

Link: <https://www.computerwoche.de/a/storage-in-der-cloud-was-anwender-beachten-muessen,1902791>

Amazon S3 & Co.

Storage in der Cloud - was Anwender beachten müssen

Datum: 10.08.2009

Autor(en):Heinrich Vaske

Cloud Computing erhitzt die Gemüter. Sind Speicherkapazitäten aus der Wolke wirklich billiger? Und wie sicher und performant sind die Angebote aus der Cloud? Wir beantworten die wichtigsten Fragen.

So manchem IT-Manager wird blümerant zumute, wenn er mit der Cloud-Diskussion konfrontiert wird. Immerhin ist er es, der den Kopf hinhalten muss, wenn es zu **Sicherheitsproblemen**¹, Performance-Verlusten oder sonstigem Ärger kommt. Solange die Cloud-Strategien der großen IT-Player nicht richtig erkennbar sind, fühlen sich die IT-Chefs oft auch gar nicht verpflichtet, unnötige Risiken einzugehen.

Und doch bleibt ein leises Unbehagen: In vielen Unternehmen wird mit Diensten wie **Amazon S3**² lebhaft experimentiert - erfolgreich, und die Kosten sinken. Wir beantworten im Folgenden wichtige Fragen zum Thema Storage in der Cloud.

Der Unterschied zwischen Public und Private Storage Cloud

Die Public Cloud ist eine von vielen Anwendern nutzbare Speicher-Utility, bei der je nach Nutzung gezahlt werden muss. Alle Komponenten befinden sich außerhalb der Unternehmens-Firewall des Kunden in einer geteilten Infrastruktur, die logisch partitioniert, **Multi-Tenant**³-fähig (mit einer Softwareinstanz werden viele Kunden bedient) und über eine sichere Internet-Verbindung zugreifbar ist.

Anbieter wie zum Beispiel Amazon.com mit seinem Simple Storage Service (S3) erheben normalerweise eine monatliche Gebühr nach genutzten Gigabytes sowie zusätzliche eine Gebühr für die Datenübertragung zu und von der Cloud. Kunden brauchen keine eigene Speichereinheiten mehr und auch kein entsprechendes Know-how im Haus. Der Serviceanbieter verwaltet die komplette Infrastruktur und organisiert seine Kapazitäten so, dass er ein möglichst breites Kundenspektrum bedienen kann.

Eine Private Cloud residiert innerhalb der Firewall und baut auf der Hardware und Software auf, die ein Unternehmen gekauft und lizenziert hat. Die Daten bleiben vollständig im Haus und werden von der eigenen IT-Mannschaft kontrolliert. Das IT-Team legt die Speicherkapazitäten in einen Pool zusammen, aus dem die einzelnen Abteilungen und Projektteams sich unabhängig von ihrem jeweiligen Standort bedienen können. Wie in der Public Cloud lassen sich Speicherkapazitäten schnell und einfach aufstocken, in dem der Pool durch weitere Server ausgebaut wird.

Lesenswertes zu Cloud Computing:

FAQ Cloud Computing;⁴

Security-Risiken bremsen Cloud Computing aus⁵;

Industrieallianz für Cloud Computing⁶

Neun Mythen um Cloud Computing⁷

Um Speichersysteme in der Private Cloud anzubieten gibt es bereits spezielle Applikationen, die beispielsweise von **ParaScale**⁸ ("CloudStorage") oder **Caringo**⁹ ("CAStor") angeboten werden. Die Software von ParaScale wurde entwickelt, um Plattenspeicher auf vielen standardisierten Linux-Servern so anzubieten, das eine skalierbare, sich weitgehend selbst verwaltende Speicher-Cloud entsteht. Produkt-Manager Mike Maxey beschreibt den Unterschied zwischen privater und öffentlicher Cloud im Gespräch mit unserer US-Schwesterpublikation "**Computerworld**¹⁰" wie folgt: "Wenn man über ein Wide-area Network angebunden ist und die Ressourcen mit anderen Kunden teilt, ist man in der öffentlichen Cloud."

Das gebe Sinn, wenn man ein extrem dezentrales, verteilt organisiertes Unternehmen habe ohne einheitliche, zentrale Infrastruktur. Vorteile ergäben sich außerdem, wenn man sehr kurzlebige, speicherintensive Daten wie beispielsweise Film-Trailer verwalten müsse, die für einige Monate laufen sollen. Die Public Cloud eignet sich also nicht zuletzt dazu, kurzfristig Lastspitzen abzufedern.

Ist Cloud-Speicher für alle Datentypen sinnvoll?

