Нов български университет

Проектиране на корпоративни мрежи

Част III
Препоръки за използване на VSS и EtherChannel

Доц. Д-р Емил Стоилов

Департамент по Информатика на НБУ

София, януари 2015

Съдържание

В [1] и [2] бяха представени архитектурата и методите за оптимизиране на дизайна на слой 2 на корпоративните мрежи. Преди да преминем към оптимизирането на дизайна на слой 3, желателно е да можем да отговорим на следните въпроси:

1. Разполагаме ли с мрежови устройства със специални възможности и как да използваме тези възможности в нашия проект?

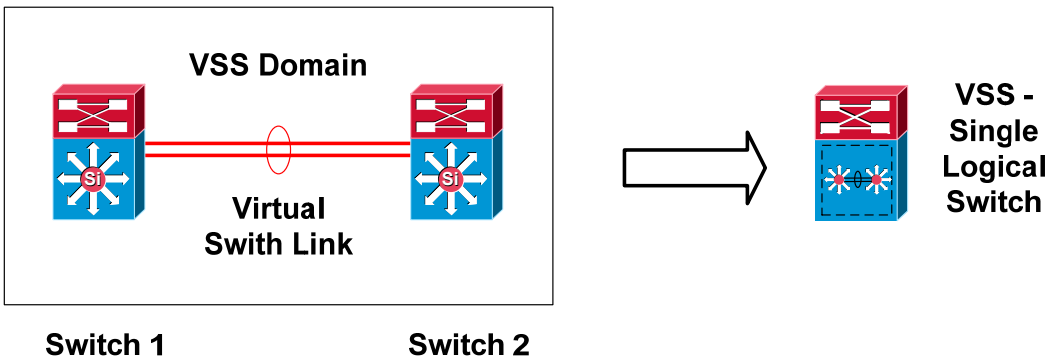
2. Съществуват ли специално разработени протоколи, които биха улеснили проектирането на мрежата и подобрили нейните характеристики?

3. Как да проектираме дизайна на границата между слой 2 и слой 3 и да го оптимизираме?

В този технически доклад ще се опитаме да дадем отговори на първите два въпроса.

1. Виртуални комутиращи системи

Както е показано на Фиг.1, виртуалната комутираща система (Virtual Switching System - VSS) представлява мрежова технология за виртуализация, която комбинира два Cisco Catalyst 6500 комутатора в една логическа единица. Двата комутатора образуват един виртуален комутатор с единно интегрирано управление, т.е. с една плоскост за управление (management plane) [3]. От гледна точка на дизайна на мрежата, VSS се разглежда като един комутатор с два надзорни модула (supervisors) [2], независимо от това, че тези модули се намират в две отделни физически шасита.



Фиг. 1 Физическа и логическа топология на VSS

Виртуалната връзка между комутаторите (Virtual Switch Link -VSL) е реализирана със специален EtherChannel канал, който продължава вътрешните магистрали до отсрещния комутатор. Това разширение позволява на надзорния модул на единия комутатор да управлява хардуера в другото шаси. По него се обменя управляваща информация като например програмиране на надзорния модул за работа в режим на горещ резерв (hot-standby), контролиране на състоянието на интерфейсните модули (line cards) [3], програмиране на модула за разпределено препредаване (Distributed Forwarding Card – DFC), управление на системата, диагностика и др. Освен това по VSL, когато е необходимо, се предават и потребителски данни. На практика този EtherChannel се състои от най-малко две 10 Gbps връзки, като всяка от тях е терминирана в отделен интерфейсен модул за постигане на максимална наличност.

От гледна точка на плоскостите за контрол и управление на мрежата [3], само един от двата надзорни модула е активен и контролира всички интерфейсни модули в двете шасита. Маршрутизиращият процесор (Route Processor – RP) във втория надзорен модул (от другото шаси) работи в режим на горещ резерв и е готов да поеме управлението, когато активният RP дефектира. Въпреки това, всички компоненти в плоскостта за предаване на данни, като различните комутиращи матрици (fabrics), PFC и DFC модулите, дори и

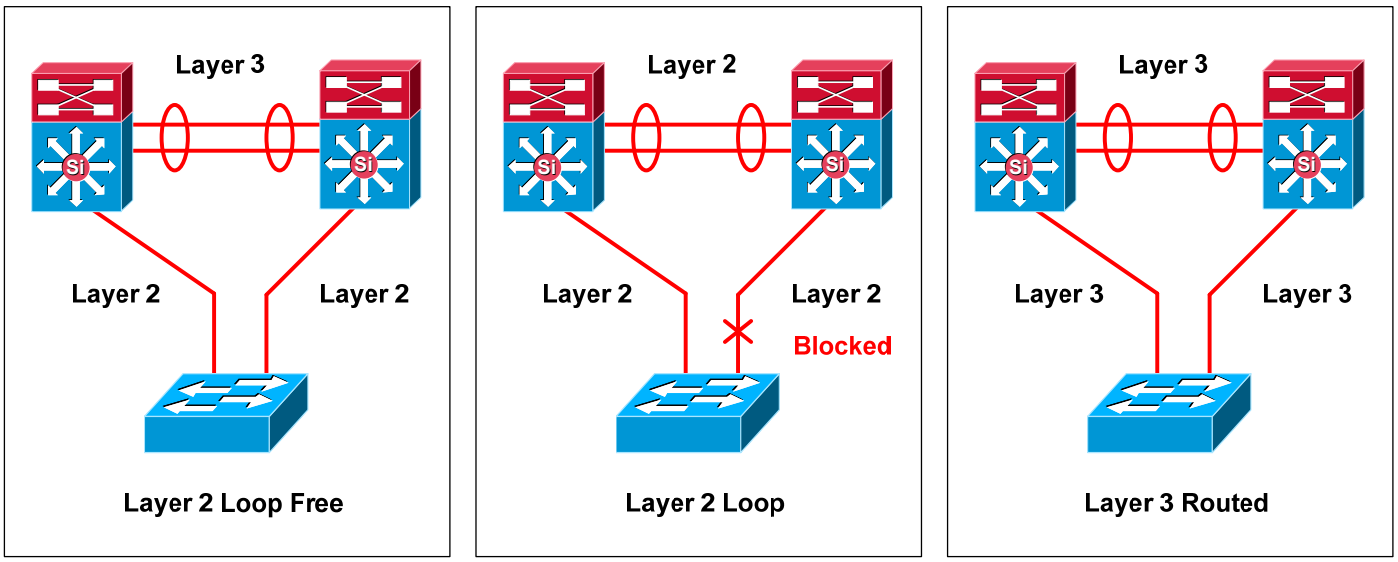
компонентите в работещия като горещ резерв надзорен модул, са активни и се използват за препредаване на трафик. В резултат на това, VSS има достъп до пълния капацитет за препредаване на съобщения на двата физически комутатора.

За постигането на висока надеждност, VSS използва съществуващите механизми за динамична защита на комутирането (Stateful Switchover – SSO) и за препращане без прекъсване (Nonstop Forwarding – NSF). Тези механизми бяха описани вече в [2]. Това гарантира бързото възстановяване след срив. Тъй като RP на втория надзорен модул се държи в състояние на горещ резерв, информацията за системата се синхронизира между двата надзорни модула по VSL. Когато активният надзорен модул дефектира, вторият модул поема активната роля. Това става, като се използват същите механизми, както ако разполагаме с два надзорни модула в едно шаси (в един комутатор).

От гледна точка на дизайна, ползата от VSS се състои в това, че можем да изградим EtherChannel, при който всяка връзка е физически терминирана в две отделни Catalyst 6500 шасита. Такава топология се нарича EtherChannel между отделни шасита (Multichassis EtherChannel - MEC). Понеже двете шасита формират една логическа единица, устройствата от другата страна на тази EtherChannel връзка използват стандартна EtherChannel технология за да се свържат към VSS. В резултат на това, VSS решението е напълно прозрачно за крайните устройства на един EtherChannel, като например комутаторите за достъп или сървърите.

1.1 Топология на връзките между слоя за достъп и разпределителния слой

На Фиг.2 е са показани различните топологии на свързване на слоя за достъп с разпределителния слой. За тази цел се използват главно три модела: дизайн без затворени контури в слой 2, дизайн със затворени контури в слой 2, и дизайн с маршрутизиране в слой 3.



