



IOT IN DER CLOUD

ERKENNTNISSE UND ERFAHRUNGEN EINES PLATTFORMVERGLEICHS.

Wussten Sie, dass man anhand des Bewegungsmusters einer Kuh erkennen kann, an welcher Krankheit sie leidet? So kann man sie gezielt mit einem Antibiotikum behandeln. Wenn Sie eine Kuhherde mit fünf Kühen betreiben, könnte Ihnen das ohne Hilfsmittel auffallen. Bei einer Größe von mehreren hundert Tieren wird es schwierig. Da könnte es eine gute Idee sein, die Tiere mit GPS-Sensoren auszustatten und die Überwachung von einer Software zu erledigen. Wenn ein paar zehntausend Sensoren zur Prüfung der Erde, des Niederschlags und der Temperatur hinzukommen, wird die Verarbeitung der Events schwierig bis unmöglich. Dann wird es Zeit, über die Einführung einer IoT-Cloud nachzudenken.

Was ist eine IoT-Cloud?

Eine IoT-Cloud ist ein fertiger Technologie-Stack, der dabei unterstützt, Dinge mit Backend-Komponenten oder aber auch wieder mit anderen Dingen kommunizieren zu lassen. Die Technologien sind dabei so stark miteinander integriert, dass Themen wie Device Management, Integration, Schnittstellen-Design bis hin zu Big Data und Analytics nahtlos konfiguriert werden können. Ein Dashboard für einen Temperatursensor oder aber auch komplexe Analysen wie das obige Beispiel mit den Kühen lassen sich somit schnell und effizient realisieren.

Natürlich kann man notwendige Technologien selbst miteinander kombinieren. Die Vorteile, die auch bei anderen Cloud-Diensten zu finden sind, liegen aber auf der Hand. Alleine die Beherrschung der Komplexität zur Ausführung von Analysen setzt einiges an Kenntnissen voraus, zum Beispiel das Aufsetzen eines eigenen Spark Clusters^[1]. Kommen weitere Themen wie die Integration der IoT Devices oder auch Machine Learning hinzu, ist in vielen Szenarien eine selbständige Umsetzung nicht tragfähig. Nicht außer Acht gelassen werden sollte auch die Skalierung, die über einen Cloud-Ansatz möglich ist.

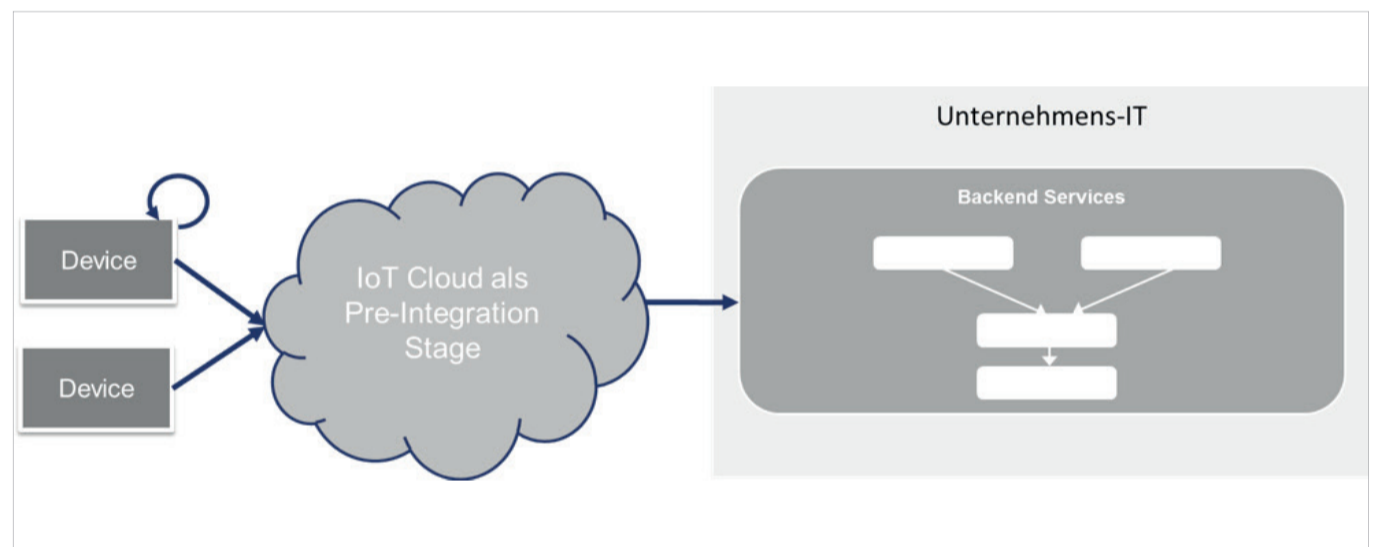


Bild 1: Integrationsebenen.

