ICM Questions

Module 1:

1. Digital data is continuously generated, collected, stored, and analyzed through ……. (Software)

2. The digital universe generates approximately 4.4 trillion GB of data annually (44 in 2020) by ……. (International Data Corporation (IDC))

3. …….. is made up of Internet-connected equipment and sensors (IOT)

4. Organizations are dependent on continuous and reliable .....

(Access to information)

5. Organizations seek to effectively store, protect, process, manage, and leverage ..... (information)

6. Organizations are increasingly implementing intelligent storage solutions To

....., ..., ... (efficiently store and manage information, gain competitive advantage, derive new business opportunities)

7. A collection of facts that is transmitted and stored in electronic form, and processed through software

a) Data Center
b) Digital Data
c) information

Note: type of digital data
-Unstructured
-Quasi-structured
-Semi-structured
-Structured

8. Textual data files with an apparent pattern, enabling analysis, E.g., Spreadsheets and XML files

a) Unstructured
b) Quasi-structured

c) Semi-structured
d) Structured

9. Data having a defined data model, format, structure, E.g., Database

a) Unstructured
b) Quasi-structured
c) Semi-structured
d) Structured

10. Data that has no inherent structure and is usually stored as different types of files, E.g., Text documents, PDFs, images, and videos

a) Unstructured
b) Quasi-structured
c) Semi-structured
d) Structured

11. Textual data with erratic formats that can be formatted with effort and software tools, E.g., Clickstream data

a) Unstructured
b) Quasi-structured
c) Semi-structured
d) Structured

12. Processed data that is presented in a specific context to enable useful interpretation and decision-making.

a) Data Center
b) Digital Data
c) information

13. Information is stored on storage devices on

a) non-volatile media
b) volatile media

Note: Types of storage devices
- Magnetic storage devices

- Optical storage devices
- Flash-based storage devices

14. Hard disk drive and magnetic tape
    a) Magnetic storage devices
    b) Optical storage devices
    c) Flash-based storage devices
15. Solid state drive, memory card, and USB thumb drive
    a) Magnetic storage devices
    b) Optical storage devices
    c) Flash-based storage devices

16. Blu-ray disc, DVD, and CD
    a) Magnetic storage devices
    b) Optical storage devices
    c) Flash-based storage devices

17. Storage devices are assembled within a..... (Storage system or "array")

18. Storage systems along with other IT infrastructure are housed in a .....
(Data center)

19. hardware component that contains a group of homogeneous/heterogeneous storage devices assembled within a cabinet

    a) Facility
    b) Data center
    c) Storage system

20. A facility that houses IT equipment including compute, storage, and network components, and other supporting infrastructure for providing centralized data-processing capabilities.

    a) Data Center
    b) Digital Data

c) information

Note: A data center comprises
- facility
- IT equipment
- supporting infrastructure

21. The building and floor space where the data center is constructed

a) facility
b) IT equipment
c) supporting infrastructure

22. Power supply, fire detection, HVAC, and security systems

a) facility
b) IT equipment
c) supporting infrastructure

23. Compute, storage, and network equipment

a) facility
b) IT equipment
c) supporting infrastructure

Note: Key Characteristics of a Data Center
- Availability
- Security
- Performance
- Scalability
- capability
- Data Integrity
- Manageability

24. Policies and procedures should be established, and control measures should be implemented to prevent unauthorized access to and alteration of information

a) Availability
b) Security
c) Performance

25. deploy additional resources such as compute systems, new applications, and databases to meet the growing requirements. Data center resources should scale to meet the changing requirements, without interrupting business operations

a) Scalability
b) capability
c) Data Integrity

26. Availability of information as and when required should be ensured

a) Availability
b) Security
c) Performance

27. Data center operations require adequate resources to efficiently store and process large and increasing amounts of data. When capacity requirements increase, additional capacity should be provided either without interrupting availability or with minimal disruption

a) Scalability
b) capability
c) Data Integrity

28. provide optimal performance based on the required service levels

a) Availability
b) Security
c) Performance

29. ensure that data is stored and retrieved exactly as it was received.

a) Scalability
b) capability
c) Data Integrity

30. A data center should provide easy, flexible, and integrated management of all its components.

a) Data Integrity
b) Manageability
c) Efficient manageability

31. can be achieved through automation for reducing manual intervention in common, repeatable tasks.

a) Data Integrity
b) Manageability
c) Efficient manageability

Note: Key Data Center Management Processes

- Monitoring
- Reporting
- Provisioning
- Planning
- Maintenance

32. Configuring and allocating resources to meet the capacity, availability, performance, and security requirements

a) Monitoring
b) Reporting
c) Provisioning

33. Continuously gathering information on data center resources

a) Monitoring
b) Reporting
c) Provisioning

34. Presenting the details on resource performance, capacity, and utilization

a) Monitoring
b) Reporting
c) Provisioning

35. Estimating the amount of resources required to support business operations

a) Provisioning
b) Planning
c) Maintenance

36. Ensuring the proper functioning of resources and resolving incidents

a) Provisioning
b) Planning
c) Maintenance

37. First platform based on

a) mainframes
b) Client-Server

38. Users connect to mainframes through

a) client program
b) web-interface

c) terminals

39. Applications and databases hosted centrally

a) first platform
b) second platform
c) third platform

Note: Challenges with mainframes

- Substantial CAPEX and OPEX
  - High acquisition costs
  - Considerable floor space and energy requirements

40. Second platform based on

a) mainframes
b) Client-Server

41. Users connect through

a) client program & web-interface
b) terminals

42. Distributed application architecture , Servers receive and process requests for resources from clients

a) first platform
b) second platform
c) third platform

Note: Challenges with client-server model

- Creation of IT silos
- Hardware and software maintenance overhead
- Scalability to meet the growth of users and workloads

