

What is a Multi-Hop Architecture?



Learning Objectives

- Structure a reliable data pipeline architecture
- Dissect its components
- Highlight its advantages



Multi-Hop Architecture

- Data pipeline architecture for batch processing and/or real-time processing
- Separates data storage and data processing in multiple layers
- Also known as Medallion Architecture



Data Sources



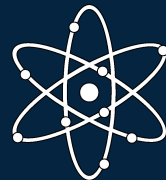
- Databases
- Data Lakes
- Real Time Data

Data Pipeline



- Multi-Hop Architecture
- Incremental Data Quality
- Incremental Business Value

Data Analytics



- BI Use Cases
- ML use Cases
- AI Use Cases

Bronze Layer

- Generally includes a copy of all raw data sources
- Provides the ability to query unprocessed data
- Provides efficient storage



Silver Layer

- Generally includes a cleaner, more polished version of the data
- It is the result of cleaning, filtering, and aggregating data
- Enforces data quality checks and quarantines corrupted data



Gold Layer

- Generally includes refined data with the highest business value
- Supports final BI, ML, and AI use cases
- Optimized for reading queries performances



Advantages of Multi-Hop Architecture

- Built on simple concepts
- Works well with batch processing, real-time processing, or both
- Provides a precise audit trail from raw data, to business critical data

