



# Протокол ТСР

# План

- Місце TCP в моделях OSI і TCP/IP
- Основні поняття TCP
- Встановлення з'єднання
- Вікно підтвердження
- Формат заголовка TCP
- Управління потоком і перевантаженням

# Місце в моделях OSI і TCP/IP

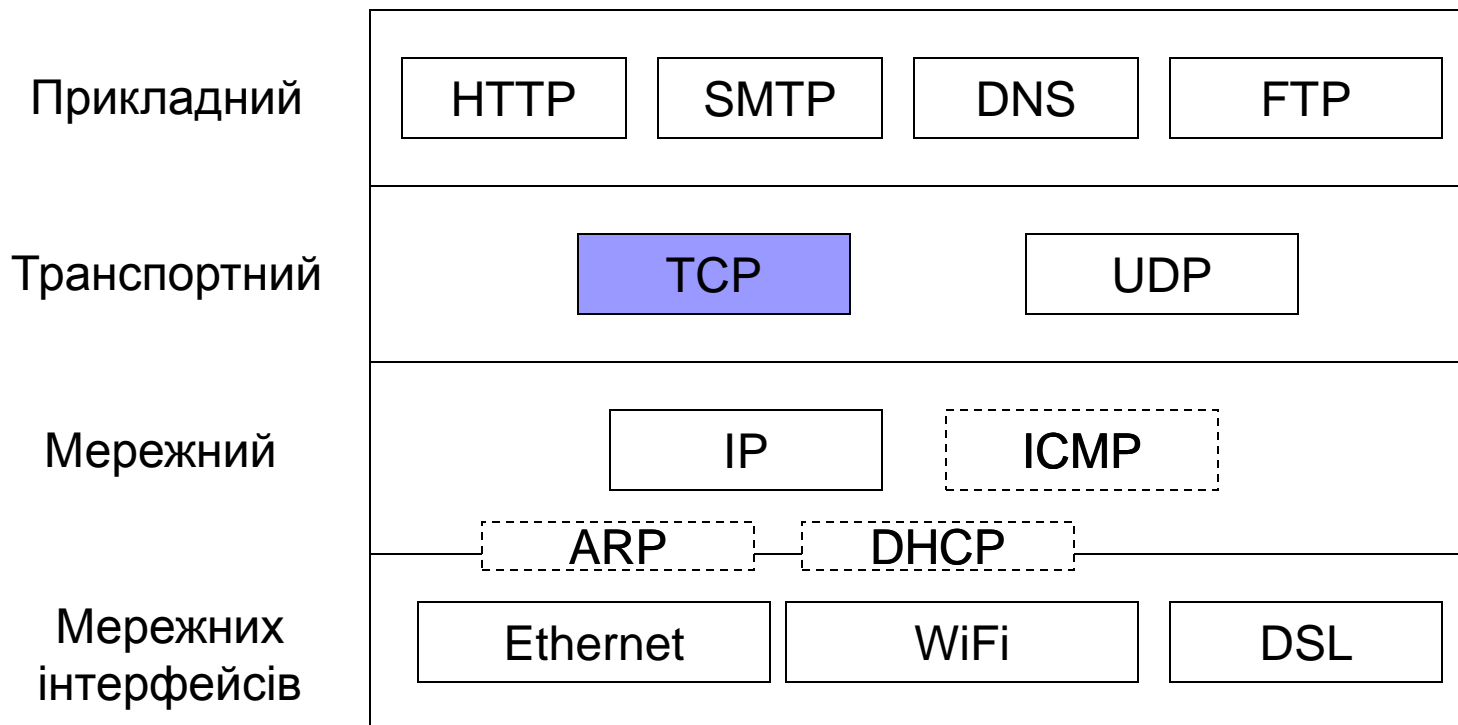
Модель OSI

Прикладний
Представлення
Сеансовий
Транспортний
Мережний
Канальний
Фізичний

Модель TCP/IP

Прикладний
Транспортний
Мережний
Мережних інтерфейсів

# Транспортні протоколи TCP/IP

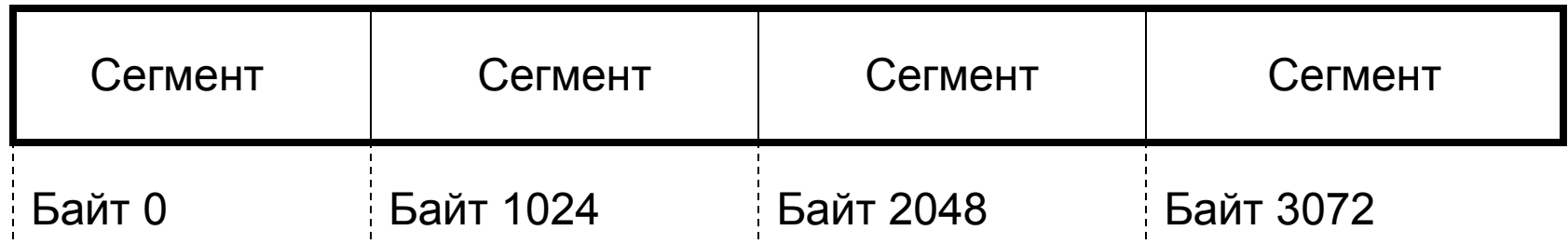


# Протокол ТСР

- Transmission Control Protocol (ТСР) - протокол керування передачею
- ТСР передає потік байт від одного процесу іншому
- Повідомлення ТСР називається сегментом
- Особливість ТСР: гарантія доставки і порядку надходження даних

