

IDŹ DO

PRZYKŁADOWY ROZDZIAŁ



SPIS TREŚCI

KATALOG KSIĄŻEK

KATALOG ONLINE

ZAMÓW DRUKOWANY KATALOG

TWÓJ KOSZYK

DODAJ DO KOSZYKA

CENNIK I INFORMACJE

ZAMÓW INFORMACJE
O NOWOŚCIACH

ZAMÓW CENNIK

CZYTELNIA

FRAGMENTY KSIĄŻEK ONLINE

Wysoko wydajne sieci TCP/IP

Autorzy: Mahbub Hassan, Raj Jain

Tłumaczenie: Witold Ziolo (słowo wstępne, rozdz. 1 – 7),

Marek Pałczyński (rozdz. 8 – 13, dod. A – D)

ISBN: 83-7361-554-7

Tytuł oryginału: [High Performance TCP/IP Networking](#)

Format: B5, stron: 464



Wszystko, co niezbędne do stworzenia sieci o maksymalnej wydajności

- Techniki badania i oceny wydajności sieci
- Działanie protokołu TCP/IP w nowych środowiskach sieciowych
- Algorytmy kontroli przeciążenia sieci
- Rozwiązania zwiększające wydajność

To co do tej pory wiedziałeś o sieciach komputerowych może okazać się kroplą w morzu, gdy zechcesz stosować najnowsze, wysoko wydajne technologie. Tradycyjne sieci są coraz częściej zastępowane sieciami bezprzewodowymi, optycznymi i satelitarnymi. Standard komunikacji pomiędzy sieciami – protokół TCP/IP, jest wykorzystywany również w sieciach nowej generacji, co wywołuje nowe wyzwania i problemy związane z wydajnością przekazywania danych. Zaprojektowanie efektywnie działającej sieci wymaga poznania zagadnień mających wpływ na wydajność protokołu TCP/IP.

Książka „Wysoko wydajne sieci TCP/IP” to obszerny przewodnik. Nie koncentruje się na szczegółach działania protokołu TCP/IP. Zawiera natomiast informacje poświęcone ocenie wydajności sieci, działaniu protokołu TCP/IP w sieciach różnego typu, metodom kontroli przeciążenia ruchu w sieciach oraz nowoczesnym implementacjom protokołu TCP/IP.

- Podstawowe informacje o protokole TCP/IP
- Pomiary wydajności sieci i stosowane do tego narzędzia
- Symulowanie działania sieci TCP/IP
- Modelowanie matematyczne sieci TCP/IP
- Wydajność protokołu TCP/IP w sieciach bezprzewodowych, mobilnych, optycznych, asymetrycznych i satelitarnych
- Nowe standardy protokołu TCP/IP
- Zarządzanie kolejkami
- Przegląd implementacji programowych protokołu TCP/IP

Jeśli chcesz się przystosować lub przyczynić do zmian w technologiach sieciowych, musisz poznać metody poprawiania wydajności działania protokołu TCP/IP. W tej książce znajdziesz wszystkie niezbędne do tego wiadomości.

Wydawnictwo Helion
ul. Chopina 6
44-100 Gliwice
tel. (32)230-98-63
e-mail: helion@helion.pl



Spis treści

Współautorzy	13
Słowo wstępne	17
Rozdział 1. Wstęp	21
1.1. Historia TCP/IP	22
1.2. Aplikacje i usługi TCP	23
1.3. Powody badania wydajności TCP/IP.....	24
1.4. Jak rozumiemy wydajność TCP?.....	25
1.5. Omówienie pozostałych rozdziałów książki.....	27
1.6. Zalecana lektura	30
1.7. Podsumowanie.....	31
1.8. Pytania kontrolne.....	31
1.9. Przykład wdrożenia — prezentacja Wireless Corporation	32
Rozdział 2. Podstawy TCP/IP	35
2.1. TCP	35
2.1.1. Usługi TCP.....	36
2.1.2. Format nagłówka.....	37
2.1.3. Enkapsulacja w IP	39
2.1.4. Mechanizm potwierdzania	39
2.1.5. Mechanizm retransmisji	40
2.1.6. Ustanowienie i zakończenie połączenia	41
2.1.7. Kontrola przepływu i przesuwanie okna	43
2.1.8. Kontrola przeciążenia.....	44
2.2. UDP.....	47
2.2.1. Usługi UDP	47
2.2.2. Format nagłówka.....	48
2.2.3. Enkapsulacja w IP	49
2.3. IP	49
2.3.1. Usługi IP	49
2.3.2. Dzielenie datagramów na fragmenty i składanie fragmentów.....	50
2.3.3. Format nagłówka.....	51
2.3.4. IP wersja 6.....	52
2.4. Zalecana lektura	53
2.5. Podsumowanie.....	53
2.6. Pytania kontrolne.....	53
2.7. Przykład wdrożenia — WCORP upowszechnia TCP/IP	54

Rozdział 3. Pomiary wydajności sieci TCP/IP	57
3.1. Powody, dla których przeprowadza się pomiary sieci	58
3.2. Czynności pomiarowe	58
3.3. Klasyfikacja narzędzi pomiarowych	59
3.4. Popularne narzędzia pomiarowe i ich zastosowanie	60
3.4.1. Tcpdump	60
3.4.2. Tcpstat	69
3.4.3. ttcp	75
3.4.4. Netperf	77
3.4.5. NetPIPE	83
3.4.6. DBS	86
3.5. Wybór właściwego narzędzia	95
3.6. Zalecana literatura	96
3.7. Podsumowanie	97
3.8. Pytania kontrolne	98
3.9. Ćwiczenia praktyczne	99
3.10. Przykład wdrożenia — WCORP monitoruje ruch w sieci	99
Rozdział 4. Symulacja sieci TCP/IP	101
4.1. Rola symulacji	101
4.2. Czynności uporządkowanego badania symulacyjnego	102
4.3. Rodzaje symulacji	105
4.3.1. Zdarzenie ciągłe a zdarzenie dyskretne	105
4.3.2. Symulacje kończące się a symulacje stanu ustalonego	106
4.3.3. Symulacje syntetyczne a symulacje z wykorzystaniem zapisu pakietów	106
4.4. Potwierdzenie i sprawdzenie symulacji	108
4.5. Poziom ufności wyników symulacji	109
4.5.1. Wzór na wyliczenie poziomu ufności	109
4.5.2. Symulacja kończąca się	111
4.5.3. Symulacja stanu ustalonego	111
4.5.4. Najczęściej popełniane błędy	114
4.6. Symulacje z użyciem ruchu samopodobnego	115
4.7. Klasyfikacja narzędzi symulacyjnych	116
4.8. Symulator ns	117
4.8.1. Tworzenie modelu oraz ustalenie parametrów	118
4.8.2. Zbieranie danych	122
4.8.3. Przeprowadzenie symulacji	123
4.8.4. Prezentacja wyników	124
4.8.5. Przykłady symulacji TCP/IP wykonywanych za pomocą programu ns	124
4.9. OPNET	130
4.9.1. Tworzenie modelu	130
4.9.2. Ustalanie parametrów	135
4.9.3. Zbieranie danych	136
4.9.4. Przeprowadzenie symulacji	137
4.9.5. Prezentacja wyników	139
4.9.6. Przykłady symulacji TCP/IP wykonywanych za pomocą programu OPNET	140
4.10. Wybór odpowiedniego narzędzia	144
4.11. Zalecana literatura	147
4.12. Podsumowanie	148
4.13. Pytania kontrolne	148
4.14. Ćwiczenia praktyczne	149
4.15. Przykład wdrożenia — określenie przez WCORP potrzebnej pojemności łącza między Sydney a Melbourne za pomocą pomiarów, analiz i symulacji	149

Rozdział 5. Modelowanie TCP	153
5.1. Powody modelowania matematycznego TCP.....	153
5.2. Podstawy modelowania TCP.....	155
5.2.1. Dynamika okna	156
5.2.2. Proces utraty pakietów	156
5.3. Galeria modeli TCP.....	157
5.3.1. Model okresowy.....	157
5.3.2. Szczegółowy model utraty pakietów	159
5.3.3. Model stochastyczny z ogólnym procesem utraty pakietów.....	166
5.3.4. Model systemu sterowania	170
5.3.5. Model systemu sieciowego	172
5.4. Zalecana lektura	177
5.5. Podsumowanie.....	178
5.6. Pytania kontrolne.....	179
5.7. Ćwiczenia praktyczne.....	180
5.8. Przykład wdrożenia — identyfikacja czynników wpływających na przepustowość TCP	180
Rozdział 6. Wydajność TCP/IP w sieciach bezprzewodowych	183
6.1. Sieci bezprzewodowe	184
6.1.1. Charakterystyka ogólna.....	184
6.1.2. Bezprzewodowe sieci lokalne (WLAN).....	185
6.1.3. Komunikacyjne systemy komórkowe (CC).....	187
6.2. Problemy z wydajnością TCP w łączach bezprzewodowych	189
6.2.1. Niepotrzebne zmniejszanie wielkości okna przeciążeniowego	189
6.2.2. Zmniejszenie się szybkości połączeń w sieciach WLAN.....	189
6.2.3. Zmniejszenie się szybkości połączeń w sieciach CC	190
6.3. Poprawa wydajności TCP w łączach bezprzewodowych	193
6.3.1. Dzielenie połączeń TCP	193
6.3.2. Podsluchiwanie TCP w stacji bazowej	194
6.3.3. Informowanie o przyczynie utraty pakietu	194
6.3.4. Użycie potwierdzania selektywnego TCP	195
6.3.5. Podsumowanie i porównanie mechanizmów usprawniających	196
6.4. Ewolucja systemów bezprzewodowych a TCP/IP.....	197
6.4.1. Trendy w komunikacyjnych systemach komórkowych.....	197
6.4.2. Trendy w bezprzewodowych sieciach lokalnych.....	198
6.4.3. TCP/IP w heterogenicznych systemach bezprzewodowych	198
6.5. Zalecana lektura	199
6.6. Podsumowanie.....	200
6.7. Pytania kontrolne.....	200
6.8. Ćwiczenia praktyczne.....	201
6.9. Przykład wdrożenia — instalacja sieci bezprzewodowej przez WCORP	202
Rozdział 7. Wydajność TCP/IP w sieciach mobilnych	203
7.1. Sieci komórkowe i sieci ad hoc	203
7.2. Wydajność TCP w sieciach komórkowych	204
7.2.1. Mobile IP	204
7.2.2. Wpływ mobilności na wydajność TCP	206
7.2.3. Rozwiązania usprawniające wydajność TCP	207
7.3. Wydajność TCP w sieciach ad hoc.....	211
7.3.1. Protokół DSR	211
7.3.2. Wpływ mobilności na wydajność TCP	212
7.3.3. Rozwiązania usprawniające wydajność TCP	213

7.4. Zalecana lektura	215
7.5. Podsumowanie.....	215
7.6. Pytania kontrolne.....	216
7.7. Ćwiczenia praktyczne.....	217
7.8. Przykład wdrożenia — walka WCORP z zanikami sygnału w sieci mobilnej	217
Rozdział 8. Wydajność TCP/IP w sieciach optycznych.....	219
8.1. Rozwój sieci optycznych.....	220
8.2. Protokół IP w systemach DWDM	221
8.3. Wieloprotokołowe przełączanie etykiet.....	223
8.4. Wieloprotokołowe przełączanie lambda.....	225
8.5. Przełączanie zbitów danych.....	226
8.6. Optyczne przełączanie pakietów	228
8.6.1. Format pakietów optycznych	231
8.6.2. Przeciżenia w optycznych przełącznikach pakietów	232
8.7. Wydajność TCP/IP w sieciach optycznych	234
8.7.1. Kompleksowa wydajność sieci optycznej	235
8.7.2. Odwzorowanie połączeń TCP w pakietach optycznych	237
8.7.3. Projekt sieci optycznej dla środowiska TCP/IP	239
8.8. Zalecana lektura	242
8.9. Podsumowanie.....	242
8.10. Pytania kontrolne.....	242
8.11. Ćwiczenia praktyczne.....	243
Rozdział 9. Wydajność TCP/IP w sieciach satelitarnych	245
9.1. Krótka historia przesyłania danych za pomocą łączy satelitarnych	246
9.2. Przyczyny stosowania systemów satelitarnych.....	246
9.3. Rodzaje satelitów	247
9.4. Architektura satelitarnej sieci internet	250
9.5. Czynniki wpływające na działanie protokołu TCP w łączach satelitarnych	252
9.5.1. Długa pętla sprzężenia zwrotnego	252
9.5.2. Niedoskonałości łącza	255
9.5.3. Iloczyn szerokości pasma i opóźnienia	256
9.5.4. Asymetria pasma	258
9.5.5. Zmienne opóźnienia	258
9.5.6. Przełączanie w systemach LEO.....	258
9.5.7. Przeciżenie widmowe	258
9.5.8. Zabezpieczenia	259
9.6. Założenia systemów poprawy wydajności TCP	259
9.7. Usprawnienia protokołu TCP w sieciach satelitarnych	260
9.7.1. Wyznaczanie wartości MTU trasy	261
9.7.2. Transakcje TCP.....	262
9.7.3. Skalowanie okna	262
9.7.4. Duży rozmiar okna początkowego	264
9.7.5. Liczenie bajtów	266
9.7.6. Opóźnione potwierdzenia w trakcie powolnego startu	267
9.7.7. Wyraźne powiadamianie o przeciążeniach	268
9.7.8. Wielokrotne połączenia	268
9.7.9. Rozkładanie segmentów TCP.....	269
9.7.10. Kompresja nagłówka TCP/IP	269
9.7.11. Problemy bezpieczeństwa	271
9.7.12. Podsumowanie metod poprawy wydajności TCP	271

