

ALIBABA CLOUD

Alibaba und die 40 Services

Neben den großen drei Cloud-Diensten – AWS, Azure und GCP – gibt es auch noch die Alibaba Cloud, die weitaus mehr Potenzial hat, vorne mitzumischen, als manche denken.

Wer Entscheidungswege in Unternehmen kennt, kennt auch die Binsenweisheit, dass nur wenige Consultants und Manager dem Unternehmen Gartner würden widersprechen wollen. Eine alte Beratungsregel besagt deshalb: Ein Gartner-Diagramm, das die eigene Argumentation stützt, ist Goldes wert.

Bild 1 visualisiert, dass es nach Gartner im Cloud-Computing-Infrastruktur-Markt drei führende Anbieter gibt: AWS führt den Markt an, Azure hat deutlich aufgeholt. Im B2B-Bereich sehen viele Experten auch Microsoft Azure bereits führend. Die Google Cloud Platform (GCP) ist leicht abgeschlagen, hat sich aber als Hersteller von Kubernetes und als entwicklerorientierter Cloud-Anbieter einen Ruf erarbeitet.

Warum also Alibaba Cloud? Geht es nach Gartner, liegt Alibaba Cloud neben Oracle und IBM etwas zurück.

Das vermutlich wesentlichste Argument ist ein politischer Aspekt: Chinesische Firmen wachsen in Ländern schneller und stärker, in denen die USA unbeliebt sind. Viele amerikanische Cloud-Dienste sind beispielsweise im Iran gesperrt, und aufgrund von Sanktionen sind die Beziehungen zwischen amerikanischen und iranischen Firmen auf Eis gelegt. Das öffnet natürlich Türen für die Alibaba Cloud (vergleiche den Kasten Politik: Wer kontrolliert die Cloud?).

Auch gilt China mittlerweile als weltweit größter Investor in Ländern wie Afrika. Dort kann sich Alibaba ebenfalls leicht

ausbreiten, denn dort wo die Wirtschaft wächst, wächst auch die IT.

Lange galt, dass „alle guten Dinge aus den USA kommen“. Mit dem Kauf der Laptop-Sparte durch Lenovo in den Millenniumjahren wurde jedoch klar, dass viele Produkte, die man mit den USA assoziierte, künftig auch aus China kommen könnten.

Dass man sich auch in Europa nicht mehr nur allein nach den USA ausrichtet, zeigt die Geschichte eines der vielleicht innovativsten Projekte aus Deutschland für den Streaming-Markt: Apache Flink wurde als Kandidat gehandelt, Apache Spark abzulösen, da es mit einer nativen Streaming-Engine technologisch in einzelnen Bereichen voraus war. Doch letztlich wurde Data Artisans, das Unternehmen, das Apache Flink vorantrieb, nicht – wie viele vielleicht erwartet hätten – von einem US-amerikanischen Unternehmen, sondern von Alibaba gekauft [1].

Ziel dieses Artikels ist es, den augenscheinlichsten Unterschied – nämlich dass Alibaba ein chinesisches Produkt ist – mit technischen Inhalten zu ergänzen.

Dienste-Matrizen

Zufall oder nicht? Alibaba wirbt tatsächlich mit 40 Produkten (siehe Bild 2). Und wer diese 40 Produkte kennenlernen will, durchläuft wie bei jeder anderen Neuregistrierung bei einem

