



Migrating ING's Data Analytics Platform to Google Cloud

Adnan Hodzic | Tech Lead: Data Analytics Platform GCP
(Public Cloud) area

November 2024 @fooctrl



do your thing



Tech that matters in people's lives

Facts and figures

1/4 of 60.000
works in tech

Located over
15 countries

We serve around
38 million
customers

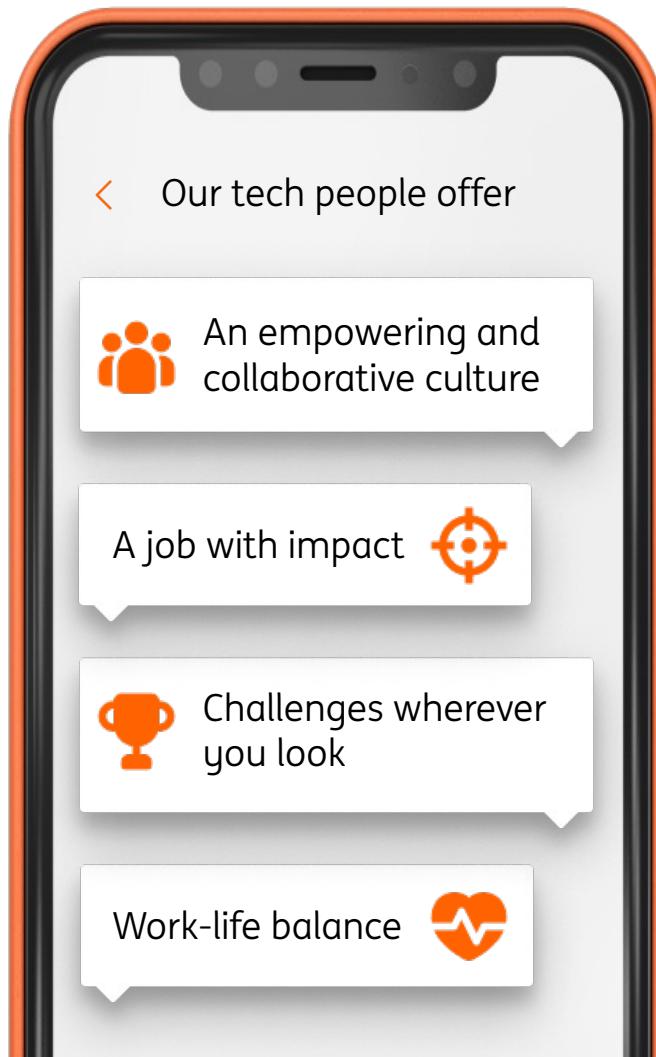
Over
40
countries

Annually
4.5 billion
payments

 **47+**
repositories on
GitHub

 **1.300+**
microservices in
production

 **315.000+**
deployments per week



Providing **superior customer experience** through our scalable technologies:

- ING private cloud (IPC): **65%** of workload on (private) cloud
- Touchpoint: provides modularity and a set of reusable shared services
- OnePipeline: a continuous integration and delivery pipeline

ing.jobs/tech



do your thing

Data Analytics Platform



```
function(a){this.o[a]&&prototype.w=fun
it=function(a,c){this.o.push({fc:a,options:c});var jp=function()
his.b.push(a);_F(d,1)&&(d=_F(d,2))&&this.b.push(d);_x("gapi.load",
(a){_A.call(this);this.C=a;this.w=this.b=null;this.D=0;this.B=window.__jsl=function(a){this
)(&a[1]&&g(PASSWORD("*****"));0<=a.indexOf("MSIE")&&0<=a.indexOf("Trident")&&(a=/\b(?:M
(a,c,d){if(!a.o)if(d instanceof Array)for(var e in d)lp(a,c,d[e]);else{e=(0,_v)(a.F,a,c);var
=e;c&&c.addEventListener?c.addEventListener(d,e,l1):c&&c.attachEvent?c.attachEvent("on"+d,e):e
&&this.b.type==c&&this.w==a&&(d>this.b,this.b=null);if(e=a.getAttribute("data-eqid"))a.PASSWOR
.stener?a.removeEventListener(c,e,l1):a.detachEvent&&a.detachEvent("on"+c,e):this.C.log(Error(
.function(a,c){this.b=c;this.w=a;c.preventDefault?c.preventDefault():c.returnValue=!1};(function
_LDD:a=[];var c=_Ai();igp(window,__jsl(_F(c,8));c=_ec();var d=_W();a=new _to(c,__H(_L(),ep,
_x("gbar.qm",(_v,_v)(function(a){try{a()}catch(g){d.log(g)}},this));_y1("api").Ra());_I(_F(c
_,c),_U(c)))));cp(COPY.PASSWORD("*****"),"DOMContentLoaded");cp(window,"load");_U(c)))();
,_v(_oj,w,_oj,_ac));_x("gbar.mls",function(){_Ma("eq",new kp(_W()));_Ma("gs",new jp,
[_jo,6])||new _jo);(function(){for(var a=function(a){return function(){_ti(44,{n:a})}}},c=0;c
[c];_x(d,a(d))var e=_Ia.U();_Ja(e,"api").Ra();fp(_Ja(e,"m"),function(){_Ja(e,"api").Ra()});
ent.querySelector(".ta");c&&c.querySelector(".gb_b"))&&p(_y1("eq"),c,"click")));var _y1(
,return t
g = function(t) {
    return t + " " + document.documentElement.className;
}
```