9.8. Zaawansowane rozwiązania oraz nowe wersje protokołu TCP	272
9.8.1. Protokół TCP szybkiego startu	273
9.8.2. HighSpeed TCP	274
9.8.3. TCP Peach	274
9.8.4. Wyraźne powiadamianie o błędach transportowych	275
9.8.5. TCP Westwood	275
9.8.6. XCP	276
9.9. Nowe protokoły transportowe sieci satelitarnych	277
9.9.1. Satelitarny protokół transportowy	277
9.9.2. Specyfikacje satelitarnych protokołów komunikacyjnych — protokół transportowy	278
9.10. Jednostki pośredniczące w poprawie wydajności	279
9.10.1. Przyczyny zastosowania systemu PEP w sieciach satelitarnych	279
9.10.2. Rodzaje jednostek PEP	280
9.10.3. Algorytmy działania jednostek PEP	283
9.10.4. Skutki stosowania jednostek PEP	284
9.10.5. Bezpieczeństwo w systemie PEP	285
9.10.6. Komercyjne systemy PEP	286
9.11. Zalecana lektura	288
9.12. Podsumowanie	288
9.13. Pytania kontrolne	289
9.14. Ćwiczenia praktyczne	291
9.15. Przykład wdrożenia — poprawa wydajności TCP w sieci satelitarnej SkyX	291
Rozdział 10. Wydajność TCP w sieciach asymetrycznych	295
10.1. Rodzaje asymetrii	296
10.1.1. Asymetria pasma	296
10.1.2. Asymetria dostępu do medium transmisyjnego	297
10.1.3. Asymetria utraty pakietów	297
10.2. Wpływ asymetrii na wydajność TCP	298
10.2.1. Asymetria pasma	298
10.2.2. Asymetria dostępu do medium transmisyjnego	300
10.3. Zwiększanie wydajności TCP w sieciach asymetrycznych	304
10.3.1. Zarządzanie pasmem łącza nadrzędnego transmisji danych	304
10.3.2. Obsługa pakietów ACK przesyłanych z mniejszą częstotliwością	307
10.4. Doświadczalna ocena poszczególnych technik poprawy wydajności	310
10.4.1. Badania asymetrii pasma	310
10.4.2. Badania asymetrii dostępu do medium transmisyjnego	311
10.5. Zalecana lektura	312
10.6. Podsumowanie	313
10.7. Pytania kontrolne	313
10.8. Ćwiczenia praktyczne	314
10.9. Przykład wdrożenia — poprawa wydajności TCP w łączach ADSL	314
Rozdział 11. Nowe standardy i odmiany TCP	317
11.1. Duplikowane potwierdzenia i szybka retransmisja	318
11.2. Szybkie odtwarzanie utraconych pakietów w TCP Reno	318
11.3. Protokół TCP NewReno	321
11.4. Protokół TCP z selektywnymi potwierdzeniami	321
11.5. Potwierdzenia generowane w przód	322
11.6. Protokół TCP Vegas	323
11.7. Przegląd innych funkcji i opcji	324
11.8. Porównanie wydajności odmian protokołu TCP	325
11.9. Zalecana lektura	334

11.10. Podsumowanie.....	334
11.11. Pytania kontrolne.....	335
11.12. Ćwiczenia praktyczne.....	336
11.13. Przykład wdrożenia — protokół TCP dla aplikacji sieciowego przetwarzania danych.....	336
Rozdział 12. Aktywne zarządzanie kolejkami w sieciach TCP/IP	339
12.1. Pasywne zarządzanie kolejkami	340
12.1.1. Odrzucanie pakietów z końca kolejki	341
12.1.2. Usuwanie pakietów z początku kolejki	341
12.1.3. Wypychanie.....	341
12.1.4. Problemy związane z pasywnym zarządzaniem kolejkami	342
12.2. Aktywne zarządzanie kolejkami.....	343
12.2.1. Wczesna losowa detekcja.....	343
12.2.2. Klasyfikacja odmian RED.....	347
12.2.3. Algorytmy RED uwzględniające sterowanie całościowe	349
12.2.4. Odmiany systemu RED z mechanizmem sterowania przepływami.....	354
12.3. Ocena wydajności i porównanie algorytmów AQM.....	359
12.3.1. Przepustowość i bezstronność	360
12.3.2. Opóźnienie i fluktuacja opóźnień	362
12.3.3. Czas odpowiedzi	363
12.3.4. Oscylacje natężenia ruchu	363
12.3.5. Podsumowanie pomiarów wydajności systemów AQM	363
12.4. Algorytmy AQM w routerach o zróżnicowanych usługach.....	364
12.5. Zalecana lektura	366
12.6. Podsumowanie.....	367
12.7. Pytania kontrolne.....	368
12.8. Ćwiczenia praktyczne.....	369
12.9. Przykład wdrożenia — system aktywnego zarządzania kolejkami w firmie WCORP	369
Rozdział 13. Implementacje programowe TCP	371
13.1. Przegląd implementacji protokołu TCP.....	372
13.1.1. Buforowanie i przenoszenie danych.....	374
13.1.2. Dostęp do pamięci użytkownika.....	376
13.1.3. Wymiana danych TCP	377
13.1.4. Retransmisje.....	381
13.1.5. Przeciążenie.....	382
13.2. Protokół TCP o wysokiej wydajności.....	383
13.2.1. Wysoka wartość iloczynu opóźnienia i szerokości pasma.....	383
13.2.2. Szacowanie czasu RTT	384
13.2.3. Wyznaczanie wartości MTU	385
13.3. Ograniczenie narzutu komunikacyjnego w stacjach końcowych.....	386
13.3.1. Narzut, wykorzystanie procesora i szerokość pasma.....	387
13.3.2. Rola przetwarzania danych w aplikacji	388
13.3.3. Źródła narzutu w komunikacji TCP/IP.....	389
13.3.4. Narzut pakietu	391
13.3.5. Przerwanie.....	392
13.3.6. Sumy kontrolne	393
13.3.7. Zarządzanie połączeniem	394
13.4. Unikanie kopiowania.....	395
13.4.1. Zmiana odwzorowania stron	396
13.4.2. Operacje wejścia-wyjścia ze scalaniem rozkazów.....	398
13.4.3. Zdalny bezpośredni dostęp do pamięci.....	399

13.5. Zmniejszenie obciążenia protokołem TCP	401
13.6. Zalecana lektura	403
13.7. Podsumowanie.....	403
13.8. Pytania kontrolne.....	404
13.9. Ćwiczenia praktyczne.....	404
Dodatek A Kolejka M/M/1.....	405
Dodatek B FreeBSD.....	409
B.1. Instalacja.....	409
B.2. Konfiguracja	410
B.2.1. Konfiguracja karty sieciowej	410
B.2.2. Uruchamianie usług sieciowych	410
B.3. Modyfikacja jądra	411
B.3.1. Modyfikacja plików konfiguracyjnych jądra	411
B.3.2. Modyfikacja plików źródłowych jądra	412
B.3.3. Kompilacja i instalacja nowego jądra	412
B.3.4. Techniki przywracania jądra.....	413
B.3.5. Przykład modyfikacji jądra.....	414
Dodatek C Samodostrajanie TCP.....	417
C.1. Przyczyny stosowania mechanizmu samodostrajania	417
C.2. Techniki i produkty związane z samodostrajaniem protokołu TCP.....	418
C.3. Wybór algorytmu samodostrajania TCP.....	419
C.4. Zalecana lektura.....	420
Bibliografia	421
Skorowidz.....	443

Rozdział 8.

Wydajność TCP/IP w sieciach optycznych

Cele rozdziału

Po przeczytaniu tego rozdziału Czytelnik:

- ◆ Będzie znał ogólne zasady działania sieci optycznych.
- ◆ Powinien znać architekturę warstwy transportowej, służącej do przenoszenia ruchu TCP/IP w sieciach optycznych.
- ◆ Zrozumie czynniki wpływające na wydajność przesyłania danych TCP/IP w sieciach optycznych.
- ◆ Będzie potrafił konfigurować optyczne przełączniki pakietów w celu uzyskania maksymalnej wydajności sieci.

Technologie optyczne są wdrażane w sieciach szkieletowych internetu, gdyż stanowią podstawę do budowania szybkich sieci internetowych nowej generacji. Wprowadzenie do architektury sieci internetowych rozwiązań optycznych stawia przed projektantami nowe wyzwania, dając jednocześnie możliwość zwiększenia wydajności sieci. Rozdział ten zawiera omówienie najważniejszych założeń sieci optycznych, ze szczególnym uwzględnieniem problemu optycznego przełączania pakietów. Została w nim zamieszczona analiza wydajności transportu TCP bazującego na optycznych przełącznikach, na podstawie której można określić wpływ technologii optycznych na efektywność przesyłania danych TCP między wybranymi punktami sieci.

8.1. Rozwój sieci optycznych

Początek komunikacji optycznej datuje się na wczesne lata siedemdziesiąte ubiegłego wieku. Powstały wówczas pierwsze systemy transmisji wykorzystujące włókna światłowodowe. Komunikacja optyczna zaczęła się rozwijać bardzo szybko, co doprowadziło do powstania niezwykle wydajnych systemów transmisyjnych, zarówno w odniesieniu do szybkości transmisji danych w poszczególnych kanałach, jak i długości łączy. Przyczyną tak szybkiego rozwoju było przede wszystkim wprowadzenie całkowicie optycznych wzmacniaczy. Jednak mimo niezwykle dużych postępów technologie optyczne nie znalazły szerokiego zastosowania w tych elementach sieci, które odpowiadają za przełączanie pakietów i zarządzanie pracą sieci.

Dopiero w kilku ostatnich latach odnotowano znaczny postęp w optycznym przełączaniu pakietów i budowaniu „przezroczystych” sieci optycznych. To z kolei pozwala na dalszy rozwój sieci szybkich transmisji danych. Potwierdzeniem może być tu fakt coraz większego zainteresowania czasopism i przeglądów technicznych szczegółowymi problemami wspomnianych koncepcji. Organizowanych jest również coraz więcej konferencji poświęconych tym zagadnieniom.

Największe zainteresowanie sieciami optycznymi jest spowodowane korzyściami wynikającymi ze stosowania spójnej transmisji optycznej, która pozwala na przeniesienie w sieciach DWDM (ang. *Dense Wavelength Division Multiplexing* — *systemy gęstego zwielokrotnienia falowego*) setek wiązek światła o różnych długościach fali w tym samym włóknie i przetwarzanie ich w pasywnych lub aktywnych elementach optycznych. Dzięki technologii DWDM przekształcenie włókna światłowodowego w wiązkę łączy o wysokiej przepustowości nie wymaga żadnego szczególnego wysiłku. Całkowita pojemność sieci i możliwość zestawiania połączeń zwiększa się wówczas o jeden lub dwa rzędy wielkości. Zintegrowane elementy optyczne sprawiają, że implementacja aktywnych komponentów sieciowych, takich jak przełączniki i multiplexery, staje się coraz bardziej realna.

