Important Instructions:

1) Open this MS-Word document and start writing answers below each respective question given on page 2.

2) Answers the question in the same sequence in which they appear.

3) Provide to the point and concrete answers. Some of the questions are open ended and therefore must be answered using your own opinion and thoughts but backed with logical reasons.

4) First read the questions and understand what is required of you before writing the answer.

5) Attempt the paper yourself and do not copy from your friends or the Internet. Students with exactly similar answers or copy paste from the Internet will not get any marks for their assignment.

6) You can contact me for help if you have any doubt in the above instructions or the assignment questions.

7) All questions must be attempted.

8) Do not forget to write your name, university ID, class and section information.

9) Rename you answer file with your university ID# before uploading to SIC.

10) When you are finished with writing your answers and are ready to submit your answer, convert it to PDF and upload it to SIC unzipped, before the deadline mentioned on SIC.

**Mid Semester Assignment
Course: - Distributed Computing**

Deadline: - Mentioned on SIC   **Marks: - 30**

Program: - MS (CS)   **Dated: 20 April 2020**

Student Name: Muhammad Zulfiqar Khan   **Student ID#: 13602**

Class and Section: MS (CS)

Question1: **Provide an example of a modern Distributed System not discussed in the course; discuss how this system solves certain challenges by employing distributed architecture.** **(5)**

Answer1: Modern Distributed System means multiple computer, software and application to form in single form or work for single task. They connect with each other by a network for information exchange.

Example: Hadoop

Hadoop is a modern distributed system. it is open software framework using for multiple types of big data storage. Applications written in Hadoop can run in huge information libraries distributed on MapReduce. It is easy to access and more sufficient for big data files distribution by different nodes.

The following challenges solves this distributed system by employing distributed architecture.

1. To handle large or huge amount data: this system can provide a distributed system based on cluster to reduce volume and store big data. Also stored any kind of data from social media.

2. improve power of computing: this system accessing large data very fast.

3. For error in system: it is distributed system and have different to connect each other. when a node fail. Then data or files already have another other node. Because they make all duplicated copies for data. To remove any fault easily. So, they reduce fault occurrence.

4. provides flexibility for all kind data and it is very important for unstructured information.

5. it is free to use on any machine or computer without license. So, it reduces the cost.

6. more sufficient and scalable to handle by a single user and sample to make more nodes for big data distribution.

So, this is the main challenges which we can solves by this modern distrusted system.

Question2: **Among the trends of Distributed Systems discussed in C1-Lec2, which trend in your opinion will be most dominant in the future and why?** **(4)**

Answer2: There many trends of distributed system discussed in CI-Lec2, but our most dominant opinion in future is **"Pervasive networking and the modern Internet".** Because as we know that digital market, e-commers, smarts devices, and other industrial area. These all area would be shift to modern internet.
There are some important components that the hit trends of future.

> The main reason is sensors, network communication, cloud computing and application. They can reduce the cost. Smart devices also increasing the performance of organization or factories.
> The cloud computing for future trends is very sufficient and fast in the term of data storage. All data will be store on cloud in future. Because they have large space and have strong security.
> The next reason "digital marketing" based on internet of thing. They shift all e-commers on internet.
> All application now moving to based open-source that increasing power of compiler and security for secrete information.

Example:
Now real example of pervasive networking and the modern internet is we conduct all education online like, Classes, Assignments and Exams. They provide very help in future for education in critical situation. Now today covid-19 pandemic is very bad situation and its very difficult to continuous education in university, colleges and school. So only one the modern internet is doing very great role in this time.
So, this point on views we accept that "pervasive networking and the modern internet" will be most dominant in the future.

Question3: **Among the challenges of Distributed Systems discussed in C1-Lec2, which problem in your opinion will accompany distributed systems into the future and why?** **(3)**

Answer3: Among the challenges of distributed system our opinion is "**Security**" will most accompany distributed system into the future. There many reasons behind security problem like, some one is using to target other computers, networks, or websites for certain important purposes, such as information gathering to disrupt network infrastructure and data information systems by using a variety of methods. Now security problem for mobile devices is very challenges. There are some important reasons which consider "**security**" will be accompany distribution systems in the future.
Because there are large challenges have with distributed system in terms of "security" that will be more dangerous for future distributed system. the following are important security problem that why we consider "Security" will be big challenge.

