



ARMAZENAMENTO

Block Storage – Amazon EBS

O Amazon Elastic Block Storage é um serviço que fornece volumes de armazenamento em blocos, e que pode ser usado com instâncias EC2. Se você desligar ou apagar uma instância do Amazon EC2, todos os dados no volume do EBS anexo permanecerão disponíveis, permitindo reanexar a uma instância.

Para criar um volume do EBS, basta definir a configuração (como tamanho e tipo do volume) e a zona de disponibilidade da região onde será provisionado.

Depois de criar um volume do EBS, ele pode ser anexado a uma instância do Amazon EC2, similar à forma como você anexa um HD externo ao seu computador. Os volumes EBS agem de forma muito parecida a um HD externo. Grande parte dos volumes do Amazon EBS só pode ser conectada a uma instância por vez. A maioria dos volumes do EBS tem uma relação um para um com instâncias do EC2, portanto, eles não podem ser compartilhados ou anexados a várias instâncias ao mesmo tempo (recentemente, a AWS anunciou o recurso multi-attach do Amazon EBS, que permite que volumes sejam anexados a várias instâncias do EC2 ao mesmo tempo. Esse recurso não está disponível para todos os tipos de instância e todas as instâncias devem estar na mesma zona de disponibilidade).

Dica

Você pode desanexar um volume do EBS de uma instância do EC2 e anexá-lo a outra instância do EC2 que esteja na mesma zona de disponibilidade para acessar os dados nela contidos.

A unidade externa é separada do computador. Isso significa que, se ocorrer um acidente e o computador cair, você ainda terá seus dados na unidade externa. Isso também vale para volumes EBS em relação a uma instância Amazon EC2.

Snapshots de EBS

Como os volumes do EBS são para dados que precisam perdurar, é importante fazer backup dos dados. Você pode fazer backups complementares de volumes do EBS criando snapshots do Amazon EBS.

Um snapshot do EBS é um backup incremental. Isso significa que o primeiro backup de um volume copia todos os dados. Nos backups subsequentes, somente os blocos de dados que foram alterados desde o snapshot mais recente serão salvos.

Exemplo

Se você tiver 10 GB de dados em um volume e apenas 2 GB de dados tiverem sido modificados desde o último snapshot, apenas os 2 GB que foram alterados serão gravados no Amazon Simple Storage Service (Amazon S3).

Quando você faz um snapshot de qualquer um dos seus volumes EBS, os backups são armazenados de forma redundante em várias zonas de disponibilidade usando o Amazon S3. Esse aspecto de armazenar o backup no Amazon S3 é tratado pela AWS e, portanto, você não precisará interagir com o Amazon S3 para trabalhar com seus snapshots do EBS. Você os gerencia no console do Amazon EBS, que faz parte do console do Amazon EC2.

Os snapshots do EBS podem ser usados para criar vários volumes, estejam eles na mesma zona de disponibilidade ou em outra.

Quando você cria um novo volume a partir de um snapshot, ele é uma cópia exata do volume original no momento em que o snapshot foi obtido.

Casos de uso do EBS

O Amazon EBS é útil quando você precisa recuperar dados rapidamente e manter os dados por um longo prazo. Os volumes são comumente usados nos seguintes cenários:

Sistemas operacionais

Volumes de inicialização/raiz para armazenar um sistema operacional. O dispositivo raiz de uma instância executada a partir de uma imagem de máquina da Amazon (AMI) geralmente é um volume do Amazon EBS. Eles são comumente referidos como AMIs com suporte de EBS.

Bancos de dados

Uma camada no Amazon EC2 de armazenamento para bancos de dados em execução que dependem de leituras e gravações transacionais.

Aplicativos corporativos

O Amazon EBS fornece armazenamento de blocos confiável para executar aplicativos essenciais aos negócios.

Aplicativos com taxa de transferência intensiva

Aplicativos que executam leituras e gravações longas e contínuas.

Tipos de EBS

Os volumes do Amazon EBS são organizados em duas categorias principais: unidades de estado sólido (SSDs) e unidades de disco rígido (HDDs). Os SSDs oferecem forte desempenho para entrada/saída aleatória (E/S), enquanto os HDDs oferecem forte desempenho para E/S sequencial. A AWS oferece dois tipos de cada um. A tabela a seguir pode ajudar você a decidir qual volume de EBS é a melhor opção para sua carga de trabalho.