# Потік байт

Потік байт від застосунку

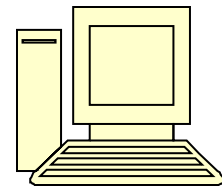
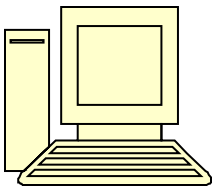


- Транспортна підсистема отримує від програми дані в вигляді потоку байт
- Потік розбивається на окремі частини - сегменти
- Протокол TCP нумерує байти в потоці
  - Сегменти не нумеруються

# Гарантія доставки

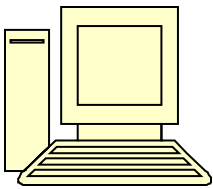
- Можливі проблеми при доставці:
  - Втрата сегментів
  - Зміна порядку доставки сегментів
  - Повторна доставка сегментів
- Механізми реалізації:
  - Нумерація повідомлень
  - Підтвердження отримання повідомлення
  - Повторна відправка при відсутності підтвердження

# Підтвердження отримання

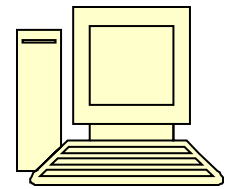




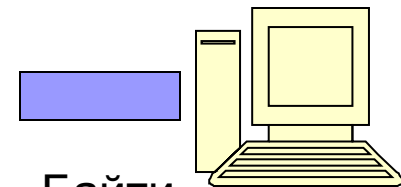
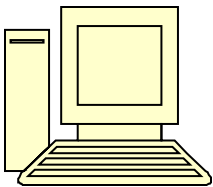
# Підтвердження отримання



Байти  
0-1024

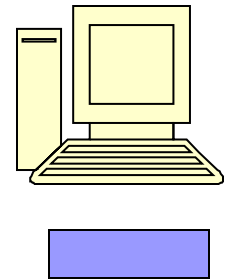
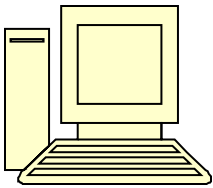


# Підтвердження отримання

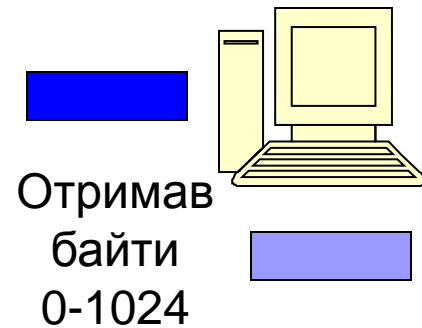
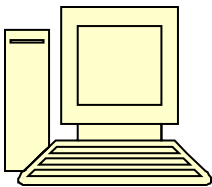


Байти  
0-1024

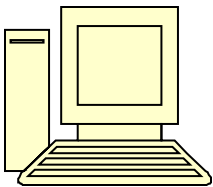
# Підтвердження отримання



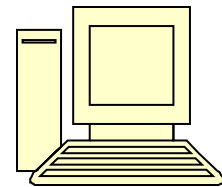
# Підтвердження отримання



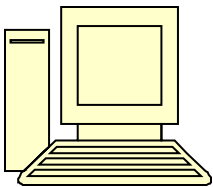
# Підтвердження отримання



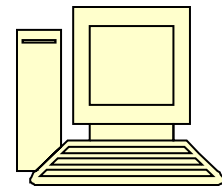
Отримав  
байти  
0-1024



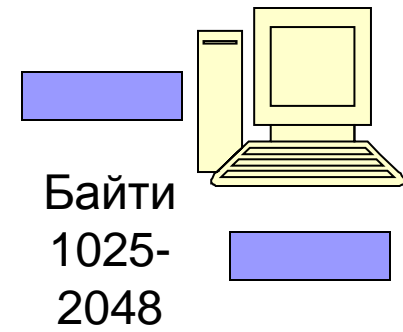
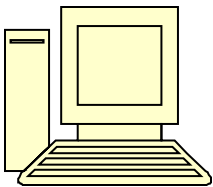
# Підтвердження отримання



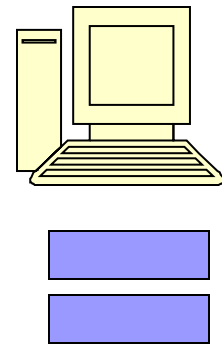
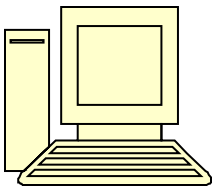
Байти  
1025-  
2048



# Підтвердження отримання

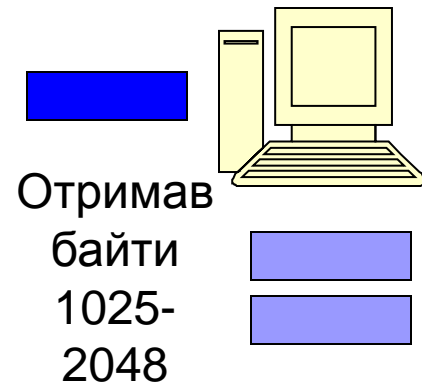
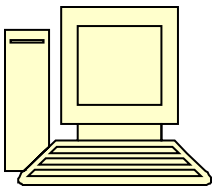


