

## 附录C RFC 1122 的有关内容

本附录总结了Net/3实现与RFC 1122[Braden 1989a]建议的兼容性。RFC 1122分4类给出了实现需求：

- 链路层
- IP层
- UDP
- TCP

我们将按照本书章节的顺序讨论这些实现要求。

### C.1 链路层的需求

本节依据RFC 1122中的2.5节总结了链路层需求和Net/3代码对这些需求的支持程度。

- 建议支持尾部封装。

部分支持：Net/3不发送带有尾部封装的IP数据报，但某些Net/3设备驱动程序能够接收此类数据报。感兴趣的读者可以阅读RFC 893和[Leffler et al. 1989]的11.8节。

- 没有协商之前，默认状态必须不发送尾部。

不支持：Net/2支持是否发送尾部启动的协商过程。Net/3忽略发送尾部请求，且不会向对端申请发送尾部。

- 必须能够发送和接收RFC 894的以太网封装。

支持：Net/3支持RFC 894的以太网封装。

- 应该能够接收RFC 1042(IEEE 802)封装。

不支持：Net/3能够处理收到的IEEE 802的封装格式，但只用于OSI协议栈。到达的按802.3封装的IP数据报将被ether\_input丢弃(图4-13)。

- 建议发送报文段实现RFC 1042封装格式，为此，系统还必须实现软件可配置的转换开关以选择合适的封装格式，且RFC 894应为默认值。

不支持：Net/3的发送报文段不支持RFC 1042封装格式。

- 必须向IP层提交链路层的广播报文。

支持：链路层通过置位mbuf数据报首部的M\_BCAST标志(或M\_MCAST多播标志)报告链路层的广播。

- 必须向链路层提交IP TOS值。

支持：Net/3没有明确地提交TOS值，而是做为链路层可利用的IP首部的一部分出现。

### C.2 IP的需求

本节总结了RFC 1122第3.5节建议的IP的需求以及本书介绍的Net/3系统对这些需求的支持程度。

- 必须实现IP和ICMP。

支持：`inetsw[0]`实现了IP，`inetsw[4]`实现了ICMP。

- 必须处理应用层的远端多接口 (multihoming) 通信。

支持：内核不区分远端多接口通信，因此，既不阻碍也不支持应用程序的这种通信方式。

- 建议支持本地的多接口通信。

支持：Net/3系统维护一个`ifnet`列表，且每个`ifnet`结构都带有一个`ifaddr`列表，即每个IP接口可配置多个IP地址，从而支持本地多接口通信。

- 如果转发IP数据报，则必须满足路由器规约。

部分支持：参见第18章，其中详细讨论了路由器的需求。

- 必须为内置的路由器功能提供使能选项，默认设置应为主机操作。

支持：`ipforwarding`变量默认值为FALSE，控制Net/3中的IP数据报转发机制。

- 必须禁止基于IP接口数的选路。

支持：`if_attach`函数并不根据系统初启时配置的接口数来修改 `ipforwarding` 变量。

- 应该记录丢弃的数据报，包括其内容，并在统计计数器中记录丢弃事件。

部分支持：Net/3没有提供一种机制，能够保存丢弃数据报的内容，但维护多种统计计数器。

- 必须丢弃IP版本号不等于4的数据报而不回显信息。

支持：`ipintr`实现此需求。

- 必须验证IP校验和，并丢弃验证失败的数据报而不回显信息。

支持：`ipintr`调用`ip_cksum`，实现此需求。

- 必须支持子网地址 (RFC 950)。

支持：在`in_ifaddr`结构中，所有的IP地址都有一个对应的子网掩码。

- 必须把主机自己的IP地址作为源IP地址，与数据报同时发送。

部分支持：如果运输层发送的IP数据报中，源地址为全0时，IP插入外出接口的IP地址作为源地址。应用进程可以把某个本地插口绑定在本地IP广播地址上，IP将其作为无效的源地址发送。

- 必须丢弃不是发往本地主机的数据报而不回显信息。

支持：如果系统没有被配置为路由器，`ipintr`丢弃目的地址差错 (无法辨认的单播、广播或多播地址) 的数据报。

- 必须丢弃源地址差错的数据报而不回显信息。

不支持：`ipintr`把数据报提交给运输层之前，不检测进入数据报的源地址。

- 必须支持重装。

支持：`ip_reass`实现重装。

- 建议为同一个IP数据报设定同样的ID。

不支持：`ip_output`为每个外出数据报赋一个新的ID，并且不允许运输层协议设定IP数据报的ID。参见第32章。

- 必须允许运输层设定TOS。

支持：`ip_output`接受运输层协议在IP首部设定的所有TOS值。运输层在默认情况下，

必须把TOS设为全0。应用进程可通过 `IP_TOS` 插口选项设定某个特定数据报或连接的TOS值。

- 必须把TOS值上交给运输层。

支持：Net/3在输入处理期间保存 TOS字段的值。当 IP 针对接收数据报的协议调用 `pr_input` 函数时，运输层可得到完整的 IP首部。不幸的是，UDP和TCP运输层协议忽略该字段。

- 应该不采用RFC 795 [Postel 1981d]中建议的TOS链路层匹配方式。

支持：Net/3没有使用这些匹配方式。

- 必须不发送TTL等于0的数据报。

部分支持：Net/3中的IP层(`ip_output`)不检查这项需求，而是依靠运输层以使得不会构造TTL等于0的IP首部。UDP、TCP、ICMP和IGMP都选择了一个非零的TTL默认值。但默认值可被 `IP_TTL` 选项忽略。

- 必须不丢弃TTL小于2的接收数据报。

支持：只要系统是数据报的目的地，`ipintr`将接受该数据报，并不测试其TTL值。只有在数据报需要被转发时，才会检测TTL值。

- 必须允许运输层设定TTL。

支持：运输层在调用 `ip_output` 之前，必须设定TTL。

- 必须允许配置一个固定的TTL值。

支持：全局整型变量 `ip_defttl` 中保存了TTL默认值，等于64(`IPDEFTTL`)。UDP和TCP都使用此默认值，除非应用进程通过 `IP_TTL` 插口选项为某个特定的插口指派了一个不同的值。通过调用 `sysctl`，通过指派 `IPCTL_DEFTTL`，可以修改 `ip_defttl`。

