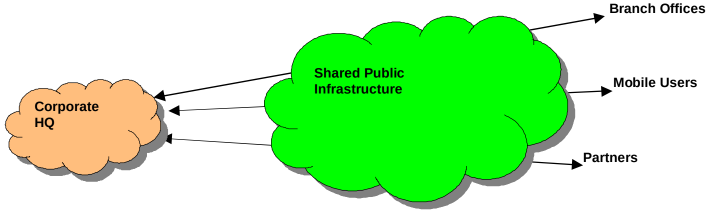
# S-38.3192 Verkkopalvelujen tuotanto

# S-38.3192 Network Service Provisioning

Lecture 7: VPN

## Virtual Private Network

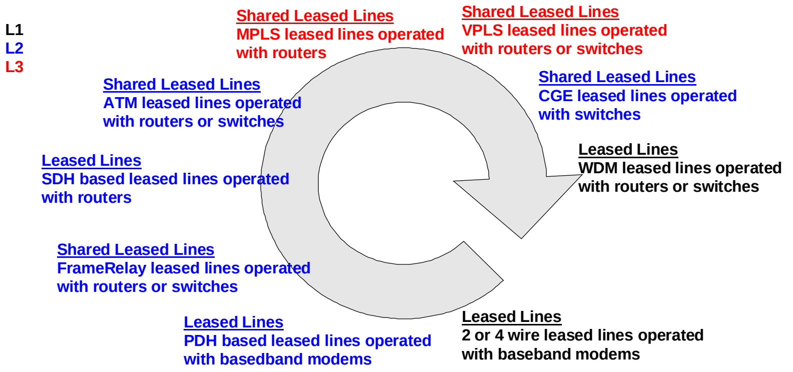
- VPN is
  - A private network constructed over a shared public infrastructure
    - ATM, FrameRelay, MPLS, IP
  - One of several network realizations on the same infrastructure
    - Each have their own routing policy



## VPN

- **Virtual**
  - Network resources used are part of a common shared resource
- **Private**
  - Separate addressing and routing – topological isolation
    - Flow of routing data is constrained to constrain the flow of user data
- **Network**
  - Devices that communicate through some arbitrary method
- **GOAL: Restricted connectivity**
  - Internet: Any to Any
  - VPN: Point to Point or Set to Set

## Development Cycle



L1
L2
L3

Shared Leased Lines
MPLS leased lines operated with routers

Shared Leased Lines
VPLS leased lines operated with routers or switches

Shared Leased Lines
ATM leased lines operated with routers or switches

Shared Leased Lines
CGE leased lines operated with switches

Leased Lines
WDM leased lines operated with routers or switches

Leased Lines
SDH based leased lines operated with routers

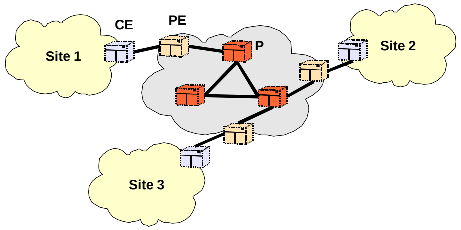
Shared Leased Lines
FrameRelay leased lines operated with routers or switches

Leased Lines
2 or 4 wire leased lines operated with baseband modems

Leased Lines
PDH based leased lines operated with basedband modems

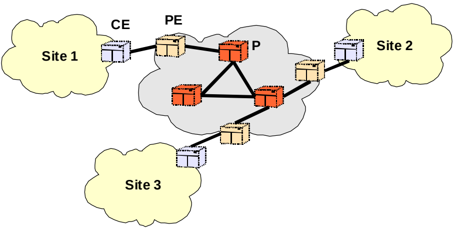
# Terminology

- **Router Types**
  - **CE: Customer Edge Router**
    - Provides connection to the provider network
  - **PE: Provider Edge Router**
    - Provider customer interface
    - Terminates routing from both sides
  - **P: Provider Router**
    - Provider core routers which should not be aware of customers



