COMPSCI 677: Distributed and Operating Systems

Homework 3

Due Date: None

Instructions: Homework 3 is last year's mid-term exam that is handed out as a practice test. Please use this exam as a way to understand the type of questions to expect on the mid-term and also to practice for the exam.

This homework will not be graded. Hence there is no need to turn it in.

Name: **Student Id:**

COMPSCI 677: Distributed and Operating Systems

Exam 1

March 22, 2016

General instructions:

- This examination booklet has 13 pages.
- Do not forget to put down your name and student number on the exam books.
- The is a closed book, closed notes exam.
- Explain your answers clearly and be concise. *Do not write long essays.*
- Be a smart test-taker. If you get stuck, move on to the next question.
- State any assumptions you make clearly.
- Good luck.

---

1. **Short answer questions** (25 points)
   Answer the following questions in a few sentences each.

   (a) (5pts) Explain the difference between a network operating system and a middleware-based operating system.

(b) (5pts) *Replication transparency:* Replication transparency hides the presence of replicas from the end-user. Consider a web site (say `http://www.foo.com`) that is replicated on two identical servers. What mechanism would you use to provide replication transparency for this replicated web site?

(c) (5pts) *Threads:* Suppose that you decide to implement a multi-threaded on your brand-new dual core laptop using threads. You are given a choice of a user-level threads library or a lightweight process-based threads library. Which one would you choose to make the best use of your laptop resources? Explain your choice in brief.

(d) (5pts) *Naming* Consider an iterative name resolution in DNS where the client caches each resolved level in the name space. Why is it desirable to cache entries from higher up in the DNS hierarchy for longer time periods?

(e) (5pts) *Load balancing:* Load balancing in distributed systems is useful only when system as a whole is moderately loaded. Explain why.

2. **Virtualuzation** (15 points)

(a) (8pts) Explain in brief why normal Type 1 hypervisors can not be implemented on older Intel (x86) hardware. Next explain how the use of paravirtualization enables Type 1 hypervisors to be used on older hardware.

(b) (7pts) Explain why OS level virtualization is more efficient that Type 2 hardware virtualization. Also argue why OS level virtualization is less secure than hardware virtualization.

3. **Process and Code migration** (10 points)

(a) (4 pts) What is the difference between strong mobility and weak mobility? Are Java applets an example of the former or the latter?

(b) (6pts) Consider the migration of a process that is accessing an local database on a machine. What type of resource is a database (i.e., unattached, fastened or fixed) and why? How should we provide access to the database after the process migrates?

4. **Communication in distributed systems** (15 points)

(a) (4pts) Pick one of the following choices and explain in one or two sentences.
It is { *sometimes* | *never* | *always* } possible to write an event-based server using blocking communication primitives.

(b) (5pts) Assume that you are watching a video on your smartphone and the cellular data connection is limited to 1Mb/s. How can we use the token bucket mechanism at the video server to ensure that the transmitted video stream is limited to 1 Mb/s?

(c) (6pts) From a communication standpoint, classify the following applications as *persistent* or *transient* and *synchronous* or *asynchronous*: (i) email, (ii) online chat, (iii) bulletin boards such as Yahoo Groups. Explain your answer in brief.

5. **Canonical Problems in Distributed Systems** (25 points)

(a) (10 pts) Explain the principle behind GPS in brief. Using your answer, explain why it is *necessary* to synchronize a receiver's clock in order to find the receiver's location? Also explain why we need at least 4 satellites (at a minimum) to locate a receiver using GPS?

(b) (10pts) *Logical clocks:* Vector clocks have the property that if $V_a < V_b$ for any two events a and b then $a \rightarrow b$. Provide an informal proof of this property.

(c) (5pts) *Mutual exclusion* Why is it not feasible to use timeouts to detect the loss of a token in the token ring mutual exclusion algorithm?

6. **Multi-tier applications** (10pts)
Consider a two-tier web site consisting of the Apache HTTP server at Tier 1 and a database server at Tier 2. Assume that the two tiers run on independent machines and collectively implement an online book store, where users can *browse* the book catalog and *buy* books. Each *browse* request undergoes HTTP processing at tier 1 and then issues a search query at the database tier, while a *buy* request undergoes HTTP processing and issues a buy transaction at the database. Suppose HTTP processing at tier 1 takes 2ms, while the search queries take 5ms, and transactions take 10ms. Under normal circumstances, the web site sees a workload of 100 requests /second, with 80% browse requests and 20% buy requests. Will the system see a bottleneck if the incoming workload doubles to 200 requests/s with a mix of 60% browse and 40% buy requests? Assume that the workload is CPU-intensive at both tiers.