Anbieter von Cloud Speicherdiensten im Überblick

Christoph Meinel, Maxim Schnjakin, Tobias Metzke,
Markus Freitag

Technische Berichte Nr. 84

des Hasso-Plattner-Instituts für
Softwaresystemtechnik
an der Universität Potsdam

Technische Berichte des Hasso-Plattner-Instituts für
Softwaresystemtechnik an der Universität Potsdam

Technische Berichte des Hasso-Plattner-Instituts für
Softwaresystemtechnik an der Universität Potsdam | 84

Christoph Meinel | Maxim Schnjakin | Tobias Metzke | Markus Freitag

Anbieter von Cloud Speicherdiensten im Überblick

Universitätsverlag Potsdam

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.dnb.de/ abrufbar.

Universitätsverlag Potsdam 2014
http://verlag.ub.uni-potsdam.de/

Am Neuen Palais 10, 14469 Potsdam
Tel.: +49 (0)331 977 2533 / Fax: 2292
E-Mail: verlag@uni-potsdam.de

Die Schriftenreihe **Technische Berichte des Hasso-Plattner-Instituts für Softwaresystemtechnik an der Universität Potsdam** wird herausgegeben von den Professoren des Hasso-Plattner-Instituts für Softwaresystemtechnik an der Universität Potsdam.

ISSN (print) 1613-5652
ISSN (online) 2191-1665

Dieser überarbeiteten Ausgabe ist ein Korrekturverzeichnis angefügt.

Online veröffentlicht auf dem Publikationsserver der Universität Potsdam
URL http://pub.ub.uni-potsdam.de/volltexte/2014/6878/
URN urn:nbn:de:kobv:517-opus-68780
http://nbn-resolving.de/urn:nbn:de:kobv:517-opus-68780

Zugleich gedruckt erschienen im Universitätsverlag Potsdam:
ISBN 978-3-86956-274-2

Inhaltsverzeichnis

Abbildungsverzeichnis **5**

Tabellenverzeichnis **7**

1 Executive Summary **10**

2 Einführung **13**

3 Cloud Storage **16**

4 Kriterienübersicht und Anbietervergleich **21**

4.1 Recht 21
4.2 Sicherheit 25
4.3 Verfügbarkeit 29
4.4 Kosten 31
4.5 Vertrauen, Zertifikate und Standards 34
4.6 Zugriffsmöglichkeiten 36
4.7 Performance 40
4.8 Übersicht 41

5 Performanz **43**

5.1 Ausgangslage 43
5.1.1 Versuchsanordnung 43
5.1.2 Zielstellung 43
5.1.3 Metriken 44
5.2 Ergebnisse 45
5.2.1 Übertragungszeit 45
5.2.2 Antwortzeit 51
5.2.3 Belastbarkeit 53
5.2.4 Verfügbarkeit 54

6 Provider im Detail **59**
6.1 Microsoft Windows Azure 59
6.2 Amazon S3 62
6.3 Google Cloud Storage 65
6.4 Rackspace Cloud Files 68
6.5 HP Cloud Object Storage 71
6.6 Nirvanix Public Cloud Storage 74

7 Fazit **77**

Literaturverzeichnis **78**

Abbildungsverzeichnis

4.1 Kosten für internen Speicher unter Einbezug aller Nebenkosten für Strom, Administration und Ähnlichem [66] 31
4.2 Kosten für Cloud Storage unter Einbezug aller Nebenkosten für Strom, Administration und Ähnlichem [66] 32
4.3 Kosten für internen Speicher ohne Betrachtung aller anfallenden Nebenkosten [66] 33
4.4 Einordnung der bewerteten 20 Cloud-Standards des BMWi [14] 35

5.1 Transferzeiten für Downloads der unterschiedlichen Anbieter hinsichtlich verschiedener Dateigrößen von 100 KB bis 1 MB. 47
5.2 Transferzeiten für Downloads der unterschiedlichen Anbieter hinsichtlich verschiedener Dateigrößen von 10 MB bis 1 GB. 48
5.3 Transferzeiten für Uploads der unterschiedlichen Anbieter hinsichtlich verschiedener Dateigrößen von 100 KB bis 1 MB. 49
5.4 Transferzeiten für Uploads der unterschiedlichen Anbieter hinsichtlich verschiedener Dateigrößen von 10 MB bis 1 GB. 50
5.5 Antwortzeiten der untersuchten Anbieter auf eine *getHash* Anfrage zu unterschiedlichen Tageszeiten. 52
5.6 Durchschnittliche Übertragungszeiten einer 10 MB-Datei bei steigender Anzahl paralleler Uploads von 10 MB-Dateien. 55
5.7 Durchschnittliche Übertragungszeiten einer 10 MB-Datei bei steigender Anzahl paralleler Uploads von 10 MB-Dateien. 56
5.8 Notwendige Übertragungszeit einer 100 MB-Datei bei steigender Segmentierung und Parallelität. 57
5.9 Notwendige Übertragungszeit einer 100 MB-Datei bei steigender Segmentierung und Parallelität. 58

6.1 Speicherkonzept von Windows Azure BLOB Storage [16] 60
6.2 Amazon Bucket System [34] 63
6.3 Speicherkonzept von Google Cloud Storage 66
6.4 Speicherkonzept von Rackspace Cloud Files 69

6.5 Speicherkonzept von HP Cloud Object Storage 72
6.6 Speicherkonzept von Nirvanix Public Cloud Storage 75

Tabellenverzeichnis

4.1 Providerübersicht hinsichtlich der *Rechtslage* 24
4.2 Providerübersicht hinsichtlich der *Sicherheit* 28
4.3 Providerübersicht hinsichtlich der *Verfügbarkeit* 30
4.4 Kostenübersicht für alle ausgewählten Dienste 33
4.5 Providerübersicht hinsichtlich der *Zugriffsmöglichkeiten* 38
4.6 Providerübersicht hinsichtlich der *Zugriffsmöglichkeiten* 39
4.7 Providerübersicht hinsichtlich Recht, Sicherheit, Verfügbarkeit und Schnittstellen 42

5.1 Gemessene Verfügbarkeit der untersuchten Dienste in einem Zeitraum von etwa 20 Tagen. 54

6.1 Kriterienübersicht für *Windows Azure* 61
6.2 Kriterienübersicht für *Amazon S3* 64
6.3 Kriterienübersicht für *Google Cloud Storage* 67
6.4 Kriterienübersicht für *Rackspace Cloud Files* 70
6.5 Kriterienübersicht für *HP Cloud Object Storage* 73
6.6 Kriterienübersicht für *Nirvanix Public Cloud Storage* 76

Durch die immer stärker werdende Flut an digitalen Informationen basieren immer mehr Anwendungen auf der Nutzung von kostengünstigen Cloud Storage Diensten [72]. Die Anzahl der Anbieter, die diese Dienste zur Verfügung stellen, hat sich in den letzten Jahren deutlich erhöht. Um den passenden Anbieter für eine Anwendung zu finden, müssen verschiedene Kriterien individuell berücksichtigt werden. In dieser Arbeit wird eine Auswahl an Anbietern etablierter Basic Storage Diensten vorgestellt und miteinander verglichen. Für die Gegenüberstellung werden Kriterien extrahiert, welche bei jedem der untersuchten Anbieter anwendbar sind und somit eine möglichst objektive Beurteilung erlauben. Hierzu gehören unter anderem Kosten, Recht, Sicherheit, Leistungsfähigkeit sowie angebotene Schnittstellen. Die vorgestellten Kriterien können genutzt werden, um Cloud Storage Anbieter bezüglich eines konkreten Anwendungsfalles zu bewerten.

1 Executive Summary

Immer mehr Unternehmen stehen vor der dem Problem, rasant wachsende Datenmengen zu verwalten. Der stetige Ausbau firmeneigener Rechenzentren zur Bewältigung der steigenden Informationsflut erfordert trotz sinkender Speicherkosten je GB erhebliche Investitionen. Cloud Storage stellt eine Alternative zum on-promise Speicher für Unternehmen dar. Die vorliegende Untersuchung vergleicht eine Auswahl etablierter Cloud Storage Provider, die eine eigene physikalische Infrastruktur aufweisen können und als *Basic Storage Provider* bezeichnet werden.

Im Wesentlichen lassen sich für den Einsatz von Cloud Storage vier Szenarien herausstellen:

- **Primärspeicher:** Die operativen Anwendungsdaten des Tagesgeschäfts werden in der Cloud gespeichert. Häufig werden auch die Anwendungen selbst in der Cloud ausgeführt, um die Nähe zwischen Programm und Daten und somit Leistungsfähigkeit zu gewährleisten.

- **Backup:** Sekundäre Kopien zur Sicherung der operativen Unternehmensdaten können in die Cloud ausgelagert werden, um Ressourcen im Unternehmen zu sparen.

- **Archiv:** Daten, die nicht geschäftskritisch sind, werden häufig im operativen Geschäft nicht benötigt und somit archiviert. Cloud Storage kann als kostengünstige Alternative zur Langzeitspeicherung dienen.

- **Content Delivery:** Die weltweite Verteilung von Medieninhalten an Nutzer kann über Content Delivery Netzwerke realisiert werden, die Daten replizieren und an verschiedenen Knotenpunkten auf der Welt vorhalten, um geringe Antwortzeiten für alle Nutzer weltweit zu gewährleisten.

Die Eignung eines Cloud Storage Dienstes für diese Einsatzszenarien hängt von vielen Faktoren ab. Die wichtigsten Kriterien für eine objektive Gegenüberstellung von Cloud Storage Diensten finden sich in den Bereichen Recht, Sicherheit, Verfügbarkeit, Kosten, Zertifikate, genutzte Standards, Zugriffsmöglichkeiten und Leistungsfähigkeit.

Im Bereich *Recht* bestimmen Fragen nach den verfügbaren Speicherstandorten, nach Datenschutzregelungen und dem Eigentumsrecht an den gespeicherten Daten die Fähigkeiten eines Dienstes. Aus Sicht europäischer Unternehmen ist dieser Bereich besonders

relevant bei der Wahl eines Dienstes, da viele Anbieter aus den USA stammen und anderen Datenschutzgesetzen unterliegen als europäische Unternehmen. Alle Provider weisen in diesem Bereich gleiche Eigenschaften auf und bieten mindestens einen Standort in der EU an (mit Ausnahme von HP), lassen den Nutzer den Speicherstandort frei wählen und sichern Kompensationen bei Nichterfüllung festgelegter Qualitätsmaße seitens des Anbieters zu. Was im Insolvenzfall des Providers mit den Daten geschieht, ist jedoch bei keinem der Anbieter vertraglich geregelt.

Bezüglich der *Sicherheit*, also dem Schutz der Daten vor unbefugtem Zugriff und Verlust, liefern alle Anbieter ebenfalls ein einheitliches Bild. Alle Provider erlauben den Zugriff auf Daten über verschlüsselte Verbindungen, ermöglichen die Verwaltung von Benutzern und deren Zugriffsrechten und bieten dem Nutzer die physische Replikation der Daten und deren Speicherung an unterschiedlichen Standorten zur erhöhten Ausfallsicherheit an. Im Gegensatz zu allen anderen untersuchten Anbietern erlauben Google und Rackspace keinen Zugriff auf die Daten nach Vertragsende. Rackspace bietet zudem keine Verschlüsselung der Daten selbst an.

Die Verfügbarkeit der gespeicherten Informationen spielt eine große Rolle im Cloud Storage. Alle Anbieter definieren auf transparente Art und Weise, was ein Ausfall bei ihnen bedeutet und welche Verfügbarkeiten sie zusichern. Nirvanix garantiert hier mit 99,999 % die mit Abstand höchste Verfügbarkeit, gefolgt von Azure und HP mit 99,95 %. Amazon, Google und Rackspace geben eine Verfügbarkeit von 99,9 % an.

Die *Kosten* der Speicherdienste sind, mit Ausnahme von Nirvanix, auf einem vergleichbaren Niveau. Die Kosten pro GB belaufen sich auf etwa 0,04 bis 0,09 US Dollar bei allen Diensten, je nach Speichermenge und -standort. Darüber hinaus können Kosten für das Herunterladen von Daten sowie für den Zugriff auf diese entstehen. Nirvanix erhebt hier bis auf die reinen Speicherkosten keine weiteren Gebühren, Rackspace stellt lediglich den Download von Daten zusätzlich in Rechnung, nicht jedoch die Anzahl der Zugriffe auf diese.

Der Einsatz von *Zertifikaten* und *Standards* im Cloud Computing ist bei allen untersuchten Anbietern aufgrund der geringen Zahl etablierter, cloud-spezifischer Zertifizierungen und Technologien noch im Aufbau. Mit *OpenStack* sticht ein cloud-spezifischer Standard heraus, der sowohl von HP als auch von Rackspace genutzt wird.

Der Einsatz von Cloud Storage in Unternehmen erfordert Schnittstellen und Tools, über die gespeicherte Daten verwaltet werden können. Alle Cloud Storage Provider bieten hierzu eine Reihe von gut dokumentierten, unterstützten Programmiersprachen, Schnittstellen und Werkzeugen an. Rackspace bietet als einziger Anbieter eine eigene mobile Applikation, Azure und Amazon ein Entwicklungspaket für mobile Applikationen an.

Amazon und Google ermöglichen zudem das Einsenden von Festplatten auf dem Postweg, um den Transfer aus oder in den Cloud Storage für große Datenmengen zu beschleunigen und zu vereinfachen.

Die *Leistungsfähigkeit* eines Providers im Cloud Storage gibt an, wie schnell ein Dienst auf eine Anfrage antwortet und wie lange das Herunter- und Hochladen von Daten bei einem Dienst dauert. Hierbei unterscheiden sich die Anbieter sehr stark je nach Größe und Anzahl der verwendeten Dateien. Die detaillierten Ergebnisse zur Bestimmung der Performanz der Provider in Kapitel 5 liefern hier Einblicke in die Leistungsfähigkeit der Dienste in bestimmten Szenarien.

Insgesamt lässt sich keiner der untersuchten Anbieter pauschal als der am besten geeignete für alle Einsatzmöglichkeiten identifizieren. Alle Provider können spezifische Vorteile für einzelne der vorgestellten Szenarien vorweisen. Die Wahl des Dienstes ist stark von dem Standort und den Bedürfnissen des Nutzers abhängig.

2 Einführung

Der weltweite Siegeszug der PCs und Smartphones, eine stark wachsende Internetnutzung in den Schwellenländern sowie die zunehmende Verbreitung vernetzter Geräte wie z. B. Überwachungskameras und intelligenter Stromzähler hat dazu geführt, dass sich das digitale Universum in den letzten zwei Jahren auf unvorstellbare 2,8 Zettabyte verdoppelt hat. Laut der aktuellen Studie von IDC [15] soll das digitale Universum in den nächsten acht Jahren auf 40 Zettabyte anwachsen. Damit stehen immer mehr Unternehmen vor dem Problem rasant wachsender Datenmenge, die verwalten werden muss.

Obwohl die Speicherkosten pro GB immer weiter sinken, müssen Firmen weiterhin regelmäßig in den Ausbau ihrer Rechenzentren, neuer Server, aktuelle Kühlung und möglichst niedrigen Stromverbrauch investieren. Cloud Computing stellt ein Modell dar, welches den bequemen, skalierbaren Netzwerkzugriff auf gemeinsame, konfigurierbare Ressourcen ermöglicht [39]. Eine dieser Ressourcen ist Speicher, der so genannte Cloud Storage. Durch Nutzung von Cloud-Storage-Diensten können Unternehmen ihre traditionell in eigenen Rechenzentren vorgehaltenen Daten an einen externen Dienstleister auslagern, der über festgelegte Schnittstellen über das Internet diese bereitgestellt und nach dem tatsächlichen Verbrauch abgerechnet. Die angebotenen Dienste stehen dabei prinzipiell jedermann zur Verfügung, der im Besitz einer gültigen Kreditkarte ist.

Die Beliebtheit und Relevanz von Cloud Storage für Unternehmen hat in den letzten Jahren vermehrt zugenommen [11]. Die Cloud als Datenspeicher gehört heute zu den beliebtesten Cloud Anwendungen. Laut Untersuchungen des Marktforschungsinstituts iHS iSuppli [65] werden bis Jahresende 2013 rund 500 Millionen Anwender Cloud Speicher Dienste nutzen. Ende Juni nutzten laut iSuppli bereits 375 Millionen Nutzer solche Dienste, womit sich abzeichnet, dass die Prognose der Experten sogar übertroffen wird. In den nächsten Jahren soll sich die Zahl der Nutzer von kostenlosen und kostenpflichtigen Cloud-Speichern mehr als verdoppeln und iSuppli zufolge 2017 auf 1,3 Milliarden ansteigen.

Die Erwartungen der Nutzer sind dabei vielfältig: Bessere Skalierbarkeit, Kosteneinsparungen, Auslagerung nicht wertschöpfender Aktivitäten und sicheres Vorhalten der Daten. Trotz wirtschaftlicher Vorteile zögern jedoch viele Unternehmen, interne Daten an externe Anbieter zu übertragen. Besonders wenn es sich dabei um vertrauliche Daten wie z. B. Kundeninformationen, Buchhaltung oder juristische Dokumente handelt.

Eine Möglichkeit, die Anbieter von Cloud Storage, Cloud Storage Provider (CSP), zu unterscheiden, ist diese in Basic Storage Provider (BSP) und Advanced Storage Provider (ASP) zu unterteilen. BSPs betreiben eine eigene physikalische Infrastruktur zum Speichern der Daten. Üblicherweise bieten sie dem Endnutzer keine grafische Oberfläche zum Zugriff auf die Daten an. Stattdessen können die Daten über ein Application Programming Interface (API) programmatisch abgerufen werden. Zu den BSPs zählen u. a. Amazon S3 , Google Storage und Rackspace Cloud Files.

Im Gegensatz zu den BSPs betreiben ASPs keine eigene Infrastruktur zum Speichern der Daten und greifen hierfür auf die Dienste der erwähnten BSPs zurück. Dafür bieten sie dem Endnutzer in der Regel eine grafische Benutzeroberfläche zum Verwalten der Daten an, z. B. als Desktopanwendung oder als Web-Interface. Außerdem warten ASPs häufig mit Zusatzfunktionen auf, die über das reine Speichern von Daten hinausgehen. Dropbox ermöglicht hier z. B. die Synchronisation ganzer Ordner über Rechnergrenzen hinweg. Neben Dropbox sind Mozy, Google Drive und Apples iCloud weitere Beispiele für ASPs. Die Anzahl der Anbieter, die diese Speicher Dienste zur Verfügung stellen, hat sich in den letzten Jahren deutlich erhöht.

Im Rahmen dieser Arbeit werden eine Reihe der auf dem Markt verfügbaren BSPs und ihre Leistungsmerkmale vorgestellt. Darauf aufbauend wird ein Versuch unternommen, objektive Vergleichsaspekte anzuleiten, um anschließend Ansatzpunkte für eine Bewertung der Leistungsfähigkeit der einzelnen Dienstanbieter zu finden. Die gewonnenen Informationen können dazu genutzt werden, die Auswahl eines geeigneten Anbieters für individuelle Anwendungsszenarien einzuschränken.

Die Wahl eines Providers erfordert dabei eine Betrachtung vieler Kriterien. Bei einem genaueren Blick wird deutlich, dass es bei den Anbietern trotz gleicher Basisfunktionalitäten große Unterschiede gibt. Die grundlegenden Funktionalitäten zur Erstellung, Manipulation und Löschung von Daten sind bei allen Anbietern zu finden. Operationen zur Transcodierung und Größenanpassung von Multimediadaten direkt im Cloud Storage sind hingegen nur bei bestimmten Diensten vorhanden.

Diese Studie beleuchtet neben den technischen auch die rechtlichen und wirtschaftlichen Aspekte der Dienste. Sie untersucht die Anbieter unter den Kriterien Recht, Sicherheit, Vertrauen, Kosten, Schnittstellen und Leistungsfähigkeit und stellt die jeweiligen Stärken und Schwächen der einzelnen Anbieter in diesen Punkten heraus.

Dropbox.
Webdienst, dessen Geschäftsmodel auf Cloud Computing basiert, um einen Storage-Service für große Mengen an Daten ohne Datencenter bereitstellen zu können. Der Dienst wird von mehr als 100 Mio. Nutzern weltweit in Anspruch genommen und von etwa 70 Mitarbeitern betrieben. Der Nutzerzuwachs beträgt etwa 1 Nutzer pro Sekunde. Das Geschäftsmodel baut auf dem "freemium"-Ansatz auf. Dabei wird der grundlegende Dienst kostenlos angeboten und Nutzer können sich weitere Features, wie z. B. mehr Speicherplatz, hinzukaufen. Der jährliche Umsatz liegt bei ca. 500 Mio. US-Dollar[1]. Dropbox war ursprünglich die Idee eines MIT-Studenten, der einen Weg suchte, Dateien überall zugänglich zu machen ohne USB-Geräte mit sich führen zu müssen. Dropbox konnte das eigene rapide Wachstum durch die Nutzung von Amazon S3-Diensten bewerkstelligen, über deren Datencenter sich nun auch die Daten von Dropbox verteilen. Auf Basis dieser Cloud Computing-Infrastruktur hat Dropbox gleichermaßen einen Synchronisations-, Backup-, und File-Sharing-Dienst geschaffen, ohne eigene Datencenter zu benötigen.

Die Arbeit ist wie folgt aufgebaut: Kapitel 3 gibt eine Übersicht zu Vor- und Nachteilen von Cloud Storage sowie gängiger Einsatzszenarien des Cloudspeichers im wirtschaftlichen Kontext. Nachfolgend beschreibt Kapitel 4 die Kriterien, nach denen die ausgewählten Anbieter untersucht und miteinander verglichen werden. Ein detaillierter Performance Vergleich der Dienste folgt in Kapitel 5. Kapitel 6 stellt darauf folgend alle ausgewählten Anbieter basierend auf den definierten Kriterien im Detail vor, bevor in Kapitel 7 die Erkenntnisse der Arbeit zusammengefasst werden.

3 Cloud Storage

Die Popularität von Cloud Storage Diensten, wie aktuelle Studien zeigen [12, 72], nimmt rasant zu. Die Untersuchungen belegen, dass besonders kleine und mittlere Unternehmen (KMU) Storage Dienste aus der Cloud entweder bereits einsetzen oder in naher Zukunft einsetzen wollen. Die Beliebtheit begründet sich sowohl in den reduzierten Aufwendungen für eigene Speicherinfrastruktur und damit auch Kosteneinsparungen, als auch in der flexiblen Skalierbarkeit der Speicheranforderungen [12, 72]. Zudem bietet Cloud Storage eine einfache Möglichkeit, einem unvorhergesehenen endgültigen Datenverlust, wie z. B. durch Naturkatastrophen, Brände oder andere Ausfälle, vorzubeugen bzw. die Eintrittswahrscheinlichkeit für einen solchen Fall zu minimieren. Die Daten werden in einer anderen geografischen Zone bei einem CSP vorgehalten und können im Schadensfall von dort wiederhergestellt werden.

Auf der einen Seite haben Unternehmen hohe Erwartungen an das Konzept des Cloud Computing, auf der anderen Seite herrschen aktuell noch große Bedenken bezüglich des Kontrollverlustes, ungewisser Verlässlichkeit, Kosten- und Performanceunsicherheit sowie rechtlicher Konformität [72].

Ein Blick auf Statistiken zur weltweiten Cloud Storage Nutzung [12] verdeutlicht drei Fakten:

1. kleine und mittlere Unternehmen (KMUs) benutzen Cloud Storage im Durchschnitt schon seit 2 Jahren
2. große Unternehmen holen in den letzten 12 Monaten deutlich auf
3. Unternehmen mit hohen Sicherheitsanforderungen (Gesundheits- und Bankwesen) benutzen kaum Cloud Storage-Lösungen und zögern bei zukünftigem Einsatz

Dabei lassen sich die grundsätzlichen Szenarien zur Nutzung von Cloud Storage im Unternehmenskontext wie folgt zusammenfassen [67, 69]:

Primärspeicherung: Der Einsatz von Cloud Storage als Primärspeicher bedeutet dessen Ersatz oder Ergänzung für on-premise Speicher für laufende operative Tätigkeiten. Primärspeicher wird hauptsächlich dafür genutzt, die Anwendungsdaten der firmeninternen Software zu verwalten. Dabei wird auf den Speicher vorrangig programmatisch zugegriffen. Das bedeutet, dass Programme den Speicher zur

Datenerstellung und -verwaltung verwenden. Endnutzer greifen auf Primärspeicher im Normalfall nur indirekt zu, indem sie beispielsweise Kundendaten in einer CRM-Anwendung[1] ändern. Die Kundendaten werden von den Nutzern hierbei nicht direkt über das Dateisystem verwaltet, sondern über eine Anwendung, die diese Daten organisisert und verändert. Primärspeicher wird von laufenden Anwendungen genutzt, wodurch die bestehenden Daten gelesen und verändert sowie neue Daten erstellt werden und Lese- und Schreibzugriffe gleichermaßen häufig stattfinden. Zudem operiert Primärspeicher hauptsächlich mit unternehmensinternen und vertraulichen Daten, weshalb Vertraulichkeit und Sicherheit besonders wichtige Anforderungen für den Anwendungsfall sind. Wird der Primärspeicher in die Cloud verlagert, geht oft auch eine Verlagerung der Anwendungen in die Cloud mit einher. Das begründet sich in der notwendigen Nähe der Daten zur Anwendung [38, 10]. Je weiter die Anwendung von ihren Daten entfernt ist, desto länger benötigt sie, um auf diese zugreifen zu können. Die Performance der Anwendung wird mit wachsender Entfernung der Daten demnach schlechter.