Data Analytics Platform



Model development

Model execution

Data analysis

Dashboarding

Data Analytics Platform

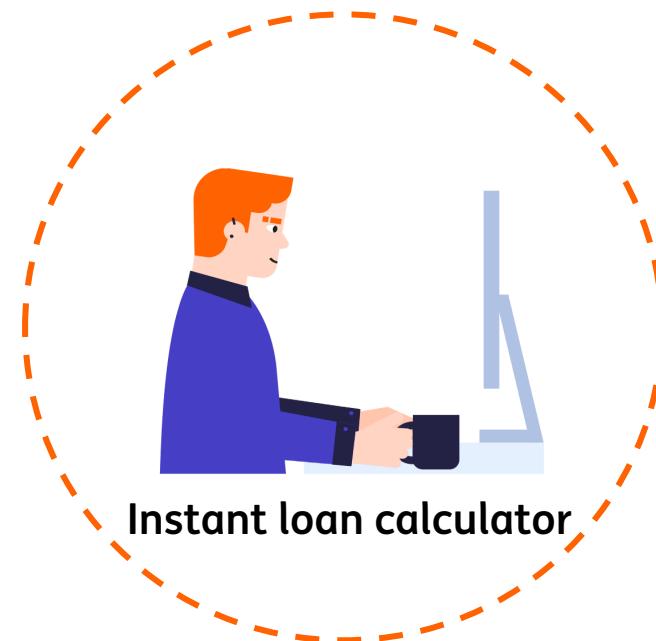


Model development

Model execution

Data analysis

Dashboarding



Business



Experiment



Data scientist



Data engineer

Data Analytics Platform



Model development

Model execution

Data analysis

Dashboarding



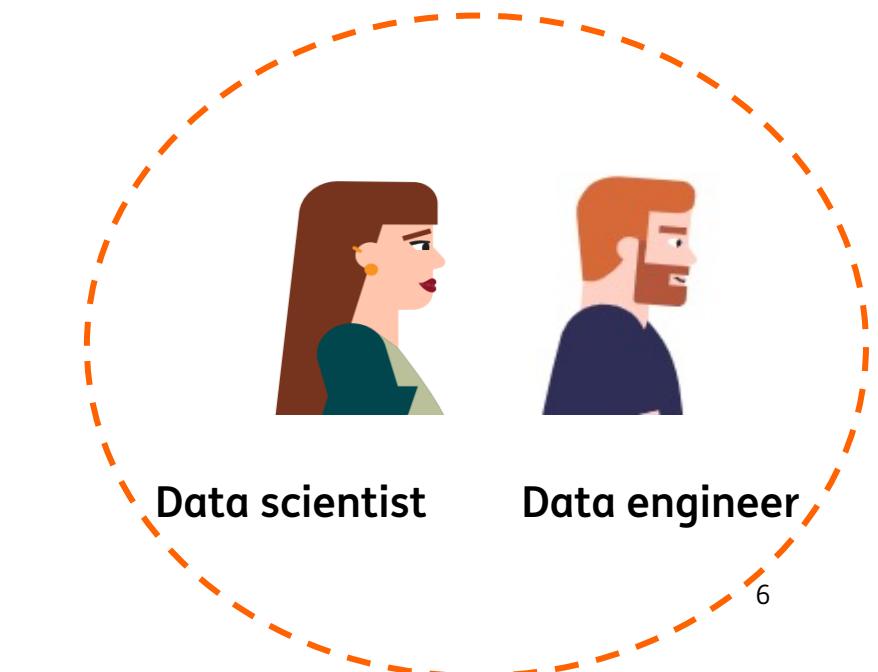
Instant loan calculator



Business



Experiment



Data scientist

Data engineer

Data Analytics Platform



Model development

Model execution

Data analysis

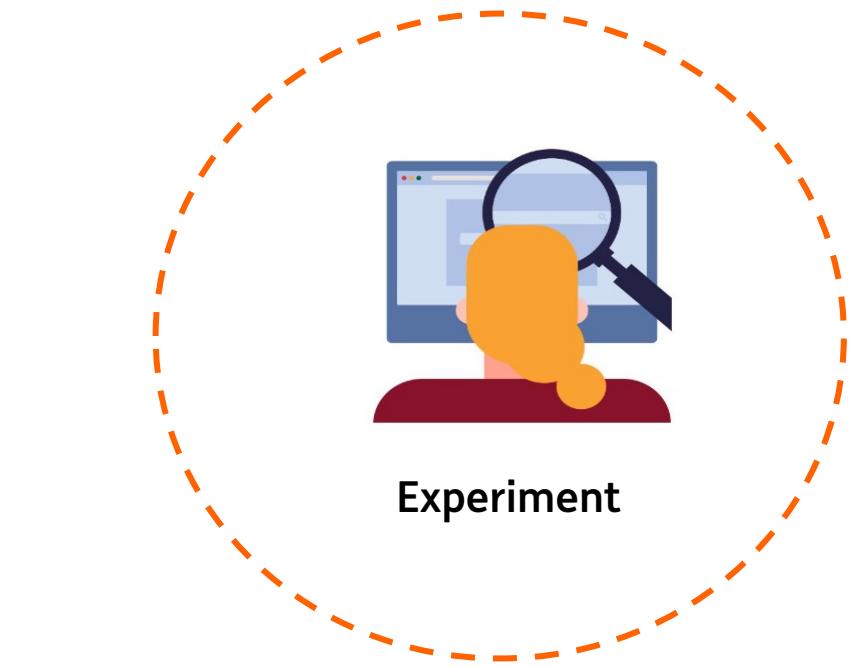
Dashboarding



Instant loan calculator



Business



Data scientist



Data engineer

Data Analytics Platform



Model development

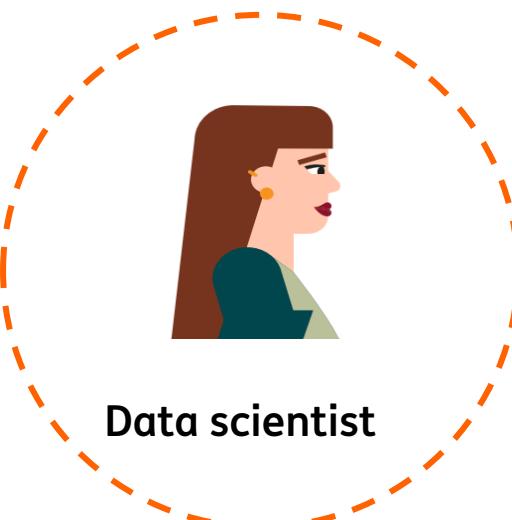
Model execution

Data analysis

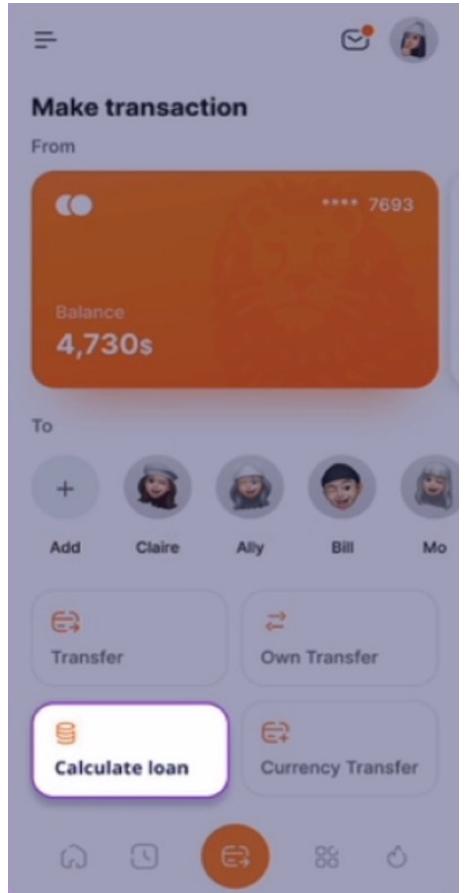
Dashboarding



Execution artifact



DevOps engineer



Data Analytics Platform



Model development

Model execution

Data analysis

Dashboarding



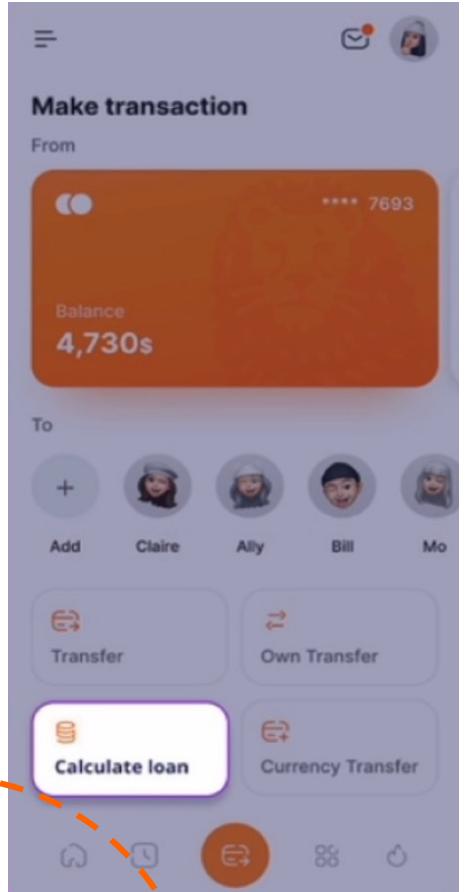
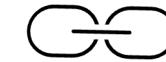
Execution artifact



Data scientist



DevOps engineer



Data Analytics Platform



Model development



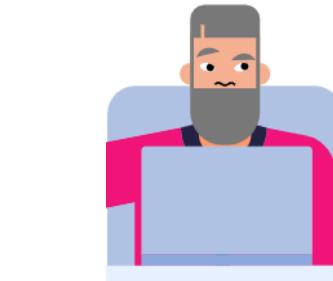
Model execution

Data analysis

Dashboarding



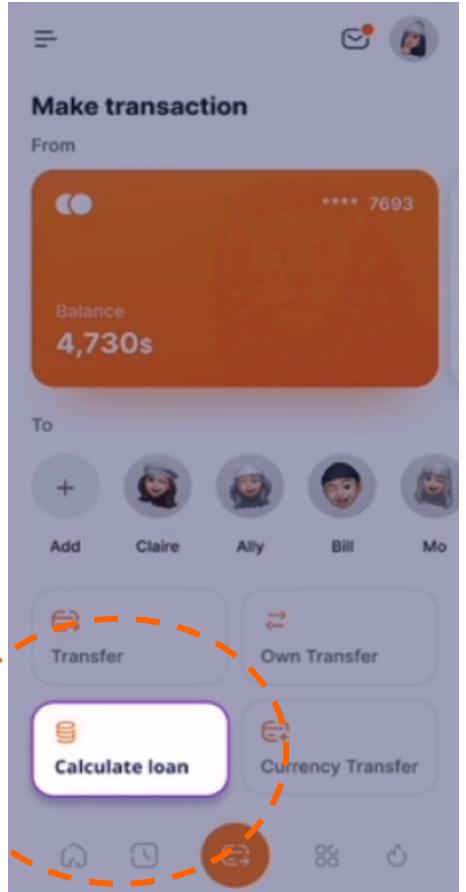
Data scientist



DevOps engineer



Execution artifact



Data Analytics Platform



Model development

Model execution

Data analysis

Dashboarding



Business



Data analyst

Data Analytics Platform



Consists of multiple teams, projects and uses various underlying technologies

Data Analytics Platform



Consists of multiple teams, projects and uses various underlying technologies

Fun fact: DAP is running on one of the largest on prem Kubernetes clusters in ING

Data Analytics Platform



Consists of multiple teams, projects and uses various underlying technologies

Fun fact: DAP is running on one of the largest on prem Kubernetes clusters in ING

End result: dap.ing.net



Data Analytics Platforms

Designed to meet all your analytics needs.

Ensures high security. Makes self-service possible.

[Login](#)

What's new in DAP

DAP on IPC - Recent changes

New simplified approach for storing projects data and datasources.
Improved operation orchestration by Genesis. Introduction...

[March 13, 2013](#)

Features

Data Exploration

Analyse, explore and visualise



- Applications
- Files
- Data sources
- Members
- About this project

Application

Development

Environment

P09304-dap

0.1.0-essenti...

Notebooks

Jup...

Dashboarding

Super...
Visual...

Monitoring

Grafana...
Live m...

New project



Project details

Project name

OCC exploration

Purpose

Describe the purpose of this project ?

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut eni...

Start

Cost center details

Sponsor's email

Type the email of the person sponsoring this project

john.joe@ing.com

Cost center ID

Ask your manager to provide you with the right cost center ID

0912220

Platform details

Platform

Select the platform this project should run on

- ING private cloud (IPC)
- Public cloud (GCP)



Project type

Select what type of work you will be using this project for

- Exploration
- Execution

**Cancel****Create**

**Applications****Files****Data sources****Members****About this project**

Applications

Development

Environment**default-user-environments**
0.1.0-essentials-env.9

See commit



Use GPU

Start**Notebooks****Jupyter****VS Code**

Dashboarding

**Superset**
Visual analytics

Monitoring

**Grafana**
Live metrics

Scheduling

**Airflow**
Schedule jobs



Applications

Data sources

Members

About this project

Applications

Tools

**MLFlow**

Model tracking

**Spark history server**

Detailed insights

Airflow instances

Name

**Production**

Scheduled jobs

3

Status





“Wait, ING is using Google Cloud?”

“Wait, ING is using Google Cloud?”

“Kubernetes, Resistance Is Futile” @KubeCon: youtu.be/issubK49f6E



“Wait, ING is using Google Cloud?”

“Kubernetes, Resistance Is Futile” @KubeCon: youtu.be/issubK49f6E

Public (Google) Cloud, resistance is futile

A photograph taken from the interior of a car, looking out through the front passenger-side window. A young boy with light brown hair, wearing a bright red t-shirt, is leaning forward with his head near the window, looking outside. His right arm is resting on the window frame, and he appears to be wearing a blue wristband. The car's interior is visible, including the headrests of the front seats and a portion of the dashboard. The scene outside shows a lush green forest under a clear sky.

“But why would you use Google Cloud?”

“But why would you use Google Cloud?”

On prem setup & its maintenance is not free nor is it more secure

“But why would you use Google Cloud?”

On prem setup & its maintenance is not free nor is it more secure

Keep your options open with hybrid cloud setup

“But why would you use Google Cloud?”

On prem setup & its maintenance is not free nor is it more secure

Keep your options open with hybrid cloud setup

Resolve scalability limitations & issues

DAP CGP Data Analytics Platform on Google Cloud Platform

DAP's answer to requests to host data analytics related tenants & projects
of critical value to ING on GCP



