



Why the Lakehouse?



Craig Porteous
Principal Consultant



<https://craigporteous.com>



[@cporteous](https://twitter.com/cporteous)



<https://github.com/cporteous>



[@ADVANCINGANALYTICS](https://www.instagram.com/advancinganalytics)



[@ADVANALYTICSUK](https://twitter.com/advanalyticsuk)



[/ADVANCING ANALYTICS](https://www.youtube.com/advancinganalytics)



Why the Lakehouse?



Craig Porteous
Principal Consultant



<https://craigporteous.com>



[@cporteous](https://twitter.com/cporteous)



<https://github.com/cporteous>



[@ADVANCINGANALYTICS](https://www.instagram.com/advancinganalytics)



[@ADVANALYTICSUK](https://twitter.com/advanalyticsuk)



[/ADVANCING ANALYTICS](https://www.youtube.com/advancinganalytics)



IS THIS A TIMESHARE?



databricks



Microsoft

ORACLE®



snowflake

CLOUDERA



Google Cloud

What problem
is it trying to
solve?



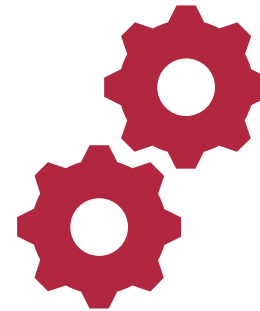
PATH TO THE LAKEHOUSE



History



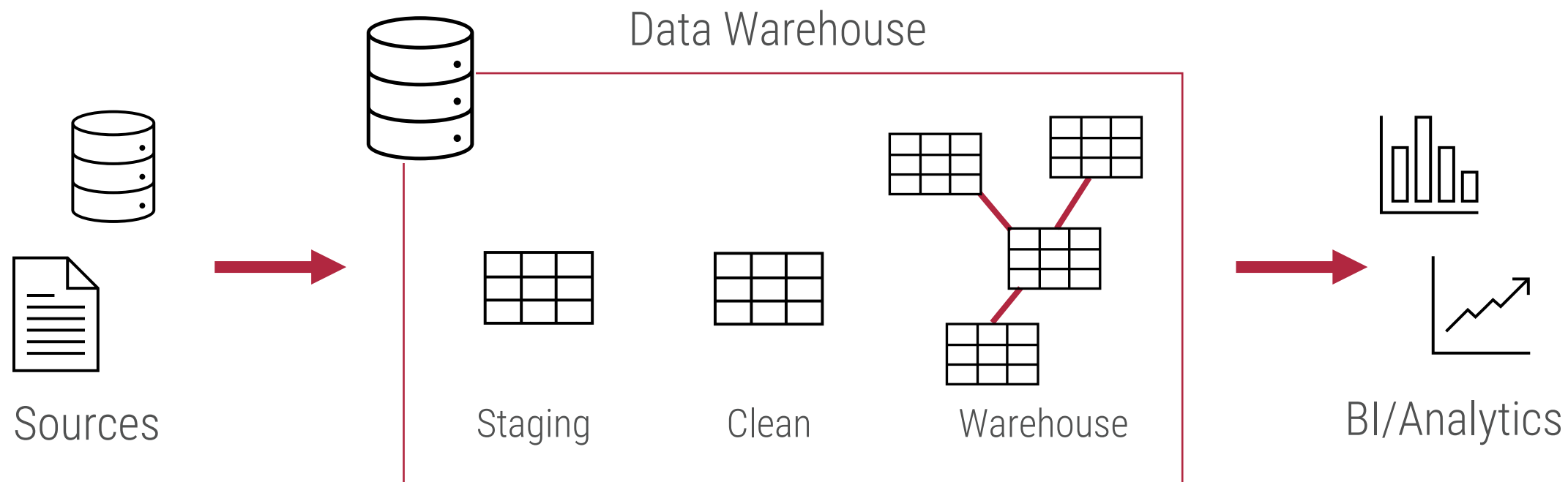
Lakehouse



Technologies

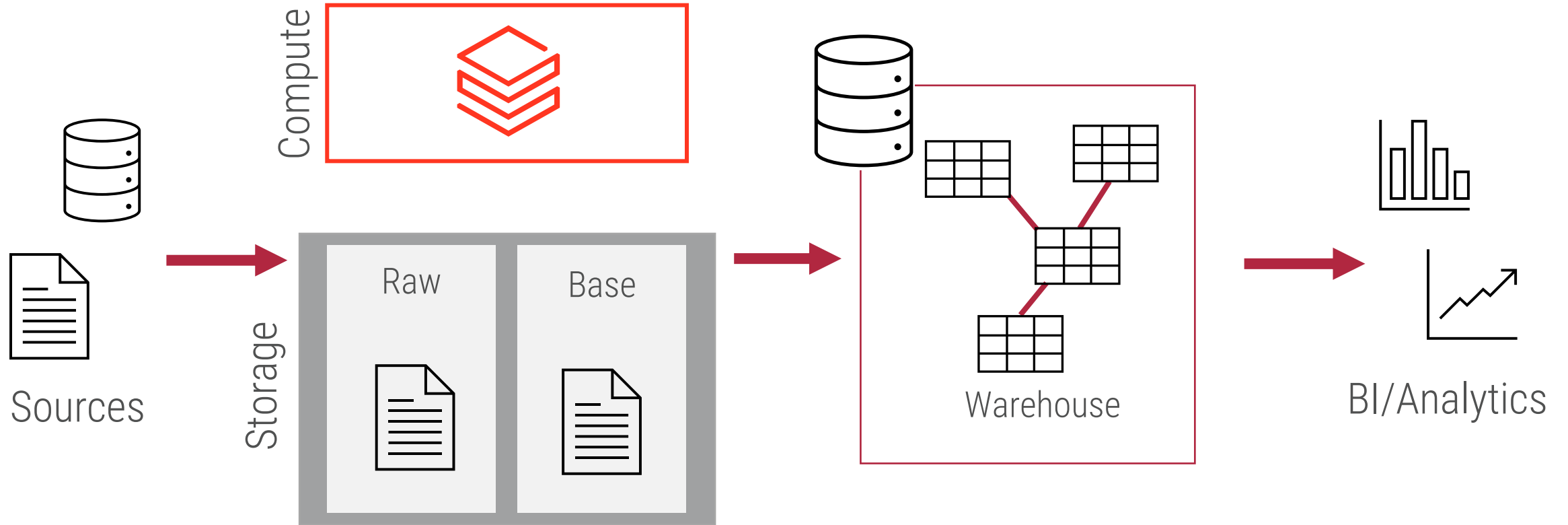


THE DATA WAREHOUSE



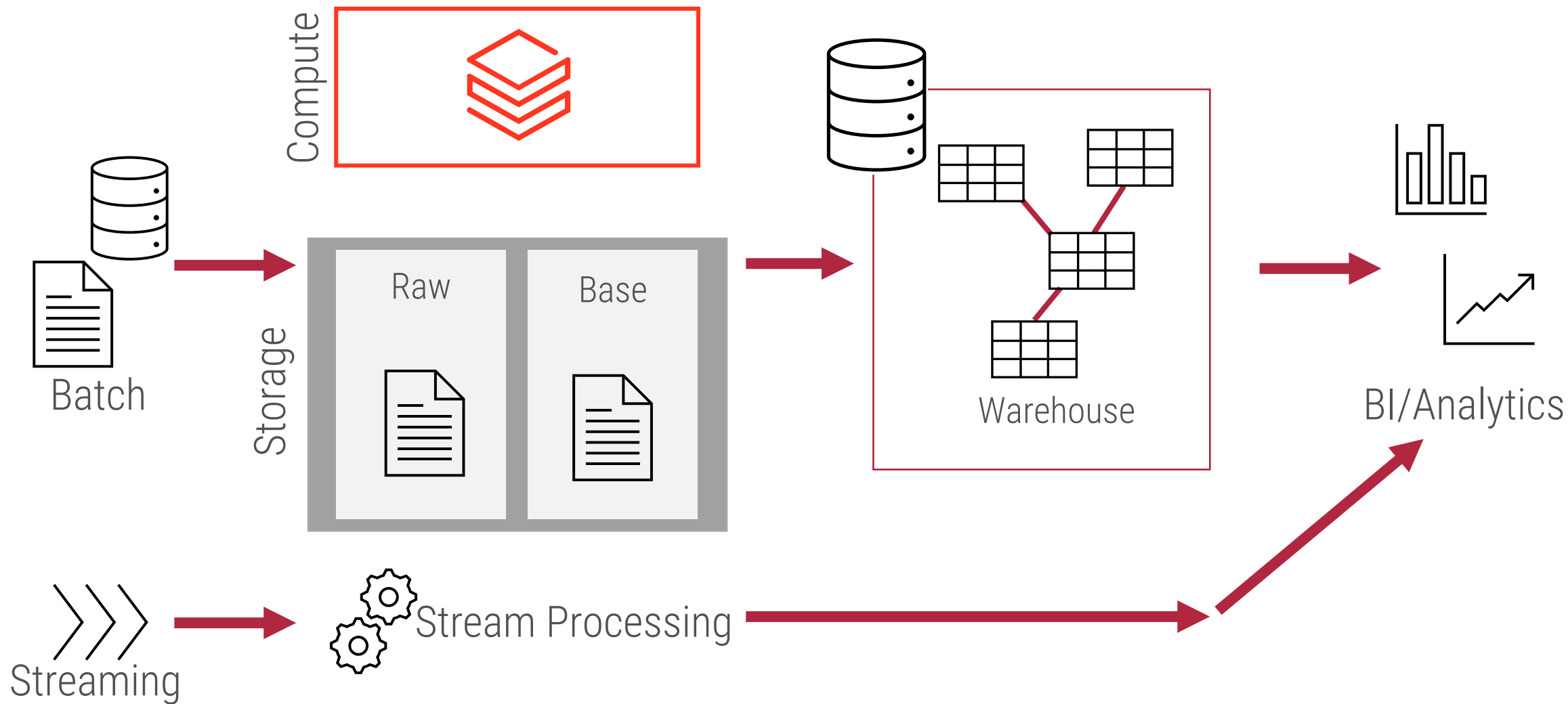


THE MODERN DATA WAREHOUSE



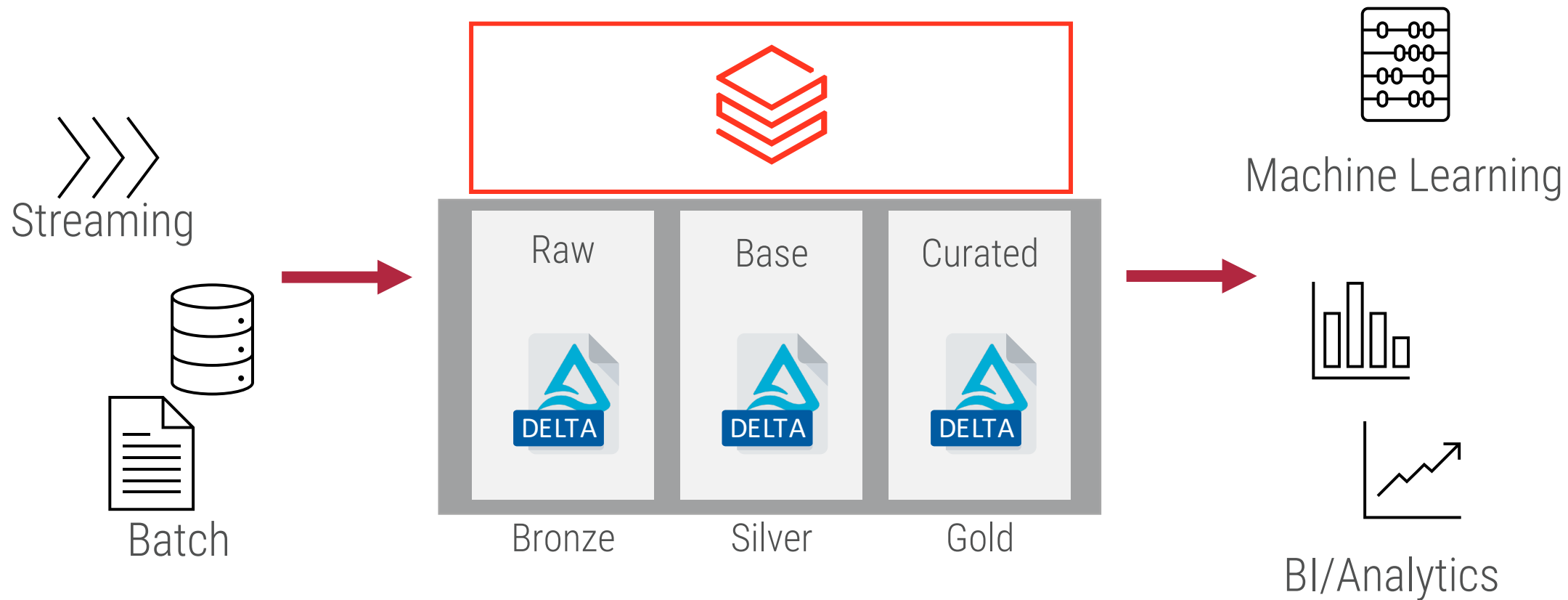


LAMBDA ARCHITECTURE



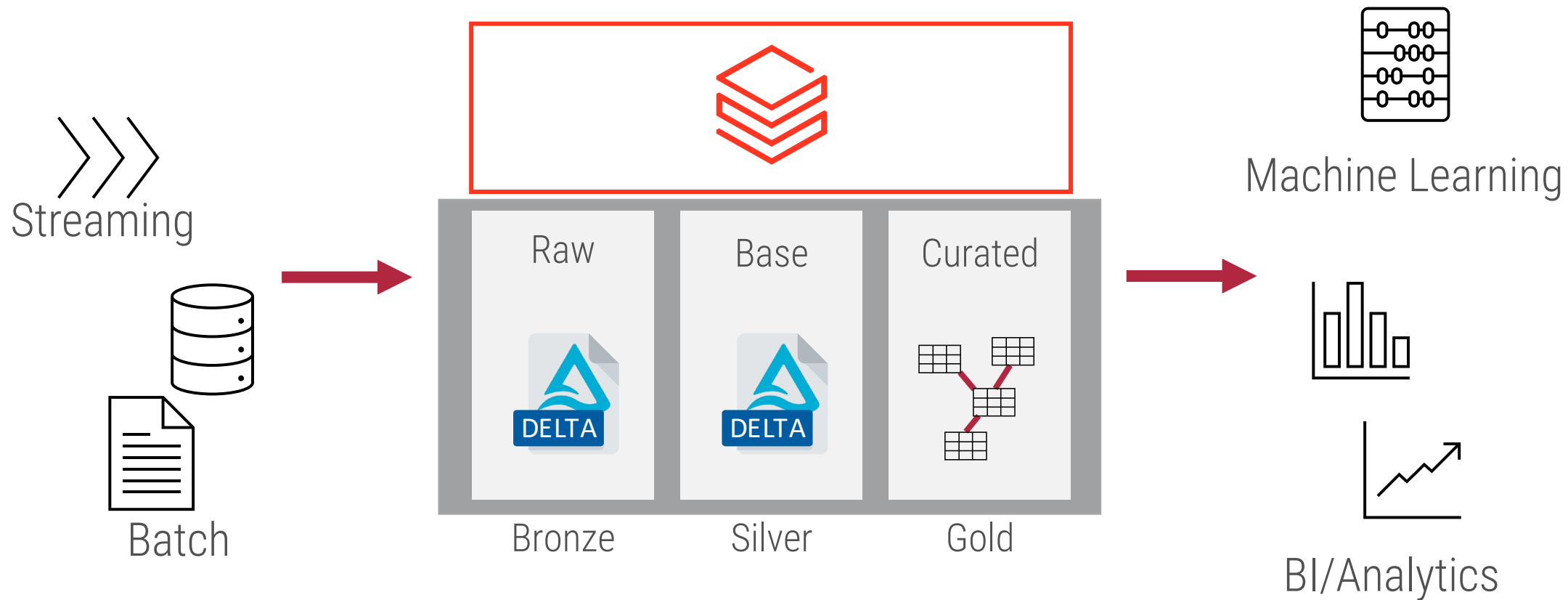


THE DATA LAKEHOUSE





