

# The Cloud vs. The Grid

## Proseminar Technische Informatik

Björn Karger

28. Januar 2009

Zusammenfassung:

Diese Ausarbeitung soll die beiden Konzepte hinter Cloud- und Grid-Computing erläutern. Hierbei sollen Gemeinsamkeiten und Unterschiede der Technologien aufgezeigt werden. Für das bessere Verständnis dieses Themengebietes wird ebenfalls auf mögliche Anwendungsgebiete eingegangen.

---

## Inhaltsverzeichnis

<b>1</b>	<b>Einführung</b>	<b>3</b>
1.1	Anwendungsgebiete . . . . .	3
1.1.1	Vorraussetzungen . . . . .	3
1.1.2	Parallele Anwendungen . . . . .	3
1.2	Anwendungsbeispiele . . . . .	4
1.2.1	Grid . . . . .	4
1.2.2	Cloud . . . . .	4
1.3	Zukunft . . . . .	5
<b>2</b>	<b>The Grid</b>	<b>5</b>
2.1	Ursprung des Namens . . . . .	5
2.2	Entstehung . . . . .	5
2.3	Definitionen & Aufbau . . . . .	6
2.4	Gemeinschaft . . . . .	7
<b>3</b>	<b>The Cloud</b>	<b>8</b>
3.1	Ursprung des Namens . . . . .	8
3.2	Entwicklung . . . . .	8
3.3	Definitionen & Aufbau . . . . .	8
3.4	Vision . . . . .	9
<b>4</b>	<b>Vergleich</b>	<b>10</b>
<b>5</b>	<b>Schlusswort</b>	<b>12</b>

# 1 Einführung

The Cloud vs. The Grid stehen für Cloud-Computing und Grid-Computing und gehören zum Gebiet des verteilten Rechnens (*Distributed Computing*). Den unterschiedlichen Techniken des verteilten Rechnens liegt die begrenzte Rechenkapazität eines einzelnen Computers zugrunde. Es existieren Aufgaben, bei denen ein einziger Computer Jahrzehnte zur Berechnung bräuhete. Um solche Aufgaben in annehmbarer Zeit lösen zu können, haben sich Menschen über die Umgehung dieser Beschränkung Gedanken gemacht und das Gebiet des verteilten Rechnens entstand. Oft ist der Betrieb eines eigenen Supercomputers, und somit das einfache Erhöhen der Rechenkapazität, für Wissenschaft und Industrie nicht praktikabel [4]. Eine Lösung für dieses Problem ist: die ungenutzten Ressourcen eines Universitäts- oder Firmennetzwerkes für rechenintensive Aufgaben zu nutzen. Cloud-Computing und Grid-Computing sind zwei unterschiedliche Konzepte, die ein verteiltes Rechnen auf vielen unterschiedlichen Computersystemen ermöglichen.

## 1.1 Anwendungsgebiete

In diesem Abschnitt sollen die Anwendungsgebiete des verteilten Rechnens beleuchtet werden. Die gemachten Aussagen treffen für alle Technologien des verteilten Rechnens zu und somit auch für Grid- und Cloud-Computing.

### 1.1.1 Voraussetzungen

Voraussetzung für die Bearbeitung einer Aufgabe auf vielen verteilten Rechnern ist eine gute Parallelisierbarkeit, was die Aufteilung einer Aufgabe in viele kleine Teilaufgaben bedeutet. Erfordert eine Aufgabe die Abarbeitung von aufeinander aufbauenden Teilschritten (seriell), so ist eine parallele Verarbeitung nicht durchführbar, da ein Teilschritt immer auf die Ergebnisse seines Vorgängers angewiesen ist. Es lassen sich also nur Aufgaben durch parallele Verarbeitung beschleunigen, in denen der serielle Anteil nicht zu groß ist.

Eine weitere Voraussetzung für das verteilte Rechnen ist die Kommunikation und Vernetzung, denn ohne diese wäre das Gebiet des verteilten Rechnens nicht existent, da der Transfer von Informationen die Voraussetzung für dieses Entwicklungsgebiet darstellt.