## 多接口

- 应该选取接收数据报中指定的目的地址作为响应的源地址。

支持：内核生成的响应 (ICMP响应报文段) 包含了正确的源地址 (C.5节)。运输层生成的响应在其各自章节中作了描述。

- 必须允许应用进程选取本地IP地址。

支持：应用程序能够把插口绑定在指派的本地IP地址上 (15.8节)。

- 建议丢弃目的地址与所到达的接口IP地址不同的数据报而不回显信息。

不支持：Net/3实现了一个简单的终端系统模型，`ipintr`接受此类的数据报。

- 建议数据报离开系统时所选接口的IP地址应与数据报的源地址一致。本需求不适用于源站选路的数据报。

不支持：Net/3允许数据报通过任意接口离开系统——另一个简单终端系统的特征。

## 广播

- 必须禁止在源地址中选用IP广播地址。

部分支持：如果应用程序明确地指派了源地址，IP层不会改变此设置。否则，IP选择与外出接口相连的IP地址作为源地址。

- 应该接受全0或全1的广播地址。

支持：ipintr接受发向上述任何一个地址的数据报。

- 建议提供选项，允许在指派接口上配置广播地址为全 0或全 1。如果提供该选项，可配置的广播地址默认值应为全 1。

不支持：应用进程必须明确地向全 0(INADDR\_ANY)或全 1(INADDR\_BROADCAST)广播地址发送数据报。没有配置默认值。

- 必须在链路层广播中使用 IP 广播地址或 IP 多播目的地址。

支持：只有当目的地址是 IP 多播地址或广播地址时，ip\_output才会置位链路层多播或广播标志。

- 应该丢弃链路层广播数据帧而不回显信息，如果它未指派某个 IP 广播地址作为其目的地址。

不支持：Net/3中，没有对输入数据报中的 M\_BCAST或M\_MCAST标志作明确测试，但ip\_forward在转发前会丢弃这些数据报。

- 对直接相连的网络，应使用受限的广播地址

部分支持：Net/3中，是否使用受限的广播地址(相对于子网广播地址和全网广播地址)由应用进程决定。

## IP 接口

- 必须允许运输层使用所有的 IP 机制(如 IP 选项、TTL 和 TOS)。

支持：Net/3 中的运输层可使用所有的 IP 机制。

- 必须向运输层提交 IP 接口号。

支持：每个保存进入数据报的 mbuf 中的 m\_pkthdr.rcvif 成员变量指向一个 ifnet 结构，其中保存了接收该数据报的接口的信息。

- 必须向运输层提交所有 IP 选项

支持：ipintr向运输层接收协议的 pr\_input 函数提交的数据报中，包含了完整的 IP 首部，包括各种 IP 选项。

- 必须允许运输层发送 “ICMP 端口不可达” 报文和其他所有 ICMP 查询报文。

支持：运输层调用 icmp\_error 可以发送任何 ICMP 差错报文；或者调用 ip\_output，构造并发送任何类型的 IP 数据报。

- 必须向运输层提交下列 ICMP 报文：目的地址不可达、源站抑制、回显回答、时间戳回答和数据报超时。

支持：ICMP 可以向其他运输层协议发送此类报文段，或通过原始 IP 机制向任何等待进程发送此类报文段。

- 必须在向运输层提交的 ICMP 报文中包括 ICMP 报文内容(IP 首部和附加数据)

支持：icmp\_input 向运输层提交包含在 ICMP 报文中的原始 IP 数据报。

- 应该在一次处理中完成所有功能。

不支持：也许在下一版 IP 中能够实现。

## C.3 IP 选项的要求

本节总结了 RFC 1122 第 3.5 节中 IP 选项处理的需求，以及本书介绍的 Net/3 系统对这些需求

的支持程度。

- 必须允许运输层设置IP选项。

支持：`ip_output`的第二个参数即为用于输出IP数据报的IP选项列表。

- 必须向高层提交收到的所有IP选项。

支持：IP首部及选项都能传递给运输层接收协议的 `pr_input` 函数。

- 必须忽略所有未知选项。

支持：`ip_dooptions`中的`default`语句跳过了所有未知选项。

- 建议支持安全选项。

不支持：Net/3不支持IP安全选项。

- 建议不发送流标识选项，并且必须忽略接收数据报中的该选项。

支持：Net/3不支持流标识选项，并忽略接收数据报中的该选项。

- 建议支持路由记录选项。

支持：Net/3支持路由记录选项。

- 建议支持时间戳选项。

部分支持：Net/3支持时间戳选项，但没有完全遵照规定的方式递增时间戳值。有时间戳请求时，源主机并未在报文段中插入时间戳，而是由对端主机在向运输层提交数据报之前记录时间戳值。时间戳值遵守 RFC 1122第3.2.2.8节中对ICMP时间戳报文段标准值的规定。

- 必须支持源站选路，必须能够成为源站选路报文段的终点。

支持：源站选路选项可作为传送给 `ip_output` 的参数，`ip_dooptions`能够正确地终止源站选路，并能保存该路由，在构造返回路由时使用。

- 向运输层提交数据报时，必须同时提交完整的源站选路路由。

支持：源站选路路由与其他数据报中可能出现的IP选项一起，提交给运输层。

- 必须构造正确的(非冗余的)返回路由。

不支持：Net/3只是简单地逆转收到的源选路路由，并不检查或纠正路由中可能存在的多余转发。

- 必须禁止在一个首部中发送多个源路由选项。

不支持：Net/3的IP层不禁止运输层在一个数据报中构造并发送多个源路由。

## 源路由转发

- 建议支持带源路由选项的数据报的转发。

支持：Net/3支持源路由选项。`ip_dooptions`实现所有功能。

- 处理源路由选项的同时，必须遵守相应的路由器原则。

支持：Net/3遵守路由器原则，无论数据报上是否包含源路由。

- 必须根据网关原则更新TTL。

支持：`ip_forward`实现本需求。

- 必须生成ICMP差错代码4和5(需要分片和源路由失败)。

支持：`ip_output`能够生成“需要分片”报文，`ip_dooptions`能够生成“源路由失败”报文。

- 必须允许带有源路由选项数据报的 IP 源地址不是转发主机的 IP 地址。

支持：ip\_output 发送此类数据报。

RFC 1122 将本需求指明为“建议”，因为地址可能不一致，但必须允许这种不一致。

- 必须更新时间戳和记录路由选项

支持：ip\_dooptions 为带有源路由选项的数据报处理这些选项。

- 必须支持一个可配置选项，用于打开或关闭“非本地源路由”。选项默认值应为关闭。

不支持：Net/3 允许非本地源选路，没有提供一个选项来关闭此功能。非本地源选路指在两个不同接口间转发数据报，而不是在同一接口接收和发送数据报。

- 非本地源选路处理中，必须满足网关接入规则。

支持：Net/3 在非本地源选路处理过程中，遵守转发规则。

- 如果无法转发源选路数据报（除了 ICMP 差错报文），应该发送 ICMP “目的地不可达” 差错（源路由失败）报文。

支持：ip\_dooptions 发送 ICMP “目的地不可达” 差错报文。如果处理的数据报是一个 ICMP 差错报文，icmp\_error 将丢弃新生成的 ICMP 差错报文。

