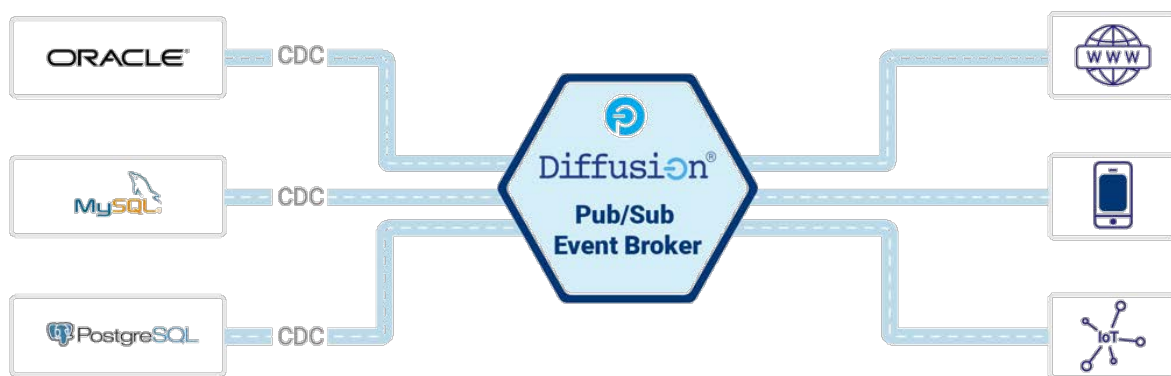


Capture & Maximize The Value Of Your Database Events – In Real Time

The world has changed since technologists began developing mission-critical applications using a monolithic database as the central location for ingesting, storing, analyzing, and exporting virtually everything going in and out of the business. Companies have become distributed and are working in the cloud as well as on premise. More types of data must be integrated in real time, and agile and responsive application architectures are required. However, transforming to an event-centric organization and processing event data streams from non-event based, source-system databases poses challenges to:

- Easily detect and capture data from source system databases,
- Integrate, enrich, and transform the data, on-the-fly,
- Assure reliable, resilient, and secure, low-latency event data delivery, in the required format, to a wide variety of target consumers.

As every technologist knows, creating an efficient, connected and fast-moving pipeline of source system(s) data to application-ready event data is complex and time consuming. Selecting the right Change Data Capture tool and data platform for your event driven architecture is more critical than ever given the need to capture, transform and propagate events (data changes) across multiple environments to a wide range of data consumers – in real time and at high volume.



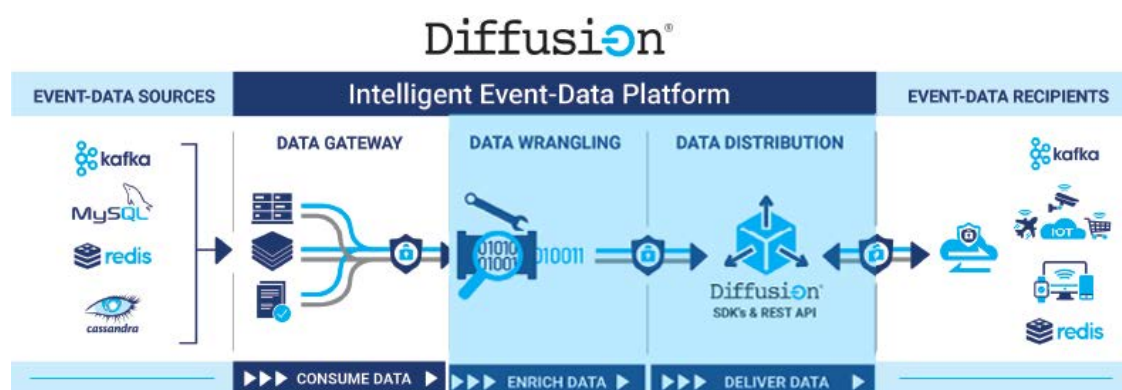
"Event notifications will form part of over 60% of new digital business solutions." Gartner

