

A CRITICAL ANALYSIS OF "FABRIC" ARCHITECTURE IN DATA CENTERS.

*Ionka Gancheva, e-mail: igancheva@nbu.bg
New Bulgarian University, Department Informatics*

Abstract: This publication aims to study and analyze "fabric" architectures in the data center, with task to implement the most effective of them in a real environment. A critical analysis of the data center design using solutions from leading technology companies: Cisco - "FabricPath", Brocade - "VCS Fabric", Juniper - "Qfabric" is made, which allow open parallel access routes to different zones / "PODs". In conclusion, with implementation of a "FabricPath" between access and distribution layers in the data centers design will be achieving flexible, reliable, fast, scalable and highly available network architecture. Ensuring efficiency through simplified provisioning of virtual machines, as well as optimized use of network connectivity is one of the main goals.

Keywords

Data center, "FabricPath" technology, multipath networks, STP, Layer 2 loop prevention mechanism, Ethernet Fabrics, TRILL.

КРИТИЧЕН АНАЛИЗ НА "FABRIC" АРХИТЕКТУРИТЕ В ЦЕНТРОВЕТЕ ЗА ДАННИ.

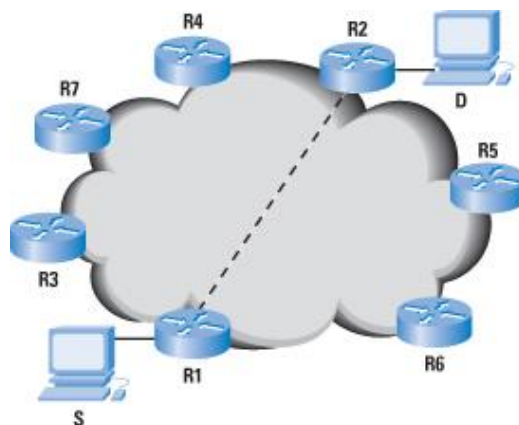
*Йонка Ганчева, e-mail: igancheva@nbu.bg
Нов български университет, департамент Информатика*

Абстракт: Настоящата публикация има за цел изследване, и анализиране на „fabric” архитектурите в центровете за данни, и задача за имплементирането на най-ефективната от тях в реална среда. Направен е критичен анализ на дизайна на центровете за данни при използването на решения на водещи технологични компании: Сиско - „FabricPath“, Brocade - „VCS Fabric“, Juniper - „Qfabric“, които дават възможност за много отворени паралелни пътища на достъп към различни зони/“PODs”. В заключение при имплементиране на „FabricPath“ между слоевете на достъп и дистрибуционен, в центровете за данни, се постига гъвкава, надежна, бърза, мащабируема и високо достъпна мрежова архитектура.

Ключови думи: център за данни, „FabricPath” технология, паралелни отворени пътища на достъп, “STP”, “Ethernet Fabrics”, “TRILL”.

Въведение

Традиционните мрежови архитектури са проектирани да осигурят висока наличност на достъп до статични приложения и сървъри. Тенденциите в ИТ технологиите като сървърната виртуализация, големи мащабируеми разпределени приложения изискват повече гъвкавост и надграждане. С виртуализацията центровете за данни станаха много динамични, с висока гъстота на сървъри и приложения. Виртуализацията и облачните услуги дефинират, че приложение, услуга, сървър, общо казано работно натоварване, може да съществува навсякъде, по всяко време, при поискване и да се премести на всяко място без прекъсване. Това е сериозна и дълбока промяна от старите мрежови архитектури, чрез

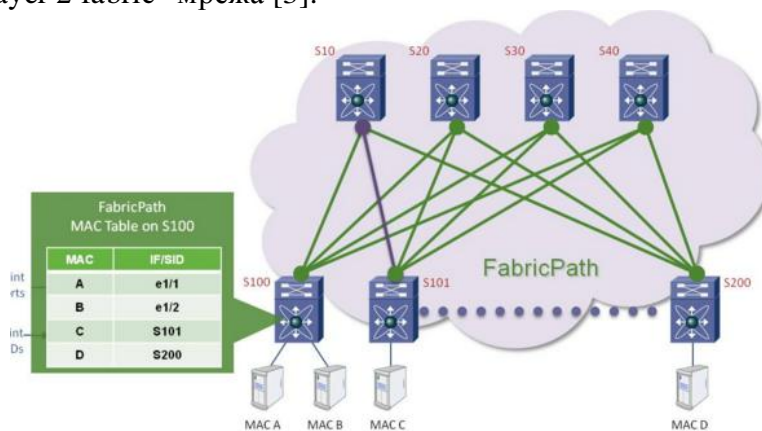


Фиг. 2 Технология TRILL с управление на "unicast" трафик.

