6. Протоколи міждоменної маршрутизації

"Протоколи міждоменної маршрутизації" вводиться поняття автономної системи і дається їх характеристика. Наведені основні відомості про протокол BGPv4 та приводиться короткий опис технології безкласової міждоменної маршрутизації (CIDR).

5.1 Технологія безкласової міждоменної маршрутизації CIDR

За останні кілька років у мережі Internet багато чого змінилося: різко зросло число вузлів і мереж, підвищилася інтенсивність трафіка, змінився характер переданих даних. Через недосконалість протоколів маршрутизації обмін повідомленнями про відновлення таблиць став іноді приводити до збоїв магістральних маршрутизаторів через перевантаження при обробці великого обсягу службової інформації. Так, в 1994 році таблиці магістральних маршрутизаторів в Internet містили до 70 000 маршрутів.

На рішення цієї проблеми була спрямована, зокрема, і технологія *безкласової міждоменної маршрутизації (Classless Inter-Domain Routing, CIDR),* уперше про яку було офіційно оголошене в 1993 році, коли були опубліковані RFC 1517, RFC 1518, RFC 1519 і RFC 1520.

Суть технології CIDR полягає в наступному. Кожному постачальнику послуг Internet повинен призначатися неперервний діапазон у просторі IP-адрес. При такому підході адреси всіх мереж кожного постачальника послуг мають загальну старшу частину – *префікс,* тому маршрутизація на магістралях Internet може здійснюватися на основі префіксів, а не повних адрес мереж. Агрегування адрес дозволить зменшити обсяг таблиць у маршрутизаторах всіх рівнів, а отже, прискорити роботу маршрутизаторів і підвищити пропускну здатність Internet.

Розподіл IP-адреси на номер мережі й номер вузла в технології CIDR відбувається не на основі декількох старших біт, що визначають клас мережі (***A***, ***B*** або ***C***), а на основі маски змінної довжини, що призначається постачальником послуг. На рис. 5.1 показаний приклад деякого простору IP-адрес, що є в розпорядженні гіпотетичного постачальника послуг. Всі адреси мають загальну частину в *k* старших розрядах – префікс. *N* розрядів, що залишилися, використовуються для доповнення незмінного префікса змінною частиною адреси. Діапазон наявних адрес у такому випадку становить 2^n.

Коли споживач послуг звертається до постачальника послуг з деякої кількості адрес, то в наявній пулі адрес «вирізьблюється» безперервна область S1, S2, S3 або S4 відповідні розміри. Причому границі цієї області вибираються такими, щоб для нумерації необхідного числа вузлів вистачило деякого числа молодших розрядів, а значення всіх, що залишилися (старших) розрядів було однаковим у всіх адрес даного діапазону. Таким умовам можуть задовольняти тільки області, розмір яких кратний ступеня двійки. А границі ділянки, що виділяється повинні бути кратні необхідному розміру.

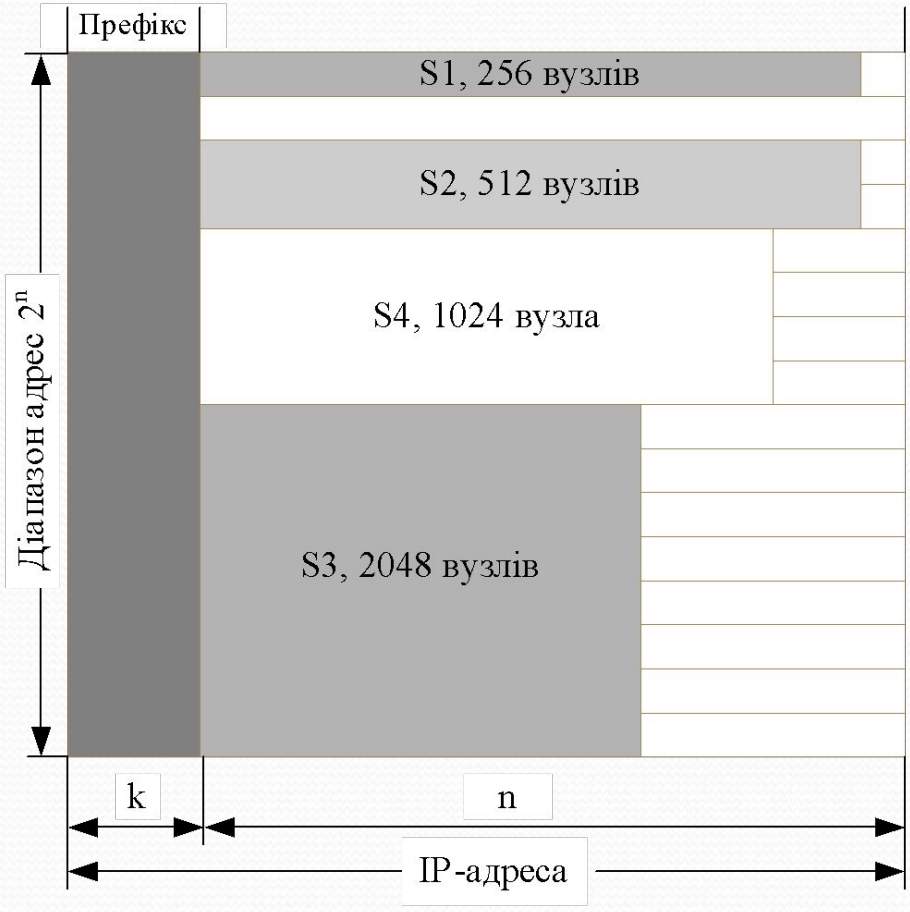

Рис. 5.1. Технологія CIDR

Розглянемо приклад. Нехай постачальник послуг Internet має у своєму розпорядженні пул адрес у діапазоні 193.20.0.0-193.23.255.255 (11000001. 00010100. 00000000. 00000000 – 11000001. 00010111. 11111111. 11111111) із загальним префіксом 193.20 (11000001. 000101) і маскою, що відповідає цьому префіксу 255.252.0.0.

Якщо абоненту цього постачальника послуг потрібно зовсім небагато адрес, наприклад 13, то постачальник міг би запропонувати йому різні варіанти: мережа 193.20.30.0, мережа 193.20.30.16 або мережа 193.21.204.48, усе з тим самим значенням маски 255.255.255.240. У всіх випадках у розпорядженні абонента для нумерації вузлів є 4 молодших біти.

Розглянемо інший варіант, коли до постачальника послуг звернувся великий замовник, який, можливо, збирається надавати послуги з доступу в Internet. Йому потрібен блок адрес в 4000 вузлів. У цьому випадку постачальник послуг міг би запропонувати йому, наприклад, діапазон адрес 193.22.160.0 –193.22.175.255 з маскою 255.255.240.0. Агрегований номер мережі (префікс) у цьому випадку буде дорівнює 193.22.160.0.

Адміністратор маршрутизатора М2 (рис. 5.2) помістить у таблицю маршрутизації тільки по одному запису на кожного клієнта, якому був виділений пул адрес, незалежно від кількості підмереж, організованих клієнтом. Якщо клієнт, що одержав мережу 193.22.160.0, через якийсь час розділить її адресний простір в 4096 адрес на 8 підмереж, то в маршрутизаторі М2 первісна інформація про виділену йому мережу не зміниться.

Для постачальника послуг верхнього рівня, що підтримує клієнтів через маршрутизатор М1, зусилля постачальника послуг нижнього рівня по поділу його адресного простору також не будуть помітні. Запис 193.20.0.0 з маскою 255.252.0.0 повністю описує мережі постачальника послуг нижнього рівня в маршрутизаторі М1.

Отже, впровадження технології CIDR дозволяє вирішити два основні завдання:

• Більш ощадлива витрата адресного простору. Дійсно, одержуючи у своє розпорядження адресу мережі, наприклад, класу ***C***, деякі організації не використовують весь можливий діапазон адрес оскільки в їхній мережі є набагато менше 255 вузлів. Технологія CIDR відмовляється від традиційної концепції поділу адрес протоколу IP на класи, що дозволяє одержувати в користування стільки адрес, скільки реально необхідно. Завдяки технології CIDR постачальники послуг одержують можливість «нарізати» блоки з виділеного їм адресного простору в точній відповідності з вимогами кожного клієнта, при цьому у клієнта залишається простір для маневру на випадок його майбутнього росту.

