

habr.com

Принцип работы протокола MSTP

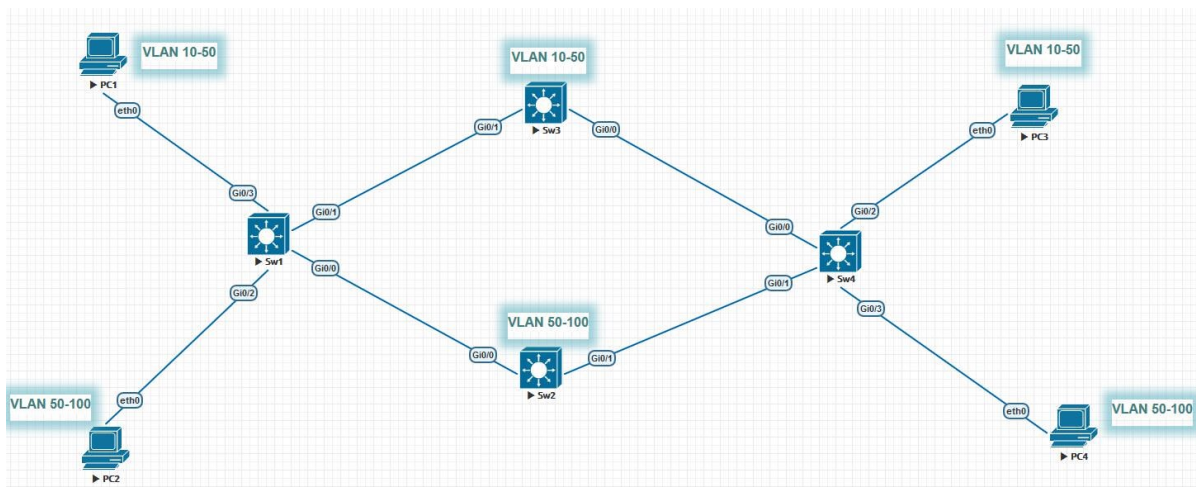
16-20 минут

Сегодня поговорим об MSTP. Перед тем, как разбираться с MSTP, надо ознакомиться с протоколами [STP](#) и [RSTP](#). MSTP является модификацией RSTP, а значит и STP. Если RSTP это тот же STP, только с более оптимизированной отправкой BPDU и в целом работы STP, то почему надо придумывать MSTP, который работает на основе RSTP? Основная фишка MSTP — это умение работать с VLAN-ми. Некоторые читатели могут сказать — «Подождите, а разве на Cisco pvst+ и rpvt+ не умеют работать с вланами?» RPVST+ и PVST+ просто запускает автономные экземпляры RSTP или STP в пределах одного влана. Но тут возникают проблемы:

- RPVST+ и PVST+ есть только на оборудовании Cisco, а на Cisco нет классического STP и RSTP. Что нам делать, если в топологии участвуют другие вендоры?
- Каждая инстанция STP и RSTP отправляет каждые две секунды BPDU. Если на транзитном порте пропускаются 100 вланов, то, значит, будет 100 сообщений в 2 секунды отправляться. Что не слишком хорошо.
- На Cisco есть ограничение в количестве экземпляров STP или RSTP на одном коммутаторе в зависимости от модели. То есть добавив 128 вланов на каком-то коммутаторе, мы столкнемся с

таким ограничением. Ссылка [тут](#)

Все эти проблемы успешно решает мультивендорный протокол MSTP. Постараемся разобраться с его работой. Схема построения топологии без петель в RSTP заключается в построении графа без пересечений вне зависимости от логической топологии, не учитывая где и как настроены вланы. В MSTP мы получаем возможность объединить определенные вланы в группу и для каждой группы построить отдельную топологию. Например, взглянем на такую топологию:



PC выполняют роль сегментов сети. PC1, PC3 — это сегмент сети, где используются вланы с 10 по 50, а PC2, PC4 - вланы с 50 по 100. На самих коммутаторах Sw1-4 созданы вланы с 10 по 100. Если мы будем использовать протокол RSTP, то, чтобы сегментам PC1 и PC2 добраться до сегментов PC3 и PC4 соответственно, необходимо использовать один общий путь, выбрав его из двух возможных: **Sw1-Sw2-Sw4** или **Sw1-Sw3-Sw4**. К сожалению, построить схему с распределением нагрузки не получится и один из этих путей будет заблокирован и пустовать. При помощи PVST+ и RPVST+, это получится сделать, но там свои проблемы, указанные выше. MSTP

приходит нам на помощь в данном вопросе, создавая RSTP инстанции для группы вланов 10-50 и 50-100. Можно настроить, что Root коммутатором для вланов 10-50 будет Sw3, а для 50-100 — Sw2. Путь **Sw1-Sw3-Sw4** будут использовать вланы 10-50, путь **Sw1-Sw2-Sw4** — вланы 50-100. Идея заключается в создании инстанций, которые будут объединять в себе вланы и строить дерево RSTP отдельно для каждой инстанции. Таким образом, объединив вланы 10-50 и 51-100 в разные инстанции, протокол MSTP позволит построить независимые деревья для каждой инстанции. Конфигурация будет выглядеть так:

```
spanning-tree mst configuration
name Note
instance 1 vlan 10-50
instance 2 vlan 51-100
```

Она будет идентична на всех 4-ех коммутаторах. Коммутаторы с идентичной конфигурацией данных параметров создают один регион. По поводу регионов будем говорить более подробно ниже, пока рассмотрим топологию в пределах одного региона. Для каждой инстанции выбирается собственный Root Bridge. На Sw3 введем команду **spanning-tree mst 1 root primary** для того, чтоб Sw3 был Root Bridge для вланов 10-50, а на Sw2 **spanning-tree mst 2 root primary** для вланов 51-100. Обобщая, каждая инстанция - **multiple spanning-tree instances (MSTI)** — объединяет в себе вланы. А каждый регион объединяет в себе коммутаторы, которые имеют одинаковые MSTI. Если говорить строго, то в регионе у коммутатора должны быть одинаковые следующие параметры:

- region name — название региона. Задается командой name.

