Carnegie Mellon

Computer Science Dept.
Carnegie Mellon University

15-744 Fall 2004

Midterm

Name: Solutions

INSTRUCTIONS:

There are 13 pages (numbered at the bottom) make sure you have all of them.

Please write your name on this cover and at the top of each page in this booklet.

If you find a question ambiguous, be sure to write down any assumptions you make.

It is advantageous to partially answer a question than not attempt it at all.

Be clear and concise. Limit your answers to the space provided.

For multiple-choice **CIRCLE ALL ANSWERS THAT APPLY.**

SCORING:

A	B	C	D	E	F	G	Total
/30	/20	/15	/15	/10	/30	/5	/125

A. Potpourri (30 pts)

1. Which of the following is true about modern fast router implementations?

   a) They use a bus-based backplane since it is simpler to implement than a switched backplane. (-2 pts)

   b) iSLIP achieves fairness across different input ports by using randomization (like Ethernet) instead of round-robin style scheduling (like Token-Ring). (-2 pts)

   c) Although line cards handle most common case processing, a central CPU still handles the routing protocol and exception processing. (+1 pt)

   d) They do not use simple output buffering since this requires the router backplane to support a large speedup. (+1 pt)

   e) Input Queuing, Output Queuing and Virtual Output Queuing all suffer from head of the line blocking. (-2 pts)

   f) The load balanced switch design in [KCY03] can achieve a theoretical throughput of 100% independent of traffic patterns. (+2 pts)

   g) The main observation on which the method in [GM99] is based is on the fact that packets do not match many rules in practice. (+2 pts)

   COMMON MISTAKES: did not mark d.

2. There are at least three approaches to designing a unicast routing protocol: distance vector (e.g., RIP), link-state (e.g., OSPF), and path-vector (e.g., BGP4). Which of these statements is true of these approaches?

   a) The use of split-horizon and poison-reverse in a distance-vector protocol like RIP always successfully prevent the count-to-infinity problem when a network partition occurs. (-3pts)

   b) Link state protocol create more routing messages than similar distance vector protocols. (+3 pts)

   c) Both link state and distance vector protocols end up storing the same state at each node – the next hop for each destination and an associated route metric (-3 pts)

   d) Link-state protocols do not suffer from the count-to-infinity problem. (+3 pts)

   COMMON MISTAKES: most people got this right.

3. A network C advertises the network number 192.3.124/22 (and no other numbers). What network numbers (all 24 bits) could AS C own?

   a) 192.3.123 (-3 pts)
   b) 192.3.124 (+3 pts)
   c) 192.3.125 (+3 pts)
   d) 192.3.128(-3 pts)
   e) 192.3.10 (-3 pts)
   f) 192.4.124(-3 pts)

   COMMON MISTAKES: most people got this right ..

4. Which of the following statements are true about area hierarchical routing and landmark routing?

   a) Area hierarchy leads to suboptimal routes because each network has only one egress point. (-3 pts)
   b) In a landmark hierarchy, nodes may have multiple valid addresses. (+3 pts)
   c) Landmark routing and source routing are equivalent. (-3 pts)
   d) In area hierarchy, a route between two nodes belonging to the same area never leaves that area. (+3 pts)

   COMMON MISTAKES: some people chose a.

5. Suppose S sends a packet to D in a packet-switched network. Suppose the network topology and the state in the routers/switches do not change. Clearly explain why each of the statements below is true or false.

   a) If datagram routing is used, correct forwarding can occur if the packet traverses the same network link (and switch pair) in opposite directions.
      -- No , this would result in a loop since routers only examine the destination address for deciding next hop. (1 point for answer, 2 points for explanation)
   b) If virtual circuit switching is used, correct forwarding can occur even if the packet traverses the same network link (and switch pair) in opposite directions.
      -- Yes, virtual circuit forwarding is based on label + interface. You can have different label/interface on either direction.
      (1 pt for answer, 2 points for explanation)

   COMMON MISTAKES: few people got b. wrong. Reasoning was unclear.

B. CSFQ, XCP and FQ (20 pts)

6. Which of the following is true about XCP?

   a) XCP uses MIMD in the efficiency controller and AIMD in the fairness controller. (+2pts)

   b) Each XCP packet carries both the congestion window and the RTT for the flow. (+2 pts)

   c) XCP does not require any cooperation from end-hosts to achieve congestion control. (-2pts)

   d) Unlike CSFQ XCP could provide RTT proportional fairness instead of mx-min fairness (+2pts)

   COMMON MISTAKES: Did not choose a.