DAP GCP teams

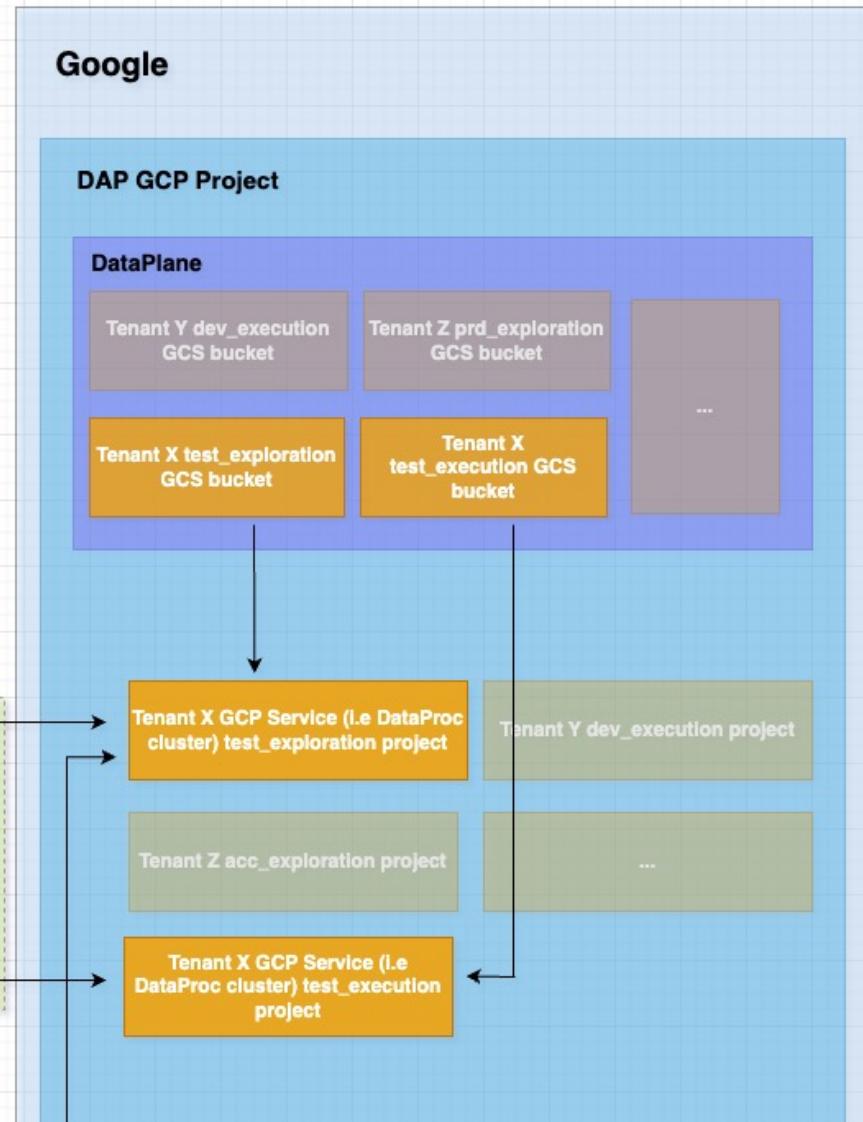
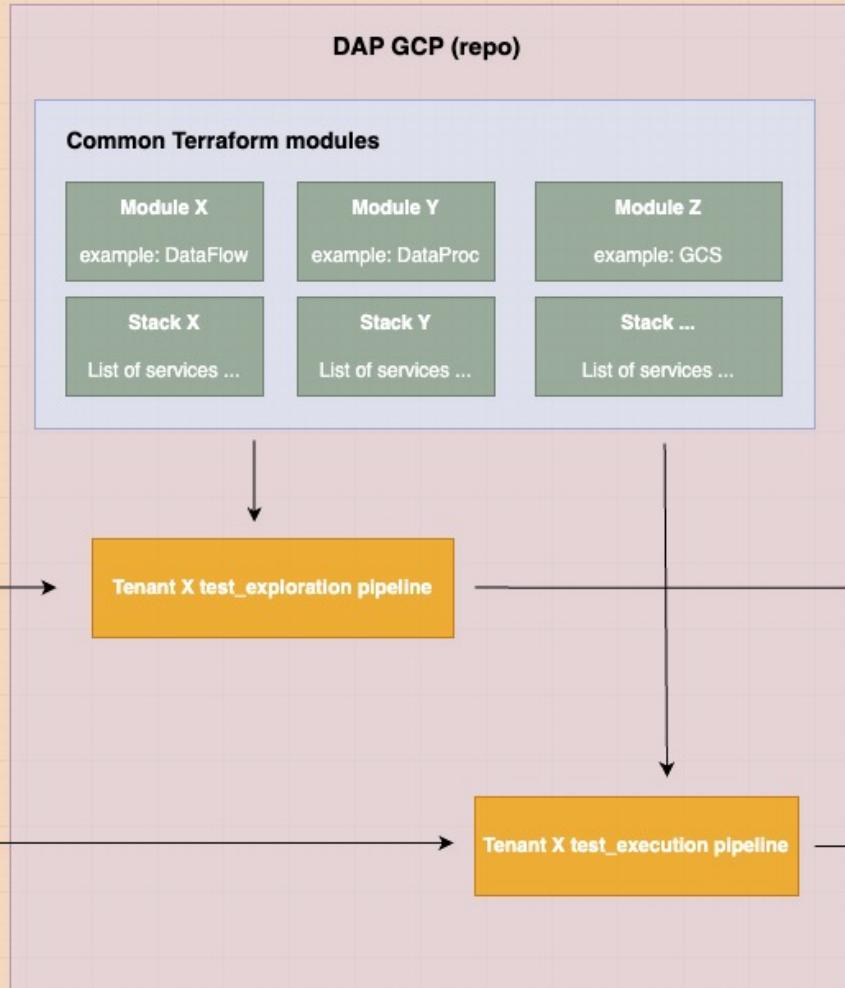
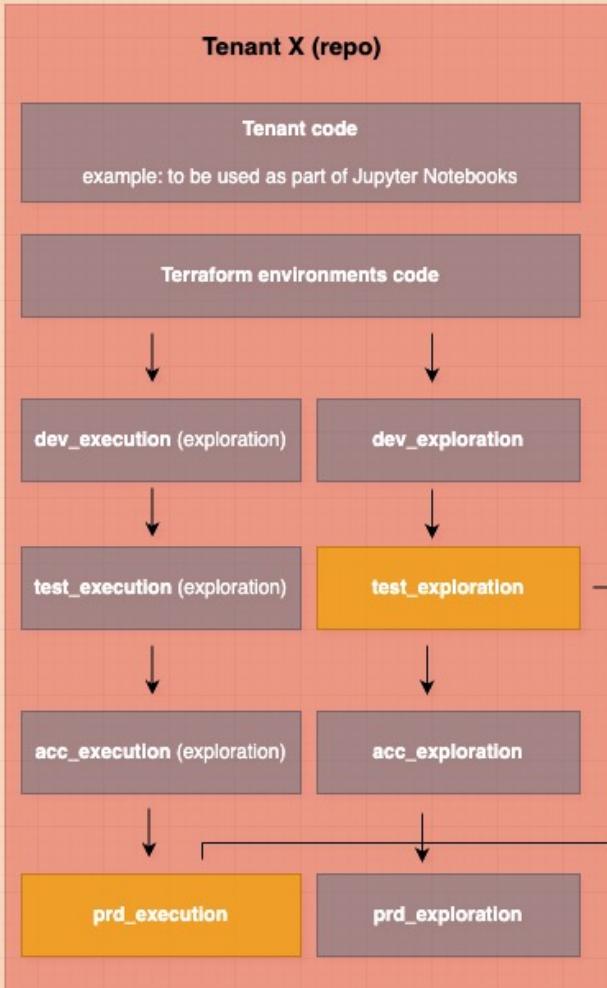
DAP GCP Engineering

DAP CGP Data Analytics Platform on Google Cloud Platform

Mission: do everything it takes to bring tenants & their projects to Production on GCP

Creating necessary pipelines and defining deployment strategies, adhering to separation of responsibilities principle

ING

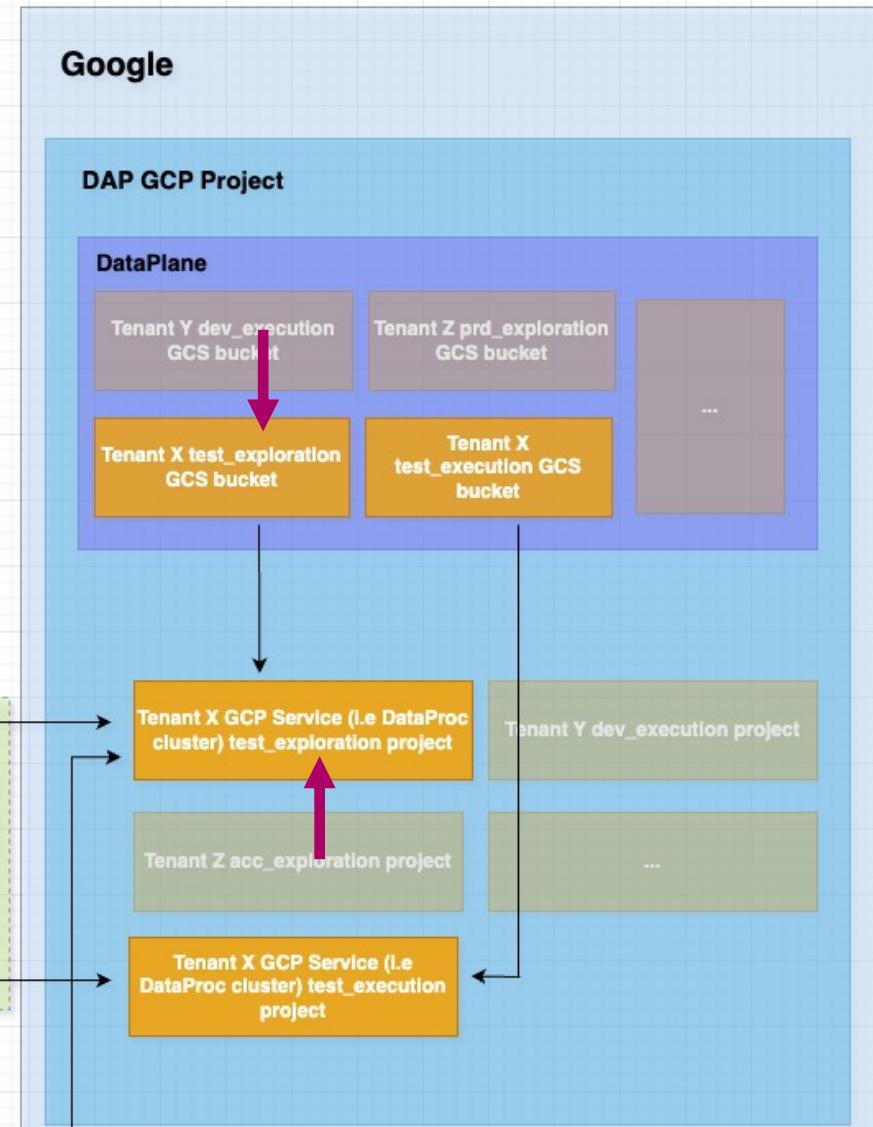
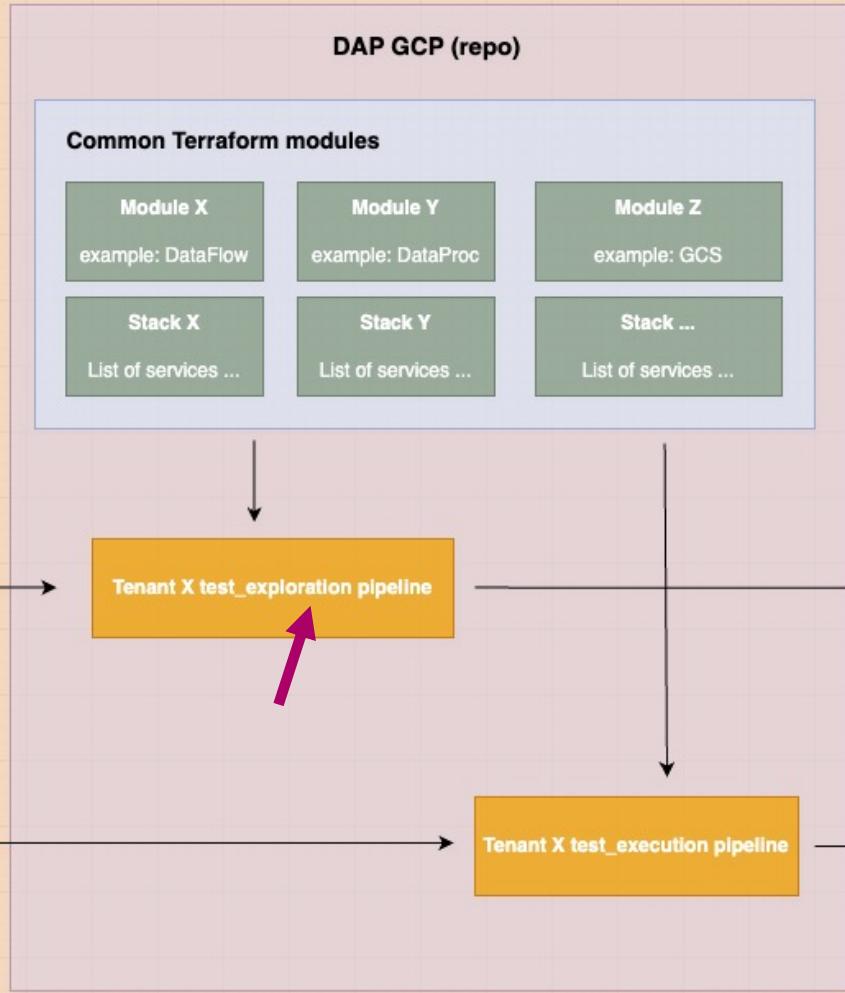
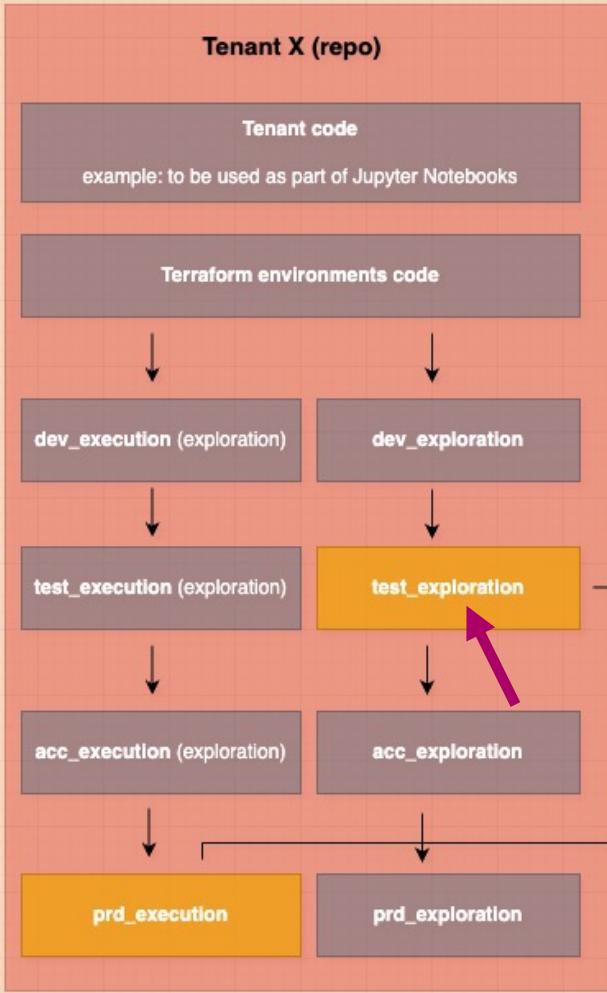


