Creazione di un'architettura ad alta disponibilità e disaster recovery con metodi nativi e ibridi per i database Microsoft SQL Server su Amazon EC2

AWS Guida prescrittiva

Copyright © 2025 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

AWS Guida prescrittiva: Creazione di un'architettura ad alta disponibilità e disaster recovery con metodi nativi e ibridi per i database Microsoft SQL Server su Amazon EC2

Copyright © 2025 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

I marchi e l'immagine commerciale di Amazon non possono essere utilizzati in relazione a prodotti o servizi che non siano di Amazon, in una qualsiasi modalità che possa causare confusione tra i clienti o in una qualsiasi modalità che denigri o discrediti Amazon. Tutti gli altri marchi non di proprietà di Amazon sono di proprietà delle rispettive aziende, che possono o meno essere associate, collegate o sponsorizzate da Amazon.

Table of Contents

Introduzione 1
SQL Server su architettura a EC2 nodo singolo Amazon 2
Tipi di istanza 4
Storage 5
Considerazioni su Amazon EBS e Amazon S3 6
SQL Server su Amazon FSx per Windows File Server 9
Opzioni e considerazioni sull'HA/DR 11
Gestione delle risorse HA/DR in AWS Backup 12
Utilizzo per HA/DR AWS DMS 13
AWS Application Migration Service Utilizzo per DR 15
Ulteriori considerazioni 16
Scenari di disaster recovery 17
Errore nella zona di disponibilità 17
Errore nella regione 18
Casi di utilizzo comune 19
Diagrammi di EC2 architettura SQL Server su Amazon 23
Architettura HA/DR a due nodi con cluster di gruppo Always On Availability (regione singola, Multi-AZ) 23
Architettura HA/DR a tre nodi (regione singola, Multi-AZ) 24
Architettura HA/DR a quattro nodi con cluster di gruppo a disponibilità distribuita Always On (multiregione, Multi-AZ) 25
Architettura HA/DR a tre nodi con un unico gruppo di disponibilità (multiregione) 26
Architettura HA/DR a tre nodi con spedizione dei log (multiregione) 27
Opzioni di ripristino 28
Uso di Amazon S3 28
Utilizzo AWS DataSync e Amazon FSx 29
Utilizzo di Amazon S3 File Gateway 30
Risorse e passaggi successivi 32
Appendice: Tipi di storage SSD Amazon EBS 34
Cronologia dei documenti 37
Glossario 38
..... 38
A 39
B 42

C 44
D 47
E 51
F 54
G 55
H 57
I 58
L 60
M 62
O 66
P 69
Q 72
R 72
S 75
T 79
U 80
V 81
W 82
Z 83
..... lxxxiv

Creazione di un'architettura ad alta disponibilità e disaster recovery con metodi nativi e ibridi per i database Microsoft SQL Server su Amazon EC2

Ram Yellapragada e Alysia Tran, Amazon Web Services ()AWS

Febbraio 2022 (cronologia del documento)

Microsoft SQL Server offre molte opzioni native per supportare l'alta disponibilità (HA) e il disaster recovery (DR), per garantire la continuità aziendale per i carichi di lavoro del database. Questa guida descrive una configurazione ideale per SQL Server su Amazon Elastic Compute Cloud (Amazon EC2) nel cloud Amazon Web Services (AWS). Il rehosting di SQL Server su Amazon EC2 offre un sistema autogestito in cui è possibile mantenere il pieno controllo sulle operazioni e sulla configurazione del database.

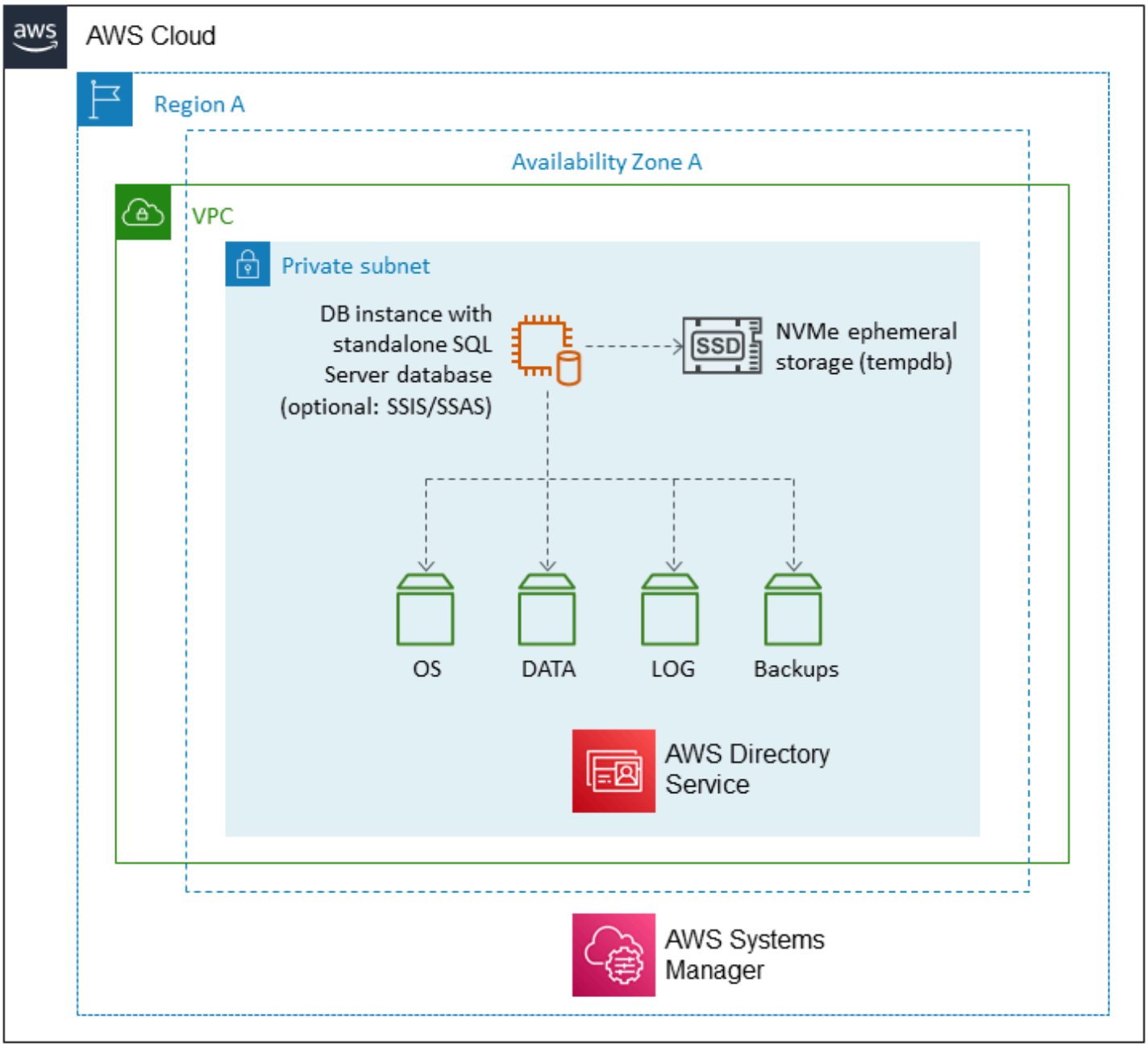
La guida illustra le opzioni HA/DR ibride di SQL Server che includono vari AWS servizi e infrastrutture e fornisce indicazioni sui componenti e le impostazioni dell'infrastruttura, tra cui classi di istanze, opzioni di archiviazione, configurazione e HA/DR setup. This document also explains how a given HA/DR strategy might fit into an example use case that has specific recovery time objective (RTO) and recovery point objective (RPO) requirements, and covers a few recovery scenarios, including relevant architecture diagrams. This guide doesn’t provide solutions designed for specific applications or requirements. It presents some HA/DR opzioni basate su RTO e RPO, in modo da poter scegliere un'architettura che soddisfi i tuoi requisiti.

Inoltre, come esercizio di dimensionamento, la guida definisce le opzioni HA/DR per un tipico carico di lavoro OLTP (Online Transaction Processing) di SQL Server e fornisce un confronto tra queste opzioni. side-by-side Per una discussione sul rehosting di SQL Server su AWS, consulta la sezione Amazon EC2 for SQL Server nella guida Migrating Microsoft SQL Server database to the AWS Cloud. Per informazioni su altre opzioni di migrazione, consulta la sezione Strategie di migrazione del database SQL Server in quella guida. Per ulteriori informazioni, consulta la Risorse e passaggi successivi sezione.

SQL Server su architettura a EC2 nodo singolo Amazon

Il diagramma seguente illustra un'architettura consigliata per un SQL Server a nodo singolo su Amazon Elastic Compute Cloud (Amazon EC2) prima di aggiungere il supporto per l'alta disponibilità (HA) e il disaster recovery (DR).

In questa architettura, il database SQL Server viene distribuito su un' EC2 istanza, utilizzando un'Amazon Machine Image (AMI) per SQL Server e volumi separati per sistema operativo, DATA, LOG e backup. Lo storage non volatile memory express (NVMe) è collegato direttamente all' EC2 istanza e utilizzato per il database tempdb di SQL Server. AWS Directory Service viene utilizzato per configurare l'autenticazione Windows per il database SQL Server. È inoltre possibile utilizzare AWS Systems Manager per rilevare e installare patch e aggiornamenti di SQL Server.



La tabella seguente riassume i consigli per configurare questa architettura. Questi consigli sono discussi in dettaglio nelle sezioni che seguono.

Tipo di istanza/AMI	• Tipo di istanza ottimizzato per le prestazioni di Amazon Elastic Block Store (Amazon EBS) • NVMe ad esempio storage (temporaneo)

	• Amazon EC2 AMIs per SQL Server
Edizione SQL Server	• Edizione SQL Server Developer (non di produzione) • Edizioni SQL Server Standard ed Enterprise (produzione)
Storage Type (Tipo di storage)	• Amazon EBS • NVMe(tempdb) (gp2//io1) io2
Volumi	• Sistema operativo • DATA • LOG • tempdb • Spazio di memoria virtuale per archiviare e scaricare i backup
Opzioni DR	• Amazon EC2 • Amazon EBS snapshots • Backup nativi di SQL Server

Tipi di istanza

AWS offre una selezione di classi di istanze per i carichi di lavoro di SQL Server. È possibile scegliere tra opzioni ottimizzate per il calcolo, per la memoria, per lo storage, per uso generico e di altro tipo, a seconda del carico di lavoro previsto sul server di database, della versione, delle opzioni HA/DR, dei core richiesti e delle considerazioni relative alle licenze. Ti consigliamo di scegliere tipi di istanza ottimizzati per Amazon EBS per SQL Server. Questi offrono il miglior throughput con volumi EBS collegati in una rete dedicata, il che è fondamentale per i carichi di lavoro di SQL Server che potrebbero avere requisiti elevati di accesso ai dati. Per i carichi di lavoro di database standard, puoi eseguire classi di istanze ottimizzate per la memoria come R5, R5b, R5d e R5n. È inoltre possibile includere lo storage o lo storage delle istanze. NVMe Sono entrambi ideali per tempdb e offrono prestazioni bilanciate per i carichi di lavoro del database.

Per i carichi di lavoro critici, l'istanza z1d ad alte prestazioni è ottimizzata per carichi di lavoro che comportano costi di licenza elevati, come SQL Server. L'istanza z1d è costruita con un processore scalabile Intel Xeon personalizzato che offre una frequenza turbo all-core sostenuta fino a 4.0, che è significativamente più veloce rispetto ad altre istanze. GHz Per i carichi di lavoro che richiedono un'elaborazione sequenziale più veloce, puoi eseguire meno core con un'istanza z1d e ottenere prestazioni uguali o migliori rispetto ad altre istanze con più core.

Amazon fornisce anche servizi dedicati AMIsper SQL Server su Microsoft Windows Server per aiutarti a ospitare le ultime edizioni di SQL Server su Amazon EC2.

