# Approcci al backup e al ripristino tramite AWS

*Giugno 2016*

# Note

Il presente documento è fornito a solo scopo informativo. In esso sono illustrate le attuali offerte di prodotti e le prassi di AWS alla data di pubblicazione del documento, offerte che sono soggette a modifica senza preavviso.
È responsabilità dei clienti effettuare una propria valutazione indipendente delle informazioni contenute nel presente documento e dell'uso dei prodotti o dei servizi di AWS, ciascuno dei quali viene fornito "così com'è", senza garanzie di alcun tipo, né esplicite né implicite. Il presente documento non dà origine a garanzie, rappresentazioni, impegni contrattuali, condizioni o assicurazioni da parte di AWS, delle sue società affiliate, dei suoi fornitori o dei licenzianti.
Le responsabilità di AWS nei confronti dei propri clienti sono definite dai contratti AWS e il presente documento non costituisce parte né modifica qualsivoglia contratto tra AWS e i suoi clienti.

# Indice

# Sintesi

Questo documento è rivolto a progettisti di soluzioni aziendali, progettisti di backup e amministratori IT responsabili della protezione dei dati negli ambienti IT aziendali. Nel documento si illustrano le architetture e i carichi di lavoro di produzione che possono essere implementati tramite AWS per potenziare o sostituire una soluzione di backup e ripristino. Questi approcci offrono costi ridotti, maggiore scalabilità e maggiore durabilità per soddisfare i requisiti di conformità, i tempi di recupero e i punti di recupero previsti.

# Introduzione

In un contesto in cui i dati aumentano velocemente, la loro protezione è sempre più impegnativa. Domande sulla durabilità e sulla scalabilità dei metodi di backup sono all'ordine del giorno, ad esempio: in che modo il cloud contribuisce a soddisfare le esigenze di backup e di archiviazione?

Questo documento copre una serie di architetture di backup (applicazioni native su cloud, ambienti ibridi e locali) e di servizi AWS correlati che possono essere utilizzati per creare soluzioni scalabili e affidabili per la protezione dei dati.

# Perché utilizzare AWS come piattaforma per la protezione dei dati?

Amazon Web Services (AWS) è una piattaforma di cloud computing sicura, ad alte prestazioni, flessibile, a costi ridotti e di facile utilizzo. AWS si occupa delle gravose attività ordinarie non distintive dell'azienda e fornisce strumenti e risorse da poter utilizzare per creare soluzioni di backup e ripristino scalabili.

L'utilizzo di AWS come parte della strategia di protezione dei dati apporta numerosi vantaggi:

- **Durabilità:** Amazon Simple Storage Service (Amazon S3) e Amazon Glacier sono progettati per garantire il 99,999999999% (11 nove) di durabilità degli oggetti in essi archiviati. Entrambe le piattaforme offrono ubicazioni affidabili per i dati di backup.

- **Sicurezza:** AWS fornisce una serie di opzioni per il controllo degli accessi e la crittografia dei dati in transito e inattivi.
- **Infrastruttura globale:** grazie alla disponibilità dei servizi AWS a livello mondiale, puoi eseguire il backup e l'archiviazione dei dati nella regione che soddisfa i tuoi requisiti di conformità.
- **Conformità:** l'infrastruttura AWS dispone di certificazioni di conformità a standard come Service Organization Controls (SOC), Statement on Standards for Attestation Engagements (SSAE) 16, International Organization for Standardization (ISO) 27001, Payment Card Industry Data Security Standard (PCI DSS), Health Insurance Portability and Accountability Act (HIPPA), SEC[1] e Federal Risk and Authorization Management Program (FedRAMP), consentendoti di inserire facilmente la soluzione di backup nel tuo regime di conformità esistente.
- **Scalabilità:** con AWS, non devi preoccuparti della capacità. Puoi aumentare o ridurre il consumo a seconda delle esigenze, senza costi amministrativi.
- **Costo totale di proprietà ridotto:** la portata delle operazioni AWS riduce i costi dei servizi e contribuisce a ridurre il costo totale di proprietà dello storage. AWS offre questi risparmi sui costi ai clienti sotto forma di riduzione del prezzo.
- **Prezzo in base al consumo:** acquista i servizi AWS in base alle esigenze e solo per il periodo che prevedi di utilizzarli. Il prezzo di AWS non prevede costi anticipati, penali in caso di recesso o contratti a lungo termine.

# Servizi di storage AWS per la protezione dei dati

Amazon S3 e Amazon Glacier sono i servizi ideali per il backup e l'archiviazione. Rappresentano entrambi piattaforme di storage durevoli e a costi contenuti. Offrono capacità illimitata e non richiedono la gestione dei volumi o dei supporti con la crescita dei set di dati di backup. Grazie al modello di pagamento in base all'utilizzo e al basso costo per GB/mese, questi servizi rappresentano un'ottima soluzione per i casi d'uso di protezione dei dati.

[1] https://aws.amazon.com/about-aws/whats-new/2015/09/amazon-glacier-receives-third-party-compliance-assessment-for-sec-rule-17a-4f-from-cohasset-associates-inc/

## Amazon S3

Amazon S3 fornisce uno storage di oggetti sicuro e scalabile.

Puoi utilizzare Amazon S3 per archiviare e recuperare qualsiasi quantità di dati, in qualunque momento e da ogni luogo tramite il Web. Amazon S3 archivia i dati sotto forma di oggetti all'interno di risorse chiamate *bucket*. AWS Storage Gateway e numerose soluzioni di backup di terze parti possono gestire gli oggetti di Amazon S3 per conto tuo. Puoi archiviare tutti gli oggetti che desideri in un bucket e scrivere, leggere e cancellare gli oggetti nel tuo bucket. I singoli oggetti possono avere una dimensione massima di 5 TB.

Amazon S3 offre diverse classi di storage concepite per diversi casi d'uso, tra cui:

- **Amazon S3 Standard** per lo storage generico di dati a cui accedi frequentemente.
- **Amazon S3 Standard - Infrequent Access** per i dati di lunga durata a cui accedi meno frequentemente.
- **Amazon Glacier** per l'archiviazione a lungo termine.

Amazon S3 offre anche la possibilità di configurare policy per gestire i dati durante il loro ciclo di vita. Dopo l'impostazione di una policy, i dati verranno spostati nella classe di storage appropriata senza alcuna modifica all'applicazione. Per maggiori informazioni, consulta Classi di storage S3.

## Amazon Glacier

Amazon Glacier è un servizio di storage nel cloud a costi estremamente bassi che fornisce uno storage sicuro e durevole per l'archiviazione e il backup online dei dati. Per mantenere bassi i costi, Amazon Glacier è ottimizzato per i dati a cui si accede frequentemente e per i quali sono accettabili tempi di recupero di diverse ore. Con Amazon Glacier, puoi archiviare in modo affidabile piccole o grandi quantità di dati per un costo di appena $ 0,007 per gigabyte al mese, un risparmio significativo rispetto alle soluzioni locali. Amazon Glacier è particolarmente indicato per lo storage dei dati di backup con requisiti di conservazione lunghi o indefiniti e per l'archiviazione dei dati a lungo termine. Per maggiori informazioni, consulta Amazon Glacier.