# Підтвердження отримання

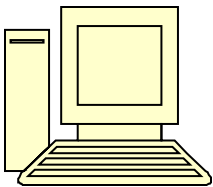




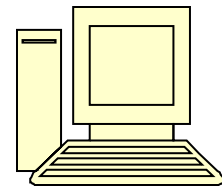
# Підтвердження отримання



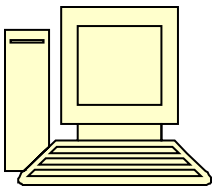
# Підтвердження отримання



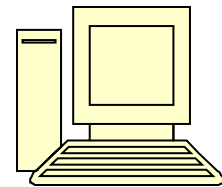
Отримав  
байти  
1025-  
2048



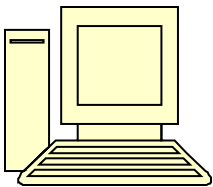
# Підтвердження отримання



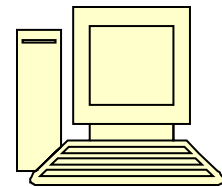
Байти  
2049-  
3072



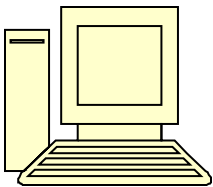
# Підтвердження отримання



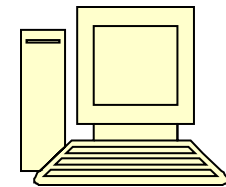
Байти  
2049-  
3072



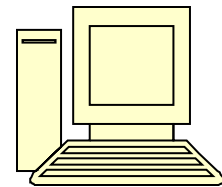
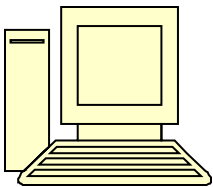
# Підтвердження отримання



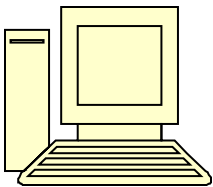
Байти  
2049-  
3072



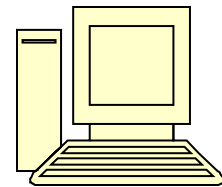
# Підтвердження отримання



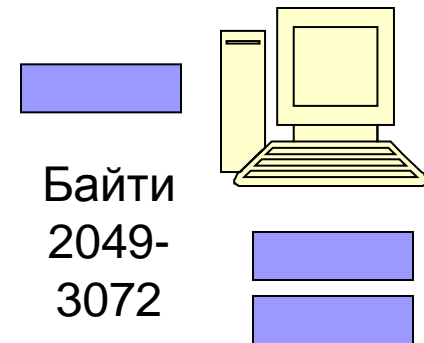
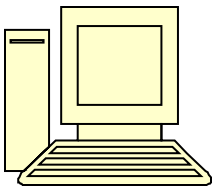
# Підтвердження отримання



Байти  
2049-  
3072

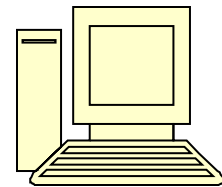
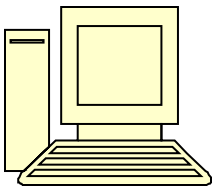


# Підтвердження отримання

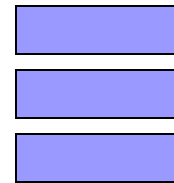
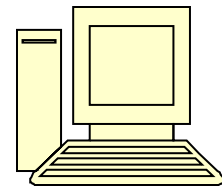
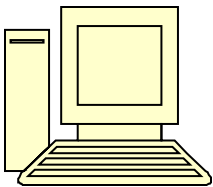




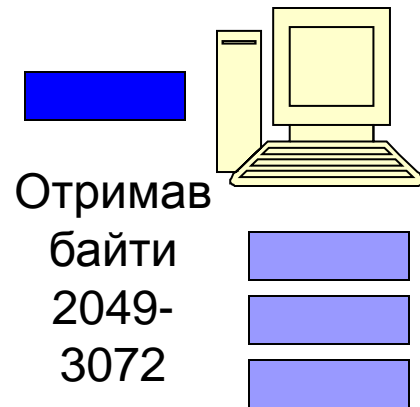
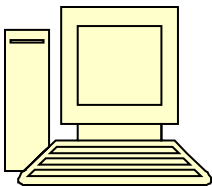
# Підтвердження отримання



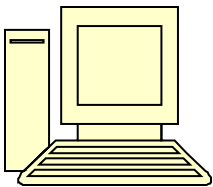
# Підтвердження отримання



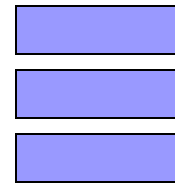
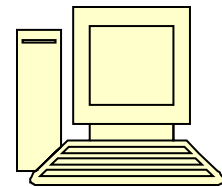
# Підтвердження отримання



# Підтвердження отримання



Отримав  
байти  
2049-  
3072



# Нумерація байтів

- Як вибрати номер для першого байту?
- Можливі проблеми:
  - Прийшов сегмент із затримкою
  - Перший сегмент втрачений, відразу прийшов другий
  - Перший сегмент висланий повторно
- Початкові номери байтів не повторюються
- Початковий номер вибирається випадковим чином, потім номери збільшуються

# З'єднання

- З'єднання - домовленість між відправником і отримувачем
- З'єднання задає:
  - Початкові номери для нумерації даних відправника і одержувача
  - Обсяг даних, які готовий прийняти одержувач
- З'єднання в ТСП дуплексне
  - Дані можуть передаватися в обидва боки
  - Підтвердження отримання і дані в одному сегменті

