

# Cloud-Technologien in der Hochschullehre – Pflicht oder Kür?

## Eine Standortbestimmung innerhalb der GI-Fachgruppe Datenbanksysteme

Stefanie Scherzinger · Andreas Thor

Received: date / Accepted: date

**Zusammenfassung** Ein eigenes Themenheft zum Datenmanagement in der Cloud dient uns als Anlass, die Präsenz von Cloud-Themen in der akademischen Datenbanklehre zu erfassen. In diesem Artikel geben wir die Ergebnisse einer Umfrage innerhalb der Fachgruppe *Datenbanksysteme* durch den Arbeitskreis *Datenmanagement in der Cloud* wieder. Dozentinnen und Dozenten von über zwanzig Hochschulen nahmen an der Umfrage teil. Es zeigt sich deutlich, dass sich das Thema „Cloud“ in der Hochschullehre zunehmend etabliert, jedoch überwiegend als ergänzendes Angebot, und seltener in der grundständigen Lehre verankert. Wir fassen die Ergebnisse unserer Umfrage zusammen und wagen Deutungsversuche.

**Schlüsselwörter** Hochschullehre · Cloud Computing

### 1 Eine Umfrage zum Stand der Lehre

Während sich die Modulbeschreibung einer „Einführung in die Datenbanken“ von Hochschule zu Hochschule kaum unterscheidet, lässt sich bei Cloud-basiertem Datenmanagement eine Vielfalt an Inhalten und Lehrformen feststellen. „Wie hältst du’s mit Cloud-Inhalten in deiner Lehre?“ ist die Gretchenfrage unter Hochschuldozentinnen und Dozenten auf den Herbst- und Frühjahrstreffen der GI-Fachgruppe *Datenbanksysteme*. In den Kaffeepausen findet ein reger Austausch von Erfahrungen statt, oft gefolgt von einem elektronischen

Austausch von Folienmaterial. Denn noch gibt es nur wenige Lehrbücher, auf die in der Vorbereitung der Lehre zurückgegriffen werden kann. Zudem schließt sich das „Marktfenster“ für eine Cloud-Vorlesung mit mühevoll gestalteten praktischen Übungen bereits im Folgesemester, so dynamisch entwickeln sich die Programmierschnittstellen. Selten findet sich etwa eine Installationsanleitung für Hadoop in gedruckter Form, die sich tatsächlich noch so ausführen lässt.

Gerade Dozentinnen und Dozenten, die in ihrer Lehre stark auf *Hands-on* Projekte setzen, bekommen die Notwendigkeit der regelmäßigen und sehr aufwändigen Pflege ihrer Projekt-Infrastruktur zu spüren. Diese stellt hohe Ressourcen-Anforderungen, die durch einen akademischen Mittelbau gestützt werden müssen. Mitfühlen können hier die Kolleginnen und Kollegen, die in den letzten Jahren auf eine Professur für mobile Anwendungen berufen wurden; sie haben mit sich ständig ändernden APIs auf dem Smartphone-Markt zu kämpfen.

Im 2012 gegründeten Arbeitskreis *Datenmanagement in der Cloud* (DMC)<sup>1</sup> des GI-Fachbereichs *Datenbanken und Informationssysteme* (DBIS) greifen wir diese Gretchenfrage auf. Mit einer online-Umfrage eruiieren wir, wie wir uns als fachliche *Community* in der Lehre zu diesen neuen Technologien positionieren.

Über den E-Mailverteiler der Fachgruppe *Datenbanksysteme*<sup>2</sup> wurde im Dezember 2013 dazu eingeladen, an der Umfrage teilzunehmen. Über den Zeitraum von zwei Monaten gingen 21 Antworten aus Deutschland und auch aus Österreich ein. Die Beiträge stammen zu gleichen Teilen von Lehrenden an Universitäten und Fachhochschulen. Es zeigt sich somit klar, dass das

---

Stefanie Scherzinger  
OTH Regensburg  
E-Mail: stefanie.scherzinger@oth-regensburg.de

Andreas Thor  
Deutsche Telekom  
Hochschule für Telekommunikation Leipzig  
E-Mail: thor@hft-leipzig.de

---

<sup>1</sup> <http://dmc.fim.uni-passau.de/>

<sup>2</sup> [fg-db@informatik.uni-rostock.de](mailto:fg-db@informatik.uni-rostock.de)

Thema Cloud an allen Hochschulformen Einzug in die Lehre gehalten hat.