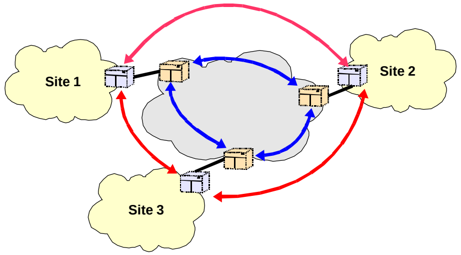
# Terminology

- **Site**
  - Is a collection of networking devices that communicate together without traveling through provider network
  - Is mapped to PE router interface
    - Separate routing table is associated for each site in PE router
- **VPN Routing and Fording Table**
  - VRF stores site specific routes learned from
    - CE with any means
    - PE with MP-IBGP



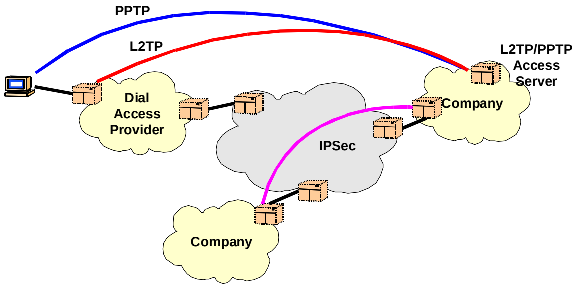
# VPN Types

- **Customer based**
  - Routing and control at the CE routers
  - L2TP, PPTP, IPSec, GRE
- **Provider based**
  - Routing and control at the PE routers
  - MPLS, VPLS



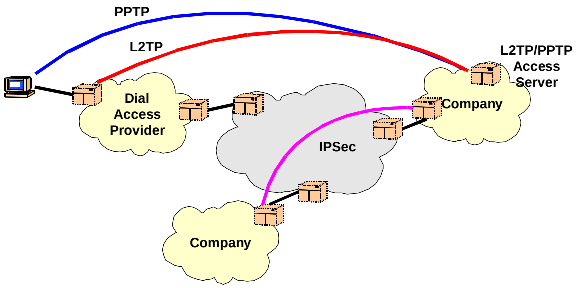
# Customer Based VPNs

- PPTP/L2TP are typical ways to build L2 VPNs from dial-up connections to company resources
  - Operate on top of TCP (PPTP) or UDP (L2TP)



# Customer Based VPNs

- IPSec is used to create L3 VPNs between location whether end host or CPE device
  - Native support for strong encryption (company confidentiality)



# Provider Based VPNs

- **L3 approach**
  - RFC 2547bis
  - Provider delivers L3 access between PE routers of customer sites
  - Customer locations are routed together using BGP as means to deliver labels and addressing information through the core
- **L2 approach**
  - Draft-martini, Draft-kompella
  - VPLS
  - Provider delivers L2 (Ethernet) access between PE routers of customer sites
  - BGP of LDP is used to distribute labels between PE routers

# RFC 2547bis

- Routed interconnection of VPN sites
- Multiprotocol BGP extensions are used to transfer routes through the core network
- Customers are separated to individual routing and forwarding tables
- Scalability is achieved by minimizing configuration
  - CE only knows interfacing PE
  - PE needs to know interfacing CE
  - P knows nothing about VPNs

# VPN-IPv4 NLRI

- **MP-BGP**
  - Multiprotocol extensions for BGP-4
  - RFC 2283
- **NLRI: AFI:1 SAFI:128**
  - Mask
  - MPLS label
  - Route distinguisher
    - Disambiguates IPv4 addresses -> Controlled duplicates of addresses
  - Subscriber IPv4 prefix

| | | | | | |
|---|---|---|---|---|---|
| Mask | Label | Type | Adm | AN | IP Address |

# VPN-IPv4 route distinguisher

- **Type**
  - 0:
    - Adm=AS number
    - AN=4 bytes (PE RID)
  - 1:
    - Adm=4 bytes (PE RID)
    - AN=Unique Number
- **Administrator**
  - Identifies the assigned number authority
    - AS -> PE RID
    - PE RID -> Unique Number
- **Assigned Number**

