

在 Windows 和 Sun 系统中调整 IP MTU、TCP MSS 和 PMTUD

目录

[简介](#)

[先决条件](#)

[要求](#)

[使用的组件](#)

[规则](#)

[背景信息](#)

[问题描述与可能原因](#)

[TCP MSS 值](#)

[PMTUD 故障](#)

[不同的 LAN 介质类型](#)

[“哑铃”网络拓扑](#)

[缺陷链路](#)

[如何在终端节点上禁止 PMTUD 并配置更小的 MTU/MSS ?](#)

[Solaris 10 \(及更低版本 \)](#)

[HP-UX 9.x、10.x 与 11.x](#)

[IBM AIX Unix](#)

[Linux](#)

[Windows 95/98/ME](#)

[Windows NT 3.1/3.51](#)

[Windows NT 4.0](#)

[Windows 2000/XP](#)

[相关信息](#)

简介

本文档将举例说明如何在 Solaris 10 (及更低版本) ， HP-UX 9.x、10.x、11.x ， IBM AIX ， Linux ， Windows 95/98/ME ， Windows NT 3.1/3.51 ， Windows NT 4.0 和 Windows 2000/XP 上配置节点来调整 IP 最大传输单元 (MTU)、TCP 最大分段大小 (MSS) 和 IP 路径 MTU 发现 (PMTUD)。

先决条件

要求

本文档的读者应掌握以下这些主题的相关知识：

- [IP 分段和路径 MTU 发现 \(PMTUD\)](#)
- [传输控制协议 \(TCP\)](#)

使用的组件

本文档不限于特定的软件和硬件版本。

规则

有关文档规则的详细信息，请参阅 [Cisco 技术提示规则](#)。

背景信息

由于网络硬件故障、配置错误或软件缺陷，您也许会发现以下情况：小容量 TCP 数据传输可以顺利进行，但大容量数据传输（最大长度的数据包）会暂停然后超时。一种解决方法是配置发送节点，以执行下面的一项或两项操作：

- 禁用 PMTUD。
- 压缩 TCP MSS 和/或 IP MTU，以减少最大数据包大小。

问题描述与可能原因

有时，TCP/IP 节点可以通过某些 IP 路径顺利发送小容量数据（一般少于 1500 字节），但在尝试传输较大容量数据时，会暂停然后超时。通常，这种大容量数据只能以一个方向传输，无法在另一个方向上传输的问题被视为单向性问题。此问题很可能由 TCP MSS 值、PMTUD 故障、不同的 LAN 介质类型、或有缺陷的链路引起。以下小节将描述这一问题：

TCP MSS 值

TCP MSS 值指定本地系统能够接受(重新召集)的单个 IP 数据包中的 TCP 数据的最大数量。在发送时，IP 数据报可以分段成多个数据包。理论上，此值最高可达 65495，但从未使用过这么大的值。通常，终端系统使用“传出接口 MTU”减去 40 后的值作为其报告的 MSS。例如，以太网 MSS 值为 1460 (1500 - 40 = 1460)。

PMTUD 故障

PMTUD 是 [RFC 1191](#) 中描述的一种算法，在最新的 TCP/IP 协议栈中得以实施。 [此算法尝试找到无需分段就可以通过 IP 路径发送的最大 IP 数据报，并最大化数据传输吞吐量。](#)

通过让 IP 发送方在 IP 报头中设置“不分段”(DF) 标记，可以实施 PMTUD。“如果一个设置这个标志位的 IP 数据包到达下一台路由器，该路由器的下一跳链路的 MTU 过小，无法发送没有分段的数据包，则该路由器会丢弃数据包，并向 IP 发送方发送 ICMP ""Fragmentation needed but DF set"" 错误消息。”当 IP 发送方收到此 Internet Control Message Protocol (ICMP) 消息时，它会使用较小的 IP MTU 将数据包发送到此目标，并使后续数据包也能顺利通过。

