

## 一、术语解释

1. **Wi-Fi:** Wi-Fi (Wireless-Fidelity) 是一种可以将个人电脑、手持设备(如 pad、手机)等终端以无线方式互相连接的技术,事实上它是一个高频无线电信号。
2. **DHCP:** DHCP(Dynamic Host Configuration Protocol, 动态主机配置协议)是一个局域网的网络协议,使用 UDP 协议工作,主要有两个用途:给内部网络或网络服务供应商自动分配 IP 地址,给用户或者内部网络管理员作为对所有计算机作中央管理的手段,在 RFC 2131 中有详细的描述。DHCP 有 3 个端口,其中 UDP67 和 UDP68 为正常的 DHCP 服务端口,分别作为 DHCP Server 和 DHCP Client 的服务端口;546 号端口用于 DHCPv6 Client,而不适用于 DHCPv4,是为 DHCP failover 服务,这是需要特别开启的服务, DHCP failover 是用来做“双机热备”的。
3. **DWDM:** DWDM 就是所谓密集波分复用(Dense Wavelength Division Multiplexing)技术,指的是一种光纤数据传输技术,这一技术利用激光的波长按照比特位并行传输或者字符串行传输方式在光纤内传送数据。
4. **M/M/1/∞/FCFS (未讲):** 排队系统的记号, FCFS 代表先来先服务的服务规程,第一个 M 代表顾客到达的规律服从 Markov,第二个 M 代表服务时间分布为 Markov,1 代表服务员的数目为 1 个,∞代表系统容量限制、顾客资源数目不限。
5. **POP:** POP 互联网电子邮件协议标准是 Post Office Protocol 的缩写
6. **SNR:** (Signal/Noise) 信噪比,又称为讯噪比,即放大器的输出信号的电压与同时输出的噪声电压的比,常常用分贝数表示。
7. **SIP:** SIP (Session Initiation Protocol, 会话初始协议)是由 IETF (Internet Engineering Task Force, 因特网工程任务组)制定的多媒体通信协议。它是一个基于文本的应用层控制协议,用于创建、修改和释放一个或多个参与者的会话。广泛应用于 CS (Circuit Switched, 电路交换)、NGN (Next Generation Network, 下一代网络)以及 IMS (IP Multimedia Subsystem, IP 多媒体子系统)的网络中,可以支持并应用于语音、视频、数据等多媒体业务,同时也可以应用于 Presence (呈现)、Instant Message (即时消息)等特色业务。
8. **The Internet of Things:** 物联网是新一代信息技术的重要组成部分,也是“信息化”时代的重要发展阶段。其英文名称是:“Internet of things (IoT)”。顾名思义,物联网就是物物相连的互联网。这有两层意思:其一,物联网的核心和基础仍然是互联网,是在互联网基础上的延伸和扩展的网络;其二,其用户端延伸和扩展到了任何物品与物品之间,进行信息交换和通信,也就是物物相息。

## 二、单项填空

1. 某 PAM 采用 4 相位、2 幅度的调制算法,在 1200 Baud 的信号传输速率下能达到的数据传输率是 (B) bps。  
A. 2400    B. 3600    C. 4800    D. 9600
2. 下列表述中,正确的是 (B)  
A. 由同一二层交换机互连的所有计算机都不在同一广播域中,但在同一冲突域中。  
B. 由同一二层交换机互连的所有计算机都不在同一冲突域中,但在同一广播域中。  
C. 由同一集线器互连的所有计算机都不在同一广播域中,但在同一冲突域中。  
D. 由同一集线器互连的所有计算机都不在同一冲突域中,但在同一广播域中。

解: HUB (集线器)的所有端口都在一个冲突域和一个广播域, SWITCH (交换机)的所有端口都在一个广播域,每个端口是一个冲突域,只有在划分 VLAN 之后才能分割广播域, ROUTER (路由器)的每个端口是一个冲突域也是一个广播域。

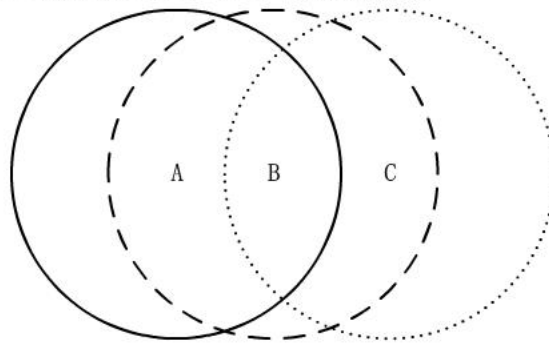
3. A、B、C、D 四个具有相同覆盖半径  $r$  ( $\frac{\sqrt{2}}{2}R \leq r < R$ ) 的无线站点,在半径为  $R$  的圆

周上依次等间隔分布。则有关站点冲突避免问题的正确表述是 (B)

- A. 对 A 来说, 其隐蔽站点是 B 与 D; 其暴露站点是 C。
- B. 对 B 来说, 其隐蔽站点是 D; 其暴露站点是 A 和 C。
- C. 对 C 来说, 其隐蔽站点是 D 和 B; 其暴露站点是 A。
- D. 对 D 来说, 其隐蔽站点是 C; 其暴露站点是 A 和 B。

