Thomas Hildmann (Hrsg.)

Clouddienste im Hochschuleinsatz 2016/2017

Proceedings der Veranstaltungen „Forum Clouddienste“ im Rahmen der DFN-Betriebstagungen am 29. September 2016 und 22. März 2017

Universitätsverlag der TU Berlin

Technische Universität Berlin

Thomas Hildmann (Hrsg.)

Clouddienste im Hochschuleinsatz 2016/2017

Clouddienste im Hochschuleinsatz 2016/2017

Proceedings der Veranstaltungen „Forum Clouddienste"
im Rahmen der DFN-Betriebstagungen am
29. September 2016 und 22. März 2017

Herausgeber:
Thomas Hildmann

Universitätsverlag der TU Berlin

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 der TU Berlin, 2018
http://verlag.tu-berlin.de

Fasanenstr. 88, 10623 Berlin
Tel.: +49 (0)30 314 76131 / Fax: -76133
E-Mail: publikationen@ub.tu-berlin.de

Druck: docupoint GmbH
Satz/Layout: J. Daniel

Umschlagfoto:
Mathieu Plourde | https://www.flickr.com/photos/mathplourde/3633883684/
CC BY 2.0 | https://creativecommons.org/licenses/by/2.0/

ISBN 978-3-7983-2928-7 (print)
ISBN 978-3-7983-2929-4 (online)

Zugleich online veröffentlicht auf dem institutionellen Repositorium der Technischen Universität Berlin:
DOI 10.14279/depositonce-5955
http://dx.doi.org/10.14279/depositonce-5955

Inhalt

Vorwort

Thomas Hildmann

Der vorliegende Tagungsband „Clouddienste im Hochschuleinsatz“ ist eine Fortsetzung der Reihe „Cloudspeicher im Hochschuleinsatz“ nach den gleichnamigen Veranstaltungen. Im Jahr 2016 entschieden wir uns diese Veranstaltung mit dem DFN-Forum Cloud-Dienste zusammenzulegen. Gleichzeitig wurde der Fokus über den Cloudspeicher hinaus erweitert.

Wie an meinem eigenen Beitrag (TU Berlin, S. 51) zu erkennen ist, entspricht die Erweiterung des Fokus der natürlichen Entwicklung. Wie bereits in den Beiträgen aus 2015 erkennbar war, ist die Integration von Lösungen immer relevanter. Die Grenzen zu einer reinen Speicherlösung verschwimmen immer mehr.

So passt der Statusbericht aus dem Block „Rechtliches und Organisatorisches“ zum Thema IaaS-Vergabeverfahren beim GÈANT (DFN, S. 21) ausgezeichnet in die aktuelle Entwicklung. Daran geknüpfte Rechtsfragen für solche Angebote behandelt Hr. Nehlsen (S. 5) und zeigt u. a. noch einmal deutlich die Vorteile, die die Kooperation unter den Hochschulen im Bereich IaaS auch aus rechtlichen Gründen bringen.

Es folgen zwei Erfahrungsberichte aus der DFN-Cloud (S. 25) und aus Münster zum Projekt sciebo (S. 35), die beide nahtlos an die Berichte aus den Vorjahren anschließen.

Den Erfahrungsberichten folgen zwei technisch orientierte Artikel: The Adoption of a New Cloud Storage Service in German Universities (Uni Münster, S. 39) und der bereits zitierte Artikel aus Berlin.

Mit Interesse aber auch Sorge beobachteten wir im letzten Jahr, wie das weit verbreitete Projekt ownCloud eine Aufspaltung erlebte und den Mitbewerber Nextcloud hervorbrachte. Nach zahlreichen Fragen aus der DFN-Community organisierten wir eine Fragerunde mit Vertretern beider Firmen deren Ergebnisse Hr. Röder (DFN, S. 59) zusammengefasst hat.

Ich möchte mich an dieser Stelle noch einmal bei allen Teilnehmerinnen und Teilnehmern sowie den Vortragenden aus den DFN-Foren für ihre Mitwirkung danken. Mein besonderer Dank gilt in diesem Zusammenhang jedoch den Autoren der Beiträge in diesem Tagungsband sowie den Kolleginnen und Kollegen, die bei der Zusammenstellung und Aufarbeitung mitgewirkt haben!

Allen Leserinnen und Lesern wünsche ich nun viel Spaß bei der interessanten Lektüre. Das DFN-Forum steht dem Diskurs zu diesen und verwandten Themen jeder Zeit offen. Insofern freue ich mich schon auf die Einreichungen und Ideen für das nächste DFN-Forum Cloud-Dienste im Herbst 2017.

Rechtsfragen zu Cloud-Angeboten für Hochschulen

Johannes Nehlsen
Stabsstelle IT-Recht
der bayerischen staatlichen Universitäten und Hochschulen
c/o Rechenzentrum Universität Würzburg
Am Hubland
97074 Würzburg
johannes.nehlsen@uni-wuerzburg.de

Abstract: Staatliche Hochschulen sind nicht auf eine Auftragsdatenverarbeitung durch Anbieter mit Sitz im Europäischen Wirtschaftraum (EWR) beschränkt. Auch auf die Anbieter aus Drittstaaten kann zurückgegriffen werden, wenn die Auftragsdatenverarbeitung den gesetzlichen Anforderungen genügt und ein angemessenes Datenschutzniveau gewährleistet ist. Die restriktiven Normen aus den deutschen Datenschutzgesetzen sind europarechtskonform auszulegen. Der bloße Abschluss einer Auftragsdatenverarbeitung allein ist aber noch keine Grundlage zur Datenerhebung, sodass es daneben immer auch einer Einwilligung oder anderen rechtlichen Grundlage bedarf. Hochschulintern sind zudem der datenschutzrechtliche Freigabeprozess, die Mitbestimmung durch den Personalrat und gesetzliche wie vertragliche Pflichten zur Geheimhaltung bei der Dienstnutzung zu beachten.

1 Hintergrund

Die Digitalisierung stellt die Hochschulen und Forschungseinrichtungen vor die große Herausforderung ihre IT im Idealfall ohne zu große zusätzliche finanzielle Belastungen zu modernisieren und weiter zu professionalisieren.

Zum einen sind die Ansprüche der Lehrenden und Studierenden gestiegen. Sie erwarten von den Hochschulen, dass sie ihnen sämtliche notwendige Software und Services komfortabel und jederzeit an jedem Ort zur Verfügung stellen. Fragen zum Studium werden als Videoantworten gewünscht und Lehre und Prüfungen sollen ortsunabhängig jederzeit online, multimedial und in ansprechender Form möglich sein.

Zum anderen werden Verwaltungsvorgänge digitalisiert, während mehr und mehr rechtliche Vorgaben es erfordern, IT-Sicherheit auf hohem Niveau zu gewährleisten.

Bei so vielen Wünschen und Anforderungen sind schnell die personellen Ressourcen und Kapazitäten der Serverräume wie auch der technischen Infrastruktur universitärer Rechenzentren erschöpft.

Um auch der politisch gewollten Digitalisierung der Hochschulen gerecht zu werden, liegt es eigentlich nahe, geeignete Dienste an Drittanbieter auszulagern. Doch genau hier stoßen Hochschulen und Forschungseinrichtungen an rechtliche Grenzen. Denn bei dem Versuch Antworten auf die Frage nach der rechtlichen Zulässigkeit bei der "Cloud" aus Drittstaaten zu erhalten, heißt es schnell: „Schwierig bis unmöglich“. Das mag der sicherste Weg sein, doch auch der Schritt in diese Cloudangebote kann rechtssicher gelingen.[1] Dafür leistet dieser Beitrag eine Hilfe.

[1] In diese Sinne auch von dem Bussche in: Plath, BDSG/DSGVO, 2. Aufl. 2016, § 4b BDSG, Rn. 18.

2 Grundvoraussetzungen

Die Datenschutzgesetze sollen verhindern, dass der Umgang mit personenbezogenen Daten zu einer Beeinträchtigung des Persönlichkeitsrechtes von Betroffenen führt.[2]

Im Anwendungsbereich liegen damit nur personenbezogene Daten Betroffener.[3] Dies sind alle Daten, die einen Personenbezug ermöglichen können, d.h. eine Person identifizierbar machen könnten. Dafür ist auch Zusatzwissen Dritter relevant, das vernünftiger Weise herangezogen werden kann, insbesondere öffentlich zugängliche Informationen und gesetzlich oder vertraglich rechtmäßig zustehende Auskunftsrechte.[4] Daher empfiehlt es sich in Zweifelsfällen einen Personenbezug anzunehmen.

Die deutschen Datenschutzgesetze verlangen bei jedem Umgang mit personenbezogenen Daten entweder eine gesetzliche Erlaubnis oder eine Einwilligung.[5] Fehlt es an einer dieser alternativen Voraussetzungen, ist der Umgang mit personenbezogenen Daten unzulässig.[6]

Darüber hinaus ist neben dem Gebot der Datensparsamkeit[7] insbesondere der Zweckbindungsgrundsatz von entscheidender Bedeutung. Dieser besagt: Bereits mit der Erhebung von personenbezogenen Daten ist im Regelfall der Zweck der Datenverarbeitung festzulegen.[8]

Betroffene haben um den Umgang mit Ihren Daten kontrollieren zu können, besondere Rechte gegenüber den Stellen, die für den Umgang mit den Daten verantwortlich sind.

Diese Rechte umfassen Auskunft, Berichtigung, Löschung oder Sperrung der Daten,[9] deren Ausübung die Datenschutzbehörden für Betroffene im Konfliktfall durchsetzen können.[10] Insbesondere vor diesem Hintergrund sind die Gesetze so ausgestaltet, dass die verantwortliche Stelle grundsätzlich nicht die Hoheit über die Daten verlieren soll. Soweit die verantwortliche Stelle im Rahmen ihrer Tätigkeit Dritte einspannt, diese aber in Kontakt mit personenbezogenen Daten kommen, bedarf es einer vertraglichen Absicherung für den Umgang mit den Daten. Diese vertragliche Absicherung ist in den Gesetzen als Auftragsdatenverarbeitung ausgestaltet.[11] Eben dieser Vertragstyp ermöglicht es in vielen Fällen auch Dienste aus der Cloud zu nutzen. Gleichzeitig bleibt die Art des Einsatzes in der Verantwortung des Auftraggebers, denn die Verarbeitung und Nutzung von Daten erfolgt durch Weisungen.

Die Alternative zur Auftragsdatenverarbeitung, die Funktionsübertragung, bringt nur dann eine Erleichterung, wenn die Daten nicht mehr von der ursprünglich verantwortlichen Stelle verarbeitet oder genutzt werden müssen, da anderenfalls jeder weitere Daten(rück)fluss einer erneuten Legitimation bedürfte,[12] und somit kaum einen Anwendungsbereich für Dienstleistungen aus einer Public-Cloud schaffen kann.

---

[2] § 1 BDSG, Art. 1 BayDSG.
[3] § 2 Abs.1 BDSG, Art. 4 Abs. 1 BayDSG.
[4] EuGH, Urteil vom 19.10.2016 - C-582/14.
[5] §. 4 Abs. 1 BDSG, Art. 15 Abs. 1 BayDSG.
[6] Statt aller BeckOK DatenSR/Bäcker BDSG § 4 Rn. 21.
[7] § 3a BDSG.
[8] Allgemeiner Grundsatz, spätestens seit BVerfGE 65, 1–71.
[9] § 6 BDSG, Art. 10–13 BayDSG.
[10] § 38 BDSG, Art. 31 BayDSG.
[11] § 11 BDSG, Art. 6 BayDSG.
[12] Gola/Schomerus/Gola/Klug/Körffer BDSG § 11 Rn. 9, beck-online; Ehman in Wilde/Ehmann/Niese/Knoblauch Datenschutz in Bayern Art. 6 BayDSG Rn. 10.

Damit wird auch deutlich, dass ein Transfer von Daten in Staaten, die Betroffenen keinen gleichartigen Datenschutz gewähren wie das Land mit dem Sitz der verantwortlichen Stelle es bietet, nur unter besonderen Vorrausetzungen zulässig sein kann.[13]

Aufgebrochen wird dieses Konzept durch die Verträge zur Europäischen Union und die darauf aufsetzende europäische Regulierung, die auch den Handel mit Daten deutlich in den Vordergrund rückt.[14] Die zukünftige Datenschutzgrundverordnung räumt zwar dem Datenschutz mehr Gewicht ein,[15] bleibt aber dem Grundsatz treu, dass durch die Regulierung Datenhandel ermöglicht und aufrechterhalten werden soll.[16] Die Mitgliedschaft in der europäischen Union hat für die Rechtsanwendung zur Folge, dass diese die bestmögliche Zur-Geltung-Bringung von Unionsrecht erzielen muss.[17]

Vor diesem Hintergrund ergibt sich in letzter Konsequenz, dass, soweit unionsrechtlich eine Verarbeitung von personenbezogen Daten in Auftrag zulässig ist, diese auch im deutschen Recht zu ermöglichen ist; gegebenenfalls unter zu Hilfenahme aller zulässigen Auslegungsmethoden.[18] Einzig, wenn das Unionsrecht den Mitgliedsstaaten eine Abweichung zubilligt oder außerhalb seines Anwendungsbereiches liegt, greift dieser Mechanismus bei der Umsetzung der Datenschutzrichtlinie nicht ein.[19]

3 Europarechtliche Betrachtung

3.1 Richtlinienbeeinflusster Bereich der Regelung

Zunächst ist festzuhalten, dass die Richtlinie 95/46/EG (künftig bezeichnet als „Datenschutzrichtlinie“) keine ausdrückliche Regelung zu der Frage hat, ob mit ihr nur ein Mindeststandard im Datenschutz sichergestellt werden soll oder eine Vollharmonisierung angestrebt ist. Mit Blick auf die häufige Verwendung von Worten wie „zumindest“ bei Informations-[20], Auskunfts-[21] und Widerspruchsrechte[22] liegt eher eine Mindestharmonisierung nahe. Ebenso ist für die Mitgliedsstaaten ein Spielraum vorgesehen.[23] Andererseits legt der Erwägungsgrund 7 der Datenschutzrichtlinie eher nahe, dass unterschiedlich hohe Schutzniveau der Zielerreichung der Richtlinie, nämlich den Handel mit Daten zu erleichtern, entgegenstehen würde.

Die Rechtsprechung des EuGH hat sich diesbezüglich seit 2003[24] hin zu einer durch die Richtlinie verfolgten umfassenden Harmonisierung positioniert.

[13] Art. 4c BDSG, Art. 21 BayBSG.
[14] Art. 16 Abs. 2 AEUV; Erwägungsgrund 56 Richtlinie 95/46/EG.
[15] Erwägungsgrund 1 Verordnung (EU) 2016/679.
[16] Erwägungsgrund 101; Art. 1 Abs. 3 Verordnung (EU) 2016/679.
[17] Art. 4 Abs. 3 EUV; Groeben, von der /Schwarze/Walter Obwexer EUV Art. 4 Rn. 116–119.
[18] Calliess/Ruffert/Ruffert, 5. Aufl. 2016, AEUV Art. 1 Rn. 24; Herresthal EuZW 2007, 396 (400).
[19] Grabitz/Hilf, Das Recht der EU, Vorbemerkung: Datenschutz und die Europäische Gemeinschaft Rn. 45–52, beck-online.
[20] Art. 11 Datenschutzrichtlinie.
[21] Art. 12 Datenschutzrichtlinie.
[22] Art. 14 Datenschutzrichtlinie.
[23] Erwägungsgrund 9 der Datenschutzrichtlinie.
[24] EuGH Urteil vom 6. November 2003 – Rs. C-101/01; EuGH Urteil vom 24. November 2011 – Rs. C‑468/10 und C‑469/10.

Damit geht einher, dass sich der den Mitgliedstaaten zur Verfügung stehende Spielraum nur auf die Ausgestaltung bezieht, jedoch nicht den Umfang einer zulässigen Datenverarbeitung begrenzen darf.[25] Die Begriffsdefinitionen der Datenschutzrichtlinie sind erst recht bindet, da anderenfalls die angestrebte Rechtsangleichung hinfällig wäre.

3.2 Datenschutzrichtlinie

Die Datenschutzrichtlinie definiert den Dritten in Art. 2:

> *„"Dritter" die natürliche oder juristische Person, Behörde, Einrichtung oder jede andere Stelle, außer der betroffenen Person, dem für die Verarbeitung Verantwortlichen, dem Auftragsverarbeiter und den Personen, die unter der unmittelbaren Verantwortung des für die Verarbeitung Verantwortlichen oder des Auftragsverarbeiters befugt sind, die Daten zu verarbeiten;"*

Nach dieser Definition ist ein Auftragsverarbeiter im Rahmen der Auftragstätigkeit niemals Dritter. Der Sitz des Auftragnehmers ist relevant für die Frage, ob eine Übermittlung in Drittstaaten erfolgt. Aus Art. 4 Abs. 2 wird ersichtlich, dass es für Anbieter aus Drittstaaten nur erforderlich ist, einen in einem Mitgliedstaat der Europäischen Union ansässigen Vertreter zu benennen, falls Daten im Hoheitsgebiet vorgehalten werden.

Wann welche nationalen Datenschutzgesetze Anwendungen finden, ist in seinen Feinheiten komplex ausgestaltet und die Auswirkungen aus der Rechtsprechung des EuGH[26] wohl noch nicht gänzlich durchdrungen. Im Grundprinzip entscheidet sich das „ob" nach dem anwendbaren Recht für die Behörde, das „wie" nach dem Datenschutzrecht der europäischen Niederlassung des Cloud-Anbieters.[27]

Die Richtlinie enthält keine Definition für den Begriff der Übermittelung.[28] Für die Frage, ob eine Übermittlung vorliegt, spielt es keine Rolle, ob Daten zu einem „Dritten" transferiert werden.[29] Der Begriff ist somit im Kontext der Definition des Verarbeitens aus Art. 2 lit. b vielmehr eher technisch zu verstehen, jedoch bedarf es Eingrenzungen, um einen uferlosen Anwendungsbereich der Richtlinie zu verhindern.[30]

Art. 25 ermöglicht die Übermittlung personenbezogener Daten in Drittstaaten, wenn dort ein angemessenes Schutzniveau gewährleistet ist.

Auch ohne dass dieses vorliegt, können Daten bei Vorliegen geeigneter Garantien in Drittstaaten übermittelt werden gemäß Art. 26 Abs. 2. Als eine solche Garantie werden auch die von der EU-

[25] Grabitz/Hilf, Das Recht der EU, Vorbemerkung: Datenschutz und die Europäische Gemeinschaft Rn. 45–52; Hören, RDV 2009, 89 (94f).
[26] Z. B. EuGH, Urteil v. 13.05.2014, Az. C-131/12.
[27] Art. 3 Abs. 1 b) Datenschutzrichtlinie.
[28] EuGH Urteil vom 6. November 2003 - Rechtssache C-101/01 Rn. 56.
[29] Grabitz/Hilf, Das Recht der EU, Art. 2 Rn. 24, beck-online.
[30] EuGH Urteil vom 6. November 2003 - Rechtssache C-101/01 Rn. 56.

Kommission beschlossen Standardvertragsklauseln angesehen.[31] Diese liegen in drei verschiedenen Fassungen vor,[32] wobei im Bereich Cloud Computing die 2010 verabschiedeten Klauseln am weitesten verbreitet sind.[33]

Eine Alternative für die Praxis zu den Standardvertragsklauseln den Datentransfer in die USA zu ermöglichen, ist das EU-U.S. Privacy Shield.[34] Ist ein Unternehmen auf der Datenschutzschild-Liste gemäß Art. 1 Abs. 3 Durchführungsbeschluss (EU) 2016/1250 aufgeführt, bedarf es zwar immer auch noch des Abschlusses einer Vereinbarung über eine Datenverarbeitung im Auftrag,[35] die Standardvertragsklauseln müssen aber dafür nicht zwingend verwendet werden.

Soweit Daten unternehmensintern auch in Drittstaaten übertragen werden, können auch sogenannte Binding Corporate Rules einen angemessenen Datenschutz gewährleisten. Die Anzahl an Unternehmen mit von den Datenschutzbehörden genehmigten Binding Corporate Rules ist überschaubar.[36]

Für die Auftragsdatenverarbeitung sieht die Datenschutzrichtlinie vor, dass diese durch Dienstleister auf der ganzen Welt erfolgen kann. Außerhalb des EWR sind aber die genannten zusätzlichen Schutzmechanismen, [37] z. B. der Einsatz von Standardvertragsklauseln, notwendig. Ein Unternehmen mit weltweiten Niederlassungen kann seinen internen Datentransfer mittels Binding Corporate Rules absichern.

3.3 Standardvertragsklauseln

Soweit der Einsatz von Standardvertragsklauseln nach dem EU-Kommissions-Beschluss 2010/87/EU beabsichtigt ist, bedürfen zwei Punkte einer besonderen Berücksichtigung.

Der Mustervertrag darf gemäß Klausel 10 Standardvertragsklauseln (Auftragsverarbeiter) grundsätzlich nicht verändert werden. Deshalb ist es hilfreich, wenn zum Beispiel im Vertrag sichergestellt ist, dass die Standardvertragsklauseln in Zweifelsfällen stets vorrangig sind.

Eine Möglichkeit sich vor unzulässigen Abänderungen der Standardvertragsklauseln zu schützen, kann ähnlich den Mustern der „bitkom“ zur Auftragsdatenverarbeitung nach BDSG gelingen mit Formulierungen im Vertrag wie:

[31] Grabitz/Hilf, Das Recht der EU, Art. 26 Rn. 21, beck-online; BeckOK DatenSR/Schantz BDSG § 4c Rn. 42–46, beck-online.

[32] EU-Kommission, Entscheidung v. 15.6.2001 (K(2001) 1539); EU-Kommission, Entscheidung v. 27.12.2004 (K(2004) 5271) und EU-Kommissions-Beschluss 2010/87/EU.

[33] So nutzen z. B. Amazon http://www.cnpd.public.lu/en/actualites/international/2015/03/AWS/index.html , Dropbox https://www.dropbox.com/privacy#business_agreement und Microsoft https://www.microsoft.com/en-us/TrustCenter/Compliance/EU-Model-Clauses diese Klauseln.

[34] Durchführungsbeschluss (EU) 2016/1250 der Kommission.

[35] So auch explizit unter 10. a. i. Anhang II „Grundsätze des EU-US-Datenschutzschilds vorgelegt vom amerikanischen Handelsministerium“ Durchführungsbeschluss (EU) 2016/1250 der Kommission.

[36] Eine Liste der Unternehmen ist abrufbar unter: http://ec.europa.eu/justice/data-protection/international-transfers/binding-corporate-rules/bcr_cooperation/index_en.htm .

[37] Art. 26 Datenschutzrichtlinie; Grabitz/Hilf, Das Recht der EU, Vorbemerkung zu Art. 26, beck-online.

Deutsch: Bei etwaigen Widersprüchen gehen Regelungen der vereinbarten EU-Standardvertragsklauseln den Regelungen des geschäftlichen Vertrages vor.

English: In case of any conflict, the regulations of the EU-Model Clauses as contractually agreed shall take precedence over the regulations of the Business Agreement.

Der zweite kritische Punkt sind die Anhänge, die für eine wirksame Auftragsdatenverarbeitung für den jeweiligen Vertrag passend und vollständig auszufüllen sind. Nicht jeder Cloud-Anbieter bietet sie bereits vollständig vorausgefüllt an. Sie können jedoch nur mit Kenntnis der technischen Infrastruktur und Sicherheitsfeatures des Anbieters vervollständigt werden.

3.4 Artikel 29-Gruppe

In der Stellungnahme[38] der Artikel 29-Gruppe, deren Aufgabe es ist, öffentliche Stellungnahmen zu Datenschutzfragen aus unionsrechtlicher Sicht abzugeben, werden nur Anforderung für Cloud Computing festgelegt. Von einer grundsätzlichen Zulässigkeit kann damit ausgegangen werden.

3.5 Datenschutzgrundverordnung

Auch die DSGVO ändert diese Systematik nicht. Im Unterschied zum bisherigen Recht, ist jedoch nicht der jeweils nationale Gesetzgeber gehalten, Auftragsdatenverarbeitung (nach dem Termini der DSGVO Verarbeitung im Auftrag) zu ermöglichen, sondern die Verarbeitung im Auftrag ist durch Unionrecht unmittelbar vorgesehen.[39]

Neu ist, dass nun die Anforderungen an die Sicherheit der Verarbeitung einheitlich vorgegeben sind[40] und zur Prüfung der Einhaltung dieser Anforderung auch Zertifizierungen als Faktor mitherangezogen werden können.[41]

Die bisherigen Entscheidungen der Kommission, wie z. B. zur Angemessenheit des Datenschutzes in Drittstaaten, zum EU-US Privacy Shield oder zu den Standardvertragsklauseln, gelten bis zum Erlass neuer Durchführungsrechtsakte fort.[42]

Auch das Gesetz zur Anpassung des Datenschutzrechts an die Verordnung (EU) 2016/679 und zur Umsetzung der Richtlinie (EU) 2016/680 (Datenschutz-Anpassungs- und -Umsetzungsgesetz EU - DSAnpUG-EU) baut auf den Definition der DSGVO auf.

---

[38] Abrufbar z. B. unter: https://www.lda.bayern.de/media/wp196_de.pdf .

