Information

- This exam is for course codes IK2215, IK2204, and 2G1701
- The duration of the exam is 4 hours (14.00-18.00).
- Answers should be well structured and readable.
- Write your name and personal-id/date-of-birth on each page.
- No help material is allowed.
- Answers will be posted on the course web within 2 weeks after the exam.
- Results will be published in Daisy no later than November 19, 2014. Graded exams can be found as PDF files in Daisy. Requests for grading re-evaluations should be done in writing, and sent to IK2215@ict.kth.se, no later than December 3, 2014.
- The exam consists of 2 parts; Part A and Part B. Part A is a set of questions with short answers. **Respect the word limits!** Answers longer than the word limit will be truncated, meaning that we will disregard from the part of your answer that exceeds the word limit during the exam marking. Part B is a smaller set of questions that require more elaborative answers. To pass the exam you need to attain a certain number of points (preliminary 75%) on Part A. Higher grades (A-C or 4-5) will be based on the total score (Part A + Part B). **Part B will not be graded for those who do not pass Part A.**
- Preliminary grading is as follows:

Points	**Grade (A-F)**
23-30 points on Part A and 45-50 points in total	A
23-30 points on Part A and 40-44 points in total	B
23-30 points on Part A and 35-39 points in total	C
23-30 points on Part A and 23-34 points in total	D
21-22 points on Part A and passed complementary assignment	E
21-22 points on Part A (complementary assignment offered)	Fx
0-20 points on Part A	F (Fail)

Points	**Grade (U-5)**
23-30 points on Part A and 42-50 points in total	5
23-30 points on Part A and 37-41 points in total	4
23-30 points on Part A and 23-36 points in total	3
21-22 points on Part A (complementary assignment offered)	U
0-20 points on Part A	U (Fail)

Good Luck!

Exam Part A (30p) (Note the word limits)

1) Various true/false statements (10p)

Mark the following statements as **true** or **false**. Don't write "t" or "f", since indistinct hand-writing makes it hard to differ between the two.
Note:

- you will get 1p for each correct answer
- you will get -1p for each wrong answer
- you will get 0p for each "no answer"
- you will **not** get less than 0p in total on this question

A. In OSPF, a node sends distance vectors to all other nodes in the same OSPF area. (1p)
B. UDP will discard a datagram with checksum errors on the receiver side without sending feedback to the sender. (1p)
C. When BGP learns the same prefix from multiple paths, it prefers the one with the longest AS_PATH length. (1p)
D. MSDP can be used to interconnect multiple IPv4 PIM-SM domains. (1p)
E. According to GeSI (Global e-Sustainability Initiative), the greenhouse gas (GHG) emissions from ICT is not far below the GHG emissions from the aviation industry. (1p)
F. In diff-serv (Differentiated Services) a PHB (per-hop behavior) defines how individual application flows should be treated by routers inside the diff-serv domain. (1p)
G. In IPv6 the IP address identifies the host instead of a network interface card as in IPv4. (1p)
H. IP multicast works only for link technologies with hardware support for multicast. (1p)
I. One idea with a CDN (Content Distribution Network) is that the CDN company replicates content to multiple servers in different geographical locations. (1p)
J. When joining a torrent in BitTorrent, a user obtains a list of peers from an infrastructure node called a tracker. (1p)

Answer:

A. False
B. True
C. False
D. True
E. True
F. False
G. False
H. False
I. True
J. True

2) Various questions with short answers (10p)

Answer the following questions with short answers.
Note:

- You will get 1p for each entirely correct answer
- Word limit per question: 30 words

A. Place the following four protocols at the correct layer in the TCP/IP protocol stack: SMTP, PPP, TCP, and ICMPv6. (1p)
B. Can you aggregate the prefixes 199.1.1.0/25, 199.1.1.128/26, and 199.1.1.192/26 to one single subnet? If so, specify (in CIDR notation) the resulting aggregated prefix. If not, explain why.(1p)
C. At a given time, a TCP sender has a congestion window of 2000 bytes and a receiver-advertised window of 1800 bytes. What is the maximum number of bytes that the sender can transmit before waiting for an ACK? (1p)
D. What mechanism do CDNs typically use to allow users to find the most suitable server? (1p)
E. Mention one multicast routing protocol using a source-based tree and one multicast protocol using a group-shared tree. (1p)
F. RTP (Real-time Transport Protocol) is typically run on top of UDP. What are the two most important features that RTP adds to UDP for real-time applications? (1p)
G. Is the EEE (Energy-Efficient Ethernet) standard based on rate switching (switch to lower tranimssion rate when possible) or low-power idle (sleep between packets when possible)? (1p)
H. Mention one advantage and one disadvantage of OSPF compared to RIP. (1p)
I. Name two transition strategies for IPv4 to IPv6, devised by IETF. (1p)
J. BGP is neither based on link state nor on distance vector. Instead BGP uses path vector. How is this different from distance vector? (1p)