解: 隐藏站点与暴露站点问题

隐藏站点: 在同一无线频率下, 彼此不能感知对方的发送信号, 这类站点互为隐藏站点。图中: A、C 互为隐藏站点, 彼此不能感知对方

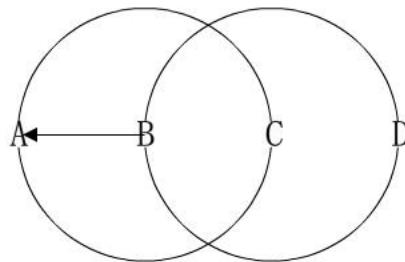


A的范围为实线, B为虚线, C为点线

A 覆盖 B, C 覆盖 B, B 覆盖 AC, 但 A 与 C 彼此不覆盖。导致: A、C 侦听不到彼此的载波 (载波侦听失效), 会出现: A、C 同时向 B 发送数据, 且彼此认为没有冲突, 实际上冲突会在 B 处发生。

隐藏终端问题: 加剧冲突, 严重影响 WLAN 的性能

暴露终端: 图中左边的圆圈对应 B 的覆盖范围, 右边的圆圈对应 C 的覆盖范围, C 和 B 互为暴露终端, 当 B 向 A 发射时, C 因侦听 B 的载波, 误认为不可以向 D 发射。其实, B 和 A 的通信并不影响 C 和 D 的通信。



暴露终端问题: 降低信道的使用效率。

4. 若发送方的 CDMA 编码是(1,-1,1,-1,1,-1,1,-1), 则输出数据 01 的对应比特串是 (D)

- A. 11111111    00000000    B. 00000000    11111111
- C. 10101010    01010101    D. 01010101    10101010

解: 向量积 (或称“叉积”) 的结果是一个向量, 点积或称“内积”的结果是“数量”, 又称“标量”。若发送 (编码) 1, 则发送自己的 m 比特码片序列, 若发送 0, 则发送码片的反码。惯例把 1 写成+1, 0 写成-1, 反码=-S, 即+1 的反码是-1, -1 的反码是+1。

5. 采用 mB/nB (m<n, 如千兆以太网的 8B/10B 和万兆以太网的 64B/66B) 编码方式的目的是, 在比特流传输中 (B)
- A. 防止过多的 1 码流    B. 消除比特序列中的直流分量
  - C. 防止过多的 0 码流    D. 防止没有冗余数据
6. 拥控是防止 (D)

- A. 发送方的发送速率比接收方的接收速率慢      B. 接收方的缓冲区溢出  
C. 发送方的缓冲区比接收方的缓冲区大      D. 发送方把过多数据发向网络

7. 华中科技大学有约 10 万师生和一个 IPV6 地址段 2001:0250:4000::/48, 每个师生约能分配到 (C) 量级的 IPV6 地址。

- A.  $10^5$     B.  $10^{10}$     C.  $10^{20}$     D.  $10^{30}$

解: IPV6 总共  $2^{128}$  为, 该 IP 地址段有 IP 数为  $2^{(128-48)}=2^{80}$  个, 华科约 10 万师生, 则每个师生能分配到约  $2^{80}/100000 \approx 10^{20}$ 。

8. 视频会议一般要使用 (C) 协议

- A. SCTP    B. RTSP、RTP    C. RTCP、RTP    D. PPLIVE

### 三、分析计算

1. 用类 C 语言编写出路由跟踪程序 TraceRouting。完成如下功能?【设最大条数 Hopmax=256, 并可忽略打印格式】

- A) 接受用户输入的目标 IP 地址, 依次打印当前跟踪到节点的 IP 地址?  
B) 依次打印从前一节点到当前节点的延时(毫秒)?

2. 嫦娥 2 号月球卫星和地球接收站相距 40 万 km, 太空电波传输速度为 30 万 km/s, 数据链路速率为 3Mbps。嫦娥 2 号每拍摄一幅 W 比特新照片(拍摄时间忽略不计)后, 一方面需要经过 X 秒钟的位移和校准时间, 才能拍摄下一幅照片; 另一方面必须同时在 X 秒钟内把刚拍摄的照片发送到地球接收站, 以便腾出缓存实施下一次拍摄。(提示 km: 千米; ps: 每秒; 且不考虑其它延时) 试求?

- A) 该太空链路的时延带宽积? 链路上每比特的宽度是多少米?  
B) 移动和校准时间 X 最多是多少秒, 照片 W 最大是多少比特?  
C) 若要拍摄每幅 W=8Mbits 的高清图片, 则数据链路速率应提高到多少?

解: 时延带宽积=bandwidth $\times$ delay。时延带宽积又称为以比特为单位的链路长度, 等于传播时延 $\times$ 带宽。

- A) 传播时延=40 万 km/30 万 km/s=4/3 s, 则时延带宽积=4/3  $\times$  3 Mbps=4Mb=0.5MB  
链路上每比特的宽度=40 万千米/4Mb=100 米。

- B) 一幅 W 比特的照片的发送时延=W/3Mbps=W/3  $\times 10^{(-6)}$  s,

3. 某台计算机有一输出文件 F Bits, 该计算机的输出链路速率是 X bps, 但受到其接入网关的令牌桶交通管制: 令牌桶始终保持 Y bps 的填充速率; 令牌桶大小是 P Bits, 假设发送开始时令牌桶已经充满, 试求?

- A) 计算机以完全速率发送的持续时间 t 之表达式?  
B) 若 X>Y, F>P, 求文件全部输出完毕所需要的时间 T 是多少?  
C) 若 X=6Mbps, Y=1Mbps, F=10.6Mbits, P=8Mbits, 求 t=?、T=?