- revision level — параметр изменения конфигурации.
- MSTI.

В каждом регионе есть инстанция MSTI 0 (instance 0), которая создана по умолчанию и в нее входят все вланы, которые не были включены в остальные инстанции. Instance 0 называется IST:

Internal Spanning Tree (IST) — специальная копия связующего дерева, которая, по умолчанию, существует в каждом MST-регионе. IST присвоен номер 0 (Instance 0). Она может отправлять и получать кадры BPDU и служит для управления топологией внутри региона. Все VLAN, настроенные на коммутаторах данного MST-региона, по умолчанию, привязаны к IST.

Root Bridge для IST называются Regional Root Bridge. Именно посредством IST передаются кадры BPDU, через которые строится дерево для каждой инстанции. Рассмотрим как выглядит BPDU в данном регионе:

```
Spanning Tree Protocol
Protocol Identifier: Spanning Tree Protocol (0x0000)
Protocol Version Identifier: Multiple Spanning Tree (3)
BPDU Type: Rapid/Multiple Spanning Tree (0x02)
BPDU flags: 0x38, Forwarding, Learning, Port Role: Root
Root Identifier: 24576 / 0 / 50:00:00:03:00:00
  Root Bridge Priority: 24576
  Root Bridge System ID Extension: 0
  Root Bridge System ID: 50:00:00:03:00:00 (50:00:00:03:00:00)
  Root Path Cost: 0
Bridge Identifier: 24576 / 0 / 50:00:00:03:00:00
  Bridge Priority: 24576
  Bridge System ID Extension: 0
  Bridge System ID: 50:00:00:03:00:00 (50:00:00:03:00:00)
Port identifier: 0x8002
Message Age: 0
Max Age: 20
Hello Time: 2
Forward Delay: 15
Version 1 Length: 0
Version 3 Length: 96
MST Extension
  MST Config ID format selector: 0
  MST Config name: Note
  MST Config revision: 0
  MST Config digest: 7ff3063c971d3b795396ed400d7fa9a2
  CIST Internal Root Path Cost: 20000
  CIST Bridge Identifier: 32768 / 0 / 50:00:00:01:00:00
    CIST Bridge Priority: 32768
    CIST Bridge Identifier System ID Extension: 0
    CIST Bridge Identifier System ID: 50:00:00:01:00:00 (50:00:00:01:00:00)
    CIST Remaining hops: 20
  MSTID 1, Regional Root Identifier 24576 / 50:00:00:03:00:00
    MSTI flags: 0x38, Forwarding, Learning, Port Role: Root
    0110 .... = Priority: 0x6
    .... 0000 0000 0001 = MSTID: 1
    Regional Root: 50:00:00:03:00:00 (50:00:00:03:00:00)
    Internal root path cost: 20000
```

```

Bridge Identifier Priority: 8
Port identifier priority: 8
Remaining hops: 20
MSTID 2, Regional Root Identifier 24576 / 50:00:00:02:00:00
MSTI flags: 0x7c, Agreement, Forwarding, Learning, Port Role: Designated
0110 .... = Priority: 0x6
.... 0000 0000 0010 = MSTID: 2
Regional Root: 50:00:00:02:00:00 (50:00:00:02:00:00)
Internal root path cost: 20000
Bridge Identifier Priority: 8
Port identifier priority: 8
Remaining hops: 19

