Netfilter架构及iptables

Jason.fang

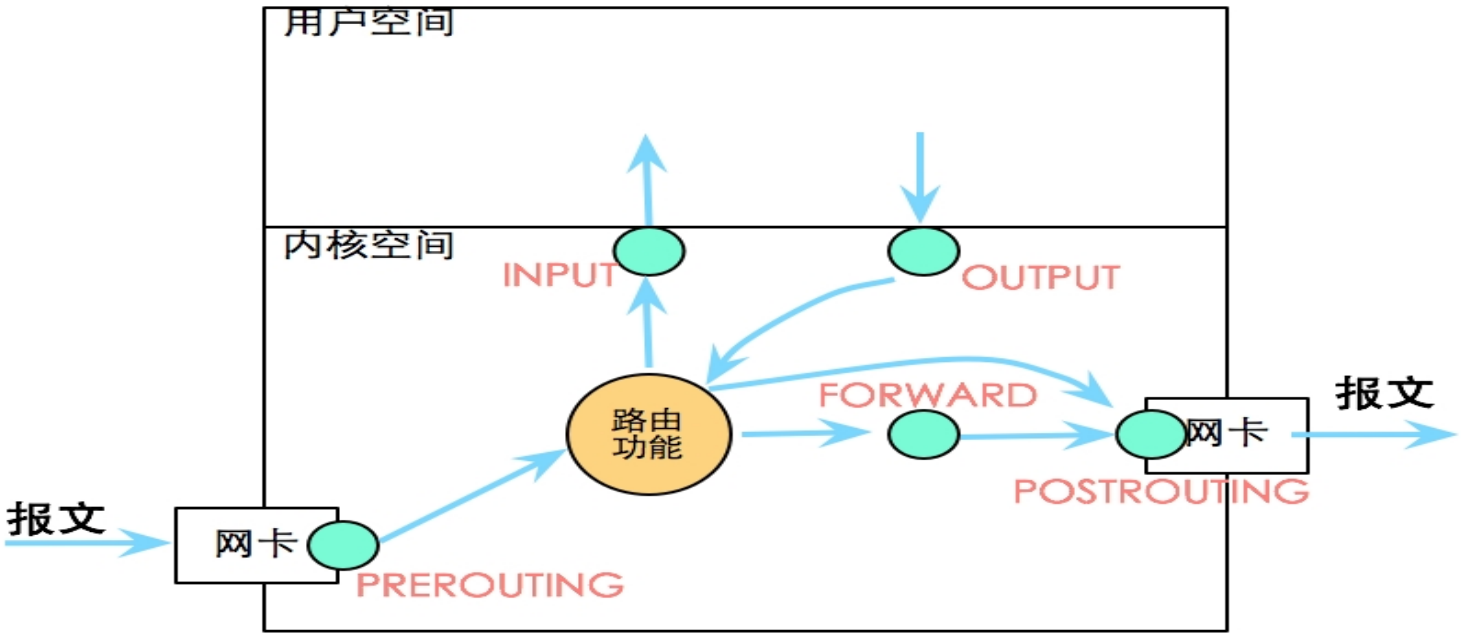
2016/12/16

Netfilter概述

- Linux内核中的netfilter架构有以太网层netfilter，主要管理工具为ebtables，有网络层的netfilter，主要管理工具为iptables

- Netfilter 是内核的一部分，由一些信息包过滤表组成，这些表包含内核用来控制信息包过滤处理的规则集。
- 通过使用用户空间，可以构建自己的定制规则，这些规则存储在内核空间的信息包过滤表中。这些规则具有 目标，它们告诉内核对来自某些源、前往某些目的地或具有某些协议类型的信息包做些什么。如果某个信息包与规则匹配，那么使用目标 ACCEPT 允许该信息包通过。还可以使用目标 DROP 或 REJECT 来阻塞并杀死信息包。对于可对信息包执行的其它操作，还有许多其它目标。