Фиг.2 Дизайн на връзките между слоя за достъп и разпределителния слой

По-долу са описани тези три модела:

- **Дизайн без затворени контури в слой 2 (Layer 2 loop-free design):** При тази топология комутаторите за достъп използват комутиране в слой 2. Връзките между слоя за достъп и разпределителния слой са оформени като магистрали (trunks) от слой 2. Връзката между комутаторите в разпределителния слой е изградена като маршрутизираща връзка от слой 3. Обикновено тя е изградена като EtherChannel с цел увеличение на наличността. При този дизайн не могат да възникнат затворени

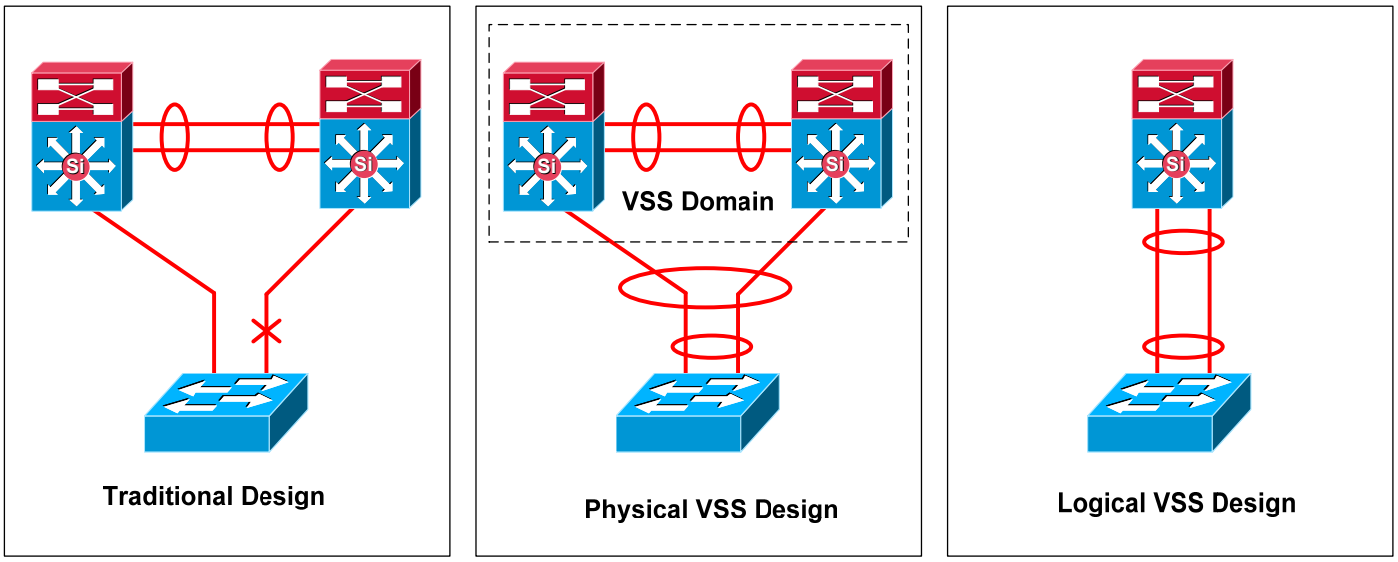
контури в слой 2, което означава, че протоколът Spanning Tree Protocol (STP) не играе никаква роля при сходимостта и балансирането на натоварването (load balancing). За STP всички портове са в режим на препредаване на съобщенията (Forwarding). За балансирането на трафика от слоя за достъп към разпределителния слой се използва протокола First Hop Router Protocol (FHRP), В случай на повреда времето за възстановяване се определя преди всичко от този протокол. Като недостатък на това решение можем да посочим, че то е оптимално само за мрежи, при които една VLAN от слоя за достъп е ограничена в рамките на един комутатор за достъп. Разпростирането на една VLAN до няколко комутатора за достъп не е препоръчително. Тази топология се смята за най-добра от практическа гледна точка. Често обаче не е възможно тя да бъде използвана, именно поради ограниченията, които налага върху виртуалните локални мрежи. По различни унаследени причини, много от съществуващите мрежи използват VLAN, които се разпростират върху множество комутатори от слоя за достъп, и това прави невъзможно използването на този сценарий.

- **Дизайн със затворени контури в слой 2 (Layer 2 looped design):** При този дизайн комутаторите за достъп също използват комутиране в слой 2 и връзките между слоевете за достъп и разпределение също са оформени като магистрали от слой 2. За разлика от предишната топология обаче, връзката между комутаторите в разпределителния слой е оформена като магистрала в слой 2. При такава конфигурация се появяват затворени контури между комутаторите в слоя за достъп и в разпределителния слой. За да елиминира тези контури, протоколът STP блокира някои връзки от комутаторите за достъп към комутаторите в разпределителния слой. Тази топология се препоръчва за мрежи, при които VLAN се разпростират върху множество комутатори за достъп. Недостатъкът е, че в случай на повреда, времето за възстановяване зависи вече от сходимостта на STP в комбинация със сходимостта на FHRP, Друг недостатък е ограниченото балансиране на натоварването.

- **Дизайн с маршрутизиране в слой 3 (Layer 3 routed design):** При този дизайн използваме маршрутизиране в слой 3 при комутаторите за достъп. Изобщо всички връзки между комутаторите са конфигуриране като маршрутизирани връзки от слой 3. Предимството на този дизайн е, че STP е елиминиран от връзките между комутаторите. Той все още може да бъде активиран на някои гранични портове с цел защита от създаване на затворени контури от страна на потребител, но не играе никаква роля при възстановяването на мрежата. FHRP също е елиминиран, понеже шлюзовете по подразбиране на крайните потребители сега се намират в комутаторите в слоя за достъп вместо в комутаторите в разпределителния слой. Времето за възстановяване на мрежата се определя само от характеристиките на използваните маршрутизиращи протоколи. Подобно на дизайна без затворени контури в слой 2, тук също виртуалните локални мрежи са ограничени до отделни комутатори в слоя за достъп. Тази конструкция освен че не позволява разпростирането на VLAN до множество комутатори, то има и изискването за използване на по-сложен хардуер при комутаторите в слоя за достъп.

1.2 EtherChannel между отделни шасита и VSS

VSS позволява използването на EtherChannel между отделни шасита (MEC). На Фиг. 3 са показани разликите между традиционния дизайн и физическия и логическия VSS дизайн. Както вече споменахме, VSS дава нови възможности за изграждане на връзка между слоя за достъп и разпределителния слой. Тъй като VSS представлява отделна логическа единица, технологията EtherChannel може сега да се използва за обединяване на две възходящи връзки (uplinks) на всеки комутатор за достъп в един логически канал, дори когато тези връзки са терминирани в две различни физически шасита на Catalyst 6500 комутаторите.



Фиг.3 VSS и MEC

Понеже STP разпознава EtherChannel като една логическа връзка, този протокол ефективно е отстранен от топологията на мрежата. Той все още може да бъде активиран на някои гранични портове с цел защита от неправилно и погрешно свързване от страна на крайния потребител, създаващо затворен контур, но не играе никаква роля в процеса на възстановяване на мрежата. Основното предимство на проекти използващи MEC е, че всички връзки между слоя за достъп и разпределителния слой се използват за предаване на трафик. Цялостният трафик е балансиран по връзките като се използват механизмите за EtherChannel хеширане. Другото предимство е, че този дизайн позволява VLAN да се разпростират до множество комутатори за достъп, без това да създава затворени контури в топологията на слой 2.

