

Apache Flink

Aplicaciones de procesamiento de datos en streaming



Bienvenidos al curso **Apache Flink Aplicaciones de procesamiento de datos en streaming**. Encontrarás esta guía en el temario del curso para que puedas consultarla en cualquier momento.

N.º Horas: 30 Fecha Inicio: X Fecha Fin: X

Objetivos del curso

Hola y bienvenido a este curso de Apache Flink. En este curso guiado aprenderás paso a paso y desde cero a trabajar con Apache Flink, una tecnología open source que te permitirá procesar datos en tiempo real de forma distribuida y escalable con muy baja latencia.



El curso contendrá partes teóricas, con las que aprenderás los conceptos clave de Apache Flink como su arquitectura, componentes y funcionamiento. El curso también tiene partes totalmente prácticas en las que aprenderás a usar directamente la herramienta y a implementar aplicaciones de procesamiento en streaming.

Estas lecciones prácticas se encuentran intercaladas con las lecciones teóricas para que te resulte fácil de seguir. También aprenderás a instalar todos los componentes del sistema y a usar las herramientas disponibles.

El curso contiene ejemplos de desarrollo de aplicaciones en streaming que interactúan con fuentes y destinos de datos con otras tecnologías. En todas las lecciones prácticas se adjunta el código usado para que te resulte más sencillo seguir las explicaciones y las pruebas.

Por último, el curso incluye un sencillo proyecto de integración con Apache Kafka y Elasticsearch en el que pondremos en uso los conocimientos que hemos aprendido.

Para seguir el curso es recomendable tener conocimientos de shell de linux y del lenguaje de programación java o scala. Este es un curso fundamental para quien desee implementar aplicaciones de procesamiento de datos en tiempo real.

¿Para quién es este curso?

- Interesados en implementar sistemas de procesamiento en tiempo real
- Interesados en aprender a implementar aplicaciones con Apache Kafka

Requisitos

Para seguir el curso es recomendable tener conocimientos de shell de linux y del lenguaje de programación java o scala. Este es un curso fundamental para quien desee implementar aplicaciones de procesamiento de datos en tiempo real.

- Instalación de Java
- Conocimientos básicos de sistemas distribuidos
- Conocimientos básicos de Java o Scala
- Conocimientos básicos de la shell de Linux

Plataforma de formación

Cada alumno ha recibido un email con los datos de acceso a la plataforma. Además de a este curso todos los alumnos matriculados con TrainingIT tienen

acceso al curso "**Conociendo la plataforma TrainingIT**". Este curso está conformado por una serie de videos cortos donde se explican las principales funcionalidades de la plataforma.

Resumen de contenidos

El curso está dividido en **6 módulos** temáticos. Puedes consultar el temario al final de esta guía.

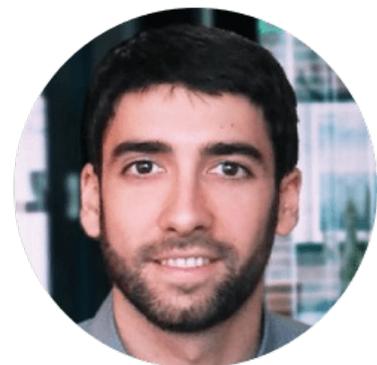
En cada módulo puede encontrar diferentes elementos:

- Apuntes de la lección**
- Videos teórico-prácticos**
- Código de la lección**
- **Test**
- **Otros recursos**

Además, el curso cuenta con un **ejercicio práctico voluntario** que se debe desarrollar durante el curso y que el profesor corregirá y evaluará al finalizar el curso.

Profesor

Hola, Soy **Óscar Fernández**, ¡me alegro de que estés aquí ya que eso significa que te interesa o que te apasiona el Big Data!



Soy un ingeniero de datos que vive en Madrid.

Me dedico a ayudar a organizaciones a sacar partido de las nuevas tecnologías relacionadas con **Big Data y Cloud**. Estudié en la Universidad Politécnica de Madrid Ingeniería Informática y posteriormente me especialicé en Sistemas Distribuidos y Big Data.

Siempre he pensado que la tecnología tiene un impacto enorme en la sociedad y en el mundo entero. No parece que nada de esto vaya a detenerse, y toda esta transformación en el tratamiento de los datos actuará como un soporte para el desarrollo de avanzada inteligencia artificial en el futuro.

