



Enhancing Ignition SCADA with contextual data and ML capabilities via cloud data lake integration

Introduction

Ignition SCADA by Inductive Automation is a comprehensive, industrial-grade software platform for building and deploying powerful HMI, SCADA, and IIoT solutions. Ignition is built on an open architecture allowing easy integration with almost any device or protocol. This makes it easy to create and maintain a powerful, unified system that can be scaled up or down depending on your needs.

Ignition's scripting and programming capabilities can be used to build sophisticated real-time IoT solutions. Machine Learning models can make these solutions more potent by providing predictive and optimization capabilities. However, machine learning models typically require additional data sources and processing power.

contextual data sources from IT or third-party vendors in addition to control loop data.

Typically, in a contemporary data architecture, operational and external data is initially stored in a data lake before becoming accessible as input for other systems

In addition, the need for low latency in the given scenario may necessitate running machine learning models before ingesting data into Ignition.

Integrating data from your data lake into Ignition's data model and building and managing machine learning models that rely on these data sources can present significant challenges, including:





External data sources may not conform to the data model within the Ignition. As a result, the data needs to be contextualized before ingestion.



The control network that Ignition might be installed on does not permit pull data requests from external sources because of security requirements.



Significant data engineering might be required before the operational data is amenable to machine learning.



Managing the machine learning model life cycle so that they are continuously refreshed and monitored requires a significant amount of custom software development.

DeepIQ enables you to establish secure and scalable data pipelines between different cloud data lake platforms and your Ignition SCADA, all without writing a single line of code. These data pipelines can be highly sophisticated, empowering you to move data seamlessly and extract insights.

Supported Data Lake Platforms and Communication Frameworks



DeepIQ DataStudio provides comprehensive support for industry-leading data lake platforms across the major cloud providers. Supported Platforms include:

- Snowflake
- Databricks Delta Lake
- AWS Redshift
- Google BigQuery
- Google BigTable
- ADLS Gen2
- AWS S3

In addition, DeepIQ supports reading and writing to streaming sources, including:

- AWS Kinesis
- Kafka
- Azure Event Hub
- Google PubSub

DeepIQ can read data from Ignition in multiple ways, as shown in Figure 1.

1. Ignition can be configured to expose all tags as an OPC Server. You can access Ignition data in real-time by subscribing to this server with DeepIQ OPC Client.
2. Ignition's SQL bridge module can sink data into an external database on a scheduled frequency. DeepIQ's Database connectors can pull data from this database on a schedule.
3. However, the simplest way to access Ignition's data is using DeepIQ's add-on web module. This module can be imported to Ignition and provides read and write access to the tags in Ignition. This API is protected using a secret token, and IP number-based restrictions, so unauthorized applications cannot access it.



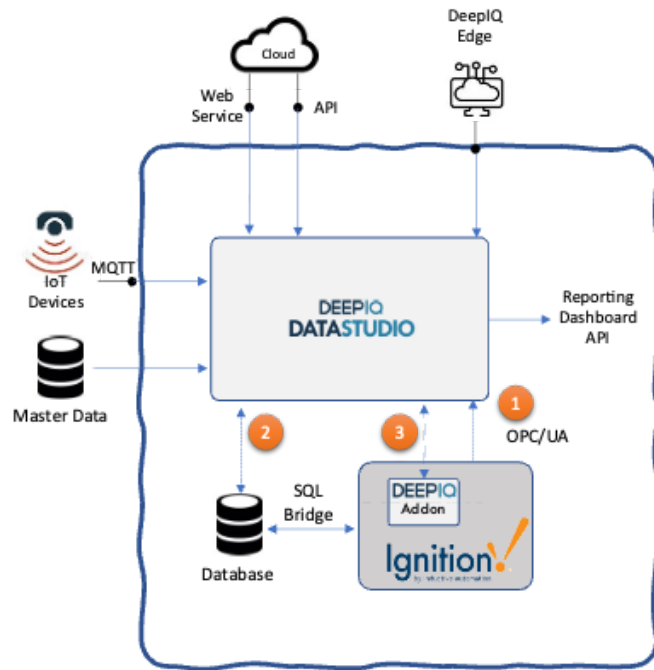


Figure 1: Typical DataStudio and Ignition Architecture

DeepIQ can be used to write data to Ignition in similar ways:

1. Ignition's SQL bridge has two-way data movement software and can synchronize tags from an external database. DeepIQ can write directly to these databases using its database connector.
2. DeepIQ's add-on module provides write access to tags when enabled by Ignition.

Security

When Ignition is used as a Tier 1 SCADA system and is within the operational network, external applications may not be allowed read access to the API or the databases that Ignition can read from.