# Встановлення з'єднання

- Найпростіша схема:
  - Запит на установку з'єднання
  - Відповідь про встановлення з'єднання (або відмову)
- Проблеми:
  - Втрата або повторна доставка сегментів
- Застосовувана схема:
  - Триразове рукостискання

# Вікно підтвердження

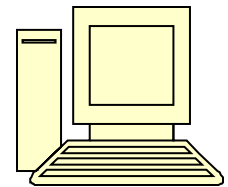
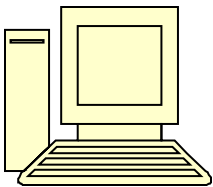
- Очікування підтвердження призводить до зниження продуктивності
- Приклад мережі:
  - Пропускна здатність 1 Гб/с
  - Час затримки передачі в один бік - 100 мс
  - Кількість сегментів в секунду - 5 шт.



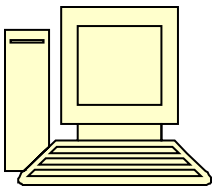
# Вікно підтвердження

- Різні варіанти підтверджень:
  - Зупинка і очікування - передача даних після отримання підтвердження кожного повідомлення (Wi-Fi, канальний рівень)
  - Вікно підтвердження - передача заданої кількості повідомлень без очікування підтвердження (TCP, транспортний рівень)
- Розмір вікна - кількість байтів даних, які можуть бути передані без отримання підтвердження
- Кумулятивне підтвердження - підтвердження прийому зазначеного байта даних і всіх попередніх

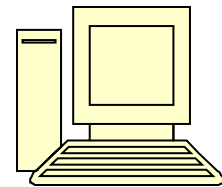
# Зупинка і очікування



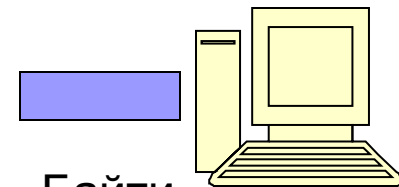
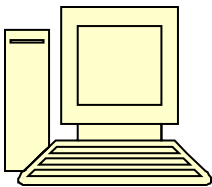
# Зупинка і очікування



Байти  
0-1024

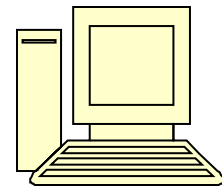
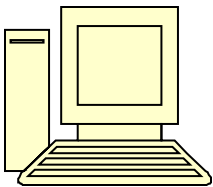


# Зупинка і очікування

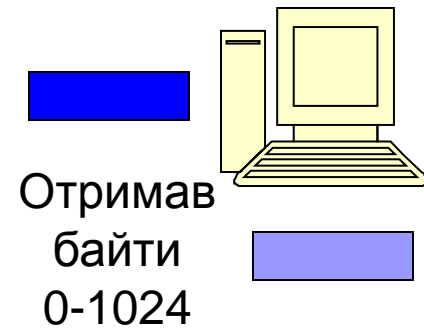
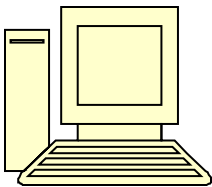


Байти  
0-1024

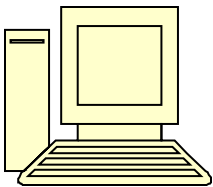
# Зупинка і очікування



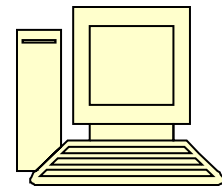
# Зупинка і очікування



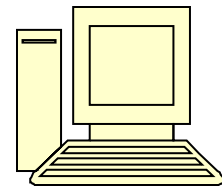
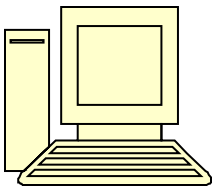
# Зупинка і очікування



Отримав  
байти  
0-1024

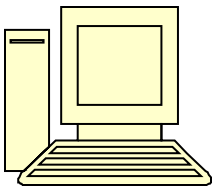


# Зупинка і очікування

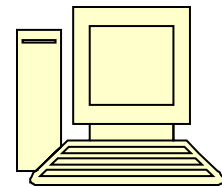




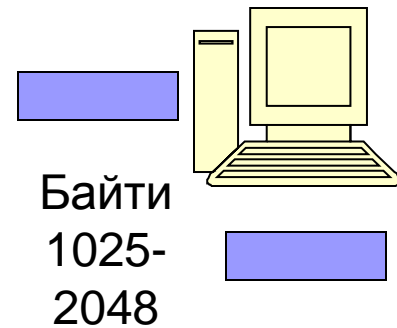
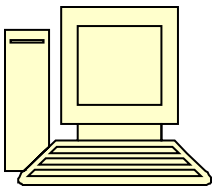
# Зупинка і очікування



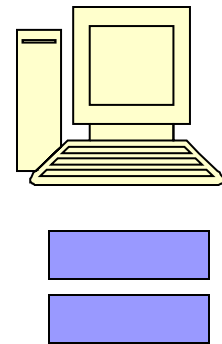
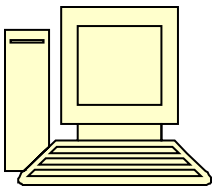
Байти  
1025-  
2048



