Amplifying the Value of Timeseries

Combining Seeq and Databricks

The Presenters

Do these guys even know *anything*?

Chris Herrera

Head of API & Interoperability

Seeq

Brent Railey

Chief Data & Analytics Officer

Chevron Phillips CHEMICAL — Performance by design. Caring by choice.™

Who is Chevron Phillips Chemical?

"Performance by design. Caring by choice." ™
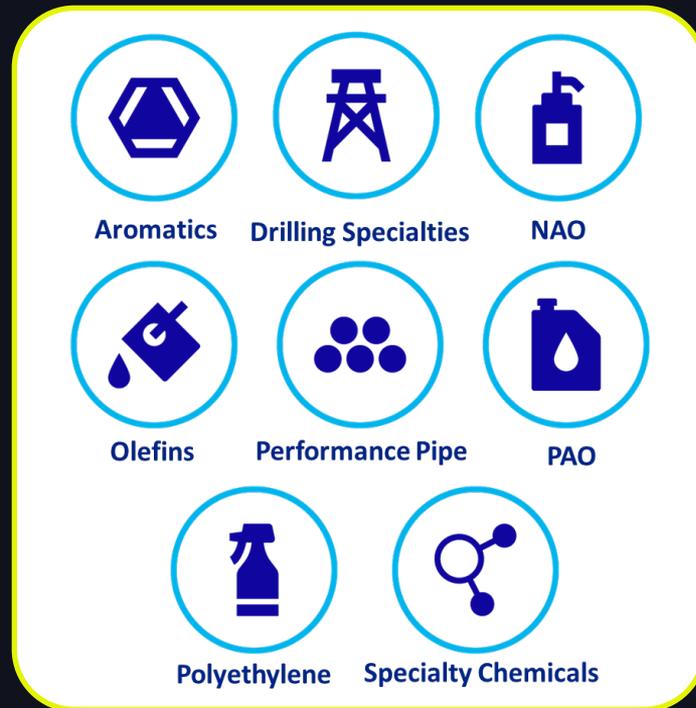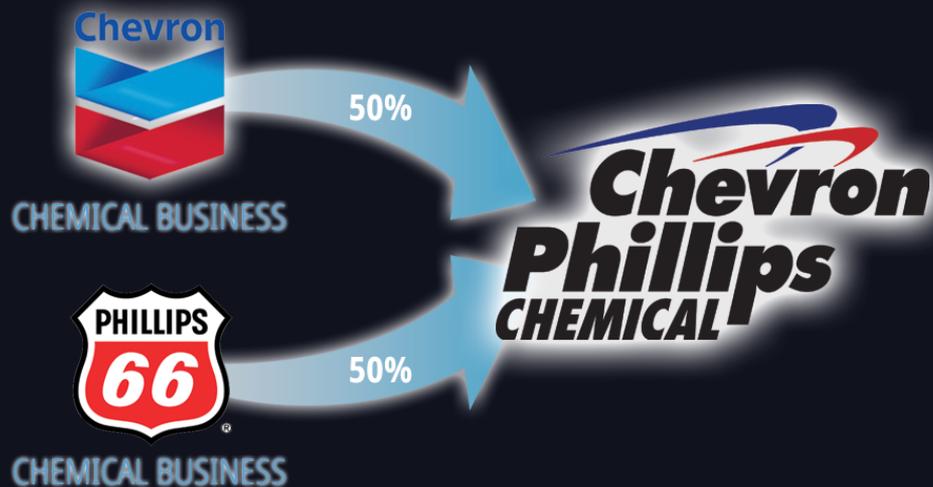
Who is Seeq?

"Accelerating a Modern, Sustainable Manufacturing Future"

2013 Year Founded by Industry Experts	**200+** Domain Experts with Deep Industry Expertise	**60+** Data Sources	**250+** Global Customers
$2B+ Validated ROI	**2.6T** Samples/Day	**1.8B** Signals	**40** Countries

What is Seeq?

Pure Timeseries Magic!

