Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: September 7, 2012

W. George
Time Warner Cable
R. Shakir
BT

March 6, 2012

IP VPN Scaling Considerations
draft-gs-vpn-scaling-01

Abstract

This document discusses scaling considerations unique to implementation of Layer 3 (IP) Virtual Private Networks, discusses a few best practices, and identifies gaps in the current tools and techniques which are making it more difficult for operators to cost-effectively scale and manage their L3VPN deployments.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on September 7, 2012.

Table of Contents

1. Introduction

As IP networking has become more ubiquitous and mature, many enterprises have begun migration away from legacy point to point or layer 2 virtual private network (VPN) implementations towards layer 3 VPNs. The VPN implementation as defined by RFC 4364 [RFC4364] enables flexible and robust implementations of IP VPNs. However, in practice, it has become clear that it suffers from significant scaling considerations beyond those discussed in RFC4364. In many cases, the limits of scale for a given platform are not in sync with the maximum physical and logical interface density supported by the platform, such that a platform may be considered "full" long before the physical slots and ports have all been filled with equipment and connections. This represents an inefficient use of space and power, as well as stranded capital assets, which increase the operator's cost to provide the service as well as the complexity of managing the platform to ensure proper service levels in a wide variety of circumstances. While these scaling considerations are somewhat similar to the scaling concerns experienced in the Global Internet, those are at best a subset of the overall problem, and may not have a great deal of overlap between solutions and best practices. The added complexity and feature set required to support today's enterprise IP networks drives additional scaling considerations for large deployments. A common response to concerns about control plane scale is simply to "throw hardware at the problem" in the form of ever-increasing amounts of memory and CPU resources. In some cases, this may be the only solution, but similarly to the concerns identified in RFC 4984 [RFC4984], there are limits to the growth curve that can be supported and cost-effectively deployed by a VPN provider such that their service remains profitable, and therefore it is necessary to explore the potential for optimization to make the existing resources stretch further.

Generally, router scale can be considered in one of three areas: forwarding capacity, interface density, and control plane capacity. This draft will focus almost exclusively on control plane capacity, because while the others are important considerations for most operators, they are less affected by the details of how L3VPN is implemented either by the router vendor or the operator. Interface density is usually a factor of the forwarding capacity of a given module or slot as well as physical packaging. In this application, interface density is interesting from the perspective of its impact to the control plane - more interfaces means more of all of the different factors that contribute to control plane load, and the operator wants to be able to strike a balance between interface density and control plane capacity such that neither grows out of pace with the other.

1.1. Intention of this Document

This document is intended to provide a discussion of the challenges that network operators face in deploying large-scale L3VPN environments at the time of writing, with two key sets of recommendations. As such, these outcomes can be divided into those that apply to network operators regarding the deployment of particular technologies, and those that apply to network protocol and operating system implementors relating to allowing better understanding of scaling characteristics in deployments of such equipment.

The best practices defined in this document are intended to allow more optimal scaling of L3VPN networks, whilst minimising the impact on end-customer network behaviour. It is intended that such guidance can be directly utilised by Service Providers to improve the scaleability of network elements. However, the guidance in this document should not be viewed as a panacea to the problems of scaling network elements. It is the intention of the authors to document a number of key problems experienced in such environments and provide information to the SP that may result in more optimal deployment of existing technologies to this audience. It is appreciated that there is a point at which the limits of hardware will be reached, and hence new network elements are required. The key intention of the recommendations provided to Service Providers within this document are intended to allow the resources that exist within existing elements to be utilised in the most efficient manner. Clearly, the optimal point in this balance is that the data-plane and control-plane scale to support similar levels of service termination, so as to result in minimal "over provisioning" of one element.

The scaling considerations presented in this document are intended to provide both network operators and network equipment implementors further guidance around the toolset, and information required to provide accurate means of capacity planning in L3VPN environments. Again, the authors consider that the scaling characteristics, and toolsets required of L3VPN PE equipment diverge somewhat from those required by Internet network equipment. In Internet deployments, relatively standardised interconnects exist across all deployments - typically utilising either static routing, or BGP-4. As such, each connected port comes with a relatively standard overhead in terms of the protocols required. Whilst there is some variance in how "chatty" each customer connection may be, this is balanced by the fact that the whole Internet routing table is typically held on such edge equipment (and hence individual customer's instability tends to be relatively small when compared to the instability of the Internet DFZ). In addition, since such instability is limited to relatively few impacts to a node (interface or BGP session flapping, and BGP

UPDATE messages) routers can be optimised to cope with such instability. Counter to this, the L3VPN environment does not have a standardised connectivity model, and typically connects to much less controlled environments. Further details of this are provided within later sections of this document. The result of this difference is that 'headline' scaling figures presented for particular equipment tends to be of limited utility to a network operator. The recommendations within this document outline some of the considerations that must be made in considering the scaling of such elements, and provide guidance as to the missing inputs and tools that are required to provide information around the capacity of such elements.