Der Fragebogen adressierte in über 20 Fragen bestehende Lehrveranstaltungen im Bereich Datenbanken, die auch Cloud-Themen behandeln. Dabei ging es um

- die Art der Lehrveranstaltungen,
- die verwendeten Lehrmaterialien,
- die konkreten Lehrinhalte,
- die verwendete technische Infrastruktur, sowie
- die „Konkurrenz“ durch die Lehrangebote anderer Fachgebiete der Informatik.

Im Folgenden fassen wir die Umfrageergebnisse zusammen. Wir schließen mit einem kurzen Ausblick.

## 2 Lehrveranstaltungen: Überwiegend Wahlpflichtfächer im Master-Studium

In welcher Form werden Cloud-Themen heute an unseren Hochschulen gelehrt? Laut unserer Umfrage werden sie vor allem im Rahmen von Vorlesungen behandelt (zu 76%). Seminare, in denen die Studierenden wissenschaftliche Arbeiten präsentieren, werden sowohl an Universitäten als auch an Fachhochschulen durchgeführt (zu 48%, Mehrfachnennungen möglich). Üblich sind auch eigene Praktika, in denen mit Cloud-Technologie *hands-on* gearbeitet wird (zu 43%). Externe Vorträge, bei denen eingeladene Experten zu ausgewählten Themen referieren, wurden seltener genannt (zu 19%).

Wenig überraschend wird das Thema bisher vor allem im Master-Studium als Wahlpflichtfach angeboten. Das liegt zunächst sicherlich an der Komplexität des Themas sowie an den notwendigen Voraussetzungen für Studierende, um Veranstaltungen zum Thema Cloud (sinnvoll) belegen zu können. Zudem bietet das Master-Curriculum oft eine größere organisatorische Flexibilität, so dass neue Themen einfacher einfließen können.

Ein interessantes Evaluationsergebnis lieferte die Frage nach der Eigenständigkeit der Vorlesung. Zu ungefähr gleichen Teilen wird das Thema in einer eigenen Veranstaltung oder als Teil einer (allgemeineren) Veranstaltung behandelt. So wird etwa in einzelne Cloud-Themen innerhalb bestehender Datenbankvorlesungen eingeführt. Man kann vermuten, dass die zunehmende Bedeutung des Cloud-Themas den Trend zu eigenständigen Lehrveranstaltungen verstärkt. Dennoch können auch einführende Datenbankvorlesungen genutzt werden, um Studierenden einen ersten Einblick in das Cloud-Datenmanagement zu geben.

Dieses „Anwerben“ von Studierenden für spätere eigenständige Lehrveranstaltungen sollte auch vor dem Hintergrund geschehen, dass es natürlich auch Konkurrenz aus anderen Fachgebieten um die Deutungs-

hoheit zum Thema Cloud – und natürlich um motivierte Studierende – gibt. So betrachten andere Fachgebiete das Thema aus ihrem eigenen Blickwinkel; Beispiele sind Software Engineering (z.B. mit Software-as-a-Service), Verteilte Systeme (z.B. skalierbare Anwendungen), Netzwerke (z.B. Software-defined Networks) oder Information Retrieval (z.B. NLP-Processing mit MapReduce).

Zusammenfassend lässt sich feststellen, dass das Thema Cloud in der Lehre präsent ist, allerdings überwiegend im „Kürprogramm“ und für ein fachlich fortgeschrittenes Publikum.

