
Johannes Mayer

studierte nach seinem Abitur 2013 am Romain-Rolland-Gymnasium Dresden zunächst Mechatronik und Mathematik (ohne Abschluss) an der TU Dresden. 2017 begann er bei seinem späteren Praxispartner deecoob GmbH zunächst als Praktikant in der Webentwicklung. Ab dem Wintersemester 2018 studierte er dual Wirtschaftsinformatik an der Berufsakademie Sachsen, Staatliche Studienakademie Dresden und schrieb 2021 seine Bachelorarbeit zum Thema: „Fallstudie zur Analyse und Bewertung von Serverless Computing im Kontext der Migration eines Elasticsearch-Clusters“.

Kontakt: jojo.mayer@gmx.net

Fallstudie zur Analyse und Bewertung von Serverless Computing im Kontext der Migration eines Elasticsearch-Clusters

Johannes Mayer

In dieser Fallstudie wurde auf Basis eines realen Systems ein Vergleich zwischen On-Premises-Betrieb und dem Betrieb in der Cloud durchgeführt. Dazu wurden entscheidungsrelevante Kriterien definiert – die Skalierbarkeit, die Sicherheit, die Performanz und die Kosten – und näher untersucht.

Aus der Analyse der gesammelten Daten konnte abgeleitet werden, dass die maximale Performanz in der Cloud-Umgebung deutlich besser ist. Es wurde versucht, Gründe für die Unterschiede zwischen den beiden Umgebungen zu finden. Bzgl. der Sicherheit und der Skalierbarkeit hat sich gezeigt, dass die Cloud-Umgebung deutliche Vorteile gegenüber einem On-Premises-Betrieb mit sich bringt. Um eine ähnliche Qualität hinsichtlich Verfügbarkeit, Ausfallsicherheit, Verschlüsselung, und Skalierbarkeit in einem On-Premises-Betrieb zu erreichen, sind die Kosten deutlich höher. Aus den Erkenntnissen wurden weitere Schritte bzw. Handlungsmöglichkeiten abgeleitet.

Based on a real system, this case study compared on-premises operation with operation in the cloud. This involved the definition of decision-relevant criteria – scalability, security, performance and costs – and their detailed examination.

The analysis of the gathered data revealed that the maximum performance is significantly better in the cloud environment. An attempt was made to identify reasons for the differences between the two environments. In terms of security and scalability, it has been shown that the cloud environment has clear advantages over on-premises operation. In order to achieve a similar quality in terms of availability, failover, encryption, and scalability in an on-premises operation, the costs are significantly higher. Further steps and possible courses of action were derived from the findings.

1. Einleitung

Der Grund für die Durchführung dieser Fallstudie ist die Idee, ein On-Premises Elasticsearch-Cluster in die Cloud zu migrieren. Im Rahmen der Fallstudie wird eine mögliche Umsetzung in der Cloud erzeugt und mit dem On-Premises-Betrieb verglichen. Als Grundlage für die Untersuchung bzw. Bewertung der beiden Betriebsumgebungen – Cloud und On-Premises – wird ein Kriterienkatalog auf Basis einer Literaturrecherche erzeugt.

Als Ziele dieser Fallstudie ergeben sich entsprechend:

1. das Erstellen eines Kriterienkataloges, mit dem der Betrieb eines Systems in der Cloud und On-Premises verglichen werden können
2. der Vergleich des Systems beim Betrieb in der Cloud und On-Premises anhand des Kriterienkatalogs
3. die Ableitung von Erkenntnissen und Empfehlungen hinsichtlich einer Migration in die Cloud

Die Grundlage für die Untersuchungsobjekte – auch Beispielsysteme genannt – ist ein Teil eines realen Systems. Das gesamte reale System dient dem Zweck, Daten zu Veranstaltungen, bei denen Musik gespielt wird, mittels Webcrawling zu erzeugen, diese dann zu bewerten, anzureichern und sie Verwertungsgesellschaften zur Verfügung zu stellen. Die angereicherten Daten können durch eine API oder eine Web-App, die Anfragen an Elasticsearch abstrahieren, durchsucht werden.

