The Enterprise Routing Guide to MPLS VPN Service Migration

MPLScon 2006

May 25, 2006

Ajay Simha (asimha@cisco.com)

Cisco Systems, Inc.

170 W. Tasman Dr. San Jose, CA 95134

http://www.cisco.com

Speaker Biography

Biography: Ajay Simha

Ajay Simha CCIE #2970, joined Cisco Systems, Inc. TAC in 1996. He then went on to support tier 1 and 2 ISPs as a part of the Cisco's ISP Expert team. Currently a member of the Metro Ethernet group in Advanced Systems Central Engineering Team. He has extensive experience in designing MPLS networks for large service providers. Ajay has been a speaker at Networkers and the MPLS International conference (Washington D.C). He is also the co-author of the Cisco Press publication Traffic Engineering with MPLS. Ajay has a M.S in computer science from New Jersey Institute of Technology.

Agenda

- Introduction
- Physical Migration to MPLS VPN Backbone
- Routing considerations using
  - BGP as PE-CE protocol
  - OSPF as PE-CE protocol
  - EIGRP as PE-CE protocol
- Default route handling in MPLS VPN
- Preventing routing Loops with SOO
- Limiting vrf routes and potential black holing
- Multi-homing Scenarios
- Summary

Introduction

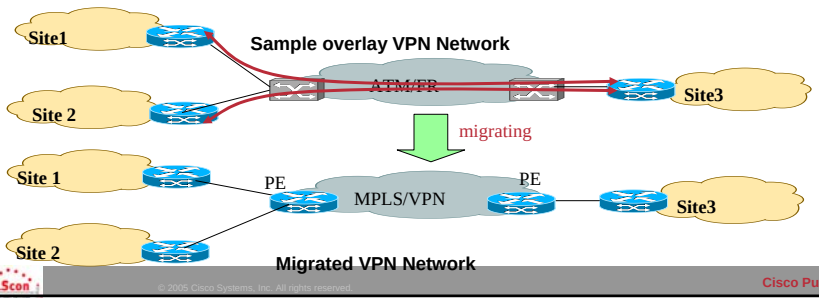
- Many enterprises are migrating to VPN services based on Layer 3 infrastructure (aka RFC 2547 based VPNs)
- In the traditional Layer 2 VPN Frame or ATM-based networks, Service provider network does not participate in the enterprise routing.
- Change in routing policies may result in network either sub-optimally utilized or even could lead to routing loops
- Enterprise network operators need to fully understand various factors that determine the overall complexity during and after migration such as
  - Internal Site routing protocols
  - Choice of PE-CE protocols
  - Multi-homing, Redundancy and load balancing options
  - Existence of backdoor links
  - Network size (large number of sites)
  - Number of Hub sites etc.
- Various network scenarios are discussed to highlight the issues and possible solutions.

Agenda

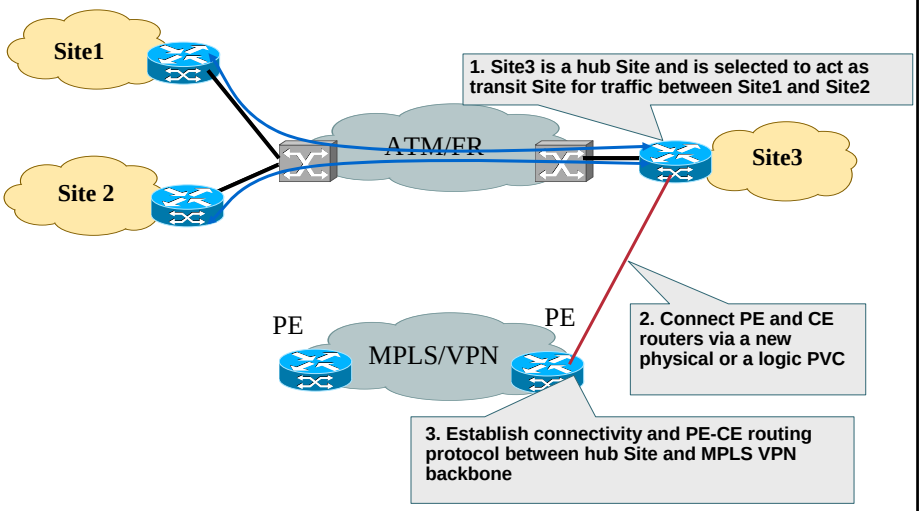
- Introduction
- **Physical Migration to MPLS VPN Backbone**
- Routing considerations using
  - BGP as PE-CE protocol
  - OSPF as PE-CE protocol
  - EIGRP as PE-CE protocol
- Default route handling in MPLS VPN
- Preventing routing Loops with SOO
- Limiting vrf routes and potential black holing
- Multi-homing Scenarios
- Summary

Migration Considerations

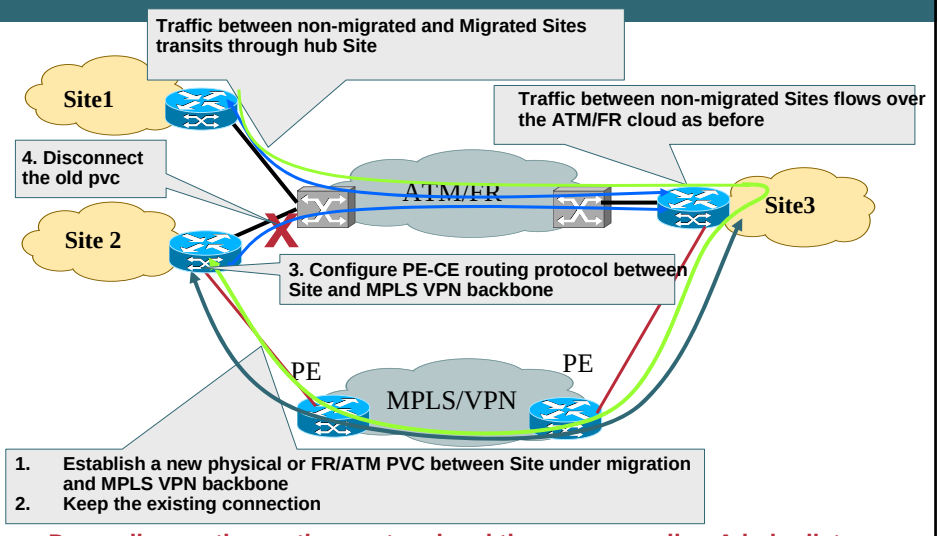
- **Minimize impact on customer connectivity and traffic forwarding as well as avoid potential Site isolation during migration.**
- **Routing interaction of PE-CE routing protocols with the Site local IGP**
  - **Customers may not use their existing internal routing protocol to exchange routing information with the provider.**
- **Need to make sure internal as well internet routing works as desired**
- **Migration of a large enterprise to MPLS VPN needs phased approach**