Backup: Ein Backup beschreibt eine sekundäre Kopie bestehender gespeicherter Daten. Diese dient vorrangig der Sicherung und dem Schutz der operativen Unternehmensdaten. Backups können aber auch zum Load Balancing[2] genutzt werden. Sind die Unternehmensdaten nicht intern redundant gespeichert, können sekundäre Kopien auch in der Cloud gehalten werden. In einem solchen Anwendungsfall müssen diese Kopien stetig aktualisiert werden, weshalb Schreibzugriffe auf den Speicher die Lesezugriffe dominieren. Aus diesem Grund erfordern Backups vorrangig eine hohe Schreibperformance. Backup-Lösungen müssen zudem ähnliche Bedingungen bezüglich Sicherheit, Vertraulichkeit und rechtlicher Konformität erfüllen wie die Primärspeicherung, da es sich um operative Daten handelt. Ein weiterer Aspekt des Backups ist die Wiederherstellung. Werden Kopien in der Cloud gehalten, muss dies im Recovery-Plan des Unternehmens berücksichtigt werden, da die Übertragung großer Datenmengen vom Cloud Backup zurück ins Unternehmen sowohl kosten- als auch zeitintensiv werden kann.

Archiv: Nicht alle Daten eines Unternehmens sind geschäftskritisch. Daten, die nicht für das operative Tagesgeschäft gebraucht werden[3], werden deshalb oft archiviert. Das notwendige Archivsystem kann intern im Unternehmen eingebunden sein. Dies erfordert Investitionen in eigene Speicherinfrastrukturen sowie die Administration des Systems. Eine Auslagerung der Archivdaten in die Cloud kann eine kosten-

[1] Customer Relationship Management Anwendung
[2] Load Balancing beschreibt das Verteilen von Last auf mehrere parallel arbeitende Instanzen
[3] hierunter fallen zum Beispiel alte Projektdaten und Kennzahlen vergangener Jahre

günstige Alternative darstellen. Allerdings müssen Archivdaten oft einer Reihe rechtlicher Anforderungen gerecht werden (siehe Kapitel 4). Die rechtliche Konformität sowie die Sicherheit der Daten und der Datenschutz spielen in diesem Anwendungsfall eine wichtige Rolle. Besonders für europäische und deutsche Unternehmen ist dieser Punkt ein entscheidender Faktor, wenn eine Archivlösung in der Cloud angestrebt wird. Die Schreibzugriffe auf Archivdaten sind von höherer Bedeutung als die Lesezugriffe, da die Daten vorrangig in das Archiv geschrieben und nur unregelmäßig daraus abgerufen werden. Die Daten müssen beständig und sicher auch über einen Zeitraum von beispielsweise 25 bis 30 Jahren archiviert werden können. Der Zugriff auf Archivdaten ist nicht geschäftskritisch, daher muss lediglich die Verfügbarkeit zu einem bestimmten Zeitpunkt, nicht aber zu jedem Zeitpunkt sichergestellt sein.

Content Delivery: Content Delivery[4] bezeichnet die Verbreitung und Verteilung von meist großen Mediendaten für eine weltweit verteilte Menge an Nutzern. Prominente Beispiele hierfür sind Video-Portale wie *Youtube*[5] und Foto-Webseiten wie *Flickr*[6]. Die Verteilung geschieht meist über ein Content Delivery Netzwerk, ein Verbund aus einer primären Speicherinstanz[7] und Replica-Instanzen[8], um Anfragen von Endnutzern möglichst ökonomisch und schnell durch naheliegende Instanzen zu bedienen. Die zu verteilenden Daten werden einmalig erstellt und zur primären Instanz hochgeladen. Anschließend werden diese von Endnutzern angesehen und heruntergeladen. Zu beachten ist dabei, dass es sich häufig um große Einzeldateien handelt. In einem solchen Anwendungsszenario dominieren die Lesezugriffe die Schreibzugriffe. Darüber hinaus kann gesagt werden, dass im Gegensatz zum Anwendungsfall der Primärspeicherung Vertrauen und Sicherheit keine ausschlaggebende Rolle spielen, da es sich für gewöhnlich nicht um vertrauliche Daten handelt. Viel mehr hat Content Delivery hohe Anforderungen im Bezug auf die Performance des Datenverkehrs und die Lesezugriffe sowie die Verfügbarkeit der Replica-Instanzen. Auch die Kosten für die Übertragung vom primären Server auf die Replica-Server sowie für Anfragen von Endnutzern spielen hier aufgrund der hohen erwarteten Nutzerzahlen eine wichtige Rolle.

[4] Verteilung oder Zustellung von Inhalten, meist Medieninhalten, über ein globales Netz
[5] `http://www.youtube.com/`
[6] `http://www.flickr.com/`
[7] dieser Server hält die Originale der zu verteilenden Daten
[8] diese Server halten eine aktuelle Kopie der Daten an anderen Standorten als dem der primären Instanz

Die beschriebenen Anwendungsfälle stellen eine Klassifizierung der möglichen Einsatzszenarien in Unternehmen dar. An diesen Szenarien können potenzielle Nutzer die eigenen Einsatzgebiete orientieren. Alle Anwendungsfälle setzen unterschiedliche Schwerpunkte bezüglich ihrer Anforderungen, welche ihrerseits von den existierenden Cloud Storage Providern unterschiedlich umgesetzt werden. Diese Studie untersucht deshalb eine Auswahl etablierter Basic Cloud Storage Dienste bezüglich der ausschlaggebenden Aspekte der Szenarien und stellt sie vergleichend gegenüber (siehe Kapitel 4). Aus Gründen der Übersichtlichkeit untersucht diese Studie die folgenden Anbieter, die stellvertretend für existierende Cloud Storage Dienste stehen.

- Amazon S3
- Google Cloud Storage
- Microsoft Windows Azure BLOB Storage
- HP Cloud Object Storage
- Rackspace Cloud Files
- Nirvanix Public Cloud Storage

In die Auswahl wurden nur die Anbieter aufgenommen, welche mindestens die nachfolgenden Kriterien [67] erfüllen:

- Vorhandensein einer API zum Datenzugriff über das Internet und Internetprotokolle zur Bereitstellung eines einfachen Zugangs zu gespeicherten Daten
- Verfügbarkeit transparenter und automatisch skalierender Speicherkapazitäten, auf die bedarfsorientiert und uneingeschränkt zugegriffen werden kann. Nutzer müssen in der Lage sein, so viele Ressourcen wie nötig in Anspruch nehmen zu können.
- Bereitstellung definierter und nachvollziehbarer Sicherheits-, Datenschutz- und Verfügbarkeitsangaben als Teil eines Service Level Agreements[9] (SLA). Letztere sichern dem Benutzer bestimmte Merkmale des Dienstes vertraglich zu und legen Strafen seitens des Providers im Falle eine Verletzung dieser Zusicherungen fest.
- nutzungsbasierte Abrechnung der Speicherkapazität und des Datentransfers auf einem granularen Level wie *pro Gigabyte pro Monat* für die Speicherkapazität, um die tatsächlich beanspruchten Ressourcen nachvollziehbar zu machen

[9] Vereinbarung zwischen zwei Vertragspartnern über den genauen Umfang der im Vertrag garantierten Leistungen

- Aufweisen einer etablierten Marktreife basierend auf Speichermenge, Kundenzahl oder Umsatz aus Cloud Storage Angeboten, die von dem Anbieter verwaltet werden, um eine gewisse Aussagekraft und Relevanz der zu vergleichenden Dienste sicherzustellen

4 Kriterienübersicht und Anbietervergleich

Wie bereits in Kapitel 3 erwähnt, gibt es neben Vorteilen wie der Kostenreduktion und flexibler Skalierung noch immer eine Reihe von Hürden und Bedenken, die Unternehmen gegenüber Cloud Storage anführen. Hierzu zählen vor allem Sicherheitsbedenken, unbekannte Leistungseinbuße, Verfügbarkeit sowie mangelnde rechtliche Konformität [72]. In diesem Kapitel werden die ausgewählten Cloud Storage Anbieter hinsichtlich dieser Bereiche mit jeweils relevanten Unterpunkten untersucht und miteinander verglichen. Der Vergleich soll Aufschlüsse darüber geben, in welchem Maße sich die Anbieter den vorhandenen Hürden angenommen haben. Zudem werden zusätzliche Bereiche beleuchtet, wodurch ein breiterer Überblick über die Dienste und deren Funktionalitäten ermöglicht werden soll. Im Folgenden werden die untersuchten Kriterien vorgestellt, welche als Grundlage des Vergleichs der ausgewählten Anbieter dienen.

Als Ausgangspunkt für diese Studie dienen die Angaben aus den offiziellen und öffentlich zugänglichen Dokumenten und Informationen der jeweiligen Anbieter. Dies umfasst die Vertragsvereinbarungen[1], die darin referenzierten Service Level Agreements[2], sowie weitere technische Informationsquellen wie Whitepaper[3], offizielle Blogs und Websites mit technischen Erläuterungen[4].

4.1 Recht

Der Bereich Recht beleuchtet die Grundlagen für vertragliche Vereinbarungen zwischen CSP und Nutzer. Das deutsche Recht der Auftragsdatenverarbeitung, wie es bei Cloud Computing vorliegt, ist noch nicht an Cloud-Verhältnisse angepasst. Auf europäischer Ebene bedarf es noch einer zeitgemäßen Weiterentwicklung in Form einer Reform des Datenschutzrechtes [9].

[1] oft auch *Konditionen*, *Terms of Service* oder *Terms and Conditions* genannt, zu finden unter [26, 63, 3, 53, 25, 46]
[2] [31, 62, 8, 52, 24, 44]
[3] [22, 7, 33, 54, 49, 48, 51, 43, 35, 50]
[4] [40, 61, 32, 29, 27, 58, 60, 57, 59, 6, 4, 5, 2, 42, 30, 28, 23]

Bei einem Vertragsschluss mit einem Dienstanbieter müssen grundsätzliche Fragen bindend geklärt werden. Die Parteien, welche den Vertrag abschließen, unterliegen nicht zwingend dem gleichen Recht bzw. haben nicht den gleichen Standort. Hier gilt die Vertragsfreiheit und es muss entschieden werden, welches Landesrecht für die Auslegung des Vertrages anzuwenden ist. Außerdem muss geklärt werden, was im Falle von Vertragsverletzungen passiert und wie bei einem möglicherweise unvorhergesehen Vertragsende mit den gespeicherten Daten umgegangen wird. Daneben kann der Dienstanbieter auch an zusätzliche Gesetze, wie z. B. den US-amerikanischen Patriot Act[5], gebunden sein. CSP mit Sitz in den USA können ihren Nutzern nach eigenen Angaben nicht garantieren, dass amerikanische Behörden keine Einsicht in die gespeicherten Daten nehmen werden [11].

Für europäische Nutzer ist ein Anbieter, dessen Dienste dem EU-Recht bzw. dem Recht eines EU-Staates zugrunde liegen, zu bevorzugen, da das EU-Recht restriktiver beim Umgang mit Daten ist und auch Datenschutz eine größere Rolle als im US-Recht spielt. So wird z. B. im Entwurf der EU-Datenschutz-Grundverordnung bereits zwischen dem "für die Verarbeitung Verantwortlichen" und dem "Auftragsverarbeiter" unterschieden [9], was eine klare rechtliche Beschreibung von Datenschutzverantwortlichkeiten darstellt. Insgesamt genügen die aktuellen Regelungen den wirtschaftlichen Anwendungsszenarien bisher aber kaum. Einen Standort in der EU zu besitzen ist also ein klares Kriterium für einen Provider. Dieses Kriterium erfüllen auch alle Anbieter mit Ausnahme von HP.

Storage-Anbieter aus den USA, die keinen Standort in der EU besitzen, lassen für die Nutzung ihrer Dienste auch ausschließlich US-Recht gelten. Anbieter mit Standorten in der EU, und der Option der freien Wahl des Standortes, können dagegen ihren Nutzern größere rechtliche Zuversicht bieten. Alle untersuchten Provider, mit Ausnahme von HP, bieten ihren Nutzern diese Möglichkeit. Kunden können also bei fast allen CSP frei wählen, an welchem Standort die Daten gespeichert werden sollen, mindestens ein europäischer Standort eingeschlossen. Hierbei ist jedoch zu beachten, dass dies nur den endgültigen Speicherort sichert. Es ist nicht ausgeschlossen, dass die Daten trotzdem über Server in den USA oder andere Länder gelangen.

Weiterhin wird untersucht, inwiefern die Verträge durch gesetzliche Regelungen wie den Patriot Act beeinflusst sein können und somit den Datenschutz gefährden. Rechtliche Sicherheit kann hier das sogenannte EU Safe Harbor Agreement[6] bieten. Unternehmen, die sich dem Agreement anschließen, garantieren die Einhaltung von Datenschutzrichtlinien äquivalent zu EU-Recht. Aus diesem Grund wurde die Teilnahme an diesem Abkommen als ein Kriterium in diese Studie aufgenommen.

[5] `http://www.fincen.gov/statutes_regs/patriot/index.html`, aufgerufen am 10.06.2013
[6] `http://export.gov/safeharbor/`, aufgerufen am 10.06.2013

Tatsächlich garantieren auch alle CSP die Einhaltung der Richtlinien im Sinne dieses Agreements. Das Kriterium ist also für alle Anbieter erfüllt. Zu beachten ist, dass aber trotzdem alle Anbieter laut ihrer Verträge immer noch US-Recht und möglichen gerichtlich verordneten Kontrollen unterliegen.

Patriot Act.
Der Patriot Act war eine Reaktion der US-amerikanischen Regierung auf die Anschläge vom 11. September 2001. Er verleiht Regierungsorganisationen neue Rechte und weicht die Grenze zwischen Judikative und Exekutive auf. Im Speziellen Geheimdienste erwarben so neue Befugnisse, u. a. Einsicht in Daten, die datenschutzrechtlich geschützt sind. [37]. Der Patriot Act verpflichtet Unternehmen aus bestimmten Industriebereichen dazu, Behörden auf Anfrage Zugriff zu ihren Daten zu verschaffen. Zugriff verschaffen bedeutet in diesem Fall, nicht nur die Herausgabe von Daten, die in den USA gespeichert sind. Es umfasst auch die Schaffung einer Möglichkeit zur Einsichtnahme in Daten, welche sich nicht physisch in den USA befinden und z. B. bei einem Tochterunternehmen in Europa gespeichert sind.

Dazu gehören auch Regelungen des Eigentumsrechtes an den Daten. Ein weiteres Kriterium sind Vertragsstrafen bzw. Regelungen bei Verstößen gegen das SLA. Wann genau ein Ausfall eines Dienstes auch nach Vertragsdefinition vorliegt, unterscheidet sich zwischen den Anbietern. Kulante und transparente Regelungen ohne aushebelnde Ausnahmefälle sind hier für den Nutzer wünschenswert. Weitere Untersuchungspunkte sind also die Transparenz und die Regelung der Vertragsstrafen, die konkrete Gefährdung durch Gesetze, die den Datenschutz betreffen, sowie das Eigentumsrecht an den Daten.

Vertragsstrafen in Form von Gutschriften[7] räumen alle Anbieter im Falle von Nichterfüllung der zugesicherten Vereinbarungen (siehe SLA) ein. Dabei werden grundsätzlich keinerlei Aus- oder Rückzahlungen eingeräumt. Die tatsächlichen Kosten, die durch einen Ausfall für ein Unternehmen entstehen, bleiben ebenso unberücksichtigt. Die Höhe der maximal möglichen Gutschriften unterscheiden sich auch stark bei den Anbietern. Bei Nirvanix, Amazon S3 und Windows Azure wird ein Maximum von 25 % des Rechnungsbetrages zugesichert. Bei HP sind es 30 % und bei Google sogar 50 %. Doch selbst wenn die Verfügbarkeit auf 0 % sinken würde, was einem Totalausfall gleich kommen würde, ist keine höhere Gutschrift möglich. Eine positive Ausnahme bietet hier Rackspace. Je nach tatsächlicher Verfügbarkeit ist hier eine Gutschrift bis zu 100 % möglich.

[7] eine Gutschrift bedeutet eine Anrechnung eines vertraglich definierten prozentualen Betrages der anstehenden Rechnung auf die folgende

Das Eigentumsrecht ist bei fast allen Anbietern genau geregelt. Lediglich HP macht im Vertrag keinerlei Angaben dazu. Sonst enthalten die Verträge Paragraphen, die zusichern, dass alle Inhalte eines Nutzers, und jegliches geistige Eigentum auch beim Nutzer bleiben. Andersrum bleiben natürlich auch die Services und die Software der Anbieter auch immer Eigentum der jeweiligen Provider.

Ein weiterer rechtlicher Aspekt ist der Fall der Zahlungsunfähigkeit des CSP. Eine Studie des Nationale Initiative für Informations- und Internet-Sicherheit e.V. (NIFIS) sieht Insolvenz als eines der größten rechtlichen Risiken im Cloud Computing [47]. 56 % der befragten Unternehmen sehen Insolvenz als Nachteil. Die größte Frage ist hier, was mit der Hardware des Cloud-Anbieters und den darauf enthaltenen Daten geschieht. Im Rahmen der Untersuchung stellte sich heraus, dass der Fall einer Insolvenz sowie die daraus resultierenden Fragen bei keinem der Anbieter vertraglich geregelt sind. Lediglich Google erwähnt den Insolvenzfall überhaupt, aber auch nur in der Form, dass im Falle einer Insolvenz des einen Vertragspartners, der andere das Recht auf Kündigung des Vertragsverhältnisses hat.

	Azure	**Amazon S3**	**Google**	**Rackspace**	**HP**	**Nirvanix**
mind. 1 Standort in der EU	✓	✓	✓	✓	–	✓
Standort frei wählbar	✓	✓	✓	✓	✓	✓
Teilnehmer am Safe Harbor Agreement	✓	✓	✓	✓	✓	✓
Vertragsstrafen bei Nichterfüllung des SLA	✓	✓	✓	✓	✓	✓
nutzerseitiges Eigentumsrecht an den Daten	✓	✓	✓	✓	–	✓
Insolvenzfall geklärt	–	–	–	–	–	–

Tabelle 4.1: Providerübersicht hinsichtlich der *Rechtslage*

4.2 Sicherheit

Sicherheit von Daten bedeutet Schutz gegen unbefugten Zugriff oder Verlust. Sicherheitsrisiken können je nach Anwendungsfall und Nutzerempfinden unterschiedlich stark ausfallen oder wahrgenommen werden. Generelle Risiken können wie folgt zusammengefasst werden [56]:

- Angriffe aus dem Internet
- Zugriff auf eigene Daten durch Unbefugte und Konkurrenten
- versteckte Datentransaktionen
- Verletzung der Datenschutzvorgaben
- Missbrauch der Kundendaten für Selbstzwecke des CSP

Analysten zufolge führt das Vorhandensein nur eines dieser Risiken bereits zur Unsicherheit bei Nutzern. Transparenz und Nachvollziehbarkeit bei der Ergreifung von Sicherheitsmaßnahmen wirken dieser Unsicherheit entgegen. Eine Nichterfüllung von Sicherheitsmaßnahmen ist als ebenso schwerwiegend zu betrachten wie die mangelnde Nachvollziehbarkeit dieser. Eine Möglichkeit zur Verhandlung von Sicherheitsauflagen sind sogenannte Security Service Level Agreements (SSLA) oder Protection Level Agreements [56]. Zusätzlich zu den grundlegenden SLAs beschreiben SSLAs weitere Sicherheitsauflagen, die der Anbieter einhalten muss. Sie beschreiben zudem, wie Nutzer die Einhaltung kontrollieren können und wie Missstände berichtet werden. SSLAs gehören allerdings noch nicht zum Status Quo und keiner der untersuchten Provider bietet sie in dieser Form an.

Auf nationaler Ebene ließen sich generell rechtliche Vorgaben umsetzen, aufgrund der international unterschiedlichen Rechtslage könnte dies aber eine einseitige Benachteiligung für bestimmte Dienstleister bedeuten und ist daher eher unrealistisch. Die tatsächlich einzige sichere Variante zum Schutz von Daten ist, diese auf technischer Ebene zu sichern, d. h. eine durchgehende und vollständige Verschlüsselung der Daten durchzuführen. Hierbei spielt die Umsetzung eine wichtige Rolle für die tatsächliche Sicherheit. Der Regelung des Zugangs zu den Schlüsseln sowie deren Aufbewahrung sind bei der Umsetzung entscheidend. In dieser Studie wird daher untersucht inwiefern die Anbieter eine Form der Verschlüsselung oder des gesicherten Zugangs zu den Daten unterstützen. Bei der Bewertung der Sicherheit wird zwischen *physischer* Sicherheit der Daten in Datenzentren des BSP und *logischer* Sicherheit beim Zugriff auf die Daten unterschieden.

Im Bereich der physischen Sicherheit soll beleuchtet werden, wie gut Daten gegen äußere Einflüsse wie Naturgewalten geschützt sind, die einen unmittelbaren Verlust der

Daten nach sich ziehen können. Aufgrund des fehlenden Zugangs zu den Rechenzentren, der zur Untersuchung der physischen Sicherheit gemäß des IT-Grundschutzes[8] nötig ist, beschränken sich die Untersuchungen in diesem Bereich auf die Replikation. Die Speicherung von Daten an nur einem Standort ist mit dem Risiko des unwiderruflichen Datenverlusts verbunden. Auch der Ausfall einzelner Server innerhalb eines Datencenters kann bei nicht vorhandener Replikation in einem zeitlich eingeschränkten Zugang zu den Daten resultieren. Hierbei spielt es auch eine Rolle, inwiefern Mitarbeiter des CSP physischen Zugang zu den Daten haben und ob eine mögliche mutwillige Schädigung der Daten möglich ist[9]. Hierüber kann jedoch aufgrund des erwähnten, fehlenden Zugangs zu den Datenzentren der Anbieter keine Aussage getroffen werden, wenngleich alle Anbieter einen beschränkten Zugang durch ihre Mitarbeiter zusichern. Daher wird bei physischer Sicherheit analysiert, ob eine Replikation der Daten stattfindet und ob die Daten dabei an unterschiedlichen Standorten gehalten werden.

Im Bereich logischer Sicherheit wird untersucht, ob eine Verschlüsselung von Daten möglich ist, ob der Zugriff verschlüsselt erfolgt und auch, ob die Daten nach einem Vertragsende noch zugänglich sind. Der letzte Punkt ist vor allem unter dem Gesichtspunkt relevant, dass alle untersuchten Anbieter sich selbst das Recht einräumen, bei Verstoß gegen Vertragsvereinbarungen, den Vertrag sofort aufzulösen und dies ein mögliches Verlustrisiko mit sich bringt. Auch die Beurteilung darüber, ob ein Verstoß tatsächlich stattfand, liegt laut Verträgen der CSP im Allgemeinen in ihrem eigenen Ermessen.

Ein Nutzer eines CSP hat idealerweise zusätzlich die Option, Zugangsrechte zu den Daten auch für die eigenen Mitarbeiter einzurichten. Dafür wird von Anbietern eine Form von Zugangskontrolle bereitgestellt, mit der sich unterschiedliche Zugangsrechte verwalten lassen, wobei nur bestimmte Daten von bestimmten Nutzern einsehbar sind. Daher wird auch das Vorhandensein eines solchen Features in dieser Studie betrachtet.

Für die Angaben in dieser Studie ist entscheidend, ob die jeweilige Funktion auch explizit vertraglich gewährt ist. So erfüllt ein Anbieter das Kriterium der Replikation von Daten im Rahmen der vorliegenden Untersuchung nur dann, sofern diese explizit als Funktion angeboten wird.

Im Bereich logischer Sicherheit schneiden die Anbieter wie folgt ab: Eine serverseitige Verschlüsselung von Daten wird von allen Anbietern, mit Ausnahme von Rackspace, unterstützt. Eine verschlüsselte Übertragung per SSL ist mit allen Anbietern möglich.

