Chapter 3 internetworking.

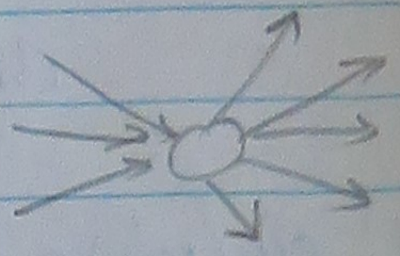
3.1 switching & bridging
- .1 datagram
- .2 virtual circuit switching.
- .3 source routing.
- .4 bridges & LAN Switches

3.2 basic internetworking IP
- .1 what is internetwork
- .2 service model
- .3 global addresses
- .4 datagram forwarding in IP
- .5 subnetting & classless addressing
- ✓ .6 address translation. (ARP)
- ✓ .7 host configuration. (DHCP)
- ✓ .8 error reporting (ICMP)
- ✓ .9 virtual networks & tunnels

3.3 routing
- ✓ .1 network as a graph.
- .2 distance-vector (RIP)
- ✓ .3 link state (OSPF)
- ✓ .4 metrics

✓ 3.4 implementation & performance
- .1 switch basics
- .2 ports
- .3 fabrics
- .4 router implementation.

3.5 summary



3.1

switch
- 定义: a device with multi input & multi output.
- 功能: switching or forwarding.
  - receive pkt & transmit to another link.
- how to decide to put pkt on which output link?
  - datagram (connectionless) approach
  - virtual circuit (connection oriented) approach.
  - source routing.

3.1.1 datagram.

forwarding table for switch 2

dest	port
⋮	⋮

定义: every pkt contans complete dest address

特点: ① connectionless : a host can send a pkt anywhere at any time.
(对比 connection-oriented: connection state need to be established before 1st pkt is sent)
② robustness to failures : good
└ a link or a switch failuer ~~不会~~ ⇒ serious effect.

3.1.2 virtual circuit switching

~~表~~ VC Table entry for swich 1

incoming interface	incoming VCI	outgoing interface	outgoing VCI.
...	...	...	...

virtual circuit identifier (每一个 link 有一个 VCI)

connection-oriented: set up virtual connection from source host to dest host, before all (1st) data is sent.
- ✓ stage ① connection set up
  - by administrator.
    - └ PVC (permanent)
  - by source node (想发送 data 给 host)
    - SVC (switched / signaled)
    - using signalling
      - └ 发送 "setup" message
- ✓ stage ② data transfer.
- stage ③ tear down connection
  - 当 A 不想再传 data 给 B.
  - 或 if a switch/link fails

(自己加的)

VCS: · if a switch / link fails ⟹ 有很大影响.
└ connection is broken.

解决方案 ─┬ ① tear down old connection
├ ② ~~clear~~ free up storage space
└ ③ set up new connection.

~~3.1.3~~

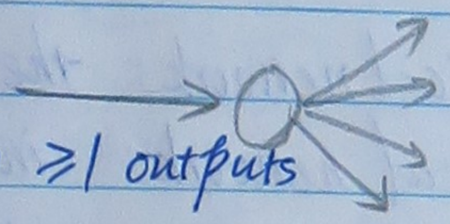
3.1.3 source routing.

定义: all info about network topology is provided by source host.
↓ to switch a pkt accross a network.

3.4.1 bridges & LAN switches.

定义: LAN switches: forward pkts between LANs (local area networks)
(switches that forward pkts between LANs)

bridge : a kind of switch.
transfer pkt from one input to ≥1 outputs



extended LAN: a collection of LANs connected by ≥1个bridges

learning bridges: ┬ learn source address of incoming frames
└ use it to choose on which port to send outgoing frames.

flood: forward the pkt to all ports 除了此pkt已经到过的 ports.

TTL: time to live.

MAC address: media access control address

~~span~~

does the learning bridge algorithm work in an extended LAN with multiple learning bridges? No ∵ frames will loop.

spanning tree algorithm.

定义: 自动 switch off some ports 使得 no loops in the network remain
but every node remains reachable.
即 reduce network topology to ⟹ a spanning tree.

root: ~~smalle~~ the switch/bridge with smallest ID.

计算从每个LAN到 root 的 cost　　cost: 经由几个 switch, 就有几个单位的cost

比较 cost 的大小.
├ 若两 costs 不相等, ⟹ 选 min cost 的 path.
└ ······ 相等 ⟹ 比较 switch 的 ID数值 ⟹ 选 min ID 的 path.

3.2 basic internetworking IP

- internet i : collection of networks interconneeted to provide some sort of host-to-host pkt delivery service
- Internet I : globale internetwork. large.

