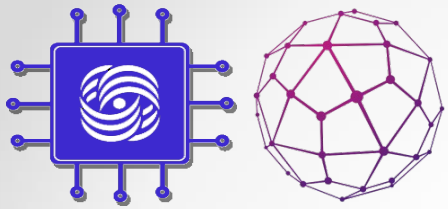
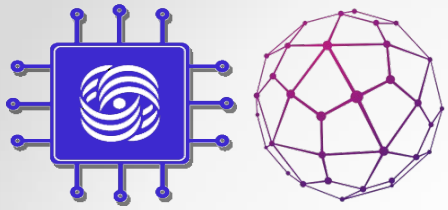


Layer 2 MPLS VPN

Introduction



Layer 2 VPN Types

- Virtual leased line (VLL)
- Pseudo-Wire Emulation Edge to Edge (PWE3)
- Virtual Private LAN Service (VPLS)

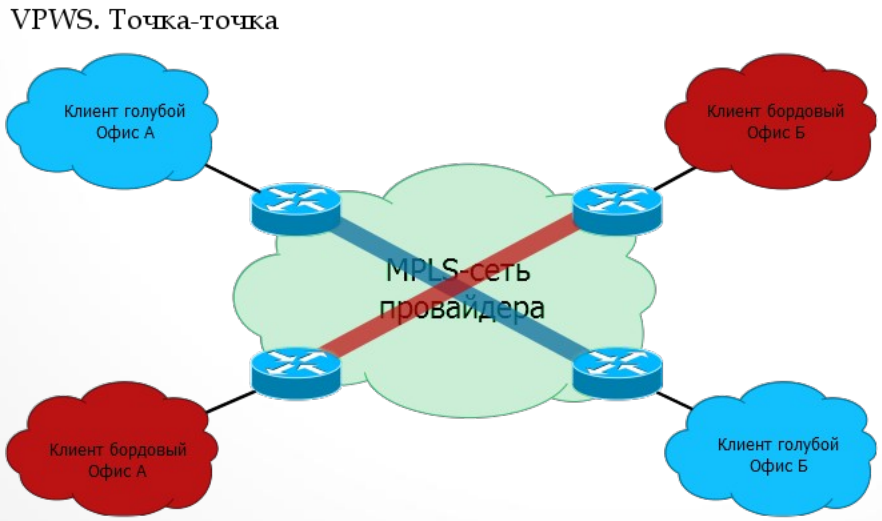
Layer 2 VPN Types

Для построения любого L2VPN существуют два концептуально разных подхода.

Point-to-Point. Применим к любым типам протоколов канального уровня

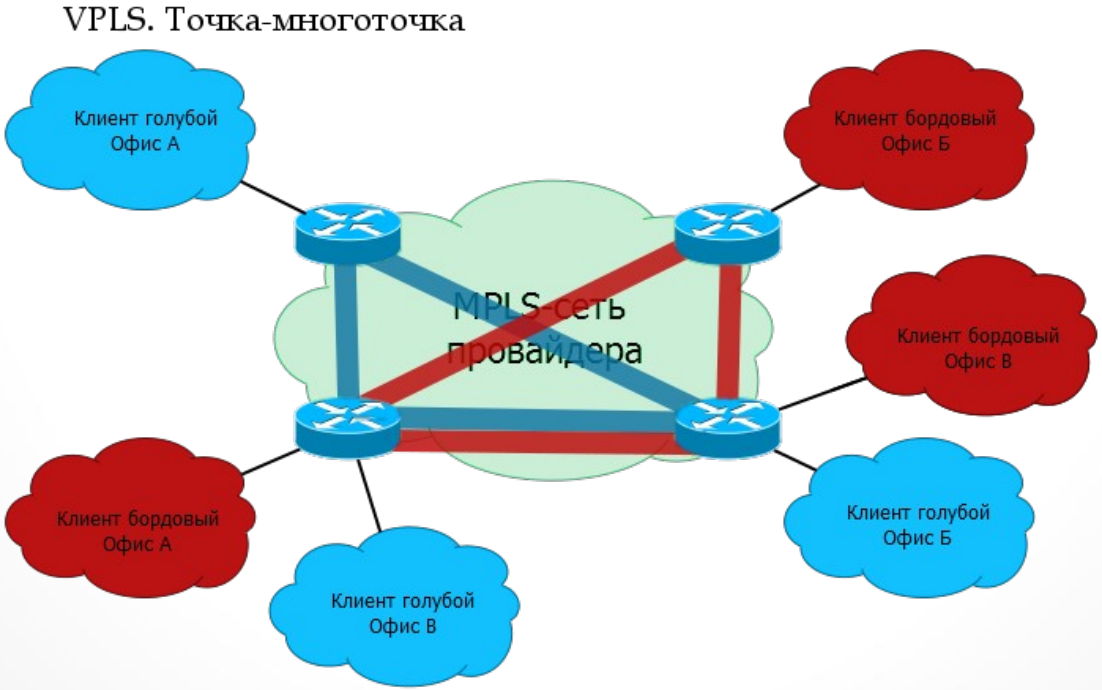
В основе лежит концепция PW — PseudoWire — псевдопровод. Соединение двух узлов друг с другом. Сеть провайдера можно рассматривать как один виртуальный кабель — то, что вошло в него на одном конце обязательно выйдет на другом без изменений.

