Quadro di analisi della resilienza

AWS Guida prescrittiva

AWS Guida prescrittiva: Quadro di analisi della resilienza

Table of Contents

Framework di analisi della resilienza

John Formento, Bruno Emer, Steven Hooper, Jason Barto e Michael Haken, Amazon Web Services (AWS)

settembre 2023(cronologia dei documenti)

Standard e processi coerenti e ripetibili sono una parte importante del miglioramento continuo. Questo vale anche per la resilienza dei sistemi distribuiti. Lo scopo di questa guida è introdurre un framework di analisi della resilienza che fornisca un modo coerente per analizzare le modalità di errore e il modo in cui potrebbero influire sui carichi di lavoro. L'utilizzo di questo framework per l'intero ciclo di vita del carico di lavoro, dalla progettazione all'esercizio, consente di migliorare continuamente la resilienza dei carichi di lavoro a una gamma più ampia di potenziali modalità di guasto in modo coerente e ripetibile. Questo aiuta a garantire il raggiungimento degli obiettivi di resilienza e il mantenimento delle proprietà di resilienza desiderate per i carichi di lavoro.

Questo framework è stato sviluppato grazie all'esperienza dei team sul campo dell'architettura delle soluzioni AWS nel loro lavoro con clienti di tutti i settori. Si rivolge a costruttori che possono ricoprire diversi ruoli, tra cui responsabili di prodotto, sviluppatori di software, ingegneri di sistema, team operativi e architetti. Queste sono le persone che conoscono meglio il sistema, il servizio o il prodotto che viene analizzato. L'uso del framework in esercizi continui può aiutarti a fare progressi incrementali e a raggiungere i tuoi obiettivi di resilienza a lungo termine.

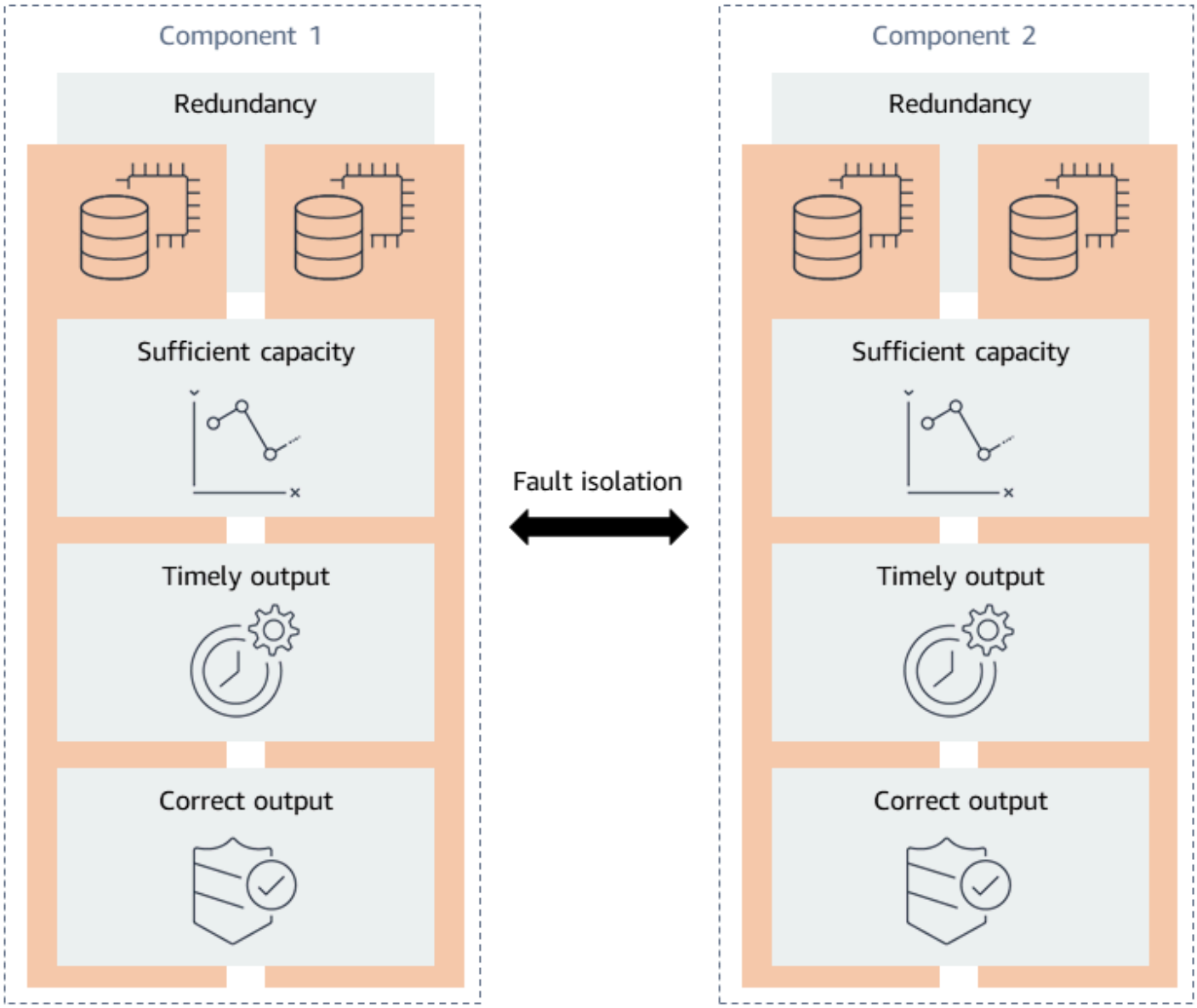
L'obiettivo del framework è identificare le potenziali modalità di guasto e i controlli preventivi e correttivi che è possibile utilizzare per mitigarne l'impatto. Anche se i guasti si verificano in componenti che non sono direttamente sotto il controllo dell'utente, ad esempio un aumento dei tassi di errore in una dipendenza, è necessario considerare in che modo tali guasti potrebbero influire sul carico di lavoro e come progettare tale carico di lavoro in modo da rispondere a tali errori. In definitiva, dovreste concentrarvi sufallimenti a cui puoi rispondereutilizzando una mitigazione che è sotto il tuo controllo.

Questa guida delinea il framework e poi spiega come identificare e documentare un carico di lavoro, come applicare il framework a quel carico di lavoro e come valutare le strategie di mitigazione per eventuali guasti riscontrati.

Indice

Panoramica del framework

Il framework di analisi della resilienza è stato sviluppato identificando le proprietà di resilienza desiderate di un carico di lavoro. Le proprietà desiderate sono le cose che volete che siano vere riguardo al sistema. La resilienza viene generalmente misurata in base alla disponibilità, quindi cinque proprietà sono le caratteristiche di un sistema distribuito ad alta disponibilità: ridondanza, capacità sufficiente, output tempestivo, output corretto e isolamento dai guasti. Queste proprietà sono illustrate nel diagramma seguente.



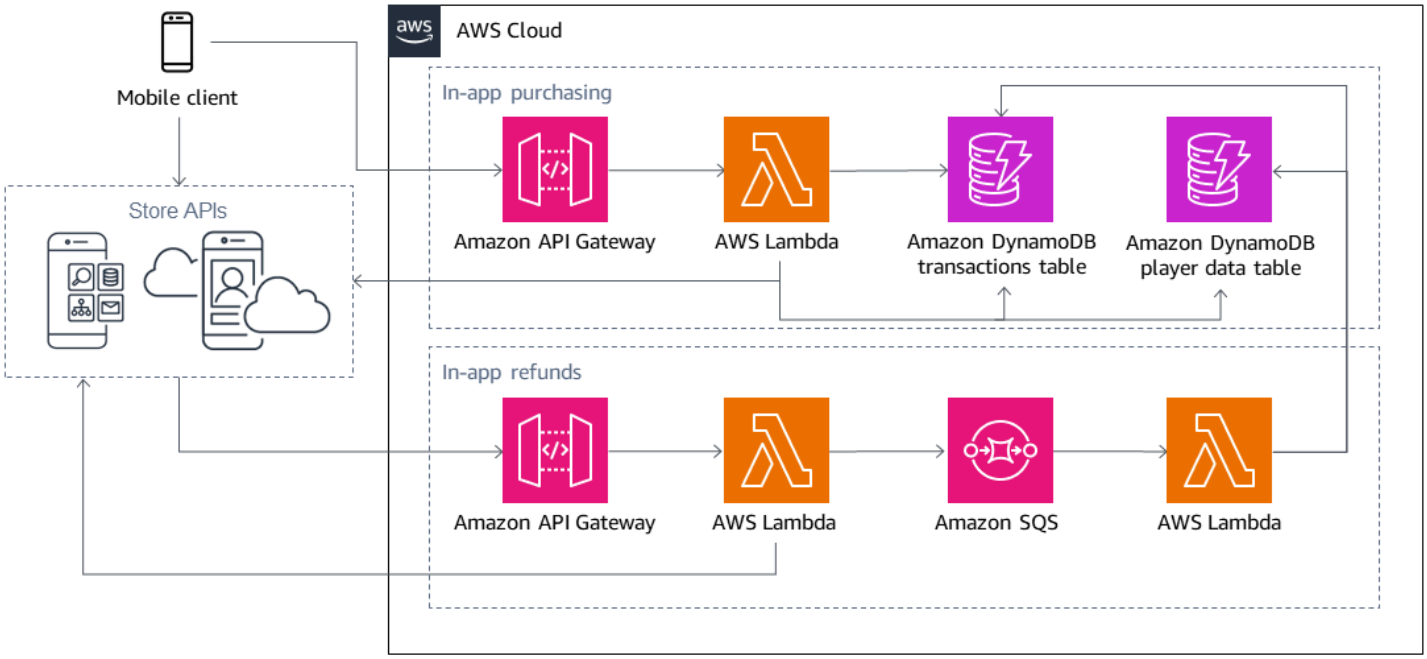
- Ridondanza— La tolleranza ai guasti si ottiene attraverso la ridondanza che elimina i singoli punti di errore (SPOF). La ridondanza può spaziare dai componenti di ricambio del carico di lavoro alle repliche complete dell'intero stack di applicazioni. Quando si considera la ridondanza delle applicazioni, è importante tenere conto del livello di ridondanza fornito dall'infrastruttura, dagli archivi di dati e dalle dipendenze utilizzate. Ad esempio, Amazon DynamoDB e Amazon Simple Storage Service (Amazon S3) forniscono ridondanza replicando i dati su più zone di disponibilità in una regione eAWS Lambdaesegue le tue funzioni su più nodi di lavoro in più zone di disponibilità. Per ogni servizio che utilizzi, prendi in considerazione ciò che viene fornito dal servizio e per cosa devi progettare.
- Capacità sufficiente— Il carico di lavoro richiede risorse sufficienti per funzionare come previsto. Le risorse includono memoria, cicli di CPU, thread, storage, velocità effettiva, quote di servizio e molte altre.
- Uscita tempestiva— Quando i clienti utilizzano il tuo carico di lavoro, si aspettano che svolga la funzione prevista entro un periodo di tempo ragionevole. A meno che il servizio non fornisca un accordo sul livello di servizio (SLA) per la latenza, le loro aspettative si basano generalmente su prove empiriche, vale a dire sulla loro esperienza. Questoesperienza media del clienteè generalmente considerata la latenza mediana (P50) del sistema. Se il carico di lavoro richiede più tempo del previsto, questa latenza può influire sull'esperienza dei clienti.
- Output corretto— L'output corretto del software del carico di lavoro è necessario affinché quest'ultimo fornisca le funzionalità previste. Un risultato errato o incompleto può essere peggio di nessuna risposta.
- Isolamento degli errori— L'isolamento dei guasti limita l'ambito di impatto al contenitore di guasto previsto quando si verifica un guasto. Garantisce che componenti specifici del carico di lavoro si guastino insieme, impedendo al contempo che un guasto si ripercuota a cascata su altri componenti indesiderati. Inoltre, aiuta a limitare l'ambito di impatto del carico di lavoro sui clienti. L'isolamento dei guasti è leggermente diverso dalle quattro proprietà precedenti, poiché accetta che si sia già verificato un errore ma deve essere contenuto. È possibile creare l'isolamento degli errori nell'infrastruttura, nelle dipendenze e nelle funzioni software.

La violazione di una proprietà desiderata può causare l'indisponibilità, o la percezione, dell'indisponibilità di un carico di lavoro. Sulla base di queste proprietà di resilienza desiderate e della nostra esperienza di lavoro con moltiAWSclienti, abbiamo identificato cinque categorie di errori comuni: singoli punti di errore, carico eccessivo, latenza eccessiva, configurazioni errate e bug e destino condiviso, che abbreviiamo in SEEMS. Questi forniscono un metodo coerente per classificare le potenziali modalità di errore e sono descritti nella tabella seguente.

Categoria di errore	Viola	Definizione
Singoli punti di guasto (SPOF)	Ridondanza	Un guasto in un singolo componente danneggia il sistema a causa della mancanza di ridondanza del componente.
Carico eccessivo	Capacità sufficiente	Il consumo eccessivo di una risorsa dovuto a una domanda o a un traffico eccessivi impedisce alla risorsa di svolgere la funzione prevista. Ciò può includere il raggiungi mento di limiti e quote, che causano la limitazione e il rifiuto delle richieste.
Latenza eccessiva	Uscita tempestiva	L'elaborazione del sistema o la latenza del traffico di rete superano il tempo, gli obiettivi a livello di servizio (SLO) o gli accordi sui livelli di servizio (SLA) previsti.
Errori di configurazione e bug	Output corretto	I bug del software o l'errata configurazione del sistema portano a un output errato.
Destino condiviso	Isolamento degli errori	Un guasto causato da una delle precedenti categorie di guasto supera i limiti di isolamento dei guasti previsti e si ripercuote a cascata su altre parti del sistema o su altri clienti.

Comprendere il carico di lavoro

Per applicare il framework, inizia a comprendere il carico di lavoro che desideri analizzare. Un diagramma dell'architettura del sistema fornisce un punto di partenza per documentare i dettagli più rilevanti del sistema. Tuttavia, cercare di analizzare un intero carico di lavoro può essere complesso, poiché molti sistemi hanno numerosi componenti e interazioni. Ti consigliamo invece di concentrarti sustorie degli utenti, quali sonospiegazioni informali e generali delle funzionalità del software scritte dal punto di vista dell'utente finale. Il loro scopo è spiegare in che modo una funzionalità software fornisce valore al cliente. È quindi possibile modellare queste storie utente con diagrammi di architettura e diagrammi di flusso di dati per facilitare la valutazione dei componenti tecnici che forniscono le funzionalità aziendali descritte. Ad esempio, una soluzione di acquisto di giochi per dispositivi mobili in-app potrebbe avere due storie utente, «acquisto di crediti in-app» e «ottenimento di rimborsi in-app», come illustrato nel diagramma seguente. (Questa architettura di esempio evidenzia come è possibile scomporre un sistema in storie utente; non è pensata per rappresentare un'applicazione altamente resiliente).