解: 令牌桶里存放的是令牌而不是数据, 包传输之前必须获得一个令牌, 传输之后就删除这个令牌。

- A) t 时间内计算机输出数据=网关输入数据,  $tX=tY+P \times 8$ , 则  $t=P \times 8/(X-Y)$ 。  
B) 最开始计算机可以以完全速率发送数据, 因为这时令牌桶之前已有 P Bits, 同时它以 Y bps 的速率填充令牌桶即,  $tX=tY+P \times 8$ , 则  $t=P \times 8/(X-Y)$ , t 之后, 因为令牌桶里令牌已耗尽, 同时, 令牌桶填充速率 Y 小于计算机链路速率 X, 此时文件还剩余  $F \times 8 - t \times X$ , 则剩余的数据传输只能以令牌桶填充速率算即  $(F \times 8 - t \times X)/Y$ , 故总计时间  $T=t+(F \times 8 - t \times X)/Y$ 。

- C) 根据 A) B) 带入数据计算可得结果。

### 四、综合求解

1. 假设某 TCP 在时刻 t 有: 1) 发送窗口大小为 W(t) 比特 (接收端的通告窗口足够大, 即

不限制发送窗口增长); 2) 往返时间为  $R(t)$  秒, 忽略其它延迟; 3) 丢包数为  $N(t)$  个。该 TCP 采用如下拥塞控制算法: 若在一个  $R(t)$  时间内正确受到应答, 则  $W(t)$  在原基础上加 1; 若发现丢包, 则把  $W(t)$  减少到原来的  $1/2$ 。试求?

A) 表示该 TCP 窗口  $W(t)$  随时间动态变化的微分方程 (或动力学模型)?

B) 在  $R(t) = 2\tau$ ,  $N(t) = \lambda t$ , ( $\lambda$ 、 $\tau$  都是常数), 且  $w(0)=0$  的条件下,  $W(t)$  的解析解 (即求  $W(t)$  的解析函数)?

C) 画出 B) 中窗口函数的图形?

2. (未讲) 数据包以泊松方式及每秒 900 个的平均速率到达路由器, 以 10Mbps 速率输出, 并服从平均长度为 1250 字节的指数分布。【提示: 系统符合 M/M/1 排队模型, 列队长为  $n$  的概率  $p_n = (1-\rho)\rho^n$ , 且  $\rho = \lambda/\mu$ ,  $W_q = \rho/(\mu - \lambda)$ ,  $L_q = \lambda W_q$ ,  $\log 3 = 0.4771$ 】, 分别求:

A) 每个包在路由器中的平均等待时间  $W_q$ ? 路由器中的平均瞪大队长  $L_q$ ?

B) 若路由器仅有 9 个包大小的队列缓冲区, 求该缓冲区溢出的概率?

C) 若使路由器丢包率在没  $10^6$  个包丢一个包以下, 需要队列缓冲区多大?

3. 某 P2P Chord 系统采用 DHT 路由方式, 其环空间大小是  $2^3$ , ID 号从 0-7, 图中给出了后继节点表 (Succ.Table2) 的格式。在此基础上按 Chord 原理试求:

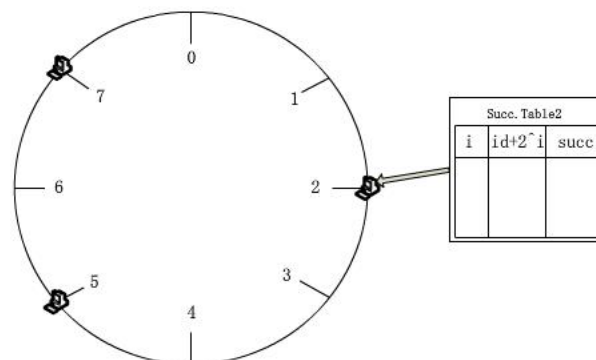
A) 节点 2 第一个加入后的 Succ.Table2?

B) 节点 5 加入后的 Succ.Table2 和 Succ.Table5?

C) 节点 7 加入后的 Succ.Table2、Succ.Table5 和 Succ.Table7?

D) 节点 5 退出后的 Succ.Table2 和 Succ.Table7?

E) 如果 Chord 上最多有  $N$  个节点, 且每个节点的路由表 (或后继节点表) 大小为  $O(\log N)$ ; 用你自己的方法求证在该 Chord 上, 由于节点加入或退出而导致的路由表更新代价是多少? (必须有证明的步骤)。



解: Chord 的目的是提供一种能在 P2P 网络快速定位资源的算法, Chord 并不关心资源是如何存储的, 只是从算法层面研究资源的取得, 因此 Chord 的 API 就简单到只有一个 set、get。Chord 是结构化的 P2P 网络。

如图所示, Chord 环有 8 个 NodeID (从 0 到 7), 最开始有三个资源, ObjectID (即 2、5、7), 也就是说只有这三个资源可被索引到, 其他 Node 节点处没有资源。

NodeID 从小到大, 顺时针排列于 1 个环上。

资源节点 (也即对象)  $k$  的后继 Successor( $k$ ) 节点: 数据对象  $k$  (即 ObjectID) 也按环上顺时针方向, 分配到资源节点  $k$  (注意是资源节点而不是 Node 节点) 或第一个比  $k$  大 ( $\text{mod } 2^m$ ) 的资源节点上。

A) 节点 2 加入,  $i$  的取值范围为 0 到  $m-1$ , 即在本题中  $i$  取值为 0、1、2, 当  $i=0$ ,  $\text{id}+2^i=2+2^0=3$ , 则 succ 即 succ(3) 为环上顺时针等于 3 或大于 3 ( $\text{mod } 2^3$ ) 的资源节点, 注意节点 2 第一个加入

