APRIL 2021

DATA ENGINEERING COE

GOOGLE CLOUD PLATFORM CAPABILITY

ACCELERATE WHAT MATTERS. NOW.

brillio

Data Engineering COE | Google Cloud Capability

brillio
ACCELERATE WHAT MATTERS. NOW.

Capabilities

- DATA WAREHOUSE MODERNIZATION
- DATA MIGRATION
- DATA LAKE IMPLEMENTATION
- MULTI-CLOUD IMPLEMENTATION

Objectives

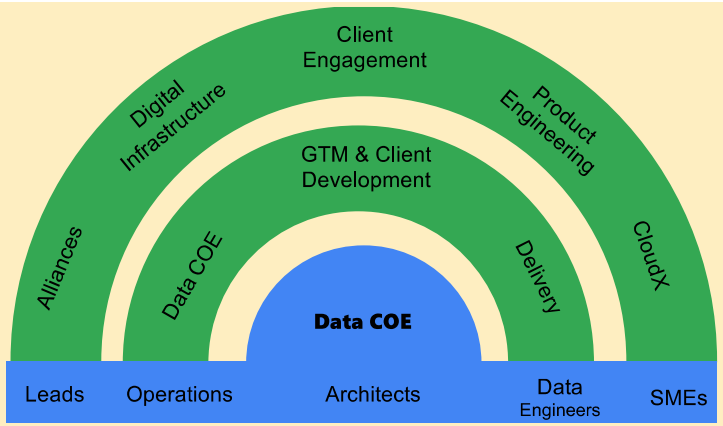
Solutions and GTM Support
- Asset Development
- Proof of Concepts
- Solution Architecture

Thought Leadership
- Point of Views
- Best Practices
- Reference Architecture

Talent Development
- Learning & Certifications
- Build a pool of Google Professional Architects and Data Engineers

External Collaboration
- Alliances
- Analyst Relations
- Webinar, White Papers, Blogs

Community



Competencies

- Dataflow
- Dataproc
- Pub/Sub
- BigQuery
- Bigtable
- Spanner
- Cloud Functions
- Looker
- Data Studio
- Cloud Storage
- IoT Core
- AI Platform
- App Engine
- Cloud Vision
- Data Catalog
- Kubernetes Engine
- Cloud Build
- Composer
- IAM & Admin
- Compute Engine

reach us @ coe.daedataengg@brillio.com

OUR DIGITAL TRANSFORMATION ENABLERS

DRIVEN BY PRODUCT MINDSET

DESIGN THINKING

CONTENT COLLABORATION

DESIGN STUDIO

PRODUCT ENGINEERING

- OMNI CHANNEL APPS
- MICROSERVICES/MESH ARCHITECTURE
- MODERN APPS & CONTAINERIZATION
- DEVOPS
- LOW/NO CODE SOLUTION
- COGNITIVE TESTING

CUSTOMER EXP PLATFORMS

- CRM IMPLEMENTATION
- MARKETING/SERVICE CLOUD
- SERVICE BOT
- HYBRID INTEGRATION
- INTELLIGENT SALES & E-COMMERCE

DATA & ANALYTICS

- MASTER DATA MANAGEMENT
- DATA MIGRATION
- DATA LAKE ON CLOUD
- AI/ML
- ANALYTICS AS A SERVICE

DIGITAL INFRASTRUCTURE

- CLOUD STRATEGY & MIGRATION
- DIGITAL OPERATIONS
- ROBOTIC PROCESS AUTOMATION
- MANAGED SERVICES
- ZERO OPS
- SECURITY & COMPLIANCE

ADVANCED TECHNOLOGY GROUP

TECH STRATEGY & CONSULTING | TECH LABS | ENTERPRISE ARCHITECTURE | BLOCKCHAIN | EDGE | SERVERLESS COMPUTING

Microsoft | Google Cloud | aws | salesforce | SAP | paloalto NETWORKS | Digital Asset | snowflake

Accelerators: brillio BOLT™ | brillio SMARTEST™ | brillio DEVOPS TOOLKIT | brillio ONE™ | brillio CLIP

brillio

EVOLVING NATURE OF DATA IS DRIVING THE NEED FOR NEW DATA SOLUTIONS WITHIN THE ENTERPRISE

NEW WAVE OF DATA

GROWTH OF DATA

Exponential growth in data volumes provides opportunities to analyze historical patterns across multiple data attributes

NEW DATA FORMATS

Multiple new data formats offer opportunities for greater innovation and capabilities for monetization

NEWER CONSUMPTION PATTERNS

New consumption patterns is driving data democratization leading to greater business agility and ease of decision making

DATA POWERED SOLUTIONS

- CUSTOMER 360
- DATA MONETIZATION
- SELF-SERVICE BI
- ADAPATIVE DATA DISCOVERY

WHY GOOGLE CLOUD

For the past 14 years, Google has been building fastest, most powerful, high quality cloud infrastructure on the planet

3X more undersea cables than any other cloud.