## C.4 IP 分片与重装的需求

本节总结了 RFC 1122 第 3.5 节中关于 IP 分片和重装的需求，以及 Net/3 对这些需求的支持程度。

- 必须能够重装输入的数据报，数据报长度至少为 576 字节。

支持：ip\_reass 支持数据报的重装，且数据报的长度不限。

- 应该不限制输入数据报的长度，或者允许配置输入数据报长度的上限。

支持：Net/3 不限制输入数据报的长度。

- 必须提供某种机制，允许运输层了解接收的最大数据报长度。

不适用：Net/3 可接收的数据报长度只受可用存储器的限制。

- 重装超时，必须能够发送 ICMP “数据报超时” 差错报文。

不支持：Net/3 不发送 ICMP “数据报超时” 差错报文。参见图 10-30 和习题 10.1。

- 应该设定一个固定的重装超时值，且不应该采用收到的 IP 分片中 TTL 的剩余值作为重装超时值。

支持：Net/3 采用编译时指派的固定值 30 秒 (IPFRAGTTL 等于 60 个慢超时时间间隔，约为 30 秒) 做为重装超时值。

- 必须向高层提供 NMS\_S (可发送的最大报文段长度)。

部分支持：TCP 首先从到达目的地的路由表项中找到最大的 MTU，或者是读取外出接口的 MTU，并根据上述 MTU 计算出 NMS\_S。UDP 应用程序无法得到此信息。

- 必须支持对外出数据报的本地分片。

支持：如果 ip\_output 发现外出数据报长度大于指定外出接口的 MTU，则对其分片。

- 如果不支持数据报的本地分片，则必须禁止运输层发送长度大于 NMS\_S 的报文段。

不适用：这条对于运输层的需求不适用于 Net/3 系统，因为系统支持数据报的本地分片。

- 如果无法确知到达远端目的地路由的最小 MTU，则不应该向远端目的地发送大于 576 字



节的报文段。

部分支持：Net/3 TCP报文段的默认大小等于 553字节(512字节数据 + 40字节首部)。Net/3 UDP应用程序无法确认目的地址位于本地，或是远端，因此，通常都将报文段大小限制在540字节以下(512 + 20 + 8)。内核中没有机制禁止发送长度超出限制的报文。

- 建议支持“全部子网MTU”配置标志。

支持：全局变量subnetsarelocal默认为TRUE。TCP向本地网络中的某个子网发送报文段时，利用该标志选择较大的报文段长度（外出接口的MTU的大小），取代默认的报文段大小。

## C.5 ICMP的需求

本节总结了RFC 1122第3.5节中关于ICMP的需求，以及Net/3系统对这些需求的支持程度。

- 必须丢弃不了解的ICMP报文而不回显信息。

部分支持：icmp\_input忽略未知的ICMP报文，将其交给rip\_input，后者将报文段交给任何等待进程，或者在没有进程接收的情况下将其丢弃而不回显信息。

- 建议携带原始数据报中至少8字节的内容。

不支持：icmp\_error在ICMP差错报文中，最多返回原始数据报中8字节的内容，参见习题11.9。

- 必须原封不动地返回接收数据报的首部和数据。

部分支持：Net/3的ipintr将接收数据报的ID、偏移量和长度字段从网络字节序转换为主机字节序，从而方便数据报的处理。但 Net/3在把偏移量和长度字段放入 ICMP差错报文时，没有将这两个字段转换回网络字节序。如果系统的主机字节序与网络字节序相同，这个差错不会引起误解。但如果系统的主机字节序与网络字节序不一致，ICMP差错报文中携带的IP首部报文段中的偏移量和长度字段都是错误的。

作者发现，Intel版的SVR4和AIX 3.2(基于Net/2)返回的长度字段的字节顺序都是错误的，而实验过的其他不基于 Net/2和Net/3的实现(Cisco、NetBlazer、VM和Solaris2.3)却没有此差错。

此外，在UDP代码中发送ICMP“端口不可达”差错报文时，还有一个差错：接收数据报的首部长度被错误地修改了(23.7节)。作者在Net/2和Net/3系统中都发现了这个差错，而Net/1版中却没有。

- 必须能够将收到的ICMP差错分用给运输层协议。

支持：icmp\_error利用原始数据报首部的协议字段选择适当的运输层协议，以响应该差错。

- 发送ICMP差错报文时，TOS字段应该等于0。

支持：icmp\_error构造的所有ICMP差错报文的TOS字段都等于0。

- 必须禁止ICMP差错报文再次引发新的ICMP差错报文。

部分支持：ICMP重定向报文可能导致 icmp\_error发送新的ICMP差错报文。RFC 1122第3.2.2节中把ICMP重定向报文划分为ICMP差错报文。

- 必须禁止IP广播或多播报文引发ICMP差错外。

不支持：`icmp_error`不进行此类检查。

最初BSD中的Deering多播处理代码的`icmp_error`函数进行此类检查。

- 必须禁止链路层广播报文引发ICMP差错报文。

支持：`icmp_error`丢弃作为链路层广播或多播报文到达的ICMP报文。

- 必须禁止IP数据报的后续分片引发ICMP差错。

支持：`icmp_error`丢弃此类情况引发的ICMP差错报文。

- 必须禁止源地址不确定的数据报引发ICMP差错报文。

支持：`icmp_reflect`检查实验性地址和多播地址。`ip_output`丢弃源地址等于广播地址的数据报。

- 不属于禁止范围之内时，必须返回ICMP差错报文。

部分支持：一般情况下，Net/3发送适当的ICMP差错报文。但某些情况下，它无法发送ICMP“分片重装超时”报文(习题10.1)。

- 应该生成ICMP“目的站不可达”报文(协议和端口)。

不支持：如果数据报所指明的协议系统不支持，则交由`rip_input`函数处理。后者检查确认系统中没有应用进程能够处理此数据报后，将数据报丢弃而不回显信息。UDP生成ICMP“端口不可达”差错报文。

- 必须向高层提交ICMP“目的站不可达”差错报文。

支持：`icmp_input`向指定协议的`pr_ctlinput`函数(例如，UDP的`udp_ctlinput`函数，TCP的`tcp_ctlinput`函数)提交此类报文。