Answer:

A. SMTP: Application layer, PPP: Link layer, TCP: Transport layer, ICMPv6: Network layer.
B. Yes, ther resulting subnet is 199.1.1.0/24.
C. 1800 bytes.
D. DNS redirection.
E. Source-based tree: DVMRP, MOSPF, PIM.DM. Group-shared tree: PIM-SM, CBT.
F. Sequence number and timestamp.
G. EEE uses low-power idle.
H. Advantage: faster convergence, more functionality. Disadvantage: uses more memory in routers, more complicated.
I. Dual stack, tunneling, header translation.
J. A distance vector give the cost (distance) to each network. Instead of just the cost, BGP adds an AS path to every network it advertises.

3) Transport layer (2p) (Word limit: 120)

How many RTTs or transmission rounds does it take to transfer a 30 KB file over a 1 Gbps link with no competing traffic using TCP? The initial congestion window ($cwnd_{init}$) is 1 KB; the size of the destination's receive buffer (rwnd) is 6 KB. Since the source node runs Linux, ssthresh has been cached from a previous session, and is 4 KB at the start of the file transfer. The receiver is able to remove packets from the receive buffer at the pace of their arrival, i.e., the receiver always advertises an rwnd of 6 KB.

Answer:

The transfer starts out in slow-start, and stays there until the cwnd reaches the ssthresh (4 KB), i.e., during three RTTs. It continues in congestion avoidance during the remaining part of its lifetime, however, when the cwnd reaches the limit of the rwnd (6 KB), the send window stops growing. Consequently, it takes another four RTTs until the file transfer completes. Thus, the file transfer takes 7 RTTs in total (1 KB + 2 KB + 4 KB + 5 KB + 6 KB+ 6 KB + 6 KB = 30 KB).

4) Multicast (2p) (Word limit: 80)

In a large subnet there could be many simultaneous receivers of the same multicast group. When a router sends an IGMP query, this could then generate a large amount of IGMP reports. How is this avoided in an IGMP implementation?

Answer:

IGMP reports are sent to the group address. The hosts can then snoop for other host's IGMP reports. So, before sending its IGMP report a host will set a random timer and suppress its IGMP report if it detects that another host on the same subnet sends the report for this group.

5) Link state routing (2p) (Word limit: 50)

You are snooping OSPF messages on a network and you know that there are five nodes (routers) in total in the OSPF area. After a while, you have collected the following LSAs (link state advertisements):

From *A*	
Node	Cost
B	2
E	5

From *B*	
Node	Cost
A	2
C	1
D	3

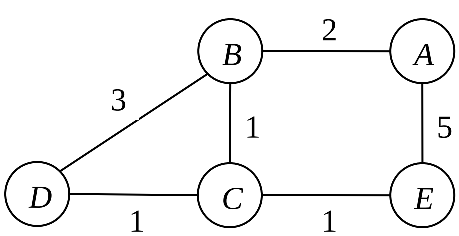
From *C*	
Node	Cost
B	1
D	1
E	1

From *D*	
Node	Cost
B	3
C	1

From *E*	
Node	Cost
A	5
C	1

What does the network look like? Draw a network graph (nodes, links, and link costs) that corresponds to these link state advertisements.

Answer:



6) IP QoS (2p) (Word limit: 120)

Int-serv (Integrated Services) and diff-serv (Differentiated Services) can be used to provide quality of service (QoS) for multimedia applications in IP networks. What are the main problems with int-serv and how are they addressed in diff-serv?

Answer:

The main problem with int-serv are that it requires end-to-end connection setup and resource reservation. Thus, for int-serv to work, all the routers along the traffic path must support it. Routers need to keep per-flow state information, which is not scalable.

In diff-serv, traffic flows are aggregated into a small number of classes. All packets belonging to the same traffic class get the same treatment in the diff-serv network. diff-serv requires no advance setup, no reservation, and no time-consuming end-to-end negotiation for each flow.