Tipos de volume EBS	Descrição	Caso de uso	Tamanho	IOPS Máximo	Throughput Máximo
EBS Provisioned IOPS SSD (io1)	SSD de desempenho mais alto, projetado para carga de trabalho transacional sensível à latência	NoSQL com uso intensivo de I/O e bancos de dados relacionais	4 GB–16 TB	64,000	1,000 MB/s
EBS Geral Purpos e SSD	SSD de uso geral que equilibra	Volumes de boot, aplicativos interativos de	1 GB–16 TB	16,000	250 MB/s

Tipos de volume EBS	Descrição	Caso de uso	Tamanho	IOPS Máximo	Throughput Máximo
(gp2/gp3)	preço e desempenho para uma ampla variedade de cargas de trabalho transacionais	baixa latência, desenvolvimento e teste			
Throughput Optimized HDD (st1)	HDD de baixo custo projetado para cargas de trabalho com alto rendimento e acesso frequente	Big data, data warehouses, processamento de logs	500 GB–16 TB	500	500 MB/s
Cold HDD (sc1)	HDD de menor custo	Dados mais frios que exigem	500 GB–	250	250 MB/s

Tipos de volume EBS	Descrição	Caso de uso	Tamanho	IOPS Máximo	Throughput Máximo
	projetado para cargas de trabalho acessadas com menos frequência	menos varreduras por dia	16 TB		

Volumes EBS e suas características.

Gustavo Ribeiro

Object storage - Amazon S3

Ao contrário do Amazon Elastic Block Store (Amazon EBS), o Amazon Simple Storage Service (Amazon S3) é uma solução de armazenamento independente, que não está vinculada à computação e permite que você recupere seus dados de qualquer lugar na web. Se você usou um serviço de armazenamento on-line para fazer backup dos dados de sua máquina local, provavelmente usou um serviço semelhante ao Amazon S3.

Nesse serviço, você armazena seus objetos em contêineres chamados de buckets (baldes).

Não é possível fazer upload de nenhum objeto, nem mesmo uma única foto, para o Amazon S3 sem criar um bucket primeiro.

Ao criar um bucket você especifica, no mínimo, dois detalhes: o nome desse bucket e a região da AWS na qual deseja que ele resida.



Para escolher uma região, você normalmente selecionará uma que tenha usado para outros recursos, como sua computação. Quando você escolhe uma região para seu bucket, todos os objetos que você coloca dentro dele serão armazenados de forma redundante em vários dispositivos, em várias zonas de disponibilidade. Esse nível de redundância foi projetado para fornecer aos clientes do Amazon S3, 99,999999999% de durabilidade e 99,99% de disponibilidade para objetos em um determinado ano.

Versionamento no S3

Conforme vimos, o Amazon S3 identifica objetos em parte usando o nome do objeto. Por exemplo, ao carregar uma foto, você pode nomear o objeto `photo.gif` e armazená-lo em uma pasta chamada `PhotosFiles`. Se você não usar o controle de versão do Amazon S3, toda vez que fizer upload de um objeto chamado `photo.gif` para a pasta `PhotosFiles`, ele substituirá o arquivo original.

Isso pode ser um problema por vários motivos como, por exemplo:

Exemplo 1

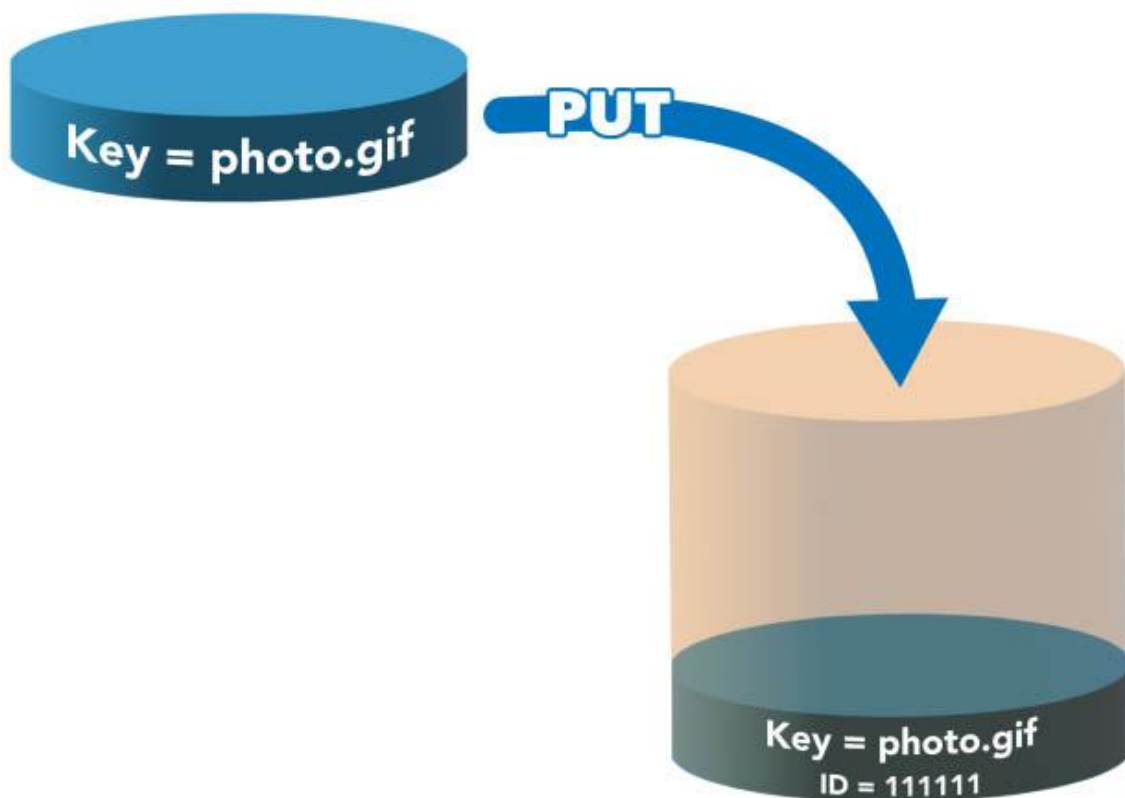
O nome do arquivo `photo.gif` é um nome comum para um objeto de foto. Você ou outra pessoa que tenha acesso ao bucket pode não ter a intenção de

substituí-lo, mas uma vez substituído, o arquivo original não pode ser acessado.

Exemplo 2

Você pode querer preservar versões diferentes de photo.gif. Sem controle de versão, se você quiser criar uma nova versão de photo.gif, precisará carregar o objeto e escolher um nome diferente para ele. Ter vários objetos, todos com pequenas diferenças nas variações de nomenclatura, pode causar confusão e desordem nos buckets do S3.

Para neutralizar esses problemas, você pode usar o controle de versão do S3. O controle de versão mantém várias versões de um único objeto no mesmo bucket. Isso preserva versões antigas de um objeto sem usar nomes diferentes, o que ajuda na recuperação de arquivos de exclusões acidentais, substituições acidentais ou falhas de aplicativos.



Bucket PhotosFiles, com versionamento habilitado.

Se você habilitar o controle de versão para um bucket, o Amazon S3 gerará automaticamente um ID de versão exclusivo para o objeto.

Exemplo

Em um bucket você pode ter dois objetos com a mesma chave, mas diferentes IDs de versão, como photo.gif (versão 111111) e photo.gif (versão 111122).

Os buckets com versionamento, habilitados para controle de versão, permitem que você recupere objetos de exclusão ou substituição acidental.

A exclusão de um objeto não o remove permanentemente.

Em vez disso, o Amazon S3 coloca um marcador no objeto que mostra que você tentou excluí-lo. Se quiser restaurar o objeto, você pode remover o marcador e ele restabelece o objeto. Se você substituir um objeto, isso resultará em uma nova versão do objeto no bucket. Mas você ainda terá acesso às versões anteriores do objeto.

Classes de armazenamento no S3

Quando você carrega um objeto no Amazon S3 e não especifica a classe de armazenamento, você o carrega na classe de armazenamento padrão. Tudo que aprendemos até aqui é sobre a classe de armazenamento padrão do Amazon S3 mesmo sem você saber!

As classes de armazenamento do Amazon S3 permitem que você altere seu nível de armazenamento quando suas características de dados mudarem. Por exemplo, se você estiver acessando suas fotos antigas com pouca frequência, talvez queira alterar a classe de armazenamento das fotos para economizar custos.

Saiba mais um pouco sobre cada uma das principais classes de armazenamento no S3:

Standard

O S3 Standard oferece um armazenamento de objetos com altos níveis de resiliência, disponibilidade e performance para dados acessados com frequência. Como fornece baixa latência e alto throughput, o S3 Standard é adequado para uma grande variedade de casos de uso, como aplicativos na nuvem, sites dinâmicos, distribuição de conteúdo, aplicativos móveis e de jogos e dados analíticos de big data.

Standard-IA

O S3 Standard-IA é indicado para dados acessados com menos frequência, mas que exigem acesso rápido, quando esses dados são necessários. A categoria S3 Standard-IA oferece os altos níveis de resiliência e throughput e uma baixa latência da categoria S3 Standard, com taxas reduzidas por GB de armazenamento e por GB de recuperação. A combinação de baixo custo e alta performance torna a classe S3 Standard-IA ideal para armazenamento de longa duração, backups e data stores para arquivos de recuperação de desastres.