Even Apple uses Google Cloud Platform for data storage for its iCloud services

Most of the Bigdata frameworks like GFS, MapReduce, Kubernetes etc. are being innovated by Google itself.

The Google security model is an end-to-end process, built on over 15 years of experience focused on keeping customers and their data safe.

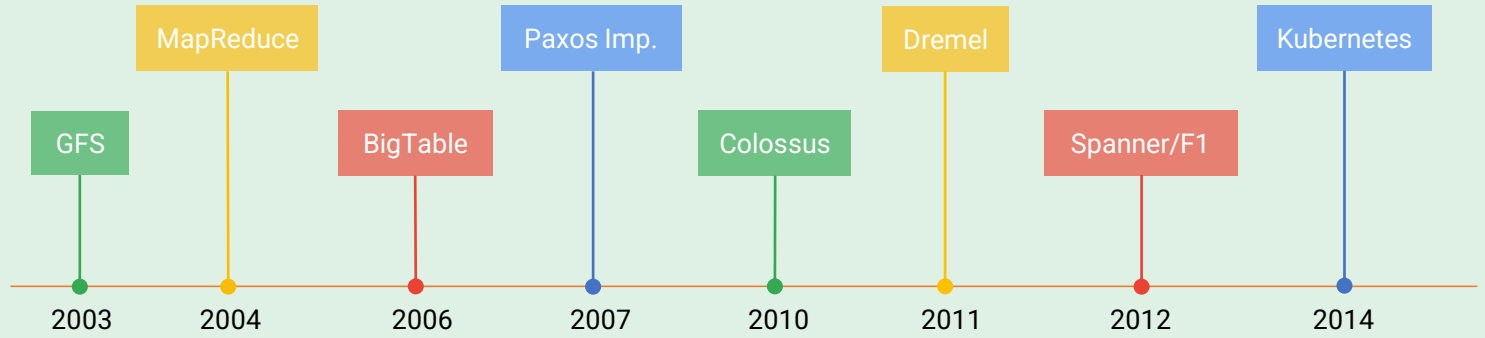
Yes, We can power that

- Web
- Mobile
- Storage & Database
- Big Data
- Highly Scalable System
- Data Mining

