DELL Technologies

**Validated Designs for Analytics**
**Solution Brief**

# Enable Your Real-Time Enterprise

## Real-Time Data Streaming with Apache Kafka and Confluent on Dell EMC Infrastructure

## Kafka with Confluent

Apache® Kafka® is an open-source distributed streaming platform capable of handling trillions of events a day. Confluent® Platform improves Kafka with additional open source and commercial features designed to enhance the streaming experience of both operators and developers in production, at massive scale.

**Optimized infrastructure allows modular integration from edge to insights.**
Businesses need real-time data insights from the core to the edge to be responsive, predictive and competitive. Transitioning from batch processing to real-time data and event-driven applications enables innovation, increases agility, and supports the creation of new digital products and services.

This journey to a real-time enterprise starts with a streaming platform for data. However, real-time data processing solutions are often complex and time-consuming to design and implement. The Dell Technologies Validated Design for Real-Time Data Streaming helps reduce the time, effort and resources spent on architecting real-time data pipelines and streaming applications.

### Turn data into insights at the point of action.

As data is continuously and infinitely generated in the enterprise and at the edge, cloud computing, even using 5G networks, introduces latency and connectivity issues that can interfere with real-time response. Data needs to be captured and processed at its source to enable use cases that demand immediate response, such as enabling autonomous vehicles to react instantly to pedestrians, addressing assembly line issues before they affect production, and detecting and preventing financial fraud. To enable groundbreaking, data-driven innovations like these, data needs to be compiled, controlled and consolidated with nanosecond precision.

Real-time data stream processing enables instant insights from endless data streams, but selecting, deploying and tuning the infrastructure can be complex. Traditional streaming data platforms are built using a vast ecosystem with numerous moving parts, creating a high barrier to entry.

The Dell Technologies Validated Design for Real-Time Data Streaming with Apache Kafka and Confluent platform is an optimized real-time data processing system that allows modular integration from edge ingestion to analytical results. Deploying a streaming platform can avoid the latency of batch processing and cloud computing, making data available in real time to applications that need it.

## Learn more

- Real-Time Data Streaming Design Guide
- DellTechnologies.com/Analytics

## Solution benefits

**Use more data** — Enable real-time data pipelines and streaming applications by integrating data from multiple sources into a single, central event-streaming platform.

**Improve productivity** — Efficiently filter and flow data to where the right people and tools can be applied to extract useful insights.

**Reduce complexity** — Leverage validated designs that simplify connecting data sources to Apache Kafka, building applications with Kafka services, and securing, monitoring and managing Kafka infrastructure.

## Configuration options

Choosing the correct infrastructure for a streaming platform is critical for success, performance, and scalability. The Validated Design is an engineering-tested and -optimized infrastructure for deploying Kafka and Confluent Enterprise in production environments. The Real-Time Data Streaming solution combines the data distribution strength of a publish-subscribe model with a storage layer and a processing layer. This makes it much easier to create data pipelines and connect them to all of your systems.

The Dell Technologies design guide describes recommended configurations for implementing Lambda and Kappa stream processing architectures using Apache Kafka and Confluent Enterprise. The design can be used to implement small or large clusters, with a clear path to scaling from small to large. It also addresses the performance, latency and reliability requirements of mission-critical production deployments.

Configuration options

| Dell EMC PowerEdge servers | Networking | Software |
|---|---|---|
| • Confluent Platform: R650<br>• Kafka brokers: R650<br>• KSQL servers: R650 | • Dell EMC PowerSwitch S3048-ON (management)<br>• Dell EMC PowerSwitch S5148-ON (rack)<br>• NVIDIA® ConnectX®-4 Lx dual port 25GbE (cluster) | • Apache Kafka<br>• Confluent Platform |

## Take the step toward real-time streaming data processing.

Take the first step toward real-time streaming data processing. Contact your Dell Technologies representative to learn more about how this Validated Design for Analytics can help you capture and act on more data, faster.