1.2. Horizontal vs. Vertical Scaling

Within this document, two forms of 'scaling' are referred to - the "throw hardware at the problem" approach outlined previously involves deploying additional network elements in order to provide further network capacity. Throughout this document, this approach is referred to a horizontal scaling - insofar as it requires parallel deployment of numerous similar elements and balancing the load across the combined capacity of all of the elements. The approach of increasing the capacity of an individual node through allowing the control plane capacity to support the maximum forwarding plane capacity (be it data forwarded, or available ports) is referred to as vertical scaling. It is obvious that at some point the approach of horizontal scaling of elements is required - due to either exhausting available port capacities, or available forwarding plane - however, it should be noted that there are a number of motivations for delaying such provisioning, some of which relate directly to the characteristics of L3VPN environments.

Since a significant proportion of the customers who purchase L3VPN services are Enterprise customers, typically, the service is utilised as a WAN for their inter-location connectivity. Clearly, as such customer base tends to be distributed based on differing factors, this implies that such customers connect in numerous geographical locations. The requirement to support service in these locations therefore results in a requirement for the service provider network architecture to support geographically distributed access into such services. A balance must be struck between the extent to which access networks are utilised to backhaul traffic to the service layer, and the geographical distribution of the service layer itself. Both scale and performance characteristics of such networks tend to result in more geographical distribution of service layer elements than in Internet deployments. This distribution results in two particular changes - primarily that the idea of a "point-of-presence" must be reconsidered - where an assumption in Internet environments

may be that there are separated core and access elements within a single location, within a distributed L3VPN environment, a point of presence may be a single PE device. The result of these small scale points of presence is that numerous core and edge functions must be collapsed onto a single device. For this reason, the approach of adding additional devices to the network may have an impact on a further subset of devices within the network (particularly due to any mesh-based protocols that are deployed), and hence result in a change in the scaling characteristics of these devices. In this case, there is further motivation to avoid large number of devices in the network where possible. Further to this, the smaller PoP profile may result in physical constraints around the deployment of additional network elements, particularly due to the availability of power and physical space to deploy such elements.

1.3. Developing Requirements for Scaled L3VPN Environments

Whilst the collected scaling considerations outlined in this document are based on the author's collective experience within various Service Provider networks, and discussions with operators of similar networks, it should be noted that the problems outlined in this document are not static. With the growth in the use of IP as the underlying transport of many services, the demand for L3VPN environments has grown. As such, this has meant that various technologies are being considered to allow growth of these networks at a lower cost point to a wider footprint than was previously required. A network operator must therefore consider the extent to which the service layer must be built - both to meet economic and technical requirements. With newer aggregation methods, the service layer edge (and hence the L3VPN PE) acquires responsibility for inter-working between newer dynamic aggregation technologies, and the existing IP network. As such, these edge functionalities result in further requirements for loading onto these network elements.

*** Author's note: Do we want to put anything about NNI for footprint extension here? Datacenter edge - perhaps Ning's problem around the L3VPN edge in his datacentres? ***

1.4. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2. PE-CE routing protocols

One of the things that makes IP VPNs so flexible and robust is their ability to participate in the encapsulated network's routing protocols, where the customer edge (CE) router has a direct neighbor relationship with its upstream provider edge (PE) router in order to exchange routing information about the Virtual Route Forwarding (VRF) instance that represents the VPN. In many cases, this is managed through a combination of static routes and BGP neighbors, but IGPs such as OSPF RFC 4577 [RFC4577] are often supported, because it enables a more complete integration into an existing enterprise network design and topology. In some single-vendor implementations, carriers sometimes support proprietary routing protocols such as EIGRP [EIGRP]. IGPs may also be chosen due to a belief that they will respond more rapidly during a failure than BGP will. In reality, this may not be true due to the fact that VRF routing information is still carried in MP-BGP from PE to PE, and the PE-CE routing protocol's characteristics are only locally significant. In fact, the increased overhead may lead to slower convergence times than a more standard BGP implementation.

IGPs often translate to a significant increase in overhead due to their inherent characteristics as link-state routing protocols requiring full topology databases and flooding of updates to all participants, and the fact that they invoke additional processes on the router when compared to simply using BGP (which is already going to be running on a router using MP-BGP for VPNs). While a router may be able to scale almost effortlessly with a few thousand routes in a single IGP plus hundreds of thousands of routes and many neighbors in BGP, it may be quickly challenged if it is also required to run multiple instances of an IGP each with a certain number of routes that must be moved into MP-BGP to be passed to the rest of the VPN infrastructure. The advent of support for IPv6 within a VPN (6VPE) [RFC4659] has the potential to make this problem worse, especially in the case of OSPF, where it now requires both OSPFv2 [RFC4577] and v3 [I-D.ietf-l3vpn-ospfv3-pece] to run as separate instances for the two address families.

Another consideration in PE-CE routing protocols is the timers used for each session. These will be discussed in greater detail in the best practices section.

2.1. Best Common Practice

Ultimately, the decision as to which PE-CE routing protocols to support is a business decision much more often than it is a technical one, because there are few use cases where something other than BGP and static routing as PE-CE routing protocols is a technical

requirement. If a provider chooses to support additional protocols, especially IGPs, they should consider the effects that these have on the overall scaling profile of the PE routers and the network as a whole when determining if and to what extent they will support other protocols.

