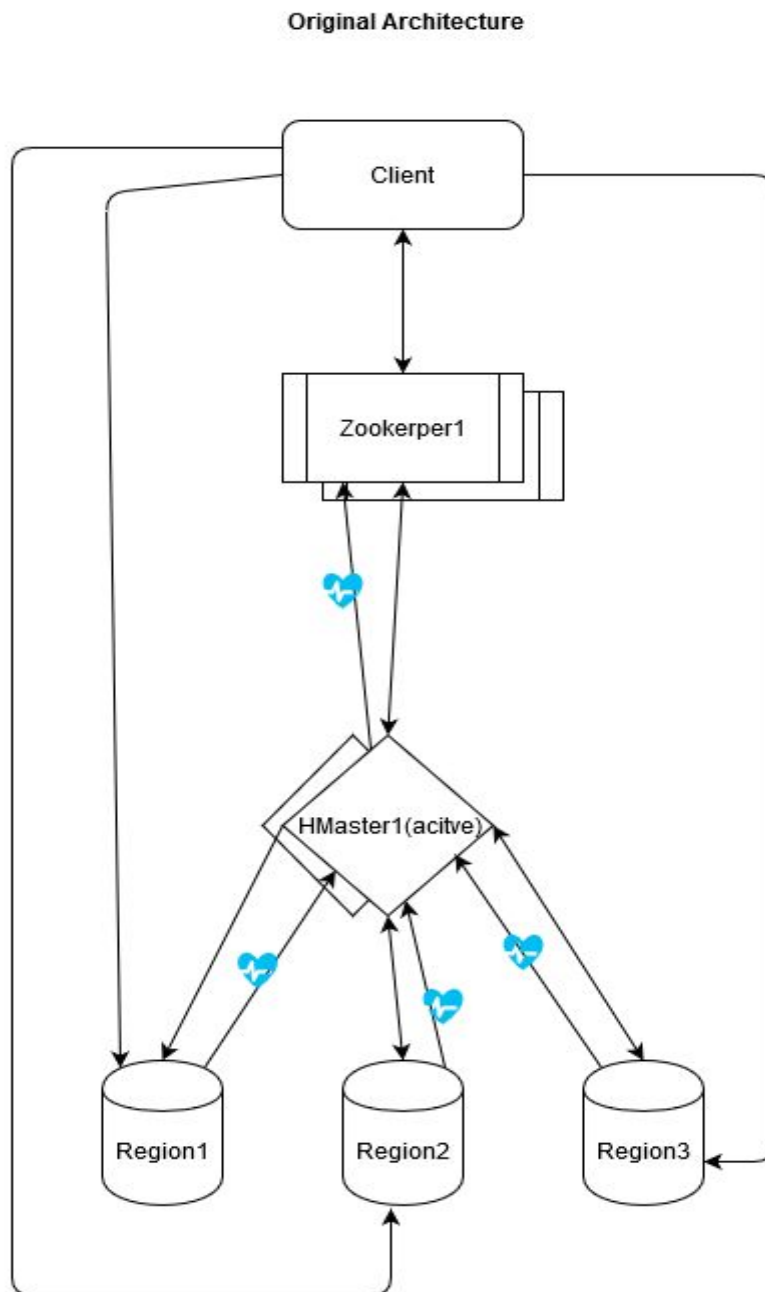
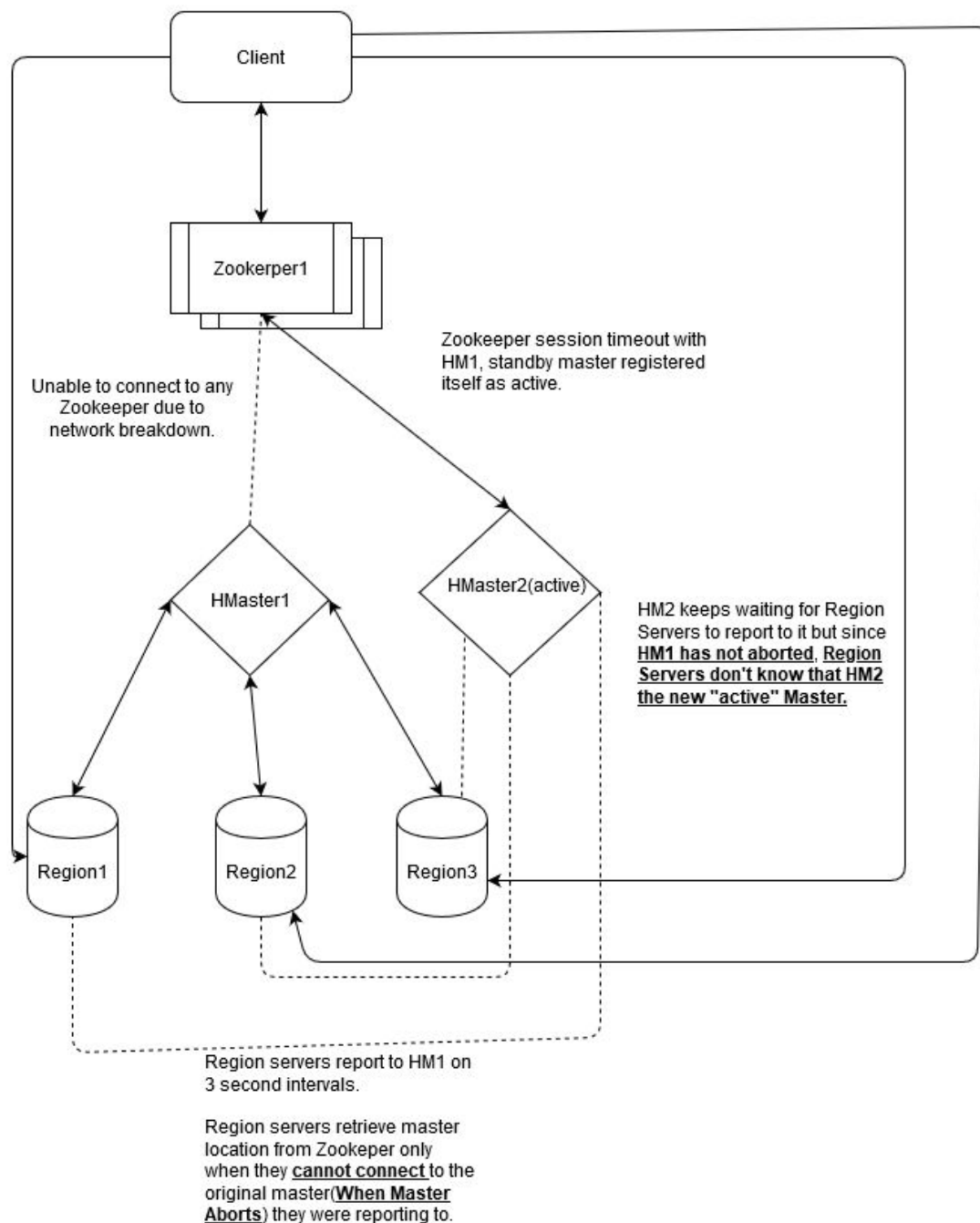


1. An overview of HBase architecture.



2. (a) Visual diagram of the bug



(b) The general reason for the occurrence of the bug is the inability of **HM1** to not **abort** when it **cannot connect** with **Zookeeper**. Since the HM1 doesn't abort, the Region servers have no reason to believe that HM1 isn't active anymore, keep reporting to HM1 and don't ask the Zookeeper for a new active HMaster(HM2). Another issue that comes up is when one of the Region servers crashes, it would not be properly detected because the notification will never come and read/write can still happen. So though may not be data loss due to WAL, but there may be data inconsistency i.e different clients getting different views of the data.

3. I used the unique_count aggregation method, with X axis buckets more than the number of records.

This data was taken on April 23,202 at 10:55pm.

I found a total of 8 unique Hostnames for over the last 1 year.

	Hostname	Count
1.	llebeat-demo-green-78xb2	44,687,753
2.	filebeat-demo-green-xtmd5	17,047,332
3.	filebeat-demo-green-zjzrx	2,127,265
4.	filebeat-demo-green-rfs9k	1,672,279
5.	siem-es	619,442
6.	raspberrypi	369,411
7.	suricata-iowa	225,579
8.	zeek-iowa	126,297