Netfilter工作原理



Netfilter的4大功能

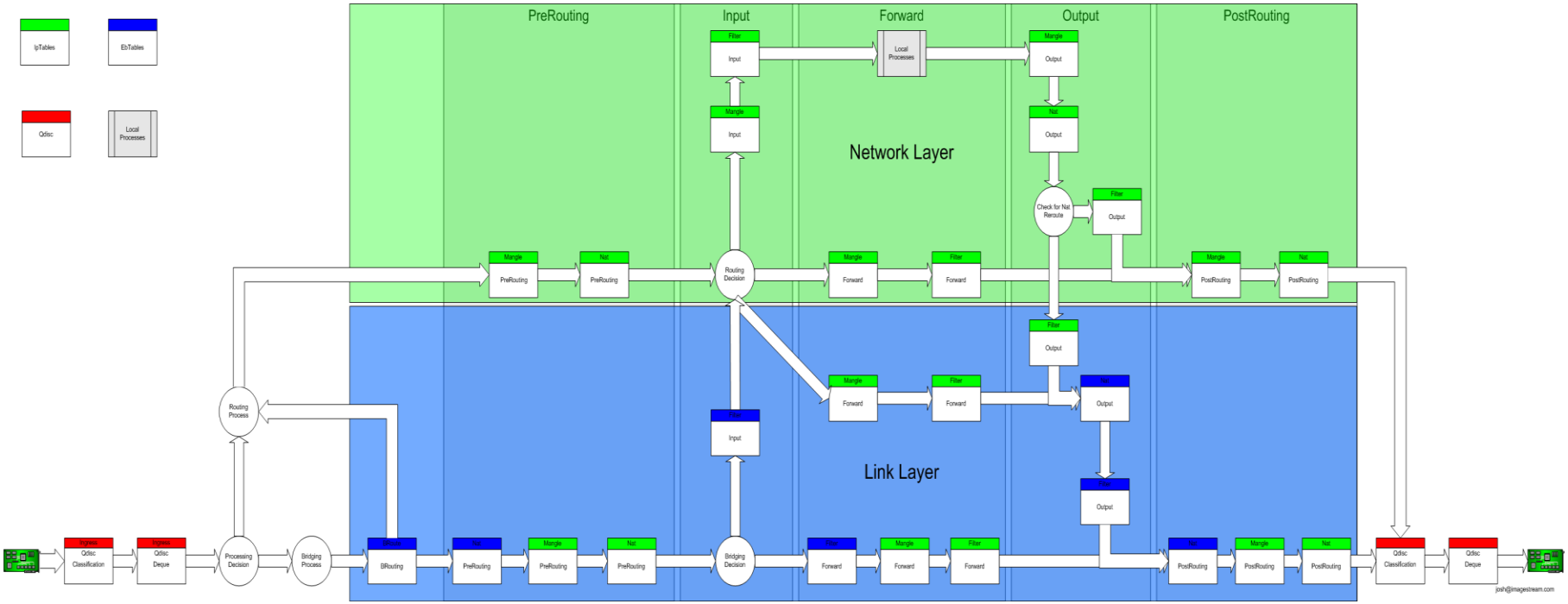
高

优先级

低

raw	动作是notrack，将跳过nat表和ip_conntrack处理
mangle	对报文进行修改，可以在5类hook_function中的任意位置
nat	分snat和dnat，作用于PREROUTING和POSTROUTING链
filter	默认机制，作用于INPUT，OUTPUT和FORWARD链