Storage

Alcuni tipi di istanze offrono volumi di archiviazione delle NVMe istanze. NVMe è un'opzione di archiviazione temporanea (effimera). Questo storage è collegato direttamente all'istanza. EC2 Sebbene NVMe lo storage sia temporaneo e i dati vadano persi al riavvio, offre le prestazioni ottimali. Pertanto, è adatto per il database tempdb di SQL Server, che presenta I/O elevati e modelli di accesso casuali ai dati. Non sono previsti costi aggiuntivi per l'utilizzo di un NVMe instance store per tempdb. Per ulteriori indicazioni, consulta la sezione Posizionare tempdb in un instance store nella guida Best practice for deploying SQL Server on Amazon. EC2

Amazon EBS è una soluzione di storage durevole che soddisfa i requisiti di SQL Server per uno storage veloce e disponibile. Microsoft consiglia di tenere separati i volumi di dati e log per prestazioni ottimali. I motivi di questa separazione includono i seguenti:

- Diversi metodi di accesso ai dati. I volumi di dati utilizzano l'accesso casuale ai dati OLTP (Online Transaction Processing), mentre i volumi di registro utilizzano l'accesso seriale.
- Opzioni di ripristino migliori. La perdita di un volume non influisce sull'altro volume e facilita il recupero dei dati.
- Diversi tipi di carico di lavoro. I volumi di dati sono destinati ai carichi di lavoro OLTP, mentre i volumi di registro sono destinati ai carichi di lavoro OLAP (Online Analytic Processing).
- Requisiti prestazionali diversi. I volumi di dati e log hanno requisiti di IOPS e latenza diversi, velocità di throughput minime e benchmark prestazionali simili.

Per selezionare il tipo di volume Amazon EBS giusto, è necessario analizzare i metodi di accesso al database, gli IOPS e il throughput. Raccogli i parametri sia durante le normali ore lavorative che durante i picchi di utilizzo. SQL Server utilizza le estensioni per archiviare i dati. L'unità di

archiviazione atomica in SQL Server è una pagina con una dimensione di 8 KB. Otto pagine fisicamente contigue costituiscono un'estensione, che ha una dimensione di 64 KB. Pertanto, su un computer SQL Server, la dimensione dell'unità di allocazione NTFS per l'hosting dei file di database SQL (incluso tempdb) deve essere di 64 KB. Per informazioni su come controllare la dimensione di allocazione NTFS delle unità, consulta la guida Best practice for deploying SQL Server on Amazon. EC2

La scelta del volume EBS dipende dal carico di lavoro, ovvero dal fatto che il database richieda un'intensa attività di lettura o scrittura, richieda IOPS elevati, storage di archiviazione e considerazioni simili. La tabella seguente mostra una configurazione di esempio.

Risorsa Amazon EBS	Tipo	Descrizione
Disco del sistema operativo	`gp3`	Archiviazione per uso generico.
Disco DATI	`io1`/`io2`	Archiviazione ad alta intensità di scrittura.
Disco LOG	`gp3` o `io2`	Storage generico per carichi di lavoro intensivi.
Disco di backup	`st1`	Storage di archiviazione meno costoso. Per prestazioni migliori, i backup possono anche essere archiviati su un disco più veloce se vengono copiati regolarmente su Amazon Simple Storage Service (Amazon S3).

Considerazioni su Amazon EBS e Amazon S3

La tabella seguente mostra un confronto tra Amazon EBS e Amazon S3 per lo storage. Utilizza queste informazioni per comprendere le differenze tra i due servizi e scegliere l'approccio migliore per il tuo caso d'uso.

Servizio	Disponibilità	Durabilità	Note
Amazon EBS	• Tutti i tipi di volume EBS offrono funzionalità di snapshot durevoli e sono progettati per una disponibilità del 99,999%. • È possibile utilizzare le istantanee per effettuare il provisioning di nuove istanze in diverse AWS regioni in caso di emergenza.	• I dati del volume EBS vengono replicati su più server in un'unica zona di disponibilità per evitare la perdita di dati dovuta al guasto di un singolo componente. • I volumi EBS sono progettati per un tasso di errore annuale (AFR) compreso tra lo 0,1 e lo 0,2 percento, in cui il fallimento si riferisce a una perdita totale o parziale del volume, a seconda delle dimensioni e delle prestazioni del volume.	• Un'istanza ottimizzata per Amazon EBS utilizza uno stack di configurazione ottimizzato e fornisce una larghezza di banda aggiuntiva dedicata per l'I/O di Amazon EBS. Questa ottimizzazione offre le migliori prestazioni per i tuoi volumi EBS riducendo al minimo la contesa tra l'I/O di Amazon EBS e altro traffico proveniente dall'istanza. • I ripristini rapidi delle istantanee sono supportati per un massimo di 50 istantanee contemporaneamente. È necessario abilitare questa funzionalità in modo esplicito per ogni singola istantanea. • Un'istanza ottimizzata per Amazon

Servizio	Disponibilità	Durabilità	Note
			EBS offre prestazioni complete al momento dell'inizializzazione, quindi non è necessario alcun tempo di riscaldamento.

Servizio	Disponibilità	Durabilità	Note

Amazon S3	• Alta disponibilità.
	• Progettata per una disponibilità del 99,99% in un determinato anno.
	• Sono disponibili più classi di storage, come S3 Standard e S3 Standard-Infrequent Access ((S3 Standard-IA)). È possibile spostare i file di backup in una classe di archiviazione in base a un periodo di conservazione.
		• Amazon S3, Amazon S3 Glacier e S3 Glacier Deep Archive sono progettati per garantire una durabilità del 99,99999% (11 nove). Sia Amazon S3 che S3 Glacier offrono un backup affidabile dei dati, con replica degli oggetti su almeno tre zone di disponibilità distribuite geograficamente.	• Puoi utilizzare Amazon S3 per backup a lungo termine a livello di file di SQL Server (inclusi backup completi e log delle transazioni).
			• Amazon S3 supporta:
			  • Controllo del tempo di replica (RTC)
			  • Replica tra regioni tramite la gestione del ciclo di vita di S3 e AWS Backup
			  • Tiering intelligente
			• Amazon S3 offre lo storage più economico. Si applicano i costi di trasferimento dei dati tra regioni.

SQL Server su Amazon FSx per Windows File Server

Amazon FSx for Windows File Server offre prestazioni elevate con un throughput di base fino a 2 GB/second per file system, hundreds of thousands of IOPS, and consistent sub-millisecond latencies. To provide the right performance for your SQL Server instances, you can choose a throughput level that

is independent of your file system size. Higher levels of throughput capacity also come with higher levels of IOPS that the file server can serve to the SQL Server instances accessing it. The storage capacity determines not only how much data you can store, but also how many I/O operazioni al secondo (IOPS) che puoi eseguire sullo storage: ogni GB di storage fornisce 3 IOPS. È possibile effettuare il provisioning di ogni file system in modo che abbia una dimensione massima di 64 TiB (rispetto ai 16 TiB di Amazon EBS). Puoi anche utilizzare FSx i sistemi Amazon come testimone della condivisione di file per le distribuzioni di Windows Server Failover Cluster (WSFC).

Opzioni e considerazioni sull'HA/DR

Sebbene la possibilità che una zona o una regione di AWS disponibilità vada completamente offline è estremamente rara, consigliamo un approccio su più fronti al backup e al ripristino in caso di emergenza per la ridondanza e per ridurre al minimo la perdita di dati. I processi di backup e ripristino devono includere il livello di granularità appropriato per soddisfare il Recovery Time Objective (RTO) e il Recovery Point Objective (RPO) per il carico di lavoro e supportare i processi aziendali e spesso dipendono dall'applicazione. Nel caso dei database, supporta AWS anche tutti i consigli di Microsoft per l'installazione e la configurazione di SQL Server per l'alta disponibilità e il disaster recovery (HA/DR). Different editions of SQL Server support various HA/DR options, and you should consider special cases such as very large databases (VLDBs) on a case-by-case basis. As with any DR configuration, testing is essential to ensure that each application meets its service-level agreements (SLAs) for HA/DR. For your test/developmentambiente), prendi in considerazione l'utilizzo dell'edizione SQL Server Developer, che è gratuita ma presenta delle limitazioni.

Per un caso d'uso che richiede un RPO di 15 minuti e un RTO di 4 ore, è possibile prendere in considerazione una combinazione delle seguenti opzioni HA/DR:

- Opzioni HA/DR native di SQL Server con standby caldo (a livello di database): per illustrazioni di alcune di queste architetture, consulta la sezione più avanti in questa guida. Diagrammi di EC2 architettura SQL Server su Amazon
  - Multi-AZ a due nodi in una singola regione (modalità di commit sincrona) o in più regioni (modalità di commit asincrona, gruppo di disponibilità di base)
  - Multi-AZ a tre nodi (o più) in più regioni (modalità di commit sincrono e commit asincrono)
  - Spedizione a due nodi, Multi-AZ e dei log in più regioni (con backup dei log ogni 5 minuti)
- Backup nativi di SQL Server su Amazon S3 (a livello di database, solo DR): backup completi (una volta al giorno)
  - Backup differenziali (ogni 2-4 ore).
  - Backup dei log (ogni 5-10 minuti).
  - I backup devono essere eseguiti e copiati su Amazon Simple Storage Service (Amazon S3) utilizzando script personalizzati o un'opzione come File Gateway per backup e trasferimento efficienti.

  - Se disponi di centinaia di database, puoi continuare a utilizzare gli strumenti di backup esistenti (come Commvault o Litespeed) per gestire in modo efficiente i backup e archiviarli direttamente in Amazon S3.
  - Usa Amazon S3 Cross-Region Replication (CRR) con S3 Replication Time Control (RTC) per controllare e monitorare la replica degli oggetti entro uno SLA di 15 minuti.
  - Per garantire la conformità e ridurre i costi, puoi anche utilizzare S3 Lifecycle Management per spostare e archiviare i backup più vecchi per lo storage a lungo termine.
  - Se esegui backup nativi di SQL Server e li sposti regolarmente su Amazon S3, in caso di emergenza, i backup saranno immediatamente disponibili nella regione di destinazione. Ciò elimina la necessità di trasferire backup o ripristinare istantanee.
  - Si consiglia di utilizzare SQL Native Backup Compression per ridurre le dimensioni dei file.
- AWS istantanee (a livello di istanza e volume, solo DR)
  - Amazon Elastic Compute Cloud (Amazon EC2) Backup di Amazon Machine Image (AMI) per ricostruire database da zero
  - Istantanee del volume Amazon Elastic Block Store (Amazon EBS) per collegare volumi EBS ad Amazon EC2

Gestione delle risorse HA/DR in AWS Backup

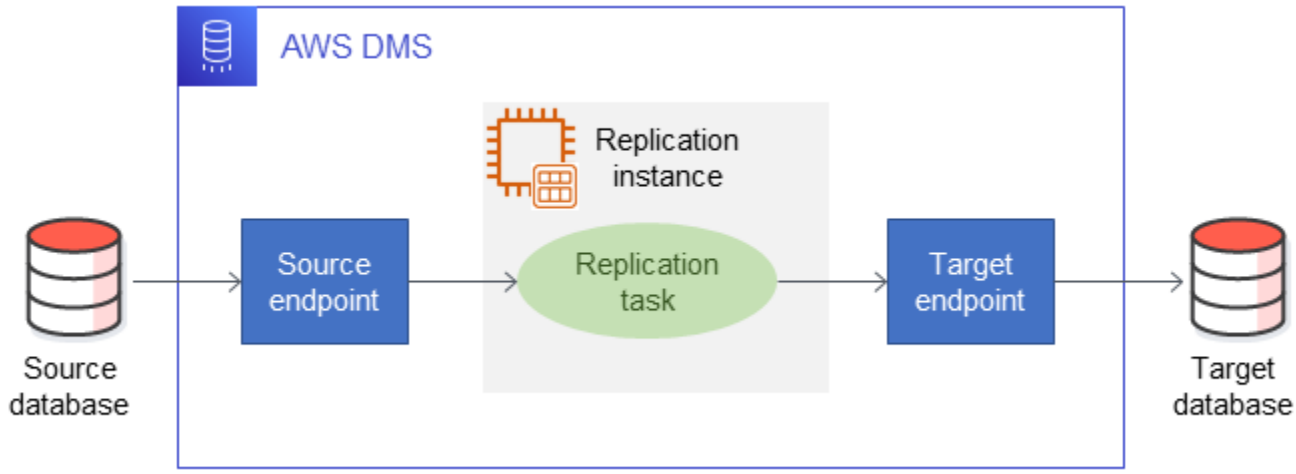
AWS Backupè un servizio completamente gestito che offre la possibilità di creare piani e pianificazioni di backup e assegnare AWS le risorse coinvolte nella configurazione HA/DR, come i volumi Amazon EBS per creare istantanee e Amazon, a questi piani di backup. EC2 AMIs Puoi anche AWS Backup utilizzarlo per pianificare copie multiregionali di questi snapshot EBS. Per un utilizzo ottimale, è AWS Backup necessario un efficiente meccanismo di etichettatura delle risorse. AWS Backup supporta anche backup coerenti con le applicazioni tramite Windows Volume Shadow Copy Service (VSS), che è possibile utilizzare per SQL Server. Per una protezione a livello di storage, consigliamo di utilizzare le istantanee EBS. Le istantanee EBS iniziali sono complete e le istantanee successive sono incrementali. Sebbene le istantanee EBS offrano una protezione a livello di storage, non sostituiscono i backup nativi basati su file di SQL Server che offrono il ripristino. point-in-time

Utilizzo per HA/DR AWS DMS

Se stai cercando un'alternativa alle opzioni di replica di SQL Server Always On o se disponi di database di origine e di destinazione eterogenei, in una configurazione ibrida o in AWS, puoi utilizzare AWS Database Migration Service ()AWS DMS nei seguenti modi.

