Sicherungs- und Wiederherstellungslösungen mit AWS

Juni 2016

Hinweise

Dieses Dokument wird nur zu Informationszwecken zur Verfügung gestellt. Es stellt das aktuelle Produktangebot und die Praktiken von AWS zum Erstellungsdatum dieses Dokuments dar. Änderungen vorbehalten. Kunden sind für ihre eigene unabhängige Einschätzung der Informationen in diesem Dokument und jedwede Nutzung der AWS-Services verantwortlich. Jeder Service wird „wie besehen“ ohne Gewähr und ohne Garantie jeglicher Art, weder ausdrücklich noch impliziert, bereitgestellt. Dieses Dokument gibt keine Garantien, Gewährleistungen, vertragliche Verpflichtungen, Bedingungen oder Zusicherungen von AWS, seinen Partnern, Zulieferern oder Lizenzgebern. Die Verantwortung und Haftung von AWS gegenüber seinen Kunden werden durch AWS-Vereinbarungen geregelt. Dieses Dokument ist weder ganz noch teilweise Teil der Vereinbarungen von AWS mit seinen Kunden und ändert diese Vereinbarungen auch nicht.

Inhalt

Übersicht 4

Einführung 4

Warum AWS als Sicherungsplattform? 4

AWS-Speicherservices für die Datensicherung 6

- Amazon S3 6
- Amazon Glacier 7
- AWS Storage Gateway 7
- AWS Transferservices 7

Entwerfen einer Sicherungs- und Wiederherstellungslösung 7

Cloud-Infrastruktur 9

- Sicherungen mit EBS-Snapshots 10
- Ansätze für Datenbanksicherungen 15

Lokale Infrastruktur und AWS-Infrastruktur 18

Hybride Umgebungen 21

- Sichern von AWS-basierten Anwendungen in Ihrem Rechenzentrum 23
- Migrieren der Sicherungsverwaltung in die Cloud zur Gewährleistung der Verfügbarkeit 23
- Beispiel für ein hybrides Szenario 24
- Archivieren von Daten mit AWS 25

Schützen von Sicherungsdaten in AWS 27

Zusammenfassung 27

Mitwirkende 28

Am Dokument vorgenommene Änderungen 28

Übersicht

Dieses Dokument wendet sich an Architekten von Enterprise- und Sicherungslösungen sowie an IT-Administratoren, die für den Schutz der Daten in den IT-Umgebungen ihres Unternehmens verantwortlich sind. Vorgestellt werden Produktions-Workloads und Architekturen, die mithilfe von AWS implementiert werden können, um eine Sicherungs- und Wiederherstellungslösung zu erweitern oder zu ersetzen. Diese Ansätze ermöglichen geringere Kosten, bessere Skalierbarkeit und höhere Zuverlässigkeit, sodass alle Anforderungen bezüglich Recovery Time Objective (RTO), Recovery Point Objective (RPO) und Compliance erfüllt werden können.

Einführung

Durch das beschleunigte Wachstum von Unternehmensdaten wird die Aufgabe, diese zu schützen, immer anspruchsvoller. Fragen über die Zuverlässigkeit und Skalierbarkeit von Sicherungsmethoden sind an der Tagesordnung, insbesondere diese: Wie kann die Cloud den Sicherungs- und Archivierungsbedürfnissen des Unternehmens gerecht werden?

Dieses Dokument behandelt eine Reihe von Sicherungsarchitekturen (reine Cloud-Anwendungen sowie hybride und lokale Umgebungen) und die zugehörigen AWS-Services, die skalierbare und zuverlässige Datenschutzlösungen ermöglichen.

Warum AWS als Sicherungsplattform?

Bei Amazon Web Services (AWS) handelt es sich um eine sichere, leistungsstarke, flexible, kosteneffiziente und leicht zu handhabende Plattform für das Cloud Computing. AWS erledigt die undifferenzierten Hauptaufgaben und stellt Tools und Ressourcen bereit, mit deren Hilfe Sie skalierbare Sicherungs- und Wiederherstellungslösungen erstellen.

Als Teil Ihrer Datenschutzstrategie bietet AWS etliche Vorteile:

- **Zuverlässigkeit:** Amazon Simple Storage Service (Amazon S3) und Amazon Glacier garantieren eine Beständigkeit von 99,999999999 % (11 Neunen) für die darin gespeicherten Objekte. Beide Plattformen bieten zuverlässige Standorte für Sicherungsdaten.
- **Sicherheit:** AWS bietet eine Reihe von Optionen für die Zugangskontrolle sowie zur Verschlüsselung der Daten bei der Übertragung und im Ruhezustand.
- **Weltweite Infrastruktur:** AWS-Services werden überall auf der Welt angeboten, sodass die Sicherung und Speicherung der Daten in einer Region erfolgen kann, die Ihre Compliance-Anforderungen erfüllt.
- **Compliance:** Die AWS-Infrastruktur ist bezüglich der Compliance mit folgenden Normen zertifiziert: Service-Organisation Controls (SOC), Statement on Standards for Attestation Engagements (SSAE) 16, International Organization for Standardization (ISO) 27001, Payment Card Industry Data Security Standard (PCI DSS), Health Insurance Portability and Accountability Act (HIPPA), SEC[1] und Federal Risk and Authorization Management Program (FedRAMP). Sie können Ihre Sicherungslösung also problemlos in ein bestehendes Compliance-Schema integrieren.
- **Skalierbarkeit:** Mit AWS gehören Kapazitätsprobleme der Vergangenheit an. Sie können Ihren Verbrauch je nach Bedarf und ohne administrativen Overhead in beide Richtungen skalieren.
- **Niedrige Gesamtbetriebskosten:** Das umfangreiche AWS-Angebot reduziert sowohl die Kosten für die Services als auch die Gesamtbetriebskosten für die Datenspeicherung. Diese Kosteneinsparungen gibt AWS in Form von Preisreduzierungen an die Kunden weiter.
- **Nutzungsabhängige Preise:** Sie können AWS-Services nach Bedarf kaufen und zahlen nur für die Nutzungsdauer. Bei AWS gibt es keine Vorabzahlungen, Kündigungsgebühren oder langfristige Verträge.

[1] https://aws.amazon.com/about-aws/whats-new/2015/09/amazon-glacier-receives-third-party-compliance-assessment-for-sec-rule-17a-4f-from-cohasset-associates-inc/

AWS-Speicherservices für die Datensicherung

Amazon S3 und Amazon Glacier sind ideale Services zur Sicherung und Archivierung von Daten. Beide sind beständige, preisgünstige Speicherplattformen. Beide bieten unbegrenzte Kapazität und erfordern kein Volume- oder Medienmanagement, wenn die Sicherungsdatensätze anwachsen. Durch das Modell der nutzungsabhängigen Preise und wegen der geringen Kosten pro GB/Monat sind diese Services die passende Wahl für Sicherungsanwendungsfälle.

Amazon S3

Amazon S3 bietet eine hochsichere, skalierbare Speicherung für Objekte.

Mit Amazon S3 können Sie jederzeit beliebige Datenmengen über das Internet speichern und abrufen. Amazon S3 speichert Daten als Objekte in Ressourcen, die *Buckets* genannt werden. AWS Storage Gateway und viele Sicherungslösungen von Drittanbietern sind in der Lage, Ihre Amazon S3-Objekte zu verwalten. Sie können beliebig viele Objekte in einem Bucket speichern sowie darin schreiben, lesen und Objekte löschen. Die maximale Größe eines Objekts beträgt 5 TB.

Amazon S3 bietet eine Reihe von Speicherklassen, die für verschiedene Anwendungsfälle ausgelegt sind. Dazu gehören:

- **Amazon S3 Standard** für die Speicherung beliebiger Daten, auf die häufig zugegriffen wird.
- **Amazon S3 Standard - Infrequent Access** für Langzeitdaten mit selteneren Zugriffen.
- **Amazon Glacier** für Langzeitarchivierung.