Enrique Pascual y **Beatriz Pino**, actuarán como gestores del curso en todo lo relacionado con la plataforma de formación. Te puedes poner en contacto vía mensaje a través de la plataforma, o directamente en epascual@trainingit.es o bpino@trainingit.es.

A todos ellos los puedes encontrar en la sección “**Participantes**” del curso con sus respectivos roles.

Calendario del curso

El curso comienza el X y la fecha de finalización será el X último día en que se podrán entregar el ejercicio del curso.

A partir de esa fecha los profesores dejarán de participar en los foros del curso.

Si no lo terminas a tiempo, seguirás teniendo acceso a la plataforma con todo el material disponible. Pero no tendrás a los profesores a tu disposición, aunque sí a otros alumnos que puedan estar en tu misma situación.

Al tratarse de un curso online no existe una programación semanal de temas que tengáis que realizar. Cada alumno puede ir a su propio ritmo. Te aconsejo que trates de avanzar día a día, semana a semana y no lo vayas dejando todo para el final.

La práctica final voluntaria la puedes entregar hasta el último día del curso.

Desde TrainingIT estaremos al tanto de tu progreso en el curso y te avisaremos si vemos que hay algún problema.

Consejos para aprovechar tu curso

1. Lee toda la documentación relacionada con el curso

Antes de comenzar, dedica tiempo a revisar toda la información proporcionada: guías, temarios, requisitos técnicos, y objetivos del curso. Esto te dará una visión clara de lo que se espera de ti y cómo organizarte.

2. Utiliza un segundo monitor

Si es posible, configura un segundo monitor. Esto te permitirá seguir las clases en uno mientras practicas o tomas apuntes en el otro, optimizando tu experiencia de aprendizaje.

3. Revisa los temas y sus elementos antes de empezar

Antes de ver las clases, explora los materiales asociados a cada tema: videos, lecturas, ejercicios, o recursos adicionales. Esto te ayudará a identificar conceptos clave y planificar tu aprendizaje.

4. Haz un primer visionado de los videos

Visualiza los videos del curso de manera inicial, prestando atención al contenido global. No te preocupes por detenerte en cada detalle; el objetivo es obtener una comprensión general del tema.

5. Realiza un segundo visionado siguiendo los pasos del profesor

En el segundo visionado, pausa el video según sea necesario y sigue los pasos del instructor. Reproduce las explicaciones en tu entorno, asegurándote de comprender y aplicar los conceptos explicados.

6. Practica lo aprendido con otros casos y busca información adicional

No te limes a los ejercicios del curso. Aplica los conocimientos adquiridos en casos prácticos nuevos. Además, complementa tu aprendizaje investigando en internet, explorando foros, blogs, y tutoriales relacionados con los temas del curso.

7. Participa activamente en el foro del curso

El foro del curso es una herramienta fundamental para resolver dudas, compartir experiencias y aprender en comunidad. Úsalo para plantear preguntas claras y específicas sobre los temas del curso, mencionando el módulo o ejercicio en cuestión para facilitar la respuesta del profesor. Antes de preguntar, revisa si alguien ya ha planteado una duda similar, y si tienes conocimientos o experiencias útiles, comparte tus propias respuestas para contribuir al aprendizaje de tus compañeros. Además, aprovecha el feedback del profesor, que estará disponible para aclarar conceptos y guiarte en tu progreso, enriqueciendo aún más la experiencia formativa.

Tutorías

El "**Foro de Dudas y Consultas**" es el centro de las tutorías del curso. Todas tus dudas relacionadas con el temario debes plantearlas en el foro ya que también ayudarán a otros compañeros, y otros compañeros podrán ayudarte a ti.

Al tratarse de un foro no existe un horario de tutorías. Se responderá en un plazo de 24 horas laborables. **Utilizad siempre el foro.**

Enrique Pascual y **Beatriz Pino**, actuarán como gestores del curso en todo lo relacionado con la plataforma de formación. Te puedes poner en contacto vía mensaje a través de la plataforma, o directamente en **epascual@trainingit.es** o **bpino@trainingit.es**.

Práctica final voluntaria y examen teórico

El curso cuenta con un **ejercicio práctico final voluntario** que será corregido por el profesor.

Durante el curso se realizarán varias **pruebas de conocimientos obligatorias** mediante un examen de autoevaluación con preguntas tipo test.

Evaluaciones

En el curso habrá dos áreas de evaluación cada una con una influencia en la nota final del curso:

- **Práctica final: 40%**
- **Examen teórico: 60%**

Las calificaciones se podrán visualizar en el apartado “**Calificaciones**” del curso.

