Computer Communication
Exam: EDA344, DIT423; Re-exam EDA343

Time and Place: Wednesday 18/3, 2020, 14.00-18.00, from home, using electronic submission

Course Responsible: Marina Papatriantafilou (Tel: 031 772 5413)

Grading:

For this exam, which is to be written from home and submitted through the Canvas system due to extra-ordinary situations, we will not base the overall evaluation on detailed points on assignments, but rather evaluate each submitted exam based on the overall impression, i.e. how well you demonstrate that you understand the subject. It might be not difficult to pass (grade 3 for CTH and G for GU), but it will be significantly harder to show that you deserve a higher grade (4, 5 for CTH and VG for GU). Good justifications that show that you understand the topic are required; therefore, try to answer the questions in detail, motivate your answers, rather than replying with just a sentence or two.

Instructions

- Submit **a SINGLE PDF file** with your answers. Please name the file **firstName-lastName.pdf**.
- **Make sure to write on the FIRST PAGE your (i) full name, (ii) identification number (personnumer) (iii) course-code (EDA344/EDA343/DIT423). Leave the rest of the fisrt page blank.**
- **Start answering each assignment on a new section, titled Question X.Y (Eg Question 1.a, Question 3.c etc); please sort and number the sections according to the question-ordering. Change page when changing Question e.g. from Question 1 to 2, from 2 to 3, etc**
- If you need to add 1-2 hand-written illustrations, you may take pictures with your phone/camera and embed them in the document. **If you include such drawings, these should be made by you, i.e. do not include others' pictures or figures. Make sure that the material is clearly readable.**
- Write in a clear manner and **motivate** (explain, justify) your answers. If some answer is not explained/justified, it will get **significantly lower or zero** marking. If you make any **assumptions** in answering any item, do not forget to clearly state what you assume. A **rule-of-thumb for the extend of detail to provide**, is to include enough information/explanation so that a person whose knowledge on computer communication is at the level of our introductory lecture, can understand what you explain.
- Please answer in English, if possible. If you have large difficulty with that and you think that your grade might be affected, feel-free to answer any question in Swedish.
- **Inspection of exam**: announcement will be sent through the canvas system.

Good Luck !!! Lycka till !!!!

Questions

1. General questions, performance, security (12p)

   (a) (6p) (i) What is the meaning of the term effective throughput of a network session?
   (ii) What is the effective throughput over a session running a stop-and-wait protocol over a link with bandwidth B bps, packet size P bits and one-way propagation delay T_p sec ?
   (iii) What is the effective throughput of such a session when running a pipelined protocol with the same B, P, T_p as above ?
   Explain and justify carefully your answers and if possible draw space-time diagram illustrations to support your arguments.

   (b) (3p) Host A and B are directly connected with a 100 Mbps link. There is one TCP connection between the two hosts, and Host A is sending to Host B an enormous file over this connection. Host A can send its application data into its TCP socket at a rate as high as 120 Mbps. Host B can read out of its TCP receive buffer at a maximum rate of 50 Mbps. Describe the effect of TCP flow control. Explain carefully your answer.

   (c) (3p) (i) Describe the main principles of a cryptographic hash algorithm and its use for validating data integrity.
   (ii) You are given a 4GB movie file and its SHA-1 hash. If you change one bit of the of the original file and recalculate the hash, do you think the new hash will be different? Why/why not?

   Hints:
   - the actual data rate going through; $P/(P/B + 2T_p)$ (actual data sent / time required for it to go through, considering transmission, propagation, acknowledgements and other protocol-related aspects) ; for N packets in the pipeline it is min (N times the above, B) the latter is when the pipeline fills the available bandwidth of the link
   - Limited throughput to 50 Mbps, Flow ctrl is caring for the capacity of the receiver.
   - one-way function that maps original data (eg message) to a fixed-size bit string, the "hash" (aka "message digest"); Compare the hash of the transmitted msg with the hash of the received one (ii) different; with SHA-1 it is intractable to to find another sequence that maps to the same hash value + by construction it is such that a small change in input will drastically change the output

2. Reliable Data Transfer and Transport Layer (12p)

   (a) (6p) (i) Explain a significant similarity and a significant difference regarding the problems of flow control and congestion control.
   (ii) Explain how TCP builds on the observations of the answers to the above question, for its flow control and congestion control mechanisms. What are the common features and what are the differences in TCP's ways of dealing with the problems?

   (b) (3p) If two end-systems are connected through multiple routers in the Internet and the Data-Link Layer between all of them ensures reliable data delivery, is Transport Layer reliability necessary? Why/why not?

   (c) (3p) Compare the addressing and de-multiplexing provided by UDP and by TCP and elaborate on what is a main advantage of the latter in client-server applications.

