



# Modernizing Your Data Historian with InfluxDB

Ben Corbett  
Solutions Engineer, InfluxData

January 2024





---

## **Ben Corbett**

Solutions Engineer, InfluxData

# Goals

*(Why I am here?)*

To provide a high level understanding of:

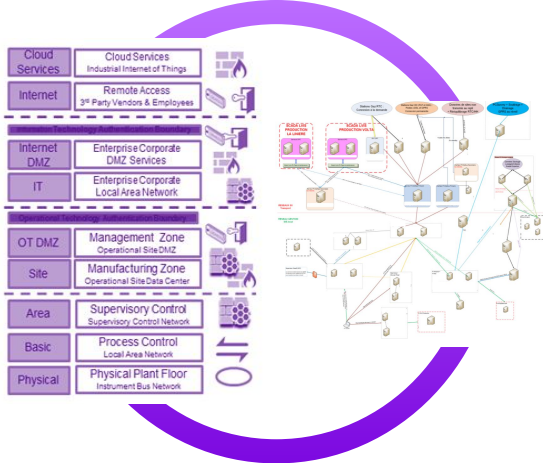
1. The key **differences** between Data Historians and TSDB's
2. **Challenges** facing historian customers today
3. What **benefits** a TSDB, and specifically InfluxDB, could give you

# Agenda

- Historian vs TSDB
- InfluxDB 3.0
- Integrations & Partners
- Customer Example
- Q&A

# Data Historian

A data historian is a **specialised database for industrial settings**, typically deployed on-prem, that's designed for collecting, storing, and retrieving high-frequency time-stamped data.



# TSDB

A TSDB is **general purpose** database for storing any data with a timestamp.

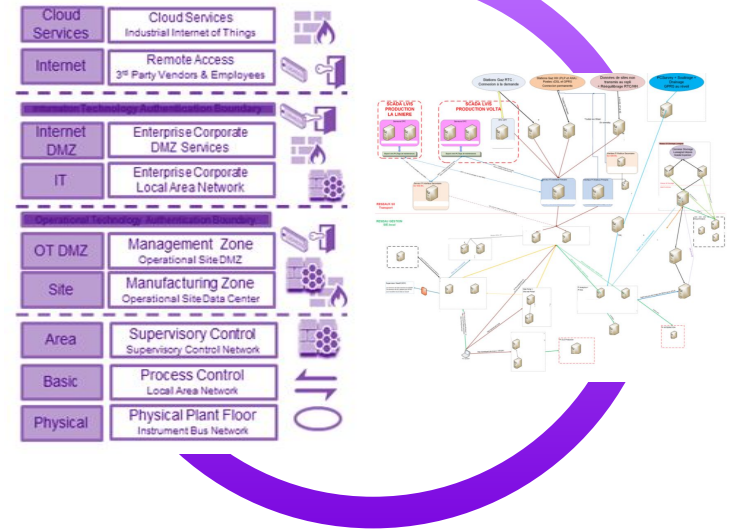


# Data Historian

## Pro's

1. **Domain specific** functionality, built from the ground up for industrial and manufacturing environments
2. Well integrated with **OT control systems** and standards
3. An **e2e solution**
  - a. Rich in functionality

## On-prem infrastructure to manage and interconnect

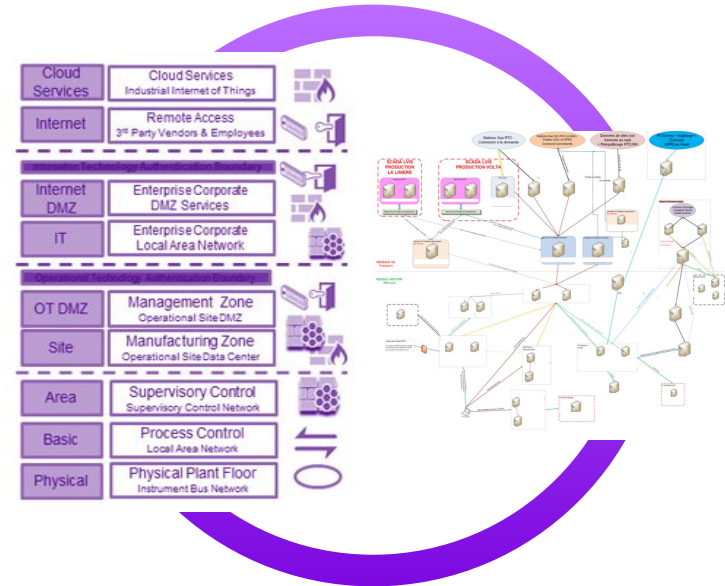


# Data Historian

## Con's

1. Rigid legacy tech - don't get left behind!
  - a. 'Walled gardens' of proprietary tech **hampers ability to adapt, innovate and grow**
  - b. How to digitally transform and **integrate with the modern data ecosystem** and services
2. On-prem / closed systems create **organisational silos** and connectivity challenges
3. **Vendor lock-ins** create unbalanced power dynamic with suppliers
4. **Cost efficiency** and model suitability