43. The four pillars are transforming the way organizations are using technology for business operations

a) first platform
b) second platform
c) third platform

Note: The four pillars

- Cloud , Big Data , Mobile , Social

Module 2:

1.A model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources, (e.g., servers, storage, networks, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction

a) Cloud computing
b) Cloud
c) Big Data

2. a collection of network-accessible hardware and software resources

a) Cloud computing
b) Cloud
c) Big Data

Note: Essential Cloud Characteristics

- On-demand self-service
- Broad Network Access
- Resource Pooling
- Rapid Elasticity
- Measured Service

Note: Cloud Service Models

- Infrastructure as a Service (IaaS)
- Platform as a Service (PaaS)
- Software as a Service (SaaS)

3. The capability provided to the consumer is to provide processing, storage, networks, and other fundamental computing resources where the consumer can deploy and run arbitrary software, which can include operating systems and applications.

a) Infrastructure as a Service (IaaS)
b) Platform as a Service (PaaS)
c) Software as a Service (SaaS)

4. The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, and deployed applications; and possibly limited control of select networking components, (e.g., host firewalls).

a) Infrastructure as a Service (IaaS)

b) Platform as a Service (PaaS)
c) Software as a Service (SaaS)

5. The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages, libraries, services, and tools supported by the provider.

a) Infrastructure as a Service (IaaS)
b) Platform as a Service (PaaS)
c) Software as a Service (SaaS)

6. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly configuration settings for the application-hosting environment

a) Infrastructure as a Service (IaaS)
b) Platform as a Service (PaaS)
c) Software as a Service (SaaS)

7. The capability provided to the consumer is to use the provider's applications running on a cloud infrastructure. The applications are accessible from various client devices through either a thin client interface, such as a web browser, (e.g., web-based email, or a program interface.

a) Infrastructure as a Service (IaaS)
b) Platform as a Service (PaaS)
c) Software as a Service (SaaS)

8. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings.

a) Infrastructure as a Service (IaaS)
b) Platform as a Service (PaaS)
c) Software as a Service (SaaS)

Note: Cloud Deployment Models

- Public Cloud
- Private Cloud
- Community Cloud
- Hybrid Cloud

Note: Hybrid Cloud Model Use Cases

- Cloud bursting
- Web application hosting
- Migrating packaged applications
- Application development and testing

9. Hosting less critical applications on the public cloud

a) Cloud bursting
b) Web application hosting
c) Migrating packaged applications
d) Application development and testing

10. Provisioning resources for a limited time from a public cloud to handle peak workloads

a) Cloud bursting
b) Web application hosting
c) Migrating packaged applications
d) Application development and testing

11. Migrating standard packaged applications such as e-mail to the public cloud

a) Cloud bursting
b) Web application hosting
c) Migrating packaged applications
d) Application development and testing

12. Developing and testing applications in the public cloud before launching them

a) Cloud bursting
b) Web application hosting
c) Migrating packaged applications
d) Application development and testing

13. Information assets whose high volume, high velocity, and high variety require the use of new technical architectures and analytical methods to gain insights and for deriving business value.

a) Cloud computing
b) Cloud
c) Big Data

14. What type of data that big data include

a) Structured

b) non-structured

c) both

Note: Big Data requires highly-scalable storage

Note: Characteristics of Big Data

- Volume
- Velocity
- Veracity
- Varity
- Variability
- Value

15. Constantly changing meaning of data

a) Varity
b) Variability
c) Value

16. Cost-effectiveness and business value

a) Varity
b) Variability
c) Value

17. Diverse data from numerous sources

a) Varity
b) Variability
c) Value

18. Challenges in transforming and trusting data

a) Volume
b) Velocity
c) Veracity

19. Varying quality and reliability of data

a) Volume
b) Velocity
c) Veracity

20. Massive volumes of data and Challenges in storage and analysis

a) Volume
b) Velocity
c) Veracity

21. Rapidly changing data

a) Volume
b) Velocity
c) Veracity

22. Challenges in integration, and analysis

a) Varity
b) Variability
c) Value

23. Challenges in gathering and interpretation

a) Varity
b) Variability
c) Value

24. Challenges in real-time analysis

a) Volume
b) Velocity
c) Veracity

25. Data for analytics typically comes from

a) data warehouses
b) data lakes
c) both

26. central repository of integrated data gathered from different sources

a) data warehouses
b) data lakes

27. collection of data that is stored as an exact or near-exact copy of the source format

a) data warehouses
b) data lakes

28. Data is classified, organized, or analyzed only when it is accessed

a) data warehouses
b) data lakes

29. Stores current and historical data in a structured format

a) data warehouses
b) data lakes

30. Designed for query and analysis to support decision making

a) data warehouses
b) data lakes

31. Enables analysts to implement their own analysis techniques

a) data warehouses
b) data lakes

Note: Components of a Big Data Analytics Solution

- Query
- MapReduce
- Storage

32. Efficient way to process, store, and retrieve data ,Platform for user-friendly analytics systems

a) Query
b) MapReduce
c) Storage

33. Distributed architecture, Non-Relational, non-structured data

a) Query
b) MapReduce
c) Storage

34. Parallel computation a cross many servers, Batch-processing model

a) Query
b) MapReduce
c) Storage

35. consist of multiple nodes collectively called a “cluster” (Storage systems)

36. Storage systems Based on (distributed file systems)

37.Each node has (processing capability) and (storage capacity)

38. (NoSQL database) may be implemented on top of the distributed file system

39. Simplifies the specification of MapReduce operations, and the retrieval and analysis of the results (Query)

40. Designed to retrieve and process massive amounts of non-structured data (Query)