Impacts of utilizing eiBGP Maximum-Path load sharing in a cloud environment.

White paper

Organizations that maintain multiple paths into data centers have a desire to utilize both paths simultaneously in a steady state. This type of configuration provides resiliency in the event of a hardware or circuit failure while ensuring that each individual circuit is fully operational. This also maximizes an organization's assets for capacity when both circuits are available. Careful consideration should be used when planning for the proper sizing of the data center circuits for growth while not over-sizing the connections and incurring undue costs.

Bandwidth resources are less constrained and deployed more rapidly as customers migrate from traditional data centers into the cloud environment. Also, the connections are more reliable since the interconnects are typically cross-connects within the same facility instead of circuits provisioned from the network provider edge to the organization's data center location.

Distributing traffic across multiple circuits

There are multiple techniques that can be utilized to distribute traffic across redundant paths in the traditional data center. eiBGP Maximum-Path is one of the more popular techniques to load-share traffic across multiple circuits into a single location.

The Provider Edge router (PE) will load-share traffic using IP prefixes with identical BGP attributes on a per-flow or session basis if the BGP metrics are equal. This will be considered a BGP tie and two equal cost routes will be injected into the routing table instead of utilizing the default behavior of injecting a single path. Note that the use of eiBGP Maximum-Path has the potential to cause undesired results when implemented in a cloud environment which will be discussed later in the article.

Diagram 1 shows how eiBGP Maximum-Path is implemented in a traditional data center environment.

Diagram 1 -EiBGP Maximum-path Design

- The cloud on the left represents a customer data center advertising the same BGP prefixes utilizing the same metrics originating from the same AS Number.
- eiBGP Maximum-Path is implemented on the Verizon remote site PE routers. This will allow the iBGP core of Private IP to accept multiple BGP paths for the same IP prefix.
- To utilize BGP Maximum-Path for load-sharing all BGP attributes must be the same including the weight, local preference, autonomous system path (entire path and not just length), origin code, and Multi-Exit Discriminator (MED). Also the BGP next-hop IP address for each path must be different.

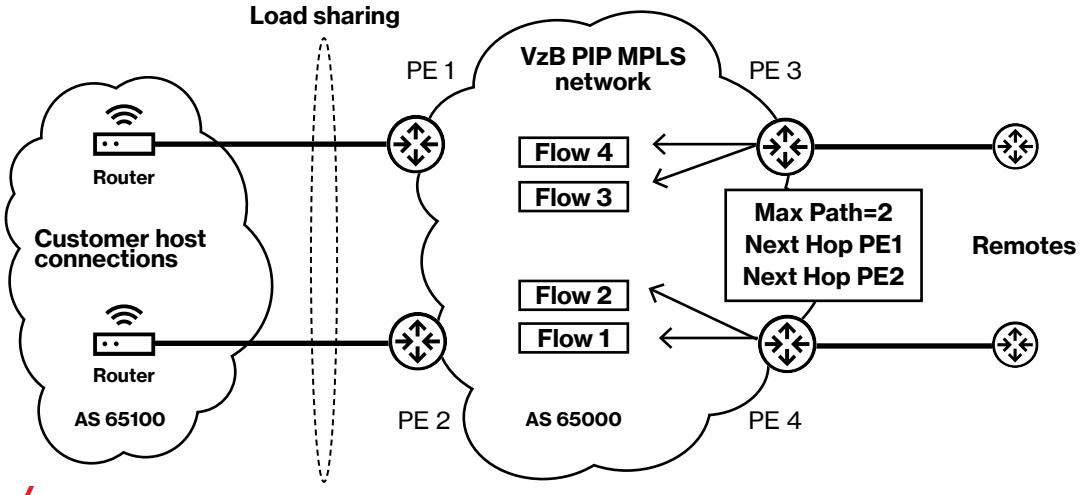

Diagram 1

Unique BGP routes based upon originating AS number.
Backbone IGP determines which data center for load-share.

