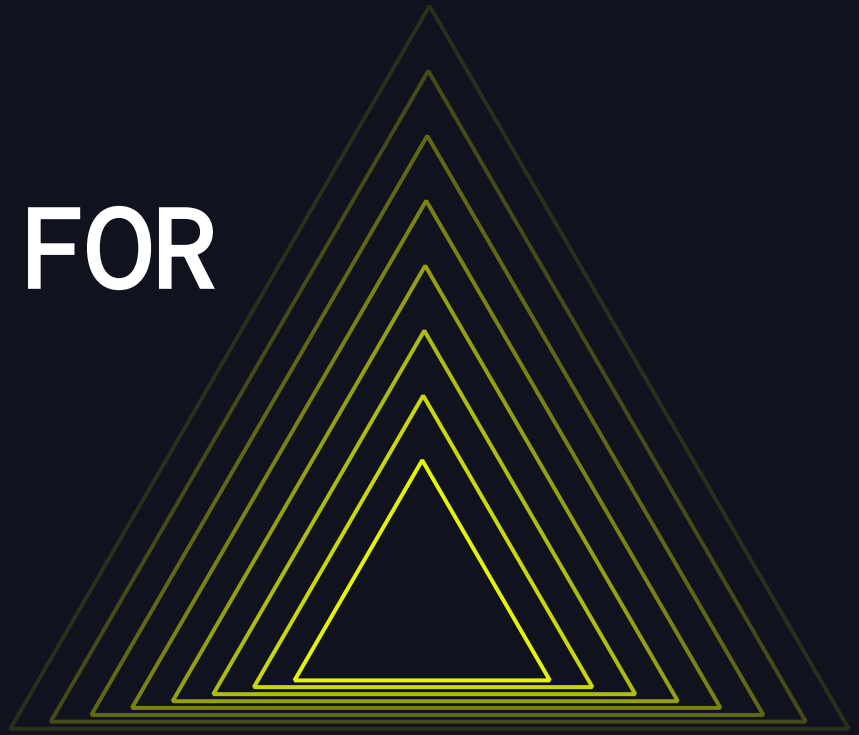
Product safe harbor statement

This information is provided to outline Databricks' general product direction and is for informational purposes only. Customers who purchase Databricks services should make their purchase decisions relying solely upon services, features, and functions that are currently available. Unreleased features or functionality described in forward-looking statements are subject to change at Databricks discretion and may not be delivered as planned or at all

DATA+AI SUMMIT BY databricks

DELTA SHARING: OPEN PROTOCOL FOR SECURE DATA SHARING (OSS)



Tianyi Huang, Databricks | Alexey Filanovskiy, Oracle
June 11, 2024

©2024 Databricks Inc. — All rights reserved

PRESENTERS

Tianyi Huang
Product Manager, Databricks

Alexey Filanovskiy
Senior Principal Product Manager, Oracle

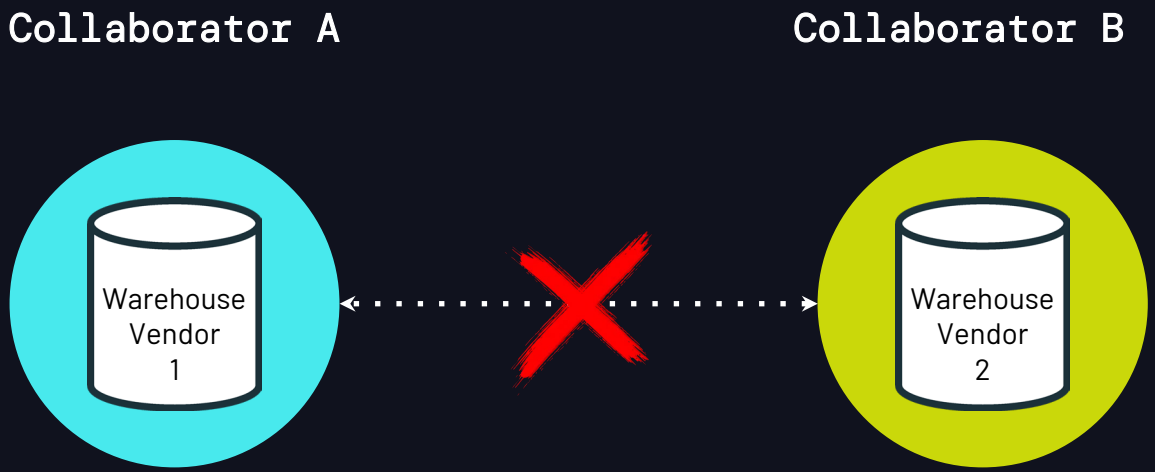
©2024 Databricks Inc. — All rights reserved

DATA NEEDS TO FLOW BEYOND PLATFORM BOUNDARIES

©2024 Databricks Inc. — All rights reserved

Proprietary platforms not built for sharing

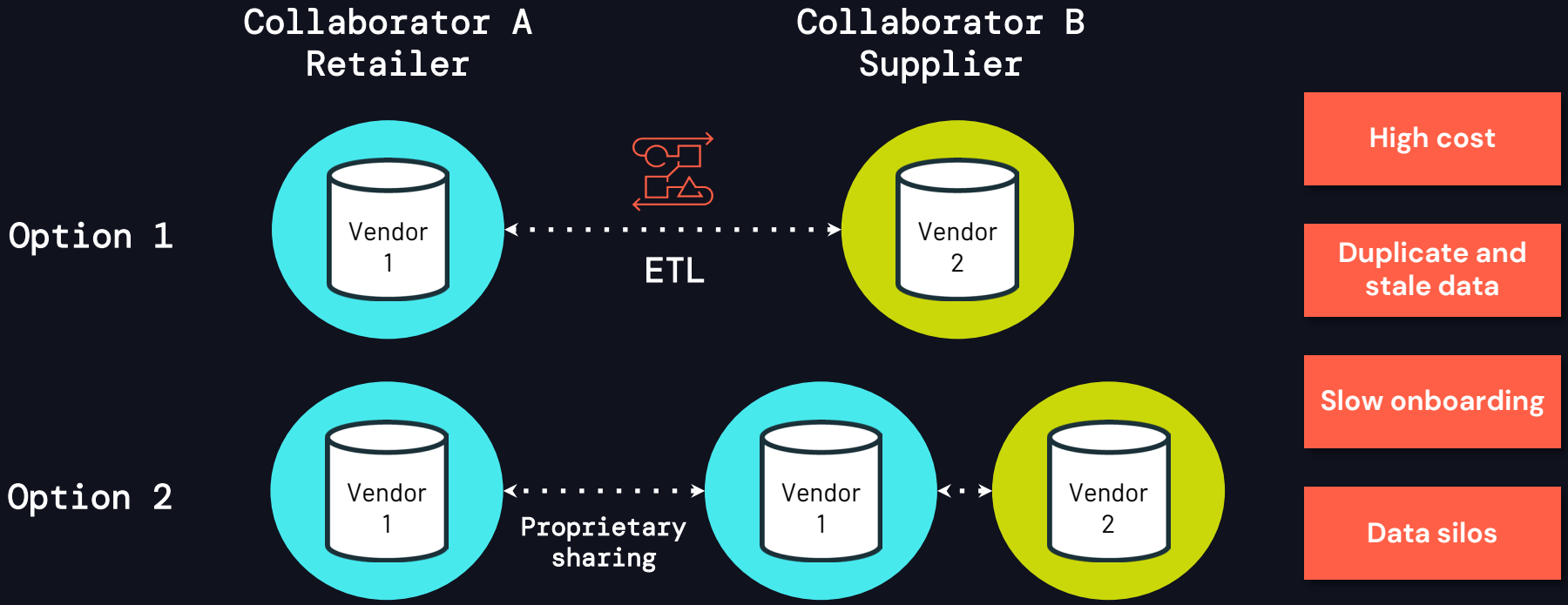
It's complex and expensive to share data with proprietary solutions



