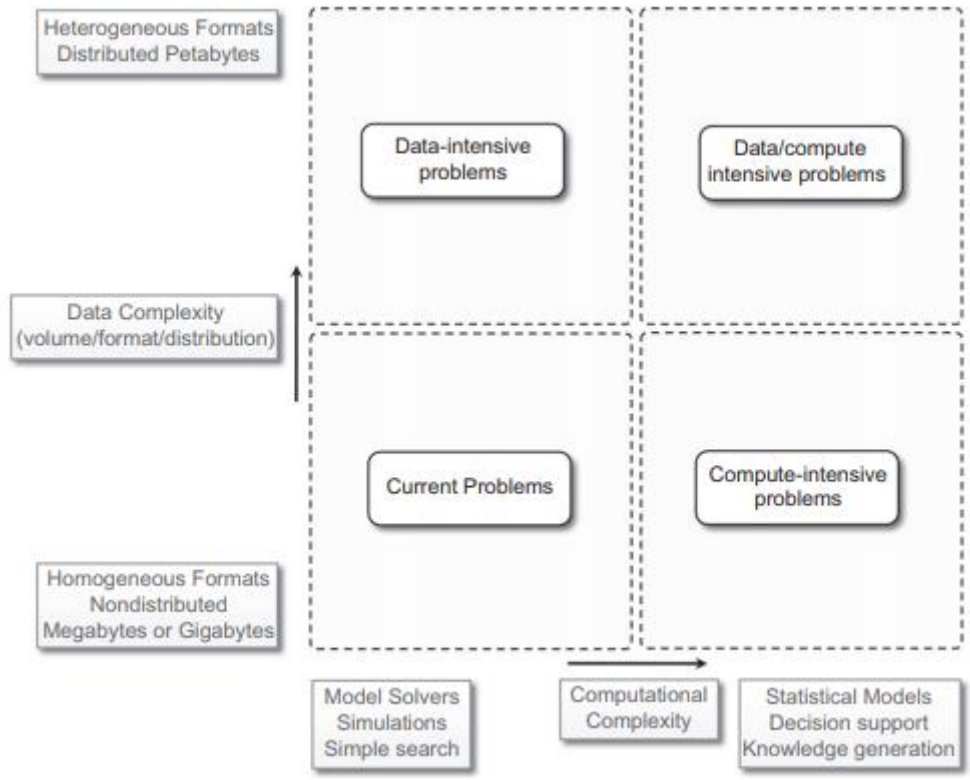
Department of Artificial Intelligence and Machine Learning

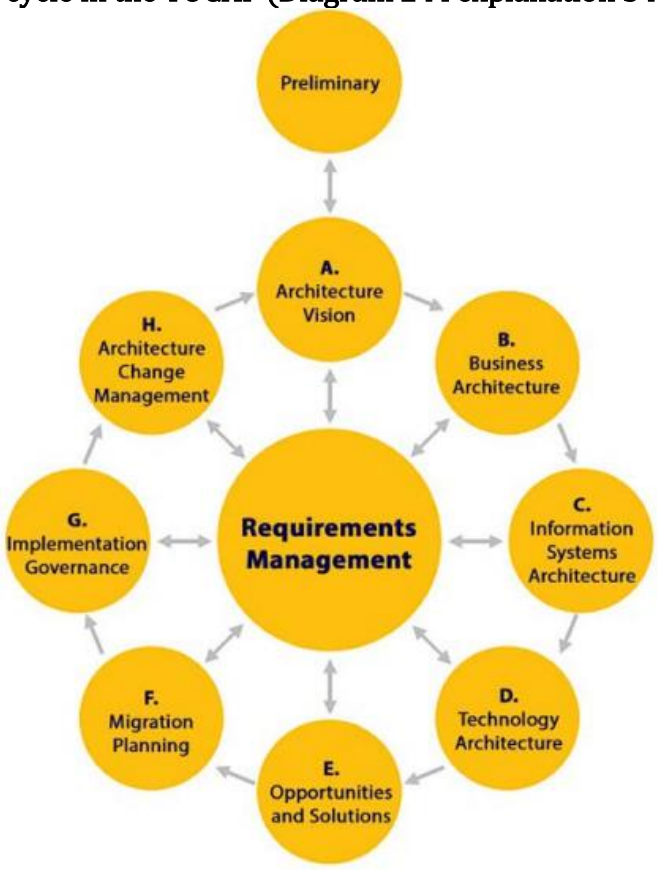
CLOUD COMPUTING AND ARCHITECTURES
Scheme and Solutions

SL. No	Questions	Marks
1 a	**With diagram discuss Data-intensive research issues (Figure 2 M, Research Issues 3M any 3)**	