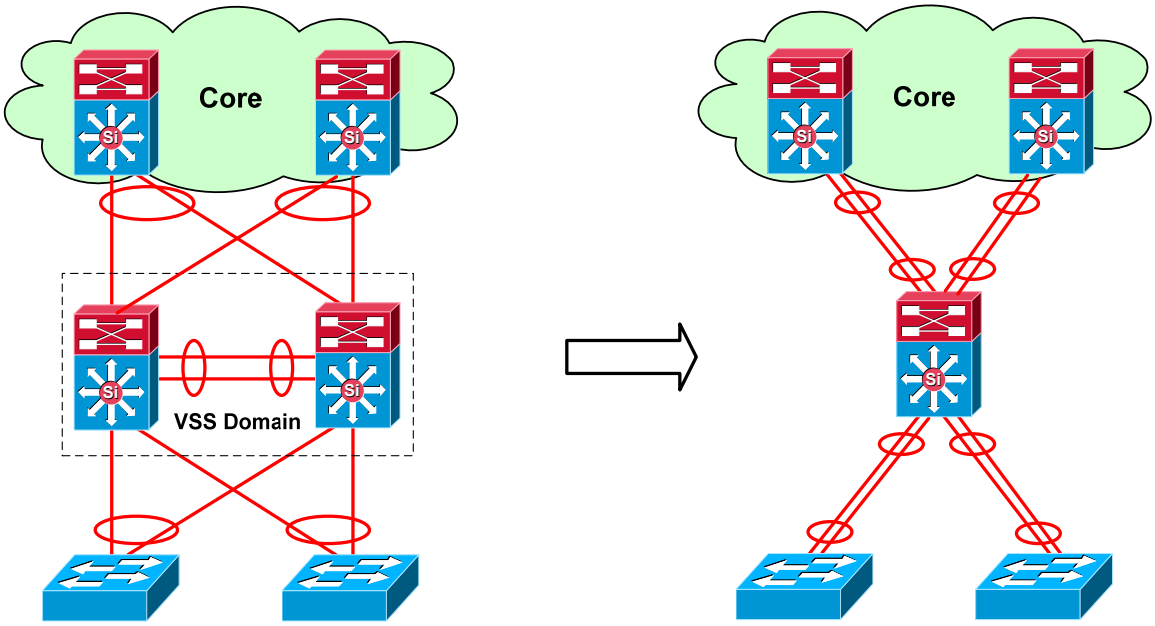
1.3 Съображения за използването на VSS дизайн

VSS позволява изграждането на логически топологии тип звезда, които запазват пълна физическа редундантност [2].

Както е показано на Фиг.4, VSS и MEC предоставят стабилен и високо достъпен дизайн без жертването на определена гъвкавост. Те също не налагат ограничения върху хардуера, който се използва в слоя за достъп.

VSS и MEC могат да се използват за създаване на логическа топология от тип звезда, като запазват пълната резервираност на базовата физическа топология. Връзките на MEC към комутаторите в слоя за достъп могат да бъдат изградени като магистрали от слой 2. Към ядрото на мрежата те могат да бъдат изградени като маршрутизирани връзки от слой 3. Към комутаторите в слоя за достъп те също могат да бъдат изградени като маршрутизирани връзки, стига там да е използван подходящия хардуер.

Обикновено в корпоративната мрежа на фирмения комплекс VSS се разгръща в разпределителния слой. Независимо от това, VSS може да се използва и в ядрото на мрежата, а дори и в слоя за достъп на комплекса от сървъри или в центъра за данни (data center), където да свържем сървъри с два интерфейса (dual-homed servers) към VSS като използваме MEC.



Фиг.4 Логическа топология на VSS

1.4 Двойно активно разпознаване и възстановяване

VSL е жизнено важен компонент на VSS. При повреда във VSL наличността на системата и нейната способност да се възстанови трябва да бъде гарантирана. Ключов момент при разсъжденията, когато използваме VSS, е дали имаме достатъчна редундантност на VSL между двете шасита. VSL се използва от надзорния модул в състояние на горещ резерв за наблюдение на състоянието на активния надзорен модул. Ако настъпи повреда във VSL, резервният надзорен модул приема, че другият надзорен модул е дефектирал и той възприема активната роля. Но може повредата да е само във VSL, а не в активния комутатор. В резултат на това два отделни комутатора стават активни. Понеже те използват една и съща конфигурация, в зависимост от използваните протоколи могат да се появят различни повреди:

- **EtherChannel в слой 2 (Layer 2 MEC):** И двата комутатора започват да изпращат служебни съобщения BPDU по MEC към съседните комутатори. Понеже те използват в тези BPDU различни физически (MAC) адреси, съседните комутатори виждат тази несъгласуваност и забраняват EtherChannel канала.

- **OSPF:** И двата надзорни модула започват да използват един и същ OSPF идентификатор на маршрутизатор. В зависимост от това дали използваме MEC от слой 3 или отделни връзки, се нарушава базата данни за съседство на този протокол или настъпва „наводнение" от пакети. В резултат на това подмрежите в слоя на достъп стават недостъпни.

- **EIGRP:** Когато използваме MEC от слой 3 може да се загуби базата данни за съседство в зависимост от това как EIGRP трафика е хеширан. В случай на използване на отделни връзки, EIGRP продължава да функционира, понеже идентификаторът на маршрутизатора играе значително по-малка роля, отколкото този при OSPF.

Най-добрия начин да се предотвратят тези повреди е да направим самата VSL колкото е възможно по-устойчива. Като минимум трябва да предвидим две връзки за VSL EtherChannel, но с използването на повече връзки увеличаваме нейната наличност.

1.5 Препоръки за използване на VSS дизайн

При проектирането и внедряването на VSS, следвайте препоръките на най-добрите практики за получаване на оптимални резултати:

- Винаги използвайте топологии тип звезда изградени върху MEC с VSS. Така можете да сте сигурни, че няма да се появят затворени контури и ще постигнете възможно най-добрите времена на сходимост

- Използвайте уникални номера на VSS домейните за всяка двойка, дори когато двойките не са директно свързани помежду си.

- Както при всяко EtherChannel свързване, винаги броят на връзките трябва да бъде четен и степен на 2 (2,4,8) за да се оптимизира балансирането на трафика по VSL.

- Не конфигурирайте превантивни действия на комутаторите (switch preemption). Превантивното действие (поемането на инициативата) осигурява един и същ надзорен модул да бъде винаги активен, когато двата комутатори са включени. Това може да предизвика при някои сценарии ненужни рестартирания и няма осезателна полза от това един конкретен надзорен модул да играе винаги активна роля.

- Настройването на протокола за управление на връзките (Link Management Protocol - LMP), на протокола за обединяване на връзките (Link Aggregation Control Protocol – LACP) и на таймерите на протокола за агрегиране на портовете (Port Aggregation Protocol – PAgP) по агресивен начин може да повлияе неблагоприятно на производителността на системата. При използването на топологии тип звезда, които се основават на VSS и MEC, протоколът STP не е включен активно в поддържането на топологията на мрежата. Въпреки това той не трябва да бъде забраняван и функциите PortFast и BPDU guard [2] трябва да се използват за защита от възникване на затворени контури на ръба на мрежата (network edge).

- Активирайте механизма за двойно активно разпознаване и възстановяване за да се защитите от повреди във VSL. Използвайте PAgP където е възможно. Ако този метод не е на разположение, използвайте метода fast hellos. Ако и този метод не може да се използва, опитайте с протокола Bidirectional Forwarding Detection (BFD).

2. Разработка на оптимален дизайн за слой 3

За да се постигне висока степен на наличност и бързо възстановяване на мрежата във фирмения комплекс с оборудване на Cisco, проектантът трябва да преследва множество разнородни цели, включително следните:

- Управление на свръх абонамента (oversubscription) и на честотната лента (bandwidth)

- Поддръжка на балансиране на натоварването на връзките (load balancing)

- Избор на маршрутизиращ протокол

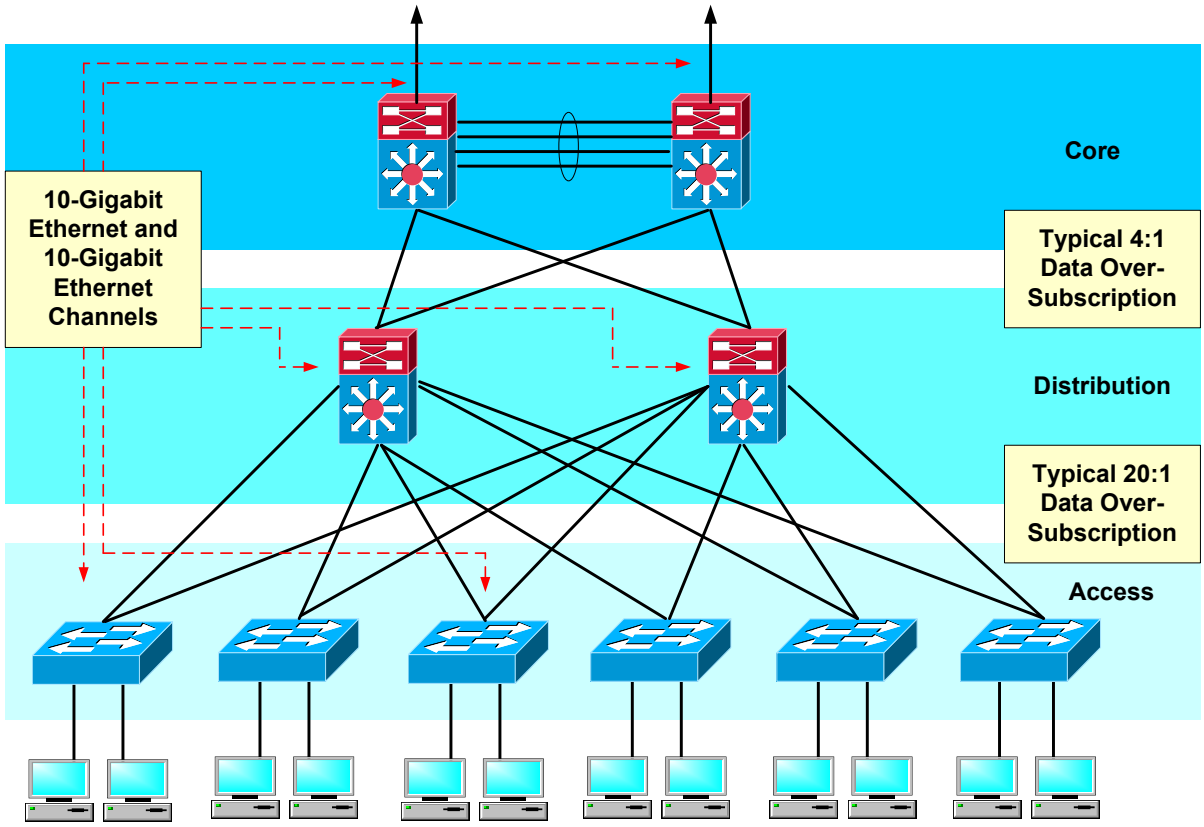
- Избор на протокол за резервиране на шлюза по подразбиране (First-hop Redundancy Protocol – FHRP).

