**Practical Questions to be confirmed with Yerri**

1. How do you monitor Spark jobs . Are you using <http://localhost:4040>
2. How do you run Spark jobs ? Is it using Autosys tools ? Is oozie used ?
3. How to you retrieve the logs in driver ? How to look at individual logs in data nodes ? Is there a place where it is consolidated
4. What log libraries is used ? Is it SL log4j ?
5. What are the monitoring tools which are used ?
6. What are the performance management tools used?
7. What is the deployment procedure ? Is there CI/CD pipeline involved
8. Apache Mesos Vs Yarn – which one should be used under which conditions