

УДК 004.052

О.В. РЫЖКОВА

Таврический национальный университет им. В.И. Вернадского, Украина

ОЦЕНКА КРИТИЧНОСТИ ОТКАЗОВ КОММУТАТОРОВ УРОВНЯ ДОСТУПА КЛАСТЕРА ЦЕНТРА ОБРАБОТКИ ДАННЫХ

Рассматриваются различные топологии построения сети уровня доступа кластера центра обработки данных. Оценивается тяжесть последствий отказов коммутаторов уровня доступа. Предлагается использование гибридной топологии кластера дата-центра и методики выбора для этой топологии необходимого количества и характеристик коммутаторов в соответствии с минимальным уровнем критичности и оценкой накладных расходов. Выполняется анализ эффективности использования различных топологий с учетом надежности сети, количества серверов в стойках кластера, ценовой компоненты.

Ключевые слова: критичность отказов, коэффициент связности сети, кластер, ЦОД.

Введение

Наилучшие показатели готовности сети или информационно-управляемой системы (ИУС) можно получить, если минимизировать тяжесть последствий отказов и увеличить показатель связности сети. Количество серверов в стойках, измеряемое в юнитах (U), нормировано в соответствии со стандартом ANSI/EIA RS-310 D [1] и, как правило, составляет 10, 20, 42 юнита. Условно, в зависимости от количества стоек и серверов на уровне доступа, по масштабу дата-центры можно разделить на 3 группы: небольшие (в среднем, 40-80 стоек, например – ЦОД Tel Hosting, Россия, Москва, 60 стоек); средние (100-500 стоек, например – ЦОД Serverius, Голландия, 200 стоек по 45 юнитов); большие (около 1000 и более стоек, например, ЦОД компании Google, США, Дуглас, порядка 1000 стоек с 45000 серверами). Анализ информации от компаний, предоставляющих в аренду вычислительные и аппаратные ресурсы своих ЦОД (в том числе, для реализации облачных сервисов), показывает, что на сегодняшний день преобладают дата-центры (ЦОД) небольших и средних масштабов.

Основные рекомендации по проектированию физической топологии ЦОД собраны в стандартах TIA/EIA-942 [2], EN 50173-5 [3]. В соответствии с этими стандартами регламентированы 3 основные физические топологии уровня доступа кластеров ЦОД: Top of Rack (ToR), End of Row (EoR), Middle of Row (MoR). Однако нет четких рекомендаций по выбору той или иной топологии для конкретных реализаций. В различных источниках [4, 5] рассматриваются преимущества и недостатки каждой из них. В работах [6, 7] рассматриваются общие концепции организации сетевых топологий, в [8] пред-

лагается гибридная топология для кластера хранения данных.

Цель данной статьи – сравнительный анализ различных топологий организации сети уровня доступа кластера дата-центра с учетом оценки критичности отказов оборудования (коммутаторов). В работе предлагается использование гибридной архитектуры для заданного уровня критичности отказов коммутаторов а также учитывается стоимость развертывания различных вариантов архитектур. Для достижения поставленной цели необходимо решить следующие задачи:

1. Выполнить оценки тяжести последствий отказов коммутаторов уровня доступа различных архитектур кластера ЦОД.
2. Оценить уровень критичности отказов коммутаторов, сопоставив тяжесть последствий отказов в соответствии со шкалой классификации.
3. Выполнить расчет необходимого количества коммутаторов, которое обеспечит заданный уровень критичности.
4. Оценить среднюю стоимость сетевого оборудования уровня доступа дата-центра для реализации рассматриваемых топологий.
5. Сделать вывод о целесообразности выбора конкретной архитектуры или перехода на нее.

1. Оценка тяжести последствий отказов распределенной ИУС

В качестве показателя для оценки тяжести последствий отказов и распределенной ИУС, в основе которой лежит компьютерная сеть в работе используется коэффициент, характеризующий степень сохранения связности сети при возникновении отказов. Этот показатель может быть представлен в виде:

$$r_i = \frac{\sum_{l=1}^{D'} S(c_l, T^l)}{\sum_{l=1}^D S(c_l, T^0)} = \frac{\sum_{l=1}^{D'} S(c_l, T^l)}{D(D-1)},$$

где D – количество конечных узлов (серверов, компьютеров) сети при начальной топологии T^0 , свободной от отказов; D' - количество конечных точек сети при топологии T' , полученной в результате возникновения отказа i -го элемента (коммутатора), S - количество компьютеров, с которыми связан l -й конечной узел с топологии T' . Тяжесть последствий каждого отказа может быть определена как разница между единицей и значением коэффициента связности: $u_i = 1 - r_i$.

