Bewährte Methoden zur Optimierung der Amazon EKS-Observability

AWS Präskriptive Leitlinien

AWS Präskriptive Leitlinien: Bewährte Methoden zur Optimierung der Amazon EKS-Observability

Table of Contents

Bewährte Methoden zur Optimierung der Amazon EKS-Observability

Ishwar Chauthaiwale, Naveen Suthar und Pratap Kumar Nanda, Amazon Web Services ()AWS

März 2026 (Geschichte der Dokumente)

Amazon Elastic Kubernetes Service (Amazon EKS) benötigt umfassende Observability-Lösungen, um containerisierte Workloads effektiv zu überwachen und Fehler zu beheben. Verteilte Systeme und Microservices haben komplexe Architekturen in Amazon EKS-Umgebungen. Daher ist die Implementierung geeigneter Beobachtbarkeitspraktiken für die Aufrechterhaltung eines zuverlässigen Betriebs von entscheidender Bedeutung. Effektive Beobachtbarkeit in Amazon EKS-Umgebungen ermöglicht es Teams, tiefe Einblicke in die Anwendungsleistung zu gewinnen, Probleme effizient zu beheben und einen optimalen Cluster-Zustand aufrechtzuerhalten.

Die Herausforderung besteht darin, sich im riesigen Ökosystem von Tools und Techniken zurechtzufinden, die für Amazon EKS-Observability verfügbar sind, und gleichzeitig bewährte Verfahren einzuhalten, die den Unternehmenszielen und Industriestandards entsprechen. Effektive Observability-Strategien müssen ein Gleichgewicht zwischen umfassender Datenerfassung und Leistungsaspekten, Kosteneffektivität und Skalierbarkeit herstellen.

Dieser Leitfaden soll Unternehmen dabei helfen, ihre Amazon EKS-Beobachtbarkeit in den folgenden Bereichen zu optimieren:

- Einrichtung effizienter Protokollierungsmechanismen
- Implementierung robuster Überwachungslösungen
- Einsatz von verteilter Ablaufverfolgung für komplexe Architekturen
- Implementierung von Strategien zur Warnung und Reaktion auf Vorfälle

Durch die Übernahme dieser Best Practices kann Ihr Unternehmen seine Fähigkeit verbessern, tiefe Einblicke in seine Amazon EKS-Umgebung zu gewinnen, was zu einer verbesserten Zuverlässigkeit, Leistung und Betriebseffizienz führt. Dieser optimierte Beobachtungsansatz hilft bei der Fehlerbehebung und Wartung und unterstützt datengestützte Entscheidungen zur kontinuierlichen Verbesserung von Kubernetes-basierten Anwendungen und Infrastrukturen. (Detaillierte Informationen zu Amazon EKS finden Sie in der Servicedokumentation.)

Dieser Leitfaden befasst sich eingehend mit jedem Aspekt der Amazon EKS-Observability und untersucht die Tools und Strategien, die Sie an die spezifischen Anforderungen Ihrer Amazon EKS-Bereitstellungen anpassen können, von kleinen Anwendungen bis hin zu großen, komplexen Microservices-Architekturen.

In diesem Leitfaden:

Ziele

Dieser Leitfaden kann Ihnen und Ihrer Organisation dabei helfen, die folgenden Geschäftsziele zu erreichen:

- Verbesserte betriebliche Transparenz — Verschaffen Sie sich mithilfe effektiver Beobachtungspraktiken umfassende Einblicke in Ihre Amazon EKS-Cluster und -Anwendungen.

  Dieses Ziel unterstreicht, wie wichtig es ist, die vollständige Transparenz Ihrer gesamten Amazon EKS-Umgebung aufrechtzuerhalten. Tools wie AWS X-RayAmazon CloudWatch Container Insights und AWS Distro OpenTelemetry helfen Ihnen dabei, das Systemverhalten zu verstehen, Probleme schnell zu identifizieren und eine optimale Leistung aufrechtzuerhalten.
- Verbesserte Effizienz bei der Fehlerbehebung — Reduzieren Sie die mittlere Zeit bis zur Erkennung (MTTD) und die mittlere Zeit bis zur Problemlösung (MTTR) durch effektive Verfolgungs- und Überwachungsstrategien.

  Dieses Ziel konzentriert sich auf die Implementierung von Beobachtungspraktiken, die eine schnelle Identifizierung und Lösung von Problemen ermöglichen. Techniken wie verteilte Rückverfolgung, effektive Protokollierung und umfassende Erfassung von Kennzahlen sind entscheidend, um dieses Ziel zu erreichen.
- Proaktives Leistungsmanagement — Ermöglichen Sie die Früherkennung potenzieller Probleme, bevor sie sich auf Endbenutzer auswirken.

Eine proaktive Überwachung ist entscheidend für die Aufrechterhaltung einer hohen Serviceverfügbarkeit und Leistung. Dieses Ziel trägt der Bedeutung der Implementierung geeigneter Warnmeldungen, Trendanalysen und vorausschauender Überwachung Rechnung, um Serviceunterbrechungen zu verhindern.

- Kostengünstige Beobachtbarkeit — Optimieren Sie die Kosten für die Beobachtbarkeit und sorgen Sie gleichzeitig für eine umfassende Systemtransparenz.

  Die Kostenoptimierung umfasst die Implementierung effizienter Probenahmestrategien, geeigneter Richtlinien zur Datenspeicherung und optimaler Instrumentierungsansätze. Ziel ist es, die Anforderungen an die Beobachtbarkeit mit Kostenerwägungen in Einklang zu bringen und gleichzeitig eine effektive Systemüberwachung sicherzustellen.
- Skalierbare Überwachungsarchitektur — Stellen Sie sicher, dass Ihre Observability-Lösungen nahtlos mit Ihrer Amazon EKS-Umgebung skaliert werden.

  Dieses Ziel konzentriert sich auf die Implementierung von Überwachungslösungen, die mit Ihrer Anwendung mitwachsen können. Ganz gleich, ob Sie einen einzelnen Cluster oder eine Bereitstellung mit mehreren Clustern und mehreren Regionen betreiben, Ihre Observability-Strategie sollte entsprechend skaliert werden

Anmeldung bei Amazon EKS

Die Protokollierung ist ein wichtiger Aspekt bei der Verwaltung und Wartung von Anwendungen, die auf Amazon EKS ausgeführt werden. Effektive Protokollierungspraktiken in Amazon EKS-Umgebungen helfen Entwicklern, Betriebsteams und Systemadministratoren, wertvolle Einblicke in das Verhalten, die Leistung und den Zustand ihrer containerisierten Anwendungen und der zugrunde liegenden Infrastruktur zu gewinnen.

Die Implementierung einer robusten Protokollierungsstrategie in Amazon EKS ist aus mehreren Gründen unerlässlich:

- Fehlerbehebung: Mithilfe von Protokollen können Probleme schnell erkannt und diagnostiziert werden, wodurch Ausfallzeiten reduziert und die allgemeine Systemzuverlässigkeit verbessert wird.
- Einhaltung von Vorschriften: In vielen Branchen ist eine umfassende Protokollierung zu Prüfungs- und Regulierungszwecken erforderlich.
- Sicherheit: Die Protokollanalyse kann Ihnen helfen, potenzielle Sicherheitsbedrohungen oder Sicherheitsverletzungen zu erkennen und zu untersuchen.
- Leistungsoptimierung: Protokolle bieten Einblicke in die Anwendungs- und Systemleistung, sodass Sie Engpässe erkennen und die Ressourcennutzung optimieren können.
- Überwachung und Alarmierung: Protokolldaten können verwendet werden, um Überwachungssysteme einzurichten und Warnmeldungen für bestimmte Ereignisse oder Bedingungen auszulösen.

In diesem Abschnitt:

Arten der Anmeldung in Amazon EKS

In Amazon EKS umfasst die Protokollierung das Erfassen, Speichern und Analysieren verschiedener Arten von Protokolldaten, die von verschiedenen Komponenten des Kubernetes-Clusters generiert werden, darunter:

- Systemprotokolle: Informationen über die zugrunde liegenden Amazon Elastic Compute Cloud (Amazon EC2) -Instances oder -Knoten AWS Fargate
- Kubernetes-Komponentenprotokolle : Daten aus Kernkomponenten von Kubernetes wie dem API-Server, dem Scheduler und dem Controller-Manager
- Container-Runtime-Logs: Informationen aus der Container-Laufzeit, wie Docker oder containerd
- Anwendungsprotokolle: Ausgabe von containerisierten Anwendungen

Um Protokolle in Ihrer Amazon EKS-Umgebung effektiv zu verwalten, verwenden Sie in der AWS-Services Regel eine Kombination aus Tools von Drittanbietern und bewährten Methoden. Dies kann die Verwendung von Amazon CloudWatch, Fluent Bit, Elasticsearch, Kibana und anderen Protokollierungs- und Analysetools zur Erfassung, Speicherung und Visualisierung von Protokolldaten beinhalten.

In den folgenden Abschnitten werden verschiedene Aspekte der Protokollierung in Amazon EKS untersucht, darunter bewährte Methoden, Tools und Techniken für die Implementierung einer umfassenden Protokollierungsstrategie in Ihren Kubernetes-Clustern auf. AWS

Systemprotokolle

Die Protokollierung für zugrunde liegende EC2-Instances oder Fargate-Knoten in Amazon EKS umfasst je nach Knotentyp unterschiedliche Ansätze.

Um die Protokollierung für EC2-Instances in Amazon EKS zu implementieren, können Sie die folgenden Tools verwenden:

- CloudWatch Agent: Installieren und konfigurieren Sie den CloudWatch Agenten auf Ihren EC2-Instances. Konfigurieren Sie ihn so, dass Systemprotokolle wie `/var/log/messages` und `/var/log/secure` gesammelt werden. Sie können Benutzerdatenskripts oder Konfigurationsmanagement-Tools verwenden, um diesen Prozess zu automatisieren.
- Fluent Bit: Stellen Sie Fluent Bit bereit, DaemonSet um Protokolle von allen Knoten zu sammeln. Konfigurieren Sie es so, dass Protokolle an CloudWatch Logs oder andere zentrale Protokollierungssysteme weitergeleitet werden.
- Container Insights: Aktivieren Sie Container Insights in Ihrem EKS-Cluster, um automatisch Metriken und Protokolle von EC2-Instances zu sammeln.
- Benutzerdefinierte Skripts: Entwickeln Sie benutzerdefinierte Skripts, um bestimmte Protokolle zu sammeln und sie an Ihr bevorzugtes Logging-Ziel zu senden.

- SSM-Agent: Verwenden Sie den AWS Systems Manager Agenten (SSM-Agent), um Protokolle zu sammeln und an Logs CloudWatch weiterzuleiten.

Verwenden Sie die folgenden Tools, um die Protokollierung für Fargate-Knoten in Amazon EKS zu implementieren:

- Fargate-Logging: Fargate sammelt `stdout` und `stderr` protokolliert automatisch Ihre Container. Konfigurieren Sie Ihr Fargate-Profil, um diese Protokolle an Logs zu CloudWatch senden.
- Fluent Bit for Fargate: AWS bietet ein Fluent Bit-Bild speziell für die Fargate-Protokollierung. Setze ihn als Beiwagencontainer in deinen Fargate-Pods ein, um Logs zu sammeln und weiterzuleiten.
- Container Insights für Fargate: Aktivieren Sie Container Insights, um Metriken und Protokolle von Fargate-Knoten zu sammeln.

Protokolle der Kubernetes-Komponenten

Das Sammeln von Protokollen aus Kubernetes-Komponenten wie dem API-Server, dem Scheduler und dem Controller-Manager in Amazon EKS erfordert einen etwas anderen Ansatz als die Anwendungsprotokollierung. Diese Komponenten werden als Teil der Amazon EKS-Steuerebene ausgeführt, die von verwaltet wird AWS. So können Sie diese Protokolle sammeln und darauf zugreifen:

- Protokollierung auf Kontrollebene aktivieren: Sie können die Protokollierung der Kontrollebene für Ihren EKS-Cluster über die Tools AWS-Managementkonsole, AWS Command Line Interface (AWS CLI) oder Infrastructure as Code (IaC) wie AWS CloudFormationoder Terraform aktivieren. Wenn Sie die Protokollierung auf der Kontrollebene aktivieren, werden die Protokolle an Amazon CloudWatch Logs gesendet. Sie können sie in der CloudWatch Konsole in der `/aws/eks/<cluster-name>/cluster` Protokollgruppe anzeigen. Innerhalb dieser Protokollgruppe hat jede Komponente der Steuerungsebene ihren eigenen Protokollstream wie folgt:

Name des Datenstroms	Description
Kube-Apiserver	Kubernetes-API-Serverprotokolle
Kube-Scheduler	Entscheidungsprotokolle des Schedulers
kube-controller-manager	Controller-Manager-Protokolle

Name des Datenstroms	Description
Authentifikator	IAM-Authentifikator-Protokolle
audit	Kubernetes-Audit-Logs (müssen explizit aktiviert werden)

Um Logs für eine bestimmte Komponente anzuzeigen, navigieren Sie zur Cluster-Log-Gruppe und filtern Sie nach dem Namen des Ziel-Log-Streams.

- CloudWatch Logs Insights verwenden: Sie können CloudWatch Logs Insights verwenden, um komplexe Abfragen in Ihren Logs durchzuführen.
- Protokolle nach Amazon S3 exportieren: Für die langfristige Speicherung oder weitere Analysen können Sie Protokolle nach Amazon Simple Storage Service (Amazon S3) exportieren.
- Verwenden Sie Tools von Drittanbietern: Sie können Tools wie Fluent Bit verwenden, um diese Protokolle zu sammeln und an andere Protokollierungssysteme wie Elasticsearch oder Splunk weiterzuleiten.
- Verwendung AWS CloudTrail: Der AWS CloudTrailService kann zusätzliche Einblicke in API-Aufrufe an Ihren EKS-Cluster bieten.

Container-Laufzeitprotokolle

Das Protokollieren von Container-Laufzeitprotokollen in Amazon EKS beinhaltet das Erfassen und Verwalten von Protokollen aus der Container-Laufzeit, was in der Regel `containerd` für Amazon EKS gilt. So können Sie die Protokollierung von Container-Laufzeitprotokollen in Amazon EKS angehen:

- Greifen Sie direkt auf die Protokolle auf Amazon EC2 EC2-Knoten zu. Bei selbstverwalteten EC2-Knoten können Sie von diesen Speicherorten aus direkt auf die Container-Laufzeitprotokolle auf dem Host zugreifen:
  - `containerd`Protokolle: `/var/log/containers/`
  - Docker-Logs (wenn Sie die Docker-Laufzeit verwenden): `/var/log/docker.log`
- Verwenden Sie a DaemonSet für die Protokollerfassung.
- Stellen Sie einen Protokollerfassungsagenten (z. B. Fluent Bit) bereit DaemonSet , um Protokolle von allen Knoten zu sammeln.

- Konfigurieren Sie den CloudWatch Agenten so, dass er Container-Laufzeitprotokolle sammelt.
- Aktivieren Sie Container Insights, um Metriken und Logs zur Container-Laufzeit zu sammeln.
- Benutze Fargate. Für Fargate-Knoten werden Container-Laufzeitprotokolle automatisch gesammelt und können über CloudWatch Logs abgerufen werden.
- Implementieren Sie benutzerdefinierte Logging-Lösungen mithilfe von Tools wie Fluent Bit oder Logstash. Richten Sie CloudWatchAlarme ein oder verwenden Sie Tools wie Prometheus, um in Container-Laufzeitprotokollen nach bestimmten Mustern oder Problemen zu suchen. Erwägen Sie die Verwendung von Protokollierungslösungen von Drittanbietern, die sich gut in Kubernetes und Amazon EKS integrieren lassen, wie Datadog, Splunk oder den Elastic Stack (ELK Stack). Verwenden Sie Tools zur Protokollaggregation, um Protokolle aus mehreren Quellen zu sammeln und sie an ein zentrales Protokollierungssystem weiterzuleiten.

Anwendungsprotokolle

Anwendungsprotokolle in Amazon EKS sind ein wichtiger Bestandteil der Wartung und Fehlerbehebung Ihrer Anwendungen. Um die Anwendungsprotokollierung in Amazon EKS zu implementieren, können Sie aus den folgenden Optionen wählen:

- Protokolle nach`stdout`/schreiben`stderr`: Die einfachste und Kubernetes-systemativste Methode, Anwendungsprotokolle zu verarbeiten, besteht darin, sie in und zu schreiben. `stdout` `stderr` Kubernetes erfasst diese Streams automatisch.
- Implementieren Sie die Protokollaggregation: Verwenden Sie einen Log-Aggregator wie Fluent Bit, um Logs von all Ihren Pods zu sammeln.
- Log-Routing konfigurieren: Konfigurieren Sie Ihren Log-Aggregator so, dass Logs an Ihr gewünschtes Ziel weitergeleitet werden (wie CloudWatch Logs oder Elasticsearch).
- CloudWatch Container Insights verwenden: Aktivieren Sie Container Insights für eine umfassende Protokollierung und Überwachung.

Bewährte Methoden für die Anmeldung bei Amazon EKS

Die folgenden bewährten Methoden helfen dabei, ein robustes, skalierbares und effizientes Protokollierungssystem für Ihre Amazon EKS-Umgebung zu erstellen und sorgen für eine bessere Fehlerbehebung, Überwachung und allgemeine Verwaltung Ihrer Kubernetes-Cluster.

- Zentralisieren Sie die Protokollerfassung: Verwenden Sie eine zentralisierte Protokollierungslösung wie CloudWatch Logs, Elasticsearch oder einen Drittanbieter-Service, um Logs aus allen Komponenten zu aggregieren. Dies bietet einen zentralen Zugriffspunkt für die Protokollanalyse und vereinfacht die Verwaltung.
- Implementieren Sie strukturierte Protokollierung: Verwenden Sie strukturierte Protokollformate wie JSON, damit Protokolle einfacher analysiert und durchsucht werden können. Fügen Sie relevante Metadaten wie Zeitstempel, Protokollebenen und Quellkennungen hinzu.
- Verwenden Sie Protokollebenen angemessen: Implementieren Sie die richtigen Protokollebenen (wie `DEBUG`,`INFO`,`WARN`, und`ERROR`) in Ihren Anwendungen. Konfigurieren Sie Produktionsumgebungen so, dass sie auf geeigneten Ebenen protokollieren, um eine übermäßige Protokollierung zu vermeiden.
- Container-Logging aktivieren: Konfigurieren Sie Ihre Container so, dass sie sich bei `stdout` und anmelden`stderr`. Auf diese Weise kann Kubernetes diese Protokolle erfassen und an die von Ihnen gewählte Logging-Lösung weiterleiten.
- Anwendungsprotokollierung aktivieren: Konfigurieren Sie Anwendungen so, dass sie Protokolle in `stdout` und `stderr` statt in Protokolldateien schreiben. Dies folgt der 12-Faktor-App-Methodik und entspricht den Cloud-nativen Best Practices.
- Verwenden Sie Kubernetes DaemonSets für die Protokollerfassung: Stellen Sie Agenten für die Protokollerfassung (wie Fluent Bit) bereit, DaemonSets um sicherzustellen, dass sie auf jedem Knoten in Ihrem Cluster ausgeführt werden.
- Implementieren Sie Aufbewahrungsrichtlinien: Definieren Sie Richtlinien zur Aufbewahrung von Protokollen und setzen Sie diese durch, um Vorschriften einzuhalten und die Speicherkosten zu kontrollieren.
- Sichere Protokolldaten: Verschlüsseln Sie Protokolle während der Übertragung und im Ruhezustand. Implementieren Sie Zugriffskontrollen, um einzuschränken, wer Protokolle einsehen und verwalten kann.
- Überwachen Sie die Protokollaufnahme: Richten Sie Warnmeldungen für Fehler oder Verzögerungen bei der Protokollaufnahme ein, um eine kontinuierliche Protokollierung sicherzustellen.
- Verwenden Sie Kubernetes-Anmerkungen und -Labels: Verwenden Sie Kubernetes-Anmerkungen und -Labels, um Metadaten zu Ihren Logs hinzuzufügen und so die Durchsuchbarkeit und Filterung zu verbessern.
- Implementieren Sie verteiltes Tracing: Verwenden Sie verteilte Tracing-Tools wie oder Jaeger, um Logs zwischen Microservices zu korrelieren. AWS X-Ray

- Optimieren Sie das Protokollvolumen: Wählen Sie sorgfältig aus, was Sie protokollieren, um unnötige Kosten und Leistungsprobleme zu vermeiden. Verwenden Sie Stichproben für Protokolle mit hohem Volumen und geringem Wert.
- Implementieren Sie die Protokollaggregation: Verwenden Sie Tools wie Logstash, um Logs aus mehreren Quellen zu aggregieren, bevor Sie sie an Ihr zentrales Logging-System senden.
- Verwenden Sie AWS-Services , wenn möglich: Dienste wie CloudWatch Logs und Container Insights bieten eine nahtlose Integration mit anderen. AWS-Services
- Implementieren Sie Protokollanalyse und -visualisierung: Verwenden Sie Tools wie CloudWatch Logs Insights, Elasticsearch mit Kibana oder Lösungen von Drittanbietern für die Protokollanalyse und -visualisierung.
- Implementieren Sie automatisierte Protokollanalysen: Verwenden Sie maschinelles Lernen und KI-gestützte Tools, um Anomalien und Muster in Ihren Logs automatisch zu erkennen.
- Dokumentieren Sie Ihre Protokollierungsstrategie: Sorgen Sie für eine klare Dokumentation Ihrer Protokollierungsarchitektur, -praktiken und -tools für Ihr Team.

Wichtige Überlegungen zur Anmeldung bei Amazon EKS

In diesem Abschnitt werden wichtige Überlegungen erläutert, die Sie bei der Implementierung der Protokollierung in Amazon EKS berücksichtigen sollten.

- Auswirkungen auf die Leistung: Eine übermäßige Protokollierung kann die Anwendungsleistung beeinträchtigen. Achten Sie auf das Volumen und die Häufigkeit der generierten Protokolle.
- Kostenmanagement: Die Speicherung und Verarbeitung von Protokollen kann zu erheblichen Kosten führen, insbesondere bei großem Umfang. Implementieren Sie Richtlinien zur Aufbewahrung von Protokollen und erwägen Sie die Verwendung von Protokollaggregation, um die Kosten zu senken.
- Sicherheit und Compliance: Stellen Sie sicher, dass die Protokolle keine vertraulichen Informationen wie Passwörter oder persönliche Daten enthalten. Implementieren Sie Verschlüsselung für Protokolle bei der Übertragung und Speicherung. Beachten Sie bei der Verarbeitung von Protokollen Compliance-Anforderungen wie die Allgemeine Datenschutzverordnung (DSGVO) oder den Health Insurance Portability and Accountability Act (HIPAA).
- Skalierbarkeit: Stellen Sie sicher, dass Ihre Logging-Lösung mit Ihrer Clustergröße und Ihrem Protokollvolumen skaliert werden kann. Erwägen Sie die Verwendung von Pufferung und Batching für die Protokollübertragung.

