# AZURE

Generics

1. How an azure account is modeled?
2. What is a resource manager? How resource procurement works?
3. Explain azure global infrastructure in detail.
   1. geography
   2. regions
   3. region pair
   4. availability zones
   5. data centers
   6. availability sets (fault domain, update domain)
   7. servers
4. How will you select a particular location for your resource?

VM

1. Why we need to use azure virtual machines?
2. When do we need to use azure virtual machines?
   1. Typical ‘iaas’ requirements for ex. to migrate legacy apps
3. Internally, how a virtual machine will be procured?
   1. virtualization
4. What is a VMSS and why we need it?

ADF

1. What is ADF?
2. What are some of the important components inside ADF?
3. What are the different ways to interact with ADF?
4. What are pipelines and activities in ADF?
5. What are linked services and datasets in ADF?
6. What are the different kinds of triggers in ADF? Explain the working of each one of them.
   1. schedule
   2. tumbling window
   3. event based
7. What are integration runtimes? Why we need it?
8. What are different types of integration runtimes supported in ADF?
   1. azure integration (default)
   2. self-hosted
   3. azure ssis
9. Can you pls explain what is azure integration runtime?
10. Can you pls explain what is managed vnet azure integration runtime?
11. Can you pls explain what is self-hosted integration runtime and why it is needed?
12. What is parameterization in ADF? Why we need it?
13. What are the different types of parameters?
14. What are different ways we can create parameters?
    1. UI
    2. json content
15. What are the different types of variables in ADF? What is the use of each one of them?
    1. system
    2. user
16. What are the different types of activities in ADF?
    1. data movement - copy, delete
    2. data transformation - script, stored procedure, data flow
    3. control flow - set variable, append variable, execute pipeline, filter, forEach, get metadata, if condition, wait, until, web, webhook, switch, validation, lookup, fail
17. What are data flows in ADF? What are the different types of data flows in ADF?
    1. mapping
    2. wrangling
18. What are the different types of transformations in data flows in ADF?
    1. in mapping data flows - filter, aggregate, join (and fuzzy join), conditional split, derived column, exists, union, lookup, sort, select, pivot, unpivot, surrogate, window, alter row, flatten, rank, parse, stringify, assert, cast
    2. in wrangling data flows - merged queries, group by
19. Can you use parameters in mapping data flow?
20. How will you perform schema validation in mapping data flow?
21. What is an inline dataset in mapping data flow and what is its use?
22. What are flowlets and how they can be useful?
23. What are user-defined functions in mapping data flow?
24. What is cache sink and how can we lookup the cached data in mapping data flow?
25. How can we reduce cluster spin up time for multiple activities in mapping data flow?
26. What are the different ways of authoring in ADF?
27. How can we setup a github repository for authoring? How does it work?
28. How will you setup ci/cd in ADF?
29. What are annotations in ADF and why they are needed?
30. What are templates in ADF and why they are needed?
31. How can you check the costs incurred in ADF?
32. How will you configure alerts for your pipelines in ADF?
33. -------------------------------------- Scenario based ------------------------------------------
34. How to read json output of one activity in the next activity?
35. How can we import tables from a website?
36. How can we return some values from child pipeline to parent pipeline?
37. How can we import data from an API which supports pagination in a page wise manner?
38. How can we deactivate an activity and why its needed?
39. How will you implement CDC for a SQL source in mapping data flow?
40. How will you implement SCD type 1 in mapping data flow?
41. How will you implement SCD type 2 in mapping data flow?

# AWS

# GCP