Rozwój wspomnianych rozwiązań powinien doprowadzić do tzw. *przezroczystych sieci optycznych*, w których sygnały przesyłane są od początku do końca w formie optycznej. Technologie wykorzystywane w chwili pisania książki wymagały przekształcania sygnałów optycznych do postaci elektrycznej w każdym węźle sieci. Taki sposób postępowania wynika z:

- ◆ konieczności przeprowadzenia regeneracji sygnału, czyli wzmocnienia i odtworzenia kształtu impulsów sygnału oraz odtworzenia sygnału zegarowego sekwencji bitowej;
- ◆ przetwarzania i sterowania przepływem, które to czynności mają gwarantować określoną jakość transmisji danych w kanale (odpowiadają na przykład za sprawdzanie błędów);
- ◆ konieczności zmiany długości fali, wymaganej przy odbiorze danych z jednego łącza na określonej długości fali i wprowadzaniu danych do drugiego łącza na innej długości fali;

- ♦ przełączania pakietów i rozkładania obciążenia ruchu, będących podstawowymi zadaniami komutatorów, umożliwiających przenoszenie danych z jednego portu wejściowego na określony port wyjściowy przy statystycznej multipleksacji dostępnych zasobów.

Konwersja sygnału optycznego do postaci elektrycznej stanowi „wąskie gardło” systemu, ograniczając całkowitą przepustowość sieci i czyniąc jej elementy bardziej skomplikowanymi. Obecnie dzięki rozwojowi technologii optycznych wspomniane funkcje można realizować bez przekształcania sygnału. Przykładem są tutaj całkowicie optyczne regeneratory 3R, a także w pełni optyczne przełączniki i konwertery długości fali. W czasie pisania książki rozwiązania te były wciąż testowane w laboratoriach i nie można ich było po prostu kupić. Należy się jednak spodziewać, że w ciągu kilku kolejnych lat będą one implementowane w sieciach szkieletowych o bardzo dużej pojemności.

Bez wątpienia jednym z najważniejszych problemów poruszanych w różnego rodzaju debatach prowadzonych w ramach społeczności naukowej jest konieczność zagwarantowania zgodności przezroczystych sieci optycznych z wykorzystywanymi dotychczas sieciami i protokołami (z których najważniejsze są chyba sieci TCP/IP). Niniejszy rozdział zawiera omówienie tych właśnie zagadnień. Prezentuje również wyniki analiz wydajności protokołu TCP w omawianych systemach.

8.2. Protokół IP w systemach DWDM

Wprowadzenie protokołu IP do sieci DWDM wymaga zainwestowania przez operatorów telekomunikacyjnych wielu środków w rozwój zaawansowanej infrastruktury transmisyjnej oraz — co się z tym wiąże — w opracowanie planów stopniowej i bezproblemowej konwergencji obu rozwiązań. W chwili pisania książki wysoko przepustowy transport danych był oparty na innych technologiach — takich jak ATM i Frame Relay. Z tego względu, w początkowej fazie rozwoju, sieci DWDM muszą obsługiwać również protokoły inne niż IP. Ponadto większość sieci optycznych gwarantuje połączenia typu punkt-punkt dzięki wykorzystaniu w warstwie fizycznej standardu SONET/SDH. Do głównych zalet stosowania standardu SONET/SDH należy zaliczyć dokładnie zdefiniowany i opracowany mechanizm ramkowania danych, ich multipleksacji i zabezpieczania przed błędami. Dzięki temu operator dysponuje stabilną i nadającą się do zarządzania infrastrukturą sieciową.

Jedno z rozwiązań mających na celu wprowadzenie protokołu IP do sieci DWDM zakłada bezpośrednie wprowadzenie dotychczasowych protokołów do infrastruktury transmisyjnej sieci DWDM. Przykład takiej architektury został przedstawiony na rysunku 8.1. Do wad omawianej propozycji trzeba zaliczyć niższą wydajność i zwiększone koszty zarządzania siecią. W praktyce wiele funkcji zarządzania jest umieszczonych w różnych warstwach stosu protokołów, co wiąże się również ze znacznym narzutem bitowym (szczegółowa analiza systemu IP/ATM/SONET znajduje się w publikacji [158]).

Rysunek 8.1.

*Przykład pełnego stosu
protokołów w rozwiązaniu
wykorzystującym standardy
ATM i SONET
pośredniczące
w przenoszeniu ruchu IP
w sieciach DWDM*

IP/MPLS
ATM
SONET/SDH
PH

Ze względu na opisane niedogodności prowadzone są prace nad przesyłaniem danych IP bezpośrednio w sieci DWDM, bez dodatkowych warstw pośrednich, przy zachowaniu niezbędnych cech funkcjonalnych lub wręcz wyższej wydajności. Przykładem mogą tu być propozycje przenoszenia ruchu IP bezpośrednio w sieciach SONET — opisane w pracy [292] — lub wyeliminowania warstwy SONET przez zastosowanie ramkowania PPP/HDLC w łączach optycznych (tzw. otoczki cyfrowej).

Inne rozwiązanie zakłada wykorzystanie nowej warstwy optycznej umieszczonej poniżej dotychczasowych warstw protokołów i dostępnej za pomocą optycznego interfejsu użytkownika (O-UNI — ang. *Optical User Network Interface*). Zadaniem wspomnianej warstwy jest udostępnianie łączy optycznych dla połączeń generowanych przez wyższe warstwy stosu i oddzielanie zasobów DWDM od protokołów wyższych warstw. Praca komponentów tej warstwy sprowadza się do rezerwowania kanałów o określonej długości fali, przez odpowiednie konfigurowanie przełączników znajdujących się na obydwu końcach łącza. Dodatkowymi zadaniami byłyby również: zagwarantowanie poprawności działania sieci (dzięki zastosowaniu mechanizmów zabezpieczania przed awariami i automatycznej rekonfiguracji), formowanie ramek danych, monitorowanie pracy sieci oraz zarządzanie adresami niezbędnymi do zestawiania kanałów optycznych w poszczególnych węzłach sieci optycznej.

Definicja warstwy optycznej bazuje na standardzie Międzynarodowej Unii Telekomunikacyjnej (ITU — ang. *International Telecommunication Union*) [177], dzięki czemu dostawcy sprzętu mogą zachować odpowiednie standardy rozwiązań. Proponowana metoda jest korzystniejsza niż stosowanie wielowarstwowego systemu przedstawionego na rysunku 8.1, choć wymaga zdefiniowania protokołu warstwy optycznej, co wiąże się ze zwiększeniem złożoności mechanizmów zarządzania siecią i wymiany danych z innymi sieciami. Kolejna wada omawianej propozycji wynika z połączeniowej natury warstwy optycznej, która może ograniczać elastyczne współdzielenie pasma i może utrudniać wprowadzanie nowych technologii, takich jak optyczne przełączanie pakietów.

We wszystkich wspomnianych rozwiązaniach dotychczasowy protokół IP znajduje się na szczycie standardów sieciowych, związanych z poszczególnymi protokołami (zasada pozostaje niezmienna, niezależnie od tego, czy zastosuje się rozwiązanie typu SONET, czy optyczną warstwę ITU), zarządzanych w sposób niezależny.

Kolejna propozycja, uwzględniająca większy poziom scalenia poszczególnych standardów, zapewnia łatwiejszą ewolucję sieci w kierunku nowoczesnych technologii optycznych, które integrują płaszczyznę zarządzania IP z płaszczyzną zarządzania siecią optycznych. Funkcje realizowane w warstwie adaptacyjnej sieci optycznej są

przesunięte do wyższych warstw stosu (podobnie do rozwiązania proponowanego w wieloprotokołowym mechanizmie przełączania etykiet MPLS — ang. *Multiprotocol Label Switching*).

W kolejnych podrozdziałach zostaną opisane różne rozwiązania sieci optycznych. Szczegółowej analizie zostanie w nich poddana możliwość integracji z systemem MPLS. Dalsze rozważania rozpoczyna jednak krótki opis samego systemu MPLS. Jako kolejna została omówiona propozycja routingu na podstawie długości fali, czyli rozwiązanie znane jako wieloprotokołowe przełączanie lambda (MPλS — ang. *Multi-protocol Lambda Switching*) oraz optyczne przełączanie pakietów (zbitok bitów).

8.3. Wieloprotokołowe przełączanie etykiet

Mechanizm wieloprotokołowego przełączania etykiet (MPLS) jest protokołem zorientowanym na połączenie (w przeciwieństwie do protokołu IP), odpowiedzialnym za ustanawianie jednokierunkowych kompleksowych połączeń między routerami. Wyznaczane w ten sposób trasy nazywa się trasami przełączania etykiet (LSP — ang. *Label Switched Paths*). Do ich identyfikacji stosowane są dodatkowe etykiety, dołączane do datagramów IP [281]. Aby dokładniej zrozumieć zasadę działania tego mechanizmu, należy przypomnieć sobie podstawowe zadania realizowane przez dotychczas stosowane routery IP.

Routing: Routery wykorzystują różne protokoły routingu w celu wymiany informacji o sposobie zarządzania połączeniami sieciowymi, które to informacje są przechowywane w ich tablicach routingu.

Przekazywanie danych: Przekazywanie danych polega na wyznaczaniu właściwego interfejsu wyjściowego dla każdego odebranego pakietu. Do realizacji zadania niezbędne są informacje zapisane w tablicy routingu.

Przełączanie: Przełączanie polega na dostarczaniu poszczególnych pakietów do odpowiednich interfejsów wyjściowych, zgodnie z wynikiem działania procedury przekazywania danych.

Buforowanie: Buforowanie sprowadza się do rozwiązywania konfliktów w przypadku jednoczesnego odbioru kilku pakietów danych.

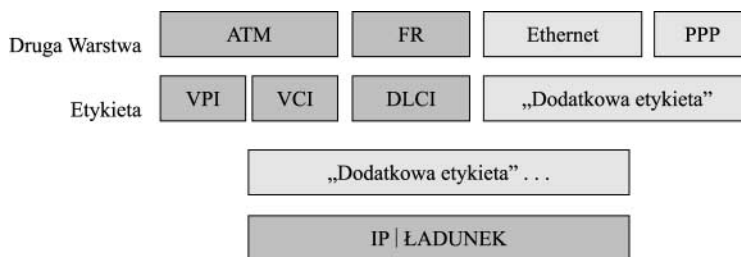
Definicja standardu MPLS ma na celu wydzielenie dwóch podstawowych funkcji warstwy sieciowej — *funkcji sterującej* i *funkcji przekazywania danych*. Element sterujący jest odpowiedzialny za routing. Jego działanie bazuje na wykorzystaniu standardowych protokołów routingu, takich jak OSPF czy BGP. Dzięki tym protokołom możliwa jest wymiana informacji z innymi routerami i utrzymanie tablicy przekazywania danych zależnej od algorytmu routingu. Tablica przekazywania danych bazuje na koncepcji klas jednakowych tras przekazywania danych (FEC — ang. *Forwarding*

Equivalence Class). Cały zbiór pakietów przekazywanych przez router jest dzielony na określoną liczbę podzbiorów stanowiących klasy FEC. Pakiety należące do jednej klasy FEC są w trakcie przekazywania przez router przetwarzane w taki sam sposób. Do jednej klasy FEC mogą należeć pakiety o różnych nagłówkach IP. Klasa FEC jest bowiem w routerze charakteryzowana przez adres kolejnego routera na drodze pakietów, który to adres jest zapisany w tablicy routingu. Za wyznaczanie klas FEC i odpowiadających im routerów kolejnego skoku jest odpowiedzialny komponent sterujący.

Element przekazywania danych odpowiada z kolei za przetwarzanie pakietów odbieranych, analizę ich nagłówków i określanie sposobu przekazywania pakietów na podstawie tablicy przekazywania danych [118]. Standard MPLS zakłada przekazywanie danych na podstawie etykiet, czyli krótkich pól o stałej długości dołączanych do pakietów i wykorzystywanych jako wartości odpowiadające informacjom zapisanym w tablicach przekazywania danych. Wspomniane pola określają klasy FEC, do których poszczególne pakiety są przypisane. Pakiety tej samej klasy FEC nie są przez mechanizm przekazywania danych w żaden sposób rozróżniane i są przekazywane z urządzenia źródłowego do urządzenia docelowego przez połączenia wyznaczone za pomocą wartości LSP. Etykieta może zostać zakodowana w nagłówku protokołu przekazywania danych — jak ma to miejsce w przypadku sieci ATM — lub dodatkowych etykiet dodawanych do pakietów IP przed sformowaniem ramki warstwy drugiej. Na rysunku 8.2 przedstawiono przykład dodatkowego nagłówka dołączonego do ramki PPP przenoszonej w sieci SONET/SDH oraz odwzorowania etykiety MPLS na wartości VPI/VCI nagłówka ATM.

Rysunek 8.2.

Przykład włączenia dodatkowej etykiety MPLS do ramki ethernetowej oraz odwzorowania etykiety MPLS na wartości VPI/VCI nagłówka ATM



Standard MPLS nie odnosi się do problemów przełączania i buforowania danych, gdyż są one zależne od konkretnych implementacji rozwiązania i nie mają wpływu na procedury routingu i przekazywania pakietów.

