INFOSOFT IT SOLUTIONS

Training | Projects | Placements

APACHE FLINK

Introduction to Apache Flink

- Overview of Apache Flink and its evolution
- Comparison with other stream processing frameworks (e.g., Apache Spark, Apache Storm)
- Use cases and scenarios suitable for Apache Flink

Apache Flink Architecture

- Understanding Flink's architecture (JobManager, TaskManager, JobGraph)
- Execution model and data flow in Flink
- Fault tolerance and checkpointing mechanisms

Flink Data Streaming Basics

- Introduction to data streams and data transformations
- Windowing concepts (time-based, count-based)
- Event time processing and watermarks

Flink APIs

- Overview of Flink APIs (DataStream API, DataSet API)
- Writing and deploying Flink applications
- Key transformations (map, flatMap, filter, reduce, etc.)

Stateful Stream Processing

- Introduction to stateful computations in Flink
- Managing state with KeyedState and OperatorState
- State backend configurations and tuning

Flink Connectors

- Working with Flink connectors (Kafka, Apache Cassandra, Elasticsearch, etc.)
- Customizing connectors and handling data sources and sinks
- Using Table API and SQL for data integration

Advanced Flink Topics

- Exactly-once processing semantics
- Dynamic scaling and resource management
- Handling late data and out-of-order events
- Flink's integration with Apache Beam

Monitoring and Operations

- Monitoring Flink applications (Web UI, metrics)
- Logging and debugging Flink jobs
- Configuration management and best practices

Performance Optimization

- Tuning Flink applications for better performance
- Memory management and JVM options
- Optimizing parallelism and throughput

Real-time Use Cases and Case Studies

- Reviewing real-world applications of Apache Flink
- Case studies from various industries (e.g., finance, telecommunications)
- Lessons learned and best practices from deployments

Flink Ecosystem and Extensions

- Overview of Flink's ecosystem (FlinkML, FlinkCEP, etc.)
- Exploring Flink extensions and community contributions
- Integrating Flink with Apache Hadoop and other data processing frameworks