

Sieci komputerowe

Wykład 5 TCP

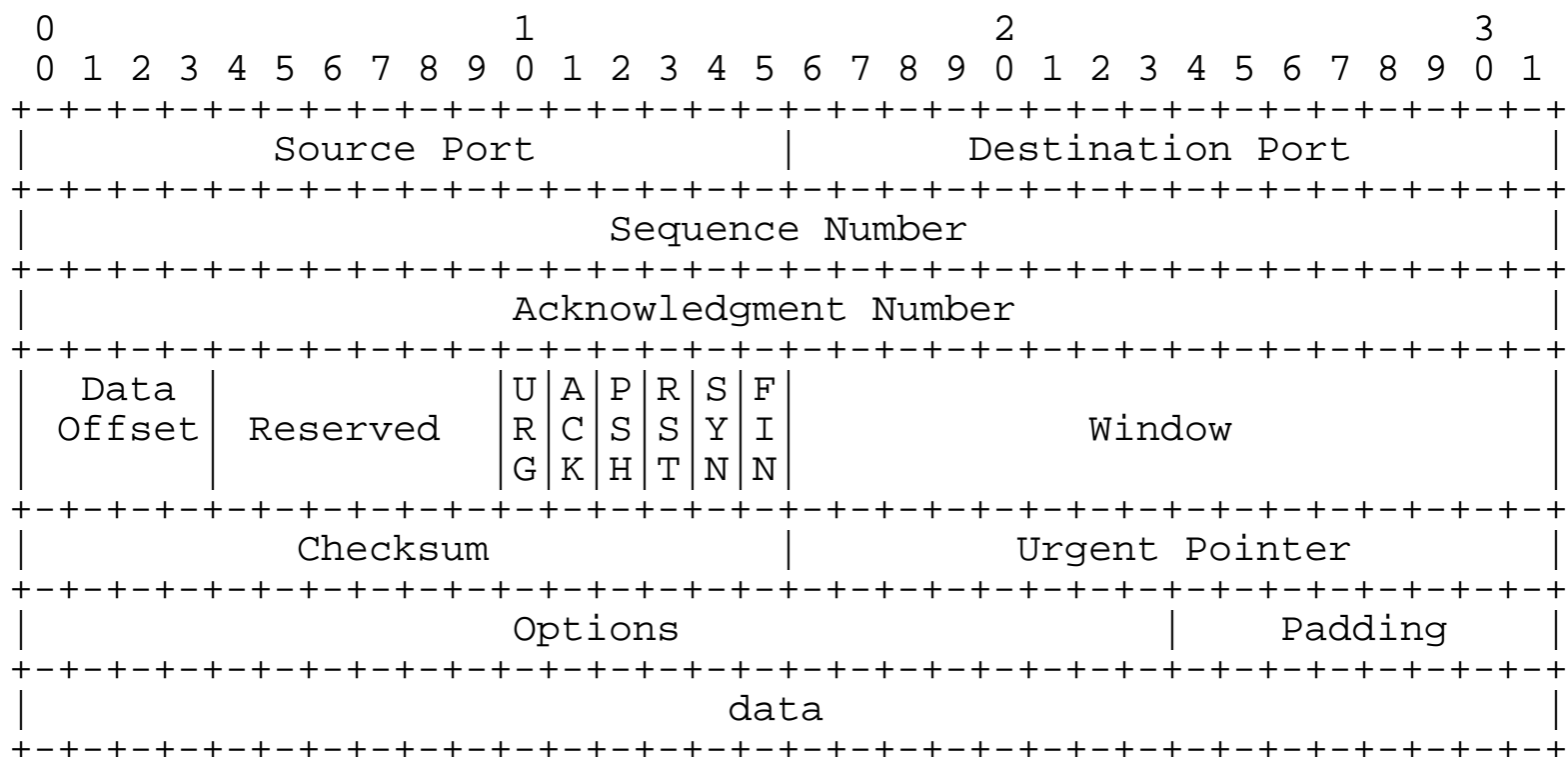
Własności TCP

- Jest zorientowany połączeniowo. Zanim zostaną przesłane jakiegokolwiek dane, musi zostać nawiązane połączenie
- Zapewnia niezawodność (dane przesyłane przez aplikację są dzielone na tzw. segmenty, które wg. TCP mają najlepszy rozmiar, po wysłaniu segmentu jest uruchamiany zegar i rozpoczyna się oczekiwanie na potwierdzenie odebrania segmentu przez drugą stronę). W przypadku nieotrzymania potwierdzenia, segment jest wysyłany ponownie
- Sortuje segmenty i w razie potrzeby odrzuca zdublowane
- Stosuje sumę kontrolną nagłówka i danych do kontroli poprawności. Jeśli zostanie wykryty błąd sumy kontrolnej potwierdzenie nie jest wysyłane
- Ponieważ TCP wykorzystuje mechanizm połączeń, nie jest możliwe zastosowanie go do transmisji typu broadcast

