



APACHE STORM

Course description

Apache Storm is a distributed stream-processing computation engine that allows the user to carry out real-time big-data processing. Like Hadoop, it executes mapping and reducing functions, but the data it processes is in a stream and not in big batches. Storm can process information at the rate of over a million records per second per node in an average-sized cluster, yet it is simple and, reportedly, fun. Storm is fault tolerant, restarting workers that die, and restarting workers on another node if a node dies. It is written in Java and Closure, but its applications can be written in any programming language. It is designed as a topology that acts as a data transformation pipeline. It is critically useful in a number of cases, such as real-time analytics, continuous computation, and online machine learning.

Student Take away

- Study Material
- Learning stuff
- Sample project for practice

APACHE STORM online training curriculum

Introduction & Getting Started

- What is Apache Storm?

Apache Storm vs Apache Hadoop

- Basics of Linux
- Basics of Java

Setting up storm cluster

- Test cluster setup
- Production ready cluster setup

Storm UI

- Overview on Storm UI

Apache Storm- (Advanced)

- Spouts & Bolts
- Interface method explanation

Storm Topologies

Types of Groupings

- Guaranteed Message Processing

Overview of Trident

Apache-Kafka installation

- Kafka Spout
- Kafka-storm

Apache Storm Practical with real world examples

BISP