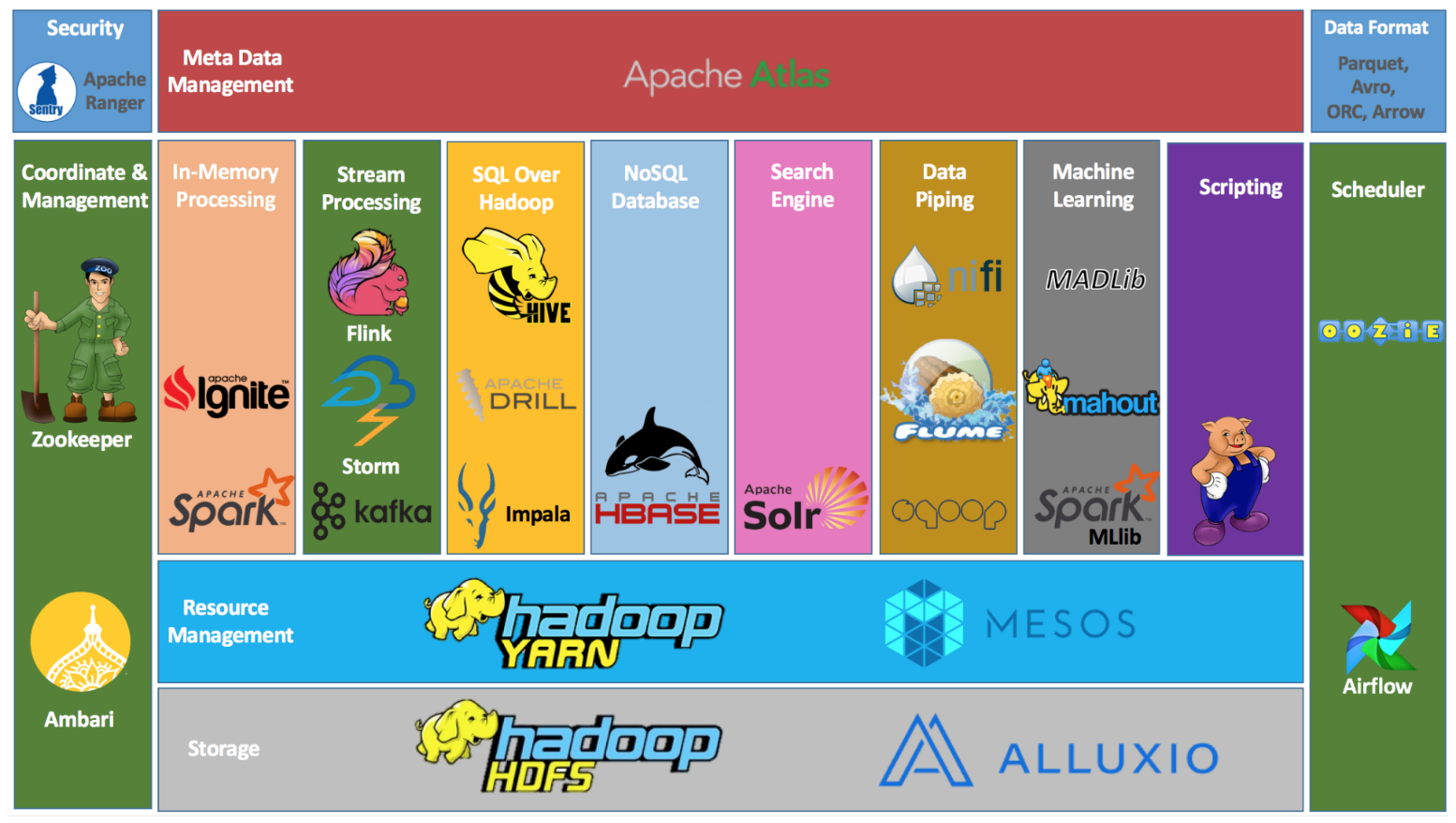
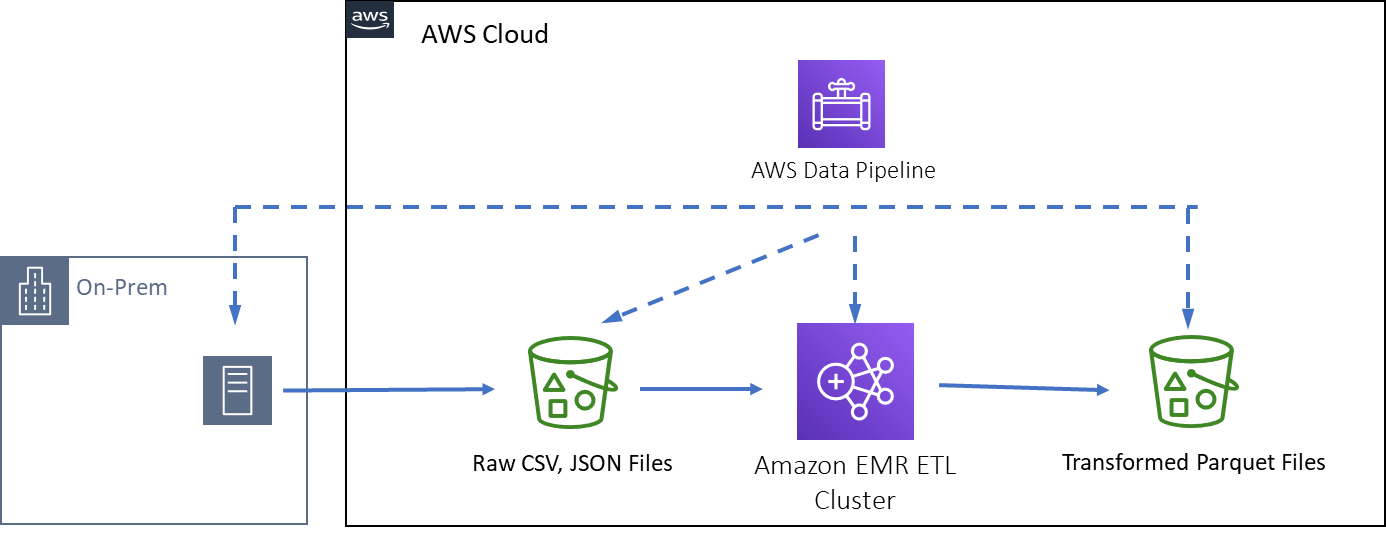
EMR Integration Patterns

