Volunteer Cloud. In: HPCWire 17.3.2011 (http://www.hpcwire.com/2011/03/17/cloud_home_goals_challenges_and_benefits_of_a_volunteer_cloud/) und in IT Professional 14 (2012) 1, S. 27-31.
- Economist (1998): The future of computing: after the PC. In: The Economist: 348 (12-18 Sept 1998) H. 8085 S. 93-95 (URL: <http://www.economist.com/node/164733>).
- Economist (2006): Lifting the bonnet. The e-commerce giant wants to be more than just a retailer. In: The Economist 5.10.2006 (<http://www.economist.com/node/8001774>).
- Economist (2012): Technology giants at war. Another game of thrones. Google, Apple, Facebook and Amazon are at each other's throats in all sorts of ways. In: Economist, 1.12.2012 (<http://www.economist.com/node/21567361>).
- Eichenwald, Kurt (2012): Microsoft's Lost Decade. In: Vanity Fair, August 2012 (<http://www.vanity-fair.com/business/2012/08/microsoft-lost-mojo-steve-ballmer>).
- Elgin, Ben (2001): The Last Days of Net Mania. In: Bloomberg Businessweek Magazine, 15.4.2001 (<http://www.businessweek.com/stories/2001-04-15/the-last-days-of-net-mania>).
- Eymann, Torsten (2006), Die Infrastrukturen des Grid – Potenziale für eine realistische Beurteilung. In: Wirtschaftsinformatik 48 (2006) 1, S. 68-70.
- Fano, Robert M. (1967): The Computer Utility and the Community, Auszug aus dem Protokoll der IEEE International Convention Record, 1967, in: Annals of the History of Computing 14 (1992) 2, S. 39.
- Fano, Robert M.; Corbató, Fernando J. (1966): Das Time-Sharing-Verfahren. In: Information, Computer und künstliche Intelligenz, Frankfurt a. M. 1966, S. 83-114.
- Farber, Dan (2004): Utility computing: What killed HP's UDC? In: zdnet.com 29.9.2004 (<http://www.zdnet.com/news/utility-computing-what-killed-hps-udc/138727>).
- Ferranti, Marc (1995): News: IT Forum: Ellison, Gates Jockey for Position on Info Highway. IDG News Service, Paris Bureau, 9.4.1995, zit. nach Javaworld 1996/3 (URL: http://www.javaworld.com/java_world/jw-03-1996/idgns.java.1995.028.html).
- Fordahl, M. (2005). Sun Microsystems unveils grid computing initiative. Technology Review 2. 2. 2005 (URL: http://www.technologyreview.com/articles/05/02/ap/ap_2020205.asp?p=1).
- Foster, Jan (2002a): The Grid: a new infrastructure for 21st century science. In: Physics Today 55 (2002) 2, S. 42-47.
- Foster, Jan (2002b): What is the Grid? A Three Point Checklist, 20. Juli 2002. In: Daily News and Information For The Global Grid Community, 1 (2002) 6, S. 8.
- Foster, Jan; Kesselman, Carl (1997): Globus: A Metacomputing Infrastructure Toolkit. In: International Journal of Supercomputer Applications 11 (1997) 2 , S. 115-128
- Foster, Jan; Kesselman, Carl (1999): Computational Grids. In: Dies. (Hrsg.): The Grid: Blueprint for a New Computing, San Francisco, CA 1999, S. 15-51.
- Fox, Geoffrey; Pierce, Marlon (2007): Grids meet too much computing: too much data and never too much simplicity (<http://grids.ucs.indiana.edu/ptliupages/publications/GridSandwich.pdf>).
- Friedewald, Michael (1999): Der Computer als Werkzeug und Medium. Berlin 1999.
- Gates, Bill (2005): Internet Software Services. Memorandum to Executive Staff and Direct Reports; Distinguished Engineers, 30. 10. 2005 (<http://scripting.com/disruption/mail.html>).
- Gates, Bill (2008): Keynote: Microsoft Tech•Ed 2008 – Developers, 3.6.2008 (<http://www.microsoft.com/en-us/news/exec/billg/speeches/2008/06-03teched.aspx>).