• Зменшення числа записів у таблицях маршрутизаторів за рахунок об'єднання маршрутів – один запис у таблиці маршрутизації може представляти велику кількість мереж. Дійсно, для всіх мереж, номера яких починаються з однакової послідовності цифр, у таблиці маршрутизації може бути передбачений один запис (див. рис. 5.2). Так, маршрутизатор М2 установлений в організації, що використовує техніку CIDR для виділення адрес своїм клієнтам, повинен підтримувати у своїй таблиці маршрутизації всі 8 записів про мережі клієнтів. А маршрутизатору М1 досить мати один запис про всі ці мережі, на підставі якої він передає пакети із префіксом 193.20 маршрутизатору М2, що їх і розподіляє по потрібних портах.

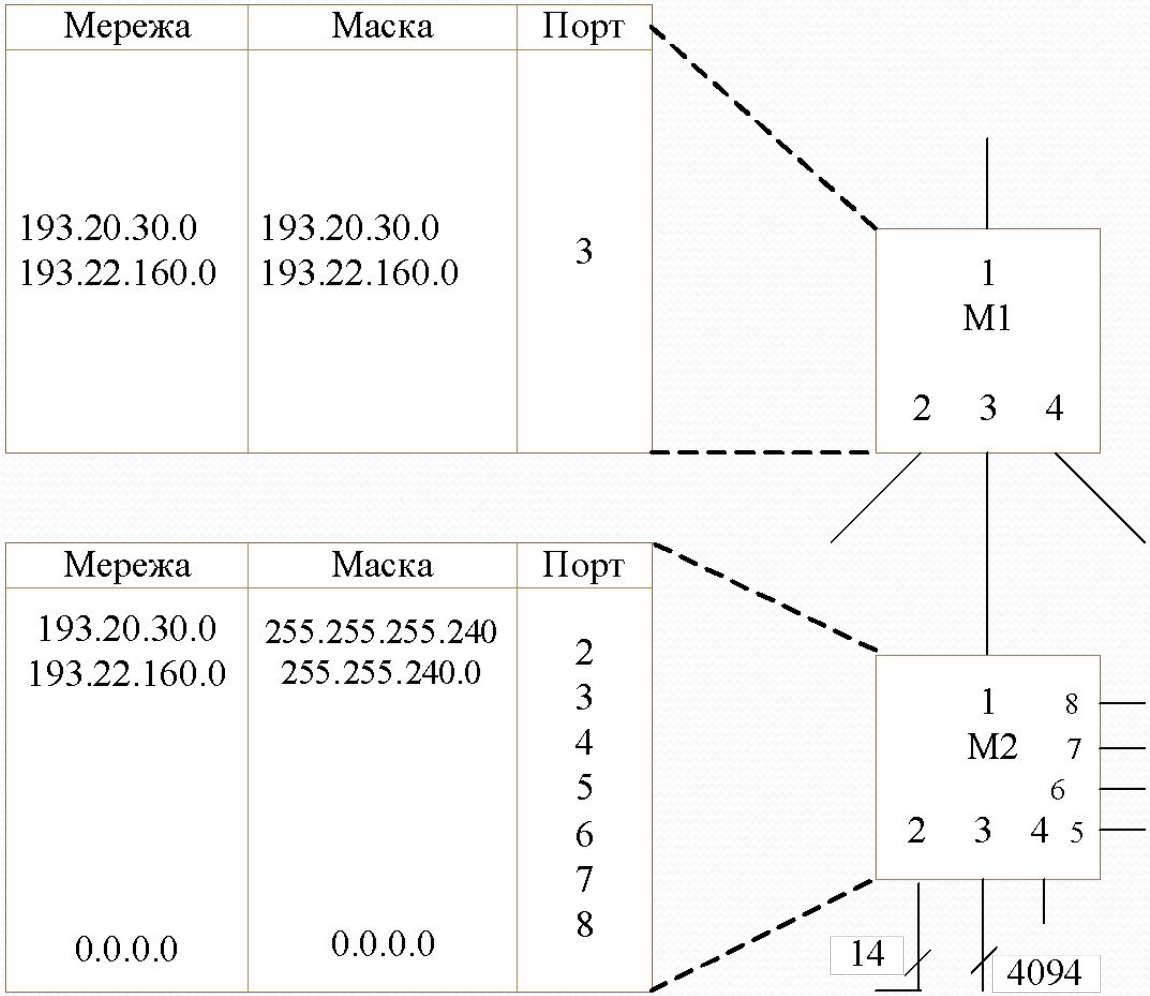

Рис. 5.2. Виграш у кількості записів у маршрутизаторі при використанні технології CIDR

Якщо всі постачальники послуг Internet будуть дотримуватися стратегії CIDR, то особливо помітний виграш буде досягатися в магістральних маршрутизаторах.

Технологія CIDR уже успішно використовується в IPv4,6 і підтримується такими протоколами маршрутизації, як OSPF, RIP-2, BGP4. Технологія CIDR підтримується магістральними маршрутизаторами Internet, а не звичайними хостами в локальних мережах.

Використання CIDR у мережах IP у загальному випадку вимагає перенумерації мереж. Оскільки ця процедура вимагає певних часових й матеріальних витрат, то для її проведення користувачів потрібно яким-небудь образом стимулювати. В якості таких стимулів розглядається, наприклад, введення оплати за рядок у таблиці маршрутизації або ж за кількість вузлів у мережі. При використанні класів мереж абонент часто не повністю займає весь припустимий діапазон адрес вузлів – 254 адреси для мережі класу ***C*** або 65534 адреси для мережі класу ***B***. Частина адрес вузлів звичайно пропадає. Вимога оплати кожної адреси вузла допоможе користувачеві зважитися на перенумерацію, щоб одержати рівно стільки адрес, скільки йому потрібно.

5.2 Автономні системи

Зовнішні протоколи маршрутизації були розроблені для керування таблицями маршрутів, що розрослися, і для підвищення структурованості мережі Internet шляхом поділу доменів маршрутизації на різні адміністративні одиниці, які називаються автономними системами (autonomous systems – AS), мають власні незалежні правила маршрутизації й використовують унікальні внутрішні протоколи IGP .

На початковому етапі розвитку в мережі Internet використовувався протокол зовнішнього шлюзу EGP8 (Exterior Gateway Protocol) (не плутати з узагальнюючою назвою протоколів зовнішнього шлюзу!). Так, у мережі NSFNET цей протокол використовувався для обміну інформацією про взаємну досяжність між магістральною й регіональною мережами. Хоча протокол EGP і застосовувався дуже широко, його обмеження по топології, неефективність у розпізнаванні петель маршрутизації та при встановленні правил маршрутизації породили потребу в новому універсальному протоколі, позбавленому цих недоліків. У цей час стандартом де-факто для організації міждоменної маршрутизації в мережі Internet є протокол BGPv4

5.2.1 Поняття автономної системи

Автономна система – це набір маршрутизаторів, що мають єдині правила маршрутизації, керовані однією технічною адміністрацією й працюючі на одному із протоколів IGP. Для всієї іншої мережі AS є кінцевим найпростішим елементом і сприймається як єдине ціле. Реєстром мережі Internet або провайдером кожної AS призначається унікальний ідентифікаційний номер. Маршрутизація між різними AS здійснюється за допомогою протоколу зовнішнього шлюзу, такого як BGPv4.

Розберемося, які переваги має розбивка великої мережі на адміністративні ділянки (з обліком того, що мережа Internet завдяки використанню протоколів OSPF або IS-IS могла б бути однією складною мережею). Більш дрібні мережі, що представляються в якості AS, здатні реалізовувати власні правила маршрутизації, які б унікально їх характеризували й описували всі послуги, надавані іншими мережами. Тепер у кожної AS можна запускати свій пакет протоколів IGP, незалежно від того, які набори IGP запущені на інших AS.

