

aws marketplace

Apache NiFi

Reviews, tips, and advice from real users



Powered by  PeerSpot

Contents

Product Recap..... 3 - 4

Valuable Features..... 5 - 10

Other Solutions Considered..... 11

Use Case..... 12 - 15

Setup..... 16 - 18

Customer Service and Support..... 19 - 20

Other Advice..... 21 - 24

Trends..... 25 - 26

About PeerSpot..... 27 - 28

Product Recap



Apache NiFi

Apache NiFi Recap

Apache NiFi is an easy to use, powerful, and reliable system to process and distribute data. It supports powerful and scalable directed graphs of data routing, transformation, and system mediation logic.

Valuable Features

Excerpts from real customer reviews on PeerSpot:



“It is highly effective for handling real-time data by working with APIs for immediate and continuous data extraction.”



Bharghava Raghavendra Beesa

Senior Developer at Infosys



“The initial setup is very easy.”



Arjun Pandey

Engineering Lead- Cloud and Platform Architecture at a financial services firm with 1,001-5,000 employees



“The initial setup is very easy. I would rate my experience with the initial setup a ten out of ten, where one point is difficult, and ten points are easy.”



SabinaZeynalova

Data Engineer Team Lead at Unibank



“The user interface is good and makes it easy to design very popular workflows.”



Bruno_Silva

Partner at Timestamp, SA



“We can integrate the tool with other applications easily.”



Moulali Naguri

Project Engineer at Wipro Limited



“Apache NiFi is user-friendly. Its most valuable features for handling large volumes of data include its multitude of integrated endpoints and clients and the ability to create cron jobs to run tasks at regular intervals.”



Teodor Muraru

Developer at Emag



“It's an automated flow, where you can build a flow from source to destination, then do the transformation in between.”



Reza Gholami

Senior Technology Architect at a tech services company with 10,001+ employees

What users had to say about valuable features:

“I use custom functions for specific features in Apache NiFi. I also use the processes available in NiFi. I can write custom functions to achieve the desired functionality, even if it is not explicitly available as a built-in NiFi feature..”

SabinaZeynalova

Data Engineer Team Lead at Unibank

[Read full review](#) 

The visual workflow aspect of Apache NiFi is an invaluable feature as it operates on a no-code platform that allows for easy drag-and-drop pipeline construction. Compared to Airflow, which requires programming before visual representation, Apache NiFi offers clarity in pipeline activities. This feature greatly aids in understanding what the pipeline is doing.

Danuphan Suwanwong

Head of Data Engineering and AI Engineering at Coraline

[Read full review](#) 

“NiFi works on data and file levels, streamlining real-time data processes. It is highly effective for handling real-time data by working with APIs for immediate and continuous data extraction. For real-time data tasks, this front-end UI-based tool is superior to back-end platforms..”

Bharghava Raghavendra Beesa

Senior Developer at Infosys

[Read full review](#) 

“The user interface is good and makes it easy to design very popular workflows.

There are nice parameters for migration.

It is very easy to schedule jobs that realize improvements and monetize..”

Bruno_Silva

Partner at Timestamp, SA

[Read full review](#) 

“The most valuable feature has been the range of clients and the range of connectors that we could use. The processing ability of the solution is also really good because it allows high volumes of data to flow in real time.

Overall, this solution is reliable. .”

Aria Amini

Data Engineer at Behsazan Mellat

[Read full review](#) 

“The good thing about Apache NiFi is that it has a concept called a flow file, and there's something called a flow file processor. The processor is the building block of your entire job. They have close to 500 processors for each purpose.

For example, for reading from Kafka, Ni-Fi has a processor called "consumer Kafka". To write to S3, they have a processor called "put S3". Now, if I read from Kafka and write my own application, I'd need to ensure the library I'm using tracks my messages. I'd also need to handle any failures by rereading messages and ensuring acknowledgment. But all this complexity is already handled by Apache processor.

They have around 500 processors, with a community investing significant effort into developing them. I can design your processor with a single click, export the entire workflow, and import it. The format is actionable, so NiFi is immediately set up.

