AWS Well-Architected Framework

Säule „Nachhaltigkeit“

Copyright © 2024 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Säule „Nachhaltigkeit“: AWS Well-Architected Framework

Copyright © 2024 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Die Handelsmarken und Handelsaufmachung von Amazon dürfen nicht in einer Weise in Verbindung mit nicht von Amazon stammenden Produkten oder Services verwendet werden, durch die Kunden irregeführt werden könnten oder Amazon in schlechtem Licht dargestellt oder diskreditiert werden könnte. Alle anderen Marken, die nicht im Besitz von Amazon sind, gehören den jeweiligen Besitzern, die möglicherweise mit Amazon verbunden sind oder von Amazon gesponsert werden.

Table of Contents

Zusammenfassung und Einführung i
Einführung 1
Cloud-Nachhaltigkeit 2
Das Modell der geteilten Verantwortung 3
Nachhaltigkeit der Cloud 4
Nachhaltigkeit in der Cloud 4
Nachhaltigkeit durch die Cloud 4
Designprinzipien für Nachhaltigkeit in der Cloud 5
Verbesserungsprozess 7
Beispielszenario 8
Identifizieren von Verbesserungszielen 8
Ressourcen 9
Bewerten spezifischer Verbesserungen 9
Proxy-Metriken 9
Geschäftsmetriken 10
Wichtige Leistungskennzahlen 10
Schätzen der Verbesserung 11
Bewerten von Verbesserungen 11
Priorisieren und Planen von Verbesserungen 13
Testen und Validieren von Verbesserungen 14
Bereitstellen von Änderungen für die Produktion 15
Messen von Ergebnissen und Wiederholen von Erfolgen 16
Nachhaltigkeit als nichtfunktionale Anforderung 19
Bewährte Methoden in Bezug auf Nachhaltigkeit in der Cloud 21
Auswahl der Region 21
SUS01-BP01 Auswählen der Region auf Grundlage von Unternehmensanforderungen und Nachhaltigkeitszielen 21
Ausrichtung am Bedarf 23
SUS02-BP01 Dynamisches Skalieren der Workload-Infrastruktur 24
SUS02-BP02 Ausrichten von SLAs an Nachhaltigkeitszielen 28
SUS02-BP03 Beenden der Erstellung und Wartung nicht verwendeter Komponenten 30
SUS02-BP04 Optimieren der geografischen Platzierung von Workloads auf der Grundlage ihrer Netzwerkanforderungen 32

SUS02-BP05 Optimieren von Ressourcen für Teammitglieder im Hinblick auf die ausgeführten Aktivitäten ... 35
SUS02-BP06 Implementierung von Pufferung oder Drosselung, um die Bedarfskurve zu verflachen ... 37
Software und Architektur ... 40
SUS03-BP01 Optimieren von Software und Architektur für asynchrone und geplante Aufträge ... 40
SUS03-BP02 Entfernen oder Faktorwechsel von Workload-Komponenten mit geringer oder keiner Nutzung ... 43
SUS03-BP03 Optimieren von Codebereichen, die die meiste Zeit oder die meisten Ressourcen verbrauchen ... 45
SUS03-BP04 Optimieren der Auswirkungen auf Geräte und Ausrüstung von Kunden ... 48
SUS03-BP05 Verwenden von Softwaremustern und Architekturen, die Datenzugriffs- und Speichermuster optimal unterstützen ... 50
Datenverwaltung ... 53
SUS04-BP01 Implementieren einer Richtlinie für die Klassifizierung von Daten ... 53
SUS04-BP02 Verwenden von Technologien, die Datenzugriff und Speichermuster unterstützen ... 55
SUS04-BP03 Verwalten des Lebenszyklus von Datensätzen mithilfe von Richtlinien ... 60
SUS04-BP04 Verwendung von Elastizität und Automatisierung zur Erweiterung des Block-Speichers oder des Dateisystems ... 62
SUS04-BP05 Entfernen nicht benötigter oder redundanter Daten ... 64
SUS04-BP06 Verwenden geteilter Dateisysteme oder Objektspeicher für den Zugriff auf allgemeine Daten ... 67
SUS04-BP07 Minimieren von Datenübertragungen zwischen Netzwerken ... 69
SUS04-BP08 Sichern von Daten nur in dem Fall, wenn ihre erneute Erstellung schwierig ist ... 71
Hardware und Services ... 73
SUS05-BP01 Verwenden der geringstmöglichen Menge an Hardware zur Erfüllung Ihrer Anforderungen ... 73
SUS05-BP02 Verwenden von Instance-Typen mit den geringsten Auswirkungen ... 76
SUS05-BP03 Verwenden verwalteter Services ... 79
SUS05-BP04 Optimieren der Nutzung von hardwarebasierten Datenverarbeitungsbeschleunigern ... 81
Prozess und Kultur ... 83

SUS06-BP01 Einführen von Methoden, die schnelle Verbesserungen für die Nachhaltigkeit ermöglichen 84
SUS06-BP02 Konstantes Aktualisieren Ihrer Workload 86
SUS06-BP03 Höhere Auslastung von Entwicklungsumgebungen 88
SUS06-BP04 Verwenden verwalteter Gerätefarmen für Tests 90
Schlussfolgerung 93
Mitwirkende 94
Weitere Informationen 95
Dokumentversionen 96
Hinweise 97
AWS-Glossar 98

Säule „Nachhaltigkeit“ – AWS Well-Architected Framework

Veröffentlichungsdatum: 27. Juni 2024 (Dokumentversionen)

In diesem Whitepaper geht es speziell um die Säule „Nachhaltigkeit“ des Well-Architected Framework von Amazon Web Services (AWS). Es enthält Designprinzipien, Betriebsanleitungen, bewährte Methoden, potenzielle Kompromisse und Verbesserungspläne, die Sie nutzen können, um die Nachhaltigkeitsziele für Ihre AWS-Workloads zu erreichen.

Einführung

Das AWS Well-Architected Framework unterstützt Sie dabei, die Vor- und Nachteile der Entscheidungen nachzuvollziehen, die Sie beim Erstellen von Workloads in AWS treffen. Das Framework hilft Ihnen, bewährte Architekturmethoden für den Entwurf und Betrieb zuverlässiger, sicherer, effizienter, kostengünstiger und nachhaltiger Workloads in der AWS Cloud zu ermitteln. Es bietet Ihnen die Möglichkeit, Ihre Architekturen konsistent auf die Einhaltung bewährter Methoden zu prüfen und Verbesserungspotenzial zu identifizieren. Mit Well-Architected-Workloads können Sie Ihre Geschäftsergebnisse wesentlich besser unterstützen.

Das Framework basiert auf den folgenden sechs Säulen:

- Operative Exzellenz
- Sicherheit
- Zuverlässigkeit
- Leistungseffizienz
- Kostenoptimierung
- Nachhaltigkeit

Dieses Dokument konzentriert sich auf die Säule der Nachhaltigkeit und dabei insbesondere auf ökologische Nachhaltigkeit. Es richtet sich an Nutzer in technologischen Rollen, z. B. CTOs (Chief Technology Officers), Architekten, Entwickler und Mitglieder von Betriebsteams.

Nach der Lektüre dieses Dokuments kennen Sie die aktuellen AWS-Empfehlungen und -Strategien, die Sie bei der Entwicklung von Cloud-Architekturen mit Blick auf die Nachhaltigkeit anwenden können. Wenn Sie die in diesem Dokument beschriebenen Praktiken anwenden, können Sie Architekturen aufbauen, die die Effizienz maximieren und Verschwendung reduzieren.

Cloud-Nachhaltigkeit

Beim Thema Nachhaltigkeit geht es um die langfristigen ökologischen, wirtschaftlichen und gesellschaftlichen Auswirkungen Ihrer geschäftlichen Aktivitäten. Die United Nations World Commission on Environment and Development definiert nachhaltige Entwicklung als „Entwicklung, die die Bedürfnisse der Gegenwart erfüllt, ohne die Möglichkeiten künftiger Generationen zu gefährden, ihre eigenen Bedürfnisse zu erfüllen“. Die Tätigkeit Ihres Unternehmens oder Ihrer Organisation kann negative Auswirkungen auf die Umwelt haben, wie etwa direkte oder indirekte Kohlendioxidemissionen, nicht-recycelbare Abfälle oder Beeinträchtigungen gemeinsam genutzter Ressourcen, etwa sauberen Wassers.

Bei der Erstellung von Cloud-Workloads bestehen nachhaltige Praktiken darin, die Auswirkungen der genutzten Services zu verstehen, diese über den gesamten Workload-Lebenszyklus hinweg zu quantifizieren, sowie konzeptionelle Grundsätze und bewährte Methoden einzusetzen, die dabei helfen, diese Auswirkungen zu reduzieren. In diesem Dokument geht es um Auswirkungen auf die Umwelt, insbesondere um Energieverbrauch und -effizienz, da diese wichtige Faktoren für Architekten sind, die ihre direkten Aktionen zur Reduzierung des Ressourcenverbrauchs beeinflussen.

Wenn Sie sich auf die Umweltauswirkungen konzentrieren, sollten Sie verstehen, wie diese Auswirkungen typischerweise verbucht werden und welche Folgen dies für die Emissionsbilanzierung Ihrer Organisation hat. Das Greenhouse Gas Protocol gliedert die Kohlenstoffemissionen in die folgenden Bereiche, zusammen mit relevanten Emissionsbeispielen innerhalb jedes Bereichs für einen Cloud-Anbieter wie AWS:

- Umfang 1: Alle direkten Emissionen aus den Aktivitäten einer Organisation oder unter deren Kontrolle. Beispiel: die Verbrennung von Treibstoff in Notstromgeneratoren von Rechenzentren.
- Umfang 2: Indirekte Emissionen durch Strom, der für den Betrieb von Rechenzentren und anderen Einrichtungen gekauft und verwendet wird. Beispiel: Emissionen aus der kommerziellen Stromerzeugung.
- Umfang 3: Alle anderen indirekten Emissionen aus Aktivitäten einer Organisation aus Quellen, die sie nicht kontrollieren kann. AWS-Beispiele sind Emissionen im Zusammenhang mit dem Bau von Rechenzentren sowie der Herstellung und dem Transport von IT-Hardware, die in Rechenzentren eingesetzt wird.

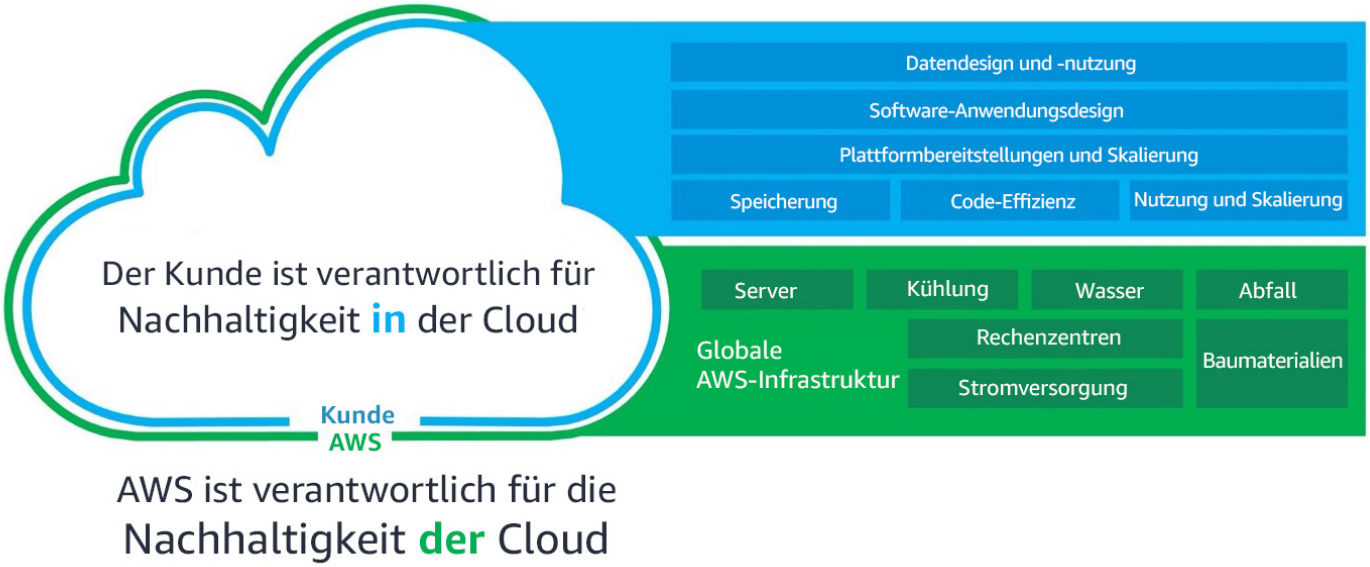
Aus der AWS-Kundenperspektive werden die Emissionen, die durch Ihre in AWS laufenden Workloads entstehen, als indirekte Emissionen verbucht und sind Teil Ihrer Umfang-3-Emissionen. Jede bereitgestellte Workload erzeugt einen Bruchteil der AWS-Gesamtemissionen aus jedem der vorherigen Bereiche. Die tatsächliche Menge variiert je nach Workload und hängt von mehreren Faktoren ab, darunter die verwendeten AWS-Services, der Energieverbrauch dieser Services, die CO_2-Intensität der Stromnetze, die die AWS-Rechenzentren versorgen, in denen sie betrieben werden, und die AWS-Beschaffung erneuerbarer Energie.

In diesem Dokument wird zunächst ein Modell der gemeinsamen Verantwortung für ökologische Nachhaltigkeit beschrieben. Anschließend werden bewährte Architekturverfahren vorgestellt, mit denen Sie die Auswirkungen Ihrer Workloads minimieren können, indem Sie die Gesamtressourcen, die für den Betrieb in AWS-Rechenzentren erforderlich sind, reduzieren.

Das Modell der geteilten Verantwortung

Die ökologische Nachhaltigkeit ist eine gemeinsame Verantwortung von Kunden und AWS.

- AWS ist verantwortlich für die Optimierung der Nachhaltigkeit der Cloud – die Bereitstellung effizienter, gemeinsam genutzter Infrastrukturen, der verantwortungsvolle Umgang mit Wasser und die Beschaffung erneuerbarer Energie.
- Kunden sind verantwortlich für Nachhaltigkeit in der Cloud – Optimierung von Workloads und Ressourcennutzung und Minimierung der Gesamtressourcen, die für Ihre Workloads bereitgestellt werden müssen.



Modell der geteilten Verantwortung

Nachhaltigkeit der Cloud

Cloud-Anbieter haben einen geringeren CO_2-Fußabdruck und sind energieeffizienter als die typischen On-Premises-Alternativen, da sie in effiziente Strom- und Kühltechnologie investieren, energieeffiziente Serverpopulationen betreiben und hohe Serverauslastungsraten erzielen. Cloud-Workloads reduzieren die Auswirkungen, indem sie gemeinsame Ressourcen wie Netzwerke, Strom, Kühlung und physische Einrichtungen nutzen. Sie können Ihre Cloud-Workloads zu effizienteren Technologien migrieren, sobald diese verfügbar werden, und Cloud-basierte Services verwenden, um Ihre Workloads nachhaltiger auszuführen.

Ressourcen

- The Carbon Reduction Opportunity of Moving to Amazon Web Services
- AWS ermöglicht nachhaltige Lösungen

Nachhaltigkeit in der Cloud

Nachhaltigkeit in der Cloud ist ein kontinuierliches Bestreben, das sich in erster Linie auf die Reduzierung des Energieverbrauchs und die Effizienz aller Komponenten einer Workload konzentriert. Dazu muss der maximale Nutzen aus den bereitgestellten Ressourcen gezogen und die insgesamt erforderlichen Ressourcen müssen minimiert werden. Diese Bemühung kann von der anfänglichen Auswahl einer effizienten Programmiersprache, der Einführung moderner Algorithmen, der Nutzung effizienter Datenspeichertechniken, der Bereitstellung einer korrekt dimensionierten und effizienten Recheninfrastruktur bis hin zur Minimierung der Anforderungen an leistungsstarke Endbenutzerhardware reichen.

Nachhaltigkeit durch die Cloud

Sie können nicht nur die Auswirkungen der von Ihnen bereitgestellten Workloads minimieren, sondern auch die AWS Cloud zur Ausführung von Workloads nutzen, die zur Unterstützung Ihrer umfassenderen Nachhaltigkeitsprobleme entwickelt wurden. Beispiele für diese Herausforderungen sind die Verringerung der CO_2-Emissionen, die Senkung des Energieverbrauchs, das Recycling von Wasser oder die Reduzierung von Abfall in anderen Bereichen Ihres Unternehmens oder Ihrer Organisation.

Nachhaltigkeit durch die Cloud bedeutet, dass Sie AWS-Technologie nutzen, um ein breiteres Nachhaltigkeitsproblem zu lösen. Sie können zum Beispiel einen Machine-Learning-Service wie Amazon Monitron nutzen, um abnormales Verhalten bei Industriemaschinen zu erkennen. Anhand dieser Erkennungsdaten können Sie eine vorbeugende Wartung durchführen, um das Risiko von Umweltschäden durch unerwartete Geräteausfälle zu verringern und sicherzustellen, dass die Maschinen weiterhin mit höchster Effizienz arbeiten.

Designprinzipien für Nachhaltigkeit in der Cloud

Wenden Sie diese Designprinzipien an, wenn Sie Ihre Cloud-Workloads entwickeln, um die Nachhaltigkeit zu maximieren und die Auswirkungen zu minimieren.

- Verstehen Ihrer Auswirkungen: Messen Sie die Auswirkungen Ihrer Cloud-Workloads und modellieren Sie diese Auswirkungen für die Zukunft. berücksichtigen Sie dabei alle relevanten Faktoren, darunter Auswirkungen durch die Verwendung Ihrer Produkte durch Kunden sowie solche durch deren Außerbetriebnahme und Entsorgung. Vergleichen Sie den produktiven Output mit den Gesamtauswirkungen Ihrer Cloud-Workloads, indem Sie die für jede Arbeitseinheit erforderlichen Ressourcen und die damit verbundenen Emissionen ermitteln. Anhand dieser Daten können Sie Leistungskennzahlen (KPIs) einrichten, Möglichkeiten zur Verbesserung der Produktivität bei gleichzeitiger Reduzierung der Auswirkungen finden und berechnen, wie sich vorgeschlagene Änderungen im Zeitverlauf auswirken werden.
- Festlegen von Nachhaltigkeitszielen: Formulieren Sie für alle Cloud-Workloads langfristige Nachhaltigkeitsziele wie etwa die Reduzierung der pro Transaktion erforderlichen Datenverarbeitungs- und Speicherressourcen. Modellieren Sie den ROI von Verbesserungen in Bezug auf die Nachhaltigkeit vorhandener Workloads. Stellen Sie den Besitzern die nötigen Ressourcen zur Verfügung, um in Nachhaltigkeitsziele investieren zu können. Planen Sie wachstumsorientiert und gestalten Sie Ihre Workloads so, dass das Wachstum mit geringeren Auswirkungen einhergeht – gemessen in einer sinnvollen Einheit, etwa pro Benutzer oder pro Transaktion. Ziele helfen Ihnen, die allgemeinen Nachhaltigkeitsziele Ihres Unternehmens oder Ihrer Organisation zu erreichen, Rückschritte zu identifizieren und Bereiche mit Verbesserungsmöglichkeiten zu priorisieren.
- Maximieren der Auslastung: Sorgen Sie für Workloads angemessenen Umfangs und nutzen Sie effiziente Designprinzipien, um eine hohe Auslastung zu gewährleisten und die Energieeffizienz der zugrunde liegenden Hardware so zu maximieren. Zwei Hosts mit 30 % Auslastung sind aufgrund des grundlegenden Energieverbrauchs pro Host weniger effizient als ein Host mit 60 % Auslastung. Gleichzeitig sollten Sie nicht genutzte Ressourcen, Verarbeitungsvorgänge und

Speicher beseitigen oder minimieren, um den Gesamtenergieverbrauch für Ihre Workloads zu senken.

- Antizipieren und Einführen neuer und effizienterer Hardware- und Software-Angebote: Unterstützen Sie die Verbesserungen, die Ihre Partner und Lieferanten in früheren Prozessphasen vornehmen, um die Auswirkungen Ihrer Cloud-Workloads zu reduzieren. Achten Sie stets auf neue und effizientere Hardware- und Software-Angebote. Planen Sie für Flexibilität, damit neue effiziente Technologien schnell eingeführt werden können.
- Verwenden von verwalteten Services: Die gemeinsame Nutzung von Services über eine breite Kundenbasis hinweg hilft dabei, die Ressourcennutzung zu maximieren und dadurch den Umfang der Infrastruktur zu verringern, der für die Unterstützung Ihrer Cloud-Workloads erforderlich ist. So können Kunden die Auswirkungen allgemeiner Rechenzentrumskomponenten wie Energieversorgung und Netzwerk teilen, indem sie Workloads zur AWS Cloud migrieren und verwaltete Services einführen, z. B. AWS Fargate für Serverless-Container. Dabei kann AWS skalierbar ausgeführt werden und ist für einen effizienten Betrieb verantwortlich. Verwenden Sie verwaltete Services, die dabei helfen können, Ihre Auswirkungen zu verringern, wie etwa die automatische Verschiebung selten genutzter Daten in „kalte“ Speicher mit Amazon S3 Lebenszykluskonfigurationen oder Amazon EC2 Auto Scaling, um Ihre Kapazitäten an die jeweiligen Anforderungen anzupassen.
- Reduzieren der nachgelagerten Auswirkungen Ihrer Cloud-Workloads: Senken Sie den Energie- oder Ressourcenverbrauch für die Nutzung Ihrer Services. Reduzieren oder beseitigen Sie die Erfordernis für Kunden, ihre Geräte zu aktualisieren, um Ihre Services nutzen zu können. Verwenden Sie in Ihren Tests Gerätefarmen, um die zu erwartenden Auswirkungen zu verstehen, und führen Sie Tests mit Kunden durch, um die tatsächlichen Auswirkungen der Nutzung Ihrer Services zu erkennen.

Verbesserungsprozess

Zur Verbesserung der Architektur gehört es, zu verstehen, was Sie haben und was Sie tun können, um Verbesserungen zu erzielen. Außerdem sollten Sie Zielbereiche für Verbesserungen auswählen, Verbesserungsmöglichkeiten testen, jene übernehmen, die sich dabei als erfolgreich erweisen, Ihren Erfolg quantifizieren und das weitergeben, was Sie gelernt haben, damit dies anderswo repliziert werden kann. Anschließend beginnt dieser Zyklus von neuem.

Zu den Zielen für Verbesserungen kann Folgendes gehören:

- Beseitigung von Verschwendung, geringer Nutzung und ungenutzten oder ungenutzten Ressourcen
- Maximierung des Werts der verbrauchten Ressourcen

> **Note**
>
> Nutzen Sie alle Ihnen zur Verfügung stehenden Ressourcen und erledigen Sie die gleiche Arbeit mit einem Minimum an Ressourcen.

Beseitigen Sie in den frühen Phasen der Optimierung zunächst Bereiche mit Verschwendung oder geringer Auslastung und gehen Sie dann zu gezielteren Optimierungen über, die Ihrer spezifischen Workload entsprechen.

Überwachen Sie die Auswirkungen auf den Ressourcenverbrauch im Zeitverlauf. Ermitteln Sie, wo akkumulierte Änderungen zu ineffizientem Ressourcenverbrauch oder zu deutlichen Anstiegen des Verbrauchs führen. Stellen Sie fest, welche Verbesserungen im Hinblick auf diese Änderungen beim Ressourcenverbrauch erforderlich sind, und implementieren Sie die von Ihnen priorisierten Verbesserungen.

Die folgenden Schritte sind als iterativer Prozess konzipiert, der auf Nachhaltigkeit ausgerichtete Verbesserungen für Cloud-Workloads evaluiert, priorisiert, testet und implementiert.

1. Identifizieren von Verbesserungszielen: Überprüfen Sie Ihre Workloads anhand der in diesem Dokument genannten bewährten Methoden für Nachhaltigkeit und ermitteln Sie Ziele für Verbesserungen.

2. Bewerten spezifischer Verbesserungen: Bewerten Sie ausgewählte Änderungen hinsichtlich potenzieller Verbesserungen, geplanter Kosten und Geschäftsrisiken.
3. Priorisieren und Planen von Verbesserungen: Setzen Sie Prioritäten für Änderungen, die die größten Verbesserungen bei geringsten Kosten und Risiken bieten, und erstellen Sie einen Plan für Tests und Implementierung.
4. Testen und Validieren von Verbesserungen: Implementieren Sie Änderungen in Testumgebungen, um deren Verbesserungspotenzial zu validieren.
5. Bereitstellen von Änderungen für die Produktion: Implementieren Sie Änderungen in Produktionsumgebungen.
6. Messen von Ergebnissen und Wiederholen von Erfolgen: Suchen Sie nach Möglichkeiten, Erfolge in allen Workloads zu wiederholen, und machen Sie Änderungen rückgängig, die zu inakzeptablen Ergebnissen geführt haben.