Общее название услуги: VPWS — Virtual Private Wire Service.



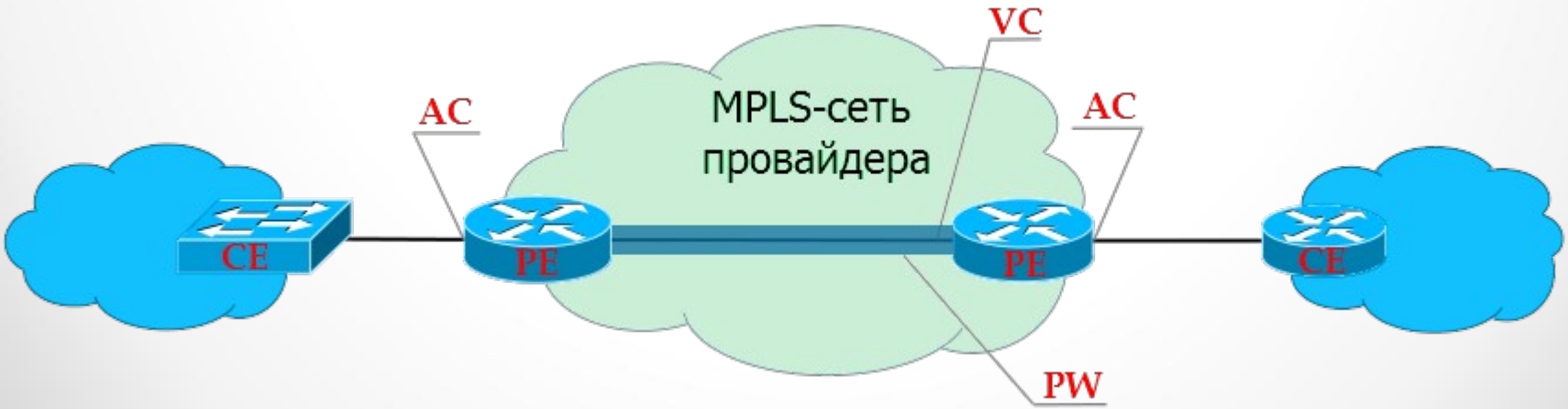
Layer 2 VPN Types

Point-to-Multipoint. Режим только для Ethernet. У клиента может быть несколько точек подключения/филиалов, и они должны передавать данные друг другу, причём, как одному конкретному филиалу, так и всем сразу. Можно рассматривать как виртуальный Ethernet-коммутатор