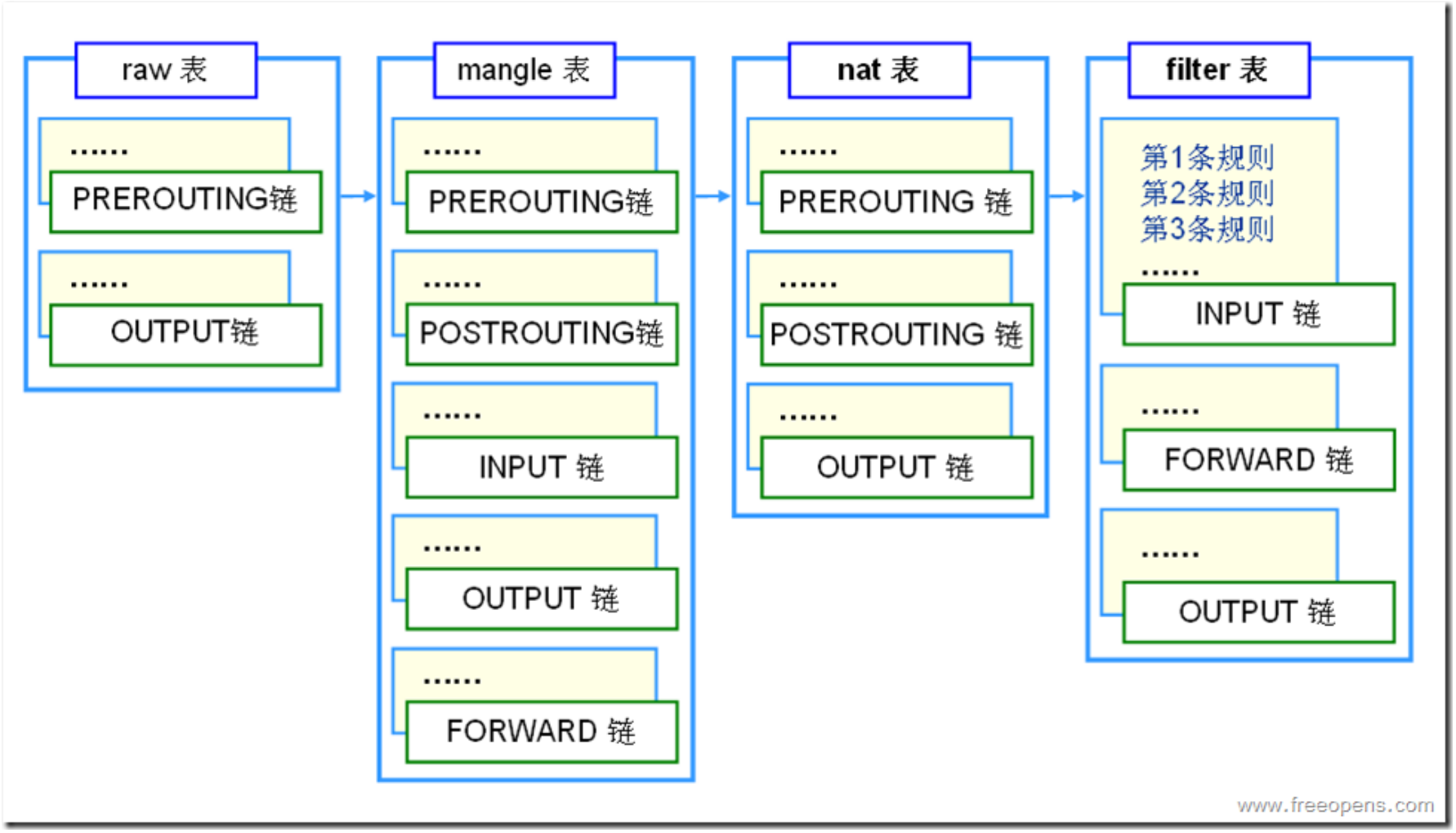
Packet flow


IpTables
EbTables
Qdisc
Local Processes
PreRouting
Input
Forward
Output
PostRouting
Network Layer
Link Layer
Routing Process
Processing Decision
Bridging Process
Routing Decision
Bridging Decision
Check for Nat Reroute
josh@imagestream.com

iptables

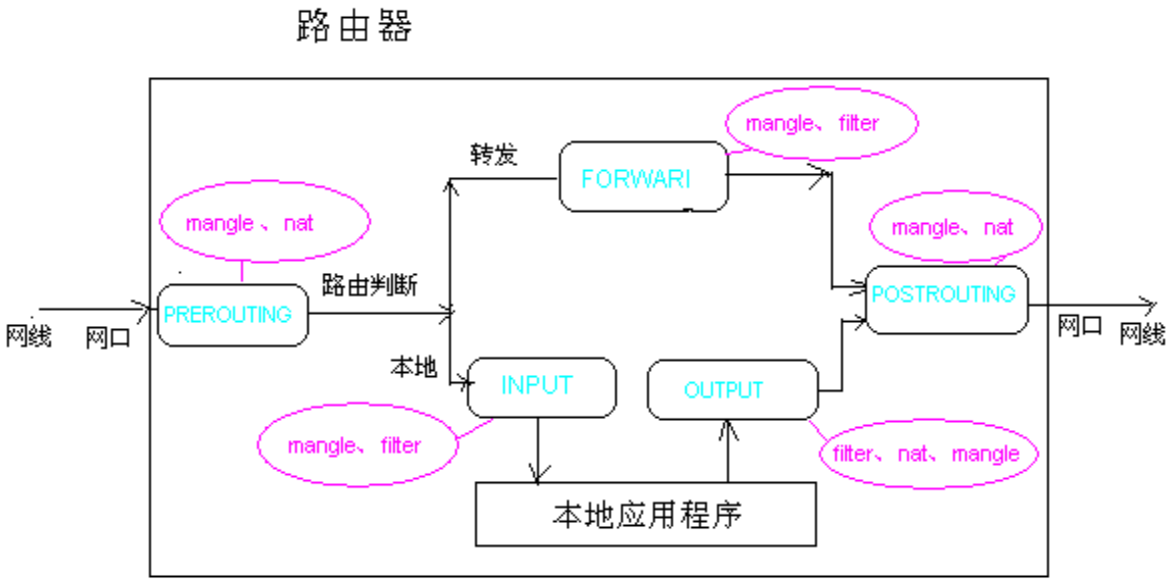
- iptables基础
  规则（rules）其实就是网络管理员预定义的条件，规则一般的定义为“如果数据包头符合这样的条件，就这样处理这个数据包”。规则存储在内核空间的信息包过滤表中，这些规则分别指定了源地址、目的地址、传输协议（如TCP、UDP、ICMP）和服务类型（如HTTP、FTP和SMTP）等。当数据包与规则匹配时，iptables就根据规则所定义的方法来处理这些数据包，如放行（accept）、拒绝（reject）和丢弃（drop）等。配置防火墙的主要工作就是添加、修改和删除这些规则。
  链（chains）是数据包传播的路径，每一条链其实就是众多规则中的一个检查清单，每一条链中可以有一条或数条规则。当一个数据包到达一个链时，iptables就会从链中第一条规则开始检查，看该数据包是否满足规则所定义的条件。如果满足，系统就会根据该条规则所定义的方法处理该数据包；否则iptables将继续检查下一条规则，如果该数据包不符合链中任一条规则，iptables就会根据该链预先定义的默认策略来处理数据包。
  表（tables）提供特定的功能，iptables内置了4个表，即filter表、nat表、mangle表和raw表，分别用于实现包过滤，网络地址转换、包重构(修改)和数据跟踪处理。