- Gentzsch, Wolfgang (2001a): Grid Computing: A New Technology for the Advanced Web. In: Grigoras, Dan u.a. (Hrsg.): IWCC 2001 Advanced Environments, Tools, and Applications for Cluster Computing (Lecture Notes in Computer Science Vol. 2326), Berlin, Heidelberg, New York 2002, S. 1-15.

- Gentzsch, Wolfgang (2001b): Sun Grid Engine: Towards Creating a Compute Power Grid. In: Proceedings of the 1st International Symposium on Cluster Computing and the Grid (CCGRID '01), New York 2001, S. 35-39.
- Gentzsch, Wolfgang (2002): Grid Computing: A New Dimension of the Internet. In: Schubert, Sigrid E.; Reusch, Bernd; Jesse, Norbert Informatik bewegt: Informatik 2002 - 32. Jahrestagung der Gesellschaft für Informatik e.v. (GI), 30. September - 3.Oktober 2002 in Dortmund, GI LNI 20,1 2002, S. 3-9.
- Geralds, John (2000): Andreessen goes public with his plans for Loudcloud. In: v3 co uk, 22 Feb 2000 (<http://www.v3.co.uk/v3-uk/analysis/2003526/andreessen-goes-public-plans-loudcloud>).
- Gilder, George (2006): The Information Factories. In: Wired 14.10, Oktober 2006 (http://archive.wired.com/wired/archive/14.10/cloudware.html?pg=1&topic=cloudware&topic_set=1).
- Giloi, Wolfgang (1968): Das Zauberwort Time-Sharing. In: Computer Praxis 1968, H. 7, S. 130-134.
- Gorst, John P. (2006): John P Gorst. CEO Interview. In: The Wallstreet Transcript, December 11, 2006 (<http://www.twst.com/interview/23736>).
- Gray, Jim; Vogels, Werner (2006): A Conversation with Werner Vogels. In: ACM Queue 30 (2006) 4, S. 14-22.
- Grimshaw, Andrew S. (1992): Metasystems: An Approach Combining Parallel Processing and Heterogeneous Distributed Computing Systems. In: Journal of Parallel and Distributed Computing 21 (1992) 3, S 257-270.
- Grimshaw, Andrew S. (1994): Enterprise-wide Computing. In: Science Bd. 265, 12.8.1994, S. 892-894.
- Grimshaw, Andrew S.; Wulf, William A.; French, James C. (1994): Legion: The next logical step toward a nationwide virtual computer. Technical Report CS-94-21, University of Virginia, 1994/6/8.
- Grimshaw, Andrew S.; Wulf, William A. (1997): The LegionVision of a Worldwide Virtual Computer. In: Communications of the ACM 40 (1997) 1, S. 39-45.
- Grosch, Herbert R. J. (1958): The Future of Computing. In: Industrial & Engineering Chemistry Research 50 (1958) 11, S. 1664-1666.
- Grosch, Herbert R. J. (2003): Computer. Bit Slices from a Life.Novato, Cal, Lancaster, Penn. 1991; 3 Aufl. 2003 (<http://www.columbia.edu/cu/computinghistory/computer.html>).
- Hall, Mark (2010): Why Startups Need Cloud Computing. In: Mark Halls Blog My Two And Half Cents, 4.5.2010 (<http://mytwoandahalfcents.com/what-is-cloud-computing/>).
- Hayes, Brian (2008): Cloud Computing. In: Communications of the ACM 51 (2008) 7, S 9-11.
- Heilmann, Till (2009): Wolkentechnik. Verfügbarkeit und Zerstreuung. Vortrag am Workshop HyperKult 18 „The Cloud“ Lüneburg, 3. Juli 2009 (URL: <http://www.tillheilmann.info/hyperkult18.php>).
- Hellige, Hans Dieter (1992): Militärische Einflüsse auf Leitbilder, Lösungsmuster und Entwicklungsrichtungen der Computerkommunikation. In: Technikgeschichte 59(1992) 4, S. 371-401.
- Hellige, Hans Dieter (1996): Leitbilder im Time-Sharing-Lebenszyklus: Vom "Multi-Access zur "Interactive On-line Community". In: Hellige, Hans Dieter (Hrsg.): Technikleitbilder auf dem Prüfstand. Das Leitbild-Assessment aus Sicht der Informatik- und Computergeschichte, Berlin 1996, S. 205-234.
