УДК 004.738.5-004.732-621.391 DOI: 10.33839/2708-731X-25-1-14-20

В.М. Кулаківський, канд. техн. наук; **О.М. Давидов**, **І.В. Скворцов,** інженери

Інститут надтвердих матеріалів ім. В. М. Бакуля НАН України, вул. Автозаводська 2, 04074, м. Київ, e-mail: ivan@ism.kiev.ua

ВИКОРИСТАННЯ РЕКУРСИВНОЇ МАРШРУТИЗАЦІЇ ДЛЯ РЕЗЕРВУВАННЯ ТА БАЛАНСУВАННЯ НАВАНТАЖЕННЯ МІЖ КАНАЛАМИ ДЛЯ РОЗПОДІЛЕНОЇ ГЕТЕРОГЕННОЇ МЕРЕЖІ ПІДПРИЄМСТВА НА ПРИКЛАДІ МЕРЕЖІ ІНМ НАН УКРАЇНИ

Предметом дослідження статті є взаємодія локальної комп'ютерної мережі Інституту надтвердих матеріалів із глобальною мережею Internet. Метою дослідження є оптимізація взаємодії локальної та глобальної комп'ютерних мереж, балансування навантаження між каналами зв'язку із глобальною мережею. Надано рекомендації щодо налаштування маршрутизації в мережі, підвищення надійності зв'язку між локальною та глобальною мережами.

Ключові слова: *рекурсивна маршрутизація, балансування навантаження, інтернет трафік, routerOS, ECMP.*

Постановка задачі. Одна з найважливіших характеристик локальної обчислювальної мережі підприємства – це надійність зв’язку локальної мережі (intranet) з глобальною мережею (internet). А забезпечення безперебійного доступу до Internet на даний час є особливо важливим. Зазвичай це питання вирішується створенням декількох незалежних підключень до глобальної мережі Інтернет. Мережа нашого інституту належить до середніх або великих локальних мереж (близько 300 користувачів). На даний час корпоративні мережі таких розмірів мають, як правило, не менше двох незалежних підключень. Надалі ми будемо використовувати назву «*Підключення1»* для основного каналу та «*Підключення2*» для резервного каналу. На перших етапах побудови мережі ІНМ НАНУ ми використовували у якості маршрутизаторів сервери на платформах x86 та x64 під керуванням *FreeBSD (4.0 4.3, 6.2, 12)*. Однак дана операційна система, як і інші серверні ОС для платформи x86/x64, не мали (і не мають досі) вбудованих засобів для зручного перемикання між каналами, отже, для перемикання у разі відмов одного з каналів було необхідно або ручне втручання, або використання написаних скриптів.

Метод вирішення. Використання апаратних маршрутизаторів. Найбільш доступні на українському ринку маршрутизатори, які дозволяють гнучке конфігурування згідно з нашими завданнями – це сімейство маршрутизаторів *MikroTik.*

Обидва способи – і ручне втручання, і використання скриптів – мають суттєві недоліки, тому зі зростанням локальної мережі ІНМ став здійснюватися поетапний перехід на апаратні маршрутизатори *MikroTik* під керуванням *RouterOS*. *RouterOS* має вбудовані засоби моніторингу шлюзів:

1) за допомогою пошуку в таблиці *arp* (таблиці відповідностей фізичних адрес мережевих пристроїв та IP адрес);

2) за допомогою ping (утиліта для перевірки з'єднань у мережах *TCP/IP*).