[8] die vom BSI bereitgestellten Kataloge und Methoden zum IT-Grundschutz stellen eine Anleitung zur Sicherstellung gewisser Sicherheitsziele in der IT dar, zu finden unter `https://www.bsi.bund.de/DE/Themen/ITGrundschutz/itgrundschutz_node.html`, aufgerufen am 10.06.2013

[9] Studien belegen, dass unzufriedene Systemadministratoren eine nicht zu unterschätzende Bedrohung darstellen können und sich auch nicht der notwendigen Sicherheitsanforderungen vertraulicher Daten bewusst sind [13, 70]

Ein Nutzermanagement oder eine erweiterte Zugangskontrolle bietet nur Nirvanix nicht. Dieser Provider bietet lediglich einzelne Accounts ohne eine individuelle Rechteverwaltung. Alle anderen Dienstleister erlauben dies z. B. in Form von Access Control Lists[10]. An dieser Stelle sei noch zusätzlich erwähnt, dass Nirvanix als einziger Anbieter darauf hinweist, dass für die Virenfreiheit der eigenen Software nicht garantiert werden kann. Ein Zugriff auf die gespeicherten Daten nach Beendigung des Vertragsverhältnisses – dies kann auch eine Kündigung seitens des Anbieters sein – gewähren die meisten Anbieter, jedoch mit unterschiedlichen Fristen. Nirvanix und HP bieten 15 Tage Zeit, Amazon 30 Tage und Azure sogar insgesamt 120 Tage, in denen auf die Daten noch zugegriffen werden kann. Google macht keine Angabe zur Verfügbarkeit der Daten nach Vertragsende. Rackspace schließt den Zugriff nach Vertragsende explizit aus.

Im Bereich physischer Sicherheit weisen alle Provider ähnliche Merkmale auf. Alle Anbieter replizieren ihre Daten. Bei Rackspace ist diese Option zwar nur gegen Aufpreis möglich, erhöht aber gleichzeitig drastisch die Verfügbarkeitsgarantie (siehe Kapitel 4.3). Die Replikation findet bei allen Anbietern geo-redundant statt, auch innerhalb von ausgewählten Verfügbarkeitszonen von z. B. Amazon oder Google, d. h. die replizierten Daten befinden sich an physisch voneinander entfernten Orten. Alle Dienstleister garantieren hohe Sicherheitsvorkehrungen beim physischen Schutz ihrer Datenzentren. Laut öffentlich verfügbarer Angaben existieren durchweg Sicherheitspersonal rund um die Uhr, Videoaufzeichnungen, Zugangskontrollen und biometrische Scanner zur Identifizierung von autorisierten Mitarbeitern, sowie schlossgesicherte Serverschränke. Im Bereich physischer Sicherheit schneiden daher alle Anbieter durchweg positiv ab. Alle Ergebnisse im Bereich Sicherheit sind in Tabelle 4.2 zusammengefasst.

[10] im Allgemeinen eine Zugriffssteuerungsliste, mit welcher der Zugriff auf Daten/Funktionen für bestimmte Nutzer und Gruppen detailliert festgelegt werden kann

	Azure	**Amazon S3**	**Google**	**Rackspace**	**HP**	**Nirvanix**
Logische Sicherheit						
Datenverschlüsselung möglich	✓	✓	✓	–	✓	✓
Verschlüsselter Zugriff/Verbindung	✓	✓	✓	✓	✓	✓
Zugangskontrolle/-management für mehrere Benutzer	✓	✓	✓	✓	✓	✓
Datenzugriff nach Vertragsende	✓	✓	–	–	✓	✓
Physische Sicherheit						
Replikation möglich	✓	✓	✓	✓	✓	✓
Daten an unterschiedlichen Standorten	✓	✓	✓	✓	✓	✓

Tabelle 4.2: Providerübersicht hinsichtlich der *Sicherheit*

4.3 Verfügbarkeit

Die Verfügbarkeit der Daten ist eine essenzielle Anforderung an einen Cloud Storage-Provider, da durch eingeschränkte Erreichbarkeit der Systeme finanzielle Schäden für den Nutzer entstehen können. Grundsätzlich gilt, je höher die Verfügbarkeit, desto zuverlässiger und – je nach Anwendungsfall – geeigneter ist ein Dienst für einen Nutzer. So sind Content Delivery Netzwerke auf hohe Verfügbarkeit maßgeblich angewiesen, wohingegen Archivierungsdienste praktisch nur eine grundlegende Erreichbarkeit zu einem bestimmten, nicht aber jedem Zeitpunkt erfordern.

Bezogen auf Cloud Storage Dienste drückt Verfügbarkeit aus, dass jede Anfrage an Informationen zu jeder Zeit beantwortet werden kann. Die tatsächliche Verfügbarkeit lässt sich über nachvollziehbare Ausfälle messen. In diesem Zusammenhang bedeutet ein Ausfall, dass eine Anfrage nicht mit einer entsprechenden Antwort oder Ressource bedient werden kann. Dabei wird von Faktoren, die der CSP nicht beeinflussen kann, abstrahiert. Die Erreichbarkeit der entsprechenden Server, angegeben in prozentualer Verfügbarkeit für einen gegebenen Zeitraum, ist hier also ein relevantes Untersuchungskriterium. Zudem ist entscheidend, wie die Nichterreichbarkeit definiert wird, wann für den CSP nach Vertragsbedingungen auch tatsächlich ein Ausfall vorliegt und ob beispielsweise Wartungsarbeiten von den Garantien ausgeschlossen sind. Selbst wenn ein Ausfall nach Vertragsdefinition stattfindet, bleibt ein selbstgesteckter Kulanzbereich, den sich die CSP in Form der angegebenen Verfügbarkeit einräumen. Grundsätzlich gilt immer ein Mindestintervall oder eine Mindestfrequenz von Ausfällen, die erfüllt sein muss, damit es sich nach Vertragsdefinitionen auch tatsächlich um einen gutschriftrelevanten Ausfall handelt.

Wie bei physischer Sicherheit kann auch die Replikation der Daten eine Rolle bei der Verfügbarkeit spielen. So können replizierte Daten bei z. B. Ausfall oder Überlastung eines Standortes durch den Zugriff auf die Daten an einem anderen Standort auch weiter erreicht werden. Die implizite Replikation der gespeicherten Daten durch einen CSP ist daher ein Indikator für potenziell höhere Verfügbarkeit.

Für den Vergleich werden in erster Linie die Verfügbarkeitswerte untersucht, die von den Dienstanbietern zugesichert werden und wie diese definiert sind. Unter den untersuchten CSP garantieren Google, Rackspace und Amazon eine Verfügbarkeit von 99,9 % oder ca. 9 Stunden Ausfallzeit innerhalb eines Jahres. Dieses Level der Verfügbarkeit wird häufig auch in der Anzahl der garantierten Neunen angegeben, hier also drei Neunen. Zum Vergleich: Großrechner (Mainframes) garantieren eine Verfügbarkeit von bis zu 99,999 % (oder fünf Neunen), d. h. ca. 6 Min. Ausfallzeit innerhalb eines Jahres [71]. Eine solche Verfügbarkeit von bis zu 99,999 % wird zum Zeitpunkt der Erhebung der Studie nur von dem Anbieter Nirvanix garantiert. Azure und HP sichern 99,95 % (oder

dreieinhalb Neunen) zu, was einer Ausfallzeit von höchstens einer Stunde pro Jahr entspricht [71]. Bei der Definition eines Ausfalls und zugehöriger Kulanz schneiden alle Anbieter gleich ab. Die in den SLA beschriebenen Berechnungen sind durchweg nachvollziehbar. Grundlegender Unterschied ist aber die Berechnungsgrundlage, die entweder in der einfachen Laufzeit der Server, oder der tatsächlich stattfindenden Requests liegt. Weiterhin zählen nur bei Nirvanix und Rackspace geplante Wartungsarbeiten explizit nicht als Ausfall. Wie im Abschnitt zur Sicherheit (siehe Kapitel 4.2) bereits beschrieben ist, garantieren alle Anbieter eine Replikation der Daten. Damit kann davon ausgegangen werden, dass damit eine höhere Verfügbarkeit der Systeme auch technisch zugesichert werden kann.

	Azure	**Amazon S3**	**Google**	**Rackspace**	**HP**	**Nirvanix**
Garantierte Verfügbarkeit (in %)	99,95	99,9	99,9	99,9	99,95	99,999
Garantierte Verfügbarkeit (Anzahl 9en)	3,5	3	3	3	3,5	5
Transparente Definition eines Ausfalls	✓	✓	✓	✓	✓	✓
Replikation der Daten	✓	✓	✓	✓	✓	✓
Wartungsarbeiten zählen als Ausfall	✓	✓	✓	–	✓	–

Tabelle 4.3: Providerübersicht hinsichtlich der *Verfügbarkeit*

4.4 Kosten

Das Auslagern von Dateien und Anwendungen samt ihrer Anwendungsdaten kann gegenüber unternehmensinternen Lösungen Kostenvorteile im Bereich von bis zu 70 % bedeuten [66], wie dem Vergleich der on-premise Kosten in Abbildung 4.1 und der Kosten für Cloud Storage in Abbildung 4.2 zu entnehmen ist. Die Analysten zeigen zudem deutlich, dass die Kosten für interne Speicherstrukturen schnell unterschätzt werden können, wie die berechneten Kosten für on-premise Speicher in Abbildung 4.3 verdeutlichen. Gegenüber den reellen Kosten für on-premise Speicher (siehe Abbildung 4.1) wird deutlich, dass Kostenfaktoren wie Administration, Strom, Replikation und Reservoirs für zukünftige Speicheranforderungen einen erheblichen Anteil an den Gesamtausgaben für den internen Speichers darstellen und daher mit berücksichtigt werden müssen.

Figure 2 Traditional File Storage Systems Are Expensive To Buy And Run

Internal storage

	Assumptions	Calculations	
TB of actual data	100	$400,000	Acquisition cost of base amount of storage
Years expected lifespan of storage	4	300	Usable TB required including data copies
$/usable GB purchase price	$4	420	Total usable TB required for primary, copies, and utilization
Copies of data for redundancy	3	$1,680,000	Acquisition cost of total storage requirement
Typical utilization of storage (excluding RAID and system resource overhead)	60%	$420,000	Annualized storage acquisition cost based on lifespan
Typical TB/FTE	150	2.8	FTE requirement for storage admin of stated TB count
Fully loaded $/FTE	$120,000	$336,000	Annual storage admin cost
Facilities and power charge (of storage acquisition cost)	5%	$84,000	Annual facilities and power charge
Years of included warranty	3	$63,000	Annualized maintenance charge
Percent of original purchase price for additional warranty years	15%	$52,500	Annualized data migration charge
Cost per usable TB data migration	$500	**$955,500**	**Total annual cost of internal storage**

57696 Source: Forrester Research, Inc.

Abbildung 4.1:
Kosten für internen Speicher unter Einbezug aller Nebenkosten für Strom, Administration und Ähnlichem [66]

Darüber hinaus wird in Abbildung 4.2 gezeigt, dass auch bei Cloud Storage-Diensten zusätzliche Kosten zu den reinen Speicherkosten hinzu kommen können und deshalb der Umgang mit Speicherstrukturen und Unternehmensanforderungen genau berücksichtigt werden muss. Die nutzungsbasierte Abrechnung des Cloud Storage-Modells erfordert deshalb eine detaillierte Auseinandersetzung mit den Anforderungen an die Speicherstrukturen, um Kriterien wie Zugriffszahlen, Speichervolumen, notwendige Redundanzen

Figure 3 Cloud Storage Is More Straightforward And A Lot Cheaper Than Traditional Storage

Cloud storage

	Assumptions	Calculations	
TB of actual data	100	$11,800	Monthly cost of cloud storage
$/GB/month cloud charge	$0.118	$141,600	Annual cost of cloud storage
Months/year	12	$10,000	Total data-transfer-in charges
GB/TB	1,000	50	TB out/month
$/GB data-transfer-in rate	$0.1	$90,000	Annual data-transfer-out charges
Initial data in, assumed in annual cost	100%	$0	Additional redundancy capacity charges
$/GB data-transfer-out rate (simplified)	$0.15	$0	Cloud gateway annualized charge
Data out/month	50%	$20,000	Incremental annual network charge
Included copies of data for redundancy	3	$251,600	**Total annual cost of cloud storage**
Cloud gateway hardware/ software charge	$0		
Years expected lifespan of gateway	4		
Incremental annual network charge	20,000		

57696 Source: Forrester Research, Inc.

Abbildung 4.2:
Kosten für Cloud Storage unter Einbezug aller Nebenkosten für Strom, Administration und Ähnlichem [66]

und Aufbewahrungszeit zu erkennen und das passende Angebot zu wählen, sofern sich der Einsatz von Cloud Storage anbietet.

Die Preismodelle der angebotenen Dienste unterscheiden sich hierbei nicht gravierend hinsichtlich der reinen Speicherkosten, aber bezüglich der Kosten für Zugriffe, Datentransfer zwischen Cloud Storage und Dienstnutzern sowie zusätzliche Features wie Content Delivery-Netzwerke. Tabelle 4.4 gibt hierzu eine Übersicht über die zum Erstellungszeitpunkt der Studie aktuellen Preise der verschiedenen Dienste. Einige Anbieter staffeln ihre Preise hierbei je nach Region und genutztem Volumen, andere bieten einen konstanten Preis an. Die angegebenen Preise bezeichnen die jeweils günstigste Variante, gestaffelte Preise sind gesondert gekennzeichnet.

Figure 1 It's Easy To Get An Incomplete And Incorrect Comparison Of Cloud And Internal Storage

Cloud storage		Internal storage	
100 TB of actual data			
Assumptions		Assumptions	
$/GB/month cloud charge	$0.118	4	Years expected lifespan of storage
Months/year	12	$4	$/usable GB purchase price
GB/TB	1,000		
Calculations		Calculations	
Monthly cost of cloud storage	$11,800	$400,000	Acquisition cost of base amount of storage
Annual cost of cloud storage	$141,600	$100,000	Annualized over life of storage
Simple annual cost of cloud storage	**$141,600**	**$100,000**	**Simple annual cost of internal storage**

57696 Source: Forrester Research, Inc.

Abbildung 4.3: Kosten für internen Speicher ohne Betrachtung aller anfallenden Nebenkosten [66]

	Azure	Amazon S3	Google	Rackspace	HP	Nirvanix
Speicherkosten pro GB/Monat	0,037*	0,055*	0,054*	0,075*	0,09	0,25
Kosten ausgehende Daten pro GB	0,05*	0,5*	0,08*	0,12	0,05*	–
Kosten eingehende Daten pro GB	–	–	–	–	–	–
Kosten für 10.000 Anfragen (GET)	0,01	0,01*	0,01	–	0,01	–
Kosten für 100.000 Anfragen (GET)	0,01	0,10*	0,10	–	0,1	–
Kosten für 10.000 Anfragen (restliche)	0,01	0,10*	0,10	–	0,01	–
Kosten für 100.000 Anfragen (restliche)	0,01	1*	1	–	0,1	–
Transfer vom Speicher zum CDN pro GB	0,05*	0,02*	–	–	–	–

Tabelle 4.4:
Kostenübersicht für alle ausgewählten Dienste, alle Angaben in US Dollar
* Angabe beschreibt den niedrigsten zu erreichenden Preis; Preis ist abhängig von geografischer Region und genutztem Speicher- und Transfervolumen

4.5 Vertrauen, Zertifikate und Standards

Die Übertragung der Verantwortung für die eigenen Daten an einen Cloud Storage-Anbieter setzt ein gewisses Maß an Vertrauen gegenüber diesem voraus. Das Vertrauen der Nutzer kann anhand verschiedener Instrumente gestärkt werden: hohe Sicherheits- und Datenschutzanforderungen an die eigene Infrastruktur, Reputation, Zertifikate und Standards. Wie umfangreich die Datenschutz- und Sicherheitsmaßnahmen eines Anbieters sind, ist in den Bereichen *Recht* (Kapitel 4.1) und *Sicherheit* (Kapitel 4.2) näher dargestellt. Die Reputation eines Cloud Storage-Anbieters ist derzeit nur bedingt zu ermitteln. Zwar gibt es Internetportale, die eine Bewertung und einen Vergleich verschiedener Dienste anstreben[11], jedoch ist der Umfang, die Zuverlässigkeit der Angaben sowie die Aktualität dieser Portale begrenzt. Eine objektive Messung der Reputation eines Anbieters ist derzeit deshalb nur schwer zu ermitteln und wird in diesem Vergleich nicht betrachtet.

Bezüglich der verwendeten Zertifikate sieht es hierbei anders aus. Zertifizierungen stellen im Idealfall ein objektives Mittel zum Nachweis der Einhaltung standardisierter Richtlinien und somit auch zur Vertrauensbildung dar. Bisher gibt es jedoch nur wenige Zertifizierungen, die sich explizit im Cloud Umfeld bewegen. Ein Vorreiter in diesem Bereich ist *Euro Cloud SaaS Star Audit*. Über einen detaillierten Fragenkatalog wird hier die Einhaltung von Sicherheitsrichtlinien überprüft. Das Zertifikat steht noch am Anfang seiner Entwicklung, daher ist keiner der untersuchten Dienstleister Euro-Cloud-zertifiziert. Die Akzeptanz des Zertifikats wird sich in nächster Zeit zeigen müssen [55, 36]. Als Grundlage für alle Datenzentren dient im IT-Bereich die *ISO 27001* Zertifizierung. Das Bundesamt für Sicherheit in der Informationstechnik (BSI) erteilt diese Zertifizierung, die als eine der vertrauenswürdigsten einzustufen und weltweit als Standard anerkannt ist [55, 36]. Die Datenzentren aller untersuchten Anbieter sind nach ISO 27001 zertifiziert. Darüber hinaus gibt es jedoch nur wenige Zertifizierungen, die zur Vertrauensbildung im Cloud Storage beitragen. Hierzu zählt beispielsweise die *SAS 70* Zertifizierung, die vor allem im US-amerikanischen Raum häufig aufzufinden ist. Diese ist jedoch sehr aufwändig, schwer nachvollziehbar, in Europa nur bedingt anerkannt und zudem nicht Cloud-spezifisch [55, 36].

Neben Zertifikaten untersucht die Studie auch, inwiefern die Anbieter offene Standards einsetzen. Deren bewusste Unterstützung fördert die Interoperabilität bei verwendeten Produkten und Technologien auf der Nutzerseite. Kunden des Dienstleisters erhalten dadurch Investitionsschutz für ihre bestehende Infrastruktur sowie mehr Flexibilität im

[11] Beispiele für solche Portale sind `http://cloud-computing.findthebest.com/`, `http://www.cloudstoragereviews.org/comparison/` und `http://www.bestbackups.com/top-10-backup-providers/`, alle aufgerufen am 10.06.2013

Hinblick auf die Gestaltung ihrer IT-Landschaft. Die Migration zu einem anderen Dienst fällt dem Nutzer hier deutlich leichter als es bei Anbietern proprietärer APIs der Fall ist. Im Cloud Storage-Bereich gibt es laut aktueller Studien des BMWi[12] mittlerweile mehrere Standards [14]. Die Reife, Qualität und erwartete Durchsetzungsfähigkeit der verschiedenen Standards variiert dabei stark. Eine hohe Bewertung in diesen Bereichen wird hier den Standards *OAuth*, *OCCI* und *OpenStack* zugesagt, wie Abbildung 4.4 verdeutlicht. OpenStack wird durch Rackspace und HP unter den untersuchten Anbietern in somit zwei von sechs Fällen umgesetzt. Allgemein ist die Verbreitung von spezifischen Cloud Standards jedoch in etwa mit der der Zertifizierungen zu vergleichen. Etablierte Standards aus dem Bereich der Rechenzentren sind auch im Cloud Umfeld vertreten, spezifische Standards stehen aber noch am Anfang ihrer Entwicklung. Mit der Unterstürzung durch unter anderem HP und Rackspace zeigt OpenStack hier aber bereits eine sehr positive Entwicklung und Verbreitung.

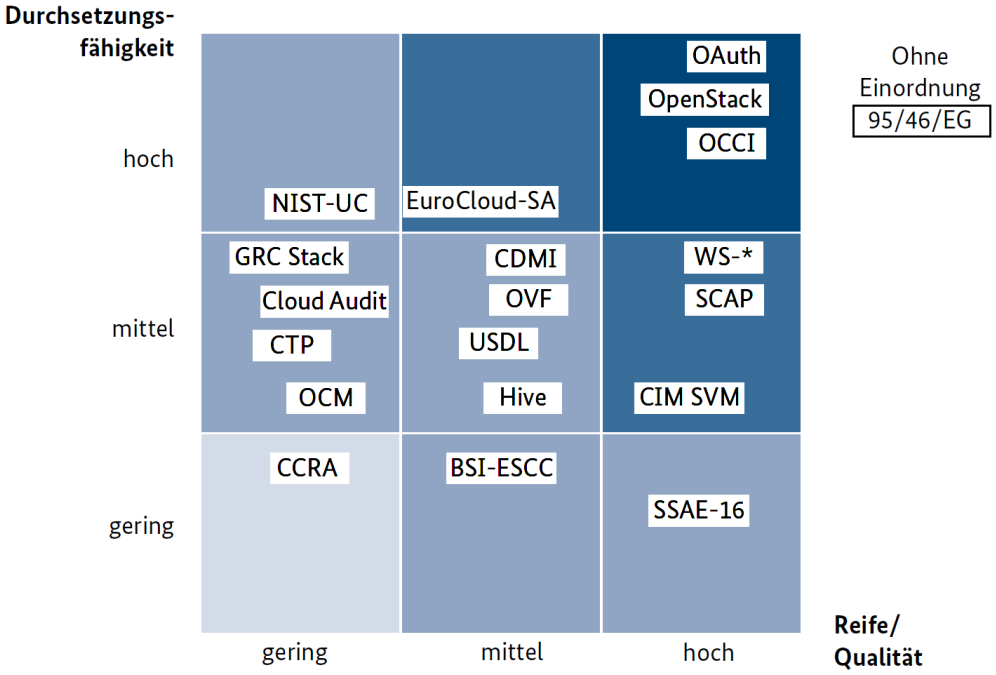


Abbildung 4.4: Einordnung der bewerteten 20 Cloud-Standards des BMWi [14]

[12] Bundesministerium für Wirtschaft und Technologie, zu finden unter https://www.bmwi.de/, aufgerufen am 10.06.2013

4.6 Zugriffsmöglichkeiten

Der Zugriff auf Speicherressourcen von BSPs erfolgt überwiegend über Programmierschnittstellen. Im Bereich der Zugriffsmöglichkeiten wird daher ein besonderes Augenmerk auf die bereitgestellten APIs sowie die verfügbaren Sprachbibliotheken gelegt. Darüber hinaus wird untersucht, ob die Dienste auch ein mobiles sowie ein Webinterface zur Datenverwaltung zur Verfügung stellen. Hierbei lässt sich feststellen, dass alle untersuchten Dienste sowohl eine REST API als auch ein Webinterface unterstützen. Alle Anbieter stellen zudem Bibliotheken für mindestens zwei gängige Programmiersprachen[13] bereit. Im Bereich der mobilen Geräte bieten derzeit Microsoft Azure und Amazon S3 APIs für iOS und Android an, während Rackspace Datenverwaltung über eine eigens entwickelte mobile App für Android und iOS ermöglicht.

Neben den angebotenen Schnittstellen und Interfaces ist es zusätzlich von Bedeutung, wie der Support beim Einbinden des Dienstes in die eigene Infrastruktur ausfällt. Das Bereitstellen von Beispielen in verschiedenen Programmiersprachen sowie detaillierte API Erläuterungen sind hierbei von Vorteil. Auch in diesen Punkten unterscheiden sich die Anbieter nur wenig voneinander. Alle APIs sind über eigene Webseiten dokumentiert, die Benutzung und Einbindung wird anhand von mehreren Beispielen erläutert. Auf alle Dokumentationen kann problemlos zugegriffen werden, auch wenn der Umfang und die Auffindbarkeit je nach Dienst und Gestaltung der verfügbaren Webseiten unterschiedlich ausfallen kann.

Ein weiterer Untersuchungspunkt ist die maximale Größe der Dateien, die übertragen und gespeichert werden können. Hierbei lassen sich große Unterschiede feststellen. Als einziger Anbieter legt Rackspace hier keine Restriktionen fest. Azure erlaubt eine maximale Dateigröße von 1 TB bzw. 200 GB (je nach Objekt, siehe Kapitel 6.1) und beschränkt einen Account, in dem Daten abgelegt werden können, zudem auf 100 TB. Sowohl Amazon als auch Google räumen 5 TB pro Objekt ein, während HP in seinem Object Storage Dateien bis zu maximal 5 GB Größe erlaubt, die mittels einer speziellen Meta-Datei aber zu einer nach außen hin größeren Datei verknüpft werden können. Bei Nirvanix gibt es keine Beschränkungen bezüglich der Dateigröße, allerdings sind die öffentlichen Accounts auf 2 TB beschränkt, womit auch die maximale Dateigröße in dieser Untersuchung auf 2 TB limitiert wird. Die Anbieter optimieren ihre Datenstrukturen auch hier für unterschiedliche Einsatzszenarien und der Nutzer muss hier genau auf seine eigenen Anforderungen achten. Kapitel 6 stellt die jeweiligen Datenmodelle der einzelnen Anbieter noch einmal genauer dar.