Available In

25	76	144	200+
Regions	Zones	Network Edge Location	Countries

Coming Soon! Doha, Toronto, Melbourne, Delhi, Paris, Milan, Santiago

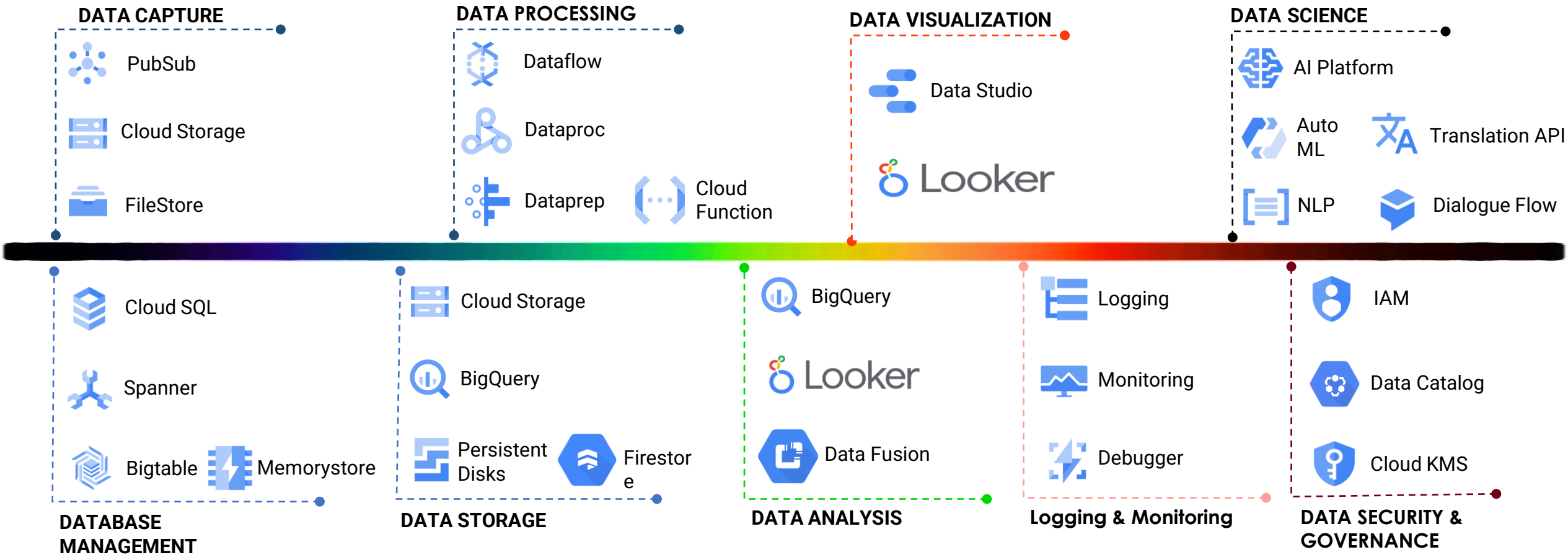


Google Cloud is built on the same infrastructure that powers Google and when you build on top of it, you are building on Google's fast, scalable, and highly reliable infrastructure.

Pricing: 15-41% less expensive than other cloud

Sample Instance Type	Monthly $	GCP Instance Type	Monthly $	GCP x% less Expensive
Standard	87.60	Custom 2 core 8 GB	54.82	37.42%
High-MEM	121.18	n1-standard-4	102.20	15.66%
High-CPU	76.65	Custom 2 core 3.75 Gb	44.66	41.74%

SPECTRUM OF GCP SERVICES MAPPED TO THE DATA SUPPLY CHAIN

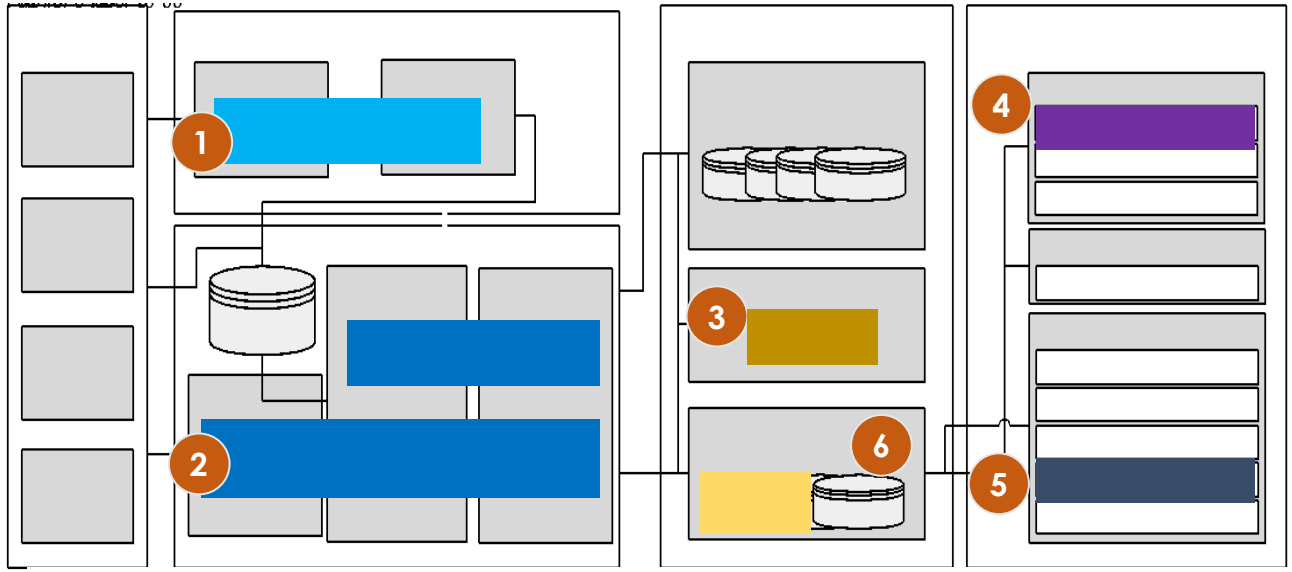


brillio

DATA WAREHOUSE MODERNIZATION ON GOOGLE CLOUD PLATFORM

TOGETHER WE KNOW HOW!!