- Hellige, Hans Dieter (2003): Zur Genese des informatischen Programmabegriffs: Begriffsbildung, metaphorische Prozesse, Leitbilder und professionelle Kulturen. In Rödiger, Karl-Heinz (Hrsg.) Algorithmik - Kunst – Semiotik, Heidelberg 2003, S. 42-73.
- Hellige, Hans Dieter (2008a): Die Geschichte des Internet als Lernprozeß, in: Kreowski, Hans-Jörg (Hrsg.) Informatik und Gesellschaft. Verflechtungen und Perspektiven (Kritische Informatik, Bd. 4), Münster, Hamburg, Berlin 2008, S. 121-170.
- Hellige, Hans Dieter (2008b): Krisen- und Innovationsphasen in der Mensch-Computer-Interaktion. In: Hellige, H. D. (Hrsg.), Mensch-Computer-Interface. Zur Geschichte und Zukunft der Computer-Bedienung, Bielefeld 2008, S. 11-92.
- Hellige, Hans Dieter (2008c): Zentralistische, verteilte und dezentrale Architekturen einer digitalen Weltbibliothek. Vortrag auf der 17. HYPERKULT Computer als Medium zum Thema "Ordnungen des Wissens" am 4.07.2008 in der Universität Lüneburg. Verfügbar unter der URL: [XXX](http://www.hyperkult.de)
- Hellige, Hans Dieter (2009): Skalenökonomische Mengeneffekte der Informationstechnik und ihr Einfluss auf den Ressourcenverbrauch: Informationstechnik als Beispiel produktionsgetriebener Wachstumsdynamik. In: Ines Weller (Hrsg.): Systems of Provision & Industrial Ecology: Neue Perspektiven für die Forschung zu nachhaltigem Konsum, artec-paper 162, September 2009, S. 135-195.
- Hellige, Hans Dieter (2010): Medienkombinatorik für selektive Interface-Kulturen. Alternativen zu Paradigmen-geleiteten HCI-Entwicklungen, artec-Paper 170, November 2010, 55 S.
- Hölzle, Urs; Hanselmann, Matthias (2009): Google: Energieverbrauch von Rechenzentren lässt sich drastisch senken. Urs Hölzle im Gespräch mit Matthias Hanselmann. In: Deutschlandradio Kultur 6.3.2009 (http://www.deutschlandradiokultur.de/google-energieverbrauch-von-rechenzentren-laesst-sich.954.de.html?dram:article_id=144089).

- Hof Robert F. (1999): Sun Power. Is the center of the computing universe shifting? In: BusinessWeek 18.1.1999 (<http://www.businessweek.com/1999/03/b3612001.htm>).
- Hoffman, Thomas (2002): Irving Wladawsky-Berger Information Utilitarian. In: Computerworld 30.9.2002
- Hoover, J. Nicholas (2008): Dell Seeks, May Receive 'Cloud Computing' Trademark. In: InformationWeek 4.8.2008 (<http://www.informationweek.com/dell-seeks-may-receive-cloud-computing-trademark/d/d-id/1070661?>).
- Hsin, Lin Hsin (1996): It has a brain, but no stomach! In: Computer Times 20.11.1996 (<http://www.lhham.com.sg/biodata/ctjs.html>).
- Hwang Kai; Fox, Geoffrey C.; Dongarra Jack J. (2012): Distributed and Cloud Computing. Amsterdam, Boston, Heidelberg 2012.
- Joseph, Joshy; Ernest, Mark; Fellenstein, Craig (2004): Evolution of grid computing architecture and grid adoption models. In: IBM Systems Journal 43 (2004) 4, S. 624-645.
- Kagermann, Henning; Lukas, Wolf-Dieter (2011): Industrie 4.0: Mit dem Internet der Dinge auf dem Weg zur 4. industriellen Revolution. In VDI-Nachrichten 1.4.2011.