Beispielszenario

Das folgende Beispielszenario wird später in diesem Dokument zur Veranschaulichung der einzelnen Schritte des Verbesserungsprozesses herangezogen.

Ihre Organisation hat eine Workload, die komplexe Image-Manipulationen für Amazon-EC2-Instances durchführt und die geänderten und ursprünglichen Dateien für den Benutzerzugriff speichert. Die Verarbeitungsaktivitäten sind rechenintensiv und die Ausgabedateien sind extrem groß.

Identifizieren von Verbesserungszielen

Lernen Sie die bewährten Methoden kennen, mit denen Sie Ihre Nachhaltigkeitsziele erreichen können. Ausführliche Beschreibungen dieser bewährten Methoden und Empfehlungen für Verbesserungen finden Sie weiter unten in diesem Dokument.

Überprüfen Sie Ihre Workloads und die verwendeten Ressourcen. Identifizieren Sie kritische Punkte, wie z. B. große Bereitstellungen und häufig genutzte Ressourcen. Untersuchen Sie diese kritischen Punkte auf Möglichkeiten zur Verbesserung der effektiven Nutzung Ihrer Ressourcen und zur Verringerung der Gesamtressourcen, die zum Erreichen Ihrer Geschäftsergebnisse erforderlich sind.

Überprüfen Sie Ihre Workload anhand bewährter Methoden und ermitteln Sie Verbesserungsmöglichkeiten.

Indem Sie diesen Schritt auf das Beispielszenario anwenden, identifizieren Sie die folgenden bewährten Methoden als mögliche Ziele für Verbesserungen:

- Verwenden der geringstmöglichen Menge an Hardware zur Erfüllung Ihrer Anforderungen
- Verwenden von Technologien, die Ihre Datenzugriffs- und Speichermuster am besten unterstützen

Ressourcen

- Optimizing your AWS Infrastructure for Sustainability, Part I: Compute
- Optimizing your AWS Infrastructure for Sustainability, Part II: Storage
- Optimizing your AWS Infrastructure for Sustainability, Part III: Networking

Bewerten spezifischer Verbesserungen

Machen Sie sich mit den Ressourcen vertraut, die Ihre Workload für die Durchführung einer Arbeitseinheit vorsieht. Bewerten Sie potenzielle Verbesserungen und schätzen Sie ihre möglichen Auswirkungen, die Kosten für die Implementierung und die damit verbundenen Risiken ab.

Um Verbesserungen im Laufe der Zeit zu messen, müssen Sie zunächst verstehen, was Sie in AWS bereitgestellt haben und wie diese Ressourcen verbraucht werden.

Verschaffen Sie sich zunächst einen vollständigen Überblick über Ihren AWS-Verbrauch und nutzen Sie AWS-Kosten- und Verbrauchsberichte, um kritische Punkte zu identifizieren. Verwenden Sie diesen AWS-Beispielcode, um Sie bei der Überprüfung und Analyse Ihres Berichts mithilfe von Amazon Athena zu unterstützen.

Proxy-Metriken

Wenn Sie bestimmte Änderungen bewerten, müssen Sie auch beurteilen, welche Metriken die Auswirkungen dieser Änderung auf die zugehörige Ressource am besten quantifizieren. Diese Metriken werden als Proxy-Metriken bezeichnet. Wählen Sie Proxy-Metriken, die die Art der Verbesserung, die Sie bewerten, und die Ressourcen, auf die die Verbesserung abzielt, am besten widerspiegeln. Diese Metriken können sich im Laufe der Zeit weiterentwickeln.

Zu den Ressourcen, die zur Unterstützung Ihrer Workload bereitgestellt werden, gehören Datenverarbeitungs-, Speicher- und Netzwerkressourcen. Bewerten Sie die bereitgestellten Ressourcen anhand Ihrer Proxy-Metriken, um zu sehen, wie diese Ressourcen verbraucht werden.

Verwenden Sie Ihre Proxy-Metriken, um die Ressourcen zu messen, die zur Erreichung der Geschäftsergebnisse bereitgestellt werden.

Ressource	Beispiele für Proxy-Metrik	Verbesserungsziele
Datenverarbeitung	vCPU-Minuten	Maximieren der Nutzung der bereitgestellten Ressourcen
Speicher	Bereitgestellte GB	Reduzieren der gesamten Bereitstellung
Netzwerk	Übertragene GB oder Pakete	Reduzieren der übertrage nen Gesamtmenge und der Übertragungsentfernung

Geschäftsmetriken

Wählen Sie Geschäftsmetriken aus, um das Erreichen von Geschäftsergebnissen zu quantifizieren. Ihre Geschäftsmetriken sollten den Wert widerspiegeln, den Ihre Workload bietet, z. B. die Anzahl der gleichzeitig aktiven Benutzer, die Anzahl der API-Aufrufe oder die Anzahl der abgeschlossenen Transaktionen. Diese Metriken können sich im Laufe der Zeit weiterentwickeln. Seien Sie vorsichtig bei der Bewertung von finanzbasierten Geschäftsmetriken, da Unstimmigkeiten beim Wert von Transaktionen Vergleiche ungültig machen.

Wichtige Leistungskennzahlen

Teilen Sie die bereitgestellten Ressourcen durch die erzielten Geschäftsergebnisse, um die bereitgestellten Ressourcen pro Arbeitseinheit zu ermitteln.

$$\textit{Ressourcen pro Arbeitseinheit bereitgestellt} = \frac{\textit{Proxy-Metrik für bereitgestellte Ressource}}{\textit{Geschäftsmetrik für Ergebnis}}$$

KPI-Formel

Verwenden Sie Ihre Ressourcen pro Arbeitseinheit als Ihre KPIs. Legen Sie Basiswerte auf der Grundlage der bereitgestellten Ressourcen als Grundlage für Vergleiche fest.

Ressource	Beispiel-KPIs	Verbesserungsziele
Datenverarbeitung	vCPU-Minuten pro Transaktion	Maximieren der Nutzung der bereitgestellten Ressourcen
Speicher	GB pro Transaktion	Reduzieren der gesamten Bereitstellung
Netzwerk	Pro Transaktion übertragene GB oder Pakete	Reduzieren der übertragenen Gesamtmenge und der Übertragungsentfernung

Schätzen der Verbesserung

Schätzen Sie die Verbesserung sowohl als quantitative Verringerung der bereitgestellten Ressourcen (wie durch Ihre Proxy-Metriken angegeben) als auch als prozentuale Veränderung im Vergleich zu Ihren Basisressourcen, die pro Arbeitseinheit bereitgestellt werden.

Ressource	Beispiel-KPIs	Verbesserungsziele
Datenverarbeitung	Prozentuale Reduzierung der vCPU-Minuten pro Transaktion	Maximieren der Auslastung
Speicher	Prozentuale Reduzierung der GB pro Transaktion	Reduzieren der gesamten Bereitstellung
Netzwerk	Prozentuale Reduzierung der pro Transaktion übertragenen GB oder Pakete	Reduzieren der übertragenen Gesamtmenge und der Übertragungsentfernung

Bewerten von Verbesserungen

Bewerten Sie mögliche Verbesserungen anhand des erwarteten Nettonutzens. Bewerten Sie den Zeit-, Kosten- und Arbeitsaufwand für die Implementierung und Wartung sowie die Geschäftsrisiken, z. B. unvorhergesehene Auswirkungen.

Gezielte Verbesserungen stellen oft einen Kompromiss zwischen den verschiedenen Arten von verbrauchten Ressourcen dar. So können Sie z. B. ein Ergebnis speichern, um den Datenverarbeitungsaufwand zu verringern, oder Sie können Daten verarbeiten, bevor Sie das Ergebnis an einen Client senden, um die übertragenen Daten zu begrenzen. Diese Kompromisse werden später noch ausführlicher erörtert.

Berücksichtigen Sie bei der Bewertung der Risiken für Ihre Workload auch nichtfunktionale Anforderungen wie Sicherheit, Zuverlässigkeit, Leistungseffizienz, Kostenoptimierung und die Auswirkungen von Verbesserungen auf Ihre Fähigkeit, Ihre Workload zu betreiben.

Indem Sie diesen Schritt auf das Beispielszenario anwenden, bewerten Sie die angestrebten Verbesserungen mit den folgenden Ergebnissen:

Bewährte Methode	Gezielte Verbesserung	Potenzial	Kosten	Risiko
Verwenden der geringstmöglichen Menge an Hardware zur Erfüllung Ihrer Anforderungen	Implementierung einer vorausschauenden Skalierung zur Verringerung von Zeiten geringer Auslastung	Mittelschwer	Niedrig	Niedrig
Verwenden von Technologien, die Ihre Datenzugriffs- und Speicherm uster am besten unterstützen	Implementierung effektiverer Komprimierungsmechanismen, um die Gesamtspeicherung und die dafür benötigte Zeit zu reduzieren	Hoch	Niedrig	Niedrig

Die Implementierung von prädikativer Skalierung reduziert die vCPU-Stunden, die von nicht oder nur unzureichend ausgelasteten Instances verbraucht werden, und bietet mit einer geschätzten

Reduzierung der verbrauchten Ressourcen um 11 % moderate Vorteile gegenüber bestehenden Skalierungsmechanismen. Die damit verbundenen Kosten sind gering und umfassen die Konfiguration der Cloud-Ressourcen und den Betrieb der vorausschauenden Skalierung für Amazon EC2 Auto Scaling. Das Risiko ist eine eingeschränkte Leistung, wenn die Aufskalierung reaktiv erfolgt, wenn die Nachfrage die Prognosen übersteigt.

Die Implementierung einer effektiveren Komprimierung wird sich erheblich auf die Dateigröße aller Original-Images und bearbeiteten Images auswirken und den Speicherbedarf in der Produktion um schätzungsweise 25 % verringern. Die Implementierung des neuen Algorithmus ist eine Substitution mit geringem Aufwand und Risiko.

Priorisieren und Planen von Verbesserungen

Legen Sie die Prioritäten für die ermittelten Verbesserungen auf der Grundlage der größten zu erwartenden Auswirkungen mit den geringsten Kosten und akzeptablen Risiken fest.

Entscheiden Sie, auf welche Verbesserungen Sie sich zunächst konzentrieren wollen, und nehmen Sie diese in Ihre Ressourcenplanung und Ihre Entwicklungs-Roadmap auf.

Indem Sie diesen Schritt auf das Beispielszenario anwenden, priorisieren Sie die gezielten Verbesserungen wie folgt:

Priorität	Verbesserung	Potenzial	Kosten	Risiko
1	Implementieren wirksamerer Komprimierungsmechanismen	Hoch	Niedrig	Niedrig
2	Implementieren von prädiktiver Skalierung	Mittelschwer	Niedrig	Niedrig

Das hohe Potenzial, die geringen Kosten und das Risiko einer Aktualisierung der Dateikomprimierung machen sie zu einem hochwertigen Ziel für Ihr Unternehmen und zu einer Priorität gegenüber der Implementierung von prädiktiver Skalierung. Sie legen fest, dass die Implementierung der prädiktiven

Skalierung mit ihrem mittleren Wirkungspotenzial, den geringen Kosten und dem geringen Risiko die vorrangige Verbesserung nach Abschluss der Dateikomprimierung sein sollte.

Sie beauftragen ein Teammitglied mit der Implementierung einer verbesserten Dateikomprimierung und fügen Ihrem Backlog prädiktive Skalierung hinzu.

Testen und Validieren von Verbesserungen

Führen Sie kleine Tests mit minimalen Investitionen durch, um das Risiko eines groß angelegten Aufwands zu verringern.

Implementieren Sie eine repräsentative Kopie Ihrer Workload in Ihrer Testumgebung, um die Kosten und Risiken für die Durchführung von Tests und Validierungen zu begrenzen. Führen Sie einen vordefinierten Satz von Testtransaktionen durch, messen Sie die bereitgestellten Ressourcen und ermitteln Sie die pro Arbeitseinheit verbrauchten Ressourcen, um eine Test-Baseline zu erstellen.

Implementieren Sie die angestrebte Verbesserung in der Testumgebung und wiederholen Sie den Test mit der gleichen Methodik unter den gleichen Bedingungen. Messen Sie dann die bereitgestellten Ressourcen und die pro Arbeitseinheit verbrauchten Ressourcen, wenn Sie Ihre Verbesserung umgesetzt haben.

Berechnen Sie die prozentuale Veränderung der pro Arbeitseinheit bereitgestellten Ressourcen im Vergleich zum Ausgangswert und bestimmen Sie die erwartete quantitative Reduzierung der in Ihrer Produktionsumgebung bereitgestellten Ressourcen. Vergleichen Sie diese Werte mit den erwarteten Werten. Stellen Sie fest, ob das Ergebnis ein akzeptables Maß an Verbesserung darstellt. Beurteilen Sie, ob der Nettonutzen der Verbesserung aufgrund von Kompromissen bei den zusätzlich verbrauchten Ressourcen inakzeptabel ist.

Bestimmen Sie, ob die Verbesserung ein Erfolg ist und ob Ressourcen in die Umsetzung der Änderung in der Produktion investiert werden sollten. Wenn die Änderung zu diesem Zeitpunkt als nicht erfolgreich bewertet wird, leiten Sie Ihre Ressourcen um, um Ihr nächstes Ziel zu testen und zu validieren, und setzen Sie Ihren Verbesserungszyklus fort.

Prozentuale Reduzierung der bereitgestellten Ressourcen pro Arbeitseinheit	Quantitative Reduzierung der bereitgestellten Ressourcen	Action (Aktion)
Erwartungen erfüllt	Erwartungen erfüllt	Fortfahren mit der Verbesser ung

Prozentuale Reduzierung der bereitgestellten Ressourcen pro Arbeitseinheit	Quantitative Reduzierung der bereitgestellten Ressourcen	Action (Aktion)
Erwartungen nicht erfüllt	Erwartungen erfüllt	Fortfahren mit der Verbesser ung
Erwartungen erfüllt	Erwartungen nicht erfüllt	Alternative Verbesserung anstreben
Erwartungen nicht erfüllt	Erwartungen nicht erfüllt	Alternative Verbesserung anstreben

Indem Sie diesen Schritt auf das Beispielszenario anwenden, führen Sie Tests durch, um den Erfolg zu überprüfen.

Nach den Tests mit dem verbesserten Komprimierungsalgorithmus entsprach die prozentuale Verringerung der pro Arbeitseinheit bereitgestellten Ressourcen (der für das ursprüngliche und das geänderte Image benötigte Speicherplatz) den Erwartungen, mit einer durchschnittlichen Verringerung des bereitgestellten Speicherplatzes um 30 % und einer vernachlässigbaren Erhöhung der Rechenlast.

Sie stellen fest, dass die zusätzlichen Rechenressourcen, die für die Anwendung des verbesserten Komprimierungsalgorithmus auf bestehende Dateien in der Produktion erforderlich sind, im Vergleich zu der erzielten Speicherplatzeinsparung unbedeutend sind. Sie haben den Erfolg der quantitativen Verringerung der erforderlichen Ressourcen (TB an Speicherplatz) bestätigt, und die Verbesserung ist für die Produktionsbereitstellung freigegeben.

Bereitstellen von Änderungen für die Produktion

Implementieren Sie getestete, validierte und genehmigte Verbesserungen in der Produktion. Führen Sie die Implementierung mit begrenzten Bereitstellungen durch, bestätigen Sie die Funktionalität Ihrer Workload, testen Sie die tatsächliche Verringerung der bereitgestellten Ressourcen und des Ressourcenverbrauchs pro Arbeitseinheit innerhalb der begrenzten Bereitstellung und prüfen Sie auf unbeabsichtigte Folgen der Änderung. Fahren Sie nach erfolgreichen Tests mit der vollständigen Bereitstellung fort.

Machen Sie Änderungen rückgängig, wenn Tests fehlschlagen oder Sie auf inakzeptable, unbeabsichtigte Folgen Ihrer Änderung stoßen.

Indem Sie diesen Schritt auf das Beispielszenario anwenden, ergreifen Sie die folgenden Maßnahmen.

Sie implementieren die Änderungen in der Produktion mit einer begrenzten Bereitstellung durch eine Blau/Grün-Bereitstellungsmethodik. Die Funktionstests für die neu installierten Instances sind erfolgreich. Sie sehen eine durchschnittliche Reduzierung des bereitgestellten Speichers für ursprüngliche und manipulierte Image-Dateien um 26 %. Bei der Komprimierung neuer Dateien gibt es keine Anzeichen für einen Anstieg der Rechenlast.

Sie bemerken eine unerwartete Verringerung der Zeit, die für die Komprimierung von Image-Dateien benötigt wird, und führen dies auf den stark optimierten Code für den neuen Komprimierungsalgorithmus zurück.

Sie fahren mit der vollständigen Bereitstellung der neuen Version fort.

Messen von Ergebnissen und Wiederholen von Erfolgen

Messen Sie die Ergebnisse und wiederholen Sie die Erfolge auf folgende Weise:

- Messen Sie die anfängliche Verbesserung der bereitgestellten Ressourcen pro Arbeitseinheit und den quantitativen Rückgang der bereitgestellten Ressourcen.
- Vergleichen Sie erste Schätzungen und Testergebnisse mit Ihren Produktionsmessungen. Ermitteln Sie Faktoren, die zu den Unterschieden beigetragen haben könnten, und aktualisieren Sie gegebenenfalls Ihre Schätz- und Testmethoden.
- Bestimmen Sie den Erfolg und den Grad des Erfolgs und teilen Sie die Ergebnisse mit den Beteiligten.
- Wenn Sie Änderungen aufgrund fehlgeschlagener Tests oder unbeabsichtigter negativer Folgen der Änderung rückgängig machen mussten, ermitteln Sie die dafür verantwortlichen Faktoren. Iterieren Sie, wo dies möglich ist, oder bewerten Sie neue Ansätze, um die Ziele der Veränderung zu erreichen.
- Wenden Sie das Gelernte an, Gebrauchen Sie, was Sie gelernt haben, legen Sie Standards fest und wenden Sie erfolgreiche Verbesserungen auf andere Systeme an, die ebenfalls davon profitieren können. Erfassen und teilen Sie Ihre Methodik, die zugehörigen Artefakte und den Nettonutzen mit anderen Teams und Organisationen, damit andere Ihren Standard übernehmen und Ihren Erfolg nachahmen können.

- Überwachen Sie die bereitgestellten Ressourcen pro Arbeitseinheit und verfolgen Sie Änderungen und Gesamtauswirkungen im Zeitverlauf. Änderungen Ihrer Workload oder der Art und Weise, wie Ihre Kunden Ihre Workload in Anspruch nehmen, können sich auf die Wirksamkeit Ihrer Verbesserungen auswirken. Bewerten Sie die Verbesserungsmöglichkeiten neu, wenn Sie kurzfristig einen erheblichen Rückgang der Effektivität Ihrer Verbesserung oder im Laufe der Zeit einen kumulierten Rückgang der Effektivität feststellen.
- Quantifizieren Sie den Nettonutzen Ihrer Verbesserung im Laufe der Zeit (einschließlich des Nutzens für andere Teams, die Ihre Verbesserung angewandt haben, falls verfügbar), um die Rentabilität Ihrer Verbesserungsaktivitäten aufzuzeigen.

Indem Sie diesen Schritt auf das Beispielszenario anwenden, messen Sie die folgenden Ergebnisse.

Ihre Workload zeigt eine anfängliche Verbesserung von 23 % bei den Speicheranforderungen, nachdem der neue Komprimierungsalgorithmus auf bestehende Image-Dateien angewandt wurde.

Der gemessene Wert stimmt weitgehend mit den anfänglichen Schätzungen (25 %) überein, und der signifikante Unterschied im Vergleich zu den Tests (30 %) ist darauf zurückzuführen, dass die bei den Tests verwendeten Image-Dateien nicht repräsentativ für die in der Produktion vorhandenen Image-Dateien sind. Sie ändern den Test-Image-Satz so, dass er den Images in der Produktion besser entspricht.

Die Verbesserung wird als voller Erfolg gewertet. Die Gesamtreduzierung des bereitgestellten Speicherplatzes liegt 2 % unter den geschätzten 25 %, aber 23 % sind immer noch eine enorme Verbesserung der Nachhaltigkeitsauswirkungen und gehen mit einer entsprechenden Kosteneinsparung einher.

Die einzigen unbeabsichtigten Folgen dieser Änderung sind die Verringerung der Zeit, die für die Komprimierung benötigt wird, und eine entsprechende Verringerung des vCPU-Verbrauchs. Diese Verbesserungen sind auf den hochgradig optimierten Code zurückzuführen.

Sie richten ein internes Open-Source-Projekt ein, in dem Sie Ihren Code, die zugehörigen Artefakte, Anleitungen zur Implementierung der Änderung und die Ergebnisse Ihrer Implementierung bereitstellen. Das interne Open-Source-Projekt macht es Ihren Teams leicht, den Code für alle Anwendungsfälle der persistenten Dateispeicherung zu übernehmen. Ihre Teams übernehmen die Verbesserung als Standard. Ein weiterer Vorteil des internen Open-Source-Projekts besteht darin, dass jeder, der die Lösung übernimmt, von den Verbesserungen der Lösung profitiert, und jeder kann Verbesserungen zum Projekt beitragen.

Sie veröffentlichen Ihren Erfolg und teilen das Open-Source-Projekt in Ihrer gesamten Organisation. Jedes Team, das die Lösung übernimmt, repliziert den Nutzen mit minimalen Investitionen und trägt zum Nettonutzen Ihrer Investition bei. Sie veröffentlichen diese Daten als eine fortlaufende Erfolgsgeschichte.

Sie überwachen die Auswirkungen der Verbesserungen im Laufe der Zeit und nehmen bei Bedarf Änderungen an dem internen Open-Source-Projekt vor.

Nachhaltigkeit als nichtfunktionale Anforderung

Die Aufnahme von Nachhaltigkeit in die Liste der geschäftlichen Anforderungen kann zu kostengünstigeren Lösungen führen. Wenn Sie sich darauf konzentrieren, die von Ihnen genutzten Ressourcen besser zu nutzen und weniger von ihnen zu verbrauchen, führt dies direkt zu Kosteneinsparungen in AWS, da Sie nur für das bezahlen, was Sie nutzen.

Das Erreichen von Nachhaltigkeitszielen erfordert möglicherweise keine gleichwertigen Kompromisse bei einer oder mehreren anderen traditionellen Metriken wie Betriebszeit, Verfügbarkeit oder Reaktionszeit. Sie können erhebliche Nachhaltigkeitsgewinne erzielen, ohne dass dies messbare Auswirkungen auf Servicelevel hat. Wenn geringfügige Kompromisse erforderlich sind, können die durch diese Kompromisse erzielten Verbesserungen der Nachhaltigkeit die Veränderung der Servicequalität überwiegen.

Ermutigen Sie Ihre Teammitglieder, bei der Entwicklung funktionaler Anforderungen kontinuierlich mit Verbesserungen der Nachhaltigkeit zu experimentieren. Teams sollten bei der Festlegung von Zielen auch Proxy-Metriken einbeziehen, um sicherzustellen, dass sie die Ressourcenintensität bei der Entwicklung ihrer Workload bewerten.

Im Folgenden finden Sie Beispiele für Kompromisse, die den Verbrauch von Cloud-Ressourcen reduzieren können:

Anpassen der Ergebnisqualität: Mithilfe von Approximate Computing können Sie bei der Ergebnisqualität Kompromisse eingehen, um die Workload-Intensität zu verringern. Beim Approximate Computing werden Optimierungsmöglichkeiten ermittelt, indem geprüft wird, was produziert wird und was Kunden brauchen. Wenn Sie Ihre Daten beispielsweise in eine feste Datenstruktur einbringen, können Sie den ORDER BY-Operator in SQL weglassen, um unnötige Verarbeitungen zu vermeiden, Ressourcen zu sparen und trotzdem eine akzeptable Antwort zu erhalten.

Anpassen der Reaktionszeit: Eine Antwort mit einer langsameren Reaktionszeit kann CO_2 reduzieren, indem der gemeinsame Overhead minimiert wird. Die Verarbeitung flüchtiger Ad-hoc-Aufgaben kann zu Startup-Overhead führen. Gruppieren und verarbeiten Sie Aufgaben in Stapeln, anstatt bei jeder neuen Aufgabe für den Overhead zu zahlen. Die Stapelverarbeitung führt zu einer verlängerten Reaktionszeit, verringert jedoch den gemeinsamen Overhead für das Hochfahren der Instance, das Herunterladen des Quellcodes und die Ausführung des Prozesses.

