**#ASSIGNMENT 34.3**

**Explain in brief with their uses.**

● Oozie Action and Decision Nodes

● Oozie Workflow Nodes

● Fork and Join

● Oozie Web Console

**Q1-**

**Action nodes:**

1. Action nodes trigger the execution of a computation/processing task.

2. An action node represents a workflow task, e.g., moving files into HDFS, running a MapReduce, Pig or Hive jobs, importing data using Sqoop or running a shell script of a program written in Java.

3. All computation/processing tasks triggered by an action node are remote to Oozie. No workflow application specific computation/processing task is executed within Oozie.

4. All computation/processing tasks triggered by an action node are executed asynchronously by Oozie. For most types of computation/processing tasks triggered by workflow action, the workflow job has to wait until the computation/processing task completes before transitioning to the following node in the workflow.

5. The exception is the fs action that is handled as a synchronous action.

6. Oozie provides recovery capabilities when starting or ending actions.

**Decision Nodes:**

1. A decision node enables a workflow to make a selection on the execution path to follow.

2. The behavior of a decision node can be seen as a switch-case statement.

3. A decision node consists of a list of predicates-transition pairs plus a default transition.

4. Predicates are evaluated in order or appearance until one of them evaluates to true and the corresponding transition is taken. If none of the predicates evaluates to true the default transition is taken.

5. The name attribute in the decision node is the name of the decision node.

6. Each case elements contains a predicate transition name. The predicate ELs are evaluated in order until one returns true and the corresponding transition is taken.

7. The default element indicates the transition to take if none of the predicates evaluates to true.

8. All decision nodes must have a default element to avoid bringing the workflow into an error state if none of the predicates evaluates to true.

**Q2-**

**Oozie Workflow Nodes:**

Workflow nodes are classified in control flow nodes and action nodes:

**1. Control flow nodes:** nodes that control the start and end of the workflow and workflow job execution path.

It is further classified into following:

**1.1 Start Control Node**

The start node is the entry point for a workflow job, it indicates the first workflow node the workflow job must transition to.

**1.2 End Control Node**

The end node is the end for a workflow job, it indicates that the workflow job has completed successfully.

**1.3 Kill Control Node**

The kill node allows a workflow job to kill itself.

**1.4 Decision Control Node**

A decision node enables a workflow to make a selection on the execution path to follow.

**1.5 Fork and Join Control Nodes**

A fork node splits one path of execution into multiple concurrent paths of execution.

A join node waits until every concurrent execution path of a previous fork node arrives to it.

**2. Action nodes:** nodes that trigger the execution of a computation/processing task.

**Q3-**

**Fork and Join Control Nodes:**

1. A fork node splits one path of execution into multiple concurrent paths of execution.

2. A join node waits until every concurrent execution path of a previous fork node arrives to it.

3. The fork and join nodes must be used in pairs. The join node assumes concurrent execution paths are children of the same fork node.

4. The name attribute in the fork node is the name of the workflow fork node.

5. The start attribute in the path elements in the fork node indicate the name of the workflow node that will be part of the concurrent execution paths.

6. The name attribute in the join node is the name of the workflow join node. The to attribute in the join node indicates the name of the workflow node that will executed after all concurrent execution paths of the corresponding fork arrive to the join node.

**Q4-**

**Oozie Web Console:**

1. Oozie web console is one which is used for monitoring the Status of the job, workflow applications status and workflow jobs status.

2. It is read only interface and we cannot create any workflow.

3. We can only read only.