[13] als gängige Programmiersprachen betrachtet diese Studie beispielsweise die ersten 20 Vertreter des RedMonk-Rankings, siehe `http://redmonk.com/sogrady/2013/02/28/language-rankings-1-13/`, aufgerufen am 10.06.2013

Eine Besonderheit stellt die Übertragung von größeren Datenmengen im Bereich von mehreren Terabyte dar. Insbesondere bei der ersten Nutzung des Cloud Storage oder im Falle der Archivierung kann diese Funktion relevant sein. Die Dauer eines solchen Uploads kann, je nach Internetanbindung des Nutzers und der zu übertragenden Datenmenge, mehrere Tage oder Wochen in Anspruch nehmen. Selbst bei einer Anbindung von 10 MB/s, wie sie private Nutzer nur selten zu Verfügung haben[14], würde das Hochladen eines 10 TB großen Archivs mehr als 11 Tage in Anspruch nehmen und dabei die Internetanbindung auslasten. Jeder Ausfall der Verbindung während dieser Zeit kann zum Abbruch des Uploads führen und ein erneutes Hochladen notwendig machen. Die Cloud Storage Angebote von Amazon und Google bieten für dieses Szenario eine Alternative: das Einsenden von Festplatten, um große Datenmengen in den Cloud Storage zu laden oder von dort zu beziehen. Bei einer durchschnittlichen Lieferzeit über Logistikunternehmen von etwa 3 Tagen[15] und einer angegebenen Upload-Geschwindigkeit von etwa 85 MB/s[16] kann es sich schnell lohnen, die Daten auf dem postalischen Weg in die Cloud zu transferieren bzw. hinaus zu bekommen, denn auch der umgekehrte Weg ist möglich. Bei einer durchschnittlichen Internetleitung von 10 MB/s bietet das Angebot von Amazon und Google bereits ab 3 TB Zeit- und Kostenvorteile. Nach öffentlich zugänglichen Informationen wird diese Art des Datentransfers von den anderen untersuchten Anbietern nicht unterstützt.

Ein weiterer Untersuchungspunkt beim Zugriff auf die Daten sind die Operationen, die von den Diensten unterstützt werden und wie diese umgesetzt sind. Die grundlegenden Operationen zum Erstellen, Beziehen, Verändern und Löschen[17] von Daten und Metadaten[18] werden hierbei von allen Anbietern in ähnlichem Maße umgesetzt. Tabelle 4.6 gibt einen Überblick über die Operationen, die von der REST API der Provider unterstützt werden. In anderen Bereichen zeigen sich jedoch Unterschiede. So kann die Übertragung von mehreren Dateien über einen Kanal, wie von Google über *Batch Requests*[19] angeboten, die Transportzeit deutlich beschleunigen. Darüber hinaus wird das Umbenennen der Objekte nicht gleichermaßen unterstützt. Nur wenige APIs wie die von

[14] nachzulesen unter `http://www.tecchannel.de/netzwerk/news/2036610/breitbandabdeckung_durchschnittliche_bandbreite_von_17_mbit_s/`, aufgerufen am 10.06.2013

[15] basierend auf Informationen aus `http://www.deutschepost.de/content/dam/dpag/images/B_b/Briefe_ins_Ausland/pdf/laender_und_laufzeiten.pdf`, aufgerufen am 10.06.2013

[16] basierend auf Angaben des Amazon Import Rechners, siehe `http://awsimportexport.s3.amazonaws.com/aws-import-export-calculator.html`, aufgerufen am 10.06.2013

[17] oft auch als CRUD Operationen für *Create*, *Retrieve*, *Update*, *Delete* bezeichnet

[18] Metadaten bezeichnen Schlüssel-Wert-Paare, die Datenobjekte genauer beschreiben; hierzu zählt unter anderem der Hash-Wert eines Datenobjekts

[19] in diesem Fall werden mehrere Anfragen in eine Anfrage verpackt, wodurch beispielsweise nur eine einmalige Authentifizierung notwendig ist für den Upload mehrerer Dateien; es muss nur eine Verbindung zum Cloud Storage aufgebaut werden, was die Übertragungszeit erheblich verkürzt

	Azure	Amazon S3	Google	Rackspace	HP	Nirvanix
Zugriff über RESTful API	✓	✓	✓	✓	✓	✓
Zugriff über Web Interface	✓	✓	✓	✓	✓	✓
SDKs für mind. 2 gängige Programmiersprachen	✓	✓	✓	✓	✓	✓
SDKs für mobile Geräte	✓	✓	–	–	–	–
mobile Applikation	–	–	–	✓	–	–
Beispiele und API Dokumentation	✓	✓	✓	✓	✓	✓
maximale Dateigröße (in TB)	100	5	5	∞	0.005	2
Festplatten einsenden möglich	–	✓	✓	–	–	–

Tabelle 4.5: Providerübersicht hinsichtlich der *Zugriffsmöglichkeiten*

Nirvanix oder Tools wie z. B. *gsutil*[20] unterstützen das direkte Umbenennen einer Datei im Cloud Storage. Viele Schnittstellen erfordern das Kopieren des Objekts unter neuem Namen direkt im Cloudspeicher und das anschließende Löschen des alten Objekts. Das Umbenennen erfordert daher in vielen Fällen manuellen Programmieraufwand, zusätzliche API-Zugriffe und einen temporären Anstieg des verbrauchten Speichervolumens, wodurch zusätzliche Kosten anfallen können.

Die unterstützten Operationen der Anbieter unterliegen zudem ständigen Veränderungen. Darüber hinaus legen die Anbieter unterschiedliche Schwerpunkte bei den angebotenen Operationen. Auf der einen Seite bietet beispielsweise Google den parallelen Upload sowie wiederaufnehmbare Up- und Downloads für Dateien an, um Benutzern beim Erstellen von Daten möglichst viel Komfort zu bieten. Auf der anderen Seite ermöglicht Nirvanix das direkte Audio- und Videotranscodieren sowie die Bildgrößenanpassung für Daten, die im Cloud Storage liegen, ohne diese zunächst herunterladen zu müssen. Auch die anderen untersuchten Anbieter bieten hier Operationen, die sie unter der Konkurrenz hervorheben sollen. Der Nutzer kann auch hier nach den für ihn passenden Vorteilen abwägen. Wie bereits erwähnt, werden die Grundfunktionalitäten von allen untersuchten Anbietern in ähnlichem Maße abgedeckt.

[20] angeboten von Google, zu finden unter `https://developers.google.com/storage/docs/gsutil`, aufgerufen am 10.06.2013

	Azure	**Amazon S3**	**Google**	**Rackspace**	**HP**	**Nirvanix**
Auflisten aller Ordner eines Accounts	✓	✓	✓	✓	✓	✓
Erstellen eines neuen Ordners in Account	✓	✓	✓	✓	✓	✓
Erstellen eines neuen Unterordners	–	–	–	–	–	✓
Löschen eines leeren Ordners	✓	✓	✓	✓	✓	✓
Löschen eines Ordners einschließlich Inhalt	✓	–	–	–	–	✓
Kopieren eines Ordners	–	–	–	–	–	✓
Anzeigen der Metadaten eines Ordners	✓	✓	✓	✓	✓	✓
Verändern der Metadaten eines Ordners	✓	✓	✓	✓	✓	✓
Auflisten aller Dateien eines Ordners	✓	✓	✓	✓	✓	✓
Erstellen einer neuen Datei in Ordner	✓	✓	✓	✓	✓	✓
Konkatenieren aller Dateien eines Ordners	–	–	✓	–	–	–
Herunterladen einer Datei	✓	✓	✓	✓	✓	✓
Löschen einer Datei	✓	✓	✓	✓	✓	✓
Löschen mehrerer Dateien	–	✓	–	✓	–	✓
Anzeigen der Metadaten* einer Datei	✓	✓	✓	✓	✓	✓
Verändern der Metadaten* einer Datei	✓	✓	✓	✓	✓	✓
Kopieren einer Datei in Zielordner	✓	✓	✓	✓	✓	✓
Verschieben einer Datei in Zielordner	–	–	–	–	–	✓
Umwandeln einer Mediendatei (Audio oder Video)	–	–	–	–	–	✓

Tabelle 4.6: Providerübersicht hinsichtlich der *Zugriffsmöglichkeiten*

4.7 Performance

Die Vorhersagbarkeit der Leistungsfähigkeit, die zuverlässige Antwortzeit und die Durchsatzrate eines Anbieters können entscheidende Faktoren bei der Wahl eines geeigneten Dienstanbieters sein. Sie sind insofern relevant, als die Werte zur Einschätzung der Ausführungsdauer IT-gesteuerter Prozesse genutzt werden können. Nutzer erwarten hohe Verfügbarkeit und Skalierbarkeit bei vergleichsweise geringen Kosten als essentiellen Vorteil von Cloud Computing. Die Technologie kann den Nutzern einzigartige Vorteile für diverse Anwendungsfälle bringen. Jedoch ist auch eine konstante Zuverlässigkeit der Performance für viele Anwendungen unentbehrlich. So müssen z. B. Anbieter multimedialer Inhalte oder auch nur einfacher Websites in kürzester Zeit ihre Nutzer oder Kunden mit Daten beliefern können. Der Anwendungsfall *Content Delivery* erfordert eine niedrige Antwortzeit unabhängig von dem Standort des Nutzers. Auch im Forschungsumfeld spielt Leistungsvorhersage eine wichtige Rolle, damit beispielsweise die Dauer der Experimente eingeschätzt werden kann [68], bzw. diese wiederholbar sind. In den Szenarien, die in Kaptiel 3 vorgestellt wurden, werden eher hohe Nutzerzahlen erwartet. Dies zieht viele Transaktionen auf einem Account nach sich und erfordert eine hohe Stressbelastbarkeit der Anbieterhardware [64].

Trotz der enormen Wichtigkeit der Vorhersagbarkeit der Leistung eines Dienstes lassen sich zugesicherte Werte bezüglich der Performance eines Dienstes nicht in den SLA der Provider finden [68]. Dies lässt sich von Seiten der Anbieter damit begründen, dass die Übertragungsleistung maßgeblich von dem physischen Standpunkt der Infrastruktur sowie der Güte der Internetanbindung der an der Kommunikation beteiligten Parteien (CSP, Dienstnutzer) beeinflusst wird. Es gilt, je weiter ein Nutzer von der Anbieterhardware entfernt ist, desto geringer können Übertragungsraten ausfallen. Bereits bei vergleichsweise kleinen Datenpaketen von wenigen hundert Megabyte können die Unterschiede in den Transferzeiten bei fast einer Stunde liegen (siehe Kapitel 5). Zusicherungen zur Leistungsfähigkeit können aus Anbietersicht daher nur schwer in das SLA aufgenommen werden.

Szenarien wie Content Delivery erfordern darüber hinaus parallele Zugriffe auf Daten oder benötigen eine hohe Anzahl von Verbindungen durch verschiedene Nutzer weltweit. Dies sollte die Performance im Idealfall nicht einschränken. Die Untersuchungspunkte im Bereich *Performance* sind daher die Antwortzeiten und Durchsatzraten der ausgewählten Anbieter sowie die Stabilität und das Verhalten bei Parallelzugriffen und hohem Verbindungsaufkommen.

Im Rahmen dieser Studie wurde die Leistungsfähigkeit ausgewählter CSPs für den Standpunkt des Hasso-Plattner-Instituts in Potsdam experimentell ermittelt. Die Testergebnisse sind hierbei abhängig von der Internetanbindung, dem Standort und der

Methodik. Die erzielten Ergebnisse sind daher nicht allgemein gültig, vermitteln jedoch einen guten Eindruck von der Leistungsfähigkeit der untersuchten Anbieter in vergleichbaren Fällen. In diesem Abschnitt werden die Ergebnisse der durchgeführten Messungen übersichtlich vorgestellt. Eine ausführliche Erklärung des Experimentaufbaus, der verwendeten Methodik und den detaillierten Ergebnissen findet sich in Kapitel 5.

Zusammenfassend lässt sich in den Performanceuntersuchungen feststellen, dass die Anbieter sich teilweise sehr stark unterschieden. Die Ergebnisse zeigen, dass die Schnittstellen der Provider für unterschiedliche Dateigrößen und Einsatzfälle optimiert sind. Während einige Anbieter sich auf die schnelle Übertragung relativ kleiner Dateien spezialisiert haben, liefern andere auch bei der Übertragung großer Dateien vergleichsweise gute Werte. Wie die Experimente belegen, lassen sich aber auch bei den langsameren Providern besonders große Dateien beispielsweise durch Segmentierung und parallele Transfers performant übertragen. Der Nutzer hat hier also die Wahl zwischen Anbietern, die für verschiedene Anwendungsfälle Vorteile mitbringen, in anderen Szenarien jedoch deutlich schlechter als die Konkurrenz abschneiden. Weitere Messungen und deren Auswertungen finden sich in Kapitel 5.

4.8 Übersicht

Nachdem alle Kriterien in den vorherigen Abschnitten im Detail vorgestellt wurden, dient dieser Abschnitt der übersichtlichen Aufarbeitung der beschriebenen Provider und ihrer Funktionalitäten. Tabelle 4.7 stellt zusammenfassend die Eigenschaften aller ausgewählten Anbieter dar. Als Grundlage dienen die beschriebenen Untersuchungspunkte der Bereiche Recht, Sicherheit, Verfügbarkeit und Zugriffsmöglichkeiten. Die Themen Kosten, Zertifikate, Standards und Vertrauen werden aus der Gründen der vorherrschenden Homogenität unter den Providern sowie der Übersichtlichkeit hier nicht aufgeführt. Eine Übersicht über die Leistungsfähigkeit der untersuchten Dienste wird in Kapitel 5 dargestellt.

	Azure	Amazon S3	Google	Rackspace	HP	Nirvanix
Recht						
mind. 1 Standort in der EU	✓	✓	✓	✓	–	✓
Standort frei wählbar	✓	✓	✓	✓	✓	✓
Teilnehmer am Safe Harbor Agreement	✓	✓	✓	✓	✓	✓
Vertragsstrafen bei Nichterfüllung des SLA	✓	✓	✓	✓	✓	✓
Eigentumsrecht an den Daten	✓	✓	✓	✓	–	✓
Insolvenzfall geklärt	–	–	–	–	–	–
Sicherheit						
Datenverschlüsselung möglich	✓	✓	✓	–	✓	✓
Verschlüsselter Zugriff/Verbindung	✓	✓	✓	✓	✓	✓
Zugangskontrolle/-management für mehrere Benutzer	✓	✓	✓	✓	✓	✓
Datenzugriff nach Vertragsende	✓	✓	–	–	✓	✓
Replikation möglich	✓	✓	✓	✓	✓	✓
Daten an unterschiedlichen Standorten	✓	✓	✓	✓	✓	✓
Begrenzter Zugriff für Mitarbeiter des CSP	✓	✓	✓	✓	✓	✓
Verfügbarkeit						
Garantierte Verfügbarkeit (Anzahl der 9en)	3,5	3	3	3	3,5	5
Transparente Definition eines Ausfalls	✓	✓	✓	✓	✓	✓
Replikation der Daten	✓	✓	✓	✓	✓	✓
Wartungsarbeiten zählen als Ausfall	✓	✓	✓	–	✓	–
Zugriffsmöglichkeiten						
Zugriff über RESTful API	✓	✓	✓	✓	✓	✓
Zugriff über Web Interface	✓	✓	✓	✓	✓	✓
SDKs für mind. 2 gängige Programmiersprachen	✓	✓	✓	✓	✓	✓
SDKs für mobile Geräte	✓	✓	–	–	–	–
mobile Applikation	–	–	–	✓	–	–
Beispiele und API Dokumentation	✓	✓	✓	✓	✓	✓

Tabelle 4.7: Providerübersicht hinsichtlich Recht, Sicherheit, Verfügbarkeit und Schnittstellen

5 Performanz

5.1 Ausgangslage

Das Experiment wurde am Hasso Plattner Institut (HPI) in Potsdam über einen Zeitraum von 480 Stunden im November 2012 durchgeführt. Darüber hinaus wurden einzelne Testreihen in April, Mai und Juli 2013 erneut ausgeführt, um die Aktualität der Ergebnisse zu verifizieren. Durch die ausgedehnte Laufzeit des Experiments kann davon ausgegangen werden, dass die Ergebnisse der Messungen geographisch verteilte Peaks zu unterschiedlichen Tageszeiten enthalten. Ferner verfügt das HPI über eine High-Speed Internetverbindung (1 GB), sodass während des Experiments am Testsystem selbst keine Flaschenhalseffekte beobachtet wurden.

5.1.1 Versuchsanordnung

Die Testanordnung wird insgesamt 10 Mal ausgeführt. Jeder Durchlauf besteht aus einer Menge vordefinierter Testkonfigurationen (im Folgenden Sequenzen genannt). Zwischen den einzelnen Anfragen wird ein Timeout von einer Sekunde festgelegt. Im Falle einer fehlgeschlagen Anfrage wird nach der definierten Pause der Aufruf wiederholt. Beim dritten (aufeinanderfolgenden) fehlgeschlagenen Anlauf wird die Anfrage als gescheitert angesehen und mit einem entsprechenden Vermerk protokolliert.

5.1.2 Zielstellung

Im Rahmen des Experiments wurden folgende Aspekte untersucht:

- Die allgemeinen Transferzeiten, die Anbieter für Dateien unterschiedlicher Größe erreichen. Hierzu werden Dateien mit den Größen 100 KB, 500 KB, 1 MB, 10 MB, 100 MB und 1 GB als Einzeldateien nacheinander zu den einzelnen Anbietern übertragen. Das Messergebnis ergibt sich aus der Zeitspanne zwischen dem Zeitpunkt des Starts und dem Zeitpunkt der Fertigstellung des Datenübertragung.
- Die Anzahl (simultaner) Anfragen, die von einem Dienstanbieter innerhalb einer Minute bearbeiten werden kann. Hierzu soll eine steigende Anzahl paralleler Anfragen an einzelne Anbieter gesendet werden.

- Der Grad gegenseitiger Beeinflussung einzelner Threads bei parallelen Übertragungen. In jedem Durchlauf wird am Testsystem n ein Datensegment fester Größe ausgewählt und an einen ausgewählten Dienstanbieter übertragen. Hierbei wird die durchschnittliche Übertragungszeit einzelner Transferoperation (Lesen und Schreiben) gemessen.
- Das Potential einer Übertragungssteigerung durch vorangehende Datensegmentierung. Hierbei wird die Gesamtdauer der Übertragung beginnend mit dem ersten Segment bis zum letzten Segment gemessen.

5.1.3 Metriken

Das Ziel des Experiments ist somit die Untersuchung verschiedener Leistungseigenschaften ausgewählter Cloud Speicher Anbieter. Dabei geht es zum Einen um die generelle Übertragungsleistung, zum Anderen um das Verhalten der Cloud-Provider bei vielen offenen parallelen Verbindungen und hoher Datenlast. Im Einzelnen werden hierbei Übertragungszeiten, Antwortzeiten, Belastbarkeit und Verfügbarkeit der APIs untersucht.

Übertragungszeit Allgemein berechnet sich die Übertragungs- oder Transferzeit aus der Zeitspanne zwischen dem Start und dem Ende der Datenübertragung. Da diese je nach Auslastung der Leitung unterschiedlich ausfallen kann, werden die Messungen mit einem festen Testdatensatz wiederholt ausgeführt und anschließend gemittelte Messwerte verwendet.

Antwortzeit Grundsätzlich hängt die Messung der Antwortzeit (Latenz) von dem Aufbau und der Lage des Netzwerks, sowie der Auslastung der Leitung ab. Antwortzeit ist hier als die Zeitverzögerung definiert, die beim Anbieter bei der Reaktion auf einen API-Zugriff auftritt. Angewendet auf das Experiment muss also die Zeit gemessen werden, die zwischen dem Start einer Datenübertragung über die jeweilige API und dem Empfang der ersten Bytes der Datei vergeht. Die Implementierung eines solchen Verfahrens erfordert einen hohen Aufwand. Stattdessen wurde die Entscheidung getroffen, die Reaktionszeit eines Dienstes durch die Beantwortung einer *getHash*-Anfrage (als Referenzwert) zu verwenden. Hierbei wird angenommen, dass bei allen Anbietern der Verarbeitungsprozess eingehender Anfragen identisch ist.

Belastbarkeit Belastbarkeit betrachtet die Fähigkeit eines Systems (Netzwerk, Service, Infrastruktur, etc.) eine akzeptable Dienstgüte trotz auftretender Fehler oder Störungen bereit zu stellen. In dieser Studie untersuchen wir im Speziellen i) das Verhalten ausgewählter Dienstleister bei parallel ausgeführten Anfragen sowie die entsprechenden Antwortzeiten, ii) die Kontinuität der Performance bei simultanen

Datenübertragungen und iii) die Zuverlässigkeit der Provider bei anhaltenden parallelen Operationen über einen längeren Zeitraum. Für den Test wird eine Reihe simultaner Lese- und Schreiboperationen mit Datensegmenten unterschiedlicher Größe (1 MB, 10 MB, 100 MB, 1 GB) durchgeführt.

Verfügbarkeit Üblicherweise bezeichnet Verfügbarkeit das Verhältnis der Uptime eines Servers zur Summe seiner Uptime und Downtime. Wie in Kapitel 4 beschrieben, wird bei den untersuchten Anbietern die Verfügbarkeit (bzw. der Ausfall eines Dienstes) unterschiedlich definiert. In der vorliegenden Studie beschreibt die Verfügbarkeit das Verhältnis erfolgreicher Anfragen zur Summe aller gesendeten Anfragen an das jeweilige System.

5.2 Ergebnisse

5.2.1 Übertragungszeit

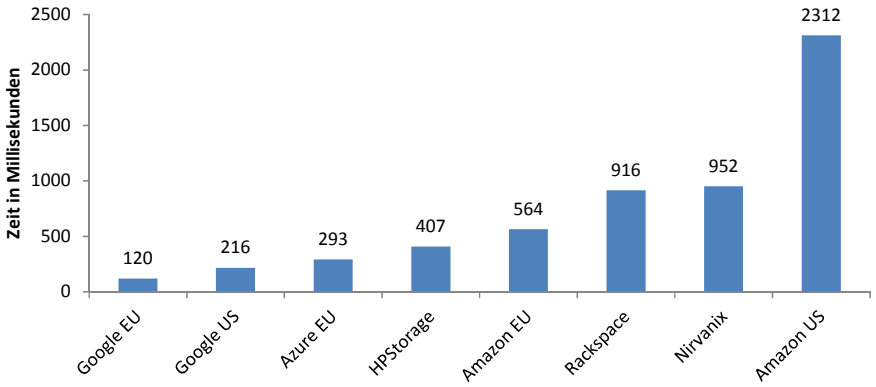
Die Messung der Übertragungszeit stellt das erste grundlegende Maß im Vergleich der Provider dar. Hierbei sollen die allgemeinen Übertragungszeiten ermittelt werden. Grundsätzlich kann festgestellt werden, dass die Güte der Dienste hinsichtlich der Performance sehr verschieden ausfällt. Besonders bei Down- und Uploadprozessen mit unterschiedlichen Dateigrößen zeigen sich große Differenzen (vlg. Abbildungen 5.1 bis 5.4). Nahezu alle getesteten Dienste zeigen im Download deutlich bessere Durchsatzwerte gegenüber dem Upload. Dieses Verhalten lässt sich damit erklären, dass die im Rahmen der Untersuchung betrachteten Anbieter die Schnittstellen ihrer Dienste für bestimmte Anwendungsfälle optimiert haben. So kann beispielsweise der Schreibvorgang einer 100 kB Datei bei dem Dienst Google US oder Google EU ca. 12 bis 19 mal länger dauern als ein Lesevorgang. Das Verhalten lässt sich auch bei Übertragung größerer Dateien beobachten, wenn auch nicht mehr mit dermaßen drastischen Schwankungen. Hier unterscheidet sich die Performance um das vier- bis fünffache mit Ausnahme von dem Anbieter Rackspace, bei dem ein Schreibzugriff bis zu 19 bzw. 49 mal langsamer ausfallen kann als ein Lesezugriff (abhängig von der Dateigröße).

Andererseits konnte auch ein Zusammenhang zwischen der Übertragungsleistung und der Datengröße beobachtet werden. So erzielen bei der Übertragung kleinerer Dateien (bis 1 MB) die Anbieter Azure und Amazon EU (beim Upload) die besten Ergebnisse. Bei 100 MB fällt der Anbieter Amazon EU zurück. Ähnlich verhält sich auch der Dienst Google US. Allerdings verbessern sich hier die Performancewerte mit wachsender Dateigröße.