Stosowanie mechanizmów MPLS nie wpływa na inżynierię ruchu IP. Nic nie stoi na przeszkodzie, żeby administrator zdefiniował stałe trasy wiodące przez sieć, które mają na celu obniżenie wykorzystania zasobów sieciowych. Możliwe jest również utworzenie odrębnych tras dla różnych klas usług. Ponadto, co jest najważniejsze w niniejszym rozdziale, mechanizm przełączania etykiet ułatwia pracę elementów routera odpowiedzialnych za przekazywanie danych, co pozwala na zwiększenie szybkości przetwarzania pakietów.

8.4. Wieloprotokołowe przełączanie lambda

Rozwiązanie MPLS [52] w bardzo istotny sposób wpływa na możliwość zintegrowania mechanizmu MPLS z całkowicie optycznymi sieciami, zawierającymi routowalne fale określonej długości. Standard zakłada wykorzystanie długości fali jako jednostki pojemności sieci optycznej. Zgodnie z tym założeniem połączenia kompleksowe są udostępniane przez sieć jako zbiór fal o różnych długościach, przesyłanych w poszczególnych włóknach światłowodowych.

Mechanizm MPLS odwzorowuje trasy LSP na określone długości fal. Dzięki temu możliwe jest zaprojektowanie elementów sterujących dla przełącznic optycznych (OXC — ang. *Optical Cross-connect*) zależnych od sposobu obsługi ruchu MPLS, co z kolei pozwala na dodanie do routerów IP funkcji multipleksacji DWDM. Standard MPLS ma na celu zapewnienie mechanizmu rezerwacji w czasie rzeczywistym (zazwyczaj w ciągu kilku sekund lub nawet milisekund) kanałów sieci optycznej. Poza samą rezerwacją kanałów optycznych rozwiązanie to powinno gwarantować poprawne działanie sieci przez zabezpieczanie i odnawianie odpowiednich zasobów sieci optycznej.

Przełącznica OXC jest dla optycznej warstwy transportowej przezroczystym elementem przełączającym. Odpowiada ona za ustanawianie routowalnych kanałów optycznych przez zestawianie kanałów optycznych portów wejściowych (wejściowych włókien światłowodowych) z portami wyjściowymi (wyjściowymi włóknami światłowodowymi). Niezależnie od tego, czy przełącznica OXC realizuje swoje zadania w odniesieniu do sygnałów elektrycznych czy optycznych, zakłada się, że przełączaniu podlegają sygnały optyczne. Z tego względu funkcje przełączania odnoszą się do długości fal, a element OXC jest nazywany *przełącznikiem routującym długości fal*. Analogicznie zagadnienie ustanawiania tras optycznych jest nazywane zagadnieniem routingu i przydziału długości fal.

Definicja sieci optycznej uwzględniającej kanały ma na celu opisanie sieci kanałów optycznych jako tras między punktami dostępowymi. Sieć tego typu powinna gwarantować realizację funkcji routingu, monitorowania, konsolidacji ruchu, zabezpieczenia i przywracania kanałów optycznych. Zapewnienie funkcji operujących w ten sposób kanałami optycznymi wymaga opracowania programowalnych urządzeń typu OXC, wyposażonych w rekonfigurowalne obwody oraz inteligentne moduły sterujące, szczególnie w przypadku, gdy problem dotyczy sieci optycznych o topologii siatki.

Mechanizm MPLS tworzy kanały optyczne typu punkt-punkt między punktami dostępowymi optycznej sieci transportowej, wykorzystując do tego celu te same protokoły sygnalizacyjne, jakie są stosowane w rozwiązaniu MPLS. Urządzenie OXC jest jednostką o określonym adresie IP, zdolną do realizacji wszystkich funkcji standardu MPLS, czyli do ustalania dostępnych zasobów, rozsyłania informacji o stanie łącza, wyboru trasy na podstawie odpowiednich mechanizmów routingu oraz zarządzania tymi trasami. Do rozpowszechniania informacji o stanie łącza można wykorzystać znane od dawna protokoły IS-IS, czy OSPF, zastosowane do wyznaczania tras optycznych. Powoływanie tras optycznych można natomiast powierzyć takim protokołom sygnalizacyjnym, jak RSVP.

Jednostki OXC nie przetwarzają pakietów na poziomie analizy danych. Oznacza to, że informacje o przekazywaniu pakietów nie mogą zostać pozyskane z samego pakietu, ale muszą zostać ustalone na podstawie wykorzystywanej długości fali kanału optycznego. W sieci optycznej wyposażonej w urządzenia OXC łącze jest określone przez fizyczne włókno światłowodowe oraz długość fali świetlnej pełniącą rolę etykiety. Przełącznice OXC zgodne z założeniami MP λ S udostępniają kanały optyczne dzięki ustanawianiu relacji między długościami fal portu wejściowego a odpowiednimi długościami fal portu wyjściowego. Opisane odwzorowanie należy do zadań odpowiedniego sterownika, którego ustawienia nie mogą zostać zmienione przez funkcje analizy danych.

Istnieją jednak pewne istotne różnice między systemami MP λ S bazującymi na przełącznicach OXC a routerami sygnałów elektrycznych MPLS. W systemach MP λ S nie ma możliwości łączenia etykiet, gdyż nie jest możliwa operacja łączenia kilku długości fali. Ponadto urządzenia OXC oferują mniejszą liczbę poziomów alokacji zasobów, gdyż wykorzystana może być jedynie niewielka liczba kanałów optycznych o określonej szerokości pasma. Routery elektryczne w zasadzie nie ograniczają liczby tras, a tym samym liczby poziomów alokacji zasobów.

8.5. Przełączanie zbitek danych

Koncepcja MP λ S zakłada wykorzystanie długości fali jako jednostki pojemności sieci i umożliwia odwzorowanie wartości LSP na wartości poszczególnych długości fali. Rozwiązanie to nie zawsze jest najefektywniejszą metodą multipleksacji, gdyż strumienie danych mają naturę zbitki bitów. Przyszła sieć internet będzie wymagała szybkich mechanizmów przydzielania długości fali, umożliwiających zestawianie kompleksowych połączeń optycznych o wysokiej przepustowości. W tym celu większość elektronicznych urządzeń sieciowych musi zostać wyeliminowana, a czas rezerwowania długości fali musi zostać skrócony. Niezbędne jest więc odstępianie od typowej komutacji łączy w sieciach DWDM na rzecz efektywniejszych statystycznie multipleksowanych i ekonomiczniejszych sieci optycznych.

Optyczne przełączanie zbitek informacji [274] stanowi pierwszy krok w tym kierunku i może być uznane za rozwiązanie przejściowe na drodze do całkowicie optycznego przełączania pakietów. Celem wprowadzenia takiego mechanizmu jest zwiększenie wykorzystania fal o różnych długościach i ich współdzielenia dzięki zastosowaniu dynamicznego zarządzania długościami fal.

Koncepcja przełączania zbitek danych polega na ustanawianiu tras na określonej długości fali w czasie rzeczywistym, po wyodrębnieniu znacznej ilości danych (zbitki). Zbitka taka jest przekazywana określoną trasą, a po zakończeniu transmisji kanał zostaje usunięty. Celem jest zwiększenie współczynnika wykorzystania określonej długości fali, przy jednoczesnym ograniczeniu sterowania jedynie do zestawienia połączenia i jego zakończenia.

Pakiet sterujący poprzedza zbitkę danych o określonej wartości czasu i przenosi informacje niezbędne do przekazania samych danych. Jedynym regulowanym parametrem w tego rodzaju sieciach jest wartość czasu przerwy. W praktyce czas przerwy może mieć dodatkowe znaczenie, gdyż pozwala na wyznaczenie efektywności transmisji danych różnych klas i przydzielanie im różnych priorytetów. Zasada jest taka — im dłuższa przerwa, tym wyższy jest priorytet strumienia danych. Niezbędne zasoby sieciowe można bowiem zarezerwować znacznie wcześniej, przed przesłaniem samych informacji. Zrozumienie mechanizmu zestawienia połączenia optycznego wymaga przeanalizowania zagadnień, takich jak: algorytm wyznaczania zbitki danych optycznych, algorytm routingu i przydzielania długości fal, mechanizm rezerwacji zasobów oraz zestawianie kompleksowych połączeń.

Algorytm wyznaczania zbitki danych optycznych jest procesem decyzyjnym, mającym na celu ustalenie, czy należy zestawić trasę optyczną między dwoma punktami dla strumienia danych o określonym charakterze. Uwzględniając czas potrzebny na utworzenie połączenia oraz opóźnienie związane z przekazywaniem informacji do punktu docelowego i z powrotem, rozwiązanie znajduje zastosowanie jedynie w przypadku przesyłania strumienia danych o znacznych rozmiarach. Np. wyznaczanie trasy optycznej dla listu elektronicznego jest nieefektywne pod względem wykorzystania zasobów sieciowych. Z tego względu aplikacje wykorzystujące sieć jedynie przez krótki czas powinny korzystać z konwencjonalnych urządzeń elektronicznych. Wynika z tego, że niektóre informacje sterujące muszą zostać przekazane z warstwy aplikacji do warstwy sieciowej i (lub) do warstwy optycznej. Muszą w nich być zawarte dane o charakterystyce generowanego strumienia. W przeciwnym przypadku pełne wykorzystanie tuneli optycznych wymagałoby zastosowania niezwykle skomplikowanych algorytmów agregacji i konsolidacji ruchu.

Po podjęciu decyzji o zestawieniu połączenia niezbędne jest przeprowadzenie operacji opisanych algorytmem routingu, które mają na celu wyznaczenie optymalnej trasy związanej z właściwym przydziałem określonych długości fal na poszczególnych etapach drogi do urządzenia końcowego. Konieczne jest więc zastosowanie wielu mechanizmów przydzielania zasobów i zarządzania nimi. Wspomniane zasoby muszą zostać zarezerwowane na wszystkich odcinkach trasy, co wymaga zestawienia kompleksowego połączenia.

Ogólne założenia architektury MPLS czyni to rozwiązanie odpowiednim dla mechanizmu optycznego przełączania zbitki danych. Każdy węzeł przełączający (OBS — ang. *optical-burst switching*) może zostać wyposażony w sterownik IP/MPLS, który czyni taką jednostkę podobną do konwencjonalnego elektronicznego routera wykorzystującego metodę przełączania etykiet. Wspomniane węzły są klasyfikowane jako węzły wewnętrzne lub brzegowe. Pierwsze z wymienionych dokonują przełączania zbitki, co oznacza, że dane odbierane na danej długości fali są przenoszone wewnątrz optycznego systemu przełączającego na inną długość fali portu wyjściowego.

Nietrudno zauważyć, że długości fal dla transmisji danych nie są jednoznacznie określone. W wewnętrznych węzłach sieci nie jest realizowana konwersja optyczno-elektryczno-optyczna, dlatego informacje sterujące muszą być przekazywane na falach o ustalonej długości, gdyż mogą być przetwarzane jedynie w sposób elektroniczny. Dzięki temu możliwe jest efektywne sterowanie za pomocą warstwy sterowania

MPLS. Zatem informacje kontrolne, takie jak etykiety, muszą być przesyłane zarezerwowanymi trasami. W sieciach wykorzystujących mechanizm przełączania zbitek danych takie zasoby, jak pasmo, są rezerwowane za pomocą algorytmów jednokierunkowych. Zwiększa to stopień wykorzystania zasobów sieci, ale jednocześnie wprowadza ryzyko zablokowania zbitki danych w nieokreślonym punkcie drogi do stacji docelowej.

Zaimplementowanie mechanizmu MPLS pozwala na przydzielenie odbieranym danym etykiety, zapisywanej w pakiecie kontrolnym i określającej, jaka długość fali jest wykorzystywana do ich transmisji. Podczas tworzenia wartości LSP warstwa sterująca kojarzy zdefiniowaną wcześniej etykietę wyjściową (tj. określoną długość fali) z etykietą wejściową. Inne rozwiązanie polega na zastosowaniu algorytmu optycznego przełączania etykietowanych zbitek danych (LOBS — ang. *Labeled Optical Burst Switching*) [114]. Zamiast przydzielenia zbitki danych do określonego portu (włókna i długości fali), etykieta przenoszona w pakiecie sterującym wyznacza kolejny odcinek trasy zapisany w LSP. Oznacza to, że w wybranym włóknie wyjściowym może zostać wykorzystana dowolna długość fali, co z kolei istotnie podnosi wydajność systemu. W praktyce zablokowanie zbitki danych może wówczas wystąpić jedynie w przypadku zajęcia wszystkich długości fali w danym włóknie światłowodowym. Jednak zapewnienie właściwych algorytmów inżynierii ruchu powoduje, że taki przypadek jest mało prawdopodobny.