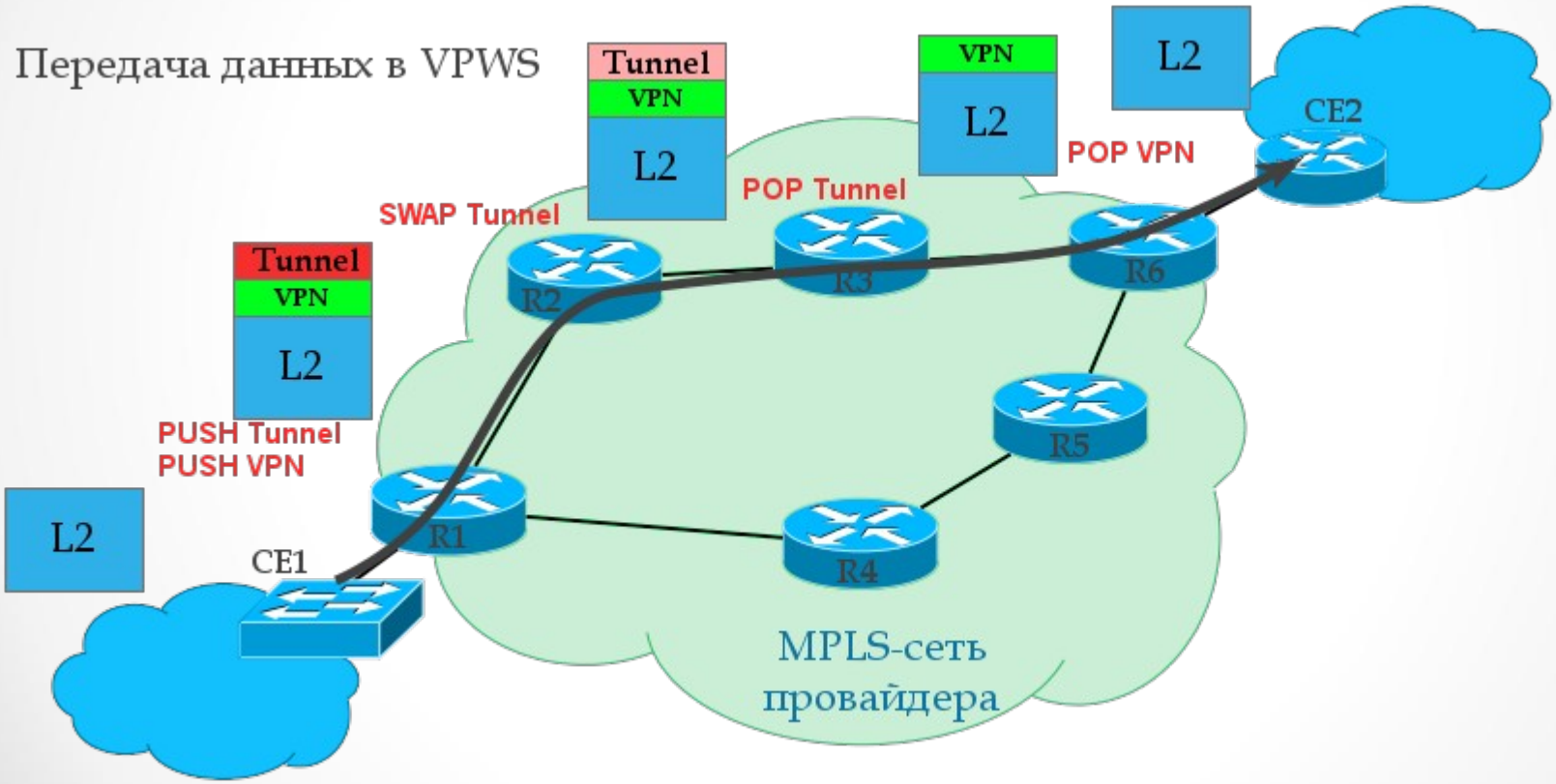


Терминология

PE — Provider Edge — граничные маршрутизаторы MPLS-сети провайдера, к которым подключаются клиентские устройства (CE).
CE — Customer Edge — оборудование клиента, непосредственно подключающееся к маршрутизаторам провайдера (PE).
AC — Attached Circuit — интерфейс на PE для подключения клиента.
VC — Virtual Circuit — виртуальное однонаправленное соединение через общую сеть, имитирующее оригинальную среду для клиента. Соединяет между собой AC-интерфейсы разных PE. Вместе они составляют канал: AC→VC→AC.
PW — PseudoWire — виртуальный двунаправленный канал передачи данных между двумя PE — состоит из пары однонаправленных VC.



