Opzioni di connettività di rete attive AWS per le offerte SaaS

AWS Guida prescrittiva

AWS Guida prescrittiva: Opzioni di connettività di rete attive AWS per le offerte SaaS

Table of Contents

Opzioni di connettività di rete attive AWS per le offerte SaaS

Tomas Sykora e Luca Schumann, Amazon Web Services

Settembre 2025 (storia del documento)

Questa guida esplora gli scenari comuni per connettere le applicazioni consumer ai provider di software as a service (SaaS). Descrive come connettersi a risorse locali, nei cloud di altri provider di servizi cloud (CSP) o in architetture ibride. Cloud AWS Questi scenari includono quanto segue:

- Esposizione di servizi Web tramite HTTPS
- Esposizione di servizi basati su TCP
- Utilizzo AWS AppSyncper implementare publish-subscribe (Pub/Sub) e GraphQL APIs
- Utilizzo di risorse AWS per l'esposizione per applicazioni in tempo reale WebSockets
- Consentire l'accesso bidirezionale per la comunicazione interattiva dei servizi

Allineandosi alle best practice illustrate in questa guida, i provider SaaS possono promuovere la fiducia dei clienti e supportare un accesso scalabile, sicuro e resiliente alle offerte SaaS.

Questa guida include anche criteri di autovalutazione per aiutarti a valutare in che misura stai soddisfacendo i requisiti di rete dei consumatori per la tua offerta SaaS. Oltre ai modelli di connettività, troverai confronti completi tra i servizi di AWS rete, diagrammi architettonici di alto livello per vari scenari di implementazione e linee guida pratiche su come selezionare l'approccio giusto in base al contesto aziendale specifico. La guida esplora le considerazioni sulla sicurezza per ciascuna opzione di rete, illustra le insidie comuni da evitare e fornisce consigli di implementazione che bilanciano i requisiti tecnici con l'efficienza operativa. Inoltre, troverai quadri strategici per allineare le tue decisioni di rete al tuo modello di business, agli obiettivi di crescita e alle esigenze di conformità normativa.

Destinatari principali

Questa guida è destinata ai provider SaaS. Aiuta gli architetti del cloud, i product manager e gli ingegneri di rete che progettano, implementano e ottimizzano la connettività di rete per le offerte SaaS nel. Cloud AWS Per comprendere i concetti e le raccomandazioni di questa guida, è necessario conoscere AWS i fondamenti, i concetti base del SaaS e i principi di rete di alto livello.

Obiettivi

Questa guida illustra le opzioni di architettura di rete e le best practice testate sul campo che aiutano i consumatori a ottimizzare l'accesso alle offerte SaaS. L'implementazione delle raccomandazioni contenute in questa guida supporta quanto segue:

- Facilità di integrazione: offrite al cliente un percorso semplice, dall'onboarding alla produzione, in modo da accelerare il time-to-value dei clienti e abbreviare il loro ciclo di riconoscimento dei ricavi.
- Adattabilità: si integra perfettamente con le infrastrutture di rete esistenti dei clienti adattandosi alle loro esigenze in evoluzione. Ciò migliora la proposta di valore del prodotto.
- Costo totale di proprietà: standardizza l'accesso alla rete per ridurre i costi di modifica e i costi per tenant. Migliorando la coerenza dell'implementazione, è inoltre possibile ridurre il tempo necessario per eseguire l'analisi o la riparazione delle cause alla radice.
- Gestione delle dipendenze: comprendi le dipendenze, le implicazioni a lungo termine e i compromessi delle diverse opzioni di accesso alla rete. Questo aiuta i leader di prodotto a prendere decisioni informate sui prodotti.
- Componibilità ed estensibilità: separa lo sviluppo delle funzionalità di base dall'infrastruttura operativa. Questo aiuta i team di sviluppo ad agire più velocemente e a concentrarsi sulla creazione di valore per i clienti.
- Promuovi la fiducia: fornendo un accesso resiliente, tollerante ai guasti, sicuro e scalabile alle offerte SaaS, puoi ridurre i rischi normativi e guadagnare fiducia nella tua capacità di supportare la crescita dei tuoi clienti.

Valutazione delle decisioni di accesso alla rete per le offerte SaaS

Comprendi il tuo mercato

Le decisioni che prendi ora sul networking determinano se la proposta di valore del tuo prodotto SaaS può essere fornita ai tuoi clienti. Nonostante l'importanza strategica di queste decisioni, fornire l'accesso all'offerta SaaS viene spesso percepito come un argomento puramente tecnologico. Il rischio che comporta questa percezione include cicli prolungati di riconoscimento dei ricavi, inefficienze operative e disallineamento con la strategia aziendale. Ad esempio, se la rapida espansione è un obiettivo aziendale strategico, il processo decisionale dovrebbe essere incentrato sulla scalabilità e la flessibilità delle soluzioni prese in considerazione per supportare l'espansione. Anche se riesci a far crescere la tua attività, le spese generali operative non devono diventare un ostacolo per le future crescite e una struttura dei costi disallineata potrebbe consumare tutti i tuoi profitti.

Ad esempio, considerate in che modo le seguenti considerazioni di mercato influiscono sugli aspetti tecnici del prodotto, come il networking:

- Se il vostro modello di business è basato sugli abbonamenti, è probabile che i vostri clienti preferiscano soluzioni con costi prevedibili e ricorrenti piuttosto che ingenti investimenti iniziali.
- Se la tua strategia aziendale si rivolge a clienti di alto valore e di livello aziendale, i criteri di sicurezza, governance e conformità normativa determinano se la tua offerta SaaS verrà presa in considerazione.
- Se il mercato di riferimento è costituito principalmente da startup, la facilità di integrazione, il time-to-value e l'adattabilità sono probabilmente fattori importanti. Le startup in genere danno priorità alla velocità e all'agilità. Poiché devono costruire un marchio e generare profitti rapidamente, è probabile che preferiscano soluzioni veloci e facili da integrare, in grado di scalare a costi contenuti, ridurre la dipendenza dagli esperti e che non blocchino cicli preziosi.
- Alcune aziende richiedono un accesso stabile, ad alta velocità e a bassa latenza. Ciò include l'industria dell'intrattenimento e dei media, la produzione e l'elaborazione delle transazioni finanziarie. Se questi sono i vostri clienti target, l'affidabilità è la loro principale preoccupazione.

In tutti questi casi, i clienti potrebbero percepire un'offerta SaaS altrimenti valida se l'accesso alla rete non è semplice. Se la rete diventa un ostacolo, ciò non supporta la vostra tesi aziendale. Se i

tuoi clienti non possono accedere in modo affidabile ai servizi che offri, la proposta di valore delle tue offerte SaaS è nulla.

Comprendere il proprio ruolo

Il vostro ruolo nel sostenere gli obiettivi aziendali dipende da chi siete, quali sono i vostri obiettivi specifici individuali e di team, da chi sono i vostri clienti e da cosa è importante per loro. Anche se non fai parte di un team che in genere interagisce con i clienti, devi preoccuparti di chi sono e di cosa hanno bisogno. I team di progettazione e sviluppo devono inoltre preoccuparsi dei propri clienti interni, in particolare di quelli con cui interagiscono regolarmente. In genere, si tratta dei team addetti alle operazioni e al successo dei clienti.

Se fai parte di un'organizzazione di vendita, è essenziale comunicare con i team di prodotto e di progettazione in merito al networking, anche se si tratta di un argomento apparentemente puramente tecnologico. Condividi approfondimenti sulla struttura del mercato di riferimento. Comunica i punti deboli e le esigenze dei tuoi clienti e partner esistenti e potenziali. Condividi dati e aneddoti sulle opportunità mancate, sulla crescita prevista per segmento e sugli eventi. Poni domande che mettano alla prova la capacità della tua organizzazione di supportare la crescita aziendale. Ciò aumenta il numero di opportunità e migliora la redditività a lungo termine della vostra azienda. In definitiva, questo aiuta la tua organizzazione a finanziare l'espansione e lo sviluppo futuri.

Se fai parte di un'organizzazione ingegneristica, comprendi la strategia aziendale della tua organizzazione prima di tentare di elaborare una soluzione. L'allineamento con la strategia aziendale consente di scegliere le metriche giuste per valutare diverse opzioni di accesso alla rete. Può anche impedire una costosa riprogettazione della rete su larga scala man mano che l'organizzazione cresce. L'allineamento aziendale aiuta il team a proteggere e conservare le risorse necessarie per le sfide future. L'organico del tuo team, il budget per lo sviluppo professionale o l'accesso a tecnologie all'avanguardia dipenderanno dalla tua capacità di dimostrare l'allineamento aziendale. Idealmente, puoi mostrare in che modo le tue decisioni hanno contribuito al successo aziendale dell'organizzazione. Pertanto, ti suggeriamo di tenere conto del processo decisionale, compresi i criteri di selezione metrica. Esamina periodicamente le tue metriche per confermare che siano in linea con gli obiettivi aziendali. Questo può aiutare il tuo team a ottenere il credito che merita. Le revisioni periodiche aiutano anche a verificare che il team non stia prendendo decisioni basate su ipotesi o ragioni storiche obsolete.

L'elenco delle metriche nelle seguenti sezioni è rilevante per l'accesso alla rete:

- Metriche commerciali e di prodotto

- Metriche ingegneristiche che influenzano le decisioni di rete

Questa guida utilizza un sottoinsieme di queste metriche per aiutarti a identificare gli approcci di accesso alla rete ottimali per le tue offerte SaaS. Scegli le metriche più importanti e pertinenti per la tua attività, quindi valuta gli approcci in base a tali metriche.

Metriche commerciali e di prodotto che influenzano le decisioni relative al networking

I team commerciali e di prodotto utilizzano criteri di successo per valutare se stanno raggiungendo gli obiettivi aziendali. Questa sezione descrive le metriche commerciali o di prodotto che possono essere influenzate positivamente o negativamente dalle decisioni di accesso alla rete prese dall'organizzazione.

Utilizzate queste metriche e domande di autovalutazione per valutare in che modo il vostro approccio all'accesso alla rete si allinea al posizionamento aziendale e alla strategia di mercato. Questa valutazione vi aiuta a determinare se le vostre attuali decisioni di rete supportano la differenziazione di mercato, i vantaggi competitivi e le esigenze del pubblico di destinazione della vostra azienda.

Questa sezione contiene metriche e domande di autovalutazione per i seguenti argomenti:

- Modello di business e posizionamento sul mercato
- Mercato totale indirizzabile, tasso di acquisizione di nuovi clienti, crescita e scalabilità
- Esperienza e conservazione dei clienti
- Efficienza e performance finanziaria
- Conformità normativa e gestione del rischio
- Strategia dei partner

Modello di business e posizionamento sul mercato

Queste metriche si riferiscono alla posizione dell'azienda sul mercato, tra cui la differenziazione competitiva, la portata del mercato e la percezione del marchio. È fondamentale valutare l'allineamento tra l'approccio di accesso alla rete e il modello di business. Esegui una valutazione indipendentemente dal fatto che sia basata su abbonamento, basata sull'utilizzo, freemium, su più livelli, marketplace, API-first o white label. Assicurati che il modello supporti gli obiettivi dell'organizzazione e gli obiettivi dei clienti.

Criteri con punteggi elevati

L'approccio all'accesso alla rete si allinea perfettamente al modello di business. Facilita l'adozione e l'erogazione del servizio. Supporta la sostenibilità finanziaria a lungo termine del modello di business e la struttura dei costi è compatibile con la crescita prevista. Riduce al minimo qualsiasi attrito per clienti o partner nell'adozione dell'offerta. Ciò migliora l'esperienza dell'utente e incoraggia una più ampia diffusione del servizio.

Indicatori a basso punteggio

L'approccio di accesso alla rete selezionato non è in linea con il modello di business che dovrebbe supportare. La struttura dei costi e i tempi di implementazione rappresentano un ostacolo all'adozione nel mercato di riferimento. I costi operativi e infrastrutturali correnti inibiscono qualsiasi potenziale profitto. Ciò impedisce la crescita aziendale e rende difficile operare alla scala prevista. In alternativa, le proprietà dell'approccio di accesso alla rete potrebbero impedire ai clienti di prendere in considerazione il servizio per motivi normativi.

Domande di autovalutazione

- Quali sono le implicazioni in termini di costi dell'approccio di accesso alla rete selezionato per l'implementazione iniziale e la distribuzione continua? Quali sono i costi fissi e variabili dell'approccio?
- L'approccio all'accesso alla rete può scalare in modo efficace ed efficiente per soddisfare le esigenze di crescita del modello di business? Considerate le dimensioni dei singoli inquilini e il numero di inquilini inseriti.
- L'approccio all'accesso alla rete impone limitazioni tecniche o operative che potrebbero limitare la flessibilità o l'adattabilità del modello di business?
- Per quanto riguarda l'approccio all'accesso alla rete, in che modo il lead time di implementazione si allinea alla velocità di commercializzazione richiesta dal modello di business?

Mercato totale indirizzabile, tasso di acquisizione di nuovi clienti, crescita e scalabilità

È fondamentale valutare l'impatto delle decisioni di rete sulla capacità dell'organizzazione di espandersi in nuovi mercati, acquisire clienti in modo efficace e mantenere la scalabilità operativa. Questi fattori influiscono sui tassi di conversione. Influiscono anche sul fatto che l'approccio

all'accesso alla rete supporti l'espansione in segmenti di mercato significativi o limiti a servire solo tipi di clienti specifici.

Criteri di punteggio più alto

L'approccio all'accesso alla rete aiuta l'organizzazione a raggiungere una parte significativa del mercato di riferimento oppure può essere efficacemente combinato con altri approcci di rete per estendere la copertura del mercato. Questo approccio dovrebbe richiedere uno sforzo di integrazione aggiuntivo minimo. L'approccio supporta tempi di consegna brevi per l'implementazione, l'ingresso rapido sul mercato e l'espansione. Consente un numero elevato di implementazioni parallele. L'integrazione è semplice per i clienti, il che riduce le barriere all'adozione e migliora l'esperienza del cliente. L'approccio riduce al minimo il sovraccarico operativo, preserva la capacità operativa e supporta le proiezioni di crescita.

Indicatori a basso punteggio

L'approccio all'accesso alla rete supporta solo una piccola parte del mercato di riferimento o è adatto principalmente a segmenti di nicchia che non hanno priorità nella strategia aziendale. Non integra efficacemente altri approcci di accesso alla rete già supportati. I tempi di implementazione rallentano le richieste del mercato, il che limita l'espansione del mercato e l'acquisizione di nuovi clienti. Il modello di implementazione è sequenziale, il che aumenta i rischi di rallentamenti del servizio man mano che la domanda aumenta. I processi di integrazione complessi scoraggiano i potenziali clienti, il che influisce negativamente sul tasso di acquisizione e sui tassi di conversione. Un notevole sovraccarico operativo riduce la capacità operativa dell'organizzazione. Questo diventa un ostacolo alla crescita prevista.

Per questi indicatori, valuta se l'introduzione di un nuovo approccio di accesso alla rete può aiutare l'organizzazione a raggiungere i suoi obiettivi aziendali strategici. Valuta se il nuovo approccio all'accesso alla rete potrebbe creare nuove dipendenze tra i prodotti o consumare risorse operative senza fornire i risultati desiderati.

Domande di autovalutazione

- Esistono lacune nell'approccio attuale che vi impediscono di raggiungere segmenti più ampi del mercato di riferimento?
- Qual è l'elenco minimo di approcci di accesso alla rete non sovrapposti e standardizzati che dovreste supportare per coprire il 70-90% del mercato di riferimento?
- Quale portata consente ciascun approccio di accesso alla rete e quali sono i relativi aumenti di parametri importanti, come i costi dell'infrastruttura, i cicli operativi e la dipendenza dagli esperti?

- In che modo le capacità di implementazione e i limiti di servizio dell'infrastruttura di rete si allineano alle aspettative di crescita nei mercati di riferimento?
- L'integrazione di rete crea barriere all'ingresso di nuovi clienti? Come possono essere risolti per migliorare i tassi di conversione?
- In che modo il sovraccarico operativo di gestione della rete influisce sulla capacità di crescita e scalabilità?
- Quali strategie è possibile implementare per ridurre i tempi di consegna per l'implementazione della rete e migliorare l'espansione del mercato e l'acquisizione di clienti?
- Esistono dipendenze da risorse esperte che potrebbero ritardare l'implementazione o l'integrazione con gli ecosistemi dei clienti?

Esperienza e conservazione dei clienti

Le metriche di questa sezione ti aiutano a comprendere la capacità della tua organizzazione di acquisire e, soprattutto, fidelizzare i clienti. Comprendere la relazione tra gli approcci all'accesso alla rete e la soddisfazione del cliente può aiutare i team di prodotto e di progettazione a prendere decisioni basate sui dati.

Criteri con punteggi elevati

L'approccio all'accesso alla rete è affidabile e facile da gestire. Contribuisce ad aumentare la soddisfazione dei clienti (CSAT) e i risultati del net promoter score (NPS). Questi punteggi sono indicativi di una solida reputazione del marchio e della fidelizzazione dei clienti. Grazie alla perfetta integrazione con gli ecosistemi esistenti dei clienti, l'attrito nell'adozione è basso e la dipendenza dagli esperti è ridotta. L'organizzazione rispetta costantemente gli accordi sui livelli di servizio (SLAs), il che rafforza la fiducia dei clienti e gli obblighi contrattuali. Poiché i clienti usufruiscono di servizi stabili e affidabili, la fidelizzazione dei clienti è elevata.

Indicatori a basso punteggio

La difficile integrazione e l'accesso incoerente ai servizi generano spesso frustrazione dei clienti e feedback negativi. Ciò danneggia la reputazione del marchio. I nuovi clienti non riescono a passare da piani gratuiti o di prova a servizi a pagamento a causa della dipendenza da esperti o a causa dei tempi prolungati di onboarding e integrazione. I frequenti inadempimenti comportano SLAs sanzioni pecuniarie e una perdita di credibilità, con una potenziale riduzione dei tassi di fidelizzazione dei clienti.

Domande di autovalutazione

- In che modo le prestazioni della rete (come velocità, uptime e latenza) influiscono direttamente sui risultati CSAT e NPS? Quali miglioramenti specifici della rete potrebbero aumentare questi punteggi?
- In che modo le attuali metriche relative alla latenza di rete e all'uptime influiscono sull'esperienza utente iniziale e sui tassi di adozione? Quali miglioramenti specifici delle prestazioni di rete sono necessari per ottimizzare queste metriche?
- Esistono problemi ricorrenti nelle configurazioni di rete o nelle impostazioni di sicurezza che complicano l'integrazione per i nuovi clienti? Come puoi semplificare questi processi?
- In che modo la facilità di configurazione dell'accesso alla rete influisce sull'esperienza di onboarding dei nuovi utenti? Esistono punti di accesso alla rete o tempi di consegna specifici che possono essere ottimizzati per migliorare le impressioni iniziali degli utenti?
- Quali sono le sfide legate all'automazione della fornitura di servizi di rete per i nuovi clienti. Come è possibile modificare questo processo per migliorare la scalabilità e l'affidabilità?
- Analizza le cause principali delle recenti violazioni degli SLA. Erano legati alla configurazione della rete, alla pianificazione della capacità o a problemi dei fornitori esterni?
- Con che frequenza i problemi di rete vi fanno mancare gli impegni SLA? Quali sono i guasti più frequenti relativi alla rete?
- Quali miglioramenti delle prestazioni di rete hanno mostrato l'impatto positivo più significativo sulla soddisfazione dei clienti in passato?

Efficienza e performance finanziaria

Questa categoria valuta gli aspetti finanziari e di redditività dell'azienda, come l'efficienza dei costi, la redditività a lungo termine, la redditività, il ritorno sull'investimento (ROI) e il costo totale di proprietà (TCO). Semplificando le operazioni di rete attraverso la standardizzazione, è possibile ridurre le spese generali operative e i costi di manutenzione. Ciò supporta gli obiettivi di crescita dell'organizzazione.

Criteri con punteggi elevati

La struttura dei costi dell'approccio di accesso alla rete è ben allineata con il modello di business. Supporta una crescita sostenibile e i significativi risparmi sui costi che si ottengono aumentano la redditività. L'accesso efficiente alla rete consente una rapida onboarding dei clienti, che riduce i

tempi di fornitura di valore e accelera la penetrazione nel mercato. Ciò riduce direttamente il ciclo di riconoscimento dei ricavi.

Indicatori a basso punteggio

I clienti si rivolgono alla concorrenza per accelerare la fornitura delle loro applicazioni e servizi. La vostra organizzazione ha registrato un aumento dei costi operativi associati a configurazioni di rete complesse e varie e a tempi di consegna prolungati. La struttura dei costi e il modello di business non sono allineati, il che potrebbe causare costi iniziali elevati per i servizi in abbonamento. I complicati processi di onboarding riducono la penetrazione del mercato e posticipano il riconoscimento dei ricavi.

Domande di autovalutazione

- Quali sono gli attuali tempi di consegna per l'implementazione di nuovi servizi e in che modo influiscono sui tempi di riconoscimento del mercato e dei ricavi?
- In che modo le operazioni di rete standardizzate riducono efficacemente le spese generali e i costi di manutenzione?
- Sono necessarie risorse esperte per completare con successo l'integrazione iniziale, operare quotidianamente, risolvere i problemi o implementare le modifiche?
- Quanto sono sostenibili gli attuali investimenti di rete in termini di progressi tecnologici? State investendo in tecnologie a prova di futuro in linea con gli sviluppi di mercato previsti?
- Con quanta efficacia allocate e tenete traccia dei costi relativi al traffico di rete e all'utilizzo da parte dei singoli tenant?

