



---

# DATA TRANSFORMATION AND ETL

---

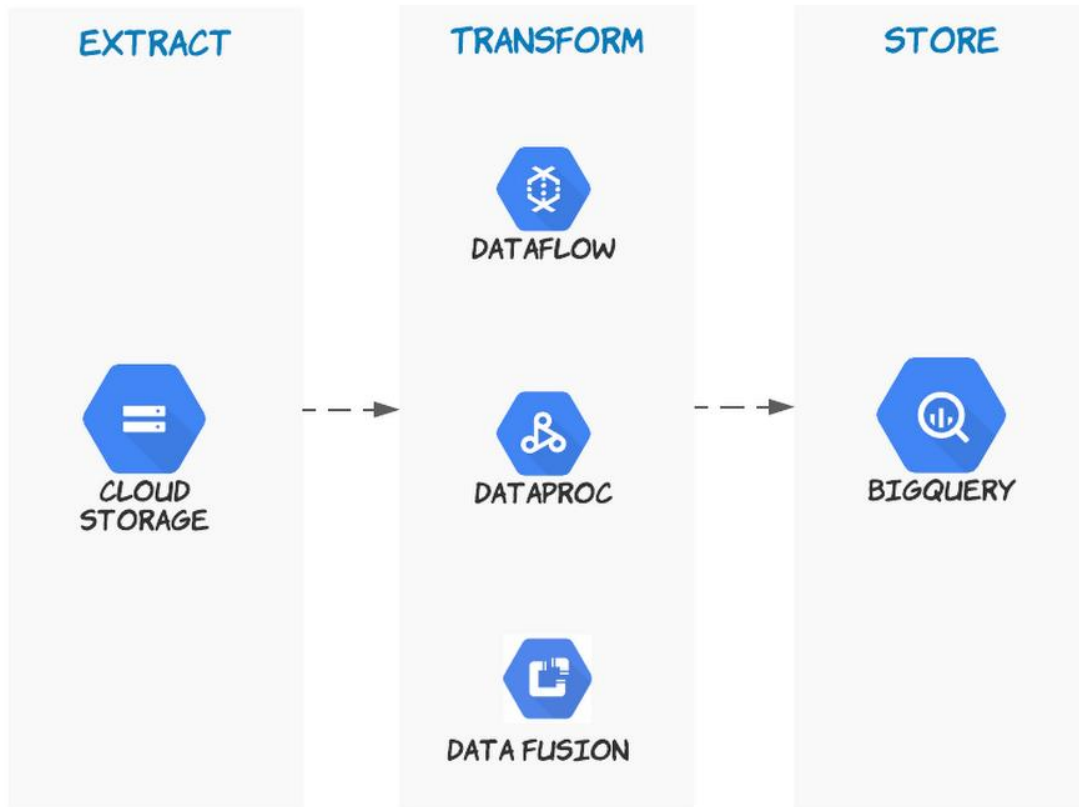


MAY 31, 2023  
BYTEWISE LIMITED

## Contents

<b>Data Transformation and ETL .....</b>	<b>2</b>
<b>Explanation GCP ETL .....</b>	<b>2</b>
<b>1. Data flow .....</b>	<b>2</b>
<b>2. Data Proc .....</b>	<b>3</b>
<b>3. Cloud Data Fusion .....</b>	<b>3</b>

## Data Transformation and ETL



*Fig 1 GCP ETL*

## Explanation GCP ETL

### 1. Data flow

- Data in real time is generated and capturing the data **analyzing and processing** data is imp to make data easily used by systems in 1 single format then data flows helps to achieve it.
- Server less, cost effective, fast data processing service for streaming and batch data.
- You read data from source transform it and write it to a back sink.
- You can process and enrich any data using data flow.

## 2. Data Proc

- **Accelerate your open-source data and analytics processing.**
- Is a managed service that supports big data processing.
- With data proc you can spin clusters within 90 sec
- Manages cluster creation, monitoring and orchestration for you.
- **Simplify your data analytics and processing.**

## 3. Cloud Data Fusion

- **Helps to build data warehouse, data lakes and data marts.**
- Fully managed cloud native enterprise integration service for ETL developers, data engineers etc.
- Design, Deploy and manage data pipelines.
- Visual point-and-click interface enabling code-free deployment of ETL/ELT data pipelines.
- Cloud Data Fusion's integration makes development and iteration fast and easy.