Workbench

Designed for *engineers* engaged in diagnostic, descriptive, and predictive analytics with process manufacturing data who need a simple, but visually powerful interface to *interactively* and iteratively with *timeseries* data.
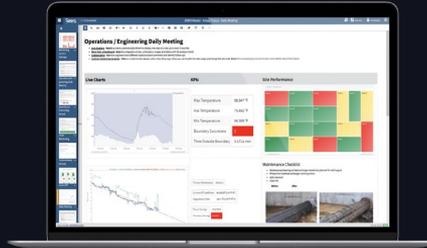
Organizer

Use the analyses accomplished in Workbench to regularly monitor KPIs related to manufacturing processes--in real-time.
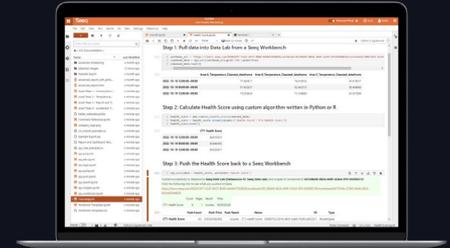
Data Lab

Give power to your engineers to experiment with machine learning with a notebook like interface that is tightly integrated with Seeq's data connectivity and analytic capacity.

Why Seeq?

An illustrative example...

Why Seeq + Databricks?

What does the combination bring to the Customer?

Business Intelligence & Analytics

MLOps

Scalable Compute for ML and Data Service

Data Scientists   **BI Experts**

+

SME-led Operational Intelligence & Analytics

Access to stranded OT data stores

Operational Experts (Process & Equipment)

=

Enterprise Intelligence

DATA+AI SUMMIT

Machine Learning

©2024 Databricks Inc. — All rights reserved

Why Seeq+Databricks for Timeseries ML?

What do Seeq and Databricks bring to the table for ML?

Seeq

Virtualization Layer
into OT Timeseries Data Sources

Cleansing & Gridding
Disparate Signals in Time
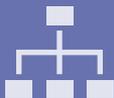
Asset Hierarchies
help Scale Efforts

databricks

mlflow™

Unity Catalog
Does this need explanation? ☺

Distributed Computing
Parallel computing to speed up research efforts.

Example: CPChem's Digital Rheometer

A Real-world Seeq/Databricks ML Implementation

Objective
Inference the Melt Flow
(Inverse of Viscosity at a 190ºC)

Hopper

Fluff

Heaters

Nozzle

Screw

Melted Polymer

Barrel
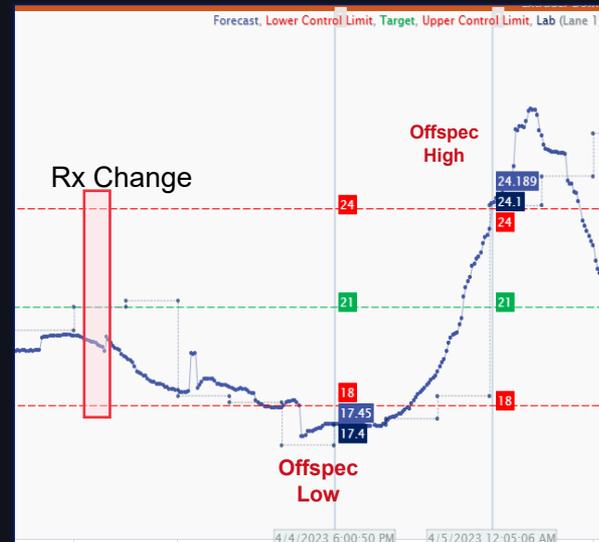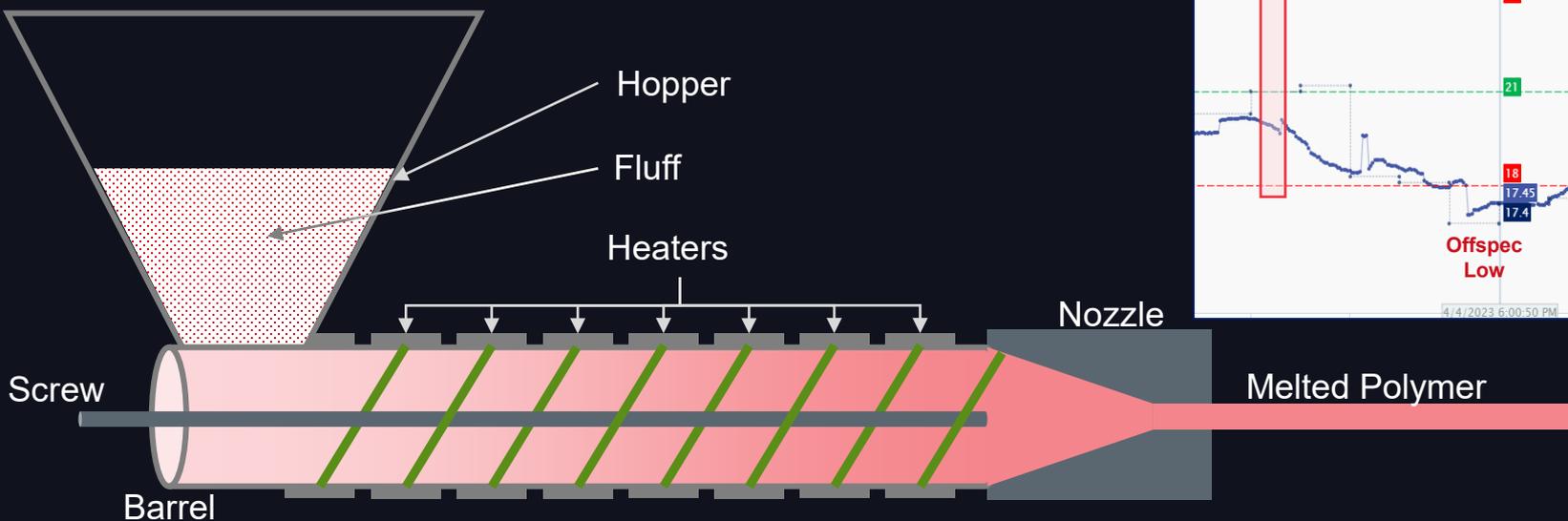
Digital Rheometer Architecture