## On-prem infrastructure to manage and interconnect



# Time Series Database

## Pro's

1. **Cloud-native, modern** and **open technology** supports the Industry 4.0 transition
  - a. Application versatility
  - b. Development agility
  - c. Easy integrations and ecosystem support (APIs, connectors, protocols, tools)
2. **Flexible query** language and **advanced analytics** capabilities
3. Excels in **real-time processing**
4. **Cost efficient** and scalable commercial model
5. **Scalability** and **storage efficiency**





# Time Series Database

## InfluxDB Additional Pro's

1. Supports **massive scale**, incl. Unlimited cardinality
2. **Hybrid deployments** (edge, on-prem, cloud)
3. Edge Data Replication (**store-forward**) capabilities
4. Flexible **schema-on write**
5. Hot and cold **storage tiers**
6. Vast **community** and network of **integrations** and partners
7. Many more...



# Time Series Database

## Con's

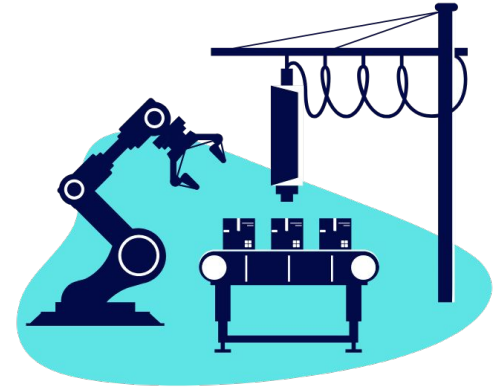
- Not domain specific
- Build vs buy
  - Effort
  - Learning curve
- Require ecosystem to fulfil industry-specific capabilities and a complete e2e solution



# The Future of Industrial Data: Industry 4.0...

- IoT
- Cloud computing
- Edge computing
- 5G networking
- AI and ML
- Cybersecurity
- Digital twins
- Real-time analytics

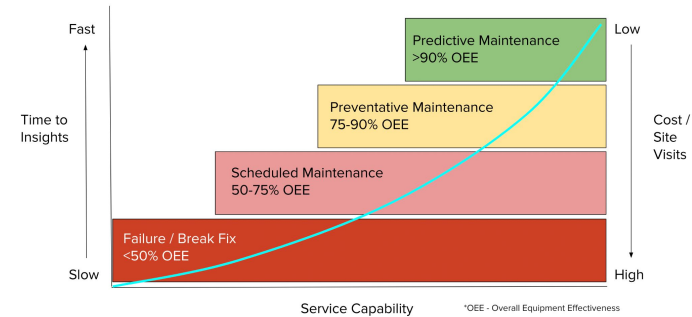
- ✓ Operational efficiency
- ✓ Data Analytics & insights
- ✓ Customisation & flexibility
- ✓ Quality control
- ✓ Supply chain optimisation
- ✓ Workplace safety
- ✓ Connectivity & collaboration
- ✓ Energy & sustainability
- ✓ Cost controls & efficiency



# The Future of Industrial Data: Industry 4.0...

- IoT
- Cloud computing
- Edge computing
- 5G networking
- AI and ML
- Cybersecurity
- Digital twins
- Real-time analytics

- ✓ Operational efficiency
- ✓ Data Analytics & insights
- ✓ Customisation & flexibility
- ✓ Quality control
- ✓ Supply chain optimisation
- ✓ Workplace safety
- ✓ Connectivity & collaboration
- ✓ Energy & sustainability
- ✓ Cost controls & efficiency