- networks: 或 < a directly connected network / a switched network. } use technology → 802.11 / Ethernet
- subnetworks = subnets
- internetworks : network of networks ∴ made up of lots of smaller networks

- physical network: the underlying network that we are interconnecting
- logical network: internet is a logical network built out of a collection of physical networks.

常见的 single-technology network:
- Ethernet network
- wireless network
- point to point link network

routers: nodes that interconnect the networks
也称 gateways.

internet protocol (IP)
- a key protocol to build scalable, 合成的 internets
- also used in Internet
- runs in all routers & hosts

3.2.2 IP service model

定义: host-to-host service you want to provide.

service model has 2 parts
- an address scheme
  - a way to identify all hosts in internetwork
- datagram (connectionless) model of data delivery.
  - besteffort. or unreliable senvice
    - pkts my be lost
    - . . . . delayed
    - . . . . delivered out of order
    - . . . . duplicated.

"more" flag = { 1 this is not the final fragment; 0 is the final fragment }

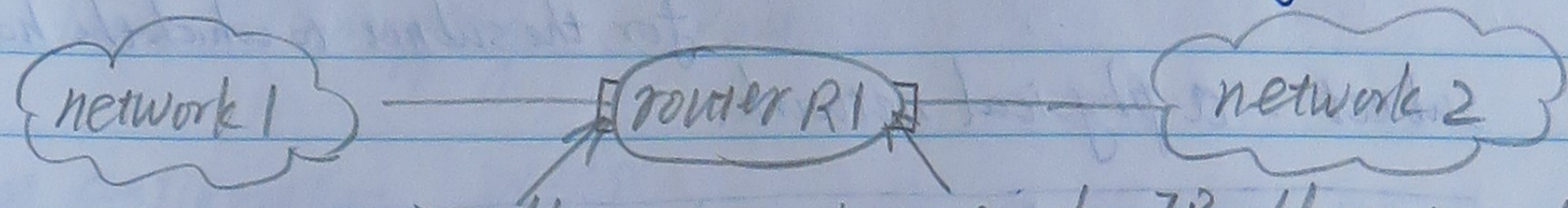
MTU = largest IP datagram that it can carry in a frame

offset { 1 offset = 8 bytes; =0 if this is 1st fragment of a frame; ≠0 . . . . not 1st . . . . }

fragmentation & reassembly
- pkt format
- "more" flag = { 1 has more; 0 this is final }
- MTU = max transmission unit.
- offset { 1 offset = 8 bytes; 1st 部分 of a datagram's offset = 0 }

3.2.3 global addresses (IPv4)

properties of IPv4:
- no 2 hosts have same ~~IP~~ address ⇒ IP uniqueness
- hierarchical: 2 parts
  - network part: which network the host is attached
  - host ... : identify each host uniquely on that particular network
- 32 bits ⇒ 2^{32} = 4 billion IP addresses
- routers (& hosts) should have ~~IP~~ one IP address for each interface.



original format
class A, B, C, D: multicast

a decimal integer
↑
a byte
↑
8 bits. 8 bits. 8 bits. 8 bits

dotted decimal notation. 171.69.210.245
└ each integer (十进制数) represents: decimal value contained in 1 byte

3.2.4 datagram forwarding in IP.

forwarding table for ~~a~~ router R2.

~~dest~~ Network Num	~~port~~ NextHop.
⋮	⋮

- every IP datagram contains dest IP address
- each router maintains a forwarding table.
- router 先已知 forwarding table. 而后, 当 router 收到 pkt, router 根据 pkt 自带的起发&终点地址, 与 forwarding table 比对, ⇒ router 知道该把 pkt ~~发~~ 从哪个 hop (=port) 发出去
- ~~if there~~ 例: for router R2
  if (R2 的某一个 interface 的 IP 地址的 network part
  == pkt 的 dest 地址的 network part) 即与 dest network 直接相连
  then forward to the dest host
  else,
  then forward to next-hop router.

3.2.5

subnetting & classless addressing

add another level to address hierarchy: subnet

subnet number: all hosts on the same physical network have the same subnet number.

subnet mask: to introduce a subnet ~~number~~ #

subnetting: a host is configured with both {an IP address and {a subnet mask} for the subnet to which the host is attached

subnet: a physical network.

subnetted address:	network #	subnet ID	host ID
	大范围	小范围	最小的单位

计算:

a host's IP address : 128.96.34.15

AND

a host's subnet mask : 255.255.255.128

---

subnet number : 128.96.34.0 ~~比较计算出的~~ subnet # 和 forwarding table 中的 subnet #, 若相等则说明在同一个 subnet 上

{ internetwork 由许多个 network 组成.

network ..... subnet ....

subnet 包含 ... hosts.

subnetting: 用于判断两个 hosts 是否在同一个 subnet 上.