THERE ARE KEY CHALLENGES OPERATING A LEGACY WAREHOUSE IN THE NEW AGE OF DATA



- Non-Relational Data Store
- Relational Data Store
- Stored procedures /OLAP cubes
- Analytics Sandbox
- BI and Reporting Layer
- Data Archive Store

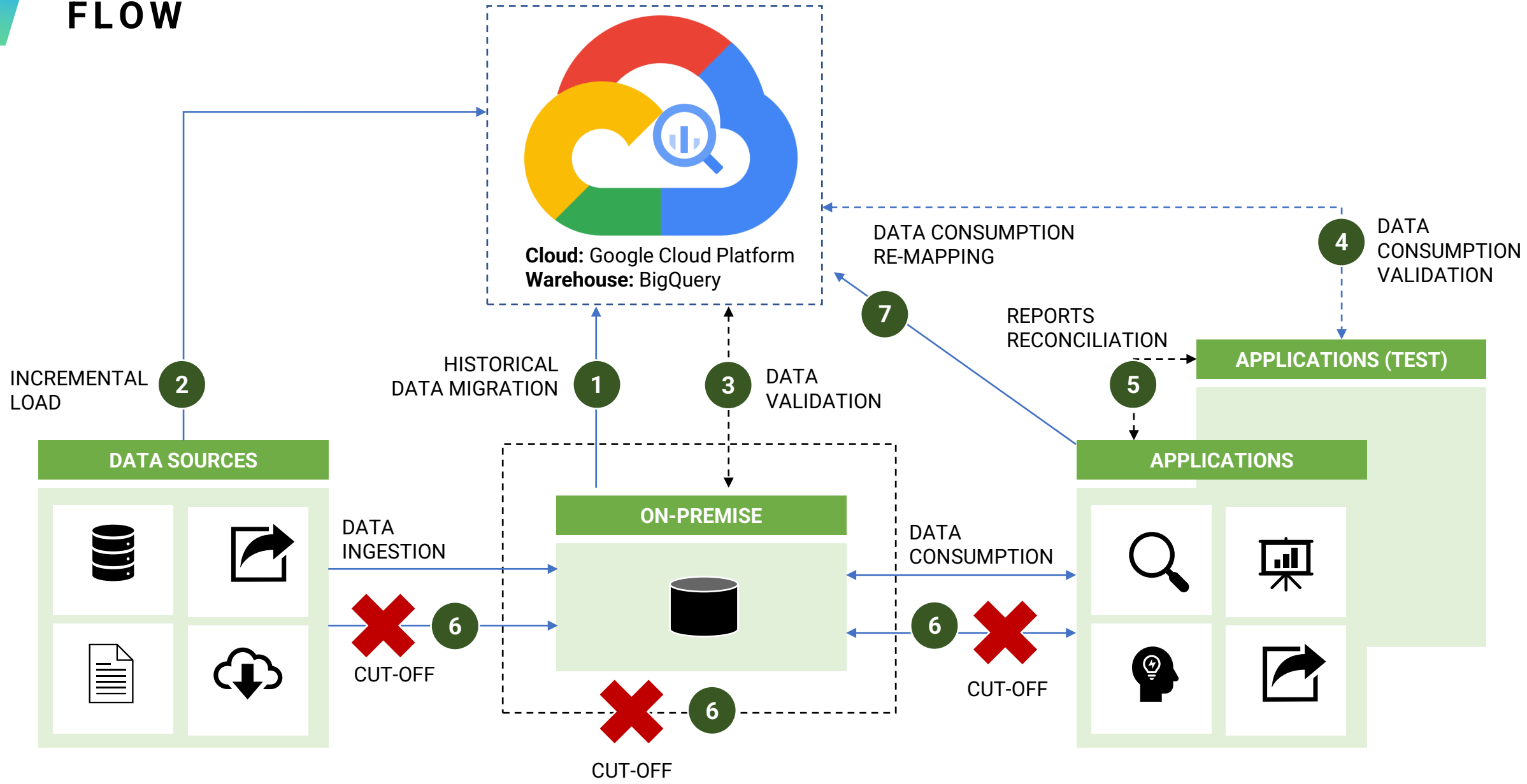
1. Difficult to maintain and expand custom extract and transformations processing non-relational data.
2. Slow running ETLs that address data ingestion, data quality, data preparation and aggregation become a very expensive scaling proposition – requiring more compute resources
3. Pushdown SQLs for relevant DW technologies take up expensive resources and starving critical workloads
4. Inflexible data for analytics as requirements evolve. Unmanageable and slow ad hoc analytical queries.
5. Slow BI reporting jobs dropping out of SLAs all the time. Unacceptable response time for queries.
6. Unused and dormant data living around taking up expensive storage space

