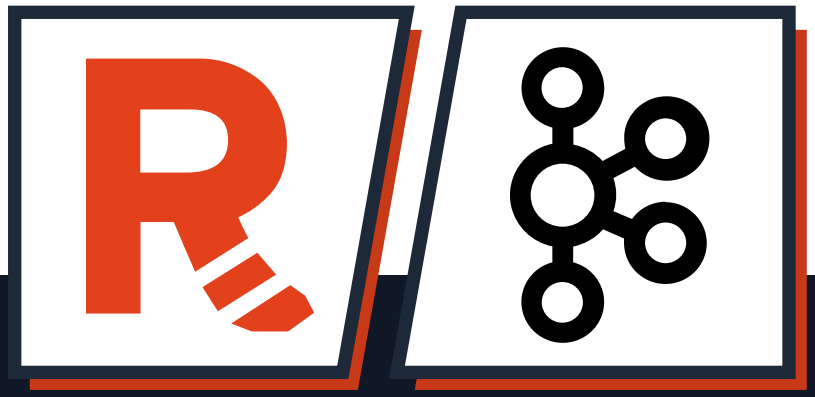
Redpanda

Redpanda vs. Kafka

A practical guide for leaders working with real-time data



If you're in the market for a real-time data streaming platform, your search likely led you to Apache Kafka® — and for good reason. Kafka is open source and widely used across the industry to build data streaming systems. However, Kafka's notorious complexity and Java-based infrastructure make it challenging (and expensive) for businesses trying to keep up with the increasing demand for real-time data and analytics.

Enter Redpanda, a powerful streaming data platform that offers all the reliability and convenience of Kafka APIs — without the Kafka complexity or limitations. Built in C++, Redpanda is leaner, faster, cost-efficient at scale, and radically simplifies building real-time data applications and AI pipelines so businesses can get up and running in no time.

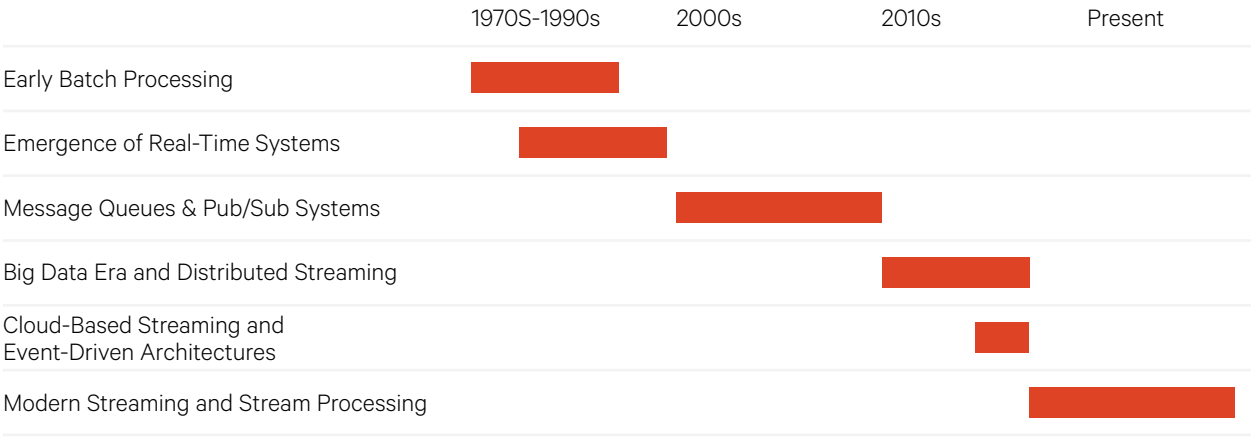
So, if you're considering Redpanda as a reliable Kafka alternative but don't want to wade through pages of documentation or talk to sales (yet), this guide gives you a practical summary of everything you need to know before taking the next step — from high-level differences to line-by-line comparisons.

The rise and plateau of Kafka

You can't make an informed decision without the proper context, so let's start with a quick tour of the data streaming landscape then and now.

From the early days of batch processing in the 1970s, data streaming technologies, as we know them today, have evolved in response to several key advancements over the past 50 years.

Evolution of data streaming



First, it was the emergence of real-time systems in the late 1990s and Pub/Sub systems in the early 2000s. The following decade saw the dawn of the Big Data era, which created an unprecedented need for distributed streaming technologies.

The birth of the modern cloud just a few years later introduced all-new possibilities for analytics, automation, and decision-making – along with an appetite for new technologies to make them a reality.