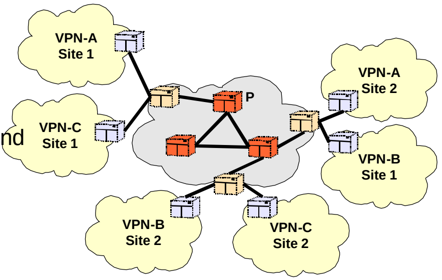
| | | | |
|---|---|---|---|
| Type | Adm | AN | IP Address |

# VPN-IPv4 Addresses

- These extended addresses appear only in control plane of PE routers
  - Route distinguisher points into a VRF where particular address should be stored for packet delivery
    - Same address can safely co-exist in two different VRFs due to full isolation between them
      - (Logical) interfaces are bound into VRFs

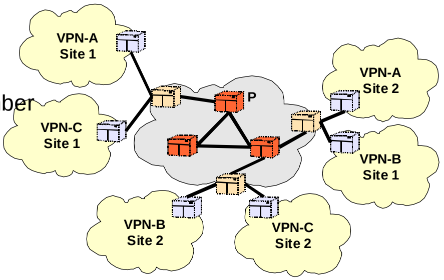
# Distribution of routes

- Distribution of customer routes through provider network is based on BGP
  - IBGP between PE routers of different customer sites
    - Full mesh of PE routers
    - All VPN routes are sent to every other PE
      - Scalability concerns
        - » IBGP peering
        - » Storing of routes and labels



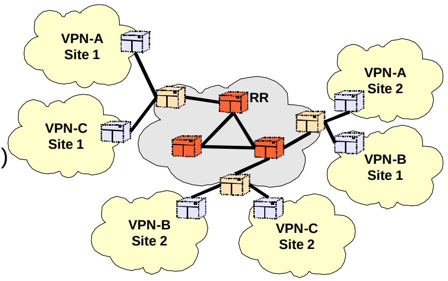
# Distribution of routes

- Route target is a BGP extended community attribute which can be used to filter routes coming from IBGP sessions
  - Identifies a set of VRFs to which a PE router wishes to distribute routes
  - Same format options as in route distinguisher
    - ASN:IPv4 Address
    - IPv4 Address:Unique Number



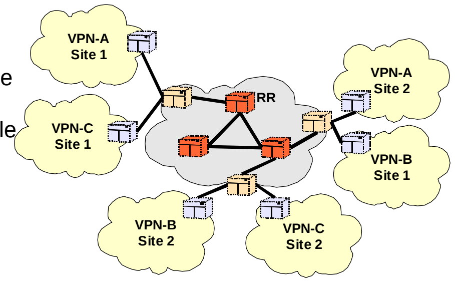
# Distribution of routes

- Route reflectors can be used to alleviate peering constraints in IBGP sessions
  - PE routers send their VPN routes to RR which stores them into VPN table
  - PE's receive all routes of other PE's
  - PE's can ask routes with certain target
    - Route target filtering (RTF)



# Distribution of routes

- Route reflector need not to be PE router as it does not have VRF tables
  - Routes from individual VRFs are stored in a single BGP routing table
  - BGP refresh capability is used to retrieve routes on non disruptive manner from the RR
  - An LSP is required from RR to every PE
    - BGP next-hop needs to be resolvable from the RR in order to make route usable

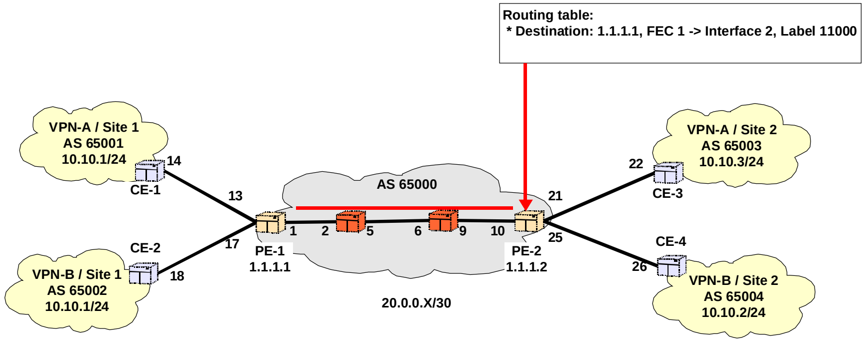


# Route Target Filtering

- Route target filtering uses an separate NLRI format
  - AFI:1 SAFI:132
    - Prefix limit
      - Maximum number of RT advertisements that can be reiceived