Для використання цих засобів достатньо встановити два (або більше) *default route* (маршрут за замовчуванням) на *Підключення1* і *Підключення2* з різними метриками. За такої конфігурації трафік відправлятиметься через основний канал – маршрут з меншою метрикою. У разі відмови основного каналу, коли *ping* на його шлюз за замовчуванням зникає, маршрут стає неактивним, і трафік перенаправляється через резервний канал. Таким чином, немає

необхідність ручної зміни *default route*. Однак це тільки часткове вирішення проблеми резервування каналів доступу до Інтернету. Проблема в тому, що така конфігурація дозволяє відслідковувати лише стан каналу до найближчого шлюзу *default route*. У разі проблем або із зовнішнім каналом самого *uplink*-провайдера, або навіть далі, на магістральному каналі Інтернет (таке хоч і дуже рідко, але все ж таки трапляється) дана конфігурація не зможе автоматично переключитися з основного каналу на резервний або навпаки, оскільки *default router* відповідатиме на запити *ping*. Зазвичай таке завдання вирішується за допомогою різних протоколів динамічної маршрутизації [1]. Однак для їх використання необхідне бажання *uplink*-провайдерів, а вони не завжди мають бажання та технічну можливість анонсування своїх маршрутних таблиць – наприклад, дана конфігурація може бути не передбачена автоматизованою системою управління політикою мережі.

Вирішення задачі. Обладнання *MikroTik* дає можливість вносити рекурсивні маршрути вручну, як статичні, і при цьому перевіряти їх працездатність за допомогою *ping*. Для реалізації цієї можливості в налаштуваннях маршрутизатора *MikroTik* передбачені два параметри: «*scope*» – «поточна область маршруту», та «*target scope*» – «область пошуку». Будь-який маршрут має свої *scope* і *target-scope*. Для рекурсивних *default route* у різних джерелах (Інтернеті або документації), зокрема тут [2–4], можна зустріти рекомендацію використовувати «високодоступні» хости, наприклад, DNS *Google* 8.8.8.8, 8.8.4.4 або *Cloudflare* DNS 1.1.1.1 . На наш погляд, дана рекомендація має певний мінус: дані DNS-сервера нерідко використовуються як у налаштуваннях сервера *DHCP* як на самому центральному маршрутизаторі, так і на «малих» маршрутизаторах підключених клієнтів і навіть на робочих станціях, і при відмові каналу одного з провайдерів визначення стане недоступними. Тому в нашому випадку було взято сервер dns для доменної зони *com.ua*, що знаходиться в США, в Санта-Кларі – *ua-ba1.na.colo.net* (74.123.224.37). Це досить «далеко» від нашої мережі, що певною мірою забезпечує показник надійності: якщо даний сервер доступний, тобто ping на нього проходить, можна вважати, що і «весь Інтернет» через даний канал підключення доступний. Крім того, хости за запитами DNS до кореневих серверів звертаються порівняно нечасто – в більшості випадків дана інформація зберігається в кеші dns, а якщо й станеться так, що цей сервер не буде доступний, тобто інші основні авторитативні сервери dns для доменів верхнього рівня. Звичайно, тільки *ping*, без додаткових параметрів, недостатньо ефективний, проте це краще, ніж взагалі нічого. Другий маршрут можна вказати і більшою метрикою, просто з моніторингом шлюзу за замовчуванням: якщо вже так станеться, що відмовить одночасно 2 канали доступу, то тут не допоможуть ніякі «хитрощі». Методика даного налаштування детально описана у різних блогах, на форумах в Інтернеті та в офіційній документації для *MikroTik* у прикладах [4], тому не будемо зупинятися на ній детально. Зазначимо лише загальне правило: значення *target scope* має бути **більшим або рівним**, ніж *scope* попереднього маршруту. Значення за замовчуванням для статичних маршрутів *scope = 30*, *target-scope = 10*. Отже, зі значеннями за замовчуванням маршрути не працюватимуть, як рекурсивні – значення *scope* та *target scope* необхідно корегувати вручну. Для маршруту до сервера dns нами було обрано значення *scope = 10*, *target-scope = 10,* для *default route – scope = 30, target-scope = 11*. Тим самим було досягнуто досить ефективне і надійне резервування каналів доступу до Інтернету: достатньо «віддалене» розташування точки моніторингу гарантує доступність, інакше основний канал переключається на резервний.