В този раздел ще направим преглед на някои модели и препоръчителни практики за постигане на висока наличност и бързо възстановяване на слой 3 на мрежата. Не трябва да се смята, че това изчерпва въпроса за цялостно оптимизиране на слой 3 на мрежата. По-

скоро ще бъде подадена информация, която впоследствие ще бъде използвана при проектирането на границата между слой 2 и слой 3 на мрежата.

2.1 Управление на свръх абонамента и на честотната лента

Типичните мрежи на фирмените комплекси се проектират със свръх абонамент (oversubscription), както е показано на Фиг. 5. Препоръката за този свръх абонамент в слоя за достъп е да бъде 20:1, т.е. такова да бъде съотношението на броя портовете за достъп до мрежата към броя на портовете за връзка към разпределителния слой. При връзките от разпределителния слой към ядрото, това съотношение на портовете се препоръчва да е 4:1. Когато използвате такива съотношения на свръх абонамент, редно е да очаквате пренатоварването на връзките да стане често явление. В тези случаи трябва да използвате методи за гарантиране на качеството на услугите (Quality of Service – QoS). Ако задръстването на връзките се случва често, то проектът не разполага с достатъчна честотна лента по връзките нагоре (към ядрото).



Фиг.5 Управление на свръх абонамента и на честотната лента

Когато честотната лента на слоя за достъп достигне 1 Gbps, няколко пъти по 1 Gbps или дори 10 Gbps, се появява необходимост от агрегиране на трафика по връзките от разпределителния слой към ядрото. Тогава трябва да се използва за този трафик EtherChannel технология, при която скоростта по отделните връзки е от 1 до 10 Gbps.

2.1.1 Увеличение на честотната лента с EtherChannel

Когато трафикът от разпределителния слой към ядрото се увеличи значително, свръх абонамента към слоя за достъп трябва да бъде управляван и трябва да се вземат определени решения по изменението на дизайна. Простото добавяне на повече възходящи връзки от разпределителния слой към ядрото води до използването на повече портове, а това е свързано със значително увеличение на разходите.

EtherChannel може да намали броя на използваните портове, като създаде единен логически интерфейс. Все пак трябва да се осмислят някои въпроси свързани с това как маршрутизиращите протоколи реагират на прекъсване на единична връзка:

- Протоколът OSPF под операционна система Cisco IOS ще забележи повредата във връзката и трафикът ще бъде пренасочен по връзка с по-висока цена. Това събитие ще предизвика преходен процес с определени времена на сходимост.

- EIGRP може да не промени цената на връзката, тъй като протоколът се интересува от цената от край до край. Той може обаче да претовари останалите връзки.

Протоколът LACP поддържа функцията EtherChannel Min-Links. Тя ви позволява да конфигурирате минималния брой портове, които трябва да са в активно състояние (link-up state) и свързани в EtherChannel, за да се активира този канал. Можете да използвате функцията EtherChannel Min-Links за да предотвратите активирането на EtherChannel, който няма да има очакваната честотна лента.

2.1.2 Увеличение на честотната лента с 10 Gigabit интерфейси

Обновяването на връзките между разпределителния слой и ядрото с 10 Gigabit Ethernet връзки е друг алтернативен начин на управление на честотната лента. Връзките с 10 Gigabit Ethernet изпълняват и строгите изисквания за скорост на прехвърляните данни.

Този дизайн се препоръчва защото:

- За разлика от използването на много ниско скоростни връзки, една 10 Gigabit Ethernet връзка не увеличава броя на участниците в маршрутизацията.

- За разлика от решението с използване на EtherChannel, маршрутизиращите протоколи детерминистично ще изберат най-добрия път между разпределителния слой и ядрото.

2.2 Балансиране на натоварването на връзките

На Фиг.6, съгласно препоръчаната топология, са осигурени множество връзки между два комутатора. Тези връзки са с еднаква цена и преминават през разпределителния слой и ядрото. От гледна точка на слоя за достъп съществуват най-малко три връзки с еднаква цена за прехвърляне на трафика към другия комутатор.

Cisco Express Forwarding (CEF) е детерминистичен алгоритъм [3]. Както е показано на Фиг.6, когато пакетите преминават през мрежата, всички те използват една и съща входна стойност на CEF хеша, „наляво” или „надясно”, когато се взима решение за преминаване по редундантен път. В резултат на това получаваме, че някои пътища са пренебрегнати или недостатъчно използвани. Тогава казваме, че мрежата изпитва CEF поляризация.

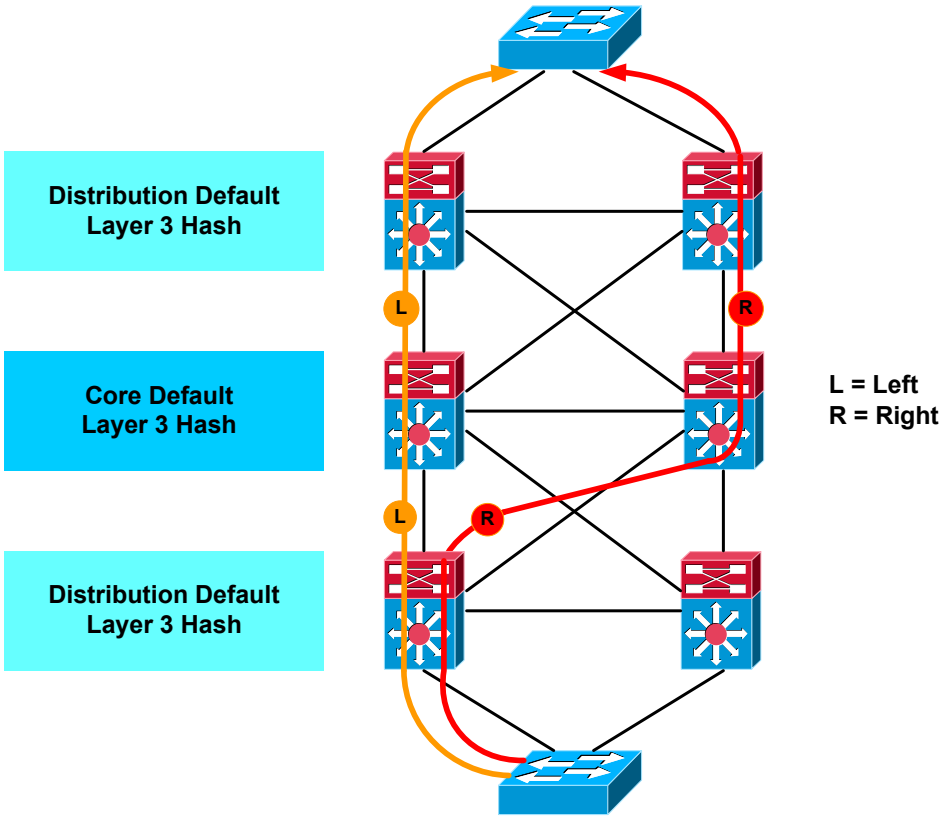
За да избегнете CEF поляризацията, можете да настройвате входа на CEF алгоритъма при преминаване през различните слоеве в мрежата. По подразбиране входната стойност е от слой 3 за източника и местоназначението. Ако промените тази входна стойност от слой 3 на слой 3 плюс слой 4, изходната хеш стойност също ще бъде променена.