Den Einsatz ein IoT-Cloud verstehen wir als eine Art Pre-Integration Stage, wie Bild 1 zeigt. Es gibt weitere Ansätze, typischerweise ist eine IoT-Cloud aber ein Konstrukt, das zwischen IoT Devices und weiteren Backend Services platziert wird. Die Cloud übernimmt dabei Aufgaben wie Anbindung, Gerätemanagement und Daten-Analyse. Aus IoT-Daten werden Business Events erzeugt, die dann den eigentlichen Backend Services zugeführt werden.

In unserem Kuh-Beispiel könnten Regensensoren genutzt werden, um Zugänge zu den Stallungen je nach Wetterlage zu steuern. Sensoren nehmen dabei den Regen an mehreren Stellen einer Weide wahr und melden dies der Cloud. Stellen mehr als zehn Sensoren den gleichen Zustand fest, wirft die Analytics-Komponente der Cloud-Lösungen ein Event, dass ein Backend System dazu veranlasst, den Weg zu den Stallungen zu öffnen.

Kandidatenauswahl

Es gibt am Markt eine Vielzahl von Cloud-Plattformen. Einige überleben wenige Monate, viele leben ein Nischendasein und wenige haben sich zu Marktführern entwickelt. Um die passende Cloud-Plattform zu finden, ist eine gute Voranalyse Pflicht. So kann eine Fehlentscheidung vermieden werden, die bei einem Plattformwechsel zu extrem hohen Zusatzaufwänden und -kosten führen kann.

Eine große Anzahl von Plattformen bis ins Detail miteinander vergleichen zu wollen, ist nicht praktikabel. Für unseren Vergleich haben wir eine Vorauswahl getroffen und uns für vier Plattformen entschieden. Für eine Eingrenzung sind Fachartikel und Berichte zu IoT-Plattformen gut geeignet. Unter anderem haben uns auch der „Crisp Vendor Universe“ [2] und die „Industrie 4.0/ IoT Rising Star Germany“ [3] die Auswahl erleichtert.

Entschieden haben wir uns für die Amazon Web Services (AWS) [4], Windows Azure [5], Oracle Cloud [6] und Q-Loud [7]. AWS und Azure gehören zu den Marktführern. Oracle hat im Cloud-Umfeld in den letzten Jahren Boden gut gemacht. Q-Loud ist ein mittelständisches Unternehmen, das gerade für Retrofitting und bei gleichzeitiger Hardware-Entwicklung interessant sein kann.

Mittelständler können bestimmte Nischen für einzelne Domänen ausfüllen, deshalb sollte man sie bei einer Auswahl nicht außer Acht lassen. Eine Betrachtung entsprechender Möglichkeiten würde aber den Rahmen dieses Artikels sprengen.

Kriterien

Um die Plattformen möglichst objektiv miteinander zu vergleichen, haben wir einen Kriterienkatalog erarbeitet. Die Kriterien haben wir dabei in die Hauptkriterien Plattformeigenschaften, Plattform-Features, Kommunikation und Sonstiges aufgeteilt (siehe Bild 2).

