**KAFKA**

1) Em linhas gerais, cite uma característica do Kafka que demonstre sua importância no contexto da CSD?

O sistema do Kafka é um serviço pub/sub, solução para plataformas que operam uma grande quantidade de troca de dados, de diferentes temas e entre diversos usuários com diferentes interesses. Essa é a realidade do atual número de transações financeiras, com volume de dados gigantesco e tipos de transações variados, seja a nível nacional ou internacional. Portanto, é coerente que a CSD assimile o Kafka para encarar essa demanda nacional de troca de dados do mercado financeiro. Além disso, o Kafka irá permitir que a empresa englobe cada vez mais atribuições, serviços e usuários com muito menos problemas se tivesse aderido a outra abordagem.

2) Levando em conta a arquitetura dirigida a eventos, pode-se dizer que o Kafka apresenta um perfil passivo em sua arquitetura. Disserte sobre essa afirmação.

Diz-se que o Kafka possui um perfil passivo porque a troca de dados pelo sistema é condicionada a requesições dos usuários ou ao envio de dados pelos produtores.

3) Cite uma vantagem sobre os dados do Kafka serem persistidos em disco?

Uma das vantagens é que os dados não são perdidos caso haja erros de conexão ou falhas de hardware ou software dos usuários.

4) Disserte sobre os seguintes elementos da Arquitetura Kafka: Tópico, Partição e Offset. Estabeleça relações entre estes conceitos em sua resposta.

Um tópico é similar a uma tabela em um banco de dados ou um diretório de um sistema. Um tópico armazena mensagens que possuem um assunto em comum. Tópicos são identificados por um nome e são composto por partições.

Partições são compostas pelas unidades que armazenam as mensagens que são ordenadas em série dentro da partição. É a maneira como o Kafka consegue escabilidade e redundância, pois é possível criar cópias de partição em outros brukers, inclusive em diferentes servidores. Elas auxiliam na rapidez da leitura do tópico de maneira paralela por diferentes consumidores.

Offsets são numeros inteiros inseridos nas unidades da partição junto com a mensagem e chave (opcionalemte). Ele é usado para ordenar os dados dentro de uma partição e permite que um usuário possa parar e reiniciar o acesso de onde parou.

5) Dê uma breve descrição sobre os seguintes conceitos:  
 Broker / Consumer Group / Bootstrap Broker / Líder de partição / Política de retenção.

Bruker: servidor Kafka, que pode ser virtual ou físico.

Consumer group: é um grupo de consumidores que são agrupados para acessar um tópico em comum. Dessa forma é garantido que cada partição vai ser consumida por apenas um dos membros do grupo.

Bootstrap broker: também chamado de broker controlador, ele é um dos broker de um cluster de brukers. Ele é eleito de forma automatica dos brukers ativos e possui funções extras como designar partições e monitorar falhas para os demais.

Líder de partição: é o bruker o qual partição é pertencente e do produtores e consumidores devem se conectar para ter acesso as mensagem daquela partição. A partir desse líder, cópias das partições são geradas.

Política de restrição: é o tempo de vida das mensagens em uma partição, que pode ser de minutos ou indeterminado.

6) Suponha a seguinte situação: em um determinado momento no tempo, guardamos Strings (valores) no Kafka. Posteriormente, é desejada a seguinte mudança: JSON’s sejam guardados ao invés das Strings Cite uma preocupação decorrente da migração de uma forma para outra.

O Kafka não se preocupa com os formatos ao armazenar os dados, pois faz isso utilizando dados em código binário. Portanto, não haveria preocupação em relação a armazenagem. Embora estejam disponíveis no Kafta tanto as opções para conversão de stings quanto JSON’s, esses formatos possuem schemaIDs distintos. Além disso, um JSON pode ou não ter um schema asociado no resultado da conversão. Por isso, para esse caso do exercício é necessário verificar se o worker (ou conector) esta configurado corretamente para strings.

7) O Kafka é conhecido como uma plataforma altamente escalável. Quais as preocupações que um desenvolvedor deve considerar ao modelar e produzir sua solução para se aproveitar dessa característica?

Ao configurar o número de consumidores, o desenvolvedor deve estar atento ao consumo do respectivo tópico. Se a demanda for alta, deve-se elevar o numero de consumidores. Se for baixa, alguns consumidores podem ficar ociosos e por isso deve-se então diminuir o número de consumidores. O ideal é que se mantenha um paralelismo entre consumidores e partições de um tópico, ou seja, numa proporção de 1:1.

8) Cite duas vantagens na utilização do Kafka Streams em uma aplicação. Lembre-se: há diferenciais do Kafka Streams com relação ao seu “irmão maior”?

O Kafka stream permite monitorar dados trocados com o Kafka, permitindo que seja detectados fraudes. Além disso, processa mensagens em modo exactly-oce, permitindo remoção de duplicatas e perda de dados.