- Kahn, Robert E.; Cerf, Vinton G. (1988): An Open Architecture for a Digital Library System and a Plan for Its Development. The Digital Library project. Vol 1: The World of Knowbots (Draft), Corporation for National research Initiatives, März. 1988: im Internet unter: http://www.cnri.reston.va.us/pub_archive.html Dokumentennummer: hdl: 4263537/2091.
- Kanakamedala, Kishore; Krishnakanthan, Vasantha; Roberts, Roger P. (2006): Two new tools that CIOs want. In: McKinsey on IT, Number 8, Summer 2006, S. 32-33.
- Kelly, Kevin (2007): [A Cloudbook for the Cloud](#). In: Kellys Blog KK*, 15.11.2007 (<http://kk.org/thetechnium/2007/11/a-cloudbook-for/>).
- Kibbe, Andre (2008): Review: The Big Switch. In: Tools for Thought, 26.8.2008 (<http://tools-for-thought.com/2008/08/26/review-the-big-switch/>).
- Kleinrock, Leonard (1961): Information Flow in Large Communication Nets. Proposal for a Ph.D. Thesis, MIT, Ms. May 1961; im Internet unter: <http://www.cs.ucla.edu/~lk/LK/Bib/REPORT/PhD/> (zuletzt gesehen: 3.7.2013).
- Kleinrock, Leonard (1974/76): Queueing Systems, 2 Bde. New York, London, Sydney, Toronto 1974/76.
- Kraemer, Kenneth L.; Dedrick, Jason (1999): Sun Microsystems: Integrating its Own Enterprise. I.T. in Business, Center for Research on Information Technology and Organizations, UC Irvine August 1999 (<http://www.escholarship.org/uc/item/8xz9r879>).
- LaMonica, Martin (2006a): Web giants lure developers. In: CNET News Web services, 1.9.2006. (http://news.cnet.com/Web-giants-lure-developers/2100-7345_3-6111465.html).
- LaMonica, Martin (2006b): Amazon: Utility computing power broker. In: CNET News Web services, 16.11.2006 (http://news.cnet.com/Amazon-Utility-computing-power-broker/2100-7345_3-6135977.html).
- Leiner, Alan L; Notz, W. A.; Smith, J. L.; Weinberger, Arnold (1957): Organizing a network of computers to meet deadlines. In: Proceedings IRE-ACM-AIEE '57 (Eastern) Papers and discussions presented at the December 9-13, 1957, Eastern Joint Computer Conference, New York 1957, S. 115-128.
- Lenk, Alexander; Klems, Markus; Nimis, Jens u.a. (2009): What's inside the Cloud? An architectural map of the Cloud landscape. In: CLOUD '09 Proceedings of the 2009 ICSE Workshop on Software Engineering Challenges of Cloud Computing. Washington, DC 2009, S. 23-31.
- Levy, Steven (2011): In The Plex: How Google Thinks, Works, and Shapes Our Lives. New York, London, Toronto 2011.
- Li, C. Jin (2006): Erasure Resilient Codes in Peer-to-Peer Storage Cloud. In: Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on. Tooulouse. Mai 2006, Bd. 4, S. 233-236.
- Licklider, Joseph C. R.(1960): Man-Computer Symbiosis. In: IRE Transactions on Human Factors in Electronics, Bd.1 März 1960, S. 4-11.
- Licklider, Joseph C. R. (1965): Libraries of the Future. Cambridge, Mass. 1965.
- Lyons, Daniel (2005): 'On Demand' Letdown. In: Forbes, 24.2.2005 (http://www.forbes.com/2005/02/24/cz_dl_0224ibmsidebar.html).
- Mann, Alan O. (1959): Computer Power: A Public Utility? In: Computers and Automation 8 (1959) 4, S. 11-16.
- Markoff, John (2001): Internet Critic Takes on Microsoft. In: New York Times 9.4.2001 (<http://www.nytimes.com/2001/04/09/technology/09HAIL.html>).
- Markoff, John (2007): Software via the Internet: Microsoft in 'Cloud' Computing. In: The New York Times, 3.9.2007 (http://www.nytimes.com/2007/09/03/technology/03cloud.html?pagewanted=all&_r=0).
- Martin, Richard (2007): The Red Shift Theory. In: Dr. Dobb's The World of Software Development, 18.8.2007 (<http://www.drdobbs.com/the-red-shift-theory/201800873>).