©2024 Databricks Inc. — All rights reserved

Proprietary platforms not built for sharing

It's complex and expensive to share data with proprietary solutions



- High cost
- Duplicate and stale data
- Slow onboarding
- Data silos

©2024 Databricks Inc. — All rights reserved

Delta Sharing: a platform-agnostic approach

Industry's first open protocol for secure data sharing

©2024 Databricks Inc. — All rights reserved

Guiding design principles

Industry's first open protocol for secure data sharing

1. Share live data directly, without the need for copying or moving
2. Support the widest range of clients
3. Enable strong security, auditing and governance
4. Scale to support massive datasets seamlessly

©2024 Databricks Inc. — All rights reserved

How does it work?

©2024 Databricks Inc. — All rights reserved

Querying shared data is as easy as …

Using the Python library as an example

```
import delta_sharing

# Create a delta sharing client.
client = delta_sharing.SharingClient(<credential_file_path>)

# List all shared tables.
client.list_all_tables()

# Load a table as a Pandas or Spark DataFrame.
delta_sharing.load_as_pandas(<table_url>)
delta_sharing.load_as_spark(<table_url>)
```


©2024 Databricks Inc. — All rights reserved

NEW: Delta Sharing Tableau connector

Accelerate data-driven decisions and innovations with open sharing

Expand the reach of your data to business analysts via Tableau

Easily tap into the rich Delta Sharing data ecosystem from Tableau

General Availability

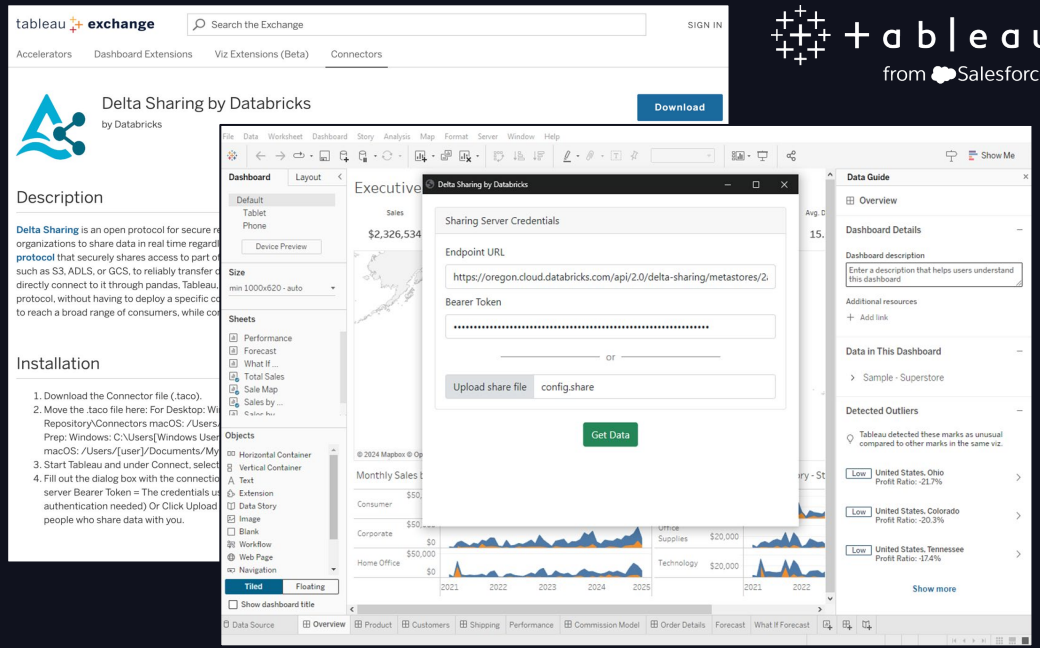


tableau from Salesforce

©2024 Databricks Inc. — All rights reserved

Delta Sharing on Databricks

Native integration with the Databricks Data Intelligent Platform

Centralized governance with Unity Catalog

Centrally manage, govern, audit and track usage of the shared data on one platform

Collaborate on all data + AI assets

Empower AI + BI workloads by sharing

- Tables and views
- Notebooks
- Unstructured data (Volumes)
- AI models

Broadest ecosystem of connectors and data

Share data with

- Open source clients
- Commercial platforms

Consume data from

- Marketplace
- SaaS partners

©2024 Databricks Inc. — All rights reserved

Databricks' open sharing ecosystem

S&P Global

Enhancing data distribution capabilities with Delta Sharing

"We are excited to continue to work with Databricks to enhance our data distribution capabilities through their Delta Sharing platform which provides an open ecosystem, making a number of S&P Global content sets more seamlessly accessible and available to our clients. This expands our collaboration with Databricks which began with S&P Global Capital IQ Workbench, leveraging their technology to create a collaborative analytics notebook environment for our users."