THE DATA LAKEHOUSE





DELTA FEATURES

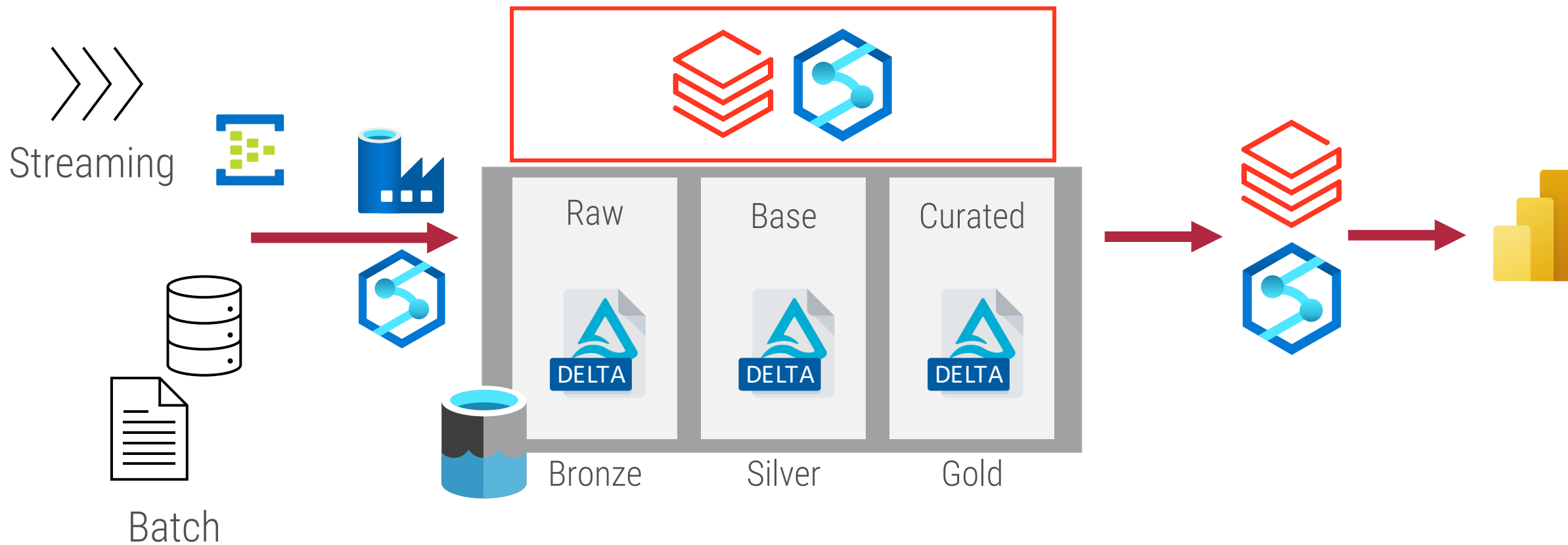


DELTA LAKE

- ◆ OPEN SOURCE
- ◆ BASED ON PARQUET
- ◆ ACID TRANSACTIONS
- ◆ TIME TRAVEL
- ◆ SCHEMA EVOLUTION
- ◆ BATCH & STREAMING SUPPORT



AN AZURE DATA LAKEHOUSE

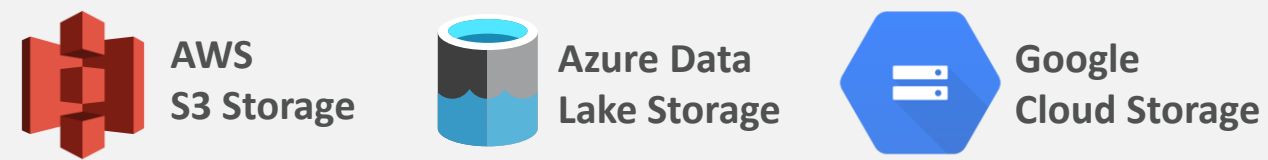


COMMON LAKEHOUSE TECHNOLOGIES

COMPUTE



STORAGE



SERVE





SKILLS GAP

Spark

SQL

PYTHON

SCALA



WHAT IS A DATA LAKEHOUSE

**DATA
WAREHOUSE**

◆ ACID TRANSACTIONS

◆ GOVERNANCE



**DATA LAKE
STORAGE**

◆ CHEAP

◆ FLEXIBLE

DATA LAKEHOUSE

◆ ANALYTICS

◆ BUSINESS INTELLIGENCE

◆ MACHINE LEARNING

◆ BIG DATA



THANK YOU

Please leave feedback



<https://craigporteous.com>



[@cporteous](https://twitter.com/cporteous)



<https://github.com/cporteous>

<https://sqlb.it/?10207>



@ADVANCINGANALYTICS



@ADVANALYTICSUK



/ADVANCING ANALYTICS