

Topics

| |
|--|
| Problems with traditional systems What is Hadoop and what does it do? Hadoop Cluster Overview Introduction to Git and IntelliJ IDE |
| MapReduce Hadoop Jobs and Tasks The Hadoop Distributed File System Understand your first Mapreduce program Hadoop 1 vs 2 |
| Installation and configuration of Hadoop in Pseudo-distributed mode Writing basic MapReduce programs in Java Accessing HDFS Programmatically |
| MapReduce job chaining Performance optimization Pig: high level map-reduce programming language Discuss Project definition |
| Hive: Batch processing SQL-like querying engine Impala: Ad-hoc query processing tool |
| Sqoop: Data transfer tool between Relational database and HDFS |
| Spark: In memory data processing framework |
| Cluster management tools Discussion on Industry use cases of tools and techniques covered so far Road map for further studies |

Contact Details:

Email: kantesariyashyam@gmail.com

LinkedIn: <https://www.linkedin.com/in/kantesariyashyam/>