目 录

1 IP 性能优化

1.1 IP性能优化简介

在一些特定的网络环境里，可以通过调整 IP 的参数，以使网络性能达到最佳。IP 性能的优化配置包括：

- 配置允许接收和发送定向广播报文；
- 配置接口 MTU；
- 配置接口的 TCP 最大报文段长度；
- 配置 TCP 连接的 Path MTU 探测功能；
- 配置 TCP 的 SYN Cookie 功能；
- 配置 TCP 连接的缓冲区大小；
- 配置 TCP 定时器；
- 配置 ICMP 差错报文发送功能；
- 配置发送 ICMP 差错报文对应的令牌桶容量和令牌刷新周期；
- 配置 ICMP 报文指定源地址功能；
- 开启 IP 分片报文本地重组功能；

1.2 配置允许接口接收和转发直连网段的定向广播报文

定向广播报文是指发送给特定网络的广播报文。该报文的目的 IP 地址中网络号码字段为特定网络的网络号，主机号码字段为全 1。

接口接收和转发直连网段的定向广播报文包括以下几种情况：

- 在接收定向广播报文的情况下，如果在接口上配置了此命令，设备允许接收此接口直连网段的定向广播报文。
- 在转发定向广播报文的情况下，如果在接口上配置了此命令，设备从其他接口接收到目的地址为此接口直连网段的定向广播报文时，会从此接口转发此类报文。

黑客可以利用定向广播报文来攻击网络系统，给网络的安全带来了很大的隐患。但在某些应用环境下，设备接口需要接收或转发这类定向广播报文，例如：

在上述情况下，用户可以通过命令配置接口允许接收和转发直连网段的定向广播报文。

1.2.1 配置允许接口接收和转发直连网段的定向广播报文

表1-1 配置允许接口接收和转发直连网段的定向广播报文

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-

操作	命令	说明
配置允许接口接收和转发面向直连网段的定向广播报文	**ip forward-broadcast**	缺省情况下，设备禁止转发直连网段的定向广播报文；设备允许接收定向广播报文

1.3 配置接口MTU

当设备收到一个报文后，如果发现报文长度比转发接口的 MTU 值大，则进行下列处理：

- 如果报文不允许分片，则将报文丢弃；
- 如果报文允许分片，则将报文进行分片转发。

为了减轻转发设备在传输过程中的分片和重组数据包的压力，更高效的利用网络资源，请根据实际组网环境设置合适的接口 MTU 值，以减少分片的发生。

表1-2 配置接口 MTU

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置接口MTU	**ip mtu** *mtu-size*	缺省情况下，未配置接口MTU

1.4 配置接口的TCP最大报文段长度

TCP 最大报文段长度（Maximum Segment Size，MSS）表示 TCP 连接的对端发往本端的最大 TCP 报文段的长度，目前作为 TCP 连接建立时的一个选项来协商：当一个 TCP 连接建立时，连接的双方要将 MSS 作为 TCP 报文的一个选项通告给对端，对端会记录下这个 MSS 值，后续在发送 TCP 报文时，会限制 TCP 报文的大小不超过该 MSS 值。当对端发送的 TCP 报文的长度小于本端的 TCP 最大报文段长度时，TCP 报文不需要分段；否则，对端需要对 TCP 报文按照最大报文段长度进行分段处理后再发给本端。

用户可以通过下面的命令配置接口的 TCP 最大报文段长度，配置后该接口接收和发送的 TCP 报文的大小都不能超过该值。

该配置仅对新建的 TCP 连接生效，对于配置前已建立的 TCP 连接不生效。

该配置仅对 IP 报文生效，当接口上配置了 MPLS 功能后，不建议再配置本功能。

表1-3 配置接口的 TCP 最大报文段长度

操作	命令	说明
进入系统视图	**system-view**	-
进入接口视图	**interface** *interface-type interface-number*	-
配置接口的TCP最大报文段长度	**tcp mss** *value*	缺省情况下，未配置接口的TCP最大报文段长度