Własności TCP c.d,

- Umożliwia przesyłanie danych w obie strony (tzw. tryb full duplex)
- Zapewnia kontrolę przepływu za pomocą mechanizmu okien
- W celu poprawy efektywności, stosuje się algorytm opóźnianych, skumulowanych potwierdzeń

Nagłówek protokołu TCP



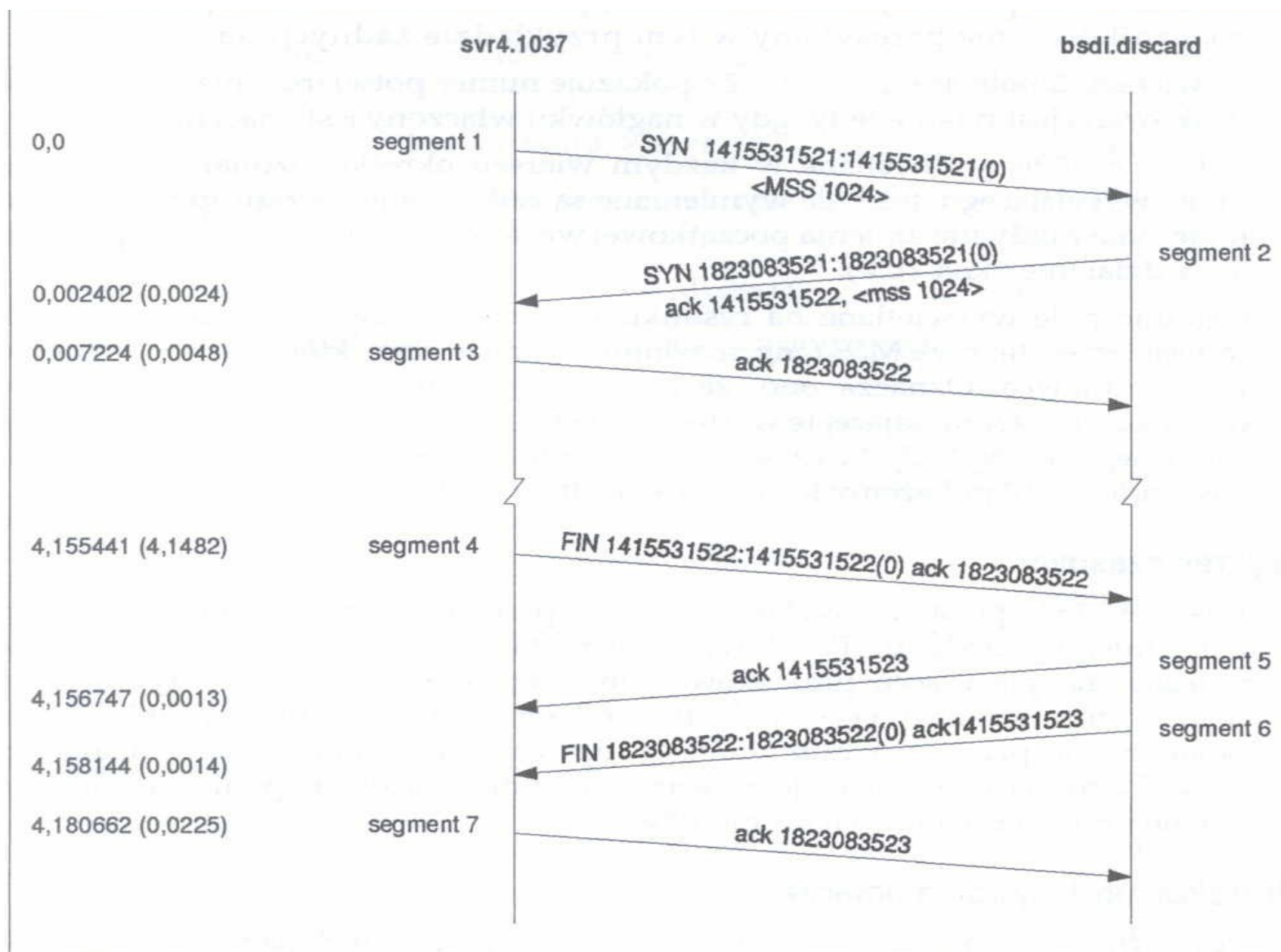
Najważniejsze pola nagłówka TCP

- Numery sekwencyjny i potwierdzenia:
 - Numer sekwencyjny służy do numerowania bajtów
 - Numer potwierdzenia jest następnym spodziewanym numerem sekwencyjnym
- Pole *rozmiar okna* służy do kontroli przepływu
- Pole *Opcje* – najważniejsza: MSS (ang. Maximum Segment Size)

