TECHNISCHE UNIVERSITEIT EINDHOVEN
Faculteit Wiskunde en Informatica

Examination Architecture of Distributed Systems (2II45),
on Monday April 14, 2014, from 09.00 to 12.00 hours.

Before you start read the entire exam carefully. Answers to all questions must be motivated and stated clearly. For each question the maximum obtainable score is indicated between parentheses. The total score sums up to 25 points. This is a closed book exam, i.e., you are not allowed to use books or other lecture material when answering the questions.

1. (2 points) Give the basic ingredients of an architectural description as specified by the ISO/IEC/IEEE 40210 standard. Illustrate your discussion with an appropriate UML model.

   Answer. For a UML-diagram see the architectural description (meta-)model (slide 40 of the introduction slide set). Of that diagram at least the boxes labelled "System of Interest", "Stakeholder", "Concern", "Viewpoint", "View", "Model", and "Architectural Description"should be present, and of course the relationships that hold between them.

2. (1.5 point) Describe the difference between broadcast, multicast and anycast and give an application of all three.

   Answer.

   broadcast message without destination reaching all reachable neighbors.
   Example: DHCP, other service discovery

   multicast message sent to a given set of destinations.
   Example: IP multicast, e.g.in UPnP.

   anycast message delivered to one destination out of a set of possible ones.
   Example: DNS requests to root servers.

3. (1 point) You are designing a client-server system and have to decide between a thin and a fat client. What criteria do you use to make your decision?

   Answer. Two important criteria are resource availability at the client side and network bandwidth between client and server. If the client has few resources, then a thin client is preferable. Then the client usually only needs to support an interface and the application logic is at the server side On the other hand, a thin client usually requires more network bandwidth between client and server.

4. For name spaces that are distributed across multiple name servers one distinguishes between iterative and recursive name resolution.

(a) (0.5 point) Explain the difference.

(b) (0.5 point) Give an argument in favor for iterative resolution.

(c) (0.5 point) Give an argument in favor for recursive resolution.

Answer. See TvS "Implementation of Name Resolution", pp 205 – 209.

5. (2 points) Describe the peer-to-peer (P2P) architectural style using the appropriate vocabulary, name the concepts involved, give a motivation for its usage and mention typical usage.

   Answer. See slide 13 of the slide set on architectural style.

6. Consider the Chord scheme for DHTs. Assume an 6-bit identifier space, and assume there is a node set N with 6 nodes whose identifiers are given by $id(N) = \{3, 8, 23, 44, 51, 53\}$.

   (a) (1 point) Give the finger table of node 51.
   Answer. Each table has 6 entries. For $1 \leq i \leq 6$, the i-th entry equals $succ(51 + 2^{i-1})$. Hence we obtain:
   $FT_{51}[1] = FT_{51}[2] = 53$, $FT_{51}[3] = FT_{51}[4] = FT_{51}[5] = 3$ and $FT_{51}[6] = 23$.

   (b) (1 point) What is the average number of hops taken to look up a resource, i.e. resolve a key, starting from node 51?
   Answer. For any key k the number of hops is between 0 and 3. To be precise:
   For $45 \leq k \leq 51$, it takes 0 hops: $(succ(k) = 51)$
   For $k = 52, 53, 3, 23$, it takes 1 hop: $(51 \rightarrow succ(k))$
   For $54 \leq k \leq 2 (= 66 \mod 64)$, it takes 2 hops: $(51 \rightarrow 53 \rightarrow succ(k))$
   For $4 \leq k \leq 8$, it takes 2 hops: $(51 \rightarrow 3 \rightarrow succ(k))$
   For $9 \leq k \leq 22$, it takes 3 hops: $(51 \rightarrow 3 \rightarrow 8 \rightarrow succ(k))$
   For $24 \leq k \leq 44$, it takes 2 hops: $(51 \rightarrow 23 \rightarrow succ(k))$
   So the average becomes ((7*0)+(4*1)+((13+5+21)*2)+(14*3))/64 = 124/64

7. (2 points) The viewpoint library of Rozanski and Woods differs from that of Kruchten. In particular, it contains an *Information Viewpoint* which is not present in Kruchten's 4+1 viewpoints. Indicate the topics addressed in this viewpoint, and at least two stakeholders including their concerns pertaining to this viewpoint.

   Answer. See slide 35 of the introductory slide set.

8. (1 point) Module architecture control (MAC) can be used to analyze conformance between various models of an architecture. Explain how this is done. In particular, explain how conformance of the code view of a system architecture to its module view is analyzed.

   Answer. For the general approach, see slide 24 of the slide set on MAC. The module view uses module diagrams to express part-of and uses-relation between modules.

In the code view modules could be given by header files and the uses-relations by include-statement within these files.

9. (2 points) Requirements for quality attributes are specified through quality scenarios.