- Aufbewahrung von Protokollen: Definieren und implementieren Sie geeignete Aufbewahrungsfristen für Protokolle. Sorgen Sie für ein ausgewogenes Verhältnis zwischen Compliance-Anforderungen und Speicherkosten.
- Zugriffskontrolle: Implementieren Sie die richtigen AWS Identity and Access Management (IAM-) Rollen und Richtlinien für den Protokollzugriff. Halten Sie sich bei der Protokollverwaltung an das Prinzip der geringsten Rechte.
- Protokollkonsistenz: Verwenden Sie konsistente Protokollformate für verschiedene Anwendungen und Dienste. Verwenden Sie die strukturierte Protokollierung für eine einfachere Analyse und Analyse.
- Zeitsynchronisierung: Synchronisieren Sie die Zeit auf allen Knoten, um konsistente Zeitstempel in den Protokollen zu erhalten.
- Ressourcenzuweisung: Weisen Sie den Logging-Agenten entsprechende Ressourcen (wie CPU und Arbeitsspeicher) zu. Überwachen Sie die Ressourcennutzung der Logging-Komponenten.
- Überlegungen zu Fargate: Fargate verfügt über spezifische Protokollierungsmechanismen, die sich von EC2-basierten Knoten unterscheiden. Machen Sie sich mit den Einschränkungen und Möglichkeiten der Fargate-Protokollierung vertraut.
- Mehrmandanten-Cluster: Stellen Sie in Umgebungen mit mehreren Mandanten sicher, dass die Protokolle ordnungsgemäß zwischen den Mandanten isoliert sind.
- Analyse und Analyse von Protokollen: Berücksichtigen Sie die Tools und Fähigkeiten, die für eine effektive Protokollanalyse erforderlich sind. Implementieren Sie die Protokollanalyse für die Extraktion strukturierter Daten.
- Überwachung des Protokollierungssystems: Richten Sie die Überwachung für die Protokollierungsinfrastruktur selbst ein. Generieren Sie Warnmeldungen zur Protokollierung von Systemausfällen oder Rückständen.
- Auswirkungen auf das Netzwerk: Beachten Sie die Netzwerkbandbreite, die für die Protokollübertragung verwendet wird. Erwägen Sie die Verwendung von Komprimierung für Protokolldaten.
- Kubernetes-Ereignisse: Übersehen Sie Kubernetes-Ereignisse nicht als Quelle wichtiger Informationen.
- Protokollierung auf Kontrollebene: Machen Sie sich mit den Auswirkungen und Kosten der Aktivierung der Protokollierung auf der Kontrollebene vertraut.
- Debugging-Funktionen: Stellen Sie sicher, dass Ihre Logging-Lösung ein einfaches Debuggen und Problembehebungen ermöglicht.

- Integration mit vorhandenen Tools: Überlegen Sie, wie sich Ihre Amazon EKS-Protokollierungslösung in bestehende Überwachungs- und Warnungstools integrieren lässt.
- Testen: Testen Sie regelmäßig Ihr Logging-Setup, insbesondere nach Cluster-Upgrades.
- Dokumentation: Sorgen Sie für eine klare Dokumentation Ihrer Protokollierungsarchitektur und -praktiken.
- Latenz bei der Protokollaggregation: Achten Sie auf etwaige Latenzen bei der Protokollaggregation und darauf, wie sich diese auf die Echtzeitüberwachung auswirken kann.

Überwachung in Amazon EKS

Die Überwachung in Amazon EKS bietet wichtige Einblicke in den Zustand, die Leistung und die Sicherheit Ihrer Kubernetes-Workloads. Ohne angemessene Überwachung riskieren Sie Serviceunterbrechungen, Sicherheitsverletzungen und eine ineffiziente Ressourcennutzung, die sich auf den Geschäftsbetrieb auswirken und die Kosten in die Höhe treiben können. Eine effektive Überwachung ermöglicht es Ihnen, Probleme proaktiv zu identifizieren und zu lösen, die Ressourcennutzung zu optimieren und die Compliance-Anforderungen für Ihre containerisierten Anwendungen einzuhalten. Durch die Implementierung umfassender Überwachungslösungen können Sie eine hohe Verfügbarkeit sicherstellen, Anomalien frühzeitig erkennen und datengestützte Entscheidungen zur Skalierung und Verbesserung Ihrer Amazon EKS-Infrastruktur treffen.

In diesem Abschnitt werden die verschiedenen Aspekte der Amazon EKS-Überwachung untersucht, darunter verschiedene Überwachungstypen, verfügbare Tools und bewährte Methoden, die Ihnen helfen, eine robuste Überwachungsstrategie für Ihre Kubernetes-Umgebung zu entwickeln.

In diesem Abschnitt:

Arten der Überwachung in Amazon EKS

Effektive Beobachtbarkeit in Amazon EKS umfasst Aktivitäten zur Infrastruktur-, Anwendungs- und Sicherheitsüberwachung.

Überwachung der Infrastruktur

Die Infrastrukturüberwachung ist eine grundlegende Komponente der Amazon EKS-Observability, die tiefe Einblicke in den Zustand und die Leistung der grundlegenden Elemente Ihres Kubernetes-Clusters bietet. Im Kern geht es darum, die Vitalparameter sowohl der Komponenten der Kontrollebene als auch der Worker-Knoten zu verfolgen und sicherzustellen, dass die zugrunde liegende Plattform stabil und effizient bleibt.

- Die Überwachung der Kontrollebene ist von entscheidender Bedeutung, da sie wichtige Komponenten wie den API-Server, die etcd-Datenbank und den Scheduler überwacht. Durch die Überwachung der API-Serverlatenz können Sie schnell Leistungsengpässe erkennen, die sich auf die Anwendungsbereitstellung oder Skalierung auswirken könnten. Die Etcd-Leistungsüberwachung überprüft, ob die Statusdatenbank des Clusters effizient funktioniert, und verhindert Datenkonsistenzprobleme, die sich auf den gesamten Cluster auswirken könnten.
- Die Überwachung auf Knotenebene ist ebenso wichtig, da sie sich auf die Rechenressourcen konzentriert, die Ihre containerisierten Workloads ausführen. Dazu gehört die Verfolgung der CPU-Auslastung, des Speicherverbrauchs, der Festplatten-I/O und der Netzwerkleistung über alle Worker-Knoten hinweg. Das Verständnis dieser Metriken hilft, eine Erschöpfung der Ressourcen zu verhindern, Entscheidungen zur Knotenskalierung zu optimieren und eine angemessene Kapazitätsplanung sicherzustellen.
- Die Netzwerküberwachung spielt eine entscheidende Rolle bei der Aufrechterhaltung einer zuverlässigen Kommunikation zwischen Pods, Diensten und externen Ressourcen. Durch die Überwachung von Netzwerkdurchsatz, Latenz und Verbindungsstatus können Sie Verbindungsprobleme frühzeitig erkennen und eine reibungslose Anwendungskommunikation sicherstellen. Die Speicherüberwachung ergänzt die Netzwerküberwachung, indem sie Volumenleistung, Kapazitätsauslastung und I/O Muster verfolgt, um datenbedingte Engpässe zu vermeiden.

Die Infrastrukturüberwachung dient als Frühwarnsystem für potenzielle Probleme, ermöglicht eine proaktive Wartung und gewährleistet eine optimale Ressourcenzuweisung. Ohne eine zuverlässige Infrastrukturüberwachung riskieren Sie unerwartete Ausfallzeiten, Leistungseinbußen und ineffiziente Ressourcennutzung, die sich erheblich auf den Geschäftsbetrieb und die Kosten auswirken können.

Anwendungsüberwachung

Die Anwendungsüberwachung ist für die Aufrechterhaltung gesunder, leistungsfähiger und zuverlässiger containerisierter Anwendungen in Ihrer Amazon EKS-Umgebung unerlässlich. Diese Überwachungsebene konzentriert sich auf die tatsächlichen Workloads, die in Ihrem Cluster ausgeführt werden, und bietet wichtige Einblicke in das Verhalten, die Leistung und die Interaktion Ihrer Anwendungen mit anderen Diensten.

Die Anwendungsüberwachung umfasst die Überwachung auf Container-Ebene, die Überwachung auf Service-Ebene und die verteilte Ablaufverfolgung.

- Auf Container-Ebene verfolgt die Anwendungsüberwachung wichtige Kennzahlen wie den Zustand des Containers, die Anzahl der Neustarts und die Muster des Ressourcenverbrauchs. Diese Metriken helfen Ihnen dabei, problematische Container zu identifizieren, die möglicherweise übermäßig viele Ressourcen verbrauchen oder häufig neu gestartet werden, was auf zugrunde liegende Probleme wie Speicherlecks oder Konfigurationsprobleme hinweisen könnte. Durch die Überwachung von Ereignissen im Container-Lebenszyklus können Sie sicherstellen, dass die Anwendung ordnungsgemäß funktioniert, und Bereitstellungsprobleme schnell beheben.
- Die Überwachung auf Service-Ebene bietet Einblick in Kennzahlen zur Leistung und Zuverlässigkeit von Anwendungen wie Reaktionszeiten, Fehlerraten und Anforderungsdurchsatz. Diese Kennzahlen sind für die Einhaltung der Service-Level-Ziele (SLOs) und die Sicherstellung eines positiven Endbenutzererlebnisses von entscheidender Bedeutung. Sie können die Latenz an verschiedenen Service-Endpunkten verfolgen, Leistungsengpässe identifizieren und Fehlermuster überwachen, um die Zuverlässigkeit von Anwendungen aufrechtzuerhalten.
- Die verteilte Ablaufverfolgung ist ein weiterer wichtiger Aspekt der Anwendungsüberwachung, insbesondere in Microservices-Architekturen. Durch die Implementierung der Ablaufverfolgung können Sie Anfragen verfolgen, während sie verschiedene Dienste durchlaufen, Abhängigkeiten verstehen und Leistungsengpässe identifizieren. Diese end-to-end Transparenz hilft Ihnen dabei, Serviceinteraktionen zu optimieren und komplexe Probleme zu beheben, die sich über mehrere Komponenten erstrecken.

Maßgeschneiderte Anwendungsmetriken spielen eine entscheidende Rolle bei der Bereitstellung geschäftsspezifischer Erkenntnisse. Dazu können Kennzahlen wie die Bearbeitungsrate von Bestellungen, die Häufigkeit der Benutzeranmeldungen oder die Erfolgsquote von Transaktionen gehören. Sie können diese benutzerdefinierten Metriken mit Infrastruktur- und Container-Metriken korrelieren, um besser zu verstehen, wie sich die Infrastrukturleistung auf den Geschäftsbetrieb auswirkt, und um datengestützte Entscheidungen zur Skalierung und Optimierung zu treffen.

Die Bedeutung der Anwendungsüberwachung liegt in ihrer Fähigkeit, einen umfassenden Überblick über den Zustand und die Leistung von Anwendungen zu bieten. Diese Überwachung ermöglicht es Ihnen, eine hohe Servicequalität aufrechtzuerhalten, Probleme schnell zu lösen und Ihre Anwendungen kontinuierlich zu optimieren, um Ihre Geschäftsziele zu erreichen.

Überwachung der Sicherheit

Die Sicherheitsüberwachung in Amazon EKS ist eine wichtige Aktivität, die Unternehmen dabei unterstützt, die Integrität, Vertraulichkeit und Konformität ihrer Kubernetes-Umgebungen aufrechtzuerhalten. Dieser umfassende Sicherheitsansatz kombiniert kontinuierliche Überwachung,

Bedrohungserkennung und Compliance-Überwachung, um containerisierte Workloads vor potenziellen Sicherheitsrisiken und unbefugtem Zugriff zu schützen. Er umfasst die Überwachung der Authentifizierung und Autorisierung, die Überwachung der Netzwerksicherheit sowie die Überwachung der Konfiguration und Einhaltung von Vorschriften.

- Die Authentifizierungs- und Autorisierungsüberwachung bildet die erste Verteidigungslinie, indem sie alle Versuche, auf den Cluster zuzugreifen, verfolgt. Dazu gehören die Überwachung von API-Serveranfragen, die Nachverfolgung erfolgreicher und fehlgeschlagener Anmeldeversuche und die Prüfung von Änderungen der rollenbasierten Zugriffskontrolle (RBAC). Durch die Führung detaillierter Auditprotokolle darüber, wer wann auf welche Ressourcen zugegriffen hat, können Sie potenzielle Sicherheitsverletzungen, unbefugte Zugriffsversuche oder Aktivitäten zur Eskalation von Rechten schnell erkennen. Dies ist besonders wichtig in Umgebungen mit mehreren Mandanten, in denen strenge Zugriffskontrollen unerlässlich sind.
- Die Überwachung der Netzwerksicherheit konzentriert sich auf die Erkennung und Verhinderung unbefugter Kommunikation zwischen Pods und Diensten. Durch die Überwachung von Verstößen gegen Netzwerkrichtlinien und ungewöhnlicher Datenverkehrsmuster können Sie potenzielle Sicherheitsbedrohungen wie Fluchtversuche in Containern oder seitliche Bewegungen innerhalb des Clusters identifizieren. Dazu gehört die Verfolgung sowohl der internen Cluster-Kommunikation als auch der externen Datenverkehrsmuster, um sicherzustellen, dass Container nur mit autorisierten Endpunkten kommunizieren und definierte Sicherheitsrichtlinien einhalten.
- Die Überwachung der Konfiguration und Einhaltung der Vorschriften ist für die Aufrechterhaltung der Sicherheitsstandards und die Einhaltung gesetzlicher Anforderungen unerlässlich. Dazu gehören das kontinuierliche Scannen von Container-Images auf Sicherheitslücken, die Überwachung der Laufzeitsicherheit und die Nachverfolgung von Konfigurationsänderungen, die sich auf den Sicherheitsstatus auswirken könnten. Regelmäßige Compliance-Audits stellen die Einhaltung von Industriestandards und organisatorischen Sicherheitsrichtlinien sicher, und die Erkennung von Konfigurationsabweichungen hilft dabei, unbefugte Änderungen zu verhindern, die Sicherheitsrisiken mit sich bringen könnten.

Die Sicherheitsüberwachung in Amazon EKS bietet die nötige Transparenz und Kontrolle, um sich vor modernen Sicherheitsbedrohungen zu schützen und gleichzeitig die Einhaltung gesetzlicher Anforderungen sicherzustellen. Durch die Implementierung einer umfassenden Sicherheitsüberwachung kann Ihr Unternehmen ein solides Sicherheitsniveau aufrechterhalten, schnell auf Sicherheitsvorfälle reagieren und die Einhaltung verschiedener regulatorischer Standards nachweisen.

Überwachungstools für Amazon EKS

In diesem Abschnitt werden drei Kategorien von Amazon EKS-Überwachungstools beschrieben: AWS Überwachungsdienste, Open-Source-Lösungen oder proprietäre Lösungen und spezielle Tools.

AWS Dienste

- Amazon CloudWatch: Umfassender Überwachungs- und Protokollierungsservice

  CloudWatch bildet das Rückgrat von AWS Überwachungslösungen und bietet umfangreiche Funktionen für Amazon EKS-Umgebungen. Es bietet Container Insights für detaillierte Container- und Cluster-Metriken, sodass Sie Leistung, Ressourcennutzung und Anwendungsintegrität überwachen können. Der Service zeichnet sich durch die Aggregation und Analyse von Protokollen aus und unterstützt die zentralisierte Protokollierung über Container und Knoten hinweg. CloudWatch integriert sich auf natürliche Weise in. AWS-Services Es bietet eine automatisierte Alarmkonfiguration und unterstützt benutzerdefinierte Metriken und Dashboards, was es zu einem unverzichtbaren Tool für die Amazon EKS-Überwachung macht.
- AWS X-Ray: Fortschrittliche verteilte Tracing-Plattform

  X-Ray verbessert die Beobachtbarkeit durch die Bereitstellung ausgeklügelter Funktionen zur verteilten Nachverfolgung. Die Visualisierung der Service Map bietet klare Einblicke in die Anwendungsarchitektur und die Abhängigkeiten, und die detaillierte Nachverfolgung von Anfragen hilft bei der Identifizierung von Leistungsengpässen zwischen Diensten. X-Ray kann Anfragen über komplexe Microservices-Architekturen verfolgen und ist daher für die Fehlerbehebung und Optimierung von unschätzbarem Wert, insbesondere in verteilten Systemen, die sich über mehrere Systeme erstrecken. AWS-Services
- AWS Distribution für: Einheitliches Observability-Framework OpenTelemetry

  Distro for OpenTelemetry bietet einheitliche Datenerfassungsfunktionen mit plattformübergreifender Unterstützung und ist daher ideal für Hybridumgebungen. Dieser Service lässt sich in andere integrieren AWS-Services, unterstützt kundenspezifische Instrumentierung und bietet Flexibilität bei der Implementierung umfassender Überwachungslösungen bei gleichzeitiger Wahrung der Kompatibilität mit Industriestandards.
- Amazon Managed Grafana: Visualisierung auf Unternehmensniveau

  Amazon Managed Grafana bietet einen vollständig verwalteten Service für Datenvisualisierung und -analyse. Es bietet eine nahtlose Integration mit anderen AWS-Services integrierten Sicherheitsfunktionen und Skalierbarkeit auf Unternehmensniveau. Der Service vereinfacht die

Erstellung und Verwaltung von Dashboards und bietet gleichzeitig erweiterte Funktionen wie den kontoübergreifenden Zugriff auf Datenquellen und die Integration mit. AWS IAM Identity Center

- Amazon Managed Service für Prometheus: Hochverfügbare, sichere, verwaltete Überwachung

  Amazon Managed Service for Prometheus ist ein vollständig verwalteter, Promethus-kompatibler Überwachungsservice. Er bietet automatische Skalierung, hohe Verfügbarkeit und sichere Erfassung und Abfrage von Metriken. Der Service lässt sich nahtlos in Amazon EKS integrieren und macht den betrieblichen Aufwand für die Verwaltung von Prometheus-Servern überflüssig.

Open-Source-Lösungen oder proprietäre Lösungen

Die im vorherigen Abschnitt beschriebenen AWS Tools bieten eine nahtlose Integration und verwaltete Dienste. Die in diesem Abschnitt aufgeführten Open-Source-Tools ergänzen das Angebot AWS-Services durch Flexibilität und umfangreiche Anpassungsmöglichkeiten. Wenn Sie die Funktionen und Anwendungsfälle der einzelnen Tools kennen, können Sie Überwachungsstrategien entwickeln, die Ihren spezifischen Anforderungen am besten entsprechen.

- Prometheus: Toolkit zur Erfassung von Metriken

  Prometheus ist eine Open-Source-Lösung für die Erfassung von Kennzahlen in Kubernetes-Umgebungen. Die Zeitreihen-Datenbank und die PromQL-Abfragesprache ermöglichen anspruchsvolle Metrikanalysen. Die Serviceerkennungsfunktionen der Plattform passen sich automatisch an dynamische Kubernetes-Umgebungen an, und das Alert-Management-System hält Sie über kritische Probleme auf dem Laufenden. Prometheus bietet umfangreiche Integrationsoptionen, die es zu einer vielseitigen Wahl für die umfassende Überwachung von Kennzahlen machen.

- Grafana: Fortschrittliche Visualisierungs-Engine

  Grafana wandelt mithilfe seiner Visualisierungsfunktionen komplexe Überwachungsdaten in umsetzbare Erkenntnisse um. Die Plattform erstellt maßgeschneiderte Dashboards, die Daten aus mehreren Quellen kombinieren und eine einheitliche Ansicht der Infrastruktur- und Anwendungsmetriken bieten. Die Unterstützung verschiedener Datenquellen und die Funktionen zur Verwaltung von Warnmeldungen ermöglichen eine umfassende Überwachung. Grafana kann Ihnen helfen, sowohl Echtzeit- als auch historische Daten zu visualisieren, sodass Sie Trends erkennen und fundierte Entscheidungen treffen können.

- Fluent Bit: Einheitliche Protokollierungsebene

Diese Protokollierungslösung ermöglicht die Erfassung und Verwaltung von Protokollen für Kubernetes-Umgebungen. Die native Kubernetes-Integration gewährleistet eine nahtlose Erfassung von Protokollen aus Containern und Knoten, und die Unterstützung mehrerer Ausgabeziele bietet Flexibilität bei der Speicherung und Analyse von Protokollen. Erweiterte Funktionen wie Log-Parsing und Filterung ermöglichen es Ihnen, Logs auf der Grundlage spezifischer Anforderungen zu verarbeiten und weiterzuleiten. Aufgrund seines geringen Gewichts eignet sich Fluent Bit besonders für containerisierte Umgebungen.

- Datadog: Beobachtbarkeit im gesamten Stack

  Datadog bietet umfassende Überwachungsfunktionen mit nativer Kubernetes-Unterstützung. Es bietet Infrastrukturüberwachung, Anwendungsleistungsüberwachung (APM), Protokollverwaltung und Echtzeitanalysen. Sie können die automatische Serviceerkennung und den umfangreichen Integrationskatalog der Plattform für die Amazon EKS-Überwachung sowie die Funktionen für maschinelles Lernen nutzen, um Anomalien zu erkennen und potenzielle Probleme vorherzusagen.

- New Relic: Überwachung der Anwendungsleistung

  New Relic bietet Einblick in die Anwendungsleistung und den Zustand der Infrastruktur. Die Kubernetes-Integration bietet detaillierte Einblicke in Container, verteiltes Tracing und benutzerdefinierte Dashboards. Die Plattform hilft Ihnen dabei, die Anwendungsleistung mit den Infrastrukturkennzahlen zu korrelieren, sodass Sie Probleme schnell identifizieren und lösen können.

- Elastic Stack (ELK Stack): Protokollanalyse und Suche

  Der ELK Stack kombiniert Elasticsearch, Logstash und Kibana, um Funktionen zur Protokollverwaltung und -analyse bereitzustellen. Er bietet erweiterte Suchfunktionen, Visualisierungstools und Funktionen für maschinelles Lernen. Sie können den Stack verwenden, um große Mengen an Protokolldaten aus Ihren Amazon EKS-Umgebungen zu verarbeiten.

Spezialisierte Tools

Sie können die folgenden Tools je nach Ihren spezifischen Überwachungsanforderungen, Ihrem Betriebsumfang und Ihren Unternehmenspräferenzen kombinieren. Der Schlüssel liegt darin, einen Monitoring-Stack zu erstellen, der umfassende Transparenz bietet und gleichzeitig überschaubar und kostengünstig bleibt.

- kube-state-metrics (KSM): Überwachung des Kubernetes-Zustands

Dieser Zusatzdienst überwacht den Kubernetes-API-Server und generiert Metriken über den Status von Objekten. Er bietet Einblicke in den Zustand von Bereitstellungen, Pods und anderen Kubernetes-Ressourcen.

- Kubernetes Metrics Server: Ressourcen-Metriken

  Dieser Metrikserver sammelt Ressourcenmetriken von Kubelets und stellt sie über die Kubernetes-Metrik-API zur Verfügung. Er bietet horizontale automatische Pod-Skalierung und grundlegende CPU- und Speichermetriken.

- Kubecost: Kostenüberwachung für Kubernetes

  Tools wie Kubecost bieten detaillierte Kostenanalysen und Optimierungsempfehlungen für EKS-Cluster. Sie helfen Ihnen dabei, die Cloud-Ausgaben für verschiedene Namespaces, Bereitstellungen und Dienste zu verstehen und zu optimieren.