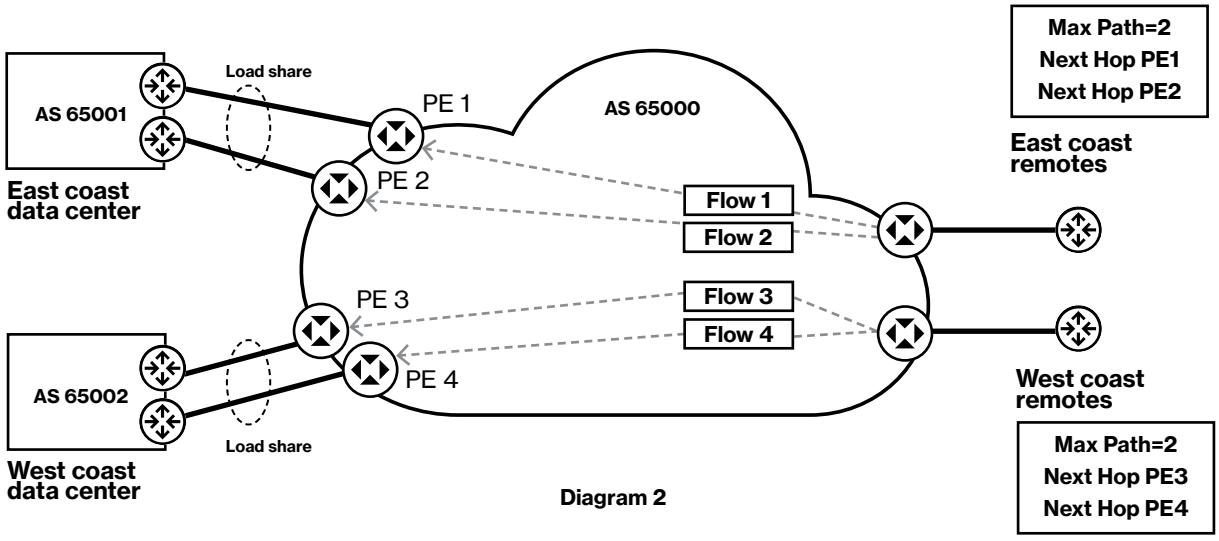

Diagram 2

eiBGP Maximum-Path Using Multiple Regional Data Centers

In a traditional environment where organizations utilize data center redundancy in addition to circuit diversity, data centers are typically distinguished by utilizing different AS Numbers for each individual location. With this configuration, organizations are able to load-share traffic across multiple circuits into a single data center while keeping the individual data centers separate from a routing perspective when advertising the same route prefixes. eiBGP Maximum-Path will differentiate the data center locations based on different AS Numbers and will load-share to the data center AS with the most attractive Interior Gateway Protocol (IGP) cost.

In the example above (Diagram 2), the remote PE router will send the traffic to the East Coast Data Center or the West Coast Data Center based upon IGP cost from each remote site PE.

eiBGP Maximum-Path in a Regional Cloud Environment with multiple locations

As organizations migrate their applications into a cloud environment, the desire is to provide the same resiliency design compared to their legacy data centers. However, developing a load-share design utilizing multiple paths into the cloud data center is not as beneficial since the network connections are more robust and load-sharing designs are more complex to manage. These connections are typically cross-connects between the network edge equipment and the cloud providers within the same campus environment.

This methodology reduces the number of network components that can fail as well as reduce the distance of these network connections. The net result is the network connections into the cloud center are more elastic compared to installing network connections into an organization's data center. The primary and backup connections are built to scale for the entire bandwidth capacity required and can easily change as application traffic grows. This reduces the need for load-sharing traffic over multiple connections.

Many organizations will duplicate data center designs in a cloud environment, including configuring services in different regions of a given Cloud Service Provider (e.g., one deployment in a cloud center in the East and a duplicate deployment in the West for geographical resiliency). Some cloud providers utilize a single BGP Autonomous System Number regardless of the cloud center location. The use of Maximum-Path with MPLS might cause undesired routing behavior if not addressed during the migration. If the organization advertises the same routes over the cloud provider's BGP peering connections using the same AS Number in multiple cloud center locations, the traffic will be shared across cloud centers, which can result in:

- Traffic being directed to cloud centers that are not geographically close to the hosted applications, causing additional latency.
- Asymmetrical routing with the traffic forwarded into one cloud center in one direction and returned through another cloud center.
- Increasing the potential for routing anomalies since the cloud service provider boundary utilizes the same AS Number.

Diagram 3 shows the unintended consequences of utilizing the same data center load-sharing design in a cloud center environment.