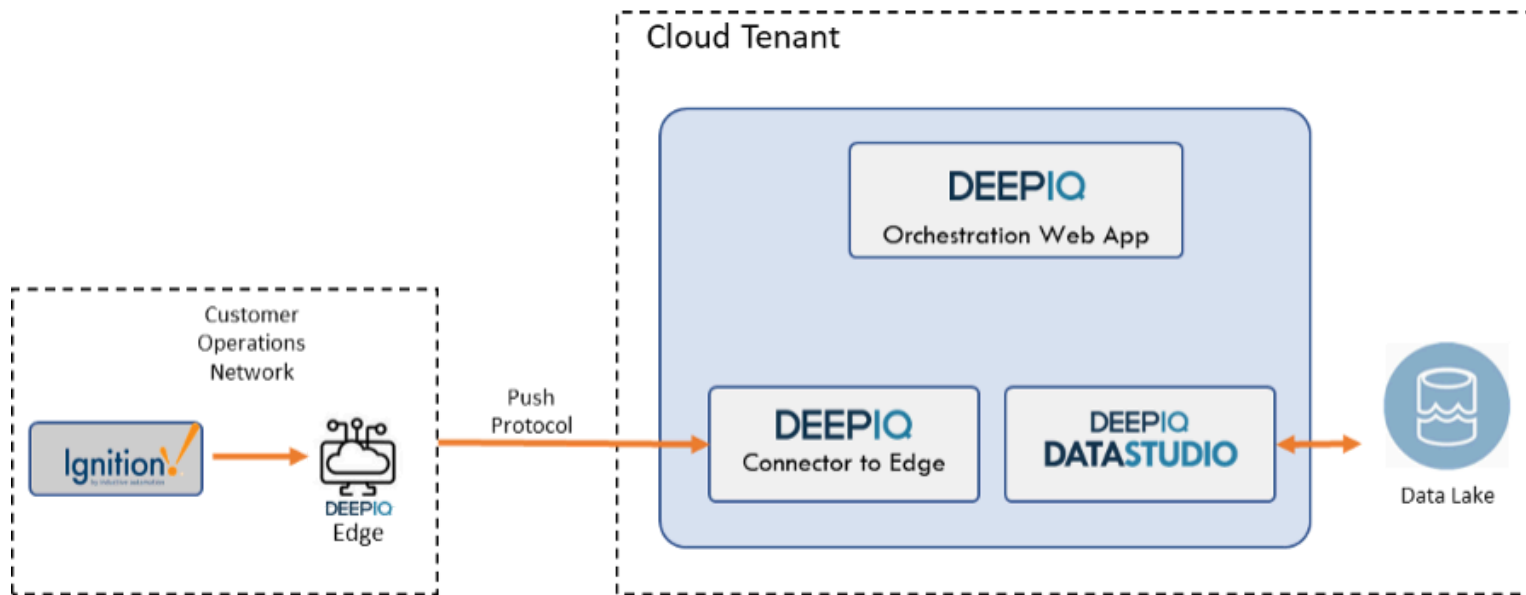


Figure 2: Edge Architecture

Similarly, DataStudio software pushes relevant data to the DeepIQ edge, which can then write to Ignition.

The edge software developed by DeepIQ utilizes a “push protocol” that conforms to the Purdue security model. Figure 2 demonstrates the data flow architecture. The DeepIQ Edge software is installed in the control network, and it employs the site-to-site protocol to monitor requests from the DataStudio server, which is installed in your cloud tenant. Afterward, it retrieves data from Ignition utilizing one of the methods specified earlier and transfers it to DataStudio, which then stores the data in the data lake. The communication between the edge and the cloud uses military-grade key-based encryption and compression.

Contextualization

Incorporating external data feeds from third-party sources or IoT devices streaming to an MQTT broker in your cloud tenant may pose challenges as they may not conform to the naming conventions used by User-Defined Types (UDTs) in your Ignition SCADA model. For instance, if your Ignition data model follows a naming hierarchy such as basin/field/well pad/well used in the oil and gas exploration and production (E&P) industry, the external data source may use a different naming convention, making it challenging to integrate the data. DeepIQ’s contextualization capabilities make data synchronization simple, even in this scenario.

While the exact details of how mastering multiple data sources or the data lake design are outside the scope of this article, the following description illustrates the key concepts.



It is possible that there may not be a precise mapping between the different hierarchies in some cases, as certain entities may not be present in one of the data sources. To handle this scenario, DataStudio workflows can be customized to identify the unmatched data that must be exported into a sink and stored in a temporary table. This temporary table is re-evaluated once the global master table is updated, allowing for the inclusion of previously unmatched entities.



Figure 2: Edge Architecture

Machine Learning

As a first step towards machine learning, it's essential to implement data engineering workflows to pre-process timeseries data. DataStudio provides robust capabilities for imputation, interpolation, outlier removal, and change-point detection that can efficiently scale to high-volume and high-velocity operational use cases. Furthermore, DeepIQ provides a flexible data flow environment that allows seamless communication between your Operational Technology (OT) and cloud networks. Once your data is curated and contextualized, you can train machine learning models on your data lake and export them to Ignition. These models can then be used to generate real-time insights with the help of Ignition.

modules. Alternatively, DeepIQ supports workflows that stream data from Ignition to your cloud data lake, which enables machine learning models to be executed in real-time on your cloud platform. Both approaches offer different advantages and feasibility; the choice between them will depend on your specific use case requirements.

Conclusion

Leveraging DeepIQ to integrate your cloud data lakes with Ignition SCADA can significantly simplify the architecture and implementation of AI-enabled IoT solutions. This integration allows you to tackle several high-ROI use cases, including predictive maintenance, anomaly detection, process optimization, and quality control, with greater ease and effectiveness.

For more information, please get in touch with us at info@deepiq.com

DeepIQ is on a mission to transform industrial processes by digitizing industrial expertise. Our vision is to drive end-to-end automation, enabling systems such as self-running power plants or drilling rigs using generative AI as the higher order reasoning layer operating over existing industrial automation technology stack.

[Privacy Policy](#) [Terms of Services](#)



FEATURES

Extract
Engineer
Explore

INDUSTRIES

Upstream
Midstream
Downstream & Chemical
Mining

SOLUTIONS

IT-OT Contextualization
Well Construction
Optimization
P & ID Digitization
Predictive Maintenance

PARTNERS

AWS
Azure
Cloudera
Google



Production Optimization

Route Optimization

Sustainability

Databricks

OSIsoft(AVEVA)

Snowflake

COMPANY

About Us

Resources

News

DataStudio Deployment
Guide

Career

Contact Us