Se si utilizza AWS DMS SQL Server in un contesto autogestito (ospitato su Amazon EC2 o in locale), supporta la replica una tantum e continua in due modalità: utilizzando MS-REPLICATION (per acquisire le modifiche alle tabelle con chiavi primarie) e MS-CDC (per acquisire le modifiche alle tabelle che non dispongono di chiavi primarie). Tuttavia, se utilizzi Amazon Relational Database Service (Amazon RDS) come sorgente AWS DMS per, è supportato solo MS-CDC. AWS DMS offre una gamma di endpoint di origine e destinazione, supporta motori di database eterogenei e offre un controllo granulare sul processo di replica. È inoltre possibile utilizzare il () with per migrazioni di database eterogenei AWS Schema Conversion Tool .AWS SCT AWS DMS AWS SCT automatizza le modifiche a livello di schema e produce anche report sulla preparazione e la pianificazione della migrazione.

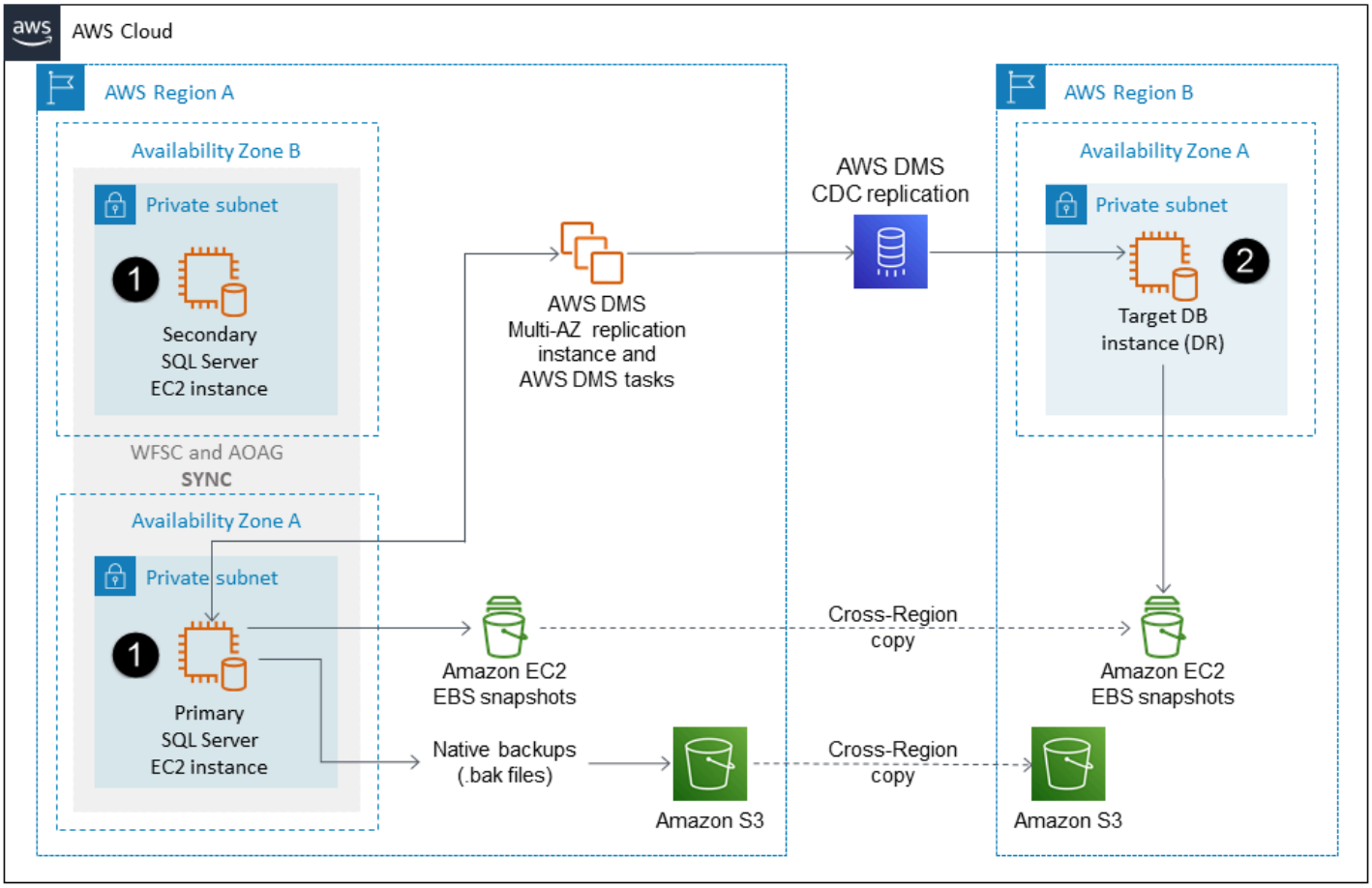
I database di origine e di destinazione vengono aggiunti come endpoint in AWS DMS, come illustrato nel diagramma seguente. Questo servizio implementa un processo di replica logica utilizzando MS-REPLICATION o MS-CDC. Se si dispone di una configurazione ibrida, è possibile configurare AWS DMS la replica continua tra sistemi locali e. AWS Durante il cutover, l'attività di AWS DMS migrazione può essere interrotta e l'applicazione sarà in grado di connettersi al database già sincronizzato con il database locale senza ulteriori ritardi. L'utilizzo AWS DMS di SQL Server come sorgente presenta alcune limitazioni, descritte nella documentazione.AWS DMS



Valuta la possibilità di utilizzare metodi HA/DR AWS DMS al posto dei metodi HA/DR nativi nei seguenti scenari:

- Quando vuoi risparmiare sui costi di licenza. Ad esempio, se utilizzi una versione avanzata come SQL Server Enterprise Edition solo per le opzioni Always On, potresti prendere in considerazione la possibilità di configurarla AWS DMS , poiché può fornire un'opzione di replica logica senza il costo di una licenza dell'edizione Enterprise.
- Quando si hanno fonti e destinazioni eterogenee. Non è necessario che le versioni di SQL Server sui nodi primari e di disaster recovery corrispondano (entro AWS DMS limiti), il che offre una flessibilità significativa.
- Per evitare il sovraccarico di Windows, il clustering di SQL Server e la configurazione e la gestione dei gruppi di disponibilità distribuita. AWS DMS offre una configurazione semplice e una facile gestione delle attività di replica.
- Per casi d'uso aziendali come il trasferimento quasi in tempo reale (a seconda dell'istanza di replica, della configurazione di rete e del volume dei dati), il mascheramento dei dati, il filtraggio selettivo, la mappatura di schema/tabelle (omogenea ed eterogenea), le valutazioni prima della migrazione e il supporto JSON.
- Per duplicare, interrompere e avviare facilmente le attività in base alle esigenze in base ai numeri di sequenza di log (), ai timestamp e a opzioni simili. LSNs

Il diagramma seguente mostra un approccio alternativo a come fornire supporto per la AWS DMS replica. In questa configurazione, l'origine è un cluster di gruppo di disponibilità Always On di SQL Server e AWS DMS utilizza l'opzione Change Data Capture (CDC) per replicare continuamente i dati su una destinazione in una regione diversa. AWS Per ottenere prestazioni ottimali, è fondamentale garantire che l'istanza di replica abbia le dimensioni corrette e rimanga nella regione di origine.



I motori di origine e di destinazione non devono necessariamente corrispondere. Nel diagramma, i nodi primari e secondari contrassegnati come (1) possono essere un cluster SQL Server in una configurazione Single-AZ o Multi-AZ. Oppure l'origine può essere un singolo nodo di SQL Server che supporta MS-CDC o MS-REPLICATION.

L'istanza DB di destinazione, contrassegnata come (2) nel diagramma, può essere qualsiasi versione di SQL Server su Amazon RDS EC2, Amazon o qualsiasi altra destinazione eterogenea. Non deve necessariamente corrispondere alle istanze primarie e secondarie o supportare i gruppi di disponibilità Always On. Ad esempio, l'origine può essere un cluster di gruppo di disponibilità SQL Server Always On e la destinazione può essere Amazon Aurora PostgreSQL Compatible Edition.

AWS Application Migration Service Utilizzo per DR

Si consiglia di utilizzare il AWS Application Migration Service per lift-and-shift le migrazioni verso. AWS Application Migration Service replica continuamente le macchine (inclusi il sistema operativo, la configurazione dello stato del sistema, i database, le applicazioni e i file) in un'area di gestione

temporanea a basso costo nell' AWS account di destinazione e nella regione preferita. In caso di emergenza, puoi utilizzare Application Migration Service per avviare automaticamente migliaia di macchine nello stato di completo provisioning in pochi minuti.

Ulteriori considerazioni

L'elenco seguente identifica i possibili ostacoli da considerare quando si progetta una strategia HA/DR.

- Larghezza di banda, latenza, complessità della rete e connettività in una configurazione di nodi multiregione.
- Dimensioni degli EC2 snapshot di Amazon EBS o Amazon e tempo necessario per copiarli utilizzando. AWS Backup
  - Le EC2 istantanee di Amazon EBS e Amazon vengono archiviate in Amazon S3 utilizzando. AWS Backup
  - Uno snapshot EBS non viene replicato nella regione di destinazione in Amazon S3 fino al completamento dello snapshot corrente. La durata della replica dipende anche dalla dimensione del volume.
  - Una volta completata l'istantanea, il tempo necessario per copiare le istantanee può essere di soli 15 minuti per il 99,99% degli oggetti. Tuttavia, sono necessari test approfonditi per casi d'uso specifici e volumi critici di grandi dimensioni.
- Tempo necessario per ripristinare i volumi EBS nella zona e nella regione di disponibilità di destinazione.
- Tempo necessario per ripristinare EC2 le immagini Amazon nella zona e nella regione di disponibilità di destinazione.
- Se si crea da zero, è necessario del tempo per fornire l'infrastruttura per l' EC2immagine Amazon o ripristinare gli snapshot EBS nella zona di disponibilità e nella regione di destinazione.
- In caso di ripristino da zero, è necessario il tempo necessario per ripristinare i backup completi, differenziali e di log nativi di SQL Server nella zona di disponibilità e nella regione di destinazione.
- Applicazioni e dipendenze esterne che devono essere disponibili in tutte le regioni.
- Limitazioni sulle dimensioni dei file per i volumi e per il caricamento su Amazon S3.