Glacier Instant Retrieval

A Amazon S3 Glacier Instant Retrieval é uma classe de armazenamento de arquivos que oferece o armazenamento de custo mais baixo para dados de longa duração, que raramente são acessados e exigem recuperação em milissegundos. Com a S3 Glacier Instant Retrieval, você pode economizar até 68% nos custos de armazenamento em comparação com o uso da classe de armazenamento S3 Standard-IA, quando seus dados são acessados uma vez por trimestre. A S3 Glacier Instant Retrieval oferece o acesso mais rápido ao armazenamento de arquivo, com a mesma taxa de transferência e acesso em milissegundos que as classes de armazenamento S3 Standard e S3 Standard – IA. O S3 Glacier Instant Retrieval é ideal para arquivar dados que precisam de acesso imediato, como imagens médicas, recursos de mídia de notícias ou arquivos de conteúdo gerado pelo usuário.

Glacier Flexible Retrieval

O S3 Glacier Flexible Retrieval oferece armazenamento de baixo custo, com custo até 10% menor (do que o S3 Glacier Instant Retrieval), para dados de arquivo que são acessados 1 a 2 vezes por ano e recuperados de forma assíncrona. Para dados de arquivo que não requerem acesso imediato, mas que precisam da flexibilidade para recuperar grandes conjuntos de dados sem nenhum custo, como casos de uso de backup ou recuperação de desastres, a S3 Glacier Flexible Retrieval é a classe de armazenamento ideal. É uma solução ideal para backup, recuperação de desastres, necessidades de armazenamento externo de dados e para quando alguns dados ocasionalmente precisam ser recuperados em minutos e você não quer se preocupar com custos.

Glacier Deep Archive

A S3 Glacier Deep Archive é a classe de armazenamento mais acessível do Amazon S3 e oferece suporte à retenção e preservação digital de longo prazo para dados que podem ser acessados uma ou duas vezes por ano. Essa

classe é projetada para clientes que mantêm conjuntos de dados por 7 a 10 anos ou mais para cumprir requisitos de conformidade regulatória, especialmente em setores altamente regulados como serviços financeiros, saúde e setores públicos. O S3 Glacier Deep Archive também pode ser usado para casos de uso de backup e recuperação de desastres, além de ser uma alternativa mais barata e fácil de gerenciar em comparação aos sistemas de fita magnética como bibliotecas on-premises ou serviços externos, e podem ser restaurados em até 12 horas.

Casos de uso

O Amazon S3 é um serviço de armazenamento amplamente utilizado, com muito mais casos de uso do que caberia em uma tela. A lista a seguir resume algumas das formas mais comuns de usar o Amazon S3:

Backup e armazenamento

O Amazon S3 é um lugar natural para fazer backup de arquivos porque é altamente redundante. Conforme mencionado anteriormente, a AWS armazena seus snapshots do EBS no S3 para aproveitar sua alta disponibilidade.

Hospedagem de mídia

Como é possível armazenar objetos ilimitados e cada objeto individual pode ter até 5 TBs, o Amazon S3 é um local ideal para hospedar uploads de vídeos, fotos e músicas.

Data lakes

O Amazon S3 é uma base ideal para um data lake devido à sua escalabilidade virtualmente ilimitada. Você pode aumentar o armazenamento de gigabytes para petabytes de conteúdo, pagando apenas pelo que usar.

Sites estáticos

É possível configurar seu bucket S3 para hospedar um site estático de HTML, CSS e scripts do lado do cliente.

File storage – Amazon EFS

No armazenamento de arquivos, vários clientes (como usuários, aplicativos, servidores e assim por diante) podem acessar dados armazenados em pastas de arquivos compartilhadas. Nessa abordagem, um servidor de armazenamento organiza os arquivos por meio do uso de armazenamento em bloco, com um sistema local de arquivos. Os clientes acessam os dados por meio de caminhos de arquivo.

Comparado ao armazenamento em blocos e ao armazenamento de objetos, o armazenamento de arquivos é ideal para casos de uso em que muitos serviços e recursos precisam acessar os mesmos dados ao mesmo tempo.



O Amazon Elastic File System (Amazon EFS) é um sistema de arquivos escalável, usado com os serviços de nuvem AWS e recursos locais. À medida que você adiciona e remove arquivos, o Amazon EFS expande e retrai automaticamente, de forma que pode dimensionar sob demanda para petabytes sem interromper os aplicativos.