The dawn of distributed data streaming

In 2011, Kafka revolutionized data streaming as the first open-source distributed platform for high-throughput, fault-tolerant real-time data processing. It addressed major limitations in scalability and performance, enabling Big Data systems to handle growing volumes efficiently through horizontal scaling (i.e., adding more brokers).

Unlike proprietary solutions, Kafka offered a cost-effective, flexible alternative with robust community support. However, over 10 years later, Kafka faces new challenges as modern demands for faster, more reliable, and simpler streaming solutions expose its limitations.

The cracks in Kafka's infrastructure

In a time of high-performance, high-throughput, low-latency applications, Kafka is no longer the streaming superpower it once was. Kafka now falls behind in a world where edge computing has redefined real-time data from seconds to milliseconds, alongside the rise of AI and the ever-increasing demand for real-time analytics.

The problem lies in its foundation. Kafka was designed to exploit commodity spinning disks, so while it can be used with modern hardware and networks, it isn't optimized for them. Kafka deployments also require a lot of infrastructure, so costs quickly escalate as you grow.

Kafka has also been slow to respond to the evolving needs of developers. Some of the most common complaints include:

- Kafka is built in Java, which makes it tough to reliably scale in production.
- Requires the JVM and external dependencies that create additional complexities.
- Impossible to customize performance for each workload, leading to resource inefficiencies and higher costs.
- Data connectors that are easy to misconfigure but hard to customize, monitor, and debug.
- Requires an in-depth understanding of numerous settings and parameters to run efficiently.
- No user interface adds to the high learning curve and slows development

Simply put, **it takes a lot of work to make Kafka work**. So unless you have a sizeable budget to burn on a fully managed Kafka solution or the luxury of dedicating a team to manage Kafka in-house, you're better off choosing a developer-friendly platform that will help not only keep pace with the speed of innovation — but accelerate it with real-time access to mission-critical data and insights.

Meet Redpanda: A fast, simple, reliable platform for real-time data and analytics

Redpanda provides an end-to-end platform with high-performance connectivity, high-throughput processing, and resource-efficient storage that's reliable for all mission-critical workloads and use cases.

Here's how:

Streamlined architecture for simpler operations

Fewer moving parts translate to easier administration and a streamlined experience for developers and admins. Redpanda is designed from the ground up to be lighter, faster, and simpler to operate than Kafka by eliminating any dependencies on:

- **ZooKeeper/KRaft:** Redpanda is a single, self-sufficient package, so say goodbye to external dependencies (like ZK or KRaft) and hello to rapid boot times, simpler CI/CD integration, and reliable production environments.
- **External schema registries:** Redpanda provides an integrated schema registry in which each node functions as both a message broker and schema registry simultaneously, simplifying deployment and reducing operational overhead.
- **Java virtual machines (JVMs):** Redpanda is built in C++ for streamlined performance, freeing teams from the dreaded JVMs and their high memory overhead.
- **Kafka Cruise Control:** Redpanda's Continuous Data Balancing constantly monitors node and rack availability and disk usage to ensure smooth operations and optimal cluster performance with self-healing clusters that dynamically balance partitions.

Developer-first to speed up innovation

Complex platforms only slow down development and delay features from reaching your end users. With a best-in-class user experience and intuitive tools for visibility, data integration, and automation, developers can hit the ground running.

Here's how Redpanda makes building real-time apps easier:

- Fully Kafka-API compatible means developers can just drop Redpanda in and the default configuration will work with existing Kafka streaming apps and tools. No code changes needed.
- **Redpanda Console** provides a simple, interactive approach for gaining visibility into your topics, masking data, managing consumer groups and exploring real-time data with time-travel debugging.

- **rpk** is an all-in-one command-line interface (CLI) that lets developers easily set up, configure, and manage their Redpanda clusters.
- **HTTP Proxy** for REST API access lets developers skip the need to understand Kafka's binary protocol.
- **WebAssembly (WASM)** engine allows developers to perform real-time, in-broker transformations without an external stream processing system. No more "data ping pong."
- **Tiered Storage** allows operators to deploy Redpanda on fewer, smaller brokers, and with less storage — reducing infrastructure costs and admin overhead.
- **Redpanda Connect** is a simplified yet powerful alternative to Kafka Connect, with over 300+ pre-built connectors to easily integrate data from different sources and power any real-time data and AI pipelines.

Kafka Connect has long been a staple for data integration, so it helps to have a detailed comparison before switching to a new tool.