### AWS Storage Gateway

AWS Storage Gateway collega un'appliance software locale allo storage basato sul cloud per fornire un'integrazione perfetta ed estremamente sicura tra l'ambiente IT locale e l'infrastruttura di storage AWS. Per maggiori informazioni, consulta AWS Storage Gateway.

### Servizi di trasferimento AWS

Per trasferire i tuoi dati, oltre ai gateway e ai connettori di terze parti, puoi utilizzare opzioni AWS come AWS Direct Connect, AWS Snowball, AWS Storage Gateway e Amazon S3 Transfer Acceleration. Per maggiori informazioni, consulta Migrazione dei dati nel cloud.

## Progettazione di una soluzione di backup e ripristino

Quando elabori una strategia completa per il backup e il ripristino dei dati, devi innanzitutto individuare le situazioni di guasto o di emergenza che potrebbero verificarsi e il loro potenziale impatto sull'attività. In alcuni settori, è necessario considerare i requisiti normativi per la sicurezza dei dati, la privacy e la conservazione dei record.

Dovresti attuare processi di backup che offrano il livello giusto di granularità per rispettare i tempi di recupero e i punti di recupero previsti dell'azienda, tra cui:

- Ripristino a livello di file
- Ripristino a livello di volume
- Ripristino a livello di applicazione (ad esempio, database)
- Ripristino a livello di immagine

Nelle sezioni seguenti vengono descritti gli approcci al backup, al ripristino e all'archiviazione in base all'organizzazione dell'infrastruttura. L'infrastruttura IT può essere suddivisa in tre grandi categorie: nativa sul cloud, locale e ibrida.

# Infrastruttura nativa sul cloud

In questo scenario viene descritto un ambiente di carichi di lavoro interamente collocato su AWS. Come illustrato nella figura seguente, lo scenario include server Web, server di applicazioni, server di monitoraggio, database e Active Directory.

Se esegui tutti i servizi da AWS, puoi utilizzare numerose caratteristiche integrate per soddisfare le esigenze di protezione e ripristino dei dati.

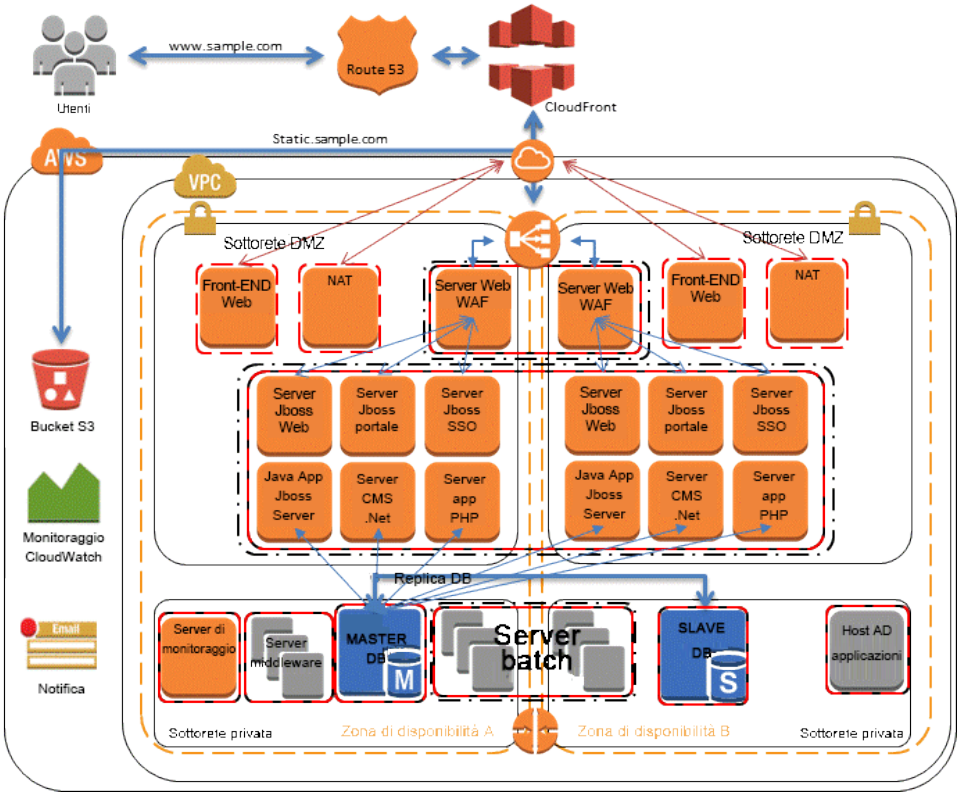


**Figura 1: scenario di un'infrastruttura nativa su AWS**

## Protezione basata su snapshot EBS

Quando i servizi vengono eseguiti su Amazon Elastic Compute Cloud[2] (Amazon EC2), le istanze di elaborazione possono utilizzare i volumi Amazon Elastic Block Store (Amazon EBS) per archiviare i dati principali e accedervi. Puoi utilizzare questo storage in blocchi per i dati strutturati, come i database, o per i dati non strutturati, come i file in un file system nel volume.

Amazon EBS offre la possibilità di creare snapshot (backup) di qualsiasi volume Amazon EBS. Una copia del volume viene posizionata in Amazon S3, dove viene archiviata in modo ridondante in più zone di disponibilità. La prima snapshot è una copia completa del volume mentre le snapshot continue archiviano solo le modifiche incrementali a livello di blocco.

Questo è un modo rapido e affidabile di ripristinare tutti i dati del volume. Se necessiti solo di un ripristino parziale, puoi collegare il volume all'istanza in esecuzione sotto il nome di un dispositivo diverso, montarlo e utilizzare i comandi di copia del sistema operativo per copiare i dati dal volume di backup al volume di produzione.

Le snapshot Amazon EBS possono anche essere copiate da una regione AWS all'altra tramite la funzionalità di copia delle snapshot di Amazon EBS disponibile nella console o dalla riga di comando, come illustrato nella Guida per l'utente di Amazon Elastic Cloud Compute.[3] Puoi utilizzare questa caratteristica per archiviare il backup in un'altra regione senza dover gestire la tecnologia di replica sottostante.

---

[2] http://aws.amazon.com/ec2/

[3] http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ebs-copy-snapshot.html

## Creazione di snapshot EBS

Con la creazione di una snapshot, proteggi i dati direttamente sullo storage durevole basato su disco. Per creare la snapshot Amazon EBS, puoi utilizzare la console di gestione AWS, l'interfaccia a riga di comando (CLI) oppure le API.

Nella console Amazon EC2, nella pagina **Elastic Block Store Volumes**, scegli **Create Snapshot** dal menu **Actions**. Nella finestra di dialogo **Create Snapshot**, scegli **Create** per creare una snapshot che verrà archiviata in Amazon S3.