### 1.1.2 Parallele Anwendungen

Als Anwendungen können alle natürlichen Phänomene dienen, bei der Sensordaten analysiert werden müssen (z. B. Sonnenflecken, Solarwinde, Gravitationswellen, Wetterdaten). Zum besseren Verständnis werden zu solchen und anderen Phänomenen noch Simulationen durchgeführt. Auch das Untersuchen von Menschen erschaffenen Produkten wie Theorien, Algorithmen (z. B. Verschlüsselung, Hashing) oder Spielen (z. B. Schach, Dame) bilden ganze Forschungszweige. In Unternehmen könnten z. B. der Newsletter-Versand oder das

Umwandeln von Dateien als Anwendungen dienen. Diese Fülle von Anwendungen benötigen eine große Rechenkapazität, lassen sich aber gut parallelisieren, wodurch die benötigte Rechenleistung durch das verteilte Rechnen zur Verfügung gestellt werden kann.

## 1.2 Anwendungsbeispiele

Ob Radio-Signale aus dem Weltraum nach Zeichen intelligenten Lebens durchsucht werden (SETI@home) [22], Protein-Faltungen simuliert werden um Krankheiten wie Alzheimer oder Parkinson zu erforschen (Folding@home) [21] oder das Weltklima simuliert wird um die Auswirkungen der globalen Erwärmung zu erforschen (climateprediction.net) [19], eines haben alle diese Projekte gemeinsam: sie benötigen eine enorme Rechenleistung und stellen somit Anwendungsbeispiele für den Bereich des verteilten Rechnens dar.

Fälschlicherweise werden diese Beispiele häufig als Grids bezeichnet, obwohl sie, wie wir später sehen werden, gegen Definitionen des Grids verstoßen. Die verwendeten Protokolle sind weder standardisiert noch offen und es findet nur die Bereitstellung von Rechenleistung statt.

### 1.2.1 Grid

Letztes Jahr sollte der LHC (Large Hadron Collider) in Betrieb gehen, ein Experiment, welches multiple Theorien der Teilchenphysik bestätigen bzw. widerlegen sollte. Der Betrieb startet nach technischen Problemen voraussichtlich dieses Jahr. Dann erzeugt dieses Projekt geschätzte 15 Petabyte (15 Millionen Gigabyte) Datenaufkommen pro Jahr. Um diese Datenmengen verarbeiten und analysieren zu können wurde schon Jahre vorher die Entwicklung einer verlässlichen Grid-Struktur voran getrieben. Mit dem „Worldwide LHC Computing Grid“ (WLCG) [26] werden mehr als 140 Rechenzentren in 33 Ländern verbunden.

### 1.2.2 Cloud

1. Ein Betreiber einer Mehrspieler-Plattform, der für hunderttausende Benutzer virtuelle Welten simuliert, kann durch das Betreiben seiner Plattform in einer Cloud die Zuverlässigkeit und Verfügbarkeit steigern. Die weltweit schwankenden Benutzerzahlen, welche durch den Tagesrhythmus der Benutzer entsteht, können durch die dynamische Verlagerung der Rechenleistung ausgeglichen werden. Der Betreiber spart hierdurch überdimensionierte Rechenzentren, die außerhalb der Spitzenzeiten nicht voll ausgelastet wären.
2. Ein Online-Händler, mit einem gefragten Produkt, kann innerhalb weniger Minuten die Verfügbarkeit seiner Internetseite, durch Erhöhung der Rechenleistung des virtuellen Servers, garantieren. Somit ist er in der Lage auf kurzfristige Schwankungen (TV-Beitrag über das Produkt) effektiv zu reagieren und langfristige Schwankungen (Weihnachtsgeschäft) einzuplanen. Der Online-Händler sorgt so für die Zufriedenheit seiner Kunden bei optimalen Kosten für die IT.

3. Die New York Times (NYT) hatte 2007 das Vorhaben, alle Artikel von 1851-1922 online als PDFs verfügbar zu machen [15]. Dafür wären mehrere Wochen Rechenzeit der Firmenrechner nötig gewesen. Derek Gottfrid wurde mit der Aufgabe betraut und nutzte dafür „Amazon Web Services“, welche in einer Amazon-Cloud liefen. Mit hundert Recheninstanzen wurden alle 11 Millionen Artikel in weniger als 24 Stunden in PDFs umgewandelt.