- Marks, Eric A.; Lozano, Bob (2010): Executive's Guide to Cloud Computing. Hoboken, NJ 2010.
- Matlack, Richard C. (1955): The Role of Communications Networks in Digital Data Systems. In: Proceedings of the Eastern Computer Conference, AFIPS 8 (1955), S. 83-87.
- Miller, Rich (2008a): The Cloud As Seen in 2002. In: Data Center Knowledge, 5.9.2008. (URL: <http://www.datacenterknowledge.com/archives/2008/09/05/the-cloud-as-seen-in-2002/>).
- McCarthy, John (1962): Time-Sharing Computer Systems. In: Martin Greenberger (Hrsg.): Management and the Computer of the Future, Cambridge, Mass., New York, London 1962, S. 220-236.
- McCarthy, John (1970): The home information terminal---a 1970 view. In: Man and Computer. Proceedings of the 1st International Conference. Bordeaux, June 1970, Basel 1972, S. 48-57.
- Meinrath, Sascha D.; Losey, James W.; Pickard, Victor W. (2011): Digital Feudalism: Enclosures and Erasures from Digital Rights Management to the Digital Divide. In: Commlaw Conspectus 19 (2011), S. 423-479.
- Metz, Cade (2012): If Xerox PARC Invented the PC, Google Invented the Internet. In: Wired 8.8.2012 (<http://www.wired.com/2012/08/google-as-xerox-parc/>).
- Moody, Glyn (2002): Getting wired: Is there anything new in Sun Microsystems' N1? In: Computer-Weekly.com, Okt. 2002 (<http://www.computerweekly.com/opinion/Getting-wired-Is-there-anything-new-in-Sun-Microsystems-N1>).
- Morales, Alexandra Weber (2011): How the cloud (kinda) changes (sorta) everything. In: Software Development Times, 12.8.2011 (<http://sdtimes.com/how-the-cloud-kinda-changes-sorta-everything/>).
- Morgan, Timothy Prickett (2008): HP squeezes 85 data centers into 'six-pack'. In: The Register 1.12.2008 (http://www.theregister.co.uk/2008/12/01/hp_data_centers_done/).
- Ozzie, Ray (2005): The Internet Services Disruption. Internal Memo, Microsoft 28.10.2005. (<http://scripting.com/disruption/ozzie/TheInternetServicesDisruptio.htm>).
- Papadopoulos, Greg (2002): The Next Big Thing. In: Sun Microsystems (Hrsg.): N1—Revolutionary IT Architecture. 2002, S. 1-3.
- Papadopoulos, Greg (2006): The World Needs Only Five Computers. In: Greg Papadopoulos's Weblog, 10.11.2006 (https://blogs.oracle.com/Gregp/entry/the_world_needs_only_five).
- Papadopoulos, Greg (2009): The Intercloud. In: Greg Papadopoulos's Weblog 20.2.2009
- Parkhill, Douglas F. (1966): The Challenge of the computer utility, Reading, Mass., Palo Alto, London 1966.
- Pattay, Walter v. (1993): Die technologischen Ursachen für die wachsende Bedeutung internationaler Normen, Diss. Bremen 1993.
- Pianese, Fabio; Bosch, Peter; Duminuco, Alessandro u.a. (2010): Toward a Cloud Operating System. In: IEEE Network Operations and Management Symposium Workshops - NOMS Workshops, 2010, S. 335 – 342.
- Popper, Karl (1973). Über Wolken und Uhren. In: ders. (Hrsg.), Objektive Erkenntnis. Ein evolutionärer Entwurf. Hamburg 1973, S. 214-267.
- Regalado, Antonio (2011): Who Coined 'Cloud Computing'? In: MIT Technology Review, 31.10.2011 (<http://www.technologyreview.com/news/425970/who-coined-cloud-computing/#>).
- Rittinghouse, John W.; Ransome, James F. (2010): Cloud Computing Implementation, Management and Security. Boca Raton, London, New York 2010.
- Revett, Michael C.; Knul, Michael; Stephens, Lee (1997): Network computing. In: BT Technology Journal 15 (April 1997) 2, S. 172-178.
