

- 1) Distributed systems are built to achieve greater fault tolerance and redundancy, availability, scalability, better resource sharing and utilization, and improved performance.
- 2)
  - a) Metcalfe's Law: The value of a telecommunication network is proportional to the square of number of users in the system.
  - b) Moore's Law: The number of transistors on a microchip doubles approximately every two years, leading to an exponential increase in computing power and a decrease in relative cost.
  - c) According to Moore's law it is safe to say that for every few years, each new node added to or replaced within a distributed system would be capable of more computation, meaning the whole system performs better. Hence, every few years a network/distributed system becomes more valuable and efficient with not much increase in its cost. Both Moore's Law and Metcalfe's Law drive the evolution of distributed systems, making them more efficient, scalable, and valuable over time.
- 3) Openness of a distributed system is the ability of it to be extended and improved both in terms of hardware and software. This enforces a plug-and-play kind of environment, ensuring that new components can integrate effectively with existing ones without causing compatibility issues.
  - a) Systems should conform to well-defined interfaces.
  - b) Systems should easily interoperate.
  - c) Systems should be easily extensible.
  - d) Systems should support portability of applications.
- 4)
  - a) Average number of requests in an M/M/1 server are:  $U/(1 - U)$ , where  $U$  is the utilization of the server. Given utilization = 0.9. Therefore, the expected number of requests in the server are 9.
  - b) Probability of having  $k$  requests in the system is:  $(1 - U)U^k$ . Therefore, the probability of having 3 requests in the system is 0.0729.

5)

- a) Node failures: Nodes abruptly stop functioning.
  - i) Can be guarded by implementing effective failure detection mechanisms, such as heartbeat signals, but can be complicated by network latency and asynchrony.
- b) Network failures: Messages are lost in transit.
  - i) Designing systems that remain operational despite partitions requires careful consideration of consistency and availability trade-offs.
- c) Performance failure: Operations do not complete within the expected time frame.
  - i) Network latency and node capacity have high correlation with such failures. Guarding those would help in reducing performance failures.
- d) Cascading failure: Failure in one component leads to subsequent failures in others, potentially collapsing the entire system.
  - i) Services have to be loosely coupled to prevent such failures.
- e) In general guarding against such failures in a distributed system is inherently challenging because of the complexity and intricate dependencies.
- f)

6) TCP vs UDP:

- a) TCP has a handshake protocol overhead to establish a connection. UDP is faster but does not have a dedicated end-to-end session between client and server.
- b) TCP is reliable through different mechanisms like retransmission and sequence control for data integrity and order. UDP inherently does not provide such reliable mechanisms.
- c) Continuous stream of data is sent over in TCP, whereas in UDP data is transmitted in-terms of datagram packets.

- d) I would choose a TCP communication protocol if secure sequential data transmission is essential for my application, where reliability is a greater priority than speed.
- e) I would choose an UDP communication protocol if real-time speed has greater priority.

7) Caching vs Replication:

- a) Purpose of caching is to reduce the latency while accessing data. On the other hand, the purpose of replication is to improve fault tolerance.
- b) Frequently requested/accessed data is pulled into local cache registers. Whole database contents are synchronized proactively and copied/pushed to create a replica.

8) In reality networks are susceptible to various issues such as packet loss, disconnections, and hardware failures and there is always some delay in data transmission, which can vary based on distance, network congestion, and other factors. Network bandwidth will also be finite with non zero transport cost and can be a limiting factor, especially in high-traffic scenarios. There will always be the possibility of security threats, including eavesdropping, man-in-the-middle attacks, and data breaches that have to be taken care of. Changes in network topology is also expected to be very common due to factors like hardware failures, maintenance, or scaling efforts. Also, it is efficient only to have multiple administrators with different need-based access-control to handle a complex distributed system. It is to be noted that large scale coherent systems can not be made completely with homogeneous systems. Given all the reasons above, each one of the eight pitfalls of designing distributed systems identified by Deutsch is actually a pitfall, which are common misconceptions that can lead to flawed system designs.

9)

- a) Issues in parameter passing:
  - i) Passing by value and by reference: While calling remote procedures, we need to either send the actual data object or the reference of the object to the computing node. When the data is huge, sending it over to another

computing node is very inefficient. On the other hand, if there is no shared memory between the nodes, sending references would potentially give wrong results. The representation of data can also be different in each node, creating issues with parameter passing.

- b) Issues in mitigating failures:
  - i) Network reliability, server availability, idempotency and error propagation are few of the major issues under this aspect.
- 10) In a distributed wireless sensor network, I would first put an end to replication transparency to maximize its availability. Replication transparency ensures that multiple copies of data across different nodes appear as a single entity to users and allowing applications to be aware of data replication can enable more efficient synchronization strategies, reducing unnecessary data transmissions and conserving energy.