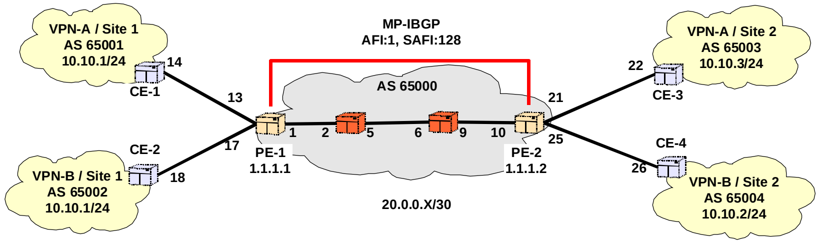
# MPLS LSP

- LSP between PE-1 and PE-2 is set up for tunneling VPN packets through the provider core


Routing table:
* Destination: 1.1.1.1, FEC 1 -> Interface 2, Label 11000

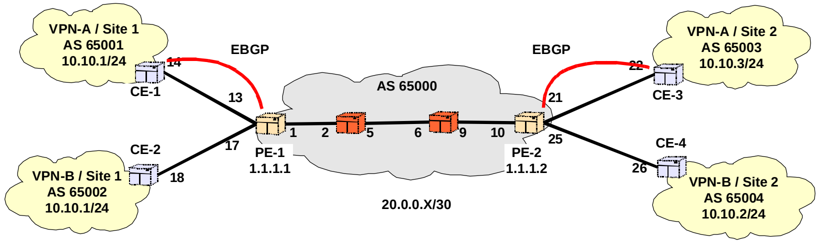
# PE-PE

- MP-IBGP session between PE routers is established
  - LSP between PE routers is required to resolve BGP next-hop



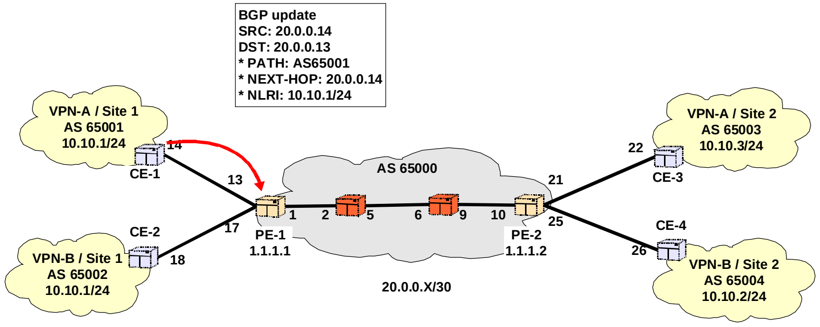
# CE-PE Communication

- BGP is native choice between two different administrative domains
- IGPs (RIP, OSPF, IS-IS) could also be used
  - Separate routing process needs to be run for each customer
    - Separation of customer and provider routing