Implementierung von Hochverfügbarkeit für Amazon EKS-Überwachungslösungen

Eine robuste Hochverfügbarkeitsstrategie (HA) für die Amazon EKS-Überwachung ist entscheidend, um einen kontinuierlichen Einblick in Ihre Kubernetes-Umgebung zu gewährleisten. In diesem Abschnitt wird ein umfassender Ansatz zur Implementierung von HA in verschiedenen Aspekten Ihrer Überwachungsinfrastruktur erörtert.

Architektonische Redundanz und Skalierbarkeit

Der Aufbau eines hochverfügbaren Überwachungssystems beginnt mit der richtigen architektonischen Gestaltung. Die Überwachungskomponenten sollten zum Schutz vor Zonenausfällen auf mehrere AWS Availability Zones verteilt werden. Dazu gehört die Implementierung einer horizontalen Skalierung für kritische Überwachungskomponenten wie Prometheus-Server, Log-Collectors und Alert Manager. Sie können AWS Managed Services wie Amazon Managed Service for Prometheus und Amazon Managed Grafana verwenden, um den Betriebsaufwand zu reduzieren und gleichzeitig eine hohe Verfügbarkeit sicherzustellen. Konfigurieren Sie automatische Failover-Mechanismen, um die Servicekontinuität bei Komponentenausfällen aufrechtzuerhalten. Dazu gehören Integritätsprüfungen und automatische Wiederherstellungsverfahren.

Zuverlässige Datenspeicherstrategie

Die Widerstandsfähigkeit der Datenspeicherung ist für die Aufrechterhaltung der Zuverlässigkeit des Überwachungssystems von grundlegender Bedeutung. Durch die Implementierung verteilter Speicherlösungen wird sichergestellt, dass metrische Daten und Protokolle auch dann zugänglich bleiben, wenn einzelne Speicherknoten ausfallen. Dazu gehören die Konfiguration einer ordnungsgemäßen Datenreplikation in mehreren Availability Zones und die Verwendung verschiedener Speicher-Backends für Redundanz. Richten Sie regelmäßige Backup-Verfahren für historische Daten mit dokumentierten Wiederherstellungsprozessen für verschiedene Ausfallszenarien ein. Bei Zeitreihendatenbanken wie Prometheus hilft die Implementierung von Remotespeicherlösungen dabei, Speicherprobleme von der Datenerfassung zu trennen und die allgemeine Systemzuverlässigkeit zu verbessern.

Redundantes Alarmmanagement

Das Alert-Management erfordert in einem HA-Setup besondere Aufmerksamkeit. Durch den Einsatz redundanter Warnmanager wird sichergestellt, dass kritische Benachrichtigungen auch bei Systemausfällen die vorgesehenen Empfänger erreichen. Konfigurieren Sie mehrere Benachrichtigungskanäle wie E-Mail, SMS, Slack und stellen PagerDuty Sie alternative Kommunikationswege bereit. Verwenden Sie Mechanismen zur Deduplizierung von Alarmen, um Alert-Storms bei teilweisen Systemausfällen zu verhindern, und alternative Benachrichtigungsmethoden, um sicherzustellen, dass kritische Warnmeldungen nie übersehen werden. Die Implementierung der Korrelation von Warnmeldungen trägt dazu bei, den Kontext während Failover-Szenarien aufrechtzuerhalten, und verhindert doppelte Benachrichtigungen von redundanten Systemen.

Lastenausgleich und Serviceerkennung

Ein ordnungsgemäßer Lastenausgleich ist für die Aufrechterhaltung stabiler Überwachungsdienste unerlässlich. AWS Application Load Balancer verteilen den eingehenden Monitoring-Verkehr auf mehrere Endpunkte, und Integritätsprüfungen stellen sicher, dass der Datenverkehr nur an fehlerfreie Instances weitergeleitet wird. Mithilfe von Diensterkennungsmechanismen können sich Überwachungskomponenten automatisch an Änderungen in der Umgebung anpassen, z. B. das Hinzufügen neuer Knoten oder Dienste. Stellen Sie Überwachungsagenten konsistent auf allen Knoten DaemonSets bereit, indem Sie bei der Skalierung des Clusters eine umfassende Abdeckung sicherstellen.

Zusätzliche Überlegungen zur Hochverfügbarkeit

Ausfallsicherheit des Netzwerks:

- Implementieren Sie redundante Netzwerkpfade.
- Konfigurieren Sie das richtige Subnetzdesign für alle Availability Zones.
- Verwenden Sie es AWS Direct Connectmit Backup-Routen.
- Konfigurieren Sie die entsprechenden Sicherheitsgruppen und Netzwerkzugriffskontrolllisten (Netzwerk ACLs).

Überwachung der Monitore:

- Setzen Sie sekundäre Überwachungssysteme ein.
- Implementieren Sie eine regionsübergreifende Überwachung.
- Konfigurieren Sie Warnmeldungen für Systeme, die nicht reagieren.
- Testen Sie regelmäßig die Failover-Verfahren.

Kapazitätsplanung:

- Überwachen Sie Trends bei der Ressourcennutzung.
- Implementieren Sie vorausschauende Skalierung.
- Testen Sie die Leistung regelmäßig.

Datenmanagement:

- Implementieren Sie Richtlinien zur Datenspeicherung.
- Konfigurieren Sie die Metrikaggregation.
- Planen Sie das Datenlebenszyklusmanagement.
- Optimieren Sie den Speicher regelmäßig.

Wiederherstellungsverfahren:

- Prozesse zur Wiederherstellung von Dokumenten.
- Testen Sie die Notfallwiederherstellung regelmäßig.

- Implementieren Sie nach Möglichkeit eine automatisierte Wiederherstellung.
- Identifizieren und implementieren Sie klare Eskalationspfade.

Durch die Implementierung dieser Hochverfügbarkeitspraktiken können Sie sicherstellen, dass Ihre Amazon EKS-Überwachungsinfrastruktur zuverlässig und belastbar bleibt und dass Sie auch bei verschiedenen Ausfallszenarien einen kontinuierlichen Einblick in Ihre Kubernetes-Umgebungen haben. Regelmäßige Tests und Aktualisierungen dieser HA-Konfigurationen stellen sicher, dass sie auch bei der Weiterentwicklung der Umgebung wirksam bleiben.

Bewährte Methoden für die Überwachung in Amazon EKS

Strategischer Implementierungsansatz

Eine erfolgreiche Amazon EKS-Überwachungsstrategie beginnt mit einem gut geplanten, schrittweisen Implementierungsansatz.

- Beginnen Sie mit der Identifizierung und Überwachung kritischer Kennzahlen, die sich direkt auf Ihren Geschäftsbetrieb und die Zuverlässigkeit Ihrer Anwendungen auswirken. Diese Grundlage sollte wichtige Infrastrukturkennzahlen, wichtige Leistungsindikatoren für Anwendungen und kritische Sicherheitsmetriken umfassen. Erweitern Sie den Umfang der Überwachung schrittweise auf der Grundlage der betrieblichen Anforderungen und der gewonnenen Erkenntnisse und stellen Sie sicher, dass jede Ergänzung einen sinnvollen Nutzen bietet.
- Implementieren Sie automatisierte Bereitstellungsprozesse mithilfe von Infrastructure-as-Code-Tools (IaC) wie Terraform oder CloudFormation um Konsistenz und Wiederholbarkeit sicherzustellen.
- Testen und validieren Sie Überwachungssysteme, um Zuverlässigkeit und Genauigkeit zu gewährleisten.
- Verfeinern Sie die Überwachungsparameter kontinuierlich, um sie an die sich ändernden Geschäftsanforderungen anzupassen.

Effektives Datenmanagement

Ein ordnungsgemäßes Datenmanagement ist entscheidend für die Aufrechterhaltung einer effizienten und kostengünstigen Überwachungslösung.

- Implementieren Sie klare Richtlinien zur Datenspeicherung, die den Bedarf an historischen Analysen mit den Speicherkosten in Einklang bringen.
- Konfigurieren Sie geeignete Abtastraten für verschiedene Metriktypen: höhere Frequenz für kritische Messwerte und niedrigere Frequenz für weniger kritische Messwerte.
- Verwenden Sie die Metrikaggregation, um das Datenvolumen zu reduzieren und gleichzeitig aussagekräftige Erkenntnisse zu gewinnen, insbesondere für langfristige Trendanalysen.
- Implementieren Sie systematische Verfahren zur Aufbewahrung und Archivierung von Protokollen für zentralisierte Protokollierungssysteme (z. B. CloudWatch Logs), um die Speicherkosten zu kontrollieren und sicherzustellen, dass der Zugriff auf wichtige Daten gewährleistet bleibt.

> **Note**
>
> Die Protokollrotation auf Containerebene wird in Amazon EKS Version 1.21 oder höher automatisch vom Kubelet abgewickelt.

- Erwägen Sie die Implementierung einer hot-warm-cold Architektur für die Protokollspeicherung, um sowohl die Zugriffsgeschwindigkeit als auch die Kosteneffizienz zu optimieren.

Konfiguration und Verwaltung von Warnmeldungen

Die Konfiguration von Warnmeldungen erfordert sorgfältige Überlegungen, um die Effektivität aufrechtzuerhalten, ohne dass es zu einer Ermüdung der Warnmeldungen kommt.

- Definieren Sie klare, umsetzbare Schwellenwerte auf der Grundlage von Service-Level-Zielen (SLOs) und historischen Leistungsmustern.
- Implementieren Sie ein System mit abgestuftem Schweregrad für Warnmeldungen, das klar zwischen kritischen Problemen, die sofortige Aufmerksamkeit erfordern, und weniger dringenden Problemen unterscheidet.
- Stellen Sie sicher, dass Warnmeldungen ausreichend Kontext und umsetzbare Informationen enthalten, um eine schnelle Problemlösung zu ermöglichen.
- Richten Sie klare Eskalationsverfahren mit definierten Zuständigkeitsbereichen und Reaktionszeiten für unterschiedliche Schweregrade von Alarmen ein.
- Überprüfen und verfeinern Sie die Warnkonfigurationen regelmäßig, um deren Relevanz und Effektivität zu gewährleisten.

Optimierung der Ressourcen

Die kontinuierliche Überwachung der Ressourcennutzung ist für die Aufrechterhaltung eines kostengünstigen Betriebs unerlässlich.

- Implementieren Sie eine umfassende Ressourcenüberwachung für alle Cluster-Komponenten, einschließlich Knoten, Pods und persistenter Volumes.
- Konfigurieren Sie die automatische Skalierung auf der Grundlage der tatsächlichen Nutzungsmuster und Leistungsanforderungen, um eine effiziente Ressourcennutzung bei gleichbleibender Leistung sicherzustellen.
- Verwenden Sie Tags zur Kostenzuweisung, um den Ressourcenverbrauch durch verschiedene Teams, Anwendungen oder Umgebungen zu verfolgen.
- Analysieren Sie regelmäßig Kennzahlen zur Ressourceneffizienz, um Optimierungsmöglichkeiten zu identifizieren und Verbesserungen umzusetzen.
- Erwägen Sie die Implementierung von Kostenmanagement-Tools, um Cloud-Ausgaben zu verfolgen und zu optimieren.

Sicherheit

Sicherheitsüberlegungen sollten integraler Bestandteil Ihrer Überwachungsstrategie sein.

- Implementieren Sie die Prinzipien des Zugriffs mit den geringsten Rechten für alle Überwachungskomponenten, um sicherzustellen, dass Benutzer und Dienste nur über die Berechtigungen verfügen, die sie benötigen.
- Ermöglichen Sie eine umfassende Auditprotokollierung, um alle Zugriffe und Änderungen an Überwachungssystemen nachzuverfolgen.
- Führen Sie regelmäßige Sicherheitsüberprüfungen der Überwachungskonfigurationen und Zugriffsmuster durch, um potenzielle Sicherheitslücken zu identifizieren.
- Implementieren Sie Verschlüsselung für sensible Überwachungsdaten sowohl bei der Übertragung als auch bei der Speicherung.
- Integrieren Sie die Sicherheitsüberwachung in bestehende SIEM-Systeme (Security Information and Event Management), um eine umfassende Sicherheitstransparenz zu gewährleisten.

Überlegungen zur erweiterten Überwachung in Amazon EKS

Leistungsoptimierung:

- Optimieren Sie die Intervalle für die Erfassung von Kennzahlen.
- Konfigurieren Sie effiziente Abfragemuster.
- Implementieren Sie die Voraggregation von Metriken.
- Verwenden Sie geeignete Speicherlösungen.

Einhaltung von Vorschriften und Unternehmensführung:

- Pflegen Sie Prüfpfade.
- Implementieren Sie die Compliance-Überwachung.
- Sorgen Sie für regelmäßige Compliance-Berichte.
- Verfahren zur Dokumentenüberwachung.

Wiederherstellung nach einem Notfall:

- Erstellen Sie regelmäßig Backups der Überwachungskonfigurationen.
- Verfahren zur Wiederherstellung von Dokumenten.
- Testen Sie die Wiederherstellungsprozesse.

Kontinuierliche Verbesserung:

- Überwachen Sie die Überprüfungssitzungen regelmäßig.
- Optimieren Sie die Leistungszyklen.
- Aktualisieren Sie die Überwachung auf der Grundlage von Vorfällen.
- Integrieren Sie das Feedback der Benutzer.

Diese Best Practices bieten einen Rahmen für die Implementierung und Wartung effektiver Überwachungslösungen für Amazon EKS-Umgebungen. Überprüfen und aktualisieren Sie diese Praktiken regelmäßig, damit sie Ihren organisatorischen Anforderungen und Industriestandards entsprechen. Bei der Überwachung handelt es sich nicht um eine einmalige Einrichtung, sondern um einen kontinuierlichen Prozess, der regelmäßige Aufmerksamkeit und Verbesserung erfordert.

Ablaufverfolgung in Amazon EKS

Die Ablaufverfolgung ist eine wichtige Komponente der Anwendungsbeobachtbarkeit in Amazon EKS. Die Ablaufverfolgung bietet einen detaillierten Einblick in Anforderungsabläufe und Serviceinteraktionen, indem sie den Pfad der Anfragen sammelt, verarbeitet und visualisiert, während sie verschiedene Microservices durchlaufen, die auf EKS-Clustern bereitgestellt werden. Diese Funktion hilft Ihnen, das Systemverhalten zu verstehen, Engpässe zu identifizieren und Probleme in Ihrer Amazon EKS-Umgebung effektiv zu beheben. Effektives Tracing macht das Debuggen verteilter Systeme weniger komplex, indem es end-to-end Einblicke in die Anforderungsabläufe bietet. Es ermöglicht die Nachverfolgung von Transaktionen über Servicegrenzen hinweg und die Identifizierung von Leistungsproblemen oder Ausfällen innerhalb von Amazon EKS-Workloads.

Die allgemeine Ablaufverfolgungsimplementierung in Amazon EKS ermöglicht es Ihnen, das Systemverhalten zu verstehen, die Leistung zu optimieren und die Zuverlässigkeit Ihrer containerisierten Anwendungen aufrechtzuerhalten. Letztlich verbessern die Funktionen der Rückverfolgung die betriebliche Transparenz und die Systemwartbarkeit in Amazon EKS-Umgebungen.

AWS X-Ray spielt eine wichtige Rolle bei der Rückverfolgung von Daten über Ihre Anwendung. Bei der Rückverfolgung werden verschiedene Aspekte der Dienstinteraktionen überwacht, darunter die folgenden:

- Anforderungspfade und Abhängigkeiten bieten wichtige Einblicke in das Verhalten Ihres verteilten Systems. Sie verfolgen den gesamten Weg von Anfragen, während sie verschiedene Microservices und Komponenten durchlaufen. Die Abbildung von Serviceabhängigkeiten hilft Ihnen dabei, Kommunikationsmuster zu verstehen und kritische Pfade in Ihrer Anwendungsarchitektur zu identifizieren. Einzelheiten zur Implementierung finden Sie unter Verwenden der AWS X-Ray Service-Trace-Map in der X-Ray-Dokumentation.
- Servicelatenzen und Engpässe sind wichtige Messgrößen für die Aufrechterhaltung einer optimalen Systemleistung. Durch die Messung und Analyse der Reaktionszeiten zwischen Diensten können Sie Leistungsprobleme effektiv identifizieren. Diese Daten ermöglichen es Ihnen, bestimmte Dienste oder Vorgänge zu ermitteln, die zu Verzögerungen in der Anforderungskette führen, und gezielte Optimierungsmaßnahmen zu ermöglichen. Weitere Informationen zur Latenzanalyse finden Sie unter Interaktion mit der Analytics-Konsole in der X-Ray-Dokumentation.
- Muster der Fehlerausbreitung helfen Ihnen dabei, die Zuverlässigkeit und Fehlertoleranz des Systems besser zu verstehen. Wenn Sie verstehen, wie Fehler im System kaskadieren, indem Sie Fehlerpfade zwischen Diensten verfolgen, können Sie Ihre Anwendungen besser konzipieren.

Diese Transparenz hilft Ihnen dabei, die Hauptursache von Fehlern und deren Auswirkungen auf abhängige Dienste zu identifizieren, was zu robusteren Systemen führt. Einzelheiten zur Implementierung finden Sie unter Traces in der X-Ray-Dokumentation.

- Die dienstübergreifende Ressourcennutzung bietet Einblicke in die Systemeffizienz und Kostenoptimierung. Sie können die CPU-, Arbeitsspeicher- und Netzwerkauslastungsmuster überwachen, die mit Trace-Daten korrelieren, um den Ressourcenbedarf zu verstehen. Diese Daten helfen Ihnen bei der Analyse von Trends beim Ressourcenverbrauch, um die Serviceleistung und die Kosten in Ihrem gesamten EKS-Cluster zu optimieren. Informationen zur Einrichtung der Überwachung finden Sie unter Überwachen der Cluster-Leistung und Anzeigen von Protokollen in der Amazon EKS-Dokumentation.
- Transaktionsabläufe für Endbenutzer sind entscheidend für das Verständnis und die Verbesserung der Benutzererfahrung. Durch die Verfolgung vollständiger Benutzerinteraktionen von Frontend- bis Backend-Services können Sie eine optimale Anwendungsleistung sicherstellen. Sie können die end-to-end Reaktionszeiten für kritische Benutzererfahrungen messen und optimieren, was sich direkt auf die Kundenzufriedenheit auswirkt. Verwenden Sie das AWS X-Ray SDK für Ihre Programmiersprache, um die Endbenutzerüberwachung zu implementieren.
- API-Gateway-Interaktionen stehen im Mittelpunkt der Leistung und Sicherheit Ihrer Anwendung. Sie können Anforderungsmuster und Leistung an API-Zugangspunkten überwachen, um eine optimale Servicebereitstellung sicherzustellen. Diese Transparenz hilft Ihnen dabei, die Auswirkungen von Authentifizierung, Autorisierung und Ratenbegrenzung auf den Anfragefluss nachzuverfolgen, um sowohl die Sicherheits- als auch die Leistungsanforderungen zu erfüllen. Erfahren Sie mehr über API-Tracing in der Dokumentation Amazon API Gateway with X-Ray.

Effektives Tracing in Amazon EKS geht über das Sammeln von Spans und Traces hinaus. Es erfordert eine gut strukturierte Strategie, die die Anforderungen an die Beobachtbarkeit mit der Systemleistung in Einklang bringt. Diese Strategie sollte sich auf Folgendes konzentrieren:

- Implementierung geeigneter Stichprobenraten: Konfigurieren Sie Stichprobenregeln auf der Grundlage von Verkehrsmustern und Geschäftsprioritäten, um die Kosten zu optimieren und gleichzeitig die Sichtbarkeit kritischer Transaktionen zu wahren. Weitere Informationen finden Sie unter Konfiguration von Probenahmeregeln in der X-Ray-Dokumentation.
- Definition kritischer Pfade und Dienste, die nachverfolgt werden sollen: Identifizieren und priorisieren Sie wichtige Dienste und Benutzerabläufe, für die eine detaillierte Nachverfolgung erforderlich ist, um eine optimale Leistungsüberwachung zu gewährleisten. Weitere Informationen finden Sie unter Senden von Metrik- und Trace-Daten mit ADOT Operator in der Amazon EKS-Dokumentation.

- Festlegung geeigneter Richtlinien zur Datenspeicherung: Richten Sie Regeln für das Datenlebenszyklusmanagement ein, um die Anforderungen an die Beobachtbarkeit mit den Speicherkosten und den Compliance-Anforderungen in Einklang zu bringen. Informationen zur Anzeige von CloudWatch Aufbewahrungsrichtlinien finden Sie in der Dokumentation zu CloudWatch Protokollen unter Arbeiten mit Protokollgruppen und Protokollströmen.
- Einrichtung effektiver Visualisierungs- und Analysetools: Stellen Sie Visualisierungstools wie die AWS X-Ray Analytics-Konsole oder Amazon Managed Grafana bereit und konfigurieren Sie sie, um Trace-Daten effektiv zu analysieren. Weitere Informationen finden Sie unter Interaktion mit der Analytics-Konsole in der X-Ray-Dokumentation.

In diesem Abschnitt:

Tools zur Ablaufverfolgung für Amazon EKS

Amazon EKS unterstützt mehrere Optionen AWS und Optionen von Drittanbietern für die Implementierung von verteilter Ablaufverfolgung.

AWS-Services

- AWS X-Ray: Fortschrittliche Plattform für verteilte Ablaufverfolgung

  X-Ray ist ein vollständig verwaltetes Programm AWS-Service , das end-to-end Tracing-Funktionen bietet. Es instrumentiert AWS-Services und bietet automatisch detaillierte Servicemaps und Analysen für Ihre Anwendungen, die auf Amazon EKS ausgeführt werden. X-Ray ist in andere Systeme AWS-Services, einschließlich Amazon CloudWatch, integriert und ermöglicht die automatische Korrelation von Traces mit AWS-Service Anrufen.
- AWS Distribution für OpenTelemetry: Einheitliches Observability-Framework

  Distro for OpenTelemetry ist eine sichere, produktionsbereite und AWS unterstützte Distribution von für Cloud-native Anwendungen. OpenTelemetry Es bietet herstellerneutrale Instrumentierungsfunktionen und behält gleichzeitig die native AWS-Service Integration bei, was es ideal für Hybrid-Cloud-Umgebungen macht. Distro for OpenTelemetry unterstützt mehrere Observability-Backends und bietet eine nahtlose Integration mit Überwachungsdiensten. AWS

Open-Source-Lösungen

- OpenTelemetry: Open-Source-Framework für Beobachtbarkeit

  OpenTelemetry bietet ein standardisiertes Observability-Framework mit umfassenden Instrumentierungsbibliotheken, die mehrere Programmiersprachen unterstützen. Aufgrund seiner flexiblen Backend-Optionen und seines herstellerneutralen Ansatzes eignet es sich ideal für Workloads, die Konsistenz in verschiedenen Umgebungen erfordern. Das umfangreiche Ökosystem des Frameworks gewährleistet eine umfassende Kompatibilität mit verschiedenen Monitoring-Lösungen.