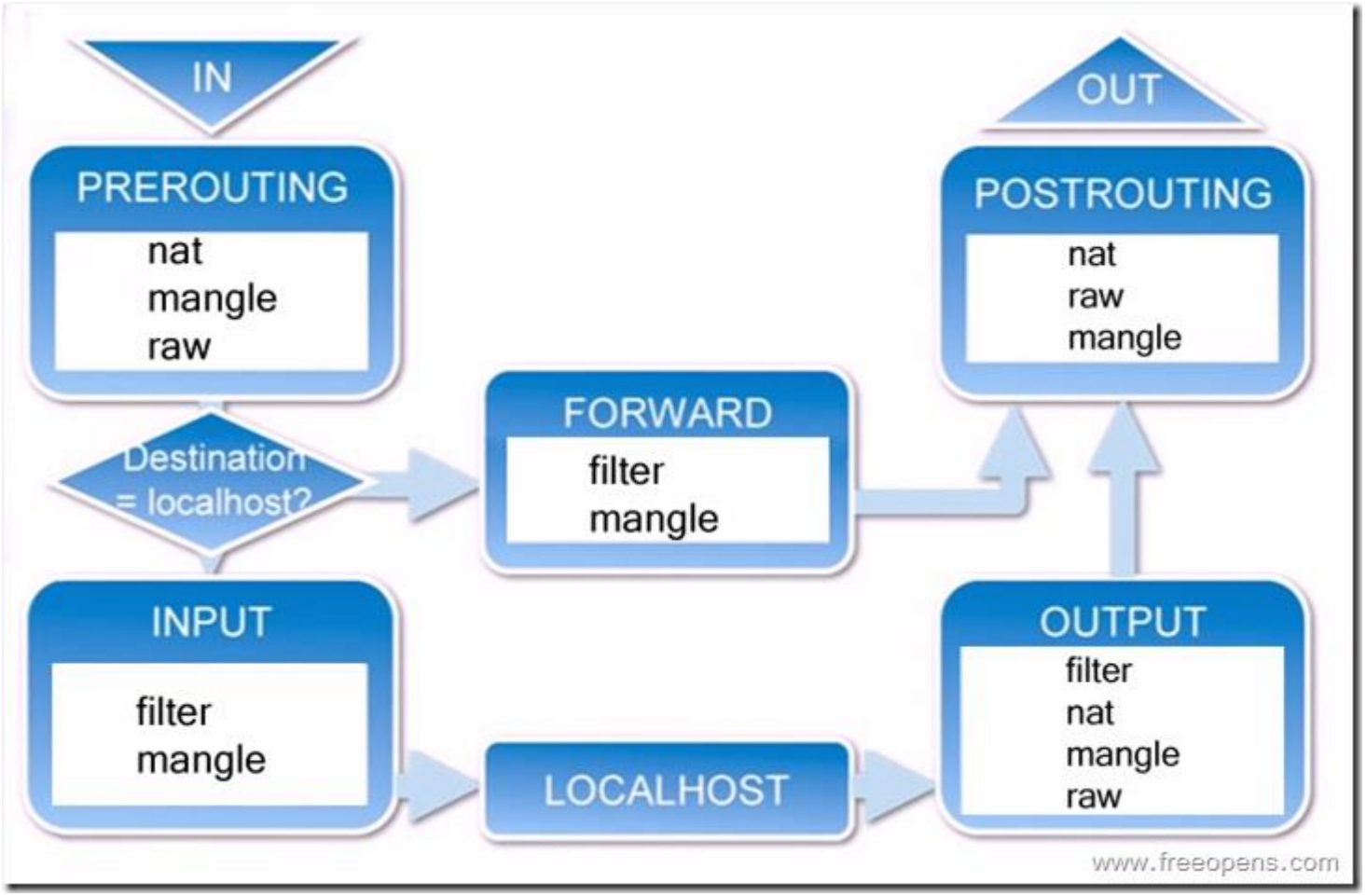
iptables表、链、规则



iptables传输数据包的过程

- iptables传输数据包的过程
  ① 当一个数据包进入网卡时，它首先进入PREROUTING链，内核根据数据包目的IP判断是否需要转送出去。
  ② 如果数据包就是进入本机的，它就会沿着图向下移动，到达INPUT链。数据包到了INPUT链后，任何进程都会收到它。本机上运行的程序可以发送数据包，这些数据包会经过OUTPUT链，然后到达POSTROUTING链输出。
  ③ 如果数据包是要转发出去的，且内核允许转发，数据包就会如图所示向右移动，经过FORWARD链，然后到达POSTROUTING链输出。（图4）