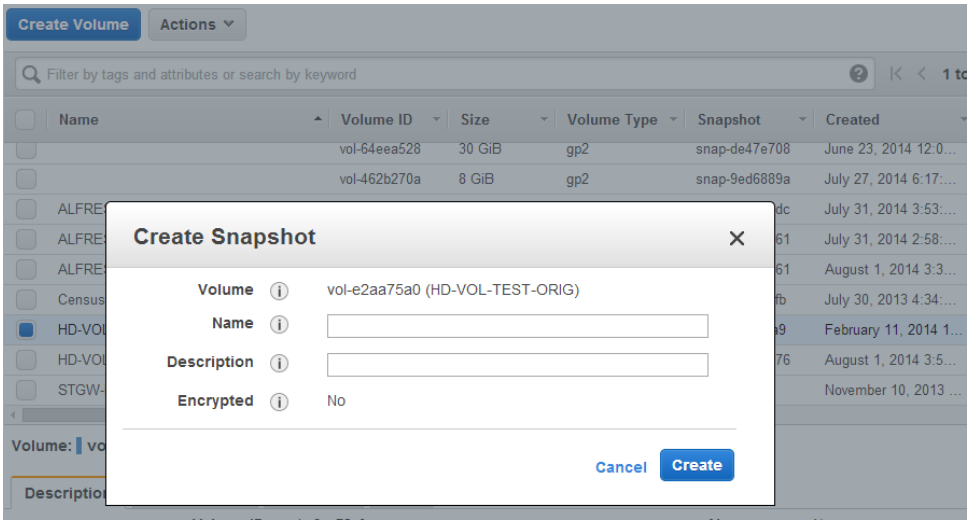


**Figura 2: utilizzo della console EC2 per creare una snapshot**

Per creare la snapshot tramite il comando della CLI, esegui il comando sottostante:

```
➢ aws ec2 create-snapshot
```

Puoi pianificare ed eseguire periodicamente i comandi `aws ec2 create-snapshot` per eseguire il backup dei dati EBS. Grazie al suo prezzo economico, Amazon S3 ti permette di conservare molte generazioni di dati. Inoltre, poiché le snapshot sono basate su blocchi, utilizzi solo lo spazio per i dati modificati dopo la creazione della prima snapshot.

## Ripristino da una snapshot EBS

Per ripristinare i dati da una snapshot, puoi utilizzare la console di gestione AWS, la CLI o le API per creare un volume da una snapshot esistente.

Per ripristinare, ad esempio, un volume a un backup point-in-time precedente, segui questi passaggi:

1. Utilizza il comando sottostante per creare un volume dalla snapshot di backup:

```
➢ aws ec2 create-volume --region us-west-1b --snapshot-
  id mysnapshot-id
```

2. Nell'istanza Amazon EC2, smonta il volume esistente.

   In Linux, utilizza `umount`. In Windows, utilizza il gestore logico dei volumi (LVM, Logical Volume Manager).

3. Utilizza il comando sottostante per scollegare il volume esistente dall'istanza:

```
➢ aws ec2 detach-volume --volume-id oldvolume-id –
  instance-id myec2instance-id
```

4. Utilizza il comando sottostante per collegare il volume creato dalla snapshot:

```
➢ aws ec2 attach-volume --volume-id newvolume-id
  --instance-id myec2instance-id --device /dev/sdf
```

5. Monta nuovamente il volume sull'istanza in esecuzione.

## Creazione di backup coerenti o a caldo

Quando esegui un backup, è preferibile che il sistema sia in uno stato in cui non esegue alcun I/O. La situazione ideale è quella in cui la macchina non accetta traffico, ma questo è un caso sempre più raro in un contesto in cui operazioni IT 24 ore su 24, 7 giorni su 7 diventano la norma.

Questo è il motivo per cui, per eseguire un backup pulito, devi disattivare il file system o il database. La procedura di disattivazione dipende dal database o dal file system.

La procedura per un database è la seguente:

- Se possibile, imposta il database nella modalità di backup a caldo.
- Esegui i comandi della snapshot Amazon EBS.
- Disattiva la modalità di backup a caldo del database o, se utilizzi una replica di lettura, chiudi l'istanza di tale replica.

La procedura per un file system è simile, ma dipende dalle capacità del sistema operativo o del file system. XFS, ad esempio, è un file system che può scaricare i dati per un backup coerente. Per maggiori informazioni, consulta xfs_freeze.[4]

Se il tuo file system non supporta la capacità di blocco, dovresti smontarlo, rilasciare il comando della snapshot e rimontare il file system. In alternativa puoi semplificare questa procedura utilizzando un gestore logico dei volumi che supporti il blocco di I/O.

Poiché l'elaborazione della snapshot continua in background e la creazione della snapshot viene eseguita rapidamente e acquisisce un punto nel tempo, i volumi di cui esegui il backup devono solo essere smontati per qualche secondo. L'intervallo di backup è il più ridotto possibile, quindi il tempo di interruzione è prevedibile e può essere pianificato.

[4] https://access.redhat.com/documentation/en-US/Red_Hat_Enterprise_Linux/6/html/Storage_Administration_Guide/xfsfreeze.html

## Esecuzione di backup multivolume

In alcuni casi, puoi eseguire lo striping dei dati su più volumi Amazon EBS utilizzando un gestore logico di volumi per aumentare il throughput potenziale. Quando utilizzi un gestore logico di volumi (ad esempio, mdadm o LVM), è importante eseguire il backup dal layer del gestore dei volumi anziché dai volumi EBS sottostanti. Ciò assicura che tutti i metadati siano allineati e i volumi dei sottocomponenti siano coerenti.

Esistono diversi modi per ottenere questo risultato. Puoi utilizzare, ad esempio, lo script creato da alestic.com.[5] I buffer di memoria dovrebbero essere scaricati sul disco, l'I/O del file system sul disco dovrebbe essere interrotto e dovrebbe essere avviata una snapshot simultaneamente per tutti i volumi che compongono il set RAID. Dopo l'avvio di una snapshot per i volumi (generalmente uno o due secondi), il file system può proseguire le sue operazioni. Le snapshot dovrebbero essere contrassegnate in modo tale da poter essere gestite collettivamente durante un ripristino.

Puoi eseguire questi backup anche dal gestore logico di volumi o a livello di file system. In questi casi, l'utilizzo di un agente di backup tradizionale abilita il backup dei dati nella rete. Una serie di soluzioni di backup basato su agente è disponibile su Internet e su AWS Marketplace.[6] Tieni presente che i software di backup basato su agente richiedono un nome del server e un indirizzo IP coerenti. Di conseguenza, la soluzione migliore per garantire affidabilità è utilizzare questi strumenti con istanze distribuite in un Virtual Private Cloud (VPC)[7] Amazon.

Un approccio alternativo è creare una replica dei volumi del sistema principale su un unico volume di grandi dimensioni. Ciò semplifica il processo di backup in quanto viene eseguito il backup di un unico grande volume e l'attività non avviene sul sistema principale. Tuttavia, dovresti stabilire innanzitutto se il singolo volume può garantire prestazioni sufficienti durante il backup e se la dimensione massima del volume è adeguata per l'applicazione.