有很多问题都可能导致 PMTUD 算法发生故障。IP 发送方不会自动使用较小的路径 MTU，而是继续不成功地重新传输过大的数据包，直到重新传输超时。问题包括：

- 路由器的下一跳路径过小，无法生成必要的 ICMP 错误信息。
- 小型 MTU 路由器和 IP 发送方之间的反向路径中的某台路由器，在 ICMP 错误信息到达 IP 发送方之前丢弃该信息。
- IP 发送方的堆栈中存在混乱，导致 IP 发送方忽略收到的 ICMP 错误信息。

针对这些问题的一种解决方法是，配置 IP 发送方以禁用 PMTUD。这会导致 IP 发送方发送的数据报中 DF 标记被清除。当大量信息包到达小型 MTU 路由器上时，该路由器将信息包分割成多个更小的片段。经过分段的较小数据在到达目标后，会被重组为原始大型数据包。

不同的 LAN 介质类型

二台主机在同一个路由网络上，但却在不同的 LAN 介质类型中(以太网、令牌环和光纤分布式数据接口 FDDI)能执行不同的操作。当令牌环和 FDDI 连接的系统发生故障时，以太网连接的系统仍能正常工作。造成此故障的原因是，以太网系统报告的 MSS 值为 1460，而令牌环和 FDDI 连接的系统报告的 MSS 值却为 4400 左右。因为远程服务器无法超过另一端报告的 MSS 值，所以当它与以太网连接的系统通信时，使用的数据包可能会比它与令牌环和 FDDI 连接的系统通信时使用的数据包小。

“哑铃”网络拓扑

PMTUD 问题在“Dumbbell”网络拓扑中比较常见，在这种拓扑中，网络路径中的内部链路的 MTU 少于通信主机的接口的 MTU。例如，如果您使用 IP (通用路由封装 (GRE)) 隧道，隧道接口的 MTU 小于相应物理接口的 MTU。如果 PMTUD 由于 ICMP 过滤或主机堆栈问题发生故障，则大型数据包无法穿过隧道。Cisco IOS 软件版本中针对 Cisco bug ID [CSCdk15279](#) (仅限[注册](#)用户) 的一个解决方法是，将隧道的 IP MTU 增加到 1500B。

缺陷链路

有时，路由器的某个链路具有较大的 (1500 字节) MTU，但路由器无法通过该链路传送同样大小的数据报。该路由器不会向发送方返回“需要分段但不需要 DF 设置”的 ICMP 错误，因为该链路的 MTU 实际上并不小。但是，大型数据报无法通过该链路。因此，PMTUD 不起作用，所有通过此链路的大型数据包传输尝试都会失败。

这有时是因为链路发生了下层问题，例如帧中继电路的 MTU 过小、缓冲过小；信道服务单元/数据服务单元 (CSU/DSU) 或转发器发生故障；电缆不符合规范；软件或固件缺陷。

下表显示了相关的软件缺陷。

- Cisco Bug ID [CSCdj11304](#) (仅限[注册](#)用户)
- Cisco Bug ID [CSCdi75411](#) (仅限[注册](#)用户)
- Cisco Bug ID [CSCdj74245](#) (仅限[注册](#)用户)

链路的另一个下层问题是因为使用了无法执行 IP 层分段的不合格“FDDI 到以太网”网桥。一种可能的解决方法是，在连接到故障链路的路由器接口上配置较小的 MTU。但是，这可能不是一个好方法，而且可能不完全有效。您可能希望在 IP 端节点上配置更小的 MTU (例如 1500)，如下一章节所述。

如何在终端节点上禁止 PMTUD 并配置更小的 MTU/MSS ?