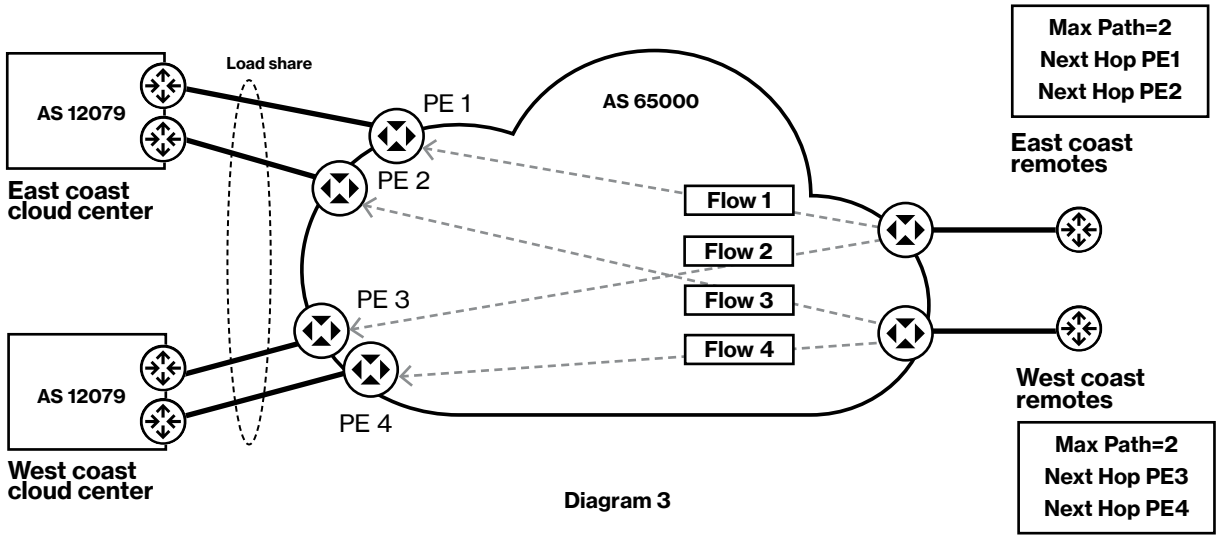

Diagram 3

Common Techniques used to address issues with load-sharing using eiBGP Maximum-Path in a cloud environment.

There are a few alternative solutions to the aforementioned load-sharing design challenges including:

- Removing the eiBGP Maximum-Path Feature from the MPLS PE routers. This will remove load-sharing for the entire enterprise and revert back to BGP decisions based upon IGP cost. The calculation of the enterprise is that the effort to design the cloud architecture to match the data center architecture isn't worth all the complexity. Some organizations decide that load-sharing traffic in their data center isn't a true requirement and remove the load-sharing feature altogether.
- Configuring a unique AS Number per Cloud Center if this is a feature that the cloud provider can support. This will allow the customer to maintain the load-sharing configuration at their legacy data center and cloud environment since a unique AS Number is being used to identify each data center and cloud location.
- The cloud operator does not allow customers to configure the AS number, but, instead, works with Verizon to configure the IP port connected to the Cloud provider with a BGP Replace-AS configuration on one or all of the cloud center connections. The Replace-AS feature will add an additional AS Hop, so some organizations will apply this feature to all cloud center connections to make the BGP paths equal.
- Utilizing unique IP prefixes per Cloud Center and differentiating the individual cloud location by IP block while in the planning stages. If each cloud center location utilizes a specific subnet that is a part of a larger summary subnet, this will provide a unique BGP path per physical location.

Once organizations are aware of the potential issues with utilizing load-sharing designs when migrating services into a cloud environment, techniques can be developed to address the problem. Each cloud provider will have different policies and tools that can be used by the organization to control routing over the network. Cloud connectivity models are changing and will continue to change as customers with more sophisticated designs migrate to cloud environments. Careful consideration will be required by IT organizations to ensure that existing data center designs can be replicated in the cloud environment without disruption or design changes.

Verizon has built a delivery model that allows organizations to connect into an ecosystem of cloud operators with techniques and best-practices that address many of the common challenges associated with cloud migrations. We continue to invest in platforms that automate the implementation process and reduce the complexity of accessing cloud environments. Our experience and capabilities position Verizon as a valuable partner as organizations navigate to the cloud.

If you would like to discuss this further with Verizon, please contact us at tech-expert@verizon.com.