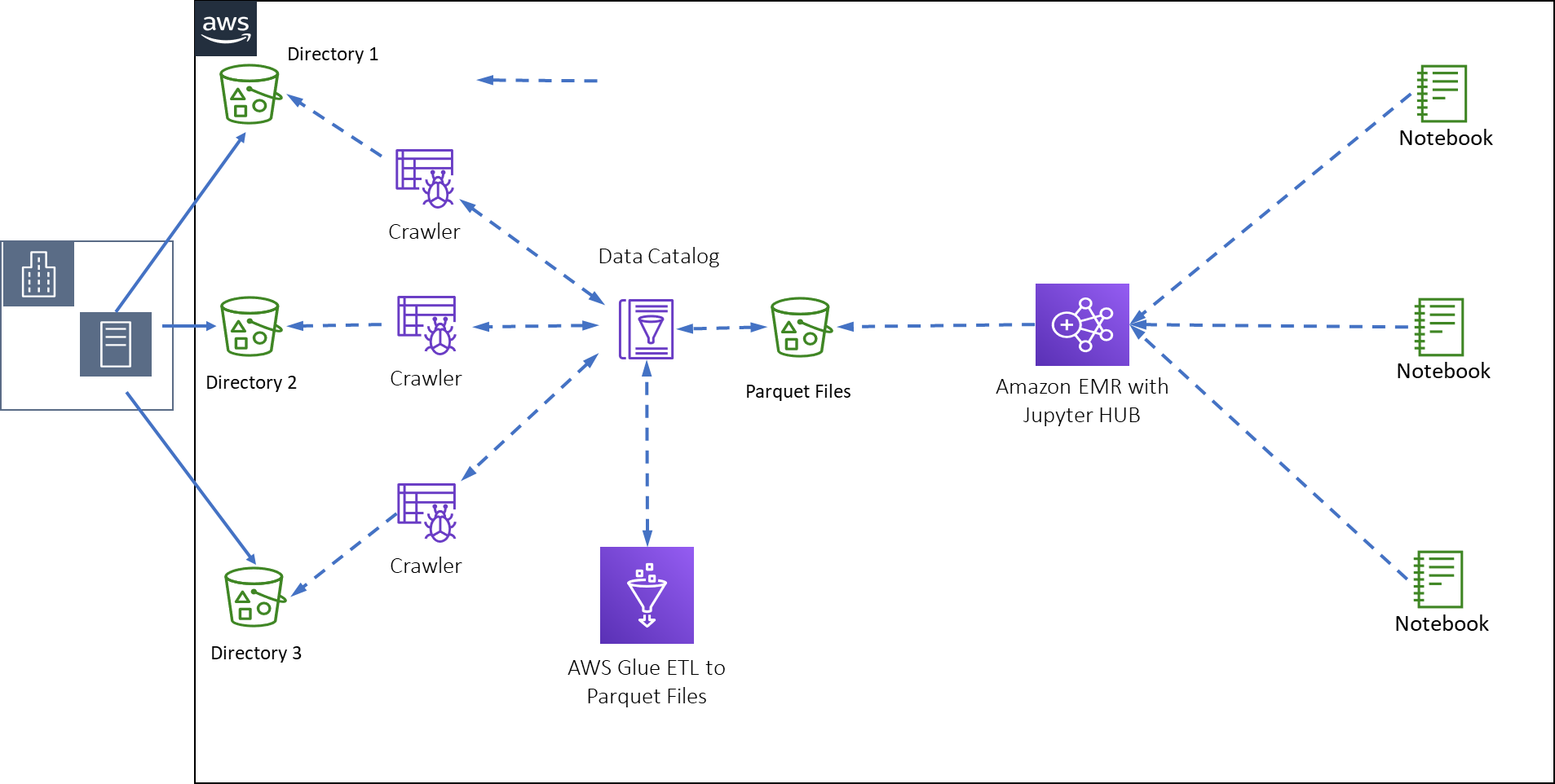
1. Resource Manager
   1. YARN (Default)
   2. Mesos
   3. Kubernetes
2. Cluster Lifetime
   1. Persistent
   2. Transient
3. Usage
   1. Notebook
      1. Zupyter
      2. Zeppelin
   2. Job Steps
4. Columnar Storage Formats for Input Files
   1. Supported By Glue
      1. Parquet (<https://parquet.apache.org>)
      2. ORC (<https://orc.apache.org/>)
   2. Not Supported By Glue
      1. CarbonData (<https://carbondata.apache.org/>)
      2. Arrow (<https://arrow.apache.org/>)
5. Stream Processing Frameworks
   1. Storm (<https://storm.apache.org/>)
   2. Flink (<https://flink.apache.org/>)



1. Transformatin of Huge Raw Data to Parquet Files



1. Batch Processing with Jupyter Hub and Parquet Files



1. Batch Processing with Jupyter Notebook and Parquet Files

