Meeting Security Requirements for Apache Beam Pipelines on Google Cloud

Lorenzo Caggioni
Google
linkedin.com/in/lcaggio/

BEAM SUMMIT
NYC 2023

Agenda

Securing a Beam Pipelines on Google Cloud

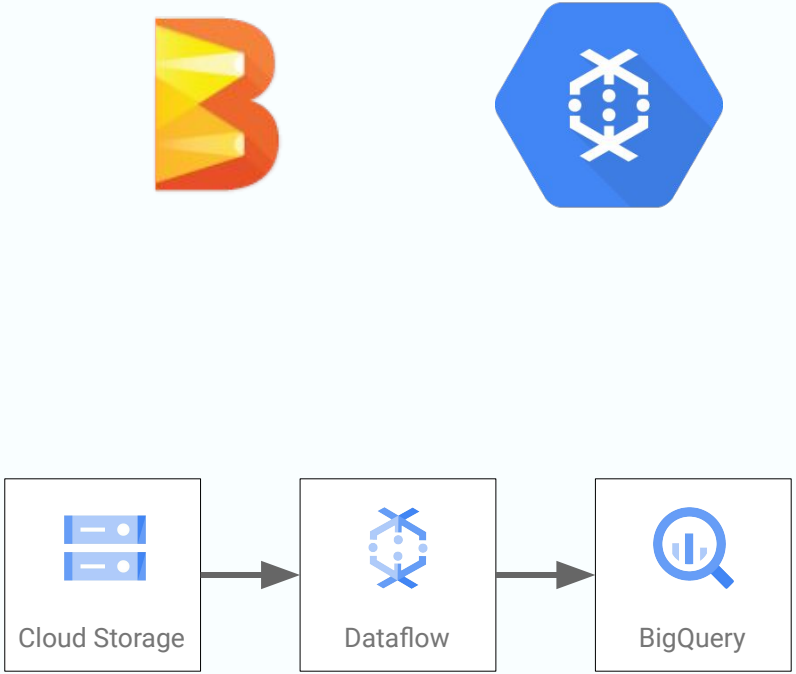
- Private resources
- Role separation and least privileges
- Data Encryption at rest

Customer requirements

1. Internal addressment of tenants must be private.

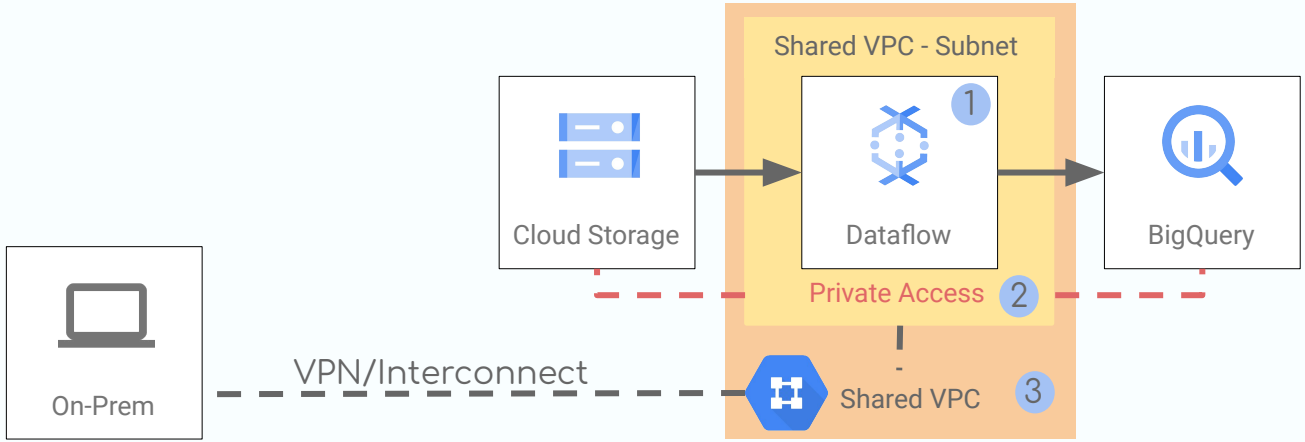
2. Every tenants must be isolated and dedicated to a specific system of services.

