# System Design

1. Latency:
   * Latency, to put it simply, is basically how long it takes for data to traverse a system. And more specifically, how long does it take for data to get from one point in a system, to another point in the system.
   * It might refer to the different kinds of the context in the system. For instance:
     + If you are referring to the network then how long does take for one request to go from a client to server and then back from server to the client.
     + If you are reading a piece of data for memory/disk then the time to read the data is referred as a latency
2. Throughput
   * Throughput is an amount work a machine can perform in a given amount of time.
   * We're really referring to how much data can be transferred from one point in a system to another point in a system, in a given amount of time.
   * It is measured in gigabits per second / megabits per second/kilobits per second. 1 gbps network can support one gigabits per second