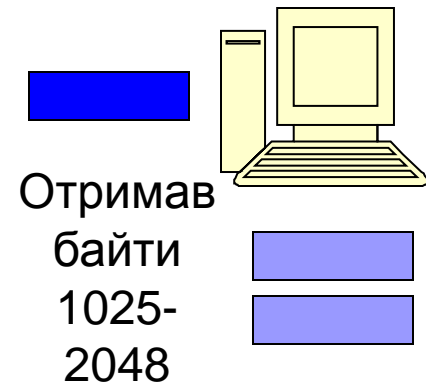
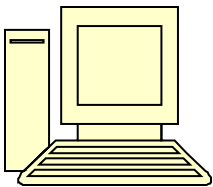
# Зупинка і очікування



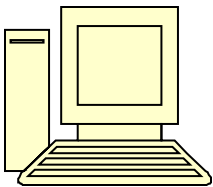
# Зупинка і очікування



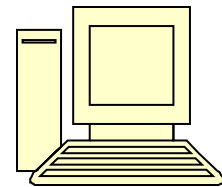
# Зупинка і очікування



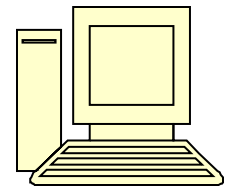
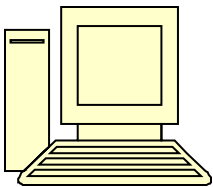
# Зупинка і очікування



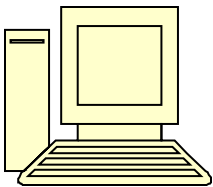
Отримав  
байти  
1025-  
2048



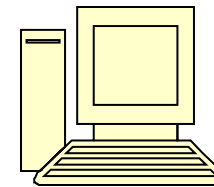
# Вікно підтвердження



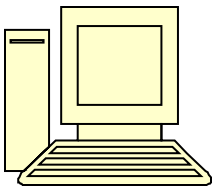
# Вікно підтвердження



Байти  
0-1024



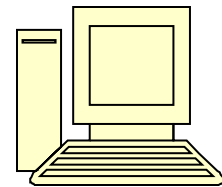
# Вікно підтвердження



Байти  
1025-  
2048

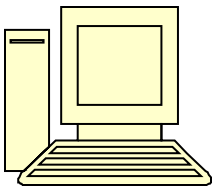


Байти  
0-1024





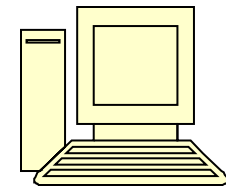
# Вікно підтвердження



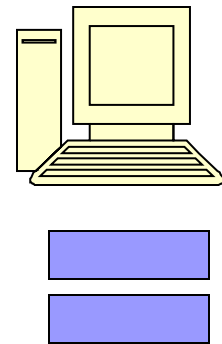
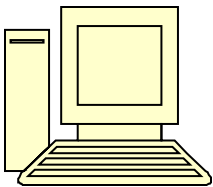
Байти  
1025-  
2048



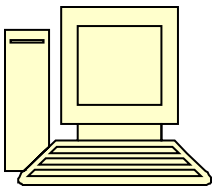
Байти  
0-1024



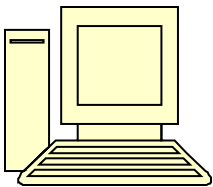
# Вікно підтвердження



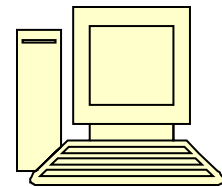
# Вікно підтвердження



# Вікно підтвердження



Отримав  
байти  
до 2048



# Формат заголовку ТСР

32 біти



# Формат заголовку ТСР

- Порядковий номер - номер байта, який пересилається в сегменті
- Номер підтвердження - номер наступного очікуваного байта
- Кумулятивне підтвердження, що всі попередні байти отримані

# Формат заголовку ТСР

- Довжина заголовка ТСР - довжина в 32-х бітних словах (4 біта)
  - Заголовок може включати параметри, тому довжина може бути різною
- 3 біта не використовується
- Прапори - 9 шт. по 1 біту

# Формат заголовку ТСР

- Розмір вікна - кількість байт, яка може бути прийняте одержувачем
- Контрольна сума - контрольна сума заголовка і даних ТСР
  - Використовується для підвищення надійності
  - Не обов'язкова
- Показчик на термінові дані - зміщення від поточного порядкового номера байта до термінових даних в сегменті



# Прапори ТСП

- URG - прапор наявності в сегменті термінових даних
  - Використовується спільно з полем «Вказівник на термінові дані»
  - Дозволяє передавати сигнали від відправника до одержувача (переривання)
- АСК - прапор підтвердження
  - Якщо прапор АСК встановлений, значить поле «Номер підтвердження» містить осмислені дані

# Прапори TCP

- PSH - прапор виштовхування (PUSH)
  - Просить одержувача відразу відправляти дані додатку, без буферизації
- Прапори RST, SYN і FIN використовуються для управління з'єднанням
  - SYN - щоб встановити з'єднання
  - FIN, RST - розрив з'єднання

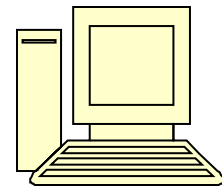
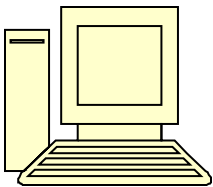
# Прапори TSP

- NS - Одноразова сума (Nonce Sum), використовується з метою покращення роботи механізму ECN
- CWR - Вікно перевантаження зменшено (Congestion Window Reduced)
- ECE - ECN-Ехо (ECN-Echo), поле показує, що відправник підтримує ECN

