Exercise 1

1. In a hard disk drive (HDD), the average seek time is 12 ms, the rotation delay is 4 ms, and the transfer rate is 4MB/sec. For simplicity, we assume in this question only 1MB equals 1000KB.

   (a) What is the seek time delay?

   (b) What is the rotation delay?

   (c) What will be the disk access time for a transfer size of 8MB? What will be the disk access time for a transfer size of 8KB?

   (d) In a solid state drive, what will be the disk access time for a transfer size of 8MB when the transfer rate is 4MB/sec? Is an SSD faster than an HDD for the same amount of data transfer (Assuming the base sequential data transfer rates are the same for the given two drives.)? Why?

2. There are two different machines where machine A has a smaller cache with an average 50% cache hit ratio (H) and the other machine (machine B) has a much larger cache with an average 90% cache hit ratio. However, the memory access time of machine A is 100C and the memory access time of machine B is 400C (i.e., memory access in machine A is faster than memory access in machine B), where C is the cache access time. Which machine has an overall faster effective memory access time?

Supplementary details on NoSQL databases.

- *Key-value storage*: A key-value database stores data as a collection of key-value pairs where a key serves as a unique identifier. All the accesses to the database are done via the key. Both keys and values can be complex.
- *Document storage*: Flexible for storing different kinds of documents, where they may not all have the same sections.
- *Column-based NoSQL databases*: These types of databases are a hybrid of RDBMSs and Key-Value stores. Values are stored in groups of zero or more columns, but in Column-Order (as opposed to Row-Order). They are fast if you are not accessing individual records but need columns.
- *Graph storage*: Graphs capture connectivity between entities. Searching and traversing by relations are very fast in such structures. A graph is a structure amounting to a set of objects (called vertices) where some pairs of objects are connected/related in some sense. A connection is called an edge. The links can be material or non-material:
  - Links between two streets can be road junctions.

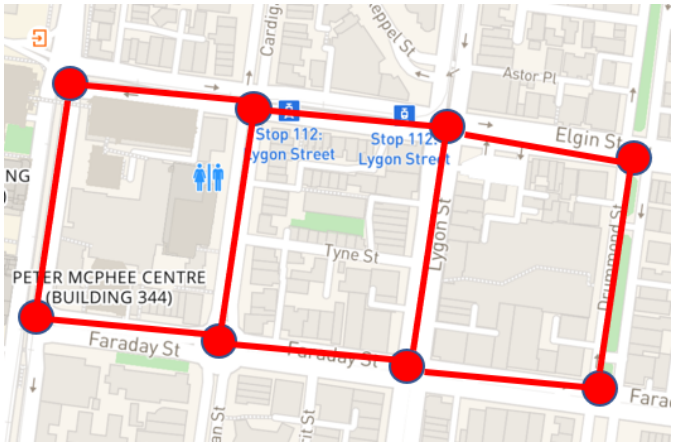

Figure 1: Streets Connecting Road Junctions in Parkville, Victoria.

  - Links between people can represent their Facebook connections (non-material links)

4. Discuss example applications of different types of NoSQL databases.

5. Consider the different scenarios below and discuss which database architecture is the most suitable choice and why:

   - FriendBook is a new startup app that will launch its operation soon. They have only one office without much budget right now, but they are expecting high growth in the scale of millions of users across the globe in a couple of years. Which of the following database architecture is the most suitable choice for this scenario?
     - Cloud storage
     - World wide web
     - Distributed database
     - Centralised database
   - FriendBook is a new social network site that will launch its operation soon. They have offices in many major cities in USA. They need a database that can handle millions of users across the globe. For preserving privacy and security, they need their own data storage system, which is not shared or owned by any other company. Which of the following database architecture is the most suitable choice for this scenario?
     - Cloud storage
     - World wide web
     - Distributed database
     - Centralised database