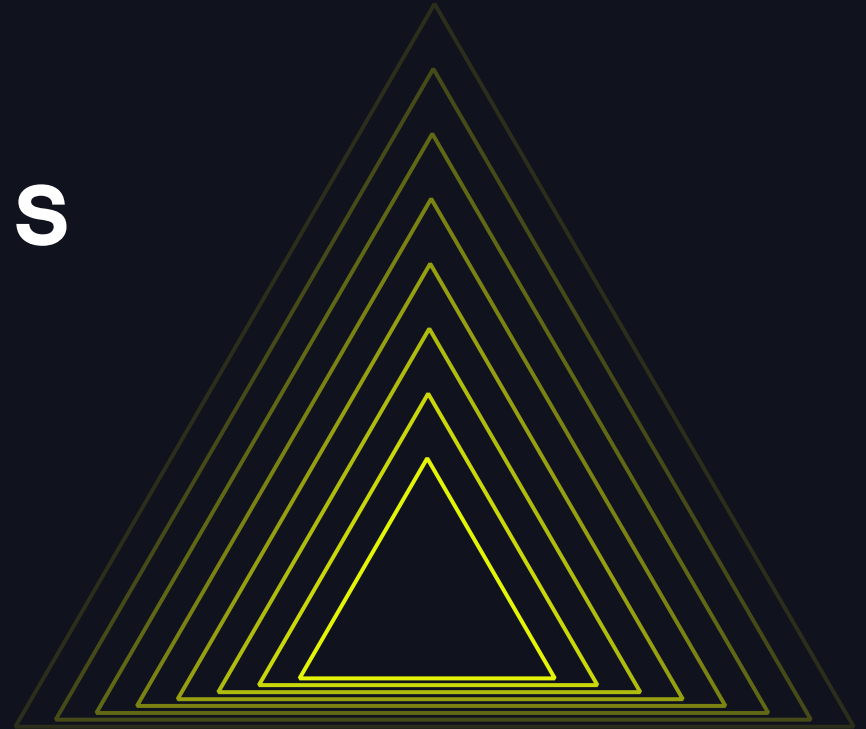


Honeywell Intelligrated's IoT Streaming Lakehouse

Mitul Desai, Quinton Stephens, Priyesh Agarwal
6/10/24



Overview By Function

The Journey (Through Different Viewpoints)



Business Development: Mitul

- Common challenges
- New services launched
- Growth trajectory



Architect & Build: Quinton

- Business pain points
- Solutions
- Upcoming innovation



Build & Optimize: Priyesh

- Development velocity
- Optimizations
- Robust integrations
- Databricks Alerting

Business Development :



Mitul Desai
6/10/24

Common Challenges

New Solutions Powered by Databricks Services

Historically data collection has been overlooked, non-standardized, limited in accessibility. To make better business decisions better data is needed.

Scalability/Performance

- Streaming data ingestion
- High velocity data exploration
- Unified Lakehouse architecture
- Universal BI solution
- Efficient data processing

Data Sharing

- Simple, secure data sharing
- Universal BI Data accessibility
- Robust query integration

Reliability

- Innate pipeline anomaly monitoring
- Robust data exploration and configuration
- Durable resilient architecture

New Services Launched

Streaming & Analytics Tailored to Customers



Distribution Center Data Collection (DC²)

- DC² connects siloed data systems, automates process, and distribution of KPI's
- Optimized data solution offering visibility into asset performance, reliability and capacity
- Universal platform for collecting, storing, leveraging and sharing data



High Fidelity Analytics (HFA)

- Standardized dashboards driving fast to market customer solutions
- Strategic KPI reports with improved accuracy, resolution and frequency
- Streamlined decision making workflows by identifying opportunities sooner

Growth Trajectory

Working Alongside Customers to Extract Greater Value From Their Data

Providing streaming data collection, partnered with standardized metrics for BI services removes data barriers and allows better business decisions to be made faster.

