香 港 中 文 大 學
The Chinese University of Hong Kong

Course Examinations 2017-18 (1st term)

Course Code & Title : **CSCI4180 Introduction to Cloud Computing and Storage**

Time allowed : 2 **hours** 0 **minutes**

Student I.D. No. : .... **Seat No. :** ....

You have two hours to complete the exam. All questions are to be completed. The full score is **100 points**. You are allowed to bring two A4-sized cheat sheets, with both sides printed or written, and use an electronic calculator approved by the University. Other electronic equipments are prohibited. Write down your **student ID** and **seat number** in the answer book. Write **all** the answers in the answer book. Write **neatly**. Anything that is unreadable will receive zero point.

Questions

1. **(40%, 4% each) Quick Questions.** Keep your answers short and precise.

   (a) What is the vendor lock-in problem in cloud computing? Explain one way that can relieve the vendor lock-in problem.

   (b) State one similarity and one difference between a private cloud and a virtual private cloud.

   (c) State one similarity (e.g., in terms of architecture) and one difference (e.g., in terms of functionalities) between MapReduce and YARN.

   (d) MapReduce emphasizes locality. What is it and why is it useful?

   (e) In practice, the rebalancing algorithm in BigTable does not work perfectly. Why?

   (f) Dynamo uses consistent hashing to support incremental scalability. Consider the case where the hash ring is in the form: Node 10 $\rightarrow$ Node 18 $\rightarrow$ Node 26 $\rightarrow$ Node 35 $\cdots$. If Node 18 leaves the system, state which objects and where the objects will be relocated.

   (g) Dropbox uses server-side deduplication instead of client-side deduplication. State one advantage and one disadvantage.

   (h) Suppose that we re-configure Facebook Haystack to store **four** images for each photo and use **five** replicas for better reliability. What is the number of I/Os that needs to be issued for writing a photo? Derive your answers.

   (i) Explain what the PERSISTENT_SEQUENTIAL flag means in Zookeeper.

   (j) One way to reduce tail latency is to send hedged requests. What does it mean? Is it a proactive or reactive approach?

2. **(18%) Design Trade-offs of HDFS.**

   There are different ways to configure HDFS to trade between performance and reliability. By default, HDFS uses 128 MB block size, three-way replication, pipelined replication, and rack-aware block placement. Explain how each of the following configurations affects the *write* performance of HDFS compared to the default configuration. For each answer, choose one of the three options: improved, degraded, or unchanged. Make sure you provide the correct explanations for your chosen option in order to gain the full mark of each answer.

   Note: There may be more than one correct option for each answer, yet whether the option is correct depends on your explanation. You only need to pick one option for each answer.

   (a) Set the block size to 128 KB.

   (b) Use five-way replication.

   (c) Let the client distribute replicas across different nodes.

   (d) Put all replicas in one rack.

   (e) Use inline deduplication.

   (f) Use Amazon Dynamo's configuration: $(N, R, W) = (3, 2, 2)$.

Go to Next Page...

3. **(10%) MapReduce.**

   Figure 1 shows the pseudo-code of the MapReduce algorithm on computing the co-occurrence counts for each word. We now want to modify the Map function to support *in-mapper combining*. Write the pseudo-code of the revised Map function, while keeping its input interface unchanged.

```
class Mapper
    method Map(docid a, doc d)
        for all term w ∈ doc d do
            H ← new AssociativeArray
            for all term u ∈ Neighbors(w) do
                H{u} ← H{u} + 1            ▷ Tally words co-occurring with w
            Emit(Term w, Stripe H)
```

```
class Reducer
    method Reduce(term w, stripes [H_1, H_2, H_3, ...])
        H_f ← new AssociativeArray
        for all stripe H ∈ stripes [H_1, H_2, H_3, ...] do
            Sum(H_f, H)                    ▷ Element-wise sum
        Emit(term w, stripe H_f)
```

Figure 1: Q3: MapReduce Pseudo-code.

4. **(10%) PageRank Algorithm.**

   Figure 2 shows a web graph containing only five webpages denoted by n1, n2, n3, n4, and n5. Our goal is to compute the PageRank of each page using MapReduce. Show how you can solve for the PageRank values for all webpages using MapReduce. We assume that there is no random jump. In your answers, you need to show (i) how you initialize the PageRank values, and (ii) the outputs of the Map and Reduce functions of the *first two* iterations. Feel free to define the notation that is necessary.

   Note: Please present your answers up to four decimal places. To keep things simple, you can assume 1/3 = 0.3333, and 2/3 = 0.6666 or 0.6667. Don't worry too much about the round-off error.

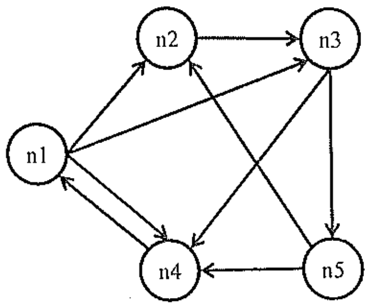

Figure 2: Q4: Topology.

Go to Next Page...

5. **(22%) Cloud storage.**

We develop a deduplication system for cloud storage. We use Rabin fingerprinting as the chunking algorithm. Let us assume that we store a file with the following byte sequence $\{t_1, t_2, \cdots\}$, where t_i denotes the i^{th} byte. Then we operate on the byte sequence and generate Rabin fingerprints based on the following formula:

$$p_s(d,q) = \begin{cases} (\sum_{i=1}^{m} t_i \times d^{m-i}) \ mod \ q, & s = 0 \\ (d \times (p_{s-1} - d^{m-1} \times t_s) + t_{s+m}) \ mod \ q, & s > 0 \end{cases}$$

for some parameters m, q, and d.

(a) (1%) Suggest one workload for which deduplication is useless.

(b) (2%) Explain one drawback of setting $m = 2$

(c) (2%) Explain one drawback of setting $q = 4000000$.

(d) (2%) Explain one drawback of setting $d = 127$.

(e) (3%) Suppose that we insist on setting $m = 2$ and $q = 4000000$ (we know they are bad choices). We modify the chunking implementation such that an anchor point is generated if $p_s == 0$ or $p_s == 1$. What is the average chunk size? Explain your answers.

(f) (3%) Some measurement studies show that fixed-size chunking is as good as (or even better than) Rabin fingerprinting for virtual machine images. State one reason.

(g) (3%) We now put the full fingerprint index on disk and deploy a Bloom filter to save the disk I/O. Explain which types of disk I/O can be saved through the Bloom filter.

(h) (3%) Extreme Binning cannot achieve exact deduplication (i.e., a duplicate chunk cannot be deduplicated). Explain how this could happen.

(i) (3%) Explain why practical deduplication systems store chunks in units of containers rather than store chunks directly.

Hope you have a fruitful winter break!

-End-