Ogni storia utente è composta da quattro componenti comuni: codice e configurazione, infrastruttura, archivi dati e dipendenze esterne. I diagrammi devono includere tutti questi componenti e riflettere le interazioni tra i componenti. Ad esempio, se il carico sull'endpoint Amazon API Gateway è eccessivo, considera in che modo tale carico si ripercuote su altri componenti del sistema, come ilAWS Lambdafunzioni o tabelle Amazon DynamoDB. Il monitoraggio di queste interazioni ti aiuta a capire in che modo la modalità di errore può influire sulla storia dell'utente. È possibile acquisire

questo flusso visivamente con un diagramma di flusso di dati o utilizzando semplici frecce di flusso in un diagramma di architettura, come nell'illustrazione precedente. Per ogni componente, prendete in considerazione l'acquisizione di dettagli come il tipo di informazioni che vengono trasmesse, le informazioni ricevute, se la comunicazione è sincrona o asincrona e quali limiti di errore vengono superati. Nell'esempio, le tabelle DynamoDB sono condivise in entrambe le storie utente, come si può vedere dalle frecce che indicano che il componente Lambda nella storia dei rimborsi in-app accede alle tabelle DynamoDB nella storia degli acquisti in-app. Ciò significa che un errore causato dalla storia utente relativa agli acquisti in-app potrebbe ripercuotersi a cascata sulla storia utente relativa ai rimborsi in-app come conseguenza di un destino condiviso.

Inoltre, è importante comprendere la configurazione di base per ogni componente. La configurazione di base identifica vincoli quali il numero medio e massimo di transazioni al secondo, la dimensione massima di un payload, il timeout del client e le quote di servizio predefinite o correnti per la risorsa. Se state modellando un nuovo progetto, vi consigliamo di documentarne i requisiti funzionali e di considerarne i limiti. Questo aiuta a capire come potrebbero manifestarsi le modalità di guasto nel componente.

Infine, è necessario dare priorità alle storie degli utenti in base al valore aziendale che forniscono. Questa prioritizzazione ti aiuta a concentrarti innanzitutto sulle funzionalità più importanti del tuo carico di lavoro. È quindi possibile concentrare l'analisi sui componenti del carico di lavoro che fanno parte del percorso critico per tale funzionalità e ottenere valore dall'utilizzo più rapido del framework. Durante l'iterazione del processo, è possibile esaminare storie utente aggiuntive con priorità diverse.

Applicazione del framework

Il modo migliore per applicare il framework di analisi della resilienza è iniziare con una serie standard di domande, organizzate per categoria di errore, da porre su ogni componente della storia utente che state analizzando. Se alcune domande non si applicano a tutti i componenti del tuo carico di lavoro, utilizza le domande più pertinenti.

Puoi iniziare a pensare alle modalità di fallimento da due punti di vista:

- In che modo l'errore influisce sulla capacità del componente di supportare la storia dell'utente?
- In che modo l'errore influisce sulle interazioni del componente con gli altri componenti?

Ad esempio, se si considerano gli archivi di dati e il carico eccessivo, si potrebbe pensare alle modalità di errore in cui il database è sottoposto a un carico eccessivo e le query scadono. Potresti anche pensare a come il client del database potrebbe sovraccaricare il database con nuovi tentativi o non riuscire a chiudere le connessioni al database, esaurendo il pool di connessioni. Un altro esempio è un processo di autenticazione, che può comprendere diversi passaggi. È necessario riflettere su come il fallimento di un'applicazione di autenticazione a più fattori (MFA) o di un provider di identità (IdP) di terze parti potrebbe influire sulla storia utente di questo sistema di autenticazione.

Nel rispondere alle seguenti domande, è necessario considerare l'origine dell'errore. Ad esempio, il sovraccarico è stato causato da un aumento di clienti o da un operatore umano che ha messo fuori servizio troppi nodi durante un'attività di manutenzione? Potresti essere in grado di identificare più fonti di errore in ogni domanda, il che potrebbe richiedere mitigazioni diverse. Mentre ponete le domande, tenete un registro delle potenziali modalità di errore che scoprite, dei componenti a cui si applicano e dell'origine di ogni errore.

Singoli punti di errore

- Il componente è progettato per la ridondanza?
- Cosa succede se il componente si guasta?
- L'applicazione è in grado di tollerare la perdita parziale o totale di una singola zona di disponibilità?

Latenza eccessiva

- Cosa succede se questo componente presenta una latenza maggiore o un componente con cui interagisce ha una latenza maggiore (o interruzioni di rete come i ripristini TCP)?

- Hai dei timeout configurati in modo appropriato con una strategia di riprova?
- Fallisci velocemente o lentamente? Esistono effetti a cascata, ad esempio l'invio involontario di tutto il traffico a una risorsa compromessa perché si guasta rapidamente?
- Quali sono le richieste più costose fatte a questo componente?

Carico eccessivo

- Cosa può sopraffare questo componente? Come può questo componente sopraffare gli altri componenti?
- Come puoi evitare di sprecare risorse in un lavoro che non avrà mai successo?
- Avete un interruttore automatico configurato per il componente?
- Qualcosa può creare un arretrato insormontabile?
- Dove può questo componente sperimentare un comportamento bimodale?
- Quali limiti o quote di servizio possono essere superati (inclusa la capacità di archiviazione)?
- Come si ridimensiona il componente sotto carico?

Errori di configurazione e bug

- Come si evita che configurazioni errate e bug vengano implementati in produzione?
- È possibile ripristinare automaticamente una distribuzione errata o spostare il traffico lontano dal contenitore di guasto in cui è stato distribuito l'aggiornamento o la modifica?
- Quali barriere avete predisposto per prevenire gli errori degli operatori?
- Quali elementi (come credenziali o certificati) possono scadere?

Destino condiviso

- Quali sono i tuoi limiti di isolamento per colpa?
- Le modifiche apportate alle unità di dispiegamento sono di dimensioni almeno pari a quelle previstelimiti di isolamento delle fagliema idealmente più piccoli, come un ambiente monoblocco (una singola istanza all'interno del limite di isolamento dei guasti)?
- Questo componente è condiviso tra le storie degli utenti o altri carichi di lavoro?
- Quali altri componenti sono strettamente collegati a questo componente?
- Cosa succede se questo componente o le sue dipendenze presentano un guasto parziale o grigio?

Dopo aver posto queste domande, puoi utilizzare SEEMS anche per sviluppare altre domande specifiche per il tuo carico di lavoro e per ciascun componente. SEEMS è meglio utilizzato come modo strutturato per pensare alle modalità di fallimento e come fonte di ispirazione quando si esegue un'analisi della resilienza. Non è una tassonomia rigida. Non perdete tempo a preoccuparvi della categoria in cui rientra una particolare modalità di guasto: non è importante. Cosaèimportante è che tu abbia pensato al fallimento e l'abbia annotato. Non ci sono risposte sbagliate; essere creativi e pensare fuori dagli schemi è utile. Inoltre, non date per scontato che una modalità di errore sia già mitigata; includete tutte le potenziali modalità di errore a cui potete pensare.

È improbabile che tu possa anticipare tutte le potenziali modalità di fallimento nel primo esercizio. Le iterazioni multiple del framework consentono di generare un modello più completo, in modo da non dover cercare di risolvere tutto al primo passaggio. È possibile eseguire l'analisi con cadenza regolare, settimanale o bisettimanale. In ogni sessione, concentrati su una modalità o un componente di errore specifico. Ciò può contribuire a compiere progressi costanti e incrementali nel miglioramento della resilienza del carico di lavoro. Dopo aver raccolto un elenco di potenziali modalità di errore per una storia utente, puoi decidere cosa fare al riguardo.

Mitigazione dei potenziali guasti

Ora che in una storia utente sono presenti potenziali guasti dei componenti, puoi concentrarti sulle mitigazioni. Innanzitutto, esaminate i potenziali compromessi in relazione al potenziale impatto e alla probabilità di ogni guasto scoperto. Quindi determinate il livello di osservabilità richiesto e selezionate una strategia di mitigazione. I compromessi dovrebbero includere lo sforzo di strumentare il giusto livello di osservabilità e strategia di mitigazione. Infine, stabilisci la giusta cadenza per condurre revisioni periodiche dell'analisi della resilienza.

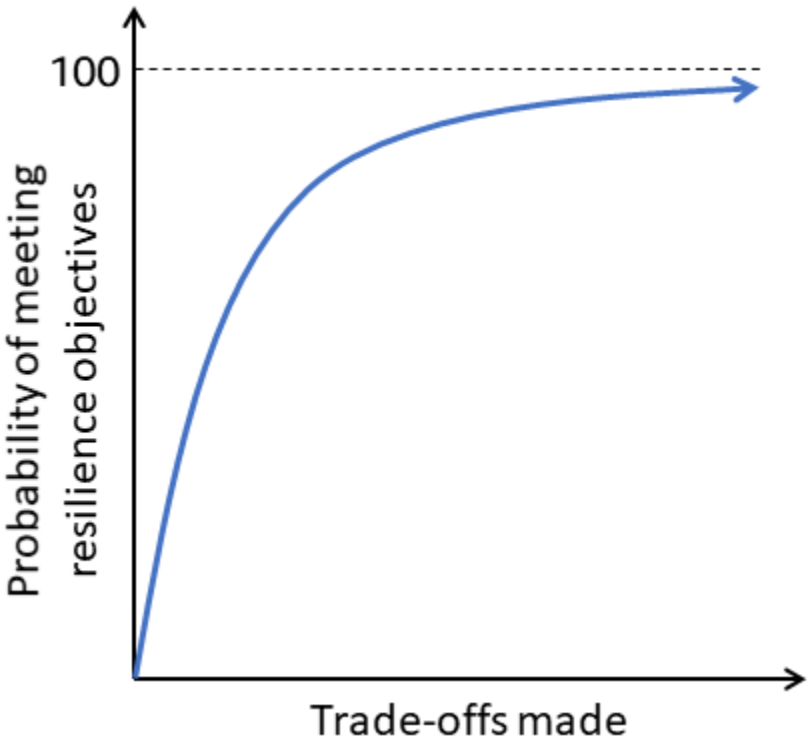
Sections

Comprendere i compromessi e i rischi

Le architetture resilienti dovrebbero utilizzare una manciata di meccanismi ben collaudati, semplici e affidabili per rispondere ai guasti. Per raggiungere i massimi livelli di resilienza, i carichi di lavoro devono rilevare e ripristinare automaticamente il maggior numero possibile di modalità di errore. Ciò richiede ingenti investimenti nell'esecuzione di analisi della resilienza. Ciò significa che il raggiungimento di livelli più elevati di resilienza implica il raggiungimento di compromessi. Tuttavia, man mano che si continuano a fare compromessi, si arriva a un punto in cui i rendimenti rispetto ai propri obiettivi di resilienza diminuiscono. Ecco i compromessi più comuni:

- Costo: componenti ridondanti, migliore osservabilità, strumenti aggiuntivi o un maggiore utilizzo delle risorse comporteranno un aumento dei costi.
- Complessità del sistema: il rilevamento e la risposta alle modalità di guasto, comprese le soluzioni di mitigazione, e il potenziale mancato utilizzo di servizi gestiti comportano un aumento della complessità del sistema.
- Impegno ingegneristico: sono necessarie ore aggiuntive per gli sviluppatori per creare soluzioni in grado di rilevare e rispondere alle modalità di guasto.

- Sovraccarico operativo: il monitoraggio e il funzionamento di un sistema che gestisce più modalità di errore possono aumentare il sovraccarico operativo, in particolare quando non è possibile utilizzare i servizi gestiti per mitigare modalità di errore specifiche.
- Latenza e coerenza: la creazione di sistemi distribuiti che favoriscono la disponibilità richiede compromessi in termini di coerenza e latenza, come descritto nel teorema PACELC.



Quando consideri le mitigazioni per le modalità di errore identificate nella storia dell'utente, considera i compromessi che è necessario adottare. Come per la sicurezza, la resilienza è un problema di ottimizzazione. È necessario decidere se evitare, mitigare, trasferire o accettare i rischi posti dal guasto identificato. Potrebbero esserci alcune modalità di errore che è possibile evitare, un set che si accetta e alcune che è possibile trasferire. Potresti scegliere di mitigare molte delle modalità di errore che identifichi. Per determinare l'approccio da adottare, esegui una valutazione ponendo due domande: Qual è la probabilità che si verifichi l'errore? Qual è l'impatto sul carico di lavoro se si verifica?

La probabilità è quanto sia plausibile che si verifichi un evento. Ad esempio, se la storia utente ha un componente che opera su una singola istanza Amazon Elastic Compute Cloud (Amazon EC2), il componente potrebbe subire interruzioni a un certo punto durante il funzionamento del sistema, magari a causa di procedure di patching o errori del sistema operativo. In alternativa, un database

gestito da Amazon Relational Database Service (Amazon RDS) che sincronizza i dati tra le istanze primarie e secondarie ha una bassa probabilità di diventare completamente non disponibile.

L'impatto è una stima del danno che un evento può causare. Deve essere valutato sia dal punto di vista finanziario che reputazionale ed è relativo al valore delle storie degli utenti su cui influisce. Ad esempio, un database sovraccarico potrebbe avere un impatto significativo sulla capacità di un sistema di e-commerce di accettare nuovi ordini. Tuttavia, la perdita di una singola istanza su un parco di 20 istanze supportate da un sistema di bilanciamento del carico avrebbe probabilmente un impatto minimo.

Puoi confrontare le risposte a queste domande con il costo dei compromessi che devi adottare per mitigare il rischio. Se si considerano queste informazioni in base alla soglia di rischio e agli obiettivi di resilienza, esse influiscono sulla decisione sulle modalità di fallimento che si intende mitigare attivamente.

Osservabilità della modalità di guasto

Per mitigare una modalità di errore, è innanzitutto necessario rilevare che sta attualmente influendo o sta per influire sul carico di lavoro. Una mitigazione è efficace solo se c'è un segnale che indica che è necessario intraprendere un'azione. Ciò significa che parte della creazione di qualsiasi mitigazione include, come minimo, la verifica di disporre o di creare l'osservabilità necessaria per rilevare l'impatto del guasto.

È necessario considerare i sintomi osservabili della modalità di guasto in due dimensioni:

- Quali sono gli indicatori principali che vi informano che il sistema si sta avvicinando a una condizione in cui presto potrebbe verificarsi un impatto?
- Quali sono gli indicatori di ritardo che possono mostrare l'impatto della modalità di guasto il più rapidamente possibile dopo che si è verificata?

