

Whitepaper zu AWS

Übersicht über Amazon-EC2-Spot-Instances



Übersicht über Amazon-EC2-Spot-Instances: Whitepaper zu AWS

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

Die Marken und Handelsmarken von Amazon dürfen nicht in einer Weise in Verbindung mit nicht von Amazon stammenden Produkten oder Services verwendet werden, die geeignet ist, die Kunden zu verwirren oder Amazon in einer Weise herabzusetzen oder zu diskreditieren. Alle anderen Marken, die nicht Eigentum von Amazon sind, sind Eigentum ihrer jeweiligen Inhaber, die mit Amazon verbunden oder nicht verbunden oder von Amazon gesponsert oder nicht gesponsert sein können.

Table of Contents

Überblick und Einführung	1
Überblick	1
Einführung	1
Verwendung von Spot-Instances	2
Starten von Spot-Instances	3
Funktionsweise von Spot-Instances	4
Verwalten von Spot-Instance-Unterbrechungen	5
Spot-Instance-Limits	6
Bewährte Methoden für Spot-Instances	7
Spot-Integration in andere AWS-Services	9
Amazon-EMR-Integration	9
EC2-Auto-Scaling-Integration	9
Amazon-EKS-Integration	9
Amazon-ECS-Integration	9
Amazon ECS mit AWS-Fargate-Spot-Integration	10
Amazon-Batch-Integration	10
Amazon-SageMaker-Integration	10
Amazon-Gamelift-Integration	10
AWS-Elastic-Beanstalk-Integration	11
Fazit	12
Ressourcen	13
Dokumentverlauf und Mitwirkende	14
Dokumentverlauf	14
Mitwirkende	15

Übersicht über Amazon-EC2-Spot-Instances

Veröffentlichungsdatum: 5. März 2021 ([Dokumentverlauf und Mitwirkende](#))

Überblick

Dieses Dokument soll Ihnen ermöglichen, den Wert Ihrer Investitionen zu maximieren, die Prognosegenauigkeit und die Vorhersagbarkeit der Kosten zu verbessern, eine Unternehmenskultur und Kostentransparenz zu schaffen und Ihren Optimierungsstatus kontinuierlich zu messen.

Dieses Whitepaper bietet eine Übersicht über Amazon-EC2-Spot-Instances sowie bewährte Methoden, um diese effektiv zu nutzen.

Einführung

Neben [On-Demand](#), [Reserved Instances](#) und [Savings Plans](#) ist das vierte Preismodell von [Amazon Elastic Compute Cloud](#) (Amazon EC2) [Spot-Instances](#).

Mit Spot-Instances können Sie Reserve-Rechenkapazität von Amazon EC2 zu Rabatten von bis zu 90 % im Vergleich zu On-Demand-Preisen nutzen. Das bedeutet, dass Sie die Kosten für die Ausführung Ihrer Anwendungen erheblich reduzieren oder die Rechenkapazität und den Durchsatz Ihrer Anwendung für dasselbe Budget skalieren können. Der einzige Unterschied zwischen On-Demand-Instances und Spot-Instances liegt darin, dass Spot-Instances mit zwei Minuten Benachrichtigungsvorlauf durch EC2 unterbrochen werden können, wenn EC2 die Kapazität zurück braucht.

Im Gegensatz zu Reserved Instances oder Saving Plans erfordern Spot-Instances keine Vertragsbindung, um Kosteneinsparungen gegenüber On-Demand-Preisen zu erzielen. Da Spot-Instances jedoch von EC2 beendet werden können, wenn im Kapazitätspool (eine Kombination aus einem Instance-Typ und einer Availability Zone), in dem sie ausgeführt werden, keine Kapazität verfügbar ist, eignen sie sich am besten für flexible Workloads.

Verwendung von Spot-Instances

Sie können Spot-Instances für verschiedene fehlertolerante und flexible Anwendungen verwenden. Beispiele sind zustandslose Webserver, API-Endpunkte, Big Data- und Analyseanwendungen, containerisierte Workloads, CI/CD-Computing mit hoher Leistung und hohem Durchsatz (High Performance/High Throughput Computing, HPC/HTC), Rendering-Workloads und andere flexible Workloads.

