

Apache Kafka Essentials (2 Days)

Day 1

Kafka Fundamentals and Client APIs

Morning Module

Part 1: Kafka Overview

- Kafka ecosystem
- Use cases
- Messaging concepts
- Topics and partitions
- Replication
- KRaft overview

Part 2: Producers

- Producer architecture
- Partition selection
- Keys and ordering
- Acknowledgements
- Retries
- Batching and compression
- Producer troubleshooting

Part 3: Consumers

- Consumer groups
- Offsets
- Rebalancing
- Group coordinator
- New Kafka 4.0 consumer protocol
- Positioning consumers
- Consumer troubleshooting

Part 4: Storage

- Retention
- Segment files
- Compaction
- State vs event topics
- GDPR considerations

Lab

- Create topics
- Produce and consume records
- Observe partition assignment
- Offset management

Kafka Architecture, Schema Registry and Security

Afternoon Module

Part 1: Kafka in Context

- When Kafka fits
- Anti-patterns
- Kafka vs Database
- Kafka vs RabbitMQ
- Kafka vs Pulsar
- Kafka Queues (Kafka 4.0)

Part 2: Schema Registry

- Why schemas matter
- Avro fundamentals
- Schema Registry architecture
- Compatibility modes
- Schema evolution
- Broker-side schema validation

Part 3: Cluster Architecture

- ZooKeeper architecture
- Controller election
- KRaft architecture
- KRaft migration process
- Listener configuration

Part 4: Security

- TLS/SSL
- SASL mechanisms
- ACL authorization
- End-to-end encryption

Lab

- Avro producer/consumer
- Schema evolution exercise

Day 2

Reliability, Performance and Kafka Connect

Morning Module

Part 1: Reliability

- Replication internals
- ISR
- Acknowledgement strategies
- Idempotent producers
- Transactions
- Exactly-once processing

Part 2: Performance Tuning

- Throughput optimization
- Latency optimization
- Durability optimization
- Availability optimization

Part 3: Kafka Connect

- Connect architecture
- Workers and tasks
- Converters
- SMTs
- JDBC connectors
- Sink connectors
- Debezium CDC

Part 4: REST Proxy

- Architecture
- Use cases
- Trade-offs

Lab

- JDBC Source Connector
- JDBC Sink Connector

Multi-Cluster Architectures, Monitoring and Stream Processing

Afternoon Module

Part 1: Multi-Cluster Kafka

- Active-Passive
- Active-Active
- Stretched Cluster
- RPO/RTO concepts

Part 2: Monitoring

- Broker metrics
- Producer metrics
- Consumer metrics
- JMX
- Prometheus
- Cruise Control
- Kafka Monitor
- Trogdor

Part 3: Stream Processing Concepts

- Event-driven analytics
- Stateless transformations
- Stateful transformations
- Event time
- Windowing

Part 4: Kafka Streams

- Streams architecture
- KStream
- KTable
- State stores
- Joins
- Aggregations
- Scaling and parallelism