### 1.3 Zukunft

„If communication is unlimited and free, then we are not restricted to using local resources to solve problems.“[11]

Die Voraussetzungen dieses Zitates sind zwar noch nicht erfüllt, offenbaren allerdings die Richtung, in welche sich die Computerverarbeitung hinentwickelt. Hierfür ausschlaggebend ist, dass die Vernetzung von Computersystemen, die Geschwindigkeit von Netzwerken und die Kapazität von Festplatten stärker wachsen als die Rechenleistung der Prozessoren [11]. Bei zunehmender Vernetzung von lokalen Netzwerken und deren zunehmend schnelleren Anbindung ans Internet spielt es keine Rolle mehr, ob eine Berechnung auf einem lokalen Rechner ausgeführt wird, oder ob die Berechnung im Internet stattfindet und nur das Ergebnis zurück übermittelt wird.

## 2 The Grid

### 2.1 Ursprung des Namens

Jeder Mensch nutzt täglich elektrische Geräte und muss sich dabei keine Gedanken über die eigentliche Stromversorgung machen. Ebenso einfach soll das Nutzen einer Infrastruktur aus vernetzten Ressourcen erfolgen, welche unter anderem Rechenleistung liefert. Die Vernetzung dieser Infrastruktur erfolgt über das Internet und ist somit ähnlich umfassend wie das elektrische Stromnetz (*Electrical Power Grid*). Basierend auf den Ähnlichkeiten beider Systeme wurde das neue System nach dem bestehenden System (Stromnetz) benannt und bekam den Namen "Grid".

### 2.2 Entstehung

Das Grid entstand als Folge von Experimenten im Bereich des Hochleistungsrechnens [3]. Diese Experimente fanden 1995 in den USA und Kanada statt und hatten zum Ziel, eine verteilte, virtuelle Anwendung über mehrere Rechen- und Entwicklungszentren hinweg zu realisieren (I-WAY-Projekt) [7]. Untersucht wurden eventuelle Probleme im Ressourcenmanagement und in der Ressourcenplanung, die durch die große Entfernung hätten auftreten können. Aus den Erkenntnissen der Wissenschaftler entstand die Definition des Grids und die genutzte Software wurde zur heutigen Grid-Middleware weiterentwickelt.

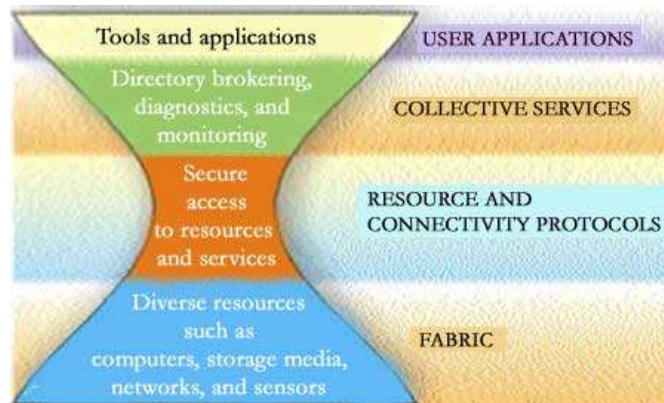


Abbildung 1: Konzeptioneller Aufbau eines Grids, Quelle: [11]

### 2.3 Definitionen & Aufbau

Ein Grid ist ein System, das Ressourcen koordiniert, die keiner zentralen Kontrollinstanz unterliegen und das standardisierte, offene, universelle Protokolle und Schnittstellen nutzt, um eine nicht triviale Dienstgüte zu liefern [12].

Die folgende Erläuterung der Definition erfolgt anhand der Quellen [18] [12] [11].

- Mit Ressourcen sind nicht nur Rechenleistung und Massenspeicher gemeint, sondern ebenfalls Datenbanken, Sensordaten wissenschaftlicher Messgeräte und vieles mehr. Die Positionen und Verwaltungseinheiten der Ressourcen sind beliebig, wodurch diese in verschiedenen Städten, Ländern und Kontinenten sowie Unternehmen, Haushalten, Universitäten, usw. (Verwaltungseinheiten) lokalisiert sein können. Die freigegebenen Ressourcen müssen sich in das jeweilige Sicherheitssystem der Verwaltungseinheit integrieren. Alle Ressourcen, die von den Benutzern zur Verfügung gestellt werden, befinden sich im konzeptionellen Aufbau eines Grids (Abbildung 1) auf unterster Ebene („fabric“). Die Vielzahl der Ressourcen wird durch den breiten Sockel einer Sanduhr repräsentiert.
- Die Koordination der Ressourcen erfolgt durch die Ebene „resource and connectivity protocols“, welche direkt überhalb der „fabric“-Ebene angesiedelt ist. Diese Ebene enthält einen Teil der standardisierten und offenen Protokolle. Die Protokolle ermöglichen Folgendes:
  - Basiskommunikation: Datenaustausch zwischen Ressourcen
  - Authentifikation: kryptographische Methoden zur Identitätsprüfung (Benutzer und Ressourcen)
  - Überwachung, Kontrolle und Inbetriebnahme der Ressourcen

Da diese grundlegenden Protokolle in jede Ressource integriert werden müssen, ist die Größe und Anzahl dieser beschränkt. Diese Einschränkung wird durch den verjüngten, mittleren Teil der Sanduhr dargestellt.

