Innopolis University

SYSTEM AND NETWORKING ENGINEERING

Distributed Systems

READING QUESTIONS 2

Architectures. Processes.

Student	**ID**
Grebennikov Sergey	47611

Lecturer
Konstantin Urysov

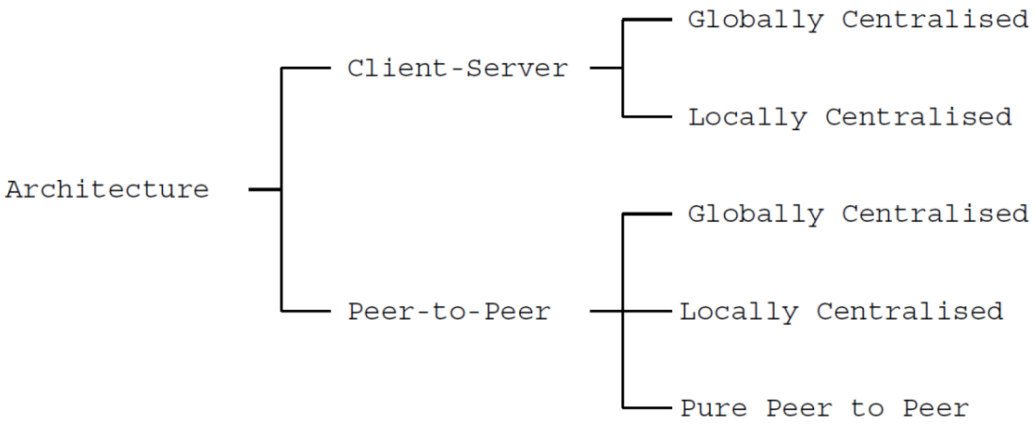
October 27, 2017

Readings:

- **[DS]** Chapter 2, 3. Distributed systems: principles and paradigms, Andrew S. Tanenbaum, Maarten Van Steen
- **[TAX]** Section 3.2 Storage Architecture; A Taxonomy of Distributed Storage Systems, Martin Placek and Rajkumar Buyya

Questions:

1. According to **[DS]**, there are three types of system architectures: centralized, decentralized and hybrid. And according to **[TAX]** there are five of them:



   How do you think to which type of architecture by **[DS]** should we correspond the types that were proposed in the **[TAX]** and why?
   Answer:
   Centralized organization in [DS] should correspond to Globally Centralised Client-Server types of a system architecture in [TAX] since client requests are centrally processed by a server. Decentralized organization in [DS] should correspond to Pure Peer-to-Peer and Globally Centralised Peer-to-Peer types of system architecture whereas the Hybrid organization in [DS] should correspond to Locally Centralised Peer-to-Peer and Locally Centralised Client-Server types in [TAX] since they combine architectural features.

2. What is the difference between a process and a thread?
   Answer:
   A process is a program in execution whereas a thread is a flow of execution through the process code, with its own program counter, system registers and stack.

3. What is an asynchronous (non-blocking) I/O operation?
   Answer:
   This operation appears when a I/O request occurs. In this case, the thread will schedule this operation, which will be interrupted later by the operating system.

4. Does it make sense to use threads on a single-core CPU?
   Answer:
   Yes, since a multithreaded solution has several advantages, such as improving performance and better the structure even on a single-core processor:

   - Starting a thread is much cheaper than starting a new process
   - Having a single-threaded server prohibits simple scale-up to a multiprocessor system
   - Hide network latency by reacting to next request while previous one is being replied
   - Using simple, well-understood blocking calls simplifies the overall structure
   - Multithreaded programs tend to be smaller and easier to understand due to simplified flow of control

5. VM images such as AMIs can be quite big. How does this impact cloud providers that have many customers creating many different virtual machines all the time?
   Answer:
   Using a distributed system of virtual servers in different regions can effectively deal with such problems.

6. Are Web servers stateless or stateful?
   Answer:
   Web server is stateless. It merely responds to incoming HTTP requests, which can be either for uploading a file to the server or (most often) for fetching a file.

7. What is the difference in request dispatching for local-area and wide-area clusters? At what point will we need a redirection policy?
   Answer:
   For local-area clusters, there is a fixed switch (such as TCP-handoff) while in wide-area clusters it is necessary to decide which switch to select, which requires additional costs. A redirection policy allows to decide which server should handle the client request in wide-area cluster implementation.

8. Wide-area redirection requires a method for measuring the distance between two IP addresses. Think of two different methods and discuss pros and cons.
   Answer:
   Several redirection mechanisms are possible: use of a Domain Name System (DNS) or a distributed server. In the case of DNS, there may be a huge additional communication cost, as the local DNS server is often not that local. In addition, it may happen that the DNS server that is deciding on which IP address to return, may be fooled by the fact that the requester is yet another DNS server acting as an intermediate between the original client and the deciding DNS server. The basic idea behind a distributed server is that clients benefit from a robust, high-performing, stable server.

9. What problems will you need to solve to allow live migration of virtual machines between different wide-area clusters?
   Answer:
   One of the potential problems with the migration of virtual machines is that it can take a considerable time. This in itself should not be bad if services running on a portable virtual machine can continue to work.

10. According to Fuggetta (Note 3.11) there are three segments in a process. Which segment do you think is typically more difficult to migrate?
    Answer:
    Probably it will be an execution segment since it contains data that is highly dependent on a specific implementation of the underlying operating system. For example, it may rely on information normally found in the operating systems process table.