1.5 配置TCP连接的Path MTU探测功能

RFC 1191 中规定的 TCP 连接的 Path MTU 探测功能，可以探测 TCP 路径上从源端到目的端的最小 MTU，其探测机制如下：

(1) TCP 源端将发送的 TCP 数据段的外层 IP 报文设置 DF（不可分片）标记。

(2) 如果 TCP 路径上某路由器的出接口 MTU 值小于该 IP 报文长度，则会丢弃报文，并给 TCP 源端发送 ICMP 差错报文，报文中会携带该出接口 MTU 值。

(3) TCP 源端通过解析该 ICMP 差错报文，可知 TCP 路径上当前最小的单向 MTU 值。

(4) 后续 TCP 源端发送数据段的长度不超过 MSS。其中，MSS＝最小 MTU 值－IP 头部长度－TCP 头部长度。

当 MSS 已经达到系统规定的最小的 32 字节后，如果再次收到减少 MSS 的 ICMP 差错报文，系统将允许该 TCP 连接发送的报文进行分片。

产生 ICMP 差错报文的路由器可能不支持 RFC 1191，其产生的 ICMP 差错报文中的出接口 MTU 字段值为 0，对于这种报文，TCP 源端将按照 RFC 1191 中规定的 MTU 表获取比当前路径 MTU 更小的值作为计算 TCP MSS 的基础。MTU 表的内容为（单位为字节）：68、296、508、1006、1280、1492、2002、4352、8166、17914、32000、65535（由于系统规定的 TCP 最小 MSS 为 32，所以对应最小的 MTU 实际为 72 字节）。

用户通过命令行开启 TCP 连接的 Path MTU 探测功能后，新建的 TCP 连接均会携带 Path MTU 探测属性，可以通过上述探测机制确定 Path MTU，按照数据路径上的最小 MTU 组织 TCP 分段长度，最大限度利用网络资源，避免 IP 分片的发生。

Path MTU 值可以老化，这样当 Path MTU 增大时可以充分利用网络资源，尽量按照转发路径可以容忍的最大报文长度发送数据。Path MTU 的老化机制如下：

- 当 TCP 源端收到 ICMP 差错报文后，除了减小 Path MTU 值，同时会为该 Path MTU 值启动老化定时器。
- 当该定时器超时后，系统将按照 RFC 1191 规定的 MTU 表依次递增 TCP 的 MSS 值。
- 如果增加一次 MSS 之后的 2 分钟内未收到 ICMP 差错报文，则继续递增，直到 MSS 增长到对端在 TCP 三次握手阶段通告的 MSS 值。

表1-4 配置 TCP 连接的 Path MTU 探测功能

操作	命令	说明
进入系统视图	**system-view**	-
开启TCP连接的Path MTU探测功能	**tcp path-mtu-discovery** [**aging** *age-time* | **no-aging**]	缺省情况下，TCP连接的Path MTU探测功能处于关闭状态

说明

TCP 连接的 Path MTU 探测功能依赖 IP 报文的 DF 标记位设置后触发 ICMP 差错报文，因此需要 TCP 路径上的所有设备打开 ICMP 差错报文发送功能（**ip unreachables enable**），以确保 ICMP 差错报文可以发送到 TCP 源端。

1.6 配置TCP的SYN Cookie功能

一般情况下，TCP 连接的建立需要经过三次握手，即：

(1) TCP 连接请求的发起者向目标服务器发送 SYN 报文；

(2) 目标服务器收到 SYN 报文后，建立处于 SYN_RECEIVED 状态的 TCP 半连接，并向发起者回复 SYN ACK 报文，等待发起者的回应；

(3) 发起者收到 SYN ACK 报文后，回应 ACK 报文，这样 TCP 连接就建立起来了。

利用 TCP 连接的建立过程，一些恶意的攻击者可以进行 SYN Flood 攻击。攻击者向服务器发送大量请求建立 TCP 连接的 SYN 报文，而不回应服务器的 SYN ACK 报文，导致服务器上建立了大量的 TCP 半连接。从而，达到耗费服务器资源，使服务器无法处理正常业务的目的。