- Die dritte Ebene „collective services“, welche auf den ersten beiden Ebenen aufbaut, beinhaltet einen Teil der standardisierten und offenen Protokolle und ermöglicht mit diesen eine ressourcenübergreifende Zusammenarbeit. Die Dienste dieser Ebene erkunden das Grid nach neuen Ressourcen und binden diese auf Bedarf ein (resource broker). Die Zugriffskontrolle für die Ressourcengruppen, also welche Personen mit welchen Rechten auf welche Ressourcen zugreifen dürfen, wird in dieser Ebene realisiert. Weitere Dienste verwalten und überwachen zusammengefasste Ressourcen, die bereits grundlegende Aufgaben bewältigen können. Bei komplexeren Aufgaben werden dem Bedarf entsprechend Ressourcen beim „resource broker“ angefordert und sobald vorhanden eingebunden. Diese Flexibilität stellt eine „nicht triviale Dienstgüte“ dar und ermöglicht es, dass der Nutzen des gesamten Systems größer ist, als die Summe seiner einzelnen Teile. Diagnose- und Überwachungsdienste erhöhen ebenfalls den Gesamtnutzen, da mit ihnen Ausfälle (Soft- und Hardwarefehler) erkannt und kompensiert werden können. Die gestiegene Ausfallsicherheit ist eine Dienstgüte.
- Auf der obersten Ebene „user applications“ laufen alle Anwendungen, für die das Grid entwickelt wurden. Deren Anzahl und Vielseitigkeit wird durch die Breite der Sanduhr ausgedrückt.

Einen detaillierteren Aufbau liefert die „Open Grid Services Architecture“ (OGSA) [1], welche 2002 von dem Open Grid Forum (OGF) veröffentlicht wurde.

Eine Software, welche das zweite und dritte Level umsetzt, nennt sich Grid-Middleware. Die Middleware stellt hierbei die Plattform dar, auf der die verteilten Anwendungen laufen. Standard-Middlewares sind die Plattformen: UNICORE, Globus Toolkit und gLite [4].

## 2.4 Gemeinschaft

Eine Schlüsselrolle bei der Entstehung eines globalen Grids bildet die Gemeinschaft. Ohne eine Abstimmung aller Beteiligten ist ein solches Vorhaben unmöglich, daher hat sich das Open Grid Forum [10] gebildet, welches aus Vertretern aus Wissenschaft und Industrie besteht und einen globalen Grid-Standard anstrebt.

Die Nutzer eines Grids teilen sich in virtuelle Organisationen (VO) auf. VO's sind dynamische Zusammenschlüsse von Personen oder Instituten mit einem gemeinsamen Ziel/Projekt [5].

Neben regionalen und nationalen Grid-Projekten existiert noch das internationale „Enabling Grids for E-sciencE“-Projekt (EGEE) [9], welches zurzeit das weltweit größte Grid bildet. Dieses Grid steht Wissenschaftlern rund um die Uhr für ihre Forschungen zur Verfügung.

## 3 The Cloud

### 3.1 Ursprung des Namens

Das Internet wird in vielen Netzwerkdiagrammen als Wolke (Cloud) dargestellt und genau in dieser Wolke liegen die Dienstleistungen, um die es beim Cloud-Computing geht. Diese Dienstleistungen werden von Unternehmen angeboten, welche eigene Rechenzentren betreiben.