Ad esempio, un errore di caricamento eccessivo applicato a un elemento del database potrebbe avere come indicatore principale il numero di connessioni. Il costante aumento del numero di connessioni può essere considerato un indicatore anticipato del fatto che il database potrebbe presto superare il limite di connessioni. In questo modo è possibile intervenire, ad esempio interrompere le connessioni utilizzate meno di recente, per ridurre il numero di connessioni. L'indicatore di ritardo indica quando il limite di connessione al database è stato superato e gli errori di connessione al database aumentano. Oltre a raccogliere i parametri delle applicazioni e dell'infrastruttura, prendi in

considerazione la possibilità di raccogliere indicatori chiave di prestazione (KPI) per rilevare quando i guasti influiscono sull'esperienza del cliente.

Quando possibile, ti consigliamo di includere entrambi i tipi di indicatori nella tua strategia di osservabilità. In alcuni casi, potresti non essere in grado di creare indicatori anticipatori, ma dovresti sempre pianificare di avere un indicatore di ritardo per ogni errore che desideri mitigare. Per scegliere la giusta mitigazione, è inoltre necessario considerare se l'errore è stato rilevato da un indicatore iniziale o da un indicatore di ritardo. Ad esempio, considera un improvviso picco di traffico verso il tuo sito web. Probabilmente vedrai solo un indicatore di ritardo. In questo caso, il ridimensionamento automatico da solo potrebbe non essere la soluzione migliore perché richiede tempo per implementare nuove risorse, mentre la limitazione potrebbe prevenire il sovraccarico quasi immediatamente e dare all'applicazione il tempo necessario per scalare o ridurre il carico. Al contrario, per un aumento graduale del traffico, si vedrebbe un indicatore anticipatore. In questo caso, la limitazione non sarebbe appropriata perché si ha il tempo di rispondere ridimensionando automaticamente il sistema.

Strategie di mitigazione comuni

Per iniziare, pensa all'utilizzo di mitigazioni preventive per evitare che la modalità di errore influisca sulla storia dell'utente. Quindi dovresti pensare a mitigazioni correttive. Le mitigazioni correttive aiutano il sistema a ripararsi da solo o ad adattarsi alle mutevoli condizioni. Di seguito è riportato un elenco di mitigazioni comuni per ogni categoria di guasto che si allineano alle proprietà di resilienza.

Categoria di errore	Proprietà di resilienza desiderate	Attenuazioni
Singoli punti di errore (SPOF)	Ridondanza e tolleranza agli errori	• Implementa la ridondanza, ad esempio utilizzando più istanze EC2 su Elastic Load Balancing (ELB). • Rimuovi le dipendenze dal piano di controllo del servizio AWS globale e assumi le dipendenze solo dai piani dati di servizio globali.

		• Se una risorsa non è disponibile, ricorri alla degradazione graduale, in modo che il sistema sia staticamente stabile fino a un singolo punto di errore.
Carico eccessivo	Capacità sufficiente	• Le principali strategie di mitigazione sono la limitazione della velocità, la riduzione del carico e la definizione delle priorità del lavoro, il lavoro costante, il backoff esponenziale e i tentativi con jitter o senza riprovare affatto, il controllo del servizio più piccolo, la gestione della profondità della coda, la scalabilità automatica, la prevenzione delle cold cache e gli interruttori automatici. • Dovresti anche prendere in considerazione il tuo piano di capacità e pensare ai limiti futuri di capacità e scalabilità, sia relativi alle risorse AWS che ai limiti all'interno del tuo sistema, che potresti raggiungere.

Latenza eccessiva	Uscita tempestiva	• Implementa timeout configurati in modo appropriato o timeout adattivi (modificando i valori di timeout in base alle condizioni di latenza attuali e previste per consentire potenzialmente a una dipendenza lenta di progredire invece di rinunciare a richieste lente). • Implementa il backoff esponenziale e riprova con jitter, hedging, utilizzando tecnologie come il protocollo TCP multipath per la connessione ai servizi cloud da ambienti locali e sperimenta la latenza su percorsi specifici, utilizzando interazioni asincrone con sistemi ad accoppiamento variabile, memorizzazione nella cache e senza sprecare il lavoro.

Errori di configurazione e bug	Output corretto	• Il modo principale per rilevare errori funzionali ripetibili nel software è eseguire test rigorosi attraverso meccanismi come analisi statica, test unitari, test di integrazione, test di regressione, test di carico e test di resilienza. • Implementa strategie come l'automazione dell'infrastruttura come codice (IaC) e l'integrazione continua e la distribuzione continua (CI/CD) per contribuire a mitigare le minacce di configurazione errata. • Utilizza tecniche di implementazione come implementazioni one-box, distribuzioni canarie, distribuzioni frazionate allineate ai limiti di isolamento dei guasti o implementazioni blu/verdi per ridurre errori di configurazione e bug.

Destino condiviso	Isolamento degli errori	• Implementa la tolleranza agli errori nel tuo sistema e utilizza limiti di isolament o logico e fisico degli errori come più cluster di calcolo o container, più account AWS, più principali AWS Identity and Access Management (IAM), più zone di disponibi lità e forse più. Regioni AWS • Anche tecniche come le architetture basate su celle e lo shuffle sharding possono migliorare l'isolame nto dei guasti. • Prendi in considerazione modelli come l'accoppiam ento allentato e la degradazi one graduale per evitare guasti a cascata. Quando si assegnano priorità alle storie degli utenti, è possibile utilizzare tale prioritizzazione anche per distinguere tra storie utente essenziali per la funzione aziendale principale e storie utente che possono essere degradate con eleganza. Ad esempio, in un sito di e-commerce, non vorresti che una disattivazione del widget per le promozioni sul sito Web influisse sulla

capacità di elaborare nuovi ordini.

Sebbene alcune di queste mitigazioni richiedano uno sforzo minimo per essere implementate, altre (come l'adozione di un'architettura basata su celle per l'isolamento prevedibile degli errori e il minimo degli errori a destino condiviso) potrebbero richiedere una riprogettazione dell'intero carico di lavoro e non solo dei componenti di una particolare storia utente. Come discusso in precedenza, è importante valutare la probabilità e l'impatto della modalità di errore rispetto ai compromessi che si adottano per mitigarla.

Oltre alle tecniche di mitigazione che si applicano a ciascuna categoria di modalità di errore, è necessario prendere in considerazione le mitigazioni necessarie per il ripristino della storia utente o dell'intero sistema. Ad esempio, un errore potrebbe interrompere un flusso di lavoro e impedire la scrittura dei dati nelle destinazioni previste. In questo caso, potrebbero essere necessari strumenti operativi per reindirizzare il flusso di lavoro o correggere manualmente i dati. Potrebbe anche essere necessario inserire un meccanismo di checkpoint nel carico di lavoro per prevenire la perdita di dati in caso di guasti. Oppure potresti dover creare un cavo Andon per mettere in pausa il flusso di lavoro e smettere di accettare nuovi lavori per evitare ulteriori danni. In questi casi, dovreste pensare agli strumenti operativi e ai guardrail di cui avete bisogno.

Infine, dovreste sempre dare per scontato che gli umani commetteranno degli errori man mano che sviluppate la vostra strategia di mitigazione. Sebbene DevOps le pratiche moderne cerchino di automatizzare le operazioni, gli esseri umani devono comunque interagire con i carichi di lavoro per vari motivi. Un'azione umana scorretta potrebbe causare un errore in qualsiasi categoria SEEMS, ad esempio rimuovendo troppi nodi durante la manutenzione e causando un sovraccarico, oppure impostando erroneamente un flag di funzionalità. Questi scenari sono davvero un fallimento nei guardrail preventivi. Un'analisi delle cause profonde non dovrebbe mai terminare con la conclusione che «un umano ha commesso un errore». Dovrebbe invece affrontare in primo luogo i motivi per cui gli errori erano possibili. Pertanto, la strategia di mitigazione dovrebbe considerare il modo in cui gli operatori umani possono interagire con i componenti del carico di lavoro e come prevenire o ridurre al minimo l'impatto degli errori degli operatori umani attraverso barriere di sicurezza.

Miglioramento continuo

La resilienza è un processo continuo. Nel corso del ciclo di vita del sistema, l'ambiente in cui opera cambierà. Per garantire che il sistema rimanga resiliente, è necessario integrare il framework nelle revisioni operative e architetturali periodiche. Potreste trovare nuove modalità di errore che non

avevate identificato la prima volta, oppure adottare soluzioni di mitigazione nuove o mai pensate in precedenza. L'analisi della resilienza dovrebbe essere un processo iterativo e non un esercizio isolato.

Dovresti testare empiricamente le tue strategie di mitigazione con processi come l'ingegneria del caos o le giornate di gioco per verificare che funzionino come previsto. Se non disponi di un meccanismo di test rigoroso, non avrai la certezza che la mitigazione funzioni come previsto quando ne avrai bisogno. Durante l'analisi della resilienza, è possibile determinare che una modalità di errore è già gestita da una specifica mitigazione, ma è importante verificare anche queste ipotesi. È necessario verificare sia le mitigazioni esistenti che le nuove mitigazioni create utilizzando il framework di analisi della resilienza.

Dovresti anche valutare l'efficacia dell'analisi attraverso retrospettive del team. Tutti sapevano a cosa stavano lavorando durante l'analisi? Il numero di modalità di fallimento rilevate attraverso l'analisi della resilienza era in linea con le aspettative del team? Potresti identificare le mitigazioni per tutte le modalità di errore che hai scoperto? Il team ha trovato utile il processo? Credi che porterà a miglioramenti nella resilienza del tuo carico di lavoro?

Quando si verificano eventi di errore reali che influiscono sulla disponibilità del carico di lavoro, registra la modalità di errore specifica, i componenti che hanno contribuito all'errore e il modello di mitigazione utilizzato. Rendi questi metadati ricercabili nel tuo strumento di analisi post-incidente in modo da poter determinare su quali modalità di guasto e componenti concentrarti in futuro. Durante tutto questo processo, puoi coinvolgere il tuo AWS account team e gli architetti delle soluzioni.

Conclusioni e risorse

Questa guida presenta un framework per eseguire l'analisi della resilienza in modo continuo e coerente. Questo framework consente di identificare in che modo singoli punti di errore, carico eccessivo, latenza eccessiva, configurazione errata, bug e destino condiviso potrebbero influire sui componenti del carico di lavoro. L'identificazione di queste modalità di errore consente di determinare una strategia di mitigazione appropriata nell'ambito della creazione di un'architettura orientata al ripristino.

Per ulteriori informazioni sull'analisi della resilienza, consulta i seguenti link:

- Quadro del ciclo di vita della resilienza (Prescriptive Guidance) AWS
- Soluzioni per la resilienza (Solutions Library) AWS
- Verso una resilienza continua (Adrian Hornsby, The Cloud Architect, 24 marzo 2021)

Cronologia dei documenti

La tabella seguente descrive le modifiche significative a questa guida. Se desideri ricevere notifiche sugli aggiornamenti futuri, puoi iscriverti a unFeed RSS.

Modifica	Descrizione	Data
Pubblicazione iniziale	—	5 settembre 2023

AWS Glossario delle linee guida prescrittive

I seguenti sono termini comunemente usati nelle strategie, nelle guide e nei modelli forniti da AWS Prescriptive Guidance. Per suggerire voci, utilizza il link Fornisci feedback alla fine del glossario.

Numeri

7 R

Sette strategie di migrazione comuni per trasferire le applicazioni sul cloud. Queste strategie si basano sulle 5 R identificate da Gartner nel 2011 e sono le seguenti:

- Rifattorizzare/riprogettare: trasferisci un'applicazione e modifica la sua architettura sfruttando appieno le funzionalità native del cloud per migliorare l'agilità, le prestazioni e la scalabilità. Ciò comporta in genere la portabilità del sistema operativo e del database. Esempio: esegui la migrazione del database Oracle on-premise ad Amazon Aurora edizione compatibile con PostgreSQL.
- Ridefinire la piattaforma (lift and reshape): trasferisci un'applicazione nel cloud e introduci un certo livello di ottimizzazione per sfruttare le funzionalità del cloud. Esempio: migra il tuo database Oracle locale su Amazon Relational Database Service (Amazon RDS) per Oracle nel cloud. AWS
- Riacquistare (drop and shop): passa a un prodotto diverso, in genere effettuando la transizione da una licenza tradizionale a un modello SaaS. Esempio: esegui la migrazione del tuo sistema di gestione delle relazioni con i clienti (CRM) su Salesforce.com.
- Eseguire il rehosting (lift and shift): trasferisci un'applicazione sul cloud senza apportare modifiche per sfruttare le funzionalità del cloud. Esempio: migra il tuo database Oracle locale su Oracle su un'istanza EC2 nel cloud. AWS
- Trasferire (eseguire il rehosting a livello hypervisor): trasferisci l'infrastruttura sul cloud senza acquistare nuovo hardware, riscrivere le applicazioni o modificare le operazioni esistenti. Questo scenario di migrazione è specifico di VMware Cloud on AWS, che supporta la compatibilità delle macchine virtuali (VM) e la portabilità del carico di lavoro tra l'ambiente locale e. AWSÈ possibile utilizzare le tecnologie VMware Cloud Foundation dai data center on-premise durante la migrazione dell'infrastruttura a VMware Cloud su AWS. Esempio: trasferisci l'hypervisor che ospita il database Oracle su VMware Cloud on. AWS
- Riesaminare (mantenere): mantieni le applicazioni nell'ambiente di origine. Queste potrebbero includere applicazioni che richiedono una rifattorizzazione significativa che desideri rimandare a

un momento successivo e applicazioni legacy che desideri mantenere, perché non vi è alcuna giustificazione aziendale per effettuarne la migrazione.

- Ritirare: disattiva o rimuovi le applicazioni che non sono più necessarie nell'ambiente di origine.

A

ABAC

Vedi controllo degli accessi basato sugli attributi.

servizi astratti

Vedi servizi gestiti.

ACIDO

Vedi atomicità, consistenza, isolamento, durata.

migrazione attiva-attiva

Un metodo di migrazione del database in cui i database di origine e di destinazione vengono mantenuti sincronizzati (utilizzando uno strumento di replica bidirezionale o operazioni di doppia scrittura) ed entrambi i database gestiscono le transazioni provenienti dalle applicazioni di connessione durante la migrazione. Questo metodo supporta la migrazione in piccoli batch controllati anziché richiedere una conversione una tantum. È più flessibile ma richiede più lavoro rispetto alla migrazione attiva-passiva.

migrazione attiva-passiva

Un metodo di migrazione di database in cui i database di origine e di destinazione vengono mantenuti sincronizzati, ma solo il database di origine gestisce le transazioni provenienti dalle applicazioni di connessione mentre i dati vengono replicati nel database di destinazione. Il database di destinazione non accetta alcuna transazione durante la migrazione.

funzione aggregata

Una funzione SQL che opera su un gruppo di righe e calcola un singolo valore restituito per il gruppo. Esempi di funzioni aggregate includono `SUM` e. `MAX`

Intelligenza artificiale

Vedi intelligenza artificiale.