[39] Hier hat die DSGVO eine dogmatische Schwäche. So kann die Verarbeitung im Auftrag gemäß Art. 28 DSGVO als unionsrechtliche Rechtsgrundlage gemäß Art. 6 Abs. 3 lit. a DSGVO für Art. 6 Abs. 1 lit. c DSGVO angesehen werden (so wohl Ehmann in: Datenschutz in Bayern, Kommentar Art. 28 DSGVO S. 1f) oder auch das jeder Form der Verarbeitung nach Art. DSGVO auch als Verarbeitung im Auftrag möglich ist (so z. B. Plath in: Plath, BDSG/DSGVO, 2. Aufl. 2016, Artikel 28 DSGVO, Rn. 3). Dass wegen der der fehlenden rechtlichen Eindeutig, die Verarbeitung im Auftrag nicht möglich sei, wird nicht ernsthaft vertreten. Dies zeigt sich auch bei der Auslegung von Schmidt/Freund ZD 2017, 14 (14–16).

[40] Art. 32 DSGVO; Eine Abweichung von diesen Anforderungen sieht auch die Öffnungsklausel in Art. 6 Abs. 2, 3 DSGVO nicht vor. Siehe auch Plath in: Plath, BDSG/DSGVO, 2. Aufl. 2016, Artikel 6 DSGVO, Rn. 25.

[41] Art. 28 Abs. 5, 6 DSGVO; siehe auch Plath in: Plath, BDSG/DSGVO, 2. Aufl. 2016, Artikel 28 DSGVO, Rn. 15.

[42] Art. 45 Abs. 9; 46 Abs. 5 DSGVO.

4 Betrachtung der Auftragsdatenverarbeitung nach BDSG

4.1 Auslegung am Gesetzestext BDSG

Das BDSG folgt einer anderen Systematik als die Datenschutzrichtlinie und enthält eigene Begriffsdefinitionen.

Ausgangspunkt sind die im Gesetz angelegten Definitionen sowie die Gesetzessystematik. Jeder Umgang mit Daten bedarf einer gesetzlichen Grundlage (Rechtsvorschrift) oder einer Einwilligung.[43] Das Gesetzt unterscheidet die Phasen des Umgangs mit personenbezogenen Daten in Erhebung, Verarbeitung und Nutzen.[44]

Übermitteln ist das Bekanntgeben gespeicherter oder durch Datenverarbeitung gewonnener personenbezogener Daten an einen Dritten, § 3 Abs. 4 Nr. 1 BDSG.

Nach § 3 Abs. 6 S. 2 BDSG sind Dritte nicht der Betroffene sowie Personen und Stellen, die im Inland, in einem anderen Mitgliedstaat der Europäischen Union oder in einem anderen Vertragsstaat des Abkommens über den Europäischen Wirtschaftsraum personenbezogene Daten im Auftrag erheben, verarbeiten oder nutzen.

Cloudcomputing ist somit möglich, wenn die Tätigkeit des Cloudanbieters im Rahmen einer Auftragsdatenverarbeitung im EWR erfolgt. Eine Privilegierung der Auftragsdatenverarbeitung in Drittstaaten ist nicht vorgesehen. Vielmehr ist die Übermittlung nur in den engen Ausnahmen der §§ 4b, 4c BDSG möglich.[45]

Neben der Frage, ob überhaupt übermittelt werden darf, muss zudem die Übermittlung auf Basis der Einwilligung der Betroffenen oder einer Rechtsvorschrift erfolgen.[46]

Da bei einer Einwilligung aber das Risiko besteht, dass diese nicht wirksam erklärt worden ist und sie zudem jederzeit widerrufen werden kann, eignet sich eine Einwilligung nur selten als taugliche Grundlage für einen Datentransfer im Rahmen einer Auftragsdatenverarbeitung.[47]

Und für Behörden gibt es, anders als für nicht öffentliche Stellen, keine Generalklausel ähnlich dem § 28 Abs. 1 S. 1 Nr. 2 BDSG, über die ein Datentransfer ohne Einwilligung möglich wäre.[48]

Im Ergebnis ist öffentlichen Stellen (soweit nur nach Wortlaut und Systematik betrachtet) die Wahl eines Cloudanbieters aus Drittstaaten verwehrt. Einzig günstig an dieser Betrachtung ist, dass der Serverstandort des Anbieters in keiner Weise maßgeblich wäre, denn innerhalb einer verantwortlichen Stelle kann nicht „übermittelt“ werden.

[43] § 4 Abs. 1 BDSG; statt aller Gola/Schomerus/Körffer/Gola/Klug BDSG § 4 Rn. 3 beck-online.
[44] § 3 Abs. 3–5 BDSG.
[45] Von dem Bussche in: Plath, BDSG/DSGVO, 2. Aufl. 2016, § 4b BDSG, Rn. 16f.
[46] § 4 Abs. 1 BDSG; statt aller Gola/Schomerus/Körffer/Gola/Klug BDSG § 4 Rn. 3 beck-online.
[47] Simitis in: Simitis Bundesdatenschutzgesetz (2014), § 4a Rn. 94; Borges, Borges/Meents Cloud Computing (2016), S. 287 Rn. 35.
[48] § 28 BDSG steht im dritten Abschnitt „Datenverarbeitung nicht-öffentlicher Stellen und öffentlich-rechtlicher Wettbewerbsunternehmen“. Der Umgang mit personenbezogen Daten soll für Unternehmen flexibler sein als für Behörden; vgl. Simitis in: Simitis Bundesdatenschutzgesetz, § 27 Rn. 2.

4.2 Auffassung der Datenschutzbehörden

Die Arbeitskreise Technik und Medien der Konferenz der Datenschutzbeauftragten des Bundes und der Länder haben eine „Orientierungshilfe – Cloud Computing“[49] erstellt.

Diese Position bleibt nahe an Wortlaut und Systematik, ermöglicht allerdings für **nicht öffentliche Stellen** die Datenübermittlungen außerhalb des EWR über § 28 Abs. 1 S. 1 Nr. 2 BDSG,[50] sofern die Anforderung des § 4c BSDG erfüllt sind; d.h., unter anderem nur mit Einwilligung oder Garantien wie nach den Standardvertragsklauseln. Als Alternative kann nach § 4b Abs. 2 BDSG ein angemessenes Datenschutzniveaus festgestellt werden, in der Regel durch Entscheidung der EU Kommission.[51]

Bezogen auf die Landesdatenschutzgesetze heißt es dort:

> „Soweit öffentliche Stellen Cloud Services in Drittstaaten anwenden, ist hier eine besonders sorgfältige Prüfung geboten, denn ein dem § 28 Abs. 1 Satz 1 Nr. 2 BDSG entsprechender Erlaubnistatbestand dürfte es in den Landesdatenschutzgesetzen nicht geben, soweit ersichtlich. Die Verfasser dieser Orientierungshilfe haben allerdings keine Prüfung aller Landesdatenschutzgesetze vorgenommen.“

Somit trifft die Orientierungshilfe keine Aussage darüber, ob Cloud-Computing für Behörden außerhalb des Europäischen Wirtschaftsraumes möglich ist.

4.3 Weitere Auffassungen

4.3.1 Restriktive Auslegungen

Einige Auffassungen bleiben beim Wortlaut und sind dabei teilweise noch einschränkender bei der Interpretation berechtigter Interessen im Rahmen von § 28 Abs. 1 S. Nr. 2 BDSG.

So sieht Simitis[52], dass § 28 Abs. 1 S. 1 Nr. 2 BDSG nur in Ausnahmenfällen eine Auftragsdatenverarbeitung rechtfertigt. Dammann[53] sieht eine europarechtliche Interpretation des Begriffes „Übermittlung“ nicht veranlasst, und lehnt eine analoge Anwendung von Art. 3 Abs. 8 BDSG ab.

Nach Weber und Voigt[54] sei Durchführung einer Verarbeitung im Auftrag von der Datenschutzrichtlinie auf den EWR beschränkt.

4.3.2 Rechtfertigung nach § 28 Abs. 1 S. 1 Nr. 2 BDSG

Ein großer Anteil der Stimmen wie die Landesdatenschutzbehörden (siehe 4.2), aber auch in der Literatur sieht die Möglichkeit der Rechtfertigung über § 28 Abs. 1 S. 1 Nr. 2 BDSG bei Cloudanbietern aus Drittstaaten. Erforderlich sind also Einwilligungen, alternativ ein angemessenes Datenschutzniveau oder ausreichenden Garantien vorliegen (siehe 3.2 und 3.5).[55]

---

[49] Abrufbar z. B. unter: https://www.datenschutz-bayern.de/technik/orient/oh_cloud.pdf .
[50] Orientierungshilfe – Cloud Computing, S. 16.
[51] Vgl. z. B. BeckOK DatenSR/Schantz BDSG § 4b Rn. 25–35, beck-online.
[52] Simitis in: Simitis Bundesdatenschutzgesetz (2014), § 28 Rn. 101.
[53] Dammann in: Simitis Bundesdatenschutzgesetz (2014), § 3 Rn. 246.
[54] Weber/Voigt, ZD 2011, 74 (77). Die sich jedoch Grundsätzlich die Möglichkeit einer internationalen Auftragsdatenverarbeitung bejahen über eine Analogie zu § 3 Abs. 8 BDSG (a.a.O. 78).
[55] Nachweise finden sich z. B. bei Borges, Borges/Meents Cloud Computing (2016), S. 232f Rn. 10.

4.3.3 Richtlinienkonforme Anwendung § 3 Abs. 8 BDSG

Neben einer unmittelbaren Anwendung der Richtlinie[56] wird im Hinblick auf die europarechtlichen Vorgaben entweder eine teleologische Reduktion vertreten oder eine analoge Anwendung der Norm bei der Verarbeitung personenbezogener Daten im Auftrag mit Sitz des Anbieters in Drittstaaten vorgeschlagen.[57]

Diese Autoren erblicken eine planwidrige oder überschießende Regelung in § 3 Abs. 8 BDSG, sowie eine vergleichbare Interessenlage, insbesondere wenn die Standardvertragsklauseln die Diensterbringung zu Grunde liegen. Diese Auslegung sei auch unionsrechtlich geboten.[58]

5 Betrachtung nach BayDSG

Die Begriffsdefinitionen entsprechen – soweit hier relevant – dem BDSG, jedoch sind die Übermittlungsmöglichkeit von Daten auch an nicht öffentliche Stellen im Art. 21 Abs. 2 S. 4 Nr. 3 BayDSG weitergefasst als im BDSG. Allerdings findet sich kein passender Anknüpfungspunkt für eine Übermittlung von Daten im Rahmen von Cloud Computing mit Bezügen zu Drittstaaten, denn mit dieser Norm sollte im Wesentlichen nur die Datenschutzrichtlinie umgesetzt werden[59], die primär für den Fall internationaler Überweisungen gedacht ist[60].

5.1 Auffassung des Bayerischen Landesbeauftragte für den Datenschutz

Im 25. (2012) Tätigkeitsbericht des Bayerischen Landesbeauftragten für den Datenschutz wird bei der Inanspruchnahme von Cloud Diensten äußerste Zurückhaltung angemahnt.[61] Diese Einschätzung wurde im 26. (2014) und 27. (2016) Tätigkeitsbericht aufrechterhalten.[62]

5.2 Weitere Auffassung

Für die Landesdatenschutzgesetze ist die Literaturlage überschaubar. Im Kommentar „Datenschutz in Bayern“ heißt es lapidar: „Gerade ein Cloud Computing unter Einschaltung von Drittstaaten ist für öffentliche Stellen in aller Regel keine realistische Option.“[63]

Die Meinungen und Argumentationen, die für das BDSG vorgebracht werden, sind aber übertragbar, da die Begriffsdefinitionen des BayDSG, soweit hier relevant, denen des BDSG entsprechen.

[56] Kahler, RDV, 2012, 167 ff.
[57] Nachweise bei Borges, Borges/Meents Cloud Computing (2016), S. 233 Rn. 11.
[58] Borges, Borges/Meents Cloud Computing (2016), S. 234 f. Rn. 12.
[59] Bayerischer Landtag, Drucksache 14/3327, S. 14.
[60] Grabitz/Hilf, Das Recht der EU, Art. 26 Rn. 8.
[61] Punkt 2.3.3. abrufbar unter https://www.datenschutz-bayern.de/ .
[62] 26. Tätigkeitsbericht 2014 Punkt 13.1. https://www.datenschutz-bayern.de/tbs/tb26/k13.html#13.1 und 27. Tätigkeitsbericht 2016 Punkt 13.3, abrufbar unter https://www.datenschutz-bayern.de/tbs/tb27/k13.html#13.3 .
[63] Ehmann in: Wilde/Ehemann/Niese/Knoblauch Datenschutz in Bayern, 26. EL – Stand Oktober 2016, Art. 6 BayDSG Rn. 3g.

6 Eigene Stellungnahme

Folgt man den restriktiven Ansichten, steht Behörden oft keine praxistaugliche Möglichkeit zur Verfügung, auf Cloud-Computing von Anbietern außerhalb des EWR zurückzugreifen. Da auch in Unteraufträgen der Schutz des Hauptvertrages nicht unterschritten werden darf, kann nicht auf Leistungen von Konzernen mit Töchtern eigener Rechtspersönlichkeiten zurückgegriffen werden, denn der Datenschutz kennt (noch) kein Konzernprivileg.

Ein Verzicht auf eine Korrektur des Begriffs des Dritten, verbunden mit einer großzügigeren Auslegung des § 28 BDSG, verändert die datenschutzrechtliche Ausgangssituation für Behörden nicht, da es für den öffentlichen Bereich an einer dem § 28 BDSG entsprechenden Vorschrift fehlt.

Ein Verharren auf dem deutschen Begriff der „Übermittlung“ führt aber zu der perplexen Situation, dass innerhalb einer jeweils zuständigen Stelle der Speicherort der Daten nicht relevant wäre, da innerhalb einer Stelle nicht übermittelt werden kann. Somit könnte eine Hochschule Server in Pjöngjang oder Minsk mieten, aber keinen Updateservice und Remotesupportvertrag mit Netzwerkausrüstern aus Drittstaaten, wie HPE oder Cisco, oder mit Softwareanbietern aus Drittstaaten, wie Microsoft oder VMWare, abschließen. Der von der Richtlinie gewollte sicherere Rechtsrahmen, dass, überall dort wo die Daten mit den Daten umgegangen werden, auch ein angemessener Datenschutz gewährleistet wäre, könnte unterlaufen werden.[64]

Um das Schutzniveau der Datenschutzrichtlinie im nationalen Datenschutzrecht umzusetzen, bleibt einzig der Weg die Definition des Dritten unionsrechtlich bedingt zu modifizieren und wie folgt anzuwenden: „Dritte sind nicht der Betroffene sowie Personen und Stellen, die ~~im Inland, in einem anderen Mitgliedstaat der Europäischen Union oder in einem anderen Vertragsstaat des Abkommens über den Europäischen Wirtschaftsraum~~ personenbezogene Daten im Auftrag erheben, verarbeiten oder nutzen.“

Gleichzeitig muss der Begriff der Übermittlung in § 4b und § 4c BDSG, bzw. z. B. Art. 21 BayDSG europarechtlich ohne Bezug auf die Definition im BDSG bzw. BayDSG verstanden werden. Das bedeutet, dass auch z. B. die Miete eines Servers außerhalb des EWR als Übermittlung aufgefasst wird.

Die Begründung für diese beiden Begriffskorrekturen liegt in der Pflicht, das Unionsrecht durch Auslegung des nationalen Rechts bestmöglich zur Geltung zu bringen.[65] Schon mit der Datenschutzrichtlinie wurde in Europa ein gleichwertiges Schutzniveau geschaffen, und die in der Richtlinie vorgesehenen Begriffe stehen im Rahmen der Umsetzung einer Richtlinie dem nationalen Gesetzgeber nicht zur Disposition.[66]

Das Ergebnis der unionsrechtskonformen Auslegung des deutschen Datenschutzrechts ist: Auch öffentliche Behörden sind nicht auf eine Auftragsdatenverarbeitung durch Anbieter mit Sitz im EWR beschränkt. Auch auf die Dienste aus Drittstaaten kann zurückgegriffen werden.

[64] Dies verdeutlicht auch Erwägungsgrund 10 Datenschutzrichtlinie.
[65] Siehe dazu bereits die Ausführungen unter 2.
[66] Gebot effektiver Umsetzung: Herdegen, Europarecht (2016), § 8 Rn. 41–45.

7 Umsetzung

Allein aus der Tatsache, dass es sich juristisch gut vertreten lässt, dass Clouddienste auch bei einem Bezug zu Drittstaaten für Behörden genutzt werden können, bleibt die Herausforderung bestehen die gesetzlichen Anforderungen der Auftragsdatenverarbeitung vollständig umzusetzen. Ohne korrekte Auftragsdatenverarbeitung liegt stets eine unzulässige Datenübermittlung vor.[67]

Im Allgemeinen gelten z. B. für bayerische staatliche Hochschulen und die Akademie der bayrischen Wissenschaften die Anforderungen aus Art. 6 BayDSG, für Universitäten des Bundes und Vereinigungen des Privatrechts die Anforderungen aus § 11 BDSG. Für Religionsgesellschaften finden, insbesondere wenn für sie als Körperschaften des öffentlichen Rechts agieren, weder Bundes- noch Landesdatenschutzgesetze Anwendung, da diese keine des Bundes oder der Länder sind.[68] In Deutschland haben sich die Bistümer der katholischen Kirche, wie auch die evangelischen Landeskirchen Datenschutzordnungen gegeben, die im Wesentlichen Regelungen des BDSG aufgreifen.[69] Im katholischen Datenschutz lassen sich Strafvorschriften bei Datenschutzverstößen wenigstens für gravierende Verstöße im Zusammenspiel von Can. 1399[70] und Can. 220 konstruieren.[71] Ein Vergleich der Landesdatenschutzgesetze zeigt eine Skepsis des Gesetzgebers, sofern die Auftragsdatenverarbeitung bei nicht-öffentlichen Stellen stattfindet.[72]

Zwar findet bereits durch die Datenschutzrichtlinie für den Auftragsdatenverarbeiter im Grundsatz nur das Datenschutzrecht an seiner Niederlassung Anwendung. Aber durch die Wahl der Standardvertragsklauseln wird vertraglich die (eingeschränkte) Anwendung des Datenschutzrechts des Auftraggebers vereinbart. Dies führt für internationale Anbieter zu allein in Deutschland mindestens 19 unterschiedlichen anwendbaren Datenschutzgesetzen.

7.1 Absicherung internationaler Datentransfers

Die Rechtsprechung des Europäischen Gerichtshofs erfordert, dass Beschlüsse für die Angemessenheit des Datenschutzniveaus in Drittstaaten regelmäßig zu überprüfen sind.[73] Da die politischen Voraussetzungen nicht immer berechenbar sind, ist es nicht ratsam einen internationalen Datentransfer nur minimalistisch abzusichern. Zudem sind nicht alle Entwicklungen der Rechtsprechung vorhersehbar.

Sollten also Datentransfers nur auf Basis der Standardvertragsklausen oder einer Auftragsdatenverarbeitung mit EU-US Privacy Shield abgesichert sein, könnte es durch ein einziges Urteil des EUGH veranlasst sein, den Datentransfer zu stoppen. Dies ist auch mit ein Grund dafür, trotz der Verwendung von Standardvertragsklauseln zusätzlich noch eine Auftragsdatenverarbeitung in Verträge aufzunehmen.

---

[67] Diesbezüglich, ergingen auch bereits Urteile: z. B. VG Wiesbaden, DuD 2015, 262–265; auch haben Datenschutzbehörden erste Bußgelder für Verträge mit inhaltlichen Mängeln festgesetzt. https://www.lda.bayern.de/media/pm2015_11.pdf .

[68] Diese Frage ist umstritten. Vgl. Gola/Schomerus/Körffer/Gola/Klug BDSG § 2 Rn. 14a m.w.N., sowie Preuß ZD 2015, 217 (218 ff.).

[69] Preuß ZD 2015, 217 (223).

[70] Ausführlich dazu Max Ortner, Die Entwertung des Gesetzlichkeitsprinzips und des Analogieverbotes durch die Generalnorm des Kanon 1399 des CIC/1983 (2017).

[71] Diese Besonderheit des kirchlichen Strafrechts wird häufig übersehen, so wohl z. B. Preuß ZD 2015, 217 (223).

[72] Vgl. z. B. § 11 Abs. 3 DSGNRW oder § 4 Abs. 4 RLP LDSG.

[73] EuGH, Urteil vom 06.10.2015, C-362/14 (Schrems) Rn. 76.

7.2 Mindestinhalte der Verträge bei internationalen Datentransfers

Durch die Standardvertragsklauseln, die auch kompatibel zu den in Deutschland anwendbaren Datenschutzgesetzen ausgestaltet werden können, hat sich ein auch bei den Auftragsdatenverarbeitern akzeptiertes Vertragsmuster etabliert. Besonderheiten der deutschen Landesdatenschutzgesetze und des BDSG können in der Anlage 1 unter dem Punkt Datenverarbeitung mit Gegenstand, Dauer, Umfang, Ort und Zweck, Unteraufträgen, Weisungsbefugnissen und Rückgabe überlassene Datenträger und Löschung von Daten abgebildet werden.[74] Durch die Integration dieser Anforderungen wird im Regelfall auch den Anforderungen der Landesdatenschutzgesetze genügt.

Eine Besonderheit für Behörden, sofern auf die Kontrollen aus der Anlage zu § 9 BDSG zurückgegriffen wird, ergibt sich mit Blick auf die Datenschutzgesetze aus dem Erfordernis einer zusätzlichen Kontrolle hinsichtlich der Organisation des Datenschutzes, so z. B. Art. 7 Abs. 2 Nr. 10 BayDSG. Auftragsdatenverarbeiter, die diese zusätzlich zu den Kontrollen, wie sie das BDSG kennt, aufnehmen, erleichtern den behördlichen Datenschutzbeauftragten so ihre Prüfung, hinsichtlich der Übereinstimmung mit den jeweils anzuwendenden Landesdatenschutzgesetzen.

Auch wenn viele Landesdatenschutzgesetze, wie auch das Standarddatenschutzmodell, sich nun an Zielen statt an allgemeinen Maßnahmenkatalogen orientieren, dürfte in vielen Fällen ein Vertrag, in dem die Kontrollen des BDSG gewissenhaft umgesetzt sind, für viele Anwendungsfälle ausreichen.

Bis es genehmigte Verhaltensregeln und Datenschutzzertifizierungen nach der DSGVO geben wird, liegt es nahe, sich ab Geltung der DSGVO an Kontrollen aus ISO 27002 zur Erfüllung der technischen und organisatorischen Maßnahmen zur Datensicherheit zu orientieren.[75]

7.3 Datenschutzbeauftragter

Für einen Auftragsdatenverarbeiter in Deutschland ist ein Datenschutzbeauftragter obligatorisch.[76] In Drittstaaten und teilweise sogar im EWR ist dieser erst mit der DSGVO für nahezu jeden Anbieter vorgeschrieben.[77] Daher ist es zu empfehlen, vertraglich gegenseitig die Datenschutzbeauftragten als Kontaktpersonen festzuhalten und über deren Wechsel und Verhinderungen zu informieren. Fehlt bei einem internationalen Auftragsdatenverarbeiter die Position des Datenschutzbeauftragten, bedarf es vor Geltung der DSGVO eines adäquaten Ersatzes durch den die Einhaltung und Durchsetzung der Pflichten gewährleistet werden kann. Ob dies dann auch den behördlichen Datenschutzbeauftragten bei seiner unabhängigen Entscheidung über die Freigabe des Verfahrens ausreicht, bleibt aber ungewiss.

[74] So von dem Bussche in: Moos, Datennutzungs- und Datenschutzverträge, 1. Aufl. 2014, III. EU-Standardvertragsklauseln für die Übermittlung personenbezogener Daten an Auftragsverarbeiter in Drittländern.

[75] Dieses Vorgehen empfiehlt z. B. die renommierte ActiveMind AG in ihrem Mustervertrag zur Auftragsverarbeitung nach DSGVO, vgl. https://www.activemind.de/datenschutz/dokumente/av-vertrag/ .

[76] § 11 Abs. 4 Nr. 2 i.V.m. § 4 f BDSG.

[77] Art. 37 DSGVO.

7.4 Stand der Technik

Als Grundsatz aus dem deutschen Vertragsrecht gilt, dass, sofern nichts vereinbart ist, nur eine Leistung mittlerer Art und Güte geschuldet ist.[78] Wenn die datenschutzrechtlichen Anforderungen des Auftragsgebers nicht nur Mittelmaß sondern den Stand der Technik verlangen, wird dies ohne vertragliche Regelung vom Auftragnehmer nicht geschuldet. Technische und organisatorische Datenschutzmaßnahmen erfordern aber stets die Berücksichtigung des Stands der Technik.[79] Ein Abweichen nach unten – was ein höheres Risiko für die Sicherheit der Daten mit sich bringt – ist akzeptierbar, wenn die gewählte Maßnahme mit Blick auf die Eintrittswahrscheinlichkeit und Schwere des Risikos für die Rechte und Freiheiten natürlicher Personen dennoch ein angemessenes Schutzniveau gewährleistet. In diese Abwägung werden auch Implementierungskosten sowie Art, Umfang, Umstände und Zwecke der Verarbeitung mit einbezogen. Entscheidend ist, dass es einer Begründung für ein Abweichen nach unten bedarf. Hier könnten ferner auch weitergehende Interessen wie Informationssicherheit oder der Schutz von Knowhow und Geheimnissen miteinfließen.