```

До начала поля MST Extension, BPDU MSTP очень трудно отделить от BPDU RSTP и, грубо говоря, IST это есть классический RSTP. MSTP лишь добавляет данные о MSTI. В BPDU хранится информация о Root Bridge для Instance 0-2. Таким образом, для всех вланов и инстанций отправляется только один BPDU, который содержит всю необходимую информацию. Это огромная экономия по сравнению с PVST+ и RPVST+. Посмотрим вывод команды `show spanning-tree mst` на коммутаторе Sw2:

```

Sw2#show spanning-tree mst
##### MST0      vlans mapped: 1-9,101-4094
Bridge          address 5000.0002.0000 priority 32768 (32768 sysid 0)
Root            address 5000.0003.0000 priority 24576 (24576 sysid 0)
                 port Gi0/1 path cost 0
Regional Root   address 5000.0003.0000 priority 24576 (24576 sysid 0)
                 internal cost 40000 rem hops 18
Operational     hello time 2, forward delay 15, max age 20, txholdcount 6
Configured      hello time 2, forward delay 15, max age 20, max hops 20

Interface       Role Sts Cost      Prio.Nbr Type
-----
Gi0/0           Altn BLK 20000    128.1   P2p
Gi0/1           Root FWD 20000    128.2   P2p

##### MST1      vlans mapped: 10-50
Bridge          address 5000.0002.0000 priority 28673 (28672 sysid 1)
Root            address 5000.0003.0000 priority 24577 (24576 sysid 1)
                 port Gi0/0 cost 40000 rem hops 18

Interface       Role Sts Cost      Prio.Nbr Type
-----
Gi0/0           Root FWD 20000    128.1   P2p
Gi0/1           Altn BLK 20000    128.2   P2p

##### MST2      vlans mapped: 51-100
Bridge          address 5000.0002.0000 priority 24578 (24576 sysid 2)
Root            this switch for MST2

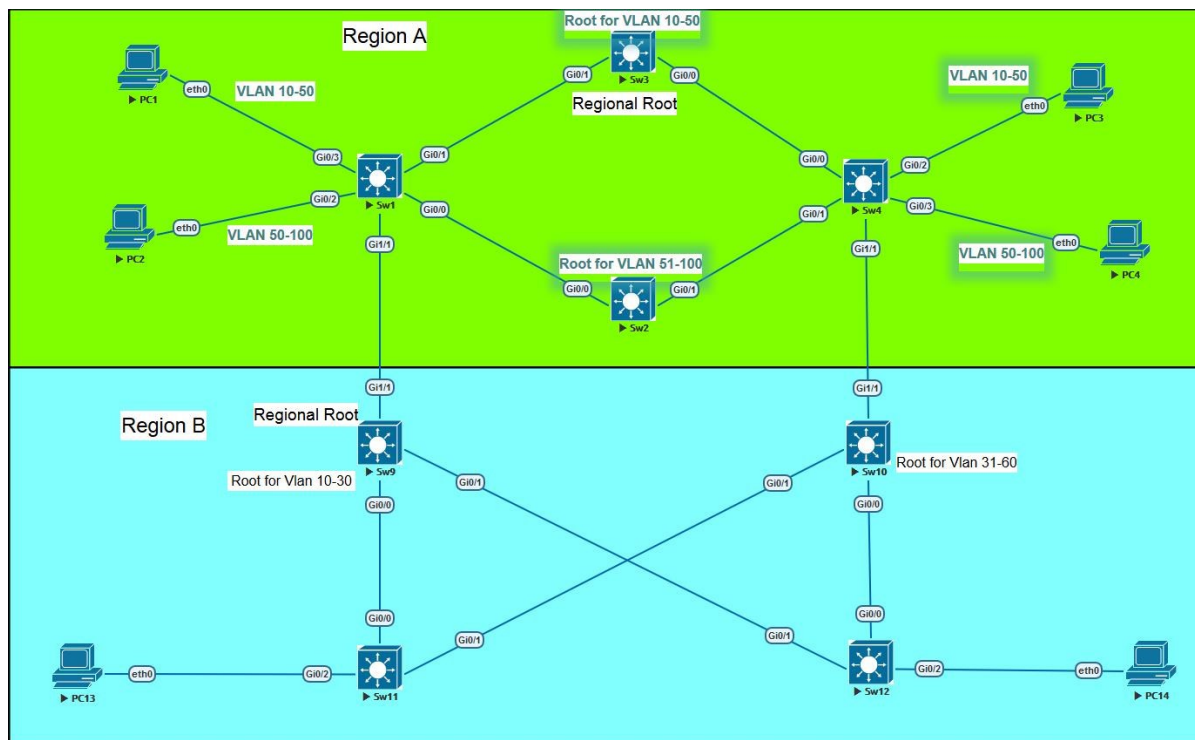
Interface       Role Sts Cost      Prio.Nbr Type
-----
Gi0/0           Desg FWD 20000    128.1   P2p
Gi0/1           Desg FWD 20000    128.2   P2p

Sw2#

```

Для instance 0 есть специальное поле — Regional Root. Regional Root мы выбрали Sw3 при помощи команды **`spanning-tree mst 0 root primary`**. Regional Root — это корневой коммутатор для MSTI 0 в пределах одного региона. Для MSTI1 Root также Sw3, а

для MSTI2 — Sw2. В плане блокирования портов и сходимости MSTP повторяет принципы RSTP на основе которого и работает, поэтому, думаю, работа MSTP в пределах одного региона довольно понятна. Рассмотрим топологию с двумя регионами:



Про Region A было сказано выше, теперь попытаемся разобраться как взаимодействуют между собой регионы. В region B у коммутаторов следующая конфигурация:

```
spanning-tree mst configuration
name RegionB
instance 1 vlan 10-30
instance 2 vlan 31-60
```

При этом Sw9 — **spanning-tree mst 1 root primary** — корневой коммутатор для Vlan 10-30, а Sw10 — **spanning-tree mst 2 root primary** — корневой коммутатор для Vlan 31-60.

Построение дерева STP в Region B аналогично Region A и было описано выше. Только скажем, так как мы не задавали Root

Bridge для MSTI 0 в Region B, то он будет выбран по наименьшему MAC-адресу среди Sw9-12. Наименьший MAC-адрес у Sw9. Вывод команды с Sw10:

```
Sw9#show spanning-tree mst
##### MST0      vlans mapped: 1-9,61-4094
Bridge          address 5000.0009.0000 priority      28672 (28672 sysid 0)
Root            address 5000.0003.0000 priority      24576 (24576 sysid 0)
                port    Gi1/1      path cost    20000
Regional Root   this switch
Operational     hello time 2 , forward delay 15, max age 20, txholdcount 6
Configured      hello time 2 , forward delay 15, max age 20, max hops 20

Interface       Role Sts Cost      Prio.Nbr Type
-----
Gi0/0           Desg FWD 20000    128.1    P2p
Gi0/1           Desg FWD 20000    128.2    P2p
Gi1/1           Root FWD 20000    128.6    P2p Bound(RSTP)

##### MST1      vlans mapped: 10-30
Bridge          address 5000.0009.0000 priority      24577 (24576 sysid 1)
Root            this switch for MST1

Interface       Role Sts Cost      Prio.Nbr Type
-----
Gi0/0           Desg FWD 20000    128.1    P2p
Gi0/1           Desg FWD 20000    128.2    P2p
Gi1/1           Mstr FWD 20000    128.6    P2p Bound(RSTP)

##### MST2      vlans mapped: 31-60
Bridge          address 5000.0009.0000 priority      28674 (28672 sysid 2)
Root            address 5000.000a.0000 priority      20482 (20480 sysid 2)
                port    Gi0/0      cost         40000      rem hops 18

Interface       Role Sts Cost      Prio.Nbr Type
-----
Gi0/0           Root FWD 20000    128.1    P2p
Gi0/1           Altn BLK 20000    128.2    P2p
Gi1/1           Mstr FWD 20000    128.6    P2p Bound(RSTP)

Sw9#
```