5.2.2 Автономні системи з заглушками

Автономна система вважається системою з заглушкою (stub) за умови, що всі маршрути з неї до інших мереж проходять через одну точку. Ці AS також називають одноканальними (single-homed) стосовно провайдерів. На рис. 5.4 наведений приклад одноканальної AS, або AS із заглушкою.

Одноканальним AS не потрібно одержувати інформацію про всі маршрути в Internet від провайдера. Через те, що для таких AS є всього один вихід у зовнішній світ, весь трафік може за замовчуванням відправлятися провайдеру. Маючи AS такої конфігурації, провайдер може використовувати різні способи для оголошення іншим провайдерам маршрутів у клієнтську мережу.

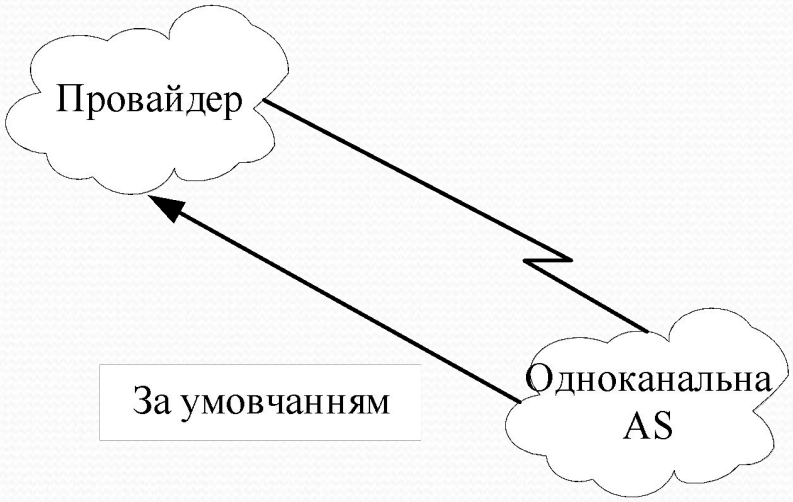

Рис. 5.4. Одноканальна AS (AS із заглушкою)

Один з таких способів – зберігання записів про статичні маршрути в підмережі клієнтів на маршрутизаторі провайдеру. Потім провайдер може за допомогою BGP оголосити всі ці статичні маршрути по мережі Internet. Цей метод забезпечує гарну масштабованість, якщо маршрути в клієнтські мережі представлені невеликими наборами об'єднаних маршрутів. Коли клієнтам виділяється занадто велике число підмереж з адресами, які випливають в різнобій, той їхній зміст у списку на маршрутизаторі стає неефективним. Найкраще, якщо клієнтам виділені неперервні блоки IP-адрес. Цим досягається найкраще об'єднання маршрутів.

Як альтернатива для оголошення маршрутів до мереж своїх клієнтів провайдер може використовувати протоколи IGP. Ці протоколи можуть використовуватися між клієнтом і провайдером для розсилання маршрутної інформації. Така схема надає переваги динамічної маршрутизації, коли будь-які зміни й інша мережна інформація динамічно посилаються на вузол провайдеру. Однак цей метод не знайшов широкого застосування через низьку масштабованість і через те, що нестабільність каналу робить нестабільною роботу самих протоколів IGP.

І третій спосіб, за допомогою якого провайдер може одержувати інформацію про маршрути й повідомляти про маршрути до клієнтських мереж, – застосування протоколу BGP між клієнтом і провайдером. При роботі в AS із заглушкою досить складно зареєструвати номер AS у реєстрі Internet, тому що правила маршрутизації в клієнтських мережах є по суті доповненням до правил маршрутизації провайдеру.

Провайдер також може привласнити клієнтської AS номер з діапазону номерів для приватних AS (65412-65535), допускаючи, що правила маршрутизації провайдеру забезпечують підтримку роботи приватного простору AS у своїх клієнтів, як це описано в RFC 227010.

Для забезпечення роботи між провайдером і клієнтом можна використовувати кілька комбінацій протоколів. На рис. 5.5 представлені можливі конфігурації протоколів між клієнтом і провайдером – як приклад використана AS із заглушкою.

Як видно з рис. 5.5, провайдери можуть переносити маршрутизатори клієнта у свої точки присутності або свої маршрутизатори – на технічні площадки клієнтів. Зверніть увагу, що не у всіх випадках потрібно, щоб клієнт із провайдером працював по протоколу BGP.

В принципі, багатоканальні AS без транзиту не мають потреби в організації роботи по протоколу BGP зі своїми провайдерами, хоча це й рекомендується, а в більшості випадків і потрібно провайдерам. Робота з провайдерами по протоколу BGPv4 має безліч переваг і для контролю за поширенням маршрутів, і для фільтрації.

5.2.3 Багатоканальні транзитні автономні системи

Багатоканальні транзитні AS також мають кілька з'єднань із зовнішнім світом і можуть використовуватися для транзиту трафіка в інтересах інших AS (рис. 5.5). Транзитним (стосовно багатоканальної AS) є будь-який трафік, відправник і одержувач якого не належать до локальної AS.

Хоча протокол BGP-4 є протоколом зовнішнього шлюзу, він може використовуватися й усередині AS для обміну оновлень маршрутів на основі протоколу BGP. З'єднання між маршрутизаторами на базі протоколу BGP усередині автономних систем відносяться до внутрішнього BGP (Internal BGP – IBGP), у той час як з'єднання між маршрутизаторами різних автономних систем відносять до зовнішнього BGP (External BGP – EBGP). Маршрутизатори, що працюють на базі IBGP, називають також транзитними маршрутизаторами. Вони займаються пересиланням транзитного трафіка, що надходить в AS.

Транзитні AS повідомляють і маршрути, отримані ними від інших AS. Таким чином, транзитна AS буде відкрита для трафіка, що адресований в іншу AS. У багатоканальних транзитних AS рекомендується використовувати протокол BGP-4 для забезпечення з'єднань із іншими AS і захищати внутрішні нетранзитні маршрутизатори від одержання маршрутів з мережі Internet. Необов'язково, щоб всі маршрутизатори в домені працювали по протоколу BGP. Внутрішні маршрутизатори без транзиту цілком можуть обійтися маршрутами за замовчуванням на маршрутизатори з BGP, що зменшує кількість маршрутів на маршрутизатори без транзиту. У мережах найбільш великих сервіс-провайдерів всі маршрутизатори звичайно обробляють весь набір BGP маршрутів самостійно.

На рис. 5.6 представлена багатоканальна транзитна автономна система AS1, підключена до двох різних провайдерів, ISP1 і ISP2. Автономна система AS1 одержує відомості про маршрути n3, n4, n5 і n6 і від ISP1, і від ISP2, а потім, додавши свої локальні маршрути, ділиться цією інформацією з провайдерами ISP1 і ISP2. У цьому випадку провайдер ISP1 може використовувати AS1 у якості транзитної AS для того, щоб відправити трафік у мережі n5 і n6, а провайдер ISP2 може використовувати AS 1 для того, щоб досягти мереж n3 і n4.