Trasy do urządzeń docelowych mogą być wyznaczane na podstawie bezpośredniego routingu (ang. *explicit routing*) oraz zawężonego routingu (ang. *constraint-based routing*). Mechanizmy sygnalizacji i zarządzania kompleksowymi połączeniami mogą być realizowane dzięki przekazywaniu etykiet lub wykorzystaniu protokołów rezerwacji zasobów. Węzły brzegowe są wyposażane w urządzenia elektroniczne, mogą więc wykonywać zadania charakterystyczne dla klasycznych routerów IP. Szczegółne znaczenie węzłów brzegowych wynika z faktu, że do ich zadań należy formowanie zbitek, złożonych często z kilku pakietów IP (funkcja asemblacji). W tym samym czasie urządzenia te muszą utworzyć pakiet sterujący, gwarantujący rezerwację zasobów kompleksowego połączenia, przydzielić strumieniowi danych odpowiednią etykietę oraz wyznaczyć trasę od punktu początkowego do końcowego. Pakiet sterujący może zostać sklasyfikowany zależnie od priorytetu przypisanego danym, przez co pozwala na różnicowanie rodzajów usług. Należy jednak pamiętać, że ze względu na niemożność łączenia różnych długości fal w jedną nie jest możliwe zaimplementowanie wszystkich funkcji algorytmu MPLS, stosowanych w odniesieniu do etykiet elektronicznych.

8.6. Optyczne przełączanie pakietów

Jak już wielokrotnie informowano, ostatecznym celem procesu dążenia do elastycznego zarządzania pasmem i multipleksacją strumieni danych jest uzyskanie systemu optycznego przełączania pakietów. Przełączanie pakietów prowadzi do optymalnego wykorzystania kanałów DWDM. Wynika to z dynamicznej alokacji zasobów, pozbawionej wad niskiej efektywności, charakterystycznych dla typowych połączeniowych

trybów pracy. Pakietowa transmisja danych jest z założenia bardziej elastyczna z punktu widzenia wykorzystania pasma. Pojemność łącza jest bowiem dzielona w czasie między poszczególne pakiety, dzięki czemu wszystkie łącza sieciowe są efektywniej wykorzystywane.

W pełni optyczne przełączanie pakietów było w czasie pisania książki wciąż problemem badawczym. Prace nad tym zagadnieniem trwają od ponad dekady. Celem działań jest uzyskanie przezroczystych routerów optycznych, zdolnych do przenoszenia ruchu TCP/IP. Największy problem, z jakim trzeba się zmierzyć, to ograniczenia wnoszone przez stosowaną technologię — brak pamięci optycznych — oraz sama charakterystyka urządzeń optycznych.

Routing i przekazywanie danych są z definicji funkcjami, które narzucają znaczne ograniczenia przepustowości. Powodem jest konieczność implementowania złożonych algorytmów, wymagających znacznego czasu przetwarzania danych. Obniża to oczywiście całkowitą szybkość przekazywania pakietów. W standardzie MPLS element sterujący (odpowiedzialny za routing) jest oddzielony od elementu przekazywania danych. Routery optyczne wymagają dalszego podziału elementu przekazywania danych na algorytm przekazywania danych (czyli algorytmu analizy tablicy routingu, umożliwiającego wskazanie następnego urządzenia na drodze datagramu do celu) oraz funkcję przełączającą, odpowiedzialną za fizyczne przekazanie datagramu do interfejsu wyjściowego, wybranego przez algorytm przekazywania danych. Głównym celem prac jest ograniczenie liczby konwersji elektryczno-optycznych i uzyskanie efektywniejszego interfejsu dla systemów transmisyjnych DWDM. Aby uzyskać zamierzone cele, wykorzystuje się następujące techniki.

- ♦ Etykieta pakietu lub jego nagłówek jest przekształcany z postaci optycznej do postaci elektrycznej, przez co algorytm przekazywania danych może operować sygnałami elektrycznymi.
- ♦ Ładunek datagramu jest przełączany optycznie i nie wymaga przekształcania do postaci elektrycznej.

Zatem efektywna implementacja routerów optycznych wymaga stosowania całkowicie optycznych komutatorów oraz algorytmu przekazywania danych zdolnego do pracy z prędkością charakterystyczną dla elementów optycznych. Spełnienie pierwszego założenia nie stanowi problemu, przykłady optycznych komutatorów pakietów zostały zaprezentowane w pracach [134, 135]. Wykorzystuje się w nich jako bramki optyczne półprzewodnikowe wzmacniacze optyczne (SOA — ang. *Semiconductor Optical Amplifier*) [85].

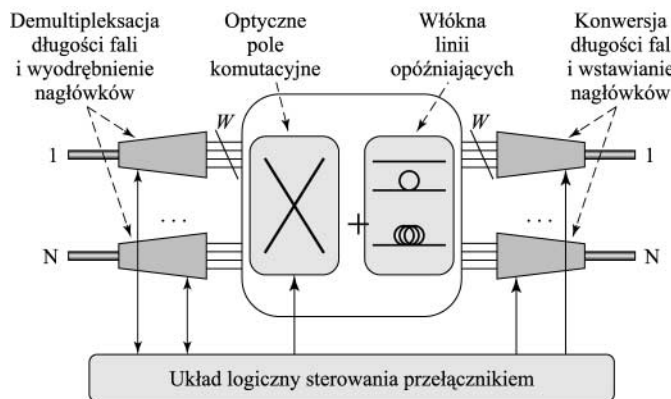
Rozwiązując drugi z wymienionych problemów, należy pamiętać, że routery wewnętrzne umieszczane w sieci szkieletowej IP są narażone na działanie przy bardzo dużym natężeniu ruchu. Z kolei przetwarzanie nagłówków IP jest procesem wymagającym wykonania setek linii kodu programu. Badania tego problemu [84] doprowadziły do stwierdzenia, że typowe przekazywanie danych IP stanowi wąskie gardło systemu, chyba że możliwe jest zagwarantowanie minimalnej długości pakietu optycznego. Z uwagi na fakt, że ładunek pakietu jest przełączany w sposób optyczny, czas przeznaczony na realizację algorytmu przekazywania danych jest zależny od średniej długości pakietu. Oznacza to, że algorytm przekazywania danych wyznacza

minimalną wartość długości pakietu. Niezachowanie jej powoduje wprowadzanie sztucznych przerw między pakietami — co obniża współczynnik wykorzystania pasma — lub prowadzi do odrzucenia niektórych pakietów ze względu na to, że router nie ma czasu na ich analizę. Franco Callegati wraz ze współpracownikami dokonał obliczeń, z których wynika, że średnia długość pakietu powinna zostać ustalona na poziomie 10 KB, co zagwarantuje zgodność z dotychczasowymi technologiami analizy nagłówków IP [84].

Rozwiązaniem problemu wąskiego gardła może być zastosowanie mechanizmu MPLS. Stosowane w nim optyczne urządzenia brzegowe (OED — ang. *Optical Edge Devices*) są odpowiedzialne za wyznaczanie trasy LSP, wprowadzanie pakietów IP do jednego lub większej liczby pakietów optycznych, wstawianie odpowiednich etykiet do nagłówków pakietów optycznych i przekazanie danych do kolejnego przełącznika MPLS. Praca przełącznika MPLS wymaga analizowania jedynie etykiet, bez konieczności przetwarzania nagłówków IP (rysunek 8.5). Wynika z tego, że mechanizm MPLS i rozwiązania optycznego przełączania pakietów mogą ze sobą doskonale współpracować. Z kolei routery optyczne łączące w sobie obydwie wymienione technologie wykorzystują najlepsze cechy przetwarzania elektrycznego i optycznego. Standardowe protokoły znalazły zastosowanie w tych komponentach routingu, które nie są silnie zależne od czasu przetwarzania. W przypadkach wymagających bardzo dużej wydajności systemu algorytm przekazywania danych bazuje na etykietach MPLS. Technologie optyczne są wykorzystywane do realizacji zadań przełączania i przesyłania danych, zapewniając wysoką szybkość transmisji danych i przepustowość.

Ogólna struktura optycznego routera pakietów DWDM została przedstawiona na rysunku 8.3. Zawiera on kilka opisanych poniżej bloków funkcjonalnych.

Rysunek 8.3.
Architektura funkcjonalna optycznego routera pakietów



- ◆ **Interfejsy wejściowe i wyjściowe** — wykorzystywane do demultipleksacji fal o różnych długościach przesyłanych we włóknie światłowodowym i przeprowadzania konwersji elektryczno-optycznych etykiet datagramów.
- ◆ **Optyczne pole komutacyjne** — zbiór pól komutacyjnych mających na celu fizyczne połączenie portów wejściowych z odpowiednimi portami wyjściowymi.

- ♦ **Bufor linii opóźniającej** — wykorzystywane do rozwiązywania ewentualnych problemów z dopasowaniem czasowym (funkcje buforowania zostały szczegółowo opisane w podrozdziale 8.6.2).
- ♦ **Sterowanie elektroniczne** — realizuje zadania przetwarzania etykiet i analizowania tabel przekazywania danych; poprzedza działanie urządzeń optycznych ustawiając odpowiednio parametry komutatora.

8.6.1. Format pakietów optycznych

Kluczowym elementem w poprawnym przełączaniu pakietów optycznych jest zachowanie odpowiedniego ich formatu. Format ten powinien uwzględniać wymagania technologii optycznych. Przykładem może tu być właściwy odstęp czasowy (zwany pasmem ochronnym) między nagłówkiem pakietu a jego polem danych, gwarantujący odpowiednią przerwę, niezbędną do wyznaczenia trasy przez komutator [135]. Czas przełączania w urządzeniu może mieć wartość dziesiątek nanosekund, ale przy bardzo wysokich szybkościach transmisji danych nie jest to czas, który można pominąć (jedna nanosekunda odpowiada dziesięciu bitom w łączu o przepływności 10 Gb/s). Zadaniem pasm ochronnych jest zabezpieczenie układu przed utratą danych. Można sobie zatem wyobrazić sieć szkieletową jako zgrupowanie routerów optycznych oraz routerów brzegowych, budujących pakiety optyczne na podstawie datagramów IP i wyznaczających wartości LSP.

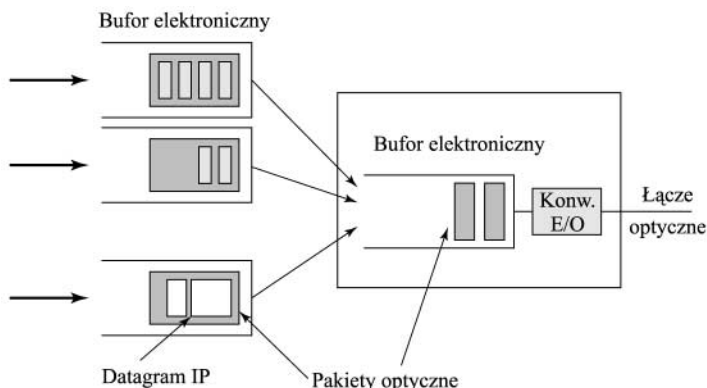
Realizację założeń rozwiązania po raz kolejny umożliwia standard MPLS. Pakiety optyczne są tworzone w routerach brzegowych przez dodanie do datagramów IP pasma ochronnego i etykiety. Jeśli zachodzi taka potrzeba, można w nich również przeprowadzać konsolidacje i (lub) segmentację pakietów [132]. Wyróżniamy dwa rodzaje formatów pakietów:

- ♦ pakiety o stałej długości przesyłane w określonych szczelinach czasowych [135];
- ♦ pakiety o zmiennej długości przesyłane w sieciach asynchronicznych [305].

Pierwsza propozycja wydaje się korzystniejsza przy założeniu implementowania optycznych komutatorów. Jej wadą jest jednak konieczność optycznej synchronizacji i nie najlepsze efekty pracy w przypadku przenoszenia informacji o niestalej długości. Przykładem może tu być format pakietu zdefiniowany w projekcie KEOPS [135], który zakładał stałą długość pakietu z 14-bajtowym nagłówkiem przesyłanym z szybkością 622 Mb/s oraz polem danych przesyłanym z szybkością od 622 Mb/s do 10 Gb/s. Druga propozycja nie wymaga synchronizacji i jest korzystniejsza w przypadku łączenia jej z sieciami TCP/IP. Niestety, przełączniki optyczne mają wówczas znacznie bardziej skomplikowaną budowę i nie cechują się taką wydajnością pracy, jaka jest charakterystyczna dla elementów poprzedniego rozwiązania.