Migration steps: Hub Site Migration



Migration steps: Individual Sites Migration



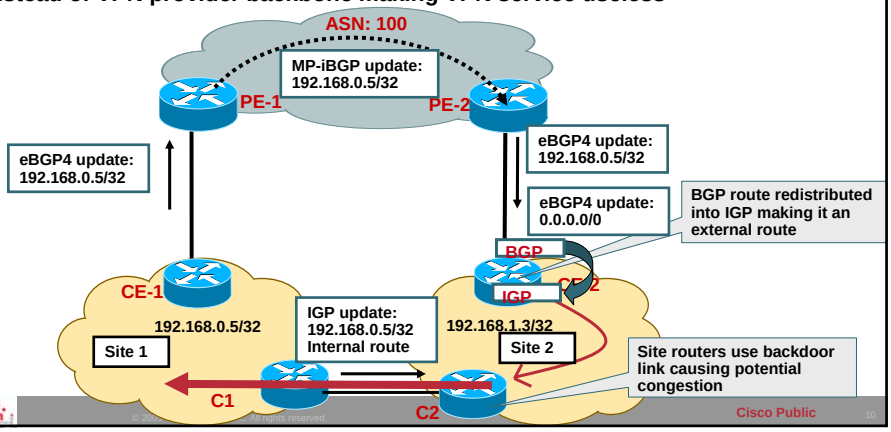
Depending on the routing protocol and the corresponding Admin distance and metrics, traffic will start flowing over MPLS VPN backbone

Agenda

- Introduction
- Physical Migration to MPLS VPN Backbone
- Routing considerations using
  - BGP as PE-CE protocol
    - BGP interaction with local Site IGPs
    - AS Considerations and VPN Topologies
  - OSPF as PE-CE protocol
  - EIGRP as PE-CE protocol
- Default route handling in MPLS VPN
- Preventing routing Loops with SOO
- Limiting vrf routes and potential black holing
- Multi-homing Scenarios
- Summary

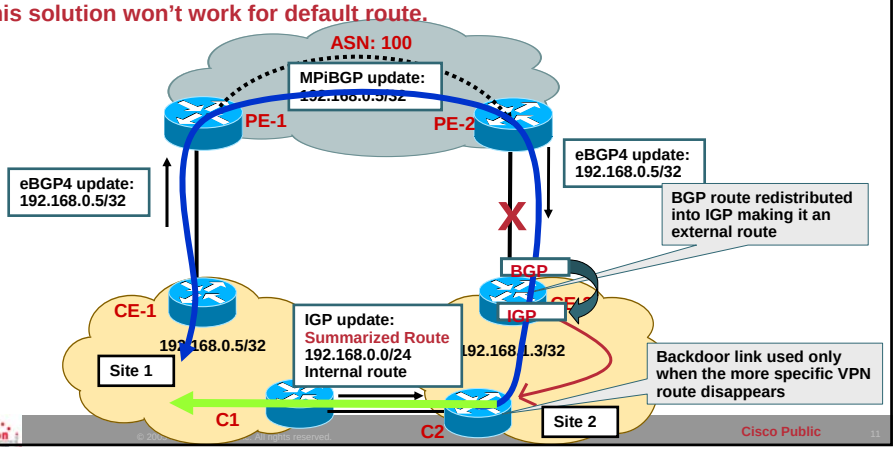
Redistributing BGP into local Site IGP
Problem - Backdoor being preferred

- BGP route redistributed in local Site IGP (such as OSPF, EIGRP) becomes external
- Backdoor link is part of the same IGP
- Site 2 for example also learns the same prefix via backdoor link as internal route
- At Site 2, internal route is preferred over external. Traffic is sent over backdoor link instead of VPN provider backbone making VPN service useless



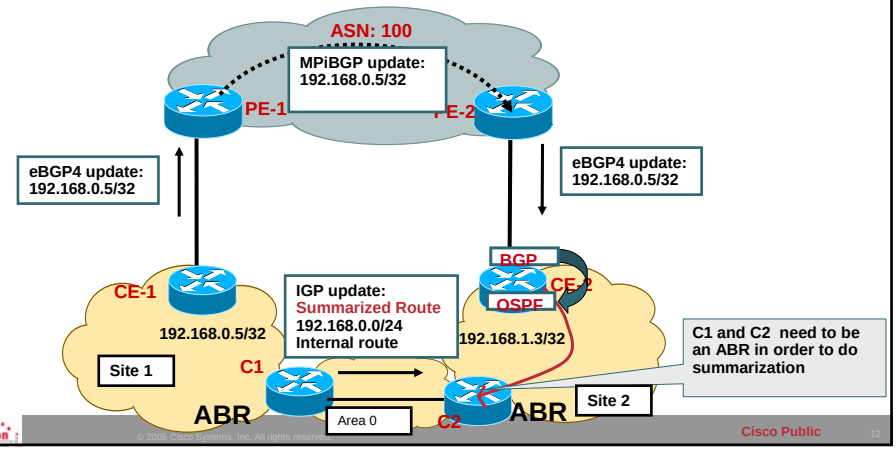
Redistributing BGP into local Site IGP
Solution – Advertise a Summary route

- Simplest solution is to remove the backdoor link ☺
- Other possible solution is to send a summarized route from Site 1 to Site 2 and vice versa over the backdoor link
- In normal conditions, at each Site more specific route learnt from the SP would be preferred over the summary route.
- This solution won't work for default route.



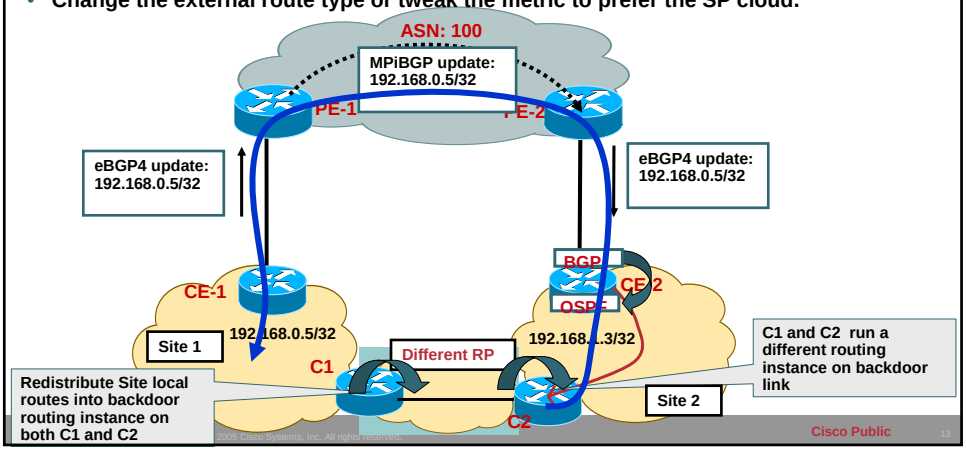
Redistributing BGP into OSPF local Site IGP
Make backdoor part of area 0