Scenari di disaster recovery

Questa sezione fornisce esempi di errori in una singola zona o AWS regione di disponibilità e illustra le opzioni per il disaster recovery (DR). Gli esempi presuppongono un Recovery Point Objective (RPO) di 15 minuti e un Recovery Time Objective (RTO) di 4 ore.

Errore nella zona di disponibilità

È possibile utilizzare una delle seguenti opzioni per il ripristino da un singolo errore della zona di disponibilità entro i parametri specificati (RPO di 15 minuti, RTO di 4 ore).

- Effettua il ripristino dell'applicazione utilizzando il backup di immagini Amazon Elastic Compute Cloud (Amazon EC2) più recente e connettiti all'istanza di database Warm Standby esistente tramite una distribuzione di gruppo di disponibilità Always On o la spedizione di log.
  - Una configurazione del gruppo di disponibilità Always On di SQL Server per DR con due o più nodi fornisce il failover automatico sul nodo secondario tramite la modalità di commit sincrono o di commit asincrono, in modo che il database sia immediatamente disponibile. Per una configurazione HA, entrambi i nodi sono disponibili per le operazioni di lettura. Questa opzione soddisfa perfettamente i requisiti RTO e RPO. Nell'edizione SQL Server Standard, è possibile utilizzare anche i gruppi di disponibilità di base, ma è limitato a due nodi, poiché un gruppo di disponibilità può includere solo un database. Tuttavia, è possibile configurare più gruppi di disponibilità all'interno di una regione o tra più regioni. Questa configurazione consente di risparmiare sui costi, poiché non vi è alcun costo aggiuntivo per il nodo secondario, che non è accessibile per le operazioni di lettura. L'edizione SQL Server Enterprise offre funzionalità complete e failover per tutti i database all'interno di un singolo gruppo di disponibilità. Per esempi di questa opzione, vedere i seguenti diagrammi di architettura:
    - Architettura HA/DR a due nodi con cluster di gruppo Always On Availability (regione singola, Multi-AZ)
    - Architettura HA/DR a tre nodi (regione singola, Multi-AZ)
    - Architettura HA/DR a quattro nodi con cluster di gruppo a disponibilità distribuita Always On (multiregione, Multi-AZ)
    - Architettura HA/DR a tre nodi con un unico gruppo di disponibilità (multiregione)
  - La spedizione dei log di SQL Server come soluzione DR richiede un failover manuale su un server di standby e dipende dalla frequenza dei backup dei log. Questa è una delle opzioni di

DR meno costose. Non è necessario che le edizioni di SQL Server per il sito DR principale e il sito DR fornito con log corrispondano. Questa opzione soddisfa l'RPO (utilizzando backup del registro delle transazioni ogni 5 minuti) e l'RTO, ma richiede la manutenzione tramite script manuali e personalizzati. Per un esempio di questa opzione, consultate il seguente diagramma di architettura:

  - Architettura HA/DR a tre nodi con spedizione dei log (multiregione)

- Se si dispone di un'applicazione come un'applicazione SQL Server Reporting Services (SSRS) con una distribuzione scalabile, il load balancer può reindirizzare tutto il traffico verso il nodo secondario.
- Puoi utilizzare Amazon EC2 base AMIs per l'applicazione e il server di database per il provisioning dell'infrastruttura. I database possono essere ripristinati in una nuova zona di disponibilità, a seconda delle dimensioni e della frequenza di backup, dai backup nativi più recenti (backup completo, backup differenziale o backup del registro delle transazioni ogni 5 minuti) o utilizzando le istantanee EBS. Questa opzione soddisfa i requisiti RPO e RTO ma richiede uno script personalizzato. È inoltre necessario considerare il tempo necessario per il provisioning dell'infrastruttura e soddisfare i requisiti RPO e RTO può essere difficile.
- EC2 Le immagini Amazon (inclusi i volumi EBS) per entrambe le applicazioni e il server di database possono essere ripristinate in una nuova zona di disponibilità. L'RPO può essere impegnativo, a seconda del backup più recente, ma questa opzione può essere combinata con i log delle transazioni più recenti per soddisfare i requisiti. Questa opzione supporta le istantanee di Windows Volume Shadow Copy Service (VSS).

Errore nella regione

È possibile utilizzare una delle seguenti opzioni per il ripristino da un errore in una singola AWS regione entro i parametri specificati (RPO di 15 minuti, RTO di 4 ore).

- Puoi utilizzare Amazon Machine Images (AMIs) di EC2 base su Amazon per l'applicazione e il server di database per il provisioning dell'infrastruttura. I database possono essere ripristinati in una nuova regione, a seconda delle dimensioni e della frequenza di backup, dai backup nativi più recenti (backup completo, backup differenziale o backup del registro delle transazioni ogni 5 minuti). Questa opzione soddisfa i requisiti RPO e RTO ma richiede uno script personalizzato.
  - La spedizione dei log di SQL Server come soluzione DR richiede un failover manuale su un server di standby e dipende dalla frequenza dei backup dei log. Questa è una delle opzioni di DR meno costose. Non è necessario che le edizioni di SQL Server per il sito DR principale e

il sito DR fornito con log corrispondano. Questa opzione soddisfa l'RPO (utilizzando i backup del registro delle transazioni ogni 5 minuti) e l'RTO, ma richiede la manutenzione tramite script manuali e personalizzati. I database di grandi dimensioni richiedono tempi di ripristino lunghi.

- Puoi utilizzare un' EC2 AMI Amazon sia per l'applicazione che per il server di database e ripristinarla su una destinazione in una nuova regione. L'RPO dipende dalla dimensione e dalla frequenza dei backup.
  - Le immagini più recenti delle applicazioni possono essere ripristinate utilizzando un'AMI. È possibile utilizzare i backup nativi recenti del registro differenziale o delle transazioni ogni 5 minuti per aggiornare il database in modo da soddisfare l'RPO.
  - L'RTO dipende dalla dimensione e dal tempo di trasferimento e ripristino delle istantanee nella nuova regione, se l'origine non è già sincronizzata con la destinazione.
- La soluzione con il minor tempo di inattività consiste nel ripristinare l'immagine di backup dell'applicazione e disporre di un nodo SQL Server in standby caldo in una regione remota utilizzando una configurazione del gruppo di disponibilità a due, tre o quattro nodi (di base, classica o distribuita) e connettersi al server del database in standby dopo un failover. La replica in modalità commit sincrona soddisfa i requisiti RPO, mentre la replica in modalità commit asincrona potrebbe subire ritardi a seconda del volume delle transazioni. È possibile utilizzare una configurazione di gruppi di disponibilità distribuiti per scalare i nodi del database in una nuova regione, se necessario. Questa configurazione riduce anche la complessità perché utilizza due gruppi di disponibilità indipendenti anziché un singolo gruppo di disponibilità distribuito tra le regioni in modalità di commit sincrono o di commit asincrono e soddisfa perfettamente i requisiti RTO e RPO. In alternativa, è possibile utilizzare anche i gruppi di disponibilità di base di SQL Server nell'edizione Standard. Tuttavia, presenta delle limitazioni perché supporta solo fino a due nodi e solo un database può appartenere a un singolo gruppo di disponibilità, sebbene siano supportati più gruppi di disponibilità. È possibile configurare l'edizione SQL Server Standard all'interno di una o più regioni. Questa edizione offre risparmi sui costi perché non prevede costi per il nodo secondario, che non è accessibile per le operazioni di lettura. L'edizione SQL Server Enterprise offre funzionalità complete e supporta il failover di tutti i database come failover di un singolo gruppo di disponibilità.

Casi di utilizzo comune

Come esercizio di dimensionamento, l'80% delle applicazioni SQL Server in esecuzione su Amazon EC2 con un normale carico di lavoro OLTP (Online Transaction Processing) può essere raggruppato in una delle tre categorie in base alla loro importanza:

- Backup di SQL Server HA/DR con SQL Server, utilizzando due repliche con commit sincrono e una replica in modalità commit asincrona
- AWS Backup HA/DR con backup di SQL Server, utilizzando un' EC2 AMI Amazon sia per l'applicazione che per il database e storage Amazon EBS
- AWS Backup HA/DR con backup di SQL Server, utilizzando un' EC2 AMI di base Amazon per il server di database, un' EC2 immagine Amazon per l'applicazione e istantanee Amazon EBS

La tabella seguente fornisce dettagli su ciascuna categoria.

	SQL Server HA/ DR con backup di SQL Server	AWS Backup HA/ DR con AMIs storage EBS e backup di SQL Server	AWS Backup HA/DR con AMIs, istantanee EBS e backup di SQL Server
Processo di ripristino in caso di emergenza	• Ripristina l'AMI Amazon EC2 base per l'applicazione da AWS Backup • Esegui il failover sull'istanza di standby nella regione (in caso di errore nella zona di disponibi lità) o sull'istanza interregionale (in caso di errore nella regione) • Soddisfa i requisiti RPO e RTO	• Ripristina EC2 le immagini Amazon dai backup sia per l'applicazione che per il database • Fornisce supporto sia a livello regionale che interregionale • Applica i backup differenziali e dei log delle transazio ni più recenti di SQL Server (ogni 15 minuti) per soddisfare i requisiti RPO e RTO per il database	• Ripristina EC2 l'immagine Amazon dal backup per l'applicazione • Ripristina l'AMI di EC2 base Amazon per il server di database • Ripristina le istantanee EBS (se presenti) • Il cluster deve essere ricostruito • Fornisce supporto sia a livello regionale che interregionale • Applica i backup differenziali e dei log delle transazio