Anpassen der Verfügbarkeit: Mit AWS können Sie Redundanz hinzufügen und Hochverfügbarkeitsziele mit nur wenigen Klicks erreichen. Sie können die Redundanz durch

Techniken wie statische Stabilität erhöhen, indem Sie ungenutzte Ressourcen bereitstellen, die immer zu einer geringeren Auslastung führen. Bewerten Sie bei der Festlegung von Zielen die Bedürfnisse des Unternehmens. Relativ geringe Abstriche bei der Verfügbarkeit können zu wesentlich größeren Verbesserungen bei der Nutzung führen. Das Architekturmuster der statischen Stabilität beinhaltet beispielsweise die Bereitstellung ungenutzter Failover-Kapazitäten, um die Last nach einem Komponentenfehler sofort zu übernehmen. Eine Lockerung der Verfügbarkeitsanforderung kann den Bedarf an ungenutzten Online-Kapazitäten beseitigen, da die Automatisierung Zeit für die Bereitstellung von Ersatzressourcen bietet. Das Hinzufügen von On-Demand-Failover-Kapazität führt zu einer höheren Gesamtauslastung ohne Geschäftsauswirkungen während des Normalbetriebs und hat den Nebeneffekt, dass die Kosten gesenkt werden.

Bewährte Methoden in Bezug auf Nachhaltigkeit in der Cloud

Optimieren Sie die Workload-Platzierung und optimieren Sie Ihre Architektur im Hinblick auf Bedarf, Software, Daten, Hardware und Prozesse, um die Energieeffizienz zu erhöhen. Jeder dieser Bereiche bietet Möglichkeiten für den Einsatz bewährter Methoden zur Reduzierung der Nachhaltigkeitsauswirkungen Ihrer Cloud-Workloads – beispielsweise durch die Maximierung der Nutzung und die Minimierung von Abfällen sowie des Einsatzes der für Ihre Workloads bereitgestellten Energie und der verbrauchenden Ressourcen.

Themen

- Auswahl der Region
- Ausrichtung am Bedarf
- Software und Architektur
- Datenverwaltung
- Hardware und Services
- Prozess und Kultur

Auswahl der Region

Welche Region Sie für Ihre Workload auswählen, hat signifikante Auswirkungen auf deren KPIs, u. a. Leistung, Kosten und CO_2-Bilanz. Um diese KPIs effizient zu verbessern, sollten Sie die Regionen für Ihre Workloads in Abhängigkeit von den Unternehmensanforderungen und Nachhaltigkeitszielen auswählen.

Bewährte Methoden

- SUS01-BP01 Auswählen der Region auf Grundlage von Unternehmensanforderungen und Nachhaltigkeitszielen

SUS01-BP01 Auswählen der Region auf Grundlage von Unternehmensanforderungen und Nachhaltigkeitszielen

Wählen Sie für Ihre Workload eine Region auf der Grundlage Ihrer Geschäftsanforderungen und Nachhaltigkeitsvorgaben aus, um KPIs wie Leistung, Kosten und CO2-Bilanz zu optimieren.

Typische Anti-Muster:

- Sie wählen die Region der Workload auf der Grundlage Ihres eigenen Standorts aus.
- Sie konsolidieren alle Workload-Ressourcen an einem einzelnen geografischen Standort.

Vorteile der Nutzung dieser bewährten Methode: Wenn Sie eine Workload in der Nähe von Amazon-Projekten für erneuerbare Energien oder in Regionen mit nachweislich niedrigen Kohlendioxidemissionen platzieren, kann die CO2-Bilanz einer Clouds-Workload gesenkt werden.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Mittel

Implementierungsleitfaden

Die AWS Cloud ist ein ständig wachsendes Netzwerk aus Regionen und Points of Presence (PoP), die durch eine globale Netzwerkinfrastruktur verbunden sind. Die Wahl der Region für Ihre Workload hat erhebliche Auswirkungen auf zugehörige KPIs wie Leistung, Kosten und CO2-Bilanz. Um diese KPIs effizient zu verbessern, sollten Sie Regionen für Ihre Workload abhängig von den Unternehmensanforderungen und Nachhaltigkeitszielen auswählen.

Implementierungsschritte

- Befolgen Sie diese Schritte, um potenzielle Regionen für Ihre Workload zu bewerten und in die engere Auswahl zu nehmen. Berücksichtigen Sie dabei die Anforderungen Ihres Unternehmens, unter anderem in Bezug auf Compliance, verfügbare Funktionen, Kosten und Latenz:
  - Vergewissern Sie sich, dass die Regionen konform sind (basierend auf Ihren erforderlichen lokalen Vorschriften).
  - Prüfen Sie anhand der Liste der verfügbaren AWS-Services nach Regionen, ob die Regionen über die für Ihre Workload erforderlichen Services und Funktionen verfügen.
  - Berechnen Sie die Kosten der Workload in jeder Region mithilfe des AWS Pricing Calculator.
  - Testen Sie die Netzwerklatenz zwischen den Standorten Ihrer Endbenutzer und jeder AWS-Region.
- Wählen Sie Regionen in der Nähe von Amazon-Projekten für erneuerbare Energien aus. Es sollte sich um Regionen handeln, in denen das Stromnetz nachweislich geringere Kohlendioxidemissionen generiert als an anderen Standorten (oder in anderen Regionen).
  - Ermitteln Sie die relevanten Nachhaltigkeitsrichtlinien, um die jährlichen CO2-Emissionen gemäß dem Greenhouse Gas Protocol nachzuverfolgen und zu vergleichen (marktbasierte und standortbasierte Verfahren).
  - Wählen Sie die Region entsprechend der Methode aus, mit der Sie CO2-Emissionen nachverfolgen. Weitere Informationen zum Auswählen einer Region anhand von

Nachhaltigkeitsrichtlinien finden Sie im Artikel zum Auswählen einer Region für Ihre Workload auf der Grundlage von Nachhaltigkeitszielen.

Ressourcen

Zugehörige Dokumente:

- Grundlegendes zu CO2-Emissionsschätzungen
- Amazon weltweit
- Methodik für erneuerbare Energien
- „Relevante Aspekte bei der Wahl einer Region für Ihre Workloads“ erläutert

Zugehörige Videos:

- AWS re:Invent 2023 – Nachhaltigkeitsinnovationen in der globalen AWS-Infrastruktur
- AWS re:Invent 2023 – Nachhaltige Architektur: Vergangenheit, Gegenwart und Zukunft
- AWS re:Invent 2022 – Bereitstellung nachhaltiger, leistungsstarker Architekturen
- AWS re:Invent 2022 – Nachhaltige Architektur und Reduzieren der AWS-CO2-Bilanz
- AWS re:Invent 2022 – Nachhaltigkeit in der globalen AWS-Infrastruktur

Ausrichtung am Bedarf

Wenn Sie berücksichtigen, wie Benutzer und Anwendungen Ihre Workloads und andere Ressourcen nutzen, können Sie auf diese Weise Verbesserungsmöglichkeiten ermitteln, um Nachhaltigkeitsziele zu erreichen. Skalieren Sie Ihre Infrastruktur so, dass Sie den Bedarf kontinuierlich anpassen können. Sorgen Sie zudem dafür, dass zur Unterstützung Ihrer Benutzer nicht mehr Ressourcen verwendet werden als unbedingt nötig. Richten Sie Service-Levels an den Kundenanforderungen aus. Positionieren Sie Ressourcen so, dass die Netzwerkkapazitäten, die für Benutzer und Anwendungen erforderlich sind, begrenzt werden. Entfernen Sie ungenutzte Komponenten. Stellen Sie Teammitgliedern Geräte zur Verfügung, die ihre Anforderungen bei geringstmöglichen Auswirkungen auf die Nachhaltigkeit erfüllen.

Bewährte Methoden

- SUS02-BP01 Dynamisches Skalieren der Workload-Infrastruktur
- SUS02-BP02 Ausrichten von SLAs an Nachhaltigkeitszielen

- SUS02-BP03 Beenden der Erstellung und Wartung nicht verwendeter Komponenten
- SUS02-BP04 Optimieren der geografischen Platzierung von Workloads auf der Grundlage ihrer Netzwerkanforderungen
- SUS02-BP05 Optimieren von Ressourcen für Teammitglieder im Hinblick auf die ausgeführten Aktivitäten
- SUS02-BP06 Implementierung von Pufferung oder Drosselung, um die Bedarfskurve zu verflachen

SUS02-BP01 Dynamisches Skalieren der Workload-Infrastruktur

Nutzen Sie die Elastizität der Cloud und skalieren Sie Ihre Infrastruktur dynamisch, um das Angebot an Cloud-Ressourcen an den Bedarf anzupassen und eine Überbereitstellung von Kapazitäten in Ihrer Workload zu vermeiden.

Typische Anti-Muster:

- Sie skalieren Ihre Infrastruktur nicht mit der Benutzerlast.
- Sie skalieren Ihre Infrastruktur immer manuell.
- Sie behalten die erhöhte Kapazität nach dem Hochskalieren bei, anstatt sie wieder herunterzuskalieren.

Vorteile der Nutzung dieser bewährten Methode: Das Konfigurieren und Testen der Workload-Elastizität trägt dazu bei, das Angebot an Cloud-Ressourcen effizient an den Bedarf anzupassen und eine Überbereitstellung von Kapazitäten zu vermeiden. Sie können die Vorteile der Elastizität in der Cloud nutzen, um die Kapazität während und nach Bedarfsspitzen automatisch zu skalieren und so sicherzustellen, dass Sie nur die Menge an Ressourcen nutzen, die für die Erfüllung Ihrer Geschäftsanforderungen erforderlich ist.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Mittel

Implementierungsleitfaden

Die Cloud bietet Ihnen die Flexibilität, Ressourcen dynamisch durch verschiedene Mechanismen zu erweitern oder zu reduzieren, um einem veränderten Bedarf gerecht zu werden. Eine optimale Abstimmung von Angebot und Bedarf führt zu den geringsten Auswirkungen auf die Umgebung für eine Workload.

Der Bedarf kann fest oder variabel sein und erfordert Metriken und Automatisierung, um sicherzustellen, dass die Verwaltung nicht zur Last wird. Anwendungen können vertikal (hoch oder herunter) und/oder horizontal (ab oder auf) skaliert werden. Bei der vertikalen Skalierung wird die Instance-Größe geändert, bei der horizontalen Skalierung die Anzahl von Instances.

Sie können verschiedene Ansätze nutzen, um das Angebot an Ressourcen auf den Bedarf abzustimmen.

- Zielverfolgungsansatz: Überwachen Sie Ihre Skalierungsmetriken und erhöhen oder verringern Sie die Kapazität automatisch nach Bedarf.
- Prädiktives Skalieren: Skalieren Sie auf der Grundlage erwarteter täglicher und wöchentlicher Trends.
- Zeitplanbasierter Ansatz: Legen Sie einen eigenen Skalierungszeitplan auf der Grundlage vorhersehbarer Laständerungen fest.
- Service-Skalierung: Wählen Sie Services (beispielsweise Serverless) aus, die nativ von Natur aus skalierbar sind oder Auto Scaling als Feature bieten.

Identifizieren Sie Zeiträume mit geringer oder gar keiner Nutzung und skalieren Sie Ressourcen, um überschüssige Kapazitäten zu entfernen und die Effizienz zu verbessern.

Implementierungsschritte

- Elastizität ermöglicht die Anpassung der verfügbaren Ressourcen an den Bedarf. Instances, Container und Funktionen bieten Mechanismen für Elastizität, entweder in Kombination mit Auto Scaling oder als Feature des Service. AWS bietet eine Reihe von Mechanismen für Auto Scaling, um sicherzustellen, dass Workloads in Zeiten geringer Benutzerlast schnell und einfach herunterskaliert werden können. Hier sind einige Beispiele für Auto-Scaling-Mechanismen:

Auto-Scaling-Mechanismus	Verwendung
Amazon EC2 Auto Scaling	Verwenden Sie diesen Mechanismus, um zu überprüfen, ob die richtige Menge an Amazon-EC2-Instances zur Verfügung steht, um die Benutzerlast für Ihre Anwendung zu bewältigen.

Auto-Scaling-Mechanismus	Verwendung
Application Auto Scaling	Verwenden Sie diesen Mechanismus, um die Ressourcen für einzelne AWS-Services über Amazon EC2 hinaus automatisch zu skalieren – beispielsweise Lambda-Funktionen oder Amazon Elastic Container Service (Amazon ECS)-Services.
Kubernetes Cluster Autoscaler	Verwenden Sie diesen Mechanismus, um Kubernetes-Cluster in AWS automatisch zu skalieren.

- Skalieren wird häufig im Zusammenhang mit Datenverarbeitungs-Services wie Amazon-EC2-Instances oder AWS Lambda-Funktionen behandelt. Ziehen Sie die Konfiguration von nicht Daten verarbeitenden Services wie Lese- und Schreibkapazitätseinheiten von Amazon DynamoDB oder Shards von Amazon Kinesis Data Streams in Betracht, um den Bedarf zu decken.
- Vergewissern Sie sich, dass die Metriken zum Hoch- oder Herunterskalieren für die jeweilige Art der bereitgestellten Workload überprüft werden. Wenn Sie eine Anwendung zur Video-Transkodierung bereitstellen, wird eine CPU-Auslastung von 100 % erwartet, weshalb dies nicht die Hauptmetrik sein sollte. Sie können bei Bedarf eine benutzerdefinierte Metrik (etwa die Speicherauslastung) für Ihre Skalierungsrichtlinie verwenden. Beachten Sie bei der Wahl geeigneter Metriken die folgenden Hinweise für Amazon EC2:
  - Es muss sich um eine gültige Nutzungsmetrik handeln, die beschreibt, wie stark eine Instance genutzt wird.
  - Der Wert der Metrik muss sich proportional zur Anzahl der Instances in der Auto-Scaling-Gruppe erhöhen oder verringern.
- Verwenden Sie für Ihre Auto-Scaling-Gruppe eine dynamische Skalierung anstelle einer manuellen Skalierung. Außerdem empfiehlt es sich, bei der dynamischen Skalierung Skalierungsrichtlinien zur Zielverfolgung zu verwenden.
- Vergewissern Sie sich, dass Workload-Bereitstellungen sowohl Hoch- als auch Herunterskalierungsereignisse behandeln können. Erstellen Sie Testszenarien für Herunterskalierungsereignisse, um sich zu vergewissern, dass sich die Workload wie erwartet verhält und die Benutzererfahrung nicht beeinträchtigt wird (etwa durch den Verlust von Sticky Sessions). Sie können den Aktivitätsverlauf verwenden, um eine Skalierungsaktivität für eine Auto-Scaling-Gruppe zu überprüfen.

- Überprüfen Sie Ihre Workload auf vorhersagbare Muster und skalieren Sie proaktiv, wenn Sie vorhergesagte und geplante Bedarfsänderungen erwarten. Mit der prädiktiven Skalierung können Sie die Notwendigkeit einer Überbereitstellung von Kapazitäten vermeiden. Weitere Einzelheiten finden Sie unter Prädiktive Skalierung mit Amazon EC2 Auto Scaling.

Ressourcen

Zugehörige Dokumente:

- Erste Schritte mit Amazon EC2 Auto Scaling
- Prädiktive Skalierung für EC2, unterstützt von Machine Learning
- Analysieren des Benutzerverhaltens mit Amazon OpenSearch Service, Amazon Data Firehose und Kibana
- Was ist Amazon CloudWatch?
- Überwachen der DB-Last mit Performance Insights in Amazon RDS
- Vorstellung von nativer Unterstützung für die prädiktive Skalierung mit Amazon EC2 Auto Scaling
- Vorstellung von Karpenter – Open-Source-Kubernetes-Cluster-Autoscaler mit hoher Leistung
- Detaillierte Einblicke in Amazon ECS Cluster Auto Scaling

Zugehörige Videos:

- AWS re:Invent 2023 – Skalierung in AWS für die ersten 10 Millionen Benutzer
- AWS re:Invent 2023 – Nachhaltige Architektur: Vergangenheit, Gegenwart und Zukunft
- AWS re:Invent 2022 – Entwickeln einer kosten-, energie- und ressourceneffizienten Computing-Umgebung
- AWS re:Invent 2022 – Container-Skalierung von einem einzelnen Benutzer auf mehrere Millionen
- AWS re:Invent 2023 – Skalierung der FM-Inferenz auf Hunderte von Modellen mit Amazon SageMaker
- AWS re:Invent 2023 – Nutzung der Leistungsfähigkeit von Karpenter für die Skalierung, Optimierung und Aktualisierung von Kubernetes

Zugehörige Beispiele:

- Auto Scaling

SUS02-BP02 Ausrichten von SLAs an Nachhaltigkeitszielen

Überprüfen und optimieren Sie die Service Level Agreements (SLA) für Workloads auf der Grundlage Ihrer Nachhaltigkeitsziele, um die für die Unterstützung Ihrer Workload erforderlichen Ressourcen zu minimieren und weiterhin die Geschäftsanforderungen zu erfüllen.

Typische Anti-Muster:

- Workload-SLAs sind unbekannt oder nicht eindeutig.
- Sie definieren Ihre SLA nur für Verfügbarkeit und Leistung.
- Sie verwenden für alle Ihre Workloads die gleichen Designmuster (wie etwa Multi-AZ-Architektur).

Vorteile der Nutzung dieser bewährten Methode: Die Abstimmung von SLAs mit Nachhaltigkeitszielen führt zu einer optimalen Ressourcennutzung bei gleichzeitiger Erfüllung der Geschäftsanforderungen.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Niedrig

Implementierungsleitfaden

SLAs definieren das von einer Cloud-Workload erwartete Service-Niveau, beispielsweise im Hinblick auf Reaktionszeit, Verfügbarkeit und Datenaufbewahrung. Sie beeinflussen die Architektur, die Ressourcennutzung und die Umweltauswirkungen einer Cloud-Workload. Prüfen Sie SLAs regelmäßig und streben Sie eine erhebliche Reduzierung der Ressourcennutzung bei akzeptabler Verringerung des Service-Niveaus an.

Implementierungsschritte

- Nachhaltigkeitsziele verstehen: Identifizieren Sie Nachhaltigkeitsziele in Ihrer Organisation – etwa eine Reduzierung des CO2-Ausstoßes oder eine bessere Ressourcennutzung.
- SLAs überprüfen: Bewerten Sie Ihre SLAs, um festzustellen, ob sie Ihre Geschäftsanforderungen unterstützen. Wenn Sie SLAs überschreiten, weiten Sie die Überprüfung aus.
- Kompromisse verstehen: Machen Sie sich ein Bild von den Kompromissen zwischen der Komplexität (zum Beispiel hohe Anzahl gleichzeitiger Benutzer), der Leistung (zum Beispiel Latenz) und den Auswirkungen auf die Nachhaltigkeit Ihrer Workloads (zum Beispiel Ressourcenbedarf). In der Regel geht die Priorisierung von zwei der Faktoren auf Kosten des dritten.
- SLAs anpassen: Passen Sie Ihre SLAs an, indem Sie eine erhebliche Reduzierung der Nachhaltigkeitsauswirkungen bei akzeptabler Verringerung des Service-Niveaus anstreben.

  - Nachhaltigkeit und Zuverlässigkeit: Workloads mit hoher Verfügbarkeit verbrauchen in der Regel mehr Ressourcen.
  - Nachhaltigkeit und Leistung: Die Nutzung von mehr Ressourcen, um die Leistung zu steigern, führt unter Umständen zu einer höheren Umweltbelastung.
  - Nachhaltigkeit und Sicherheit: Übermäßig sichere Workloads haben möglicherweise eine höhere Umweltbelastung zur Folge.
- Nachhaltigkeits-SLAs definieren, wenn möglich: Schließen Sie Nachhaltigkeits-SLAs für Ihre Workload ein. Definieren Sie beispielsweise ein Mindestnutzungsniveau als Nachhaltigkeits-SLA für Ihre Computing-Instances.
- Effiziente Designmuster verwenden: Nutzen Sie Designmuster wie Microservices in AWS, die geschäftskritische Funktionen priorisieren, und lassen Sie für nicht kritische Funktionen niedrigere Service Level zu (etwa für Reaktions- und Wiederherstellungszeiten).
- Verantwortlichkeiten festlegen und kommunizieren: Teilen Sie die SLAs mit allen relevanten Stakeholdern, einschließlich Ihres Entwicklungsteams und Ihrer Kunden. Verwenden Sie Berichte, um die SLAs zu verfolgen und zu überwachen. Weisen Sie Verantwortlichkeiten zu, um die Nachhaltigkeitsziele Ihrer SLAs zu erreichen.
- Anreize und Prämien nutzen: Nutzen Sie Anreize und Prämien, um SLAs zu erreichen oder zu übertreffen, die auf Nachhaltigkeitsziele ausgerichtet sind.
- Überprüfen und wiederholen: Überprüfen Sie Ihre SLAs regelmäßig und passen Sie sie an, um sicherzustellen, dass sie mit den sich entwickelnden Nachhaltigkeits- und Leistungszielen in Einklang stehen.

Ressourcen

Zugehörige Dokumente:

- Grundlegendes zu Resilienzmustern und Kompromissen, um eine effiziente Architektur in der Cloud zu entwickeln
- Bedeutung von Service Level Agreements für SaaS-Anbieter

Zugehörige Videos:

- AWS re:Invent 2023 – Kapazität, Verfügbarkeit, Kosteneffizienz: Wählen Sie drei Optionen aus
- AWS re:Invent 2023 – Nachhaltige Architektur: Vergangenheit, Gegenwart und Zukunft

- AWS re:Invent 2023 – Fortschrittliche Integrationsmuster und Kompromisse für lose gekoppelte Systeme
- AWS re:Invent 2022 – Bereitstellung nachhaltiger, leistungsstarker Architekturen
- AWS re:Invent 2022 – Entwickeln einer kosten-, energie- und ressourceneffizienten Computing-Umgebung

SUS02-BP03 Beenden der Erstellung und Wartung nicht verwendeter Komponenten

Nehmen Sie nicht verwendete Ressourcen in Ihrer Workload außer Betrieb, um die Anzahl der Cloud-Ressourcen zu verringern, die zur Unterstützung Ihres Bedarfs und zur Minimierung von Verschwendung erforderlich sind.

Typische Anti-Muster:

- Sie analysieren Ihre Anwendung nicht auf Ressourcen, die redundant sind oder nicht mehr benötigt werden.
- Sie entfernen keine redundanten oder nicht mehr benötigten Ressourcen.

Vorteile der Nutzung dieser bewährten Methode: Das Entfernen nicht genutzter Ressourcen setzt Kapazitäten frei und verbessert die allgemeine Effizienz der Workload.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Niedrig

Implementierungsleitfaden

Nicht verwendete Ressourcen verbrauchen Cloud-Kapazitäten wie Speicherplatz oder Rechenleistung. Wenn Sie solche Ressourcen identifizieren und eliminieren, können Sie diese Kapazitäten freisetzen, was zu einer effizienteren Cloud-Architektur führt. Analysieren Sie Anwendungsressourcen (wie vorab kompilierte Berichte, Datensätze, statische Bilder) sowie Zugriffsmuster für Komponenten, um Redundanzen, eine zu geringe Auslastung und mögliche Kandidaten für die Außerbetriebnahme zu identifizieren. Entfernen Sie diese redundanten Ressourcen, um die Ressourcenverschwendung in Ihrer Workload zu reduzieren.

Implementierungsschritte

- Bestandsaufnahme durchführen: Führen Sie eine umfassende Bestandsaufnahme durch, um alle Komponenten innerhalb Ihrer Workload zu identifizieren.

- Nutzung analysieren: Verwenden Sie die kontinuierliche Überwachung, um statische Komponenten zu identifizieren, die nicht mehr benötigt werden.
- Ungenutzte Komponenten entfernen: Entwickeln Sie einen Plan, um Komponenten zu entfernen, die nicht mehr benötigt werden.
  - Prüfen Sie vor dem Entfernen einer Ressource die Auswirkungen dieser Maßnahme auf die Architektur.
  - Konsolidieren Sie sich überschneidende generierte Komponenten, um eine redundante Verarbeitung zu entfernen.
  - Aktualisieren Sie Ihre Anwendungen, damit diese nicht mehr benötigte Ressourcen nicht weiter produzieren und speichern.
- Mit Dritten kommunizieren: Weisen Sie Dritte an, die Erstellung und Speicherung von Komponenten einzustellen, die in Ihrem Auftrag verwaltet und nicht mehr benötigt werden. Bitten Sie darum, dass redundante Komponenten konsolidiert werden.
- Lebenszyklusrichtlinien verwenden: Verwenden Sie Lebenszyklusrichtlinien, damit ungenutzte Komponenten automatisch gelöscht werden.
  - Mit Amazon-S3-Lebenszyklen können Sie Ihre Objekte während ihres gesamten Lebenszyklus verwalten.
  - Mit Amazon Data Lifecycle Manager lassen sich das Erstellen, Aufbewahren und Löschen von Amazon-EBS-Snapshots und Amazon-EBS-gestützten AMIs automatisieren.
- Überprüfen und optimieren: Überprüfen Sie regelmäßig Ihre Workload, um ungenutzte Komponenten zu identifizieren und zu entfernen.