It's also distributed in nature so that I can scale it across nodes based on the workload. These nodes share their state. If one node goes down during processing, that data might be lost, but any subsequent data is safe. Such occurrences are rare.

In essence, if you want a quick solution, Apache NiFi is a strong contender. There are other solutions like AirFlow and some paid pipeline options.

AirFlow is open-source but can be complicated. For ETL or ERT solutions, there are pricier options. But if I need a pipeline that I can monitor step by step, Apache NiFi is a good choice. It integrates with Prometheus metrics, allowing me to embed them in my workflow.

There's also a processor for integration with Slack, and I can receive notifications when the workflow is completed or fails.

Another feature I appreciate is "back pressure," which NiFi handles automatically. It maintains its own queue and addresses back-pressure issues. If, for instance, an upstream entity isn't fast enough, items get stored in a queue, managed internally

by NiFi's back pressure algorithm..”

Arjun Pandey

Engineering Lead- Cloud and Platform Architecture at a financial services firm with 1,001-5,000 employees

[Read full review](#) 

Other Solutions Considered

“We have used similar solutions like Apache Airflow, StreamSets, and others. The ramp-up for NiFi is a lot bigger and you have to spend a lot of time to understand how to use it..”

Verified user

Senior Solutions Architect/ Software Architect at a comms service provider with 51-200 employees

[Read full review](#) 

“We were previously using Apache Flume for our data ingestion. Flume was a solid product, but it just did not offer the visual representations and full start-to-end processing that our organization was looking for..”

Verified user

Infrastructure Engineer at Zirous, Inc.

[Read full review](#) 

“We looked at some proprietary solutions and also, evaluated StreamSets. The proprietary solutions were expensive and often didn’t suit our use cases. StreamSets didn’t have the same level of adoption..”

Verified user

Software Engineer at a consultancy with 1,001-5,000 employees

[Read full review](#) 

Use Case

I am implementing the ETL workflow using Apache NiFi to prepare data and upload it to the cloud. Our use case involves importing data from on-premise and private servers to build a data hub and data mart. The data mart is then published on the cloud.

Danuphan Suwanwong

Head of Data Engineering and AI Engineering at Coraline

[Read full review](#) 

“Our company uses the solution to ingest raw data. We have five repositories with a huge amount of data. We normalize the data to previously structured files, prepare it, and ingest it to devices.

The size of any project team depends on the workflow or management activities but typically includes two to five users..”

Bruno_Silva

Partner at Timestamp, SA

[Read full review](#) 

“I use Apache NiFi to build workflows.

It's an event that is used for distributed messaging. You need to transfer the message that comes into Kafka Broker Topic.

You get the messages in the Kafka queue topic, you transform it and send it to other entities for storage, or you can return it back to Kafka to send to the consumer..”

Reza Gholami

[Read full review](#) 

Senior Technology Architect at a tech services company with 10,001+ employees

“As a DevOps engineer, my day-to-day task is to move files from one location to another, doing some transformation along the way. For example, I might pull messages from Kafka and put them into S3 buckets. Or I might move data from a GCS bucket to another location.

NiFi is really good for this because it has very good monitoring and metrics capabilities. When I design a pipeline in NiFi, I can see how much data is being processed, where it is at each stage, and what the total throughput is.

I can see all the metrics related to the complete pipeline. So, I personally like it very much..”

Arjun Pandey

[Read full review](#) 

Engineering Lead- Cloud and Platform Architecture at a financial services firm with 1,001-5,000 employees

“One example is how Apache NiFi has helped us to create data pipelines to migrate data from Oracle to Postgres, Oracle to Oracle, Oracle to Minio, or other databases, such as relational databases, NoSQL databases, or object storage. We create templates for these pipelines so that we can easily reuse them for different data migration projects.

For example, we have a template for migrating data from Oracle to Postgres. This template uses an incremental load process. The template also checks the source and destination databases for compatibility and makes any necessary data transformations.

If our data is not more than ten terabytes, then NiFi is mostly used. But for a heavy table setup, I don't use NiFi for customers or enterprise solutions..”