Conformità normativa e gestione del rischio

È di fondamentale importanza convalidare la conformità alle normative relative alla rete. Ciò conferma che state operando legalmente e che potete mantenere la fiducia dei clienti. La standardizzazione delle operazioni di rete semplifica il processo di conformità e promuove la coerenza tra diverse giurisdizioni e aree geografiche. Queste misure ti aiutano a espandere i tuoi servizi.

Criteri con punteggi elevati

Le operazioni di rete rispettano costantemente gli standard legali senza complicazioni, il che contribuisce all'espansione del mercato, riduce gli attriti nell'adozione e migliora la fiducia dei clienti.

La comprovata conformità ai quadri normativi fondamentali, come il Digital Operational Resilience Act (DORA) e il National Institute of Standards and Technology (NIST), consente di conquistare clienti sensibili alla conformità normativa. La visibilità continua sullo stato di conformità riduce il tempo necessario per completare un audit.

Indicatori a basso punteggio

Le lacune nella conformità della rete causano forti difficoltà nell'adozione, ritardi nel lancio del servizio, problemi legali e potenziali multe. Queste sfide portano a ritardi o annullamenti dei piani di espansione in nuovi mercati. È difficile mantenere pratiche di conformità standard in diverse giurisdizioni e ciò influisce sull'efficienza operativa e sulla reputazione sul mercato.

Domande di autovalutazione

- In che misura le operazioni di rete sono allineate alle linee guida normative o di settore applicabili? Cosa ha rivelato il vostro ultimo audit di conformità?
- Come state mantenendo la conformità alle normative emergenti nel campo della sicurezza digitale e di rete?
- Quanto è efficace il vostro processo di documentazione e rendicontazione nel soddisfare i requisiti dei diversi organismi di regolamentazione?
- Quali strategie di gestione del rischio avete messo in atto per identificare e affrontare i potenziali rischi di conformità prima che portino a problemi legali?
- Di che livello di formazione e consapevolezza sulla conformità richiedono i team di gestione della rete per supportare i vostri approcci di accesso alla rete?

Strategia dei partner

Valuta in che misura l'approccio di accesso alla rete si allinea a un ecosistema di partner, piattaforme e marketplace riconosciuti. Questo è essenziale, soprattutto se la tua strategia di crescita dipende dalla scalabilità attraverso i partner.

Criteri di punteggio elevato

L'approccio all'accesso alla rete è integrato nell'ecosistema di partner. La sua struttura dei costi si allinea bene ai modelli di business dei vostri partner principali. I partner possiedono le competenze di rete necessarie per una perfetta integrazione delle tue offerte SaaS e possono fornire accesso e funzionalità duraturi.

Indicatori a basso punteggio

L'approccio di accesso alla rete selezionato richiede competenze, risorse o apparecchiature specializzate che sono scarse o difficili da reperire. Si differenzia dai protocolli di accesso alla rete standard comunemente utilizzati da piattaforme e mercati. Ciò si traduce in una struttura dei costi imprevedibile e difficile da conciliare. L'approccio all'accesso alla rete non è in linea con i modelli di business dei partner chiave.

Domande di autovalutazione

- Quali sono le implicazioni in termini di costi dell'approccio di accesso alla rete per i partner. In che modo questi costi si allineano ai loro modelli di business? Quale aspetto dell'integrazione sostiene la maggior parte dei costi e quanti cicli operativi devono essere investiti?
- Per quanto riguarda l'approccio all'accesso alla rete, esistono barriere all'integrazione o alla manutenzione che potrebbero influire sulle relazioni con i partner o sulla scalabilità dell'ecosistema?
- Come si può ottimizzare l'approccio all'accesso alla rete per migliorare la compatibilità e la facilità di integrazione in tutto l'ecosistema?

Metriche ingegneristiche che influenzano le decisioni di rete

Come i team di prodotto e commerciale, anche i team di progettazione utilizzano criteri di successo per valutare se stanno raggiungendo gli obiettivi aziendali. Tuttavia, queste metriche sono diverse e si concentrano sulla capacità del team di sviluppare, gestire e soddisfare i requisiti di sicurezza e conformità. Questa sezione descrive le metriche ingegneristiche che possono essere influenzate positivamente o negativamente dalle decisioni di accesso alla rete prese dall'organizzazione.

Utilizzate queste metriche e domande di autovalutazione per valutare il vostro attuale approccio all'accesso alla rete rispetto ai requisiti aziendali e alle capacità tecniche. Questa valutazione vi aiuta a identificare le lacune nella vostra architettura e a dare priorità ai miglioramenti in linea con i vostri obiettivi strategici. Esaminando regolarmente questi criteri, potete assicurarvi che la vostra strategia di accesso alla rete continui a supportare sia le esigenze dei clienti che i piani di crescita dell'organizzazione.

Questa sezione contiene metriche e domande di autovalutazione per le seguenti categorie e argomenti:

- Metriche di sviluppo

Metriche di sviluppo relative all'accesso alla rete per le offerte SaaS

Questa sezione contiene le seguenti metriche:

Frequenza di implementazione, tempo di implementazione e velocità di sprint

Per ottimizzare l'efficienza del ciclo di sviluppo, è essenziale comprendere l'influenza del provisioning dello stack di rete sulla velocità di sprint.

Criteri con punteggi elevati

Il provisioning dello stack di rete è semplificato e automatizzato e richiede un intervento manuale minimo. Non influisce in modo significativo sulla velocità di sprint. Il provisioning e la ridistribuzione dello stack di rete possono essere eseguiti da qualsiasi membro del team. Ciò riduce i colli di bottiglia e la dipendenza da risorse specializzate.

Indicatori a basso punteggio

È necessario un numero elevato di story point per il provisioning dello stack di rete. Ciò suggerisce un processo complesso e dispendioso in termini di tempo che sminuisce lo sviluppo di nuove funzionalità. La frequente ridistribuzione dello stack di rete comporta notevoli spese generali in termini di tempo e costi. Le attività di fornitura della rete richiedono competenze ingegneristiche specializzate, il che crea colli di bottiglia e rallenta il ciclo di sviluppo.

Domande di autovalutazione

- Quali passaggi manuali, se del caso, sono coinvolti nel processo di implementazione. In che modo influiscono sulla frequenza e sui tempi di implementazione?
- Come vengono gestiti i rollback in caso di errori di implementazione. Qual è il loro impatto sulla frequenza di implementazione e sui tempi di ripristino?
- Quanti punti di riferimento sono necessari per il provisioning dello stack di rete quando si configurano nuovi ambienti?
- Quanto tempo e costi aggiuntivi sono associati alla frequente ridistribuzione dello stack di rete durante il processo di sviluppo?
- Il provisioning dello stack di rete dipende da competenze ingegneristiche specializzate o è un'attività che può essere gestita da qualsiasi membro del team?

Flessibilità e fornitura di funzionalità

L'approccio all'accesso alla rete può influenzare la capacità del team di ingegneri di innovare e implementare nuove funzionalità in modo efficiente.

Criteri con punteggio elevato

L'approccio all'accesso alla rete offre la flessibilità necessaria per un'implementazione rapida e senza interruzioni delle funzionalità. Supporta un'ampia gamma di protocolli di comunicazione, comunicazioni unidirezionali e bidirezionali e dimensioni dei messaggi. Non impone vincoli significativi ai processi di sviluppo o all'innovazione.

Indicatori a basso punteggio

L'approccio all'accesso alla rete limita la capacità del team di implementare nuove funzionalità a causa della mancanza di protocolli di comunicazione supportati, della mancanza di flessibilità nelle dimensioni dei messaggi o della dipendenza da tecnologie specifiche e dalle relative risorse di esperti. Ciò può portare a cicli di sviluppo più lenti e ostacolare l'evoluzione del servizio.

Domande di autovalutazione

- In che modo l'approccio all'accesso alla rete influisce sull'agilità del team nello sviluppo e nell'implementazione di nuove funzionalità?
- Esistono limitazioni nell'approccio all'accesso alla rete che limitano il supporto di determinati protocolli o tecnologie di comunicazione?
- In che modo l'approccio facilita o limita l'integrazione di nuove tecnologie e innovazioni nel servizio?
- In che modo l'approccio all'accesso alla rete influisce sulle tempistiche di sviluppo e sulla roadmap del prodotto?

Modifica il tasso di fallimento

L'approccio di accesso alla rete scelto può influire sulla variazione del tasso di errore durante l'implementazione di nuovi servizi o funzionalità. Un maggiore controllo spesso significa maggiore flessibilità, ma aumenta anche il rischio di configurazioni errate, ad esempio quando si gestisce una configurazione di routing complessa.

Criteri di punteggio elevato

È possibile implementare modifiche allo stack di rete con un rischio minimo di guasto. Sono presenti meccanismi di test sufficienti, esistono meccanismi di rollback efficienti e un monitoraggio efficace consente di identificare e risolvere rapidamente i problemi.

Indicatori a basso punteggio

L'approccio all'accesso alla rete è soggetto a guasti durante le modifiche. Le opzioni di test sono limitate, le strategie di implementazione sono complicate o le funzionalità di monitoraggio e risoluzione dei problemi sono insufficienti. Per partecipare alle sessioni di risoluzione dei problemi sono necessarie più parti. Ciò può comportare un aumento dei tempi di inattività e ridurre la disponibilità dell'offerta SaaS.

Domande di autovalutazione

- Quali misure sono in atto per mitigare il rischio di errori di modifica durante l'aggiornamento dello stack di rete?
- Esistono processi di test e convalida completi?

- Quanto velocemente può essere ripristinato il sistema dopo una modifica non riuscita? È in atto un processo di rollback efficiente?
- Esistono sistemi di monitoraggio e avviso proattivi per rilevare e risolvere rapidamente i problemi durante e dopo le modifiche allo stack di rete?
- Qual è il tasso storico di errore delle modifiche per le implementazioni degli stack di rete. Quali lezioni sono state tratte dagli incidenti passati?
- In che modo l'approccio all'accesso alla rete facilita o limita l'implementazione delle modifiche. L'approccio riduce al minimo le interruzioni del servizio?
- Qual è il rischio di influire sulla disponibilità dell'offerta SaaS nell'ambiente di produzione quando si implementano modifiche che coinvolgono l'approccio all'accesso alla rete?

Qualità del codice e prestazioni del team di progettazione

Gli approcci di accesso alla rete possono influire indirettamente sulla qualità del codice per le offerte SaaS. La mancanza di standardizzazione nell'accesso alla rete può costringere il team di progettazione a supportare diversi approcci di integrazione, il che può portare a una base di codice gonfia. Ciò, a sua volta, può ostacolare la capacità del team di sviluppare la profondità e il controllo sulla qualità del codice necessari per mantenere team di progettazione ad alte prestazioni.

Criteri con punteggi elevati

Il team di ingegneri rimane concentrato grazie alla modularità e alla riusabilità del codice attraverso gli approcci di accesso alla rete supportati. Gli approcci di accesso alla rete sono compatibili con le pipeline di implementazione esistenti e le strategie di test automatizzate.

Indicatori a basso punteggio

Le prestazioni del team di progettazione sono ridotte a causa del sovraccarico associato all'integrazione e alla manutenzione di troppi approcci di accesso alla rete. Alcuni approcci aumentano in modo significativo la complessità, generano debito tecnologico o richiedono lo sviluppo di soluzioni alternative per risolvere il problema delle funzionalità mancanti o insufficienti.

Domande di autovalutazione

- In che modo l'approccio all'accesso alla rete gestisce la variabilità della rete?
- È necessario sviluppare codice aggiuntivo per gestire le interruzioni della connettività?
- Un nuovo approccio di accesso alla rete si integra perfettamente con gli approcci esistenti o richiede uno sviluppo personalizzato significativo?

- Qual è la portata del cambiamento necessario per adottare un nuovo approccio di accesso alla rete? La base di codice esistente e i test automatici possono essere utilizzati in modo efficace?
- Quanto è facile o difficile implementare o ridistribuire il servizio con l'approccio di accesso alla rete selezionato? È possibile eseguire questa operazione frequentemente? Esistono delle dipendenze dalle risorse degli esperti?
- L'approccio all'accesso alla rete facilita o complica l'adesione agli standard di codifica e alle migliori pratiche?
- In che modo l'approccio influisce sulle time-to-market nuove funzionalità o correzioni?

Riduzione tecnica del debito

Una valutazione dell'impatto di un approccio di accesso alla rete sul debito tecnico dovrebbe prendere in considerazione la sua scalabilità, osservabilità e capacità di sicurezza.

Criteri con punteggi elevati

L'approccio semplifica efficacemente la gestione dell'infrastruttura man mano che la base clienti si espande. Offre solide capacità di osservabilità. out-of-the-box Ciò favorisce un monitoraggio e una manutenzione efficienti.

Indicatori a basso punteggio

L'approccio all'accesso alla rete protegge in modo inadeguato i canali di comunicazione e non dispone di strumenti sufficienti per l'osservazione metrica qualitativa. Potrebbe inoltre richiedere uno sviluppo aggiuntivo per la gestione dell'infrastruttura man mano che la base clienti aumenta, oppure potrebbe richiedere soluzioni alternative per problemi di affidabilità.

Domande di autovalutazione

- In che modo l'approccio all'accesso alla rete influenza la scalabilità a lungo termine dell'infrastruttura? Facilita una crescita senza interruzioni con un investimento aggiuntivo minimo?
- Quanto sono completi gli strumenti di osservabilità inclusi? Consentono il monitoraggio proattivo e la risoluzione dei problemi?
- Qual è l'impatto previsto dell'approccio all'accesso alla rete sulla manutenzione e l'evoluzione della codebase nel tempo?
- L'approccio si integra bene con l'infrastruttura esistente e pianificata? Richiede modifiche o aggiunte significative?

Scalabilità, capacità e prestazioni

Per determinare l'idoneità di un approccio di accesso alla rete per un'offerta SaaS, è essenziale analizzare in che modo mantiene prestazioni ottimali all'aumentare della domanda.

Criteri di punteggio elevato

L'approccio all'accesso alla rete facilita senza problemi l'espansione. Mantiene una bassa latenza durante l'elaborazione delle richieste e gestisce in modo efficiente i picchi di traffico. Fornisce prestazioni costanti indipendentemente dall'aumento dei livelli di traffico e non impone limiti operativi alla crescita.

Indicatori a basso punteggio

L'approccio all'accesso alla rete non è scalabile in modo efficace, probabilmente a causa delle limitazioni intrinseche della larghezza di banda o dell'insufficiente capacità dell'infrastruttura.
Il provisioning e la gestione delle risorse aumentano la complessità o creano dipendenze. Le prestazioni del servizio sono ridotte a causa dell'aumento della latenza, del jitter e della variabilità del throughput, in particolare in condizioni di rete congestionata.

Domande di autovalutazione

- In che modo l'approccio all'accesso alla rete soddisfa un numero crescente di inquilini e i relativi volumi di dati?
- È intrinsecamente scalabile per soddisfare le esigenze future?
- Quali misure sono in atto per garantire che le prestazioni siano costanti, anche durante i periodi di picco di traffico o gli eventi di rapida scalabilità?
- In che modo l'approccio gestisce la latenza e il jitter della rete? Esistono meccanismi per ottimizzare la velocità di trasmissione dei dati e ridurre al minimo i ritardi?
- L'approccio di accesso alla rete può adattarsi alle diverse condizioni di rete? Può fornire un'esperienza single-tenant per ogni cliente?
- Qual è l'impatto dell'approccio di accesso alla rete sull'infrastruttura sottostante? Richiede aggiornamenti o modifiche significative ai sistemi esistenti?

Metriche di eccellenza operativa relative all'accesso alla rete per le offerte SaaS

Questa sezione contiene le seguenti metriche:

Resilienza operativa e disaster recovery

L'approccio all'accesso alla rete dovrebbe aiutare l'offerta SaaS a resistere a vari tipi di interruzioni e a riprendersi rapidamente da eventuali disastri.

Criteri di punteggio elevato

I piani di disaster recovery consolidati e testati dimostrano costantemente che l'approccio di accesso alla rete soddisfa i requisiti di disaster recovery. L'approccio di accesso alla rete supporta configurazioni ad alta disponibilità e supporta meccanismi di failover automatici, rapidi e affidabili.

Indicatori a basso punteggio

L'approccio di accesso alla rete rende difficile la creazione di una strategia coerente di disaster recovery. Si osservano tempi di ripristino prolungati dopo le interruzioni. I frequenti guasti operativi dell'infrastruttura di rete influiscono sull'erogazione dei servizi.

Domande di autovalutazione

- Quando è stata l'ultima esercitazione di disaster recovery e quali sono stati i risultati?
- Quanto tempo occorre per ripristinare i servizi critici dopo un'interruzione? Quale parte dell'infrastruttura di rete deve essere ridistribuita?
- Quali miglioramenti è possibile apportare all'infrastruttura di rete per semplificare i piani di disaster recovery?
- Esistono ridondanze per i componenti di rete più critici?
- Hai automatizzato la potenziale ridistribuzione dell'infrastruttura di rete dopo un'interruzione critica?
- In che modo l'approccio all'accesso alla rete supporta la tolleranza agli errori e l'affidabilità? Esistono meccanismi integrati per gestire le interruzioni di rete e mantenere l'integrità dei dati?

Monitoraggio delle prestazioni dei servizi e delle applicazioni

L'approccio all'accesso alla rete può influire sugli strumenti di monitoraggio delle prestazioni utilizzati per convalidare il funzionamento ottimale e l'operatività del servizio. A seconda del servizio, potresti avere accesso a metriche di basso livello (come le percentuali di caduta dei pacchetti) o a metriche di livello superiore (come la durata della sessione). Le metriche di basso livello forniscono informazioni

tecniche dettagliate sul comportamento della rete, ma possono essere complesse da interpretare. Al contrario, le metriche di livello superiore offrono spesso un modo più diretto e semplice per valutare l'esperienza complessiva dell'utente. Questo perché aggregano l'impatto delle condizioni di rete sottostanti in chiari indicatori della qualità del servizio.

Criteri di punteggio elevato

Sono immediatamente disponibili strumenti di monitoraggio completi che forniscono informazioni quasi in tempo reale. Disponi di avvisi e sistemi di risposta automatici che risolvono i problemi di prestazioni. È possibile prevedere potenziali rallentamenti o guasti del servizio prima che si ripercuotano sugli utenti.

Indicatori a basso punteggio

Le interruzioni frequenti del servizio o i problemi di prestazioni si verificano senza essere osservati o risolti. La mancanza di visibilità sulle prestazioni del servizio comporta una risposta lenta ai rallentamenti delle prestazioni. Sono necessari team composti da più parti per risolvere i problemi dell'infrastruttura di rete.

Domande di autovalutazione

- Quali strumenti di monitoraggio e metriche dell'infrastruttura di rete sono attualmente disponibili? Quanto sono efficaci nel rilevare le anomalie del servizio?
- Con quale rapidità è possibile identificare e risolvere i problemi di prestazioni?
- Disponete di meccanismi che prevedono potenziali problemi di prestazioni?
- Quali miglioramenti potete apportare per potenziare le capacità di osservabilità?

Metriche di sicurezza e governance relative all'accesso alla rete per le offerte SaaS

Questa sezione contiene le seguenti metriche:

- Gestione della sicurezza, della conformità e delle vulnerabilità

Gestione della sicurezza, della conformità e delle vulnerabilità

È fondamentale valutare gli aspetti di sicurezza dell'approccio all'accesso alla rete, inclusa la conformità agli standard di sicurezza e la gestione delle vulnerabilità.

Criteri di punteggio elevato

L'approccio all'accesso alla rete aiuta il team a rispettare i framework di sicurezza, come l'International Organization for Standardization (ISO) 27001, System and Organization Controls 2 (SOC 2) o NIST. Semplifica l'esecuzione di controlli di sicurezza regolari. Sono in atto solidi meccanismi di crittografia e autenticazione. Le reti sono isolate e solo le risorse necessarie sono esposte all'infrastruttura del cliente. È possibile individuare le anomalie di rete quasi in tempo reale, senza sovraccarichi eccessivi.

Indicatori a basso punteggio

L'approccio all'accesso alla rete è soggetto a violazioni o vulnerabilità ricorrenti della sicurezza e non è conforme ai principali standard di sicurezza. Si riscontrano spesso ritardi nel rilevamento e nelle risposte agli incidenti di sicurezza.

Domande di autovalutazione

- Vi sono state recenti violazioni della sicurezza legate all'approccio selezionato per l'accesso alla rete e cosa abbiamo imparato da esse?
- In che modo il vostro approccio all'accesso alla rete è conforme agli standard di sicurezza globali?
- Quanto tempo occorre per rilevare e rispondere alle minacce alla sicurezza? In che modo l'accesso alla rete aiuta o limita questa capacità?
- Con quale frequenza vengono condotte valutazioni di sicurezza sugli approcci di accesso alla rete? È possibile utilizzare strumenti comuni per valutare la sicurezza dell'approccio di accesso alla rete o è necessario un software specializzato?
- Quale livello di sicurezza è intrinseco all'approccio all'accesso alla rete e in che modo si allinea alle migliori pratiche del settore e ai requisiti normativi?

Panoramica dei servizi di AWS rete per le offerte SaaS

Questa sezione descrive i servizi AWS di rete a cui si fa riferimento in questa guida. Inoltre, confronta le loro capacità e descrive le considerazioni sulla sicurezza per ogni servizio.