Следователно препоръката е да използвате променливи хешове в комутаторите на ядрото и разпределителния слой:

- В слоя на ядрото използвайте настройките по подразбиране, които се опират на информация само от слой 3.

- В разпределителния слой използвайте информацията от слой 3 и слой 4 като вход на CEF хеш алгоритъма, като въведете командата

  Dist2-6500 *(config)# mls ip cef load-sharing full*.



Фиг.6 Балансиране на натоварването при CEF (по подразбиране)

Този алтернативен подход помага да елиминирате решенията „винаги надясно” или „винаги наляво” и да балансирате трафика в мрежата по редундантни връзки с еднаква цена.

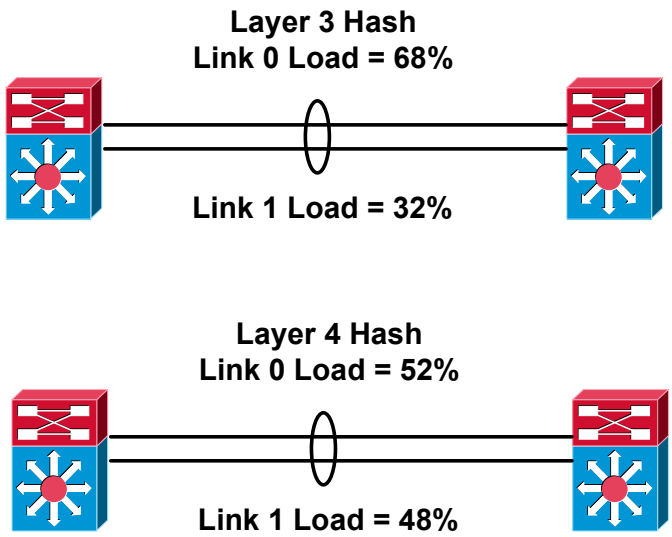
2.2.1 Балансиране на натоварването с EtherChannel

EtherChannel позволява разпределение на трафика между всички връзки на канала и резервираност в случай, че една или повече връзки в него дефектират. Можете да настроите хеш алгоритъма използван за избор на конкретна EtherChannel връзка по която пакетът да бъде предаден. За целта можете да използвате информацията по подразбиране от слой 3 за източника и местоназначението на пакета, или можете да добавите още информация от слой 4 (за TCP портовете) като вход за алгоритъма.

Фиг.6 илюстрира резултати от експерименти в тестова среда на Cisco, където е използвана една типична IP адресна схема с една подмрежа за VLAN, с два VLAN в комутатор в слоя за достъп и частно адресно пространство съгласно RFC 1918. Както се вижда с параметри по подразбиране (от слой 3) алгоритмът изпраща една трета от пакетите по втората връзка. Когато към него се подава и информация за слой 4 получаваме почти пълно балансиране на трафика между двете връзки.

Препоръчителната практика е да се използва за балансирането на трафика колкото се може повече информация от слой 3 и слой 4 като вход на EtherChannel алгоритъма с което се постига равномерно използване на връзките. Използва се командата *port-channel load-balance* за захранване на хеш алгоритъма с по-уникални стойности. Пример за това е:

dist1-6500 (config)*#port-channel load-balance srcdst-port.*



Фиг. 6 Балансиране на натоварването с EtherChannel

За постигане на по-добро разпределение на натоварването се използват две, четири или осем връзки в канала. При по-старите комутатори на Cisco има изискване портовете да са към една и съща ASIC интегрална схема.

Недостатък на балансирането с EtherChannel е, че то не се извършва за отделните пакети. Ако имаме генериран голям трафик от приложение, което използва една и съща информация до слой 4 включително, то целият този трафик (всички пакети) ще преминат по един и същи път.

2.2.2 EtherChannel или Equal-Cost Multipath?

В някои мрежови проекти можете да бъдете поставен пред дилемата да изберете конфигурация с EtherChannel или да използвате маршрутизиращия протокол Equal Cost Multipath (ECMP) за да балансирате натоварването между два комутатора от слой 3.

Кой метод е по-добър зависи от специфичните изисквания на дизайна. Тук ще направим само някои общи разсъждения:

- Опциите на алгоритмите за хеширане при EtherChannel и ECMP се различават и възможностите, които се предлагат са платформено зависими. Методите за хеширане и за смесване на трафика определят това доколко трафикът ще бъде балансиран по отделните връзки.

- При EtherChannel връзките между комутаторите се възприемат от маршрутизиращия протокол като една логическа връзка. Установява се само едно съседство между двата комутатора от слой 3. При ECMP се установява съседство за всяка връзка поотделно, което увеличава информацията, която трябва да обработва маршрутизиращия протокол.

- ECMP използва механизмите на маршрутизиращ протокол за да избере връзки, да добави или премахне пътища, както и да балансира трафика по връзките. EtherChannel използва LACP за управление на връзките. По принцип маршрутизиращите протоколи позволяват по-точен избор на път, имат по-добра сходимост и по-добри характеристики на балансиране на трафика.

- При ECMP, само маршрутизиращият протокол се занимава с пренасочването на трафика по връзките. EtherChannel използва както маршрутизиращ протокол, така и LACP. В тази връзка ECMP е по-лесен за конфигуриране и поддръжка.

- Поради по-големите възможности за контролиране и наблюдение, при ECMP е по-лесно да се откриват неизправности и да се отстраняват, отколкото при EtherChannel.

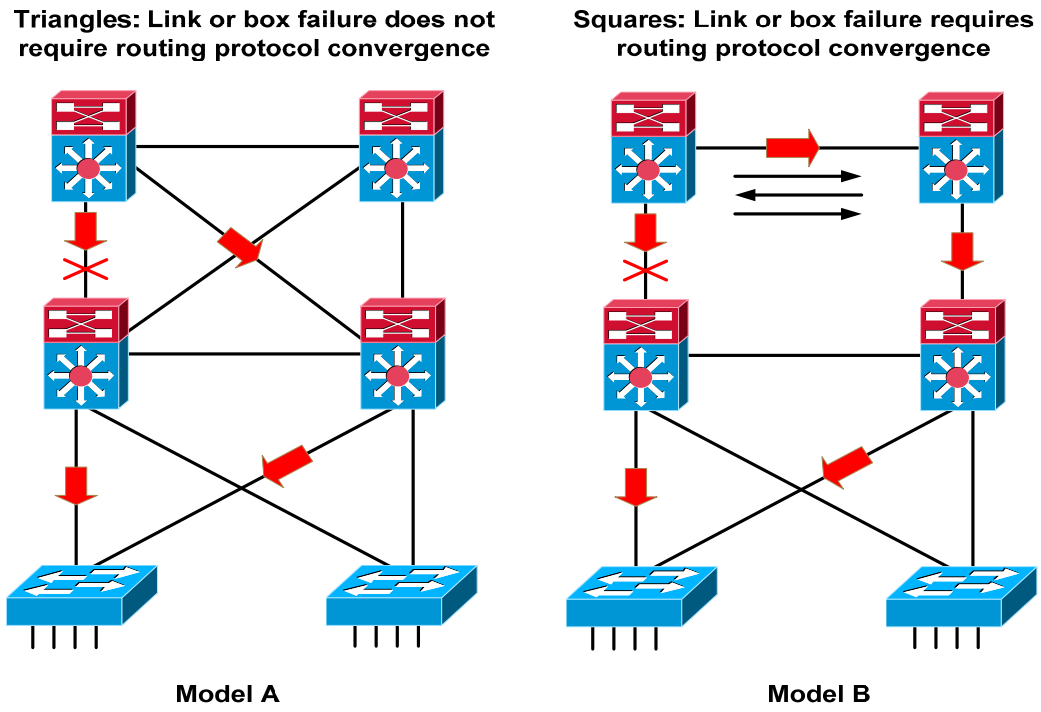
2.3 Избор на маршрутизиращ протокол

В този раздел се правят препоръки за избор на маршрутизиращ протокол в мрежата на фирмения комплекс на предприятието. Маршрутизиращите протоколи обикновено се използват във връзките между разпределителния слой и ядрото. както и в самото ядро. Комутиране от слой 3 се среща понякога и в слоя за достъп, но това не е честа практика. Маршрутизиращите протоколи се използват за бързо пренасочване по обходни пътища в случай на повреда в някои възли или връзки, като същевременно осигуряват балансиране на натоварването използвайки резервните маршрути.