## 3 Lehrmaterial: Woher nehmen, wenn nicht „stehlen“?

Laut unserer Umfrage werden für die Lehre von Cloud-Technologien zu 67% wissenschaftliche Publikationen eingesetzt. Zu 86% werden eigene Foliensätze erstellt. Es ist nahe liegend, dass sich die Vorlesungsvorbereitung dadurch sehr aufwändig gestaltet.

Der Markt an Sachbüchern zu Cloud-Technologien boomt, der Markt an geeigneten Lehrbüchern für die Hochschullehre ist dagegen noch überschaubar. So geben in unserer Umfrage nur 24% an, ein Lehrbuch einzusetzen. Sachbücher hingegen wurden zu 48% genannt. Das zeigt nach Ansicht der Autoren einen klaren Bedarf an geeigneter Literatur auf, insbesondere wenn Cloud-Themen sich auch in der grundständigen Lehre etablieren sollen.

An dieser Stelle wollen die Autoren auf Lehrbücher aus der deutschen Datenbank-*Community* verweisen, ohne einen Anspruch auf Vollständigkeit zu erheben.

- Im deutschsprachigen Lehrbuch „Datenbanksysteme – eine Einführung“ von Alfons Kemper und André Eickler [6] finden sich einzelne Kapitel zu NoSQL-Datenbanken, Multitenancy sowie dem MapReduce-Konzept. Dadurch lässt sich im Rahmen einer grundständigen Datenbank-Vorlesung ein erster Einblick in diese aktuelle Themen geben.
- Auch das „Taschenbuch Datenbanken“ wird in seiner aktuellen Neuauflage um Cloud-Themen ergänzt. In einem eigenen Kapitel zu NoSQL-Datenbanken erläutert Uta Störl die Herausforderungen in der Anwendungsentwicklung mit NoSQL-Datenbanken, insbesondere bei der Datenmodellierung [8].
- Von Stefan Edlich et al. stammt ein sehr frühes deutschsprachiges Buch mit dem Titel „NoSQL: Einstieg in die Welt nicht-relationaler Web 2.0 Datenbanken“ [4]. Das Buch ist sehr angewandt und eignet sich als Referenz für Projektarbeiten.

- Wolfgang Lehner und Kai-Uwe Sattler haben mit „Web-Scale Data Management for the Cloud“ [7] ein englischsprachiges Buch herausgegeben, das die fachliche Tiefe für eine Veranstaltung im Master-Studium bietet.

Auch in fachlich verwandten Bereichen, wie etwa der Wirtschaftsinformatik, finden sich relevante Lehrbücher, so etwa „E-Business mit Cloud Computing“ von Thomas Barton [2]. Erwartungsgemäß liegt hier der Fokus auf betriebswirtschaftlichen Aspekten.

Im handlichen Taschenbuchformat bietet der Titel „Cloud Computing – Web-basierte dynamische IT-Services“ [3] von Christian Braun et al. einen Überblick über wesentliche Begrifflichkeiten und zentrale Anbieter. Das Buch leistet somit als eine erste Orientierungshilfe im *Buzzword*-Dschungel gute Dienste.

#### 4 Inhalte: MapReduce ist immer mit dabei

Ein weiterer Teil unserer Umfrage beschäftigte sich mit den konkreten Lehrinhalten, wo wir Themenbereiche zur Auswahl vorgaben, aber auch die Möglichkeit der Nennung nicht aufgeführter Themen vorsahen.

Ein fester Bestandteil von Cloud-Lehrveranstaltungen ist das MapReduce-Konzept, welches von nahezu allen Teilnehmerinnen und Teilnehmern der Umfrage genannt wurde (zu 90%). Man kann also davon ausgehen, dass Studierende nach Absolvierung einer Cloud-Veranstaltung das Prinzip von MapReduce-Programmen verinnerlicht haben.

Auffällig ist, dass in diesem Zusammenhang High-Level-Sprachen (u.a. Hive und Pig) einen wesentlichen Anteil in der Lehre einnehmen (zu 67%). Hier bleibt die Datenbank-*Community* offenbar dem Ansatz treu, dass Datentransformationen deklarativ zu beschreiben sind.

