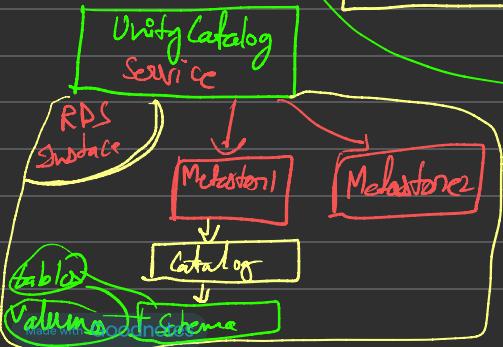


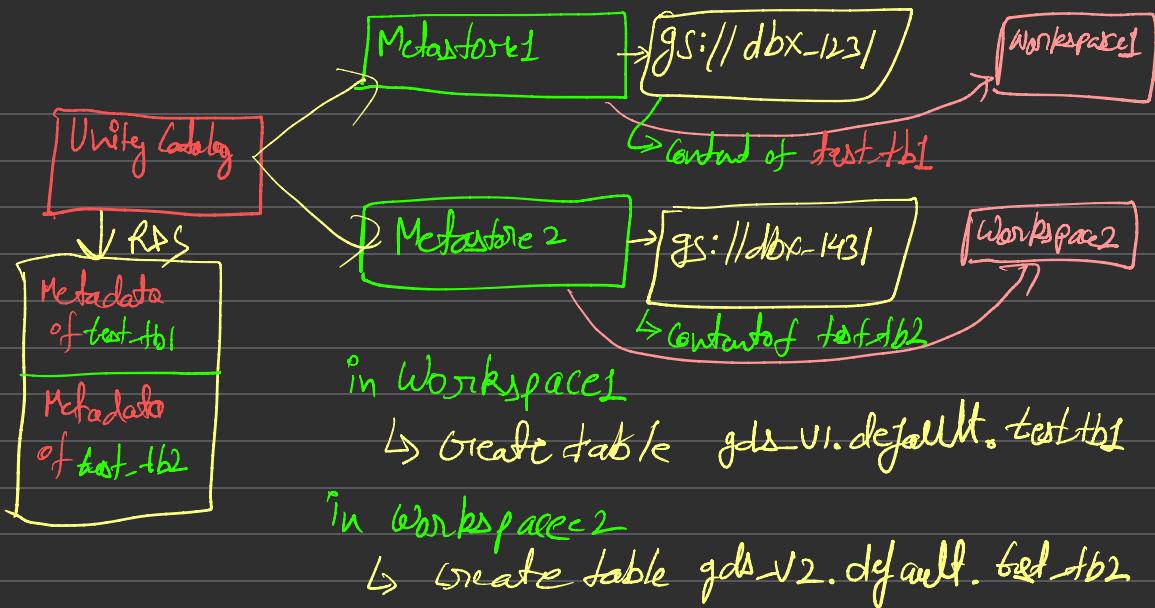
Unity Catalog (Framework)

→ Databricks' governance layer

- ↳ Manage metadata
- ↳ Actual data
- ↳ Security
- ↳ Permissions
- ↳ Access
- ↳ Lineage

