

Терабітний Ethernet:

що це і навіщо



За прогнозами, терабіт з'явиться вже до кінця десятиліття, хоча, звісно, не в персональних комп'ютерах.

Торік виповнилось 40 років першому стандарту Ethernet. Відтоді технологія пішла далеко вперед, того ж таки 2020 року організація Ethernet Technology Consortium затвердила чергову ітерацію, яка забезпечує сукупну швидкість 800 Гбіт/с на порт, і вже обговорюється наступна, яка дасть усі 1,6 Тбіт/с.

Головною сферою застосування Ethernet майбутнього називають центри обробки даних, котрі мають справу з постійно зростаючими потоками трафіка. «СіБ» спробував розібратися, як працюють близькі до терабіту технології Ethernet і що вже можуть запропонувати виробники.

Технологія для операторів і дата-центрів

Terabit Ethernet — це взагалі умовна назва для всіх технологій, які забезпечують інтерфейс понад 1000 Гбіт/с. Наразі стандартизовано 400 Гбіт/с (400G Ethernet), який IEEE офіційно ратифікувала у грудні 2017 року. В останній версії дорожньої карти, яка опублікована на сайті

організації Ethernet Alliance, передбачається, що вже за кілька років (у середині 2020-х) будуть ухвалені стандарти для 800 Гбіт/с і 1,6 Тбіт/с

(рис. 1). За прогнозами Del'Oro Group, до 2024 року на 400GbE і ще більш швидкісні технології припадатиме чверть усіх проданих портів Ethernet.