2.3.1 Изграждане на редундантни триъгълници

За да се възползваме от предимствата на редундантните пътища с еднаква цена и да получим бърза детерминирана сходимост на протоколите между разпределителния слой и ядрото, препоръчително е да изградим редундантни триъгълници, а не квадрати.

Топологията свързваща комутаторите на разпределителния слой с тези в ядрото трябва да бъде изградена с помощта на триъгълници и пътища с еднаква цена към всички редундантни възли. Този дизайн с триъгълници е показан на Фиг.7, Model A, и използва по два пътя с еднаква цена, като по този начин се избягва базираната на таймери, недетерминистична сходимост на протоколите. Вместо непрякото откриване на пропадане на маршрут с използване на служебните пакети “hello” и изтичане на таймаутите на различните таймери, то при този дизайн загубата на физическата връзка се установява хардуерно, връзката веднага се означава като неизползваема и трафикът се пренасочва по алтернативния път. В този случай няма нужда маршрутизиращите протоколи OSPF или EIGRP да извършват изчисления и да търсят нов път.



Фиг.7 Изграждане на редундантни триъгълници

За разлика от този дизайн, топологията с квадрати, показана на фиг.7, Model B, изисква при повреда на връзките да се извърви целия път на сходимост на алгоритъма. Да отбележим, че е възможно да се изгради и топология, която не разчита на резервни пътища с еднаква цена. Това обикновено се прави за да се компенсира ограничения брой оптични връзки или за намаление на разходите. При нея обаче не може да се постигне същата детерминистична сходимост и поради тази причина този дизайн не се препоръчва като оптимален от гледна точка на висока надеждност.

2.3.2 Установяване на партньорство само по транзитните връзки.

Друга препоръчителна практика е да ограничите партньорството и актуализациите в слоя за достъп само по транзитните връзки.

По подразбиране, комутаторите в разпределителния слой изпращат маршрутизиращи актуализации и се опитват да установят партньорство с комутаторите в слоя за достъп за всяка VLAN. Това не е необходимо и води до загуба на процесорно време.

На Фиг.8 е показана мрежа с 3 комутатора в слоя за достъп, във всеки от тях по 4 VLAN, което означава, че са формирани 12 ненужни съседства. Необходимо е само партньорство по връзката между двата комутатора в разпределителния слой. От гледна точка за висока достъпност, това само затормозява протокола харчейки излишно процесорно време, памет и трафик от актуализации. Да не говорим за това, че системата е станала твърде сложна. Да отбележим, че в случай на повреда, трафикът може да се насочи през съседния комутатор за достъп, което е нежелано.

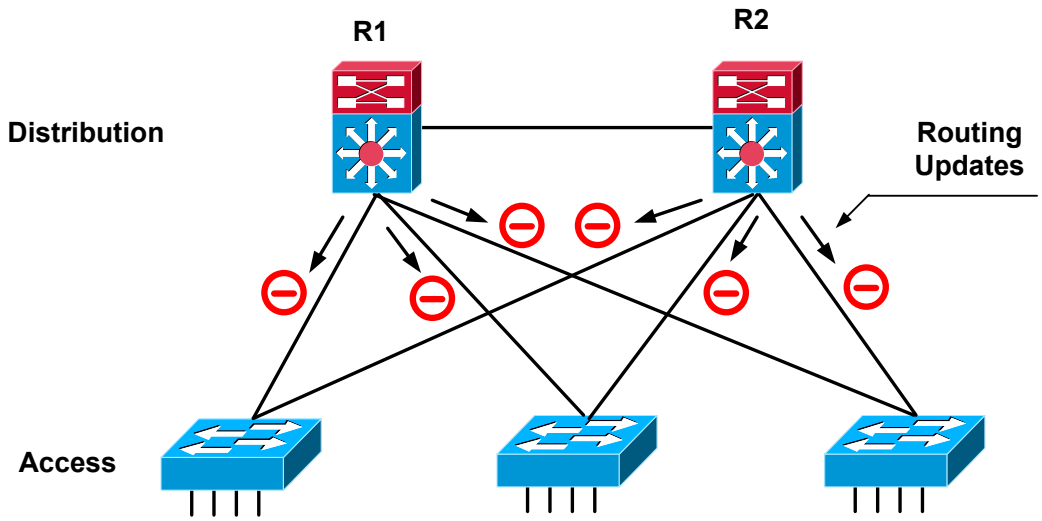


Figure 8 Използване на пасивни интерфейси в триъгълниците на слоя за достъп

Следователно може да се препоръча да конфигурираме портовете към комутаторите от слой 2 в слоя за достъп като пасивни и по този начин да потиснем рекламирането на маршрутна актуализация. Ако комутаторът от разпределителния слой не получава маршрутни актуализации за потенциален партньор по даден интерфейс, то той няма какво да обработва и няма да установява съседство с потенциални партньори по този интерфейс.

Съществуват два подхода за конфигуриране на пасивни интерфейси към комутаторите в слоя за достъп:

- Използвайте командата *passive-interface default*, а след това избирателно използвайте командата *no passive-interface* за да разрешите установяването на съседство където това е необходимо.

- Използвайте командата *passive-interface* за да направите определени интерфейси пасивни.

Пример за конфигуриране на пасивни интерфейси за OSPF:

R1(config)# ***router ospf 1***
R1(config-router)# ***passive-interface Vlan 99***

или

R1(config)# ***router ospf 1***
R1(config-router)# ***passive-interface default***
R1 (config-router)# ***no passive-interface Vlan 99***

Пример за конфигуриране на пасивни интерфейси за EIGRP:

R1(config)# ***router EIGRP 1***
R1(config-router)# ***passive-interface Vlan 99***

или

R1(config)# ***router EIGRP 1***
R1(config-router)# ***passive-interface default***
R1(config-router)# ***no passive-interface Vlan 99***

Това са само няколко команди, но с тях значително можете да опростите трафика и да получите желаните резултати.

2.3.3 Обобщаващи маршрути в разпределителния слой

Използването на йерархична маршрутизацията намалява трафика от актуализации на маршрути и избягва ненужните маршрутни изчисления. Такава йерархия се постига като организираме IP мрежите в съседни блокове, които лесно могат да се обобщят от динамичния маршрутизиращ протокол.

Препоръчителната практика е да се конфигурира един обобщен маршрут в разпределителния слой, и той да се рекламира за достъп до много IP мрежи в сградата (до блок от комутатори). В резултат на това се рекламират само няколко маршрута през ядрото. Ако маршрутизиращата информация не е обобщена в ядрото, то при повреда в някой възел, EIGRP и OSPF ще трябва да взаимодействат потенциално с голям брой участници. а това се отразява на времето за сходимост.

Обобщаването на връзките в разпределителния слой оптимизира процеса на пренасочване. Ако една връзка към устройство в слоя за достъп пропадне, обратният трафик от разпределителния слой към това устройство се отстранява докато не се постигне сходимост на протокола. Когато възел в разпределителния слой използва обобщени връзки, то той не рекламира към ядрото загубата на свързаност към определена подмрежа или VLAN. Това означава, че ядрото не знае че не трябва да изпраща трафик към комутатора в разпределителния слой където имаме дефектирала връзка за достъп. Обобщенията ограничават броя на възлите които EIGRP трябва да разпитва или броя на рекламираните за състоянието на връзките, които OSPF трябва да обработва, следователно ускоряват процеса на пренасочване.

Обобщаването трябва да се извършва на границата, където разпределителния слой на всяка сграда се свързва с ядрото на мрежата. Методът за конфигуриране на обобщени връзки е различен в зависимост от използвания протокол. Дизайнът изисква

съществуването на връзки от слой 3 между разпределителните комутатори. Това позволява на разпределителния комутатор нямащ достъп до дадена VLAN или подмрежа да пренасочи трафика към съседния разпределителен комутатор именно по тази връзка. За да бъде процесът ефективен, избраното адресно пространство за връзките между комутаторите в разпределителния слой трябва да бъде в адресното пространство на обобщените връзки. Следователно обобщаването трябва да се предхожда от солиден анализ на използваната схема за адресиране.

