

 Generalitat de Catalunya Centre de Telecomunicacions i Tecnologies de la Informació	PTD		N. revisió doc.: 1.0
	Gestió de l'entorn de manteniment		
	N. versió aplicació: 1.0	Build: 1.0	Pàg. 1 / 11

Revisió	Redactat per	Revisat per	Aprobat per	Data aprovació	Data publicació
1.0	Minsait	Minsait			

Registre de canvis del document

Revisió	Apartat	Data Modificació	Motiu del canvi

RESPONSABLE DEL DOCUMENT:

Í N D E X

1. INTRODUCCIÓ	2
1.1 Objectiu	2
1.2 Context de l'arquitectura	2
1.3 Principis de disseny	3
2. ESTRUCTURA DE L'ENTORN DE MANTENIMENT	3
2.1 Flux de gestió després d'una incidència	3
2.1.1 Job de clonatge	4
2.1.2 Job de neteja	5
2.1.3 Identitat per a l'execució i seguretat	6
2.2 Flux operatiu després d'una incidència	6
2.2.1 Configurar projecte Bundle pel manteniment	7
2.2.2 Crear la branca manteniment com a còpia de master	7
2.2.3 Editar el fitxer databricks.yml y _init_.py	7
2.2.4 Desplegament en manteniment.....	9
2.2.5 Pujar la correcció a producció (máster).....	11

 Generalitat de Catalunya Centre de Telecomunicacions i Tecnologies de la Informació	PTD	N. revisió doc.: 1.0
	Gestió de l'entorn de manteniment	
	N. versió aplicació: 1.0	Build: 1.0 Pàg. 2 / 11

1. INTRODUCCIÓ

La gestió d'incidències relacionades amb dades forma part de l'operativa habitual de qualsevol plataforma analítica.

Quan es detecten inconsistències, comportaments inesperats o dubtes sobre la integritat de la informació, és necessari disposar de mecanismes que permetin analitzar la situació amb detall, de manera controlada i sense posar en risc l'estabilitat de l'entorn principal.

Amb aquest objectiu, s'ha definit un entorn de manteniment dins de la PTD. Aquest entorn proporciona un espai específic per a dur terme tasques de diagnòstic i anàlisis sobre dades productives de manera aïllada, permetent als equips tècnics treballar amb més seguretat i traçabilitat.

El present document descriu l'ús d'aquest entorn, quin és el seu paper dins de l'arquitectura global i quins processos automatitzats el gestionen.

1.1 Objectiu

L'entorn de manteniment s'ha dissenyat com un espai controlat dins de la PTD amb el propòsit de permetre la investigació, anàlisi i resolució d'incidències que afecten dades en producció, sense intervenir directament sobre els actius productius.

L'objectiu principal no és dur a terme desenvolupaments ni proves funcionals, sinó donar suport tècnic a producció. La seva funció és permetre que els equips puguin:

- Analitzar l'estat de les dades en el moment de la incidència.
- Validar hipòtesis sobre possibles causes.
- Executar consultes i comprovacions sense afectar els sistemes en operació.

D'aquesta manera, es redueix el risc operatiu, s'evita la modificació accidental de dades productives i es millora la traçabilitat de les accions realitzades durant el procés de diagnòstic.

1.2 Context de l'arquitectura

La PTD disposa actualment d'una arquitectura distribuïda per dominis de dades on:

- Cada domini de dades té el seu propi workspace de Databricks.
- Cada workspace conté un o més catàlegs de Unity Catalog, que poden ser o no accessibles des d'altres workspaces.
- Cada catàleg està associat a una storage location al seu propi Storage Account.

Per tant, l'entorn de manteniment s'implementa com un workspace addicional dins de la subscripció de producció, que no conté dades pròpies persistents, sinó còpies controlades de dades productives generades sota demanda.

El workspace de Databricks que correspon a aquest entorn de manteniment és l'anomenat **pgc-tra-dbs-man-pro-001** (<https://adb-727588810011327.7.azuredatabricks.net/>), dins de la subscripció de producció (PTD Pro).