Amazon S3 bietet außerdem konfigurierbare Lebenszyklusrichtlinien zur Verwaltung Ihrer Daten im gesamten Lebenszyklus. Sobald eine Richtlinie eingerichtet ist, werden Ihre Daten automatisch in die passende Speicherklasse migriert, ohne dass Änderungen an Ihrer Anwendung erforderlich sind. Weitere Informationen finden Sie unter S3 Speicherklassen.

Amazon Glacier

Amazon Glacier ist ein äußerst kostengünstiger Speicherservice, der in der Cloud einen sicheren und beständigen Speicher für Datenarchivierung und Online-Datensicherung bereitstellt. Zur Kostenminimierung ist Amazon Glacier für Daten optimiert, auf die selten zugegriffen wird und für die eine Abrufzeit von mehreren Stunden annehmbar ist. Mit Amazon Glacier können Sie zuverlässig große oder kleine Datenmengen für nur 0,007 USD pro Gigabyte und Monat speichern, wodurch sich signifikante Einsparungen im Vergleich zu lokalen Lösungen ergeben. Amazon Glacier ist bestens geeignet zur Speicherung von Sicherungsdaten mit langen oder unbestimmten Aufbewahrungsfristen sowie zur Langzeitarchivierung von Daten. Weitere Informationen finden Sie unter Amazon Glacier.

AWS Storage Gateway

AWS Storage Gateway verbindet eine lokale Software-Appliance mit Cloud-basiertem Speicher. Dieser Service sorgt für ein nahtloses, hochsicheres Zusammenspiel zwischen Ihrer lokalen IT-Umgebung und der Speicherinfrastruktur von AWS. Weitere Informationen finden Sie unter AWS Storage Gateway.

AWS Transferservices

Zusätzlich zu Gateways und Konnektoren von Drittanbietern stehen Ihnen AWS-Optionen wie AWS Direct Connect, AWS Snowball, AWS Storage Gateway und Amazon S3 Transfer Acceleration für den schnellen Transfer Ihrer Daten zur Verfügung. Weitere Informationen finden Sie unter Datenmigration in die Cloud.

Entwerfen einer Sicherungs- und Wiederherstellungslösung

Wenn Sie eine umfassende Strategie für die Sicherung und Wiederherstellung von Daten entwickeln, müssen Sie zuerst die möglichen Fehler- oder Notfallsituationen und deren potenzielle geschäftliche Auswirkungen identifizieren. In einigen Branchen sind zudem gesetzliche Regelungen bezüglich Datensicherheit, Datenschutz und Aufbewahrungsdauer zu berücksichtigen.

Sie müssen Datensicherungsprozesse implementieren, die das richtige Maß an Granularität bieten, um den RTO- und RPO-Vorgaben des Unternehmens zu entsprechen. Dazu gehören:

- Wiederherstellung auf Dateiebene
- Wiederherstellung auf Volume-Ebene
- Wiederherstellung auf Anwendungsebene (beispielsweise Datenbanken)
- Wiederherstellung auf Abbildebene

In den nächsten Abschnitten werden die Sicherungs-, Wiederherstellungs- und Archivierungsansätze für die drei Kategorien einer IT-Infrastruktur beschrieben. Eine IT-Infrastruktur kann entweder in der Cloud, lokal oder in hybrider Form realisiert sein.

Cloud-Infrastruktur

Dieses Szenario beschreibt ein vollständig in AWS realisiertes Arbeitsumfeld. Wie die folgende Abbildung zeigt, gehören dazu Webserver, Anwendungsserver, Überwachungsserver, Datenbanken und Active Directory.

Wenn Sie alle Services auf AWS ausführen, stehen Ihnen viele integrierte Funktionen zur Verfügung, mit denen Sie Ihren Anforderungen bezüglich Datensicherung und Wiederherstellung gerecht werden.

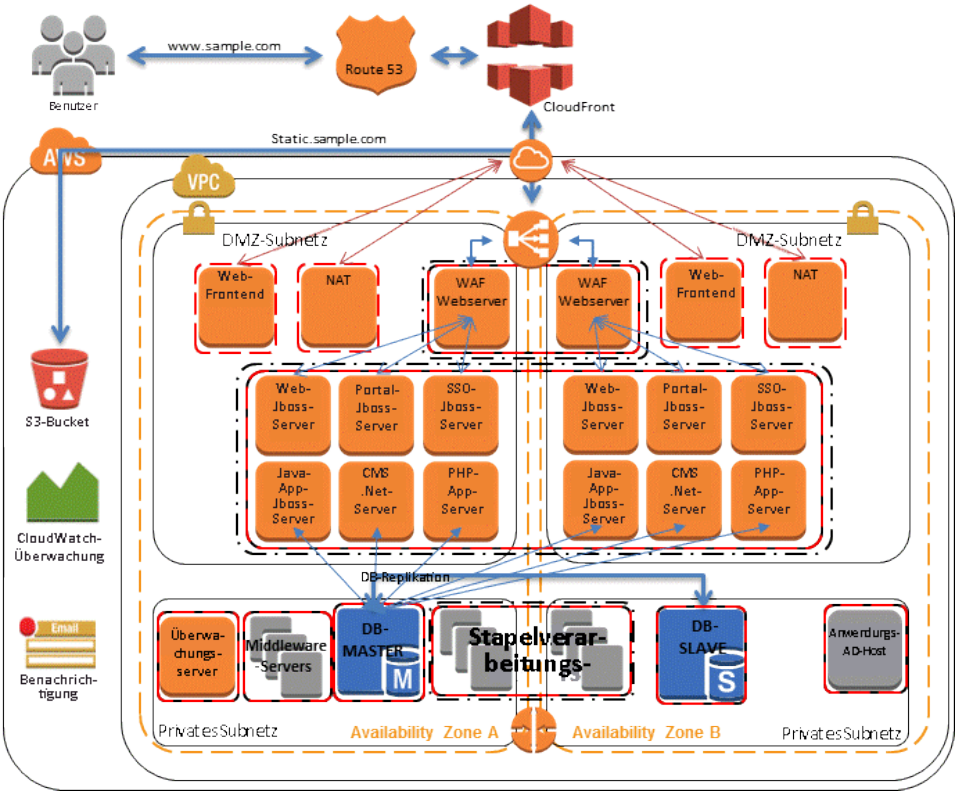

Abbildung 1: Szenario für eine reine AWS-Infrastruktur

Sicherungen mit EBS-Snapshots

Wenn Services in der Amazon Elastic Compute Cloud[2] (Amazon EC2) ausgeführt werden, können Datenverarbeitungs-Instances die Amazon Elastic Block Store (Amazon EBS)-Volumes verwenden, um Daten zu speichern und abzurufen. Diese Blockspeicher sind sowohl für strukturierte Daten (wie Datenbanken) als auch für unstrukturierte Daten (beispielsweise Dateien in einem Dateisystem auf dem Volume) geeignet.

Amazon EBS bietet die Möglichkeit, Snapshots (Sicherungen) von jedem Amazon EBS-Volume zu erstellen. Dabei wird eine Kopie des Volumes nach Amazon S3 übertragen und dort redundant in mehreren Availability Zones gespeichert. Der erste Snapshot ist eine vollständige Kopie des Volumes, die weiteren enthalten nur inkrementelle Änderungen auf Blockebene.

Dies ist eine schnelle und zuverlässige Methode zur Wiederherstellung der Daten des gesamten Volumes. Wenn Sie lediglich eine teilweise Wiederherstellung vornehmen möchten, fügen Sie das Volume der aktiven Instance unter einem anderen Gerätenamen hinzu, aktivieren es und verwenden dann Betriebssystembefehle zum Kopieren der Daten aus dem Sicherungs-Volume in das Produktions-Volume.

Amazon EBS-Snapshot können auch zwischen AWS-Regionen kopiert werden, entweder in der Konsole oder über die Befehlszeile. Dieser Vorgang ist im Handbuch Amazon Elastic Compute Cloud User Guide beschrieben.[3] Mit dieser Funktion speichern Sie Ihre Sicherung in einer anderen Region, ohne sich dabei mit der zugrundeliegenden Datenreplizierungstechnologie auseinandersetzen zu müssen.