CDC Is Easy With The Diffusion Intelligent Event-Data Platform

Client applications and systems need event data when the events occur. However, the development effort required to capture, transform, and distribute the event data from your backend databases to your consumers, in a reliable, scalable, and secure manner, is a substantial development challenge. The Diffusion Intelligent Data Platform solves this challenge, speeds Event-Driven Architecture development, and simplifies ongoing operations. Diffusion handles:

- **Data Consumption** – The Diffusion platform Gateway makes it simple to capture and ingest data from any source, saving development time and effort. The Gateway has an array of adapters to capture all forms of data feeds, including a pre-built CDC adapter to stream updates from source database(s) such as MySQL and PostgreSQL into Diffusion. The Diffusion Gateway also includes adapters for: Kafka, JMS, and REST.
- **Data Wrangling** – Raw event data streams rarely contain the exact content or are in the correct format required by consumers. Therefore, the data must be transformed to suit the consumers' needs. Since event data is time sensitive, the transformations must be done on-the-fly and in real time. Diffusion's low-code data wrangling capabilities reduce the amount of coding required using DSL to handle in-flight protocol translation, data transformation, and payload conversion, to assure consumers receive the exact data needed – every time, in real time.
- **Data Delivery** – Diffusion distributes the desired data reliably, efficiently, and in real time, which is critical to the business value of event data. The data required and allowed for each consumer is assured. In addition, event data can be buffered if the rate of event production exceeds the capacity of a particular event consumer or consumer group.

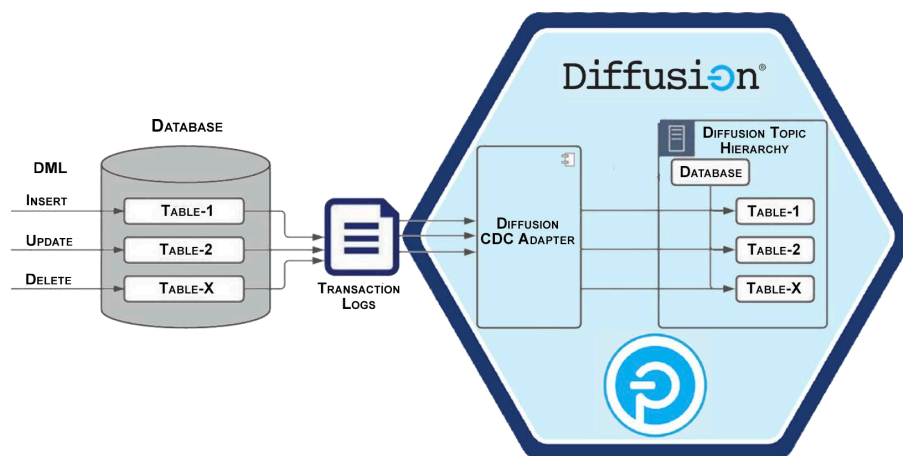
Importantly, security is maintained because it is an integral part of the Diffusion platform with fine grained role-based access control via security policies applied instantaneously to all connections. Authentication options include OAuth, JWT or custom authentication.



The Diffusion CDC Adapter In Action

As a low-code, intelligent event-data platform with capabilities spanning all of the Event Data Architecture broker types (queue, log, and subscription), Diffusion handles the entire data journey from data capture to delivery.

The Diffusion Data Gateway CDC adapter captures change data plus its schema from non-event based, source-system databases. The adapter then processes and publishes the change data events to JSON Diffusion topics according to the configuration you choose. Depending upon the business use case, consumers can subscribe to topics containing the database event changes.



The Diffusion CDC Gateway adapter is based on Debezium, a popular open-source change data capture tool, that monitors database logs/tables to identify changes and pushes these changes into the Diffusion Intelligent Event-Data platform.

Debezium is the optimal technology choice because it simplifies change data capture from databases in a highly scalable way and it provides the real-time latency of a push-based system with minimal changes to the source system(s). The only change usually required to the source database is a configuration change to ensure its database logs include the correct level of detail for Debezium to capture the desired changes.

Simple Configuration

Using the Diffusion Console Network Tab, you can easily configure the CDC adapter to define its behavior. The configuration can also be modified/updated during runtime via the console. Configuration to setup and share data between databases and Diffusion can be done using either the included configuration.json file or environment variables. Every event (insert/update/delete) from the database is captured and published to Diffusion.