Nein. In Clouds lassen sich am besten große Mengen unstrukturierter Daten und Archivmaterial verwalten. Nach heutigem Stand werden Public Clouds noch nicht besonders gut mit hochtransaktionalen Dateien oder Datenbanken fertig, die schnelle Netzverbindungen verlangen. Online-Transaktionsverarbeitung funktioniert definitiv nicht performant.

[Hinweis auf Bildergalerie: **Was Anwender von Cloud Computing halten**] ^{gal1}

Cloud-Speicher ist auch ungeeignet für Block-basierten Datenspeicher. Wer beispielsweise ein SAN hinter einem transaktionsorientierten CRM-System gegen eine Cloud-Lösung eintauschen möchte, macht einen Fehler. Wenn man Block-level-Storage in der Cloud abbilden wollte, käme es auch aufgrund der Latenzzeiten im Web zu deutlich verzögerten Antwortzeiten. Cloud-Systeme können Anwendern in etwa den Datenzugang von NAS-Systemen gewähren, nicht mehr.

Typische Nutzungsszenarien für die Cloud sind demnach Backup, Archivierung und Disaster Recovery. "Man wird vernünftigerweise keine Datenbank in der Cloud anlegen, aber man kann dort durchaus die Kopie einer älteren Datenbank vorhalten, anstatt sie auf teurer SAN- oder NAS-Technologie zu verwalten", sagt John Engate, Cheftechniker der **Rackspace**¹¹ Hosting Inc. Er empfiehlt, ruhende Daten in der Wolke abzulegen, durchaus auch in sehr großen Mengen.

Eine andere Daumenregel in Expertenkreisen besagt, man soll Cloud-Speicher nur für Anwendungen in Erwägung ziehen, bei denen die Zugriffszeiten nicht ganz so wichtig sind. Das gilt für Backup-, Archiv- und sonstige Massendaten, bei denen es auf die Zugriffszeiten nicht so ankommt. Datenbanken und Performance-sensitive Daten sind demnach nicht geeignet.

Vor allen Cloud-Erwägungen sollten sich Anwender allerdings erst einmal generell fragen, warum ihr Datenbestand eigentlich so ausufert. Die private oder öffentliche Cloud mit Daten zu füllen, deren Zustandekommen niemand wirklich nachvollziehen kann und die weder gespeichert werden müssen noch jemals wieder gebraucht werden, ist eine eher suboptimale Herangehensweise.

Wer nutzt in der Praxis auf welche Art Cloud-Speicher?

Start-ups und Web-2.0-Unternehmen machen vorläufig die Mehrheit der Nutzer aus. Ein Beispiel ist die **Cloudize**¹² Inc., Anbieter eines netzbasierenden Collaboration-Tools für kleine und mittlere Unternehmen, mit dem sich unter anderem Dateien zentral in der Cloud verwalten lassen. Anwender von **Salesforce.com**¹³ können dort beispielsweise Sales-Präsentationen oder Videos miteinander teilen. Die Public-Storage-Cloud dahinter stellt **Nirvanix**¹⁴, ein Anbieter aus San Diego, bereit. Für Cloudize liegt dieses Geschäftsmodell nahe, zumal Salesforce.com-Kunden längst an Cloud Computing gewöhnt sind und einen Gutteil ihrer sensiblen Kontaktdaten in der Cloud verwalten.

Im Segment der Großunternehmen gibt es bislang weniger Beispiele. "Wir sind noch in einem sehr frühen Stadium", räumt Nirvanix-CEO Jim Ziernick gegenüber der "Computerworld" ein. Die Kunden, mit denen er derzeit redet, seien klassische "Early Adopters". Interessanterweise befänden sich aber beispielsweise auch Finanzdienstleister darunter, die gesetzlich verpflichtet seien, Audiodateien mit Kundengesprächen sowie jede Menge Dateien aufzuheben. Auch Anbieter von Internet-Content, die Bild- und Multimediadateien streamen müssen, gehören zu den typischen Kunden. Zudem gebe es in vielen Unternehmen Anwender, die Cloud-Storage für Pilotprojekte oder einfach testweise ausprobierten.

Die **Schumacher Group**¹⁵, ein amerikanischer Lieferant von Management-Services und Personal rund um die Notfall-Medizin, zählt zu den Ausnahmen, die sich bezüglich Cloud Computing weit aus dem Fenster lehnen. Das Unternehmen speichert eine Reihe von Dokumenten, darunter auch Verträge und Reports, bei **Force.com**¹⁶, der Cloud-Plattform von Salesforce.com. Große Teile der Buchhaltungsdaten liegen zudem in einem **gehosteten Peoplesoft-System**¹⁷, Informationen aus dem Personalbereich beim Cloud-Dienstleister **Workday**¹⁸. Wir bilden derzeit mehr als 50 Prozent unserer Prozesse in der Cloud ab, sagt CIO Douglas Menefee. "Es handelt sich um sehr große Datenbestände mit Tausenden von Transaktionen täglich." Ende 2009 sollen 75 Prozent der Unternehmensprozesse via Cloud Computing abgewickelt werden.

