**Question**

- How does Docker Cloud balance my nodes among different availability zones?
- What is Docker Cloud's high availability schema on AWS?

**Answer**

By default, Docker Cloud tries to deploy your node cluster using a high availability strategy. To do this, it places every instance one by one in the less populated availability zone for that node cluster. We can see this behavior with some examples:

### Allowing Docker Cloud to manage VPC/subnets

Suppose it's the first time you're deploying a node cluster with Docker Cloud in a region, say Sao Paulo (South America, sa-east-1) and you delegated deployment management to Docker Cloud. You didn't send any `provider_options` using the API, and you left `VPC/subnet/security groups/IAM role` values set to their defaults on the dashboard. Docker Cloud will:

1. Docker Cloud looks for a VPC called `dc-vpc`. The VPC does not exist on the first try, so Docker Cloud must create it and a `dc-gateway`, which attaches to the VPC.
2. Docker Cloud retrieves all subnets in the VPC. No subnets exist on the first try, so Docker Cloud creates them. For every availability zone (AZ), Docker Cloud splits the VPC CIDR IP space in (# of AZs + 1) blocks and tries to create (# of AZs) subnets. Remember, we left space for custom subnets. For every subnet, Docker Cloud tries to dry-run an instance of the selected type and creates it if the operation succeeds, creating and associating a `dc-route-table` to the subnet.
3. Once all subnets have been created, Docker Cloud deploys every node of the cluster using a round-robin pattern. (Note that there can be fewer subnets than AZs due to failures when dry-running the instances.)

### Scaling a node cluster

Following the example described in the previous section, you have a node cluster deployed and want to scale it up. Docker Cloud will:

1. Look for `dc-vpc` Found!
2. Look for `dc-subnets` Found!
3. Count the nodes in every subnet. Now Docker Cloud chooses the less populated subnet and deploys the next node there.
4. Go to 3 until all nodes are deployed.

### Choosing where to deploy

Suppose we have another VPC for other purposes, with its components already created, and we want to deploy a node cluster in that VPC. Docker Cloud will:

1. Look for the selected VPC. Found!
2. Look for selected subnets. If you do not select any subnets, Docker Cloud tries to create them using the rules previously described.
3. If you selected more than one subnet, Docker Cloud distribute the nodes in the cluster among those subnets. If not, all nodes are placed in the same subnet.