Для MSTI 0 Regional Root 5000.0009.0000 (Sw9). Ниже будет обговорено почему именно Sw9 и почему приоритет или мак-адрес тут не причем. Сейчас нас интересует больше вопрос, что будет происходить на границе. Чтоб дерево STP строилось корректно между регионами, посредством MST0 (IST) создается общее дерево для всех регионов. Такое дерево называется CIST. Давайте введем понятие также CST. Итак, сначала вспомним IST:

- **Internal Spanning Tree (IST)** – специальная копия связующего дерева, которая по умолчанию существует в каждом MST-регионе. IST присвоен номер 0 (Instance 0). IST может

отправлять и получать кадры BPDU и служит для управления топологией внутри региона. По умолчанию все VLAN одного региона привязаны к IST. Если в регионе создано несколько MSTI, то не ассоциированные с ними VLAN остаются привязанными к IST. В наших регионах A и B строится независимое дерево STP для вланов, которые не попали в инстанции 1 и 2. Root Bridge в IST называется Regional Root.

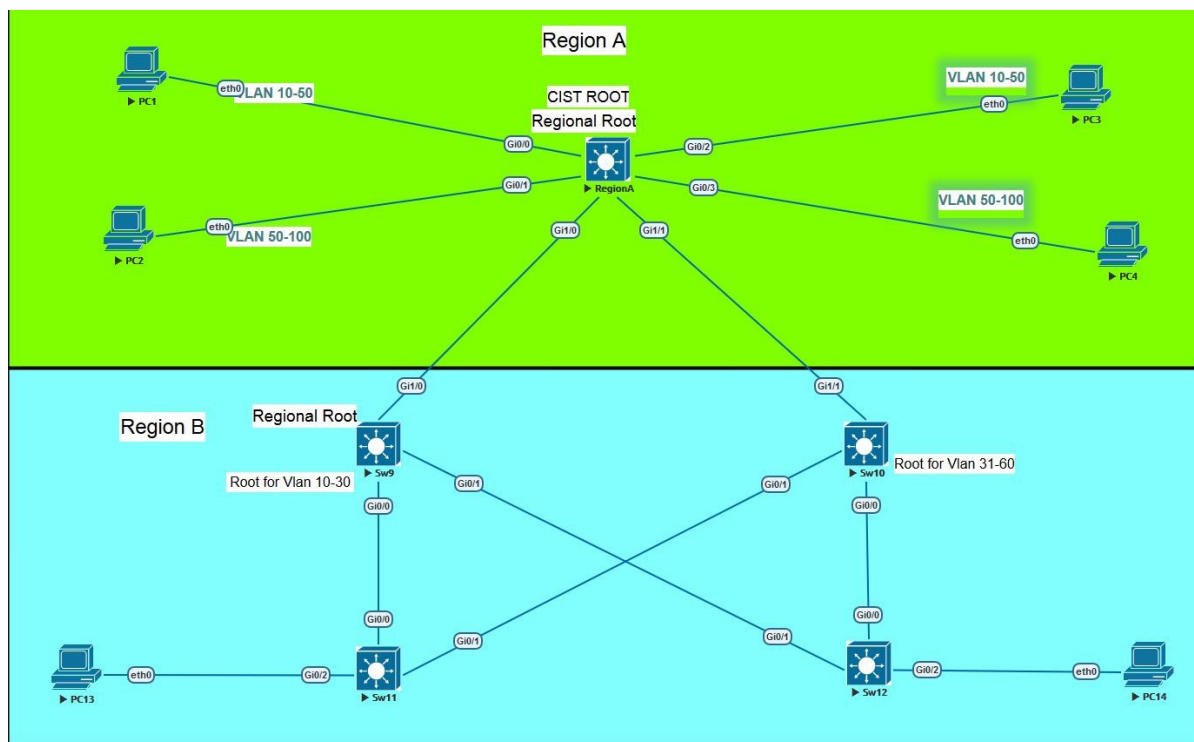
- **Common Spanning Tree (CST)** — дерево, соединяющее между собой регионы и все коммутаторы STP, RSTP, PVST+, RPVST+ (не MST коммутаторы) .
- **Common and Internal Spanning Tree (CIST)** — единое связующее дерево, объединяющее CST и IST каждого MST-региона.

Для понимания границ каждого дерева советую следующую [статью](#).

IST — это дерево в пределах одного региона, CIST — это дерево между регионами, CST — это дерево, объединившее в себе и деревья внутри региона, и дерево для соединения регионов.