---

[5] https://github.com/alestic/ec2-consistent-snapshot

[6] https://aws.amazon.com/marketplace/

[7] http://aws.amazon.com/vpc/

# Approcci al backup dei database

AWS offre molte opzioni per i database. Puoi eseguire il tuo database su un'istanza EC2 o utilizzare una delle opzioni di database con servizi gestiti fornite da Amazon Relational Database Service[8] (Amazon RDS). Se esegui il tuo database su un'istanza EC2, puoi eseguire il backup dei dati su file utilizzando strumenti nativi (ad esempio, MySQL[9], Oracle[10], MSSQL[11], PostgreSQL[12]) oppure creare una snapshot dei volumi che contengono i dati utilizzando uno dei metodi illustrati in "Protezione basata su snapshot EBS".

## Utilizzo di backup basati sulla replica del database

Per i database creati sui set RAID dei volumi Amazon EBS, puoi evitare lo sforzo di eseguire i backup sul database principale creando una replica di lettura del database. Si tratta di una copia aggiornata del database, eseguita su un'istanza Amazon EC2 separata. L'istanza della replica del database può essere creata utilizzando più dischi simili all'origine oppure i dati possono essere consolidati in un unico volume EBS. A questo punto, per creare la snapshot dei volumi EBS, puoi utilizzare una delle procedure illustrate in "Protezione basata su snapshot EBS". Questo approccio è spesso utilizzato per i database di grandi dimensioni che devono essere operativi 24 ore su 24, 7 giorni su 7. In questo caso, il backup richiede troppo tempo e il database di produzione non può essere interrotto per periodi così lunghi.

## Utilizzo di Amazon RDS per i backup

Amazon RDS offre caratteristiche per l'automazione dei backup dei database. Amazon RDS crea una snapshot dei volumi di storage dell'istanza database, eseguendo il backup dell'intera istanza database anziché dei singoli database.

Amazon RDS fornisce due diversi metodi di backup e ripristino delle istanze database:

---

[8] https://aws.amazon.com/rds/

[9] http://dev.mysql.com/doc/refman/5.7/en/backup-and-recovery.html

[10] http://docs.oracle.com/cd/E11882_01/backup.112/e10642/rcmbckba.htm#BRADV8003

[11] http://msdn.microsoft.com/en-us/library/ms187510.aspx

[12] http://www.postgresql.org/docs/9.3/static/backup.html

- **I backup automatizzati** consentono il ripristino point-in-time dell'istanza database. I backup automatizzati vengono attivati per impostazione predefinita quando crei una nuova istanza database. Amazon RDS esegue un backup giornaliero completo dei dati in un intervallo di tempo da te stabilito al momento della creazione dell'istanza database. Puoi configurare un periodo di conservazione fino a 35 giorni per il backup automatizzato. Amazon RDS utilizza questi backup di dati periodici insieme ai registri delle transizioni per consentirti di ripristinare l'istanza database a qualsiasi secondo durante il periodo di conservazione, fino al `LatestRestorableTime` (generalmente, gli ultimi cinque minuti). Per trovare il punto di ripristino più recente per le istanze database, puoi utilizzare la chiamata API `DescribeDBInstances` o cercare il database nella scheda **Description** nella console di gestione AWS.

  Quando avvii un ripristino point-in-time, i registri delle transazioni vengono applicati al backup giornaliero più appropriato per recuperare l'istanza database nel punto richiesto.

- Le **snapshot DB** sono backup avviati dall'utente che ti permettono di eseguire il backup dell'istanza database a uno stato noto con la frequenza che desideri e ripristinare successivamente quello stato in qualsiasi momento. Per creare snapshot DB, puoi utilizzare la console di gestione AWS o la chiamata API `CreateDBSnapshot`. Queste snapshot possono essere conservate per un periodo di tempo illimitato. Vengono mantenute finché non sarai tu a cancellarle esplicitamente tramite la console o la chiamata API `DeleteDBSnapshot`.

Quando ripristini un database in un punto nel tempo o da una snapshot DB, verrà creata una nuova istanza di database con un nuovo endpoint. In questo modo puoi creare più istanze di database da una specifica snapshot DB o da un punto nel tempo.

Per cancellare le vecchie istanze di database, puoi utilizzare la console di gestione AWS o una chiamata `DeleteDBInstance`.

## Utilizzo di AMI per il backup di istanze EC2

AWS archivia immagini di sistema nelle cosiddette Amazon Machine Image (AMI). Queste immagini sono formate dal modello del volume radice, necessario per l'avvio di un'istanza. Per eseguire il backup del volume radice come AMI, puoi utilizzare la console di gestione AWS o il comando CLI `aws ec2 create-image`.

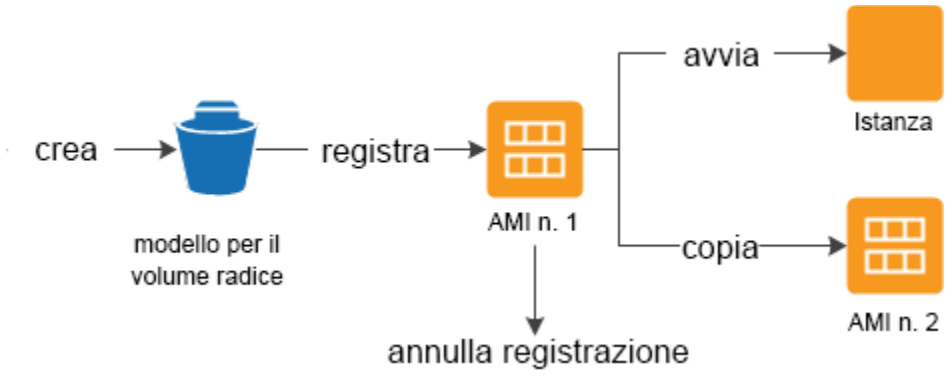


**Figura 3: utilizzo di un'AMI per il backup e l'avvio di un'istanza**

Quando registri un'AMI, questa viene archiviata nel tuo account tramite snapshot Amazon EBS. Queste snapshot si trovano in Amazon S3 e sono estremamente durevoli.