## Targeted Personas

### Typical users

OT (Operational Technology) Engineers

OT Site Managers

IT Architects

IT Project Managers

DevOps Engineer

Software Engineer

### 'Newer' users

Data Scientists

Data Engineers

Business Analysts

## Challenges

Operational Efficiency, Actionable Insights

Sustainability, Quality Control

Cost Efficiency

Connectivity, Accessibility, Data Siloes

Diversity of data sources, Inconsistent Schema

Data Volume, Dimensionality, Resolution

## Business Outcomes

Improve OEE, Proactive Operations

Improve Energy Management, wastage, products & processes

Lower Cost of Storage, Lower TCO

'Single Pane of Glass', Durable & secure data capture/sync

Accelerate Time to Value

Real-time OT Monitoring

## InfluxDB Capabilities

Automate Predictive Analytics & OT processes

Real-time analytical queries at scale

Interoperability with BI & data science tools

Compressed data storage on cloud object store

Zero retrieval cost to query historical data

Intelligent Edge with InfluxDB Edge

Edge Data Replication

Handle late data arrivals and duplicate data

Telegraf with 300+ plugins

Open & Flexible Data ingestion

Schema on Write

Purpose built for time series data

High speed writes & queries at scale

Unlimited data cardinality

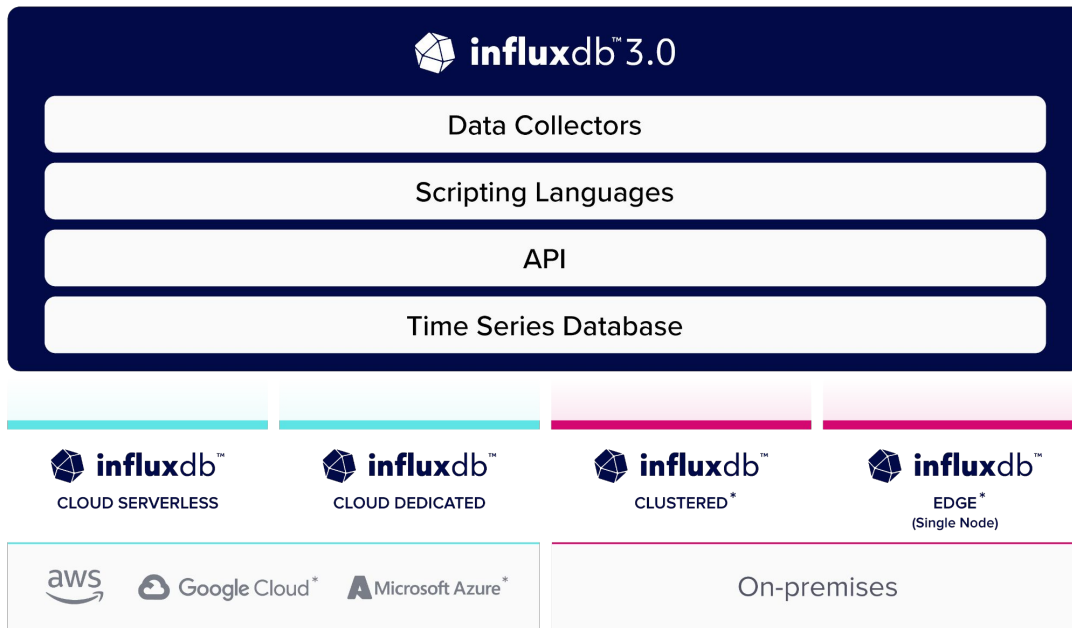
# InfluxDB 3.0

InfluxDB 3.0 is the new database & storage engine launched in 2023, and forms the new core of the platform itself

# What is InfluxDB 3.0?

Database & platform for handling time series data at massive scale

We encourage that customers evaluate the **InfluxDB version** that is desired for production



\* Availability to be announced

Setup specifications corresponding to the edition can follow...

# Fast time series platform with built-in analytics



**One data store  
for metrics,  
events & traces**

**Faster time to  
learning and  
insights**



**Designed to  
deliver  
sub-second  
query responses**

**Deliver awesome  
end user  
experiences**



**Keep data  
forever on  
low-cost object  
store**

**Lower overall  
costs of  
ownership (TCO)**



**Faster Time to  
Awesome® with  
SQL, InfluxQL**

**Improve  
developer  
productivity**



**Open &  
Interoperable  
with Data  
Ecosystem**

**Improve data  
efficiency**



# Unlimited Cardinality

Unlimited Cardinality with InfluxDB allows Customers to capture all of the required metadata as tags for their sensors and OT devices without limitations and Improves Efficiency.