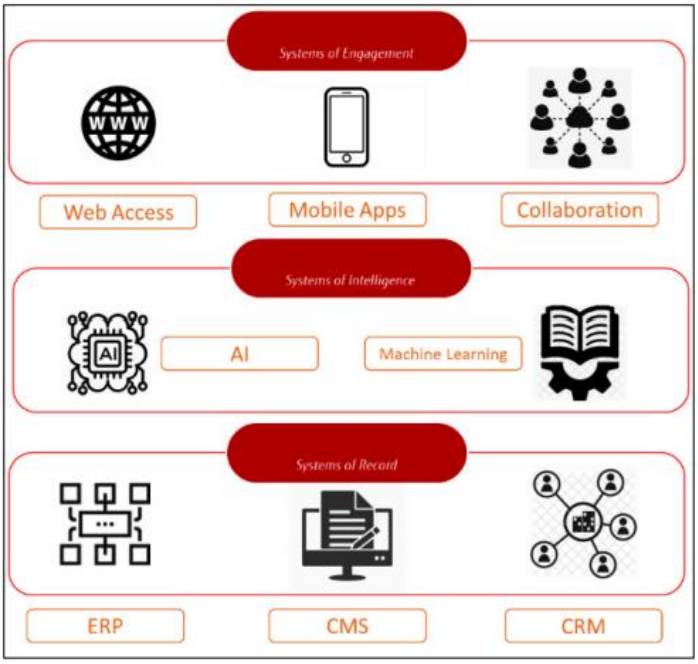


- Scalable algorithms that can search and process massive datasets
- New metadata management technologies that can scale to handle complex, heterogeneous, and distributed data sources
- Advances in high-performance computing platforms aimed at providing a better support for accessing in-memory multiterabyte data structures
- High-performance, highly reliable, petascale distributed file systems • Data signature-generation techniques for data reduction and rapid processing
- New approaches to software mobility for delivering algorithms that are able to move the computation to where the data are located
- Specialized hybrid interconnection architectures that provide better support for filtering multigigabyte datastreams coming from high-speed networks and scientific instruments
- Flexible and high-performance software integration techniques that facilitate the combination of software modules running on different platforms to quickly form analytical pipelines

1 b	**Describe the characteristics of Amazon Simple Storage Service (S3). (each Characterstics 1 mark)** 1. **The storage is organized in a two-level hierarchy.** S3 organizes its storage space into buckets that cannot be further partitioned. This means that it is not possible to create directories or other kinds of physical groupings for objects stored in a bucket 2. **Stored objects cannot be manipulated like standard files.** S3 has been designed to essentially provide storage for objects that will not change over time 3. **Content is not immediately available to users.** The main design goal of S3 is to provide an eventually consistent data store. As a result, because it is a large distributed storage facility, changes are not immediately reflected. 4. **Requests will occasionally fail.** Due to the large distributed infrastructure being managed, requests for object may occasionally fail. Under certain conditions, S3 can decide to drop a request by returning an internal server error 5. **Access to S3 is provided with REST full Web services**. These express all the operations that can be performed on the storage in the form of HTTP requests (GET, PUT, DELETE, HEAD, and POST),	
2 a	**Discuss the different ways to address the bucket in Amazon Cloud with an example (Explanation 3 M example 2M)** **Amazon offers three different ways of addressing a bucket:** ▪ **Canonical form:** http://s3.amazonaws.com/bukect_name/. The bucket name is expressed as a path component of the domain name s3.amazonaws.com. This is the naming convention that has less restriction in terms of allowed characters, since all the characters that are allowed for a path component can be used. ▪ **Subdomain form:** http://bucketname.s3.amazon.com/. Alternatively, it is also possible to reference a bucket as a subdomain of s3.amazonaws.com. ▪ **Virtual hosting form:** http://bucket-name.com/. Amazon also allows referencing of its resources with custom URLs. This is accomplished by entering a CNAME record into the DNS that points to the subdomain form of the bucket URI. **Example:** • Canonical form: http://s3.amazonaws.com/bukect_name/object_name • Subdomain form: http://bucket-name/s3.amzonaws.com/object_name • Virtual hosting form: http://bucket-name.com/object_name	
2 b	**What is a role? What types of roles can be used in Microsoft Azure? (2M for defining role Each role 1M with explanation)** ▪ A role is a runtime environment that is customized for a specific compute task. ▪ Roles are managed by the Azure operating system and instantiated on demand in order to address surges in application demand. **Currently, there are three different roles:**	

	Web role Worker role Virtual Machine (VM) role. **Web role** ▪ The Web role is designed to implement scalable Web applications. ▪ Web roles represent the units of deployment of Web applications within the Azure infrastructure. They are hosted on the IIS 7 (Internet Information Server) Web Server, which is a component of the infrastructure that supports Azure **Worker role** ▪ Worker roles are designed to host general compute services on Azure. ▪ The **Azure SDK** provides developers with convenient APIs and libraries that allow connecting the role with the service provided by the runtime and easily controlling its startup as well as being notified of changes in the hosting environment. **Compute services** ▪ Virtual Machine (VM) role. ▪ The Virtual Machine role allows developers to fully control the computing stack of their compute service by defining a custom image of the Windows Server 2008 R2 operating system and all the service stack required by their applications	
3 a	**Illustrate the ADM cycle in the TOGAF (Diagram 2 M explanation 3 M)**  The core of TOGAF is the ADM cycle, short for Architecture Development Method.	

