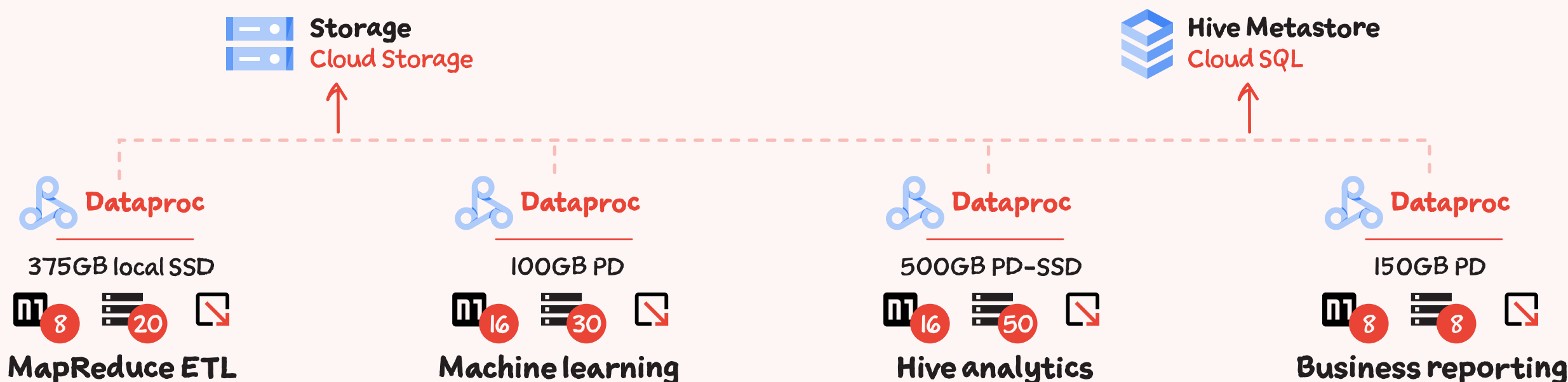




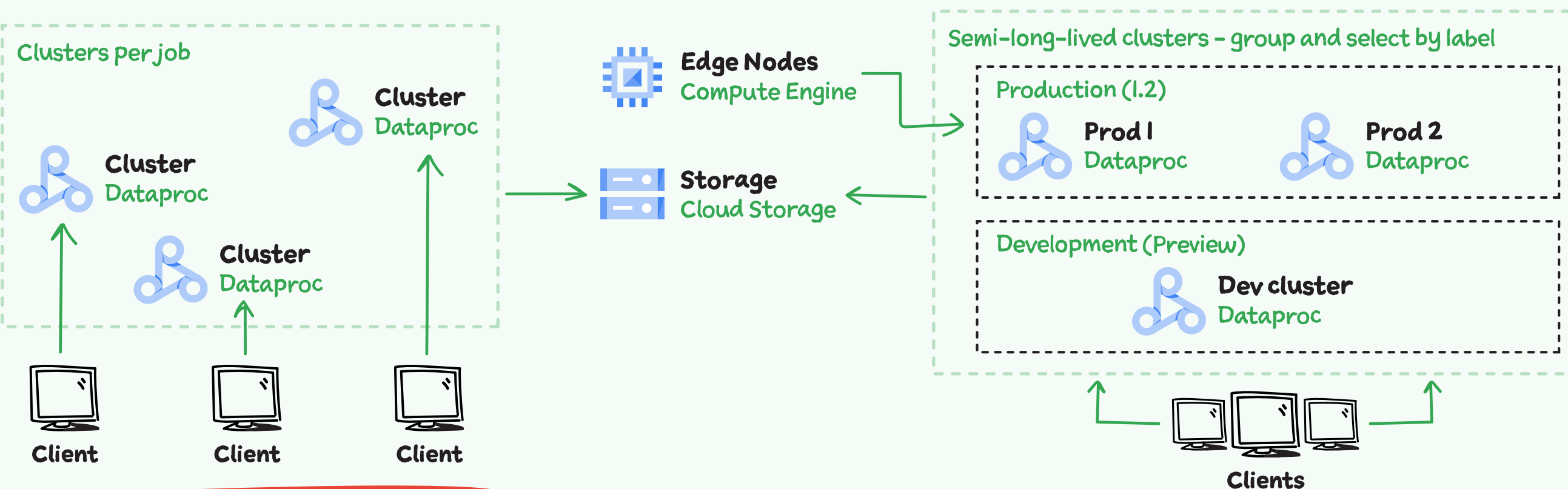
# Dataproc #GCPsketchnote

@PVERGADIA THECLOUDGIRL.DEV 12.07.2020

Can I right-size clusters per use case?