Often, those designing VPN solutions attempt to use extremely aggressive routing protocol timer and keepalive values as a means of rapid failure detection and reconvergence. This tends to make PE-CE routing protocols more fragile and increase the load on the PE router with questionable benefit. This is especially common in scenarios where the network designer is attempting to replicate native IGP-like failure detection and reroute capabilities using BGP. In order to avoid this, the preferred values should be set to something that is appropriate for large-scale implementations (*** do we want to make a specific recommendation? ***). Further, because timer and keepalive values are often negotiated based on the more aggressive neighbor, it is a good idea to set a minimum acceptable value, so that instead of being forced to support negotiated timer values that are too aggressive for the scale that a given PE router is expected to support, the neighbor session will simply stay down until the remote end timers are reconfigured to a more acceptable value. This acts as a safety valve against abuse that can destabilize a router used by multiple customers. Because aggressive timers may be unavoidable in certain situations, it may be advisable to track the number of sessions which are provisioned with aggressive timers vs how many are using more conservative timers on a per-router basis, so that effort can be made to balance aggressive and conservative timers on each router. This will help to prevent "hot-spots" where given a similar port and VRF density, some routers have significantly higher CPU usage in steady-state than others.

It is important to realize that while use of aggressive routing protocol timers is not a scalable way to do fast failure detection, fast failure detection is still a requirement for many customers. Because this is becoming such a table-stakes requirement, the provider must consider other alternatives such as Bidirectional Forwarding Detection ([RFC5880]), Ethernet OAM 802.1ag [IEEE802.1], ITU-T &.1731 [Y.1731] LACP 802.3ad [IEEE802.3] and the like. These extensions often come with their own scaling considerations, but more and more they are implemented in a distributed fashion so that instead of affecting the main router CPU like a routing protocol might, they offload that processing to the linecard CPU, and therefore can support more aggressive scale. The general philosophy is that these lower-layer detection mechanisms should serve as the primary detection and failure point, with the upper layer routing protocols only serving as a backstop if the failure is not detected by the lower level protocols for some period of time.

2.2. Common Problems at Scale Limits

Two common problems when working on a heavily-loaded system:

CPU cycle constraints, even before the system reaches the point of scheduler thrashing often lead to one or more routing protocol neighbor hello drops. If several consecutive drops occur, the remote neighbor may declare the session dead, which triggers a restart of the connection and a resync of the routing data. Because this connection initialization requires dedicated CPU cycles to generate, receive, acknowledge, and process the updates, it increases the CPU utilization further, which may trigger additional hello failures and neighbor resets, resulting in a snowball effect where a relatively minor event rapidly becomes a major one due to interactions between multiple scaling limitations. This problem is made worse by extremely aggressive timer values, because they raise the baseline CPU load with more frequent hellos and responses, and are more sensitive to drops caused by increased CPU load. Further, because failures brought on by loss of hello packets are unlikely to invoke any graceful restart [RFC4781] machinery that the system may support, it is unlikely that the session reset will be able to take advantage of optimizations like only synching the changes that occured while the session was dead, thus increasing the outage time and the CPU cycles to get things back into sync.

Another potential issue during times of high-CPU operation is related to process prioritization. This is applicable in different ways for both multithreaded and interrupt-driven OS architectures. In each case, the scheduling algorithm that the router uses to prioritize different CPU cycle work items and manage the timeslices individual tasks are given to complete may require significant tuning and prioritization in order to ensure the desired behavior during high CPU usage. Improperly tuned or prioritized processes may significantly delay completion of routing table/update processing such that it may take an excessive amount of time for the routing table to converge properly. This issue is further exacerbated if the VRF instance has a large amount of routes, or is prone to frequent event-driven route churn. In some cases, the routing table in a given VRF may never fully converge, leading to routing loops, traffic loss, inconsistent latency, and a generally adverse customer experience.

It worth noting that these items also have a cascade effect on other routers in the system that participate in a given VRF that is being affected by this type of scaling issue. Not only is the local PE router affected, but any upstream Route reflectors, as well as other PEs, and even CEs participating in this VRF will see increased CPU cycles in order to receive and process the increased flow of updates

driven by the local churn.

specific items related to different PE-CE protocols?

3. Multicast

Multicast support within a VPN [RFC6513] has become an increasingly popular feature, but comes with its own scaling considerations. Depending on the application, the frequency at which multicast state changes within a given VPN (e.g. PIM joins and prunes) will contribute to the CPU load on the router, and any instability in the network can potentially increase these as remote sites flap. In extreme cases, PIM neighborships can be lost during events, disrupting the flow of multicast traffic.

It should be noted that, in some cases, dynamic action is required by a PE device to support the transition of flooding of multicast data from a non-optimal distribution tree (the default MDT in [RFC6037], or the I-PMSI) onto a more optimal one (a data MDT or S-PMSI). Where such a transition is required, consideration is required of the nature of the traffic sourced by an end user of the L3VPN service. The net result of this consideration is that it becomes increasingly difficult to reliably gauge the scaling impact of specific end-site deployments. Additional scaling considerations around Multicast in a VPN are related to the size and number of multicast streams. While this is a consideration whenever Multicast is used even outside of a VPN because of the bandwidth utilization it may generate in the core, the additional overhead of implementing multicast within a VPN makes this a more siginificant consideration in this case. Related to the previous consideration is the stream fanout - the amount of P and PE router paths in the network that could potentially carry a given multicast stream based on the number of PEs that are configured with a given Multicast-enabled VRF, and the number that actually do carry the stream based on actual receivers joining the stream behind that PE.

