**Non Functional Requirements Implemented in Spark and Scala 301**

* **Performance**
  + Application will be loaded with in 5 seconds once it is launched.
  + Maximum time taken to respond to query is in 100ms from Cassandra table.
  + Maximum time taken during initial load of data into Hive tables is around 600 seconds for SFO and 300 seconds for open\_flight with given A2 VM.
  + Application can interact with 3rd party tools only after authentication in Ambari.
  + Application is highly scalable with increased demand of storage and processing by adding additional nodes.
  + For faster retrieval of query results, the input data is stored as RC file format and Partitioning and Clustering is applied for the tables.
* **Security**
  + Different group level access is provided to developers, operational team and administration team using Ambari.
  + While accessing a remote window the session timeout will occur after 300 seconds of inactivity.
  + Users will be provided with respective username and password.
  + Knox provides the single point of authentication over the cluster.
* **Availability**
  + Application will be available 24/7 with Cassandra cluster facing users.
  + It is deployed on Azure Cloud on South American Region.
  + Almost zero downtime for Cassandra access.
* **Integrity**
  + Number of records in the raw data are same as the number of records loaded to HDFS.
  + Incorrect records are sent to client for correction.
  + Number of records in the HDFS are same as number of records in the Hive Tables.
  + Data is matching in Hive and HDFS.
  + After processing the data in hive, number of records in the Dataframes is equal to number of records in the respective Cassandra table.
  + Hive tables internally use RC file format for reliable storage and faster access of data.
  + No referential integrity is maintained in Cassandra Data Model.
* **Reliability**
  + Application is highly reliable with statistics monitored over Ambari.
  + Application is deployed on cloud. After application failure, it can be deployed to another (VM) within 2 hours.
  + Mean time between failures is almost 1 year.
* **Compatibility**
  + Ubuntu is compatible for Azure Cloud VM installation.
  + Hive is compatible with Spark and Cassandra installation with respective versions.
  + All third party applications can access through the authentication from Ambari.
* **Maintainability**
  + Development is adhered to coding standards.
  + Design is adhered to standards.
  + Code is well documented.
  + All integration is adhered to architecture standards.