Gegenüber großen Cloudanbietern prüfen die Aufsichtsbehörden die Verträge. Bei Services sehen sie die Anforderungen des europäischen Datenschutzes hinsichtlich der Auftragsdatenverarbeitung oft als erfüllt an, sofern umfassende informationssicherheitsbezogene Zertifizierungen bestehen.[80] Bei Anbietern von Infrastructure as a Service (Iaas) sind jedoch teilweise die technischen und organisatorischen Maßnahmen nicht Gegenstand der Prüfung der Aufsichtsbehörden.

Auch der neue Standard für Bundesbehörden für die Auswahl von Cloudanbietern C5, der durch das BSI gemäß § 8 Abs. 1 S. 1 BSIG festgelegt worden ist, bietet anderen Behörden eine verlässliche Orientierung, dass bei Anbietern mit diesem Testat, die Datenverarbeitung unter der erforderlichen Berücksichtigung des Stands der Technik erfolgen kann.

8 Private Nutzung von Clouddiensten am Beispiel Office 365

Soweit Hochschulen Studierenden oder Beschäftigten zu privaten Zwecken Clouddienste von Anbieter aus Drittstaaten wie z. B. Microsoft Office 365 ohne oder gegen nur geringe Entgelte zur Verfügung stellen, bedarf es keiner Auftragsdatenverarbeitung, da in diesem Fall nur die Nutzenden die Betroffenen sind und somit eine Einwilligung gegenüber dem Anbieter ausreichend ist.[81]

Die Freiwilligkeit der Einwilligung der Studierenden steht nicht in Zweifel, sie können sowohl auf andere Anbieter als auch auf die „Offline“ Variante Microsoft Office 2016 ausweichen.[82]

Die Universität Bamberg bietet ihren Studierenden Exchange Online und Office 365 mit allen Funktionen an. Sofern jedoch Studierende dies nicht wünschen, steht ein eigenes universitäres E-Mail-Postfach zur Verfügung.[83]

[78] Meents, Borges/Meents Cloud Computing (2016), S. 100 f. Rn. 124.
[79] Erwägungsgrund 46 und Art. 17 Abs. 1 Datenschutzrichtlinie.
[80] Siehe z. B. https://www.blog.google/topics/google-cloud/eu-data-protection-authorities-confirm-compliance-google-cloud-commitments-international-data-flows/ .
[81] So auch Borges, Borges/Meents Cloud Computing (2016), S. 286 Rn. 33.
[82] Spindler/Nink in Spindler/Schuster Recht der elektronischen Medien (2015) BDSG § 4a Rn. 6, beck-online.
[83] https://www.uni-bamberg.de/rz/dienstleistungen/mail/studium/altmailstud/ .

Ferner können Hochschulen die verfügbaren Features von Office 365 begrenzen, wie auch Einfluss darauf nehmen, welche Informationen über Nutzende übertragen werden.

So kann Office 365 auf die Möglichkeit beschränkt werden, dass nur die Installationsdateien der Desktopversion und die Nutzung der Apps auf Tablets und Smartphones möglich ist. Nutzeraccounts beinhalten nur zufällige Buchstaben und Zahlen sowie die Domain; weder Vor- noch Nachname werden an Microsoft übermittelt.[84] Ähnlich datenschutzsparsam kann über diesen Ansatz auch z. B. Azure oder AWS weiterverfolgt werden.

Es zeigt sich also, dass für einzelne Anwendungsszenarien das Übermitteln von Daten sehr gut mit Einwilligungen gelingen kann und zugleich datensparsame Modifikationen möglich sind.

9 Anforderungen bei dienstlicher Nutzung von Clouddiensten

Soweit jedoch ein dienstlicher Einsatz von Office 365 beabsichtigt ist, kann, selbst wenn man unterstellt es lägen wirksame Einwilligungen der Nutzenden vor, nicht ausgeschlossen werden, dass personenbezogene Daten Dritter an den Cloudanbieter übermittelt werden (z. B. Bilder Dritter bei der Nutzung von Adobe Creative Suite, Adressdaten in Briefentwürfen zu Office Online, Kontakte zu Google- oder Appleddiensten).

Die Frage, ob bei der Einführung und Nutzung eines Clouddienstes der mitbestimmungspflichtige Teil des Personals an Hochschulen betroffen ist,[85] wird im Regelfall zu bejahen sein, soweit Rechenzentren z. B. für Administration mitverantwortlich sind oder Verwaltungskräfte den Dienst mitnutzen.[86] Es bietet sich daher an, eine Dienstvereinbarung über die Einführung und Nutzung von Clouddiensten abzuschließen, da diese dann den Umgang mit den personenbezogenen Daten in der Cloud anstelle einer Einwilligung rechtfertigen kann.[87] Diese erstreckt sich jedoch nicht auf den von der Mitbestimmung ausgenommen Personenkreis. Ohne eine Einwilligung können deren personenbezogene Daten für Clouddienste nur im Rahmen des BayDSG oder anderer Rechtsvorschriften erhoben, verarbeitet und genutzt werden.

Zwar können Satzungen von Körperschaften des öffentlichen Rechts eine solche Rechtsvorschrift darstellen, jedoch muss bereits aus der Ermächtigungsgrundlage der Eingriff in die Grundrechte der Betroffenen erkennbar sein.[88] Soweit sich der Umgang mit personenbezogenen Daten für die Aufgaben z. B. eines Rechenzentrums bei der Nutzung von IT aufdrängt, dürfte z. B. die Ermächtigungsgrundlage des Bayerischen Hochschulgesetztes[89] für Satzungen tauglich sein um auch die damit verbundenen Eingriffe in Grundrechte als Erlaubnisnorm rechtfertigen. Die Eingriffsintensität wird erheblich durch die Pflichten des Diensteanbieters zu technischen und organisatorischen

[84] https://www.rz.uni-wuerzburg.de/dienste/shop/studierende/software_fuer_studierende/microsoft_office/ .

[85] Art. 4 Abs. 4 BayPVG über den Professoren und Professorinnen (Art. 2 Abs. 1 S. 1 Nr. 1 BayHSchPG), Juniorprofessoren und Juniorprofessorinnen (Art. 2 Abs. 1 S. 1 Nr. 2 BayHSchPG), Wissenschaftliche Mitarbeiter und Mitarbeiterinnen mit Weiterqualifizierungsaufgaben (Art. 22 Abs. 3 BayHSchPG).

[86] Soweit es sich nicht um (bayerische) staatliche Hochschulen handelt, kann der Kreis der mitbestimmungspflichtigen Personen auch größer sein, und wissenschaftlicher Mitarbeiter miterfassen.

[87] Ehman in WILDE/EHMANN/NIESE/ KNOBLAUCH Datenschutz in Bayern Art. 15 Rn. 12 BayDSG.

[88] So Bäcker in BeckOK DatenSR 18. Ed. 1.5.2016, BDSG § 4 Rn. 12. Vgl. auch zur ähnlichen Situation bei Informationsfreiheitssatzungen BayVGH, Beschluss vom 27.02.2017, Az. 4 N 16.461.

[89] Art. 19 Abs. 5 S.5 BayHSchG. Eine solche Satzung oder Ordnung kann jedoch nicht dazu eingesetzt werden, die Mitbestimmungspflicht des Personalrates zu umgehen.

Datenschutzmaßnahmen[90], Datenschutzbeauftragen, Freigaben und Verfahrensverzeichnissen[91] sowie inzwischen auch Informationssicherheitskonzepten[92] abgemildert.

Die Legitimationswirkung von Dienstvereinbarung und Hochschulsatzung endet im Regelfall aber stets dort, wo auch personenbezogene Daten nicht satzungsmäßiger Mitglieder der Hochschule zu einem Clouddienst verlagert werden. Der Umgang mit diesen personenbezogenen Daten bedarf einer eigenständigen Legitimation. Die Fiktion, dass an einen Auftragsdatenverarbeiter im Rahmen der Auftragsdatenverarbeitung die Daten nicht übermittelt werden, bleibt dann der einzige Weg den zusätzlichen Datenfluss zum Auftragsdatenverarbeiter zu legitimieren. Aus diesem Grund ist grundsätzlich eine Auftragsdatenverarbeitung erforderlich, es sei denn, die Nutzung z. B. bei Office 365 wird auf die Möglichkeit beschränkt, dass nur die Installationsdateien der Desktopversion verfügbar sind und die Nutzung der Apps auf Tablets und Smartphones ohne Zugriff auf OneDrive möglich ist.

10 Geheimnisschutz

10.1 Aktuelle Rechtslage

Aus dem Blick kann auch geraten, dass viele Informationen aus der Hochschule dem Dienstgeheimnis[93], dem Amtsgeheimnis[94] oder dem Geheimnisschutz für Dritte[95] unterliegen. Ein Lösungsansatz kann sein, dass Daten nur im verschlüsselten Zustand in der Cloud liegen, sofern die Verschlüsselung dem Stand der Technik entspricht und die Schlüsselverwaltung allein in der Verfügungsgewalt der Hochschule bleibt.[96] Die Alternative für die öffentliche Verwaltung ist, dass das Personal des Anbieters förmlich verpflichtet wird und die gesetzliche Pflicht zum Geheimnisschutz besteht.[97] Der teilweise vertretene großzügigere Gehilfenbegriff[98] löst die Probleme nur für Berufsgeheimnisträger, erfasst aber nicht den Bereich des Geheimnisschutzes bei Tätigkeiten für die öffentliche Verwaltung. Bei einer mutigeren Gesetzesauslegung könnte angenommen werden, dass keine unbefugte Offenbarung von Dienstgeheimnissen im Rahmen einer Auftragsdatenverarbeitung erfolgt.[99] Für eine praktikablere, rechtssichere Lösung, welche die Reichweite des Tatbestandes beschränkt, hat der Bundestag am 29. Juni 2017 eine Reform des § 203 StGB auf Grundlange des Regierungsentwurfs [100] auf den Weg gebracht. Demnächst können Vertraulichkeitsvereinbarungen mit den Anbietern Strafbarkeitsrisiken beseitigen.

[90] Art. 7 BayDSG.
[91] Art. 25–28 BayDSG.
[92] Art. 8 Abs.1 BayEGovG.
[93] § 353b StGB.
[94] Art. 30 BayVwVfG; § 30 VwVfG.
[95] § 203 Abs. 2 StGB, ggf. auch § 17 UWG.
[96] Thalhofer in Auer-Reinsdorff/Conrad, IT- und Datenschutzrecht, Teil C. Software-, Hardware- und Providerverträge § 19 Outsourcing-Verträge Rn. 227–230; Spickhoff in: Spickhoff Medizinrecht (2014) § 205 StGB Rn. 23.
[97] § 11 Abs. 1 Nr. 2 c StGB; Schünemann in Leipziger Kommentar (2009) § 203 StGB Rn. 44.
[98] Preuß, DuD 2016, 802 (806).
[99] So z. B. Pohle, Kommunikation und Recht, 2013, 34 (35).
[100] BT-Drs 18/11936.

10.2 Rechtslage unter der DSGVO

Eine explizite Vorschrift für den Umgang mit durch besondere Geheimnispflichten geschützte, personenbezogene Daten findet sich in der DSGVO abgesehen von satzungsmäßigen Berufsgeheimnisträgern[101] nicht. Allerdings werden diese erwähnt.[102] Die Mitgliedstaaten können regeln, welche Einrichtung oder Behörde Aufsicht über die Einhaltung des Datenschutzes dieser speziellen personenbezogenen Daten wahrnimmt. Für den behördlichen Bereich wird anzunehmen sein, dass § 203 Abs. 2 StGB den zulässigen Tätigkeitsbereich von Auftrags(daten)verarbeitern weiterhin einschränkt,[103] da Geheimnisschutz und Datenschutz zwar Überschneidungen aufweisen, in ihrer Zielsetzung aber unterschiedlich sind.[104]

10.3 Fazit zum Geheimnisschutz

Die Verantwortung Geheimnisschutz zu gewährleisten liegt bei der Behörde selbst und ist eine vom Datenschutz unabhängige Pflicht. Während es gesetzliche Vorgaben für die Verlagerung von personenbezogenen Daten zu Auftragsdatenverarbeitern gibt, ist zum Schutz vor dem Offenbaren von Geheimnissen im gesetzlichen Grundfall zukünftig eine Vertraulichkeitsvereinbarung vorgesehen. Sind personenbezogene Daten zugleich auch Geheimnisse, nimmt die DSGVO ferner an, dass oft ein höherer Schaden im Falle einer Verletzung des Schutzes personenbezogener Daten vorliegen wird, sodass solche Daten besonderer Schutzmaßnahmen bedürfen.[105]

11 Fazit

Auch öffentliche Behörden sind oft nicht auf eine Auftragsdatenverarbeitung durch Anbieter mit Sitz im EWR beschränkt. Es kann auch auf die Dienste aus Drittstaaten zurückgegriffen werden, wenn die Auftragsdatenverarbeitung vertraglich richtig und umfassend geregelt ist. Zertifizierte Dienstleister, die bereit sind, ihre Leistung mit für öffentliche Stellen geeigneten Auftragsdatenverarbeitungsverträge oder mit vollständigen EU-Standardvertragsklauseln anzubieten, ebnen einen rechtssichern Weg in die Cloud.

Daneben bleiben interne Herausforderungen durch Freigabe, Mitbestimmung und Geheimnisschutz. Dies zeigt, dass weder kostenfreie Dienste noch bei Verbrauchern favorisierte Produkte stets eine Lösung für den Einsatz an Hochschulen sind sobald eine dienstliche Nutzung erfolgen soll. Um einen guten Kompromiss zu finden, sind die Hochschulen gefragt ihre Anforderungen zur Auftragsdatenverarbeitung klar zu kommunizieren. Auch nach Innen ist eine Sensibilisierung nötig um Bewusstsein für Dienste zu schaffen, die der Gesetzmäßigkeit der Verwaltung genügen (Neudeutsch auch „Compliance“ genannt).

Inhalt, Form und Kontrolle der Auftragsdatenverarbeitung geben die Gesetze vor. Wirksamen Verträgen stehen jedoch oft formelle und inhaltliche Fehler in den Vertragsmustern der Dienstleanbieter entgegen. Hiervor können auch Zertifizierungen nicht schützen. Ein erster Schritt zur Fehlervermeidung ist das GÉANT IaaS Vergabeverfahren, ein zweiter die Rechtsvereinheitlichung durch die DSGVO.

---

[101] Art. 14 Abs. 5 d DSGVO.
[102] Erwägungsgrund 50 DSGVO.
[103] So Preuß, DuD 2016, 802 (808).
[104] Dies wird sehr deutlich in den Erwägungsgründen 18 und 35 Richtlinie (EU) 2016/943 und 63 DSGVO.
[105] Erwägungsgrund 76 DSGVO; Plath in: Plath, BDSG/DSGVO, 2. Aufl. 2016, Artikel 24 DSGVO Rn. 7.

Neues aus der DFN-Cloud: GÈANT IaaS-Vergabeverfahren

Michael Röder
DFN-Verein
Berliner Geschäftsstelle
Alexanderplatz 1
10178 Berlin
roeder@dfn.de
csdm@dfn.de

Abstract: Der DFN-Verein organisiert Cloudlösungen speziell für seine Anwender aus Wissenschaft und Forschung. Mit Rücksicht auf deren besondere Bedarfe wird das Portfolio der Cloud-Services ständig weiterentwickelt. Diese Arbeit fasst den aktuellen Stand der Wissenschaftscloud zusammen und beschreibt den Ansatz, das Szenario Infrastructure-as-a-Service (IaaS) darin zu integrieren.

1 Die DFN-Cloud: Eine Cloud für die Wissenschaft

Dieses Thema ist sowohl im Forum Cloud-Dienste während der 65. als auch während der 66. DFN-Betriebstagung vorgestellt worden. Der vorliegende Beitrag kombiniert beide Vorträge miteinander und lässt dabei aktuelle Erkenntnisse einfließen.

1.1 Föderierte Dienste in der DFN-Cloud

Seit 2014 werden föderierte Dienste in der DFN-Cloud angeboten. Dabei handelt es sich um Dienstangebote, die von Rechenzentren innerhalb der Community aus Wissenschaft und Forschung für andere Einrichtungen aus der Community erbracht werden.

Einrichtungen, die selbst einen föderierten Dienst in der DFN-Cloud anbieten, nehmen die Rolle der Forschungspartner ein. Forschungspartner öffnen ihre Infrastruktur über den eigenen Wirkungskreis hinaus auch für andere Anwender im Wissenschaftsnetz (X-WiN) mit dem Ziel der Erprobung und Weiterentwicklung eines föderierten Dienstes. Ein Forschungspartner kann einen oder mehrere föderierte Dienste anbieten. Einrichtungen, die einen föderierten Cloud-Dienst anwenden und dadurch aktiv zu dessen Weiterentwicklung beitragen wollen, treten in der Rolle des Erprobungspartners auf. Der DFN-Verein übernimmt als zentraler Ansprechpartner die Aufgabe, das Vertragswerk zu entwickeln und den Dialog zwischen Erprobungs- und Forschungspartnern zu koordinieren.

In der DFN-Cloud sind zum Zeitpunkt der Verfassung dieses Beitrages vier Forschungspartner aktiv. Insgesamt werden fünf föderierte Dienste angeboten. Dabei handelt es sich zur Zeit ausschließlich um Sync-&-Share-Services mit deren Hilfe die Kollaboration über Filesynchronisation unabhängig von Ort und Betriebssystem vereinfacht wird. Aktuell wenden 27 Erprobungspartner einen föderierten Dienst an. Über die Einrichtungsgrenzen hinweg werden dadurch rund 21.500 Endnutzer durch Clientlizenzen mit insgesamt circa 123 TB Storage versorgt.

1.2 Weiterentwicklung föderierter Dienste

Die aktuell angebotenen föderierten Dienste werden in der Cloud-Service-Klasse „Software-as-a-Service“ (SaaS) eingeordnet. SaaS charakterisiert sich dadurch, dass dem Anwender ein definiertes Nutzungsszenario erbracht wird, welches sich eindeutig durch ein begrenztes Featureset beschreiben lässt und sich dabei dem Anwender gegenüber frei von Wartungsaspekten darstellt.

Im Falle von Sync-and-Share bedeutet das, dass – vereinfacht gesagt – lokale Inhalte, die in einem bestimmten Ordner im Dateisystem abgelegt werden, durch einen Client im Hintergrund automatisch über die gewünschte Anzahl registrierter Endgeräte hinweg synchronisiert werden. Für den Nutzer belanglos sind dabei Komponenten wie beispielsweise Serverhardware/-standort des Dienstleisters, Softwareentwicklung des Clients oder welche RAID-Philosophie beim Dienstleister für die notwendige Redundanz sorgt. Der Dienst erscheint dem Nutzer wie eine wartungsfreie Software und definiert sich über sein äußeres Erscheinungsbild und über die Funktionalität. Grundsätzlich geeignet für Cloud-Services à la SaaS sind Server-Client-Systeme wie beispielsweise browserbasierte Office-, Projektmanagement- oder Chatlösungen. An genau diesen Projekten arbeiten die Forschungspartner zum Zeitpunkt der Erstellung dieses Artikels, um Anwendern aus der Community die Kollaboration basierend auf der Technologie und Verfügbarkeit des Wissenschaftsnetzes orts- und betriebssystemunabhängig zu erleichtern.

Eine weitere populäre Klasse der Cloud-Services stellt „Infrastructure-as-a-Service“ (IaaS) dar. IaaS äußert sich dem Endanwender gegenüber ein Level abstrakter als SaaS: er bekommt hier keine Software sondern eine Ressource als Dienstleistung zur Verfügung gestellt. Diese Ressource charakterisiert sich durch technische Rahmenbedingungen wie bspw. Anzahl CPU-Kerne, RAM, Storage, GPU, gewähltes Betriebssystem oder Netzanbindung. Entsprechend der vom Nutzer definierten Parameter wird eine virtuelle Ressource erzeugt, die im Verbund mit vielen weiteren virtualisierten Instanzen anderer Nutzer – abgesichert gegenüber gegenseitigem Zugriff – auf einer gemeinsamen leistungsfähigen physischen Hardware im Rechenzentrum des Betreibers läuft. Die Performance der Ressource richtet sich ausschließlich entlang derjenigen Anforderungen aus, die der Nutzer definiert, um seinen Rechenprozess durchführen zu können. Denkbare Anwendungsszenarien für solche virtualisierten entfernten Ressourcen sind beispielsweise besonders CPU-lastige, speicherhungrige, grafisch aufwändige oder I/O-intensive Prozesse, deren Bearbeitungsdauer die Lebenszeit einer physischen Hardware deutlich unterschreitet. In anderen Anwendungsfällen werden virtualisierte Ressourcen generiert und im StandBy-Betrieb belassen bis sie zum Abfedern von Lastspitzen benötigt werden. Sie werden erst aktiviert, wenn die Leistungsfähigkeit der physischen Hardware im eigenen Rechenzentrum nicht mehr ausreicht, um alle Clientanfragen zu versorgen; typischerweise zu Spitzenzeiten wie Semesteranfang oder -ende.

Allen virtuellen Ressourcen gemein ist, dass sie der Endnutzer via Remote Access wie ein physisches Gerät im eigenen Rechenzentrum administriert. Für den Anwender ergibt sich der wesentliche Vorteil, dass Kosten für die Ressource nur in dem Zeitraum anfallen, in denen das Gerät tatsächlich in Benutzung ist – im Gegensatz zum Kauf, der nicht nur initial hohe Kosten verursacht, sondern ggf. auch durch Vergabeprozesse Personal bindet. Liegt das zu berechnende Ergebnis vor, wird die Ressource freigegeben und gelöscht. Sie verursacht anschließend keine Kosten mehr. Umgebungsbedingungen wie Klimatisierung, Stellfläche, Energieversorgung oder Wartung, die beim Self-Hosting unabdingbar sind, sind bereits im Preis für die virtualisierte Ressource integriert und müssen vom Anwender nicht zusätzlich berücksichtigt werden.

Einige Forschungspartner arbeiten derzeit daran, IaaS ebenfalls als föderierten Dienst in der DFN-Cloud anzubieten.

2 Externe Services in der DFN-Cloud

Im zurückliegenden Jahr ist durch die GÉANT Association, dem Dachverband der europäischen NRENs[1], ein europaweites Vergabeverfahren angestoßen worden. Kommerzielle IaaS-Anbieter sind darum gebeten worden, durch ein eigenes Angebot teilzunehmen und darin die besonderen Bedarfe der europäischen Community aus Wissenschaft und Forschung einzubeziehen.

Als Voraussetzung für ein erfolgreiches Angebot hat GÉANT beispielsweise darum gebeten, die Strukturen der Abrechnungsprozesse öffentlicher Auftraggeber (z. B. langfristige Vorauszahlungen und Zahlung gebuchter Ressourcen auf Rechnung) zu berücksichtigen. Da die europäischen NRENs ihre Anwender bereits mit leistungsfähigen Netzinfrastrukturen versorgen, wurden Bieter außerdem dazu angehalten, keine zusätzlichen Kosten für den Netzwerktraffic zu erheben und leistungsfähige Übergänge zu den Wissenschaftsnetzen zu realisieren. Bieter mussten sich zudem dazu bereit erklären, dass sie die Vorgaben der europäischen Datenschutzrichtlinie berücksichtigen und dass sie den Wechsel von einem Dienstanbieter zum anderen für Anwendereinrichtungen möglichst unkompliziert gestalten. Zusätzlich setzt eine erfolgreiche Teilnahme am Vergabeverfahren voraus, dass das Dienstangebot kompatibel zu der in Wissenschaft und Forschung weit verbreiteten SSO[2]-Infrastruktur auf Basis des SAML2[3]-Standards ist.

Für den Fall, dass teilnehmende Unternehmen den Radius ihrer Diensterbringung regional begrenzen wollten, hatten alle Bieter die Möglichkeit, gemeinsam mit ihren Unterlagen anzugeben, in welchen Ländern Europas sie ihren IaaS-Dienst anbieten wollten.

2.1 Grundlagen des Vergabeverfahrens

Es handelt sich bei der gewählten Form des Verfahrens um ein zweistufiges Verfahren. Das bedeutet, dass die ausschreibende Instanz (GÉANT) allen erfolgreichen Bietern eine Rahmenvereinbarung („Framework Agreement“) anbietet. Diese Rahmenvereinbarung besitzt eine zeitlich begrenzte Gültigkeit von 4 Jahren. Einen Zuschlag erhalten alle diejenigen Teilnehmer, die den Ausschreibungsprozess auf Grund ihrer eingereichten Unterlagen erfolgreich absolvieren konnten und die die Rahmenvereinbarung anerkennen.

Auf Basis der Rahmenvereinbarung können Anwendereinrichtungen – sofern das NREN, in welchem sie organisiert sind, vorher am Vergabeverfahren teilgenommen hat – einzelne Dienste im Rahmen von Einzelaufträgen („Call-Off-Agreement“) abrufen. Die Ausschreibungspflicht teilnehmender Einrichtungen ist bereits erfüllt.

Die Unterlagen zum Einzelauftrag sind standardisiert worden. Soll ein Einzelauftrag in einzelnen Punkten von der Vorlage abweichen, steht der Anwendereinrichtung das Werkzeug des Miniwettbewerbs („Mini Competition“) zur Verfügung. Dadurch können selbst gewählte Kriterien von besonderer Bedeutung auf Basis des Frameworks für den eigenen Wirkungskreis verhandelt werden.