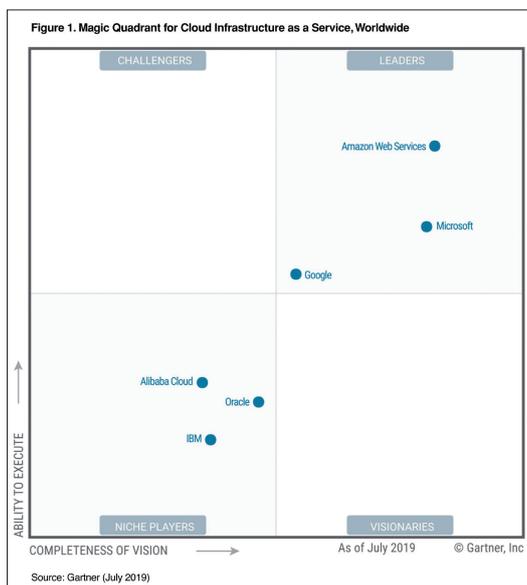
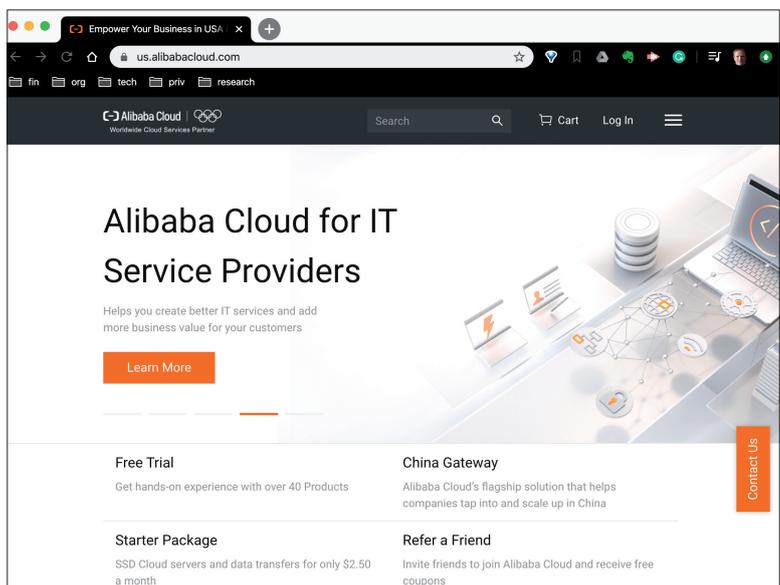


Bild: Gartner, Juli 2019

Magic Quadrant von Gartner (Bild 1)



Einstiegssseite von Alibaba (Bild 2)

● Politik: Wer kontrolliert die Cloud?

Beeinflusst Politik tatsächlich die Entscheidung für einen Cloud-Anbieter? Im Artikel wurde der Iran als Beispiel genannt, ein Land, in dem 80 Millionen Menschen leben. Die Banken, Telcos, Versicherungen und viele weitere Unternehmen haben Bedarfe an Softwarelösungen, und nicht jedes Unternehmen wird diese ignorieren wollen, nur weil die USA Sanktionen verhängt haben. Für ein Unternehmen, das also seine Dienste in diesen Ländern anbieten will, kann Alibaba Cloud eine interessante Alternative sein.

Spricht man mit Top-Entscheidungssträgern in einem Unternehmen, kommt oft das Argument auf, dass in der Frage der Cloud tatsächlich ein Wirtschaftskrieg zu spüren ist. Einem internationalen Konzern einen chinesischen Cloud-Anbieter statt eines US-amerikanischen anzubieten, kann Fragen aufwerfen.

Eines darf man auch nie vergessen: Wenn wir alles politisch sehen, dürfen wir auch nicht die Argumente der Cloud-Gegner übersehen, die bereits argumentierten, dass sie kein gutes Gefühl dabei haben, wenn Daten auf Servern abgelegt werden, die einem US-amerikanischen Unternehmen gehören. Mag sein, dass das Vertrauen in chinesische Unternehmen noch geringer ist, vor allem dann, wenn das Bewusstsein für den Schutz von personenbezogenen Daten infrage gestellt wird.

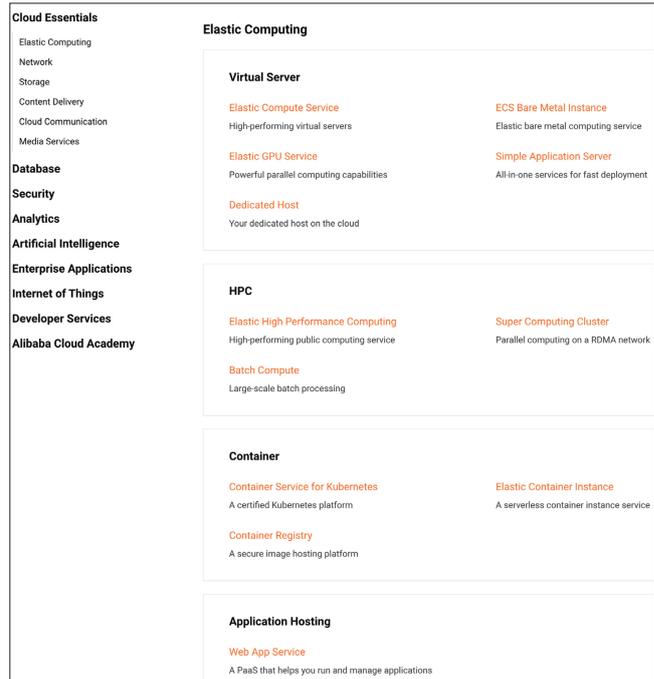
Cloud-Provider das identische Prozedere, samt Eingabe einer Kreditkarte.