SYN Cookie 功能用来防止 SYN Flood 攻击。在服务器上配置此功能后，当服务器收到 TCP 连接请求时，不建立 TCP 半连接，而直接向发起者回复 SYN ACK 报文。服务器接收到发起者回应的 ACK 报文后，建立连接，并进入 ESTABLISHED 状态。通过这种方式，可以避免在服务器上建立大量的 TCP 半连接，防止服务器受到 SYN Flood 攻击。

表1-5 配置 TCP 的 SYN Cookie 功能

操作	命令	说明
进入系统视图	**system-view**	-
使能SYN Cookie功能	**tcp syn-cookie enable**	缺省情况下，SYN Cookie功能处于关闭状态

1.7 配置TCP连接的缓冲区大小

表1-6 配置 TCP 连接的缓冲区大小

操作	命令	说明
进入系统视图	**system-view**	-
配置TCP连接的接收和发送缓冲区的大小	**tcp window** *window-size*	缺省情况下，TCP连接的接收和发送缓冲区大小为63KB

1.8 配置TCP定时器

可以配置的 TCP 定时器包括：

- synwait 定时器：当发送 SYN 报文时，TCP 启动 synwait 定时器和重传 SYN 报文定时器，当 synwait 定时器超时且 SYN 报文重传未达到最大次数时，如果设备未收到回应报文，则 TCP 连接建立不成功；当 synwait 定时器未超时但是 SYN 报文重传达到最大次数时，如果设备未收到回应报文，则 TCP 连接建立不成功。
- finwait 定时器：当 TCP 的连接状态为 FIN_WAIT_2 时，启动 finwait 定时器，如果在定时器超时前未收到报文，则 TCP 连接终止；如果收到 FIN 报文，则 TCP 连接状态变为 TIME_WAIT 状态；如果收到非 FIN 报文，则从收到的最后一个非 FIN 报文开始重新计时，在超时后中止连接。

表1-7 配置 TCP 定时器

操作	命令	说明
进入系统视图	**system-view**	-
配置TCP的synwait定时器超时时间	**tcp timer syn-timeout** *time-value*	缺省情况下，synwait定时器超时时间为75秒
配置TCP的finwait定时器超时时间	**tcp timer fin-timeout** *time-value*	缺省情况下，finwait定时器超时时间为675秒

1.9 配置ICMP差错报文发送功能

发送差错报文是 ICMP（Internet Control Message Protocol，互联网控制消息协议）的主要功能之一。ICMP 报文通常被网络层或传输层协议用来在异常情况发生时通知相应设备，从而便于进行控制管理。

重定向报文、超时报文、目的不可达报文是 ICMP 差错报文中的三种。下面分别介绍这三种差错报文发送的条件及作用。

(1) ICMP 重定向报文发送功能

主机启动时，它的路由表中可能只有一条到缺省网关的缺省路由。当满足一定的条件时，缺省网关会向源主机发送 ICMP 重定向报文，通知主机重新选择正确的下一跳进行后续报文的发送。

满足下列条件时，设备会发送 ICMP 重定向报文：

- 接收和转发数据报文的接口是同一接口；
- 被选择的路由本身没有被 ICMP 重定向报文创建或修改过；
- 被选择的路由不是到默认目的地（0.0.0.0）的路由；
- 数据报文中没有源路由选项。

ICMP 重定向报文发送功能可以简化主机的管理，使具有很少选路信息的主机逐渐建立较完善的路由表，从而找到最佳路由。

(2) ICMP 超时报文发送功能

ICMP 超时报文发送功能是在设备收到 IP 数据报文后，如果发生超时差错，则将报文丢弃并给源端发送 ICMP 超时差错报文。

设备在满足下列条件时会发送 ICMP 超时报文：

- 设备收到 IP 数据报文后，如果报文的目的地不是本地且报文的 TTL 字段是 1，则发送“TTL 超时”ICMP 差错报文；
- 设备收到目的地址为本地的 IP 数据报文的第一个分片后，启动定时器，如果所有分片报文到达之前定时器超时，则会发送“重组超时”ICMP 差错报文。

