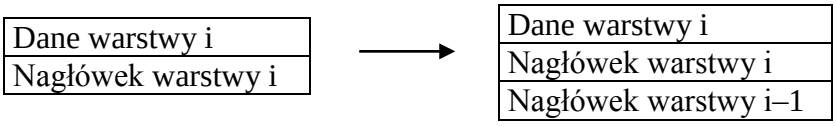
Model warstwowy komunikacji sieciowej. Budowa nagłówka ramki Ethernet II i pakietu IP.

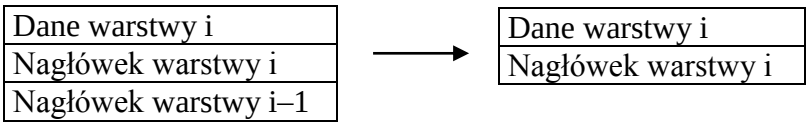
Model warstwowy OSI: (Open Systems Interconnection)

Aplikacji
Prezentacji
Sesji
Transportu
Sieci
Łącza danych
Fizyczna

Każdy protokół sieciowy można opisać przy pomocy powyższego modelu, przy czym nie zawsze do opisu protokołu używane są wszystkie warstwy. Np. protokół IP składa się tylko z warstwy sieci, natomiast nie może funkcjonować bez wsparcia dwóch warstw niższych. Ogólna zasada stanowi, że protokół umiejscowiony w warstwie n wymaga wsparcia wszystkich warstw od 1 do n–1. Poszczególne warstwy (oprócz fizycznej) realizowane są przez oprogramowanie. W trakcie wysyłania dane są przekazywane z warstw wyższych do niższych, przy czym każda warstwa dodaje swój nagłówek. W rezultacie dane przekazywane przez warstwę i do warstwy i–1 składają się z danych i nagłówka warstwy i, co przedstawia następujący rysunek:



Po dotarciu do celu, dane przekazywane są w odwrotną stronę, mianowicie od warstw niższych do wyższych, przy czym każda warstwa usuwa swój nagłówek. Jest to przedstawione na kolejnym rysunku ilustrującym przekazywanie danych z warstwy i–1 do warstwy i:



Po dotarciu do warstwy n, usuwany jest jej nagłówek i dane ulegają przetworzeniu przez oprogramowanie tej warstwy.

Warstwy MAC i LLC

Warstwa łącza danych podzielona jest na dwie podwarstwy – MAC (Media Access Control) i LLC (Logical Link Control). Funkcje pierwszej z nich to: odczytywanie i zapis adresów sprzętowych (MAC), definiowanie formatu ramki, realizowanie metody dostępu do medium transmisyjnego (żeton, CSMA/CD), kontrola błędów. Podstawową funkcją podwarstwy LLC jest tzw. multipleksowanie i de-multipleksowanie protokołów warstwy sieci.

Model warstwowy TCP/IP:

Jest uproszczeniem modelu ISO, zawiera tylko 5 warstw:

Aplikacji
Transportu
Sieci
Łącza danych
Fizyczna

Ethernet II:

W sieciach Ethernet stosowane jest kodowanie typu Manchester. Polega ono na tym, że bit 1 jest kodowany zmianą napięcia z wyższego na niższe, a bit 0 – odwrotnie. Zmiana zachodzi w połowie czasu trwania bitu. Ethernet II jest najczęściej stosowaną odmianą protokołu Ethernet. Oto budowa ramki Ethernet II:

Preambuła + SFD	Adres odb.	Adres nad.	Typ	Dane	FCS
8 oktetów	6 oktetów	6 oktetów	2 oktety	46-1500	4 oktety

Preambuła – ciąg 56 bitów (na przemian jedynki i zera) umożliwiający synchronizację nadawcy i odbiorcy. Chociaż karty sieciowe po obu stronach łącza ustawione są na tę samą prędkość transmisji, zazwyczaj między ich rzeczywistymi prędkościami jest niewielka różnica, która musi być zniwelowana w wyniku synchronizacji.