DAP GCP Architecture (simplified)

Showcasing how Tenant X code would be deployed to Tenant X GCP service (i.e DataProc cluster) on test_exploration project in Google Cloud

ING employee i.e: data scientist
acces to i.e Jupyter Notebooks **only**
accessible on **exploration** envs

ING



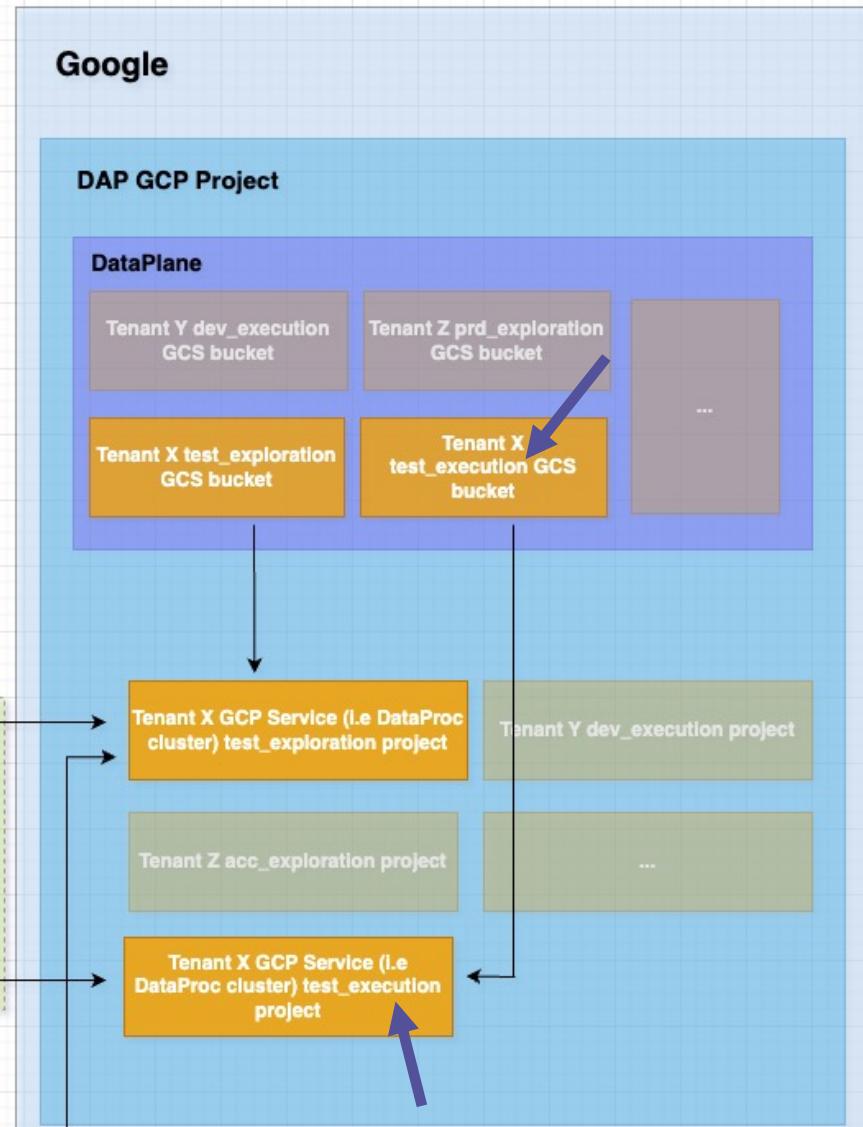
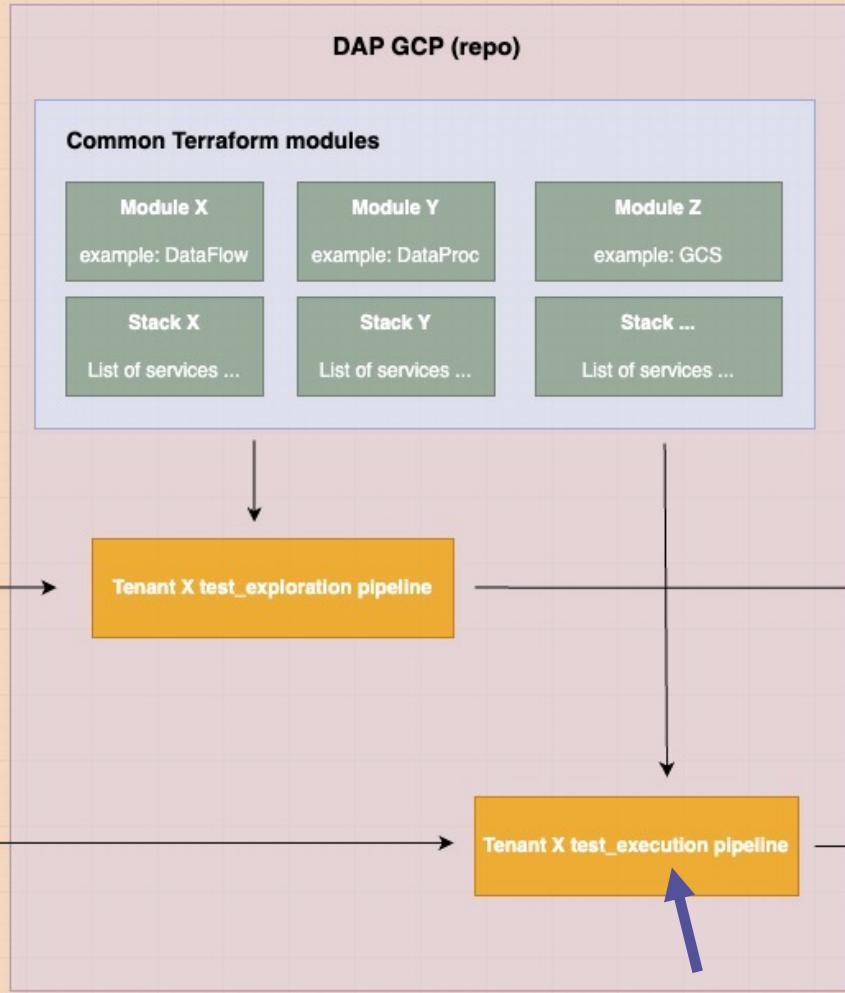
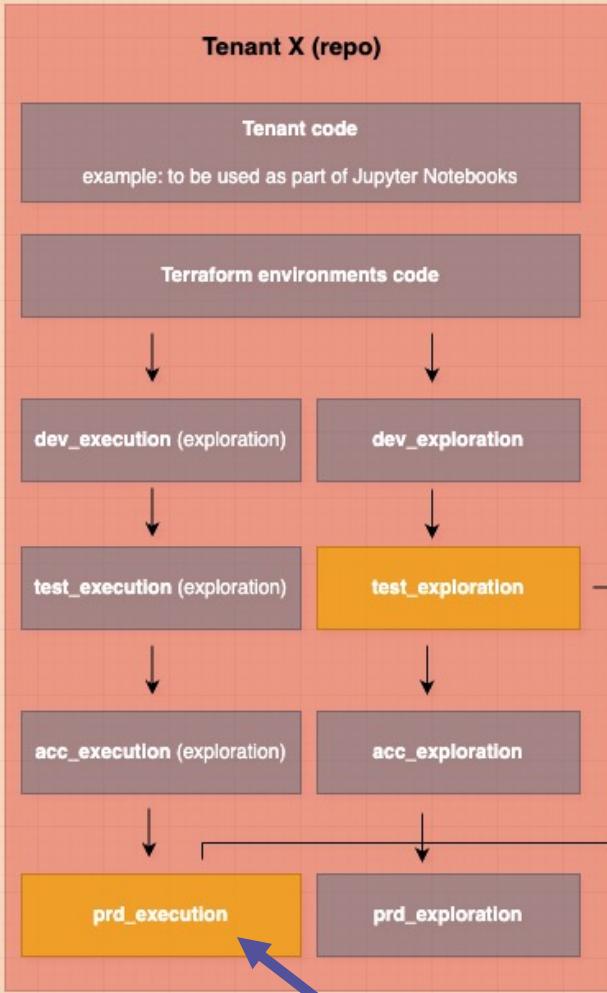
DAP GCP Architecture (simplified)

Showcasing how Tenant X code would be deployed to Tenant X GCP service (i.e DataProc cluster) on test_exploration project in Google Cloud