AIOps

Guarda le operazioni di intelligenza artificiale.

anonimizzazione

Il processo di eliminazione permanente delle informazioni personali in un set di dati. L'anonimizzazione può aiutare a proteggere la privacy personale. I dati anonimi non sono più considerati dati personali.

anti-modello

Una soluzione utilizzata di frequente per un problema ricorrente in cui la soluzione è controproducente, inefficace o meno efficace di un'alternativa.

controllo delle applicazioni

Un approccio alla sicurezza che consente l'uso solo di applicazioni approvate per proteggere un sistema dal malware.

portfolio di applicazioni

Una raccolta di informazioni dettagliate su ogni applicazione utilizzata da un'organizzazione, compresi i costi di creazione e manutenzione dell'applicazione e il relativo valore aziendale. Queste informazioni sono fondamentali per il processo di scoperta e analisi del portfolio e aiutano a identificare e ad assegnare la priorità alle applicazioni da migrare, modernizzare e ottimizzare.

intelligenza artificiale (IA)

Il campo dell'informatica dedicato all'uso delle tecnologie informatiche per svolgere funzioni cognitive tipicamente associate agli esseri umani, come l'apprendimento, la risoluzione di problemi e il riconoscimento di schemi. Per ulteriori informazioni, consulta la sezione Che cos'è l'intelligenza artificiale?

operazioni di intelligenza artificiale (AIOps)

Il processo di utilizzo delle tecniche di machine learning per risolvere problemi operativi, ridurre gli incidenti operativi e l'intervento umano e aumentare la qualità del servizio. Per ulteriori informazioni su come viene utilizzato AIOps nella strategia di migrazione AWS , consulta la guida all'integrazione delle operazioni.

crittografia asimmetrica

Un algoritmo di crittografia che utilizza una coppia di chiavi, una chiave pubblica per la crittografia e una chiave privata per la decrittografia. Puoi condividere la chiave pubblica perché non viene utilizzata per la decrittografia, ma l'accesso alla chiave privata deve essere altamente limitato.

atomicità, consistenza, isolamento, durabilità (ACID)

Un insieme di proprietà del software che garantiscono la validità dei dati e l'affidabilità operativa di un database, anche in caso di errori, interruzioni di corrente o altri problemi.

Controllo degli accessi basato su attributi (ABAC)

La pratica di creare autorizzazioni dettagliate basate su attributi utente, come reparto, ruolo professionale e nome del team. Per ulteriori informazioni, consulta ABAC for AWS nella documentazione AWS Identity and Access Management (IAM).

fonte di dati autorevole

Una posizione in cui è archiviata la versione principale dei dati, considerata la fonte di informazioni più affidabile. È possibile copiare i dati dalla fonte di dati autorevole in altre posizioni allo scopo di elaborarli o modificarli, ad esempio anonimizzandoli, oscurandoli o pseudonimizzandoli.

Zona di disponibilità

Una posizione distinta all'interno di un edificio Regione AWS che è isolata dai guasti in altre zone di disponibilità e offre una connettività di rete economica e a bassa latenza verso altre zone di disponibilità nella stessa regione.

AWS Cloud Adoption Framework (CAF)AWS

Un framework di linee guida e best practice AWS per aiutare le organizzazioni a sviluppare un piano efficiente ed efficace per passare con successo al cloud. AWS CAF organizza le linee guida in sei aree di interesse chiamate prospettive: business, persone, governance, piattaforma, sicurezza e operazioni. Le prospettive relative ad azienda, persone e governance si concentrano sulle competenze e sui processi aziendali; le prospettive relative alla piattaforma, alla sicurezza e alle operazioni si concentrano sulle competenze e sui processi tecnici. Ad esempio, la prospettiva relativa alle persone si rivolge alle parti interessate che gestiscono le risorse umane (HR), le funzioni del personale e la gestione del personale. In questa prospettiva, AWS CAF fornisce linee guida per lo sviluppo delle persone, la formazione e le comunicazioni per aiutare a preparare l'organizzazione all'adozione del cloud di successo. Per ulteriori informazioni, consulta il sito web di AWS CAF e il white paper AWS CAF.

AWS Workload Qualification Framework (WQF)AWS

Uno strumento che valuta i carichi di lavoro di migrazione dei database, consiglia strategie di migrazione e fornisce stime del lavoro. AWS WQF è incluso in (). AWS Schema Conversion Tool AWS SCT Analizza gli schemi di database e gli oggetti di codice, il codice dell'applicazione, le dipendenze e le caratteristiche delle prestazioni e fornisce report di valutazione.

B

bot difettoso

Un bot che ha lo scopo di interrompere o causare danni a individui o organizzazioni.

BCP

Vedi la pianificazione della continuità operativa.

grafico comportamentale

Una vista unificata, interattiva dei comportamenti delle risorse e delle interazioni nel tempo. Puoi utilizzare un grafico comportamentale con Amazon Detective per esaminare tentativi di accesso non riusciti, chiamate API sospette e azioni simili. Per ulteriori informazioni, consulta Dati in un grafico comportamentale nella documentazione di Detective.

sistema big-endian

Un sistema che memorizza per primo il byte più importante. Vedi anche endianness.

Classificazione binaria

Un processo che prevede un risultato binario (una delle due classi possibili). Ad esempio, il modello di machine learning potrebbe dover prevedere problemi come "Questa e-mail è spam o non è spam?" o "Questo prodotto è un libro o un'auto?"

filtro Bloom

Una struttura di dati probabilistica ed efficiente in termini di memoria che viene utilizzata per verificare se un elemento fa parte di un set.

distribuzioni blu/verdi

Una strategia di implementazione in cui si creano due ambienti separati ma identici. La versione corrente dell'applicazione viene eseguita in un ambiente (blu) e la nuova versione dell'applicazione nell'altro ambiente (verde). Questa strategia consente di ripristinare rapidamente il sistema con un impatto minimo.

bot

Un'applicazione software che esegue attività automatizzate su Internet e simula l'attività o l'interazione umana. Alcuni bot sono utili o utili, come i web crawler che indicizzano le informazioni su Internet. Alcuni altri bot, noti come bot dannosi, hanno lo scopo di disturbare o causare danni a individui o organizzazioni.

botnet

Reti di bot infettate da malware e controllate da un'unica parte, nota come bot herder o bot operator. Le botnet sono il meccanismo più noto per scalare i bot e il loro impatto.

ramo

Un'area contenuta di un repository di codice. Il primo ramo creato in un repository è il ramo principale. È possibile creare un nuovo ramo a partire da un ramo esistente e quindi sviluppare funzionalità o correggere bug al suo interno. Un ramo creato per sviluppare una funzionalità viene comunemente detto ramo di funzionalità. Quando la funzionalità è pronta per il rilascio, il ramo di funzionalità viene ricongiunto al ramo principale. Per ulteriori informazioni, consulta Informazioni sulle filiali (documentazione). GitHub

accesso break-glass

In circostanze eccezionali e tramite una procedura approvata, un mezzo rapido per consentire a un utente di accedere a un sito a Account AWS cui in genere non dispone delle autorizzazioni necessarie. Per ulteriori informazioni, vedere l'indicatore Implementate break-glass procedures nella guida Well-Architected AWS .

strategia brownfield

L'infrastruttura esistente nell'ambiente. Quando si adotta una strategia brownfield per un'architettura di sistema, si progetta l'architettura in base ai vincoli dei sistemi e dell'infrastruttura attuali. Per l'espansione dell'infrastruttura esistente, è possibile combinare strategie brownfield e greenfield.

cache del buffer

L'area di memoria in cui sono archiviati i dati a cui si accede con maggiore frequenza.

capacità di business

Azioni intraprese da un'azienda per generare valore (ad esempio vendite, assistenza clienti o marketing). Le architetture dei microservizi e le decisioni di sviluppo possono essere guidate dalle capacità aziendali. Per ulteriori informazioni, consulta la sezione Organizzazione in base alle funzionalità aziendali del whitepaper Esecuzione di microservizi containerizzati su AWS.

pianificazione della continuità operativa (BCP)

Un piano che affronta il potenziale impatto di un evento che comporta l'interruzione dell'attività, come una migrazione su larga scala, sulle operazioni e consente a un'azienda di riprendere rapidamente le operazioni.

C

CAF

Vedi AWS Cloud Adoption Framework.

implementazione canaria

Il rilascio lento e incrementale di una versione agli utenti finali. Quando sei sicuro, distribuisci la nuova versione e sostituisci la versione corrente nella sua interezza.

CoE

Vedi Cloud Center of Excellence.

CDC

Vedi Change Data Capture.

Change Data Capture (CDC)

Il processo di tracciamento delle modifiche a un'origine dati, ad esempio una tabella di database, e di registrazione dei metadati relativi alla modifica. È possibile utilizzare CDC per vari scopi, ad esempio il controllo o la replica delle modifiche in un sistema di destinazione per mantenere la sincronizzazione.

ingegneria del caos

Introduzione intenzionale di guasti o eventi dirompenti per testare la resilienza di un sistema. Puoi usare AWS Fault Injection Service (AWS FIS) per eseguire esperimenti che stressano i tuoi AWS carichi di lavoro e valutarne la risposta.

CI/CD

Vedi integrazione continua e distribuzione continua.

classificazione

Un processo di categorizzazione che aiuta a generare previsioni. I modelli di ML per problemi di classificazione prevedono un valore discreto. I valori discreti sono sempre distinti l'uno dall'altro. Ad esempio, un modello potrebbe dover valutare se in un'immagine è presente o meno un'auto.

crittografia lato client

Crittografia dei dati a livello locale, prima che il destinatario li Servizio AWS riceva.

centro di eccellenza del cloud (CCoE)

Un team multidisciplinare che guida le iniziative di adozione del cloud in tutta l'organizzazione, tra cui lo sviluppo di best practice per il cloud, la mobilitazione delle risorse, la definizione delle tempistiche di migrazione e la guida dell'organizzazione attraverso trasformazioni su larga scala. Per ulteriori informazioni, consulta i post di CCoE sul blog AWS Cloud Enterprise Strategy.

cloud computing

La tecnologia cloud generalmente utilizzata per l'archiviazione remota di dati e la gestione dei dispositivi IoT. Il cloud computing è generalmente connesso alla tecnologia di edge computing.

modello operativo cloud

In un'organizzazione IT, il modello operativo utilizzato per creare, maturare e ottimizzare uno o più ambienti cloud. Per ulteriori informazioni, consulta Building your Cloud Operating Model.

fasi di adozione del cloud

Le quattro fasi che le organizzazioni in genere attraversano quando migrano al AWS cloud:

- Progetto: esecuzione di alcuni progetti relativi al cloud per scopi di dimostrazione e apprendimento
- Fondamento: effettuare investimenti fondamentali per dimensionare l'adozione del cloud (ad esempio, creazione di una zona di destinazione, definizione di un CCoE, definizione di un modello operativo)
- Migrazione: migrazione di singole applicazioni
- Reinvenzione: ottimizzazione di prodotti e servizi e innovazione nel cloud

Queste fasi sono state definite da Stephen Orban nel post sul blog The Journey Toward Cloud-First & the Stages of Adoption on the AWS Cloud Enterprise Strategy. Per informazioni su come si relazionano alla strategia di AWS migrazione, consulta la guida alla preparazione alla migrazione.

CMDB

Vedi database di gestione della configurazione.

repository di codice

Una posizione in cui il codice di origine e altri asset, come documentazione, esempi e script, vengono archiviati e aggiornati attraverso processi di controllo delle versioni. Gli archivi cloud più comuni includono GitHub o AWS CodeCommit. Ogni versione del codice è denominata ramo. In una struttura a microservizi, ogni repository è dedicato a una singola funzionalità. Una singola pipeline CI/CD può utilizzare più repository.

cache fredda

Una cache del buffer vuota, non ben popolata o contenente dati obsoleti o irrilevanti. Ciò influisce sulle prestazioni perché l'istanza di database deve leggere dalla memoria o dal disco principale, il che richiede più tempo rispetto alla lettura dalla cache del buffer.

dati freddi

Dati a cui si accede raramente e che in genere sono storici. Quando si eseguono interrogazioni di questo tipo di dati, le interrogazioni lente sono in genere accettabili. Lo spostamento di questi dati su livelli o classi di storage meno costosi e con prestazioni inferiori può ridurre i costi.

visione artificiale (CV)

Un campo dell'intelligenza artificiale che utilizza l'apprendimento automatico per analizzare ed estrarre informazioni da formati visivi come immagini e video digitali. Ad esempio, AWS Panorama offre dispositivi che aggiungono CV alle reti di telecamere locali e Amazon SageMaker fornisce algoritmi di elaborazione delle immagini per CV.

deriva della configurazione

Per un carico di lavoro, una modifica della configurazione rispetto allo stato previsto. Potrebbe causare la non conformità del carico di lavoro e in genere è graduale e involontaria.

database di gestione della configurazione (CMDB)

Un repository che archivia e gestisce le informazioni su un database e il relativo ambiente IT, inclusi i componenti hardware e software e le relative configurazioni. In genere si utilizzano i dati di un CMDB nella fase di individuazione e analisi del portafoglio della migrazione.

Pacchetto di conformità

Una raccolta di AWS Config regole e azioni correttive che puoi assemblare per personalizzare i controlli di conformità e sicurezza. È possibile distribuire un pacchetto di conformità come singola entità in una regione Account AWS and o all'interno di un'organizzazione utilizzando un modello YAML. Per ulteriori informazioni, consulta i Conformance Pack nella documentazione. AWS Config

integrazione e distribuzione continua (continuous integration and continuous delivery, CI/CD)

Il processo di automazione delle fasi di origine, creazione, test, gestione temporanea e produzione del processo di rilascio del software. Il processo CI/CD è comunemente descritto come una pipeline. CI/CD può aiutare ad automatizzare i processi, migliorare la produttività, migliorare

la qualità del codice e velocizzare le distribuzioni. Per ulteriori informazioni, consulta Vantaggi della distribuzione continua. CD può anche significare continuous deployment (implementazione continua). Per ulteriori informazioni, consulta Distribuzione continua e implementazione continua a confronto.

CV

Vedi visione artificiale.