Niezależnie od analizowanego przypadku routery brzegowe muszą być wyposażone w odpowiedni interfejs dostępowy. Przykład implementacji takiego interfejsu został przedstawiony na rysunku 8.4. Obejmuje on przypadek transmisji danych o stałej długości pakietu (np. ATM) oraz o zmiennej długości (np. datagramy IP). Ze względu na brak pamięci optycznych funkcje *generowania pakietów optycznych* i *transmisji danych*

Rysunek 8.4.
*Przykład optycznego
 urządzenia
 brzegowego (OED)*



muszą być realizowane jednocześnie. Oznacza to, że pakiet optyczny jest generowany i wysyłany w tym samym czasie. Oczywiście wszystkie składające się na dany pakiet bity muszą być wcześniej dostępne w interfejsie. Dlatego buforowanie danych jest wykonywane przez elementy elektroniczne, wchodzące w skład interfejsu. Ogólne zasady działania interfejsu są następujące.

1. Bity odbierane z łącza wejściowego muszą zostać zapisane w pamięci elektronicznej przed przekazaniem ich do układu formowania pakietu optycznego.
2. Pakiety optyczne są pobierane z buforów i wysyłane zgodnie z założeniami polityki dostępu do medium.

Algorytm formowania pakietów musi oczywiście uwzględniać problemy efektywności przesyłania, nie dopuszczając do przekroczenia maksymalnego czasu pakietyzacji optycznego pola danych. Ze względu na konieczność zachowania wspomnianego parametru niezbędne jest wyznaczenie odrębnych sposobów postępowania w przypadku pakietów o stałej i zmiennej długości. W obydwu przypadkach koncepcja jest relatywnie prosta — przed przesłaniem pakietu optycznego należy zgromadzić możliwie największą liczbę pakietów przychodzących, ale jednocześnie trzeba określić dopuszczalny czas przechowywania datagramu w buforze interfejsu.

8.6.2. Przeciążenia w optycznych przełącznikach pakietów

W tej części rozdziału zostanie omówiony problem zapobiegania przeciążeniom oraz różnicowania usług w optycznych przełącznikach pakietów. Przeciążenie może wystąpić w przypadku czasowego przeładowania danego portu wyjściowego. Rozwiązanie problemu wymaga podjęcia działań w dziedzinie czasu — kolejkovanie — oraz w dziedzinie długości fal — właściwe multipleksowanie długości fal.

Kolejkowanie. Efekt kolejkovania jest osiągany przez zastosowanie linii opóźniających (zwojów włókna), które opóźniają transmisję pakietów w taki sam sposób, jak umieszczenie ich na pewien czas w pamięci, a następnie pobranie i wysłanie. Bufor

linii opóźniającej jest podobny do standardowej kolejki, choć nie jest identyczny. Najważniejsza różnica polega na tym, że w pamięci elektronicznej pakiet może być przechowywany przez dowolnie długi czas. W buforze optycznym pakiet może być przechowywany jedynie przez pewien maksymalny, wstępnie określony czas (czas potrzebny na „przejście” przez najdłuższe włókno linii opóźniającej). Ponadto w przypadku pakietów o zmiennej długości wydajność buforowania silnie zależy od tzw. skali czasowej bufora, czyli od długości poszczególnych linii opóźniających. Szczegółowe analiza tego problemu została opisana w pracy [83]. Zdaniem Franco Callegati zapewnienie jakości transmisji danych wyrażonej przez współczynnik utraty pakietów na poziomie niższym niż 10^{-3} (akceptowalny dla aplikacji internetowych) wymaga zastosowania kilkuset linii opóźniających. Niestety, ze względu na straty w rozdzielaczach i utratę mocy bufor o tak dużej liczbie linii opóźniających nie jest możliwy do wykonania przy technologii optycznej dostępnej w chwili pisania książki.

Multipleksacja długości fali. Kolejowanie nie jest jedyną metodą unikania przeciążeń. Znaczną poprawę wydajności systemu można uzyskać dzięki połączeniu kolejowania z multipleksowaniem długości fali. Pojęcie *multipleksacja długości fali* odpowiada strategii przypisywania poszczególnych długości fal pakietom przesyłanym w łączach DWDM. Wyróżnia się dwa rozwiązania tego typu [132, 133]:

1. Obwód określonej długości fali (WC — ang. *Wavelength Circuit*).

Poszczególne trasy wiodące przez sieć, odwzorowywane za pomocą wartości LSP, są wyznaczane przez długości fali. Wynika z tego, że pakiety należące do jednego połączenia MPLS są przesyłane na jednej długości fali.

2. Pakiet o określonej długości fali (WP — ang. *Wavelength Packet*). Długości fal są wykorzystywane jako współdzielony zasób sieci. Obciążenie jest więc rozkładane na całym zbiorze długości fal, zależnie od ich dostępności. Pakiety jednego połączenia MPLS mogą być przekazywane za pomocą fal o różnych długościach.

Obwody o określonej długości fali charakteryzują się łatwiejszym zarządzaniem. Pakiety o określonej długości fali oferują w zamian niezwykłą wydajność systemu. W drugim z omawianych rozwiązań pakiet przesyłany na silnie obciążonej długości fali może zostać przeniesiony na długość fali o mniejszym obciążeniu, w tym samym włóknie światłowodowym. Zachowana zostaje wówczas jedna ścieżka sieciowa (jedno włókno), ale wykorzystywany jest mechanizm multipleksacji długości fali. Omawiane rozwiązanie dowiodło swojej efektywności w sieci bezpołączeniowej, w której inteligentne algorytmy doboru długości fal zmniejszyły prawdopodobieństwo utraty pakietu o kilka rzędów wielkości [86]. Uzyskanie zadowalającej jakości systemu wymaga zastosowania zaledwie kilkudziesięciu włókien linii opóźniających.

Niestety, w sieciach zorientowanych na połączenie, takich jak MPLS, przenoszenie pakietów należących do jednej trasy LSP na różne długości fali powoduje często nieskuteczne dostarczanie pakietów oraz uaktualnień tablicy przekazywania danych. Problem wymaga zastosowania bardziej złożonych interfejsów na obrzeżach sieci optycznej odpowiedzialnych za odtwarzanie sekwencji i wiąże się z ewentualnym przeciążeniem funkcji nadzorowania optycznych przełączników pakietów. Konieczne jest więc wyznaczenie pewnego rozwiązania pośredniego między ograniczaniem przeciążeń za pomocą zmian długości fal oraz przesyłaniem pakietów o tych samych

wartościach LSP na jednej długości fali. Badania nad multipleksacją długości fal w sieciach zorientowanych na połączenie (takich jak MPLS) zostały już rozpoczęte. Dowiodły również, że możliwe jest uzyskanie za pomocą mechanizmów multipleksowania długości fal wielokrotnie wyższej wydajności systemu — obniżenie współczynnika utraty pakietów. Do osiągnięcia prawdopodobieństwa utraty pakietu na poziomie 10^{-4} w przełączniku 4×4 o szesnastu długościami fal na włókno wystarczy osiem włókien linii opóźniających w jednym buforze. Ponadto wykorzystanie inteligentnego algorytmu doboru długości fal, który ma za zadanie ograniczenie liczby zmian długości fal w odniesieniu do całej trasy LSP, pozwala na zmniejszenie liczby nie dostarczonych pakietów do poziomu dziesięciu procent pierwotnej ich liczby.

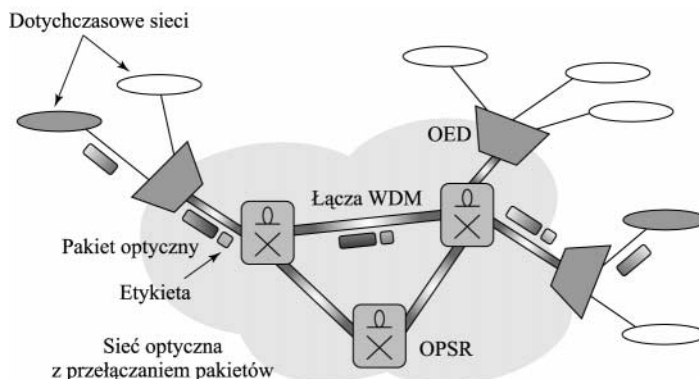
8.7. Wydajność TCP/IP w sieciach optycznych

Zgodnie z prezentowanymi wcześniej informacjami nakładkowe rozwiązanie mające na celu przekazywanie protokołu IP w sieci optycznej musi gwarantować końcowym aplikacjom TCP kompleksowy przezroczysty kanał (na który składa się jedna lub większa liczba długości fal), zarządzany za pomocą elementów optycznych, który może być rozważany na poziomie protokołu IP jako łącze typu punkt-punkt. Z kolei zastosowanie standardu MPLS w sieciach wykorzystujących routing długości fal musi zapewniać protokołowi IP możliwość wyboru tras przez sieć optyczną na podstawie wartości etykiet. Cechą charakterystyczną rozwiązania jest wyznaczenie po utworzeniu trasy stałego opóźnienia, które musi być znane we wszystkich węzłach sieci.

W obydwu przypadkach wydajność protokołu TCP jest związana z efektywnością transmisji w łączach optycznych, których parametry nie zmieniają się w trakcie trwania sesji TCP. Na działanie protokołu TCP wpływa parametr opóźnienia, który jest pewną odmianą okna TCP. W rozważanym przypadku zestawienie trasy optycznej nie powoduje powstawania przeciążeń. Zatem najważniejszym parametrem wpływającym na wydajność protokołu TCP jest wartość opóźnienia między punktami krańcowymi trasy.

Inna sytuacja ma miejsce w przypadku optycznego przełączania zbitek danych lub pakietów. W tych przypadkach, a szczególnie podczas przełączania pakietów, datagramy IP są narażone na opóźnienia i straty, które mogą mieć zmienny charakter, zależny od obciążenia sieci oraz procedur zestawiania tras. Wymagają zatem bardziej szczegółowej analizy wydajności. W końcowej części rozdziału znajduje się analiza odnosząca się do optycznego przełączania pakietów. Rozważana jest tam architektura sieci zaproponowana w publikacji [132]. Opis dotyczy optycznej sieci szkieletowej złożonej z routerów optycznych zdolnych do przetwarzania pakietów o stałej długości na podstawie etykiet (rysunek 8.5). Zakłada się, że datagramy są wprowadzane do pakietów optycznych oznaczonych zgodnie z klasami FEC datagramów. Do przekazywania danych do jednostki docelowej jest wykorzystywana sieć optyczna zawierająca przełączniki pakietów.

Rysunek 8.5.
*Przykładowa
struktura sieci
optycznej
zawierającej
przełączniki pakietów*



Z uwagi na pakietową naturę sieci wzrost obciążenia może powodować wzrost współczynnika utraty pakietów oraz opóźnienia. Wartości wymienionych parametrów zależą od sposobu działania węzłów sieci. W dalszej części rozdziału zostaną przedstawione wyniki analiz odnoszące się do wydajności połączeń między jej punktami końcowymi. Uzyskane rezultaty zostaną zestawione z analizami działania protokołu TCP.

8.7.1. Kompleksowa wydajność sieci optycznej

W sieciach DWDM zawierających przełączniki pakietów — mimo ograniczeń nakładanych na pojemności buforów — można opracować architekturę węzłów i techniki multipleksacji długości fal, które zapewnią bardzo niskie prawdopodobieństwo utraty pakietu optycznego. Uzasadnione jest więc założenie, że przeciążenie sieci optycznej występuje bardzo rzadko. W konsekwencji najważniejszym parametrem wpływającym na wydajność transmisji TCP/IP staje się kompleksowe opóźnienie.

Cechą charakterystyczną sieci optycznych jest to, że w każdym jej węźle pakiety są opóźniane o pewną wartość czasu. Dodatkowe opóźnienia występują w urządzeniach brzegowych, które są odpowiedzialne za budowę pakietu. Z architektury sieci wynika zatem, że kompleksowe opóźnienie jest przede wszystkim funkcją trzech następujących czynników:

1. Opóźnienia w interfejsach dostępowych (wejściowych i wyjściowych).
2. Opóźnienia w węzłach sieci.
3. Opóźnienia propagacyjnego w włóknach światłowodowych łączących węzły.

Opóźnienia wnoszone przez interfejsy. Interfejs brzegowy można podzielić na dwa bloki logiczne. Pierwszy z nich gromadzi datagramy IP zgodnie z wartościami FEC przypisanymi do poszczególnych pakietów optycznych. Drugi z nich odpowiada za kolejkovanie pakietów przeznaczonych do wysłania za pomocą łącza optycznego. Z tego względu opóźnienia wnoszone przez interfejsy dostępowe można podzielić na dwie grupy — opóźnienia pakietyzacji oraz opóźnienia transmisyjne. Pierwsza z nich odpowiada czasowi potrzebnemu na sformowanie pakietu optycznego po pojawieniu się datagramu w interfejsie. Drugi z parametrów jest związany z czasem kolejkovania danych w kolejce FIFO oraz z czasem przekazania danych przez łącze wyjściowe.

