

姓名: by P^}i^A班级: Unknown学号: N/A

《计算机网络》2007 年期末试题

以下已经预备好表格的题目请直接在试卷上填表作答，其余问题则在答题纸上回答，
试卷必须同答题纸一并交回；答题纸封面请务必也写明姓名、班级和学号。

(2007 年 12 月 30 日 19:00 – 21:00)

第一部分 性能和可扩展性

1. (本题有三个小题，共 10 分)

考虑一个文件分发的应用。服务器上有一个 100KB 的文件，要分发给 31 个接收者。假定所有主机拥有双向 40Kbps 的带宽连接到互联网上，即在它们都以这个速率工作的时候，网络不会发生拥塞；忽略传播带来的延迟以及各种协议头的开销。

(1) 如果所有接收者都直接从服务器下载该文件，则最少需要多长时间完成这个分发任务？

$$100 \text{ KB} / 40 \text{ Kbps} = 20 \text{ s}$$

$$20 \text{ s} \cdot 31 = \underline{\underline{620 \text{ s}}}$$

有同学精确区分了作为
数据量单位的 KB 和作为
带宽单位的 Kbps 之间“K”
的含义的不同，自然不矛盾
但换个角度讲，1 Mbps 来
来就是
很粗略
的近似
值

(2) 如果以 P2P 的方式工作，假如已经全部收到整个文件的接收者也可以作为数据源进一步把文件分发给其它的接收者，最少需要多长时间完成分发任务？

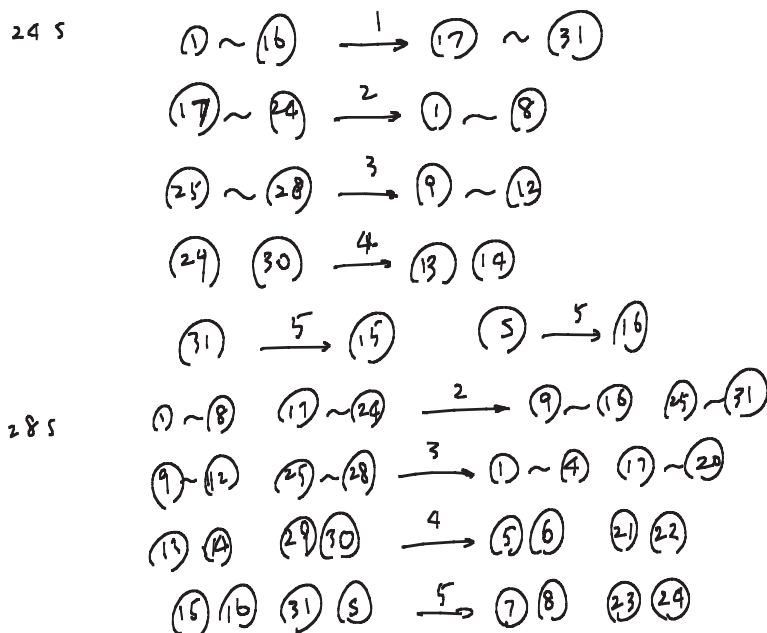
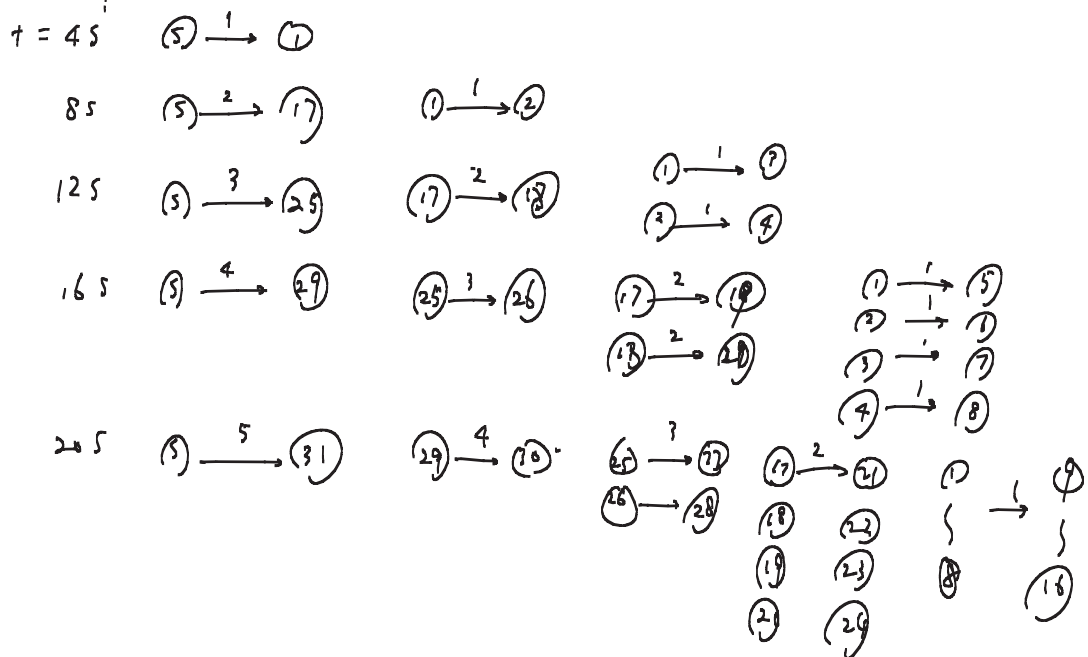
当一个接收者完成接收后，它就可以作为一个新的源
发送数据。即每经过 20s，源的数量会倍增长，经
过 5 轮，源的数量恰好达到 $1 + 2 + 4 + 8 + 16 = 31$ 个
耗时 100 s。