2. Основные физические топологии построения кластера дата-центра

ToR топология [4] (рис. 1) предполагает размещение отдельного коммутатора (либо двух для обеспечения избыточности) на вершине каждой стойки серверов.

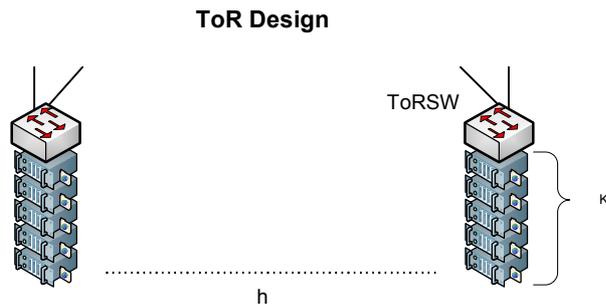


Рис. 1. Топология ToR

Если в кластере m серверов, которые распределены в одинаковом количестве k на h стоек, то коэффициент связности сети с такой топологией при отказе любого коммутатора ToRSW может быть рассчитан следующим образом:

$$r_i = \frac{(m-k)(m-k-1)}{m(m-1)} = \frac{(h-1)(kh-k-1)}{h(kh-1)}.$$

Следовательно, тяжесть последствий таких отказов будет определяться выражением

$$u_i = 1 - \frac{(h-1)(kh-k-1)}{h(kh-1)} = \frac{2kh+k+1}{h(kh-1)}.$$

Выберем для выполнения расчетов следующие данные. Пусть количество стоек $h=50$, в каждой из которых располагается $k=10$ серверов, тогда

$$r_i = 0,9603, u_i = 0,039.$$

В этом случае видно, что тяжесть последствий отказов коммутаторов ToRSW не критична, так как

этот показатель может варьироваться в пределах от 0 до 1 и значение 0,039 не является критичным для систем высокой готовности.

Другая топология (рис. 2) предполагает использование одного (двух для избыточности) коммутатора в конце (EoR) или в середине (MoR) ряда стоек с серверами, к которому подсоединены все сервера массива стоек [5].

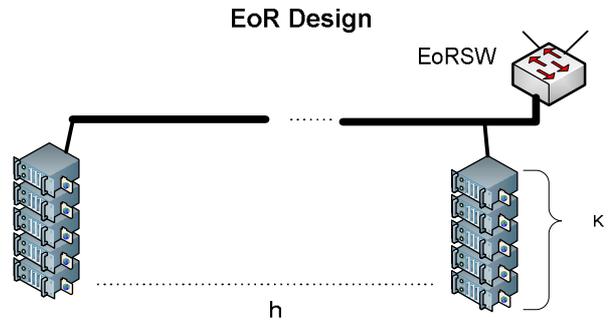


Рис. 2. Топология EoR

В таком исполнении вся нагрузка ложится на свитч EoR. Если произойдет отказ недублированного коммутатора EoRSW, то

$$r_i = 0, u_i = 1.$$

С точки зрения связности сети, топологии ToR и EoR/MoR представляют две крайности: в первом случае связность сети максимальна, но требуются большие расходы на оборудование и большее количество портов на коммутаторе верхнего уровня (агрегации), к которому непосредственно подключены ToR-коммутаторы уровня доступа. Во втором – требуется меньшее количество портов на коммутаторе уровня агрегации и меньшие затраты на коммутатор уровня доступа, но связность сети при этом будет минимальна. В работе анализируется промежуточный (гибридный) вариант между топологиями ToR и EoR, который будет являться своеобразным балансом с учетом накладных расходов, экономии портов и, в то же время, обеспечит минимально-приемлемый уровень критичности отказов коммутаторов уровня доступа (рис. 3).

Такая топология предполагает подключение нескольких стоек серверов к отдельным коммутаторам уровня доступа, которые могут быть расположены в середине ряда. Актуальной задачей построения этой архитектуры является определение количества коммутаторов уровня доступа K_{sw} и количества стоек x , подключенных к этим коммутаторам, которое обеспечит приемлемые оценки критичности отказов коммутаторов уровня доступа и тяжести последствий отказов, а также характеристик этих коммутаторов, таких, как средняя наработка на отказ (MTBF).

