SCIGILITY
SOLUTIONS FOR BIG DATA

Dear Candidate,

For this interview we have a technical challenge for you to work on. The goal is to see how far you can use technologies we are applying often at the customer's side, and/or how fast you can adapt base principles of new technologies.

Please prepare for the next interview the following tasks:

1. Provide a 3-node fully functioning Kafka cluster on either a public cloud provider of your choice or locally on your workstation. You can use any Kafka distribution such as Apache Kafka, Strimzi or the Confluent stack.

   You do not need to use additional tools like e.g SchemaRegistry/REST-Proxy etc. just focus on the base functionality.

   Goal is to deploy your own Kafka cluster on your infrastructure of choice. Hence do not use any SaaS offering like Confluent-Cloud, AWS MSK or Azure Event Hub.

2. The deployment and setup of the Kafka platform should be as much as possible automated.

3. Configure Kafka Brokers to listen to both
   - PLAINTEXT (unauthenticated) and
   - one of the authenticated listeners
     - Kerberos, or
     - SSL certificates

   If time allows you can also implement both of them. For Kerberos it is sufficient to use a local installation of e.g. MIT-Kerberos.

   Communication between the brokers themselves shall be SSL authenticated.

4. Create a test-topic and start a console-producer and a console-consumer (or a consumer/producer in a language of your choice)

   - one connected to PLAINTEXT (unauthenticated)
   - the other connected to the listener which requires authentication (either SSL and/or Kerberos)

   Start the Kafka clients directly from within the cluster to produce and consume some test messages.

5. How would you monitor the Kafka cluster in production? This is a conceptual question which you can outline during the interview. Hence nothing to implement here.

Enjoy and good luck !