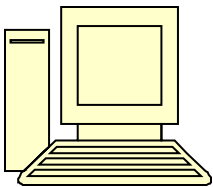
# Встановлення з'єднання

- ТСП використовує схему «Триразового рукостискання»
- Прапор SYN - ознака встановлення з'єднання
  - SYN = 1, ACK = 0 - запит встановлення з'єднання (CONNECTION REQUEST)
  - SYN = 1, ACK = 1 - підтвердження встановлення з'єднання (CONNECTION ACCEPT)
  - SYN = 0, ACK = 1 - завершення встановлення з'єднання

# Встановлення з'єднання



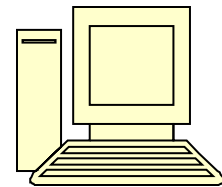
# Встановлення з'єднання



SYN = 1

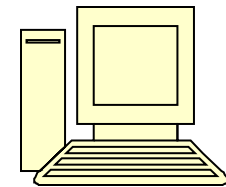
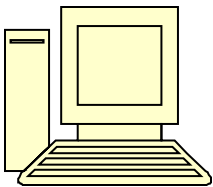
ACK = 0

Порядковий номер 1024



Відправляю байт № 1024

# Встановлення з'єднання



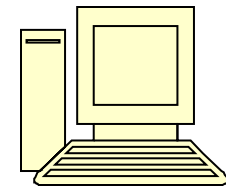
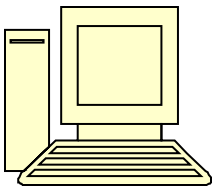
SYN = 1

ACK = 0

Порядковий номер 1024

Відправляю байт № 1024

# Встановлення з'єднання



SYN = 1

ACK = 0

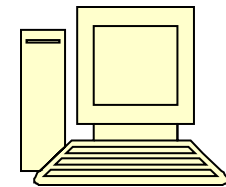
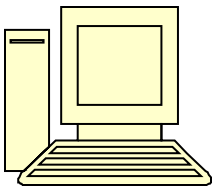
Порядковий номер: 1024

Відправляю байт № 1024

Отримав байт № 1024  
Відправляю байт № 3072



# Встановлення з'єднання



SYN = 1

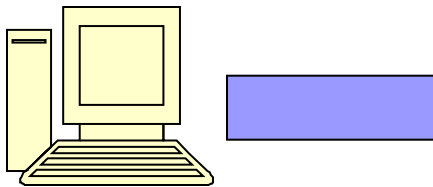
ACK = 1

Порядковий номер: 3072  
Номер підтвердження: 1025

Отправляю байт № 1024

Отримав байт № 1024  
Відправляю байт № 3072

# Встановлення з'єднання



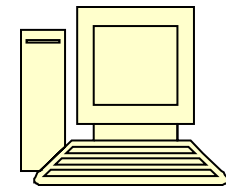
SYN = 1

ACK = 1

Порядковий номер: 3072

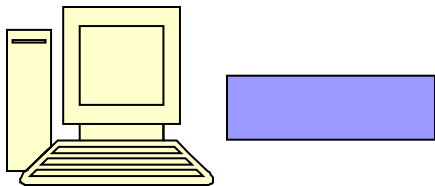
Номер підтвердження: 1025

Отправляю байт № 1024



Отримав байт № 1024  
Відправляю байт № 3072

# Встановлення з'єднання



SYN = 1

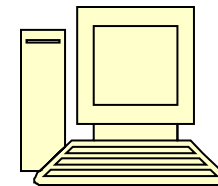
ACK = 1

Порядковий номер: 3072

Номер підтвердження: 1025

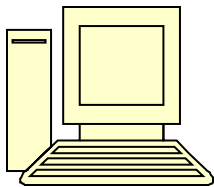
Відправляю байт № 1025

Отримав байт № 3072



Отримав байт № 1024  
Відправляю байт № 3072

# Встановлення з'єднання



SYN = 0

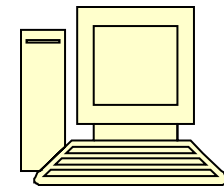
ACK = 1

Порядковий номер: 1025

Номер підтвердження: 3073

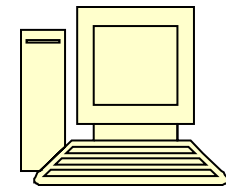
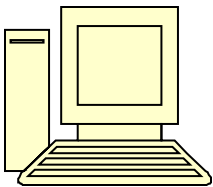
Відправляю байт № 1025

Отримав байт № 3072



Отримав байт № 1024  
Відправляю байт № 3072

# Встановлення з'єднання



SYN = 0

ACK = 1

Порядковий номер: 1025

Номер підтвердження: 3073

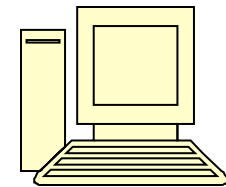
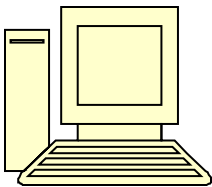
Відправляю байт № 1025