Однак у процесі спостереження за трафіком у мережі ми дійшли висновку, що цього недостатньо. Недоліки такої конфігурації:

1) у разі відмови зовнішнього магістрального каналу на США uplink-провайдера, у нашому випадку це *Підключення1*, або навіть за його межами, або проблем на самому сервері *ua-ba1.na.colo.net* (74.123.224.37) цей маршрут як непрацездатний і переключить увесь трафік на резервний канал, незважаючи на те, що сама мережа *Підключення1* залишалася б робочою;

2) швидкість резервного каналу вчетверо менша, ніж основного, проте у цій конфігурації він буде “простоювати”, бо через нього будуть відправлятися лише контрольні ping-запити на сервер, а незважаючи на меншу швидкість, було б бажано, щоб він також використовувався, а не просто простоював в очікуванні відмови основного каналу.

Зазвичай для досить великих мереж, що мають власні, зареєстровані в *RIPE*, номери автономної системи (AS), таке завдання також вирішується за допомогою протоколу *BGP*, точніше *E-BGP* – протоколу граничного шлюзу [5], зовнішньої версії, для маршрутизації між *AS* [1, 6]. Однак зовнішній блок IP-адрес, що використовується в корпоративних мережах середніх масштабів, як правило, становить до 30 адрес (у *CIDR* підмережа /27); реєструвати окрему AS для такого блоку неможливо. Альтернативні ж рішення – використання «приватних» номерів AS, або маршрутизація всередині самої *AS* з використанням внутрішньої версії протоколу *BGP*, *I-BGP* – виявилися не передбаченими автоматизованою системою управління мережею uplink-провайдера *Підключення1*, тому основний uplink-провайдер відмовився анонсувати маршрути *BGP-4*.

Проте операційна система *RouterOS*, використовувана на маршрутизаторах *MikroTik*, як було зазначено у прикладі з *default route*, дозволяє реалізувати статичну маршрутизацію з допомогою рекурсивних статичних маршрутів. Отже, є можливість запровадити статичні маршрути для потрібних мереж, причому так, щоб працездатність даних статичних маршрутів також перевірялася за допомогою *ping*. Цей підхід і був реалізований під час заповнення статичних маршрутів на центральному, тобто безпосередньо підключеним до обох мереж uplink-провайдерів.

Перший і найбільш очевидний крок – це прописування статичних маршрутів до мереж, що входять в AS безпосередньо підключених uplink-провайдерів. Така інформація доступна за допомогою on-line служб як самої *RIPE*, так і інших – зокрема, *2ip.ua* або *db-ip.com*. Оскільки анонсування даних таблиць «зовні» не передбачається, можливе використання досить великих агрегатів маршрутів [6, с. 185].

При цьому на виборі IP-адреси хоста для перевірки доступності маршрутів слід зупинитися більш детально. Звичайно, в ідеалі було б варто для перевірки доступності конкретної мережі використовувати IP-адресу максимально доступного хоста з цієї мережі. Однак таке не завжди є можливим. Якщо в даній мережі за допомогою мережевих утиліт *nslookup* або *dig* вдається виявити сервер *WWW*, *DNS* або інших Інтернет-служб, то з досить великою ймовірністю можна вважати, що цей хост буде доступним практично 24/7, і його можна використовувати для перевірки доступності мережі. Найчастіше такий хост виявити не вдається. Крім того, для такого рішення необхідно попередньо вказувати, якщо, звичайно, це вже не було зроблено шляхом агрегації маршрутів, окремий маршрут до даного хосту, що збільшує розмір таблиці маршрутизації і, відповідно, обсяг необхідної ручної роботи. В цьому випадку доводиться перевіряти маршрут за допомогою утиліти *traceroute* та за результатами трасування вирішувати, чи перевіряти найближчий до цієї мережі маршрутизатор, або навіть магістральний вузол найближчої мережі uplink-провайдера. Звичайно, у разі використання вузлів самого uplink-провайдера існує ймовірність того, що в разі відмови «далі» хоста, що перевіряється, дана мережа виявиться недоступною у разі проблем у мережі самого uplink-провайдера: *ping* на даний проміжний хост буде проходити, проте мережа за ним буде недоступною, якщо до цієї мережі немає інших альтернативних маршрутів у тій самій мережі uplink-провайдера. Якщо ця мережа є «шлейфною» [1, 7], тобто має лише одне з'єднання з глобальною мережею Інтернет, це не має значення, оскільки при відмові єдиного з'єднання мережа буде недоступною у будь-якому випадку, за будь-яких записів таблиці маршрутів. Однак у випадку, якщо у цієї мережі все ж таки є альтернативні маршрути, і ці маршрути не анонсуються *AS* uplink-провайдера, то можливі проблеми з доступом. Якщо ж як хост для ping вказати більш «віддалений» хост, то є ймовірність «хибної відмови» даного маршруту: якщо через відмову каналу за межами мережі маршрут до цієї мережі виявляється неактивним.