SFD – znacznik początku ramki (ang. Start-of-Frame Delimiter), czyli ciąg następujących 8 bitów: 1 0 1 0 1 0 1 0 1 1

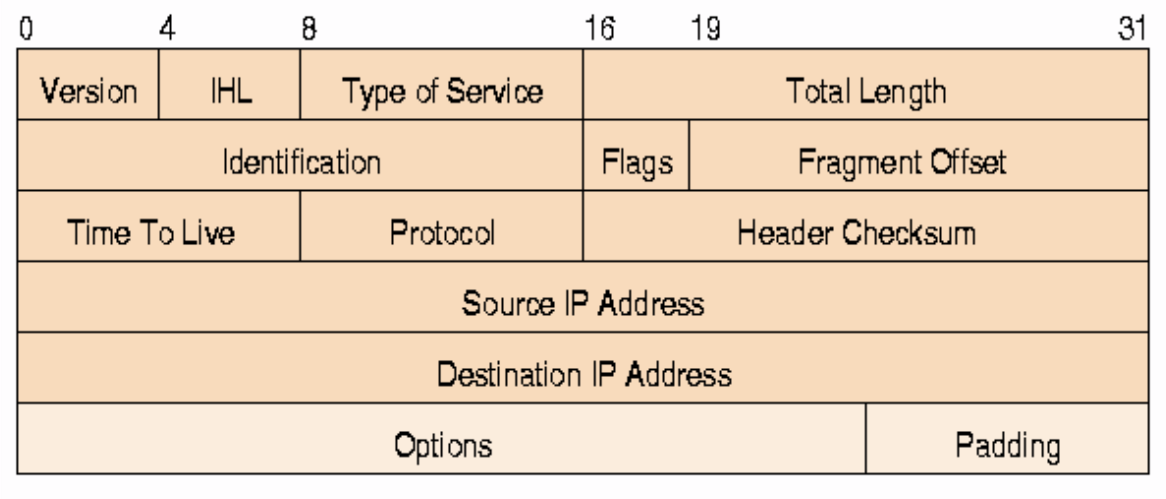
Adres odbiorcy/nadawcy – docelowy/źródłowy adres MAC składający się z 6 oktetów

Typ – jeśli wartość w tym polu jest większa od 1500 (0x5DC), jest w nim kod protokołu warstwy sieci. Dla IP jest to 0x800, dla ARP – 0x806, dla IPX – 0x8137. W przeciwnym przypadku pole zawiera informację o **długości** pola danych i ramka nie jest typu Ethernet II, ale należy do typu Raw (standard IEEE 802.3) albo typu LLC (standard IEEE 802.2). Typ Raw nie zawiera informacji o protokole warstwy sieci i jest stosowany w sieciach Novell. W przypadku LLC, dwa pierwsze bajty znajdujące się za polem „Typ" zawierają informację o protokole warstwy sieci (DSAP, SSAP). Rozszerzeniem typu LLC jest typ SNAP.

Dane – ze względu na fizyczne parametry sieci Ethernet, wprowadzone są ograniczenia na całkowitą długość ramki. Jeśli w trakcie nadawania wystąpi kolizja, to stacja musi mieć możliwość stwierdzenia tego faktu jeszcze przed zakończeniem nadawania. Wynika stąd dolne ograniczenie na długość ramki. Ramki nie mogą być też zbyt długie, m.in. ze względu na możliwość utraty synchronizacji między stacją nadającą i odbierającą.

FCS – suma kontrolna (Frame Control Sequence), wykorzystywana wówczas, jeśli oprogramowanie warstwy łącza danych zawiera mechanizmy sprawdzania poprawności transmisji. Jest obliczana za pomocą algorytmu CRC (ang. Cyclic Redundancy Check).

Budowa nagłówka IPv4:



Opis pól:

Version – numer wersji protokołu IP (4 – IPv4, 6 – IPv6)

IHL – długość nagłówka IPv4 (Internet Header Length) w słowach 32-bitowych (czyli 4-bajtowych). Jest konieczne, ponieważ w skład nagłówka IP wchodzi pole opcji o zmiennej długości. Minimalna długość nagłówka IP wynosi 20 bajtów (brak opcji), natomiast maksymalna – 15*4=60 bajtów. Górne ograniczenie na długość nagłówka IP wynika z faktu, że pole IHL składa się z 4 bitów i 15 jest największą liczbą, jaką można w nim zapisać.

