Maximum transmission unit

From Wikipedia, the free encyclopedia

In computer networking, the **maximum transmission unit** (**MTU**) is the size of the largest network layer protocol data unit that can be communicated in a single network transaction.[1]:25 Fixed MTU parameters usually appear in association with a communications interface or standard. Some systems may decide MTU at connect time. The MTU relates to, but is not identical with the maximum frame size that can be transported on the data link layer, e.g. Ethernet frame.

Larger MTU is associated with reduced overhead. Smaller values can reduce network delay. In many cases MTU is dependent on underlying network capabilities and must be or should be adjusted manually or automatically so as not to exceed these capabilities.

Contents

Applicability

MTUs apply to communications protocols and network layers. The MTU is specified in terms of bytes or octets of the largest protocol data unit that the layer can pass onwards. MTU parameters usually appear in association with a communications interface (NIC, serial port, etc.). Standards (Ethernet, for example) can fix the size of an MTU; or systems (such as point-to-point serial links) may decide MTU at connect time.

Underlying data link and physical layers usually add overhead to the network layer data to be transported, so for a given maximum frame size of a medium one needs to subtract the amount of overhead to calculate that medium's MTU. For e.g. Ethernet, the maximum frame size is 1518 bytes, 18 bytes of which are overhead (header and FCS), resulting in an MTU of 1500 byte.

Tradeoffs

A larger MTU brings greater efficiency because each network packet carries more user data while protocol overheads, such as headers or underlying per-packet delays, remain fixed; the resulting higher efficiency means an improvement in bulk protocol throughput. A larger MTU also means processing of fewer packets for the same amount of data. In some systems, per-packet-processing can be a critical performance limitation.

However, this gain is not without a downside. Large packets occupy a slow link for more time than a smaller packet, causing greater delays to subsequent packets, and increasing lag and minimum latency. For example, a 1500-byte packet, the largest allowed by Ethernet at the network layer (and hence over most of the Internet), ties up a 14.4k modem for about one second. Large packets are also problematic in the presence of communications errors. If no forward error correction is used, corruption of a single bit in a packet requires that the entire packet be retransmitted, which can be very costly. At a given bit error rate, larger packets are more likely to be corrupt. Their greater payload makes retransmissions of larger packets take longer. Despite the negative effects on retransmission duration, large packets can still have a net positive effect on end-to-end TCP performance.[2]

MTUs for common media

The MTUs in this section are given as the maximum size of an IP packet that can be transmitted without fragmentation over a given medium – including IP headers but excluding headers from lower levels in the protocol stack. The MTU must not be confused with the minimum datagram size that all hosts must be prepared to accept, which has a value of 576 bytes for IPv4[3] and of 1280 bytes for IPv6.[4] It must also not be confused with the maximum size for the physically transmitted *frame*. In the case of an Ethernet frame this adds an overhead of 18 bytes, or 22 bytes with an IEEE 802.1Q tag for VLAN or quality of service.

Media for IP transport	Maximum transmission unit (bytes)	Notes
Internet IPv4 path MTU	At least 68,[5] max of 64KB[6]	Practical path MTUs are generally higher. Systems may use Path MTU Discovery[7] to find the actual path MTU.
Internet IPv6 path MTU	At least 1280,[8] max of 64KB, but up to 4GB with optional jumbogram[9]	Practical path MTUs are generally higher. Systems must use Path MTU Discovery[10] to find the actual path MTU.
Ethernet v2	1500[11]	Nearly all IP over Ethernet implementations use the Ethernet V2 frame format.
Ethernet with LLC[12] and SNAP[12]	1492[13]	
Ethernet jumbo frames	1501 – 9198 or more[14]	The limit varies by vendor. For correct interoperation, the whole Ethernet network must have the same MTU.[15] Jumbo frames are usually only seen in special-purpose networks.
PPPoE over Ethernet v2	1492[16]	= Ethernet v2 MTU (1500) - PPPoE header (8)
DS-Lite over PPPoE	1452	= Ethernet v2 MTU (1500) - PPPoE header (8) - IPv6 header (40)
PPPoE over Ethernet jumbo frames	1493 – 9190 or more[17]	= Ethernet Jumbo Frame MTU (1501 - 9198) - PPPoE header (8)
WLAN (802.11)	2304	The maximum MSDU size is 2304 before encryption. WEP will add 8 bytes, WPA-TKIP 20 bytes, and WPA2-CCMP 16 bytes.
Token Ring (802.5)	4464	
FDDI	4352[7]	

IP (Internet protocol)

DARPA designed the Internet protocol suite to work over many different networking technologies, each of which may use packets of different size. While a host will know the MTU of its own interface and possibly that of its peers (from initial handshakes), it will not initially know the lowest MTU in a chain of links to other peers. Another potential problem is that higher-level protocols may create packets larger than a particular link supports.