- 若在 → 直接传送 pkt
- 不在 → 传 pkt 到 default router 上.

forwarding table with subnetting :

~~dest~~ subnet #	subnet mask	~~port~~ nexthop
...	...	...

Classless addressing

provider ——→ customer.

大范围 小范围

例 /21 /24

3.2.6 address translation ARP.

enable each host on a network to buid up a table of mapping between IP addresses & link-level addresses

IP addresses → used in IP datagram

link-level addresses → used in physical interface hardware

IP MAC address

mapping < IP 地址, MAC 地址, TTL >

learn the mapping : P229.

3.2.7 host configuration DHCP.

goal: when a host joins network, DHCP allows this host to dynamically obtain its IP address from network server.

可以 - allow reuse of 地址
- 更新旧地址的租使用期限.

DHCP can return 超 more than just *allocated IP 地址 on subnet:

- network mask
- name & IP 地址 of DNS server
- default router

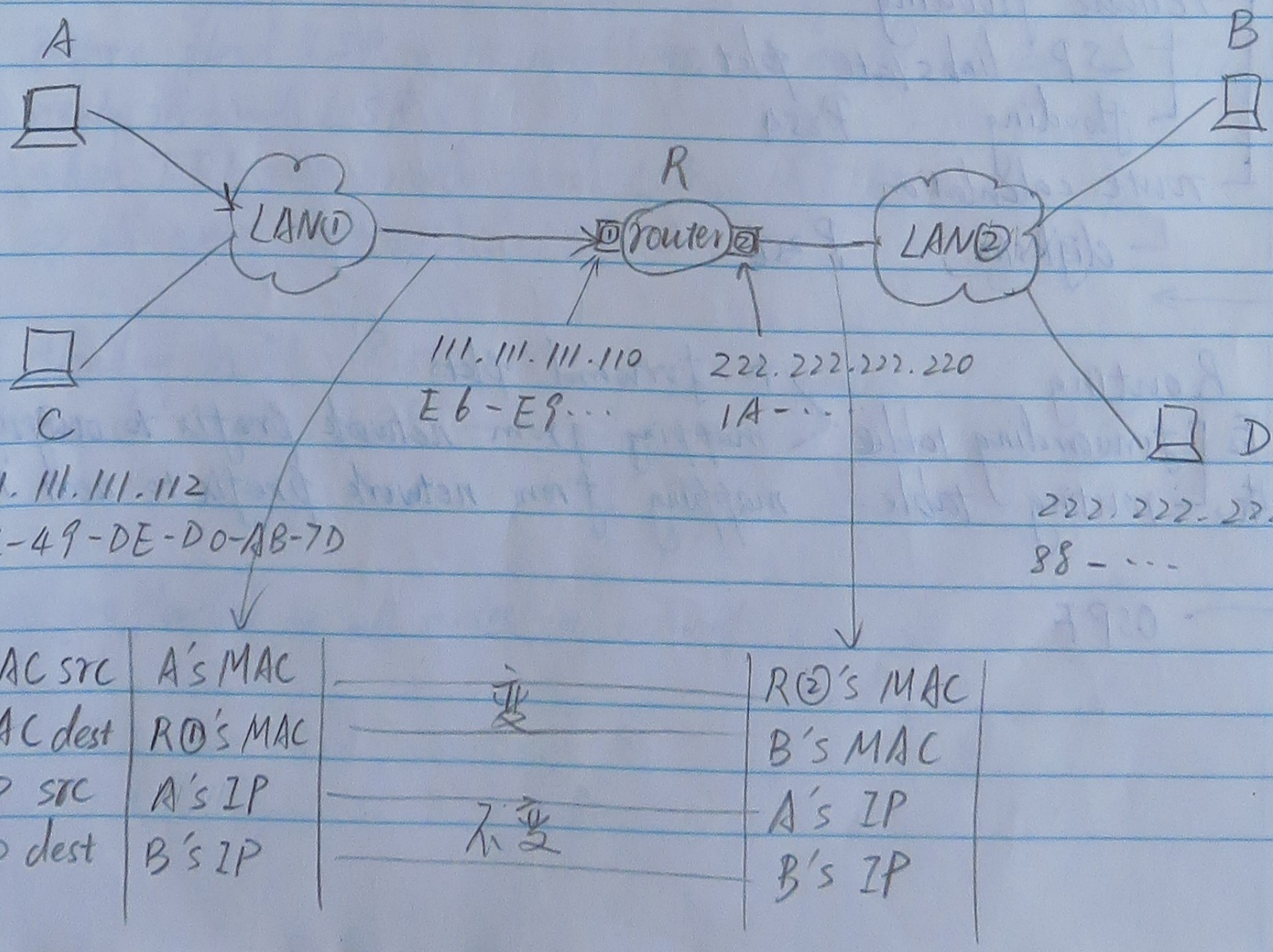
difference IP 地址 & MAC 地址 when transmiting a pkt.