Wer bereits mit einem Cloud-Provider gearbeitet hat, wird es leichter haben, auch andere Cloud-Provider zu verstehen. Ein populärer Dienst, der auf einer Plattform läuft, ähnelt Diensten auf anderen Plattformen.

Compute-Dienste sind das beste Beispiel. Mit EC2 hatte Amazon einen Dienst geschaffen, der Hardware virtualisiert und diese virtualisierte Hardware mit einem Betriebssystem bereitstellt. Ein neuer Anbieter muss nun Compute nicht neu erfinden, sondern er schafft eine eigene Virtualisierung, die der von AWS ähnelt. Und ob sich ein Entwickler nun mit SSH zu Amazon EC2 oder zu Alibaba ECS verbindet, ist keinen vergleichenden Screenshot wert. Jeder weiß, wie ein Linux-Terminal in einem Fenster aussieht.

Wie die Produkt-Seite zeigt (siehe Bild 3), weicht Alibaba in seiner Kategorisierung von zum Beispiel AWS ab [2]. Was bleibt, ist, dass jeder, der bereits eine Cloud kennt, irgendwo Dienste wiederfindet, die ihm vertraut sind. Oft ist es nur ein Mapping von Diensten eines Anbieters zu einem anderen.

Das bewirkt zugleich auch Ernüchterung, wenn man versucht, einen Entscheidungsbaum aufzubauen, in dem Kriterien festgelegt sind, anhand derer man sich für den einen oder den anderen Anbieter entscheidet. Die detaillierten technischen Unterschiede zwischen Diensten von Cloud-Anbietern sind für viele nicht relevant. ECS hat etwa mehr Instanztypen [3], EC2 hat mehr Datacenter. Mag zwar sein, dass der eine oder andere SysOps-Ingenieur bemängelt, dass es diverse Instanztypen eines Systems bei dem anderen An-

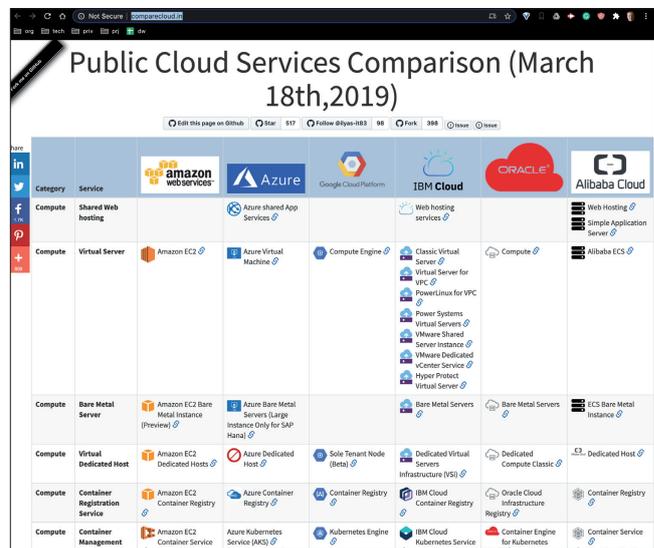


Produkte auf Alibaba im Überblick (Bild 3)

bieter nicht gibt, doch wiegen diese Gründe das Killerargument Preis nicht auf: Ist ECS billiger als EC2, so können selbst wenige Cent Unterschied zu hohen Beträgen hochskalieren.