- The summary route solution will not work if OSPF is the local IGP
- Summary generated only if C1 and C2 routers are OSPF ABRs or (ASBRs if routes are external)



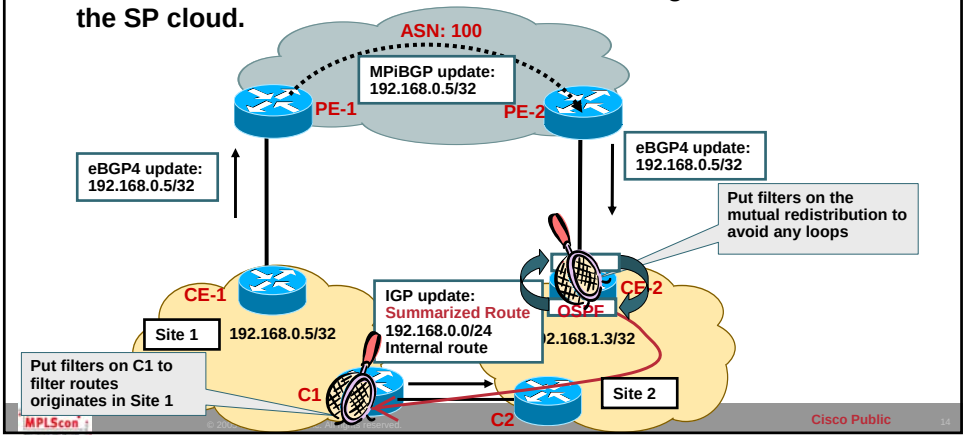
Redistributing BGP into OSPF local Site IGP
Make backdoor part of a different Routing Protocol

- Run a different routing protocol or different IGP instance on the backdoor link
- Redistribute Site local IGP routes into the backdoor routing protocol instance
- Now routes from SP cloud learnt via BGP and the route learnt over back door are both external
- Change the external route type or tweak the metric to prefer the SP cloud.



Redistributing BGP into local Site IGP
Filtering considerations

- Because of mutual redistribution on CE routers at each Site, routing loops are possible
- Need to apply filters to advertise only locally sourced routes from each Site and block Site local routes being received from the SP cloud.

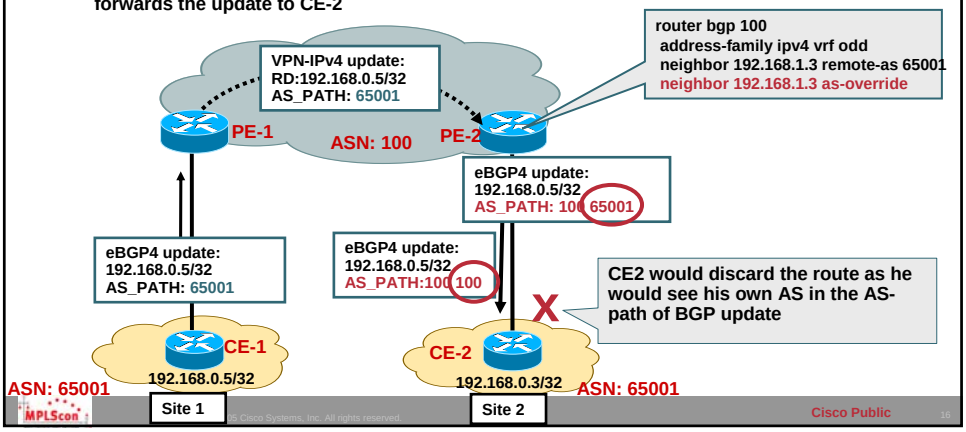


Agenda

- Introduction
- Physical Migration to MPLS VPN Backbone
- Routing considerations using
  - BGP as PE-CE protocol
    - BGP interaction with local Site IGPs
    - AS Considerations and VPN Topologies
  - OSPF as PE-CE protocol
  - EIGRP as PE-CE protocol
- Default route handling in MPLS VPN
- Preventing routing Loops with SOO
- Limiting vrf routes and potential black holing
- Multi-homing Scenarios
- Summary

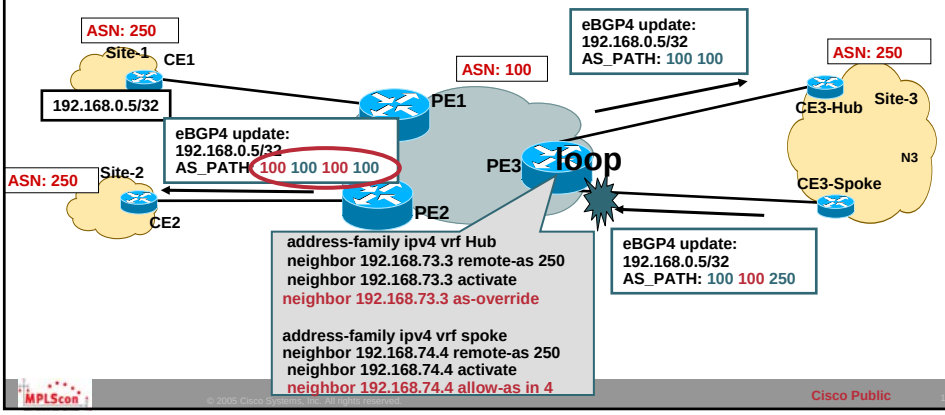
BGP AS Considerations
VPN Sites belong to same ASN

- Customer may have same AS number in all its Sites
- Default BGP behaviour would force the CE to drop the routing update because of the AS-path loop detection
- "Allow-as in" can be used on the CE to accept the update even if it contains its own AS.
- Service provider can re-write the customer AS using "AS- override" feature
- PE-2 replaces all occurrences of customer ASN in the AS-Path with its own ASN and forwards the update to CE-2



VPN Topology considerations
Hub and Spoke Model

- PE3 sees its own AS in the AS-Path and rejects the update
- "Allow-as in" if configured at spoke Site, will allow the update at PE3 if it contains SP's ASN

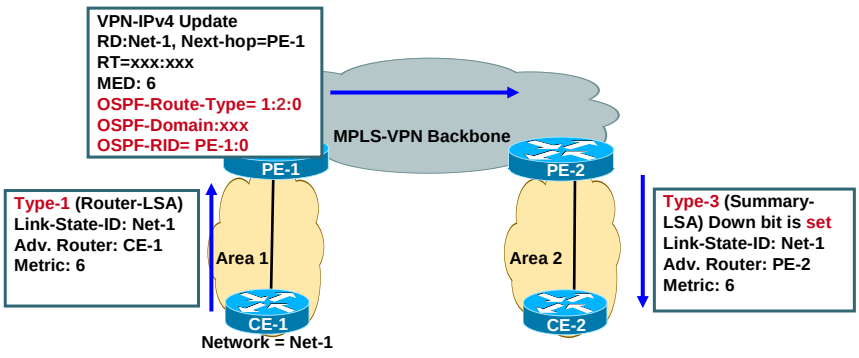


Agenda

- Introduction
- Physical Migration to MPLS VPN Backbone
- Routing considerations using
  - BGP as PE-CE protocol
  - **OSPF as PE-CE protocol**
  - EIGRP as PE-CE protocol
- Default route handling in MPLS VPN
- Preventing routing Loops with SOO
- Limiting vrf routes and potential black holing
- Multi-homing Scenarios
- Summary