Ressourcen

Zugehörige Dokumente:

- Optimizing your AWS Infrastructure for Sustainability, Part II: Storage
- How do I terminate active resources that I no longer need on my AWS-Konto?

Zugehörige Videos:

- AWS re:Invent 2.023 - Sustainable architecture: Past, present, and future
- AWS re:Invent 2.022 - Preserving and maximizing the value of digital media assets using Amazon S3

- AWS re:Invent 2.023 - Optimize costs in your multi-account environments

SUS02-BP04 Optimieren der geografischen Platzierung von Workloads auf der Grundlage ihrer Netzwerkanforderungen

Wählen Sie Cloud-Standorte und -Services für Ihre Workload, die die Entfernungen reduzieren, über die Netzwerkdatenverkehr übertragen werden muss, um die Zahl der Netzwerkressourcen zu verringern, die zur Unterstützung Ihre Workload erforderlich sind.

Typische Anti-Muster:

- Sie wählen die Region der Workload auf der Grundlage Ihres eigenen Standorts aus.
- Sie konsolidieren alle Workload-Ressourcen an einem geografischen Standort.
- Der gesamte Datenverkehr fließt durch Ihre bestehenden Rechenzentren.

Vorteile der Nutzung dieser bewährten Methode: Die Platzierung von Workloads in der Nähe der Benutzer bietet die geringstmögliche Latenz und verringert gleichzeitig die Bewegung der Daten durch das Netzwerk und damit die Umweltauswirkungen.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Mittel

Implementierungsleitfaden

Die AWS Cloud-Infrastruktur basiert auf Standortoptionen wie etwa Regionen, Availability Zones, Platzierungsgruppen und Edge-Standorten wie AWS Outposts und AWS Local Zones. Diese Standortoptionen stellen die Konnektivität zwischen Anwendungskomponenten, Cloud-Services, Edge-Netzwerken und On-Premises-Rechenzentren sicher.

Analysieren Sie die Netzwerkzugriffsmuster in Ihrer Workload, um festzustellen, wie diese verwendet werden können, um die Entfernungen für den Netzwerkdatenverkehr zu reduzieren.

Implementierungsschritte

- Analysieren Sie die Netzwerkzugriffsmuster in Ihrer Workload, um zu ermitteln, wie die Benutzer Ihre Anwendung verwenden.
  - Verwenden Sie Überwachungstools wie Amazon CloudWatch und AWS CloudTrail, um Daten zu Netzwerkaktivitäten zu sammeln.

  - Analysen Sie die Daten, um das Netzwerkzugriffsmuster zu identifizieren.
- Wählen Sie die Regionen für Ihre Workload-Bereitstellung auf der Grundlage der folgenden zentralen Elemente aus:
  - Ihr Nachhaltigkeitsziel: wie unter Regionsauswahl erklärt.
  - Standort Ihrer Daten: Für datenintensive Anwendungen (wie etwa Big Data oder Machine Learning) sollte der Anwendungscode so nahe wie möglich zu den Daten ausgeführt werden.
  - Standort Ihrer Benutzer: Wählen Sie für benutzerseitige Anwendungen eine Region (oder Regionen) in der Nähe der Benutzer der Workload.
  - Weitere Einschränkungen: Berücksichtigen Sie Einschränkungen wie Kosten und Compliance, wie unter Relevante Aspekte bei der Wahl einer Region für Ihre Workloads erläutert.
- Verwenden Sie lokale Zwischenspeicherung oder AWS-Zwischenspeicherung für häufig genutzte Ressourcen zur Verbesserung der Leistung, zur Verringerung von Datenverschiebungen und zur Reduzierung der Umweltauswirkungen.

Service	Wann sollte dies verwendet werden?
Amazon CloudFront	Verwenden Sie dies für die Zwischens peicherung statischer Inhalte wie Bilder, Skripts und Videos sowie dynamischer Inhalte wie API-Antworten oder Webanwendungen.
Amazon ElastiCache	Verwenden Sie dies für die Zwischens peicherung von Inhalten für Webanwend ungen.
DynamoDB Accelerator	Verwenden Sie dies für die Add-in-Speicher-Beschleunigung für Ihre DynamoDB-Tabellen.

- Nutzen Sie Services, die Ihnen dabei helfen können, Code näher an den Benutzern Ihrer Workload auszuführen:

Service	Wann sollte dies verwendet werden?
Lambda@Edge	Verwenden Sie dies für rechenintensive Anwendungen, die initiiert werden, wenn sich Objekte nicht im Zwischenspeicher befinden.

Service	Wann sollte dies verwendet werden?
Amazon-CloudFront-Funktionen	Verwenden Sie diese für einfache Anwendungsfälle wie HTTP(s)-Anfragen oder Antwortmanipulationen, die von kurzlebigen Funktionen initiiert werden können.
AWS IoT Greengrass	Verwenden Sie dies für die Ausführung lokaler Rechenoperationen, Messaging sowie die Datenzwischenspeicherung für verbundene Geräte.

- Nutzen Sie Verbindungspooling, um die erneute Nutzung von Verbindungen zu ermöglichen und die Zahl der erforderlichen Ressourcen zu reduzieren.
- Verwenden Sie verteilte Datenspeicher, die nicht auf persistente Verbindungen und synchrone Updates angewiesen sind, um regionale Benutzergruppen zu unterstützen.
- Ersetzen Sie vorab bereitgestellte statische Netzwerkkapazität durch geteilte dynamische Kapazitäten und teilen Sie die Auswirkungen von Netzwerkkapazitäten auf die Nachhaltigkeit mit anderen Abonnenten.

Ressourcen

Zugehörige Dokumente:

- Optimizing your AWS Infrastructure for Sustainability, Part III: Networking
- Dokumentation zu Amazon ElastiCache
- Was ist Amazon CloudFront?
- Hauptfunktionen von Amazon CloudFront
- Globale AWS-Infrastruktur
- AWS Local Zones and AWS Outposts, choosing the right technology for your edge workload
- Platzierungsgruppen
- AWS Local Zones
- AWS Outposts

Zugehörige Videos:

- Demystifying data transfer on AWS
- Scaling network performance on next-gen Amazon EC2 instances
- AWS Local Zones Explainer Video
- AWS Outposts: Overview and How it Works
- AWS re:Invent 2.023 - A migration strategy for edge and on-premises workloads
- AWS re:Invent 2.021 - AWS Outposts: Bringing the AWS experience on premises
- AWS re:Invent 2.020 - AWS Wavelength: Run apps with ultra-low latency at 5G edge
- AWS re:Invent 2.022 - AWS Local Zones: Building applications for a distributed edge
- AWS re:Invent 2.021 - Building low-latency websites with Amazon CloudFront
- AWS re:Invent 2.022 - Improve performance and availability with AWS Global Accelerator
- AWS re:Invent 2.022 - Build your global wide area network using AWS
- AWS re:Invent 2020: Global traffic management with Amazon Route 53

Zugehörige Beispiele:

- AWS Networking-Workshops
- Architecting for sustainability - Minimize data movement across networks

SUS02-BP05 Optimieren von Ressourcen für Teammitglieder im Hinblick auf die ausgeführten Aktivitäten

Optimieren Sie die Ressourcen, die Teammitgliedern zur Verfügung gestellt werden, um negative Auswirkungen auf die Nachhaltigkeit zu minimieren und gleichzeitig ihre Anforderungen zu erfüllen.

Typische Anti-Muster:

- Sie berücksichtigen nicht die Auswirkungen der von Ihren Teammitgliedern verwendeten Geräte auf die Gesamteffizienz Ihrer Cloud-Anwendung.
- Sie verwalten und aktualisieren die von Teammitgliedern verwendeten Ressourcen manuell.

Vorteile der Nutzung dieser bewährten Methode: Die Optimierung der Teammitglieder-Ressourcen verbessert die allgemeine Effizienz Cloud-fähiger Anwendungen.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Niedrig

Implementierungsleitfaden

Verstehen Sie die Ressourcen, mit denen Ihre Teammitglieder Ihre Services nutzen, deren erwartete Lebensdauer sowie die finanziellen und nachhaltigkeitsbezogenen Auswirkungen. Implementieren Sie Strategien zur Optimierung dieser Ressourcen. Beispielsweise können Sie komplexe Vorgänge wie Rendering und Kompilierung auf intensiv genutzter und skalierbarer Infrastruktur anstatt auf weniger ausgelasteten Einzelbenutzersystemen mit hohem Energieverbrauch ausführen.

Implementierungsschritte

- Energieeffiziente Workstations verwenden: Stellen Sie den Teammitgliedern energieeffiziente Workstations und Peripheriegeräte zur Verfügung. Verwenden Sie effiziente Energiemanagementfeatures (wie den Energiesparmodus) auf diesen Geräten, um ihren Energieverbrauch zu reduzieren.
- Virtualisierung verwenden: Verwenden Sie virtuelle Desktops und Anwendungs-Streaming, um Upgrade- und Geräteanforderungen zu begrenzen.
- Remote-Zusammenarbeit fördern: Ermutigen Sie die Teammitglieder, Tools für die Remote-Zusammenarbeit wie Amazon Chime oder AWS Wickr zu verwenden, um den Reisebedarf und die damit verbundenen CO_2-Emissionen zu reduzieren.
- Energieeffiziente Software verwenden: Stellen Sie den Teammitgliedern energieeffiziente Software zur Verfügung, indem Sie nicht benötigte Features und Prozesse entfernen oder deaktivieren.
- Lebenszyklen verwalten: Evaluieren Sie die Auswirkungen von Prozessen und Systemen auf die Lebenszyklen von Geräten. Wählen Sie Lösungen aus, die den Bedarf für Geräteaustauschvorgänge minimieren und gleichzeitig die geschäftlichen Anforderungen erfüllen. Pflegen und aktualisieren Sie regelmäßig Workstations oder Software, um die Effizienz aufrechtzuerhalten und zu verbessern.
- Remote-Verwaltung für Geräte: Implementieren Sie die Remote-Verwaltung für Geräte, um die Anzahl der erforderlichen Geschäftsreisen zu reduzieren.
  - AWS Systems Manager Fleet Manager ist eine vereinheitlichte UI-Umgebung, mit der Sie Ihre in AWS oder On-Premises ausgeführten Knoten aus der Ferne überwachen können.

Ressourcen

Zugehörige Dokumente:

- What is Amazon WorkSpaces?

- Cost Optimizer für Amazon WorkSpaces
- Dokumentation zu Amazon AppStream 2.0
- NICE DCV

Zugehörige Videos:

- Managing cost for Amazon WorkSpaces on AWS

SUS02-BP06 Implementierung von Pufferung oder Drosselung, um die Bedarfskurve zu verflachen

Pufferung und Drosselung verflachen die Bedarfskurve und reduzieren die erforderliche bereitgestellte Kapazität für Ihre Workload.

Typische Anti-Muster:

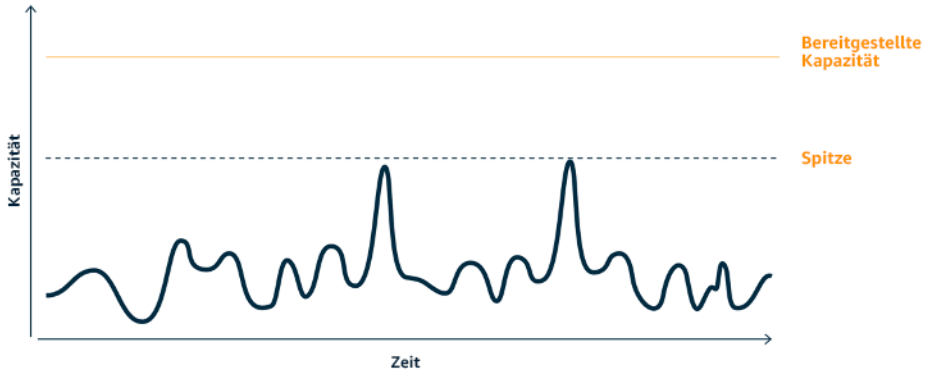
- Sie verarbeiten die Client-Anfragen sofort, obwohl dies nicht erforderlich ist.
- Sie analysieren die Anforderungen für Client-Anfragen nicht.

Vorteile der Nutzung dieser bewährten Methode: Das Verflachen der Bedarfskurve reduziert die erforderliche bereitgestellte Kapazität für die Workload. Die Reduzierung der bereitgestellten Kapazität bedeutet geringeren Energieverbrauch und geringere Umweltauswirkungen.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Niedrig

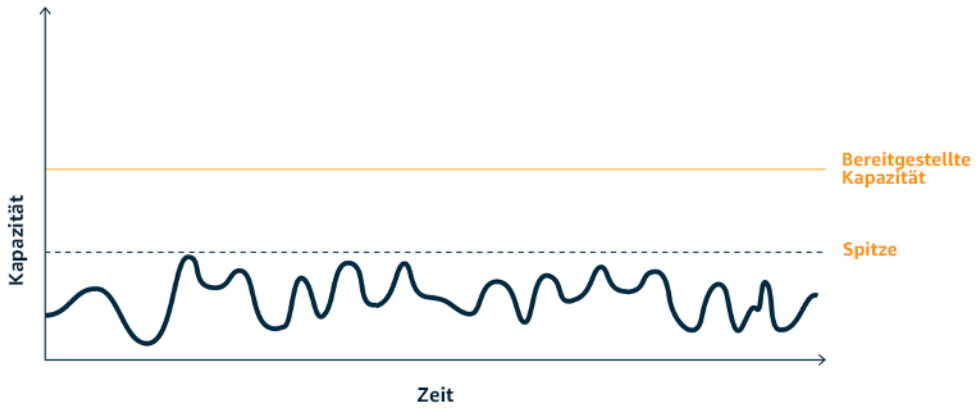
Implementierungsleitfaden

Die Verflachung der Bedarfskurve kann Ihnen dabei helfen, die bereitgestellte Kapazität für eine Workload zu verringern und dessen Umweltauswirkungen zu reduzieren. Nehmen wir eine Workload mit der nachfolgend gezeigten Bedarfskurve. Diese Workload hat zwei Spitzen und um damit umzugehen, wird die Ressourcenkapazität bereitgestellt, die hier durch die orangefarbene Linie angezeigt wird. Die für diese Workload aufgewendeten Ressourcen und die eingesetzte Energie werden nicht durch die Fläche unter der Bedarfskurve, sondern von der Linie für die bereitgestellte Kapazität angezeigt, da für den Umgang mit den beiden Spitzen bereitgestellte Kapazität erforderlich ist.



Bedarfskurve mit zwei deutlichen Spitzen, die hohe bereitgestellte Kapazität erfordern

Sie können Pufferung oder Drosselung verwenden, um die Bedarfskurve zu beeinflussen und die Spitzen abzumildern, was weniger bereitgestellte Kapazität und einen geringeren Energieverbrauch bedeutet. Implementieren Sie Drosselung, wenn Ihre Clients wiederholte Versuche durchführen können. Implementieren Sie die Pufferung, um die Anforderung zu speichern und die Verarbeitung auf einen späteren Zeitpunkt zu verschieben.



Auswirkung der Drosselung auf die Nachfragekurve und die bereitgestellte Kapazität

Implementierungsschritte

- Analysieren Sie die Client-Anfragen, um festzulegen, wie darauf zu reagieren ist. Wichtige Faktoren dabei sind:
  - Kann diese Anfrage in asynchroner Weise verarbeitet werden?
  - Kann der Client die Anfrage erneut versuchen?

- Wenn dies der Fall ist, können Sie Drosselung verwenden, die der Quelle mitteilt, dass wenn sie die Anfrage zum aktuellen Zeitpunkt nicht bedienen kann, es später erneut versucht werden sollte.
  - Sie können Amazon API Gateway verwenden, um Drosselung zu implementieren.
- Für Clients, die Anfragen nicht erneut versuchen können, muss zur Verflachung der Bedarfskurve ein Puffer implementiert werden. Ein Puffer verschiebt die Anforderungsverarbeitung, so dass Anwendungen, die mit unterschiedlichen Raten ausgeführt werden, effektiv kommunizieren können. Bei der Pufferung werden Nachrichten von Produzenten in eine Warteschlange oder einen Stream gestellt. Nachrichten können dadurch von Verbrauchern in der für ihre Geschäftsanforderungen passenden Geschwindigkeit gelesen und verarbeitet werden.
  - Amazon Simple Queue Service (Amazon SQS) ist ein verwalteter Service, der Warteschlangen bietet, die es einem einzelnen Verbraucher ermöglichen, individuelle Nachrichten zu lesen.
  - Amazon Kinesis stellt einen Stream bereit, mit dem viele Verbraucher dieselben Nachrichten lesen können.
- Analysieren Sie den Gesamtbedarf, die Änderungsrate und die erforderliche Reaktionszeit, um die korrekte Größe der erforderlichen Drosselung oder des Puffers zu bestimmen.

Ressourcen

Zugehörige Dokumente:

- Getting started with Amazon SQS
- Application integration Using Queues and Messages
- Managing and monitoring API throttling in your workloads
- Throttling a tiered, multi-tenant REST API at scale using API Gateway
- Application integration Using Queues and Messages

Zugehörige Videos:

- AWS re:Invent 2.022 - Application integration patterns for microservices
- AWS re:Invent 2.023 - Smart savings: Amazon EC2 cost-optimization strategies
- AWS re:Invent 2.023 - Advanced integration patterns & trade-offs for loosely coupled systems

Software und Architektur

Implementieren Sie Muster für den Lastausgleich und die Wahrung einer konsistent hohen Nutzung der bereitgestellten Ressourcen, um die Zahl der genutzten Ressourcen zu minimieren. Komponenten werden möglicherweise aufgrund von Änderungen des Benutzerverhaltens über die Zeit nicht mehr genutzt. Prüfen Sie Muster und Architekturen, um nicht ausreichend genutzte Komponenten zu konsolidieren und so die Nutzung insgesamt zu erhöhen. Nehmen Sie Komponenten außer Betrieb, die nicht mehr benötigt werden. Identifizieren Sie die Leistung Ihrer Workload-Komponenten und optimieren Sie die Komponenten, die die meisten Ressourcen verbrauchen. Achten Sie auf die Geräte, mit denen Ihre Kunden auf Ihre Services zugreifen, und implementieren Sie Muster, um den Bedarf für Geräte-Upgrades zu minimieren.

Bewährte Methoden

- SUS03-BP01 Optimieren von Software und Architektur für asynchrone und geplante Aufträge
- SUS03-BP02 Entfernen oder Faktorwechsel von Workload-Komponenten mit geringer oder keiner Nutzung
- SUS03-BP03 Optimieren von Codebereichen, die die meiste Zeit oder die meisten Ressourcen verbrauchen
- SUS03-BP04 Optimieren der Auswirkungen auf Geräte und Ausrüstung von Kunden
- SUS03-BP05 Verwenden von Softwaremustern und Architekturen, die Datenzugriffs- und Speichermuster optimal unterstützen

SUS03-BP01 Optimieren von Software und Architektur für asynchrone und geplante Aufträge

Verwenden Sie effiziente Software- und Architekturmuster wie warteschlangenbasierte Systeme, um eine durchgängig hohe Auslastung von bereitgestellten Ressourcen zu erzielen.

Typische Anti-Muster:

- Sie stellen zu viele Ressourcen in der Cloud-Workload bereit, um auf unerwartete Nachfragesteigerungen reagieren zu können.
- In Ihrer Architektur werden Absender und Empfänger von asynchronen Nachrichten nicht durch eine Messaging-Komponente entkoppelt.

Vorteile der Nutzung dieser bewährten Methode:

- Durch effiziente Software- und Architekturmuster werden ungenutzte Ressourcen in Ihrer Workload minimiert und die allgemeine Effizienz gesteigert.
- Sie können die Verarbeitung unabhängig vom Empfang asynchroner Nachrichten skalieren.
- Durch eine Messaging-Komponente gelten weniger strenge Verfügbarkeitsanforderungen, die mit weniger Ressourcen erfüllt werden können.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Mittel

Implementierungsleitfaden

Verwenden Sie effiziente Architekturmuster wie eine ereignisgesteuerte Architektur, die zu einer gleichmäßigen Nutzung der Komponenten führen und die Überbereitstellung in Ihrer Workload minimieren. Durch die Verwendung effizienter Architekturmuster werden ungenutzte Ressourcen, die aufgrund von Änderungen der Nachfrage im Laufe der Zeit nicht genutzt werden, minimiert.

Analysieren Sie die Anforderungen Ihrer Workload-Komponenten und führen Sie Architekturmuster ein, mit denen die allgemeine Auslastung der Ressourcen gesteigert wird. Nehmen Sie Komponenten außer Betrieb, die nicht mehr benötigt werden.

Implementierungsschritte

- Analysieren Sie die Nachfrage für Ihre Workload, um zu bestimmen, wie diese erfüllt werden kann.
- Verwenden Sie für Anfragen oder Aufträge, für die keine synchronen Antworten erforderlich sind, warteschlangenbasierte Architekturen und Worker mit Auto Scaling, durch die die Auslastung maximiert wird. Hier finden Sie einige Beispiele für Situationen, in denen Sie eine warteschlangenbasierte Architektur in Erwägung ziehen sollten:

Warteschlangenmechanismus	Beschreibung
AWS Batch-Auftragswarteschlangen	AWS Batch-Aufträge werden an eine Auftragswarteschlange gesendet, in der sie bleiben, bis ihre Ausführung in einer Datenverarbeitungsumgebung geplant werden kann.

Warteschlangenmechanismus	Beschreibung
Amazon Simple Queue Service und Amazon EC2 Spot Instances	Durch das Koppeln von Amazon SQS und Spot Instances lassen sich fehlertolerante und effiziente Architekturen erstellen.

- Verwenden Sie für Anfragen oder Aufträge, die jederzeit verarbeitet werden können, Planungsmechanismen zur Auftragsverarbeitung in Batches, um die Effizienz zu steigern. Hier sind einige Beispiele für Planungsmechanismen in AWS:

Zeitplanungsmechanismus	Beschreibung
Amazon EventBridge Scheduler	Eine Funktion von Amazon EventBridge, mit der Sie in großem Umfang geplante Aufgaben erstellen, ausführen und verwalten können.
Zeitbasierter AWS Glue-Plan	Hiermit definieren Sie einen zeitbasierten Plan für Crawler und Aufträge in AWS Glue.
Geplante Aufgaben von Amazon Elastic Container Service (Amazon ECS)	Amazon ECS unterstützt das Erstellen geplanter Tasks. Geplante Tasks verwenden Amazon EventBridge Regeln, um Tasks entweder nach einem Zeitplan oder als Antwort auf ein EventBridge-Ereignis auszuführen.
Instance Scheduler	Konfigurieren Sie Zeitpläne zum Starten und Beenden Ihrer Instances von Amazon EC2 und Amazon Relational Database Service.

- Wenn Sie Abfrage- und Webhook-Mechanismen in Ihrer Architektur verwenden, ersetzen Sie diese durch Ereignisse. Erstellen Sie mit ereignisgesteuerten Architekturen hocheffiziente Workloads.
- Nutzen Sie Serverless in AWS, um eine übermäßige Bereitstellung in einer Infrastruktur zu eliminieren.
- Wählen Sie die richtige Größe für Ihre Architektur, um zu vermeiden, dass ungenutzte Ressourcen auf Eingaben warten.
  - Sie können die Empfehlungen zur Dimensionierung in AWS Cost Explorer oder AWS Compute Optimizer zur Identifizierung von Dimensionierungsmöglichkeiten verwenden.

- Weitere Informationen finden Sie unter Größenanpassung: Bereitstellung von an Workloads angepassten Instances.