以下示例将 [Solaris 10 \(及先前版本\)](#)，[HP-UX 9.x、10.x 和 11.x](#)，[IBM AIX](#)，[Linux](#)，[Windows 95/98/ME](#)，[Windows NT 3.1/3.51](#)，[Windows NT 4.0](#) 和 [Windows 2000/XP](#) 的 IP MTU 设置为

1500，或 TCP MSS 设置为 1460。因为 TCP 分段的 IP/TCP 报头通常为 40 字节，所以将 IP MTU 值设置为 1500 和将 TCP MSS 值设置为 1460 产生的效果通常相同。

注意：如果更改了接口 MTU (路由器或端节点)，则连接到相同广播域(电线和集线器)的所有系统必须运行相同的 MTU。如果同一广播域中的二个系统没有使用相同的 MTU 值，当信息包(比小型 MTU 大，但比大型 MTU 小)从 MTU 较大的系统发送到 MTU 较小的系统时，它们的通信将遇到一些困难。

[Solaris 10 \(及更低版本 \)](#)

禁用 PMTUD：

```
$ ndd -set /dev/ip ip_path_mtu_discovery 0
```

将最大 MSS 设置为 1460：

```
$ ndd -set /dev/tcp tcp_mss_max 1460
```

来源：TCP/IP Illustrated：The Protocols，Vol. 1，附录 E，作者：W. Richard Stevens 和 Gary R. Wright。

[HP-UX 9.x、10.x 与 11.x](#)

禁用 PMTUD：

HP-UX 9.X 不支持路径 MTU 发现。

HP-UX 10.00、10.01、10.10、10.20 和 10.30 支持路径 MTU 发现。对于 TCP，该功能默认为开启 (1)；对于 UDP，该功能默认为关闭 (0)。开启/关闭可以通过 **net tune** 命令切换。

```
# net tune -s tcp_pmtu 0 # net tune -s udp_pmtu 0
```

HP-UX 11 支持 PMTU 发现，并且默认启用该功能。这可以通过 **ndd setting ip_pmtu_strategy** 命令来控制。

```
# ndd -h ip_pmtu_strategy 0
```

设置路径 MTU 发现策略：0 禁用路径 MTU 发现；1 启用策略 1；2 启用策略 2。有关详细信息，请在 HP-UX 11 系统中使用 **ndd -h** 命令。

来源：[Hewlett Packard](#)

将最大 MSS 设置为 1460：

HP-UX 10.x：

```
# lanadmin -M 1460 <NetMgmtID> /usr/sbin/lanadmin [-a] [-A station_addr] [-m] [-M mtu_size] [-R] [-s] [-S speed] NetMgmtID -M mtu_size
```

为与 NetMgmtID 对应的接口设置新的 MTU 大小。mtu_size 值必须在链路指定范围内，并且您必须有超级用户特权。

来源：HP-UX 版本 10.2 的手册页

HP-UX 11.x :

```
# ndd -set /dev/tcp tcp_mss_max 1460
```

有关详细信息，请参阅 HP-UX 11 系统上有关 `ndd` 的手册页。

[IBM AIX Unix](#)

禁用 PMTUD :

路径 MTU 发现被添加到 AIX 4.2.1 中，默认状态为关闭。从 AIX4.3.3 起，默认状态为开启。

```
# no -o tcp_pmtu_discover=0
```

来源：[IBM](#)

设置最大 MSS :

对于 AIX 4.2.1 或更新版本，只有路径 MTU 发现未启用或者路径 MTU 发现不能发现路径 MTU 时，才能使用 `tcp_mssdflt`。默认：512 字节；范围：1 到 1448。

```
# no -o tcp_mssdflt=1440
```

如果这里有几个带不同 MTU 大小的适配器，那么只有一个值可以设置。此更改是系统范围内的更改。

来源：[IBM](#)

[Linux](#)

禁用 PMTUD :

您可以通过将文件 `ip_no_pmtu_disc` 的内容更改为“0”或“1”来相应启用或禁用路径 MTU 发现。要禁用 PMTUD，请使用以下命令：

```
# echo 1 >/proc/sys/net/ipv4/ip_no_pmtu_disc
```

设置接口 MTU :

当您编辑 `ifcfg-<name>` (`<name>` 指配置文件控制的设备的名称) 文件并更改“MTU”参数时，可以修改接口的 MTU 值。例如，要修改以太网接口的配置，请修改名称为“`ifcfg-eth0`”的文件。此文件用于控制系统中的第一个网络接口卡 (NIC)