# Exchange of routing information

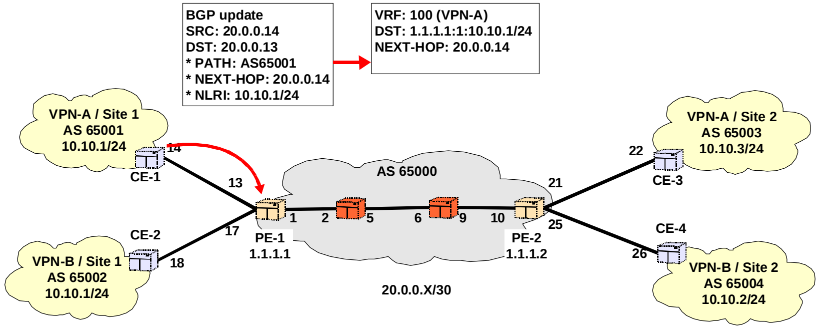
- CE-1 sends a BGP update to PE-1



BGP update
SRC: 20.0.0.14
DST: 20.0.0.13
* PATH: AS65001
* NEXT-HOP: 20.0.0.14
* NLRI: 10.10.1/24

# Exchange of routing information

- PE-1 checks that it has BGP-next hop in IGP and install routes in correct VRF



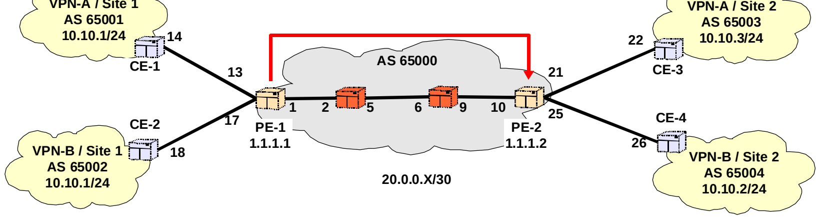
BGP update
SRC: 20.0.0.14
DST: 20.0.0.13
* PATH: AS65001
* NEXT-HOP: 20.0.0.14
* NLRI: 10.10.1/24

VRF: 100 (VPN-A)
DST: 1.1.1.1:1:10.10.1/24
NEXT-HOP: 20.0.0.14

# Exchange of routing information

- PE-1 sends a MP-IBGP update to peers (PE-2)
  - Next hop is pointing to PE-1 as it does routing table change
  - MPLS label is pointing to VRF 100

MP-IBGP update
SRC: 1.1.1.1
DST: 1.1.1.2
* PATH: 65001
* NEXT-HOP: 1.1.1.1
* LABEL: 10003
* NLRI: 1.1.1.1:1:10.10.1/24
* ROUTE TARGET: 65000:100

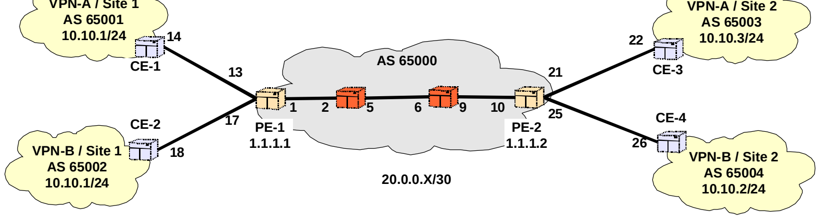


# Exchange of routing information

- PE-2 checks for proper import filter (route target) and installs routes to correct VRF

MP-IBGP update
SRC: 1.1.1.1
DST: 1.1.1.2
* PATH: 65001
* NEXT-HOP: 1.1.1.1
* LABEL: 10003
* NLRI: 1.1.1.1:1:10.10.1/24
* ROUTE TARGET: 65000:100

VRF: 100 (VPN-A)
DST: 1.1.1.1:1:10.10.1/24
NEXT-HOP: 1.1.1.1
LABEL:
* INNER: 10003
* OUTER: 11000

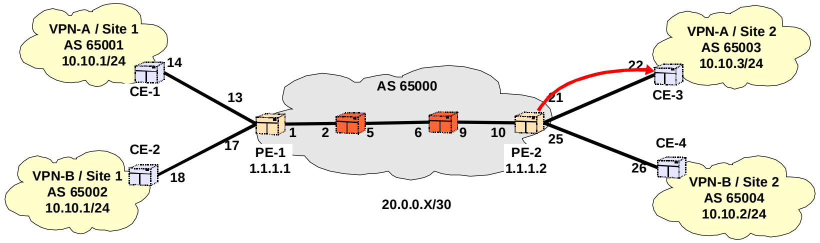


# Exchange of routing information

- PE-2 generates an BGP update for CE-3
  - Note that PATH is changed from 65001 -> 65000, 65000 due to loop detection of BGP