- 应该响应“目的站不可达”差错报文。

参见23.9节和27.6节。

- 必须把“目的站不可达”差错解释为一种暗示，可能只是一种临时状态。

参见23.9节和27.6节。

- 如果配置为主机，则必须禁止发送ICMP“重定向”报文段。

支持：`ip_forward`，唯一的检测和发送“重定向”报文的函数，只有在系统配置为路由器时才会被调用。

- 收到ICMP“重定向”报文时，必须更新路由表缓存。

支持：`ipintr`调用`rtredirect`，处理此报文。

- 必须能够处理“主机重定向”和“网络重定向”报文段。此外，必须把“网络重定向”报文作为“主机重定向”报文进行处理。

支持：`ipintr`调用`rtredirect`，处理这两类报文。

- 应该丢弃非法的重定向报文

支持：`rtredirect`丢弃非法的重定向报文(第19.7节)。

- 存储器不足时，建议发送“源站抑制”报文。

支持：如果`ip_output`返回`ENOBUFS`，`ip_forward`发送“源站抑制”报文。如果`mbuf`不足，或者接口输出队列已满时，会出现这种情况。

- 必须向高层提交“源站抑制”报文。

支持：`icmp_input`向运输层提交“源站抑制”差错报文。



- 高层应该响应“源站抑制”差错报文。  
详见23.9节和27.6节，UDP和TCP的处理逻辑。ICMP和IGMP都不接受ICMP差错报文（它们没有定义自己的`pr_ctlinput`函数）。这种情况下，IP将丢弃ICMP差错报文。
- 必须向运输层提交“数据报超时”差错报文。  
支持：`icmp_input`向运输层提交此类差错报文。
- 应该发送“数据报参数错”报文段。  
支持：`ip_dooptions`发现选项构造差错时，会发送此报文段。
- 必须向运输层报告出现数据报参数差错。  
支持：`icmp_input`向运输层报告此类差错。
- 建议向应用进程报告出现参数差错。  
详见23.9节和27.6节，UDP和TCP的处理逻辑。ICMP和IGMP都不接受ICMP差错报文。
- 必须支持回显服务器，应该支持回显客户。  
支持：`icmp_input`实现回显服务器，`ping`程序利用原始IP插口实现回显客户。
- 建议丢弃发往广播地址的回显请求报文。  
不支持：`icmp_reflect`发送应答。
- 建议丢弃发往多播地址的回显请求报文。  
不支持：Net/3响应发往多播地址的回显请求报文。`icmp_reflect`和`ip_output`都允许多播目的地址。
- 必须使用确定的目的地址作为回显回答报文的源地址。  
支持：`icmp_reflect`将广播地址或多播地址转换为接收接口的IP地址，并将转换后的地址用于回显回答报文的源地址。
- 必须在回显回答报文中返回回显请求数据。  
支持：`icmp_reflect`不更改回显请求报文的数据部分。
- 必须向高层提交回显回答报文。  
支持：ICMP回显回答报文提交给`rip_input`，进而交给指明的应用进程。
- 必须响应ICMP回显请求报文中携带的记录路由和时间戳选项。  
支持：`icmp_reflect`在回显回答报文中包括记录路由和时间戳选项。
- 必须逆转并响应源路由选项。  
支持：`icmp_reflect`调用`ip_srcroute`，获取逆转的源路由，并放入外出的回显回答报文中。
- 应该不支持ICMP信息请求和信息回答报文段。  
部分支持：内核不生成或响应这两类报文段，但应用进程可能会通过原始IP插口发送或接收这两类报文段。
- 建议实现ICMP时间戳请求和时间戳回答报文段。  
支持：`icmp_input`实现时间戳服务器的功能。时间戳客户的功能可通过原始IP机制来实现。
- 必须最小化时间戳延时偏移量(如果实现时间戳报文)。  
部分支持：接收时间戳在报文从IP输入队列中取出时加入，发送时间戳在报文放入接口输出队列前加入。

- 建议丢弃发往广播地址的时间戳请求而不回显信息。  
不支持：`icmp_input`响应发向广播地址的时间戳请求。
- 必须使用确定的目的地址作为时间戳回答报文的源地址。  
支持：`icmp_reflect`将广播地址和或多播地址转换为接收接口的IP地址，并将转换后的地址用于时间戳回答报文的源地址。
- 应该响应ICMP时间戳请求报文中携带的记录路由和时间戳选项。  
支持：`icmp_reflect`在时间戳回答报文中包括记录路由和时间戳选项。
- 必须逆转并响应ICMP时间戳请求报文中携带的源路由选项。  
支持：`icmp_reflect`调用`ip_srcroute`，获取逆转的源路由，并放入外出的时间戳回答报文中。
- 必须向高层提交时间戳回答报文。  
支持：ICMP时间戳回答报文提交给`rip_input`，进而交给指定的应用进程。
- 必须遵守有关标准时间戳值的规定。  
支持：`icmp_input`调用`iptime`，后者可返回标准时间戳值。
- 必须能够通过配置改变接口的地址掩码。  
不支持：Net/3通过`ifconfig`程序，只支持地址掩码的静态配置。
- 必须支持地址掩码的静态配置。  
支持：Net/3间接实现了这一功能。典型情况下，通过`/etc/netstart`批处理文件执行系统初始化，调用`ifconfig`程序配置接口时，可设定静态信息。
- 建议系统初始化时动态获取地址掩码。  
不支持：Net/3不支持利用BOOTP或DHCP，获取地址掩码信息。
- 建议通过ICMP地址掩码请求和回答报文获取地址掩码信息。  
不支持：Net/3不支持通过ICMP地址掩码请求和回答报文获取地址掩码信息。
- 如果没有响应，必须重传地址掩码请求。  
不适用：Net/3不支持此项功能。
- 如果没有响应，建议使用假定的默认地址掩码  
不适用：Net/3不支持此项功能。
- 只允许在收到第一个响应时更新地址掩码。  
不适用：Net/3不支持此项功能。
- 建议对所有已安装地址掩码进行合理的检测。  
不支持：Net/3不对地址掩码进行检测。
- 必须禁止响应未确认的地址掩码请求报文，且必须被明确地配置为代理。  
支持：`icmp_input`只有当`icmpmaskrepl`非零时（默认为0），才会响应地址掩码请求报文。
- 应该为每个静态配置的地址掩码设定相应的地址掩码确认标志。  
不支持：Net/3只维护一个全局的确认标志（`icmpmaskrepl`），发送任何接口的地址掩码回答报文之前都要查询同一个全局变量。
- 必须在初始化时广播地址掩码回答报文。  
不支持：Net/3配置接口时，不广播地址掩码回答报文。