Spot-Instances sind nicht für Workloads geeignet, die unflexibel, zustandsbehaftet, fehlerintolerant oder eng zwischen Instance-Knoten verbunden sind. Sie werden auch nicht für Workloads empfohlen, die intolerant gegenüber gelegentlichen Zeiträumen sind, in denen die Zielkapazität nicht vollständig verfügbar ist. Wir warnen nachdrücklich davor, Spot-Instances für diese Workloads zu verwenden oder ein Failover auf On-Demand-Instances zu versuchen, um Unterbrechungen zu bewältigen.

Starten von Spot-Instances

Der am häufigsten empfohlene Service zum Starten von Spot-Instances ist [Amazon EC2 Auto Scaling](#), da Sie damit eine gewünschte Kapazität starten und verwalten und automatisch Ressourcen anfordern können, um beeinträchtigte oder manuell beendete Ressourcen zu ersetzen. Wenn Sie eine Auto-Scaling-Gruppe konfigurieren, müssen Sie nur die Instance-Typen und die gewünschte Kapazität basierend auf Ihren Anwendungsanforderungen angeben. Weitere Informationen finden Sie unter Auto-Scaling-Gruppen im [Amazon-EC2-Auto-Scaling-Benutzerhandbuch](#).

Wenn Sie mehr Flexibilität benötigen, Ihre eigenen Workflows für den Start von Instances entwickelt haben oder einzelne Aspekte der Instance-Starts oder der Skalierungsmechanismen steuern möchten, empfehlen wir Ihnen, die Verwendung von [EC2 Fleet](#) im Sofortmodus als Alternative zu EC2 Auto Scaling in Betracht zu ziehen. Diese synchrone API ermöglicht es Ihnen, eine Liste von Instance-Typen und Startanforderungen anzugeben, und bietet flexiblere Funktionen als der [EC2-RunInstances](#)-API-Aufruf zum Starten von Spot-Instances oder On-Demand-Instances.

Wenn Sie AWS-Services für die Ausführung Ihrer Cloud-Workloads verwenden, können Sie sie auch zum Starten von Spot-Instances nutzen. Beispiele dafür sind [Amazon EMR](#), [Amazon EKS](#), [Amazon ECS](#), [AWS Batch](#) und [AWS Elastic Beanstalk](#). Sie können Spot-Instances auch mithilfe von Tools von Drittanbietern starten, die in die AWS-Cloud integriert sind.

Sie können Spot-Instance-Starts automatisieren, indem Sie Infrastructure-as-Code-Tools ([AWS CloudFormation](#), [AWS CDK](#)) oder die AWS-API, CLI oder SDKs verwenden. [Spot-Vorlagen](#) bieten einen angeleiteten Assistenten, mit dem Sie Infrastructure-as-Code-Vorlagen für AWS CloudFormation und Hashicorp-Terraform generieren können, die den bewährten Methoden von Spot entsprechen.

Funktionsweise von Spot-Instances

Spot-Instances funktionieren genauso wie andere ausgeführte EC2-Instances. Sie können jedoch von Amazon EC2 unterbrochen werden, wenn EC2 die Kapazität wieder benötigt.

Wenn EC2 Ihre Spot-Instance unterbricht, wird sie entweder beendet, angehalten oder in den Ruhezustand versetzt, je nachdem, welches Unterbrechungsverhalten Sie wählen.

Wenn EC2 Ihre Spot-Instance in der ersten Stunde, also vor einer vollen Stunde Laufzeit, unterbricht, wird Ihnen die angefangene Stunde nicht in Rechnung gestellt. Wenn Sie Ihre Spot-Instance jedoch anhalten oder beenden, zahlen Sie für jede angefangene Stunde (wie bei On-Demand- oder Reserved-Instances). Informationen darüber, wie Ihnen unterbrochene Spot-Instances auf verschiedenen Betriebssystemen in Rechnung gestellt werden, finden Sie unter [Fakturierung für unterbrochene Spot-Instances](#) im EC2-Benutzerhandbuch.

