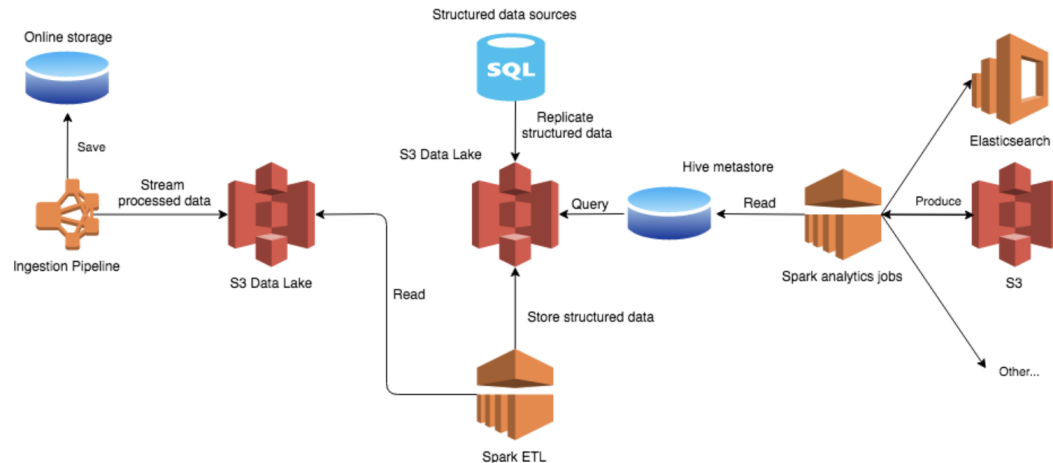


Here is how our data lake infrastructure roughly looks:



- The activity ingestion pipeline processes data and stores the data in online storage. Simultaneously, the pipeline streams data to a Kinesis