---

[2] http://aws.amazon.com/ec2/

[3] http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ebs-copy-snapshot.html

Erstellen von EBS-Snapshots

Beim Erstellen eines Snapshots speichern Sie Ihre Daten direkt in einem dauerhaften, festplattenbasierten Speicher. Ein Amazon EBS-Snapshot lässt sich mithilfe der AWS Management Console, der Befehlszeile oder eines API-Aufrufs erstellen.

In der Amazon EC2-Konsole wählen Sie auf der Seite **Elastic Block Store Volumes** im Menü **Actions** die Option **Create Snapshot**. Im Dialogfeld **Create Snapshot** wählen Sie dann **Create**, um einen Snapshot zu erstellen, der in Amazon S3 gespeichert wird.

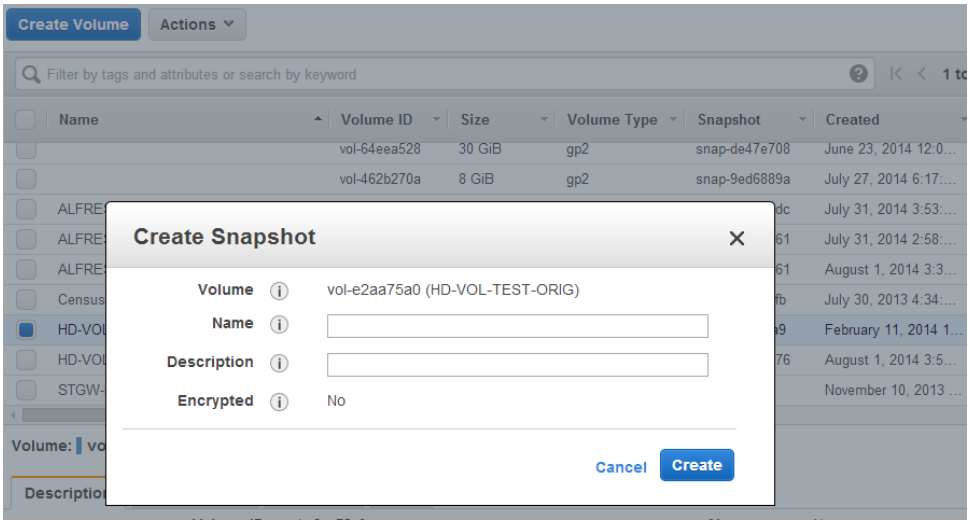

Abbildung 2: Snapshot-Erstellung mit der EC2-Konsole

In der Befehlszeile erstellen Sie mit folgendem Befehl einen Snapshot:

```
➢  aws ec2 create-snapshot
```

Sie können den `aws ec2 create-snapshot`-Befehl regelmäßig zur Sicherung von EBS-Daten verwenden. Aufgrund der wirtschaftlichen Preisstruktur von Amazon S3 lassen sich auf diese Weise viele Datengenerationen aufbewahren. Und da Snapshots blockbasiert sind, verbrauchen sie nach Erstellung des ersten Snapshots zusätzlichen Speicherplatz nur für geänderte Daten.

Wiederherstellung aus einem EBS-Snapshot

Zur Wiederherstellung der Daten aus einem Snapshot erzeugen Sie mithilfe der AWS Management Console, der Befehlszeile oder eines API-Aufrufs ein neues Volume aus dem Snapshot.

Um beispielsweise ein Volumen aus einer früheren Sicherung wiederherzustellen, gehen Sie wie folgt vor:

1. Erstellen Sie ein Volume aus dem Sicherungs-Snapshot mit diesem Befehl:

```
➢ aws ec2 create-volume –region us-west-1b –snapshot-
  id mysnapshot-id
```

2. Deaktivieren Sie das vorhandene Volume in der Amazon EC2 Instance.

   In Linux verwenden Sie dazu den Befehl `umount`. In Windows verwenden Sie den Logical Volume Manager (LVM).

3. Mit dem folgenden Befehl trennen Sie das vorhandene Volume von der Instance:

```
➢ aws ec2 detach-volume –volume-id oldvolume-id –
  instance-id myec2instance-id
```

4. Mit dem folgenden Befehl fügen Sie das aus dem Sicherungs-Snapshot erstellte Volume hinzu:

```
➢ aws ec2 attach-volume –volume-id newvolume-id
  –instance-id myec2instance-id –device /dev/sdf
```

5. Reaktivieren Sie das Volume in der aktiven Instance.

Erstellen konsistenter oder dynamischer Sicherungen

Während eine Sicherung erstellt wird, sollte das System keine Ein- oder Ausgaben vornehmen. In diesem Idealfall verarbeitet das System keinen Datenverkehr. Das ist aber kaum noch realisierbar, da Daten immer häufiger rund um die Uhr verarbeitet werden.

Folglich muss zur Erstellung einer „sauberen“ Sicherung das Dateisystem bzw. die Datenbank stillgelegt werden. Wie Sie dazu vorgehen, ist von Ihrer Datenbank bzw. Ihrem Dateisystem abhängig.

Für eine Datenbank sind folgende Schritt notwendig:

- Aktivieren Sie für die Datenbank nach Möglichkeit den dynamischen Sicherungsmodus, der eine Sicherung während des Betriebs gestattet.
- Führen Sie die Befehle für einen Amazon EBS-Snapshot aus.
- Deaktivieren Sie den dynamischen Sicherungsmodus für die Datenbank. Sollten Sie mit einer Read Replica gearbeitet haben, beenden Sie die Read Replica-Instance.

Für ein Dateisystem sind ähnliche Schritte erforderlich, die allerdings von den Fähigkeiten des Betriebs- oder Dateisystems abhängig sind. Beispielsweise ist XFS ein Dateisystem, das seine Daten für eine konsistente Sicherung auslagern kann. Weitere Informationen finden Sie unter xfs_freeze.[4]

Wenn Ihr Dateisystem eine solche Freeze-Funktion nicht unterstützt, sollten Sie es deaktivieren, den Snapshot-Befehl ausführen und das Dateisystem anschließend wieder aktivieren. Dieser Vorgang lässt sich mit einem Logical Volume Manager vereinfachen, der die Ein- und Ausgabeaktivitäten unterbricht.

Da der Snapshot-Prozess im Hintergrund ausgeführt wird und die Snapshot-Erstellung sehr schnell erfolgt, müssen die Volumes, von denen Sie gerade eine Sicherung erstellen, lediglich für ein paar Sekunden deaktiviert werden. Wegen des minimalen Sicherungsfensters ist die Ausfallzeit vorhersehbar und kann geplant werden.

---

[4] https://access.redhat.com/documentation/en-US/Red_Hat_Enterprise_Linux/6/html/Storage_Administration_Guide/xfsfreeze.html

Sichern mehrerer Volumes

In einigen Fällen kann es angebracht sein, Daten über mehrere Amazon EBS-Volumes hinweg zu verteilen, wobei in der Regel zwecks Erhöhung des Durchsatzes ein Logical Volume Manager zum Einsatz kommt. Wenn Sie einen Logical Volume Manager (beispielsweise mdadm oder LVM) verwenden, ist es wichtig, die Sicherung auf dessen Ebene auszuführen und nicht auf der Ebene der zugrunde liegenden EBS-Volumes. Dadurch wird gewährleistet, dass alle Metadaten konsistent und die verschiedenen Teilkomponenten-Volumes kohärent sind.

Es gibt eine Reihe von Möglichkeiten, um dies zu erreichen. Zum Beispiel können Sie das von alestic.com[5] erstellte Skript verwenden. Die Speicherpuffer werden auf die Festplatte geschrieben. Dann wird die Ein-/Ausgabe des Dateisystems auf die Festplatte gestoppt. Anschließend wird ein Snapshot gleichzeitig für alle Volumes des RAID-Verbunds initialisiert. Nachdem der Snapshot für die Volumes eingeleitet wurde (in der Regel nach ein oder zwei Sekunden), kann das Dateisystem seine Operationen fortsetzen. Die Snapshots sollten markiert werden, sodass sie bei einer Wiederherstellung gemeinsam verwalten werden können.