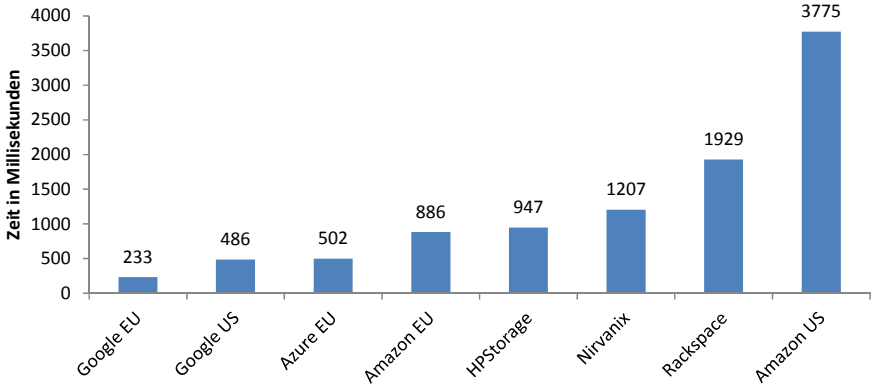
Das beobachtete Verhalten kann damit zusammenhängen, dass die relativ große Reaktionszeit des Dienstes (welche in erster Linie auf die lange Entfernung zwischen unserem Testsystem und dem Zielknoten zurückzuführen ist) bei größeren Dateien immer weniger die Messung der Übertragungsdauer beeinflusst.

Noch deutlicher lässt sich jedoch dieses Verhalten bei dem Dienst Google EU beobachten. Bei der Übertragung kleinerer Dateien zeigt der Dienst eine relativ schlechte Performance im direkten Vergleich zu anderen Anbietern. Mit wachsender Dateigröße rückt der Dienst jedoch immer weiter nach Vorne, bis er schließlich bei der Übertragung von 100 MB Dateien die Spitzenposition erreicht und damit den Dienst Azure ablöst.

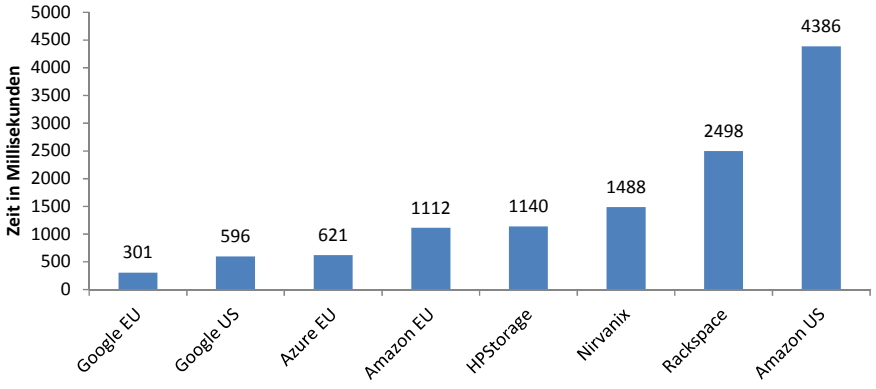
Ähnliche Zusammenhänge lassen sich auch beim Download beobachten. Bei der Übertragung kleinerer Dateien, gehört der Anbieter Azure zu den führenden Diensten. Mit wachsender Dateigröße fällt Azure jedoch immer weiter zurück. Auch der Dienst Rackspace zeigt bei kleineren Dateien wesentlich bessere Performance. Bei der Übertragung von Dateien mit einer Größe von 100 kB gehört der Dienst noch zu den führenden Anbietern. Bereits bei einer Dateigröße von 500 kB fällt der Dienst auf einen der letzten Plätze zurück.



Dateigröße 100 KB

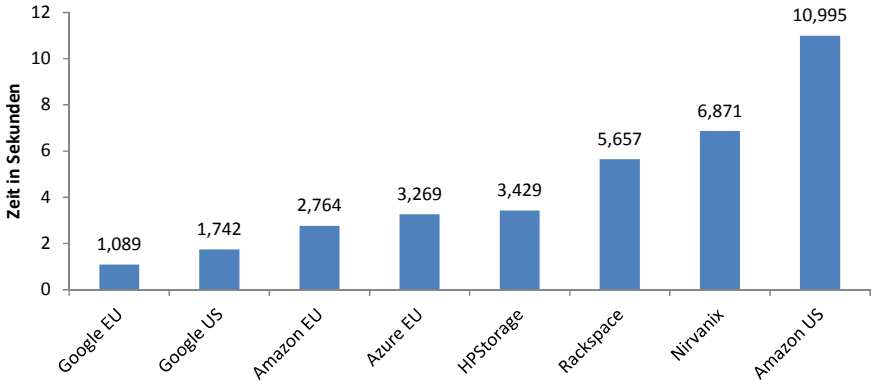


Dateigröße 500 KB

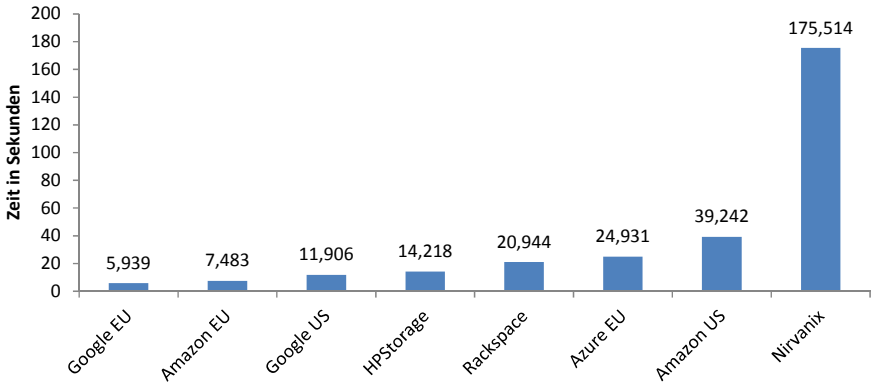


Dateigröße 1 MB

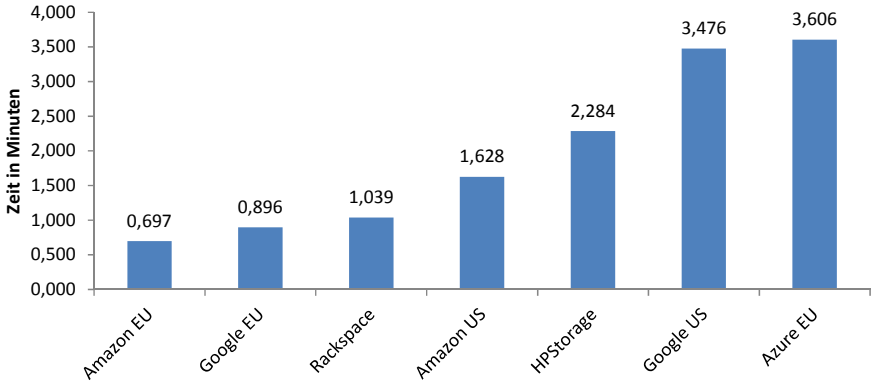
Abbildung 5.1:
Transferzeiten für Downloads der unterschiedlichen Anbieter hinsichtlich verschiedener Dateigrößen von 100 KB bis 1 MB.



Dateigröße 10 MB

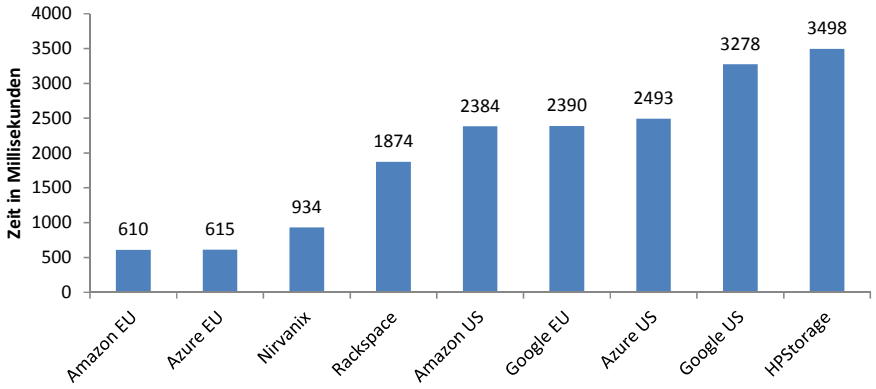


Dateigröße 100 MB

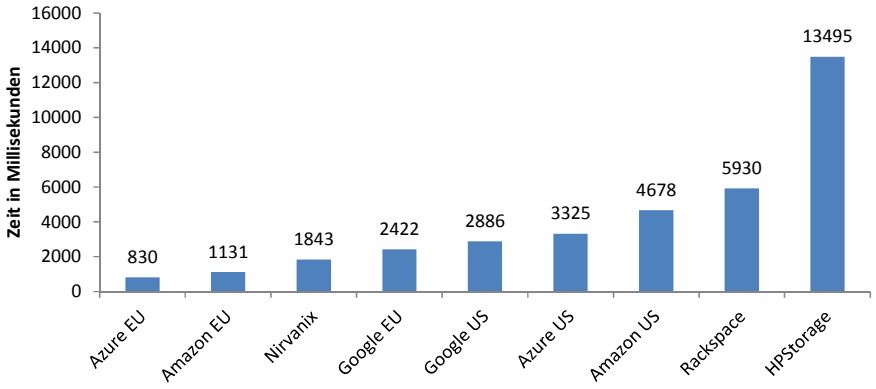


Dateigröße 1 GB

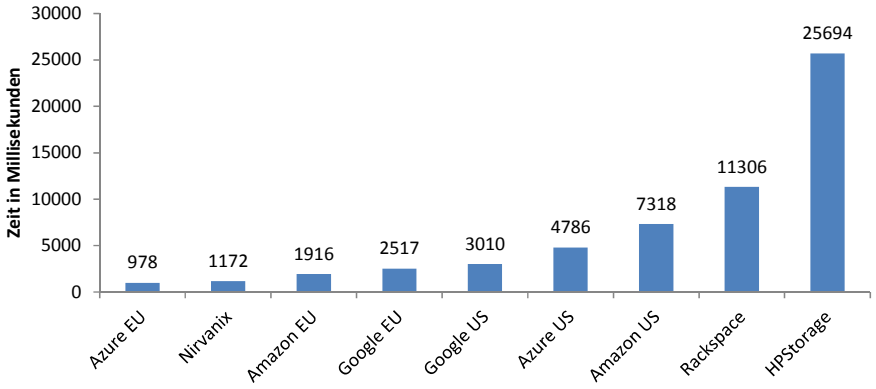
Abbildung 5.2:
Transferzeiten für Downloads der unterschiedlichen Anbieter hinsichtlich verschiedener Dateigrößen von 10 MB bis 1 GB.



Dateigröße 100 KB

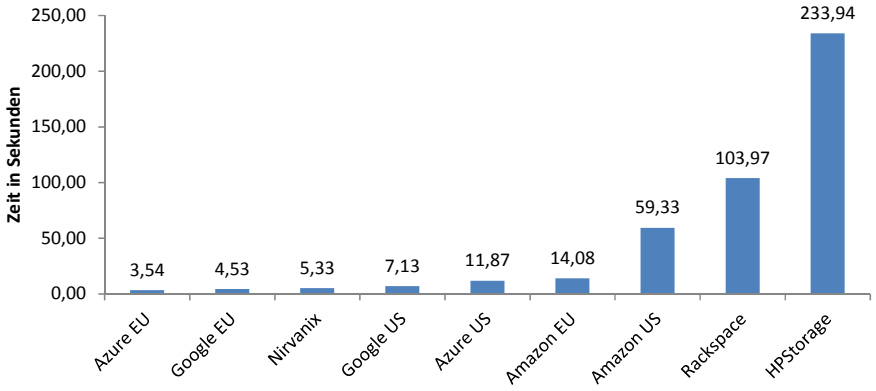


Dateigröße 500 KB

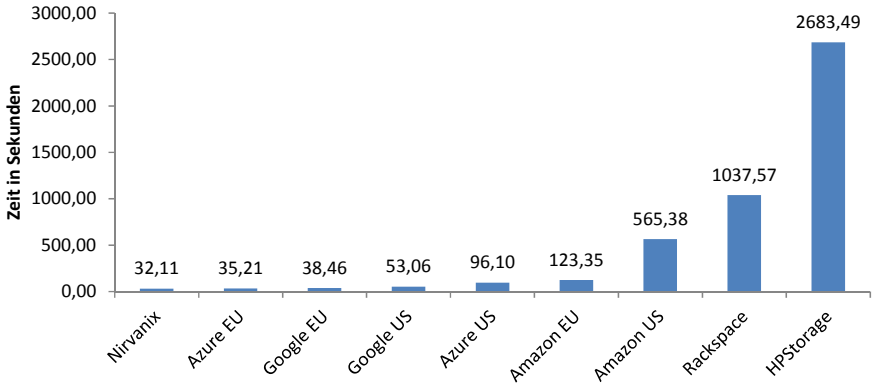


Dateigröße 1 MB

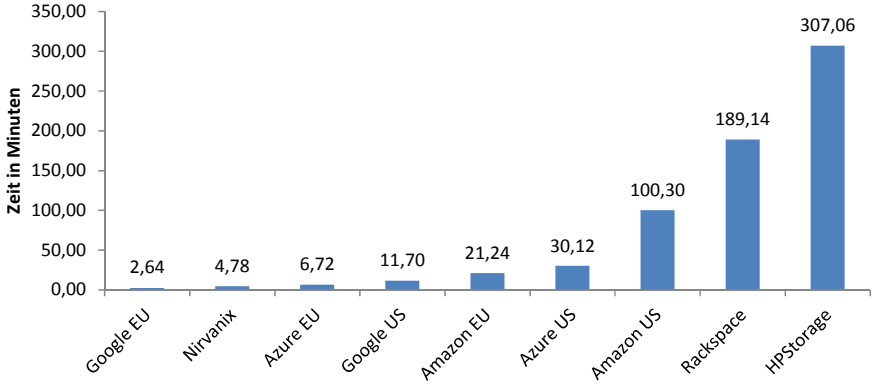
Abbildung 5.3:
Transferzeiten für Uploads der unterschiedlichen Anbieter hinsichtlich verschiedener Dateigrößen von 100 KB bis 1 MB.



Dateigröße 10 MB



Dateigröße 100 MB



Dateigröße 1 GB

Abbildung 5.4:
Transferzeiten für Uploads der unterschiedlichen Anbieter hinsichtlich verschiedener Dateigrößen von 10 MB bis 1 GB.

5.2.2 Antwortzeit

Die Vorhersagbarkeit der Leistung eines Online-Speicher Anbieters beinhaltet neben der Übertragungszeit auch die Reaktionszeit der Dienste. Aus diesem Grund wurde im Rahmen der Untersuchung die Antwortzeit ausgeführter Anfragen beobachtet. Anhand der Ergebnisse können die untersuchten Anbieter (für den Standort HPI) in folgende Kategorien eingeteilt werden:

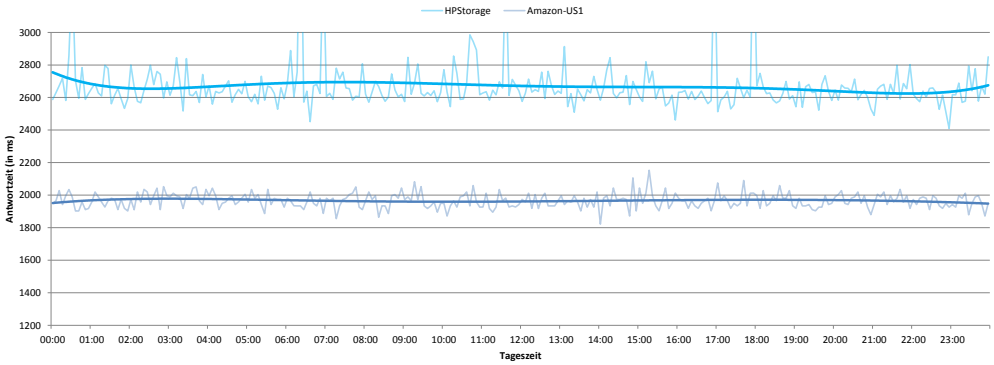
- Schnell (Antwortzeit liegt unter 200 ms);
- Mittelmäßig (Antwortzeit variiert zwischen 200 und 1500 ms); und
- Langsam (Antwortzeit liegt über 1500 ms)

Abbildung 5.5 zeigt, wie schnell Anbieter auf eine *getHash* Anfrage reagieren. Zu unterschiedlichen Zeitpunkten des Experiments konnte eine erhöhte Reaktionszeit beobachtet werden, die sich auf plötzlich erhöhten Datenverkehr auf dem Zielserver zurückführen lassen kann (vergl. Google (USA) in Abbildung 5.5).

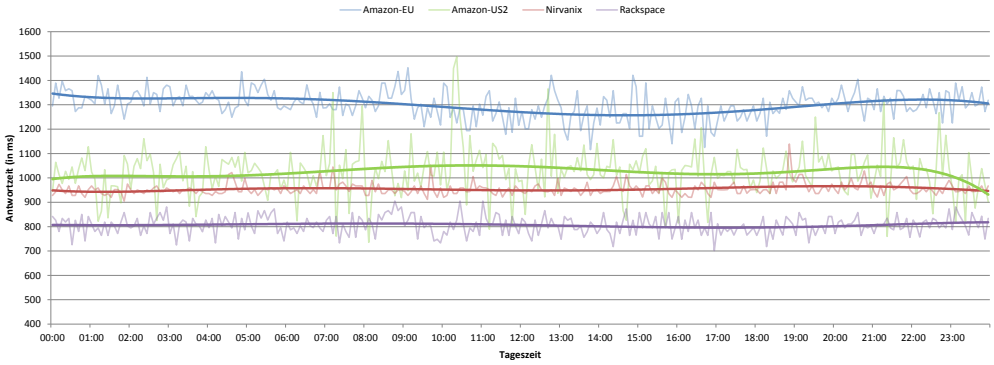
Insgesamt liefert Azure hier die besten und beständigsten Resultate für unseren Standort. Im Durchschnitt benötigt der Dienst etwa 50 ms für die Antwort auf eine Anfrage. Die längste Reaktionszeit benötigt der Dienst HP (USA). Dies lässt sich leicht auf die große Entfernung zwischen den Testservern in Deutschland (am HPI) und den Zielservern in den USA zurückführen.

Die größten Abweichungen in den Reaktionszeiten lassen sich bei den Dienstanbietern Google und Amazon beobachten. Hier können die Abweichungen bis zu 100 ms betragen. Über die Gründe für dieses Verhalten kann nur spekuliert werden. Eine mögliche Erklärung ließe sich darin finden, dass die Provider auf den Knoten mehrere Cloud Dienste betreiben.

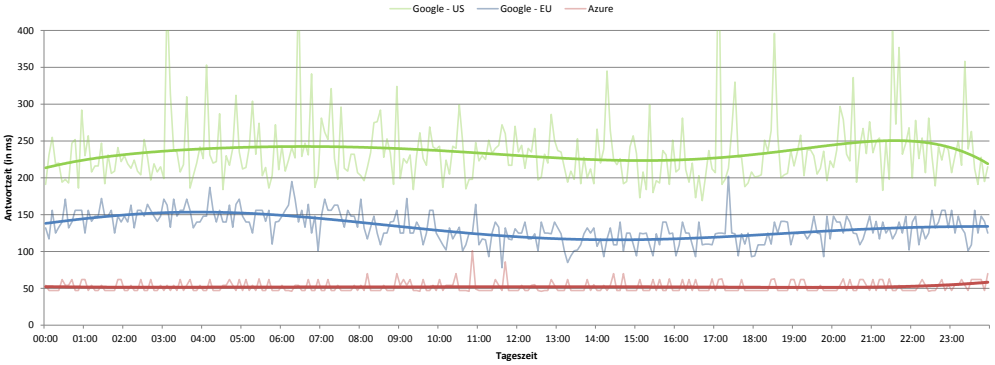
Die Ergebnisse in Abbildung 5.5 erlauben verlässliche Aussagen über die zu erwartenden Reaktionszeiten der untersuchten Anbieter (für den Standort HPI), da sie auf kontinuierlichen Messungen von über 20 Tagen beruhen.



HP, Amazon US 1



Amazon EU, Amazon US 2, Nirvanix, Rackspace



Google US, Google EU, Azure

Abbildung 5.5:
Antwortzeiten der untersuchten Anbieter auf eine *getHash* Anfrage zu unterschiedlichen Tageszeiten.

5.2.3 Belastbarkeit

Im Unternehmenskontext sind Anwendungen oft parallelen Zugriffen ausgesetzt. Aus diesem Grund untersuchten wir auch die CSPs hinsichtlich der Fähigkeit, parallele Zugriffe zu verarbeiten. Besonderes Augenmerk legten wir dabei auf das Verhalten der Transferzeiten einzelner Threads bei parallelen Up- & Downloads (bei gleicher Dateigröße).

Die Evaluierung durchschnittlicher Transferzeiten aller Threads bei parallelen Übertragungen zeigt deutliche Unterschiede zwischen den einzelnen Anbietern auf. Die Abbildungen 5.6 und 5.7 veranschaulichen, dass die Anzahl paralleler Upload- und Downloads mit Dateien gleicher Größe auf einige Anbieter einen größeren Einfluss hat als auf andere. Grundsätzlich lässt sich festhalten, dass mit zunehmender Zahl paralleler Übertragungen die Performance einzelner Threads deutlich abnimmt. Allerdings zeigt der vierte Teil des Experiments, dass der Gesamtdurchsatz einer Übertragung durch parallele Threads deutlich erhöht werden kann, sofern die eigene Anbindung besser ist als die des verwendeten BSPs.

Die Ergebnisse zeigen: durch eine Segmentierung einer 1 GB Datei in einzelne Datenobjekte gleicher Größe mit anschliessender parallelisierten Übertragung, kann eine 70-fache Übertragungssteigerung erreicht werden. Für eine 100 MB Datei kann sogar eine Steigerung um den Faktor 120 erzielt werden. Die Ergebnisse hängen dabei stark von der Ausgangsgröße der hochzuladenden Datei, dem Segmentierungsstufe und dem Anbieter ab. Die Abbildungen 5.8 und 5.9 geben eine Übersicht über die zu erreichenden Leistungsgewinne bei untersuchten Anbietern mit verschiedenen Segmentierungsstufen. Generell lässt sich beobachten, dass eine deutliche Verringerung der Gesamtübertragungszeit bereits bei geringer Segmentierung erreichen lässt. Wie zuvor angedeutet, fällt der Zeitgewinn bei jedem Anbieter unterschiedlich aus. Die Ergebnisse der Versuchsreihe lassen sich wie folgt zusammenfassen:

- Die als *langsam* eingestuften Provider haben ihre Schnittstellen für die Übertragung kleinerer Dateien optimiert. Diese Anbieter erreichen durch die Segmentierung großer Dateien einen deutlich größeren Leistungsgewinn, da die zu übertragenden Dateien nun deutlich kleiner sind und parallel übertragen werden können. Benötigt der Upload einer einzelnen 100 MB Datei bei HP etwa 45 Minuten, kann die Datei segmentiert und parallelisiert in 23,6 Sekunden übertragen werden. Die übrigen Anbieter erreichen zwar ebenfalls deutliche Steigerungen durch Segmentierung, jedoch nicht in dem Ausmaß.

- Verringert sich mit kontinuierlich steigender Parallelisierung die Performance einzelner Transfers ebenso kontinuierlich und ist die Übertragungszeit für alle Dateigrößen verhältnismäßig etwa gleich schnell, kann insgesamt nur ein gleichbleibender Performancegewinn erzielt werden, unabhängig von der Segmentierung (siehe Google, hier kann lediglich eine etwa 5-fache Steigerung erreicht werden)

- Bei allen untersuchten Anbietern, ausgenommen Nirvanix, lässt sich unabhängig von der gewählten Segmentierungsstufe eine Leistungssteigerung feststellen. Lediglich bestimmte Segemntierungslevel haben eine Erhöhung des Durchsatzes bei Nirvanix zur Folge. Verglichen mit der unsegmentierten Übertragung erhöhen einige Konfiguration hier gar die Übertragungszeit.

5.2.4 Verfügbarkeit

Im Laufe des gesamten Experiments haben wir bei den untersuchten Anbietern mehr als 3.5 Millionen Operationen (Lese- und Schreibzugriffe) ausgeführt. Während dieser Zeit konnten nur wenige Ausfälle und fehlschlagende Anfragen festgestellt werden. Nach einer eingehenden Überprüfung der aufgetretenen Fehler stellte sich raus, dass der Großteil der fehlgeschlagenen Anfragen auf Implementierungsfehler in unserem System zurückzuführen war. Nichtsdestotrotz konnten einige Fehler auch auf Seiten der CSP (beispielsweise *readTimeout* und *peerNotAuthenticated*) beobachten werden. Tabelle 5.1 fasst die ermittelten Verfügbarkeitswerte über die gesamte Dauer des Experiments zusammen (berechnet aus $\frac{erfolgreiche\ Operationen}{versuchte\ Operationen}$). Wie bereits mehrfach betont, handelt es dich bei den Werte nicht um allgemein gültigen Zahlen, sondern vielmehr um Erfahrungswerte, die im Rahmen unserer Untersuchungen innerhalb einen begrenzten Zeitraum festgehalten wurden.