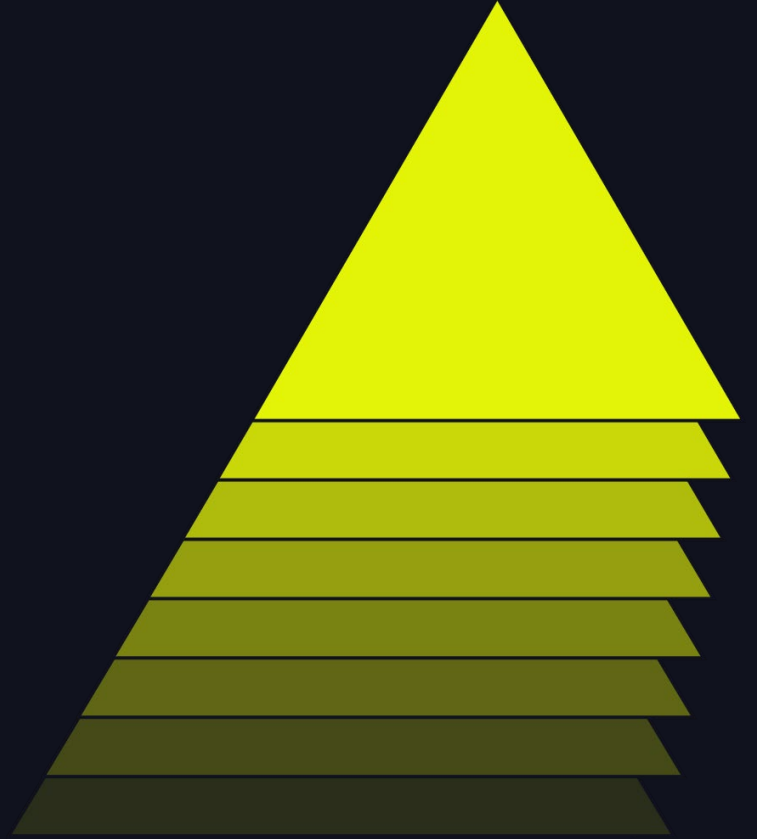
Expansion

- Unique Pilots
 - Customer variation
 - System variation
 - Application variation
- Contract offering
- Aftermarket offering
- Forecasted 40% Revenue Growth YoY

Transparency

- Data sharing (Delta Sharing)
- Driving Data Insight
- Lineage Metrics
- Processing efficiency

Architect & Build:



Quinton Stephens
6/10/24

Business Pain Points

Challenges to Solve on Behalf of the Team and Our Customers

High-Level Challenge Patterns

- Lack of standardization
- Need more scalable platform
- Reduce technical debt

Scalability/Performance

- Legacy infrastructure not designed to scale well
- Temperate/non-durable legacy ingestion processes
- Slow dashboard performance (HFA)

Self-Service Data Sharing

- Need to share data back to customers
- Previous solutions more involved than expected

Solutions

Laying the Foundations for New Innovation & Capabilities

High-Level Patterns

- Migrate & standardize
- Build on scalable, durable services
- Minimize tech-debt via standardization

Scalability/Performance

- New architecture designed to scale with expansion
- Highly durable ingestion & processing
- Increased dashboard performance

Self-Service Data Sharing

- Share data securely and efficiently with customers
- Delta Sharing enables higher-velocity sharing of data

Demo :

DLT Streaming

Demo :

Delta Sharing

Demo :

Federated Querying

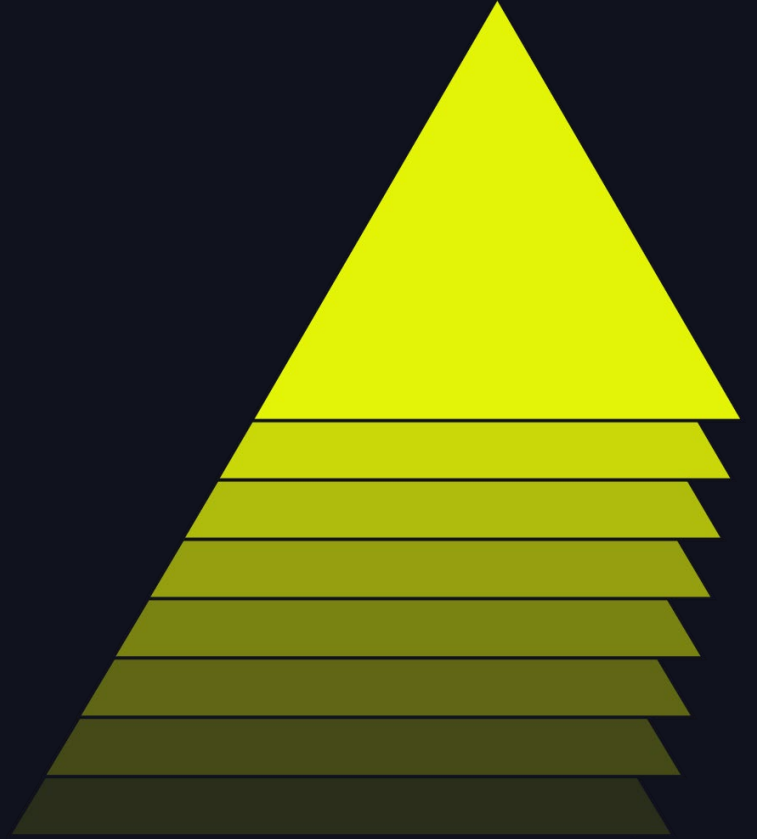
Incoming Innovation

Currently in Development:

- Domain-specific ML insights
- Function-specific GenAI apps
- Domain-specific ML apps

- This spot is great for a diagram, image, small table, code snippet or screenshot

Build & Optimize:



Priyesh Agarwal
6/10/24

Development Velocity

How Databricks has helped us develop faster



Before Databricks

- Data load time 6-7 Hours(Full Refresh)
- Setting up incremental refresh needed to be done manually(Watermarking)
- Fixed CPU usage
- No AI Assistant for reviewing bugs/issues

After Databricks

- Data load time 20-30 minutes (Full Refresh)
- Incremental refresh handled by Databricks automatically via Delta Live Tables
- Photon Acceleration can be used when required
- Databricks AI assistant for reviewing bugs/issues in notebook or getting suggestions

Optimizations

Databricks tools for better pipelining

- Use Materialized views for calculated Fields
- Use separate tasks for faster processing
- Development and Production modes for DLT Pipeline
- Photon Acceleration

The screenshot displays the Databricks interface for a job run. The main heading is "Sample job run". Below it, the breadcrumb path is ".Jobs > Job Example > Run 6". A notification at the top right states "Free trial ends in 3 days. Continue with a pay-as-you-go subscription by providing your billing information." The job title is "Job Example run" with a "Delete job run" button. The central part of the interface shows a task dependency graph with four tasks: Task1, Task2, Task3, and Task4. Task1 and Task2 are parallel tasks that both feed into Task3, which then feeds into Task4. All tasks are marked as "Succeeded". Task1 and Task2 have a duration of "4m 10s", Task3 has "9s", and Task4 has "11s". Each task card includes the notebook path "...s/yuz120@gmail.com/Notebook1" through "Notebook4" and the cluster name "Job_Example_cluster". On the right side, the "Job run details" panel shows: Job ID (17), Job run ID (6), Started (2022-03-13 12:11:47 EDT), Duration (4m 33s), and Status (Succeeded). Below this, the "Clusters" section lists "Job_Example_cluster" with details: Driver: i3.xlarge, Workers: i3.xlarge, 8 workers, On-Demand and Spot, fail back to On-Demand, 9.1 LTS (includes Apache Spark 3.1.2, Scala 2.12). At the bottom of the details panel are buttons for "View cluster", "Spark UI", "Logs", and "Metrics". A sidebar on the left contains navigation icons for home, jobs, clusters, recent, search, and workspace.

Robust Integrations

Integrating Databricks with Tableau, Azure for Clients

Business Intelligence

- Databricks Connector- Tableau
- ODBC/JDBC Connectivity- Tableau
- Faster Dashboard Performance

Azure

- Azure Databricks JDBC Connector
- Azure Data Factory

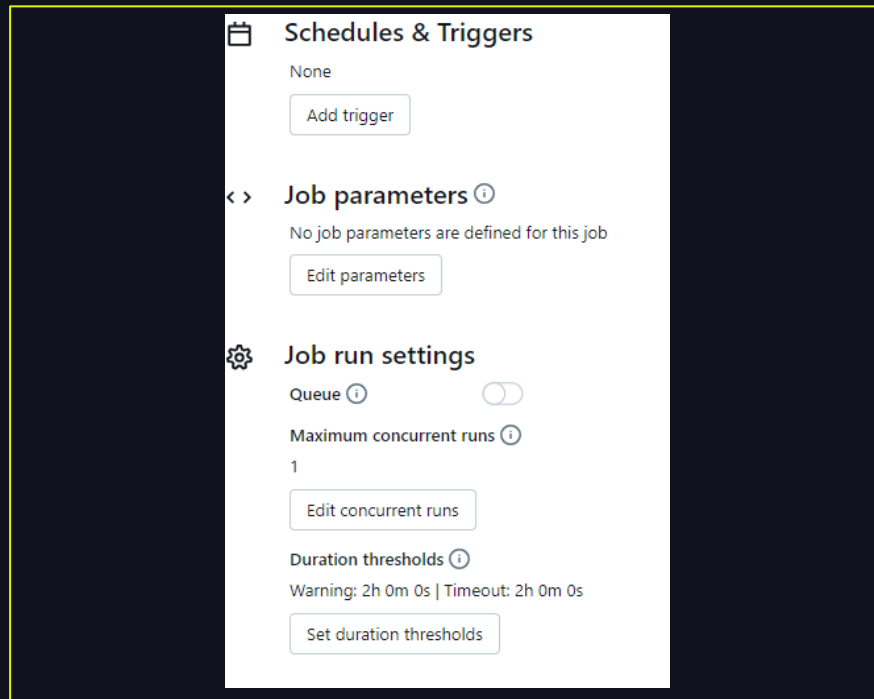
Delta Sharing

- Real-time Data Sharing
- Unified Data Access

Databricks Alerting

Notifications for Jobs and Pipelines

- Job duration warning notification
- Custom SQL Alerts
- Success/Failure notification



Demo :

Optimization and Alerting

Thank you!

— Mitul.Desai@honeywell.com,
Quinton.Stephens@honeywell.com,
Priyesh.Agarwal@honeywell.com
6/10/24

