

David Torres

Hadoop Developer

SUMMARY

Middle-level Hadoop engineer with 3.5 years of commercial experience across batch ETL, real-time ingestion, and cluster operations. Maintains a 48-node HDFS cluster processing 1.6 TB daily; migrated 11 MapReduce jobs to Spark, saving 6.5 hours of nightly compute. Leads platform documentation and junior engineer onboarding.

[Recruiter's Note - Hannah]: The summary opens with years of experience, a cluster-scale number, and a concrete migration outcome. That is the right density for a mid-level profile: enough to confirm seniority, short enough to read in 10 seconds.

EXPERIENCE

Middle Hadoop Developer

DataGrid Iberia

2023 - Present, Madrid

- Own HDFS and YARN operations for a 48-node cluster handling 1.6 TB daily ingestion from 7 upstream source systems
- Rewrote 23 Hive queries with partition pruning and ORC conversion, reducing dashboard latency from 14 minutes to 2.8 minutes
- Migrated 11 MapReduce batch jobs to Spark on YARN, cutting total nightly compute from 9 hours to 2.5
- Built Ambari alert rules and HDFS block scanner thresholds that reduced weekly under-replicated block count from 1,200 to under 30

[Recruiter's Note - Hannah]: Four bullets, each with a before-and-after metric or a production scale anchor. This is the minimum standard for a developer resume. Every line answers the question: what changed because of you?

Junior Data Engineer

StreamPath Solutions

2021 - 2023, Valencia

- Developed Sqoop import jobs for 14 relational source tables and managed incremental loads with watermark columns, achieving 99.97% row-level accuracy
- Built Airflow DAGs to schedule 36 daily ingestion and transformation pipelines, replacing manual cron scripts
- Assisted in Kafka consumer group setup for real-time clickstream data; processed peak 80K events per second with no consumer lag exceeding 2 seconds

[Recruiter's Note - Hannah]: The junior role shows clear progression: ingestion and scheduling in the first position, full cluster ownership in the second. Hiring managers look for that arc when evaluating a Hadoop admin resume for 3 years experience.

EDUCATION

B.Sc. Computer Engineering, Universidad Politecnica de Madrid, 2021

LANGUAGES

English Proficient ●●●●●●
Spanish Native ●●●●●●

d.torres@email.com

000 - 00 - 000 - 00

Madrid, Spain

linkedin.com/in/davidtorres-data

github.com/dtorres-Hadoop

TECHNICAL SKILLS

Hadoop Core:

HDFS, YARN, MapReduce, ZooKeeper

Query and Modeling:

Hive, HBase, Pig, Presto

Processing:

Apache Spark (PySpark, Scala API), Kafka, Flink basics

Orchestration:

Apache Oozie, Airflow, cron-based scheduling

Ingestion::

Sqoop, Flume, REST connectors, S3-to-HDFS pipelines

Ops and Monitoring:

Ambari, Grafana, Linux shell, log aggregation

Languages:

SQL, Python, Java, Bash

[Recruiter's Note - Hannah]: Skills grouped by function. The ops line signals that this candidate can manage the cluster, not just write jobs against it. That distinction matters when a team needs someone to cover both engineering and platform stability.

CERTIFICATIONS

- Cloudera Certified Associate: Data Analyst, 2024
- IBM Data Engineering Professional Certificate, 2023
- Databricks Certified Associate Developer for Apache Spark, 2023

[Recruiter's Note - Hannah]: Three certifications across analyst, engineering, and processing domains. That range confirms the candidate can work across the full Hadoop data stack, not just one layer of it.