Per accedir a aquest workspace, serà necessari estar connectat a la VPN del CTTI, utilitzant Cisco AnyConnect.

 Generalitat de Catalunya Centre de Telecomunicacions i Tecnologies de la Informació	PTD		N. revisió doc.: 1.0
	Gestió de l'entorn de manteniment		
	N. versió aplicació: 1.0	Build: 1.0	Pàg. 3 / 11

1.3 Principis de disseny

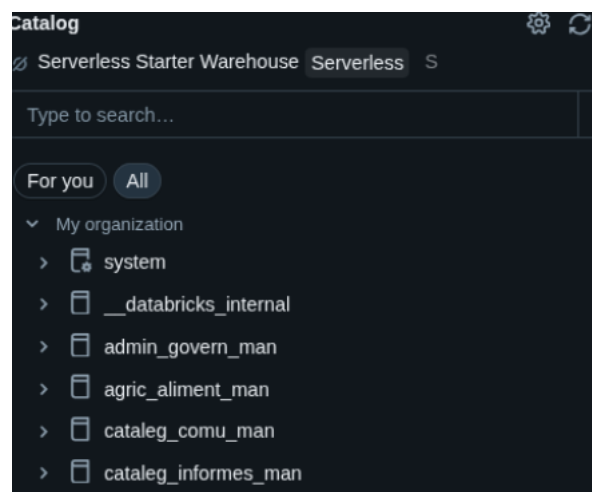
L'entorn es regeix pels següents principis:

1. **Aïllament operatiu:** Les accions d'anàlisi no es realitzen mai directament sobre taules de producció.
2. **No duplicació innecessària de dades:** S'utilitza la funcionalitat Shallow Clone de Delta Lake, que permet replicar l'estructura i els metadades d'una taula sense copiar físicament les dades.
3. **Reproduïbilitat de l'estat de la dada:** Els clons es poden generar sobre una versió concreta de la taula, fet que permet analitzar exactament l'estat en el moment de la incidència.
4. **Governança i seguretat:** Tots els processos s'executen sota un Service Principal tècnic, evitant dependències d'usuaris individuals.
5. **Entorn efímer:** Un cop resolta la incidència, el catàleg de manteniment es neteja completament.

2. ESTRUCTURA DE L'ENTORN DE MANTENIMENT

Per a cada catàleg productiu existeix un catàleg equivalent en manteniment, identificat amb el sufix “_man”. Per exemple, el catàleg “A” en producció té el seu equivalent “A_man” en manteniment.

Aquests catàlegs actuen com a contenidors de treball. No s'han d'utilitzar com a repositoris de dades permanents. El seu contingut depèn de les incidències actives i pot variar en el temps.



2.1 Flux de gestió després d'una incidència

En primer lloc, un cop detectada la incidència, s'identifiquen els catàlegs, esquemes i taules que caldrà replicar a l'entorn de manteniment.

A continuació, s'obre un tiquet a ACOPTD indicant els catàlegs, esquemes i taules que s'han de replicar a l'entorn de manteniment.

Permisos de seguretat a les entitats (catàlegs, esquemes i taules)

Posteriorment, l'equip de PTD genera l'entorn de treball en manteniment mitjançant l'execució del job corresponent.

 Generalitat de Catalunya Centre de Telecomunicacions i Tecnologies de la Informació	PTD	N. revisió doc.: 1.0
	Gestió de l'entorn de manteniment	
	N. versió aplicació: 1.0	Build: 1.0 Pàg. 4 / 11

Name 	Type
clean-up-catalog	 Job
shallow-clone-tables	 Job

Quan aquest job hagi finalitzat l'execució, es disposa d'una còpia de les taules indicades. No obstant això, aquest clonatge inclou únicament les dades i l'estructura de les taules i no aplica els permisos necessaris sobre les taules i els esquemes.