Questa sezione contiene i seguenti argomenti:

- AWS servizi di rete
- Confronto delle funzionalità del servizio
- Caratteristiche e considerazioni sulla sicurezza

AWS servizi di rete

Di seguito sono riportati Servizi AWS gli argomenti trattati in modo coerente in questa guida.

AWS PrivateLink

AWS PrivateLinkè un servizio nativo del cloud che può fornire l'accesso alla tua offerta SaaS se i tuoi clienti operano già in. Cloud AWS Il tuo cliente si connette all'offerta SaaS tramite un endpoint VPC di interfaccia. Si tratta di un'interfaccia di rete endpoint fornita in una o più sottoreti del cliente. Account AWS Negli scenari di questa guida, il traffico viaggia attraverso l'interfaccia VPC endpoint e arriva a un Network Load Balancer del tuo account. Il Network Load Balancer inoltra il traffico all'applicazione SaaS, che hai registrato come servizio endpoint. Tramite gli endpoint VPC di risorse, AWS PrivateLink può anche aiutarti ad accedere ad altre risorse, come i database.

Amazon VPC Lattice

Amazon VPC Lattice è un servizio di rete di applicazioni che aiuta i provider SaaS a offrire i propri servizi in modo sicuro ed efficiente ai clienti che operano su più piattaforme. VPCs Account AWS I clienti accedono alla tua offerta SaaS tramite VPC Lattice, che offre connettività di rete coerente, solidi controlli degli accessi e gestione avanzata del traffico. In questi scenari, il traffico fluisce attraverso VPC Lattice verso i servizi applicativi registrati. Fornisce comunicazioni scalabili e sicure, indipendentemente dal servizio di elaborazione utilizzato.

Peering VPC

Il peering VPC è una connessione di rete tra due cloud privati virtuali (VPCs) che indirizza il traffico tra di essi utilizzando indirizzi o IPv4 indirizzi privati. IPv6 Il peering VPC viene in genere utilizzato

tra entità attendibili, come quelle all'interno della stessa organizzazione. Il cliente crea una richiesta di peering a uno dei tuoi. VPCs Una volta accettata, il traffico può fluire tra i due VPCs in entrambe le direzioni. Questo approccio di connessione si distingue per la sua unicità perché implica la comunicazione diretta tra due persone VPCs senza alcun servizio o infrastruttura di intermediazione da gestire.

AWS Transit Gateway

AWS Transit Gatewayè un hub di transito di rete centralizzato in grado di connettere VPCs connessioni di rete privata virtuale (VPN), AWS Direct Connect gateway, dispositivi virtuali di terze parti in un VPC e altri gateway di transito. Un gateway di transito può avere una tabella di routing diversa per ogni allegato. Ciò offre la massima flessibilità per il routing e aiuta a isolare le reti. Viene spesso utilizzato per collegarne molte tra VPCs loro o per l'ispezione centralizzata.

AWS Site-to-Site VPN

AWS Site-to-Site VPNpuò utilizzare la tecnologia Internet Protocol Security (IPsec) per stabilire connessioni tra reti locali, uffici remoti, fabbriche, altri provider di servizi cloud e la AWS rete globale. La connessione viene stabilita da un gateway privato virtuale o da un gateway di transito in un VPC Cloud AWS a un gateway cliente fisico o basato su software, che può essere in locale o nel Cloud AWS cloud di un altro CSP. La connessione può avvenire tramite Internet o tramite una connessione fisica. AWS Direct Connect È anche possibile avere una connessione Site-to-Site VPN accelerata utilizzando AWS Global Accelerator. Una connessione accelerata indirizza il traffico verso una posizione AWS periferica e offre una latenza ridotta e prestazioni migliorate.

AWS Direct Connect

AWS Direct Connectstabilisce una connessione privata ad alta velocità tra un data center locale e. Cloud AWS Bypassando la rete Internet pubblica, Direct Connect fornisce una connessione a bassa latenza più affidabile, sicura e coerente a. Cloud AWS I clienti si connettono a una delle Direct Connect sedi e quindi scelgono una connessione ospitata o dedicata a. AWS Sebbene si tratti di una scelta di architettura non comune per le offerte SaaS, può essere adatta ai provider SaaS che hanno pochi ma grandi clienti aziendali.

Confronto delle funzionalità del servizio

La tabella seguente illustra le funzionalità supportate di Servizi AWS quelle illustrate in questa guida. Di seguito sono riportate le descrizioni delle funzionalità incluse in questa tabella:

- Intervalli CIDR sovrapposti: può connettere due o più reti con intervalli CIDR uguali o sovrapposti
- Comunicazione bidirezionale: può supportare un canale di comunicazione bidirezionale in modo che il consumatore SaaS possa esporre risorse interne, come un database, al provider SaaS
- IPv6— Può supportare uno o due stack IPv6
- Jumbo frame: può supportare frame jumbo con una dimensione del frame fino a 8.500 byte
- Cloud ibrido: può supportare una connessione con una rete locale
- Multi-cloud: può supportare una connessione tra reti su diversi provider di servizi cloud

Servizio o approccio	Intervalli CIDR sovrapposti	Comunicazione bidirezionale	IPv6	Cornice Jumbo	Cloud ibrido	Multicloud
Peering VPC	No	Sì	Sì	Sì [5]	No	No
AWS PrivateLink	Sì	Sì [1]	Sì	Sì	Sì [6]	No [6]
Amazon VPC Lattice	Sì	Sì [1]	Sì	Sì	Sì [6]	No [6]
AWS Transit Gateway	No	Sì	Sì	Sì	Sì [3]	Sì [3]
AWS Site-to-Site VPN	No	Sì	Sì	No	Sì	Sì
AWS Direct Connect	No	Sì	Sì	Sì [2]	Sì	Sì

Accesso pubblico a Internet [4]	Non applicabile	No	Sì	Sì	Sì	Sì

1. Con risorse VPC in Amazon VPC Lattice
2. Solo per interfacce virtuali private e di transito
3. Con Site-to-Site VPN o allegati AWS Direct Connect
4. Come termine generale per le AWS risorse che rendono un'applicazione accessibile al pubblico, come un Application Load Balancer
5. Solo per connessioni peering all'interno di una Regione AWS
6. Possibile tramite una connessione Layer 3 preesistente tra gli ambienti

Caratteristiche e considerazioni sulla sicurezza

La tabella seguente descrive le funzionalità di sicurezza Servizi AWS illustrate in questa guida.

- Mezzi di autenticazione: come puoi assicurarti che solo i tuoi clienti possano connettersi al tuo servizio. Di solito è ancora necessario un altro livello di autenticazione per le richieste in entrata, specialmente negli ambienti con tenant condivisi.
- Crittografia in transito: descrive se la crittografia in transito è fornita per impostazione predefinita. La crittografia nativa descrive la crittografia che AWS fornisce tutto il traffico all'interno VPCs VPCs, tra o tra i data center. La crittografia supplementare descrive la crittografia controllata dall'utente e che può essere interrotta dal rispettivo servizio.

Servizio o approccio	Mezzi di autenticazione	Crittografia in transito
Peering VPC	Avviate una richiesta di peering verso il Account AWS VPC del vostro cliente o accettate una richiesta da lui avviata. Vedi Accettare o rifiutare una connessione peering VPC.	Solo crittografia nativa

AWS PrivateLink	Sei tu a scegliere quali Account AWS sono autorizza ti a creare endpoint per il tuo servizio. Questi account sono noti come indirizzi principal i consentiti. Vedi Accettare o rifiutare le richieste di connessione.	Solo crittografia nativa
Amazon VPC Lattice	Condividete un servizio o una rete di assistenza VPC Lattice con i vostri clienti. Account AWS Vedi Condividi le tue entità VPC Lattice.	Crittografia nativa e crittografia TLS supplementare
AWS Transit Gateway	Il cliente crea una richiesta di peering allegato a partire da lui oppure sei tu Account AWS ad avviare la richiesta. Vedi gli allegati di peering del gateway Transit in Amazon VPC Transit Gateway.	Crittografia nativa e crittografia supplementare con IPsec un allegato VPN
AWS Site-to-Site VPN	Utilizzi chiavi IPsec precondiv ise o un certificato privato sul dispositivo del cliente. Vedi le opzioni di autenticazione AWS Site-to-Site VPN del tunnel.	Crittografia supplementare IPsec
AWS Direct Connect	Il cliente crea una richiesta di interfaccia virtuale dal proprio Account AWS. Visualizza le interfacce Direct Connect virtuali e le interfacce virtuali ospitate.	Crittografia supplementare di livello 2 possibile in siti selezionati. Vedi Direct Connect Sedi.

Accesso pubblico a Internet [1]	È richiesta l'autenticazione personalizzata.	È possibile una crittografia TLS supplementare

1. Come termine generale per le AWS risorse che rendono un'applicazione accessibile al pubblico, come un Application Load Balancer

Valutazione delle opzioni di accesso alla rete per le offerte SaaS

Le metriche importanti per la tua organizzazione dipenderanno da chi sono i tuoi clienti, dalla tua strategia aziendale e dai tuoi obiettivi organizzativi. Questa guida presenta le metriche che potete utilizzare per scegliere un approccio di accesso alla rete, ma dovreste dare la priorità a quelle che soddisfano i requisiti specifici del vostro caso d'uso.

Questa sezione contiene i seguenti argomenti:

Metriche di valutazione

Alcune metriche sono coerenti tra le organizzazioni e i casi d'uso e queste sono le metriche che possiamo aiutarti a valutare. Le seguenti sono queste metriche:

- Facilità di integrazione: con quale rapidità e facilità è possibile acquisire nuovi clienti?
- Costo totale di proprietà (TCO): qual è la struttura dei costi? Oltre ai costi fissi e variabili dell'infrastruttura, vi sono importanti considerazioni sui costi aggiuntivi associati al sovraccarico operativo, alla dipendenza dagli esperti, al costo di implementazione delle modifiche e alla conformità. Per ulteriori informazioni, consulta la sezione Costo totale di proprietà.
- Scalabilità: il vostro approccio all'accesso alla rete è scalabile per supportare la crescita della vostra azienda? La scalabilità della base clienti comporta importanti considerazioni architettoniche e organizzative. Considerate come potreste scalare per soddisfare un numero di clienti da 5 a 100 volte superiore a quello che supportate oggi.
- Adattabilità: è possibile implementare facilmente le modifiche? Le modifiche possono includere una nuova applicazione, una nuova funzionalità, una piattaforma diversa o una rete diversa.
- Isolamento della rete: quanta parte dell'infrastruttura di rete esponete ai vostri clienti? State fornendo il giusto grado di accesso o state esponendo intere reti? Se si isolano tempestivamente le risorse di rete, sarà più facile fornire garanzie di sicurezza, privacy e conformità in un secondo momento.

- Osservabilità: qual è la tua capacità di rilevare guasti o deterioramenti del servizio? Quanto è facile e veloce identificare il problema? In quanto tempo (e con quali costi generali) potete aiutare i vostri clienti a comprendere i loro punti di errore e aiutarli a risolverli?
- Tempo di riparazione: qual è il tempo che intercorre tra il rilevamento di un guasto o di un deterioramento del servizio e la ripresa delle operazioni? Quali sono i fattori che influiscono su questa capacità?

Altre metriche sono specifiche della vostra organizzazione o della vostra offerta perché si riferiscono alle operazioni, alla strategia o agli obiettivi aziendali. Solo tu puoi valutare queste metriche. Le seguenti sono queste metriche:

- Allineamento del modello di business: qual è il vostro modello di business e in che misura gli approcci di accesso individuali si allineano ad esso?
- Mercato indirizzabile totale (TAM): qual è il vostro mercato attuale e futuro e in che misura è coperto dall'approccio di accesso alla rete?
- Ritorno sull'investimento (ROI): quali miglioramenti vi aspettate in termini di redditività e margini? I vantaggi finanziari attesi sono sufficienti a soddisfare le vostre esigenze di accesso ai servizi adattabile e flessibile?
- Conformità normativa: che tipo di requisiti normativi si applicano e in quale mercato?
- Accordi sui livelli di servizio (SLAs): i clienti hanno bisogno che la tua offerta SaaS sia altamente disponibile? Che tipo di impegni siete contrattualmente obbligati a rispettare?

Costo totale di proprietà

Questa sezione esplora il costo totale di proprietà (TCO), che è una delle metriche di valutazione utilizzate per confrontare gli approcci di accesso alla rete. Il TCO è una metrica composita composta da costi di infrastruttura fissi e variabili, costi generali operativi, dipendenza specialistica, costo della modifica e costi di conformità.

La classificazione del TCO per ogni approccio di accesso alla rete può variare in base al caso d'uso. Ad esempio, il costo del cambiamento per un provider SaaS con un semplice servizio web e cinque tenant è diverso da quello di un provider SaaS con un portafoglio di prodotti complesso e interconnesso e centinaia o migliaia di tenant. Inoltre, non tutti i componenti hanno lo stesso peso. Ad esempio, assumere uno specialista di rete è spesso più costoso dei costi di infrastruttura necessari per supportare un'implementazione individuale del servizio. Utilizzate i valori riportati nella tabella seguente per orientarvi inizialmente e come punto di riferimento per ulteriori discussioni.

Approccio di accesso	Costi fissi dell'infrastruttura	Costi di infrastruttura variabili	Sovraccarico operativo	Dipendenza specialistica	Costo della modifica	Costi di conformità
Peering VPC	Nessuno	Nessuno	Elevata	Bassa	Elevata	Media
AWS PrivateLink	Bassa	Bassa	Bassa	Nessuno	Bassa	Bassa
Amazon VPC Lattice	Media	Media	Bassa	Bassa	Bassa	Bassa
AWS Transit Gateway	Media	Media	Bassa	Bassa	Bassa	Media
AWS Site-to-Site VPN	Media	Elevata	Elevata	Media	Media	Bassa
AWS Direct Connect	Elevata	Media	Media	Elevata	Elevata	Bassa
Accesso pubblico a Internet	Bassa	Elevata	Media	Bassa	Bassa	Elevata

Costi di peering VPC

Non vi sono costi di infrastruttura diretti associati a una connessione peering VPC. Quando il traffico rimane all'interno della stessa zona di disponibilità, non è previsto alcun costo per il trasferimento dei dati. Tuttavia, il sovraccarico operativo può essere significativo perché la gestione e la complessità aumentano esponenzialmente con ogni connessione peering aggiuntiva. Una conoscenza di base del networking è sufficiente per configurare una connessione peering, ma le modifiche alla rete sono

difficili da implementare con più di una manciata di connessioni peering. I costi di conformità sono leggermente più elevati perché entrambe le parti espongono l'un intero VPC l'una all'altra, anziché i singoli servizi.

AWS PrivateLink costi

AWS PrivateLink è spesso una soluzione economica con costi operativi ridotti. Questo perché il provider SaaS deve gestire solo un Network Load Balancer e il consumatore deve gestire solo gli endpoint VPC. È possibile apportare modifiche su entrambi i lati in modo trasparente, il che riduce la collaborazione interorganizzativa costosa e dispendiosa in termini di risorse. I costi di conformità tendono ad essere bassi perché il provider SaaS espone solo i servizi che desidera e non l'intera rete.

Costi di Amazon VPC Lattice

Amazon VPC Lattice offre una struttura di costi bilanciata con costi di infrastruttura fissi e variabili moderati. Essendo una rete di servizi completamente gestita, riduce in modo significativo il sovraccarico operativo automatizzando l'individuazione dei servizi, la gestione del traffico e il controllo degli accessi su più piattaforme. VPCs Ciò semplifica sia l'implementazione iniziale che la gestione continua rispetto alle configurazioni di rete manuali. È possibile implementare le modifiche tramite controlli basati su policy senza complessi aggiornamenti di routing, il che riduce la dipendenza dagli specialisti di rete. I costi di conformità tendono ad essere inferiori rispetto agli approcci di rete tradizionali perché VPC Lattice offre controlli di accesso granulari e visibilità completa attraverso funzionalità di monitoraggio e registrazione integrate. Ciò può semplificare la dimostrazione della conformità normativa.

AWS Transit Gateway costi

AWS Transit Gateway ha costi orari e per l'elaborazione dei dati superiori a quelli AWS PrivateLink, ma ha costi operativi simili. È necessario avere una conoscenza più approfondita del AWS Transit Gateway servizio e del routing per configurare correttamente tutte le tabelle di routing. AWS Le modifiche all'infrastruttura potrebbero richiedere aggiornamenti di routing o DNS. I costi di conformità sono simili a quelli del peering VPC perché entrambe le parti espongono potenzialmente sottoreti o intere reti l'una all'altra. VPCs AWS Transit Gateway Inoltre, le tabelle di routing devono essere gestite con attenzione perché sono condivise da più utenti e non è necessario consentire alcun traffico tra di esse.

AWS Site-to-Site VPN costi

Poiché la Site-to-Site VPN invia essenzialmente traffico a Internet, il costo variabile è più elevato rispetto ai costi di trasferimento dei dati. Sebbene si tratti di un servizio di rete privata virtuale (VPN) gestito, comporta un notevole sovraccarico operativo, in particolare sul gateway del cliente. Il provisioning e le operazioni richiedono una conoscenza avanzata del networking e le modifiche spesso richiedono l'intervento di entrambe le parti. I costi di conformità sono generalmente bassi perché i team di sicurezza spesso approvano preventivamente i IPsec tunnel senza ulteriori controlli.

AWS Direct Connect costi

AWS Direct Connect comporta il costo di infrastruttura fisso più elevato in quanto si tratta di una connessione fisica privata direttamente al Cloud AWS. Sono necessarie conoscenze specialistiche per configurare e gestire una sessione BGP (Border Gateway Protocol) (se necessario), per gestire una connessione VPN ed eseguire l'ingegneria del traffico. Questo servizio riduce l'impegno dei team di sicurezza perché combina la connettività privata con la possibilità di aggiungere Media Access Control Security (MACsec) e crittografia. IPsec

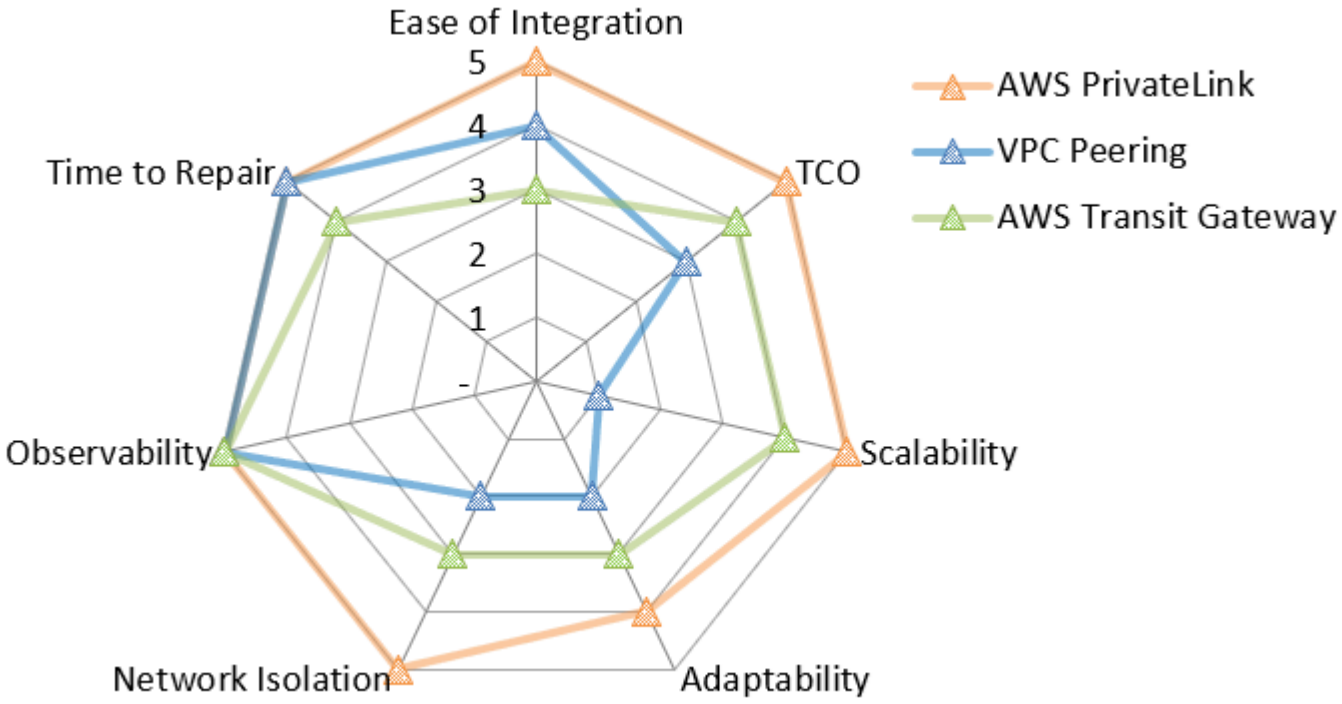
Costi di accesso pubblico a Internet

L'accesso pubblico a Internet si riferisce alle AWS risorse che è possibile utilizzare per rendere un'applicazione accessibile al pubblico, ad esempio un Application Load Balancer. Per questo approccio, vi sono costi variabili legati alla fornitura dell'accesso ai servizi, compresi i costi per il trasferimento dei dati su Internet. I costi operativi generali e di conformità possono essere significativi, in quanto si espone il servizio a Internet e sono necessari meccanismi di sicurezza e autenticazione aggiuntivi. Tuttavia, non è necessario un routing complesso e nessuna delle parti deve conoscere i dettagli dell'infrastruttura dell'altra.