Dagegen werden die (manuelle) Implementation von Algorithmen als MapReduce-Programme (zu 57%), MapReduce-basierte Join-Algorithmen (zu 43%) oder MapReduce-spezifische Optimierungstechniken (zu 29%) nur zu einem kleineren Teil behandelt.

Bei dem Themenbereich Data Stores wurden einfache Key-Value-Stores, Document Stores und Record Stores zu jeweils 71% genannt. Seltener thematisiert werden Graph-Datenbanken (zu 43%) und RDF-Stores (zu 10%). Allgemein finden das Thema der „Data Cloud“ (z.B. Linked Open Data) und die damit verbundenen stark vernetzten Datensätze kaum Eingang in Lehrveranstaltungen. Diese Beobachtung steht im Einklang mit der Aussage von Gerhard Weikum [10], dass Datensätze, d.h. deren Erstellung und Verfügbarmachung, im Vergleich zu anderen *Communities* vergleichsweise wenig gewürdigt werden. Hier kann man vermuten, dass

die Datenbank-*Community* das Potenzial von Big Data eventuell nicht komplett ausschöpft.

Es scheint fast, dass wir unseren Studierenden anhand „langweiliger“ Daten die Techniken der effizienten Datenverarbeitung in der Cloud vermitteln; wir zeigen ihnen aber nicht den *Impact* anhand konkreter Szenarien oder Anwendungen auf.

Diesen Blickwinkel bzw. Fokus auf Verfahren und Techniken bestätigt auch das Umfrageergebnis, wonach klassische Datenbank-Themen ebenfalls häufig in Cloud-Lehrveranstaltungen auftreten. So haben (verteilte) relationale DBMS in der Cloud (zu 62%) als auch (verteilte) In-Memory-Datenbanken (zu 57%) ihren Platz. In diesem Zusammenhang werden auch Verfahren zu verteilten Transaktionen sowie zu Synchronisation vorgestellt (zu 38% bzw. 33%).

#### 5 Infrastruktur: XaaS predigen vs. praktizieren

Zu den Herausforderungen in der Lehre von Cloud-Technologien gehört es, Cloud Computing für Studierende erfahrbar zu machen. Zwar geben über 81% der Befragten an, dass *as-a-Service*-Modelle zu den Kerninhalten ihrer Lehrveranstaltungen zählen, doch nur 19% setzen die Infrastruktur der großen Cloud-Anbieter in ihrer Lehre ein. Diese Beobachtung verleitet uns dazu, dieses Kapitel mit einem provokativen Titel zu versehen: Es scheint, als würden wir zwar *as-a-Service*-Dienste in unserer Lehre predigen, aber in den vorlesungsbegleitenden Übungen nicht praktizieren.

An dieser Stelle wollen wir zwei Möglichkeiten aufzeigen, wie Studierende ihre Projekte tatsächlich in einer öffentlichen Cloud laufen lassen können:

- Mit kostenlosen Kontingenten auf der *Google Cloud Platform* [5] können eine Vielzahl von Systemen und Technologien ausprobiert werden: Dazu zählen die *Google AppEngine* (ein PaaS-Dienst), *Google Cloud Datastore* (ein NoSQL Document Store), und *Blobstore* (ein klassischer Key-Value Store). Sogar MapReduce-Prozesse können auf *Google App Engine* gestartet werden.
- Dabei beschränkt sich der Installationsaufwand auf wenige Eclipse-Plugins, die Studierenden benötigen zusätzlich einen Google Account. Seit dem großen Re-Launch der *Google Cloud Platform* im März 2014 sind die angebotenen Systeme zudem noch besser miteinander integriert.
- Amazon bietet im Rahmen seines *AWS in Education*-Programms [1] \$100-Gutscheine für alle teilnehmenden Studierenden einer Lehrveranstaltung an. Die Bewerbung ist unkompliziert und führt nach der Erfahrung der Autoren auch zum Erfolg.