Sie können Sicherungen dieser Art auch auf der Ebene des Logical Volume Manager bzw. Dateisystems durchführen. In diesen Fällen ist es unter Verwendung eines traditionellen Sicherungs-Agent möglich, Daten über das Netzwerk zu sichern. Etliche solcher Sicherungslösungen, die auf einem Agent basieren, sind im Internet und im AWS Marketplace erhältlich.[6] Bei derartiger Sicherungssoftware ist zu beachten, dass Servername und IP-Adresse einheitlich sein müssen. Am besten lässt sich Zuverlässigkeit folglich gewährleisten, wenn diese Tools gemeinsam mit Instances verwendet werden, die in einer virtuellen privaten Cloud (VPC)[7] von Amazon bereitstehen.

Alternativ kann ein Replikat der Volumes des Primärsystems in einem einzelnen großen Volume erstellt werden. Das erleichtert den Sicherungsvorgang, da lediglich ein großes Volume gesichert werden muss und die Sicherung nicht auf dem Primärsystem erfolgt. Allerdings muss zuvor festgestellt werden, ob das einzelne Volume leistungsmäßig für die Sicherung geeignet ist und ob die maximale Volume-Größe für die Anwendung ausreicht.

---

[5] https://github.com/alestic/ec2-consistent-snapshot

[6] https://aws.amazon.com/marketplace/

[7] http://aws.amazon.com/vpc/

Ansätze für Datenbanksicherungen

AWS bietet zahlreiche Optionen für Datenbanken. Sie können Ihre eigene Datenbank auf einer Amazon EC2 Instance ausführen oder eine der Optionen für verwaltete Services aus Amazon Relational Database Service [8](Amazon RDS) wählen. Wenn Sie Ihre eigene Datenbank auf einer Amazon EC2 Instance ausführen, können Sie Daten mithilfe eigenständiger Tools (z. B. MySQL[9], Oracle[10], MSSQL[11], PostgreSQL[12]) in Dateien sichern, oder Sie erstellen mithilfe der im Abschnitt „Sicherungen mit EBS- SnapShots“ genannten Methoden einen Snapshot der Volumes, in denen die Daten enthalten sind.

Verwenden von Datenbankreplikaten als Sicherungen

Bei Datenbanken, die auf RAID-Sets von Amazon EBS-Volumes ausgeführt werden, können Sie sich die Mühe zur Sicherung der primären Datenbank ersparen, indem eine Read Replica der Datenbank zu erstellen. Dies ist eine Kopie der Datenbank, die auf einer separaten Amazon EC2 Instance ausgeführt wird. Die Instance des Datenbankreplikats kann, der Quelle entsprechend, unter Verwendung mehrerer Festplatten erstellt werden. Oder Sie fassen die Daten in einem einzigen EBS-Volume zusammen. Dann haben Sie die Möglichkeit, mithilfe der im Abschnitt „Sicherungen mit EBS-Snapshots“ genannten Prozeduren einen Snapshot für das EBS-Volume zu erstellen. Dieser Ansatz wird oft für große Datenbanken verwendet, die rund um die Uhr in Betrieb sein müssen. In einem solchen Fall würde eine Sicherung so lange dauern, dass die Produktionsdatenbank für diesen Zeitraum nicht deaktiviert werden kann.

Verwenden von Amazon RDS für Sicherungen

Amazon RDS umfasst Funktionen für automatisierte Datenbanksicherungen. Amazon RDS erstellt einen Snapshot für das Volume mit Ihrer Datenbank-Instance, sodass die gesamte DB-Instance gesichert wird und nicht nur einzelne Datenbanken.

---

[8] https://aws.amazon.com/rds/

[9] http://dev.mysql.com/doc/refman/5.7/en/backup-and-recovery.html

[10] http://docs.oracle.com/cd/E11882_01/backup.112/e10642/rcmbckba.htm#BRADV8003

[11] http://msdn.microsoft.com/en-us/library/ms187510.aspx

[12] http://www.postgresql.org/docs/9.3/static/backup.html

Amazon RDS stellt zwei verschiedene Methoden zur Sicherung und Wiederherstellung Ihrer DB-Instances bereit:

- **Automatisierte Sicherungen** ermöglichen eine zeitpunktbezogene Wiederherstellung Ihrer DB-Instance. Automatisierte Sicherungen werden standardmäßig aktiviert, wenn Sie eine neuen DB-Instance erstellen. Amazon RDS führt täglich eine vollständige Sicherung Ihrer Daten in einem Zeitfenster aus, das Sie bei der Definition der DB-Instance vorgeben. Die Aufbewahrungsfrist für die automatische Sicherung kann auf maximal 35 Tage festgelegt werden. Amazon RDS verwendet diese periodischen Datensicherungen in Verbindung mit Ihren Transaktionsprotokollen, um Ihnen zu ermöglichen, Ihre DB-Instance zu jeder Sekunde innerhalb der Aufbewahrungsfrist wiederherzustellen, wobei der letztmögliche Zeitpunkt durch den Wert `LatestRestorableTime` bestimmt wird (in der Regel die letzten fünf Minuten). Sie finden diese Zeitangabe für Ihre DB-Instances mithilfe des API-Aufrufs `DescribeDBInstances` oder in der AWS Management Console auf der Registerkarte **Beschreibung** für die Datenbank.

  Beim Einleiten einer zeitpunktbezogenen Wiederherstellung werden die Transaktionsprotokolle auf die am besten geeignete tägliche Sicherung angewendet, damit Ihre DB-Instance genau auf dem von Ihnen gewünschten Stand wiederhergestellt wird.

- **DB-Snapshots** werden vom Benutzer initiiert und ermöglichen Ihnen, Ihre DB-Instance jederzeit und beliebig oft in einem bestimmten Zustand zu sichern und diesen bei Bedarf wiederherzustellen. DB-Snapshots können mithilfe der AWS Management Console oder des API-Aufrufs `CreateDBSnapshot` erstellt werden. Für diese Snapshots gilt eine unbegrenzte Aufbewahrungsdauer. Sie bleiben erhalten, bis sie von Ihnen über die Konsole oder den API-Aufruf `DeleteDBSnapshot` gelöscht werden.

Wenn Sie eine zeitpunktbezogene Wiederherstellung einer Datenbank vornehmen oder eine Datenbank aus einem DB-Snapshot wiederherstellen, wird eine neue Datenbank-Instance mit einem neuen Endpunkt erstellt. Auf diese Weise können Sie, ausgehend von einem bestimmten DB-Snapshot oder Zeitpunkt, mehrere Datenbank-Instances erstellen.

Zum Löschen der alten Datenbank-Instance verwenden Sie die AWS Management Console oder den Aufruf `DeleteDBInstance`.

Verwenden eines AMI zur Sicherung von EC2 Instances

AWS speichert ein Systemabbild in einem sog. Amazon Machine Image (AMI). Diese Abbilder bestehen aus der Vorlage für das Stamm-Volume, das zum Starten einer Instance benötigt wird. Um das Stamm-Volume als AMI zu sichern, verwenden Sie die AWS Management Console oder in der Befehlszeile die Anweisung `aws ec2 create-image`.

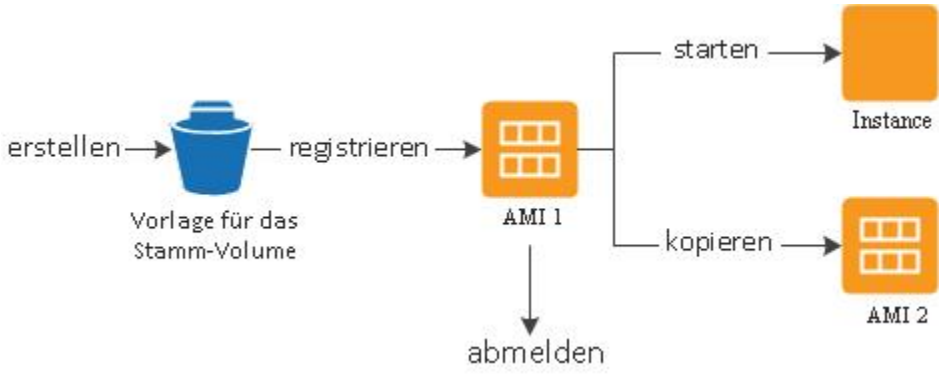

Abbildung 3: Verwenden eines AMI zum Sichern und Starten einer Instance

Ein von Ihnen registriertes AMI wird unter Verwendung von Amazon EBS-Snapshots automatisch in Ihrem Konto gespeichert. Diese Snapshots sind in Amazon S3 gespeichert und höchst beständig.