不少同学少算了一轮。:(

(3) 如果以 P2P 的方式工作，而且文件可以被分成 20KB 大小的单元，一旦一个接收者收到 20KB 数据，它就可以把这 20KB 作为一个整体分发给其它接收者，这样又最少需要多长时间？

如何最大程度利用带宽是这个问题之关键。一个比较简单的做法
的办法是把这个最小分发时间过程构造出来。以 $20KB/40Kbps$

$= 4s$ 为单位。下面给出一个典型的过程。



32s.

(1)~(8), (9)~(16), (17)~(24), (25)~(32)

(1)~(8), (13)~(16), (21)~(24), (29)~(32)

(5) (6) (13) (14) (21) (22) (29) (30) $\xrightarrow{4}$ (1) (2) (9) (10) (17) (18) (25) (26)

(7) (8) (15) (16) (23) (24) (31) (32) $\xrightarrow{5}$ (3) (4) (11) (12) (19) (20) (27) (28)

36s.

(1) (2) (5) (6) (9) (10) (13) (14) (17) (18) (21) (22) (25) (26) (29) (30)

(3) (4) (7) (8) (11) (12) (15) (16) (19) (20) (23) (24) (27) (28) (31) (32)

稍微联想一点在脑中，则是考虑到如下情况

① 经过 20s 后 每个接收者都可以得到一个 chunk.

② 之后在传递是 full pipe in, 所以每个接收者
都还需要 $80\text{KB}/40\text{Kbps} = 16\text{s}$.

总共的时间是 $20 + 16 = 36\text{s}$.

只要能确定 20s 后可以 fully piped 这件事情,

却是一些挑战性的.

这个题目看起来
答得很不错!

这个题目实际是 Jennifer Rexford 教授设计的. 考察学生
对于带宽知识的应用, 特别是是分析 C/S 和 P2P
的效率问题。

2. Dennis Ritchie 说: "Never underestimate the bandwidth of a station wagon full of tapes." 而今我校出口带宽受限, 每个主机允许的带宽为 1Mbps。你有一个大文件, 试图从校内传送到约 1000 公里外南京大学的另一计算机上, 可以采用 100 公里/小时的载重汽车运载或者采用互联网来传输。考虑 TCP/IP 一般情况下的协议头开销, 假设接收方窗口最大值为 64KB, 传输线路中的光速为 2×10^8 米/秒, 试计算文件多大时采用载重汽车更为优越, 列出计算过程。(5 分)

有同学分析了
sws (物理和商
口综合班) 的情况

但 sws 主要和
tcp 交互数据
流有关。对于
bulk transfer
似乎不应该有
类似问题。
需要再讨论。



$$RTT = \frac{2 \times 10^6}{2 \times 10^8} = 10^{-3} \text{ (s)}$$

$$64 \text{ KB} / 1 \text{ Mbps} = 256 \text{ ms} > 10 \text{ ms}$$

full pipelined \therefore 10 小时最多可传送

$$1 \text{ Mbps} \times 3600 \text{ s} = 4.5 \text{ GB}$$

协议头开销约 40/1500, \therefore 10 小时能传递的
最大文件是 $4.5 \times (1 - \frac{40}{1500}) = 4.38 \text{ GB}$.