时，只有节点 2 处有资源节点，其他 Node 节点处均没有资源节点，故刚开始节点 2 的 succ 均为它本身即 2；以此类推，得节点 2 的 Succ.Table2 如下：

i	$id+2^i$	succ
0	3	2
1	4	2
2	6	2

B) 节点 5 加入后，环上有两个资源节点，即 2 和 5，根据 A) 可推得：

Succ.Table2: (注意变化)

i	$id+2^i$	succ
<u>0</u>	<u>3</u>	<u>5</u>
1	4	5
2	6	2

Succ.Table5: 注意当  $id+2^i$  超过  $2^m-1$  时取余 ( $\text{mod } 2^m$ )。

i	$id+2^i$	succ
0	6	2
1	7	2
2	1	2

C) 节点 7 加入之后，当前环上存在三个资源节点，即 2、5、7，推得：

Succ.Table2:

i	$id+2^i$	succ
0	3	5
1	4	5
<u>2</u>	<u>6</u>	<u>7</u>

Succ.Table5:

i	$id+2^i$	succ
<u>0</u>	<u>6</u>	<u>7</u>
<u>1</u>	<u>7</u>	<u>7</u>
2	1	2

Succ.Table7:

i	$id+2^i$	succ
0	$((7+2^0)\text{mod}(2^3))=0$	2
1	1	2
2	3	5

D) 节点 5 退出之后，环上资源节点还剩 2、7。此时其他资源节点的后继为 5 的节点均需要更新。

Succ.Table2:

i	$id+2^i$	succ
<u>0</u>	<u>3</u>	<u>7</u>
<u>1</u>	<u>4</u>	<u>7</u>
<u>2</u>	<u>6</u>	<u>7</u>

Succ.Table7:

i	$id+2^i$	succ
0	$((7+2^0)\text{mod}(2^3))=0$	2
1	1	2



<u>2</u>	<u>3</u>	<u>7</u>
----------	----------	----------

E) 维护代价：节点的进/出，自适应到最新状态需 $O(\log^2 N)$ ，即 $O((\log N)^2)$ 。

## 2012

### 一、术语解释

1. **RTP**: 实时传输协议 RTP (Real-time Transport Protocol) 是一个网络传输协议，它是由 IETF 的多媒体传输工作小组 1996 年在 RFC 1889 中公布的，后在 RFC3550 中进行更新。
2. **EoS**: 基于 SDH 的以太网，POS (Packet Over SDH, 基于 SDH 的包交换) 和 EOS 是 SDH 网络承载 IP 业务的两种不同的实现方式。POS 技术出现得较早一些，EoS 是后来出现，技术上做了一些演进。和 POS 的主要差异是先将以太网帧进行封装后再映射到 SDH/SONET 的 VC (虚容器) 中，然后按照 SDH 的交叉映射系统处理信息。
3. **NDN**: Named Data Networking, 命名数据网络。
4. **ASON**: ASON 的概念是国际电联在 2000 年 3 月提出的，基本设想是在光传送网中引入控制平面，以实现网络资源的按需分配从而实现光网络的智能化。使未来的光传送网能发展为向任何地点和任何用户提供连接的网，成为一个由成千上万个交换接点和千万个终端构成的网络，并且是一个智能化的全自动交换的光网络。
5. **LTE**: LTE (Long Term Evolution, 长期演进) 是由 3GPP (The 3rd Generation Partnership Project, 第三代合作伙伴计划) 组织制定的 UMTS (Universal Mobile Telecommunications System, 通用移动通信系统) 技术标准的长期演进，于 2004 年 12 月在 3GPP 多伦多 TSG RAN#26 会议上正式立项并启动。
6. **POP**: 见上
7. **DHT**: DHT (Distributed Hash Table, 分布式哈希表) 类似 Tracker 的根据种子特征码返回种子信息的网络。DHT 全称叫分布式哈希表 (Distributed Hash Table), 是一种分布式存储方法。在不需要服务器的情况下，每个客户端负责一个小范围的路由，并负责存储一小部分数据，从而实现整个 DHT 网络的寻址和存储。
8. **The Internet of Things**: 见上

### 二、选择题

1. 见上 (2010 选择题 1)
2. 下列表述中，正确的是 (A) (注意与 2010 年的区分)
  - A. 由同一二层交换机互连的所有计算机都不在同一冲突域中，但在同一广播域中。
  - B. 由同一集线器互连的所有计算机都不在同一冲突域中，但在同一广播域中。
  - C. 由同一二层交换机互连的所有计算机都不在同一广播域中，但在同一冲突域中。
  - D. 由同一集线器互连的所有计算机都不在同一广播域中，但在同一冲突域中。

解：HUB (集线器) 的所有端口都在一个冲突域和一个广播域，SWITCH (交换机) 的所有端口都在一个广播域，每个端口是一个冲突域，只有在划分 VLAN 之后才能分割广播域，ROUTER (路由器) 的每个端口是一个冲突域也是一个广播域。

3. 具有相同覆盖半径  $r$  ( $\sqrt{2}R/2 \leq r < R$ ) 的 A、B、C、D 四个无线 AP，依次等间隔分布在半径为 R 的圆周上，每个 AP (接入点) 只分配 1 个 802.11n 信道 (频率)，共有 14 个可用信道，按其频率增序依次从 (01) 到 (14) 编号。则避免频率冲突的正确分配方案是 (D)。
  - A. A:(07); B:(08); C:(12); D:(13)。
  - B. A:(02); B:(03); C:(05); D:(10)。
  - C. A:(11); B:(06); C:(01); D:(06)。

D. A:(04); B:(09); C:(14); D:(08)。

解: 避免相邻 AP 信道干扰, 最多可以在同一个覆盖区域设置 4 个互不重叠的信道(相差 20MHz)。实际室内部署时相邻 AP 最少相隔 5 个信道, 所以一个覆盖区域内只能有 3 个信道可用。