VRF: 100 (VPN-A)
DST: 1.1.1.1:1:10.10.1/24
NEXT-HOP: 1.1.1.1
LABEL:
* INNER: 10003
* OUTER: 11000

BGP update
SRC: 20.0.0.21
DST: 20.0.0.22
* PATH: AS65000,AS65000
* NEXT-HOP: 20.0.0.21
* NLRI: 10.10.1/24

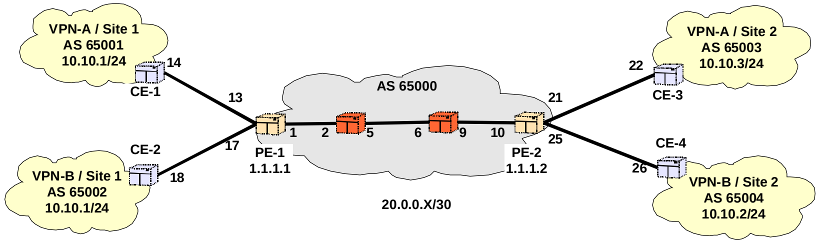


# Exchange of routing information

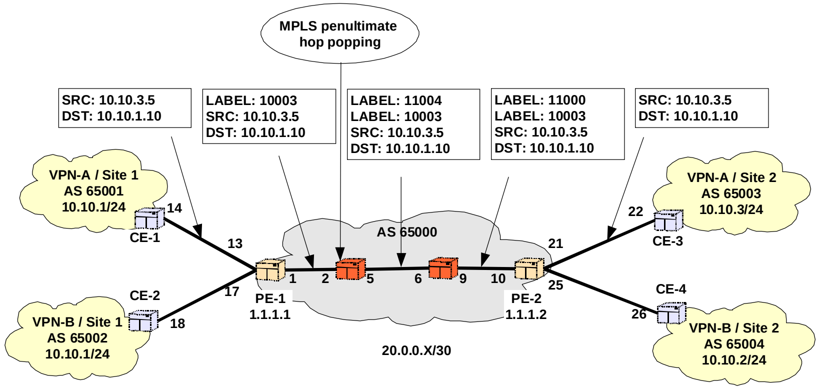
- CE-3 checks from IGP validity of BGP-next hop and installs routes to RIB

BGP update
SRC: 20.0.0.21
DST: 20.0.0.22
* PATH: AS65000,AS65000
* NEXT-HOP: 20.0.0.21
* NLRI: 10.10.1/24

RIB
DST: 10.10.1/24
NEXT-HOP: 20.0.0.21



# Dataflow



# L2 MPLS VPN

- PE router maps circuit IDs (VLAN ID, FrameRelay DLCI, ATM VPI/VCI) to label
- Decouple of customer facing technology from core technology
- Simplify provisioning of customer services
- Each site has own circuit from CE to PE
- Interconnection happens at Ces (routing)

- Draft-Martini
  - Communication between PE routers is based on LDP

- Draft-Kompella
  - Communication between PE routers is based on BGP

# L2VPN NLRI

- Length of the NLRI
- Route Distinguisher
- Site ID (Identifies the CE)
  - Unique ID withing VPN
- Label Base
  - First label in label range
- Label Block Offset
  - If multiple label blocks are used defines the offset from the base label

- Circuit Status
  - Signals the L2 status of PE-CE link to the other end of the link
    - Simultaneous carrier loss at both ends
      - L2 detection for OSPF
  - Also carries Label range value

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Length | Type | Adm | AN | Site ID | Offset | Label Base | Circuit Status |

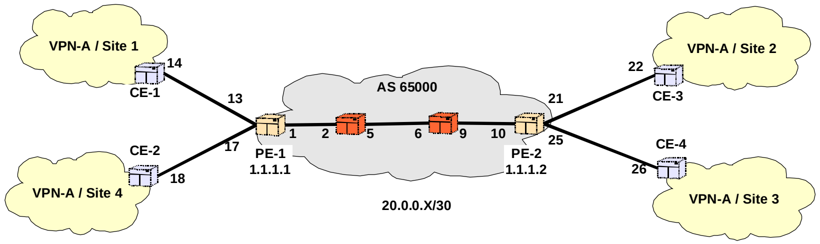
# L2VPN VFT

- Route Target
  - Community for forming L2 VPN
- Site ID
  - Unique ID withing VPN
- Label Range
  - Number of possible peer CE
- Label Base
  - First label in label range

