网络知识十全大补丸

刘楠

目录

网络硬件介绍

- 双绞线(twisted pair)，绝缘铜导线两两绞合在一起聚合而成的线缆

类型	五类线CAT5	超五类线CAT5e	六类线CAT6	超六类线CAT6e
最大带宽	100Mbps	1000Mbps	1000Mbps	10Gbps
最大长度	100米	100米	100米	55米

- 双绞线线序定义

EIA/TIA 568A的线序定义依次为绿白、绿、橙白、蓝、蓝白、橙、棕白、棕，其标号如下表所示：

绿白	绿	橙白	蓝	蓝白	橙	棕白	棕
1	2	3	4	5	6	7	8

EIA/TIA 568B的线序定义依次为橙白、橙、绿白、蓝、蓝白、绿、棕白、棕，其标号如下表所示：

橙白	橙	绿白	蓝	蓝白	绿	棕白	棕
1	2	3	4	5	6	7	8

- **集线器 Hub**　工作在物理层，复制转发多路数字信号，外形和交换机非常像，逐步被低端交换机淘汰

- **以太网供电PoE(Power over Ethernet)**，通过网线给网络设备供电，比如IP电话机、无线AP、部分交换机可以通过网线供电不需要额外电源

- **交换机 Switch**，包括用于小型局域网的二层交换机和大型局域网的三层交换机

- **网桥 Bridge**，与交换机非常类似，工作于二层网络用于连接多个局域网，通常端口数量较少，常见于虚拟网桥应用场景

- **路由器 Router**，工作于三层用于连接不同的以太网，通过路由协议为报文提供路由

- **防火墙 Firewall**，用于网络之间的风险隔离，传统防火墙工作于三层，更前沿的防火墙可以工作在四层和七层，如支持DPI的深度包检测防火墙

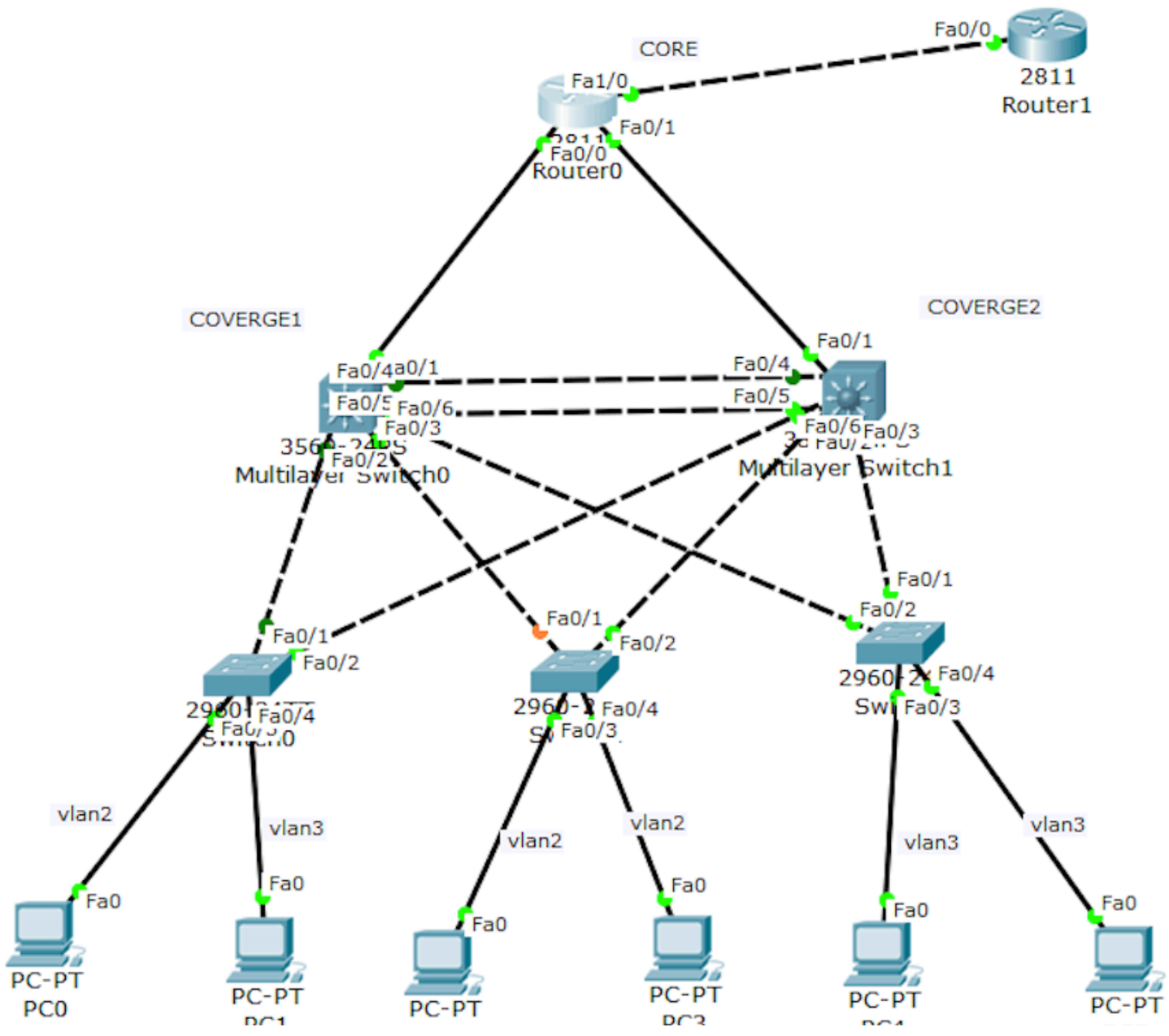
- **虚拟专用网 VPN**，在公网上架设的虚拟加密通信网络，常见的两种形式是：终端远程接入私有网络，两个私有网络的互联互通