Der Teil des Gesamtsystems, der aus Web-App und Elasticsearch besteht, wird mit einem System verglichen, das in der Cloud betrieben wird und die gleichen Aufgaben absolvieren kann. Als Cloud Provider für diese Untersuchung wurde sich auf die Amazon Web Services (AWS) festgelegt.

2. Kriterienkatalog

Entscheidungen von Unternehmen basieren immer auf Informationen – so auch bei einer Entscheidung in die Cloud zu migrieren. Die Migration in Cloudumgebungen wie AWS versprechen unter anderem:

- höhere Flexibilität hinsichtlich der Skalierung
- garantierte Servicequalität durch mehrere Standorte und bessere Katastrophewiederherstellungsprozesse
- Kosteneffizienz durch verbrauchsgesteuerte Bezahlmodelle

Ob diese technischen und betrieblichen Versprechen erfüllt werden, ist ausschlaggebend für die Evaluierung, ob die Migration eines Systems in die Cloud tatsächlich vorteilhaft ist.¹ Im Rahmen dieser Fallstudie sind diese Versprechen den Begriffen Skalierbarkeit, Sicherheit und Kosten untergeordnet.

Zusätzlich zu diesen Aspekten, ist bei jeder Migration entscheidend, dass eine funktionale Äquivalenz zwischen Altsystem und Zielsystem

gewährleistet wird. Diese funktionale Äquivalenz kann durch den Begriff der Performanz abgedeckt werden. Mit Performanz ist in diesem Zusammenhang gemeint, dass das Zielsystem die gleiche Leistung wie das Altsystem erbringen kann.²

Durch die Schwierigkeit des Erfassens der Eigenschaften der der Cloudumgebung zu Grunde liegenden Hardware³ bietet es sich an, die tatsächlich erbrachte Leistung der Systeme durch Benchmarking zu ermitteln. Dadurch kann auch sichergestellt werden, dass eine äquivalente Qualität, aus Sicht der späteren Benutzer des Systems, gewährleistet wird.⁴

Die hohe Flexibilität in der Skalierung wird oft als größter Vorteil von Cloudumgebungen genannt.⁵ Bei On-Premises-Systemen ist das Skalieren von Systemkomponenten immer durch die zur Verfügung stehende Hardware begrenzt, während in Cloudumgebungen die Illusion von unendlichen Kapazitäten dazu führt, dass Nutzer deutlich weniger Planung in die Bereitstellung von Hardware stecken müssen. Die Cloud bietet nicht nur die Möglichkeit, das System nach oben zu skalieren, sondern auch, sie wieder nach unten zu skalieren, wodurch eine hohe Performanz gewährleistet werden kann. Zusätzlich dazu bietet eine automatisierte und bedarfsgesteuerte Skalierung Kosteneinsparungen.

Die Isolation von Daten und Ressourcen bei cloudbasierten Diensten ist die Voraussetzung für deren Adaption. Und selbst bei ausreichender Isolation sind, durch das Teilen der zugrundeliegenden Hardware, Angriffszenarien denkbar.⁶ Eine weitere Sorge ist, dass Daten und betriebskritische Prozesse auf Hardware ausgeführt werden, die sich unter der Kontrolle einer anderen Organisation befindet.⁷ Eine On-Premises-Umgebung bietet den Vorteil, dass betriebskritische Prozesse auf Hardware, die nicht mit dem Internet verbunden ist, ausgeführt werden kann. Neben den Aspekten der Absicherung des Zugriffs auf Daten und Prozesse, sind die Verfügbarkeit, die Verteilung der Systeme und die Katastrophewiederherstellung von Bedeutung. In der Cloud können Systemkomponenten dank automatischer Backups im Katastrophenfall automatisch neu gestartet werden.⁸ Des Weiteren werden durch die Nutzung mehrerer Verfügbarkeitszonen für das Betreiben der Systeme eine hohe Verfügbarkeit⁹ und eine kundennähere Bereitstellung der Systeme ermöglicht, um Netzwerkkommunikationszeiten zu verkürzen.

2 vgl. Rai, R.; Mehruz, S.; Sahoo, G. (2013), 58ff; vgl. Ginnich, R.; Winter, A. (2005), 24.

3 vgl. Kousiouris, G. et al. (2014), 714f.

