

Homework

Zhijia Chen

October 17, 2018

1 HW1

5. Hash table:

Can I use hash key as an id? Specify the reason.

No. A hash function is a multiple-to-one mapping which means for a certain item, its hash value is not unique, while an id by its definition should be unique.

How to find the difference between two list of strings?

For each pair of corresponding strings in the two list, first compare their hash values, if their hash values are different, then they are different. Otherwise, since having the same hash value does not guarantee the two strings are identical, we still need to compare the two strings character by character.

How to handle the duplication of the hash code?

We can do rehash by using new hash functions. We can also use open hash addressing such as chaining, linear probing.

How to deal with an shared hash table being used on the web server?

1. When updating the harsh table, lock the table. Each time, there are only one request can update the table. Save space because we only need to keep one copy of the table. But waste time, other requests need to wait.

2. Each request works on its own copy of the table, as long as they are not updating the same entry, the changes distributed among those copies can be merged. But if there multiple requests are updating the same entry, a lock is still needed. It saves time but wastes space.

2 HW2

1. What are the transport layer protocols used for live video, le transfer, DNS and email, respectively?

UDP, TCP, UDP, TCP

2. Consider a TCP implementation with the Additive Increase (linear) and Multiplicative Decrease (AIMD) algorithm, ignoring the rst phase where the ssthreshold is detected, assume the window size at the start of the slow start phase is 1 MSS and the ssthreshold at the start of the rst transmission is 8 MSS.

Assume that a timeout occurs during the 4th transmission. Find the congestion window size at the end of the 8th transmission.

Suppose we are using the TCP reno implementation.

Phase 1: slow start. Double the window size after each successful transmission. Window size of 1st transmission: 1 MSS. Window size of 2nd transmission: 2MSS. Window size of 3rd transmission: 4 MSS. Window size of 4th transmission: 8 MSS. A timeout occurs, set ssthreshold to half of current window size, i.e., 4 MSS, and set window size to 1. Window size of 5th transmission: 1 MSS. Window size of 6th transmission: 2 MSS. Window size of 7th transmission: 4 MSS. The window size reaches the ssthreshold, leaving Phase 1 and entering phase 2.

Phase 2: congestion avoidance. Increase window size by one after each successful transmission. Window size of 8th transmission: 5 MSS.

So by the end of 8th transmission, the congestion window size is 5 MSS.

3. UDP: Look into the following IP packet enclosing a UDP datagram. UDP checksum includes three sections: a pseudoheader, the UDP header, and payload – the data coming from the application layer. Pseudoheader and padding are only used to calculate checksum. After checksum is calculated, they are dropped. The figure. 1 is not a complete IP packet. It has the major info of the IP packet that has the UDP datagram and these info are what UDP checksum need to calculate. Now, assume we have Source IP field: 10011001 00010010 00001000 01101001

The sum of the first 16 bits and the last 16 bits of the source IP: 10100001 01111011

Destination IP field: 1010 1011 0000 0010 0000 1110 0000 1010

Protocol field: 0001 0001

The length of the pseudoheader is 12 bytes.

The length of the head is 8 bytes.

The length of the payload after padding is 8 bytes.

UDP total length (28 bytes) field: 0000 0000 0001 1100

Source port field: 0000 0100 0011 1111

Destination port field: 0000 0000 0000 1101

Data field: 01010100 01000101 01010011 01010100 01001001 01001110 01000111 00000000