   Hints:
   - similarity: control rates of data flows; difference: flow ctrl: sender-receiver problem, congestion control: network core problem. Similarity: TCP Limits

sending rate at the sender, through window-variables (different receiver's window and congestion window)
- yes because routers can drop messages in out-queues; messages can also be re-ordered, hence ordering them back is needed.
- TCP multiplexes and demultiplexes; assign separate sockets for (sID,dIP,sPort,dPort), UDP uses only one IP and port number; TCP's advantage: the application does not need to keep track of the info, it can eg associate one thread of a server per client.

3. Network Layer (12p)

   (a) (4p) Describe the difference between (i) routing and forwarding, as well as control-plane and data-plane functions. (ii) Is it possible to have virtual-circuit functionality using Software-Defined Networking? If yes, explain how. If no, explain why not.

   (b) (4p) Describe an advantage and a disadvantage when using a centralized routing algorithm and when using a distributed one. Provide an example algorithm for each paradigm.

   (c) (4p) Consider a network that is assigned the IPv4 prefix 33.22.20.0/23. The network shall optimally and fully, be divided into three subnets, one large and two equal-size smaller subnets. The large subnetwork must have double address space compared to each of the two smaller subnets.
   When answering the following questions, please explain your answers and calculations carefully, giving both the binary and decimal formats. For each subnet give: (i) address and subnet mask (ii) the number of valid host addresses that each subnet has room for (iii) the first and last valid host address of each subnet.

Hints:
- deciding about routes (control plane, @ routers or SDN servers/centers) vs actually moving the data (data plane, each router is responsible for its own forwarding); yes, in principle it is possible to make forwarding tables that forward based on virtual circuit numbers. In practice it can be very costly to calculate and maintain such forwarding tables; it is easier to calculate routes for aggregated traffic flows for different types of traffic (eg video, text applications, machine 2 machine traffic)
- consistent updates, apply changes concurrently, but update-server is single point of failure/bottleneck; Dijkstra's algo Link State routing; D: self-stabilising, needs time to convergence; Distance vector algo
- straightforward calculation as in the example in slides of corresponding lecture; make sure addresses are reserved for fixed stuff: broadcast, self-address

4. Data Link Layer and Wireless (12p)

   (a) (4p) Compare and justify the choices made to handle (i) bit errors and (ii) medium sharing in Ethernet and in 802.11 Data Link Layer protocols.

   (b) (4p) Compare the approaches of handling mobility using direct and indirect routing.

   (c) (4p) Consider a situation in which you work in a project, which is about designing the local network for machine-to-machine communication for a production environment, e.g. a factory that assembles vehicle components. Describe your line of thought for the following issues: what criteria are important for deciding about the type of medium and the type of medium access for the network to deploy and what methods could be applicable?

Hints:
-bit errors: in Ethernet, few errors so only CRC; in Wifi, higher bit error rates, so CRC + retransmission; medium sharing CSMA/CD vs CSMA/CA due to how collisions are experienced at the receiver
- description: see section 7.5; indirect routing better handles mobility but causes "triangle routing problem"; direct routing solves triangle routing but needs extra mechanism to handle mobility, to re-fix the direct route when location change requires that.
- Wireless vs wired: is there mobility?; Medium access: is the communication real-time or not? Predictable delays vs predictability of load: can lead to different choices, eg reservation based vs allowing collisions and having lower guarantees about successful transmission times. Harsh environment: error control? retransmissions?

5. Multimedia, Congestion Control and Internet in Evolution (12p)

   (a) (6p) (i) Why are TCP's error control and congestion control limiting factors when it comes to supporting multimedia applications in the Internet?
   (ii) How are these limitations addressed in the Internet? Explain and justify your answer carefully, including the TCP-friendliness perspectives.
   (iii) Why is TCP-friendliness important?

   (b) (6p) Explain one similarity and one difference between peer-to-peer file sharing applications and Content Distribution Networks. Reason carefully about your reply and explain using examples.

Hints:
- (i) Both ack-based; Retransmissions introduce extra latency and varying sending rate (due to CC) introduces jitter; (ii) buffering, FEC, UDP+QUIC, non-ack-based rate adjustments; (iii) to care for congestion control
- Similarities: distribute media from different sources (ie not a single server); need a directory structure or a procedure to ask where is file X stored (eg Gnutella, Kazaa, Distributed Hash tables); content is delivered from another entity than the one that maintains the directory info; also possible to have different sources in the same fetch (eg analogy bitorrent-DASH) Differences: directory structure and content is maintained among peers in the former (eg overlays in Gnutella, Kazaa, Distributed Hash tables), while the latter relies on infrastructure using servers and DNS (example as the Netflix or Spotify ones discussed in class) to translate addresses and redirecting; then content is delivered from content provider