7. Which of the following is true about fair queuing?

   a) Once assigned, the finishing round number does not depend on future packet arrivals and departures of either the same or a different flow. (-3 pts)

   b) The rate of change of the round number with time (dr/dt) depends on the number of active flows. (+3 pts)

   c) The starting round number of a packet depends on its size. (-3 pts)

   d) The finishing round number of a packet depends on its size. (+3 pts)

   COMMON MISTAKES: many marked a.

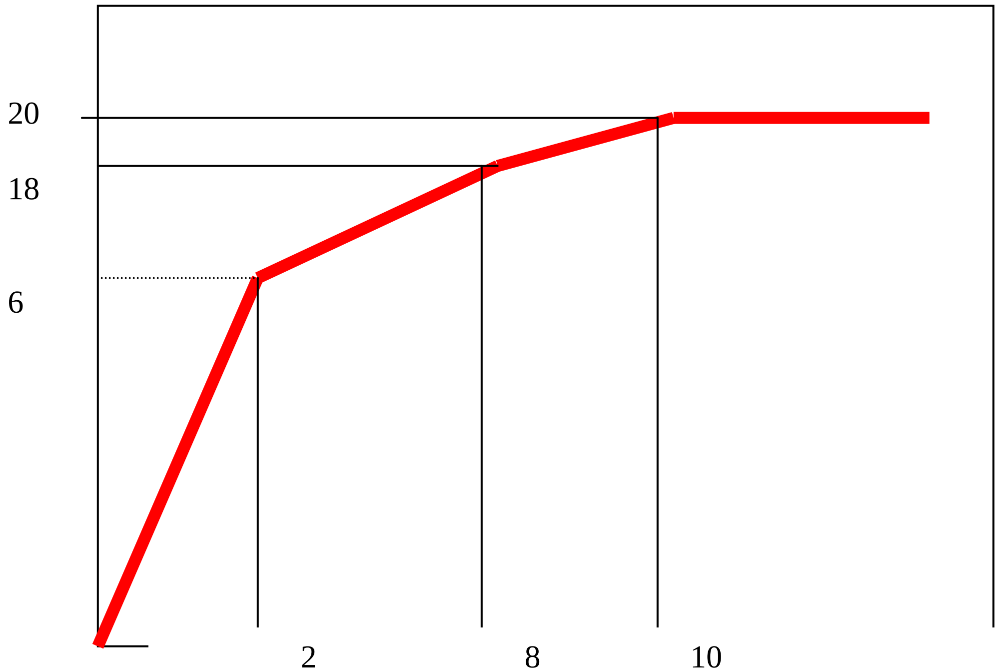
For the following, assume that a core router has 3 flows: flow 1, 2 and 3 with bandwidths of 2, 8, and 10 Mbps respectively.

8. a. The output link for these flows has a bandwidth of 10 Mbps. What is the rates achieved by the flows under max-min fair sharing.

2,4,4 (1pt for each value)

COMMON MISTAKES: most people got this right

8 b.This question relates to the Core-stateless fair-queuing paper. Draw a plot of F, the acceptance rate, (on the Y axis) against α, the fair-share parameter (on the X axis) (points 2, 8, 10 are marked on the X axis, provide the values for F at these points).



1 point for concave curve, 1 point for piecewise linear, 3 points for y coordinates

F(alpha) = sum(i) Min(C_i, alpha)
F(2) = Min(2,2) + Min (8,2) + Min (10,2) = 6
F(8) = Min(2,8) + Min (8,8) + Min(10,8) = 18
F(10) = Min(2,10) + Min (8,10) + Min(10,10) = 20

COMMON MISTAKES: Very few got this right !
Either the curve was not concave, not piecewise linear, or the y coordinates were wrong

C. AQM & TCP (15 pts)

Consider a network of RED gateways and a network of fair-queuing gateways. Both networks have congested links/gateways. Consider two long-running TCP connections in each network. In which of the following situations do the two connections achieve roughly the same throughput? For each question, give a short reason for your answer. (3 points each, 1 pt for answer 2 pts for explanation)

9. In the RED network, the two connections share a congested gateway, traverse no other congested gateways and have different round-trip times.

No – RED gives all flow the same loss rate, the flow with lower RTT will have higher throughput

10. In the RED network, the two connections share all their congested gateways and have equal round-trip times.

    Yes – expected throughput will be the same if they have the same RTT, and share all congested gateways.