6. Access inferences in Seeq

5. Model Management

User Interface

Edge | Cloud

Seeq

Seeq Remote Agent
Connect, Compress, Secure

Seeq
Transform, Contextualize, Cleanse, and UI to Results

Seeq SPY MODULE

Databricks Spark
Incremental ingestion and ETL into Delta Tables

MLOps in Unity Catalog

Databricks Serverless SQL

3. Further Cleansing, Training, Inference, Environment Control

4. Storage of Training Data & Inferences

1. Process and Laboratory Data Sources

2. Asset Hierarchies, Calculations/Feature Engineering

DELTA LAKE

IIoT
Self-Service
Analytics

Why Seeq for Self-Service Analytics?

What does Databricks do well?

You have **unconventional** timeseries data, or **IIoT** data that is not integrated with your process control network that you need to analyze. You need to store it and serve it.

DELTA LAKE

databricks

Streaming Data Processing

Live Tables receive, ingest, process, and store data.

Efficient and Scalable Storage

Liquid Clustering organizes data for efficient retrieval

Distributed Computing

Parallel computing to divide and conquer upon retrieval (and ingestion/storage optimization).

Why Seeq for Self-Service Analytics?

What does Seeq do well?
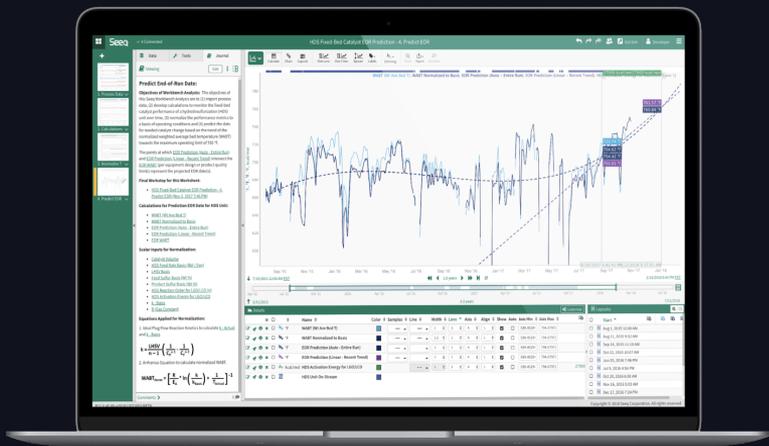
Built for Timeseries

Engineering and manufacturing problems live in this space and nothing beats Seeq at this.

Instant, Iterative Interaction

Allows engineers and analysts to easily interact with timeseries data and quickly iterate refinements.

Scale Analyses

Allows users to quickly scale and refine the same analytics to similar equipment in the field.