- DoS and DDoS Attacks
- Man-in-the-Middle (MitM)
- Spear Phishing Attacks
- Drive-by Attacks
- Password Attack
- SQL injection Attack

- Cross Site Scripting (XSS) Attack
- Eaves-Dropping Attack
- Malware Attack

So, these are main reason we consider "Security" is very big problem in future.

Question4: **The design of distributed systems can be described and discussed in three ways i.e Physical Model, Architectural Model and Fundamental Model. Describe the example of distributed system in Question1 with respect to these three models.**

(5)

Answer4: In **question1** we explained "Hadoop" is distributed system. the following are three models of this distributed system

Physical Model for Hadoop Distributed system: For Hadoop distributed system required low disk space for storage. Total of 500 GB enough for cluster nodes file storage. 8 to 16 CPU cores with 64 to 96 GB RAM required based on cluster sizes and nodes. Hadoop connectivity is very same and easy on network because it is working on open source web-based application. So, it reduces the cost of application and physical model.

Architectural Model for Hadoop Distributed System: the architecture of Hadoop is a collection of file system like MapReduce engine. It is task tracker for user or client. When they request for a specific file. If the file fail, they rescheduled it. another is HDFS means Hadoop distributed file system for cluster that consists of master and slave. This architecture single NameNode for performs for master cluster that contains all information about operations. They slave contains multiple DataNodes of block for storage data and using data.

Fundamental Model for Hadoop Distributed System: there many fundamental features of Hadoop distributed system some of the following.

- It is open source and available free of cost. That allows user and client to interact easy and performs any task as per requirement. They do not need any license.
- Fast Processing because they provide cluster file distributed system that improve performance of system.
- Provides high availability because when a data node is down. The same data available on another different data nodes. That means reduce faults.

Question5: **What is the purpose of Inter Process Communication (IPC) in distributed systems? Given the choice which protocol out of UDP and TCP will you use for your own distributed system and why?** **(5)**

IPC is very important for the design process of blocks and nan blacks because it reduces the number of functions provided by the blocks. These functions are obtained by communicating with the server through inter process communication. The communication volume is greatly increased. The inter process communication interface usually contains a variable analysis framework structure. These processes can ensure the compatibility between the multi-vector protocols on which the inter process communication model depends.

Both protocols are system of rules using for communication and exchanging information through a medium.

- Our choice is TCP
- It is slower than UDP but TCP is connection oriented while UDP connection less.

- TCP is reliable for sending and receiving messages.

- TCP require ordering by sequence and also more reliable for modern internet distributed system, information sharing. For example. World Wide Web.
- While UDP is sufficient for streaming data.

So, our choice for distributed system is TCP.

Question6: **The following are some of the threats and attacks on Distributed Systems. Provide potential solutions as how may be these threats and attacks be mitigated?** **(8)**
1. **Leakage**
2. **Tampering**
3. **Vandalism**
4. Eavesdropping.

Answer6: **1. Leakage**

There are different solutions to leaks our data. Some of the following.
Identify most important and critical data and identify what data most needs to protect. Then use data loss prevention application and software to protect data from treads and attacks. First target sensitive data based on organization requirement and classify then in main sensitive data and protect it by using data loss prevention software. Next monitor access and activity of traffic. Utilize encryption and lock network with end point.

2. Tampering

We can use web base application firewall. which can provide some protection against parameter tampering. Overall, you can minimize the vulnerability of computers or networks to parameter tampering by implementing strict application security routines and ensuring that they are kept up to date.

3. Vandalism

For this anti-malware software is essential to protect computers, mobile devices, and data from computer damage, viruses, worms, Trojan horses, and other malicious software**.** Other prevention tips are use a strategy for protection like use a security light, look security.

4. Eavesdropping.

By using encryption, we can prevent network-based eavesdropping attacks. For example: if the connection between the client and the server is Web-based, use an HTTPS encrypted connection to prevent eavesdroppers from reading the communication content.