CSCI-1380: Distributed Computer Systems
Homework #5
Assigned: 03/31/2021
Due: 04/04/2021

1 System Design

In general distributed systems are about understanding and picking appropriate design choices for your scenarios. Ultimately, your knowledge of novel design choices and interesting combination will be based on internal documents or publicly released descriptions of existing techniques.

Recall, HW#1 and HW#5 are both related HWs! Both of these assignments are not technical in nature. They are meant to provide you with a mechanism for reading and understanding internal design docs and external system descriptions which you may have to deal with as a new member of a distributed systems team.

For HW#1, you read the original 6-page Cassandra paper from Facebook. When you read this paper, many terms and concepts were potentially unfamiliar. However, now we have cover many of these terms.

Now for HW#5, you will revisit your responses to HW#1: for each term that you identify as being unfamiliar, specify if you now understand the term or if you guess correctly.

What to do: Re-read the PDF that you turned in for HW#1. Create a table with three columns:

1. Column 1: The term you were unfamiliar with.
2. Column 2: Yes, No – do you know understand the term.?
3. Column 3: Yes, No – did you guess correctly?

Moreover, now that you understand significantly more about distributed systems principles, you may find some of the design choices that were made by the developer of Cassandra to be odd or peculiar. Re-read section 1, 3, 5 of the Cassandra paper. Identify any design choices, e.g., Consistent Hashing, that you do not agree with or that you find surprising.

What to Turn in: You should create (and submit) a PDF that includes your answers from HW#1 and on a separate page include the answers from HW#5 (i.e., table and paragraph).

1. The table described above.
2. A paragraph describing how the additional knowledge has changed your understanding of the Cassandra system.
   (a) In particular, list out the design choices that you think are questionable (sub-optimal, inefficient).
   (b) for each design choice, that you question, why do you question them?
   (c) for each design choice that you question, what is an alternative technique that you would use?

Note: Given the varying levels of expertise, we do not expect everyone to have the same list of terms. Moreover, given some of the private questions at least one or two of you are familiar with material in the first third of the class.