TOS – typ obsługi (Type of Service), może zawierać wskazania dla routerów odnośnie wyboru trasy dla pakietu, ale zazwyczaj składa się z samych zer. Pierwsze sześć bitów tego pola tworzy tzw. pole DSCS (ang. Differentiated Services Code Point). Bity 0, 1 i 2 oznaczają ważność pakietu (precedence), bit 3 – żądanie małego opóźnienia (delay), bit 4 – żądanie dużej przepustowości (throughput), bit 5 – żądanie dużej niezawodności (reliability). W tabeli routingu też występuje rubryka TOS. Router wybiera daną trasę, jeśli zawartość pola TOS pakietu jest zgodna z zawartością rubryki TOS dla tej trasy. Bity 7 i 8 tworzą tzw. pole ECN (Explicit Congestion Notification) i są używane do powiadamiania o zatorach w sieci.

Total Length – całkowita długość pakietu IP (łącznie z nagłówkiem) mierzona w bajtach (słowach 8-bitowych). Pole to ma długość 16 bitów, z czego wynika, że maksymalna długość pakietu IP wynosi 2^16 – 1 = 65535 oktetów.

Trzy kolejne pola wykorzystywane są przez mechanizmy fragmentacji i składania pofragmentowanych pakietów IP. Fragmentacja polega na dzieleniu pakietu na części o długości nie przekraczającej MTU (maksymalna długość pola danych ramki w bajtach) tej sieci, do której pakiet ma być wysłany. Przykładowo, MTU wynosi 1500 dla sieci Ethernet, a 4470 dla FDDI. Fragmentacji mogą dokonywać zarówno hosty jak i routery. Jeśli jest to konieczne, router fragmentuje pakiety przekazywane z sieci o większym MTU do sieci o mniejszym MTU. Łączenie pakietu w całość odbywa się tylko w komputerze docelowym.

Identification – liczba identyfikująca pakiet, jest kopiowana z nagłówka całego pakietu do nagłówka każdego fragmentu, który powstaje w wyniku podziału danego pakietu. Informuje komputer docelowy, które fragmenty składają się na dany pakiet.

Flags – pierwszy bit tego pola jest zawsze zerem. Drugi bit (DF – do not fragment) służy do informowania routerów, czy mogą fragmentować pakiet; jeśli ten bit jest jedynką, to w przypadku konieczności fragmentowania, router odrzuca pakiet i wysyła do jego nadawcy odpowiedni komunikat. Trzeci bit (MF – more fragments) informuje, czy dany fragment jest ostatni (wartość 0), czy nie (wartość 1).

Fragment Offset – odsunięcie początku pola danych fragmentu od początku pola danych całego pakietu. Jest podawane w słowach 8-bajtowych (64-bitowych). Zgodnie z definicją, dla pierwszego fragmentu ma wartość zero.

Time to Live – oznacza maksymalny czas w sekundach, przez jaki pakiet może przebywać w sieci. Pole to jest wypełniane wartością początkową przez nadawcę pakietu. Wartość ta zależy od systemu operacyjnego. Każdy router na trasie pakietu sprawdza wartość TTL i przed wysłaniem pakietu zmniejsza ją o liczbę większą lub równą 1 (czas w sekundach, przez jaki pakiet był przetwarzany). Jeśli router otrzyma pakiet z TTL równym 1, to odrzuca go i wysyła do nadawcy komunikat o przekroczeniu TTL. W ten sposób z sieci usuwane są np. pakiety krążące w pętli, która może powstać w wyniku błędnej konfiguracji routerów.

Protocol – określa protokół następnej warstwy, zgodnie z opisem w dokumencie RFC 1060. Przykładowe wartości to 6 dla TCP, 17 dla UDP, 2 dla IGMP, 1 dla ICMP.