Interessierte Einrichtungen sind dazu eingeladen, Fragen rund um das Verfahren und seine Ergebnisse an cloud@dfn.de zu stellen.

[1] National Research and Education Network („Forschungsnetz“)
[2] Single-Sign-On
[3] Security Assertion Markup Language

2.2 Ergebnisse des Verfahrens

Im Anschluss an das Verfahren stellte sich heraus, dass die unter 2 genannten Voraussetzungen erheblichen Aufwand bei den bietenden Unternehmen verursachten. Deshalb erwies sich der Ansatz, diese Forderungen durch GÉANT zentral zu formulieren, als hilfreich. Die hinter GÉANT vereinte potentielle Anzahl anwendender Einrichtungen konnte eine große Zahl kommerzieller Dienstanbieter dazu motivieren, ihre bereits am Markt verfügbaren Dienstangebote um die besonderen Bedarfe von Wissenschaft und Forschung zu erweitern. In der Folge steht Anwendereinrichtungen in jedem teilnehmenden NREN mindestens ein IaaS-Dienst zur Verfügung.

Die folgende Tabelle zeigt alle Dienstanbieter, die sich für den Erhalt der Rahmenvereinbarung qualifiziert haben und ihren Service in Deutschland anbieten:

Originäre Provider [4]	**Reseller**	
	Amazon Web Services	*Microsoft Azure*
CloudSigma	Arcus Global	Atea
Dimension Data	Comparex	Comparex
Interoute	Telecom Italia	SoftwareOne
T-Systems		
Vancis		

Tabelle 1: Auflistung der Dienstanbieter für DFN-Anwender

3 Neue Services in der DFN-Cloud

Parallel zu den föderierten Diensten in der DFN-Cloud sollen in Zukunft die Ergebnisse des europaweiten Ausschreibungsverfahrens als „externe Dienste“ in der DFN-Cloud angeboten werden.

In der Anordnung und den Rollen der an den externen Diensten beteiligten Parteien existieren im Vergleich zu den föderierten Diensten teilweise erhebliche Unterschiede. Deshalb sind hier voraussichtlich weitere administrative und organisatorische Fragestellungen zu klären bevor der Produktivbetrieb aufgenommen werden kann.

Zunächst können interessierte Einrichtungen einen Pilotbetrieb starten. Dieser Pilotbetrieb ist zeitlich begrenzt und wird anschließend nahtlos in den Produktivbetrieb überführt.

[4] Im Kontext des Vergabeverfahrens auch: „OIP“ (Original IaaS Provider)

Aus dem Leben eines DFN-Cloud Providers
Menschen, Verträge, Technik

Dr.-Ing. Thomas Hildmann
tubIT - IT Service Center
Technische Universität Berlin
Einsteinufer 17
10587 Berlin
thomas.hildmann@tu-berlin.de

Abstract: Der DFN Verein bietet mit der DFN-Cloud seinen Mitglieder eine unkomplizierte Möglichkeit zur Kooperation, insbesondere zur Erforschung von Cloud-basierten Diensten. Der vorliegende Erfahrungsbericht wirft ein Schlaglicht auf die verschiedenen Aspekte der Bereitstellung eines Dienstes für andere DFN-Mitglieder an Hand von ausgesuchten Beispielen. Die Erfahrungen können gleichermaßen für (potentielle) Betreiber, wie auch für (potentielle) Nutzer und den DFN-Verein selbst interessant sein.

1 Einleitung

Nachdem die Kooperationen auf Basis der DFN-Cloud zunächst sehr zögerlich anliefen, gibt es mittlerweile mehr als 15 Einrichtungen, die den DFN-Cloud Dienst der TU Berlin nutzen. Jede Einrichtung bringt dabei ihre eigenen Anwendungsfälle, technischen und administrativen Voraussetzungen und Ideen mit.

Dies bietet Möglichkeiten zur Weiterentwicklung auch des eigenen Cloud-Speicher Angebots aber auch Herausforderungen, vor allem was die Skalierung eines solchen Dienstes angeht.

Es ist absehbar, dass die Nutzerzahlen der DFN-Cloud Teilnehmer bald die Zahl der eigenen Nutzer überschreiten wird. Damit verabschieden wir uns von einem Dienst, den wir ohnehin für die eigene Universität anbieten und den wir für viele kleine Einrichtungen mitbetreiben und entwickeln uns zu einem Anbieter eines Dienstes für viele Nutzer unterschiedlichster Einrichtungen.

Der technische Aspekt eines solchen Angebots ist dabei nur der augenscheinlichste (Oder der, der einem Techniker zuerst einfällt). Was aber bedeutet dies für die eigenen Mitarbeiterinnen und Mitarbeiter? Wer sind die Partner, mit denen wir zusammenarbeiten? Welchen Aufwand macht der vertraglich-/organisatorische Teil aus? Und wie sehen wir die Zukunft der DFN-Cloud und des eigenen Angebots?

Dieser Bericht behandelt Hauptaspekte eines DFN-Cloud Angebots aus Betreibersicht:

1. Die Menschen hinter dem Dienst
2. Die Vertragslage
3. Die Technische Entwicklung

Er ist nicht nur gedacht als Information für andere (potentielle) Anbieter der DFN-Cloud, sondern auch als Information für aktuelle oder künftige Nutzer unseres Dienstes und für den DFN-Verein als Initiator der DFN-Cloud. Die DFN-Cloud wurde als Forschungsvorhaben ins Leben gerufen. Vor diesem Hintergrund ist das Sammeln von Erfahrungen und Erkenntnissen und deren Verbreitung und Diskussion eine der Kernaufgaben des Vorhabens.

2 Menschen

Bei der Einrichtung eines Cloudspeicher-Dienstes für andere DFN-Einrichtungen haben wir zunächst an die Bereitstellung von Technik für andere Einrichtungen gedacht. Die Bereitstellung eines solchen Cloud-Dienstes hat zunächst einmal jedoch sehr viel mehr mit Menschen und Organisation zu tun, als mit Technik.

Dabei darf die Komponente der menschlichen Netzwerke nicht unterschätzt werden. Die Kooperation im Bereich der DFN-Cloud Dienste hat auch für eine engere Zusammenarbeit auf anderen Gebieten gesorgt.

Menschen bedarf es aber vor allem auch intern mit unterschiedlichen Arbeitsgebieten, um ein solches Vorhaben stemmen zu können.

2.1 Partner

Kunden bzw. Nutzer unseres Cloud-Storage Dienstes sind zunächst erst einmal die Mitglieder der TU Berlin. Etwa 20.000 Nutzer zählt der tubCloud-Dienst aktuell. Fast 3.000 nutzen den Dienst täglich, über 6.000 mindestens monatlich. Die Tendenz ist weiterhin steigend. Hinzu kommen nun noch über 15 weitere Einrichtungen, die den Dienst aktuell an der TU Berlin nutzen oder testen. Diese Nutzer haben 20 bis 5.000 Konten gebucht mit einer absoluten Quota zwischen 500 und 50.000 GB. Die Authentisierung erfolgt über die lokale Nutzerverwaltung (ggf. über die API gesteuert), eine LDAP-Anbindung oder über Shibboleth (üblicherweise die DFN-AAI).

Dabei gibt es unterschiedliche Voraussetzungen, bevor die Partner das Angebot der TU Berlin nutzen:

1. Einrichtung eines neuen Dienstes: An der Einrichtung gab es bislang noch keinen Sync-n-Share Dienst. Dieser soll über die DFN-Cloud nun aufgebaut und angeboten werden.
2. Ersetzen eines evtl. externen Dienstes: Aus unterschiedlichen Gründen (Kosten/Nutzen/Aufwand, rechtliche Aspekte, Stabilität, …) soll ein bestehender Dienst abgelöst werden. Häufig wird dann die Funktionalität des alten Dienstes in der DFN-Cloud nachgebildet und gemäß der neuen Anforderungen erweitert.
3. Migration einer eigenen Installation: Einige Einrichtung hatten bereits eine eigene ownCloud/Nextcloud-Installation, die dann zur TU Berlin migriert werden sollte. Häufig sind solche Migrationen am anspruchsvollsten.

Alle Einrichtungen haben gemeinsam, dass sie sich um den 1st-Level Support selbst kümmern und mit der TU Berlin über 2 bis 3 Ansprechpartner in den Kontakt gehen. Dabei laufen die Anfragen bei uns in der Regel in einem Ticketsystem auf und werden intern verteilt. Nach außen treten ebenfalls wieder 2 bis 3 Personen auf, so dass man sich kennenlernt und auch ein Verständnis für die Bedürfnisse der Endnutzer in den Partnereinrichtungen aufbaut.

Nicht selten führt die Bekanntschaft über einen solchen Kontakt auch zu einem Austausch zu angrenzende Themen. Selbstverständlich werden auch Ideen über die Weiterentwicklung des Dienstes ausgetauscht. Trotz der steigenden Nutzerzahl geht es erfreulicher Weise nur selten um die Behebung von Störungen oder die Beseitigung von Problemen. Der häufigste Kontaktgrund war in der vergangenen Zeit die Absprache zu Updates und damit Wartungsfenstern oder der Wunsch von neuen Funktionen (z. B. Aktivierung von Plugins).

2.2 Mitarbeiter/innen

Eine häufig gestellt Frage im Zusammenhang mit unserem Dienst ist die nach dem Betreuungsaufwand im Sinne von Personentagen, die ein solcher Dienst kostet. Der Mehraufwand für die einzelnen Erprobungspartner tritt eigentlich nur dann zutage, wenn ein Update ansteht oder ein kundenspezifisches Problem aufgetreten ist. In der Regel verhalten sich alle Dienste gleich und werden von uns auf einem ähnlichen Architektur- und Patchstand gehalten. Damit gilt grob: Wenn die tubCloud läuft, laufen auch die 16 anderen Instanzen.

Bei tubIT gibt es nur einen für den Dienst hauptverantwortlichen Administrator, der gemeinsam mit mir als Leiter der Abteilung Infrastruktur die strategische Ausrichtung und die organisatorischen Aspekte koordiniert. Unterstützung gibt es vom Linux-Team, das aus sechs weiteren Personen besteht und die Basis für den Dienst bereitstellt (Betriebssystem, Deployment, Cluster-Filesystem, …). Das Netzwerkteam unterstützt bei Bedarf mit bis zu zwei weiteren Personen bei Fragen des Load-Balancing, Firewall-Konfiguration etc. In der Abteilung IDM und im Linux-Team gibt es ferner Sozialisten zu Themen wie LDAP und Shibboleth, die bei der Grundkonfiguration und bei Problemen hinzugezogen werden können.

Gerade, wenn es darum geht eine neue Version der Server-Software oder der Apps und Sync-Clients zu testen können wir auf die Mitarbeiterinnen und Mitarbeiter des tubIT-Supports zurückgreifen, die nicht nur aus ihrem eigenem Erfahrungsschatz als Nutzer/innen testen können sondern auch vor dem Hintergrund der bearbeiteten Support-Anfragen.

Je nach Bedarf arbeiten so also 1,5 bis 10 Personen für die tubCloud oder den DFN-Cloud Dienst. Die meisten haben jedoch weitere Tätigkeitsfelder und werden bei Bedarf hinzugezogen.

Eine genaue Berechnung der in die Dienste gesteckten Arbeitszeit ist sehr schwierig zu erstellen, da viele Arbeit für die Weiterentwicklung (z. B. OpenStack und Container-Technologie) auch vielen weiteren Diensten zu Gute kommt und nur am Beispiel der tubCloud erarbeitet wird. Selbst Aufträge eines DFN-Cloud Partners werden selten nur von diesem genutzt, sondern kommen in der Regel allen, auch den eigenen tubCloud-Nutzern früher oder später zu Gute.

3 Verträge

Die Konstruktion der DFN-Cloud erfordert Verträge auf verschiedenen Seiten. Die jeweiligen Projektpartner schließen Verträge mit dem DFN und Verträge untereinander.

3.1 DFN Cloud Verträge

Alle Verträge, die mit dem DFN oder von den Hochschulen untereinander geschlossen werden, sind vom DFN vorbereitet. Das Vorgehen für die DFN-Cloud ist in [Rö16] genau beschrieben. Somit haben alle Seiten mit diesem Teil keinen nennenswerten Aufwand.

Gleiches gilt für die Erstellung der SSL-Zertifikate, die unkompliziert über die DFN-PKI erstellt werden können. Hier können bestehende Workflows weiter genutzt werden.

Während die Rahmenvereinbarung mit dem DFN geprüft und unterzeichnet wird, kann bereits mit dem Test des Dienstes bei der TU Berlin begonnen werden. Entschließt sich der Tester schließlich für die Nutzung des Dienstes, findet eine Beratung durch uns statt an dessen Ende wir die relevanten technischen Daten abfragen und eine Instanz konfigurieren und u. a. Mit SSL-Schlüsseln ausstatten (Da der Dienst bei der TU Berlin läuft, muss das Schlüsselpaar auch hier hinterlegt werden).

Eine andere Möglichkeit ist die Bereitstellung eines transparenten Proxy, der im Kontext der Partnereinrichtung läuft und seinerseits eine SSL-Verbindung zum Dienst bei der TUB herstellt.

3.2 Auftragsdatenverarbeitungsvereinbarung

Die meisten Änderungswünsche entstanden bei der Ausarbeitung der Auftragsdatenverarbeitungsvereinbarung (ADVV). Im Laufe der Zeit gingen uns unzählige Verbesserungen von verschiedenen Datenschutzbeauftragten bzw. Administratoren anderer Rechenzentren zu. Da bei technischen Änderungen die ADVVs immer zu prüfen sind, hatten wir uns nach einigen Monaten schließlich entschlossen, die Änderungen zusammenführen und eine neue modulare ADVV zu erstellen und auch bestehende Vereinbarungen zu aktualisieren.

Für alle Vertragspartner gehen so die Verbesserungen der anderen Einrichtungen mit ein und für uns vereinfacht dies die technischen Prüfungen z. B. bei Updates. Der Erkenntnisgewinn für uns an dieser Stelle war, dass nicht nur technische Erfahrungen mit einem Partner eine Verbesserung des Dienstes für alle anderen bedeuten kann, sondern sehr wohl auch organisatorische und rechtliche Erkenntnisse für alle nutzbar gemacht werden können.

Abgesehen davon brachte die Auseinandersetzung mit den datenschutzrechtlichen Fragen mit verschiedenen Partnern auch viele Erkenntnisse für den Betrieb des Dienstes und die technische Ausgestaltung dessen. Viele Anmerkungen gingen schließlich auch die Entwicklung des Cloud-Dienstes mit ein.

Die Hauptunterscheidung bei den ADVVs basiert auf der verwendeten Authentisierungsmethode. Die höchste Datensparsamkeit bietet hier die Anbindung via Shibboleth. Eine LDAP-Anbindung belässt die Primärquelle bei der jeweiligen Einrichtung, wohingegen die lokale Benutzerverwaltung alle Daten inkl. Passwörter etc. bei der TU Berlin hält. Welche Variante die jeweils Beste für den Anwendungsfall ist, wird mit dem Auftraggeber gemeinsam besprochen.

Aber auch die Nutzung verschiedener Plugins kann Einfluss auf die ADVV haben. Einige Plugins dienen der Ablage weiterer personenbezogener Daten (z. B. das Adressbuch). Je nach zusätzlich genutzter Anwendung muss die ADVV um weitere Punkte erweitert werden.

4 Technik

Die Weiterentwicklung der Technik für einen DFN-Cloud Dienst wie den unsrigen ist eine Daueraufgabe, die nicht zu unterschätzen ist. Gründe für die Weiterentwicklung sind:

- Neue oder geänderte Anforderungen der Nutzer z. B. durch neue Use-Cases
- Erweiterung / Skalierung der Infrastruktur z. B. durch Erprobungspartner mit vielen Nutzern
- Reaktion auf technische Anforderungen z. B. durch neue Versionen und geänderte Abhängigkeiten
- Verbesserung der Performanz oder des Kosten/Leistungsfaktors z. B. Einsparung von Lizenzkosten
- Reduktion des Wartungsaufwands bzw. der Fehleranfälligkeit

Im Folgenden sind einige Beispiele für eine nötige Weiterentwicklung beschrieben.

4.1 Infrastruktur

Die Cloud-Infrastruktur basierte lange Zeit auf den folgenden Komponenten:

1. Cluster-Filesystem: GPFS via. IBM GSS-Cluster
2. Datenbankcluster: Galera-Cluster
3. Frontendserver: Webserver mit PHP
4. Load-Balancer: Cisco ACE

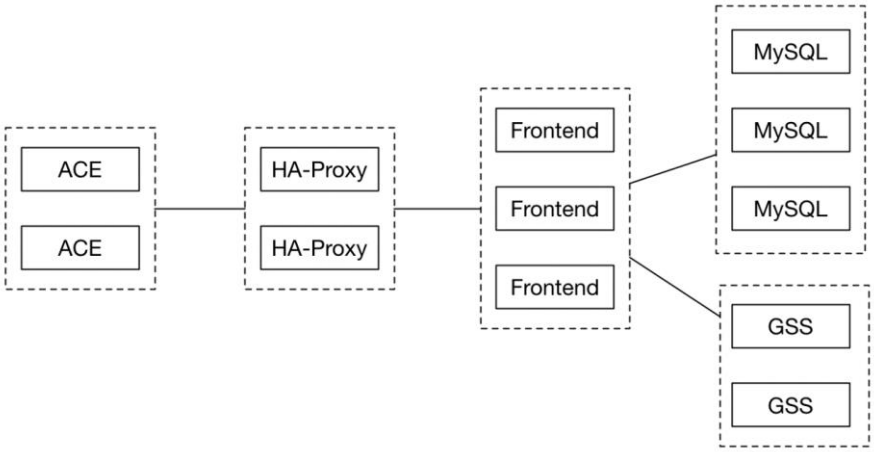

Abbildung 1: Übergangsarchitektur mit HA-Proxy

Mit der Ankündigung des Cisco ACE Load-Balancers musste ein Ersatz implementiert werden. Noch während der Umstellungsphase zwangen uns verschiedene Sicherheitslücken im SSL-Protokoll, das Offloading der SSL-Verbindung auf gesonderte HA-Proxies zu verlagern. Unserer Erfahrung nach ist es ratsam, die SSL-Verbindungen auf wenige leistungsstarke Maschinen (am besten mit Hardware-Crypto) zu verlagern und die eigentlichen Frontend-Server nicht mit der Ver- und Entschlüsselung von SSL-Verkehr zu belasten.

Nach Implementierung der HA-Proxies musste dann in einem zweiten Schritt das Konstrukt Cisco ACE / HA-Proxy von der neu angeschafften F5 Big IP abgelöst werden.

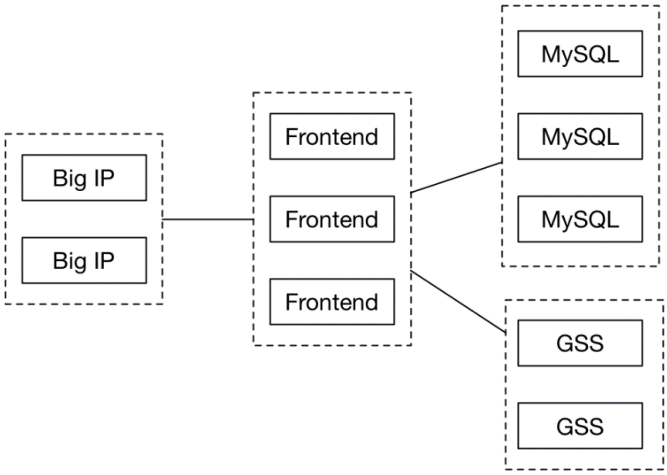

Abbildung 2: Architektur mit F5 Big IP

Eine detaillierte Darstellung der technischen Weiterentwicklung des Dienstes ist in [Hi17] zu finden.

4.2 Updates

In unseren Update-Zyklen sind jeweils zwei reguläre Updates pro Jahr eingeplant. Davon ist in der Regel mindestens eines ein „Feature-Update“, bei dem es hauptsächlich darum geht, den Nutzern weitere Funktionalitäten zur Verfügung zu stellen. Das zweite Update ist dann häufig ein technisches, bei dem der Grundstock für neue Funktionen gelegt wird oder bei dem hauptsächlich Fehler beseitigt und z. B. die Wartbarkeit verbessert wird.

In den Jahren 2015/16 kamen ferner einige kurzfristige Security-Updates hinzu, die von uns jeweils einzeln bewertet werden:

Nachdem vom Hersteller über eine Sicherheitslücke und einen Patch informiert wird, wird die Sicherheitslücke vor dem Hintergrund unserer Infrastruktur und vor dem Hintergrund unserer Anwendungsfälle betrachtet. Im Ergebnis kommen wir zu einer eigenen Gefährdungseinschätzung. In Abhängigkeit der Gefährdung treten wir dann mit den Erprobungspartnern in Kontakt und vereinbaren Wartungsfenster.

Gleiches gilt für die „branded Clients“, also die Sync-Clients und die mobilen Apps, die fertig konfiguriert und mit eigenem Logo bereitgestellt werden. Diese unterliegen eigenen Wartungszyklen und müssen ebenfalls regelmäßig aktualisiert werden. In der Regel warten unsere Partner die „branded Clients“ selbst. Wir unterstützen dabei aber auch gerne.

4.3 Bugs

Fehler gibt es sowohl technische, wie auch organisatorische. Exemplarisch greifen wir zwei Fehler heraus, die 2016 behandelt werden musste:

1. Ein Problem im Zusammenspiel mit dem MySQL-Cluster
2. Das Mysterium um einen verschwundenes Fachgebietsshare

4.3.1 MySQL-Casting

Es waren ein paar Anführungszeichen, die den tubCloud für einige Tage extrem verlangsamten und schließlich unbenutzbar machten. Zu beobachten war eine stetig wachsende Zahl von Datenbankverbindungen. Diese machten den Datenbankcluster immer langsamer und sorgten schließlich dafür, dass keine weiteren Verbindungen mehr angenommen werden konnten. Einziges Mittel zur Reduzierung des Problems war die Beschränkung der eingehenden Verbindungen auf die Frontendserver, was allerdings zur Folge hatte, dass Nutzer sehr lange auf Verbindungen warten mussten oder sogar Timeouts bekamen.

Die Analyse des Problems gemeinsam mit dem Support von ownCloud ergab schließlich folgendes:

Auf dem MySQL Galeracluster konnte man erkennen, dass es Anfragen eines Types gab die als „Langläufer“ markiert waren; also Datenbankanfragen, die außergewöhnlich lange brauchen. Diese Anfragen waren von der Form:

```
mysql> SELECT `configkey`, `configvalue` FROM `oc_appconfig`
WHERE (`appid` = 'files_sharing') AND (`configkey` IN
(70311, 231042, 312174));
```

Aus der Tabelle „oc_appconfig“ die wie der Name sagt die Konfiguration der Apps genannten Plugins von ownCloud beinhaltet sollen zwei Werte ermittelt werden: Ein „configkey“ (Schlüsselfeld) und ein „configvalue“ also der dazugehörige Wert. Dabei wird die App mit der „appid = 'file_sharing'“ gesucht.

Spannend ist der zweite Teil der Bedingung, der nämlich alle „configkeys“ sucht, die in einer Menge von mehreren Zahlen (in unserem Beispiel drei Werte) stehen.

Das Ergebnis sind dann drei Paare aus Schlüssel und Wert, wobei die Schlüssel eben genau die drei Schüssel 70311, 231042 und 312174 sind.

Auf den ersten Blick ist die Abfrage sehr simpel und sollte einen Datenbankcluster, wie den unserigen, der die meisten Tabellen komplett im RAM hält und hinreichend viele Prozessorkerne und SSDs zur Verfügung hat keine nennenswerte Zeit kosten.

Der Test überrascht dann allerdings:

```
Empty set, 3 warnings (6.15 sec)
```

Das Ergebnis ist eine leere Menge, was kein Problem sein muss, sondern ggf. ein erwartetes Verhalten. Spannend ist, dass für diese leere Menge mehr als 6 Sekunden benötigt werden.

Der Grund ist so einfach, wie überraschend. Die Spalte „configkey“ kann nicht nur Zahlen enthalten, sondern beliebige Zeichenketten. Gibt man in der gesuchten Menge („IN“) nun Zahlen an, werden diese zunächst von MySQL in den richtigen Datentypen überführt (casting) und dann verglichen. Warum dies bei reichlich ausgestatteten Servern 6 Sekunden dauert, bleibt das Geheimnis von MySQL.

Der Bugfix sah am Ende wie folgt aus:

```
mysql> SELECT `configkey`, `configvalue` FROM `oc_appconfig`
WHERE (`appid` = 'files_sharing') AND (`configkey` IN
('70311', '231042', '312174'));

Empty set (0.01 sec)
```

OwnCloud übernimmt das Casting selbst und liefert statt 70311 den Wert '70311', also die Zahl als Zeichenfolge. Die Abfrage dauert dann 0.01 Sekunden statt der vorherigen 6.

Dieser Fehler war vermutlich schon sehr lange im ownCloud-Code. Aufgefallen ist er erst durch den Betrieb in unserer Konfiguration. Vor allem die Größe (Menge der Nutzer) brachte ihn zum Vorschein. 10 Nutzer auf einem Heimserver hätten damit vermutlich nie ein Problem bekommen. Schützen kann man sich vor solchen Fehlern nicht. Wichtig ist hier eine funktionierende Support-Infrastruktur und ein guter Draht zu den Entwicklern, die die Fehler dann gemeinsam mit den Betreibern schnell finden und auch sofort beheben können.