ENTERPRISES ARE MOVING TO BIGQUERY FOR MODERNIZING THEIR LEGACY DATA WAREHOUSE ON CLOUD

#	Feature	Snowflake	BigQuery
1	Setup	Need additional efforts to setup storage and compute.	BigQuery follows serverless architecture- No additional efforts for setup
2	Scalability	Managed by users through auto-scaling	Users don't need to worry about scaling at all- everything is handled under the hood
3	Machine Learning	No native ML functionality in database	Provides the ability to train and use machine learning models right there in the database
4	Streaming	No option to stream data directly into table	Streaming data directly into table is possible
5	Cross-region replication	Comparatively difficult to enable.	Very easy to setup.
6	Storage cost	it doesn't matter how old your data is; you pay a single rate	Low storage cost. $0.01 per GB for any table or table partition that has not been modified for 90 consecutive days
7	Compute cost	Time-based, users are charged for execution time	query-based, users are only billed for data that scanned for their queries

- Serverless architecture
- Native ML, Streaming functionality
- Low storage cost
- Query over administration
- Granular level of security- Table, Row and column based

DATA WAREHOUSE MODERNIZATION ON GCP – PROCESS FLOW



MODERNIZATION ON GOOGLE CLOUD - KEY MIGRATION STEPS

Client Stakeholder

DISCOVERY

- Technical discovery of the warehouse
- Assess workloads readiness for cloud
- Define target arch. and technical pattern

ETL Process Migration

- Identify required transformations
- Identify target system configuration
- Optimize process as per target DDL
- Perform ETL process conversion
- Apply and orchestrate target ETL

SCHEMA OPTIMIZATION

- Extract Source DDL
- DDL Conversion
- Optimize and denormalize (if required) Schema
- Create Target DDL
- Apply Target DDL

QUERY CONVERSION

- Identify target SQL format
- Perform query conversion
- ANSI SQL Conversion and determine the mapping file
- Perform manual conversion, wherever applicable

DATA VALIDATION

- Connect source & target DBs
- Connect tables as flat files
- Row by row comparison using MD5
- Identify column level mismatch

DWH MOVEMENT

- Classify tables as per usage
- Perform data extraction
- Identify targets for DB objects
- Load data into targets

Towards successful migration

Google Cloud

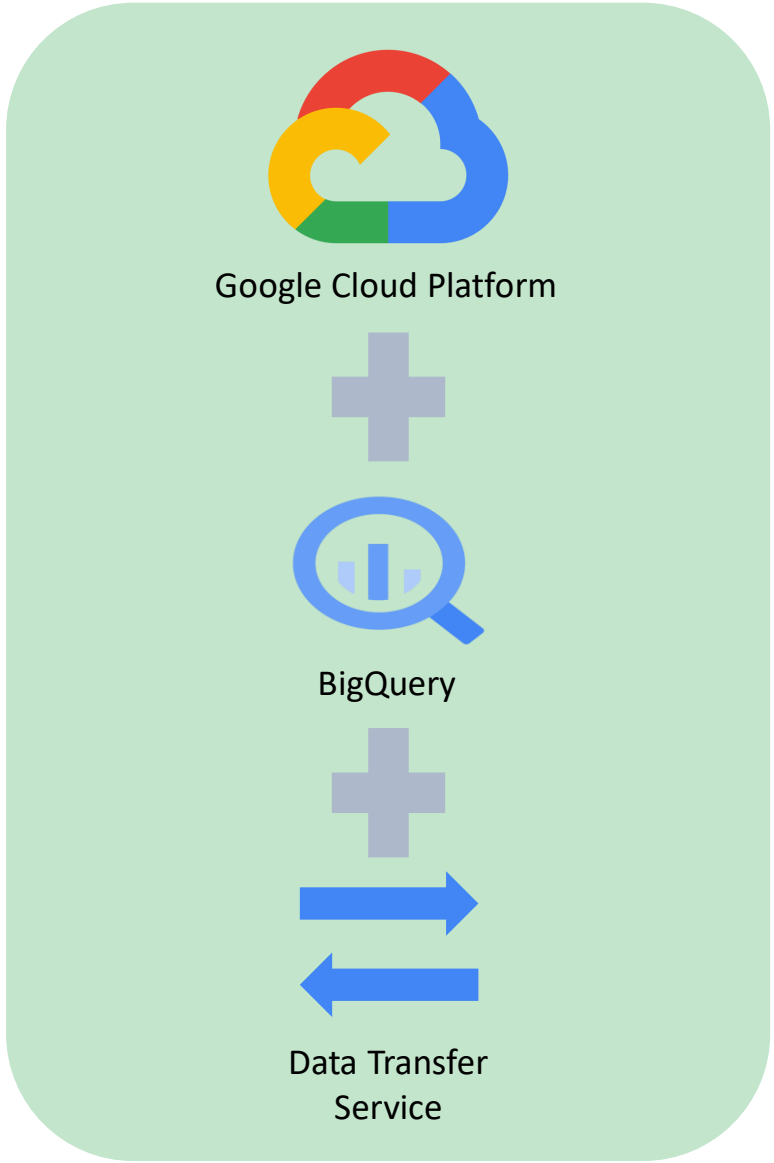
brillio

brillio

DATA MIGRATION STRATEGIES FOR GOOGLE CLOUD PLATFORM

TOGETHER WE KNOW HOW!!