D

dati a riposo

Dati stazionari nella rete, ad esempio i dati archiviati.

classificazione dei dati

Un processo per identificare e classificare i dati nella rete in base alla loro criticità e sensibilità. È un componente fondamentale di qualsiasi strategia di gestione dei rischi di sicurezza informatica perché consente di determinare i controlli di protezione e conservazione appropriati per i dati. La classificazione dei dati è un componente del pilastro della sicurezza nel AWS Well-Architected Framework. Per ulteriori informazioni, consulta Classificazione dei dati.

deriva dei dati

Una variazione significativa tra i dati di produzione e i dati utilizzati per addestrare un modello di machine learning o una modifica significativa dei dati di input nel tempo. La deriva dei dati può ridurre la qualità, l'accuratezza e l'equità complessive nelle previsioni dei modelli ML.

dati in transito

Dati che si spostano attivamente attraverso la rete, ad esempio tra le risorse di rete.

rete di dati

Un framework architettonico che fornisce la proprietà distribuita e decentralizzata dei dati con gestione e governance centralizzate.

riduzione al minimo dei dati

Il principio della raccolta e del trattamento dei soli dati strettamente necessari. Praticare la riduzione al minimo dei dati in the Cloud AWS può ridurre i rischi per la privacy, i costi e l'impronta di carbonio delle analisi.

perimetro dei dati

Una serie di barriere preventive nell' AWS ambiente che aiutano a garantire che solo le identità attendibili accedano alle risorse attendibili delle reti previste. Per ulteriori informazioni, consulta Building a data perimeter on. AWS

pre-elaborazione dei dati

Trasformare i dati grezzi in un formato che possa essere facilmente analizzato dal modello di ML. La pre-elaborazione dei dati può comportare la rimozione di determinate colonne o righe e l'eliminazione di valori mancanti, incoerenti o duplicati.

provenienza dei dati

Il processo di tracciamento dell'origine e della cronologia dei dati durante il loro ciclo di vita, ad esempio il modo in cui i dati sono stati generati, trasmessi e archiviati.

soggetto dei dati

Un individuo i cui dati vengono raccolti ed elaborati.

data warehouse

Un sistema di gestione dei dati che supporta la business intelligence, come l'analisi. I data warehouse contengono in genere grandi quantità di dati storici e vengono generalmente utilizzati per interrogazioni e analisi.

linguaggio di definizione del database (DDL)

Istruzioni o comandi per creare o modificare la struttura di tabelle e oggetti in un database.

linguaggio di manipolazione del database (DML)

Istruzioni o comandi per modificare (inserire, aggiornare ed eliminare) informazioni in un database.

DDL

Vedi linguaggio di definizione del database.

deep ensemble

Combinare più modelli di deep learning per la previsione. È possibile utilizzare i deep ensemble per ottenere una previsione più accurata o per stimare l'incertezza nelle previsioni.

deep learning

Un sottocampo del ML che utilizza più livelli di reti neurali artificiali per identificare la mappatura tra i dati di input e le variabili target di interesse.

defense-in-depth

Un approccio alla sicurezza delle informazioni in cui una serie di meccanismi e controlli di sicurezza sono accuratamente stratificati su una rete di computer per proteggere la riservatezza, l'integrità e la disponibilità della rete e dei dati al suo interno. Quando si adotta questa strategia AWS, si aggiungono più controlli a diversi livelli della AWS Organizations struttura per proteggere le risorse. Ad esempio, un defense-in-depth approccio potrebbe combinare l'autenticazione a più fattori, la segmentazione della rete e la crittografia.

amministratore delegato

In AWS Organizations, un servizio compatibile può registrare un account AWS membro per amministrare gli account dell'organizzazione e gestire le autorizzazioni per quel servizio. Questo account è denominato amministratore delegato per quel servizio specifico. Per ulteriori informazioni e un elenco di servizi compatibili, consulta Servizi che funzionano con AWS Organizations nella documentazione di AWS Organizations .

implementazione

Il processo di creazione di un'applicazione, di nuove funzionalità o di correzioni di codice disponibili nell'ambiente di destinazione. L'implementazione prevede l'applicazione di modifiche in una base di codice, seguita dalla creazione e dall'esecuzione di tale base di codice negli ambienti applicativi.

Ambiente di sviluppo

Vedi ambiente.

controllo di rilevamento

Un controllo di sicurezza progettato per rilevare, registrare e avvisare dopo che si è verificato un evento. Questi controlli rappresentano una seconda linea di difesa e avvisano l'utente in caso di eventi di sicurezza che aggirano i controlli preventivi in vigore. Per ulteriori informazioni, consulta Controlli di rilevamento in Implementazione dei controlli di sicurezza in AWS.

mappatura del flusso di valore dello sviluppo (DVSM)

Un processo utilizzato per identificare e dare priorità ai vincoli che influiscono negativamente sulla velocità e sulla qualità nel ciclo di vita dello sviluppo del software. DVSM estende il processo di

mappatura del flusso di valore originariamente progettato per pratiche di produzione snella. Si concentra sulle fasi e sui team necessari per creare e trasferire valore attraverso il processo di sviluppo del software.

gemello digitale

Una rappresentazione virtuale di un sistema reale, ad esempio un edificio, una fabbrica, un'attrezzatura industriale o una linea di produzione. I gemelli digitali supportano la manutenzione predittiva, il monitoraggio remoto e l'ottimizzazione della produzione.

tabella delle dimensioni

In uno schema a stella, una tabella più piccola che contiene gli attributi dei dati quantitativi in una tabella dei fatti. Gli attributi della tabella delle dimensioni sono in genere campi di testo o numeri discreti che si comportano come testo. Questi attributi vengono comunemente utilizzati per il vincolo delle query, il filtraggio e l'etichettatura dei set di risultati.

disastro

Un evento che impedisce a un carico di lavoro o a un sistema di raggiungere gli obiettivi aziendali nella sua sede principale di implementazione. Questi eventi possono essere disastri naturali, guasti tecnici o il risultato di azioni umane, come errori di configurazione involontari o attacchi di malware.

disaster recovery (DR)

La strategia e il processo utilizzati per ridurre al minimo i tempi di inattività e la perdita di dati causati da un disastro. Per ulteriori informazioni, consulta Disaster Recovery of Workloads su AWS: Recovery in the Cloud in the AWS Well-Architected Framework.

DML

Vedi linguaggio di manipolazione del database.

progettazione basata sul dominio

Un approccio allo sviluppo di un sistema software complesso collegandone i componenti a domini in evoluzione, o obiettivi aziendali principali, perseguiti da ciascun componente. Questo concetto è stato introdotto da Eric Evans nel suo libro,Domain-Driven Design: Tackling Complexity in the Heart of Software (Boston: Addison-Wesley Professional, 2003). Per informazioni su come utilizzare la progettazione basata sul dominio con il modello del fico strangolatore (Strangler Fig), consulta la sezione Modernizzazione incrementale dei servizi Web Microsoft ASP.NET (ASMX) legacy utilizzando container e il Gateway Amazon API.

DOTT.

Vedi disaster recovery.

rilevamento della deriva

Tracciamento delle deviazioni da una configurazione di base. Ad esempio, puoi utilizzarlo AWS CloudFormation per rilevare la deriva nelle risorse di sistema oppure puoi usarlo AWS Control Tower per rilevare cambiamenti nella tua landing zone che potrebbero influire sulla conformità ai requisiti di governance.

DVSM

Vedi la mappatura del flusso di valore dello sviluppo.

E

EDA

Vedi analisi esplorativa dei dati.

edge computing

La tecnologia che aumenta la potenza di calcolo per i dispositivi intelligenti all'edge di una rete IoT. Rispetto al cloud computing, l'edge computing può ridurre la latenza di comunicazione e migliorare i tempi di risposta.

crittografia

Un processo di elaborazione che trasforma i dati in chiaro, leggibili dall'uomo, in testo cifrato.

chiave crittografica

Una stringa crittografica di bit randomizzati generata da un algoritmo di crittografia. Le chiavi possono variare di lunghezza e ogni chiave è progettata per essere imprevedibile e univoca.

endianità

L'ordine in cui i byte vengono archiviati nella memoria del computer. I sistemi big-endian memorizzano per primo il byte più importante. I sistemi little-endian memorizzano per primo il byte meno importante.

endpoint

Vedi service endpoint.

servizio endpoint

Un servizio che puoi ospitare in un cloud privato virtuale (VPC) da condividere con altri utenti. Puoi creare un servizio endpoint con AWS PrivateLink e concedere autorizzazioni ad altri Account AWS o a AWS Identity and Access Management (IAM) principali. Questi account o principali possono connettersi al servizio endpoint in privato creando endpoint VPC di interfaccia. Per ulteriori informazioni, consulta Creazione di un servizio endpoint nella documentazione di Amazon Virtual Private Cloud (Amazon VPC).

pianificazione delle risorse aziendali (ERP)

Un sistema che automatizza e gestisce i processi aziendali chiave (come contabilità, MES e gestione dei progetti) per un'azienda.

crittografia envelope

Il processo di crittografia di una chiave di crittografia con un'altra chiave di crittografia. Per ulteriori informazioni, vedete Envelope encryption nella documentazione AWS Key Management Service (AWS KMS).

ambiente

Un'istanza di un'applicazione in esecuzione. Di seguito sono riportati i tipi di ambiente più comuni nel cloud computing:

- ambiente di sviluppo: un'istanza di un'applicazione in esecuzione disponibile solo per il team principale responsabile della manutenzione dell'applicazione. Gli ambienti di sviluppo vengono utilizzati per testare le modifiche prima di promuoverle negli ambienti superiori. Questo tipo di ambiente viene talvolta definito ambiente di test.
- ambienti inferiori: tutti gli ambienti di sviluppo di un'applicazione, ad esempio quelli utilizzati per le build e i test iniziali.
- ambiente di produzione: un'istanza di un'applicazione in esecuzione a cui gli utenti finali possono accedere. In una pipeline CI/CD, l'ambiente di produzione è l'ultimo ambiente di implementazione.
- ambienti superiori: tutti gli ambienti a cui possono accedere utenti diversi dal team di sviluppo principale. Si può trattare di un ambiente di produzione, ambienti di preproduzione e ambienti per i test di accettazione da parte degli utenti.

epica

Nelle metodologie agili, categorie funzionali che aiutano a organizzare e dare priorità al lavoro. Le epiche forniscono una descrizione di alto livello dei requisiti e delle attività di implementazione.

Ad esempio, le epopee della sicurezza AWS CAF includono la gestione delle identità e degli accessi, i controlli investigativi, la sicurezza dell'infrastruttura, la protezione dei dati e la risposta agli incidenti. Per ulteriori informazioni sulle epiche, consulta la strategia di migrazione AWS , consulta la guida all'implementazione del programma.

ERP

Vedi pianificazione delle risorse aziendali.

analisi esplorativa dei dati (EDA)

Il processo di analisi di un set di dati per comprenderne le caratteristiche principali. Si raccolgono o si aggregano dati e quindi si eseguono indagini iniziali per trovare modelli, rilevare anomalie e verificare ipotesi. L'EDA viene eseguita calcolando statistiche di riepilogo e creando visualizzazioni di dati.

F

tabella dei fatti

Il tavolo centrale in uno schema a stella. Memorizza dati quantitativi sulle operazioni aziendali. In genere, una tabella dei fatti contiene due tipi di colonne: quelle che contengono misure e quelle che contengono una chiave esterna per una tabella di dimensioni.

fallire velocemente

Una filosofia che utilizza test frequenti e incrementali per ridurre il ciclo di vita dello sviluppo. È una parte fondamentale di un approccio agile.

limite di isolamento dei guasti

Nel Cloud AWS, un limite come una zona di disponibilità Regione AWS, un piano di controllo o un piano dati che limita l'effetto di un errore e aiuta a migliorare la resilienza dei carichi di lavoro. Per ulteriori informazioni, consulta AWS Fault Isolation Boundaries.

ramo di funzionalità

Vedi filiale.

caratteristiche

I dati di input che usi per fare una previsione. Ad esempio, in un contesto di produzione, le caratteristiche potrebbero essere immagini acquisite periodicamente dalla linea di produzione.

importanza delle caratteristiche

Quanto è importante una caratteristica per le previsioni di un modello. Di solito viene espresso come punteggio numerico che può essere calcolato con varie tecniche, come Shapley Additive Explanations (SHAP) e gradienti integrati. Per ulteriori informazioni, vedere Interpretabilità del modello di machine learning con:AWS.

trasformazione delle funzionalità

Per ottimizzare i dati per il processo di machine learning, incluso l'arricchimento dei dati con fonti aggiuntive, il dimensionamento dei valori o l'estrazione di più set di informazioni da un singolo campo di dati. Ciò consente al modello di ML di trarre vantaggio dai dati. Ad esempio, se suddividi la data "2021-05-27 00:15:37" in "2021", "maggio", "giovedì" e "15", puoi aiutare l'algoritmo di apprendimento ad apprendere modelli sfumati associati a diversi componenti dei dati.

FGAC

Vedi il controllo granulare degli accessi.

controllo granulare degli accessi (FGAC)

L'uso di più condizioni per consentire o rifiutare una richiesta di accesso.

migrazione flash-cut

Un metodo di migrazione del database che utilizza la replica continua dei dati tramite l'acquisizione dei dati delle modifiche per migrare i dati nel più breve tempo possibile, anziché utilizzare un approccio graduale. L'obiettivo è ridurre al minimo i tempi di inattività.

G

blocco geografico

Vedi restrizioni geografiche.

limitazioni geografiche (blocco geografico)

In Amazon CloudFront, un'opzione per impedire agli utenti di determinati paesi di accedere alle distribuzioni di contenuti. Puoi utilizzare un elenco consentito o un elenco di blocco per specificare i paesi approvati e vietati. Per ulteriori informazioni, consulta Limitare la distribuzione geografica dei contenuti nella CloudFront documentazione.

Flusso di lavoro di GitFlow

Un approccio in cui gli ambienti inferiori e superiori utilizzano rami diversi in un repository di codice di origine. Il flusso di lavoro Gitflow è considerato obsoleto e il flusso di lavoro basato su trunk è l'approccio moderno e preferito.

strategia greenfield

L'assenza di infrastrutture esistenti in un nuovo ambiente. Quando si adotta una strategia greenfield per un'architettura di sistema, è possibile selezionare tutte le nuove tecnologie senza il vincolo della compatibilità con l'infrastruttura esistente, nota anche come brownfield. Per l'espansione dell'infrastruttura esistente, è possibile combinare strategie brownfield e greenfield.

guardrail

Una regola di livello elevato che consente di governare risorse, policy e conformità tra le unità organizzative (OU). I guardrail preventivi applicano le policy per garantire l'allineamento agli standard di conformità. Vengono implementati utilizzando le policy di controllo dei servizi e i limiti delle autorizzazioni IAM. I guardrail di rilevamento rilevano le violazioni delle policy e i problemi di conformità e generano avvisi per porvi rimedio. Sono implementati utilizzando Amazon AWS Config AWS Security Hub GuardDuty AWS Trusted Advisor, Amazon Inspector e controlli personalizzati AWS Lambda .