2.4 Резервиране на първия скок

Резервирането на първия скок (first-hop redundancy) или както още е известно като резервиране на шлюза по подразбиране (default-gateway redundancy), е важен компонент в процеса на сходимост при йерархичния дизайн на високо достъпните мрежи. Този компонент позволява мрежата да се възстанови при дефект в устройство, функциониращо като шлюз по подразбиране за възлите в един физически сегмент. Когато слоят за достъп е изграден с комутатори от слой 2, комутаторите от разпределителния слой действат като шлюзове по подразбиране за целия домейн от слой 2 който поддържат (както е показано на Фиг.9).

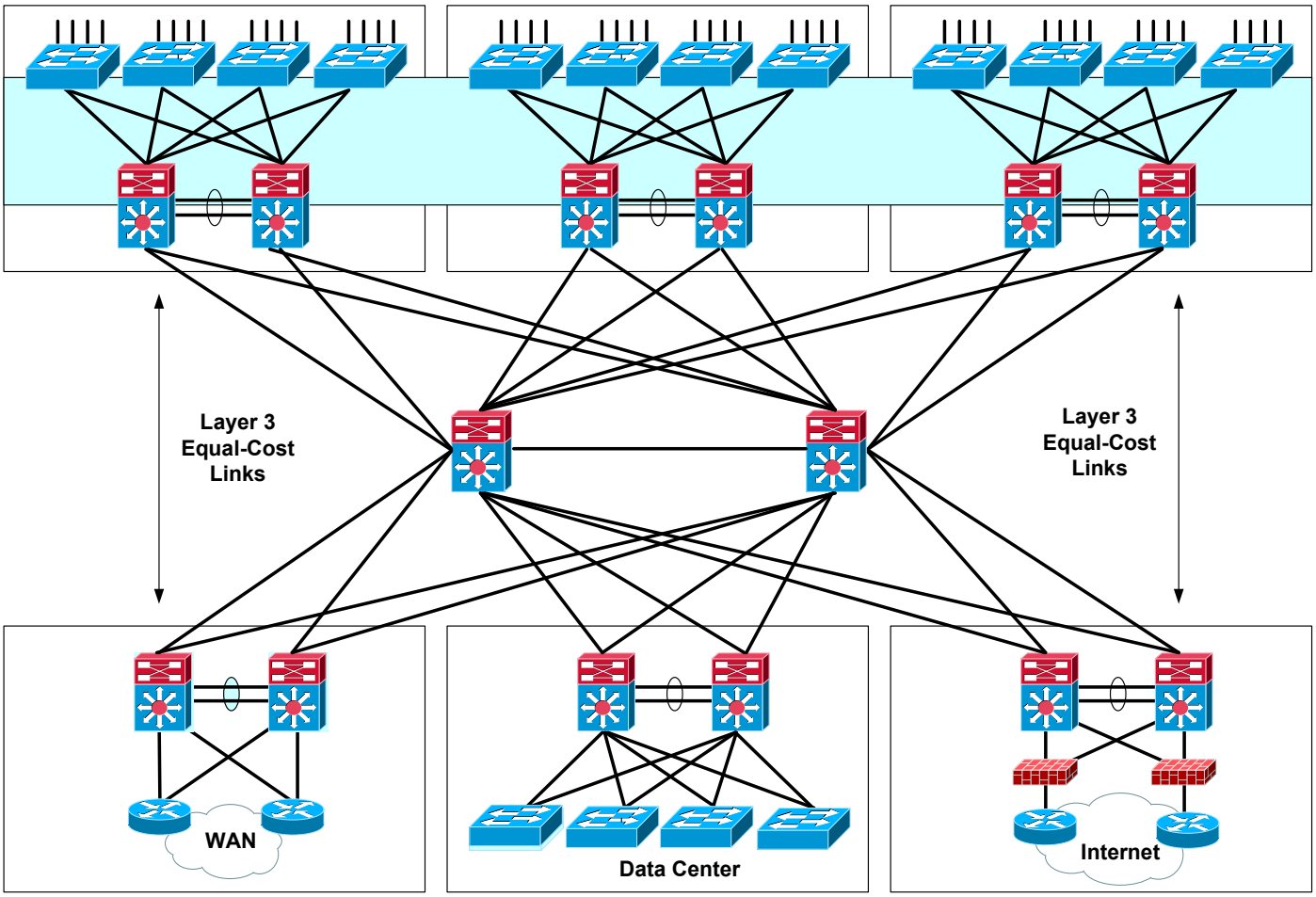
Нужда от протокол за резервиране на първия скок имаме само ако дизайнът използва връзки от слой 2 между комутаторите от слоя за достъп и комутаторите в разпределителния слой. Ако използваме слой 3 в комутаторите за достъп, то шлюзът по подразбиране за крайните устройства се намира в слоя за достъп, в тези комутатори.

В схемите на Cisco обикновено използваме като протокол за резервиране на първия скок (First Hop Redundancy Protocol – FHRP) протоколът Hot Standby Router Protocol (HSRP) [4]. Друг протокол – Virtual Router Redundancy Protocol (VRRP) е стандарт на Internet Engineering Task Force (IETF) за резервиране на шлюза по подразбиране [5]. При някои внедрявания започнаха да използват GLBP [6], при който по-лесно се получава балансиране на натоварването по връзките между слоя за достъп и разпределителния слой, резервирането на шлюза по подразбиране, както и защитата от повреди.

Когато са настроени правилно, и двата протокола, HSRP и VRRP, предоставят ефективен метод за резервиране на шлюза по подразбиране и осигуряват превключване към резервния комутатор в разпределителния слой за по-малко от секунда. В обкръжение на Cisco, HSRP е за препоръчване, тъй като по-бързо в него се внедряват нови функционалности. VRRP е логическият избор когато е необходима оперативна съвместимост с устройства на други доставчици.

Таймерите на HSRP и GLBP трябва да бъдат внимателно настроени, за да се постигне сходимост в рамките на 800 ms при повреда на връзка или възел на границата между слоя за достъп и разпределителния слой. Следният фрагмент от конфигурация показва как да променим параметрите по подразбиране на HSRP за да постигнем бърза сходимост:

```
interface Vlan5
ip address 10.1.5.3 255.255.255.0
ip helper-address 10.5.10.20
standby 1 ip 10.1.5.1
standby 1 timers msec 200 msec 750
standby 1 priority 150
standby 1 preempt delay minimum 180
```



Фиг.9 Триъгълници при резервирането на първия скок.

2.4.1 Настройка на времето за поемане на инициативата

Един важен фактор, който трябва да се вземе предвид когато използваме HSRP или друг протокол за настройка на резервирането на шлюза по подразбиране, е поемането на инициативата. Вследствие на поемането на инициативата (preemption), основният HSRP комутатор придобива отново главната роля след възстановяване от повреда. Такова възстановяване на водещата роля е желателно, тъй за дадена подмрежа или VLAN коренното устройство на протокола RSTP и основният шлюз на HSRP трябва да съвпадат. Ако те не съвпадат, т.е. HSRP и RSTP не са синхронизиране след повреда, връзката между разпределителните комутатори може да стане транзитна връзка и трафикът ще преминава през няколко възела до шлюза по подразбиране.

Времето за поемане на инициативата при HSRP трябва да бъде съгласувано с времето за зареждане и свързване на комутатора към останалата част от мрежата. Инициативата трябва да бъде поета едва след като:

- Интерфейсните модули (line cards) на комутатора започнат да предават пакети.
- Настъпила е сходимост на STP протокола и всички портове са в установено правилно състояние.
- Настъпила е сходимост на маршрутизиращия протокол, всички актуализации са завършени и правилните маршрути са фиксирани в маршрутизиращата таблица.

Възможно е инициативата да бъде поета от основния комутатор преди той да се свърже с ядрото. Ако това се случи, трафикът генериран в слоя за достъп ще бъде отстранен докато комутаторът не установи пълната връзка с ядрото.

Препоръчителната практика е да се измери времето за зареждане на системата, след което да увеличим това време с 50% и с командата *standby preempt delay minimum* да го въведем в комутатора. Това ни гарантира, че този първичен комутатор в разпределителния слой се е свързал с всички части на мрежата преди да поеме инициативата.

2.4.2 Елиминиране на FHRP в проектите използващи VSS

VSS комбинира дава физически комутатора в една логическа единица, и по този начин елиминира FHRP от разпределителния слой.