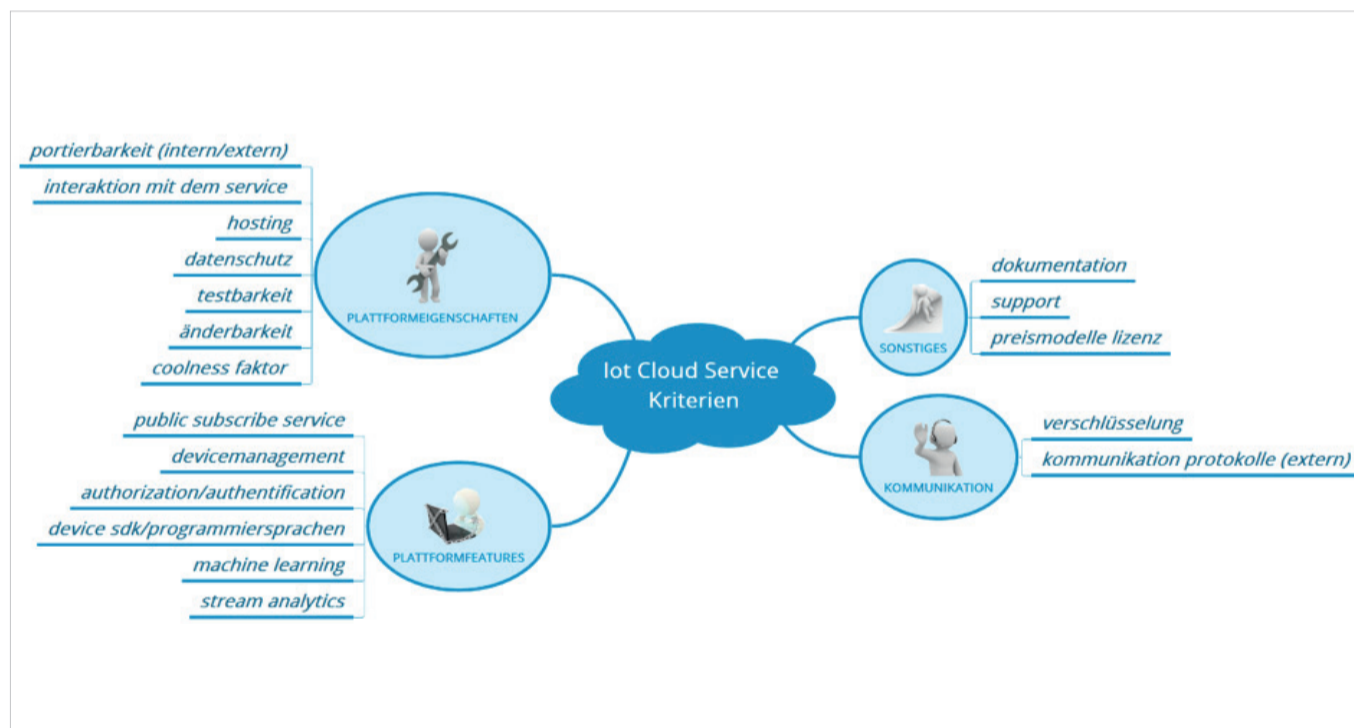


Bild 2: Kriterien für den Plattformvergleich.

Für die Bewertung haben wir sowohl die Dokumentation, Gespräche und Ergebnisse aus der Softwareentwicklung herangezogen.

Gemeinsamer Showcase

Theoretische Überlegungen werden schnell von der Praxis eingeholt und häufig widerlegt. Deshalb haben wir einen Beispiel-Showcase entwickelt und diesen einmal mit jeder der vier Plattformen implementiert. Im Showcase, den wir gemeinsam mit einem Kunden entwickelt haben, werden die Verkäufe von Verkaufsautomaten überwacht mit dem Ziel, Verkaufsautomaten bedarfsgerecht befüllen zu lassen, wenn



ein Produkt zur Neige geht. Statistiken über die Verkaufszahlen können aber zum Beispiel auch dabei helfen, defekte Automaten zu erkennen. Ein Automat ist mit hoher Wahrscheinlichkeit kaputt, wenn über einen längeren Zeitraum keine Waren verkauft werden. Im Folgenden werden die einzelnen Kriterien und die gewonnenen Erkenntnisse vorgestellt.

Plattformeigenschaften

Steigen wir mit den Kriterien zu den Plattformeigenschaften ein.

Hosting

Hosting ist die Klärung, in welchem Land die Cloud-Landschaft effektiv betrieben wird. Der Hintergrund: Je nachdem, in welchem Land die Daten gespeichert werden, unterliegt der Betreiber der Gerichtsbarkeit des Host-Landes. Bei den großen Anbietern Microsoft, Oracle und Amazon kann das Hosting der Daten auf Servern auf fast allen Kontinenten konfiguriert werden. Aber Vorsicht: Gegebenenfalls lassen sich bestimmte Konfigurationen nicht in allen Ländern nutzen. Q-Cloud als mittelständisches Unternehmen beschränkt sich dabei auf Deutschland als Hosting-Land. Seit dem NSA-Skandal wird bewusster hinterfragt, in welchem Land die Daten gehostet werden. Das führt direkt zum Thema Datenschutz.