H

AH

Vedi disponibilità elevata.

migrazione di database eterogenea

Migrazione del database di origine in un database di destinazione che utilizza un motore di database diverso (ad esempio, da Oracle ad Amazon Aurora). La migrazione eterogenea fa in genere parte di uno sforzo di riprogettazione e la conversione dello schema può essere un'attività complessa. AWS offre AWS SCT che aiuta con le conversioni dello schema.

alta disponibilità (HA)

La capacità di un carico di lavoro di funzionare in modo continuo, senza intervento, in caso di sfide o disastri. I sistemi HA sono progettati per il failover automatico, fornire costantemente prestazioni di alta qualità e gestire carichi e guasti diversi con un impatto minimo sulle prestazioni.

modernizzazione storica

Un approccio utilizzato per modernizzare e aggiornare i sistemi di tecnologia operativa (OT) per soddisfare meglio le esigenze dell'industria manifatturiera. Uno storico è un tipo di database utilizzato per raccogliere e archiviare dati da varie fonti in una fabbrica.

migrazione di database omogenea

Migrazione del database di origine in un database di destinazione che condivide lo stesso motore di database (ad esempio, da Microsoft SQL Server ad Amazon RDS per SQL Server). La migrazione omogenea fa in genere parte di un'operazione di rehosting o ridefinizione della piattaforma. Per migrare lo schema è possibile utilizzare le utilità native del database.

dati caldi

Dati a cui si accede frequentemente, come dati in tempo reale o dati di traduzione recenti. Questi dati richiedono in genere un livello o una classe di storage ad alte prestazioni per fornire risposte rapide alle query.

hotfix

Una soluzione urgente per un problema critico in un ambiente di produzione. A causa della sua urgenza, un hotfix viene in genere creato al di fuori del tipico DevOps flusso di lavoro di rilascio.

periodo di hypercare

Subito dopo la conversione, il periodo di tempo in cui un team di migrazione gestisce e monitora le applicazioni migrate nel cloud per risolvere eventuali problemi. In genere, questo periodo dura da 1 a 4 giorni. Al termine del periodo di hypercare, il team addetto alla migrazione in genere trasferisce la responsabilità delle applicazioni al team addetto alle operazioni cloud.

I

IaC

Considera l'infrastruttura come codice.

Policy basata su identità

Una policy associata a uno o più principi IAM che definisce le relative autorizzazioni all'interno dell' Cloud AWS ambiente.

applicazione inattiva

Un'applicazione che prevede un uso di CPU e memoria medio compreso tra il 5% e il 20% in un periodo di 90 giorni. In un progetto di migrazione, è normale ritirare queste applicazioni o mantenerle on-premise.

IIoT

Vedi Industrial Internet of Things.

infrastruttura immutabile

Un modello che implementa una nuova infrastruttura per i carichi di lavoro di produzione anziché aggiornare, applicare patch o modificare l'infrastruttura esistente. Le infrastrutture immutabili sono intrinsecamente più coerenti, affidabili e prevedibili delle infrastrutture mutabili. Per ulteriori informazioni, consulta la best practice Deploy using immutable infrastructure in Well-Architected AWS Framework.

VPC in ingresso (ingress)

In un'architettura AWS multi-account, un VPC che accetta, ispeziona e indirizza le connessioni di rete dall'esterno di un'applicazione. Nel documento Architettura di riferimento per la sicurezza di AWS si consiglia di configurare l'account di rete con VPC in entrata, in uscita e di ispezione per proteggere l'interfaccia bidirezionale tra l'applicazione e Internet in generale.

migrazione incrementale

Una strategia di conversione in cui si esegue la migrazione dell'applicazione in piccole parti anziché eseguire una conversione singola e completa. Ad esempio, inizialmente potresti spostare solo alcuni microservizi o utenti nel nuovo sistema. Dopo aver verificato che tutto funzioni correttamente, puoi spostare in modo incrementale microservizi o utenti aggiuntivi fino alla disattivazione del sistema legacy. Questa strategia riduce i rischi associati alle migrazioni di grandi dimensioni.

Industria 4.0

Un termine introdotto da Klaus Schwab nel 2016 per riferirsi alla modernizzazione dei processi di produzione attraverso progressi in termini di connettività, dati in tempo reale, automazione, analisi e AI/ML.

infrastruttura

Tutte le risorse e gli asset contenuti nell'ambiente di un'applicazione.

infrastruttura come codice (IaC)

Il processo di provisioning e gestione dell'infrastruttura di un'applicazione tramite un insieme di file di configurazione. Il processo IaC è progettato per aiutarti a centralizzare la gestione dell'infrastruttura, a standardizzare le risorse e a dimensionare rapidamente, in modo che i nuovi ambienti siano ripetibili, affidabili e coerenti.

Internet delle cose industriale (IIoT)

L'uso di sensori e dispositivi connessi a Internet nei settori industriali, come quello manifatturiero, energetico, automobilistico, sanitario, delle scienze della vita e dell'agricoltura. Per ulteriori informazioni, consulta Creazione di una strategia di trasformazione digitale dell'Internet delle cose industriale (IIoT).

VPC di ispezione

In un'architettura AWS multi-account, un VPC centralizzato che gestisce le ispezioni del traffico di rete tra VPC (uguali o diversi Regioni AWS), Internet e reti locali. Nel documento Architettura di riferimento per la sicurezza di AWS si consiglia di configurare l'account di rete con VPC in entrata, in uscita e di ispezione per proteggere l'interfaccia bidirezionale tra l'applicazione e Internet in generale.

Internet of Things (IoT)

La rete di oggetti fisici connessi con sensori o processori incorporati che comunicano con altri dispositivi e sistemi tramite Internet o una rete di comunicazione locale. Per ulteriori informazioni, consulta Cos'è l'IoT?

interpretabilità

Una caratteristica di un modello di machine learning che descrive il grado in cui un essere umano è in grado di comprendere in che modo le previsioni del modello dipendono dai suoi input. Per ulteriori informazioni, consulta la sezione Interpretabilità dei modelli di machine learning con AWS.

IoT

Vedi Internet of Things.

libreria di informazioni IT (ITIL)

Una serie di best practice per offrire servizi IT e allinearli ai requisiti aziendali. ITIL fornisce le basi per ITSM.

gestione dei servizi IT (ITSM)

Attività associate alla progettazione, implementazione, gestione e supporto dei servizi IT per un'organizzazione. Per informazioni sull'integrazione delle operazioni cloud con gli strumenti ITSM, consulta la guida all'integrazione delle operazioni.

ITIL

Vedi la libreria di informazioni IT.

ITSM

Vedi Gestione dei servizi IT.

L

controllo degli accessi basato su etichette (LBAC)

Un'implementazione del controllo di accesso obbligatorio (MAC) in cui agli utenti e ai dati stessi viene assegnato esplicitamente un valore di etichetta di sicurezza. L'intersezione tra l'etichetta di sicurezza utente e l'etichetta di sicurezza dei dati determina quali righe e colonne possono essere visualizzate dall'utente.

zona di destinazione

Una landing zone è un AWS ambiente multi-account ben progettato, scalabile e sicuro. Questo è un punto di partenza dal quale le organizzazioni possono avviare e distribuire rapidamente carichi di lavoro e applicazioni con fiducia nel loro ambiente di sicurezza e infrastruttura. Per ulteriori informazioni sulle zone di destinazione, consulta la sezione Configurazione di un ambiente AWS multi-account sicuro e scalabile.

migrazione su larga scala

Una migrazione di 300 o più server.

BIANCO

Vedi controllo degli accessi basato su etichette.

Privilegio minimo

La best practice di sicurezza per la concessione delle autorizzazioni minime richieste per eseguire un'attività. Per ulteriori informazioni, consulta Applicazione delle autorizzazioni del privilegio minimo nella documentazione di IAM.

eseguire il rehosting (lift and shift)

Vedi 7 R.

sistema little-endian

Un sistema che memorizza per primo il byte meno importante. Vedi anche endianità.

ambienti inferiori

Vedi ambiente.

M

machine learning (ML)

Un tipo di intelligenza artificiale che utilizza algoritmi e tecniche per il riconoscimento e l'apprendimento di schemi. Il machine learning analizza e apprende dai dati registrati, come i dati dell'Internet delle cose (IoT), per generare un modello statistico basato su modelli. Per ulteriori informazioni, consulta la sezione Machine learning.

ramo principale

Vedi filiale.

malware

Software progettato per compromettere la sicurezza o la privacy del computer. Il malware potrebbe interrompere i sistemi informatici, divulgare informazioni sensibili o ottenere accessi non autorizzati. Esempi di malware includono virus, worm, ransomware, trojan horse, spyware e keylogger.

servizi gestiti

Servizi AWS per cui AWS gestisce il livello di infrastruttura, il sistema operativo e le piattaforme e si accede agli endpoint per archiviare e recuperare i dati. Amazon Simple Storage Service (Amazon S3) Simple Storage Service (Amazon S3) e Amazon DynamoDB sono esempi di servizi gestiti. Questi sono noti anche come servizi astratti.

sistema di esecuzione della produzione (MES)

Un sistema software per tracciare, monitorare, documentare e controllare i processi di produzione che convertono le materie prime in prodotti finiti in officina.

MAP

Vedi Migration Acceleration Program.

meccanismo

Un processo completo in cui si crea uno strumento, si promuove l'adozione dello strumento e quindi si esaminano i risultati per apportare le modifiche. Un meccanismo è un ciclo che si rafforza e si migliora man mano che funziona. Per ulteriori informazioni, consulta Creazione di meccanismi nel AWS Well-Architected Framework.

account membro

Tutti gli account Account AWS diversi dall'account di gestione che fanno parte di un'organizzazione in. AWS Organizations Un account può essere membro di una sola organizzazione alla volta.

MEH

Vedi sistema di esecuzione della produzione.

Message Queuing Telemetry Transport (MQTT)

Un protocollo di comunicazione machine-to-machine (M2M) leggero, basato sul modello di pubblicazione/sottoscrizione, per dispositivi IoT con risorse limitate.

microservizio

Un piccolo servizio indipendente che comunica tramite API ben definite ed è in genere di proprietà di piccoli team autonomi. Ad esempio, un sistema assicurativo potrebbe includere microservizi che si riferiscono a funzionalità aziendali, come vendite o marketing, o sottodomini, come acquisti, reclami o analisi. I vantaggi dei microservizi includono agilità, dimensionamento flessibile, facilità di implementazione, codice riutilizzabile e resilienza. Per ulteriori informazioni, consulta Integrazione dei microservizi utilizzando servizi serverless. AWS

architettura di microservizi

Un approccio alla creazione di un'applicazione con componenti indipendenti che eseguono ogni processo applicativo come microservizio. Questi microservizi comunicano tramite un'interfaccia ben definita utilizzando API leggere. Ogni microservizio in questa architettura può essere aggiornato, distribuito e dimensionato per soddisfare la richiesta di funzioni specifiche di un'applicazione. Per ulteriori informazioni, vedere Implementazione dei microservizi su. AWS

Programma di accelerazione della migrazione (MAP)

Un AWS programma che fornisce consulenza, supporto, formazione e servizi per aiutare le organizzazioni a costruire una solida base operativa per il passaggio al cloud e per contribuire a compensare il costo iniziale delle migrazioni. MAP include una metodologia di migrazione per eseguire le migrazioni precedenti in modo metodico e un set di strumenti per automatizzare e accelerare gli scenari di migrazione comuni.

migrazione su larga scala

Il processo di trasferimento della maggior parte del portfolio di applicazioni sul cloud avviene a ondate, con più applicazioni trasferite a una velocità maggiore in ogni ondata. Questa fase utilizza le migliori pratiche e le lezioni apprese nelle fasi precedenti per implementare una fabbrica di migrazione di team, strumenti e processi per semplificare la migrazione dei carichi di lavoro attraverso l'automazione e la distribuzione agile. Questa è la terza fase della strategia di migrazione AWS.

fabbrica di migrazione

Team interfunzionali che semplificano la migrazione dei carichi di lavoro attraverso approcci automatizzati e agili. I team di Migration Factory includono in genere operazioni, analisti e proprietari aziendali, ingegneri addetti alla migrazione, sviluppatori e DevOps professionisti che lavorano nell'ambito degli sprint. Tra il 20% e il 50% di un portfolio di applicazioni aziendali è costituito da schemi ripetuti che possono essere ottimizzati con un approccio di fabbrica. Per ulteriori informazioni, consulta la discussione sulle fabbriche di migrazione e la Guida alla fabbrica di migrazione al cloud in questo set di contenuti.

metadati di migrazione

Le informazioni sull'applicazione e sul server necessarie per completare la migrazione. Ogni modello di migrazione richiede un set diverso di metadati di migrazione. Esempi di metadati di migrazione includono la sottorete, il gruppo di sicurezza e l'account di destinazione. AWS

modello di migrazione

Un'attività di migrazione ripetibile che descrive in dettaglio la strategia di migrazione, la destinazione della migrazione e l'applicazione o il servizio di migrazione utilizzati. Esempio: riorganizza la migrazione su Amazon EC2 AWS con Application Migration Service.

Valutazione del portfolio di migrazione (MPA)

Uno strumento online che fornisce informazioni per la convalida del business case per la migrazione al Cloud. AWS MPA offre una valutazione dettagliata del portfolio (dimensionamento

corretto dei server, prezzi, confronto del TCO, analisi dei costi di migrazione) e pianificazione della migrazione (analisi e raccolta dei dati delle applicazioni, raggruppamento delle applicazioni, prioritizzazione delle migrazioni e pianificazione delle ondate). Lo strumento MPA (richiede l'accesso) è disponibile gratuitamente per tutti i AWS consulenti e i consulenti dei partner APN.

valutazione della preparazione alla migrazione (MRA)

Il processo di acquisizione di informazioni sullo stato di preparazione al cloud di un'organizzazione, l'identificazione dei punti di forza e di debolezza e la creazione di un piano d'azione per colmare le lacune identificate, utilizzando il CAF. AWS Per ulteriori informazioni, consulta la guida di preparazione alla migrazione. MRA è la prima fase della strategia di migrazione AWS.

strategia di migrazione

L'approccio utilizzato per migrare un carico di lavoro nel cloud. AWS Per ulteriori informazioni, consulta la voce 7 R in questo glossario e consulta Mobilita la tua organizzazione per accelerare le migrazioni su larga scala.