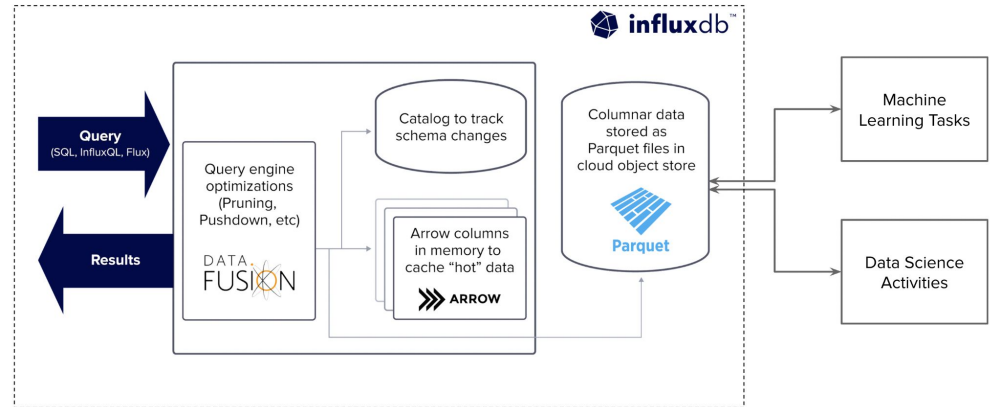
# API integration



InfluxDB APIs enable interoperability with other data and ML tools. Customers can automate Predictive Maintenance and other processes Improving OEE (Overall Equipment Effectiveness)

# Interoperability with data tools

InfluxDB persists data as Apache Parquet files which allows interoperability with machine learning and data science tools and therefore enabling Customers to Improve Process Efficiency.



# In Memory Columnar Store

InfluxDB caches the recently ingested or queried data in “in-memory” hot tier. With this capability, Customers can get real time insights on “live” incoming data [Improving Their Ability to perform Proactive Operations.](#)



## Hot

---

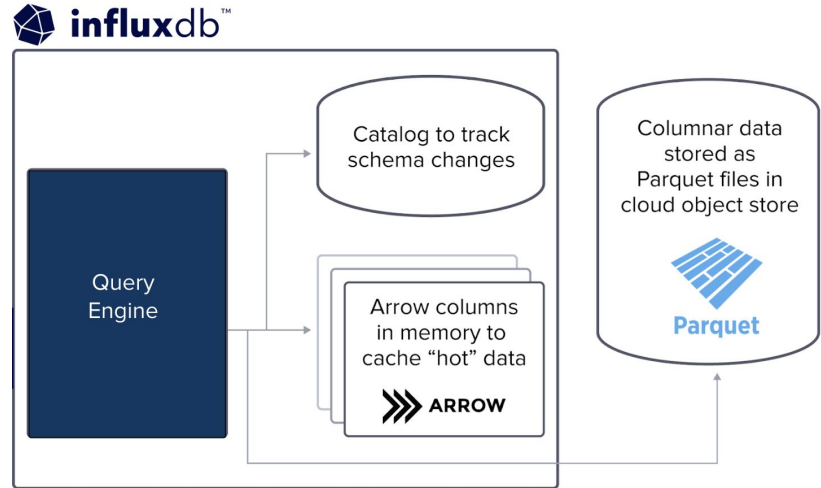
**Data in memory**

*seconds, minutes, hours*

Optimized for low  
latency analytical  
queries

# Persist Data on Cloud Object Store

InfluxDB persists aged data with maximum compression on an inexpensive cloud object store. With this capability, Customers can meet their long term data retention requirements while Lowering Storage Costs.



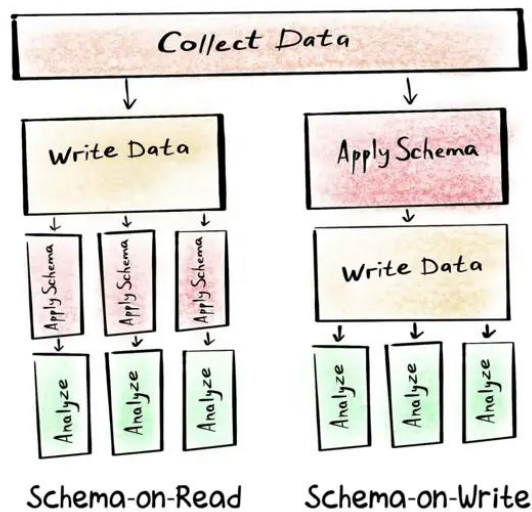
# Zero Retrieval Cost Capability

Unlike other solutions where there is an additional effort to load historical data archived in low cost locations and make it available for normal queries, InfluxDB allows Customers to make no additional effort and just **query the historical data like any other data**, Lowering their Total Cost of Ownership



