Estratégias de Otimização de Custos na Nuvem e Instâncias Spot

Termos de Uso





Propriedade Growdev

Todo o conteúdo deste documento é propriedade da Growdev. O mesmo pode ser utilizado livremente para estudo pessoal.

É proibida qualquer utilização desse material que não se enquadre nas condições acima sem o prévio consentimento formal, por escrito, da Growdev. O uso indevido está sujeito às medidas legais cabíveis.



Por que otimizar custos na nuvem é essencial?

A computação em nuvem oferece flexibilidade, escalabilidade e inovação, mas sem um controle adequado, os custos podem se tornar altos e inesperados. Estratégias de otimização são fundamentais para equilibrar eficiência e economia, permitindo que empresas obtenham o máximo retorno de seus investimentos.

Pontos para análise:

- Flexibilidade: Pague apenas pelo que usar.
- **Economia:** Reduza custos significativos com estratégias inteligentes.
- Crescimento Sustentável: Invista recursos no que realmente importa.



O que são Instâncias Spot?

Instâncias spot são recursos computacionais vendidos com grande desconto quando a capacidade está ociosa. São interruptíveis a critério do provedor.

Quem Oferece:

- **AWS:** Spot Instances
- **Azure:** Azure Spot Virtual Machines
- Google Cloud: Preemptible VMs (ou Spot VMs no modelo mais recente)



Vantagens e Desvantagens das Instâncias Spot

Instâncias spot são vantajosas, mas é importante considerar seus riscos.

Vantagens:

- Custos até 90% menores.
- Excelente para workloads não críticos e temporários.
- Disponibilidade em múltiplas regiões e tamanhos.

Desvantagens:

- Possibilidade de interrupção sem aviso.
- Não é recomendada para workloads críticos ou de alta disponibilidade.



Quando Usar Instâncias Spot?

As instâncias spot são mais adequadas para cenários tolerantes a falhas.

- Processamento batch (ex.: ETL, big data).
- Treinamento de modelos de machine learning.
- Renderização de vídeos e gráficos.
- Ambientes de teste ou desenvolvimento.



Instâncias Spot vs Instâncias Padrão

A escolha entre spot e on-demand depende da necessidade de custo e criticidade.

Característica	Instâncias Spot	Instâncias On-demand
Custo	Até 90% mais barato	Preço fixo
Interrupção	Sim (sem aviso prévio)	Não
Cargas adequadas	Tolerantes a falhas	Críticas e persistentes
Disponibilidade	Quando há capacidade ociosa	Sempre disponível



Como as Clouds Oferecem Instâncias Spot?

AWS Spot Instances

 Oferece integração com Auto Scaling e EC2 Fleet, ideal para cargas ajustáveis.

Azure Spot VMs

 Suporte para escalabilidade dinâmica e integração com políticas de desalocação.

• Google Preemptible/Spot VMs

Recomendadas para workloads curtos, com limite de 24h por execução (*preemptibles*).



Boas Práticas para Usar Instâncias Spot

Use as instâncias spot de forma estratégica para aproveitar os benefícios.

- Planeje sua aplicação para lidar com interrupções (ex.: checkpoints).
- Configure alertas para monitorar interrupções e custos.
- Use ferramentas nativas, como Auto Scaling (AWS) e escalabilidade dinâmica (Azure/Google).

Resumo da ópera

- **Definição:** Instâncias spot oferecem economia significativa com riscos de interrupção.
- Casos de Uso: Workloads tolerantes a falhas, como batch e machine learning.
- **Disponibilidade:** AWS, Azure e Google Cloud possuem opções de instâncias spot.
- Boas Práticas: Planejamento, automação e ferramentas nativas são essenciais para maximizar o uso.