ML

Vedi machine learning.

modernizzazione

Trasformazione di un'applicazione obsoleta (legacy o monolitica) e della relativa infrastruttura in un sistema agile, elastico e altamente disponibile nel cloud per ridurre i costi, aumentare l'efficienza e sfruttare le innovazioni. Per ulteriori informazioni, vedere Strategia per la modernizzazione delle applicazioni in. Cloud AWS

valutazione della preparazione alla modernizzazione

Una valutazione che aiuta a determinare la preparazione alla modernizzazione delle applicazioni di un'organizzazione, identifica vantaggi, rischi e dipendenze e determina in che misura l'organizzazione può supportare lo stato futuro di tali applicazioni. Il risultato della valutazione è uno schema dell'architettura di destinazione, una tabella di marcia che descrive in dettaglio le fasi di sviluppo e le tappe fondamentali del processo di modernizzazione e un piano d'azione per colmare le lacune identificate. Per ulteriori informazioni, consulta la sezione Valutazione della preparazione alla modernizzazione per le applicazioni nel cloud AWS.

applicazioni monolitiche (monoliti)

Applicazioni eseguite come un unico servizio con processi strettamente collegati. Le applicazioni monolitiche presentano diversi inconvenienti. Se una funzionalità dell'applicazione registra un

picco di domanda, l'intera architettura deve essere dimensionata. L'aggiunta o il miglioramento delle funzionalità di un'applicazione monolitica diventa inoltre più complessa man mano che la base di codice cresce. Per risolvere questi problemi, puoi utilizzare un'architettura di microservizi. Per ulteriori informazioni, consulta la sezione Scomposizione dei monoliti in microservizi.

MAPPA

Vedi Migration Portfolio Assessment.

MQTT

Vedi Message Queuing Telemetry Transport.

classificazione multiclasse

Un processo che aiuta a generare previsioni per più classi (prevedendo uno o più di due risultati). Ad esempio, un modello di machine learning potrebbe chiedere "Questo prodotto è un libro, un'auto o un telefono?" oppure "Quale categoria di prodotti è più interessante per questo cliente?"

infrastruttura mutabile

Un modello che aggiorna e modifica l'infrastruttura esistente per i carichi di lavoro di produzione. Per migliorare la coerenza, l'affidabilità e la prevedibilità, il AWS Well-Architected Framework consiglia l'uso di un'infrastruttura immutabile come best practice.

O

OAC

Vedi Origin Access Control.

QUERCIA

Vedi Origin Access Identity.

OCM

Vedi gestione delle modifiche organizzative.

migrazione offline

Un metodo di migrazione in cui il carico di lavoro di origine viene eliminato durante il processo di migrazione. Questo metodo prevede tempi di inattività prolungati e viene in genere utilizzato per carichi di lavoro piccoli e non critici.

OI

Vedi l'integrazione delle operazioni.

OLA

Vedi accordo a livello operativo.

migrazione online

Un metodo di migrazione in cui il carico di lavoro di origine viene copiato sul sistema di destinazione senza essere messo offline. Le applicazioni connesse al carico di lavoro possono continuare a funzionare durante la migrazione. Questo metodo comporta tempi di inattività pari a zero o comunque minimi e viene in genere utilizzato per carichi di lavoro di produzione critici.

OPC-UA

Vedi Open Process Communications - Unified Architecture.

Comunicazioni a processo aperto - Architettura unificata (OPC-UA)

Un protocollo di comunicazione machine-to-machine (M2M) per l'automazione industriale. OPC-UA fornisce uno standard di interoperabilità con schemi di crittografia, autenticazione e autorizzazione dei dati.

accordo a livello operativo (OLA)

Un accordo che chiarisce quali sono gli impegni reciproci tra i gruppi IT funzionali, a supporto di un accordo sul livello di servizio (SLA).

revisione della prontezza operativa (ORR)

Un elenco di domande e best practice associate che aiutano a comprendere, valutare, prevenire o ridurre la portata degli incidenti e dei possibili guasti. Per ulteriori informazioni, vedere Operational Readiness Reviews (ORR) nel Well-Architected AWS Framework.

tecnologia operativa (OT)

Sistemi hardware e software che interagiscono con l'ambiente fisico per controllare le operazioni, le apparecchiature e le infrastrutture industriali. Nella produzione, l'integrazione di sistemi OT e di tecnologia dell'informazione (IT) è un obiettivo chiave per le trasformazioni dell'Industria 4.0.

integrazione delle operazioni (OI)

Il processo di modernizzazione delle operazioni nel cloud, che prevede la pianificazione, l'automazione e l'integrazione della disponibilità. Per ulteriori informazioni, consulta la guida all'integrazione delle operazioni.

trail organizzativo

Un percorso creato da noi AWS CloudTrail che registra tutti gli eventi di un'organizzazione per tutti Account AWS . AWS Organizations Questo percorso viene creato in ogni Account AWS che fa parte dell'organizzazione e tiene traccia dell'attività in ogni account. Per ulteriori informazioni, consulta Creazione di un percorso per un'organizzazione nella CloudTrail documentazione.

gestione del cambiamento organizzativo (OCM)

Un framework per la gestione di trasformazioni aziendali importanti e che comportano l'interruzione delle attività dal punto di vista delle persone, della cultura e della leadership. OCM aiuta le organizzazioni a prepararsi e passare a nuovi sistemi e strategie accelerando l'adozione del cambiamento, affrontando i problemi di transizione e promuovendo cambiamenti culturali e organizzativi. Nella strategia di AWS migrazione, questo framework si chiama accelerazione delle persone, a causa della velocità di cambiamento richiesta nei progetti di adozione del cloud. Per ulteriori informazioni, consultare la Guida OCM.

controllo dell'accesso all'origine (OAC)

In CloudFront, un'opzione avanzata per limitare l'accesso per proteggere i contenuti di Amazon Simple Storage Service (Amazon S3). OAC supporta tutti i bucket S3 in generale Regioni AWS, la crittografia lato server con AWS KMS (SSE-KMS) e le richieste dinamiche e dirette al bucket S3. `PUT` `DELETE`

identità di accesso origine (OAI)

Nel CloudFront, un'opzione per limitare l'accesso per proteggere i tuoi contenuti Amazon S3. Quando usi OAI, CloudFront crea un principale con cui Amazon S3 può autenticarsi. I principali autenticati possono accedere ai contenuti in un bucket S3 solo tramite una distribuzione specifica. CloudFront Vedi anche OAC, che fornisce un controllo degli accessi più granulare e avanzato.

O

Vedi la revisione della prontezza operativa.

- NON

Vedi la tecnologia operativa.

VPC in uscita (egress)

In un'architettura AWS multi-account, un VPC che gestisce le connessioni di rete avviate dall'interno di un'applicazione. Nel documento Architettura di riferimento per la sicurezza di AWS si consiglia di configurare l'account di rete con VPC in entrata, in uscita e di ispezione per proteggere l'interfaccia bidirezionale tra l'applicazione e Internet in generale.

P

limite delle autorizzazioni

Una policy di gestione IAM collegata ai principali IAM per impostare le autorizzazioni massime che l'utente o il ruolo possono avere. Per ulteriori informazioni, consulta Limiti delle autorizzazioni nella documentazione di IAM.

informazioni di identificazione personale (PII)

Informazioni che, se visualizzate direttamente o abbinate ad altri dati correlati, possono essere utilizzate per dedurre ragionevolmente l'identità di un individuo. Esempi di informazioni personali includono nomi, indirizzi e informazioni di contatto.

Informazioni che consentono l'identificazione personale degli utenti

Visualizza le informazioni di identificazione personale.

playbook

Una serie di passaggi predefiniti che raccolgono il lavoro associato alle migrazioni, come l'erogazione delle funzioni operative principali nel cloud. Un playbook può assumere la forma di script, runbook automatici o un riepilogo dei processi o dei passaggi necessari per gestire un ambiente modernizzato.

PLC

Vedi controllore logico programmabile.

PLM

Vedi la gestione del ciclo di vita del prodotto.

policy

Un oggetto in grado di definire le autorizzazioni (vedi politica basata sull'identità), specificare le condizioni di accesso (vedi politicabasata sulle risorse) o definire le autorizzazioni massime per tutti gli account di un'organizzazione in (vedi politica di controllo dei servizi). AWS Organizations

persistenza poliglotta

Scelta indipendente della tecnologia di archiviazione di dati di un microservizio in base ai modelli di accesso ai dati e ad altri requisiti. Se i microservizi utilizzano la stessa tecnologia di archiviazione di dati, possono incontrare problemi di implementazione o registrare prestazioni

scadenti. I microservizi vengono implementati più facilmente e ottengono prestazioni e scalabilità migliori se utilizzano l'archivio dati più adatto alle loro esigenze. Per ulteriori informazioni, consulta la sezione Abilitazione della persistenza dei dati nei microservizi.

valutazione del portfolio

Un processo di scoperta, analisi e definizione delle priorità del portfolio di applicazioni per pianificare la migrazione. Per ulteriori informazioni, consulta la pagina Valutazione della preparazione alla migrazione.

predicate

Una condizione di interrogazione che restituisce o, in genere, si trova in una clausola`WHERE`. `true` `false`

predicato pushdown

Una tecnica di ottimizzazione delle query del database che filtra i dati della query prima del trasferimento. Ciò riduce la quantità di dati che devono essere recuperati ed elaborati dal database relazionale e migliora le prestazioni delle query.

controllo preventivo

Un controllo di sicurezza progettato per impedire il verificarsi di un evento. Questi controlli sono la prima linea di difesa per impedire accessi non autorizzati o modifiche indesiderate alla rete. Per ulteriori informazioni, consulta Controlli preventivi in Implementazione dei controlli di sicurezza in AWS.

principale

Un'entità in AWS grado di eseguire azioni e accedere alle risorse. Questa entità è in genere un utente root per un Account AWS ruolo IAM o un utente. Per ulteriori informazioni, consulta Principali in Termini e concetti dei ruoli nella documentazione di IAM.

Privacy fin dalla progettazione

Un approccio all'ingegneria dei sistemi che tiene conto della privacy durante l'intero processo di progettazione.

zone ospitate private

Un container che contiene informazioni su come si desidera che Amazon Route 53 risponda alle query DNS per un dominio e i relativi sottodomini all'interno di uno o più VPC. Per ulteriori informazioni, consulta Utilizzo delle zone ospitate private nella documentazione di Route 53.

controllo proattivo

Un controllo di sicurezza progettato per impedire l'implementazione di risorse non conformi. Questi controlli analizzano le risorse prima del loro provisioning. Se la risorsa non è conforme al controllo, non viene fornita. Per ulteriori informazioni, consulta la guida di riferimento sui controlli nella AWS Control Tower documentazione e consulta Controlli proattivi in Implementazione dei controlli di sicurezza su. AWS

gestione del ciclo di vita del prodotto (PLM)

La gestione dei dati e dei processi di un prodotto durante l'intero ciclo di vita, dalla progettazione, sviluppo e lancio, attraverso la crescita e la maturità, fino al declino e alla rimozione.

Ambiente di produzione

Vedi ambiente.

controllore logico programmabile (PLC)

Nella produzione, un computer altamente affidabile e adattabile che monitora le macchine e automatizza i processi di produzione.

pseudonimizzazione

Il processo di sostituzione degli identificatori personali in un set di dati con valori segnaposto. La pseudonimizzazione può aiutare a proteggere la privacy personale. I dati pseudonimizzati sono ancora considerati dati personali.

pubblica/iscriviti (pub/sub)

Un pattern che consente comunicazioni asincrone tra microservizi per migliorare la scalabilità e la reattività. Ad esempio, in un MES basato su microservizi, un microservizio può pubblicare messaggi di eventi su un canale a cui altri microservizi possono abbonarsi. Il sistema può aggiungere nuovi microservizi senza modificare il servizio di pubblicazione.

Q

Piano di query

Una serie di passaggi, come le istruzioni, utilizzati per accedere ai dati in un sistema di database relazionale SQL.

regressione del piano di query

Quando un ottimizzatore del servizio di database sceglie un piano non ottimale rispetto a prima di una determinata modifica all'ambiente di database. Questo può essere causato da modifiche a statistiche, vincoli, impostazioni dell'ambiente, associazioni dei parametri di query e aggiornamenti al motore di database.

R

Matrice RACI

Vedi responsabile, responsabile, consultato, informato (RACI).

ransomware

Un software dannoso progettato per bloccare l'accesso a un sistema informatico o ai dati fino a quando non viene effettuato un pagamento.

Matrice RASCI

Vedi responsabile, responsabile, consultato, informato (RACI).

RCAC

Vedi controllo dell'accesso a righe e colonne.

replica di lettura

Una copia di un database utilizzata per scopi di sola lettura. È possibile indirizzare le query alla replica di lettura per ridurre il carico sul database principale.

riprogettare

Vedi 7 Rs.

obiettivo del punto di ripristino (RPO)

Il periodo di tempo massimo accettabile dall'ultimo punto di ripristino dei dati. Ciò determina quella che viene considerata una perdita di dati accettabile tra l'ultimo punto di ripristino e l'interruzione del servizio.

obiettivo del tempo di ripristino (RTO)

Il ritardo massimo accettabile tra l'interruzione del servizio e il ripristino del servizio.

rifattorizzare

Vedi 7 R.

Regione

Una raccolta di AWS risorse in un'area geografica. Ciascuna Regione AWS è isolata e indipendente dalle altre per fornire tolleranza agli errori, stabilità e resilienza. Per ulteriori informazioni, consulta Specificare cosa può usare Regioni AWS il tuo account.

regressione

Una tecnica di ML che prevede un valore numerico. Ad esempio, per risolvere il problema "A che prezzo verrà venduta questa casa?" un modello di ML potrebbe utilizzare un modello di regressione lineare per prevedere il prezzo di vendita di una casa sulla base di dati noti sulla casa (ad esempio, la metratura).

riospitare

Vedi 7 R.

rilascio

In un processo di implementazione, l'atto di promuovere modifiche a un ambiente di produzione.

trasferisco

Vedi 7 Rs.

ripiattaforma

Vedi 7 Rs.

riacquisto

Vedi 7 Rs.

resilienza

La capacità di un'applicazione di resistere o ripristinare le interruzioni. L'elevata disponibilità e il disaster recovery sono considerazioni comuni quando si pianifica la resilienza in. Cloud AWSPer ulteriori informazioni, vedere Cloud AWS Resilience.

policy basata su risorse

Una policy associata a una risorsa, ad esempio un bucket Amazon S3, un endpoint o una chiave di crittografia. Questo tipo di policy specifica a quali principali è consentito l'accesso, le azioni supportate e qualsiasi altra condizione che deve essere soddisfatta.

matrice di assegnazione di responsabilità (RACI)

Una matrice che definisce i ruoli e le responsabilità di tutte le parti coinvolte nelle attività di migrazione e nelle operazioni cloud. Il nome della matrice deriva dai tipi di responsabilità definiti nella matrice: responsabile (R), responsabile (A), consultato (C) e informato (I). Il tipo di supporto (S) è facoltativo. Se includi il supporto, la matrice viene chiamata matrice RASCI e, se la escludi, viene chiamata matrice RACI.

controllo reattivo

Un controllo di sicurezza progettato per favorire la correzione di eventi avversi o deviazioni dalla baseline di sicurezza. Per ulteriori informazioni, consulta Controlli reattivi in Implementazione dei controlli di sicurezza in AWS.

retain

Vedi 7 R.

andare in pensione

Vedi 7 Rs.

rotazione

Processo di aggiornamento periodico di un segreto per rendere più difficile l'accesso alle credenziali da parte di un utente malintenzionato.

controllo dell'accesso a righe e colonne (RCAC)

L'uso di espressioni SQL di base e flessibili con regole di accesso definite. RCAC è costituito da autorizzazioni di riga e maschere di colonna.