„TRILL“ изчислява топологията на центъра за данни и препраща „layer 2“ пакетите като използва „IS-IS“ протокола. „TRILL“ включва функционалността на ефективно препращане на „unicast“ пакети, едновременно им препращане по много паралелни пътища и на „unicast“, и „multicast“ трафик. Целта на използването на „TRILL“ в центровете за данни е за да има на „Layer 2“ отворени паралелни повече пътища на достъп и да се елиминира използването на „Spanning tree“ протокол от гръбнака на мрежата. В допълнение предимството на технологията, е че „TRILL“ намалява и времето на закъснение на трафика. Той е и основата на създадените „fabric“ технологии днес. Целта на настоящата публикация е да се разгледат основните характеристики, анализират предимствата и недостатъците при използването на архитектурите „Fabric“ в центровете за данни и да се имплементира в реална среда най-ефективната от тях [1].

1. Сиско архитектура „FabricPath“.

Технологията „FabricPath“, базирана на „TRILL“ обединява предимствата на маршрутизирането с опростеността на комутирането, а именно: при комутирането: лесно конфигуриране, лесно свързване на нови устройства, гъвкаво разрастване; при маршрутизирането: едновременно използване на много пътища на достъп, бърза конвергенция, стабилност на мрежата и възможност за разрастване, за да изгради гъвкава и мащабируема „Layer 2 fabric“ мрежа [3].

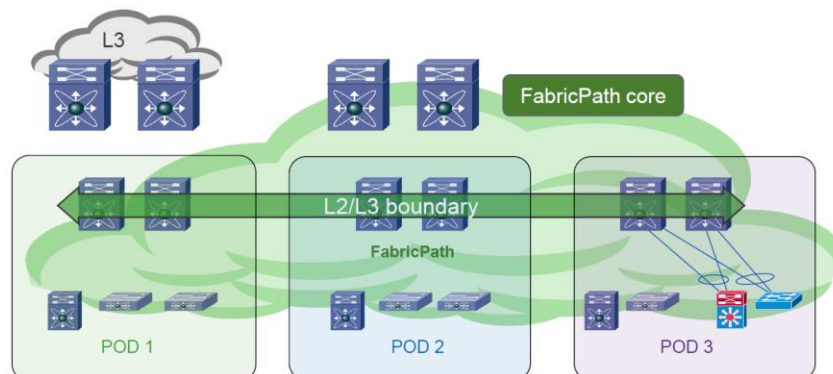


Фиг. 3 Изглед на топологията на център за данни, използваща „FabricPath“.

С използването на технологията „FabricPath“ се елиминира нуждата от „Spanning Tree“ протокол, постига се по-добра стабилност и конвергенция в мрежата, има опростена

конфигурация, използват се всички паралелни пътища на достъп между две точки, намалява се времето на закъснение на пакетите в сравнение с използването на типичния дизайн, постига се гъвкавост при планирането на “vlangs”, и не на последно място се постига висока мобилност на виртуалните машини [4].

Предимство при “layer 2” трафика технологията “FabricPath” може да използва 16 еднакви пътища на достъп между две точки. Ако пътищата са логически обединени линкове / “port channel” (16 links по 10Gb/s) се достига до 2,56 Tb/s “bandwidth” за комуникация. Пакетите се препращат по най-краткия възможен път до тяхната дестинация намалявайки по този начин времето за закъснение до максимум, в сравнение със “spaning tree” решението [2].



Фиг. 4 Технология Cisco “FabricPath” в топология разделена на зони.

