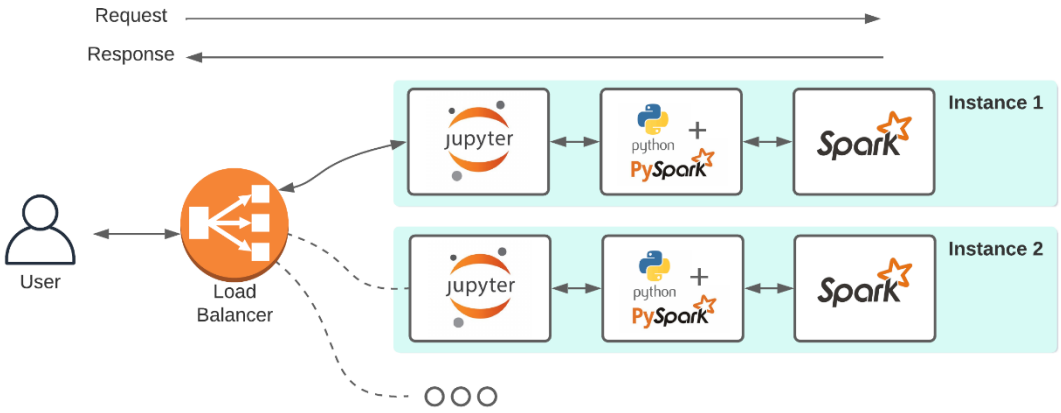

Figure 1: Software Architecture. Interactions between the software components in your cloud. A load balancer forwards communications between the user and the Jupyter servers. Jupyter, Python, PySpark, and Spark only interact within their instance. For example, Jupyter interacts with the Python kernel, and the Python kernel interacts with Spark via PySpark. There are no interactions between different instances.