Die AWS-Plattform bietet eine Vielzahl von Systemen an, insbesondere lassen sich MapReduce-Programme einfach und doch eindrucksvoll ausführen. Allerdings setzt die AWS-Plattform voraus, dass die Studierenden eine Kreditkarte hinterlegen (für den Fall, dass das Guthaben überschritten wird). Die Erfahrung zeigt mitunter, dass längst nicht alle Studierenden dazu bereit sind oder gar eine eigene Kreditkarte besitzen. Alternativ kann auch ein gemeinsames Gruppenkonto verwaltet werden, doch ist dann ebenfalls eine Kreditkarte zu hinterlegen, die womöglich belastet wird. Zudem ist fraglich, inwiefern bei einem Gruppenkonto ein Haftungsrisiko für die Handlungen einzelner Studierender besteht. Somit sind AWS-Kontingente für die Lehre nicht an allen Hochschulen praktikabel.

In unserer Umfrage geben 38% an, den Studierenden für ihre Projekte einen eigenen Rechencluster zur Verfügung zu stellen. Dies gilt nach unseren Daten gleichermaßen für Universitäten als auch für Fachhochschulen. So betreibt etwa die *Hochschule Darmstadt* einen eigenen *Big Data Cluster* für Lehrzwecke [9].

Ein alternativer Ansatz für Praktika ist das Bereitstellen von virtuellen Maschinen (zu 43%), die die Studierenden auf ihren Rechnern laufen lassen. Mitunter werden die Studierenden auch gebeten, sich selbst eine Infrastruktur aufzusetzen (zu 38%), wodurch wertvolle praktische Kompetenzen erworben werden.

## 6 Ausblick

Zu den Zielen des Arbeitskreises *Datenmanagement in der Cloud* zählt explizit auch der Erfahrungsaustausch in der Hochschullehre. Mit unserer Umfrage und der Präsentation der Ergebnisse laden wir alle Interessierten dazu ein, die Diskussion voranzutreiben, und Cloud-Themen auch in der Datenbank-Lehre zu verankern. Insbesondere erwarten wir, dass Cloud-Themen langfristig nicht mehr nur zur „Kür“ in der akademischen Informatik-Ausbildung zählen, sondern Teil des Pflichtprogramms werden.

Wir bedanken uns an dieser Stelle bei allen, die sich an unserer Umfrage beteiligt haben, und mit ihren Beiträgen diese erste Standortbestimmung erst möglich gemacht haben.

**Danksagung** Wir danken Edwin Schicker, Günther Specht und Uta Störl für ihre Kommentare zu einer früheren Version dieses Artikels.

## Literatur

1. Amazon Web Services: AWS in Education (Grants). <http://aws.amazon.com/de/grants/>, Stand April 2014
2. Barton, T.: E-Business mit Cloud Computing. Springer Vieweg (2014)
3. Baun, C., Kunze, M., Nimis, J., Tai, S.: Cloud Computing – Web-Based Dynamic IT Services, 2. Auflage. Springer (2011)
4. Edlich, S., Friedland, A., Hampe, J., Brauer, B.: NoSQL: Einstieg in die Welt nichtrelationaler Web 2.0 Datenbanken, 2. Auflage. Carl Hanser Verlag (2011)
5. Google Inc.: Google Cloud Platform. <https://cloud.google.com/>, Stand April 2014
6. Kemper, A., Eickler, A.: Datenbanksysteme – Eine Einführung, 9. Auflage. Oldenbourg (2013)
7. Lehner, W., Sattler, K.U.: Web-Scale Data Management for the Cloud. Springer Publishing Company, Incorporated (2013)
8. Störl, U.: NoSQL-Datenbanksysteme. In: T. Kudraß (ed.) Taschenbuch Datenbanken. Carl Hanser Verlag (2014)
9. Störl, U.: Big Data Cluster an der Hochschule Darmstadt. <https://www.fbi.h-da.de/organisation/personen/stoerl-uta/big-data-cluster.html>, Stand Februar 2014
10. Weikum, G.: Where's the Data in the Big Data Wave? ACM Sigmod Blog (2013). <http://wp.sigmod.org/?p=786>, March 6, 2013