Czytaj lub słuchaj artykuł online

Zeskanuj kod QR, aby otworzyć artykuł na XO.PL

Źródło artykułu:

https://xo.pl/it-na-wesolo-net-mtu

SIECI KOMPUTEROWE

MTU i MSS, czyli co się dzieje, gdy ciężarówka nie mieści się pod wiaduktem

Na Autostradzie Pakietów wydarzyło się coś niezwykłego. Kierownik firmy transportowej kupił nową ciężarówkę. Tak wielką, że stojący obok autobus wyglądał jak samochód dostawczy. Na jej widok dział księgowości zemdlał, a dział marketingu natychmiast napisał: „Przełom w logistyce przyszłości”.

— Panowie! Koniec z małymi samochodami!
— Dlaczego, panie kierowniku?
— Bo jedna wielka ciężarówka przewiezie tyle co dziesięć małych!
— Ale nie wszędzie nią przejedziemy.
— Bzdura! Wszystkie drogi są połączone! ŁADUJEMY!

Załadowali. Ciężarówka minęła magazyn, dwa skrzyżowania, trzy bramki i zatrzymała się pod wiaduktem. Na znaku widniało ograniczenie wysokości.

— Panie kierowniku, ciężarówka się nie mieści.
— To niech pan pojedzie drugim pasem!
— Tam jest ten sam wiadukt.
— To może szybciej?
— PANIE KIEROWNIKU, ONA JEST ZA DUŻA!
— Próbował pan wyłączyć i włączyć silnik?
— …
— Halo?
— Niestety, słyszę pana doskonale.

🤣

MTU, czyli ile może przejechać naraz

Sieci komputerowe też mają swoje wiadukty. Na jednym odcinku można przesłać większy pakiet, na innym mniejszy. **MTU (Maximum Transmission Unit)** określa maksymalny rozmiar pakietu warstwy sieciowej, który dany odcinek może przenieść bez dzielenia go po drodze.

W typowej sieci Ethernet **MTU wynosi 1500 bajtów**. Mowa tutaj o rozmiarze pakietu IP, razem z jego nagłówkiem, ale bez nagłówka i zakończenia ramki Ethernet. Nie mieszajmy więc wielkości ciężarówki z wymiarami całej bramki.

— Mam 1501 bajtów. To tylko jeden więcej!
— Limit wynosi 1500.
— To może pan przymknąć oko?
— Proszę pana, to jest urządzenie sieciowe, nie kontrola bagażu podręcznego.

Najważniejsze: **o przejeździe decyduje najciaśniejszy odcinek całej trasy**. Szeroka autostrada na początku nie pomoże, jeśli dwie miejscowości dalej stoi niższy wiadukt.

A może rozładować ciężarówkę? Fragmentacja IPv4

Pracownik magazynu patrzy na wiadukt i wpada na pomysł.

— A gdyby przewieźć ten sam towar w kilku mniejszych samochodach?
— W KILKU?! Przecież kupiliśmy tę wielką ciężarówkę!
— Ale mniejsze przejadą.
— To po co kupowaliśmy wielką?
— To pytanie do działu zakupów.

W **IPv4** istnieje **fragmentacja**: zbyt duży pakiet może zostać rozdzielony na mniejsze fragmenty. Każdy jedzie osobno, a odbiorca próbuje złożyć z nich oryginalny pakiet. Nie jest to jednak darmowa magia: fragmenty zwiększają narzut, a utrata jednego z nich może uniemożliwić odtworzenie całości.

Router **IPv4** może fragmentować pakiet tylko wtedy, gdy wolno mu to zrobić. Gdy w nagłówku ustawiona jest flaga **DF (Don't Fragment)**, takiego pakietu po drodze dzielić nie wolno.

— Za duży. Proszę podzielić ładunek.
— Nie mogę, mam dokument „NIE DZIELIĆ”.
— To kto go wystawił?
— Nadawca.
— TO NIECH NADAWCA PRZYŚLE MNIEJSZĄ CIĘŻARÓWKĘ!

Router odrzuca wtedy pakiet i powinien odesłać do nadawcy komunikat **ICMP Destination Unreachable — Fragmentation Needed**, najlepiej z informacją o dopuszczalnym MTU.