7) Peer-to-peer networking (2p) (Word limit: 100)

Assume two skype users, where each user is connected to his/her own private network behind a NAT (Network Address Translation) box. The NAT arrangement prevents hosts from outside the private network to initiate connections to hosts inside the private network. Still, skype will work without requiring the users to configure the NAT boxes for incoming calls. Describe the technique that skype uses to achieve this.

Answer:

Skype uses super peers and relay nodes to solve the problem. When a skype user signs in on skype from his/her private network the user is assigned a non-NATed super peer. Sessions can then be initiated to the super peers, who inform each other and the calling parties. If the call is accepted, a public non-NATed relay node is selected: Bob—relay node—Alice.

Exam Part B (20p)

8) IPv6 (5p)

Briefly explain the process (including the messages involved) of how an IPv6 host obtains its IPv6 address using stateless address autoconfiguration assuming there is an IPv6 router to hand out a global prefix.

Answer:

First the host creates an IPv6 with the link local prefix and its own MAC address. Then:

1. **A Multicast Listener report.** The host starts to listen on the multicast address associated with the link local address it intends to use.
2. **A Neighbor Solicitation request** sent to that multicast address for duplicate address detection (DAD).
3. **A Router Solicitation request.** It looks for a router on the link to advertise a link prefix (in order to form a global IPv6 address).
4. The router replies with **a router advertisement** (thereby the client can form a global IPv6 address).
5. The host sends another DAD message (as a **neighbor solicatition request)** to the associated multicast address. But this time it sends ICMPv6 with the global address as target address.

9) Transport protocol (5p)

Multipath TCP (MPTCP) is a set of extensions to standard TCP that are currently underway to become standardized by the Internet Engineering Task Force (IETF).

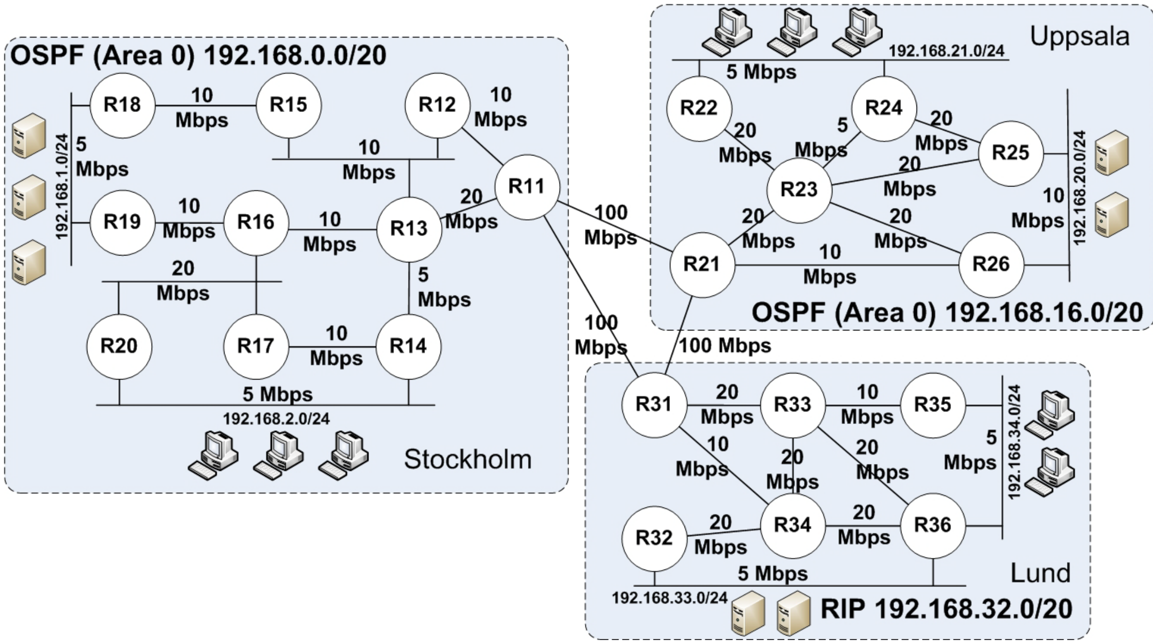
A. Give two design goals of MPTCP. (2p)
B. Explain the resource pooling principle and how MPTCP adheres to this principle. (2p)
C. Why does MPTCP employ coupled congestion control? (1p)

Answer:

A. At least two of the following goals: *improve throughput*, i.e., should perform at least as good as a single-path flow over the best of the available network paths; *improve resilience*, i.e., should permit packets to be sent and re-sent on any available network path; *do not harm*, i.e., should not impede performance on other regular TCP flows; *balance congestion*, i.e., move traffic from the most to the least congested network paths; *application compatible*, i.e., provide the same API and service model to applications as standard TCP.
B. The resource pooling principle states that the network's resources should behave as though they make up a single pooled resource. MPTCP adheres to this principle by making it possible for a single TCP connection to consist of several network paths.