Ressourcen

Zugehörige Dokumente:

- What is Amazon Simple Queue Service?
- What is Amazon MQ?
- Scaling based on Amazon SQS
- Was ist AWS Step Functions?
- Was ist AWS Lambda?
- Using AWS Lambda with Amazon SQS
- What is Amazon EventBridge?
- Managing Asynchronous Workflows with a REST API

Zugehörige Videos:

- AWS re:Invent 2.023 - Navigating the journey to serverless event-driven architecture
- AWS re:Invent 2.023 - Using serverless for event-driven architecture & domain-driven design
- AWS re:Invent 2.023 - Advanced event-driven patterns with Amazon EventBridge
- AWS re:Invent 2.023 - Sustainable architecture: Past, present, and future
- Asynchronous Message Patterns | AWS Events

Zugehörige Beispiele:

- Event-driven architecture with AWS Graviton Processors and Amazon EC2 Spot Instances

SUS03-BP02 Entfernen oder Faktorwechsel von Workload-Komponenten mit geringer oder keiner Nutzung

Entfernen Sie ungenutzte Komponenten, die nicht mehr benötigt werden, und führen Sie einen Faktorwechsel für Komponenten mit geringer Nutzung durch, um die Verschwendung von Ressourcen in Ihrer Workload zu begrenzen.

Typische Anti-Muster:

- Sie prüfen den Nutzungsgrad der einzelnen Komponenten Ihrer Workload nicht regelmäßig.
- Sie prüfen und analysieren nicht die Empfehlungen von AWS-Dimensionierungstools wie etwa AWS Compute Optimizer.

Vorteile der Nutzung dieser bewährten Methode: Das Entfernen nicht genutzter Komponenten minimiert Ausschuss und verbessert die allgemeine Effizienz Ihrer Workload.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Mittel

Implementierungsleitfaden

Prüfen Sie Ihre Workload, um nicht oder wenig genutzte Komponenten zu identifizieren. Dies ist ein sich wiederholender Verbesserungsprozess, der von Änderungen beim Bedarf oder der Einführung eines neuen Cloud-Services ausgelöst werden kann. Beispielsweise kann ein deutliches Zurückgehen der Laufzeit der AWS Lambda-Funktion darauf hindeuten, dass die Speichergröße reduziert werden muss. Außerdem können sich die optimalen Services und die Architektur für Ihre Workload ändern, wenn AWS neue Services und Features veröffentlicht.

Überwachen Sie kontinuierlich die Workload-Aktivität und suchen Sie nach Möglichkeiten zur Verbesserung des Nutzungsgrads einzelner Komponenten. Wenn Sie nicht genutzte Komponenten entfernen und Dimensionierungsaktivitäten durchführen, erreichen Sie Ihre geschäftlichen Ziele mit der geringstmöglichen Menge von Cloud-Ressourcen.

Implementierungsschritte

- Machen Sie eine Bestandsaufnahme Ihrer AWS-Ressourcen. In AWS können Sie AWS Ressourcen Explorer einschalten, um Ihre AWS-Ressourcen zu erkunden und zu organisieren. Weitere Informationen finden Sie unter AWS re:Invent 2.022 - How to manage resources and applications at scale on AWS.
- Überwachen und erfassen Sie die Nutzungsmetriken für kritische Komponenten Ihrer Workloads (etwa CPU-Nutzung, Speichernutzung oder Netzwerkdurchsatz in Amazon-CloudWatch-Metriken).
- Identifizieren Sie ungenutzte oder zu wenig genutzte Komponenten in Ihrer Architektur.
  - Prüfen Sie für stabile Workloads regelmäßig AWS-Dimensionierungstools wie AWS Compute Optimizer, um nicht oder wenig genutzte Komponenten zu identifizieren.
  - Prüfen Sie für kurzzeitige Workloads die Nutzungsmetriken, um nicht oder wenig genutzte Komponenten zu identifizieren.

- Nehmen Sie nicht mehr benötigte und dazugehörige Ressourcen (wie etwa Amazon-ECR-Images) außer Betrieb.
  - Automatisierte Bereinigung von nicht verwendeten Images in Amazon ECR
  - Löschen von ungenutzten Amazon Elastic Block Store (Amazon EBS)-Volumes mit AWS Config und AWS Systems Manager
- Führen Sie einen Faktorwechsel für nicht ausreichend genutzte Ressourcen durch oder konsolidieren Sie sie mit anderen Ressourcen, um die Nutzungseffizienz zu verbessern. Sie können beispielsweise mehrere kleine Datenbanken auf einer einzelnen Amazon-RDS-Datenbank-Instance bereitstellen, anstatt Datenbanken auf einzelnen sehr wenig ausgenutzten Instances auszuführen.
- Machen Sie sich mit den Ressourcen vertraut, die durch Ihre Workload bereitgestellt werden, um eine Arbeitseinheit zu erledigen.

Ressourcen

Zugehörige Dokumente:

- AWS Trusted Advisor
- Was ist Amazon CloudWatch?
- Größenanpassung: Bereitstellung von an Workloads angepassten Instances
- Optimizing your cost with Rightsizing Recommendations

Zugehörige Videos:

- AWS re:Invent 2.023 - Capacity, availability, cost efficiency: Pick three

Zugehörige Beispiele:

- Optimize Hardware Patterns and Observe Sustainability KPIs

SUS03-BP03 Optimieren von Codebereichen, die die meiste Zeit oder die meisten Ressourcen verbrauchen

Optimieren Sie den Code, der innerhalb der verschiedenen Komponenten Ihrer Architektur ausgeführt wird, um die Ressourcennutzung zu minimieren und die Leistung zu maximieren.

Typische Anti-Muster:

- Sie versäumen die Optimierung Ihres Codes für die Ressourcennutzung.
- Sie reagieren auf Leistungsprobleme normalerweise mit Erhöhung des Ressourceneinsatzes.
- Ihr Code-Prüfungs- und -Entwicklungsprozess verfolgt keine Leistungsänderungen.

Vorteile der Nutzung dieser bewährten Methode: Die Verwendung effizienten Codes minimiert die Ressourcennutzung und verbessert die Leistung.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Mittel

Implementierungsleitfaden

Es ist sehr wichtig, jeden funktionalen Bereich, einschließlich des Codes einer für die Cloud erstellten Anwendung, zu untersuchen, um ihre Ressourcennutzung und Leistung zu optimieren. Überwachen Sie kontinuierlich die Leistung Ihrer Workload in Build-Umgebungen und Produktionsbereichen und suchen Sie nach Möglichkeiten, Codeausschnitte zu verbessern, die einen besonders hohen Ressourcenverbrauch haben. Führen Sie einen regelmäßigen Prüfungsprozess ein, um Fehler oder Anti-Muster in Ihrem Code zu identifizieren, die Ressourcen in ineffizienter Weise nutzen. Nutzen Sie einfache und effiziente Algorithmen, die dieselben Ergebnisse für Ihre Anwendungsfälle liefern.

Implementierungsschritte

- Effiziente Programmiersprache verwenden: Verwenden Sie das jeweils effizienteste Betriebssystem und die optimale Programmiersprache für die Workload. Weitere Informationen zu energieeffizienten Programmiersprachen (einschließlich Rust) finden Sie unter Sustainability with Rust.
- KI-Programmierungsbegleiter verwenden: Erwägen Sie die Verwendung eines Begleiters zur KI-Programmierung wie Amazon CodeWhisperer, um Code effizient zu schreiben.
- Code-Überprüfungen automatisieren: Führen Sie bei der Entwicklung Ihrer Workloads einen automatischen Code-Prüfungsprozess ein, um die Qualität zu verbessern sowie Fehler und Anti-Muster zu identifizieren.
  - Automatisieren von Code-Reviews mit Amazon CodeGuru Reviewer
  - Erkennen von Parallelitätsfehlern mit Amazon CodeGuru
  - Verbessern der Codequalität für Python-Anwendungen mit Amazon CodeGuru

- Code-Profiler verwenden: Verwenden Sie einen Code-Profiler für Code-Prüfungen, um die Codebereiche als Optimierungsziele zu identifizieren, die die meiste Zeit oder die meisten Ressourcen verwenden.
  - Reduzieren des CO_2-Fußabdrucks Ihrer Organisation mit Amazon CodeGuru Profiler
  - Verständnis der Speichernutzung in Ihrer Java-Anwendung mit Amazon CodeGuru Profiler
  - Verbessern des Kundenkomforts und Senken von Kosten mit Amazon CodeGuru Profiler
- Überwachen und optimieren: Verwenden Sie Ressourcen für die kontinuierliche Überwachung, um Komponenten mit hohem Ressourcenbedarf oder suboptimaler Konfiguration zu identifizieren.
  - Ersetzen Sie rechenintensive Algorithmen durch einfachere und effizientere Versionen, die dieselben Ergebnisse liefern.
  - Entfernen Sie unnötigen Code und überflüssige Formatierungen.
- Code-Faktorwechsel oder -Transformation verwenden: Erkunden Sie die Möglichkeiten der Amazon-Q-Codetransformation für die Wartung und Aktualisierung von Anwendungen.
  - Sprachversionen mit Amazon-Q-Codetransformation aktualisieren
  - AWSre:Invent 2023 – Automatisierung von App-Upgrades und -Wartung mithilfe von Amazon Q-Codetransformation

Ressourcen

Zugehörige Dokumente:

- What is Amazon CodeGuru Profiler?
- FPGA instances
- Die AWS-SDKs Tools zum Entwickeln in AWS

Zugehörige Videos:

- Improve Code Efficiency Using Amazon CodeGuru Profiler
- AWS re:Invent 2.023 - Best practices for Amazon CodeWhisperer
- Automate Code Reviews and Application Performance Recommendations with Amazon CodeGuru

Zugehörige Beispiele:

- Optimizing Code with Amazon CodeGuru

SUS03-BP04 Optimieren der Auswirkungen auf Geräte und Ausrüstung von Kunden

Verstehen Sie die in Ihrer Architektur verwendeten Geräte und nutzen Sie Strategien, um ihre Nutzung zu reduzieren. Dies kann die Umweltauswirkungen Ihrer Cloud-Workload insgesamt verringern.

Typische Anti-Muster:

- Sie ignorieren die Umweltauswirkungen der Geräte, die Ihre Kunden verwenden.
- Sie verwalten und aktualisieren die von Kunden verwendeten Ressourcen manuell.

Vorteile der Nutzung dieser bewährten Methode: Die Implementierung von Softwaremustern und Features, die für Kundengeräte optimiert sind, können die Umweltauswirkungen von Cloud-Workloads insgesamt verringern.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Mittel

Implementierungsleitfaden

Die Implementierung für Kundengeräte optimierter Softwaremuster und Features können die Umweltauswirkungen auf unterschiedliche Weise reduzieren:

- Die Implementierung neuer abwärtskompatibler Features kann die Anzahl der Hardwareaustauschvorgänge verringern.
- Die Optimierung einer Anwendung, so dass sie effizient auf Geräten ausgeführt werden kann, kann bei der Reduzierung des Energieverbrauchs helfen und die Batterielaufzeit verlängern (falls Batterien zum Einsatz kommen).
- Die Optimierung einer Anwendung für Geräte kann auch Datenübertragungen über das Netzwerk verringern.

Verstehen Sie die in Ihrer Architektur verwendeten Geräte, ihre erwartete Lebensdauer und die Auswirkungen des Austauschs dieser Komponenten. Implementieren Sie Softwaremuster und Features, die dabei helfen, den Energieverbrauch von Geräten zu senken, und den Austausch von Geräten sowie manuelle Upgrades durch Kunden seltener erforderlich machen.

Implementierungsschritte

- Bestandsaufnahme durchführen: Inventarisieren Sie die in ihrer Architektur verwendeten Geräte. Dabei kann es sich um Mobilgeräte, Tablets, IOT-Geräte, Smart Light- oder auch Smartgeräte in einer Fabrik handeln.
- Energieeffiziente Geräte verwenden: Erwägen Sie den Einsatz energieeffizienter Geräte in Ihrer Architektur. Verwenden Sie Energieverwaltungskonfigurationen auf Geräten, um in den Energiesparmodus zu wechseln, wenn sie nicht verwendet werden.
- Effiziente Anwendungen ausführen: Optimieren Sie die Anwendung, die auf den Geräten ausgeführt wird:
  - Verwenden Sie Strategien wie die Ausführung von Aufgaben im Hintergrund, um den Energieverbrauch zu verringern.
  - Berücksichtigen Sie beim Erstellen von Nutzlasten Netzwerkbandbreite und Latenz und implementieren Sie Funktionen, mit denen Ihre Anwendungen auch über Verbindungen mit geringer Bandbreite und hoher Latenz gut funktionieren.
  - Wandeln Sie Nutzlasten und Dateien in von den Geräten benötigte optimierte Formate um. Sie können beispielsweise Amazon Elastic Transcoder oder AWS Elemental MediaConvert verwenden, um große, qualitativ hochwertige Digitalmediendateien in Formate umzuwandeln, die Benutzer auf Mobilgeräten abspielen können.
  - Führen Sie rechenintensive Aktivitäten (z. B. das Rendern von Bildern) serverseitig aus oder nutzen Sie Anwendungs-Streaming, um den Benutzerkomfort auf älteren Geräten zu verbessern.
  - Segmentieren und paginieren Sie Ausgaben, besonders für interaktive Sitzungen, um Nutzlasten zu verwalten und lokale Speicheranforderungen zu begrenzen.
- Anbieter einbeziehen: Arbeiten Sie mit Geräteanbietern zusammen, die nachhaltige Materialien verwenden und für Transparenz in ihren Lieferketten und Umweltzertifizierungen sorgen.
- Over-the-Air (OTA)-Updates verwenden: Verwenden Sie den automatisierten Over-the-Air (OTA)-Mechanismus, um Updates auf einem oder mehreren Geräten bereitzustellen.
  - Mit einer CI/CD-Pipeline können Sie mobile Anwendungen aktualisieren.
  - Mit AWS IoT Device Management können Sie verbundene Geräte in großem Umfang aus der Ferne verwalten.
- Verwaltete Gerätefarmen verwenden: Verwenden Sie zum Testen neuer Features und Updates verwaltete Gerätefarmen mit repräsentativen Sätzen von Hardwaregeräten, um den Umfang der unterstützten Geräte zu maximieren. Weitere Details finden Sie unter SUS06-BP04 Verwenden verwalteter Gerätefarmen für Tests.

- Kontinuierliche Überwachung und Verbesserung: Verfolgen Sie den Energieverbrauch von Geräten, um Verbesserungsmöglichkeiten zu identifizieren. Verwenden Sie neue Technologien oder bewährte Methoden, um die Umweltauswirkungen dieser Geräte zu verbessern.

Ressourcen

Zugehörige Dokumente:

- Was ist AWS Device Farm?
- AppStream 2.0-Dokumentation
- NICE DCV
- OTA-Tutorial zum Aktualisieren der Firmware auf Geräten, auf denen FreeRTOS ausgeführt wird
- Optimizing Your IoT Devices for Environmental Sustainability

Zugehörige Videos:

- AWS re:Invent 2.023 - Improve your mobile and web app quality using AWS Device Farm

SUS03-BP05 Verwenden von Softwaremustern und Architekturen, die Datenzugriffs- und Speichermuster optimal unterstützen

Identifizieren Sie, wie Daten in Ihrer Workload verwendet, von Benutzern genutzt, übertragen und gespeichert werden. Verwenden Sie Softwaremuster und Architekturen, die den Datenzugriff und die Speicherung optimal unterstützen, um die zur Unterstützung der Workload erforderlichen Datenverarbeitungs-, Netzwerk- und Speicherressourcen zu reduzieren.

Typische Anti-Muster:

- Sie gehen davon aus, dass für alle Workloads ähnliche Datenspeicher- und Zugriffsmuster gelten.
- Sie verwenden nur eine Speicherebene, vorausgesetzt, dass alle Workloads in diese Ebene passen.
- Sie gehen davon aus, dass Datenzugriffsmuster im Laufe der Zeit konsistent bleiben.
- Ihre Architektur unterstützt potenzielle hohe Bursts beim Datenzugriff, was dazu führt, dass die Ressourcen die meiste Zeit ungenutzt bleiben.

Vorteile der Nutzung dieser bewährten Methode: Die Auswahl und Optimierung Ihrer Architektur auf der Grundlage von Datenzugriffs- und Speichermustern hilft bei der Reduzierung der Entwicklungskomplexität und der Steigerung der allgemeinen Nutzung. Das Verständnis, wann globale Tabellen, Datenpartitionen und Caching verwendet werden sollen, hilft Ihnen dabei, den Betriebsaufwand zu verringern und basierend auf Ihren Workload-Anforderungen zu skalieren.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Mittel

Implementierungsleitfaden

Verwenden Sie Software- und Architekturmuster, die optimal zu den Eigenschaften Ihrer Daten und den Zugriffsmustern passen. Verwenden Sie etwa eine moderne Datenarchitektur in AWS, die die Nutzung speziell erstellter Services ermöglicht, die für Ihre ganz speziellen Analytikanwendungsfälle optimiert sind. Diese Architekturmuster ermöglichen die effiziente Datenverarbeitung und verringern die Ressourcennutzung.

Implementierungsschritte

- Analysieren Sie die Eigenschaften ihrer Daten und Ihre Zugriffsmuster, um die korrekte Konfiguration für Ihre Cloud-Ressourcen zu identifizieren. Zu den berücksichtigenden Schlüsselmerkmalen gehören:
  - Datentyp: strukturiert, semistrukturiert, unstrukturiert
  - Datenwachstum: begrenzt, unbegrenzt
  - Lebensdauer von Daten: anhaltend, flüchtig, vorübergehend
  - Zugriffsmuster: Lese- oder Schreibzugriff, Häufigkeit von Aktualisierungen, schwankend oder konsistent
- Verwenden Sie Architekturmuster, die Datenzugriffs- und Speichermuster optimal unterstützen.
  - Muster zur Aktivierung der Datenpersistenz
  - Let’s Architect! Moderne Datenarchitekturen
  - Datenbanken in AWS: Das richtige Tool für jede Aufgabe
- Nutzen Sie Technologien, die nativ mit komprimierten Daten funktionieren.
  - Dateiformate von Athena Compression Support
  - Mögliche Formate für ETL-Eingaben und -Ausgaben in AWS Glue
  - Laden komprimierter Datendateien aus Amazon S3 mit Amazon Redshift

- Verwenden Sie zweckgerichtet erstellte Analytikservices für die Datenverarbeitung in Ihrer Architektur. Ausführlichere Informationen zu speziell entwickelten Analytikservices von AWS finden Sie unter AWS re:Invent 2.022 - Building modern data architectures on AWS.
- Verwenden Sie die Datenbank-Engine, die das dominierende Abfragemuster jeweils am besten unterstützt. Verwalten Sie Ihre Datenbankindizes so, dass sie die effiziente Ausführung von Abfragen unterstützen. Weitere Detaila finden Sie unter AWS-Datenbanken und AWS re:Invent 2.022 - Modernize apps with purpose-built databases.
- Wählen Sie Netzwerkprotokolle aus, die die Menge der genutzten Netzwerkkapazitäten in Ihrer Architektur reduzieren.

Ressourcen

Zugehörige Dokumente:

- COPY from columnar data formats with Amazon Redshift
- Converting Your Input Record Format in Firehose
- Improve query performance on Amazon Athena by Converting to Columnar Formats
- Monitoring DB load with Performance Insights on Amazon Aurora
- Monitoring DB load with Performance Insights on Amazon RDS
- Speicherklasse Amazon S3 Intelligent-Tiering
- Build a CQRS event store with Amazon DynamoDB

Zugehörige Videos:

- AWS re:Invent 2.022 - Building data mesh architectures on AWS
- AWS re:Invent 2.023 - Deep dive into Amazon Aurora and its innovations
- AWS re:Invent 2.023 - Improve Amazon EBS efficiency and be more cost-efficient
- AWS re:Invent 2.023 - Optimizing storage price and performance with Amazon S3
- AWS re:Invent 2.023 - Building and optimizing a data lake on Amazon S3
- AWS re:Invent 2.023 - Advanced event-driven patterns with Amazon EventBridge

Zugehörige Beispiele:

- AWS Purpose Built Databases Workshop

- AWS Modern Data Architecture Immersion Day
- Build a Data Mesh on AWS

Datenverwaltung

Implementieren Sie Verfahren für die Datenverwaltung, die den zur Unterstützung Ihrer Workload bereitgestellten Speicher und die für dessen Nutzung erforderlichen Ressourcen reduzieren. Identifizieren Sie Ihre Daten und verwenden Sie Speichertechnologien und Konfigurationen, die den Unternehmenswert und die Nutzung der Daten optimal unterstützen. Verschieben Sie die Daten während des Lebenszyklus zu effizienteren Speichern mit geringerer Leistung, wenn die Anforderungen abnehmen. Löschen Sie Daten, die nicht mehr benötigt werden.

Bewährte Methoden

- SUS04-BP01 Implementieren einer Richtlinie für die Klassifizierung von Daten
- SUS04-BP02 Verwenden von Technologien, die Datenzugriff und Speichermuster unterstützen
- SUS04-BP03 Verwalten des Lebenszyklus von Datensätzen mithilfe von Richtlinien
- SUS04-BP04 Verwendung von Elastizität und Automatisierung zur Erweiterung des Blockspeichers oder des Dateisystems
- SUS04-BP05 Entfernen nicht benötigter oder redundanter Daten
- SUS04-BP06 Verwenden geteilter Dateisysteme oder Objektspeicher für den Zugriff auf allgemeine Daten
- SUS04-BP07 Minimieren von Datenübertragungen zwischen Netzwerken
- SUS04-BP08 Sichern von Daten nur in dem Fall, wenn ihre erneute Erstellung schwierig ist

SUS04-BP01 Implementieren einer Richtlinie für die Klassifizierung von Daten

Klassifizieren Sie die Daten, um zu verstehen, wie wichtig sie für die Geschäftsergebnisse sind, und wählen Sie die richtige energieeffiziente Speicherebene zur Speicherung der Daten.

Typische Anti-Muster:

- Sie identifizieren keine Datenbestände mit ähnlichen Merkmalen (z. B. Sensibilität, Geschäftskritikalität oder gesetzliche Anforderungen), die verarbeitet oder gespeichert werden.
- Sie haben keinen Datenkatalog zur Inventarisierung Ihrer Datenbestände eingeführt.

Vorteile der Nutzung dieser bewährten Methode: Durch die Implementierung einer Datenklassifizierungsrichtlinie können Sie die energieeffizienteste Speicherebene für Daten bestimmen.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Mittel

Implementierungsleitfaden

Bei der Datenklassifizierung wird identifiziert, welche Arten von Daten in einem Informationssystem verarbeitet und gespeichert werden, das einer Organisation gehört oder von ihr betrieben wird. Dazu gehört auch die Bestimmung der Kritikalität der Daten und der wahrscheinlichen Auswirkungen von Preisgaben, Verlusten oder Missbrauch von Daten.

Implementieren Sie Richtlinien zur Datenklassifizierung, indem Sie von der kontextuellen Verwendung der Daten ausgehen und ein Kategorisierungsschema erstellen, das den Grad der Kritikalität eines bestimmten Datensatzes für die Abläufe einer Organisation berücksichtigt.

Implementierungsschritte

- Bestandsaufnahme vornehmen: Führen Sie eine Bestandsaufnahme der verschiedenen Datentypen durch, die für Ihre Workload vorhanden sind.
- Daten gruppieren: Bestimmen Sie die Kritikalität, Vertraulichkeit, Integrität und Verfügbarkeit von Daten auf der Grundlage des Risikos für die Organisation. Verwenden Sie diese Anforderungen, um Daten in eine der von Ihnen gewählten Datenklassifizierungsebenen einzuteilen. Ein Beispiel finden Sie unter Vier einfache Schritte zur Klassifizierung Ihrer Daten und zur Sicherung Ihres Startups.
- Datenklassifizierungsebenen und Richtlinien definieren: Definieren Sie für jede Datengruppe die Datenklassifizierungsebene (z. B. öffentlich oder vertraulich) und die Verarbeitungsrichtlinien. Kennzeichnen Sie Daten entsprechend. Einzelheiten zu den Kategorien für die Datenklassifizierung finden Sie im Whitepaper zur Datenklassifizierung.
- Regelmäßige Überprüfung: Überprüfen und kontrollieren Sie Ihre Umgebung regelmäßig auf nicht markierte und nicht klassifizierte Daten. Verwenden Sie die Automatisierung, um diese Daten zu identifizieren und die Daten entsprechend zu klassifizieren und zu markieren. Ein Beispiel finden Sie unter Data Catalog and crawlers in AWS Glue.
- Datenkatalog einrichten: Richten Sie einen Datenkatalog mit Prüfungs- und Governance-Funktionen ein.
- Dokumentation: Dokumentieren Sie Datenklassifizierungsrichtlinien und Verarbeitungsverfahren für jede Datenklasse.