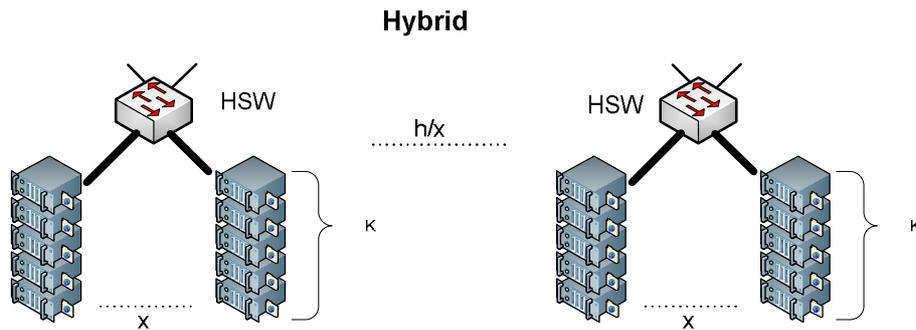


Рис. 3. Топологія Hybrid

3. Влияние критичности отказов на выбор количества коммутаторов уровня доступа

Согласно стандарту ГОСТ 27.310-95 (международный стандарт МЭК 812) [9], тяжести последствий отказов можно оценить по 10-бальной шкале, в соответствие которой сопоставить 4-х уровневую шкалу уровня критичности этих отказов: «высокий», «средний», «низкий», «очень низкий», как показано в табл. 1:

Для обеспечения высокой готовности системы предполагаем, что уровень критичности отказов компонентов (коммутаторов) должен соответствовать значению «очень низкий» или «низкий». Пусть требуемое значение тяжести последствия отказа коммутатора HSW $u_i = g$. Тогда коэффициент связности

$$\tau_i = 1 - g, \text{ т.е. } x = \frac{2kh - 1 - \sqrt{1 - 4kh(1-g)(1-kh)}}{2k}, \text{ а}$$

необходимое количество коммутаторов HSW уровня доступа $K_{sw} = \frac{h}{k} = \frac{2kh}{2kh - 1 - \sqrt{1 - 4kh(1-g)(1-kh)}}$.

Пусть требуемое значение тяжести последствия отказа коммутатора HSW $g = 0,1$, что соответствует степени критичности «очень низкий». Для рассматриваемого примера при $h=50, k=10$ необходимое количество коммутаторов $K_{sw} \approx 20, x \approx 2,5$. Следовательно, возможен вариант подключения двух стоек

на каждый коммутатор HSW. Для такой реализации потребуется 25 коммутаторов, что по сравнению с вариантом топологии ToR в 2 раза меньше. Вариант такой замены и перехода от топологии ToR к гибридной без ущерба надежности всей системы возможен только в случае, если уменьшение количества коммутаторов уровня доступа скомпенсируется их более высокими надежностными характеристиками, такими, как средняя наработка на отказ. В рассматриваемом примере в качестве ToR коммутаторов уровня доступа могут применяться модели Cisco Catalyst 3750-16TD, которые, согласно технической документации, имеют среднее время наработки на отказ 184422 часов. Для реализации гибридной архитектуры можно выбрать Cisco WS-C2960S-24TD-L с заявленным производителем MTBF=332958 ч. Сравнительная оценка расходов на коммутаторы уровня доступа топологий ToR и Hybrid представлена в табл. 2:

Цена на оборудование взята как среднее значение диапазона цен из каталогов интернет-магазинов на 15.02.2013. Видно, что применение гибридной топологии может существенно уменьшить расходы компании. Если на серверах используется технология NIC-Teaming и два сетевых интерфейса, такая избыточность предполагает дублирование коммутаторов уровня доступа. Следовательно, разница в стоимости рассматриваемых двух вариантов архитектуры будет еще более значительной. Также при выборе и оптимизации физической топологии уровня доступа стоит учесть еще несколько факторов:

Таблица 1

Шкала уровня критичности отказов по ГОСТ 27.310-95

Описание последствий отказов	Оценка последствий в баллах	Тяжесть последствий	Коэффициент связности	Уровень критичности
Отказ не приводит к заметным последствиям, пользователи вероятно не обнаружат наличие неисправности	1	0-0,1	1-0,9	«очень низкий»
Последствия отказа незначительны, но пользователи могут выразить неудовольствие его появлением	2-3	0,1-0,3	0,7-0,9	«низкий»
Отказ приводит к заметному для пользователей снижению эксплуатационных характеристик и/или к неудобству использования услуг	4-6	0,3-0,8	0,3-0,8	«средний»
Высокая степень недовольства пользователей, но угрозы безопасности отказ не представляет	7-8	0,8-1	0-0,2	«высокий»

Таблица 2

Сравнительная оценка расходов на коммутаторы уровня доступа

ToR			Hybrid		
Модель коммутатора	Кол-во	Цена за ед., \$	Модель коммутатора	Кол-во	Цена за ед., \$
Cisco Catalyst 3750-16TD	50	1300	Cisco WS-C2960S-24TD-L	25	1708
Итого 65000			42700		

1. При выборе количества и характеристик коммутаторов Hybrid-топологии необходимо учитывать пропускную способность, которая будет распределяться между серверами уровня доступа, и функции, такие как поддержка VLANs, LAG, протоколов MSTP и др.