Der Spot-Preis für die einzelnen Instance-Typen in der jeweiligen Availability Zone berechnet sich nach den langfristigen Trends bei Angebot und Nachfrage sowie nach den EC2-Reservekapazitäten. Sie zahlen den geltenden Spot-Preis, der sekundengenau abgerechnet wird.

Optional können Sie einen Höchstpreis für Ihre Spot-Instance angeben. Wenn Sie keinen Höchstpreis angeben, entspricht der Standardhöchstpreis dem On-Demand-Preis. Beachten Sie, dass Sie niemals mehr als den Spot-Preis zahlen, der zu dem Zeitpunkt gilt, zu dem Ihre Spot-Instance ausgeführt wird. Wir empfehlen, keinen Höchstpreis anzugeben, sondern als Höchstpreis standardmäßig den On-Demand-Preis zu verwenden. Ein hoher Höchstpreis erhöht nicht Ihre Chancen, eine Spot-Instance zu starten, und verringert auch nicht die Wahrscheinlichkeit, dass Ihre Spot-Instance unterbrochen wird (weil EC2 Ihre Spot-Instance immer noch unterbrechen kann, wenn sie die Kapazität wieder benötigt).

Der Spot-Preis für einen Instance-Typ in einer Availability Zone kann sich jederzeit ändern, dies geschieht aber im Allgemeinen nicht sehr häufig. AWS veröffentlicht den aktuellen Spot-Preis und die historischen Preise für Spot-Instances über die [DescribeSpotPriceHistory](#)-API sowie in der AWS-Managementkonsole, die die Daten aus der API widerspiegelt. So können Sie Höhe und Zeitpunkte der Schwankungen des Spot-Preises im Verlauf der Zeit einschätzen.

Verwalten von Spot-Instance-Unterbrechungen

Der beste Weg für Sie, um Spot-Instance-Unterbrechungen elegant zu bewältigen und die Auswirkungen auf Ihre Leistung oder Verfügbarkeit zu minimieren, besteht darin, Ihre Anwendung fehlertolerant zu gestalten. Um dies zu erreichen, können Sie die Empfehlungen zum Neuausgleich von EC2-Instances und Spot-Instance-Unterbrechungenbenachrichtigungen nutzen.

Eine EC2-Instance-Empfehlung zum Neuausgleich ist eine neue Funktion, die Sie benachrichtigt, wenn eine Spot-Instance einem erhöhten Risiko einer Unterbrechung ausgesetzt ist. Das Signal gibt Ihnen die Möglichkeit, die Spot-Instance vor der zweiminütigen Spot-Instance-Unterbrechungenbenachrichtigung proaktiv zu verwalten. Sie können entscheiden, Ihren Workload auf neue oder bestehende Spot-Instances auszugleichen, die nicht einem erhöhten Risiko einer Unterbrechung ausgesetzt sind. Wir haben es Ihnen leicht gemacht, dieses Signal zu verwenden, indem wir die Kapazitätsausgleichsfunktion in EC2-Auto Scaling-Gruppen verwenden. Weitere Informationen finden Sie unter [Amazon-EC2-Auto-Scaling-Kapazitätsausgleich](#).

Eine Benachrichtigung über die Unterbrechung der Spot-Instance ist eine Warnung, die zwei Minuten vor der Unterbrechung einer Spot-Instance durch Amazon EC2 ausgegeben wird. Wenn Ihr Workload zeitlich flexibel ist, können Sie Ihre Spot-Instances so konfigurieren, dass sie angehalten oder in den Ruhezustand versetzt werden. Sie werden also nicht beendet, wenn sie unterbrochen werden. Amazon EC2 stoppt Ihre Spot-Instances bei einer Unterbrechung automatisch oder versetzt sie in den Ruhezustand. Die Instances werden automatisch fortgesetzt, wenn wir verfügbare Kapazität haben.