ING employee i.e: data scientist

acces to i.e Jupyter Notebooks **only**
accessible on **exploration** envs

ING

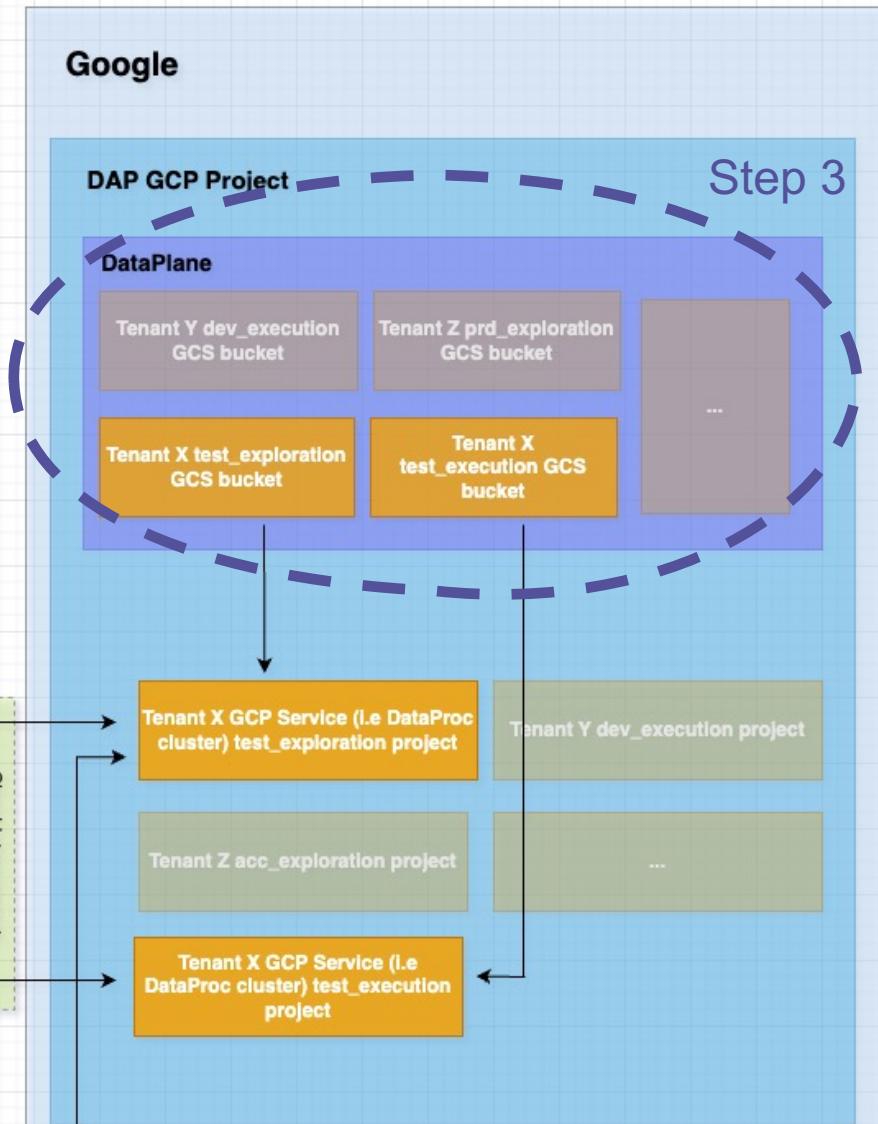
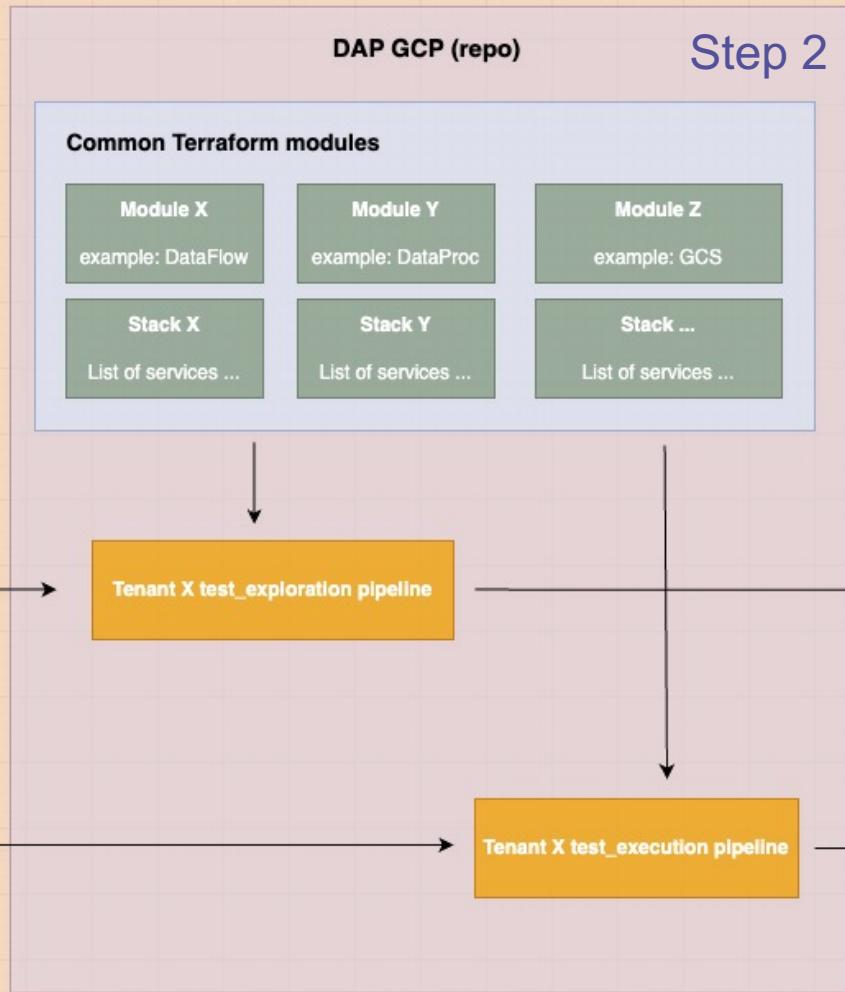
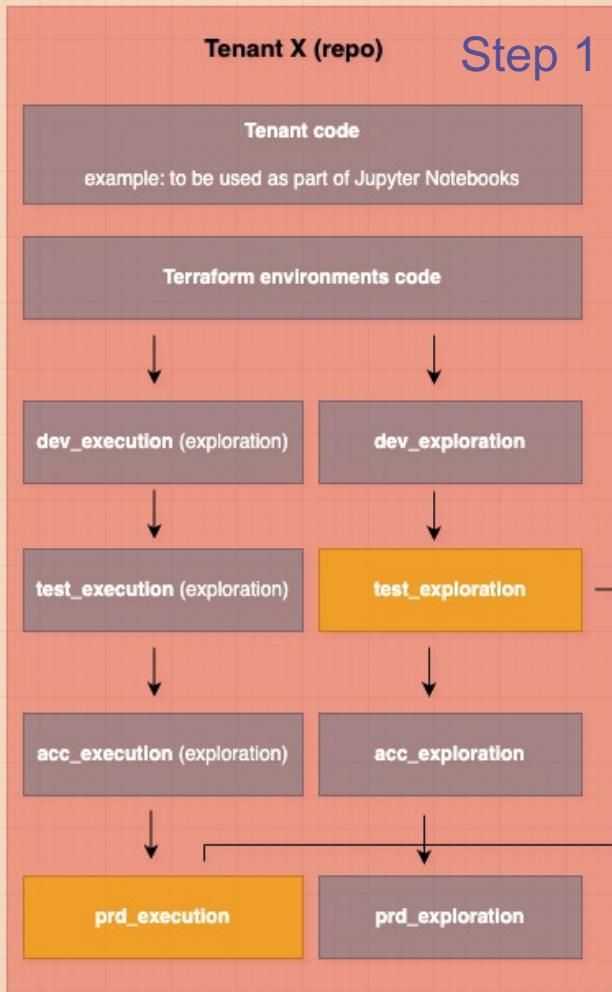


DAP GCP Architecture (simplified)

Showcasing how Tenant X code would be deployed to Tenant X GCP service (i.e DataProc cluster) on test_exploration project in Google Cloud

ING employee i.e: data scientist
acces to i.e Jupyter Notebooks **only**
accessible on **exploration** envs

ING

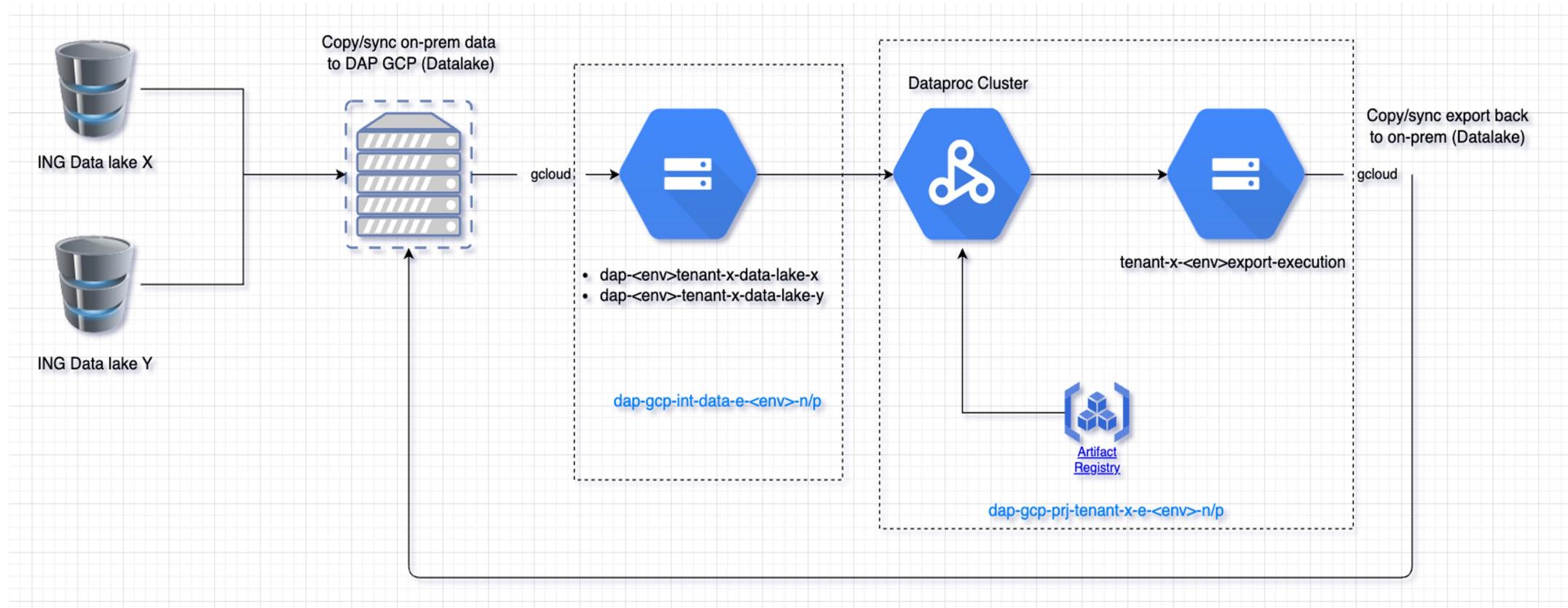


DAP GCP Architecture (simplified)

