

Problem przeciążenia

Termin „przeciążenie” odnosi się do sytuacji, kiedy sumaryczne zapotrzebowanie na zasoby sieciowe przekracza aktualne możliwości sieci. Występowanie przeciążenia w sieciach z przełączaniem pakietów, jest rezultatem stosowania multipleksacji statystycznej, której celem jest maksymalizować wykorzystanie zasobów sieciowych. Przeciążenie może być także spowodowane awarią wewnątrz sieci, ale ponieważ przypadek ten występuje bardzo rzadko, został on pominięty w niniejszej pracy.

Jako przykład przyczyny występowania przeciążenia przeanalizuję pracę przełącznika ATM o N portach wejściowych (oczywiście mamy równocześnie N portów wyjściowych), do których wpływa N strumieni komórek. Jeżeli założymy, że pojemność buforów wyjściowych wynosi M to możemy stwierdzić, że w danej chwili do portu wyjściowego (określonego na podstawie pola VPI/VCI i informacji zawartej w tablicy połączeń) może być skierowanych nie więcej niż M komórek. W przypadku kiedy więcej niż M komórek jest skierowanych do danego portu wyjściowego, to część z nich musi pozostać w buforach wejściowych do czasu uzyskania dostępu do danego wyjścia przełącznika. Ponieważ bufor wejściowy i wyjściowy mają ograniczoną pojemność, to łatwo zauważyć, że nadchodzące do przełącznika komórki, które zastają pełny bufor wejściowy są tracone.

Kontrola przeciążenia w sieci jest przedmiotem wielu publikacji, albowiem obecnie stosowane w sieciach pakietowych mechanizmy są nieefektywne dla zastosowań w ATM

Istnieje kilka błędnych przekonań, mówiących że problem przeciążenia może być rozwiązany automatycznie poprzez rozwój nowej technologii i jej zastosowanie np. wymianę urządzeń sieciowych na bardziej wydajne.

Przedstawię dwa takie poglądy:

Przeciążenia spowodowane jest zbyt małą pojemnością buforów. Problem ten zostanie rozwiązany, kiedy pamięci staną się na tyle tanie, aby można było stosować bufory o bardzo dużych pojemnościach.

Niestety większa pojemność buforów nie rozwiąże problemu przeciążenia. Sieci skonstruowane z przełączników o nieskończonej pojemności buforów są podatne tak samo na przeciążenia jak sieci z przełącznikami o małych buforach. Dla tych ostatnich zbyt duży ruch spowoduje przepełnienie buforów i stratę komórek (Rysunek 3 a). W sieci z przełącznikami o nieskończonej pojemności buforów (Rysunek 3 b) kolejka i opóźnienie może się stać na tyle długie, że za nim komórki wyjdą z bufora, większość z nich jest już „time-out” i są jeszcze raz retransmitowane przez wyższe warstwy sieci, co powoduje jeszcze większe przeciążenie.

Przeciążenie spowodowane jest wolnymi łączami. Problem zostanie rozwiązany, kiedy szybkie łącza staną się ogólnie dostępne.

Stwierdzenie to nie zawsze jest prawdziwe, czasami zwiększanie przepustowości łącza może zwiększyć problem przeciążenia. Nowe szybkie łącza muszą współpracować ze starszymi i wolniejszymi łączami. Następujący eksperyment pokazuje, że wdrażanie szybkich łącz, bez odpowiedniej kontroli przeciążenia może obniżyć wydajność całej sieci.[2]. Rysunek 4 pokazuje cztery węzły połączone ze sobą szeregowo trzema łączami o przepustowości 19.2 kbit na sekundę. Czas transferu zwykłego pliku wynosił pięć minut. Zamianę łącza pomiędzy dwoma pierwszymi węzłami na łącze o przepustowości 1Mbit/s spowodowała zwiększenie czasu transmisji plik do siedmiu godzin. Z szybszym łączem dane przychodziły do pierwszego routera z większą prędkością niż przepustowość wyjścia, prowadziło to do powstawania długich kolejek, przepełnienia bufora i stratę komórek, powodując konieczność retransmisji,

co zwiększało czas transmisji.

Wymiana wszystkich łączy na szybsze także nie rozwiąże problemu przeciążenia. Przedstawiona przykładowa konfiguracja na Rysunek 5 pokazuje ten problem. Jeżeli węzły A i B zaczną nadawać do węzła C w tym samym czasie spowoduje to powstanie przeciążenia.

Przeciążenie jest problemem dynamicznym, żadne statyczne rozwiązanie nie będzie wystarczające do jego rozwiązania. Strata pakietów na skutek małej pojemności bufora, jest symptomem, nie przyczyną przeciążenia. Wzrastająca ilość szybkich sieci prowadzi do coraz większego zróżnicowania współistniejących sieci, powodując, że problem kontroli przeciążenia staje się coraz ważniejszym problemem. Odpowiednia metody zarządzania zasobami sieci ATM i kontrola przeciążenia ruchu w sieci zwiększy jej efektywność i wydajność.

