

Azure DevOps Assignment - 3

Mitushi Vishwakarma

Notes :

Continuous Integration / Continuous Deployment

A CI / CD pipeline is a concept central to software. It spans a whole field of processes, testing methods and tooling, all facilitated by the Git code versioning process.

Continuous Integration :-

If we add a new pipeline, then we test by running data through it.

This ensures the condition connection is okay.

NOVEMBER						
MONDAY			WEEK 45 • DAY 310-055			
M	T	W	T	F	S	S
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31				

Money should not and must not ever change our Commitment

Continuous deployment :-

Once pipeline works properly,
it is deployed in production environment

→ CI checks and tests every new piece
of code, we add to data pipeline.

→ CD ensures that once tested and approved.
this code gets added to live system
without manual intervention.

For a data engineer,

- it means faster, more reliable updates
- data processes
- high quality data delivery consistently.
- if any issue, fixed swiftly.

CI / CD in data pipelines

CI / CD, in the context of data pipeline deployment, focuses on automating data operations and transformations

WEEK	M	T	W	T	F	S	S
44	30	31	1	2	3	4	5
45	6	7	8	9	10	11	12
46	13	14	15	16	17	18	19
47	20	21	22	23	24	25	26
48	27	28	29	30			

04-11-2023
NOVEMBER
2023 SATURDAY WK 44 • DAY 308-057
O4

I attribute my success to this, I never gave or took any excuse

This merges development, testing and operations workflows into a unified, automated process, ensuring that data assets are consistently high quality and that data infrastructure evolves smoothly, even at scale.

CI / CD characteristics :-

Automated Testing

Automated Tests check the integrity and quality of data transformations, ensuring the data is processed as expected and any error is spotted early.

Version Control

Data pipeline code is stored in repositories like Git, allowing tracking and managing changes.

Consistent environment

CI tools can run tests in environments that mirror production, ensuring that differences in configuration or dependencies don't introduce errors.

NOVEMBER							DECEMBER						
SUN	MON	TUE	WED	THU	FRI	SAT	SUN	MON	TUE	WED	THU	FRI	SAT
WK 44 • DAY 307-058	4	5	6	7	8	9	10	11	12	13	14	15	16
17	18	19	20	21	22	23	24	25	26	27	28	29	30
31	51	52	53	54	55	56	57	58	59	60	61	62	63

The truly creative person is one who can think crazy

Data Quality Checks

These might include checks for null values, data range violations, data type mismatches, or other custom quality rules.

Azure Synapse Analytics

- limitless analytics service
- brings together enterprise data warehousing and Big data analytics.
- gives freedom to query data, either using serverless or provisioned resources
- unified experience to ingest, prepare, manage and serve data for immediate business intelligence and machine learning needs

SQL Analytics

- best in class price per performance
- Industry leading security
- Workload aware query execution.
- Data flexibility
- developer productivity

44	45	46	47	48
4	5	6	7	8
12	13	14	15	16
17	18	19	20	21
22	23	24	25	26
27	28	29	30	

02 - 11 - 2023
2023 THURSDAY NOVEMBER WK 44 • DAY 306-059 09.00 10.00 11.00 12.00

02

A small group of committed citizens can change the world

Leveraging ISV partners with Azure Synapse Analytics

- full backward compatibility with Azure SQL Data Warehouse for delta integration and orchestration.
- Additional analytics capabilities in Azure Synapse unlock new ISV scenarios.
- Azure Synapse + ISV can bring data continuity with Azure Machine Learning and Power BI.
- Reduce migration effort by reusing existing partner platforms.

OLTP workloads are not suitable
and Data preparation not suitable

- | OLTP | Data preparation |
|---------------------------------------|------------------------------------|
| → high frequency reads and writes | → Row - by - row processing needs. |
| → large no. of singleton selects | → incompatible XML |
| → high volumes of single row inserts. | |

MON	TUE	WED	THU	FRI	SAT	SUN	
44	45	46	47	48	27	28	29
5	6	7	8	9	10	11	12
12	13	14	15	16	17	18	19
19	20	21	22	23	24	25	26
26	27	28	29	30			

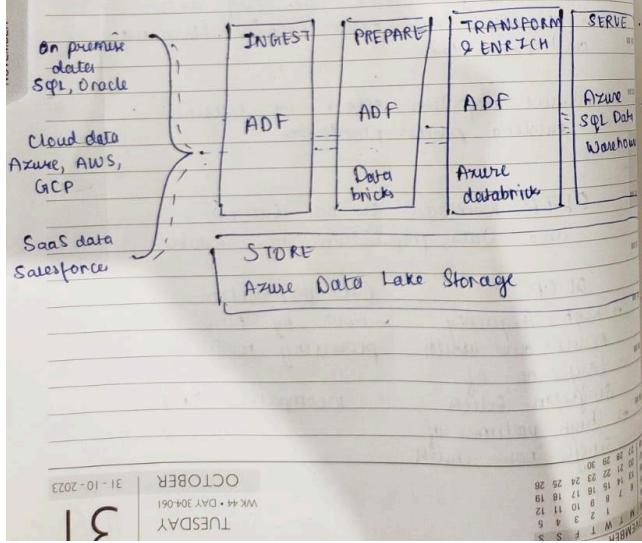
2023 NOVEMBER WK 44 • DAY 305-060

WESENDSDAY 01-11-2023 OI

Which workloads are suitable

- Analytics
- Store large volumes of data
- Consolidate disparate data into a single location
- Shape, model, transform and aggregate data
- Batch / Micro-batch loads

Modern Data Warehouse



Azure Synapse Analytics



This is how
modern data warehouse
is replaced by Azure
Synapse Analytics.

MONDAY
30 OCTOBER 2023
WEEK 44 • DAY 30462
30.10.2023

OCTOBER 2023						
MONDAY		TUESDAY	WEDNESDAY	THURSDAY	FRIDAY	SATURDAY
30	31	1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31	1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30
31	1	2	3	4	5	6