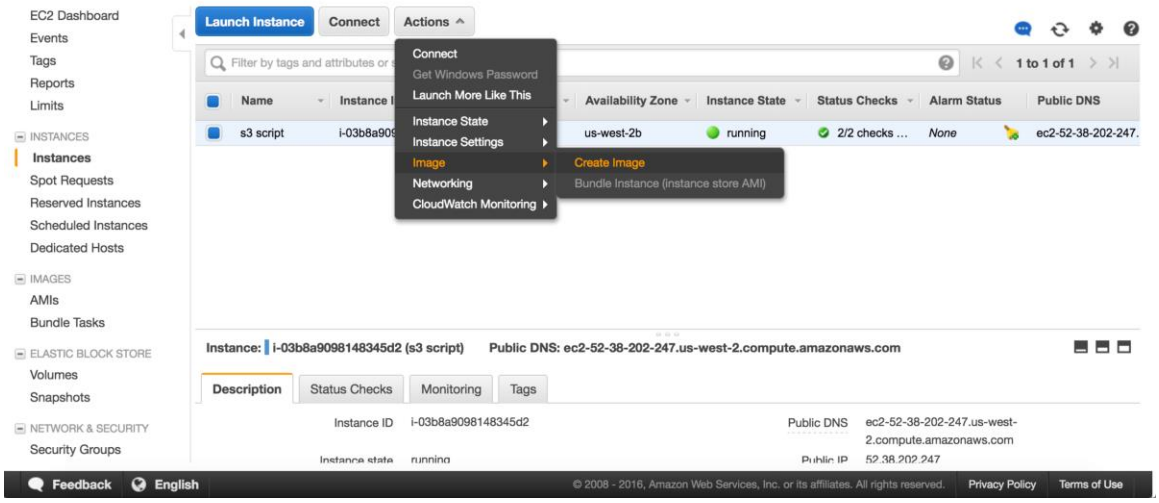


**Figura 4: utilizzo della console EC2 per creare una Machine Image**

Dopo aver creato un'AMI dell'istanza Amazon EC2, puoi utilizzare l'AMI per ricreare l'istanza o avviare più copie dell'istanza stessa. Inoltre, puoi copiare le AMI da una regione all'altra per la migrazione delle applicazioni o per il disaster recovery.

# Dall'infrastruttura locale all'infrastruttura AWS

In questo scenario viene descritto un ambiente di carichi di lavoro senza componenti nel cloud. Tutte le risorse, tra cui server Web, server di applicazioni, server di monitoraggio, database, Active Directory e altro ancora, sono ospitate nel data center del cliente o tramite colocation.

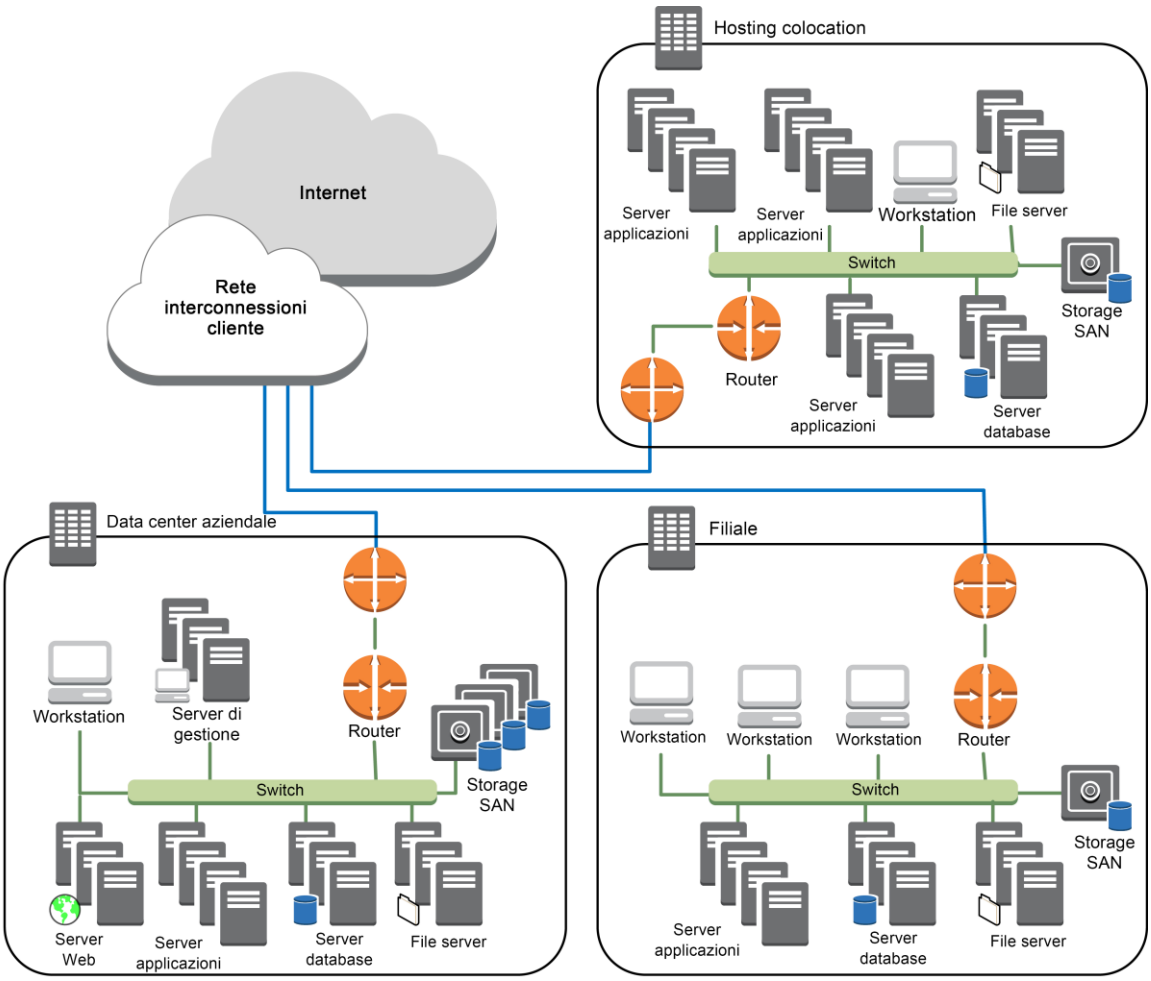


**Figura 5: ambiente locale**

In questo scenario, utilizzando i servizi di storage AWS, puoi concentrarti sulle attività di backup e archiviazione. Per completare l'attività di backup, non devi preoccuparti della scalabilità dello storage o della capacità dell'infrastruttura.

Amazon S3 e Amazon Glacier sono basati su API in modo nativo e disponibili tramite Internet. Ciò permette ai fornitori di software di backup di integrare direttamente le proprie applicazioni nelle soluzioni di storage AWS, come illustrato nella figura sottostante.

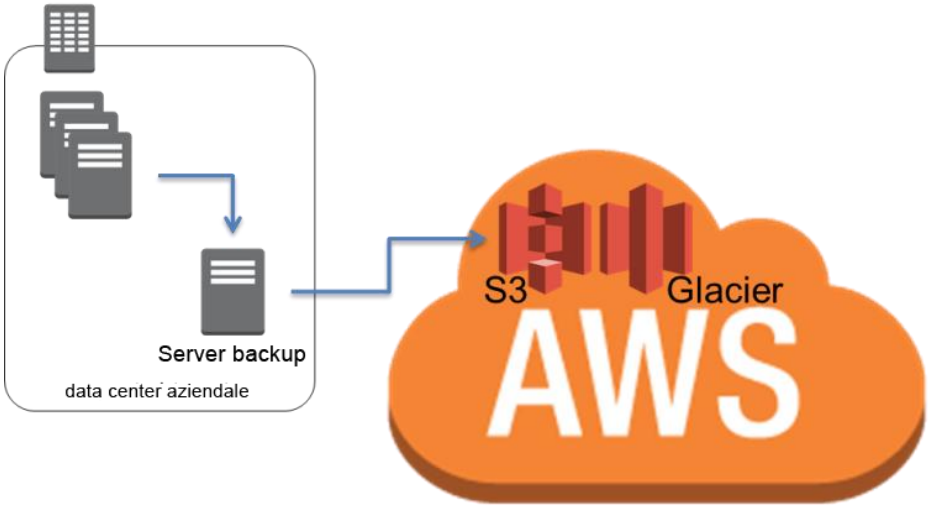


**Figura 6: connettore di backup verso Amazon S3 o Amazon Glacier**

In questo scenario, il software di backup e archiviazione si interfaccia direttamente con AWS tramite le API. Il software di backup, che riconosce AWS, esegue il backup dei dati dai server locali direttamente ad Amazon S3 o Amazon Glacier.