4. 若发送方的 CDMA 码片为  $(-1, 1, -1, 1, -1, 1, -1, 1)$ , 则输出数据 10 的对应比特为 (D)。  
A. 11111111    00000000    B. 00000000    11111111  
C. 10101010    01010101    D. 01010101    10101010
5. 影响当前互联网路由可扩展的主要问题是 (B)  
A. IP 地址的双重语义    B. IP 地址位数不够  
C. 路由器处理能力不够大    D. 光纤传播速率不够快
6. 流控是防止 (B)  
A. 发送方的发送速率比接收方的接收速率慢    B. 发送方的发送速率比接收方的接收速率快  
C. 发送方的缓冲区比接收方的缓冲区大    D. 发送方把过多数据发向网络
7. 见 2010 年
8. 见 2010 年

### 三、分析计算

1. 嫦娥 2 号月球卫星和地球接收站相距 40 万 km, 太空电波传输速度为每秒 30 万 km, 数据链路速率为 3Mbps。嫦娥 2 号每拍摄一幅 W 比特新照片后, 立即开始向地球接收站发送已拍的照片; 一旦发送完毕一幅照片的最后 1 比特, 无需等待接收站回答就开始下一次同样的拍摄和发送。(相对太空的传播时延, 设备的发送和接收时间可忽略) 试求:  
A) 该太空链路的时延带宽积? 链路上每比特的宽度是多少米?  
B) 照片的拍摄周期 X 是多少秒? 照片尺寸 W 最大是多少比特?  
C) 若要拍摄每幅 W=50Mbits 的非压缩高清图片, 则数据链路速率应提高到多少?

解: A) 见 2010 年 A)

B) 设备的发送时延忽略, 那么整个传输过程只需考虑传播时延, 故而照片的拍摄周期即为照片的传播时延。传播时延=40 万 km/30 万 km=4/3 s。

照片的发送时延为  $W/3\text{Mbps}=W/3\text{ us}$

可知照片的发送时延必须小于等于传播时延, 即有  $W/3 \leq 4/3 \times 10^6$ , 有  $W \leq 4\text{Mbits}$ 。

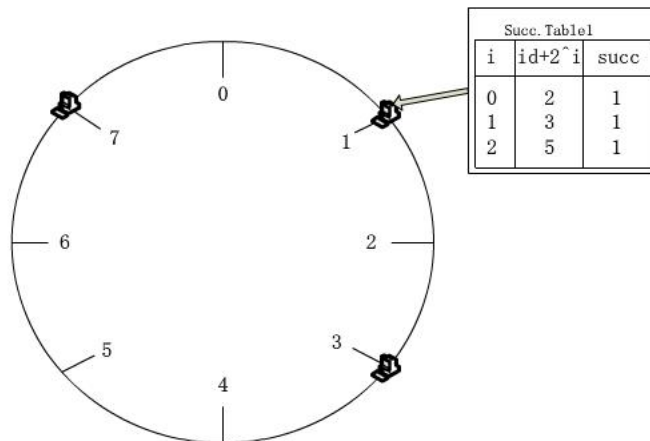
C) 当拍摄照片为 W=50Mbits 时, 设数据链路速率为 V, 则有  $50\text{Mbits}/V \leq 4/3\text{ s}$ , 得出  $V \geq 37.5\text{Mbps}$ , 故数据链路速率应提高到 37.5Mbps。

2. (未讲) 平均大小为 1250 字节的包以平均每毫秒 (ms) 80 个的速率到达交换机 M/M/1 的缓冲队列 (可缓存 n 个包), 交换机的输出吞吐率为 1Gbps 【提示:  $\log_2=0.3010$ 】, 试求:  
A) 在交换机中滞留的平均包个数? 以及包在交换机中等待的平均时间 (ms)?  
B) 若要使交换机的丢包率在  $10^{-6}$  以下, 则队列缓冲区的大小 n 应为多少?
3. 客户机通过一条速率为 R bps、往返时延为 RTT 的链路 (假设无拥塞或丢包) 向服务器读取容量为 Q 比特的网页。假设该 TCP 协议采用静态不变的拥塞窗口 (例如长度为 W 传输, 即经过三次握手后, 服务器就立即连续向客户机发送 W 个报文段, 以后服务器每当收到一个来自客户机的确认报文段, 就再向网络发送一个数据报文段, 这样反复直到所有 Q 比特发送完毕, 并假设服务器在完成当前窗口的所有传输之前就已经收到了当前窗口的第一个报文段的确认)。试求:

- A) 画出客户机与服务器的 TCP 连接及报文段数据传输的时线图
- B) 计算从客户机发起 TCP 连接到收到全部网页对象数据的时延 T
- C) 当  $RTT=100ms$ ,  $Q=125KB$ ,  $R=100Kbps$  的时延 T

#### 四、综合求解

1. 有 3 个网页 X、Y 和 Z。X 与 Y、X 与 Z 分别互相指向对方，且 Y 指的 Z；假设其阻尼系数  $d=0.5$ ，试求？
  - A) 用 PageRank 算法列出这 3 个网页的级别方程组
  - B) 用代数方程求解这 3 个网页的级别数？
2. 图中所示某 P2P 系统采用 DHT 文件路由方式，其 Chord 空间大小是  $2^3$ ，图中给出了节点 1 ( $id=1$ ) 中的后继节点表 (Succ.Table1)。在此基础上按 Chord 原理试求：
  - A) 节点 3 加入后的 Succ.Table1 和 Succ.Table3？
  - B) 节点 7 加入后的 Succ.Table1、Succ.Table3 和 Succ.Table7？
  - C) 节点 1 退出后的 Succ.Table3 和 Succ.Table7？
  - D) 如果 Chord 上最多有 N 个节点，且每个节点的路由表（或后继节点表）大小为  $O(\log N)$ ；证明在该 Chord 上找到指定文件最多只需  $O(\log N)$  跳。（必须要有证明步骤）