第二部分 协议和体系结构

以下 a~d.是在网络中的某个节点上用 tcpdump 监听到的一段连续报文用 -vv 命令行选项处理得到的摘要, 以此为依据, 回答后面的问题 3~7。

- IP (tos 0x0, ttl 56, id 8644, offset 0, flags [+], length: 1492) 125.34.47.220 > 202.38.118.4: icmp 1472: echo request seq 0
- IP (tos 0x0, ttl 63, id 0, offset 0, flags [DF], length: 1232) 202.38.17.186 > 125.34.47.220: icmp 1212: 202.38.118.4 unreachable - need to frag (mtu 1472) for IP (tos 0x0, ttl 56, id 8644, offset 0, flags [+], length: 1492) 125.34.47.220 > 202.38.118.4: icmp 1472: echo request seq 0
- IP (tos 0x0, ttl 56, id 8644, offset 1472, flags [none], length: 56) 125.34.47.220 > 202.38.118.4: icmp
- IP (tos 0x0, ttl 56, id 8666, offset 0, flags [DF], length: 64) 125.34.47.220.50980 > 202.38.118.4.22: S [tcp sum ok] 4223701445:4223701445(0) win 65535 <mss 1452,nop,wscale 0,nop,nop,timestamp 731544427 0,sackOK,eol>

3. 前三个报文 (a~c.) 涉及一个由 UNIX 主机 125.34.47.200 发出的一个 ICMP echo request 消息, 这个消息实际是由一个 ping 命令产生的, 这个 ping 命令所指定的消息载荷是多大? (5 分)

a. 和 c. 是 125.34.47.200 发出的报文, 注意到它们的 id 相同而 offset 不一样, flag 也不相同, 所以是同 - datagram in 两个 frag. 总长度是 1492 + 56 bytes. 扣除两个 IP 头以及一个 ICMP 头, 数据载荷是 $(1492 - 20 - 8) + (56 - 20) = 1500$ byte.

表 1 (第 4 题用)

4	5	0	0	0	0	4	0	2	1	D	A	4	0	0	0
3	8	0	6	3	3	B	5	7	D	2	2	2	F	D	C
C	A	2	6	7	6	0	4	C	7	2	E	0	0	1	6
F	B	C	0	9	1	C	5	0	0	0	0	0	0	0	0
B	0	0	2	F	F	F	F	4	E	1	A	0	0	0	0
0	2	0	4	0	5	A	C	0	1	0	3	0	3	0	0
0	1	0	1	0	8	0	A	2	B	9	A	7	B	6	B
0	0	0	0	0	0	0	0	0	4	0	2	0	0	0	0

$$\begin{array}{r} 4223701445 \\ 263981340 \dots 5 \\ \hline 16498833 \dots C \\ 1031177 \dots 1 \\ \hline 64448 \dots 9 \\ \hline 4028 \dots 0 \\ \hline 251 \dots C \\ \hline 15 \dots B \\ \hline \downarrow \\ F \end{array}$$

4. 试将上面摘要 d. 翻译成用十六进制表示的 IP 分组。已知 IP 头不包含选项。上面表格包含整个分组。为简单起见, 有的信息已经给出, 请写出其它部分, 每个格子填写一个十六进制数 (请填写清晰, 明显模棱两可的字迹不算对)。 (10 分)

在表基础上, 今年这个题目只要求象征意义了。

5. 采集到这几个报文的节点是否有可能是 202.38.118.4, 为什么? (5 分)

根据 b. 只有中间节点才能给源返回 ICMP need to frag 消息。这一消息目标端是不可能收到的。(严格说有一种情况可以令该端点收到, 取巧于一种极特殊的网络连接)

6. 当分组 d. 到达 202.38.118.4 的时候, 202.38.118.4 响应了一个标志字段为 0x12 的 TCP segment, 这一事件刚刚发生的时候, 125.34.47.220 和 202.38.118.4 的相关进程分别处于什么状态? (5 分)

125.34.47.220 — SYN-SENT
202.38.118.4 — SYN-RCVD

7. (本题有两个小题, 共 10 分)

- (1) 滑动窗口的取值范围会限制端到端的最大吞吐率, 如果 125.34.49.16 和 202.38.118.4 之间的往返延迟是 500ms, 而带宽没有限制, 则在 window 字段以字节为单位的情况下, TCP 的最大吞吐率是多少?

