

Amplifying the Value of Timeseries

Combining Seq and Databricks

The Presenters

Do these guys even know *anything*?

Chris Herrera



Head of API & Interoperability



Brent Railey



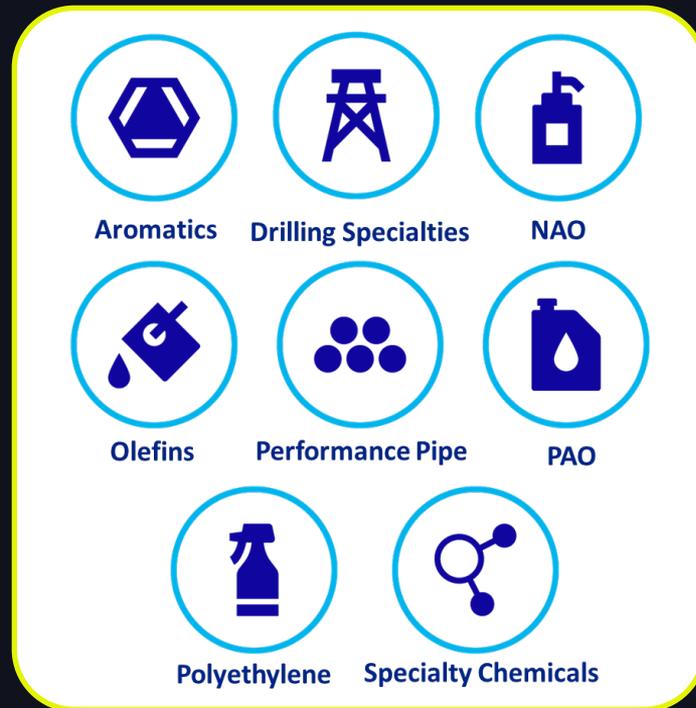
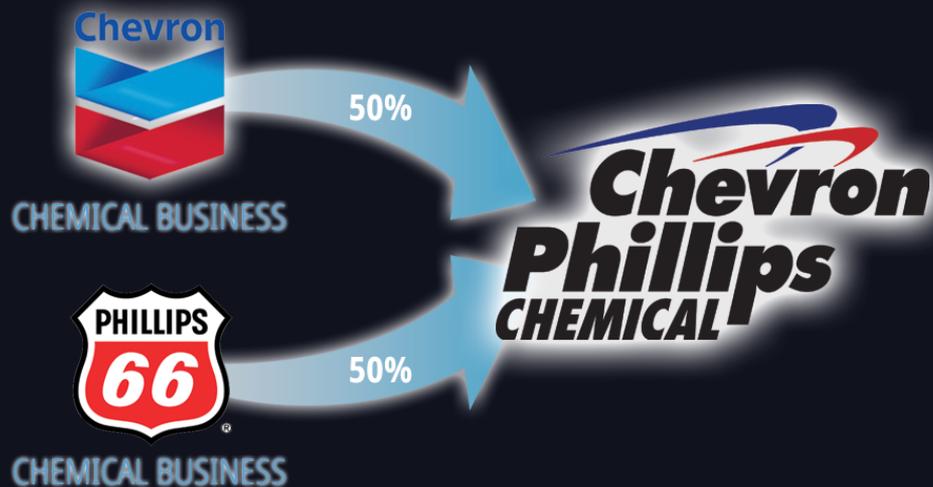
Chief Data & Analytics Officer



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



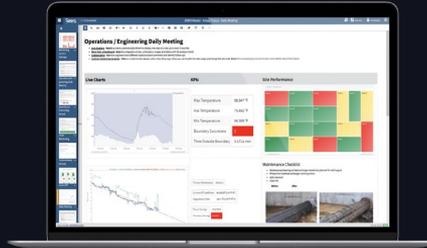
What is Seeq?

Pure Timeseries Magic!