*** This section is quite weak. We're looking for contributors who can assist in fleshing this out ***

3.1. Best Common Practices

Multicast BCPs???

3.2. Common Problems at Scale Limits

Multicast tree interruptions

PIM neighbor adjacency drops

4. Network Events

Network events are an important scaling consideration because they can have wide-ranging impacts far beyond the individual VRF or even PE router that experiences the event. At high scale, a seemingly innocuous event on one router or VRF can trigger secondary impacts and outages on remote routers elsewhere in the network. Correlating these events for root cause analysis can be challenging by itself, and trying to characterize the impacts as they relate to scale in a way that informs the provider's decisions is even more difficult. Different types of Network Events that can contribute are: Interface flaps, hardware and software outages (both planned and unplanned), externally driven route-churn events (such as those that originate on an NNI partner's network) and configuration changes.

4.1. Best Common Practices

While this document suggests that lower layer failure detection protocols like BFD and Ethernet OAM be more aggressive so that routing protocol timers can be more conservative, it is still important to remember that this can generate false positives or excessive churn that will cascade into a scaling problem at other parts of the system, so the timers should not automatically be configured to their minimum supported values. Rather, each application may be slightly different, and the timers should only be set as aggressively as necessary to ensure acceptable performance of the applications in question. It may be appropriate to set limits (e.g. in provisioning logic/rules) as to the number of interfaces per router and per VRF that can use aggressive, moderate, and conservative interface timers.

Even with timers set as conservatively as the application will allow, churn is unavoidable. For this reason, it is also a good idea to use interface-level dampening such as hold-down timers or event dampening in order to ensure that interfaces that flap too rapidly will not telegraph that churn into the upper-layer routing protocols any more than necessary. This helps to ensure that problems are localized to a single PE or even a single interface, rather than causing instability and routing churn throughout the VRF and the provider network.

In addition to interface dampening, it may be advisable to consider implementing some manner of route flap dampening to assist in reducing the impact that route churn may have on the SP's network infrastructure. This is currently fairly uncommon within VPN

environments, and is not without controversy. While it may help with scaling, it also requires each PE to maintain more state to store and compute the per-prefix penalty values, which may reduce the benefits gained by implementing RFD. Further, customers typically expect a fair amount of transparency in the provider's participation in their routing instances. Many providers and customers view a VPN or VRF as a part of the customer's internal network and therefore compartmentalized so that the customer can only affect their own routing if they have a problem with excessive route flaps. Further, if routes are dampened it requires intervention from the SP to clear the dampening, which can potentially add to the outage time that a customer experiences once the issue that triggered the dampening is resolved. Implementing RFD may even drive the need for a customer-accessible looking glass, which is far more complex in the VPN space owing to the requirement to prevent one customer from looking at another's VRF routes on a common platform.

4.2. Common Problems at Scale Limits

Network events are both a cause and a symptom of a system running at or near its scaling limits. As noted above, event-driven routing table churn or routing protocol interactions can significantly drive up CPU usage on the locally connected PE as well as on other PEs and CEs participating in the VRF. If routes are constantly changing due to a preferred path repeatedly being added and removed, latency and jitter numbers can be affected in a way that adversely effects applications sensitive to this sort of change. Network events can also be triggered by routers with high CPU, because similarly to systems which may have aggressive routing protocol timers for enhanced failure detection, systems with centralized CPU-based implementations for lower-layer protocols (such as HDLC [ISO13239] PPP [RFC1661], LACP, BFD/EOAM) may start losing keepalives and declaring outages that result in physical interfaces being torn down and restored. Again, implementations that choose timer and multiplier values or numbers of sessions at or near the maximum rated scaling for the device put the operator in a position where there is very little headroom to deal with an event that momentarily spikes CPU usage, meaning that the likelihood of a cascade failure dramatically increases.

As above, these network events may be something that occurs elsewhere in the network, and may trigger a failure on a completely different PE or CE router. The danger with this is that it is extremely difficult to troubleshoot and correlate root causes when the outage observed isn't caused by an event on the same router. Failures become increasingly non-deterministic and difficult for operators to manage and address.

5. General Route Scale

PE routers in a carrier network can have many different implementation scenarios. Some carriers implement a dedicated PE router that is only responsible for carrying VPN routes and therefore may only carry IGP routes in its global routing table, rather than a full internet routing table. Others use combined edge routers that carry full routes plus a complement of customer VPN routes, and some even place the full internet routing table into one or more VRF instances. The issue here is that the weight of all of these routes and paths must be combined when considering the maximum scale of the router, both in terms of memory footprint and in terms of convergence times. The addition of an 8-byte RD appended to the IP address to ensure uniqueness means that each VPN prefix takes up incrementally more physical space in memory than an equivalent non-VPN route. Further, the greater number of Address-families running simultaneously on the same router, the more sensitive it will be to event-induced churn since each address-family (and VRF) often has its own independent computation/SPF run. The addition of IPv6 support within both the global routing table and within a VPN adds yet another source for routing table bloat. A PE router can be running a combination of any of the following address-families:

- Global IPv4 unicast
- Global IPv4 multicast
- VPN IPv4 unicast
- VPN IPv4 multicast
- Global IPv6 unicast
- Global IPv6 multicast
- VPN IPv6 unicast
- VPN IPv6 multicast

On high-scale PE routers, the VPN routing tables are often as large as or larger than the equivalent global routing table in both number of routes and number of paths, i.e. if the IPv4 unicast table is 350,000 routes and 1M paths, the IPv4 VPN unicast table may be 400,000 routes and 750,000 paths (**** verify numbers ****). This is at least partially due to the fact that there are no constraints on the customer addressing plan within a VPN other than they cannot conflict within a given VRF, or with any extranet with which the VRF interconnects. As such, they may not necessarily adhere to any best

practices to control the deaggregation of the routing table such as hierarchical addressing, aggregation and summarization of announcements, and minimum prefix lengths. It's also quite likely that connected interfaces will be redistributed, and little or no route filtering may take place. Most PE routers use the absence of a given VRF instance (or RD/RT filtering) to limit the number of routes that they must actually carry, but this is sometimes of limited utility for a couple of reasons. First, it leads to an inconsistent routing table footprint from one PE router to the next, and it can change with every new customer turned up on the router. This leads to non-deterministic performance and scale. Second, many customer VPNs are so large and have such stringent diversity requirements that they have a presence on nearly every PE router in a provider's network, meaning that one cannot rely heavily on statistical distribution to reduce the percentage of VRFs that must be installed on a specific PE router. In addition, customers may request the use of BGP multipath for faster failover or better load balancing, which has the net effect of installing more active routes into the table, rather than simply selecting the single best path.

In addition to such intended behaviour, within many L3VPN networks, a balance must be struck between complexity in OSS such as provisioning and inventory systems, and complexity in network deployments. One such example of this is the assignment of route distinguisher (RD) attributes. Where it may be possible to assign a single RD per L3VPN instance, and hence achieve some level of route aggregation on BGP speakers within the solution, this has some consequences for both convergence in the VPN (due to BGP convergence being relied upon) and in its potential to exacerbate geographic distance between PE and Route-reflector and is therefore undesirable in some circumstances. In order to avoid this, multiple RDs are then required, which requires OSS and inventory support to control the namespace. As such, due to this requirement, often each VRF instance is deployed with a specific RD - which, whilst achieving the desired convergence effect, places load on all BGP control-plane elements of the provider network.

Total supportable route scale on a given PE router will be driven by multiple different variables, which have a roughly inverse relationship to one another: Number of VRFs per router, number of routes per VRF, number of neighbors per VRF. For example, a router can support a low number of VRFs per router if each VRF has a large number of routes per VRF and/or a large number of neighbors per VRF. Conversely, a router can support a relatively high number of VRFs if each VRF is kept to a much lower number of routes per VRF, and/or lower numbers of neighbors per VRF. This provides a baseline that then must be reduced based on the expected level of event-driven churn, the type of protocol chosen, etc. In short, this is a

difficult problem from a modeling and capacity planning perspective.

It is fairly common for the contract or Service Level Agreement between SP and customer to include a maximum limit as to how many routes can be carried in a VRF. At its most basic, this maximum can be used as a method to estimate the number of VRFs that can be present on a PE given its scaling limitations. However, there is a wide gulf between a contractual limitation of no more than N routes per VRF with a corresponding configured limit and the fact that many customers will not carry nearly that many routes. This leads to the potential for significant stranded capacity. Therefore the provider needs a way to have different tiers of "maximum routes allowed" so that the capacity management can be done in such a way as to enable better loading of PE routers to take this relationship into account (e.g. populating a PE with a combination of high-scale and low-scale VRFs). The alternative to this method would be to assume a standard maximum routes per VRF, and then similarly to the way that carriers use statistical multiplexing and oversubscription to assume that not all customers will have their pipes full of bandwidth at the same time, make some assumptions about control plane capacity. This may come in the form of an average that is calculated based on the actual size of the routes in each VRF. This has many challenges. Among them- Should it be calculated per-PE? Network-wide? What happens when there are too many VRFs that exceed the average on a given PE? How does one add control plane capacity to a "full" router? This may be a manageable model in a network with a robust and flexible provisioning system, such that high-scale VRFs can be moved between PE routers to balance the load, but each of these moves likely represents an outage for the customer and the potential for other errors to creep in, and is not likely to be attractive due to the operational costs of managing the network. In other words, it doesn't scale, but for a completely different reason.

5.1. Best Common Practices

A number of things can be done to improve the general route scaling. Most BGP sessions can be configured with a similar set of protections as they would be if they were global Internet eBGP sessions, such as maximum prefix limits, inbound and outbound prefix filtering, etc. Prefix filtering is less common within VPNs because it is treated more like iBGP, where filtering is typically not recommended (***reference?***), or as noted above, it's part of the customer's network and therefore not the SP's business/problem to do filtering in an application that can only break that customer's network. What is often more important in the case of individual VRFs is to configure an acceptable maximum number of routes that the VRF is permitted to carry. This allows the SP to control their exposure to sudden increases in the memory footprint of the routing table,

especially if a misconfiguration on the CE side leads to significant amounts of route leakage, such as to suddenly leak a significant amount of the Global Internet Routing Table into their VRF. However, it can also be used to enforce the assumptions on number of routes per VRF that the SP has used to determine what the other max scaling values such as number of VRFs per router, number of sessions per router, etc.