(3) ICMP 目的不可达报文发送功能

ICMP 目的不可达报文发送功能是在设备收到 IP 数据报文后，如果发生目的不可达的差错，则将报文丢弃并给源端发送 ICMP 目的不可达差错报文。

设备在满足下列条件时会发送目的不可达报文：

- 设备在转发报文时，如果在路由表中未找到对应的转发路由，且路由表中没有缺省路由，则给源端发送“网络不可达”ICMP 差错报文；
- 设备收到目的地址为本地的数据报文时，如果设备不支持数据报文采用的传输层协议，则给源端发送“协议不可达”ICMP 差错报文；
- 设备收到目的地址为本地、传输层协议为 UDP 的数据报文时，如果报文的端口号与正在使用的进程不匹配，则给源端发送“端口不可达”ICMP 差错报文；
- 源端如果采用“严格的源路由选择”发送报文，当中间设备发现源路由所指定的下一个设备不在其直接连接的网络上，则给源端发送“源站路由失败”的 ICMP 差错报文；
- 设备在转发报文时，如果转发接口的 MTU 小于报文的长度，但报文被设置了不可分片，则给源端发送“需要进行分片但设置了不分片比特”ICMP 差错报文。

ICMP 差错报文的发送虽然方便了网络的控制管理，但是也存在缺限：发送大量的 ICMP 报文，增大网络流量;如果有用户发送 ICMP 差错报文进行恶意攻击,会导致设备性能下降或影响正常工作。

为了避免上述现象发生，可以关闭设备的 ICMP 差错报文发送功能，从而减少网络流量、防止遭到恶意攻击。

表1-8 配置 ICMP 差错报文发送功能

操作	命令	说明
进入系统视图	**system-view**	-
开启ICMP重定向报文发送功能	**ip redirects enable**	缺省情况下，ICMP重定向报文发送功能处于关闭状态
开启ICMP超时报文发送功能	**ip ttl-expires enable**	缺省情况下，ICMP超时报文发送功能处于关闭状态
开启ICMP目的不可达报文发送功能	**ip unreachables enable**	缺省情况下，ICMP目的不可达报文发送功能处于关闭状态

说明

- 关闭 ICMP 超时报文发送功能后，设备不会再发送“TTL 超时”ICMP 差错报文，但“重组超时”ICMP 差错报文仍会正常发送。
- 设备开启 DHCP 服务后，在未发送 ICMP 回显请求（ECHO-REQUEST）报文情况下，收到非法 ICMP 回显应答（ECHO-REPLY）报文，此时设备不会回应“协议不可达”ICMP 差错报文。关于 DHCP 的详细介绍，请参见“三层技术-IP 业务配置指导”中的“DHCP”。

1.10 配置发送ICMP差错报文对应的令牌桶容量和令牌刷新周期

如果网络中短时间内发送的 ICMP 差错报文过多，将可能导致网络拥塞。为了避免这种情况，用户可以控制设备在指定时间内发送 ICMP 差错报文的最大数目，目前采用令牌桶算法来实现。

用户可以设置令牌桶的容量，即令牌桶中可以同时容纳的令牌数；同时可以设置令牌桶的刷新周期，即每隔多长时间发放一个令牌到令牌桶中，直到令牌桶中的令牌数达到配置的容量。一个令牌表示允许发送一个 ICMP 差错报文，每当发送一个 ICMP 差错报文，则令牌桶中减少一个令牌。如果连续发送的 ICMP 差错报文超过了令牌桶的容量，则后续的 ICMP 差错报文将不能被发送出去，直到按照所设置的刷新频率将新的令牌放入令牌桶中。