Mappa dei valori della rete

Per aiutarti a vedere il quadro generale e prendere decisioni informate, questa guida include una mappa dei valori di rete per ogni scenario. Poiché le valutazioni differiscono da scenario a scenario, lo stesso servizio potrebbe ottenere punteggi diversi per due scenari. Le mappe dei valori sono grafici radar, in cui un ipotetico punteggio perfetto sarebbe un cinque in tutte le categorie.

Ad esempio, l'immagine seguente mostra un esempio di grafico radar. Include solo le metriche che possiamo aiutarti a valutare. Ti consigliamo di creare la tua mappa dei valori che includa le metriche aggiuntive che solo tu puoi valutare.


Ease of Integration
5
4
3
2
1
-
TCO
Scalability
Adaptability
Network Isolation
Observability
Time to Repair
AWS PrivateLink
VPC Peering
AWS Transit Gateway

Scenari di accesso alla rete per le offerte SaaS in Cloud AWS

Questa sezione descrive diverse opzioni di accesso alla rete per le tue offerte SaaS in. Cloud AWS Descrive gli approcci dal punto di vista del consumatore, che potrebbe avere esigenze di connettività all' Cloud AWS interno dei data center locali o di altri provider di servizi cloud (). CSPs Inoltre, potrebbe essere necessario supportare l'accesso da diversi tipi di ambienti consumer.

Comprendere i requisiti di connettività di rete in questi diversi ambienti è essenziale per creare una strategia di accesso completa. Le decisioni relative all'architettura devono tenere conto dei diversi modelli di sicurezza, delle aspettative prestazionali e dei vincoli tecnici, mantenendo al contempo l'efficienza operativa. L'approccio giusto offre una connettività sicura e affidabile che si adatta alla crescita aziendale e riduce al minimo sia la complessità dell'implementazione che il sovraccarico di gestione continuo.

Nel valutare le opzioni di accesso alla rete, considerate l'impatto di ogni approccio sul costo totale di proprietà, che include non solo i costi dell'infrastruttura ma anche i costi operativi e i requisiti di conformità. Alcuni approcci eccellono in termini di scalabilità ma possono introdurre complessità, mentre altri danno priorità alla facilità di integrazione a scapito dell'isolamento della rete. Anche le capacità e le risorse tecniche dei vostri consumatori svolgono un ruolo importante nella determinazione della soluzione più appropriata.

Per i consumatori in tutto il mondo Cloud AWS, servizi come questi AWS PrivateLink offrono vantaggi significativi in termini di sicurezza e scalabilità. Gli utenti locali potrebbero trarre vantaggio da AWS Direct Connect prestazioni costanti o trarre vantaggio dalla Site-to-Site VPN per una connettività conveniente. Gli scenari multi-cloud spesso richiedono un'attenta considerazione delle sfide di interoperabilità e potresti utilizzare architetture VPC di transito per standardizzare i modelli di accesso. In tutti i casi, il design deve anticipare le future crescite dei consumatori e del traffico in modo che l'architettura di rete rimanga resiliente e adattabile man mano che l'offerta SaaS si evolve.

Questa sezione contiene i seguenti scenari:

- Consumatori SaaS che operano su AWS
- Consumatori di servizi che operano in sede
- Consumatori SaaS che operano su altri provider di servizi cloud
- Supporto per ambienti ibridi

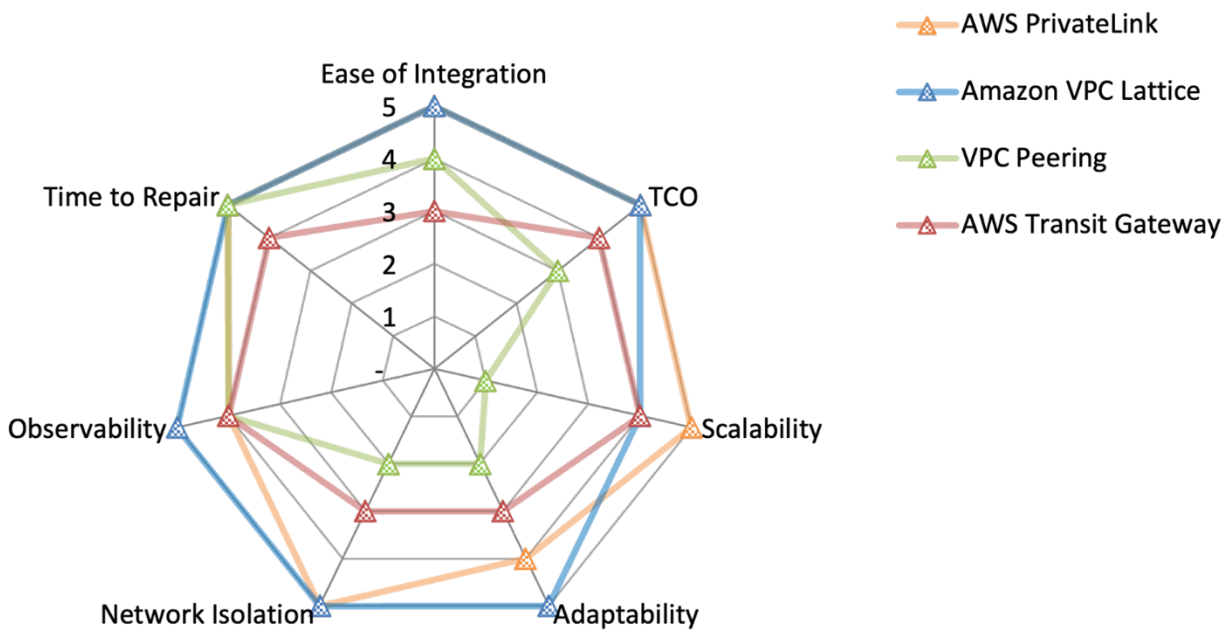
Consumatori SaaS che operano su AWS

Questa sezione illustra le opzioni di connettività se sia tu che i tuoi consumatori operate in. Cloud AWS Questo scenario offre la massima flessibilità perché molte si integrano Servizi AWS in modo nativo e perché entrambe le parti hanno accesso all'intero Servizio AWS portafoglio.

Questa sezione illustra i seguenti approcci di accesso alla rete:

- Integrazione con AWS PrivateLink
- Condivisione di un servizio Amazon VPC Lattice
- Creazione di connessioni peering VPC
- Connessione VPCs con AWS Transit Gateway

La seguente mappa dei valori di rete riassume il punteggio di ciascuna di queste opzioni per ogni metrica di valutazione. Per ulteriori informazioni sulle metriche di valutazione, consulta Metriche di valutazione in questa guida. Nella mappa, un cinque rappresenta il punteggio migliore, ad esempio il TCO più basso, il miglior isolamento di rete o il minor tempo di riparazione. Per ulteriori informazioni su come leggere questo grafico radar, Mappa dei valori della rete consulta questa guida.



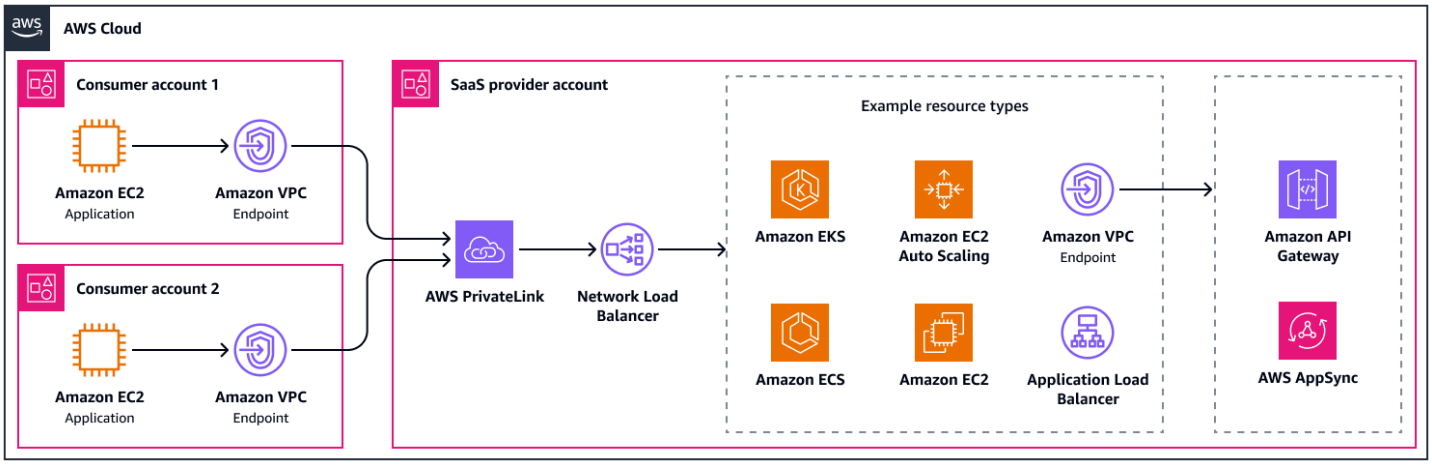
Il grafico radar mostra i seguenti valori.

Metrica di valutazione	AWS PrivateLink	Amazon VPC Lattice	Peering VPC	AWS Transit Gateway
Facilità di integrazione	5	5	4	3
TCO	5	5	3	4
Scalabilità	5	4	1	4
Adattabilità	4	5	2	3
Isolamento della rete	5	5	2	3
Osservabilità	4	5	4	4
È ora di riparare	5	5	5	4

Integrazione con AWS PrivateLink

AWS PrivateLinkè il modo più nativo per il cloud per integrare un'offerta SaaS. I provider SaaS possono ospitare le proprie applicazioni utilizzando un Network Load Balancer. Il Network Load Balancer si integra direttamente con un Application Load Balancer, Amazon Elastic Container Service (Amazon ECS), Amazon Elastic Kubernetes Service(Amazon EKS) e gruppi Auto Scaling. È anche possibile instradare il traffico dal Network Load Balancer per interfacciare gli endpoint VPC nell'account del provider SaaS. Questo ti aiuta a utilizzare un'API per raggiungere le applicazioni, ad esempio tramite Amazon API Gateway o AWS AppSync. Se la tua applicazione richiede l'accesso a risorse nell'ambiente del cliente che non sono bilanciate dal carico, come un database, puoi utilizzare gli endpoint VPC delle risorse.

AWS PrivateLink supporta una larghezza di banda fino a 100 Gbps per zona di disponibilità. Il diagramma seguente mostra una configurazione di base con alcune possibili integrazioni. Collega due account consumer all'account del provider SaaS tramite. AWS PrivateLink Sono presenti endpoint di servizio negli account dei consumatori e un Network Load Balancer nell'account del provider SaaS.



I vantaggi di questo approccio sono i seguenti:

- Facilità di integrazione: non sono richieste modifiche alla tabella delle rotte
- Facilità di integrazione: puoi offrire servizi endpoint tramite Marketplace AWS
- Facilità di integrazione: gli endpoint VPC supportano nomi DNS intuitivi
- Scalabilità: è scalabile fino a migliaia di consumatori SaaS
- Adattabilità: Supporto per intervalli CIDR sovrapposti
- Adattabilità: Support per IPv6
- Adattabilità: supporto interregionale
- TCO: AWS PrivateLink è un servizio completamente gestito, quindi richiede meno sforzi operativi
- Isolamento della rete: vantaggio in termini di sicurezza per il consumatore SaaS perché il traffico non può essere avviato dal provider SaaS
- Isolamento della rete: vantaggio in termini di sicurezza per il provider SaaS perché non espone un'intera sottorete o VPC

Di seguito sono riportati gli svantaggi di questo approccio:

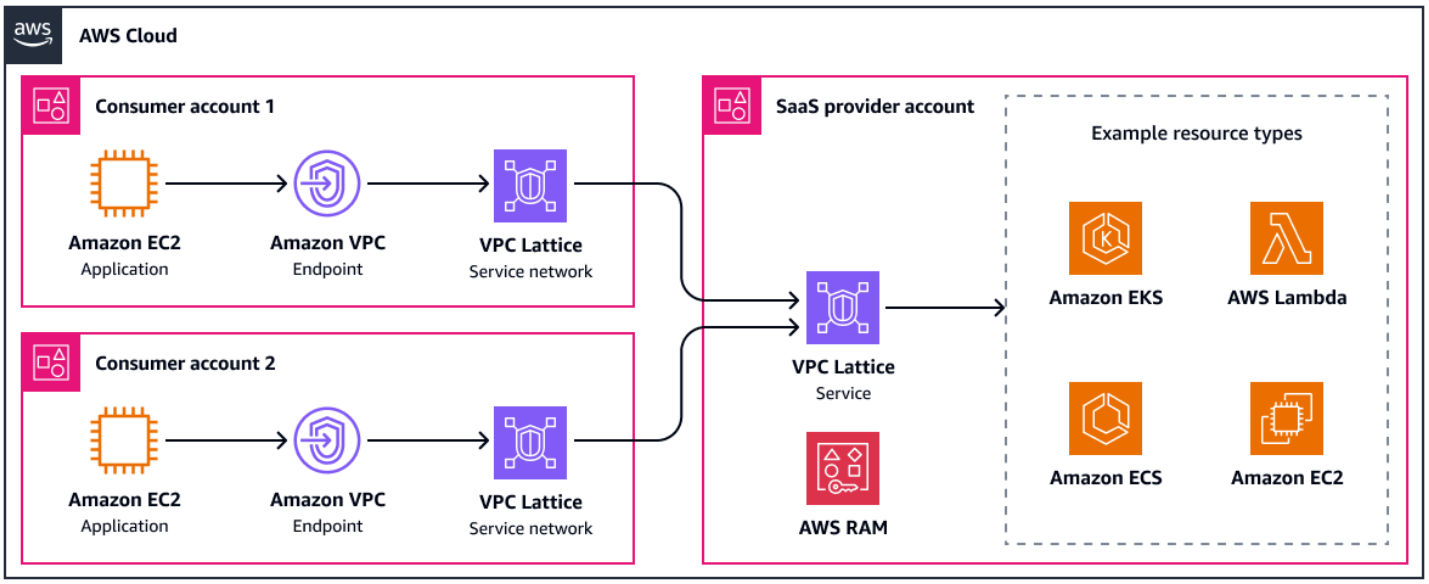
- Adattabilità: il provider SaaS deve utilizzare le stesse zone di disponibilità del consumatore
- Adattabilità: il supporto solo per le connessioni avviate dal client e gli endpoint VPC di risorse sono necessari per la comunicazione avviata dal servizio
- Adattabilità: Network Load Balancer è l'unica integrazione diretta per AWS PrivateLink

Condivisione di un servizio Amazon VPC Lattice

Per utilizzare Amazon VPC Lattice come opzione di connettività per la tua applicazione SaaS, devi prima creare uno o più servizi VPC Lattice che rappresentano i componenti dell'applicazione SaaS. Puoi configurare listener e regole di routing per indirizzare il traffico verso i tuoi obiettivi di backend, come istanze, contenitori o funzioni di Amazon EC2. AWS Lambda Per ulteriori informazioni, vedere Connessione dei servizi Saas all'interno di una rete di servizi VPC Lattice AWS (post del blog). Dal punto di vista concettuale, è quasi la stessa cosa che configurare un Application Load Balancer. Quindi, condividi il tuo servizio SaaS in modo sicuro con clienti Account AWS o organizzazioni utilizzando AWS Resource Access Manager (AWS RAM), specificando le autorizzazioni di cui dispongono. Dopo che i clienti hanno accettato la condivisione delle risorse, possono associare il servizio SaaS alle loro reti di servizi VPC Lattice esistenti o di nuova creazione per consentire la comunicazione. service-to-service

Ogni servizio VPC Lattice può supportare fino a 10 Gbps e 10.000 richieste al secondo per zona di disponibilità. Implementando le politiche di autenticazione, i tuoi clienti possono avere un controllo granulare su quali servizi e risorse possono accedere all'applicazione SaaS. È possibile utilizzare i gateway di risorse per accedere a risorse che richiedono una connessione TCP. Ad esempio, potrebbe trattarsi di un cluster Amazon EKS che gestisci o di una risorsa gestita dal cliente a cui l'applicazione deve accedere. Per ulteriori informazioni sull'utilizzo dei gateway di risorse per le offerte SaaS, consulta Estendere le funzionalità SaaS Account AWS utilizzando il AWS PrivateLink supporto per le risorse VPC (post del blog).AWS

Il diagramma seguente mostra una configurazione VPC Lattice di alto livello con alcuni esempi di integrazioni. Utilizza reti di servizi gestite dal cliente per accedere all'applicazione SaaS.



I vantaggi di questo approccio sono i seguenti:

- Facilità di integrazione: non sono richieste modifiche alla tabella di routing
- Facilità di integrazione: individuazione dei servizi pronta all'uso
- Scalabilità: è scalabile fino a migliaia di consumatori SaaS
- Adattabilità: Supporto per intervalli CIDR sovrapposti
- Adattabilità: Support per IPv6
- Adattabilità: si integra con qualsiasi servizio di AWS elaborazione come servizio VPC Lattice
- TCO: VPC Lattice è un servizio completamente gestito, quindi richiede meno sforzi operativi
- TCO: bilanciamento del carico integrato con routing avanzato del traffico
- Isolamento della rete: autorizzazione granulare con politiche di autenticazione
- Isolamento della rete: vantaggio in termini di sicurezza per il consumatore SaaS perché il traffico non può essere avviato dal provider SaaS
- Isolamento della rete: vantaggio in termini di sicurezza per il provider SaaS perché non si espone un'intera sottorete o VPC

Di seguito sono riportati gli svantaggi di questo approccio:

- Adattabilità: il supporto è richiesto solo per le connessioni avviate dal client e i gateway di risorse sono necessari per la comunicazione avviata dal servizio
- Adattabilità: nessun supporto interregionale

Creazione di connessioni peering VPC

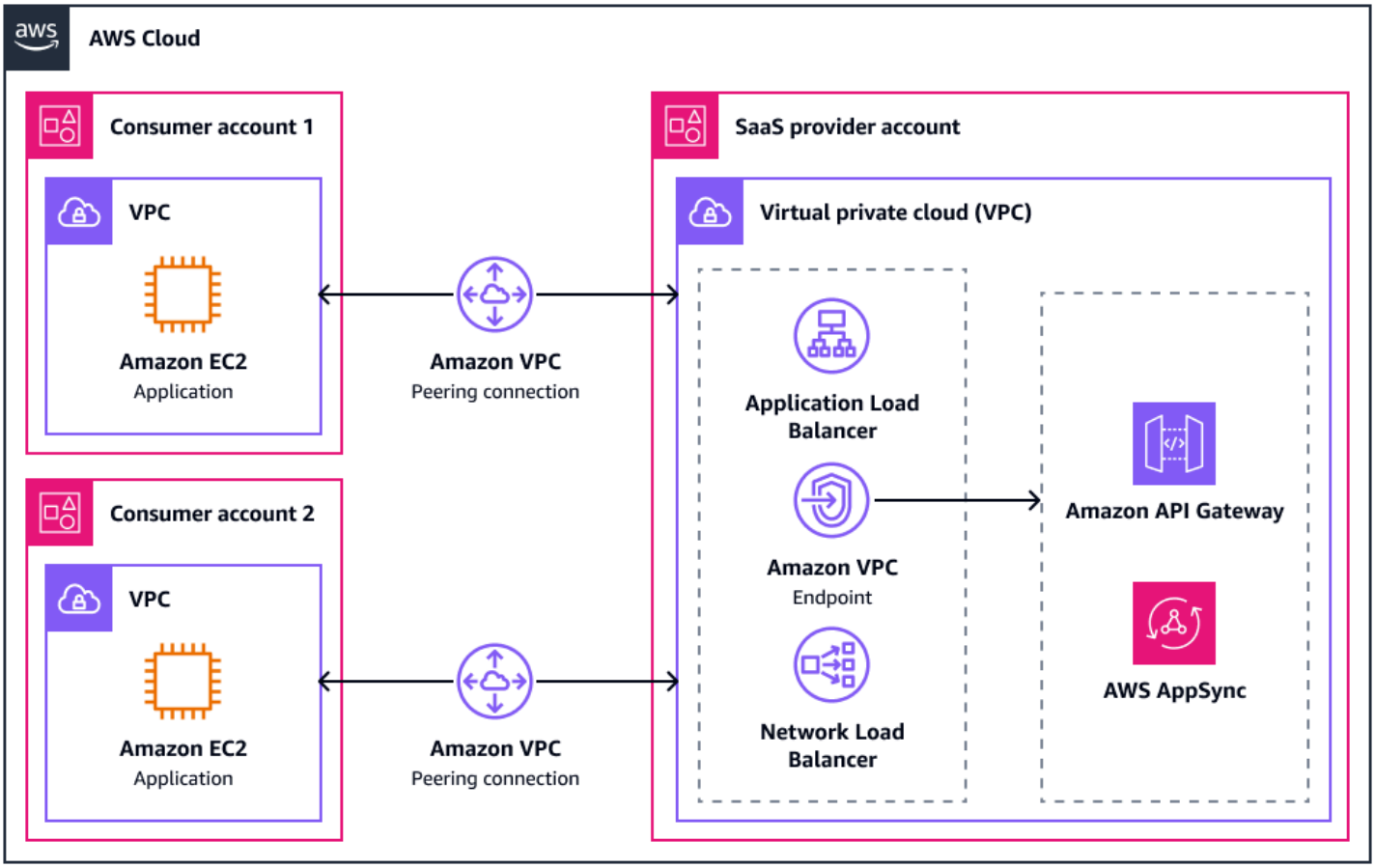
Quando utilizzi il peering VPC per connettere il VPC del provider SaaS con il VPC del consumatore, entrambe le parti sono in grado di avviare connessioni. Ciò richiede una corretta configurazione dei gruppi di sicurezza, dei firewall e delle liste di controllo degli accessi alla rete () in entrambi gli account. NACLs In caso contrario, il traffico indesiderato potrebbe entrare nella rete tramite la connessione peering. È possibile utilizzare i gruppi di sicurezza per fare riferimento ai gruppi di sicurezza provenienti dal VPCs peering. Questo può aiutarvi a controllare l'accesso all'applicazione, perché i gruppi di sicurezza che inseriscono gli elenchi consentiti forniscono un controllo degli accessi più esplicito e granulare rispetto agli indirizzi IP che elencano gli indirizzi IP consentiti.