2. При оценке стоимости аппаратного обеспечения следует учитывать количество портов, занятых на коммутаторах уровня агрегации, так как стоимость коммутатора напрямую зависит от количества портов. На уровне агрегации, как правило, функционируют устройства с оптическими портами 10Gbps для обмена потоками данных с уровнем доступа, на котором, в свою очередь, чаще всего используются порты 1Gbps (медные или оптические) [10]. Средняя цена за оптический порт со скоростью передачи данных 10 Гбит в секунду варьируется в зависимости от производителя и модели коммутатора. На коммутаторах Cisco она составляет около 800-1000\$ за порт, что существенно влияет на общую стоимость аппаратных ресурсов.

Рассмотрим следующий пример. Количество стоек $h=60$, количество серверов в стойках $k=14$. Коэффициент связности и тяжесть последствий отказов коммутаторов топологии ToR будут соответственно равны $r_1 = 0,966$, $u_1 = 0,033$. Рассчитаем количество коммутаторов гибридной топологии с сохранением показателя тяжести последствий отказов на уровне критичности «очень низкий» $g = 0,1$. Тогда, следуя вышеописанной методике расчета, $K_{sw} = 20$, $x = 3$. Следовательно, для сохранения необходимого коэффициента связности возможен вариант подключения 3 стоек к каждому коммутатору ($x=3$), т.е. 42 сервера на коммутатор. С учетом того, что исходящие порты (uplinks) коммутатора уровня доступа имеют пропускную способность 10 Гбит/с и для каждого сервера требуется скорость 1 Гбит/с (на серверах установле-

ны сетевые интерфейсы со скоростью передачи данных 1000 Мбит/с), общая пропускная способность 420 Гбит/с будет обеспечена минимум 4 портами uplinks 10 Gps на каждом коммутаторе. Но в этом случае гарантированная пропускная способность на сервер будет несколько занижена – 0,95 Gps. Так как ситуация одновременной загрузки всех каналов до максимального значения маловероятна, можно принять эти данные для выбора коммутатора, либо рассмотреть другой вариант подключения – двух стоек к каждому коммутатору. Для такой реализации будет достаточно 3 портов uplinks 10 Gps на коммутатор, чтобы гарантировано обеспечить 1 Гбит/с на сервер, при этом коэффициент связности сети будет увеличен до значения 0,92, а тяжесть последствий отказов коммутаторов уменьшена до 0,08. Одним из рекомендуемых для дата-центров ToR коммутаторов, который подходит для рассматриваемого примера, является модель Cisco Nexus 2224TP Fabric Extender. С учетом рассматриваемых характеристик, количества портов и требуемых скоростей передачи данных, увеличения показателей надежности и средней наработки на отказ коммутатора для перехода на Hybrid-топологию можно применить модель Cisco Nexus 2248TP GE Fabric Extender. Оценка расходов с применением различных топологий представлена в табл. 3.

Таким образом, экономически более целесообразным будет переход на Hybrid -топологию с подключением каждой трех стоек серверов к отдельным коммутаторам уровня доступа Cisco Nexus 2248TP, но с точки зрения показателей связности сети, надежности и избыточности полосы пропускания предпочтительнее вариант Hybrid -топологии с подключением каждой двух стоек серверов к отдельным коммутаторам. При этом тяжесть последствий отказов коммутаторов уровня доступа для всех рассмотренных вариантов реализации физической топологии соответствует уровню «очень низкий».