As noted above, the number of VRFs per router, number of routes per VRF, and number of sessions per router and per VRF are all inter-related values in the way that they contribute to overall router scale. The more of this information is known in advance based on the design of the customer's network, the more it can be used as input to the provisioning system to determine the best available PE router on which to terminate the connections for consistent loading. Since these values are usually estimates, and considerations like diverse router terminations may drive a specific choice, this is not by any means fool-proof, but is a valuable optimization to improve the density of customers on a given router and maximize the return on investment for the capacity deployed. It is worth noting, however, that many SP VPN networks have a different geographic spread than do their Internet service counterparts, where there will be more POPs with fewer routers, as it is important to provide more local handoffs to customers. This may limit the SP's flexibility in terms of homing locations and router choices, and thus may be of limited value when controlling scale impacts on individual PE routers.

*** Discuss incremental SPF, next-hop tracking, SPF timer tuning (By protocol and AF), prefix prioritization, etc? All of these are generally thought of as convergence optimizations, and may be applicable here as a way to both reduce the CPU load and ensure that behavior is more deterministic, but I'm not sure how much depth we want to get into here, especially since some are vendor-specific or FIB-specific optimizations... ***

5.2. Common problems at scale limits

As mentioned above, systems that are carrying a large number of VRFs and/or VRFs with large numbers of routes tend to be more sensitive during events due to the increased amount of periodic and event-driven processing that must be done to complete a walk of the routing table to process updates. While optimization techniques may reduce the overhead of (re)programming the FIB after an update, there are less tricks to be employed in managing the RIB, and they are often vendor-specific, which leads to a lowest-common-denominator threshold in multivendor environments.

In addition to CPU constraints, it's common for route memory

footprint to be a consideration if there are large numbers of VRFs with large numbers of routes. Similarly to the way that high scale reduces the cushion of available CPU resources to absorb temporary peaks, as memory use reaches its high threshold, allocation of the remaining memory becomes less efficient and more fragmented, such that memory allocations may begin to fail well before the available memory is actually exhausted. Depending on the specific implementation, the "largest free" may be more important than the "total free" and it may be difficult or impossible to coalesce the free memory to reduce fragmentation to an acceptable level. As with other scaling problems, a failure of this type has the nasty habit of causing a cascade of problems. Depending on how robust the system is at recovering from memory allocation failures, it may trigger restarts of critical routing processes or even the entire system. These may or may not be graceful and hitless, and even if they are locally a fairly low impact, these may trigger events on other routers due to the ripple effect of the network event itself. It is also worth noting that there are hardware and software limits to how much memory a given system can use - if the router in question does not use a 64-bit OS, then it is unable to address more than 4GB of RAM, for example. This may make an otherwise robust system incapable of scaling to the necessary level, and make memory usage an even more significant consideration.

6. Known issues and gaps

6.1. PE-CE routing protocols

While support for route flap dampening in BGP as a PE-CE routing protocol is equivalent to its support in non-VPN applications, the addition of IGP routing protocols such as OSPF creates a new problem, in that there is not really a way to manage route dampening, either by configuring it within the context of the IGP itself, or by configuring it in the translation point where the IGP's routing information is moved into the MP-BGP control plane infrastructure to be exchanged between participating PEs across the VPN network. This means that in the case where IGPs are used, which is often more CPU-intensive and performance-conscious to start with, the route flaps associated with an unstable network will make a bad problem even worse. It may be advisable for the IETF to document updates to standards managing use of IGPs as PE-CE routing protocols to explicitly define the use of RFD in this application.

There are also not clear guidelines based on testing and real-world experience for recommended timer values or appropriate use cases for an IGP vs BGP as a PE-CE routing protocol. In other words, rather than enterprises simply defaulting to whatever IGP is already in use

or they are most comfortable with, there may be certain cases where use of an IGP is recommended, and those where it is not. Guidance in this area may be very useful to both the SPs supporting these networks and the engineers designing the corporate networks that make use of them.

6.2. Multicast

Issues in multicast VPN scale?

6.3. Network Events

Guidance on interface event dampening values (research and testing), correlation tools to help determine root cause in a cascade failure,

6.4. General Route Scale

Discuss Virtual Aggregation as a potential solution here?

Route flap dampening may potentially be a best practice, but it has a number of shortcomings. First, there is no systematic way for end customers to view and clear dampening without some sort of advanced-functionality looking glass that allows them to view only the routes in their authorized VRFs. Also, allowing customers to make unattended clears of dampened routes may defeat the purpose of having dampening enabled at all, since customers may clear the dampening without addressing the underlying cause of the problem. In addition, as noted in [I-D.ymbk-rfd-usable] and [I-D.shishio-grow-isp-rfd-implement-survey] , Route flap Dampening is not widely used even within the Global Internet routing table, and its values probably need to be tweaked. Due to the differences in the characteristics of VPN routes compared with the global routing table, additional study and recommendations as to appropriate RFD values within a VPN are likely required. Additionally, it is not possible to configure RFD on IGPs, either natively within the PE-CE routing protocol or upstream where the learned routes are carried in MP-BGP. This means that in some cases, there is no way to insulate the SP network from the adverse impacts of rapid route churn.