Per aquest motiu, un cop creat, l'equip de Seguretat s'ha d'encarregar d'aplicar els permisos necessaris per poder treballar amb aquestes taules i esquemes.

2.1.1 Job de clonatge

Aquest job, anomenat shallow-clone-tables, s'encarrega de generar les taules de treball necessàries al catàleg de manteniment.

La seva funció és crear els esquemes requerits i clonar les taules indicades des de producció mitjançant shallow clone. Aquestes còpies conserven l'estructura i els metadades de les taules originals.

El job rep com a paràmetre un bloc de configuració en format JSON que descriu els elements que s'han de clonar. En aquest bloc s'hi indiquen els catàlegs origen i destí, els esquemes implicats i les taules concretes sobre les quals es necessita treballar.

catalogs: Array de configuracions de migració de catàlegs

- source_catalog: Catàleg d'origen des d'on es copiaran les taules
- target_catalog: Catàleg de destinació on es copiaran les taules
- schemas: Objecte que mapeja esquemes a llistes de taules
- Clau (ej: "cat_one"): Nom de l'esquema d'origen
- Valor (ej: ["maintenance_tasks"]): Array de noms de taules a copiar des d'aquest esquema

```
{
  "catalogs": [
    {
      "source_catalog": "manteniment_catalog",
      "target_catalog": "manteniment_catalog_man",
      "schemas": {
        "cat_one": [ "maintenance_tasks" ],
        "cat_two": [ "reports", "report_runs" ]
      }
    },
    {
      "source_catalog": "manteniment_catalog",
      "target_catalog": "manteniment_catalog_man",
      "schemas": {
```

Gestió de l'entorn de manteniment_V1.1.docx
05/07/2026

23:40:44

```
"cat_three": [ "ingestion_status" ]
}
}
}
}
```

Eixample

```
{
  "catalogs": [
    {
      "source_catalog": "finops_pro",
      "target_catalog": "finops_man",
      "schemas": {
        "gld_finops": [ "azure_cost_management", "denodo_invoice" ],
        "slv_finops": [ "azure_cost_management" ]
      }
    }
  ]
}
```

schemas a clonar tablas a clonar

Settings

Select tasks to run
Click individual task to select or select multiple tasks using the hover task controls

shallow_clone_task

..._dbs/prod/files/src/shallow_clone_job

shallow_clone_cluster

Compute

Performance-optimized mode improves startup and execution speed. Disabling performance optimization gives you startup times similar to classic infrastructure and reduces your cost. Currently only applies to certain Serverless workloads. [Learn more](#)

☐ Performance optimized

Parameters

Override job parameters for this run. Any parameters that are not overridden here will have their default values. [Learn More](#)

Key	Value
finops_project	finops
payload	{ "catalogs": [{ " " }] }

2.1.2 Job de neteja

Aquest segon job, anomenat clean-up-catalog, té com a objectiu retornar el catàleg de manteniment a un estat buit un cop finalitzada la incidència.

Rep com a paràmetre el nom del catàleg de manteniment i elimina tots els objectes que conté, incloent-hi taules, vistes, volums i esquemes.