Так как мы ввели команду **spanning-tree mst 0 root primary** на Sw3, то CIST Root Bridge для обоих регионов будет Sw3. Если во всей топологии всего один регион, то Regional Root Bridge и CIST Root Bridge совпадают. Если регионов много, то выбирается лучший Regional Root среди всех регионов. Также в выборах на роль CIST Root Bridge могут использоваться коммутаторы, которые используют протоколы отличные от MSTP. Если попытаться построить общую картину, то объяснить взаимодействие регионов можно так: Каждый регион

представляется объединение некоторого количества коммутаторов и представляется другим коммутаторам как один большой виртуальный коммутатор. То есть, если рассмотреть нашу топологию глазами региона В, то для него получится такая картина:



Аналогично будет и для региона А, регион В будет представляться одним коммутатором. В каждом регионе, у каждого коммутатора есть Root Port, подключающий его к Regional Root. Также у каждого региона выбирается один Root Port для MSTI 0, который ведет к общему CIST Root Bridge. Такими портами могут быть порты Gi1/1 на Sw9 и Sw10, так как они соединяют регионы. В нашей топологии, Sw9 обладает лучшим Bridge ID, то Root Port выбирается на нем, а на Sw10 порт Gi1/1 блокируется. На Sw9 для MSTI 0 порт Gi1/1 он является Root Port, но, например, для MSTI 1 и 2, есть Root Port для Instance Root Bridge и порт, который ведет к CIST Root Bridge, получает новую роль — Master. От одного региона к

другому может быть только один работающий канал или другими словами только один Master порт и только на одном коммутаторе. Вот информация о MST на коммутаторе Sw9, на котором будет выбран Master порт, обратите внимание на порт Gi1/1:

```
Sw9#show spanning-tree mst
##### MST0      vlans mapped: 1-9,61-4094
Bridge          address 5000.0009.0000 priority 28672 (28672 sysid 0)
Root            address 5000.0003.0000 priority 24576 (24576 sysid 0)
                port Gi1/1 path cost 20000
Regional Root   this switch
Operational     hello time 2 , forward delay 15, max age 20, txholdcount 6
Configured      hello time 2 , forward delay 15, max age 20, max hops 20

Interface      Role Sts Cost      Prio.Nbr Type
-----
Gi0/0          Desg FWD 20000    128.1    P2p
Gi0/1          Desg FWD 20000    128.2    P2p
Gi1/1          Root FWD 20000    128.6    P2p Bound(RSTP)

##### MST1      vlans mapped: 10-30
Bridge          address 5000.0009.0000 priority 24577 (24576 sysid 1)
Root            this switch for MST1

Interface      Role Sts Cost      Prio.Nbr Type
-----
Gi0/0          Desg FWD 20000    128.1    P2p
Gi0/1          Desg FWD 20000    128.2    P2p
Gi1/1          Mstr FWD 20000    128.6    P2p Bound(RSTP)

##### MST2      vlans mapped: 31-60
Bridge          address 5000.0009.0000 priority 28674 (28672 sysid 2)
Root            address 5000.000a.0000 priority 20482 (20480 sysid 2)
                port Gi0/0 cost 40000 rem hops 18

Interface      Role Sts Cost      Prio.Nbr Type
-----
Gi0/0          Root FWD 20000    128.1    P2p
Gi0/1          Altn BLK 20000    128.2    P2p
Gi1/1          Mstr FWD 20000    128.6    P2p Bound(RSTP)

Sw9#
```

Данный порт для MSTI 0, как мы говорили, имеет роль Root, а для MSTI 1-2 Master. Также вводится новый тип канала — P2p Bound (RSTP). Тип Boundary присваивается тем портам, которые стоят на границе с другим регионом или другой разновидностью протокола STP. Через Master порт в данный регион передается информация о CIST Root Bridge, отличительная черта заключается также в том, что данный порт сам по себе не отправляет BPDU, а только принимает, в отличие от типа порта P2P в RSTP. Исключением является только BPDU с флагом TC (изменение топологии). Посмотрим, как

коммутатором в одном регионе обрабатывается BPDU из другого региона. Как мы сказали на Master порте Gi1/1 Sw9 будут приниматься BPDU от Sw1, при этом сам Sw9 отправлять не будет, рассмотрим его еще раз:

```
Spanning Tree Protocol
  Protocol Identifier: Spanning Tree Protocol (0x0000)
  Protocol Version Identifier: Multiple Spanning Tree (3)
  BPDU Type: Rapid/Multiple Spanning Tree (0x02)
  BPDU flags: 0x7c, Agreement, Forwarding, Learning, Port Role: Designated
  Root Identifier: 24576 / 0 / 50:00:00:03:00:00
  Root Path Cost: 0 External Path Cost
  Bridge Identifier: 24576 / 0 / 50:00:00:03:00:00
  Port identifier: 0x8006
  Message Age: 0
  Max Age: 20
  Hello Time: 2
  Forward Delay: 15
  Version 1 Length: 0
  Version 3 Length: 96
MST Extension
  MST Config ID format selector: 0
  MST Config name: Note
  MST Config revision: 0
  MST Config digest: 7ff3063c971d3b795396ed400d7fa9a2
  CIST Internal Root Path Cost: 20000
  CIST Bridge Identifier: 32768 / 0 / 50:00:00:01:00:00
  CIST Remaining hops: 19
  MSTID 1, Regional Root Identifier 24576 / 50:00:00:03:00:00
    MSTI flags: 0x7c, Agreement, Forwarding, Learning, Port Role: Designated
    0110 .... = Priority: 0x6
    .... 0000 0000 0001 = MSTID: 1
    Regional Root: 50:00:00:03:00:00 (50:00:00:03:00:00)
    Internal root path cost: 20000 Internal Path Cost
    Bridge Identifier Priority: 8
    Port identifier priority: 8
    Remaining hops: 19
  MSTID 2, Regional Root Identifier 24576 / 50:00:00:02:00:00
    MSTI flags: 0x7c, Agreement, Forwarding, Learning, Port Role: Designated
    0110 .... = Priority: 0x6
    .... 0000 0000 0010 = MSTID: 2
    Regional Root: 50:00:00:02:00:00 (50:00:00:02:00:00)
    Internal root path cost: 20000 Internal Path Cost
    Bridge Identifier Priority: 8
    Port identifier priority: 8
    Remaining hops: 19
```