	Also, in architecting multi-cloud environments, ADM is applicable. The ground principle of ADM is B-D-A-T: the cycle of **business**, **data**, **applications**, and **technology**. Every business should have the goal of generating revenue and earning money. That's not really a strategy. The strategy is defined by how it generates money with the products the business makes or the services that it delivers. A good strategy comprises a well-thought-out balance between timing, access to and use of data, and something that has to do with braveness—daring to make decisions at a certain point in time. That decision has to be based on—you guessed it—proper timing, planning, and the right interpretation of data that you have access to.	
3 b	**Discuss the five most important data principles in Multicloud. (Each Principle carries 1 mark)** The most often used data principles are related to data confidentiality and, from that, protecting data. The five most important data principles are: **1. Accuracy:** Data should be accurate, complete, and reliable. Inaccurate data can lead to flawed insights and decisions and can have serious consequences. **2. Relevance:** Data should be relevant to the problem or question at hand. Unnecessary or irrelevant data can add noise to the analysis and make it harder to extract meaningful insights. **3. Timeliness:** Data should be timely and up to date. Outdated or stale data can lead to incorrect or misleading conclusions. **4. Consistency:** Data should be consistent in terms of format, definitions, and units of measurement. Inconsistent data can create confusion and make it difficult to combine or compare different sources of information. **5. Privacy and security:** Data should be protected from unauthorized access and use. Sensitive data, such as personal or financial information, should be handled with care and stored securely to avoid breaches or leaks.	
4 a	**Illustrate how the Intelligent layer used to manage the big data** **(Diagram 2 M Explantion 3m)** Two important technology terms that have become quite common in cloud environments. **1. Systems of record:** Systems of record are data management or information storage systems; that is, systems that hold data. **2. Systems of engagement:** Systems of engagement are systems that are used to collect or access data. This can include a variety of systems: think of email, collaboration platforms, and content management systems	

	 A growing number of architects believe that there will be a new layer in the model. That layer will hold **"systems of intelligence"** using machine learning and artificial intelligence (AI).	
4 b	**Explain the below quality attributes used in the Architecture Principle (Each carries 2M)** **i. Configurability** **ii. Scalability** **Configurability** Companies use cloud technology to be able to respond quickly to changes, and to become agile. As a result, systems are not static in the cloud. Systems in the cloud must be easy to configure They might include: • CPU and memory allocation • Storage settings and disk usage • Boot sequence • Security settings such as hardening the system • Access management • Operating system configurations **Scalability** Scaling out is also referred to as horizontal scaling. When we scale out an environment, we usually add systems such as virtual machines to that environment. In scaling up—or vertical scaling—we add resources to a system, typically CPUs, memory, or disks in a storage system.	
5	**Assume the friends are stored as Person->[List of Friends], our friends list is then:** **A -> B C D** **B -> A C D** **C -> A B D** **D -> A B C** **Each line will be an argument to a mapper. For every friend in the list of friends, the**	

mapper will output a key-value pair. The key will be a friend along with the person. The value will be the list of friends. The key will be sorted so that the friends are in order, causing all pairs of friends to go to the same reducer.

Apply map reduce to find the common friend of D and A
(Write the map reduce function, Write both the map set and reducer set, and conclude)
(Formula 2 M, Steps 3M, Mapping 2M, Reducing 2M, result 1)

More precisely, the MapReduce model is expressed in the form of two functions, which are defined as follow:

$$map\ (k1, v1) \rightarrow list(k2, v2)$$
$$reduce(k2, list(v2)) \rightarrow list(v2)$$

Mapping: For each user, emit key-value pairs where the key is the user and the value is a list of their friends.
Shuffling and Sorting: Group together the lists of friends for each user.
Reducing: For each pair of users, find the intersection of their friend lists.

Let's apply these steps:

Mapping:

- For each user, emit key-value pairs where the key is the user and the value is a list of their friends.

mathematicaCopy code

```
A -> [B, C, D]
B -> [A, C, D]
C -> [A, B, D]
D -> [A, B, C]
```

Shuffling and Sorting:

- This step is handled by the MapReduce framework.

Reducing:

- For each pair of users, find the intersection of their friend lists.
- In this case, we need to find the intersection of the friend lists of D and A.

mathematicaCopy code

```
User D's friends: [A, B, C]
User A's friends: [B, C, D]
```

The intersection of D's and A's friend lists is [B, C], which means B and C are the common friends of D and A.

Therefore, the common friends of user's D and A are users B and C.