5. 网络层

- 将数据包从发送主机传输到接收主机
- 每个主机中的网络层协议,路由器
- 路由器检查通过它的所有 IP 数据报中的报头字段

两个关键的网络层功能

路由:确定数据包从源到目标的路由——路由算法

转发: 将数据包从路由器的输入移动到适当的路由器输出

虚电路和数据报

路由器

两个关键功能

- 运行路由算法/协议 (RIP、OSPF、BGP)
- 将数据报从传入链路切换到传出链路

输入端口功能

分散交换

- 给定数据报目标,使用输入端口内存中的路由表查找输出端口
- 目标: 以"线速"完成输入端口处理
- 排队: 如果数据报到达速度比转发速率快到达交换机结构

输入端口组合的结构速度较慢 - > 排队可能发生在输入队列中,由于输入缓冲区溢出而导致的排队延迟和丢失

队头 (HOL) 阻塞: 队列前面的排队数据报阻止队列中的其他人向前移动

总线模式

交换机互联模式

输出端口

- 当通过交换机的到达速率超过输出线速度时进行缓冲
- 排队(延迟)和由于输出端口缓冲区溢出而导致有损失

路由算法

Dijkstra 算法

广播路由

泛洪:将数据包从源传送到所有其他节点

- 受控泛洪
 - 序号受控泛洪:
 - 在广播数据包中加入 ID + 广播序号
 - 反向路径转发, RPF
- 能够避免广播风暴
- 生成树广播
 - 能够避免传输冗余广播数据包

链路状态

广播路由 + Dijkstra 算法

- 让每个节点都向所有其他节点广播链路状态数据包 → 所有节点都拥有相同且完整的网络视图
- 使用 Dijkstra 算法计算从一个节点到网络中所有其他节点的最低成本路径
- 如果链路成本发生变化,则重新广播并重新计算
- 每次迭代: 需要检查所有节点, 算法复杂度较高

距离向量

基本思想

- 每个节点定期向邻居发送自己的距离向量估计值
- 当节点 x 从邻居那里收到新的 DV 估计值时,它会使用 B-F 方程 Dx (y) ← minv{c (x, v)
 +Dv (y) }更新每个节点

• , 估计值 Dx (y) 收敛于实际的最小成本 dx (y)

链路成本变化

- 节点检测本地链路成本变化
- 更新路由信息,重新计算距离向量
- 如果 DV 发生变化,则通知邻居

好消息传播得很快,坏消息传播得很慢-无穷计数问题—毒性反转

分层路由

管理自主权 Internet = 网络的网络 每个网络管理员可能都希望控制自己网络中的路由

自治系统

将路由器聚合到多个区域,即"自治系统"(AS)

- 同一 AS 中的路由器运行相同的路由协议 "AS 内"路由协议
- 不同 AS 中的路由器可以运行不同的 AS 内路由协议

Internet的路由

- 两级路由
 - Intra-AS: 管理员负责选择
 - Inter-AS: 唯一标准

LS和DV的复杂度

- 消息复杂度
 - 。 LS: 有 n 个节点, E 链路, 发送 O (nE) 消息
 - 。 DV: 仅在邻居之间交换
- 收敛速度
 - 。 LS: O (n2) 算法需要 O (nE) 消息
 - 可能有振荡
 - 。 DV: 收敛时间不同

- 可能有无穷计数问题
- 鲁棒性——如果路由器出现故障怎么办?
 - o LS
 - 节点可能公布不正确的链路成本
 - 每个节点只计算自己的表
 - o DV
 - DV 节点可能公布不正确的路径成本
 - 其他节点使用每个节点的表——错误累积

Internet 中的路由

Intra-AS Routing

也称为**内部网关协议 (IGP)**

- 最常见的 AS 内部路由协议
 - 。 RIP: 路由信息协议
 - 。 OSPF: 开放最短路径优先
 - IGRP: 内部网关路由协议
- RIP
 - 。 DV算法
 - 。 每30s交换一次
 - 。 路由最多可到25个网络
 - 。 运行UDP数据包协议
- OSPF
 - 。 LS算法
 - IP数据包
- IGRP (内部网关路由协议)

- 。 DV算法
- o TCP数据包

Internet inter-AS routing

BGP (边界网关协议)