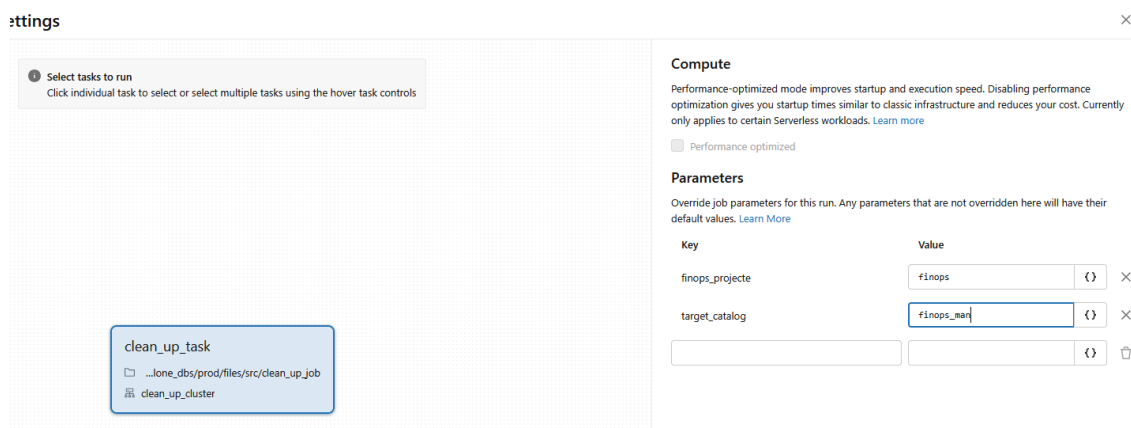
El procés es duu a terme de manera controlada i seguint l'ordre necessari per evitar dependències entre objectes. Com a mesura de seguretat, aquest job només es pot executar sobre catàlegs el nom dels quals acaba en “_man”, evitant qualsevol risc sobre catàlegs productius. Un cop executat, el catàleg queda preparat per a futures incidències.

Aquest document s'ha basat en la plantilla publicada al MQS Disseny Detallat v1.2

 Generalitat de Catalunya Centre de Telecomunicacions i Tecnologies de la Informació	PTD	N. revisió doc.: 1.0
	Gestió de l'entorn de manteniment	
	N. versió aplicació: 1.0	Build: 1.0 Pàg. 6 / 11

Paràmetres d'entrada:

Aquest *job*, igual que l'anterior, també s'executa sota demanda, i cal assegurar-se que, en fer-ho, s'ha indicat en el paràmetre **target_catalog** el nom del catàleg de manteniment sobre el qual es vol realitzar la neteja i el *finops-projecte*.



Settings

Select tasks to run
Click individual task to select or select multiple tasks using the hover task controls

- clean_up_task
 - ...lone_dbs/prod/files/src/clean_up.job
 - clean_up_cluster

Compute
Performance-optimized mode improves startup and execution speed. Disabling performance optimization gives you startup times similar to classic infrastructure and reduces your cost. Currently only applies to certain Serverless workloads. [Learn more](#)

☐ Performance optimized

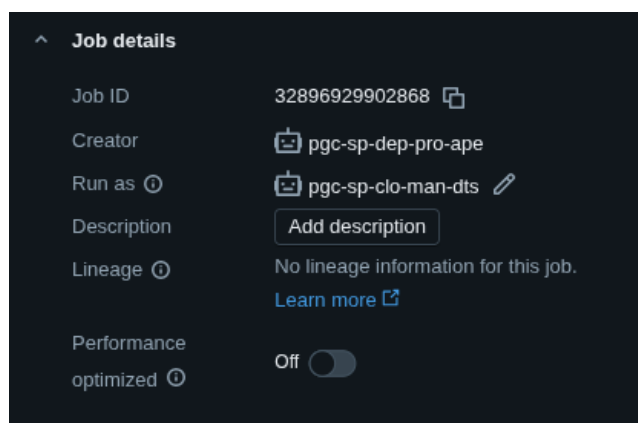
Parameters
Override job parameters for this run. Any parameters that are not overridden here will have their default values. [Learn More](#)

Key	Value
finops_project	finops {} ×
target_catalog	finops_man {} ×

2.1.3 Identitat per a l'execució i seguretat

Ambdós jobs s'executen sota un Service Principal tècnic. Això garanteix que el seu funcionament no depengui d'usuaris individuals i facilita l'auditoria i el control de permisos.

Aquest principal disposa únicament dels permisos necessaris sobre els catàlegs de manteniment.



Job details

Job ID: 32896929902868

Creator: pgc-sp-dep-pro-ape

Run as: pgc-sp-clo-man-dts

Description: Add description

Lineage: No lineage information for this job. [Learn more](#)

Performance optimized: Off

2.2 Flux operatiu després d'una incidència

En primer lloc, un cop detectada la incidència, s'identifiquen els catàlegs, esquemes i taules que caldrà replicar a l'entorn de manteniment.