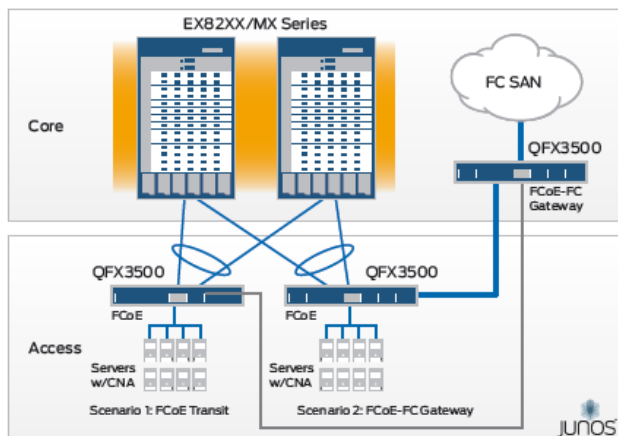
Имплементирането на Сиско “FabricPath” може да стане между “distribution” “core” слоевете на мрежата. На “core” слоя се маршрутизират пакетите между мрежите от различните зони/PODs. „Vlan” трафика свършва на дистрибуционния слой, не отива до “core”. В допълнение може да се използва еднаква “vlan” номерацията за различните “PODs”. Предимствата са, че „spaning tree“ протокола не работи между „POD“, чрез този дизайн отдалечени зони/PODs, дори и центрове за данни могат да бъдат свързвани, дава се възможност за увеличаване на честотната лента при нужда.

2. Juniper архитектура „Qfabric“

„Qfabric“ свива традиционният многослоен модел в еднослоен, където всички комутатори са свързани към едно устройство. Тази архитектурата е просто начин за свързване на количество “ethernet” комутатори в топология всеки със всеки / “full mesh” /. Предимство, е че комутаторите работят като едно логическо устройство с физическо разделение на трафика за управление и потребителския трафик, което елиминира нуждата от “spaning tree” протокола и “TRILL“. Всички комутатори знаят за всички крайни точки на останалите комутатори, което позволява да изпращат директно един към друг трафик, намалявайки значително времето на закъснение при пренос на пакета. Може да се мисли за “QFabric” като едно шаси, което може да се разрастне до 6000 10 Gb/s порта[8].

Елементи на архитектурата: в „QFabric“ архитектурата мрежовите карти са ситуирани в рамките на едно шаси и се наричат QF/Node. Те осигуряват достъп до и от фабриката. QF/Interconnect е шаси, което осигурява свързаността между картите или QF/Node. QFabric Предимството, е че фабриката се управлява от едно място, тоест едно устройство. Недостатък на технологията, е че ако това устройство падне, пропада цялата мрежа. Недостатък е, че няма ясна формулировка как се делегират права на достъп или права

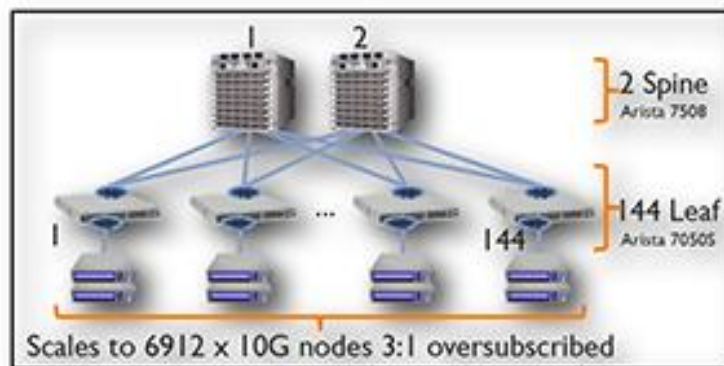
върху различни устройства от фабриката на различните мрежови администратори. Дали идеята за препращане на пакети от всеки към всеки комутатор е добра, но как се вземат решенията на къде да се препратят? Какво предотвратява безкрайното движение на пакети в мрежата? Как промените в топологията и маршрутизирането ще се отразят на фабриката? Въпроси на които се очаква отговор от производителя, за да може технологията да бъде приложена на практика.