- 路径向量协议——和DV算法相似
- 最短路径算法
- 使用 TCP 交换的 BGP 消息

BGP信息

- OPEN
- UPDATE
- KEEPALIVE
- NOTIFICATION

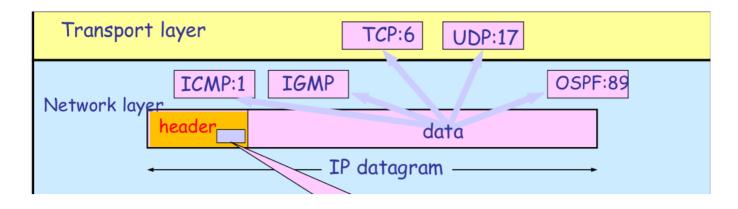
IP: 互联网协议

IP数据报格式

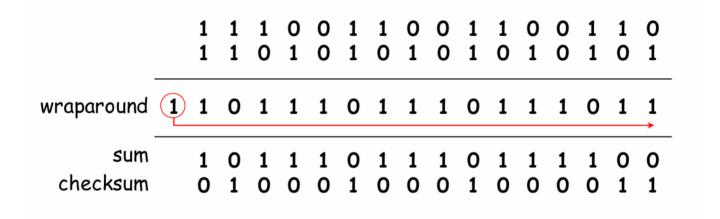
IP 协议版本号	Head len 以4	数据类型	总数据报的长度			
	个字节为单位					
16bit 的表示			Flags 3 位	Fragment offset 13 位		
Time to live		上层协议	头部校验和			
32bits 的源 IP 地址						
32bits 的目的 IP 地址						
可选:例如:时间戳、记录所走的路由、指定要访问的路由器列表。						
data (可变长度,通常为 TCP 或 UDP 段)						

上层协议字段

该字段指示数据应**传递到哪个上层协议**



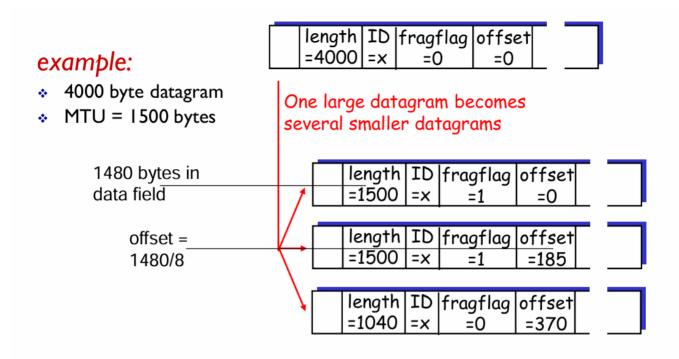
头部校验和



进位放后面, 然后再取反

IP 分段

MTU: 最大传输长度



flag=0: 表示分帧的最后一帧

IP地址

部分基础自己看ppt

网络号	主机号	地址类型	用途
全0	全0	本机	启动时使用
网络ID	全0	网络	标识一个网络
网络ID	全1	直接广播	在指定网上广播
全1	全1	有限广播	在本地网上广播
127	任意	回送	环回测试(127.0.0.1)

无类别域间路由

• 两级地址, 只有两部分

• 地址格式: a.b.c.d/x, 其中 x 是地址网络部分, 网络部分通常称为前缀

DHCP: 动态主机配置协

动态获取地址:"即插即用"

- Discover
- Offer
- Request
- ACK

分层寻址: 路由聚合

NAT: 网络地址转换

动机: 就外部世界而言, 本地网络只使用一个 IP 地址

LAN 地址和 ARP

LAN 上的每个 IP 节点 (主机、路由器) 都有 ARP 模块,表

ARP 表:某些 LAN 节点的 IP/MAC 地址映射

ARP协议

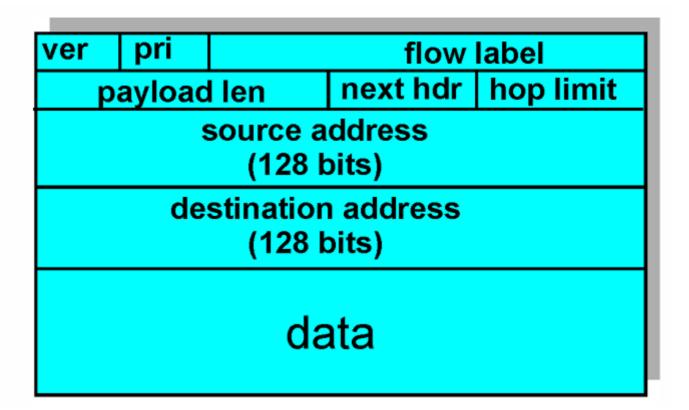
ICMP

Internet Control Message Protocol 互联网控制消息协议

- ICMP 是 IP 的控制同级
- ICMP 由 IP 使用,并使用 IP 作为网络层协议
- ICMP 用于 ping、traceroute 和路径 MTU 发现

IPV6

- IPv6 数据报格式
 - 。 固定长度 40 字节标头
 - 。 不允许分段
- Priority: 确定流中数据报的优先级流标签 Flow Label:识别同一"流"中的数据报 Next header: 识别数据的上层协议



• 其他变化

校验和: 完全删除以减少每个跃点的处理时间

选项:允许,但在报头之外,由"下一个报头"字段指示

ICMPv6: 新版本的 ICMP

其他消息类型,例如"数据包太大" 多播组管理功能

• 地址的三种类型

三种类型: 单播、多播、任播

From IPv4 To IPv6

提出的两种方法

- 双堆栈
 - 一些具有双栈 (V6、V4) 的路由器可以在格式之间"转换"
- 隧道
 - IPv6 在 IPv4 路由器之间作为 IPv4 数据报的负载传输