Work clearly. Read the entire exam before you start. Maximal grades per question are given between parentheses. The total sums up to 20 points and there are 12 questions. Multiple choice questions may have one or more correct answers. You are expected to mark all correct answers. In case of multiple correct answers the points are evenly distributed. Answers that are wrong count negatively (i.e., they cancel a correct answer but scores per question will not be negative).

1. (1) IoT refers, amongst others, to (select all that apply)
   a) the reach of electronics and software into the physical world
   b) the ease of use of electronic equipment
   c) the generation of massive and detailed data about embedded states and events
   d) massive data processing of heterogeneous data
   e) the automation of standard tasks
   f) the increase of compute power
   Answer: a, c, d

2. (3) Suppose you want to monitor the occurrence of some event (e.g. a fire hazard) and need to use a battery operated sensor $S$ with a wireless connection for that. Between sensing operations the sensor goes to sleep to save battery life. When the event occurs, $S$ reports it to node $D$ that is connected to the powered wired infrastructure.

   a) (0.5) Which of the following two interactions is preferable (motivate):
      i. When the event occurs let $S$ report it to $D$ and let $D$ always listen.
      ii. Let $D$ pull information about events from $S$ at regular moments in time.
   b) (0.5) Explain in both cases the effect on event reporting latency and energy.
   c) (1.0) Suppose that the software of $S$ needs to be reconfigured via communication. How could this be done?
   d) (1.0) Suppose that $D$ is duty cycled. Explain how $S$ and $D$ could synchronize in order to communicate. Can $S$ transmit at arbitrary moments?
   Answer: a): i. (no time synchronization required and wakeup periods, no delay upon event reporting, simpler for $S$); b) i. has the lowest latency possible while ii. has a max
latency of the synchronization period. ii also spends more energy as $S$ has to be prepared for information pull (hence, be on at regular times) c) Let $S$ send a message at a low frequency (e.g. once an hour or day, a 'keep-alive' message) to which $D$ responds with a command that brings $S$ into a mode to do this update. d) Sender initiated, and $S$ must know the period of $D$. Receiver initiated does not work since it is not known when $S$ wakes up. $S$ can send at arbitrary times but $D$ receives only when it is ready which is guaranteed to happen within the period of $D$, hence, $S$ must transmit repeatedly: long enough to cover this interval. Another correct answer is: no, $S$ must send around the time $D$ receives. Alternatively, $S$ and $D$ can apply strict time synchronization but this requires messaging as well (could be piggybacked) and some means of starting this up. An always-on broker between the two also works but shifts the problem to the broker.

3. (2) Consider a network consisting of nodes, some of which have CoAP services while all of them implement mDNS. Consider the following statements and indicate whether they are true or false with a mandatory and correct one line explanation.

   a) mDNS is required to discover CoAP services.
   b) mDNS can be used to discover arbitrary services.
   c) mDNS can be used instead of DNS.
   d) mDNS uses broadcast messaging for initiating service discovery.

   Answer: a) no, a discovery (GET /.well-known/core) from a multicast address (All-CoAP-Nodes) will yield all CoAP services; b) yes, provided that these services register themselves with the mDNS subsystem; c) no, mDNS is a separate protocol that uses the DNS message format; d) no, multicast messaging (to address 224.0.0.251 (IPv4) or FF02::FB (IPv6)). [broadcast would have no destination]

4. (2) Consider the following statements and indicate whether they are true or false with a mandatory and correct one line explanation.

   a) CoAP is an application protocol.
   b) CoAP requires a full URI as parameter to ‘GET’.
   c) Message IDs in CoAP couple Requests (like GET) and Replies (like 2.05)
   d) CoAP requires a reliable carrier protocol like TCP.

   Answer: true, it uses UDP as carrier, false, the specific parts of a URI can be given as parameters, false, message IDs couple CONs and ACKs, false, it can use any carrier

5. (2) Consider the following statements and indicate whether they are true or false with a mandatory and correct one line explanation.

   a) LWM2M is implemented using CoAP messaging.
   b) A LWM2M server always acts like a CoAP server.
   c) LWM2M provides security.
   d) LWM2M commands refer to objects with a standardized numbering scheme.