Sw1 рассылает BPDU один и тот же, независимо от того коммутатору внутри региона или за пределами региона отправляется BPDU. Sw9 приняв этот BPDU будет обрабатывать информацию только до начала поля MST Extension, только эта информация используется для построение дерева в инстанции

0. Здесь можно заметить интересный факт — Root Identifier (50:00:00:03:00:00) совпадает с Bridge Identifier (50:00:00:03:00:00), хотя отправляет пакет Sw1 с Bridge Identifier — 50:00:00:01:00:00. Это подтверждает нашу теорию о том, что каждый регион для другого представляет одним большим виртуальном коммутатором. Также можно сказать, что между регионами у нас работает классический RSTP. Но поговорим о Root path cost. Различают два вида Path Cost — внутренний и внешний. Root Path Cost, который указывает для MSTI 0, является внешним. И несмотря на то, что от Sw1 до Root Bridge (Sw3) есть еще один канал, он все равно указывается как нулевой. Как будто его отправил сам Root Bridge. Внутренние Root Path Cost указаны в полях ниже MST Extension. CIST Internal Root Path Cost указывает стоимость пути до Regional Root Bridge для MSTI 0, а в полях ниже MSTID 1 и 2 указаны Internal Root Path Cost до Root Bridge для каждой инстанции, соответственно. У внешнего Root Path Cost особая и очень важная роль, именно он будет определять Regional Root Bridge, а не приоритет и мак-адрес, об этом мы делали замечание выше. Тот коммутатор, у которого наименьший Root Path Cost до CIST Root Bridge и станет Regional Root Bridge! Например, Sw9 и Sw10 имеют Root Path Cost — 20000, равный стоимости порта Gi1/1. И только после этого начинают сравнивать свои Bridge ID. Давайте проверим это, изменим стоимость линка на Sw10 и при этом увеличим его приоритет до максимального, чтоб исключить его вариант выбора Regional Root по приоритету. Введем следующие команды:

```
Sw10(config-if)# spanning-tree mst 0 priority 61440  
Sw10(config-if)# interface gigabitEthernet 1/1
```



```
Sw10(config-if)# spanning-tree mst 0 cost 10000
```

И мы получим, что Sw10, несмотря на свой ужасный приоритет, стал Regional Root Bridge:

```
Sw10#show spanning-tree mst
##### MST0
Bridge      vlans mapped: 1-9,61-4094
address 5000.000a.0000 priority 61440 (61440 sysid 0)
Root      address 5000.0003.0000 priority 24576 (24576 sysid 0)
port Gi1/1 path cost 10000
Regional Root this switch
Operational hello time 2 , forward delay 15, max age 20, txholdcount 6
Configured  hello time 2 , forward delay 15, max age 20, max hops 20

Interface    Role Sts Cost      Prio.Nbr Type
-----
Gi0/0        Desg FWD 20000    128.1    P2p
Gi0/1        Desg FWD 20000    128.2    P2p
Gi1/1        Root FWD 10000    128.6    P2p Bound(RSTP)

##### MST1
Bridge      vlans mapped: 10-30
address 5000.000a.0000 priority 28673 (28672 sysid 1)
Root      address 5000.0009.0000 priority 24577 (24576 sysid 1)
port Gi0/1 cost 40000 rem hops 18

Interface    Role Sts Cost      Prio.Nbr Type
-----
Gi0/0        Altn BLK 20000    128.1    P2p
Gi0/1        Root FWD 20000    128.2    P2p
Gi1/1        Mstr FWD 20000    128.6    P2p Bound(RSTP)

##### MST2
Bridge      vlans mapped: 31-60
address 5000.000a.0000 priority 20482 (20480 sysid 2)
Root      this switch for MST2

Interface    Role Sts Cost      Prio.Nbr Type
-----
Gi0/0        Desg FWD 20000    128.1    P2p
Gi0/1        Desg FWD 20000    128.2    P2p
Gi1/1        Mstr FWD 20000    128.6    P2p Bound(RSTP)

Sw10#
```