Header Checksum – suma kontrolna nagłówka. Przy jej obliczaniu uwzględniane są wyłącznie pola nagłówka, bez pola danych. Musi być aktualizowana na każdym routerze, ponieważ zawartość pola TTL ulega, a niektórych innych pól (np. TOS, Options) – może ulec zmianie, po przejściu pakietu przez router.

Następne dwa pola zawierają źródłowy i docelowy adres IP

Options – opcjonalne pole opcji. Każda opcja składa się z oktetu, w którym zapisany jest jej kod, opcjonalnego oktetu określającego jej długość, oraz opcjonalnego ciągu oktetów zawierających dane opcji. Pierwszy bit kodu opcji określa, czy opcje powinny być kopiowane do wszystkich fragmentów (wartość 1), czy tylko do pierwszego fragmentu (wartość 0). Dwa kolejne bity określają klasę opcji, natomiast pięc pozostałych – numer opcji. W poniższej tabeli przedstawione są najważniejsze opcje IP.

Klasa	Numer	Długość	Opis
0	0	1	Koniec listy opcji. Zajmuje tylko 1 oktet.
0	3	zmienna	Swobodne trasowanie według nadawcy.
0	9	zmienna	Rygorystyczne trasowanie według nadawcy
0	7	zmienna	Zapisywanie trasy pakietu.
2	4	zmienna	Zapisywanie stempli czasowych wzdłuż trasy.

Padding – wypełnienie do pełnego słowa 32-bitowego. Składa się z samych zer.

Trasowanie według nadawcy (ang. source routing) polega na określeniu trasy pakietu (adresy kolejnych routerów) przez jego nadawcę. Trasa jest więc z góry narzucona, a nie wytyczana na każdym jej etapie przez kolejne routery. Trasowanie swobodne polega na podaniu adresów tych routerów, przez które pakiet musi przejść na swojej trasie, ale oprócz nich może też przechodzić przez inne. Z kolei trasowanie rygorystyczne polega na podaniu adresów wszystkich routerów na trasie pakietu, czyli musi on przejść przez wszystkie podane routery i nie może przechodzić przez inne. Trasowanie wg nadawcy jest wykorzystywane w celach diagnostycznych (np. sprawdzenie możliwości przesłania pakietu daną trasą), albo w celu ominięcia pewnych routerów mogących znaleźć się na trasie pakietu.

Fragmentacja przy wysyłaniu danych w sieć Ethernet:

Struktura ramki Eth. II przenoszącej pakiet IP z nagłówkiem w podstawowej wersji (1 kreska to 1 bajt):

Jeśli pakiet IP nie mieści się w jednej ramce, to jest dzielony na fragmenty, z których każdy jest przesyłany w osobnej ramce. Parametry fragmentu znajdują się w drugim słowie nagłówka IP. Składa się ono z 3 pól: 16-bitowego pola Identyfikacja, 3-bitowego pola flag, oraz 13-bitowego pola Offset.

W pole Identyfikacja każdego fragmentu jest wstawiana zawartość pola Identyfikacja całego pakietu. Pozwala to stacji docelowej rozpoznać fragmenty pochodzące z tego samego pakietu.

Flagi: pierwsza to 0, druga to DF (ang. don't fragment), trzecia to MF (ang. more fragments)

MF=1: dany fragment nie jest ostatni
MF=0: dany fragment jest ostatni

MTU - maximum transfer unit, maksymalna długość (w bajtach) pola danych ramki protokołu realizującego komunikację wewnątrzsieciową (MTU=1500 dla Ethernet II)

Router znajdujący się na trasie pakietu może łączyć sieci o różnych MTU. Jeśli pakiet ma być przekazany z sieci o większym MTU do sieci o mniejszym MTU, to może być konieczna jego fragmentacja. Jeśli DF=1, to router nie przekaże dalej pakietu (bo nie zezwala na to wartość flagi DF), a do jego nadawcy wyśle odpowiedni komunikat ICMP.