Na zakończenie tego punktu omówię niektóre funkcje i procedury przedstawione przez ATM Forum, które powinny znaleźć zastosowanie w zarządzaniu zasobami sieci ATM.

a) sterowanie przyjęciem zgłoszenia (*Connection Admission Control*) –Kiedy nowe zgłoszenie przybywa do węzła sieci ATM, użytkownik deklaruje zbiór parametrów ruchowych i wymagany poziom jakości obsługi (QOS). Wykorzystując te informacje oraz znając stan sieci, blok realizujący funkcję CAC decyduje o zaakceptowaniu lub odrzuceniu nowo przybywającego zgłoszenia.

b) kontrola parametrów użytkownika (*Usage Parameter Control*), zapewnia wymuszenie zgodności parametrów zgłoszenia zadeklarowanych na etapie akceptacji z tymi, które występują w trakcie transmisji.

c) sterowanie priorytetem (*Priority Control*) Końcowy węzeł sieci ATM może generować komórki o różnym priorytecie używając bitu CLP. Urządzenia sieci mogą selektywnie odrzucać komórki o niższym priorytecie, jeżeli np. w ten sposób zapobiegają

przeciążeniu.

d) *Traffic Shaping* –kształtowanie charakterystyki przepływu informacji na podstawie danych uzyskanych z sieci lub wyżej wymienionych funkcji.

Problem przeciążenia może być częściowo rozwiązany poprzez zastosowanie wyżej wymienionych funkcji. Jednak w większości przypadkach występowania przeciążenia w sieciach ATM należy zastosować wspólnie wyżej wymienione funkcje i algorytmy kontroli przeciążenia.

Obecnie ATM Forum jest w trakcie standaryzacji algorytmów. Algorytmy, które zostały przedstawione w ATM Forum i uzyskały największe zainteresowanie zostaną omówione w dalszej części pracy.

Celem zastosowania algorytmów kontroli przeciążenia i sterowania ruchem jest zapewnienie utrzymania wynegocjowanych parametrów QOS dla wszystkich rodzajów połączeń w sieci ATM i efektywne wykorzystanie zasobów sieciowych. Algorytmom kontroli przeciążenia stawia się szereg wymagań, którymi powinny się charakteryzować. Głównymi wymaganiami są:

Skalowalność

Algorytm nie może być ograniczony przepustowością łączy, odległością, ilością przełączników lub ilością wirtualnych kanałów. Ten sam algorytm powinien być stosowany zarówno w lokalnych sieciach komputerowych (LAN) jak i w rozległych sieciach (WAN)[1].

Optymalność(sprawiedliwość)

W dzielonym środowisku przepustowość jednego źródła zależy od żądań innych źródeł. Algorytm powinien „sprawiedliwie” rozdzielić dostępne pasmo pomiędzy aktywne kanały wirtualne. Żadne z realizowanych równorzędnych połączeń nie może być faworyzowane.

Odporność

Algorytm powinien być „nieczuły” na niewielkie odchylenia parametrów. Na przykład niezgodność jednego z parametrów lub strata ramki kontrolnej nie powinna zmieniać gwałtownie stanu sieci. Algorytm powinien także izolować użytkowników i chronić ich zasoby.

Implementowalność

Algorytm nie powinien być zbyt skomplikowany, czasochłonny. Algorytm nie może dyktować szczegółowo architektury przełącznika

Algorytmy przeciążenia mogą być klasyfikowane w zależności od czasu, kiedy działają w sytuacji przeciążenia. *Wychodzenia z przeciążenia* jest mechanizmem, który próbuje zmienić już przeciążony zasób do stanu nie przeciążonego tak szybko jak jest to tylko możliwe. *Zapobieganie przeciążeniu* jest metodą dla której wystąpienie przeciążenia jest niemożliwe. Jednak metoda ta powoduje nieefektywne wykorzystanie dostępnego pasma przepustowego w sieci. *Unikanie przeciążenia* jest metodą, która stara się stale utrzymywać zasoby sieci w stanie nie przeciążonym, ale jednocześnie zapewniając efektywność wykorzystania jej zasobów. Użytkownicy sieci wymagają od sieci efektywności, wydajności, dlatego też większość proponowanych metod kontroli przeciążenia należy do metod unikających przeciążenia. Jednak, szybkie zmiany pomiędzy stanami sieci (przeciążona – nie przeciążona) są nieuniknione, dlatego też, metody wychodzenia z przeciążenia są więc zwykle integralną częścią metod kontroli przeciążenia. Jedną z metod wychodzenia z przeciążenia jest metoda odrzucania całych pakietów (*packet discard*).

Mechanizmy kontroli przepływu danych, które dostosowują prędkość transmisji zapobiegając występowaniu zatłoczenia w sieciach ATM mogą być ogólnie klasyfikowane jako: *open-loop* i *close-loop*. Kontrola typu *open-loop* jest stosowana dla

połączeń, których charakterystykę ruchu można precyzyjnie określić a wymagania jakościowe znane są z góry. Połączenie w kontroli typu *open-loop* specyfikuje swoje wymagania co do zasobów sieci, które, jeśli są dostępne, następnie są rezerwowane na czas trwania połączenia.