Фиг 5 Архитектура Juniper QFabric

3. Arista - Архитектура

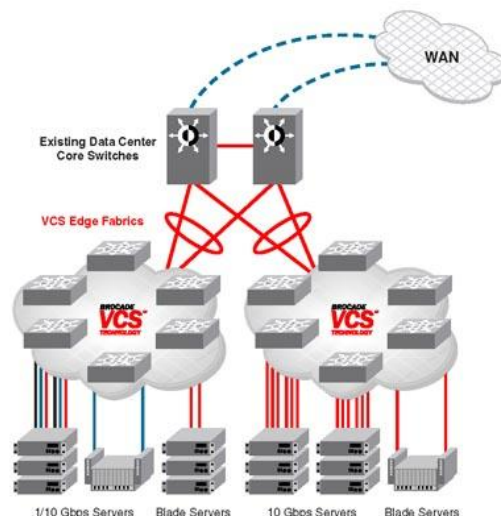
За архитектура на мрежа на център за данни Arista поддържа двуслоен модел, с фиксирани комутатори на слоя за достъп и модуларни комутатори на дистрибуционния слой. Модуларният комутатор има модул фабрик, чрез който се свързват всички мрежови карти. Той осигурява свързаност от всеки порт към всеки порт без загуба на пакети. Използват технолологията „MLAG / Multi-Chassis Link Aggregation /“, начин за обединение на паралелни свързани линкове към две шасита с едно управление. На дистрибуционния слой комутаторите използват технологията „VOQ / virtual output queue /“ за буферизиране на пакети, чрез която не се премахват пакети дори и при най-натовареният трафик. Според анализите времето на закъснение при препращане на пакетите е 600 наносекунди или 16 пъти повече ако се използва архитектурата на Сиско, където е 16 наносекунди[9].



Фиг. 6 Двуслоен модел на Arista за дизайн на център за данни.

4. Brocade архитектура – „VCS Fabric“

„VCS (Virtual Cluster Switching)“ технологията е базирана на „TRILL“, позволява създаването на мрежова фабрика, която конвергира потребителски трафик и трафик от масивите за данни използвайки обща комуникационна среда. Използва разширена архитектура за добавяне на нови услуги и възможности. Архитектурата осигурява гъвкава мрежова свързаност между индивидуални комутатори, с общо наименование фабрика[10]. Физическите комутаторите създават виртуален клъстер, видим отвън като единствен комутатор. Отново основният недостатък е при падането на клъстера, какво става с мрежата? Отново се очаква отговор от производителя на технологията.



Фиг. 7 Изглед на Brocade „VCS“ архитектура.

Заклучение

В статията са анализирани основните предимства и недостатъци на „fabric” архитектурите в центровете за данни на големите производители. След критичният анализ, в проекта за изграждане на виртуален център в НБУ ще бъде имплементирана Сиско фабричната архитектура като най-гъвкава, надеждна, най-ниска латенция, даваща възможност за много отворени паралелни пътища на достъп към различните зони/“Poll Of Devices”. Очаква се получената мрежова инфраструктура да бъде бърза, мащабируема, високо достъпна, с възможност за предоставяне на хостинг услуги до различните зони, включващи приложения, сървъри и други виртуални машини. Той също ще осигури ефективност чрез опростено провизиране на виртуалните машини, както и оптимизирано използване на мрежовата свързаност.

References

- [1] <http://datatracker.ietf.org/wg/trill/documents/> Интернет, 15.07.2015 г.
- [2] http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9402/white_paper_c11-687554.html Интернет, посетен на 15.07.2015 г.

- [3] http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9402/white_paper_c11-709336.html Интернет, посетен на 22.07.2015 г.
- [4] http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9670/guide_c07-690079.html Интернет, посетен на 25.08.2015 г.
- [5] http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9402/white_paper_c11-605488.html Интернет, посетен на 15.08.2015 г.
- [6] http://www.cisco.com/en/US/netsol/ns743/networking_solutions_program_home.html Интернет, посетен на 13.09.2015 г.
- [7] http://www.cisco.com/en/US/docs/switches/datacenter/nexus5000/sw/configuration_limits/limits_521/nexus_5000_config_limits_521.html Интернет, посетен на 10.10.2015 г.
- [8] <http://www.juniper.net/us/en/> Интернет, посетен на 15.08.2015 г.
- [9] <https://www.arista.com/en/company/news/press-release> Интернет, посетен на 15.08.2015 г.
- [10] http://community.brocade.com/t5/Design-Build/Data-Center-Infrastructure-Base-Reference-Architecture/ta-p/36885#_Toc384393794, Интернет, посетен на 15.08.2015 г.
- [11] <http://www.arista.com/en/>, Интернет, посетено на 15.09.2015 г.