VPWS Data Plane



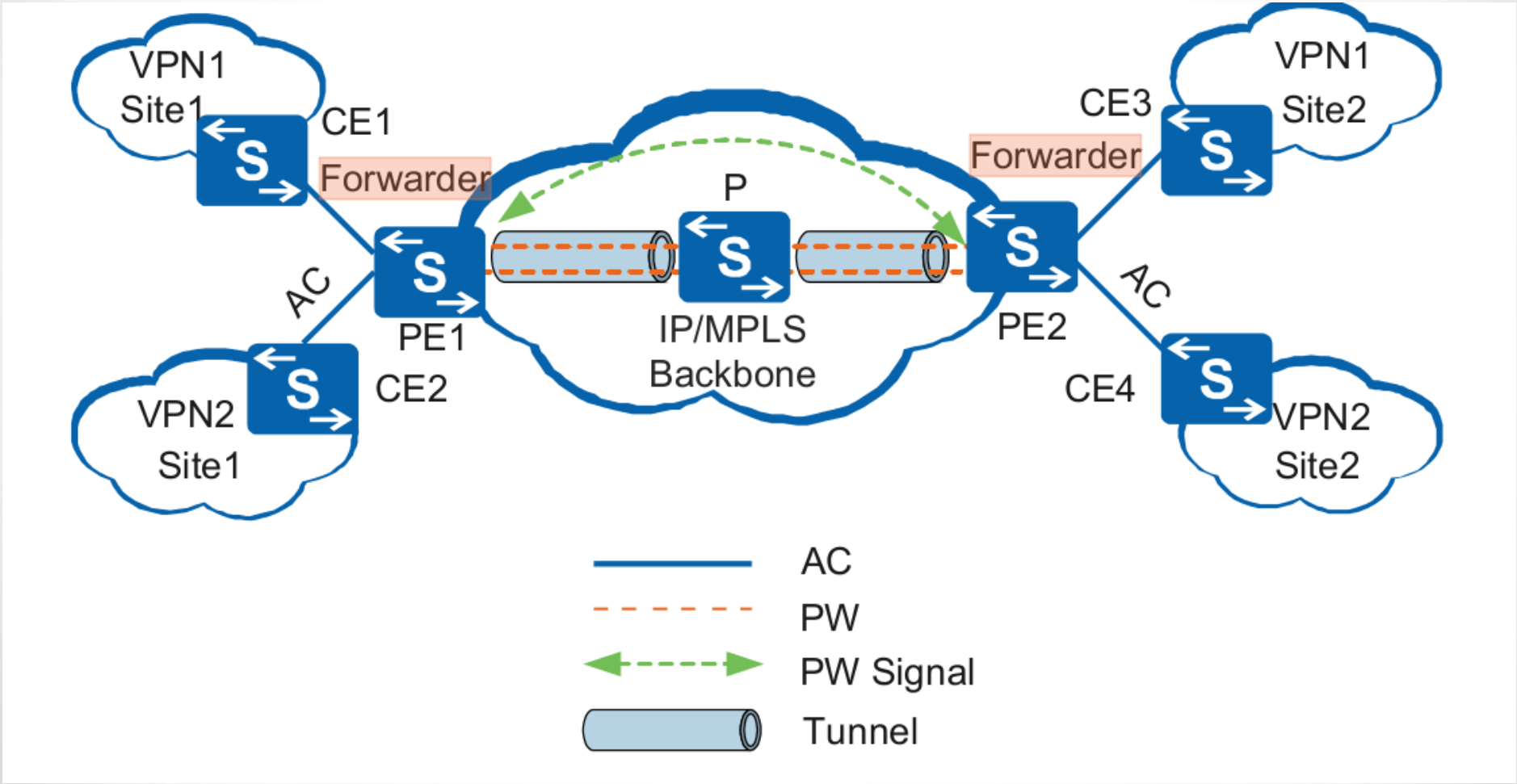
VPWS Data Plane

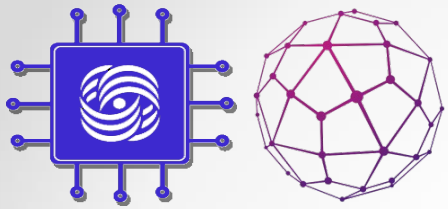
a) LSP уже построены. Для построения LSP могут использоваться протоколы LDP или RSVP-TE
b) Когда входной маршрутизатор R1 получает кадр от устройства клиента (CE1), он уже знает транспортную метку и выходной интерфейс на пути к R6
c) При этом интерфейс (AC), к которому подключен CE1 должен быть привязан к идентификатору клиента — VC ID (аналог VRF в L3VPN). На основании этой информации R1 даёт кадру сервисную метку, которая сохранится до конца пути неизменной.
d) R1 знает точку назначения — IP-адрес удалённого PE-маршрутизатора — R6, выясняет транспортную метку и вставляет её в стек меток MPLS. Это будет внешняя — транспортная метка.
e) Пакет MPLS передается по сети оператора через P-маршрутизаторы. Транспортная метка меняется на новую на каждом узле, сервисная остаётся без изменений.
f) На предпоследнем маршрутизаторе снимается транспортная метка — происходит PHP. На R6 пакет приходит с одной сервисной VPN-меткой.
g) Выходной маршрутизатор PE (R6), анализирует сервисную метку и определяет, в какой интерфейс нужно передать распакованный кадр.

tLDP

LDP	tTLDP
Соседями могут быть только непосредственно подключенные маршрутизаторы	Соседом может быть любой маршрутизатор в сети, с которым есть IP-связность.
Поиск всех возможных соседей	Соседи уже определены конфигурацией
Широковещательная рассылка сообщений Discovery	Адресная отправка сообщения Discovery конкретным соседям.
В качестве FEC обычно выступает IP-адрес	В качестве FEC обычно выступает VC ID

How Layer 2 VPN Works (Cont.)