Znaczniki w nagłówku TCP

- Znaczniki
 - URG – znacznik ważności pola „wskaźnik pilności”
 - ACK – znacznik ważności pola „numer potwierdzenia”
 - PSH – znacznik ten, jeśli ustawiony, oznacza, że odbiorca powinien przekazać dane do aplikacji tak szybko, jak to możliwe
 - RST – przerwanie połączenia
 - SYN – synchronizacja numerów sekwencyjnych w celu inicjalizacji połączenia
 - FIN – nadawca zakończył wysyłanie danych

Nawiązywanie i kończenie połączenia

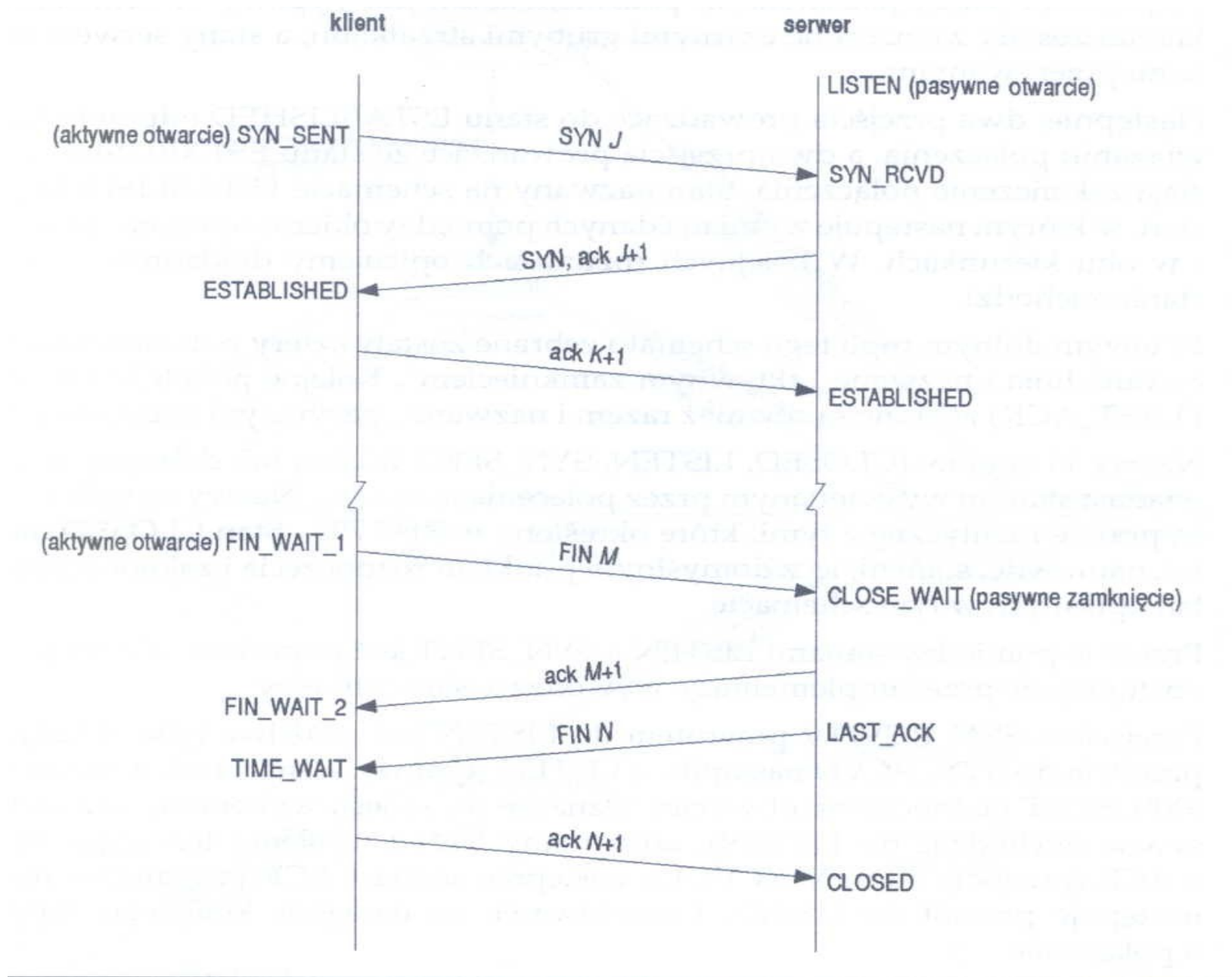


Rysunek 18.3 Wykres czasowy przedstawiający nawiązywanie i zakończenie połączenia