Redpanda Connect vs. Kafka Connect

Feature	Redpanda Connect	Kafka Connect
Ease of Deployment	✓ Easy and fast	✗ Complex setup
Resource Efficiency	✓ Lightweight	✗ Higher overhead
Data Transformation	✓ Pre-built or invoke custom processors via YAML	⚠ SMTs require Java for customization
Connector Options	✓ 300+ mostly open source	✓ 200+ mostly open source
Error Handling	✓ Built-in recovery	⚠ Requires customization
Scalability on Kubernetes	✓ Stateless, resilient	⚠ Mix of stateful & stateless, complex
Programming Flexibility	✓ Low-code, Go, or Bloblang (DSL)	✗ Java-only
Cluster Replication & Migration	✓ One-command Redpanda Migrator tool	⚠ MirrorMaker 2, multiple components
AI Processing	✓ Connectors for AI models & vector DBs	✗ No AI connectors
CDC Support	⚠ Specializes in Postgres, MySQL, MongoDB	✓ Uses the Debezium ecosystem
Community Support	✓ Active & growing	✓ Mature
Kafka Dependency	✓ Use standalone or with other pub/sub systems	✗ Requires Kafka for large-scale use

Bottom line: Redpanda Connect is faster to deploy, easier to scale, more customizable, has data sovereignty built-in, and simplifies migrating from Kafka using just one command.

To dig a little deeper, read our blog post on Kafka Connect vs. Redpanda Connect.

Highly performant and reliable for mission-critical workloads

Engineered in C++ with a thread-per-core architecture, Redpanda achieves 10x lower latencies than Kafka and supports GB/s+ workloads with a smaller hardware footprint, so you can do more with less — without sacrificing data reliability or durability.

With cloud-first storage and the Raft consensus protocol as its foundation, Redpanda is safe by default, delivering up to 6x faster performance while ensuring zero data loss (Jepsen-verified). In addition, its intelligent memory handling and autotuning ensure cluster integrity and high availability at petabyte scale. Just add more nodes as requirements grow.

Furthermore, with built-in observability via Prometheus, it's easy to monitor the health and performance of Redpanda clusters using your team's favorite DevOps and open-source tooling. Redpanda-provided Grafana dashboards get you up and running fast.

Safe, secure streaming that keeps you compliant

Meet data security requirements without the headaches. Redpanda supports end-to-end encryption, Kerberos authentication, Kafka-compatible ACLs, cloud IAM roles, Redpanda Console SSO, RBAC, and more.

Platform-wide audit logging of all cluster administration and produce/consume events provide data provenance and proof of controls to auditors, easing compliance efforts. A FIPS-compliant binary allows Redpanda Enterprise customers to deploy within regulated industries like finance and public sector organizations.

For organizations handling sensitive data, Redpanda BYOC's privacy-first architecture allows you to run everything within your own Virtual Private Cloud (VPC), so your data never leaves your own environment. This gives you full ownership over your data and allows you to stay compliant, even in the cloud.

Now that you know each platform better, it's time for the showdown. Below is a detailed comparison of the key features and capabilities that set Redpanda apart from Kafka.

Redpanda vs. Kafka

	Redpanda Enterprise	Apache Kafka®
Architecture		
Streamlined Components	✓ Offers single-binary deployment with brokers, schema registry, HTTP proxy, and Raft consensus built in	✗ Imposes external dependencies on ZooKeeper/KRaft, JVM, schema registry, and HTTP proxy
Engine Configurations	✓ Ability to dynamically optimize individual workloads within clusters for better security and availability, greater consistency, and lower latency	✗ Less flexible with configurations that affect clusters globally and unsafe with an ISR replication strategy that creates potential for data loss
Cloud Storage Options	✓ Production-proven tiered storage: Redpanda Cloud Topics provides direct-to-object-storage persistence (coming soon)	✓ Unproven tiered storage: production-ready as of version 3.9.0. No option to use object storage directly
Provision		
Cluster Sizing	✓ Consultative guidance on sizing and deployment is included	✗ DIY effort
Automation	✓ Continuous data and partition balancing built in	✗ Requires separate system (Kafka Cruise Control)
Monitoring	✓ Available natively via Prometheus	✗ Requires 3rd-party tools
Ecosystem		
Data Connectors	✓ Redpanda Connect: 300+ data and AI connectors, config-based and stateless for ease of use	✓ Kafka Connect: 200+ integrations, none for AI, includes a mix of stateful/stateless connectors
Kafka API	✓ Redpanda natively supports the Kafka API	✓ Kafka API is the industry standard for data streaming
UI	✓ Redpanda Console is open source and designed to work with Apache Kafka	✗ Kafka requires 3rd-party tools (like Redpanda Console)
Schema Registry	✓ Built into the Redpanda binary	✓ Requires separate infrastructure
Transactions	✓ Part of Redpanda's Kafka compatibility	✓ Kafka is the industry standard for data streaming
Support		
Access Experts 24/7	✓ 24/7/365 coverage for production outages	✗ Community-based, best-effort support
Security	✓ Enterprise-grade features such as SSO, RBAC, audit-logging & more	✗ DIY security, often requires 3rd-party tools