## C.6 多播的需求

本节总结了RFC 1122第3.5节关于IP多播功能的需求，以及Net/3系统对这些需求的支持程度。

- 应该支持本地IP多播(RFC 1122)。

支持：Net/3支持IP多播。

- 应该在启动时加入全主机组。

支持：in\_ifinit初始化接口时加入全主机组。

- 应该为高层提供一种机制，使其能够了解接口的IP多播功能。

支持：内核代码能够直接访问接口ifnet结构中的IFF\_MULTICAST标志，应用进程通过SIOCGIFFLAGS命令也能做到这一点。

## C.7 IGMP的需求

本节总结了RFC 1122第3.5节关于IGMP功能的需求，以及Net/3系统对这些需求的支持程度。

- 建议支持IGMP(RFC 1122)

支持：Net/3支持IGMP。

## C.8 选路的需求

本节总结了RFC 1122第3.5节关于IP选路功能的需求，以及Net/3系统对这些需求的支持程度。请注意，RFC 中的这些需求只适用于主机，而非内核的实现。Net/3内核的选路函数没有明确实现其中的某些条款，但这些功能都包括在后台选路进程如routed或gated中。

- 必须使用地址掩码来确定数据报的目的地址是否位于直接相连的网络中。

支持：在配置连通某个网络(如以太网)的接口时，同时也配置了接口的地址掩码(或根据IP地址的类别选择一个默认的地址掩码)，保存在路由表表项中。rn\_match查找网络匹配时，会用到已配置的地址掩码。

- 不存在路由器(所有网络都直接相连)时，必须能在最小环境中运行正常，

支持：这种情况下，系统管理员不允许配置默认路由。

- 必须在缓存中保存到达下一跳路由器的路由。

支持：路由表位于缓存中。

- “网络重定向”报文的处理方式应该等同于“主机重定向”报文。

支持：详见19.7节。

- 必须使用默认路由器，如果路由表中没有到达目的地址的路由记录。

支持：条件是路由表中已配置有默认路由。

- 必须支持多个默认路由器。

内核不支持多条默认路由。完成选路的后台进程可能支持此功能。

- 建议实现静态路由表。

支持：可在系统初始化时通过route命令实现。

- 建议为每条静态路由指派一个标志，说明它是否能被重定向报文修改。

不支持。

- 建议采用完整的主机地址，而不是网络地址作为路由表的表项

支持：主机路由比到达同一网络的网络路由具有优先权。

- 应该在路由表表项中包括TOS值。

不支持：第21章中描述的sockaddr\_inarp结构中定义了TOS字段，但目前未使用。

- 必须能够检测路由表中出现在网关域的下一跳路由器的故障，并能选择其他的下一跳路由器。

消极的建议。in\_losing生成的RTM\_LOSING报文，将被上交给从选路插口读取数据的所有进程，从而允许应用进程(如，某个选路后台进程)处理该事件。

- 不应该假定一条路由永远正常。

支持：除了ARP生成的表项外，Net/3内核路由表中的其余表项没有超时字段。UNIX系统标准的后台选路进程负责对路由表项定时，超时后，在可能的情况下，选择另一条路由替代已超时的路由。

- 必须禁止连续ping(ICMP回应请求)路由器。

支持：Net/3内核不会这样做。后台选路进程也不会生成ICMP回应请求报文。

- 只有在需要向路由器发送报文时，才允许ping路由器。

Net/3内核绝不会ping下一跳路由器。

- 应该实现某种机制，允许高层或低层向路由模块报告正常或差错。

部分支持：其他层向Net/3选路函数传递信息的唯一方式是通过in\_losing，而in\_losing只被TCP调用。选路层采取的唯一动作是生成RTM\_LOSING报文段。

- 默认路由器出现故障时，必须切换到另一个默认路由器上。

支持：尽管Net/3内核不实现此功能，但得到后台选路进程支持。

- 必须能够手工配置路由表中的下述信息：IP地址、网络掩码和默认路由表。

支持：但内核只支持一条默认路由。

## C.9 ARP的需求

本节总结了RFC 1122第2.5节关于ARP功能的需求，以及Net/3系统对这些需求的实现程度。

- 必须提供某种机制，能够清除过时的ARP记录。如果利用超时，时限应是可配置的。

支持：arptimer提供所要求的机制。时限是可配置的(arp\_prune和arp\_keep全局变量)，但改变时限值的唯一方式是重新编译内核，或通过调试器修改内核。

- 必须提供某种机制，防止ARP泛滥

支持：详见图21-24。

- 应该保存(而非丢弃)至少一个(最后一个)，发往同一个未解析的IP地址的数据报，并且在IP地址解析后发送保存的数据报。

支持：这就是定义llinfo\_arp结构中la\_hold成员变量的目的。

## C.10 UDP的需求

本节总结了RFC 1122第4.1.5节关于UDP功能的需求，以及Net/3系统对这些需求的实现程度。

- 应该发送ICMP“端口不可达”差错报文。

支持：`udp_input`完成此功能。

- 必须向应用程序提交接收的IP选项。

不支持：`udp_input`中实现此功能的代码被注释掉了。也就是说，即使应用进程收到的UDP报文段中带有源选路选项，也无法采用逆向路由发送响应。

- 必须允许应用进程指派发送的IP选项。

支持：`IP_OPTIONS`插口选项实现此功能。指派的选项保存在PCB中，并由`ip_output`放入输出的IP数据报中。

- 必须能向下向IP层传递IP选项。

支持：前面已提到，IP把选项放入IP数据报中。

- 必须向应用程序提交收到的ICMP报文。

支持：我们必须阅读RFC原文：“基于UDP的应用程序，如果希望接收ICMP差错报文，则应负责维护必需的状态，从而在报文段到达时能够正确处理。例如，应用程序可能需要为此保存一个挂起的接收操作。”基于Berkeley系统所需的状态，是指插口已连接到远端地址和插口上。如同图23-26起始处的注释所指出的，某些应用程序为指派的远端端口同时创建连接插口和非连接插口，利用连接插口接收异步差错。

- 必须能够生成并验证UDP检验和。

支持：`udp_input`在全局变量`udpcksum`的基础上实现此功能。

- 必须丢弃检验和验证失败的UDP报文段而不回显信息。

支持：只要`udpcksum`非零，Net/3丢弃该报文段。我们前面曾提到，该变量同时控制发送时检验和的产生以及接收检验和的验证。如果它等于0，则内核对收到的非零检验和不做验证。