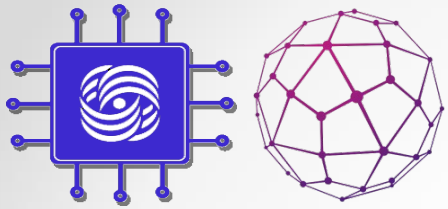




How Layer 2 VPN Works (Cont.)

- CE1 sends Layer 2 packets to PE1 through an AC.
- After PE1 receives the packets, the forwarder selects a PW to forward the packets.
- PE1 generates double MPLS labels (private and public network labels) based on the forwarding entry of the PW. The private network label is used to identify the PW, and the public network label identifies the tunnel to PE2 on the public network.
- After Layer 2 packets arrive at PE2 through the tunnel on the public network, the Penultimate Hop Popping (PHP) device (a P device) pops out the public network label, and PE2 pops out the private network label.
- The forwarder of PE2 selects an AC to forward the Layer 2 packets to CE3.

Forwarder (FWRD) - A PE subsystem that selects the PW to use in order to transmit a payload received on an AC.
Similar to a **forwarding table**. After a PE receives packets from an AC, the forwarder of the PE selects a PW to forward these packets.

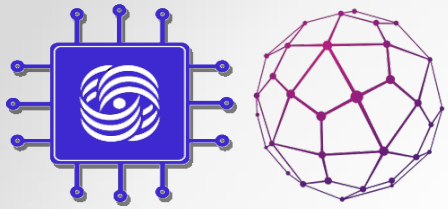


How Layer 2 VPN Works (Cont.)

As discussed in [RFC3985], a pseudowire can be thought of as connecting two "**forwarders**". The protocol used to set up a pseudowire must allow the forwarder at one end of a pseudowire to identify the forwarder at the other end.

We use the term "attachment identifier", or "AI", to refer to the field that the protocol uses to identify the forwarders. In the PWid FEC, the PWid field serves as the AI.

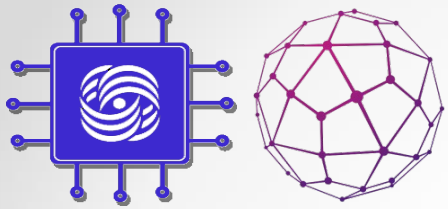
Every Forwarder in a PE must be associated with an Attachment Identifier (AI), either through configuration or through some algorithm. The Attachment Identifier must be unique in the context of the PE router in which the Forwarder resides. The combination <**PE router IP address, AI**> must be globally unique.



How Layer 2 VPN Works (Cont.)

when PE receives a Label Mapping message (LDP), PE interprets the message as a request to set up a PW whose endpoint (at PE) is the Forwarder identified by the TAI (Target AI). **From the perspective of the signaling protocol, exactly how PE maps AIs to Forwarders is a local matter.** In some Virtual Private Wire Services (VPWS) provisioning models, the TAI might, for example, be a string that identifies a particular Attachment Circuit, such as "ATM3VPI4VCI5", or it might, for example, be a string, such as "Fred", that is associated by configuration with a particular Attachment Circuit. In VPLS, the AGI could be a VPN-id, identifying a particular VPLS instance.

If PE cannot map the TAI to one of its Forwarders, then PE2 sends a Label Release message to the sender of Label Mapping message PE, with a Status Code of "Unassigned/Unrecognized TAI", and the processing of the Label Mapping message is complete.



Virtual Private LAN Service

VPLS-Virtual Private LAN Service. Можно рассматривать как эмуляцию коммутатора.
Задача транспортной сети провайдера — обеспечить корректную коммутацию кадров, используется изучении MAC-адресов.

VPLS domain — изолированная виртуальная L2-сеть. Два разных клиента — два разных VPLS-домена.

VSI-Virtual Switching Instance. Виртуальный коммутатор в пределах одного узла. Для каждого клиента (или сервиса) он свой. Трафик одного VSI не может передаваться в другой VSI.
Аналог VRF / VPN-instance в L3VPN.

VE-VPLS Edge-PE node, участник VPLS-домена.