Ein weiteres Argument, um Kunden zu erreichen, ist die Frage, in welchen Regionen Datacenter stehen. Meist wird argumentiert, dass US-amerikanische Lösungen eine bessere englische Dokumentation haben, weil natürlich einiges bei Alibaba Cloud aus dem Chinesischen übersetzt wurde. Offensichtlich wird das Sprachenproblem, sobald man im Internet nach spezifischen Alibaba-Produkten sucht: Viele Ergebnisse kommen in chinesischer Sprache zurück.

Wer den Matrizenvergleich sucht, findet auf der Seite comparecloud [4] (siehe Bild 4) einen exzellent aufbereiteten ▶



Cloud-Anbieter-Vergleich bei Comparecloud.in (Bild 4)

Database Services		
Database Type	Product	Scenario
Cloud Database Dedicated Cluster (CDDC) Autonomous and controllable database service cluster on Alibaba Cloud.	ApsaraDB Dedicated Cluster New	A highly cost-effective cloud database service on dedicated cluster with grants overprovisioning, hybrid deployment, resource scheduling and more authority.
	ApsaraDB for PolarDB	New generation of "Super MySQL, PostgreSQL or compatible Oracle" supporting up to 100 TB storage.
On-Line Transaction Processing (OLTP) Comprehensive relational transactional database products on Alibaba Cloud.	ApsaraDB RDS for MySQL	An availability and performance enhanced MySQL service supporting versions 5.5, 5.6, 5.7 and 8.0.
	ApsaraDB RDS for PostgreSQL	An on-demand PostgreSQL service with automated monitoring, backup and disaster recovery capabilities.
	ApsaraDB RDS for SQL Server	A cost effective cloud hosted SQL Server service with automated monitoring, backup and disaster recovery capabilities.
	ApsaraDB RDS for MariaDB TX	Cloud hosted MariaDB Enterprise Edition.
	Distributed Relational Database Service (DRDS)	A database sharding service on top of RDS and PolarDB to support extremely large transactional business data.
	ApsaraDB for Redis	Performance enhanced and cost effective Redis offering in-memory caching and high-speed access to applications.
NoSQL Semi-structure or non-structured databases built for superior performance.	ApsaraDB for MongoDB	A secure, reliable, and elastically scalable MongoDB service.
	Time Series Database (TSDB)	A stable, reliable, and cost-effective online high performance time series database service.
	GDB Coming soon	Fast and reliable Graph database with TinkerPop Gremlin interface and ACID property.
	ApsaraDB for HBase New	Performance and reliability enhanced HBase service also serving Alibaba core e-Commerce systems.
	ApsaraDB for Cassandra New	Highly reliable Cassandra with multi-active across available zones and CQL interface.

Datenbanken von Alibaba (Bild 5)

wollen, so will er – oder sie – natürlich im Idealfall die Datenbank mit einem Lift-and-shift in die Cloud verschieben. Auf den Punkt gebracht: Wenn die einzige notwendige Anpassung in der Applikation der Connection-String ist, so ist das Lift-and-shift gelungen. Aus diesem Grund gibt es bei der Alibaba Cloud die getrennten Ausprägungen ApsaraDB RDS for PostgreSQL, ApsaraDB RDS for MySQL, ApsaraDB RDS for SQL Server und ApsaraDB RDS for MariaDB TX.

PolarDB

Jede Cloud hat ihre eigene Cloud-Native-Lösung. Bei AWS ist es Aurora, bei Azure SQL Server und bei Alibaba

Vergleich aller großen Anbieter, der zeigt, dass wirklich jeder Dienst eines Cloud-Anbieters in irgendeiner Form auf den eines anderen Anbieters gemappt werden kann.

Datenbanken

Um nicht mit der grundsätzlichen Realität schließen zu müssen – dass es nämlich in vielen Fällen tatsächlich keinen Unterschied machen wird, welche technischen Features ein Cloud-Anbieter bringt –, soll ein Dienst der Alibaba Cloud näher betrachtet werden, der hervorsteht und sich teilweise vom Wettbewerb abhebt.

Wer sich Database-Benchmarks ansieht, erkennt bald, dass die Alibaba Cloud in einigen Kategorien führend ist [5]. PolarDB gilt auch als das am schnellsten wachsende Produkt innerhalb der Alibaba Cloud. Was zeichnet also relationale Datenbanksysteme bei Alibaba Cloud aus?

