Solutions

NATIONAL UNIVERSITY OF SINGAPORE
SCHOOL OF COMPUTING

Midterm Examination
for Semester 1, 2012/2013

CS5248 – Systems Support for Continuous Media

31 October 2012 Time Allowed: 2 Hours

INSTRUCTION TO CANDIDATES

1. This examination paper contains **TEN (10)** questions and comprises **ELEVEN (11)** printed pages including this page. The total mark is 100.
2. This is an **OPEN BOOK** examination. You may use any approved calculators but not any handphone or laptop, especially those capable of external connectivity or communication.
3. Answer **ALL** questions within the space provided in this **QUESTION AND ANSWER SCRIPT**.
4. Please **write your matriculation number below** and **do NOT** write your name.

YOUR MATRICULATION NUMBER: ________________________

(This portion is for examiner's use only)

Question	Marks	Score	Check
Q1	10		
Q2	8		
Q3	4		
Q4	12		
Q5	12		
Q6	12		
Q7	12		
Q8	8		
Q9	12		
Q10	10		
Total	100		

Question 1 Warm-up Questions
Please indicate for each statement whether it is True or False. Each question counts for 1 mark.
Answer (10 × 1 = 10 marks):

a) In DASH streaming the scheduling of the segment retrievals is done at the server side.
☐ True ☒ False

b) It is a good idea to use retransmissions for packet loss recovery on a channel with a long transmission delay.
☐ True ☒ False

c) The run-length encoding in an MPEG or H.264 compression algorithm generally results in a size reduction.
☒ True ☐ False

d) The GOP structure in a media file (e.g., IBBPBB) must remain the same within the whole file.
☐ True ☒ False

e) We must always use three different qualities (low, medium, and high) in a DASH system.
☐ True ☒ False

f) Delay jitter is caused by variable propagation delay in a network. ☐ True ☒ False

g) In an ongoing RTP transmission (e.g., video conferencing) the compressor (e.g., MPEG-2 or H.264) can be changed during the transmission. ☒ True ☐ False

h) The compression ratio of H.265 will be about 4 times higher than the compression ratio of MPEG-2.
☒ True ☐ False

i) Since all codec algorithms are based on intellectual property, licensing fees must be paid for all to use them.
☐ True ☒ False

j) The additive increase, multiplicative decrease in the rate adaptation of a protocol helps to keep the Internet stable. ☒ True ☐ False

Question 2 Compression.
Many of the compression algorithms, especially MPEG, use the concept of slices when encoding a video frame. (a) Describe what the reason is to use slices in encoded frames. (b) Describe one advantage and one disadvantage when using one slice per frame versus using many slices per frame.
Answer (2 × 4 = 8 marks):

(a) Slices are useful in error recovery when some parts of a frame have been lost due to packet loss. A slice has its own header. If the bitstream contains an error, the decoder can skip to the start of the next slice. Having more slices in the bitstream allows better error concealment, but uses bits that could otherwise be used to improve picture quality.

(b.1) If a frame has more slices, then less damage will be visible if a small part of a frame has been lost, i.e., more of the frame can still be successfully decoded.

(b.2) Each slice incurs some overhead as it includes an additional slice header. Therefore, if we use a lot of slices per frame then the compression ratio will be less than with fewer slices.

Question 3 RTP Protocol
Many voice streaming systems (VoIP) for interactive conferencing applications encode 20 milliseconds of audio into one packet each. Assume that encoding 20 ms of audio results in 50 bytes of compressed data. (a) Give a detailed reason why 20 ms is a good value, as opposed to for example 5 ms or 100 ms.
Answer (1 × 4 = 4 marks):

(a) An audio packet size of 20 ms is usually a good compromise (tradeoff) between latency and overhead. If the data in each packet is short (e.g., 5 ms) then the amount of data will be less and the fixed RTP header overhead of 12 bytes represents a larger overhead in terms of size. Additionally, if more small packets are generated then more interrupts are generated in the network stack at the receiver (the system will be more busy).