Отримав байт № 3072

Отримав байт № 1024

Відправляю байт № 3072

# Встановлення з'єднання



SYN = 0

ACK = 1

Порядковий номер: 1025

Номер підтвердження: 3073

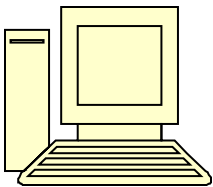
Відправляю байт № 1025

Отримав байт № 3072

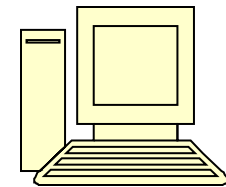
Отримав байт № 1025

Відправляю байт № 3073

# Встановлення з'єднання



З'єднання встановлено



Відправляю байт № 1025  
Отримав байт № 3072

Отримав байт № 1025  
Відправляю байт № 3073

# Розрив з'єднання

- З'єднання в ТСП дуплексное
  - Дані можуть передаватися в обидва боки
- Схема розриву з'єднання
  - Одночасне (обидві сторони розірвали з'єднання)
  - Одностороннє (сторона припиняє передавати дані, але може приймати)
- Прапор FIN - одностороннє закриття з'єднання
  - З'єднання закривається, коли обидві сторони відправляють сегмент з встановленим прапором FIN
- Прапор RST - розрив з'єднання через критичну ситуацію
  - Одночасний розрив з'єднання обома сторонами



# Параметри ТСР

- Параметри в заголовку ТСР є необов'язковими, але деякі з них використовуються часто
- Приклади параметрів:
  - Максимальний розмір сегмента (Maximum Segment Size, MSS)
  - Масштаб вікна - дозволяє збільшити розмір вікна до 1 ГБ, що ефективно для швидких каналів
  - Мітки часу
  - Вибіркове підтвердження (Selective Acknowledgment, SACK) - підтвердження діапазонів прийнятих байт

# Керування потоком в ТСП

- Керування потоком дозволяє регулювати швидкість передачі даних
  - Запобігання «затоплення» швидким відправником повільного одержувача
  - Мережа може бути вільна, але додаток не готовий отримати дані
- Транспортна підсистема працює з додатками:
  - Додаток не повинен забирати дані, як тільки вони з'явилися
  - Транспортна підсистема не зобов'язана передавати дані додатку або в мережу, як тільки вона їх отримала

# Керування потоком в ТСР

- Для керування потоком ТСР використовує механізм розсувного вікна
- Одержувач записує в поле заголовка ТСР «Розмір вікна» обсяг даних, які він готовий прийняти
- Розмір вікна може змінюватися динамічно
  - Процес читає дані з буфера швидко - вікно зростає
  - Процес читає повільно, буфер заповнений - вікно зменшується

# Керування потоком в ТСП

- Одержувач може встановити вікно нульового розміру
  - Передача даних припиняється незалежно від навантаження на мережу
- Продовження передачі:
  - Одержувач повторно відправляє підтвердження з ненульовим розміром вікна
  - Відправник надсилає запит «window probe» - прохання повторити підтвердження і розмір вікна

# Продуктивність TSP

- Деякі сервіси читають і пишуть дані маленькими порціями
- Емулятори терміналу telnet або ssh
  - При натисканні кожної клавіші дані передаються на сервер - 1 байт даних
  - Для передачі 1 байта даних потрібно передати IP-пакет довжиною 41 байт (20 байт заголовок IP, 20 байт заголовок TSP, 1 байт даних)
  - Високі накладні витрати

# Продуктивність ТСР

- Відкладені підтвердження
  - Затримка відправлення підтвердження до 500 мс в надії отримати дані
  - Термінал за 500 мс видає ехо, дані відправляються разом з підтвердженням
- Алгоритм Нагля (Nagle's algorithm)
  - Одержувачу відправляється тільки перша порція маленьких даних
  - Інші дані буферизується, поки не прийде підтвердження
  - Дані з буфера відправляються в одному сегменті
  - Триває накопичення даних в буфері, поки не прийде нове підтвердження

# Продуктивність ТСР

- Синдром «безглузлого вікна»
  - Програма читає дані з буфера по символах
  - Буфер заповнений, розмір вікна 0
  - Програма прочитала один байт - розмір вікна збільшився до 1
  - Відправник передав 1 байт даних (IP-пакет 41 байт)
  - Буфер заповнений, розмір вікна 0

# Таймери ТСП

- Таймер повторної передачі
  - Час очікування підтвердження отримання сегмента
  - Якщо підтвердження немає, сегмент відправляється знову
- Таймер наполегливості
  - Час, через яке відправляється запит «window probe»
- Таймер перевірки активності
  - Використовується при тривалому простої з'єднання
  - Задає час, через який повинна виконатися перевірка працездатності з'єднання
- Таймер закриття з'єднання
  - Задає час очікування рівний подвійному часу життя сегмента
  - За цей час всі сегменти з'єднання повинні залишити мережу



# Контроль перевантаження в ТСП

- Швидкість передачі даних по мережі визначається не тільки можливостями одержувача, але і навантаженням на мережу
- Механізми регулювання швидкості:
  - Вікно керування потоком
  - Вікно перевантаження

# Вікна перевантаження і керування потоком

- Вікно керування потоком:
  - Встановлюється одержувачем (поле «Розмір вікна» в заголовку TCP)
  - Розмір визначається можливостями додатку читати дані з буфера
- Вікно перевантаження:
  - Встановлюється відправником
  - Розмір визначається завантаженням мережі
- Розмір розсувного вікна визначається меншим з вікон перевантаження або керування потоком

# Вікно перевантаження

- Розмір вікна перевантаження визначається навантаженням на мережу
- Сигнал про перевантаження - втрата пакетів
  - Вважається, що пакети рідко губляться через помилки передачі
  - Якщо помилки в середовищі передачі даних зустрічаються часто, то це вирішується на канальному рівні (наприклад, Wi-Fi)
  - Пакети відкидаються маршрутизаторами при перевантаженнях