Sie können die Empfehlung zur Neuverteilung der EC2-Instance und/oder den Hinweis zur Unterbrechung der Spot-Instance befolgen, um Ihren Workload unter Berücksichtigung der Fehlertoleranz zu konstruieren. So können Sie Benachrichtigungen erfassen und den Status eines Jobs speichern (z. B. Amazon S3, Amazon EFS oder Amazon FSx), Protokolldateien von der Instance dauerhaft speichern (oder sie kontinuierlich für einen fehlertoleranteren Ansatz streamen), Verbindungen von einer Lastenverteilung leeren usw.

Einige AWS- und Drittanbieter-Services kümmern sich bereits für Sie um Spot-Unterbrechungen, damit die Auswirkungen auf Ihre Anwendung verringert werden. Amazon EKS beispielsweise, das [verwaltete Knotengruppen mit Spot-Instances](#) ausführt, startet automatisch Reserve-Kubernetes-Knoten, wenn eine Empfehlung zur Neuverteilung oder ein Hinweis zur Unterbrechung für einen bestehenden Knoten ausgegeben wird.

Spot-Instance-Limits

Die Anzahl der laufenden und angeforderten Spot-Instances pro AWS-Konto und -Region ist begrenzt. Spot-Instance-Limits werden in Bezug auf die Anzahl der virtuellen Zentralverarbeitungseinheiten (Virtual Central Processing Units, vCPUs) verwaltet, die von Ihren ausgeführten Spot-Instances oder bis zur Erfüllung offener Spot-Instance-Anforderungen verwendet werden. Wenn Sie Ihre Spot-Instances beenden, die Spot-Instance-Anforderungen nicht abbuchen, werden die Anforderungen auf Ihr Spot-Instance-vCPU-Limit angerechnet, bis Amazon EC2 die Beendigung der Spot-Instance erkennt und die Anforderungen schließt.

Es gibt sechs Spot-Instance-Limits:

- Alle Standard-Spot-Instance-Anforderungen (A, C, D, H, I, M, R, T, Z)
- Alle F-Spot-Instance-Anforderungen
- Alle G-Spot-Instance-Anforderungen
- Alle Inf-Spot-Instance-Anforderungen
- Alle P-Spot-Instance-Anforderungen
- Alle X-Spot-Instance-Anforderungen

Jedes Limit gibt das vCPU-Limit für eine oder mehrere Instance-Familien an. Weitere Informationen zu den unterschiedlichen Instance-Familien, Generationen und Größen finden Sie unter [Amazon-EC2-Instance-Typen](#).

Bei vCPU-Limits können Sie Ihr Limit als Anzahl von vCPUs verwenden, die zum Starten einer beliebigen Kombination von Instance-Typen erforderlich sind, die Sie für Ihre sich ändernden Anwendungen benötigen. Angenommen, Ihr Limit für alle Standard-Spot-Instance-Anforderungen liegt bei 256 vCPUs, könnten Sie 32 `m5.2xlarge`-Spot-Instances (32 x 8 vCPUs) oder 16 `c5.4xlarge`-Spot-Instances (16 x 16 vCPUs) oder eine Kombination aus Standard-Spot-Instance-Typen und -Größen mit insgesamt 256 vCPUs anfordern.

Weitere Informationen finden Sie unter [Limits und Nutzung von Spot-Instances überwachen](#) und [Erhöhung des Spot-Instance-Limits anfordern](#) im Amazon-EC2-Benutzerhandbuch für Linux-Instances.

Bewährte Methoden für Spot-Instances

Ihre Budget- und Instance-Typ-Anforderungen sowie Ihr Anwendungskonzept bestimmen, wie die bewährten Methoden für Ihre Anwendung umgesetzt werden:

- Seien Sie flexibel hinsichtlich der Instance-Typen. Ein Spot-Instance-Pool ist ein Satz nicht verwendeter EC2-Instances mit demselben Instance-Typ (z. B. m5.large) sowie derselben Availability Zone (z. B. us-east-1a). Sie sollten flexibel darin sein, welche Instance-Typen Sie anfordern und in welchen Availability Zones Sie Ihren Workload bereitstellen können. Dies gibt Spot eine bessere Chance, die erforderliche Menge an Rechenkapazität zu finden und zuzuweisen. Fragen Sie zum Beispiel nicht nur nach c5.large, wenn Sie bereit wären, Larges aus den Familien c4, m5 und m4 zu verwenden.
- Verwenden Sie die kapazitätsoptimierte Zuweisungsstrategie. Zuweisungsstrategien in EC2-Auto-Scaling-Gruppen helfen Ihnen, Ihre Zielkapazität bereitzustellen, ohne manuell nach den Spot-Instance-Pools mit Reservekapazität suchen zu müssen. Es wird empfohlen, die kapazitätsoptimierte Strategie zu verwenden, da diese Strategie automatisch Instances aus den am häufigsten verfügbaren Spot-Instance-Pools bereitstellt. Da Ihre Spot-Instance-Kapazität aus Pools mit optimaler Kapazität bezogen wird, verringert dies die Wahrscheinlichkeit, dass Ihre Spot-Instances unterbrochen werden. Weitere Informationen zu Zuweisungsstrategien finden Sie unter [Spot-Instances](#) im Amazon-EC2-Auto-Scaling-Benutzerhandbuch.
- Verwenden Sie den proaktiven Kapazitätsausgleich. Der Kapazitätsausgleich hilft Ihnen, die Verfügbarkeit von Workloads aufrechtzuerhalten, indem Sie Ihre Auto-Scaling-Gruppe proaktiv um eine neue Spot-Instance erweitern, bevor eine laufende Spot-Instance die zweiminütige Unterbrechungsbenachrichtigung erhält. Wenn der Kapazitätsausgleich aktiviert ist, versucht Auto Scaling proaktiv Spot-Instances zu ersetzen, die eine Empfehlung zum Neuausgleich erhalten haben, und bietet die Möglichkeit, Ihren Workload auf neue Spot-Instances auszugleichen, die nicht einem erhöhten Unterbrechungsrisiko ausgesetzt sind.
- Verwenden Sie integrierte AWS-Services zur Verwaltung Ihrer Spot-Instances. Andere AWS-Services werden in Spot integriert, um die gesamten Rechenkosten zu senken, ohne dass einzelne Instances oder Flotten verwaltet werden müssen. Wir empfehlen Ihnen, die folgenden Lösungen für Ihre jeweiligen Workloads in Betracht zu ziehen: Amazon EMR, Amazon ECS, AWS Batch, Amazon EKS, SageMaker, AWS Elastic Beanstalk und Amazon GameLift. Weitere Informationen zu bewährten Methoden für Spot-Instances mit diesen Services finden Sie auf der [Amazon-EC2-Spot-Instances-Workshops-Website](#).

- Wählen Sie das moderne und korrekte Starttool für Spot-Instances aus. Wenn einer der integrierten AWS-Services nicht gut zu Ihrem Workload passt und Sie Ihre Anwendung dennoch mit Kontrolle über den Start von Spot-Instances erstellen müssen, verwenden Sie das richtige Tool. Für die meisten Workloads sollten Sie EC2 Auto Scaling verwenden, da es einen umfassenderen Funktionssatz für eine Vielzahl von Workloads bietet, z. B. ELB-gestützte Anwendungen, containerisierte Workloads und Warteschlangen-Verarbeitungsaufträge. Wenn Sie mehr Kontrolle über einzelne Anforderungen haben möchten und nach einem Tool suchen, das nur zum Start verwendet werden kann, verwenden Sie EC2 Fleet im Sofortmodus als Drop-In-Ersatz für RunInstances, aber mit einem breiteren Spektrum an Funktionen, wie Diversifizierung von Instance-Typen und Zuweisungsstrategien.

Spot-Integration in andere AWS-Services

Amazon-EC2-Spot-Instances sind in mehrere AWS-Services integriert.