Opóźnienie pakietyzacji określa się jako czas między odbiorem całego datagramu (tj. odbioru ostatniego bitu datagramu) a przesłaniem datagramu do kolejki nadawczej. Podczas formowania pakietu wykorzystywane są bufory, w których datagramy IP są przechowywane do czasu sformowania całego pakietu optycznego lub upływu dopuszczalnego czasu tej operacji. Zastosowanie odpowiedniego zegara ogranicza maksymalny czas opóźnienia w interfejsach dostępowych. Zakłada się jednocześnie, że generowanie nagłówków pakietów zajmuje znacznie mniej czasu niż procedura opisana powyżej, przez co może ono zostać pominięte w analizie.

W pracy [87] został przedstawiony przykład obliczenia opóźnienia pakietyzacji. Wykorzystano w nim symulację dla ruchu przychodzącego zgodnego z rozkładem wykładniczym i rozkładem Pareto. Badania zostały przeprowadzone z wykorzystaniem klasycznych sieci 100 Mb/s oraz sieci optycznej z łączami 10 Gb/s. Symulacja i analizy odnoszą się do pakietów optycznych o czasie trwania $T = 2,8 \mu\text{s}$. Uzyskane wyniki pozwalają na wyznaczenie górnej granicy opóźnienia rejestrowanego w różnych konfiguracjach sieci.

Badania dowiodły, że istnieje duże prawdopodobieństwo, że opóźnienie pakietyzacji będzie miało wartość na poziomie kilkuset mikrosekund. Można zatem przyjąć, że maksymalny czas przerwy między pakietami optycznymi może mieć wartość około 500 μs . Przeprowadzenie pełnej analizy opóźnień transmisyjnych nie jest łatwe do przeprowadzenia. Można jednak określić pewne granice wartości tego parametru. Zgodnie z informacjami publikowanymi w pracy [87] przyjmuje się, że całkowita wartość stanowi ułamek milisekundy.

Opóźnienie wnoszone przez węzły sieci. Opóźnienia występujące w routerach wewnętrznych zostały wyznaczone dla optycznych przełączników pakietów z kolejkowaniem na wyjściu. Opóźnienia pakietów podczas przesyłania ich przez sieć optyczną wynikają z konieczności przetwarzania nagłówków, ustawiania komutatorów i opóźniania pakietów we włóknach światłowodowych linii opóźniających. W przypadku zastosowania mechanizmów przekazywania danych MPLS i szybko rekonfigurowalnych przełączników (takich jak półprzewodnikowe wzmacniacze optyczne SOA — ang. *Semiconductor Optical Amplifiers*) przetwarzanie nagłówków może zostać pominięte. Opóźnienie z nimi związane nie przekracza jednej mikrosekundy. Wartość ta jest pomijalnie mała w porównaniu z innymi parametrami, które trzeba uwzględnić w analizie. Opóźnienie kolejkowania jest funkcją natężenia ruchu i może zostać wyznaczone za pomocą klasycznych narzędzi analitycznych. Zakładając występowanie łańcucha przełączników, za pomocą aproksymacji addytywnej można określić górną granicę całkowitego opóźnienia przy przekazywaniu pakietu przez dziesiątki przełączników.

Opóźnienie propagacyjne. Opóźnienie propagacyjne nie jest zależne od architektury sieci, ale od jej fizycznej topologii, mediów transmisyjnych oraz całkowitej odległości między węzłami. Typowe wartości tego parametru są określane na poziomie od 4 do 5 mikrosekund na kilometr. Takie też współczynniki zostały przyjęte podczas omawiania odwzorowania pakietów TCP/IP w sieci szkieletowej.

Zsumowanie przybliżonych wartości kompleksowych opóźnień wyznaczonych za pomocą opisanych wcześniej reguł prowadzi do ustalenia górnej granicy kompleksowych opóźnień na poziomie milisekundy.

8.7.2. Odwzorowanie połączeń TCP w pakietach optycznych

W tej części rozdziału omówione zostanie odwzorowanie połączeń TCP w pakietach optycznych. Ze względu na bardzo wysokie wartości szybkości transmisji dostępnych w sieciach optycznych konieczne jest bardzo staranne oszacowanie iloczynu szerokości pasma i opóźnienia [213]. Opóźnienie jest wyznaczane jako czas przesłania pakietu do jednostki docelowej i z powrotem — zostało obliczone w poprzedniej części rozdziału.

Udział sieci w wartości opóźnienia sprowadza się do jednej milisekundy. Do tego trzeba dodać opóźnienie propagacji. Przy 1000-kilometrowym połączeniu opóźnienie podczas transmisji danych w jednym kierunku wynosi 5 ms (ze względu na czas propagacji) plus około 1 ms. Zatem całkowite opóźnienie przy transmisji w obydwie strony wynosi 12 ms. W łączach 2,5 Gb/s pojemność kanału danych wynosi zaledwie 30 Mb. Wartość ta jest znacznie mniejsza niż maksymalny rozmiar okna TCP. Kanał staje się zatem efektywny jedynie przy multipleksowaniu wielu połączeń TCP.

Skoro wiadomo, że niezbędne jest współdzielenie kanałów optycznych między kilka połączeń TCP, należy zdefiniować pewną politykę wprowadzania segmentów TCP do pakietów optycznych. Wynika z tego, że interfejsy dostępne pełnią niezwykle istotną rolę, gdyż odpowiadają za przekształcanie danych do postaci pakietów optycznych i wyodrębniają dane z pakietów optycznych zgodnie z obowiązującą polityką. Ustalamy na przykład, jak wiele datagramów IP będzie umieszczanych w pakiecie optycznym i które z nich zostaną wprowadzone do tego samego pakietu optycznego. Liczba segmentów TCP, które mogą zostać odwzorowane w jednym pakiecie, zależy od maksymalnego rozmiaru segmentu TCP, długości pakietu optycznego oraz szybkości transmisji danych w sieci.

Można rozważyć przykład trzech różnych szybkości transmisji danych typowych dla linii naziemnych: 2,5; 10 i 40 Gb/s. Jeśli czas trwania pakietu wynosi dwie mikrosekundy, przy łączach 2,5; 10 i 40 Gb/s składa się on z 5 Kb, 20 Kb i 80 Kb danych. Zatem zależnie od wartości maksymalnego rozmiaru segmentu (MMS — ang. *Maximum Segment Size*), w jednym pakiecie optycznym można umieścić różną liczbę segmentów. W tabeli 8.1 zestawiono liczbę segmentów, jakie mogą być umieszczone w pakiecie optycznym zależnie od szybkości transmisji danych. Dane dotyczą dwóch wartości MSS — 460 i 960 bajtów (co odpowiada datagramom IP o rozmiarze 500 i 1000 bajtów) — oraz trzech wartości czasu trwania pakietu.

Aby zdefiniować politykę formowania pakietów, należy uwzględnić mechanizm zwalczania przeciążeń sieci protokołu TCP. W sieciach optycznych przeciążenia występują niezwykle rzadko, a jak wykazują analizy, prawdopodobieństwo utraty pakietu można obniżyć do wartości 10^{-10} . W protokole TCP zostały zaimplementowane dwa mechanizmy wykrywania strat — upływanie maksymalnego czasu retransmisji i odbieranie trzech zduplikowanych potwierdzeń ACK. W przypadku wystąpienia przeciążenia uruchamiane są algorytmy szybkich retransmisji i szybkiego przywracania stanu normalnego, które pozwalają na uniknięcie negatywnych skutków wiążących się

Tabela 8.1. Maksymalna liczba segmentów, które mogą być umieszczane w jednym pakiecie optycznym

Łącze [Gb/s]	MSS [bajty]		Czas trwania [μ s]
	460	960	
2,5	1	0	2
10	5	2	2
40	20	5	2
2,5	2	1	4
10	10	5	4
40	13	20	4
2,5	4	1	6
10	16	7	6
40	65	31	6

z upływem maksymalnego czasu retransmisji. Cohen i Hamo [110] udowadniają, że aby uniknąć skutków upływu maksymalnego czasu (oznaczających rozpoczęcie procedury powolnego startu), nie mogą zaginać więcej niż trzy segmenty z danego okna. Fakt ten trzeba wziąć pod uwagę odwzorowując segmenty TCP w pakietach optycznych. Z drugiej strony, jeśli zaginie pakiet optyczny zawierający jeden lub większą liczbę datagramów z komunikatem ACK należących do różnych połączeń, nie jest to wystarczający powód do wywołania procedur retransmisji wynikających z upływu dopuszczalnego czasu. Po raz kolejny realizowana jest szybka retransmisja i szybkie przywrócenie normalnego stanu.

Oznacza to, że wyeliminowanie czynników negatywnie wpływających na przepływność danych między punktami końcowymi wymaga niedopuszczenia do utraty kilku segmentów umieszczanych w tym samym oknie. Zatem podczas budowania pakietu w interfejsie brzegowym konieczne jest ustalenie, do jakich połączeń należą poszczególne segmenty TCP, i zapobieżenie umieszczaniu kilku segmentów tego samego połączenia w jednym pakiecie optycznym. Innymi słowy, interfejs brzegowy musi działać w taki sposób, żeby segmenty jednego połączenia były oddzielone od siebie odpowiednią przerwą lub odwzorowywane w odrębnych pakietach optycznych.

Ponadto interfejsy brzegowe są wyposażane w elementy sterujące, odpowiedzialne za zarządzanie etykietami i przydzielanie etykiet zgodnie z wartościami FEC. Wynika z tego, że podczas formowania pakietów optycznych w urządzeniu brzegowym konieczne jest umieszczanie w nich segmentów o jednakowym priorytecie i adresowanych do tego samego zdalnego urządzenia brzegowego. Jeśli więc jest to możliwe, dwa następujące po sobie segmenty tego samego strumienia powinny być zapisywane w oddzielnych pakietach optycznych o tej samej wartości FEC.

Ostatnim zagadnieniem wartym szczególnej uwagi jest związanie z działaniem protokołu TCP w fazie powolnego startu (ang. *slow-start phase*). W jej trakcie rozmiar wysyłanych po sobie segmentów rośnie w sposób wykładniczy. Z kolei ze względu na mechanizm unikania przeciążeń są one wysyłane w równomiernych odstępach czasu. W przypadku tradycyjnych routerów procedura taka może prowadzić do przeciążeń, gdyż całkowita ilość danych adresowanych do łącza wyjściowego może czasowo

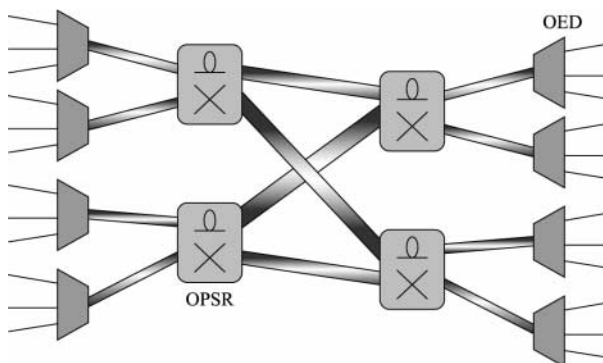
przekroczyć jego pojemność. Konieczne jest więc określenie właściwych rozmiarów buforów wyjściowych. Z drugiej strony interfejsy brzegowe współdziałają z bardzo szybkimi łączami wyjściowymi, które nigdy nie są przeciążone (przynajmniej przy obecnych warunkach pracy). Zatem opisywany problem nie dotyczy w żadnym stopniu interfejsów sieci optycznych.

8.7.3. Projekt sieci optycznej dla środowiska TCP/IP

Po zapoznaniu się z ogólnymi uwagami na temat pakietów optycznych przedstawionymi w poprzednich częściach rozdziału warto przeanalizować wyniki badań dotyczących wpływu optycznego przełączania pakietów na wydajność protokołu TCP. W dalszej części rozdziału szczególnie dokładnie zostanie przedstawiony problem zachowania właściwych relacji między parametrami formowania pakietów optycznych (takimi jak długość pakietu) oraz maksymalnym czasem retransmisji określanym w interfejsie brzegowym. Wydajność systemu jest wyrażana za pomocą wartości efektywności formowania pakietów i opóźnienia pakietyzacji oraz opisywana sposobem działania okna przeciążeniowego. Do szacowania wydajności zostało wykorzystane narzędzie symulacji *ns* opisane w rozdziale 4.