- 建议允许发送程序指定是否需要计算输出报文段的检验和，默认操作必须为需要计算

不支持：应用程序不能控制UDP检验和。默认情况下，Net/3计算UDP检验和，除非内核编译时定义了4.2BSD的兼容功能，或者系统管理员通过`sysctl(8)`关闭了UDP检验和功能。

- 建议允许接收进程指定是丢弃收到的不带检验和的UDP报文段(例如，收到的检验和等于0)，还是将此类报文段提交给应用进程。

不支持：即使接收报文段的检验和字段等于0，也会被提交给应用进程。

- 必须向应用进程提交目的IP地址。

支持：应用程序可以调用`recvmsg`，并指派`IP_RECVDSTADDR`插口选项。尽管在图23-25中的讨论中曾指出，如果目的地址是广播地址或多播地址，4.4BSD不遵守此规定。

- 必须允许应用进程指派发送UDP报文段时所使用的本地IP地址。

支持：应用程序调用`bind`，为UDP插口指派本地IP地址。在第22.8节结尾处，我们已经讨论了源IP地址和输出接口IP地址间的差别。Net/3不允许应用程序指派输出接口——`ip_output`负责根据到达目的地址的路由选取本地输出接口。

- 必须允许应用程序指派本地IP地址的通配地址。

支持：如果`bind`调用中指派的IP地址为`IPADDR_ANY`，`in_pcbconnect`将根据到达目的地址的路由选取本地IP地址。

- 应该允许应用程序了解选定的本地 IP 地址。

支持：应用程序必须调用 `connect`。如果插口未建立连接，且绑定时指派的是本地通配地址，在该插口上发送报文段时，`ip_output` 选择输出接口，并把输出接口的 IP 地址作为源地址。但在 `sendto` 返回前，`udp_output` 结尾处的代码会把 PCB 中的 `inp_laddr` 成员变量，重置为通配地址。因此，`getsockname` 将返回空值。但应用程序可以调用 `connect`，把 UDP 插口连接到指定目的地，强迫 `in_pcbconnect` 选择输出接口，并把接口地址保存到 PCB 中。应用程序随后调用 `getsockname`，可得到本地接口的 IP 地址。

- 必须丢弃收到的源地址差错（广播地址或多播地址）的 UDP 报文段而不回显信息。

不支持：即使收到的 UDP 报文段源地址差错，但如果有插口绑定在指派的目的端口上，则报文段也会被提交给该插口。

- 必须发送有效的 IP 源址。

支持：如果通过 `bind` 指派本地 IP 地址，`bind` 会检查地址的有效性。如果指派了本地通配地址，`ip_output` 选择本地地址。

- 必须实现 RFC 1122 第 3.4 节定义的完整的 IP 接口。

参见 C.2 节

- 必须允许应用进程为输出报文段指派 TTL、TOS 和 IP 选项。

支持：应用程序可使用 `IP_TTL`、`IP_TOS` 和 `IP_OPTIONS` 插口选项。

- 建议向应用程序提交 TOS。

不支持：应用程序无法得到接收报文段 IP 首部的 TOS 值。请注意，调用 `getsockopt`，参数为 `IP_TOS` 时得到的返回值是输出报文段，而非接收报文段的 TOS 值。接收报文段的 TOS 对 `udp_input` 是可见的，但后者在处理过程中，将它与整个 IP 首部一起丢弃。

## C.11 TCP 的需求

本节总结了 RFC 1122 第 4.2.5 节关于 TCP 功能的需求，以及 Net/3 系统对这些需求的实现程度。

### PSH 标志

- 建议累积用户发送的不带 PSH 标志的数据。

部分支持：Net/3 没有为应用进程提供机制，能够在调用 `write` 的同时指派 PSH 标志。但 Net/3 能够累积用户在多次 `write` 操作中发送的数据。

- 建议把收到的不带 PSH 标志的数据放入队列。

不支持：接收报文段中是否带有 PSH 标志不会影响 Net/3 的处理。接收报文段处理过程中，收到的数据都放入插口的接收队列。

- 发送方在对数据打包时，应该合并连续的 PSH 标志。

不支持。

- 建议在 `write` 调用中指派 PSH 标志。

不支持：插口 API 不提供此项功能。

- 由于 PSH 标志不是 `write` 调用的一部分，必须防止无限期地缓存数据，且必须在最后一个缓存报文段中置位 PSH 标志。



支持：基于 Berkeley 的系统都采用此方法。

- 建议向应用程序提交收到的 PSH 标志。

不支持：插口 API 不提供此项功能。

- 在可能的情况下，应该发送最大长度报文段，以提高性能。

支持。

## 窗口

- 必须把窗口大小视为无符号整数。应该把窗口大小视为 32 bit 数值。

支持：图 24-13 中的所有窗口大小的数据类型都是 `unsigned long`。RFC 1323 关于窗口大小选项的说明中也有此要求。

- 不允许接收方缩小窗口。

支持：详见图 26-29。

- 发送方必须非常灵活，在对端缩小窗口时也能正常运作。

支持：详见图 29-15。

- 建议无限期关闭提供的接收窗口。

支持。

- 发送方必须能够探测零窗口。

支持：这也是设置持续定时器的目的。

- 在窗口因为 RTO 而关闭时，应该发送第一次零窗口探测报文段。

不支持：Net/3 把持续定时器的下限设为 5 秒，一般情况下都大于 RTO。

- 应该线性递增连续探测报文段间的时间间隔。

支持：详见图 25-14。

- 必须允许对端窗口无限期地关闭。

支持：TCP 会一直向关闭窗口发送探测报文段。

- 不允许由于接收方一直发送零窗口通告，发送方就超时关闭连接。

支持。

## 发送数据

- 紧急指针必须指向紧急数据的最后一个字节。

不支持：基于 Berkeley 的系统都将紧急指针解释为指向紧急数据结束后的第一个字节。

- 必须支持任何长度的紧急数据。

支持：条件是修订了习题 26.6 中讨论的差错。

- 必须通知接收进程：(1) TCP 收到紧急指针，并且没有正等待处理的紧急数据；或者 (2) 紧急指针出现在数据流之前。

支持：详见图 29-17。

- 必须允许应用进程以某种方式确定剩余的紧急数据量，或者至少能确定是否还有需要读取的紧急数据。

支持：这就是设置带外数据标志 `SIOCATMARK` 的目的。

## TCP选项

- 必须能够接收任何报文段中的TCP选项。  
支持。
- 必须忽略所有不支持的选项。  
支持：详见第28.3节。
- 必须处理非法的选项长度。  
支持：详见第28.3节。
- 必须能够发送并接收MSS选项。  
支持：图28-10中的代码处理收到的MSS选项，图26-23中的代码总是在SYN中发送MSS选项。
- 如果接收MSS不等于536，则应该在响应SYN报文段中发送MSS选项；建议在所有SYN报文段中发送MSS选项。  
支持：前面已提到，Net/3在所有SYN报文段中发送MSS选项。
- 如果收到的SYN报文段中未携带MSS选项，必须假定MSS默认值等于536。  
不支持：MSS的默认值等于512，而非536。