- Jaeger: Verteilte Open-Source-Tracing-Plattform

  Jaeger bietet umfassende Tracing-Funktionen mit verteilter Kontextverbreitung in Echtzeit. Es bietet eine Ursachenanalyse und Leistungsoptimierung durch eine detaillierte Visualisierung der Serviceabhängigkeiten. Die Architektur von Jaeger ist auf hohe Skalierbarkeit ausgelegt und unterstützt verschiedene Speicher-Backends, sodass sie sich für groß angelegte Amazon EKS-Bereitstellungen eignet. Sehen Sie sich die Einrichtung von Jaeger für EKS an

- Grafana Tempo: Verteilte Rückverfolgung

  Tempo ist eine Lösung von Grafana Labs, die eine umfangreiche Speicherung von Traces und eine nahtlose Integration mit Prometheus-Metriken bietet. Aufgrund seines kostengünstigen Trace-Retention-Modells und der nativen Integration mit Grafana eignet es sich für Unternehmen, die Grafana bereits für die Visualisierung verwenden. Die Architektur von Tempo wurde speziell für Cloud-native Umgebungen wie Amazon EKS entwickelt.

Bewährte Methoden für die Ablaufverfolgung in Amazon EKS

Dieser Abschnitt enthält eine umfassende Liste mit bewährten Methoden und Techniken für die Erstellung eines effektiven Ablaufverfolgungssystems, das die Beobachtbarkeit und Fehlerbehebung für Ihre Kubernetes-basierten Anwendungen in Amazon EKS verbessert.

- Strategisches Sampling: Konfigurieren Sie verschiedene Sampling-Raten, die auf den Datenverkehrsmustern Ihrer Anwendung und der Bedeutung der von Ihnen verwendeten Dienste basieren. Implementieren Sie höhere Abtastraten für kritische Pfade und reduzieren Sie gleichzeitig die Anzahl der Probenahmen für weniger kritische Routen mit hohem Volumen, um die Kosten zu optimieren. Eine Anleitung finden Sie in der AWS X-Ray Dokumentation unter Konfiguration von Probenahmeregeln.

- Einrichtung der Instrumentierung: Verwenden Sie automatische Instrumentierungstools wie das X-Ray SDK oder AWS Distro for OpenTelemetry Collectors, um den manuellen Instrumentierungsaufwand zu minimieren. Sorgen Sie für eine bessere Trace-Korrelation bei gleichbleibenden Benennungskonventionen und kontextübergreifender Verteilung zwischen Diensten. Weitere Informationen finden Sie in der Dokumentation zu Distro for OpenTelemetry Collector.
- Datenmanagement: Implementieren Sie geeignete Aufbewahrungsfristen und Komprimierungsstrategien, um die Speicherkosten mit Ihren Observability-Anforderungen in Einklang zu bringen. Richten Sie klare Datenschutzkontrollen und Backup-Verfahren ein, um sensible Trace-Daten zu schützen. Weitere Informationen finden Sie in der Protokolldokumentation unter CloudWatch Protokolle unter Aufbewahrung von CloudWatch Protokolldaten ändern.
- Leistungsoptimierung: Überwachen und optimieren Sie den Tracing-Overhead, um die Auswirkungen auf die Anwendungsleistung zu minimieren. Verwenden Sie effiziente Pufferung und asynchrone Verarbeitung, um die Auswirkungen auf die Latenz zu reduzieren. Weitere Informationen finden Sie in der X-Ray-Dokumentation unter Konfiguration des AWS X-Ray Daemons.
- Sicherheitskontrollen: Implementieren Sie mithilfe von IAM-Rollen und -Richtlinien angemessene Zugriffskontrollen und Datenschutzmaßnahmen. Regelmäßige Sicherheitsaudits und Compliance-Überprüfungen tragen dazu bei, dass die Trace-Daten sicher bleiben. Weitere Informationen finden Sie unter Sicherheit AWS X-Ray in der X-Ray-Dokumentation.
- Überwachung und Warnmeldungen: Richten Sie eine umfassende Überwachung des Zustands der Trace-Erfassung ein und konfigurieren Sie Warnmeldungen für Probleme mit der Erfassung. Verfolgen Sie die Abtastraten und Kennzahlen zur Systemleistung, um einen optimalen Betrieb sicherzustellen. Weitere Informationen finden Sie in der CloudWatch Dokumentation unter Container Insights.
- Hohe Verfügbarkeit: Stellen Sie redundante Collectors in allen Availability Zones bereit und konfigurieren Sie die richtigen Failover-Mechanismen. Regelmäßige Tests des Hochverfügbarkeits-Setups gewährleisten eine zuverlässige Trace-Erfassung. Weitere Informationen finden Sie unter Using AWS Distro for OpenTelemetry as a Collector in der Dokumentation zu Amazon Managed Service for Prometheus.

Wenn Sie diese bewährten Methoden befolgen, können Sie ein robustes, effizientes und effektives Ablaufverfolgungssystem für Ihre Amazon EKS-Umgebung erstellen. Dies trägt dazu bei, eine umfassende Beobachtbarkeit, eine effiziente Fehlerbehebung und eine optimale Leistung Ihrer Kubernetes-basierten Anwendungen sicherzustellen.

Warnmeldungen in Amazon EKS

Warnmeldungen sind eine wichtige Komponente bei der Verwaltung und Wartung von Anwendungen, die auf Amazon EKS ausgeführt werden. Es dient als Frühwarnsystem, das Betreiber und Entwickler über potenzielle Probleme, Anomalien oder Leistungseinbußen informiert, bevor diese zu schwerwiegenden Problemen eskalieren, die die Serviceverfügbarkeit oder das Benutzererlebnis beeinträchtigen könnten. Bei der Alarmierung werden verschiedene Aspekte des Kubernetes-Clusters überwacht, darunter:

- Zustand der Infrastruktur
- Leistung der Anwendung
- Containermetriken
- Benutzerdefinierte Geschäftskennzahlen

Effektive Benachrichtigungen in Amazon EKS gehen über das einfache Einrichten von Benachrichtigungen hinaus. Es erfordert eine well-thought-out Strategie, die den Bedarf an zeitnahen Informationen mit dem Risiko einer Übermüdung der Warnmeldungen in Einklang bringt. Diese Strategie sollte:

- Definieren Sie aussagekräftige Schwellenwerte und Bedingungen.
- Priorisieren Sie Warnmeldungen nach Schweregrad und Auswirkung.
- Implementieren Sie die richtigen Weiterleitungs- und Eskalationsverfahren.
- Integrieren Sie es in Tools für das Vorfallmanagement und die Kommunikation.

In diesem Abschnitt:

Warntools für Amazon EKS

Amazon EKS unterstützt mehrere Optionen AWS und Optionen von Drittanbietern für die Implementierung von Warnmeldungen. Wenn Sie sich für ein Tool für Amazon EKS-Benachrichtigungen entscheiden, sollten Sie Faktoren wie Integrationsfähigkeit, Skalierbarkeit,

Benutzerfreundlichkeit, Kosten und spezifische Funktionen berücksichtigen, die Ihren Überwachungs- und Warnungsanforderungen entsprechen. Viele Unternehmen verwenden eine Kombination dieser Tools, um eine umfassende Überwachungs- und Warnlösung für ihre Amazon EKS-Umgebungen zu erstellen.

- Amazon CloudWatch: AWS-Service zur Überwachung und Beobachtbarkeit

  CloudWatch bietet Metriken, Protokolle und Alarme für EKS-Cluster und lässt sich gut in andere AWS-Services integrieren.
- Prometheus: Open-Source-Überwachungs- und Warnungstool für Kubernetes

  Prometheus bietet eine leistungsstarke Abfragesprache (PromQL) zur Definition von Alarmbedingungen.
- Alertmanager: Ergänzung zu Prometheus für den Umgang mit Alarmen

  Alertmanager ermöglicht die Deduplizierung, Gruppierung und Weiterleitung von Warnmeldungen. Es unterstützt verschiedene Benachrichtigungskanäle, darunter E-Mail, Slack und. PagerDuty
- Grafana: Open-Source-Plattform für Überwachung und Beobachtbarkeit

  Grafana bietet Visualisierungs- und Alarmfunktionen. Es kann in verschiedene Datenquellen integriert werden, darunter Prometheus und. CloudWatch
- Elastic Stack (ELK Stack): Kombination aus Elasticsearch, Logstash und Kibana

  Dieses Tool ist nützlich für die Aggregation, Analyse und Alarmierung von Protokollen. Es kann mit den Observability-Funktionen von Elastic erweitert werden.
- Lösungen von Drittanbietern

  Auf dem Markt sind viele Tools erhältlich, darunter Datadog, New Relic, Sysdig, Dynatrace, Zabbix, Nagios, Splunk, IBM Instana und. AppDynamics

Bewährte Methoden für Warnmeldungen in Amazon EKS

In diesem Abschnitt werden die bewährten Methoden für die Erstellung eines robusten Warnsystems beschrieben, das die Zuverlässigkeit und Leistung Ihrer Kubernetes-basierten Anwendungen in Amazon EKS verbessert.

Definieren Sie klare Schwellenwerte für Warnmeldungen:

- Legen Sie aussagekräftige Schwellenwerte auf der Grundlage historischer Daten und Geschäftsanforderungen fest.
- Verwenden Sie gegebenenfalls dynamische Schwellenwerte, um unterschiedlichen Workloads Rechnung zu tragen.

Implementieren Sie die Priorisierung von Warnmeldungen:

- Kategorisieren Sie Warnmeldungen nach Schweregrad (z. B. kritisch, hoch, mittel, niedrig).
- Passen Sie die Prioritäten der Warnmeldungen an die Auswirkungen auf das Unternehmen an.

Vermeiden Sie Alarmermüdung:

- Reduzieren Sie den Geräuschpegel, indem Sie redundante oder geringwertige Warnmeldungen eliminieren.
- Korrelieren Sie Warnmeldungen mit gruppenbezogenen Problemen.

Verwenden Sie mehrstufige Warnmeldungen:

- Implementieren Sie Warnschwellen, bevor kritische Werte erreicht werden.
- Verwenden Sie unterschiedliche Benachrichtigungskanäle für unterschiedliche Warnschweregrade.

Implementieren Sie die richtige Weiterleitung von Alarmen:

- Stellen Sie sicher, dass Benachrichtigungen an die richtigen Teams oder Einzelpersonen gesendet werden.
- Nutzen Sie Bereitschaftszeiten und Rotationen, um den ganzen Tag und jeden Tag abzudecken.

Nutzen Sie native Kubernetes-Metriken:

- Überwachen Sie die Kernkomponenten von Kubernetes (Knoten, Pods, Dienste).
- Verwenden Sie kube-state-metrics (KSM) für zusätzliche Kubernetes-Objektmetriken.

Überwachen Sie sowohl die Infrastruktur als auch die Anwendungen:

- Richten Sie Warnmeldungen für den Zustand des Clusters, den Knotenstatus und die Ressourcenauslastung ein.
- Implementieren Sie anwendungsspezifische Warnmeldungen wie Fehlerraten und Latenz.

Verwenden Sie Prometheus und Alertmanager:

- Verwenden Sie Prometheus für die Erfassung von Metriken und PromQL, um Alarmbedingungen zu definieren.
- Verwenden Sie Alertmanager für die Weiterleitung und Deduplizierung von Alarmen.

Integrieren Sie mit Amazon CloudWatch:

- Verwenden Sie CloudWatchContainer Insights für Amazon EKS-spezifische Metriken.
- Richten Sie CloudWatchAlarme für kritische AWS Ressourcenmetriken ein.

Implementieren Sie kontextreiche Warnmeldungen:

- Nehmen Sie relevante Informationen wie Clustername, Namespace und Pod-Details in Warnmeldungen auf.
- Stellen Sie in Warnmeldungen Links zu relevanten Dashboards oder Runbooks bereit.

Verwenden Sie die Erkennung von Anomalien:

- Implementieren Sie auf maschinellem Lernen basierende Anomalieerkennung für komplexe Muster.
- Verwenden Sie Dienste wie die Erkennung von CloudWatch Anomalien oder Tools von Drittanbietern.

Implementieren Sie die Unterdrückung und Stummschaltung von Alarmen:

- Erlaubt die vorübergehende Unterdrückung bekannter Probleme.
- Implementieren Sie Wartungsfenster, um den Geräuschpegel bei geplanten Ausfallzeiten zu reduzieren.

Überwachen Sie die Leistung von Warnmeldungen:

- Verfolgen Sie Kennzahlen wie Warnungshäufigkeit, Lösungszeit und Falsch-Positiv-Raten.
- Überprüfen und verfeinern Sie die Warnregeln regelmäßig auf der Grundlage dieser Kennzahlen.

Implementieren Sie Eskalationsverfahren:

- Definieren Sie klare Eskalationspfade für ungelöste Alarme.
- Verwenden Sie Tools wie PagerDuty oder Opsgenie für automatisierte Eskalationen.

Testen Sie die Warnsysteme regelmäßig:

- Führen Sie regelmäßige Tests Ihrer Alarm-Pipeline durch.
- Schließen Sie Warntests in Notfallwiederherstellungsübungen ein.

Verwenden Sie Vorlagen für einheitliche Warnmeldungen:

- Erstellen Sie standardisierte Warnungsvorlagen für gängige Szenarien.
- Sorgen Sie für eine einheitliche Formatierung und Information in allen Warnmeldungen.

Implementieren Sie eine Ratenbegrenzung:

- Beugen Sie Stürmen vor, indem Sie eine Ratenbegrenzung für häufig ausgelöste Alarme einrichten.

Verwenden Sie benutzerdefinierte Metriken:

- Implementieren Sie benutzerdefinierte Metriken für die anwendungsspezifische Überwachung.
- Verwenden Sie die Kubernetes-API für benutzerdefinierte Metriken für die automatische Skalierung auf der Grundlage dieser Metriken.

Implementieren Sie die Protokollierungsintegration:

- Korrelieren Sie Warnmeldungen mit relevanten Protokollen, um die Fehlerbehebung zu beschleunigen.
- Verwenden Sie Tools wie Grafana Loki oder den ELK Stack in Verbindung mit Ihrem Warnsystem.

Ziehen Sie Kostenwarnungen in Betracht:

- Richten Sie Benachrichtigungen für unerwartete Spitzen beim Ressourcenverbrauch oder bei den Kosten ein.
- Verwenden Sie Kostenmanagement-Tools AWS Budgetsoder Tools von Drittanbietern.

Verwenden Sie verteiltes Tracing:

- Integrieren Sie verteilte Tracing-Tools wie Jaeger oder. AWS X-Ray
- Richten Sie Warnmeldungen für abnormale Ablaufverfolgungsmuster oder Latenzen ein.

Runbooks für Dokumentwarnungen:

- Erstellen Sie klare, umsetzbare Runbooks für jeden Warnungstyp.
- Nehmen Sie Schritte zur Fehlerbehebung und Eskalationsverfahren in Runbooks auf.

Wenn Sie diese bewährten Methoden befolgen, können Sie ein robustes, effizientes und effektives Warnsystem für Ihre Amazon EKS-Umgebung einrichten. Dies trägt dazu bei, eine hohe Verfügbarkeit, eine schnelle Problemlösung und eine optimale Leistung Ihrer Kubernetes-basierten Anwendungen sicherzustellen.

Nächste Schritte

Dieser Leitfaden bot ein umfassendes Framework für die Implementierung robuster Observability in Amazon EKS-Umgebungen, wobei der Schwerpunkt auf der Erfassung von Metriken, der Protokollierungsinfrastruktur, der verteilten Ablaufverfolgung und der Kostenoptimierung lag. Wenn Sie diese Kernkomponenten verstehen und anwenden, können Sie eine äußerst beobachtbare, wartungsfreundliche und kostengünstige Container-Umgebung aufbauen, die tiefe Einblicke in das Anwendungs- und Infrastrukturverhalten bietet. Die Integration von AWS-Services Amazon CloudWatch Container Insights und AWS X-Ray, kombiniert mit Open-Source-Lösungen wie Prometheus und OpenTelemetry, schafft eine leistungsstarke Grundlage für die Überwachung und Fehlerbehebung von containerisierten Anwendungen.

Der Erfolg der Implementierung hängt von einem schrittweisen Ansatz ab, der mit der Erfassung der wichtigsten Kennzahlen beginnt und schrittweise auf umfassende Funktionen für Protokollierung und verteilte Ablaufverfolgung ausgeweitet wird. Wir empfehlen Ihnen, zunächst Ihre aktuellen Überwachungskapazitäten zu bewerten, Lücken zu identifizieren und geeignete Tool-Kombinationen auszuwählen, die Ihren betrieblichen Anforderungen und der Expertise Ihres Teams entsprechen. Dieser methodische Ansatz stellt sicher, dass jede Komponente des Observability-Stacks ordnungsgemäß implementiert und integriert wird, während die Teams die erforderlichen Fähigkeiten und Prozesse entwickeln, um diese Tools effektiv zu nutzen.

Die langfristige Nachhaltigkeit der Amazon EKS-Observability hängt von der regelmäßigen Optimierung von Kosten, Ressourcen und Prozessen ab. Sie sollten Ihre Observability-Infrastruktur, einschließlich Richtlinien zur Datenspeicherung, Stichprobenraten und Ressourcenzuweisung, kontinuierlich überprüfen und anpassen, um das richtige Gleichgewicht zwischen umfassender Überwachung und betrieblicher Effizienz zu wahren. Dieser iterative Verbesserungsansatz in Kombination mit kontinuierlichen Teamschulungen und Aktualisierungen der Dokumentation ermöglicht es Ihrem Unternehmen, eine effektive Beobachtbarkeit aufrechtzuerhalten und gleichzeitig das Geschäftswachstum zu unterstützen und sich an sich entwickelnde Anwendungsarchitekturen anzupassen.

Ressourcen

AWS Dokumentation

- Leitfaden zu bewährten Methoden für Amazon EKS
- Einblicke in Amazon CloudWatch Container
- Amazon Managed Service for Prometheus
- Amazon-verwaltetes Grafana
- AWS Distribution für und OpenTelemetry AWS X-Ray
- OpenSearch Amazon-Dienst

AWS Blog-Beiträge

- Amazon EKS verbessert die Beobachtbarkeit der Kubernetes-Steuerebene
- Automatisieren der Erfassung von Metriken auf Amazon EKS mit Amazon Managed Service für von Prometheus verwaltete Scraper
- Automatisieren Sie die Überwachung Ihres Amazon EKS-Clusters mit CloudWatch Container Insights
- Verbesserung der Beobachtbarkeit mit einer verwalteten Überwachungslösung für Amazon EKS

Sonstige Ressourcen

- OpenTelemetry-Dokumentation
- Prometheus-Dokumentation
- Fluent Bit-Dokumentation
- Dokumentation zur Überwachung, Protokollierung und Debugging in Kubernetes

Dokumentverlauf

In der folgenden Tabelle werden wichtige Änderungen in diesem Leitfaden beschrieben. Um Benachrichtigungen über zukünftige Aktualisierungen zu erhalten, können Sie einen RSS-Feed abonnieren.

Änderung	Beschreibung	Datum
Aktualisierungen	Wir haben das Kapitel Logging in Amazon EKS aktualisiert.	17. März 2026
Erste Veröffentlichung	—	10. April 2025

AWS Glossar zu präskriptiven Leitlinien

Die folgenden Begriffe werden häufig in Strategien, Leitfäden und Mustern von AWS Prescriptive Guidance verwendet. Um Einträge vorzuschlagen, verwenden Sie bitte den Link Feedback geben am Ende des Glossars.

Zahlen

7 Rs

Sieben gängige Migrationsstrategien für die Verlagerung von Anwendungen in die Cloud. Diese Strategien bauen auf den 5 Rs auf, die Gartner 2011 identifiziert hat, und bestehen aus folgenden Elementen:

- Faktorwechsel/Architekturwechsel – Verschieben Sie eine Anwendung und ändern Sie ihre Architektur, indem Sie alle Vorteile cloudnativer Feature nutzen, um Agilität, Leistung und Skalierbarkeit zu verbessern. Dies beinhaltet in der Regel die Portierung des Betriebssystems und der Datenbank. Beispiel: Migrieren Sie Ihre lokale Oracle-Datenbank auf die Amazon Aurora PostgreSQL-kompatible Edition.
- Plattformwechsel (Lift and Reshape) – Verschieben Sie eine Anwendung in die Cloud und führen Sie ein gewisses Maß an Optimierung ein, um die Cloud-Funktionen zu nutzen. Beispiel: Migrieren Sie Ihre lokale Oracle-Datenbank zu Amazon Relational Database Service (Amazon RDS) für Oracle in der. AWS Cloud
- Neukauf (Drop and Shop) – Wechseln Sie zu einem anderen Produkt, indem Sie typischerweise von einer herkömmlichen Lizenz zu einem SaaS-Modell wechseln. Beispiel: Migrieren Sie Ihr CRM-System (Customer Relationship Management) zu Salesforce.com.
- Hostwechsel (Lift and Shift) – Verschieben Sie eine Anwendung in die Cloud, ohne Änderungen vorzunehmen, um die Cloud-Funktionen zu nutzen. Beispiel: Migrieren Sie Ihre lokale Oracle-Datenbank zu Oracle auf einer EC2-Instanz in der. AWS Cloud
- Verschieben (Lift and Shift auf Hypervisor-Ebene) – Verlagern Sie die Infrastruktur in die Cloud, ohne neue Hardware kaufen, Anwendungen umschreiben oder Ihre bestehenden Abläufe ändern zu müssen. Sie migrieren Server von einer lokalen Plattform zu einem Cloud-Dienst für dieselbe Plattform. Beispiel: Migrieren Sie eine Microsoft Hyper-V Anwendung zu AWS.
- Beibehaltung (Wiederaufgreifen) – Bewahren Sie Anwendungen in Ihrer Quellumgebung auf. Dazu können Anwendungen gehören, die einen umfangreichen Faktorwechsel erfordern und

die Sie auf einen späteren Zeitpunkt verschieben möchten, sowie ältere Anwendungen, die Sie beibehalten möchten, da es keine geschäftliche Rechtfertigung für ihre Migration gibt.

- Außerbetriebnahme – Dekommissionierung oder Entfernung von Anwendungen, die in Ihrer Quellumgebung nicht mehr benötigt werden.

A

ABAC

Siehe attributbasierte Zugriffskontrolle.

abstrahierte Dienste

Siehe Managed Services.

ACID

Siehe Atomarität, Konsistenz, Isolierung und Haltbarkeit.

Aktiv-Aktiv-Migration

Eine Datenbankmigrationsmethode, bei der die Quell- und Zieldatenbanken synchron gehalten werden (mithilfe eines bidirektionalen Replikationstools oder dualer Schreibvorgänge) und beide Datenbanken Transaktionen von miteinander verbundenen Anwendungen während der Migration verarbeiten. Diese Methode unterstützt die Migration in kleinen, kontrollierten Batches, anstatt einen einmaligen Cutover zu erfordern. Es ist flexibler, erfordert aber mehr Arbeit als eine aktiv-passive Migration.

Aktiv-Passiv-Migration

Eine Datenbankmigrationsmethode, bei der die Quell- und Zieldatenbanken synchron gehalten werden, aber nur die Quelldatenbank verarbeitet Transaktionen von verbindenden Anwendungen, während Daten in die Zieldatenbank repliziert werden. Die Zieldatenbank akzeptiert während der Migration keine Transaktionen.

Aggregatfunktion

Eine SQL-Funktion, die mit einer Gruppe von Zeilen arbeitet und einen einzelnen Rückgabewert für die Gruppe berechnet. Beispiele für Aggregatfunktionen sind `SUM` und`MAX`.

