CSI3344 Distributed Systems

Workshop Solution 02

Q1.   If a client and a server are placed far apart, we may see network latency dominating overall performance. How can we tackle this problem?

A: It really depends on how the client is organized. It may be possible to divide the client-side code into smaller parts that can run separately. In that case, when one part is waiting for the server to respond, we can schedule another part. Alternatively, we may be able to rearrange the client so that it can do other work after having sent a request to the server. This last solution effectively replaces the synchronous client-server communication with asynchronous one-way communication.

Q2.   What is a three-tiered client-server architecture?

A: A three-tiered client-server architecture consists of three logical layers, where each layer is, in principle, implemented at a separate machine. The highest layer consists of a client user interface, the middle layer contains the actual application, and the lowest layer implements the data that are being used.

Q3.   What is the difference between a vertical distribution and a horizontal distribution?

A: Vertical distribution refers to the distribution of the different layers in a multitiered architectures across multiple machines. In principle, each layer is implemented on a different machine. Horizontal distribution deals with the distribution of a single layer across multiple machines, such as distributing a single database.

Q4.   Consider a chain of processes P_1, P_2, ..., P_n implementing a multitiered client-server architecture. Process P_i is client of process P_{i+1}, and P_i will return a reply to P_{i-1} only after receiving a reply from P_{i+1}. What are the main problems with this organization when taking a look at the request-reply performance at process P_1?

A: Performance can be expected to be bad for large n. The problem is that each communication between two successive layers is, in principle, between

two different machines. Consequently, the performance between P_1 and P_2 may also be determined by n - 2 request-reply interactions between the other layers. Another problem is that if one machine in the chain performs badly or is even temporarily unreachable, then this will immediately degrade the performance at the highest level.

<u>END OF THE WORKSHOP SOLUTION</u>