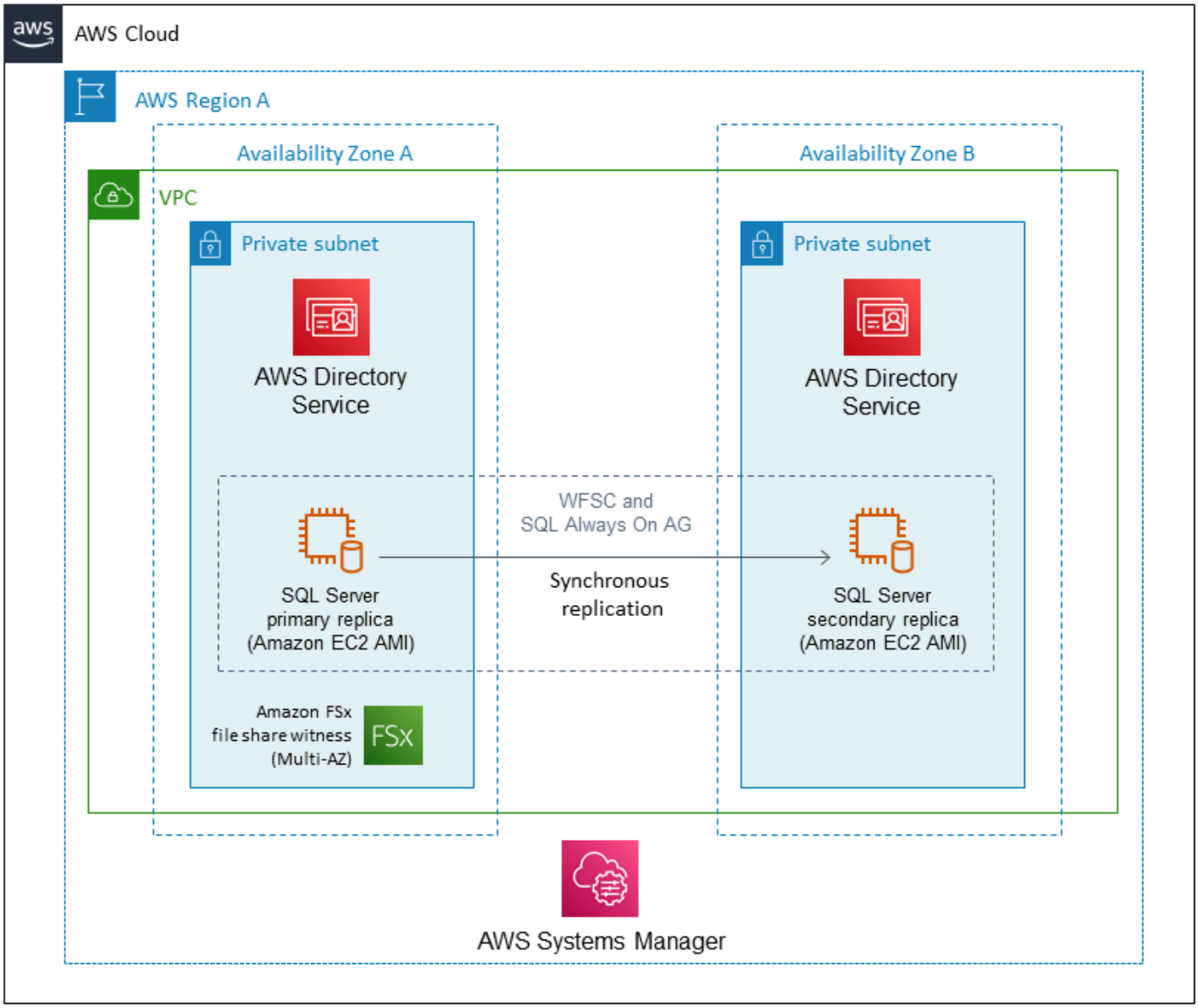
	SQL Server HA/ DR con backup di SQL Server	AWS Backup HA/ DR con AMIs storage EBS e backup di SQL Server	AWS Backup HA/DR con AMIs, istantanee EBS e backup di SQL Server
			ni più recenti al database per soddisfare i requisiti RPO, ma l'RTO potrebbe non essere soddisfatto
Risorse primarie	• Tre licenze SQL Server Enterprise edition (la licenza passiva per nodi HA e DR è gratuita se si dispone di un contratto di licenza Software Assurance esistente con Microsoft; vedi annuncio) • Spazio EC2 di backup Amazon su Amazon Simple Storage Service (Amazon S3) • Trasferimento di dati tra regioni	• Una licenza SQL Server (qualsiasi edizione). • Spazio EC2 di backup Amazon su Amazon S3 • Backup di SQL Server (file differenziali e di registro) su Amazon S3 • Trasferimento di dati tra regioni	• Una licenza SQL Server (qualsiasi edizione). • Spazio EC2 di backup Amazon su Amazon S3 • Backup di SQL Server (file differenziali e di registro) su Amazon S3 • Trasferimento di dati tra regioni
HA/DR	Offre HA e DR	Offre solo DR	Offre solo DR

	SQL Server HA/ DR con backup di SQL Server	AWS Backup HA/ DR con AMIs storage EBS e backup di SQL Server	AWS Backup HA/DR con AMIs, istantanee EBS e backup di SQL Server
RPO	Il failover è gestito dal gruppo di disponibilità di SQL Server (il DR è manuale)	Con script manuale o personalizzato	Con script manuale o personalizzato
RTO	Da secondi a minuti	Da minuti a ore	Più ore
Rischio di scomparsa SLAs	Bassa	Media	Elevata
Gestibilità	Semplice	Media	Media
Dimensionamento	Semplice	Media	Media
Limiti alle dimensioni dei file per i caricamenti su Amazon S3 o il trasferimento tra regioni	N/A: gestito in modalità di commit sincrono o in modalità di commit asincrono in modalità standby a caldo	Sì	Sì
Perdita di dati	Quasi zero (dipende dal carico di lavoro e dall'infrastruttura predisposta)	Dipende dalla frequenza delle immagini di EC2 backup di Amazon e dei backup di SQL Server	Dipende dalla frequenza delle immagini di EC2 backup di Amazon o degli snapshot EBS e dei backup di SQL Server
Costo	Media	Basso — medio	Basso — medio

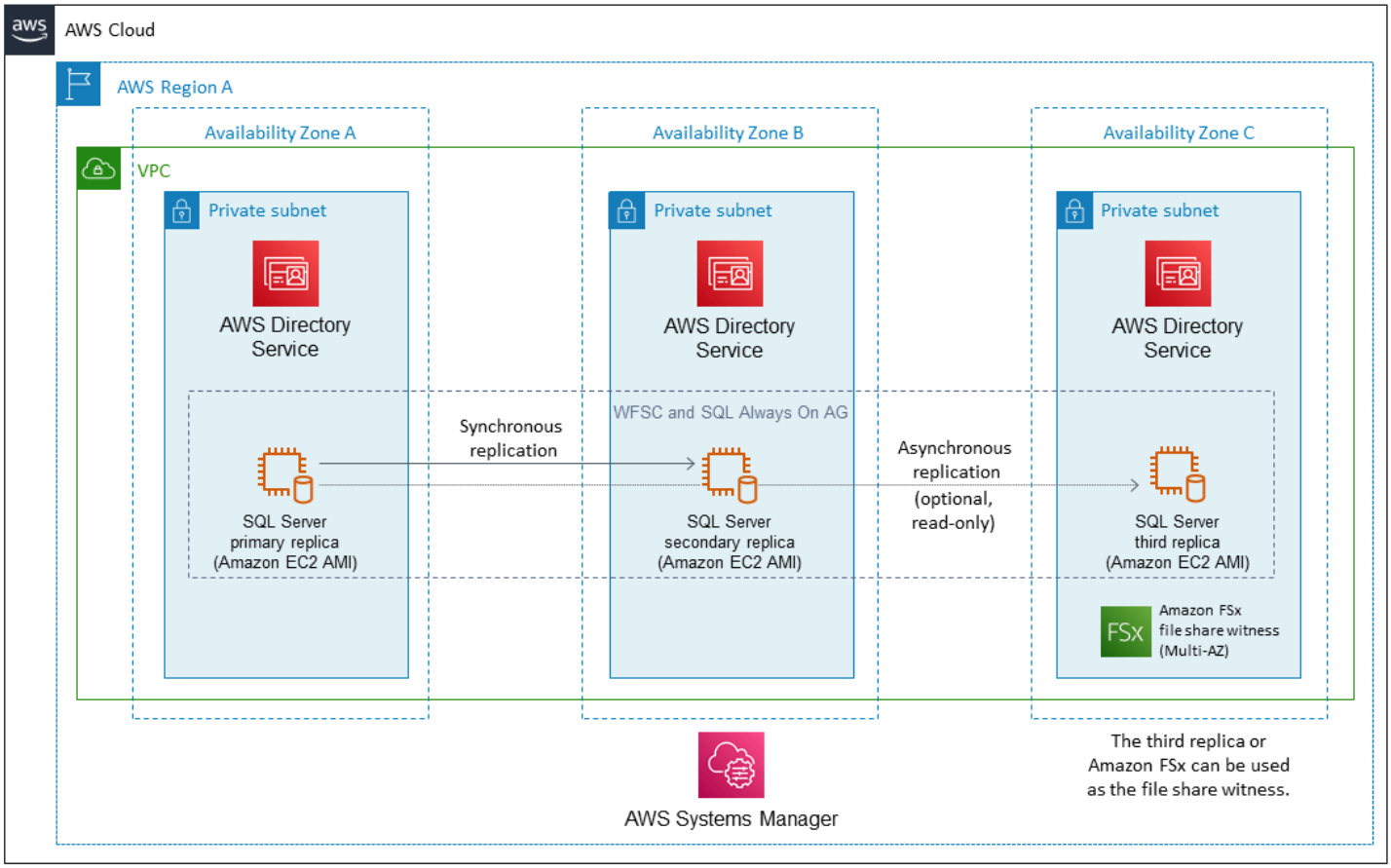
Diagrammi di EC2 architettura SQL Server su Amazon

Questa sezione fornisce diagrammi di architettura che illustrano le strategie HA/DR descritte nelle sezioni precedenti.

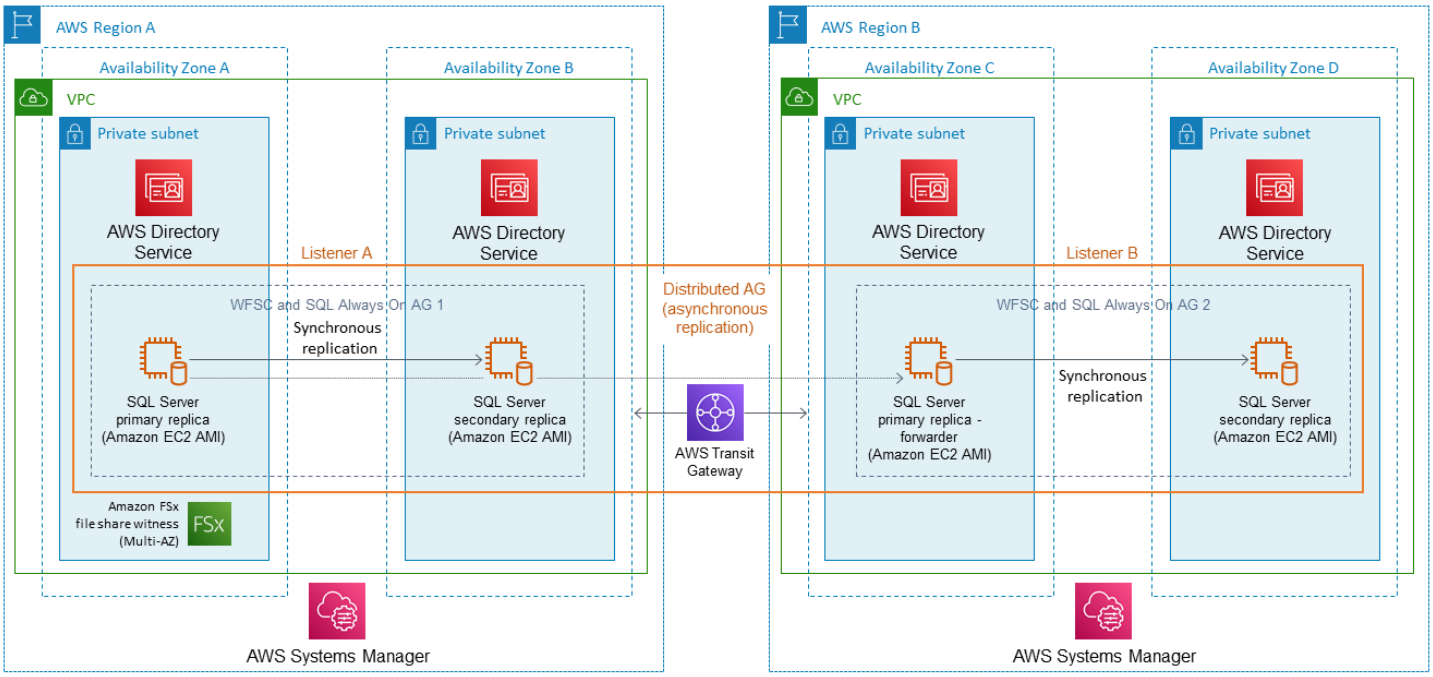
Architettura HA/DR a due nodi con cluster di gruppo Always On Availability (regione singola, Multi-AZ)