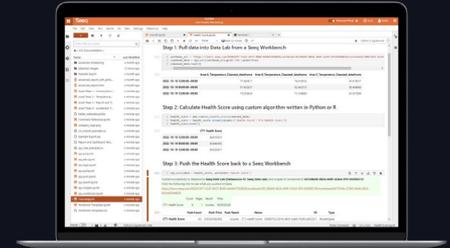
Workbench

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...

The screenshot displays the Seeq software interface. The top navigation bar includes 'Data', 'Tools', and 'Journal' tabs. The main workspace is titled 'Weight Drawdown Catalyst - databricks demo'. A callout box in the center of the workspace reads: 'Add data to this display via the Data tab. You can search for signals or find them in an Asset Tree'. The left sidebar contains an 'Objective' section with the text: 'Use the catalyst weight in the tank to estimate monthly consumption of catalyst (as it is drawn down)'. Below this is a 'Process' section with 18 numbered steps: 1. Bring in the measured process variable. (41WIS700.PV) 2. Move to a wider time frame. 3. Analyze the variable in small timesteps. 4. You can see the noise, so let's investigate further. 5. Given this information, smoothing would be beneficial. So lets go back to a wider timeframe to test the smoothing. 6. Let's smooth the signal using agileFilter. 7. Combine into a single axis, and contrast colors. 8. Zoom in a smaller time range. 