AI

Siehe künstliche Intelligenz.

AIOps

Siehe Operationen im Bereich künstliche Intelligenz.

Anonymisierung

Der Prozess des dauerhaften Löschens personenbezogener Daten in einem Datensatz. Anonymisierung kann zum Schutz der Privatsphäre beitragen. Anonymisierte Daten gelten nicht mehr als personenbezogene Daten.

Anti-Muster

Eine häufig verwendete Lösung für ein wiederkehrendes Problem, bei dem die Lösung kontraproduktiv, ineffektiv oder weniger wirksam als eine Alternative ist.

Anwendungssteuerung

Ein Sicherheitsansatz, bei dem nur zugelassene Anwendungen verwendet werden können, um ein System vor Schadsoftware zu schützen.

Anwendungsportfolio

Eine Sammlung detaillierter Informationen zu jeder Anwendung, die von einer Organisation verwendet wird, einschließlich der Kosten für die Erstellung und Wartung der Anwendung und ihres Geschäftswerts. Diese Informationen sind entscheidend für den Prozess der Portfoliofindung und -analyse und hilft bei der Identifizierung und Priorisierung der Anwendungen, die migriert, modernisiert und optimiert werden sollen.

künstliche Intelligenz (KI)

Das Gebiet der Datenverarbeitungswissenschaft, das sich der Nutzung von Computertechnologien zur Ausführung kognitiver Funktionen widmet, die typischerweise mit Menschen in Verbindung gebracht werden, wie Lernen, Problemlösen und Erkennen von Mustern. Weitere Informationen finden Sie unter Was ist künstliche Intelligenz?

Operationen mit künstlicher Intelligenz (AIOps)

Der Prozess des Einsatzes von Techniken des Machine Learning zur Lösung betrieblicher Probleme, zur Reduzierung betrieblicher Zwischenfälle und menschlicher Eingriffe sowie zur Steigerung der Servicequalität. Weitere Informationen zur Verwendung in der AWS Migrationsstrategie finden Sie im Operations Integration Guide. AIOps

Asymmetrische Verschlüsselung

Ein Verschlüsselungsalgorithmus, der ein Schlüsselpaar, einen öffentlichen Schlüssel für die Verschlüsselung und einen privaten Schlüssel für die Entschlüsselung verwendet. Sie können den

öffentlichen Schlüssel teilen, da er nicht für die Entschlüsselung verwendet wird. Der Zugriff auf den privaten Schlüssel sollte jedoch stark eingeschränkt sein.

Atomizität, Konsistenz, Isolierung, Haltbarkeit (ACID)

Eine Reihe von Softwareeigenschaften, die die Datenvalidität und betriebliche Zuverlässigkeit einer Datenbank auch bei Fehlern, Stromausfällen oder anderen Problemen gewährleisten.

Attributbasierte Zugriffskontrolle (ABAC)

Die Praxis, detaillierte Berechtigungen auf der Grundlage von Benutzerattributen wie Abteilung, Aufgabenrolle und Teamname zu erstellen. Weitere Informationen finden Sie unter ABAC AWS in der AWS Identity and Access Management (IAM-) Dokumentation.

maßgebliche Datenquelle

Ein Ort, an dem Sie die primäre Version der Daten speichern, die als die zuverlässigste Informationsquelle angesehen wird. Sie können Daten aus der maßgeblichen Datenquelle an andere Speicherorte kopieren, um die Daten zu verarbeiten oder zu ändern, z. B. zu anonymisieren, zu redigieren oder zu pseudonymisieren.

Availability Zone

Ein bestimmter Standort innerhalb einer AWS-Region , der vor Ausfällen in anderen Availability Zones geschützt ist und kostengünstige Netzwerkkonnektivität mit niedriger Latenz zu anderen Availability Zones in derselben Region bietet.

AWS Framework für die Einführung der Cloud (AWS CAF)

Ein Framework mit Richtlinien und bewährten Verfahren, das Unternehmen bei der Entwicklung eines effizienten und effektiven Plans für die erfolgreiche Umstellung auf die Cloud unterstützt. AWS AWS CAF unterteilt die Leitlinien in sechs Schwerpunktbereiche, die als Perspektiven bezeichnet werden: Unternehmen, Mitarbeiter, Unternehmensführung, Plattform, Sicherheit und Betrieb. Die Perspektiven Geschäft, Mitarbeiter und Unternehmensführung konzentrieren sich auf Geschäftskompetenzen und -prozesse, während sich die Perspektiven Plattform, Sicherheit und Betriebsabläufe auf technische Fähigkeiten und Prozesse konzentrieren. Die Personalperspektive zielt beispielsweise auf Stakeholder ab, die sich mit Personalwesen (HR), Personalfunktionen und Personalmanagement befassen. Aus dieser Perspektive bietet AWS CAF Leitlinien für Personalentwicklung, Schulung und Kommunikation, um das Unternehmen auf eine erfolgreiche Cloud-Einführung vorzubereiten. Weitere Informationen finden Sie auf der AWS -CAF-Webseite und dem AWS -CAF-Whitepaper.

AWS Workload-Qualifizierungsrahmen (AWS WQF)

Ein Tool, das Workloads bei der Datenbankmigration bewertet, Migrationsstrategien empfiehlt und Arbeitsschätzungen bereitstellt. AWS WQF ist in () enthalten. AWS Schema Conversion Tool AWS SCT Es analysiert Datenbankschemas und Codeobjekte, Anwendungscode, Abhängigkeiten und Leistungsmerkmale und stellt Bewertungsberichte bereit.

B

schlechter Bot

Ein Bot, der Einzelpersonen oder Organisationen stören oder ihnen Schaden zufügen soll.

BCP

Siehe Planung der Geschäftskontinuität.

Verhaltensdiagramm

Eine einheitliche, interaktive Ansicht des Ressourcenverhaltens und der Interaktionen im Laufe der Zeit. Sie können ein Verhaltensdiagramm mit Amazon Detective verwenden, um fehlgeschlagene Anmeldeversuche, verdächtige API-Aufrufe und ähnliche Vorgänge zu untersuchen. Weitere Informationen finden Sie unter Daten in einem Verhaltensdiagramm in der Detective-Dokumentation.

Big-Endian-System

Ein System, welches das höchstwertige Byte zuerst speichert. Siehe auch Endianness.

Binäre Klassifikation

Ein Prozess, der ein binäres Ergebnis vorhersagt (eine von zwei möglichen Klassen). Beispielsweise könnte Ihr ML-Modell möglicherweise Probleme wie „Handelt es sich bei dieser E-Mail um Spam oder nicht?“ vorhersagen müssen oder „Ist dieses Produkt ein Buch oder ein Auto?“

Bloom-Filter

Eine probabilistische, speichereffiziente Datenstruktur, mit der getestet wird, ob ein Element Teil einer Menge ist.

Blau/Grün-Bereitstellung

Eine Bereitstellungsstrategie, bei der Sie zwei separate, aber identische Umgebungen erstellen. Sie führen die aktuelle Anwendungsversion in einer Umgebung (blau) und die neue

Anwendungsversion in der anderen Umgebung (grün) aus. Mit dieser Strategie können Sie schnell und mit minimalen Auswirkungen ein Rollback durchführen.

Bot

Eine Softwareanwendung, die automatisierte Aufgaben über das Internet ausführt und menschliche Aktivitäten oder Interaktionen simuliert. Manche Bots sind nützlich oder nützlich, wie z. B. Webcrawler, die Informationen im Internet indexieren. Einige andere Bots, sogenannte bösartige Bots, sollen Einzelpersonen oder Organisationen stören oder ihnen Schaden zufügen.

Botnetz

Netzwerke von Bots, die mit Malware infiziert sind und unter der Kontrolle einer einzigen Partei stehen, die als Bot-Herder oder Bot-Operator bezeichnet wird. Botnetze sind der bekannteste Mechanismus zur Skalierung von Bots und ihrer Wirkung.

branch

Ein containerisierter Bereich eines Code-Repositorys. Der erste Zweig, der in einem Repository erstellt wurde, ist der Hauptzweig. Sie können einen neuen Zweig aus einem vorhandenen Zweig erstellen und dann Feature entwickeln oder Fehler in dem neuen Zweig beheben. Ein Zweig, den Sie erstellen, um ein Feature zu erstellen, wird allgemein als Feature-Zweig bezeichnet. Wenn das Feature zur Veröffentlichung bereit ist, führen Sie den Feature-Zweig wieder mit dem Hauptzweig zusammen. Weitere Informationen finden Sie unter Über Branches (GitHub Dokumentation).

Zugang durch Glasbruch

Unter außergewöhnlichen Umständen und im Rahmen eines genehmigten Verfahrens ist dies eine schnelle Methode für einen Benutzer, auf einen Bereich zuzugreifen AWS-Konto , für den er normalerweise keine Zugriffsrechte besitzt. Weitere Informationen finden Sie unter dem Indikator Implementation break-glass procedures in den AWS Well-Architected-Leitlinien.

Brownfield-Strategie

Die bestehende Infrastruktur in Ihrer Umgebung. Wenn Sie eine Brownfield-Strategie für eine Systemarchitektur anwenden, richten Sie sich bei der Gestaltung der Architektur nach den Einschränkungen der aktuellen Systeme und Infrastruktur. Wenn Sie die bestehende Infrastruktur erweitern, könnten Sie Brownfield- und Greenfield-Strategien mischen.

Puffer-Cache

Der Speicherbereich, in dem die am häufigsten abgerufenen Daten gespeichert werden.

Geschäftsfähigkeit

Was ein Unternehmen tut, um Wert zu generieren (z. B. Vertrieb, Kundenservice oder Marketing). Microservices-Architekturen und Entwicklungsentscheidungen können von den Geschäftskapazitäten beeinflusst werden. Weitere Informationen finden Sie im Abschnitt Organisiert nach Geschäftskapazitäten des Whitepapers Ausführen von containerisierten Microservices in AWS.

Planung der Geschäftskontinuität (BCP)

Ein Plan, der die potenziellen Auswirkungen eines störenden Ereignisses, wie z. B. einer groß angelegten Migration, auf den Betrieb berücksichtigt und es einem Unternehmen ermöglicht, den Betrieb schnell wieder aufzunehmen.

C

CAF

Weitere Informationen finden Sie unter Framework AWS für die Cloud-Einführung.

Bereitstellung auf Kanaren

Die langsame und schrittweise Veröffentlichung einer Version für Endbenutzer. Wenn Sie sich sicher sind, stellen Sie die neue Version bereit und ersetzen die aktuelle Version vollständig.

CCoE

Weitere Informationen finden Sie im Cloud Center of Excellence.

CDC

Siehe Erfassung von Änderungsdaten.

Erfassung von Datenänderungen (CDC)

Der Prozess der Nachverfolgung von Änderungen an einer Datenquelle, z. B. einer Datenbanktabelle, und der Aufzeichnung von Metadaten zu der Änderung. Sie können CDC für verschiedene Zwecke verwenden, z. B. für die Prüfung oder Replikation von Änderungen in einem Zielsystem, um die Synchronisation aufrechtzuerhalten.

Chaos-Technik

Absichtliches Einführen von Ausfällen oder Störungsereignissen, um die Widerstandsfähigkeit eines Systems zu testen. Sie können AWS Fault Injection Service (AWS FIS) verwenden, um Experimente durchzuführen, die Ihre AWS Workloads stress, und deren Reaktion zu bewerten.

CI/CD

Siehe Continuous Integration und Continuous Delivery.

Klassifizierung

Ein Kategorisierungsprozess, der bei der Erstellung von Vorhersagen hilft. ML-Modelle für Klassifikationsprobleme sagen einen diskreten Wert voraus. Diskrete Werte unterscheiden sich immer voneinander. Beispielsweise muss ein Modell möglicherweise auswerten, ob auf einem Bild ein Auto zu sehen ist oder nicht.

clientseitige Verschlüsselung

Lokale Verschlüsselung von Daten, bevor das Ziel sie AWS-Service empfängt.

Cloud-Exzellenzzentrum (CCoE)

Ein multidisziplinäres Team, das die Cloud-Einführung in der gesamten Organisation vorantreibt, einschließlich der Entwicklung bewährter Cloud-Methoden, der Mobilisierung von Ressourcen, der Festlegung von Migrationszeitplänen und der Begleitung der Organisation durch groß angelegte Transformationen. Weitere Informationen finden Sie in den CCoE-Beiträgen im AWS Cloud Enterprise Strategy Blog.

Cloud Computing

Die Cloud-Technologie, die typischerweise für die Ferndatenspeicherung und das IoT-Gerätemanagement verwendet wird. Cloud Computing ist häufig mit Edge-Computing-Technologie verbunden.

Cloud-Betriebsmodell

In einer IT-Organisation das Betriebsmodell, das zum Aufbau, zur Weiterentwicklung und Optimierung einer oder mehrerer Cloud-Umgebungen verwendet wird. Weitere Informationen finden Sie unter Aufbau Ihres Cloud-Betriebsmodells.

Phasen der Einführung der Cloud

Die vier Phasen, die Unternehmen bei der Migration in der Regel durchlaufen AWS Cloud:

- Projekt – Durchführung einiger Cloud-bezogener Projekte zu Machbarkeitsnachweisen und zu Lernzwecken
- Fundament — Tätigen Sie grundlegende Investitionen, um Ihre Cloud-Einführung zu skalieren (z. B. Einrichtung einer landing zone, Definition eines CCo E, Einrichtung eines Betriebsmodells)

- Migration – Migrieren einzelner Anwendungen
- Neuentwicklung – Optimierung von Produkten und Services und Innovation in der Cloud

Diese Phasen wurden von Stephen Orban im Blogbeitrag The Journey Toward Cloud-First & the Stages of Adoption im AWS Cloud Enterprise Strategy-Blog definiert. Informationen darüber, wie sie mit der AWS Migrationsstrategie zusammenhängen, finden Sie im Leitfaden zur Vorbereitung der Migration.

CMDB

Siehe Datenbank für das Konfigurationsmanagement.

Code-Repository

Ein Ort, an dem Quellcode und andere Komponenten wie Dokumentation, Beispiele und Skripts gespeichert und im Rahmen von Versionskontrollprozessen aktualisiert werden. Zu den gängigen Cloud-Repositorys gehören GitHub oderBitbucket Cloud. Jede Version des Codes wird Zweig genannt. In einer Microservice-Struktur ist jedes Repository einer einzelnen Funktionalität gewidmet. Eine einzelne CI/CD-Pipeline kann mehrere Repositorien verwenden.

Kalter Cache

Ein Puffer-Cache, der leer oder nicht gut gefüllt ist oder veraltete oder irrelevante Daten enthält. Dies beeinträchtigt die Leistung, da die Datenbank-Instance aus dem Hauptspeicher oder der Festplatte lesen muss, was langsamer ist als das Lesen aus dem Puffercache.

Kalte Daten

Daten, auf die selten zugegriffen wird und die in der Regel historisch sind. Bei der Abfrage dieser Art von Daten sind langsame Abfragen in der Regel akzeptabel. Durch die Verlagerung dieser Daten auf leistungsschwächere und kostengünstigere Speicherstufen oder -klassen können Kosten gesenkt werden.

Computer Vision (CV)

Ein Bereich der KI, der maschinelles Lernen nutzt, um Informationen aus visuellen Formaten wie digitalen Bildern und Videos zu analysieren und zu extrahieren. Amazon SageMaker AI bietet beispielsweise Bildverarbeitungsalgorithmen für CV.

Drift in der Konfiguration

Bei einer Arbeitslast eine Änderung der Konfiguration gegenüber dem erwarteten Zustand. Dies kann dazu führen, dass der Workload nicht mehr richtlinienkonform wird, und zwar in der Regel schrittweise und unbeabsichtigt.

Verwaltung der Datenbankkonfiguration (CMDB)

Ein Repository, das Informationen über eine Datenbank und ihre IT-Umgebung speichert und verwaltet, inklusive Hardware- und Softwarekomponenten und deren Konfigurationen. In der Regel verwenden Sie Daten aus einer CMDB in der Phase der Portfolioerkennung und -analyse der Migration.

Konformitätspaket

Eine Sammlung von AWS Config Regeln und Abhilfemaßnahmen, die Sie zusammenstellen können, um Ihre Konformitäts- und Sicherheitsprüfungen individuell anzupassen. Mithilfe einer YAML-Vorlage können Sie ein Conformance Pack als einzelne Entität in einer AWS-Konto AND-Region oder unternehmensweit bereitstellen. Weitere Informationen finden Sie in der Dokumentation unter Conformance Packs. AWS Config

Kontinuierliche Bereitstellung und kontinuierliche Integration (CI/CD)

Der Prozess der Automatisierung der Quell-, Build-, Test-, Staging- und Produktionsphasen des Softwareveröffentlichungsprozesses. CI/CD wird allgemein als Pipeline beschrieben. CI/CD kann Ihnen helfen, Prozesse zu automatisieren, die Produktivität zu steigern, die Codequalität zu verbessern und schneller zu liefern. Weitere Informationen finden Sie unter Vorteile der kontinuierlichen Auslieferung. CD kann auch für kontinuierliche Bereitstellung stehen. Weitere Informationen finden Sie unterKontinuierliche Auslieferung im Vergleich zu kontinuierlicher Bereitstellung.

CV

Siehe Computer Vision.

D

Daten im Ruhezustand

Daten, die in Ihrem Netzwerk stationär sind, z. B. Daten, die sich im Speicher befinden.

Datenklassifizierung

Ein Prozess zur Identifizierung und Kategorisierung der Daten in Ihrem Netzwerk auf der Grundlage ihrer Kritikalität und Sensitivität. Sie ist eine wichtige Komponente jeder Strategie für das Management von Cybersecurity-Risiken, da sie Ihnen hilft, die geeigneten Schutz- und Aufbewahrungskontrollen für die Daten zu bestimmen. Die Datenklassifizierung ist ein Bestandteil

der Sicherheitssäule im AWS Well-Architected Framework. Weitere Informationen finden Sie unter Datenklassifizierung.

Datendrift

Eine signifikante Abweichung zwischen den Produktionsdaten und den Daten, die zum Trainieren eines ML-Modells verwendet wurden, oder eine signifikante Änderung der Eingabedaten im Laufe der Zeit. Datendrift kann die Gesamtqualität, Genauigkeit und Fairness von ML-Modellvorhersagen beeinträchtigen.

Daten während der Übertragung

Daten, die sich aktiv durch Ihr Netzwerk bewegen, z. B. zwischen Netzwerkressourcen.

Datennetz

Ein architektonisches Framework, das verteilte, dezentrale Dateneigentum mit zentraler Verwaltung und Steuerung ermöglicht.

Datenminimierung

Das Prinzip, nur die Daten zu sammeln und zu verarbeiten, die unbedingt erforderlich sind. Durch Datenminimierung im AWS Cloud können Datenschutzrisiken, Kosten und der CO2-Fußabdruck Ihrer Analysen reduziert werden.

Datenperimeter

Eine Reihe präventiver Schutzmaßnahmen in Ihrer AWS Umgebung, die sicherstellen, dass nur vertrauenswürdige Identitäten auf vertrauenswürdige Ressourcen von erwarteten Netzwerken zugreifen. Weitere Informationen finden Sie unter Aufbau eines Datenperimeters auf. AWS

Vorverarbeitung der Daten

Rohdaten in ein Format umzuwandeln, das von Ihrem ML-Modell problemlos verarbeitet werden kann. Die Vorverarbeitung von Daten kann bedeuten, dass bestimmte Spalten oder Zeilen entfernt und fehlende, inkonsistente oder doppelte Werte behoben werden.

Herkunft der Daten

Der Prozess der Nachverfolgung des Ursprungs und der Geschichte von Daten während ihres gesamten Lebenszyklus, z. B. wie die Daten generiert, übertragen und gespeichert wurden.

betroffene Person

Eine Person, deren Daten gesammelt und verarbeitet werden.

Data Warehouse

Ein Datenverwaltungssystem, das Business Intelligence wie Analysen unterstützt. Data Warehouses enthalten in der Regel große Mengen historischer Daten und werden in der Regel für Abfragen und Analysen verwendet.

Datenbankdefinitionssprache (DDL)

Anweisungen oder Befehle zum Erstellen oder Ändern der Struktur von Tabellen und Objekten in einer Datenbank.

Datenbankmanipulationssprache (DML)

Anweisungen oder Befehle zum Ändern (Einfügen, Aktualisieren und Löschen) von Informationen in einer Datenbank.

DDL

Siehe Datenbankdefinitionssprache.

Deep-Ensemble

Mehrere Deep-Learning-Modelle zur Vorhersage kombinieren. Sie können Deep-Ensembles verwenden, um eine genauere Vorhersage zu erhalten oder um die Unsicherheit von Vorhersagen abzuschätzen.

Deep Learning

Ein ML-Teilbereich, der mehrere Schichten künstlicher neuronaler Netzwerke verwendet, um die Zuordnung zwischen Eingabedaten und Zielvariablen von Interesse zu ermitteln.

defense-in-depth

Ein Ansatz zur Informationssicherheit, bei dem eine Reihe von Sicherheitsmechanismen und -kontrollen sorgfältig in einem Computernetzwerk verteilt werden, um die Vertraulichkeit, Integrität und Verfügbarkeit des Netzwerks und der darin enthaltenen Daten zu schützen. Wenn Sie diese Strategie anwenden AWS, fügen Sie mehrere Steuerelemente auf verschiedenen Ebenen der AWS Organizations Struktur hinzu, um die Ressourcen zu schützen. Ein defense-in-depth Ansatz könnte beispielsweise Multi-Faktor-Authentifizierung, Netzwerksegmentierung und Verschlüsselung kombinieren.

delegierter Administrator

In AWS Organizations kann ein kompatibler Dienst ein AWS Mitgliedskonto registrieren, um die Konten der Organisation und die Berechtigungen für diesen Dienst zu verwalten. Dieses Konto

wird als delegierter Administrator für diesen Service bezeichnet. Weitere Informationen und eine Liste kompatibler Services finden Sie unter Services, die mit AWS Organizations funktionieren in der AWS Organizations -Dokumentation.

Einsatz

Der Prozess, bei dem eine Anwendung, neue Feature oder Codekorrekturen in der Zielumgebung verfügbar gemacht werden. Die Bereitstellung umfasst das Implementieren von Änderungen an einer Codebasis und das anschließende Erstellen und Ausführen dieser Codebasis in den Anwendungsumgebungen.

Entwicklungsumgebung

Siehe Umgebung.

Detektivische Kontrolle

Eine Sicherheitskontrolle, die darauf ausgelegt ist, ein Ereignis zu erkennen, zu protokollieren und zu warnen, nachdem ein Ereignis eingetreten ist. Diese Kontrollen stellen eine zweite Verteidigungslinie dar und warnen Sie vor Sicherheitsereignissen, bei denen die vorhandenen präventiven Kontrollen umgangen wurden. Weitere Informationen finden Sie unter Detektivische Kontrolle in Implementierung von Sicherheitskontrollen in AWS.