IPv6: proszę wrócić do magazynu

W **IPv6** zasada jest bardziej stanowcza: **routery pośrednie nie fragmentują pakietów**. Gdy pakiet jest za duży na kolejny odcinek, router go odrzuca i wysyła komunikat **ICMPv6 Packet Too Big**. Nadawca musi dostosować rozmiar następnych pakietów. Jeśli fragmentacja **IPv6** jest potrzebna, wykonuje ją urządzenie źródłowe.

— Może pan rozłożyć towar tutaj, na poboczu?
— Nie. W **IPv6** robimy to u nadawcy.
— Ale magazyn jest 800 kilometrów stąd!
— To proszę następnym razem sprawdzić trasę.
— Czy pan wie, ile kosztował transport?!
— Nie. Ja pilnuję wiaduktu.

IPv6 wymaga, aby łącza potrafiły obsłużyć MTU co najmniej 1280 bajtów. To jednak *nie* oznacza, że cała trasa ma dokładnie taki limit.

Path MTU Discovery: sprawdzamy najniższy wiadukt

Kierownik zatrudnił specjalistę od tras.

— Pierwszy wiadukt ma limit 1500, drugi 1492, trzeci 1400.
— To jaką ciężarówkę wysyłamy?
— Taką, która zmieści się pod najniższym: 1400.
— A NIE MOŻNA PODNIEŚĆ WIADUKTU?!
— Można, jeśli jest nasz.
— A jeśli należy do innego operatora?
— Możemy napisać, że nam się Internet nie mieści.

🤣

Path MTU Discovery (PMTUD) służy do poznawania maksymalnego rozmiaru pakietów, które przejdą przez całą ścieżkę. W klasycznym wariancie nadawca korzysta z informacji zwrotnych ICMP: w **IPv4** o konieczności fragmentacji, w **IPv6** o zbyt dużym pakiecie. Istnieją też metody sprawdzające to aktywnie, na poziomie transportu, np. **PLPMTUD**.

Jeśli komunikaty o zbyt dużych pakietach są bezmyślnie blokowane, powstaje **PMTU black hole**: nadawca nie dowiaduje się, dlaczego większy ruch znika.

— Panie kierowniku, **ping** działa!
— Czyli Internet działa!
— Ale strona się nie otwiera.
— JAK TO?!
— Małe samochody przejeżdżają, duże nie.
— To dlaczego kierowcy dużych nie zgłaszają problemu?
— Zgłaszają. Tylko ich wiadomości też nie dochodzą.

🤣🤣

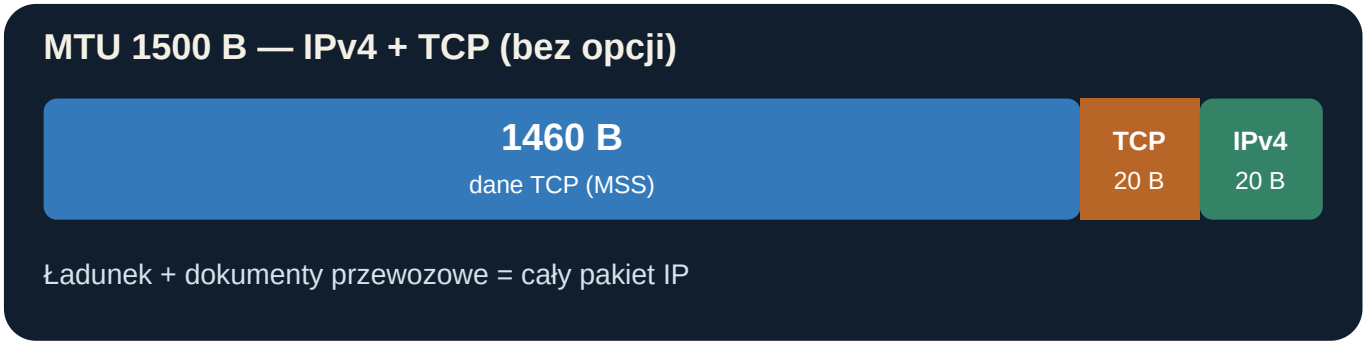
Mały **ping** nie dowodzi więc, że większe pakiety przechodzą. Przy diagnostyce sprawdza się rozmiary prób i reakcje na ograniczenia MTU, pamiętając, że nie każdy router odpowiada na ICMP tak samo.

MSS, czyli towar to nie cała ciężarówka

Po aferze z wiaduktem kierownik spojrzał na nową fakturę.