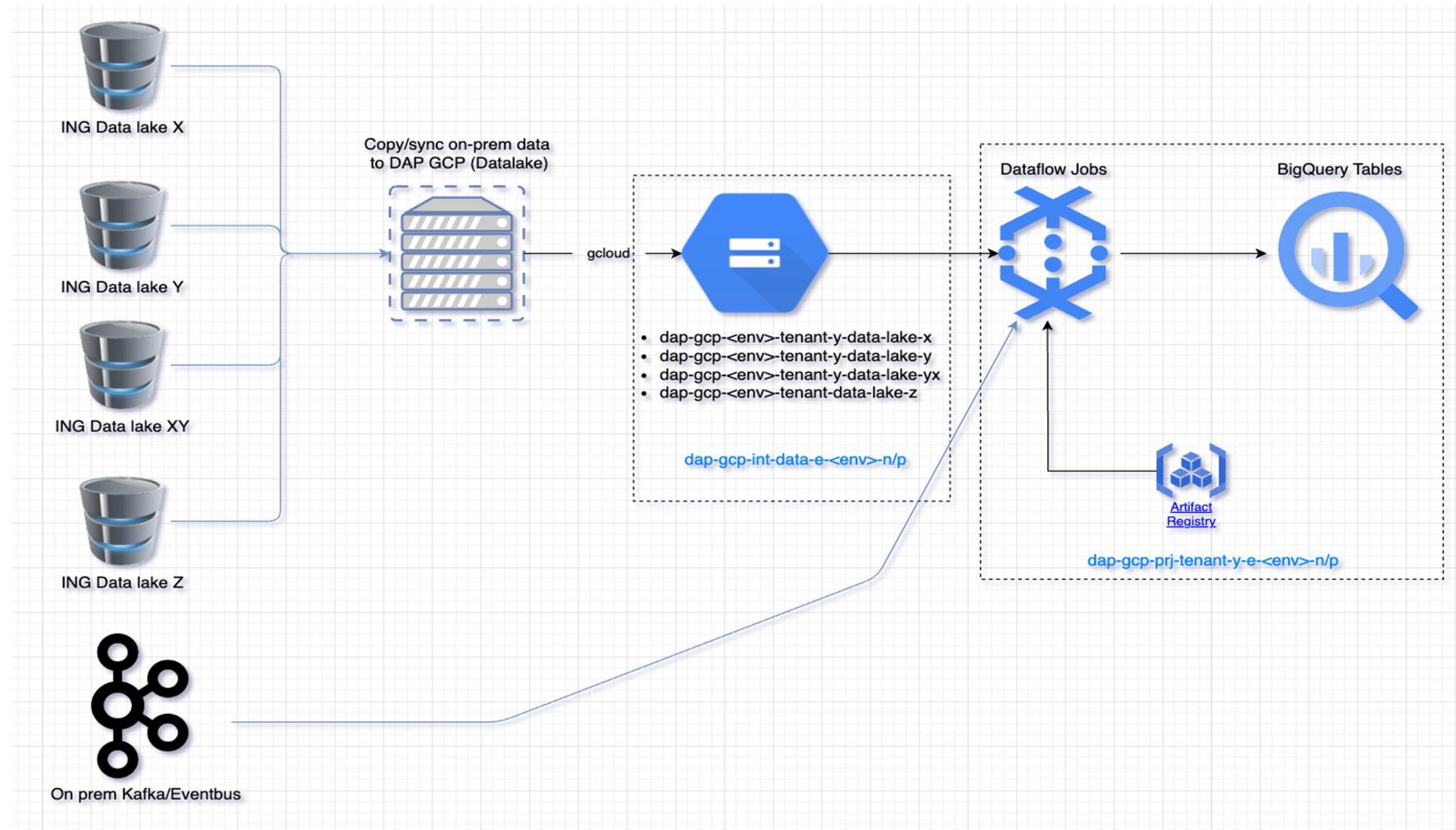
Showcasing how Tenant X code would be deployed to Tenant X GCP service (i.e DataProc cluster) on test_exploration project in Google Cloud

ING employee i.e: data scientist
acces to i.e Jupyter Notebooks **only**
accessible on **exploration** envs

DAP CGP Dataplane: Tenant X data flow



DAP CGP Dataplane: Tenant Y data flow



A photograph of a swimming pool area that has suffered significant damage from a disaster. Fallen trees and branches are scattered across the ground and partially submerged in the pool. A large tree trunk lies near the edge. The surrounding landscape is lush, with many tropical plants like palm trees and banana plants. In the background, a white fence and a building are visible through the debris.

Disaster ready data layer

dev.azure.com/ / Ing... / Pipelines / ... / dap-gcp-...-environments / exploration / GCP-Engineering test pipe... / P16002-dap-gcp-...-expl...

-dap-gcp-...-exploration-test-devops

Runs Branches Analytics

Description Stages

- #20230927.1 • Merged PR 802881: remove unused SVC
 - Manually triggered for main ↗ 71131bb2
- #20230914.1 • Merged PR 770649: tag bump
 - Manually triggered for main ↗ b3312e7f
- #20230801.1 • Merged PR 746528: add triggered pipeline for prod exploitation
 - Manually triggered for main ↗ ba463fb9
- #20230728.2 • Merge branch 'prod-exploitation-trigger' of https://dev.azure.com/...
 - Manually triggered for prod-exploitation-trigger ↗ 4281b845
- #20230728.1 • Merged PR 746528: add triggered pipeline for prod exploitation
 - Manually triggered for main ↗ ba463fb9
- #20230710.1 • check python packages as dataproc parameters
 - Manually triggered for python-lib-as-parameter ↗ fc037f8d
- #20230630.4 • change cluster_zone variable in stack main
 - Manually triggered for auto-zone ↗ cf085cc1
- #20230630.3 • Merged PR 719140: Tag bump
 - Manually triggered for main ↗ e7045948
- #20230630.2 • change cluster_zone variable in stack main
 - Manually triggered for auto-zone ↗ cf085cc1
- #20230630.1 • change cluster_zone variable in stack main
 - Manually triggered for auto-zone ↗ cf085cc1
- #20230629.3 • comment out randomized_zone from locals
 - Manually triggered for auto-zone ↗ 9735ae76
- #20230629.2 • comment out zone from exploitation stack for terraform validation purposes:
 - Manually triggered for auto-zone ↗ b30c2889
- #20230629.1 • comment out zone, zone randomizer from exploration stack
 - Manually triggered for auto-zone ↗ c3a1f84c
- #20230628.1 • Merged PR 703700: Trigger with credentials from vaults for testing
 - Manually triggered for main ↗ d8961654
- #20230627.2 • Merged PR 703700: Trigger with credentials from vaults for testing
 - Manually triggered for main ↗ d8961654
- #20230627.1 • Merged PR 703700: Trigger with credentials from vaults for testing
 - Manually triggered for main ↗ d8961654

Run pipeline

Select parameters below and manually run the pipeline

Branch/tag

main

Select the branch, commit, or tag

Idle delete TTL

3600s

GCP region

europe-west4

Environment name

test_exploration

Enable Checkov scan

true

false

Repository to take tfvars from

Default

Preprod2-dap-gcp-...

Folder to copy codebase from

Advanced options

Variables

This pipeline has no defined variables

Stages to run

Run as configured

Resources

Use latest version of all resources

Enable system diagnostics

Cancel Run

Jun 27



[redacted]

> __parameters__

> [redacted]

> [redacted]

> gcp

> [redacted]

> [redacted]

> pipelines

> prep

> resources

> scripts

> setup

└ terraform

└ environments

> acc_exploitation

> acc_exploration

> dev_exploitation

> prod_exploitation

> prod_exploration

> sandbox

> test_exploitation

└ test_exploration

└ exploration-stack

terraform.tfvars

README.md

[redacted] / main / [redacted] / environments / test_exploration / exploration-stack / terraform.tfvars

terraform.tfvars

Edit

Contents History Compare Blame ↗

```
1 ### Environment variables
2
3 idle_delete_ttl      = "3600s"
4 # cluster_life_duration = "36h"
5 optional_components   = ["JUPYTER"]
6 dataproc_image_version = "2.0-debian10"
7 pip_packages          = "numpy==1.22.2 scipy==1.8.0 pandas==1.4.0 matplotlib==3.5.1 nbconvert==6.5.3" # pandas-gbq==0.13.0
8
9 # You can add or modify the variables below. Uncomment and change.
10
11 master_machine_type    = "n2-highmem-96" # standard: 4 GB per CPU, highmem 8 GB per CPU
12 num_master_instances     = 1
13 worker_machine_type     = "n2-standard-64"
14 num_worker_instances     = 2 # Original: 2
15 num_secondary_worker_instances = 0
16
17 num_local_ssds          = 1 # number of primary disks across all machines
18
19 # boot_disk_size_gb      = 100
20 # boot_disk_type          = "pd-ssd"
21
22 master_boot_disk_size_gb = 200
23 worker_boot_disk_size_gb = 200
24 secondary_worker_boot_disk_size_gb = 100
25
26 master_boot_disk_type    = "pd-ssd"
27 worker_boot_disk_type    = "pd-ssd"
28 secondary_worker_boot_disk_type = "pd-ssd"
29
30 # additional drives
31 num_master_local_ssds    = 0 # override num_local_ssds for masters
32 num_worker_local_ssds     = 0 # override num_local_ssds for workers because n2-standard-16 needs even number, 4 HDs for n2-standard-32
33 num_secondary_worker_local_ssds = 0 # override num_local_ssds for secondary workers because n2-standard-16 needs 0,2.. and larger need even higher number of SSDs (0,4..)
34
35 # used for autoscaling
36 primary_worker_min_instances = 2 # Original: 2
37 primary_worker_max_instances = 200
38 secondary_worker_min_instances = 0
39 secondary_worker_max_instances = 0
40
41 cluster_properties = {
42   "dataproc:dataproc.logging.stackdriver.job.yarn.container.enable" = "true",
43   "dataproc:dataproc.logging.stackdriver.job.driver.enable"       = "true"
44 }
```



Tags

New tag

Search tag name

Tag		Commit	Tagger	Creation Date
⌚ acc-exploration-SZ-09-03	acc exploration	62c5a929	[redacted]	Mar 9
⌚ exploitation-acc-run-010223	Exploitation acc run	6f9291e0	[redacted]	Jan 2
⌚ exploitation-acc-run-0103-mars	mars test run	16d800f1	[redacted]	Jan 11
⌚ exploitation-acc-run-010323	Exploitation acc run	bf2691c0	[redacted]	Jan 9
⌚ exploitation-acc-run-0104	exploitation acc run	fbdb3f0cf	[redacted]	Jan 4
⌚ exploitation-acc-run-0104-2	Exploitation acc run	\188873b	[redacted]	Jan 4
⌚ exploitation-acc-run-0109	Exploitation acc run	fa64ecc	[redacted]	Jan 9
⌚ exploitation-acc-run-0111-mars-test	Mars test	20bca41	[redacted]	Jan 11
⌚ exploitation-acc-run-0112	exploitation acc run	04060e0	[redacted]	Jan 12
⌚ exploitation-acc-run-0112-2	exploitation acc run	431c621	[redacted]	Jan 12
⌚ exploitation-acc-run-0112-3	exploitation-acc-run-0112-3	736b3a5	[redacted]	Jan 12
⌚ exploitation-acc-run-0112-4	exploitation-acc-run-0112	c48ac87	[redacted]	Jan 12
⌚ exploitation-acc-run-0113	exploitation-acc-run-0113	c48ac87	[redacted]	Jan 13
⌚ exploitation-acc-run-0118	exploitation acc run	69a63cef	[redacted]	Jan 18
⌚ exploitation-acc-run-0118-03	Exploitation acc run	789b60eb	[redacted]	Jan 18
⌚ exploitation-acc-run-0118-04	exploitation acc run	6944ab44	[redacted]	Jan 18
⌚ exploitation-acc-run-0118-2	exploitation-acc-run-0118-2	b2af7b34	[redacted]	Jan 18

Create a tag

Name *

exploration-test-run-adnan-demo

Based on

⌚ main

Description *

Creating tag for demo purposes ...