При традиционния дизайн, FHRP се разгръща в разпределителния слой, за да се гарантира, че шлюзът по подразбиране за крайните устройства остава достъпен, дори когато единият от двата разпределителни комутатора се повреди. За да се постигнат оптималните времена на възстановяване на мрежата в случай на повреда, трябва съответно да се настроят протоколите FHRP и STP, както и използвания маршрутизиращ протокол. Настройките на тези протоколи трябва да бъдат координирани помежду си.

При използването на VSS в разпределителния слой всъщност двата комутатора представляват един логически елемент, следователно вече не се нуждаем от FHRP. За изпълнение на неговите функции се грижат механизмите за възстановяване от срив, които се явяват вътрешни за VSS и MEC.

Ако един от двата разпределителни комутатора се повреди, това се равнява на загубата на половината от интерфейсните модули и на един надзорен модул във виртуалния комутатор. В плоскостта на предаване на данни, повредата се изразява в загуба на една от двете връзки на EtherChannel. Трафикът продължава да преминава по останалата връзка. В плоскостта за контрол на мрежата резервният надзорен модул играе активна роля. Тъй като той използва NFS и SSO, той поема всички отговорности на дефектиралия надзорен модул, включително и ролята на шлюз по подразбиране.

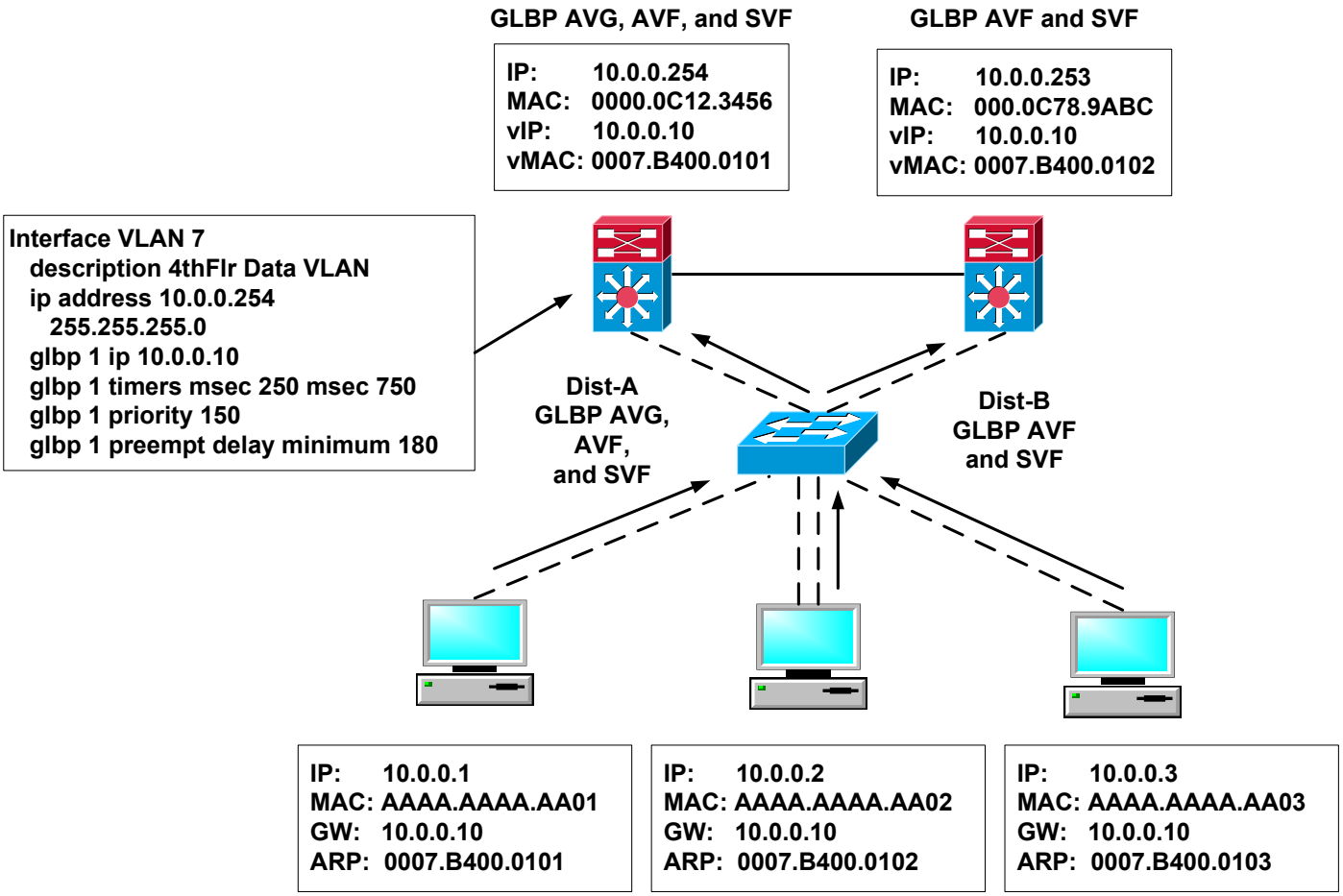
Елиминирането на FHRP от дизайна е още една полза от прилагането на VSS при свързването на слоя за достъп с разпределителния слой.

2.4.3 Преглед на Gateway Load Balancing Protocol (GLBP)

GLBP е протокол за резервиране на първия скок, който е разработен от Cisco, и позволява поток от пакети да бъде разпределян в група от резервни маршрутизатори.

Когато използваме HSRP или VRRP за да осигурим резервен шлюз по подразбиране, резервните комутатори не са активни и очакват да настъпи повреда, за да се намесят и да започнат активното предаване на трафик. Съществуват методи за използване на резервните връзки с HSRP или VRRP, но те са трудни за имплементиране. Няма да ги описваме тук, тъй като гледна точка на тяхното конфигуриране, поддръжка и управление срещаме изключителни затруднения.

GLBP осигурява всички предимства на HSRP, като допълнително включва балансиране на натоварването. При HSRP се задава един виртуален MAC адрес на крайните точки, когато те използват протокола ARP за да научат физическия MAC адрес на техния шлюз по подразбиране. GLBP позволява група от маршрутизатори да функционира като един виртуален маршрутизатор, като разделят един виртуален IP адрес и в същото време използват много виртуални MAC адреси за препредаване на трафика. На Фиг.10 е показан пример на конфигурация подкрепяща GLBP и неговата роля.



Фиг.10 GLBP триъгълници

Когато едно крайно устройство използва ARP за своя шлюз по подразбиране, то получава виртуален MAC адрес от GLBP активния виртуален шлюз (Active Virtual Gateway – AVG) на базата на кръгово избиране (round-robin). Тези шлюзове, които поемат отговорността за препращането на пакетите изпратени към техния виртуалния MAC адрес са известни като активни виртуални препращачи (Active Virtual Forwarders – AVF). Понеже трафика от една подмрежа преминава през резервни шлюзове, то всички връзки към разпределителния слой се използват.

Преодоляването на срив и конвергенцията при GLBP са подобни на тези в HSRP. Вторичният виртуален препращач (Secondary Virtual Forwarder – SVF) поема трафика предназначен за един виртуален MAC адрес засегнат от повредата и започва да го препраща. Крайният резултат е, че се постига по-равномерно използване на възходящите връзки , при това с минимално конфигуриране.

Да отбележим, че когато използваме GLBP в топологии, при които STP е блокирал възходяща връзка, можем да получим път в два скока за трафика от слой 2.

В среда, където една VLAN преминава през комутатори в разпределителния слой, HSRP е предпочитан протокол за резервиране на шлюз по подразбиране.

3. Литература

[1] Стоилов Емил, Проектиране на корпоративни мрежи. Част I Архитектура, Technical Report, Научен електронен архив на НБУ, 2014, http://eprints.nbu.bg/2219/

[2] Стоилов Емил, Проектиране на корпоративни мрежи. Част II Оптимизиране на слой 2, Technical Report, Научен електронен архив на НБУ, 2014, http://eprints.nbu.bg/2231/

[3] Стоилов Емил, Общи архитектурни концепции използвани в IP маршрутизаторите, Technical Report, Научен електронен архив на НБУ, 2013, http://eprints.nbu.bg/1770/

[4] Hot Standby Router Protocol Features and Functionality, http://www.cisco.com/c/en/us/support/docs/ip/hot-standby-router-protocol-hsrp/9234-hsrpguidetoc.html

[5] Virtual Router Redundancy Protocol (VRRP) https://tools.ietf.org/html/rfc3768

[6] GLBP – Gateway Load Balancing Protocol http://www.cisco.com/en/US/docs/ios/12_2t/12_2t15/feature/guide/ft_glbp.html