two sigma：

basic problems:
1.difference between thread and process
 1） processes are typically independent, while threads exist as subsets of a process

 2） processes carry considerably more state information than threads, whereas multiple threads within a process share process state as well as memory and other resources

 3） processes have separate address spaces, whereas threads share their address space

 4） processes interact only through system-provided inter-process communication mechanisms

 5） Context switching between threads in the same process is typically faster than context switching between processes.

 In computing, a process is an instance of a computer program that is being executed. It contains the program code and its current activity. Depending on the operating system (OS), a process may be made up of multiple threads of execution that execute。

instructions concurrently.

1.how do threads communicate
        1) shared memory or object
          2) message queue (http://en.wikipedia.org/wiki/Message_queue)

https://developer.apple.com/library/ios/documentation/Cocoa/Conceptual/Multithreading/AboutThreads/AboutThreads.html
1.(ipc) how do inter－process communicate
http://en.wikipedia.org/wiki/Inter-process_communication

1.hashmap hash collisions :linear probing hashtable

2.Design pattern 的 coding detail

定义：a general reusable solution to a commonly occurring problem within a given context

6. hashtable的insertion和deletion还有search的具体实现；

7.throughput & latency

Latency is the time required to perform some action or to produce some result. Latency is measured in units of time -- hours, minutes, seconds, nanoseconds or clock periods.

Throughput is the number of such actions executed or results produced per unit of time. This is measured in units of whatever is being produced (cars, motorcycles, I/O samples, memory words, iterations) per unit of time. The term "memory bandwidth" is sometimes used to specify the throughput of memory systems

example：

http://community.cadence.com/cadence_blogs_8/b/sd/archive/2010/09/13/understanding-latency-vs-throughput

4, use hash table to store data, but there is much more data than the machine's RAM, how to deal with that?
　答案：add one more machine, rehash and reconstruct the hash table

5, A application involves with multiple machines and it is slow, figure out why
　答案：can use matrix to measure the latency and throughput of each machine.

coding:
1.external sort 1tb data, 1gb memory sort~~
2.tree: 给你n个stock的price，然后你要找到每天的median, runtime
　　　n(logn +logn
3.等差数列求和公式及其推到，估计这个大家都会，下一个平方数列求
　　　和公式及其推倒
http://zhidao.baidu.com/question/16413317.html
http://zh.wikipedia.org/wiki/%E7%AD%89%E5%B7%AE%E6%95%B0%E5%8

[8%97](#)

立方数列求和

等比数列求和

 等差数列**arithmetic progression** (AP) or **arithmetic sequence**

4.下面一个算法题是给你一个很大的数据流如何**track median**，相信看过**cracking the interview**的都能答上来，两个**heap** 一个**max** 一个**min**接下来又问，如果我想得到**(n/10)th element**怎么办。一时糊涂脑袋没反应过来，就说加几个**heap**分段**track**。放下电话脑袋才转过来。 其实**median**就是求**(n/2)th element**。 只要改变**min heap** 和**max heap** 的**size**比例就可以了。让**max heap**的大小是**min heap**的**1/9,**就可以了

5.clone graph bfs 和**dfs**

http://www.1point3acres.com/bbs/forum.php?mod=viewthread&tid=135020&extra=page%3D1%26filter%3Dsortid%26sortid%3D311%26sortid%3D311

http://www.1point3acres.com/bbs/forum.php?mod=viewthread&tid=108193&highlight=sigma

http://www.1point3acres.com/bbs/forum.php?mod=viewthread&tid=110665&extra=&highlight=sigma&page=1

http://www.1point3acres.com/bbs/thread-42043-1-1.html