Datenschutz

Datenschutz ist eines der Kernthemen dieser Zeit. Schließlich sollen weder private Daten noch Unternehmensdaten in die falschen Hände geraten. Alle Hersteller legen hohen Wert auf Datenschutz. Bei Amazon AWS ist es beispielsweise möglich, neben dem Standort der Daten die Verschlüsselung oder Anonymisierung der Daten einzustellen. Darüber hinaus können bei allen Anbietern Identitäts- und Zugriffsmanagement als auch Sicherheitsanmeldeinformationen festgelegt werden.

Oracle bietet unter anderem eine Konfiguration an, die darüber entscheidet, ob Daten für Marketinginformationen verwendet werden dürfen.

Microsoft Azure führt kein Data Mining auf den Kundendaten durch und betont, dass bei einer Offenlegung der Daten an Regierungs- oder Justizbehörden ihre Kunden unverzüglich informiert werden.

Portierbarkeit

Bei der Portierbarkeit haben alle Hersteller einen kleinen Haken: Vendor-Lockin. Die Migration der Daten von einer Plattform zur anderen Plattform ist bei allen generell möglich, aber mit hohen Aufwänden verbunden. Die Hersteller möchten ihre Kunden lange an ihre Cloud-Plattform binden und erschweren daher den direkten Wechsel von einer Plattform in eine andere. Das macht entsprechende Migrationsprojekte notwendig. Schon bei der Auswahl einer Plattform sollte man dies im Hinterkopf haben. Die hohen Kosten für einen Wechsel können dazu führen, dass man lieber weiter „in den sauren Apfel beißt“ als die Plattform zu wechseln.

Proof of Concepts können dabei helfen, Plattformen frühzeitig zu evaluieren und gegebenenfalls sogar mit einem Minimal Viable Product im Produktionsbetrieb zu starten, um Erfahrungen zu Kosten und Betrieb zu sammeln, bevor man sich bindet.

Interaktion: Oberfläche und Schnittstellen

Die großen Hersteller haben eine ähnliche Vorgehensweise, mit dem Cloud Service zu interagieren. Sie bieten webbasierte Portale an, auf denen die ermittelten Daten in Echtzeit angezeigt, analysiert und verarbeitet werden können.



Daniel Rock,
Mitglied des Competence Centers
„Internet of Things“ bei Opitz Consulting

Q-Loud bietet dabei eine Besonderheit. Hier ist die Schnittstelle eine REST-API. Da Entwickler als Zielpublikum für die Plattform gedacht sind, ist es eine Schnittstelle von Entwicklern für Entwickler. So können die Daten zielgerichtet abgegriffen und verarbeitet werden. Die REST-API ist also der Standard-Weg. Die API ist gut dokumentiert und durch die Nutzung entstehen direkt automatisierbare Artefakte in Form von REST Requests.

Testbarkeit

Es ist wichtig, die Plattform zu testen, um ein Gefühl für Machbarkeit, Lösungsansätze und Entwicklungsprozesse zu bekommen. So erkennen Sie schnell, ob Plattform den Anforderungen entspricht und es genügend Konfigurationsmöglichkeiten für den geplanten Einsatz gibt. Es ist bei allen Plattformen möglich, einen Test-Account zu erhalten. Bei Azure erhalten Sie eine Gutschrift über 170 Euro, wenn Sie die Plattform ausprobieren. Zusätzlich sind 3000 Nachrichten pro Tag kostenlos. AWS bietet 250.000 kostenfreie Nachrichten im ersten Jahr. Bei Oracle ist der Account für 30 Tage gültig. Für Q-Loud muss man sich an den Vertrieb wenden.

Coolness-Faktor

Jeder Hersteller unterliegt einem gewissen „Coolness-Faktor“, der ausschlaggebend für die Auswahl sein kann. So ist je nach Zielpublikum die Reaktion auf die Plattform unterschiedlich. Manche bevorzugen einen eher konservativen Anbieter. Andere wollen besonders angesagte Dinge machen und benötigen eine passende Community im Hintergrund. Amazon und Azure sind hip, sodass dies erheblichen Einfluss auf die Mitarbeitergewinnung haben kann. Der Coolness-Faktor ist folglich nicht außer Acht zu lassen.

Plattform-Features

Unter den Plattform-Features haben wir die am häufigsten geforderten Features untersucht. Dazu gehören Pub/Sub-Services, das Device Management und Authentifizierung/Autorisierung, SDKs und unterstützte Programmiersprachen, Machine Learning und Stream Analytics.