Der Bugfix stand uns nach wenigen Minuten zur Verfügung und der Dienst lief wieder einige Zeit absolut problemlos.

4.3.2 Das verschwundene Fachgebiets-Share

Aus einem Fachgebiet meldete eine Person (hier Bob genannt), dass alle Fachgebietsdaten plötzlich verschwunden sein. Eine solche Meldung ist selbstverständlich alarmierend, da die Zuverlässigkeit und Datenintegrität selbstverständlich obersterstes Ziel eines Cloudspeicher-Dienstes ist und jeder Anhaltspunkt, Daten könnten verloren gehen sofort eingehend untersucht wird.

Die Nachforschungen ergaben, dass auch Alice, die im selben Fachgebiet arbeitete nicht mehr auf die Fachgebietsdaten zugreifen konnte. Seltsamer Weise konnte jedoch nicht festgestellt werden, dass es jemals irgendwelche Daten im „Teamshare“ des Fachgebiets gab.

In der tubCloud wurden sogenannte „Teamshares“ implementiert. Eine Einrichtung an der Universität (z. B. ein Fachgebiet) wird in der tubCloud durch einen Benutzer repräsentiert, für den niemand ein Passwort o. ä. besitzt. Dieser „Service-Nutzer“ wird vom System gesteuert und dient letztlich nur dazu, den Mitgliedern des jeweiligen „Teams“ ein Verzeichnis mit Schreib-Lese-Berechtigung zu teilen. Dies geschieht über die APIs der Anwendung. Alle Nutzer eines Teams oder einer Einrichtung finden so einen geteilten Ordner in ihrer tubCloud, in die sie Dateien ablegen oder dort finden können. Die Steuerung der Teams geschieht über Teamfunktionen, die schon länger im Self-Service-Bereich des Rechenzentrums bekannt waren und lediglich um eine tubCloud-Ressource erweitert wurden.

Nachforschungen ergaben schließlich, dass das „Teamshare“ niemals genutzt wurde. Stattdessen hatte ursprünglich die Nutzerin Carol ein Verzeichnis an alle Fachgebietsmitglieder geteilt. Dass die Daten alle zu Lasten der Quota von Carol gingen, war dabei nicht weiter wichtig. Carol stand ohnehin kurz vor dem Ausscheiden aus dem Dienst an der TU Berlin. Und genau hierin bestand das Problem: Nicht zuletzt aus Gründen des Datenschutzes werden nach einer vereinbarten Übergangsfrist alle Daten eines ausgeschiedenen TU-Mitglieds gelöscht. Nachdem nun Carol in den Ruhestand gegangen war, standen die Daten noch eine gewisse Zeit zur Verfügung. Carol bekam verschiedene Warnungen per E-Mail, dass ihre Daten bald gelöscht werden würden, womit Carol auch gar kein Problem hatte. Dass die vermeintlichen Fachgebietsdaten jedoch allein Carol zugeordnet waren und somit mit dem Weggang von Carol ebenfalls verschwinden würden, war niemandem im Fachgebiet klar.

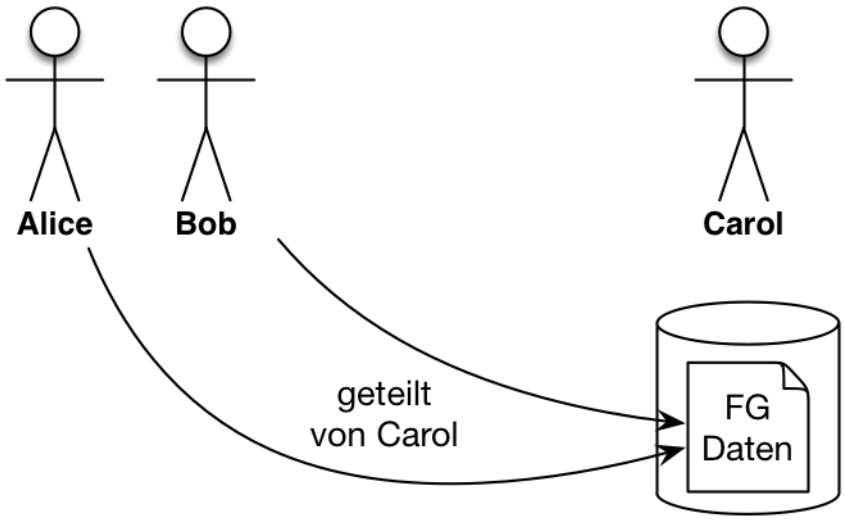

Abbildung 3: Problem persönliche vs. Teamshares

Dies führt zu einem allgemein immer dringender werdenden Problem von IT-Systemen und IT-Sicherheit. Alle Nutzerinnen und Nutzer von Diensten müssen über die sachgerechte Nutzung der Dienste aufgeklärt werden. Viele Seiteneffekte sind Nutzern nicht klar. Ein bestimmtes Verhalten der IT wird angenommen, aber nicht überprüft. Der Ausweg besteht darin, einfache Leitlinien an die Hand zu geben und diese gemäß der Erfahrungen mit den Systemen immer weiter auszubauen.

Im Gegensatz zu persönlichen Daten bleiben Teamshares so lange bestehen, wie das Team bzw. die Einrichtung an der TU Berlin existiert. Die klare Empfehlung ist daher, möglichst alles, was nicht persönlich ist und was Bestand haben soll, in diese Teamshares zu packen, wo es auch direkt von den jeweiligen Teams zugegriffen werden kann. Ein weiterer Vorteil besteht darin, dass die Mitgliedschaft z. B. von Fachgebietsshares, auch durch die Zuordnung zu dem Fachgebiet in der Personalstelle gesteuert wird. Neu eingestellte Mitglieder haben somit automatisch Zugriff auf die Daten, wechselt eine Kollegin/ein Kollege, wird auch der Zugriff automatisch angepasst.

Solche Funktionen sind Sync-n-Share Diensten klassischer Weise nicht implementiert. Sie können (und sollten) jedoch durch Kopplung an das IDM bzw. durch Integration in bereits vorhandene Self-Service-Funktionen implementiert werden. Unter anderem hieraus ergibt sich der Mehrwert, der eine Nutzung von z.T. rechtlich bedenklichen Diensten, wie Dropbox, iCloud, Google Drive und Co. unattraktiv macht.

5 Ausblick

Nach einer längeren Anlaufphase kommt die Nutzung der DFN-Cloud langsam in Schwung. Die Zusammenarbeit mit dem DFN und den Erprobungspartnern ist dabei über die Erprobung von Sync-n-Share Diensten hinaus effektiv. Die Ideen der verschiedenen Partner befruchten sich gegenseitig. So gibt es neben Skalierungsvorteilen eines großen Dienstanbieters auch weitere inhaltliche Vorteile für uns als Anbieter eines solchen Dienstes. Im Jahr 2017 stehen das erste Mal Erweiterungen an, die perspektivisch für DFN-Cloud Kunden geplant sind. Bislang konnten die Ressourcen aus nicht (mehr) genutzter Infrastruktur der TU Berlin bezogen werden.

Unser Schwerpunkt wird weiter auf der Verbesserung des Service und der Qualität des Dienstes liegen. Aber auch die Funktionalität wird in den nächsten Monaten erweitert werden, um den wachsenden Anforderungen der Partner gerecht zu werden.

6 Literatur

[Rö16] Michael Röder, Das erste Jahr der DFN-Cloud, Verträge, Statistiken und Tendenzen, Cloudspeicher im Hochschuleinsatz 2015, Universitätsverlag der TU Berlin, 2016

[Hi17] Thomas Hildmann, Evolution der ownCloud-Installation an der TU Berlin, Clouddienste im Hochschuleinsatz 2016/2017, Universitätsverlag der TU Berlin, 2017

Ein Jahr sciebo: Wie schlägt sich die Campuscloud gegen Dropbox, iCloud, Google Drive & Co.? Ergebnisse einer hochschulübergreifenden Befragung

Dominik Rudolph, Anne Thoring, Raimund Vogl
Zentrum für Informationsverarbeitung
Westfälische Wilhelms-Universität Münster
Röntgenstraße 7-13
d.rudolph@uni-muenster.de
r.vogl@uni-muenster.de
a.thoring@uni-muenster.de

Abstract: Das vorliegende Paper präsentiert die Ergebnisse einer hochschulübergreifenden Befragung zur Nutzung von Cloud-Diensten im Hinblick auf die Bewertung des ein Jahr zuvor im Februar 2015 gestarteten Dienstes sciebo. Sciebo wird von den Hochschulen selbst betrieben. Trotz der starken Konkurrenz durch etablierte kommerzielle Dienste zeigen die Ergebnisse, dass sciebo aus Sicht der Nutzer eine mindestens gleichwertige Alternative darstellt.

1 Einleitung

Knapp ein Jahr nach dem Start des Cloudspeicherdienstes sciebo wurde an den Teilnehmerhochschulen des betreibenden Sync & Share NRW-Konsortiums eine Onlinebefragung zur Nutzung von Cloud-Diensten durchgeführt, um den Erfolg des Projektes messen zu können und um potentielles Verbesserungspotential zu identifizieren. Insgesamt beteiligten sich über 18.000 Studierende und Beschäftigte im Dezember 2015 an der Befragung.

Ziel des Dienstes „sciebo – die Campuscloud“ ist die Schaffung einer sicheren Alternative zu kommerziellen, aus Datenschutzgründen kritisch zu sehenden Produkten. Da die Hochschulangehörigen eine hohe Autonomie genießen und ihnen die Nutzung eines bestimmten Dienstes nur sehr bedingt vorgeschrieben werden kann, muss sciebo auch in Puncto Nutzerkomfort und Funktionsumfang mit den etablierten Angeboten konkurrieren können, obwohl dem Dienst deutlich weniger Ressourcen zur Verfügung stehen. Um einen Überblick über die Cloud-Nutzung insgesamt zu bekommen und sciebo im Vergleich zur kommerziellen Konkurrenz einordnen zu können, wurden sämtliche Hochschulangehörigen befragt, also auch Nutzer kommerzieller Dienste und solche, die gar keine Clouddienste nutzen.

2 Ergebnisse

Rund 86 Prozent der Befragten setzen Clouddienste ein. sciebo konnte mit einer Nutzungsquote von 31 Prozent bereits ein knappes Jahr nach dem Start den zweiten Platz der genutzten Clouddienste hinter Platzhirsch Dropbox erobern und damit so namhafte Anbieter wie Apple (iCloud), Google (Drive), Microsoft (OneDrive) und Amazon (Cloud Drive) hinter sich lassen, auch der Bekanntheitsgrad ist mit 65 Prozent schon recht hoch. Unter Beschäftigten kommt sciebo sogar bereits auf 65 Prozent. Maß aller Dinge bleibt aber Dropbox mit einer Nutzungsrate von 77 Prozent. Interessanterweise nutzen die meisten Nutzer mehrere Dienste parallel, im Schnitt 2,3.

Das spricht für unterschiedliche Stärken und sich ergänzende Nutzungsweisen: So dürfen etwa bei sciebo nur hochschulbezogene Daten gespeichert werden, für private Daten ist ein weiterer Anbieter erforderlich. Ein weiterer Grund könnte in der Datenmengenbeschränkung vieler Anbieter liegen.

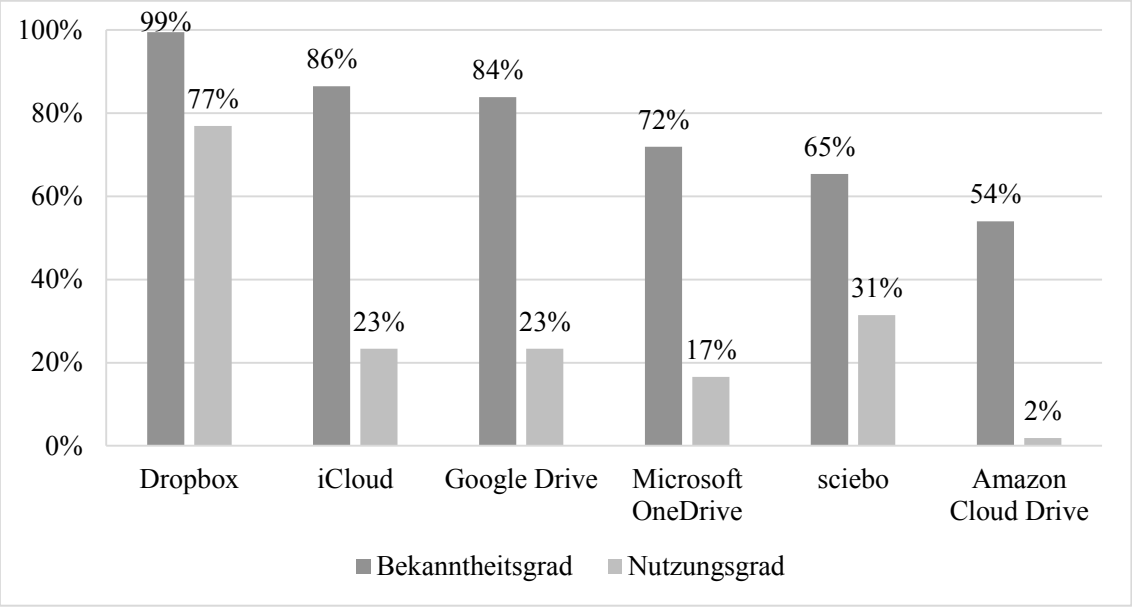

Abbildung 1: Bekanntheits- und Nutzungsgrad verschiedener Clouddienste

Als wichtigste Informationsquelle sind insbesondere Informationen der Hochschule, z. B. Rundmails, aber auch Kommilitonen bzw. Kollegen zu nennen. Wesentliche Funktion aus Nutzersicht ist das Synchronisieren von verschiedenen Endgeräten, die Sicherung einzelner Dateien sowie das Teilen mit Angehörigen der eigenen Hochschule. Die angebotene Speichergröße von 30 GB (für Studenten) reicht den meisten Nutzern aus, rund 83 Prozent nutzen sogar nach eigener Einschätzung weniger als 10 GB. Auch die Beschäftigten, die potentiell 500 GB nutzen können, nutzen nur zu 8 Prozent mehr als 30 GB.

Ziel des Sync & Share-Projektes ist es, die Nutzung kommerzieller Angebote aus dem Hochschulkontext zu verdrängen. Dies scheint gelungen zu sein: von den sciebo-Nutzern geben 47 Prozent an, nun andere Dienste weniger zu nutzen, weitere 12 Prozent sind vollständig zu sciebo gewechselt. Insbesondere auf dienstlichen Rechnern wird die Campuscloud deutlich häufiger genutzt als die kommerzielle Konkurrenz. Auf Smartphones und Tablets sind dagegen andere Dienste noch im Vorteil. 72 Prozent nutzen den Synchronisationsclient, d.h. ein Viertel der Nutzer greift ausschließlich über das Webinterface auf die Daten zu. Rund 38 Prozent der Clientnutzer wählen selektiv die zu synchronisierenden Ordner aus.

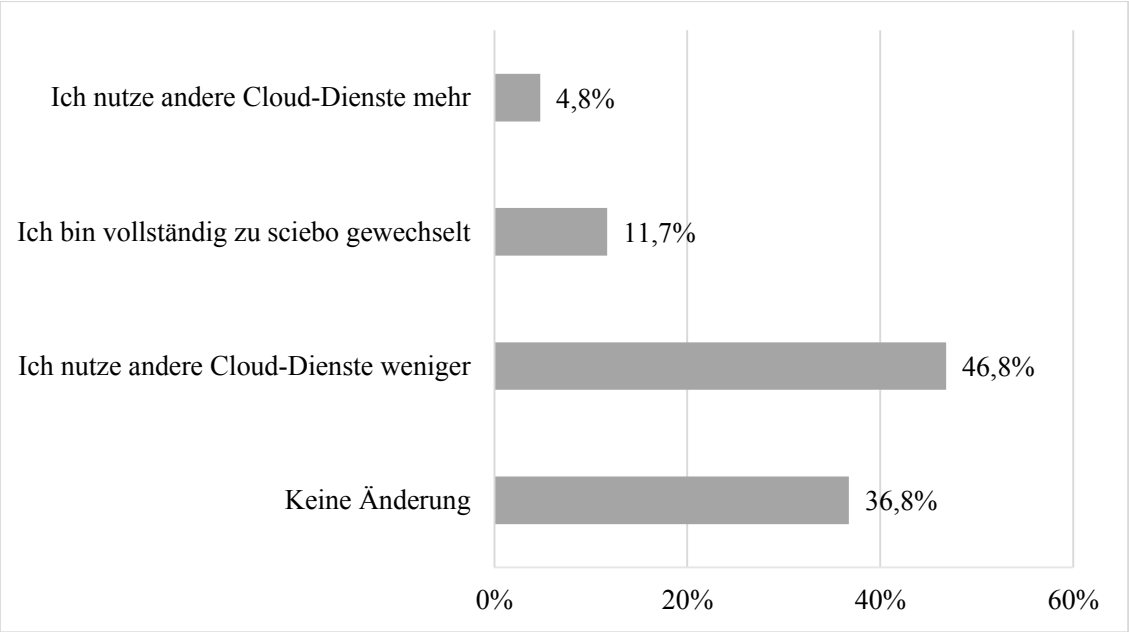

Abbildung 2: Auswirkungen von sciebo auf die Nutzung anderer Cloud-Dienste

Nutzungsgründe von sciebo sind insbesondere die hohe Sicherheit, das große Datenvolumen und die Nutzung durch Kommilitonen bzw. Kollegen, aber auch, dass es ein Angebot der eigenen Hochschule ist. Gleichzeitig wird sciebo von seinen Nutzern besser beurteilt als alle anderen Clouddienste von deren jeweiligen Nutzern. In fünf von sieben Kriterien (Sicherheit, Kosten, Datenvolumen, Zuverlässigkeit und Gesamtbewertung) liegt sciebo vor seinen Konkurrenten, lediglich beim Funktionsumfang und der Handhabung liegen andere Dienste vorne. So überrascht es nicht, dass vier von fünf Personen, die sciebo kennen, den Dienst zukünftig nutzen wollen. 72 Prozent der sciebo-Nutzer haben den Dienst bereits weiterempfohlen oder wollen dies noch tun.

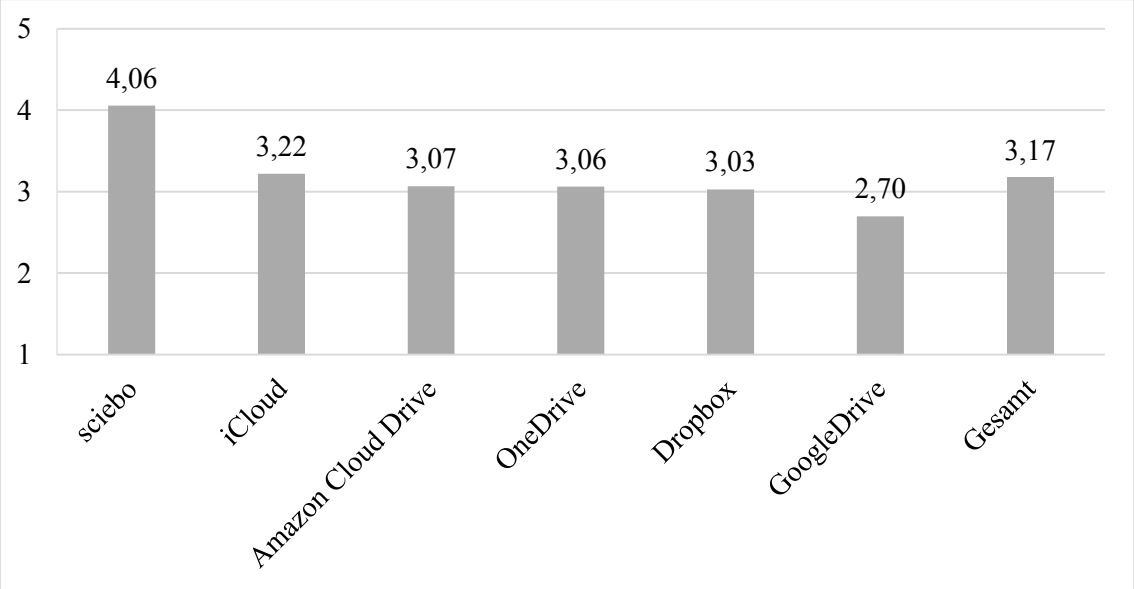

Abbildung 3: Vertrauen von Dienste-Nutzern zu den jeweiligen Betreibern;
Mittelwerte 5-stufige Likert-Skala.

Bemerkenswert: über 90 Prozent vertrauen den Betreibern von sciebo, also dem ZIV der WWU Münster und seinen Partnern im Sync & Share NRW-Konsortium. Zum Vergleich: nur 35 Prozent der Dropbox-Nutzer vertraut den Dropbox-Betreibern. Im Gegensatz zu den kommerziellen Anbietern werden die Daten bei sciebo ausschließlich von den Hochschulen selbst an Standorten in NRW (Münster, Bonn, Essen) gespeichert und unterliegen damit dem strengen deutschen Datenschutz. Auf den guten Bewertungen kann sich sciebo aber nicht ausruhen, da die Mehrheit der Nutzer den Wechselaufwand zu einem anderen Dienst als sehr gering einstuft.

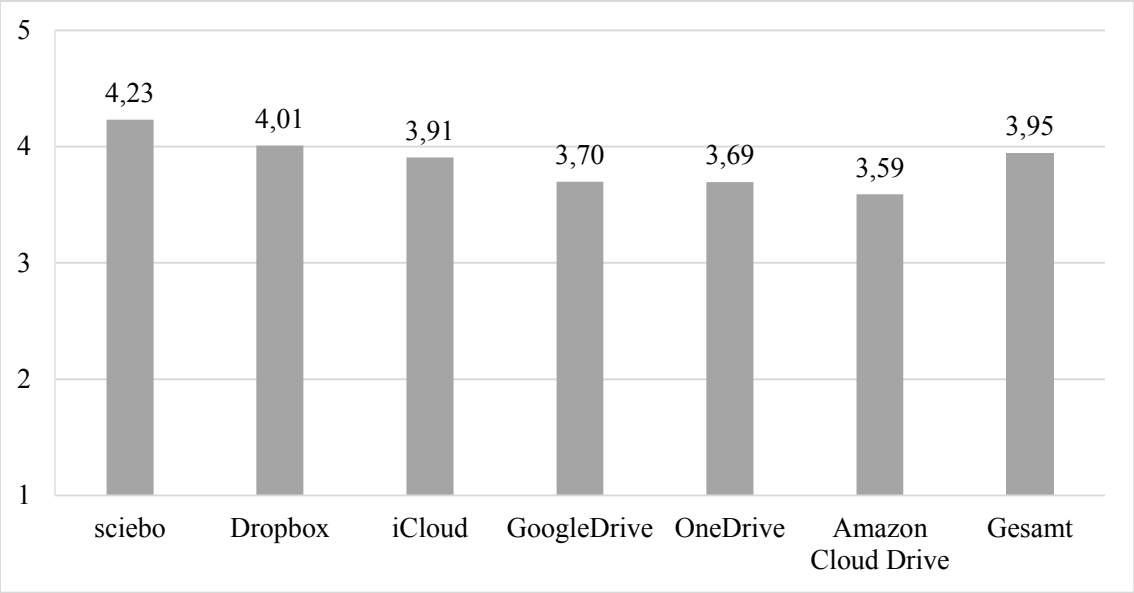

Abbildung 4: Gesamtbewertung; Mittelwerte 5-stufige Likert-Skala.

Das große Interesse von Studierenden und Mitarbeitern an sciebo demonstriert vor allem auch die große Zahl an Anregungen, die im Rahmen der Umfrage eingegangen ist: Mehr als 2.500 teils sehr ausführliche Kommentare liefern neben Kritik und positiven Rückmeldung auch allerlei Ideen und Vorschläge zur Optimierung der Campuscloud, die in die Weiterentwicklung von sciebo einfließen werden.

3 Zusammenfassung

Das Ziel des Sync & Share-Projektes mit dem Dienst "Sciebo – Die Campuscloud ist die Etablierung eines Cloudspeicherdienstes, der dank der lokalen Speicherung der Daten in der Hand der Hochschulen selbst unter der Hoheit des deutschen Datenschutzgesetzes eine sichere Alternative zu kommerziellen Anbietern wie Dropbox und Google darstellt. Anders als in Unternehmen lassen sich die potentiellen Nutzer – Studierende und Hochschulmitarbeiter – wegen der Freiheit von Forschung und Lehre kaum zur Nutzung einer bestimmten Software zwingen, wenn diese nicht ihren Bedürfnissen entspricht. Die sciebo-Betreiber stehen also vor der großen Herausforderung, einen Dienst anzubieten, der den etablierten und mit großen Ressourcen ausgestatteten kommerziellen Angeboten nicht nur in puncto Sicherheit überlegen ist, sondern auch in der Nutzerfreundlichkeit mindestens ebenbürtig ist. Dies scheint gelungen zu sein, dafür sprechen nicht nur die mittlerweile rund 80.000 Nutzer (Stand Juni 2017), sondern auch die Ergebnisse der hier diskutierten Befragung.