Con il peering VPC, l'offerta SaaS può essere raggiunta tramite un servizio o una risorsa distribuiti nel VPC. La maggior parte delle applicazioni SaaS si basa su un Application Load Balancer o un

Network Load Balancer. AWS AppSync private APIs o Amazon API Gateway private APIs sono altri punti di accesso comuni alle applicazioni SaaS, in quanto possono fungere da destinazione tramite una connessione peering tramite endpoint VPC di interfaccia.

Dopo aver stabilito una connessione peering, è necessario aggiornare le tabelle di routing per entrambi gli account per definire la VPCs connessione peering come hop successivo per il rispettivo intervallo CIDR. Questa soluzione è consigliata solo ai provider SaaS che hanno pochi consumatori perché la gestione di più connessioni peering diventa rapidamente troppo complessa.

Il diagramma seguente mostra una configurazione di base con alcune possibili integrazioni. VPCs su due account consumer è presente una connessione peering con un VPC nell'account del provider SaaS.



I vantaggi di questo approccio sono i seguenti:

- Tempo di riparazione: nessun singolo punto di errore per la comunicazione
- Scalabilità: nessuna limitazione di larghezza di banda rispetto al peering VPC

- TCO: nessun costo per la connessione peering o il traffico sulla connessione peering all'interno della stessa zona di disponibilità
- TCO: nessuna infrastruttura da gestire
- Adattabilità: Support per IPv6
- Adattabilità: è supportato il peering interregionale

Di seguito sono riportati gli svantaggi di questo approccio:

- Adattabilità: nessun supporto per il routing transitivo
- Adattabilità: nessun supporto per intervalli CIDR sovrapposti
- Scalabilità: scalabilità limitata (massimo 125 connessioni peering per VPC)
- TCO: la complessità cresce esponenzialmente con ogni connessione peering aggiuntiva
- TCO: sovraccarico derivante dalla gestione delle tabelle delle rotte, dal peering stesso delle connessioni, dalle regole dei gruppi di sicurezza e dall'ispezione del traffico
- Isolamento della rete: sono necessari controlli di sicurezza rigorosi poiché entrambe le parti sono esposte VPCs

Connessione VPCs con AWS Transit Gateway

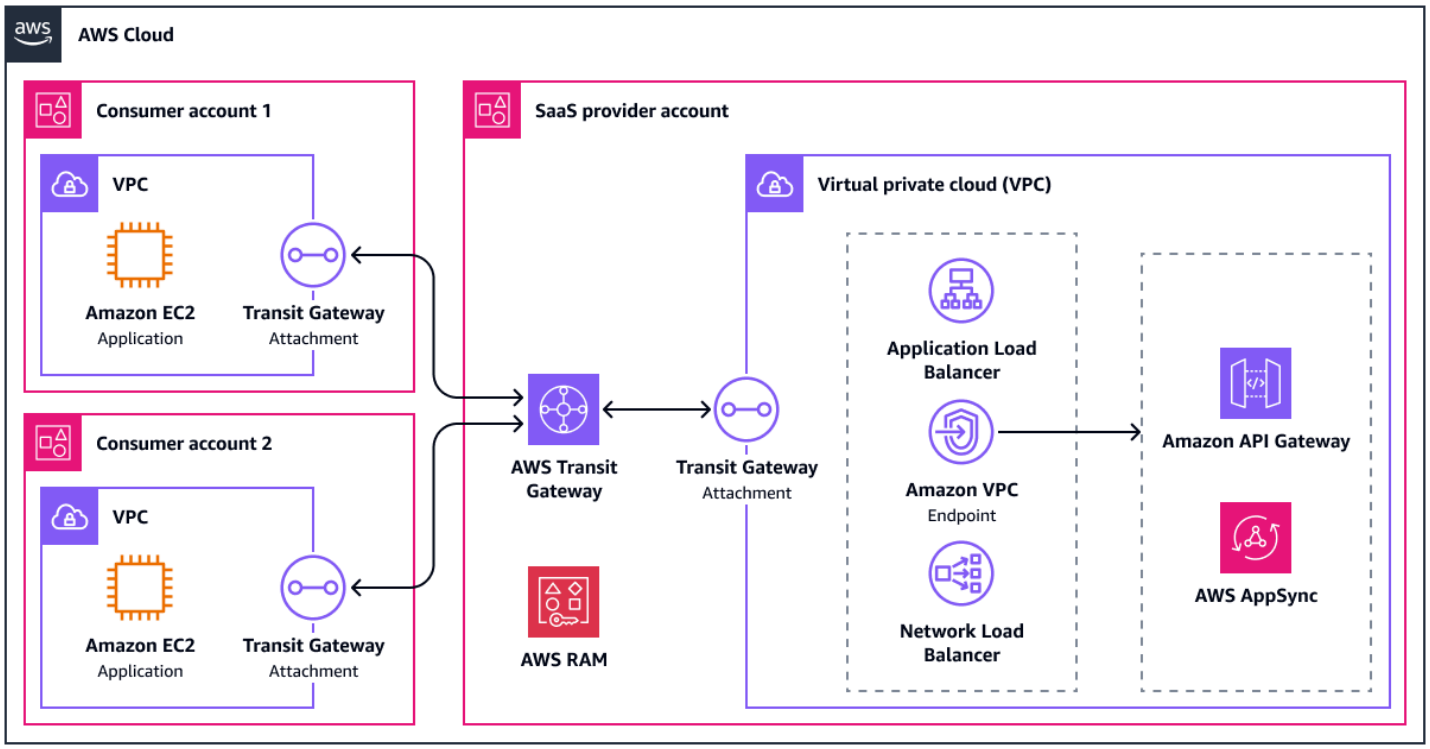
Quando ti connetti VPCs AWS Transit Gateway, crea allegati VPC e distribuisce interfacce di rete nelle sottoreti di ogni zona di disponibilità che dovrebbero instradare il traffico da e verso il VPC. Si consiglia di disporre di una `/28` sottorete dedicata in ogni zona di disponibilità per l'allegato VPC. Per ulteriori informazioni, consulta le best practice di progettazione di Amazon VPC Transit Gateway. VPCs È necessaria una tabella di routing aggiornata per inviare il traffico attraverso l'interfaccia di rete distribuita e le tabelle di routing Transit Gateway devono essere aggiornate di conseguenza. In una configurazione multi-tenant, si desidera che il VPC del provider SaaS disponga di un percorso verso tutti i consumatori. VPCs Il consumatore VPCs dovrebbe avere un percorso solo verso il VPC del provider SaaS.

Transit Gateway è altamente disponibile fin dalla sua progettazione. Supporta il monitoraggio con VPC Flow Logs e la larghezza di banda massima per un allegato Transit Gateway è di 100 Gbps per zona di disponibilità. Come il peering VPC, questo approccio consente il riferimento ai gruppi di sicurezza tra VPC, semplificando il controllo degli accessi tra gli ambienti.

Esistono due opzioni principali per connettere i consumatori alla tua offerta SaaS con Transit Gateway.

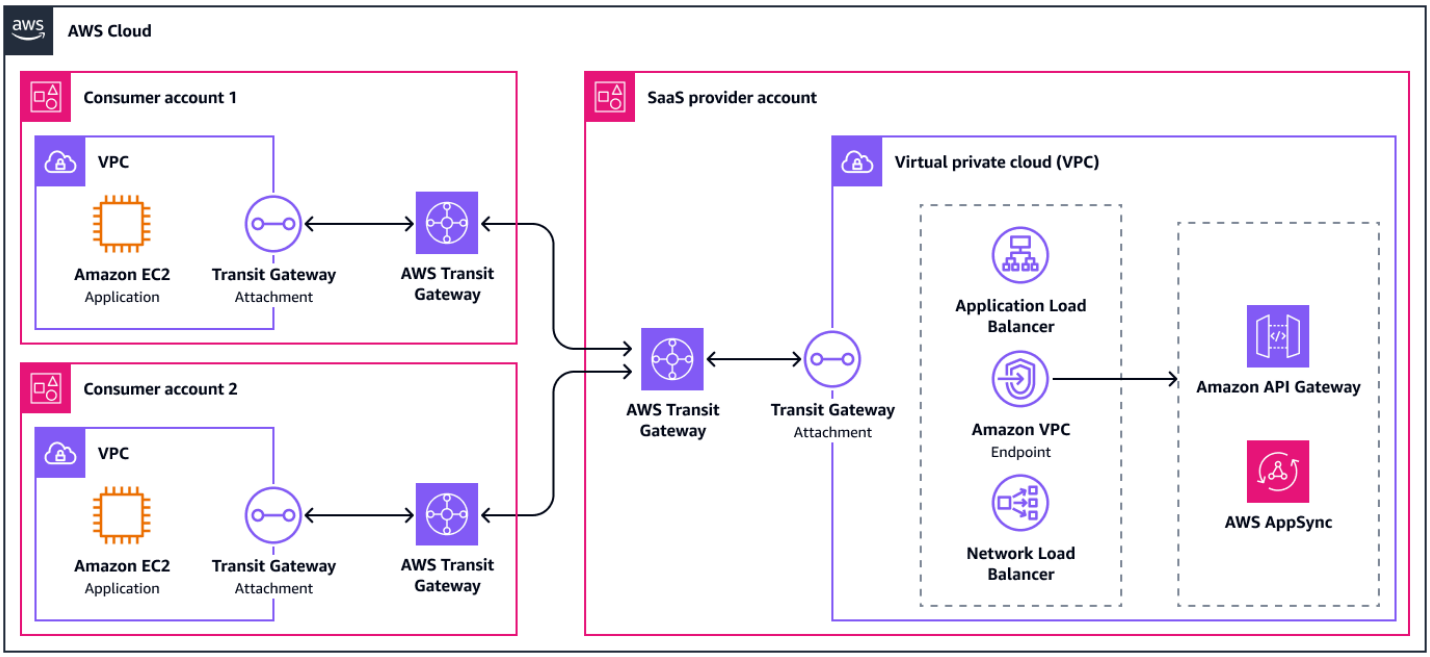
Opzione 1: utilizzo della RAM

Nella prima opzione, il fornitore di servizi condivide il Transit Gateway con i consumatori utilizzando AWS Resource Access Manager (AWS RAM). Ciò consente ai consumatori di implementare gli allegati VPC nei propri account. Il diagramma seguente mostra questa opzione a un livello elevato.



Opzione 2: gateway di transito peer-to-peer

La seconda opzione consiste nel collegare il gateway di transito con un gateway di transito negli account dei consumatori. Ciò offre ai consumatori una maggiore flessibilità perché ora possono controllare completamente le tabelle di percorso all'interno del loro gateway di transito. Ad esempio, potrebbero impostare un'ispezione centralizzata tra il servizio e i loro carichi di lavoro. Uno svantaggio di questa opzione è che è supportato solo il routing statico tra i gateway di transito. Il diagramma seguente mostra questa opzione a un livello elevato.



I vantaggi di questo approccio sono i seguenti:

- Scalabilità: Support per un massimo di 5.000 allegati
- Scalabilità: un unico posto per gestire e monitorare tutti i dispositivi connessi VPCs
- Adattabilità: Transit Gateway può essere collegato anche a VPNs Direct Connect gateway e dispositivi SD-WAN di terze parti
- Adattabilità: architettura flessibile, come l'aggiunta di un VPC di ispezione
- Adattabilità: Supporto per il routing transitivo
- Adattabilità: può collegare gateway di transito intra-regionali e interregionali
- Adattabilità: Support per IPv6
- TCO: AWS Transit Gateway è un servizio completamente gestito, quindi richiede meno sforzi operativi
- TCO: il TCO cresce in modo lineare con ogni collegamento aggiuntivo al gateway di transito

Di seguito sono riportati gli svantaggi di questo approccio:

- Facilità di integrazione: la configurazione del routing richiede conoscenze di rete avanzate
- Adattabilità: nessun supporto per intervalli CIDR sovrapposti
- TCO: sovraccarico derivante dalla gestione delle voci delle tabelle dei percorsi, delle regole dei gruppi di sicurezza e dell'ispezione del traffico

- Sicurezza: sono necessari controlli di sicurezza rigorosi poiché entrambe le VPCs parti sono esposte

Consumatori di servizi che operano in sede

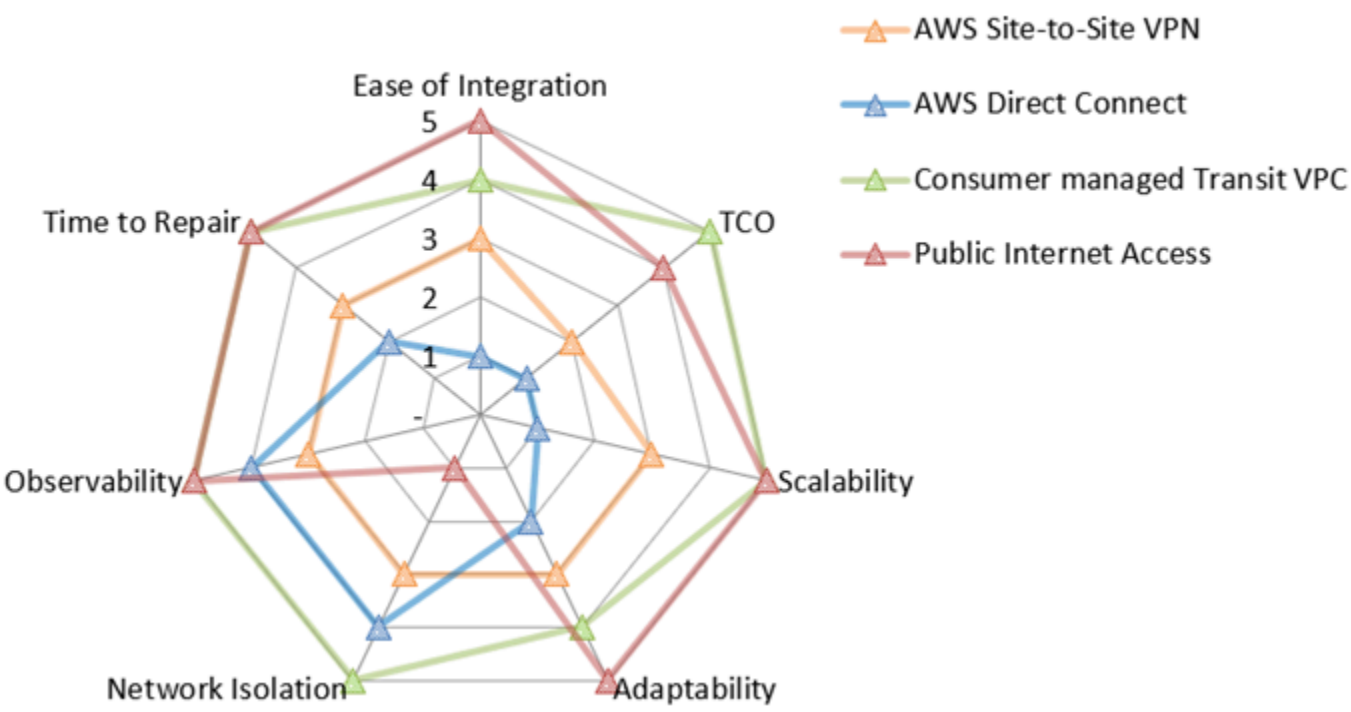
Questa sezione illustra le opzioni di connettività tra i carichi di lavoro SaaS nei data center locali Cloud AWS e quelli locali. Molti consumatori con requisiti locali, in particolare a livello aziendale, vedono il cloud come un'estensione della loro rete fisica e vogliono rifletterli nella loro architettura. Ciò significa connettività privata all'offerta SaaS nel cloud, tramite tunnel logici o anche tramite una connessione fisica privata. Altri consumatori accetteranno la connettività tramite la rete Internet pubblica, argomento trattato anche in questa sezione.

In questa sezione vengono descritti i seguenti approcci di accesso alla rete:

- Connettersi con AWS Site-to-Site VPN
- Connessione con AWS Direct Connect
- Connessione con un'architettura VPC di transito
- Connessione tramite Internet pubblico

La seguente mappa dei valori di rete riassume il punteggio di ciascuna di queste opzioni per ogni metrica di valutazione. Per ulteriori informazioni sulle metriche di valutazione, consulta Metriche di valutazione in questa guida. Nella mappa, un cinque rappresenta il punteggio migliore, ad esempio il TCO più basso, il miglior isolamento di rete o il minor tempo di riparazione. Per ulteriori informazioni su come leggere questo grafico radar, Mappa dei valori della rete consulta questa guida.

> **Note**
>
> L'opzione VPC di transito gestito dal provider è esclusa perché i punteggi dipendono in larga misura dai servizi gestiti.



Il grafico radar mostra i seguenti valori.

Metrica di valutazione	AWS Site-to-Site VPN	AWS Direct Connect	VPC di transito gestito dai consumatori	Accesso pubblico a Internet
Facilità di integrazione	3	1	4	5
TCO	2	1	5	4
Scalabilità	3	1	5	5
Adattabilità	3	2	4	5
Isolamento della rete	3	4	5	1
Osservabilità	3	4	5	5
È ora di riparare	3	2	5	5

Connettersi con AWS Site-to-Site VPN

AWS Site-to-Site VPNle connessioni possono terminare su un gateway privato virtuale o su un gateway di transito. Un gateway privato virtuale è l'endpoint VPN sul AWS lato della connessione Site-to-Site VPN che può essere collegato a un singolo VPC. Un gateway di transito è un hub di transito che può essere utilizzato per interconnettere più reti locali VPCs . Può essere utilizzato anche come endpoint VPN per il AWS lato della Site-to-Site connessione VPN. Questa sezione illustra entrambe le opzioni.

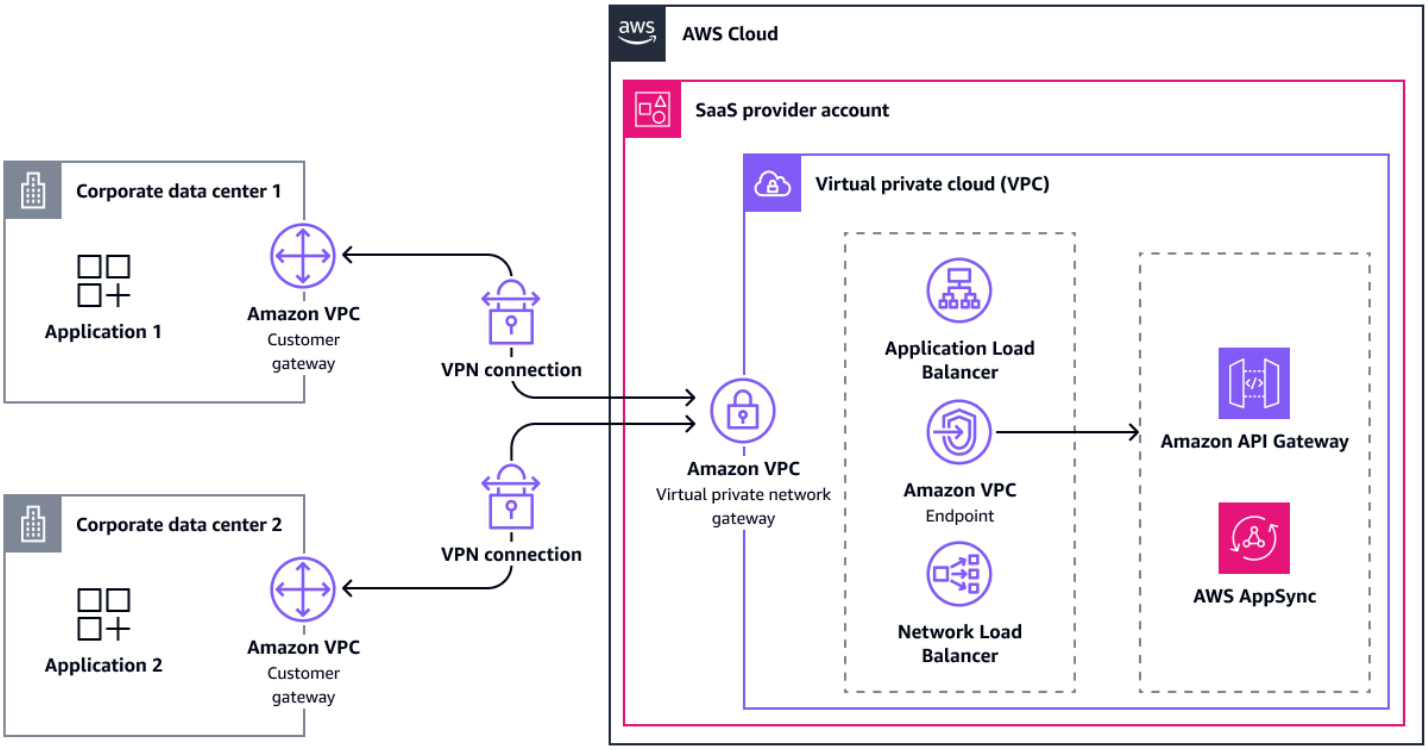
Connessione tramite un gateway privato virtuale

Dopo aver creato un gateway privato virtuale, lo colleghi al VPC che contiene la tua offerta SaaS. Quindi, abiliti la propagazione delle rotte per propagare le rotte VPN alla tabella delle rotte VPC. Questi percorsi possono essere statici o dinamici pubblicizzati da BGP.