11. In the fair queuing network, the two connections share a congested gateway, traverse no other congested gateways and have different round-trip times.

    Yes – if they share a congested router , then the throughput of both flows will be the max-min fairshare of that router

12. In the fair queuing network, the two connections share ***some*** congested gateways (but not others).

    No, it is possible that one of the flows is through a congested router with a lower fairshare allocation (more competing flows)

13. In the fair queuing network, the two connections share the slowest [raw] bandwidth link on their paths. However, the connections may have different round-trip times and traverse other congested gateways.
    no .. even though this link may be the true bottleneck in terms of capacity , it may be the case that some other congested gateway is the bottleneck for a particular flow in terms of available bandwidth
    COMMON MISTAKES: Answered yes to 12/13.

D. TCP Tuning (15pts)

Harry Bovik is given the responsibility of configuring the packet queuing component of a new router. The link speed of the router is 100 Megabit/s and he expects the average Internet round-trip time of connections through the router to be 80ms. Harry realizes that he needs to size the buffers appropriately.

You should assume the following:

- You're dealing with exactly one TCP connection.
- The source is a long-running TCP connection implementing additive-increase (increase window size by 1 packet after an entire window has been transmitted) and multiplicative-decrease (factor-of-two window reduction on congestion).
- The advertised window is always much larger than the congestion window.
- The loss recovery is perfect and has no impact on performance.
- The overhead due to headers can be ignored.

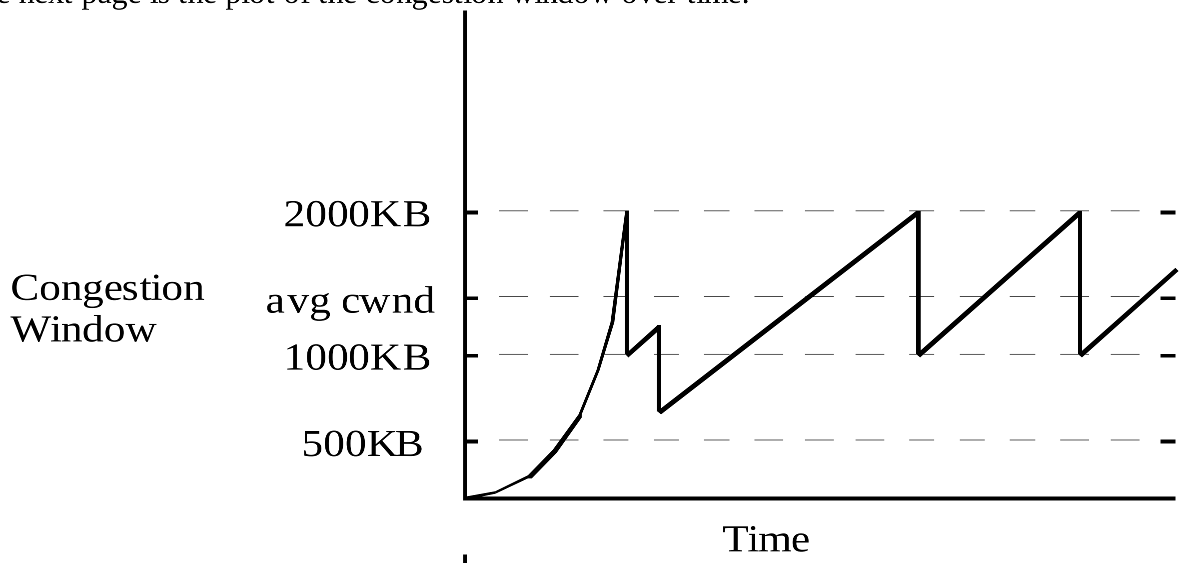
Harry argues that because the average RTT is 80ms, the average one-way delay is 40ms. Therefore, the amount of buffering he needs for high link utilization is 100 Mbps * 40 ms or 500 KBytes.
Approximately what bandwidth will TCP achieve with this buffering? (6 pts)

14.
Very Simple Linear Approximation:
100% at 1000KB and 75% at 0KB, therefore 100+75/2 = 87.5% at 500KB

Simple Linear Approximation:
The time it takes cwnd to go from 750KB to 1000KB is half the time it takes to go from 1000KB to 1500KB (Aside: This is not exactly true. Queuing makes it take longer to go from 1000KB to 1500KB). Whenever, cwnd > 1000KB (the delay- BW product) the link is fully utilized. Therefore the link is fully utilized 2/3 of the time. During the other 1/3 of the time, the average window is (750 + 1000)/2 = 875KB. This means that the link is on utilized approximately 87.5% during this period. Given that the link is underutilized by 12.5% approx. 1/3 of the time - this means that it is underutilized 12.5/3 = 4.16% over the entire period. The through put achieved is 93.84Mbps.