9. Let's adjust the agileFilter. 10. Let's take the derivative of this signal. 11. Let's look at a wider view where a fill takes place. 12. You can see that the derivative shows when a fill takes place. 13. Identify fill periods. 14. Invert to draw periods. 15. Check longer time range. 16. Create a monthly condition. 17. Integrate over the derivative to get a running consumption, resetting every month and excluding fills. 18. Give me a value each month for last value in the running total. The bottom of the interface shows a 'Comments' section with 0 comments and a 'Feedback' button. The footer includes 'R63.1.0-4+202403060645' and 'Databricks | Server load: 0%'.

Why Seeq + Databricks?

Seeq®



databricks

What does the combination bring to the Customer?



Machine Learning

Why Seeq+Databricks for Timeseries ML?

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

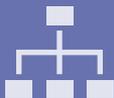
Seeq[®]



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.



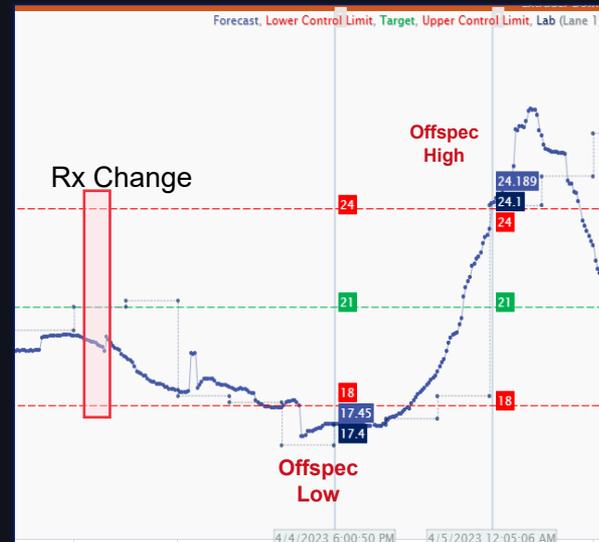
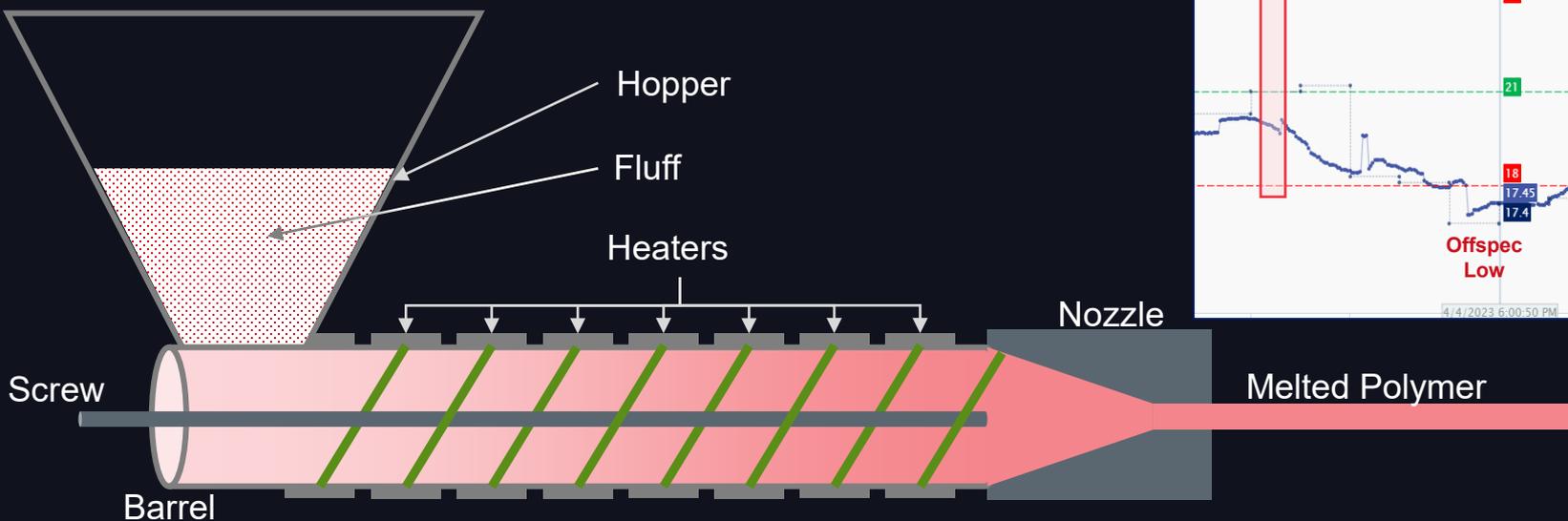
Example: CPChem's Digital Rheometer