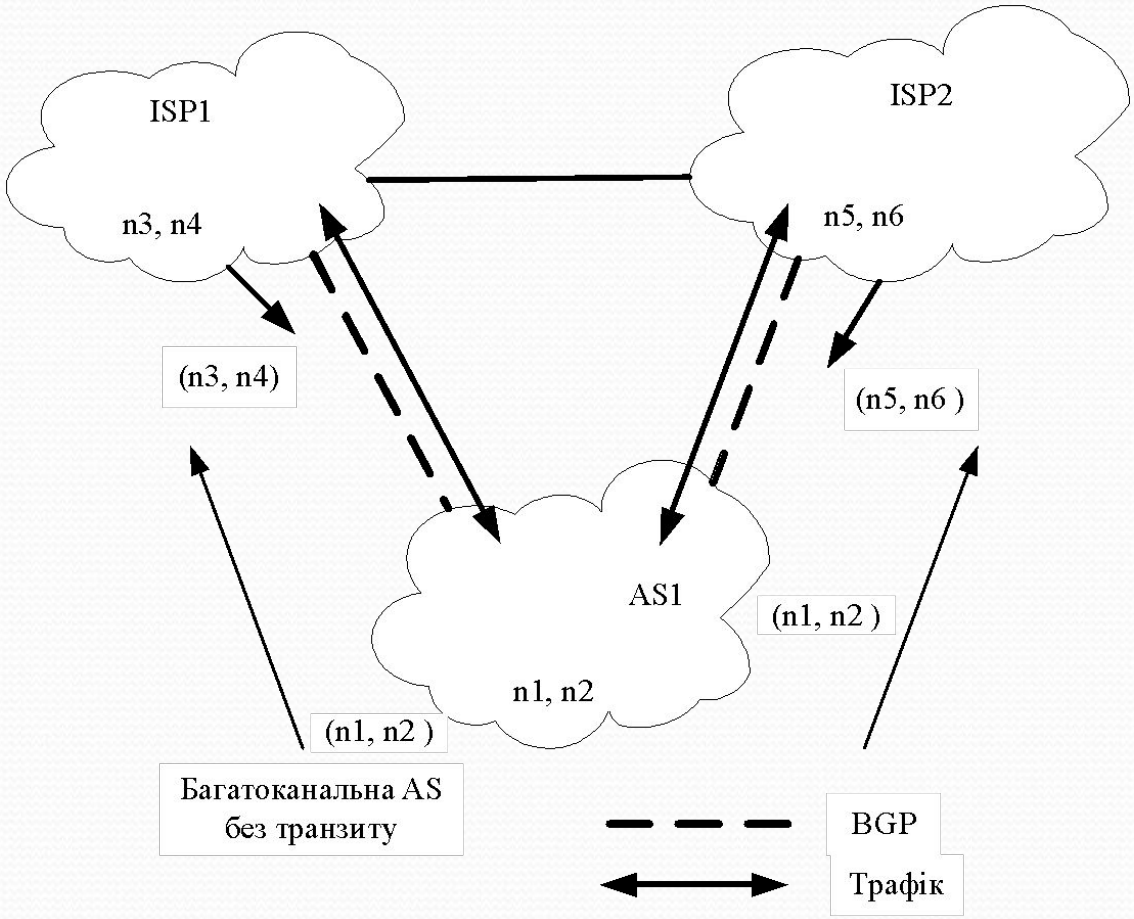

Рис. 5.6. Приклад багатоканальної AS без транзиту

5.3 Протокол BGP

Протокол граничного шлюзу (Border Gateway Protocol – BGP) являє собою протокол маршрутизації, що використовується при передачі даних між автономними системами. Протокол BGP використовується для обміну маршрутною інформацією в мережі Internet і є протоколом, використовуваним між провайдерами послуг Internet (Internet Service provider – ISP). У мережах користувачів, таких як університети й корпорації, для обміну маршрутною інформацією між мережами звичайно застосовуються протоколи внутрішнього шлюзу (Interior Gateway Protocol – IGP), таких як RIP або OSPF. Користувачі підключаються до ISP-Провайдерів, а останні використовують BGP для обміну маршрутною інформацією між користувачем і провайдером ISP. Коли протокол BGP використовується для обміну між автономними системами, він називається зовнішнім BGP (External BGP – EBGP). Якщо провайдер служб Internet використовує протокол BGP для обміну маршрутами усередині автономної системи AS, то цей протокол називається внутрішнім (Interior BGP – IBGP). Це розходження проілюстроване на рис. 5.7.

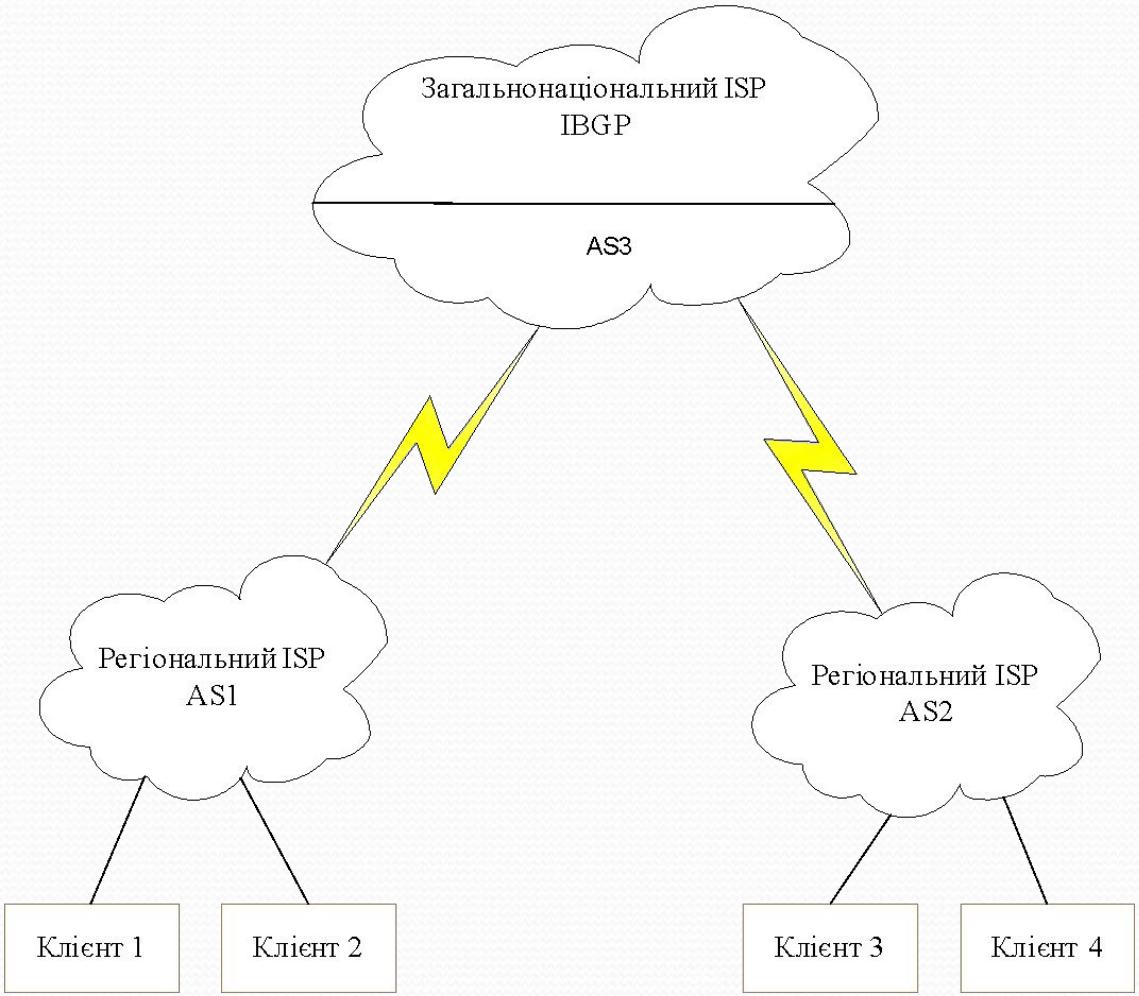

Рис. 5.7. Зовнішній і внутрішній протоколи BGP

5.3.1 Атрибути протоколу

Маршрути, отримані з використанням протоколу BGP, мають деякі властивості, які використовуються для визначення найкращого маршруту в тих випадках, коли є кілька маршрутів до пункту призначення [33]. Ці властивості називаються атрибутами протоколу BGP, і розуміння їхнього впливу на вибір маршруту необхідно для розробки стійкої мережі. У даному розділі описуються наступні атрибути, які BGP використовує при виборі маршруту:

- Weight;
- Local preference;
- Multi-exit discriminator;
- Origin;
- AS_ path;
- Next-hop;
- Community.

- **Атрибут Weight.** Атрибут *Weight (вага)* являє собою атрибут, уведений корпорацією Cisco і є локальним для конкретного маршрутизатора. Він не анонсується сусіднім маршрутизаторам. Якщо маршрутизатор виявляє кілька маршрутів до пункту призначення, то вибирається маршрут з найбільшою вагою.