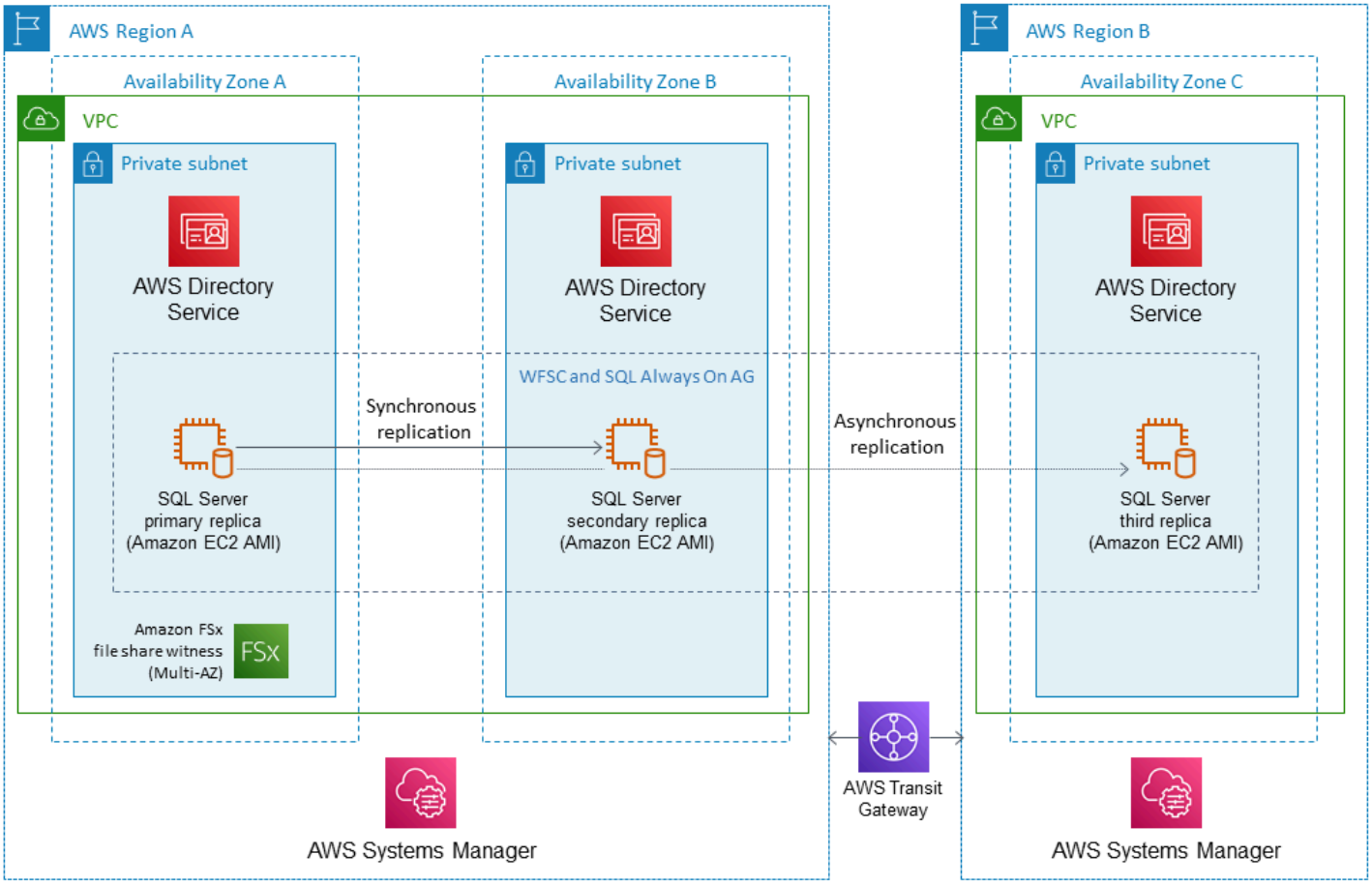
Architettura HA/DR a tre nodi (regione singola, Multi-AZ)



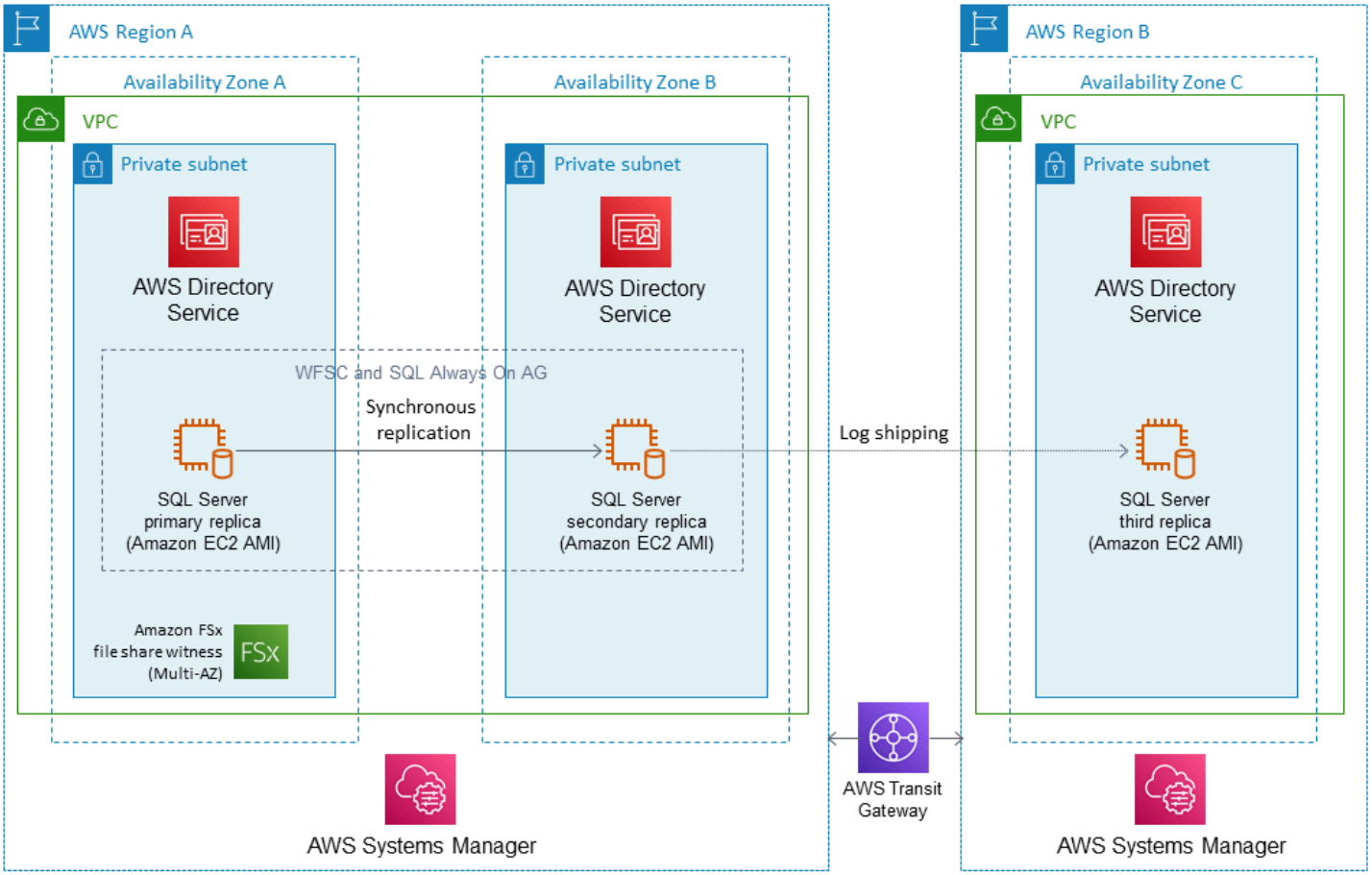
Architettura HA/DR a quattro nodi con cluster di gruppo a disponibilità distribuita Always On (multiregione, Multi-AZ)



Architettura HA/DR a tre nodi con un unico gruppo di disponibilità (multiregione)



Architettura HA/DR a tre nodi con spedizione dei log (multiregione)

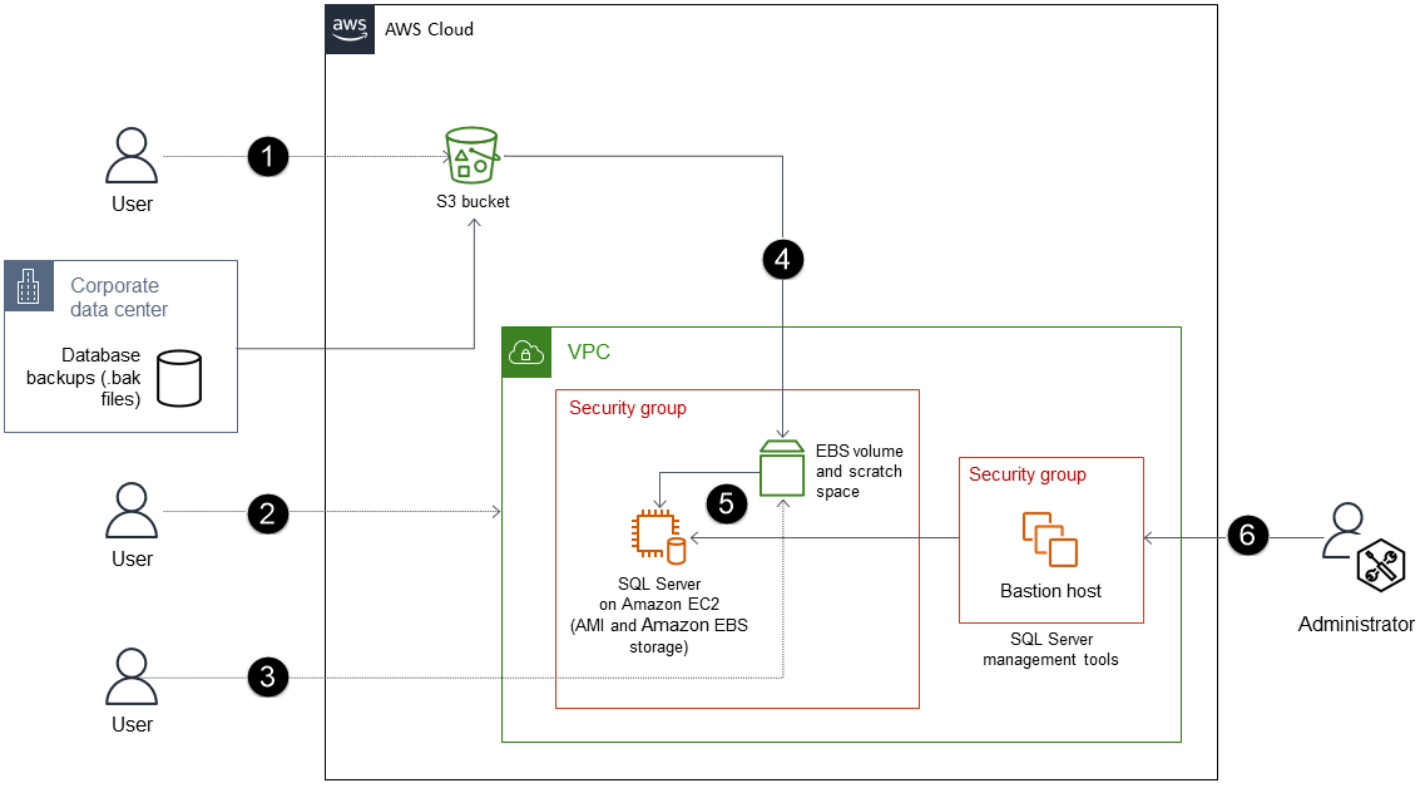


Opzioni di ripristino

Le seguenti sezioni forniscono due opzioni di ripristino del database per SQL Server su Amazon Elastic Compute Cloud (Amazon EC2), quando i backup sono locali.

Uso di Amazon S3

Questo approccio di ripristino del database SQL Server utilizza i comandi Amazon Simple Storage Service (Amazon S3) per AWS CLI() o AWS Command Line Interface l'API Amazon S3 per caricare i file di backup direttamente in un bucket S3.