$$64 \text{ KB} / 500 \text{ ms} = 128 \text{ KB/s}$$

- (2) 为了提高最大可能的吞吐率, TCP 可以采用 Window Scale 选项, 该选项在连接建立的时候协商一个 wscale 值, 表示整个会话中所通告窗口的实际大小是 TCP 头里 window 字段左移 wscale 位后的字节数。如果主机 202.38.118.4 在建立连接时通告 wscale = 2。那么, 它发出一个 window 字段值为 5112 的报文所通告的实际窗口大小是多少字节?

$$5112 \ll 2 = 20448 \text{ Bytes}$$

有问 5112 是十进制还是十六进制的。原则上写 0x 开头才表示十六进制。

第三部分 共享介质和网络

8. 注意到 802.11 协议里使用了 ACK 机制而有线的 Ethernet 却没有设计 ACK, 有哪些原因会导致这样的一种差别? (5 分)

有线 Ethernet 通过载波侦听和冲突检测能够确定传输

考虑到两种情况: 一种数据发送成功, 不需要 ACK。无线环境中由于隐藏站下误码率等问题导致发送方不能确知传输是否受到干扰; 不同也是阻碍暴露节点问题导致发送方对传输受到干扰的情况称赞心, 这一差别发生误判。所以无线局域网协议里 ACK 是有意义的。使得 ACK 9. (本题有两个小题, 10 分)

必要性显著不同。

(1) 考虑在一个共享介质的 CSMA/CD 网络中，存在某个节点在检测到碰撞后不做后退。那么这样一个情况对该节点本身和其它节点的性能会有什么影响？

该节点会总是在冲突后赢得传输机会
而其他节点传输会被延迟，性能下降。

(2) 考虑网络中有一个主机，在有明显的迹象表明有分组丢失的时候，其 TCP 却并不调整发送的速率，而是只考虑 advertised window 的影响。这样一个情况对该主机上的进程本身和网络中的其它主机的 TCP 性能会有什么影响？

如果丢包是拥塞造成的，不减少发送速率的节点会占用别的节点宝贵的资源，造成不公平。另一方面，过于积极的重传会使该节点本身也频繁重传，恶化网络负担，降低吞吐量。

第四部分 交换和路由：存储转发的算法和数据结构

H 和 C 之间的网络包含若干个转发设备，同时还连接有主机 K，其拓扑结构如下面图 1。回答问题 10.~12.。

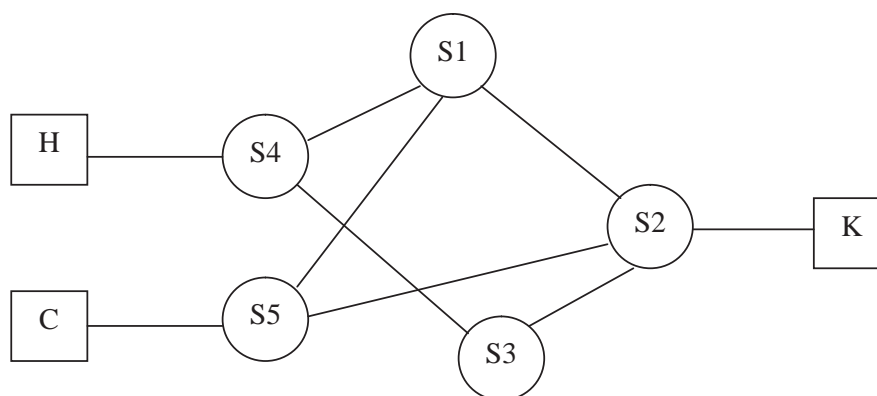
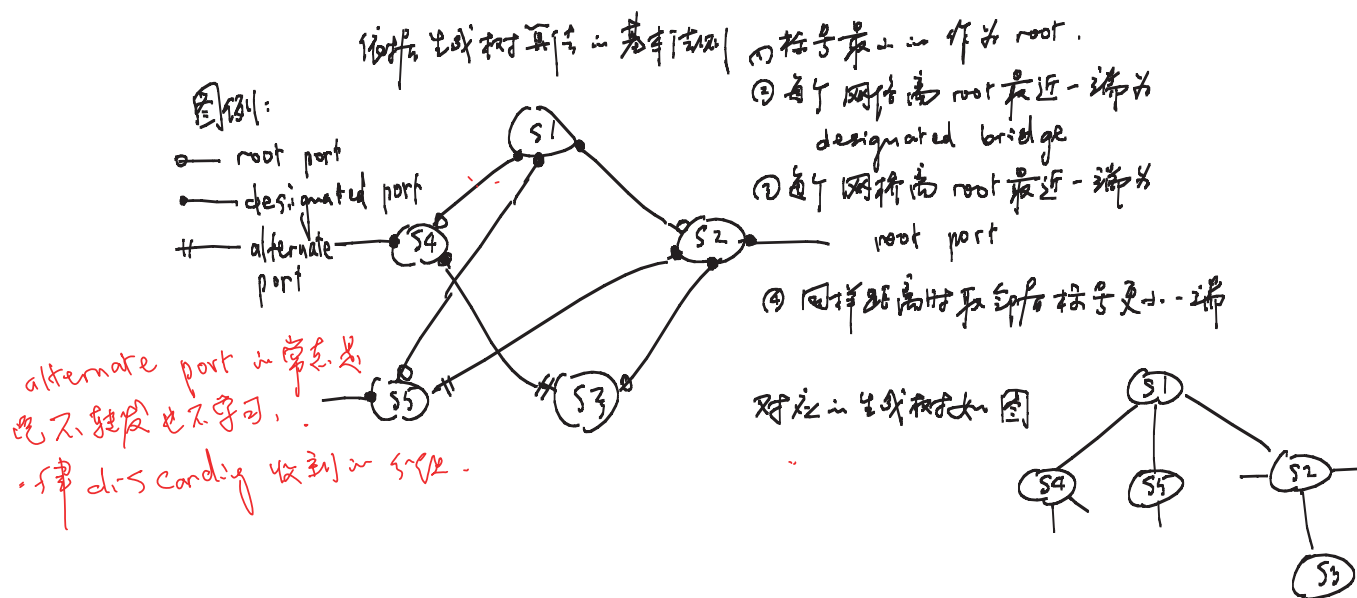