VPLS Data Plane

1. PE-маршрутизатор считывает заголовок Ethernet-фрейма и проверяет MAC-адрес отправителя.
   - Если он уже есть в таблице MAC-ов данного VSI, PE переходит к шагу 2.
   - Если этого адреса ещё нет — он записывает соответствие MAC-порт в таблицу и переходит к шагу 2.
2. PE-маршрутизатор проверяет MAC-адрес получателя.

   а) если он присутствует в таблице MAC-адресов данного VSI:
   - PE ищет выходной интерфейс для кадра с данным MAC'ом. Это может быть физический интерфейс или PW.
   - Если порт назначения — физический интерфейс — просто отправляет фрейм в этот порт.
   - Если это PW, то добавляет соответствующую метку — сервисную. Эта метка будет неизменна до конца пути.
   - PW — это канал между двумя IP-узлами, поэтому зная IP-адрес удалённого PE, локальный PE из таблицы меток извлекает транспортную и ставит её сверху стека — она будет меняться на каждом P-маршрутизаторе.

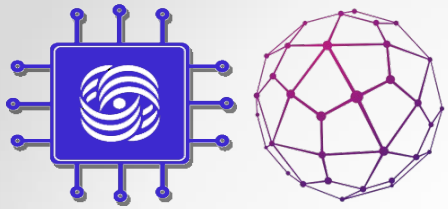
VPLS Data Plane

а) Если же MAC-адрес неизвестен, то PE должен выполнить широковещательную рассылку кадра по всем PE данного VSI.

- Локальный PE составляет список всех удалённых PE этого VSI, и, создав копии этого кадра, вставляет в них сервисные метки — каждому присваивается своя.
- Далее на каждую копию кадра добавляется транспортная метка (тоже своя для каждого PE).
- Все кадры рассылаются по сети провайдера.
- Также копии широковещательного кадра отправляются в AC-интерфейсы, если такие есть, без заголовков MPLS.

VPLS Data Plane

3. Удалённый PE после получения кадра и снятия меток (то есть когда он уже определил VSI) действует как коммутатор:
   a) Если MAC-адрес **источника** ему не известен, вносит его в таблицу. В качестве входного интерфейса будет указан PW к Ingress PE
   b) Если MAC-адрес **назначения** ему известен, отсылает кадр без заголовков MPLS в тот порт, за которым он изучен.
   c) Если этот MAC-адрес **назначения** ему не известен, выполняется широковещательная рассылка по всем AC-портам этого VSI. **PE не будет рассылать этот кадр в PW данного VSI**, т.к. все другие PE уже получили копию этого кадра от входного PE. Т.е. применяется правило расщепления горизонта (Split Horizon).

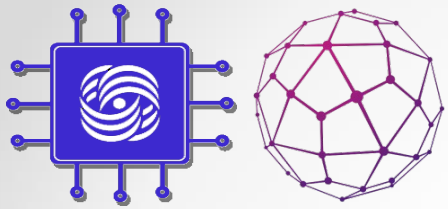


VPLS Data Plane

Как и в обычном коммутаторе записи в MAC-таблице VSI контролируется время существования записи и «старые» записи удаляются.

Клиентский кадр нужно отправить в правильный PW, т.е. правильному соседу.
Для этой цели каждому соседу выдаётся личная метка, с которой тот будет отправлять кадр этому PE в данном VPLS-домене.
В дальнейшем по VPN-метке, и записи в LFIB, PE узнает, от какого соседа пришёл кадр.

В L3VPN не имеет значения, откуда пришёл IP-пакет, поэтому для префикса в VRF всем соседям сообщается одна и та же метка.



VPLS Control Plane

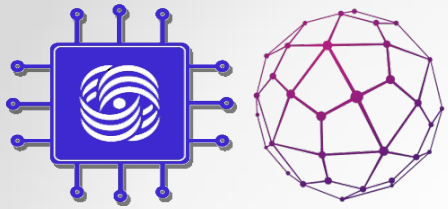
Для VPLS требуется полносвязная топология для каждого VSI.
Соседями, будут только те PE, где есть этот же VSI.

Обнаружение PE, куда подключены клиенты данного VSI:

- **ручная настройка** (драфт Мартини)
- **автоматическое обнаружение** (драфт Компелла).

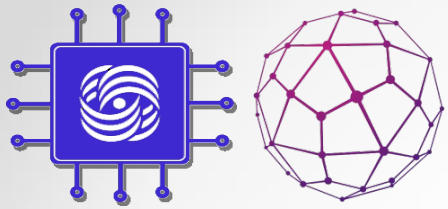
VPLS представляет собой группу point-to-point PW.