来源：[Redhat Linux 指南](#)

[Windows 95/98/ME](#)

注意： 修改 Windows 95 TCP/IP 参数涉及注册表的编辑。此操作只应由经验丰富的系统管理员进行，因为如果发生错误，可能会导致系统无法启动。在这些注册表更改完成后，请重新启动以应用更改。

禁用 PMTUD :

将以下注册表值添加到注册表项中：

```
Hkey_Local_Machine\System\CurrentControlSet\Services\VxD\MSTCP PMTUDiscovery = 0 or 1 Data Type: DWORD
```

此值用于指定 Microsoft TCP/IP 是否按照 [RFC 1191](#) 中所规定，尝试执行路径 MTU 发现。[“1”启用发现而“0”禁用发现。默认值是 1。](#)

注意：在 Windows 98 中，数据类型是字符串值。

将接口 MTU 设置为 1500：

此部分中的条目必须添加到此注册表项中，其中“n”表示特定的 TCP/IP 到网络适配器绑定。

```
Hkey_Local_Machine\System\CurrentControlSet\Services\Class\netTrans\000n MaxMTU = 16-bit integer Data Type: String
```

此注册表项用于指定能够传送给媒体驱动器的数据报 IP 的最大大小。子网访问协议(SNAP)和源路由报头(如果在媒体使用)没有包含在该值中。例如，在以太网中，MaxMTU 的默认值为 1500。使用的实际值为，利用此参数指定的值和媒体驱动器报告的大小中的较小值。默认值为媒体驱动器报告的大小。

来源：Microsoft 知识库文章 [Q158474](#)

[Windows NT 3.1/3.51](#)

注意：修改 Windows NT TCP/IP 参数涉及注册表的编辑。此操作只应由经验丰富的系统管理员进行，因为如果发生错误，可能会导致系统无法启动。在这些注册表更改完成后，请重新启动以应用更改。

禁用 PMTUD：

PMTU 发现默认为启用状态，但可以通过在注册表中添加以下值进行控制：

```
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\tcpip\parameters \EnablePMTUDiscovery PMTU Discovery: 0 or 1 (Default = 1) Data Type: DWORD
```

“1”启用发现而“0”禁用发现。如果禁用 PMTU 发现，系统会为所有非本地的目标 IP 地址使用 576 字节的 MTU。TCP MSS 为 536。

来源：Microsoft 知识库文章 [Q136970](#)

将接口 MTU 设置为 1500：

TCP/IP 的这些参数专用于个人网络适配器卡。这些参数显示在以下注册表路径下，其中“adapterID”是指特定适配器卡的 Services 子项：

```
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\adapterID\Parameters\Tcpip MTU: REG_DWORD (Number in octets) Default: 0 (That is, use the value supplied by the adapter.)
```

此值用于指定接口的 MTU 大小。可以为 TCP/IP 使用的每个接口指定一个不同的 MTU 值。系统通常通过与低层驱动程序协商来决定 MTU 值。但是，使用的低层驱动程序值可能会被覆盖。

RouterMTU REG_DWORD Number in octets Default: 0 (That is, use the value supplied by the lower interface.)

此值用于指定当目标 IP 地址处于其他子网时，需要使用的 MTU 的大小。可以为 TCP/IP 使用的每个接口指定一个不同的 RouterMTU 值。在许多实施中，RouterMTU 值被设置为 576 个八位位组。这是所有 IP 节点必须支持的最小大小。由于更新的路由器通常能够处理大于 576 八位位组的 MTU，因此此参数的默认值与 MTU 使用的值完全相同。

来源：Microsoft 知识库文章 [Q102973](#)

Windows NT 4.0

注意：修改 Windows NT TCP/IP 参数涉及注册表的编辑。此操作只应由经验丰富的系统管理员进行，因为如果发生错误，可能会导致系统无法启动。在这些注册表更改完成后，请重新启动以应用更改。