### 3.2 Entwicklung

Cloud-Computing wird besonders von Unternehmen vorangetrieben. In den letzten Jahren ist in den Unternehmen ein richtiger Hype um Cloud-Computing entstanden. Erste Errungenschaften dieses Bemühens sind z. B. mehrere Serviceangebote von Amazon (Elastic Cloud Computing - EC2, SimpleDB, Simple Storage Service - S3, Simple Queue Service SQS) die alle in einer Amazon-Cloud angesiedelt sind. Mit diesen Services lassen sich innerhalb von wenigen Minuten eine Vielzahl von virtuellen Computern oder mehrere Gigabyte Festplattenspeicher buchen. Viele weitere Unternehmen wie z. B. IBM und Microsoft betreiben Clouds oder haben den Bau von Clouds angekündigt [5]. HP, Intel und Yahoo, haben zur Förderung von Clouds zusammen mit wissenschaftlichen Instituten eine Kooperation gebildet und betreiben Cloud-System- und Anwendungsentwicklung sowie Cloud-Forschung. Jedoch entwickeln die meisten großen Unternehmen ihre eigenen Cloud-Angebote, wodurch die Umsetzung der Vision einer globalen Cloud erschwert wird [20] [24].

### 3.3 Definitionen & Aufbau

Es existiert keine genaue Definition was eine Cloud genau ist oder wie sie aufgebaut ist. Cloud-Computing bezeichnet eher einen Überbegriff, der ein System von Webtechnologien, Architekturen, Serviceangeboten und Businessmodellen beschreibt [20]. Cloud-Computing versteht sich als Weiterentwicklung des Grid-Computing und verspricht eine einfachere Benutzbarkeit, schnelleren Zugang und geringere Kosten als das Grid-Computing [4].

Eine allgemeine und oft zitierte Definition des Clouds stammt von James Staten, einem Technologieanalyst eines Marktforschungsunternehmens (Forrester Research). Demnach ist eine Cloud ein „Pool aus abstrahierter, hochskalierbarer und verwalteter IT-Infrastruktur, die Kundenanwendungen vorhält und nach Verbrauch abgerechnet wird“ [23]. Der Zugriff auf die Cloud erfolgt dabei orts- und geräteunabhängig über Standard-Internetprotokolle.

Abhängig davon, wer eine Cloud definiert, besteht eine Cloud zu bestimmten Anteilen aus folgenden Komponenten [8] [14].

- Software als Service (SaaS - Software as a service): Software wird als Dienstleistung angeboten. Der Dienstleister sorgt für die ständige Verfügbarkeit der Software über



das Internet. Die Infrastruktur, auf der die Software betrieben wird, ist unbekannt und wird vom Dienstleister gewartet. Auf Kundenseite ist nur eine minimale Hardwareausstattung nötig, da die Hardware nur zur Bedienung der entfernten Software dient. Ein Beispiel hierfür stellt Google Docs dar.

- Plattform als Service (PaaS - Platform as a service): Eine Web-Plattform, die als Dienstleistung angeboten wird und als Weiterentwicklung von SaaS die unterschiedlichen SaaS-Dienste bündelt. Die Plattform dient Programmierern als Entwicklungsplattform für Web-Anwendungen. Dabei werden die Programmierer, durch die Vielzahl von Grundfunktionen, entlastet und können sich auf das Wesentliche konzentrieren. Beispiele hierfür sind Google App Engine, Amazon Web Services und IBMs „Blue Cloud“.
- Infrastruktur als Service (IaaS - Infrastructure as a service): Liefert eine Infrastruktur als Dienstleistung, die auf jeden Kunden zugeschnitten ist. IaaS ermöglicht, dank Virtualisierung, eine skalierbare Infrastruktur, die jederzeit verändert werden kann und nach Verbrauch abgerechnet wird. Der Zugang zur Infrastruktur erfolgt auch hier über das Internet. Ein Beispiel hierfür ist Amazons Speicherdienst S3.

Die Architektur jeder Cloud basiert auf der Zusammenstellung dieser drei Komponenten. Je nach Unternehmen variieren die Schwerpunkte, daher setzt jedes große Internetunternehmen seine Vorstellungen einer Cloud in eigenen Datenzentren um. Der konzeptionelle Aufbau sollte dem des Grids ähneln. Durch das Betreiben von Clouds in Datenzentren kommt es zu einer Zentralisierung der Ressourcen, denn die Ressourcen stammen meist nur aus unternehmenseigenen Datenzentren, von denen es, der Effizienz wegen, nur wenige geben wird [25].