3. All data must have encryption at-rest with keys managed by ACME's security team.



1. Internal addressment of tenants must be private.

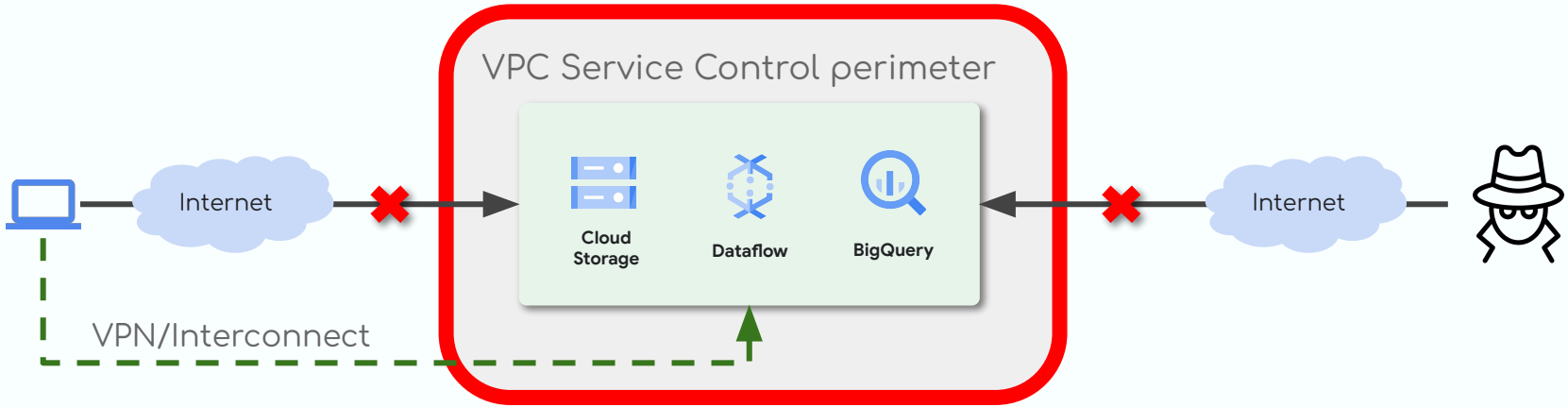
1. Set `disable-public-ips` when deploying the pipeline
2. Enable `Private Access` on the subnet to access GoogleAPIs
3. Network: shared-VPC



1. Mitigate Data Exfiltration

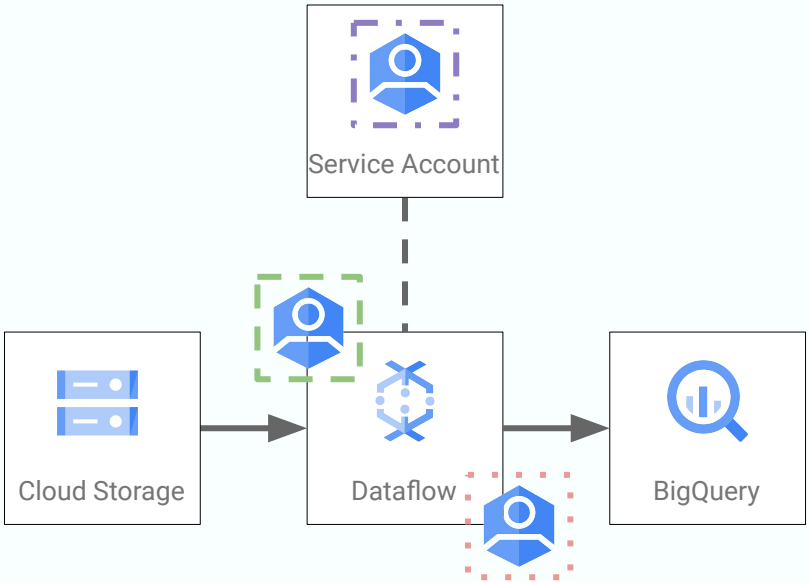
VPC Service Controls helps preventing **data exfiltration** and controlling access to **Google APIs**.

Isolate resources of **multi-tenant Google Cloud** services to mitigate data exfiltration risks.



2. Tenants must be isolated

IAM and Service Accounts



Dataflow Service Agent

```
roles/dataflow.serviceAgent
roles/compute.networkUser
```