— David Coluccio, Head of Distribution Solutions, S&P Global Market Intelligence

©2024 Databricks Inc. — All rights reserved

Sharing data as a provider …

Reach all your clients wherever they are

1. Out-of-box Databricks Delta Sharing
2. Build on Databricks Delta Sharing
3. Host and manage your own Delta Sharing server from OSS (Oracle)

©2024 Databricks Inc. — All rights reserved

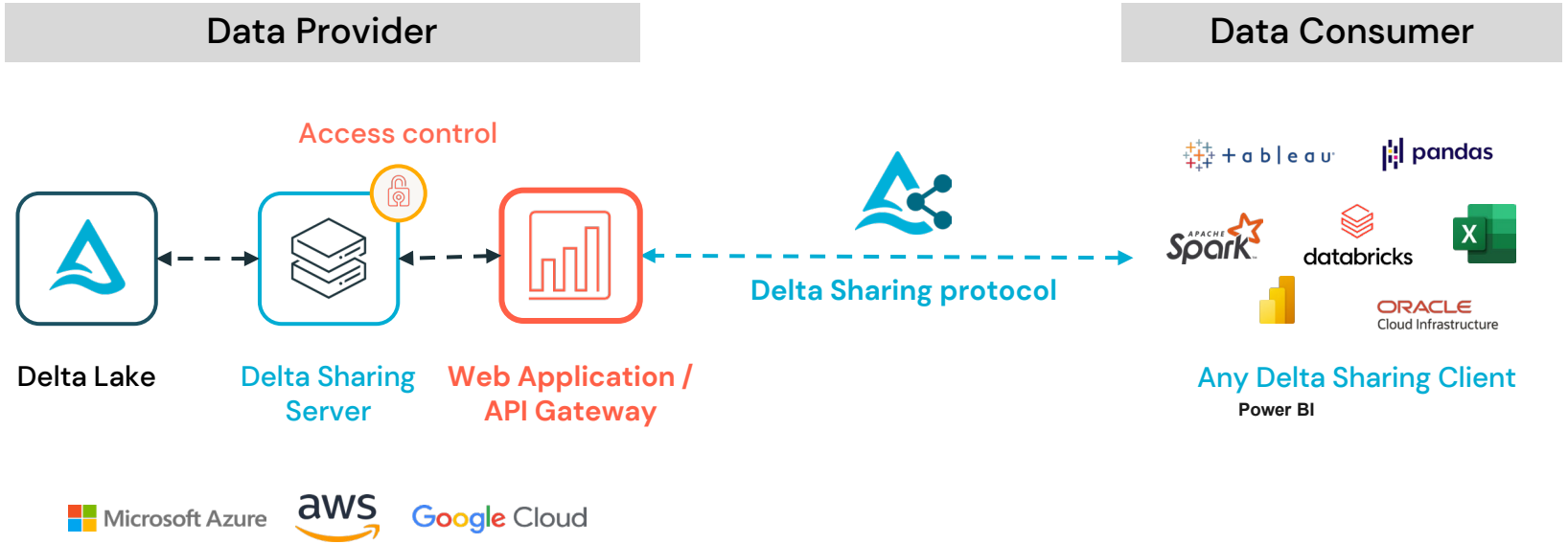
Sharing data as a provider …

Reach all your clients wherever they are

1. Out-of-box Databricks Delta Sharing
2. Build on Databricks Delta Sharing
3. Host and manage your own Delta Sharing server from OSS (Oracle)

©2024 Databricks Inc. — All rights reserved

Build on Databricks Delta Sharing



©2024 Databricks Inc. — All rights reserved

ATLASSIAN

Driving insights with a flexible, open ecosystem

"Atlassian Analytics recently launched Data Shares, leveraging Delta Sharing from Databricks, to boost flexibility and accelerate customers' time-to-insight. Whether users choose to work within Atlassian Analytics or continue using dashboards they're already familiar with, Delta Sharing's open ecosystem of connectors, including Tableau, PowerBI, and Spark, enables customers to easily power their environments with data directly from the Atlassian Data Lake."

— Suresh Raman, Head of Product, Data Platform, Atlassian

©2024 Databricks Inc. — All rights reserved

Build on Databricks Delta Sharing

Hear from Atlassian and Nasdaq about their success stories

ATLASSIAN

Session: Empowering Enterprise Grade Customers with Delta Sharing - an Atlassian Analytics Story

Tuesday, Jun 11
5:20 PM - 5:40 PM PDT
North, Expo Level, Theater 2

Nasdaq

Session: Delta Sharing Unlocks the Value of Your Data to Partners and Customers

Thursday, Jun 13
2:50 PM - 3:30 PM PDT
South, Level 3, Rm 302

©2024 Databricks Inc. — All rights reserved

Sharing data as a provider …

Reach all your clients wherever they are

1. Out-of-box Databricks Delta Sharing
2. Build on Databricks Delta Sharing
3. Host and manage your own Delta Sharing server from OSS (Oracle)

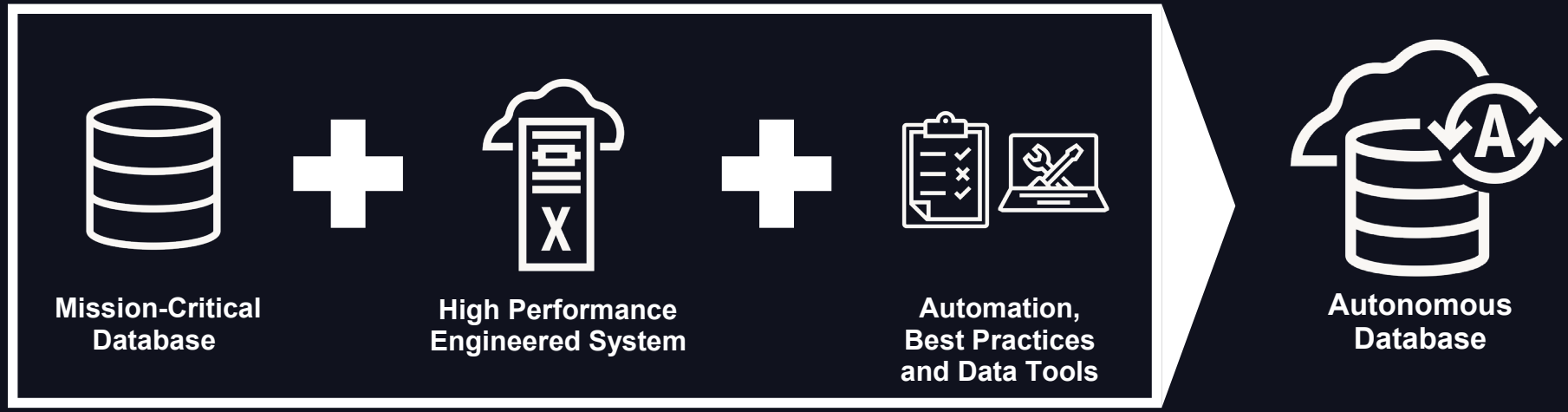
©2024 Databricks Inc. — All rights reserved

Oracle

©2024 Databricks Inc. — All rights reserved

Oracle Autonomous Database delivers the best data experience

With the lowest total cost of ownership



©2024 Databricks Inc. — All rights reserved

Lower costs and reduced risk through automation

Top autonomous capabilities that drive benefit

1 Auto-Provisioning
Automatically deploys mission-critical databases (RAC on Exadata infrastructure) which are fault-tolerant and highly available. Enables seamless scale-out, protection in case of a server failure and allows updates to be applied in a rolling fashion, while apps continue to run.