Proces nawiązywania połączenia

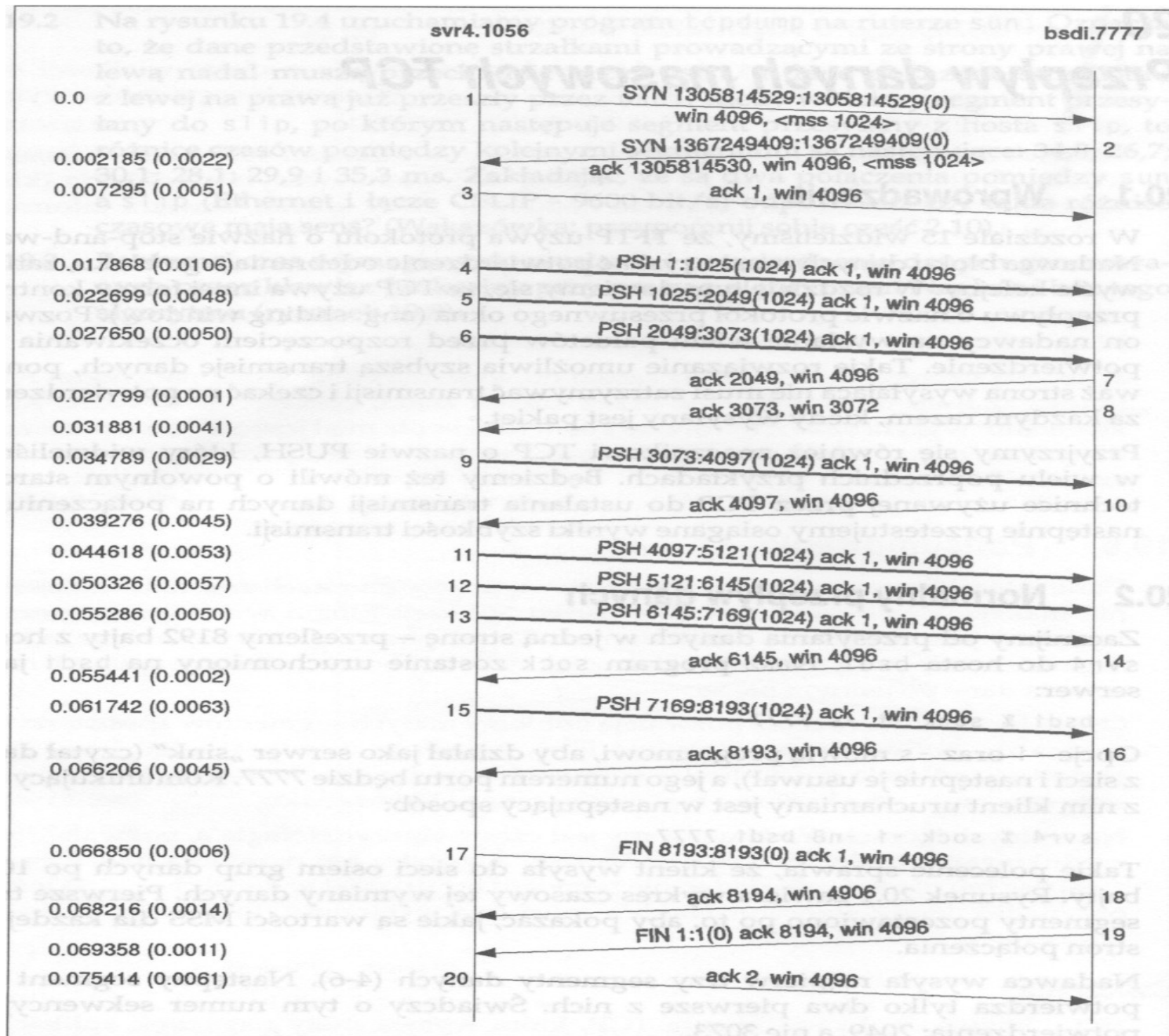
- Proces nawiązywania połączenia polega na synchronizacji numerów sekwencyjnych (trójstopniowy handshake), tak aby było wiadomo jak numerować bajty
- Po nawiązania połączenia możliwa jest komunikacja full-duplex
- Istotny jest wybór początkowego numeru sekwencyjnego (numeru ISN), tak aby był unikalny dla danego połączenia (dba o to implementacja TCP)