On the other hand, if more data is put into one packet (100 ms worth), then the end-to-end delay will increase because the data must first be collected from the soundcard before it can be transmitted to the receiver.

Question 4 Rate Adaptation
DASH systems usually use a fairly coarse-grained rate adaptation mechanism with, for example, three different bandwidth levels (equivalent to three different video qualities: low, medium and high). A DASH scheduler can either be aggressive or conservative in switching between those levels, i.e., it may try to switch frequently or rarely. Describe 1 advantage (positive aspect) and 2 disadvantages (negative aspects) of switching aggressively.
Answer (3 × 4 = 12 marks):

(a) (Positive) Switching aggressively, especially to higher quality, will often result in a higher overall (average) video quality. The scheduler will quickly use higher quality if the bandwidth is available.

Please note that switching aggressively does **not** mean to switch more frequently. In DASH, we can only switch after each streamlet. Switching aggressively means to change between the qualities more often whereas a passive policy would just keep requesting the same quality for a long time.

(b) (Negative) Switching aggressively can result in frequent quality changes for the users. Especially when it is clearly visible that the quality changed, it may be somewhat annoying or disrupting if the quality constantly changes.

(c) (Negative) With an aggressive policy the client may more easily run out of data when there is congestion in the network. If we assume that the maximum buffer size is fixed at the client, then if we fill it up with high-quality segments (which are larger) we will be able to “cover” less display time if the network connection breaks.

Question 5 VoD Rate Smoothing
Methods such as the Multi-Threshold Flow Control (MTFC) are used to smooth the transmission rate between a server and its clients (see Lecture 6). In the lecture slides three different possible optimization criteria are mentioned for a rate smoothing algorithm: (1) minimize the number of rate changes, (2) minimize the client buffer requirements, and (3) minimize the peak data rate. (a) Select 2 of the above criteria and describe one reason each why a designer should use them. (b) Come up with one other (your own) optimization criteria and give a reason of why to use it.
Answer (3 × 4 = 12 marks):

(Note: MTFC was implemented in a system with RTSTP/RTP/RTCP streaming protocols (server push), not in a DASH system (client pull).)

(a.1) Minimizing the number of rate changes will minimize the control traffic that is sent between the client and the server. Each rate change requires a message to be sent to the server. If there are no changes, then the server will just keep sending at the previous rate.

(a.2) Improving the utilization of the buffer in this case means to have a buffer that has a small size while still providing good performance. If we implement a large buffer which most of the time is not used much, then we are wasting memory cost. This may not be a problem in a desktop computer. However, if we implement our streaming client in a set-top box then having less memory means less cost of the device.

(a.3) Reducing the peak data rate of a stream in a network will most likely reduce congestion and allow for more concurrent streams to be carried in the network.

(b) For example, minimize the change in the data rate from one segment to the next. By reducing the “step-size” of each rate change we ensure that there will be gradual changes and network congestions are reduced.

Question 6 Adaptive Playout

Adaptive audio playout is a technique (see Lecture 6) where voice conversations are divided into talkspurts and the time between talkspurts is dynamically adjusted. Assume that you are the designer of a video conferencing system and you would like to use a similar technique for video transmissions. (a) Describe how you would divide the video (i.e., what is the equivalent of an audio talkspurt in video?), and (b) describe some of the challenges that you see for the implementation of such a technique for video.

Answer (2 × 6 = 12 marks):

(a) The key for audio talkspurts is that there is silence between them. A talkspurt is the part where people are talking. Such segments can be quite long (several hundred milliseconds or even seconds). In audio it is relatively easy to do silence detection. With video it is much more difficult to detect periods that are “useless”, i.e., which could be cut from the transmission. Also, in video there may be non-verbal communication, e.g., gestures, which should not be lost. We could try to do some kind of “activity detection” and then cut video parts where no activities occur.

(b) As mentioned under (a), doing activity detection in video is difficult. Gestures may be detectable. However, small face expressions (e.g. a frown), etc., may be difficult to detect. Therefore it will be difficult to find periods where there is no activity. As an alternate solution we may be able to speed up or slow down the video during low-activity periods.