Доводиться шукати якийсь компроміс, враховуючи належність мережі – цю інформацію можна отримати або через он-лайн службу 2ip.ua, або, наприклад, за іменами зворотнього перетворення dns в процесі виконання трасування, а також враховуючи довжину префікса мережі – відсутність порівняно невеликої мережі з префіксом /24 менш критична, ніж більшою.

Далі заповнення таблиці маршрутизації здійснювалось виходячи з політики анонсування маршрутних таблиць, взятої з *2ip.ua*. При цьому, з метою зменшення числа записів, застосовувалася більша агрегація, аж до префікса /3 – незважаючи на те, що при цьому включалися можливі "чужі" підмережі. В даному випадку це навіть корисно, оскільки як default route вказано gateway «резервного» uplink, і на ті «чужі» підмережі трафік буде направлено також через "основний" канал. Крім того, у разі великих агрегатів маршрутів вказують 2 або більше альтернативних маршрутів з різними метриками, вказуючи при цьому як *next-hop* для ping 2 різних, що знаходяться досить «далеко», за даними traceroute, хоста – це гарантує досить надійний «захист» від хибної відмови даних маршрутів. Також окремо вказані рекурсивні статичні маршрути до найбільш відомих джерел трафіку – зокрема, позитивною виявилася обставина, що *cdn*-сервери відеохостингу *YouTube* знаходяться в межах AS основного uplink-провайдера (*Підключення1*). Також вказані маршрути до відомих мереж: UA-IX (українська точка обміну Інтернет-трафіком), мережі доставки контенту *CloudFlare*, *Clarivate* (*Web of Science*) та інших. Для контролю доступності мереж у сумнівному випадку краще вказувати *ping next-hop* далі, бо при перемиканні на резервний канал мережа все одно залишиться доступною, в той час як якщо *next-hop* буде досить близьким, є ймовірність того, що маршрут виявиться активним, проте через проблеми далі IP-адреси next-hop мережа виявиться недоступною.

ECMP протокол. Окремо слід зупинитися на технології *ECMP*, точніше, на її реалізації в *RouterOS MikroTik*. Технологія ECMP (Equal Cost Multipath Routing) дозволяє теоретично балансувати трафіком між двома або більше шлюзами за алгоритмом «round robin» або «по пакетах», або «по потоках», детальніше [8]. Однак при цьому в реалізації *ECMP RouterOS* не працює перевірка працездатності каналу за допомогою arp або ping [8, 9, 10]. Однак цей мінус також можна компенсувати саме встановленням рекурсивних маршрутів, тому що в разі падіння одного з каналів роутер автоматично переключить рекурсивний маршрут на інший інтерфейс. Однак використання даної технології на більшості маршрутів недоцільне з таких причин:

1) внаслідок використання технології *NAT* та протоколу *TCP*, коли з'єднання встановлюється між конкретними хостами, трафік просто не може балансуватися між різними інтерфейсами, доки сеанс *TCP* не буде завершений і не буде ініційований новий.

2) пропускна спроможність каналів *Підключення1* (основне) і *Підключення2* (резервне) помітно відрізняються, тому ділити весь трафік «наполовину» за алгоритмом «*round robin*» також недоцільно.