4 vgl. Kousiouris, G. et al. (2014), 716ff; vgl. Costa, P. J.; Cruz, A. M. (2012), 101.

5 vgl. Mukherjee, S. (2019), 4.

6 vgl. Mundada, Y.; Ramachandran, A.; Feamster, N. (2011), 1f; vgl. Ristenpart, T. et al. (2009), 199ff.

7 vgl. Garg, S. K.; Versteeg, S.; Buyya, R. (2013), 1013.

8 vgl. Paavolainen, S. (2013), 27.

9 vgl. Paavolainen, S. (2013), 27.

1 vgl. Mukherjee (2019), 1ff

Viele Entscheidungen im Unternehmen lassen sich auf eine Gegenüberstellung von Kosten und Nutzen zurückführen. Bei dem Betrieb von Systemen in der Cloud erwarten Unternehmen eine höhere Kosteneffizienz und -transparenz durch verbrauchsgesteuerte Modelle. Für einen Vergleich zwischen On-Premises und Cloud reicht es nicht nur, die reinen Betriebskosten zu vergleichen, sondern es müssen die Gesamtkosten des Betriebs – total cost of ownership (TCO) – verglichen werden.

3. Versuchsaufbau

Abbildung 1 zeigt den Aufbau des lokalen Systems, was auch die Basis für das Cloud System (siehe Abbildung 2) darstellt. Der Unterschied zwischen den Systemen ist lediglich, dass beim lokalen System ein Application Load Balancer verwendet wird, der wie ein Reverse-Proxy funktioniert. Dieser ist notwendig, um das lokale Elasticsearch über das Internet erreichbar zu machen. Für das Cloud System ist dies nicht notwendig, da der Amazon Elasticsearch Service schon so konfiguriert werden kann, dass dieser über das Internet erreichbar ist.