2 Auto-Configuration
Automatically configures the database to optimize for specific workloads. Everything from the memory configuration, the data formats, and access structures are optimized to improve performance.
Customers can simply load data and go.

3 Auto-Backups
Automatic daily backup of database or on-demand. Restore or recover a database to any point-in-time you specify in the last 60 days.

4 Auto-Scaling
Automatically scales compute resources when needed by workload. All scaling occurs online, while the application continuously runs. Enables true pay per use.

5 Auto-Indexing
Automatically monitors workload and detects missing indexes that could accelerate applications. It validates each index to ensure its benefit, before implementing it and uses machine learning to learn from its own mistakes.

6 Automated Security
Automatic encryption for the entire database, backups and all network connections. No access to OS or admin privileges prevents phishing attacks. Protects the system from both cloud operations and any malicious internal users.

7 Automatic Failover
Automatic failover with zero-data loss to standby. It's completely transparent to end-user applications. Provides 99.995% SLA.

8 Automated Detection and Resolution
Using pattern recognition, hardware failures are automatically predicted without long timeouts. IOs are immediately redirected around unhealthy devices to avoid database hangs. Continuous monitoring for each database automatically generates service requests for any deviation.

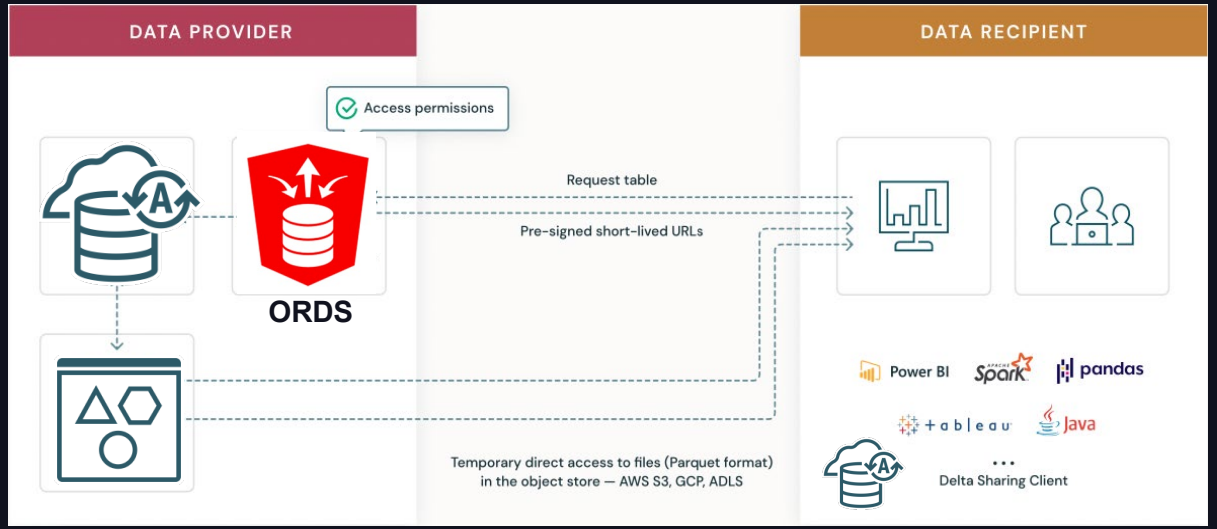
9 Automated Hardware and Software Upgrades
Hardware upgrades occur automatically and transparently. Major database upgrades are automated and scheduled by the customer.

10 Auto-Patching
Automatically patch or upgrade with zero downtime. Applications continue to run as patching occurs in a round-robin fashion across RAC nodes or servers.

©2024 Databricks Inc. — All rights reserved

Share Oracle data to any Delta client

Data Sharing through open standard protocol Delta Sharing



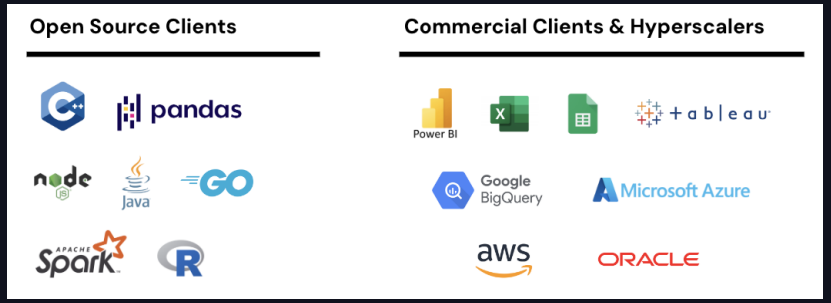
- Delta Sharing is part of Autonomous Database
- Available upon provisioning
- Autonomous Database can consume and provide Delta Sharing
- Nothing has to be managed or installed

©2024 Databricks Inc. — All rights reserved

Summary: Data Sharing in Autonomous Data Warehouse

Share with anyone in any cloud

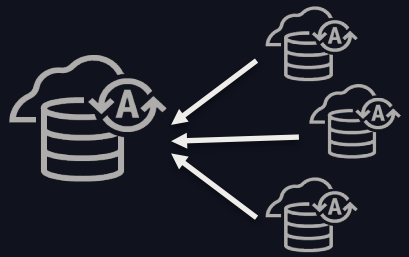
- Implemented via **open-source Delta Sharing** API
- Share data with any supported recipient: other databases, data analytics tools, open-source platforms



Optimized sharing within Oracle Cloud

Implemented via **Cloud Links** and OCI infrastructure and metadata

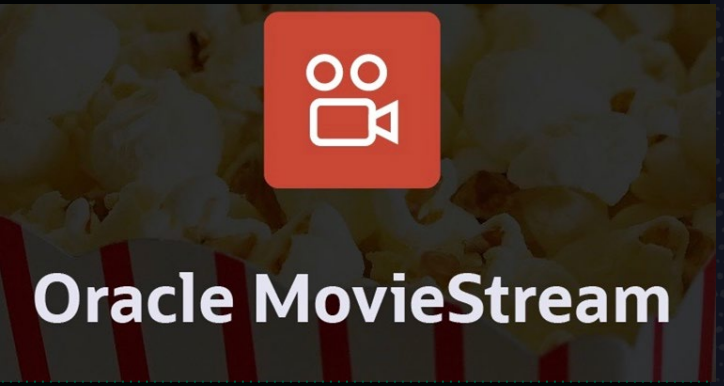
Automatic recipient notification and share discovery



©2024 Databricks Inc. — All rights reserved

Demo

©2024 Databricks Inc. — All rights reserved



MovieStream is a fictitious online movie streaming company

The customer base has expanded rapidly, generating massive volumes of structured, semi-structured and unstructured data

IT infrastructure consists of multiple products, run on a different platforms

Check this Live Lab

“ I am Mike. I am a Data Analyst.
I know something about IT, but I do not have strong background.”