### 3.4 Vision

In der Zukunft stellt man sich das Cloud-Computing wie folgt vor: Das Internet entwickelt sich weiter zur Cloud [24], in der Dienste, die heutzutage noch lokal in den Unternehmen liegen, ausgelagert in der Cloud laufen. Das Unternehmen zahlt nur für die tatsächlich benötigten Leistungen und muss keine eigene Hardware betreiben. Unter den Diensten versteht man z. B. Betreiben eines virtuellen Web-, E-Mail- oder Datenbankservers, Speicherung von Daten oder Ausführen von leistungsintensiven Berechnungen. Die Dienste stehen natürlich auch Privatpersonen zur Verfügung.

## 4 Vergleich

Der folgende Vergleich basiert auf einem Artikel vom „Karlsruhe Institute of Technology“ [5] :

- Zielsetzung: Gemeinsam haben beide Konzepte, dass es um das Bereitstellen von verteilten Ressourcen über Netzwerkdienste geht. Das Grid ermöglicht eine gemeinsame Nutzung verteilter Ressourcen und das Ausführen von Anwendungen auf diesen. Bei Clouds werden die Ressourcen nach den Wünschen des Kunden bereitgestellt. Die Cloud sieht sich als Nachfolger des Grids und versucht die Schwächen dieses zu beseitigen.
- Infrastruktur: Der größte Unterschied liegt hier in der Form der Kontrollinstanz. Clouds unterliegen der zentralen Kontrolle des Betreibers, wohingegen bei Grids keine zentrale Kontrolle vorliegt. Beide Systeme haben Vor- und Nachteile, allerdings ergibt sich durch die zentrale Kontrolle der Cloud eine höhere Flexibilität, da ohne zentrale Kontrolle um Ressourcen verhandelt werden muss.

Ein weiterer Unterschied ist die Zentralisierung einer Cloud. Wenige vernetzte Rechenzentren lassen sich effizient betreiben und bilden heutzutage Clouds. Die Verteilung der Ressourcen in einem Grid ist aufgrund ihrer Vielzahl größer.

- Middleware: Die unterschiedlichen Middlewares sind bei Grids frei. Was bedeutet, dass die Middleware quelloffen ist und zu jedem Zweck genutzt werden darf. Die Middleware von Clouds erfüllen nicht alle diese Bedingungen und sind daher nicht frei. Die Grid-Middleware muss lokal installiert werden um Ressourcen zu nutzen. In einer Cloud werden alle Ressourcen durch Netzwerkdienste zur Verfügung gestellt, somit ist keine lokale Installation einer Software nötig.
- Standardisierung: Bei Clouds mangelt es an standardisierten Protokollen und Schnittstellen, denn jedes große Internetunternehmen entwickelt seine eigene Cloud. Zusammenarbeiten zwischen Unternehmen und Wissenschaft existieren erst kürzlich. Beim Grid hingegen existieren schon lange Standardisierungsgremien wie das OGF oder seine Vorgänger. Grid-Middlewares halten sich an diese Standards, was allerdings nicht die Zusammenarbeit der Grids/Middlewares garantiert [6].
- Anwendungen: Das Anwendungsgebiet von Cloud ist größer als das von Grids. Die virtuellen Maschinen der Clouds unterstützen nahezu jedes Betriebssystem und somit auch die darauf laufenden Programme. Um die Rechenleistung von mehreren Instanzen zu nutzen, muss das Programm parallele Verarbeitung unterstützen. Ein Beispiel hierfür ist das freie und kostenlose Programm Hadoop [16], welches die Verarbeitung auf verteilten Ressourcen unterstützt und womit z. B. die Umwandlung der PDFs der New York Times stattfand. Auf Grids hingegen laufen meist spezielle, wissenschaftliche Anwendungen, die für bestimmte Projekte entwickelt wurden.

- Benutzerschnittstelle: Schon bei der Entwicklung von Clouds wurde auf einfache Bedienbarkeit geachtet, daher ist das Nutzen einer Cloud mit geringer Einarbeitung möglich. Die Bedienung erfolgt meist über Netzwerkdienste und ermöglicht ein Ausführen der Anwendungen nach wenigen Minuten [13]. Bei speziellen Anforderungen können Clouds durch ihre Schnittstellen an die Bedürfnisse angepasst werden.