Reduziert Cloud-Computing den Bedarf an eigener IT-Infrastruktur?

Öffentliche Clouds zu nutzen, heißt, den Bedarf an internen Servern und Speicherequipment herunterzufahren. Allerdings gibt es weiterhin technische Ressourcen, die intern gebraucht werden. So nutzen die meisten Storage-Angebote aus der Cloud neuere Protokolle wie **WebDav**¹⁹ oder **REST**²⁰. Wenn die Inhouse-Anwendungen des Kunden diese Protokolle nicht unterstützen, kommt auf das technische Personal Anpassungsarbeit zu. Probleme gibt es diesbezüglich vor allem mit älteren Anwendungen. Neue, modular aufgebaute Applikationen kommen mit Cloud-Speicher besser zurecht.

Anwender, die sich Cloud-Angeboten bedienen, haben meistens nicht die Reduzierung ihres IT-Personals im Hinterkopf. Sie möchten - im Gegenteil - ihre Ressourcen so nutzen, dass sie einen Differenzierungsbeitrag im Business bieten, also die Wettbewerbsfähigkeit erhöhen.

Warum sind viele Unternehmen noch skeptisch?

Die Kontrolle über die unternehmenseigenen Daten liegt im Kernverantwortungsbereich eines CIOs. Diese Daten aus der Hand zu geben, trauen sich viele IT-Verantwortlichen nicht - zumal die Anbieter ihnen - mit Ausnahme vielleicht von Amazon.com - oft nicht wirklich vertrauenswürdig erscheinen. Allerdings wird dieser Verzicht zunehmend bedauert: Aspekte wie bedarfsbezogene Abrechnung, keine Vorleistungen oder keine Planung künftiger Speicherkapazitäten klingen auch für konservative CIOs verlockend. Gene Ruth, Analyst der Burton Group, berichtet, dass keiner seiner 100 größten Kunden derzeit Cloud-Speicherdienste für die Verwaltung von Live-Daten nutze. Allerdings sei das Interesse gewaltig.

Schwierigkeiten macht den Anwendern, dass die Cloud ihrem Namen gerecht wird: Man weiß nicht, was man kauft. Gehen die Services von einem verteilten Standard-Speicher aus, von einem traditionellen Midrange-System oder von einer zuverlässigen High-end-Maschine? Oft ist nicht einmal sicher, ob der Dienstleister selbst überhaupt die Speichertechnik besitzt oder selbst Kunde eines Cloud-Anbieters ist. Die Frage der Zuverlässigkeit beziehungsweise des Risikos ist also zentral.

Ebenso die Sicherheitsthematik: In einer Speicher-Cloud liegen die Unternehmensdaten oft auf derselben Platte wie die von anderen Nutzern. Was aber, wenn beispielsweise Ermittlungsbehörden oder Geheimdienste das Medium konfiszieren? Sind dann die Daten immer noch geschützt? Die Gesetzeslage ist hier nicht eindeutig. Manche Anbieter versuchen das Problem zu lösen, indem sie jeden Kunden auf einer physikalisch anderen Platte verwalten - im Zeitalter der Virtualisierungstechniken kein sehr wirtschaftliches Modell. Andere verschlüsseln die Daten ihrer Kunden unterschiedlich, dürften damit aber den Ermittlern kein allzu großes Problem bereiten.

Manche Anwender möchten auf jeden Fall sichergehen, dass ihre Daten nicht verloren gehen können. Sie verlangen deshalb Kopien, die regelmäßig repliziert an das Data Center zurückgehen. Damit reduziert sich die Cloud auf die Funktion eines Zweitspeichers und manchem Kunden stellt sich die Frage, warum er sich überhaupt noch einer Public Cloud anvertrauen sollte.

Auch die Themen Herstellerabhängigkeit und fehlende Standards stellen sich massiv. Jeder Anbieter von Speicher-Services hat seine proprietären Programmierschnittstellen (APIs). Erst allmählich beginnen die Anbieter, über gemeinsame Standards nachzudenken. Da die meisten Cloud-Dienstleister aber Startups mit unsicherer Zukunft sind, fürchten die Kunden, dass im Falle einer Pleite das Equipment als Konkursmasse einfach beschlagnahmt werden könnte.