Provider	Anfragen	Fehler	Quote	zugesichert
HP	380422	17	99,99996 %	99,95 %
Google	766717	433	99,9994 %	99,9 %
Amazon	864044	1139	99,999 %	99,9 %
Rackspace	554359	4006	99,993 %	99,9 %
Nirvanix	483869	4282	99,991 %	99,999 %
Azure	396516	4025	99,99 %	99,95 %

Tabelle 5.1: Gemessene Verfügbarkeit der untersuchten Dienste in einem Zeitraum von etwa 20 Tagen.

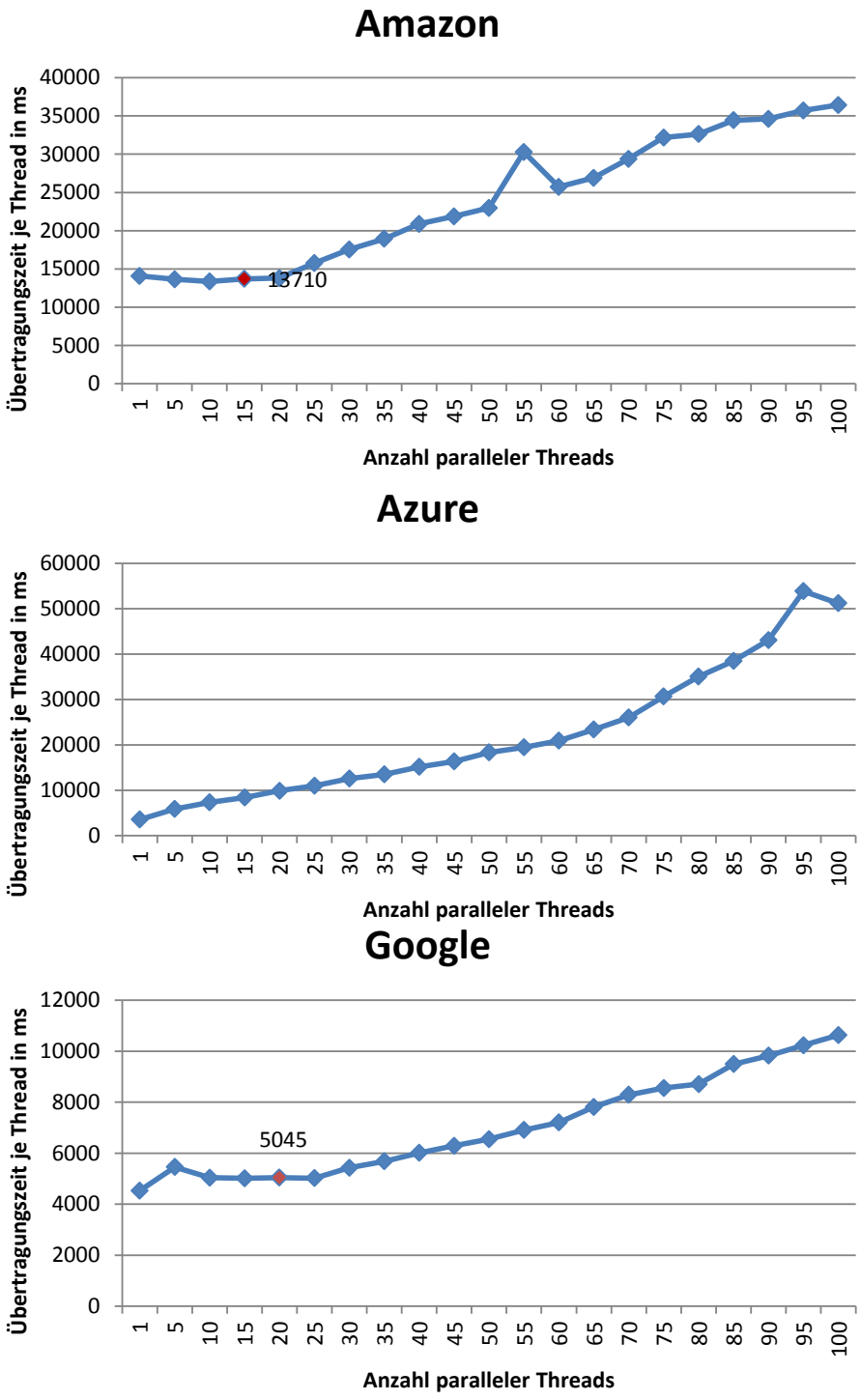


Abbildung 5.6:
Durchschnittliche Übertragungszeiten einer 10 MB-Datei bei steigender Anzahl paralleler Uploads von 10 MB-Dateien.

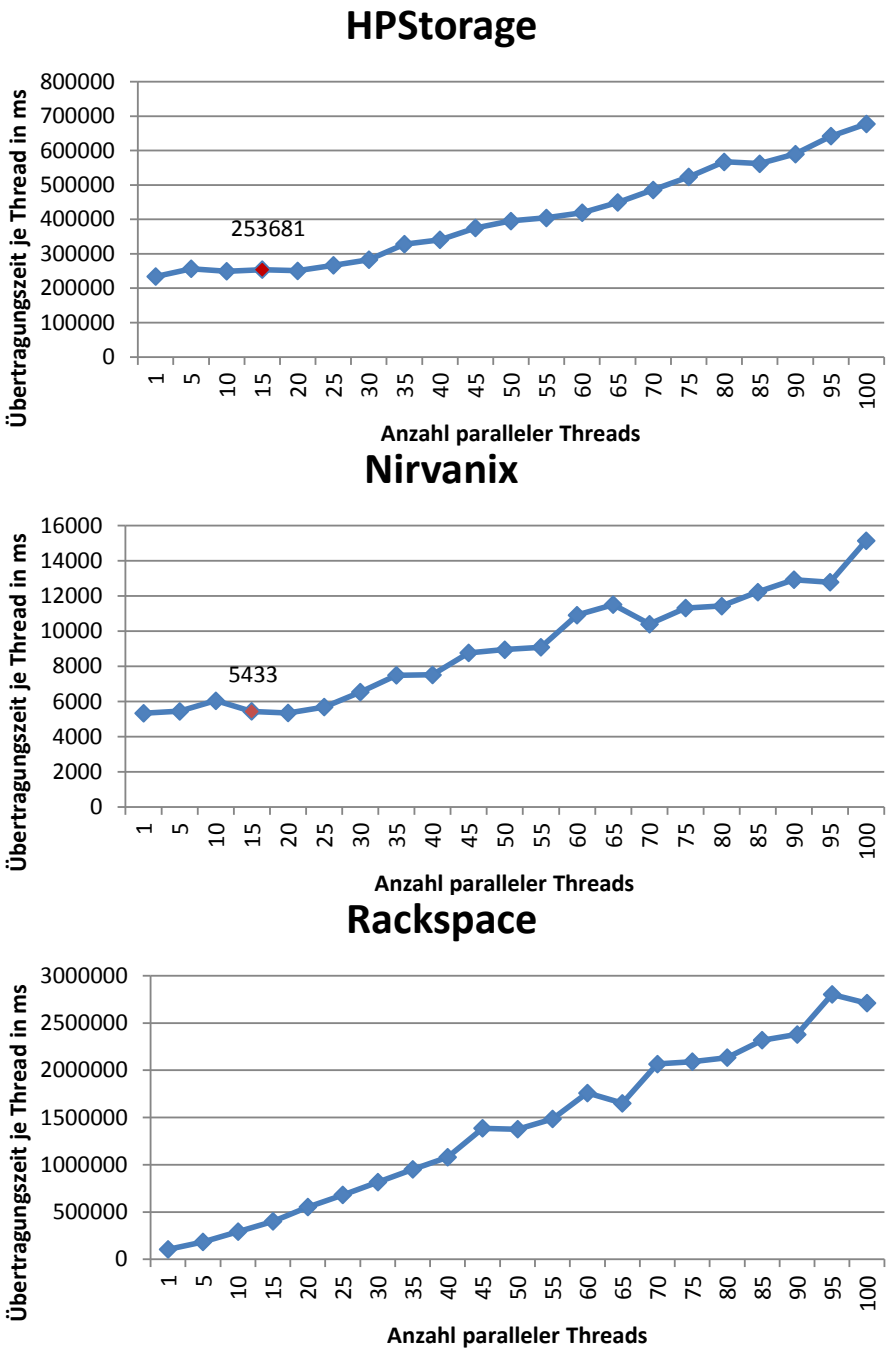


Abbildung 5.7:
Durchschnittliche Übertragungszeiten einer 10 MB-Datei bei steigender Anzahl paralleler Uploads von 10 MB-Dateien.

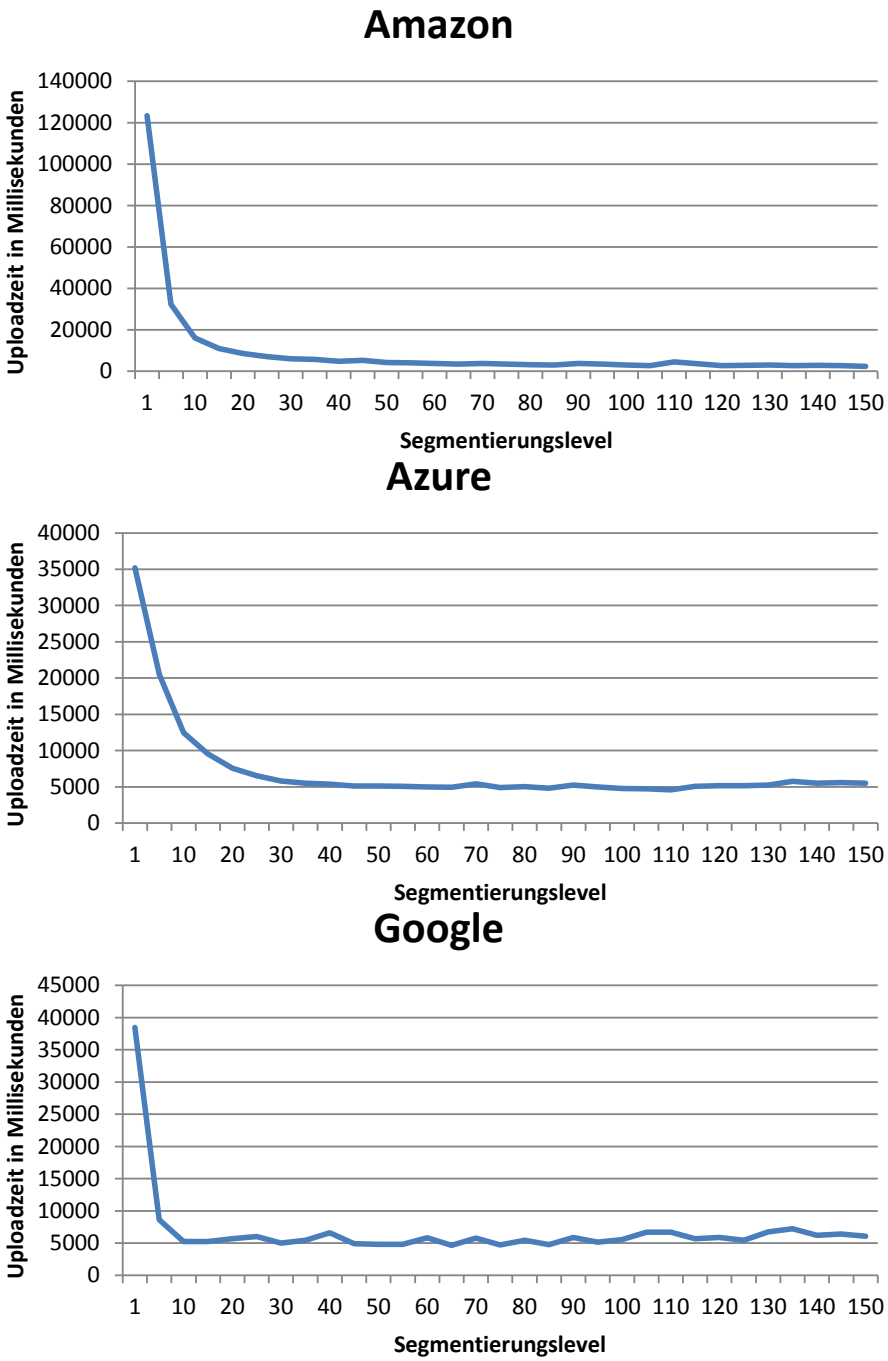


Abbildung 5.8:
Notwendige Übertragungszeit einer 100 MB-Datei bei steigender Segmentierung und Parallelität.

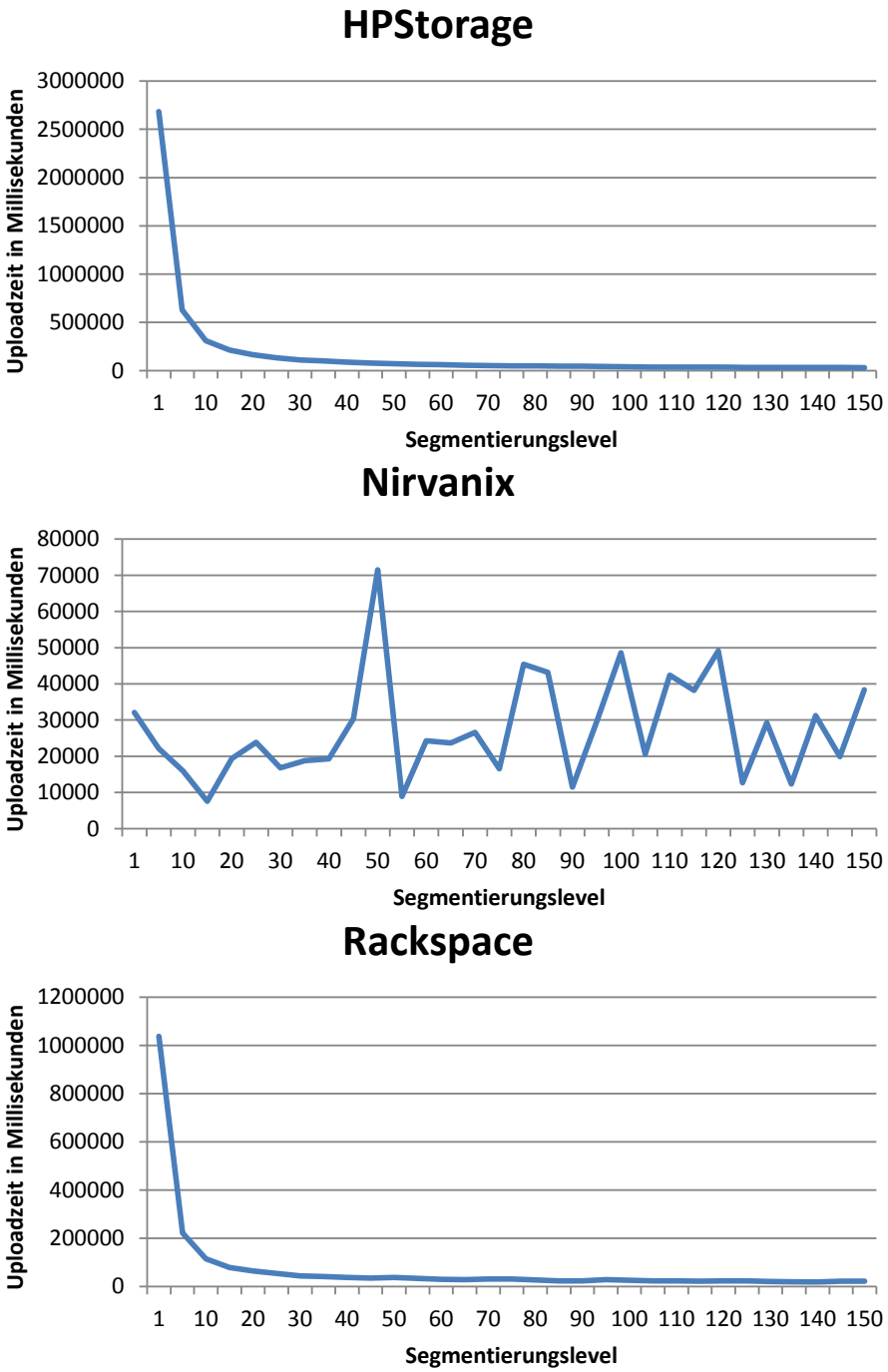


Abbildung 5.9:
Notwendige Übertragungszeit einer 100 MB-Datei bei steigender Segmentierung und Parallelität.

6 Provider im Detail

Die vorhergehenden Kapitel führen Cloud Storage relevante Szenarien für Unternehmen ein und definieren Kriterien, die bei der Auswahl eines Cloud Storage Providers beachtet werden sollten. Dieses Kapitel stellt die ausgewählten Provider und ihre Eigenschaften übersichtlich zusammengefasst im Einzelnen vor.

6.1 Microsoft Windows Azure

Windows Azure	Name	Microsoft Windows Azure BLOB Storage
	Standorte	USA (4), Europa (2), Asien (2)
	Website	`http://www.windowsazure.com/en-us/develop/net/how-to-guides/blob-storage/`

Microsoft bietet unter dem Namen Windows Azure mehrere Cloud Dienste an. Die angebotenen Storage-Dienste unterteilen sich in die nachfolgenden vier Angebote: Der Table Storage bietet eine Möglichkeit zum strukturierten Speichern von nicht-relationalen Daten. Windows Azure Drives erlaubt das Einbinden von NTFS, sodass virtuelle Maschinen darauf zugreifen können. Der Queue-Service ermöglicht Nachrichtenaustausch für Anwendungen in Windows Azure mit Hilfe von Warteschlangen umzusetzen [43]. Der Dienst, der im Zentrum der Betrachtungen der Studie steht, ist der BLOB Storage, der Speicher für binäre und textuelle Daten. Das Konzept hinter der Organisation von Speicher in diesem Dienst stellt Abbildung 6.1 dar. Die Daten eines Nutzers werden in Container strukturiert und unter einem zentralen Storage Account zusammengefasst. Ein Container wiederum enthält BLOBs (Binary Large Objects), die mit Metadaten verbunden sein können. Container bieten eine flache Speicherstruktur an, wodurch Container nicht ineinander verschachtelt werden und ausschließlich BLOBs enthalten können. Storage Accounts werden vom Nutzer angelegt und mit einer von sechs angebotenen Regionen verknüpft. Unter anderem werden hier Regionen in Nord- und Westeuropa, aber auch Regionen für die Vereinigten Staaten und Südostasien angeboten.

Je nach Anwendungsfall kann entschieden werden, ob die zu speichernden Daten *Block BLOBs* oder *Page BLOBs* sein sollen [45]. Werden Block BLOBs verwendet, wird eine

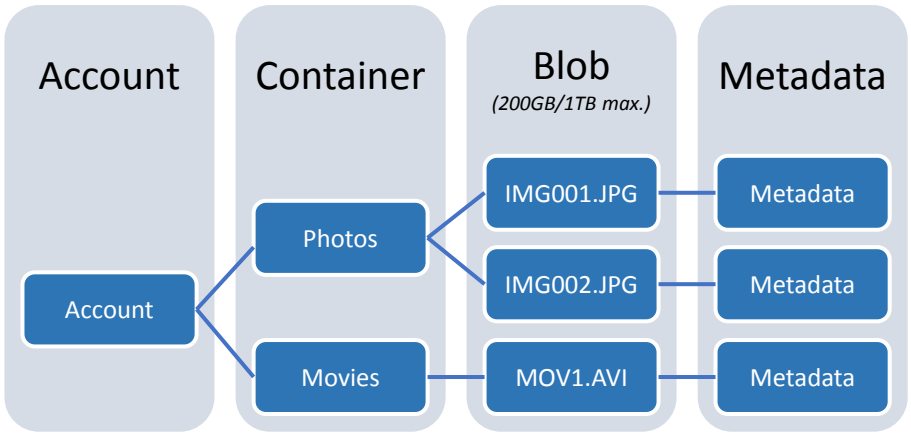


Abbildung 6.1: Speicherkonzept von Windows Azure BLOB Storage [16]

Datei in Blöcke von jeweils bis zu 4 MB Größe aufgeteilt. Block BLOBs sind für Anforderungen ausgelegt, die beim Umgang mit besonders großen Datenmengen auftreten. So ist es zum Beispiel möglich, parallel Blöcke in einen Block BLOB zu übertragen oder bei Übertragungsfehlern nur gezielt den fehlerhaften Block neu zu übermitteln. Block BLOBs können dabei maximal 50.000 Blöcke enthalten und maximal 200 GB groß sein. Page BLOBs hingegen sind besonders geeignet, um Dateien zu speichern, auf denen ein wahlfreier Zugriff erlaubt werden soll [41]. Hierbei kann auf jede Seite (Page) des BLOBs direkt zugegriffen werden. Ein Page BLOB ist auf eine maximale Größe von 1 TB beschränkt, welche zudem ein Vielfaches von 512 Byte sein muss. Es kann also passend zur konkreten Nutzung der Daten ein Speichertyp gewählt werden, der für diesen Zweck optimiert ist. Windows Azure speichert alle Daten geo-redundant, um diese vor Ausfällen von Rechenzentren zu schützen. Diese Form der Replikation lässt sich auch ausschalten. In diesem Fall wird nur lokal im Rechenzentrum repliziert, die entstehenden Kosten sind hierbei geringer. Tabelle 6.1 fasst darüber hinaus die untersuchten Kriterien der Bereiche *Recht*, *Sicherheit*, *Verfügbarkeit* und *Schnittstellen* noch einmal für Windows Azure zusammen.

Kriterium	erfüllt
Recht	
mind. 1 Standort in der EU	✓
Standort frei wählbar	✓
Teilnehmer am Safe Harbor Agreement	✓
Vertragsstrafen bei Nichterfüllung des SLA	✓
nutzerseitiges Eigentumsrecht an den Daten	✓
Sicherheit	
Datenverschlüsselung möglich	✓
Verschlüsselter Zugriff/Verbindung	✓
Zugangskontrolle/-management für mehrere Benutzer	✓
Datenzugriff nach Vertragsende	✓
Replikation möglich	✓
Daten an unterschiedlichen Standorten	✓
Begrenzter Zugriff für Mitarbeiter des CSP	✓
Verfügbarkeit	
Garantierte Verfügbarkeit in %	99.95
Transparente Definition eines Ausfalls	✓
Wartungsarbeiten zählen als Ausfall	✓
Zugriffsmöglichkeiten	
Zugriff über RESTful API	✓
Zugriff über Web Interface	✓
SDKs für mind. 2 gängige Programmiersprachen	✓
SDKs für mobile Geräte	✓
mobile Applikation	–
Beispiele und API Dokumentation	✓
maximale Dateigröße	1 TB/200 GB

Tabelle 6.1: Kriterienübersicht für *Windows Azure*

6.2 Amazon S3

amazon web services™ S3 Simple Storage Service	Name	Amazon Simple Storage Service (S3)
	Standorte	USA (3), Europa (1), Asien (3), Südamerika (1)
	Website	`http://aws.amazon.com/de/s3/`

Der Amazon Simple Storage Service (S3) ist ein Cloud Speicherdienst, der von Amazon Web Services (AWS) bereitgestellt wird. Der Dienst wurde in den USA im März 2006 gestartet und war ab November 2007 auch in Europa verfügbar. Entwickler haben hier Zugriff auf dieselbe Speicher-Infrastruktur, welche Amazon zum Ausführen seines eigenen globalen Website-Netzwerks verwendet. Die Infrastruktur wurde darauf ausgelegt, dass die Zugriffe auf die Daten vom Kunden komplett kontrolliert werden können. Des Weiteren soll eine hohe Zuverlässigkeit bei optimaler Skalierbarkeit und Geschwindigkeit gewährleistet werden. Diese Anforderungen sollen dem Kunden möglichst komfortabel und kostengünstig zur Verfügung gestellt werden. Als Teil des Amazon Web Services kann Amazon S3 mit anderen Angeboten von Amazon Web Services verbunden werden, um eine optimale Ausnutzung der Ressourcen zu erreichen.

Zur Speicherung der Objekte benutzt Amazon S3 ein Bucket-System, welches in Abbildung 6.2 skizziert wird. Hierbei wird jedes Objekt (1 Byte bis 5 Terabyte) in einem Bucket gespeichert [5]. Auch Amazon bietet, wie der Großteil der untersuchten Dienste, eine flache Bucket-Struktur an. Buckets können aus diesem Grund ausschließlich Datenobjekte, aber keine weiteren Buckets enthalten. Jeder Bucket kann vom Nutzer einzeln verwaltet werden. So können Buckets auf einzelne Regionen beschränkt und der Zugriff auf bestimmte Daten aus anderen Regionen unterbunden werden. Dies ist sehr vorteilhaft beim Umgang mit Nutzerdaten, die spezielle rechtliche Auflagen erfüllen müssen (siehe Kapitel 4.1). Für zusätzliche Kontrolle über die Buckets können Tags erstellt werden, die unter anderem auch für die Kostenkontrolle nutzbar sind. Zusätzlich zum normalen Abrechnungsmodell bietet Amazon die Möglichkeit die Requester Pays Option zu aktivieren, wodurch die Person die Kosten für die Anfrage übernimmt, die auch die Daten anfordert.