- **网关 Gateway**，网络的出口设备，工作于三层，可以集合路由、防火墙、流控、VPN等功能与一体

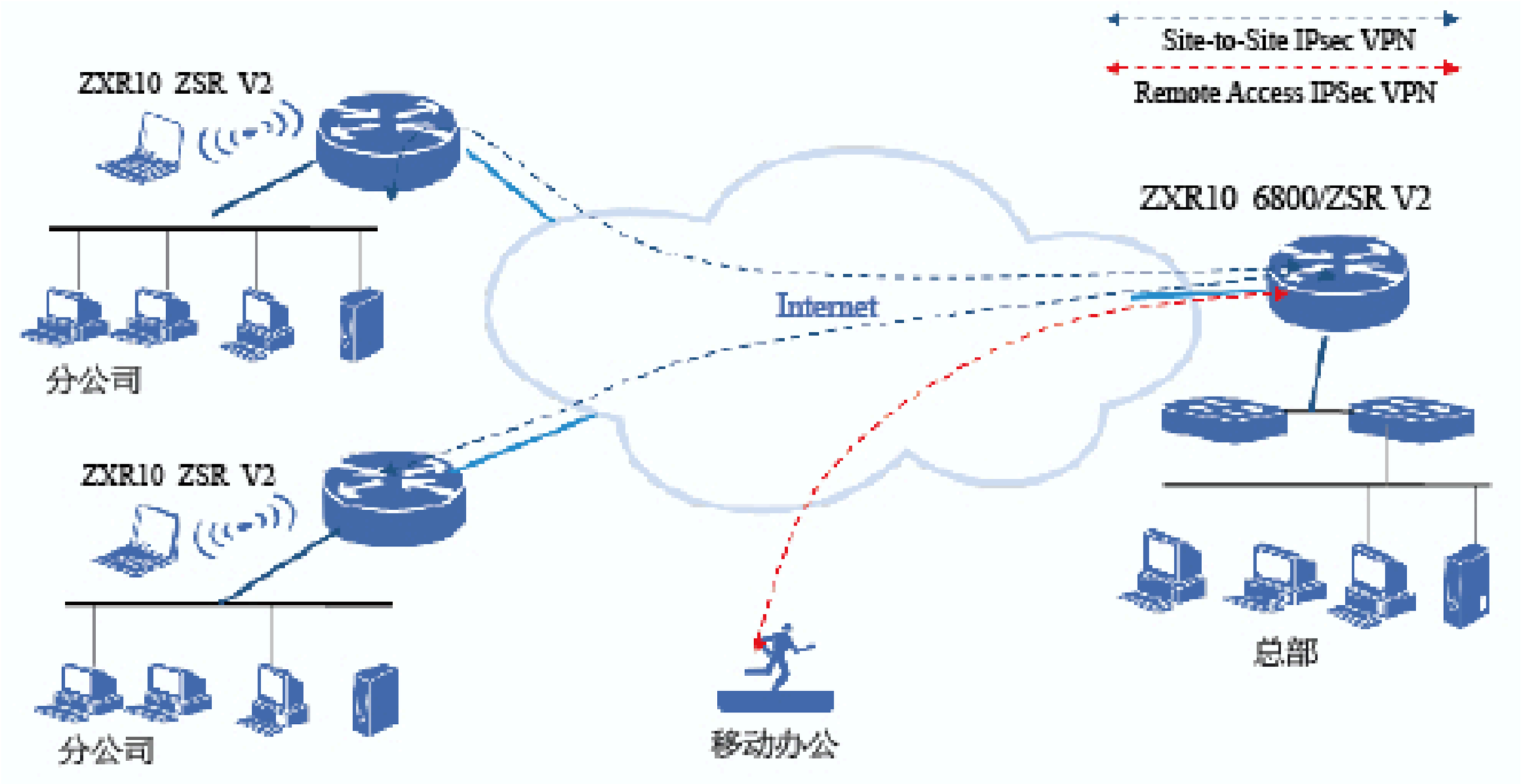
- **无线接入点 AP**，运行WiFi系列协议802.11.x的二层网络设备，相当于无线网络的交换机，可以支持移动终端在多个接入点之间无缝漫游

网络部署模式

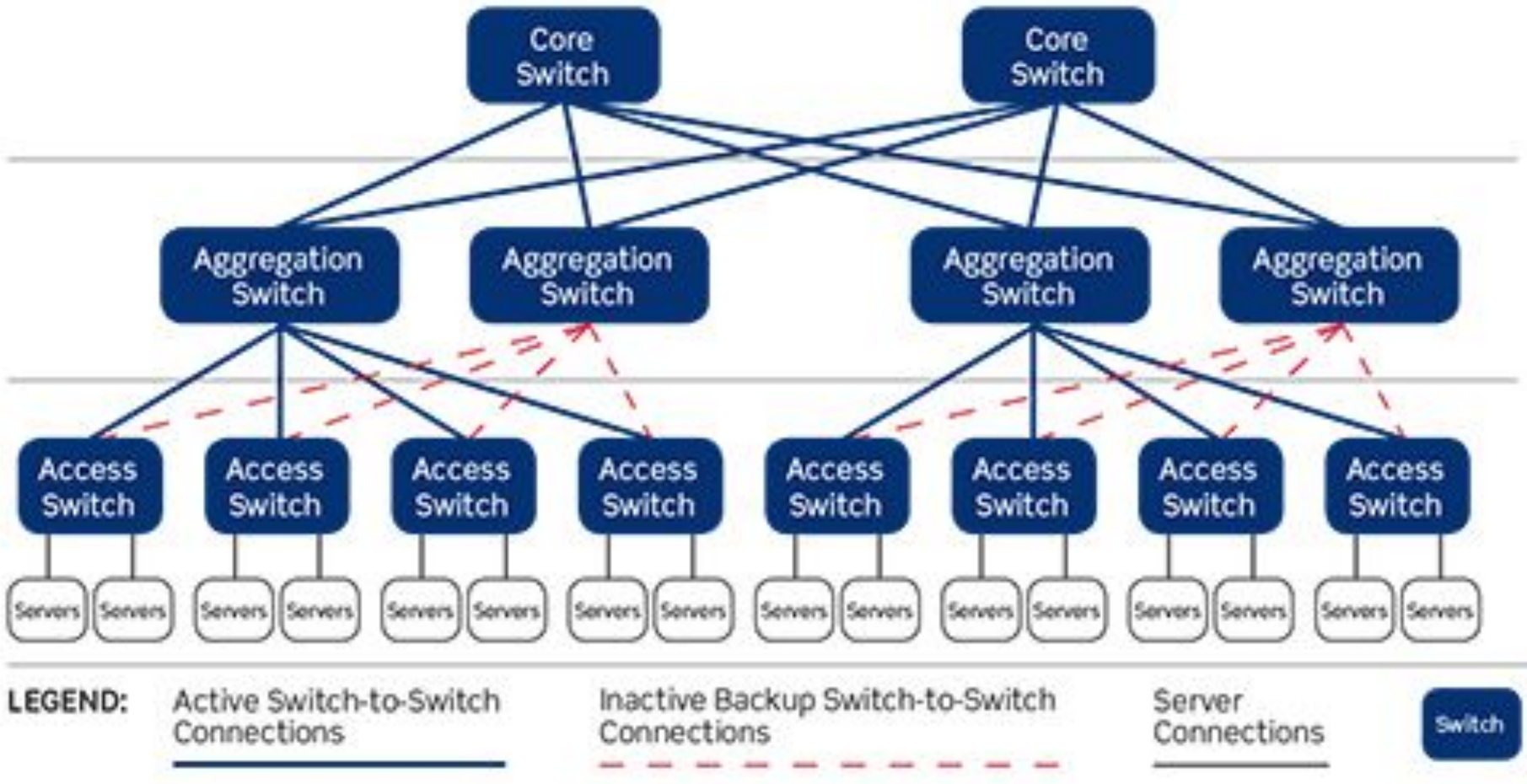
- 小型企业网络



- 带VPN的企业网


Site-to-Site IPsec VPN
Remote Access IPSec VPN
ZXR10 ZSR V2
ZXR10 6800/ZSR V2
Internet
ZXR10 ZSR V2
分公司
分公司
移动办公
总部