In going from a window
of 750 packets to 1500 packets, approximately A= (1500 + 750)*(1500-750)/2 packets are sent. However, during this same period, the source could have sent about B= 1000*250 + (1500+1000)(1500-1000)/2. Thus the utilization is A/B = 96.4 %.

Harry raises the buffering in the router to 1000KBytes in hope of fully utilizing his link. After Harry installs the router, he notices that packets are dropped due to both corruption and congestion. Therefore, he decides to make some observations about the observed TCP performance. He makes a transfer to some distant site and records the congestion window whenever it changes. The RTT to this distant site was exactly 100ms throughout the transfer. On the next page is the plot of the congestion window over time.



15. Assuming that Harry is using something like TCP NewReno or SACK, mark an X on the Time axis of the congestion window plot whenever a loss event occurred.
everytime there is a vertical drop.. (3 pts)

16. Circle the losses (X's from the previous question) that you think are likely due to corruption and not congestion. (3 pts)
congestion – when w reaches 2000 ..
corruption – second drop from left ?

17. Highlight the periods of time during which the link is fully utilized. (3 pts)

window > 1000kb (assuming RTT of 80 ms)
or window > 1250 kb (assuming RTT of 100 ms)

COMMON MISTAKES: most people got 15,16,17 right
No one got the real right answer for 14. If you answered 87.5 you were given full credit.

E. Congestion Control (10 points)

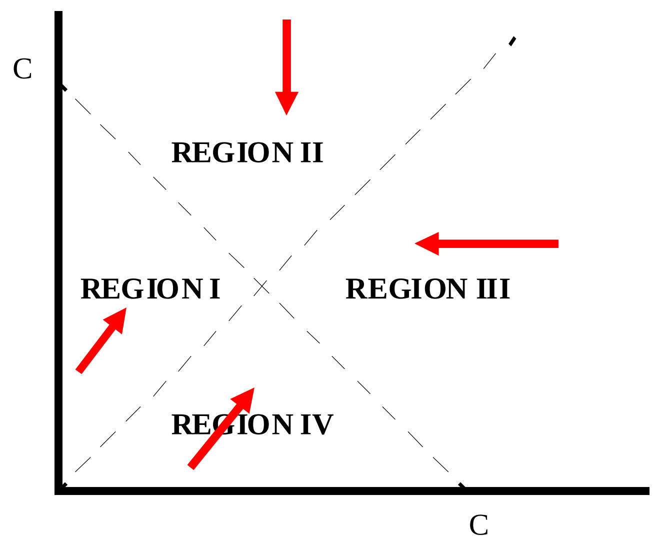
Harry Bovik is building a new type of router. The router uses three explicit feedback messages - DECREASE, STAY and INCREASE - to communicate with sources in the network. It uses these messages in the following ways:

- If the network is underutilized, the router sends an INCREASE message to all sources.
- If the network is over utilized, the router sends a DECREASE to the sources with the highest allocation and STAY messages to all other sources.
- If the network is perfectly utilized, the router sends a DECREASE to sources with the highest allocations.

We will assume that the number of streams using the router is always fixed, feedback messages are never lost, and the round-trip times for all connections are the same.

Despite having learned from the Chiu and Jain paper that additive-increase/multiplicative decrease (AIMD) is the "only good scheme", Harry decides that sources in this network should use additive increase/*additive decrease* (AIAD). In this scheme, each source will increase its sending rate additively by a small amount in response to a INCREASE message, decrease additively by a small amount on a DECREASE message, and not adjust its sending rate on a STAY message. Harry claims that the system should still converge to fairness and efficiency.

18. Harry decides to use phase-plots to check his intuition for the two-user case. The figure below shows a simple phase plot. In the graph, the state of the system is represented by a point (x1, x2), where x1 is stream 1's current rate and x2 is stream 2's current rate. The total capacity of the link is C bits/s. On the figure, for each of the labeled regions, I, II, III, and IV, draw vectors indicating the direction in which the system will move after feedback from the router.