Stany TCP



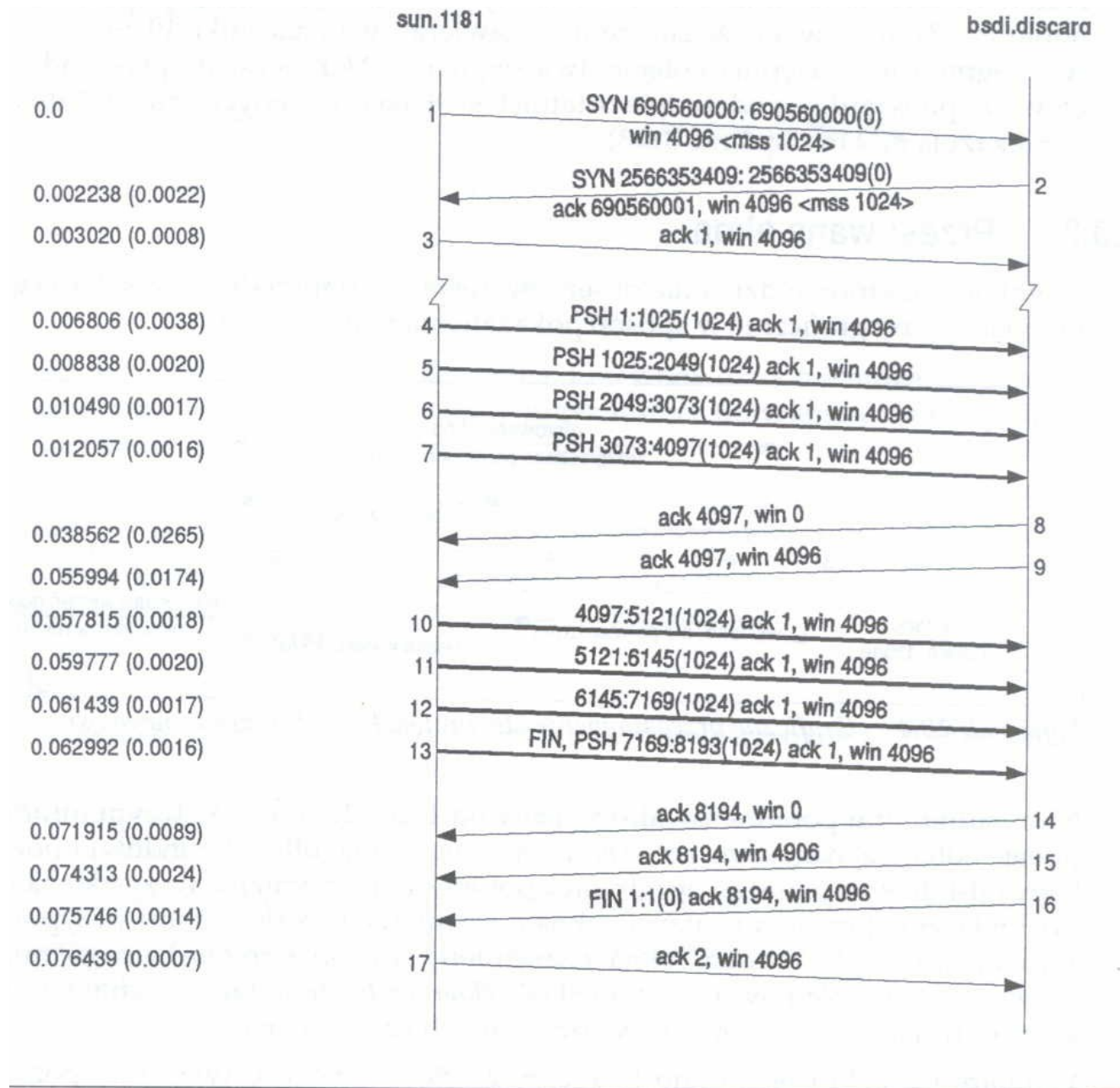
rysunek 18.13 Stany TCP odpowiadające typowemu przebiegowi nawiązywania i zakończenia połączenia TCP