IN
PREROUTING
nat
mangle
raw
Destination
= localhost?
INPUT
filter
mangle
FORWARD
filter
mangle
LOCALHOST
OUT
POSTROUTING
nat
raw
mangle
OUTPUT
filter
nat
mangle
raw
www.freeopens.com

iptables命令格式

- iptables的命令格式较为复杂，一般的格式如下：
  iptables [-t table] 命令 [chain] [rules] [-j target]
  table——指定表明
  命令——对链的操作命令
  chain——链名
  rules——规则
  target——动作如何进行

表选项

- ．表选项
  表选项用于指定命令应用于哪个iptables内置表，iptables内置包括filter表、nat表、mangle表和raw表。

命令选项iptables命令格式

- 命令选项iptables命令格式
- 命令　　说明
  -P或–policy <链名>　定义默认策略
  -L或–list <链名>　查看iptables规则列表
  -A或—append <链名>　在规则列表的最后增加1条规则
  -I或–insert <链名>　在指定的位置插入1条规则
  -D或–delete <链名>　从规则列表中删除1条规则
  -R或–replace <链名>　替换规则列表中的某条规则
  -F或–flush <链名>　删除表中所有规则
  -Z或–zero <链名>　将表中数据包计数器和流量计数器归零

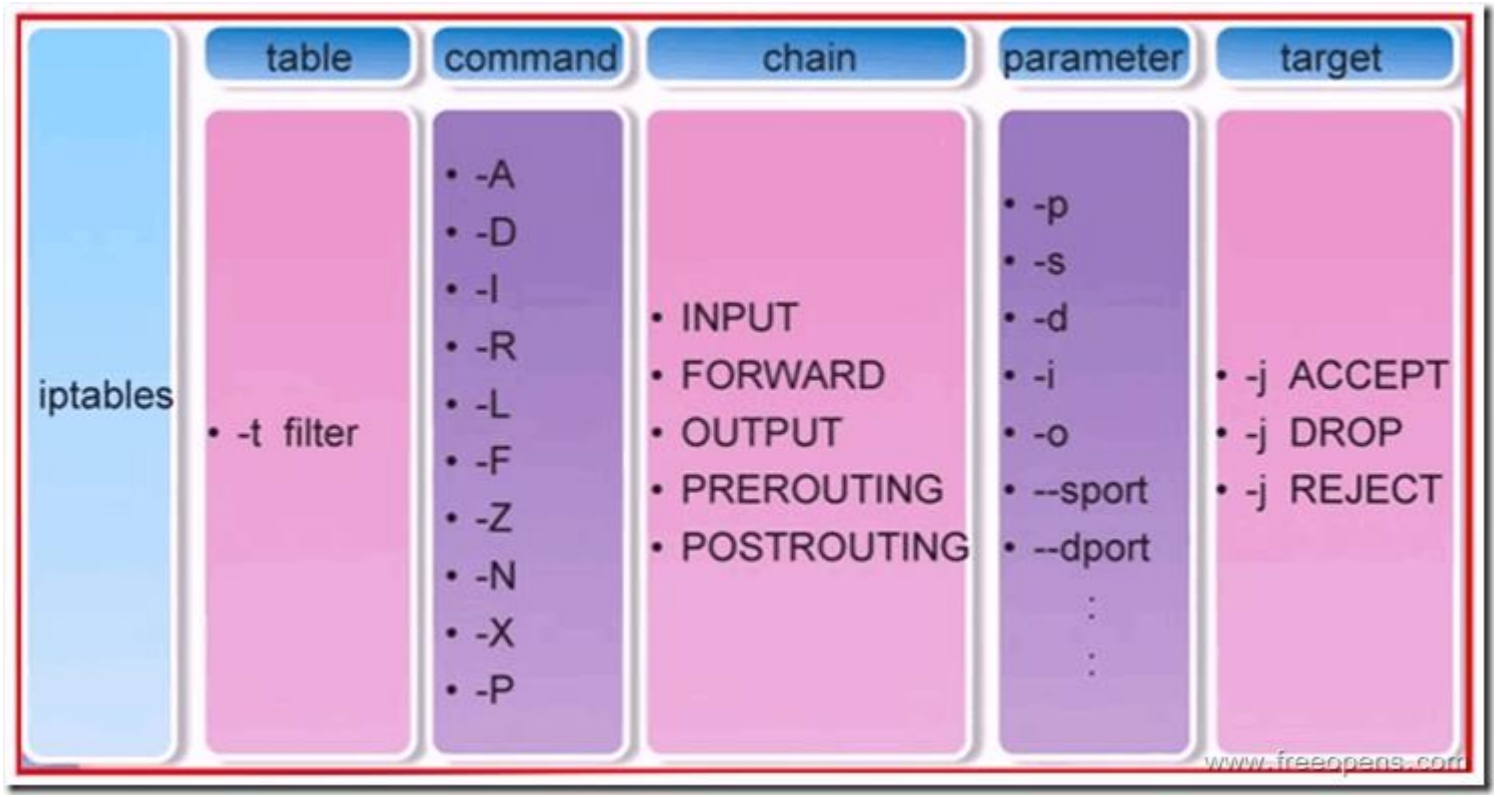
匹配选项

- 匹配　　说明
  -i或–in-interface <网络接口名>　指定数据包从哪个网络接口进入，如ppp0、eth0和eth1等
  -o或–out-interface <网络接口名>　指定数据包从哪块网络接口输出，如ppp0、eth0和eth1等
  -p或—proto协议类型 < 协议类型>　指定数据包匹配的协议，如TCP、UDP和ICMP等
  -s或–source <源地址或子网>　　指定数据包匹配的源地址
  –sport <源端口号>　　指定数据包匹配的源端口号，可以使用“起始端口号:结束端口号”的格式指定一个范围的端口
  -d或–destination <目标地址或子网>　指定数据包匹配的目标地址
  –dport目标端口号　　指定数据包匹配的目标端口号，可以使用“起始端口号:结束端口号”的格式指定一个范围的端口