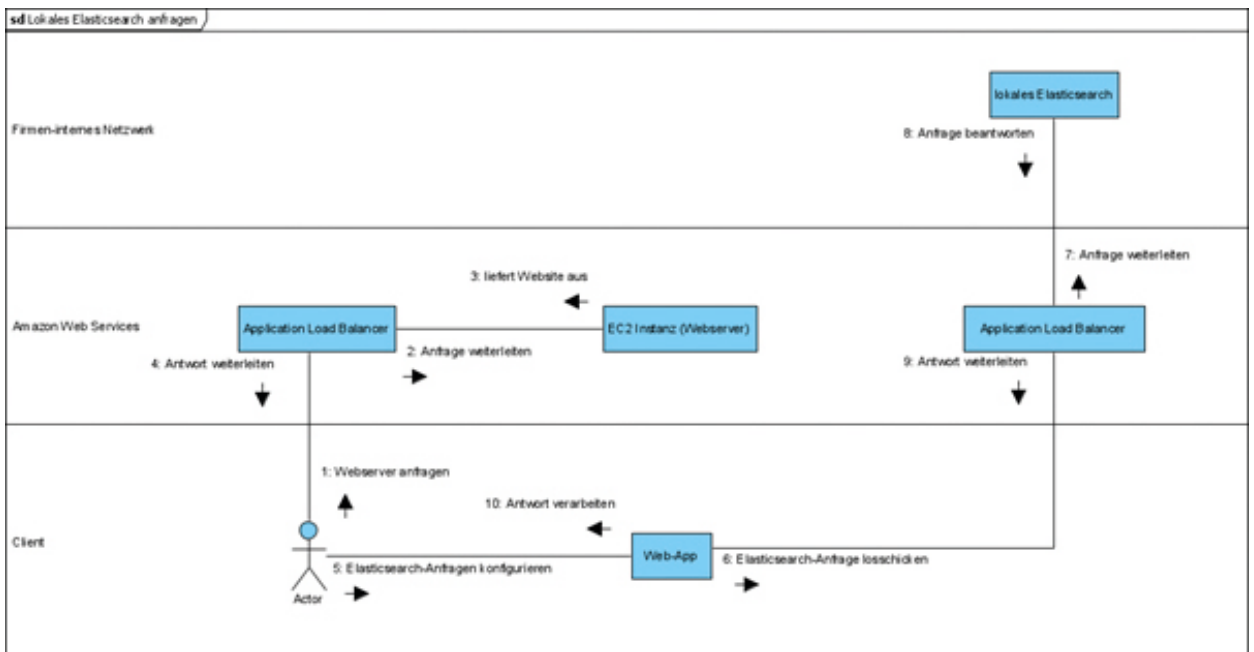


Abbildung 1: Kommunikationsdiagramm – lokales System

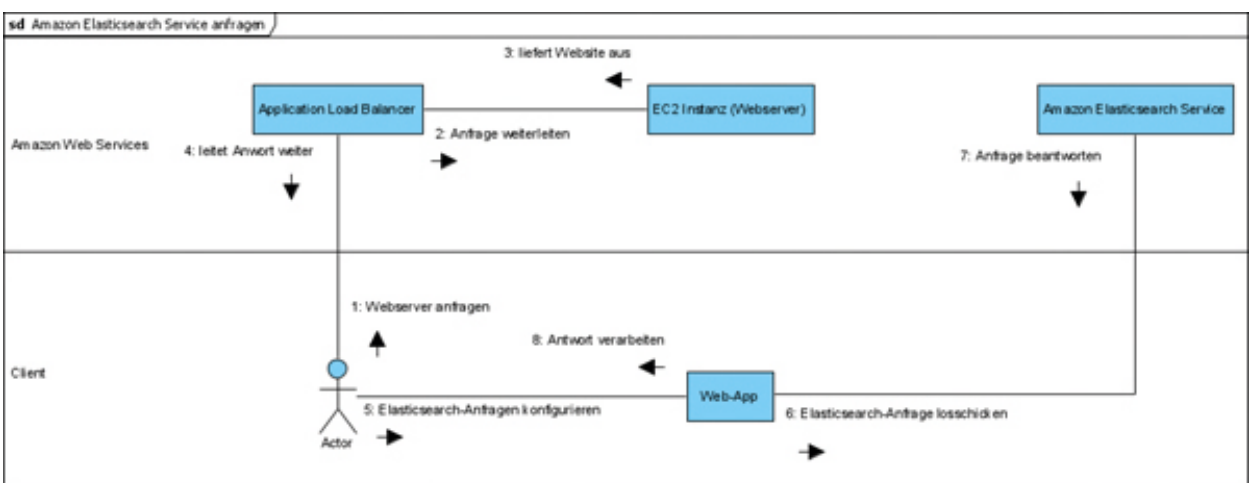


Abbildung 2: Kommunikationsdiagramm – Cloud System

4. Datenerhebung

Das hauptsächliche Messinstrument für die Performanz ist eine mit create-react-app¹⁰ selbstentwickelte Web-App. Die Schritte 1 bis 4 (siehe Abbildung 2) zeigen, wie die Web-App aufgerufen wird. Die Anfragen an das Elasticsearch werden mittels der Web-App konfiguriert, gesendet und die Antworten präsentiert. Jede einzelne Anfrage durchläuft demnach Schritte 5 bis 8 (siehe Abbildung 2).

Das wesentliche Merkmal der Web-App ist eine „Random Query Sender“-Komponente, die es erlaubt, eine bestimmte Anzahl von vordefinierten Anfragen zufallsbasiert und in einem bestimmten Intervall zu senden. Die vordefinierten Anfragen wurden aus dem Ursprungssystem ausgewählt und bilden typische Nutzerinteraktionen ab. Die Anfragen werden zufallsbasiert und mehr als einmal gesendet, um die Interaktionen ausgehend von mehreren Benutzern zu simulieren und natürliche Schwankungen in der Anfragebearbeitungsdauer zu kompensieren. Die Messung der Bearbeitungsdauer führt Elasticsearch selbst, defaultmäßig, durch.

Die Ausprägungen bezüglich Sicherheit und Skalierbarkeit wurden literaturgestützt und durch Experimente ermittelt. Die Kosten für den Betrieb des Cloud Systems wurden anhand der Preislisten der Amazon Web-Services ermittelt.¹¹ Für die Kosten des lokalen Systems wurde ein Experteninterview mit einem Mitarbeiter, der für die Systemadministration verantwortlich ist, durchgeführt.