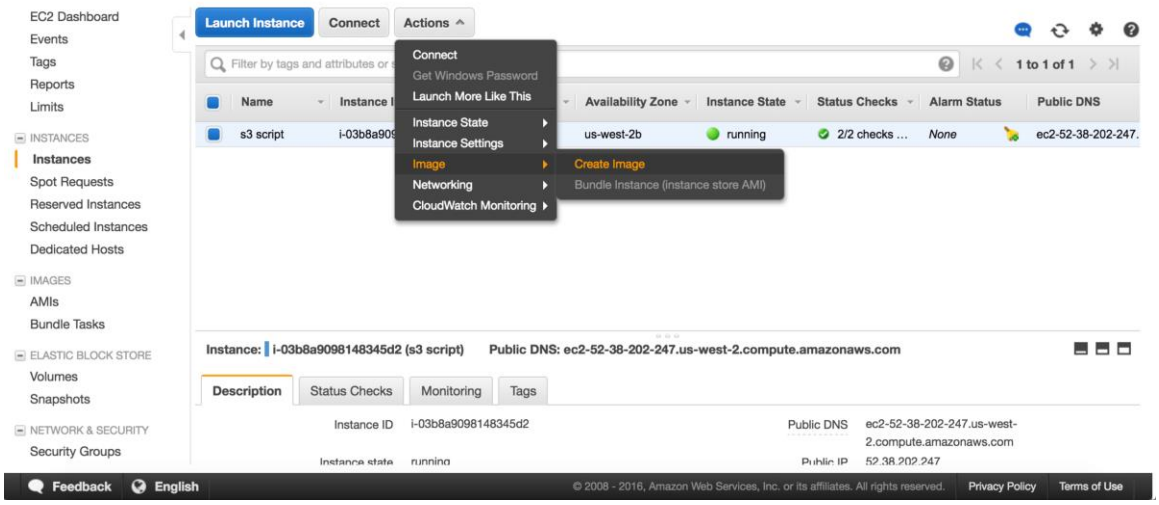

Abbildung 4: Verwenden der EC2-Konsole zum Erstellen eines Computerabbilds

Mit einem AMI Ihrer Amazon EC2 Instance können Sie die Instance erneut erstellen oder weitere Kopien der Instance starten. Es ist auch möglich, AMIs aus einer Region in eine andere zu kopieren, beispielsweise zur Migration von Anwendungen oder für eine Notfallwiederherstellung.

Lokale Infrastruktur und AWS-Infrastruktur

Dieses Szenario beschreibt ein Arbeitsumfeld ohne Cloud-Komponenten. Alle Ressourcen, darunter Webserver, Anwendungsserver, Überwachungsserver, Datenbanken und Active Directory, werden entweder im Rechenzentrum des Kunden oder in einer Colocation-Umgebung gehostet.

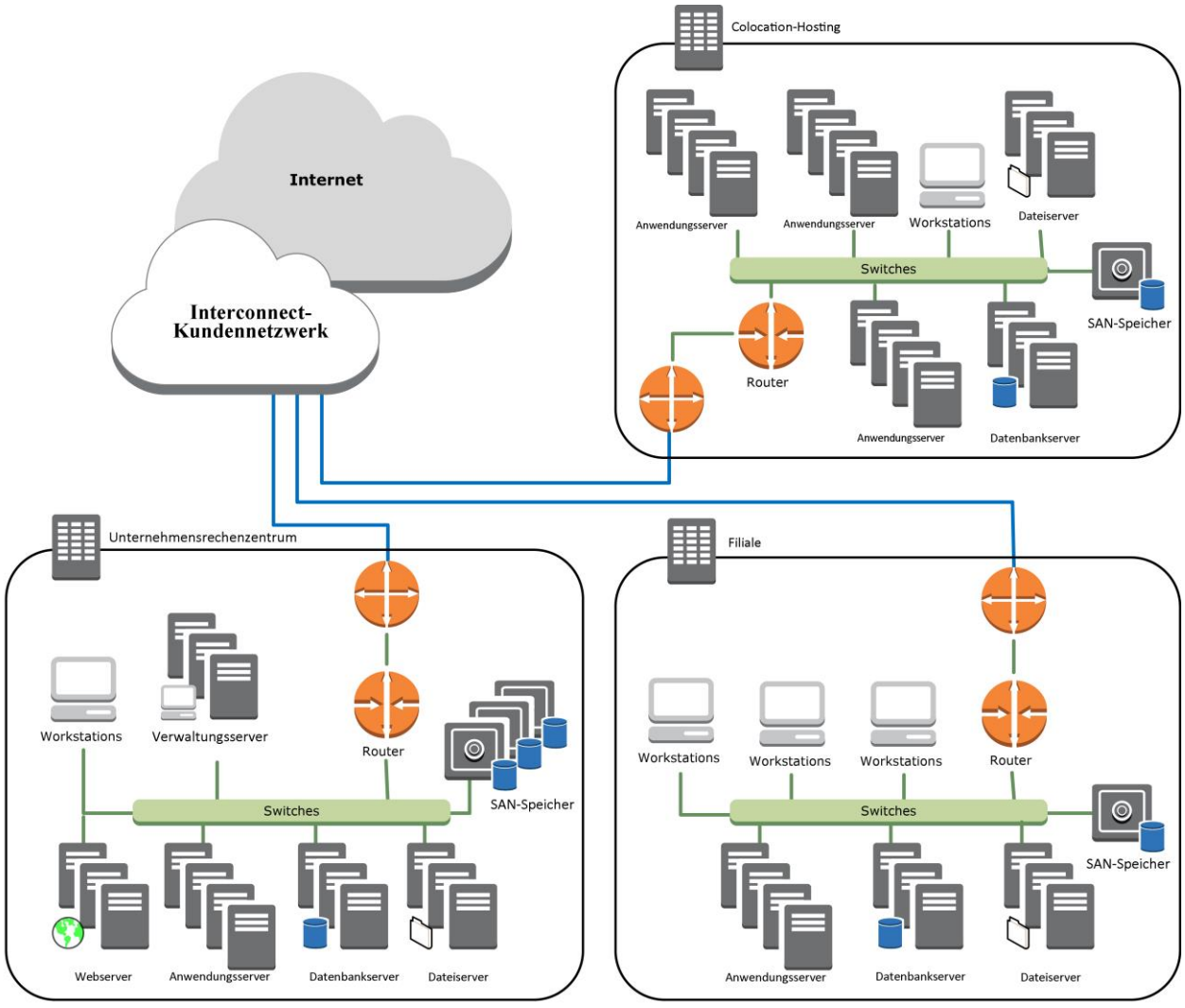

Abbildung 5: Lokale Umgebung

Durch die Verwendung der Speicherservices von AWS in diesem Szenario können Sie sich auf die Sicherungs- und Wiederherstellungsaufgaben konzentrieren. Sie müssen sich keine Gedanken mehr über die Skalierung der Speicher- oder Infrastrukturkapazität machen.

Amazon S3 und Amazon Glacier sind API-basiert und über das Internet verfügbar. Dies ermöglicht es Anbietern von Sicherungssoftware, ihre Anwendungen direkt in Speicherlösungen von AWS zu integrieren (s. folgende Abbildung).