Abbildung des Wertstroms in der Entwicklung (DVSM)

Ein Prozess zur Identifizierung und Priorisierung von Einschränkungen, die sich negativ auf Geschwindigkeit und Qualität im Lebenszyklus der Softwareentwicklung auswirken. DVSM erweitert den Prozess der Wertstromanalyse, der ursprünglich für Lean-Manufacturing-Praktiken konzipiert wurde. Es konzentriert sich auf die Schritte und Teams, die erforderlich sind, um durch den Softwareentwicklungsprozess Mehrwert zu schaffen und zu steigern.

digitaler Zwilling

Eine virtuelle Darstellung eines realen Systems, z. B. eines Gebäudes, einer Fabrik, einer Industrieanlage oder einer Produktionslinie. Digitale Zwillinge unterstützen vorausschauende Wartung, Fernüberwachung und Produktionsoptimierung.

Maßtabelle

In einem Sternschema eine kleinere Tabelle, die Datenattribute zu quantitativen Daten in einer Faktentabelle enthält. Bei Attributen von Dimensionstabellen handelt es sich in der Regel um Textfelder oder diskrete Zahlen, die sich wie Text verhalten. Diese Attribute werden häufig zum Einschränken von Abfragen, zum Filtern und zur Kennzeichnung von Ergebnismengen verwendet.

Katastrophe

Ein Ereignis, das verhindert, dass ein Workload oder ein System seine Geschäftsziele an seinem primären Einsatzort erfüllt. Diese Ereignisse können Naturkatastrophen, technische Ausfälle oder das Ergebnis menschlichen Handelns sein, z. B. unbeabsichtigte Fehlkonfigurationen oder ein Malware-Angriff.

Notfallwiederherstellung (DR)

Die Strategie und der Prozess, mit denen Sie Ausfallzeiten und Datenverluste aufgrund einer Katastrophe minimieren. Weitere Informationen finden Sie unter Disaster Recovery von Workloads unter AWS: Wiederherstellung in der Cloud im AWS Well-Architected Framework.

DML

Siehe Sprache zur Datenbankmanipulation.

Domainorientiertes Design

Ein Ansatz zur Entwicklung eines komplexen Softwaresystems, bei dem seine Komponenten mit sich entwickelnden Domains oder Kerngeschäftszielen verknüpft werden, denen jede Komponente dient. Dieses Konzept wurde von Eric Evans in seinem Buch Domaingesteuertes Design: Bewältigen der Komplexität im Herzen der Software (Boston: Addison-Wesley Professional, 2003) vorgestellt. Informationen darüber, wie Sie domaingesteuertes Design mit dem Strangler-Fig-Muster verwenden können, finden Sie unter Schrittweises Modernisieren älterer Microsoft ASP.NET (ASMX)-Webservices mithilfe von Containern und Amazon API Gateway.

DR

Siehe Disaster Recovery.

Erkennung von Driften

Verfolgung von Abweichungen von einer Basiskonfiguration. Sie können es beispielsweise verwenden, AWS CloudFormation um Abweichungen bei den Systemressourcen zu erkennen, oder Sie können AWS Control Tower damit Änderungen in Ihrer landing zone erkennen, die sich auf die Einhaltung von Governance-Anforderungen auswirken könnten.

DVSM

Siehe Abbildung des Wertstroms in der Entwicklung.

E

EDA

Siehe explorative Datenanalyse.

EDI

Siehe elektronischer Datenaustausch.

Edge-Computing

Die Technologie, die die Rechenleistung für intelligente Geräte an den Rändern eines IoT-Netzwerks erhöht. Im Vergleich zu Cloud Computing kann Edge Computing die Kommunikationslatenz reduzieren und die Reaktionszeit verbessern.

elektronischer Datenaustausch (EDI)

Der automatisierte Austausch von Geschäftsdokumenten zwischen Organisationen. Weitere Informationen finden Sie unter Was ist elektronischer Datenaustausch.

Verschlüsselung

Ein Rechenprozess, der Klartextdaten, die für Menschen lesbar sind, in Chiffretext umwandelt.

Verschlüsselungsschlüssel

Eine kryptografische Zeichenfolge aus zufälligen Bits, die von einem Verschlüsselungsalgorithmus generiert wird. Schlüssel können unterschiedlich lang sein, und jeder Schlüssel ist so konzipiert, dass er unvorhersehbar und einzigartig ist.

Endianismus

Die Reihenfolge, in der Bytes im Computerspeicher gespeichert werden. Big-Endian-Systeme speichern das höchstwertige Byte zuerst. Little-Endian-Systeme speichern das niedrigwertigste Byte zuerst.

Endpunkt

Siehe Service-Endpunkt.

Endpunkt-Services

Ein Service, den Sie in einer Virtual Private Cloud (VPC) hosten können, um ihn mit anderen Benutzern zu teilen. Sie können einen Endpunktdienst mit anderen AWS-Konten oder AWS Identity and Access Management (IAM AWS PrivateLink -) Prinzipalen erstellen und diesen

Berechtigungen gewähren. Diese Konten oder Prinzipale können sich privat mit Ihrem Endpunktservice verbinden, indem sie Schnittstellen-VPC-Endpunkte erstellen. Weitere Informationen finden Sie unter Einen Endpunkt-Service erstellen in der Amazon Virtual Private Cloud (Amazon VPC)-Dokumentation.

Unternehmensressourcenplanung (ERP)

Ein System, das wichtige Geschäftsprozesse (wie Buchhaltung, MES und Projektmanagement) für ein Unternehmen automatisiert und verwaltet.

Envelope-Verschlüsselung

Der Prozess der Verschlüsselung eines Verschlüsselungsschlüssels mit einem anderen Verschlüsselungsschlüssel. Weitere Informationen finden Sie unter Envelope-Verschlüsselung in der AWS Key Management Service (AWS KMS) -Dokumentation.

Umgebung

Eine Instance einer laufenden Anwendung. Die folgenden Arten von Umgebungen sind beim Cloud-Computing üblich:

- Entwicklungsumgebung – Eine Instance einer laufenden Anwendung, die nur dem Kernteam zur Verfügung steht, das für die Wartung der Anwendung verantwortlich ist. Entwicklungsumgebungen werden verwendet, um Änderungen zu testen, bevor sie in höhere Umgebungen übertragen werden. Diese Art von Umgebung wird manchmal als Testumgebung bezeichnet.
- Niedrigere Umgebungen – Alle Entwicklungsumgebungen für eine Anwendung, z. B. solche, die für erste Builds und Tests verwendet wurden.
- Produktionsumgebung – Eine Instance einer laufenden Anwendung, auf die Endbenutzer zugreifen können. In einer CI/CD Pipeline ist die Produktionsumgebung die letzte Bereitstellungsumgebung.
- Höhere Umgebungen – Alle Umgebungen, auf die auch andere Benutzer als das Kernentwicklungsteam zugreifen können. Dies kann eine Produktionsumgebung, Vorproduktionsumgebungen und Umgebungen für Benutzerakzeptanztests umfassen.

Epics

In der agilen Methodik sind dies funktionale Kategorien, die Ihnen helfen, Ihre Arbeit zu organisieren und zu priorisieren. Epics bieten eine allgemeine Beschreibung der Anforderungen und Implementierungsaufgaben. Zu den Sicherheitsepen AWS von CAF gehören beispielsweise Identitäts- und Zugriffsmanagement, Detektivkontrollen, Infrastruktursicherheit, Datenschutz und

Reaktion auf Vorfälle. Weitere Informationen zu Epics in der AWS -Migrationsstrategie finden Sie im Leitfaden zur Programm-Implementierung.

ERP

Siehe Enterprise Resource Planning.

Explorative Datenanalyse (EDA)

Der Prozess der Analyse eines Datensatzes, um seine Hauptmerkmale zu verstehen. Sie sammeln oder aggregieren Daten und führen dann erste Untersuchungen durch, um Muster zu finden, Anomalien zu erkennen und Annahmen zu überprüfen. EDA wird durchgeführt, indem zusammenfassende Statistiken berechnet und Datenvisualisierungen erstellt werden.

F

Faktentabelle

Die zentrale Tabelle in einem Sternschema. Sie speichert quantitative Daten über den Geschäftsbetrieb. In der Regel enthält eine Faktentabelle zwei Arten von Spalten: Spalten, die Kennzahlen enthalten, und Spalten, die einen Fremdschlüssel für eine Dimensionstabelle enthalten.

schnell scheitern

Eine Philosophie, die häufige und inkrementelle Tests verwendet, um den Entwicklungslebenszyklus zu verkürzen. Dies ist ein wichtiger Bestandteil eines agilen Ansatzes.

Grenze zur Fehlerisolierung

Dabei handelt es sich um eine Grenze AWS Cloud, z. B. eine Availability Zone AWS-Region, eine Steuerungsebene oder eine Datenebene, die die Auswirkungen eines Fehlers begrenzt und die Widerstandsfähigkeit von Workloads verbessert. Weitere Informationen finden Sie unter Grenzen zur AWS Fehlerisolierung.

Feature-Zweig

Siehe Zweig.

Features

Die Eingabedaten, die Sie verwenden, um eine Vorhersage zu treffen. In einem Fertigungskontext könnten Feature beispielsweise Bilder sein, die regelmäßig von der Fertigungslinie aus aufgenommen werden.

Bedeutung der Feature

Wie wichtig ein Feature für die Vorhersagen eines Modells ist. Dies wird in der Regel als numerischer Wert ausgedrückt, der mit verschiedenen Techniken wie Shapley Additive Explanations (SHAP) und integrierten Gradienten berechnet werden kann. Weitere Informationen finden Sie unter Interpretierbarkeit von Modellen für maschinelles Lernen mit AWS.

Featuretransformation

Daten für den ML-Prozess optimieren, einschließlich der Anreicherung von Daten mit zusätzlichen Quellen, der Skalierung von Werten oder der Extraktion mehrerer Informationssätze aus einem einzigen Datenfeld. Das ermöglicht dem ML-Modell, von den Daten profitieren. Wenn Sie beispielsweise das Datum „27.05.2021 00:15:37“ in „2021“, „Mai“, „Donnerstag“ und „15“ aufschlüsseln, können Sie dem Lernalgorithmus helfen, nuancierte Muster zu erlernen, die mit verschiedenen Datenkomponenten verknüpft sind.

Eingabeaufforderung mit wenigen Klicks

Bereitstellung einer kleinen Anzahl von Beispielen, die die Aufgabe und das gewünschte Ergebnis veranschaulichen, bevor das LLM aufgefordert wird, eine ähnliche Aufgabe auszuführen. Bei dieser Technik handelt es sich um eine Anwendung des kontextbezogenen Lernens, bei der Modelle anhand von Beispielen (Aufnahmen) lernen, die in Eingabeaufforderungen eingebettet sind. Bei Aufgaben, die spezifische Formatierungs-, Argumentations- oder Fachkenntnisse erfordern, kann die Eingabeaufforderung mit wenigen Handgriffen effektiv sein. Siehe auch Zero-Shot Prompting.

FGAC

Siehe detaillierte Zugriffskontrolle.

Feinkörnige Zugriffskontrolle (FGAC)

Die Verwendung mehrerer Bedingungen, um eine Zugriffsanfrage zuzulassen oder abzulehnen.

Flash-Cut-Migration

Eine Datenbankmigrationsmethode, bei der eine kontinuierliche Datenreplikation durch Erfassung von Änderungsdaten verwendet wird, um Daten in kürzester Zeit zu migrieren, anstatt einen schrittweisen Ansatz zu verwenden. Ziel ist es, Ausfallzeiten auf ein Minimum zu beschränken.

FM

Siehe Fundamentmodell.

Fundamentmodell (FM)

Ein großes neuronales Deep-Learning-Netzwerk, das mit riesigen Datensätzen generalisierter und unbeschrifteter Daten trainiert wurde. FMs sind in der Lage, eine Vielzahl allgemeiner Aufgaben zu erfüllen, z. B. Sprache zu verstehen, Text und Bilder zu generieren und Konversationen in natürlicher Sprache zu führen. Weitere Informationen finden Sie unter Was sind Foundation-Modelle.

G

Generative KI

Eine Untergruppe von KI-Modellen, die mit großen Datenmengen trainiert wurden und mit einer einfachen Textaufforderung neue Inhalte und Artefakte wie Bilder, Videos, Text und Audio erstellen können. Weitere Informationen finden Sie unter Was ist Generative KI.

Geoblocking

Siehe geografische Einschränkungen.

Geografische Einschränkungen (Geoblocking)

Bei Amazon eine Option CloudFront, um zu verhindern, dass Benutzer in bestimmten Ländern auf Inhaltsverteilungen zugreifen. Sie können eine Zulassungsliste oder eine Sperrliste verwenden, um zugelassene und gesperrte Länder anzugeben. Weitere Informationen finden Sie in der Dokumentation unter Beschränkung der geografischen Verteilung Ihrer Inhalte. CloudFront

Gitflow-Workflow

Ein Ansatz, bei dem niedrigere und höhere Umgebungen unterschiedliche Zweige in einem Quellcode-Repository verwenden. Der Gitflow-Workflow gilt als veraltet, und der Trunk-basierte Workflow ist der moderne, bevorzugte Ansatz.

goldenes Bild

Ein Snapshot eines Systems oder einer Software, der als Vorlage für die Bereitstellung neuer Instanzen dieses Systems oder dieser Software verwendet wird. In der Fertigung kann ein Golden Image beispielsweise zur Bereitstellung von Software auf mehreren Geräten verwendet werden und trägt zur Verbesserung der Geschwindigkeit, Skalierbarkeit und Produktivität bei der Geräteherstellung bei.

Greenfield-Strategie

Das Fehlen vorhandener Infrastruktur in einer neuen Umgebung. Bei der Einführung einer Neuausrichtung einer Systemarchitektur können Sie alle neuen Technologien ohne Einschränkung der Kompatibilität mit der vorhandenen Infrastruktur auswählen, auch bekannt als Brownfield. Wenn Sie die bestehende Infrastruktur erweitern, könnten Sie Brownfield- und Greenfield-Strategien mischen.

Integritätsschutz

Eine allgemeine Regel, die dazu beiträgt, Ressourcen, Richtlinien und die Einhaltung von Vorschriften in allen Unternehmenseinheiten zu regeln (OUs). Präventiver Integritätsschutz setzt Richtlinien durch, um die Einhaltung von Standards zu gewährleisten. Sie werden mithilfe von Service-Kontrollrichtlinien und IAM-Berechtigungsgrenzen implementiert. Detektivischer Integritätsschutz erkennt Richtlinienverstöße und Compliance-Probleme und generiert Warnmeldungen zur Abhilfe. Sie werden mithilfe von AWS Config, AWS Security Hub CSPM, Amazon GuardDuty AWS Trusted Advisor, Amazon Inspector und benutzerdefinierten AWS Lambda Prüfungen implementiert.

H

HEKTAR

Siehe Hochverfügbarkeit.

Heterogene Datenbankmigration

Migrieren Sie Ihre Quelldatenbank in eine Zieldatenbank, die eine andere Datenbank-Engine verwendet (z. B. Oracle zu Amazon Aurora). Eine heterogene Migration ist in der Regel Teil einer Neuarchitektur, und die Konvertierung des Schemas kann eine komplexe Aufgabe sein. AWS bietet AWS SCT, welches bei Schemakonvertierungen hilft.

hohe Verfügbarkeit (HA)

Die Fähigkeit eines Workloads, im Falle von Herausforderungen oder Katastrophen kontinuierlich und ohne Eingreifen zu arbeiten. HA-Systeme sind so konzipiert, dass sie automatisch ein Failover durchführen, gleichbleibend hohe Leistung bieten und unterschiedliche Lasten und Ausfälle mit minimalen Leistungseinbußen bewältigen.

historische Modernisierung

Ein Ansatz zur Modernisierung und Aufrüstung von Betriebstechnologiesystemen (OT), um den Bedürfnissen der Fertigungsindustrie besser gerecht zu werden. Ein Historian ist eine Art von Datenbank, die verwendet wird, um Daten aus verschiedenen Quellen in einer Fabrik zu sammeln und zu speichern.

Daten zurückhalten

Ein Teil historischer, beschrifteter Daten, der aus einem Datensatz zurückgehalten wird, der zum Trainieren eines Modells für maschinelles Lernen verwendet wird. Sie können Holdout-Daten verwenden, um die Modellleistung zu bewerten, indem Sie die Modellvorhersagen mit den Holdout-Daten vergleichen.

Homogene Datenbankmigration

Migrieren Sie Ihre Quelldatenbank zu einer Zieldatenbank, die dieselbe Datenbank-Engine verwendet (z. B. Microsoft SQL Server zu Amazon RDS für SQL Server). Eine homogene Migration ist in der Regel Teil eines Hostwechsels oder eines Plattformwechsels. Sie können native Datenbankserviceprogramme verwenden, um das Schema zu migrieren.

heiße Daten

Daten, auf die häufig zugegriffen wird, z. B. Echtzeitdaten oder aktuelle Translationsdaten. Für diese Daten ist in der Regel eine leistungsstarke Speicherebene oder -klasse erforderlich, um schnelle Abfrageantworten zu ermöglichen.

Hotfix

Eine dringende Lösung für ein kritisches Problem in einer Produktionsumgebung. Aufgrund seiner Dringlichkeit wird ein Hotfix normalerweise außerhalb des typischen DevOps Release-Workflows erstellt.

Hypercare-Phase

Unmittelbar nach dem Cutover, der Zeitraum, in dem ein Migrationsteam die migrierten Anwendungen in der Cloud verwaltet und überwacht, um etwaige Probleme zu beheben. In der Regel dauert dieser Zeitraum 1–4 Tage. Am Ende der Hypercare-Phase überträgt das Migrationsteam in der Regel die Verantwortung für die Anwendungen an das Cloud-Betriebsteam.

I

IaC

Sehen Sie Infrastruktur als Code.

Identitätsbasierte Richtlinie

Eine Richtlinie, die einem oder mehreren IAM-Prinzipalen zugeordnet ist und deren Berechtigungen innerhalb der AWS Cloud Umgebung definiert.

Leerlaufanwendung

Eine Anwendung mit einer durchschnittlichen CPU- und Arbeitsspeicherauslastung zwischen 5 und 20 Prozent über einen Zeitraum von 90 Tagen. In einem Migrationsprojekt ist es üblich, diese Anwendungen außer Betrieb zu nehmen oder sie On-Premises beizubehalten.

IIoT

Siehe Industrielles Internet der Dinge.

unveränderliche Infrastruktur

Ein Modell, das eine neue Infrastruktur für Produktionsworkloads bereitstellt, anstatt die bestehende Infrastruktur zu aktualisieren, zu patchen oder zu modifizieren. Unveränderliche Infrastrukturen sind von Natur aus konsistenter, zuverlässiger und vorhersehbarer als veränderliche Infrastrukturen. Weitere Informationen finden Sie in der Best Practice Deploy using immutable infrastructure im AWS Well-Architected Framework.

Eingehende (ingress) VPC

In einer Architektur AWS mit mehreren Konten ist dies eine VPC, die Netzwerkverbindungen von außerhalb einer Anwendung akzeptiert, überprüft und weiterleitet. Die AWS Security Reference Architecture empfiehlt, Ihr Netzwerkkonto mit eingehendem und ausgehendem Datenverkehr und Inspektion einzurichten, VPCs um die bidirektionale Schnittstelle zwischen Ihrer Anwendung und dem Internet im weiteren Sinne zu schützen.

Inkrementelle Migration

Eine Cutover-Strategie, bei der Sie Ihre Anwendung in kleinen Teilen migrieren, anstatt eine einziges vollständiges Cutover durchzuführen. Beispielsweise könnten Sie zunächst nur einige Microservices oder Benutzer auf das neue System umstellen. Nachdem Sie sich vergewissert haben, dass alles ordnungsgemäß funktioniert, können Sie weitere Microservices oder Benutzer

schrittweise verschieben, bis Sie Ihr Legacy-System außer Betrieb nehmen können. Diese Strategie reduziert die mit großen Migrationen verbundenen Risiken.

Industrie 4.0

Ein Begriff, der 2016 von Klaus Schwab eingeführt wurde und sich auf die Modernisierung von Fertigungsprozessen durch Fortschritte in den Bereichen Konnektivität, Echtzeitdaten, Automatisierung, Analytik und KI/ML bezieht.

Infrastruktur

Alle Ressourcen und Komponenten, die in der Umgebung einer Anwendung enthalten sind.

Infrastructure as Code (IaC)

Der Prozess der Bereitstellung und Verwaltung der Infrastruktur einer Anwendung mithilfe einer Reihe von Konfigurationsdateien. IaC soll Ihnen helfen, das Infrastrukturmanagement zu zentralisieren, Ressourcen zu standardisieren und schnell zu skalieren, sodass neue Umgebungen wiederholbar, zuverlässig und konsistent sind.

industrielles Internet der Dinge (T) IIo

Einsatz von mit dem Internet verbundenen Sensoren und Geräten in Industriesektoren wie Fertigung, Energie, Automobilindustrie, Gesundheitswesen, Biowissenschaften und Landwirtschaft. Weitere Informationen finden Sie unter Aufbau einer digitalen Transformationsstrategie für das industrielle Internet der Dinge (IIoT).

Inspektions-VPC

In einer Architektur AWS mit mehreren Konten eine zentralisierte VPC, die Inspektionen des Netzwerkverkehrs zwischen VPCs (in demselben oder unterschiedlichen AWS-Regionen), dem Internet und lokalen Netzwerken verwaltet. In der AWS Security Reference Architecture wird empfohlen, Ihr Netzwerkkonto mit eingehendem und ausgehendem Datenverkehr sowie Inspektionen einzurichten, VPCs um die bidirektionale Schnittstelle zwischen Ihrer Anwendung und dem Internet im weiteren Sinne zu schützen.

Internet of Things (IoT)

Das Netzwerk verbundener physischer Objekte mit eingebetteten Sensoren oder Prozessoren, das über das Internet oder über ein lokales Kommunikationsnetzwerk mit anderen Geräten und Systemen kommuniziert. Weitere Informationen finden Sie unter Was ist IoT?

Interpretierbarkeit

Ein Merkmal eines Modells für Machine Learning, das beschreibt, inwieweit ein Mensch verstehen kann, wie die Vorhersagen des Modells von seinen Eingaben abhängen. Weitere Informationen finden Sie unter Interpretierbarkeit des Modells für maschinelles Lernen mit. AWS

IoT

Siehe Internet der Dinge.

IT information library (ITIL, IT-Informationsbibliothek)

Eine Reihe von bewährten Methoden für die Bereitstellung von IT-Services und die Abstimmung dieser Services auf die Geschäftsanforderungen. ITIL bietet die Grundlage für ITSM.

T service management (ITSM, IT-Servicemanagement)

Aktivitäten im Zusammenhang mit der Gestaltung, Implementierung, Verwaltung und Unterstützung von IT-Services für eine Organisation. Informationen zur Integration von Cloud-Vorgängen mit ITSM-Tools finden Sie im Leitfaden zur Betriebsintegration.

BIS

Siehe IT-Informationsbibliothek.

ITSM

Siehe IT-Servicemanagement.

L

Labelbasierte Zugangskontrolle (LBAC)

Eine Implementierung der Mandatory Access Control (MAC), bei der den Benutzern und den Daten selbst jeweils explizit ein Sicherheitslabelwert zugewiesen wird. Die Schnittmenge zwischen der Benutzersicherheitsbeschriftung und der Datensicherheitsbeschriftung bestimmt, welche Zeilen und Spalten für den Benutzer sichtbar sind.