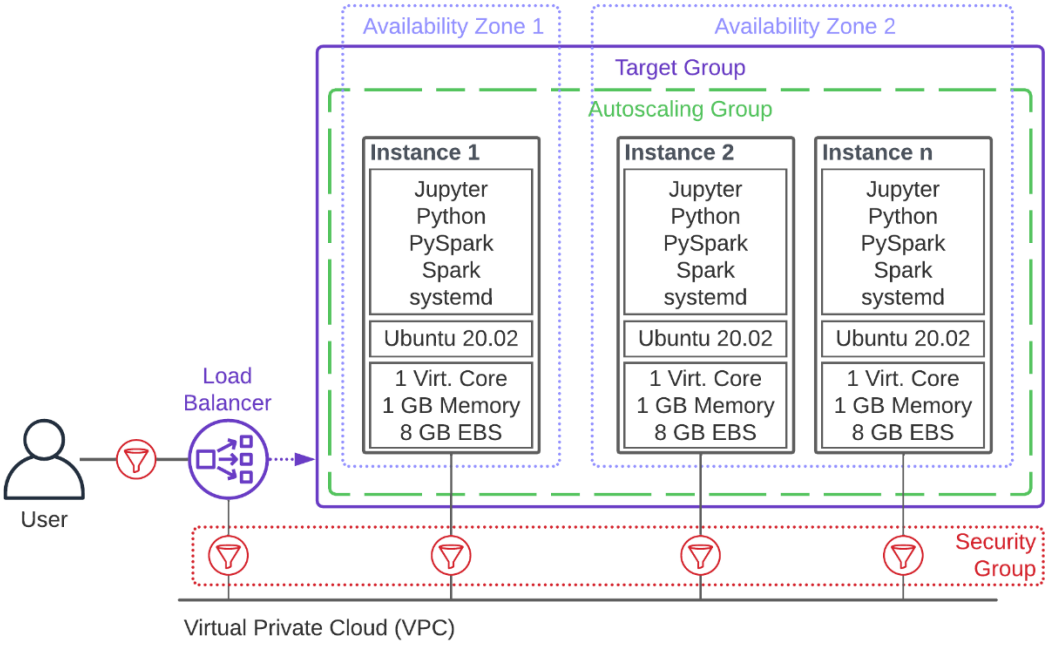

Figure 2: Software Architecture. The organization and networking of resources in your cloud. Users interact only with the load balancer, which forwards requests (responses) to (from) instances in the VPC, and all communications are filtered through security groups (e.g., only packets directed at port 8888 should be allowed). Each instance is composed of applications (top box), an operating system (middle box), and compute/storage/memory resources (bottom box). Note that the number of instances is not always 3; the autoscaling group should request (and delete) new instances when computational demand rises (or falls).

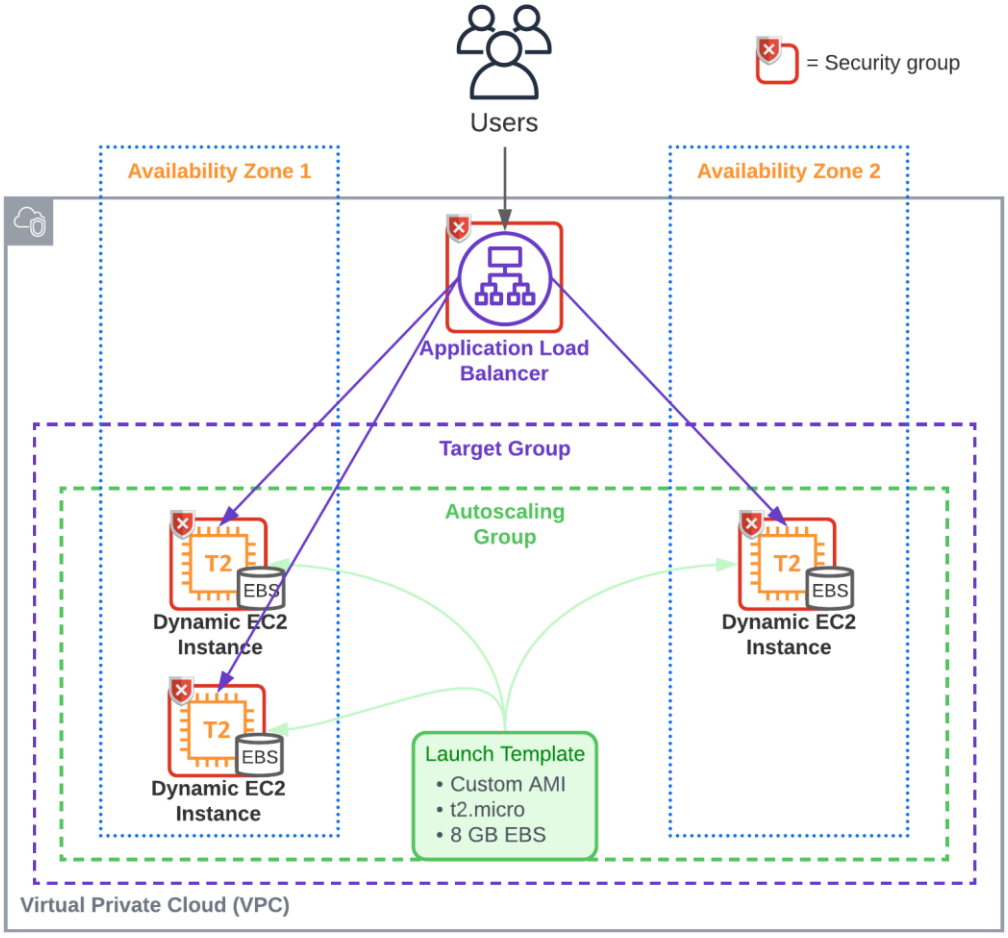

Figure 3: Systems Architecture. Alternative visualization of the system in figure 2 in terms of the modules provided by Amazon Web Services (AWS).