Решение о передаче кадра принимает Ingress PE (выбирает нужный PW).



VPLS Martini Mode (LDP)

- VPLS Martini Mode - RFC 4762
- Для сигнализации меток используется LDP
- В отличие от VPWS, удалённые LDP сессии создаются для каждого VSI не с одним соседом, а с несколькими
- Удалённые сессии с каждым соседом в VPLS-домене настраиваются вручную.
  Каждому участнику VPLS, выделяется индивидуальная метка (передается в сообщении LDP Label Mapping Message.
- Если в VPLS-домен добавляется новый PE, необходимо настроить LDP-соседство со всеми существующими PE этого VPLS. После чего с каждым из них новый PE обменятся метками.

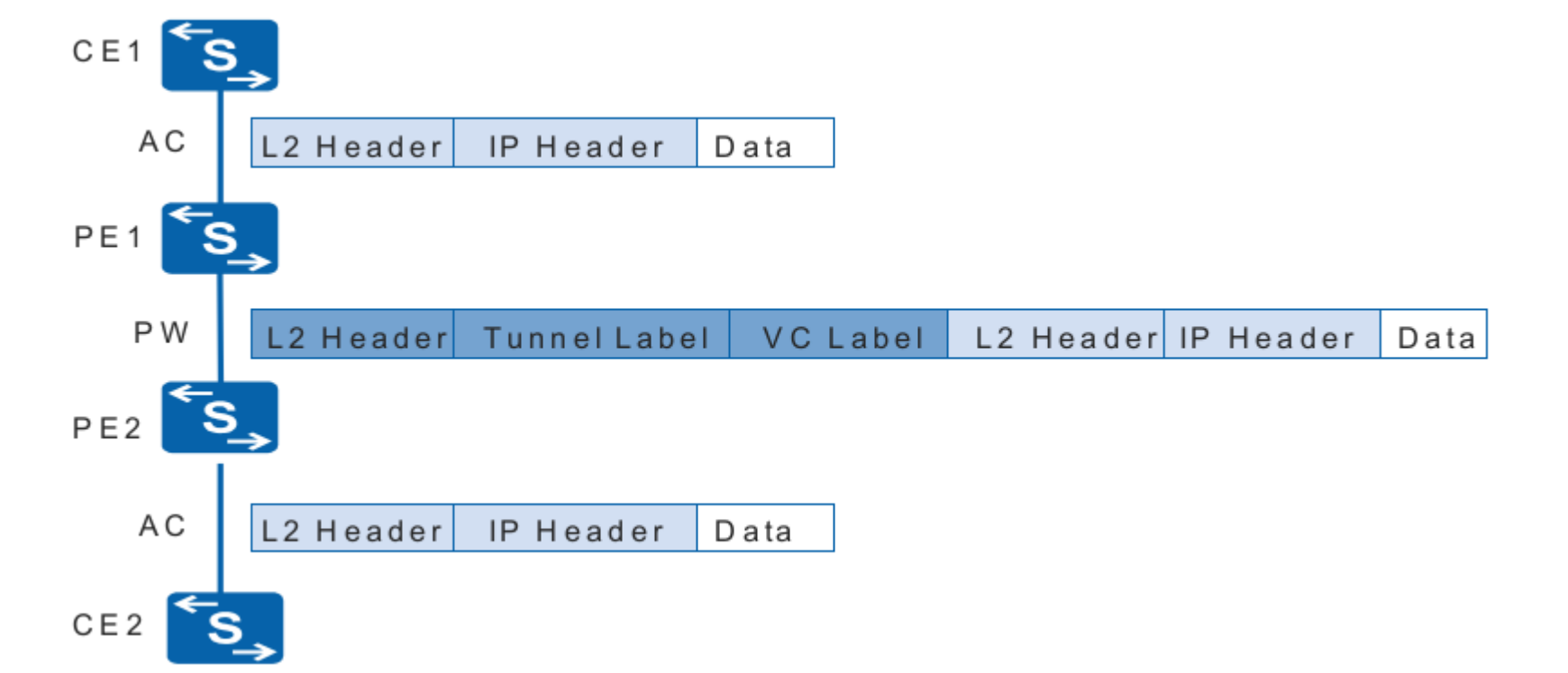


VPLS Martini Mode (LDP)