表1-9 配置发送 ICMP 差错报文对应的令牌桶容量和令牌刷新周期

操作	命令	说明
进入系统视图	**system-view**	-
配置发送ICMP差错报文对应的令牌桶容量和令牌刷新周期	**ip icmp error-interval** *interval* [*bucketsize*]	缺省情况下，令牌桶容量为10，令牌刷新周期为100毫秒 刷新周期为0时，表示不限制ICMP差错报文的发送

1.11 配置ICMP报文指定源地址功能

在网络中 IP 地址配置较多的情况下，收到 ICMP 报文时，用户很难根据报文的源 IP 地址判断报文来自哪台设备。为了简化这一判断过程，可以配置 ICMP 报文指定源地址功能。用户配置特定地址（如环回口地址）为 ICMP 报文的源地址，可以简化判断。

设备发送 ICMP 差错报文（TTL 超时、端口不可达和参数错误等）和 ping echo request 报文时，都可以通过上述命令指定报文的源地址。

表1-10 配置 ICMP 报文指定源地址功能

操作	命令	说明
进入系统视图	**system-view**	-
开启设备的ICMP报文指定源地址功能	**ip icmp source** [**vpn-instance** *vpn-instance-name*] *ip-address*	缺省情况下，ICMP报文指定源地址功能处于关闭状态

说明

用户发送 ping echo request 报文时，如果 **ping** 命令中已经指定源地址，则使用该源地址，否则使用 **ip icmp source** 配置的源地址。

1.12 开启IP分片报文本地重组功能

多台设备组成的 IRF 环境下，当某成员设备收到目的为本 IRF 设备的 IP 分片报文时，需要把分片报文送到主设备进行重组，这样会导致报文重组性能较低的问题。当开启 IP 分片报文本地重组功能后，分片报文会在该成员设备上直接进行报文重组，这样就能提高分片报文的重组性能。开启 IP 分片报文本地重组功能后，如果分片报文是从设备上不同的成员设备进入的，会导致 IP 分片报文本地无法重组成功。

表1-11 开启 IP 分片报文本地重组功能

操作	命令	说明
进入系统视图	**system-view**	-
开启IP分片报文本地重组功能	**ip reassemble local enable**	缺省情况下，IP分片报文本地重组功能处于关闭状态

1.13 IP性能优化显示和维护

在完成上述配置后，在任意视图下执行 **display** 命令可以显示配置 IP 性能优化功能后的运行情况，通过查看显示信息验证配置的效果。

在用户视图下执行 **reset** 命令清除 IP、TCP 和 UDP 的流量统计信息。

表1-12 IP 性能优化显示和维护

操作	命令
显示RawIP连接摘要信息	**display rawip** [**slot** *slot-number*]
显示RawIP连接详细信息	**display rawip verbose** [**slot** *slot-number* [**pcb** *pcb-index*]]
显示TCP连接摘要信息	**display tcp** [**slot** *slot-number*]
显示TCP代理连接的简要信息	**display tcp-proxy slot** *slot-number*
显示TCP代理非保留端口的使用信息	**display tcp-proxy port-info slot** *slot-number*
显示TCP连接详细信息	**display tcp verbose** [**slot** *slot-number* [**pcb** *pcb-index*]]
显示UDP连接摘要信息	**display udp** [**slot** *slot-number*]
显示UDP连接详细信息	**display udp verbose** [**slot** *slot-number* [**pcb** *pcb-index*]]
显示IP报文统计信息	**display ip statistics** [**slot** *slot-number*]
显示TCP连接的流量统计信息	**display tcp statistics** [**slot** *slot-number*]

操作	命令
显示UDP流量统计信息	**display udp statistics** [**slot** *slot-number*]
显示ICMP流量统计信息	**display icmp statistics** [**slot** *slot-number*]
清除IP报文统计信息	**reset ip statistics** [**slot** *slot-number*]
清除TCP连接的流量统计信息	**reset tcp statistics**
清除UDP流量统计信息	**reset udp statistics**