- 数据中心网络


Core Switch
Core Switch
Aggregation Switch
Aggregation Switch
Aggregation Switch
Aggregation Switch
Access Switch
Access Switch
Access Switch
Access Switch
Access Switch
Access Switch
Access Switch
Access Switch
Servers
LEGEND:
Active Switch-to-Switch Connections
Inactive Backup Switch-to-Switch Connections
Server Connections
Switch
Servers

Linux 协议栈

- Linux 收包流程

1 数据包到达网卡NIC (Network Interface Card)
2 NIC校验MAC(网卡非混杂模式)和帧的校验字段FCS
3 NIC通过DMA将数据包放入提前映射好的内存区域
4 NIC将数据包的引用放入接收的ring buffer队列rx中
5 NIC等待rx-usecs的超时时间或者rx队列长度达到rx-frames后触发硬件中断IRQ
6 CPU执行硬件中断和网卡的驱动程序
7 驱动程序清理硬中断并触发软中断NET_RX_SOFTIRQ
8 软中断对网卡进行轮询收包
9 数据包被放入qdisc队列
10 将数据包送入协议栈，调用ip_recv
11 调用netfilter的PREROUTING链
12 查找路由表，进行转发或者投递到local
13 对投递到local的数据包调用netfilter的LOCAL_IN链
14 调用四层协议栈，如tcp_v4_rcv
15 查找到对应的socket，运行TCP的状态机
16 将数据放入TCP的接受缓冲区中
17 通过epoll或者其他轮询方式通知应用程序
18 应用程序读取数据

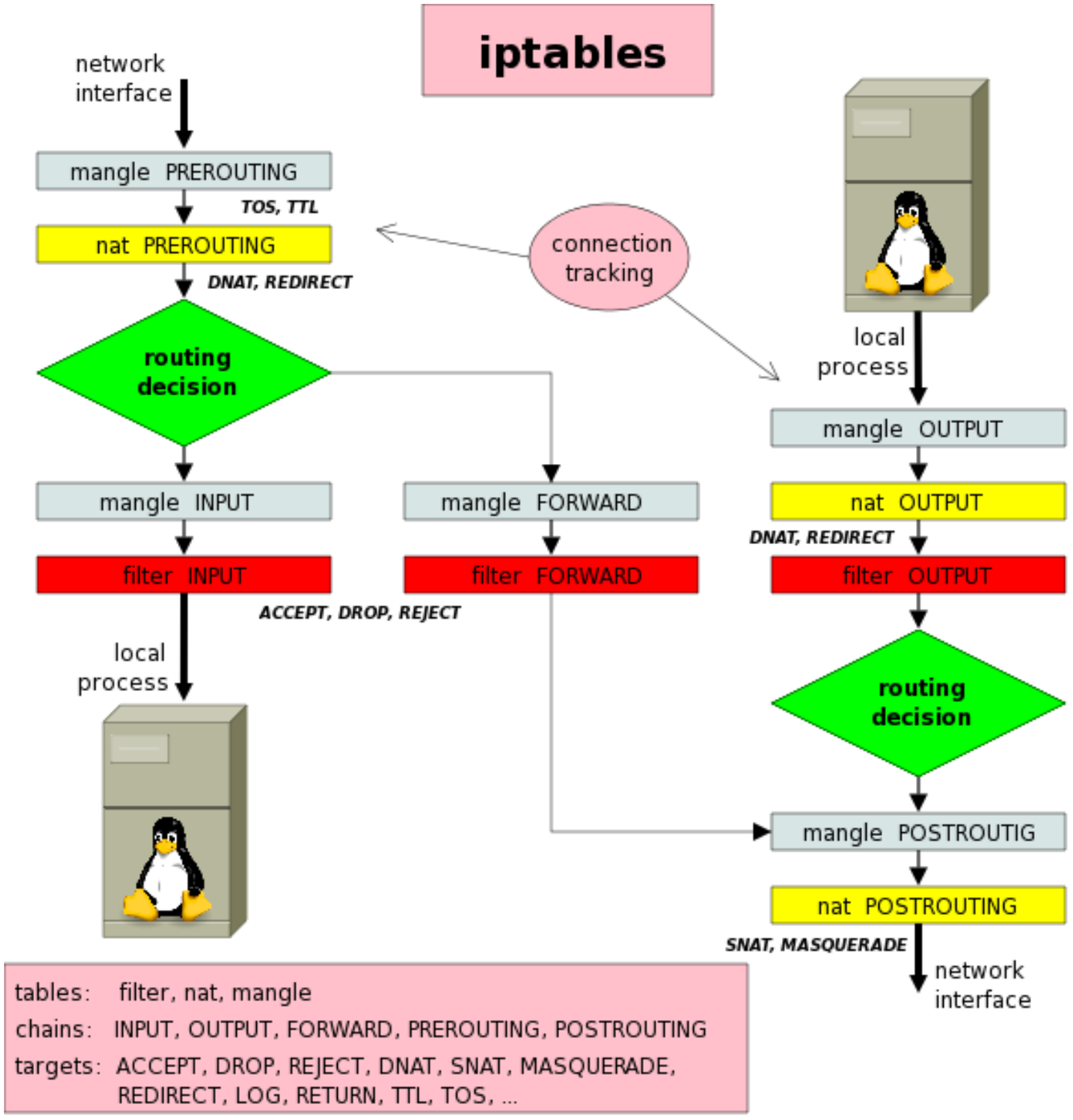
- Linux 发包流程

1 应用程序发送数据
2 TCP为发送的数据申请skb
3 构建TCP头部，如src和dst的port, checksum
4 调用第三层协议，构建IP头部，调动netfilter的LOCAL_OUT链
5 查找路由表
6 调用netfilter的POST_ROUTING链
7 对超过MTU的报文进行分片(fragment)
8 调用而成的发包函数dev_queue_xmit
9 将待发数据包放入输出的QDisc队列
10 调用网卡驱动程序，将数据包放入循环缓冲队列tx
11 驱动程序在tx-usecs的超时时间后，或者积累tx-frames个待发数据包后触发软中断
12 驱动程序启用网卡的硬件中断
13 驱动程序将数据包映射到DMA内存
14 网卡从DMA中取数据并发送
15 网卡发送完毕后触发硬件中断
16 硬中断清理中断信号后触发软中断
17 软中断释放已经发送完的数据包的内存

Linux 防火墙

- iptables　应用层规则管理工具和内核中的table模块(如filter, nat)