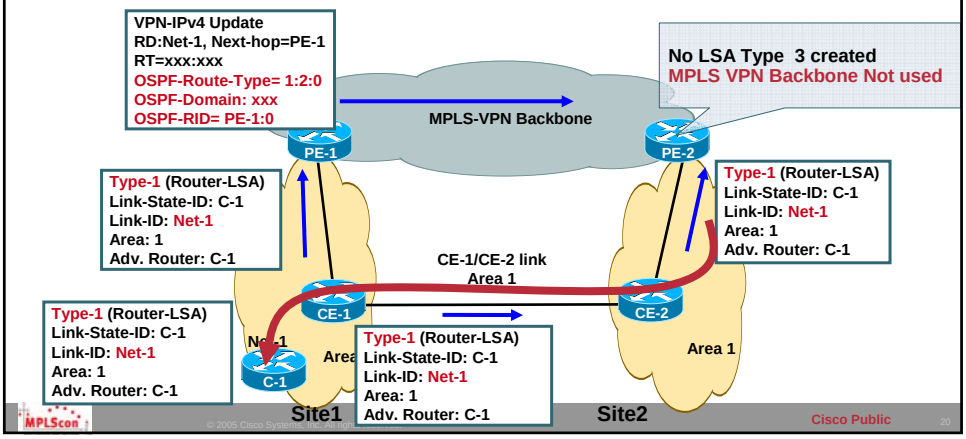
Common Design Consideration- OSPF Area placement
OSPF Sites belong to different areas

- Area 0 is not mandatory when migrating to MPLS VPN service
- VPN sites may have different Sites configured for different areas
- If Area 0 exists, it must touch MPLS VPN PE routers.



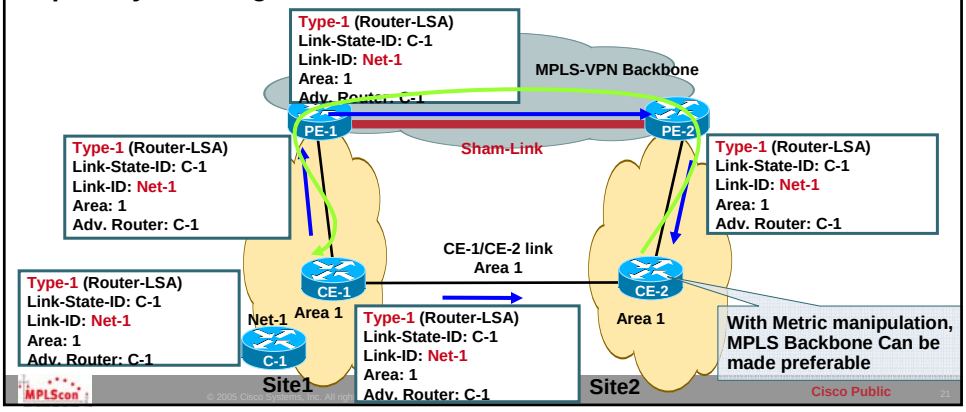
Common Design Consideration- OSPF Area placement
Sites are in the same Area- Backdoor exists

- Customers Sites are in the same area and there is a backdoor link
- Route is advertised to MPLS VPN backbone
- Same prefix is learnt as intra-area route via backdoor link
- PE2 does not generate Type3 LSA once type-1 LSA is received from the site
- Traffic is sent over backdoor link instead of MPLS VPN cloud.



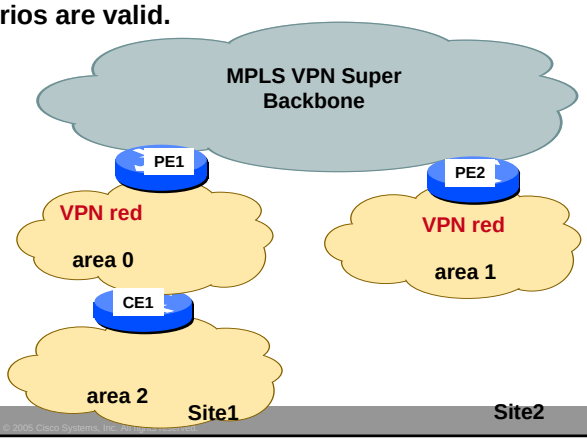
Common Design Consideration- OSPF Area placement
Sites are in the same Area- Backdoor with Sham link

- The sham link is treated as a virtual-link : unnumbered, ptp, DC link
- The sham link is reported in the router LSA's type 1 originated by the two routers connecting to the sham link
- The MPLS VPN backbone or the backdoor link can be made preferred path by tweaking the metrics



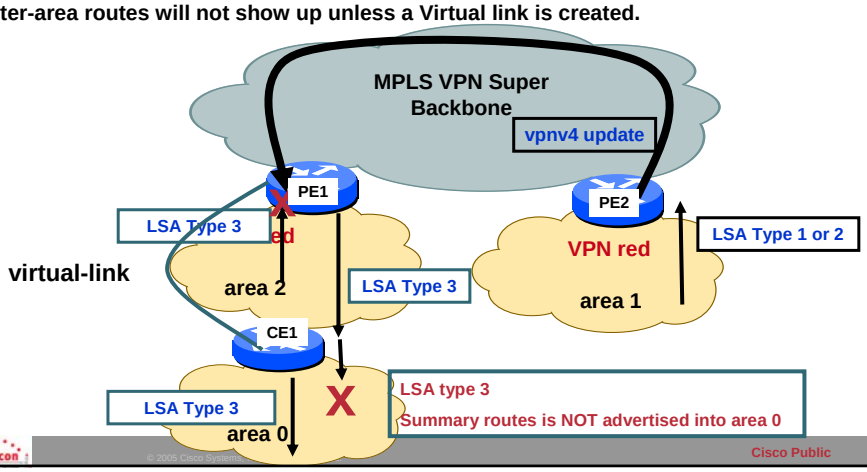
Common Design Consideration- OSPF Area placement
Other scenarios

- Some OSPF sites entirely belong to area 0 and some other sites can belong to non area 0
- Some sites may consist of hierarchical OSPF topology consisting of area 0 as well as non-zero areas.
- Both scenarios are valid.