Per garantire un'elevata disponibilità, una connessione Site-to-Site VPN dispone di due tunnel VPN che terminano lateralmente in due zone di disponibilità. AWS Se uno diventa non disponibile, il secondo tunnel può prendere il sopravvento. Un singolo tunnel consente una larghezza di banda massima di 1,25 Gbps. Poiché i gateway privati virtuali non supportano il routing multipercorso (ECMP) a costo uguale, è possibile utilizzare solo un tunnel alla volta.

Per aumentare la tolleranza agli errori, è possibile configurare una seconda connessione VPN a un secondo gateway fisico per il cliente. Una volta stabilita la connessione, il consumatore può accedere alle risorse nel VPC del provider SaaS.

Il diagramma seguente mostra questa architettura.



I vantaggi di questo approccio sono i seguenti:

- Tempo di riparazione: failover gestito sul tunnel VPN secondario
- Osservabilità: integrazione per il monitoraggio attivo gestito utilizzando Network Synthetic Monitor
- Facilità di integrazione: supporto per il routing dinamico tramite BGP
- Adattabilità: compatibilità con la maggior parte delle apparecchiature di rete locali
- Adattabilità: supporto IPv6
- TCO: AWS Site-to-Site VPN è un servizio completamente gestito, quindi richiede meno sforzi operativi
- TCO: nessun costo per i gateway virtuali, sebbene siano previsti costi per i due indirizzi pubblici IPv4 su ciascuno
- Isolamento della rete: consente comunicazioni private sicure tramite Internet

Di seguito sono riportati gli svantaggi di questo approccio:

- Facilità di integrazione: il consumatore deve configurare il proprio gateway per il cliente
- Scalabilità: la mancanza del supporto ECMP limita la larghezza di banda a 1,25 Gbps per gateway virtuale

- Scalabilità: scalabilità limitata dovuta alla maggiore complessità della rete e al sovraccarico operativo
- Adattabilità: IPv6 supporto solo per gli indirizzi IP interni dei tunnel VPN
- Adattabilità: nessun routing transitivo
- TCO: sovraccarico operativo per mantenere, gestire e configurare numerose connessioni VPN per il provider SaaS

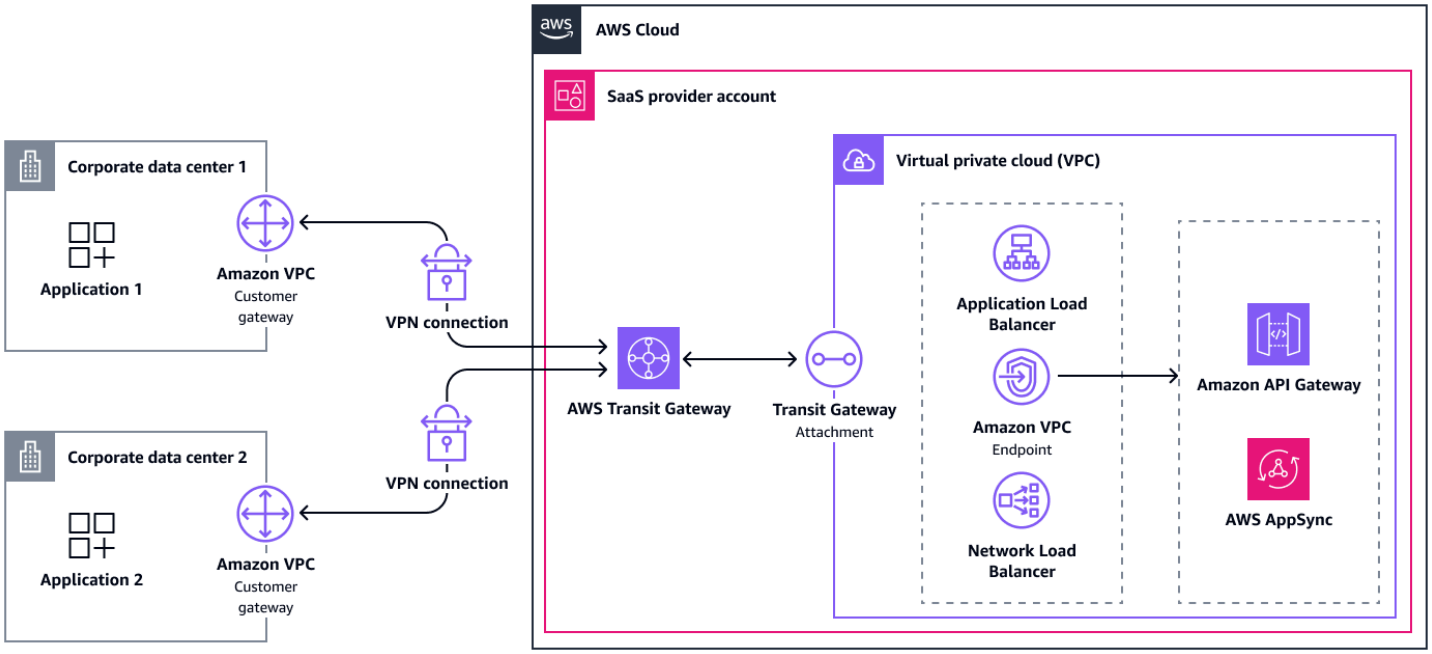
Connessione tramite un gateway di transito

Le connessioni tramite gateway di transito sono simili ai gateway virtuali. Tuttavia, ci sono alcune differenze da tenere a mente.

Innanzitutto, le route per l'allegato VPN possono essere propagate automaticamente all'interno della tabella delle rotte del gateway di transito, ma è necessario aggiungere manualmente le rotte all'allegato VPCs.

Rispetto a un gateway virtuale, Transit Gateway supporta ECMP. Se il gateway del cliente supporta ECMP, può utilizzare entrambi i tunnel per raggiungere un throughput massimo totale di 2,5 Gbps. È possibile stabilire più connessioni tra la stessa rete locale e il gateway di transito. Utilizzando questo approccio, è possibile aumentare la larghezza di banda massima fino a 2,5 Gbps per connessione.

Il diagramma seguente mostra questa architettura.



I vantaggi di questo approccio sono i seguenti:

- Tempo di riparazione: failover gestito sul tunnel VPN secondario
- Osservabilità: integrazione per il monitoraggio attivo gestito utilizzando Network Synthetic Monitor
- Facilità di integrazione: supporto per il routing dinamico tramite BGP
- Scalabilità: il supporto ECMP consente di scalare il throughput VPN per soddisfare requisiti di ampia larghezza di banda
- Scalabilità: gran numero di connessioni VPN supportate da un unico gateway di transito (fino a quasi 5.000)
- Scalabilità: un unico posto per gestire e monitorare tutte le connessioni VPN
- Adattabilità: compatibilità con la maggior parte delle apparecchiature di rete locali
- Adattabilità: supporto IPv6
- Adattabilità: eredita la flessibilità di AWS Transit Gateway
- TCO: AWS Transit Gateway è un servizio completamente gestito, quindi richiede meno sforzi operativi
- TCO: nessun costo per i gateway virtuali, sebbene siano previsti costi per i due indirizzi pubblici IPv4 su ciascuno
- Isolamento della rete: consente comunicazioni private sicure tramite Internet

Di seguito sono riportati gli svantaggi di questo approccio:

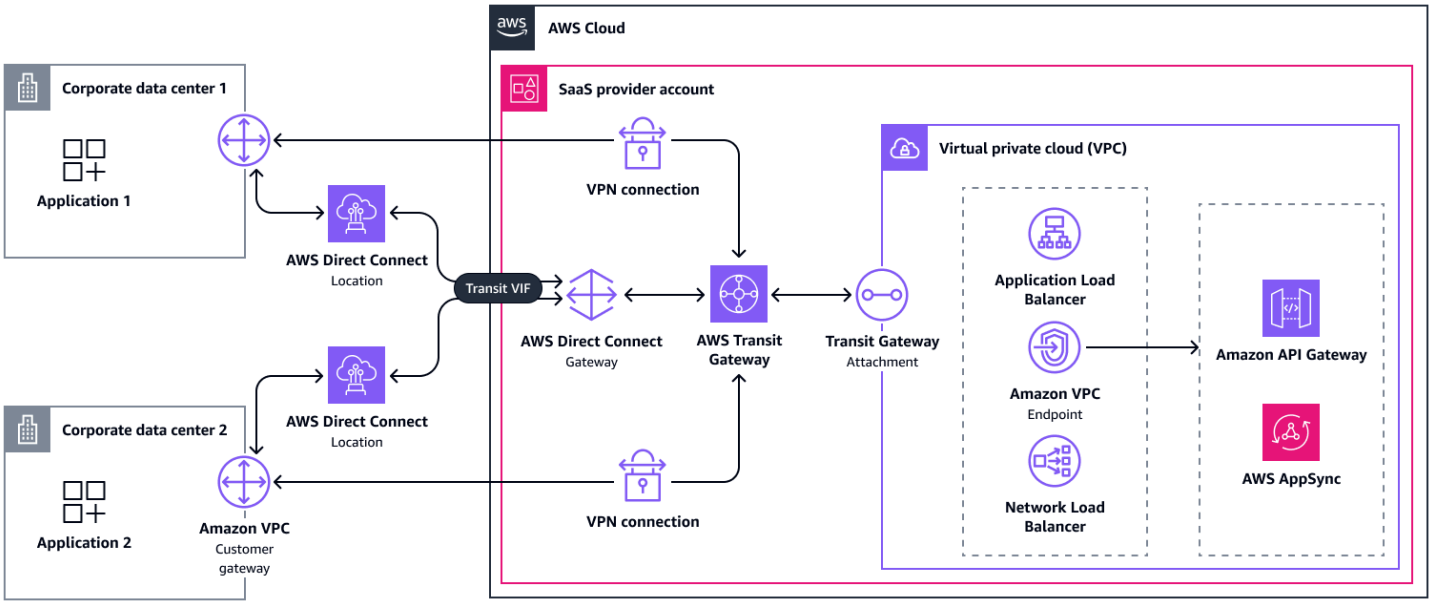
- Facilità di integrazione: il consumatore deve configurare il proprio gateway per il cliente
- Scalabilità: scalabilità limitata dovuta alla maggiore complessità della rete e al sovraccarico operativo
- Adattabilità: IPv6 supporto solo per gli indirizzi IP interni dei tunnel VPN
- TCO: sovraccarico operativo per mantenere, gestire e configurare numerose connessioni VPN per il provider SaaS
- TCO: costi aggiuntivi per l'utilizzo di AWS Transit Gateway
- TCO: ulteriore complessità nella gestione delle tabelle di routing dei gateway di transito

Connessione con AWS Direct Connect

AWS Direct Connectcollega la rete interna a una Direct Connect posizione tramite un cavo Ethernet standard in fibra ottica. A differenza delle altre opzioni di architettura, non è possibile stabilire una connessione dedicata in pochi minuti. Al contrario, questo processo può richiedere fino a diversi giorni se tutti i requisiti sono soddisfatti. In caso contrario, potrebbe volerci più tempo. Pertanto, ti suggeriamo di contattare il team del tuo AWS account o di Supporto AWS chiedere aiuto in merito a questo approccio. Facoltativamente, puoi scegliere una connessione ospitata fornita da un AWS partner e condivisa con altri clienti. L'architettura è la stessa a prescindere. Puoi scegliere Direct Connect perché riduce la latenza, migliora la larghezza di banda o è conforme ai requisiti normativi.

Per utilizzare la Direct Connect connessione, i consumatori devono creare un'interfaccia virtuale pubblica, privata o di transito. Sono disponibili diverse opzioni di architettura. La più flessibile a cui connettere più sedi locali Cloud AWS è un'interfaccia virtuale di transito connessa a un Direct Connect gateway. Un Direct Connect gateway è un componente logico globale che consente al fornitore di servizi di collegare fino a sei gateway di transito ad esso. Inoltre, è possibile connettere fino a 30 interfacce virtuali al gateway. Per motivi di scalabilità, puoi creare Direct Connect gateway aggiuntivi. Nell'account del provider SaaS, i gateway di transito si collegano quindi a VPCs, come descritto in precedenza.

I consumatori possono connettersi utilizzando da una a quattro Direct Connect connessioni da un totale di una o due Direct Connect postazioni, a seconda del livello di resilienza desiderato. Per ulteriori informazioni, consulta Configurazione Direct Connect per la massima resilienza. Una AWS Site-to-Site VPN connessione via Internet può anche fungere da percorso di backup a basso costo per una connessione. Direct Connect Le connessioni Direct Connect dedicate supportate possono essere utilizzate MACsecper crittografare il collegamento sul livello 2 tra la Direct Connect posizione e il data center. È comune disporre di una connessione Site-to-Site VPN per una maggiore riservatezza dei dati. La connessione Site-to-Site VPN può terminare sul gateway di transito utilizzando un normale allegato VPN. Il diagramma seguente mostra questa architettura.



I vantaggi di questo approccio sono i seguenti:

- Osservabilità: integrazione per il monitoraggio attivo gestito utilizzando Network Synthetic Monitor
- Scalabilità: Supporto per una maggiore velocità di trasmissione della larghezza di banda
- Adattabilità: supporto IPv6
- TCO: possibilità di ridurre il trasferimento dei dati
- TCO: esperienza di rete coerente
- Isolamento della rete: connettività privata in grado di soddisfare i requisiti normativi

Di seguito sono riportati gli svantaggi di questo approccio:

- Facilità di integrazione: tempo e impegno manuale per la configurazione
- Scalabilità: scalabilità limitata oltre le decine di Direct Connect connessioni perché ci sono più quote da tracciare
- Adattabilità: le opzioni di configurazione dipendono dalle posizioni disponibili Direct Connect
- TCO: la Direct Connect manutenzione programmata può causare tempi di inattività che richiedono un intervento

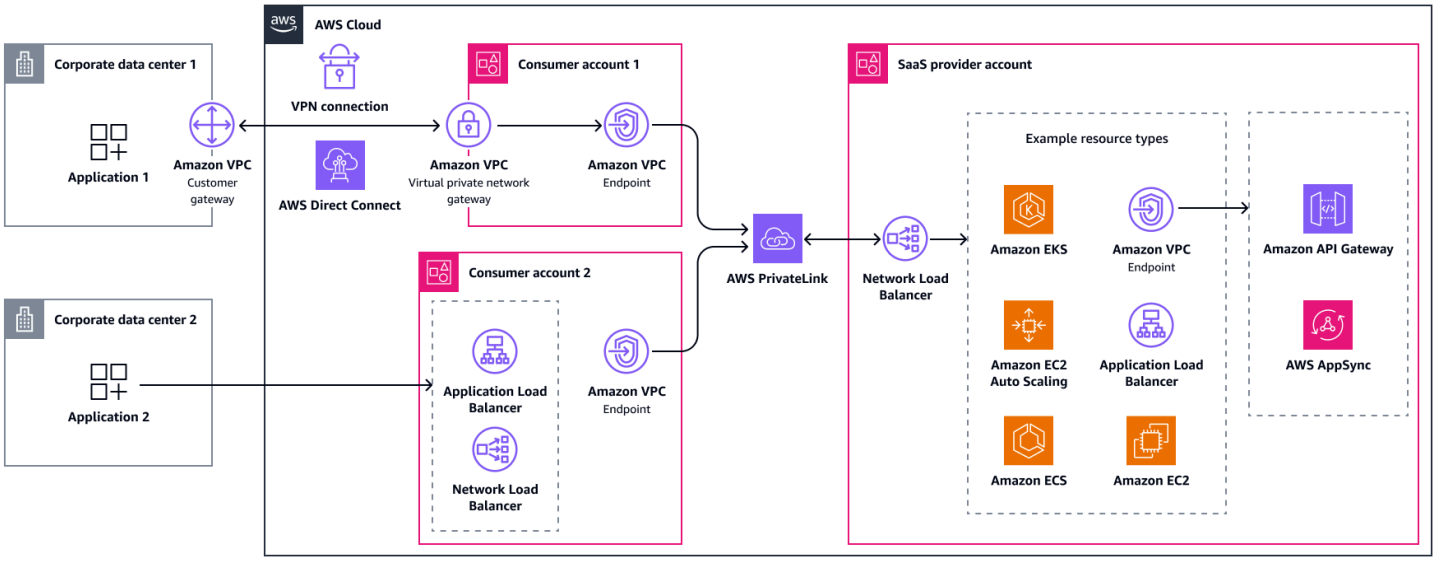
Connessione con un'architettura VPC di transito

Transit VPC è un'opzione di architettura che offre flessibilità ai consumatori su come connettersi e consente ai AWS provider SaaS di trarre vantaggio dall'accesso unificato al proprio servizio tramite. AWS PrivateLink Il consumatore si connette dall'ambiente locale a un VPC di transito che contiene solo un punto di ingresso (come un gateway privato virtuale) e un endpoint VPC di interfaccia, che è una risorsa. AWS PrivateLink Il transito VPCs dovrebbe essere di proprietà del provider SaaS o dei consumatori. Questa sezione illustra entrambe le opzioni.

È possibile creare il VPC di transito e le sottoreti con intervalli CIDR compatibili con il data center locale. Se richiedono una connettività privata, i consumatori possono connettersi a quel VPC tramite AWS Direct Connect o. AWS Site-to-Site VPN Puoi anche configurare l'accesso all'account di transito dalla rete Internet pubblica utilizzando un Application Load Balancer o un Network Load Balancer che punti all'endpoint VPC.

VPC di transito gestito dai consumatori

In questo approccio, il provider SaaS lascia la gestione del VPCs transito ai consumatori. Da un punto di vista tecnico, l'architettura del provider SaaS è la stessa di quando si connette ai consumatori in modalità through. Cloud AWS AWS PrivateLink Dal punto di vista delle vendite e del prodotto, si tratta di uno sforzo aggiuntivo, perché alcuni consumatori non lo hanno Account AWS ancora fatto. Potrebbero essere riluttanti ad aprire e gestire un account. Il provider SaaS dovrebbe fornire indicazioni ai propri consumatori su come creare Account AWS e connettere il proprio data center locale. Il diagramma seguente mostra una combinazione di accesso pubblico e privato, in cui il transito è di proprietà dei consumatori. VPCs



I vantaggi di questo approccio sono i seguenti:

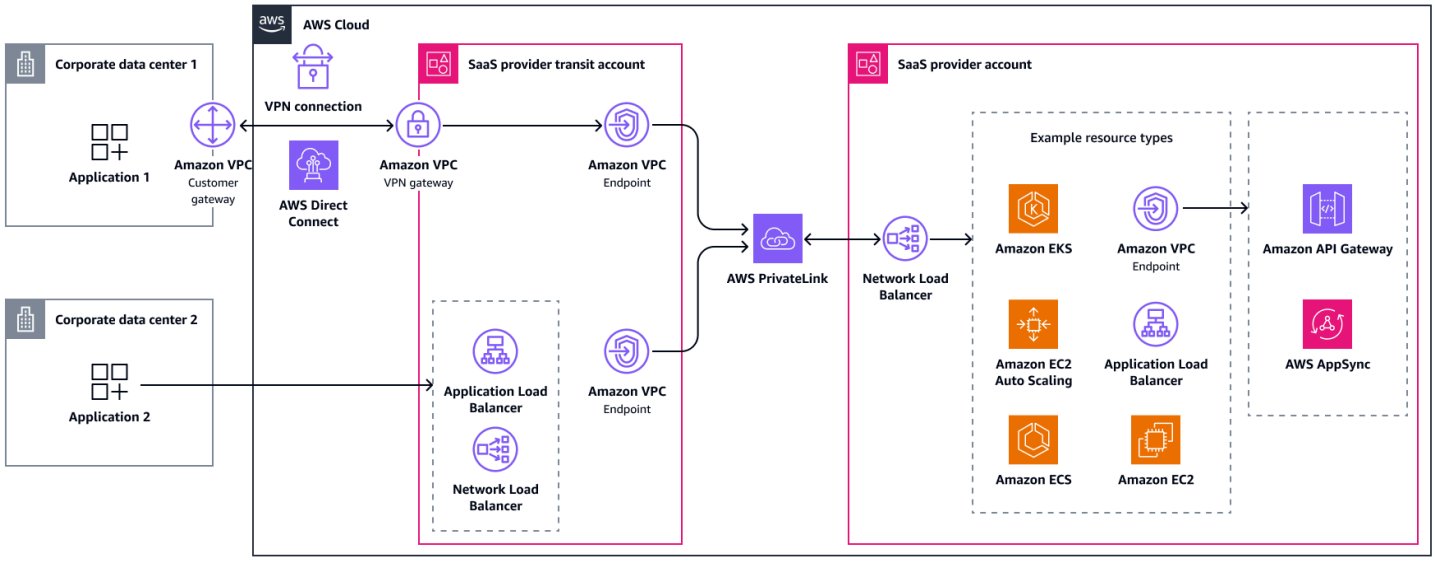
- Tempo di riparazione: il sovraccarico operativo viene in gran parte trasferito ai consumatori SaaS
- Adattabilità: i consumatori SaaS possono scegliere tra diverse opzioni di accesso
- Adattabilità: nessun conflitto nell'intervallo CIDR, anche quando si utilizza VPN o Site-to-Site Direct Connect
- Tutte le metriche: il fornitore di servizi eredita i vantaggi AWS PrivateLink

Di seguito sono riportati gli svantaggi di questo approccio:

- Facilità di integrazione: i consumatori SaaS ne richiedono almeno uno Account AWS
- TCO: un VPC di transito è un'architettura, non un servizio completamente gestito, quindi richiede un maggiore impegno operativo

VPC di transito gestito dal provider

Questo approccio utilizza le stesse tecnologie, ma i confini e le responsabilità degli account cambiano. Qui, il provider SaaS è proprietario del transito VPCs, preferibilmente in un account separato dall'offerta SaaS. Questo disaccoppiamento riduce i costi, riduce i rischi e consente all'account di transito di scalare in modo indipendente. Per gli ambienti che richiedono un elevato grado di isolamento, è possibile creare una separazione aggiuntiva tra i tenant utilizzando una sottorete o creando un VPC di transito separato per ogni consumatore. I consumatori possono quindi scegliere come connettersi al VPC di transito. Questo approccio offre più opzioni per espandere il mercato indirizzabile totale, ma ha un TCO più elevato per il provider SaaS a causa della necessità di gestire e monitorare componenti architettonici aggiuntivi.