BIGQUERY DATA TRANSFER SERVICE



About

The BigQuery Data Transfer Service automates data movement into BigQuery on a scheduled, managed basis. Your analytics team can lay the foundation for a BigQuery data warehouse without writing a single line of code.
After configuring a data transfer, the BigQuery Data Transfer Service automatically loads data into BigQuery on a regular basis. Data backfills can also be initiated to recover from any outages or gaps.

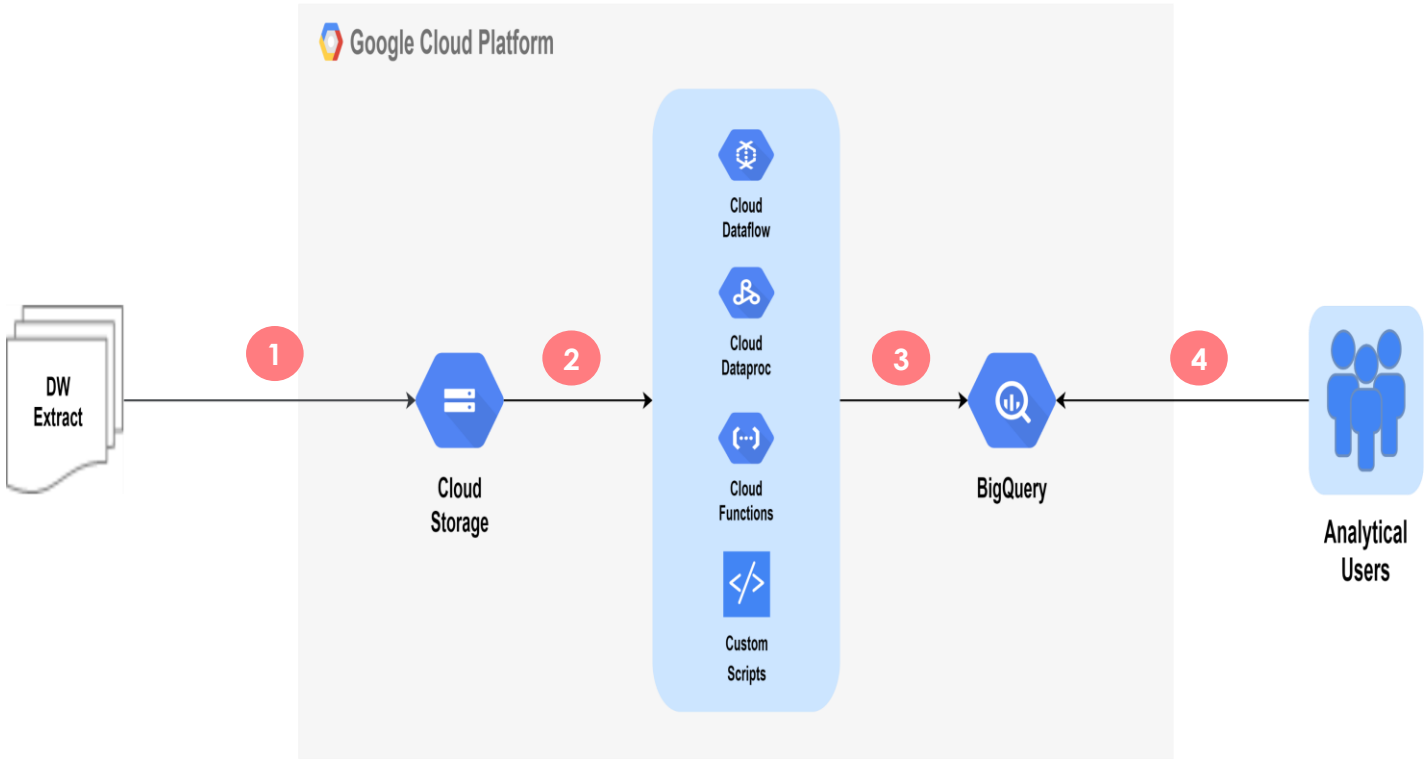
Supported data sources

- **Google SaaS apps** (Campaign Manager, Cloud Storage, Ad Manager, Ads, Merchant Center, Play, Search Ads 360, YouTube)
- **External cloud storage providers** (Amazon S3)
- **Data warehouses** (Teradata, Redshift)

Way of Access

- Cloud Console
- bq command-line tool
- BigQuery Data Transfer Service API

DATA MIGRATION ON GCP – TRANSITION APPROACH



1. Create data extract from various source systems. Send data extract to GCS buckets periodically.

2. GCS serves the purpose of Data Lake. Various data processing may take place in GCS using Dataflow or Dataproc-Spark or Cloud Functions or Custom Scripts for data transformation based on business requirement.