Common Design Consideration- OSPF Area placement
Area 0 Placement

- As before some sites may consist of hierarchical OSPF topology consisting of area 0 as well as non-zero areas.
- If site contains area 0, it must touch provider PE router.
- OSPF RULE: Summary LSAs from non-zero area's are not injected into backbone area 0
- Inter-area routes will not show up unless a Virtual link is created.

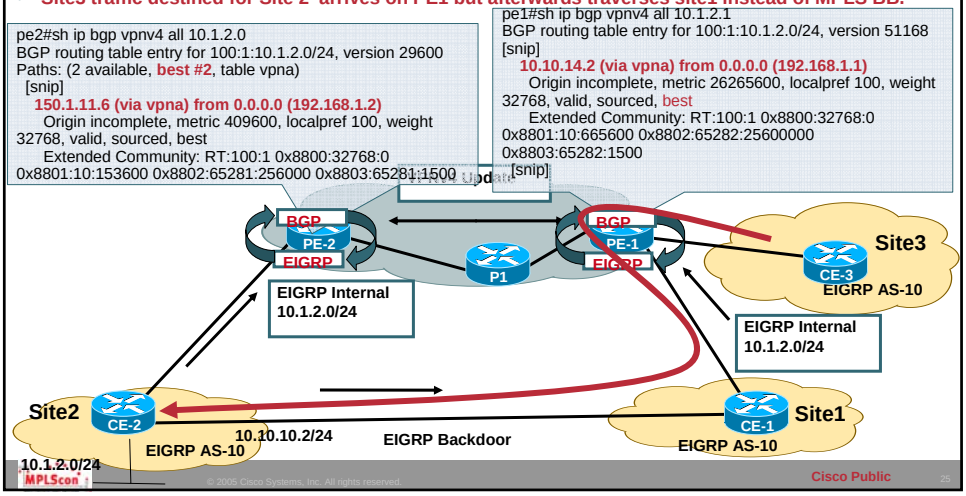


Agenda

- Introduction
- Physical Migration to MPLS VPN Backbone
- Routing considerations using
  - BGP as PE-CE protocol
  - OSPF as PE-CE protocol
  - EIGRP as PE-CE protocol
- Default route handling in MPLS VPN
- Preventing routing Loops with SOO
- Limiting vrf routes and potential black holing
- Multi-homing Scenarios
- Summary

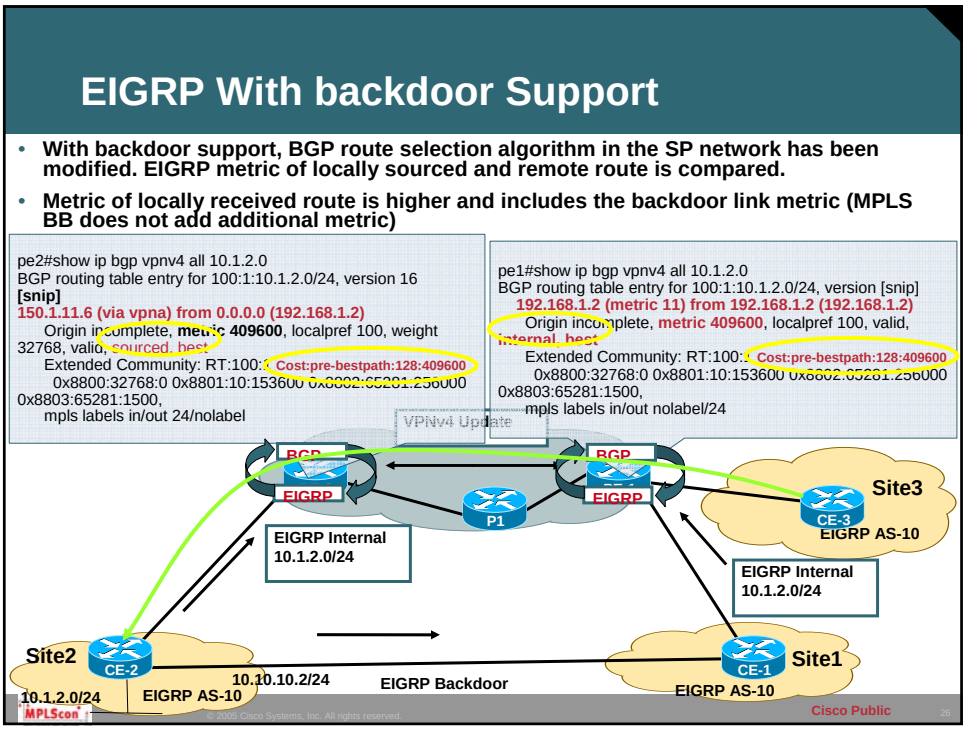
EIGRP Without backdoor Support

- Site 1 and Site3 are connected to PE1. In addition a backdoor link exists between site1 and site2.
- PE1 learns the route via EIGRP and also received the same route via iBGP from PE2.
- EIGRP route redistributed in BGP becomes locally sourced and is preferred over iBGP learnt route
- Site3 traffic destined for Site 2 arrives on PE1 but afterwards traverses site1 instead of MPLS BB.



EIGRP With backdoor Support

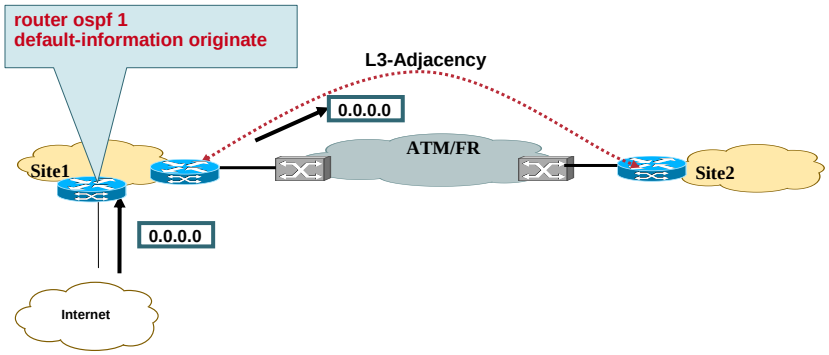
- With backdoor support, BGP route selection algorithm in the SP network has been modified. EIGRP metric of locally sourced and remote route is compared.
- Metric of locally received route is higher and includes the backdoor link metric (MPLS BB does not add additional metric)



