



Apache Flink Developer Training

Level: Beginner

Description

This course is a hands-on introduction to Apache Flink for Java and Scala developers who want to learn to build streaming applications. The presentations will focus on the core concepts of distributed streaming dataflows, event time, and key-partitioned state. The exercises will give you a chance to see how these concepts are reflected in the API, and to understand how the pieces fit together to solve real problems.

Course Content

- Introduction to Stream Processing and Apache Flink
- Foundations of the DataStream API
- Getting Setup for Flink Development (incl. exercise)
- Stateful Stream Processing (incl. exercise)
- Time, Timers, and ProcessFunctions (incl. exercise)
- Testing (incl. exercise)
- Fault Tolerance

Prerequisites

No prior knowledge of Apache Flink is required.

You will need your own laptop with these tools installed:

- Git
- Java 8 JDK (a JRE is not sufficient, and newer versions of Java will not work)
- Maven 3.x
- An IDE for Java (or Scala) development

Any laptop capable of running an IDE, such as IntelliJ (MacOS, Linux, or Windows), should be fine.

To save time during the event, we would also like you to get Flink running before you come to the training. We will send an email with detailed setup instructions a week or two before the conference.



Apache Flink Runtime & Operations Training

Level: Beginner

Description

This course is a hands-on introduction to key topics relating to putting Flink applications into production, including configuring and tuning your job, deployment, operations, and maintenance. Also included is a solid introduction to the organization of the Flink runtime to help you tune and troubleshoot your applications. The intended audience includes both developers and operations staff. The exercises will provide a hands-on introduction to these same topics.

Course Content

- Runtime Architecture
- Fault Tolerance + hands-on
- Scaling: state backends and checkpointing + hands-on
- Capacity Planning
- Managing Flink
- Metrics, Monitoring, and Alerting + hands-on
- Deploying Flink
- Application Evolution + hands-on
- Network stack: latency and backpressure + hands-on

Prerequisites

No prior knowledge of Apache Flink is required, but the course assumes familiarity with issues arising in distributed systems.

You will need a notebook with at least 8 GB RAM and Docker installed.

We will send an email with more detailed setup instructions a week or two before the conference.



Apache Flink SQL Developer Training

Level: Beginner

Description

Apache Flink supports SQL as a unified ecosystem for stream and batch processing. Flink SQL can be used for a variety of use cases and leads to solutions that are easier to build and to maintain than applications built with Flink's lower-level APIs.

In this hands-on training you will learn how to create end-to-end SQL pipelines and leverage the potential of running SQL queries on data streams with Apache Flink. We will look at different use cases for streaming SQL, including enriching and joining streaming data, computing windowed aggregations, maintaining materialized views, and defining and detecting patterns using the `MATCH_RECOGNIZE` clause.

Course Content

- Basics of Apache Flink
- Introduction to Flink SQL
- Connecting to External Systems & Ecosystem
- Querying Dynamic Tables
- Joining Dynamic Tables
- Detecting Patterns using `MATCH_RECOGNIZE`

Prerequisites

No prior knowledge of Apache Flink is required.

Basic knowledge of SQL is required.

You will need a notebook with at least 8 GB RAM and Docker installed.

To save time during the event, we would also like you set up the training environment beforehand, i.e. to download all required Docker containers. We will send an email with detailed setup instructions about a week before the conference.



Apache Flink Tuning & Troubleshooting

Level: Advanced

Description

Working with numerous Flink users over the last years, we have learned a lot about the most common challenges when bringing a streaming application from early PoC stages up into production. In this training, we will focus on eliminating a few of these challenges together. We will provide a starting point for a useful troubleshooting toolset and present best practices, tips, and tricks in fields such as monitoring, watermarking, serialization, state back-ends and more. During hands-on sessions between the talks, participants will have the opportunity to apply the newly acquired knowledge to tackle some of the usual suspects in the context of an ill-performing Flink job: We will identify reasons for jobs not making progress or not performing as expected in either throughput or latency.

Prerequisites

In order to make the most of this training, you should already have Flink knowledge equivalent to our two-day developer training. In particular, you should be familiar with:

- time and watermarking,
- state handling and state backends,
- Flink's fault-tolerance guarantees,
- checkpoints and savepoints,
- DataStream API and ProcessFunction.

You will need your own laptop with these tools installed:

- Git
- Java 8 JDK (a JRE is not sufficient, and newer versions of Java will not work)
- Maven 3.x
- an IDE for Java development

Any laptop capable of running an IDE, such as IntelliJ (MacOS, Linux, or Windows), should be fine.

To save time during the event, we would also like you to get Flink running before you come to the training. We will send an email with detailed setup instructions a week or two before the conference.