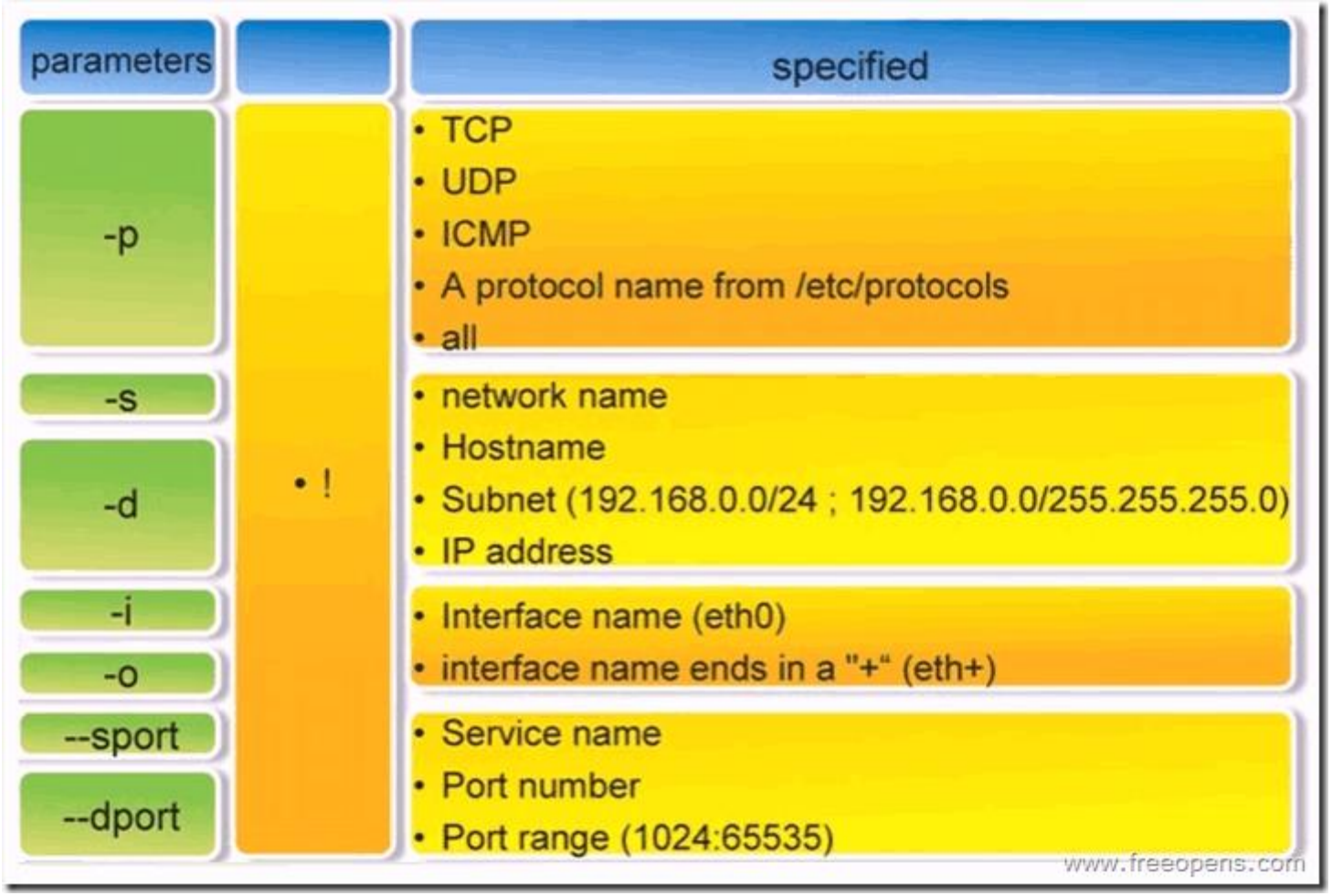
动作选项

- 动作　　说明
  ACCEPT　　接受数据包
  DROP　　丢弃数据包
  REDIRECT　与DROP基本一样，区别在于它除了阻塞包之外，还向发送者返回错误信息。
  SNAT　　源地址转换，即改变数据包的源地址
  DNAT　　目标地址转换，即改变数据包的目的地址
  MASQUERADE IP伪装，即是常说的NAT技术，MASQUERADE只能用于ADSL等拨号上网的IP伪装，也就是主机的IP是由ISP分配动态的；如果主机的IP地址是静态固定的，就要使用SNAT
  LOG　　日志功能，将符合规则的数据包的相关信息记录在日志中，以便管理员的分析和排错

iptables命令格式



iptables过滤条件



iptables的语法

- 1．定义默认策略
  当数据包不符合链中任一条规则时，iptables将根据该链预先定义的默认策略来处理数据包，默认策略的定义格式如下。
  iptables [-t表名] <-P> <链名> <动作> ?参数说明如下。
  [-t表名]：指默认策略将应用于哪个表，可以使用filter、nat和mangle，如果没有指定使用哪个表，iptables就默认使用filter表。
  <-P>：定义默认策略。
  <链名>：指默认策略将应用于哪个链，可以使用INPUT、OUTPUT、FORWARD、PREROUTING、OUTPUT和POSTROUTING。
  <动作>：处理数据包的动作，可以使用ACCEPT（接受数据包）和DROP（丢弃数据包）。

查看iptables规则

- 2．查看iptables规则
  查看iptables规则的命令格式为：
  iptables [-t表名] <-L> [链名]
  参数说明如下。
  [-t表名]：指查看哪个表的规则列表，表名用可以使用filter、nat和mangle，如果没有指定使用哪个表，iptables就默认查看filter表的规则列表。
  <-L>：查看指定表和指定链的规则列表。