Connecting Seeq to Databricks

Example: Ethylene RTO

From the edge to the cloud!   (RTO = Real-time Optimization)


```
Process Data  -->  aspentech
                   Realtime Optimization  -->  Model Outputs
First-Principle  -->                       -->  Setpoint Recommendations
Models
Economic Data  -->
```

DATA'AI SUMMIT
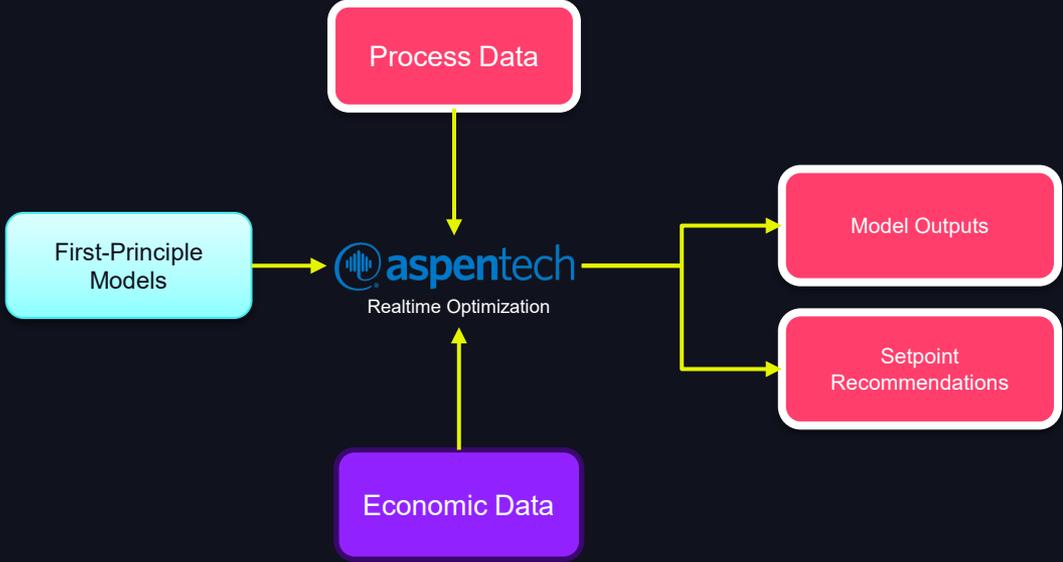
Ethylene RTO Data Architecture
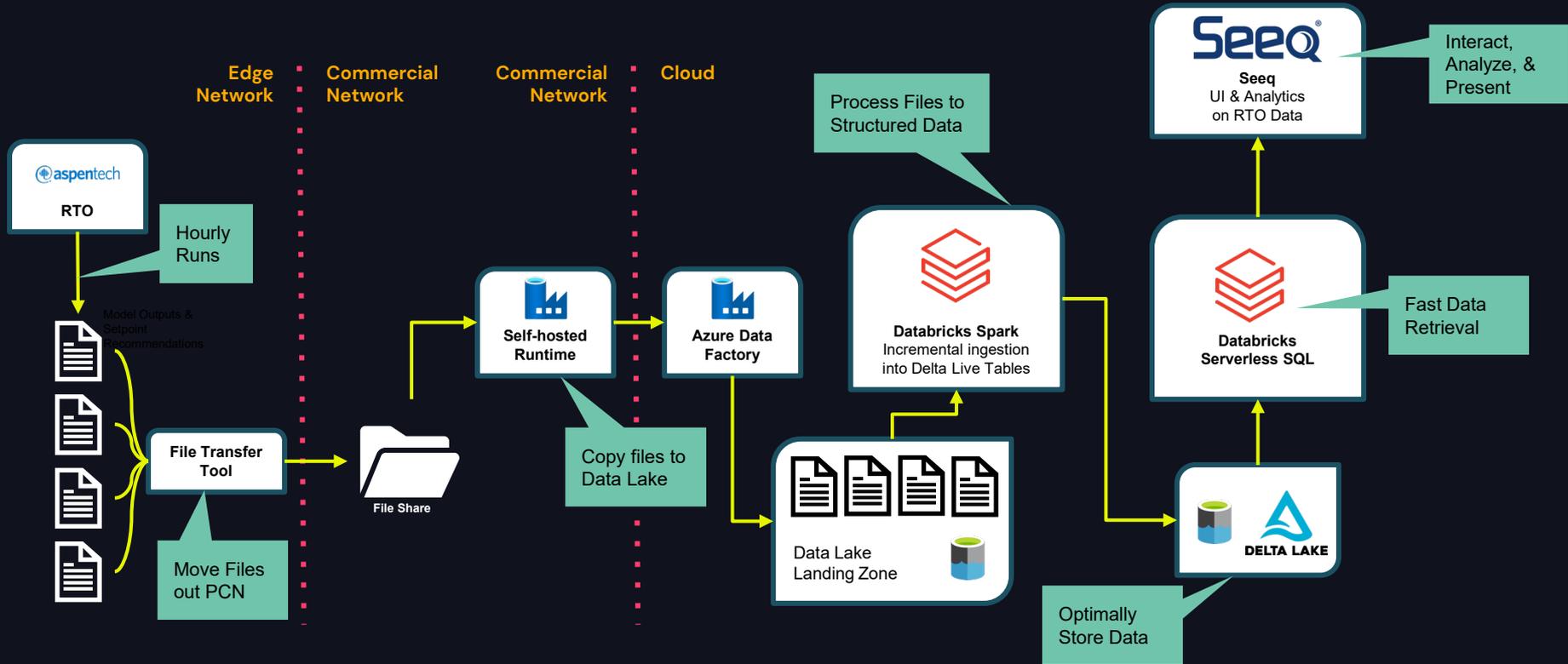
Enterprise Reporting