Choosing a Redpanda deployment

Many customers have noted Redpanda's flexible deployment options as a reason for switching over from Kafka. Redpanda runs everywhere your systems and applications can: at the edge, on bare metal, in the cloud, in Kubernetes and beyond — all without breaking the bank on escalating infrastructure and storage costs.

Here's a brief overview of each deployment option.

Redpanda Self-Managed - leverage your existing infrastructure

Redpanda Self-Managed offers two options for businesses that need full control over their data environment:

Redpanda Community Edition

Redpanda's Community Edition is a free version source-available under the BSL license, which allows you to install the single binary on your machine for use in your development and test operations.

Redpanda Enterprise

For organizations with stringent security or compliance requirements that mandate on-premises data storage, Redpanda Enterprise is a more robust, self-managed solution with full commercial support, enterprise-grade security, and Enterprise features, including data balancing, read replicas, and intelligent tiered storage for near-infinite data retention.

Both can be hosted on-premises or in your private cloud and, while very developer-friendly, are best suited for organizations with strong DevOps/SRE teams.

Redpanda Cloud – a fully managed, end-to-end streaming data platform

Redpanda Cloud offers three options: Serverless, Bring Your Own Cloud (BYOC), and Redpanda Dedicated. Check out the table below to explore which one might be the best fit for you.

	Serverless*	BYOC	Redpanda Dedicated
Provision			
Hosting Model	Multi-tenant (AWS)	In your cloud (AWS, Azure, GCP)	Single-tenant (AWS, Azure, GCP)
Networking	Public internet	Public or private (PrivateLink, Private Service Connect, etc.)	Public or private (PrivateLink, Private Service Connect, etc.)
Elastic Scaling	✓ Scale automatically, deploy or decommission in seconds	✓ Scale up, plus Tiered Storage, Follower Fetching, Leader Pinning, etc.	✓ Scale up, plus Tiered Storage, Follower Fetching, Leader Pinning, etc.
Data Sovereignty	✗ Data sovereignty is available with Redpanda's BYOC option	✓ Data resides in the customer's cloud environment	✗ Data sovereignty is available with Redpanda's BYOC option
Ecosystem			
Security	⚠ SSO (Google, GitHub), Kafka ACLs	✓ SSO (Google, GitHub, OIDC), RBAC, audit logs	✓ SSO (Google, GitHub, OIDC), RBAC, audit logs
Data Connectors	✓ Supports hundreds of data connectors available with Redpanda Connect**	✓ Supports hundreds of data connectors available with Redpanda Connect**	✓ Supports hundreds of data connectors available with Redpanda Connect**
Schema Registry	✓ Schema Registry support for Avro, JSON, and Protobuf	✓ Schema Registry support for Avro, JSON, and Protobuf	✓ Schema Registry support for Avro, JSON, and Protobuf
Transactions	✓ Transactional support included, compatible with Kafka's transaction API	✓ Transactional support included, compatible with Kafka's transaction API	✓ Transactional support included, compatible with Kafka's transaction API
Redpanda Console	✓ The leading UI for Kafka-compatible streaming	✓ The leading UI for Kafka-compatible streaming	✓ The leading UI for Kafka-compatible streaming
Support			
Max Guaranteed Throughput	Write: 100 MB/sec Read: 300 MB/sec	Write: 2 GB/sec Read: 4 GB/sec	Write: 400 MB/sec Read: 800 MB/sec
Max Partitions***	5,000	112,500	22,800
SLA	✓ 99.9% uptime	✓ 99.99% uptime with 24x7 monitoring and alerting	✓ 99.99% uptime with 24x7 monitoring and alerting
Access Experts 24/7	⚠ Enterprise support available with Annual Commitment	✓ Enterprise support and dedicated account manager	✓ Enterprise support and dedicated account manager

*Currently in Limited Availability (LA) period.
**Currently in free Beta period.
***The number of logical partitions before replication. All Redpanda Cloud services use a replication factor of 3

Redpanda Dedicated

Redpanda Dedicated offers an out-of-the-box solution for businesses that want a hassle-free, scalable, and reliable data streaming service with minimal management overhead.