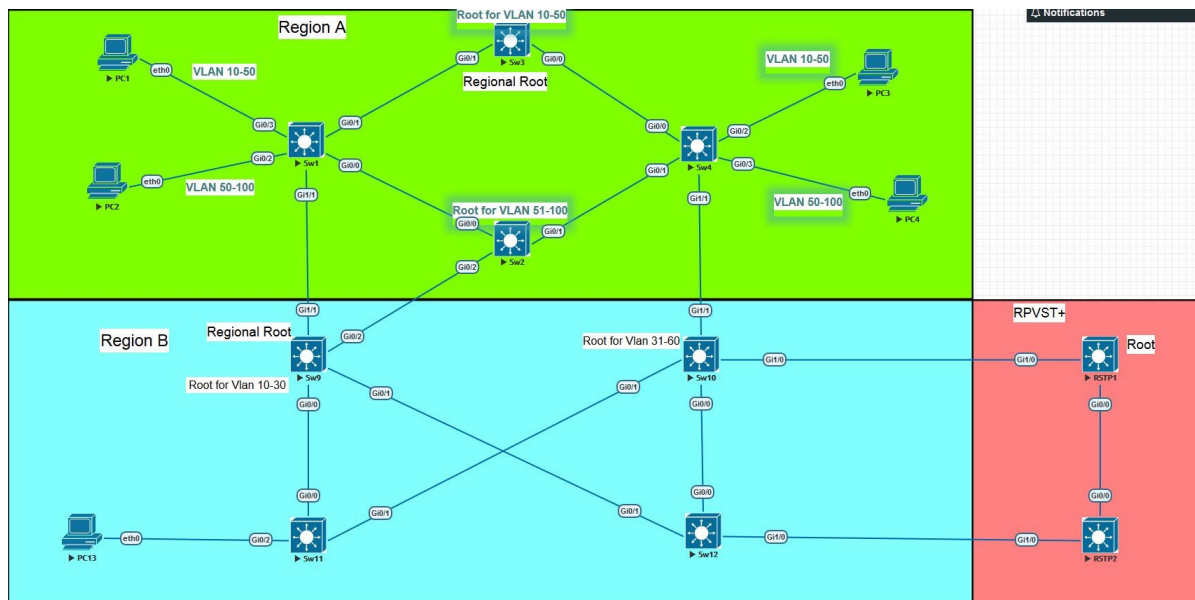
Таким образом, выбор Regional Root Bridge происходит в такой последовательности и никогда не может Regional Root Bridge быть коммутатор без граничных портов:

- Lowest External path cost to CIST Root bridge.
- Lowest Regional Bridge Identifier.
- CIST Bridge Identifier из BPDU от региона с CIST Root Bridge.

Если также, например, мы добавим линк между Sw9 и Sw4 и задумаемся какой линк станет Master, то при всех равных параметрах, определяющим станет поле CIST Bridge Identifier, которое ниже MST Extension. Мы немного соврали, когда говорили, что Sw9 не смотрит данные ниже MST Extension. В

данном случае, CIST Bridge Identifier из MST Extension будет учитываться.

Теперь посмотрим поведение MST в случае, когда мы добавим RPVST коммутаторы:



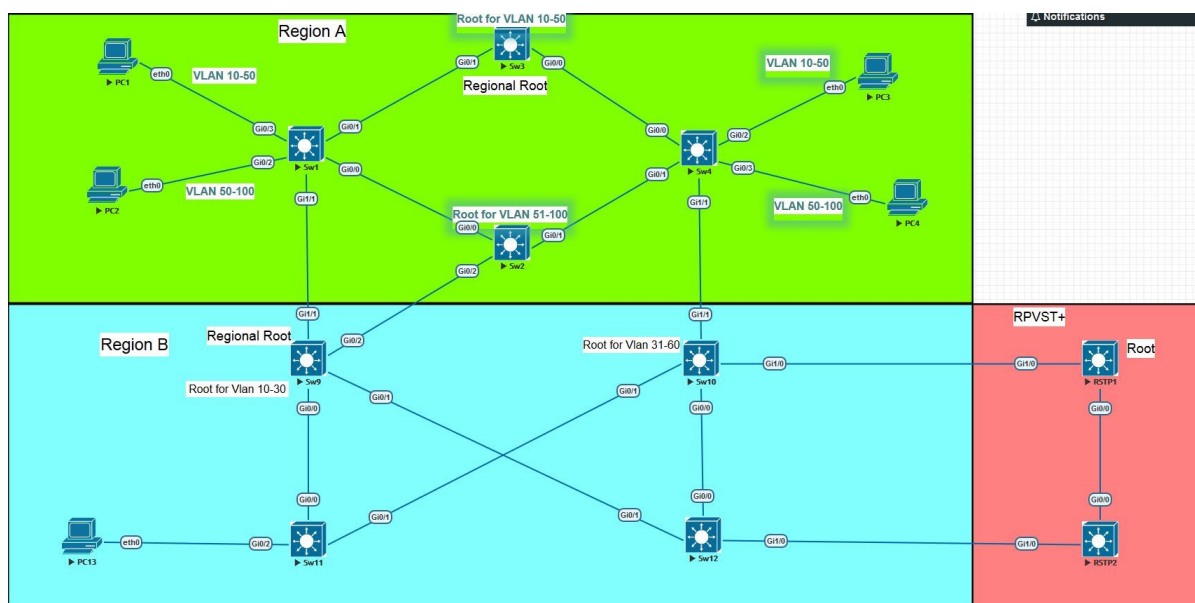
Возможны два случая:

1. Root Bridge принадлежит RPVST.
2. Root Bridge принадлежит MSTP.

Рассмотрим сначала случай, когда Sw3 продолжает быть CIST Root Bridge (то есть случай 2). В этом случае, Sw10 и Sw12 как только получают BPDU от RSTP1 и RSTP2 по каждому VLAN, то они изменяют port type Gi1/0 в **P2p Bound(PVST)**. Данный тип означает, что Sw10 и Sw12 на данном порту подключены к RPVST коммутаторам и начнут посылать копии BPDU для MSTI0 (то есть без полей ниже MST Extension) по всем VLAN, разрешенным на данном порту. Тем самым, RSTP1 и RSTP2 не заметят того, что Sw10 и Sw12 используют иной протокол.

Данная технология в MSTP называется **PVST Simulation** и при помощи нее мы получаем согласование между различными коммутаторами.