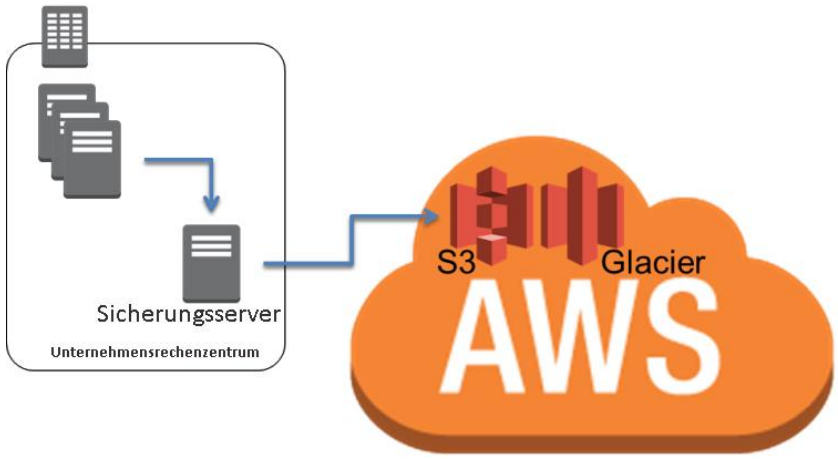

Abbildung 6: Sicherungskonnektor für Amazon S3 oder Amazon Glacier

In diesem Szenario bilden die APIs die direkte Schnittstelle zwischen der Sicherungs- bzw. Archivierungssoftware und AWS. Da die Sicherungssoftware mit AWS zusammenarbeitet, werden die auf den lokalen Servern gespeicherten Daten direkt in Amazon S3 oder Amazon Glacier gesichert.

Sollte Ihre Sicherungssoftware die AWS Cloud nicht unterstützen, stehen Ihnen die AWS Storage Gateway-Produkte zu Verfügung. AWS Storage Gateway[13] ist eine virtuelle Anwendung, die eine nahtlose und sichere Integration Ihres Rechenzentrums in die AWS-Speicherinfrastruktur ermöglicht. Mit diesem Service können Sie Daten skalierbar und kosteneffizient in der AWS Cloud speichern. Storage Gateway unterstützt Speicherprotokolle nach Branchenstandard, die reibungslos mit Ihren bestehenden Anwendungen zusammenarbeiten und alle Ihre Daten sicher und verschlüsselt in Amazon S3 oder Amazon Glacier speichern.

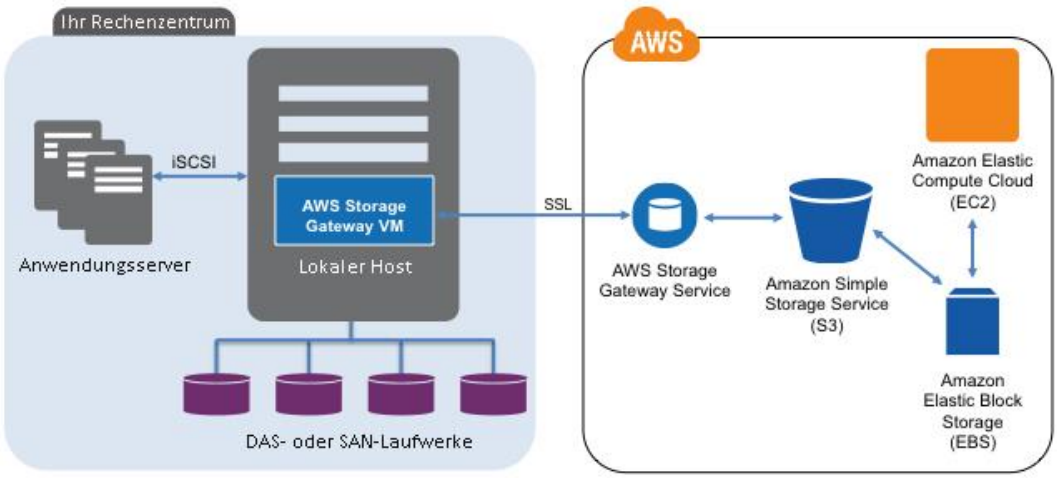

Abbildung 7: Verbindung der lokalen Umgebung mit AWS-Speicher

AWS Storage Gateway unterstützt die folgenden Konfigurationen:

- **Volume-Gateways:** Volume-Gateways bieten Cloud-basierte Speicher-Volumes, die Sie als iSCSI-Geräte (Internet Small Computer System Interface) über Ihre lokalen Anwendungsserver aktivieren können. Das Gateway unterstützt die folgenden Volume-Konfigurationen:
  - **Gateway-Cached-Volumes:** Sie speichern Ihre primären Daten in Amazon S3 und häufig aufgerufene Daten lokal. Gateway-Cached-Volumes bieten erhebliche Kosteneinsparungen bei Primärspeichern. Zudem ist keine Skalierung der lokalen Speicher erforderlich, und der schnelle Zugriff auf Daten, die häufig genutzt werden, bleibt gewährleistet.

[13] http://aws.amazon.com/storagegateway/

- **Gateway-Stored-Volumes:** Sollten Sie schnellen Zugriff auf alle Ihre Daten benötigen, können Sie Ihr lokales Daten-Gateway so konfigurieren, dass Ihre primären Daten lokal und zeitpunktbezogene Sicherungs-Snapshots dieser Daten asynchron in Amazon S3 gespeichert werden. Gateway-Stored-Volumes ermöglichen externe, dauerhafte und kostengünstige Sicherungen, die Sie lokal oder über Amazon EC2 wiederherstellen können.

- **Gateway-Virtual Tape Library (Gateway-VTL):** Mit Gateway-VTL können Sie eine Sammlung beliebig vieler virtueller Bänder anlegen. Jedes virtuelle Band kann in einer virtuellen Bandbibliothek (Virtual Tape Library, VTL) in Amazon S3 oder einem virtuellen Bandregal (Virtual Tape Shelf, VTS) in Amazon Glacier gespeichert werden. Die VTL stellt eine branchenübliche iSCSI-Schnittstelle bereit, die Ihrer Sicherungsanwendung Onlinezugriff auf die virtuellen Bänder ermöglicht. Sobald Sie nicht mehr sofort oder häufig auf die Daten eines virtuellen Bands zugreifen müssen, verschieben Sie es mit Ihrer Sicherungsanwendung aus der VTL in Ihr VTS und reduzieren auf diese Weise nochmals Ihre Speicherkosten.

Diese Gateways fungieren als Plug-and-Play-Geräte und stellen standardmäßige iSCSI-Geräte bereit, die sich in Ihr Sicherungs- oder Archivierungs-Framework integrieren lassen. Sie können die iSCSI-Geräte wie Speicherpools für Ihre Sicherungssoftware nutzen und die Gateway-VTL verwenden, um bandbasierte Sicherungen oder Archive direkt in Amazon S3 oder Amazon Glacier zu übertragen.

Mit dieser Methode speichern Sie Ihre Sicherungen und Archive automatisch extern (zur Gewährleistung der Compliance) und auf dauerhaften Medien, sodass die Komplexität und die Sicherheitsgefahren einer Bandverwaltung an externen Standorten vermieden wird.

Hybride Umgebungen

Die beiden bisher behandelten Infrastrukturimplementierungen, die reine Cloud-Lösung und die lokale Lösung, lassen sich in einem hybriden Szenario zusammenfassen, in dem das Arbeitsumfeld lokale und AWS-Infrastrukturkomponenten enthält. Ressourcen, darunter Webserver, Anwendungsserver, Überwachungsserver, Datenbanken und Active Directory, werden entweder im Rechenzentrum des Kunden oder in AWS gehostet. In der AWS Cloud ausgeführte Anwendungen sind mit den in der lokalen Umgebung des Kunden ausgeführten Anwendungen verbunden.