Przepływ danych masowych



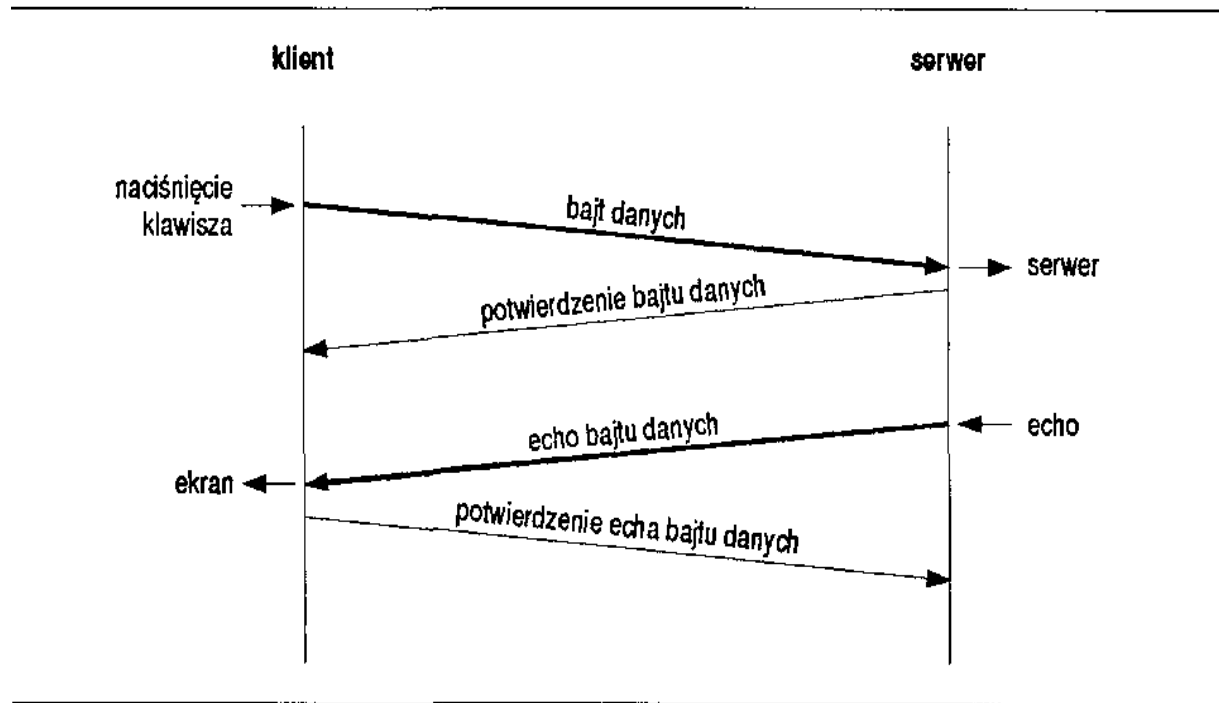
Rysunek 20.1 Przesyłanie 8192 bajtów z svr4 do bsdi