Случай же, когда Root Bridge находится среди RPVST коммутаторов более сложный и нерекommenованный. Представим, что у нас RSTP1 является Root Bridge для всех вланов. Для простоты будем считать, что созданы вланы 1-3 и введена команда **spanning-tree vlan 1-3 priority 12288**, самый низкий приоритет среди RSTP и MSTP коммутаторов. Sw10 начнет получать BPDU по каждому из вланов. Очень важно понимать, что MSTP коммутатором будет обработан только BPDU для Vlan 1. Как это происходит?



В зависимости от того, сколько вланов сконфигурировано, столько RPVST+ BPDU пакетов будет отправляться, но помимо этого, также будет отправляться по native vlan BPDU с мак-адресом назначения — 01:80:c2:00:00:00. Именно такой мак-адрес используют MSTP, PVST и RPVST коммутаторы используют другой мак-адрес — 01:00:0c:cc:cc:cd. Независимо от того, какой vlan сконфигурирован как native, в данном BPDU будет передаваться информация для первого влана. Только этот пакет будет обрабатываться для построения дерева коммутаторами MSTP. Остальные же BPDU для определенных вланов будут использоваться для проверки защиты корневого

узла при симуляции PVST (PVST simulations root-guard-based consistency check). Что за проверка? Как только мы настроим RSTP1 при помощи такой команды **spanning-tree vlan 1-3 priority 12288**, то на Sw10 у нас сразу же появится ошибка:

```
Sw10#
*Jan 27 19:20:14.561: %SPANNTREE-2-PVSTSIM_FAIL: Blocking root port Gi1/0: Inconsitent inferior PVST BPDU received on
VLAN 4, claiming root 12292:5000.000f.0000
Sw10#show spanning-tree mst 0

#### MST0    vlans mapped: 1-9,61-4094
Bridge      address 5000.000a.0000 priority 32768 (32768 sysid 0)
Root        address 5000.000f.0000 priority 12289 (12288 sysid 1)
            port Gi1/0 path cost 20000
Regional Root this switch
Operational hello time 2 , forward delay 15, max age 20, txholdcount 6
Configured  hello time 2 , forward delay 15, max age 20, max hops 20

Interface    Role Sts Cost      Prio.Nbr Type
-----
Gi0/0        Desg FWD 20000    128.1   P2p
Gi0/1        Desg FWD 20000    128.2   P2p
Gi1/0        Root BKN*20000    128.5   P2p Bound(PVST) *PVST_Inc
Gi1/1        Desg FWD 20000    128.6   P2p

Sw10#
```

Попытаемся объяснить. Sw10 принял по native vlan BPDU для vlan 1 с приоритетом 12288+1. Обработал и решил, что Gi1/0 его Root Port. Потом пришли PVST BPDU для остальных вланов (1-3), он их изучил для проверки целостности корневого коммутатора и там оказались приоритеты 12288+2, 12288+3 для вланов 2-3, которые больше чем 12288+1. Целостность разрушается — с одной стороны это должен быть Root Port, а получение других более больших приоритетов заставляет порт перейти в роль Designated. Такая двусмысленность непозволительна для таких протоколов и MST блокирует данный порт переводя его в состояние BKN, сообщая об ошибке — **Blocking root port Gi1/0: Inconsitent inferior PVST BPDU received**. Чтоб предотвратить такое, необходимо, чтоб ни один vlan, BPDU которого будут передаваться по данному порту, не имел приоритет хуже (больше), чем у vlan 1. То есть, если мы уменьшим приоритеты вланов 2-3 до 4096, заведомо меньше, чем у влана 1, то исправим данную проблему.

```
Sw10#
*Jan 27 19:35:00.056: %SPANNTREE-2-PVSTSIM_OK: PVST Simulation inconsistency cleared on port GigabitEthernet1/0.
Sw10#show spanning-tree mst 0

#### MST0    vlans mapped: 1-9,61-4094
Bridge      address 5000.000a.0000 priority 32768 (32768 sysid 0)
Root        address 5000.000f.0000 priority 12289 (12288 sysid 1)
            port Gi1/0 path cost 20000
Regional Root this switch
Operational hello time 2 , forward delay 15, max age 20, txholdcount 6
Configured  hello time 2 , forward delay 15, max age 20, max hops 20

Sw10#
```

```
Interface      Role Sts Cost      Prio.Nbr Type
-----
Gi0/0          Desg FWD 20000    128.1   P2p
Gi0/1          Desg FWD 20000    128.2   P2p
Gi1/0          Root FWD 20000    128.5   P2p Bound (PVST)
Gi1/1          Desg FWD 20000    128.6   P2p
Sw10#
```

Как видим, появилось сообщение, которое восстанавливает правильную работу порта — **PVST Simulation inconsistency cleared on port GigabitEthernet 0/1.**

На этом, думаю, завершить разговор о MSTP. Полезные ссылки внизу:

1. www.cisco.com/c/en/us/support/docs/lan-switching/multiple-instance-stp-mistp-8021s/116464-configure-pvst-00.html
2. habrahabr.ru/post/128172
3. www.cisco.com/c/ru_ru/support/docs/lan-switching/multiple-instance-stp-mistp-8021s/116464-configure-pvst-00.pdf
4. www.cisco.com/c/en/us/td/docs/switches/datacenter/nexus5000/sw/configuration/guide/cli_rel_4_0_1a/CLIConfigurationGuide/MST.html
5. networkengineering.stackexchange.com/questions/28716/multiple-spanning-tree-terminology-cst-ist-cist-and-exact-behavior
6. blog.sbolshakov.ru/11-3-mstp
7. blog.ine.com/2010/02/22/understanding-mstp
8. networklessons.com/spanning-tree/multiple-spanning-tree-mst