Pub/Sub-Service

Oracle, Microsoft und Amazon sind in diesem Bereich unschlagbar. Sie bieten einen Push- oder Notifikationsservice an, der über Ereignisse informieren und Nachrichten versenden kann. Zusätzlich bietet der Amazon Simple Notification Service (SNS) die Unterstützung von verschiedenen Abonnement-Typen an, um Nachrichten an Warteschlangen, AWS-Lambda-Funktionen und HTTP-Endpunkte weiterzuleiten. Bei Microsoft Azure können verschiedene Push-Benachrichtigungsdienste von Apple, Google und Microsoft angebunden werden. Q-Loud kann diese Fülle an Features nicht abbilden.

Device Management

Alle vier Plattformen bieten eine Übersicht und die Verwaltung der verwendeten Geräte an. So können eindeutige IDs für Geräte vergeben und ihnen Metadaten hinzugefügt werden. Auf diese Weise wird verhindert, dass keine unberechtigten Geräte Daten senden oder empfangen können.

Authentifizierung/Autorisierung

Alle Plattformen benutzen ein ähnliches Konzept zur Geräte-Authentifizierung. Über eine Geräte-Id und einen Key werden IoT Device und -Applikation identifiziert. Nachfolgende Berechnungen können gestartet werden.

Eine Besonderheit hat Q-Loud im Hinblick auf die Geräte-IDs. Wird die Hardware direkt über Q-Loud bezogen oder über Q-Loud produziert, wird die Geräte Id normalerweise direkt mit auf die Hardware geschrieben. Infolgedessen wird es für Angreifer schwieriger, eine Geräte Id von der Hardware abzugreifen, um eine „Man in the Middle“-Attacke“ über ein IoT Device durchzuführen.



Dominik Bial,
Solution Architect
bei Opitz Consulting

Device SDK und Programmiersprachen

Wie bei den anderen Kriterien unterscheiden sich die drei großen Hersteller auch bei Device SDK und Programmiersprachen nicht sonderlich. Sie bieten alle SDKs an, damit Kunden eigene Clients und Gateways schreiben können. Die gängigen Programmiersprachen Java, C, JavaScript, .NET etc. werden dabei abgedeckt.

Machine Learning

Bis auf Q-Loud wird Machine Learning von allen Herstellern angeboten. Es verursacht weitere Kosten, da es als zusätzlicher Service hinzugebucht werden muss. Microsoft bietet beispielsweise das ML Studio an, in dem per Drag and Drop der gesamte Machine-Learning-Prozess modelliert werden kann.

Stream Analytics

Oracle bietet Stream Analytics inhärent als Teil der Plattform an, bei Microsoft und Amazon muss der Service kostenpflichtig hinzugebucht werden. Die Verwendung der Daten erfolgt über EQL, eine SQL-ähnliche Sprache, die zum Beispiel um Zeitintervalle betrachtet wurde. Beispielsweise wären Abfragen möglich wie „Wirf ein Event, wenn in den letzten 3 Stunden keine Kuh trinken war“. Q-Loud stellt eine regelbasierte Mustererkennung bereit.

Kommunikation: Protokolle und Verschlüsselung

Alle Hersteller haben gemeinsam, dass die Kommunikation über TLS und eine Zwei-Wege-Authentifizierung stattfindet. Sie bieten alle HTTP als Protokoll an. Die großen Hersteller bieten zusätzlich MQTT an, Q-Loud würde dies auf Nachfrage auch möglich machen. Azure stellt des Weiteren AMQP zu Verfügung. Bei Amazon können darüber hinaus weitere Protokolle über das eigene SDK hinzugefügt werden.

Security und Cloud-Kommunikation

Die großen Hersteller bringen Security-Features mit. Daten von Gerät und zur Cloud sind verschlüsselt. Der Weg vom Gerät in die Cloud hinein wird jedoch nicht abgedeckt, da Cloud-Angebote im Normalfall keine Hardware umfassen. Man ist selbst dafür verantwortlich, Geräte und Daten an die Cloud anzubinden. Entsprechende Hardware-Anbieter versuchen in die Bresche zu springen und mittels ihrer Produkte oder entsprechender Serviceleistungen zu unterstützen.

Q-Loud fährt hier ebenfalls ein interessantes Konstrukt. Bezieht man die IoT Devices von Q-Loud, ist es möglich Kosten, für die Cloud auf die Hardware umzulegen. So erhöhen sich nur die Stückkosten für das Produkt, und es entstehen keine Kosten für die Nutzung durch Endkunden.