Виходячи з цього, *ECMP*-маршрути встановлювали не на всі мережі, а на деякі, зокрема на мережу *YouTube*. За спостереженнями за характером трафіку, багато користувачів переглядають відео *YouTube* за допомогою мобільних гаджетів – при цьому використовується протокол *udp*, а не *tcp*, а отже, передача пакетів може здійснюватися через різні канали за алгоритмом «round robin», оскільки протокол udp працює без встановлення з'єднання. Також вони були встановлені на AS сервера конференцій *Zoom*, оскільки даний сервіс став стандартом «де-факто» для дистанційного проведення різноманітних зустрічей – голосовий та відеотрафік також передається за протоколом udp [11]. Для компенсації ж відмінності у пропускній спроможності каналів можна також вказати кілька «рекурсивних» шлюзів, маршрути на які вказують на інтерфейс основного каналу *Підключення1* – у разі відмови одного з каналів він автоматично переключиться на резервний канал *Підключення2*.

Висновок

Таким чином, можна зробити висновок, що встановлення статичних рекурсивних маршрутів, хоча повною мірою і не замінює протокол міждоменної динамічної маршрутизації E-BGP, є простішим в реалізації і не вимагає прямого узгодження з технічною службою uplink-провайдерів та реєстрації унікальної AS, дозволяючи при цьому організувати резервування маршрутів та балансування трафіку між різними каналами. Крім того, цей спосіб має достатню масштабованість, бо досить просто організувати резервування та балансування трафіку більш ніж між двома підключеннями.

Ця технологія була реалізована при підключенні розподіленої гетерогенної корпоративної мережі *Інституту надтвердих матеріалів НАН України* [12], а також дочірніх підприємств.

V. M. Kulakivskyy, O. M. Davydov , I. V. Skvortsov
V. Bakul Institute for Superhard Materials of the National Academy of Sciences of Ukraine

APPLYING OF RECURSIVE ROUTING FOR REDUNDANCY AND LOAD BALANCING FOR A DISTRIBUTED HETEROGENEOUS NETWORK OF AN ENTERPRISE ON THE EXAMPLE OF V.BAKUL INSTITUTE FOR SUPERHARD MATERIALS

The article's subject of the study is the interaction of the local computer network of the Institute of Superhard Materials with the global Internet network. The purpose of the study is to optimize the interaction of local and global computer networks, as well as to balance the load between channels of communication with the global network. Recommendations are given for setting up routing in the network, increasing the reliability of communication between local and global networks.

Key words: *recursive routing, load balancing, internet traffic, routerOS, ECMP.*

Література

1. Хелеби С., Мак-Ферсон Д. Принципы маршрутизации в Internet. – М: Издательский дом «Вильямс», 2001. – 448 с.
2. Основы статической маршрутизации в Mikrotik RouterOS [Електронний ресурс]. – Режим доступу: https://temofeev.ru/info/articles/osnovy-staticheskoy-marshrutizatsii-v-mikrotik-routeros/.
3. Резервирование каналов в Mikrotik при помощи рекурсивной маршрутизации [Електронний ресурс]. – Режим доступу: https://interface31.ru/tech_it/2020/04/rezervirovanie-kanalov-v-mikrotik-pri-pomoshhi-rekursivnoy-marshrutizacii.html.
4. Manual: Using scope and target-scope attributes [Електронний ресурс]. – Режим доступу: https://wiki.mikrotik.com/wiki/Manual:Using_scope_and_target-scope_attributes.
5. A Border Gateway Protocol 4 (BGP-4) [Електронний ресурс]. – Режим доступу: https://datatracker.ietf.org/doc/html/rfc4271.
6. Манн С., Крелл М. Linux. Администрирование сетей TCP/IP. – М.: ООО «Бином-Пресс», 2003. – 656 с.
7. Остерлох Х. Маршрутизация в IP-сетях. Принципы, протоколы, настройка. СПб.: ООО «ДиаСофтЮП», 2002. – 512 с.
8. Сибгатулин М. ECMP и превратности балансировки на сетевом оборудовании [Електронний ресурс]. – Режим доступу: https://nag.ru/material/36217.
9. Mikrotik. Failover. Load Balancing [Електронний ресурс]. – Режим доступу: https://habr.com/ru/post/244385/.
10. ECMP load balancing with masquerade [Електронний ресурс]. -- Режим доступу: https://wiki.mikrotik.com/wiki/ECMP_load_balancing_with_masquerade.
11. Koryachko V., Perepelkin, D., Perepelkin D., Saprykin A. Ivanchikova M. Development of Cloud Video Conferencing System Based on Two-Phase Routing Networks / 2021 10th

