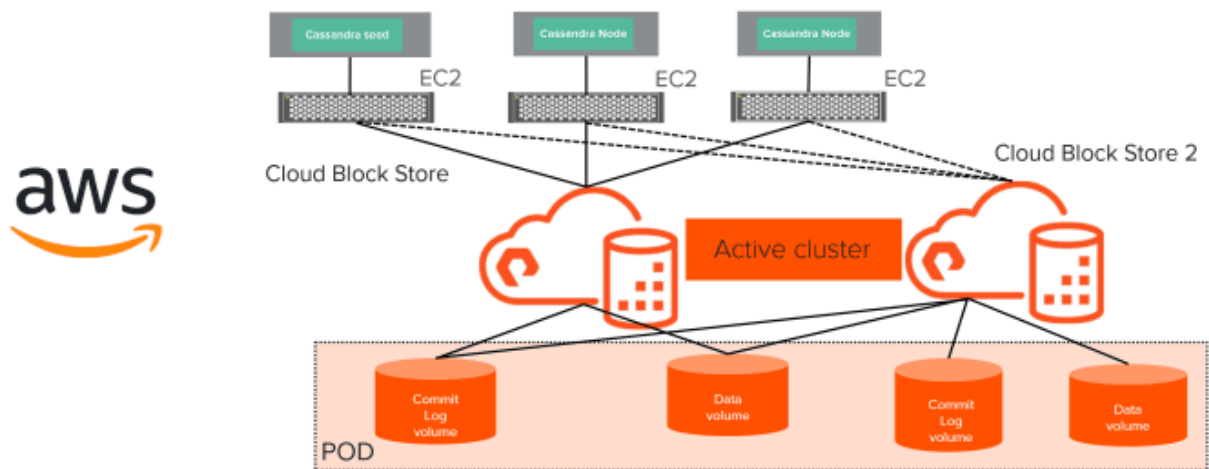


Benefits of Running Apache Cassandra on Pure Storage Cloud Block Store

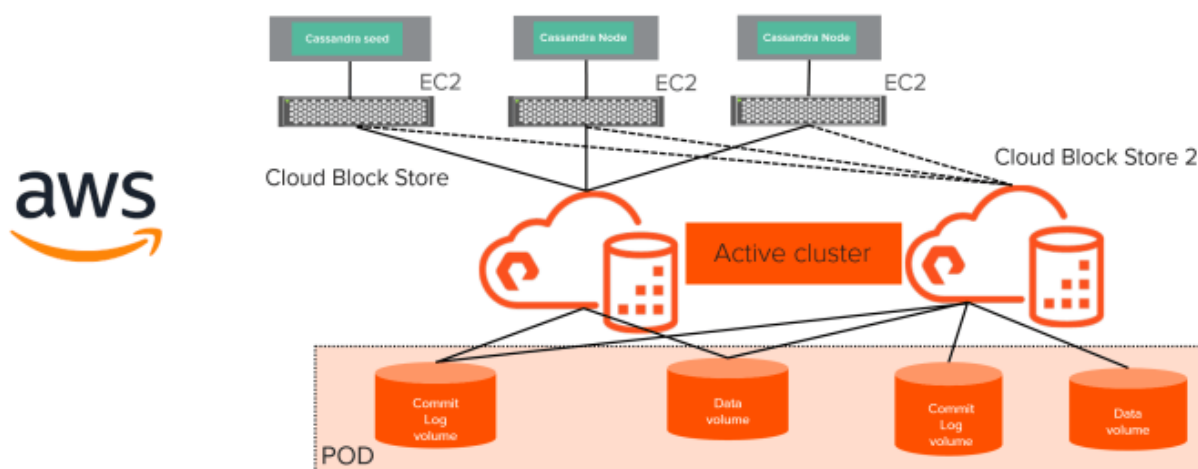
Cassandra setup on Cloud Block Store



Pure Storage® launch of Cloud Block Store for Amazon Web Services (AWS) offers a way for Pure platforms to be extended into the cloud. In this blog/demo, see the benefits of running Cassandra on Pure Storage Cloud Block Store for AWS. I chose to demo Cassandra because it is one of the top 5 NoSQL databases in AWS [based on this survey](#).

To demonstrate this, I built a 3-node cluster on Cloud Block Store, which is connected to another Cloud Block Store instance via ActiveCluster to provide maximum resiliency. The 3 node Cassandra cluster has data and commit logs on Cloud Block Store volumes.

Cassandra setup on Cloud Block Store



I built another Cassandra cluster directly on EBS for comparison. AWS even has a [best practices guide that recommends flash storage](#).

The following value propositions are covered in this demo:

1. Data reduction of Cassandra data on Cloud Block Store
2. Maximum Resiliency of Cassandra cluster on Cloud Block Store using ActiveCluster.
3. Using Cloud Block Store instantaneous snapshots to protect or clone Cassandra cluster.

Data Reduction

The data reduction on Cloud Block Store is around 2.4 to 3.0 for the data and commit logs volumes. The data on Cassandra table was not compressed. This reduces the storage footprint by a big factor on Cloud Block Store. On AWS Elastic Block Store there is no data reduction.

Resiliency on Cloud Block Store

For maximum resiliency of the Cassandra on Cloud Block Store, it can be deployed on two instances of Cloud Block Store. These two Cloud Block Store instances are connected via ActiveCluster. The Cassandra data and log volumes on Cloud Block Store are added to ActiveCluster stretched POD. In the demo, to show resiliency one instance of Cloud Block Store is shutdown to simulate failure and then we check on the cluster status. As the instances are connected by ActiveCluster the data is synchronously replicated to the secondary Cloud Block Store instance. In case of failure of a Cloud Block Store instance, the Cassandra cluster will be up and running as if nothing happened. Pure Storage Cloud Block Store ActiveCluster here offers lot more than enterprise-grade resiliency.

Cloud Block Store snapshots

Cloud Block Store snapshots are instantaneous, which makes it a great feature to protect or clone your

Cassandra clusters. In the demo I used them to recover the Cassandra cluster—also as mentioned before, it can be used to copy or clone the Cassandra cluster for development and testing purposes. It will be very difficult to do this using EBS snapshots. Cassandra offers native snapshots but it is very difficult to use. Most of the customers cannot take more than 2-3 native Cassandra snapshots before they run out of storage space, whereas Pure Storage Cloud Block Store snapshots are instantaneous and do not consume much space to start with. Also, we can take thousands of Pure Storage Cloud Block Store snapshots without any issues.