

DIFFUSION® & Kafka



The Diffusion Intelligent Data Platform consumes, enriches and delivers data among applications, devices, & systems – via web, mobile, & satellite networks.

The Diffusion platform extends the reliability, scalability, & functionality of in-place data streaming & messaging products.

Kafka Options

Apache Kafka is an Open-Source messaging broker that provides real-time pub/sub data streaming for back-end systems with reliability and scalability. Due to Kafka's flexibility and performance, it has wide adoption as the back-bone for event processing and data pipelines across various market sectors.

Confluent Inc. offers a product suite built upon the core open-source Kafka product that includes added functionality for the needs of Enterprise architectures, including comprehensive tooling for monitoring, resilience, and management of Kafka deployments. In addition, Confluent integrates with Kafka Connect, providing a variety of “connectors” for integrating 3rd party products with Kafka. This ecosystem is designed to provide a simple, “out of the box” experience for using Kafka in any environment.

While Kafka has a proven capability for high-throughput data handling within the data center, it is not designed for last-mile streaming – the delivery of data over the

public Internet to Web, Mobile, or IoT devices. When dealing with potentially hundreds of thousands of concurrent connections, the qualities that make Kafka extremely powerful in the back-end and within the data center do not scale to address the unique challenges of edge networks.

Diffusion can connect to both traditional and Confluent-specific Kafka installations. The Confluent Kafka Adapter connects one Kafka server to one Diffusion server. The Apache Kafka Adapter allows you to connect many to many Kafka and Diffusion servers. No code is required - simply complete a few text fields and Diffusion will consume, enrich and deliver topic views. You don't need to learn how to use Diffusion and your data flows.

Diffusion Intelligent Data Platform

The Diffusion platform is a real-time data platform for web, mobile, and IoT applications. It is designed to offer similar pub/sub functionality over last-mile network connections such as mobile or satellite, while retaining high-throughput and Quality-of-Service at scale. As a result, Diffusion reduces application code complexity and speeds time-to-market.

By combining Kafka and Diffusion, developers get the benefits of a well-integrated back-end data pipeline, plus the critical reliability, resilience, and high-performance front-end data delivery. The Diffusion Kafka adapter makes this process simple.

The Diffusion Kafka Adapter

The Diffusion Kafka Adapter seamlessly integrates with Apache Kafka as well as Kafka Connect and Confluent's Control Center. By simply specifying ingest and broadcast configuration to any Apache Kafka server, or placing the adapter in your Confluent's Kafka deployment's plugin directory, you can use the Control Center UI to easily add a new Source or Sink connector for ingest or broadcast of events. The Diffusion Kafka Adapter automatically maps between Kafka messages types and JSON, allowing for web, mobile, or IoT clients to consume data that is stored as rich data structures within Kafka.

Bandwidth

One of the key challenges with running Kafka over the Internet is that the Kafka protocol consumes a lot of bandwidth. This is not typically an issue when running within a data center because bandwidth is cheap and constant. Unfortunately, these same qualities do not apply over the Internet. Mobile connections constantly vary in available bandwidth, either due to signal strength, regional carriers, or local contention – even when the end device is connected over a modern 4G network. The same is true of satellite connections, where bandwidth is often tightly constrained by design. Diffusion is uniquely designed to optimize data delivery for precisely these scenarios, being able to reduce the amount of data sent over the network by up to 99% - in a completely transparent manner.

Security

Unlike Kafka, Diffusion provides fine-grained, role-based permission access control which can be changed dynamically on-the-fly without disrupting operations. Granular access controls determine

who can see specific subsets of the live data model, while application-level controls over data & session life cycles enable total management of system events.

Scalability

Diffusion is purpose built to easily and reliably scale to meet the high-traffic demands of modern web, mobile, and IoT applications. While Kafka provides the ability to run consumers within separate consumer groups, this is difficult to manage across multiple partitions when dealing with hundreds of thousands of concurrent connections.

The Kafka connection model design is predicated on the notion that slow-consuming clients – i.e. clients that are unable to receive data as fast as it is produced - can be handled by adding more clients to the same consumer group. This model works for back-end systems, where nodes can be added according to need. However, for user-facing applications, this model is neither efficient nor effective. For any given application, each user requires a unique connection that must perform well, securely, and independently of all other connections.

Diffusion is built to address these stringent requirements, with functionality such as delta streaming, conflation, throttling, and adaptive flow-control. Diffusion assures that every user/device connection operates at peak performance. Diffusion also exposes data at a far more granular level than Kafka. This granularity allows data to be spread across hundreds of thousands of unique topics which permits individual clients to only subscribe to the data that the application needs at any given moment. This Diffusion functionality assures that a high volume of clients will receive data while retaining sub-second latencies regardless of connection quality.

Web & Mobile

Diffusion also includes SDKs designed and optimized for web, mobile, and IoT data management and distribution, with full feature support across all platforms. Integration of custom authentication – such as SSO (Single Sign-On) or LDAP – combined with a role-based permission system makes it simple to ensure that front-end clients only receive data for which they are allowed access.

Presence notifications and session events allow back-end systems to react to new connections and perform higher-level application logic on behalf of front-end clients. All Diffusion Client SDKs make it simple for developers to connect and subscribe dynamically to data, while providing automatic and reliable reconnect in the event of unexpected and inevitable disconnections.

Diffusion reduces the complexity of business-critical application development, substantially reduces delivery time-to-market, and eliminates the challenges of data delivery over last-mile connections.

Diffusion & Kafka Benefits

Diffusion radically simplifies the complexities of building Internet-based applications. By integrating Diffusion with Kafka, developers can:

- Easily and quickly extend existing data pipelines across the web, without incurring the costs of solving the challenges of data delivery and security access at scale.
- Use Diffusion to act as the edge-broker for dispersing data to large numbers of end-user applications & devices, or efficiently and reliably consume data from disparate sources including IoT devices, without any changes to the pre-existing Kafka architecture.
- Increase reliability and resiliency of web & mobile applications, especially when coping with congested or unreliable networks.
- Expand and scale operations on-premise, in-the-cloud, or in hybrid environments to match rapidly increasing or fluctuating demand, with minimal investment in infrastructure.