- **Атрибут Local Preference.** Атрибут *Local Preference* використовується для вибору точки виходу з локальної автономної системи. На відміну від атрибута Weight, атрибут Local Preference анонсується у всій локальній автономній системі. Якщо в автономній системи є кілька точок виходу, то цей атрибут використовується при виборі точки виходу для певного маршруту.

- **Атрибут Multi-exit Discriminator.** Атрибут *Multi-Exit Discriminator* (MED), також називаний *атрибутом метрики (metric attribute),* використовується як пропозиція зовнішній автономній системі AS вибрати маршрут до AS, що анонсує дану метрику. Термін *"пропозиція"* застосовується тому, що зовнішня AS, що одержує атрибут MED, може використовувати для вибору маршруту інші атрибути протоколу BGP.

- **Атрибут Origin.** Атрибут *Origin* указує, яким способом протокол BGP довідається про конкретний маршрут. Цей атрибут може приймати одне з наступних трьох значень.

• **IGP.** Маршрут є внутрішнім стосовно вихідної автономної системи AS. Це значення встановлюється в тих випадках, коли для впровадження маршруту до протоколу BGP використовується команда конфігурування мережного маршрутизатора.

• **EGP.** Про маршрут повідомляється по протоколу зовнішнього граничного шлюзу (Exterior Border Gateway Protocol – EBGP).

• **Incomplete (неповний).** Джерело маршруту невідоме або про нього повідомляється яким-небудь іншим способом. Атрибут приймає це значення, коли маршрут перерозподіляється до протоколу BGP.

- **Атрибут AS_path.** Коли оголошення маршруту проходить через автономну систему, її номер заноситься в упорядкований список номерів автономних систем AS, який пройшов цим маршрутним оголошенням.

- **Атрибут Next-Hop.** Атрибут протоколу EBGP *Next-Hop (вузол наступного переходу)* являє собою IP-адресу, що використовується для досягнення маршрутизатора, який анонсує маршрут. Для однорангових пристроїв протоколу EBGP адресою вузла наступного переходу є IP-адреса з'єднання між одноранговими вузлами. У протоколі IBGP адреса вузла наступного переходу протоколу EBGP передається в локальну автономну систему AS.

- **Атрибут Community.** Цей атрибут забезпечує спосіб групової адресації одержувачів, що називається співтовариством (community), до якого можуть відноситися рішення про вибір маршруту (такі, як прийняття, перевага й перерозподіл). Для установки даного атрибуту використовуються перетворення маршрутів. Нижче перераховані стандартні значення атрибута Community.

- no-export (не експортується). Такий маршрут не анонсується одноранговим вузлам протоколу EBGP.

- no-advertise (не анонсується). Цей маршрут не анонсується ніяким одноранговим вузлам.

- Internet. Про цей маршрут сповіщається співтовариство Internet; до цього співтовариства належать всі маршрутизатори мережі.

Вибір маршруту по протоколу BGP

Протокол BGP може одержати повідомлення про один і той самий маршрут з декількох джерел, однак вибирає тільки один, найкращий. BGP записує обраний маршрут в таблицю IP-маршрутизації й поширює його серед своїх сусідніх маршрутизаторів.

Для вибору маршруту до одержувача протокол BGP використовує наведені нижче критерії в зазначеному порядку.

- Якщо вузол наступного переходу недоступний, то повідомлення про відновлення маршруту відкидається.
- Перевага віддається маршруту з найбільшою вагою.
- Якщо ваги однакові, то перевага віддається шляху з найбільшим значенням атрибута Local Preference.
- Якщо значення атрибутів Local Preference однакові, то перевага віддається шляху, що був ініційований процесом протоколу BGP, що виконується на цьому маршрутизаторі.
- Якщо жоден маршрут не був ініційований протоколом BGP, що виконується на даному маршрутизаторі, то перевага віддається маршруту із самим коротким атрибутом AS_path.
- Якщо всі маршрути мають однакову довжину атрибута AS_path, то перевага віддається маршруту з найнижчим значенням типу джерела (вважається, що IGP більше низький у порівнянні з EGP, що, у свою чергу, нижче, ніж неповне джерело).
- При однакових типах джерела перевага віддається маршруту з найменшим значенням атрибута MED.
- При рівних значеннях атрибута MED перевага віддається зовнішньому маршруту (у порівнянні із внутрішнім).
- Якщо й ці характеристики збігаються, то перевага віддається маршруту через найближчий сусідній IGP-пристрій.
- Кращим є маршрут з найменшою IP-адресою, що визначається ідентифікатором (ID) BGP-маршрутизатора.

5.3.2 Формат заголовка повідомлення протоколу BGP

Формат заголовка повідомлення в BGP складається з поля маркера довжиною 16 байт, поля довжини (2 байти) і поля типу (1 байт). На рис. 5.11 представлений формат заголовка повідомлення протоколу BGP.

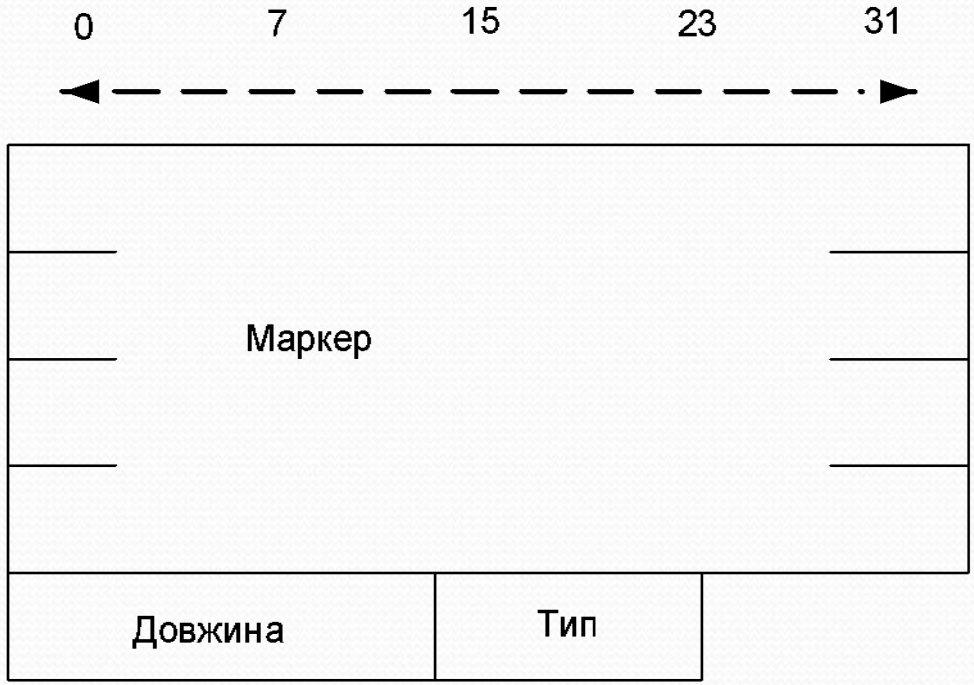

Рис. 5.11. Формат заголовка повідомлення BGP

Залежно від типу повідомлення в повідомленні протоколу BGP за заголовком може випливати або не випливати блок даних. Так, наприклад, повідомлення KEEPALIVE складаються тільки із заголовка й ніяких даних не передають.

Поле маркера довжиною 16 байт використовується для автентифікації вхідних повідомлень BGP або для детектування втрати синхронізації між двома взаємодіючими по BGP маршрутизаторами. Поле маркера буває двох форматів.

- Якщо послане повідомлення типу OPEN або в ньому відсутня інформація про автентифікацію, то в поле маркера на всі позиції виставляються «1».
- В іншому випадку значення поля маркера обчислюється відповідно до використаного механізму автентифікації.

Поле довжини розміром 2 байти використовується для відображення повної довжини повідомлення BGP, включаючи заголовок. Найменша довжина повідомлення BGP становить 19 байт (16+2+1), а найбільша – 4096 байт.

Поле типу розміром 1 байт визначає тип повідомлення. Можливі наступні значення:

- OPEN (відкриття з'єднання);
- UPDATE (відновлення маршрутної інформації);
- NOTIFICATION (повідомлення про помилку);
- KEEPALIVE (перевірка стана з'єднання).