Wer relationale Datenbanksysteme auf Alibaba Cloud durchsucht, findet wie auch bei den anderen Cloud-Anbietern mehrere Varianten (vergleiche Bild 5). Doch warum gibt es in einer relationalen Datenbank (OLTP) in der Alibaba Cloud sechs Systeme zur Auswahl?

Einige dieser angebotenen Datenbanken sind Wrapper über bestehende Datenbanksysteme, die bereits in der On-Premises-Welt verbreitet sind. Würde nun jemand seine MySQL- oder PostgreSQL-Datenbank in die Cloud migrieren

Cloud eben PolarDB. Der Grund dafür ist, dass Datenbanksysteme wie MySQL nicht für die Cloud entwickelt wurden, sondern aus einer Zeit stammen, in der Einzelrechner der Standard waren. High-Concurrency-Zugriffe sind nicht Teil der Kernarchitektur. Auch bringt jedes bestehendes Datenbanksystem Legacy mit sich, das mit in die Cloud migriert werden muss. Eine Datenbank komplett neu und Cloud-Native-geeignet zu entwerfen und dabei kompatibel mit Standards (zum Beispiel MySQL) zu sein, ist daher naheliegend.

PolarDB kann hier als ein Flaggschiff von Alibaba Cloud verstanden werden. Auf der untersten Ebene findet sich das verteilte Filesystem, das die Daten auf mehrere Knoten spiegelt und dabei mit RDMA und NVMe arbeitet, was dieses Dateisystem pfeilschnell macht und auch die CPU auf den Storage-Knoten schont. Der Parallel-Raft-Algorithmus bringt außerdem Konsistenz über mehrere Datenreplika. Auch eine Multiversion Concurrency Control (MVCC) hilft dabei, Snapshots über einzelne Compute Nodes zu isolieren. Im Ergebnis gilt: Wenn ein Eintrag auf dieses Filesystem geschrieben wird, sind die Daten außerordentlich schnell über mehrere Server repliziert.

Die Performance-Zahlen des PolarFS-Paper unter [6] künden es zu einem der schnellsten verteilten Dateisysteme am Markt. Die Bedeutung eines Hochleistungsdateisystems mit diesen Zahlen wird erst so richtig klar, wenn man sich ande-

re Aspekte der Architektur ansieht: PolarDB trennt in Compute- und Storage-Knoten. Compute Nodes konzentrieren sich auf SQL-Parsing und Optimierung und Verarbeitung von parallelen Abfragen und belasten somit Storage nicht, was zu hoch performanten Transaktionen führt.

Das Prinzip ist in Bild 6 gezeigt. Es gibt nur einen Leseschreib-Server, während die anderen Server auf Daten rein lesend zugreifen. Alle Server verwenden das verteilte Filesystem, das auf performanten Zugriff optimiert ist. Polar kann 15 Leseknoten haben und brüstet sich damit, auch AWS Aurora zu übertrumpfen, das nur sechs Leseknoten unterstützt. Kennzahlen von 1000 vCPUs, die parallel Abfragen rechnen, beeindrucken ebenfalls. PolarDB gibt an, über dieses System 87 Millionen Abfragen pro Sekunde zu schaffen.

Auch im Sicherheitsbereich braucht sich PolarDB nicht bescheiden zu geben. Das Modul AllEncrypted verheiratet den Datenbank-Kernel mit Security-Software, wodurch ein Kunde alle Daten mit einem Private Key verschlüsseln kann. Das System baut auf einem System wie Intel Software Guard Extensions (SGX) [7] auf. Weitere Details zu PolarDB sind unter [8] zu finden.

Fazit

Manche mögen die Strategie von Alibaba Cloud „typisch chinesisch“ nennen: Erfolgreiche Produkte werden geklont, bis man in der Lage ist, sie in einzelnen Bereichen zu übertreffen. Dank niedriger Kosten kann man mit guter Qualität auch eine Preisschlacht führen. Ein Vorgehen, das China zu einer führenden Wirtschaftsmacht werden ließ, scheint auch im Cloud-Bereich funktionieren zu können.