— Zaraz! Ciężarówka mieści 1500 jednostek, a wy pakujecie tylko 1460?!
— Tak.
— KTO UKRADŁ CZTERDZIEŚCI?!
— Nikt. To miejsce na dokumenty przewozowe.
— CZTERDZIEŚCI JEDNOSTEK PAPIEROLOGII?!
— Dwadzieścia na **IPv4**, dwadzieścia na **TCP**.
— Kiedyś to się przewoziło towar, a nie formularze!

Właśnie dlatego **MTU nie jest tym samym co MSS**. MTU dotyczy całego pakietu IP. **MSS (Maximum Segment Size)** określa maksymalną ilość danych **TCP** w jednym segmencie — bez nagłówków IP i **TCP**.



Dla typowego **IPv4** z MTU 1500, nagłówkiem **IPv4** 20 B i podstawowym nagłówkiem **TCP** 20 B otrzymujemy **MSS 1460 bajtów**. Przy **IPv6**, którego podstawowy nagłówek zajmuje 40 B, analogiczna wartość wynosi zwykle **1440 bajtów**. To przykłady bez dodatkowych nagłówków IP; rzeczywiste parametry zależą od połączenia.

TCP podczas ustanawiania połączenia może ogłaszać swój MSS w pakietach SYN. Każda ze stron deklaruje rozmiar danych **TCP**, jaki chce odbierać. To nie to samo, co sprawdzenie, czy każdy wiadukt na całej trasie jest odpowiednio wysoki.

— Czy **UDP** też negocjuje MSS?
— Nie. MSS to pojęcie **TCP**.
— To jak **UDP** wie, jak dużą paczkę wysłać?
— Musi uwzględniać własny sposób pakowania i ograniczenia drogi.
— Czyli znowu trzeba myśleć?
— Niestety, panie kierowniku.

MSS clamping: ostrzegamy magazyniera przed wyjazdem

Administrator postanowił, że koniec z wysyłaniem ciężarówek na ślepo. Gdy przewoźnicy **TCP** uzgadniają maksymalny ładunek, urządzenie po drodze może ograniczyć ogłaszany MSS do wartości pasującej do trasy. To **MSS clamping**.

— Od dzisiaj pakujemy najwyżej 1360 jednostek!
— Ale wcześniej mieściło się 1460!
— Wcześniej nie jechaliśmy przez tunel.
— A dlaczego tunel jest węższy?
— Bo przewozimy samochód wewnątrz drugiej ciężarówki.
— CO?!
— **VPN**.
— CHOLERA, KTO TO WYMYŚLIŁ?!
— Dział bezpieczeństwa.

🤣🤣

Tunel, **VPN** lub PPPoE dokładają własne nagłówki i mogą zmniejszyć miejsce dostępne dla pakietu wewnętrznego. Nie ma jednej uniwersalnej wartości MSS dla wszystkich tuneli. Trzeba znać rzeczywiste ograniczenia i narzut danej technologii.

MSS clamping dotyczy TCP, więc nie naprawia automatycznie dużych datagramów **UDP** ani całego ruchu IP. I nie zastępuje poprawnej obsługi ICMP, PMTUD oraz konfiguracji MTU.

Co tak naprawdę sprawdzić, gdy pod wiaduktem stoi kolejka?

Jeśli małe żądania działają, ale większe transmisje się zawieszają, a problem pojawia się szczególnie przez **VPN** lub PPPoE, warto sprawdzić: MTU na kolejnych odcinkach, narzut tuneli, komunikaty ICMP dotyczące MTU, zachowanie PMTUD oraz MSS uzgadniany przez **TCP**. Samo „**ping** odpowiada” to za mało.

Nie każdy problem z ładowaniem strony jest problemem MTU. Ale kiedy pakiety określonego rozmiaru znikają, ten podejrzany zasługuje na przesłuchanie.

— Panie kierowniku, znaleźliśmy przyczynę!
— Co było zepsute?
— Nic. Po prostu wysyłaliśmy za duże pakiety.
— A wiadukt?
— Nadal stoi.
— Czyli można było kupić mniejszą ciężarówkę?
— Można.
— To proszę nie mówić księgowości.

🤣

Dyrekcja Autostrady Pakietów przypomina: znak „maksymalna wysokość” obowiązuje również pojazdy z napisem „JESTEM BARDZO WAŻNY”. Kierownik złożył wniosek o przemianowanie MTU na „sugestię rozmiaru”. Wniosek odrzucono.