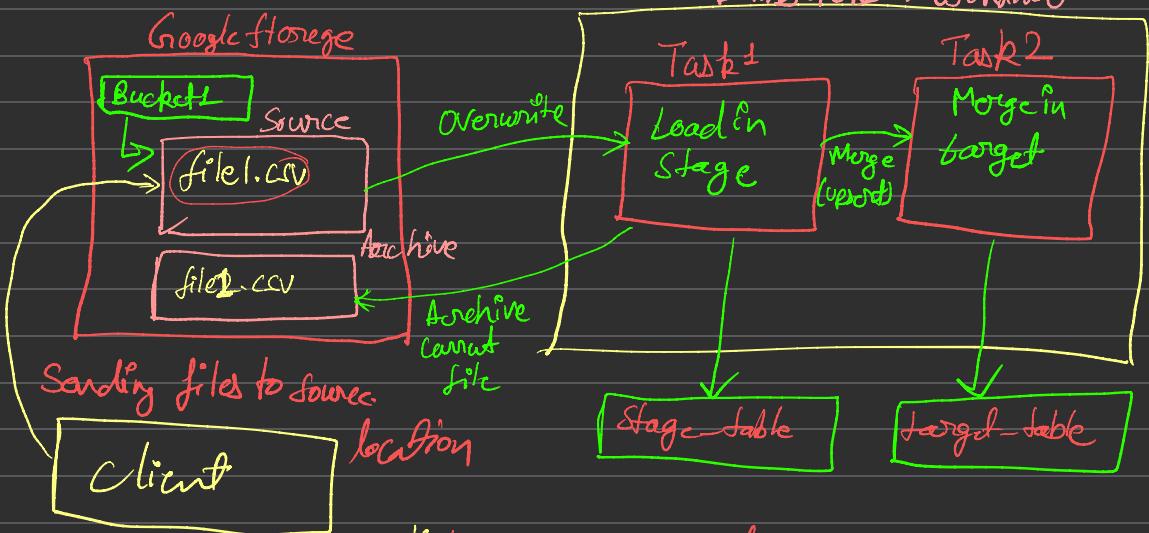
GCP





Project-2 : Incremental data load in DeltaTable with event driven architecture

Databricks Workflow



Merge Query

Source						Target					
e_id	name	dept	salary	status		e_id	name	dept	salary	status	
1	A	HR	500	X		1	A	HR	100	Y	
2	B	IT	600	Y		2	B	IT	200	Y	
3	C	HR	300	N		3	C	HR	300	Y	
4	D	FI	500	N							

Merge Into target

Using [Source] | [selected from source]
ON
source.e_id = target.e_id

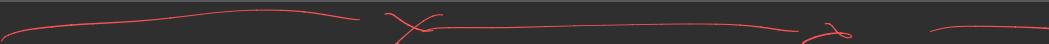
Target					
e_id	name	dept	salary	status	
1	A	HR	100	Y	
2	B	IT	200	Y	
3	C	HR	300	Y	

When matched AND source.status='Y' AND source.dept='HR' then update set
target.salary = source.salary + 1000

When matched AND source.status='Y' AND source.dept='IT' then update set
target.salary = source.salary + 500

When matched AND source.status='N' then Delete

When NOT matched then Insert *



m_id	transamt	status
M001	100	Success
M001	200	Success
M001	100	Failed
M002	50	Success
M002	100	Success
M001	50	Failed

m_id	successam	failedam
m001	300	150
m002	50	0

Sum (when status=success
then transaction_amount
else 0) → success amount

Sum (when status=failed
then amount else 0) → failed am

Source

m_id	s_am	r_am	net
m001	100	50	50

target

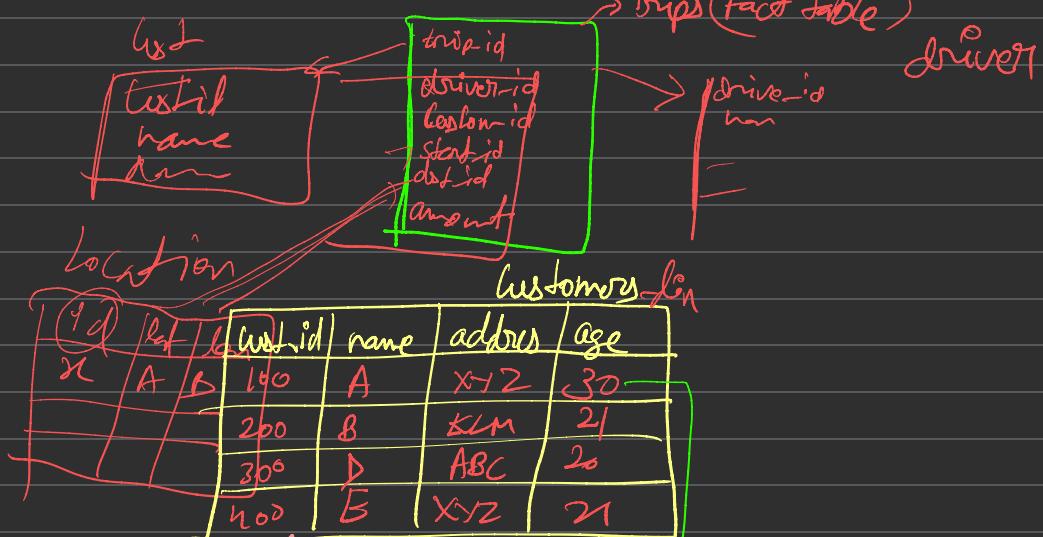
m_id	s_am	r_am	net
m001	5.00	100	400

$\text{target.s_am} = \text{target.s_am} + \text{src.s_am}$

$m001, 600, 150, 450$

SCD2 Merge On Delta Table

→ Fact Tables } Data Warehousing
 → Dimension }



SCD1 ⇒ Update Overwrite after SCD1

change	x+y	z
100	X+Y	32

Customer.dim (1)

cust_id	name	address	age	active	from	to
100	A	XYZ	30	N	12-Jan-2024	15-Jan-2024
200	B	KLM	21	Y	12-Jan-2024	25-Jan-2024
300	C	ABC	20	Y	12-Jan-24	9999
400	E	XYZ	21	N	12-Jan-24	9999
100	A	XYZ	32	Y	15-Jan-2024	15-Jan-9999

Change

100 JA XYZ 32

SCD2 → Maintain history in Dim table

t

15-Jan-2024

15-Jan-

100, A, 12-Jan-2024, 12-Jan-9999

100, A, 15-Jan-2024, 15-Jan-9999

9999

cust_id = fact_cust

AND

Year (9999)

target.valid_to = source.valid_from

Medallion Architecture