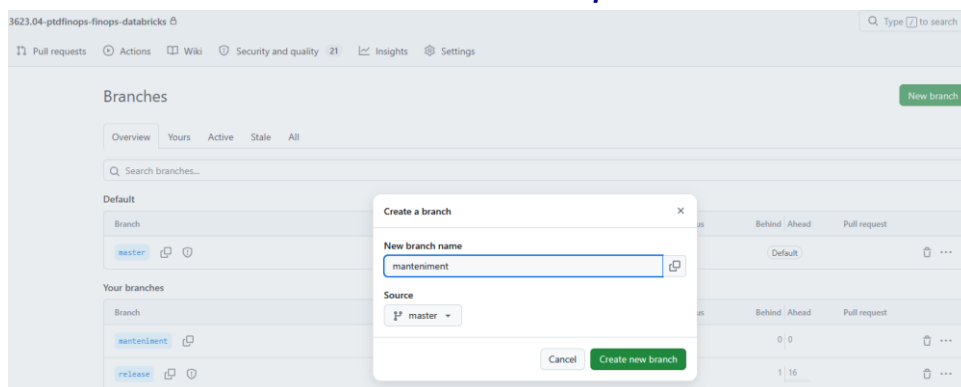
A continuació, es demana un tiquet a ACOPTD indicant els catàlegs, esquemes i taules que s'han de replicar a l'entorn de manteniment.

 Generalitat de Catalunya Centre de Telecomunicacions i Tecnologies de la Informació	PTD	N. revisió doc.: 1.0
	Gestió de l'entorn de manteniment	
	N. versió aplicació: 1.0	Build: 1.0 Pàg. 7 / 11

Quan aquesta petició estigui llesta, es disposarà en aquest entorn de totes les entitats necessàries per a la resolució del correctiu

2.2.1 Configurar projecte Bundle pel manteniment

2.2.2 Crear la branca manteniment com a còpia de master



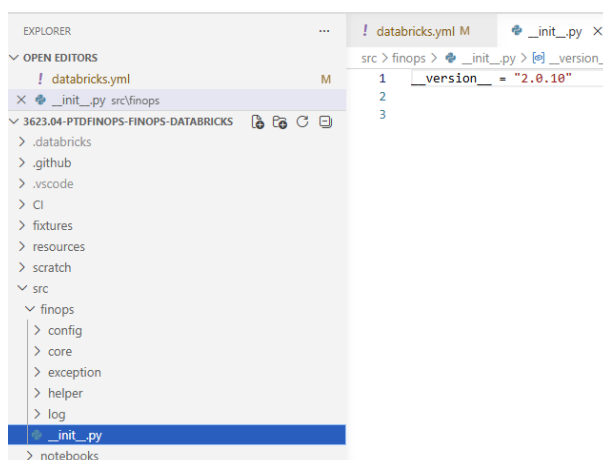
2.2.3 Editar el fitxer databricks.yml y _init_.py

Afegir l'url del workspace de manteniment:

<https://adb-727588810011327.7.azuredatabricks.net>



Modificar la versió canviant el tercer número (ex.: 2.0.11).



Amb aquests dos canvis, juntament amb el codi correctiu, es fa un “**puch**” a la branca de *manteniment* i es genera una **PR** des de la branca de *manteniment* cap a *master*.

• PR manteniment → master

3623.04-ptdfinops-finops-databricks

Pull requests

Actions

Wiki

Security and quality

21

Insights

Settings

Comparing changes

Choose two branches to see what's changed or to start a new pull request. If you need to, you can also [compare across forks](#) or [learn more about diff comparisons](#).

base: master

compare: manteniment

✓ Able to merge. These branches can be automatically merged.