Dieses Szenario tritt für die Workloads in Unternehmens immer häufiger auf. Viele Unternehmen besitzen eigene Rechenzentren und nutzen AWS zur

Ausweitung der Kapazität. Diese Rechenzentren der Kunden verfügen oftmals über hochkapazitive Netzwerkverbindungen zu AWS. Beispielsweise können Sie mit AWS Direct Connect[14] eine private, dedizierte Verbindung zwischen Ihrer lokalen Umgebung und AWS herstellen. Diese bietet die erforderliche Bandbreite und konsistente Latenz, um schützenswerte Daten in die Cloud hochzuladen, und eine konsistente Leistung und Latenz für hybride Workloads.

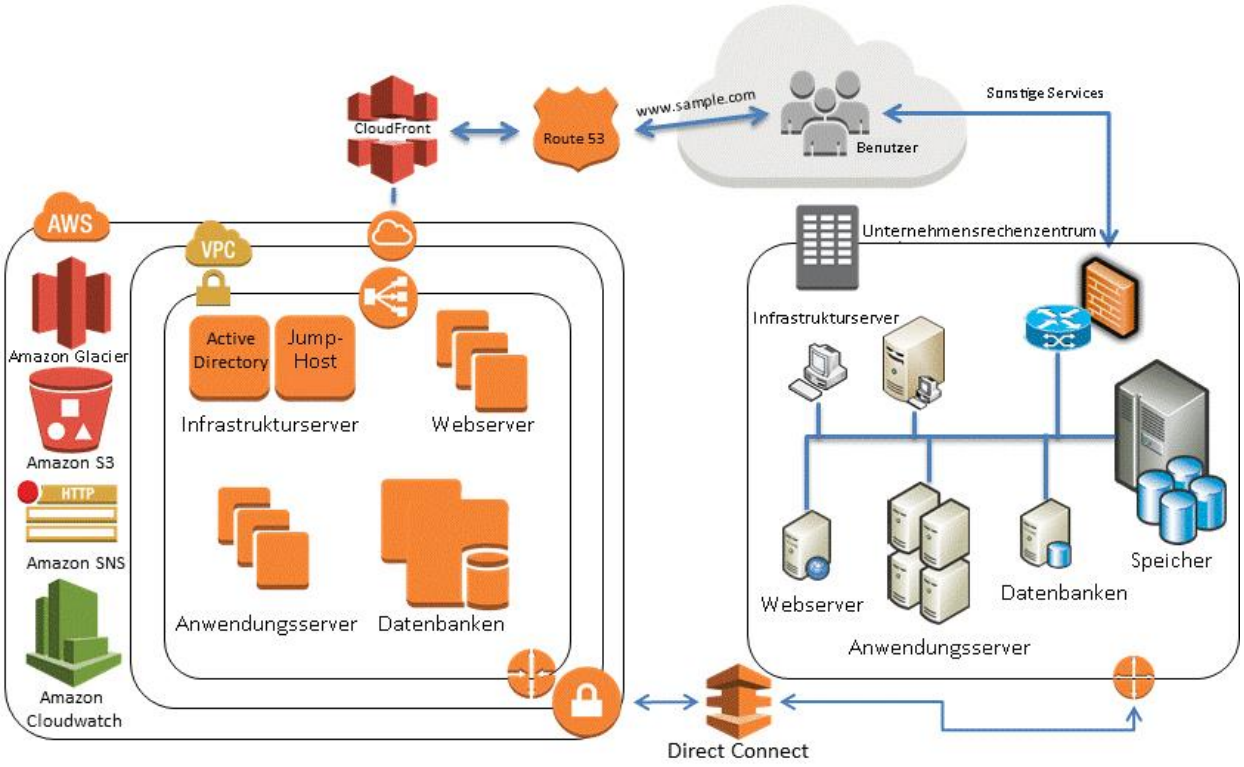

Abbildung 8: Ein Szenario mit hybrider Infrastruktur

Gute Lösungen zur Datensicherung verwenden üblicherweise eine Kombination der Methoden, die in den Abschnitten für reine Cloud-Lösungen und für lokale Lösungen beschrieben wurden.

[14] http://aws.amazon.com/directconnect/

Sichern von AWS-basierten Anwendungen in Ihrem Rechenzentrum

Wenn Sie bereits über ein bestehendes Framework zum Sichern von Daten auf Ihren lokalen Servern verfügen, ist es einfach, dieses Framework über eine VPN-Verbindung oder AWS Direct Connect um Ihre AWS-Ressourcen zu erweitern. Sie installieren dazu lediglich den Sicherungs-Agent auf den Amazon EC2 Instances und sichern diese gemäß Ihren Sicherungsrichtlinien.

Migrieren der Sicherungsverwaltung in die Cloud zur Gewährleistung der Verfügbarkeit

Ihre Sicherungsarchitektur enthält möglicherweise einen Master-Sicherungsserver und lokal einen oder mehrere Medien- bzw. Speicherserver sowie die zu sichernden Services. In diesem Fall sollten Sie in Erwägung ziehen, den Master-Sicherungsserver in eine Amazon EC2 Instance zu verschieben, um ihn vor lokalen Gefahren zu schützen und eine hochverfügbare Sicherungsinfrastruktur zu garantieren.

Um den Sicherungsdatenverkehr zu verwalten, bietet sich an, einen oder mehrere Medienserver auf Amazon EC2 Instances zu erstellen. Mit solchen Medienservern sparen Sie das Geld für Internetübertragen und erhöhen, sofern Sie auf S3 oder Amazon Glacier sichern, die gesamte Sicherungs- und Wiederherstellungsleistung.