A Real-world Seeq/Databricks ML Implementation

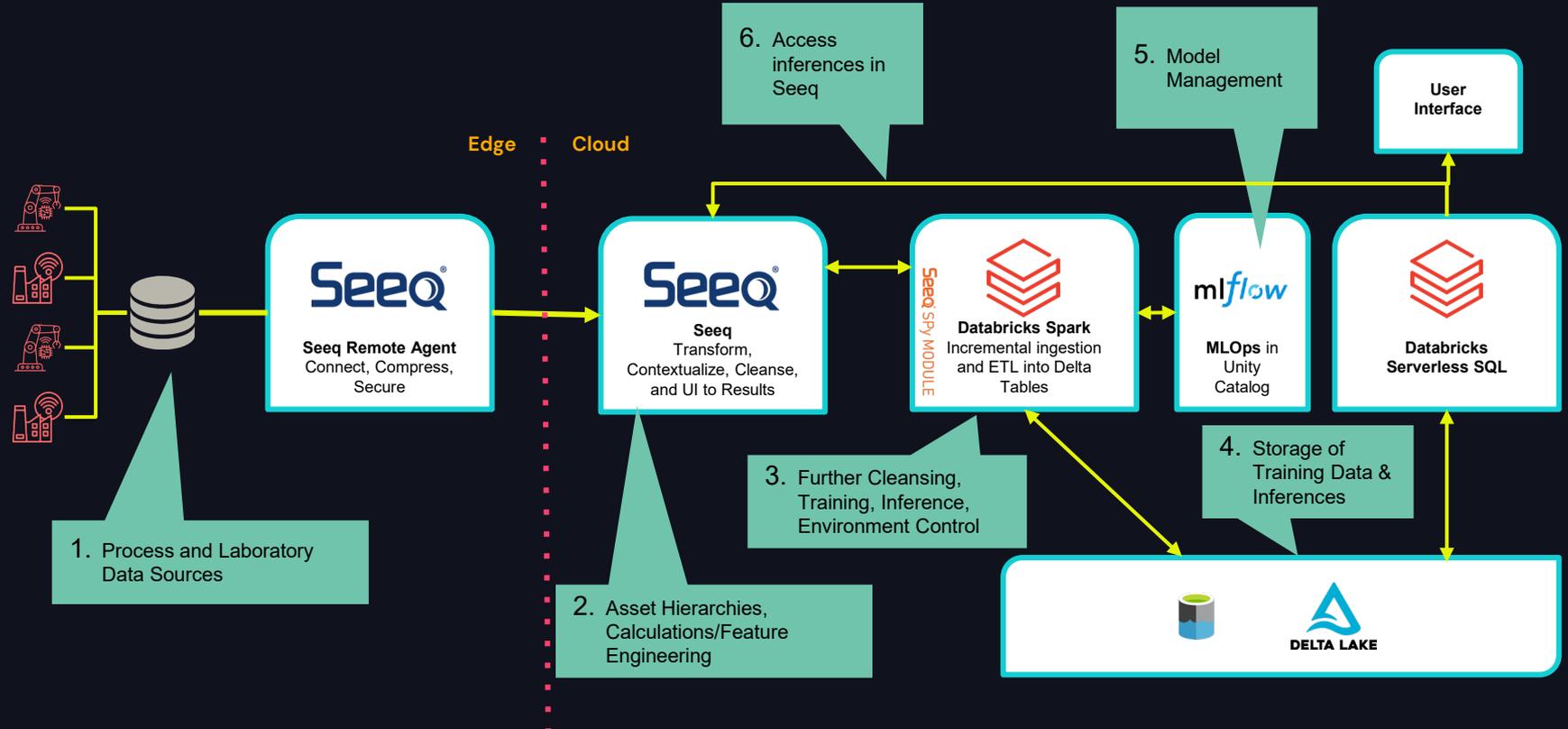


Performance by design.
Caring by choice.™

Objective
Infer the Melt Flow
(Inverse of Viscosity at a 190°C)



Digital Rheometer Architecture



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.



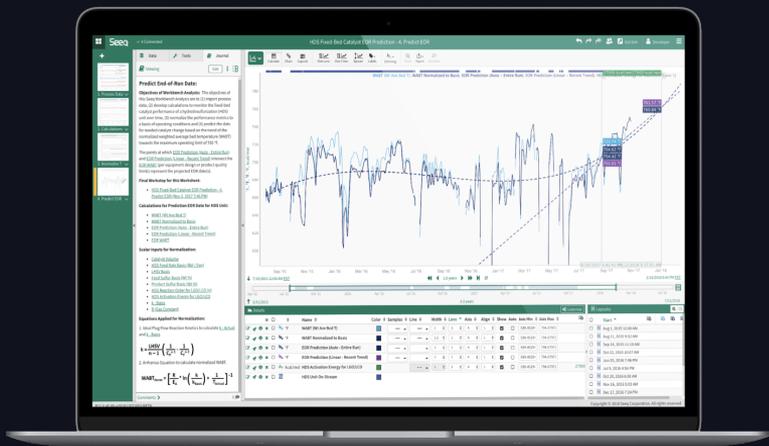
IOIO **Streaming Data Processing**
IOIO **Live Tables** receive, ingest, process, and store data.

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

The screenshot shows the Databricks Catalog Explorer interface. On the left, a sidebar lists various Databricks features like Workspace, Recents, Catalog, Workflows, Compute, SQL, SQL Editor, Queries, Dashboards, Alerts, Query History, SQL Warehouses, Data Engineering, Job Runs, Data Ingestion, Delta Live Tables, Machine Learning, Playground, Experiments, Features, Models, and Serving. The main area is titled 'Catalog Explorer' and shows a tree view of catalogs. The selected catalog is 'main.dbx_seeq_perf.databricksdatagen_tagdim'. Below the tree, there's a search bar for columns and an 'AI generate' button. A table lists the columns of the selected table:

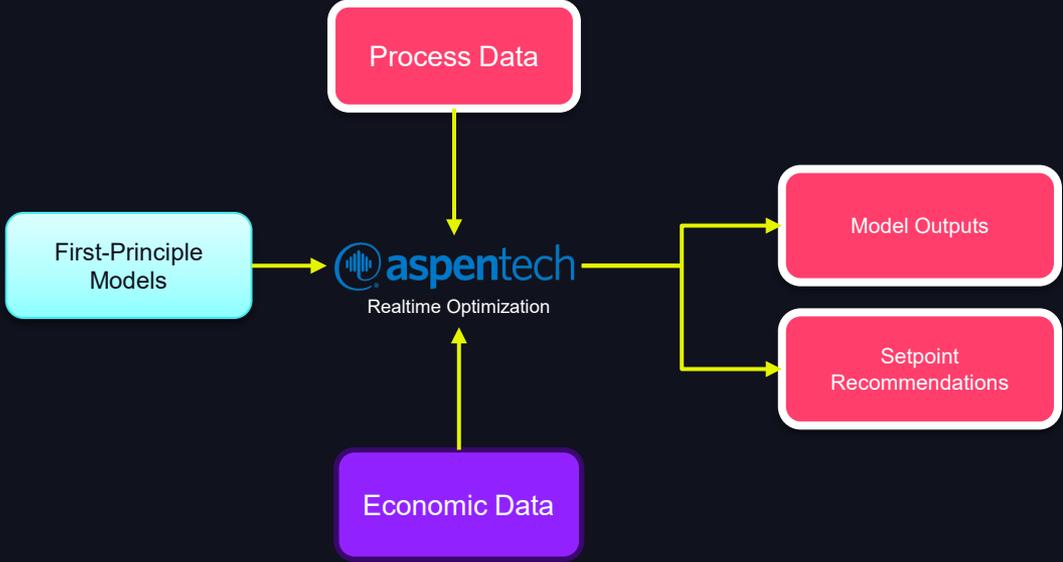
Column	Type	Comment	Tags
tag_id	int		
region_name	string		
facility_name	string		
equipment_name	string		
tag_name	string		
unit_of_measure	string		

On the right side, there's a 'About this table' section with metadata like Owner, Data source format, Last Updated, Popularity, and Size. Below that, there's an 'AI Suggested Comment' section with a preview of the comment and buttons for 'Accept', 'Edit', and 'Send feedback'.