“ MovieStream has a few data platforms running in a different clouds. I’m working with ADB, but I hear there other guys use Databricks.

MovieStreams Architecture

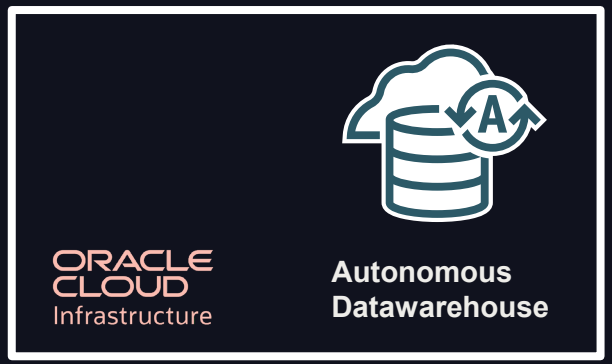
- Running Big Data Platform using Databricks (on Azure)
- Autonomous Datawarehouse has sales and customers data
- Analytics team heavily use PowerBI
- Data Scientists of course up to Python

Just another day…

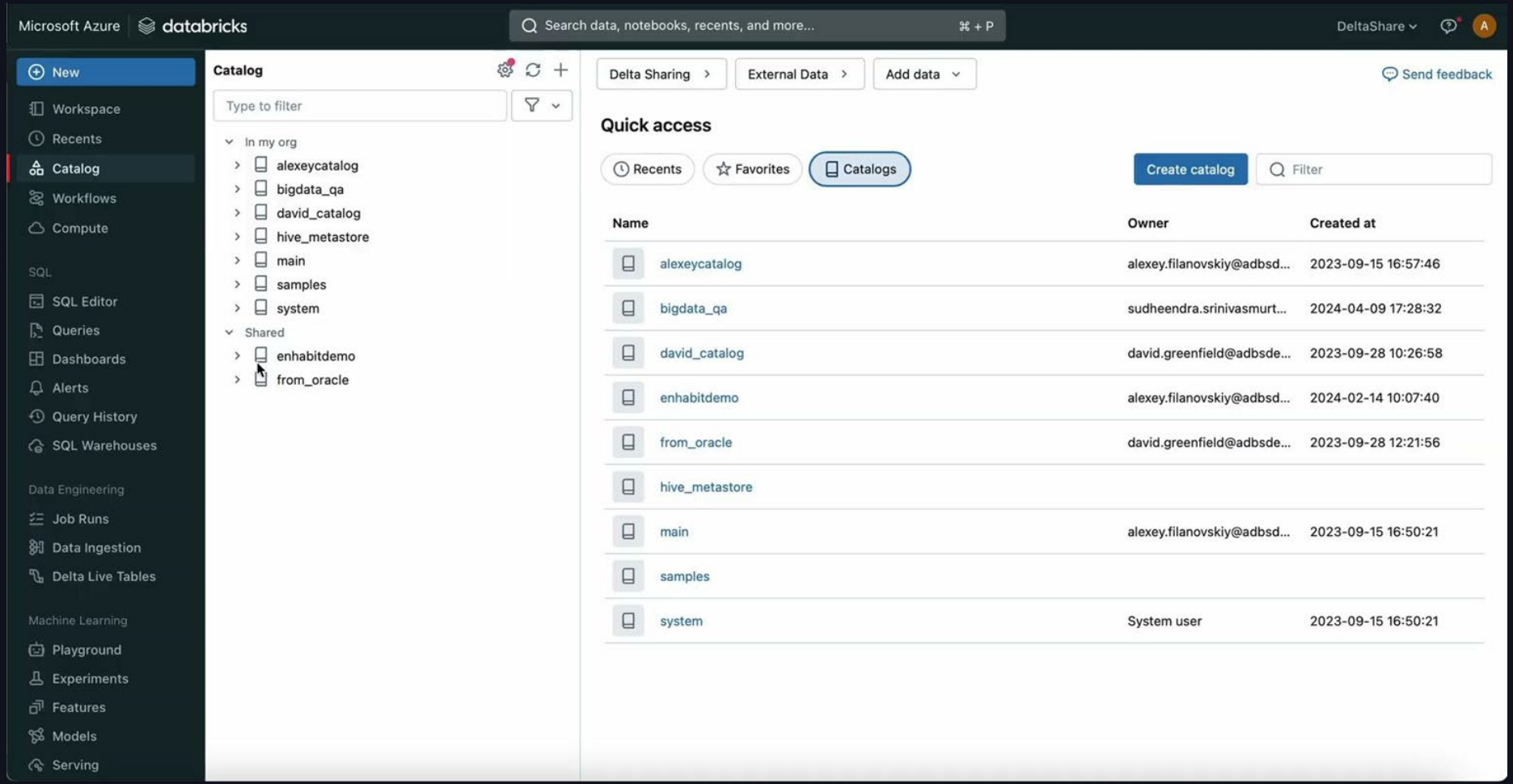
Instead of morning coffee

- This morning, I saw an annoying email. Our business is losing customers.
- The Analyst group and Data Science group wants to help, but they need data.
- Before, another group within my organization made a system to identify customers likely to leave. I need to see this data.