解：A) B) C) 见 2010 年试题。

- D) 查询代价：m 项的平均间隔  $\delta = [2^0 + 2^1 + \dots + 2^{m-1}] / m$ ；设经过 J 跳命中，应总跳数应  $\leq 2m$ ；故  $J * \delta \leq 2m$ ；故有  $J \leq m = O(\log N)$ ；一次查询的路由平均跳数为  $O(\log N)$ 。
3. 从以下三个方面论述你对互联网的认识（每个小问题不超过 300 字）
    - A) 当前互联网存在哪些主要问题？
    - B) 当前国内外下一代互联网研究前沿的动向是什么？
    - C) 你认为未来互联网将朝哪个方向发展，理由是什么？

## OpenFlow

### OpenFlow 的特点

开发者自行制造设备的方法一般是使用 PC 服务器或专用硬件搭建自己的交换路由设备，受限于主机能装备的网卡数量，这种方法不能获得足够大密度的端口（一般交换机很容易达到 48 或者更多的端口，而主机即使插上多块网卡也很难有这么多的端口），而且研究设备的交换性能一般也远不如价格的商用设备。在这种情况下，OpenFlow 论坛提出新的交换设备解决方案必须具有以下四点性质：



第一，设备必须具有商用设备的高性能和低价格的特点；

第二，设备必须能支持各种不同的研究范围；

第三，设备必须能隔绝实验流量和运行流量；

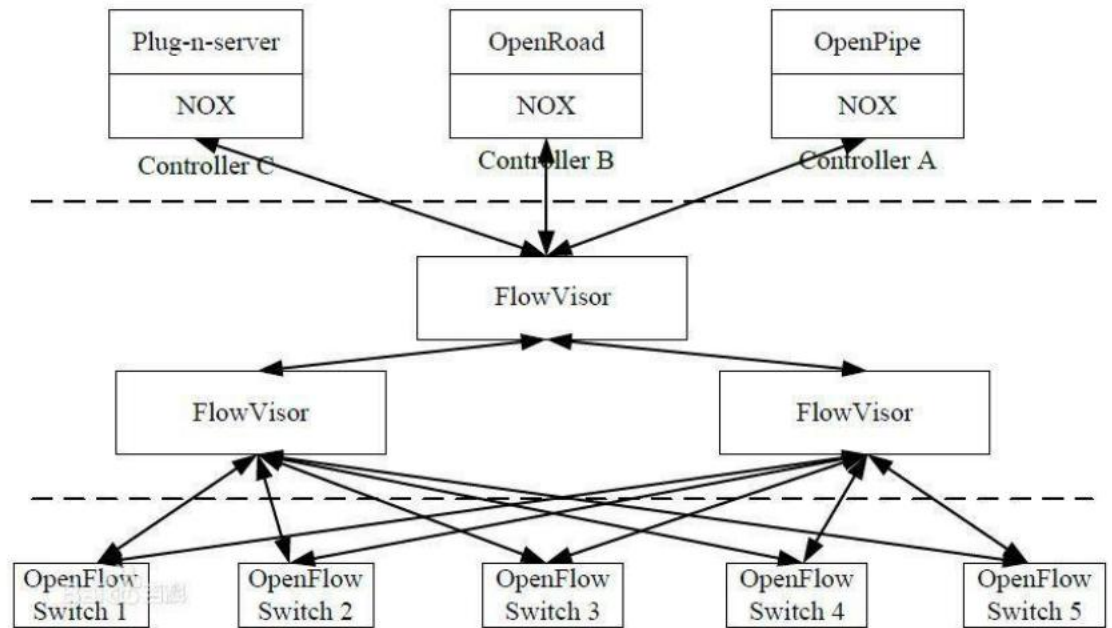
第四，设备必须满足设备制造商封闭平台的要求。

### **OpenFlow 对网络创新的影响**

由于现在的网络暴露出了越来越多的弊病以及人们对网络性能需求的提高，于是研究人员不得不把很多复杂功能加入到路由器的体系结构当中，例如 OSPF，BGP，组播，区分服务，流量工程，NAT，防火墙，MPLS 等等。这就使得路由器等交换设备越来越臃肿而且性能提升的空间越来越小。然而与网络领域的困境截然不同的是，计算机领域实现了日新月异的发展。仔细回顾计算机领域的发展，不难发现其关键在于计算机领域找到了一个简单可用的硬件底层(x86 指令集)。由于有了这样一个公用的硬件底层，所以在软件方面，不论是应用程序还是操作系统都取得了飞速的发展。很多主张重新设计计算机网络体系结构的人士认为：网络可以复制计算机领域的成功经验来解决现在网络所遇到的所有问题。在这种思想的指导下，将来的网络必将是这样的：底层的数据通路（交换机、路由器）是“哑的、简单的、最小的”，并定义一个对外开放的关于流表的公用的 API，同时采用控制器来控制整个网络。未来的研究人员就可以在控制器上自由的调用底层的 API 来编程，从而实现网络的创新。