Example: Enterprise Mfg Data Reporting

Deliver manufacturing data to traditional business intelligence tools.

DATA AI SUMMIT
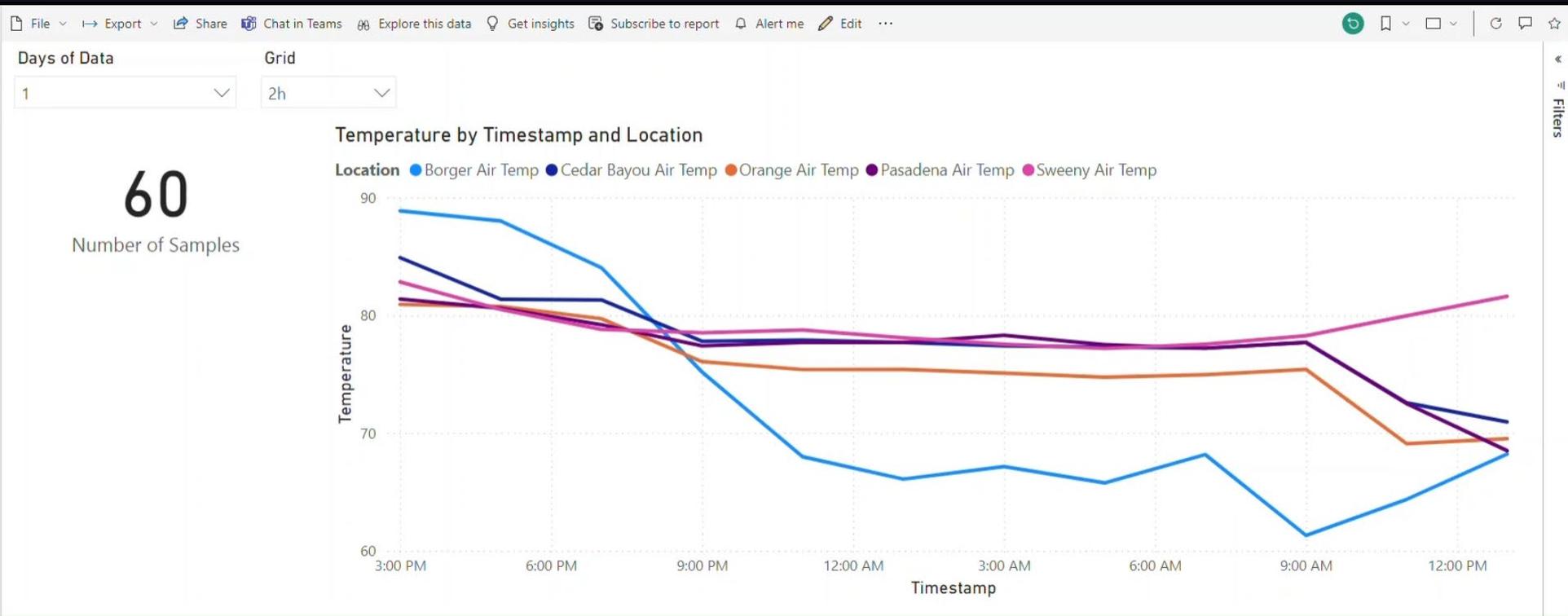
Reporting Architecture

Edge   **Cloud**

Seeq Remote Agent
Connect, Compress, Secure

Seeq Cortex
Transform, Contextualize, Cleanse

API Access

Databricks
SQL Compute via Tabular SQL UDF Calling Python UDF

Power BI

1. Connect to internal time-series, industrial datasources

2. Caching, Retrieval, API, Cleansing by SME

3. SQL compute, work distributed across cluster

4. Reporting Layer