Se i tuoi software di backup esistenti non supportano il cloud AWS in modo nativo, puoi utilizzare i prodotti AWS Storage Gateway. AWS Storage Gateway[13] è un'appliance virtuale che fornisce un'integrazione perfetta e sicura tra il tuo data center e l'infrastruttura di storage AWS. Il servizio ti permette di archiviare i dati nel cloud AWS in tutta sicurezza, per uno storage scalabile e a costi contenuti. Storage Gateway supporta i protocolli di storage standard del settore compatibili con le tue applicazioni esistenti e archivia in modo sicuro tutti i dati crittografati in Amazon S3 o Amazon Glacier.

[13] http://aws.amazon.com/storagegateway/

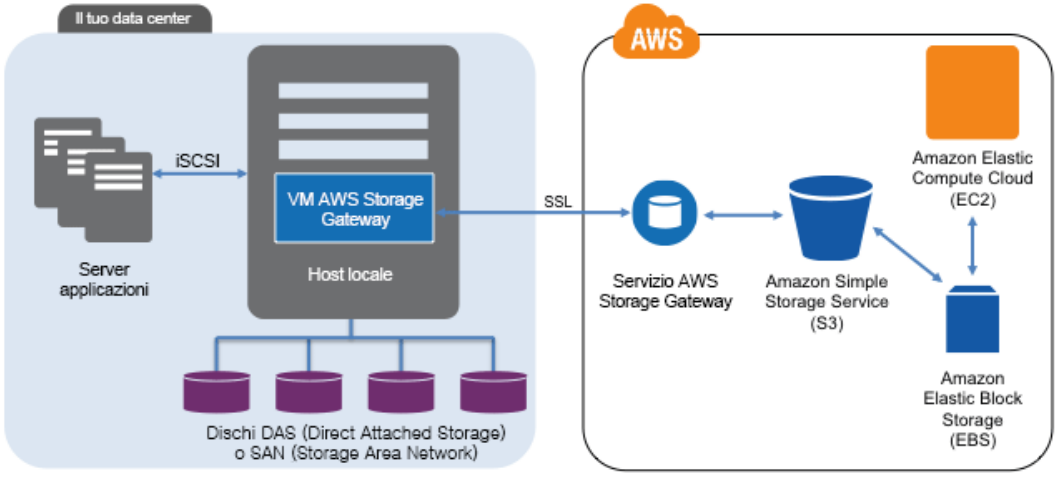


**Figura 7: connessione dell'ambiente locale allo storage AWS**

AWS Storage Gateway supporta le configurazioni seguenti:

- **Gateway di volumi:** i gateway di volumi forniscono volumi di storage basati sul cloud che puoi montare come dispositivi iSCSI (Internet Small Computer System Interface) dai server di applicazioni locali. Il gateway supporta le configurazioni di volumi seguenti:
  - **Volumi nella cache del gateway:** puoi archiviare i dati principali in Amazon S3 e conservare a livello locale i dati a cui accedi frequentemente. I volumi nella cache del gateway offrono un notevole risparmio sui costi associati allo storage principale, riducono al minimo la necessità di dimensionare lo storage locale e conservano un accesso a bassa latenza ai dati utilizzati frequentemente.
  - **Volumi nella storage del gateway:** se necessiti di un accesso a bassa latenza a tutto il set di dati, puoi configurare il gateway di dati locale in modo tale che archivi i dati principali a livello locale ed esegua il backup point-in-time asincrono delle snapshot di questi dati su Amazon S3. I volumi nello storage del gateway forniscono backup fuori sede durevoli ed economici che puoi ripristinare a livello locale o da Amazon EC2.

- **Libreria di nastri virtuali del gateway (Gateway-VTL):** con l'opzione Gateway-VTL, disponi di una raccolta illimitata di nastri virtuali. Ciascun nastro virtuale può essere archiviato in una libreria di nastri virtuali supportata da Amazon S3 o in uno scaffale di nastri virtuali supportato da Amazon Glacier. La libreria di nastri virtuali dispone di un'interfaccia iSCSI standard del settore che fornisce all'applicazione di backup accesso online ai nastri virtuali. Quando non necessiti più di accesso immediato o frequente ai dati contenuti in un nastro virtuale, puoi utilizzare la tua applicazione di backup per spostarli dalla libreria di nastri virtuali allo scaffale di nastri virtuali per ridurre ulteriormente i costi associati allo storage.

Questi gateway operano come dispositivi plug-and-play e rappresentano dispositivi iSCSI standard che possono essere integrati nel framework di backup o archiviazione. Puoi utilizzare i dispositivi disco iSCSI come pool di storage per il software di backup oppure l'opzione Gateway-VTL per eseguire l'offload del backup o dell'archiviazione basata su nastro direttamente su Amazon S3 o Amazon Glacier.

Con questo metodo, il backup e gli archivi vengono automaticamente collocati fuori sede (per scopi di conformità) e archiviati su supporti durevoli, eliminando la complessità e i rischi per la sicurezza associati alla gestione dei nastri fuori sede.

# Ambienti ibridi

Le due distribuzioni di infrastrutture illustrate finora, nativa su cloud e locale, possono essere combinate in uno scenario ibrido in cui l'ambiente dei carichi di lavoro presenta componenti dell'infrastruttura locale e dell'infrastruttura AWS. Le risorse, tra cui server Web, server di applicazioni, server di monitoraggio, database, Active Directory e altro ancora, sono ospitate nel data center del cliente o in AWS. Le applicazioni in esecuzione nel cloud AWS sono collegate alle applicazioni in esecuzione a livello locale.

Questo scenario è sempre più comune per i carichi di lavoro aziendali. Numerose aziende dispongono di data center propri e utilizzano AWS per aumentare la capacità. Questi data center dei clienti sono spesso collegati alla rete AWS tramite collegamenti di rete ad alta capacità. Ad esempio, con [AWS Direct Connect](),[14] puoi stabilire una connettività privata e dedicata dall'ambiente locale ad AWS. Ciò fornisce la larghezza di banda e una latenza costante per caricare i dati nel cloud per scopi di protezione dei dati e livelli coerenti di prestazioni e latenza per i carichi di lavoro ibridi.