I vantaggi di questo approccio sono i seguenti:

- Adattabilità: i consumatori SaaS possono scegliere tra diverse opzioni di accesso
- Adattabilità: i consumatori SaaS non hanno bisogno di avere un Account AWS
- Adattabilità: nessun conflitto nell'intervallo CIDR, anche quando si utilizza VPN o Site-to-Site Direct Connect

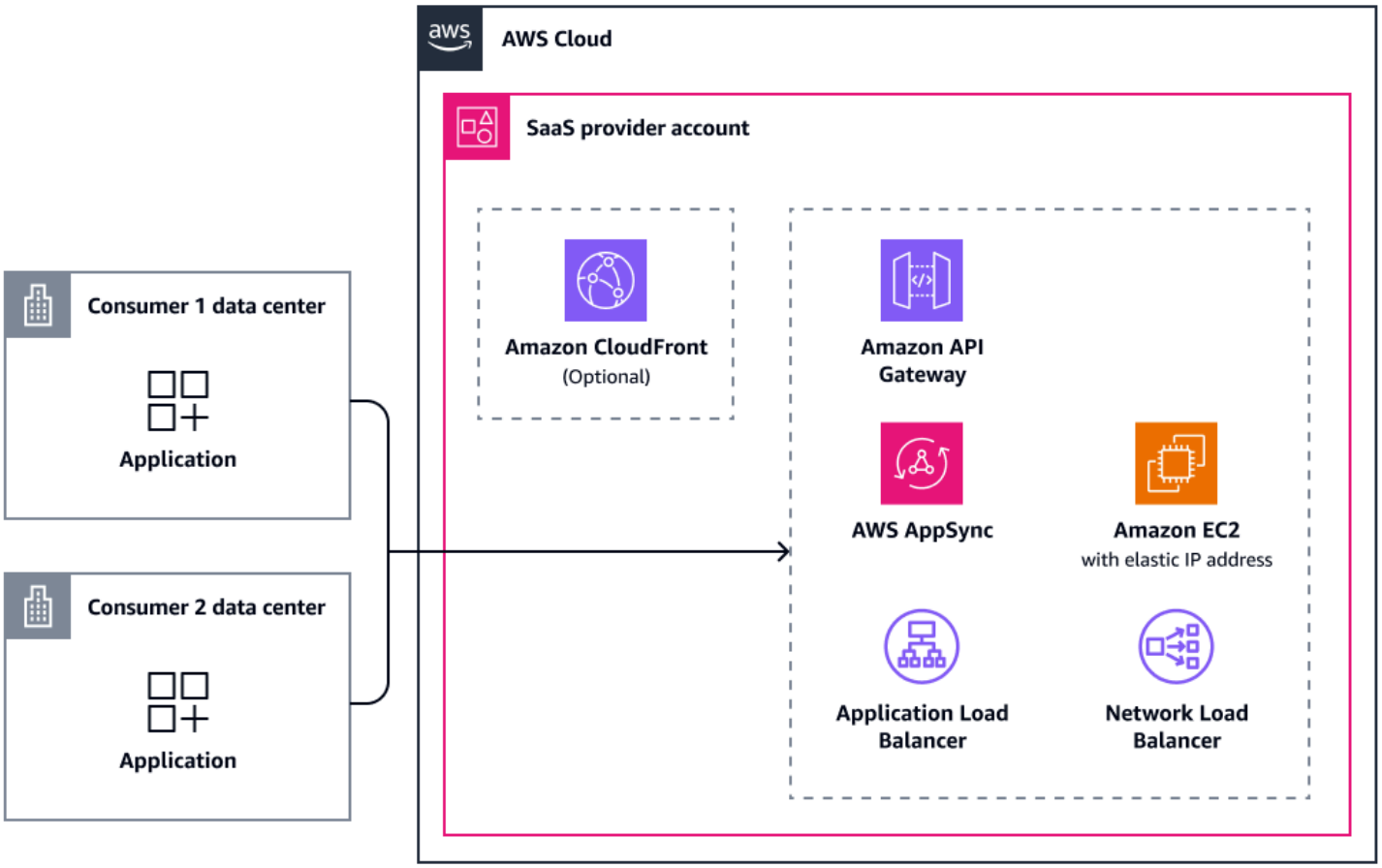
Di seguito sono riportati gli svantaggi di questo approccio:

- TCO: un VPC di transito è un'architettura, non un servizio completamente gestito, quindi richiede un maggiore impegno operativo
- TCO: il provider SaaS deve gestire e monitorare componenti architettonici aggiuntivi

Connessione tramite Internet pubblico

L'accesso pubblico a Internet è anche un'opzione valida per fornire l'accesso a un'offerta SaaS, sebbene non offra connettività privata nel senso tradizionale. Alcuni consumatori potrebbero comunque preferire un approccio ad accesso pubblico perché non richiede un'infrastruttura di rete aggiuntiva tra loro e il provider SaaS. Riduce la complessità, i costi e i tempi di integrazione in cambio di una maggiore superficie di attacco. Potenti meccanismi di autenticazione e autorizzazione possono aiutare a mitigare l'aumento del livello di minaccia e dovresti sempre crittografare il traffico. Si consiglia comunque di disporre di un ulteriore livello di sicurezza in questo scenario, ad esempio utilizzando. AWS WAF

L'architettura in questo scenario è semplice. Il consumatore si connette a un host pubblico (il provider SaaS) tramite Internet. L'applicazione può essere ospitata direttamente su un'istanza pubblica di Amazon Elastic Compute Cloud (Amazon EC2) con un indirizzo IP elastico. L'opzione preferita è ospitarlo dietro un Application Load Balancer o un servizio simile. Per prestazioni migliori e memorizzazione nella cache delle risorse statiche, puoi utilizzare una rete di distribuzione di contenuti, come Amazon CloudFront. Per servire un'applicazione con una latenza minima su due indirizzi IP Anycast statici globali, puoi posizionarla davanti AWS Global Acceleratora un'istanza Amazon EC2, Network Load Balancer o Application Load Balancer. Inoltre CloudFront, Application Load Balancer e Amazon API Gateway si integrano tutti con. AWS AppSync AWS WAF Il diagramma seguente fornisce una panoramica delle opzioni di connettività per l'accesso pubblico a Internet.



La tabella seguente descrive i protocolli e le integrazioni supportati per questo scenario.

Servizio o risorsa	IPv6	AWS WAF integration	Può essere un endpoint di Global Accelerator

Amazon CloudFront	Supportata	Supportata	Non supportata
Gateway Amazon API	Supportata	Supportata	Non supportata
AWS AppSync	Supportato parzialmente	Supportata	Non supportata
Amazon EC2 con un indirizzo IP elastico	Supportata	Non supportata	Supportata
Application Load Balancer	Supportata	Supportato	Supportata
Network Load Balancer	Supportata	Non supportata	Supportata

I vantaggi di questo approccio sono i seguenti:

- Facilità di integrazione: semplicità e accessibilità
- Scalabilità: scalabilità illimitata
- Adattabilità: non sono possibili conflitti di intervallo CIDR
- Adattabilità: supporto CloudFront

Di seguito sono riportati gli svantaggi di questo approccio:

- Isolamento della rete: nessuna connettività privata
- Isolamento della rete: sono necessarie forti misure di sicurezza

Si applicano altri vantaggi e svantaggi, a seconda dei servizi scelti.

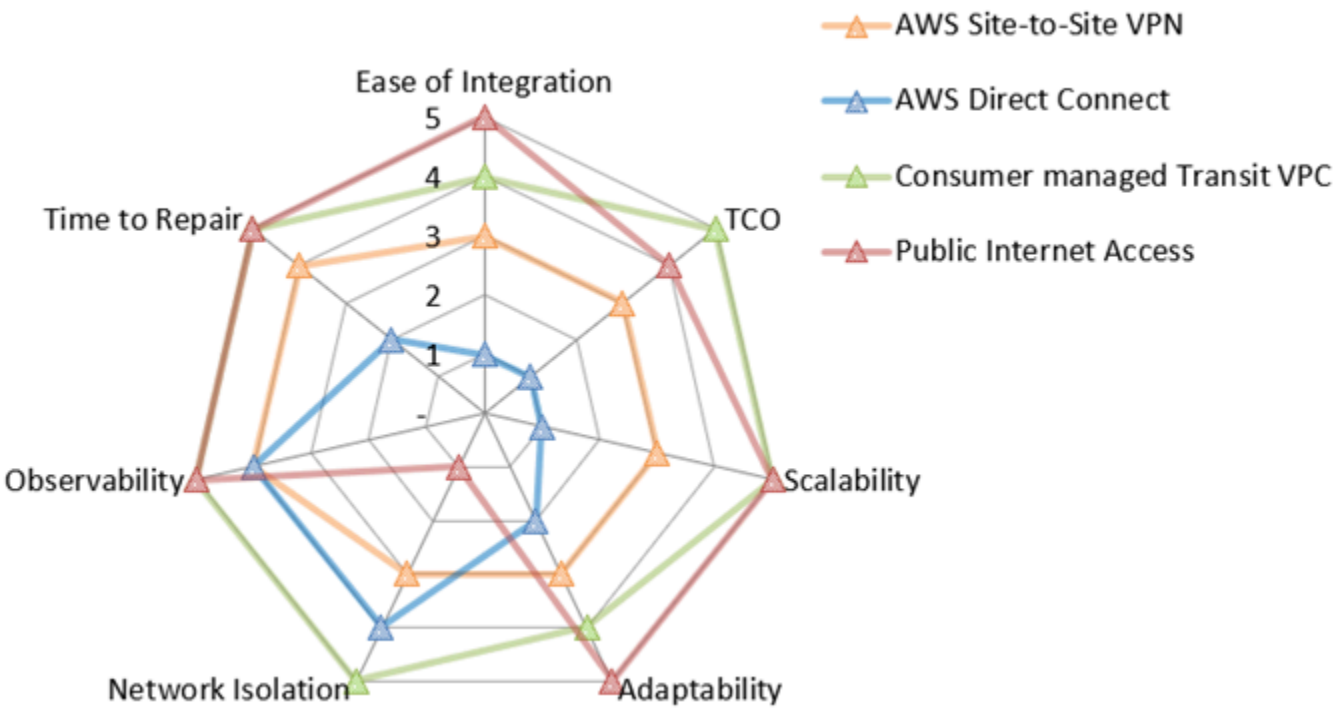
Consumatori SaaS che operano su altri provider di servizi cloud

Questo scenario descrive le soluzioni per i consumatori di altri provider di servizi cloud (CSPs). Questo scenario condivide alcuni punti in comune con le connessioni ai data center locali. In effetti, tutte le opzioni di connettività per gli ambienti locali sono ugualmente valide per i consumatori CSPs, in altri casi AWS Direct Connect è possibile anche una connessione privata con. CSPs La maggior

parte CSPs offre documentazione e supporto su come connettersi alla rete Cloud AWS tramite AWS Site-to-Site VPN o AWS Direct Connect.

Quando scelgono Site-to-Site una VPN, i consumatori possono trarre vantaggio dai gateway gestiti o da risorse simili fornite dai rispettivi CSP. I consumatori non devono necessariamente configurarli da soli, come nello scenario locale. Ciò influenza alcune metriche della Site-to-Site VPN, come il miglioramento dei tempi di riparazione e dell'osservabilità. Questo perché entrambe le estremità della connessione sono ora gestite.

La seguente mappa dei valori di rete riassume il punteggio di ciascuna di queste opzioni per ogni metrica di valutazione. È molto simile alla mappa dei valori di rete per le connessioni locali, sebbene i valori per la Site-to-Site VPN siano diversi. Per ulteriori informazioni sulle metriche di valutazione, consulta questa Metriche di valutazione guida. Nella mappa, un cinque rappresenta il punteggio migliore, ad esempio il TCO più basso, il miglior isolamento della rete o il minor tempo necessario per la riparazione. Per ulteriori informazioni su come leggere questo grafico radar, Mappa dei valori della rete consulta questa guida.



Il grafico radar mostra i seguenti valori.

Metrica di valutazione	AWS Site-to-Site VPN	AWS Direct Connect	VPC di transito gestito dai consumatori	Accesso pubblico a Internet
Facilità di integrazione	3	1	4	5
TCO	3	1	5	4
Scalabilità	3	1	5	5
Adattabilità	3	2	4	5
Isolamento della rete	3	4	5	1
Osservabilità	4	4	5	5
È ora di riparare	4	2	5	5

Supporto per ambienti ibridi

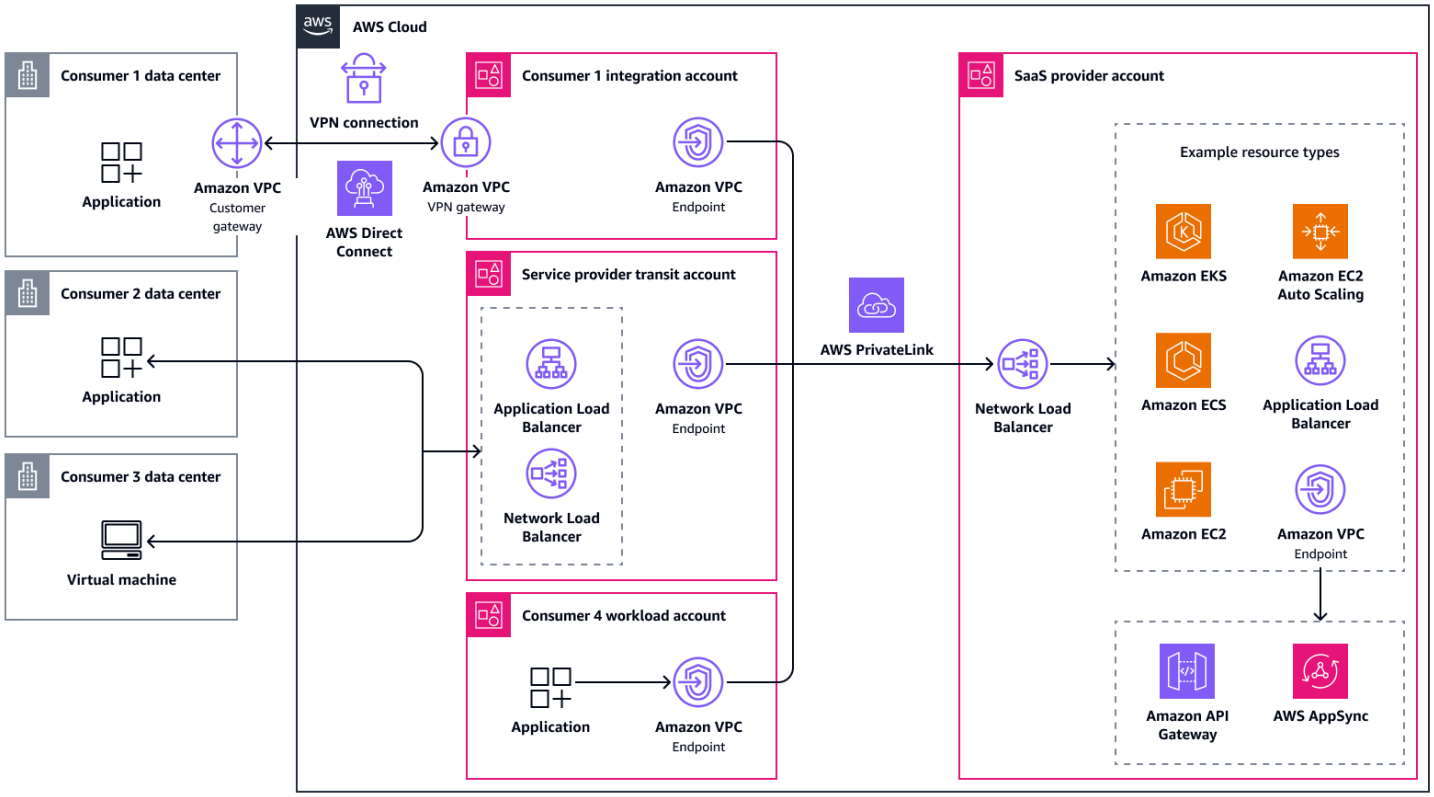
È normale che i consumatori provengano da ambienti diversi, ognuno con i propri vincoli tecnici e di sicurezza. Alcuni clienti possono operare interamente da data center locali che richiedono una connettività sicura via Internet o tramite collegamenti di rete dedicati. Altri potrebbero già eseguire carichi di lavoro interni AWS e aspettarsi percorsi di rete privati a bassa latenza. Un terzo gruppo potrebbe affidarsi ad altri CSPs, in cui la connettività deve collegare diverse reti cloud.

Indipendentemente da ciò, dovresti puntare a un accesso di rete standardizzato alla tua applicazione SaaS per semplificare l'architettura e ridurre la complessità operativa. Due degli approcci presentati in precedenza, l'accesso pubblico a Internet e il transito VPCs, funzionano bene in questi scenari. L'accesso pubblico a Internet offre il percorso di onboarding più veloce con una configurazione minima per i vostri clienti. Transit VPCs offre un accesso più controllato e privato, spesso utilizzato. AWS PrivateLink

Durante la progettazione dell'offerta SaaS, è possibile adottare un unico modello di accesso alla rete o combinare più approcci in un'offerta a più livelli. Ad esempio, potresti offrire un livello di implementazione con accesso pubblico per i clienti che danno priorità alla facilità di connessione e

all'onboarding rapido e potresti offrire un livello di implementazione dell'accesso privato per i clienti che hanno requisiti rigorosi di conformità o controllo della sicurezza. Questi livelli presentano profili di costo, prestazioni e rischio diversi. È anche possibile combinare entrambi gli approcci in un'unica architettura. In tal caso, assicurati di disporre di solide misure di sicurezza in modo che i percorsi pubblici e privati rimangano isolati.

Il diagramma seguente mostra un approccio di accesso ibrido, in cui i consumatori hanno la possibilità di connettersi privatamente dal proprio data center o CSP, pubblicamente o direttamente AWS PrivateLink (se hanno carichi di lavoro all'interno di). Cloud AWS



Scenari avanzati di accesso alla rete per le offerte SaaS in Cloud AWS

Le architetture illustrate nella Scenari di accesso alla rete per le offerte SaaS in Cloud AWS sezione dovrebbero aiutarti a trovare una soluzione per la maggior parte dei casi d'uso. Tuttavia, alcuni scenari presentano requisiti tecnici specifici. Molti non rientrano nell'ambito di questa guida.

In questa sezione vengono descritti i seguenti requisiti e considerazioni tecniche avanzati:

- Comunicazione bidirezionale
- TCP, UDP e protocolli proprietari

Comunicazione bidirezionale

In alcuni casi, le applicazioni richiedono traffico bidirezionale per funzionare come previsto. I casi d'uso più comuni sono i webhook o i servizi di notifica. In genere, è possibile ottenere ciò mediante una WebSocket connessione tra il server e il client. Questa connessione mantiene aperta la sessione TCP e consente a entrambi i partecipanti di inviare traffico tramite la connessione. La maggior parte dei servizi descritti in questa guida supporta nativamente WebSocket, inclusi Network Load Balancer, Application Load Balancer, Amazon API Gateway e AWS AppSync (tramite endpoint privati AWS PrivateLink in tempo reale).

In altri casi, un'applicazione sul lato del provider SaaS potrebbe richiedere l'accesso a risorse sul lato consumer, come un database. Quando ci si connette tramite canali bidirezionali, ad esempio una AWS Site-to-Site VPN connessione, questo non è un problema.

D'altra parte, AWS PrivateLink Elastic Load Balancing supporta solo il traffico unidirezionale. Se utilizzi questi servizi, devi configurare un altro percorso di rete per il traffico che parte dalla tua offerta SaaS. Ad esempio, potrebbe trattarsi di una AWS PrivateLink connessione aggiuntiva che va nella direzione opposta.

