

www.geekuniversity.com.br

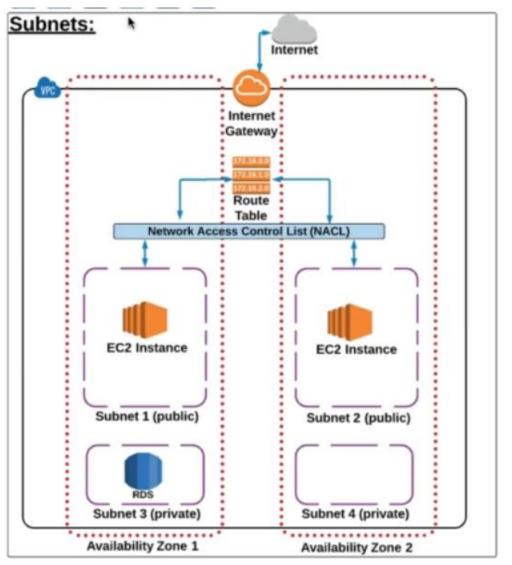


Qualquer serviço AWS que você utiliza, como EC2, RDS, etc, estarão alocados em uma subrede presente em uma zona de disponibilidade de uma região AWS onde a sua VPC está alocada.

Você pode, e deveria, utilizar múltiplas zonas de disponibilidade para criar redundância na sua arquitetura. Desta forma você estará aplicando a Alta Disponibilidade e Tolerância a Falhas aos seus sistemas.

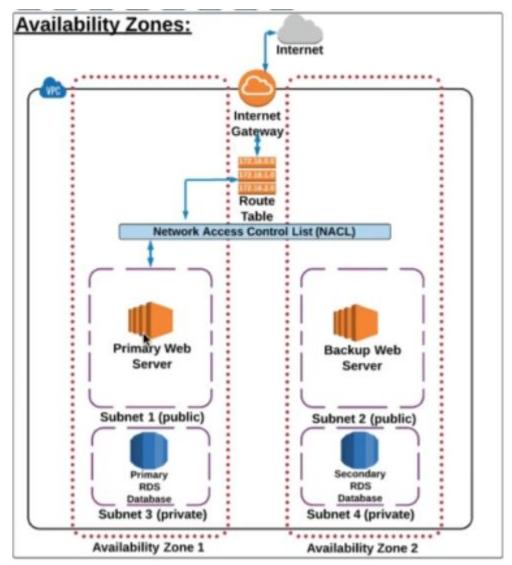
Alocando recursos em diferentes zonas de disponibilidade caso aconteça algo em uma destas zonas, a outra irá garantir que seu sistema continue online.

Claro que alocar recursos não é suficiente, pois temos que configurar os serviços para realizar a replicação de dados.

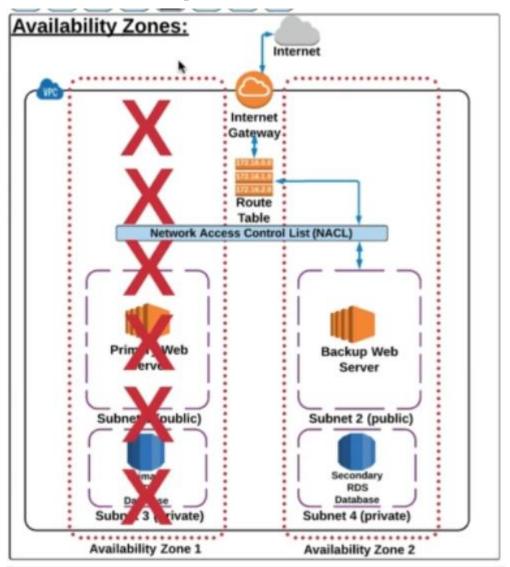


Note neste exemplo que temos recursos replicados em 2 diferentes zonas de disponibilidade.

Note também que a comunicação por padrão está sendo realizada somente nos recursos primários.



Se os recursos primários ficarem indisponíveis nossos recursos redundantes assumirão as respostas fazendo com que os sistemas continuem respondendo.



Alta Disponibilidade:

Criar uma arquitetura de forma que as aplicações estejam sempre disponíveis, ou pelo menos com um baixíssimo tempo de inatividade. É comum sistemas de alta disponibilidade prometer 99,98% de disponibilidade.

Tolerância a Falhas:

É a habilidade do seu sistema de se manter disponível por mais que ocorram falhas em algum dos pontos de risco.

Sobre tolerância a falhas você pode ouvir algo como:

- Um dos meus servidores web falhou, mas meu servidor de backup assumiu o trabalho imediatamente.
- Se algo ocorrer no meu sistema, ele irá se auto reparar.

Em resumo, zonas de disponibilidade são datacenters isolados em uma região AWS.

Em uma zona de disponiblidade podemos ter 1 ou mais subredes.

É altamente recomendado que façamos uso de múltiplias zonas de disponibilidade para manter nossa arquitetura com alta disponibilidade e tolerante a falhas.



www.geekuniversity.com.br