The Adoption of a New Cloud Storage Service in German Universities

Raimund Vogl, Dominik Rudolph, Holger Angenent,
Anne Thoring, Andreas Wilmer, Christian Schild

Zentrum für Informationsverarbeitung
Westfälische Wilhelms-Universität Münster
Röntgenstraße 7-13
r.vogl@uni-muenster.de
d.rudolph@uni-muenster.de
holger.angenent@uni-muenster.de
a.thoring@uni-muenster.de
a.wilmer@uni-muenster.de
schild@uni-muenster.de

Abstract: Cloud services like Dropbox or Google Drive are widely used in the academic community because of their obvious advantages for data availability and collaboration. However, in terms of data protection their establishment in this sector is a sensitive issue: Most commercial providers' terms and conditions are in conflict with universities' data protection regulations. As a solution, "sciebo" was launched at the beginning of 2015, a safe and easy-to-use cloud storage service operated by 25 universities in the German state of North Rhine-Westphalia (NRW). Bearing in mind that the great success of commercial providers creates a tough competitive environment for a new service, but also represents a potential user base that is aware of and familiar with cloud services, this paper opposes diffusion theory to reality using the example of sciebo.

1 Introduction

Making it possible to easily share documents with others and to synchronize data across multiple devices, cloud storage services like Dropbox or Google Drive have become quite popular in last five years – not least in the academic context, among students and researchers. Commercial services are very comfortable in use, but security concerns about their data utilization arise, especially after the Snowden disclosures. In 2013, as a consequence, the majority of the public research and applied science universities in the German state of North Rhine-Westphalia (NRW) formed a consortium to start a jointly operated private cloud service for the academic community. This sync and share storage platform should be free of charge, easy to use and, most importantly, it should be hosted on premise at several university data centers to be fully compliant with German data protection regulations [1]. With respect to the software functionality and the required hardware setup for potentially 500,000 users, the system design was grounded on empirical user studies.

A first exploratory survey on the demand for a university operated alternative to Dropbox etc. was conducted among potential users at Münster University in 2012 and extended to a multi-site survey with more than 10,000 participants from three major universities in late 2013 [2]. Both surveys focused on the participants' intention to use such a university operated cloud service, their demand for storage space and client platforms, the type of content (file types) they intended to store, and the communities they wanted to collaborate with using the service's file sharing functionalities. The procurement of the software solution as well as the sizing of the hardware

platform were based on the adoption and usage estimates derived from these surveys. In February 2015, after extensive preparatory work done for the funding proposal, the procurement process, and the system setup and commissioning, the sync and share cloud storage service was launched under the brand name "sciebo" (being short for "science box") with three university data centers (Bonn, Duisburg-Essen and Münster) hosting the system platforms on premise.

The case of sciebo is unique because it allows us to observe the diffusion of a technical innovation from the beginning in a well-controlled setting. There is plenty of literature about the adoption of cloud systems in organizations like SMEs [3–5] or special industries [6–13], but only little is known about the adoption behavior of end-users who can decide freely if they want to use a new cloud service or not [14]. Universities are a special case: On the one hand, they are organizations with a quite uniform population and a manageable size. On the other hand, because of the principle of freedom of research and teaching held high in Germany, there is no possibility to command the use of a system, so users have to be convinced.

After one year and three month of operation (as of 9 May 2016), exactly 47,647 users from 24 universities (out of 33 in NRW) and one public research center have signed up for sciebo through the self-enrollment web portal. Now is the right time to review how the initial expectations on service adoption and usage as well as system performance and availability correspond with reality.

2 Predictions

In preparing the sciebo project and applying for substantial funding, reliable predictions on the user adoption of this new service were crucial for the system design and amount of hardware to procure. For the university data centers volunteering to host the platform, estimates of the internet bandwidth to be dedicated to sciebo was important, and for the universities that had to decide if they wanted to join the sciebo project consortium it was necessary to know what quality of service, especially with respect to system availability, they could expect. Thus, based on empirical research, predictions were made on the required storage volume for the sciebo system platform and the required internet bandwidth – both directly connected to the adoption of the new service by its eligible users. System availability scores were estimated based on the analysis of three years of well documented operation incidents at the University of Münster.

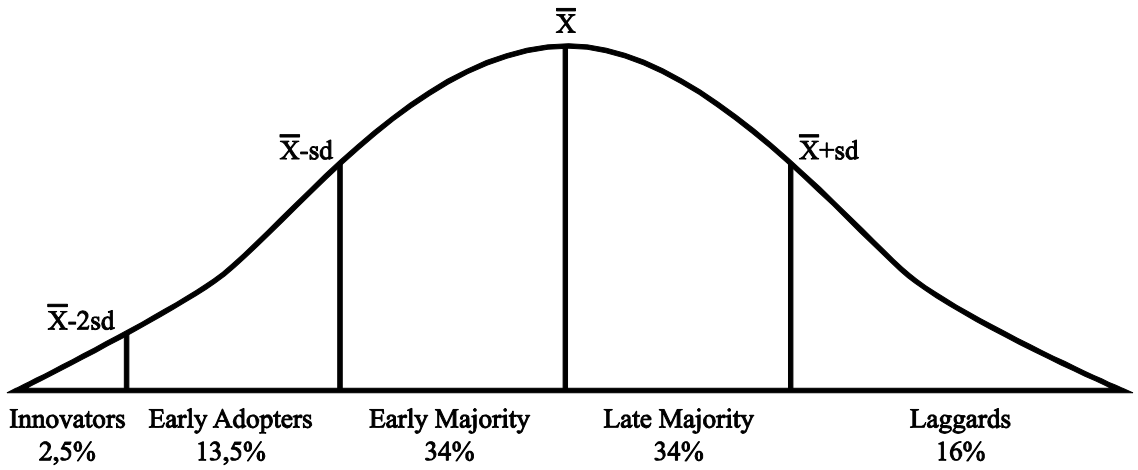

Fig. 1. Adopter categories according to Rogers [15]

Predictions on the diffusion of sciebo were made on the basis of Rogers' Diffusion of Innovations Theory. According to Rogers [16], innovativeness, i. e. the readiness and the degree to which a person or an organization adopts an innovation (e. g. a new product) compared with the other members of his population, follows the Gaussian distribution. He identifies five adopter categories with different characteristics referring to their innovativeness: innovators, early adopters, early majority, late majority and laggards (Fig. 1). If you accumulate the adoption decisions of all adopters over time, you get an S-shaped curve, the diffusion curve. The faster the innovation is adopted the more steeply this curve will rise. The speed of diffusion depends on the characteristics of the innovation, in particular its relative advantage compared to other existing products, compatibility with existing values and practices, simplicity and ease of use, trial ability and observable results.

As our data from a large user survey conducted in 2013 [2] show, Dropbox is the most used storage service among members of the universities with about 80 percent of market share. This value recurs in another survey conducted in 2015 at the same universities (unpublished work), so we conclude that Dropbox obviously has reached the saturation of demand five years after its inception in 2007 and two years after the release of the first stable version 1.0 in 2010. Examining Dropbox's worldwide diffusion, a flat growth is visible in the first two years and a take-off in the third year (Fig. 2).

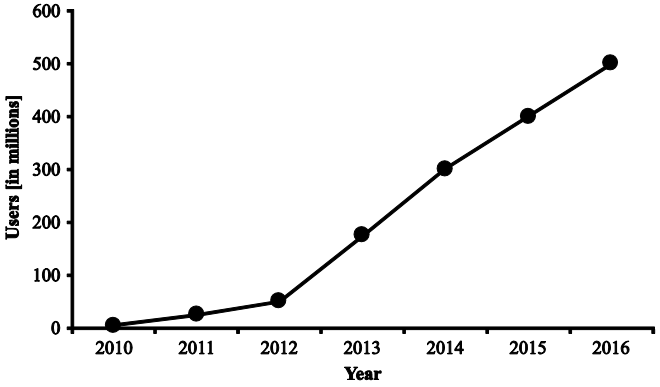

Fig. 2. Diffusion of Dropbox [15]

For sciebo, we predicted an even faster diffusion because the technology is already known from Dropbox. Moreover, sciebo's high security standards and bigger free storage space seem to be significant relative advantages as stated by participants of the survey in 2013 [2]. According to Diffusion Theory, market potential is not the total of all potential users (i. e. all members of the participating universities), but the total of all those persons who will realistically use a new service [17]. In the survey, 92.5 percent of the participants stated that they wanted to use sciebo. Being informed that their usage authorization would be revoked when leaving university, the count dropped to 65 percent. Thus, 65 percent of all members of the participating universities – that is about 252,000 individuals – constitute the estimated market potential of sciebo.

Based on the distribution of per user storage demands from the survey, we could refine the initial assumption that each user would need the planned 30 GB quota to the max. We were also able to predict an average storage volume of 8 GB (pessimistic scenario) to 16 GB (optimistic scenario) per user and could ascertain that a maximum storage space of 30 GB should fit most users. Assuming that users would switch their academic data from another platform to sciebo in the first days after the registration, we expected a quite linear growth with a 30 percent basis synchronization at the beginning and just small gain of 3 percent a month.

Considering the predictions on service adoption and storage demand, different scenarios were derived to estimate the size of storage systems to be procured and the internet bandwidth required. The total storage volume required for the operation of sciebo in the long term was estimated at 1.7 PB (pessimistic) to 5 PB (optimistic), and the internet connection bandwidth requirement for service operation was estimated at 3 Gbps in the optimistic scenario.

The predictions on system availability (based on three years of incident logs at Münster) resulted in an agreement amongst the sciebo consortium partners that the availability scores have to be 99.5 percent per year for each of the sites with a minimum of 98 percent per month.

3 Findings

3.1 User Diffusion

One year and three month after the official launch (9 May 2016), sciebo approaches another milestone with now approximately 48,000 users – this means an actual market share of 19 percent. In terms of the Diffusion Theory, this implies that sciebo's diffusion has reached the early majority phase.

However, diffusion speed varies significantly at the different universities. Fig. 3 shows the state of diffusion at the 14 universities that started sciebo in February 2015. The spectrum ranges from 8.2 percent at the University of Paderborn to 37,7 percent at RWTH Aachen. University size might serve as one explanation, as information should flow very fast at a small campus with a manageable number of departments. As stated by Rogers [16], diffusion can be seen as a communication process: In smaller and spatially closer populations, communication between the members is much more likely and easier than in a complex university with lots of different departments distributed over the whole city.

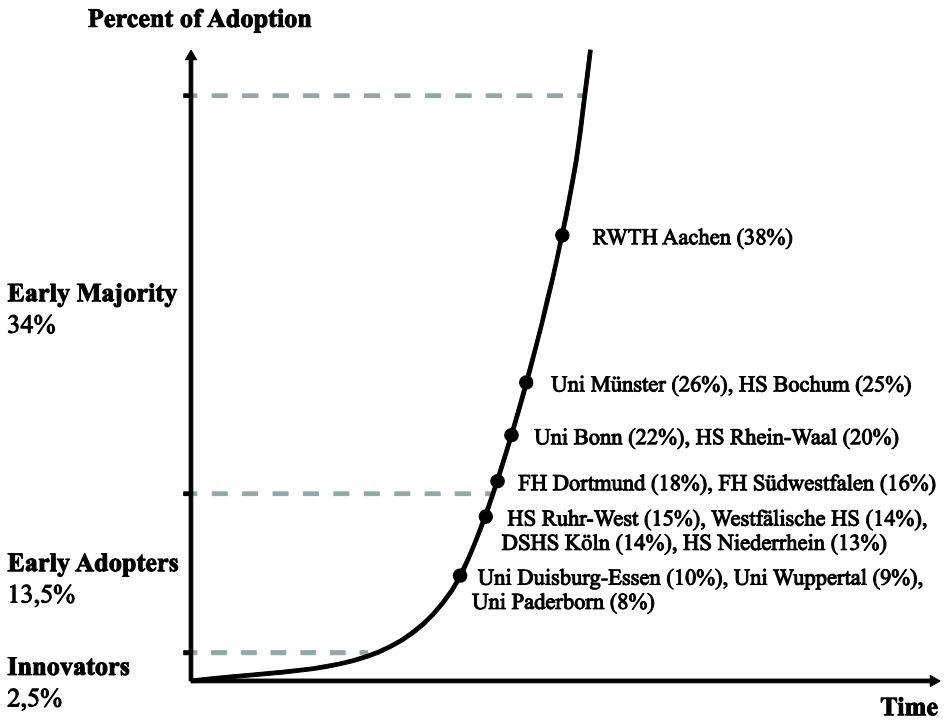

Fig. 3. Diffusion of sciebo on 9 May 2016 (taking into account only universities starting in February 2015)

Though in theory size suggests itself as a reason for the different diffusion speeds, it does not seem to be a good explanation in our case: Comparing same-sized universities – e.g. the Universities of Münster, Duisburg-Essen and RWTH Aachen with about 44,000 to 49,000 members each – the differences in market share are still evident (Fig. 4). Results show a remarkable variance of 27.2 percent between RWTH Aachen (37.7 %) and the University of Duisburg-Essen (10.5 %), with the University of Münster (26.2 %) ranking mid.

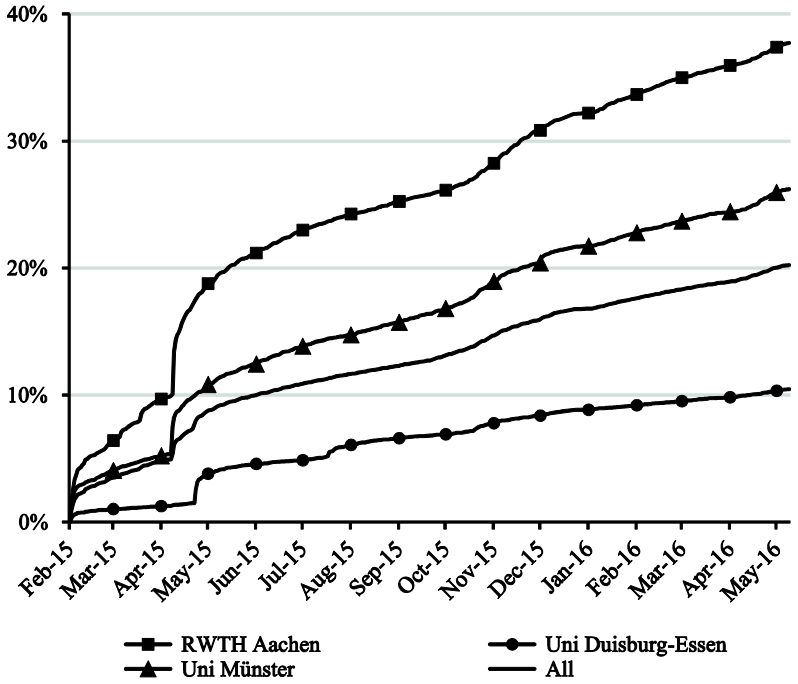

Fig. 4. Diffusion curves of selected universities with the same size compared to the diffusion curve of all universities starting in February 2015

Taking all universities starting in February 2015 into account, RWTH Aachen appears to be an outlier with its high market share. Both, the University of Münster (26.2 %) and the University of Duisburg-Essen (10.5 %), rank much closer to the overall average. One possible explanation for RWTH's high performance is that, unlike the Universities of Münster and Duisburg-Essen, RWTH is a technical university with many technophiles. They resemble the innovators described by Rogers and are the first to adopt new technologies. Logically, a technical innovation like sciebo diffuses faster in a technophile environment than in other populations.

The low performance of the University of Duisburg-Essen, compared with the same-sized universities and the overall average diffusion, is similarly interesting. A closer look reveals that the universities' commitment in terms of marketing activities might be another decisive factor. While RWTH Aachen and the University of Münster, in particular, performed a variety of marketing activities (e.g. direct mailings to all members), the University of Duisburg-Essen did not to that extent. Therefore, it is likely that only innovators and early adopters who are interested in innovations and actively search for information on their own account for their share of sciebo users. Further monitoring will show if an early majority can be reached with no marketing and just word of mouth, or if the number of users will be stagnating. According to some authors there is a gap between the early adopters and the early majority which has to be bridged by marketing activities [18, 19], while Rogers considers both groups as a continuum [16].

Examining the diffusion curves of the different universities (Fig. 4), deviations from the ideal S-curve of the diffusion model are clearly visible. Usually, they are caused by special events. The first boost in February 2015 is the official launch of the service. In the run-up we realized a large

Facebook campaign with posts in over 400 user groups related to the participating universities. Also, test users were added to the statistics. The second and largest user increase in April, at the start of the summer term in Germany, is triggered by direct mailings which most participating universities sent to their members. The diffusion curves of those universities passing up this opportunity show no such steep rise. In October most universities welcome their largest share of new students for the winter term which explains the next boost. In December, some universities used direct mailings to promote an online survey related to sciebo, again gaining attention and an additional boost for sciebo.

As regards storage space, we initially expected 9 GB (30 % of the intended per user quota limit of 30 GB) right after registration and a monthly growth of 3 percent (until the quota limit is reached). Currently, the average volume needed by an active user (i. e. a user who uploaded some data) is 3.6 GB, amounting to a total of 133.5 TB storage space used in sciebo.

3.2 Data Storage

In Fig. 5 we analyzed the storage load on an individual user basis. In particular, we looked at the dependency between the consumed storage space of a single account and its age. Shown is the mean used disk storage for user accounts in dependency of the account lifetime (solid black line), the 0.05-quantile (lower grey line) and the 0.95-quantile (upper grey line) in a logarithmic plot. The broken black line represents the expected and the dotted black line the observed linear model of the user behavior.

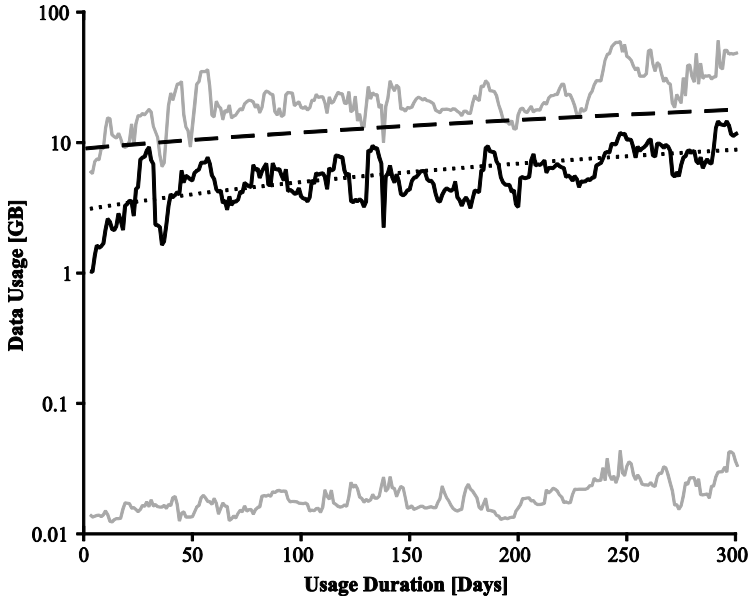

Fig. 5. Storage load on individual user basis per time vs. model (broken line)

Altogether 9,807 user accounts were analyzed on a day-wise basis. The statistical values were computed across an ensemble of user accounts for a specific account age. We restricted the analysis to active users with a used storage capacity of at least 10 MB. This eliminates seasonal side effects, such as the beginning of a new semester, and excludes subscripted, but unused accounts. Additionally, a moving average with a window size of 7 days accumulates the number of accounts to N=228±97 each day.

We expected that an account initially requires 30 percent of its full 30 GB quota and grows in a linear fashion with 3 percent of its quota per month on average. Our prediction can be written as a linear equation of the form f(x)=A+Bx. Here, the function f describes the consumed storage in

dependency of the time x and the coefficient A as the initial offset, B as the slope of the function, i. e. we hypothesized Aexp=9000 [MB] and Bexp=29.6 [MB/Day].

As a result we observed an offset Aobs=3052.5 ± 230.5 [MB] and a slope Bobs=19.2 ± 1.3 [MB/Day] with a linear Least-Squares Fit ($p<.001$ and adjusted R-Squared 0.415). The observed results show that on average a user synchronizes less data directly after the subscription than expected, but more than the 30 percent of the average storage space per user of 8 GB in the pessimistic scenario deduced from the survey findings. The growth of the synchronized data over time is also lower than in our expected model. The asymmetry between the quantiles and the mean is characteristic for a positively skewed underlying distribution. The exponential distribution is a good approximation for our data, which can be confirmed by a Lilliefors test ($p<.001$). It basically describes a random amount of files summing up to a total user quota for independent user accounts.

Furthermore, our findings suggest a wide variety of individual usage scenarios. Our service is capable to span the full spectrum from casual users synchronizing only very few files to really extensive usage cases on the scale of synchronized consumer hard drives. The mean storage usage is less than primarily expected but significantly above our pessimistic scenario. One reason might be that we discriminated active from inactive accounts just by storage usage and not by last login or access time. Another bias to less storage usage is simply given by the fact that share files or folders just charge the quota of the share creator and not the recipient.

3.3 Bandwith

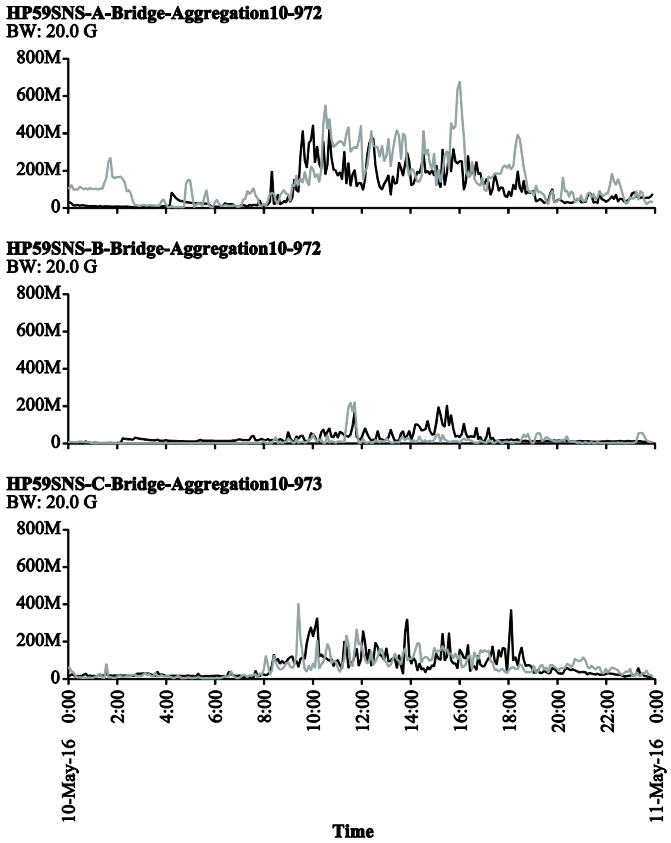

Fig. 6. Internet bandwidth (in=black, out=grey) for end-user access to sciebo consumed by the three sciebo sites at Bonn (BN), Duisburg-Essen (DU) and Münster (MS) on a typical day after one year and three month of operation

The initial estimates of bandwidth requirements were essential to make sure that the internet connection bandwidth of the three university data centers hosting the sciebo platform was not entirely consumed by the new sciebo service. Based on simply models of service utilization (up- and downloads) an overall limit of 3 Gbps sustained for the whole sciebo system, thus approximately 1 Gbps for each of the data center sites, was predicted as being sufficient.

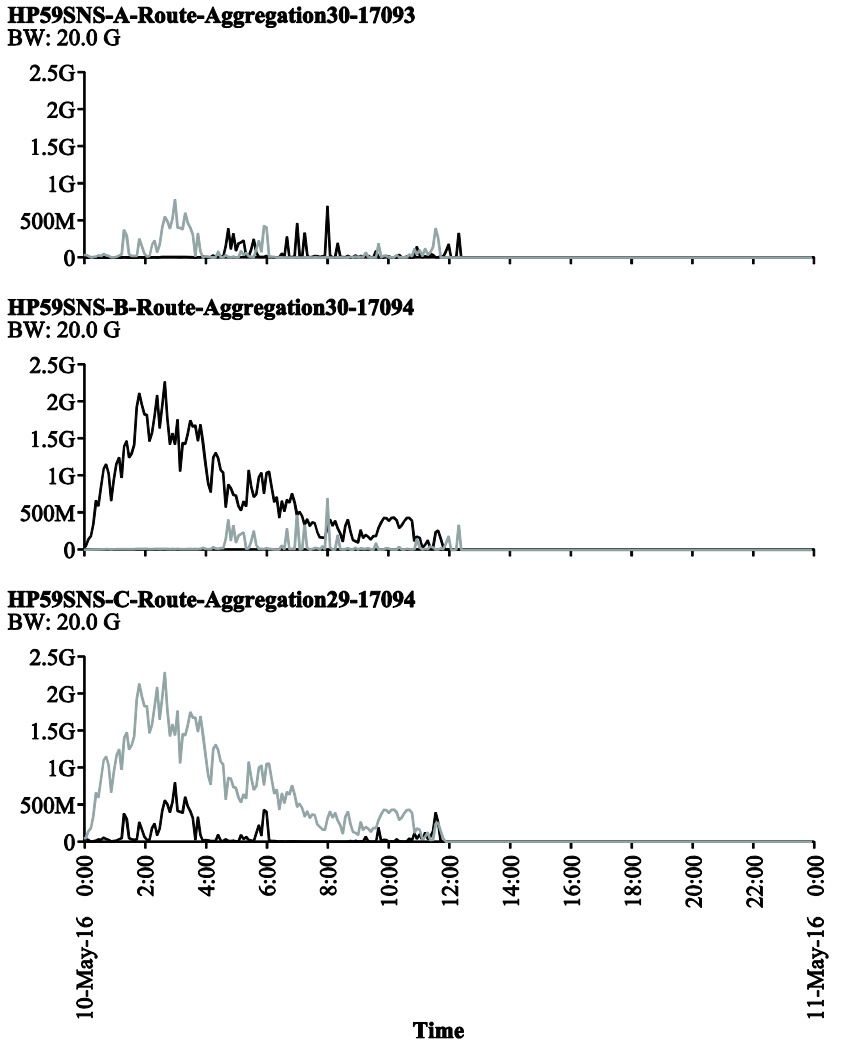

Fig. 7. Internet bandwidth (in=black, out=grey) for mutual data backup/replication between the three sciebo sites in the 12am to 12pm timeframe

One year and three month after the start of operation, peak data rates for end-user access to sciebo around 1 Gbps (adding IN and OUT traffic) can be frequently found on a typical day of operation for the site with the highest user load (Fig. 6). With continuous growth of the sciebo user base and storage volume, bandwidth demands will necessarily grow further, but negative effects on the internet connectivity of the hosting universities (each currently has a 10 Gbps internet link) are, as initially predicted, not to be expected, especially since traffic policies limiting the bandwidth allocated to individual connection could still be imposed. The mutual data backups (rsync replication) between the three sites are scheduled in the 12am to 12pm timeframe where end-user service utilization is low. This traffic is routed through a separate Layer3 VPN connecting the data center backend systems of the three sciebo sites (Fig. 7), which shares its bandwidth with the main internet connectivity of the hosting universities, though. For an extended period of time in the night, sustained bandwidth consumption of 1.5 Gbps to 2.0 Gbps can be seen here.