- Sub-int ID:Label pairs
  - Sub-interfaces in PE/CE to handle connections to different sites
  - Labels are assigned by PE based on
    - Label base
    - Label range
    - Remote-site-ID
      - Auto assignement

# L2 MPLS VPN

- PE maps incoming packets based connection ID's to LSP having label stack dependent on remote-site-ID and MPLS connection label
  - As in L3VPN case

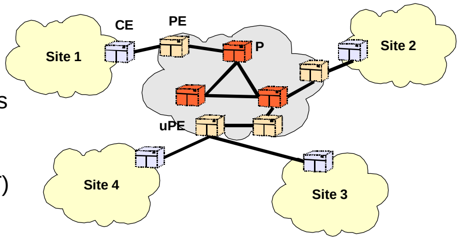


# Virtual Private LAN Service

- The idea behind VPLS is to map provider infrastructure to a virtual bridge
  - Remember the idea from Carrier Grade Ethernet -lecture
    - E-LAN service, where network looks like a distributed bridge
    - VPLS is a method for provider to offer CGE type of E-LAN service
- Two versions:
  - BGP based by Kireeti Kompella (Juniper)
    - Some scalability benefits over the other
  - LDP based by Vach Kompella (Alcatel)

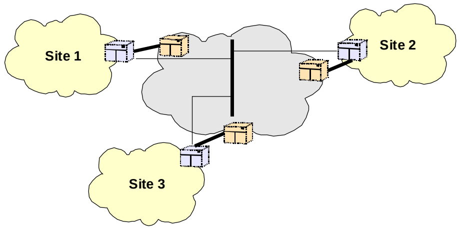
# Terminology

- **Router Types**
  - CE: Customer Edge Device
    - Router or Ethernet bridge
  - PE: Provider Edge Router
    - There can also be uPE which is L2 aggregation device in front of PE
    - Also called VE device
      - Perform MAC address learning
      - Contains VPN forwarding table (VFT)
  - P: Provider Router
    - Provider core routers which should not be aware of customers



# CE

- CE is in major role in VPLS
  - CE's form direct relationship as if there is no provider network in between
    - IP routing adjacency
    - Ethernet spanning tree adjacency
  - Same L2 configuration on all sites
    - VLAN ID



# uPE PE

- Provider edge devices are the ones which are aware of VPLS service
  - uPE does L2 aggregation in front of PE router
    - Economics of law: interfaces at L2 device are much cheaper than on the L3 device

# VFT / VCT

- VFT contains
  - Local VCT
    - Local site ID
    - Site's Layer 2 encapsulation (Ethernet, VLAN, etc)
    - Logical interfaces provisioned to the local CE
    - Label base used to associate received traffic with one of the logical interfaces
  - VCT from other PE
    - Site ID (VE ID)
    - Label

# VFT

- Route Target
  - Community for forming VPLS
- Site ID
  - Unique ID withing VPLS
- Label Range
  - Number of possible peer CE
- Label Base
  - First label in label range
- Offset

- Remote site:Label pairs
  - Other possible sites and labels that are used to communicate with peers
    - Populated with MP-IBGP

# VPLS NLRI

- Similar to L2VPN NLRI
- AFI (1), SAFI 65
- VE ID <-> Site ID
- VE Block Offset <-> Label offset
- VE Block Size <-> Label range
- Label Base <-> Label base

- No circuit staty

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Length | Type | Adm | AN | VE ID | VE<br>Block Offset | VE<br>Block Size | Label Base |

# L2 Extended Community

- Community type
  - L2 Information
- Encapsulation Type
  - 19: VPLS
- MTU
  - All sites must use  same MTU size
    - Single LAN emulation
- Flags
  - MBZ: 6 zeros
  - C: Control word required
  - S: Sequencing required

| | | | | |
|---|---|---|---|---|
| **Com Type** | **Encap Type** | **Flags** | **L2 MTU** | **Reserved** |