5. Datenanalyse

Für die Analyse der Performance wurden aus den erhobenen Daten

index_query	AVERAGE of took	MIN of took	MAX of took
merged_events; 1000 results	110	65	295
merged_events; basic	4	3	8
merged_events; ids	37	14	276
merged_events; inverted sorted + filtered	43	22	163
merged_events; sorted	678	471	1568
merged_events; sorted + filter	39	24	231

Abbildung 3: Cloud System - 5 Anfragen pro Sekunde

ex_query	AVERAGE of took	MIN of took	MAX of took
ged_events; 1000 results	136	83	
ged_events; basic	6	2	
ged_events; ids	40	18	
ged_events; inverted sorted + filter	52	28	
ged_events; sorted	1037	617	
ged_events; sorted + filter	46	31	

Abbildung 4: Lokales System - 5 Anfragen pro Sekunde

Zusätzlich wurden dabei die Grenzen des Systems, ab dem es überlastet wird, ermittelt (siehe Tabelle 1).

System	Knoten				Grenzen des Systems
	Anzahl	vCPU	RAM (GiB)	Speicherkapazität	
Cloud 2	8	4	30,5	950 GB	5-10 Anfragen/s
Lokal	8	12	64	4 TB	>20 Anfragen/s

Tabelle 1: Grenzen der Systeme

An diesen Grenzen stauen sich die Abfragen und können nicht mehr bearbeitet werden, bis Kapazitäten frei werden. Die Anzahl der parallelen Anfragen lässt sich dabei auf die Anzahl der vCPU zurückführen und die Größe des RAM zusammen mit der Festplattengeschwindigkeit auf die Anfragebearbeitungsdauer.

Hinsichtlich Skalierbarkeit lässt sich horizontal bis 200 Datenknoten innerhalb von 30 Minuten skalieren.¹² Lokal dagegen dauert das Hinzufügen eines neuen Datenknotens mindestens 4 Stunden, wenn Serverkapazität überhaupt verfügbar ist.

Bezüglich der Sicherheit unterscheiden sich lokal und in der Cloud lediglich in der Verantwortlichkeit und Komplexität der Umsetzung – in der Cloud wird eine bestimmte Sicherheit durch Verschlüsselung vorausgesetzt und es gibt die Möglichkeit, mehrere Standorte zu verwenden, ohne Mehrkosten.

Berechnet man für die Kosten nur die Betriebskosten, ohne zu beachten, dass in der Cloud eine höhere Basissicherheit ermöglicht wird, dann sind die Kosten im lokalen Betrieb deutlich günstiger. Werden die Absicherung durch mehrere Standorte, eine differenzierte Zugriffskontrolle und die Vorteile durch bedarfsgerechte Skalierung beachtet, so stellt sich der Betrieb in der Cloud vorteilhaft dar (siehe Tabelle 2).

Umgebung	Reiner Betrieb (EUR)	Inkl. 3 Standorte (EUR)
Lokal	29.979	58.420
Cloud	55.823	55.823

Tabelle 2: Kosten nach Umgebung

¹⁰ vgl. Facebook (Hg.) (2021)

¹¹ vgl. Amazon Web Services (Hg.) (2021a); vgl. Amazon Web Services (Hg.) (2021b)

¹² vgl. Amazon Web Services (Hg.) (2021c).

6. Fazit

Anhand der Analyse wurde gezeigt, dass die Implementierung des Amazon Elasticsearch Service mindestens genauso leistungsfähig wie eine On-Premises-Implementierung von Elasticsearch sein kann. Über den Wartungsaufwand, das Ausfallrisiko und die Wiederherstellungszeit des Cloud-Systems nach Katastrophen konnten keine genauen Aussagen getroffen werden, allerdings wird ein großer Teil des Wartungsaufwandes durch eine Cloudmigration an den Provider abgegeben – mindestens 10 bis 20 Arbeitstage im Jahr. Automatische Backups, ein automatisches Neustarten von abgestürzten Knoten und die Verteilung des Clusters auf drei Verfügbarkeitszonen führen zu einer hohen Uptime des Systems. Anhand einer vereinfachten Gesamtkostenkalkulation wurde gezeigt, dass die Gesamtkosten des Betriebes im Cloud-Umfeld für das Beispielsystem deutlich niedriger sind als bei einem vergleichbaren On-Premises System.