- Röhle, Theo (2010): Der Google-Komplex. Über Macht im Zeitalter des Internets. Bielefeld 2010.
- Roth, Daniel (2009): Time Your Attack: Oracle's Lost Revolution. In: Wired 18.01, 21.12.2009.
- Roush, Wade (2006): Servers for Hire. In: MIT Technology Review 28.9.2006 (<http://www.technology-review.com/news/406593/servers-for-hire/>).
- Ryan, Patrick S.; Merchant, Ronak; Falvey, Sarah (2011): Regulation of the Cloud in India. In: Journal of Internet Law, 15 (2011) 4, S. 7-17.
- Scantlebury Roger A.; Wilkinson, Peter T. (1974): The National Physical Laboratory Data Communication Network. In: Proceeding of the ICCC Conference, Stockholm August 1974, 223-228.
- Sándor, Nacsá (2012): The cloud experience vision of .NET by Microsoft 12 years ago and its delivery now with Windows Azure, Windows 8/RT, Windows Phone, iOS and Android among others. In: lazure2 wordpress. Experiencing the Cloud, 16.9.2012 (<http://lazure2.wordpress.com/2012/09/16/the-cloud-experience-vision-of-net-by-microsoft-12-years-ago-and-its-delivery-now-with-windows-azure-windows-8rt-windows-phone-ios-and-android-among-others/>).
- Saracini, Andrea (2010): IBM - Smarter Planet: Using cloud computing to deliver innovation and efficiency. IBM-Präsentation 17.2.2010, Folie 2 (<http://de.slideshare.net/innovationlabevents/ibm-smarter-planet-using-cloud-computing-to-deliver-innovation-and-efficiency>).

- Savvas, Antiny (2008): Dell fails to trademark 'cloud computing'. In: ComputerWeekly.com 19.8.2008 (<http://www.computerweekly.com/news/2240086667/Dell-fails-to-trademark-cloud-computing>).
- Selipsky, Adam (2010): Adam Selipsky of Amazon Web Services describes the origins of AWS and the role modular services play in a cloud-oriented business model. In: PricewaterhouseCoopers Technology Forecast 2010/4 (Driving growth with cloud computing), S. 22-25.
- Sheff, David (2003): Crank It Up. In: Wired 8.08, August 2000 (http://www.wired.com/wired/archive/8.08/loudcloud_pr.html).
- Sitaram, Dikar; Manjunath, Geetha (2012): Moving To The Cloud: Developing Apps in the New World of Cloud Computing. Amsterdam, Boston, Heidelberg 2012.
- Smarr, Larry (1997): Computational Infrastructure: Toward the 21st Century. In: Communications of the ACM 40 (1997) 1, S. 29-32.
- Smarr, Larry; Catlett, Charles E. (1992): Metacomputing. In: Communications of the ACM 35 (1992) 6, S. 44-52.
- Smith, Roger (2004): Grid Computing: A Brief Technology Analysis. CTOnet.org 2004 (URL: (http://www.ctonet.org/documents/GridComputing_analysis.pdf)).
- Sprague, Richard E. (1969): Information Utilities, Englewood Cliffs, N.J. 1969
- Stevens, Rick; Woodward, Paul; DeFanti, Tom; Catlett, Charlie (1997): From the I-WAY to the National Technology Grid. In: Communications of the ACM 40 (1997) 11, S. 51-60.
- Stokes, Adrian V.; Higginson, Peter L. (1975): The problems of connecting hosts into ARPANET. In: Communications Networks. Online 1975, Uxbridge 1975, S. 25-33.
- Strassmann, Paul A. (2005b): The Next I.T. Tsunami. Lecture auf der IP.4.IT Conference and Expo, Las Vegas November 2005 (Podcast unter <http://www.strassmann.com/pubs/talks/2005-11-ip4it.mp3>).
- Strassmann, Paul A. (2005c): Google: Model for the Systems Architecture of the Future. In: Lecture Series Presented by the George Mason University Volgenau School of Information Technology and Engineering, 5.12.2005 (<https://www.youtube.com/watch?v=TM7jrWiFeoQ>).