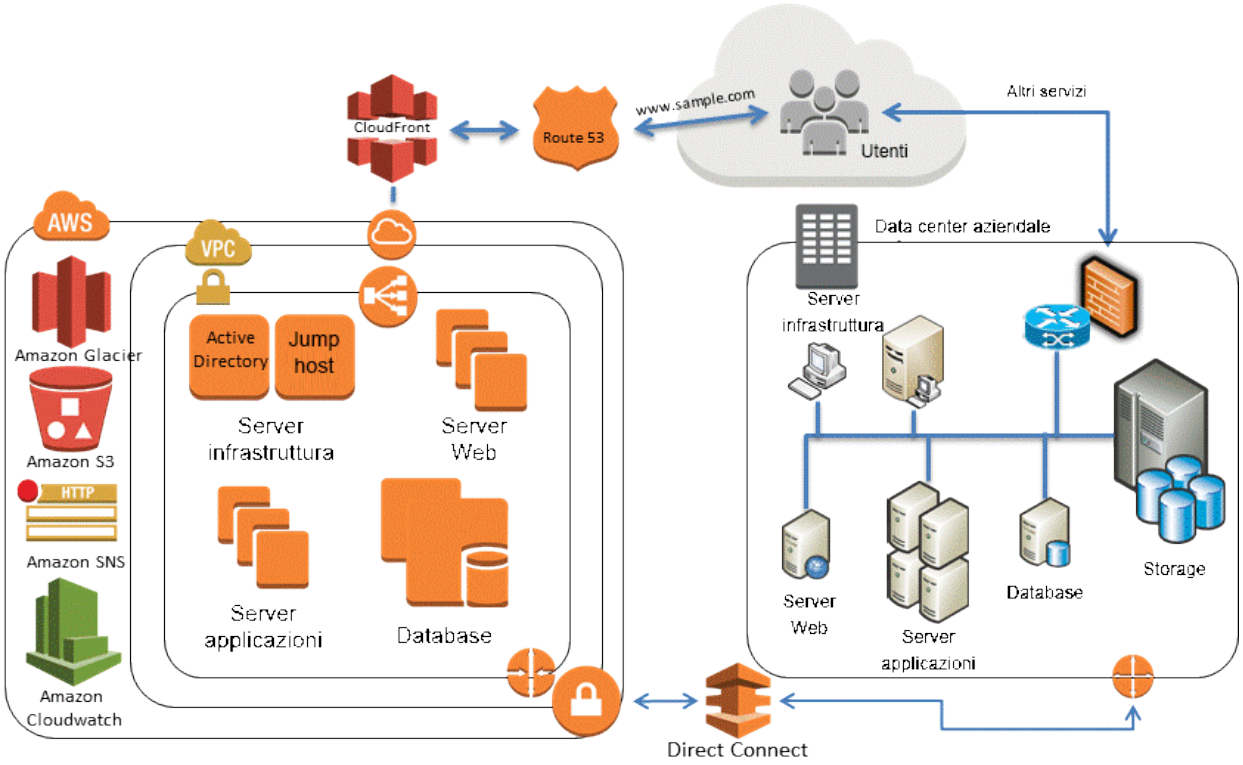


**Figura 8: scenario di un'infrastruttura ibrida**

Le soluzioni ben progettate per la protezione dei dati si basano generalmente su una combinazione dei metodi descritti nelle soluzioni native su cloud e locali.

[14] http://aws.amazon.com/directconnect/

## Backup di applicazioni basate su AWS nel data center

Se già disponi di un framework che esegue il backup dei dati per i server locali, estenderlo alle risorse AWS su una connessione VPN o tramite AWS Direct Connect è un'operazione semplice. Puoi installare l'agente di backup sulle istanze Amazon EC2 ed eseguirne il backup in base alle tue policy di protezione dei dati.

## Migrazione della gestione dei backup al cloud per ottenere disponibilità

A seconda dell'architettura di backup, potresti disporre di un server di backup master e di uno o più server multimediali o di storage a livello locale con i servizi che proteggono. In questo caso, dovresti spostare il server di backup master su un'istanza Amazon EC2 per proteggerlo da emergenze locali e ottenere un'infrastruttura di backup ad alta disponibilità.

Per gestire i flussi di dati di backup, dovresti inoltre creare uno o più server multimediali su istanze Amazon EC2. I server multimediali vicini alle istanze Amazon EC2 ti permettono di risparmiare denaro sui trasferimenti su Internet e, in caso di backup su S3 o Amazon Glacier, aumentano le prestazioni complessive di backup e ripristino.