图 1

10. (本题有两个小题, 共 15 分)

假定图 1 中所有转发设备 S1~S5 都是网桥。

(1) 标出生成树选举算法收敛后所有端口的状态 (root port, designated port, alternate port), 并画出对应的生成树。



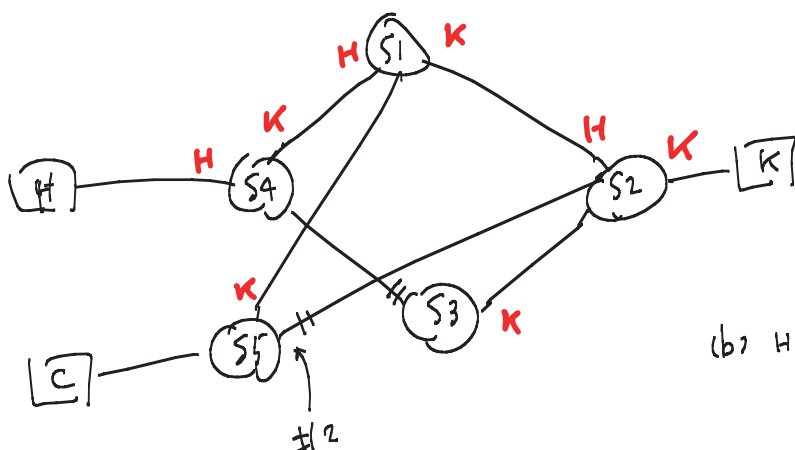
(2) 如果初始状态下所有网桥的转发表为空, 网络中依次发生如下事件序列 (a) ~ (f), 则在 S5 和 S2 之间的链路上可以听到哪些事件?

- (a) K 发送消息给 H;
- (b) H 发送消息给 K;
- (c) C 发送消息给 H;
- (d) S1 和 S2 之间的链路损坏;
- (e) H 发送消息给 C;
- (f) C 发送消息给 K。

这个题目比去年难一点, 因为中间有链路变化, 对这个变化该如何考虑呢?

有的同学把刷新不刷新都讨论到了, 自然是很全面的。

关键问题在于生成树更新后后同学得到在转发表还能用不能用, 是否需要刷新。这一点我们上课没有讲, 但是如果分析发现不刷新可能发生环路, 与生成树算法的初衷相悖, 就不难得出是要刷新的。



(a) $K \rightarrow H$

S2 floods, S5-S2 听到
 诸网桥学到 K 的方向
 注意 S5 在 H 的端口都
 能听到 K, 但 #2 处于
 alternate 状态, 不学习.