6.5. Modeling and Capacity planning

There is a significant lack of multidimensional scale guidance and modeling for capacity planning and troubleshooting large-scale VPN deployments. This has a number of contributing factors. First, behavior at scale becomes increasingly non-deterministic the more variables you're working with simultaneously, so this is classically a difficult problem to model. Even worse, it's difficult to account in a model for latent design/implementation flaws: things that work

well enough at moderate scale, but are not efficient enough for high scale, or suffer some sort of secondary impact due to dependencies, race conditions, etc. These problems are often only found through extensive testing or even escape into production. Second, it is difficult to characterize an "average" implementation in such a way that it can be tested to failure in mulitple permutations to provide a reasonably accurate multidimensional model. Consequently, the guidance available normally takes the form of multiple uni-dimensional scale thresholds plus some very conservative multi-dimensional thresholds. These conservative recommendations avoid risk to both the vendor and the implementer by catering to the lowest common denominator, but they have the adverse effect of leaving a lot of capacity sitting idle. Some vendors make an effort to characterize their customers' large scale implementations such that they can better replicate real-world conditions, but gathering this information and devising ways to replicate the behavior in a lab is problematic and time-consuming.

This leads to a follow-on issue, which is that there is a lack of instrumentation on critical scaling vectors. Some routers have very limited abilities to provide useful data about critical scaling vectors (routing updates per second, changes in multicast state, sources of internal bottlenecks, etc), either for use in a model or for use as additional capacity monitoring thresholds. While most routers can provide information about CPU usage and memory thresholds, and even which processes are consuming large amounts of resources, it often takes special instrumented versions of the OS to provide a window into what is actually causing some sort of failure at scale. Because these are not routinely monitored, it means that the provider may be blind to one or more early warning signs that the router is nearing its scaling limits and cannot take action to prevent exceeding those limits before it causes customer impacts.

Additionally, even if this information is available, the provisioning systems used by most providers do not currently have the intelligence or visibility to make a decision regarding which PE to provision new customers on to evenly load the available PE routers. The provisioning system is often aware of the available physical or logical port capacity on a given router or site, and uses this as a key input to its port choice for newly provisioned customers. However, these additional capacity and scale vectors are based on real-time statistics from the router (CPU, memory load, etc) and there is no interaction or feedback loop between the provisioning system and these types of real-time router scale stats. As a result, manual intervention is often required to either remove busy routers from the available capacity pool, move spare port capacity from a busy router to a full one, or even to reprovision customers to move them from one device to another to rebalance the load on each router.

6.6. Performance issues

In many ways, it's difficult to define a hard-and-fast scale limit, because each provider and customer have a differing view on what is an acceptable performance envelope both in steady state and during recovery from outages, whether planned or unplanned. In the most extreme sorts of network events, such as a heavily loaded PE router undergoing a cold restart, the scale considerations may take something like boot time and convergence from what the involved parties consider acceptable and extend them to the point where they significantly prolong the pain that to which an end customer is exposed. They often have the added problem of making it difficult to predict the duration of an outage, because individual customer VRFs may be affected for differing amounts of time based on all of the factors that contribute to scaling. For example, if a customer has one critical route that happens to be among the last to converge, they perceive the outage to be ongoing until that last route converges, even if the entire rest of their network has been functional for a significant amount of time prior to that point.

When dealing with scheduled outages, customers obviously prefer that they never are impacted. Since this is not really possible, they expect the provider to give them very clear and accurate guidance on what the impacts will be, when they will occur, and for what duration, so that they can set expectations for their customers. VPNs are often carrying mission-critical services and data, so any downtime is bad downtime. While a customer may be understanding of a scheduled maintenance with a 15-30 minute traffic interruption while a router reloads, they may be less so if the outage actually stretches for 60-90 minutes while the router runs at 100% CPU trying to deal with this worst-case sort of load or suffers intermittent cascade problems while any remaining cushion is used up dealing with the results of the event. These impacts may be largely invisible to the provider unless they have probes within each VRF or other means to verify that traffic is no longer impacted for a given customer. It's often difficult or impossible for a provider to tell the difference between a router that is fully converged but running near 100% CPU after a reload from one that is thrashing and causing delays in convergence and customer traffic impacts while it runs at 100% CPU after a reload. Even worse, a scheduled or known outage on one router may trigger unplanned outages on other high-CPU devices. Even in unplanned outages, communication regarding impacts and duration is key, and these sorts of scale issues make it difficult to predict the impacts.