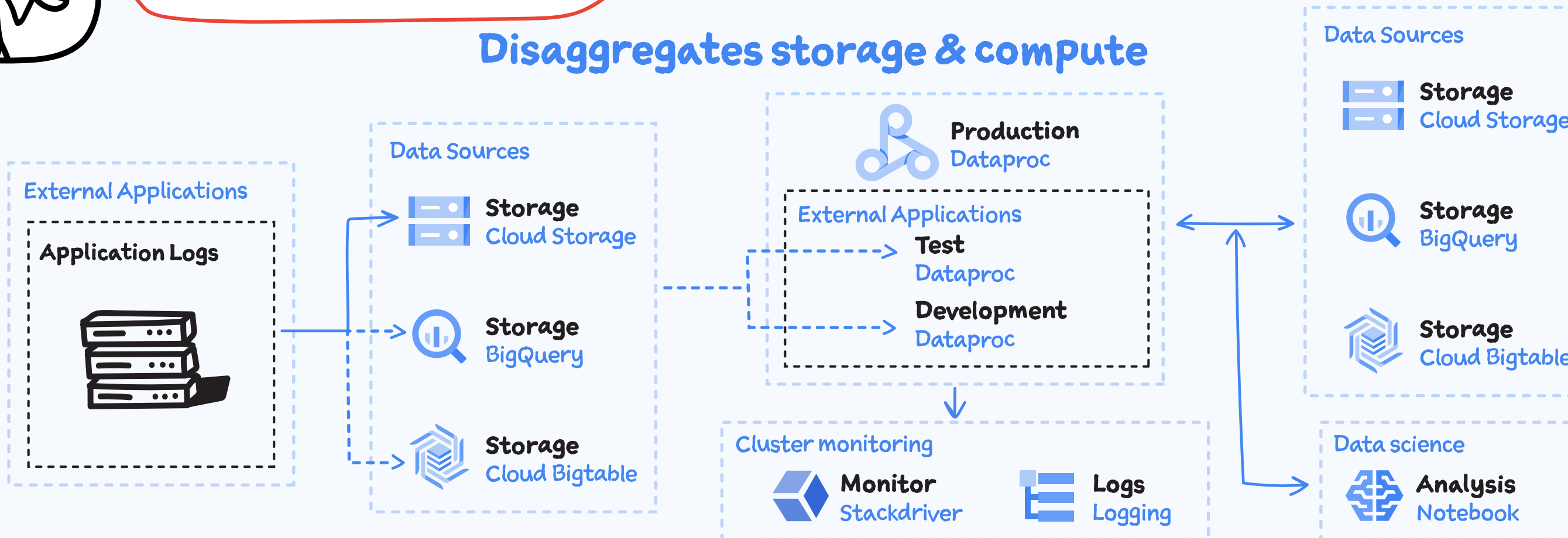


How do I use ephemeral & long-lived clusters to save cost?



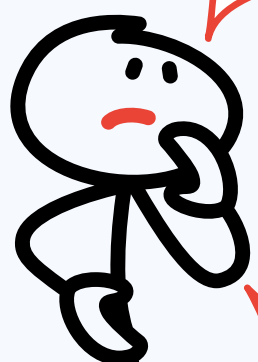
How does it work?

## Disaggregates storage & compute



Managing Apache Hadoop cluster is hard!

ERIN



We need a tool that make it easy to gain insights from our data.

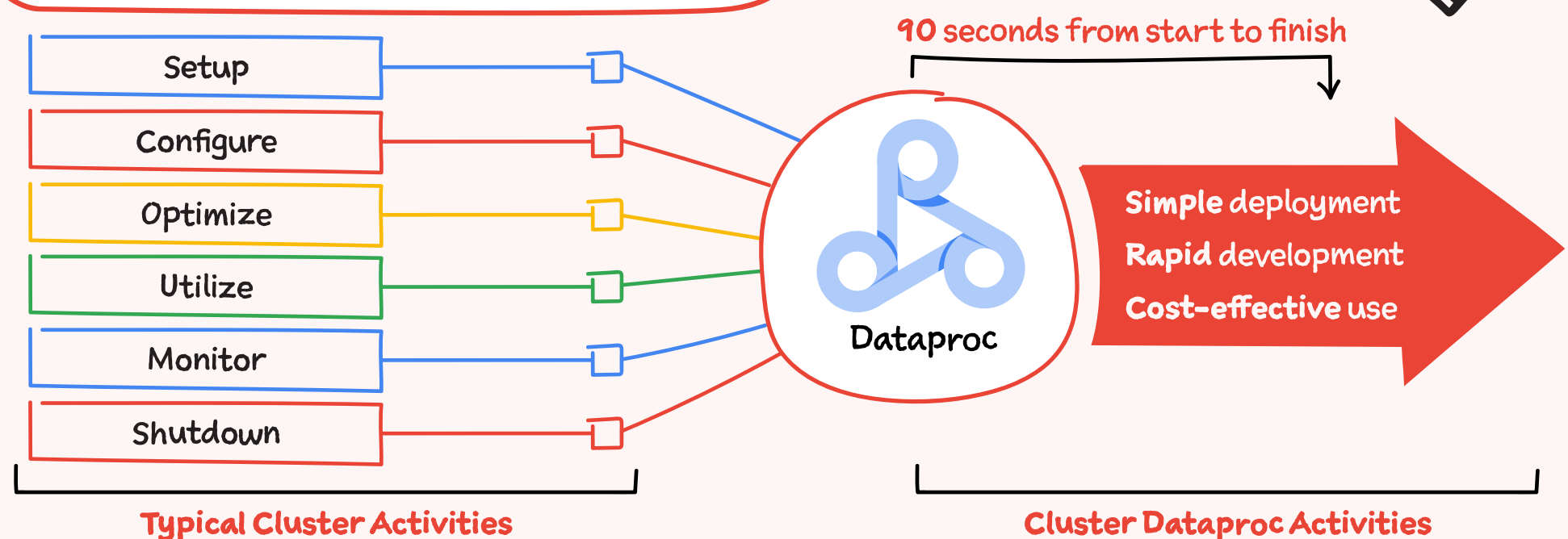
Yes, it is slow, difficult & expensive!



SAM

Dataproc!

How is a typical Hadoop/Spark cluster installation different from Dataproc?



Familiar open source tools



Customizable machines  
Easily autoscale up or down

Ephemeral clusters on-demand  
Save cost

Tightly Integrated with  
Google Cloud Platform

Rapid cluster creation  
Spin up a cluster in 90 seconds

How do I move Spark/Hadoop jobs to Dataproc?

