

CSci555 Final Exam Spring 2020

Name: _____

This exam is open book and open note. You may use electronic devices to consult materials stored on the devices, but you may not use them to access material through the net, or for communication during the 120 minutes in which you are completing the exam. You have **120 minutes** to complete the exam. The exam begins at 2PM PDT on May 8th. You must submit the completed exam through email to csci555@usc.edu AND bcn@isi.edu by 4:30PM PDT.

Type your answers in the exam itself using word, or if you prefer a different editor using the text version of the exam that is provided. The filled out exam document will be what you will return to me as described above. In answering the questions, please **TYPE** your answers rather than importing large quantities of text using cut and paste in hopes that the cut and pasted text might include an answer. **Pasted text in your responses will be ignored and you will not receive credit for words included in the pasted text.**

Be sure to include your **name in the exam document**. Ideally, please rename the document to a **file name that includes your name (e.g. csci555-s20-final-FIRSTNAME-LASTNAME)**.

There are **100 points** in all and **3 questions**.

	Q1	Q2	Q3		Total Score
Score					

Complete the following statement:

I, **(replace with your first and last name)** attest to the fact that I completed this exam within the designated time allocated (e.g. in less than two hours), that I did not access external material (e.g. web sites) or use the internet during completion of the exam, and that I completed the exam on my own without accepting or providing assistance to anyone else.

Signed: (type you name here). Date: 5/8/2020.

1. The Cloud (30 points)

a) Cloud Storage (10 points)

Discuss caching and cache consistency mechanisms in consumer cloud storage services such as drop box and google drive. Of the file caching techniques covered in the readings, which do you think is closest to these cloud storage services? What is the basis for your answer?

b) Virtualization (10 points)

Cloud Computing (in the form of Infrastructure as a Service) relies heavily on virtualization. When virtualization is used for cloud deployments a hypervisor runs in place for the Host OS which might be used for a desktop-based deployment. What is the role of the hypervisor in these systems? What are the advantages of using a hypervisor over running a virtualization application such as VMWARE desktop or VirtualBox within Windows or Linux?

c) Multiprocessing (10 points)

Amazon's EC2 (Elastic Compute Cloud) is a service that provides resizable compute capacity in the cloud. To achieve this "elasticity", parts of your application must be capable of running concurrently. Discuss the relationship between parallelism and concurrency in an application. Which of these (parallelism and concurrency) is provided by Amazon's ECC and which by the programmer of the application?

2. **Kernels and Distributed Computing Environments (40 points)**

- a. What makes a distributed system distributed? (15 points)

One aspect of a distributed system that distinguishes it from a network of connected computer is shared state. List the things that constitute shared state in each of the following systems (and why do you consider the state to be shared as compared with transferred or remotely accessed): Andrew, Athena, Amoeba, the World Wide Web.

- b. What should be in the Kernel? (15 points)

The systems we discussed in class differ significantly in terms of what services are provided by the Kernel, which services are left to the application, and which are provide by “servers” that run-in user space. Please list those functions that require at least a minimal level of support within the Kernel (e.g. they would be provided by the Kernel in a microkernel architecture)? Discuss the advantages of a micro-kernel architecture as compared with a monolithic kernel. What are the disadvantages?

- c. Fault Tolerance in Distributed Systems (10 points)

Discuss the fault tolerance techniques employed in certain distributed system services, specifically the domain name system and quorum consensus (weighted voting). Some of these techniques are effective against some kinds of failures, but ineffective against others. Against which kinds of failures are some of the techniques less effective, and why? As a little bit of a hint, ask yourself what it means to **you** for the service to be down.

3. Internet of Things (30 Points)

You have been hired to develop an IoT command center (service hub) for home users. This service hub will provide infrastructure within in the home (such as data storage, a coordinated dashboard, and an interface to remote cellphone apps).

a. Autonomy? (10 points)

An important feature/goal of the service hub is to improve the autonomy of a consumer's home IoT activities. Autonomy is the ability of a system to function in the absence of central infrastructure. Most consumer IoT today depends on cloud infrastructure (for example, the Ring doorbell cameras store recordings in the cloud, many other devices depend on similar infrastructure to communicate with smartphones, etc). In this case you are replacing such "central" infrastructure with an instance that runs on the consumer's home network. In general, for a distributed system, what are the advantages of this kind of autonomy, but also what are the potential problems that are introduced?

b. Simplifying Characteristics (10 points)

What are some of the characteristics (or absence of certain characteristics) regarding access to data in such a system (consumer IoT) which will simplify the correct implementation of the service hub.

c. Naming (10 points)

What are the items to be named in your system, and for each class of items, what do you think is the best kind of naming to be used? Why?