- netfilter　Linux包过滤框架，提供数据包过滤和处理的基础设施



- iptables 命令

格式：iptables [-t table] command [chain] [match][target]

例如：iptables -t filter -A INPUT -p tcp --sport 80 -j ACCEPT
|-table-|-cmd&chain-|————match————|—target—|

系统自带的tables包括filter, nat, mangle。每个 table包含了一些系统自带的chain或者用户自建的chain。默认使用filter，这个表中包含了INPUT, FORWARD, OUTPUT三条链

常用的target是ACCEPT和DROP，DROP和REJECT的区别就是DROP直接丢包，而REJECT会返回一个ICMP错误报文

简单命令：

1.查看设置，iptables -L -n [-t tab_name]

2.清除filter表中的规则，iptables -F

3.设置默认策略，iptables -p [INPUT | OUTPUT | FORWARD] [DROP | ACCEPT]

- iptables匹配规则

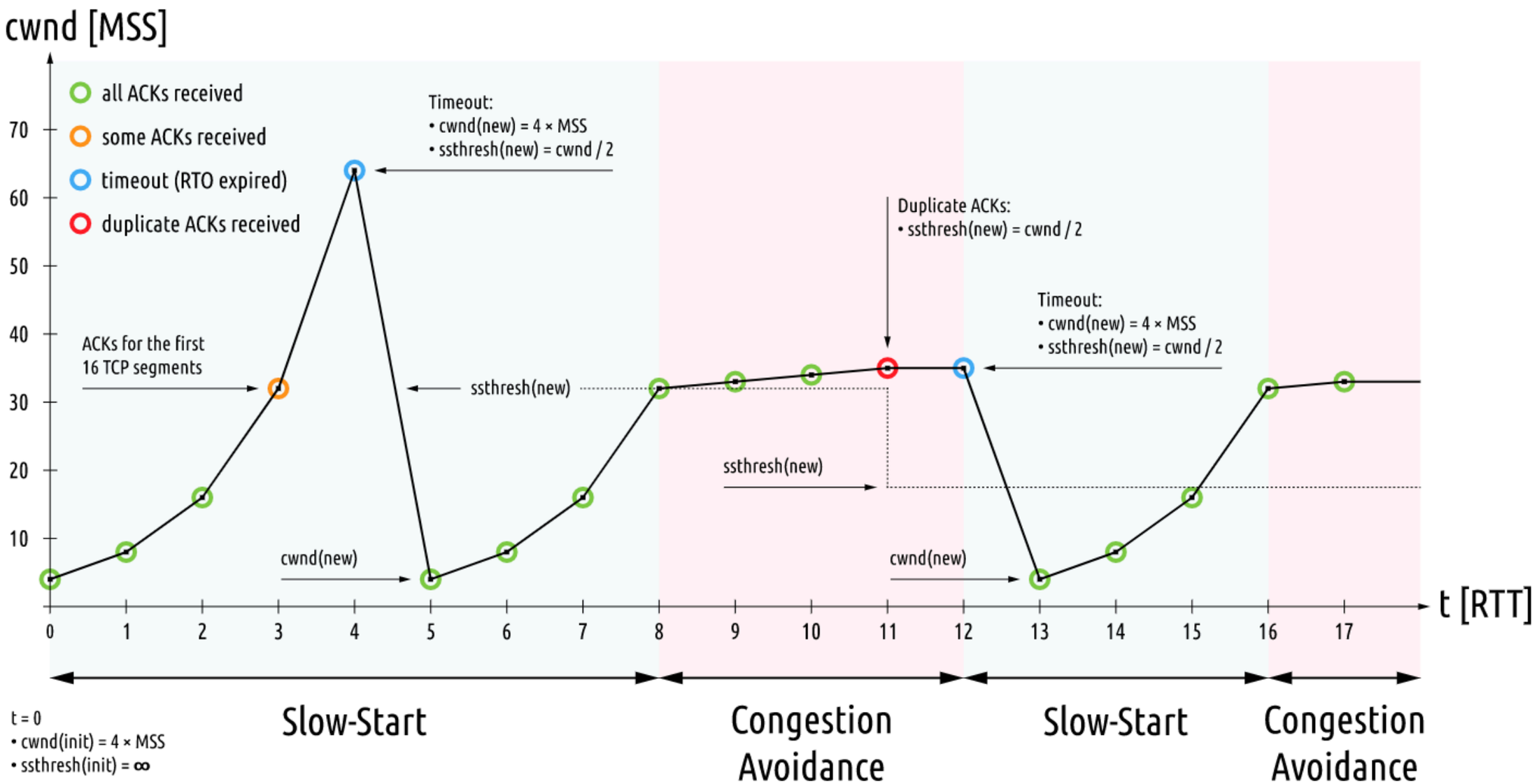
1. 匹配IP地址， source (`-s', `--source' or `--src') , destination (`-d', `--destination' or `--dst'),
   例如： iptables -A INPUT -s 10.10.10.0/24 -j DROP
2. 逻辑取反， ‘！’表示not，例如`-s！localhost'表示所有不是来自本机的数据包
3. 指定源和目的网卡接口，`-i' (or `--in-interface') ， `-o' (or `--out-interface')
4. 高级扩展匹配，使用-p或-m加载协议模块和特殊功能模块，使用模块提供的更多匹配细节，可以 使用-h或--help获取帮助，如：iptables -p tcp -h

   4.1 TCP扩展匹配
   --tcp-flags，例如：iptables -A INPUT --protocol tcp --tcp-flags ALL SYN,ACK -j DROP
   --source-port或--sport，对源端口匹配
   --destination-port或 --dport，对目的端口匹配