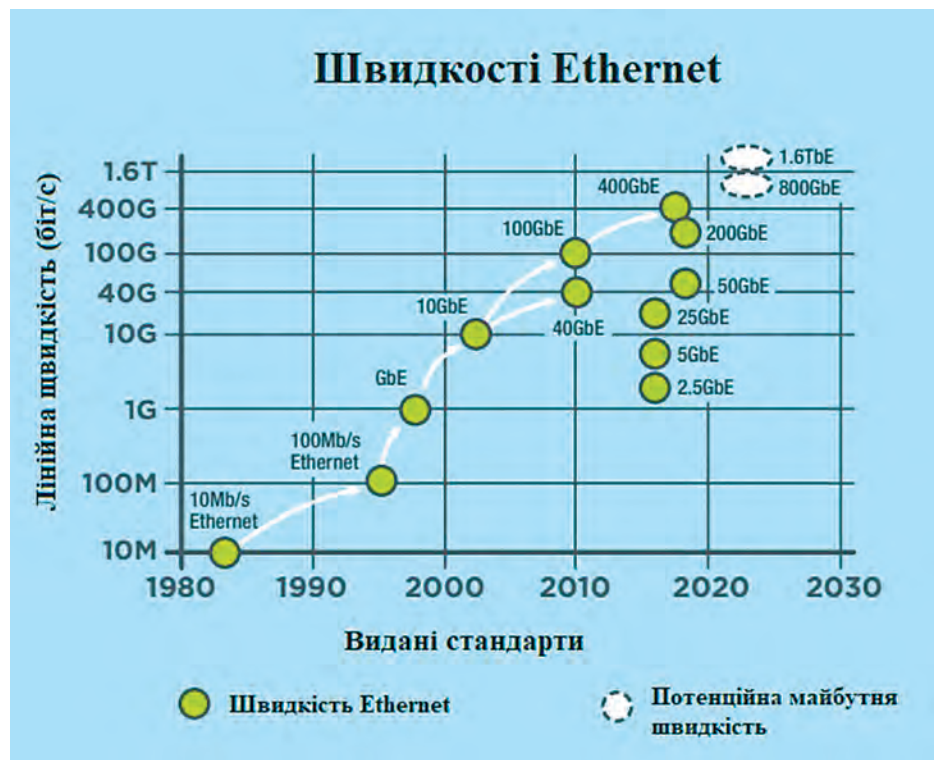


Рис. 1. Дорожня карта Ethernet Alliance

Кому можуть знадобитись швидкості понад 1000 Гбіт/с? Насамперед провайдерам хмарних послуг, значається у дорожній карті Ethernet Alliance. Вони були першими, хто ще 2010 року почав масово переходити на сервери з портами 10GbE для своїх гіпермасштабованих ЦОД. Відтоді через ненаситні апетити таких технологій, як штучний інтелект і машинне навчання, гіпермасштабовані сервери перейшли до 25GbE, а зараз рухаються до 50GbE. Сама унікальна мережева архітектура цих дата-центрів «складського розміру» стимулювала розвиток рішень класу 100GbE і вище, і найближчими роками їхня потреба у пропускній здатності зростатиме експоненційно, а отже, там не буде зайвим ні 800G, ані 1,6T.

В подібному напрямку рухаються оператори телекомунікацій, які вже десятки років використовують високошвидкісні рішення Ethernet на клієнтському боці оптичної транспортної мережі, а також для транспорту трафіка. Розгортання мереж 5G призводить до радикального збільшення трафіка між базовою станцією і ядром (backhaul), а також і всередині рознесеної БС між радіомодулями і централізованими блоками формування сигналів (fronthaul). Одночасно продовжує зростати глобальне споживання відеотрафіку.

Зараз основна маса поставок портів Ethernet припадає на сегмент корпоративних і кампусних мереж, звідки, власне, і «виросла» технологія. Щороку вони споживають мільйони портів, а за останні 15 років було прокладено понад 70 млн км LAN-мереж. Проте корпоративні ЦОД дуже ощадливі і зараз впроваджують сервери з 10GbE, а в найближчому майбутньому переходитимуть на 25GbE, загалом же на перспективу межа їхніх потреб — 400GbE.

Нарешті, найсвіжішою історією успіху Ethernet стало використання у мережах всередині автомобілів. На 2021 рік прогнозувалося постачання понад 500 млн портів для понад 100 млн автомобілів. Передбачається, що потреба у пропускній здатності для підключених

автівок може стати однією з рушійних сил, які спонукають розвиток Ethernet далі за 400GbE.

Чотириста і далі

Стандарт IEEE 802.3bs передбачає швидкість 400GbE і, до певної міри, є розвитком попередньої ітерації 100GbE, яка формувалася з 4 ліній (оптичних каналів чи волокон) по 25 Гбіт/с або 10 по 10 Гбіт/с. Для 400GbE прописані різні комбінації швидкостей, але очевидно, що схема 16×25 Гбіт/с потребуватиме надто багато волокон, і тому її передбачено лише на коротких відстанях. Для більшості ж варіантів за основу було взято технологію 50 Гбіт/с, на базі якої досягається сукупна швидкість 400 Гбіт/с за схемою 8×50 Гбіт/с.

Для підвищення лінійної швидкості було замінено схему фізичного кодування — замість коду без повернення до нуля (NRZ), який досі використовувався в Ethernet, перейшли на чотирирівневу імпульсно-амплітудну модуляцію — PAM-4. У ній сигнал має, відповідно, вдвічі більше рівнів напруги, і кожна посылка кодує два біти, а код Грея допомагає зменшити бітові помилки, викликані шумом. Таким чином фізична швидкість 25 Гбіт/с перетворюється у 50 Гбіт/с.

Специфікація 400GBASE-SR16 визначає передачу інформації по 16 багатомодових волокнах зі швидкістю 25 Гбіт/с з використанням коду NRZ на відстань до 100 м. Специфікації 400GBASE-FR16 і 400GBASE-LR16 описують передавання по одномодовому волокну за схемою 8×50 Гбіт/с з використанням WDM на відстань відповідно до 2 км і до 10 км. У 400GBASE-DR4 схема 4×100 Гбіт/с забезпечує ту саму швидкість по чотирьох волокнах, але лише до 500 м. Паралельно було стандартизовано 200GbE як перехідний етап у міграції на 400GbE.