To get around this issue, IPv4 allows fragmentation: dividing the datagram into pieces, each small enough to pass over the single link that is being fragmented for, using the MTU parameter configured for that interface. This fragmentation process takes place at the IP layer (OSI layer 3) and marks the packets it fragments as such, so that the IP layer of the destination host knows it should reassemble the packets into the original datagram. This method implies a number of possible drawbacks:

- All fragments of a packet must arrive for the packet to be considered received. If the network drops any

fragment, the entire packet is lost.
- When the size of most or all packets exceed the MTU of a particular link that has to carry those packets, almost everything has to be fragmented. In certain cases this can cause unreasonable or unnecessary overhead. For example, various tunneling situations may exceed the MTU by very little as they add just a header's worth of data. The addition is small, but each packet now has to be sent in two fragments, the second of which carries very little payload. The same amount of payload is being moved, but every intermediate router has to do double the work in terms of header parsing and routing decisions.
- As it is normal to maximize the payload in every fragment, in general as well as when fragmenting, any further fragmentation that turns out to be necessary will increase the overhead even more.
- There is no simple method to discover the MTU of links beyond a node's direct peers.

The Internet Protocol requires that hosts must be able to process IP datagrams of at least 576 bytes (for IPv4) or 1280 bytes (for IPv6). However, this does not preclude Data Link Layers with an MTU smaller than IP's minimum MTU from conveying IP data. For example, according to IPv6's specification, if a particular Data Link Layer physically cannot deliver an IP datagram of 1280 bytes in a single frame, then the link layer must provide its own fragmentation and reassembly mechanism, separate from IP's own fragmentation mechanism, to ensure that a 1280-byte IP datagram can be delivered, intact, to the IP layer.

Path MTU Discovery

The Internet Protocol defines the "Path MTU" of an Internet transmission path as the smallest MTU of any of the IP hops of the "path" between a source and destination. Put another way, the path MTU is the largest packet size that can traverse this path without suffering fragmentation.

RFC 1191 (https://tools.ietf.org/html/rfc1191) (IPv4) and RFC 1981 (https://tools.ietf.org/html/rfc1981) (IPv6) describe "Path MTU Discovery", a technique for determining the path MTU between two IP hosts. It works by setting the DF (Don't Fragment) option in the IP headers of outgoing packets. Any device along the path whose MTU is smaller than the packet will drop such packets and send back an ICMP "Destination Unreachable (Datagram Too Big)" message containing its MTU. This information allows the source host to reduce its assumed path MTU appropriately. The process repeats until the MTU becomes small enough to traverse the entire path without fragmentation.

Unfortunately, increasing numbers of networks drop ICMP traffic (for example, to prevent denial-of-service attacks), which prevents path MTU discovery from working. One often detects such blocking in the cases where a connection works for low-volume data but hangs as soon as a host sends a large block of data. For example, with IRC a connecting client might see the initial messages up to and including the initial ping (sent by the server as an anti-spoofing measure), but get no response after that. This is because the large set of welcome messages are sent out in packets bigger than the real MTU. Also, in an IP network, the path from the source address to the destination address often gets modified dynamically, in response to various events (load-balancing, congestion, outages, etc.) - this could result in the path MTU changing (sometimes repeatedly) during a transmission, which may introduce further packet drops before the host finds a new reliable MTU.

Most Ethernet LANs support an MTU of 1500 bytes (modern LANs can use Jumbo frames, allowing for an MTU up to 9000 bytes); however, border protocols like PPPoE will reduce this. The difference between the MTU seen by end-nodes (e.g. 1500) and the Path MTU causes Path MTU Discovery to come into effect, with

the possible result of making some sites behind badly configured firewalls unreachable. One can possibly work around this, depending on which part of the network one controls; for example one can change the MSS (maximum segment size) in the initial packet that sets up the TCP connection at one's firewall.

RFC 4821 (https://tools.ietf.org/html/rfc4821), Packetization Layer Path MTU Discovery, describes a Path MTU Discovery technique which responds more robustly to ICMP filtering.

MTU in other contexts

Sometimes the term is used for maximum PDU sizes in communication layers other than the network layer (L3).

- Cisco Systems use 'L2 MTU' for the maximum frame size.[18]
- Dell/Force10 use 'MTU' for the maximum frame size.[19]
- Former Hewlett Packard used just 'MTU' for the maximum frame size including the optional IEEE 802.1Q tag.[20]
- Juniper Networks use 'Physical Interface MTU' (L3 MTU plus some unspecified protocol overhead), 'Logical Interface MTU' (consistent with IETF MTU) and 'Maximum MTU' (maximum configurable frame size for jumbo frames).[21]

The transmission of a packet on a physical network segment that is larger than the segment's MTU is known as **jabber**. This is almost always caused by faulty devices.[22] Network switches and repeater hubs have a built-in capability to detect when a device is jabbering.[23][24]