   (a) Name the ingredients of a general quality scenario.
   Answer. The inputs: stimulus and source of stimulus; the preconditions: state of the system that describes the environment of the scenario, and the resources of the systems (the artifacts) that are the subject of the scenario; the outputs: the response to the stimulus and the response measure which quantifies the extent of the response.

   (b) Give a specific quality scenario for the real-time performance of a video conference system. For each ingredient mentioned under (a), make explicit by which part of your answer it is addressed.
   Answer. When, upon start of the video conference (environment) the chairman (source) welcomes (stimulus) the other participants, he wants them all to receive (response) that message approximately simultaneously (measure, expressed as a two-sided bound on the transmission delay). The scenario addresses the (latency of) the network (resources) over which the video conference is held.

10. Indicate for the following statements whether they are true or false. Motivate your answer with a short argument.

    (a) (0.5 point) DNS cannot be used for service discovery.
    Answer. False.
    Standardly, DNS contains MX-records to locate mail servers. In addition SRV-records can be used to discover custom services

    (b) (0.5 point) Deferred RPCs are used for latency hiding.
    Answer. True.
    The mechanism consists of two asynchronous RPC calls; first a request by the client, followed by a response by the server. Usually the request is made by the client to prefetch some information from the server that will speed-up subsequent synchronous RPCs by the client that without the prefetched information would block for a considerable amount of time.
    (see TvS pp. 134–135 for more information)

    (c) (0.5 point) For Web server clusters, TCP handoff and content-aware request distribution are incompatible.
    Answer. False.
    The handoff can be done in two stages. First, any server of the cluster is contacted by the "TCP-switch" according to its scheduling policy, which is not content-aware. Instead of dealing with the client's request, however this server invokes a dispatcher that determines, based on the message content, the most suitable server in the cluster to handle the request. The TCP-connection is then

handed-over to that server and the switch is informed of this fact. (see TvS p. 559 for details).

(d) (0.5 point) Gossiping can be used to obtain *eventually-consistent* replicas.
Answer. True.
Provided the selection of gossip partners is such that eventually each peer has been engaged in a gossip session to exchange state information with all of its neighbors. Random selection of gossip partners such that each neighbor has a positive probability to be selected will do so.

11. (2 points) Name at least two scaling techniques. For each of the techniques you mention, indicate a concrete system, or an architectural pattern, in which it is applied.

Answer.

(a) Hiding communication latencies which can be achieved through asynchronous communication

(b) Distribution, e.g. the file chunks in GFS are distributed over chunk servers, or the name servers of DNS that take care of the zones into which the DNS name space is divided.

(c) Replication, e.g. a cluster of game servers for an online game. Each server has a replica of the common game world. The players are distributed over the servers, thus achieving load balancing.

12. (1 point) In distributed systems we distinguish between temporally coupled and temporally decoupled communication. Explain what this distinction is about. Give an example of both types of communication.

Answer. In temporally coupled interaction the communication/processing entities should co-exists (be active simultaneously). In temporally decoupled interaction parties do not necessarily coexist and therefore the communication medium should provide persistent storage.

Coupled interaction: Synchronous RPCs.

Decoupled interaction: Email, or Publish-subscribe systems using e.g. message queues.

13. (1 point) Describe in detail how the URL

`http://ieeeexplore.ieee.org/xpl/conferences.jsp`

is resolved by your internet browser.

Answer.

- The name of the Web server `ieeeexplore.ieee.org` is resolved into its IPaddress using DNS.

- The protocol (http) is resolved (by the Web browser) to a local implementation that is used to contact the Webserver with a request for the file `xpl/conferences.jsp`
- At the Web server, the OS is invoked to resolve the file name starting at the appropriate closure (the root of the local file system) and the file is returned to the browser where it is displayed. B.t.w., the file extension `jsp` signals that the browser needs a plugin that can handle Javascript, for the latter.

14. In the Open Service Architecture for Sensors (OSAS) a technique called content-based addressing is used.

    - (1 point) Explain this technique.
      Answer. The destination of a message is defined by a predicate on the message content and the state of a node.
    - (1 point) What is the benefit of this technique in the selective distribution of virtual machine code?
      Answer. The machine code is distributed as a single sequence with content-based addresses indicating which nodes need to install it. This allows to send just a single sequence into the network rather than specialized codes, per node.

15. (2 points) Describe at least two ways in which a RPC may fail and for each of those indicate a tactic that deals with it.

    Answer. The failures are described by a failure model and the tactics may depend on RPC-semantics.

    Lost request message. The client sets a timer for the reply. When it expires before the reply is received the client re-sends the message. This works provided the requested operation is idempotent.

    Client crash. This leaves an orphan process on the server that at some time in the future has to be killed. This can be done by a tactic called reincarnition. At every boot-up, so also after a crash, a client sends out a broadcast with a number telling which reincarnation (boot-up) this is. Upon receipt of this broadcast, every server kills all processes (orphans) left over from a previous reincarnation of the client (in case the network can also become partitioned, a more detailed tactic is necessary, see TvS).

    see TvS pp. 337-342 for other examples.