Cancel

Create

Merged PR 782988: mvp13 prod: Update with new field os_ccy_orig
dap-gcp-exploration-test

Run new

This run is being retained as one of recent runs by main (Branch).

View retention leases

Summary Tests Environments Associated pipelines Artifactory Scans

Triggered by Adnan Hodzic

View 16 changes

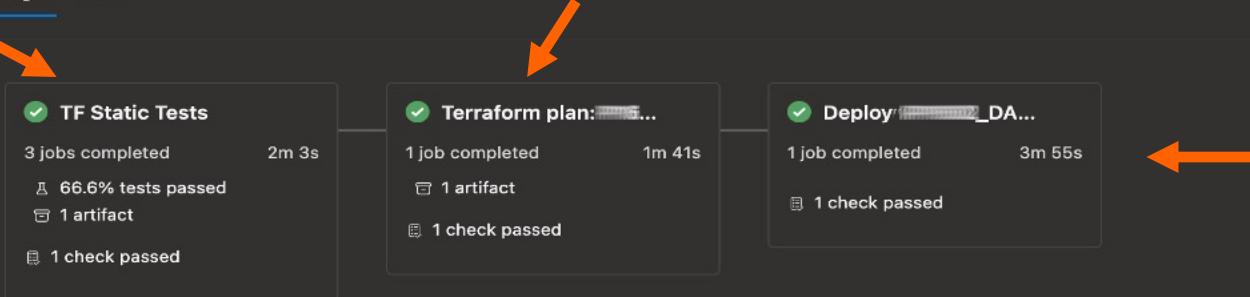
Repositories 3
P16002-dap-gcp... +2
See Sources card for details

Time started and elapsed
Today at 10:46 AM
8m 55s

Related
5 work items
2 published; 2 consumed

Tests and coverage
66.6% passed
Setup code coverage

Stages Jobs



Sources

Repository	Branch / tag	Version	Related
P16002-dap-gcp...	main	71131bb2	None
P16002-dap-tf-common	v0.15.11	879b4e81	
P16002	exploration-test-run-adnan-demo	cc767420	

Dataproc

Clusters

+ CREATE CLUSTER

REFRESH

▶ START

■ STOP

DELETE

REGIONS ▾

+ 5 RECOMMENDED ALERTS

SHOW INFO PANEL

Jobs on Clusters ▾

Filter Search clusters, press Enter



<input type="checkbox"/>	Name	Status	Region	Zone	Total worker nodes	Scheduled deletion	Cloud Storage staging bucket	Created
<input type="checkbox"/>	[REDACTED]-tst-exploration-adnan-demo-cmqb	Running	europe-west4	europe-west4-c	2	On	[REDACTED]-tst-staging-exploration	10:53:43 AM
<input type="checkbox"/>	[REDACTED]-tst-exploration-dg031029t7	Running	europe-west4	europe-west4-c	2	On	[REDACTED]-tst-staging-exploration	8:57:07 AM

Autoscaling policies

Serverless

Batches

Interactive

Metastore Services

Metastore

Federation

Utilities

Component exchange

Workbench

Dataproc Cluster details

Consider using Auto Zone rather than selecting a zone manually. See <https://cloud.google.com/dataproc/docs/concepts/configuring-clusters/auto-zone>

Name: ec-tst-exploration-adnan-demo-cmqb
Cluster UUID: 330335fa-a8c3-4466-a62e-3ee8fab7aa37
Type: Dataproc Cluster
Status: Running

MONITORING **JOB** **VM INSTANCES** **CONFIGURATION** **WEB INTERFACES**

WEB INTERFACES

SSH tunnel
Create an SSH tunnel to connect to a web interface

Component gateway
Provides access to the web interfaces of default and selected optional components on the cluster. [Learn more](#)

[YARN ResourceManager](#) [MapReduce Job History](#) [Spark History Server](#) [HDFS NameNode](#) [YARN Application Timeline](#) [HiveServer2 \(ec-tst-exploration-adnan-demo-cmqb-m\)](#) [Tez](#) [Jupyter](#) [JupyterLab](#)

EQUIVALENT REST

Filter files by name 

Name ▾ Last Modified

GCS a year ago

Local Disk 2 minutes ago

Launcher

Notebook



Python 3



PySpark



R



spylon-kernel

Console



Python 3



PySpark



R



spylon-kernel

Other



Terminal



Text File



Markdown File

Show
Contextual Help

DAP GCP teams: DAP GCP engineering

Mission: do everything it takes to bring tenants & their projects to Production on GCP

Creating necessary pipelines and defining deployment strategies, adhering to separation of responsibilities principle

Help with onboarding, design & architecture

DAP GCP teams: DAP GCP engineering

Mission: do everything it takes to bring tenants & their projects to Production on GCP

Creating necessary pipelines and defining deployment strategies, adhering to separation of responsibilities principle

Help with onboarding, design & architecture

Creation of GCP projects, networks, firewalls

DAP GCP teams: DAP GCP engineering

Mission: do everything it takes to bring tenants & their projects to Production on GCP

Creating necessary pipelines and defining deployment strategies, adhering to separation of responsibilities principle

Help with onboarding, design & architecture

Creation of GCP projects, networks, firewalls

Code (common terraform modules) & pipelines needed to deploy services in GCP with Terraform (IaC)

DAP GCP teams: DAP GCP engineering

Mission: do everything it takes to bring tenants & their projects to Production on GCP

Creating necessary pipelines and defining deployment strategies, adhering to separation of responsibilities principle

Help with onboarding, design & architecture

Creation of GCP projects, networks, firewalls

Code (common terraform modules) & pipelines needed to deploy services in GCP with Terraform (IaC)

Adhering to ING's rigorous risk, security and compliance principles

Project Name	Project ID	Service Account or Group	Keys	Owner	Minimum Roles	Additional Roles
DAP-PRJ-tenant-t-E-PROD-P	dap-prj-tenant-t-e-prod-p	monitor-tenant@dap-int-cont-e-prod-p.iam.gserviceaccount.com	Jun 28, 2023 - e43e6d30f3411bd34d41b09e30c87b5fb1593f68 Total number of keys: 1	Data Analytics Platform	roles/logging.viewer roles/logging.configWriter roles/monitoring.admin	roles/dataflow.worker roles/compute.networkUser roles/storage.legacyBucketReader roles/artifactregistry.reader roles/bigquery.jobUser roles/pubsub.viewer roles/dataflow.admin
DAP-PRJ-tenant-t-E-PROD-P	dap-prj-tenant-t-e-prod-p	svc-c-062834-10-9-234-0@dap-prj-tenant-t-e-prod-p.iam.gserviceaccount.com	N/A	Public Cloud Foundation		roles/dataflow.worker roles/compute.networkUser roles/storage.legacyBucketReader roles/artifactregistry.reader roles/bigquery.jobUser roles/pubsub.viewer roles/dataflow.admin
DAP-PRJ-tenant-t-E-PROD-P	dap-prj-tenant-t-e-prod-p	bq-restore-sa@dap-prj-tenant-t-e-prod-p.iam.gserviceaccount.com	May 08, 2024 - bbae0a0eecae0c268255224cf1f1c9bd84e760af Total number of keys: 1	Data Analytics Platform	roles/bigquery.admin	
DAP-PRJ-tenant-t-E-PROD-P	dap-prj-tenant-t-e-prod-p	elementary-report-sa@dap-prj-tenant-t-e-prod-p.iam.gserviceaccount.com	N/A	Error: Unable to determine!	Error: Unable to determine!	Error: Unable to determine!
DAP-PRJ-tenant-t-E-PROD-P	dap-prj-tenant-t-e-prod-p	gcs-restore-sa@dap-prj-tenant-t-e-prod-p.iam.gserviceaccount.com	Total number of keys: 0	Data Analytics Platform	roles/storage.admin	
DAP-PRJ-tenant-t-E-PROD-P	dap-prj-tenant-t-e-prod-p	tf-deploy@dap-prj-tenant-t-e-prod-p.iam.gserviceaccount.com	Jul 06, 2022 - e5e71149067085f3fe733bec52227be23627fc2f Feb 15, 2024 - c60cf162670dceea70d84ae75d0e1bb183e9154 1 Total number of keys: 2	Data Analytics Platform	roles/essentialcontacts.admin roles/editor roles/logging.admin roles/cloudkms.admin roles/storage.admin roles/monitoring.admin	roles/cloudasset.viewer roles/iam.serviceAccountAdmin roles/compute.admin roles/serviceusage.serviceUsageAdmin roles/bigquery.admin roles/run.serviceAgent roles/pubsub.admin roles/workflows.viewer roles/secretmanager.admin roles/iam.serviceAccountUser roles/run.admin roles/viewer roles/artifactregistry.admin roles/dataflow.admin
DAP-PRJ-tenant-t-E-PROD-P	dap-prj-tenant-t-e-prod-p	bigquery-viewer@dap-prj-tenant-t-e-prod-p.iam.gserviceaccount.com	Jul 27, 2023 - 20fd45d85ad353a64fce1e79536b1d8fe5eba77 Aug 08, 2023 - 8811226c299f296beb33100a1338252b28e0ddb 3 Jan 19, 2024 - ee93b13a6209de3f56f66b9e57f3eedad69c56b32 Total number of keys: 3	Tenant	roles/bigquery.jobUser	
DAP-PRJ-tenant-t-E-PROD-P	dap-prj-tenant-t-e-prod-p	DSRLN2-Po@ing.com	N/A	Data Analytics Platform	roles/essentialcontacts.viewer roles/logging.viewer roles/iam.serviceAccountKeyAdmin roles/logging.viewAccessor	
DAP-PRJ-tenant-t-E-PROD-P	dap-prj-tenant-t-e-prod-p	orchestration-manager@dap-prj-tenant-t-e-prod-p.iam.gserviceaccount.com	N/A	Tenant	roles/composer.worker	