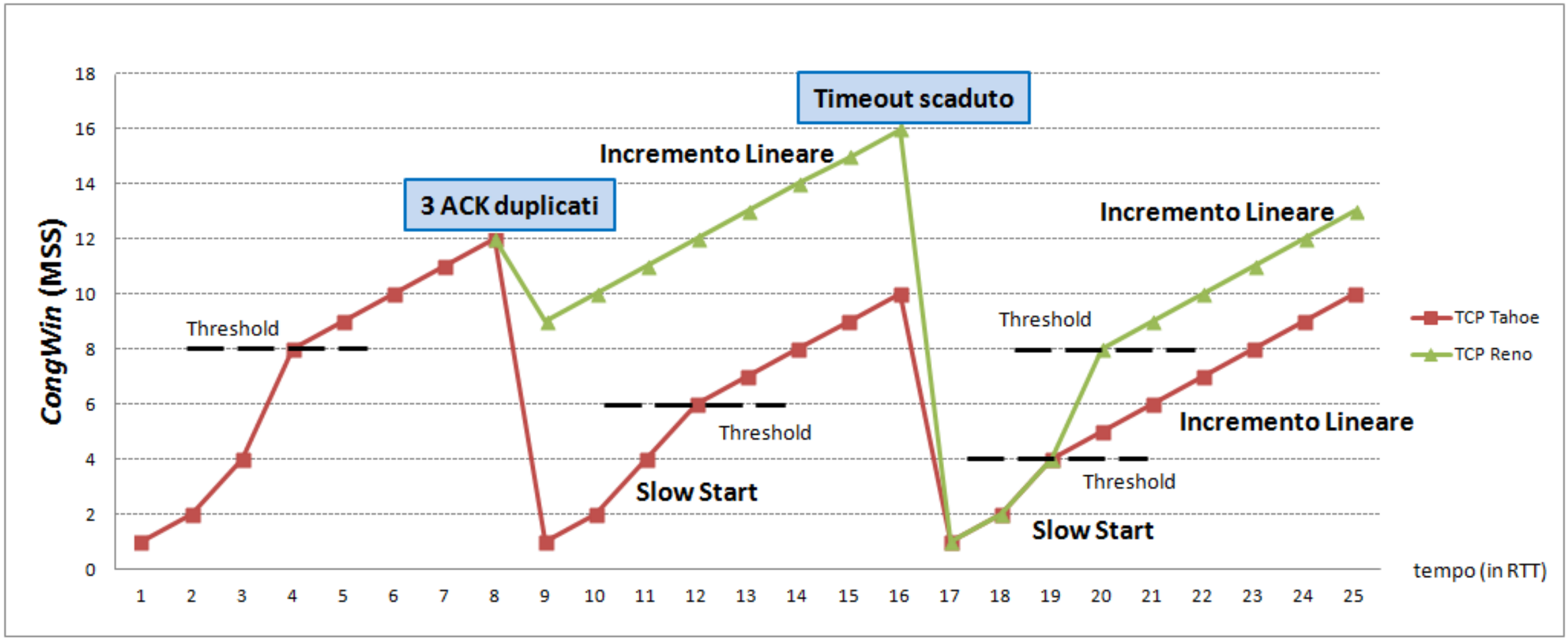
   4.2 UDP扩展匹配
   提供`--source-port', `--sport', `--destination-port' ， `--dport'，与TCP相同