Discuss and review the changes in this comparison with others. [Learn about pull requests](#)

Create pull request

1 commit

2 files changed

1 contributor

Commits on Jun 3, 2026

correctivo 2.0.11

GSPNUN0006 committed 3 minutes ago

6f56241

<>

call-reusable-workflow / databricks-ci-on-commit summary

▼ JFrog Job Summary

JFrog Job Summary

[View package details on the JFrog platform](#)

▼ Published JFrog Build Info

Build Info

[DataBricks CI on Commit 69](#)

▼ Modules published to Artifactory b

Published Modules

DataBricks CI on Commit

raw-private

3623.04

finops

2.0.11

[finops-2.0.11.zip](#)

Aquest document s'ha basat en la plantilla publicada al MQS Disseny Detallat v1.2

 Generalitat de Catalunya Centre de Telecomunicacions i Tecnologies de la Informació	PTD	N. revisió doc.: 1.0
	Gestió de l'entorn de manteniment	
	N. versió aplicació: 1.0	Build: 1.0 Pàg. 9 / 11

2.2.4 Desplegament en manteniment

L'equip de Seguretat ha de revisar els permisos de grups i rols. Es clonaran els mateixos permisos sobre les dades existents a PRO, associats al cas d'ús de les dades a clonar.

L'entorn de Manteniment tindrà una còpia de les dades de Producció per poder investigar els possibles problemes i errors que sorgeixin en producció amb dades reals. Per tant, s'hi faran servir els mateixos grups de dades que ja existeixen a l'entorn de Producció, així com els usuaris.

Es crearan grups de workspace de Manteniment per accedir al workspace i executar recursos de computació, *jobs*, etc. D'aquesta manera, es controlarà l'accés als recursos de l'entorn de Manteniment, tenint en compte que els seus usuaris podran veure dades de producció.

La còpia de dades es farà a nivell de taula, emprant la utilitat SHALLOW CLONE de Databricks i es destruirà en acabar la investigació. Per tant, **el procés que executa la còpia ha de tenir els permisos següents:**

- Sobre l'origen:
 - Catàleg origen: USE CATALOG
 - Esquema origen: USE SCHEMA
 - Taules origen: SELECT
- Sobre la destinació:
 - Catàleg destí: USE CATALOG, CREATE SCHEMA
 - Esquema destí: USE SCHEMA, CREATE TABLE, MANAGE (per poder esborrar esquema i taules)
 - Si és una taula externa, a la ubicació externa de destinació: CREATE EXTERNAL TABLE, MANAGE (per poder esborrar taules externes)

Els workflows es desplegaran utilitzant SIC+, igual que a l'entorn de Producció. Els usuaris de SIC+ i deployer hauran de tenir els mateixos permisos a l'entorn de Manteniment que a l'entorn de Producció. La configuració d'usuaris i grups dels workflows serà igual que a Producció, per tant, no es requerirà cap canvi en aquest sentit, excepte en la partícula d'entorn dels grups de workspace, que serà "MAN".

Els recursos de computació es crearan ad hoc a l'entorn de Manteniment i tindran els mateixos permisos que els equivalents de Producció. Els grups que s'utilitzaran seran els equivalents als de producció, però amb la partícula MAN, i inicialment amb els mateixos usuaris que a Producció.

 Generalitat de Catalunya Centre de Telecomunicacions i Tecnologies de la Informació	PTD		N. revisió doc.: 1.0
	Gestió de l'entorn de manteniment		
	N. versió aplicació: 1.0	Build: 1.0	Pàg. 10 / 11

Event ▾ Status ▾ Branch ▾ Actor ▾

Run workflow ▾

Use workflow from

Branch: master ▾

Artifact version in semver format, i.e:

1.0.4-RC *

2.0.11

Environment *

pro ▾

Runner OS *

{ "group": "databricks", "labels": ["self-hosted"], " ▾

ITSM ID Change Coordinator

X9253128K

Run workflow

ctti-dev / 3623.04-ptdfinops-finops-databricks

Code Issues Pull requests Actions Wiki Security and quality 21 Insights Settings

DataBridcs CD

✓ DataBridcs CD #69

Summary

All jobs

call-reusable-workflow ▾

Run details

Usage

Workflow file

Manually triggered 7 minutes ago

GSPN/N0006 62482a master

Status Success

Total duration 4m 54s

Artifacts -

databricks-cd.yaml

on: workflow_dispatch

call-reusable-workflow / itm 19s

call-reusable-workflow / wait_for_approve 3s

call-reusable-workflow / databricks-cd 2m 6s

call-reusable-workflow / execution_indicators 1m 46s

call-reusable-workflow / cancel-itm 0s

databricks

Search data, notebooks, revents, and more...

CTRL + P

pgc-tra-dbs-man-pro-001

Jobs & Pipelines

Jobs & pipelines Runs

Create new

Ingestion pipeline Ingest data from apps, databases and files

ETL pipeline Build ETL pipelines using SQL and Python

Job Orchestrate notebooks, pipelines, queries and more

Filter by name or ID substring

All Jobs Pipelines Owned by me Accessible by me Favorites Tags Run as Create

Name	Type	Tags	Run as	Trigger	Recent runs	
clean-up-catalog	Job		pgc-sp-clo-man-dts		---	
FinOpsDailyProcess	Job	pro finops	pgc-sp-fps-pro-ape	Scheduled	---	
FinOpsDailyReprocess	Job	finops	pgc-sp-fps-pro-ape		---	
FinOpsSourceExtractor	Job	finops	pgc-sp-fps-pro-ape	Scheduled	---	
FinOpsSystemTablesSync	Job	pro finops	pgc-sp-fps-pro-ape	Scheduled	---	
NotebookExecutor	Job	finops	pgc-sp-dep-pro-ape		---	
shallow-clone-tables	Job		pgc-sp-clo-man-dts		---	

Previous Next

 Generalitat de Catalunya Centre de Telecomunicacions i Tecnologies de la Informació	PTD	N. revisió doc.: 1.0
	Gestió de l'entorn de manteniment	
	N. versió aplicació: 1.0	Build: 1.0 Pàg. 11 / 11

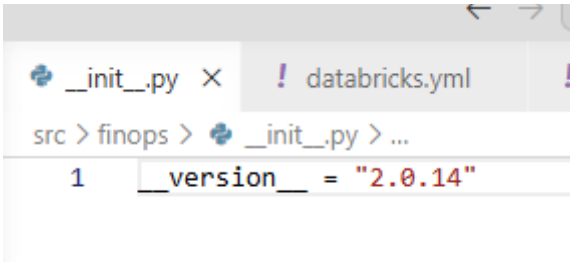
En aquest punt es pot continuar iterant sobre la incidència, sempre fent una PR des de la branca de manteniment cap a la branca master, tenint en compte que cal modificar la versió amb cada PR cap a master. Un cop finalitzada l'anàlisi i disposant de la correcció, s'ha de portar la correcció a la branca master i desplegar-la.

2.2.5 Pujar la correcció a producció (máster)

Cal tornar a editar el target de PRO al fitxer databricks.yml amb les dades de producció.

```
pro:
  mode: production
  workspace:
    host: https://adb-415924557256886.6.azuredatabricks.net
    ##host: https://adb-727588810011327.7.azuredatabricks.net
```

Modificar la versió:



Amb aquests canvis, cal fer una **PR** cap a *master* i, a continuació, desplegar a **PRO** com es fa habitualment.

Important

Després de pujar la correcció a PRO, cal baixar el fix cap a develop. Per fer-ho, s'han de seguir els passos següents:

- Crear una branca còpia de develop o bé utilitzar una branca feature que estigui actualitzada amb develop.
- Des de la nova branca (còpia de develop o feature actualitzada), fer un merge amb la branca master per incorporar tots els canvis.
- Fer la PR cap a develop.

Si no es segueixen aquests passos, es produiran conflictes en les PR cap a master.