Questions:  
1. Describe the purpose of Oozie ( 4 marks)  
2. Describe a scenario where using Oozie is needed ( 4 marks)  
3. What is a Fork and Join nodes in Apache Oozie? ( 2 marks)  
4. What is a Workflow in Apache Oozie? (2 marks)

1. Oozie is a workflow scheduler system that manages Apache Hadoop Jobs. The purpose of Oozie is to simplify the process of Hadoop jobs by allowing users to build and execute workflows of Hadoop jobs a lot easier.
2. A scenario where using Oozie would be needed would be when dealing with complex data processing workflows that involve more than one Hadoop jobs. Oozie would make the processing of workflows a lot easier when dealing with more then one Hadoop job.
3. A Fork node is used to divide the path of execution into multiple paths of execution whereas the join node is used when waiting for every execution path. The fork node and join node should always be paired up with each other.
4. A Workflow in Oozie is a collection of actions organized in a Directed Acyclic Graph(DAG). A workflow can contain a variety of actions such as Hadoop MapReduce jobs and many others. With decision, fork, and join nodes, Oozie controls the workflow path.