Worker Service Account

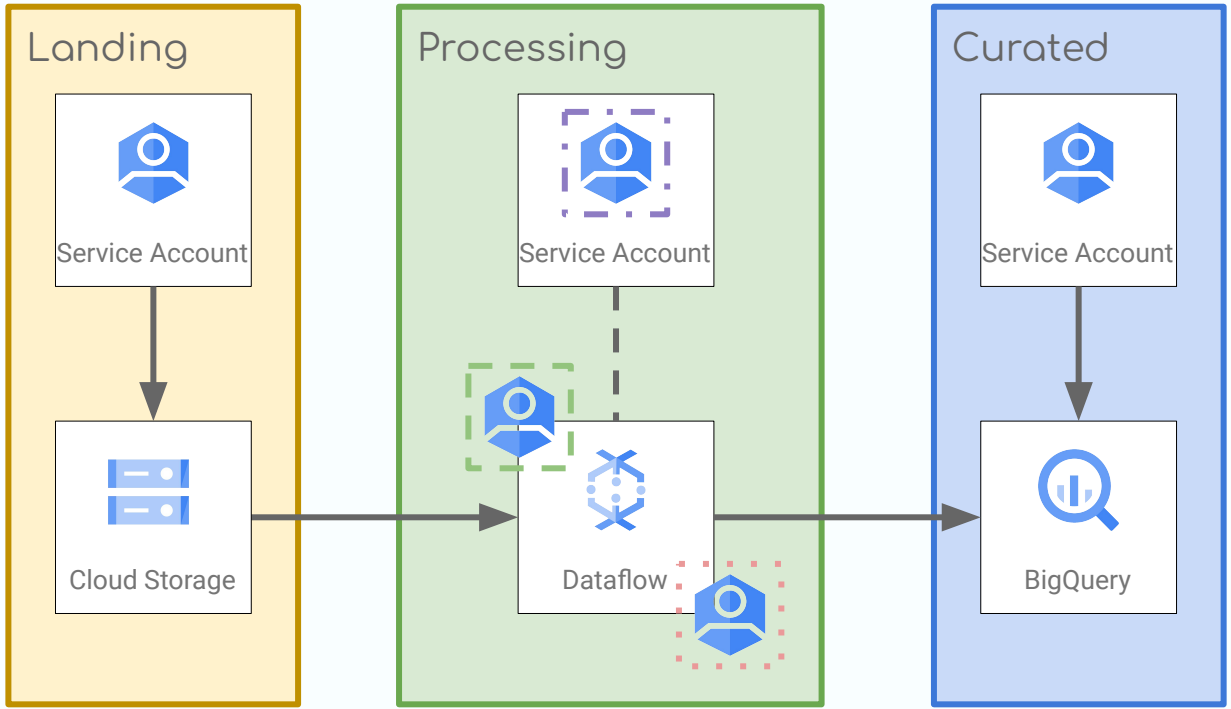
```
roles/storage.objectAdmin
roles/dataflow.worker
roles/bigquery.dataEditor
```