Landing Zone

Eine landing zone ist eine gut strukturierte AWS Umgebung mit mehreren Konten, die skalierbar und sicher ist. Dies ist ein Ausgangspunkt, von dem aus Ihre Organisationen Workloads und Anwendungen schnell und mit Vertrauen in ihre Sicherheits- und Infrastrukturumgebung starten

und bereitstellen können. Weitere Informationen zu Landing Zones finden Sie unter Einrichtung einer sicheren und skalierbaren AWS -Umgebung mit mehreren Konten..

großes Sprachmodell (LLM)

Ein Deep-Learning-KI-Modell, das anhand einer riesigen Datenmenge vorab trainiert wurde. Ein LLM kann mehrere Aufgaben ausführen, z. B. Fragen beantworten, Dokumente zusammenfassen, Text in andere Sprachen übersetzen und Sätze vervollständigen. Weitere Informationen finden Sie unter Was sind. LLMs

Große Migration

Eine Migration von 300 oder mehr Servern.

SCHWARZ

Siehe Labelbasierte Zugriffskontrolle.

Geringste Berechtigung

Die bewährte Sicherheitsmethode, bei der nur die für die Durchführung einer Aufgabe erforderlichen Mindestberechtigungen erteilt werden. Weitere Informationen finden Sie unter Geringste Berechtigungen anwenden in der IAM-Dokumentation.

Lift and Shift

Siehe 7 Rs.

Little-Endian-System

Ein System, welches das niedrigwertigste Byte zuerst speichert. Siehe auch Endianness.

LLM

Siehe großes Sprachmodell.

Niedrigere Umgebungen

Siehe Umgebung.

M

Machine Learning (ML)

Eine Art künstlicher Intelligenz, die Algorithmen und Techniken zur Mustererkennung und zum Lernen verwendet. ML analysiert aufgezeichnete Daten, wie z. B. Daten aus dem Internet der

Dinge (IoT), und lernt daraus, um ein statistisches Modell auf der Grundlage von Mustern zu erstellen. Weitere Informationen finden Sie unter Machine Learning.

Hauptzweig

Siehe Filiale.

Malware

Software, die entwickelt wurde, um die Computersicherheit oder den Datenschutz zu gefährden. Malware kann Computersysteme stören, vertrauliche Informationen durchsickern lassen oder sich unbefugten Zugriff verschaffen. Beispiele für Malware sind Viren, Würmer, Ransomware, Trojaner, Spyware und Keylogger.

verwaltete Dienste

AWS-Services für die die Infrastrukturebene, das Betriebssystem und die Plattformen AWS betrieben werden, und Sie greifen auf die Endgeräte zu, um Daten zu speichern und abzurufen. Amazon Simple Storage Service (Amazon S3) und Amazon DynamoDB sind Beispiele für Managed Services. Diese werden auch als abstrakte Dienste bezeichnet.

Manufacturing Execution System (MES)

Ein Softwaresystem zur Verfolgung, Überwachung, Dokumentation und Steuerung von Produktionsprozessen, bei denen Rohstoffe in der Fertigung zu fertigen Produkten umgewandelt werden.

MAP

Siehe Migration Acceleration Program.

Mechanismus

Ein vollständiger Prozess, bei dem Sie ein Tool erstellen, die Akzeptanz des Tools vorantreiben und anschließend die Ergebnisse überprüfen, um Anpassungen vorzunehmen. Ein Mechanismus ist ein Zyklus, der sich im Laufe seiner Tätigkeit selbst verstärkt und verbessert. Weitere Informationen finden Sie unter Aufbau von Mechanismen im AWS Well-Architected Framework.

Mitgliedskonto

Alle AWS-Konten außer dem Verwaltungskonto, die Teil einer Organisation in sind. AWS Organizations Ein Konto kann jeweils nur Mitglied einer Organisation sein.

MES

Siehe Manufacturing Execution System.

Message Queuing-Telemetrietransport (MQTT)

Ein leichtes machine-to-machine (M2M) -Kommunikationsprotokoll, das auf dem Publish/Subscribe-Muster für IoT-Geräte mit beschränkten Ressourcen basiert.

Microservice

Ein kleiner, unabhängiger Dienst, der über genau definierte Kanäle kommuniziert APIs und in der Regel kleinen, eigenständigen Teams gehört. Ein Versicherungssystem kann beispielsweise Microservices beinhalten, die Geschäftsfunktionen wie Vertrieb oder Marketing oder Subdomains wie Einkauf, Schadenersatz oder Analytik zugeordnet sind. Zu den Vorteilen von Microservices gehören Agilität, flexible Skalierung, einfache Bereitstellung, wiederverwendbarer Code und Ausfallsicherheit. Weitere Informationen finden Sie unter Integration von Microservices mithilfe serverloser Dienste. AWS

Microservices-Architekturen

Ein Ansatz zur Erstellung einer Anwendung mit unabhängigen Komponenten, die jeden Anwendungsprozess als Microservice ausführen. Diese Microservices kommunizieren mithilfe von Lightweight über eine klar definierte Schnittstelle. APIs Jeder Microservice in dieser Architektur kann aktualisiert, bereitgestellt und skaliert werden, um den Bedarf an bestimmten Funktionen einer Anwendung zu decken. Weitere Informationen finden Sie unter Implementierung von Microservices auf. AWS

Migration Acceleration Program (MAP)

Ein AWS Programm, das Beratung, Unterstützung, Schulungen und Services bietet, um Unternehmen dabei zu unterstützen, eine solide betriebliche Grundlage für die Umstellung auf die Cloud zu schaffen und die anfänglichen Kosten von Migrationen auszugleichen. MAP umfasst eine Migrationsmethode für die methodische Durchführung von Legacy-Migrationen sowie eine Reihe von Tools zur Automatisierung und Beschleunigung gängiger Migrationsszenarien.

Migration in großem Maßstab

Der Prozess, bei dem der Großteil des Anwendungsportfolios in Wellen in die Cloud verlagert wird, wobei in jeder Welle mehr Anwendungen schneller migriert werden. In dieser Phase werden die bewährten Verfahren und Erkenntnisse aus den früheren Phasen zur Implementierung einer Migrationsfabrik von Teams, Tools und Prozessen zur Optimierung der Migration von Workloads durch Automatisierung und agile Bereitstellung verwendet. Dies ist die dritte Phase der AWS -Migrationsstrategie.

Migrationsfabrik

Funktionsübergreifende Teams, die die Migration von Workloads durch automatisierte, agile Ansätze optimieren. Zu den Teams in der Migrationsabteilung gehören in der Regel Betriebsabläufe, Geschäftsanalysten und Eigentümer, Migrationsingenieure, Entwickler und DevOps Experten, die in Sprints arbeiten. Zwischen 20 und 50 Prozent eines Unternehmensanwendungsportfolios bestehen aus sich wiederholenden Mustern, die durch einen Fabrik-Ansatz optimiert werden können. Weitere Informationen finden Sie in Diskussion über Migrationsfabriken und den Leitfaden zur Cloud-Migration-Fabrik in diesem Inhaltssatz.

Migrationsmetadaten

Die Informationen über die Anwendung und den Server, die für den Abschluss der Migration benötigt werden. Für jedes Migrationsmuster ist ein anderer Satz von Migrationsmetadaten erforderlich. Beispiele für Migrationsmetadaten sind das Zielsubnetz, die Sicherheitsgruppe und AWS das Konto.

Migrationsmuster

Eine wiederholbare Migrationsaufgabe, in der die Migrationsstrategie, das Migrationsziel und die verwendete Migrationsanwendung oder der verwendete Migrationsservice detailliert beschrieben werden. Beispiel: Rehost-Migration zu Amazon EC2 mit AWS Application Migration Service.

Migration Portfolio Assessment (MPA)

Ein Online-Tool, das Informationen zur Validierung des Geschäftsszenarios für die Migration auf das bereitstellt. AWS Cloud MPA bietet eine detaillierte Portfoliobewertung (richtige Servergröße, Preisgestaltung, Gesamtbetriebskostenanalyse, Migrationskostenanalyse) sowie Migrationsplanung (Anwendungsdatenanalyse und Datenerfassung, Anwendungsgruppierung, Migrationspriorisierung und Wellenplanung). Das MPA-Tool (Anmeldung erforderlich) steht allen AWS Beratern und APN-Partnerberatern kostenlos zur Verfügung.

Migration Readiness Assessment (MRA)

Der Prozess, bei dem mithilfe des AWS CAF Erkenntnisse über den Cloud-Bereitschaftsstatus eines Unternehmens gewonnen, Stärken und Schwächen identifiziert und ein Aktionsplan zur Schließung festgestellter Lücken erstellt wird. Weitere Informationen finden Sie im Benutzerhandbuch für Migration Readiness. MRA ist die erste Phase der AWS - Migrationsstrategie.

Migrationsstrategie

Der Ansatz, der verwendet wurde, um einen Workload auf den AWS Cloud zu migrieren. Weitere Informationen finden Sie im Eintrag 7 Rs in diesem Glossar und unter Mobilisieren Sie Ihr Unternehmen, um umfangreiche Migrationen zu beschleunigen.

ML

Siehe maschinelles Lernen.

Modernisierung

Umwandlung einer veralteten (veralteten oder monolithischen) Anwendung und ihrer Infrastruktur in ein agiles, elastisches und hochverfügbares System in der Cloud, um Kosten zu senken, die Effizienz zu steigern und Innovationen zu nutzen. Weitere Informationen finden Sie unter Strategie zur Modernisierung von Anwendungen in der AWS Cloud.

Bewertung der Modernisierungsfähigkeit

Eine Bewertung, anhand derer festgestellt werden kann, ob die Anwendungen einer Organisation für die Modernisierung bereit sind, Vorteile, Risiken und Abhängigkeiten identifiziert und ermittelt wird, wie gut die Organisation den zukünftigen Status dieser Anwendungen unterstützen kann. Das Ergebnis der Bewertung ist eine Vorlage der Zielarchitektur, eine Roadmap, in der die Entwicklungsphasen und Meilensteine des Modernisierungsprozesses detailliert beschrieben werden, sowie ein Aktionsplan zur Behebung festgestellter Lücken. Weitere Informationen finden Sie unter Evaluierung der Modernisierungsbereitschaft von Anwendungen in der AWS Cloud.

Monolithische Anwendungen (Monolithen)

Anwendungen, die als ein einziger Service mit eng gekoppelten Prozessen ausgeführt werden. Monolithische Anwendungen haben verschiedene Nachteile. Wenn ein Anwendungs-Feature stark nachgefragt wird, muss die gesamte Architektur skaliert werden. Das Hinzufügen oder Verbessern der Feature einer monolithischen Anwendung wird ebenfalls komplexer, wenn die Codebasis wächst. Um diese Probleme zu beheben, können Sie eine Microservices-Architektur verwenden. Weitere Informationen finden Sie unter Zerlegen von Monolithen in Microservices.

MPA

Siehe Bewertung des Migrationsportfolios.

MQTT

Siehe Message Queuing-Telemetrietransport.

Mehrklassen-Klassifizierung

Ein Prozess, der dabei hilft, Vorhersagen für mehrere Klassen zu generieren (wobei eines von mehr als zwei Ergebnissen vorhergesagt wird). Ein ML-Modell könnte beispielsweise fragen: „Ist dieses Produkt ein Buch, ein Auto oder ein Telefon?“ oder „Welche Kategorie von Produkten ist für diesen Kunden am interessantesten?“

veränderbare Infrastruktur

Ein Modell, das die bestehende Infrastruktur für Produktionsworkloads aktualisiert und modifiziert. Für eine verbesserte Konsistenz, Zuverlässigkeit und Vorhersagbarkeit empfiehlt das AWS Well-Architected Framework die Verwendung einer unveränderlichen Infrastruktur als bewährte Methode.

O

OAC

Siehe Origin Access Control.

EICHE

Siehe Zugriffsidentität von Origin.

COM

Siehe organisatorisches Change-Management.

Offline-Migration

Eine Migrationsmethode, bei der der Quell-Workload während des Migrationsprozesses heruntergefahren wird. Diese Methode ist mit längeren Ausfallzeiten verbunden und wird in der Regel für kleine, unkritische Workloads verwendet.

OI

Siehe Betriebsintegration.

OLA

Siehe Vereinbarung auf operativer Ebene.

Online-Migration

Eine Migrationsmethode, bei der der Quell-Workload auf das Zielsystem kopiert wird, ohne offline genommen zu werden. Anwendungen, die mit dem Workload verbunden sind, können während

der Migration weiterhin funktionieren. Diese Methode beinhaltet keine bis minimale Ausfallzeit und wird in der Regel für kritische Produktionsworkloads verwendet.

OPC-UA

Siehe Open Process Communications — Unified Architecture.

Offene Prozesskommunikation — Einheitliche Architektur (OPC-UA)

Ein machine-to-machine (M2M) -Kommunikationsprotokoll für die industrielle Automatisierung. OPC-UA bietet einen Interoperabilitätsstandard mit Datenverschlüsselungs-, Authentifizierungs- und Autorisierungsschemata.

Vereinbarung auf Betriebsebene (OLA)

Eine Vereinbarung, in der klargestellt wird, welche funktionalen IT-Gruppen sich gegenseitig versprechen zu liefern, um ein Service Level Agreement (SLA) zu unterstützen.

Überprüfung der Betriebsbereitschaft (ORR)

Eine Checkliste mit Fragen und zugehörigen bewährten Methoden, die Ihnen helfen, Vorfälle und mögliche Ausfälle zu verstehen, zu bewerten, zu verhindern oder deren Umfang zu reduzieren. Weitere Informationen finden Sie unter Operational Readiness Reviews (ORR) im AWS Well-Architected Framework.

Betriebstechnologie (OT)

Hardware- und Softwaresysteme, die mit der physischen Umgebung zusammenarbeiten, um industrielle Abläufe, Ausrüstung und Infrastruktur zu steuern. In der Fertigung ist die Integration von OT- und Informationstechnologie (IT) -Systemen ein zentraler Schwerpunkt der Industrie 4.0-Transformationen.

Betriebsintegration (OI)

Der Prozess der Modernisierung von Abläufen in der Cloud, der Bereitschaftsplanung, Automatisierung und Integration umfasst. Weitere Informationen finden Sie im Leitfaden zur Betriebsintegration.

Organisationspfad

Ein Pfad, der von erstellt wird und in AWS CloudTrail dem alle Ereignisse für alle AWS-Konten in einer Organisation protokolliert werden. AWS Organizations Diese Spur wird in jedem AWS-Konto , der Teil der Organisation ist, erstellt und verfolgt die Aktivität in jedem Konto. Weitere Informationen finden Sie in der CloudTrail Dokumentation unter Einen Trail für eine Organisation erstellen.

Organisatorisches Veränderungsmanagement (OCM)

Ein Framework für das Management wichtiger, disruptiver Geschäftstransformationen aus Sicht der Mitarbeiter, der Kultur und der Führung. OCM hilft Organisationen dabei, sich auf neue Systeme und Strategien vorzubereiten und auf diese umzustellen, indem es die Akzeptanz von Veränderungen beschleunigt, Übergangsprobleme angeht und kulturelle und organisatorische Veränderungen vorantreibt. In der AWS Migrationsstrategie wird dieses Framework aufgrund der Geschwindigkeit des Wandels, der bei Projekten zur Cloud-Einführung erforderlich ist, als Mitarbeiterbeschleunigung bezeichnet. Weitere Informationen finden Sie im OCM-Handbuch.

Ursprungszugriffskontrolle (OAC)

In CloudFront, eine erweiterte Option zur Zugriffsbeschränkung, um Ihre Amazon Simple Storage Service (Amazon S3) -Inhalte zu sichern. OAC unterstützt alle S3-Buckets insgesamt AWS-Regionen, serverseitige Verschlüsselung mit AWS KMS (SSE-KMS) sowie dynamische PUT und `DELETE` Anfragen an den S3-Bucket.

Ursprungszugriffsidentität (OAI)

In CloudFront, eine Option zur Zugriffsbeschränkung, um Ihre Amazon S3 S3-Inhalte zu sichern. Wenn Sie OAI verwenden, CloudFront erstellt es einen Principal, mit dem sich Amazon S3 authentifizieren kann. Authentifizierte Principals können nur über eine bestimmte Distribution auf Inhalte in einem S3-Bucket zugreifen. CloudFront Siehe auch OAC, das eine detailliertere und verbesserte Zugriffskontrolle bietet.

ORR

Weitere Informationen finden Sie unter Überprüfung der Betriebsbereitschaft.

NICHT

Siehe Betriebstechnologie.

Ausgehende (egress) VPC

In einer Architektur AWS mit mehreren Konten eine VPC, die Netzwerkverbindungen verarbeitet, die von einer Anwendung aus initiiert werden. Die AWS Security Reference Architecture empfiehlt die Einrichtung Ihres Netzwerkkontos mit eingehendem und ausgehendem Datenverkehr sowie Inspektion, VPCs um die bidirektionale Schnittstelle zwischen Ihrer Anwendung und dem Internet im weiteren Sinne zu schützen.

P

Berechtigungsgrenze

Eine IAM-Verwaltungsrichtlinie, die den IAM-Prinzipalen zugeordnet ist, um die maximalen Berechtigungen festzulegen, die der Benutzer oder die Rolle haben kann. Weitere Informationen finden Sie unter Berechtigungsgrenzen für IAM-Entitys in der IAM-Dokumentation.

persönlich identifizierbare Informationen (PII)

Informationen, die, wenn sie direkt betrachtet oder mit anderen verwandten Daten kombiniert werden, verwendet werden können, um vernünftige Rückschlüsse auf die Identität einer Person zu ziehen. Beispiele für personenbezogene Daten sind Namen, Adressen und Kontaktinformationen.

Personenbezogene Daten

Siehe persönlich identifizierbare Informationen.

Playbook

Eine Reihe vordefinierter Schritte, die die mit Migrationen verbundenen Aufgaben erfassen, z. B. die Bereitstellung zentraler Betriebsfunktionen in der Cloud. Ein Playbook kann die Form von Skripten, automatisierten Runbooks oder einer Zusammenfassung der Prozesse oder Schritte annehmen, die für den Betrieb Ihrer modernisierten Umgebung erforderlich sind.

PLC

Siehe programmierbare Logiksteuerung.

PLM

Siehe Produktlebenszyklusmanagement.

policy

Ein Objekt, das Berechtigungen definieren (siehe identitätsbasierte Richtlinie), Zugriffsbedingungen spezifizieren (siehe ressourcenbasierte Richtlinie) oder die maximalen Berechtigungen für alle Konten in einer Organisation definieren kann AWS Organizations (siehe Dienststeuerungsrichtlinie).

Polyglotte Beharrlichkeit

Unabhängige Auswahl der Datenspeichertechnologie eines Microservices auf der Grundlage von Datenzugriffsmustern und anderen Anforderungen. Wenn Ihre Microservices über dieselbe Datenspeichertechnologie verfügen, kann dies zu Implementierungsproblemen oder zu

Leistungseinbußen führen. Microservices lassen sich leichter implementieren und erzielen eine bessere Leistung und Skalierbarkeit, wenn sie den Datenspeicher verwenden, der ihren Anforderungen am besten entspricht.

Portfoliobewertung

Ein Prozess, bei dem das Anwendungsportfolio ermittelt, analysiert und priorisiert wird, um die Migration zu planen. Weitere Informationen finden Sie in Bewerten der Migrationsbereitschaft.

predicate

Eine Abfragebedingung, die `true` oder zurückgibt`false`, was üblicherweise in einer Klausel vorkommt. `WHERE`

Prädikat Pushdown

Eine Technik zur Optimierung von Datenbankabfragen, bei der die Daten in der Abfrage vor der Übertragung gefiltert werden. Dadurch wird die Datenmenge reduziert, die aus der relationalen Datenbank abgerufen und verarbeitet werden muss, und die Abfrageleistung wird verbessert.

Präventive Kontrolle

Eine Sicherheitskontrolle, die verhindern soll, dass ein Ereignis eintritt. Diese Kontrollen stellen eine erste Verteidigungslinie dar, um unbefugten Zugriff oder unerwünschte Änderungen an Ihrem Netzwerk zu verhindern. Weitere Informationen finden Sie unter Präventive Kontrolle in Implementierung von Sicherheitskontrollen in AWS.

Prinzipal

Eine Entität AWS , die Aktionen ausführen und auf Ressourcen zugreifen kann. Diese Entität ist in der Regel ein Root-Benutzer für eine AWS-Konto, eine IAM-Rolle oder einen Benutzer. Weitere Informationen finden Sie unter Prinzipal in Rollenbegriffe und -konzepte in der IAM-Dokumentation.

Datenschutz von Natur aus

Ein systemtechnischer Ansatz, der den Datenschutz während des gesamten Entwicklungsprozesses berücksichtigt.

Privat gehostete Zonen

Ein Container, der Informationen darüber enthält, wie Amazon Route 53 auf DNS-Abfragen für eine Domain und deren Subdomains innerhalb einer oder mehrerer VPCs Domains antworten soll. Weitere Informationen finden Sie unter Arbeiten mit privat gehosteten Zonen in der Route-53-Dokumentation.

proaktive Steuerung

Eine Sicherheitskontrolle, die den Einsatz nicht richtlinienkonformer Ressourcen verhindern soll. Diese Steuerelemente scannen Ressourcen, bevor sie bereitgestellt werden. Wenn die Ressource nicht der Kontrolle entspricht, wird sie nicht bereitgestellt. Weitere Informationen finden Sie im Referenzhandbuch zu Kontrollen in der AWS Control Tower Dokumentation und unter Proaktive Kontrollen unter Implementierung von Sicherheitskontrollen am AWS.

Produktlebenszyklusmanagement (PLM)

Das Management von Daten und Prozessen für ein Produkt während seines gesamten Lebenszyklus, vom Design, der Entwicklung und Markteinführung über Wachstum und Reife bis hin zur Markteinführung und Markteinführung.

Produktionsumgebung

Siehe Umgebung.

Speicherprogrammierbare Steuerung (SPS)

In der Fertigung ein äußerst zuverlässiger, anpassungsfähiger Computer, der Maschinen überwacht und Fertigungsprozesse automatisiert.

schnelle Verkettung

Verwendung der Ausgabe einer LLM-Eingabeaufforderung als Eingabe für die nächste Aufforderung, um bessere Antworten zu generieren. Diese Technik wird verwendet, um eine komplexe Aufgabe in Unteraufgaben zu unterteilen oder um eine vorläufige Antwort iterativ zu verfeinern oder zu erweitern. Sie trägt dazu bei, die Genauigkeit und Relevanz der Antworten eines Modells zu verbessern und ermöglicht detailliertere, personalisierte Ergebnisse.

Pseudonymisierung

Der Prozess, bei dem persönliche Identifikatoren in einem Datensatz durch Platzhalterwerte ersetzt werden. Pseudonymisierung kann zum Schutz der Privatsphäre beitragen. Pseudonymisierte Daten gelten weiterhin als personenbezogene Daten.

publish/subscribe (pub/sub)

Ein Muster, das asynchrone Kommunikation zwischen Microservices ermöglicht, um die Skalierbarkeit und Reaktionsfähigkeit zu verbessern. In einem auf Microservices basierenden MES kann ein Microservice beispielsweise Ereignismeldungen in einem Kanal veröffentlichen, den andere Microservices abonnieren können. Das System kann neue Microservices hinzufügen, ohne den Veröffentlichungsservice zu ändern.