Auch scheint die Rückführung der Daten schwierig, wenn die Internet-Verbindung nicht genügend schnell ist. Das Cloud-Startup **Zetta**²¹ beispielsweise muss 25 Prozent seiner Gesamtkosten für eine angemessen dicke Pipe ausgeben, um die Kunden zufriedenzustellen. Bevor Anwender also einem Anbieter vertrauen, ist eine Due Diligence unerlässlich, aus der hervorgeht, wo die Daten liegen, ob sie in ausreichender Schnelligkeit zurückgeführt werden können und ob sie wirklich sicher gelagert sind.

Was sind die Hausaufgaben für CIOs?

IT-Entscheider sollten genau prüfen, bevor sie sich binden. Dazu gehört, die Data Centers zu besuchen, um zu sehen, was wo gespeichert ist. Ein Service-Level-Agreement mit den maßgeblichen Metriken ist einzurichten, durch regelmäßige Audits sollte dabei geprüft werden, ob der Anbieter seinen Pflichten nachkommt. Verlässt sich der Servicelieferant auf einen Drittanbieter, der den Speicher bereitstellt, ist auch dieser genauestens zu prüfen.

Sinken die Kosten?

Die meisten Anwender, die heute Storage-Clouds nutzen, beantworten diese Frage mit einem eindeutigen Ja. Die gemeinsame Nutzung einer großen Infrastruktur, die alle Skalenvorteile ausschöpft und von Virtualisierung profitiert, wirkt sich auf den Geldbeutel des Kunden aus. Außerdem können Anwender auf eigenes Equipment verzichten und so Kapital- in Betriebskosten umwandeln. Der größte Vorteil liegt aber darin, dass verbrauchsabhängig abgerechnet wird.

Immer mehr Unternehmen ächzen derzeit unter der rasant wachsenden Datenmenge, die sie in ihren Rechenzentren verwalten müssen. Obwohl die Speicherkosten pro Gigabyte theoretisch im freien Fall sind, müssen die Firmen doch immer wieder in den Ausbau des Data Center, die Anschaffung neuer Server, aktuelle Kühltechnik und einen möglichst niedrigen Stromverbrauch investieren. Der Kauf von Speicherequipment ist nur ein kleiner Faktor in der Gesamtrechnung. Insgesamt geht das Verwalten der rasant wachsenden Datenmengen mächtig ins Geld. Von sinkenden Speicherkosten kann nicht die Rede sein.

Anwender können von den Vorteilen, die spezialisierte Storage-Dienstleister nutzen, unmittelbar profitieren. Wer bei Amazon.com heute mehr als 50 Terabyte Speicher reserviert, zahlt pro Gigabyte monatlich nur noch zwölf (Dollar-)Cent. Vor einigen Wochen waren es noch 15 Cent. Auch die Transferkosten wurden auf drei Cent je Gigabyte gesenkt. Insgesamt pendeln die Angebote derzeit zwischen zwölf und 25 Cent je Gigabyte. Ein weiterer Vorteil liegt in der Kalkulierbarkeit, da nur noch über bezogene Gigabyte und nicht mehr über Data Center und Speicher-Equipment nachgedacht werden muss.

Links im Artikel:

- ¹ <https://www.computerwoche.de/hardware/data-center-server/1899581/>
- ² <http://aws.amazon.com/s3/>
- ³ <http://en.wikipedia.org/wiki/Multitenancy>
- ⁴ <https://www.computerwoche.de/management/cloud-computing/1881211/>
- ⁵ <https://www.computerwoche.de/hardware/data-center-server/1899581/>
- ⁶ <https://www.computerwoche.de/management/cloud-computing/1891583/>
- ⁷ <https://www.computerwoche.de/hardware/data-center-server/1898918/>
- ⁸ <http://www.parascale.com/>
- ⁹ <http://www.caringo.com/>
- ¹⁰ <http://www.computerworld.com/>
- ¹¹ <http://www.rackspace.com/index.php>
- ¹² <http://cloudize.net/>
- ¹³ <http://www.salesforce.com/>
- ¹⁴ <http://www.nirvanix.com/>
- ¹⁵ <http://schumachergroup.com/>
- ¹⁶ <http://www.force.com/>
- ¹⁷ <http://www.peoplesoft-planet.com/PeopleSoftHosting.html>
- ¹⁸ <http://www.workday.com/>
- ¹⁹ <http://de.wikipedia.org/wiki/WebDAV>
- ²⁰ <http://de.wikipedia.org/wiki/ReStructuredText>
- ²¹ <http://www.zetta.com/>

Bildergalerien im Artikel:

gal¹ **Was Anwender von Cloud Computing halten**