   Answer: true, see the mapping in the slides, also correct is: false: currently, there is only a CoAP binding. Bindings to other carriers are possible in principle; false, sometimes it has the client role, sometimes the server, false, it prescribes security using security of the carrier protocols, true, OMA maintains this database

6. (1) Mention two reasons why the RESTful architecture is beneficial in the context of IoT.
7. (1) Describe the difference between pull communication and push communication, and an example of when to choose either one of them.

**Answer:** Push: control flow and data flow go in same direction (‘call-back’, ‘event driven’); Pull: control flow and data flow go opposite (‘call’, ‘polling’); Push is better used when a communication must be non-blocking and not synchronized, e.g. a sensor delivering its information upon an event; pull is used when the information receiver must be in control of timing, e.g. when a controller wants to pick up information.

8. (3) CSMA is a protocol to use a shared channel by several nodes. There are three variants: CSMA/CA, CSMA/CD, CSMA/CR.
   a) (1.5) Describe them briefly, clearly exposing their differences.
   b) (1.5) For each protocol indicate whether it can be used for wireless channels or not. Briefly motivate your answers.

**Answer:** slides on networks, 28-30. CSMA/CA is used in 802.11 networks. /CR cannot be used since the detection and priority based continuation cannot be implemented in wireless networks; /CD does not work either as the sending node cannot detect a collision during transmission. /CA is indeed used widely in wireless LAN, wireless PAN and LoRA (IEEE 802.15.4 and IEEE 802.11)

9. (1) Explain the difference between horizontal and vertical analytics, and how cloud storage plays a role in both.

**Answer:** see architecture slides. Horizontal analytics analyses data about multiple entities, which is greatly simplified by bringing this data together in cloud storage. For vertical analytics (analyzing data about a single entity) cloud storage is similar to any other storage. However, the data from that single entity might still come from many sources. Then also cloud storage might help although one might argue that it should be stored in the managerial domain of the entity owner.

10. (2) Consider the following statements about Intelligent Transportation Systems, and indicate whether they are true or false, each with a mandatory and correct one line explanation.
   a) Basic Safety Messages implement safety applications.
   b) The European variant of ITS (ITS-G5) supports message routing.
   c) The US variant of ITS (WAVE) relies on periodic multicasting.
   d) Interference is a source of message loss in WAVE.

**Answer:** all true; a) periodic broadcasting of BSM messages is used to inform neighboring vehicles; b) ITS-G5 is based on IEEE 802.11p and implements geographical
routing; c) the US variant of ITS include the BSM messaging of a); WAVE suffers the standard problems of CSMA/CA: hidden nodes and collisions.

11. (1) Describe in one line the goal of k-means clustering.

\textbf{Answer}: The goal of k-means clustering is to partition all the data samples into k clusters of similar samples.

12. (1) Bob is interested in predicting his gas bill at the end of the year. A smart meter at his home publishes the total gas consumption via a webserver. The gas consumption measurements are provided at the end of each month, stating the total consumption until that measurement. Bob has collected a data set containing the following samples for the first 5 months:

<table>
<thead>
<tr>
<th>Month</th>
<th>Gas meter (m$^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2200</td>
</tr>
<tr>
<td>2</td>
<td>2600</td>
</tr>
<tr>
<td>3</td>
<td>2750</td>
</tr>
<tr>
<td>4</td>
<td>3100</td>
</tr>
<tr>
<td>5</td>
<td>3150</td>
</tr>
</tbody>
</table>

Bob assumes the following linear gas consumption model:

\text{Gas meter} = 2000 + 250 \times \text{Month}

a) What is the mean squared error of Bob’s model on the data set above?

\text{Answer:} (\frac{(2200m^3 - 2250m^3)^2 + (2500m^3 - 2600m^3)^2 + (2750m^3 - 2750m^3)^2 + (3100m^3 - 3000m^3)^2 + (3250m^3 - 3150m^3)^2}{5}) = \frac{6500}{5} = 6500 \text{ m}^6

b) What is the anticipated gas consumption during the 12 months of the year?

\text{Answer:} 12 \times 250 \text{ m}^3 = 3000 \text{ m}^3