SabinaZeynalova

Data Engineer Team Lead at Unibank

[Read full review](#) 

“I use NiFi as a tool for ETL, which stands for extract, transform, and load. It is particularly effective for integration methodologies.

The tool is useful for designing ETL pipelines and is an open-source product. Data is often stored in different forms and locations. If I want to integrate and transform it, NiFi can help load data from one place to another while making transformations.

I can handle stream or batch data and identify various data types on different platforms. NiFi can integrate with tools like Slack and perform required transformations before loading to the desired downstream.

It is primarily a pipeline-building tool with a graphical UI, however, I can also write custom JARs for specific functions. NiFi is an open-source tool effective for data migration and transformations, helping improve data quality from various sources..”

Bharghava Raghavendra Beesa

Senior Developer at Infosys

[Read full review](#) 

Setup

The setup process involves configuring and preparing the product or service for use, which may include tasks such as installation, account creation, initial configuration, and troubleshooting any issues that may arise. Below you can find real user quotes about the setup process.

“The initial setup was straightforward. We can deploy the tool easily on a single node. It won’t take much time. If it is a multi-node cluster, it will take two to three hours..”

MoulaliNaguri

Project Engineer at Wipro Limited

[Read full review](#) 

“I tried to install the tool on my work laptop, and while it worked initially, it started to run slowly after some time. The department that handles the company's databases uses Apache NiFi on proper servers. I tried using it on my laptop to see if it worked, but it ran very slowly and consumed many resources from my machine..”

Teodor Muraru

Developer at Emag

[Read full review](#) 

“The initial setup is fast, especially for communication stabilization. Although the product is open source, it functions as a cluster. For single-node environments, installation is simple. For company-wide or enterprise-level clusters, the initial stages may present issues with authentication and access. Stabilization, such as port communication, may not be immediately effective..”

Bharghava Raghavendra Beesa

Senior Developer at Infosys

[Read full review](#) 

“The initial setup is very easy, especially for users who are familiar with EDL or EMT.

NiFi is one of the easiest tools on the market to learn and use. It is also a quick-win solution, which is good for first-time users who are developing data pipelines for EMT. NiFi makes it easy to track and trace the status of your pipelines, so you can be sure that they are working properly..”

Arjun Pandey

Engineering Lead- Cloud and Platform Architecture at a financial services firm with 1,001-5,000 employees

[Read full review](#) 

“The initial setup is very easy. I would rate my experience with the initial setup a ten out of ten, where one point is difficult, and ten points are easy.

But if you want its custom mode and control, it's five out of ten.

For the initial setup, if you configure to custom mode, it's five points. But if you use its single-mode configuration and installation, it's ten..”

SabinaZeynalova

Data Engineer Team Lead at Unibank

[Read full review](#) 

“You have to install it on your platform using the Windows and Linux Operating System, or any Mac OS.

The Apache documentation is very straightforward. You can do it independently of any other guidance.

It's self-explanatory.

You can do it yourself without any hiccups, It's straightforward..”

Reza Gholami

Senior Technology Architect at a tech services company with 10,001+ employees

[Read full review](#) 

Customer Service and Support

The technical support from the official Apache team is rated a three out of ten. Issues often require self-resolution or community help, as the support isn't effectively managed.

Danuphan Suwanwong

Head of Data Engineering and AI Engineering at Coraline

[Read full review](#) 

“We contacted technical support by phone once and they were helpful, but there was some delay with their response. We would have preferred a quicker response time.

After a bit of study, we can support most of our needs so don't really need technical support. .”

Bruno_Silva

Partner at Timestamp, SA

[Read full review](#) 

Customer Service:

NiFi is open source through Apache, customer service is at a minimum, but their website offers answers to most questions my team encountered.

Technical Support:

NiFi is open source through Apache, technical support is at a minimum, but their website offers answers to most questions my team encountered..”

Verified user

Infrastructure Engineer at Zirous, Inc.

[Read full review](#) 

“I have not needed Apache support. With my seniority, I have not used anything that requires their support.