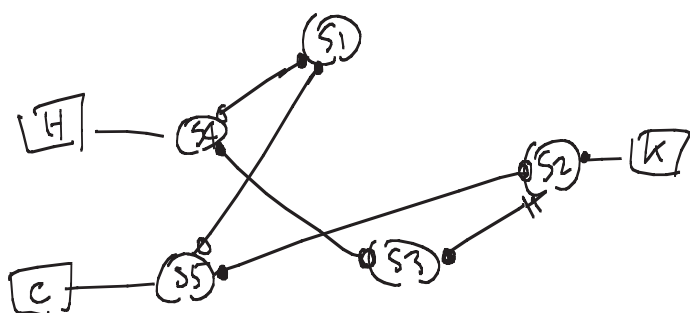
(b) $H \rightarrow K$

S4 transfers to S1
 S1 transfers to S2
 S5-S2 听不到

而 S3 不 flood
 给 S4.

(c) $C \rightarrow H$

S5 floods to S1
 S1 transfers to S4
 S5-S2 听不到



(d) S1-S2 损坏, 重新计算生成树
 端口状态变化, 注意这时所有
 转发表要清空, 如不这么做, 就
 可能出现环路, 失去生成树的意义

所以, 总结下来,
 S5-S2 链路上能听到

(a), (c), (f)

至于 (d), 可以算作“听到”
 却不是严格意义上的“听”

(e) $H \rightarrow C$

S4 floods, S1 floods.
 S5 floods, S5-S2 听到

(f) $C \rightarrow K$

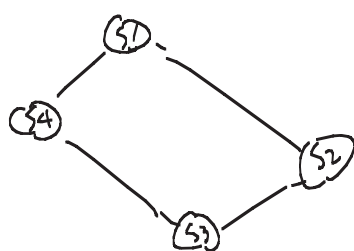
S5 floods, S5-S2 听到

11. 假定所有这些转发设备都是运行同一距离向量路由协议的路由器, 并且假定开始时已经达到收敛。考虑带“毒性逆转的水平分割” (Split Horizon with Poison Reverse) 的情况。如果在 K 到 C 发送数据的过程当中, S5 突然损坏, 试在表 2 中分析这之后 S1~S4 各路由器上到 C 的路由变化, 如果达到了收敛或者下面表格填满, 则分析中止。表项以“距离/下一跳”形式表达, 距离以 hop 为单位; 备注栏用于填写是否产生了环路、是否达到了收敛等。(15 分)

表 2 (第 11 题用)

路由器 时刻	S1	S2	S3	S4	备注
初始时刻	2/S5	2/S5	3/S2	3/S1	S5 损坏
1 个周期后	3/S2	3/S1	3/S2	3/S1	S1, S2 之间环路
2 个周期后	∞	∞	4/S2	4/S1	
3 个周期后	∞	∞	5/S4	5/S3	S3, S4 之间环路
4 个周期后	6/S4	6/S3	∞	∞	
5 个周期后	7/S2	7/S1	∞	∞	S1, S2 之间环路
6 个周期后	∞	∞	8/S2	8/S1	
7 个周期后	∞	∞	9/S4	9/S3	S3, S4 之间环路 不收敛

有些同学分析对了路由表却没有给出环路的情况。很可惜！



这一题与去年类似，同样考查 DV 算法分析问题之能力。只要思路清晰，把握 DV 算法更新路由的条件以及水平分割和毒性逆转的规则，不难得到上面的结果。

12. (本题有三个小题，共 10 分)

现在假定：S1~S5 之间都运行一个链路状态协议作为 IGP，并且它们都是自治域 AS1 的边界网关；而 H、C、K 则分别是自治域 AS2、AS3、AS4 的边界网关。边界网关之间运行 BGP 协议。如果 AS1 的政策允许其它自治域之间相互访问的流量穿越其网络，并且假定 AS1、AS2、AS3、AS4 分配得到的地址分别是 1.1.0.0/16、3.3.16.0/20、3.3.0.0/20、2.2.0.0/16。则关于目标地址 3.3.30.230，

(1) 写出 S2 的 BGP 表里对应于该目标地址的 network-prefix, next-hop 以及 AS-path 信息;

3.3.16.0/20, H, (2)

(2) 写出 S2 的转发表里对应于该目标地址的 network-prefix, next-hop 信息;

3.3.16.0/20, S1

(3) 如果这时候 S1-S2 链路故障, 上述两个表里对应于该目标地址的信息会如何变化?

BGP 表不变

转发表: 3.3.16.0/20, S3

这个题对于做 A 卷的同学来说, 应该是一个

Happy Ending! ^-^