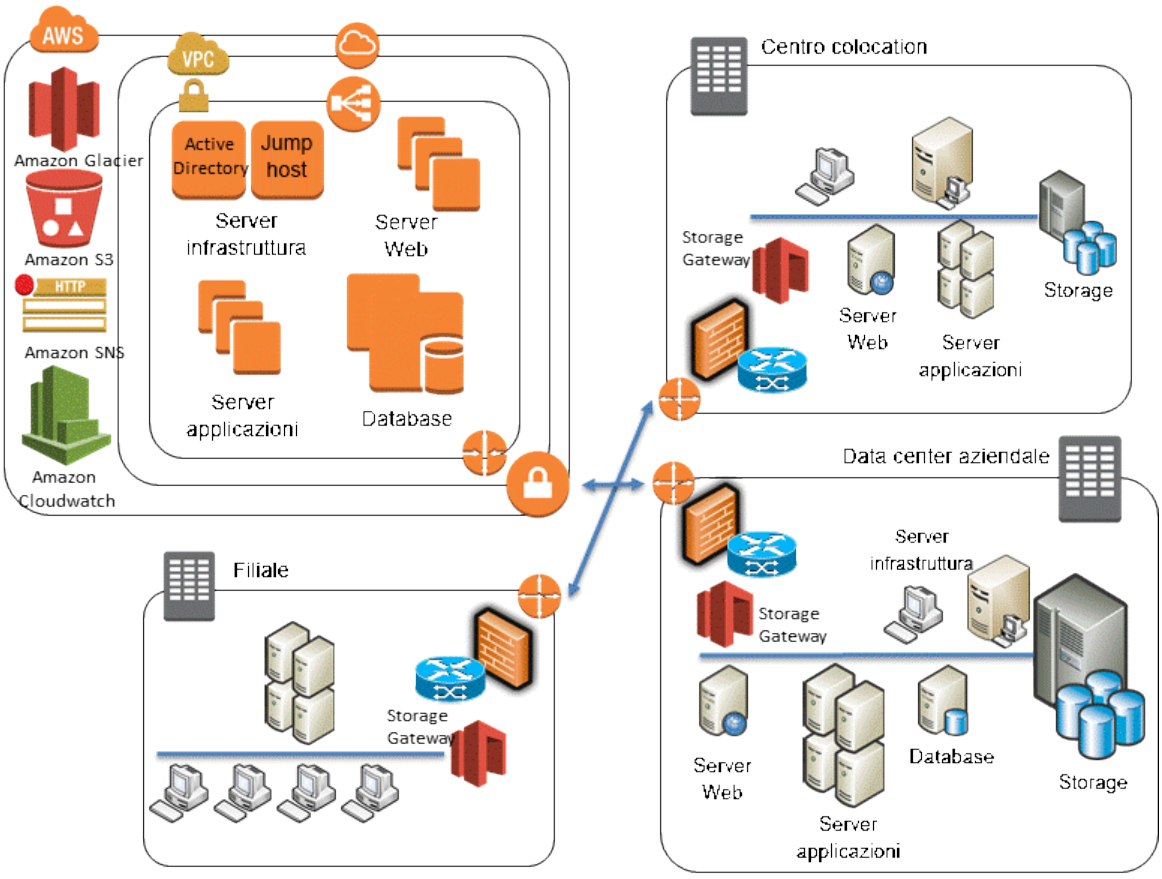


**Figura 9: utilizzo di gateway nello scenario ibrido**

## Esempio di scenario ibrido

Supponi di gestire un ambiente in cui esegui il backup di istanze Amazon EC2, server standalone, macchine virtuali e database. L'ambiente dispone di 1.000 server ed esegui il backup del sistema operativo, dei dati dei file, delle immagini delle macchine virtuali e dei database. I database di cui eseguire il backup sono 20 (una combinazione di MySQL, Microsoft SQL Server e Oracle).

Il tuo software di backup dispone di agenti che eseguono il backup di sistemi operativi, immagini di macchine virtuali, volumi di dati, database SQL Server e database Oracle (tramite RMAN). Per le applicazioni come MySQL, per le quali il tuo software di backup non ha agenti, potresti utilizzare l'utility client mysqldump per creare un file dump del database sul disco dove gli agenti di backup standard possono proteggere i dati.

Per proteggere questo ambiente, il tuo software di backup di terze parti molto probabilmente avrà un server di catalogo globale o un server master che controlla le attività di backup, archiviazione e ripristino, nonché più server multimediali collegati allo storage basato su disco, alle unità nastro LTO (Linear Tape-Open) e ai servizi di storage AWS.

Il modo più semplice di potenziare la soluzione di backup con i servizi di storage AWS è utilizzare il supporto del fornitore di backup per Amazon S3 o Amazon Glacier. Ti consigliamo di collaborare con il tuo fornitore per conoscere le opzioni relative a integrazione e connettori. Per un elenco dei fornitori di software di backup che collaborano con AWS, consulta la nostra partner directory.[15]

Se il tuo software di backup esistente non supporta in modo nativo lo storage nel cloud per il backup o l'archiviazione, puoi utilizzare un dispositivo gateway di storage come ponte tra il software di backup e Amazon S3 o Amazon Glacier.

Esistono numerose soluzioni di gateway di terze parti. Inoltre, in questo caso, puoi utilizzare le appliance virtuali AWS Storage Gateway poiché utilizzano tecniche generiche come i volumi basati su iSCSI e le librerie di nastri virtuali (VTL). Questa configurazione richiede un hypervisor supportato (VMware o Microsoft Hyper-V) e lo storage locale per ospitare l'appliance.

## Archiviazione dei dati con AWS

I dati vengono archiviati quando occorre conservarli per motivi aziendali o di conformità. A differenza dei backup, generalmente eseguiti per mantenere una copia dei dati di produzione per un breve periodo di tempo per il ripristino in caso di danneggiamento o perdita dei dati, l'archiviazione mantiene tutte le copie dei dati fino alla scadenza della policy di conservazione.

Un archivio efficace rispetta i criteri seguenti:

- Durabilità dei dati per un'integrità a lungo termine
- Sicurezza dei dati

[15] http://www.aws-partner-directory.com/PartnerDirectory/PartnerSearch?type=ISV

- Semplicità di ripristino
- Costi ridotti

Archivi di dati non modificabili possono essere un altro requisito normativo o di conformità.

Amazon Glacier fornisce archivi a basso costo, crittografia nativa dei dati inutilizzati, durabilità del 99,9999999% e capacità illimitata.

Amazon S3 Standard - Infrequent Access è una valida scelta per i casi d'uso che richiedono il rapido recupero dei dati. Amazon Glacier è un'ottima scelta per i casi d'uso in cui si accede frequentemente ai dati e in cui è possibile accettare tempi di recupero di diverse ore.

Gli oggetti possono essere suddivisi su più livelli in Amazon Glacier tramite le regole del ciclo di vita in S3 o tramite l'API Amazon Glacier. La caratteristica Vault Lock di Amazon Glacier ti permette di distribuire e applicare facilmente i controlli di conformità per le singole camere di sicurezza Amazon Glacier con una policy di blocco della camera di sicurezza. In una policy di blocco della camera di sicurezza puoi specificare controlli come WORM (Write Once, Read Many; una scrittura, più letture) e bloccare la policy per impedire modifiche future. Per maggiori informazioni, consulta Amazon Glacier.

# Protezione dei dati di backup in AWS

La sicurezza dei dati è una preoccupazione comune. Per AWS, la sicurezza è molto importante ed è la base di ogni servizio offerto. Servizi di storage come Amazon S3 forniscono solide funzionalità per il controllo degli accessi e per la crittografia dei dati sia in transito che inattivi. Tutti gli endpont di Amazon S3 e dell'API Amazon Glacier supportano la crittografia SSL per i dati in transito. Amazon Glacier esegue la crittografia di tutti i dati inattivi per impostazione predefinita. Con Amazon S3, i clienti possono scegliere la crittografia lato server per gli oggetti inattivi lasciando ad AWS la gestione delle chiavi di crittografia, fornendo le proprie chiavi durante il caricamento di un oggetto o utilizzando l'integrazione AWS Key Management Service (AWS KMS) [16]per le chiavi di crittografia. In alternativa, i clienti possono sempre crittografare i dati prima di caricarli su AWS. Per maggiori informazioni, consulta Amazon Web Services: panoramica sulle procedure di sicurezza.

# Conclusioni

Gartner ha riconosciuto AWS come leader in ambito di servizi di storage nel cloud pubblico[17]. AWS offre ottime soluzioni per aiutare le organizzazioni a spostare i carichi di lavoro su piattaforme basate su cloud: la nuova generazione di backup. AWS fornisce soluzioni a basso costo e scalabili per consentire alle organizzazioni di equilibrare i requisiti di backup e archiviazione. Questi servizi si integrano perfettamente con le tecnologie che utilizzi oggi.

---

[16] http://docs.aws.amazon.com/AmazonS3/latest/dev/UsingKMSEncryption.html

[17] http://www.gartner.com/technology/reprints.do?id=1-1WWKTQ3&ct=140709&st=sb

# Collaboratori

Le persone indicate di seguito hanno collaborato alla stesura di questo documento:

- Pawan Agnihotri, solutions architect, Amazon Web Services
- Lee Kear, solutions architect, Amazon Web Services
- Peter Levett, solutions architect, Amazon Web Services

# Revisioni del documento

Aggiornato a maggio 2016