References

1. RFC 791 (https://tools.ietf.org/html/rfc791)
2. Murray, David; Terry Koziniec; Kevin Lee; Michael Dixon (2012). "Large MTUs and internet performance" (http://www.kevin-lee.co.uk/work/research/murray_HSPR_2012.pdf) (PDF). *13th IEEE Conference on High Performance Switching and Routing (HPSR 2012)*.
3. RFC 791 (https://tools.ietf.org/html/rfc791), p. 24 "Every internet destination must be able to receive a datagram of 576 octets either in one piece or in fragments to be reassembled."
4. RFC 2460 (https://tools.ietf.org/html/rfc2460), p. 13
5. RFC 791 (https://tools.ietf.org/html/rfc791), p. 24, "Every internet module must be able to forward a datagram of 68 octets without further fragmentation."
6. RFC 791 (https://tools.ietf.org/html/rfc791), p. 12, "Total Length is the length of the datagram, measured in octets, including internet header and data. This field allows the length of a datagram to be up to 65,535 octets."
7. RFC 1191 (https://tools.ietf.org/html/rfc1191)
8. RFC 2460 (https://tools.ietf.org/html/rfc2460)
9. RFC 2675 (https://tools.ietf.org/html/rfc2675), p. 1, "The IPv6 header [IPv6] has a 16-bit Payload Length field and, therefore, supports payloads up to 65,535 octets long. This document specifies an IPv6 hop-by-hop option, called the Jumbo Payload option, that carries a 32-bit length field in order to allow transmission of IPv6 packets with payloads between 65,536 and 4,294,967,295 octets in length. Packets with such long payloads are referred to as 'jumbograms'."

10. RFC 6145 (https://tools.ietf.org/html/rfc6145)
11. Network Working Group of the IETF, RFC 894 (https://tools.ietf.org/html/rfc894): A Standard for the Transmission of IP Datagrams over Ethernet Networks, Page 1, "The maximum length of the data field of a packet sent over an Ethernet is 1500 octets, thus the maximum length of an IP datagram sent over an Ethernet is 1500 octets.", ERRATA (http://www.rfc-editor.org/errata_search.php?rfc=894)
12. IEEE 802.2
13. IEEE 802.3
14. Scott Hogg (2013-03-06), *Jumbo Frames* (http://www.networkworld.com/community/blog/jumbo-frames), Network World, retrieved 2013-08-05, "Most network devices support a jumbo frame size of 9216 bytes."
15. Joe St Sauver (2003-02-04). "Practical Issues Associated With 9K MTUs" (https://www.stsauver.com/~joe/jumbos/jumbo-frames.pdf) (PDF). uoregon.edu. p. 19–21. Retrieved 2016-12-15.
16. RFC 2516 (https://tools.ietf.org/html/rfc2516) with the standard Ethernet MTU of 1500 bytes; extensions exist
17. RFC 4638 (https://tools.ietf.org/html/rfc4638)
18. "Configure and Verify Maximum Transmission Unit on Cisco Nexus Platforms" (http://www.cisco.com/c/en/us/support/docs/switches/nexus-9000-series-switches/118994-config-nexus-00.html#anc6). Cisco. 2016-11-29. Document ID:118994. Retrieved 2017-01-04.
19. "How to configure MTU (Maximum Transmission Unit) for Jumbo Frames on Dell Networking Force10 switches" (http://www.dell.com/Support/Article/en/us/debsdt1/HOW10713). Dell. 2016-06-02. Article ID: HOW10713. Retrieved 2017-01-06.
20. "Jumbo Frames". *HP Networking 2910al Switches Management and Configuration Guide*. Hewlett Packard. November 2011. P/N 5998-2874.
21. "SRX Series Services Gateways for the Branch Physical Interface Modules Reference: *MTU Default and Maximum Values for Physical Interface Modules*" (http://www.juniper.net/techpubs/en_US/release-independent/junos/topics/reference/specifications/mtu-values-mini-pims-srx-series-services-gateway.html). Juniper. 2014-01-03. Retrieved 2017-01-04.
22. *jabber* (http://www.thenetworkencyclopedia.com/entry/jabber/), The Network Encyclopedia, retrieved 2016-07-28
23. *show interfaces* (http://www.juniper.net/documentation/en_US/junos15.1/topics/reference/command-summary/show-interfaces-10-gigabit-ethernet.html), Juniper Networks, retrieved 2016-07-28
24. IEEE 802.3 *27.3.1.7 Receive jabber functional requirements*

Further reading

- Marc Slemko (January 18, 1998). "Path MTU Discovery and Filtering ICMP" (http://alive.znep.com/~marcs/mtu/index.html). Retrieved 2007-09-02.

External links

- Tweaking your MTU / RWin for Orange Broadband Users (http://www.orangeproblems.co.uk/kitz/)
- How to set the TCP MSS value using iptables (http://www.frozentux.net/iptables-tutorial/iptables-tutorial.html#TCPMSSTARGET)
- mturoute (http://www.elifulkerson.com/projects/mturoute.php) – a console utility for debugging mtu problems

- MSS Initiative (http://www.phildev.net/mss/)
- MTU Path (http://www.iea-software.com/products/mtupath.cfm) MTU discovery tool for IPv4 and IPv6 networks

Retrieved from "https://en.wikipedia.org/w/index.php?title=Maximum_transmission_unit&oldid=805467445"