Нижче призначення й формат кожного типу повідомлень будуть розглянуті більш детально.

5.3.3 Повідомлення OPEN

На рис. 5.12 представлений формат повідомлення OPEN.

Рис. 5.12. Формат повідомлення OPEN

5.3.4 Модель кінцевих станів

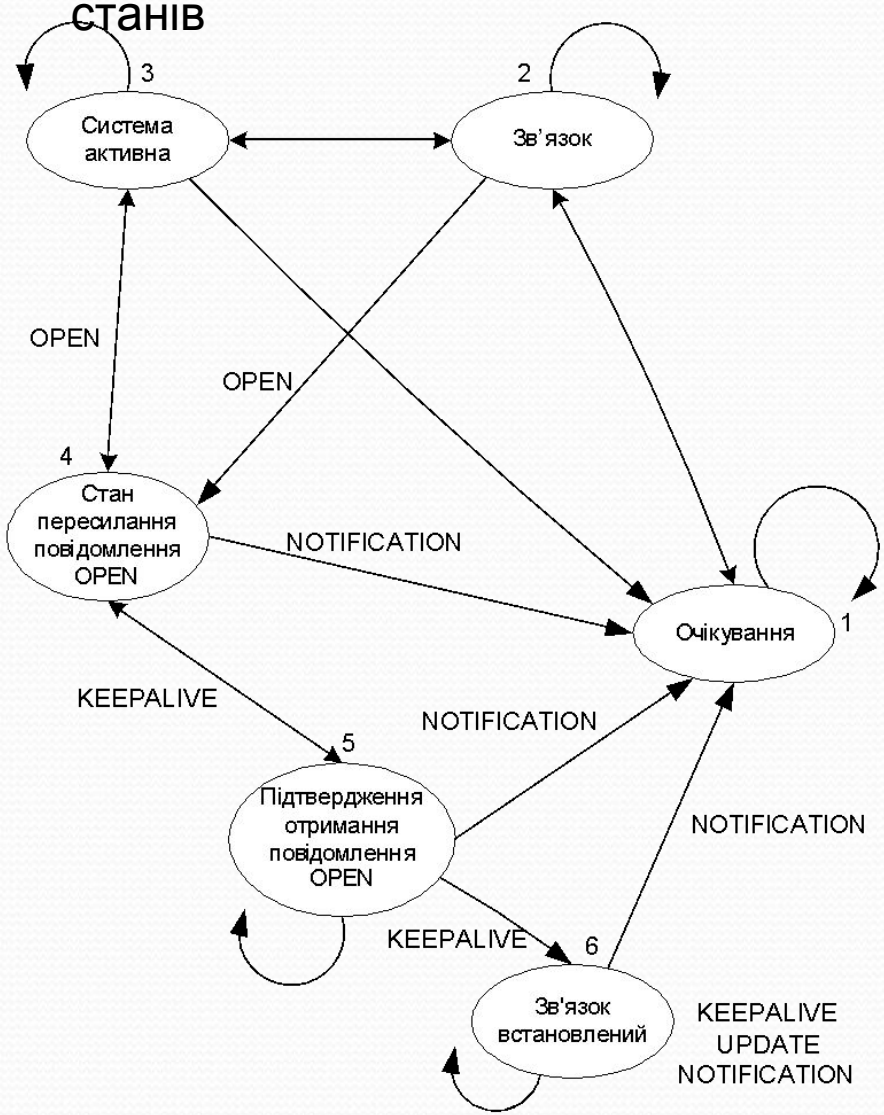

Рис. 5.13. Модель кінцевого стану при переговорах по протоколу BGP між сусідніми вузлами

5.3.5 Повідомлення NOTIFICATION

При виявленні помилки при установленні з'єднання, іншій стороні посилається повідомлення про помилку – повідомлення NOTIFICATION. Після цього, вузол, що послав повідомлення, розриває з'єднання. Мережні адміністратори повинні вміти аналізувати зміст повідомлення NOTIFICATION, щоб визначити причини, які викликали помилку протоколу маршрутизації. На рис. 5.14 наведений формат повідомлення NOTIFICATION.

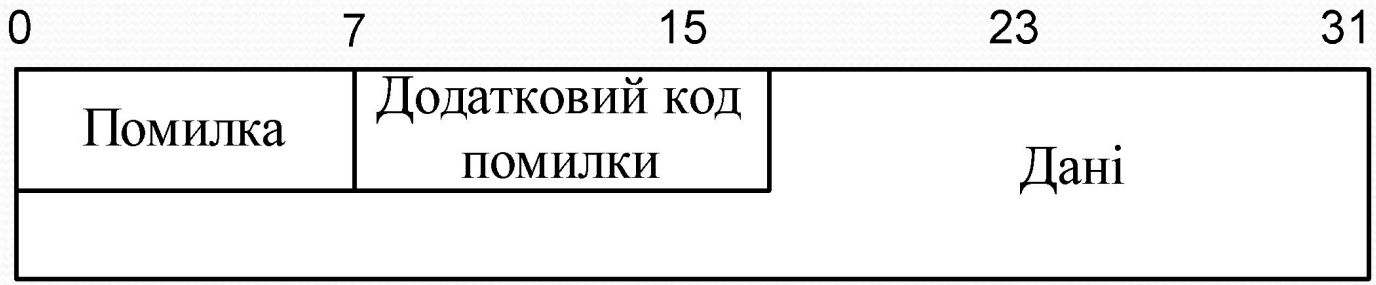

Рис. 5.14. Формат повідомлення NOTIFICATION

5.3.6 Повідомлення KEEPALIVE та UPDATE

Сторони, які беруть участь у сеансі зв'язку, періодично обмінюються повідомленнями типу KEEPALIVE для того, щоб визначити наявність каналу зв'язку й можливість досягнення по ньому вилученого вузла. Як вже відзначалося, час утримання визначає максимальний інтервал часу між успішним прийомом двох повідомлень типу KEEPALIVE або UPDATE. Повідомлення типу KEEPALIVE посилають звичайно із частотою, меншої часу, установленого таймером утримання, на підставі чого робиться висновок про нормальний хід сеансу. Рекомендований інтервал часу для посилки повідомлень KEEPALIVE – 1/3 від значення таймера утримання. Якщо ж таймер утримання встановлений в 0, то обмін повідомленнями KEEPALIVE не ведеться. Повідомлення типу KEEPALIVE представляє собою 19-байтовий заголовок протоколу BGP, без яких-небудь значень у полі даних.

Відновлення маршрутів несуть у собі всю необхідну інформацію, що використовується в протоколі BGP для побудови мережі без петель маршрутизації. В повідомлення UPDATE, як правило, входять три основних блоки:

- інформація мережного рівня про доступність мережі (Network Layer Reachability Information – NLRI);
- атрибути маршруту;
- недосяжні маршрути.

На рис. 5.15 показані усі компоненти повідомлення UPDATE.

Список атрибутів маршруту дозволяє протоколу BGP виявляти петлі в маршрутизації й надає йому додаткову гнучкість при визначенні локальних і глобальних правил маршрутизації. Як приклад атрибутів маршруту можна привести атрибут AS_PATH, за допомогою якого визначається послідовність номерів AS, що входять в маршрут до маршрутизатора BGP.

Наприклад, на рис. 5.16 автономна система AS3 одержує повідомлення UPDATE від AS2, де вказується, що в мережу 10.10.1.0/24 (NLRI) можна потрапити через два проміжних вузли – AS2 й AS1. На основі цієї інформації система AS3 може направляти трафік у мережу 10.10.1.0/24 через транзитний вузол AS2 у пункт призначення, підключений до AS1.