这可能是一个历史遗留问题。因为 VAX系统物理存储器页大小为 512字节，而 trailer协议只能处理长度为512倍数的数据。

- 必须计算“有效发送MSS”。
- 支持：详见27.5节。

## TCP检验和

- 必须在输出报文段中生成TCP检验和，必须验证收到的检验和。
- 支持：Net/3支持TCP检验和的生成和验证。

## 初始序号选择

- 必须使用RFC 793中规定的时钟驱动的选择机制。
- 不支持：RFC 793中规定的时钟，每半秒递增 125 000，而Net/3 ISN(全局变量tcp\_iss)每半秒递增 64 000，约为规定速率的一半。

## 打开连接

- 必须支持同时打开。  
支持：尽管基于Berkeley的系统，4.4BSD以前的版本，不支持此功能。第28.9节中讨论了这个问题。
- 必须能区分是从LISTEN状态，还是从SYN\_SENT状态变迁到SYN\_RCVD状态。  
支持：结果相同，方法不同。提出此需求的目的是允许被动打开的一方，在收到RST时能返回到LISTEN状态(如图24-15所示)，而主动打开的一方，如果在SYN\_RCVD状态时收到RST，应被强迫终止。如图28-36所示。
- 被动打开必须不影响系统中已建立的连接。

支持。

- 必须允许绑定在同一本地端口上的监听插口和另一插口同时处于 `SYN_SENT` 状态或 `SYN_RCVD` 状态。

支持：提出本需求的目的在于允许应用进程能同时接受多个连接请求。Berkeley系统采用的方法是，收到SYN时，系统根据处于LISTEN状态的插口创建一条完全相同的连接。

- 如果在多接口主机上，执行主动打开的应用进程没有指派源IP地址，则必须要求IP层选择一个本地IP地址作为源IP地址。

支持：由 `in_pcbconnect` 实现。

- 同一连接上发送的所有报文段必须使用相同的源IP地址。

支持：只要 `in_pcbconnect` 选定了源IP地址，就不会再改变。

- 执行主动打开时，对端地址不允许是广播地址或多播地址。

部分支持：TCP不会向广播地址发送报文段，因为图 26-32中调用 `ip_output` 时，不会指定 `SO_BROADCAST` 选项。但Net/3允许应用程序试图与多播地址建立连接。

- 必须忽略收到无效源地址的SYN报文段。

支持：图28-16中的代码检查无效源地址。

## 关闭连接

- 应该允许RST携带数据。

不支持：图28-36中对RST的处理，结束时跳转到 `drop`，略过了图 29-22中对报文段数据的所有处理。

- 必须通知应用进程对端是正常关闭连接（例如，发送了FIN），还是通过RST异常中止了连接

支持：如果收到FIN，`read`系统调用返回0（文件结束）；如果收到RST，`read`系统调用返回-1，差错代码为 `ECONNRESET`。

- 建议实现半关闭。

支持：应用进程调用 `shutdown`，令第二个参数等于1，可发送FIN。此后，应用进程仍能从连接读取数据。

- 如果应用进程完全关闭了连接（不是半关闭），但接收数据还没有被读取，或者关闭操作后，又有新的数据到达，则TCP应该发送RST，说明有数据丢失。

部分支持：如果应用进程调用 `close`，且没有读取插口接收缓存中的数据，则TCP不发送RST。但如果插口被关闭后，又有新的数据到达，则TCP将发送RST报文段。

- 必须在 `TIME_WAIT` 状态等待2MSL。

支持：尽管Net/3 MSL等于30秒，远小于RFC 793中建议的2分钟的时间长度。

- 如果在 `TIME_WAIT` 状态收到对端新发送的SYN，应允许直接建立新连接。

支持：详见图28-28。

## 重传

- 必须实现Van Jacobson的慢启动和拥塞避免算法。

支持。

- 如果重传报文段与原始报文段相同，建议使用同一个 IP 标识符。  
不支持：`ip_output`把全局变量 `ip_id` 的当前值填入 IP 数据报的 ID 字段。每发送一个 IP 数据报，`ip_id` 加 1。TCP 不负责 IP 标识符的赋值。
- 必须实现 Jacobson 的 RTO 算法和选取的 RTT 测量值 Karn 算法。  
支持：请注意，如果采用 RFC 1323 定义的时间戳选项，重传报文段二义性将不再存在，Karn 算法的一半问题也解决了。在图 29-6 中我们已讨论过这个问题。
- 对于连续的 RTO 值，必须有指数退避机制。  
支持：详见图 25-22。
- SYN 报文段的重传算法应该与数据报文段的重传算法相同。  
支持：详见图 25-16。
- 应该初始化往返时间估值参数，保证计算得到的 RTO 的初始值为 3 秒。  
不支持：`tcp_newtcpcb` 计算得到的 `t_rxtcur` 初始值等于 6 秒。详见图 25-16。
- RTO 的下限应为几分之一秒，上限应等于 2MSL。  
不支持：RTO 的下限设为 1 秒，上限等于 64 秒(图 25-3)。