# Schema on Write Capability

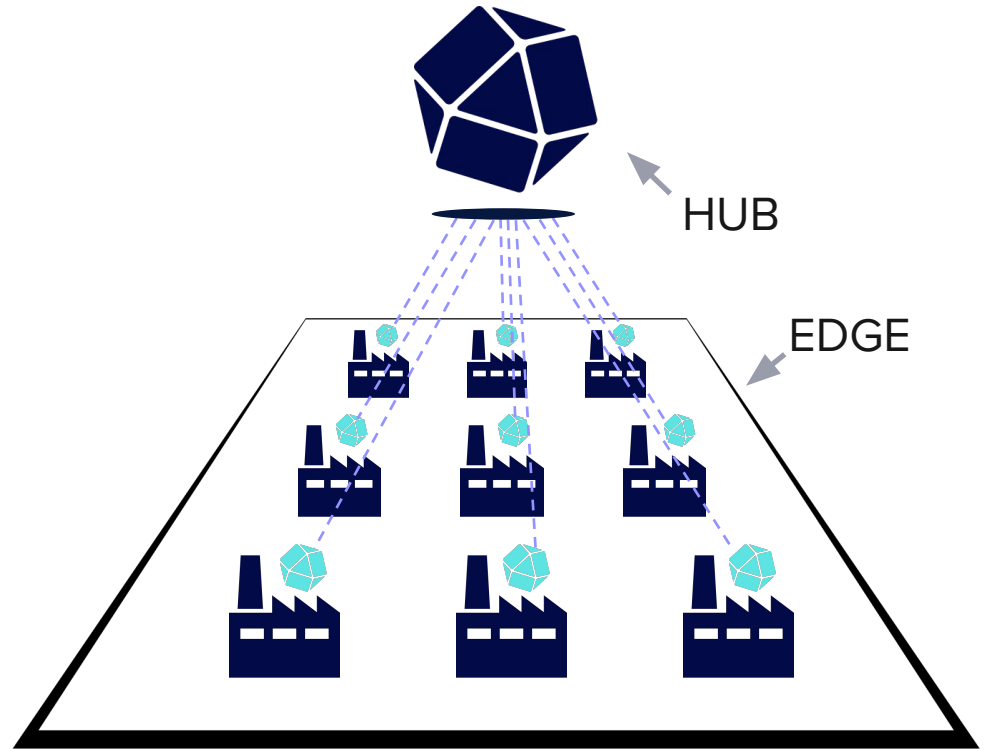
With InfluxDB's Schema on Write capability, Customers do not need to pre-define a schema in order to ingest the data. This massively Increases Developer Productivity, especially at scale with a large volume and variety of disparate sensors and OT devices involved.



# Edge Data Replication

Edge Data Replication (EDR) makes it possible to securely replicate data from InfluxDB at the Edge to a hub. EDR removes a lot of complexity around setting up, and maintaining Edge to hub replication and thus enabling Customers to bring OT and IT closer and eliminate data silos

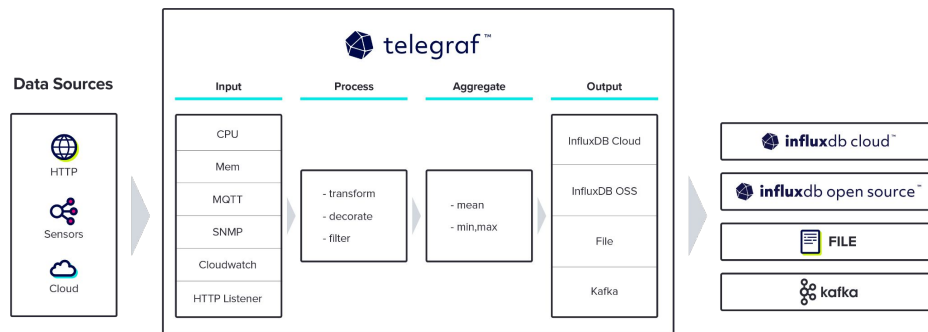
<https://killercoda.com/influxdata/course/Training/influxdb-hybrid-iiot>





# Telegraf Plugins

With our Open Source Telegraf Agent, Customers are able to Accelerate Time to Value by utilizing MQTT, OPC-UA or any of the 300+ plugins offered to collect metrics from their sensors, OT machines and devices, at granular frequencies.