Bei Grids hingegen ist ein höherer Einarbeitungsaufwand nötig. Als Beispiel soll hier die Deutsche Grid-Initiative (D-Grid) dienen [6]. Für die Nutzung ist ein gültiges Grid-User-Zertifikat nötig, dieses erhält man, indem man einen Antrag bei einer Grid-Registration Authority stellt. Weiterhin ist die Gründung einer VO oder der Beitritt in eine VO notwendig. Beide Möglichkeiten erfordern eine Absprache mit anderen Personen. Eine erste Grid-Nutzung benötigt also mehrere Tage Zeit, bevor Anwendungen betrieben werden können. Zusätzlich muss für die Nutzung einer Ressource die lokale Grid-Middleware mit der Middleware der Ressource übereinstimmen. Dieses Beispiel zeigt, dass trotz standardisierter Schnittstellen eine Middleware bzw. Grid übergreifende Nutzung nicht garantiert ist.

- Entwicklung: Cloud-Computing wird relativ einseitig von Unternehmen entwickelt, eingesetzt und gefördert, während die Forschung sich erst langsam mit diesem Gebiet auseinandersetzt. Beim Grid-Computing verhält es sich genau andersherum. Nutzung und Entwicklung findet hauptsächlich in der Wissenschaft statt und die Förderung erfolgt durch staatliche Mittel [2].

Ein Grund für das geringe Engagement von Unternehmen in Grid sind folgende Nachteile:

- mangelnde Flexibilität (Software und Betriebssysteme)
  - schwierige Bedienbarkeit
  - geringes Vertrauen in die Sicherheit
  - geringes Vertrauen in die Beständigkeit der Infrastruktur
- Sicherheit: In der Wissenschaft und demnach auch im Grid gibt es kein großes Bedürfnis nach absoluter Sicherheit [3]. Ob ein paar Megabyte der 15 Petabyte jährlich produzierten Datenmenge des LHC abgehört werden oder nicht, interessiert die beteiligten Wissenschaftler eher wenig.

In der Wirtschaft hingegen, z. B. in Banken oder im Gesundheitswesen, wäre eine solche Sicherheitspolitik inakzeptabel. Daher ziehen die Unternehmen selbstentwickelte Clouds dem Grid vor. Dort kann die Sicherheitstechnik den Bedürfnissen entsprechend aufgebaut werden.

Unternehmen die Informationen in einer unternehmensfremden Cloud speichern, können nicht sagen, wo genau ihre Informationen liegen [17]. Es ist also möglich, dass die Daten in einem Land gespeichert werden, in dem der Datenschutz weitaus geringer ist als der deutsche Datenschutz. Je nach Rechtslage eines Landes kann es hierbei zu rechtlichen Problemen kommen.

Die unternehmensübergreifende Nutzung einer Cloud ist auch eine Frage des Vertrauens zwischen den Unternehmen. Denn die Sicherheit kann zwar als Dienstgüte vertraglich geregelt werden, aber, aufgrund der unbekannteren Bauart der Cloud, nicht überprüft werden.

- Geschäftsmodell: Das Geschäftsmodell von Clouds ist kommerziell ausgerichtet und eine Abrechnung erfolgt anhand der Zeit und der genutzten Ressourcen. Grids werden staatlich gefördert und ermöglichen eine kostenlose Nutzung.

„During the past 10 years, we have seen hundreds of grid projects come and go, passing away after government funding ran dry.“ [2]

Dieses Zitat zeigt ein Problem bei Grids auf. Sobald die staatlichen Mittel für Projekte gestrichen werden, zeigt sich, dass viele Projekte eine unzureichende Zukunftsstrategie besitzen [2].

- Dienstgütevereinbarungen (Quality of Service): In einer Cloud kann vertraglich die Dienstgüte vereinbart werden. Auch in einem Grid lässt sich die Dienstgüte beeinflussen, ist aber aufgrund geringerer Unterstützung schwerer zu realisieren.
- Ressourcenbereitstellung On-demand (auf Anfrage): Die sofortige Bereitstellung von Ressourcen beherrscht nur die Cloud, denn bei der gemeinschaftlichen Nutzung der Ressourcen im Grid muss immer gewartet werden, bis die benötigten Ressourcen ungenutzt sind und zusammengefasst wurden. Die Cloud hingegen wurde genau für diesen Zweck entwickelt und ermöglicht es Kunden jederzeit die gewünschten Ressourcen zu buchen.

## 5 Schlusswort

Die Zukunft wird zeigen, ob die jeweiligen Probleme der beiden Systeme überwunden werden können. Bei Clouds könnte es zu einer schnellen Entwicklung und starken Verbreitung kommen, sobald die Sicherheitsbedenken der Kunden ausgeräumt sind. Eine Standardisierung könnte zur Verwirklichung der Vision von einer globalen Cloud führen und somit das ganze Potenzial von Clouds ausschöpfen. Ohne eine Standardisierung könnte es allerdings auch zu Insellösungen der einzelnen Internetunternehmen führen. Die jetzt startende Zusammenarbeit mit der Wissenschaft ermöglicht eine Erforschung und eine Weiterentwicklung von Clouds.