Agenda

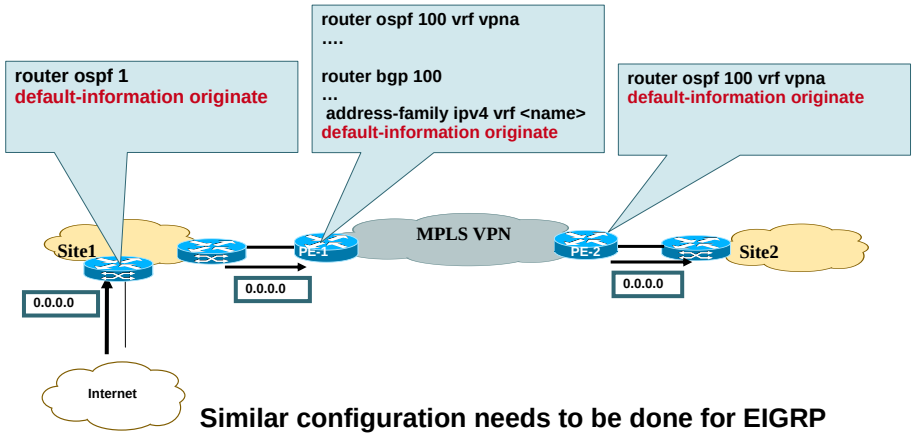
- Introduction
- Physical Migration to MPLS VPN Backbone
- Routing considerations using
  - BGP as PE-CE protocol
  - OSPF as PE-CE protocol
  - EIGRP as PE-CE protocol
- **Default route handling in MPLS VPN**
- Preventing routing Loops with SOO
- Limiting vrf routes and potential black holing
- Multi-homing Scenarios
- Summary

Default Route Origination in the Traditional Environment



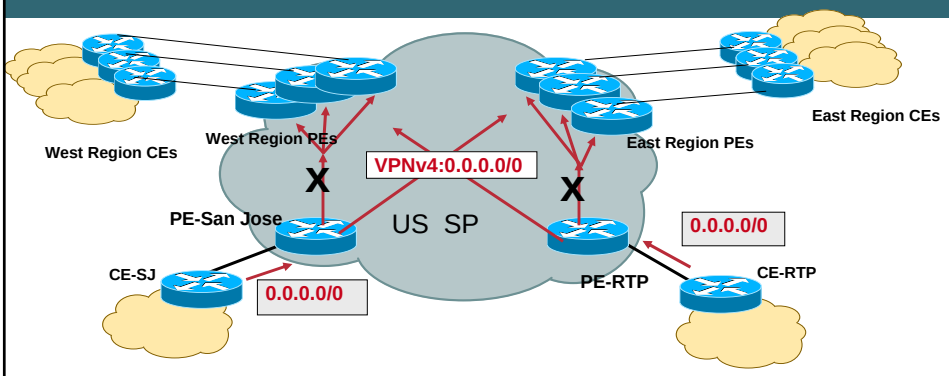
Default Route origination (OSPF/EIGRP)

BGP by default does not redistribute 0.0.0.0/0



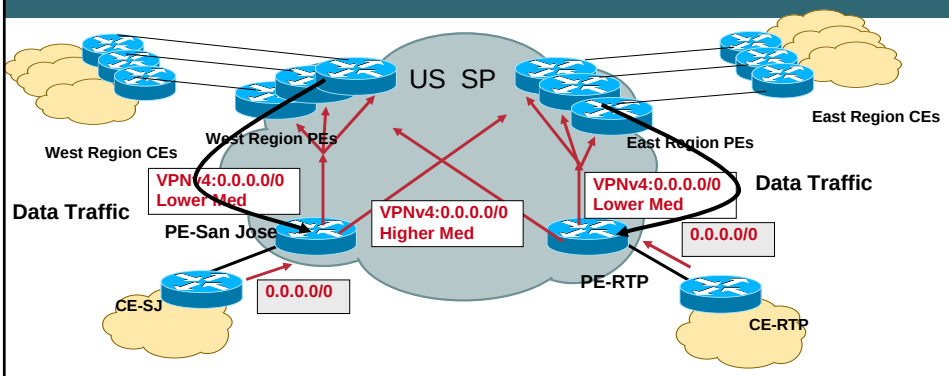
Similar configuration needs to be done for EIGRP

Default Route in Multi-hub Environment Design Objective



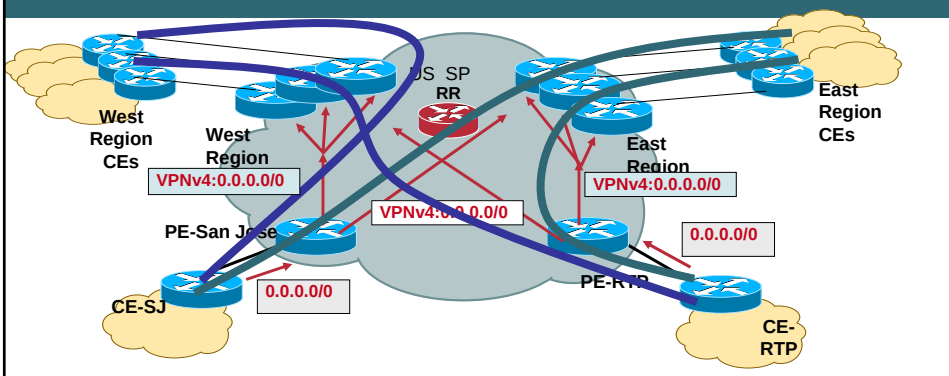
- Both San Jose and RTP advertise Default routes to the spoke Sites
- Satellite Sites in West Coast Region should take the default route to SJ and East Coast Sites should use RTP for default route
- In case of failure, spoke Sites should take the non-preferred default route

Default Route in Multi-hub Environment Possible Solution



- Over here it is proposed that when we advertise default route it would be in such a way that West Region PEs receives a lower med from SJ and higher med from RTP .
- Similarly East Region PEs receives a default route with a lower med from RTP and higher med from SJ
- In this way if SJ lost is route West Coast can then revert to the RTP
- Note: We have used med as an example any other BGP attribute can be used

Default Route in Multi-Hub Environment Possible Solution



- GRE tunnel between the sites to prefer one over the other exit site

Agenda

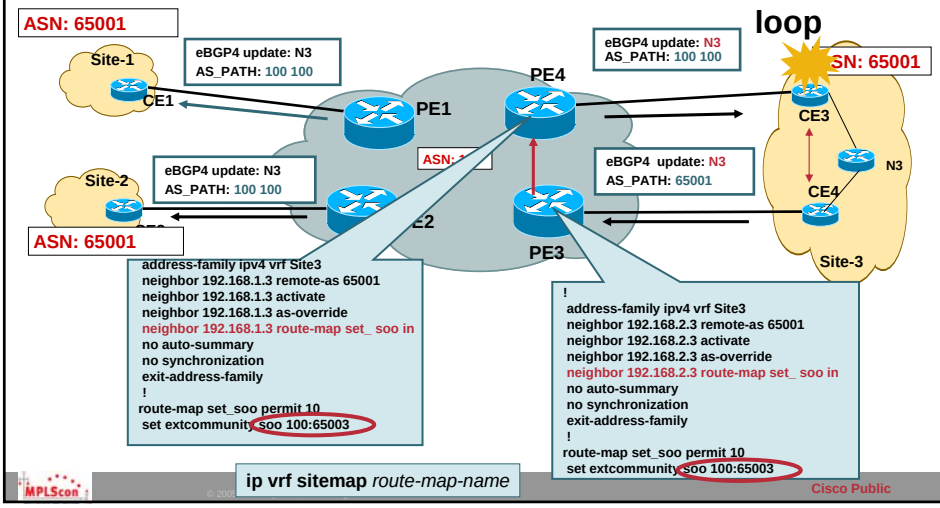
- Introduction
- Physical Migration to MPLS VPN Backbone
- Routing considerations using
  - BGP as PE-CE protocol
  - OSPF as PE-CE protocol
  - EIGRP as PE-CE protocol
- Default route handling in MPLS VPN
- **Preventing routing Loops with SOO**
- Limiting vrf routes and potential black holing
- Multi-homing Scenarios
- Summary