Ressourcen

Zugehörige Dokumente:

- Nutzung von AWS Cloud zur Unterstützung der Datenklassifizierung
- Tag-Richtlinien von AWS Organizations

Zugehörige Videos:

- AWS re:Invent 2.022 - Enabling agility with data governance on AWS
- AWS re:Invent 2.023 - Data protection and resilience with AWS storage

SUS04-BP02 Verwenden von Technologien, die Datenzugriff und Speichermuster unterstützen

Nutzen Sie Speichertechnologien, die den Zugriff auf Ihre Daten und ihre Speicherung jeweils optimal unterstützen, um die Zahl der bereitgestellten Ressourcen zu minimieren und gleichzeitig den Workload zu unterstützen.

Typische Anti-Muster:

- Sie gehen davon aus, dass für alle Workloads ähnliche Datenspeicher- und Zugriffsmuster gelten.
- Sie verwenden nur eine Speicherebene, vorausgesetzt, dass alle Workloads in diese Ebene passen.
- Sie gehen davon aus, dass Datenzugriffsmuster im Laufe der Zeit konsistent bleiben.

Vorteile der Nutzung dieser bewährten Methode: Die Auswahl und Optimierung Ihrer Speichertechnologien auf der Grundlage von Datenzugriffs- und Speichermustern hilft Ihnen, die erforderlichen Cloud-Ressourcen zu reduzieren, um Ihre Geschäftsanforderungen zu erfüllen und die Gesamteffizienz der Cloud-Workload zu verbessern.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Niedrig

Implementierungsleitfaden

Wählen Sie für maximale Leistungseffizienz die für Ihre Zugriffsmuster geeignete Speicherlösung, oder passen Sie Ihre Zugriffsmuster an die Speicherlösung an.

Implementierungsschritte

- Daten- und Zugriffsmerkmale bewerten: Bewerten Sie Ihre Datenmerkmale und Zugriffsmuster, um die wichtigsten Merkmale Ihres Speicherbedarfs zu erfassen. Zu den berücksichtigenden Schlüsselmerkmalen gehören:
  - Datentyp: strukturiert, semistrukturiert, unstrukturiert
  - Datenwachstum: begrenzt, unbegrenzt
  - Lebensdauer von Daten: anhaltend, flüchtig, vorübergehend
  - Zugriffsmuster: Lese- oder Schreibzugriff, Häufigkeit, schwankend oder konsistent
- Die richtige Speichertechnologie auswählen: Migrieren Sie Daten zur geeigneten Speichertechnologie, die Ihre Datenmerkmale und Zugriffsmuster unterstützt. Hier sind einige Beispiele für AWS-Speichertechnologien und ihre Schlüsselmerkmale:

Typ	Technologie	Schlüsselmerkmale
Objektspeicher	Amazon S3	Ein Objektspeicherservice mit unbegrenzter Skalierbarkeit, hoher Verfügbarkeit und mehreren Zugriffsoptionen. Für die Übertragung von Objekten in und aus Amazon S3 und den Zugriff auf diese Objekte können Sie einen Service wie z. B. Transfer Acceleration oder Zugangspunkte verwenden, um Ihren Standort, Ihre Sicherheitsanforderungen und Zugriffsmuster zu unterstützen.
Archivieren von Speichern	Amazon S3 Glacier	Speicherklasse von Amazon S3 für die Datenarchivierung.
Gemeinsames Dateisystem	Amazon Elastic File System (Amazon EFS)	Bereitstellbares Dateisystem, auf das mehrere Arten

Typ	Technologie	Schlüsselmerkmale
		von Datenverarbeitungs lösungen zugreifen können. Amazon EFS erweitert und verringert Speicher automatis ch und ist leistungsoptimiert , um durchgängig niedrige Latenzen zu bieten.
Gemeinsames Dateisystem	Amazon FSx	Basiert auf den aktuellen AWS-Datenverarbeit ungslösungen und unterstüt zt vier gängige Dateisyst eme: NetApp ONTAP, OpenZFS, Windows File Server und Lustre. Die Amazon FSx-Latenz, der Durchsatz und die IOPS variieren je nach Dateisystem und sollten bei der Auswahl des richtigen Dateisystems für Ihre Workload-Anforderu ngen berücksichtigt werden.
Blockspeicher	Amazon Elastic Block Store (Amazon EBS)	Skalierbarer, hochleist ungsfähiger Blockspei cherservice für Amazon Elastic Compute Cloud (Amazon EC2). Amazon EBS umfasst SSD-gestü tzten Speicher für transakti ons- und IOPS-intensive Workloads und HDD-gestü tzten Speicher für durchsatz intensive Workloads.

Typ	Technologie	Schlüsselmerkmale
Relationale Datenbank	Amazon Aurora, Amazon RDS, Amazon Redshift	Sie unterstützt AKID-Transaktionen (Atomarität, Konsistenz, Isolation und Dauerhaftigkeit) und gewährleistet die referentielle Integrität sowie eine starke Datenkonsistenz. Bei zahlreichen herkömmlichen Anwendungen, Enterprise Resource Planning (ERP), Customer Relationship Management (CRM) und E-Commerce-Systemen werden relationale Datenbanken zum Speichern der Daten verwendet.
Schlüssel-Werte-Datenbank	Amazon-DynamoDB	Für gängige Zugriffsmuster optimiert, üblicherweise zum Speichern und Abrufen großer Datenmengen. Web-Apps mit hohem Datenverkehr, E-Commerce-Systeme und Gaming-Anwendungen sind typische Anwendungsfälle für Schlüssel-Werte-Datenbanken.

- Speicherzuweisung automatisieren: Überwachen Sie bei Speichersystemen mit einer festen Größe, z. B. Amazon EBS oder Amazon FSx, den verfügbaren Speicherplatz und automatisieren die Speicherzuweisung bei Erreichen eines Schwellenwertes. Sie können mithilfe von Amazon CloudWatch verschiedene Metriken für Amazon EBS und Amazon FSx erfassen und analysieren.
- Die richtige Speicherklasse wählen: Wählen Sie die passende Speicherklasse für Ihre Daten.
  - Amazon-S3-Speicherklassen können auf Objektebene konfiguriert werden. Ein einzelner Bucket kann Objekte enthalten, die in allen Speicherklassen gespeichert sind.

- Sie können Amazon-S3-Lebenszyklusrichtlinien verwenden, um Objekte automatisch zwischen Speicherklassen zu wechseln oder Daten zu entfernen, ohne dass die Anwendung geändert werden muss. Im Allgemeinen müssen Sie bei diesen Speichermechanismen einen Kompromiss zwischen Ressourceneffizienz, Zugriffslatenz und Zuverlässigkeit eingehen.

Ressourcen

Zugehörige Dokumente:

- Amazon-EBS-Volume-Typen
- Amazon-EC2-Instance-Speicher
- Amazon S3 Intelligent-Tiering
- Amazon-EBS-E/A-Merkmale
- Verwenden von Amazon-S3-Speicherklassen
- Was ist Amazon S3 Glacier?

Zugehörige Videos:

- AWS re:Invent 2.023 - Improve Amazon EBS efficiency and be more cost-efficient
- AWS re:Invent 2.023 - Optimizing storage price and performance with Amazon S3
- AWS re:Invent 2.023 - Building and optimizing a data lake on Amazon S3
- AWS re:Invent 2.022 - Building modern data architectures on AWS
- AWS re:Invent 2.022 - Modernize apps with purpose-built databases
- AWS re:Invent 2.022 - Building data mesh architectures on AWS
- AWS re:Invent 2.023 - Deep dive into Amazon Aurora and its innovations
- AWS re:Invent 2.023 - Advanced data modeling with Amazon DynamoDB

Zugehörige Beispiele:

- Amazon-S3-Beispiele
- AWS Purpose Built Databases Workshop
- Databases for Developers
- AWS Modern Data Architecture Immersion Day

- Build a Data Mesh on AWS

SUS04-BP03 Verwalten des Lebenszyklus von Datensätzen mithilfe von Richtlinien

Verwalten Sie den Lebenszyklus aller Daten und setzen Sie automatisch Löschen durch, um den für Ihre Workload benötigten Speicher insgesamt zu minimieren.

Typische Anti-Muster:

- Sie löschen Daten manuell.
- Sie löschen keine Workload-Daten.
- Sie verschieben Daten nicht abhängig von den Aufbewahrungs- und Zugriffsanforderungen in energieeffizientere Speicherebenen.

Vorteile der Einführung dieser bewährten Methode: Durch Richtlinien für den Lebenszyklus wird die Effizienz des Datenzugriffs und der Datenaufbewahrung für eine Workload sichergestellt.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Mittel

Implementierungsleitfaden

Datensätze verfügen während ihres Lebenszyklus normalerweise über unterschiedliche Aufbewahrungs- und Zugriffsanforderungen. So kann eine Anwendung z. B. für einen bestimmten Zeitraum häufig Zugriff auf einige Datensätze benötigen. Danach wird nur noch unregelmäßig darauf zugegriffen.

Um Datensätze während ihres Lebenszyklus effizient zu verwalten, konfigurieren Sie Lebenszyklusrichtlinien, d. h. Regeln, die den Umgang mit den Datensätzen definieren.

Mit Lebenszyklus-Konfigurationsregeln können Sie einen bestimmten Speicherservice anweisen, einen Datensatz in energieeffizientere Speicherebenen zu verschieben, ihn zu archivieren oder zu löschen.

Implementierungsschritte

- Klassifizieren Sie Datensätze in Ihrer Workload.
- Definieren Sie Bearbeitungsverfahren für jede Datenklasse.

- Legen Sie automatisierte Lebenszyklusrichtlinien zur Durchsetzung von Lebenszyklusregeln fest. Hier finden Sie einige Beispiel zum Einrichten von automatisierten Lebenszyklusrichtlinien für unterschiedliche AWS-Speicherservices:

Storage Service	Festlegen von automatisierten Lebenszyklusrichtlinien
Amazon S3	Mit Amazon-S3-Lebenszyklen können Sie Ihre Objekte während ihres gesamten Lebenszyklus verwalten. Wenn die Zugriffsmuster unbekannt oder nicht prognostizierbar sind oder sich ändern, können Sie Amazon S3 Intelligent-Tiering verwenden. Hiermit werden Zugriffsmuster überwacht und Objekte, auf die nicht zugegriffen wurde, automatisch in kostengünstigere Zugriffsebenen verschoben. Anhand der Metriken von Amazon S3 Storage Lens können Sie Optimierungsmöglichkeiten und Lücken im Lebenszyklusmanagement ermitteln.
Amazon Elastic Block Store	Mit Amazon Data Lifecycle Manager lassen sich das Erstellen, Aufbewahren und Löschen von Amazon-EBS-Snapshots und Amazon-EBS-gestützten AMIs automatisieren.
Amazon Elastic File System	Das Amazon-EFS-Lebenszyklusmanagement verwaltet den Dateispeicher für Ihre Dateisysteme automatisch.
Amazon Elastic Container Registry	Amazon-ECR-Lebenszyklusrichtlinien automatisieren die Bereinigung von Container-Images, indem Images abhängig von Alter oder Anzahl ablaufen.

Storage Service	Festlegen von automatisierten Lebenszyklusrichtlinien
AWS Elemental MediaStore	Sie können eine Objektlebenszyklus-Richtlinie verwenden, die steuert, wie lange Objekte im MediaStore-Container gespeichert werden sollen.

- Löschen Sie nicht genutzte Volumes, Snapshots und Daten, deren Aufbewahrungszeitraum abgelaufen ist. Nutzen Sie zum Löschen native Service-Features wie die Gültigkeitsdauer von Amazon DynamoDB oder die Protokollaufbewahrung von Amazon CloudWatch.
- Aggregieren und komprimieren Sie Daten wenn möglich auf der Basis von Lebenszyklusregeln.

Ressourcen

Zugehörige Dokumente:

- Optimieren Ihrer Amazon-S3-Lebenszyklusregeln mit der Amazon-S3-Speicherklassenanalyse
- Evaluating Resources with AWS-Config-Regeln

Zugehörige Videos:

- AWS re:Invent 2.021 - Amazon S3 Lifecycle best practices to optimize your storage spend
- AWS re:Invent 2.023 - Optimizing storage price and performance with Amazon S3
- Simplify Your Data Lifecycle and Optimize Storage Costs With Amazon S3 Lifecycle
- Reduce Your Storage Costs Using Amazon S3 Storage Lens

SUS04-BP04 Verwendung von Elastizität und Automatisierung zur Erweiterung des Blockspeichers oder des Dateisystems

Verwenden Sie Elastizität und Automatisierung, um den Block-Speicher oder das Dateisystem zu erweitern, wenn das Datenvolumen zunimmt, um den bereitgestellten Gesamtspeicher zu minimieren.

Typische Anti-Muster:

- Sie unterhalten einen großen Block-Speicher oder ein großes Dateisystem für künftige Anforderungen.
- Sie stellen zu viele Input- und Output-Vorgänge pro Sekunde (IOPS) in Ihrem Dateisystem bereit.
- Sie überwachen die Nutzung Ihrer Daten-Volumes nicht.

Vorteile der Nutzung dieser bewährten Methode: Die Minimierung der übermäßigen Bereitstellung für das Speichersystem reduziert ungenutzte Ressourcen und verbessert die Gesamteffizienz Ihrer Workload.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Mittel

Implementierungsleitfaden

Erstellen Sie Block-Speicher und Dateisysteme mit Größenzuweisung, Durchsatz und Latenz, die den Anforderungen Ihrer Workloads entsprechen. Verwenden Sie Elastizität und Automatisierung, um den Block-Speicher oder das Dateisystem zu erweitern, wenn das Datenvolumen zunimmt, ohne dass diese Speicherservices übermäßig bereitgestellt werden.

Implementierungsschritte

- Stellen Sie bei Speichersystemen mit einer festen Größe wie Amazon EBS sicher, dass Sie die Menge des verwendeten Speichers im Vergleich zur Gesamtspeichergröße überwachen und nach Möglichkeit die Speichergröße beim Erreichen eines Schwellenwerts automatisch erhöhen.
- Verwenden Sie elastische Volumes und verwaltete Blockdaten-Services, um automatisch zusätzlichen Speicher zuzuweisen, wenn die Menge der persistenten Daten wächst. Sie können beispielsweise Amazon EBS Elastic Volumes verwenden, um Volume-Größe, Volume-Typ oder die Leistung Ihrer Amazon-EBS-Volumes zu modifizieren.
- Wählen Sie die korrekte Speicherklasse sowie den korrekten Leistungs- und Durchsatz-Modus für Ihr Dateisystem für Ihre geschäftlichen Anforderungen und überschreiten Sie diese nicht.
  - Amazon-EFS-Leistung
  - Leistung von Amazon-EBS-Volumes auf Linux-Instances
- Legen Sie Zielstufen für die Nutzung Ihrer Daten-Volumes fest und passen Sie die Größe von Volumes an, die außerhalb der erwarteten Bereiche liegen.
- Passen Sie die Größe schreibgeschützter Volumes an die Datenmenge an.
- Migrieren Sie Daten zu Objektspeichern, um zu vermeiden, dass die überschüssige Kapazität aus Volumes mit fester Größe im Blockspeicher bereitgestellt wird.

- Überprüfen Sie elastische Volumes und Dateisysteme, beenden Sie nicht genutzte und verkleinern Sie zu große Volumes, um sie an den aktuellen Datenumfang anzupassen.

Ressourcen

Zugehörige Dokumente:

- Extend the file system after resizing an EBS volume
- Modify a volume using Amazon EBS Elastic Volumes
- Dokumentation zu Amazon FSx
- What is Amazon Elastic File System?

Zugehörige Videos:

- Deep Dive on Amazon EBS Elastic Volumes
- Amazon EBS and Snapshot Optimization Strategies for Better Performance and Cost Savings
- Optimizing Amazon EFS for cost and performance, using best practices

SUS04-BP05 Entfernen nicht benötigter oder redundanter Daten

Entfernen Sie nicht benötigte oder redundante Daten, um die zum Speichern Ihrer Datensätze benötigten Speicherressourcen zu minimieren.

Typische Anti-Muster:

- Sie duplizieren Daten, die leicht abgerufen oder erneut erstellt werden können.
- Sie sichern alle Daten, ohne ihre Kritikalität zu berücksichtigen.
- Sie löschen Daten nur unregelmäßig, nur bei bestimmten Ereignissen oder gar nicht.
- Sie speichern Daten redundant, unabhängig von der Stabilität des Speicherservices.
- Sie aktivieren die Amazon-S3-Versionsverwaltung, ohne dass dies geschäftlich gerechtfertigt ist.

Vorteile der Einführung dieser bewährten Methode: Durch das Entfernen nicht benötigter Daten werden die für Ihre Workload benötigte Speichergröße und die Umweltbelastungen durch die Workload reduziert.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Mittel

Implementierungsleitfaden

Speichern Sie keine Daten, die Sie nicht benötigen. Automatisieren Sie das Löschen von nicht benötigten Daten. Verwenden Sie Technologien, die Daten auf Datei- und Blockebene deduplizieren. Nutzen Sie native Servicefeatures für Replikation und Redundanz.

Implementierungsschritte

- Bewerten Sie, ob Sie das Speichern von Daten vermeiden können, indem Sie vorhandene, öffentlich verfügbare Datensätze in AWS Data Exchange und offene Daten in AWS verwenden.
- Verwenden Sie Mechanismen, die Daten auf Block- und Objektebene deduplizieren können. Hier finden Sie einige Beispiele zum Deduplizieren von Daten in AWS:

Storage Service	Deduplizierungsmechanismus
Amazon S3	Verwenden Sie AWS Lake Formation FindMatches und das neue FindMatches ML Transform, um übereinstimmende Einträge in einem Datensatz zu finden (auch solche ohne Bezeichner).
Amazon FSx	Verwenden Sie die Datendeduplizierung in Amazon FSx für Windows.
Snapshots von Amazon Elastic Block Store	Snapshots sind inkrementelle Backups, d. h., es werden nur die Blöcke des Geräts gespeichert, die sich seit der letzten Snapshot-Speicherung geändert haben.

- Analysieren Sie den Datenzugriff, um nicht benötigte Daten zu identifizieren. Automatisieren Sie Lebenszyklusrichtlinien. Nutzen Sie zum Löschen native Service-Features wie die Gültigkeitsdauer von Amazon DynamoDB, Amazon-S3-Lebenszyklen oder die Protokollaufbewahrung von Amazon CloudWatch.
- Verwenden Sie Virtualisierungsfunktionen in AWS, um Daten an der Quelle beizubehalten und eine Duplikation zu vermeiden.
  - Cloudnative Datenvirtualisierung in AWS

  - Optimierung von Datenmustern mithilfe von Amazon Redshift Data Sharing
- Verwenden Sie Sicherungstechnologien, mit denen inkrementelle Sicherungen möglich sind.
- Nutzen Sie zum Erfüllen der Stabilitätsziele die Stabilität von Amazon S3 und Replikation von Amazon EBS anstelle von selbst verwalteten Technologien wie redundanten Arrays unabhängiger Datenträger (Redundant Array Of Independent Disks, RAID).
- Zentralisieren Sie Protokoll- und Nachverfolgungsdaten, deduplizieren Sie identische Protokolleinträge und richten Sie Mechanismen für die Anpassung der Ausführlichkeit ein, wenn notwendig.
- Füllen Sie Caches nur vorab aus, wenn dies begründet werden kann.
- Richten Sie Überwachung und Automatisierung für den Cache ein, um seine Größe entsprechend anzupassen.
- Entfernen Sie veraltete Bereitstellungen und Komponenten aus Objektspeichern und Edge-Zwischenspeichern, wenn Sie neue Versionen Ihrer Workload veröffentlichen.

Ressourcen

Zugehörige Dokumente:

- Ändern der Protokoll-Datenaufbewahrung in CloudWatch Logs
- Datendeduplizierung in Amazon FSx für Windows File Server
- Amazon FSxs von Amazon FSx für ONTAP einschließlich Datendeduplizierung
- Aufheben der Gültigkeit von Dateien in Amazon CloudFront
- Sichern und Wiederherstellen von Amazon-EFS-Dateisystemen mit AWS Backup
- What is Amazon CloudWatch Logs?
- Arbeiten mit Sicherungen in Amazon RDS
- Integrate and deduplicate datasets using AWS Lake Formation

Zugehörige Videos:

- Amazon Redshift Data Sharing Use Cases

Zugehörige Beispiele:

- Wie verwende ich Amazon Athena, um meine Amazon-S3-Serverzugriffsprotokolle zu analysieren?

SUS04-BP06 Verwenden geteilter Dateisysteme oder Objektspeicher für den Zugriff auf allgemeine Daten

Verwenden Sie geteilte Dateisysteme oder Speicher, um Datenduplizierungen zu vermeiden und eine effizientere Infrastruktur für Ihre Workload zu ermöglichen.

Typische Anti-Muster:

- Sie stellen für jeden einzelnen Client Speicher bereit.
- Sie trennen Daten-Volumes von inaktiven Clients nicht ab.
- Sie ermöglichen keinen Zugriff auf Speicher über Plattformen und Systeme hinweg.

Vorteile der Nutzung dieser bewährten Methode: Die Verwendung geteilter Dateisysteme oder Speicher ermöglicht die gemeinsame Nutzung von Daten für mehrere Verbraucher, ohne dass diese dazu kopiert werden müssen. Dies reduziert den Umfang der erforderlichen Speicherressourcen für die Workload.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Mittel

Implementierungsleitfaden

Wenn Sie mehrere Benutzer oder Anwendungen haben, die auf dieselben Datensätze zugreifen müssen, ist die Verwendung geteilter Speichertechnologien wichtig für eine effiziente Infrastruktur für Ihre Workload. Solche Technologien bieten einen zentralen Speicherort für die Speicherung und Verwaltung von Datensätzen und zur Vermeidung von Datenduplizierungen. Dazu wird die Konsistenz der Daten über verschiedene Systeme hinweg durchgesetzt. Hinzu kommt, dass geteilte Speicher die effizientere Nutzung der Datenverarbeitungsleistung ermöglichen, da mehr Datenverarbeitungsressourcen gleichzeitig auf Daten zugreifen und diese verarbeiten können.

Rufen Sie Daten von diesen geteilten Speicherservices nur bei Bedarf ab und trennen Sie nicht genutzte Volumes, um Ressourcen freizugeben.

Implementierungsschritte

- Migrieren Sie Daten in einen geteilten Speicher, wenn die Daten mehrfach genutzt werden. Hier sind einige Beispiele für geteilte Speichertechnologien in AWS:

Speicheroption	Wann sollte dies verwendet werden?
Amazon EBS Multi-Attach	Amazon EBS Multi-Attach ermöglicht die Anfügung eines einzelnen bereitgestellten IOPS SSD (io1 oder io2)-Volumes an mehrere Instances in derselben Availability Zone.
Amazon EFS	Siehe When to Choose Amazon EFS.
Amazon FSx	Siehe Auswählen eines Amazon-FSx-Dateisystems.
Amazon S3	Anwendungen, die keine Dateisystemstruktur benötigen und zur Arbeit mit Objektspeichern gedacht sind, können Amazon S3 als massive, skalierbare, dauerhafte und kostengünstige Speicherlösung nutzen.

- Kopieren Sie Daten bzw. rufen Sie sie nur dann von geteilten Dateisystemen ab, wenn Sie sie benötigen. Sie können beispielsweise ein Dateisystem von Amazon FSx für Lustre mit Unterstützung durch Amazon S3 erstellen und nur die Teilmenge der Daten laden, die für die Verarbeitung von Aufgaben zu Amazon FSx benötigt werden.
- Löschen Sie Daten entsprechend Ihren Nutzungsmustern, wie in SUS04-BP03 Verwalten des Lebenszyklus von Datensätzen mithilfe von Richtlinien erläutert.
- Trennen Sie Volumes von Clients, die sie nicht aktiv verwenden.

Ressourcen

Zugehörige Dokumente:

- Linking your file system to an Amazon S3 bucket
- Using Amazon EFS for AWS Lambda in your serverless applications
- Amazon EFS Intelligent-Tiering Optimizes Costs for Workloads with Changing Access Patterns
- Using Amazon FSx with your on-premises data repository

Zugehörige Videos:

- Storage cost optimization with Amazon EFS
- AWS re:Invent 2.023 - What's new with AWS file storage
- AWS re:Invent 2.023 - File storage for builders and data scientists on Amazon Elastic File System

SUS04-BP07 Minimieren von Datenübertragungen zwischen Netzwerken

Verwenden Sie gemeinsam genutzte Dateisysteme oder Objektspeicher zum Zugriff auf häufig genutzte Daten und minimieren Sie die zur Unterstützung von Datenverschiebungen für Ihre Workload benötigten Netzwerkressourcen.