前沿网络传输技术

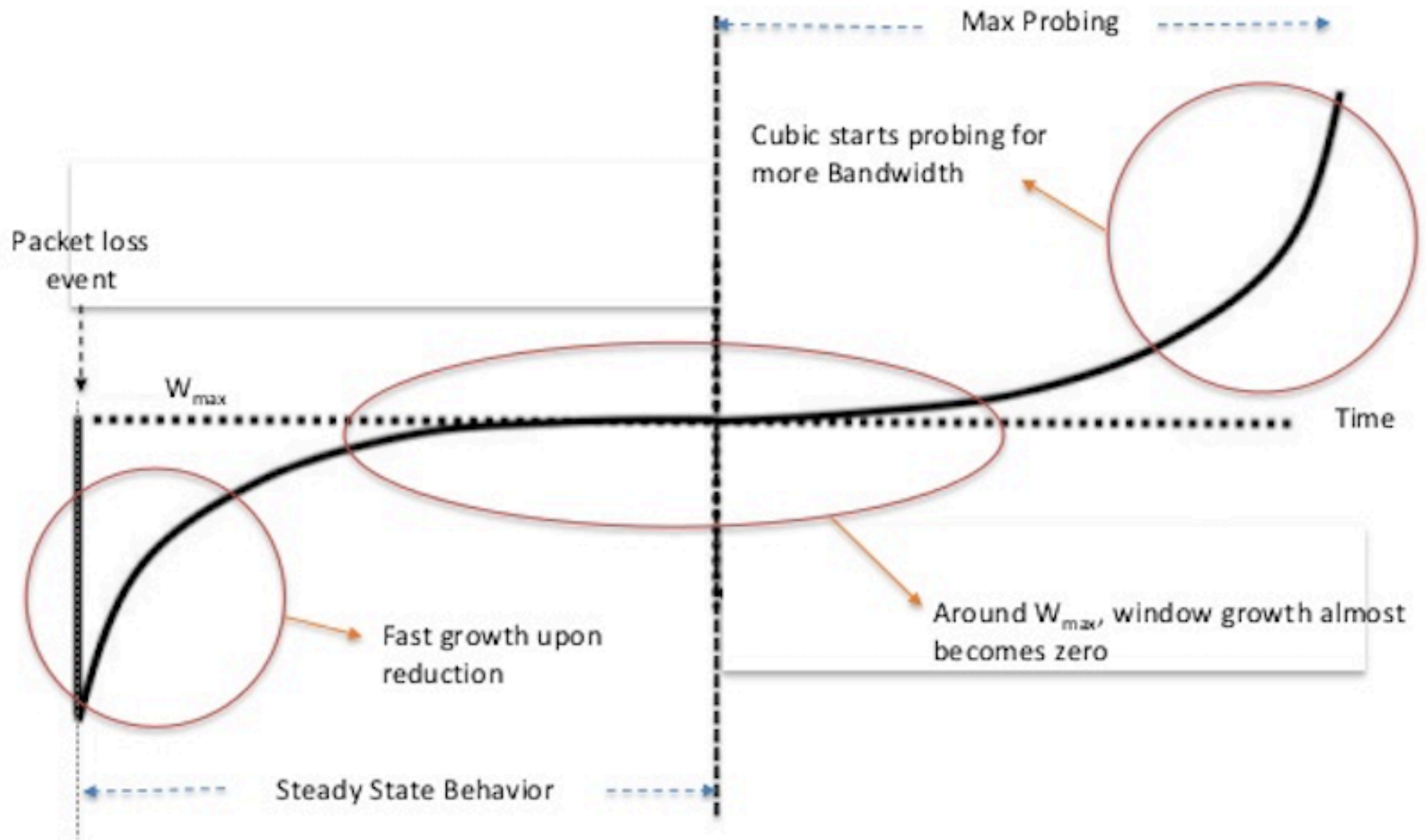
- 从TCP的演进开始



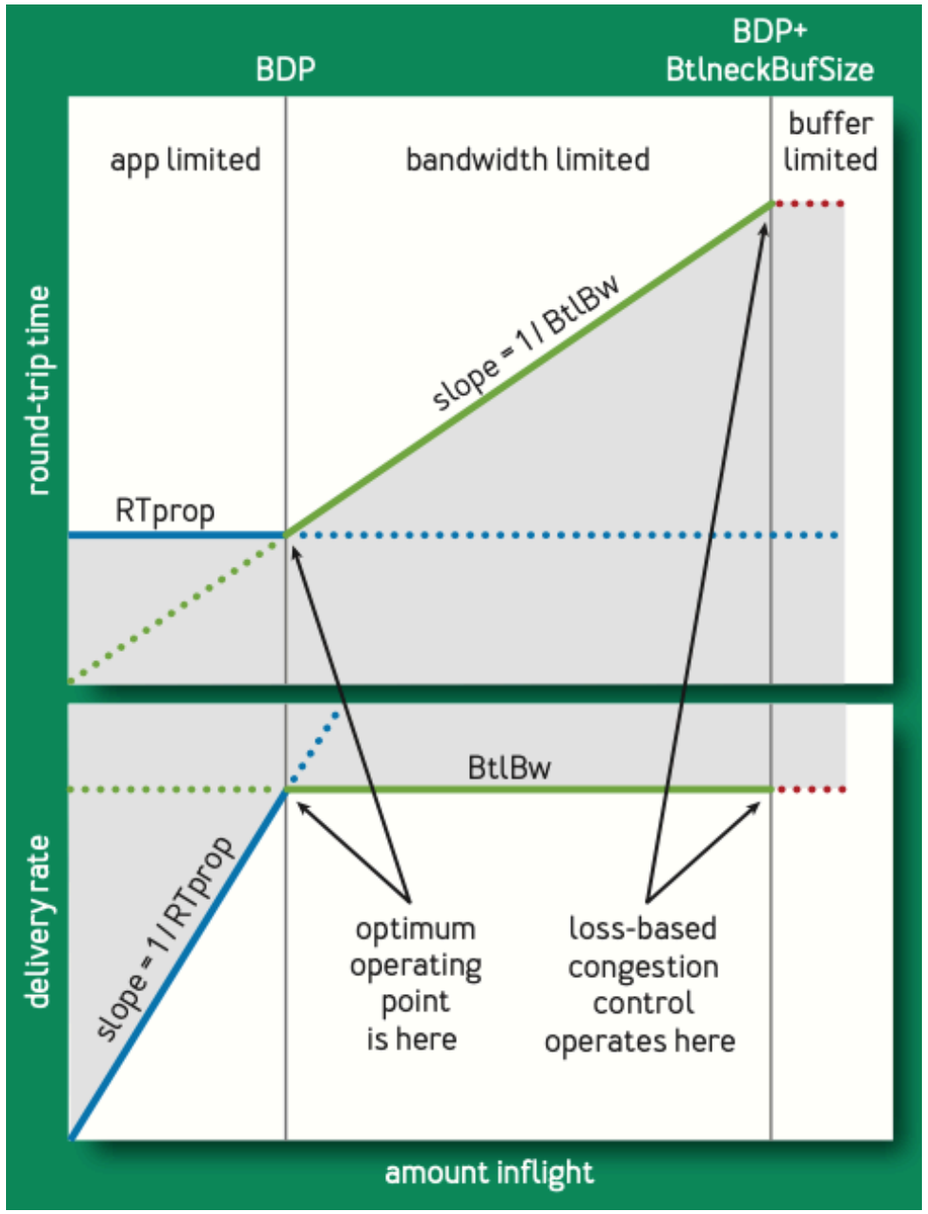
Tahoe太怂？Reno上场

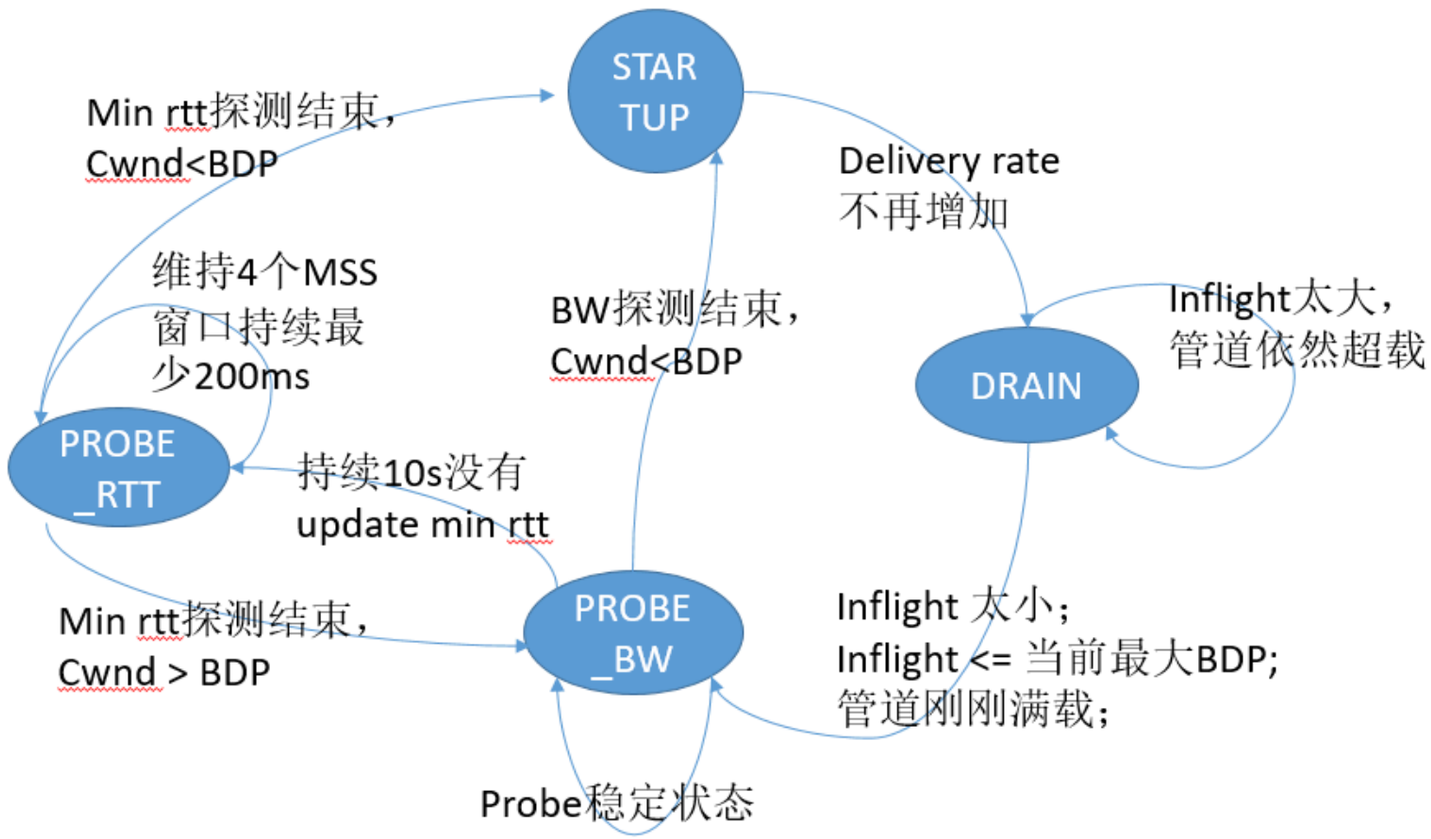


Reno还是太直男？ Cubic上场

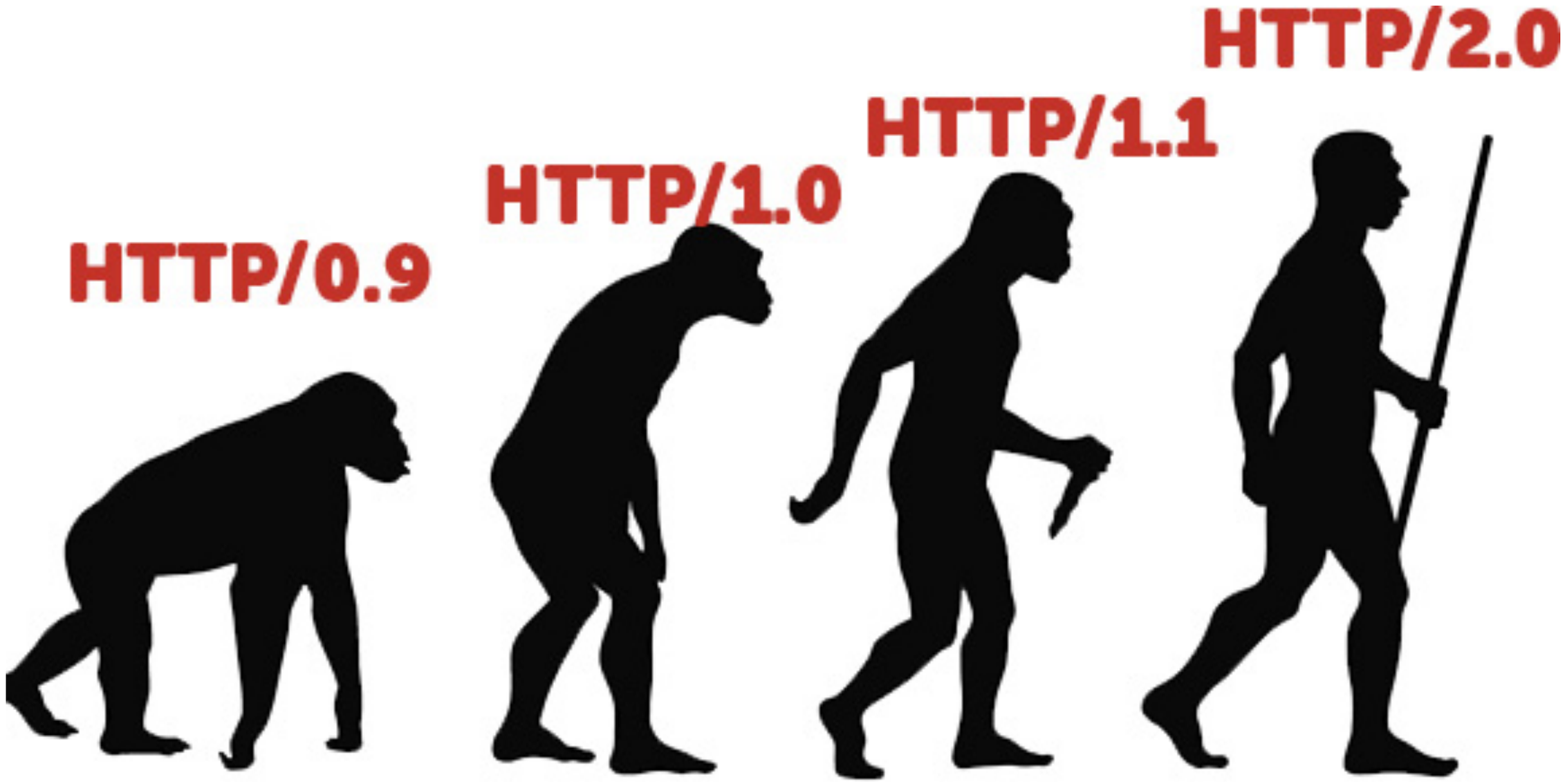


BBR：不是我针对谁，在座的各位都是…

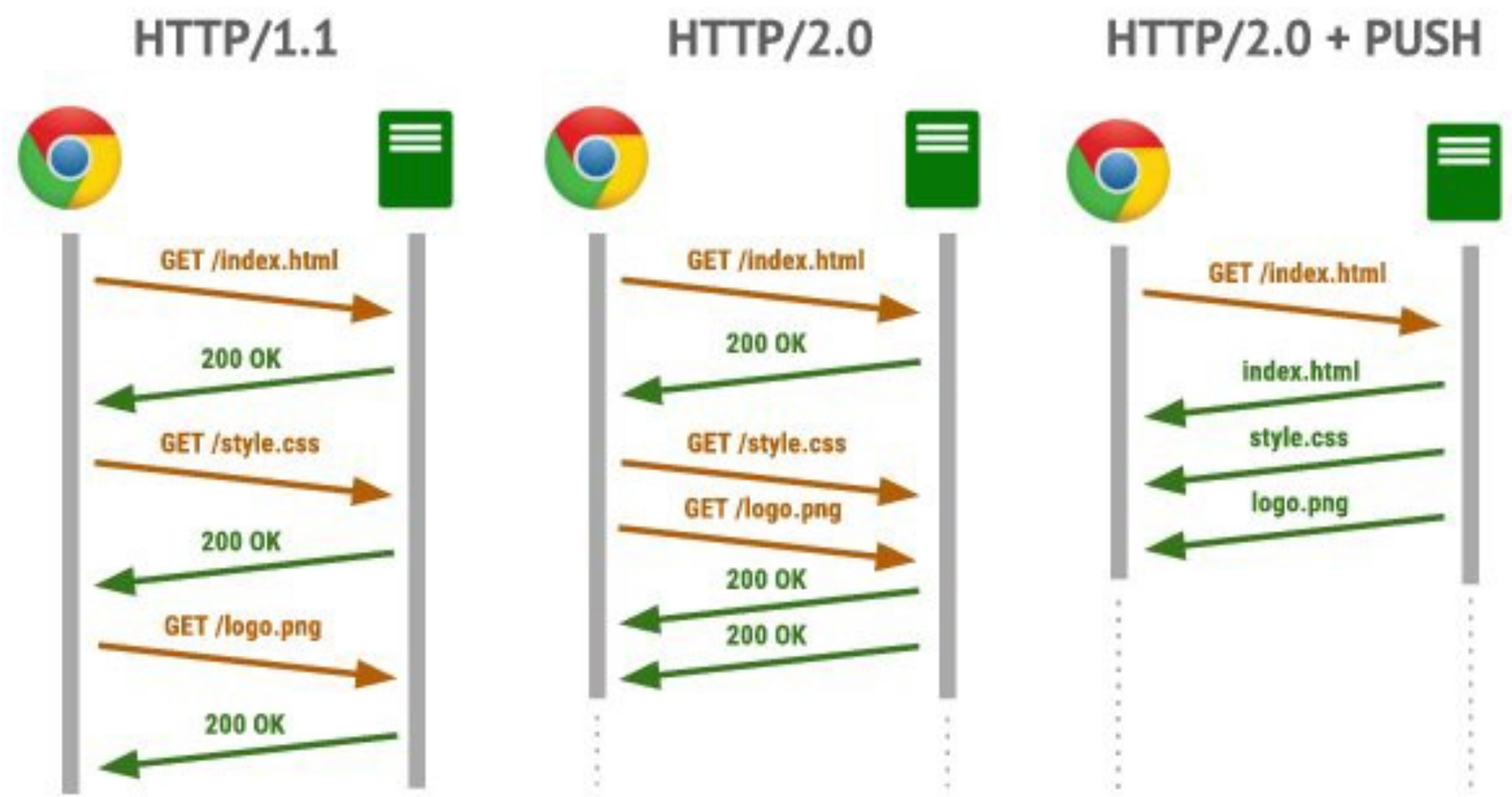