Für eine Cloudmigration sprechen vor allem die Möglichkeit, das Bezahlschema frei zu wählen – vorgeladene Investition, Verpflichtung zur Nutzung für 1 / 3 Jahre oder nur auf Bedarf – und die Möglichkeit, das Cluster schnell zu skalieren, da diese beide große Flexibilität in der Planung mit sich bringen.¹³

Literatur

- Amazon Web Services (Hg.) (2021a): Amazon Elasticsearch Service – Preise, in: <https://aws.amazon.com/de/elasticsearch-service/pricing/> (abgerufen am 13.07.2021).
- Amazon Web Services (Hg.) (2021b): Elastic Load Balancing – Preise, in: <https://aws.amazon.com/de/elasticloadbalancing/pricing/> (abgerufen am 14.07.2021).
- Amazon Web Services (Hg.) (2021c): Häufig gestellte Fragen zu Amazon Elasticsearch Service, in: <https://aws.amazon.com/de/elasticsearch-service/faqs/> (abgerufen am 13.07.2021).
- Costa, P. J.; Cruz, A. M. (2012): Migration to Windows Azure – Analysis and Comparison, *Procedia Technology*, 5, 93-102, https://www.researchgate.net/publication/256546997_Migration_to_Windows_Azure_-_Analysis_and_Comparison (abgerufen am 11.07.2021).
- Facebook (Hg.) (2021): Create a New React App, Recommended Toolchains, in: <https://reactjs.org/docs/create-a-new-react-app.html#create-react-app> (abgerufen am 13.07.2021).
- Garg, S. K.; Versteeg, S.; Buyya, R. (2013): A framework for ranking of cloud computing services, in: *Future Generation Computer Systems*, 29(4), 1012-1023, <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.448.5450&rep=rep1&type=pdf> (abgerufen am 12.07.2021).
- Gimnich, R.; Winter, A. (2005): Workflows der Software-Migration, *Softwaretechnik-Trends*, 25(2), 22-24, <https://userpages.uni-koblenz.de/~ist/documents/Gimnich2005WDS.pdf> (abgerufen am 11.07.2021).
- Kousiouris, G. et al. (2014): A Multi-Cloud Framework for Measuring and Describing Performance Aspects of Cloud Services Across Different Application Types, in: *Proceedings of the 4th International Conference on Cloud Computing and Services Science - MultiCloud*, (CLOSER 2014), 714-721, <https://www.scitepress.org/Papers/2014/49754/49754.pdf> (abgerufen am 11.07.2021).
- Mukherjee, S. (2019): Benefits of AWS in modern Cloud, in: <https://arxiv.org/ftp/arxiv/papers/1903/1903.03219.pdf> (abgerufen am 11.07.2021).
- Mundada, Y.; Ramachandran, A.; Feamster, N. (2011): SilverLine: Data and Network Isolation for Cloud Services, in: *Proceedings of the 3rd USENIX conference on Hot topics in cloud computing (HotCloud'11)*, https://www.usenix.org/legacy/events/hotcloud11/tech/final_files/Mundada.pdf (abgerufen am 12.07.2021).
- Paavolainen, S. (2013): Analysis of the Availability of Amazon Web Services' Cloud Infrastructure Services, in: *Cloud-Based Software Engineering*, 26-33, https://helda.helsinki.fi/bitstream/handle/10138/42910/cbse13_proceedings.pdf?sequence=2#page=29 (abgerufen am 13.07.2021).
- Rai, R.; Mehruz, S.; Sahoo, G. (2013): Efficient migration of application to clouds: analysis and comparison, *GSTF Journal on Computing (JoC)*, 3(3), 58-62, <http://dl6.globalstf.org/index.php/joc/article/viewFile/507/2051> (abgerufen am 11.07.2021).
- Ristenpart, T. et al. (2009): Hey, you, get off of my cloud: exploring information leakage in third-party compute clouds, in: *Proceedings of the 16th ACM conference on Computer and communications security*, 199-212, http://ecee.colorado.edu/~ekeller/classes/fall2017_advsec/papers/cloudsec-ccs09.pdf (abgerufen am 12.07.2021).

¹³ vgl. Singh, B. et al. (2013), 387