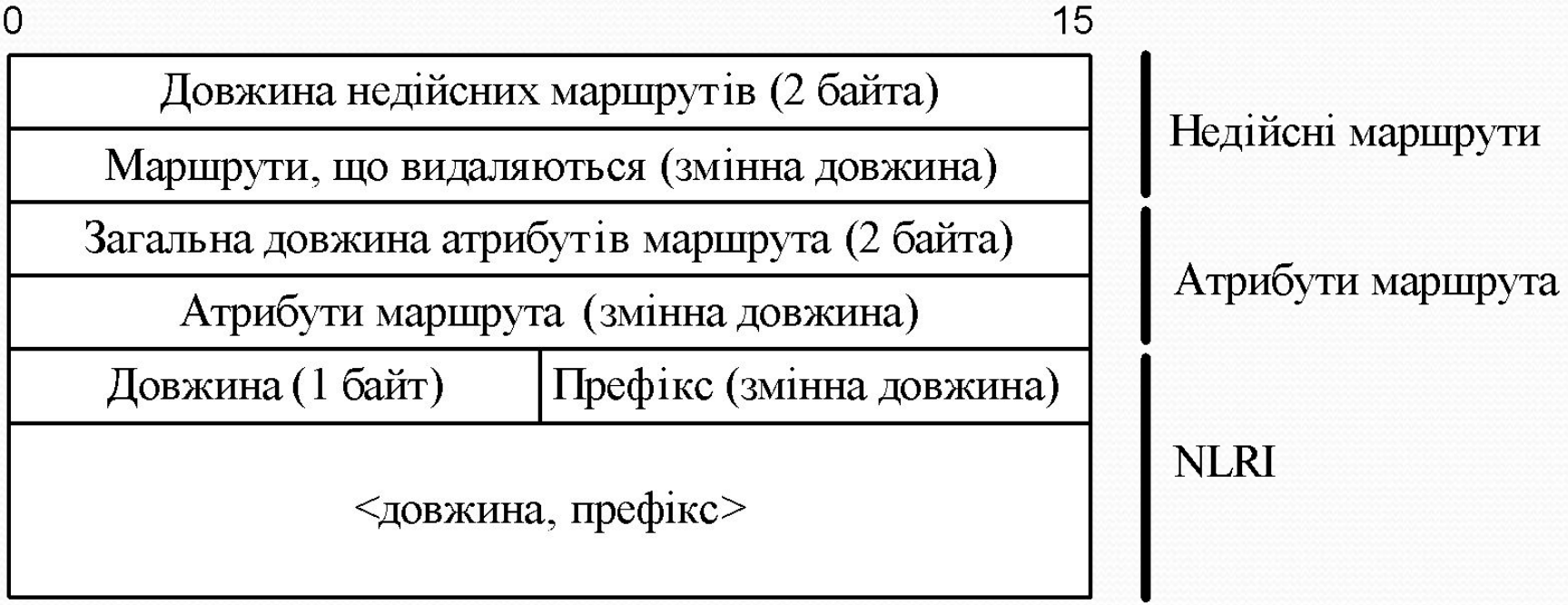

Рис. 5.15. Формат повідомлення UPDATE протоколу BGP

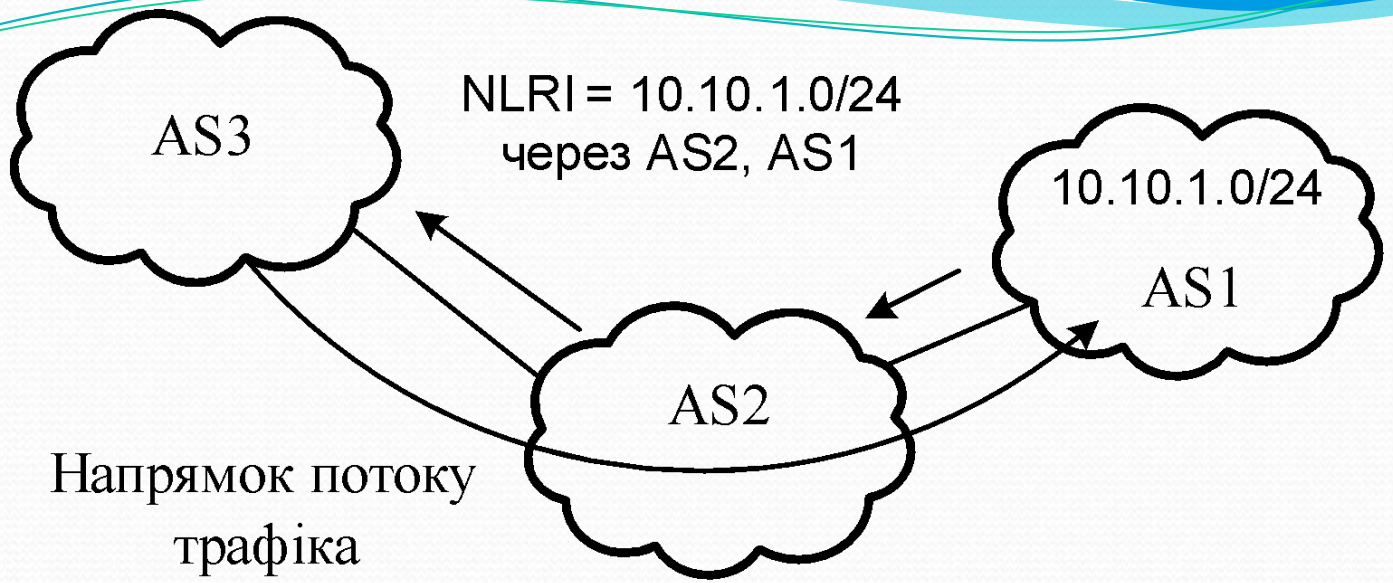

Рис. 5.16. Приклад відновлення маршрутної інформації в BGP

Третя частина повідомлення UPDATE являє собою список маршрутів, які стали недійсними, або, відповідно термінології, прийнятої в BGP, вилученими. З рис. 5.16 видно, що якщо мережа 10.10.1.0/24 стає недосяжною або в інформацію про атрибути маршруту внесені будь-які зміни, протокол BGP у всіх трьох AS може видалити маршрут і розіслати повідомлення UPDATE зі списком нової інформації про атрибути або про неможливість потрапити по заявленому раніше маршруту в задану мережу.

Принцип роботи протоколу BGP

Протокол BGP є протоколом вектора маршруту й використовується для обміну маршрутною інформацією між автономними системами. *Вектор маршруту (path vector)* пояснює принцип дії BGP: маршрутна інформація містить послідовності номерів AS, через які пройшов пакет із заданим префіксом мережі. Маршрутна інформація, пов'язана із префіксом, використовується для профілактики утворення петель у маршрутах.

Як транспортний протокол в BGP використовується протокол TCP (порт 179). Таким чином вся надійність доставки (включаючи повторну передачу) покладається на протокол TCP і не вимагає окремої реалізації в самому BGP, що спрощує механізми надійності в BGP.

Маршрутизатори, які працюють із протоколом BGP, часто називають *спікерами BGP (*BGP speakers*)*. Два спікери BGP, що утворюють сполучення – TCP-з'єднання один з одним для обміну маршрутною інформацією, називають *сусідніми (neighbors)* або *взаємодіючими (peers)*. На рис. 5.8 показана схема такої взаємодії. Взаємодіючі маршрутизатори спочатку обмінюються відкритими повідомленнями для того, щоб визначити параметри з'єднання. Ці повідомлення використовуються для узгодження параметрів, таких як номер версії BGP й ін.