RPO

Vedi l'obiettivo del punto di ripristino.

RTO

Vedi l'obiettivo del tempo di ripristino.

runbook

Un insieme di procedure manuali o automatizzate necessarie per eseguire un'attività specifica. In genere sono progettati per semplificare operazioni o procedure ripetitive con tassi di errore elevati.

S

SAML 2.0

Uno standard aperto utilizzato da molti provider di identità (IdPs). Questa funzionalità abilita il single sign-on (SSO) federato, in modo che gli utenti possano accedere AWS Management Console o chiamare le operazioni AWS API senza che tu debba creare un utente in IAM per tutti i membri dell'organizzazione. Per ulteriori informazioni sulla federazione basata su SAML 2.0, consulta Informazioni sulla federazione basata su SAML 2.0 nella documentazione di IAM.

SCADA

Vedi controllo di supervisione e acquisizione dati.

SCP

Vedi la politica di controllo del servizio.

Secret

In AWS Secrets Manager, informazioni riservate o riservate, come una password o le credenziali utente, archiviate in forma crittografata. È costituito dal valore segreto e dai relativi metadati. Il valore segreto può essere binario, una stringa singola o più stringhe. Per ulteriori informazioni, consulta Secret nella documentazione di Secrets Manager.

controllo di sicurezza

Un guardrail tecnico o amministrativo che impedisce, rileva o riduce la capacità di un autore di minacce di sfruttare una vulnerabilità di sicurezza. Esistono quattro tipi principali di controlli di sicurezza: preventivi, investigativi, reattivi e proattivi.

rafforzamento della sicurezza

Il processo di riduzione della superficie di attacco per renderla più resistente agli attacchi. Può includere azioni come la rimozione di risorse che non sono più necessarie, l'implementazione di best practice di sicurezza che prevedono la concessione del privilegio minimo o la disattivazione di funzionalità non necessarie nei file di configurazione.

sistema di gestione delle informazioni e degli eventi di sicurezza (SIEM)

Strumenti e servizi che combinano sistemi di gestione delle informazioni di sicurezza (SIM) e sistemi di gestione degli eventi di sicurezza (SEM). Un sistema SIEM raccoglie, monitora e analizza i dati da server, reti, dispositivi e altre fonti per rilevare minacce e violazioni della sicurezza e generare avvisi.

automazione della risposta alla sicurezza

Un'azione predefinita e programmata progettata per rispondere o porre rimedio automaticamente a un evento di sicurezza. Queste automazioni fungono da controlli di sicurezza investigativi o reattivi che aiutano a implementare le migliori pratiche di sicurezza. AWS Esempi di azioni di risposta automatizzate includono la modifica di un gruppo di sicurezza VPC, l'applicazione di patch a un'istanza Amazon EC2 o la rotazione delle credenziali.

Crittografia lato server

Crittografia dei dati a destinazione, da parte di chi li riceve. Servizio AWS

Policy di controllo dei servizi (SCP)

Una policy che fornisce il controllo centralizzato sulle autorizzazioni per tutti gli account di un'organizzazione in AWS Organizations. Le SCP definiscono i guardrail o fissano i limiti alle azioni che un amministratore può delegare a utenti o ruoli. Puoi utilizzare le SCP come elenchi consentiti o elenchi di rifiuto, per specificare quali servizi o azioni sono consentiti o proibiti. Per ulteriori informazioni, consulta le politiche di controllo del servizio nella AWS Organizations documentazione.

endpoint del servizio

L'URL del punto di ingresso per un Servizio AWS. Puoi utilizzare l'endpoint per connetterti a livello di programmazione al servizio di destinazione. Per ulteriori informazioni, consulta Endpoint del Servizio AWS nei Riferimenti generali di AWS.

accordo sul livello di servizio (SLA)

Un accordo che chiarisce ciò che un team IT promette di offrire ai propri clienti, ad esempio l'operatività e le prestazioni del servizio.

indicatore del livello di servizio (SLI)

Misurazione di un aspetto prestazionale di un servizio, ad esempio il tasso di errore, la disponibilità o la velocità effettiva.

obiettivo a livello di servizio (SLO)

Una metrica target che rappresenta lo stato di un servizio, misurato da un indicatore del livello di servizio.

Modello di responsabilità condivisa

Un modello che descrive la responsabilità condivisa AWS per la sicurezza e la conformità del cloud. AWS è responsabile della sicurezza del cloud, mentre tu sei responsabile della sicurezza nel cloud. Per ulteriori informazioni, consulta Modello di responsabilità condivisa.

SIEM

Vedi il sistema di gestione delle informazioni e degli eventi sulla sicurezza.

punto di errore singolo (SPOF)

Un guasto in un singolo componente critico di un'applicazione che può disturbare il sistema.

SLAM

Vedi il contratto sul livello di servizio.

SLI

Vedi l'indicatore del livello di servizio.

LENTA

Vedi obiettivo del livello di servizio.

split-and-seed modello

Un modello per dimensionare e accelerare i progetti di modernizzazione. Man mano che vengono definite nuove funzionalità e versioni dei prodotti, il team principale si divide per creare nuovi team di prodotto. Questo aiuta a dimensionare le capacità e i servizi dell'organizzazione, migliora la produttività degli sviluppatori e supporta una rapida innovazione. Per ulteriori informazioni, vedere Approccio graduale alla modernizzazione delle applicazioni in. Cloud AWS

SPOF

Vedi punto di errore singolo.

schema a stella

Una struttura organizzativa di database che utilizza un'unica tabella dei fatti di grandi dimensioni per archiviare i dati transazionali o misurati e utilizza una o più tabelle dimensionali più piccole per memorizzare gli attributi dei dati. Questa struttura è progettata per l'uso in un data warehouse o per scopi di business intelligence.

modello del fico strangolatore

Un approccio alla modernizzazione dei sistemi monolitici mediante la riscrittura e la sostituzione incrementali delle funzionalità del sistema fino alla disattivazione del sistema legacy. Questo modello utilizza l'analogia di una pianta di fico che cresce fino a diventare un albero robusto e alla fine annienta e sostituisce il suo ospite. Il modello è stato introdotto da Martin Fowler come metodo per gestire il rischio durante la riscrittura di sistemi monolitici. Per un esempio di come applicare questo modello, consulta Modernizzazione incrementale dei servizi Web legacy di Microsoft ASP.NET (ASMX) mediante container e Gateway Amazon API.

sottorete

Un intervallo di indirizzi IP nel VPC. Una sottorete deve risiedere in una singola zona di disponibilità.

controllo di supervisione e acquisizione dati (SCADA)

Nella produzione, un sistema che utilizza hardware e software per monitorare gli asset fisici e le operazioni di produzione.

crittografia simmetrica

Un algoritmo di crittografia che utilizza la stessa chiave per crittografare e decrittografare i dati.

test sintetici

Test di un sistema in modo da simulare le interazioni degli utenti per rilevare potenziali problemi o monitorare le prestazioni. Puoi usare Amazon CloudWatch Synthetics per creare questi test.

T

tags

Coppie chiave-valore che fungono da metadati per l'organizzazione delle risorse. AWS Con i tag è possibile a gestire, identificare, organizzare, cercare e filtrare le risorse. Per ulteriori informazioni, consulta Tagging delle risorse AWS.

variabile di destinazione

Il valore che stai cercando di prevedere nel machine learning supervisionato. Questo è indicato anche come variabile di risultato. Ad esempio, in un ambiente di produzione la variabile di destinazione potrebbe essere un difetto del prodotto.

elenco di attività

Uno strumento che viene utilizzato per tenere traccia dei progressi tramite un runbook. Un elenco di attività contiene una panoramica del runbook e un elenco di attività generali da completare. Per ogni attività generale, include la quantità stimata di tempo richiesta, il proprietario e lo stato di avanzamento.

Ambiente di test

Vedi ambiente.

training

Fornire dati da cui trarre ispirazione dal modello di machine learning. I dati di training devono contenere la risposta corretta. L'algoritmo di apprendimento trova nei dati di addestramento i pattern che mappano gli attributi dei dati di input al target (la risposta che si desidera prevedere). Produce un modello di ML che acquisisce questi modelli. Puoi quindi utilizzare il modello di ML per creare previsioni su nuovi dati di cui non si conosce il target.

Transit Gateway

Un hub di transito di rete che è possibile utilizzare per collegare i VPC e le reti on-premise. Per ulteriori informazioni, consulta Cos'è un gateway di transito nella AWS Transit Gateway documentazione.

flusso di lavoro basato su trunk

Un approccio in cui gli sviluppatori creano e testano le funzionalità localmente in un ramo di funzionalità e quindi uniscono tali modifiche al ramo principale. Il ramo principale viene quindi integrato negli ambienti di sviluppo, preproduzione e produzione, in sequenza.

Accesso attendibile

Concessione delle autorizzazioni a un servizio specificato dall'utente per eseguire attività all'interno dell'organizzazione AWS Organizations e nei suoi account per conto dell'utente. Il servizio attendibile crea un ruolo collegato al servizio in ogni account, quando tale ruolo è necessario, per eseguire attività di gestione per conto dell'utente. Per ulteriori informazioni, consulta Utilizzo AWS Organizations con altri AWS servizi nella AWS Organizations documentazione.

regolazione

Modificare alcuni aspetti del processo di training per migliorare la precisione del modello di ML. Ad esempio, puoi addestrare il modello di ML generando un set di etichette, aggiungendo etichette e quindi ripetendo questi passaggi più volte con impostazioni diverse per ottimizzare il modello.

team da due pizze

Una piccola DevOps squadra che puoi sfamare con due pizze. Un team composto da due persone garantisce la migliore opportunità possibile di collaborazione nello sviluppo del software.

U

incertezza

Un concetto che si riferisce a informazioni imprecise, incomplete o sconosciute che possono minare l'affidabilità dei modelli di machine learning predittivi. Esistono due tipi di incertezza: l'incertezza epistemica, che è causata da dati limitati e incompleti, mentre l'incertezza aleatoria è causata dal rumore e dalla casualità insiti nei dati. Per ulteriori informazioni, consulta la guida Quantificazione dell'incertezza nei sistemi di deep learning.

compiti indifferenziati

Conosciuto anche come sollevamento di carichi pesanti, è un lavoro necessario per creare e far funzionare un'applicazione, ma che non apporta valore diretto all'utente finale né offre vantaggi competitivi. Esempi di attività indifferenziate includono l'approvvigionamento, la manutenzione e la pianificazione della capacità.

ambienti superiori

Vedi ambiente.

V

vacuum

Un'operazione di manutenzione del database che prevede la pulizia dopo aggiornamenti incrementali per recuperare lo spazio di archiviazione e migliorare le prestazioni.

controllo delle versioni

Processi e strumenti che tengono traccia delle modifiche, ad esempio le modifiche al codice di origine in un repository.

Peering VPC

Una connessione tra due VPC che consente di instradare il traffico tramite indirizzi IP privati. Per ulteriori informazioni, consulta Che cos'è il peering VPC? nella documentazione di Amazon VPC.

vulnerabilità

Un difetto software o hardware che compromette la sicurezza del sistema.

W

cache calda

Una cache del buffer che contiene dati correnti e pertinenti a cui si accede frequentemente. L'istanza di database può leggere dalla cache del buffer, il che richiede meno tempo rispetto alla lettura dalla memoria dal disco principale.

dati caldi

Dati a cui si accede raramente. Quando si eseguono interrogazioni di questo tipo di dati, in genere sono accettabili query moderatamente lente.

funzione finestra

Una funzione SQL che esegue un calcolo su un gruppo di righe che si riferiscono in qualche modo al record corrente. Le funzioni della finestra sono utili per l'elaborazione di attività, come il calcolo di una media mobile o l'accesso al valore delle righe in base alla posizione relativa della riga corrente.

Carico di lavoro

Una raccolta di risorse e codice che fornisce valore aziendale, ad esempio un'applicazione rivolta ai clienti o un processo back-end.

flusso di lavoro

Gruppi funzionali in un progetto di migrazione responsabili di una serie specifica di attività. Ogni flusso di lavoro è indipendente ma supporta gli altri flussi di lavoro del progetto. Ad esempio, il flusso di lavoro del portfolio è responsabile della definizione delle priorità delle applicazioni, della pianificazione delle ondate e della raccolta dei metadati di migrazione. Il flusso di lavoro del portfolio fornisce queste risorse al flusso di lavoro di migrazione, che quindi migra i server e le applicazioni.

VERME

Vedi scrivere una volta, leggere molti.

WQF

Vedi AWS Workload Qualification Framework.

scrivi una volta, leggi molte (WORM)

Un modello di storage che scrive i dati una sola volta e ne impedisce l'eliminazione o la modifica. Gli utenti autorizzati possono leggere i dati tutte le volte che è necessario, ma non possono modificarli. Questa infrastruttura di archiviazione dei dati è considerata immutabile.

Z

exploit zero-day

Un attacco, in genere malware, che sfrutta una vulnerabilità zero-day.

vulnerabilità zero-day

Un difetto o una vulnerabilità assoluta in un sistema di produzione. Gli autori delle minacce possono utilizzare questo tipo di vulnerabilità per attaccare il sistema. Gli sviluppatori vengono spesso a conoscenza della vulnerabilità causata dall'attacco.

applicazione zombie

Un'applicazione che prevede un utilizzo CPU e memoria inferiore al 5%. In un progetto di migrazione, è normale ritirare queste applicazioni.

Le traduzioni sono generate tramite traduzione automatica. In caso di conflitto tra il contenuto di una traduzione e la versione originale in Inglese, quest'ultima prevarrà.