Question 7 Proxy Caching for Streaming Media
One important input factor for streaming media caching algorithms is the file and segment popularity. Describe (a) one method of how file and/or segment popularity could be realistically measured. Describe (b) one advantage and one disadvantage of the method that you described as answer to (a).
Answer (3 × 4 = 12 marks):

(a) A realistic method to measure file and segment popularity is through access frequency. Every time a segment or a file is requested by a client it will be accessed in the caching node. By counting how many accesses during a fixed time period occur we can get a measure of popularity relative to other segments and files.

(b.1) (Advantage) One advantage of the above method is that it is quite simple and can be implemented with little processing and memory overhead.

(b.2) (Disadvantage) One drawback of the method is that we can only measure the past popularity. We don’t know the future popularity of a segment or a file, which is really what we would need to optimize the system.

Question 8 Proxy Caching for Streaming Media
Any caching algorithm must have a replacement policy that decides which segment or file needs to be evicted from the cache if it is full and space is needed to bring in a new segment or file. One simple policy is Least Recently Used (LRU), which basically evicts the "oldest" segment or file when space is needed. Describe (a) one reason why LRU might be a good policy and (b) describe one reason why LRU might not be a good policy for streaming media.
Answer (2 × 4 = 8 marks):

(a) The LRU policy uses just one timestamp – when the file or segment was last accessed – as the metric to decide what object to evict. There is a good chance (however, not a guarantee) that popular segments have been recently accessed. As such, the policy is simple and efficient to implement.

(b) However, one simple number may not always capture the popularity or other factors precisely. In some cases an object may periodically be "hot", say always every day in the evening. Therefore, the LRU might replace it during the day with some cold segments or files and then it needs to be cached again in the evening. Also, LRU usually means that every segment that is accessed will be stored in the cache until its timestamp becomes old and it is thrown out. In some cases, if a file is very unpopular (e.g., it is only accessed once a year), it should never be put into the cache, i.e., it should "by-pass" the cache. However, LRU does not usually allow this.

Question 9 Application Level Multicast
With the ***Narada*** algorithm (see Lecture 10) once in a while a node randomly selects links to other nodes to examine whether they should be added or dropped. Describe (a) why this is done and (b) describe 2 important issues that need to be considered to assure that the overall system remains stable.
Answer (3 × 4 = 12 marks):

(a) The reason for periodically testing new links and removing other links is to be adaptive and improve performance of the overlay network over time. The system is built randomly at the beginning, which is not optimal. Furthermore, network conditions can change over time, so the system should adapt to those changes.

(b.1) It is important to select "good" cost functions that accurately reflect how much an added link or a removed link would improve the overall system. The cost should be proportional to the utility it brings to the system. Otherwise the system will run the optimization algorithm, but in effect end up with a sub-optimal solution.

(b.2) The cost of removing a link should be higher than the cost of adding the same link again. This guarantees that the system will not get into a situation where it will remove and add the same link over and over again (a so called oscillation).

Question 10 Error Recovery
Retransmission of packets is one error recovery technique. Instead of always doing retransmissions when a packet does not arrive, it is also possible to implement selective retransmissions. For example, packets may only be retransmitted when there is enough time available before the playout deadline. Describe (a) 1 other decision policy when to selectively transmit a lost packet. Another factor to affect the error recovery effectiveness is to allow only 1 or more than 1 retransmissions per lost packet. As a designer, describe (b) your own policy of how to treat I, P, and B frame retransmissions.
Answer (2 × 5 = 10 marks):

(a) A packet may be selectively retransmitted based on its priority or usefulness to the display quality of a media stream. For example, if a lost packet would result in a large propagation of errors, then it would be very useful to retransmit it. Generally, I-frames would result in larger propagation errors than P-frames (and none for B-frames).

(b) One possible policy is to allow up to 2 retransmissions for I frames, up to 1 retransmission for P frames, and no retransmissions for B frames.

- END OF PAPER -