  [链名]：指查看指定表中哪个链的规则列表，可以使用INPUT、OUTPUT、FORWARD、PREROUTING、OUTPUT和POSTROUTING，如果不指明哪个链，则将查看某个表中所有链的规则列表。

增加、插入、删除和替换规则

- 3．增加、插入、删除和替换规则
相关规则定义的格式为：
iptables [-t表名] <-A | I | D | R> 链名 [规则编号] [-i | o 网卡名称] [-p 协议类型] [-s 源IP地址 | 源子网] [--sport 源端口号] [-d目标IP地址 | 目标子网] [--dport目标端口号] <-j动作>
参数说明如下。
[-t表名]：定义默认策略将应用于哪个表，可以使用filter、nat和mangle，如果没有指定使用哪个表，iptables就默认使用filter表。
-A：新增加一条规则，该规则将会增加到规则列表的最后一行，该参数不能使用规则编号。
-I：插入一条规则，原本该位置上的规则将会往后顺序移动，如果没有指定规则编号，则在第一条规则前插入。
-D：从列表中删除一条规则，可以输入完整规则，或直接指定规则编号加以删除。
-R：替换某条规则，规则被替换并不会改变顺序，必须要指定替换的规则编号。
<链名>：指定查看指定表中哪个链的规则列表，可以使用INPUT、OUTPUT、FORWARD、PREROUTING、OUTPUT和POSTROUTING。
[规则编号]：规则编号用于插入、删除和替换规则时用，编号是按照规则列表的顺序排列，规则列表中第一条规则的编号为1。
[-i | o 网卡名称]：i是指定数据包从哪块网卡进入，o是指定数据包从哪块网卡输出。网卡名称可以使用ppp0、eth0和eth1等。
[-p 协议类型]：可以指定规则应用的协议，包含TCP、UDP和ICMP等。
[-s 源IP地址 | 源子网]：源主机的IP地址或子网地址。
[--sport 源端口号]：数据包的IP的源端口号。
[-d目标IP地址 | 目标子网]：目标主机的IP地址或子网地址。
[--dport目标端口号]：数据包的IP的目标端口号。
<-j动作>：处理数据包的动作，各个动作的详细说明可以参考前面的说明。