Im Zusammenhang mit dem Simple Storage Service bietet Amazon zwei weitere Dienste als Speicheroptionen an [6]: Zum Einen gibt es den Reduced Redundancy Storage (RRS)[1], zum anderen den Amazon Glacier[2]. Tabelle 6.2 fasst die untersuchten Bereiche *Recht*, *Sicherheit*, *Verfügbarkeit* und *Schnittstellen* noch einmal für Amazon S3 zusammen.

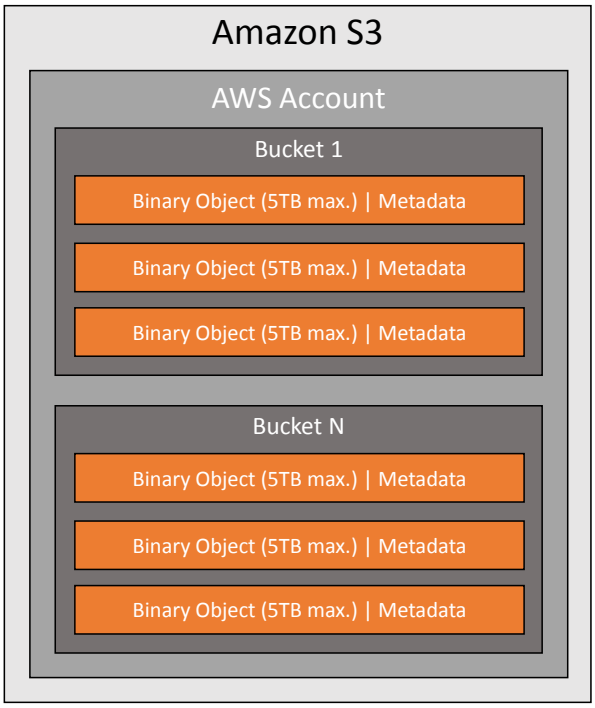


Abbildung 6.2: Amazon Bucket System [34]

[1] Durch eine geringere Redundanz werden hier die Kosten für den Nutzer gesenkt. Die Zuverlässigkeit wird von Amazon mit 99,99 % angegeben und es wird vermerkt, dass diese Option nur mit Daten genutzt werden sollte, welche reproduzierbar und nicht geschäftskritisch sind.

[2] Glacier ist auf die Langzeitspeicherung von Daten ausgelegt. Dieser Service bietet monatlich kostengünstigen Speicherplatz und wurde für Daten optimiert, auf die selten zugegriffen wird und die für Abrufzeiten von mehreren Stunden geeignet sind. Neben den sehr geringen Kosten gibt es hier Einschränkungen zur Verfügbarkeit und den Wiederherstellungszeitpunkten der Daten. Im Anwendungsfall der Archivierung (siehe Kapitel 3) kann dieser Dienst eine kostengünstige Alternative darstellen.

Kriterium	erfüllt
Recht	
mind. 1 Standort in der EU	✓
Standort frei wählbar	✓
Teilnehmer am Safe Harbor Agreement	✓
Vertragsstrafen bei Nichterfüllung des SLA	✓
nutzerseitiges Eigentumsrecht an den Daten	✓
Sicherheit	
Datenverschlüsselung möglich	✓
Verschlüsselter Zugriff/Verbindung	✓
Zugangskontrolle/-management für mehrere Benutzer	✓
Datenzugriff nach Vertragsende	✓
Replikation möglich	✓
Daten an unterschiedlichen Standorten	✓
Begrenzter Zugriff für Mitarbeiter des CSP	✓
Verfügbarkeit	
Garantierte Verfügbarkeit in %	99.9
Transparente Definition eines Ausfalls	✓
Wartungsarbeiten zählen als Ausfall	✓
Zugriffsmöglichkeiten	
Zugriff über RESTful API	✓
Zugriff über Web Interface	✓
SDKs für mind. 2 gängige Programmiersprachen	✓
SDKs für mobile Geräte	✓
mobile Applikation	–
Beispiele und API Dokumentation	✓
maximale Dateigröße	5 TB

Tabelle 6.2: Kriterienübersicht für *Amazon S3*

6.3 Google Cloud Storage

Name	Google Cloud Storage
Standorte	USA, EU
Website	`https://cloud.google.com/products/cloud-storage`

Der Dienst Google Cloud Storage ist ein Bestandteil der Google Cloud Plattform, der z. B. auch Google App Engine und Google Compute Engine angehören. Zusätzlich bietet Google einen Dienst namens Durable Reduced Availability (DRA) speziell für Szenarien wie der Archivierung von Firmendaten (siehe Kapitel 3) an. Hierbei handelt es sich um einen experimentellen Dienst, bei dem keine garantierte Verfügbarkeit gewährleistet wird. Google nennt für die zu erwartende Verfügbarkeit einen Wert von etwa 99 % [19]. Die möglichen Einbußen an Verfügbarkeit, die ein Kunde bei diesem Dienst akzeptiert, werden durch geringere Speicherkosten um bis zu 29 % ausgeglichen [21]. Die Struktur des Speichers, wie in Abbildung 6.3 dargestellt, ist analog zu der des Speicherdienstes *Amazon S3* aufgebaut. Innerhalb eines Projektes können mehrere Buckets in einer flachen Struktur angelegt werden. Buckets können daher auch hier keine weiteren Buckets, sondern lediglich Dateien von maximal 5 TB Größe enthalten. Alle Daten müssen dabei innerhalb einer Woche in den Cloud Storage übertragen sein, wodurch die effektive maximale Dateigröße eingeschränkt sein kann [18].

Derzeit findet eine kostenlose und automatische Live-Replikation der Daten statt [20]. Daten werden dabei in zwei Rechenzentren an unterschiedlichen Standorten kopiert. Der Speicherstandort eines Buckets kann hierbei vom Nutzer auf Europa oder die USA festgelegt werden [17]. Einen Überblick über das Abschneiden des Google Cloud Storage in den Bereichen *Recht*, *Sicherheit*, *Verfügbarkeit* und *Schnittstellen* gibt Tabelle 6.3.

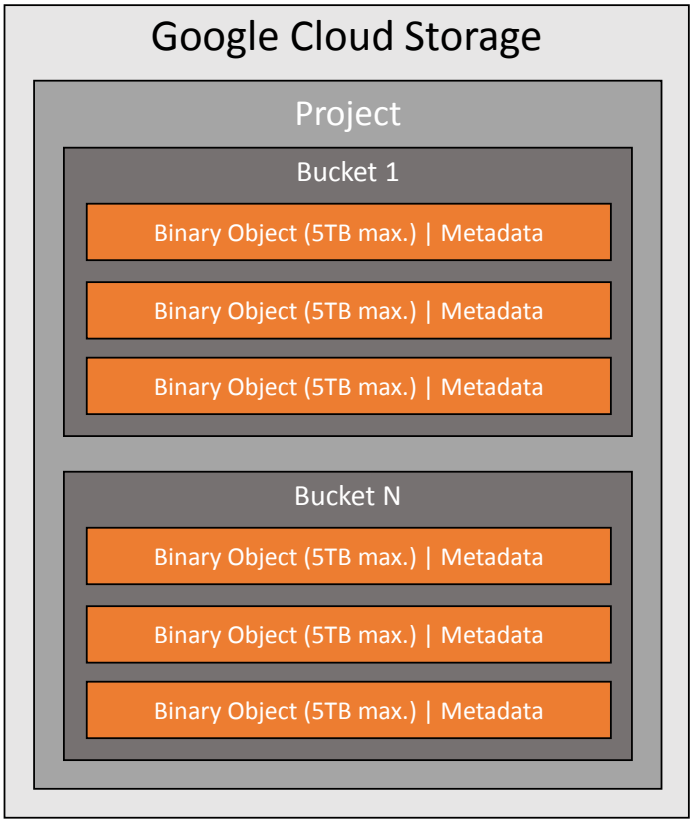


Abbildung 6.3: Speicherkonzept von Google Cloud Storage

Kriterium	erfüllt
Recht	
mind. 1 Standort in der EU	✓
Standort frei wählbar	✓
Teilnehmer am Safe Harbor Agreement	✓
Vertragsstrafen bei Nichterfüllung des SLA	✓
nutzerseitiges Eigentumsrecht an den Daten	✓
Sicherheit	
Datenverschlüsselung möglich	✓
Verschlüsselter Zugriff/Verbindung	✓
Zugangskontrolle/-management für mehrere Benutzer	✓
Datenzugriff nach Vertragsende	–
Replikation möglich	✓
Daten an unterschiedlichen Standorten	✓
Begrenzter Zugriff für Mitarbeiter des CSP	✓
Verfügbarkeit	
Garantierte Verfügbarkeit in %	99.9
Transparente Definition eines Ausfalls	✓
Wartungsarbeiten zählen als Ausfall	✓
Zugriffsmöglichkeiten	
Zugriff über RESTful API	✓
Zugriff über Web Interface	✓
SDKs für mind. 2 gängige Programmiersprachen	✓
SDKs für mobile Geräte	–
mobile Applikation	–
Beispiele und API Dokumentation	✓
maximale Dateigröße	5 TB

Tabelle 6.3: Kriterienübersicht für *Google Cloud Storage*

6.4 Rackspace Cloud Files

RACKSPACE CLOUD FILES	Name	Rackspace Cloud Files
	Standorte	USA, Großbritannien
	Website	`http://www.rackspace.com/cloud/files/`

Rackspace Cloud Files wurde im Oktober 1998 mit Fokus auf Support und Service als Alleinstellungsmerkmal gestartet. Als Mitbegründer von OpenStack implementiert Rackspace seine Dienste gemäß dessen Vorgaben. Rackspace ermöglicht es so den Benutzern, ihre Anwendungen und Daten bei anderen Anbietern zu benutzen, die ebenfalls OpenStack unterstützen.

Cloud Files von Rackspace werden in Kombination mit dem Akamai Content Delivery Network (CDN) [1] angeboten, um eine möglichst effektive Bereitstellung der Daten zu ermöglichen [59]. Analog zur Speicherstruktur der Dienste *Amazon S3* und *Google Cloud Storage* können in einem Rackspace Account (siehe Abbildung 6.4) beliebig viele Container angelegt werden, die aufgrund vorgeschriebener, flacher Strukturen keine weiteren Container, aber Dateien und deren Metadaten enthalten können. Um beliebig große Dateien zu unterstützen, gibt es die Möglichkeit, mehrere kleinere Dateien über eine Metadatei miteinander zu verbinden, so dass diese bei der Bereitstellung nur als eine einzige Datei dargestellt werden. Die Daten werden von Rackspace auf mindestens drei Servern gespeichert, welche sich an verschiedenen Standorten befinden. Die global verteilten Server von Rackspace ermöglichen eine schnelle Bereitstellung der Daten überall auf der Welt. Es werden private und öffentliche Container angeboten, wobei die privaten Container über SSL verfügbar und durch Benutzername und Passwort gesichert sind und Daten in öffentlichen Containern eine Web-Ready URL erhalten, um diese möglichst schnell im Internet verbreiten zu können. Des Weiteren können Daten auch zeitweilig über temporäre URLs und zeitlich begrenzte Objekte geteilt werden. Hierbei wird nach einer bestimmten Zeit der Link auf das Objekt ungültig oder sogar das ganze Objekt gelöscht.

Eine Übersicht über weitere Untersuchungspunkte der Bereiche *Recht*, *Sicherheit*, *Verfügbarkeit* und *Schnittstellen* und deren Ausprägung für Rackspace Cloud Files bietet Tabelle 6.4.

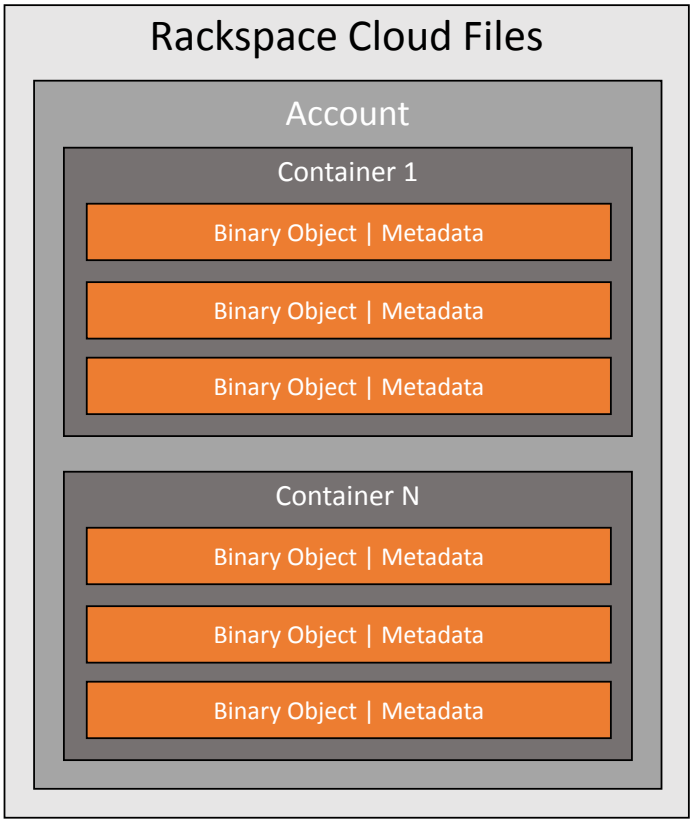


Abbildung 6.4: Speicherkonzept von Rackspace Cloud Files

Kriterium	erfüllt
Recht	
mind. 1 Standort in der EU	✓
Standort frei wählbar	✓
Teilnehmer am Safe Harbor Agreement	✓
Vertragsstrafen bei Nichterfüllung des SLA	✓
nutzerseitiges Eigentumsrecht an den Daten	✓
Sicherheit	
Datenverschlüsselung möglich	–
Verschlüsselter Zugriff/Verbindung	✓
Zugangskontrolle/-management für mehrere Benutzer	✓
Datenzugriff nach Vertragsende	–
Replikation möglich	✓
Daten an unterschiedlichen Standorten	✓
Begrenzter Zugriff für Mitarbeiter des CSP	✓
Verfügbarkeit	
Garantierte Verfügbarkeit in %	99.9
Transparente Definition eines Ausfalls	✓
Wartungsarbeiten zählen als Ausfall	–
Zugriffsmöglichkeiten	
Zugriff über RESTful API	✓
Zugriff über Web Interface	✓
SDKs für mind. 2 gängige Programmiersprachen	✓
SDKs für mobile Geräte	–
mobile Applikation	✓
Beispiele und API Dokumentation	✓
maximale Dateigröße	∞

Tabelle 6.4: Kriterienübersicht für *Rackspace Cloud Files*

6.5 HP Cloud Object Storage

hp Cloud	Name	HP Cloud Object Storage
	Standorte	USA (2)
	Website	`https://www.hpcloud.com/products/object-storage`

HP bietet im Rahmen seiner Public Cloud Produkte zwei Storage-Dienste an. Der Spezifikation aus dem OpenStack Projekt folgend wird mit HP Cloud Object Storage ein Object Store und mit HP Cloud Block Storage ein Block Storage angeboten. Die verschiedenen Produkte sind vorrangig als Speicher für Instanzen von virtuellen Maschinen des HP Cloud Computer-Services gedacht, können aber auch unabhängig von einer solchen genutzt werden und sind insbesondere durch die Implementierung von OpenStack auch für alle anderen Dienste nutzbar, welche die Compute-Schnittstelle ebenfalls unterstützen [27, 33].

Der HP Object Storage steht bei den Betrachtungen dieser Studie im Vordergrund. Der primäre Anwendungsfall ist die Speicherung von statischen Daten, die über das Internet bereitgestellt werden. Die Daten sind über das Internet für eine unbegrenzte Anzahl an Nutzern verfügbar. Der Zugriff kann durch Caching über ein Content Distribution Network beschleunigt werden. Es wird eine erhöhte Sicherheit für die Daten durch Replikation auf drei physisch getrennten Servern erreicht [33, 29]. HP repliziert im Rahmen seiner Cloud Storage Dienste automatisch und kostenlos. Die Daten werden in drei Rechenzentren in verschiedenen Verfügbarkeitszonen einer Region vorgehalten. Verfügbare Regionen bei HP sind US EAST und US WEST, eine Speicherung von personengebundenen Daten innerhalb der EU ist somit nicht möglich. Weitere rechtliche, Sicherheits- und Verfügbarkeitspunkte sowie vorhandene Schnittstellen des HP Cloud Storage sind in Tabelle 6.5 zusammengefasst.

Der Object Storage unterstützt nativ Objekte von maximal 5 GB Größe, wie in Abbildung 6.5 dargestellt. Diese können in Buckets gespeichert werden, welche wiederum zu einem Account assoziiert sind. Buckets sind bei HP in flachen Strukturen geordnet und können demnach keine weiteren Buckets, sondern lediglich Dateien und deren Metadaten enthalten. Zusätzlich zum gängigen Upload von Dateien ist es mittels Segmentierung beim Hochladen möglich, beliebig große Dateien zu speichern. Beim Download dieser großen Dateien werden transparent für den Nutzer die Segmente konkateniert [29].

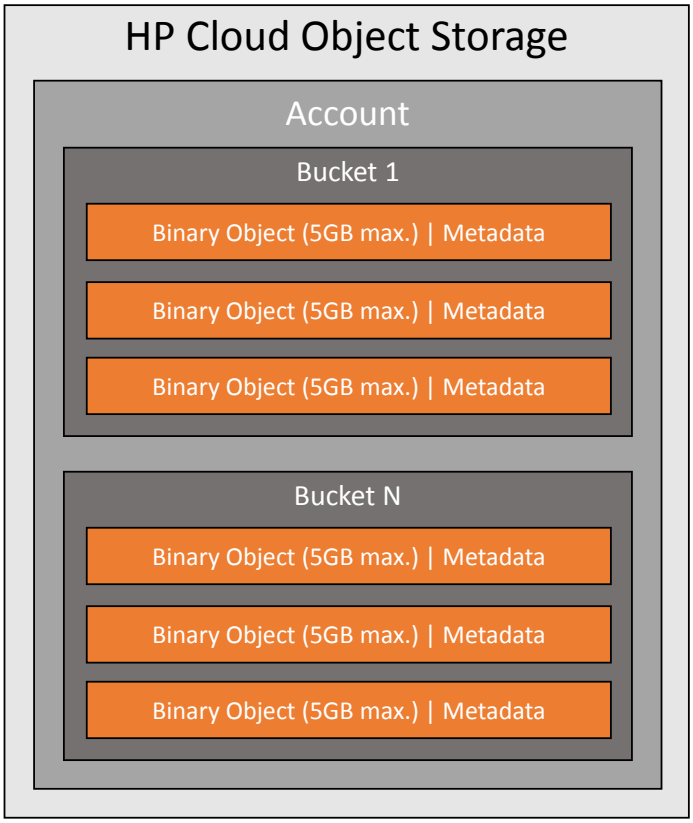


Abbildung 6.5: Speicherkonzept von HP Cloud Object Storage

Kriterium	erfüllt
Recht	
mind. 1 Standort in der EU	–
Standort frei wählbar	✓
Teilnehmer am Safe Harbor Agreement	✓
Vertragsstrafen bei Nichterfüllung des SLA	✓
nutzerseitiges Eigentumsrecht an den Daten	–
Sicherheit	
Datenverschlüsselung möglich	✓
Verschlüsselter Zugriff/Verbindung	✓
Zugangskontrolle/-management für mehrere Benutzer	✓
Datenzugriff nach Vertragsende	✓
Replikation möglich	✓
Daten an unterschiedlichen Standorten	✓
Begrenzter Zugriff für Mitarbeiter des CSP	✓
Verfügbarkeit	
Garantierte Verfügbarkeit in %	99.9
Transparente Definition eines Ausfalls	✓
Wartungsarbeiten zählen als Ausfall	✓
Zugriffsmöglichkeiten	
Zugriff über RESTful API	✓
Zugriff über Web Interface	✓
SDKs für mind. 2 gängige Programmiersprachen	✓
SDKs für mobile Geräte	–
mobile Applikation	–
Beispiele und API Dokumentation	✓
maximale Dateigröße	5 GB

Tabelle 6.5: Kriterienübersicht für *HP Cloud Object Storage*

6.6 Nirvanix Public Cloud Storage

	Name	Nirvanix Public Cloud Storage
	Standorte	USA (3), Deutschland (1), Schweiz (2), Japan (1)
	Website	`http://www.nirvanix.com/products-services/cloudcomplete-public-cloud-storage/`

Nirvanix ist ein Anbieter für Enterprise-Class Cloud Storage Services, der im Juli 2007 gegründet wurde. Der Dienstanbieter legt einen besonderen Fokus auf Sicherheit und Zuverlässigkeit, was sich in höheren Kosten im Vergleich zu anderen Anbietern niederschlägt.

Das Cloud Storage Network von Nirvanix ist die Basis für den Cloud Speicher des Anbieters und besteht aus sieben Knoten, die auf der ganzen Welt strategisch verteilt sind, um optimalen Zugriff und maximale Verfügbarkeit garantieren zu können. Zur Gewährleistung der Sicherheit setzt Nirvanix auf höchste Standards. In der Spezifikation zur Sicherheit, die von Nirvanix zur Verfügung gestellt wird, werden mögliche Angriffspunkte des Cloud Storage Networks samt getroffener Gegenmaßnahmen aufgelistet und erläutert [54]. Nirvanix bietet zudem an, die Daten gegen einen Aufpreis auf zusätzliche Knoten zu verteilen. Dadurch wird eine höhere Verfügbarkeit durch die Dienstgütevereinbarung garantiert.

Im Gegensatz zu allen anderen untersuchten Anbietern erlaubt Nirvanix eine hierarchische Speicherstruktur, wie in Abbildung 6.6 verdeutlicht. In einem Child Account können somit beliebig viele Ordner angelegt werden, die ihrerseits sowohl Dateien als auch Ordner und deren Metadaten enthalten können.

Tabelle 6.6 fasst die definierten Untersuchungspunkte für die Bereiche *Recht*, *Sicherheit*, *Verfügbarkeit* und *Schnittstellen* noch einmal für den Cloud Storage von Nirvanix zusammen.

Zwei weitere Angebote von Nirvanix sind neben dem Public Cloud Storage der Private und der Hybrid Cloud Storage. Bei der Private Cloud wird ein Cloud Knoten im Datenzentrum des Kunden angelegt [49]. Das Hybrid Cloud Angebot kombiniert das Private Cloud mit dem Public Cloud Angebot, um eine möglichst hohe Flexibilität zu erreichen. Dabei werden von Nirvanix die Überwachung, Wartung, Metadaten- und Kontenverwaltung übernommen.