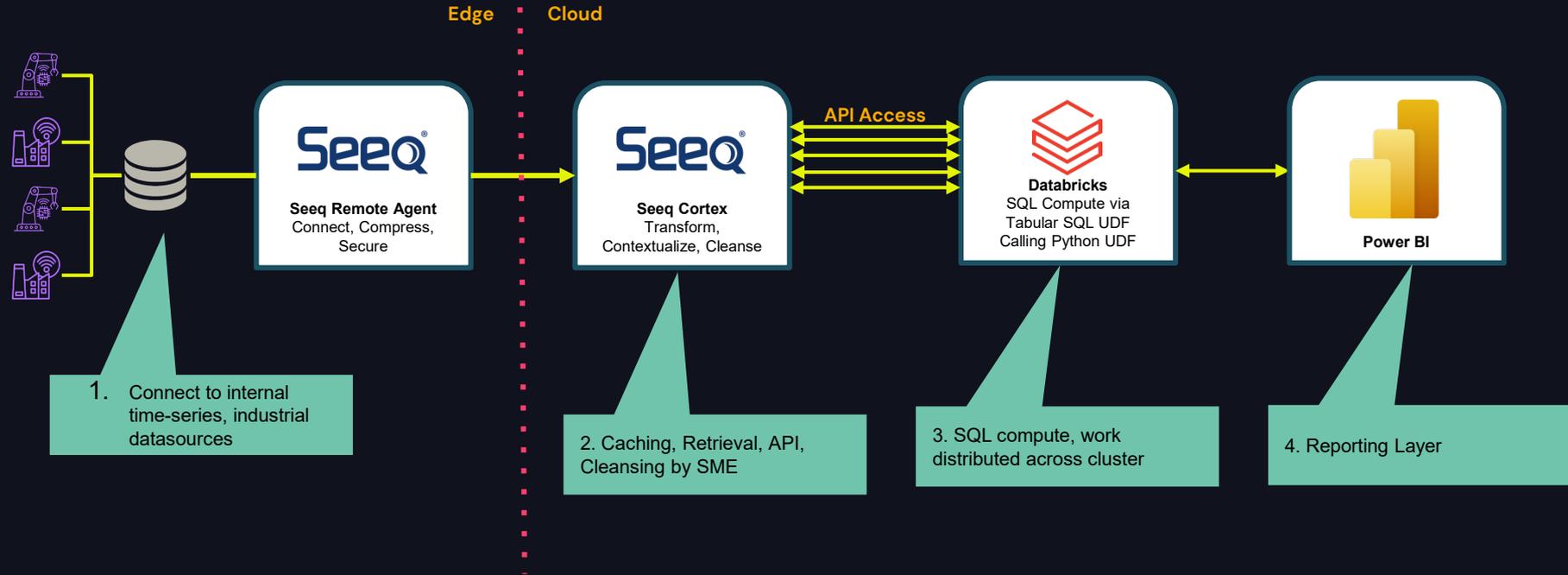
Conclusion: Seeq & Databricks

Better...Together ☺

Seeq

- The Best Timeseries Analytics UI Out There
- Simplifies Access to Hard-to-Connect Data Sources
- Excellent for "Gridding" Timeseries for Machine Learning Purposes
- Formula Language Suitable for Signal Processing and Training Periods

+

databricks

- IIoT Ingestion, Storage, Retrieval Power
- **Elastic**, Distributed Computing Capacity for ML and Data Processing
- **MLOps**: Model Track & Trace and Governance through Unity Catalog
- **Scaled Data Servicing** for Business Intelligence Platforms