INFOSOFT IT SOLUTIONS

Training | Projects | Placements

Revathi Apartments, Ameerpet, 1st Floor, Opposite Annapurna Block, Info

soft it solutions Software Training& Development 905968394,918254087

KAFKA STREAMS TRAINING

1: Introduction to Kafka Streams

- Overview of Kafka Streams
- Key concepts and terminology
- Use cases and benefits of Kafka Streams
- Comparison with other stream processing frameworks

2: Kafka Streams Architecture

- Architecture components (Stream Processor, State Stores, etc.)
- Processor Topology and DAG (Directed Acyclic Graph)
- Kafka Streams Threads and Tasks
- Integration with Kafka (brokers, topics, partitions)

3: Streams and Tables in Kafka Streams

- Understanding Streams vs. Tables
- Time semantics (event-time vs. processing-time)
- Windowing and session windows

4: Stateful Processing

- Introduction to state stores
- Types of state stores (in-memory, persistent)
- Fault tolerance and state restoration
- Handling out-of-order events

5: Kafka Streams DSL (Domain Specific Language)

- Overview of DSL operations
- Transformations (map, flatMap, filter, etc.)
- Aggregations (count, reduce, aggregate)
- Joins (inner join, outer join, left join)

6: Processor API

- Low-level API for fine-grained stream processing
- Processor Topology and custom processors
- Integration with external systems

7: Time Handling and Windowing

- Windowed operations (tumbling, hopping, sliding windows)
- Event-time processing
- Handling late data and watermarks

8: Fault Tolerance and Exactly Once Semantics

- Kafka's transactional guarantees
- Configuring Kafka Streams applications for fault tolerance
- Exactly-once processing semantics

9: Testing and Debugging Kafka Streams Applications

- Unit testing and integration testing
- Monitoring and debugging techniques
- Using Kafka Connect for data integration

ADVANCE TOPICS ;-

1: Advanced Stream Processing Concepts

- **Exactly-Once Semantics**
  - Understanding Kafka transactional guarantees
  - Implementing exactly-once processing patterns
  - Handling idempotency and transactional state

- **State Stores Optimization**
  - Advanced state store configurations (e.g., RocksDB settings)
  - Caching strategies for improving performance
  - State store scaling and partitioning
- **Interactive Queries**
  - Implementing interactive queries for real-time access
  - Caching strategies and query optimization
  - Building interactive applications using state stores

2: Advanced DSL Operations

- **Custom Processors and Transformers**
  - Implementing custom processor APIs
  - Handling complex transformations and aggregations
  - Integrating external systems and APIs
- **Global KTables**
  - Use cases and benefits of global tables
  - Implementing global KTables for global lookups
  - Managing global state and consistency

3: Scaling and Fault Tolerance

- **Scaling Kafka Streams Applications**
  - Horizontal scaling strategies
  - Partitioning and repartitioning topics
  - Load balancing and elasticity
- **Handling State Restarts**
  - Techniques for efficient state recovery
  - State migration strategies
  - Reducing downtime during application upgrades

4: Performance Optimization

- **Optimizing Kafka Configurations**
  - Tuning Kafka and Kafka Streams parameters
  - Configuring buffer sizes and timeouts
  - Network optimizations and throughput tuning
- **Serialization and Deserialization**
  - Custom Serdes implementation
  - Efficient data serialization techniques (e.g., Avro, Protobuf)
  - Schema evolution and compatibility
- **Monitoring and Metrics**
  - Utilizing Kafka Streams metrics
  - Monitoring application performance
  - Implementing custom monitoring solutions

5: Deployment Strategies

- **Cloud Deployment**
  - Deploying Kafka Streams on cloud platforms (AWS, Azure, GCP)
  - Kubernetes deployment strategies
  - Containerization and orchestration

- **Integration Patterns**
  - Best practices for integrating Kafka Streams with other systems (databases, microservices)
  - Building resilient and fault-tolerant pipelines
  - Data consistency and transactional integrity

6: Advanced Topics and Use Cases

- **Real-time Analytics and Machine Learning Integration**
  - Building real-time analytics pipelines
  - Integrating machine learning models with Kafka Streams
  - Handling complex event processing (CEP)
- **Security and Authentication**
  - Securing Kafka Streams applications
  - Integration with authentication systems (OAuth, JWT)
  - Role-based access control (RBAC)