Q

Abfrageplan

Eine Reihe von Schritten, wie Anweisungen, die für den Zugriff auf die Daten in einem relationalen SQL-Datenbanksystem verwendet werden.

Abfrageplanregression

Wenn ein Datenbankserviceoptimierer einen weniger optimalen Plan wählt als vor einer bestimmten Änderung der Datenbankumgebung. Dies kann durch Änderungen an Statistiken, Beschränkungen, Umgebungseinstellungen, Abfrageparameter-Bindungen und Aktualisierungen der Datenbank-Engine verursacht werden.

R

RACI-Matrix

Siehe verantwortlich, rechenschaftspflichtig, konsultiert, informiert (RACI).

RAG

Siehe Erweiterte Generierung beim Abrufen.

Ransomware

Eine bösartige Software, die entwickelt wurde, um den Zugriff auf ein Computersystem oder Daten zu blockieren, bis eine Zahlung erfolgt ist.

RASCI-Matrix

Siehe verantwortlich, rechenschaftspflichtig, konsultiert, informiert (RACI).

RCAC

Siehe Zugriffskontrolle für Zeilen und Spalten.

Read Replica

Eine Kopie einer Datenbank, die nur für Lesezwecke verwendet wird. Sie können Abfragen an das Lesereplikat weiterleiten, um die Belastung auf Ihrer Primärdatenbank zu reduzieren.

neu strukturieren

Siehe 7 Rs.

Recovery Point Objective (RPO)

Die maximal zulässige Zeitspanne seit dem letzten Datenwiederherstellungspunkt. Damit wird festgelegt, was als akzeptabler Datenverlust zwischen dem letzten Wiederherstellungspunkt und der Serviceunterbrechung gilt.

Wiederherstellungszeitziel (RTO)

Die maximal zulässige Verzögerung zwischen der Betriebsunterbrechung und der Wiederherstellung des Dienstes.

Refaktorierung

Siehe 7 Rs.

Region

Eine Sammlung von AWS Ressourcen in einem geografischen Gebiet. Jeder AWS-Region ist isoliert und unabhängig von den anderen, um Fehlertoleranz, Stabilität und Belastbarkeit zu gewährleisten. Weitere Informationen finden Sie unter Geben Sie an, was AWS-Regionen Ihr Konto verwenden kann.

Regression

Eine ML-Technik, die einen numerischen Wert vorhersagt. Zum Beispiel, um das Problem „Zu welchem Preis wird dieses Haus verkauft werden?“ zu lösen Ein ML-Modell könnte ein lineares Regressionsmodell verwenden, um den Verkaufspreis eines Hauses auf der Grundlage bekannter Fakten über das Haus (z. B. die Quadratmeterzahl) vorherzusagen.

rehosten

Siehe 7 Rs.

Veröffentlichung

In einem Bereitstellungsprozess der Akt der Förderung von Änderungen an einer Produktionsumgebung.

umziehen

Siehe 7 Rs.

neue Plattform

Siehe 7 Rs.

Rückkauf

Siehe 7 Rs.

Ausfallsicherheit

Die Fähigkeit einer Anwendung, Störungen zu widerstehen oder sich von ihnen zu erholen. Hochverfügbarkeit und Notfallwiederherstellung sind häufig Überlegungen bei der Planung der Ausfallsicherheit in der. AWS Cloud Weitere Informationen finden Sie unter AWS Cloud Resilienz.

Ressourcenbasierte Richtlinie

Eine mit einer Ressource verknüpfte Richtlinie, z. B. ein Amazon-S3-Bucket, ein Endpunkt oder ein Verschlüsselungsschlüssel. Diese Art von Richtlinie legt fest, welchen Prinzipalen der Zugriff gewährt wird, welche Aktionen unterstützt werden und welche anderen Bedingungen erfüllt sein müssen.

RACI-Matrix (verantwortlich, rechenschaftspflichtig, konsultiert, informiert)

Eine Matrix, die die Rollen und Verantwortlichkeiten aller an Migrationsaktivitäten und Cloud-Operationen beteiligten Parteien definiert. Der Matrixname leitet sich von den in der Matrix definierten Zuständigkeitstypen ab: verantwortlich (R), rechenschaftspflichtig (A), konsultiert (C) und informiert (I). Der Unterstützungstyp (S) ist optional. Wenn Sie Unterstützung einbeziehen, wird die Matrix als RASCI-Matrix bezeichnet, und wenn Sie sie ausschließen, wird sie als RACI-Matrix bezeichnet.

Reaktive Kontrolle

Eine Sicherheitskontrolle, die darauf ausgelegt ist, die Behebung unerwünschter Ereignisse oder Abweichungen von Ihren Sicherheitsstandards voranzutreiben. Weitere Informationen finden Sie unter Reaktive Kontrolle in Implementieren von Sicherheitskontrollen in AWS.

Beibehaltung

Siehe 7 Rs.

zurückziehen

Siehe 7 Rs.

Retrieval Augmented Generation (RAG)

Eine generative KI-Technologie, bei der ein LLM auf eine maßgebliche Datenquelle verweist, die sich außerhalb seiner Trainingsdatenquellen befindet, bevor eine Antwort generiert wird. Ein

RAG-Modell könnte beispielsweise eine semantische Suche in der Wissensdatenbank oder in benutzerdefinierten Daten einer Organisation durchführen. Weitere Informationen finden Sie unter Was ist RAG.

Drehung

Der Vorgang, bei dem ein Geheimnis regelmäßig aktualisiert wird, um es einem Angreifer zu erschweren, auf die Anmeldeinformationen zuzugreifen.

Zugriffskontrolle für Zeilen und Spalten (RCAC)

Die Verwendung einfacher, flexibler SQL-Ausdrücke mit definierten Zugriffsregeln. RCAC besteht aus Zeilenberechtigungen und Spaltenmasken.

RPO

Siehe Recovery Point Objective.

RTO

Siehe Ziel für die Erholungszeit.

Runbook

Eine Reihe manueller oder automatisierter Verfahren, die zur Ausführung einer bestimmten Aufgabe erforderlich sind. Diese sind in der Regel darauf ausgelegt, sich wiederholende Operationen oder Verfahren mit hohen Fehlerquoten zu rationalisieren.

S

SAML 2.0

Ein offener Standard, den viele Identitätsanbieter (IdPs) verwenden. Diese Funktion ermöglicht föderiertes Single Sign-On (SSO), sodass sich Benutzer bei den API-Vorgängen anmelden AWS-Managementkonsole oder die AWS API-Operationen aufrufen können, ohne dass Sie einen Benutzer in IAM für alle in Ihrer Organisation erstellen müssen. Weitere Informationen zum SAML-2.0.-basierten Verbund finden Sie unter Über den SAML-2.0-basierten Verbund in der IAM-Dokumentation.

SCADA

Siehe Aufsichtskontrolle und Datenerfassung.

SCP

Siehe Richtlinie zur Dienstkontrolle.

Secret

Interne AWS Secrets Manager, vertrauliche oder eingeschränkte Informationen, wie z. B. ein Passwort oder Benutzeranmeldeinformationen, die Sie in verschlüsselter Form speichern. Es besteht aus dem geheimen Wert und seinen Metadaten. Der geheime Wert kann binär, eine einzelne Zeichenfolge oder mehrere Zeichenketten sein. Weitere Informationen finden Sie unter Was ist in einem Secrets Manager Manager-Geheimnis? in der Secrets Manager Manager-Dokumentation.

Sicherheit durch Design

Ein systemtechnischer Ansatz, der die Sicherheit während des gesamten Entwicklungsprozesses berücksichtigt.

Sicherheitskontrolle

Ein technischer oder administrativer Integritätsschutz, der die Fähigkeit eines Bedrohungsakteurs, eine Schwachstelle auszunutzen, verhindert, erkennt oder einschränkt. Es gibt vier Haupttypen von Sicherheitskontrollen: präventiv, detektiv, reaktionsschnell und proaktiv.

Härtung der Sicherheit

Der Prozess, bei dem die Angriffsfläche reduziert wird, um sie widerstandsfähiger gegen Angriffe zu machen. Dies kann Aktionen wie das Entfernen von Ressourcen, die nicht mehr benötigt werden, die Implementierung der bewährten Sicherheitsmethode der Gewährung geringster Berechtigungen oder die Deaktivierung unnötiger Feature in Konfigurationsdateien umfassen.

System zur Verwaltung von Sicherheitsinformationen und Ereignissen (security information and event management – SIEM)

Tools und Services, die Systeme für das Sicherheitsinformationsmanagement (SIM) und das Management von Sicherheitsereignissen (SEM) kombinieren. Ein SIEM-System sammelt, überwacht und analysiert Daten von Servern, Netzwerken, Geräten und anderen Quellen, um Bedrohungen und Sicherheitsverletzungen zu erkennen und Warnmeldungen zu generieren.

Automatisierung von Sicherheitsreaktionen

Eine vordefinierte und programmierte Aktion, die darauf ausgelegt ist, automatisch auf ein Sicherheitsereignis zu reagieren oder es zu beheben. Diese Automatisierungen dienen als detektive oder reaktionsschnelle Sicherheitskontrollen, die Sie bei der Implementierung bewährter

AWS Sicherheitsmethoden unterstützen. Beispiele für automatisierte Antwortaktionen sind das Ändern einer VPC-Sicherheitsgruppe, das Patchen einer Amazon EC2 EC2-Instance oder das Rotieren von Anmeldeinformationen.

Serverseitige Verschlüsselung

Verschlüsselung von Daten am Zielort durch denjenigen AWS-Service , der sie empfängt.

Service-Kontrollrichtlinie (SCP)

Eine Richtlinie, die eine zentrale Steuerung der Berechtigungen für alle Konten in einer Organisation in ermöglicht AWS Organizations. SCPs Definieren Sie Leitplanken oder legen Sie Grenzwerte für Aktionen fest, die ein Administrator an Benutzer oder Rollen delegieren kann. Sie können sie SCPs als Zulassungs- oder Ablehnungslisten verwenden, um festzulegen, welche Dienste oder Aktionen zulässig oder verboten sind. Weitere Informationen finden Sie in der AWS Organizations Dokumentation unter Richtlinien zur Dienststeuerung.

Service-Endpunkt

Die URL des Einstiegspunkts für einen AWS-Service. Sie können den Endpunkt verwenden, um programmgesteuert eine Verbindung zum Zielservice herzustellen. Weitere Informationen finden Sie unter AWS-Service -Endpunkte in der Allgemeine AWS-Referenz.

Service Level Agreement (SLA)

Eine Vereinbarung, in der klargestellt wird, was ein IT-Team seinen Kunden zu bieten verspricht, z. B. in Bezug auf Verfügbarkeit und Leistung der Services.

Service-Level-Indikator (SLI)

Eine Messung eines Leistungsaspekts eines Dienstes, z. B. seiner Fehlerrate, Verfügbarkeit oder Durchsatz.

Service-Level-Ziel (SLO)

Eine Zielkennzahl, die den Zustand eines Dienstes darstellt, gemessen anhand eines Service-Level-Indikators.

Modell der geteilten Verantwortung

Ein Modell, das die Verantwortung beschreibt, mit der Sie gemeinsam AWS für Cloud-Sicherheit und Compliance verantwortlich sind. AWS ist für die Sicherheit der Cloud verantwortlich, während Sie für die Sicherheit in der Cloud verantwortlich sind. Weitere Informationen finden Sie unter Modell der geteilten Verantwortung.

SIEM

Siehe Sicherheitsinformations- und Event-Management-System.

Single Point of Failure (SPOF)

Ein Fehler in einer einzelnen, kritischen Komponente einer Anwendung, der das System stören kann.

SLA

Siehe Service Level Agreement.

SLI

Siehe Service-Level-Indikator.

ALSO

Siehe Service-Level-Ziel.

split-and-seed Modell

Ein Muster für die Skalierung und Beschleunigung von Modernisierungsprojekten. Sobald neue Features und Produktversionen definiert werden, teilt sich das Kernteam auf, um neue Produktteams zu bilden. Dies trägt zur Skalierung der Fähigkeiten und Services Ihrer Organisation bei, verbessert die Produktivität der Entwickler und unterstützt schnelle Innovationen. Weitere Informationen finden Sie unter Schrittweiser Ansatz zur Modernisierung von Anwendungen in der AWS Cloud.

SPOTTEN

Siehe Single Point of Failure.

Sternschema

Eine Datenbank-Organisationsstruktur, die eine große Faktentabelle zum Speichern von Transaktions- oder Messdaten und eine oder mehrere kleinere dimensionale Tabellen zum Speichern von Datenattributen verwendet. Diese Struktur ist für die Verwendung in einem Data Warehouse oder für Business Intelligence-Zwecke konzipiert.

Strangler-Fig-Muster

Ein Ansatz zur Modernisierung monolithischer Systeme, bei dem die Systemfunktionen schrittweise umgeschrieben und ersetzt werden, bis das Legacy-System außer Betrieb

genommen werden kann. Dieses Muster verwendet die Analogie einer Feigenrebe, die zu einem etablierten Baum heranwächst und schließlich ihren Wirt überwindet und ersetzt. Das Muster wurde eingeführt von Martin Fowler als Möglichkeit, Risiken beim Umschreiben monolithischer Systeme zu managen. Ein Beispiel für die Anwendung dieses Musters finden Sie unter Schrittweises Modernisieren älterer Microsoft ASP.NET (ASMX)-Webservices mithilfe von Containern und Amazon API Gateway.

Subnetz

Ein Bereich von IP-Adressen in Ihrer VPC. Ein Subnetz muss sich in einer einzigen Availability Zone befinden.

Aufsichtskontrolle und Datenerfassung (SCADA)

In der Fertigung ein System, das Hardware und Software zur Überwachung von Sachanlagen und Produktionsabläufen verwendet.

Symmetrische Verschlüsselung

Ein Verschlüsselungsalgorithmus, der denselben Schlüssel zum Verschlüsseln und Entschlüsseln der Daten verwendet.

synthetisches Testen

Testen eines Systems auf eine Weise, die Benutzerinteraktionen simuliert, um potenzielle Probleme zu erkennen oder die Leistung zu überwachen. Sie können Amazon CloudWatch Synthetics verwenden, um diese Tests zu erstellen.

Systemaufforderung

Eine Technik, mit der einem LLM Kontext, Anweisungen oder Richtlinien zur Verfügung gestellt werden, um sein Verhalten zu steuern. Systemaufforderungen helfen dabei, den Kontext festzulegen und Regeln für Interaktionen mit Benutzern festzulegen.

T

tags

Schlüssel-Wert-Paare, die als Metadaten für die Organisation Ihrer Ressourcen dienen. AWS Mit Tags können Sie Ressourcen verwalten, identifizieren, organisieren, suchen und filtern. Weitere Informationen finden Sie unter Markieren Ihrer AWS -Ressourcen.

Zielvariable

Der Wert, den Sie in überwachtem ML vorhersagen möchten. Dies wird auch als Ergebnisvariable bezeichnet. In einer Fertigungsumgebung könnte die Zielvariable beispielsweise ein Produktfehler sein.

Aufgabenliste

Ein Tool, das verwendet wird, um den Fortschritt anhand eines Runbooks zu verfolgen. Eine Aufgabenliste enthält eine Übersicht über das Runbook und eine Liste mit allgemeinen Aufgaben, die erledigt werden müssen. Für jede allgemeine Aufgabe werden der geschätzte Zeitaufwand, der Eigentümer und der Fortschritt angegeben.

Testumgebungen

Siehe Umgebung.

Training

Daten für Ihr ML-Modell bereitstellen, aus denen es lernen kann. Die Trainingsdaten müssen die richtige Antwort enthalten. Der Lernalgorithmus findet Muster in den Trainingsdaten, die die Attribute der Input-Daten dem Ziel (die Antwort, die Sie voraussagen möchten) zuordnen. Es gibt ein ML-Modell aus, das diese Muster erfasst. Sie können dann das ML-Modell verwenden, um Voraussagen für neue Daten zu erhalten, bei denen Sie das Ziel nicht kennen.

Transit-Gateway

Ein Netzwerk-Transit-Hub, über den Sie Ihre Netzwerke VPCs und Ihre lokalen Netzwerke miteinander verbinden können. Weitere Informationen finden Sie in der Dokumentation unter Was ist ein Transit-Gateway. AWS Transit Gateway

Stammbasierter Workflow

Ein Ansatz, bei dem Entwickler Feature lokal in einem Feature-Zweig erstellen und testen und diese Änderungen dann im Hauptzweig zusammenführen. Der Hauptzweig wird dann sequentiell für die Entwicklungs-, Vorproduktions- und Produktionsumgebungen erstellt.

Vertrauenswürdiger Zugriff

Gewährung von Berechtigungen für einen Dienst, den Sie angeben, um Aufgaben in Ihrer Organisation AWS Organizations und in deren Konten in Ihrem Namen auszuführen. Der vertrauenswürdige Service erstellt in jedem Konto eine mit dem Service verknüpfte Rolle, wenn diese Rolle benötigt wird, um Verwaltungsaufgaben für Sie auszuführen. Weitere Informationen finden Sie in der AWS Organizations Dokumentation unter Verwendung AWS Organizations mit anderen AWS Diensten.

Optimieren

Aspekte Ihres Trainingsprozesses ändern, um die Genauigkeit des ML-Modells zu verbessern. Sie können das ML-Modell z. B. trainieren, indem Sie einen Beschriftungssatz generieren, Beschriftungen hinzufügen und diese Schritte dann mehrmals unter verschiedenen Einstellungen wiederholen, um das Modell zu optimieren.

Zwei-Pizzen-Team

Ein kleines DevOps Team, das Sie mit zwei Pizzen ernähren können. Eine Teamgröße von zwei Pizzen gewährleistet die bestmögliche Gelegenheit zur Zusammenarbeit bei der Softwareentwicklung.

U

Unsicherheit

Ein Konzept, das sich auf ungenaue, unvollständige oder unbekannte Informationen bezieht, die die Zuverlässigkeit von prädiktiven ML-Modellen untergraben können. Es gibt zwei Arten von Unsicherheit: Epistemische Unsicherheit wird durch begrenzte, unvollständige Daten verursacht, wohingegen aleatorische Unsicherheit durch Rauschen und Randomisierung verursacht wird, die in den Daten liegt.

undifferenzierte Aufgaben

Diese Arbeit wird auch als Schwerstarbeit bezeichnet. Dabei handelt es sich um Arbeiten, die zwar für die Erstellung und den Betrieb einer Anwendung erforderlich sind, aber dem Endbenutzer keinen direkten Mehrwert bieten oder keinen Wettbewerbsvorteil bieten. Beispiele für undifferenzierte Aufgaben sind Beschaffung, Wartung und Kapazitätsplanung.

höhere Umgebungen

Siehe Umgebung.

V

Vacuuming

Ein Vorgang zur Datenbankwartung, bei dem die Datenbank nach inkrementellen Aktualisierungen bereinigt wird, um Speicherplatz zurückzugewinnen und die Leistung zu verbessern.

Versionskontrolle

Prozesse und Tools zur Nachverfolgung von Änderungen, z. B. Änderungen am Quellcode in einem Repository.

VPC-Peering

Eine Verbindung zwischen zwei VPCs , die es Ihnen ermöglicht, den Verkehr mithilfe privater IP-Adressen weiterzuleiten. Weitere Informationen finden Sie unter Was ist VPC-Peering? in der Amazon-VPC-Dokumentation.

Schwachstelle

Ein Software- oder Hardwarefehler, der die Sicherheit des Systems beeinträchtigt.

W

Warmer Cache

Ein Puffer-Cache, der aktuelle, relevante Daten enthält, auf die häufig zugegriffen wird. Die Datenbank-Instance kann aus dem Puffer-Cache lesen, was schneller ist als das Lesen aus dem Hauptspeicher oder von der Festplatte.

warme Daten

Daten, auf die selten zugegriffen wird. Bei der Abfrage dieser Art von Daten sind mäßig langsame Abfragen in der Regel akzeptabel.

Fensterfunktion

Eine SQL-Funktion, die eine Berechnung für eine Gruppe von Zeilen durchführt, die sich in irgendeiner Weise auf den aktuellen Datensatz beziehen. Fensterfunktionen sind nützlich für die Verarbeitung von Aufgaben wie die Berechnung eines gleitenden Durchschnitts oder für den Zugriff auf den Wert von Zeilen auf der Grundlage der relativen Position der aktuellen Zeile.

Workload

Ein Workload ist eine Sammlung von Ressourcen und Code, die einen Unternehmenswert bietet, wie z. B. eine kundenorientierte Anwendung oder ein Backend-Prozess.

Workstream

Funktionsgruppen in einem Migrationsprojekt, die für eine bestimmte Reihe von Aufgaben verantwortlich sind. Jeder Workstream ist unabhängig, unterstützt aber die anderen Workstreams

im Projekt. Der Portfolio-Workstream ist beispielsweise für die Priorisierung von Anwendungen, die Wellenplanung und die Erfassung von Migrationsmetadaten verantwortlich. Der Portfolio-Workstream liefert diese Komponenten an den Migrations-Workstream, der dann die Server und Anwendungen migriert.

WURM

Sehen Sie einmal schreiben, viele lesen.

WQF

Siehe AWS Workload-Qualifizierungsrahmen.

einmal schreiben, viele lesen (WORM)

Ein Speichermodell, das Daten ein einziges Mal schreibt und verhindert, dass die Daten gelöscht oder geändert werden. Autorisierte Benutzer können die Daten so oft wie nötig lesen, aber sie können sie nicht ändern. Diese Datenspeicherinfrastruktur gilt als unveränderlich.

Z

Zero-Day-Exploit

Ein Angriff, in der Regel Malware, der eine Zero-Day-Sicherheitslücke ausnutzt.

Zero-Day-Sicherheitslücke

Ein unfehlbarer Fehler oder eine Sicherheitslücke in einem Produktionssystem. Bedrohungsakteure können diese Art von Sicherheitslücke nutzen, um das System anzugreifen. Entwickler werden aufgrund des Angriffs häufig auf die Sicherheitsanfälligkeit aufmerksam.

Eingabeaufforderung ohne Zwischenfälle

Bereitstellung von Anweisungen für die Ausführung einer Aufgabe an einen LLM, jedoch ohne Beispiele (Schnappschüsse), die ihm als Orientierungshilfe dienen könnten. Der LLM muss sein vortrainiertes Wissen einsetzen, um die Aufgabe zu bewältigen. Die Effektivität von Zero-Shot Prompting hängt von der Komplexität der Aufgabe und der Qualität der Aufforderung ab. Siehe auch Few-Shot-Prompting.

Zombie-Anwendung

Eine Anwendung, deren durchschnittliche CPU- und Arbeitsspeichernutzung unter 5 Prozent liegt. In einem Migrationsprojekt ist es üblich, diese Anwendungen außer Betrieb zu nehmen.

Die vorliegende Übersetzung wurde maschinell erstellt. Im Falle eines Konflikts oder eines Widerspruchs zwischen dieser übersetzten Fassung und der englischen Fassung (einschließlich infolge von Verzögerungen bei der Übersetzung) ist die englische Fassung maßgeblich.