Certificado

El curso ofrece dos tipos de certificados:

- **Certificado de asistencia:** para aquellos alumnos que realicen el curso, pero no entreguen el ejercicio práctico voluntario.
- **Diploma acreditativo:** con evaluación positiva en las actividades realizadas en el curso para aquellos alumnos que realicen el ejercicio práctico y cuya calificación final de curso (Test+ Práctica) sea superior a 6.

Los certificados se entregarán una vez finalice el plazo del curso y se revisen que se cumplen todas las condiciones.

Bonificación

Para la evaluación y seguimiento del alumnado, la plataforma de Teleformación provee dos tipos de mecanismos distintos:

Sistemas de control internos del Aula Virtual. La plataforma registra la actividad del alumno dentro de la plataforma, obteniendo datos como:

- **Acceso del alumno a los distintos módulos del curso**

- **Días de acceso y clics realizados dentro del curso**
- **Tiempo total empleado en el curso**
- **Seguimiento de las lecciones visualizadas**
- **Seguimiento de los recursos utilizados**
- **Evaluación obtenida en cuestionarios de autoevaluación**

Los alumnos que vayan a bonificar el curso a través de sus empresas deberán cumplir lo siguiente:

- Completar al menos el **75% de las actividades** del curso.
- Completar al menos el **75% de las pruebas** obligatorias del curso (test)
- Los **tiempos de conexión** a la plataforma deben ser iguales o superiores al **75%** del tiempo de estudio estimado del curso (el tiempo estimado de estudio se indica en la cabecera de esta guía y en la cabecera del propio curso).

Los alumnos pueden observar su progreso en el curso dentro de la plataforma en los bloques “**Dedicación al curso**” y “**Estado de Finalización**” que aparece en los menús laterales de la derecha de la pantalla.



The screenshot displays the course progress interface. On the left, there is a sidebar menu with two items: 'Introducción a Jira' (with a progress bar at 0%) and 'Test Prueba Certificación'. The main content area is divided into two sections:

- Dedicación al curso:** Shows the estimated time spent on the course as '4 horas 38 minutos', the last update as '10/01/2025 01:10', and a status of 'Informe completo'.
- Estado de Finalización:** Shows a progress bar at 0%, a play button icon, and the text '1.6 Resumen Global 0:01:58 No finalizado' with a red error icon. Below this is a button labeled 'Vista general de estudiantes'.

On the right side of the main content area, there is a button labeled 'Desplegar menú lateral'.

| Acerca de TrainingIT

TrainingIT es una iniciativa para ofrecer formación especializada IT de alta calidad y bonificable por Fundae.

Descubre nuestros cursos online creados por especialistas en sus materias en www.trainingit.es.

Queremos formar a los mejores profesionales para que no se diga que en España no hay talento.

Te agradecemos que hayas confiado en TrainingIT para tu formación. Esperamos que este curso sea de tu agrado y que te ayude en tu carrera profesional.

Un saludo,

Enrique Pascual

epascual@trainingit.com

Gestor de Curso

trainingIT


Temario

1. Introducción a Apache Flink

- 1.1 Bienvenida
- 1.2 Presentación del curso
- 1.3 Visión general de Apache Flink

2. Conceptos básicos de Apache Flink

- 2.1 ¿Qué es el procesamiento en Streaming?
- 2.2 Tecnologías de Streaming
- 2.3 Introducción a Apache Flink y Casos de Uso
- 2.4 Arquitectura y Componentes de Flink

3. Primeros pasos con Apache Flink

- 3.1 Instalación de Apache Flink
- 3.2 Instalación de Apache Flink en Windows
- 3.3 Estructura de código de un programa
- 3.4 Hola Mundo
- 3.5 Aprendiendo a usar la interfaz gráfica

4. Pipelines y ETL

- 4.1 Data stream API
- 4.2 Operadores
- 4.3 Estado en Flink
- 4.4 Ejemplo estado en Flink
- 4.5 Ventanas de procesamiento
- 4.6 Control temporal de eventos
- 4.7 Ejemplos control temporal de eventos

5. Analítica con Apache Flink

- 5.1 Garantías de entrega y transacciones
- 5.2 Checkpointing y Snapshots
- 5.3 API de tablas y SQL

6. Proyecto con Apache Flink

6.1 Introducción y Prerrequisitos

6.2 Integración con Apache Kafka (Fuente)

6.3 Integración con Elasticsearch (Sink)

6.4 Implementación y pruebas