● Bis zum letzten Cent

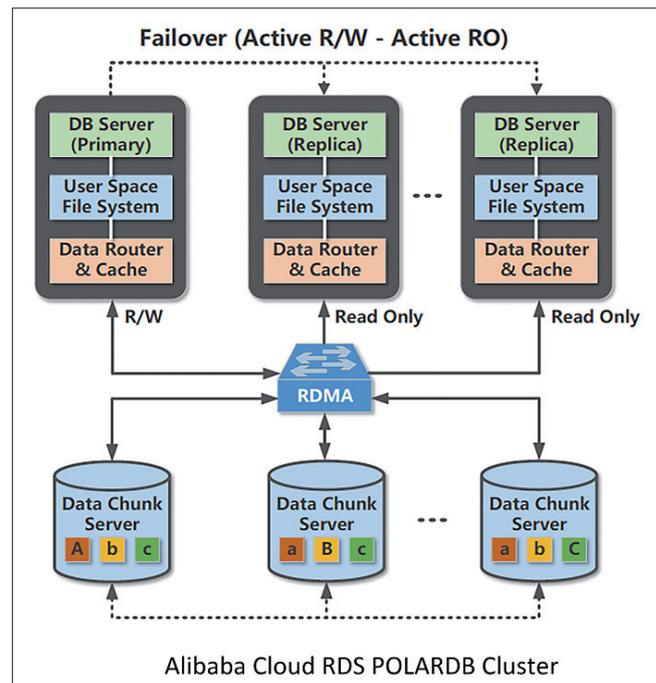
Im letzten Jahrzehnt war es die Frage „Wie hältst du es mit der Cloud?“, mit der Verkäufer von IT-Dienstleistern das Verkaufsgespräch zur Migration von On-Premises auf die Cloud einleiteten. Ein Kriterium machte den Gang in die Cloud besonders schmackhaft: der Preis. Cloud-Native-Lösungen waren kostengünstiger als On-Premises.

In den 2020er-Jahren wird die Frage nicht mehr „ob Cloud“ lauten, sondern „welche Cloud“. Abgesehen natürlich von ein paar wenigen Firmen, die sich berufen fühlen, ein „gallisches Dorf“ zu sein.

Damit ist auch die Preisschlacht eingeleitet. In Extremszenarien wird bis zum letzten Cent gekämpft. Doch gerade in der Cloud kann jeder Cent in Tausende Euro hochskalieren.

Cloud-Anbieter werden mit einfachen Migrationen locken. Gleichzeitig werden sie ihre eigene Domäne dadurch absichern, indem sie die Kosten für den Datentransfer aus der Cloud deutlich erhöhen.

Wer hier in den Dimensionen eines „Wirtschaftskriegs“ denkt und vielleicht auch nicht das allzu größte Vertrauen in die Menschheit hat, dem ist auch klar, dass das richtig schmutzig werden kann.



Architektur von PolarDB (Bild 6)

Datenbanksysteme wie PolarDB beeindrucken, und wer sich das PolarFS-Paper unter [6] durchliest und auch die ganzen Konzepte dahinter verinnerlicht, versteht auch, dass die Alibaba Cloud sich technologisch nicht verstecken muss.

Was am Ende jedoch bleibt, und das ist vor allem für Leser, die Technik lieben, leider ernüchternd: Letztlich werden wohl Preis und Politik entscheiden, wie es am Cloud-Markt weitergeht, und nicht die technischen Lösungen. ■

[1] Why Did Alibaba Choose Apache Flink Anyway?,

www.dotnetpro.de/SL2104Alibaba1

[2] Alibaba Cloud Products & Services,

www.dotnetpro.de/SL2104Alibaba2

[3] ECS-Instanztypen, www.dotnetpro.de/SL2104Alibaba3

[4] Comparecloud, <http://comparecloud.in>

[5] Aktuelle TPC-C-Benchmarks,

www.dotnetpro.de/SL2104Alibaba4

[6] PolarDB-Filesystem, www.dotnetpro.de/SL2104Alibaba5

[7] Intel SGX bei Wikipedia,

www.dotnetpro.de/SL2104Alibaba6

[8] ApsaraDB for PolarDB,

www.dotnetpro.de/SL2104Alibaba7



Stefan Papp

arbeitet als freiberuflicher Big-Data-Architekt und -Evangelist und begleitet zahlreiche Kunden dabei, sich mit der Analyse von neuen Datenquellen selbst neu zu definieren.

dnCode

A2104Alibaba