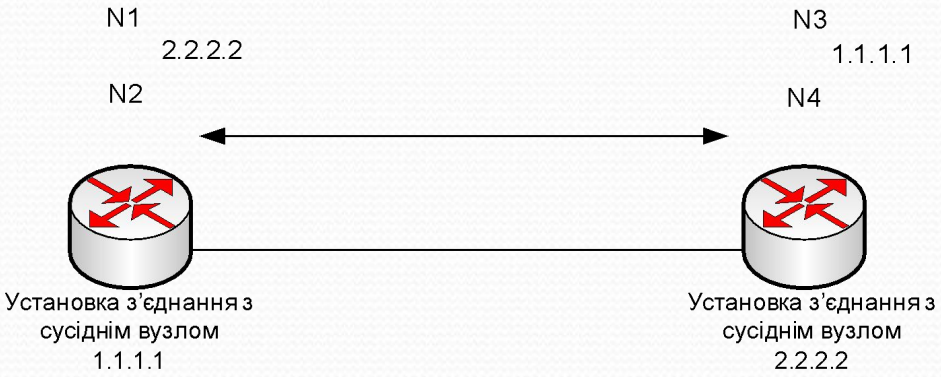

Рис. 5.8. Маршрутизатори BGP стають сусідами

Протокол BGP також забезпечує дуже витончений механізм закриття з'єднання з сусіднім маршрутизатором. Інакше кажучи, у випадку негативного результату переговорів між взаємодіючими маршрутизаторами, що може бути результатом несумісності їхніх конфігурацій, втручання оператора або викликано іншими причинами, генерується й посилає повідомлення про помилку NOTIFICATION. Одержавши це повідомлення, друга сторона повинна припинити спроби встановити з'єднання або розірвати його, якщо воно було встановлено раніше. Перевага такого механізму полягає в тім, що обидві сторони повідомляються про неможливість установки з'єднання й не витрачають свої потужності на обслуговування цього з'єднання або спроби повторно встановити зв'язок. Процедура закриття також гарантує, що обидві сторони до закриття сеансу TCP, одержать всі повідомлення про помилки, зокрема повідомлення NOTIFICATION.

На початку сеансу BGP між декількома спікерами BGP ведеться обмін всіма маршрутами, які можуть далі використатися в роботі за протоколом BGP (рис. 5.9). Після того як з'єднання встановлене й проведений початковий обмін маршрутами, по мережі розсилається лише інформація про нові маршрути – так звані *інкрементні відновлення (incremental updates)*. Застосування інкрементних відновлень, у порівнянні з періодичним відновленням маршрутів, що використовувалося в інших протоколах, таких як EGP, дозволило багаторазово збільшити продуктивність центральних процесорів на маршрутизаторах і розвантажити смугу пропускання.

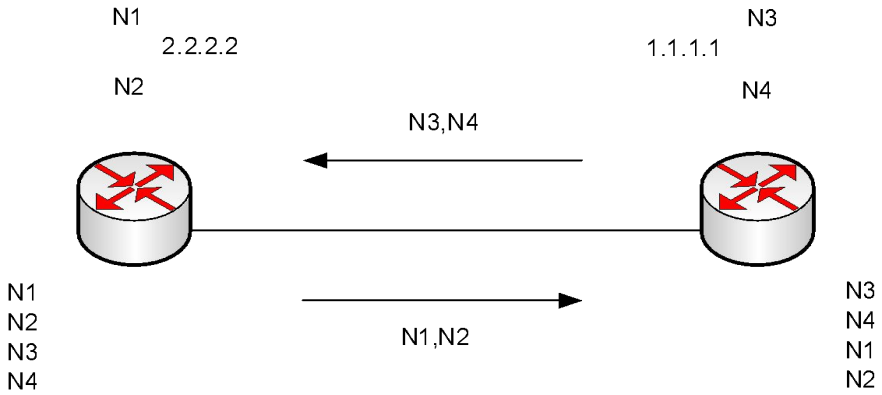

Рис. 5.9. Обмін відновленнями маршрутної інформації

Відповідно протоколу BGP, пара маршрутизаторів повідомляється про маршрути й зміни в них за допомогою повідомлення UPDATE. Повідомлення UPDATE, крім іншої корисної інформації, містить список записів типу <довжина, префікс> (<length, prefix>), що вказують на список вузлів, на які можна доставити трафік через спікер BGP.

В повідомлення UPDATE також включені атрибути маршруту. До них відноситься: ступінь переваги певного маршруту й список AS, через які пролягає маршрут.

У випадку якщо маршрут стає недійсним, тобто по ньому неможливо досягти пункту призначення, спікер BGP інформує про цьому своїх сусідів і видаляє недійсний маршрут. Як показано на рис. 5.10, маршрути, що видаляються також включаються в повідомлення UPDATE. Таким чином, ці маршрути вже не можна використовувати. Якщо ж інформація про маршрут змінилася або для того ж префікса обраний новий маршрут, то процедура видалення не виконується; у цьому випадку досить лише оголосити про заміну маршруту.

На рис. 5.10 показана система в *урівноваженому стані (steady state):* якщо немає ніяких змін у структурі маршрутів, то маршрутизатори обмінюються тільки пакетами KEEPALIVE.

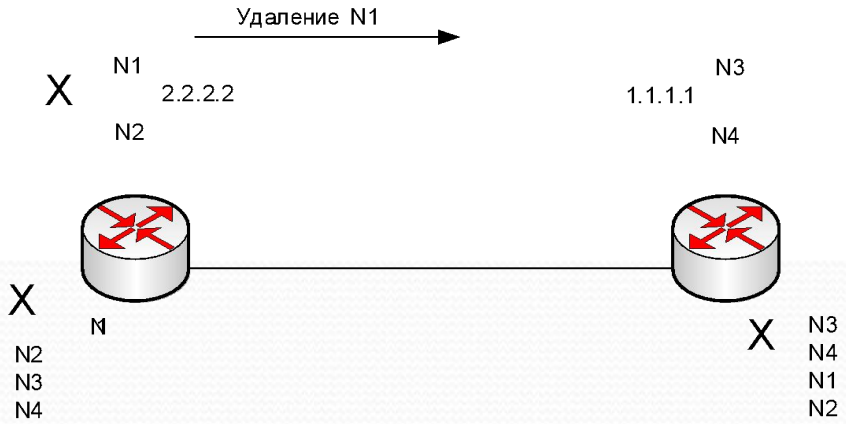

Рис. 5.10. Маршрут NJ стає несправним. Посилається часткове відновлення

Повідомлення KEEPALIVE періодично посилають між сусідніми маршрутизаторами BGP, щоб переконатися, що з'єднання перебуває в нормальному стані. Пакети KEEPALIVE (довжиною 19 байт кожен) не створюють практично ніякого навантаження на процесор маршрутизатора й смугу пропускання, тому що їм потрібно дуже незначна смуга пропускання (близько 2,5 байт/с).

У протоколі BGP враховується номер версії таблиці маршрутів, щоб відслідковувати зміни маршрутів. Якщо в таблицю маршрутів вносяться будь-які зміни, то BGP автоматично збільшує номер версії таблиці. Швидко зростаючі номери версії таблиці звичайно вказують на те, що в мережі є нестабільно працююча ділянка (хоча це досить звичайна ситуація для мереж великих провайдерів Internet). Нестабільність мереж, підключених до Internet по усьому світу, приводить до росту номерів версій таблиць маршрутів на кожному спікері BGP, де є відомості про всі маршрутні таблиці мережі Internet. Для зниження впливу цих неоднорідностей в Internet були розроблені механізми комутації маршрутів й інші заходи.

5.4 Питання для самоконтролю

1. Поясність суть технології CIDR.
2. Які завдання дозволяє вирішити впровадження технології CIDR?
3. Опишіть, що розуміють під поняттям AS.
4. Поясніть призначення автономних систем.
5. Поясніть відмінності AS з заглушкою від транзитної AS.
6. Який трафік вважається транзитним?
7. Які протоколи маршрутизації використовуються для обміну інформацією між автономними системами; в середині автономної системи?
8. Яка версія протоколу BGP використовується в цей час?
9. Яким чином протокол BGP зменшує розміри таблиць маршрутизації?
10. Назвіть TCP-порт, використовуваний для сеансів рівноправного зв'язку BGP.
11. Як забезпечується обмін маршрутною інформацією в протоколі BGP?
12. Який параметр спільно використовується рівноправними маршрутизаторами одного й того ж домену маршрутизації BGP?
13. Який тип взаємин встановлюється між двома маршрутизаторами різних AS?
14. Після того як за допомогою TCP установлений сеанс зв'язку на транспортному рівні між рівноправними маршрутизаторами, за допомогою якого повідомлення відправляється запит на запуск однорангового сеансу?
15. Назвіть основні атрибути протоколу BGP.

16. Назвіть основні типи повідомлень BGP.

17. Який тип повідомлення BGP використовується для попередження маршрутизаторів про помилки?

18. Назвіть три типи AS в BGP.

19. Назвіть і коротко охарактеризуйте основні стани моделі кінцевих станів. Поясніть для чого ця модель використовується.