Create Delta Share in Databricks

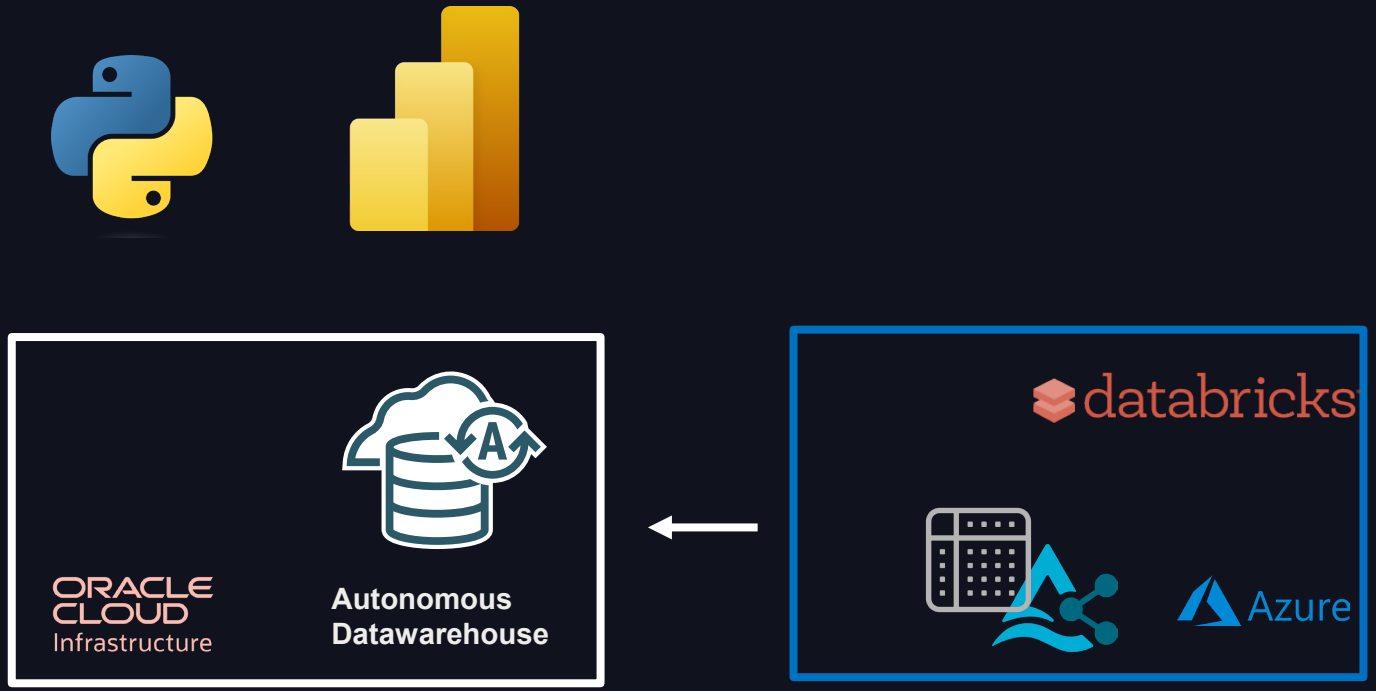


©2024 Databricks Inc. — All rights reserved



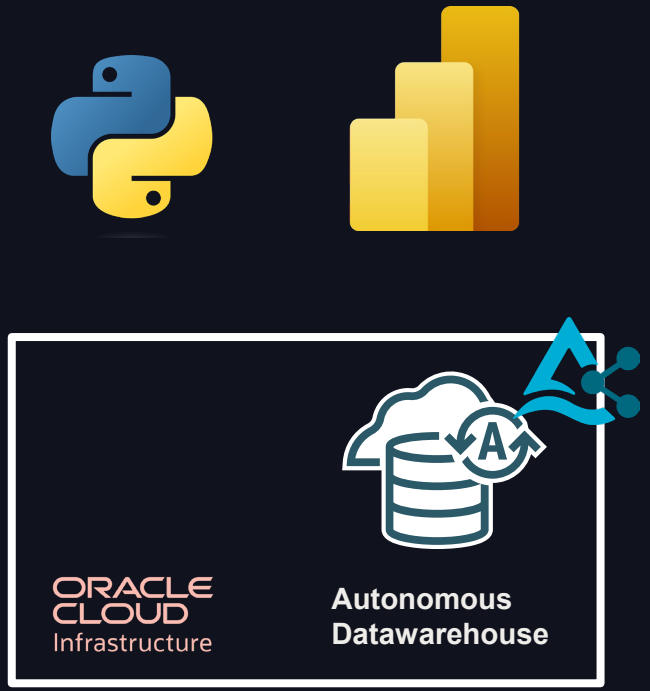
©2024 Databricks Inc. — All rights reserved