It's a good thing to be able to go back in the documentation and the information is enough to get you going.

Juniors or people who have less experience may have a lot of different questions, but I have not tested their support..”

Reza Gholami

Senior Technology Architect at a tech services company with 10,001+ employees

[Read full review](#) 

Other Advice

“The solution must be improved to compete with Kafka. As it is an open-source tool, it will take time to get all the functions. I would recommend the product to others. Overall, I rate the product a seven out of ten..”

MoulaliNaguri

Project Engineer at Wipro Limited

[Read full review](#) 

“Think about your data flows as the directed graphs between low-level processing modules, so you can re-use as much of the path as possible for different data streams. Don’t create entirely separate flows for new data sources, i.e., if you can avoid it..”

Verified user

Software Engineer at a consultancy with 1,001-5,000 employees

[Read full review](#) 

“The architect needs to evaluate the entire architecture with this platform so eventually, we are left with our architects and we need to get approval from them to do that.

I would rate this solution an eight out of ten..”

Bruno_Silva

Partner at Timestamp, SA

[Read full review](#) 

“You have to independently develop a use case for using NiFi, then test it and do the integration to have a stable workflow. After this, you jump in and integrate with Kafka distributed messaging and also connecting to your MDM database.

I would rate Apache NiFi an eight out of ten..”

Reza Gholami

Senior Technology Architect at a tech services company with 10,001+ employees

[Read full review](#) 

“I recommend the product for its data privacy features. It allows secure data handling because the data is stored on my nodes. However, a skilled technician is necessary due to the reliance on Java, especially for back-end operations and error debugging.

Enterprise versions may offer easier troubleshooting. As an open-source solution, good support is crucial.

I rate the overall product as eight out of ten..”

Bharghava Raghavendra Beesa

Senior Developer at Infosys

[Read full review](#) 

“If I were to advise someone, I would ask the user what endpoints they want to touch. If I want to read something from Kafka and I want to put this thing on the S3 bucket, what is the alternative I have?

I have Kafka Connect, where I can connect Kafka with one Kafka, and I can put it into an S3 bucket. Is this scalable? No. Is this monitoring No.

We can't monitor it. We can't scale it. It's going to be a complete black box. The person who knows Kafka Connect, or Kafka, can understand what is happening there while using Kafka Connect. But if I compare it, I literally don't need to understand what Kafka is.

I know, "Okay, this is Kafka. These are the endpoints, and this is the URL I have to point to." That's it. My job is done. I will create a complete flow pipeline within, let's say, thirty minutes or something without having any current knowledge. I can read, I can Google it, and I can just implement it.

For people who are new to big data technologies like Kafka and BigQuery, I would give this solution an eight out of ten.

Let's say you need to build a solution to read from Kafka and write to an S3 bucket. You could use Kafka Connect, but if your requirements change and you need to start reading from a database instead, Kafka Connect will not work. With Apache NiFi, you can easily modify your flow pipeline to start reading from the database instead..”