- LDP проверяет доступность своих соседей. Если какой-то из соседей выходит из группы или становится недоступным, сессия разрывается и все изученные MAC'и в PW к этому соседу очищаются.
- Если какой-либо из AC-портов VPLS-домена переходит в состояние Down, либо происходит другое событие, заставляющее очистить MAC-адреса с AC-порта, то PE сообщает об этом всем своим соседям в сообщении LDP MAC Withdraw (зависит от реализации)

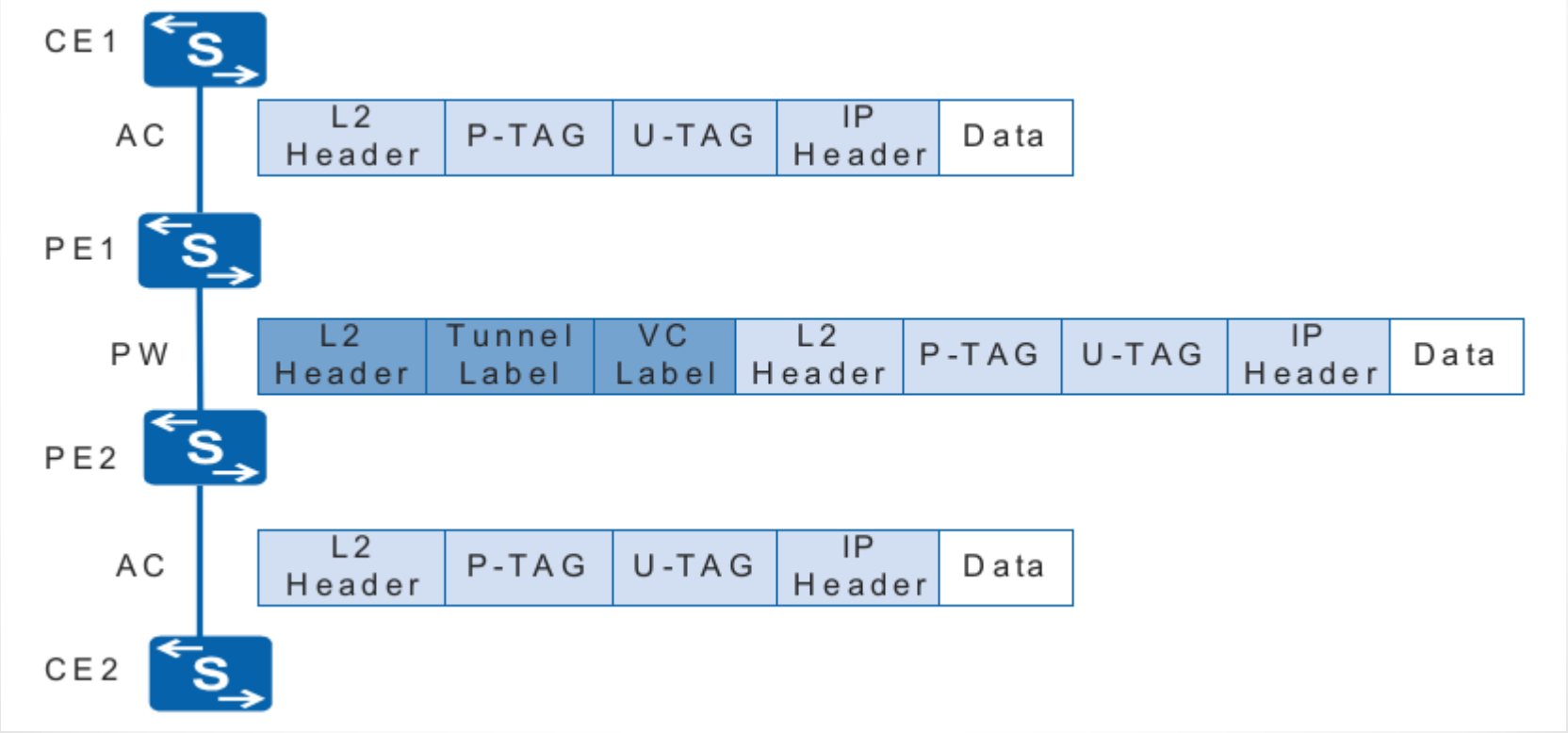
Packet Encapsulation

Ethernet + raw encapsulation (without U-Tag)



Packet Encapsulation

VLAN + tagged encapsulation (with U-Tag).



PW Signaling Protocols

MAC address learning and flooding on a PE. PC1 and PC2 both belong to VLAN10. PC1 pings IP address 10.1.1.2. PC1 does not know the MAC address corresponding to this IP address and advertises an Address Resolution Protocol (ARP) Request packet.