OpenFlow 正是这种网络创新思想的强有力的推动者。OpenFlow 交换机将原来完全由交换机/路由器控制的报文转发过程转化为由 OpenFlow 交换机（OpenFlow Switch）和控制服务器（Controller）来共同完成，从而实现了数据转发和路由控制的分离。控制器可以通过事先规定好的接口操作来控制 OpenFlow 交换机中的流表，从而达到控制数据转发的目的。因此，OpenFlow 开启了一条网络创新的道路。如果 OpenFlow 得到广泛的应用和推广，则未来的网络将如曾经的计算机一样取得日新月异的发展。基于 OpenFlow 实现 SDN（Software Defined Network）。在 SDN 中，交换设备的数据转发层和控制层是分离的，因此网络协议和交换策略的升级只需要改动控制层。OpenFlow 在 OpenFlow 交换机上实现数据转发，而在控制器上实现数据的转发控制，从而实现了数据转发层和控制层的分离。基于 OpenFlow 实现 SDN，则在网络中实现了软硬件的分离以及底层硬件的虚拟化，从而为网络的发展提供了一个良好的发展平台。

OpenFlow 网络由 OpenFlowswitch（OpenFlow 交换机）、FlowVisor（网络虚拟化层）和 Controller（控制器）三部分组成。OpenFlow 交换机进行数据层的转发；FlowVisor 对网络进行虚拟化；Controller 对网络进行集中控制，实现控制层的功能。OpenFlow 网络的结构示意图如下：



## OpenFlow 交换机

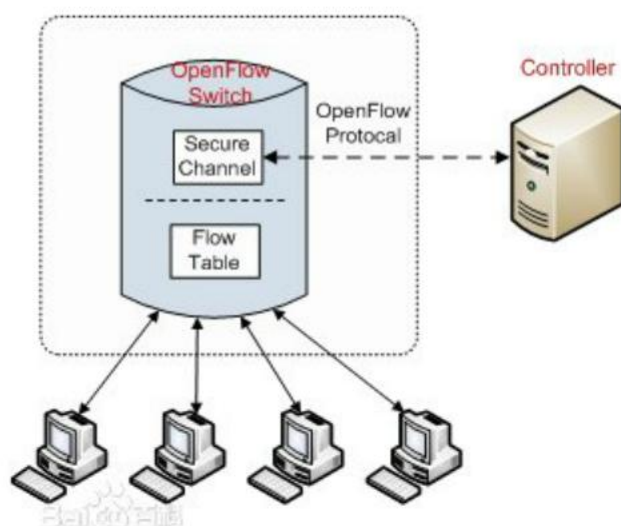
OpenFlow 交换机是整个 OpenFlow 网络的核心部件，主要管理数据层的转发。

OpenFlowSwitch 拥有一个 FlowTable（流表），它只按照流表进行转发，FlowTable 的生成、维护和下发由外置的 Controller 来实现。这里的 FlowTable 并非是指 IP 五元组（IP 源地址、IP 目的地址、协议号、源端口、目的端口），OpenFlow1.0 规范定义了包括输入端口、MAC 源地址、MAC 目的地址、以太网类型、VLANID、IP 源地址、IP 目的地址、IP 端口、TCP 源端口、TCP 目的端口在内的 10 个关键字（十元组）。FlowTable 中的每个关键字都可以通配，网络的运营商可以决定使用何种粒度的流，比如运营商只需要根据目的 IP 进行路由，那么 FlowTable 中就可以只有 IP 目的地址字段是有效的，其它全为通配。传统网络中数据包的流向是人为指定的，虽然交换机、路由器拥有控制权，却没有数据流的概念，只进行数据包级别的交换；而在 OpenFlow 网络中，统一的 Controller 取代路由，决定了所有数据包在网络中传输路径。

OpenFlow 采用控制和转发分离的架构，意味着 MAC 地址的学习由 Controller 来实现，VLAN 和基本的路由配置也由 Controller 下发给 OpenFlowSwitch。对于三层网络设备，各类路由器运行在 Controller 之上，Controller 根据需要下发给相应的路由器。当一个 Controller 同时控制多个 OpenFlowSwitch 时，它们看起来就像一个大的逻辑交换机。FlowTable 的下发可以是主动的，也可以是被动的。

**主动模式：**Controller 将自己收集的 FlowTable 信息主动下发给 OpenFlowSwitch，随后 OpenFlowSwitch 可以直接根据 FlowTable 进行转发。

**被动模式：**OpenFlowSwitch 收到数据包后，首先在本地的 FlowTable 上查找转发目标端口，如果没有匹配，则把数据包转发给 Controller，由控制层决定转发端口，并下发相应的 FlowTable。被动模式的好处是网络设备无需维护全部的 FlowTable，只有当实际的流量产生时才向 Controller 获取 FlowTable 记录并存储，当记录老化时可以删除相应的 FlowTable，故可以大大节省存储器空间。