3. Selected data are sent to BigQuery to serve warehouse and analytical purposes

4. Users can run standard SQL and other reporting tools including data visualization tools for data analysis.

brillio

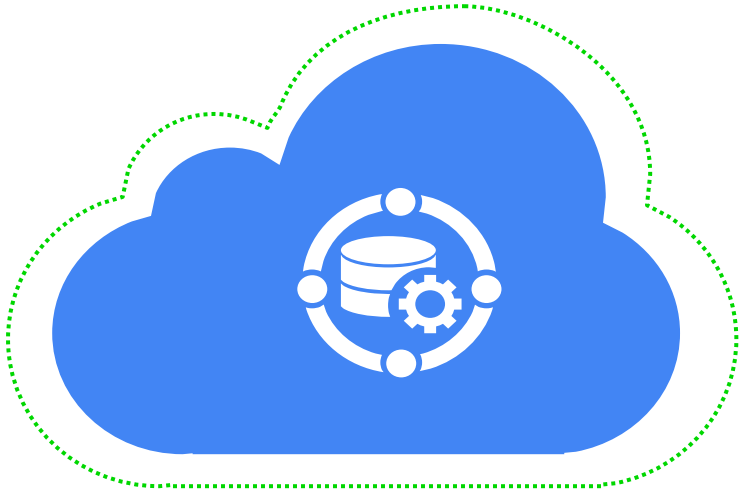
BRILLIO OFFERINGS FOR GOOGLE CLOUD PLATFORM

TOGETHER WE KNOW HOW!!

BRILLIO DATA OFFERINGS FOR GCP

DATA ON GCP

- On-Premise Data warehouse modernization on GCP
- On-Premise Data Lake Industrialization on GCP



DATA FOUNDATION ON GCP

- Design and Implement enterprise data solutions leveraging GCP data services
- Greenfield Data warehouse and Data Lakes at speed and scale leveraging GCP data services

DATA ON GCP CLOUD

MODERNIZING ON-PREMISE DATA WAREHOUSE AND DATA LAKES ON GCP

SCALE TO GROW

Organization could not scale its on-premise data lake to respond to business needs.