Поступовий перехід на 400 Гбіт/с вже між тим розпочався, і головним застосуванням технології є ЦОД і сполучення між ними (Data Center Interconnect, DCI). Як вважає онлайн-видання **The Next Platform**,

цьому сприяють кілька чинників. По-перше, одна спеціалізована мікросхема (ASIC), що здатна керувати портами 400GbE, коштує суттєво дешевше, ніж ASIC для портів 100 і 200 Гбіт/с. У багатьох випадках один ASIC може замінити шість або більше мікросхем попереднього покоління і коштуватиме наполовину менше, аніж шість цих пристроїв. Це робить мережу рівня ЦОД доступнішою по ціні. По-друге, комутатор з портами 400 Гбіт/с можна перетворити на кореневий вузол, розділивши ці порти по 100 Гбіт/с, що спростить топологію мережі і дозволить обійтися меншою кількістю комутаторів для сполучення усіх пристроїв. Наприклад, на комутаторі з 32 інтерфейсами 400GbE можна організувати 128 інтерфейсів 100GbE. Для багатьох застосувань стогігабітної швидкості вистачить ще дуже надовго.

Що стосується DCI, то тут ключову роль зіграла поява когерентної технології 400ZR, яка забезпечує передавання даних зі швидкістю 400 Гбіт/с по одній хвилі на відстань до 120 км з використанням змінних модулів у формфакторі QSFP-DD. Останні дозволяють заводити оптику безпосередньо у комутатор або маршрутизатор Ethernet. 400GbE і 400ZR разом дають змогу позбутися пари транспондерів, ROADM і лінійних систем.

Cisco підрахувала, що відмова від DWDM на користь прямого передавання трафіку по IP між комутаторами дає суттєву економію. Умови передбачали, що мережа складається з 80 вузлів, з яких кожен безпосередньо сполучений з 40% інших, а трафік, що генерується вузлами, стартує з позначки 2,4 Тбіт/с і щороку зростає на 15%. Моделювання показало, що впродовж 5 років економія у капітальних витратах складе 35% завдяки меншій вартості 400GbE, відмові від транспондерів і спрощенню рівня DWDM. Скорочення операційних витрат на 57% є прямим наслідком того, що стає непотрібним коштовне і складне обслуговування DWDM-мережі. IP-мережа також споживатиме менше електроенергії, аніж DWDM. Загалом



Рис. 2. Комутатор Cisco 8101–32FH з 32 портами 400GbE

сукупна ж вартість володіння скоротиться на 46%.

У самої Cisco нову технологію підтримують магістральні комутатори серії 8000, які, залежно від моделі (одну з яких наведено на **рис. 2**), можуть мати від 12 до 36 таких портів. Модульні шасі 8800 можуть налічувати до 648 таких портів. Влітку цього року компанія додала 400GbE у NCS 540 — високощільний комутатор доступу операторського класу.

Також потужні комутатори для ядра ЦОД є у компанії **Arista Networks** — зокрема, модель 7816R3, яка займає об'єм у 31 RU, може мати до 575 портів 400 GbE.

Juniper ще наприкінці 2019 року продемонстрував роботу 400GbE на рекордній відстані 1300 миль (майже 2100 км), сполучивши міста Денвер і Чикаго за допомогою магістрального маршрутизатора PTX10003. Ця платформа постачається в двох конфігураціях і може мати 16 або 32 400-гігабітних порти.

У **NVIDIA** в сімейство комутаторів Spectrum входять моделі з 32 портами 400GbE. Ці пристрої побудовані на основі власного чіпа, мають загальну пропускну здатність до 12,8 Тбіт/с і призначені для багатоорендних хмар та програмно-визначуваних систем зберігання даних.