# IIoT Partnerships & Integrations

Applications



Grafana | Seeq | Tableau | Factory | Clarify | Apache Superset

Data Persistence



InfluxDB | Edge | Data Centre | Cloud

Middleware



Telegraf | Kepware | Hivemq | HighByte | Apache NiFi | WinCC Open Architecture Siemens | Cogent DataHub

Processes & Assets

ICS / SCADA | PLCs | Robotics | Sensors / Devices | Plants / Factories



io-base  
value-added data

thingworx®

Ignition!  
by inductive automation

ctrlX  
AUTOMATION

akenza.io

InfluxDB

Platforms



# Customer Example



TERĒGA  
SOLUTIONS

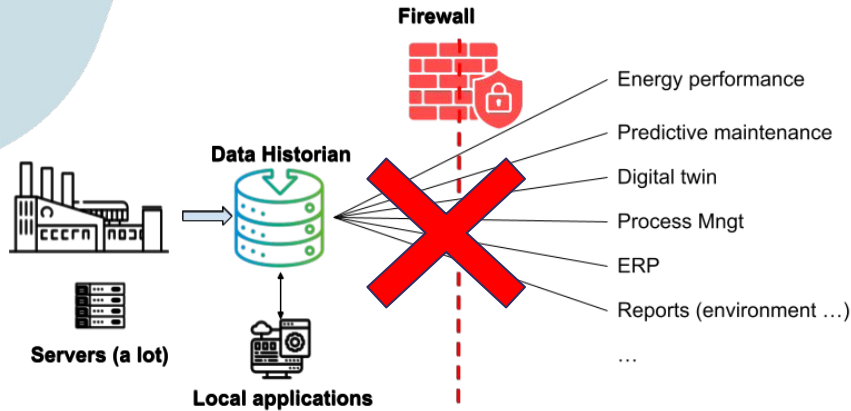
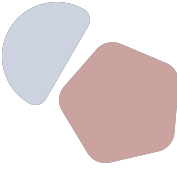


influxdb™



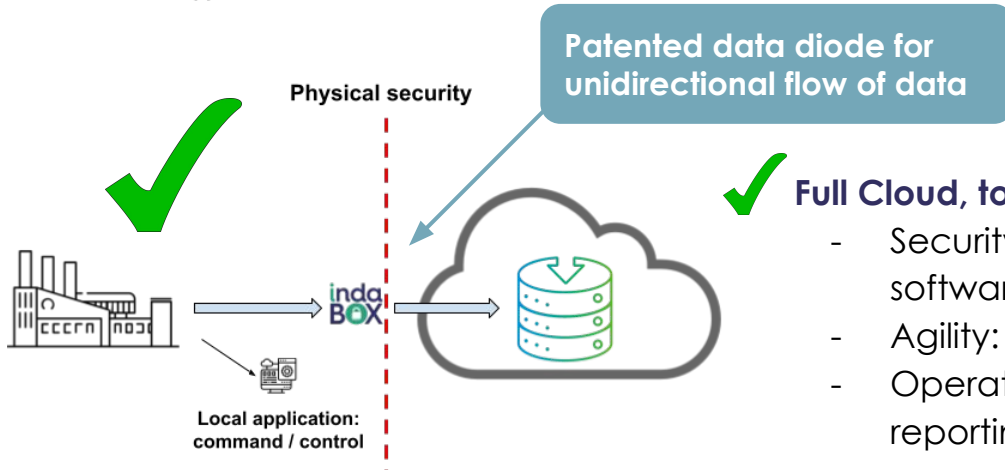
io-base  
value-added data

# Paradigm shift: data historian in the cloud



## Data cherry picking

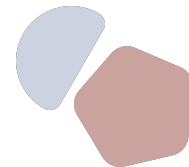
- Security issue
- Bandwidth
- Data scattering/discrepancy
- Impossible to maintain