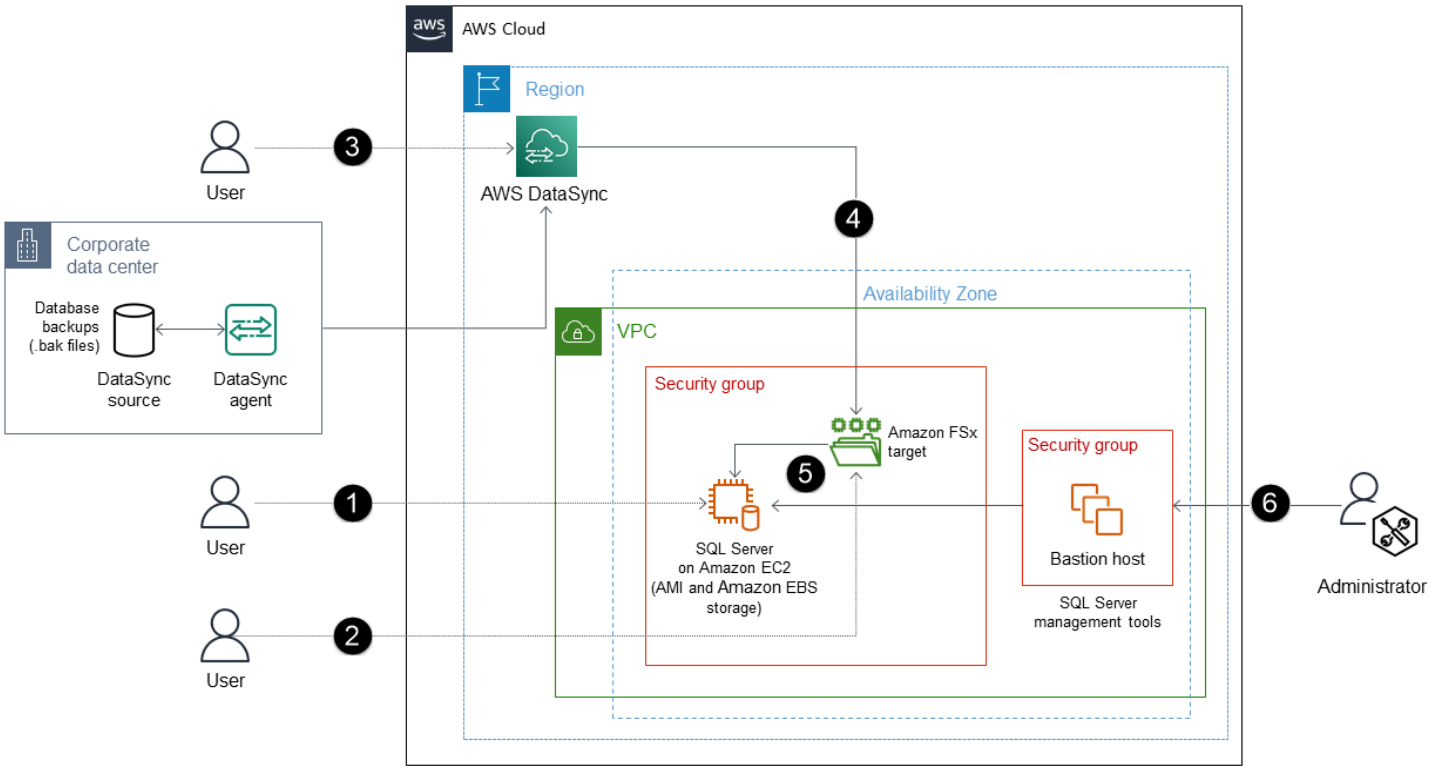
Il processo prevede i seguenti passaggi:

1. Crea un bucket S3 (o usa un bucket esistente) per archiviare i file di backup e trasferisci i file di backup (.bak) dal database locale al bucket S3 utilizzando la CLI o l'API Amazon S3. AWS
2. Distribuisci SQL Server su un' EC2 istanza ottimizzata per EBS, utilizzando un'Amazon Machine Image (AMI) di SQL Server. Questo AMI deve contenere volumi EBS configurati con una partizione del sistema operativo, una partizione DATA, una partizione LOG, uno storage tempdb (NVMe) e uno spazio scratch.

3. (Facoltativo) Collegate un volume EBS non root all'istanza. EC2
4. Copia i file di backup nel volume EBS non root.
5. Ripristina i file di backup dal volume EBS a SQL Server sull'istanza. EC2
6. Usa gli strumenti di gestione di SQL Server per gestire il tuo database.

Utilizzo AWS DataSync e Amazon FSx

Questo approccio di ripristino del database di SQL Server utilizza AWS DataSync il trasferimento dei file di backup su Amazon FSx for Windows File Server.



Il processo prevede i seguenti passaggi:

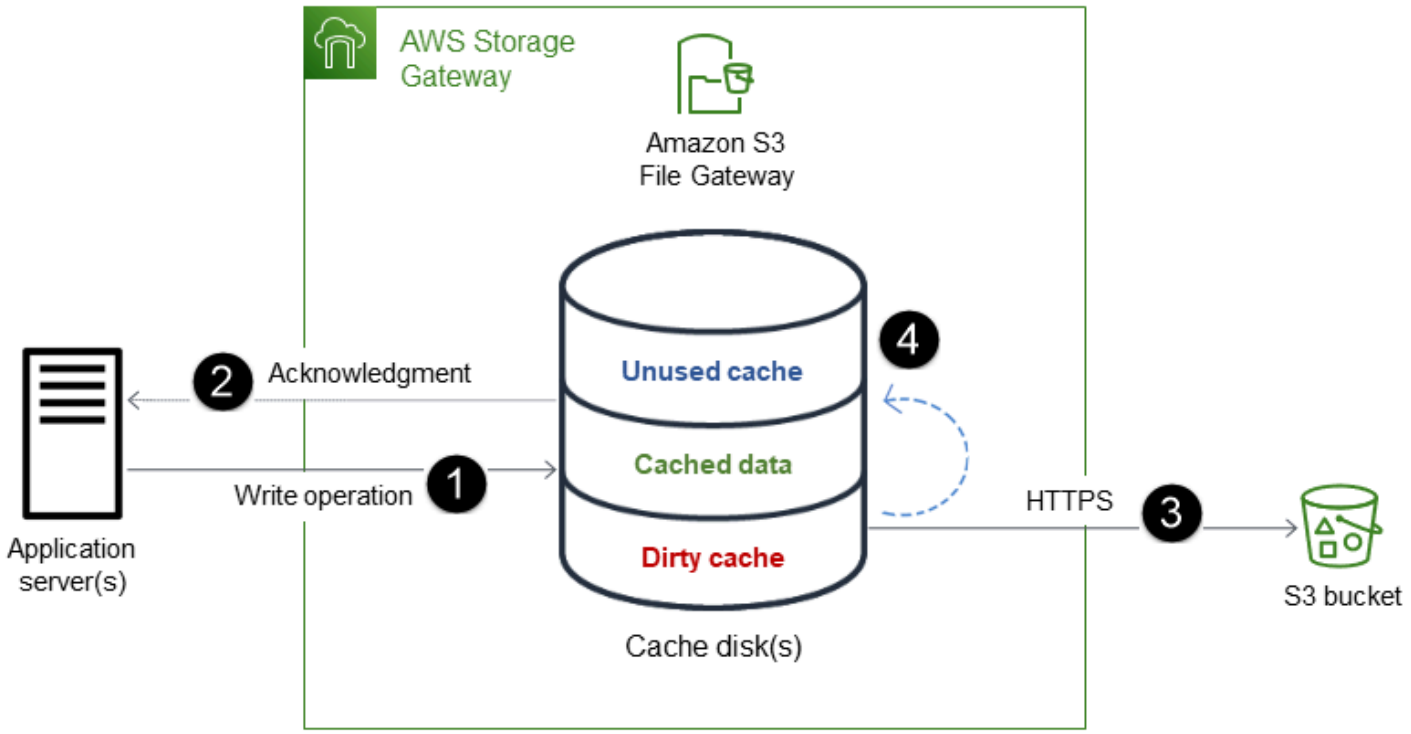
1. Distribuisci SQL Server su un' EC2 istanza ottimizzata per EBS con allegato NVMe, utilizzando un'AMI che contiene volumi EBS configurati con OS, DATA, LOG e tempdb. (Ad esempio, puoi utilizzare la classe di istanza ottimizzata per la memoria.) `r5d.large`
2. Utilizzare FSx per Windows File Server per creare un file server. Può essere utilizzato come posizione di archiviazione temporanea per scaricare i file di backup di SQL Server (.bak) dall'ambiente locale.

3. Crea un DataSync endpoint e un agente per il FSx file server Amazon.
4. DataSync automatizza la sincronizzazione dei dati tra lo storage locale e il FSx file server Amazon senza richiedere Amazon S3.
5. Ripristina i file di backup dal FSx file server Amazon a SQL Server sull' EC2istanza.
6. Usa gli strumenti di gestione di SQL Server per gestire il tuo database.

> **Note**
>
> Amazon EC2 offre Microsoft SQL Server su Microsoft Windows Server AMIs per più edizioni di SQL Server.

Utilizzo di Amazon S3 File Gateway

Puoi utilizzare Amazon S3 File Gateway per archiviare backup nativi di SQL Server su Amazon S3, come illustrato nel diagramma seguente. In alternativa, esistono strumenti come Commvault LiteSpeedche consentono di gestire i backup a livello di file su larga scala e di archiviarli direttamente in Amazon S3. Puoi anche utilizzare uno strumento come DataKeeperSIOS per il backup/ripristino e la configurazione del DR.



Il processo prevede i seguenti passaggi:

1. I dati vengono scritti sul disco di cache locale del gateway di file.
2. Dopo che i dati sono stati salvati in modo sicuro nella cache locale, il file gateway conferma il completamento dell'operazione di scrittura nell'applicazione client.
3. Il file gateway trasferisce i dati al bucket S3 in modo asincrono. Ottimizza il trasferimento dei dati e utilizza HTTPS per crittografare i dati in transito.
4. Una volta caricati nel bucket S3, i dati rimangono nella cache locale del file gateway fino a quando non vengono eliminati.

Risorse e passaggi successivi

Questa guida illustra le migliori pratiche per il ripristino di emergenza rapido dei database SQL Server. I consigli includono l'utilizzo di immagini per ripristinare l'istanza dell'applicazione e l'uso di metodi SQL nativi per ripristinare il database o, preferibilmente, per eseguire il failover del database. A differenza dei ripristini di database di grandi dimensioni che possono richiedere ore, l'utilizzo dei backup di Amazon Elastic Compute Cloud (EC2Amazon) di Amazon Machine Image (AMI) in combinazione con i log delle transazioni più recenti consente di soddisfare i requisiti di Recovery Point Objective (RPO) e Recovery Time Objective (RTO) mantenendo bassi i costi complessivi. L'approccio ottimale dipende dalla dimensione del database, dal numero e dalla natura dei backup e dalla frequenza dei backup dei log delle transazioni per i quali è necessario progettare una strategia di disaster recovery. Consulta i seguenti link per ulteriori informazioni, best practice, guide Quick Start e linee guida prescrittive sulla migrazione e l'hosting di SQL Server su Amazon. EC2

Documentazione

- Best practice e consigli per il clustering di SQL Server su Amazon EC2 (EC2documentazione Amazon)
- Amazon EC2 instance store (EC2 documentazione Amazon)
- Replica di oggetti (documentazione Amazon S3)
- Ripristino rapido degli snapshot di Amazon EBS (documentazione Amazon EC2)
- SQL Server con replica Always On su Cloud AWS(distribuzione di riferimento Quick Start)
- Tipi di volume Amazon EBS (EC2 documentazione Amazon)
- Utilizzo FSx per Windows File Server con Microsoft SQL Server (FSx documentazione Amazon)
- Che cos'è AWS Backup? (AWS Backup documentazione)
- AWS Windows AMIs (EC2 documentazione Amazon)

AWS Guida prescrittiva

- Le migliori pratiche per la distribuzione di Microsoft SQL Server su Amazon EC2
- EC2 Backup e ripristino Amazon con istantanee e AMIs
- Posiziona tempdb in un instance store

Post e notizie sul blog

- Archivia facilmente i tuoi backup di SQL Server in Amazon S3 utilizzando File Gateway
- Monitora i costi di trasferimento dei dati relativi alla replica di Amazon S3
- Distribuzione di SQL Server in più regioni utilizzando gruppi di disponibilità distribuiti
- Note sul campo: Creazione di un'architettura multiregionale per SQL Server utilizzando FCI e gruppi di disponibilità distribuiti
- Amazon offre EC2 ora Microsoft SQL Server su Microsoft Windows Server 2022 AMIs

Documentazione di SQL Server

- Edizioni e funzionalità supportate di SQL Server

Appendice: Tipi di storage SSD Amazon EBS

Amazon Elastic Block Store (Amazon EBS) fornisce i seguenti volumi supportati da unità a stato solido (SSD). Per le informazioni più recenti, consulta i tipi di volume Amazon EBS nella EC2 documentazione di Amazon.