Offset: odsunięcie pola danych fragmentu od początku pola danych całego pakietu, podawane w słowach 8-bajtowych

Uwaga: Liczba bitów pola Offset to 13, więc największa liczba, którą można w nim zapisać to 2^13 – 1 = 8191. Z tego względu offset jest podawany w słowach 8-bajtowych, bo całkowita długość pakietu IP może wynosić 2^16 – 1 = 65535 bajtów, więc musi być możliwe zapisywanie odstępów większych niż 8191 bajtów. W konsekwencji, długość w bajtach pola

danych każdego fragmentu, z wyjątkiem ostatniego, musi być całkowitą wielokrotnością ośmiu!

Przykłady fragmentacji:

Przykład 1
Przedstawić fragmenty powstałe przy wysyłaniu w sieć Ethernet (MTU=1500) pakietu IP o maksymalnej długości (2^16 – 1 = 65 535 bajtów razem z nagłówkiem), jeśli nagłówek IP ma standardową długość wynoszącą 20 bajtów.

Ponieważ
długość nagłówka IP + Długość pola danych fragmentu ≤ MTU,
więc
długość pola danych fragmentu ≤ 1480.

Czy pole danych fragmentu może mieć długość 1480 bajtów?
Tak, bo 1480 jest całkowitą wielokrotnością ośmiu (1480 = 185*8).

Z kolei
długość pola danych całego pakietu = 65535 – długość nagłówka IP = 65535 – 20 = 65515,
więc przy wysyłaniu pakietu powstaną 44 fragmenty o długości pola danych 1480 bajtów, oraz 45-ty (ostatni) fragment o długości pola danych 395 bajtów (65515 = 44×1480 + 395). W standardowej notacji, czyli „długość @ przesunięcie MF/LF" (długosc i przesuniecie bajtach), rezultat powyższej fragmentacji jest zapisywany następująco:

1 fragment: 1480 @ 0 MF
2 fragment: 1480 @ 1480 MF
3 fragment: 1480 @ 2960 MF
...
44 fragment: 1480 @ 63640 MF
45 fragment: 395 @ 65120 LF

Uwaga: Jeśli transmisja odbywa się w sieci Ethernet, wówczas minimalna długość pola danych ramki wynosi 46 oktetów. Przy założeniu, że nagłówek IP zajmuje 20 oktetów, dane fragmentu muszą mieć długość co najmniej 26 oktetów. **Zatem, dla pakietu IP ze standardowym (brak opcji) nagłówkiem, minimalna długość pola danych fragmentu IP wysyłanego w sieć Ethernet wynosi 26.** W razie konieczności pole danych ostatniego fragmentu uzupełniane jest bitami zerowymi.

Przykład 2

W łącze PPP (MTU=296) wysyłany jest segment TCP zawierający 1200 bajtów danych. Zakładamy, że nagłówki IP i TCP mają po 20 oktetów. Przedstaw powstałe fragmenty używając notacji „długość @ przesunięcie MF/LF". Długość i przesunięcie mają być podane w bajtach (w polu offset przesunięcie jest podawane w słowach 8-bajtowych).

Segment TCP = Nagłówek TCP + Dane TCP

Pakiet IP = Nagłówek IP + Dane IP = Nagłówek IP + Nagłówek TCP + Dane TCP

Długość pola danych pakietu IP = Długość nagłówka TCP + Długość pola danych TCP

Długość pola danych całego pakietu IP = 20 + 1200 = 1220

Z kolei każdy fragment zawiera nagłówek IP i pole danych, a ponieważ MTU=296, więc
długość nagłówka IP + długość pola danych fragmentu ≤ 296,
skąd wynika, że
długość pola danych fragmentu ≤ 276
Czy pole danych fragmentu może mieć długość 276 bajtów?
Nie, bo 276 = 34*8 + 4, więc 276 nie jest całkowitą wielokrotnością ośmiu.
Maksymalna długość pola danych fragmentu to 272 = 34*8.

Powstanie zatem 5 fragmentów, czyli 4 fragmenty z polem danych o długości 272 bajty oraz piąty fragment z polem danych o długości 132 (1220 = 4*272 + 132). W standardowej notacji rezultat powyższej fragmentacji zapisuje się w następujący sposób:

272 @ 0 MF
272 @ 272 MF
272 @ 544 MF
272 @ 816 MF
132 @ 1088 LF