Amazon-EMR-Integration

Sie können Amazon-EMR-Cluster auf Spot-Instances ausführen und die Kosten für die Verarbeitung großer Datenmengen für Ihre Analyse-Workloads erheblich senken. Sie können Ihre EMR-Cluster ausführen, indem Sie Spot-Instances einfach mit On-Demand- und Reserved Instances mischen. Verwenden Sie dafür die [EMR-Instance-Flotten](#)-Funktion. Sie können [EMR-Zuweisungsstrategien](#) verwenden, um Spot-Instances aus den am häufigsten verfügbaren Kapazitätspools zu starten.

EC2-Auto-Scaling-Integration

Sie können [Amazon-EC2-Auto-Scaling](#)-Gruppen verwenden, um Spot-Instances zu starten und zu verwalten, die Anwendungsverfügbarkeit aufrechtzuerhalten, die Auswahl für Instance-Typ und Kaufoption (On-Demand/Spot) zu diversifizieren und Ihre Amazon-EC2-Kapazität mithilfe dynamischer, geplanter und vorausschauender Skalierungsrichtlinien zu skalieren. Weitere Informationen finden Sie unter [Anfordern von Spot-Instances für fehlertolerante und flexible Anwendungen](#) im Amazon-EC2-Auto-Scaling-Benutzerhandbuch.

Amazon-EKS-Integration

Sie können die Kosten für Ihre Kubernetes-basierten Workloads mithilfe von Amazon EKS optimieren, indem Sie Spot-Instances in von EKS verwalteten Knotengruppen starten. Von EKS verwaltete Knotengruppen verwalten den gesamten Spot-Instance-Lebenszyklus, indem sie Spot-Instances, die bald unterbrochen werden, durch neu gestartete Instances ersetzen. Damit wird die Wahrscheinlichkeit, dass sich eine Unterbrechung der Spot-Instances auf die Leistung oder Verfügbarkeit Ihrer Anwendung auswirkt, verringert (wenn EC2 die Kapazität zurück benötigt). Weitere Informationen finden Sie unter [Verwaltete Knotengruppen](#) im Amazon-EKS-Benutzerhandbuch.

Amazon-ECS-Integration

Sie können Amazon-ECS-Cluster auf Spot-Instances ausführen, um die Betriebskosten für die Ausführung containerisierter Anwendungen zu senken. Amazon ECS unterstützt den automatischen

Ausgleich von Spot-Instances, die bald unterbrochen werden. Weitere Informationen finden Sie unter [Verwenden von Spot-Instances](#) im Entwicklerhandbuch zum Amazon Elastic Container Service.

Amazon ECS mit AWS-Fargate-Spot-Integration

Wenn Ihre containerisierten Aufgaben unterbrechbar und flexibel sind, können Sie Ihre ECS-Aufgaben mit dem AWS-Fargate-Spot-Kapazitätsanbieter ausführen. Das bedeutet, dass Ihre Aufgaben auf AWS Fargate, einer Plattform für Serverless-Container, ausgeführt werden und Sie von Kosteneinsparungen durch Fargate Spot profitieren. Weitere Informationen finden Sie unter [AWS-Fargate-Kapazitätsanbieter](#) im Entwicklerhandbuch zum Amazon Elastic Container Service.

Amazon-Batch-Integration

[AWS Batch](#) plant Batchverarbeitungs-Workloads auf AWS, legt den Zeitplan für diese fest und führt sie aus. AWS Batch fordert in Ihrem Auftrag auch dynamisch Spot-Instances an und trägt somit zur Reduzierung der Kosten für das Ausführen von Batchaufträgen bei.

Amazon-SageMaker-Integration

Amazon SageMaker vereinfacht das Schulen von Machine-Learning-Modellen mithilfe von verwalteten Spot-Instances. Mithilfe von Managed Spot Training können die Kosten für die Schulung von Modellen über On-Demand-Instances um bis zu 90 % optimiert werden. SageMaker verwaltet Spot-Unterbrechungen für Sie. Weitere Informationen finden Sie unter [Managed Spot Training in Amazon SageMaker](#) im Entwicklerhandbuch zu Amazon SageMaker.