	General Purpose SSD		Provisioned IOPS SSD		
Tipo di volume	`gp3`	`gp2`	`io2 Block Express`[1]	`io2`	`io1`
Durabilità	Durabilità dal 99,8% al 99,9% (tasso di fallimento annuo dallo 0,1% allo 0,2%)	Durabilità del 99,8% - 99,9% (tasso di fallimento annuo dello 0,1% - 0,2%)	Durata del 99,999% (tasso di fallimento annuo dello 0,001%)	Durata del 99,999% (tasso di fallimento annuo dello 0,001%)	Durabilità del 99,8% - 99,9% (tasso di fallimento annuo dello 0,1% - 0,2%)
Casi d'uso	• App interattive a bassa latenza • Ambienti di sviluppo e test	• App interattive a bassa latenza • Ambienti di sviluppo e test	Carichi di lavoro che richiedono: • Latenza media inferiore al millisecondo • Prestazioni IOPS sostenute • Oltre 64.000 IOPS o 1.000	• Carichi di lavoro che richiedono prestazioni IOPS sostenute o superiori a 16.000 IOPs • Carichi di lavoro del database con uso intensivo di I/O	• Carichi di lavoro che richiedono prestazioni IOPS sostenute o superiori a 16.000 IOPs • Carichi di lavoro del database con uso intensivo di I/O

	General Purpose SSD		Provisioned IOPS SSD		
			MiB/s di throughput		
Volume size (Dimensione dei volumi)	Da 1 GiB a 16 TiB	Da 1 GiB a 16 TiB	4 GiB — 64 TiB	4 GiB — 16 TiB	4 GiB — 16 TiB
Numero massimo di IOPS per volume (16 KiB I/O)	16,000	16,000	256.000	64.000 2	64.000 2
Velocità effettiva massima per volume	1.000 MiB/s	250 MiB/s 3	4.000 MiB/s	1.000 MiB/s 2	1.000 MiB/s 2
Multi-Attacchi Amazon EBS	Non supportato	Non supportato	Supportato	Supportato	Supportata
Volume di avvio	Supportato	Supportato	Supportato	Supportato	Supportato

[1] `io2 Block Express` volume è supportato solo con istanze R5b. `io2`i volumi collegati a un'`R5`bistanza durante o dopo l'avvio vengono eseguiti automaticamente su Block Express. Per ulteriori informazioni, consulta i volumi io2 Block Express nella EC2 documentazione di Amazon.

[2] Gli IOPS e il throughput massimi sono garantiti solo sulle istanze basate sul sistema Nitro con più di 32.000 IOPS. Altre istanze garantiscono fino a 32.000 IOPS e 500 MiB/s. I volumi `io1` creati prima del 6 dicembre 2017 e che non sono stati modificati dopo la creazione potrebbero non raggiungere prestazioni complete a meno che non si modifichi il volume.

[3] Il limite di throughput è compreso tra 128 indipendentemente dai crediti burst. MiB/s and 250 MiB/s, depending on the volume size. Volumes smaller than or equal to 170 GiB deliver a maximum throughput of 128 MiB/s. Volumes larger than 170 GiB but smaller than 334 GiB deliver a maximum

throughput of 250 MiB/s if burst credits are available. Volumes larger than or equal to 334 GiB deliver 250 MiB/s gp2i volumi creati prima del 3 dicembre 2018 e che non sono stati modificati dopo la creazione potrebbero non raggiungere le prestazioni complete a meno che non si modifichi il volume.

Cronologia dei documenti

La tabella seguente descrive le modifiche significative apportate a questa guida. Per ricevere notifiche sugli aggiornamenti futuri, puoi abbonarti a un feed RSS.

Modifica	Descrizione	Data
Pubblicazione iniziale	—	28 febbraio 2022

AWS Glossario delle linee guida prescrittive

I seguenti sono termini di uso comune 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: migra il tuo database Oracle locale all'edizione compatibile con Amazon Aurora 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 ad Amazon Relational Database Service (Amazon RDS) per Oracle in. 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: migra il 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 database Oracle locale su Oracle su un'istanza in. EC2 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. Si esegue la migrazione dei server da una piattaforma locale a un servizio cloud per la stessa piattaforma. Esempio: migra un'applicazione suMicrosoft Hyper-V. 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 AIOps viene utilizzato nella strategia di AWS migrazione, 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 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.

CCoE

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 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 gli CCoE post sull' Cloud AWS Enterprise Strategy Blog.

cloud computing

La tecnologia cloud generalmente utilizzata per l'archiviazione remota di dati e la gestione dei dispositivi IoT. Il cloud computing è generalmente collegato 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 verso Cloud AWS:

- Progetto: esecuzione di alcuni progetti relativi al cloud per scopi di dimostrazione e apprendimento
- Fondamento: effettuare investimenti fondamentali per scalare l'adozione del cloud (ad esempio, creazione di una landing zone, definizione di una CCo E, 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 Enterprise Strategy. Cloud AWS 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 oBitbucket Cloud. 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, Amazon SageMaker AI 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, compilazione, test, gestione temporanea e produzione del processo di rilascio del software. CI/CD is commonly described as a pipeline. CI/CDpuò aiutarvi ad automatizzare i processi, migliorare la produttività, migliorare la qualità del codice e velocizzare le consegne. 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.

MODIFICA

Vedi scambio elettronico di 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.

scambio elettronico di dati (EDI)

Lo scambio automatizzato di documenti aziendali tra organizzazioni. Per ulteriori informazioni, vedere Cos'è lo scambio elettronico di dati.

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 con 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, consulta 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.

prompt con pochi scatti

Fornire a un LLM un numero limitato di esempi che dimostrino l'attività e il risultato desiderato prima di chiedergli di eseguire un'attività simile. Questa tecnica è un'applicazione dell'apprendimento contestuale, in cui i modelli imparano da esempi (immagini) incorporati nei prompt. I prompt con pochi passaggi possono essere efficaci per attività che richiedono una formattazione, un ragionamento o una conoscenza del dominio specifici. Vedi anche zero-shot prompting.

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.

FM

Vedi il modello di base.

modello di fondazione (FM)

Una grande rete neurale di deep learning che si è addestrata su enormi set di dati generalizzati e non etichettati. FMs sono in grado di svolgere un'ampia varietà di attività generali, come comprendere il linguaggio, generare testo e immagini e conversare in linguaggio naturale. Per ulteriori informazioni, consulta Cosa sono i modelli Foundation.

G

AI generativa

Un sottoinsieme di modelli di intelligenza artificiale che sono stati addestrati su grandi quantità di dati e che possono utilizzare un semplice prompt di testo per creare nuovi contenuti e artefatti, come immagini, video, testo e audio. Per ulteriori informazioni, consulta Cos'è l'IA generativa.

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.

immagine dorata

Un'istantanea di un sistema o di un software che viene utilizzata come modello per distribuire nuove istanze di quel sistema o software. Ad esempio, nella produzione, un'immagine dorata può essere utilizzata per fornire software su più dispositivi e contribuire a migliorare la velocità, la scalabilità e la produttività nelle operazioni di produzione dei dispositivi.

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 alto livello che aiuta a governare le risorse, le politiche e la conformità tra le unità organizzative (). OUs 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.

dati di esclusione

Una parte di dati storici etichettati che viene trattenuta da un set di dati utilizzata per addestrare un modello di apprendimento automatico. È possibile utilizzare i dati di holdout per valutare le prestazioni del modello confrontando le previsioni del modello con i dati di holdout.

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. La AWS Security Reference Architecture consiglia di configurare l'account di rete con funzionalità in entrata, in uscita e di ispezione VPCs per proteggere l'interfaccia bidirezionale tra l'applicazione e la rete 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.

IIoInternet delle cose industriale (T)

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, vedere Creazione di una strategia di trasformazione digitale per l'Internet of Things (IIoT) industriale.

VPC di ispezione

In un'architettura AWS multi-account, un VPC centralizzato che gestisce le ispezioni del traffico di rete tra VPCs (nello stesso o in modo diverso Regioni AWS), Internet e le reti locali. La AWS

Security Reference Architecture consiglia di configurare l'account di rete con informazioni in entrata, in uscita e di ispezione VPCs 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, vedere Interpretabilità del modello 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.

modello linguistico di grandi dimensioni (LLM)

Un modello di intelligenza artificiale di deep learning preaddestrato su una grande quantità di dati. Un LLM può svolgere più attività, come rispondere a domande, riepilogare documenti, tradurre testo in altre lingue e completare frasi. Per ulteriori informazioni, consulta Cosa sono. LLMs

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.

LLM

Vedi modello linguistico di grandi dimensioni.

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 servizio piccolo e indipendente che comunica tramite canali ben definiti ed è in genere di proprietà di piccoli team autonomi. APIs 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 attraverso un'interfaccia ben definita utilizzando sistemi leggeri. APIs 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 in genere includono addetti alle 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 con AWS Application Migration Service.

Valutazione del portfolio di migrazione (MPA)

Uno strumento online che fornisce informazioni per la convalida del business case per la migrazione a. 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 verso. 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, vedere Valutazione della preparazione alla modernizzazione per le applicazioni in. 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.

ORR

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. La AWS Security Reference Architecture consiglia di configurare l'account di rete con funzionalità in entrata, in uscita e di ispezione VPCs 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 di ingegneria dei sistemi che tiene conto della privacy durante l'intero processo di sviluppo.

zone ospitate private

Un contenitore che contiene informazioni su come desideri che Amazon Route 53 risponda alle query DNS per un dominio e i relativi sottodomini all'interno di uno o più. VPCs 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.

concatenamento rapido

Utilizzo dell'output di un prompt LLM come input per il prompt successivo per generare risposte migliori. Questa tecnica viene utilizzata per suddividere un'attività complessa in sottoattività o per perfezionare o espandere iterativamente una risposta preliminare. Aiuta a migliorare l'accuratezza e la pertinenza delle risposte di un modello e consente risultati più granulari e personalizzati.

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.

publish/subscribe (pub/sub)

Un modello 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).

STRACCIO

Vedi Retrieval Augmented Generation.

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. Questo determina ciò che si considera 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.

Retrieval Augmented Generation (RAG)

Una tecnologia di intelligenza artificiale generativa in cui un LLM fa riferimento a una fonte di dati autorevole esterna alle sue fonti di dati di formazione prima di generare una risposta. Ad esempio, un modello RAG potrebbe eseguire una ricerca semantica nella knowledge base o nei dati personalizzati di un'organizzazione. Per ulteriori informazioni, consulta Cos'è il RAG.

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 Cosa c'è in un segreto di Secrets Manager? nella documentazione di Secrets Manager.

sicurezza fin dalla progettazione

Un approccio di ingegneria dei sistemi che tiene conto della sicurezza durante l'intero processo di sviluppo.

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 EC2 Amazon 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 politica che fornisce il controllo centralizzato sulle autorizzazioni per tutti gli account di un'organizzazione in. AWS Organizations SCPs definire barriere o fissare limiti alle azioni che un amministratore può delegare a utenti o ruoli. È possibile utilizzarli SCPs 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.

prompt di sistema

Una tecnica per fornire contesto, istruzioni o linee guida a un LLM per indirizzarne il comportamento. I prompt di sistema aiutano a impostare il contesto e stabilire regole per le interazioni con gli utenti.

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 puoi utilizzare per interconnettere le tue reti VPCs e quelle locali. 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 VPCs che consente di indirizzare il traffico utilizzando 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.

prompt zero-shot

Fornire a un LLM le istruzioni per eseguire un'attività ma non esempi (immagini) che possano aiutarla. Il LLM deve utilizzare le sue conoscenze pre-addestrate per gestire l'attività. L'efficacia del prompt zero-shot dipende dalla complessità dell'attività e dalla qualità del prompt. Vedi anche few-shot prompting.

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.