SOLVE END OF LIFE

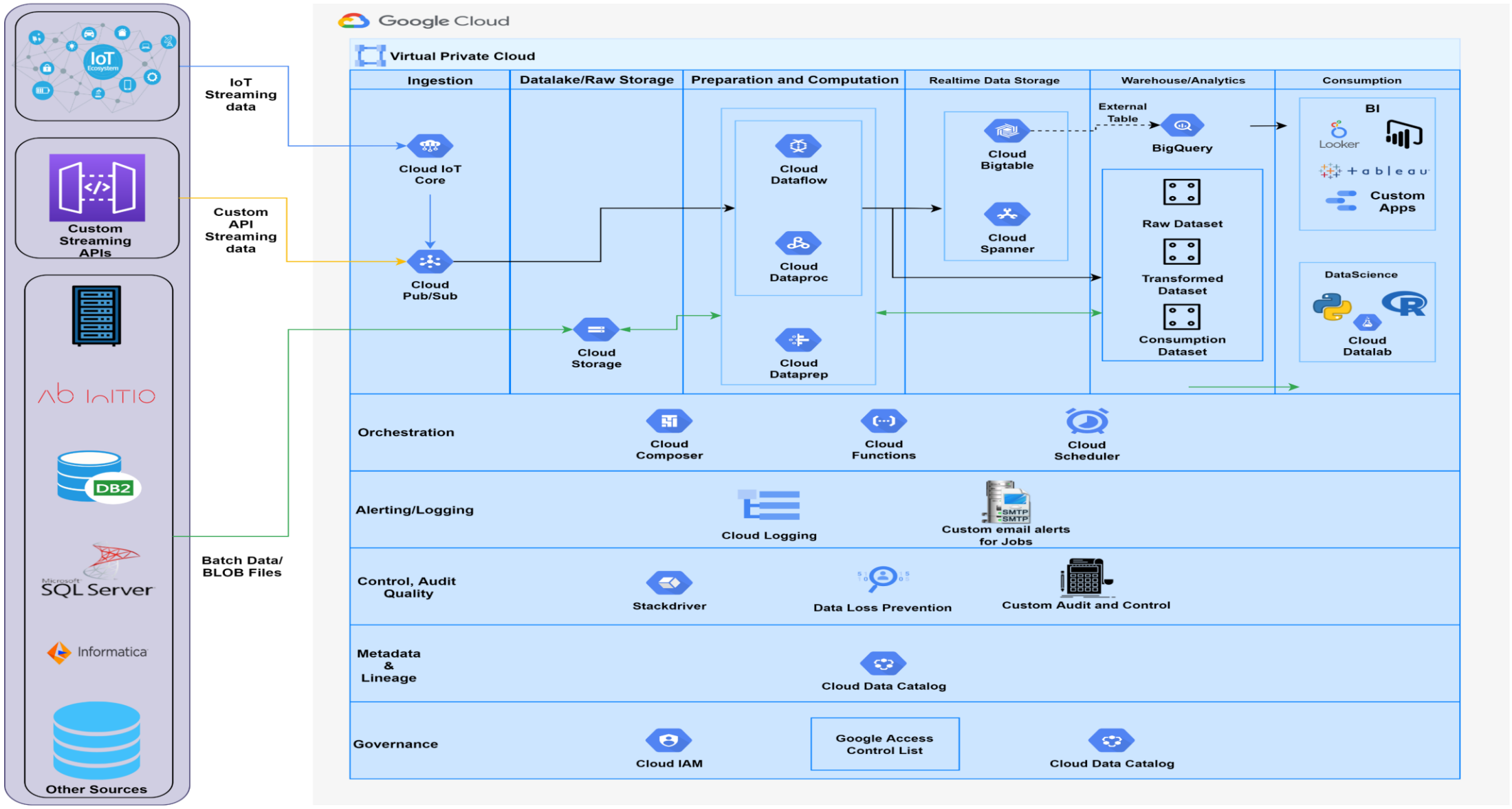
The on-premise data warehouse support was expected to end soon.

SINGLE VERSION OF THE TRUTH

Multiple on-premise data repositories forced analytics & business team to work in silos.

DATA FOUNDATION ON GCP

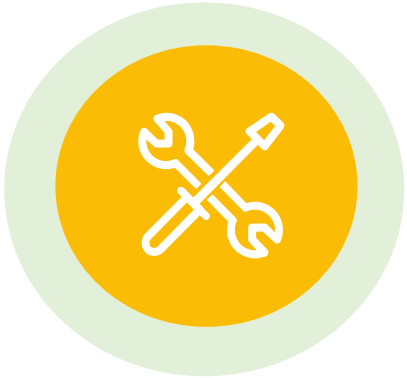
LEVERAGE GCP DATA SERVICES TO BUILD NATIVE SOLUTIONS ACROSS THE DATA SUPPLY CHAIN



HOW CAN BRILLIO HELP YOU ?

DATA BUSINESS CASE FOR MODERNIZATION & NATIVE DEVELOPMENT

MIGRATION & MODERNIZATION STRATEGY



TECHNOLOGY PROOF OF CONCEPT

AGILE EXECUTION FOR TRANSFORMATION AT SCALE

CLOUD WAREHOUSE OPERATIONS / MANAGEMENT

COMBINING BRILLIO AND GCP CAPABILITIES AND STRENGTHS

INTELLIGENT DATA ASSETS

+

INDUSTRY SOLUTIONS

+

Google Cloud

+

ANALYTICS AUTOMATION

+

CLIENT INTIMACY & INDUSTRY CONTEXT

brillio

WE RECOMMEND A PHASED APPROACH

PHASE 1
DESIGN: DATA DISCOVERY & DESIGN

PHASE 2
TRANSITION: DATA MIGRATION TO GCP

PHASE 3
TRANSITION: MANAGED DATA SERVICES ON GCP

ARCHITECTURE STRATEGY:

What data to move and where to move it?

DESIGN ARCHITECTURE:

Design data supply chain on GCP cloud

TRANSITION ARCHITECTURE:

Move the right data to GCP Cloud leveraging Brillio tools & assets

CONSUMPTION SERVICE ARCHITECTURE:

Enable agile data services on GCP with GCP native services and 3rd party consumption tools

EXECUTION ARCHITECTURE

Derive on-going value from Data on GCP Cloud

LET US CREATE SOMETHING AMAZING!

TOGETHER.

A BRILLIO PRESENTATION

brillio