Typische Anti-Muster:

- Sie speichern alle Daten in derselben AWS-Region, unabhängig davon, wo sich deren Benutzer befinden.
- Sie optimieren Datenumfang und -format nicht vor der Verschiebung über das Netzwerk.

Vorteile der Nutzung dieser bewährten Methode: Die Optimierung der Datenverschiebung über das Netzwerk reduziert den Umfang der für die Workload benötigten Netzwerkressourcen und verringert die Umweltauswirkungen.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Mittel

Implementierungsleitfaden

Das Verschieben von Daten in der gesamten Organisation erfordert Datenverarbeitungs-, Netzwerk- und Speicherressourcen. Verwenden Sie Techniken zur Minimierung von Datenverschiebungen und verbessern Sie die Gesamteffizienz Ihrer Workloads.

Implementierungsschritte

- Berücksichtigen Sie die Nähe zu den Daten oder Benutzer für die Entscheidung bei der Auswahl einer Region für Ihre Workload.
- Partitionieren Sie regional genutzte Services so, dass regionsspezifische Daten in der Region gespeichert werden, in der sie genutzt werden.
- Verwenden Sie effiziente Dateiformate (wie etwa Parquet oder ORC) und komprimieren Sie die Daten, bevor Sie sie über das Netzwerk verschieben.

- Verschieben Sie keine nicht genutzten Daten. Einige Beispiele, die Ihnen helfen können, das Verschieben ungenutzter Daten zu vermeiden:
  - Beschränken Sie API-Antworten nur auf relevante Daten.
  - Aggregieren Sie Daten, wenn keine detaillierten Informationen auf Datensatzebene benötigt werden.
  - Siehe Well-Architected Lab - Optimize Data Pattern Using Amazon Redshift Data Sharing.
  - Erwägen Sie die kontoübergreifende Datenfreigabe in AWS Lake Formation.
- Nutzen Sie Services, die Ihnen dabei helfen können, Code näher an den Benutzern Ihrer Workload auszuführen:

Service	Wann sollte dies verwendet werden?
Lambda@Edge	Verwenden Sie dies für rechenintensive Anwendungen, die ausgeführt werden, wenn sich Objekte nicht im Zwischenspeicher befinden.
CloudFront-Funktionen	Verwenden Sie diese für einfache Anwendungsfälle wie HTTP(s)-Anfragen oder Antwortmanipulationen, die von kurzlebigen Funktionen initiiert werden können.
AWS IoT Greengrass	Führen Sie lokale Datenverarbeitungsvorgänge, Messaging sowie die Datenzwischenspeicherung für verbundene Geräte aus.

Ressourcen

Zugehörige Dokumente:

- Optimizing your AWS Infrastructure for Sustainability, Part III: Networking
- Globale AWS-Infrastruktur
- Hauptfeatures von Amazon CloudFront einschließlich CloudFront Global Edge Network
- Compressing HTTP requests in Amazon OpenSearch Service
- Intermediäre Datenkompression mit Amazon EMR

- Loading compressed data files from Amazon S3 into Amazon Redshift
- Bereitstellen komprimierter Dateien mit Amazon CloudFront

Zugehörige Videos:

- Demystifying data transfer on AWS

Zugehörige Beispiele:

- Architecting for sustainability - Minimize data movement across networks

SUS04-BP08 Sichern von Daten nur in dem Fall, wenn ihre erneute Erstellung schwierig ist

Vermeiden Sie das Sichern von Daten ohne geschäftlichen Wert, um die Anforderungen an Speicherressourcen für Ihre Workload zu minimieren.

Typische Anti-Muster:

- Sie haben keine Sicherungsstrategie für Ihre Daten.
- Sie sichern Daten, die problemlos erneut erstellt werden können.

Vorteile der Nutzung dieser bewährten Methode: Das Vermeiden der Sicherung nichtkritischer Daten reduziert den Umfang der benötigten Speicherressourcen für die Workload und verringert die Umweltauswirkungen.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Mittel

Implementierungsleitfaden

Die Vermeidung der Sicherung nicht benötigter Daten kann Kosten senken und die von der Workload verwendeten Speicherressourcen verringern. Sichern Sie nur Daten, die einen geschäftlichen Wert haben oder zur Erfüllung von Compliance-Anforderungen benötigt werden. Prüfen Sie Sicherungsrichtlinien und vermeiden Sie einen flüchtigen Speicher, der in einem Wiederherstellungsszenario keinen Wert bietet.

Implementierungsschritte

- Implementieren Sie eine Richtlinie für die Klassifizierung von Daten wie in SUS04-BP01 Implementieren einer Richtlinie für die Klassifizierung von Daten erläutert.
- Nutzen Sie die Wichtigkeit Ihrer Datenklassifizierung und entwerfen Sie eine Sicherungsstrategie auf der Grundlage Ihres Recovery Time Objective (RTO) und Recovery Point Objective (RPO). Vermeiden Sie die Sicherung nichtkritischer Daten.
  - Schließen Sie Daten aus, die problemlos erneut erstellt werden können.
  - Schließen Sie flüchtige Daten von Sicherungen aus.
  - Schließen Sie lokale Kopien von Daten aus, es sei denn, die für die Wiederherstellung dieser Daten von einem gemeinsamen Standort benötigte Zeit überschreitet Ihre Service Level Agreements (SLAs).
- Verwenden Sie eine automatisierte Lösung oder einen verwalteten Service zur Sicherung geschäftskritischer Daten.
  - AWS Backup ist ein vollständig verwalteter Service, der das Zentralisieren und Automatisieren des Datenschutzes in AWS-Services, in der Cloud sowie On-Premises erleichtert. Praktische Anleitungen zur Erstellung automatisierter Sicherungen mit AWS Backup finden Sie unter Well-Architected Labs - Testing Backup and Restore of Data.
  - Automate backups and optimize backup costs for Amazon EFS using AWS Backup.

Ressourcen

Zugehörige bewährte Methoden:

- REL09-BP01 Ermitteln und Sichern aller zu sichernden Daten oder Reproduzieren der Daten aus Quellen
- REL09-BP03 Automatische Daten-Backups
- REL13-BP02: Verwenden von definierten Wiederherstellungsstrategien, um die Wiederherstellungsziele zu erreichen

Zugehörige Dokumente:

- Sichern und Wiederherstellen von Amazon-EFS-Dateisystemen mit AWS Backup
- Amazon EBS snapshots
- Working with backups on Amazon Relational Database Service
- APN-Partner: Partner, die Sie bei der Sicherung unterstützen können

- AWS Marketplace: Für die Sicherung geeignete Produkte
- Backing Up Amazon EFS
- Backing Up Amazon FSx for Windows File Server
- Backup and Restore for Amazon ElastiCache (Redis OSS)

Zugehörige Videos:

- AWS re:Invent 2.023 - Backup and disaster recovery strategies for increased resilience
- AWS re:Invent 2.023 - What's new with AWS Backup
- AWS re:Invent 2.021 - Backup, disaster recovery, and ransomware protection with AWS

Zugehörige Beispiele:

- Well-Architected Lab - Backup data

Hardware und Services

Suchen Sie nach Möglichkeiten, die Auswirkungen auf die Nachhaltigkeit Ihrer Workloads durch Änderungen der Methoden für die Hardwareverwaltung zu reduzieren. Minimieren Sie den Umfang der für die Bereitstellung erforderlichen Hardware und wählen Sie die jeweils effizienteste Hardware und den effizientesten Service für die jeweilige Workload aus.

Bewährte Methoden

- SUS05-BP01 Verwenden der geringstmöglichen Menge an Hardware zur Erfüllung Ihrer Anforderungen
- SUS05-BP02 Verwenden von Instance-Typen mit den geringsten Auswirkungen
- SUS05-BP03 Verwenden verwalteter Services
- SUS05-BP04 Optimieren der Nutzung von hardwarebasierten Datenverarbeitungsbeschleunigern

SUS05-BP01 Verwenden der geringstmöglichen Menge an Hardware zur Erfüllung Ihrer Anforderungen

Verwenden Sie die geringstmögliche Menge an Hardware für Ihre Workload, um Ihre geschäftlichen Anforderungen in effizienter Weise zu erfüllen.

Typische Anti-Muster:

- Sie überwachen die Ressourcenauslastung nicht.
- Sie haben Ressourcen mit geringer Auslastung in Ihrer Architektur.
- Sie prüfen die Nutzung statischer Hardware nicht, um festzustellen, ob sie neu dimensioniert werden muss.
- Sie formulieren keine Ziele für die Hardwarenutzung in Ihrer Datenverarbeitungsinfrastruktur auf der Grundlage geschäftlicher KPIs.

Vorteile der Nutzung dieser bewährten Methode: Die korrekte Dimensionierung Ihrer Cloud-Ressourcen hilft dabei, die Umweltauswirkungen von Workloads zu reduzieren, Geld zu sparen und Leistungsbenchmarks einzuhalten.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Mittel

Implementierungsleitfaden

Wählen Sie die optimale Anzahl von Hardwaregeräten für Ihre Workload aus, um die allgemeine Effizienz zu verbessern. Die AWS Cloud bietet die Flexibilität, Ressourcen dynamisch durch verschiedene Mechanismen wie etwa AWS Auto Scaling zu erweitern oder zu reduzieren, um einem veränderten Bedarf gerecht zu werden. Dazu kommen APIs und SDKs, mit denen Ressourcen mit minimalem Aufwand angepasst werden können. Verwenden Sie diese Möglichkeiten für häufige Änderungen an Ihren Workload-Implementierungen. Verwenden Sie dazu Dimensionierungsanleitungen von AWS-Tools für den effizienten Betrieb Ihrer Cloud-Ressourcen und die Erfüllung Ihrer geschäftlichen Anforderungen.

Implementierungsschritte

- Auswahl des Instance-Typs: Wählen Sie den Instance-Typ aus, der Ihren Anforderungen am besten entspricht. Weitere Informationen zur Auswahl von Instances von Amazon Elastic Compute Cloud und zur Verwendung von Mechanismen wie der attributbasierten Auswahl des Instance-Typs finden Sie im Folgenden:
  - Wie wähle ich den geeigneten Amazon-EC2-Instance-Typ für meine Workload aus?
  - Attributbasierte Auswahl von Instance-Typen für Amazon-EC2-Flotte
  - Erstellen einer Auto-Scaling-Gruppe mit attributbasierter Auswahl des Instance-Typs
- Skalierung: Skalieren Sie variable Workloads in kleinen Schritten.

- Verwendung mehrerer Einkaufsoptionen für Datenverarbeitung: Kombinieren Sie Instance-Flexibilität, Skalierbarkeit und Kosteneinsparungen mit mehreren Einkaufsoptionen für Datenverarbeitung.
  - On-Demand-Instances von Amazon EC2 eignen sich am besten für neue, zustandsbehaftete Workloads mit Spitzen, die hinsichtlich Instance-Typ, Standort oder Zeit nicht flexibel sein können.
  - Amazon EC2 Spot Instances eignen sich hervorragend zur Ergänzung der anderen Optionen für Anwendungen, die fehlertolerant und flexibel sind.
  - Nutzen Sie Compute Savings Plans für stabile Workloads, die Flexibilität ermöglichen, wenn sich Ihre Anforderungen (wie AZ, Region, Instance-Familien oder Instance-Typen) ändern.
- Nutzung der Vielfalt von Instances und Availability Zones: Maximieren Sie die Anwendungsverfügbarkeit und nutzen Sie überschüssige Kapazitäten, indem Sie Ihre Instances und Availability Zones diversifizieren.
- Korrekte Dimensionierung von Instances: Verwenden Sie die Empfehlungen zur Dimensionierung in AWS-Tools, um Anpassungen an Ihrer Workload vorzunehmen. Weitere Informationen finden Sie unter Kostenoptimierung mit Empfehlungen zur richtigen Dimensionierung und Richtige Dimensionierung: Bereitstellen von Instances entsprechend den Workloads.
  - Verwenden Sie die Empfehlungen zur Dimensionierung in AWS Cost Explorer oder AWS Compute Optimizer zur Identifizierung von Dimensionierungsmöglichkeiten.
- Verhandlung von Service Level Agreements (SLAs): Verhandeln Sie SLAs, die eine vorübergehende Reduzierung der Kapazität ermöglichen, während die Automatisierung Ersatzressourcen bereitstellt.

Ressourcen

Zugehörige Dokumente:

- Optimizing your AWS Infrastructure for Sustainability, Part I: Compute
- Attribute based Instance Type Selection for Auto Scaling for Amazon EC2 Fleet
- AWS Compute Optimizer-Dokumentation
- Ausführen von Lambda: Leistungsoptimierung
- Dokumentation zu Auto Scaling

Zugehörige Videos:

- AWS re:Invent 2.023 - What's new with Amazon EC2
- AWS re:Invent 2.023 - Smart savings: Amazon Elastic Compute Cloud cost-optimization strategies
- AWS re:Invent 2.022 - Optimizing Amazon Elastic Kubernetes Service for performance and cost on AWS
- AWS re:Invent 2.023 - Sustainable compute: reducing costs and carbon emissions with AWS

SUS05-BP02 Verwenden von Instance-Typen mit den geringsten Auswirkungen

Überwachen und nutzen Sie kontinuierlich neue Instance-Typen, um Verbesserungen bei der Energieeffizienz zu nutzen.

Typische Anti-Muster:

- Sie verwenden lediglich eine Familie von Instances.
- Sie verwenden nur x86-Instances.
- Sie geben einen Instance-Typ in Ihrer Konfiguration von Amazon EC2 Auto Scaling an.
- Sie verwenden AWS-Instances in einer Weise, für die sie nicht gedacht sind (beispielsweise für Datenverarbeitung optimierte Instances für speicherintensive Workloads).
- Sie evaluieren nicht regelmäßig neue Instance-Typen.
- Sie prüfen nicht die Empfehlungen von AWS-Dimensionierungstools wie etwa AWS Compute Optimizer.

Vorteile der Nutzung dieser bewährten Methode: Durch die Verwendung energieeffizienter und korrekt dimensionierter Instances können Sie die Umweltauswirkungen und die Kosten Ihrer Workloads deutlich reduzieren.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Mittel

Implementierungsleitfaden

Die Verwendung effizienter Instances für Cloud-Workloads ist von entscheidender Bedeutung für eine geringere Ressourcennutzung und die Kosteneffizienz. Überwachen Sie kontinuierlich die Einführung neuer Instance-Typen und nutzen Sie Verbesserungen bei der Energieeffizienz, einschließlich Instance-Typen, die zur Unterstützung spezifischer Workloads bestimmt sind, wie z. B. Machine-Learning-Trainings und -Inferenzen und Videotranskodierung.

Implementierungsschritte

- Kennenlernen der Instance-Typen: Finden Sie Instance-Typen, mit denen Sie die Umweltbelastung Ihrer Workloads verringern können.
  - Abonnieren Sie Neuerungen bei AWS, um sich über die aktuellen AWS-Technologien und -Instances auf dem Laufenden zu halten.
  - Informieren Sie sich über die verschiedenen AWS-Instance-Typen.
  - Informieren Sie sich über AWS-Graviton-basierte Instances, die die beste Leistung pro Watt in Amazon EC2 bieten. Sehen Sie sich re:Invent 2.020 - Deep dive on AWS Graviton2 processor-powered Amazon EC2 instances und Deep dive into AWS Graviton3 and Amazon EC2 C7g instances an.
- Verwendung von Instance-Typen mit den geringsten Auswirkungen: Planen Sie Ihre Workload und stellen Sie sie auf Instance-Typen mit den geringsten Auswirkungen um.
  - Definieren Sie einen Prozess zur Evaluierung neuer Features oder Instances für Ihre Workloads. Nutzen Sie die Agilität in der Cloud, um schnell zu testen, wie neue Instance-Typen die ökologische Nachhaltigkeit Ihrer Workloads verbessern können. Nutzen Sie Proxy-Metriken, um zu messen, wie viele Ressourcen Sie für eine Arbeitseinheit benötigen.
  - Modifizieren Sie Ihre Workload nach Möglichkeit so, dass sie mit unterschiedlichen Zahlen von vCPUs und Arbeitsspeichergrößen kompatibel ist, um die größtmögliche Auswahl an Instance-Typen zu erhalten.
  - Erwägen Sie die Übertragung Ihrer Workload zu auf Graviton basierenden Instances, um die Leistungseffizienz Ihrer Workload zu verbessern. Weitere Informationen zum Verschieben von Workloads zu AWS Graviton finden Sie unter AWS Graviton Fast Start und Considerations when transitioning workloads to AWS Graviton-based Amazon Elastic Compute Cloud instances.
  - Erwägen Sie die Auswahl der Option AWS Graviton, wenn Sie verwaltete AWS-Services verwenden.
  - Migrieren Sie Ihre Workload zu Regionen mit Instances, die die geringsten nachhaltigkeitsbezogenen Auswirkungen bieten und dennoch Ihre geschäftlichen Anforderungen erfüllen.
  - Nutzen Sie für Machine-Learning-Workloads spezielle Hardware, die auf Ihre Workload abgestimmt ist, z. B. AWS Trainium, AWS Inferentia und Amazon EC2 DL1. AWS Inferentia-Instances wie Inf2-Instances bieten eine um bis zu 50 % bessere Leistung pro Watt als vergleichbare Amazon-EC2-Instances.

  - Verwenden Sie Amazon SageMaker Inference Recommender für die korrekte Dimensionierung des ML-Inferenz-Endpunkts.
  - Verwenden Sie für Workloads, bei denen es gelegentlich zu zusätzlichen Kapazitätsanforderungen kommt, Instances mit Spitzenlastleistung.
  - Verwenden Sie für zustandslose und fehlertolerante Workloads Amazon EC2 Spot Instances, um die allgemeine Auslastung der Cloud zu verbessern und die Nachhaltigkeitsauswirkungen ungenutzter Ressourcen zu reduzieren.
- Betrieb und Optimierung: Betreiben und optimieren Sie Ihre Workload-Instance.
  - Prüfen Sie für kurzzeitige Workloads die Amazon-CloudWatch-Metriken der Instance wie `CPUUtilization`, um festzustellen, ob die Instance gar nicht oder zu wenig genutzt wird.
  - Prüfen Sie für stabile Workloads in regelmäßigen Intervallen AWS-Dimensionierungstools wie etwa AWS Compute Optimizer, um Möglichkeiten zur Optimierung und zur korrekten Dimensionierung der Instances zu erkennen. Weitere Beispiele und Empfehlungen finden Sie in den folgenden Labs:
    - Well-Architected Lab - Rightsizing Recommendations
    - Well-Architected Lab - Rightsizing with Compute Optimizer
    - Well-Architected Lab - Optimize Hardware Patterns and Observice Sustainability KPIs

Ressourcen

Zugehörige Dokumente:

- Optimizing your AWS Infrastructure for Sustainability, Part I: Compute
- AWS Graviton
- Amazon EC2 DL1
- Amazon-EC2-Flotten zur Kapazitätsreservierung
- Amazon-EC2-Spot-Flotte
- Funktionen: Lambda-Funktionskonfiguration
- Attributbasierte Auswahl von Instance-Typen für Amazon-EC2-Flotte
- Building Sustainable, Efficient, and Cost-Optimized Applications on AWS
- How the Contino Sustainability Dashboard Helps Customers Optimize Their Carbon Footprint

Zugehörige Videos:

- AWS re:Invent 2.023 - AWS Graviton: The best price performance for your AWS workloads
- AWS re:Invent 2.023 - New Amazon Elastic Compute Cloud generative AI capabilities in AWS Management Console
- AWS re:Invent 2.023 = What's new with Amazon Elastic Compute Cloud
- AWS re:Invent 2.023 - Smart savings: Amazon Elastic Compute Cloud cost-optimization strategies
- AWS re:Invent 2.021 - Deep dive into AWS Graviton3 and Amazon EC2 C7g instances
- AWS re:Invent 2.022 - Build a cost-, energy-, and resource-efficient compute environment

Zugehörige Beispiele:

- Lösung: Guidance for Optimizing Deep Learning Workloads for Sustainability on AWS
- Datenbankmigration von Amazon Relational Database Service zu Graviton

SUS05-BP03 Verwenden verwalteter Services

Verwenden Sie verwaltete Services für effizientere Betriebsabläufe in der Cloud.

Typische Anti-Muster:

- Sie verwenden Amazon-EC2-Instances mit geringer Ausnutzung für die Ausführung Ihrer Anwendungen.
- Ihr internes Team verwaltet nur die Workload, ohne Zeit zu haben, sich auf Innovation oder Vereinfachungen zu konzentrieren.
- Sie nutzen und verwalten Technologien für Aufgaben, die effizienter auf verwalteten Services ausgeführt werden können.

Vorteile der Nutzung dieser bewährten Methode:

- Durch die Verwendung verwalteter Services geht die Verantwortung auf AWS über, mit Erkenntnissen zu Millionen von Kunden, was Innovationen und neue Effizienzen ermöglicht.
- Ein verwalteter Service verteilt die Umweltauswirkungen des Services durch Multi-Tenet-Steuerebenen auf mehrere Benutzer.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Mittel

Implementierungsleitfaden

Verwaltete Services übertragen die Verantwortung für die Wahrung einer hohen durchschnittlichen Nutzung und die Optimierung der Nachhaltigkeit der bereitgestellten Hardware auf AWS. Verwaltete Services eliminieren dazu den betrieblichen und administrativen Aufwand für die Wartung eines Service, sodass Ihr Team mehr Zeit hat und sich auf Innovationen konzentrieren kann.

Prüfen Sie Ihre Workload, um die Komponenten zu identifizieren, die von verwalteten AWS-Services ersetzt werden können. Amazon RDS, Amazon Redshift und Amazon ElastiCache bieten beispielsweise einen verwalteten Datenbankservice. Amazon Athena, Amazon EMR und Amazon OpenSearch Service bieten einen verwalteten Analytikservice.

Implementierungsschritte

1. Inventarisieren Ihrer Workload: Inventarisieren Sie Ihre Workload für Services und Komponenten.
2. Identifizieren von Kandidaten: Bewerten und identifizieren Sie Komponenten, die durch verwaltete Services ersetzt werden können. Hier finden Sie einige Beispiele für Situationen, in denen Sie einen verwalteten Service in Erwägung ziehen sollten:

Aufgabe	AWS-Service
Hosten einer Datenbank	Verwenden Sie verwaltete Instances von Amazon Relational Database Service (Amazon RDS), anstatt Ihre eigenen Amazon-RDS-Instances in Amazon Elastic Compute Cloud (Amazon EC2) zu verwalten.
Hosten einer Container-Workload	Verwenden Sie AWS Fargate, anstatt Ihre eigene Container-Infrastruktur zu implement ieren.
Hosten von Web-Apps	Verwenden Sie AWS Amplify Hosting als vollständig verwalteten CI/CD- und Hosting-S ervice für statische Websites und serverseitig gerenderte Web-Apps.

3. Erstellen eines Migrationsplans: Identifizieren Sie Abhängigkeiten und erstellen Sie einen Migrationsplan. Aktualisieren Sie Runbooks und Playbooks entsprechend.

- Der AWS Application Discovery Service erfasst und präsentiert automatisch detaillierte Informationen über Anwendungsabhängigkeiten und die Nutzung von Anwendungen. Damit treffen Sie bei der Planung Ihrer Migration fundiertere Entscheidungen.

4. Tests: Testen Sie den Service vor der Migration zum verwalteten Service.
5. Ersetzen selbst gehosteter Services: Verwenden Sie Ihren Migrationsplan, um selbst gehostete Services durch verwaltete Services zu ersetzen.
6. Überwachen und anpassen: Überwachen Sie den Service nach der Migration kontinuierlich, um erforderliche Anpassungen vorzunehmen und den Service zu optimieren.

Ressourcen

Zugehörige Dokumente:

- AWS Cloud-Produkte
- AWS-Rechner für Gesamtbetriebskosten (TCO)
- Amazon DocumentDB
- Amazon Elastic Kubernetes Service (EKS)
- Amazon Managed Streaming for Apache Kafka (Amazon MSK)

Zugehörige Videos:

- AWS re:Invent 2.021 - Cloud operations at scale with AWS Managed Services
- AWS re:Invent 2.023 - Best practices for operating on AWS

SUS05-BP04 Optimieren der Nutzung von hardwarebasierten Datenverarbeitungsbeschleunigern

Sie können die Nutzung von beschleunigten Computing-Instances optimieren, um die Anforderungen Ihrer Workload an die physische Infrastruktur zu reduzieren.

Typische Anti-Muster:

- Sie überwachen die GPU-Nutzung nicht.
- Sie verwenden eine Allzweck-Instance für die Workload, während eine speziell erstellte Instance eine höhere Leistung, geringere Kosten und eine bessere Leistung pro Watt bieten kann.

