BigData Test Process

Process followed in Black box testing for BigData domain includes the phases as mentioned below-

**A)Test Requirement Analysis :-**

1. **Scenario Mindmap** **analysis** – This captures the functional and non functional aspects of a feature. The non-functional aspects include performance, compatibility, memory usage/leak as well as trustworthy test types such as reliability, security, privacy, resilience and availability. This document is prepared by the test lead/architect and shared to the developers to incorporate the points in design, implementation and UT-Functional test.
2. **Test Analysis and Design** – The techniques followed here capture the impact analysis , interaction scenario analysis as well as the test type analysis for a feature. Impact analysis captures the impact of the base features on the new feature. Interaction scenario analysis captures the interaction points related to functional and non functional aspects of the new feature with respect to the base features. Test type analysis captures the test type related scenarios of a feature. The output of these analysis techniques are high level test scenarios/specifications that are used as input for test strategy and test case design.
3. **Test Strategy** – This captures the information related to “how to test”, “what to test” and “when to test”. The details captured in test strategy doc are – test scope, version build test plan, test topology and hardware/environments, test tools, test automation framework/strategy and also risks(resource/environment/technical).

**B)Test Design :-**

1. **Equivalence Partitioning, Boundary Value analysis and Pair wise analysis** – These techniques help to identify test items for the test cases.
2. **Test case design** – Using the inputs from the test requirement analysis phase test cases are derived/designed for the different test types in a excel format and imported to test management tool.
3. **Automation framework and script design** –The test cases designed are tried to be automated/scripted using the defined automation framework.

**C)Test Execution :-**

1. **Manual Test Execution** – The non-automated test cases designed are then executed on the supported environments and the test execution results are updated in the test management tool.
2. **Automated Test Execution** – The automated test cases are executed using the Continuous Integration tool such as Jenkins or the other supported test tools/frameworks and the test execution results are analysed.
3. **Tests executed** – The UT and Sanity/Basic verification test cases are executed by developers while the testers execute the Functional test cases and Non Functional test cases related to performance, memory leak/usage, security, reliability, compatibility etc.

**D)Test Reporting , Assessment and Configuration Management :-**

1. **Test Assessment** – The test execution results are monitored and assessed for test sufficiency and gaps are filled.
2. **Test Reports** – The test reports capture the information related to overall stability and quality of the software that has to be delivered to the customers.
3. **Configuration Management** – The test documents created during the test process are maintained in the configuration management tool repository while the test code/scripts are maintained in the test code repository.

**BigData sub test types and aspects considered –**

1. **Functional 🡪Output Comparison test** – The output from the queries executed on the software under test are compared with the master output from the references and compliance to the master output ensured.
2. **Reliability 🡪BigData test** – The test cases are executed on large volume of data (atleast 30k+ records)
3. **Reliability 🡪Concurrent test** – The test cases are executed from concurrent terminals using small or large volume data.
4. **Reliability 🡪Scalability test** – The test cases are executed with increasing/decreasing number of parameters such as nodes, executors, resources etc.
5. **Reliability 🡪Availability test** – The test cases are executed related to high availability wherein the impact of the switch of services/resources across nodes of the cluster on the data operations are verified.
6. **Reliability 🡪Resilience test** – The test cases are executed related to recovery wherein large data inputs on the system lead to resource overloading or memory usage increase or leak are verified.
7. **Reliability 🡪Fault tolerance test** – The test cases are executed related to scenarios where the disk space is full, network access patchy, power on/off as well as scenarios where the related services are down/restarted while the data operations are in progress.
8. **Reliability 🡪Long Run test** – Standard set of reliability test cases are executed daily for 5-7 days and the functionality, performance, memory usage and stability of the operations are monitored .
9. **Compatibility 🡪 Data compatibility(Forward/Backward) test** – The test cases are executed related to the scenarios where data created in one software version is used in another version and the operations are verified on the data.
10. **Security 🡪 Data Security Privacy protection test** – The test cases are executed related to the scenarios where the loaded data is exposed to the user through the interfaces such as command line exceptions or logs.
11. **Performance 🡪 Competitor Benchmark Comparison test** – The test cases are executed related to competitor benchmark comparison for query response time for different operations such as select queries, data load, update, delete or read/write operations executed from single interface or concurrent interfaces.
12. **Memory 🡪 Memory usage/leak test** – The normal user operations are performed in a instrumented setup with memory usage/leak identification and monitoring tools configured and the memory usage/leaks are analysed for the operations from the reports/logs generated from the tools.