COMMON MISTAKES: most people got this right

19. In each of the following regions, does the response to the router messages increase, reduce or keep fairness the same?

Fairness = {sum_i (x_i) }^ 2 / sum_i (x_i^2)

Region I: increase .. effective ratio increases eg: consider the case when it moves from 1,2 -> 2,3) old fairness = 9/5= 1.8 new fairness = 25/13 = 1.92 , so the fairness increases

Region II: increases fairness

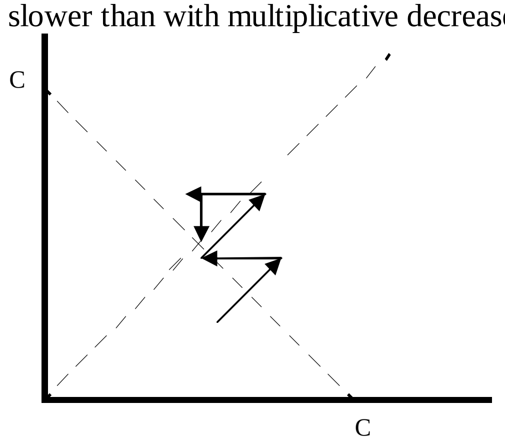
Region III: increases fairness

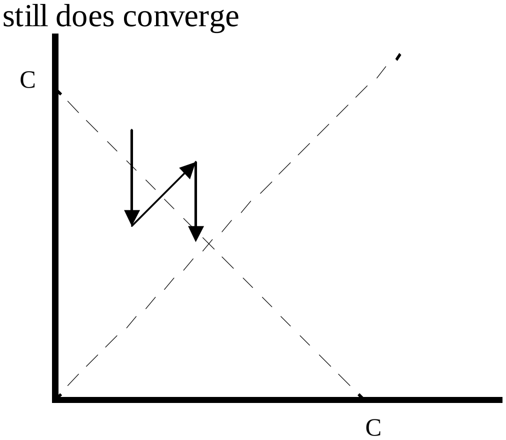
Region IV: increase .. effective ratio increases .. similar to the reasoning in Region I

On the 45 – remain same

20. Does Harry's scheme converge to a fair allocation? Explain your answer.

Yes .. From any of the four regions it is possible to reach the point of intersection.. convergence may be slower than with multiplicative decrease.. but it still does converge





Here are a couple of examples of how it may converge

F. BGP (30 pts)

21. List any 3 of the 7 BGP route selection metrics used for picking the best path to a particular destination. – should be 3 out of 6 !!

Local pref, MED, ASPATH, eBGP > iBGP, IGP cost to egress, Routerid

COMMON MISTAKES: most people got this right.