禁用 PMTUD：

PMTU 发现默认为启用状态，但可以通过在注册表中添加以下值进行控制：

```
HKEY_LOCAL_MACHINE\System\CurrentControlSet\Services\Tcpip\Parameters \EnablePMTUDiscovery PMTU  
Discovery: 0 or 1 (Default = 1) Data Type: DWORD
```

“1”启用发现而“0”禁用发现。如果禁用 PMTU 发现，系统会为所有非本地的目标 IP 地址使用 576 字节的 MTU。TCP MSS 为 536。

如果将此参数设置为 1 (True)，TCP 将尝试在通往远程主机的路径上发现最大传输单元 (MTU 或最大数据包大小)。通过路径 MTU 发现，以及将 TCP 分段大小限制为此值，TCP 能消除连接网络与不同 MTU 的路径上路由器的分段。分段会对 TCP 吞吐量和网络拥塞造成负面影响。

将接口 MTU 设置为 1500：

TCP/IP 的这些参数专用于个人网络适配器卡。这些参数显示在以下注册表路径下，其中“adapterID”是指特定适配器卡的 Services 子项：

```
HKEY_LOCAL_MACHINE\System\CurrentControlSet\Services\AdapterID\Tcpip\Parameters MTU: Set it to  
equal the required MTU size in decimal (default 1500) Data Type: DWORD
```

此参数将覆盖网络接口的默认 MTU。MTU 是在基础网络上传输的最大数据包大小（以字节为单位）。这个大小包括传输报头。一个 IP 数据报可以跨多个数据包。对于基础网络来说，如果该值大于默认值，会使用网络的默认 MTU 值进行传输。如果该值小于 68，则会使用 MTU 68 进行传输。

来源：Microsoft 知识库文章 [Q120642](#)

Windows 2000/XP

注意：修改 Windows NT TCP/IP 参数涉及注册表的编辑。此操作只应由经验丰富的系统管理员进行，因为如果发生错误，可能会导致系统无法启动。在这些注册表更改完成后，请重新启动以应用更改。

禁用 PMTUD：

PMTU 发现默认为启用状态，但可以通过在注册表中添加以下值进行控制：

HKEY_LOCAL_MACHINE\System\CurrentControlSet\Services\Tcpip\Parameters \EnablePMTUDiscovery PMTU Discovery: 0 or 1 (Default = 1) Data Type: DWORD

“1”启用发现而“0”禁用发现。如果禁用 PMTU 发现，系统会为所有非本地的目标 IP 地址使用 576 字节的 MTU。TCP MSS 为 536。

如果将此参数设置为 1 (True)，TCP 将尝试在通往远程主机的路径上发现最大传输单元 (MTU 或最大数据包大小)。通过路径 MTU 发现，以及将 TCP 分段大小限制为此值，TCP 能消除连接网络与不同 MTU 的路径上路由器的分段。分段会对 TCP 吞吐量和网络拥塞造成负面影响。

将接口 MTU 设置为 1500 :

TCP/IP 的这些参数专用于个人网络适配器卡。这些参数显示在以下注册表路径下，其中“adapter ID”是指特定适配器卡的 Services 子项：

HKEY_LOCAL_MACHINE\System\CurrentControlSet\Services\Tcpip\Parameters\ Interfaces\[Adapter ID] MTU: Set it to equal the required MTU size in decimal (default 1500) Data Type: DWORD

此参数将覆盖网络接口的默认 MTU。MTU 是在基础网络上传输的最大数据包大小 (以字节为单位)。这个大小包括传输报头。请注意，一个 IP 数据报可以跨多个数据包。对于基础网络来说，如果该值大于默认值，会使用网络的默认 MTU 值进行传输。如果该值小于 68，则会使用 MTU 68 进行传输。

来源：Microsoft 知识库文章 [Q314053](#)

相关信息

- [IP 分段和 PMTUD](#)
- [与路径 MTU 发现有关的 TCP 问题](#)
- [IP 路由协议技术支持页](#)
- [技术支持和文档 - Cisco Systems](#)