Job orchestrator

```
role/iam.serviceAccountUser
role/dataflow.admin
```

2. Tenants must be isolated

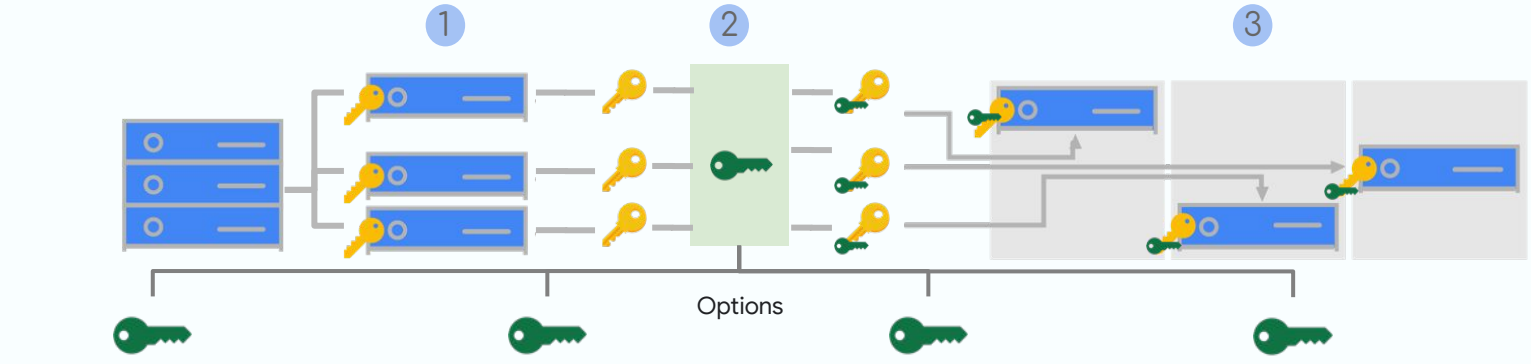
Project separation



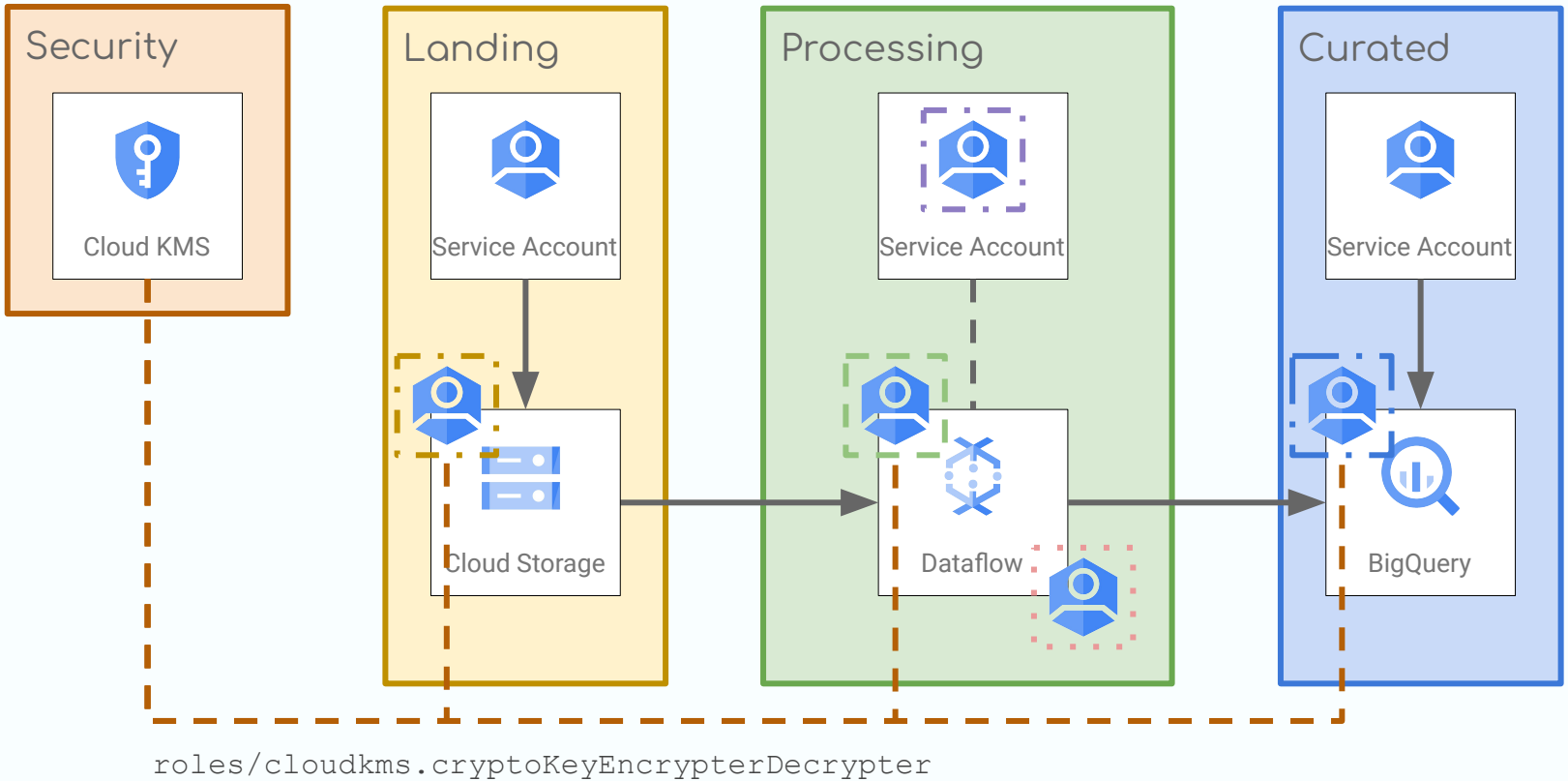
3. At rest encryption

Data at rest are encrypted on GCP:

1. Data split in chunk and encrypted with a key: Data Encryption Key (DEK)
2. DEK encrypted with Key Encryption Key (KEK)
3. Chunk stored with encrypted DEK

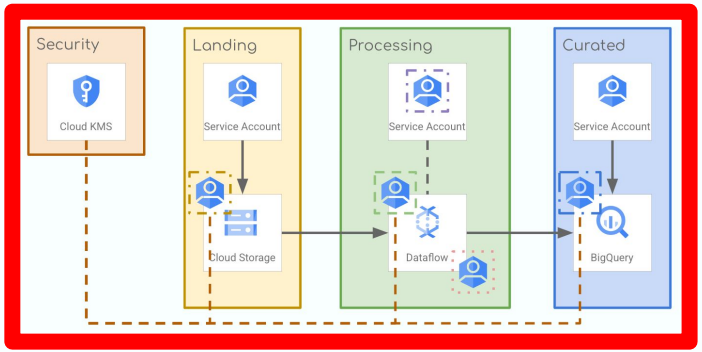


3. At rest encryption



Recap

1. ✓ Every tenants must be isolated and dedicated to a specific system of services.
2. ✓ Internal addressment of tenants must be private.
3. ✓ All data must have encryption at-rest with keys managed by ACEME's security team.



SCAN ME

End to end example

Lorenzo Caggioni

QUESTIONS?

Contact info
https://twitter.com/lcaggio
https://www.linkedin.com/in/lcaggio
https://github.com/lcaggio