



# Apache Flume: Distributed Log Collection for Hadoop - Second Edition

*Steve Hoffman*

Download now

[Click here](#) if your download doesn't start automatically

# Apache Flume: Distributed Log Collection for Hadoop - Second Edition

*Steve Hoffman*

**Apache Flume: Distributed Log Collection for Hadoop - Second Edition** Steve Hoffman  
**Design and implement a series of Flume agents to send streamed data into Hadoop**

## About This Book

- Construct a series of Flume agents using the Apache Flume service to efficiently collect, aggregate, and move large amounts of event data
- Configure failover paths and load balancing to remove single points of failure
- Use this step-by-step guide to stream logs from application servers to Hadoop's HDFS

## Who This Book Is For

If you are a Hadoop programmer who wants to learn about Flume to be able to move datasets into Hadoop in a timely and replicable manner, then this book is ideal for you. No prior knowledge about Apache Flume is necessary, but a basic knowledge of Hadoop and the Hadoop File System (HDFS) is assumed.

## What You Will Learn

- Understand the Flume architecture, and also how to download and install open source Flume from Apache
- Follow along a detailed example of transporting weblogs in Near Real Time (NRT) to Kibana/Elasticsearch and archival in HDFS
- Learn tips and tricks for transporting logs and data in your production environment
- Understand and configure the Hadoop File System (HDFS) Sink
- Use a morphline-backed Sink to feed data into Solr
- Create redundant data flows using sink groups
- Configure and use various sources to ingest data
- Inspect data records and move them between multiple destinations based on payload content
- Transform data en-route to Hadoop and monitor your data flows

## In Detail

Apache Flume is a distributed, reliable, and available service used to efficiently collect, aggregate, and move large amounts of log data. It is used to stream logs from application servers to HDFS for ad hoc analysis.

This book starts with an architectural overview of Flume and its logical components. It explores channels, sinks, and sink processors, followed by sources and channels. By the end of this book, you will be fully equipped to construct a series of Flume agents to dynamically transport your stream data and logs from your systems into Hadoop.

A step-by-step book that guides you through the architecture and components of Flume covering different approaches, which are then pulled together as a real-world, end-to-end use case, gradually going from the simplest to the most advanced features.

 [Download Apache Flume: Distributed Log Collection for Hadoo ...pdf](#)

 [Read Online Apache Flume: Distributed Log Collection for Had ...pdf](#)

## **Download and Read Free Online Apache Flume: Distributed Log Collection for Hadoop - Second Edition Steve Hoffman**

---

### **From reader reviews:**

#### **George Gomez:**

Book is to be different for each grade. Book for children till adult are different content. As you may know that book is very important for us. The book Apache Flume: Distributed Log Collection for Hadoop - Second Edition was making you to know about other knowledge and of course you can take more information. It is rather advantages for you. The reserve Apache Flume: Distributed Log Collection for Hadoop - Second Edition is not only giving you a lot more new information but also to get your friend when you experience bored. You can spend your spend time to read your publication. Try to make relationship using the book Apache Flume: Distributed Log Collection for Hadoop - Second Edition. You never sense lose out for everything should you read some books.

#### **Maria Hughes:**

As people who live in the particular modest era should be update about what going on or details even knowledge to make these individuals keep up with the era and that is always change and advance. Some of you maybe will update themselves by reading books. It is a good choice for yourself but the problems coming to a person is you don't know which you should start with. This Apache Flume: Distributed Log Collection for Hadoop - Second Edition is our recommendation to cause you to keep up with the world. Why, since this book serves what you want and wish in this era.

#### **David Gonzales:**

Reading a book can be one of a lot of exercise that everyone in the world really likes. Do you like reading book and so. There are a lot of reasons why people enjoyed. First reading a reserve will give you a lot of new info. When you read a guide you will get new information because book is one of numerous ways to share the information or even their idea. Second, reading a book will make you actually more imaginative. When you reading through a book especially tale fantasy book the author will bring you to definitely imagine the story how the character types do it anything. Third, you can share your knowledge to other individuals. When you read this Apache Flume: Distributed Log Collection for Hadoop - Second Edition, you could tells your family, friends as well as soon about yours reserve. Your knowledge can inspire average, make them reading a reserve.

#### **Donna Robinson:**

Within this era which is the greater person or who has ability in doing something more are more important than other. Do you want to become considered one of it? It is just simple method to have that. What you are related is just spending your time not much but quite enough to enjoy a look at some books. One of many books in the top collection in your reading list will be Apache Flume: Distributed Log Collection for Hadoop - Second Edition. This book that is qualified as The Hungry Hills can get you closer in getting precious person. By looking right up and review this publication you can get many advantages.

**Download and Read Online Apache Flume: Distributed Log  
Collection for Hadoop - Second Edition Steve Hoffman  
#G5QCUKZEMXO**

## **Read Apache Flume: Distributed Log Collection for Hadoop - Second Edition by Steve Hoffman for online ebook**

Apache Flume: Distributed Log Collection for Hadoop - Second Edition by Steve Hoffman Free PDF d0wnl0ad, audio books, books to read, good books to read, cheap books, good books, online books, books online, book reviews epub, read books online, books to read online, online library, greatbooks to read, PDF best books to read, top books to read Apache Flume: Distributed Log Collection for Hadoop - Second Edition by Steve Hoffman books to read online.

### **Online Apache Flume: Distributed Log Collection for Hadoop - Second Edition by Steve Hoffman ebook PDF download**

**Apache Flume: Distributed Log Collection for Hadoop - Second Edition by Steve Hoffman Doc**

**Apache Flume: Distributed Log Collection for Hadoop - Second Edition by Steve Hoffman Mobipocket**

**Apache Flume: Distributed Log Collection for Hadoop - Second Edition by Steve Hoffman EPub**