C. Otherwise, an MPTCP multi-subflow connection would compete unfairly with competing standard TCP flows when the MPTCP subflows and the TCP flow go over the same link.

10) Large-scale network scenario (10p)



The above figure illustrates Company A’s network topology. Company A has three branch offices in different cities, each with different networks. The first office is located in Stockholm using the 192.168.0.0/20 subnet. The second office is located in Uppsala using the 192.168.16.0/20 subnet. The last office is located in Lund using the 192.168.32.0/20 subnet. All routers are interconnected with different link bandwidths as shown in the figure. All routers are Cisco routers with default parameters set.

Each office has designed its own internal network and runs its own routing protocol internally. The Stockholm office uses OSPF as its sole routing protocol within its network. All routers (R11-R20) are running in the backbone area (OSPF area 0). Similar to the Stockholm office, the Uppsala office uses OSPF as its sole routing protocol within its network (on routers R21-R26). All routers are running in the backbone area (OSPF area 0). The Lund office uses RIPv2 on all its routers (R31-36). These routing protocols are not running on the links between the offices! (no OSPF or RIP on R11-R21 link, R21-R31 link, and R31-R11 link). However, different routing schemes are used for communication between branch offices.

Static routes are configured on the following routers as follows:
On R21 there are two static routes:
1) all traffic destined to 192.168.32.0/20 forwarded to R31
2) Set default route to forward to R11
On R31, traffic to 192.168.16.0/20 forwards to R21.

BGP is used between Stockholm’s R11 and Lund’s R31 (over R11-R31 link). Both routers run BGP with private AS 65001 and the following route advertisements:

On R11, advertise a single aggregate address 192.168.0.0/**19** to R31 (all other routes are filtered and NOT advertised via BGP)
On R31, advertise only an aggregate address 192.168.32.0/20 to R11 (all other routes are filtered and NOT advertised via BGP)

In addition, the Stockholm office has configured R11 to always originate default route in OSPF. The Uppsala office has configured R21 to redistribute all static routes into OSPF. The Lund office has configured to redistribute BGP routes into RIP on R31 (IMPORTANT! RIP routes on R31 are NOT redistributed into BGP in Lund).

The Stockholm office has one server network (192.168.1.0/24) and one user network (192.168.2.0/24). HSRP (Hot Standby Router Protocol) is used to provide fault-tolerance default gateway for each network. For the server network, R19 is active router and R18 is passive router. For the user network, R14 is active router and R20 is passive router.

The Uppsala office has one server network (192.168.20.0/24) and one user network (192.168.21.0/24). HSRP is used to provide fault-tolerance default gateway for each network. For the server network, R26 is active router and R25 is passive router. For the user network, R24 is active router and R22 is passive router.

The Lund office has one server network (192.168.33.0/24) and one user network (192.168.34.0/24). HSRP is used to provide fault-tolerance default gateway for each network. For the server network, R36 is active router and R32 is passive router. For the user network, R35 is active router and R36 is passive router.

Assume that default cost models are used for OSPF, RIP, as well as for static routes (a static route has a fixed cost of 1, OSPF cost formula is *cost = <100,000,000 bps>/<bandwidth in bps>*). When a route is redistributed from one protocol to the other, the original cost of the route will be inherited to the new protocol and accumulated with the new cost before the route is forwarded to other routers. For BGP originated routes, the cost will be set to 10. In addition, if a router learns the same route from different routing protocols, it will prefer the route from the routing protocol in the following order: static, OSPF, RIP, and BGP. If a router learns a route with equal cost from the same routing protocol through multiple routers, it will prefer to use the route from the router with the lowest number. For example, if R51 learns an OSPF route 10.0.0.0/24 from R55 and R60, it will prefer to use the OSPF route learned from R55.

