IBM Data Observability by Databand

Detect data incidents earlier and resolve them faster to deliver trusted data.

"Before Databand, 60% of our pipelines had at least one data incident. Now less than 1% of pipelines have incidents. This resulted in a 3X increase in our customers since we can now manage our ML deep learning models at scale."

Tzoof Hemed
AI-Engineering Team Leader
Trax Retail

As data volumes continue to grow at an unprecedented pace, data teams struggle to manage their data health and quality. Every year poor data quality costs organizations an average of $12.9 million.

Data engineering and platform teams constantly complain about firefighting data reliability incidents, making it difficult to focus on strategic initiatives. Ultimately, SLAs are unmet, and data consumers experience downstream results like inaccurate analytics, underperforming ML products, and unreliable client data. Data teams are under enormous pressure to deliver reliable data, and they need help. Enter data observability.
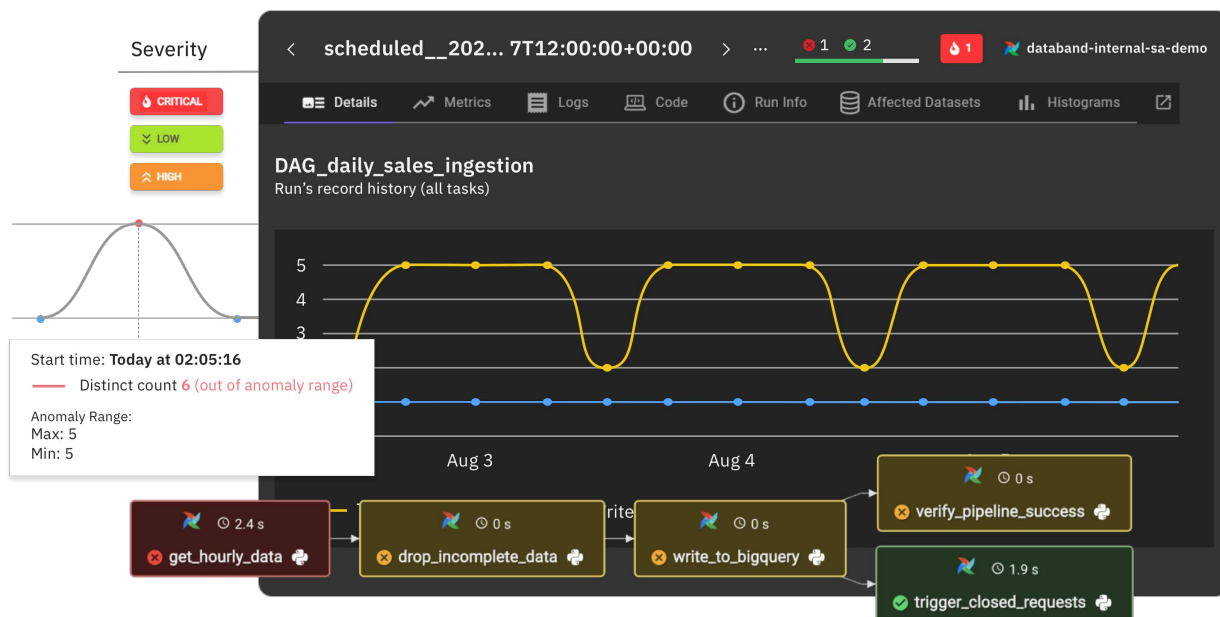
IBM Data Observability by Databand helps modern data engineering and platform teams deliver more reliable and trustworthy data by using a proactive approach to data observability. Databand helps remediate data issues such as broken pipelines, missing data, and schema changes by detecting and resolving them before they create costly business impacts.

Data observability elevates traditional data operations by using historical trends to compute statistics about data workloads and data pipelines directly at the source, to determine if they are working, and to pinpoint where problems might exist.

Key benefits

Detect earlier
Pinpoint unknown data incidents and reduce mean time to detection (MTTD) from days to minutes.

Resolve faster
Improve mean time to resolution (MTTR) with incident alerts and routing from weeks to hours.

Deliver trustworthy data
Enhance reliability and data delivery SLAs and provide visibility into pipeline quality issues.