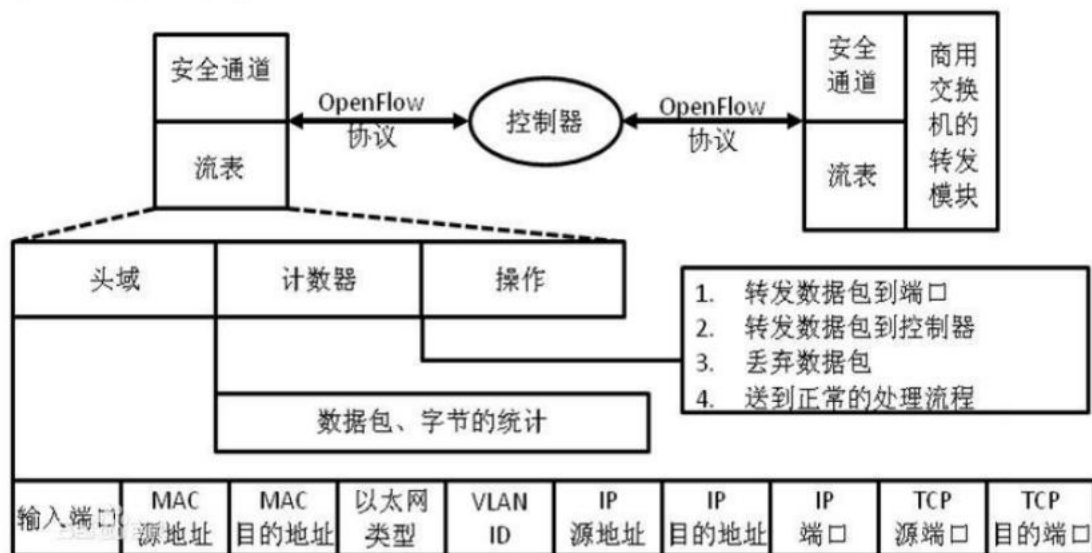


OpenFlow 交换机框架

### OpenFlow 交换机的组成

OpenFlow 交换机由 FlowTable（流表）、SecureChannel（安全通道）和 OpenFlowProtocol（协议）三部分组成。

**FlowTable 流表：**由很多个流表项组成，每个流表项就是一个转发规则。进入交换机的数据包通过查询流表来获得转发的目的端口。流表项由头域、计数器和操作组成；其中头域是个十元组，是流表项的标识；计数器用来计算流表项的统计数据；操作标明了与该流表项匹配的数据包应该执行的操作。



**Secure Channel：**安全通道是连接 OpenFlow 交换机到控制器的接口。控制器通过这个接口控制和管理交换机，同时控制器接收来自交换机的事件并向交换机发送数据包。交换机和控制器通过安全通道进行通信，而且所有的信息必须按照 OpenFlow 协议规定的格式来执行。

**OpenFlow 协议：**用来描述控制器和交换机之间交互所用信息的标准，以及控制器和交换机的接口标准。协议的核心部分是用于 OpenFlow 协议信息结构的集合。

OpenFlow 协议支持三种信息类型: Controller-to-Switch, Asynchronous 和 Symmetric, 每一个类型都有多个子类型。Controller-to-Switch 信息由控制器发起并且直接用于检测交换机的状态。Asynchronous 信息由交换机发起并通常用于更新控制器的网络事件和改变交换机的状态。Symmetric 信息可以在没有请求的情况下由控制器或交换机发起。

### OpenFlow 交换机的分类

按照对 OpenFlow 的支持程度, OpenFlow 交换机可以分为两类: 专用的 OpenFlow 交换机和支持 OpenFlow 的交换机。

专用的 OpenFlow 交换机是专门为支持 OpenFlow 而设计的。它不支持现有的商用交换机上的正常处理流程, 所有经过该交换机的数据都按照 OpenFlow 的模式进行转发。专用的 OpenFlow 交换机中不再具有控制逻辑, 因此专用的 OpenFlow 交换机是用来在端口间转发数据包的一个简单的路径部件。

支持 OpenFlow 的交换机是在商业交换机的基础上添加流表、安全通道和 OpenFlow 协议来获得了 OpenFlow 特性的交换机。其既具有常用的商业交换机的转发模块, 又具有 OpenFlow 的转发逻辑, 因此支持 OpenFlow 的交换机可以采用两种不同的方式处理接收到的数据包。

按照 OpenFlow 交换机的发展程度来分, OpenFlow 交换机也可以分为两类: “Type0”交换机和“Type1”交换机。

“Type0”交换机仅仅支持十元组以及以下四个操作: 转发这个流的数据包给一个给定的端口(或者几个端口); 压缩并转发这个流的数据包给控制器; 丢弃这个流的数据包; 通过交换机的正常处理流程来转发这个流的数据包。显然“Type0”交换机的这些功能是不能满足复杂试验要求的, 因此我们将要定义“Type1”交换机来支持更多的功能, 从而支持复杂的网络试验。“Type1”交换机将具有一个新的功能集合。

### 支持网络虚拟化的 FlowVisor

类比计算机的虚拟化, FlowVisor 就是位于硬件结构元件和软件之间的网络虚拟层。FlowVisor 允许多个控制器同时控制一台 OpenFlow 交换机, 但是每个控制器仅仅可以控制经过这个 OpenFlow 交换机的某一个虚拟网络(即 slice)。因此通过 FlowVisor 建立的试验平台可以在不影响商业流的转发速度的情况下, 允许多个网络试验在不同的虚拟网络上同时进行。FlowVisor 与一般的商用交换机是兼容的, 而不需要使用 FPGA 和网络处理器等可编程硬件。

### Controller 控制器

OpenFlow 实现了数据层和控制层的分离, 其中 OpenFlow 交换机进行数据层的转发, 而 Controller 实现了控制层的功能。Controller 通过 OpenFlow 协议这个标准接口对 OpenFlow 交换机中的流表进行控制, 从而实现对整个网络进行集中控制。Controller 的这一切功能都要通过运行 NOX 来实现, 因此 NOX 就像是 OpenFlow 网络的操作系统。此外, 在 NOX 上还可以运行 Plug-n-serve、OpenRoads 以及 OpenPipes 等应用程序。Plug-n-Serve 通过规定数据传输路径来控制网络以及服务器上的负载, 从而使得负载均衡并降低响应时间。OpenRoads 是支持 OpenFlow 无线网络移动性研究的框架。OpenPipes 可以在网络系统中通过移动每个子模块来测试每个子模块, 并可以决定如何划分设计单元。