The screenshot displays the Diffusion Management Console interface. On the left, the 'Network' tab is selected, showing a list of gateway clients under the 'CDC (cdcLocalConnector)' service. A modal window is open for configuring a 'DATABASE_CONNECTOR' service, showing details for 'publisher1', 'publisher2', and 'publisher3'. The main panel shows the 'Gateway' configuration page, which includes a description of the gateway and a configuration JSON snippet. On the right, a 'Cluster' overview panel shows the 'diffusion-primary-0-CDC' node as 'Connected' with details like 'Diffusion 6.7.2', 'OpenJDK 64-Bit Server VM', and 'Linux'. A large text overlay on the right side of the console reads: 'Connect to all your data sources using Gateways' followed by 'CDC ✓', 'Kafka ✓', 'REST ✓', and 'and more'.

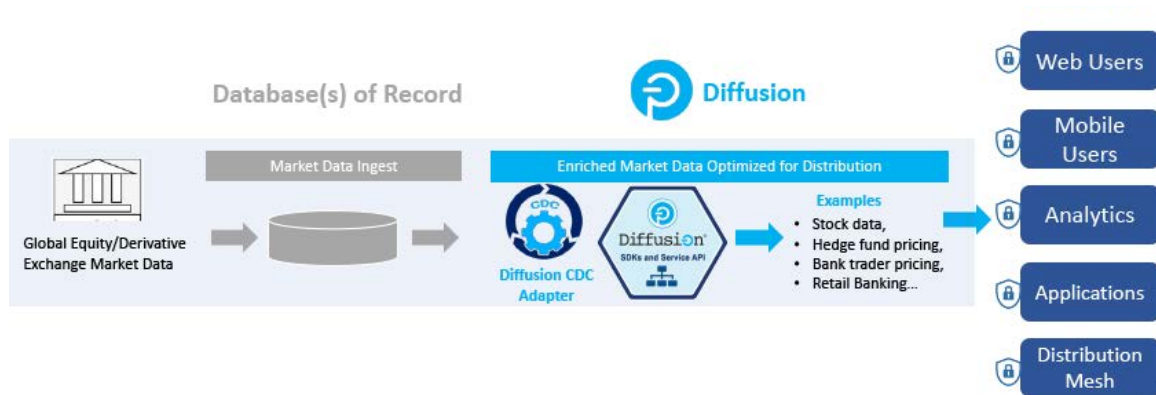
The CDC adapter provides four convenient views of the data as configuration options, for maximum flexibility in defining how your event data is published to Diffusion Topics:

- **Object:** A table is mapped to a JSON Topic. Each record is keyed by the table's primary key.
- **Array:** All records are entities in a JSON array.
- **Row:** Each table row is mapped to individual JSON Topics : database/table/pk1.
- **None:** Topic contents are exactly as received. Schema Information. Before Change Table Row Data. After Change Table Row Data.

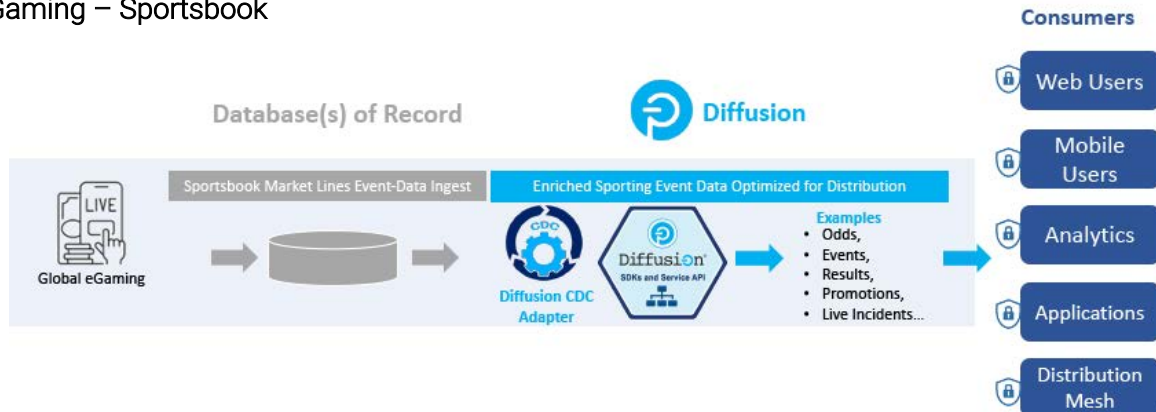
In addition, because the Diffusion Gateway CDC adapter uses Debezium to capture the change data events, all configuration options supported by Debezium are allowed. The CDC Adapter is part of the Diffusion Gateway included in your Diffusion license, and requires Java.