Таблица 3

Сравнительная оценка расходов на коммутаторы уровня доступа

ToR			Hybrid (x =2)			Hybrid (x =3)		
Компонент	Кол-во	Цена за ед., \$	Компонент	Кол-во	Цена за ед., \$	Компонент	Кол-во	Цена за ед., \$
Cisco Nexus 2224TP	60	7000	Cisco Nexus 2248TP	30	10000	Cisco Nexus 2248TP	20	10000
SFP порт 10Gbps уровня агрегации	120	800	SFP порт 10Gbps уровня агрегации	90	800	SFP порт 10Gbps уровня агрегации	80	800
Итого 516000			372000			264000		

Выводы

Рассмотренная методика анализа и оптимизации физической топологии позволяет сделать выбор в пользу развертывания конкретной архитектуры сети уровня доступа кластера ЦОД с учетом критичности отказов коммутаторов и ценовой компоненты. По результатам исследований видно, что практически всегда экономически более целесообразен переход на гибридную топологию, но с точки зрения надежности сети более оправданным будет выбор топологии ToR. Поставленные задачи были решены. Данная методика эффективна для вариантов размещения не более 20-24 серверов в стойках кластера. Если в стойке 42U полноценно используется все пространство для размещения порядка 40 серверов, дизайн ToR будет являться оптимальным решением. Направлением дальнейших исследований является анализ применения данной методики расчета количества и подбора характеристик коммутаторов уровня доступа для различных типов сетей с распределенной инфраструктурой.

Литература

1. ANSI/EIA RS-310 D. REV EIA-310-C Cabinets, racks, panels and associated equipment [Text] - CE-5.0. – 1992.

2. TIA/EIA-942. Telecommunications Infrastructure Standard for Data Centers [Text]. – SP-3-0092, 2005. – 151 p.

3. EN 50173-5. Information technology - generic cabling systems - part 1: general requirements [Text]. - European Standards(EN), 2008. – 140 p.

4. Data Center Networking – Connectivity and Topology Design Guide. [Electronic resource]. – Available to: <http://www.enterasys.com/solutions/DataCenter.aspx> – 2011. – 26 p. – 20.01.2013 y.

5. Cloud Ready Data Center - Network DESIGN GUIDE [Text]. – Juniper Networks, Inc. – Nov., 2012. – 45 p.

6. Scheer, G.W. Selecting, designing, and installing modern Ethernet or serial digital data networks in electrical substations [Text] / G.W. Scheer, D.J. Dolezilek. Relay Protection and Substation Automation of Modern Power Systems, Cheboksary. – September 9-13, 2007.

7. Roberts, S. Understanding Ethernet Communication – Topologies [Text] / S. Roberts // KE2 Therm Solutions, Washington. – 2011. – P. 1 – 4.

8. Andrade, A. Hybrid Host/Network Topologies for Massive Storage Clusters [Text] / A. Andrade, U. Mun, D. H. Chung, A. Mohr // MSST 2007/ – P. 287-292

9. ГОСТ 27.310-95. Анализ видов, последствий и критичности отказов. Основные положения [Текст]. – Введ. 01.01.1997.- 14 с.

10. Data Center Design Considerations [Text]. – Cisco Data Center Infrastructure 2.5 Design Guide.

Поступила в редакцию 26.02.2013, рассмотрена на редколлегии 27.03.2013

Рецензент: канд. техн. наук, доцент С.А. Зуев, Таврический национальный университет им. В.И. Вернадского, Симферополь, Украина.

АНАЛІЗ КРИТИЧНОСТІ ВІДМОВ КОМУТАТОРІВ РІВНЯ ДОСТУПУ КЛАСТЕРА ЦЕНТРУ ОБРОБКИ ДАНИХ

О.В. Рыжкова

Розглядаються різні топології побудови мережі рівня доступу кластера центра обробки даних. Оцінюються тяжкість наслідків відмов комутаторів рівня доступу. Пропонується використання гібридної топології кластера дата-центру та методики вибору необхідної кількості і характеристик комутаторів у відповідності з мінімальним рівнем критичності і оцінкою накладних витрат. Виконується аналіз ефективності використання різних топологій з урахуванням надійності мережі, кількості серверів в стійках кластера, цінової компоненти.

Ключові слова: критичність відмов, коефіцієнт зв'язності мережі, кластер, ЦОД.

CRITICALITY FAILURES ASSESMENT OF SWITCHES IN CLUSTER OF DATA CENTER ACCESS LAYER

O.V. Ryzhkova

The various network topologies of the cluster of data center access layers are discussed. The severe consequences of the failure of access layer switches are evaluated. A selecting method in hybrid topology for the required number of switches and their properties in accordance with the minimum level of criticality and overhead assessments is proposed. Effectiveness of using various topologies taking into account reliability of the network, the number of servers in a cluster rack, cost is analyzed.

Key words: criticality failure, factor of network connectivity, cluster, Data Center.

Рыжкова Ольга Владимировна – ассистент кафедры компьютерной инженерии и моделирования Таврического национального университета им. В.И. Вернадского «ТНУ», Симферополь, Украина.