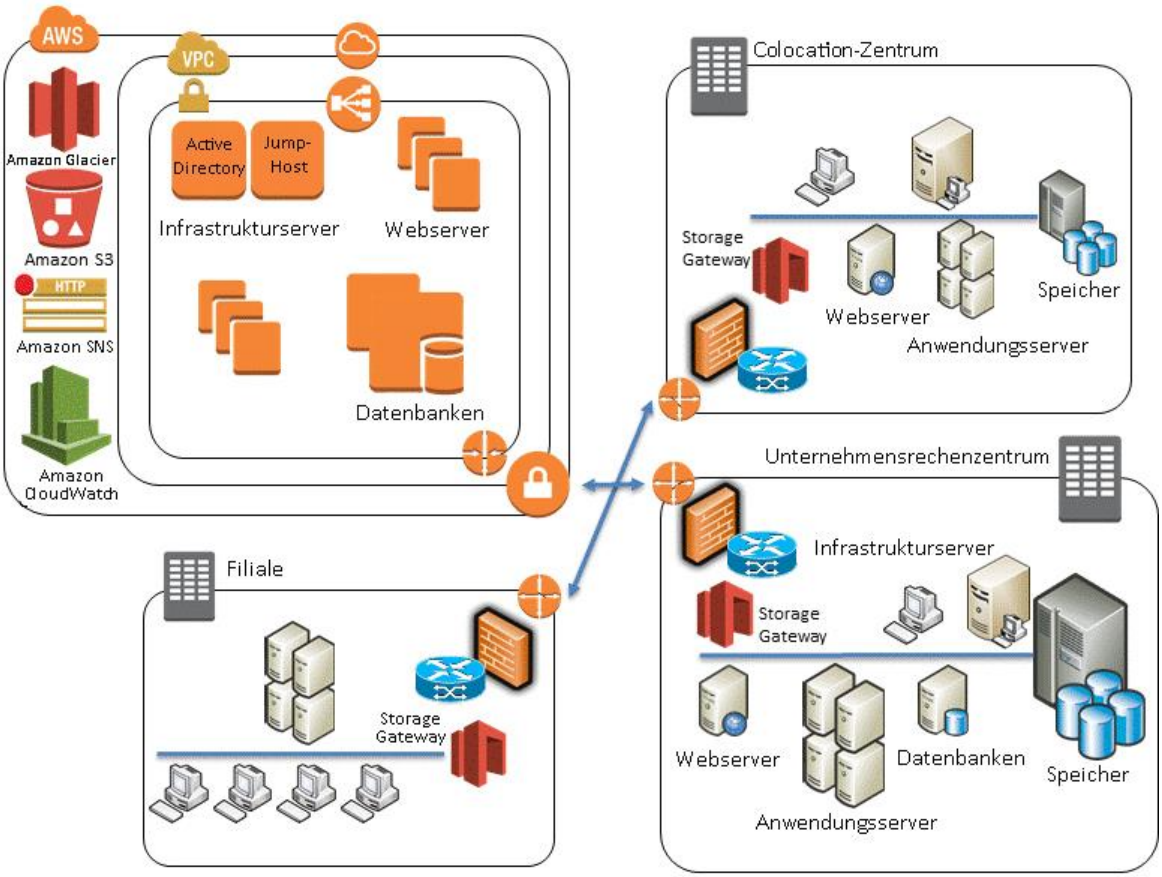

Abbildung 9: Verwendung von Gateways in einem hybriden Szenario

Beispiel für ein hybrides Szenario

Stellen Sie sich vor, Sie verwalten eine Umgebung, in der Sie verschiedene Amazon EC2 Instances, eigenständige Server, virtuelle Maschinen und Datenbanken sichern. Diese Umgebung umfasst 1000 Server, und Sie sichern das Betriebssystem, Dateidaten, Abbilder virtueller Maschinen und Datenbanken. Zu sichern sind 20 Datenbanken, darunter MySQL, Microsoft SQL Server und Oracle.

Ihre Sicherungssoftware verfügt über Agents, die das Betriebssystem, die Abbilder der virtuellen Maschinen, die Daten-Volumes, die SQL Server-Datenbanken und die Oracle-Datenbanken sichern (letztere mithilfe von RMAN). Bei Anwendungen wie MySQL, für die Ihre Sicherungssoftware keinen Agent besitzt, verwenden Sie das Hilfsprogramm mysqldump, um auf der Festplatte eine Datei mit einem Datenbankauszug zu erstellen, die dann von den standardmäßigen Sicherungs-Agents verarbeitet wird.

Um diese Umgebung zu sichern, verfügt Ihre Sicherungssoftware höchstwahrscheinlich über einen globalen Katalogserver oder Masterserver zur Steuerung der Sicherung, Archivierung und Wiederherstellung sowie der verschiedenen Medienserver, die mit festplattenbasierten Speichern, Linear Tape Open (LTO)-Bandlaufwerken und AWS-Speicherservices verbunden sind.

Der einfachste Weg, um Ihre Sicherungslösung mit AWS-Speicherservices zu erweitern, ist die Verwendung der Funktionen, mit denen der Anbieter Ihrer Sicherungssoftware Amazon S3 oder Amazon Glacier unterstützt. Wir empfehlen Ihnen, sich diese Optionen für Integration und Anbindung von Ihrem Anbieter erklären zu lassen. In unserem Partnerverzeichnis[15] finden Sie die Anbieter von Sicherungssoftware, die mit AWS zusammenarbeiten.

Sollte Ihre Sicherungssoftware keinen Cloud-Speicher zur Datensicherung und Datenarchivierung unterstützen, können Sie ein Storage Gateway-Gerät, beispielsweise eine Brücke, als Verbindung zwischen der Sicherungssoftware und Amazon S3 oder Amazon Glacier verwenden.

Gateway-Lösungen sind von vielen Drittanbietern erhältlich. Sie können diese Lücke auch mit virtuellen Appliances von AWS Storage Gateway schließen, die generische Techniken wie iSCSI-basierte Volumes und virtuelle Bandbibliotheken bereitstellen. Diese Konfiguration erfordert einen unterstützten Hypervisor (VMware oder Microsoft Hyper-V) sowie lokalen Speicher zum Hosten der Appliance.

Archivieren von Daten mit AWS

Daten, die aus Compliance- oder betrieblichen Gründen zu erhalten sind, werden in der Regel archiviert. Sicherungen werden üblicherweise ausgeführt, um für kurze Zeit eine Kopie der Produktionsdaten zu besitzen, sodass beschädigte oder verloren gegangene Daten wiederherstellbar sind. Dagegen werden bei einer Archivierung alle Kopien der Daten so lange aufbewahrt, wie es die Aufbewahrungsrichtlinien vorschreiben.

[15] http://www.aws-partner-directory.com/PartnerDirectory/PartnerSearch?type=ISV

Ein gutes Archiv erfüllt die folgenden Kriterien:

- Dauerhaftigkeit der Daten zur Sicherstellung langfristiger Integrität
- Datensicherheit
- Einfache Wiederherstellungsmöglichkeiten
- Kostengünstig

Eine weitere regulatorische oder Compliance-Anforderung kann sein, dass die Datenspeicher unveränderlich sind.

Amazon Glacier bietet Archivierung zu geringen Kosten, Verschlüsselung der gespeicherten Daten, eine Dauerhaftigkeit mit 11 Neunen sowie unbegrenzte Kapazität.

Die geeignete Wahl für Anwendungsfälle, die eine schnelle Datenwiederherstellung erfordern, ist Amazon S3 Standard - Infrequent Access. Amazon Glacier ist eine gute Wahl für Anwendungsfälle, in denen selten auf Daten zugegriffen wird und in denen eine Abrufzeit von mehreren Stunden annehmbar ist.

Objekte können in Amazon Glacier entweder durch S3-Lebenszyklusregeln oder die Amazon Glacier-API eingebunden werden. Die Vault Lock-Funktion in Amazon Glacier ermöglicht Ihnen, Compliance-Kontrollen für einzelne Amazon Glacier-Speicher mit einer entsprechenden Richtlinie auf einfache Weise bereitzustellen und durchzusetzen. Sie können in einer Vault Lock-Richtlinie beispielsweise eine WORM-Bedingung (write once, read many) festlegen und die Richtlinie vor zukünftigen Änderungen schützen. Weitere Informationen finden Sie unter Amazon Glacier.

Schützen von Sicherungsdaten in AWS

Datensicherheit ist ein allgemeines Anliegen. AWS nimmt Sicherheit sehr ernst. Sie ist die Grundlage für jeden Service, den wir bereitstellen. Speicherservices wie Amazon S3 bieten leistungsstarke Funktionen für die Zugriffskontrolle und die Verschlüsselung von Daten im Ruhezustand oder während der Übertragung. Alle API-Endpunkte in Amazon S3 und Amazon Glacier unterstützen SSL-Verschlüsselung für Daten bei der Übertragung. Amazon Glacier verschlüsselt standardmäßig alle Daten im Ruhezustand. Mit Amazon S3 können Kunden die serverseitige Verschlüsselung für Objekte im Ruhezustand wählen, indem sie AWS die Schlüsselverwaltung überlassen, eigene Schlüssel bei Hochladen eines Objekts bereitstellen oder AWS Key Management Service (AWS KMS) [16]für die Schlüssel verwenden. Alternativ haben Kunden die Möglichkeit, ihre Daten immer in verschlüsselter Form in AWS hochzuladen. Weitere Informationen finden Sie unter Amazon Web Services: Overview of Security Processes.

Zusammenfassung

Gartner hat AWS als führenden Anbieter für öffentliche Cloud-Speicherservices anerkannt.[17] AWS ist gut aufgestellt, um Organisationen dabei zu helfen, ihre Arbeitslasten auf Cloud-basierte Plattformen zu übertragen und so die neue Generation von Sicherungslösungen anzuwenden. AWS bietet kosteneffiziente und skalierbare Lösungen, sodass Organisationen den Anforderungen an die Sicherung und Archivierung von Daten gerecht werden können. Diese Services lassen sich ohne Probleme in bestehende Technologien integrieren.

[16] http://docs.aws.amazon.com/AmazonS3/latest/dev/UsingKMSEncryption.html

[17] http://www.gartner.com/technology/reprints.do?id=1-1WWKTQ3&ct=140709&st=sb

Mitwirkende

Dieses Dokument ist unter der Mitarbeit folgender Personen entstanden:

- Pawan Agnihotri, Solutions Architect, Amazon Web Services
- Lee Kear, Solutions Architect, Amazon Web Services
- Peter Levett, Solutions Architect, Amazon Web Services

Am Dokument vorgenommene Änderungen

Aktualisiert Mai 2016