DELL Technologies

Validated Designs for Analytics
Solution Brief

Leverage the Power of Data and Edge-based Analytics

Edge Analytics for Industry 4.0 with Confluent Platform on Dell EMC Infrastructure

There is one undeniable fact regarding Industry 4.0 — data is cheap and plentiful. The challenge today is how to derive value from that data to produce a positive ROI.

The ingesting, moving, processing and storing of data incurs incremental costs that are sizeable for any project with the potential for significant return on investment (ROI). If data was streamed for all sources from every piece of equipment and sensor in a modern industrial facility, disk storage appliances could fill up faster than they could be ordered and installed. But to what end? It is unlikely that creating a fully populated data lake in the context of Industry 4.0 would produce a positive ROI.

This understanding leads to being able to define perhaps the biggest challenge of Industry 4.0 — data curation. How can an organization filter the sea of data already available in order to identify the streams of information that will help produce a better ROI? The areas of possibility can include reducing cost, reducing safety risk or improving productivity. However, most organizations have not yet identified an efficient method to get data from the vast population of industrial Internet "things" to a destination where the right people and tools can be applied to extract useful insights.

There are no silver bullets in an environment as complex as Industry 4.0. Therefore, no single technology or solution exists that solves the challenges of building and deploying cyber-physical systems for every class of problem. The best solutions are built with technology that is enough to be used for multiple applications and use cases. This Dell Technologies Validated Design focuses on the use of Kafka® and the Confluent® Platform.

Kafka is a proven technology platform that has found success in many large-scale enterprise and Internet scale applications. Although the early Kafka successes were mainly in the realm of enterprise IT and application integration, the use of Kafka in both Industry 4.0 and consumer Internet of Things (IoT) environments has been getting more attention in the last few years. Dell Technologies believes that the Confluent Platform can have a significant role for building a complete end-to-end data analysis value chain for Industry 4.0 applications.

The Confluent Platform enables data science and IT practitioners to collaborate on building real-time data pipelines and streaming applications, by integrating data from multiple sources and locations into a single, central event streaming platform. Confluent Platform simplifies connecting data sources to Kafka; building applications with Kafka services; and securing, monitoring and managing Kafka infrastructure.

Resources

- Review the design guide.
- Get Validated Designs and performance testing information at the AI & Data Analytics InfoHub.
- Explore the HPC & AI Innovation Lab.

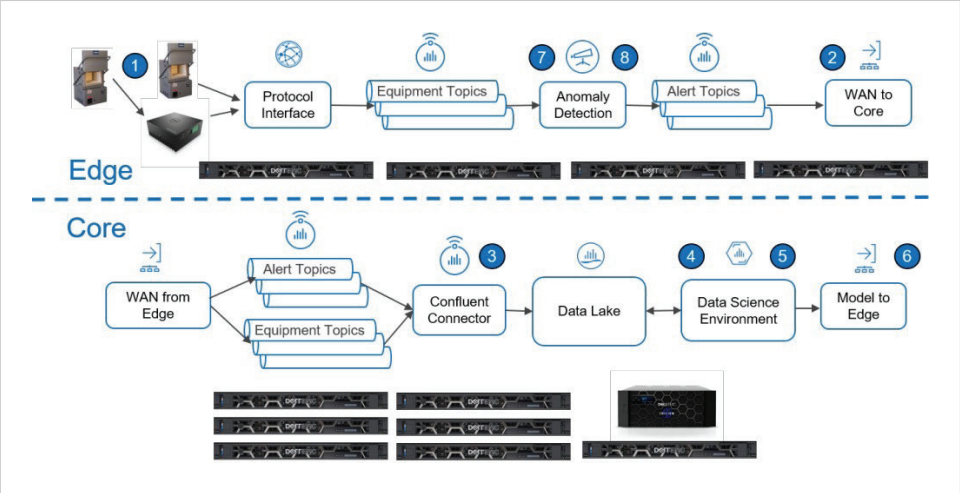
Learn more

- Validated Designs for Analytics
- Apache® Kafka
- Confluent

Validated Design

The Dell Technologies Validated Design for Analytics — Edge Analytics for Industry 4.0 with Confluent Platform is designed to support Industry 4.0 deployment scenarios in industrial facilities. This design uses an edge and core model to address the following requirements: support for multiple edge locations and multiple pieces of equipment of varying types at each edge location, anomaly detection at the edge, while disconnected from the core, flexible handling of alert notifications, support for a data science environment integrated with the data collection system, and scalable data collection and storage.

The following shows the overall edge-to-core design. Confluent Platform is used to provide the data-streaming capabilities from the edge to the core and the management of that system. Red Hat® OpenShift® Container Platform is used for the runtime environment at the core, including Confluent Platform. The data science environment uses multiple components running under OpenShift, and a Dell EMC Isilon storage array provides the data lake storage.



Summary

Successful analytics strategies for Industry 4.0 require dynamically selecting and routing required application data from a subset of devices. In turn this data must flow through processing at the edge or upstream at a core or cloud facility. In addition to managing data flows, Industry 4.0 practitioners need the ability to flow data model artifacts to infrastructure that hosts useful, problem-solving applications such as improving equipment uptime, managing worker health and safety and improving yields.

The Dell Technologies Validated Design for Analytics — Edge Analytics for Industry 4.0 with Confluent Platform design shows how the Confluent Platform on top of Dell EMC infrastructure can be applied to the challenges of industrial IoT and Industry 4.0, enabling organizations to leverage the power of data to drive greater efficiencies and better decision-making.