Kontrola typu *close-loop* jest odpowiednia dla sieci dla których zasoby nie mogą zostać zarezerwowane lub kiedy ruch w sieci nie może być sprecyzowany. W kontroli tej dostępne zasoby sieci muszą być „sprawiedliwie” i efektywnie dzielone między wieloma użytkownikami. W kontroli tej źródło stale próbuje stan sieci i dostosowuje swoją prędkość transmisji w zależności od otrzymanego sprzężenia zwrotnego. Ważną cechą kontroli typu *close-loop* jest to, że wydajność, efektywność tej metody zależy od opóźnienia sprzężenia zwrotnego. Różnica czasu pomiędzy zmianą prędkości transmisji źródła a czasem kiedy źródło zobaczy efekt tej zmiany nazwana jest *round-trip-time*. Reakcja systemów opartych na metodach ze sprzężeniem zwrotnym ograniczona jest opóźnieniem sprzężenia zwrotnego i może tylko kontrolować przeciążenia, które trwają dłużej niż *round-trip-time* w danej sieci.

Metody kontroli przeciążenia można jeszcze klasyfikować ze względu na typ kontroli jaki wprowadzają: *credit-based* lub *rate-based*. W przypadku wykorzystania kontroli typu *rate-based*, komórki sterujące zawierają informacje o bieżącym dostępnym paśmie przepustowym dla danego typu połączenia. Jeśli którykolwiek z przełączników ATM, biorący udział w realizacji danego połączenia, wykryje fakt przeciążenia tego połączenia (tzn. nastąpi odrzucenie przynajmniej jednej komórki ATM zawierającej dane użytkownika), to zawartość komórki sterującej zostanie zmieniona tak, aby zmniejszyć prędkość transmisji źródła i nie dopuścić do dalszego gubienia komórek danych. Metoda *rate-based* pozwala więc na bardzo efektywne wykorzystanie dostępnego pasma.

Odmienne natomiast działa metoda typu *credit-based*. Komórki sterujące przenoszą informacje nie o dostępnym paśmie

przepustowym, lecz o stopniu zapełnienia buforów odbiorczych przełączników oraz urządzeń brzegowych realizujących dane połączenia. Komórki te noszą nazwę kredytów. Urządzenie odbierające strumień komórek ATM na bieżąco informuje urządzenie nadawcze, które ten strumień transmituje, o stanie swojego wewnętrznego bufora wejściowego. Jeśli bufor ten zostanie wyczerpany, to urządzenie nadawcze nie otrzyma „kredytów” i w ten sposób nie będzie mogło dalej wysyłać danych.

Do efektywnego działania, obydwie metod potrzebują znać stan sieci. Typ sprzężenia jaki jest używany do badania stanu sieci jest następnym kryterium klasyfikacji metod kontroli przeciążenia. Z jawnym sprzężeniem zwrotnym (explicit feedback) elementy sieci wysyłają komunikaty wskazujące ich stan (zajętość buforów, wolne pasmo itp.) do punktów kontrolnych. Jeżeli metody kontroli przeciążenia poznają stan sieci poprzez pomiary transmisji (np. porównanie zmierzonej przepustowości i przepustowości oczekiwanej) lub poznają stan sieci przez zdarzenia takie jak time-out, informacje zawarte w potwierdzeniach, to sprzężenie zwrotne jest „ukryte” (implicit feedback). Kontrola z jawnym sprzężeniem zwrotnym pozwala na bardziej precyzyjne dostosowywanie parametrów transmisji, w odróżnieniu od kontroli z „ukrytym” sprzężeniem zwrotnym, która jest bardziej skomplikowana i czasochłonna.

Kiedy kontrola przeciążenia ma miejsce tylko w końcach systemu, jest nazwana *end-to-end control*. W kontroli tej pośrednie elementy sieci generują i wysyłają potrzebne informacje kontrolne do końca systemu. W metodach *link-by-link*, pośrednie elementy sieci wymieniają pomiędzy sobą informacje i dopasowują swoją prędkość w zależności od otrzymanej informacji (sprzężenia zwrotnego) od sąsiada. Na przykład jeżeli w przełączniku odbierającym transmisję bufor zbliża się do przepełnienia, przełącznik może wysłać do przełącznika wysyłającego żądanie zmniejszenie prędkości transmisji zapobiegając tym samym przepełnieniu bufora. W

metodzie *end-to-end* tylko końcówki systemu mogą regulować charakterystykę ruchu.

W serwisie dyplom.com.pl prezentujemy obronione prace dyplomowe, które mogą służyć za wzór do napisania własnej pracy - gdyby potrzebowali jeszcze Państwo konsultacji to polecamy stronę [pisanie prac](#) - fachowa pomoc w pisaniu prac.