3.4 System Availability

To ensure high system availability for sciebo according to the agreed on availability scores, a set of measures was taken with respect to resilient system design and has proven very effective in the first year of operation. Availability monitoring of the complete sciebo service stack through periodic automated ownCloud file access operations from probes at all three sites checking the respective other two sites with event correlation using NAGIOS and Check_MK. The recorded values from Check_MK (Table 1) show that the monthly availability scores (exemplified here by Jan-Apr 2016) for Sites B and C are well in line with the availability scores previously agreed on amongst the consortium partners of a minimum of 98 percent per month and the agreed target value of 99.5 percent availability per year for each of the sites is reached. But for Site A, due to severe software problems that were mostly impacting this site which has the highest user load, the March 2016 monthly availability fell below the agreed 98 percent target and even the one year availability score March 2015 to April 2016 was slightly down below 99.5 percent – thus a plan for communication and mitigation had to be implemented.

TABLE I
AVAILABILITY SCORES (IN %) FOR THE THREE SITES HOSTING SCIEBO (ANONYMIZED)

	Site A	Site B	Site C
Jan 2016	99.76	100.00	99,93
Feb 2016	99.16	99.66	99.56
Mar 2016	95.69	99.91	98,52

These results are also comparable with numbers publicly disclosed on Dropbox, where 99.63 percent to 99.85 percent availability (where the lower number includes unconfirmed downtimes) for the service was monitored by independent sources for the July to December 2012 period.

3.5 Additional Findings

Apart from those findings related to our predictions, some additional outcomes are worth mentioning. The first finding broaches the issue of user activity: 38 percent of the registered users are inactive, i. e. they have not uploaded any data to sciebo yet. Based on Rogers' Diffusion Theory [16], this inactivity of a substantial user fraction could be interpreted as either a prolonged phase of decision making or as discontinuance (without having used the service apart from signing up) [20, 21]. This finding needs further research.

The second finding focuses on the key collaboration feature of sciebo – sharing data with other sciebo users or externals (share via hyperlink). With an overall average of 2.5 shares per active user, this feature is not used very strongly yet. Folders (66.4 %) are shared more often than files (33.6 %). Approximately 51 percent of all shares are performed via link (primarily intended for external exchange), contrary to expectations from the survey [2], where 65 percent of the participants intended to share within their university and only 21 percent intended to share with externals.

4 Conclusion

These first results show that the predictions on service adoption and system availability made in the design phase of the sciebo service do well conform to the reality of one year of operation. Especially the prognoses on required system platform parameters phrased in the aftermath of the 2013 survey [2] are – up to now – in line with the service's adoption, and, more-over, Rogers' Diffusion Theory [16] has proved to be an adequate model. We could identify two factors influencing the speed of diffusion of the sciebo cloud-service:

1. Share of technophiles in the organization
2. Use of marketing measures

Both findings are supported by the diffusion model. As known from the diffusion literature, an innovation is more likely to be adopted if it is not too complex and consistent with known products. Consequently, technophiles who understand a technical innovation much better and usually find it less complex than other people, will be more likely to adopt an innovation quickly. As noted by some authors, there might be a gap – in terms of missing peer-to-peer connections – between innovators and early adopters on the one hand and the early majority on the other hand, because of the significant differences between those groups [19]. Marketing measures like direct mailings, Facebook posts, YouTube videos etc. can bridge this gap by informing the early majority about a new service, and thus speed up the diffusion process. According to our data, organization size does not influence the diffusion speed.

Finally, the universities' heterogeneous rate of adoption and the high fraction of inactive users leave a wide field for further research. In the upcoming months, analyzing the reasons for discontinuance of use will be a key focus.

5 References

[1] R. Vogl *et al,* "Designing a Large Scale Cooperative Sync&Share Cloud Storage Platform for the Academic Community in Northrhine-Westfalia," in *ICT Role for Next Generation Universities - 19th European University Information Systems*, Riga: Riga Technical University, 2013.

[2] S. Stieglitz, C. Meske, R. Vogl, and D. Rudolph, "Demand for Cloud Services as an Infrastructure in Higher Education," in *ICIS 2014 Proceedings*, 2014.

[3] Y. Alshamaila, S. Papagiannidis, and F. Li, "Cloud computing adoption by SMEs in the north east of England: A multi-perspective framework," *Journal of Enterprise Information Management*, vol. 26, no. 3, pp. 250–275, 2013.

[4] S. R. Tehrani and F. Shirazi, "Factors influencing the adoption of cloud computing by small and medium size enterprises (SMEs)," in *Human Interface and the Management of Information. Information and Knowledge in Applications and Services*: Springer, 2014, pp. 631–642.

[5] S. Trigueros-Preciado, D. Pérez-González, and P. Solana-González, "Cloud computing in industrial SMEs: identification of the barriers to its adoption and effects of its application," *Electronic Markets*, vol. 23, no. 2, pp. 105–114, 2013.

[6] C. G. Cegielski, L. Allison Jones-Farmer, Y. Wu, and B. T. Hazen, "Adoption of cloud computing technologies in supply chains: An organizational information processing theory approach," *The international journal of logistics Management*, vol. 23, no. 2, pp. 184–211, 2012.

[7] P. Gupta, A. Seetharaman, and J. R. Raj, “The usage and adoption of cloud computing by small and medium businesses,” *International Journal of Information Management*, vol. 33, no. 5, pp. 861–874, 2013.

[8] M. Ivanova and G. Ivanov, “Cloud computing for authoring process automation,” (da), *Procedia-Social and Behavioral Sciences*, vol. 2, no. 2, pp. 3646–3651, 2010.

[9] S. Khanagha, H. Volberda, J. Sidhu, and I. Oshri, “Management innovation and adoption of emerging technologies: The case of cloud computing,” *European Management Review*, vol. 10, no. 1, pp. 51–67, 2013.

[10] J.-W. Lian, D. C. Yen, and Y.-T. Wang, “An exploratory study to understand the critical factors affecting the decision to adopt cloud computing in Taiwan hospital,” *International Journal of Information Management*, vol. 34, no. 1, pp. 28–36, 2014.

[11] C. Low, Y. Chen, and M. Wu, “Understanding the determinants of cloud computing adoption,” *Industrial management & data systems*, vol. 111, no. 7, pp. 1006–1023, 2011.

[12] H. Moryson and G. Moeser, “Consumer Adoption of Cloud Computing Services in Germany: Investigation of Moderating Effects by Applying an UTAUT Model,” *International Journal of Marketing Studies*, vol. 8, no. 1, p. 14, 2016.

[13] T. Oliveira, M. Thomas, and M. Espadanal, “Assessing the determinants of cloud computing adoption: An analysis of the manufacturing and services sectors,” *Information & Management*, vol. 51, no. 5, pp. 497–510, 2014.

[14] J. Shin, M. Jo, J. Lee, and D. Lee, “Strategic management of cloud computing services: Focusing on consumer adoption behavior,” *Engineering Management, IEEE Transactions on*, vol. 61, no. 3, pp. 419–427, 2014.

[15] Dropbox, *Stats.* [Online] Available: https://www.dropbox.com/. Accessed on: Jun. 12 2015.

[16] E. M. Rogers, *Diffusion of innovations,* 5th ed. New York: Free Press, 2003.

[17] M. Kleinaltenkamp and W. Plinke, *Markt- und Produktmanagement: Die Instrumente des technischen Vertriebs*. Berlin: Springer, 1999.

[18] C. M. Christensen, *The innovator's dilemma: The revolutionary book that will change the way you do business*. New York, NY: Harper Business, 2011.

[19] G. A. Moore, *Crossing the chasm: Marketing and selling disruptive products to mainstream customers,* 3rd ed. New York, NY: Harper Business, 2014.

[20] W. Black, “Discontinuance and Diffusion: Examination of the Post Adoption Decision Process,” *Advances in Consumer Research*, vol. 10, no. 1, 1983.

[21] M. Parthasarathy and A. Bhattacherjee, “Understanding Post-Adoption Behavior in the Context of Online Services,” *Information Systems Research*, vol. 9, no. 4, pp. 362–379, 1998.

Evolution der ownCloud-Installation an der TU Berlin

Dr.-Ing. Thomas Hildmann
tubIT - IT Service Center
Technische Universität Berlin
Einsteinufer 17
10587 Berlin
thomas.hildmann@tu-berlin.de

Abstract: Die Entwicklung der Cloudspeicher-Installation an der TU Berlin folgt den allgemeinen Trends im Rechenzentrum. Mangels frei zugänglicher Erfahrungsberichte wurde die ownCloud-Installation zunächst auf Basis von Bare-Metal Servern gestartet, die im Laufe der Zeit auf Basis von VMware virtualisiert wurden. Zur Zeit wird die Migration auf eine Plattform auf Basis von OpenStack, LXD und Docker vorbereitet. Wir versprechen uns hiervon eine noch bessere Hardwareausnutzung und wollen eine dynamische Bereitstellung von Instanzen implementieren. Die Isolation der Dienste wird dabei auch im Sinne der Datensicherheit vorangetrieben. Bei der Migration wird die Tatsache ausgenutzt, dass in den Übergängen gemischte Installationen möglich sind.

Mittlerweile konnten wir an der TU Berlin Erfahrungen mit drei verschiedenen Load-Balancern vor den ownCloud-Frontends sammeln. Stabil blieb bislang das zugrundeliegende Clusterfilesystem und der Galera-Cluster im Backend. An diesen Stellen werden jedoch Grenzen der Skalierbarkeit sichtbar. Aus diesem Grund experimentieren wir mit einem alternativem Ansatz, den Hr. Karlitschek auf der CS3 in Amsterdam vorgestellte und an dessen Konzeption wir mitwirken.

1 Einleitung

Das IT-Dienstleistungszentrum der TU Berlin tubIT bietet seit mittlerweile mehr als fünf Jahren ein Sync-n-Share Dienst (tubCloud) für seine Mitglieder, sowie für über 15 weitere Einrichtungen (DFN-Cloud) an. Die tubCloud enthält aktuell über hundert Millionen Dateien.

Im Laufe der Zeit waren verschiedene Anpassungen der Infrastruktur nötig oder wurden erprobt, um die Performanz, Verfügbarkeit, Wartbarkeit oder Sicherheit des Dienstes zu verbessern. Dabei bleiben einige „Building-Blocks“ immer erhalten. Andere Elemente mussten oder sollten im Laufe der Zeit ausgetauscht werden.

Zunächst werden wir betrachten, wie die Entwicklung im Datacenter-Bereich im Allgemeinen verläuft. Dann werden wir betrachten, welche Bausteine es zur Erbringung eines typischen Cloud-Speicher Dienstes auf Basis von own-Cloud/Nextcloud gibt, um dann einige Evolutionsschritte in der tubCloud-Architektur zu betrachten.

Der Artikel wird von einem Ausblick abgeschlossen, der sowohl betrachtet, wie eine mögliche Weiterentwicklung der Infrastruktur aussieht, wie auch die Ausrichtung des Dienstes insgesamt.

1.1 Evolution im Datacenter allgemein

Vor einigen Jahren erreichte das tubIT Datacenter eine ca. 90 %ige Virtualisierung der Dienste. Die zuvor betriebenen „Bare-Metal“ Server wurden nach und nach durch VMware virtualisierte Maschinen ersetzt. Nur einige wenige Maschinen, wie TSM-Backup-Server, Logserver für die Virtualisierer oder GSS-Cluster (GPFS-Fileserver) blieben auf reiner Hardware.

Die Vor- und Nachteile der Virtualisierung sind bekannt. In unserem Kontext waren vor allem die folgenden Vorteile entscheidend:

- Reduktion von Hardware im Rechenzentrum
    - Geringerer Strom- und Kühlbedarf
    - Verringerung von Hardware-Wartung
    - Vereinfachung von Verkabelung etc.
- Verschieben / Evakuieren von Hardware
    - Erhöhung der Redundanzen
    - Minimierung der Ausfallzeiten
    - Ermöglichen von Wartung im laufenden Betrieb
- Kurzfristiges Erzeugen neuer virtueller Maschinen
    - Schnelle Reaktion auf Last, geänderte Anforderungen
    - Bereitstellung geeigneter Testinfrastruktur
    - Hohe Flexibilität in Projekten

Ein eindrucksvoller Moment in Bezug auf die virtualisierte Serverinfrastruktur war die Migration sämtlicher Solaris-basierter Produktionsserver auf Linux RedHat Enterprise an nur einem Wochenende. Die Maschinen wurden alle im Vorfeld vorbereitet und getestet und dann deaktiviert. Am Tag der Migration wurde ein Solaris-Server nach dem anderen außer Betrieb genommen und von seinem RedHat-Pendant ersetzt. Dabei mussten nur unwesentlich zusätzliche Hardware bereitgehalten werden. Im Falle einer Bare-Metal Migration hätten wir 100 % mehr Server oder Monate für die Umstellung benötigt.

Nicht zuletzt aus Gründen der Lizenzkosten aber auch, um die zur Verfügung stehende Hardware besser auszunutzen und weniger redundante Software-Schichten vorhalten zu müssen, geht der allgemeine Trend heute hin zu Container-basierten Infrastrukturen. Zwar bieten die aktuellen Container-Lösungen nicht alle Funktionen an, die mit den eingesetzten Virtualisierungslösungen möglich wären, für definierte Anwendungsfälle spielen diese fehlenden Funktionen jedoch keine große Rolle.

1.2 Elemente einer Cloud-Speicher Lösung

In unserem Beispiel betrachten wir die tubCloud, die auf den Softwareprodukten ownCloud bzw. Nextcloud basiert(e). Es handelt sich hierbei um eine klassische LAMP-Anwendung (Linux, Apache, MySQL, PHP). Solche Installationen können auf sehr kleinen Servern bereits ohne Probleme monolithisch betrieben werden. Für sehr kleine Installationen kann man sogar auf ein eigenes Datenbankmanagementsystem verzichten. Eine große Installation mit mehreren tausenden oder zehntausende Nutzern wird jedoch in der Regel in folgende Komponenten zerlegt:

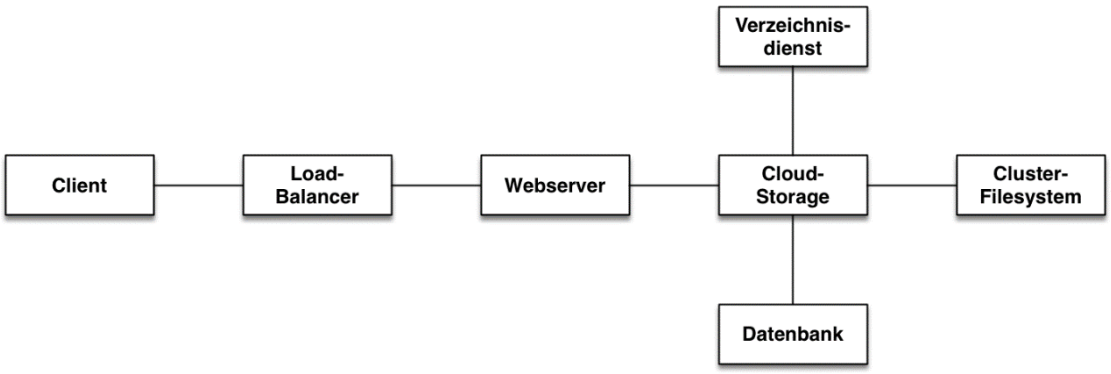

Abbildung 1: Elemente einer Cloud-Speicher Lösung

Loadbalancer: Eine zentrale, ausfallsicher aufgebaute Komponente, die die HTTP-Verbindungen zum Dienst entgegennimmt und auf die dahinter liegenden Web-Frontends verteilt. Dabei werden auch defekte oder gerade deaktivierte Frontend-Server nicht weiter bedient und die Verbindungen auf andere Server umgeleitet. Häufig übernimmt der Loadbalancer auch das SSL-Offloading, übernimmt also Client-seitig die SSL-Verbindungen und leitet den Verkehr als unverschlüsselten HTTP-Verkehr an die dahinterliegenden lokalen Server weiter.

Frontend-Server: Hier ein klassischer Webserver mit Apache oder NGINX und PHP. Er kann noch weitere lokale Komponenten enthalten, um z. B. Selbst eine Verteilung der Datenbankanfragen vornehmen zu können. Die Frontend-Server laufen alle unabhängig voneinander. Zur Verteilung der Last können theoretisch so lange weitere Frontend-Server hinzugefügt werden, bis diese die Anzahl der Zugriffe schnell genug abarbeiten können.

Datenbank-Cluster: In unserem Fall verwenden wir einen Galera-Cluster, der über mehrere Rechenzentren verteilt ist. Der Cluster dient sowohl der Ausfallsicherheit, wie auch der Lastenverteilung.

Cluster-Filesystem: Für den Dienst wird ein Dateisystem benötigt, das es erlaubt, Dateien von allen Frontend-Servern parallel zu Lesen und zu Schreiben - und das unter Einhaltung der Datenkonsistenz und wiederum ausfallsicher und lastverteilt.

2 Evolution der tubCloud / DFN-Cloud

Als wir den tubCloud Dienst für potentiell 30.000 Nutzer im Mai 2012 starteten, lagen uns leider noch keine Erfahrungswerte von anderen Universitäten vor. Wir kannten die Auswirkungen weder auf die Virtualisierungsinfrastruktur, noch auf das SAN oder das Netzwerk. Aus diesem Grund nutzten wir eine Reihe frei gewordener Server, um darauf direkte (Bare-Metal) die Frontend-Server zu installieren. GSS-Cluster für das Cluster-Filesystem und leistungsstarke Datenbankserver, auf denen die MySQL-Server ebenfalls nicht-virtualisiert laufen konnten wurden ebenfalls angeschafft.

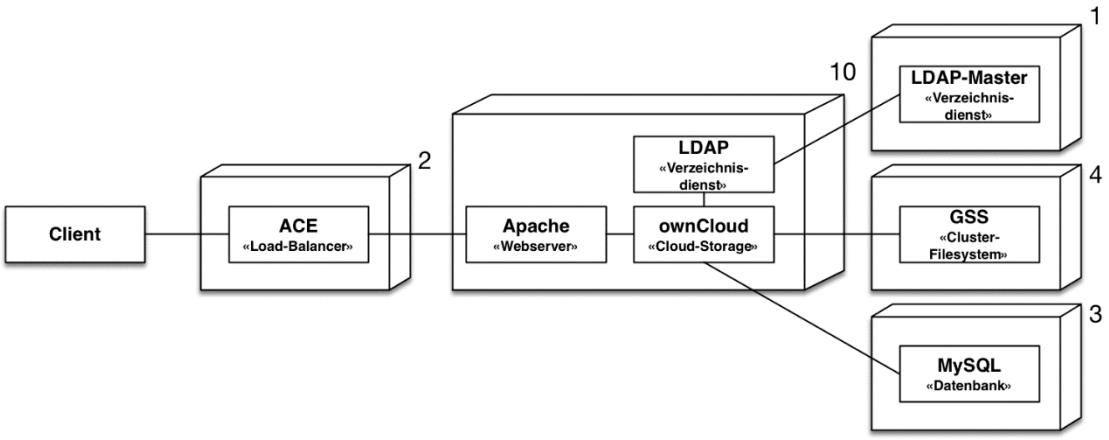

Abbildung 2: Ursprüngliche Architektur des tubCloud-Dienstes

Die DFN-Cloud Kunden wurden schließlich auf vHosts (eine Apache-Webserver Technik) in eigenen User-Spaces, d.h. mit eigenen Zugriffsrechten ausschließlich auf ihre Dateien und Datenbanken eingerichtet. Damit nutzten die DFN-Cloud Kunden dieselbe Infrastruktur, wie die TU Mitarbeiterinnen und Mitarbeiter.

Wie in [Hil17] beschrieben, mussten aus Gründen der Support-Kündigung die Cisco ACE-Loadbalancer zunächst durch HA-Proxy Server ergänzt und schließlich durch eine neue Lösung auf Basis von F5 Big IP abgelöst werden.

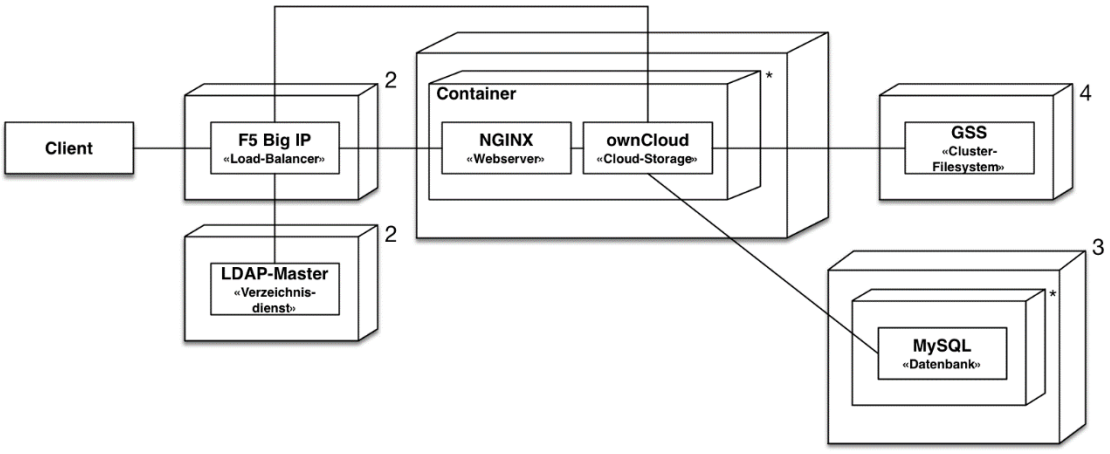

Abbildung 3: Aktuell angestrebte Architektur

Aktuell laufen sechs Nextcloud Frontend Server für die tubCloud unter NGINX und einige DFN-Cloud Kunden unter ownCloud oder Nextcloud auf eigenen virtuellen Servern.

Neben dem MySQL-Server für die tubCloud existiert ein eigener MySQL-Cluster für die DFN-Cloud Kunden. Beide Cluster sind bereits für den Einsatz mit Containern vorgesehen bzw. laufen bereits als Container. Dies sorgt für eine größere Unabhängigkeit der Dienste. Eine Überlast in der tubCloud-Installation hat keine Auswirkungen mehr auf die DFN-Cloud Kunden.

Um das Konzept konsequent zu Ende zu denken, wird jede DFN-Cloud-Instanz zukünftig *2..n* eigene Frontends in eigenen Open-Stack gesteuerten Containern bekommen. Dabei sollen die Frontends automatisiert gestartet werden, wenn die Last dies erfordert und sollen abgebaut werden, wenn diese nicht mehr benötigt werden. Container mit dysfunktionalen Frontends können leicht abgeschaltet und durch neue ersetzt werden. Dies geschieht in der Regel automatisch.

Größere DFN-Cloud Instanzen können ferner eigene MySQL-Cluster bis hin zu Clustern auf eigener Hardware bekommen. Eine Skalierung ist über die Technik problemlos möglich.

Ferner laufen bereits ProxySQL-Server in der Container-Infrastruktur, sowie Redis-Cluster für die Session-Verwaltung und das File-Locking.

Über Erfahrungen mit der Infrastruktur kann frühestens in einem Jahr berichtet werden, wenn alle Dienste umgestellt sind und auch z. B. Update-Zyklen sowohl auf Betriebssystem, wie auch auf Anwendungsebene erprobt wurden.