Beim Grid-Computing ist die nahe Zukunft erst einmal gesichert, denn alleine der LHC wird mehr als ein Jahrzehnt Daten für das WLCG produzieren. Grid-Computing hat sich in der Wissenschaft bereits etabliert und weitere Projekte werden das Grid-Computing in den nächsten Jahren vorantreiben. Es gibt auf diesem Gebiet noch viele interessante Forschungsthemen, welche von Forschern erschlossen werden wollen.

---

## Literatur

- [1] Open Grid Services Architecture. [http://forge.gridforum.org/sf/go/projects.ogsa-wg/docman.root.published\\_documents.ogsa\\_1\\_5](http://forge.gridforum.org/sf/go/projects.ogsa-wg/docman.root.published_documents.ogsa_1_5) (28.01.2008).
- [2] Grids are Dead! Or are they? [www.on-demandenterprise.com/features/26060699.html](http://www.on-demandenterprise.com/features/26060699.html) (28.01.2008).
- [3] Christian Baun. Bewegliche Ziele, Anspruch und Wirklichkeit des Grid Computing. *c't*, 21, 2008.
- [4] Christian Baun. Tonangebend, Grid-, Cloud-, Cluster- und Meta-Computing. *c't*, 21, 2008.
- [5] Dr. Marcel Kunze Christian Baun. Cloud Computing – Infrastruktur als Dienst, December 2008.
- [6] D-Grid. <http://www.d-grid.de/index.php?id=212> (28.01.2008).
- [7] Thomas A. Defanti, Ian Foster, Michael E. Papka, Rick Stevens, and Tim Kuhfuss. Overview of the i-WAY: Wide-area visual supercomputing. *The International Journal of Supercomputer Applications and High Performance Computing*, 10(2/3):123–131, Summer/Fall 1996.
- [8] Torsten Eymann. Cloud Computing - Enzyklopaedie der Wirtschaftsinformatik. September 2008.
- [9] Enabling Grids for E-science. [www.eu-egee.org](http://www.eu-egee.org) (28.01.2008).
- [10] Open Grid Forum. [www.ogf.org](http://www.ogf.org) (28.01.2008).
- [11] Ian Foster. The Grid: A New Infrastructure for 21st Century Science. *Physics Today*, February 2002.
- [12] Ian Foster. What is the Grid? A Three Point Checklist. July 2002.
- [13] Anleitung für Amazons EC2. <http://docs.amazonwebservices.com/AWSEC2/2008-12-01/GettingStartedGuide/> (28.01.2008).
- [14] Andreas Göldi. Platform-as-a-Service: Das nächste IT-Schlachtfeld. *Netzwertig.com*, May 2008.
- [15] Derek Gottfrid. Self-service, prorated super computing fun!, November 2007.
- [16] Hadoop. <http://hadoop.apache.org/core/index.html> (28.01.2008).
- [17] Wolfgang Herrmann. Cloud Computing - das Buzzword des Jahres? *Computerwoche*, April 2008.
- [18] Instant-Grid. [www.instant-grid.org/?q=en/node/177](http://www.instant-grid.org/?q=en/node/177) (28.01.2008).

- [19] Klimasimulation. [www.climateprediction.net](http://www.climateprediction.net) (28.01.2008).
- [20] Johann Kurz. Cloud Computing - IT in der «Wolke». *netzwoche*, 17, 2008.
- [21] Protein-Faltung. [folding.stanford.edu](http://folding.stanford.edu) (28.01.2008).
- [22] SETI@home. [setiathome.berkeley.edu](http://setiathome.berkeley.edu) (28.01.2008).
- [23] Bill Snyder. Cloud computing begins to emerge from the haze. March 2008.
- [24] Kemal A. Delic und Martin Anthony Walker. Emergence of the academic computing clouds. *ACM Ubiquity*, 9, August 2008.
- [25] Aaron Weiss. Computing in the Clouds. *netWorker*, 11, 2007.
- [26] WLCG. [lcg.web.cern.ch/LCG/](http://lcg.web.cern.ch/LCG/) (28.01.2008).