Implementing SOO for Loop Prevention

- **The SOO (extended BGP community) can be used to prevent loops in these scenarios.**
- **The SOO is needed only for multihomed sites.**
- **When EBGP is run between PE and CE routers, the SOO is configured through a route map command.**
- **For other routing protocols, the SOO can be applied to routes learned through a particular VRF interface during the redistribution into BGP.**

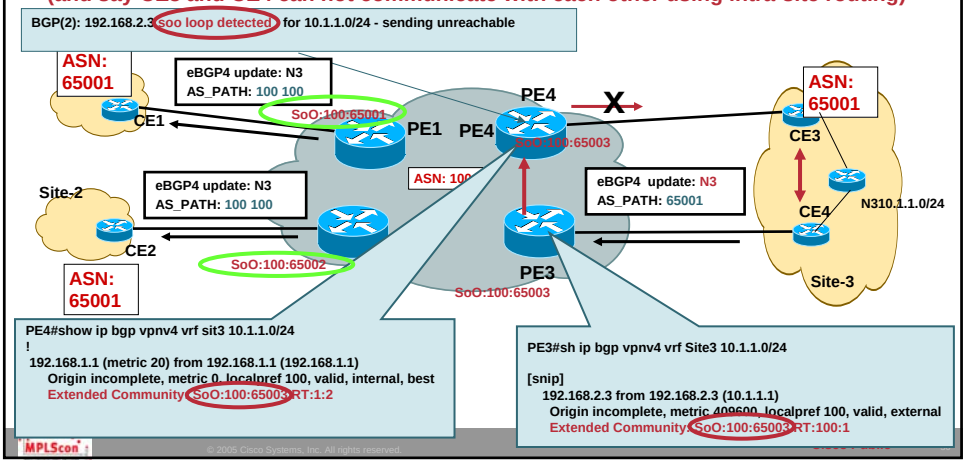
Avoiding loops with SOO

- Not a hub and spoke scenario
- You don't want the routes sent from site3 CE4 to be sent back to site3 via PE4



Avoiding loops with SOO

- PE3 and PE4 are configured with the same SoO value
- If SoO in the BGP update matches with the configured value, update will not be forwarded to CE3
- Note: In fact PE4 will never forward the update to CE3 even if the site-3 is segmented (and say CE3 and CE4 can not communicate with each other using intra-site routing)



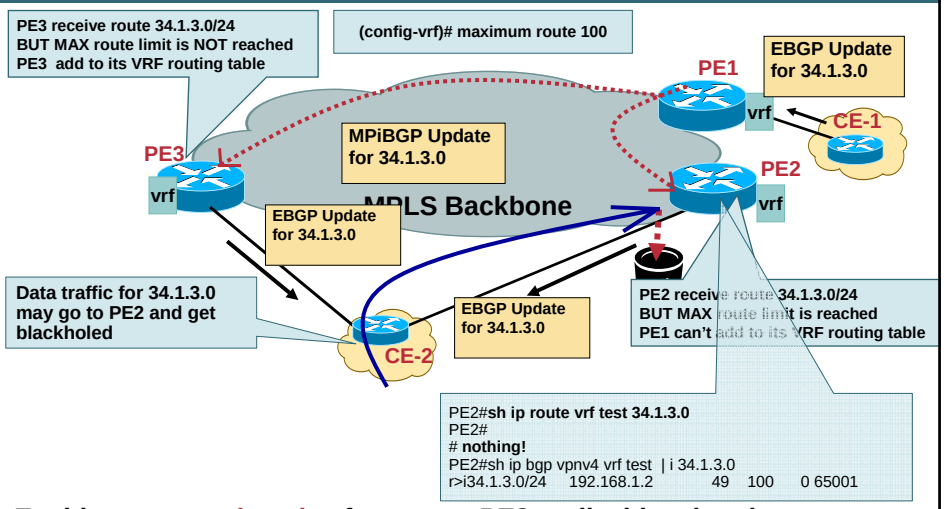
Agenda

- Introduction
- Physical Migration to MPLS VPN Backbone
- Routing considerations using
  - BGP as PE-CE protocol
  - OSPF as PE-CE protocol
  - EIGRP as PE-CE protocol
- Default route handling in MPLS VPN
- Preventing routing Loops with SOO
- **Limiting vrf routes and potential black holing**
- Multi-homing Scenarios
- Summary

VRF route limit

- **VRF route limit allows the Service Provider to protect his PE routers from uncontrolled route advertisements from CE routers**
- **VRF route-limit allows to limit the number of routes that are imported into a VRF**
  - **Routes coming from CE routers**
  - **Routes coming from other PEs (imported routes)**
- **The route limit is configured for each VRF**
- **If the number of routes exceed the route-limit**
  - **Syslog message is generated**
  - **Routes are not inserted into VRF anymore**
    - **Optional**

Max Routes exceeded- Route propagation Potential Blackholing



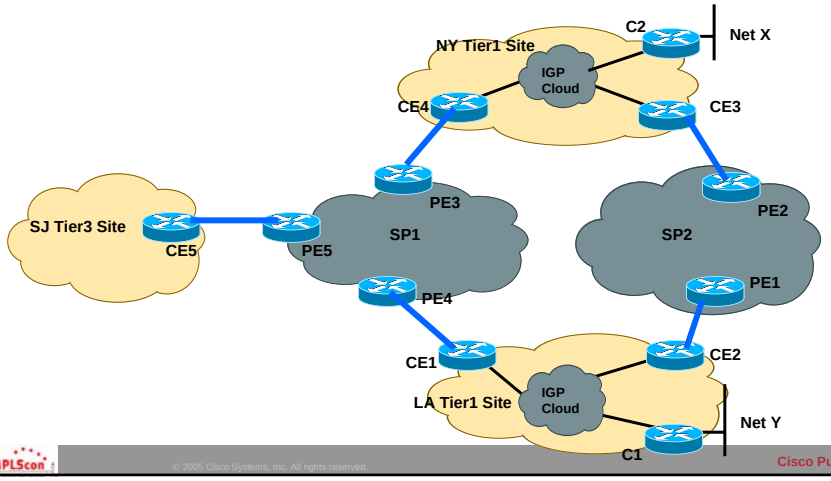
- Enable suppress-inactive feature on PE2 to disable advertisements of BGP routes that don't make it to the routing table