3 Blick in die Zukunft

Die Nutzerzahlen sind aktuell weiter steigend. Das betrifft sowohl die TU-eigenen Nutzerinnen und Nutzer wie auch die Auslastung der DFN-Cloud Kunden und die Zahl der DFN Partner selbst. Aber nicht nur die Quantität der Nutzung steigt. Immer neue Anwendungsfälle werden im Rahmen der DFN-Cloud erprobt und einige Wünsche mussten in der Vergangenheit abgelehnt werden, weil z. B. die zugrundeliegende PHP-Version zu alt aber noch nicht jede Instanz in der Lage war mit einer neuen zu arbeiten etc.

Die Unabhängigkeit der einzelnen DFN-Cloud Instanzen auf Basis der Container wird es zukünftig ermöglichen auch spezielle Konfigurationen für einen DFN Partner bereitzustellen. Dies beinhaltet z. B. auch die Nutzung speziell geforderter Plugins (bei ownCloud / Nextcloud Apps genannt) auf dem Server.

3.1 Zukünftige Architektur

Während die Frontend-Server leicht in die Breite skaliert werden können, bleiben Datenbank-Cluster und Cluster-Filesystem aktuell als relativ große, monolithische Installation in der Architektur. Über diese Einschränkung sprach mit uns Frank Karlitschek, der auf der CS3 [NeCl17] eine neue Architektur vorstellte, die auch Dateien und Datenbanken in kleine Nextcloud-Einheiten verschiebt. Nutzerdaten sind in dieser Architektur über mehrere Server verteilt gespiegelt vorhanden. Es ist jedoch möglich jeweils nur einige 100 Nutzer auf eine kleine Einheit aus Webserver, Datenbank und lokalem Dateisystem zu provisionieren und mit neuen Nutzern, dann automatisch weitere Server zu erzeugen.

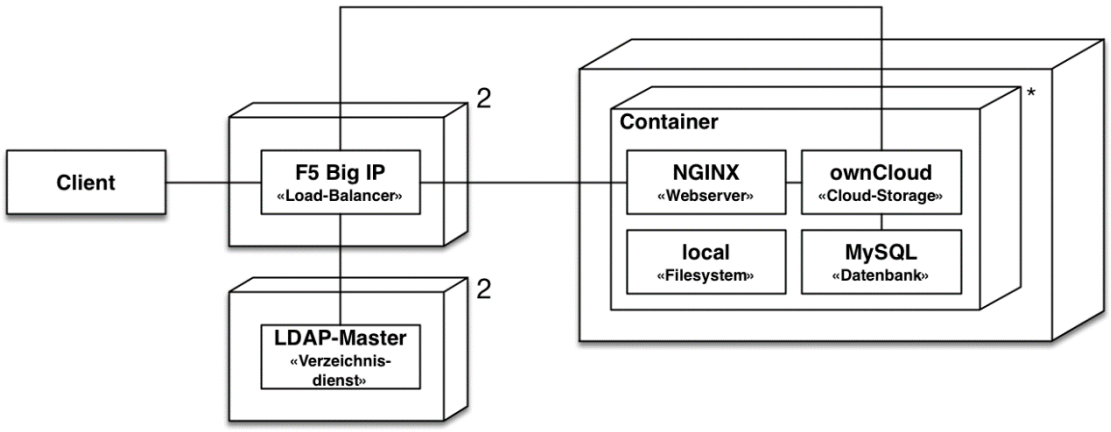

Abbildung 4: Mögliche Architektur basierend auf Nextcloud Global Scale

Zwar erhöht dieses Modell die Zahl von Diensten und fordert weitere Infrastrukturkomponenten zur Verwaltung. Auf der anderen Seite werden aber sowohl potentielle Probleme auf wenige Nutzer isoliert als auch Kosten für extrem leistungsfähige Hardware oder kostspielige Cluster-File-Lösungen reduziert. Auch werden Update-Prozesse so vereinfacht. Eine Downtime am Wochenende könnte mit dieser Architektur bald der Vergangenheit angehören, weil Nutzer jeweils einzeln und sequentiell auf aktualisierte Server verschoben werden könnten, ohne dass diese einen Betriebsausfall bemerken.

3.2 Zukünftige Ausrichtung des Dienstes

Neben der infrastrukturellen Weiterentwicklung, stellt sich ferner die Frage, wie der Dienst zukünftig ausgerichtet sein soll, welche Anwendungsfelder er beispielsweise neben dem klassischem Sync-n-Share bieten soll.

Aus Gesprächen mit unseren Nutzern kristallisieren sich drei Felder, die wir für die weitere Entwicklung beleuchten wollen:

1. Kooperations-Lösung: Neben dem Austausch oder dem gemeinsamen Arbeiten an Dateien ist auch die Kommunikation über diese Dateien bzw. die gemeinsame Arbeit daran immer wichtiger. Während aktuell Kommunikationswege und Sync-n-Share Dienst unabhängig voneinander laufen, ist eine sinnvolle Weiterentwicklung hier die Integration von Text-, Audio- und Video-Chat Kanälen, die in die Lösung integriert ist. Die Integration ist vor allem für Nicht-Techniker interessant, da so die Hürden für die Nutzung eines solchen Dienstes deutlich herabgesetzt werden.

2. Office-Funktionalität: Sehr häufig werden Office (Texte, Präsentationen, Tabellen) gemeinsam bearbeitet. Die Versionsverwaltung in der Cloud ist dabei ein Schritt um zu verhindern, dass verschiedene Autoren die Versionen von anderen unwiederbringlich überschreiben. Interessanter wäre jedoch, wenn ein gemeinsames Arbeiten an den Dokumenten (möglichst online und im Browser) möglich wäre. Für diese Funktionalität wird heute noch häufig Google Docs trotz aller rechtlichen Fragestellungen in Bezug auf die Nutzung dieses Dienstes verwendet.

3. Spezielle Dienste für Erprobungspartner: Einige Einrichtungen wünschen sich auf ihre Anwendungsfelder zugeschnittene Plugins oder Anwendungen, die sie gemeinsam mit ihrem Sync-n-Share Dienst nutzen wollen. Die beschriebene Container-Technik bietet die Möglichkeit solche Konfigurationen zu ermöglichen.

4 Literatur

[Hil17] T. Hildmann, Aus dem Leben eines DFN-Cloud Providers, Clouddienste im Hochschuleinsatz 2016/2017, S. 25

[NeCl17] Nextcloud GmbH, Nextcloud Global Scale – Architecture Whitepaper, https://nextcloud.com/globalscale/, 2017

Fragen der Community zur Trennung zwischen ownCloud und Nextcloud

Michael Röder
DFN-Verein
Berliner Geschäftsstelle
Alexanderplatz 1
10178 Berlin
roeder@dfn.de

Abstract: Für die Erbringung der Sync-&-Share-Dienste in der DFN-Cloud kommen neben diversen Eigenentwicklungen auch zahlreiche Module kommerzieller Softwareanbieter zum Einsatz. Im Jahr 2016 deuteten sich Veränderungen im Hause eines der Softwareanbieter an. Diese wurden auch unter den Anwendern innerhalb der DFN-Cloud diskutiert und während einer Publikumsdiskussion gemeinsam erörtert. Dieser Beitrag erläutert den Zusammenhang aus Sicht der DFN-Cloud-Anwender und zitiert die Diskussionsbeiträge.

1 Anforderungen von Wissenschaft und Forschung an die Cloud

Die Bedarfe von Wissenschaft und Forschung generieren unterschiedliche Anforderungen an Cloud-Dienste. Unter anderem die Sensibilität der zu verarbeitenden Daten fächert sich auf in ein weites Spektrum über alle Forschungsdisziplinen und darin definierter Sicherheitsklassen. Aus diesem Grund haben sich Einrichtungen im X-WiN[1] dazu entschieden, Cloud-Dienste speziell entlang der Bedürfnisse ihres Wirkungskreises zu entwickeln. Um Synergieeffekte zu nutzen und von der Skalierbarkeit sowohl bei der Erbringung als auch bei der Anwendung von Cloud-Diensten zu profitieren, haben sich einige Einrichtungen darüber hinaus dazu entschlossen, die eigene Cloud-Infrastruktur für andere Anwender aus Wissenschaft und Forschung zu öffnen. Dabei entstehen Cloud-Services aus der Community für die Community: die sogenannten föderierten Dienste. Einrichtungen, die einen oder mehrere föderierte Dienste anbieten, werden in diesem Kontext als Forschungspartner bezeichnet. Wenden Einrichtungen föderierte Dienste an und tragen damit aktiv zu deren Weiterentwicklung bei, handelt es sich dabei um sogenannte Erprobungspartner.

1.1 Abgrenzungsmerkmale föderierter Dienste

In allen Rechenzentren, in denen föderierte Dienste erbracht werden, sorgen diverse technische und organisatorische Maßnahmen für die Einhaltung hoher Sicherheitsstandards. Dennoch gibt es unter den Erprobungspartnern bereits Kriterien bezüglich der datenverarbeitenden Standorte, die bei der Auswahl des Forschungspartners von Bedeutung sind. Entscheidend können dabei neben technischen Aspekten wie Klimatisierung, Architektur des Rechenzentrums und Redundanzkonzepten beispielsweise auch organisatorische Vorsorgemaßnahmen für den Havariefall oder Auditierungsmechanismen sein.

[1] Wissenschaftsnetz

Neben den physischen bzw. infrastrukturellen Eigenschaften grenzen sich die föderierten Dienste auch durch den Einsatz unterschiedlicher Client- und Serversoftware voneinander ab. Dadurch ergibt sich eine Bandbreite unterschiedlicher Features, die bei einem Forschungspartner selbstverständlich sind, während ein anderer Forschungspartner sich bewusst dagegen entschieden hat. Beispielsweise ist eine Weboberfläche mit dem Ziel, betriebssystem- und ortsunabhängig Zugriff auf den Service und die hochgeladenen Daten zu erhalten, für manche Anwender eine Funktion, die unter keinen Umständen im Anforderungskatalog eines Sync-&-Share-Dienstes fehlen darf. Andere Einrichtungen sehen gerade darin ein potenzielles Risiko, denn sie möchten alle abgelegten Daten aufgrund der hohen Sensibilität garantiert Ende-zu-Ende-verschlüsselt ablegen. Der Zugriff über den Webbrowser suggeriert jedoch, dass mindestens der Webserver auf die unverschlüsselten Daten zugreifen kann.

Folglich finden sich mit Nextcloud[2], ownCloud[3], PowerFolder[4] und TeamDrive[5] derzeit vier unterschiedliche Softwareanbieter in der DFN-Cloud wieder, die bei den Forschungspartnern als Bestandteil der Sync-&-Share-Dienste zum Einsatz kommen.

Insbesondere diejenigen föderierten Dienste, die mithilfe des Produktes der ownCloud GmbH umgesetzt wurden, bewegten sich im Jahr 2016 in einem besonderen Spannungsfeld. Diverse Gerüchte rund um ownCloud sorgten für Verunsicherung bei den Erprobungspartnern. Sie machten sich Sorgen um die Qualität des Produktes, dessen Support und auch bezüglich der Weiterentwicklung falls sich die Gerüchte bestätigen sollten.

2 Nextcloud und ownCloud gehen getrennte Wege

Als der Mitbegründer der ownCloud GmbH, Frank Karlitschek, am 27. April 2016 öffentlich bekannt gab, dass er die Firma ownCloud verlassen werde[6], sorgte das für diverse Fragen seitens der Erprobungspartner.

Bereits am 2. Juni 2016 folgte die Meldung durch Herrn Karlitschek, dass die Firma Nextcloud GmbH gegründet wurde und dass sich diese im selben Marktsegment platzieren wird[7] wie ownCloud.

2.1 Redebedarf auf der 65. DFN-Betriebstagung in Berlin

Unmittelbar danach entbrannte eine Debatte in der Open-Source-Community und auch die Anwender der DFN-Cloud zeigten sich zunehmend beunruhigt. Schnell wurde deutlich, dass weder der DFN-Verein noch die Forschungspartner in der Position waren, um die Spekulationen zu entkräften oder den Befürchtungen der Anwender aussagefähige Argumente entgegenzusetzen. Deshalb wurde nach einer Möglichkeit gesucht, bei der sowohl Erprobungs- als auch Forschungspartnern und selbstverständlich den Firmen Nextcloud GmbH und ownCloud GmbH auf fachlicher Ebene und unabhängigem Terrain miteinander ins Gespräch kommen können. Ausgewählt wurde dafür die 65. DFN-Betriebstagung, die im September des Jahres in Berlin stattfand und regelmäßig durch über 250 Teilnehmer besucht wird. Die Betriebstagung setzt sich zusammen aus einer Plenumsveranstaltung und mehreren, teilweise gleichzeitig stattfinden, fachlich speziell ausgerichteten Foren. Das Forum Cloud-Dienste wurde ausgewählt, um ein Publikum mit

[2] ownCloud GmbH (https://owncloud.com/)
[3] Nextcloud GmbH (https://nextcloud.com/)
[4] dal33t GmbH (https://www.powerfolder.com/)
[5] TeamDrive Systems GmbH (https://www.teamdrive.com/)
[6] Literaturverweis [FK_16-27-04]
[7] Literaturverweis [FK_16-02-06]

fachlichem Hintergrund zu adressieren, damit sich dieses direkt an der Diskussion beteiligen konnte.

Im Vorfeld der Betriebstagung haben sich die Veranstalter gemeinsam mit der Community auf einen Fragenkatalog geeinigt. Dieser Fragenkatalog ist Vertretern beider betroffener Firmen ausgehändigt worden. Die Fragen der Community zur Trennung zwischen ownCloud und Nextcloud wurden, neben anderen Fachvorträgen, im Rahmen einer Publikumsdiskussion im Forum Cloud-Dienste am 29. September zwischen 14:00 und 18:00 Uhr ausführlich besprochen.

Im Anschluss an die Veranstaltung ist das Gesprächsprotokoll mit der Bitte um Freigabe für die Veröffentlichung ebenfalls beiden Firmen ausgehändigt worden. Dieses Protokoll ist öffentlich zugänglich[8]. Der vorliegende Beitrag zitiert auf den folgenden Seiten die zurück erhaltenen Fragen und Antworten in unbehandelter Form als Ergebnis der Diskussion.

3 Publikumsdiskussion (mit Firmenvertretern) im Forum Cloud-Dienste während der 65. DFN-Betriebstagung

Am Gespräch beteiligten sich sowohl die NextCloud GmbH als auch die ownCloud GmbH. Die Firma Nextcloud war vor Ort vertreten durch Andreas Rode (Sales Manager) und Arthur Schiwon (Senior Software Engineer). Die ownCloud GmbH nahm durch Holger Dyroff (Geschäftsführung) schriftlich Stellung zum Fragenkatalog. Die Stellungnahme wurde durch den Moderator verlesen und parallel dazu auf die Leinwand projiziert.

3.1 Frage: „Wer ist Eigentümer des Terena-Vertrages im Sinne der Rechtsnachfolge?"

- Antwort (ownCloud):
  „Die ownCloud GmbH.
  Alle europäischen Verträge wurden immer mit der ownCloud GmbH geschlossen und dort gibt es auch in der Geschäftsführung keinen Wechsel, außer dass Herr Karlitschek ausgeschieden ist."

  Nachtrag 28. September 2016:
  „Herr Gerlinger ist jetzt neu in die Geschäftsführung der ownCloud GmbH eingetreten.
  Auf Seiten Terena ist Geant in den Vertrag eingetreten."

3.2 Frage: „Wird es Änderungen im Support für die Anwender geben aufgrund des Weggangs von Personal? Wenn ja, welche?"

- Antwort (ownCloud):
  „Nein. Unser Nürnberger Support Team ist komplett und voll im Einsatz. Kräfte zum debugging stehen weiterhin zur Verfügung.
  Im Engineering Team gibt es Veränderungen. Neuer Leiter ist Klaas Freitag (bisher Teamleiter Desktop Entwicklung), neuer Chief Architect ist Herr Thomas Müller. Herr Müller wird sich insbesondere um die architektonische Koordinieung von Features über das Gesamtprodukt hinweg kümmern."

[8] Literaturverweis [DFN_16-09-29]

3.3 Frage: „Was ist Ziel bzw. Hintergrund der neu gegründeten Foundation? Welche Rolle wird diese ggf. auch in Hinblick auf uns als Hochschulkunden spielen?"

- Antwort (ownCloud):
  „Herr Rex hat auf der ownCloud Conference als Update gegeben:
  https://s3.owncloud.com/owncloud/index.php/s/zSMBPWvfuNeGFrM
  Die ownCloud Foundation ist die strategische Entkoppelung des Open-Source Projektes von individuellen Firmen oder Personen. Dazu wird ein Board gebildet welcher die Gesamtheit des Ökosystems ownCloud repräsentiert. Die Firma ownCloud GmbH wird in diesem Aufsichtsrat mit nur einem Sitz repräsentiert, ohne jegliche besondere Rechte.
  Für Hochschulkunden ist dies strategisch zu Begrüßen da ein relvanter Anteil von Aufsichtsratsmandaten an Vertreter aus dem Forschungs- und Hochschulumfeld vergeben wird, unter anderem auch an bekannte deutsche Vertreter."

3.4 Frage: „Wie unterscheidet sich aus Ihrer Sicht Support, Roadmap bzw. Produktausrichtung der Nextcloud gegenüber der ownCloud?"

- Antwort (Nextcloud):
  „Nextcloud verfolgt eine 100 % Open Source Philosophie, wie etwa Red Hat.
  Wir haben keine proprietären Komponenten.
  Zu Beginn haben wir uns darauf fokussiert, die proprietären Enterprise Funktionen von ownCloud als Open Source zu reimplementieren. Dabei konnten wir auch Schwächen, die durch diese künstliche Kapselung entstanden sind, auszumerzen. Beispielsweise arbeitet unsere SAML App vollständig auf PHPEbene, anstelle der zuhilfenahme des Apache Moduls. Somit sind wir wesentlich flexibler und auch Einschränkungen wie ein erzwungener Logout nach 15 Minuten sind ausgemerzt.
  Uns geht es darum, dass Produkt weiter zu entwickeln. Wir haben bereits Funktionen zu Nextcloud hinzugefügt, die uns differenzieren. Insbesondere auf Security und Hardening Funktionen liegt der Fokus, beispielsweise die AntiBrute Force Protection oder SameSiteCookies sind Nextcloud Erweiterungen. Dies ist nun auch der Schwerpunkt für künftige Nextcloud Versionen. Des weiteren findet die Weiterentwicklung von Nextcloud vollständig offen auf Github statt. Die Ziele für künftige Nextcloud Versionen sind dort ebenfalls festgehalten."

- Antwort (ownCloud):
  „Die Kundenbasis von ownCloud ist geprägt von großen Installationen mit hunderten bis mehreren zehntausend aktiven Nutzern.
  Daher fokussiert sich die Entwicklung und Roadmap von ownCloud auf den zuverlässigen und skalierbaren Betrieb von Deployments entsprechender Größe.
  Weder ist von ownCloud aktuell ein Device für den Heimanwender geplant noch liegt auf diesem Bereich der Fokus. Bei Produktivitätserweiterungen wie Collabora oder der Voll Text Suche geht es um skaliebare Angebote für unsere Kunden mit Funktionalitäten für die Wissenschaft, Regierungen und regulierte Industrien.
  Die Roadmap Präsentation von der Nextcloud Konferenz scheint bisher nicht öffentlich als Video oder Slides verfügbar. Die wesentlichen Weiterentwicklungen die wir sehen sind Kopien von ownCloud Funktionalitäten. Falls sinnvolle Verbesserungen erfolgen, werden wir diese auch für ownCloud Kunden in kurzen Zeiträumen zur Verfügung stellen.
  Der ownCloud Support betreut 350+ Kunden weltweit. Als nächstes großes Projekt wird gerade der Rollout der Niedersachsen Cloud beim GWDG begleitet.

Alle ownCloud Kunden erhalten weiterhin kompletten und transparenten Zugang zur ownCloud Entwicklung, inklusive aller Enterprise Module.
Alle Teile von ownCloud: Server, Desktop und Mobile Apps sind im Quellcode auf github.com verfügbar und können durchgehend gebrandet werden.
Die Nextcloud iOS app steht bis heute nicht offen zur Verfügung."
Roadmap @ ownCloud Conference öffentl abrufbar:
https://www.youtube.com/watch?v=a0yB_aIIjxw

3.5 Frage (Moderator) an Nextcloud: „Gibt es eine iOS App von Nextcloud?"

- Antwort (Nextcloud):
  „Ja, es gibt eine Nextcloud App für iOS."

3.6 Frage: „Für die Wissenschaft-Community ist das Open Cloud-Mesh Projekt essentiell wichtig. Welche Erfolge oder nächste Schritte gibt es aus diesem Projekt zu berichten? Mit welchen Produkten wird OCM kurz/mittel/langfristig kompatibel sein?"

- Antwort (Nextcloud):
  „Open CloudMesh wird von Nextcloud, Pydio und ownCloud unterstützt. Der Erfolg manifestiert sich durch die Beliebtheit und Annahme dieser Funktionen in der Nutzerschaft der Universitäten.
  Die Open CloudMesh Spezifikation und Implementierung wurde ursprünglich von Frank Karlitschek, Björn Schießle und Lukas Reschke durchgeführt, die allesamt bei Nextcloud tätig sind.
  Gegenwärtig wird ein Projekt mit einem Dienstleister durchgeführt, der die Spezifikation in den OpenAPI Standard umschreibt. Dazu gab es kürzliche ein Kickoff Telefonat, bei dem von unserer Seite aus ebenfalls Frank Karlitschek, Björn Schießle und Lukas Reschke teilgenommen haben. Dort geht es auch darum, wie die OCM Spezifikation fortgeschrieben und verbessert werden kann. CERN, zum Beispiel, hatte Wünsche geäußert, aber auch der Dienstleiter hat einige Vorschläge mitgebracht.
  Es gibt aber auch Verbesserungen beim Federated Sharing, die Björn Schießle bei Nextcloud eingebracht hat.
  Dazu zählt das Auflösen von Sharingketten, welches in die OCM Spezifikation fließen kann und einen weiternen API Call benötigt. Er macht sich aber auch Gedanken, wie das Auffinden von Personen, wie beim vorheringen Vortrag thematisiert, besser funktionieren kann. Bereits jetzt können durch Personen durch bestimmte Einträge im Adressbuch im Dialog gefunden werden. Darüber hinaus tauschen Server, die als vertrauenswürdig eingestuft wurden, ihre Adressbücher aus. Zur Zeit arbeitet Björn an einem Lookup-Server an denen verfizierte Benutzer bekannt gemacht werden können, und langfristig soll dies auch dezentral möglich sein. Das sind ebenso Kandidaten für die OpenCloudMesh Spezifikation."

- Antwort (ownCloud):
  „OpenCloudMesh ist, soweit uns bekannt, von ownCloud, Pydio sowie Nextcloud in den aktuellen Versionen supported.
  Um anderen Herstellern die Unterstützung für OCM zu erleichtern wird aktuell das Protokoll durch einen Dienstleister formalisiert.
  Aktuelle Information sind immer unter
  https://wiki.geant.org/display/OCM/Open+Cloud+Mesh abrufbar."

  Nachtrag 26. Juni 2017:
  „ownCloud arbeitet natürlich aktiv an der Weiterentwicklung von OpenCloudMesh mit."

3.7 Nachfrage aus dem Publikum (Christian Sprajc, PowerFolder): „Wer von ownCloud hat an dem Kickoff Call teilgenommen?"

- Antwort (Nextcloud):
  „Niemand."

3.8 Frage: „Die ownCloud Enterprise Edition unterschied sich von der ownCloud Community Edition u. a. durch spezielle Enterprise Apps. Wird es eine EE auch für Nextcloud geben und worin besteht hier der Unterschied?"

- Antwort (Nextcloud):
  „Es gibt bei Nextcloud keine Unterscheidung zwischen Enterprise und Community Edition."

4 Literaturverzeichnis

[FK_16-27-04] Karlitschek, F.; Blog Post: „big changes: I am leaving ownCloud, Inc. Today", 27. April 2016; http://karlitschek.de/2016/04/big-changes-i-am-leaving-owncloud-inc-today/

[FK_16-06-02] Karlitschek, F.; Blog Post: „Nextcloud", 2. Juni 2016; http://karlitschek.de/2016/06/nextcloud/

[DFN_16-09-29] DFN-Verein; „Fragen der Community zur Trennung zwischen ownCloud und Nextcloud", 29. September 2016; https://www.dfn.de/fileadmin/3Beratung/Betriebstagungen/bt65/BT65_Cloud_Publikumsdiskussion_OC-NC.pdf

Universitätsverlag der TU Berlin

Clouddienste im Hochschuleinsatz 2016/2017

Proceedings der Veranstaltungen „Forum Clouddienste“ im Rahmen der DFN-Betriebstagungen am 29. September 2016 und 22. März 2017

Diese Veröffentlichung ist eine Fortsetzung der Reihe „Cloudspeicher im Hochschuleinsatz“. Die gleichnamige Tagung wurde mit dem DFN-Forum Clouddienste zusammengelegt. Das Themenspektrum wurde über das Hauptthema Cloudspeicher hinaus um rechtliche, vertragliche und weitere Clouddienste erweitert. Dieser Tagungsband fasst die beiden Veranstaltungen im Rahmen der DFN-Betriebstagungen am 29. September 2016 und 22. März 2017 zusammen.

ISBN 978-3-7983-2928-7 (print)
ISBN 978-3-7983-2929-4 (online)

ISBN 978-3-7983-2928-7

http://verlag.tu-berlin.de