- Strassmann, Paul A. (2006): Who is best at net-centricity? Here's a hint: Google it. In: Defense Systems Magazine, February 2006 (<http://www.strassmann.com/pubs/defense/2006-2.html>).
- Sullivan, Danny (2006): Conversation with Eric Schmidt hosted by Danny Sullivan. In: Google Press Center 9.8.2006 (<http://www.google.com/press/podium/ses2006.html>).
- Sun Microsystems, Inc (2009): Introduction to Cloud Computing architecture. White Paper 1st Edition, June 2009
- Sweeny, Douglas M. (1998): Global Market Trends in the Networked Era. In: Long Range Planning, 31 (1998) 5, S. 672-683.
- Vaidhyanathan, Siva (2011): The Googlization of Everything (and Why We Should Worry), Berleley, Los Angeles 2011.
- Vencat, Girish (2009): Loudcloud: Early light on cloud computing. In: CNET News 23.2009. (<http://www.cnet.com/news/loudcloud-early-light-on-cloud-computing/>).
- Venters, Will; Whitley, Edgar A. (2012): A critical review of cloud computing: researching desires and realities. In: Journal of Information Technology 27 (2012), S. 179-197
- Vickers, Paul (2003): HP , HP Labs and Grid Computing. HP Resaerch Laboratories 29.9.2003.
- Vogels, Werner (2009): "Ahead in the Clouds" Lisa '09 Keynote by Werner Vogels. In: USENIX Update 5.11.2009 (<https://www.usenix.org/blog/ahead-clouds-lisa-09-keynote-werner-vogels>).
- Wang, Pi-Chung; Chan, Chia-Tai; Chen, Yaw-Chung (2000): An intelligent buffer management approach for GFR services in IP/ATM internetworks. In: Computer Communications 23 (August 2000), S. 1389-1399.
- Wang, Lizhe; Ranjan, Rajiv; Chen, Jinjun; Benatallah, Boualem (Hrsg.) (2012): Cloud Computing: Methodology, Systems, and Applications. Boca Raton, London, New York 2012.
- Waters, Richaerd (2009): Cloud Control. In: Financial Times 25.3. 2009 (<http://www.ft.com/intl/cms/s/0/c9e3bf12-1973-11de-9d34-0000779fd2ac.html#axzz3EnA7BuhD>).
- Youseff, Lamia; Butrica, Maria; Da Silva, Dilma (2008): Toward a Unified Ontology of Cloud Computing. In: Grid Computing Environments Workshop, 2008. GCE '08, November 2008, S. 1-10.
- Weiss, Aaron (2001): Microsoft's .NET: platform in the clouds. In: netWorker 5 (2001) 4, S. 26-31.
- Wiener, Norbert (1950): The Human Use of Human Beings - Cybernetics and Society. Cambridge, Mass. 1950.; zit. nach der Ausgabe von Free Association Books, London 1989.
- Wladawsky-Berger, Irving (2005c): The IBM Internet Division. In: Irving Wladawsky-Berger-Blog, 28.11.2005 (http://blog.irvingwb.com/blog/2005/11/the_ibm_interne.html#more).
- Wladawsky-Berger, Irving (2005d): On Demand Three Years Later - A Personal Reflection. In: Irving Wladawsky-Berger-Blog, 31.10.2005 (http://blog.irvingwb.com/blog/2005/10/on_demand_three.html).
- Wladawsky-Berger, Irving (2008a): Reflections on Cloud Computing. In: Irving Wladawsky-Berger-Blog, 31.3.2008 (<http://blog.irvingwb.com/blog/2008/03/reflections-on.html#more>).

- Wladawsky-Berger, Irving (2008b): The Promise and Reality of Cloud Computing. In: Irving Wladawsky-Berger-Blog, 14.7.2008 (<http://blog.irvingwb.com/blog/2008/07/the-promise-and.html#more>).
- Wolf, Gary (2003): The Great Library of Amazonia. In: Wired 11.12, 2003 (http://archive.wired.com/wired/archive/11.12/amazon_pr.html).
- Zeger, Hans G. (2012): Sicherheit um jeden Preis. Vortrag Initiative Menschen-Rechte, Feldkirch 3.10.2012. In: Newsletter Initiative Menschen-Rechte, 3.10.2012.