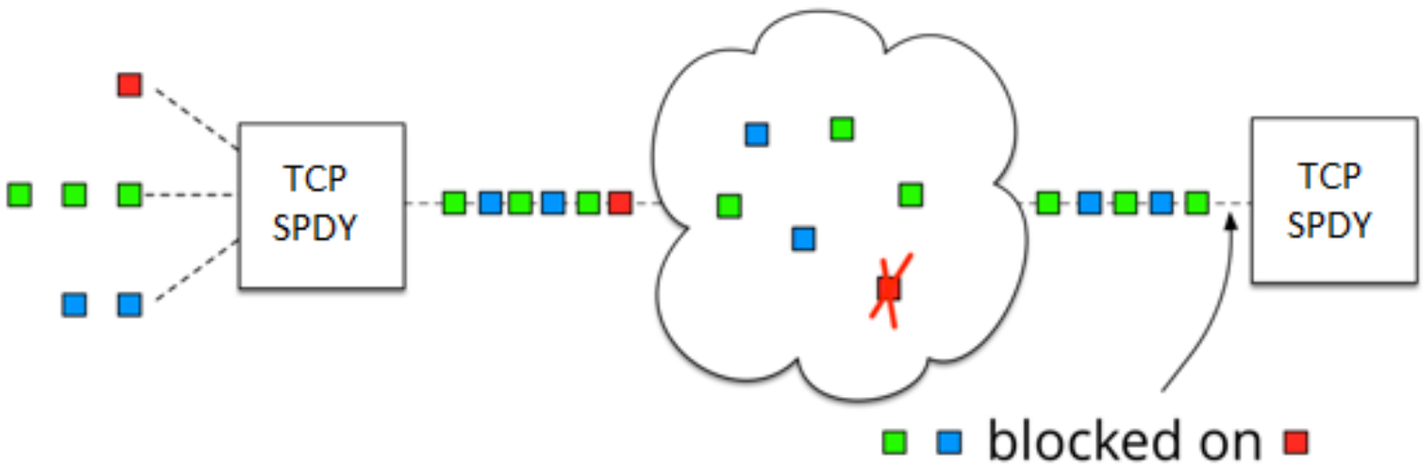
HTTP的猿族崛起

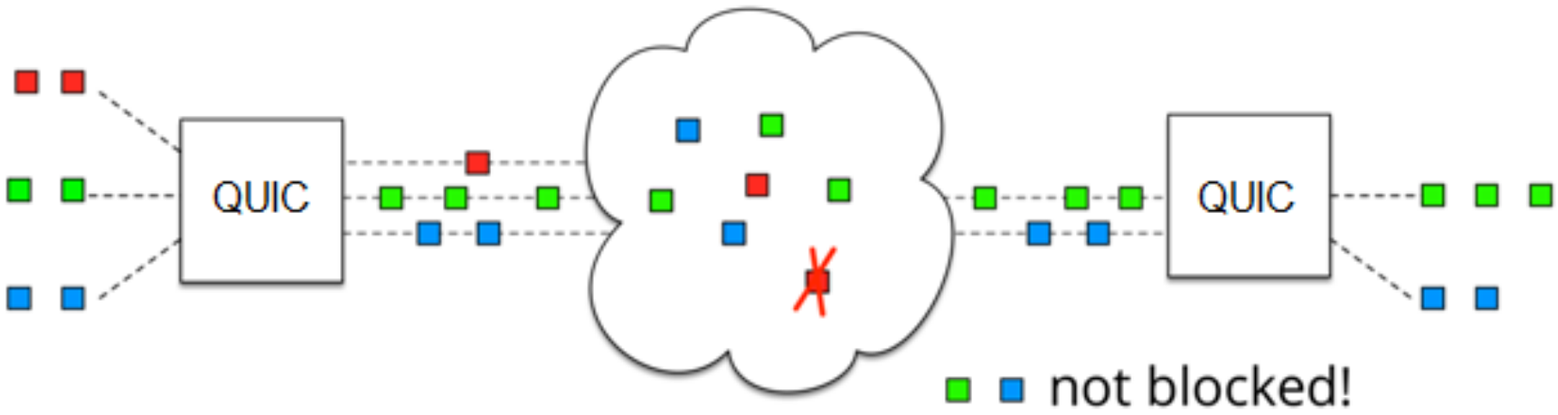


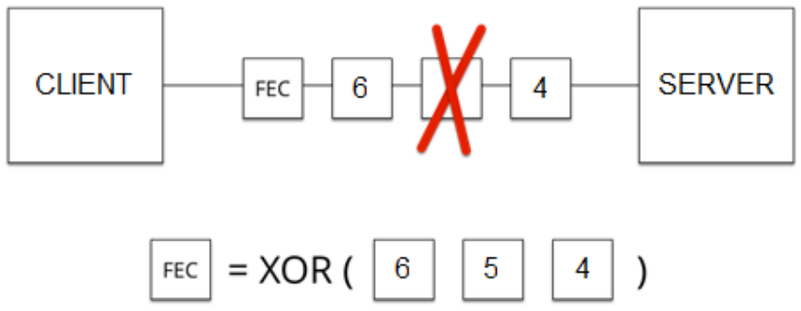
HTTP 从1.1到2.0


HTTP/1.1
GET /index.html
200 OK
GET /style.css
200 OK
GET /logo.png
200 OK
HTTP/2.0
GET /index.html
200 OK
GET /style.css
GET /logo.png
200 OK
200 OK
HTTP/2.0 + PUSH
GET /index.html
index.html
style.css
logo.png

更快的HTTP 3.0


TCP
SPDY
TCP
SPDY
blocked on


QUIC
QUIC
not blocked!


CLIENT
FEC
6
4
SERVER
FEC = XOR (6 5 4)

Thank You !

公众号：NandyTalk

E-mail: nandyliu@outlook.com