Szybki nadawca, wolny odbiorca



ysunek 20.3 Przesyłanie 8192 bajtów z szybkiego nadawcy do wolnego odbiorcy

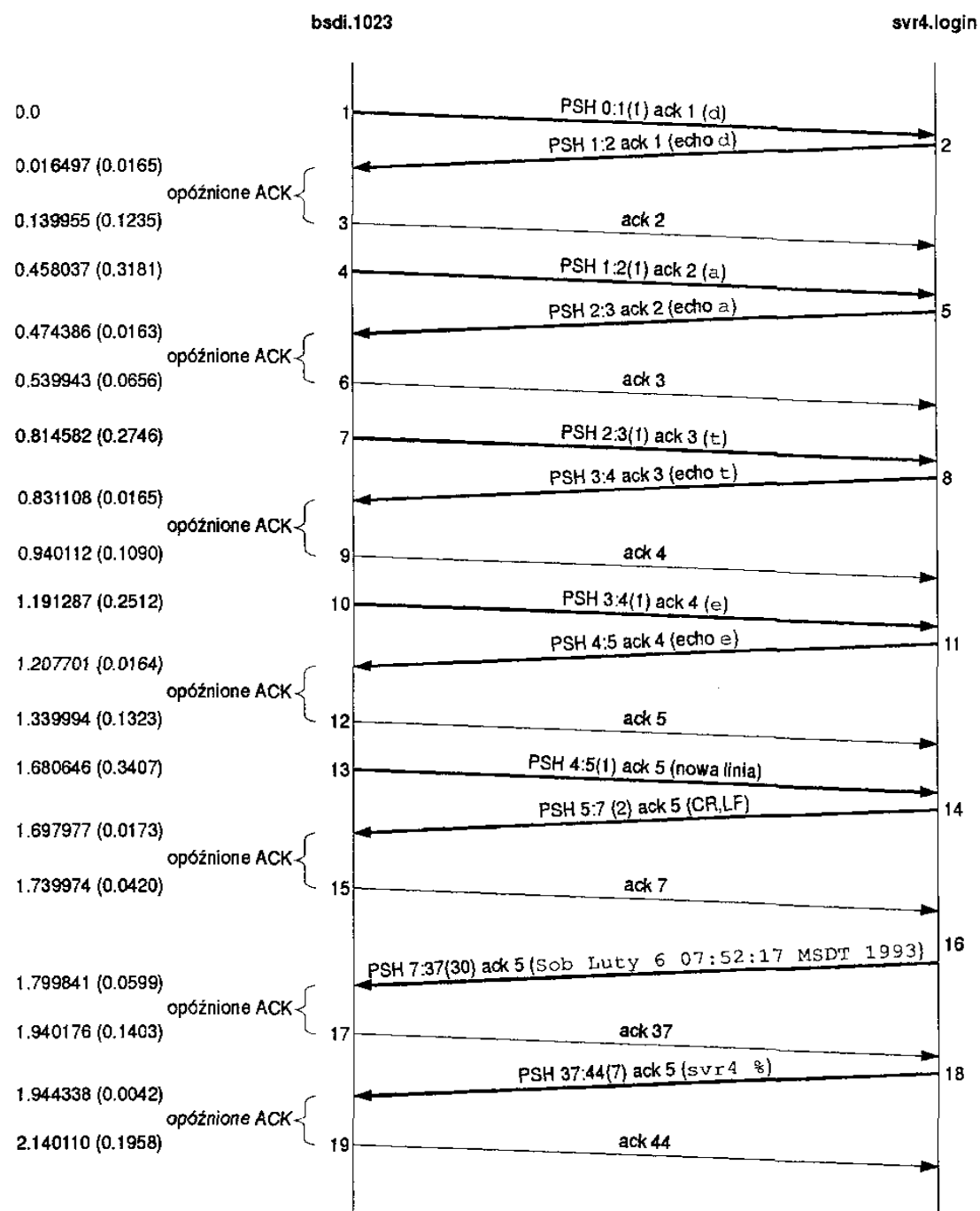
Przepływ danych interaktywnych



rysunek 19.1 Jeden z możliwych sposobów wymiany echa wpisywanych znaków

- Przesyłane są pojedyncze znaki, pakiet IP dla przesłania 1 znaku ma długość 41 bajtów!
- Dla jednego przesłanego znaku tworzone są zwykle 4 segmenty