Mediterranean Conference on Embedded Computing (MECO), 07-10 June 2021, Budva, Montenegro. – IEEE, 2021. – INSPEC Accession Number: 20738312; DOI: 10.1109/MECO52532.2021.9460189.

12. Інститут надтвердих матеріалів ім. В.М. Бакуля НАН України [Електронний ресурс]. – Режим доступу: http://www.ism.kiev.ua.

Надійшла 15.09.22

References

1. Helebi, S. & Mak-Ferson, D. (2001). *Principy marshrutizatsii v Internet [Internet Routing Principles]*. Moskva: Vilyams [in Russian].
2. Osnovy staticheskoi marshrutizatsii v Mikrotik RouterOS [Basics of Static Routing in Mikrotik RouterOS]. (n.d.). *temofeev.ru*. Retrieved from https://temofeev.ru/info/articles/osnovy-staticheskoy-marshrutizatsii-v-mikrotik-routeros/ [in Russian].
3. Rezervirovanie kanalov v Mikrotik pri pomoshchi rekursivnoi marshrutizatsii [Reserving channels in Mikrotik using recursive routing]. (n.d.). *interface31.ru*. Retrieved from https://interface31.ru/tech_it/2020/04/rezervirovanie-kanalov-v-mikrotik-pri-pomoshhi-rekursivnoy-marshrutizacii.html [in Russian].
4. Manual: Using scope and target-scope attributes. (n.d.). *wiki.mikrotik.com*. Retrieved from https://wiki.mikrotik.com/wiki/Manual:Using_scope_and_target-scope_attributes.
5. A Border Gateway Protocol 4 (BGP-4). (n.d.). *datatracker.ietf.org*. Retrieved from https://datatracker.ietf.org/doc/html/rfc4271.
6. Mann, S. & Krell, M. Linux. (2003). *Administrirovanie setei TCP/IP [TCP/IP network administration]*. Moskva: «Binom-Press» [in Russian].
7. Osterloh, H. (2002). *Marshrutizatsiia v IP-setsakh. Printsipy, protokoly, nastroika [Routing in IP networks. Principles, protocols, customization]*. Sankt-Peterburh: DiaSoftYuP [in Russian].
8. Sibgatulin, M. *ECMP i prevratnosti balansirovki na setevom oborudovanii [ECMP and the vicissitudes of balancing on network equipment]*. (n.d.). *nag.ru*. Retrieved from https://nag.ru/material/36217 [in Russian].
9. Mikrotik. Failover. Load Balancing. (n.d.). *habr.com*. Retrieved from https://habr.com/ru/post/244385/
10. ECMP load balancing with masquerade. (n.d.). *wiki.mikrotik.com.* Retrieved from https://wiki.mikrotik.com/wiki/ECMP_load_balancing_with_masquerade.
11. Koryachko V., Perepelkin, D., Perepelkin D., Saprykin A. Ivanchikova M. (2021). Development of Cloud Video Conferencing System Based on Two-Phase Routing Networks. Proceedings from 10th Mediterranean Conference on Embedded Computing (MECO)'21. (4 p.). IEEE.
12. Institut nadtverdikh materialiv im. V.M. Bakulia NAN Ukrainy [V.N. Bakul Institut for superhard materials of NAS of Ukraine]. (n.d.). *ism.kiev.ua*. Retrieved from http://www.ism.kiev.ua [in Ukrainian].