Sonstiges: Community und Dokumentation

Hier spiegelt sich der bereits erwähnte Coolness-Faktor wider. Alle großen Hersteller bieten online Dokumentationen an. Für Microsoft und Oracle gibt es viele Einträge in Benutzerforen wie etwa Stack Overflow. Hingegen sind bei Oracle die herstellereigenen Seiten die größten Informationsquellen.

Bei Q-Loud läuft das anders. Neben der API-Beschreibung und dem Schnellstart-Material, das es bei allen Herstellern gibt, ist eine direkte Kommunikation mit den Entwicklern möglich, die auf Anfragen dann auch tatsächlich relativ schnell antworten.

Support

Bei Q-Loud und Oracle ist der Support im Service enthalten. Bei Amazon und Azure muss er hinzugebucht werden. Dabei sind verschiedene Stufen (zum Beispiel Developer, Standard oder Enterprise) auswählbar.

Kosten und Preismodelle

Unsere erste Intuition war es, einen einfachen Use Cases zu definieren, anhand dessen wir die einzelnen Cloud-Angebote messen und vergleichen. Also definierten wir ein imaginäres, mittelständiges Unternehmen, das in der Landwirtschaft Telemetriedaten wie Temperatur und Niederschlagsmengen sammelt. Mit 25.000 Geräten, die alle 10 Minuten eine Nachricht pro Gerät versenden, errechneten wir eine Menge von 3,6 Millionen Nachrichten am Tag.

Im ersten Moment ist das Ergebnis erschreckend. Die Preispanne rangierte von einem dreistelligen bis hin zu einem fünfstelligen Betrag. Schaut man genauer hin, erkennt man aber schnell, dass die Preise nicht miteinander vergleichbar sind.

Bei Oracle und Q-Loud sind alle Kosten abgehandelt. Bei Microsoft und Amazon muss auf Zusatzfunktionen und Features geachtet werden, die mitbenutzt werden sollen. Hier macht man schnell den Fehler, nur auf die Kosten für die Nachrichten zu achten. Mehrwerte werden aber durch weitere, hinzuzubuchende Services erbracht, die zusätzlich Geld kosten. Für die Berechnung der Kosten bräuchte man einen Use Case in einem sehr hohen Detaillierungsgrad, um verlässliche Aussagen treffen zu können. Außerdem sollten zukünftige Änderungen am Use Case bedacht werden. In vielen Projekten sind die finalen Kosten nicht absehbar. Besonders interessant wird es, wenn die Kosten abhängig von den Benutzern sind. Wird beispielsweise davon ausgegangen, dass Informationen über die Kühe zweimal am Tag angefragt werden, es in der Realität jedoch im Durchschnitt sechsmal geschieht, ist die ursprüngliche Schätzung schnell nicht mehr haltbar. Selbst wenn die Abfragen nur dem Spieltrieb des Menschen geschuldet sind.

Zusammenfassung und Fazit

Die Auswahl der passenden IoT-Cloud-Plattform ist keine einfache Entscheidung. Bei den großen Herstellern sind die Features alle recht ähnlich, der Teufel steckt jedoch im Detail. Eine Gewichtung der Kriterien kann dabei helfen, die richtige Plattform auszuwählen.

Auch einen kleinen Mittelständler wie Q-Loud sollte man dabei nicht aus den Augen verlieren. Bei ihm haben wir beispielsweise bei Hardwarestrecken oder Security-Themen herausragende Features entdeckt. Diese können für die Auswahl einer Plattform ausschlaggebend sein.

Ebenfalls zu bedenken ist, dass derzeit ein Boom auf dem Markt herrscht. Es bleibt noch abzuwarten, wie sich die Plattformen etablieren. Noch könnte es passieren, dass die gewählte Plattform im nächsten Monat vom Markt verschwunden ist.

Daniel Rock, Dominik Bial

Links

- [1] Apache Spark: <https://spark.apache.org>
- [2] Crisp Vendor Universe: <https://www.crisp-research.com/vendor-universe/cloud-computing>
- [3] I4.0/IoT Vendor Benchmark 2016: <http://research.isg-one.de/research/studien/industrie-40-iot-vendor-benchmark-2017/ueberblick.html>
- [4] Amazon Web Services: <https://aws.amazon.com>
- [5] Windows Azure: <https://azure.microsoft.com>
- [6] Oracle Cloud: <https://cloud.oracle.com>
- [7] Q-Loud: <https://www.q-loud.de>