Agenda

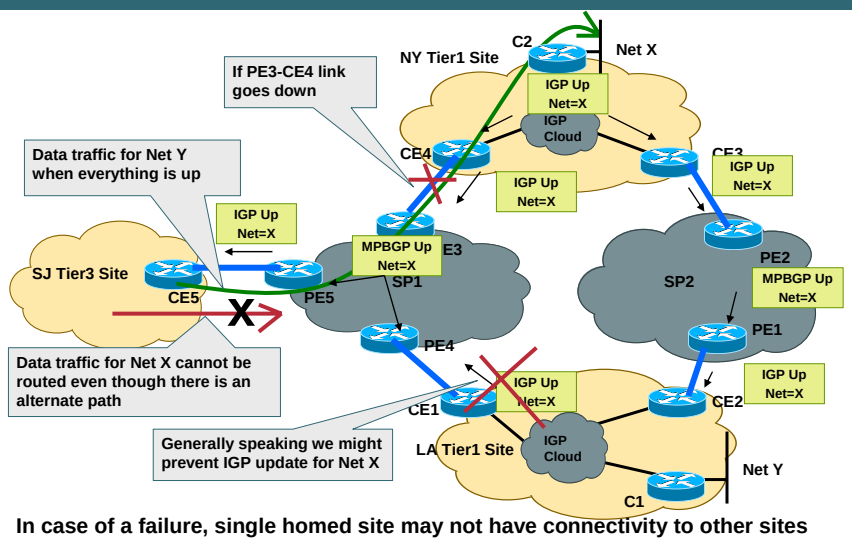
- Introduction
- Physical Migration to MPLS VPN Backbone
- Routing considerations using
  - BGP as PE-CE protocol
  - OSPF as PE-CE protocol
  - EIGRP as PE-CE protocol
- Default route handling in MPLS VPN
- Preventing routing Loops with SOO
- Limiting vrf routes and potential black holing
- Multi-homing Scenarios
- Summary

Multi-tier Sites in Multi-homed Enterprise

- An enterprise might choose multiple providers for their L3VPN services
- It is possible that some of the enterprise satellite sites might be single homed.
- Unpredictable routing behavior may occur in the steady state or after a failure

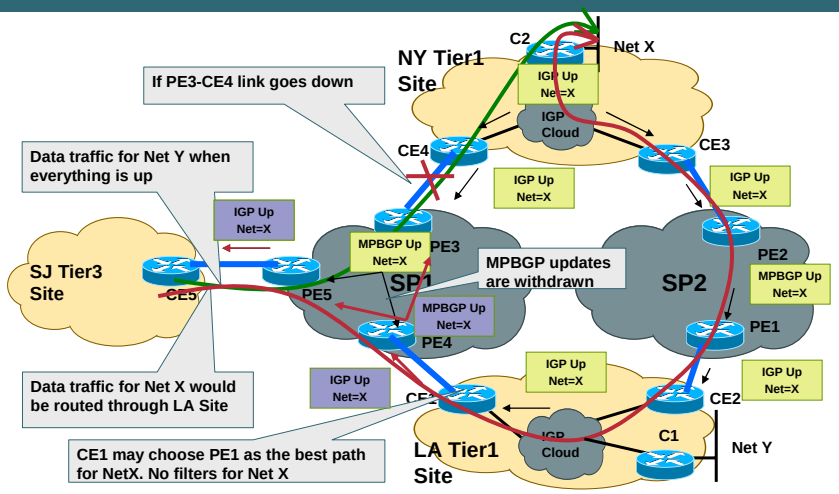


Tier3 Site transiting Tier 1 Site- Problem



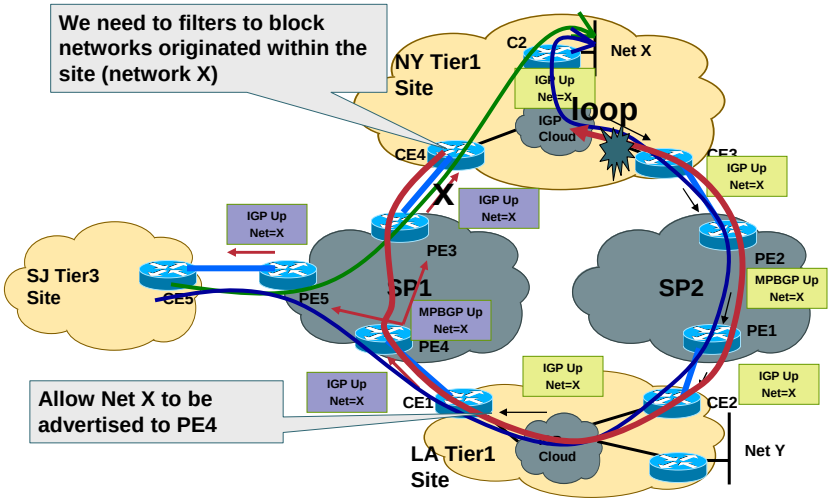
- In case of a failure, single homed site may not have connectivity to other sites
- Even though an alternate path exists but update was blocked to ensure traffic doesn't take sub-optimal path by transiting the enterprise site

Tier3 Site transiting Tier 1 Site – Possible Solution



- Don't filter the routes that do not belong to the site
- SP cloud now sees two routes. With appropriate metric manipulation, PE5 can choose path via PE3 as the primary path.
- In case of failure, an alternate valid path will be available via PE4

Tier3 Site transiting Tier 1 Site Suboptimal Routing and Routing Loops - Caveat



- CE4 can possibly choose PE3 as the best path for Net X which can result in suboptimal routing and possible routing loops

Agenda

- Introduction
- Physical Migration to MPLS VPN Backbone
- Routing considerations using
  - BGP as PE-CE protocol
  - OSPF as PE-CE protocol
  - EIGRP as PE-CE protocol
- Default route handling in MPLS VPN
- Preventing routing Loops with SOO
- Limiting vrf routes and potential black holing
- Multi-homing Scenarios
- **Summary**

Summary

- For large enterprises, migration to L3VPN service requires a phased approach so that disruption to existing services is minimal
- Existing site local routing protocols policies and their interaction with PE-CE routing protocols should be carefully analyzed
- Topological considerations such as backdoor links, multi-homing scenarios, OSPF areas placement and BGP AS number scheme etc should be taken into account to avoid sub-optimal routing or loops.
- Default route and Summarization is important for proper routing to the internet or to the central sites and could be coordinated with the service provider for optimal results.
- Site-to-site VPN routing convergence should be kept in mind while deploying delay sensitive application
- Redundancy and Multi-provider topologies may result in loops if not properly implemented.

Q and A

CISCO SYSTEMS