Aby możliwe było określenie zależności między efektywnością pakietyzacji, długością pakietu i czasem retransmisji, musi być znany rozkład długości pakietów. Rozkładem, który najdokładniej przybliża rzeczywiste długości pakietów, jest rozkład dwumodalny — w tym przypadku oznacza to pakiety o długości od 64 bajtów (45%) do 1536 bajtów (30%) [107]. Środowisko odniesienia uwzględniane podczas symulacji zostało przedstawione na rysunku 8.6. Cztery występujące na rysunku przełączniki optyczne zostały połączone łączami o przepustowości 2,5 Gb/s. Każdy interfejs brzegowy jest obciążony 80 agentami TCP o średnim obciążeniu 0,2 dla każdego agenta TCP. Grupy czterech agentów są przyłączone do każdego z 20 łączy o przepustowości 100 Mb/s. Zatem efektywne obciążenie łącza zewnętrznego interfejsu brzegowego wynosi 0,8 przy przepustowości 2,5 Gb/s. Wartości odniesienia dla liczników czasu wynoszą $t = 1,6510^{-6}$ s, $t' = 10t$, a $t'' = 100t$, gdzie t jest czasem odniesienia dla pakietu, stosowanym we wcześniej przeprowadzanych badaniach nad optycznymi sieciami pakietowymi [135]. Analiza wyników została zamieszczona poniżej.

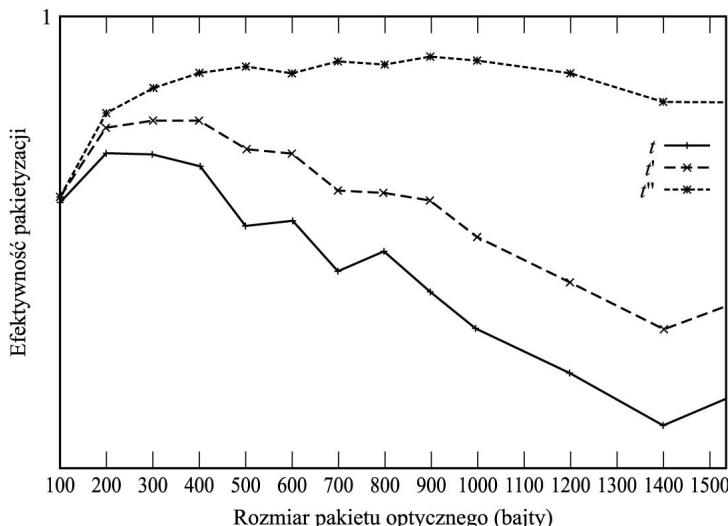
Rysunek 8.6.
Konfiguracja sieci
wykorzystana
w symulatorze *ns*



Efektywność formowania pakietów. Na rysunku 8.7 przedstawiono efektywność formowania pakietów jako funkcję rozmiaru pakietu optycznego dla trzech różnych wartości maksymalnego czasu pakietyzacji. Poza przypadkiem największej wartości dopuszczalnego czasu (który negatywnie wpływa na wartość opóźnienia) rozmiar pakietu optycznego powinien być ustalany na poziomie kilkuset bajtów. Efektywność formowania pakietów osiąga wówczas maksimum wartości. Optymalna wartość efektywności rośnie wraz ze wzrostem wartości maksymalnego czasu pakietyzacji.

Rysunek 8.7.

Efektywność formowania pakietów jako funkcja rozmiaru pakietu optycznego, dla różnych wartości maksymalnego czasu pakietyzacji: $t = 1,65 \cdot 10^{-6}$, $t' = 1,65 \cdot 10^{-5}$, $t'' = 1,65 \cdot 10^{-4}$

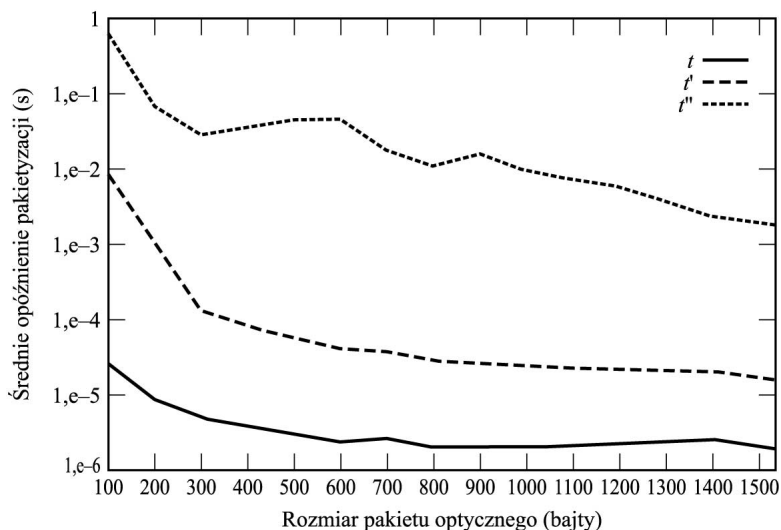


Opóźnienie pakietyzacji. Opóźnienie pakietyzacji zostało przedstawione na rysunku 8.8. Wykres sporządzono dla tych samych wartości maksymalnego czasu pakietyzacji. Można na nim zaobserwować wzrost opóźnienia odpowiadający wzrostowi efektywności formowania pakietów. Wartość opóźnienia pakietyzacji jako element składowy kompleksowego opóźnienia powinna być utrzymana na niskim poziomie, aby nie wpływała negatywnie na działania mechanizmu kontroli przeciążeń TCP. Rzeczą niezwykle istotną jest utrzymywanie otwartego okna przeciążeniowego, tak aby pasmo optyczne mogło być efektywnie wykorzystywane. Na wykresie opóźnienia wyraźnie można zaobserwować przegięcie, które powinno zostać wykorzystane podczas projektowania systemu. Wartość opóźnienia powinna przyjmować wartości znajdujące się po prawej stronie przegięcia. Oznacza to stosowanie pakietów o rozmiarze większym niż 300 bajtów.

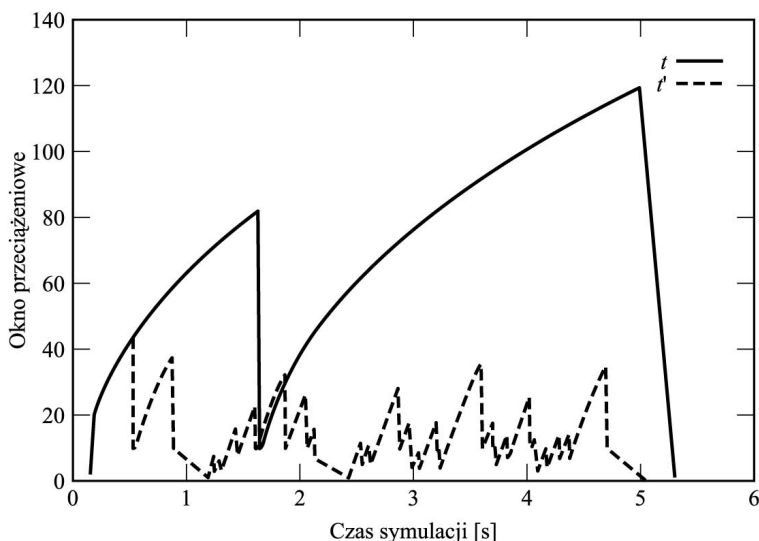
Okno przeciążeniowe. Na rysunku 8.9 zostało przedstawione zachowanie okna przeciążeniowego dla dwóch różnych wartości dopuszczalnych czasów pakietyzacji. Przyjęto, że pakiet ma rozmiar 400 bajtów, z uwzględnieniem jego nagłówka, co przy łączu 2,5 Gb/s odpowiada czasowi trwania pakietu t . W przypadku wyższych wartości czasu wartości okna przeciążeniowego są znacznie niższe. Korzystniejsze działanie protokołu można zaobserwować dla niższych wartości dopuszczalnego czasu pakietyzacji. Dynamika rozmiaru okna jest znacznie wyższa, a samo okno jest zamykane bardzo rzadko.

Rysunek 8.8.

Opóźnienie pakietyzacji jako funkcja rozmiaru pakietu optycznego, dla różnych wartości maksymalnego czasu formowania pakietów: $t = 1,65 \cdot 10^{-6}$, $t' = 1,65 \cdot 10^{-5}$, $t'' = 1,65 \cdot 10^{-4}$

**Rysunek 8.9.**

Przykład zachowania okna przeciążeniowego protokołu TCP dla pakietów o rozmiarze 400 bajtów i różnych wartości maksymalnego czasu pakietyzacji: $t = 1,65 \cdot 10^{-6}$, $t' = 1,65 \cdot 10^{-5}$



Podsumowanie wyników symulacji. Z przeprowadzonych badań można wyciągnąć dwa wnioski.

1. W przypadku przenoszenia ruchu TCP/IP rozmiar pakietu optycznego powinien być ustalany na poziomie 300 do 500 bajtów.
2. Maksymalny dopuszczalny czas formowania pakietów nie powinien osiągać bardzo dużych wartości.

8.8. Zalecana lektura

Szczegółowe omówienie systemów DWDM zostało zamieszczone w publikacji [214]. Dokładny opis systemów MPLS został zawarty w publikacji [279].

8.9. Podsumowanie

Sieci optyczne zostały opracowane z myślą o wykorzystaniu ich w sieciach szkieletowych internetu o bardzo dużej przepustowości. W rozdziale tym zostały omówione mechanizmy przenoszenia ruchu TCP/IP w sieciach optycznych. Wspomniane algorytmy odnoszą się zarówno do technik optycznego przełączania etykiet, jak i optycznego przełączania pakietów. W przypadku transmisji danych trasami wyznaczonymi za pomocą przełączania etykiet wydajność protokołu TCP nie jest przedmiotem badań, gdyż w sieciach tego typu nie występują przeciążenia — trasa o określonej długości fali jest ustalana na cały czas trwania sesji TCP. Z kolei w sieciach bazujących na optycznym przełączaniu pakietów występuje kilka problemów związanych z wydajnością, wartych poświęcenia im szczególnej uwagi. Niektóre parametry projektowe, takie jak rozmiar pakietu optycznego i dopuszczalny czas formowania pakietów, muszą podlegać optymalizacji, gdyż mogą negatywnie wpływać na wydajność transmisji TCP w sieciach optycznych.

8.10. Pytania kontrolne

1. Czym są całkowicie optyczne sieci i jakie są powody ich wdrażania?
2. Jaka jest podstawowa różnica między systemem MPLS a optycznym przełączaniem pakietów?
3. W przypadku wystąpienia przeciążenia lub przeładowania tradycyjne routery buforują pakiety w kolejkach odpowiednich portów wyjściowych. Czy buforowanie tego typu można zastosować w przełącznikach optycznych? Uzasadnij odpowiedź.
4. Z jakimi ograniczeniami należy się liczyć podczas buforowania pakietów optycznych? Wyjaśnij, w jaki sposób można uzyskać efekt buforowania w przełącznikach optycznych.
5. Na czym polega multipleksowanie długości fali? W jaki sposób może ono wspomagać zarządzanie przeciążeniami w przełącznikach optycznych?
6. Czym jest pakiet o określonej długości fali (WP)? Dlaczego rozwiązania typu WP są efektywniejsze w sieciach bezpołączeniowych niż w sieciach zorientowanych na połączenie?

7. Jakie względy wydajnościowe związane z protokołem TCP przemawiają na korzyść standardu MP λ S w porównaniu z sieciami bazującymi na routingu długości fal?
8. Od czego zależy wydajność TCP w sieciach o optycznym przełączaniu pakietów?
9. Opisz elementy składowe opóźnienia pakietu w sieciach optycznych. Dlaczego opóźnienie w urządzeniach brzegowych jest większe niż w węzłach pośrednich?
10. Co wyznacza maksymalną liczbę segmentów TCP, które podlegają pakietyzacji do postaci pojedynczego pakietu optycznego?
11. Jakie czynniki związane z przekształcaniem datagramów IP do postaci pakietów optycznych należy wziąć pod uwagę w celu zmniejszenia strat w kompleksowej przepustowości?

8.11. Ćwiczenia praktyczne

1. Powtórz symulację *ns* opisaną w podrozdziale 8.7.3 dla sieci odniesienia złożonej z ośmiu przełączników optycznych. Porównaj wyniki testu z rezultatami prezentowanymi na rysunkach 8.7, 8.8 i 8.9. Czy można zauważyć jakieś zasadnicze różnice?
2. Powtórz opisaną powyżej symulację dla innych odmian TCP (np. TCPVegas lub TCP Reno). Czy można zauważyć jakieś zasadnicze różnice? Sporządź raport opisujący uzyskane wyniki.