- Sie verwenden hardwarebasierte Datenverarbeitungsbeschleuniger für Aufgaben, bei denen CPU-basierte Alternativen effizienter sind.

Vorteile der Nutzung dieser bewährten Methode: Durch den optimalen Einsatz hardwarebasierter Beschleuniger können Sie die Anforderungen an die physische Infrastruktur Ihrer Workload reduzieren.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Mittel

Implementierungsleitfaden

Wenn Sie eine hohe Verarbeitungsleistung benötigen, können Sie beschleunigte Computing-Instances verwenden. Diese bieten Zugriff auf hardwarebasierte Datenverarbeitungsbeschleuniger wie Grafikprozessoren (Graphics Processing Units, GPUs) und Field-Programmable Gate Arrays (FPGAs). Diese Hardwarebeschleuniger führen bestimmte Funktionen wie die Grafikverarbeitung oder Datenmusterzuordnung effizienter aus als CPU-basierte Alternativen. Viele beschleunigte Workloads, wie Rendering, Transkodierung und Machine Learning, sind sehr variabel im Bezug auf die Ressourcennutzung. Betreiben Sie diese Hardware nur so lange wie nötig und nehmen Sie sie automatisch außer Betrieb, wenn sie nicht mehr benötigt wird, um den Ressourcenverbrauch zu minimieren.

Implementierungsschritte

- Ermitteln Sie, welche beschleunigten Computing-Instances für Ihre Anforderungen geeignet sind.
- Nutzen Sie für Machine-Learning-Workloads spezielle Hardware, die auf Ihre Workload abgestimmt ist, z. B. AWS Trainium, AWS Inferentia und Amazon EC2 DL1. AWS Inferentia-Instances wie Inf2-Instances bieten eine um bis zu 50 % bessere Leistung pro Watt als vergleichbare Amazon-EC2-Instances.
- Erfassen Sie Nutzungsmetriken für Ihre beschleunigten Computing-Instances. Sie können den CloudWatch-Agent beispielsweise verwenden, um Metriken wie `utilization_gpu` und `utilization_memory` für Ihre GPUs zu sammeln, wie im Artikel zum Erfassen von NVIDIA-GPU-Metriken mit Amazon CloudWatch beschrieben.
- Optimieren Sie Code, Netzwerkbetrieb und die Einstellungen von Hardwarebeschleunigern, um sicherzustellen, dass die zugrunde liegende Hardware optimal genutzt wird.
  - Optimieren der GPU-Einstellungen
  - GPU-Überwachung und -Optimierung im Deep-Learning-AMI

- Optimieren von E/A für die GPU-Leistungsoptimierung von Deep Learning-Training in Amazon SageMaker)
- Verwenden Sie die aktuellen leistungsstarken Bibliotheken und GPU-Treiber.
- Automatisieren Sie die Freigabe nicht genutzter GPU-Instances.

Ressourcen

Zugehörige Dokumente:

- Beschleunigte Datenverarbeitung
- Let’s Architect! Architecting with custom chips and accelerators
- Wie wähle ich den geeigneten EC2-Instance-Typ für meine Workload aus?
- VT1-Instances von Amazon EC2
- Choose the best AI accelerator and model compilation for computer vision inference with Amazon SageMaker

Zugehörige Videos:

- AWS re:Invent 2.021 - How to select Amazon EC2 GPU instances for deep learning
- AWS Online Tech Talks - Deploying Cost-Effective Deep Learning Inference
- AWS re:Invent 2.023 - Cutting-edge AI with AWS and NVIDIA
- AWS re:Invent 2.022 - [NEW LAUNCH!] Introducing AWS Inferentia2-based Amazon EC2 Inf2 instances
- AWS re:Invent 2.022 - Accelerate deep learning and innovate faster with AWS Trainium
- AWS re:Invent 2.022 - Deep learning on AWS with NVIDIA: From training to deployment

Prozess und Kultur

Reduzieren Sie nachhaltigkeitsbezogene Auswirkungen, indem Sie Ihre Entwicklungs-, Test- und Bereitstellungsmethoden ändern.

Bewährte Methoden

- SUS06-BP01 Einführen von Methoden, die schnelle Verbesserungen für die Nachhaltigkeit ermöglichen

- SUS06-BP02 Konstantes Aktualisieren Ihrer Workload
- SUS06-BP03 Höhere Auslastung von Entwicklungsumgebungen
- SUS06-BP04 Verwenden verwalteter Gerätefarmen für Tests

SUS06-BP01 Einführen von Methoden, die schnelle Verbesserungen für die Nachhaltigkeit ermöglichen

Nutzen Sie Methoden und Prozesse zur Validierung potenzieller Verbesserung, zur Minimierung von Testkosten und zur Bereitstellung kleinerer Verbesserungen.

Typische Anti-Muster:

- Die Prüfung Ihrer Anwendung auf Nachhaltigkeitsaspekte erfolgt nur einmal zu Beginn des Projekts.
- Ihre Workload stagniert, da der Freigabeprozess zu komplex ist, um kleinere Verbesserungen für die Ressourceneffizienz umzusetzen.
- Sie verfügen über keine Mechanismen zur Verbesserung Ihrer Workload unter Nachhaltigkeitsaspekten.

Vorteile der Nutzung dieser bewährten Methode: Durch die Einrichtung eines Prozesses für die Einführung und Nachverfolgung von Nachhaltigkeitsverbesserungen können Sie kontinuierlich neue Funktionen einführen, Probleme beseitigen und die Workload-Effizienz verbessern.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Mittel

Implementierungsleitfaden

Testen und validieren Sie potenzielle Verbesserungen in Bezug auf die Nachhaltigkeit, bevor Sie sie in der Produktion bereitstellen. Berücksichtigen Sie die Testkosten bei der Berechnung des potenziellen zukünftigen Nutzens einer Verbesserung. Entwickeln Sie kostengünstige Testmethoden, um kleinere Verbesserungen einzuführen.

Implementierungsschritte

- Kenntnis und Kommunikation der Nachhaltigkeitsziele Ihrer Organisation: Machen Sie sich mit den Nachhaltigkeitszielen Ihrer Organisation vertraut, z. B. zur Reduzierung der CO_2-Emissionen oder zum verantwortungsvollen Umgang mit Wasser. Übersetzen Sie diese Ziele in

Nachhaltigkeitsanforderungen für Ihre Cloud-Workloads. Kommunizieren Sie diese Anforderungen an wichtige Stakeholder.

- Ergänzung des Backlogs mit Nachhaltigkeitsanforderungen: Fügen Sie Ihrem Entwicklungs-Backlog Anforderungen zur Verbesserung der Nachhaltigkeit hinzu.
- Iterieren und verbessern: Verwenden Sie einen iterativen Verbesserungsprozess, um diese Verbesserungen zu identifizieren, zu bewerten, zu priorisieren, zu testen und bereitzustellen.
- Tests unter Verwendung des Minimum Viable Product (MVP): Entwickeln und testen Sie potenzielle Verbesserungen unter Verwendung der Minimum-Viable-Komponenten, um die Kosten und die Umweltauswirkungen der Tests zu reduzieren.
- Verbessern und optimieren Sie Ihre Entwicklungsprozesse kontinuierlich. Sie können beispielsweise Ihren Softwarebereitstellungsprozess mit Pipelines für die Continuous Integration und Continuous Delivery (CI/CD) automatisieren, um potenzielle Verbesserungen zu testen und bereitzustellen und so den Aufwand zu reduzieren und Fehler durch manuelle Prozesse zu minimieren.
- Schulung und Sensibilisierung: Führen Sie Schulungsprogramme für Ihre Teammitglieder durch, um sie über Nachhaltigkeit und die Auswirkungen ihrer Aktivitäten auf die Nachhaltigkeitsziele Ihrer Organisation aufzuklären.
- Bewerten und anpassen: Bewerten Sie kontinuierlich die Auswirkungen von Verbesserungen und nehmen Sie bei Bedarf Anpassungen vor.

Ressourcen

Zugehörige Dokumente:

- AWS ermöglicht nachhaltige Lösungen
- Scalable agile development practices based on AWS CodeCommit

Zugehörige Videos:

- AWS re:Invent 2.023 - Sustainable architecture: Past, present, and future
- AWS re:Invent 2.022 - Delivering sustainable, high-performing architectures
- AWS re:Invent 2.022 - Architecting sustainably and reducing your AWS carbon footprint
- AWS re:Invent 2.022 - Sustainability in AWS global infrastructure
- AWS re:Invent 2.023 - What’s new with AWS observability and operations

Zugehörige Beispiele:

- Well-Architected Lab – Kosten- und Nutzungsberichte in Effizienzberichte umwandeln

SUS06-BP02 Konstantes Aktualisieren Ihrer Workload

Halten Sie Ihre Workload auf neustem Stand, um effiziente Features zu übernehmen, Probleme zu beseitigen und die allgemeine Effizienz der Workload zu wahren.

Typische Anti-Muster:

- Sie gehen davon aus, dass Ihre aktuelle Architektur statisch ist und im Laufe der Zeit nicht aktualisiert wird.
- Sie haben keine Systeme oder regelmäßigen Besprechungen zur Prüfung, ob aktualisierte Software und Pakete mit Ihrer Workload kompatibel sind.

Vorteile der Nutzung dieser bewährten Methode: Wenn Sie einen Prozess einrichten, um Ihre Workload auf neustem Stand zu halten, können Sie neue Features und Kapazitäten nutzen, Probleme lösen und die Workload-Effizienz verbessern.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Niedrig

Implementierungsleitfaden

Aktuelle Betriebssysteme,Laufzeiten, Middleware, Bibliotheken und Anwendungen können die Workload-Effizienz verbessern und die Nutzung effizienterer Technologien unterstützen. Aktuelle Software kann darüber hinaus Features für eine genauere Messung der Auswirkungen Ihrer Workload bereitstellen, da die Anbieter mit ihren Features ebenfalls Nachhaltigkeitsziele erfüllen müssen. Sorgen Sie für Regelmäßigkeit bei der Aktualisierung Ihrer Workload mit den neuesten Features und Versionen.

Implementierungsschritte

- Definieren eines Prozesses: Definieren Sie einen Prozess und einen Zeitplan, um neue Features oder Instances für Ihre Workloads zu evaluieren. Nutzen Sie die Agilität in der Cloud, um schnell zu testen, wie neue Features Ihre Workloads verbessern können:
  - Reduzierung von Auswirkungen auf die Nachhaltigkeit
  - Erzielen von Leistungseffizienzen

  - Beseitigen von Hindernissen für geplante Verbesserungen
  - Verbesserung Ihrer Fähigkeit für die Messung von und den Umgang mit Nachhaltigkeitsauswirkungen
- Inventarisierung: Inventarisieren Sie Ihre Workload-Software und -Architektur und identifizieren Sie Komponenten, die aktualisiert werden müssen.
  - Sie können AWS Systems Manager Inventory verwenden, um Betriebssystem (BS)-, Anwendungs- und Instance-Metadaten von Ihren Amazon-EC2-Instances zu erfassen und so schnell zu verstehen, welche Instances die Software und die Konfigurationen ausführen, die Ihre Softwarerichtlinie erfordert, und welche Instances aktualisiert werden müssen.
- Kennenlernen des Aktualisierungsverfahrens: Erfahren Sie, wie die Komponenten Ihrer Workload aktualisiert werden.

Workload-Komponente	Aktualisierung
Machine Images	Verwenden Sie EC2 Image Builder zur Verwaltung von Updates für Amazon Machine Images (AMIs) für Linux- oder Windows Server-Images.
Container-Images	Verwenden Sie Amazon Elastic Container Registry (Amazon ECR) mit Ihrer vorhanden en Pipeline zur Verwaltung von Amazon Elastic Container Service (Amazon ECS)-Imag es.
AWS Lambda	AWS Lambda enthält Features zur Versionsv erwaltung.

- Verwendung von Automatisierung: Verwenden Sie Automatisierung für den Aktualisierungsvorgang, um den Aufwand für die Bereitstellung neuer Features zu reduzieren und Fehler zu begrenzen, die durch manuelle Prozesse verursacht werden.
  - Sie können CI/CD verwenden, um AMIs, Container-Images und andere Artefakte im Zusammenhang mit Ihrer Cloud-Anwendung automatisch zu aktualisieren.
  - Sie können Tools wie den AWS Systems Manager Patch Manager verwenden, um den Systemaktualisierungsprozess zu automatisieren und die Aktivitäten mit AWS Systems Manager Maintenance Windows zu planen.

Ressourcen

Zugehörige Dokumente:

- AWS-Architekturzentrum
- Neuerungen bei AWS
- AWS-Entwicklungstools

Zugehörige Videos:

- AWS re:Invent 2.022 - Optimize your AWS workloads with best-practice guidance
- All Things Patch: AWS Systems Manager

Zugehörige Beispiele:

- Well-Architected Labs: Bestands- und Patch-Verwaltung
- Lab: AWS Systems Manager

SUS06-BP03 Höhere Auslastung von Entwicklungsumgebungen

Erhöhen Sie die Ausnutzung von Ressourcen zum Entwickeln, Testen und Erstellen Ihrer Workloads.

Typische Anti-Muster:

- Sie stellen Ihre Build-Umgebungen manuell bereit oder beenden sie manuell.
- Sie lassen Ihre Build-Umgebungen unabhängig von Test-, Build- oder Freigabeaktivitäten laufen (dazu gehört etwa der Betrieb einer Umgebung außerhalb der Arbeitszeit der Mitglieder Ihres Entwicklungsteams).
- Sie stellen übermäßig viele Ressourcen für Ihre Build-Umgebung bereit.

Vorteile der Nutzung dieser bewährten Methode: Durch die Steigerung der Ausnutzung von Build-Umgebungen können Sie die allgemeine Effizienz Ihrer Cloud-Workload verbessern, da die Ressourcen in effizienter Weise Entwicklungs-, Test- und Build-Aktivitäten zugewiesen werden.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Niedrig

Implementierungsleitfaden

Verwenden Sie Automatisierung und Infrastructure as Code, um Build-Umgebungen in Betrieb zu nehmen, wenn sie gebraucht werden, und sie andernfalls zu deaktivieren. Eine typische Vorgehensweise besteht in der Planung von Verfügbarkeitszeiten, die mit den Arbeitszeiten der Entwicklungsteams übereinstimmen. Ihre Testumgebungen sollten der Produktionskonfiguration sehr stark ähneln. Suchen Sie aber nach Möglichkeiten, Instance-Typen mit Burst-Kapazität, Amazon EC2 Spot Instances, automatisch skalierenden Datenbankservices, Containern und Serverless-Technologien zu verwenden, um die Entwicklungs- und Testkapazität an der Nutzung auszurichten. Begrenzen Sie das Datenvolumen auf die Testanforderungen. Wenn Sie Produktionsdaten für einen Test verwenden, sollten Sie nach Möglichkeiten suchen, Daten aus der Produktion gemeinsam zu nutzen, anstatt Daten hin- und herzuschieben.

Implementierungsschritte

- Infrastructure as Code verwenden: Verwenden Sie Infrastructure as Code, um Ihre Entwicklungsumgebungen bereitzustellen.
- Automatisierung verwenden: Nutzen Sie Automatisierungen, um den Lebenszyklus Ihrer Entwicklungs- und Testumgebungen zu verwalten und die Effizienz Ihrer Entwicklungsressourcen zu maximieren.
- Nutzung maximieren: Verwenden Sie Strategien zur Maximierung der Nutzung von Entwicklungs- und Testumgebungen.
  - Verwenden Sie die geringstmögliche Zahl repräsentativer Umgebungen, um mögliche Verbesserungen zu entwickeln und zu testen.
  - Nutzen Sie nach Möglichkeit Serverless-Technologien.
  - Verwenden Sie On-Demand-Instances, um Entwicklergeräte zu ergänzen.
  - Verwenden Sie Instance-Typen mit Burst-Kapazität, Spot Instances und andere Technologien, um die Entwicklungskapazität an der Nutzung auszurichten.
  - Nutzen Sie native Cloud-Services für den sicheren Instance-Shell-Zugriff, statt Bastion-Host-Flotten bereitzustellen.
  - Skalieren Sie Ihre Build-Ressourcen automatisch je nach Build-Aktivität.

Ressourcen

Zugehörige Dokumente:

- AWS Systems Manager Session Manager
- Amazon EC2 Burstable performance instances
- Was ist AWS CloudFormation?
- Was ist AWS CodeBuild?
- Instance Scheduler in AWS

Zugehörige Videos:

- AWS re:Invent 2.023 - Continuous integration and delivery for AWS

SUS06-BP04 Verwenden verwalteter Gerätefarmen für Tests

Verwenden Sie verwaltete Gerätefarmen zum effektiven Testen neuer Features auf einer repräsentativen Auswahl von Hardwaregeräten.

Typische Anti-Muster:

- Sie testen Ihre Anwendung manuell und stellen sie auf einzelnen physischen Geräten bereit.
- Sie verwenden keinen App-Testservice zum Testen und zum Interagieren mit Ihren Apps (beispielsweise Android, iOS und Web-Apps) auf realen physischen Geräten.

Vorteile der Nutzung dieser bewährten Methode: Die Verwendung verwalteter Gerätefarmen zum Testen cloud-fähiger Anwendungen bringt eine Reihe von Vorteilen mit sich:

- Dazu gehören effizientere Features zum Testen von Anwendungen auf einer breiten Palette von Geräten.
- Sie machen hausinterne Infrastruktur zum Testen überflüssig.
- Sie bieten unterschiedliche Gerätetypen, darunter ältere und weniger verbreitete Hardware, was unnötige Geräte-Upgrades eliminiert.

Risikostufe, wenn diese bewährte Methode nicht eingeführt wird: Niedrig

Implementierungsleitfaden

Die Verwendung verwalteter Gerätefarmen kann Ihnen dabei helfen, Ihre Testprozesse für neue Features auf einer repräsentativen Auswahl von Hardwaregeräten zu optimieren. Verwaltete

Gerätefarmen stellen verschiedene Gerätetypen bereit, unterstützen auch ältere und weniger verbreitete Hardware und vermeiden nachhaltigkeitsbezogene Auswirkungen auf Kunden durch unnötige Geräte-Upgrades.

Implementierungsschritte

- Testanforderungen definieren: Definieren Sie Ihre Testanforderungen und Ihren Testplan (z. B. Testtyp, Betriebssysteme und Testzeitplan).
  - Sie können Amazon CloudWatch RUM verwenden, um clientseitige Daten zu erfassen und zu analysieren und Ihren Testplan zu entwerfen.
- Verwaltete Gerätefarm auswählen: Wählen Sie eine verwaltete Gerätefarm, die Ihre Testanforderungen unterstützen kann. Sie können beispielsweise die AWS-Gerätefarm verwenden, um die Auswirkungen Ihrer Änderungen auf eine repräsentative Auswahl von Hardwaregeräten zu testen und zu verstehen.
- Automatisierung verwenden: Verwenden Sie Automatisierung und kontinuierliche Integration/Bereitstellung (CI/CD) für die Planung und Durchführung Ihrer Tests.
  - Integration der AWS-Gerätefarm mit Ihrer CI/CD-Pipeline zur Durchführung browserübergreifender Selenium-Tests
  - Erstellen und Testen von iOS- und iPadOS-Apps mit AWS DevOps und mobilen Services
- Prüfen und Anpassen: Prüfen Sie kontinuierlich Ihre Testergebnisse und nehmen Sie die erforderlichen Verbesserungen vor.

Ressourcen

Zugehörige Dokumente:

- AWS Device Farm-Geräteliste
- Viewing the CloudWatch RUM dashboard

Zugehörige Videos:

- AWS re:Invent 2.023 - Improve your mobile and web app quality using AWS Device Farm
- AWS re:Invent 2.021 - Optimize applications through end user insights with Amazon CloudWatch RUM

Zugehörige Beispiele:

- AWS Device Farm-Beispiel-App für Android
- AWS Device Farm-Beispiel-App für iOS
- Appium-Webtests für AWS Device Farm

Schlussfolgerung

Immer mehr Organisationen setzen sich Nachhaltigkeitsziele als Reaktion auf veränderte staatliche Vorschriften, mögliche Wettbewerbsvorteile und die Nachfrage von Kunden, Mitarbeitern und Investoren. CTOs, Architekten, Entwickler und Mitglieder von Betriebsteams suchen nach Möglichkeiten, wie sie direkt zu den Nachhaltigkeitszielen ihrer Organisation beitragen können. Mithilfe dieser Designprinzipien und bewährten Methoden, die von AWS-Services unterstützt werden, können Sie fundierte Entscheidungen treffen, die ein Gleichgewicht zwischen Sicherheit, Kosten, Leistung, Zuverlässigkeit und operativer Exzellenz sowie Nachhaltigkeit für Ihre AWS Cloud-Workloads herstellen. Jede Maßnahme, die Sie ergreifen, um den Ressourcenverbrauch zu verringern und die Effizienz Ihrer Workloads zu steigern, trägt zu einer Verringerung der Umweltbelastung bei und unterstützt die umfassenderen Nachhaltigkeitsziele Ihrer Organisation.

Mitwirkende

Zu den Mitwirkenden an diesem Dokument gehören:

- Sam Mokhtari, Senior Efficiency Lead Solutions Architect, Amazon Web Services
- Brendan Sisson, Principal Sustainability Solutions Architect, Amazon Web Services
- Margaret O'Toole, Sustainability Tech Leader, Amazon Web Services
- Steffen Grunwald, Principal Sustainability Solutions Architect, Amazon Web Services
- Ryan Eccles, Principal Engineer, Sustainability, Amazon
- Rodney Lester, Principal Architect, Amazon Web Services
- Adrian Cockcroft, VP Sustainability Architecture, Amazon Web Services
- Ian Meyers, Director of Technology, Solutions Architecture, Amazon Web Services

Weitere Informationen

Zusätzliche Informationen finden Sie unter:

- AWS Well-Architected
- AWS-Architekturzentrum
- Sustainability in the Cloud
- AWS ermöglicht nachhaltige Lösungen
- The Climate Pledge
- Ziele für nachhaltige Entwicklung der Vereinten Nationen
- Greenhouse Gas Protocol

Dokumentversionen

Abonnieren Sie den RSS-Feed, um über Aktualisierungen des Whitepapers benachrichtigt zu werden.

Änderung	Beschreibung	Datum
Leitfäden zu bewährten Methoden aktualisiert	Kleine Änderungen in der gesamten Säule.	27. Juni 2024
Risikostufen aktualisiert	Geringfügige Aktualisierungen der Risikostufen bei bewährten Methoden.	3. Oktober 2023
Leitfäden zu bewährten Methoden aktualisiert	Bewährte Methoden wurden mit neuen Leitfäden in den folgenden Bereichen aktualisiert: Ausrichtung am Bedarf, Software und Architektur, Daten sowie Hardware und Services.	13. Juli 2023
Für neues Framework aktualisiert	Bewährte Methoden mit verbindlichen Anleitungen aktualisiert und neue bewährte Methoden hinzugefügt.	10. April 2023
Whitepaper aktualisiert	Bewährte Methoden mit neuen Implementierungsanleitungen aktualisiert.	15. Dezember 2022
Whitepaper aktualisiert	Weitere bewährte Methoden und Verbesserungspläne hinzugefügt.	20. Oktober 2022
Erste Veröffentlichung	Säule „Nachhaltigkeit“ – AWS Well-Architected Framework veröffentlicht.	2. Dezember 2021

Hinweise

Kunden sind dafür verantwortlich, Ihre eigene unabhängige Bewertung der Informationen in diesem Dokument vorzunehmen. Dieses Dokument: (a) dient nur zu Informationszwecken, (b) stellt die aktuellen AWS-Produktangebote und -praktiken dar, die ohne Vorankündigung geändert werden können, und (c) begründet keine Verpflichtungen oder Zusicherungen seitens AWS und der mit ihr verbundenen Unternehmen, Lieferanten oder Lizenzgeber. AWS-Produkte oder -Services werden ohne Mängelgewähr bereitgestellt, ohne ausdrückliche oder stillschweigende Garantien, Zusicherungen oder Bedingungen jeglicher Art. Die Verantwortung und Haftung von AWS gegenüber seinen Kunden werden durch AWS-Vereinbarungen geregelt. Dieses Dokument gehört, weder ganz noch teilweise, nicht zu den Vereinbarung von AWS mit seinen Kunden und ändert diese Vereinbarungen auch nicht.

© 2023, Amazon Web Services, Inc. oder Tochterfirmen. Alle Rechte vorbehalten.

AWS-Glossar

Die neueste AWS-Terminologie finden Sie im AWS-Glossar in der AWS-Glossar-Referenz.