Event-Driven CDC Use Case Examples On The Diffusion Platform

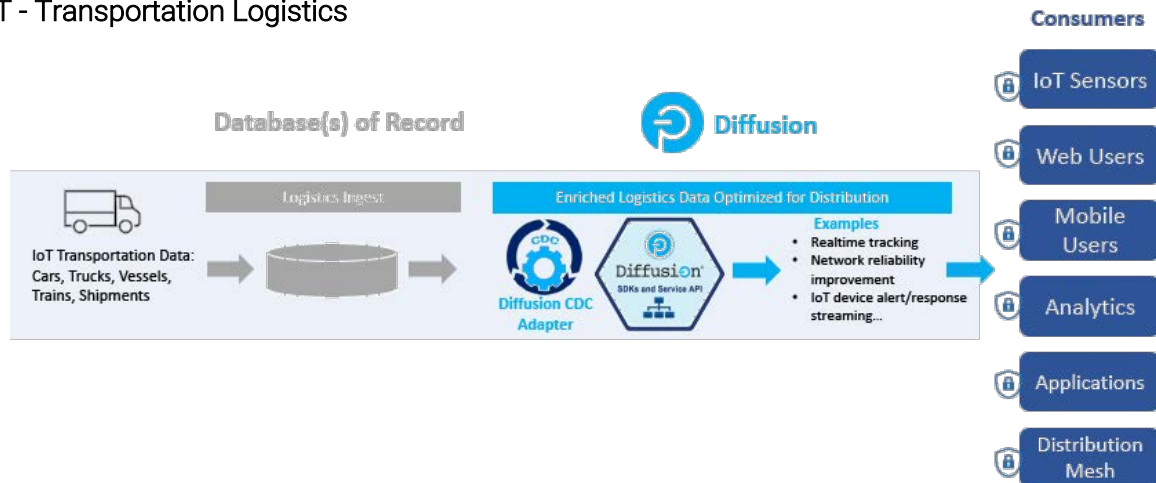
Financial Services – Market Data



eGaming – Sportsbook



IoT - Transportation Logistics



Bottom Line

Event data is only useful if you can capture, analyze, and act on it quickly. Failure to handle data in real time prevents you from taking advantage of your valuable data assets.

Change Data Capture plays a vital role in technologists toolboxes to ensure that data in non-event based source systems is captured, transformed, and delivered to applications, dashboards, analytics tools, etc., in real time.

The Diffusion platform's CDC capabilities speed Event-Driven Architecture development, simplify ongoing operations, and deliver undeniable business benefits to:

- ✓ Maximize the value of your data assets by eliminating data silos,
- ✓ Leverage your data for a variety of purposes,
- ✓ Maintain multiple applications and systems in synch,
- ✓ Help your business make faster data-driven decisions.

The Diffusion Intelligent Event-Data Platform *Future Proofs* your architecture. In addition to capturing, processing, and streaming CDC data from source databases, the Diffusion platform can consume raw event data from a wide variety of data sources and feeds in any size, format, or velocity, to handle all of your data consumption, transformation, and delivery needs. Plus, Diffusion easily scales as you require – be it the amount of data, number of events, or number of consumers to be served – to meet your needs today and in the future.

DIFFUSION DATA

[Get Started with
Diffusion Today!](#)

Additional Resources:

[Details on the Diffusion CDC Adapter](#)

[Understanding Debezium](#)

www.DiffusionData.com

UK +44 (0) 20 3588 0900

US +1 844 449-DATA (3282)

Ireland+44 (0) 20 3588 0900