清除规则和计数器

- 清除规则和计数器
  在新建规则时，往往需要清除原有的、旧的规则，以免它们影响新设定的规则。如果规则比较多，一条条删除就会十分麻烦，这时可以使用iptables提供的清除规则参数达到快速删除所有的规则的目的。
  定义参数的格式为：
  iptables [-t表名] <-F | Z>
  参数说明如下。
  [-t表名]：指定默认策略将应用于哪个表，可以使用filter、nat和mangle，如果没有指定使用哪个表，iptables就默认使用filter表。
  -F：删除指定表中所有规则。
  -Z：将指定表中的数据包计数器和流量计数器归零。

NAT的定义

- NAT的定义
  NAT英文全称是Network Address Translation，称是网络地址转换，它是一个IETF标准，允许一个机构以一个地址出现在Internet上。NAT将每个局域网节点的地址转换成一个IP地址，反之亦然。它也可以应用到防火墙技术里，把个别IP地址隐藏起来不被外界发现，使外界无法直接访问内部网络设备，同时，它还帮助网络可以超越地址的限制，合理地安排网络中的公有Internet 地址和私有IP地址的使用。
- NAT的类型
- 静态NAT(Static NAT)
  静态NAT设置起来最为简单和最容易实现的一种，内部网络中的每个主机都被永久映射成外部网络中的某个合法的地址。
- 动态地址NAT(Pooled NAT)
  动态地址NAT是在外部网络中定义了一系列的合法地址，采用动态分配的方法映射到内部网络。
  动态地址NAT只是转换IP地址，它为每一个内部的IP地址分配一个临时的外部IP地址，主要应用于拨号，对于频繁的远程联接也可以采用动态NAT。
- 网络地址端口转换NAPT（Port－Level NAT）
  NAPT是把内部地址映射到外部网络的一个IP地址的不同端口上。
  最熟悉的一种转换方式。NAPT普遍应用于接入设备中，它可以将中小型的网络隐藏在一个合法的IP地址后面。NAPT与动态地址NAT不同，它将内部连接映射到外部网络中的一个单独的IP地址上，同时在该地址上加上一个由NAT设备选定的TCP端口号。

iptables实例

- iptables实例
- 禁止客户机访问不健康网站

  【例1】添加iptables规则禁止用户访问域名为www.sexy.com的网站。

  iptables -I FORWARD -d www.sexy.com -j DROP

  【例2】添加iptables规则禁止用户访问IP地址为20.20.20.20的网站。

  iptables -I FORWARD -d 20.20.20.20 -j DROP

禁止某些客户机上网

- 禁止某些客户机上网
  【例1】添加iptables规则禁止IP地址为192.168.1.X的客户机上网。
  iptables -I FORWARD -s 192.168.1.X -j DROP
  【例2】添加iptables规则禁止192.168.1.0子网里所有的客户机上网。
  iptables -I FORWARD -s 192.168.1.0/24 -j DROP

禁止客户机访问某些服务

- 禁止客户机访问某些服务
  【例1】禁止192.168.1.0子网里所有的客户机使用FTP协议下载。
  iptables -I FORWARD -s 192.168.1.0/24 -p tcp –dport 21 -j DROP
  【例2】禁止192.168.1.0子网里所有的客户机使用Telnet协议连接远程计算机。
  iptables -I FORWARD -s 192.168.1.0/24 -p tcp –dport 23 -j DROP

强制访问指定的站点

- 强制访问指定的站点
  【例】强制所有的客户机访问192.168.1.x这台Web服务器。
  iptables -t nat -I PREROUTING -i eth0 -p tcp –dport 80 -j DNAT –to-destination 192.168.1.x:80

禁止使用ICMP协议

- 禁止使用ICMP协议
  【例】禁止Internet上的计算机通过ICMP协议ping到NAT服务器的ppp0接口，但允许内网的客户机通过ICMP协议ping的计算机。
  iptables -I INPUT -i ppp0 -p icmp -j DROP

发布内部网络服务器

- 发布内部网络服务器
  【例1】发布内网10.0.0.3主机的Web服务，Internet用户通过访问防火墙的IP地址即可访问该主机的Web服务。
  iptables -t nat -I PREROUTING -p tcp –dport 80 -j DNAT –to-destination 10.0.0.3:80
  【例2】发布内网10.0.0.3主机的终端服务（使用的是TCP协议的3389端口），Internet用户通过访问防火墙的IP地址访问该机的终端服务。
  iptables -t nat -I PREROUTING -p tcp –dport 3389 -j DNAT –to-destination 10.0.0.3:3389

Thanks!