22. Which of the following statements are true about BGP route export/import policies?

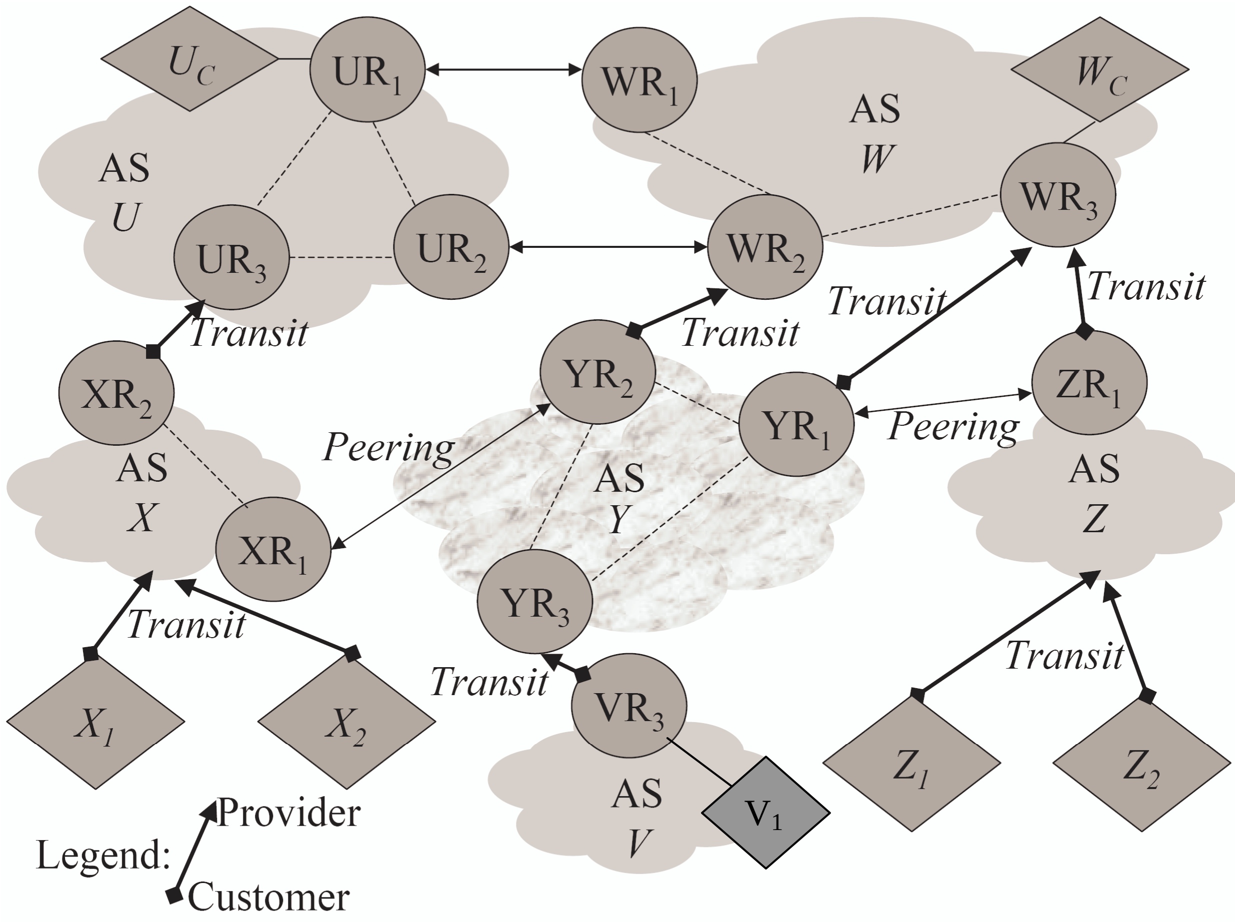
a) All sibling and customer routes are exported to peers, providers, and other customers/siblings. (+3 pts)

b) When exporting routes to a customer, an AS should export all possible routes that it learns for a particular destination (i.e. not just the route it has chosen), because the policy decision at a downstream AS may be substantially different from its decision. (-3 pts)

c) While learning routes, an AS usually prefers routes learned from peers over routes learned from providers. (+ 3 pts)

d) There may be more than one peer-to-peer edge in a in-use AS-PATH.(-3 pts)

COMMON MISTAKES: some marked b.



23. There are six AS's shown in the picture above (U, V, W, X, Y, Z). The diamond-shaped boxes are customers of the AS's (ISP's) they're connected to. Some relationships are marked as peering and transit. The circles with names like WR_1 stand for BGP routers. The first letter of such router names indicates the ISP they belong to. Within each AS, the dotted lines show IBGP connections.

In the network depicted below, circle all the paths that packets may take between a pair of clients?

a) X_1 → AS X → AS Y → AS Z → Z_1 (-2 pts)

b) X_1 → AS X → AS Y → AS W → Z → Z_1 (-2 pts)

c) X_1 → AS X → AS U → AS W → AS Z → Z_1 (+2 pts)

d) V_1 → AS V → AS Y → AS Z → Z_1 (+2 pts)

e) X_1 → AS X → AS Y → AS V → V_1 (+2 pts)

f) X_1 → AS X → AS Y → AS V → AS Y → AS Z → Z_1 (-2 pts)

COMMON MISTAKES: most got this right.

24. Which AS does not have correct IBGP interconnections?

W. It does not have a full iBGP mesh.

COMMON MISTAKES: most got this right.

25. U wants to ensure that packets sent to U_c from W are sent to it via UR_1 and not UR_2

a) Clearly explain how it might try to do this.

Use MED.
COMMON MISTAKES: do not export routes through UR2,

b) Can it always ensure that the desired behavior happens? Why or why not?

MED are not always honored by peering ISPs.

COMMON MISTAKES: Most people who got a. right also got b right.

26. W would like to ensure that packets sent to W_c from X_1 reach it via AS's X and U, and packets sent to W_c from X_2 reach it via AS Y. Can this be done with BGP? If so, how?

General guideline : you cannot guarantee control over the route selection of external ases here in fact x would prefer the peering link rather than the provider link

COMMON MISTAKES: most got this right.

THE END!!