Amazon-GameLift-Integration

Amazon GameLift ist eine Hosting-Lösung für Game-Server, die Cloud-Server für Multiplayer-Spiele bereitstellt, betreibt und skaliert. Durch die Unterstützung von Spot-Instances in Amazon GameLift können Sie Ihre Hosting-Kosten erheblich senken. Beim Erstellen von Flotten von Hosting-Ressourcen können Sie zwischen On-Demand-Instances oder Spot-Instances wählen. Während Spot-Instances innerhalb von zwei Minuten nach einer entsprechenden Benachrichtigung unterbrochen werden können, minimiert FleetiQ von Amazon GameLift das Risiko von Unterbrechungen. Weitere Informationen finden Sie unter [Verwenden von Spot-Instances mit GameLift](#) im Entwicklerhandbuch zu Amazon GameLift.

AWS-Elastic-Beanstalk-Integration

AWS Elastic Beanstalk ist ein benutzerfreundlicher Service zum Bereitstellen und Skalieren von Webanwendungen und -Services, die mit Java, .NET, PHP, Node.js, Python, Ruby, Go und Docker auf vertrauten Servern wie Apache, Nginx, Passenger und IIS entwickelt werden. Sie laden Ihren Code einfach hoch und Elastic Beanstalk übernimmt automatisch die Bereitstellung – von der Kapazitätsbereitstellung über die Lastenverteilung und automatische Skalierung bis zur Statusüberwachung der Anwendung. Sie können Spot-Instances in Ihren Elastic Beanstalk-Umgebungen zur Kostenoptimierung der zugrunde liegenden Infrastruktur Ihrer Webanwendungen verwenden. Informationen zur Verwendung von Spot-Instances mit Elastic Beanstalk finden Sie unter [Spot-Instance-Unterstützung](#) im AWS Elastic BeanstalkEntwicklerhandbuch.

Fazit

Egal, ob Sie flexible Datenverarbeitungsanforderungen haben oder die Kapazität erweitern möchten, ohne Ihr Budget zu erhöhen – Spot-Instances können eine hervorragende Möglichkeit sein, Ihre AWS-Kosten zu optimieren und/oder bei der Entwicklung auch die Skalierbarkeit zu berücksichtigen. Durch die richtige Architektur Ihrer Workloads können Sie Spot-Instances für eine Vielzahl von Anforderungen nutzen. Weitere Informationen finden Sie unter [Amazon-EC2-Spot-Instances](#).

Ressourcen

- [AWS-Architekturzentrum](#)
- [AWS-Whitepaper](#)
- [AWS Architecture Monthly](#)
- [AWS-Architekturblog](#)
- [„This Is My Architecture“-Videos](#)
- [AWS-Dokumentation](#)

Dokumentverlauf und Mitwirkende

Dokumentverlauf

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

Update-Historie-Änderung	Update-Historie-Beschreibung	Update-Historie-Datum
Kleineres Update	Seitenlayout angepasst.	30. April 2021
Kleineres Update	Inhalt wurde aktualisiert, um aktuelle bewährte Methoden wiederzugeben. Der Name des Whitepapers wurde von „Nutzen von Amazon-EC2-Spot-Instances nach Maß“ in „Übersicht über Amazon-EC2-Spot-Instances“ geändert, um den Inhalt besser wiederzugeben.	5. März 2021
Kleineres Update	Spot-Instance-Limits wurden aktualisiert.	3. Februar 2021
Erstveröffentlichung	Nutzen von Amazon-EC2-Spot-Instances nach Maß veröffentlicht.	1. März 2018

Note

Um RSS-Aktualisierungen zu abonnieren, muss für den von Ihnen verwendeten Browser ein RSS-Plug-In aktiviert sein.

Mitwirkende

Dieses Dokument ist unter der Mitarbeit folgender Personen und Organisationen entstanden:

- Amilcar Alfaro, Sr. Product Marketing Manager, AWS
- Erin Carlson, Marketing Manager, AWS
- Keith Jarrett, WW BD Lead – Cost Optimization, AWS Business Development
- Ran Sheinberg, Principal Solutions Architect, AWS