6.7. High Availability and Network Resiliency

In many cases, L3VPN services are carrying significant amounts of business-critical data. Customers and carriers design their networks to be robust enough to absorb single and sometimes even dual faults with little or no impact to the network as a whole. However, the expectations as to the frequency and duration of outages due to either scheduled or unscheduled events continue to go higher and higher. This is leading more providers to adopt features such as Non-Stop Forwarding and Non-Stop Routing, as well as things like In-Service Software Upgrades to improve the chances that outages will be transparent to the underlying customers, networks, and applications using the network elements. As these become more common within the L3VPN space, they must be considered when evaluating PE scale. Often, the machinery necessary to make these reliability enhancements work requires duplication and sharing of state between multiple elements. At its most basic level, this state sharing takes more resources and more time the more state there is to be shared, so increases in the different scaling vectors discussed in this document will cause proportional increases in the complexity and resource requirements necessary for the combined feature set. In more complex scenarios and implementations, it may contribute to the complexity associated with capacity planning, and may make the response even more non-deterministic as scale increases.

7. To-Do list

RFC EDITOR: Please remove this section before publication.

Still not discussed in the document:

use of centralized/dedicated Route-reflectors vs more distributed RRs for scale - geographic artifacts/suboptimal routing based on RR placement

Inter-AS VPN NNI scaling considerations (separate discussions on 10A, 10B/hybrid, 10C?) - include discussion on number of VRFs per NNI, routes per VRF, NNIs per router

Label Exhaustion

BGP Fast External Fallover

Comparison/disussion about horizontal scaling- pros/cons, limits, gaps

additional scaling considerations if using L2TPv3 or RSVP-TE

tunneling for PE-PE transport

Future scaling considerations (MPLS-TP at the edge, interworking with L2 technologies, significant increases in density, etc)

8. Acknowledgements

The idea for this draft came from a presentation made by Ning So during the CDNI working group meeting at IETF 81 in Quebec City where some of these same scaling considerations are discussed. Thanks also to Yakov Rekhter, Luay Jalil and Jeff Loughridge for their reviews and comments.

9. IANA Considerations

This draft makes no request to IANA..

10. Security Considerations

Security considerations for IP VPNs are covered in the protocol definitions. This draft does not introduce any new security considerations, but it is worth noting that attack vectors that result in minor impacts in a low-scale environment may make the problems observed in a high-scale or resource-constrained environment worse, thereby magnifying the potential for impacts.

11. References

11.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

11.2. Informative References

[EIGRP] Wikipedia.org, "Enhanced Interior Gateway Routing Protocol", <http://en.wikipedia.org/wiki/Enhanced_Interior_Gateway_Routing_Protocol>.

[I-D.ietf-l3vpn-ospfv3-pece] Pillay-Esnault, P., Moyer, P., Doyle, J., Ertekin, E., and M. Lundberg, "OSPFv3 as a PE-CE routing protocol", draft-ietf-l3vpn-ospfv3-pece-11 (work in progress), January 2012.

[I-D.shishio-grow-isp-rfd-implement-survey]
Pelsser, C., Bush, R., Kawamura, S., and S. Tsuchiya, "Route Flap Damping Deployment Status Survey", draft-shishio-grow-isp-rfd-implement-survey-03 (work in progress), December 2011.

[I-D.ymbk-rfd-usable]
Pelsser, C., Patel, K., Maennel, O., Mohapatra, P., and R. Bush, "Making Route Flap Damping Usable", draft-ymbk-rfd-usable-02 (work in progress), December 2011.

[IEEE802.1]
IEEE, "Connectivity Fault Management", <http://standards.ieee.org/getieee802/download/802.1ag-2007.pdf>.

[IEEE802.3]
IEEE, "Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications", <http://standards.ieee.org/about/get/802/802.3.html>.

[ISO13239]
ISO, "High-level Data Link Control protocol", <http://read.pudn.com/downloads79/doc/comm/306220/ISO%2013239.pdf>.

[RFC1661] Simpson, W., "The Point-to-Point Protocol (PPP)", STD 51, RFC 1661, July 1994.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, February 2006.

[RFC4577] Rosen, E., Psenak, P., and P. Pillay-Esnault, "OSPF as the Provider/Customer Edge Protocol for BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4577, June 2006.

[RFC4659] De Clercq, J., Ooms, D., Carugi, M., and F. Le Faucheur, "BGP-MPLS IP Virtual Private Network (VPN) Extension for IPv6 VPN", RFC 4659, September 2006.

[RFC4781] Rekhter, Y. and R. Aggarwal, "Graceful Restart Mechanism for BGP with MPLS", RFC 4781, January 2007.

[RFC4984] Meyer, D., Zhang, L., and K. Fall, "Report from the IAB Workshop on Routing and Addressing", RFC 4984, September 2007.

[RFC5880] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD)", RFC 5880, June 2010.

[RFC6037] Rosen, E., Cai, Y., and IJ. Wijnands, "Cisco Systems' Solution for Multicast in BGP/MPLS IP VPNs", RFC 6037, October 2010.

[RFC6513] Rosen, E. and R. Aggarwal, "Multicast in MPLS/BGP IP VPNs", RFC 6513, February 2012.

[Y.1731] ITU-T, "OAM functions and mechanisms for Ethernet based networks", <http://www.itu.int/rec/T-REC-Y.1731/en>.

Authors' Addresses

Wesley George
Time Warner Cable
13820 Sunrise Valley Drive
Herndon, VA 20171
US

Phone: +1 703-561-2540
Email: wesley.george@twcable.com

Rob Shakir
BT
London,
UK

Phone: +
Email: rob.shakir@bt.com