#### 生成 ACK

- 应该把乱序报文段放入重组队列。  
支持：`tcp_reass` 实现此功能。
- 发送 ACK 之前，必须先处理队列中的所有报文段。  
支持：但只适用于顺序到达的报文段。`ipintr` 处理 IP 接收队列中的数据报，如果携带的是 TCP 报文段，则调用 `tcp_input`。对于顺序到达的报文段，`tcp_input` 设定延迟 ACK 标志，控制返回 `ipintr`。如果 IP 输入队列中还有其他携带 TCP 报文段的数据报，`ipintr` 依次调用 `tcp_input`。只有 IP 输入队列中的数据报全部处理完毕后，才能调用 `tcp_fasttimo` 生成延迟 ACK，确认 `tcp_input` 处理过的全部报文段中最高的数据字节序号。  
处理乱序报文段遇到的问题是：把控制返回给 `ipintr` 之前，`tcp_input` 会直接调用 `tcp_output`，生成对乱序报文段的确认。如果此时 IP 输入队列中还有剩余报文段，它们与当前处理的乱序报文段合在一起完全有可能组成一个顺序数据序列，但它们只能等到立即 ACK 发送完毕后才会被处理。
- 建议立即发送乱序报文段的 ACK。  
支持：这也是快速重传和快速恢复算法的要求(第 29.4 节)。
- 应该实现延迟 ACK，且延迟时间应小于 0.5 秒。  
支持：`tcp_fasttimo` 每隔 200 ms，检查一次 `TF_DELACK` 标志。
- 每隔一个报文段，至少应该发送一次 ACK。  
支持：图 26-9 中的代码每隔一个报文段生成一个 ACK。我们也讨论过，只有当接收数据的应用进程在数据到达时立即读取，才会出现这种处理过程。因为只有在 `PRU_RCVD` 的处理代码中，才会调用 `tcp_output`，每隔一个报文段发送一次 ACK。
- 接收方必须实现糊涂窗口综合症避免算法。  
支持：详见图 26-29。

## 发送数据

- TCP报文段中的TTL值必须是可配置的。  
支持：`tcp_newtcpcb`初始化TTL为64(`IPDEFTTL`)，但应用进程可通过`IP_TTL`插口选项修改该值。
- 发送方必须实现糊涂窗口综合症避免算法。  
支持：详见图26-8。
- 应该实现Nagle算法。  
支持：详见图26-8。
- 必须允许应用进程对于指定连接禁止Nagle算法。  
支持：通过`TCP_NODELAY`插口选项。

## 连接失败

- 如果某报文段的重传次数超过指定的门限R1，必须向IP层告警。  
支持：图25-26中，R1等于4。如果重传次数超过4，则调用`in_losing`。
- 如果某报文段的重传次数超过R2，必须关闭连接。  
支持：R2等于12(图25-26)。
- 必须允许应用进程配置R2的值。  
不支持：在图25-26中，R2的值是定死的。
- 如果重传次数超过R1，而小于R2，应该通知应用进程。  
不支持。
- R1的默认值最小应该为3次，R2最小应该为100秒。  
支持：R1等于4，最小RTO等于1秒。`tcp_backoff`数组(第25.9节)确保R2的最小值大于500秒。
- SYN报文段重传的处理方式必须与数据报文段重传的处理方式相同。  
支持：但一般情况下，SYN重传次数不会超过R1(图25-16)。
- 对于SYN报文段，R2的最小值必须设为3分钟。  
不支持：连接建立定时器把SYN的R2限定为75秒(图25-16)。

## 连接探测报文段

- 建议实现连接探测报文段。  
支持：Net/3提供此功能。
- 必须允许应用进程打开或关闭保活功能，默认值为关闭。  
支持：默认值为关闭。应用进程必须通过`SO_KEEPALIVE`插口选项打开此功能。
- 只有当连接空闲时间超过限制时，才允许发送连接探测报文段。  
支持。
- 必须允许配置保活的时间间隔，默认值必须大于2小时v。  
部分支持：发送连接探测报文段前的时间间隔很难配置，但默认值等于2小时。如果默认的空间时间间隔被更改(修改全局变量`tcp_keepidle`)，它会影响主机上设置了保活

选项的所有用户——无法如许多用户所希望的，为每条连接单独进行配置。

- 即使对端未响应特定探测报文段，不允许立即认为连接已中断。

支持：在确认连接中断之前，Net/3系统会发送9个探测报文段。

## IP选项

- 必须忽略接收报文段中不支持的IP选项。

支持：IP层实现此功能。

- 建议支持接收报文段中的时间戳选项和记录路由选项。

不支持：Net/3只响应ICMP报文中的这些选项，把这些选项反转给发送方（icmp\_reflect）。tcp\_input调用ip\_stripoptions，丢弃所有收到的IP选项，如图28-2所示。

- 主动打开连接时，应用进程必须能够指派源路由，而且该路由应该优先于在该连接上收到的源路由。

支持：应用进程可通过IP\_OPTIONS插口选项，指派源路由。如果连接是主动打开的，tcp\_input不会查看接收报文段中的源路由。

- 如果连接是被动打开，且发送报文段时，必须使用收到的源路由的逆转路由，则必须保存连接上收到的源路由。如果后续报文段中携带的源路由与当前保存的路由不同，新路由将取代原有路由。

部分支持：只有在监听插口上收到SYN时，图28-7中的代码才会调用ip\_srcroute。如果后续报文段携带了新路由，则被忽略。

## 接收IP层提交的ICMP报文

- 收到ICMP源站抑制报文段后，应该执行慢启动。

支持：tcp\_ctlinput调用tcp\_quench函数。

- 收到ICMP网络不可达、主机不可达或源路由失败的报文段后，必须禁止TCP终止连接，并应该通知应用进程。

部分支持：如图27-12所示，Net/3完全忽略已建立连接上收到的网络不可达和主机不可达报文段。

- 收到ICMP协议不可达、端口不可达和需要分片但DF置位的报文段后，应该中断当前连接。

不支持：tcp\_notify只在t\_softerror中记录这些ICMP差错。如果连接最终被丢弃，则向应用进程报告。

- 收到处理数据报超时和数据报参数错的报文段后，应该采用与前述网络不可达和主机不可达报文段同样的处理方式。

支持：tcp\_notify只在t\_softerror中记录ICMP数据报参数错。tcp\_ctlinput忽略ICMP数据报超时报文段。两种报文段都不会导致连接被丢弃。

## 应用程序编程接口

- 必须提供某种方式，向应用进程报告软差错，一般应通过异步方式。



不支持：连接被丢弃时，通过返回值向应用进程报告软差错。

- 必须允许应用进程为连接上发送的报文段指派 TOS。应该允许应用进程在连接的生存期内，动态更改TOS。

支持：应用程序通过IP\_TOS选项可完成上述功能。

- 建议向应用进程提交最近收到的 TOS值。

不支持：插口 API 不提供此项功能。调用 `getsockopt`，参数为 `IP_TOS`，只能返回用于发送的TOS当前值。它无法返回最近收到的TOS值。

- 建议实现 “flush” 调用。

不支持：TCP尽可能快地发送应用进程数据。

- 必须允许应用进程在主动打开或被动打开之前，指派本地 IP地址。

支持：应用进程可在调用 `connect` 或 `accept` 之前，调用 `bind`。