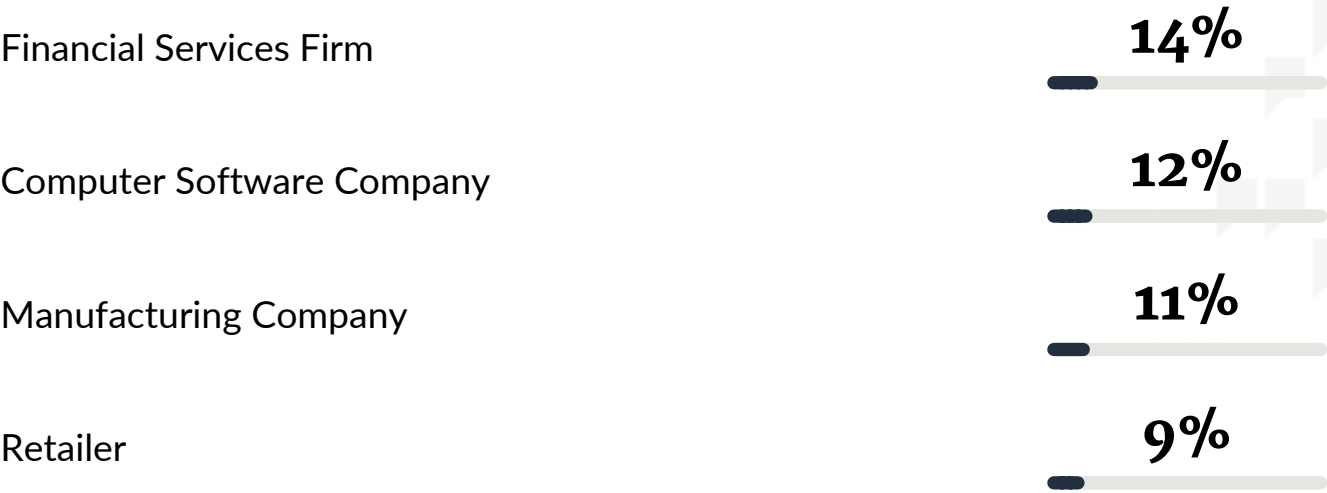
Arjun Pandey

Engineering Lead- Cloud and Platform Architecture at a financial services firm with 1,001-5,000 employees

[Read full review](#) 

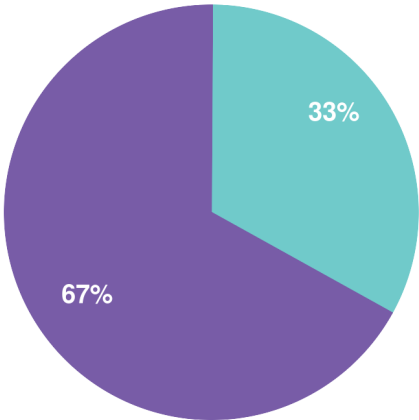
Top Industries

by visitors reading reviews

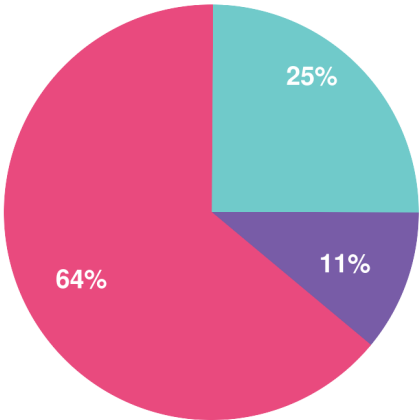


Company Size

by reviewers



by visitors reading reviews



Large Enterprise Midsize Enterprise Small Business

About this buyer's guide

Thanks for downloading this PeerSpot report.

The summaries, overviews and recaps in this report are all based on real user feedback and reviews collected by PeerSpot's team. Every reviewer on PeerSpot has been authenticated with our triple authentication process. This is done to ensure that every review provided is an unbiased review from a real user.

Get a custom version of this report... Personalized for you!

Please note that this is a generic report based on reviews and opinions from the collective PeerSpot community. We offer a [customized report](#) of solutions recommended for you based on:

- Your industry
- Company size
- Which solutions you're already considering

The customized report will include recommendations for you based on what other people like you are using and researching.

Answer a few questions in our short wizard to get your customized report.

[Get your personalized report here](#)

About PeerSpot

PeerSpot is the leading review site for software running on AWS and other platforms. We created PeerSpot to provide a trusted platform to share information about software, applications, and services. Since 2012, over 22 million people have used PeerSpot to choose the right software for their business.

PeerSpot helps tech professionals by providing:

- A list of products recommended by real users
- In-depth reviews, including pros and cons
- Specific information to help you choose the best vendor for your needs

Use PeerSpot to:

- Read and post reviews of products
- Access over 30,000 buyer's guides and comparison reports
- Request or share information about functionality, quality, and pricing

Join PeerSpot to connect with peers to help you:

- Get immediate answers to questions
- Validate vendor claims
- Exchange tips for getting the best deals with vendor

Visit PeerSpot: www.peerspot.com

PeerSpot

244 5th Avenue, Suite R-230 • New York, NY 10001

reports@peerspot.com

+1 646.328.1944