- 与 host
- 与 interface of router

有且仅有一个 unique global 地址

IP : 111.111.111.111

MAC: 74-29-94-E8-FF-55



MAC src	A's MAC	变	R②'s MAC
MAC dest	R①'s MAC		B's MAC
IP src	A's IP	不变	A's IP
IP dest	B's IP		B's IP

9.3.1 DNS domain name system

goal: to map between [IP 地址] & [host (or server) name]

=resolve= translate used for 传输 pkt

例子 www.utwente.nl

used by humans

DNS: distributed database 数据库 which stores 存储 resource records RR

host —DNS query→ local DNS server.

each ISP internet service provider has a local DNS server.

DNS name resolution

3.3.1 Network as a Graph.

basic problem of routing: find lowest-cost path between any 2 nodes.

反义词 distributed

centralized.

2 classes of routing protocols — [distance vector RIP
link state OSPF.

3.3.3 link state.

- two mechanisms — reliable dissemination of link-state info
  - calculation of routes from the sum of all accumulated link state knowledge
- reliable flooding.
  - LSP link state pkt.
  - flooding P254
- route calculation
  - dijkstra P256

3.3 Routing

反义词 [forwarding table < to forward pkts
mapping from network prefix to outgoing interface
routing table. mapping from network prefix to next hops

OSPF

3.3.3 reliable flooding.

when to generate new LSP
- when TTL=0
- or when topology is changed 即 a ~~neto~~ router / link fails.

how to detect failure ?
- each node 每隔 · ~~time~~ interval —"hello" msg→ its neighbour router.
  - if no failure, each node receives ←reply - - - - -
  - if there is failure, · · · can not receive reply in long time

sequence #: to make sure that old info is replaced by new info.
每当 a node generate a new LSP, it 1' LSP's seq # by 1. (periodically)

1) if a node was down for a long time (failure)
   then all old LSPs for that node will time out.
2) if ~~this~~ node 一切正常
   then { this node receives a copy of its own LSP
   且 seq # of received LSP is higher.
3) if a node goes down & comes back up
   then seq # =0 ← initialization.

TTL time to live: to ensure that old link state info is eventually removed from network.
- Before flood LSP to its neighbours, a node always ↓ TTL of a newly received LSP.
- when TTL =0, node flood LSP with TTL=0
  ⇒ all nodes delete this LSP.

flooding 机制 P254.

1) each node creates LSP ——
   - ID of node
   - cost
   - seq #
   - TTL
2) LSP are flooded to all nodes in network.

node X receives LSP from node Y., X 执行:

① "do I have a stored a copy of Y's LSP ?"
- if No → stores LSP
- if Yes → compare seq #s.
  - if new LSP seq # of Y > old LSP seq # of Y ↴
    - → store new LSP means new LSP is most recent
    - → remove old LSP
    - → send a copy of the new LSP to all neighbours
  - if new seq ⩽ old seq ...
    - → discard this new LSP
    - → means old LSP is most recent

9.3.1 DNS name resolution.

resolution mechanism: when 调用 invoked with a name, returns the corresponding value.

name server → a specific implementation of a resolution mechanism
- can be queried 请求 by sending a message.

name space: a set of possible names.
- flat : not divisible into components.
- hierarchical : divisible into components. (如 DNS)
- fundamental unit of implementation in DNS

naming system: maintains a collection of bindings of names to values.

name resolution: 1) look up a host's name in the local copy of table
2) return corresponding address.

DNS: map domain names into values.

in hierarchy, each node in the tree 对应于 a domain.

(RR) a resource records:
- each name server implements the zone info as a collection of resource records.
- a name-to-value binding
- RR format <Name, Value, Type, Class, TTL>

zone →
- subtree in the hierarchy
- 对应于 some administrative authority (that is responsible for that portion of the hierarchy)

client —queries (I want to get IP of domain name "...")→ name server
client ←returns all records that are related to this record (domain name "..." 的 IP 地址 is ...)— name server