Key use cases

– **Data incident management**
  Don't just observe data incidents. Resolve them fast. Now you can alert, respond, and resolve all your data incidents in one location.
– **Data pipeline monitoring**
  Manage the health of 100s to 1000s of data pipelines. Detect missing operations, failed jobs, and run durations so you can handle pipeline growth.
– **Data quality monitoring**
  Ensure better data quality by monitoring data SLAs, unexpected column changes, and null records before they get to your consumers.
– **Impact analysis and lineage**
  Break siloes and get the whole data story with end-to-end data lineage. Understand the impact of data incidents on upstream and downstream data flows.
– **Data anomaly detection**
  The worst data incidents are unknown. Anomaly detection removes bad data surprises by automatically detecting deviant behavior in your data pipelines.
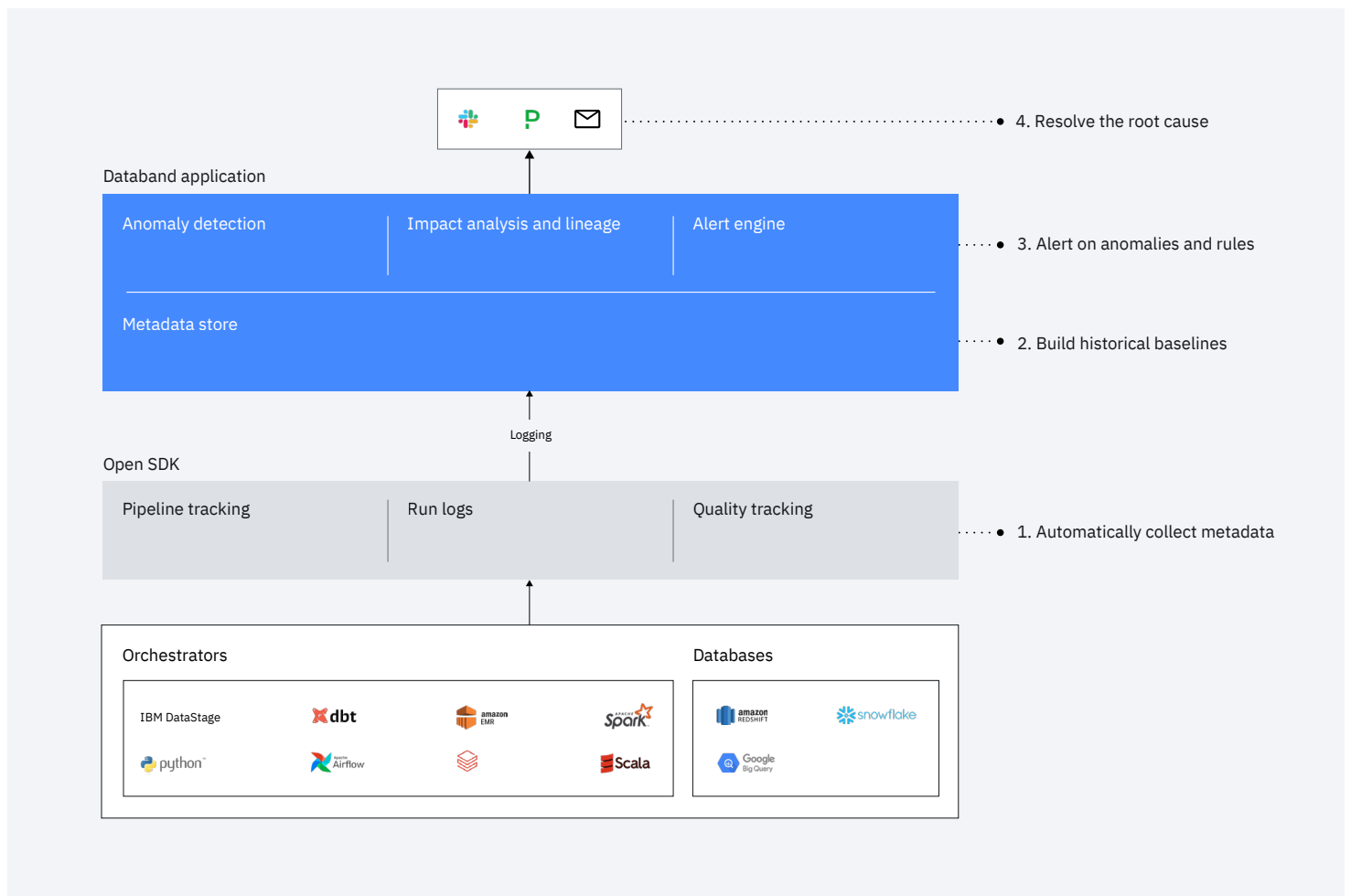
Infrastructure and deployment options

■

Databand provided maximum
flexibility for deployment options:

– Cloud SaaS subscription
– Self-hosted subscription

How Databand works

1. **Collect metadata**
   Automatically collect metadata from your modern data stack like IBM
   DataStage®, Airflow, Spark, Databricks, Redshift, dbt, and Snowflake.
2. **Profile behavior**
   Build historical baselines based on common data pipeline behavior and
   get visibility into every data flow, from source to destination.
3. **Detect and alert data incidents**
   Detect high severity data reliability errors that impact your most critical
   pipelines and alert impacted teams.
4. **Resolve through triage**
   Create smart workflows that remediate data quality incidents and keep
   SLAs on track.

Why IBM?

Leverage IBM's data fabric architecture to deliver reliable data to your business with Databand. Databand deployed through the architecture ensures your organization can:

– Automatically observe dynamic data pipelines
– Proactively address data quality and reliability
– Continuously monitor AI/ML reliability across data and model

To learn more about IBM Data Observability by Databand, please contact your IBM representative or IBM Business Partner, or visit databand.ai/request-demo to book a demo.