Dedicated offers robust support and guaranteed uptime with clusters hosted in a single-tenant environment within Redpanda's cloud infrastructure in AWS, Azure, or GCP. With Redpanda Dedicated, clusters are managed, monitored, and maintained by Redpanda, so you can focus on building your Kafka applications and data pipelines.

Redpanda Bring Your Own Cloud (BYOC)

With Redpanda BYOC, clusters are hosted on the customer's private cloud while Redpanda remotely manages provisioning, monitoring, and maintenance. This option is designed for organizations that need to keep sensitive data within their own cloud for privacy, IP, or compliance reasons.

With BYOC, sensitive data and credentials never leave your environment, while platform engineering teams have complete visibility and control and can offload operations and maintenance to Redpanda. In short, you get full data sovereignty along with the convenience of a fully managed solution. Win-win.

Redpanda Serverless

Redpanda Serverless is ideal for those who need to start small but have big plans. It offers a fully self-service experience that makes creating clusters as easy as clicking a button. You don't have to talk to Redpanda to get started, so you can literally start streaming data in seconds.

There's no need to choose between different capabilities, pricing structures, or availability so you can choose the "right" cluster from the start. Redpanda Serverless is just one product — completely eliminating the hassle of choosing the right size. With one price, you only pay for what you use.

It's the fastest, easiest, and most cost-effective way to start data streaming. To learn more about our latest Serverless offering, check out the announcement blog.

Get started with simpler, faster data streaming

With the winning combination of simplicity, scale, and cost savings compared to Kafka, Redpanda has made a name for itself by helping organizations worldwide tap into the power of real-time data while freeing their teams to focus on innovating — not troubleshooting.

If your workloads are already on Kafka, don't worry; we even made migrating easy. Redpanda Migrator is a tool within Redpanda Connect designed to simplify migrations from any Kafka system to Redpanda *with a single command*, eliminating the multiple components and performance-tuning headaches of MirrorMaker 2. Now, developers can easily move Kafka messages, schemas, and ACLs without digging into Kafka or Redpanda internals.

If you're only just getting started with real-time data streaming, it helps to see what other Kafka converts are saying before making any big decisions.

> **Also need to add this quote: "Working with Redpanda allowed us to start with event streaming without thinking about it, and with a fully managed service, our team can focus on innovation rather than maintenance."**

Juliette Tisseyre,
VP of Engineering, Deepomatic

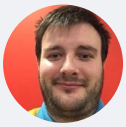
> **We have reduced our cloud infrastructure spend by about 70%, saving millions of USD annually. We are now scaling up new use cases on our event streaming architecture without worrying about spiralling infrastructure costs."**

Arya Ketan
Senior Principal Engineer, ShareChat

> **Redpanda takes the responsibility of managing Kafka off our team's plate, while allowing us to keep sensitive customer data in our own infrastructure. It's the best of both worlds."**

Ted O'Hayer
DevOps Manager, HealthEdge

> **Redpanda was five to six times more efficient, greatly reducing our hardware costs, support and licensing costs."**



Michael Pearce,
Sr. Architect, StoneX

> **Redpanda just works. We didn't need to change our code, and suddenly, our day two operations were dead simple."**

Shahir Daya,
CTO, Zafin

So if you're ready to switch to the simpler side of data streaming and learn how Redpanda can help you win in your industry, book a meeting and our team will lead the way. If you'd rather dig in and see it for yourself first, try Redpanda for free.

If, however, you're still on the fence, here are a few more resources:

Redpanda | Customer Success Stories
Redpanda | Documentation
Redpanda | Streamcasts, Tech Talks, and Masterclasses
Web page | Redpanda vs. Amazon MSK
Blog | When to choose Redpanda vs. Apache Kafka?
Blog | Kafka Connect vs. Redpanda Connect
Blog | Redpanda vs. Kafka with KRaft: Performance comparison
Report | Redpanda vs. Confluent: A Performance and TCO Benchmark
Slack | Ask our team in the Redpanda Community

Get started with Redpanda for free

Explore all the capabilities of Redpanda Serverless, Redpanda Cloud or Redpanda Enterprise with our free trial options!

START FREE

Connect with us