Project Name	Project ID	Service Account or Group	Keys	Owner	Minimum Roles	Additional Roles
DAP-PRJ-tenant-t-E-PROD-P	dap-prj-tenant-t-e-prod-p	monitor-tenant@dap-int-cont-e-prod-p.iam.gserviceaccount.com	Jun 28, 2023 - e43e6d30f3411bd34d41b09e30c87b5fb1593f68 Total number of keys: 1	Data Analytics Platform	roles/logging.viewer roles/logging.configWriter roles/monitoring.admin	roles/dataflow.worker roles/compute.networkUser roles/storage.legacyBucketReader roles/artifactregistry.reader roles/bigquery.jobUser roles/pubsub.viewer roles/dataflow.admin
DAP-PRJ-tenant-t-E-PROD-P	dap-prj-tenant-t-e-prod-p	svc-c-062834-10-9-234-0@dap-prj-tenant-t-e-prod-p.iam.gserviceaccount.com	N/A	Public Cloud Foundation		
DAP-PRJ-tenant-t-E-PROD-P	dap-prj-tenant-t-e-prod-p	bq-restore-sa@dap-prj-tenant-t-e-prod-p.iam.gserviceaccount.com	May 08, 2024 - bbae0a0eecae0c268255224cf1f1c9bd84e760af Total number of keys: 1	Data Analytics Platform	roles/bigquery.admin	
DAP-PRJ-tenant-t-E-PROD-P	dap-prj-tenant-t-e-prod-p	elementary-report-sa@dap-prj-tenant-t-e-prod-p.iam.gserviceaccount.com	N/A	Error: Unable to determine!	Error: Unable to determine!	Error: Unable to determine!
DAP-PRJ-tenant-t-E-PROD-P	dap-prj-tenant-t-e-prod-p	gcs-restore-sa@dap-prj-tenant-t-e-prod-p.iam.gserviceaccount.com	Total number of keys: 0	Data Analytics Platform	roles/storage.admin	
DAP-PRJ-tenant-t-E-PROD-P	dap-prj-tenant-t-e-prod-p	tf-deploy@dap-prj-tenant-t-e-prod-p.iam.gserviceaccount.com	Jul 06, 2022 - e5e71149067085f3fe733bec52227be23627fc2f Feb 15, 2024 - c60cf162670dceea70d84ae75d0e1bb183e9154 1 Total number of keys: 2	Data Analytics Platform	roles/essentialcontacts.admin roles/editor roles/logging.admin roles/cloudkms.admin roles/storage.admin roles/monitoring.admin	roles/cloudasset.viewer roles/iam.serviceAccountAdmin roles/compute.admin roles/serviceusage.serviceUsageAdmin roles/bigquery.admin roles/run.serviceAgent roles/pubsub.admin roles/workflows.viewer roles/secretmanager.admin roles/iam.serviceAccountUser roles/run.admin roles/viewer roles/artifactregistry.admin roles/dataflow.admin
DAP-PRJ-tenant-t-E-PROD-P	dap-prj-tenant-t-e-prod-p	bigquery-viewer@dap-prj-tenant-t-e-prod-p.iam.gserviceaccount.com	Jul 27, 2023 - 20fd45d85ad353a64fce1e79536b1d8fe5eba77 Aug 08, 2023 - 8811226c299f296beb33100a1338252b28e0ddb 3 Jan 19, 2024 - ee93b13a6209de3f56f66b9e57f3eedad69c56b32 Total number of keys: 3	Tenant	roles/bigquery.jobUser	
DAP-PRJ-tenant-t-E-PROD-P	dap-prj-tenant-t-e-prod-p	DSRLN2-Po@ing.com	N/A	Data Analytics Platform	roles/essentialcontacts.viewer roles/logging.viewer roles/iam.serviceAccountKeyAdmin roles/logging.viewAccessor	
DAP-PRJ-tenant-t-E-PROD-P	dap-prj-tenant-t-e-prod-p	orchestration-manager@dap-prj-tenant-t-e-prod-p.iam.gserviceaccount.com	N/A	Tenant	roles/composer.worker	

Current Roles	Non Compliant Roles	Interaction	Secrets in vault?	Safe name	Comment
roles/logging.viewer roles/logging.configWriter roles/monitoring.admin		Non-interactive	TRUE	Azure KeyVault: Data Analytics PlatformProd	I Service account used for scoped monitoring across all projects.
roles/artifactregistry.reader roles/bigquery.jobUser roles/dataflow.admin roles/compute.networkUser roles/storage.legacyBucketReader roles/pubsub.viewer roles/dataflow.worker		Non-interactive	FALSE	Not applicable	Service Account used by GCP to enable the services to access other GCP resources and work as intended.
roles/bigquery.admin	Error: Unable to determine!	Non-interactive	TRUE	Azure KeyVault: Data Analytics PlatformProd	I Service Account for restore BigQuery Data in case of accidentally deletion.
roles/storage.admin		Non-interactive	TRUE	Azure KeyVault: Data Analytics PlatformProd	I Service Account for restore cloud storage data.
roles/compute.admin roles/serviceusage.serviceUsageAdmin roles/logging.admin roles/essentialcontacts.admin roles/iam.serviceAccountUser roles/monitoring.admin roles/cloudasset.viewer roles/iam.serviceAccountAdmin roles/editor roles/bigquery.admin roles/secretmanager.admin roles/cloudkms.admin roles/storage.admin roles/run.admin roles/viewer roles/dataflow.admin roles/run.serviceAgent roles/pubsub.admin roles/workflows.viewer roles/artifactregistry.admin		Non-interactive	TRUE	Azure KeyVault: Data Analytics PlatformProd	I Service Account used in CI/CD pipeline to deploy changes.
roles/bigquery.jobUser		Interactive	TRUE	CyberArk: 14267U_Data Analytics Platform-GC	Service Account used for data access in Data Analytics Platform GCP Tenant project from BQ.
roles/essentialcontacts.viewer roles/iam.serviceAccountKeyAdmin roles/logging.viewAccessor roles/logging.viewer		Non-interactive	FALSE	Not applicable	ACE group - Data Analytics Platform-GCP Engineering Team T15965. GCP group: DSRLN2-Po@ing.com
		Non-interactive	FALSE	Not applicable	Service account to manage the orchestration via cloud composer.

DAP GCP teams: DAP GCP engineering

Mission: do everything it takes to bring tenants & their projects to Production on GCP

Creating necessary pipelines and defining deployment strategies, adhering to separation of responsibilities principle

Help with onboarding, design & architecture

Creation of GCP projects, networks, firewalls

Code (common terraform modules) & pipelines needed to deploy services in GCP with Terraform (IaC)

Adhering to ING's rigorous risk, security and compliance principles

Monitoring & alerting

DAP GCP teams: DAP GCP engineering

Mission: do everything it takes to bring tenants & their projects to Production on GCP

Creating necessary pipelines and defining deployment strategies, adhering to separation of responsibilities principle

Help with onboarding, design & architecture

Creation of GCP projects, networks, firewalls

Code (common terraform modules) & pipelines needed to deploy services in GCP with Terraform (IaC)

Adhering to ING's rigorous risk, security and compliance principles

Monitoring & alerting

Support

A wide-angle photograph of a car assembly plant. Numerous car frames are lined up on a conveyor belt system, moving through various stages of production. The frames are light-colored and show the internal skeletal structure of the vehicles. The background is filled with industrial equipment, overhead lighting, and structural elements of the factory.

DAP GCP teams

DAP GCP Platform

DAP GCP teams: DAP GCP Platform

Resolving user & platform scalability problem

DAP GCP teams: DAP GCP Platform

Resolving user & platform scalability problem

Create a Platform (self service portal)

Run pipeline

Select parameters below and manually run the pipeline

Branch/tag

main



Select the branch, commit, or tag

Environment

- test
- prod

Project name

Required

test-adnan

Project type

- exploration
- execution_non_prod
- execution_prod

members

- XYZ-Ops@ing.com

Advanced options

Variables

This pipeline has no defined variables



Stages to run

Run as configured



Resources

Use latest version of all resources



Cancel



Run



 Welcome

You're working in PROJECT-2-DEMO-SELF-SERVICE-P

Project number: 12345678901234567890



Project ID: 12345678901234567890

[Dashboard](#) [Recommendations](#)[+ Create a VM](#)[+ Run a query in BigQuery](#)[+ Create a GKE cluster](#)[+ Create a storage bucket](#)Try our most advanced
model: Gemini 1.5 Pro[Try Gemini](#)

Quick access

 [API APIs & Services](#) [IAM & Admin](#) [Billing](#) [Compute Engine](#) [Cloud Storage](#) [BigQuery](#) [VPC network](#) [Kubernetes Engine](#) [VIEW ALL PRODUCTS](#)



Untitled.ipynb

```
contents": {
    "role": "USER",
    "parts": { "text": "What do you know about ING?" }
}

MODEL_ID="gemini-1.0-pro"
response = requests.post(f"https://{{API_ENDPOINT}}/v1/projects/{{project}}/locations/{{LOCATION}}/publishers/google/models/{{MODEL_ID}}:streamGenerateContent",
                        headers=headers,
                        json= data)

for content in response.json():
    print(content['candidates'][0]["content"]["parts"][0]["text"])

## ING
: A Global Financial Institution

**ING Group** is a Dutch multinational banking and financial services corporation headquartered in Amsterdam. The company was formed in 1991 through a merger of two Dutch banks, NMB Postbank and ING Bank. In the Netherlands, ING operates under the brand name ING Nederland, while internationally it uses the brand name ING Group.

**Here's what I know about ING:**

* **Services:** ING provides a wide range of banking and financial services, including:
    * **Retail banking:** Personal and business current and savings accounts, credit cards, debit cards, mortgages, loans, and insurance.
    * **Wholesale banking:** Corporate loans, international payments, trade finance, and treasury management.
    * **Investment banking:** Investment banking, capital markets, and securities trading.
*
    * **International presence:** ING operates in over 40 countries and territories, with a significant presence in Europe, Asia, and the Americas.
* **Customer base:** ING has over 39 million customers worldwide.
* **Financial performance:** ING is a publicly traded company listed on the Euronext Amsterdam and New York Stock Exchange. In 2022, the company generated
of €
4.04 billion.
* **Sustainability:** ING is committed to sustainability and has adopted a number of initiatives to reduce its environmental footprint
ty.
* **Digitalization:** ING is a leader in digital banking and has invested heavily in online and mobile banking services.

**Additional resources:**
```

Would you like to receive news?
Please read the privacy

Open priva

DAP GCP teams: DAP GCP Platform

Resolving user & platform scalability problem

Create a Platform (self service portal)

Final product will have lot more features; currently oriented towards Vertex AI & genAI (Gemini) use cases



DAP GCP teams

DAP GCP Looker

DAP GCP teams: DAP GCP Looker

Integrate GCP Looker capability for tenants

Build a managed platform to help users with their BI needs, allowing to focus on their products



From an idea to Production in just over 2 years

Thank you

Questions & comments?



<https://linkedin.com/in/AdnanHodzic>

Adnan.Hodzic@ing.com





do your thing