# Керування розміром вікна перевантаження

- Транспортна система відправника не знає, які мережеві з'єднання зустрінуться по дорозі до одержувача
- Як вибрати розмір вікна перевантаження?
  - Занадто маленький розмір призведе до низької швидкості передачі даних через постійне очікування підтверджень
  - Занадто великий розмір вікна призведе до низької швидкості передачі даних через перевантаження мережі

# Керування розміром вікна перевантаження

- ТСР використовує такі механізми визначення розміру вікна перевантаження:
  - Адитивне збільшення мультиплікативне зменшення (AIMD)
  - повільний старт
- ТСР починає роботу з повільного старту, потім переходить на AIMD

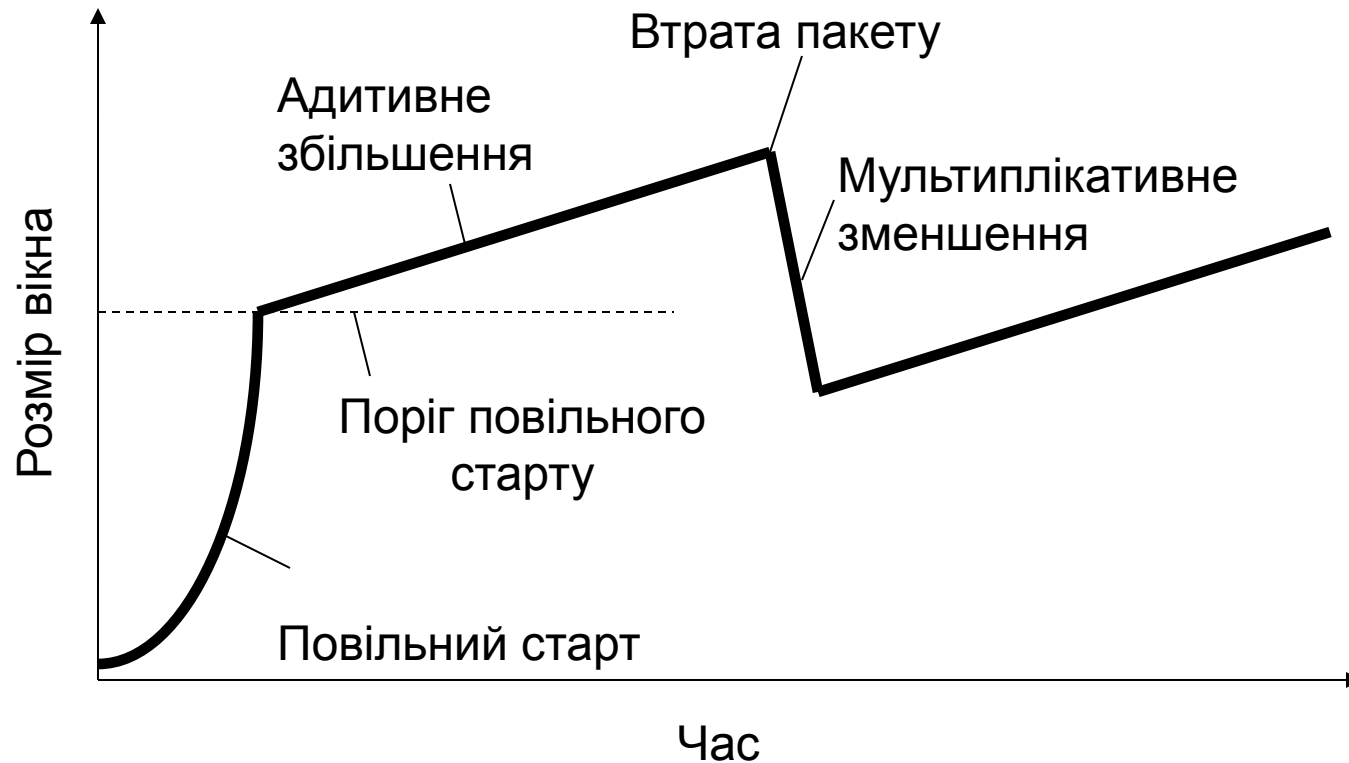
# Повільний старт

- Особливості повільного старту:
  - Спочатку розмір вікна перевантаження встановлюється маленьким (1 або 4 сегмента)
  - При кожному отриманні підтвердження відправляється в 2 рази більше сегментів
- Експоненціальне зростання розміру вікна (1 сегмент, 2, 4, 8 і т.д.)
- При втраті сегмента повільний старт починається заново

# Метод AIMD

- Особливості методу AIMD
  - При отриманні підтвердження розмір вікна перевантаження збільшується на 1 (адитивне збільшення)
  - При втраті сегмента розмір вікна перевантаження зменшується в 2 рази (мультиплікативне зменшення)
- Відмінності від повільного старту:
  - Розмір вікна росте повільніше
  - При втраті сегмента не потрібно починати все з початку

# Розмір вікна перевантаження TCP





# Визначення розміру розсувного вікна

- Додаток просить багато даних, але мережа перевантажена:
  - Вікно керування потоком: 40Кбайт
  - Вікно перевантаження: 20 Кбайт
  - Розсувне вікно: 20 Кбайт
- Мережа вільна, але додаток обмежує швидкість:
  - Вікно керування потоком: 20Кбайт
  - Вікно перевантаження: 40 Кбайт
  - Розсувне вікно: 20 Кбайт