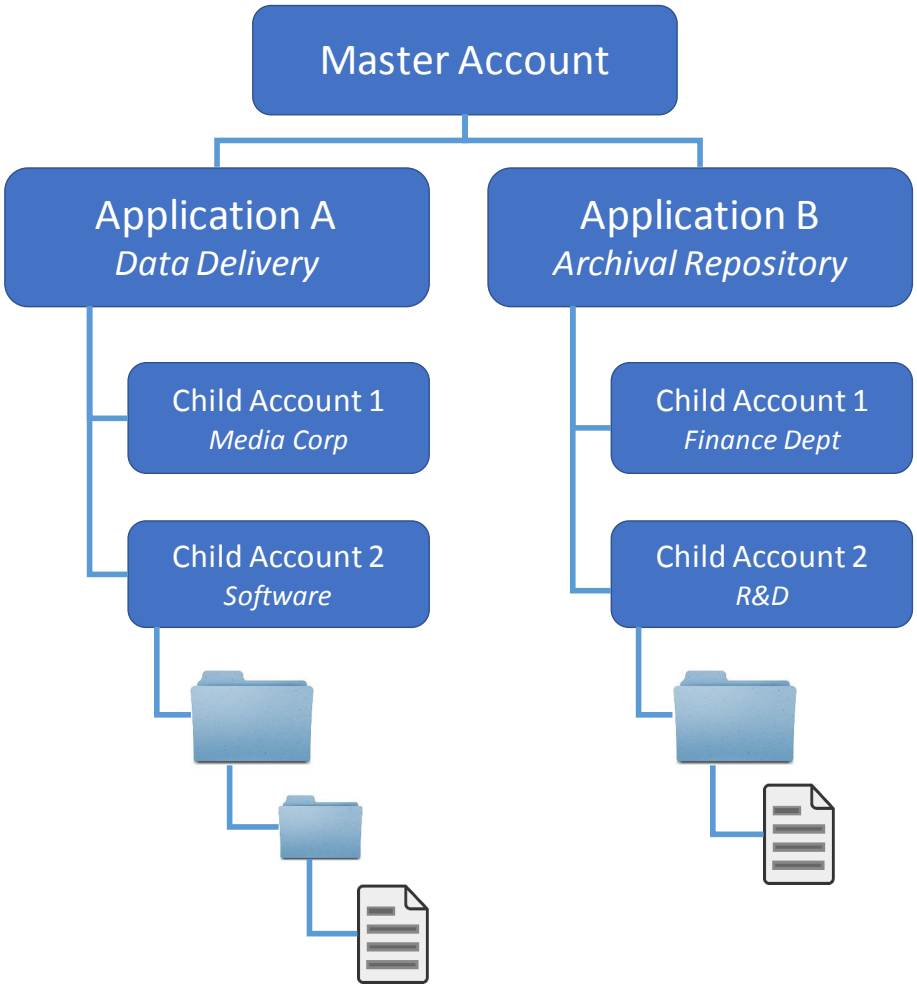


Abbildung 6.6: Speicherkonzept von Nirvanix Public Cloud Storage

Kriterium	erfüllt
Recht	
mind. 1 Standort in der EU	✓
Standort frei wählbar	✓
Teilnehmer am Safe Harbor Agreement	✓
Vertragsstrafen bei Nichterfüllung des SLA	✓
nutzerseitiges Eigentumsrecht an den Daten	✓
Sicherheit	
Datenverschlüsselung möglich	✓
Verschlüsselter Zugriff/Verbindung	✓
Zugangskontrolle/-management für mehrere Benutzer	✓
Datenzugriff nach Vertragsende	✓
Replikation möglich	✓
Daten an unterschiedlichen Standorten	✓
Begrenzter Zugriff für Mitarbeiter des CSP	✓
Verfügbarkeit	
Garantierte Verfügbarkeit in %	99.999
Transparente Definition eines Ausfalls	✓
Wartungsarbeiten zählen als Ausfall	–
Zugriffsmöglichkeiten	
Zugriff über RESTful API	✓
Zugriff über Web Interface	✓
SDKs für mind. 2 gängige Programmiersprachen	✓
SDKs für mobile Geräte	–
mobile Applikation	–
Beispiele und API Dokumentation	✓
maximale Dateigröße	2 TB

Tabelle 6.6: Kriterienübersicht für *Nirvanix Public Cloud Storage*

7 Fazit

Die Zahl der Angebote und Dienste im Cloud Computing wächst stetig an. Der Cloud Storage Bereich bildet hier keine Ausnahme. Dabei weist die Mehrheit der Angebote auf den ersten Blick häufig ähnliche Merkmale auf. Im Fall der in dieser Studie betrachteten Basic Storage Provider gilt dies auch bei näherer Betrachtung für die Bereiche Kosten, Zugriffsmöglichkeiten, Zertifikate und Standards. Die Provider bieten ähnliche Schnittstellen, wenige unterschiedliche Kostenmodelle und einen gewissen Basissatz an Standards und Zertifikaten.

Die Cloud Storage Angebote setzen sich jedoch aus mehr Bereichen als nur diesen zusammen: vor allem rechtliche Anforderungen, Sicherheitsaspekte sowie Performance und Verfügbarkeit der Dienste spielen eine wichtige Rolle. Hierbei unterscheiden sich die Anbieter in der Ausprägung der einzelnen Merkmale. Die Gewichtung dieser Merkmale für ein Unternehmen richtet sich vor allem nach dem Anwendungsfall. Die vorgestellten Szenarien *Primärspeicherung*, *Backup*, *Archivierung* und *Content Delivery* in Kapitel 3 bieten hier eine Grundlage zur Einordnung des eigenen Anwendungsfalls und der eigenen Anforderungen. Die Wichtigkeit einzelner Kriterien für ein Unternehmen hängt aber neben dem Einsatzszenario zusätzlich von Firmenrichtlinien und -strategien ab. Diese Vielfalt an möglichen Gewichtungen einzelner Merkmale führt dazu, dass eine generelle objektive Einschätzung der Eignung eines Providers für ein bestimmtes Szenario sehr schwierig ist. Die vorliegende Studie gibt deshalb Hinweise darauf, wie die Provider in bestimmten Bereichen des Cloud Storage abschneiden, welche Merkmale sie mitbringen und wie sie im Vergleich zueinander stehen. Die Arbeit liefert zudem ein Grundvokabular an Kriterien, die bei der Auswahl eines konkreten Cloud Storage Anbieters und dem Vergleich von hier nicht explizit untersuchten Diensten beachten werden sollten. Die konkrete Wahl eines Anbieters muss jeder Nutzer jedoch individuell abwägen. Die Kriterien und die detaillierten Untersuchungen dieser Studie geben hierzu einen Leitfaden, der für weitere Untersuchungen und Entscheidungsfindungen genutzt werden kann.

Literaturverzeichnis

[1] Akamai. Website, 2013. Erreichbar unter `http://www.akamai.com/`; aufgerufen am 24.04.2013.

[2] Amazon. Amazon Simple Storage Service (Amazon S3). Website, 2010. Erreichbar unter `http://aws.amazon.com/de/s3/`; abgerufen am 02.12.2010.

[3] Amazon. AWS Customer Agreement. Website, 2013. Erreichbar unter `http://alexiskold.sys-con.com/node/233855`; aufgerufen am 10.06.2013.

[4] Amazon. Amazon Web Services Sample Code and Libraries. Website, 2013. Erreichbar unter `http://aws.amazon.com/code`; aufgerufen am 24.04.2013.

[5] Amazon. Amazon S3 Managing. Website, 2013. Erreichbar unter `http://aws.amazon.com/de/s3/#managing`; aufgerufen am 24.04.2013.

[6] Amazon. Amazon Web Services Übersicht. Website, 2013. Erreichbar unter `http://aws.amazon.com/`; aufgerufen am 24.04.2013.

[7] Amazon. AWS Security Whitepaper. Whitepaper, 2013. Erreichbar unter `http://awsmedia.s3.amazonaws.com/pdf/AWS_Security_Whitepaper.pdf`; aufgerufen am 10.06.2013.

[8] Amazon. Amazon S3 SLAs. Website, 2013. Erreichbar unter `http://aws.amazon.com/de/s3-sla`; aufgerufen am 24.04.2013.

[9] Arbeitsgruppe 'Rechtsrahmen des Cloud Computing'. Datenschutzrechtliche Lösungen für Cloud Computing. Website, 2012. Erreichbar unter `http://www.fraunhofer.de/de/presse/presseinformationen/2012/februar/vertrauensvolle-cloud.html`; aufgerufen am 19.03.2013.

[10] Michael Armbrust, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy Katz, Andy Konwinski, Gunho Lee, David Patterson, Ariel Rabkin, Ion Stoica, and Matei Zaharia. Above the clouds: A berkeley view of cloud computing. *"Technical Report UCB/EECS-2009, EECS Department, University of California, Berkeley"*, 2009.

[11] Moritz Borgmann, Tobias Hahn, Michael Herfert, Thomas Kunz, Marcel Richter, Ursula Viebeg, and Sven Vowé. On the Security of Cloud Storage Services. Technical report, 2012.

[12] Dirk Csaplar. How Much of Your Data Should be in the Public Cloud. White paper, Aberdeen Group, November 2011.

[13] Cyber Ark. 2012 Trust, Security & Passwords Survey . Whitepaper, 2012. Erreichbar unter `http://www.dit.co.jp/products/cyber-ark/pdf/2012Cyber-Ark_TrustSecurityPasswordReport.pdf`; aufgerufen am 10.06.2013.

[14] Bundesministerium für Wirtschaft und Technologie. Das Normungs- und Standardisierungsumfeld von Cloud Computing. Whitepaper, 02 2012. Erreichbar unter `http://www.bmwi.de/BMWi/Redaktion/PDF/Publikationen/Studien/normungs-und-standardisierungsumfeld-von-cloud-computing,property=pdf,bereich=bmwi2012,sprache=de,rwb=true.pdf`; aufgerufen am 29.05.2013.

[15] John Gantz and David Reinsel. The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east. *IDC iView: IDC Analyze the Future*, 2012.

[16] Jonathan Gao. Windows Azure and SQL Database Tutorials - Tutorial 3: Using Windows Azure BLOB Service. Website, 2013. Erreichbar unter `http://social.technet.microsoft.com/wiki/contents/articles/2153.windows-azure-and-sql-database-tutorials-tutorial-3-using-windows-azure-blob-service-en-us.aspx`; aufgerufen am 01.05.2013.

[17] Google. Ort für Speicherung von Buckets festlegen. Website, 2013. Erreichbar unter `https://developers.google.com/storage/docs/developer-guide#specifyinglocations`; aufgerufen am 01.05.2013.

[18] Google. Google Developer Ressourcen. Website, 2013. Erreichbar unter `https://developers.google.com/storage/docs/developer-guide`; aufgerufen am 01.05.2013.

[19] Google. Google DRA Storage. Website, 2013. Erreichbar unter `https://developers.google.com/storage/docs/durable-reduced-availability`; aufgerufen am 01.05.2013.

[20] Google. Google Durability Diskussion. Website, 2013. Erreichbar unter `https://groups.google.com/forum/?fromgroups=#\!topic/gs-discussion/NwoZhPcvTWI`; aufgerufen am 01.05.2013.

[21] Google. Google Storage Preise. Website, 2013. Erreichbar unter `https://developers.google.com/storage/docs/pricingandterms`; aufgerufen am 01.05.2013.

[22] Google. Google Cloud Storage – a simple way to store, protect, and share data, 2013. Erreichbar unter `https://cloud.google.com/files/CloudStorage.pdf`; aufgerufen am 10.06.2013.

[23] Google. Daten und Sicherheit. Website, 2013. Erreichbar unter `http://www.google.com/about/datacenters/inside/data-security/index.html`; aufgerufen am 10.06.2013.

[24] Google. Service Level Agreements für Google Cloud Storage. Website, 2013. Erreichbar unter `https://developers.google.com/storage/docs/sla?hl=de`; aufgerufen am 01.05.2013.

[25] Google. Google Cloud Storage Terms of Use. Website, 2013. Erreichbar unter `https://developers.google.com/storage/docs/terms`; aufgerufen am 01.05.2013.

[26] Hewlett Packard. Hp Cloud Storage Customer Agreement. Website, 2013. Erreichbar unter `https://www.hpcloud.com/customer_agreement`; aufgerufen am 10.06.2013.

[27] Hewlett Packard. HP Block Storage Overview. Website, 2013. Erreichbar unter `https://www.hpcloud.com/products/block-storage`; aufgerufen am 24.04.2013.

[28] Hewlett Packard. Enterprise-Grade Cloud for Production Workloads. Website, 2013. Erreichbar unter `https://www.hpcloud.com/enterprise-grade-cloud`; aufgerufen am 10.06.2013.

[29] Hewlett Packard. Hp Object Storage Overview. Website, 2013. Erreichbar unter `https://www.hpcloud.com/products/object-storage`; aufgerufen am 24.04.2013.

[30] Hewlett Packard. HP Cloud Services - Security Overview. Whitepaper, 2013. Erreichbar unter `https://www.hpcloud.com/sites/default/files/HPCS%20Security%20Overview%20%2011-15-12%20v1.pdf`; aufgerufen am 24.04.2013.

[31] Hewlett Packard. Hp Cloud Storage SLAs. Website, 2013. Erreichbar unter `https://www.hpcloud.com/SLA`; aufgerufen am 24.04.2013.

[32] Hewlett Packard. Sprachen für HP Cloud Storage. Website, 2013. Erreichbar unter `https://docs.hpcloud.com/bindings`; aufgerufen am 24.04.2013.

[33] Hewlett Packard. HP cloud services - what cloud storage is right for you? Whitepaper, 2013. Erreichbar unter `http://www.papershare.com/paper/hp-cloud-services-what-cloud-storage-is-right-for-you`; aufgerufen am 24.04.2013.

[34] Alex Iskold. How Amazon S3 is going to change the world. Website, 2012. Erreichbar unter `http://alexiskold.sys-con.com/node/233855`; aufgerufen am 24.04.2013.

[35] Charlie Kaufman and Ramanathan Venkatapathy. Windows azure security overview. *go.microsoft.com*, 2010.

[36] Simone Königs. Cloud-Zertifikate als Entscheidungshilfe. Website, 10 2012. Erreichbar unter `http://www.ftd.de/it-medien/:guetesiegel-cloud-zertifikate-als-entscheidungshilfe/70105762.html`; aufgerufen am 29.05.2013.

[37] Michael T. McCarthy. Recent Developments: USA Patriot Act. *Harvard Journal on Legislation*, 39, 2002.

[38] Christoph Meinel, Christian Willems, Sebastian Roschke, and Maxim Schnjakin. *Virtualisierung und Cloud Computing, Konzepte, Technologiestudie, Marktübersicht.* Technische Berichte des Hasso-Plattner-Instituts für Softwaresystemtechnik an der Universität Potsdam; 44. Universitätsverlag Potsdam, 2011. ISBN 9783869561134.

[39] Peter Mell and Tim Grance. The NIST definition of cloud computing. National Institute of Standards and Technology. *Information Technology Laboratory, Version*, 15(10.07):2009, 2009.

[40] Microsoft. Introducing Geo-Replication for Windows Azure Storage. Website, 2011. Erreichbar unter `http://blogs.msdn.com/b/windowsazurestorage/archive/2011/09/15/introducing-geo-replication-for-windows-azure-storage.aspx`; aufgerufen am 10.06.2013.

[41] Microsoft. Windows Azure BLOB Storage: Vergleich von Block und Page BLOBs. Website, 2013. Erreichbar unter

http://msdn.microsoft.com/en-us/library/windowsazure/ee691964.aspx; aufgerufen am 01.05.2013.

[42] Microsoft. Windows Azure Storage API. Website, 2013. Erreichbar unter http://msdn.microsoft.com/en-us/library/windowsazure/dd135733.aspx; aufgerufen am 10.06.2013.

[43] Microsoft. Überblick über Windows Azure Speicherdienste. Website, 2013. Erreichbar unter http://msdn.microsoft.com/en-us/library/windowsazure/gg433040.aspx; aufgerufen am 01.05.2013.

[44] Microsoft. Windows Azure Cloud Services, Virtual Machines, and Virtual Network SLA. Website, 2013. Erreichbar unter http://www.microsoft.com/en-us/download/details.aspx?id=38427; aufgerufen am 10.06.2013.

[45] Microsoft. Übersicht über Windows Azure BLOB Storage Bestandteile. Website, 2013. Erreichbar unter http://www.windowsazure.com/en-us/develop/net/how-to-guides/blob-storage/; aufgerufen am 01.05.2013.

[46] Microsoft. Windows Azure-Vertrag. Website, 2013. Erreichbar unter http://www.windowsazure.com/en-us/support/legal/subscription-agreement/?country=de&language=de; aufgerufen am 10.06.2013.

[47] NIFIS. IT-Sicherheit und Datenschutz 2013. Website, 2013.

[48] Nirvanix. Nirvanix cloudcomplete portfolio: Hybrid cloud storage. Whitepaper, 2013. Erreichbar unter http://www.nirvanix.com/downloads/datasheets/hybrid-cloud-storage-datasheet.pdf; aufgerufen am 01.05.2013.

[49] Nirvanix. Nirvanix cloudcomplete portfolio: Private cloud storage. Whitepaper, 2013. Erreichbar unter http://www.nirvanix.com/downloads/datasheets/private-cloud-storage-datasheet.pdf; aufgerufen am 01.05.2013.

[50] Nirvanix. Nirvanix cloudcomplete portfolio: Public cloud storage. Whitepaper, 2013. Erreichbar unter http://www.nirvanix.com/downloads/datasheets/public-cloud-storage-datasheet.pdf; aufgerufen am 10.06.2013.

[51] Nirvanix. Nirvanix Secure Cloud Storage for the Enterprise. Whitepaper, 2013. Erreichbar unter `http://www.nirvanix.com/downloads/security-briefs/secure-cloud-storage-security-brief.pdf`; aufgerufen am 10.06.2013.

[52] Nirvanix. Service Level Agreements. Website, 2013. Erreichbar unter `http://www.nirvanix.com/sla.aspx`; aufgerufen am 01.05.2013.

[53] Nirvanix. Terms and Conditions. Website, 2013. Erreichbar unter `http://www.nirvanix.com/how-to-buy/terms.aspx`; aufgerufen am 10.06.2013.

[54] Nirvanix. Security on Demand. Whitepaper, 2013. Erreichbar unter `http://www.nirvanix.com/downloads/white-papers/security-on-demand-whitepaper.pdf`; aufgerufen am 01.05.2013.

[55] Hans Paulini. Was taugen Cloud-Zertifikate? Website, 09 2011. Erreichbar unter `http://www.computerwoche.de/a/was-taugen-cloud-zertifikate,2487626,2`; aufgerufen am 29.05.2013.

[56] Sachar Paulus. Standards für Trusted Clouds: Anforderungen an Standards und aktuelle Entwicklungen. *Datenschutz und Datensicherheit*, 5, 2011.

[57] Rackspace. Rackspace Cloud Files API. Website, 2013. Erreichbar unter `http://www.rackspace.com/cloud/files/api/`; aufgerufen am 24.04.2013.

[58] Rackspace. Rackspace Cloud Load Balancer. Website, 2013. Erreichbar unter `http://www.rackspace.com/cloud/loadbalancers/`; aufgerufen am 24.04.2013.

[59] Rackspace. Rackspace Cloud Files. Website, 2013. Erreichbar unter `http://www.rackspace.com/cloud/files`; aufgerufen am 24.04.2013.

[60] Rackspace. Rackspace Rackconnect. Website, 2013. Erreichbar unter `http://www.rackspace.com/cloud/hybrid/dedicatedcloud/rackconnect/`; aufgerufen am 24.04.2013.

[61] Rackspace. Security Practices. Website, 2013. Erreichbar unter `http://www.rackspace.com/information/legal/securitypractices`; aufgerufen am 10.06.2013.

[62] Rackspace. Rackspace SLAs. Website, 2013. Erreichbar unter `http://www.rackspace.com/cloud/legal/sla/`; aufgerufen am 24.04.2013.

[63] Rackspace. Rackspace Cloud Terms of Service. Website, 2013. Erreichbar unter `http://www.rackspace.com/information/legal/cloud/tos`; aufgerufen am 10.06.2013.

[64] Rajiv Ranjan, Rajkumar Buyya, and Manish Parashar. Special section on autonomic cloud computing: technologies, services, and applications. *Concurrency and Computation: Practice and Experience*, 24(9):935–937, 2012.

[65] Jagdish Rebello. Consumers Aggressively Migrate Data to Cloud Storage in First Half of 2012. Website, 2012. Erreichbar unter `http://www.isuppli.com/mobile-and-wireless-communications/news/pages/consumers-aggressively-migrate-data-to-cloud-storage-in-first-half-of-2012.aspx`, aufgerufen am 01.08.2013.

[66] Andrew Reichman. File Storage Costs Less In The Cloud Than In-House. White paper, Forrester, August 2011.

[67] Gene Ruth and Arun Chandrasekaran. Critical Capabilities for Public Cloud Storage Services. Technical report, Gartner, 2009.

[68] Jörg Schad, Jens Dittrich, and Jorge-Arnulfo Quiané-Ruiz. Runtime Measurements in the Cloud: Observing, Analyzing, and Reducing Variance. *Proceedings of the VLDB Endowment*, 3(1-2):460–471, 2010.

[69] SNIA. Cloud Storage Use Cases, 2009.

[70] Symantec. What's Yours Is Mine. Whitepaper, 2013. Erreichbar unter `http://www.symantec.com/content/de/de/about/downloads/press/Zusammenfassung_de_Studie_ENG.pdf`; aufgerufen am 10.06.2013.

[71] Techtarget. The Holy Grail of five-nines reliability, 2005. Erreichbar unter `http://www.fraunhofer.de/de/presse/presseinformationen/2012/februar/vertrauensvolle-cloud.html`; aufgerufen am 19.03.2013.

[72] TwinStrata. A Snapshot into Cloud Storage Adoption. White paper, 2012.

ERRATA

Meinel, Christoph et al.: Anbieter von Cloud Diensten im Überblick (Technische Berichte des Hasso-Plattner-Instituts für Softwaresystemtechnik an der Universität Potsdam; 84)

Universitätsverlag Potsdam 2014, ISBN: 978-3-86956-274-2

Stand: 13.03.2014

Seite, Zeile	falsch	richtig
11, 14 f.	[…] und Rackspace keinen Zugriff auf die Daten nach Vertragsende und bieten auch keine Verschlüsselung der Daten selbst an.	[…] und Rackspace keinen Zugriff auf die Daten nach Vertragsende. Rackspace bietet zudem keine Verschlüsselung der Daten selbst an.
26, 2 v.u. f.	[…] Verschlüsselung von Daten wird von den Diensten Nirvanix, Amazon, Azure und HP unterstützt.	[…] Verschlüsselung von Daten wird von allen Anbietern, mit Ausnahme von Rackspace, unterstützt.
28, 2	Logische Sicherheit ✓✓—	Logische Sicherheit ✓✓✓
42, 10	Datenverschlüsselung möglich ✓✓—	Datenverschlüsselung möglich ✓✓✓
67, 9	Datenverschlüsselung möglich —	Datenverschlüsselung möglich ✓
73, 3	mind. 1 Standort in der EU ✓	mind. 1 Standort in der EU —

Aktuelle Technische Berichte des Hasso-Plattner-Instituts

Band	ISBN	Titel	Autoren / Redaktion
83	978-3-86956-273-5	**Proceedings of the 7th Ph.D. Retreat of the HPI Research School on Service-oriented Systems Engineering**	Christoph Meinel, Hasso Plattner, Jürgen Döllner, Mathias Weske, Andreas Polze, Robert Hirschfeld, Felix Naumann, Holger Giese; Patrick Baudisch (Hrsg.)
82	978-3-86956-266-7	**Extending a Java Virtual Machine to Dynamic Object-oriented Languages**	Tobias Pape, Arian Treffer, Robert Hirschfeld
81	978-3-86956-265-0	**Babelsberg: Specifying and Solving Constraints on Object Behavior**	Tim Felgentreff, Alan Borning, Robert Hirschfeld
80	978-3-86956-264-3	**openHPI: The MOOC Offer at Hasso Plattner Institute**	Christoph Meinel, Christian Willems
79	978-3-86956-259-9	**openHPI: Das MOOC-Angebot des Hasso-Plattner-Instituts**	Christoph Meinel, Christian Willems
78	978-3-86956-258-2	**Repairing Event Logs Using Stochastic Process Models**	Andreas Rogge-Solti, Ronny S. Mans, Wil M. P. van der Aalst, Mathias Weske
77	978-3-86956-257-5	**Business Process Architectures with Multiplicities: Transformation and Correctness**	Rami-Habib Eid-Sabbagh, Marcin Hewelt, Mathias Weske
76	978-3-86956-256-8	**Proceedings of the 6th Ph.D. Retreat of the HPI Research School on Service-oriented Systems Engineering**	Hrsg. von den Professoren des HPI
75	978-3-86956-246-9	**Modeling and Verifying Dynamic Evolving Service-Oriented Architectures**	Holger Giese, Basil Becker
74	978-3-86956-245-2	**Modeling and Enacting Complex Data Dependencies in Business Processes**	Andreas Meyer, Luise Pufahl, Dirk Fahland, Mathias Weske
73	978-3-86956-241-4	**Enriching Raw Events to Enable Process Intelligence**	Nico Herzberg, Mathias Weske
72	978-3-86956-232-2	**Explorative Authoring of ActiveWeb Content in a Mobile Environment**	Conrad Calmez, Hubert Hesse, Benjamin Siegmund, Sebastian Stamm, Astrid Thomschke, Robert Hirschfeld, Dan Ingalls, Jens Lincke
71	978-3-86956-231-5	**Vereinfachung der Entwicklung von Geschäftsanwendungen durch Konsolidierung von Programmier-konzepten und -technologien**	Lenoi Berov, Johannes Henning, Toni Mattis, Patrick Rein, Robin Schreiber, Eric Seckler, Bastian Steinert, Robert Hirschfeld
70	978-3-86956-230-8	**HPI Future SOC Lab - Proceedings 2011**	Christoph Meinel, Andreas Polze, Gerhard Oswald, Rolf Strotmann, Ulrich Seibold, Doc D'Errico
69	978-3-86956-229-2	**Akzeptanz und Nutzerfreundlichkeit der AusweisApp: Eine qualitative Untersuchung**	Susanne Asheuer, Joy Belgassem, Wiete Eichorn, Rio Leipold, Lucas Licht, Christoph Meinel, Anne Schanz, Maxim Schnjakin
68	978-3-86956-225-4	**Fünfter Deutscher IPv6 Gipfel 2012**	Christoph Meinel, Harald Sack (Hrsg.)

ISBN 978-3-86956-274-2
ISSN 1613-5652