Consume share by Autonomous database

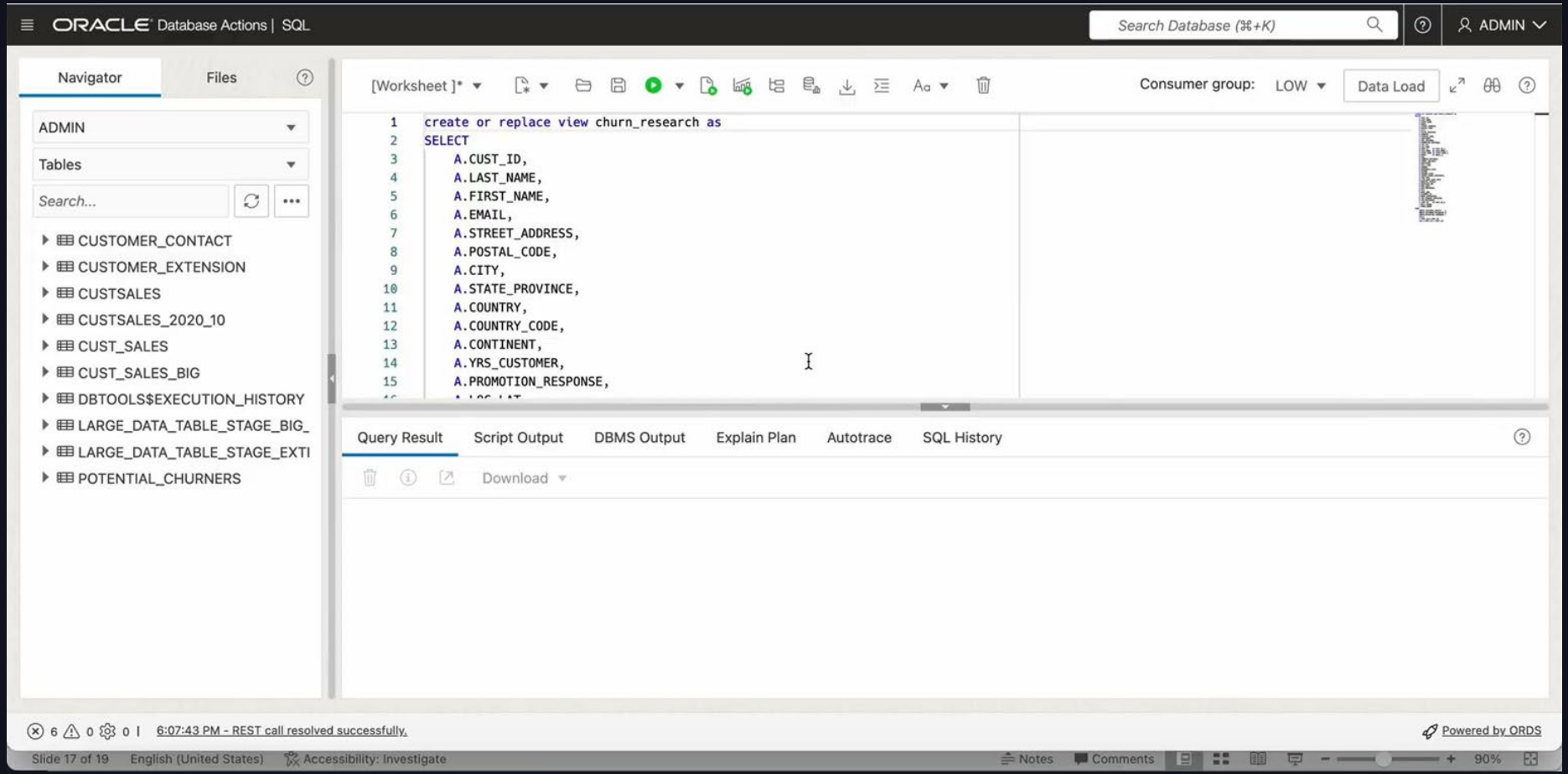


©2024 Databricks Inc. — All rights reserved

Create Delta Share on Autonomous Database

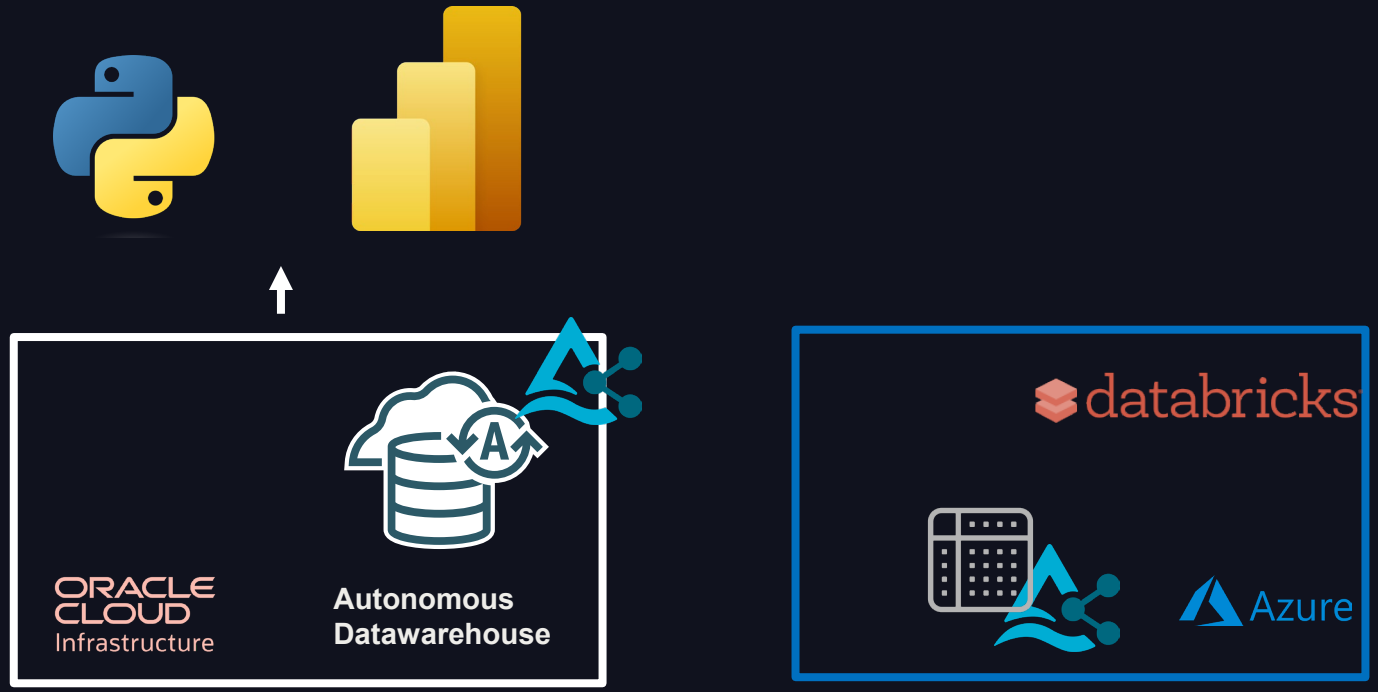


©2024 Databricks Inc. — All rights reserved

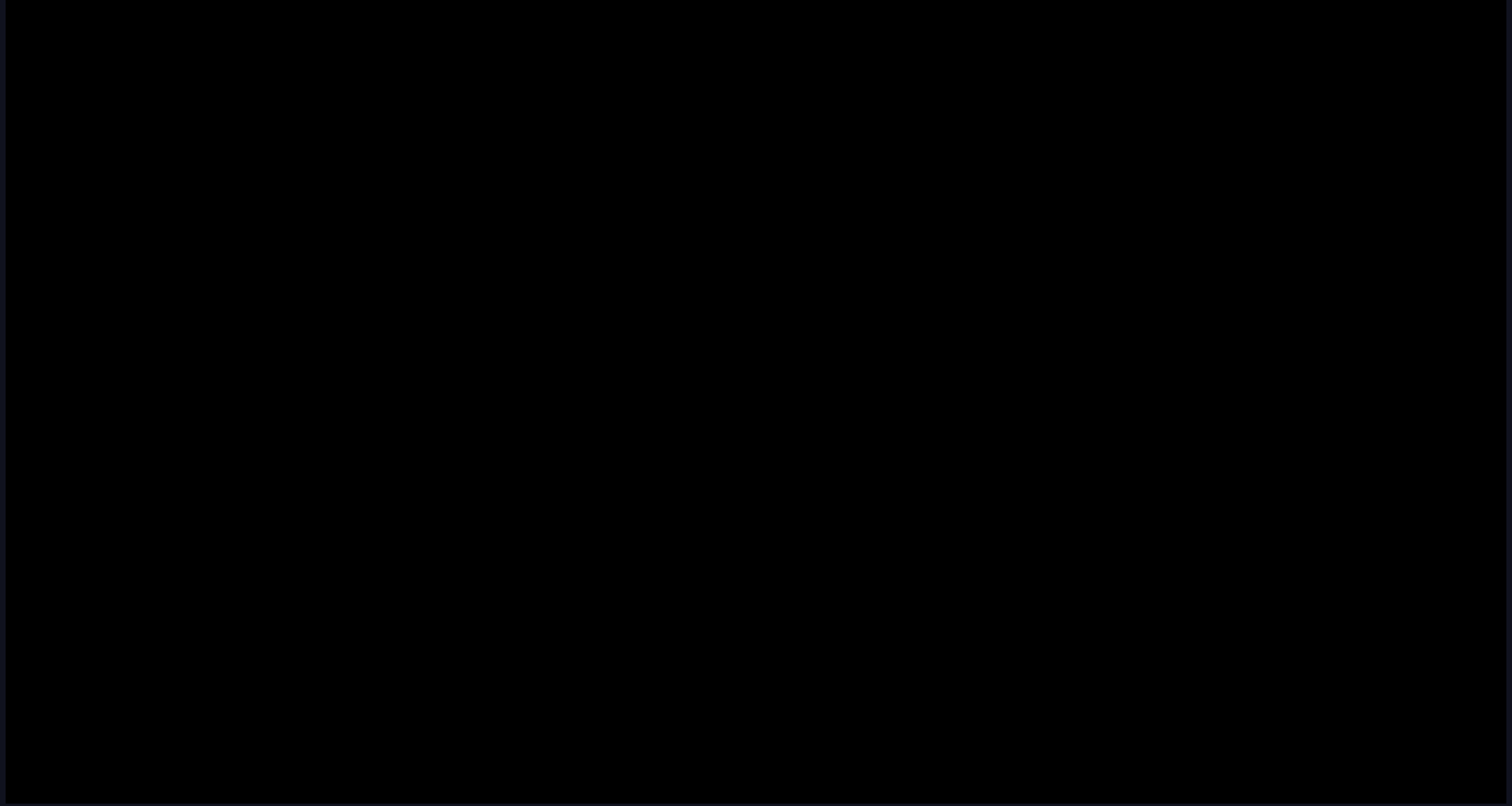


©2024 Databricks Inc. — All rights reserved

Consume Autonomous Database data using Delta Sharing protocol

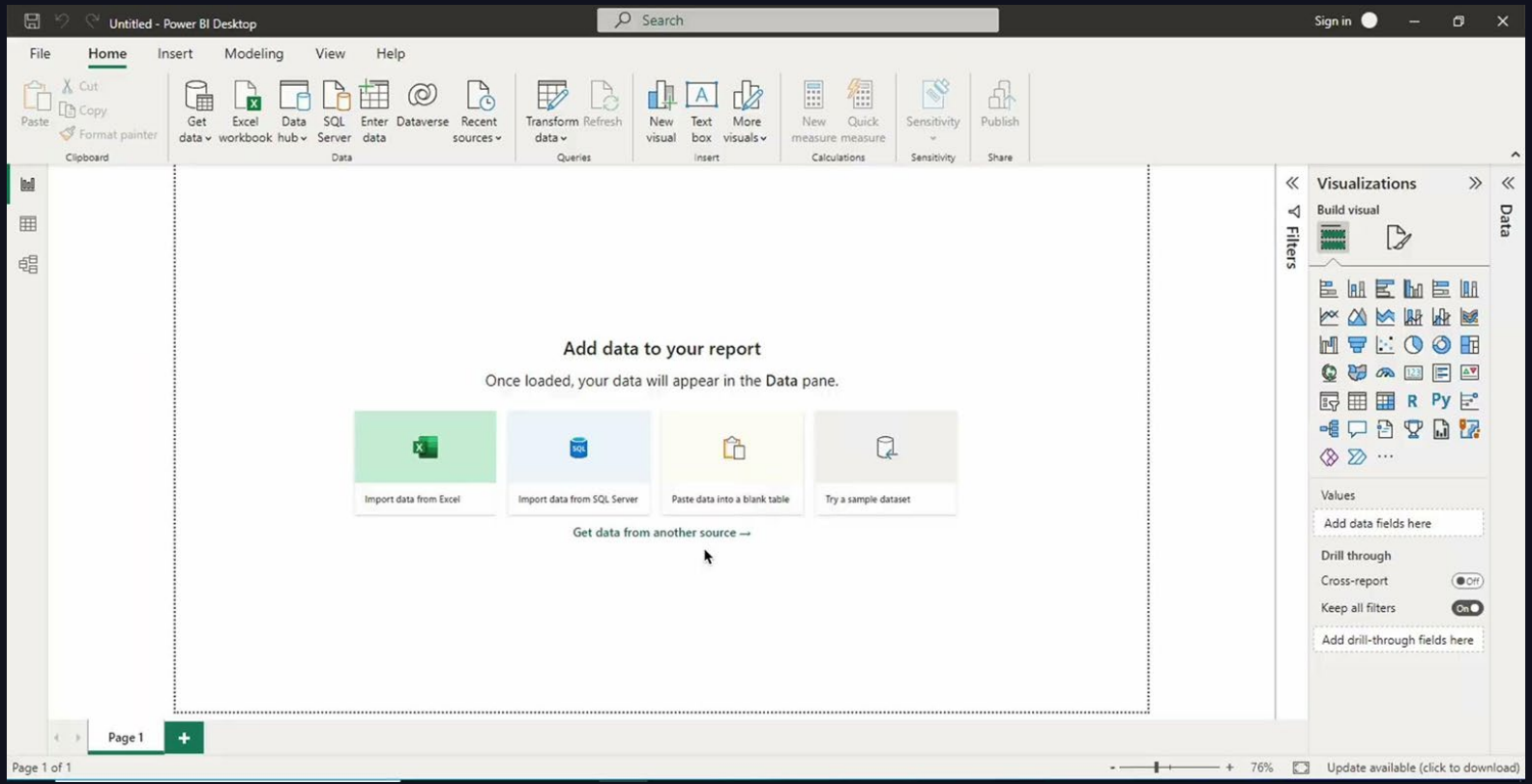


©2024 Databricks Inc. — All rights reserved



©2024 Databricks Inc. — All rights reserved

Read Delta Share with PowerBI



©2024 Databricks Inc. — All rights reserved

Autonomous Database and Delta Sharing

Takeaways

- Oracle Autonomous Database can consume delta share from Databricks
- Oracle Autonomous Database can provide delta share to other clients

Q&A

Later sessions from Delta Sharing ecosystem partners

Data partners

- **AccuWeather**: Turn Risky Weather Obstacles into Business Opportunities
- **Deutsche Börse**: Deutsche Börse's New Gateway to Data Across the Group
- **S&P Global** et al. (panel discussion): Accelerating Financial Innovation: Unlock Insights with Databricks Marketplace
- **Tetrascience**: Unlocking Scientific Data and Advancing the Scientific AI Revolution

SaaS & Technology partners

- **Amperity** et al. (panel discussion): Driving Customer Growth & Monetization Through Industry Collaboration
- **Cloudflare**: Cost Efficient Data Collaboration with Cloudflare R2 and Delta Sharing
- **Tableau**: Data Viz Excellence: Tableau's Edge with Delta Sharing & Roadmap

Built-on partners

- **Atlassian**: Empowering Enterprise Grade Customers with Delta Sharing - an Atlassian Analytics Story
- **Nasdaq**: Delta Sharing Unlocks the Value of Your Data to Partners and Customers

©2024 Databricks Inc. — All rights reserved