Opóźnione potwierdzenia w przesyłaniu interaktywnym



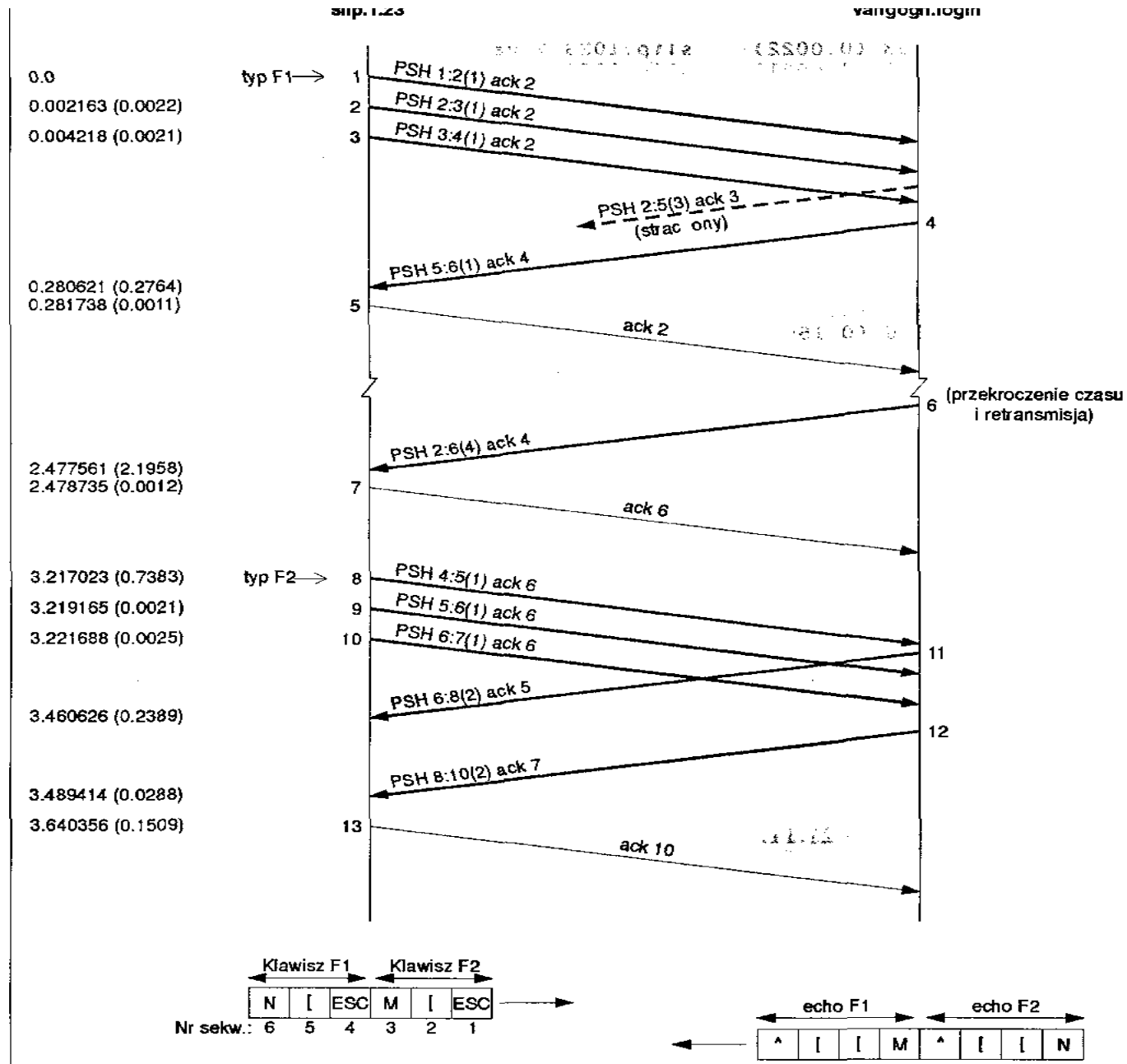
Algorytm Nagle'a

- Służy do minimalizacji liczby wysyłanych segmentów
 - Jeśli nadawca miałby wysłać dwa niepełne segmenty TCP pod rząd, to przed wysłaniem drugiego oczekuje na ACK pierwszego
 - Może więc występować wstrzymanie wysyłania danych, tak, aby później wysłać więcej w jednym segmencie
 - Istnieje możliwość wyłączenia algorytmu Nagle'a, za pomocą odpowiedniej opcji API gniazd

Retransmisja

- Retransmisja następuje w momencie, gdy TCP nie otrzymał potwierdzenia dla któregoś z segmentów
- Konieczne jest wyznaczenie czasu oczekiwania, po którym ma nastąpić retransmisja
- Aby wyznaczyć czas oczekiwania, TCP musi mierzyć czas podróży segmentów (tzw. RTT-round trip time)
 - Czas podróży: $R = \alpha R_p + (1 - \alpha)M$, $\alpha = 0.9$
 - Czas oczekiwania: $RTO = R\beta$, $\beta = 2$

Retransmisja, c.d.



Repakietyzacja

- Po zakończeniu odliczania czasu oczekiwania następuje retransmisja, ale TCP nie musi retransmitować identycznych segmentów
- Może wykonać repakietyzację, czyli wysłać segment większy

Algorytm powolnego startu

- Służy do kontroli przepływu związanego z obciążeniem sieci (inaczej niż w przypadku ogłaszania wielkości okna)
 - Problem jest wykrywany, gdy pojawiają się straty segmentów
- Nie następuje transmisja mająca na celu wypełnienie okna ogłaszanego przez odbiorcę
- Liczba wysyłanych segmentów bez potwierdzenia zaczyna się od 1, i jest stopniowo zwiększana wraz z otrzymywaniem kolejnych ACK

Powolny start c.d.

- Algorytm powolnego startu wymaga wprowadzenia i obliczania dla każdego połączenia zmiennej *cwnd* (ang. congestion window) – okna przeciążenia
 - wartość *cwnd* jest ustawiona na początku na rozmiar jednego segmentu (MSS) i stopniowo zwiększana
 - wysyłana liczba bajtów nie może przekroczyć wartości *cwnd* i wartości okna ogłaszanego przez odbiorcę
 - za każdym potwierdzonym segmentem, wielkość *cwnd* wzrasta o MSS, powoduje to wykładniczy wzrost liczby wysyłanych segmentów (1,2,4,8...)

Algorytm zapobiegania zatorom

- Wprowadza zmienną *ssthresh* (próg powolnego stratu)
- Gdy wystąpi zator (wykryty poprzez przekroczenie czasu) $ssthresh := 1/2 * cwnd$; $cwnd := MSS$ (powolny start)
- Gdy $cwnd \geq ssthresh$, zaczyna działać algorytm zapobiegania zatorom:
 - powoduje linowy wzrost liczby wysyłanych segmentów
 - kilka konkretnych algorytmów: TCP Tahoe, Reno, Vegas, BIC, CUBIC

Algorytm zapobiegania zatorom c.d.

- Algorytm powolnego startu i zapobiegania zatorom służą do kontroli przepływu ze strony nadawcy

Łączy o dużych przepustowościach i opóźnieniach

- Pojemność potoku można zdefiniować jako:
 - $\text{pojemność}(\text{bity}) = \text{szerokość pasma (b/s)} * \text{czas podróży(s)}$
- Potok powinien być wypełniony, aby uzyskać oczekiwaną przepustowość, okno określone liczbą 16 bitową nie wystarcza
- Stosuje się opcję skalowania okna

Opcja skalowania okna

- Dla sieci o dużych przepustowościach i dużych opóźnieniach okno opisane liczbą 16 bitową może być zbyt małe
- Stosowana jest opcja TCP skalowania okna
 - stosowany jest licznik jedno bajtowy, może przyjmować wartości od 0 do 14.
 - Największy rozmiar okna to $65535 * 2^{14}$
- Opcja skalowania może się pojawiać jedynie w segmentach SYN, więc jest stała dla danego połączenia
 - Jeśli strona, która wykonuje aktywne otwarcie umieści tę opcję w segmencie SYN, ale nie dostanie jej w segmencie SYN przesłanym przez drugą stronę, to opcja ta nie może być używana