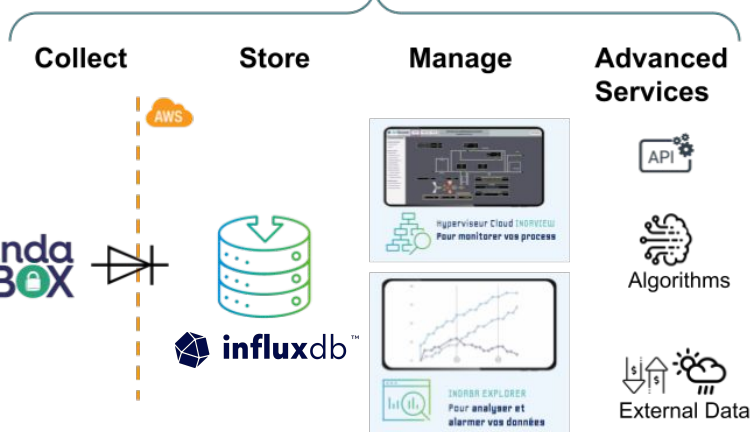
## Full Cloud, to improve

- Security: no local IT, firewalls exceptions or software to install
- Agility: data access simplified
- Operations: monitoring, on-call, analytics, reporting

# IO-Base now



## Secure, Performant, Agile, Industrial Data Twin platform



- Centralized master data
- Minimum onsite infrastructure: only collection/transmission - no maintenance
- Ease of data sharing
- Hardware / network agnostic
- Highest level of cyber-security