Навіщо нам терабіт

Між тим у квітні минулого року організація Ethernet Technology Consortium оголосила про фіналізацію специфікації 800GbE (800BASE-R). Документ наново визначає рівень управління доступом до середовища передавання (MAC) і фізичний підрівень кодування (PCS), проте його основною метою було максимальне пристосування існуючої логіки 400GbE

для передавання вдвічі більшої швидкості, а саме восьми потоків 106 Гбіт/с при забезпеченні сумісності з існуючими багатшвидкісними Ethernet-портами. Фактично 800-гігабітний інтерфейс утворюється з двох з'єднань по 400GbE (The Next Platform підозрює, що таким самим методом зрештою дійде до 1,6 і 3,2 Тбіт/с).

Тим часом група дослідників з американської компанії **Lumentum** створила альтернативне рішення, у якому потік 800 GbE формується з чотирьох ліній по 200 Гбіт/с. У рішенні, представленому у червні цього року, використовуються лазери, здатні передавати дані зі швидкістю 224 Гбіт/с на відстань у 2 км. (Це, своєю чергою, дозволить так само з 8 ліній влаштувати інтерфейс 1,6 Тбіт/с).

То навіщо нам 800 Гбіт/с і більше? У дослідженні групи IEEE 802.3, присвяченому обґрунтуванню необхідності терабітного Ethernet, зазначається, що 800GbE навіть близько не задовольнить поточні потреби у пропускну здатності, ба навіть і 1,6ТбЕ буде замало.

Як ідеться у матеріалі компанії **Xena Networks**, яка спеціалізується на тестуванні Ethernet, причина — зростання обсягів відео, на яке у 2017 році припадало 55% усього Інтернет-трафіка, а в 2022-му буде вже 70%. Через цю тенденцію провайдери посувають контент ближче до споживача, що призводить до появи великої кількості дрібних дата-центрів на краю мережі. А це, своєю чергою, збільшує вимоги до пропускну здатності мереж міського рівня (MAN).

З третього боку, зростають і вимоги до пропускну здатності в самих дата-центрах, оскільки за периметр мережі виноситься лише найбільш

популярний контент, решта ж і надалі подається з гіпермасштабованих ЦОД. А потокове відео потребує не лише швидкостей, але й пам'яті, тим більше з переходом до важчих форматів (4K, 8K). За обрахунками Cisco, на які посилається звіт Xena, з 2017 року по 2022-й обсяг трафіка, що припадає на HD-відео, зростає у п'ятеро. Відповідно мають збільшитись ресурси для зберігання цього відео, а також і канали зв'язку між серверами в дата-центрі. Результатом буде щільна Mesh-структура, яка й потребуватиме терабітних з'єднань.

Проблемою, однак, є те, що зі зростанням швидкості передачі збільшується і енергоспоживання. За оцінкою Марка Файлера, головного інженера апаратної архітектури **Microsoft Azure**, комутатори з 400GbE потребують утричі більше електроенергії, ніж 100GbE. Через цей стримуючий фактор не вдасться просто лінійно нарощувати пропускну здатність інтерфейсів, і ймовірно, що 100GbE на найближчу перспективу залишатимуться будівельними блоками для енергоощадного дата-центру. Майбутні ж мережі потребуватимуть інновацій у фотоніці, оптимізації архітектури і досконалішої апаратної реалізації.

Таким чином, наступні покоління Ethernet вже потрібні, але на шляху їх появи стоять серйозні труднощі. Як зазначає у своєму звіті на сесії IEEE Джон Д'Амброзія з **Futurewti** (американської структури **Huawei**), 1,6ТбЕ треба займатися вже зараз, щоб не спізнитися (за оцінками, специфікація буде готова лише за 8 років після 800GbE).

Між тим просування в напрямку терабіта вже почалося. У березні 2021 року Cisco представила процесор Silicon One G100, виконаний за техпроцесом 7 нм. Він забезпечує загальну продуктивність 25,6 Тбіт/с і має інтерфейс, що конфігурується у тому числі як 800GbE і 1,6ТбЕ.

Але в будь-якому разі, як було сказано, терабітний Ethernet ми побачимо не раніше кінця десятиліття.

Василь ТКАЧЕНКО, Сиб