A host in Lund with IP address 192.168.34.5 is trying to ping a server in Stockholm with IP address 192.168.1.2. Answer the following questions:

A. What path does an ICMP echo request take? (1p)
B. What path does an ICMP echo reply take? (1p)

Example answer
10.0.0.1 -> RTX -> RTY -> RTZ -> 10.0.0.2

A host in Stockholm with IP address 192.168.2.3 is trying to ping a server in Uppsala with IP address 192.168.20.2. Answer the following questions:

C. What path does an ICMP echo request take? (1p)
D. What path does an ICMP echo reply take? (1p)

Example answer
10.0.0.1 -> RTX -> RTY -> RTZ -> 10.0.0.2

Assume that the physical link between Uppsala and Lund (on R21-R31 link) was taken down and the border routers lose all routing information on this link (static routes on this link are removed from the routing tables). Assume that the topology has converged, answer the following questions:

E. What path does an ICMP echo request take if a host in Lund with IP address 192.168.34.5 is trying to ping a server in Uppsala with IP address 192.168.20.1? (1p)
F. What path does an ICMP echo request take if a host in Uppsala with IP address 192.168.21.3 is trying to ping a server in Lund with IP address 192.168.33.3? (1p)

IMPORTANT: if the ICMP echo request cannot reach the destination then identify what happens to the request.
Example answer
10.0.0.1 -> RTX -> RTY -> RTZ -> 10.0.0.2
10.0.0.1 -> RTQ -> RTR -> RTS -> RTS drops the packet
10.0.0.1 -> RTM -> RTN -> RTO <-> RTN (Loop between RTO and RTN!)

Assume that the R21-R31 link is back to normal and the original topology in the figure has converged. A network administrator would like to run a multicast routing protocol in order to distribute recorded multimedia, stored on a streaming server (IP 192.168.1.9/24) in the Stockholm office, to all users in all branch offices.
Answer the following questions:

G. Assume that you are using PIM sparse mode to distribute your multicast stream, and that R16 is selected as a rendezvous point (RP). If the SPT-threshold is set very high and never exceeded, identify which path that will be used for streaming from the streaming server to the different hosts in each office according to a-c below.
   To avoid confusion caused by having two routers on the network (HSRP routers), assume that the links between the passive routers and the server and client networks in each office are removed from the topology (R18-192.168.1.0/24 link, R20-192.168.2.0/24 link, R25-192.168.20.0/24 link, R22-192.168.21.0/24 link, R32-192.168.33.0/24 link, and R36-192.168.34.0/24 link).
   The hosts in each office are as follows:
   a. 192.168.2.10 in Stockholm office (1p)
   b. 192.168.21.10 in Uppsala office (1p)
   c. 192.168.34.10 in Lund office (1p)

H. BGP usually has a slower convergent time compared to OSPF. Assuming that there are only a small number of routes in the network (and scalability is not an issue here), what is the main benefit in using iBGP instead of OSPF. (1p)

Answer:

A. 192.168.34.5 -> R35 -> R33 -> R31 -> R11 -> R13 -> R15 -> R18 -> 192.168.1.2
B. 192.168.1.2 -> R19 -> R16 -> R13 -> R11 -> R31 -> R33 -> R35 -> 192.168.34.5
C. 192.168.2.3 -> R14 -> R13 -> R11 -> R11 drops the packet!
D. 192.168.20.2 -> R26 -> R21 -> R11 -> R13 -> R16 -> R20 -> 192.168.2.3
   It is also acceptable if you answer that there is no ICMP echo reply because the ICMP echo request in C was dropped.
E. 192.168.34.5 -> R35 -> R33 -> R31 -> R11 -> R11 drops the packet!
F. 192.168.21.3 -> R24 -> R25 -> R23 -> R21 -> R11 -> R31 -> R33 -> R36 -> 192.168.33.3
G. The paths used for sending the stream are as follow:
   a. 192.168.1.9 -> R19 –> R16 -> R17 -> R14 -> 192.168.2.10
   b. 192.168.1.9 -> R19 -> R16 –> R13 -> R11 -> R21 -> R23 -> R25 -> R24 ->192.168.21.10
   c. 192.168.1.9 -> R19 -> R16 –> R13 -> R11 -> R31 -> R33 -> R35 -> 192.168.34.10
H. Apart from having better scalability, BGP also offers better routing policy control (through BGP attribute manipulation).