Non-hackable Industrial Site

Indabox

Secure data collection

inda  
BOX  
by TEREGA



secure  
by design  
bringing your existing  
equipment online

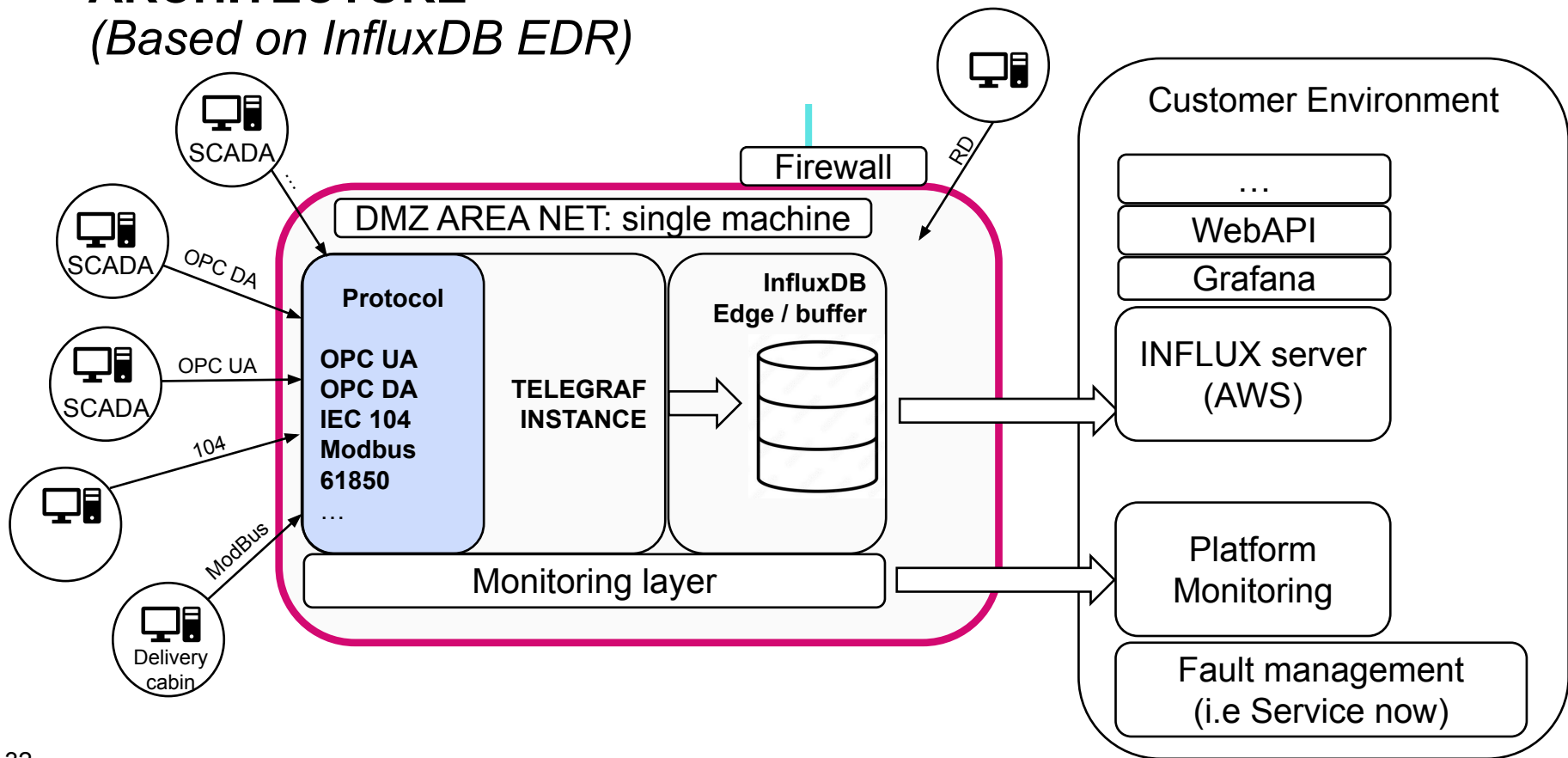




# FTSE 500 Energy Company

# ARCHITECTURE

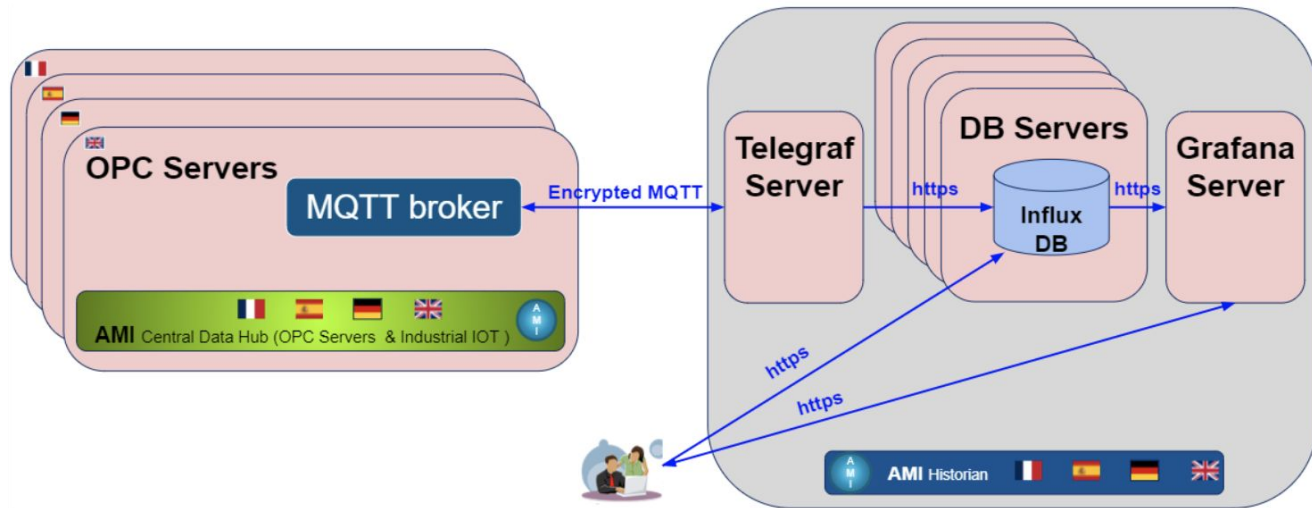
(Based on InfluxDB EDR)







# **FTSE 500 Aerospace Company**



- Smart factory IIoT use case
- Globally distributed plant network
- with scalable MQTT-based ingestion architecture

# Getting started

## Sign up

Influxdata.com

Get InfluxDB

Via cloud marketplace



## Learn

- ✓ Self-service content
- ✓ Documentation
- ✓ InfluxDB University

# InfluxDB Community Resources



Slack: [influxdata.com/slack](https://influxdata.com/slack)



Community Forum: [community.influxdata.com](https://community.influxdata.com)



Docs: [docs.influxdata.com](https://docs.influxdata.com)



InfluxDB University: [university.influxdata.com](https://university.influxdata.com)

# InfluxDB Resources

## Webinar: Gain Better Observability with OpenTelemetry and InfluxDB

*Leverage OpenTelemetry and InfluxDB to collect and analyze metrics, logs, and traces, enabling better anomaly detection, root-cause analysis, and alerting.*

Watch now

[bit.ly/3qhemCw](https://bit.ly/3qhemCw)

## Save 96% on Data Storage Costs:

Learn more

[bit.ly/3NJEcGZ](https://bit.ly/3NJEcGZ)

## Run a Proof of Concept:

Learn more

[bit.ly/3puRsal](https://bit.ly/3puRsal)





---

THANK YOU