TCP, UDP e protocolli proprietari

Molte applicazioni vengono servite tramite HTTP o HTTPS, ma non tutte. Alcuni possono utilizzare altri protocolli Layer 7 oltre al TCP, come Message Queuing Telemetry Support (MQTT). Altri potrebbero persino utilizzare UDP per servire i consumatori. In rari casi, i servizi utilizzano protocolli

proprietari che devono essere trasmessi all'interno di pacchetti (Layer 3). Per questi scenari, è importante capire quali servizi supportano la tua offerta SaaS.

Per i servizi Layer 3, puoi utilizzare Network Load Balancer, che supportano entrambi tutto il traffico TCP AWS PrivateLink e UDP.

Per i servizi Layer 7, Application Load Balancers e Amazon CloudFront supportano HTTP WebSocket, HTTPS e Google Remote Procedure Calls (gRPC). Allo stesso modo, Amazon API Gateway e AWS AppSync entrambi supportano HTTP, HTTPS e WebSocket. Amazon CloudFront è l'unico servizio che attualmente supporta HTTP/3.

Puoi usare Amazon VPC Lattice per connettere applicazioni Layer 7 e risorse Layer 3. Supporta il passthrough HTTP, HTTPS, gRPC, TCP e TLS.

Se l'applicazione è in grado di gestire il traffico solo su Layer 3, è fondamentale utilizzare i servizi AWS di rete di base, come AWS Transit Gateway, AWS Direct Connect AWS Site-to-Site VPN, e il peering VPC. Il traffico dovrebbe quindi essere indirizzato direttamente dal consumatore SaaS al livello di elaborazione dell'offerta SaaS.

Anti-pattern per l'accesso alla rete in Cloud AWS

Un anti-pattern è una soluzione utilizzata frequentemente per un problema ricorrente in cui la soluzione è controproducente, inefficace o meno efficace di un'alternativa. Le opzioni di progettazione menzionate in questa sezione di solito funzionano, ma presentano notevoli svantaggi. Se possibile, dovrebbero essere evitate perché sono disponibili alternative migliori.

Questa sezione illustra i seguenti antipatterni e sfide:

- Mancata corrispondenza della zona di disponibilità con AWS PrivateLink
- AWS Site-to-Site VPN connessioni tra Account AWS

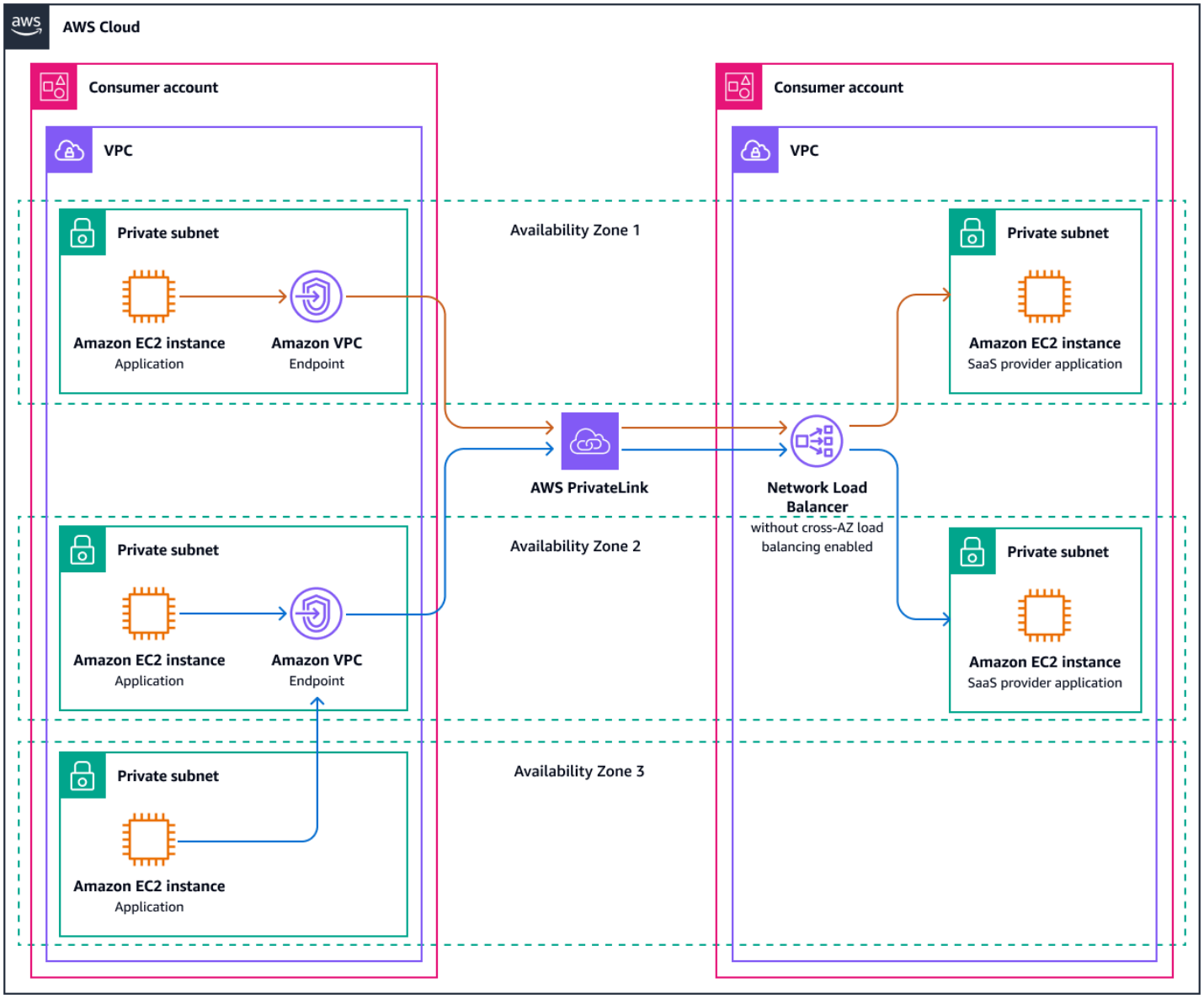
Mancata corrispondenza della zona di disponibilità con AWS PrivateLink

Quando forniscono l'accesso a un'applicazione tramite AWS PrivateLink, i consumatori SaaS possono creare endpoint VPC di interfaccia solo nelle zone di disponibilità in cui è distribuita l'applicazione. Ad esempio, se l'applicazione viene distribuita in `use1-az1` e`use1-az2`, il consumatore non può implementare un endpoint VPC in. `use1-az3` Ti consigliamo di implementare l'offerta SaaS in ogni zona di disponibilità. La maggior parte Regioni AWS ha tre zone di disponibilità, anche se alcune ne hanno di più. Per un elenco completo, consulta Regioni e zone di disponibilità. Considera il numero di zone di disponibilità quando scegli un Regione AWS.

> **Note**
>
> I nomi delle zone di disponibilità sono diversi da quelli delle zone di disponibilità IDs. Per ulteriori informazioni, consulta la sezione Zona di disponibilità IDs per le AWS risorse disponibili.

Se un provider SaaS sceglie di non effettuare l'implementazione in tutte le zone di disponibilità, ci sono alcune conseguenze. Supponiamo che l'offerta SaaS sia implementata in `use1-az1` and`use1-az2`, ma che il consumatore utilizzi tutte e tre le zone di disponibilità, incluse. `use1-az3` Gli endpoint VPC di interfaccia vengono implementati sul lato consumer `use1-az1` e `use1-az2` ora l'applicazione `use1-az3` deve accedere a uno di questi endpoint. Innanzitutto, il traffico deve essere consentito dalle sottoreti nelle zone di disponibilità senza pari verso i rispettivi endpoint VPC.

Il consumatore può decidere di utilizzare il nome AWS PrivateLink DNS regionale, che può essere risolto in entrambi gli endpoint VPC e che distribuisce uniformemente il traffico tra i due. Oppure il consumatore potrebbe scegliere di inviare il traffico direttamente a un endpoint, ad esempio. `use1-az2` Ciò comporta che il 67% del traffico arrivi dal lato provider `use1-az2` e il 33% in entrata. `use1-az1` La figura seguente illustra questo scenario.



Con un numero significativo di consumatori e una distribuzione non uniforme del traffico, un carico di lavoro potrebbe riscontrare problemi di capacità in una zona di disponibilità e risultare insufficiente in un'altra. Per risolvere questo problema, il provider SaaS può decidere di bilanciare il carico in modo uniforme sul proprio lato abilitando il bilanciamento del carico tra zone sul Network Load Balancer. Ciò comporta costi aggiuntivi.

Se il fornitore di servizi corrisponde a una sola zona di disponibilità, tutto il traffico entrerà in un unico endpoint. Ciò crea uno squilibrio ancora maggiore. Di conseguenza, l'offerta SaaS non è più altamente disponibile per il consumatore. Per il consumatore non importa se l'applicazione viene fornita su zone di disponibilità aggiuntive che non utilizza autonomamente. Nel peggiore dei casi, un provider SaaS potrebbe non essere in grado di servire un consumatore che non utilizza nessuna delle stesse zone di disponibilità.

Nel raro caso in cui non sia possibile per il provider SaaS fornire la propria applicazione su tutte le zone di disponibilità, è anche possibile creare una sottorete solo nelle zone di disponibilità mancanti e quindi estendere il servizio a quelle zone di disponibilità vuote. Il bilanciamento del carico tra zone può quindi distribuire il traffico in entrata sugli endpoint effettivi dell'applicazione nelle altre zone di disponibilità.

AWS Site-to-Site VPN connessioni tra Account AWS

Le aziende che migrano da ambienti locali al cloud a volte cercano di sollevare e spostare l'intera rete. Ciò può causare problemi perché esistono differenze significative tra le pratiche di rete locali e quelle di rete cloud. Se questo cambiamento di mentalità non avviene, possono verificarsi cose come le AWS Site-to-Site VPN connessioni da un VPC a un altro VPC. Questo approccio non riesce a sfruttare i servizi di rete appositamente creati in the Cloud AWS, che semplificano la gestione e migliorano le prestazioni. L'adattamento ai design nativi del cloud aiuta a ridurre il sovraccarico operativo e si traduce in una connettività più affidabile e scalabile tra. VPCs

Se state pensando di fornire questa opzione di connettività come provider SaaS, chiedete a voi stessi o al consumatore perché AWS Site-to-Site VPN dovrebbe essere utilizzata. Quindi, procedi a ritroso rispetto a tali requisiti per trovare un'opzione di connettività migliore. La sezione relativa al confronto delle funzionalità dei servizi di questa guida contiene una matrice che è possibile utilizzare per identificare le opzioni. Quindi, puoi consultare le sezioni pertinenti di questa guida per trovare un approccio architettonico adatto al tuo caso d'uso.

Fasi successive

Questa guida descrive vari approcci di accesso alla rete in diversi scenari e descrive i vantaggi e gli svantaggi di ciascuna architettura. È necessario comprendere perché la scelta di un approccio di accesso alla rete non dovrebbe essere una discussione puramente tecnologica. L'allineamento tra business e tecnologia è essenziale. I passaggi e le raccomandazioni seguenti possono aiutarti a valutare e standardizzare la tua strategia di architettura di rete valutando le funzionalità attuali, analizzando le esigenze del mercato e implementando i controlli di governance.

Questa sezione contiene i seguenti argomenti:

- Valutazione dell'architettura e delle funzionalità attuali
- Analisi del mercato e dei clienti
- Allineamento strategico
- Standardizzazione
- Governance
- Ripetizione

Valutazione dell'architettura e delle funzionalità attuali

Esamina l'attuale architettura di rete confrontandola con le fonti di dati pertinenti, ad esempio il quadro di autovalutazione riportato in questa guida, gli attuali requisiti normativi e lo stato attuale del mercato (sia in termini di clienti che di analisi della concorrenza). Ad esempio, prendi in considerazione l'utilizzo del AWS Well-Architected Framework, che si basa su decenni di esperienza nella gestione di sistemi di produzione su larga scala nel. Cloud AWS

Esamina eventuali eccezioni potenziali, decisioni una tantum e decisioni storiche relative ai prodotti. Siate curiosi, sfidatele e non assumete automaticamente la loro validità. Le esigenze dei clienti di anni fa potrebbero non essere più valide. I presupposti sfidanti creano l'opportunità di semplificare e ridurre la complessità dell'architettura.

In termini semplici, documentate le osservazioni in modo che possano essere consultate e comprese dai diversi ruoli dell'organizzazione. Registra dove lo stato attuale differisce dallo stato bersaglio, qual è lo stato obiettivo, l'impatto e quando sono state effettuate le osservazioni. La registrazione di queste informazioni aiuta le organizzazioni a prendere decisioni sulla base di dati aggiornati.

Analisi del mercato e dei clienti

Raccogli informazioni sulle tendenze del mercato. Qual è il modo attualmente preferito dai consumatori per accedere a offerte SaaS come la tua? Riesci ancora a soddisfare i tuoi clienti dove si trovano? Le coorti o il comportamento dei clienti sono cambiati? I vostri dirigenti hanno orientato la rotta verso un nuovo mercato, una geografia con requisiti normativi specifici o un nuovo livello di clienti? La tua azienda o il tuo modello operativo sono cambiati? Ad esempio, state considerando l'idea di etichettare i vostri servizi con etichette bianche? Il vostro piano di crescita prevede la collaborazione con i partner in modo che il vostro servizio sia disponibile per i clienti quando entrano in contatto con tali partner?

Allineamento strategico

Dopo aver compreso le capacità attuali, l'architettura attuale, il mercato e i clienti, convoca una riunione di allineamento strategico. Con i principali stakeholder di prodotto, business e tecnologia, chiedete quali requisiti sono ancora validi e quali nuovi requisiti devono essere presi in considerazione. Trova opportunità per ridurre la complessità eliminando i requisiti che non sono più necessari. Non si tratta di un progetto elaborato da un comitato; il team di ingegneri deve preparare e gestire l'architettura e i dettagli di implementazione effettivi. Tuttavia, questo incontro dovrebbe chiarire perché questo è l'insieme di requisiti che massimizza i vantaggi per i clienti e l'organizzazione.

Standardizzazione

Per attirare clienti, potreste essere tentati di lasciare che ognuno scelga liberamente come connettersi al vostro servizio. Dopotutto, qualsiasi soluzione potrebbe funzionare tecnicamente e potresti anche avere il know-how e le risorse per gestirle e utilizzarle tutte. Fino a un certo punto può funzionare bene, ma man mano che l'azienda cresce, diventa difficile da gestire. Il vostro stack di osservabilità deve supportare le metriche provenienti da più soluzioni e anche i tecnici addetti all'affidabilità del sito devono essere in grado di comprenderle. È necessaria la up-to-date documentazione per ogni approccio di connettività. Le principali modifiche all'applicazione devono essere valutate rispetto a ciascun approccio di accesso offerto. È necessario scrivere e gestire automazioni e infrastrutture come codice (IaC) per ogni approccio di accesso. Il sovraccarico aggiuntivo derivante dalla non standardizzazione dell'accesso al servizio deve essere confrontato con la flessibilità che si desidera offrire ai clienti.

Se avete bisogno di una stella polare che guidi il vostro processo decisionale, vi suggeriamo la standardizzazione. La standardizzazione del modo in cui i clienti interagiscono con i servizi che offri è in genere l'azione più efficace che puoi intraprendere per migliorare molte metriche di successo all'interno dell'organizzazione. La standardizzazione consente ai team di prodotto di comprendere più facilmente la struttura dei costi dei servizi e prendere decisioni sui prodotti basate sui dati. È più facile per i team operativi risolvere i problemi e automatizzare parti del processo di risoluzione dei problemi in un ambiente sviluppato, implementato e gestito secondo standard predefiniti. Può aiutarti a rilevare anomalie, comportamenti imprevisti o azioni di un malintenzionato. La standardizzazione riduce anche il debito tecnico. I team di ingegneri impiegano meno cicli per testare e implementare le modifiche alla produzione. Può anche aumentare la velocità di commercializzazione, migliorare il successo dell'onboarding self-service e ridurre i rischi normativi.

Pertanto, ti suggeriamo di esaminare anche eventuali soluzioni una tantum che potrebbero essere in vigore oggi. Quantifica il numero di cicli operativi che dedichi al supporto dei clienti esistenti. Confronta i risultati con i dati storici e valuta se il tuo approccio attuale è scalabile per gli anni a venire. Ogni volta che è necessario allontanarsi dagli standard, contestate i requisiti alla base di tali richieste. Valuta l'impatto e bilancia i benefici immediati con impegni a lungo termine.

Nei casi in cui la personalizzazione è inevitabile ma è in conflitto con i vostri standard, prendete in considerazione un modello di responsabilità condivisa. In questo modello, i prodotti sono ampiamente protetti dalle modifiche richieste e la personalizzazione avviene in un ambiente minimalista e dedicato. Per un esempio, consulta la sezione. Connessione con un'architettura VPC di transito

Governance

Per la conformità ai requisiti normativi e agli standard interni, la governance è essenziale. Con una governance adeguata, è possibile controllare dove e come applicare gli standard. Inoltre, stabilisci i controlli per rilevare le divergenze dagli standard e informare i proprietari delle risorse sulle azioni correttive necessarie. AWS Organizations, AWS ConfigAWS CloudTrail, e AWS Control Towersono solo alcuni dei tanti Servizi AWS che possono aiutarti a gestire e governare i carichi di lavoro in. Cloud AWS

Ripetizione

Sfruttando gli insegnamenti tratti dai tuoi sforzi iniziali, imposta un processo leggero e ripetibile per rimanere allineato nelle future generazioni. Definisci da quali ruoli hai bisogno degli input, con quale frequenza, quanto devono essere accurati i dati, come verranno condivisi e chi agirà di conseguenza.

Risorse

AWS documentazione

- Integrazione di servizi di terze parti nella Cloud AWS(GuidaAWS prescrittiva)
- Autorizzazione SaaS multi-tenant e controllo dell'accesso alle API (Prescriptive Guidance)AWS
- Gestisci i tenant su più prodotti SaaS su un unico piano di controlloAWS (Prescriptive Guidance)
- AWS Direct Connect Che cos'è? (Direct Connect documentazione)
- Che cos'è AWS PrivateLink? (documentazione Amazon VPC)
- Che cos'è AWS Site-to-Site VPN? (AWS Site-to-Site VPN documentazione)
- Che cos'è AWS Transit Gateway? (documentazione Amazon VPC)
- Cos'è il peering VPC? (documentazione Amazon VPC)

Altre risorse AWS

- Opzioni di connettività Amazon Virtual Private Cloud (AWS white paper)
- AWS re:Invent 2021 - Come scegliere il bilanciamento del carico giusto per i tuoi carichi di lavoro () AWS YouTube
- Che cos'è il SaaS? (AWS sito web)
- AWS Programma SaaS Factory (programma)AWS Partner
- Guida per le architetture multi-tenant su (Libreria di soluzioni) AWSAWS

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	—	12 settembre 2025

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 tuo database Oracle locale a Oracle su un'istanza EC2 in. 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. Esegui 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 del 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 di aggregazione

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 frequentemente 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 utilizzata 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.

implementazione blu/verde

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 Cloud Adoption AWS 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 viene comunemente descritto come una pipeline. CI/CD può 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, è possibile AWS CloudFormation utilizzarlo per rilevare deviazioni nelle risorse di sistema oppure AWS Control Tower per rilevare cambiamenti nella 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 CI/CD pipeline, l'ambiente di produzione è l'ultimo ambiente di distribuzione.

- ambienti superiori: tutti gli ambienti a cui possono accedere utenti diversi dal team di sviluppo principale. Si può trattare di un ambiente di produzione, ambienti di preproduzione e ambienti per i test di accettazione da parte degli utenti.

epica

Nelle metodologie agili, categorie funzionali che aiutano a organizzare e dare priorità al lavoro. Le epiche forniscono una descrizione di alto livello dei requisiti e delle attività di implementazione. Ad esempio, le epopee della sicurezza AWS CAF includono la gestione delle identità e degli accessi, i controlli investigativi, la sicurezza dell'infrastruttura, la protezione dei dati e la risposta agli incidenti. Per ulteriori informazioni sulle epiche, consulta la strategia di migrazione AWS , consulta la guida all'implementazione del programma.

ERP

Vedi pianificazione delle risorse aziendali.

analisi esplorativa dei dati (EDA)

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

F

tabella dei fatti

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

fallire velocemente

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

limite di isolamento dei guasti

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

ramo di funzionalità

Vedi filiale.

caratteristiche

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

importanza delle caratteristiche

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

IA generativa

Un sottoinsieme di modelli di intelligenza artificiale che sono stati addestrati su grandi quantità di dati e che possono utilizzare un semplice messaggio 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 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 CSPM 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 blocco

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

Vedi 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 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 AWS con Application Migration Service.

Valutazione del portfolio di migrazione (MPA)

Uno strumento online che fornisce informazioni per la convalida del business case per la migrazione 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.

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).

RAG

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.

Region

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 alle interruzioni o di ripristinarle. 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 obiettivo del punto di ripristino.

VERSO

Vedi 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 Console di gestione AWS 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 Gestione dei segreti AWS, 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 Amazon EC2 o la rotazione delle credenziali.

Crittografia lato server

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

Policy di controllo dei servizi (SCP)

Una 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

tag

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.