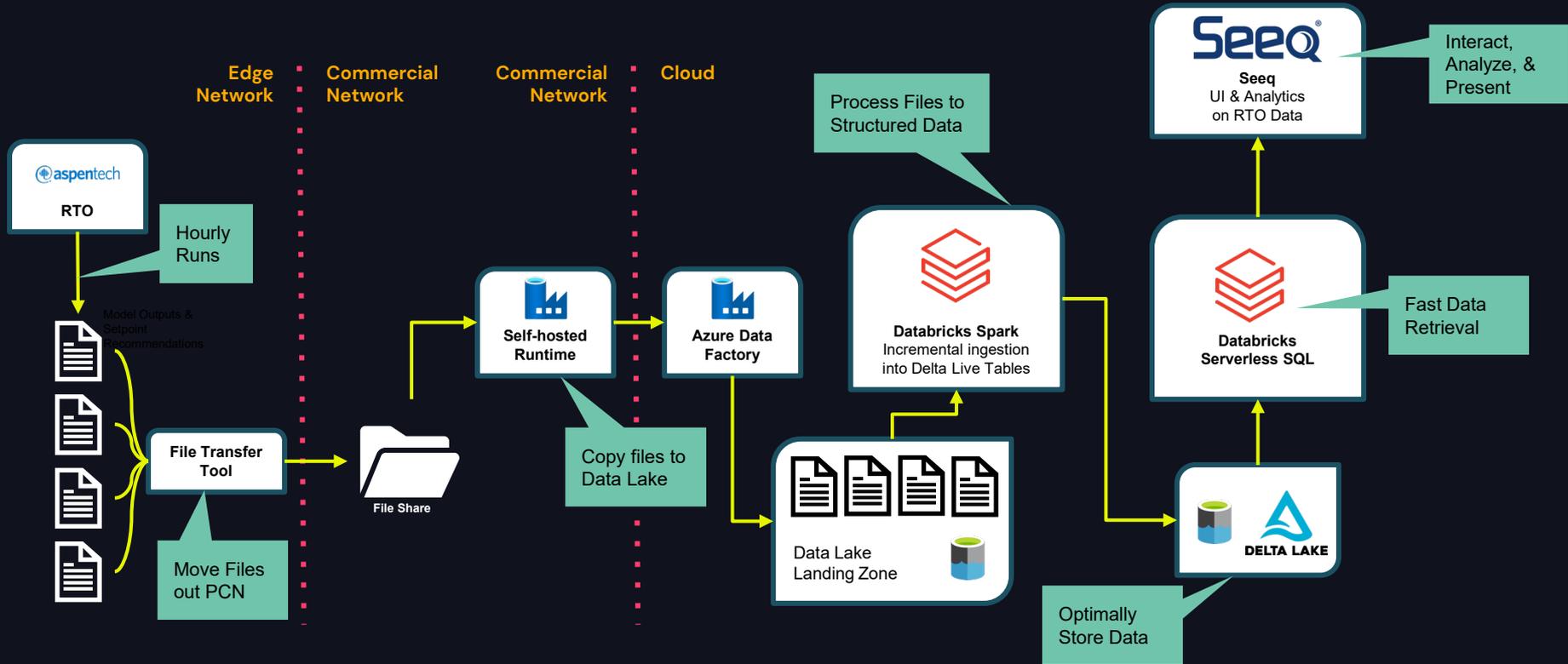


Example: Ethylene RTO

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



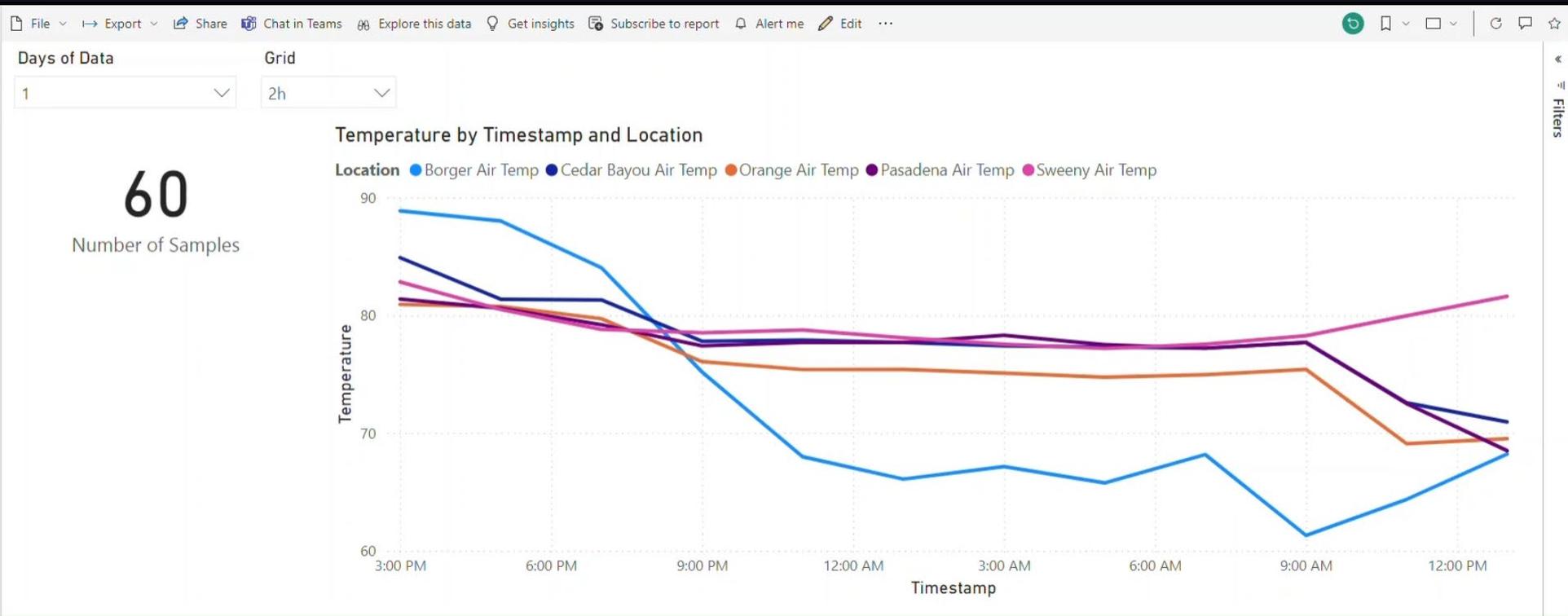
Ethylene RTO Data Architecture



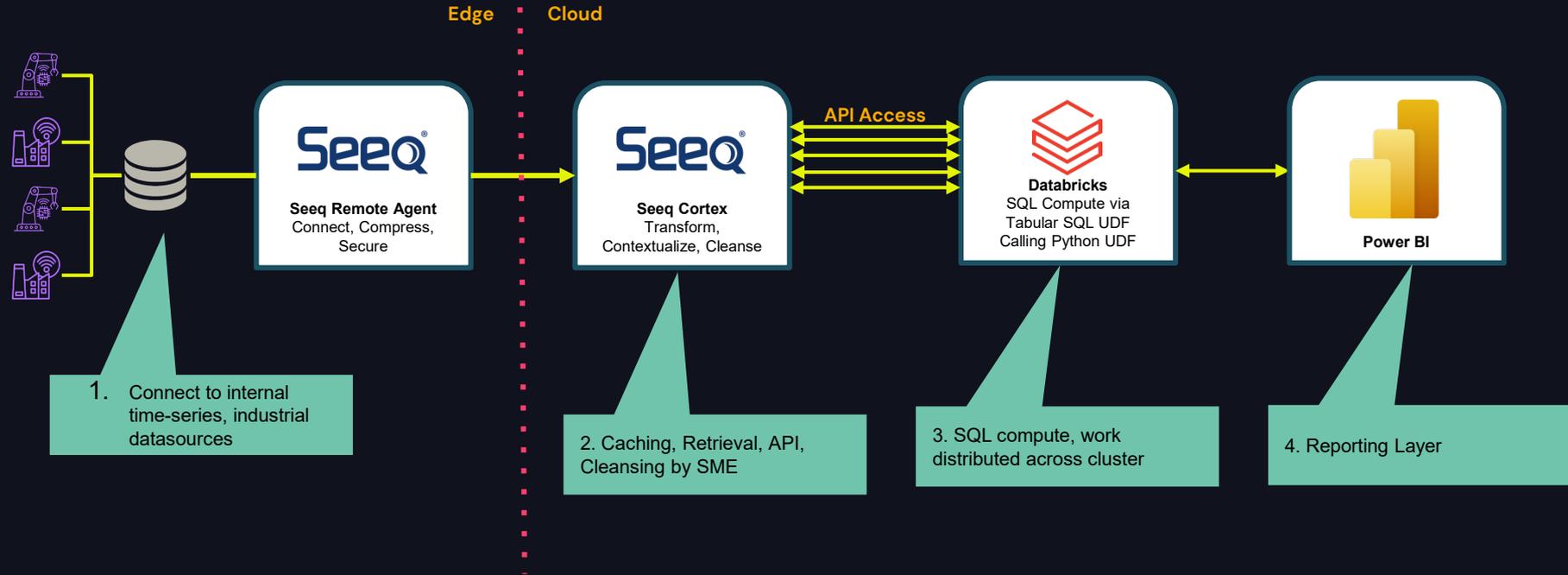
Enterprise Reporting

Example: Enterprise Mfg Data Reporting

Deliver manufacturing data to traditional business intelligence tools.



Reporting Architecture



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



IoT 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